# MALLE'S CONJECTURE FOR THE COMPOSITUM OF NUMBER FIELDS

by

Jiuya Wang

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Mathematics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2018

Date of final oral examination: 06/11/2018

The dissertation is approved by the following members of the Final Oral Committee:

Melanie Matchett Wood, Professor, Mathematics
Nigel Boston, Professor, Mathematics
Jordan Ellenberg, Professor, Mathematics
Daniel Erman, Assistant Professor, Mathematics

# ACKNOWLEDGMENTS

Firstly, I would like to thank my advisor, Prof. Melanie Matchett Wood. She has been a most wonderful advisor. Without her this work would be impossible. I am extremely grateful to her for introducing this beautiful area to me, and for her constant encouragement and guidance during the whole process. She has established a role model for me in almost every aspect in life and in math.

I am lucky to be in our number theory group, which is a friendly and supportive environment all the time. I thank Nigel Boston for teaching and showing me many interesting math. I would like to thank Jordan Ellenberg for giving such an illuminating course in elliptic curve. I want to thank Daniel Erman to teach me how to teach. I would like to thank Naser Talebizadeh Sardari for answering many of my questions in analytic number theory. I would also like to thank Tonghai Yang for giving good general suggestions.

I have benefited a lot from discussions with Manjul Bhargava, Jürgen Klüners, Arul Shankar, Takashi Taniguchi, Frank Thorne and Jacob Tsimerman during the completion of this work. I would like to express my gratitude for their generous sharing of knowledge and suggestions.

Thank you to all of my friends, for always being caring and sincere. It has been a very good time having you around. I would like to thank Mao Li for answering my random questions in algebraic geometry. I would also like to thank Yuan Liu, for spending all the hard working summers together, and for the enjoyable time reading and discussing together.

Lastly, I would like to thank my husband for taking good care of me and for making our life fun all the time. I am indebted to him for the time I spent with math and for many of his wise advices in both life and math.

# TABLE OF CONTENTS

Page

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

Counting number fields is one of the biggest questions in arithmetic statistics. Malle's conjecture gives a prediction on the asymptotic behavior of the distribution function. Aside from being a natural question to ask, it also has important applications, for example, in determining the asymptotic distribution of class groups. This thesis introduces new results on Malle's conjecture and approaches to prove Malle's conjecture for the compositum of number fields based on proven results of Malle's conjecture.

In the first part of this thesis we focus on giving uniform estimates of ramified extensions. This will be the key input we need to develop later theorems. We use class field theory to prove uniform estimates for abelian extensions. We combine techniques in geometry of numbers and class field theory to prove new uniform estimates for $S_n$ extensions for $n = 3, 4, 5$.

In the second part of this thesis we give a framework to prove Malle's conjecture for the compositum of two number fields based on proven results of Malle's conjecture and good uniformity estimates on ramified number fields. Then we prove Malle's conjecture for $S_n \times A$ over any number field $k$ for $n = 3$ with $A$ an abelian group of order relatively prime to 2, for $n = 4$ with $A$ an abelian group of order relatively prime to 6 and for $n = 5$ with $A$ an abelian group of order relatively prime to 30. As a consequence, we prove that Malle's conjecture is true for $C_3 \wr C_2$ in its $S_9$ representation, whereas its $S_6$ representation is the first counter-example of Malle's conjecture given by Klüners.

In the third part of this thesis we focus on error estimates and secondary terms in the asymptotic distribution of number fields. We take advantage of error estimates results for $S_3$ extensions. After combining a sieve method together with good uniformity estimates, we prove a secondary term for the asymptotic estimate of $S_3 \times A$ extensions over $\mathbb{Q}$ when $A$ is an odd abelian group with minimal prime divisor greater than 5. At the same time, we prove the existence of a power saving error when $A$ is any odd abelian group.

Finally, in the conclusion part, we give a summary on the general picture for this topic and bring forward several questions of interest.

# Chapter 1

# Introduction

## 1.1 Malle's Conjecture

### 1.1.1 Statement

There are only finitely many number fields with bounded discriminant, therefore it makes sense to ask how many there are. Malle's conjecture aims to answer the asymptotic question for number fields with prescribed Galois group. Let $k$ be a number field and $K/k$ be a degree $n$ extension with Galois closure $\tilde{K}/k$, we define $\mathrm{Gal}(K/k)$ to be $\mathrm{Gal}(\tilde{K}/k)$ as a transitive permutation subgroup of $S_n$ where the permutation action is defined by its action on the $n$ embeddings of $K$ into $\bar{k}$. Let $N_k(G, X)$ be the number of isomorphism classes of extensions of $k$ with Galois group isomorphic to $G$ as a permutation subgroup of $S_n$ and absolute discriminant bounded by $X$.

**Conjecture 1.1** (Malle's conjecture, [Mal02, Mal04])**.** *Given $G \subset S_n$ the Galois group, then*

$$N_k(G, X) \sim X^{1/a} \ln^{b-1} X,$$

*where $a$ and $b$ are integers depending on $G$ and $k$.*

### 1.1.2 Application

Simple as the statement might be, Malle's conjecture has been discovered to relate to many questions, and it has become a main source of proving results in arithmetic statistics. The distribution of class groups is another big question in arithmetic statistics. The *Cohen-Lenstra heuristics* attempt to answer this question for quadratic fields. The only case we can prove is the average 3-class number which is derived from $S_3$ field counting [DH71, DW88]. This heuristic has been generalized in two ways. If we consider the class group over non-abelian field extensions instead of abelian extensions, then the *Cohen-Lenstra-Martinet heuristics* suggest an answer. The only case that is proven is the average 2-class number of $S_3$ cubic fields, derived from Bhargava's work on $S_4$ field counting [Bha05]. Instead of considering the distribution of class groups, which relates to unramified abelian extensions, Wood [Woo17] gives the *non-abelian Cohen-Lenstra conjecture* on non-abelian unramified extensions over quadratic fields. The only case that is known to give a finite average number is unramified $A_n$ extensions over quadratic fields such that the resulting Galois group is $S_n$, for $n = 4, 5$, and this is also derived from counting of $S_n$ number fields for $n = 4, 5$ [Bha14].

Instead of asking the distribution of class groups in average, Malle's conjecture is also a key input in considering the distribution of class numbers pointwisely. Results of counting $S_n$ number fields for $n = 3, 4, 5$ have been applied by Ellenberg, Pierce and Wood [EPW] to get unconditional upper bound on $\ell$-torsion for almost all $S_n$ number fields when $n = 3, 4, 5$. On the other hand, the work of Pierce, Turnage-Butterbaugh and Wood [PTBW] uses a completely different method to incorporate Malle's conjecture as the key input in proving a new Chebotarev density theorem, with which they can then bound $\ell$-torsion for almost all $G$ number fields with $G$ in an infinite family of arbitrarily large Galois groups.

### 1.1.3 Known Results

Malle's conjecture has been proven for abelian extensions over $\mathbb{Q}$ [Mäk85] and over arbitrary bases [Wri89]. However, for non-abelian groups, there are only a few cases known. The first case is $S_3$ cubic fields proved by Davenport and Heilbronn [DH71] over $\mathbb{Q}$ and later proved by Datskovsky and Wright [DW88] over any $k$. Bhargava and Wood [BW08] and Belabas and Fouvry [BF10] independently proved the conjecture for $S_3$ sextic fields. The cases of $S_4$ quartic fields [Bha05] and $S_5$ quintic fields [Bha10] over $\mathbb{Q}$ are also proved by Bhargava. In [BSW17], these cases are generalized to arbitrary $k$ by Bhargava, Shankar and Wang. The case of $D_4$ quartic fields over $\mathbb{Q}$ is proved by Cohen, Diaz y Diaz and Olivier

[CyDO02]. Later Klüners proved the conjecture for groups in the form of $C_2 \wr H$ [Klü12] under mild conditions on $H$. The main result of this thesis is to prove Malle's conjecture for $S_n \times A$ in its $S_{n|A|}$ representation for $n = 3, 4, 5$ with certain families of $A$.

**Theorem 1.1.1.** *Let $A$ be an abelian group and let $k$ be any number field. Then there exists $C$ such that the asymptotic distribution of $S_n \times A$-number fields over $k$ by absolute discriminant is*

$$N_k(S_n \times A, X) \sim CX^{1/|A|}$$

*in the following cases:*

1. *$n = 3$, if $2 \nmid |A|$;*
2. *$n = 4$, if $2, 3 \nmid |A|$;*
3. *$n = 5$, if $2, 3, 5 \nmid |A|$.*

However, Malle's conjecture has been shown to be not generally correct. Klüners [Klü05] shows that the conjecture does not hold for $C_3 \wr C_2$ number fields over $\mathbb{Q}$ in its $S_6$ representation, where Malle's conjecture predicts a smaller power for $\ln X$ in the main term. See [Klü05] and [Tur08] for suggestions on how to fix the conjecture. And by relaxing the precise description of the power for $\ln X$, weak Malle's conjecture states that $N_k(G, X) \sim CX^{1/a(G)+\epsilon}$. Klüners and Malle proved weak Malle's conjecture for nilpotent groups [KM04].

Notice that for Klüners' counter example, $C_3 \wr C_2 \simeq S_3 \times C_3$, we have the following corollary.

**Corollary 1.2.** *Malle's conjecture holds for $C_3 \wr C_2$ in its $S_9$ representation over any number field $k$.*

## 1.2 Secondary Term

Once the main term is understood, it is natural to ask if we could understand the error terms better. What is striking is in some cases, we can even determine a secondary term for the asymptotic distribution. The most celebrated example we know in this aspect is $S_3$ cubic fields.

The main term is due to Davenport and Heilbronn [DH71] and is generalized by Datskovsky and Wright [DW88] to any global field with characteristic not equal to 2 or 3, along with the average class number result. Their results state as following:

**Theorem 1.2.1** ([DH71, DW88])**.** *There exist a constant $C$ such that*

$$N_k(S_3, X) \sim CX,$$

*where $k$ is any global field with characteristic not equal to 2 or 3.*

Interestingly, this counting $N_{\mathbb{Q}}(S_3, X)$ has a secondary term in the order of $X^{5/6}$. The existence of this secondary term, called Roberts' conjecture, is conjectured in both [DW88] and [Rob01]. This conjecture was proved independently by Bhargava, Shankar and Tsimerman [BST13] and by Taniguchi and Thorne [TT13] at the same time, but with very different methods. A secondary term for the average class number is also proved in both papers. By combination of these two methods, Bhargava, Taniguchi and Thorne [BTT16] are able to prove this result with a better error term. Moreover in both [TT13] and [BTT16], the asymptotic distributions of $S_3$ cubic extensions with local conditions are obtained with an explicit dependency of local parameter in the error term, which the second part of this thesis heavily depends on.

**Theorem 1.2.2** ([BTT16], Theorem 4.3)**.** *There exists constant $A$ and $B$ such that the asymptotic distribution of $S_3$ cubic extensions over $\mathbb{Q}$ is*

$$N_{\mathbb{Q}}(S_3, X) = AX + BX^{5/6} + O(X^{2/3+\epsilon}).$$

Thorne has a summary [Tho11] on all approaches to understand the secondary term for cubic fields, including Hough's [Hou10] and Zhao's work [Zha13] on variations of Roberts' conjecture from different perspective, aside from the results mentioned above. However there is still little understanding towards the secondary term in general. No conjecture on the secondary term is known from the author's knowledge.

It is surely beneficial if more examples of secondary terms for asymptotic estimates of $N_k(G, X)$ are presented. It is natural to look at the asymptotic distribution of $S_4$ quartic fields $N_{\mathbb{Q}}(S_4, X)$, of which the main term is proved in [Bha05], since they are also parametrized by orbits in a pre-homogeneous vector space and give average 2-class number of $S_3$ cubic fields. In [CyDO06], the authors record a conjectural secondary term in the order of $X^{5/6}$ of quartic fields by Yukie, along with a third term in the order of $X^{3/4}\ln X$ and even a fourth term $X^{3/4}$. However no proof on the secondary term in the quartic case is known. On the other hand, Taniguchi and Thorne [TT14] conjectured a secondary term with precise constant on $S_3$ sextic fields, and a third term is also conjectured. It would be possible to prove the secondary term in the sextic case if both the exponent of $X$ and the dependency of the local parameters could be improved a lot in the error term of the distribution of cubic fields with local conditions.

The main result of the second part of this thesis is to prove the secondary term for the asymptotic distribution of $S_3 \times A$ number fields with degree $3|A|$ for $A$ with minimal prime divisor greater than 5. This provides a second example of a secondary term in distribution of number fields, and actually infinitely many such examples.

**Theorem 1.2.3.** *Let $A$ be an abelian group with minimal prime divisor greater than 5. Then there exist $C_1$, $C_2$ and $\delta > 0$ such that the asymptotic distribution of $S_3 \times A$ number fields with degree $3|A|$ over $\mathbb{Q}$ by absolute discriminant is*

$$N_{\mathbb{Q}}(S_3 \times A, X) = C_1 X^{1/|A|} + C_2 X^{5/6|A|} + O(X^{5/6|A|-\delta}).$$

The constants $C_1$ and $C_2$ are all finite sum of Euler products. As an example, in section 4.7 we give the precise constants $C_1$ and $C_2$ when $A = C_l$ is cyclic group with prime order $l > 5$. For $A$ with minimal prime divisor 3 or 5, we prove a weaker result, i.e., a power saving error is obtained.

**Theorem 1.2.4.** *Let $A$ be any odd abelian group. Then there exist $C$ and $\delta > 0$ such that the asymptotic distribution of $S_3 \times A$-number fields over $\mathbb{Q}$ by absolute discriminant is*

$$N_{\mathbb{Q}}(S_3 \times A, X) = C X^{1/|A|} + O(X^{1/|A|-\delta}).$$

The amount of power saving $\delta$ in both Theorem 1.2.3 and 1.2.4 are computed in section 4.8.

## 1.3 Outline of this thesis

In the rest of this thesis, we go through some of the author's results: Chapter 2 is the author's work on uniform estimates for $S_n$ extensions for $n = 3, 4, 5$ and abelian extensions, which is the common crucial input for both Chapter 3 and 4; Chaper 3 is the author's work on proving Malle's conjecture for $S_n \times A$ extensions for $n = 3, 4, 5$; Chapter 4 is the author's work on further determining the secondary term and power saving error for $S_3 \times A$ extension over $\mathbb{Q}$; Some of the conclusions are in Chapter 5.

# Chapter 2

# Uniform Estimates for Ramified Extensions

In this chapter we include and prove some uniform upper bound on number of $S_n$ and $A$ extensions for $n = 3, 4, 5$ that are ramified at finitely many places. It is crucial input for both Chapter 3 and the Chapter 4.

## 2.1 Local uniformity for Abelian extensions

It has been proved [Wri89] that Malle's conjecture is true for all abelian groups over any number field $k$.

**Theorem 2.1.1.** *Let $A$ be a finite abelian group and $k$ be a number field, the number of $A$-extensions over $k$ with the absolute discriminant bounded by $X$ is*

$$N(A, X) \sim CX^{1/a(A)}(\ln X)^{b(k,A)-1}.$$

We will need to prove a uniformity estimate for $A$ extensions with certain local conditions. For an arbitrary integral ideal $q$ in $O_k$, define

$$N_q(A, X) = \sharp\{K \mid \mathrm{Disc}(K/k) \leq X, \mathrm{Gal}(K/k) = A, q | \,\mathrm{disc}(K/k)\}.$$

**Theorem 2.1.2.** *Let $A$ be a finite abelian group and $k$ be a number field, then*

$$N_q(A, X) \leq O(C^{\omega(q)})(\frac{X}{|q|})^{1/a(A)}(\ln X)^{b(k,A)-1}$$

*for an arbitrary integral ideal $q$ in $O_k$, where $C$ and the implied constant depends only on $k$ .*

*Proof.* We will follow the notation and the language of [Woo10] to describe abelian extensions. To get an upper bound of $A$-number fields, it suffices to bound on the number of continuous homomorphisms from the idèle class group $C_k \to A$. Similarly, for $A$-number fields with certain local conditions, it suffices to bound on the number of continuous homomorphisms from the idèle class group $C_k \to A$ satisfying certain local conditions.

Let $S$ be a finite set of primes such that $S$ generates the class group of $k$, including infinite primes and possibly wildly ramified primes, i.e., primes above the prime divisors of $|A|$. Denote $J_k$ to be the idèle group of $k$, $J_S$ to be the idèle group with component $O_v^\times$ for all $v \notin S$ and $O_S^*$ to be $k^* \cap J_S$. By lemma 2.8 in [Woo10], the idèle class group $C_k = J_k/k^\times \simeq J_S/O_S^\times$. Therefore to bound the number of continuous homomorphisms $C_k \to A$, we can choose to bound the number of continuous homomorphisms $J_S \to A$. The Dirichlet series for $J_S \to A$ with respect to absolute discriminant is an Euler product, see [Woo10] section 2.4,

$$F_{S,A}(s) = \prod_{p\in S}\Big(\sum_{\rho_p:k_p^*\to A} |p|^{-d(\rho_p)s}\Big)\prod_{p\notin S}\Big(\sum_{\rho_p:O_p^*\to A} |p|^{-d(\rho_p)s}\Big) = \sum_n \frac{a_n}{n^s} \tag{2.1}$$

where $d(\rho_p)$ is the exponent of $p$ in the relative discriminant and can be determined by the tame inertia group at $p$, which is the image of $O_p^*$ in $A$. Lemma 2.10 [Woo10] shows that $F_{S,A}(s)$ has exactly the same right most pole with Dirichlet series for $A$-number fields at $s = \frac{1}{a(A)}$ with the same order $b(k, A)$.

$F_{S,A}(s)$ is a nice Euler product: for all $p$-factor there is a uniform bound $M$ on the magnitude of coefficient $a_{p^r}$ and a uniform bound $R$ on $r$ such that $a_{p^r}$ is zero for $r > R$. Denote the counting function of $F_{S,A}(s)$ by $B(X) = \sum_{n\le X} a_n$. Then for a certain integer $q = \prod_i p_i^{r_i}$, denote $B_q(X) = \sum_{q|n<X} a_n$. Let $q_0 = \prod_i p_i^R$ then

$$\begin{aligned} B_q(X) &= \sum_{q|d|q_0} a_d \sum_{k,(d,k)=1,dk<X} a_k \le \sum_{q|d|q_0} a_d B(\frac{X}{d}) \le \sum_{q|d|q_0} M^{\omega(q)}(\frac{X}{d})^{1/a(A)} \ln^{b(A)-1} X \\ &= M^{\omega(q)} X^{1/a(A)} \ln^{b(A)-1} X \sum_{q|d|q_0} \frac{1}{d^{1/a(A)}} \\ &\le (MR)^{\omega(q)} X^{1/a(A)} \ln^{b(A)-1} X \frac{1}{q^{1/a(A)}} = O(C^{\omega(q)})(\frac{X}{q})^{1/a(A)} \ln^{b(A)-1} X. \end{aligned} \tag{2.2}$$

We have $N_q(A, X)$ bounded by $B_{|q|}(X)$ for an arbitrary integral ideal $q$. □

## 2.2 Uniformity Estimate for $S_n$ Number Fields

In this section, we are going to include and prove some necessary uniformity results we need for $S_3$ cubic, $S_4$ quartic, $S_5$ quintic number fields over arbitrary global field $k$.

### 2.2.1 Uniformity for $S_n$ Extensions via Class Field Theory

We will include the uniformity estimates for $S_3$ and $S_4$ extensions with certain ramification behavior at finitely many places. Both results are deduced by class field theory.

For totally ramified $S_3$ cubic extensions, we have Proposition 6.2 from [DW88]:

**Theorem 2.2.1** ([DW88], Proposition 6.2)**.** *The number of non-cyclic cubic extensions over $k$ which are totally ramified at a product of finite places $q = \prod p_i$ is:*

$$N_q(S_3, X) = O(\frac{X}{|q|^{2-\epsilon}}),$$

*for any number field $k$ and any square-free integral ideal $q$. The constant is independent of $q$, and only depends on $k$.*

For discussions about overramified $S_4$ quartic extensions, we will follow the definition of [Bha05]: $p$ is overramified if $p$ factors into $P^4$, $P^2$ or $P_1^2P_2^2$ for a finite place $p$ and if $p$ factors into a product of two ramified places for infinite place. Equivalently, this means the inertia group at $p$ contains $\langle(12)(34)\rangle$ or $\langle(1234)\rangle$. The uniformity estimate for overramified $S_4$ extensions over $\mathbb{Q}$ is given in [Bha05], see Proposition 23. And we are going to prove the same uniformity over an arbitrary number field $k$ by the same method. Let $K_{24}$ be an $S_4$ extension over $k$. Denote $K_6$ and $K_3$ to be the subfields corresponding to the subgroup $E = \{(e, (12), (34), (12)(34))\}$ and $H = \langle E, (1234)\rangle$.

**Theorem 2.2.2.** *The number of $S_4$ quartic extensions over $k$ which are overramified at a product of finite places $q = \prod p_i$ is:*

$$N_q(S_4, X) = O(\frac{X}{|q|^{2-\epsilon}}),$$

*for any number field $k$ and any square free integral ideal $q$. The constant is independent of $q$, and only depends on $k$.*

*Proof.* We can apply the class field theory argument in [Bha05]. On one hand, over arbitrary $k$ we still have that $\mathrm{Nm}_{K_3/k}(\mathrm{disc}(K_6/K_3))$ is a square ideal in $k$ for any $S_4$ extension. Actually

$$\mathrm{Nm}_{K_3/k}(\mathrm{disc}(K_6/K_3)) = \mathrm{Disc}(K_6)/\mathrm{Disc}(K_3)^2,$$

which is the Artin conductor associated to the character $\chi = \mathrm{Ind}_E^G - 2 \cdot \mathrm{Ind}_H^G$ where $E$ and $H$ are corresponding subgroups of $K_6$ and $K_3$. Here $\mathrm{Ind}_E^G$ is the induced character of the identity character of $E$ as a subgroup of $G = S_4$. By computation, the character $\chi$ has value $-4$ at the conjugacy class of (12)(34), and $-2$ at (1234). The character values are even and so the Artin conductor is always a square. On the other hand, we still have the result on the mean 2-class number of non-cyclic cubic extensions over any number field $k$ in [BSW17]. It follows that the summation of 2-class number is $O(X)$ over non-cyclic cubic extensions with bounded discriminant. □

### 2.2.2 Local uniformity for $S_n$ Extensions via Geometric Sieve

In this section, we are going to prove the uniformity of $S_5$ extensions by geometry of numbers based on previous works [Bha10, Bha14, BSW17]. We will use slightly different notation just for this section. Denote $K$ to be an arbitrary number field with degree $d = \deg(K)$. For a certain scheme $Y \in \mathbb{A}_{\mathbb{Z}}^n$, let $k$ be its codimension. One example of the uniform estimates that we could prove and input to get $S_5 \times A$ counting is the following:

**Theorem 2.2.3.** *The number of $S_5$ quintic extensions over $K$ which are totally ramified at a product of finite places $q = \prod p_i$ is:*

$$N_q(S_5, X) = O(\frac{X}{|q|^{4/15-\epsilon}}),$$

*for any number field $K$ and any square free integral ideal $q$. The constant is independent of $q$, and only depends on $k$.*

We could prove some nontrivial estimates via this method for every ramification type of $S_n$ extensions for $n = 3, 4, 5$.

The proof is an application of Bhargava's geometric sieve method [Bha14]. By [Bha14], the points in the prehomogenous space with certain ramification at a finite place $p$ are $O_K/pO_K$-points on a certain scheme $Y$, which is cut out by partial derivatives of the discriminant polynomial. And to get a power saving error, we can apply the averaging technique like in [BBP10, BST13, ST] as suggested in Remark 4.2 in [Bha14]. Instead of considering points that have extra ramification at primes greater than $M$, we only need to look at the number of points that have extra ramification at specified primes $q = \prod p$. So we will first determine the number of $O_K/qO_K$-points of a scheme $Y$ in an expanding ball and then compute the number of lattice points in the fundamental domain by averaging technique. We first look at the case when $K$ is $\mathbb{Q}$. Corresponding to Theorem 3.3 in [Bha14], we have the following theorem.

**Theorem 2.2.4.** *Let $B$ be a compact region in $\mathbb{R}^n$ having finite measure. Let $Y$ be any closed subscheme of $\mathbb{A}^n_{\mathbb{Z}}$ of codimension $k$. Let $r$ be a positive real number and $q$ be a square free integer. Then we have*

$$\sharp\{a \in rB \cap \mathbb{Z}^n \mid a(mod\ q) \in Y(\mathbb{Z}/q\mathbb{Z})\} = O(r^{n-k}) \cdot C^{\omega(q)} \cdot \max\{1, (\frac{r}{q})^k\},$$

*where the implied constant depends only on $B$ and $Y$, and $C$ is an absolute constant only depending on $Y$.*

*Proof.* The case when $k = 0$ is trivial since the number of lattice points in the box is $O(r^n)$. So the initial case is $k = 1$ with $n = 1$ . Then there is only one polynomial $f(x)$ for $n = 1$. The number of points is $O(C^{\omega(q)} \cdot \max\{1, \frac{r}{q}\})$ where we could choose $C$ to be the degree of $f$ and the implied constant depends on $f$ and $B$.

We will apply induction on $n$ and $k$. Let $\pi : \mathbb{A}^n_{\mathbb{Z}} \to \mathbb{A}^{n-1}_{\mathbb{Z}}$ be the projection onto the first $n-1$ coordinates. By dimension formula, the image $\bar{Y}$ of $Y$ in $\mathbb{A}^{n-1}_{\mathbb{Z}}$ is a closed subscheme with codimension at least $k - 1$. And we can choose $\pi$ carefully so that for each $y =$

$(a_1, \cdots, a_{n-1}) \in \mathbb{Z}^{n-1}$ that $y(\text{mod } q) \in \bar{Y}(\mathbb{Z}/q\mathbb{Z})$, the number of lattice points lying in the fiber is

$$\sharp\{a = (a_1, \ldots, a_{n-1}, b) \in rB \cap \mathbb{Z}^n \mid a(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\},$$

and is bounded by $C^{\omega(q)} \cdot \max\{1, \frac{r}{q}\}$. Indeed suppose $f \in \mathbb{Z}[x_1, \cdots, x_n]$ vanishes on $Y$, and $s$ is the direction of projection, then $f(v + st)$ as a polynomial in $t$ has leading coefficients as a polynomial in $s$. So if we choose $s$ such that the leading coefficients is non-zero, then aside from finitely many $p$, the number of solutions in $\mathbb{Z}/p\mathbb{Z}$ at a fixed $v$ is bounded by the degree of $f$. Therefore, the number of solutions in $\mathbb{Z}/q\mathbb{Z}$ is at most $O(C^{\omega(q)})$ where $C$ is the degree of $f$ and the implied constant depends on the bad primes. And the number of lattice points follows by the induction to $n = 1$ case.

By induction, the number of $y \in \mathbb{Z}^{n-1}$ in the projection of $rB$ and in $\bar{Y}(\mathbb{Z}/q\mathbb{Z})$ is $O(r^{n-k}) \cdot C^{\omega(q)} \cdot \max\{1, (\frac{r}{q})^{k-1}\}$, and the number of $x_n$ for each $y$ is $C^{\omega(q)} \cdot \max\{1, \frac{r}{q}\}$. So the totaly estimate is

$$\begin{aligned} &\sharp\{a \in rB \cap \mathbb{Z}^n \mid a(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\} \\ =&O(r^{n-k}) \cdot C^{\omega(q)} \cdot \max\{1, (\frac{r}{q})^{k-1}, \frac{r}{q}, (\frac{r}{q})^k\} = O(r^{n-k}) \cdot C^{\omega(q)} \cdot \max\{1, (\frac{r}{q})^k\}. \end{aligned} \tag{2.3}$$

□

Notice that although Theorem 3.3 in [Bha14] deals with all $p > M$, it can also give an upper bound for counting at a single prime. On one hand, our statement includes the cases where finitely many ramification conditions are specified. On the other hand, as suggested by Bhargava, we can get a slightly better error of order $r^{n-k}$ instead of $r^{n-k+1}$.

In order to apply the averaging technique, we also need to consider the number of lattice points in the box $mrB$ that is not necessarily expanding homogeneously in each direction. Here $m$ is a lower triangle unipotent transformation in $GL_n(\mathbb{Q})$ which does not change the estimate much. And $r = (r_1, \ldots, r_n)$ is the scaling factors and the estimate will depend on $r_i$.

**Theorem 2.2.5.** *Let $B$ be a compact region in $\mathbb{R}^n$ having finite measure. Let $Y$ be any closed subscheme of $\mathbb{A}^n_{\mathbb{Z}}$ of codimension $k$. Let $r = (r_1, \ldots, r_n)$ be a diagonal matrix of positive real* *number where $r_i \geq \kappa$ for a certain $\kappa$, $q$ be a square free integer, and $m$ be a lower triangle unipotent transformation in $GL_n(\mathbb{R})$. Then we have*

$$\sharp\{a \in mrB \cap \mathbb{Z}^n \mid a(mod\ q) \in Y(\mathbb{Z}/q\mathbb{Z})\} = O(\frac{\prod_{i=1}^n r_i}{q^k}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{r_i}, \frac{q^2}{r_i r_j}, \dots, \frac{q^k}{\prod_{i=i_1}^{i_k} r_i}\},$$

*where the implied constant depends only on $B$, $Y$ and $\kappa$, and $C$ is an absolute constant only depending on $Y$.*

*Proof.* For case $k = 0$, we can get the result $O(\prod_{i=1}^n r_i)$ directly because the total count of lattice points in $mrB$ only differs with those in $rB$ by lower dimension projections of $rB$ which could be bounded by $O(\prod_{i=1}^n r_i)$ where the implied constant depends on $\kappa$.

The initial case when $k = 1$, $n = 1$ is estimated to be $O(\frac{r_1}{q}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{r_1}\}$. It is the same with Theorem 2.2.4 since there is no non-trivial unipotent action. For general $n$ and $k$, we will still consider the projection to the first $n-1$ coordinates. By induction, the number of points in $\bar{Y}$ is at most $O(\frac{\prod_{i=1}^{n-1} r_i}{q^{k-u}}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{r_i}, \frac{q^2}{r_i r_j}, \dots, \frac{q^{k-u}}{\prod_{i=i_1}^{i_{k-u}} r_i}\}$. And for a fixed $y = (a_1, \dots, a_{n-1}) \in \mathbb{Z}^{n-1}$, the number of lattice points lying in the fiber is

$$\begin{aligned}&\sharp\{a = (a_1, \dots, a_{n-1}, b) \in mrB \cap \mathbb{Z}^n \mid a(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\}\\ =&\sharp\{b \in P_y(mrB) \cap \mathbb{Z} \mid a(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\}.\end{aligned} \tag{2.4}$$

Here $P_y(R)$ means the section of $R$ with $y = (a_1, \dots, a_{n-1})$ fixed where $R$ is any compact region. A lower triangle unipotent transformation $m$ has the property that once $x_i$ is fixed for $i < k$, then the action on $x_k$ is just a translation. Therefore there exists $y'$ such that $P_y(mR)$ and $P_{y'}(R)$ only differ by a constant translation, i.e., $P_y(mR) = P_{y'}(R) + b_0$ where $b_0$ is a constant vector. Since the estimate only depends on the compact region in terms of its low dimension projection, constant translation will not affect the estimate, so we can look at instead

$$\begin{aligned}&\sharp\{b \in P_{y'}(rB) \cap \mathbb{Z}^k \mid a(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\}\\ =&O(\frac{r_n}{q}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{r_n}\}.\end{aligned} \tag{2.5}$$

The implied constant in the last equality could be bound uniform for all $y$ by similar argument in Theorem 2.2.4. Therefore by taking the product, we get

$$\begin{aligned}&\sharp\{a \in mrB \cap \mathbb{Z}^n \mid a(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\} \\ =&O(\frac{\prod_{i=1}^n r_i}{q^k}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{r_i}, \frac{q^2}{r_i r_j}, \ldots, \frac{q^k}{\prod_{i=i_1}^{i_k} r_i}\},\end{aligned} \tag{2.6}$$

and the implied constant depends only on $B$, $Y$ and $\kappa$. □

**Remark 2.1.** *We can consider the above theorem as an improvement on Theorem* 26 *[BST13] in this special case. Indeed, the cubic rings $K$ that are ramified at $p$ with $p^k|$ $\mathrm{Disc}(K)$ are a union of $O(p^{4-k})$ translation of lattices. So we basically prove that when we count these lattice points in the expanding ball $mrB$, we do not get those error terms at the tail in line* (29) *in [BST13].*

***Proof of Theorem 2.2.3 over*** $\mathbb{Q}$. We first prove this statement over $\mathbb{Q}$ and then will show that the computation over other number field $K$ should give the same answer. Recall that the quintic order is parametrized by $G(\mathbb{Z})$-orbits in $V(\mathbb{Z})$ where $G = GL_4 \times GL_5$ and $V$ is the space of quadruples of skew symmetric $5 \times 5$ matrices. Denote the fundamental domain of $G(\mathbb{R})/G(\mathbb{Z})$ by $\mathcal{F}$ and $B$ is a compact region in $V(\mathbb{R})$. Let $S$ be any $G(\mathbb{Z})$-invariant subset of $V_{\mathbb{Z}}^{(i)}$ which specifies a certain property of quintic orders, $S^{irr}$ be the subset of irreducible points in $S$, and $N(S;X)$ denotes the number of irreducible-$G(\mathbb{Z})$ orbits in $S$ with discriminant less than $X$. Then by formula (20) in[BST13], the averaging integral for a certain signature $i$ is

$$N(S;X) = \frac{1}{M_i}\int_{g\in\mathcal{F}} \sharp\{x \in S^{irr} \cap gB \cap V_{\mathbb{R}}^{(i)} : |\mathrm{Disc}(x)| < X\}dg \tag{2.7}$$

where $M_i$ is a constant depending on $B$.

Here for our purpose, $S = S_q$ should be the set of maximal orders that are totally ramified at all primes $p|q$. In order to apply Theorem 2.2.5, we can replace the condition $x \in S^{irr}$ by $x \in Y(\mathbb{Z}/q\mathbb{Z})$ where $Y$ is a codimension $k = 4$ variety in a 40 dimensional space defined by $f^{(j)} = 0$ for all partial derivatives of the discriminant polynomial with order $j < 4$. See [Bha14] for the definition of $Y$.

For $g \in G(\mathbb{R})$, we have $g = mak\lambda$ as the Iwasawa decomposition [Bha10]. Here $m$ is an lower triangle unipotent tranformation, $a = (t_1, \dots, t_n)$ is a diagonal element with determinant 1 and $k$ is an orthogonal transformation in $G(\mathbb{R})$ and $\lambda = \lambda I$ is the scaling factor. We will choose $B$ such that $KB = B$, so $gB = ma\lambda B = mrB$, in which $r = \lambda(t_1, \dots, t_n)$ satisfies that $\prod_1^n t_i = 1$. Lastly, the requirement $|\mathrm{Disc}(x)| < X$ could be dropped as long as we take $\lambda \leq O(X^{1/d})$ where this implied constant depends only on $B$. So we have

$$\sharp\{x \in S^{irr} \cap gB \cap V_{\mathbb{R}}^{(i)} : |\mathrm{Disc}(x)| < X\} \leq \sharp\{x \in mrB \cap \mathbb{Z}^n \mid a(\mathrm{mod}\ q) \in Y(\mathbb{Z}/q\mathbb{Z})\}.$$

We are going to apply Theorem 2.2.5 to estimate the integral in (**??**bstave). By [Bha10], all $S_5$ orders are parametrized by quadruples of skew symmetric $5 \times 5$ matrices. So there are 40 variables and therefore the dimension for the whole space is $n = 40$. Let's call those variables $a_{ij}^l$ where $1 \leq l \leq 4$ means the $m$-th matrix, $1 \leq i \leq 4$ is the row index of a skew-symmetric $5 \times 5$ matrix, $2 \leq j \leq 5$ is the column index. We can define the partial order among all 40 entries: $a_{jk}^i$ is smaller than $a_{mn}^l$ if $i \leq l$, $j \leq m$ and $k \leq n$. The scaling factor $t_i$ in our situation could be described by a pair of diagonal matrices $(A, B)$ where

$$A = \mathrm{diag}(s_1^{-3}s_2^{-1}s_3^{-1}, s_1 s_2^{-1}s_3^{-1}, s_1 s_2 s_3^{-1}, s_1 s_2 s_3^3)$$

and

$$B = \mathrm{diag}(s_4^{-4}s_5^{-3}s_6^{-2}s_7^{-1}, s_4 s_5^{-3}s_6^{-2}s_7^{-1}, s_4 s_5^2 s_6^{-2}s_7^{-1}, s_4 s_5^2 s_6^3 s_7^{-1}, s_4 s_5^2 s_6^3 s_7^4).$$

Then $t_{lij} = A_l B_i B_j$ is the scaling factor for the $a_{ij}^l$ entry. Since the fundamental domain requires that all $s_i \geq C$, this partial order also gives the partial order on the magnitude of $r_{lij} = \lambda t_{lij}$.

There are many regions in the fundamental domain that provides irreducible $S_5$-orders. We will consider the biggest region first, i.e., the points with $a_{12}^1 \neq 0$. This region requires that $\lambda s_1^{-3}s_2^{-1}s_3^{-1}s_4^{-3}s_5^{-6}s_6^{-4}s_7^{-2} \geq \kappa$, therefore $r_{lij} \geq \kappa$ for all $l, i, j$. Let us denote this region in $\mathcal{F}$ to be $D_\lambda = \{s_i \geq C_i \mid s_1^3 s_2 s_3 s_4^3 s_5^6 s_6^4 s_7^2 \leq \lambda/\kappa\}$. So we could apply Theorem 2.2.5 directly. Let's call this count $N^1(Y; X)$. The corresponding integrand, i.e., the number of lattice points in the expanding ball $gB$ where $g \in D_\lambda$ is bounded by

$$\begin{aligned} L^1 =& \sharp\{x \in mrB \cap V_{\mathbb{Z}}^{(i)} \mid x(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\} \\ =& O(\frac{\lambda^n}{q^k}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{\lambda t_i}, \frac{q^2}{\lambda^2 t_i t_j}, \dots, \frac{q^k}{\lambda^k \prod_{i=i_1}^{i_k} t_i}\} \\ =& O(\frac{\lambda^{40}}{q^4}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{q}{\lambda t_{112}}, \frac{q^2}{\lambda^2 t_{112}t_{113}}, \frac{q^2}{\lambda^2 t_{112}t_{212}}, \frac{q^3}{\lambda^3 t_{112}t_{113}t_{123}}, \frac{q^3}{\lambda^3 t_{112}t_{113}t_{114}}, \\ & \frac{q^3}{\lambda^3 t_{112}t_{113}t_{212}}, \frac{q^3}{\lambda^3 t_{112}t_{212}t_{312}}, \frac{q^4}{\lambda^4 t_{112}t_{113}t_{114}t_{123}}, \frac{q^4}{\lambda^4 t_{112}t_{113}t_{114}t_{212}}, \frac{q^4}{\lambda^4 t_{112}t_{113}t_{123}t_{212}}, \\ & \frac{q^4}{\lambda^4 t_{112}t_{113}t_{212}t_{213}}, \frac{q^4}{\lambda^4 t_{112}t_{113}t_{212}t_{312}}, \frac{q^4}{\lambda^4 t_{112}t_{212}t_{312}t_{412}}\}. \end{aligned} \tag{2.8}$$

To integrate $L^1$ over $D_\lambda$ and then against $\lambda$, we just need to focus on the inner integral over $D_\lambda$, and see whether the integral of those product of $t_{lij}$ over $D_\lambda$ produces $O(1)$ or $\lambda^r$ for some $r \geq 0$ as the result. If it is $O(1)$, then we just need to integrate against $\lambda$ and get the expected estimate, i.e., $\frac{X^{40-i}}{q^i}$ for $0 \leq i \leq 4$ where $i$ is the number of $t_{lij}$ factors in the product; if it is $\lambda^r$ for some power $r > 0$, then we will get a bigger power of $X$.

For example, $t_{112}^{-1} = s_1^3 s_2 s_3 s_4^3 s_5^6 s_6^4 s_7^2$ and $dg = \delta_5 ds^\times = s_1^{-8} s_2^{-12} s_3^{-8} s_4^{-20} s_5^{-30} s_6^{-30} s_7^{-20} ds^\times$, therefore $t_{112}^{-1}\delta_5$ contains $s_i$ with negative power for each $i$. So after integrating over $D_\lambda$, it is $O(1)$. Same thing holds for all other products listed as above except: $t_{112}t_{113}t_{123}$, $t_{112}t_{113}t_{114}$, $t_{112}t_{113}t_{114}t_{123}$, $t_{112}t_{113}t_{114}t_{212}$, $t_{112}t_{113}t_{123}t_{212}$. All these products have at most 4 $t_{lij}$ factors, so the biggest power we could get for $s_4$, $s_5$, $s_6$ and $s_7$ should be $(B_1B_2)^4 = s_4^{-12} s_5^{-24} s_6^{-16} s_7^{-8}$, so those later $s_i$ is never a problem.

Among the product with 3 factors, the $s_i$ part for small $i$ in $t_{112}t_{113}t_{123}$ and $t_{112}t_{113}t_{114}$ is $s_1^{-9} s_2^{-3} s_3^{-3}$. Since $s_1 \leq O(\lambda^{1/3})$, the integral over $D_\lambda$ should be $O(\lambda^{1/3})$. Among the product with 4 factors, $t_{112}t_{113}t_{114}t_{212}$ and $t_{112}t_{113}t_{123}t_{212}$ has factor $s_1^{-8} s_2^{-4} s_3^{-4}$, while $t_{112}t_{113}t_{114}t_{123}$ has a bigger term $s_1^{-12} s_2^{-4} s_3^{-4}$, whose integral ends up being $O(\lambda^{4/3})$.

So the whole result is:

$$\begin{aligned} N^1(Y;X) &\leq \frac{1}{M_i}\int_{\lambda=O(1)}^{O(X^{1/40})}\int_{D_\lambda} L^1 s_1^{-8}s_2^{-12}s_3^{-8}s_4^{-20}s_5^{-30}s_6^{-30}s_7^{-20}\mathrm{d}s^\times\mathrm{d}\lambda^\times \\ &= O(C^{\omega(q)})\cdot\max\{\frac{X}{q^4},\frac{X^{39/40}}{q^{4-1}},\frac{X^{38/40}}{q^{4-2}},\frac{X^{(37+1/3)/40}}{q^{4-3}},\frac{X^{(36+4/3)/40}}{q^{4-4}}\} \\ &= O(C^{\omega(q)})\cdot\max\{\frac{X}{q^4},\frac{X^{38/40}}{q^{4-2}},\frac{X^{(36+4/3)/40}}{q^{4-4}}\}. \end{aligned} \tag{2.9}$$

We know that there are a lot of regions containing irreducible points for $S_5$ extensions. However notice that the last term above is $X^{(37+1/3)/40}$, therefore we will not compute for those regions with a total counting smaller than this. They must contribute an even smaller counting when we consider this restriction in those regions. By [Bha10] Table 1, we can see that there are still three left to be considered when $a_{12}^1 = 0$:

2. $a_{13}^1 \neq 0$, $a_{12}^2 \neq 0$;
3. $a_{13}^1 = 0$ but $a_{14}^1, a_{23}^1, a_{12}^2 \neq 0$;
4. $a_{12}^2 = 0$, but $a_{13}^1, a_{12}^3 \neq 0$.

For 2, $D_\lambda = \{s_i \geq C_i \mid s_1^3s_2s_3s_4^3s_5s_6^4s_7^2 \leq \lambda/\kappa, s_1^{-1}s_2s_3s_4^3s_5^6s_6^4s_7^2 \leq \lambda/\kappa\}$. The definition of $D_\lambda$ makes it clear that for all $t_{lij} \geq t_{113}, t_{212}$ in the partial order we define, we have $t_{lij} \geq \kappa$. And $t_{112}$ could be arbitrarily small. So we will assume $t_{112}$ to be 1 when we plug into Theorem 2.2.5 and get an upper bound on $L^2$:

$$\begin{aligned} L^2 =&O\left(\frac{\prod_{i=2}^{40} r_i}{q^k}\right)\cdot C^{\omega(q)}\cdot\max\{1,q,\frac{q^2}{r_i},\ldots,\frac{q^k}{\prod_{i=i_1}^{i_{k-1}} r_i}\} \\ =&O(\frac{\lambda^{40}}{q^4})\cdot C^{\omega(q)}\cdot\max\{\frac{q}{\lambda t_{112}},\frac{q^2}{\lambda^2 t_{112}t_{113}},\frac{q^2}{\lambda^2 t_{112}t_{212}},\frac{q^3}{\lambda^3 t_{112}t_{113}t_{123}},\frac{q^3}{\lambda^3 t_{112}t_{113}t_{114}}, \\ &\frac{q^3}{\lambda^3 t_{112}t_{113}t_{212}},\frac{q^3}{\lambda^3 t_{112}t_{212}t_{312}},\frac{q^4}{\lambda^4 t_{112}t_{113}t_{114}t_{123}},\frac{q^4}{\lambda^4 t_{112}t_{113}t_{114}t_{212}},\frac{q^4}{\lambda^4 t_{112}t_{113}t_{123}t_{212}}, \\ &\frac{q^4}{\lambda^4 t_{112}t_{113}t_{212}t_{213}},\frac{q^4}{\lambda^4 t_{112}t_{113}t_{212}t_{312}},\frac{q^4}{\lambda^4 t_{112}t_{212}t_{312}t_{412}}\}. \end{aligned} \tag{2.10}$$

The list $L^2$ contains everything in $L^1$ except the first term $O(\frac{\lambda^{40}}{q^k})\cdot C^{\omega(q)}$. As considered before, we only need to focus on those difficult terms and it suffices to see that $s_1 \leq O(\lambda^{1/3})$ again in this $D_\lambda$.

For 3 and 4, things can be done similarly. In case 3, $a_{114} \neq 0$ and $a_{123} \neq 0$ together implies that $t_{114}^{-1}t_{123}^{-1} = s_1^6 s_2^2 s_3^2 s_4 s_5^2 s_6^3 s_7^4 \leq O(\lambda^2)$, so $s_1 \leq O(\lambda^{1/3})$. In case 4, $a_{113} \neq 0$ implies that $s_1^3 s_2 s_3 s_4^3 s_5 s_6^4 s_7^2 \leq O(\lambda)$, so $s_1 \leq O(\lambda^{1/3})$.

Therefore, we get the uniformity result for $N_q(S_5, X) = O(\frac{X}{q^{4/15-\epsilon}})$. □

In order to prove Theorem 2.2.3 over arbitrary number field $K$, we will need to prove the analogue of Theorem 2.2.5 over an arbitrary number field $K$. The setup is a bit more complex than the case over $\mathbb{Q}$. The variety that describes points with extra ramification is defined over $O_K$. Since $\rho : O_K \hookrightarrow \mathbb{R}^r \bigoplus \mathbb{C}^s$ is a full lattice, an $O_K$-point on the variety corresponds to a lattice point in $\mathbb{R}^{dn} \simeq (\mathbb{R}^r \bigoplus \mathbb{C}^s)^n$ where $d$ is the degree of $K/\mathbb{Q}$. Denote $\mathbb{R}^r \bigoplus \mathbb{C}^s$ by $F$. The scaling vector is $r = (r_1, \dots, r_n)$ where $r_i \in F$ for each $i$. Define $|\cdot|_\infty$ to be the norm in $F$: $|v|_\infty = \prod_r |v_i|_i \prod_s |v_j|_j$ where $|\cdot|_i$ means standard norm in $\mathbb{R}$ at real places and square of standard norm in $\mathbb{C}$ at complex places.

**Theorem 2.2.6.** *Let $B$ be a compact region in $F^n \simeq \mathbb{R}^{nd}$ with finite measure. Let $Y$ be any closed subscheme of $\mathbb{A}^n_{O_K}$ of codimension $k$. Let $r = (r_1, \dots, r_n)$ be a diagonal matrix of non-zero elements where $|r_i|_\infty \geq \kappa$ for a certain $\kappa$. Let $q$ be a square free prime ideal in $O_K$ and $m$ be a lower triangle unipotent transformation in $GL_n(F)$. Then we have*

$$\begin{aligned}&\sharp\{a \in mrB \cap (O_K)^n \mid a(mod\ q) \in Y(O_K/qO_K)\}\\ =&O(\frac{\prod_{i=1}^n |r_i|_\infty}{|q|^k}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{|q|}{|r_i|_\infty}, \frac{|q|^2}{|r_i r_j|_\infty}, \dots, \frac{|q|^k}{\prod_{i=i_1}^{i_k} |r_i|_\infty}\}\end{aligned} \tag{2.11}$$

*where the implied constant depends only on $B$, $Y$ and $\kappa$, and $C$ is an absolute constant only depending on $Y$.*

In order to prove this analogue, we need the following lemma on the regularity of shapes of the ideal lattices for a fixed number field $K$. Given an integral ideal $I \subset O_K$, we can embed it to $F$ as a full lattice with covolume compared with $O_K$ to be $[O_K : I] = \mathrm{Nm}_{K/\mathbb{Q}}(I)$.

**Lemma 2.2.** *Let $K$ be a number field and $I \subset O_K$ be an arbitrary ideal. Given $\lambda = (\lambda_i) \in F = \mathbb{R}^r \bigoplus \mathbb{C}^s$, then*

$$\sharp\{a \in I \mid \forall i, |\sigma_i(a)|_i \leq |\lambda_i|_i\} = O(\frac{|\lambda|_\infty}{|I|}) + 1$$

*where $\sigma_i$ for $i = 1, \dots, r+s$ are the Archimedean valuations of $K$ and $|\cdot|_i$ is the usual norm in $\mathbb{R}$ for real embeddings and square of the usual norm in $\mathbb{C}$ for complex embeddings . The implied constant depends only on $K$.*

*Proof.* Given $I$ in the ideal class $R$ in the class group of $K$, denote $[a]$ to be the equivalence class of non-zero $a$ in $I$ where $a \sim a'$ if $a = ua'$ for some unit $u$. Then we have [Lan94]

$$\sharp\{[a] \in I \mid ||[a]||_\infty \leq |I|X\} = \sharp\{\alpha \subset O_K \mid \alpha \in R^{-1}, |\alpha| < X\} = O(X). \tag{2.12}$$

To take advantage of the equality above, we cover the set $W = \{a \in I \mid \forall i, |\sigma_i(a)|_i \leq |\lambda_i|_i\}\backslash\{0\}$ by a disjoint union of subsets $W_k$:

$$W = \bigcup_{k\geq 1}\{a \in I \mid \forall i, |\sigma_i(a)|_i \leq |\lambda_i|_i, \frac{|\lambda|_\infty}{2^k} \leq |a|_\infty \leq \frac{|\lambda|_\infty}{2^{k-1}}\} = \cup_k W_k. \tag{2.13}$$

For $a \in W_k$, we have that

$$\frac{|\lambda_i|_i}{2^k} \leq |\sigma_i(a)|_i \leq |\lambda_i|_i,$$

and if $ua$ is in $W$, it must be also in the same $W_k$ since $|ua|_\infty = |a|_\infty$. So the magnitude of $u$ is bounded as $2^{-k} \leq |\sigma_i(u)|_i \leq 2^k$ by the above inequality. By Dirichlet's unit theorem, the units of $K$ aside from roots of unity after taking logarithm form a lattice of rank $r+s-1$ satisfying $\sum_i \ln|\sigma_i(u)|_i = 0$, therefore

$$\sharp\{u \in O_K^\times \mid |\ln|\sigma_i(u)|_i| \leq k\} = O(k^{r+s-1}).$$

So for each $[a] \in W_k$, the multiplicity is bounded by $O(k^{r+s-1})$, and the number of equivalence classes in $W_k$ is bounded by

$$\sharp\{[a] \in I \mid |a|_\infty < \frac{|\lambda|_\infty}{2^{k-1}}\} \leq O(\frac{|\lambda|_\infty}{|I|} \cdot \frac{1}{2^{k-1}}). \tag{2.14}$$

Therefore

$$|W_k| \leq O(\frac{|\lambda|_\infty}{|I|}) \cdot \frac{k^{r+s-1}}{2^{k-1}}. \tag{2.15}$$

The total counting by summation over all $k$ is

$$\sharp\{a \in I \mid \forall i, |\sigma_i(a)|_i \leq |\lambda_i|_i\}\backslash\{0\} = \sum_k |W_k| \leq O(\frac{|\lambda|_\infty}{|I|}) \sum_k \frac{k^{r+s-1}}{2^{k-1}} \leq O(\frac{|\lambda|_\infty}{|I|}).$$

So the total counting with the origin is

$$\sharp\{a \in I \mid \forall i, |\sigma_i(a)|_i \leq |\lambda_i|_i\} = O(\frac{|\lambda|_\infty}{|I|}) + 1.$$

$\square$

A corollary of this lemma is that the shape of the ideals lattices inside $O_K$ cannot be too skew. We will make this precise in the following lemma and prove it by a more direct approach.

**Lemma 2.3.** *Given a number field $K$ with degree $d$, for any integral ideal $I \subset O_K$, denote $\mu_i$ to be the successive minimum for the Minkowski reduced basis for $I$ as a lattice in $\mathbb{R}^d$. Then $\mu_i$ is bounded by*

$$\mu_i \leq O(|I|^{1/d})$$

*for all $1 \leq i \leq d$. The implied constant only depends on the degree of $K$, the number of complex embeddings of $K$ and the absolute discriminant of $K$.*

*Proof.* Given an integral ideal $I$, and an arbitrary non-zero element $\alpha \in I$, we have $(\alpha) \subset I$, so $|(\alpha)| \geq |I|$. The length of $\alpha$ in $\mathbb{R}^d$ is

$$\begin{aligned} &\sqrt{|\alpha|_1^2 + \cdots + |\alpha|_r^2 + |\alpha|_{r+1} + \cdots + |\alpha|_{r+s}} \\ \geq &\sqrt{d(\prod_{1\leq i\leq r} |\alpha_i|^2 \prod_{r+1\leq i\leq r+s} \frac{|\alpha|_i^2}{4})^{1/d}} \\ \geq &\sqrt{d} 2^{-s/d} |(\alpha)|^{1/d} \\ \geq &\sqrt{d} 2^{-s/d} |I|^{1/d}. \end{aligned} \tag{2.16}$$

The first inequality comes from the fact that the arithmetic mean is greater than the geometric mean. While Minkowski's first theorem guarantees that $\mu_1 \leq O(|I|^{1/d})$, we can bound $\mu_1$ by $O(|I|^{1/d})$ in the other direction. This amounts to saying that the first minimum $\mu_1$ of Minkowski's reduced basis is exactly at the order of the diameter $O(|I|^{1/d})$. Moreover Minkowski's second theorem states that

$$\prod_{1\leq i\leq d} \mu_i \leq 2^d D_K^{1/2} |I|,$$

therefore for all $i \leq d$,

$$\mu_i \leq O(|I|^{1/d})$$

where the implied constant only depends on $d$, $s$ and $D_k$. □

**Remark 2.4.** *By Lemma 2.2, if we pick $\lambda$ with $|\lambda|_\infty = O(|I|)$ such that $|\lambda_i|_i = O(|I|^{1/d})$ for real places and $|\lambda_i|_i = O(|I|^{2/d})$ for complex places, we get a square box with side length $O(|I|^{1/d})$ in $\mathbb{R}^d$. Since the first term in Lemma 4.7 could be bounded by $O(\frac{|\lambda|_\infty}{|I|}) = O(1)$, we can find a uniform upper bound of $C(|I|^{1/d})$ on the side length such that the only lattice point in a smaller square box is the origin. Therefore the first successive minimum $\mu_1$ is greater than the upper bound.*

On the other hand, the Minkowski's reduced basis generates the whole lattice with covolume $|I|D_K^{1/2}$, so the angle among the vectors in the basis is away from zero. This basically means that among the family of lattices of all integral ideals of $K$ under Minkowski's reduced basis all look like square boxes, and we can find a fundamental domain within the square box.

**Corollary 2.5.** *Given a number field $K$ with degree $d$, for any integral ideal $I \subset O_K$ and any residue class $c \in O_K/IO_K$, denote $c_i$ to be the $i$-th coordinate in $\mathbb{R}^d$. Then we can find a representative $c$ such that each*

$$|c_i| \leq O(|I|^{1/d})$$

*for all $1 \leq i \leq d$. The implied constant depends only on $K$.*

***Proof of Theorem 2.2.6.*** The case where $k = 0$ is trivial since the number of lattice points in the box is $O(\prod_{i=1}^n |r_i|_\infty)$. It suffices to prove the statement for the initial case when $k = 1$ and $n = 1$. The induction procedure works similarly with Theorem 2.2.5.

There is only one polynomial $f(x)$ to be considered for $n = 1$ and $k = 1$. Since $q$ is square free, the number of solution in $O_K/qO_K$ is bounded by $C^{\omega(q)}$ by Chinese remainder theorem. Therefore the solutions of $f(\text{mod } q)$ in $O_K$ is a union of $C^{\omega(q)}$ translations $q + c$ of the lattice $q$ where $c$ is a certain residue class in $O_K/qO_K$ that is also a solution.

Lemma 2.2 states that for arbitrary $r \in F$,

$$\sharp\{a \in rB \cap O_K \mid a \in 0 + q\} = O(\max\{\frac{|r|_\infty}{|q|}, 1\})$$

when $B$ is a unit square in $F$. It follows that the equality is true for any general compact set $B$ since it could be covered by a square and then the implied constant will also depend on $B$. For other nontrivial translations by a root $c$, we have

$$\sharp\{a \in rB \cap O_K \mid a \in c + q\} = \sharp\{a \in (rB - c) \cap O_K \mid a \in q\}. \tag{2.17}$$

So it is equivalent to consider the number of lattice points in a translation of the box. We could cover $B$ by $2^n$ sub-boxes $B_s$ which is defined by sign in each $\mathbb{R}$ space. Then $rB - c$ could be covered by $rB_s - c$. It suffices to count the lattice points in each $rB_s - c$ and add them up. For each $s$, if there exists one lattice point $P \in rB_s - c$, then we can cover $rB_s - c$ by $P + rB_s$, and the number of lattice points in $rB_s + P$ is equivalent to that in $rB_s$ which is

$$\sharp\{(P + rB_s) \cap q\} = \sharp\{rB_s \cap q\} \leq O(\max\{\frac{|r|_\infty}{|q|}, 1\}).$$

If there are no lattice points in $B_s$, then there is nothing to add. Altogether we have that for any residue class $c$ and any compact set $B$,

$$\sharp\{a \in rB \cap O_K \mid a \in c + q\} \leq O(2^n \max\{\frac{|r|_\infty}{|q|}, 1\}) = O(\max\{\frac{|r|_\infty}{|q|}, 1\}).$$

Here the implied constant depends only on $B$ and $K$. Adding up all solutions of $f$, we get

$$\sharp\{a \in rB \cap O_K \mid f(a) \equiv 0 \bmod q\} = O(\frac{|r|_\infty}{|q|}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{|q|}{|r|_\infty}\}.$$

This finishes the proof for the case $k = 1$, $n = 1$. □

Finally, based on Theorem 2.2.6, we can prove Theorem 2.2.3 over a number field $K$.

***Proof of Theorem 2.2.3 over*** $K$. We will follow the notation [BSW17] in this proof. Counting $S_n$-number fields for $n = 3, 4, 5$ over a number field $K$ is different from that over $\mathbb{Q}$ mostly in two aspects.

Firstly, the structure of finitely generated $O_K$-module is more complex than that of $\mathbb{Z}$, therefore the parametrization of $S_n$ number fields over $K$ will involve other orbits aside from $G(O_K)$-orbits of $V(O_K)$ points. Actually finitely generated $O_K$-modules with rank $n$ are classified in correspondence to the ideal class group $\mathrm{Cl}(K)$ of $K$. So for each ideal class $\beta$, we get a lattice $\mathcal{L}_\beta$ corresponding to $S_n$ extensions $L$ with $O_L$ corresponding to $\beta$. We just need to count the number of orbits in $\mathcal{L}_\beta$ under the action of $\Gamma_\beta$ where $\Gamma_\beta$ is commensurable with $G(O_K)$ and $\mathcal{L}_\beta$ is commensurable with $V(O_K)$. See section 3 in [BSW17] for more details.

Secondly, the reduction theory over a number field $K$ is slightly different in that the description of fundamental domain requires the introduction of units, and this effect of units is especially beneficial for summation over fundamental domain. The most significant difference is at the description of the torus. Originally over $\mathbb{Q}$, we have $G(\mathbb{R})/G(\mathbb{Z}) = NA\mathcal{K}\Lambda$ [Bha10] where $A$ is an $l$-dimensional torus ($l = 7$ for $S_5$) embedded into $\mathrm{GL}_n(\mathbb{R})$ ($n = 40$ for $S_5$) as diagonal elements

$$T(c) = \{t(s_1, \dots, s_l) \in T(\mathbb{R}) = \mathbb{G}_m^l(\mathbb{R}) \mid \forall i, s_i \geq c\}.$$

Given a number field $K$, recall that $\rho : O_K \hookrightarrow F = \mathbb{R}^r \bigoplus \mathbb{C}^s$ is the embedding of $O_K$ as a full lattice in $\mathbb{R}^d$. Then $A$ could be described as a subset of

$$T(c, c') = \{t = t(s_1, \dots, s_l) \in T(F) = \mathbb{G}_m^l(F) \mid \forall i, |s_i|_\infty \geq c, \forall j, k, \ln \frac{|s_i|_j}{|s_i|_k} \leq c'\}.$$

Here $|s_i|_j \leq O(|s_i|_k)$ for all $j, k$ guarantees that $|s_i|_j \sim |s_i|_k$, thus $|s_i|_v \sim |s_i|_\infty^{1/(r+s)}$. Therefore, if we have a bound that $|s_i|_\infty \leq A$, then we can get the bound $|s_i|_v \leq O(A^{1/r})$. See section 4 [BSW17] for more details.

Recall that we need to compute

$$N(S; X) = \frac{1}{M_i} \int_{g \in \mathcal{F}} \sharp\{x \in S^{irr} \cap gB \cap V_F^{(i)} : |\mathrm{Disc}(x)|_\infty < X\} dg \tag{2.18}$$

where $V_F^{(i)}$ is a subspace of $V_F$ with a certain signature, and $B$ is a compact ball in the space $V_F$ that is invariant under the action of the orthogonal group $K$. By Theorem 2.2.6, the integrand is

$$
\begin{aligned}
&\sharp\{x \in S^{irr} \cap gB \cap V_F^{(i)} : |\mathrm{Disc}(x)|_\infty < X\} \leq \sharp\{x \in m\lambda t B \cap \mathcal{L} \mid x(\mathrm{mod}\ q) \in Y(\mathbb{Z}/q\mathbb{Z})\}\\
=&O(\frac{|\lambda|_\infty^n}{|q|^k}) \cdot C^{\omega(q)} \cdot \max\{1, \frac{|q|}{|\lambda t_i|_\infty}, \frac{|q|^2}{|\lambda^2 t_i t_j|_\infty}, \dots, \frac{|q|^k}{|\lambda^k \prod_{i=i_1}^{i_k} t_i|_\infty}\}.
\end{aligned} \tag{2.19}
$$

Here in order to present the result in a similar form with that over $\mathbb{Q}$, for each $\lambda \in \mathbb{R}^+$ we denote $\lambda$ to be the diagonal matrix such that $|\mathrm{Disc}(\lambda v)|_\infty = |\lambda|_\infty^n |\mathrm{Disc}(v)|_\infty$ where $n = 40$ for $S_5$.

The first case is to compute $G(O_K)$-orbits in $V(O_K)$, which corresponds to the trivial class in $\mathrm{Cl}(K)$. Denote $\mathcal{F}$ to be $G(F)/G(O_K)$ and $\mathcal{L}$ to be the image of $V(O_K)$ in $V(F)$. We first look at the case where $a_{12}^1 \neq 0$. Since $\mathcal{L}$ is a lattice, $x$ with non-zero $a_{12}^1$ is away from zero and $|a|_\infty$ could be bounded from below by $\kappa$, so we would only integrate over

$$
D_\lambda = \{t = t(s_i) \in T(c, c') \mid |s_1^3 s_2 s_3 s_4^3 s_5^6 s_6^4 s_7^2|_\infty \leq \lambda/\kappa\}.
$$

The integral over $F = \mathbb{R}^d$ gives the same result as over $\mathbb{Q}$

$$
\int_{O(1)}^{A} |s|_\infty^u \mathrm{d}s^\times \leq \prod_{1 \leq i \leq r} \int_{O(1)}^{O(A^{1/(r+s)})} s_i^u \mathrm{d}s_i^\times \prod_{r+1 \leq i \leq r+s} \int_{O(1)}^{O(A^{1/2(r+s)})} r_i^{2(u-1)} r_i \mathrm{d}r_i = O(A^u). \tag{2.20}
$$

So we will end up with the same result over $K$.

For fields corresponding to other ideal class $\beta \in \mathrm{Cl}(K)$, we can similarly compute the average number of lattice points in $\mathcal{F}v$ for $v \in B$ with bounded discriminant. Denote $\mathcal{F}_\beta = \Gamma_\beta \backslash G(F)$. By [BSW17], we can cover $\mathcal{F}_\beta$ by finitely many $g_i\mathcal{F}$ where $g_i \in G(O_K)$ are representatives of $G(O_K)/(G(O_K) \cap \Gamma_\beta)$. Let's call $\mathcal{D}_i = \mathcal{F}_\beta \cap g_i\mathcal{F}$, then we just need to sum up

$$
\begin{aligned}
&\frac{1}{M_i} \int_{g \in \mathcal{D}_i} \sharp\{x \in S^{irr} \cap gB \cap V_F^{(i)} : |\mathrm{Disc}(x)|_\infty < X\} dg\\
\leq &\frac{1}{M_i} \int_{g \in g_i\mathcal{F}} \sharp\{x \in S^{irr} \cap gB \cap V_F^{(i)} : |\mathrm{Disc}(x)|_\infty < X\} dg\\
\leq &\frac{1}{M_i} \int_{g \in \mathcal{F}} \sharp\{x \in g_i^{-1} S^{irr} \cap gB \cap V_F^{(i)}\} dg.
\end{aligned} \tag{2.21}
$$

As in [BSW17] section 3,

$$\mathcal{L}_\beta := V_n(K) \cap \beta^{-1} \prod_{p \nmid \infty} V(O_p) \prod_{p \mid \infty} V(F_p)$$

where $\beta$ is a representative of the double coset $\mathrm{cl}_S = (\prod_{p \nmid \infty} G(O_p)) \backslash G(\mathbb{A}_f)/G(K)$. Here $\mathbb{A}_f$ is the restricted product of $K_p^\times$ for all finite places $p$. So given a class $\beta$, we can choose a representative such that $\beta_p$ is the identity element in $G(O_p)$ except at a finite set of places $S$. At $p \in S$, $\beta_p$ is in $G(K_p)$. Given $v \in \mathcal{L}_\beta$, we have

$$v_p \in \beta_p^{-1} V(O_p).$$

Since $\beta_p^{-1}$ can be regarded as a linear action, there must exist $r$ large enough such that

$$v_p \pi^r \in \beta_p^{-1} \pi^r V(O_p) \in V(O_p)$$

and $(\pi^r) = (a_p)$ is a principle integral ideal in $O_K$ where $\pi$ is a uniformizer for $O_p$. Glue all the $a_p$ and we get $a = \prod_{p \in S} a_p$. By the way it is defined, we have that $a\mathcal{L}_\beta$ is in $O_K$ and $a \in O_p^\times$ at $p \notin S$. So for $p$ outside $S$, $v \in \mathcal{L}_\beta$ is in $Y(O_K/p)$, if and only if, $av \in O_K$ is in $Y(O_K/p)$. Therefore we can consider $a\mathcal{L}_\beta$ inside $O_K$ instead and do not lose the information of ramification at all but finitely many places. Since there are only finitely many ideal classes it will not affect the form of the uniformity estimate but only the implied constant. From now on, we will assume $\mathcal{L}_\beta$ to be in $O_K$.

In (2.21), $S^{irr}$ denotes the set of totally ramified points at $q$ in $\mathcal{L}_\beta$. If $q$ is a square free integral ideal away from $S$ and $x \in S^{irr}$ satisfies $x \in O_K$ and $x \in Y(O_K/q)$, then $g_i^{-1} v \in O_K$ and $g_i^{-1} v \in g_i^{-1} Y(O_K/q)$. Denoting $g_i^{-1} Y = Y_i$, then it suffices to count

$$\sharp\{x \in g_i^{-1} \mathcal{L}_\beta \cap gB \cap Y_i(O_K/q)\}. \tag{2.22}$$

Since $g_i^{-1} Y$ only differs with $Y$ by a linear transformation on coordinates, $Y_i$ has the same codimension. Apply Theorem 2.2.6 to get the same estimates. To consider arbitrary square free ideal $q = q_1 q_2$ with $q_2$ containing the involved factors in $S$, we can estimate with $q_1$ and replace $|q_1|$ by $|q|$ with a difference of at most $O(1)$ since there are only finitely many $p \in S$.

□

### 2.2.3 Codimension 1 Example

Another new uniformity estimates we could prove is on partially ramified $S_3$ cubic extensions at finitely many primes. We first use geometric sieve to get estimates on ramified extension, and then use class field theory to merge this uniformity result with previous known uniformity estimates on totally ramified $S_3$ cubic fields. Although we only verify these results over $\mathbb{Q}$, readers could compare with last subsection to get same result over arbitrary number fields.

Let $k$ be a number field and $q$ be a square-free integral ideal in $\mathcal{O}_k$. Let us deonte $N_{q,r}(S_3, X)$ to be the number of $S_3$ cubic extensions over $k$ that are partially ramified at all places $p|q$, and totally ramified at all places $p|r$. Then recall that we have that

$$N_{1,r}(S_3, X) = O(\frac{X}{|r|^{2-\epsilon}})$$

from Theorem 2.2.1.

On the other hand, by the argument we introduced in the last subsection based on the geometric sieve method [Bha14], we will prove the following uniformity estimates on partially ramified extensions.

**Theorem 2.2.7.** *The number of non-cyclic cubic extensions over $k$ which are partially ramified at a product of finite places $q = \prod p_i$ is:*

$$N_{q,1}(S_3, X) = O(\frac{X}{|q|^{1/6-\epsilon}}),$$

*for any number field $k$ and any square-free integral ideal $q$. The constant is independent of $q$, and only depends on $k$.*

This result comes from Theorem 2.2.6 and the observation that if we just focus on the number of cubic orders ramified at a fixed finite set of places, then we can improve the power saving error in the geometric sieve[Bha14] and therefore drop the codimension 2 condition. We could similarly get the uniformity result for ramified $S_4$ and $S_5$ extensions. As a corollary of Theorem 2.2.7, we get the corresponding estimates on the average 3-class number over quadratic fields ramified at $q = \prod p_i$. Given $F$ a quadratic extension over $k$, denote $h_3^*(F/k)$ to be the relative 3-class number of $F$ over $k$.

**Corollary 2.6.** *Given a square-free integral ideal $q$, the 3-class number summed over quadratic extensions $F/k$ with $q|\operatorname{disc}(F/k)$ is bounded by*

$$\sum_{\substack{[F:k=2]\\ q|\operatorname{disc}(F),\operatorname{Disc}(F)\leq X}} h_3^*(F/k) = O(\frac{X}{|q|^{1/6-\epsilon}}).$$

*Proof.* By [DW88], there is a one-to-one correspondence between the unramified abelian cubic extensions $L/F$ such that the resulting Galois group of $\tilde{L}/k$ is $S_3$ and the isomorphism classes of nowhere totally ramified non-cyclic cubic extensions $K_3/k$. Moreover, in this correspondence, we have $\operatorname{Disc}(F) = \operatorname{Disc}(K_3)$. If $q|\operatorname{disc}(F)$, then the cubic field $K_3$ is partially ramified at $q$. Therefore

$$\sum_{\substack{[F:k=2]\\ q|\operatorname{disc}(F),\operatorname{Disc}(F)\leq X}} \frac{h_3^*(F/k)-1}{2} = O(\frac{X}{|q|^{1/6-\epsilon}}). \tag{2.23}$$

Indeed the left-hand side corresponds to the number of nowhere totally ramified $S_3$ cubic extensions which are partially ramified at $q$, and it is a subset of $S_3$ cubic extensions that are partially ramified at $q$. The right-hand side gives the upper bound on this number by Theorem 2.2.7. Rearranging the expression, and applying Theorem 4.2 [Wan17] on quadratic extensions

$$\sum_{\substack{[F:k=2]\\ q|\operatorname{disc}(F),\operatorname{Disc}(F)\leq X}} 1 = O(\frac{X}{|q|^{1-\epsilon}}),$$

we have that

$$\sum_{\substack{[F:k=2]\\ q|\operatorname{disc}(F),\operatorname{Disc}(F)\leq X}} h_3^*(F/k) = O(\frac{X}{|q|^{1/6-\epsilon}}) + O(\frac{X}{|q|^{1-\epsilon}}) = O(\frac{X}{|q|^{1/6-\epsilon}}).$$

□

And by combining the Theorem 2.2.1 and 2.2.7 using class field theory, we prove the following theorem.

**Theorem 2.2.8.** *The number of non-cyclic cubic extensions over $k$ that are partially ramified at $q = \prod p_i$ and totally ramified at $r = \prod p_j$ is bounded by*

$$N_{q,r}(S_3, X) = O(\frac{X}{|q|^{1/6-\epsilon}|r|^{2-\epsilon}}),$$

*for any number field $k$ and any square-free integral ideal $qr$. The constant is independent of $q$ and $r$, and only depends on $k$.*

*Proof.* Let $F$ be a quadratic extension over $k$ and $q$ be an integral ideal that divides $\mathrm{disc}(F)$. Let $f$ be an integral ideal in $k$ and denote the conductor of an abelian cubic extension of $F$. We would like to count $S_3$ extensions that are partially ramified at $q$, so it suffices to look at quadratic fields $F$ with $q|\,\mathrm{disc}(F)$. We would also like to count $S_3$ extensions that are totally ramified at $r$, so it suffices to look at cubic abelian extensions over $F$ with conductor divided by $r$. By Lemma 6.2 [DW88], the number of cubic extensions over $F$ with conductor $f$ such that the resulting Galois group over $k$ is $S_3$, could be bounded by $O(4^{\omega(f)}h_3^*(F/k))$ where $\omega(f)$ is the number of prime divisors of $f$, and the implied constant only depends on $k$. So we just need to bound

$$\begin{aligned}
&\sum_{\substack{[F:k]=2 \\ q|\,\mathrm{disc}(F)}} \sum_{\substack{r|f \\ |f|^2\,\mathrm{Disc}(F)\le X}} 4^{\omega(f)} h_3^*(F/k) \\
=&4^{\omega(r)} \sum_{f} 4^{\omega(f)} \sum_{\substack{[F:k]=2 \\ q|\,\mathrm{disc}(F), \mathrm{Disc}(F)\le \frac{X}{|f|^2|r|^2}}} h_3^*(F/k) \\
\le&4^{\omega(r)} \sum_{f} 4^{\omega(f)} \frac{X}{|f^2 r^2||q|^{1/6-\epsilon}} \\
\le&O(\frac{X}{|q|^{1/6-\epsilon}|r|^{2-\epsilon}}) \sum_{f} \frac{4^{\omega(f)}}{|f|^2} \le O(\frac{X}{|q|^{1/6-\epsilon}|r|^{2-\epsilon}}).
\end{aligned} \tag{2.24}$$

□

***Proof of Theorem 2.2.7.*** We will prove over $\mathbb{Q}$, and the result holds equally when the base field $k$ is an arbitrary number field by Theorem 4.7 in [Wan17].

Firstly, recall that cubic orders are parametrized as $\mathrm{GL}_2(\mathbb{Z})$-orbits of the space of binary cubic forms $V(\mathbb{Z}) = \{ax^3 + bx^2y + cxy^2 + dy^3 \mid (a, b, c, d) \in \mathbb{Z}^4\}$. Please see details in section 2 and 3 in [BST13]. By Theorem 4.5 in [Wan17], let us denote $Y$ to be the variety that describes the ramification type introduced in [Bha14], we just need to integrate the the following integrand

$$L^1 = \sharp\{x \in mrB \cap V_{\mathbb{Z}}^{(i)} \mid x(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\} = O(C^{\omega(q)}) \cdot \max\{\frac{\lambda^4}{q}, \lambda^3 t^3\}, \tag{2.25}$$

over the fundamental domain of GL($\mathbb{R}$)/GL($\mathbb{Z}$) where $t \geq \sqrt[4]{3}/\sqrt{2}$. Please see section 5 in [BST13] for more details on the description of the fundamental domain. Let's denote $S$ to be the set of cubic orders that are ramified at $q$, then

$$\begin{aligned} N(S;X) &\leq O(C^{\omega(q)})\frac{1}{M_i}\int_{\lambda=O(1)}^{O(X^{1/4})}\int_{t=\sqrt[4]{3}/\sqrt{2}}^{O(\lambda^{1/3})} \max\{\frac{\lambda^4}{q}, \lambda^3 t^3\} t^{-2} \mathrm{d}t^{\times} \mathrm{d}\lambda^{\times} \\ &= O(C^{\omega(q)})\frac{1}{M_i}\int_{\lambda=O(1)}^{O(X^{1/4})} \max\{\frac{\lambda^4}{q}, \lambda^3 \lambda^{1/3}\} \mathrm{d}\lambda^{\times} \\ &= O(C^{\omega(q)}) \cdot \max\{\frac{X}{q}, X^{5/6}\} = O(C^{\omega(q)}) \cdot \max\{\frac{X}{q}, X^{5/6}\}. \end{aligned} \tag{2.26}$$

Since $|q| < X$, we have the number bounded by $O(\frac{X}{|q|^{1/6-\epsilon}})$. The global case follows similarly.

□

# Chapter 3

# Malle's Conjecture for Compositum of Number Fields - Main Term

In this chapter we go through the author's work on proving main theorem Theorem 1.1.1. In section 3.1, we analyze the discriminant of a compositum in terms of each individual discriminant, and then compute the case explicitly for $S_n \times A$. Then we check that our computation agrees with Malle's prediction. In section 3.2, we prove the product argument in two different cases. Finally, in section 3.3, we prove our main theorems based on what we have developed before.

## 3.1 Discriminant of Compositum

### 3.1.1 General Description

We will describe the relation between $\mathrm{Disc}(KL)$ and $\mathrm{Disc}(K)$, $\mathrm{Disc}(L)$ when $\tilde{K}$ and $\tilde{L}$ have trivial intersection.

**Theorem 3.1.1.** *Let $K/k$ and $L/k$ be extensions over $k$ which intersect trivially, then $\mathrm{Disc}(KL) \leq \mathrm{Disc}(K)^n \mathrm{Disc}(L)^m$, where $n = [L : k]$, $m = [K : k]$.*

*Proof.* If $k = \mathbb{Q}$, then the ring of integers $O_K$ and $O_L$ are free $\mathbb{Z}$-modules with rank $m$ and $n$. Then $\mathrm{Disc}(O_K O_L) = \mathrm{Disc}(K)^n \mathrm{Disc}(L)^m$ and $O_K O_L \subset O_{KL}$. Over arbitrary $k$, we have $\mathrm{disc}(S^{-1}O_K/S^{-1}O_k) = S^{-1}\,\mathrm{disc}(O_K/O_k)$ as an $O_k$-module, see e.g. Theorem 2.9 [Neu99]. We take $S = O_k \backslash p$ for some prime ideal $p \subset O_k$ to look at $\mathrm{disc}_p(K/k)$. Now $S^{-1}O_k \subset k$ is a discrete valuation ring with the unique maximal ideal $S^{-1}p$, and $S^{-1}O_K$ is a finitely generated

$S^{-1}O_k$-module, therefore admits an integral basis. Notice that $S^{-1}(O_K)$ intersects trivially with $S^{-1}O_L$, so it follows $\operatorname{disc}_p(KL) \leq \operatorname{disc}_p(K)^n \operatorname{disc}_p(L)^m$ similarly. □

This gives an upper bound of $\operatorname{Disc}(KL)$. To be more precise, we focus on the study of $\operatorname{Disc}(KL)$ at tamely ramified primes over arbitrary number field $k$. Firstly, any tame inertia group is cyclic, therefore it could be described by the generator. Secondly, suppose $I = \langle g \rangle$ at a certain finite place $p$, then the index of $g \in G \subset S_n$,

$$\operatorname{ind}(g) = n - \sharp\{\text{orbits}\} = \sum (e_i - 1) f_i,$$

is exactly the exponent for the $p$-part of the relative discriminant ideal. So we can determine the discriminant at $p$ by looking at the cycle type of $g$.

If $\tilde{K} \cap \tilde{L} = k$, then $\operatorname{Gal}(\tilde{K}\tilde{L}/k) \simeq \operatorname{Gal}(\tilde{K}/k) \times \operatorname{Gal}(\tilde{L}/k)$, where the isomorphism is a product of the restrictions to $\tilde{K}$ and $\tilde{L}$. Say $\operatorname{Gal}(\tilde{K}/\mathbb{Q}) = G_1 \subset S_m$ and $\operatorname{Gal}(\tilde{L}/\mathbb{Q}) = G_2 \subset S_n$, then $G = G_1 \times G_2$ has a natural permutation representation in $S_{mn}$. Suppose $\tilde{K}$ and $\tilde{L}$ are both tamely ramified at $p$ with $I_i = \langle g_i \rangle \subset G_i$ , for $i = 1, 2$, then $\tilde{K}\tilde{L}$ is also tamely ramified since tamely ramified extensions are closed under taking compositum. And the inertia group is $I = \langle g \rangle = \langle (g_1, g_2) \rangle$ for $\tilde{K}\tilde{L}$ because the inertia group for a sub-extension behaves naturally as quotient.

**Theorem 3.1.2.** *Let $K$ and $L$ be given above, and let $e_i$, for $i = 1, 2$, be the ramification indices of $\tilde{K}$ and $\tilde{L}$ at a tamely ramified $p$. If $(e_1, e_2) = 1$, then $\operatorname{ind}(g) = \operatorname{ind}(g_1) \cdot n + \operatorname{ind}(g_2) \cdot m - \operatorname{ind}(g_1) \cdot \operatorname{ind}(g_2)$.*

*Proof.* Suppose $g_1 \in G_1 \subset S_m$ is a product of disjoint cycles $\prod c_k$, then $e_1$ will be the least common multiple of $|c_k|$, the length of cycles $c_k$ for all $k$. Similarly for $g_2$ as a product of cycles $\prod d_l$. Now embed $(g_1, g_2)$ to $S_{mn}$, the permutation action is naturally defined to be mapping $a_{i,j}$ to $a_{g_1(i),g_2(j)}$ for $1 \leq i \leq m$, $1 \leq j \leq n$. If $(e_1, e_2) = 1$, then for any $k, l$, $(|c_k|, |d_l|) = 1$ and $(c_k, d_l)$ forms a single cycle of length $|c_k||d_l|$ in $S_{mn}$. So the number of orbits in $g$ is the product of number of orbits in $g_i$. Therefore $\operatorname{ind}(g) = mn - (m - \operatorname{ind}(g_1))(n - \operatorname{ind}(g_2)) = \operatorname{ind}(g_1) \cdot n + \operatorname{ind}(g_2) \cdot m - \operatorname{ind}(g_1) \cdot \operatorname{ind}(g_2)$. □

This gives a nice description of $\text{disc}_p(KL)$ independent of the cycle type when the ramification indices are relatively prime. In general, to know $\text{ind}(g)$ requires more information on the cycle type of $g_i$.

**Theorem 3.1.3.** *Let $K$ and $L$ be as given above, $g_1$ be a product of disjoint cycles $\prod c_k$ and $g_2$ be a product of disjoint cycles $\prod d_l$ where $g_i$ is the generator for a tame ramified $p$ for $\tilde{K}$and $\tilde{L}$, then* $\text{ind}(g) = mn - \sum_{k,l} \gcd(|c_k|, |d_l|)$.

*Proof.* Notice that we can write $\text{ind}(g_1) = \sum_k(|c_k| - 1)$. In general, $(c_k, d_l)$ is no longer a single orbit in $S_{mn}$. Instead, it splits into $\gcd(|c_k|, |d_l|)$ many orbits. So the summation is $\text{ind}(g) = \sum_{k,l}(|c_k||d_l| - \gcd(|c_k|, |d_l|)) = mn - \sum_{k,l} \gcd(|c_k|, |d_l|)$. □

### 3.1.2 Discriminant for $S_n \times A$

We will describe the example of $S_n \times A$ for our interests in detail here. We will only consider the cases where $n = 3, 4, 5$ and $A$ is an odd order abelian group.

Firstly, we take the example of $S_3 \times A$ where $A = C_{l^k}$ is cyclic with odd prime power order $l^k$. Possible tame inertia generators in $S_3$ could be (12), (123). For $A \subset S_{|A|}$, possible generators are of the form $g = (123...l^k)$ or powers of $g$, i.e., a single cycle of length $l^k$ or a product of $l^r$ cycles of length $l^{k-r}$. So $\text{ind}(g)$ is minimized when $g$ is $l^{k-1}$ product of cycles of length $l$, therefore $\text{ind}(A)$ is $l^k - l^{k-1}$, and $\frac{|A|}{\text{ind}(A)} = \frac{l}{l-1}$. If $l \neq 3$, then we can apply Theorem 3.1.2 to get Table 1. The numbers in the table give the exponent for $p$ in $\text{disc}_p$ for each field.

| $S_3$ | $C_{l^k}$ | $S_3 \times C_{l^k}$ |
|---|---|---|
| (12) | $l^k - l^r$ | $3l^k - 2l^r$ |
| (123) | $l^k - l^r$ | $3l^k - l^r$ |

Table 3.1 Table of $\text{Disc}_p$ for $S_3 \times C_{l^k}$, $l \neq 3$

If $l = 3$, we apply Theorem 3.1.3 to get Table 2.

We do not include in the table the cases where one of the inertia groups is trivial since

| $S_3$ | $C_{l^k}$ | $S_3 \times C_{l^k}$ |
|---|---|---|
| $(12)$ | $l^k - l^r$ | $3l^k - 2l^r$ |
| $(123)$ | $l^k - l^r$ | $3l^k - 3l^r$ |

Table 3.2 Table of $\mathrm{Disc}_p$ for $S_3 \times C_{l^k}$, $l = 3$

$\mathrm{disc}_p(KL) = \mathrm{disc}_p(K)^n \,\mathrm{disc}_p(L)^m$ at these $p$ from previous computation. To compute the precise table for general $A$, we can compute the table for all abelian $l$-groups and then apply Theorem 3.1.2 inductively to combine different $l$-parts. The general pattern we need for the proof of the main theorems is:

**Lemma 3.1.** *Let $A$ be an abelian group of odd order $m$ and $(12)$, $(123)$ be elements in $S_3$. Then for all $c \in A$, $\mathrm{ind}((12), c)/m > 2$, $\mathrm{ind}((123), c)/m > 1$.*

*Proof.* For any abelian group $A$, $\frac{|A|}{\mathrm{ind}(A)} = \frac{p}{p-1}$ where $p$ is the minimal prime divisor of $|A|$, and $\frac{p}{p-1} < 2$ if $p \neq 2$. This can be seen by combining the different $l$-parts of $A$ inductively. The value $\mathrm{ind}((12), c) = m + 3 \cdot \mathrm{ind}(c) - \mathrm{ind}(c) = m + 2 \cdot \mathrm{ind}(c) \geq m + 2 \cdot \mathrm{ind}(A) > 2m$ because $\frac{|A|}{\mathrm{ind}(A)} < 2$.

For $\mathrm{ind}((123), c)$, if $3 \nmid |A|$, then $\mathrm{ind}((123), c) = 2m + 3 \cdot \mathrm{ind}(c) - 2 \cdot \mathrm{ind}(c) = 2m + \mathrm{ind}(c) > m$ with no problem. If $3 || A|$, we separate 3-part of $A$ to compute $\mathrm{ind}((123), c)$. Let $A = A_3 \times A_{>3}$ where $A_3$ is the 3-part of $A$ and $A_{>3}$ contains all $p > 3$ part. Let $c = (c_3, c_{>3})$ be any element in $A$, then $\mathrm{ind}((123), c) = \mathrm{ind}((123), c_3, c_{>3}) = \mathrm{ind}(((123), c_3), c_{>3})$ where $((123), c_3)$ is an element in $S_3 \times A_3$. Say $\mathrm{ind}((123), c_3) = i$, then

$$\begin{aligned} \mathrm{ind}((123), c_3, c_{>3}) &= i|A_{>3}| + (3|A_3| - i) \cdot \mathrm{ind}(c_{>3}) \\ &= i(|A_{>3}| - \mathrm{ind}(c_{>3})) + 3|A_3| \cdot \mathrm{ind}(c_{>3}). \end{aligned} \tag{3.1}$$

Therefore the minimal value of $\mathrm{ind}((123), c)$ is obtained when both $i$ and $\mathrm{ind}(c_{>3})$ are smallest possible. The smallest possible $\mathrm{ind}(c_{>3})$ is $\mathrm{ind}(A_{>3})$. The smallest $\mathrm{ind}((123), c_3)$ is $\mathrm{ind}((123), e) = 2|A_3|$. Therefore, if $A = A_3$, then $2|A_3|/m = 2 > 1$. If $A_{>3}$ is non-trivial, then by (3.1), $\mathrm{ind}((123), c) \geq 2m + |A_3| \cdot \mathrm{ind}(A_{>3}) > m$. □

**Lemma 3.2.** *Let $A$ be an abelian group of odd order and $2, 3 \nmid |A| = m$ and $(12)$, $(123)$, $(1234)$, $(12)(34)$ be elements in $S_4$. Then for all $c \in A$, $\operatorname{ind}((12), c)/m > 2$, $\operatorname{ind}((12)(34), c)/m > 1$, $\operatorname{ind}((123), c)/m > 3$, $\operatorname{ind}((1234), c)/m > 2$.*

*Proof.* We can apply Theorem 3.1.2 since $2, 3 \nmid m$. Then $\operatorname{ind}((12), c) = m + 3 \cdot \operatorname{ind}(c) \geq m+3\cdot\operatorname{ind}(A) > 2m$, $\operatorname{ind}((12)(34), c) = 2m+2\cdot\operatorname{ind}(c) > m$, $\operatorname{ind}((1234), c) = 3m+\operatorname{ind}(c) > 2m$, $\operatorname{ind}((123), c) = 2m + 2 \cdot \operatorname{ind}(c) \geq 2m + 2 \cdot \operatorname{ind}(A) \geq 2m + 2 \cdot \frac{4}{5}m > 3m$. □

**Lemma 3.3.** *Let $A$ be an odd abelian group and $2, 3, 5 \nmid |A| = m$. Then $\forall c \in A$ and $k \in S_5$ , $\operatorname{ind}(k, c)/m \geq 1 + \operatorname{ind}(k) - 1/7$.*

*Proof.* We can apply Theorem 3.1.2 since $2, 3 \nmid m$. Then $\operatorname{ind}(k, c) = m \operatorname{ind}(k) + 5 \operatorname{ind}(c) - \operatorname{ind}(k)\operatorname{ind}(c) = m \operatorname{ind}(k) + (5 - \operatorname{ind}(k)) \operatorname{ind}(c) = (m - \operatorname{ind}(c)) \operatorname{ind}(k) + 5 \operatorname{ind}(c)$. So for a certain $k$, the value is smallest when $\operatorname{ind}(c) = \operatorname{ind}(A)$. And at this time $\operatorname{ind}(k, c)/m = \operatorname{ind}(k) + (5 - \operatorname{ind}(k))\frac{\operatorname{ind}(A)}{m} = \operatorname{ind}(k) + (5 - \operatorname{ind}(k))\frac{p-1}{p}$ where $p$ is the smallest divisor of $m$ and $p \geq 7$. So $\operatorname{ind}(k)/m - \operatorname{ind}(k) = (5 - \operatorname{ind}(k))\frac{p-1}{p} \geq (5 - 4)\frac{6}{7} = \frac{1}{7}$. □

### 3.1.3 Malle's Prediction for $S_n \times A$

In this section we compute the value of $a(G)$ and $b(k, G)$ for $S_n \times A$. A similar discussion on $a(G)$ for a direct product of two Galois groups in general is in [Mal02]. We include here for the convenience of the reader. Recall that given $G \subset S_n$ a permutation group, for each element $g \in G$, $\operatorname{ind}(g) = n - \sharp\{\text{orbits of g}\}$. We define $a(G)$ to be the minimum value of $\operatorname{ind}(g)$ among all $g \neq e$. The absolute Galois group $G_k$ acts on the conjugacy classes of $G$ via its action on the character table of $G$. We define $b(k, G)$ to be the number of orbits within all conjugacy classes with minimal index.

Let $G_i \subset S_{n_i}, i = 1, 2$ be two permutation groups. Consider $G = G_1 \times G_2 \subset S_{n_1 n_2}$. Suppose that $g_i \in G_i$ gives minimal index, then for $G \subset S_{n_1 n_2}$, the minimal index will either come from $g_1 \times e$ or $e \times g_2$ since for any $g \in G_2$, $\operatorname{ind}(g_1, e) \leq \operatorname{ind}(g_1, g)$. One can compute $\operatorname{ind}(g_1 \times e) = n_2 \operatorname{ind}(g_1)$. Therefore $a(G) = \min\{n_2 \cdot a(G_1), n_1 \cdot a(G_2)\} = n_1 n_2 \min\{\frac{a(G_1)}{n_1}, \frac{a(G_2)}{n_2}\}$.

If $\frac{a(G_1)}{n_1} < \frac{a(G_2)}{n_2}$, then $g \times e$ for all $g$ with $\mathrm{ind}(g) = a(G_1)$ are exactly the elements with minimal index in $G$. Irreducible representations of $G_1 \times G_2$ are $\rho_1 \otimes \rho_2$ where $\rho_i$ are irreducible representations of $G_i$ with character $\chi_i$. The corresponding character is $\chi_1 \cdot \chi_2$. Therefore the $G_k$ action on $g \times e$ has the same orbit as its action on $g$. So $b(k, G) = b(k, G_1)$.

Our case $S_n \times A$ satisfies the above condition, therefore $a(S_n \times A) = nm \min\{\frac{1}{n}, \frac{p-1}{p}\} = m$ where $p$ is the smallest prime divisor of $|A| = m$ and $n = 3, 4, 5$. And $b(k, S_n \times A) = b(k, S_n) = 1$.

## 3.2 Product Lemma

This section answers the question: given two distributions $F_i$, $i = 1, 2$, each describes the asymptotic distribution of some multi-set of positive integers $S_i$, i.e., $F_i(X) = \sharp\{s \in S_i \mid s \leq X\}$, what is the product distribution $P_{a,b}(X) = \sharp\{(s_1, s_2) \mid s_i \in S_i, s_1^a s_2^b \leq X\}$ where $a, b > 0$. We will split the discussion into two cases.

**Lemma 3.4.** *Let $F_i(X)$, $i = 1, 2$, be as given above, $F_i(X) \sim A_i X^{n_i} \ln^{r_i} X$ where $0 < n_i \leq 1$ and $r_i \in \mathbb{Z}_{\geq 0}$. If $\frac{n_1}{a} - \frac{n_2}{b} = 0$, then*

$$P_{a,b}(X) \sim \frac{A_1 A_2}{a^{r_1} b^{r_2}} \frac{r_1! r_2!}{(r_1 + r_2 + 1)!} \frac{n_1}{a} X^{\frac{n_1}{a}} \ln^{r_1 + r_2 + 1} X.$$

*Proof.* We will prove this in three steps.

**Case 1:** $n_i = 1$**,** $F_1(X) = A_1 X \ln^{r_1} X + o(X \ln^{r_1} X)$**,** $F_2(X) = A_2 X \ln^{r_2} X + O(1)$. We can assume $a = b = 1$. Define $a_n$ to be the number of copies of $n$ in $S_1$, then

$$F_1(X) = \sum_{n \leq X} a_n.$$

To simplify, we denote the main term of $F_i(X)$ by $M_i(X)$, then

$$\begin{aligned} P_{1,1}(X) &= \sum_{s_1 \in S_1} F_2(\frac{X}{s_1}) = \sum_{n \leq X} a_n F_2(\frac{X}{n}) \\ &= \sum_{n \leq X} a_n M_2(\frac{X}{n}) + \sum_{n \leq X} a_n O(1). \end{aligned} \tag{3.2}$$

The last term is easily shown to be small

$$\sum_{n\leq X} a_n O(1) \leq O(\sum_{n\leq X} a_n) = O(X \ln^{r_1} X). \tag{3.3}$$

Assuming $X$ is an integer, we apply summation by parts to compute the first sum

$$\sum_{n\leq X} a_n M_2(\frac{X}{n}) = F_1(X)M_2(1) - \int_1^X F_1(t)\frac{\mathrm{d}}{\mathrm{d}t}(M_2(\frac{X}{t}))\,\mathrm{d}t. \tag{3.4}$$

If $r_2 = 0$, the boundary term is

$$A_1A_2X\ln^{r_1} X + o(X\ln^{r_1} X),$$

otherwise it is 0. The derivative in the integral is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(M_2(\frac{X}{t})) &= -A_2X\frac{1}{t^2}(\ln^{r_2}\frac{X}{t} + r_2\ln^{r_2-1}\frac{X}{t})\\ &= X(\sum_{0\leq i\leq r_2} P_i(t)\ln^i X).\end{aligned} \tag{3.5}$$

So the integral is

$$\sum_{0\leq i\leq r_2} X\ln^i X\int_1^X F_1(t)P_i(t)\,\mathrm{d}t. \tag{3.6}$$

It is standard in analysis that if $f$ and $g$ are positive and $\lim_{X\to\infty}\int_1^X f(t)g(t)\,\mathrm{d}t = \infty$, then $\int_1^X o(f(t))g(t)\,\mathrm{d}t = o(\int_1^X f(t)g(t)\,\mathrm{d}t)$. Therefore we can plug in $M_1(t)$ for $F_1(t)$ to estimate each integral up to a small error. One can check that for each $i$ the integral of $M_1(t)P_i(t)$ together with $X\ln^i X$ has a main term in the order $X\ln^{r_1+r_2+1}X$. So we can replace $F_1(t)$ by $M_1(t)$ in (3.3) with an error in the order of $o(X\ln X^{r_1+r_2+1})$. Denote the following integral $I$,

$$\begin{aligned}I &= \int_1^X M_1(t)\frac{\mathrm{d}}{\mathrm{d}t}(M_2(\frac{X}{t}))\,\mathrm{d}t\\ &= -A_1A_2X\int_1^X \ln^{r_1} t\cdot(\ln^{r_2}\frac{X}{t} + r_2\ln^{r_2-1}\frac{X}{t})\frac{\mathrm{d}t}{t}.\end{aligned} \tag{3.7}$$

Using the substitution $u = \frac{\ln t}{\ln X}$, we reduce the integral

$$\int_1^X \ln^{r_1} t\cdot\ln^{r_2}\frac{X}{t}\frac{\mathrm{d}t}{t} = \ln^{r_1+r_2+1} X\int_0^1 u^{r_1}(1-u)^{r_2}\,\mathrm{d}u \tag{3.8}$$

to Beta function[WW96] $B(r_1+1, r_2+1)$, therefore

$$-I = A_1A_2B(r_1+1, r_2+1)X\ln^{r_1+r_2+1}X + o(X(\ln X)^{r_1+r_2+1}). \tag{3.9}$$

This is always of greater order than the boundary term, and hence finishes the proof of the first case.

**Case 2:** $n_i = 1$**,** $F_i(X) = A_iX\ln^{r_i}X + o(X\ln^{r_i}X)$.

For any $\epsilon$, we can bound $F_i(X)$ by $A_iX\ln^{r_i}X(1+\epsilon) + O_\epsilon(1)$. Therefore we can bound

$$\limsup_{X\to\infty}\frac{P_{1,1}(X)}{X\ln^{r_1+r_2+1}X} \le (1+\epsilon)^2A_1A_2B(r_1+1, r_2+1),$$

by Case 1. Similarly we can bound

$$\liminf_{X\to\infty}\frac{P_{1,1}(X)}{X\ln^{r_1+r_2+1}X} \ge (1-\epsilon)^2A_1A_2B(r_1+1, r_2+1).$$

So the limit exists and has to be $A_1A_2B(r_1+1, r_2+1)$. In case where some $A_i = 0$, we only need the upper bound to show the limit is 0.

**General case:**

Generally, we consider all possible $a$ and $b$. The condition $s_1^as_2^b \le X$ is equivalent to $s_1^{n_1}s_2^{n_2} \le X^{n_1/a} = X^{n_2/b}$. The distribution of $s_i^{n_i}$ is

$$F_i(X^{1/n_i}) = \frac{A_i}{n_i^{r_i}}X\ln^{r_i}X + o(X\ln^{r_i}X), \tag{3.10}$$

and we can regard $\frac{A_i}{n_i^{r_i}}$ as the new coefficients. The general distribution is the product distribution in Case 2 when one plugs in $X^{n_1/a}$,

$$\begin{aligned} P_{a,b}(X) &= \frac{A_1}{n_1^{r_1}}\frac{A_2}{n_2^{r_2}}B(r_1+1, r_2+1)(\frac{n_1}{a})^{r_1+r_2+1}X^{n_1/a}(\ln X)^{r_1+r_2+1} + o(X^{n_1/a}(\ln X)^{r_1+r_2+1}) \\ &\sim \frac{A_1}{a^{r_1}}\frac{A_2}{b^{r_2}}B(r_1+1, r_2+1)\frac{n_1}{a}X^{n_1/a}(\ln X)^{r_1+r_2+1}. \end{aligned} \tag{3.11}$$

□

**Lemma 3.5.** *Let $F_i(X)$, $i = 1, 2$ be as given above, $F_i(X) \sim A_iX^{n_i}\ln^{r_i}X$ where $0 < n_i \le 1$ and $r_i \in \mathbb{Z}_{\ge 0}$. If $\frac{n_1}{a} - \frac{n_2}{b} > 0$, then there exists a constant $C$ such that*

$$P_{a,b}(X) \sim CX^{\frac{n_1}{a}}\ln^{r_1}X.$$

*Furthermore if* $F_i(X) \leq A_i X^{n_i} \ln^{r_i} X$*, then we have*

$$P_{a,b}(X) \leq A_1 A_2 \frac{r_2!}{b^{r_2} a^{r_1}} \frac{1}{(\frac{n_1}{a} - \frac{n_2}{b})^{r_2+1}} \frac{n_1}{a} X^{\frac{n_1}{a}} \ln^{r_1} X.$$

*Proof.* We first prove the existence of $C$ in two steps.

**Case 1:** $F_1(X) = A_1 X^{n_1} \ln^{r_1} X + O(1)$**,** $F_2(X) = A_2 X^{n_2} \ln^{r_2} X + o(X^{n_2} \ln^{r_2} X)$.

As in Lemma 3.4, we need to bound the sum

$$\begin{aligned} P_{a,b}(X) &= \sum_{n^a m^b \leq X} a_n b_m = \sum_{m^b \leq X} b_m F_1(\frac{X^{1/a}}{m^{b/a}}) \\ &= \sum_{m^b \leq X} b_m A_1 (\frac{X^{1/a}}{m^{b/a}})^{n_1} \ln^{r_1}(\frac{X^{1/a}}{m^{b/a}}) + \sum_{m^b \leq X} b_m O(1) \qquad (3.12) \\ &= \frac{A_1}{a^{r_1}} X^{n_1/a} \ln^{r_1} X \sum_{m^b \leq X} \frac{b_m}{m^{b n_1/a}} (1 - \frac{\ln m^b}{\ln X})^{r_1} + O(X^{n_2/b} \ln^{r_2} X). \end{aligned}$$

It suffices to show the sum

$$C(X) = \sum_{m^b \leq X} \frac{b_m}{m^{b n_1/a}} (1 - \frac{\ln m^b}{\ln X})^{r_1},$$

converges to a constant $C'$, i.e., $C(X) = C' + o(1)$. Notice that $C(X)$ is monotonically increasing, so it suffices to show $C(X)$ is bounded. We will assume $X$ to be integral for simplicity, by summation by parts,

$$\begin{aligned} C(X) \leq \sum_{m^b \leq X} \frac{b_m}{m^{b n_1/a}} &= \frac{F_2(X^{1/b})}{X^{n_1/a}} + \frac{b n_1}{a} \int_1^{X^{1/b}} F_2(t) t^{-b n_1/a - 1} \, \mathrm{d}t \\ &\leq O(X^{n_2/b - n_1/a}) + \frac{b n_1}{a} \int_1^{X^{1/b}} (M t^{n_2} \ln^{r_2} t + M) t^{-b n_1/a - 1} \, \mathrm{d}t, \end{aligned} \qquad (3.13)$$

is bounded by a constant. The first term is $o(1)$ since $\frac{n_1}{a} - \frac{n_2}{b} > 0$. For the second term, we can always find $M$ such that $F_2(t) \leq M t^{n_2} \ln^{r_2} t + M$, where the constant term $M$ is a technical modification when $t = 1$. One can compute the integral to see that it is bounded by a constant. Therefore, we have proved that $C(X) = C' + o(1)$ and

$$P_{a,b}(X) \sim \frac{A_1 C'}{a^{r_1}} X^{n_1/a} \ln^{r_1} X.$$

**Case 2:** $F_i(X) = A_i X^{n_i} \ln^{r_i} X + o(X^{n_i} \ln^{r_i} X)$.

Notice that $C(X)$ is purely dependent on $F_2(X)$ and independent of $F_1(X)$ once we have decided on these constants $r_i$, $n_i$ and $a$, $b$. Therefore the coefficient of the main term of $P_{a,b}$ is linearly dependent on $A_1$.

To get the upper bound, we can bound $F_1(X) \leq A_1(1+\epsilon)X^{n_1} \ln^{r_1} X + O_\epsilon(1)$ by definition and compute the upper bound of $P_{a,b}(X)$,

$$\limsup_{X\to\infty} \frac{P_{a,b}(X)}{X^{n_1/a}\ln^{r_1} X} \leq (1+\epsilon)\frac{A_1}{a^{r_1}}C'$$

by Case 1. Similarly, we can deal with the lower bound. Therefore,

$$\lim_{X\to\infty} \frac{P_{a,b}(X)}{X^{n_1/a}\ln^{r_1} X} = \frac{A_1}{a^{r_1}}C'$$

which proves the general case with $C = \frac{A_1 C'}{a^{r_1}}$.

**Bound on $C$:**

Next we assume further that $F_i(X)$ are bounded by $M_i(X) = A_i X^{n_i} \ln^{r_i} X$. We want to show the constant $C$ can be bounded by $O(A_1 A_2)$. By summation by parts,

$$\begin{aligned} P_{a,b}(X) &\leq \sum_{n\leq X^{1/a}} a_n M_2(\frac{X^{1/b}}{n^{a/b}}) \\ &\leq F_1(\lfloor X^{1/a}\rfloor)M_2(1) - \int_1^{\lfloor X^{1/a}\rfloor} M_1(t)\frac{\mathrm{d}}{\mathrm{d}t}(M_2(\frac{X^{1/b}}{t^{a/b}}))\,\mathrm{d}t. \end{aligned} \tag{3.14}$$

If $r_2 = 0$, the boundary term is bounded by

$$\frac{A_1 A_2}{a^{r_1}} X^{n_1/a} \ln^{r_1} X,$$

otherwise it is 0. Consider the following integral

$$\begin{aligned} -I &= -\int_1^{\lfloor X^{1/a}\rfloor} M_1(t)\frac{\mathrm{d}}{\mathrm{d}t}(M_2(\frac{X}{t}))\,\mathrm{d}t \\ &= A_1 A_2 X^{\frac{n_2}{b}}(\frac{a}{b})\int_1^{\lfloor X^{1/a}\rfloor} t^{n_1-\frac{a}{b}n_2} \ln^{r_1} t \cdot (\frac{n_2}{b^{r_2}}\ln^{r_2}\frac{X}{t^a} + \frac{r_2}{b^{r_2-1}}\ln^{r_2-1}\frac{X}{t^a})\frac{\mathrm{d}t}{t} \\ &\leq A_1 A_2 X^{\frac{n_2}{b}}(\frac{1}{a^{r_1}b^{r_2}})\int_1^{X} t^{\frac{n_1}{a}-\frac{n_2}{b}} \ln^{r_1} t \cdot (\frac{n_2}{b}\ln^{r_2}\frac{X}{t} + r_2\ln^{r_2-1}\frac{X}{t})\frac{\mathrm{d}t}{t}. \end{aligned} \tag{3.15}$$

The integral is a sum of multiple pieces in the form of

$$I_{n,r_1,r_2} = \int_1^X t^n \ln^{r_1} t \ln^{r_2} \frac{X}{t} \frac{\mathrm{d}t}{t}.$$

It satisfies an induction formula

$$I_{n,r_1,r_2} = -\frac{r_1}{n} I_{n,r_1-1,r_2} + \frac{r_2}{n} I_{n,r_1,r_2-1} \tag{3.16}$$

with initial data

$$I_{n,r_1,0} \leq \frac{1}{n} X^n \ln^{r_1} X$$

$$I_{n,0,r_2} \leq \frac{r_2!}{n^{r_2+1}} X^n.$$

Notice that $I_{n,r_1,r_2}$ is always positive, by the induction formula one can show

$$I_{n,r_1,r_2} \leq \frac{r_2!}{n^{r_2+1}} X^n \ln^{r_1} X. \tag{3.17}$$

If $r_2 = 0$, $-I$ together with the boundary term is bounded,

$$P_{a,b}(X) \leq \frac{A_1 A_2}{a^{r_1}} \frac{n_1}{a} \frac{1}{\frac{n_1}{a} - \frac{n_2}{b}} X^{\frac{n_1}{a}} \ln^{r_1} X. \tag{3.18}$$

When $r_i \neq 0$, we have

$$P_{a,b}(X) \leq A_1 A_2 \frac{r_2!}{b^{r_2} a^{r_1}} \frac{n_1}{a} \frac{1}{(\frac{n_1}{a} - \frac{n_2}{b})^{r_2+1}} X^{\frac{n_1}{a}} \ln^{r_1} X. \tag{3.19}$$

This formula is compatible with the special cases where $r_i$ could be 0. □

Malle considered the compatibility of the conjecture under taking compositum in his original paper [Mal02] and estimates both the lower bound and upper bound of asymptotic distribution for compositum when the two Galois groups have no common quotient. By working out a product argument, we show a better lower bound in general.

**Corollary 3.6.** *Let $k$ be an arbitrary number field, and $G_1 \subset S_n$ and $G_2 \subset S_m$ be two Galois groups with nontrivial isomorphic quotient. Suppose Malle's conjecture holds for both groups, then there is a lower bound on $N(G_1 \times G_2 \subset S_{mn}, X)$ that*

$$N(G_1 \times G_2 \subset S_{mn}, X) \geq C X^a \ln^r X + o(X^a \ln^r X),$$

*where $a = max\{a(G_1)/m, a(G_2)/n\}$. If $a(G_1)/m = a(G_2)/n$, then $r = b(G_1, k)+b(G_2, k)-1$; if $a(G_1)/m > a(G_2)/n$, then $r = b(G_1, k) - 1$.*

A lower bound $X^a$ is also obtained in [Mal02] Proposition 4.2. Here we improve on the lower bound by adding a $\ln^r X$ factor. By analyzing the behavior of the discriminant carefully and applying good uniformity results, we show a better upper bound for our cases $S_n \times A$, see Theorem 1.1.1, which gives the same order of main term and actually matches Malle's prediction.

## 3.3 Proof of the Main Theorem

In this section, we prove our main results Theorem 1.1.1.

**Lemma 3.7.** *For $n = 3, 4, 5$, let $A$ be an abelian group satisfying the corresponding condition on $m = |A|$ in Theorem 1.1.1. Then $\forall c \in A$ and $k \in S_n$ ,*

$$\operatorname{ind}(k, c)/m - \operatorname{ind}(k) + r_k \geq 1 \tag{3.20}$$

*where the uniformity $O(X/|q|^{r_k})$ holds for $S_n$ degree $n$ extensions with $k$ as the inertia group at $p|q$.*

*Proof.* This can be checked by Lemma 3.1, 3.2 and 3.3 with Theorem 2.2.1, 2.2.2 and 2.2.3. □

Then we are going to prove the main results.

***Proof of Theorem 1.1.1.*** We will describe $S_n \times A$ number fields by pairs of $S_n$ degree $n$ field $K$ and $A$-number fields $L$

$$N(S_n \times A, X) = \sharp\{(K, L)|\operatorname{Gal}(K/k) \simeq S_n, \operatorname{Gal}(L/k) \simeq A, \operatorname{Disc}(KL) < X\}.$$

We will write $N(X)$ for short and omit the conditions $\operatorname{Gal}(K/k) \simeq S_n$ and $\operatorname{Gal}(L/k) \simeq A$ when there is no confusion. The equality holds since $S_n$ and odd abelian group have no isomorphic quotient. We will prove this result by three steps.

1. **Estimate pairs by** $\mathrm{Disc}(O_K O_L)$.

By Theorem 3.1.1, we can get a lower bound for $N(S_n \times A, X)$ by counting the number of pairs by $\mathrm{Disc}(O_K O_L)$. Denote $|A| = m$,

$$\begin{aligned} &N(S_n \times A, X) \\ &\geq \sharp\{(K,L) |\, \mathrm{Gal}(K/k) \simeq S_n, \mathrm{Gal}(L/k) \simeq A, \mathrm{Disc}(O_K O_L) = \mathrm{Disc}(K)^m \,\mathrm{Disc}(L)^n < X\}. \end{aligned} \tag{3.21}$$

By Lemma 3.5, there exists $C_0$ such that $N(S_n \times A, X) \geq C_0 X^{1/m}$ asymptotically. We can get a better understanding of the constant $C_0$ in view of Dirichlet series. Let $f(s)$ be the Dirichlet series of $S_n$ cubic number fields, and $g(s)$ be the Dirichlet series of $A$-number fields. Then the Dirichlet series for $\{(K,L)\}$ with respect to $\mathrm{Disc}(K)^m \,\mathrm{Disc}(L)^n$ is $f(ms)g(ns)$. The analytic continuation and pole behavior of $f$ and $g$ are both well studied [TT13, Wri89, Woo10]. It has been shown that $f(s)$ has the right most pole at $s = \frac{1}{\mathrm{ind}(S_n)} = 1$ and $g(s)$ has the right most pole at $s = \frac{1}{\mathrm{ind}(A)}$. Recall that for $A$ arbitrary abelian group, $\frac{m}{\mathrm{ind}(A)} = \frac{p}{p-1}$ where $p$ is the minimal prime divisor of $|A|$, so $\frac{1}{m} > \frac{1}{n\,\mathrm{ind}(A)}$. Therefore the right most pole of $f(ms)g(ns)$ is at $s = \frac{1}{m}$, and the order of the pole is exactly the order of the pole of $f(s)$ at $s = 1$, which is 1. By Tauberian Theorem[Nar83],

$$\liminf_{X\to\infty} \frac{N(S_n \times A, X)}{X^{1/m}} \geq \mathrm{Res}_{s=1} f \cdot g(\frac{n}{\mathrm{ind}(S_n)\cdot m}) = \mathrm{Res}_{s=1} f \cdot g(\frac{n}{m}). \tag{3.22}$$

2. **Estimate pairs by** $\mathrm{Disc}_Y(KL)$.

By using the idea of interpolation of discriminant in [BW08], we define $\mathrm{Disc}_Y$ to approximate Disc as follows:

$$\mathrm{Disc}_Y(KL) = \begin{cases} \mathrm{Disc}_p(KL) & |p| \leq Y \\ \mathrm{Disc}_p(K)^m \,\mathrm{Disc}_p(L)^n & |p| > Y. \end{cases} \tag{3.23}$$

Recall that $\mathrm{Disc}_p$ means the norm of $p$-factor in the discriminant, while $\mathrm{Disc}_Y$, as described above, is an approximation of Disc. The notation would be distinguished by whether the lower index is capital or little letter.

Define $N_Y(X) = \sharp\{(K, L) | \operatorname{Disc}_Y(KL) < X\}$. Since $\operatorname{Disc}_Y(KL) \geq \operatorname{Disc}(KL)$, as $Y$ gets larger, we get $N_Y(X) \leq N(X)$ which is an increasingly better lower bound for $N(X)$.

To compute $N_Y(X)$, denote the set of primes smaller than $Y$ to be $\{p_i\}$ with $i = 1, \cdots, n$. Let $\Sigma_1$ be a set containing a local étale extension over $k_{p_i}$ of degree $n$ for each $|p_i| < Y$ and $\Sigma = (\Sigma_1, \Sigma_2)$ contains a pair of local étale extension for each $p_i$. There are finitely many local étale extensions of degree $n$ and $m$, so there are finitely many different $\Sigma_i$'s and thus finitely many $\Sigma$'s for a certain $Y$. We will write $K \in \Sigma_1$ if for all $|p| \leq Y$ $K_p$ as a local étale extension is in $\Sigma_1$.

For each $\Sigma_1$, we know counting result of $S_n$ cubic field [BSW17] with finitely many local conditions

$$N_{\Sigma_1}(S_n, X) = \sharp\{K | \operatorname{Gal}(K/k) \simeq S_n, K \in \Sigma_1\}$$

and similarly for abelian extensions with in $\Sigma_2$[Mäk85, Wri89, Woo10].

We can relate $\operatorname{Disc}_Y(KL)$ and $\operatorname{Disc}(KL)$ for pairs $(K, L) \in \Sigma$,

$$\begin{aligned}\operatorname{Disc}_Y(KL) &= \prod_{|p|\leq Y} \operatorname{Disc}_p(KL) \prod_{|p|>Y} \operatorname{Disc}_p(K)^m \operatorname{Disc}_p(L)^n \\ &= \operatorname{Disc}(K)^m \operatorname{Disc}(L)^n \prod_{|p|\leq Y} \operatorname{Disc}_p(KL) \operatorname{Disc}_p(K)^{-m} \operatorname{Disc}_p(L)^{-n} \\ &= \frac{\operatorname{Disc}(K)^m \operatorname{Disc}(L)^n}{d_\Sigma}\end{aligned} \tag{3.24}$$

where $d_\Sigma$ is a factor only depending on $\Sigma$. We have seen in section 2 that at tamely ramified primes, $\operatorname{Disc}_p(KL)$ can be determined by inertia groups of $\tilde{K}$ and $\tilde{L}$, therefore it depends on $\Sigma$ at $p$. For wildly ramified primes, it suffices to see that $\operatorname{Disc}_p(KL)$ could be determined by $K_p$ and $L_p$. This is always true under taking product: if $\tilde{K}$ and $\tilde{L}$ have trivial intersection, we can get the map from absolute local Galois group $G_{k_p}$ to $S_n \times A$ by taking the product of such maps to $S_n$ and $A$. Then we get the precise local information for $KL$ including $\operatorname{Disc}_p(KL)$.

Therefore $\operatorname{Disc}_Y(KL) \leq X$ is equivalent to $\operatorname{Disc}(K)^m \operatorname{Disc}(L)^n \leq d_\Sigma X$ for $(K, L) \in \Sigma$. Apply Lemma 3.5 to $N_{\Sigma_1}(S_n, X)$ and $N_{\Sigma_2}(A, X)$, we get

$$\lim_{X\to\infty} \frac{N_Y(X)}{X^{1/m}} = C_Y. \tag{3.25}$$

For each $Y$, $N_Y(X) \leq N(X)$, therefore

$$\lim_{Y\to\infty}\lim_{X\to\infty}\frac{N_Y(X)}{X^{1/m}} = \lim_{Y\to\infty} C_Y \leq \liminf_{X\to\infty}\frac{N(X)}{X^{1/m}}. \tag{3.26}$$

By definition of $N_Y$, $C_Y$ is monotonically increasing as $Y$ increases and will be shown to be uniformly bounded in next step. So this limit does exist and gives a lower bound.

**3. Bound** $N(X) - N_Y(X)$

Our goal is to prove the other direction of the inequality 3.25.

$$\lim_{Y\to\infty} C_Y \geq \limsup_{X\to\infty}\frac{N(X)}{X^{1/m}}, \tag{3.27}$$

and thus

$$\lim_{X\to\infty}\frac{N(X)}{X^{1/m}} = \lim_{Y\to\infty}\lim_{X\to\infty}\frac{N_Y(X)}{X^{1/m}} = \lim_{Y\to\infty} C_Y. \tag{3.28}$$

To get an upper bound of $N(X)$ via $N_Y(X)$, we need to bound on $N(X)-N_Y(X)$. It suffices to show the difference is $o(X^{1/m})$. There are only finitely many wildly ramified primes, so they would only affect the constant but not the order.

$$\begin{aligned} N(X) - N_Y(X) &= \sharp\{(K,L)|\operatorname{Disc}(KL) < X < \operatorname{Disc}_Y(KL)\} \\ &= \sum_{\Sigma'}\sharp\{(K,L)\in\Sigma'|\operatorname{Disc}(KL) < X < \operatorname{Disc}_Y(KL)\} \end{aligned} \tag{3.29}$$

where the local condition $\Sigma'$ is a little bit different from $\Sigma$ in last part. Each $\Sigma'$ specifies a finite set of primes $S = \{p_j\}$ and a pair of inertia groups at tame $p$ and a pair of ramified local étale extensions at wildly ramified $p$ for each $p$ in S. Denote the pair of local information by $(h_j, g_j)$ for each $p_j$. We will not write the index $j$ each time when there is no confusion. We write $(K,L)\in\Sigma'$ if $K_p$ and $L_p$ are in $\Sigma'$ for each $p\in S$, and are not ramified simultaneously outside $S$. Denote $\exp(\cdot)$ to be the corresponding exponent of $p$ in discriminant. At tame place, $\exp(\cdot)$ is equal to $\operatorname{ind}(\cdot)$ as described before. For $(K,L)\in\Sigma'$, we can relate precise $\operatorname{Disc}(KL)$ to the product,

$$\begin{aligned} \operatorname{Disc}(KL) =& \operatorname{Disc}(K)^m \operatorname{Disc}(L)^n \prod_{p\in S}|p|^{\exp(h_j,g_j)-m\cdot\exp(h_j)-n\cdot\exp(g_j)} \\ =& \frac{\operatorname{Disc}(K)^m \operatorname{Disc}(L)^n}{d_{\Sigma'}}. \end{aligned} \tag{3.30}$$

Each $\Sigma'$ summand is

$$\begin{aligned}
&\sharp\{(K,L)\in\Sigma' |\operatorname{Disc}(KL)<X<\operatorname{Disc}_Y(KL)\}\\
\leq&\sharp\{(K,L)\in\Sigma' |\operatorname{Disc}(KL)<X\}\\
=&\sharp\{(K,L)\in\Sigma' |\operatorname{Disc}(K)^m\operatorname{Disc}(L)^n<Xd_{\Sigma'}\}\\
=&\sharp\{(K,L)\in\Sigma' |\prod_{p\notin S}\operatorname{Disc}_p(K)^m\operatorname{Disc}_p(L)^n<\frac{X}{\prod_{p\in S}|p|^{\exp(h_j,g_j)}}\}.
\end{aligned} \tag{3.31}$$

Notice only $\Sigma'$ summand where $\prod_{p\in S}|p|>Y$ is non-zero. Denote $\prod_{p\notin S}\operatorname{Disc}_p(K)$ by $\operatorname{Disc}_{res}(K)$. For a certain $\Sigma'$, define $q_k=\prod'_{p\in S,I_p=<k>}p$ where $\prod'$ means the product is taken only over tamely ramified $p$ in $\Sigma'$. Similarly we write $K\in\Sigma'$ if $K$ satisfies the local conditions specified at $S$ in $\Sigma'$. Then we can bound the number of $K\in\Sigma'$

$$\begin{aligned}
&\sharp\{K|K\in\Sigma',\operatorname{Disc}_{res}(K)\leq X\}\\
=&\sharp\{K|K\in\Sigma',\operatorname{Disc}(K)\leq X\prod_{p\in S}|p|^{\exp(h_j)}\}\\
=&O_\epsilon\left(\prod_k|q_k|^{-r_k}\prod_{p\in S}|p|^{\exp(h_j)}\right)X\\
=&O_\epsilon\left(\prod_k|q_k|^{-r_k+\operatorname{ind}(k)}\right)X.
\end{aligned} \tag{3.32}$$

Here we can ignore wildly ramified primes since there are only finitely many wildly ramified primes and finitely many wildly ramified local étale extensions. Hence the discriminant at those primes are uniformly bounded by some constant. Similarly,

$$\begin{aligned}
&\sharp\{L|L\in\Sigma',\operatorname{Disc}_{res}(L)\leq X\}\\
=&\sharp\{L|L\in\Sigma',\operatorname{Disc}(L)\leq X\prod_{p\in S}|p|^{\exp(g_j)}\}\\
=&O_\epsilon\left((\prod_{p\in S}|p|^{\exp(g_j)})^\epsilon\right)X^{1/a(A)}\ln^{b(A)}.
\end{aligned} \tag{3.33}$$

Now apply Lemma 3.5 to (3.31),

$$\begin{aligned}
&\sharp\{(K,L)\in\Sigma' \mid \operatorname{Disc}_{res}(K)^m \operatorname{Disc}_{res}(L)^n < \frac{X}{\prod_{p\in S}|p|^{\exp(h_j,g_j)}}\} \\
\leq &O_\epsilon\left(\prod_k |q_k|^{-r_k+\operatorname{ind}(k)+\epsilon}\right)\left(\frac{X}{\prod_{p\in S}|p|^{\exp(h_j,g_j)}}\right)^{1/m} \\
\leq &O_\epsilon\left(\prod_k |q_k|^{-r_k+\operatorname{ind}(k)+\epsilon-\operatorname{ind}(k,g_j)/m}\right)X^{1/m}.
\end{aligned} \tag{3.34}$$

Each $\Sigma'$ gives a list of $(q_k)$ of relatively prime ideals. Conversely, for each list $(q_k)$, there are at most $M^{\omega(\prod_k q_k)} = O_\epsilon(\prod_k q_k)^\epsilon$ many $\Sigma'$s, where $M$ is an upper bound of the number of possible tame inertia groups for $A$-extensions, then

$$\begin{aligned}
N(X)-N_Y(X) &\leq \sum_{\Sigma'} \sharp\{(K,L)\in\Sigma' \mid \operatorname{Disc}_{res}(K)^m \operatorname{Disc}_{res}(L)^n \leq \frac{X}{\prod_{p\in S}|p|^{\exp(h_i,g_i)}}\} \\
&\leq X^{1/m} O_\epsilon\left(\sum_{(q_k),\prod_k|q_k|>Y}\ \prod_k |q_k|^{\delta}\right) \\
&\leq X^{1/m} O_\epsilon\left(\sum_{|q|>Y} |q|^{\delta+\epsilon}\right)
\end{aligned} \tag{3.35}$$

where the for every $k$, the exponent $\delta$ is strictly smaller than $-1$ by Lemma 3.7 and $\epsilon$ is arbitrary small. Therefore the summation is convergent and $N(X)-N_Y(X)$ is $O(X^{1/m})$ which proves the boundedness of $C_Y$. Moreover,

$$\lim_{Y\to\infty}\limsup_{X\to\infty}\frac{N(X)-N_Y(X)}{X^{1/m}} \leq \lim_{Y\to\infty}\sum_{|q|>Y} O_\epsilon(|q|^{\delta+\epsilon}) = 0, \tag{3.36}$$

therefore it proves that

$$\begin{aligned}
\limsup_{X\to\infty}\frac{N(X)}{X^{1/m}} &\leq \lim_{Y\to\infty}\left(\lim_{X\to\infty}\frac{N_Y(X)}{X^{1/m}} + \limsup_{X\to\infty}\frac{N(X)-N_Y(X)}{X^{1/m}}\right) \\
&= \lim_{Y\to\infty} C_Y.
\end{aligned} \tag{3.37}$$

□

# Chapter 4

# Asymptotic Distribution of $S_3 \times A$ Extensions over $\mathbb{Q}$-Power Saving Error and Secondary Term

In this chapter we go through the author's work on proving the secondary term for $S_3 \times A$ extensions over $\mathbb{Q}$ for most abelian groups $A$. For all odd abelian groups $A$ where we get a main term, we prove a power saving error.

We will give the outline of the proof in section 4.1. Then we will compute carefully what the error terms are for each step of summation in section 4.2 and 4.3. In section 4.4 we will determine the tail estimates based on the uniformity estimates. In section 4.5 we will put all the estimates together and balance between the small range and the large range to optimize the exponent of the power saving error. In section 4.6, we compute the group theory data required as the final input to prove Theorem 1.2.3 and 1.2.4. In section 4.7, as an example, we give the precise expression of the constant in the main term and the secondary term for $S_3 \times C_l$ extensions where $l$ is a prime number. In section 4.8, we describe the amount of power saving away from the secondary term for cases in Theorem 1.2.3, and the amount of power saving away from the main term for cases in Theorem 1.2.4.

## 4.1 Framework

In this section, we are going to give a framework of the proof. Let $K$ be an $S_3$ cubic extension over $\mathbb{Q}$, and $L$ be an $A$ extension over $\mathbb{Q}$. Let $T$ be the set of all primes that divide $6|A|$. Define $\Sigma_p$ as follows: if $p \notin T$, let $\Sigma_p$ be the set of all possible non-trivial inertia groups for an $S_3$ cubic extensions up to conjugation; if $p \in T$, then let $\Sigma_p$ be the set of all possible local étale extensions over $\mathbb{Q}_p$ for an $S_3$ cubic extension. Similarly, we define $\Lambda_p$ for $A$-extensions at $p \notin T$ and $p \in T$ separately. Therefore define $\mathcal{A} = \{\langle(12)\rangle, \langle(123)\rangle\}$ and $\mathcal{B} = \{\langle a\rangle \mid a \neq e \in A\}$, then $\Sigma_p = \mathcal{A}$ and $\Lambda_p = \mathcal{B}$ for $p \notin T$. We will write $K \in \sigma_p$ for a certain $\sigma_p \in \Sigma_p$: if $K_p$ is isomorphic to $\sigma_p$ at $p \in T$, or if $\tilde{K}$ has $\sigma_p$ as the inertia group at $p \notin T$. Similarly for $L$. By the way $\Sigma_p$ and $\Lambda_p$ are defined, all $K \in \sigma_p$ have the same discriminant, $\operatorname{Disc}_p(K)$, so we could denote this number $\operatorname{Disc}(\sigma_p)$. Similarly for $\operatorname{Disc}(\lambda_p)$ for $A$ extensions.

Given a pair of extensions $(K, L)$ where $\operatorname{Gal}(K) = S_3$ and $\operatorname{Gal}(L) = A$, by section 2.1 and Theorem 3.1.3 we would be able to determine $\operatorname{Disc}_p(KL)$. At a certain $p$, say $K \in \sigma_p$ and $L \in \lambda_p$, then denote $\operatorname{Disc}(\sigma_p, \lambda_p)$ to be the local discriminant determined by the pair, and define $e(\sigma_p, \lambda_p)$ as

$$p^{e(\sigma_p,\lambda_p)} = \frac{\operatorname{Disc}(\sigma_p)^m \operatorname{Disc}(\lambda_p)^n}{\operatorname{Disc}(\sigma_p, \lambda_p)}.$$

The exponent $e(\sigma_p, \lambda_p)$ for $p \notin T$ could be determine by Theorem 3.1.3, and in such cases $e(\sigma_p, \lambda_p)$ is independent of $p$ and only depends on the permutation presentation of $\sigma_p$ and $\lambda_p$.

Denote the set $\mathcal{S} = \mathcal{A} \times \mathcal{B} = \{s_{ij} \mid s_{ij} = (a_i, b_j), a_i \in \mathcal{A}, b_j \in \mathcal{B}\}$ to be the direct product of $\mathcal{A}$ and $\mathcal{B}$, and the set $\mathcal{W} = \prod_{p\in T}(\Sigma_p \times \Lambda_p) = \{w \mid \forall p \in T, w_p = (\sigma_p, \lambda_p) \in \Sigma_p \times \Lambda_p\}$. Here $\mathcal{S}$ lists all possible ramification types for a pair $(K, L)$ at tamely ramified places, and $\mathcal{W}$ lists all possible local étale extensions for a pair at wildly ramified places. Denote $\rho = (w, q_{ij})$ to be one element $w = \prod_{p\in T}(\sigma_p, \lambda_p) \in \mathcal{W}$ and a tuple of square-free numbers $\rho = (q_{ij})$, where for each $1 \leq i \leq |\mathcal{A}|$ and $1 \leq j \leq |\mathcal{B}|$, and each $p|q_{ij}$ we have $p \notin T$, and $\prod_{i,j} p_{ij}$ is also square-free. For each $\alpha = (f_{ij}) \in (\mathbb{Z}/2\mathbb{Z})^{|\mathcal{S}|}$, we define $(K, L) \in \rho^\alpha$ as follows: 1) the pair $(K, L)$ satisfies the condition $w$ at all $p \in T$; 2) at each $p|q_{ij}$, we require $K \in a_i$ and $L \in b_j$; 3) if $f_{ij} = 0$, then we require further that $p|q_{ij}$ are the only primes that $(K, L)$ are simultaneously in $a_i$ and $b_j$.

Define

$$B(\rho^\alpha, X) = \sharp\{(K, L) \mid (K, L) \in \rho^\alpha, \operatorname{Disc}(K)^m \operatorname{Disc}(L)^3 \leq X\},$$

where $m = |A|$. If $\alpha = 0 \in (\mathbb{Z}/2\mathbb{Z})^{|S|}$, then we get for $(K, L) \in \rho^0$ that

$$\mathrm{Disc}(KL) = \frac{\mathrm{Disc}(K)^m \,\mathrm{Disc}(L)^3}{\prod_{p\in T} p^{e(\sigma_p,\lambda_p)} \prod_{i,j} q_{ij}^{e(a_i,b_j)}} = \frac{\mathrm{Disc}(K)^m \,\mathrm{Disc}(L)^3}{L_\rho}.$$

Therefore if we could get an estimation of $B(\rho^0, X)$ for every $\rho$, we just need to sum $B(\rho^0, XL_\rho)$ over all $\rho$ in this form to get the final counting

$$G(X) = \sharp\{(K, L) \mid \mathrm{Disc}(KL) \le X\} = \sum_{\rho} B(\rho^0, XL_\rho).$$

In order to get $B(\rho^0, X)$, we apply a sieve method. We will say that $\rho_1$ *divides* $\rho_2$ if they contain the same $w \in \mathcal{W}$, and for each $i$ and $j$, we have $q_{ij}^{(1)} | q_{ij}^{(2)}$, where $q_{ij}^{(k)}$ is the associated square-free number at the $i, j$-th position in $\rho_k$. Given $\rho_1$ and $\rho_2$ with the same $w \in \mathcal{W}$, we can also multiply to get a new tuple $(\rho_1\rho_2)_{ij} = (q_{ij}^{(1)} q_{ij}^{(2)})$ when it is legal, i.e., when the product $\prod_{i,j} q_{ij}^{(1)} q_{ij}^{(2)}$ is still square-free. By inclusion-exclusion, we have the following relation

$$B(\rho^0, X) = \sum_{\varrho=\rho\eta} \mu(\eta) B(\varrho^1, X), \tag{4.1}$$

where we define $\mu(\eta)$ to be $\prod_{i,j} \mu(q_{ij})$ with $q_{ij}$ the $i, j$-th square-free integer in $\eta$. Here we write $\varrho^1$ in short for $\varrho^{(1,1,\cdots,1)}$, which means that we require no condition on places outside $q_{ij}$ in $\varrho$. So we can apply product argument to distributions of $S_3$ cubic extensions and $|A|$-extensions with local conditions to get $B(\varrho^1, X)$. When $\rho$ involves some big primes, we will apply uniformity estimates to get a tail estimation on $B(\rho^0, X)$ and use that instead.

## 4.2 Estimates of $B(\varrho^1, X)$

In this section, we are going to compute the product distribution of $S_3$ cubic extensions and $A$-extensions, in addition with local conditions on ramification. Our computation heavily relies on the following theorem in [BTT16], which improves previous results on distribution of $S_3$ cubic extensions with local density [TT13]. On one hand it reduces the exponent of $X$ in error terms, and on the other hand, it also reduces the dependency of local parameters in the constant of the error terms.

**Theorem 4.2.1** ([BTT16], Theorem 4.3)**.** *The number of $S_3$ cubic extension that are partially ramified at $q = \prod p_i$ and totally ramified at $r = \prod p_j$ are estimated to be*

$$N_{q,r}(S_3, X) = AA_qA_{r^2}X + BB_qB_{r^2}X^{5/6} + O(C_qC_{r^2}X^{2/3+\epsilon}),$$

*where the constants $A_n$, $B_n$ and $C_n$ are some multiplicative arithmetic functions.*

We record the above densities in the form as we need them. For a complete table of every local condition, please see (6.8) and page 2487 in [TT13]. However we only need the local density on ramified cubic fields. In these cases, for each prime number $p$,

$$A_p = \frac{C_p^{-1}}{p}, \quad A_{p^2} = \frac{C_p^{-1}}{p^2},$$

where $C_p = 1 + p^{-1} + p^{-2}$ is the normalizing factor, and

$$B_p = \frac{K_p^{-1}(1 + p^{-1/3})^2}{p}, \quad B_{p^2} = \frac{K_p^{-1}(1 + p^{-1/3})}{p^2},$$

where $K_p = \frac{(1-p^{-5/3})(1+p^{-1})}{1-p^{-1/3}}$ is the normalizing factor, and

$$C_p = p^{4/5}, \quad C_{p^2} = p^{4/5}.$$

The constants are $A = \frac{1}{3\zeta(3)}$ and $B = (1 + \sqrt{3})\frac{4\zeta(1/3)}{5\Gamma(2/3)^3\zeta(5/3)}$.

We will not need the precise expression of these constants until section 4.7 where we compute explicit expression of the constants. The important input from this theorem for us is that the order of $C_qC_{r^2} \le O(\prod_{e|q} e^{4/5} \prod_{e|r} e^{4/5}) \le O(qr)^{4/5+\epsilon}$. We will keep this fact in mind, but write $C_q$ and $C_{r^2}$ on the way.

Another input we need is counting results from abelian extensions. Number of abelian extensions with local density are studied in [Mäk85, Wri89, Woo10]. We will mainly use the estimate of abelian extension in the form of the uniformity estimates. Denote $N_q(A, X)$ to be the number of $A$-extensions $L$ over $\mathbb{Q}$ such that $q|\operatorname{Disc}(L)$, then recall from Theorem 2.1.2 we have the following estimate

$$N_q(A, X) \le O(\frac{C^{\omega(q)}}{q^{1/a(A)}})X^{1/a(A)}(\ln X)^{b(k,A)-1}$$

where $\omega(q)$ is the number of prime divisors of $q$.

As for the notation, we denote the Dirichlet series $f(s)$ for $S_3$ cubic fields

$$f(s) = \sum_{\mathrm{Gal}(K/\mathbb{Q})=S_3} \frac{1}{\mathrm{Disc}(K)^s},$$

and denote $F_1(X) = \sum_{\mathrm{Disc}(K)\leq X} 1$. Given a local condition $\Sigma$ which contains $\sigma_p$ at finitely many primes, we write $K \in \Sigma$ if $K_p \in \Sigma$ at those places. We denote

$$f_\Sigma(s) = \sum_{\mathrm{Gal}(K/\mathbb{Q})=S_3, K\in\Lambda} \frac{1}{\mathrm{Disc}(K)^s},$$

and denote $F_{1,\Sigma}(X) = \sum_{\mathrm{Disc}(K)\leq X, K\in\Sigma} 1$. Similarly for $A$ extensions, we will use $F_2(X)$, $F_{2,\Lambda}(X)$, $g(s)$ and $g_\Lambda(s)$. Then Theorem 2.1.2 says that given a local condition $\Lambda$ on ramification behavior, we have that

$$F_{2,\Lambda}(X) = O_\epsilon(D_\Lambda)X^{1/a(A)+\epsilon}, \tag{4.2}$$

where $D_\Lambda$ can be bounded by the order of $O(\frac{C^{\omega(q)}}{|q|^{1/a(A)}})$ with $q$ associated with $\Lambda$. For brevity we will write $O$ instead of $O_\epsilon$ since $O_\epsilon$ only depends on $\epsilon$ and is independent of $\Lambda$.

Notice that given a tuple $\varrho$ of local conditions as defined before, there will be naturally induced local condition on $K$ and $L$, called $\Sigma(\varrho)$ and $\Lambda(\varrho)$. The local conditions that come from the same $\varrho$ have the same support outside $T$ with non-trivial ramification restriction. Indeed, say $q_{ij}$ are the $i, j$-th square-free number in $\varrho$, then $\Sigma(\varrho)$ restricts the counting to $S_3$ cubic extensions that are partially ramified at $q_1 = \prod_j q_{ij}$ and are totally ramified at $q_2 = \prod_j q_{2j}$. Similarly for $A$ extensions we have $q_j$ for $1 \leq j \leq |\mathcal{B}|$. In addition, at primes in $T$, $\Sigma(\rho)$ (and $\Lambda(\rho)$) also contains the corresponding $\sigma_p$ (and $\lambda_p$) in $w \in \mathcal{W}$. For technical reasons, if we do not include the conditions at $T$ into $\Sigma$, then we denote the smaller local condition $\Sigma'$. For brevity, we will call the corresponding coefficients depending on $\Sigma$ and $\Lambda$ in Theorem 4.2.1 and 2.1.2 by $A_\Sigma$, $B_\Sigma$, $C_\Sigma$ and $D_\Lambda$ in short. If we restrict the local condition to places outside $T$, we get the corresponding coefficient $A_{\Sigma'}$, $B_{\Sigma'}$, $C_{\Sigma'}$ and $D_{\Lambda'}$. Then if $\varrho = \rho\eta$, then $A_{\Sigma(\varrho)} = A_{\Sigma(\rho)}A_{\Sigma'(\eta)}$.

Recall that

$$B(\varrho^1, X) = \sharp\{(K, L) \mid (K, L) \in \varrho^1, \mathrm{Disc}(K)^m \,\mathrm{Disc}(L)^3 \leq X\},$$

and $\varrho^1$ naturally gives a set of local specification $\Sigma(\varrho)$ for $K$ and $\Lambda(\varrho)$ for $L$, so equivalently

$$B(\varrho^1, X) = \sharp\{(K, L) \mid K \in \Sigma(\varrho), L \in \Lambda(\varrho), \mathrm{Disc}(K)\,\mathrm{Disc}(L)^{3/m} \leq X^{1/m}\},$$

which is the product distribution of $F_{1,\Sigma}(X)$ and $F_{2,\Lambda}(X)$.

Let's say $f_\Sigma(s) = \sum_n a_n \cdot n^{-s}$ and $g_\Lambda(s) = \sum_k b_k \cdot k^{-s}$. Then we have that

$$\begin{aligned}
B(\varrho^1, X) = &\sum_{\substack{K\in\Sigma, L\in\Lambda \\ \mathrm{Disc}(K)^m\,\mathrm{Disc}(L)^3 \leq X}} 1 = \sum_{k^3 \leq X} b_k F_{1,\Sigma}(\frac{X^{1/m}}{k^{3/m}}) \\
= &\sum_{k \leq X^{1/3}} b_k \left( AA_\Sigma \frac{X^{1/m}}{k^{3/m}} + BB_\Sigma(\frac{X^{1/m}}{k^{3/m}})^{5/6} + O(C_\Sigma(\frac{X^{1/m}}{k^{3/m}})^{2/3}) \right) \\
= &AA_\Sigma X^{1/m} \sum_{k \leq X^{1/3}} \frac{b_k}{k^{3/m}} + BB_\Sigma X^{5/6m} \sum_{k \leq X^{1/3}} \frac{b_k}{k^{3/m\cdot 5/6}} \\
&+ O(C_\Sigma X^{2/3m} \sum_{k \leq X^{1/3}} \frac{b_k}{k^{3/m\cdot 2/3}}) \\
= &AA_\Sigma \cdot g_\Lambda(\frac{3}{m}) X^{1/m} + BB_\Sigma \cdot g_\Lambda(\frac{5}{2m}) X^{5/6m} + O(C_\Sigma \cdot g_\Lambda(\frac{2}{m})) X^{2/3m} \\
&+ AA_\Sigma X^{1/m} \sum_{k \geq X^{1/3}} \frac{b_k}{k^{3/m}} + BB_\Sigma X^{5/6m} \sum_{k \geq X^{1/3}} \frac{b_k}{k^{3/m\cdot 5/6}} \\
&+ O(C_\Sigma X^{2/3m} \sum_{k \geq X^{1/3}} \frac{b_k}{k^{3/m\cdot 2/3}}).
\end{aligned} \tag{4.3}$$

Notice that in the last equality, we can take those values of $g_\Lambda(s)$ at $s = 3/m, 5/2m, 2/m$ since the right most pole of $g_\Lambda(s)$ is at $s = \frac{1}{\mathrm{ind}(A)}$, which is smaller than $\frac{2}{m}$. Aside from the first two precise terms which will be the main term and the secondary term, we will denote the following errors $E_1$, $E_a$, $E_b$ and $E_c$ and analyze them one by one.

### 4.2.1 Bound on $E_i$ for $i = a, b, c$

Let's first look at $E_a$. It suffices to bound the following weighted sum of $b_k$. By Abel summation,

$$\begin{aligned} E_a = AA_\Sigma X^{1/m} \sum_{k \geq X^{1/3}} \frac{b_k}{k^{3/m}} &= AA_\Sigma X^{1/m} \left( -\frac{F_{2,\Lambda}(X^{1/3})}{X^{1/m}} + \frac{3}{m} \int_{X^{1/3}}^{\infty} \frac{F_{2,\Lambda}(t)}{t^{3/m+1}} dt \right) \\ &= O(A_\Sigma D_\Lambda) X^{1/3a(A)+\epsilon}. \end{aligned} \tag{4.4}$$

Similarly, we get for $E_b$ that

$$E_b = O(B_\Sigma D_\Lambda) X^{1/3a(A)+\epsilon}, \tag{4.5}$$

and for $E_c$ that

$$E_c = O(C_\Sigma D_\Lambda) X^{1/3a(A)+\epsilon}. \tag{4.6}$$

By Theorem 4.2.1, $A_\Sigma$ and $B_\Sigma$ are precise constants determined and are the local densities at $s = 1$ and $s = 5/6$, while $C_\Sigma$ is the upper bound of the dependency for the error in the order of $(qr)^{4/5+\epsilon}$. So $E_c$ is the biggest one among $E_a$, $E_b$ and $E_c$, and we can combine them

$$E_2 = E_a + E_b + E_c \leq O(C_\Sigma D_\Lambda) X^{1/3a(A)+\epsilon}. \tag{4.7}$$

### 4.2.2 Bound on $E_1$

To bound

$$E_1 = O(C_\Sigma \cdot g_\Lambda(\frac{2}{m})) X^{2/3m},$$

it suffices to give a bound on $g_\Lambda(\frac{2}{m})$. From now on, we denote $b_j$ to be the generator of a tamely ramified inertia group, i.e. $\langle b_j \rangle \in \Lambda_p = \mathcal{B}$ for $p \notin T$. When we write $\operatorname{ind}(b_j)$, we mean the index of the group element.

**Lemma 4.1.** *Let $\Lambda$ be a local condition on ramification for $A$ extensions and let $g_\Lambda$ be the corresponding Dirichlet series, we have that at $s > 1/a(A)$, the value of $g_\Lambda(s)$ is bounded by*

$$g_\Lambda(s) \leq O(\prod_j \prod_{p|q_j} \frac{C}{p^{\operatorname{ind}(b_j)s}}) \leq O(D_\Lambda)^{a(A)s},$$

*where $q_j$ is the product of primes where the inertia group is $\langle b_j \rangle \in \mathcal{B}$. The implied constant depends on $s$ but not on $\Lambda$.*

*Proof.* Denote $J_{\mathbb{Q}}$ to be the idèle group of $\mathbb{Q}$. Notice that we can bound the number of A-extensions by the number of continous homomorphisms $J_{\mathbb{Q}}/\mathbb{Q}^* \to A$, and it is equivalent to consider maps $\rho : \prod_p \mathbb{Z}_p^* \to A$ [Woo16]. Local conditions on the abelian extensions could also be formulated by local conditions on $\rho$. At places outside $T$, the condition of $\Lambda$ is equivalent to the condition that the image of $\mathbb{Z}_p^\times$ under $\rho$ in $A$ is exactly specified as $\lambda_p \in \Lambda_p$. So we can also write $\rho \in \Lambda$ to specify the local condition on $\rho$. We then have

$$g_\Lambda(s) = \sum_{K \in \Lambda} \frac{1}{\mathrm{Disc}(K)^s} \leq \sum_{\rho:\rho\in\Lambda} \frac{1}{\mathrm{Disc}(\rho)^s} = \prod_p \Big( \sum_{\rho_p:\mathbb{Z}_p^\times \to A, \rho_p \in \Lambda} \frac{1}{\mathrm{Disc}(\rho_p)^s} \Big) = \tilde{g}_\Lambda(s).$$

If $s > 1/a(A)$, then $g(s)$ and $\tilde{g}_\Lambda(s)$ are both convergent by [Mäk85, Wri89, Woo10]. Also since $\tilde{g}$ and $\tilde{g}_\Lambda$ are both multiplicative, we can get the estimate for $\tilde{g}_\Lambda$ easily,

$$\tilde{g}_\Lambda(s) = \tilde{g}(s) \cdot \frac{\prod_p (\sum_{\rho_p:\mathbb{Z}_p^\times \to A, \rho_p\in\Lambda} \frac{1}{\mathrm{Disc}(\rho_p)^s})}{\prod_p (\sum_{\rho_p:\mathbb{Z}_p^\times \to A} \frac{1}{\mathrm{Disc}(\rho_p)^s})} \leq \tilde{g}(s) \cdot O(D_\Lambda)^{a(A)s}.$$

□

Plugging in the value from Lemma 4.1, we get that

$$E_1 = O(C_\Sigma \cdot D_\Lambda^{2a(A)/m}) X^{2/3m}. \tag{4.8}$$

Comparing with (4.7), we have that

$$E = E_2 + E_1 \leq O(C_\Sigma \cdot D_\Lambda^{2a(A)/m}) X^{2/3m}.$$

## 4.3 Estimates of $B(\rho^0, XL_\rho)$ for Small $\rho$

In this section, we are going to compute the error for $B(\rho^0, X)$ which only involves small primes. Recall in (4.1) that

$$B(\rho^0, X) = \sum_{\varrho = \rho\eta} \mu(\eta) B(\varrho^1, X),$$

where we define $\mu(\eta)$ to be $\prod_{i,j}\mu(k_{ij})$ with $k_{ij}$ is the $i,j$-th square-free integer in $\eta$. We expect the main terms from $B(\varrho^1, X)$ to contribute to the main term, so we will only look at the error terms. Denote $\Sigma(\rho)$ to be $\Sigma$ induced by $\rho$ and similarly for $\Lambda(\rho)$. Like we define $\operatorname{Disc}(\sigma_p)$, we could also define $\operatorname{Disc}(\Sigma(\rho))$ to be $\prod_p \operatorname{Disc}(\sigma_p)$ where the product is over all $p \in T$ and $p|\rho_{ij}$ for all $i$ and $j$, and define $\operatorname{Disc}(\Sigma'(\rho))$ to be $\prod_p \operatorname{Disc}(\sigma_p)$ where the product is over all $p|\rho_{ij}$ for all $i$ and $j$, then $\operatorname{Disc}(\Sigma(\rho\eta)) = \operatorname{Disc}(\Sigma(\rho))\operatorname{Disc}(\Sigma'(\eta))$.

Notice that $B(\varrho^1, X) = 0$ when $\eta$ involves primes that are too large since

$$\operatorname{Disc}(\Sigma'(\eta))^m \operatorname{Disc}(\Lambda'(\eta))^3 = \prod_{i,j} k_{ij}^{m\operatorname{ind}(a_i)+3\operatorname{ind}(b_j)} = k^\beta \leq X^* = \frac{X}{\operatorname{Disc}(\Sigma(\rho))^m \operatorname{Disc}(\Lambda(\rho))^3}. \tag{4.9}$$

For brevity we write $k = (k_{ij})$ as a vector and $\beta = (\beta_{ij}) = (m\operatorname{ind}(a_i) + 3\operatorname{ind}(b_j))$ as a vector of exponent. So there are two sources of error: one comes from the small $\eta$ where we apply the sieve; and the other one comes from the big $\eta$ where we pretend to have precise terms.

For small $\eta$, by the inclusion-exclusion, the error is

$$\begin{aligned} W_1 &= \sum_{\substack{\varrho=\rho\eta \\ k^\beta < X^*}} \mu(\eta) O(C_{\Sigma(\varrho)} \cdot D_{\Lambda(\varrho)}^{2a(A)/m}) X^{2/3m} \\ &\leq O(C_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{2a(A)/m}) \cdot X^{2/3m} \sum_{k^\beta < X^*} C_{\Sigma(\eta)} D_{\Lambda(\eta)}^{2a(A)/m}. \end{aligned} \tag{4.10}$$

The last inequality comes from the fact that $C_\Sigma$ and $D_\Lambda$ are multiplicative up to $O(1)$ at most.

For big $\eta$, although $B(\varrho^1, X) = 0$, we would still like to use the main term and the secondary term in the same form. In order to compensate for that, we have the error coming from the main term

$$W_2 = \sum_{\substack{\varrho=\rho\eta \\ k^\beta > X^*}} O(A_{\Sigma(\varrho)} \cdot g_{\Lambda(\varrho)}(\frac{3}{m}) X^{1/m}) \leq O(A_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{3a(A)/m}) \cdot X^{1/m} \sum_{k^\beta > X^*} A_{\Sigma(\eta)} D_{\Lambda(\eta)}^{3a(A)/m}, \tag{4.11}$$

and similarly for the secondary term,

$$W_3 = \sum_{\substack{\varrho=\rho\eta \\ k^\beta > X^*}} O(B_{\Sigma(\varrho)} \cdot g_{\Lambda(\varrho)}(\frac{5}{2m})) X^{5/6m} \leq O(B_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{5a(A)/2m}) \cdot X^{5/6m} \sum_{k^\beta > X^*} B_{\Sigma(\eta)} D_{\Lambda(\eta)}^{5a(A)/2m}. \tag{4.12}$$

### 4.3.1 Bound on $W_1$

We look into the following sum

$$R_1 = \sum_{k^\beta < X^*} C_{\Sigma(\eta)} D_{\Lambda(\eta)}^{2a(A)/m}.$$

To be more precise, recall that $k_{ij}$ is the $i, j$-th square-free number for $\eta$, then

$$C_{\Sigma(\eta)} = O(\prod_j k_{1j}^a \prod_j k_{2j}^b), \tag{4.13}$$

where $a$ and $b$ are such numbers that $C_p = p^a$ and $C_{p^2} = p^b$ in Theorem 4.2.1. We know from Theorem 4.2.1 that we can take $a = b = 4/5$. We will keep $a$ and $b$ to see how much we need from them. For $\Lambda$,

$$D_{\Lambda(\eta)}^{2a(A)/m} = O\left(\prod_j \prod_i \prod_{p|k_{ij}} C' p^{-2\operatorname{ind}(b_j)/m}\right) \leq O_\epsilon\left(\prod_j (\prod_i k_{ij})^{-2\operatorname{ind}(b_j)/m+\epsilon}\right), \tag{4.14}$$

where $C' = C^{-2a(A)/m}$ is a new absolute constant depending only on $A$. So the sum $R_1$ could be bounded by a sum of multi-variable polynomial over a bounded region,

$$R_1 \leq O(\sum_{k^\beta \leq X^*} k^\gamma). \tag{4.15}$$

Here $\beta$ and $\gamma$ could be determined by (4.13), (4.14) and (4.9). The summation is considered in the following elementary calculus result. It can be proved by direct computation.

**Lemma 4.2.** *Given a vector of component $\beta$ and $\gamma$ such that $\beta_i > 0$ for all $1 \leq i \leq n$, if there exists i such that $\gamma_i \geq -1$, then the following summation is bounded*

$$\sum_{k^\beta \leq X} k^\gamma \leq O_\epsilon(X^{a(\beta,\gamma)+\epsilon}), \tag{4.16}$$

*where $a(\beta, \gamma) = \max_{1\leq i\leq n}\{\frac{\gamma_i+1}{\beta_i}\}$. If $\gamma_i < -1$ for all $i$, then the sum is bounded by $O(1)$.*

In our case, $\beta$ and $\gamma$ are indexed by $i$ and $j$. The quotient is computed to be

$$\frac{\gamma_{ij}+1}{\beta_{ij}} = \frac{a - 2\operatorname{ind}(b_j)/m + 1}{m\operatorname{ind}(a_i) + 3\operatorname{ind}(b_j)},$$

for $i = 1$, and similarly for $i = 2$

$$\frac{\gamma_{ij} + 1}{\beta_{ij}} = \frac{b - 2\operatorname{ind}(b_j)/m + 1}{m\operatorname{ind}(a_i) + 3\operatorname{ind}(b_j)},$$

after plug in (4.13), (4.14) and (4.9). Observe that if the numerator is positive, then this quantity is largest when $\operatorname{ind}(b_j) = \operatorname{ind}(A)$, i.e.,

$$a(\beta, \gamma)m = \max\left\{\frac{a - 2\operatorname{ind}(A)/m + 1}{1 + 3\operatorname{ind}(A)/m}, \frac{b - 2\operatorname{ind}(A)/m + 1}{2 + 3\operatorname{ind}(A)/m}\right\}.$$

Since we have in Theorem 4.2.1 that $C_p = p^a = C_{q^2}$ for $a = b = 4/5$, in our situations, the quantity is also largest when $\operatorname{ind}(a_i) = \operatorname{ind}(A)$, i.e.,

$$a(\beta, \gamma)m = \frac{a - 2\operatorname{ind}(A)/m + 1}{1 + 3\operatorname{ind}(A)/m}.$$

It is possible that the above expression is negative for some $A$. In that case, the summation $R_1$ is $O(1)$, so we define $a(\beta, \gamma) = 0$ for such $A$.

Plugging in $R_1$, we get $W_1$ for $B(\rho^0, X)$ that

$$W_1 \leq O(C_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{2a(A)/m}) \cdot X^{2/3m}(X^*)^{a(\beta,\gamma)+\epsilon}. \tag{4.17}$$

### 4.3.2 Bound on $W_2$ and $W_3$

In this subsection, we look into $W_2$ and $W_3$ in a similar way, and we will show that they are small. Therefore only the error from small $\eta$ makes main contribution to the error of $B(\rho^0, X)$.

Denote

$$R_2 = \sum_{k^\beta > X^*} A_{\Sigma(\eta)} D_{\Lambda(\eta)}^{3a(A)/m}.$$

We will need a similar lemma to deal with $R_2$.

**Lemma 4.3.** *Given a vector of component $\beta$ and $\gamma$ such that $\beta_i > 0$ for all $1 \leq i \leq n$, if $\gamma_i < -1$ for all $i$, then the following summation is bounded*

$$\sum_{k^\beta \geq X} k^\gamma \leq O_\epsilon(X^{a(\beta,\gamma)+\epsilon}), \tag{4.18}$$

*where $a(\beta, \gamma) = \max_{1 \leq i \leq n}\{\frac{\gamma_i + 1}{\beta_i}\}$.*

The exponent $\beta'_{ij}$ is the same as $\beta_{ij}$ in $R_1$, but the exponent $\gamma'_{ij}$ is different. By description of $A_\Sigma$ and $D_\Lambda$, the quotient is

$$\frac{\gamma'_{ij}+1}{\beta'_{ij}} = \frac{-\operatorname{ind}(a_i) - 3\operatorname{ind}(b_j)/m + 1}{m\operatorname{ind}(a_i) + 3\operatorname{ind}(b_j)} = \frac{1}{m}(-1 + \frac{1}{\operatorname{ind}(a_i) + 3\operatorname{ind}(b_j)/m}) \leq \frac{-3\operatorname{ind}(A)/m}{m + 3\operatorname{ind}(A)},$$

where in the last inequality we take $\operatorname{ind}(a_i) = \operatorname{ind}(S_3)$ and $\operatorname{ind}(b_j) = \operatorname{ind}(A)$. Therefore

$$a(\beta', \gamma')m = \frac{-3\operatorname{ind}(A)/m}{1 + 3\operatorname{ind}(A)/m}.$$

By Lemma 4.3 and description of $A_{\Sigma(\rho)}$, $D_{\Lambda(\rho)}$, we have

$$W_2 = O(A_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{3a(A)/m}) \cdot X^{1/m} \cdot (X^*)^{a(\beta',\gamma')+\epsilon} \leq O(X^*)^{1/m + a(\beta',\gamma')+\epsilon}. \tag{4.19}$$

Similarly for $W_3$, the exponent

$$a(\beta'', \gamma'')m = \frac{-5\operatorname{ind}(A)/2m}{1 + 3\operatorname{ind}(A)/m} = \frac{5}{6} \cdot a(\beta', \gamma')m,$$

and

$$W_3 = O(B_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{5a(A)/2m}) \cdot X^{5/6m}(X^*)^{a(\beta'',\gamma'')+\epsilon} \leq O(\operatorname{Disc}(\Sigma(\rho))^{-1/6}(X^*)^{5/6m + a(\beta'',\gamma'')+\epsilon}). \tag{4.20}$$

Therefore the bound on $W_3$ is smaller than that of $W_2$, so it suffices to compare that of $W_2$ with $W_1$. Notice that

$$-\frac{1}{3} + a(\beta, \gamma)m - a(\beta', \gamma')m \geq \frac{a + 2/3}{1 + 3\operatorname{ind}(A)/m} > 0,$$

we have

$$W_2 \leq O(C_{\Sigma(\rho)} \operatorname{Disc}(\Sigma(\rho))^{2/3})(X^*)^{2/3m + a(\beta,\gamma)+\epsilon} = O(C_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{2a(A)/m}) \cdot X^{2/3m}(X^*)^{a(\beta,\gamma)+\epsilon},$$

so

$$W_1 + W_2 + W_3 \leq O(C_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{2a(A)/m}) \cdot X^{2/3m}(X^*)^{a(\beta,\gamma)+\epsilon}.$$

### 4.3.3 Error for $B(\rho^0, XL_\rho)$

Finally, we replace $X$ with $XL_\rho$ in $B(\rho^0, XL_\rho)$ and denote the error by $E_\rho$. Plugging in

$$X^* = \frac{XL_\rho}{\operatorname{Disc}(\Sigma(\rho))^m \operatorname{Disc}(\Lambda(\rho))^3},$$

we get

$$\begin{aligned} E_\rho &\leq O(C_{\Sigma(\rho)} \cdot D_{\Lambda(\rho)}^{2a(A)/m}) \cdot (XL_\rho)^{2/3m} O_\epsilon(\frac{XL_\rho}{\operatorname{Disc}(\Sigma(\rho))^m \operatorname{Disc}(\Lambda(\rho))^3})^{a(\beta,\gamma)+\epsilon} \\ &\leq O_\epsilon(X^{2/3m+a(\beta,\gamma)+\epsilon}) O(C_{\Sigma(\rho)} D_{\Lambda(\rho)}^{2a(A)/m} L_\rho^{2/3m+a(\beta,\gamma)+\epsilon})(\operatorname{Disc}(\Sigma(\rho))^m \operatorname{Disc}(\Lambda(\rho))^3)^{-a(\beta,\gamma)-\epsilon} \\ &\leq O_\epsilon(X^{2/3m+a(\beta,\gamma)+\epsilon}) \prod_{i,j} k_{ij}^{e(i,j)}, \end{aligned} \tag{4.21}$$

where

$$e(1,j) = a + 2/3 - \operatorname{ind}((12)(3), b_j) \cdot (2/3m + a(\beta,\gamma)) + \epsilon,$$

and

$$e(2,j) = b + 4/3 - \operatorname{ind}((123), b_j) \cdot (2/3m + a(\beta,\gamma)) + \epsilon.$$

Here $\operatorname{ind}((12)(3), bj)$ means the index of the group element $((12)(3), b_j) \in S_3 \times A$.

## 4.4 Estimates of $B(\rho^0, XL_\rho)$ for Large $\rho$

In this section, we will use uniformity estimates to determine the tail estimate for $B(\rho^0, XL_\rho)$ for $\rho$. The expression will hold uniformly for all $\rho$, but it will be especially helpful when $\rho$ involves relatively larger prime numbers.

Recall in Theorem 2.2.8 and 2.1.2, we get uniformity estimates for $S_3$ cubic extension and $A$-extension with restriction on ramification, which states that

$$\sharp\{K \mid \operatorname{Gal}(K) = S_3, K \in \Sigma(\rho), \operatorname{Disc}(K) \leq X\} = O(\frac{X}{\prod_j k_{1j}^{1/6-\epsilon} \prod_j k_{2j}^{2-\epsilon}}),$$

and

$$\sharp\{L \mid \operatorname{Gal}(L) = A, L \in \Lambda(\rho), \operatorname{Disc}(K) \leq X\} = O(\frac{X^{1/a(A)+\epsilon}}{\prod_j (k_{1j}k_{2j})^{\operatorname{ind}(b_j)/a(A)-\epsilon}}).$$

Since $\rho^0$ requires more restriction on places outside $k_{ij}$, we have that

$$B(\rho^0, XL_\rho) \leq B(\rho^1, XL_\rho).$$

Applying Theorem 3.5 on $\mathrm{Disc}_{res}(K) = \frac{\mathrm{Disc}(K)}{\mathrm{Disc}(\Sigma(\rho))}$ and $\mathrm{Disc}_{res}(L) = \frac{\mathrm{Disc}(L)}{\mathrm{Disc}(\Lambda(\rho))}$, we get

$$\begin{aligned}
&B(\rho^1, XL_\rho)\\
=&\sharp\{(K,L) \mid K \in \Sigma(\rho), L \in \Lambda(\rho), \mathrm{Disc}_{res}(K)^m \,\mathrm{Disc}_{res}(L)^3 \leq \frac{XL_\rho}{\mathrm{Disc}(\Sigma(\rho))^m \,\mathrm{Disc}(\Lambda(\rho))^3}\}\\
=&\sharp\{(K,L) \mid K \in \Sigma(\rho), L \in \Lambda(\rho), \mathrm{Disc}_{res}(K)^m \,\mathrm{Disc}_{res}(L)^3 \leq \frac{X}{\mathrm{Disc}(\Sigma(\rho), \Lambda(\rho))}\}\\
\leq& O(\prod_j k_{1j}^{5/6+\epsilon} \prod_j k_{2j}^{\epsilon})(\frac{X}{\mathrm{Disc}(\Sigma(\rho), \Lambda(\rho))})^{1/m}\\
=&O(X^{1/m}) \prod_{i,j} k_{ij}^{d(i,j)},
\end{aligned} \tag{4.22}$$

where

$$d(1,j) = \frac{5}{6} + \epsilon - \mathrm{ind}((12)(3), b_j)/m,$$

and

$$d(2,j) = \epsilon - \mathrm{ind}((123), b_j)/m.$$

These tail estimates will all be error terms, and we will denote it by

$$D_\rho = B(\rho^0, XL_\rho) \leq B(\rho^1, XL_\rho) \leq O(X^{1/m}) \prod_{i,j} k_{ij}^{d(i,j)}.$$

## 4.5 Optimization

In this section, we will combine the error estimates in previous sections, and balance between errors in the small range and the large range to optimize the error overall.

Recall that in the small range we get the error

$$E_\rho = O(X^{2/3m+a(\beta,\gamma)+\epsilon}) \prod_{i,j} k_{ij}^{e(i,j)},$$

and in the large range we get the error

$$D_\rho = O(X^{1/m}) \prod_{i,j} k_{ij}^{d(i,j)}.$$

So to take advantage of both estimate, we will use the sieve argument when

$$E_\rho \leq D_\rho,$$

which is equivalent to

$$\prod_{i,j} k_{ij}^{\delta(i,j)} \leq X^{1/3m-a(\beta,\gamma)-\epsilon} = Q,$$

where $\delta(i,j) = e(i,j) - d(i,j)$.

So the error overall will be

$$\begin{aligned} E &= \sum_{\substack{\rho \\ \prod_{i,j} k_{i,j}^{\delta(i,j)} \leq Q}} E_\rho + \sum_{\substack{\rho \\ \prod_{i,j} k_{i,j}^{\delta(i,j)} \geq Q}} D_\rho \\ &= E_S + E_L \end{aligned} \tag{4.23}$$

1. **Estimates for $E_S$**

   The sum $E_S$ for the small range is

   $$\begin{aligned} E_S &= O_\epsilon(X^{2/3m+a(\beta,\gamma)+\epsilon}) \sum_{\substack{\rho \\ \prod_{i,j} k_{ij}^{\delta(i,j)} \leq Q}} \prod_{i,j} k_{ij}^{e(i,j)} \\ &= O_\epsilon(X^{2/3m+a(\beta,\gamma)+\epsilon}) \cdot Q^{\max\{\frac{e(i,j)+1}{\delta(i,j)}\}}. \end{aligned} \tag{4.24}$$

   For the second equality, we apply Lemma 4.2 since there exists $e(i,j) > -1$ and for all $i$ and $j$, $\delta(i,j) > 0$.

2. **Estimates for $E_L$**

   The sum $E_L$ for the large range is

   $$\begin{aligned} E_L &= O(X^{1/m}) \sum_{\substack{\rho \\ \prod_{i,j} k_{ij}^{\delta(i,j)} \geq Q}} \prod_{i,j} k_{ij}^{d(i,j)} \\ &= O(X^{1/m}) \cdot Q^{\max\{\frac{d(i,j)+1}{\delta(i,j)}\}}. \end{aligned} \tag{4.25}$$

   For the second equality, we apply Lemma 4.3 since $d(i,j) < -1$ and $\delta(i,j) > 0$ for all $i$ and $j$.

To sum up, for the small range, we use estimates with precise first term, secondary term and an error term $E_\rho$; for large range, we use estimates which is purely error term $D_\rho$. Since the first term and secondary term are both small comparing to $D_\rho$,

$$O(B_\Sigma \cdot g_\Lambda(\frac{5}{2m}))(XL_\rho)^{5/6m} \leq O(A_\Sigma \cdot g_\Lambda(\frac{3}{m}))(XL_\rho)^{1/m} \leq O(X^{1/m}) \prod_{i,j} k_{ij}^{d(i,j)},$$

by comparing the exponent for each $k_{ij}$, we could pretend that we use estimates with a precise main term and a secondary term with the error $D_\rho$ without harm. Finally we get that the error is

$$E = O_\epsilon(X^{2/3m+a(\beta,\gamma)+\epsilon}) \cdot Q^{\max\{\frac{e(i,j)+1}{\delta(i,j)}\}} + O(X^{1/m}) \cdot Q^{\max\{\frac{d(i,j)+1}{\delta(i,j)}\}}, \tag{4.26}$$

where $Q = X^{1/3m-a(\beta,\gamma)-\epsilon}$, and $a(\beta,\gamma)$, $e(i,j)$, $d(i,j)$ and $\delta(i,j)$ are constants depending on $A$.

Therefore finally it reduces to the question if we could show for $A$ that

$$\frac{2}{3m} + a(\beta,\gamma) + \epsilon + (\frac{1}{3m} - a(\beta,\gamma) - \epsilon) \cdot \max_{i,j}\{\frac{e(i,j)+1}{\delta(i,j)}\} < \frac{1}{m}, \tag{4.27}$$

and

$$\frac{1}{m} + (\frac{1}{3m} - a(\beta,\gamma) - \epsilon) \cdot \max_{i,j}\{\frac{d(i,j)+1}{\delta(i,j)}\} < \frac{1}{m}. \tag{4.28}$$

If we could show the above inequalities for $A$, then we succeed in proving a power saving error for $N(S_3 \times A, X)$. Moreover, if we could show the two inequalities with the right hand side replaced by $\frac{5}{6m}$,

$$\frac{2}{3m} + a(\beta,\gamma) + \epsilon + (\frac{1}{3m} - a(\beta,\gamma) - \epsilon) \cdot \max_{i,j}\{\frac{e(i,j)+1}{\delta(i,j)}\} < \frac{5}{6m}, \tag{4.29}$$

and

$$\frac{1}{m} + (\frac{1}{3m} - a(\beta,\gamma) - \epsilon) \cdot \max_{i,j}\{\frac{d(i,j)+1}{\delta(i,j)}\} < \frac{5}{6m}, \tag{4.30}$$

then we will succeed in saving the secondary term in the order of $X^{5/6m}$. Since the inequality are all strict, we could totally ignore those $\epsilon$ since they could be arbitrarily small.

## 4.6 Proof of the Main Theorem

In this section, we will prove the main theorem by verifying (4.27),(4.28), (4.29) and (4.29). To do that, we will compute explicitly the quantities of these parameters of an abelian group $A$: $a(\beta,\gamma)$, $e(i,j)$, $d(i,j)$ and $\delta(i,j)$. In the following discussion, let us denote an important quantity associated to $A$ by

$$\Delta = \frac{\operatorname{ind}(A)}{m} = \frac{p-1}{p},$$

in which $p$ is the smallest prime divisor of $m$.

Firstly, recall that if the following quantity is positive then

$$a(\beta,\gamma) = \frac{a - 2\operatorname{ind}(A)/m + 1}{m + 3\operatorname{ind}(A)},$$

otherwise,

$$a(\beta,\gamma) = 0.$$

By solving for $a(\beta,\gamma) = 0$ and plugging in $a = 4/5$, we get that if $p > 7$, then $a(\beta,\gamma) = 0$. On the other hand, if $p = 3, 5, 7$, then

$$a(\beta,\gamma)m = \frac{a - 2\Delta + 1}{1 + 3\Delta}.$$

Secondly, recall that we have

$$e(1,j) = a + 2/3 - \operatorname{ind}((12)(3), b_j) \cdot (2/3m + a(\beta,\gamma)) + \epsilon,$$

$$e(2,j) = b + 4/3 - \operatorname{ind}((123), b_j) \cdot (2/3m + a(\beta,\gamma)) + \epsilon,$$

$$d(1,j) = \frac{5}{6} + \epsilon - \operatorname{ind}((12)(3), b_j)/m,$$

$$d(2,j) = \epsilon - \operatorname{ind}((123), b_j)/m,$$

where $a = b = 4/5$ as in Theorem 4.2.1.

***Proof of Theorem 1.2.3 and Theorem 1.2.4.*** It suffices to prove the inequality (4.29) and (4.30) for $p > 5$ and (4.27) and (4.28) for $p = 3, 5$. The key quantity is the maximum of

$\frac{e(i,j)+1}{\delta(i,j)}$ and $\frac{d(i,j)+1}{\delta(i,j)}$ over $j$ for $i = 1, 2$. We will call them $U_i$ and $V_i$ for $i = 1, 2$ correspondingly. Notice that for each fixed $i, j$

$$\frac{2}{3m} + a(\beta,\gamma) + (\frac{1}{3m} - a(\beta,\gamma)) \cdot \frac{e(i,j)+1}{\delta(i,j)} = \frac{1}{m} + (\frac{1}{3m} - a(\beta,\gamma)) \cdot \frac{d(i,j)+1}{\delta(i,j)},$$

and in all of our cases, we can check that the maximum value of $U_i$ and $V_i$ are obtained when $\operatorname{ind}(b_j) = \operatorname{ind}(A)$. Therefore to check (4.29) is equivalent to check (4.30) and similarly for (4.27) and (4.28). It suffices to check for $U_i$.

When $p > 7$, by Theorem 3.1.3 we have

$$\frac{e(1,j)+1}{\delta(1,j)} = \frac{1 + a - 4\operatorname{ind}(b_j)/3m + \epsilon}{a + 1/6 + 2\operatorname{ind}(b_j)/3m} \leq \frac{1 + a - 4\Delta/3 + \epsilon}{1/6 + a + 2\Delta/3} = U_1,$$

where the maximum is taken when $\operatorname{ind}(b_j) = \operatorname{ind}(A)$. Similarly,

$$\frac{e(2,j)+1}{\delta(2,j)} \leq \frac{1 + b - 2\Delta/3 + \epsilon}{2 + b + \Delta/3} = U_2.$$

So the inequality (4.29) becomes purely dependent on $\Delta$:

$$\frac{2}{3} + \epsilon + (\frac{1}{3} - \epsilon) \cdot U_i < \frac{5}{6}, \tag{4.31}$$

for $i = 1, 2$. We can check this holds for $p > 7$.

When $p = 5, 7$, we have

$$U_1 = \frac{1 + a - 4\Delta/3 - (a - 2\Delta + 1)(1 + 2\Delta)(1 + 3\Delta)^{-1} + \epsilon}{1/6 + a + 2\Delta/3 - (a - 2\Delta + 1)(1 + 2\Delta)(1 + 3\Delta)^{-1}},$$

$$U_2 = \frac{1 + b - 2\Delta/3 - (a - 2\Delta + 1)(2 + \Delta)(1 + 3\Delta)^{-1} + \epsilon}{2 + b + \Delta/3 - (a - 2\Delta + 1)(2 + \Delta)(1 + 3\Delta)^{-1}}.$$

It suffices to check the following holds

$$\frac{2}{3} + \frac{a - 2\Delta + 1}{1 + 3\Delta} + \epsilon + (\frac{1}{3} - \frac{a - 2\Delta + 1}{1 + 3\Delta} - \epsilon) \cdot U_i < \frac{5}{6}, \tag{4.32}$$

for $p = 7$, and when $p = 5$ the inequality holds when the right hand side is 1, which finishes the proof of Theorem1.2.3 and Theorem 1.2.4 for $A$ with $p > 3$.

When $p = 3$, we just need to compute more carefully. Now

$$a(\beta,\gamma)m = \frac{a - 1/3}{3} = \frac{7}{45},$$

and the inequality for $U_1$ remains the same since $(12)(3)$ has order 2, which is relatively prime to order of $g$ for any $g \in A$. So it suffices to check (4.32) holds for $U_1$ when $\Delta = 2/3$ and the right hand side is 1. For $U_2$, plugging $a(\beta, \gamma)m = 7/45$ into (4.32) and rearranging the terms, it suffices to check that

$$\begin{aligned} \frac{e(2,j)+1}{\delta(2,j)} &< 1, \\ \frac{d(2,j)+1}{\delta(2,j)} &< 0, \end{aligned} \tag{4.33}$$

which is equivalent to

$$d(2,j) = \epsilon - \mathrm{ind}((123), b_j)/m < -1.$$

It follows from the Lemma 3.1. □

## 4.7 Constants for the Main Term and the Secondary Term

In this section, we are going to compute the precise constants for the main term and the secondary term when $A$ is a cyclic group with prime order $m = l$ for $l > 5$.

We will first consider all continuous homomorphisms $G_{\mathbb{Q}} \to C_l$ instead of $C_l$ extensions of $\mathbb{Q}$ for simplicity of computation of main term. The two quantity differ by a trivial map up to an action of $\mathrm{Aut}(C_l)$. The generating series for such maps is

$$g(s) = \left(1 + (l+1)l^{-2(l-1)s}\right) \prod_{p \neq l, p \equiv 1 \bmod l} \left(1 + (l-1)p^{-(l-1)s}\right).$$

Recall that the precise terms for $B(\varrho^1, XL_\rho)$ is

$$AA_{\Sigma(\varrho)} \cdot g_{\Lambda(\varrho)}(\frac{3}{m})(XL_\rho)^{1/m} + BB_{\Sigma(\varrho)} \cdot g_{\Lambda(\varrho)}(\frac{5}{2m})(XL_\rho)^{5/6m},$$

so the main term for $B(\rho^0, XL_\rho)$ is

$$A(XL_\rho)^{1/m} A_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)}(\frac{3}{m}) \sum_\eta \mu(\eta) A_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)}(\frac{3}{m}),$$

and the secondary term for $B(\rho^0, XL_\rho)$ is

$$B(XL_\rho)^{5/6m} B_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)}(\frac{5}{2m}) \sum_\eta \mu(\eta) B_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)}(\frac{5}{2m}),$$

where both sums are over all $\eta$ that is relatively prime to $\rho$. Finally we sum over all $\rho$ and get the main term for the whole counting

$$AX^{1/m}\sum_{\rho} L_{\rho}^{1/m} A_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)}(\frac{3}{m}) \sum_{\eta} \mu(\eta) A_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)}(\frac{3}{m}),$$

with a secondary term

$$BX^{5/6m}\sum_{\rho} L_{\rho}^{5/6m} B_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)}(\frac{5}{2m}) \sum_{\eta} \mu(\eta) B_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)}(\frac{5}{2m}),$$

where we sum over all possible $\rho$.

In this specific case, an extension $L$ with the Galois group $C_l$ could only be wildly ramified at $l$. At $l > 3$, an $S_3$ cubic extension $K$ could be tamely ramified so $l$ is the only place we need be careful about wildly ramification. On the other hand $C_l$ is cyclic with prime order, so there is only one type of tamely ramification, and $S_3$ has two types of tamely ramification, so the tamely ramification part in a local condition $\rho$, could be parametrized by a pair of relatively prime square-free integers, say $q$ and $r$. Similarly for $\eta$, say $k$ and $l$, with $klpq$ square-free. So plugging in all the constant we get the coefficient for the main term:

$$\begin{aligned}
\mathcal{C}_1 =& A\sum_{\rho} L_{\rho}^{1/m} A_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)}(\frac{3}{m}) \sum_{\eta} \mu(\eta) A_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)}(\frac{3}{m}) \\
=& \frac{1}{3\zeta(3)} \cdot c_l \cdot \sum_{\substack{q,r,k,l \\ \forall p|qrkl, p\equiv 1 \bmod l}} \mu(k)\mu(l) \frac{1}{q^{\Delta} r^{2\Delta}} \cdot \prod_{p|qk} \frac{C_p^{-1}}{p} \cdot \prod_{p|rl} \frac{C_p^{-1}}{p^2} \cdot g_{(}\frac{3}{l}) \prod_{p|qrkl} \frac{(l-1)p^{-3\Delta}}{1+(l-1)p^{-3\Delta}} \\
=& \frac{g(3/l)}{3\zeta(3)} \cdot c_l \cdot \prod_{p\equiv 1 \bmod l} \left\{ 1 + p^{-\Delta} \cdot \frac{C_p^{-1}}{p} \cdot \frac{(l-1)p^{-3\Delta}}{1+(l-1)p^{-3\Delta}} + p^{-2\Delta} \cdot \frac{C_p^{-1}}{p^2} \cdot \frac{(l-1)p^{-3\Delta}}{1+(l-1)p^{-3\Delta}} \right. \\
& \left. - \frac{C_p^{-1}}{p} \cdot \frac{(l-1)p^{-3\Delta}}{1+(l-1)p^{-3\Delta}} - \frac{C_p^{-1}}{p^2} \cdot \frac{(l-1)p^{-3\Delta}}{1+(l-1)p^{-3\Delta}} \right\},
\end{aligned} \tag{4.34}$$

where $C_p = 1 + p^{-1} + p^{-2}$ is the normalizing factor for the local density at $s = 1$ for $S_3$ extensions, and

$$g(3/l) = (1 + (l+1)l^{-6\Delta}) \prod_{p\neq l, p\equiv 1 \bmod l} (1 + (l-1)p^{-3\Delta}),$$

and the local factor at $l$

$$c_l = \sum_{(\sigma_l, \lambda_l)} \frac{\mathrm{Disc}(\sigma_l)\,\mathrm{Disc}(\lambda_l)^{3/l}}{\mathrm{Disc}(\sigma_l, \lambda_l)^{1/l}} \cdot \frac{g_{\lambda_l}(3/l)}{g(3/l)} \cdot A_{\sigma_l}.$$

Similarly, we can compute the constant for the secondary term

$$\begin{aligned}
\mathcal{C}_2 =& B \sum_{\rho} L_\rho^{5/6m} B_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)}(\frac{5}{2m}) \sum_{\eta} \mu(\eta) B_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)}(\frac{5}{2m}) \\
=& B \cdot g(\frac{5}{2l}) \cdot d_l \cdot \sum_{\substack{q,r,k,l \\ \forall p | qrkl \\ p \equiv 1 \bmod l}} \left\{ \mu(k)\mu(l)(\frac{1}{q^{\Delta} r^{2\Delta}})^{5/6} \cdot \prod_{p|qk} \frac{(1+p^{-1/3})^2}{K_p \cdot p} \right. \\
& \left. \cdot \prod_{p|rl} \frac{(1+p^{-1/3})}{K_p \cdot p^2} \cdot \prod_{p|qrkl} \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right\} \qquad (4.35) \\
=& B \cdot g(5/2l) \cdot d_l \cdot \prod_{p \equiv 1 \bmod l} \left\{ 1 + p^{-5\Delta/6} \cdot \frac{(1+p^{-1/3})^2}{K_p \cdot p} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right. \\
& + p^{-5\Delta/3} \cdot \frac{1+p^{-1/3}}{K_p \cdot p^2} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \\
& \left. - \frac{(1+p^{-1/3})^2}{K_p \cdot p} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} - \frac{1+p^{-1/3}}{K_p \cdot p^2} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right\},
\end{aligned}$$

where

$$B = (1+\sqrt{3}) \frac{4\zeta(1/3)}{5\Gamma(2/3)^3 \zeta(5/3)},$$

is the constant for the secondary term for $S_3$ extensions, and

$$K_p = \frac{(1-p^{-5/3})(1+p^{-1})}{1-p^{-1/3}},$$

is the normalizing factor for the local density at $s = 5/6$ for $S_3$ extensions, and

$$d_l = \sum_{(\sigma_l, \lambda_l)} \frac{\mathrm{Disc}(\sigma_l)^{5/6}\,\mathrm{Disc}(\lambda_l)^{5/2l}}{\mathrm{Disc}(\sigma_l, \lambda_l)^{5/6l}} \cdot \frac{g_{\lambda_l}(5/2l)}{g(5/2l)} \cdot B_{\sigma_l}.$$

Notice that we are counting continous homomorphisms from $G_{\mathbb{Q}}$ to $S_3 \times C_l$ which are surjective onto the $S_3$ component up to an action of $\mathrm{Aut}(C_l)$ on the $C_l$ component, therefore the true value for the constant of the main term is

$$C_1 = \frac{1}{l-1} \cdot (\mathcal{C}_1 - A),$$

and the value for the constant of the secondary term is

$$C_2 = \frac{1}{l-1} \cdot (\mathcal{C}_2 - B).$$

## 4.8 The Amount of Power Saving

In this subsection, we are going to compute the amount of power saving error away from the secondary term in the order of $X^{5/6m}$ when $p > 5$ and the amount of power saving from the main term for $p = 3, 5$.

Recall that in section 4.5, we have specified the exponent of $X$ in the error term to be the maximum value among (4.30) and (4.29), therefore the amount of power saving is

$$\frac{5}{6m} - \frac{2}{3m} - a(\beta,\gamma) - \epsilon - (\frac{1}{3m} - a(\beta,\gamma) - \epsilon) \cdot \max_{i,j}\{\frac{e(i,j)+1}{\delta(i,j)}\},$$

and

$$\frac{5}{6m} - \frac{1}{m} - (\frac{1}{3m} - a(\beta,\gamma) - \epsilon) \cdot \max_{i,j}\{\frac{d(i,j)+1}{\delta(i,j)}\}.$$

Recall $a = b = 4/5$ is the proved dependency of error for $S_3$ extensions. For $p > 7$, we can compute the amount of power saving is

$$\delta = \frac{1}{6m} \cdot \min\{\frac{10\Delta/3 - a - 11/6}{1/6 + a + 2\Delta/3}, \frac{5\Delta/3 - b}{2 + b + \Delta/3}\} - \epsilon = \frac{1}{6m} \cdot \frac{5\Delta/3 - b}{2 + b + \Delta/3} - \epsilon,$$

where for the second equality we use that $a = 4/5$ and $b = 4/5$. For $p = 7$, the amount of power saving is $\delta = 23/(1254m) - \epsilon$. For $p = 5$, the amount of power saving is $\delta = 322/(2061m) - \epsilon$. For $p = 3$, the amount of power saving is $\delta = 24/(283m) - \epsilon$.

# Chapter 5

# Conclusion

We have discussed approaches to understand both the main term and the error terms for counting number fields with bounded discriminant when the Galois group is direct product of groups where counting results are previously known. Due to the lack of results in Malle's conjecture and uniform estimates for ramified extensions, most of the work in this thesis are devoted to prove theorems for number fields with special Galois groups. However, in most situations, the skeleton of both methods would work if counting results for smaller fields are known in advance.

To sum up, given two permutation groups $G_1$ and $G_2$, the general picture for proving Malle's conjecture is determined by the two factors:

1. The gap between $\frac{\text{deg}}{\text{ind}}(G_i)$ for $i = 1, 2$.

   Usually speaking, the larger the gap is, the easier the question is. The winner of this game will determine the behavior of the distribution function more. The following picture demonstrate the quantity $\frac{\text{deg}}{\text{ind}}$ for some common seen permutation groups we know. This quantity basically describes how dense and dominant certain extensions are. (In the following picture, the Group $S_n$ denotes $S_n$ in its natural permutation representation.) The approach in this thesis will always work if the two groups are with different $\frac{\text{deg}}{\text{ind}}$ on this axis. However, when the two groups have the same $\frac{\text{deg}}{\text{ind}}$, this method will work in some cases and won't work in some cases. For example, this method will also work when $G_1 = G_2 = S_3(3)$, if only we could improve the uniform estimates for $S_3(3)$ extensions a lot.

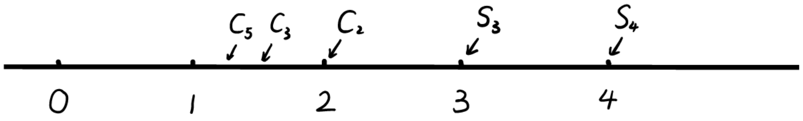

Figure 5.1 Degree/Index

2. Uniform estimates for ramified $G_i$ extensions.

   It is worth noting that improvement on uniform upper bounds on both $G_i$ ramified extensions will help. However, we seem to have very limited tools to attack this problems. The only tool we have for this question is class field theory previously. In the situation where class field theory could be applied, usually we get pretty good estimates. In this thesis we applied geometric sieve [Bha14] to give new uniform estimates for $S_n$ extensions for $n = 3, 4, 5$. This approach is good in the sense that it gives some non-trivial upper bounds for all types of ramification, but the improvement from the trivial bound is very limited.

Therefore it is a crucial question:

**Question 5.1.** *How to prove more results on uniform upper bounds of ramified number fields?*

In terms of the secondary term, we could similarly place the secondary term on the axis by replacing $1/\operatorname{ind}(G_i)$ with the exponent of the secondary term. Whether a secondary term could be proved depends heavily on how good the previous estimates are. In terms of what we expect to be true, we give the following conjecture based on this reasoning:

**Conjecture 5.2.** *Given $G_i \subset S_{n_i}$ for $i = 1, 2$, if the asymptotic distribution of $G_1$ extensions has a secondary term in the order of $X^{c_1}$ such that $\operatorname{ind}(G_2) > \frac{n_2}{n_1 c_1}$, then the asymptotic distribution of $G_1 \times G_2$ extensions has a secondary term in the order of $X^{c_1/n_2}$.*

However, in general we have no idea on:

**Question 5.3.** *When should we expect there to be a secondary term? What order do we expect for the secondary term for a general Galois group?*

One of the examples we considered , $S_3 \times C_3$, is of special interest, which is the first counter-example Klüners discovered for Malle's conjecture. Counting non-Galois number fields could be considered as counting Galois number fields by discriminant of certain subfields. We show that although this group is a counter-example when ordered by the discriminant of the degree 6 fields, the counting matches Malle's prediction when ordered by the discriminant of the degree 9 fields. A natural question thus will be:

**Question 5.4.** *What kind of subfields provide the discriminant as an invariant by which the asymptotic estimate is as predicted by Malle?*

# LIST OF REFERENCES

[BBP10] K. Belabas, M. Bhargava, and C. Pomerance. Error terms for the Davenport-Heilbronn theorems. *Duke Math. J.*, 153(1):173–210, 2010.

[BF10] K. Belabas and E. Fouvry. Discriminants cubiques et progressions arithmetiqués,. *Int. J. Number Theory*, 6(7):1491–1529, 2010.

[Bha05] M. Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math.*, 162(2):1031–1063, September 2005.

[Bha10] M. Bhargava. The density of discriminants of quintic rings and fields. *Ann. of Math. (2)*, 172(3):1559–1591, 2010.

[Bha14] M. Bhargava. The geometric sieve and the density of squarefree values of polynomial discriminants and other invariant polynomials. *http://arxiv.org/abs/1402.0031*, 2014.

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

[BSW17] M. Bhargava, A. Shankar, and X. Wang. Geometry-of-numbers methods over global fields I: Prehomogeneous vector spaces. *preprint*, 2017.

[BTT16] M. Bhargava, T. Taniguchi, and F. Thorne. Secondary terms in counting functions for cubic fields, II. *preprint*, 2016.

[BW08] M. Bhargava and M. M. Wood. The density of discriminants of $S_3$-sextic number fields. *Proc. Amer. Math. Soc.*, 136(5):1581–1587, 2008.

[CyDO02] H. Cohen, F. Diaz y Diaz, and M. Olivier. Enumerating quartic dihedral extensions of $\mathbb{Q}$. *Compositio Math.*, 133(1):65–93, 2002.

[CyDO06] H. Cohen, F. Diaz y Diaz, and M. Olivier. Counting discriminants of number fields. *J. Théor. Nombres Bordeaux*, 18(3):573–593, 2006.

[DH71] H. Davenport and H. Heilbronn. On the density of discriminants of cubic fields. II. *Proc. Roy. Soc. London. Ser. A*, 322(1551):405–420, 1971.

[DW88] B. Datskovsky and D. J. Wright. Density of discriminants of cubic extensions. *J. Reine Angew. Math*, (386):116–138, 1988.

[EPW] J. Ellenberg, L. B. Pierce, and M. M. Wood. On $\ell$-torsion in class groups of number fields. *arXiv: 1606.06103.*

[Hou10] B. Hough. Equidistribution of Heegner points associated to the 3-part of the class group. *preprint*, 2010.

[Klü05] J. Klüners. A counter example to Malle's conjecture on the asymptotics of discriminants. *C. R. Math. Acad. Sci. Paris*, 340(6):411–414, 2005.

[Klü12] J. Klüners. The distribution of number fields with wreath products as Galois groups. *Int. J. Number Theory*, (8):845–858, 2012.

[KM04] J. Klüners and G. Malle. Counting nilpotent Galois extensions. *J. Reine Angew. Math*, 572:1–26, 2004.

[Lan94] S. Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics.* Springer-Verlag New York, 1994.

[Mäk85] S. Mäki. On the density of abelian number fields. *Ann. Acad. Sci. Fenn. Diss. Series A I. Mathematica Dissertationes*, 54(104), 1985.

[Mal02] G. Malle. On the distribution of Galois groups. *J. Number Theory*, 92(2):315–329, 2002.

[Mal04] G. Malle. On the distribution of Galois groups, II. *Experiment. Math.*, 13(2):129–135, 2004.

[Nar83] W. Narkiewicz. *Number theory.* World Scientific Publishing Co., Singapore, 1983.

[Neu99] J. Neukirch. *Algebraic number theory*, volume 322. Springer-Verlag, 1999.

[PTBW] L. B. Pierce, C. Turnage-Butterbaugh, and M. M. Wood. An effective Chebotarev density theorem for families of number fields, with an application to $\ell$-torsion in class groups. *arXiv: 1709.09637.*

[Rob01] D. Roberts. Density of cubic field discriminants. *Math. Comp.*, 70(236):1699–1705, 2001.

[ST] A. Shankar and J. Tsimerman. Counting $S_5$-fields with a power saving error term. *http://arxiv.org/abs/1310.1998.*

[Tho11] F. Thorne. Four perspectives on secondary terms in the Davenport-Heilbronn theorems. *Integers Volume 12 B, Proceedings of the Integers Conference 2011*, 2011.

[TT13] T. Taniguchi and F. Thorne. Secondary terms in counting functions for cubic fields. *Duke Math. J.*, 162(13):2451–2508, 2013.

[TT14] T. Taniguchi and F. Thorne. An error estimate for counting $s_3$-sextic number fields. *Int. J. Number Theory*, 10(04):935–948, 2014.

[Tur08] S. Turkelli. Connected components of Hurwitz schemes and Malle's conjecture. *arXiv: 0809.0951*, September 2008.

[Wan17] J. Wang. Malle's conjecture for $S_n \times A$ for $n = 3, 4, 5$. *arXiv: 1705.00044*, 2017.

[Woo10] M. M. Wood. On the probabilities of local behaviors in abelian field extensions. *Compositio Math.*, 146(1):102–128, 2010.

[Woo16] M. M. Wood. Asymptotics for number fields and class groups. In *Directions in Number Theory*, pages 291–339. Springer International Publishing, 2016.

[Woo17] M. M. Wood. Nonabelian Cohen-Lenstra moments. *arXiv: 1702.04644*, 2017.

[Wri89] D. J. Wright. Distribution of discriminants of abelian extensions. *Proc. of London Math. Soc. (3)*, 58(1):1300–1320, 1989.

[WW96] E. T. Whittaker and G. N. Watson. *A course of modern analysis.* Cambridge university press, 1996.

[Zha13] Y. Zhao. On sieve methods for varieties over finite fields. *preprint*, 2013.