

Secondary Term of Asymptotic Distribution of $S_3 \times A$ Extensions over \mathbb{Q}

Jiuya Wang

December 2, 2017

Abstract

We combine a sieve method together with good uniformity estimates to 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.

Key words. Malle's conjecture, secondary term, uniformity estimates, power saving error.

1 Introduction

Davenport and Heilbronn [DH71] proved a celebrated theorem on the asymptotic distribution of S_3 cubic extensions over \mathbb{Q} , which gives the average 3-class number of quadratic extensions. This is the first and only proven case for Cohen-Lenstra heuristics over quadratic fields up till now. This work was 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. Given $G \subset S_n$ a permutation group, denote $N_k(G, X)$ to be the number of extensions over k with $\text{Gal}(K/k)$ isomorphic to G as a permutation group and with the absolute discriminant bounded by X . Then their results state as following:

Theorem 1.1 ([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.

What is more striking about this counting $N_{\mathbb{Q}}(S_3, X)$ is that it 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 our paper heavily depends on.

Theorem 1.2 ([BTT16], Theorem 4.3). *There exists constant A and B such that the asymptotic distribution of S_3 cubic extensions over \mathbb{Q} is*

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

Malle brought forward his conjecture [Mal02] and [Mal04] on the order of the main term in $N_k(G, X)$, however there is still little understanding towards the secondary term. 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.

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.

In a recent paper of the author [Wan17], the asymptotic distribution of $S_3 \times A$ fields are obtained in terms of a precise main term when A is any odd abelian group. The main result of our paper 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 1.3. *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 1.4. *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*

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

The amount of power saving δ in both Theorem 1.3 and 1.4 are computed in section 4.8.

To prove these results, we apply a sieve method for the range of small primes, building on the distribution of cubic fields with local conditions, and we prove a new uniformity estimate of ramified cubic fields for the range of large primes. For these cases in Theorem 1.3 and 1.4, this method provides another method to prove Malle's conjecture. However, the method in [Wan17] does not require any information about the error from both S_n extensions for $n = 3, 4, 5$ and A extensions, therefore we could get the main term proven in more cases there. Neither method subsumes the other.

Based on these new examples, we give the following conjecture.

Conjecture 1. *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 $\text{ind}(G_2) > \frac{n_2}{n_1 c_1}$, then the asymptotic distribution of $G_1 \times G_2$ extensions has a secondary term in the order of X^{c_1/n_2} .*

We organize the paper as following. In section 2, we give a summary on necessary results as preliminaries. This includes the description of discriminant of the compositum of two disjoint number fields, and the product lemma. In section 3, we prove a new uniformity estimate on partially ramified cubic fields by geometric sieve. In section 4, we apply the sieve method to prove the main theorem.

Notations

p : a finite place in base field k or a prime number

$|\cdot|$: absolute norm $\text{Nm}_{k/\mathbb{Q}}$

$\text{Disc}(K)$: absolute norm of $\text{disc}(K/\mathbb{Q})$

$\text{Disc}_p(K)$: p part of $\text{Disc}(K)$

\tilde{K} : Galois closure of K over \mathbb{Q}

$\text{ind}(\cdot)$: the index $n - \#\{\text{orbits}\}$ for a cycle or minimum value of index among non-identity elements for a group

$N_k(G, X)$: the number of isomorphic classes of G -extension over k with Disc bounded by X

2 Preliminaries

In this section, we will give a summary on the discussion of the discriminant of compositum and the product argument in the author's previous paper [Wan17].

2.1 Discriminant of $S_3 \times A$

Throughout the section, K and L are finite field extensions but not necessarily Galois. We will denote $\text{Gal}(K/\mathbb{Q})$ to be the Galois group $\text{Gal}(\tilde{K}/\mathbb{Q}) \subset S_n$ as a permutation subgroup which acts on the n embeddings of K into $\bar{\mathbb{Q}}$ where $n = \text{deg}(K/\mathbb{Q})$.

Given a pair of extensions (K, L) over \mathbb{Q} where $\text{Gal}(K) = G_1 \subset S_n$ and $\text{Gal}(L) = G_2 \subset S_m$ which intersect trivially, our goal is to determine the discriminant $\text{Disc}(KL)$ completely. If we know the ramification information of K and L completely, we would be able to pin down the absolute discriminant of the compositum $\text{Disc}(KL)$. Indeed, isomorphism classes of étale extensions of degree n over \mathbb{Q} (or \mathbb{Q}_p) are in one-to-one correspondence to $G_{\mathbb{Q}} \rightarrow S_n$ ($G_{\mathbb{Q}_p} \rightarrow S_n$) up to relabeling the letters, therefore we get the map $\phi_1 : G_{\mathbb{Q}_p} \hookrightarrow G_{\mathbb{Q}} \rightarrow S_n$ for K and similarly ϕ_2 for L that record the complete local information. The direct product of $\phi_1 \times \phi_2 : G_{\mathbb{Q}_p} \rightarrow S_n \times S_m \subset S_{mn}$ corresponds to the local étale extension $(KL)_p := (KL) \otimes_{\mathbb{Q}} \mathbb{Q}_p$ with $\text{Disc}((KL)_p/\mathbb{Q}_p) = \text{Disc}_p(KL)$. Therefore the data of ϕ_1 and ϕ_2 is sufficient to determine the discriminant.

Moreover at a prime p where both K and L are tamely ramified, we can determine $\text{Disc}_p(KL)$ with less data. Such primes are all but finitely many. In this case, the exponent of p in the discriminant could be determined by the inertia group as a permutation subgroup. A tamely ramified inertia group could be generated by a single element, so let's denote the inertia group of K and L by $I_1 = \langle g_1 \rangle$ and $I_2 = \langle g_2 \rangle$.

Theorem 2.1 ([Wan17], Theorem 2.2, 2.3). *Let K and L be as given above which are both tamely ramified at p . Let g_i for $i = 1, 2$, with order e_i , be the generator of inertia group at p for*

\tilde{K} and \tilde{L} . The generator g_1 is a product of k disjoint cycles $\prod c_k$ and g_2 is a product of l disjoint cycles $\prod d_l$. Then the exponent for p in $\text{Disc}_p(KL)$ is $mn - \sum_{k,l} \gcd(|c_k|, |d_l|)$. Moreover, if $(e_1, e_2) = 1$, then the exponent for p in $\text{Disc}_p(KL)$ is $mn - kl$.

At wildly ramified primes, in most cases we do not have a direct way to compute, but there are only finitely many extensions over \mathbb{Q}_p with bounded degree, so there are only finitely many possibilities for ϕ_1 and ϕ_2 for a certain pair of (G_1, G_2) . It will be shown in the discussion that they only affect the coefficient of the main term and the secondary term.

2.2 Product Argument

In this section, we are going to include the product lemma on two distributions with different order of growth. We will apply this lemma on tail estimates. Denote $F_i, i = 1, 2$, to be asymptotic distribution of some multi-set of positive integers S_i , i.e., $F_i(X) = \#\{s \in S_i \mid s \leq X\}$, and denote the product distribution $P_{a,b}(X) = \#\{(s_1, s_2) \mid s_i \in S_i, s_1^a s_2^b \leq X\}$ where $a, b > 0$.

Lemma 2.2 ([Wan17], Lemma 3.2). *Let $F_i(X), i = 1, 2$ be as given above and $F_i(X) \sim A_i X^{n_i} \ln^{r_i} X$ 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 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

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

3 Uniformity

In this section we prove a new uniformity result on partially ramified S_3 cubic extensions at finitely many primes and merge this uniformity result with previous known uniformity estimates on totally ramified S_3 cubic fields. Although we only need these results over \mathbb{Q} , all of the results hold over arbitrary number field k .

Let k be a number field and q be a square-free integral ideal in \mathcal{O}_k . Let us denote $N_{q,r}(S_3, X)$ 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 we have Proposition 6.2 from [DW88]:

Theorem 3.1 ([DW88], Proposition 6.2). *The number of non-cyclic cubic extensions over k which are totally ramified at a product of finite places $r = \prod p_i$ is:*

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

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

On the other hand, by an argument in the author's previous work [Wan17] based on the geometric sieve method introduced in [Bha14], we get the following uniformity estimates on partially ramified extensions.

Theorem 3.2. *The number of non-cyclic cubic extensions over k which are partially ramified at a product of finite places $q = \prod p_i$ is:*

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

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 4.5 in [Wan17] 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 by the same way. As a corollary of Theorem 3.2, 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 3.3. *Given a square-free integral ideal q , the 3-class number summed over quadratic extensions F/k with $q \mid \text{disc}(F/k)$ is bounded by*

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

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 $\text{Disc}(F) = \text{Disc}(K_3)$. If $q \mid \text{disc}(F)$, then the cubic field K_3 is partially ramified at q . Therefore

$$\sum_{\substack{[F:k=2] \\ q \mid \text{disc}(F), \text{Disc}(F) \leq X}} \frac{h_3^*(F/k) - 1}{2} = O\left(\frac{X}{|q|^{1/6-\epsilon}}\right). \quad (3.1)$$

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 3.2. Rearranging the expression, and applying Theorem 4.2 [Wan17] on quadratic extensions

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

we have that

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

□

And by combining the Theorem 3.1 and 3.2 using class field theory, we prove the following theorem.

Theorem 3.4. *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*

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

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 $\text{disc}(F)$. 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 $q|\text{disc}(F)$. 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

$$\begin{aligned}
& \sum_{\substack{[F:k]=2 \\ q|\text{disc}(F)}} \sum_{\substack{r|f \\ |f|^2 \text{Disc}(F) \leq X}} 4^{\omega(f)} h_3^*(F/k) \\
&= 4^{\omega(r)} \sum_f 4^{\omega(f)} \sum_{\substack{[F:k]=2 \\ q|\text{disc}(F), \text{Disc}(F) \leq \frac{X}{|f|^2|r|^2}}} h_3^*(F/k) \\
&\leq 4^{\omega(r)} \sum_f 4^{\omega(f)} \frac{X}{|f^2 r^2| |q|^{1/6-\epsilon}} \\
&\leq O\left(\frac{X}{|q|^{1/6-\epsilon}|r|^{2-\epsilon}}\right) \sum_f \frac{4^{\omega(f)}}{|f|^2} \leq O\left(\frac{X}{|q|^{1/6-\epsilon}|r|^{2-\epsilon}}\right).
\end{aligned} \tag{3.2}$$

□

Proof of Theorem 3.2. 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 $\text{GL}_2(\mathbb{Z})$ -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

$$L^1 = \#\{x \in mrB \cap V_{\mathbb{Z}}^{(i)} \mid x(\text{mod } q) \in Y(\mathbb{Z}/q\mathbb{Z})\} = O(C^{\omega(q)}) \cdot \max\left\{\frac{\lambda^4}{q}, \lambda^3 t^3\right\}, \tag{3.3}$$

over the fundamental domain of $\text{GL}(\mathbb{R})/\text{GL}(\mathbb{Z})$ 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

$$\begin{aligned}
N(S; X) &\leq O(C^{\omega(q)}) \frac{1}{M_i} \int_{\lambda=O(1)}^{O(X^{1/4})} \int_{t=\sqrt[4]{3}/\sqrt{2}}^{O(\lambda^{1/3})} \max\left\{\frac{\lambda^4}{q}, \lambda^3 t^3\right\} t^{-2} dt^\times d\lambda^\times \\
&= O(C^{\omega(q)}) \frac{1}{M_i} \int_{\lambda=O(1)}^{O(X^{1/4})} \max\left\{\frac{\lambda^4}{q}, \lambda^3 \lambda^{1/3}\right\} d\lambda^\times \\
&= O(C^{\omega(q)}) \cdot \max\left\{\frac{X}{q}, X^{5/6}\right\} = O(C^{\omega(q)}) \cdot \max\left\{\frac{X}{q}, X^{5/6}\right\}.
\end{aligned} \tag{3.4}$$

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

4 Main Proof

In this section we are going to prove Theorem 1.3 and Theorem 1.4. 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 1.3 and 1.4. 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 1.3, and the amount of power saving away from the main term for cases in Theorem 1.4.

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 Σ_p as follows: if $p \notin T$, let Σ_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 Σ_p be the set of all possible local étale extensions over \mathbb{Q}_p for an S_3 cubic extension. Similarly, we define Λ_p for A -extensions at $p \notin T$ and $p \in T$ separately. Therefore define $\mathcal{A} = \{((12)), ((123))\}$ 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 σ_p at $p \in T$, or if \tilde{K} has σ_p as the inertia group at $p \notin T$. Similarly for L . By the way Σ_p and Λ_p are defined, all $K \in \sigma_p$ have the same discriminant, $\text{Disc}_p(K)$, so we could denote this number $\text{Disc}(\sigma_p)$. Similarly for $\text{Disc}(\lambda_p)$ for A extensions.

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

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

The exponent $e(\sigma_p, \lambda_p)$ for $p \notin T$ could be determine by Theorem 2.1, and in such cases $e(\sigma_p, \lambda_p)$ is independent of p and only depends on the permutation presentation of σ_p and λ_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 $1 \leq i \leq |A|$ and $1 \leq j \leq |B|$, 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

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

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

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

Therefore if we could get an estimation of $B(\rho^0, X)$ for every ρ , we just need to sum $B(\rho^0, XL_\rho)$

over all ρ in this form to get the final counting

$$G(X) = \#\{(K, L) \mid \text{Disc}(KL) \leq X\} = \sum_{\rho} B(\rho^0, XL_{\rho}).$$

In order to get $B(\rho^0, X)$, we apply a sieve method. We will say that ρ_1 *divides* ρ_2 if they contain the same $w \in \mathcal{W}$, and for each i and j , we have $q_{ij}^{(1)} \mid q_{ij}^{(2)}$, where $q_{ij}^{(k)}$ is the associated square-free number at the i, j -th position in ρ_k . Given ρ_1 and ρ_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

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

where we define $\mu(\eta)$ to be $\prod_{i,j} \mu(q_{ij})$ with q_{ij} the i, j -th square-free integer in η . Here we write ϱ^1 in short for $\varrho^{(1,1,\dots,1)}$, which means that we require no condition on places outside q_{ij} in ϱ . 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 ρ 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 4.1 ([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*

$$N_{q,r}(S_3, X) = AA_q A_{r,2} X + BB_q B_{r,2} X^{5/6} + O(C_q C_{r,2} X^{2/3+\epsilon}),$$

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 $C_q C_{r^2} \leq O(\prod_{e|q} e^{4/5} \prod_{e|r} e^{4/5}) \leq O(qr)^{4/5+\epsilon}$. 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 $q | \text{Disc}(L)$, then we have the following estimate.

Theorem 4.2 ([Wan17], Theorem 4.13). *Let A be a finite abelian group and q be an integer, then*

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

where C and the implied constant only depends on k but not on q .

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

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

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

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

and denote $F_{1,\Sigma}(X) = \sum_{\text{Disc}(K) \leq X, K \in \Sigma} 1$. Similarly for A extensions, we will use $F_2(X)$, $F_{2,\Lambda}(X)$, $g(s)$ and $g_\Lambda(s)$. Then Theorem 4.2 says that given a local condition Λ on ramification behavior, we have that

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

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

Notice that given a tuple ϱ 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 ϱ have the same support outside T with non-trivial ramification restriction. Indeed, say q_{ij} are the i, j -th square-free number in ϱ , 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 σ_p (and λ_p) in $w \in \mathcal{W}$. For technical reasons, if we do not include the conditions at T into Σ , then we denote the smaller local condition Σ' . For brevity, we will call the corresponding coefficients depending on Σ and Λ in Theorem 4.1 and 4.2 by A_Σ , B_Σ , C_Σ and D_Λ 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

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

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

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

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

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

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 $s = \frac{1}{\text{ind}(A)}$, 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 4.1, A_Σ and B_Σ are precise constants determined and are the local densities at $s = 1$ and $s = 5/6$, while C_Σ 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 $\text{ind}(b_j)$, we mean the index of the group element.

Lemma 4.3. *Let Λ be a local condition on ramification for A extensions and let g_Λ be the corresponding Dirichlet series, we have that at $s > 1/a(A)$, the value of $g_\Lambda(s)$ is bounded by*

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

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 Λ .

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 continuous homomorphisms $J_\mathbb{Q}/\mathbb{Q}^* \rightarrow A$, and it is equivalent to consider maps $\rho : \prod_p \mathbb{Z}_p^* \rightarrow A$ [Woo16]. Local conditions on the abelian extensions could also be formulated by local conditions on ρ . At places outside T , the condition of Λ is equivalent to the condition that the image of \mathbb{Z}_p^\times under ρ 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 ρ . We then have

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

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}_Λ are both multiplicative, we can get the estimate for \tilde{g}_Λ easily,

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

□

Plugging in the value from Lemma 4.3, we get that

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

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 ρ

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 η . 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 Σ induced by ρ and similarly for $\Lambda(\rho)$. Like we define $\text{Disc}(\sigma_p)$, we could also define $\text{Disc}(\Sigma(\rho))$ to be $\prod_p \text{Disc}(\sigma_p)$ where the product is over all $p \in T$ and $p|\rho_{ij}$ for all i and j , and define $\text{Disc}(\Sigma'(\rho))$ to be $\prod_p \text{Disc}(\sigma_p)$ where the product is over all $p|\rho_{ij}$ for all i and j , then $\text{Disc}(\Sigma(\rho\eta)) = \text{Disc}(\Sigma(\rho)) \text{Disc}(\Sigma'(\eta))$.

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

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

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

For small η , by the inclusion-exclusion, the error is

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

The last inequality comes from the fact that C_Σ and D_Λ are multiplicative up to $O(1)$ at most.

For big η , 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

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

and similarly for the secondary term,

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

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 η , then

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

where a and b are such numbers that $C_p = p^a$ and $C_{p^2} = p^b$ in Theorem 4.1. We know from Theorem 4.1 that we can take $a = b = 4/5$. We will keep a and b to see how much we need from them. For Λ ,

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

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,

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

Here β and γ 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 4.4. *Given a vector of component β and γ 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*

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

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, β and γ 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 4.1 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

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

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 η 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 4.5. *Given a vector of component β and γ 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*

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

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

The exponent β'_{ij} is the same as β_{ij} in R_1 , but the exponent γ'_{ij} is different. By description of A_Σ and D_Λ , the quotient is

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

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

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

By Lemma 4.5 and description of $A_{\Sigma(\rho)}$, $D_{\Lambda(\rho)}$, we have

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

Similarly for W_3 , the exponent

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

and

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

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

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

we have

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

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_ρ in $B(\rho^0, XL_\rho)$ and denote the error by E_ρ . Plugging in

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

we get

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

where

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

and

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

Here $\text{ind}((12)(3), b_j)$ 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 ρ

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

Recall in Theorem 3.4 and 4.2, we get uniformity estimates for S_3 cubic extension and A -extension with restriction on ramification, which states that

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

and

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

Since ρ^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 2.2 on $\text{Disc}_{res}(K) = \frac{\text{Disc}(K)}{\text{Disc}(\Sigma(\rho))}$ and $\text{Disc}_{res}(L) = \frac{\text{Disc}(L)}{\text{Disc}(\Lambda(\rho))}$, we get

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

where

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

and

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

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

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

1. Estimates for E_S

The sum E_S for the small range is

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

For the second equality, we apply Lemma 4.4 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

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

For the second equality, we apply Lemma 4.5 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_ρ ; for large range, we use estimates which is purely error term D_ρ . Since the first term and secondary term are both small comparing to D_ρ ,

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

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_ρ without harm. Finally we get that the error is

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

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

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

and

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

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 $\frac{5}{6m}$,

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

and

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

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 ϵ 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

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

in which p is the smallest prime divisor of m .

Firstly, recall that if the following quantity is positive then

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

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

$$\begin{aligned} e(1, j) &= a + 2/3 - \text{ind}((12)(3), b_j) \cdot (2/3m + a(\beta, \gamma)) + \epsilon, \\ e(2, j) &= b + 4/3 - \text{ind}((123), b_j) \cdot (2/3m + a(\beta, \gamma)) + \epsilon, \\ d(1, j) &= \frac{5}{6} + \epsilon - \text{ind}((12)(3), b_j)/m, \\ d(2, j) &= \epsilon - \text{ind}((123), b_j)/m, \end{aligned}$$

where $a = b = 4/5$ as in Theorem 4.1. We will need the following lemma.

Lemma 4.6 ([Wan17], Lemma 2.4). *Let A be an abelian group of odd order m and (12) , (123) be elements in S_3 . Then for all $c \in A$, $\text{ind}((12), c)/m > 2$, $\text{ind}((123), c)/m > 1$.*

Proof of Theorem 1.3 and Theorem 1.4. 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

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

and in all of our cases, we can check that the maximum value of U_i and V_i are obtained when $\text{ind}(b_j) = \text{ind}(A)$. 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 2.1 we have

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

where the maximum is taken when $\text{ind}(b_j) = \text{ind}(A)$. 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 Δ :

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

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

When $p = 5, 7$, we have

$$\begin{aligned} U_1 &= \frac{1 + a - 4\Delta/3 - (a - 2\Delta + 1)(1 + 2\Delta)(1 + 3\Delta)^{-1} + \epsilon}{1/6 + a + 2\Delta/3 - (a - 2\Delta + 1)(1 + 2\Delta)(1 + 3\Delta)^{-1}}, \\ U_2 &= \frac{1 + b - 2\Delta/3 - (a - 2\Delta + 1)(2 + \Delta)(1 + 3\Delta)^{-1} + \epsilon}{2 + b + \Delta/3 - (a - 2\Delta + 1)(2 + \Delta)(1 + 3\Delta)^{-1}}. \end{aligned}$$

It suffices to check the following holds

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

for $p = 7$, and when $p = 5$ the inequality holds when the right hand side is 1, which finishes the proof of Theorem 1.3 and Theorem 1.4 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

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

which is equivalent to

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

It follows from the Lemma 4.6. \square

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 $G_{\mathbb{Q}} \rightarrow C_l$ 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 $\text{Aut}(C_l)$. The generating series for such maps is

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

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

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

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

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

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

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

where both sums are over all η that is relatively prime to ρ . Finally we sum over all ρ and get the main term for the whole counting

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

with a secondary term

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

where we sum over all possible ρ .

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 ρ , could be parametrized by a pair of relatively prime square-free integers, say q and r . Similarly for η , say k and l , with $klpq$ square-free. So plugging in all the constant we get the coefficient for the main term:

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

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

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

and the local factor at l

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

Similarly, we can compute the constant for the secondary term

$$\begin{aligned}
C_2 &= B \sum_{\rho} L_{\rho}^{5/6m} B_{\Sigma(\rho)} \cdot g_{\Lambda(\rho)} \left(\frac{5}{2m} \right) \sum_{\eta} \mu(\eta) B_{\Sigma'(\eta)} \cdot g_{\Lambda'(\eta)} \left(\frac{5}{2m} \right) \\
&= B \cdot g \left(\frac{5}{2l} \right) \cdot d_l \cdot \sum_{\substack{q,r,k,l \\ \forall p|qrkl \\ p \equiv 1 \pmod{l}}} \left\{ \mu(k) \mu(l) \left(\frac{1}{q^{\Delta} r^{2\Delta}} \right)^{5/6} \cdot \prod_{p|qk} \frac{(1+p^{-1/3})^2}{K_p \cdot p} \right. \\
&\quad \left. \cdot \prod_{p|rl} \frac{(1+p^{-1/3})}{K_p \cdot p^2} \cdot \prod_{p|qrkl} \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right\} \tag{4.35} \\
&= B \cdot g(5/2l) \cdot d_l \cdot \prod_{p \equiv 1 \pmod{l}} \left\{ 1 + p^{-5\Delta/6} \cdot \frac{(1+p^{-1/3})^2}{K_p \cdot p} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right. \\
&\quad \left. + p^{-5\Delta/3} \cdot \frac{1+p^{-1/3}}{K_p \cdot p^2} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right. \\
&\quad \left. - \frac{(1+p^{-1/3})^2}{K_p \cdot p} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} - \frac{1+p^{-1/3}}{K_p \cdot p^2} \cdot \frac{(l-1)p^{-5\Delta/2}}{1+(l-1)p^{-5\Delta/2}} \right\},
\end{aligned}$$

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

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

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

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

and the value for the constant of the secondary term is

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

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

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

and

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

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

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

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 $\delta = 24/(283m) - \epsilon$.

5 Acknowledgement

I am extremely grateful to my advisor Melanie Matchett Wood for many helpful discussions. I would like to thank Frank Thorne for helpful conversations, thank Manjul Bhargava, Takashi Taniguchi and Frank Thorne for providing a preprint of their work, and thank Robert Harron for helpful communications and references. I would also like to thank David Roberts, Takashi Taniguchi and Frank Thorne for suggestions on an earlier draft. This work is partially supported by National Science Foundation grant DMS-1301690 and Vilas Early Career Investigator Award.

References

- [Bha05] M. Bhargava. The density of discriminants of quartic rings and fields. *Ann. of Math.*, 162(2):1031–1063, September 2005.
- [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.
- [BTT16] M. Bhargava, T. Taniguchi, and F. Thorne. Secondary terms in counting functions for cubic fields, II. *preprint*, 2016.
- [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.
- [Hou10] B. Hough. Equidistribution of Heegner points associated to the 3-part of the class group. *preprint*, 2010.
- [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.
- [Rob01] D. Roberts. Density of cubic field discriminants. *Math. Comp.*, 70(236):1699–1705, 2001.
- [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.
- [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.
- [Wri89] D. J. Wright. Distribution of discriminants of abelian extensions. *Proc. of London Math. Soc. (3)*, 58(1):1300–1320, 1989.
- [Zha13] Y. Zhao. On sieve methods for varieties over finite fields. *preprint*, 2013.

Jiuya Wang, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, 480 LINCOLN DR., MADISON, WI 53706, USA

E-mail address: `jiuyawang@math.wisc.edu`