Lagarias and Weiss *Research in Number Theory* (2015) 1:7
DOI 10.1007/s40993-015-0006-6

Research in Number Theory
a SpringerOpen Journal

RESEARCH ARTICLE Open Access

Splitting behavior of S_n-polynomials

Jeffrey C Lagarias[1*] and Benjamin L Weiss[2]

*Correspondence:
lagarias@umich.edu
[1] Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, USA
Full list of author information is available at the end of the article

Abstract

We analyze the probability that, for a fixed finite set of primes S, a random, monic, degree n polynomial $f(x) \in \mathbb{Z}[x]$ with coefficients in a box of side B satisfies: (i) $f(x)$ is irreducible over $\mathbb{Q}$, with splitting field $K_f/\mathbb{Q}$ over $\mathbb{Q}$ having Galois group S_n; (ii) the polynomial discriminant $Disc(f)$ is relatively prime to all primes in S; (iii) $f(x)$ has a prescribed splitting type $(\mathrm{mod}\ p)$ at each prime p in S.

The limit probabilities as $B \to \infty$ are described in terms of values of a one-parameter family of measures on S_n, called z-splitting measures, with parameter z evaluated at the primes p in S. We study properties of these measures.

We deduce that there exist degree n extensions of $\mathbb{Q}$ with Galois closure having Galois group S_n with a given finite set of primes S having given Artin symbols, with some restrictions on allowed Artin symbols for $p < n$. We compare the distributions of these measures with distributions formulated by Bhargava for splitting probabilities for a fixed prime p in such degree n extensions ordered by size of discriminant, conditioned to be relatively prime to p.

Mathematics Subject Classification: Primary 11R09; Secondary 11R32; 12E20; 12E25

1 Background

By an *S_n-polynomial* we mean a degree n monic polynomial $f(x) \in \mathbb{Z}[x]$ whose splitting field $K_f/\mathbb{Q}$, obtained by adjoining all roots of $f(x)$ has Galois group S_n. It is well known that with high probability a "random" degree n monic polynomial with integer coefficients independently drawn from a box $[-B, B]^n$ is irreducible and is an S_n-polynomial. In 1936, van der Waerden [44] showed that this probability approaches 1 as the box size $B \to \infty$. For such a polynomial, adjoining one root of $f(x)$ gives an S_n-number field. Later authors obtained quantitative versions giving explicit bounds for the cardinality of the exceptional set; see Section 5.1.

This paper considers a refinement of this problem: to study the set of polynomials with coefficients in a box $[-B, B]^n$ which are S_n-polynomials prescribed to have a given splitting behavior at a given finite set of primes $\{p_k : 1 \le k \le r\}$. It shows the existence of limiting splitting densities as $B \to \infty$, conditional on the discriminant, $\mathrm{Disc}(f)$, of the polynomial f being relatively prime to $\prod_{i=1}^{r} p_i$. This conditioning imposes a non-ramification condition, requiring the polynomials to have square-free factorizations $(\mathrm{mod}\ p_i)$ with $1 \le i \le r$. This conditioning has two important consequences:

(1) The square-free assumption permits the limiting splitting densities to be interpreted as a set of probability distributions on the symmetric group S_n, which depend on the prime p. These distributions are constant on conjugacy classes of S_n.
(2) The resulting limit of distributions possess an interpolation property as p varies. The splitting densities are the values at $z = p$ of a one-parameter family of complex-valued measures $\nu^*_{n,z}$ on the symmetric group S_n which we call *z-splitting measures.* The interpolation property is: the values $\nu^*_{n,z}(g)$ on fixed elements $g \in S_n$ are rational functions in the parameter z.

These limiting splitting densities at $z = p$ have a simple origin. They are inherited from corresponding densities for splitting of polynomials in p-adic fields recently studied by the second author [45], which in turn arise from splitting probabilities for polynomials over finite fields. The latter probabilities are evaluated by counting the monic polynomials over $\mathbb{F}_p$ having various square-free factorization types in $\mathbb{F}_p[X]$, for which there are explicit combinatorial formulas.

The first contribution of this paper is to introduce and study the z-splitting measures on S_n, and show that for parameter values $z = p$ they are limiting splitting distributions for S_n-polynomials above as the box size $B \to \infty$. A second contribution is to to compare and contrast the limiting probabilities of the model of this paper to a recent probability model of Bhargava [2], which considers algebraic number fields of degree n, called S_n-number fields, whose normal closure has Galois group S_n. Bhargava's model concerns limiting splitting probabilities of a fixed prime p taken over S_n-number fields having discriminant bounded by a parameter D, as $D \to \infty$. The interesting feature is that the limiting probabilities of the two models do not agree. We now describe these two contributions in more detail.

1.1 Existence and properties of *z*-splitting measures

The paper directly defines the z-splitting measures as rational functions of z by a combinatorial formula given in Definition 2.2, and studies their basic properties in Section 4. Only later in the paper do we show that for z a prime power these p^k-splitting densities coincide with the limiting densities for splitting of S_n-polynomials, doing this for $k = 1$ in Section 5 over the rational field $\mathbb{Q}$ and for general k in Section 6 for polynomials with coefficients over general number fields.

The splitting types of a square-free monic polynomial (mod p) of degree n are described by partitions μ of n, which are identified with conjugacy classes on the symmetric group S_n. For each $n \geq 1$, and for each prime p in Section 2.1 we define *p-splitting measures* $\nu^*_{n,p}(\cdot)$ on S_n which are constant on conjugacy classes C_μ of S_n. We show the following results, whose precise statements are given in Section 2.

(i) For a fixed prime p, the limiting probabilities as $B \to \infty$ for degree n monic polynomials $f(x) \in \mathbb{Z}[x]$ conditioned on $p \nmid \mathrm{Disc}(f)$ to have a given splitting type μ exist and are given by the values $\nu^*_{n,p}(C_\mu)$ (See Theorem 2.4). For a fixed splitting type μ the values $\nu^*_{n,p}(C_\mu)$ as functions of the prime p are interpolated by a rational function $R_\mu(z) \in \mathbb{C}(z)$, where we have $R_\mu(p) = \nu^*_{n,p}(C_\mu)$ holding for each p. This *rational function interpolation property* yields a parametric family of (complex-valued) measures $\nu^*_{n,z}$ on S_n for $z \in \mathbb{P}^1(\mathbb{C}) \setminus \{0, 1\}$, termed *z-splitting measures.*

(ii) The z-splitting measure is a positive probability measure whenever n is an integer greater than 1 and $z = t$ is a real number greater than $n-1$. The uniform distribution on S_n is the z-splitting measure for $z = \infty \in \mathbb{P}^1(\mathbb{C})$ (See Theorem 2.3).

(iii) There exist infinitely many S_n-number fields having prescribed splitting types (p_i, μ_i) at a given finite set of primes $S = \{p_1, \ldots, p_r\}$, provided that all the splitting types have $\nu^*_{n,p_i}(C_{\mu_i}) > 0$. The latter conditions are satisfied if and only if there exists an S_n-number field K with a subring of algebraic integers that is a monogenic order with discriminant relatively prime to $\prod_i p_i$. The existence of one such number field K certifies that the associated probability $\nu^*_{n,p_i}(\mu_i) > 0$ (See Theorem 2.5).

(iv) For each $n \geq 2$ there is a finite set of *exceptional pairs* (p_i, μ_i) having $\nu^*_{n,p_i}(C_{\mu_i}) = 0$. The exceptional primes p_i necessarily satisfy $2 \leq p_i \leq n-1$, and this set is nonempty for $n \geq 3$. The exceptional pairs correspond to the condition that all S_n-number fields having such a splitting type (p_i, μ_i) have the prime p_i as an *essential discriminant divisor* (a notion defined in Section 2.3) (Theorem 2.6). The phenomenon of essential discriminant divisors was first noted in 1878 by Dedekind [14].

The z-splitting measures $\nu^*_{n,z}$ seem of intrinsic interest, and arise in contexts not considered in this paper. First, the measure $\nu^*_{n,k}$ may also have an interesting representation theoretic interpretation for integer values $z = k$, viewing the measure as specifying a rational character of S_n. The first author will show elsewhere that this is the case for $z = 1$, where the measure is a signed measure supported on the Springer regular elements of S_n [28]. Second, for $z = p^k$ these measures arise in a fundamental example in the theory of representation stability being developed by Church, Ellenberg and Farb ([5],[6]), see Section 7.2.

1.2 Bhargava S_n-number field splitting model

The probability model for polynomial factorization (mod$p$) studied in this paper has strong parallels with a probability model developed by Bhargava [2] for the splitting of primes in certain number fields $K/\mathbb{Q}$ of degree n.

Bhargava defines an *S_n-number field* $K/\mathbb{Q}$ to be a number field with $[K : \mathbb{Q}] = n$ whose Galois closure L over $\mathbb{Q}$ has Galois group S_n. Thus $[L : \mathbb{Q}] = n!$ while $[K : \mathbb{Q}] = n$. An S_n-number field K is a non-Galois extension of $\mathbb{Q}$ for $n \geq 3$. Bhargava's probability model takes as its sample space, with parameter D, the set of all S_n-number fields K of discriminant $|D_K| \leq D$ with the uniform distribution; his results and conjectures concern limiting behavior of the splitting densities at a fixed prime p as $D \to \infty$, conditioned on the restriction that the field K be unramified at (p), i.e. $p \nmid D_K$, the (absolute) field discriminant of K. He formulates conjectures about these limiting distributions for splitting of a fixed prime (p) and proves them for $n \leq 5$. These conjectures are unproved for $n \geq 6$.

There is a close connection of S_n-number fields with S_n-polynomials, which relates the two models. Any primitive element θ of an S_n-field that is an algebraic integer has θ being a root of an S_n-polynomial. Conversely, adjoining a single root of an S_n-polynomial $f(x)$ always yields a field $K = \mathbb{Q}(\theta)$ that is an S_n-extension in Bhargava's sense. In the case that $p \nmid \mathrm{Disc}(f)$, where $\mathrm{Disc}(f)$ is the polynomial discriminant, the splitting type of the polynomial $f(x)$ (mod$p$) determines the splitting type of the prime ideal (p) in $K/\mathbb{Q}$, and also the Artin symbol $\left[\frac{K_f/\mathbb{Q}}{(p)}\right]$ (which is a conjugacy class

in S_n). The probability model of this paper can then be interpreted as studying pairs (K, α) in which K is an S_n-number field, given with an element $\alpha \in O_K$ such that $K = \mathbb{Q}(\alpha)$, with a finite sample space specified by size restrictions on the coefficients that the (monic) minimal polynomial of α satisfies. Bhargava's model samples fields K with one distribution, while the model of this paper samples (K, α) with another distribution.

We discuss Bhargava's model in detail in Section 3. Our main observation is that the limiting probability distributions of the two models do not agree: the p-splitting measures depend on p, while Bhargava's measures are the uniform measure on S_n, which is independent of p. We also observe that Bhargava's limit measure, the uniform measure on S_n, arises as the $p \to \infty$ limit of the z-splitting measures. In Section 3.2 we present a detailed comparison of the structural features of the models, and identify differences. However we do not have a satisfying conceptual explanation that accounts for the differences of the limiting probabilities in the two models, and leave finding one as an open question.

1.3 Plan of paper

Section 2 states the main results. Section 3 discusses Bhargava's number field splitting model and compares its predicted probability distributions with the model of this paper. Section 4 derives basic properties of the splitting probabilities. Section 5 obtains the limiting distributions of splitting probabilities for polynomials with integer coefficients in a box. These splitting probabilities are essentially inherited from the analogous splitting probabilities for random monic polynomials over finite fields, see Section 4.4. We also establish result (iii) above on existence of infinitely many S_n-number fields having given splitting types at a finite set of primes, avoiding exceptional pairs. Section 6 extends the splitting results of this paper to monic polynomials with coefficients in rings of integers of a fixed number field, choosing boxes based on a fixed $\mathbb{Z}$-basis of the ring of integers. The answer involves the splitting measures $\nu^*_{n,q}(C_\mu)$ for $q = Np = p^f$, with $f \geq 1$. This generalization is an application of results of S. D. Cohen [11]. Section 7 discusses generalizations of the splitting problem to random matrix ensembles, as well as other appearances of z-splitting densities.

Notation. Our notation for partitions differs from Macdonald [33]. We denote partitions of n by $\mu = (\mu_1, \ldots, \mu_k)$, with $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k$, where Macdonald uses λ; and the multiplicity of part i of μ is denoted $c_i(\mu) := |\{j : \mu_j = i\}|$, where Macdonald uses $m_i(\lambda)$. We sometimes write a partition of n in bracket notation as $\mu = \langle 1^{c_1}, 2^{c_2}, \cdots, n^{c_n} \rangle$, with only $c_i = c_i(\mu) > 0$ included, following Stanley [41].

2 Results

2.1 Splitting measures

The results in this paper are expressible in terms of a discrete family of probability distributions on the symmetric group S_n indexed by $q = p^k$. These distributions belong to a one-parameter family of complex-valued measures on S_n, depending on a parameter $z \in \mathbb{C} \smallsetminus \{0, 1\}$ given below, which we call *z-splitting measures.* Restricting the parameter to real values $z = t \in \mathbb{R} \smallsetminus \{0, 1\}$ we obtain signed measures of total mass 1, and all the parameter values $t = q = p^k$ which are prime powers give nonnegative probability measures on S_n; these measures originally arose in statistics involving the factorization of random square-free polynomials over $\mathbb{F}_q[X]$, see Section 4.4.

Definition 2.1. For each degree $m \geq 1$ the m-th *necklace polynomial* $M_m(X)$ by

$$M_m(X) := \frac{1}{m} \sum_{d|m} \mu(d) X^{m/d}.$$

where $\mu(d)$ is the Möbius function.

The necklace polynomial takes integer values at integers n, its values at positive integers have an enumerative interpretation that justifies its name, given in Section 4.1. These polynomials arise in our context because for $X = q = p^f$ a prime power, $M_m(q)$ counts the number of irreducible monic degree m polynomials in $\mathbb{F}_q[X]$, where $\mathbb{F}_q$ is the finite field with q elements, see Lemma 4.1.

For a given element $g \in S_n$, denote its cycle structure (lengths of cycles) by $\mu = \mu(g) =: (\mu_1, \mu_2, \ldots \mu_k)$ with $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_k$. Here we regard μ as an *unordered partition* of n, though for convenience we have listed its elements in decreasing order, and we denote it $\mu \vdash n$. The conjugacy classes on S_n consist of all elements g with a fixed cycle structure and we denote them C_μ. For a partition $\mu \vdash n$ we let

$$c_i = c_i(\mu) := |\{j : \mu_j = i\}|$$

count its number of parts of size i, and we sometimes denote it by the bracket notation $\mu = \langle 1^{c_1}, 2^{c_2}, \cdots, n^{c_n} \rangle$, with only $c_i > 0$ included.

Definition 2.2. *The* z-splitting measure $\nu_{n,z}(g)$ for $g \in S_n$ *is given by*

$$\nu^*_{n,z}(g) := \frac{1}{n!} \cdot \frac{1}{z^{n-1}(z-1)} \prod_{i=1}^{n} i^{c_i} c_i! \binom{M_i(z)}{c_i(\mu)}, \tag{2.1}$$

where for a complex number w we interpret $\binom{w}{k} := \frac{(w)_k}{k!} = \frac{w(w-1)\cdots(w-k+1)}{k!}$.

For each fixed $g \in S_n$, the quantity $\nu^*_{n,z}(g)$ is a rational function of z, and is well-defined away from the polar set, which is contained in $z = 0, 1$. The splitting measure of an individual element g depends only on its cycle type $\mu = \mu(g)$, so is constant on conjugacy classes C_μ of S_n. Using the well known formula

$$|C_\mu| = n! \prod_{i=1}^{n} \frac{i^{-c_i(\mu)}}{c_i(\mu)!}, \tag{2.2}$$

for the size of conjugacy classes ([41], Proposition 1.3.2), we obtain

$$\nu^*_{n,z}(C_\mu) := \sum_{g \in C_\mu} \nu^*_{n,z}(g) = \frac{1}{z^{n-1}(z-1)} \prod_{i=1}^{n} \binom{M_i(z)}{c_i(\mu)}. \tag{2.3}$$

Properties of these measures are studied in Section 4. The measures are defined by the right side of (2.3) as complex-valued measures for all z on the Riemann sphere,

excluding $z = 0$. The definition implies that they have total mass one, in the sense that

$$\sum_{g \in S_n} \nu^*_{n,z}(g) = 1.$$

In this paper we restrict to real values $z = t$, in which case $\nu^*_{n,t}(g)$ in general defines a signed measure on S_n.

In Section 4.5 we prove results specifying positive real z-values where the z-splitting measure is nonnegative. In particular we show nonnegativity of the measure holds for all positive integers $t = m \geq 2$.

Theorem 2.3. *Let $n \geq 2$. The z-splitting measures $\nu^*_{n,z}$ have the following properties, for positive real parameters $z = t > 1$.*

1. *For all real $t > n - 1$, one has*

$$\nu^*_{n,t}(g) > 0 \ \text{ for all } \ g \in S_n,$$

For these parameter values $\nu_{n,t}(\cdot)$ is a probability measure with full support on S_n.

2. *For integers $k = 2, 3, \ldots, n - 1$, one has*

$$\nu^*_{n,k}(g) \geq 0 \ \text{ for all } \ g \in S_n,$$

so that $\nu^*_{n,k}(\cdot)$ is a probability measure on S_n. For these parameter values this measure does not have full support on S_n. It is zero on the conjugacy class of the identity element $C_{\langle 1^n \rangle}$.

3. *As $t \to \infty$ through positive real values, one has*

$$\lim_{t \to \infty} \nu^*_{n,t}(g) = \frac{1}{n!}.$$

In Section 4.6 we prove a complementary result specifying negative real z-values where the z-splitting measure is nonnegative. In particular, nonnegativity holds for all negative integers $m \leq -1$ (Theorem 4.7). This result is not used elsewhere in the paper.

We also note that a later result (Theorem 2.6) below refines case (1) of Theorem 2.3 to characterize for each n all pairs (p, μ) with p a prime and $\nu^*_{n,p}(C_\mu) = 0$.

2.2 Prime splitting densities of S_n-polynomials

Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial. Consider for a prime p the splitting of such polynomials (mod p), viewed in $\mathbb{F}_p[X]$.

More generally for $q = p^f$, any monic $f(x) \in \mathbb{F}_q[x]$ factors uniquely as $f(x) = \prod_{i=1}^{k} g_i(x)^{e_i}$, where the e_i are positive integers and the $g_i(x)$ are distinct, monic, irreducible, and non-constant. We may define the *splitting type* of such a polynomial (following Bhargava [2]) to be the formal symbol

$$\mu_q(f) := \left(\deg(g_1)^{e_1},\ \deg(g_2)^{e_2},\ \ldots, \deg(g_k)^{e_k}\right)$$

where k is the number of distinct irreducible factors of $f(x)$. Here we order the degrees in decreasing order. We let T_n denote the set of all possible formal symbols for degree n polynomials, which we call *splitting symbols*. Thus $T_3 = \{(111), (21), (3), (1^2 1), (1^3)\}$.

Using this definition, given any monic $f(x) \in \mathbb{Z}[x]$ and any prime p, we may assign to it a splitting type $\mu_p(f) \in T_n$.

This paper mainly restricts to *square-free splitting types*, which are those having all $e_i = 1$. We define $T_n^* \subset T_n$ to denote the set of such splitting types. Thus $T_3^* = \{(111), (21), (3)\}$. Each element $\mu := (\mu_1, \mu_2, \ldots \mu_k) \in T_n^*$ with $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_k$ specifies a partition of n, to which there is associated a unique conjugacy class $C_\mu \subset S_n$. The conjugacy class C_μ is the set of all elements of S_n whose cycle lengths are equal to the (unordered) numbers $\mu_1, \ldots, \mu_k$. In this case we will refer also to C_μ as a *splitting type*, and if $\mu_p(f) \in T_n^*$ then we will write $\mu_p(f) = C_\mu$.

Given any positive integer n and a positive number B we let $\mathcal{F}_n(B)$ denote the collection of all degree n monic polynomials with integer coefficients,

$$f(x) = x^n + \sum_{j=0}^{n-1} c_j x^j \in \mathbb{Z}[x],$$

having coefficients bounded by $-B < c_j \leq B$, for $0 \leq j \leq n-1$. Then let $\mathcal{F}_n(B;p)$ be the subset of monic polynomials in $\mathcal{F}_n(B)$ having the following properties:

(i) The polynomial discriminant $(\mathrm{Disc}(f), p) = 1$.

(ii) All coefficients of $f(x)$ are contained in $[-B+1, B]$. This implies that the polynomial discriminant $|\mathrm{Disc}(f)| \leq (4B)^{n(n-1)}$.

(iii) $f(x)$ is irreducible over $\mathbb{Q}$ and the degree n number field $K_f = \mathbb{Q}(\theta_f)$ generated by one root has normal closure with Galois group S_n.

The allowed splitting types $(\mathrm{mod}\, p)$ of polynomials in $\mathcal{F}_n(B;p)$ are constrained by the requirement (i) on the discriminant to belong to T_n^*, i.e. to be square-free $(\mathrm{mod}\ p)$. For this case we show:

Theorem 2.4. *(Limiting Splitting Densities) Let $n \geq 2$ be given. Then:*
(1) For each prime p, there holds

$$\lim_{B\to\infty} \frac{\#\{f(x) \in \mathcal{F}_n(B;p)\}}{\#\{f(x) \in \mathcal{F}_n(B)\}} = 1 - \frac{1}{p}. \tag{2.4}$$

(2) For each (square-free) splitting type $\mu \in T_n^*$, there holds

$$\lim_{B\to\infty} \frac{\#\{f(x) \in \mathcal{F}_n(B;p) \mid (p) \text{ has splitting type } C_\mu\}}{\#\{f(x) \in \mathcal{F}_n(B;p)\}} = \nu_{n,p}^*(C_\mu), \tag{2.5}$$

where $\nu_{n,p}^*$ is the splitting measure for n with parameter $t = p$, i.e.

$$\nu_{n,p}^*(C_\mu) = \frac{1}{p^{n-1}(p-1)} \prod_{i=1}^{n} \binom{M_i(p)}{c_i(\mu)}. \tag{2.6}$$

This result is proved in Section 5; it is a special case $r = 1$ of Theorem 5.2 which applies more generally to finite sets $\mathcal{S} = \{p_1, p_2, \ldots, p_r\}$ of primes. In Section 6 we give a further generalization of the result to algebraic number fields.

2.3 Existence of S_n-Number Fields with Prescribed Prime Splitting

We also show there are infinitely many S_n-number fields with prescribed prime splitting at any finite set $\mathcal{S}$ of primes, of those types allowed by the splitting measures. The splitting measures impose some extra constraints associated to the existence of monogenic orders in the S_n-number fields having discriminants relatively prime to given elements.

Theorem 2.5. *Let $n \geq 2$ be given, let $\mathcal{S} = \{p_1, \ldots, p_r\}$ denote a finite set of (distinct) primes, and let $\mathfrak{U} = \{\mu_1, \ldots, \mu_r\}$ be a prescribed set of (not necessarily distinct) splitting symbols for these primes. Then the following conditions are equivalent.*

(1) The positive measure condition

$$\nu^*_{n,p_i}(C_{\mu_i}) > 0 \text{ for } 1 \leq i \leq r$$

holds.

(2) There exists an S_n-number field K having the following two properties:

- *(P1) The field K contains a monogenic order $O = \mathbb{Z}[\,1, \theta, \ldots \theta^{n-1}]$ whose discriminant is relatively prime to $p_1 p_2 \cdots p_r$.*
- *(P2) The Galois closure K^{spl} of $K/\mathbb{Q}$ is unramified at all prime ideals above those in $\mathcal{S}$ and the primes in $\mathcal{S}$ have prescribed Artin symbols*

$$\left[\frac{K^{spl}/\mathbb{Q}}{(p_i)}\right] = C_{\mu_i}, \qquad 1 \leq i \leq r.$$

(3) There exist infinitely many S_n-number fields K having properties (P1) and (P2).

The condition (1) automatically holds when all $p_i \geq n$, because the probability measure $\nu^*_{p_i,n}$ then has full support on the group S_n. However for primes $2 \leq p < n$ there are restrictions on the allowed splitting behavior. This restriction has to do with the non-existence of monogenic maximal orders satisfying (P1) for S_n-number fields having specific prime factorization at small prime ideals. The polynomials $f(x)$ generating such fields have *essential discriminant divisors*, as defined in Cohn ([12], Defn. 9.55, Lemma 10.44c) and Cohen ([8], p. 197). A famous example due to Dedekind [14] (see [12], Exercise 9.4; Lemma 10.44c) is an S_3-number field K for which the prime ideal (2) splits completely in K; all monogenic orders then have an even index, and correspondingly $\nu^*_{3,2}([\,1^3]) = 0$. However it is known that infinitely many S_3-number fields K exist in which the ideal (2) splits completely in the maximal order. This result follows from results of Bhargava for $n = 3$ discussed in Section 3.1. Such fields are not covered by Theorem 2.5. (Related concepts include the *inessential discriminant divisor* $I(K)$ of a field K (Tormhein [42]), also called the *non-essential discriminant divisor* of K (Sliwa [40]). Here $I(K) = \gcd_{\theta \in O_K} i(\theta)$, where $i(\theta) := \left[O_K : \mathbb{Z}\left[1, \theta, \ldots, \theta^{n-1}\right]\right]$. The *essential discriminant divisors* are exactly the prime divisors of $I(K)$.)

Theorem 2.5 allows us to characterize the splitting measures for prime values $t = p \geq 2$ having probability 0 in terms of field-theoretic data.

Theorem 2.6. *For p a prime, and a splitting type $\mu \vdash n$, for fixed $n \geq 2$, the following three conditions are equivalent.*

(C1) *The splitting measure at $t = p$ has*

$$\nu^*_{n,p}(C_\mu) = 0.$$

(C2) *There are no degree n monic polynomials $f(x) \in \mathbb{Z}[x]$ with $f(x) \pmod p$ having a square-free factorization of splitting type C_μ.*

(C3) *All S_n-number fields K in which (p) is unramified and has splitting type μ necessarily have p as an essential discriminant divisor.*

This result is proved in Section 5.4. The condition (C1) is vacuous for $n = 1, 2$. This theorem provides the easy-to-check criterion (C1) for an S_n-number field K to have (p) as an essential discriminant divisor, via the splitting type μ of (p) in K. Condition (C2) is a statement about all $f(x) \in \mathbb{Z}[x]$; it does not require $f(x)$ to be an S_n-polynomial or to be irreducible over $\mathbb{Q}$. Our proof does not show the existence of even a single field satisfying (C3) for any given pair (p, μ) satisfying (C1). Conjecture 3.2 of Bhargava below would imply that infinitely many such fields exist, and this conjecture is known to be true for $n \leq 5$.

In Section 6 we establish generalizations of Theorem 2.5 and Theorem 2.6 above in which the base field $\mathbb{Q}$ is replaced by an algebraic number field K. These generalizations are stated as Theorems 6.2 and 6.3, respectively. These generalizations are more complicated to state, but their proofs are straightforward, using results of S. D. Cohen [11].

3 Bhargava number field splitting model

Recall from Section 1.2 that Bhargava defines an *S_n-number field* $K/\mathbb{Q}$ to be a number field with $[K : \mathbb{Q}] = n$ whose Galois closure L over $\mathbb{Q}$ has Galois group S_n. Bhargava's number field splitting model has sample space the set of all S_n-number fields K of discriminant $|D_K| \leq D$ with the uniform distribution, and his results and conjectures concern the limiting behavior of splitting densities at a fixed prime p as $D \to \infty$, conditioned on the property that the field K be unramified at (p), i.e. $p \nmid D_K$, the (absolute) field discriminant of K.

3.1 Bhargava's conjectures for prime splitting in S_n-number fields

In 2007 Bhargava [2] formulated conjectures about the splitting of primes averaged over S_n-number fields ordered by the size of their field discriminants. Bhargava developed his conjectures based on the following principle ([2], p. 10):

> The expected (weighted) number of global S_n-number fields of discriminant D is simply the product of the (weighted) number of local extensions of $\mathbb{Q}_v$ that are discriminant-compatible with D, where v ranges over all places of $\mathbb{Q}$, (finite and infinite).

In this statement a *local extension of* $\mathbb{Q}_v$ means a degree n étale algebra E over $\mathbb{Q}_v$ (not necessarily a field) and *discriminant-compatible* means that the valuation of the discriminant of E matches that of D and that, in the archimedean case, the signs of the discriminants match. We state two of his conjectures below in order to later compare them with our results.

Firstly, given any positive integer n and a positive number B we let $\mathcal{G}_n(B)$ denote the collection of S_n-number fields K that have discriminants $|D_K| \le B$. Secondly, given any positive integer n, prime p and positive number B we let $\mathcal{G}_n(B;p)$ denote the collection of all degree n number fields K such that:

(i) The ideal (p) is unramified in K;
(ii) The field discriminant $|D_K| \le B$;
(iii) The degree n field K over $\mathbb{Q}$ has normal closure having Galois group S_n.

The first conjecture of Bhargava concerns which fraction of S_n-number fields have field discriminant K relatively prime to p ([2], Conj. 1.4).

Conjecture 3.1 (Bhargava). *Fix a prime p and a positive integer n. Then*

$$\lim_{B\to\infty} \frac{\#\{K \in \mathcal{G}_n(B;p)\}}{\#\{K \in \mathcal{G}_n(B)\}} = 1 - \rho_n(p). \tag{3.1}$$

where $\rho_n(p)$ is the "probability of ramification", given by

$$\rho_n(p) := \frac{\sum_{k=1}^{n-1} q(k, n-k)p^{n-1-k}}{\sum_{k=0}^{n-1} q(k, n-k)p^{n-1-k}}, \tag{3.2}$$

in which $q(k,n)$ denotes the number of partitions of k into at most n parts.

By convention we set $q(0,n) = 1$ for $n \ge 1$. This distribution $\rho_n(p)$ depends on both n and p and is a rational function of p. For fixed n, $\rho_n(p) = \frac{1}{p} + O\left(\frac{1}{p^2}\right)$ as $p \to \infty$.

Bhargava proves Conjecture 3.1 for $n \le 5$. For these cases, the probabilities are $\rho_1(p) = 0$ and $\rho_2(p) = \frac{1}{p+1}$, $\rho_3(p) = \frac{p+1}{p^2+p+1}$, $\rho_4(p) = \frac{p^2+2p+1}{p^3+p^2+2p+1}$, and $\rho_5(p) = \frac{p^3+2p^2+2p+1}{p^4+p^3+2p^2+2p+1}$, respectively. In another conjecture, Bhargava ([2], Conjecture 5.2) further relates these probabilities to the distribution of splitting types in T_n having repeated factors.

Bhargava's second conjecture about prime splitting in S_n-number fields is as follows ([2], Conj. 1.3).

Conjecture 3.2 (Bhargava). *Fix a prime p, a positive integer n, and $\mu \in T_n^*$. Then*

$$\lim_{B\to\infty} \frac{\#\{K \in \mathcal{G}_n(B,p)\} \mid p \text{ has Artin symbol in } C_\mu\}}{\#\{K \in \mathcal{G}_n(B;p)\}} = \nu_n(C_\mu), \tag{3.3}$$

where $\nu_n(\cdot)$ denotes the Chebotarev density distribution on conjugacy classes of S_n, which is

$$\nu_n(C_\mu) := \frac{|C_\mu|}{|S_n|}$$

Conjecture 3.2 predicts that the limiting density exists and agrees with that predicted by the Chebotarev density theorem for conjugacy classes (see [29], [37], Chap. 7, §3, [38], Theorem 13.4); this measure corresponds to the uniform distribution on S_n. This limiting distribution depends on n but is independent of p. It is proved for $n \le 5$. The case $n = 3$ is deducible from results of Davenport and Heilbronn [13], see also Cohen et al. [9]. Bhargava proved the result for $n = 4$ and $n = 5$ using his earlier results for discriminant density in quartic and quintic fields [1,3].

For a general viewpoint on Bhargava's conjectures, see Venkatesh and Ellenberg ([43], Section 2.3). Bhargava's conjectures on local mass formulas, were reinterpreted in connection with Galois representations in Kedlaya [26] and further cases were considered by Wood [47,48].

3.2 Random polynomial model versus random number field model

We compare the distributions for prime splitting in S_n number fields in the random polynomial model against those of the random number field model given in Bhargava's conjectures. These splitting distributions differ.

We summarize the comparison in Table 1. A main feature is that for each $n \geq 1$ the densities random monic polynomial model in the $p \to \infty$ limit approaches the uniform density distribution conjectured in Bharagava's model.

Both model predictions assign a weighted contribution of S_n-number fields K having discriminant prime to p, which depend on the parameter B (resp. D), and consider the limiting distribution as the corresponding parameter grows. In each model the splitting density is a conditional probability based on conditioning against an "unramifiedness" condition. There is a difference of scale in the cutoffs in the B and D parameters between the two models, in that polynomial discriminants D_f grow proportionally to B^n. However the limit as the parameters go to infinity, this scale differences play no role.

The main differences in the predicted probabilities in the models are the following.

(1) In Bhargava's conjectures the probability of ramification $\rho_n(p)$ depends on both the prime p and the degree n. One has

$$\theta_n(p) := 1 - \rho_n(p) = \frac{1}{1 + \sum_{k=1}^{n-1} q(k, n-k)p^{-k}},$$

This formula implies that for fixed p and variable n the function $\rho_n(p)$ increases to the limit $\rho_\infty(p) := 1 - \frac{1}{P(1/p)}$ where

$$P(x) := \sum_{n=0}^{\infty} p_n x^n = \prod_{n=1}^{\infty} \left(\frac{1}{1-x^n} \right).$$

Table 1 Comparison of polynomial splitting model and random S_n-number field model probabilities

Probability model	Random S_n-polynomial model	Random S_n-number field model (Bhargava)
Sample space	Degree n, monic polynomials with integer coefficients $\lvert c_i \rvert \leq B$, generating an S_n-number field	S_n-number fields K with field discriminant $\lvert D_K \rvert$ bounded by D
Limit procedure	Box size $B \to \infty$	Discriminant $D \to \infty$
Ramification probability at (p)	Prob[p divides $Disc(f)$] equals $\frac{1}{p}$, which is independent of n	Prob[p divides $Disc(K)$] is a quantity $\theta_n(p)$ which depends on both n and p (Conjecture 3.1)
Limiting distribution on S_n of splitting types	p-splitting distribution $\nu^*_{n,p}(C_\mu)$ on conjugacy classes, whose probabilities depend on both n and p	Chebotarev distribution $\nu_n(C_\mu) = \frac{\lvert C_\mu \rvert}{n!}$, which is independent of p (Conjecture 3.2)
Limit $p \to \infty$ of ramification probability	0	0
Limit $p \to \infty$ of distribution densities	Uniform distribution $\nu^*_{n,\infty} = \nu_n$ on elements of S_n	Uniform distribution $\nu^*_{n,\infty} = \nu_n$ on elements of S_n

(Conjectures 3.1 and 3.2 are theorems for $n \leq 5$.)

In contrast, in the random polynomial model the probability of ramification $\frac{1}{p}$ is independent of n, according to Theorem 2.4 (1). The formula above implies that for fixed n one has $\rho_n(p) = \frac{1}{p} + O\left(\frac{1}{p^2}\right)$ as $p \to \infty$, so both ramification probabilities go to 0 as $p \to \infty$ at the same rate.

(2) In Bhargava's conjectures the splitting probabilities are independent of both p and n. In contrast, in the random polynomial model the probabilities $\nu^*_{n,p}(C_\mu)$ depend on both n and p.

What features of the models account for the differing answers in the two models? The models themselves have structural differences.

(D1) The (irreducible) polynomial f is associated algebraically not with the ring of integers O_K of the field $K = \mathbb{Q}(\theta)$ generated by a root θ of f, but with the particular monogenic order $O_f = \mathbb{Z}[1, \theta, \theta^2, \cdots, \theta^{n-1}]$. In particular discriminant $\mathrm{Disc}(f) = D_K c^2$, where $c = [O_K : O_f]$ is the index of O_f inside f. In particular $\mathrm{Disc}(f)$ may be divisible by primes which do not divide D_K, so the "unramified" conditions of the two probability models differ. For some S_n-number fields K the ring of integers O_K is not monogenic. The number of monogenic orders of a given index in the maximal order O_K (isomorphism up to an additive shift of a variable) is known to depend on the index within a given field K, cf. Evertse [16].

(D2) Many different polynomials in $\mathcal{F}_n(B; p)$ generate the same S_n-number field $K = K_f$. Thus each field K that occurs is weighted by the number of polynomials in the box that generate it (and which satisfy the discriminant co-primeness condition). The weights depend in a complicated way on K and B and change as $B \to \infty$.

The difference (D1) of the ramification conditions in the two models presumably accounts for much of the mismatch. The S_n-number fields detected by the random polynomial model are always unramified in the field sense, but the random monic polynomial models do not detect some S_n-number fields not ramified at (p). We should really replace the p-part of $\mathrm{Disc}(f)$ with the p-part of D_K, with $K = \mathbb{Q}(\theta)$, which involves studying the p-adic coefficients of $f(x)$. The model of Bhargava is based on a mass formula counting p-adic étale extensions with weights, and the weights matter. However from the viewpoint of the difference (D2) it is not immediately clear that such weighted sums will conspire to produce the nice limiting values given in Theorem 2.4. To understand difference (D2) better it might be interesting to study an auxiliary question: for each pair of S_n-number fields K_1, K_2 what is the behavior as $B \to \infty$ of the ratio of the number of $f(x)$ in the box of size B that generate the field K_1 (resp. K_2) and satisfy $p \nmid \mathrm{Disc}(f(x))$. Does this quantity have a limiting value and if so, how does it depend on K_1 and K_2?

We conclude that there are observable structural differences between the two models. We do not currently have a conceptual explanation how these structural differences account for and quantitatively explain the differences in the limiting densities of the two models.

4 Splitting measures

In this section we define and study the one-parameter family of splitting measures $\nu^{\star}_{n,z}$ ($z \in \mathbb{C}$) on the symmetric group S_n, for each n. We relate this measure at $z = q = p^k$ to finite field factorization of degree n monic polynomials over $\mathbb{F}_q$.

4.1 Necklace polynomials

The number of monic irreducible polynomials of degree m over finite fields $\mathbb{F}_q$ for $q = p^f$ are well known to be interpolatable by universal polynomial $M_m(X)$ evaluated at value $X = q$. Recall that for $m \geq 1$ the *necklace polynomial* of degree m is $M_m(X) \in \mathbb{Q}[X]$ by

$$M_m(X) := \frac{1}{m}\left(\sum_{d|m} \mu\left(\frac{m}{d}\right) X^d\right) = \frac{1}{m}\left(\sum_{d|m} \mu(d) X^{\frac{m}{d}}\right), \tag{4.1}$$

where $\mu(d)$ is the Möbius function. For $m = 0$ we set $M_0(X) = 1$. We note that $M_1(X) = X$ and $M_2(X) = \frac{1}{2}X(X-1)$. Clearly $M_m(X) \in \frac{1}{m}\mathbb{Z}[X]$, for $m \geq 1$.

The name "necklace polynomial" was proposed by Metropolis and Rota [35], because the value $M_m(k)$ for positive integer k has a combinatorial interpretation as counting the number of necklaces of m distinct colored beads formed using k colors which have the property of being *primitive* in the sense that their cyclic rotations are distinct (Moreau [36]). In 1937, Witt ([46], Satz 3) showed that $M_m(k)$ counts the number of basic commutators of degree m in the free Lie algebra on k generators. See Hall [24] and the discussion in Hazewinkel ([23], Sect. 17).

For later use we give some basic properties of $M_m(X)$.

Lemma 4.1. *(1) Let $q = p^k$ be a prime power and let $N_m^{irred}(\mathbb{F}_q)$ count the number of irreducible monic polynomials in $\mathbb{F}_q[X]$ of degree m. Then*

$$M_m(q) = N_m^{irred}(\mathbb{F}_q).$$

(2) The polynomial $M_m(X) \in \mathbb{Q}[X]$ is an integer-valued polynomial, i.e. one has $M_m(k) \in \mathbb{Z}$ for all $k \in \mathbb{Z}$.

Proof. (1) The well known formula $N_m^{irred}(\mathbb{F}_q) = M_m(q)$ was found by Gauss[a] in the unpublished Section 8 of *Disquisitiones Arithmeticae*, Articles 342 to 347, see Gauss ([21], pp. 212–240), cf. Maser [34]. A proof is given in Rosen ([39], p. 13).

(2) To verify that a polynomial in $\frac{1}{m}\mathbb{Z}[X]$ is integer-valued, it suffices to check the integrality property holds at m consecutive integer values of X. The integrality property at positive integers follows from the counting interpretation of the values $M_m(j)$ of Moreau [36], see also [35]. □

We next obtain bounds on the size of $M_m(X)$ which will be used in Sections 4.5 and 4.6 to establish non-negativity properties of the z-splitting distributions for certain parameter ranges.

Lemma 4.2. *(1) The necklace polynomial $M_m(X)$ has $M_m(0) = 0$ for $m \geq 1$. In addition*

$$M_m(1) = \begin{cases} 1 & \text{for } m = 1, \\ 0 & \text{for } m \geq 2. \end{cases}$$

One has $(X-1)^2 \nmid M_m(X)$ for all $m \geq 2$.

(2) One has

$$M_m(t) > 0, \quad \textit{for all real } \ t \geq 2.$$

In addition, for $1 \leq j \leq m$ *there holds for real* $t > m-1$,

$$M_j(t) > \left\lfloor \frac{m}{j} \right\rfloor - 1. \tag{4.2}$$

(3) For $m \geq 1$ *one has*

$$(-1)^m M_m(-t) > 0, \quad \textit{for all real } \ t \geq 2. \tag{4.3}$$

In addition, for each $m \geq 2$ *and* $t > 0$ *with* $t(t+1) > m-2$, *there holds for* $1 \leq j \leq m/2$,

$$M_{2j}(t) > \left\lfloor \frac{m}{2j} \right\rfloor - 1. \tag{4.4}$$

Proof. (1) We have $M_m(0) = 0$ since it has no constant term for $m \geq 1$. For $m \geq 1$ we have $M_m(1) = \sum_{d|m} \mu(d)$, which yields $M_m(1) = 0$ for $m \geq 2$. Thus $X(X-1)|M_m(X)$ for $m \geq 2$. The relation $(X-1)^2$ does not divide $M_m(X)$ follows from

$$M_m'(X)|_{X=1} = \frac{1}{m}\left(\sum_{d|m} \mu(d)\frac{m}{d}\right) = \prod_{p|m}\left(1 - \frac{1}{p}\right) > 0.$$

(2) For $m \geq 2$ and real $t \geq 2$, one has

$$m\,M_m(t) \geq t^m - \left(\sum_{j=1}^{\lfloor m/2 \rfloor} t^j\right) = t^m - \left(\frac{t^{\frac{m}{2}+1} - t}{t-1}\right) \geq t^m - t^{\frac{m}{2}+1} + t > 0. \tag{4.5}$$

For the second part, suppose $m \geq 2$ and $1 \leq j \leq m$. We have for $j = 1$ and $t > m-1$ that $M_1(t) = t > m-1$. For $j = 2$ and $t > m-1$ we have

$$M_2(t) = \frac{1}{2}t(t-1) \geq \left\lfloor \frac{m}{2} \right\rfloor - 1,$$

the last inequality being immediate for $m = 2$ and easy for $m \geq 3$. Finally, for $3 \leq j \leq m$, and $t > m-1$, we have $t^j - t^{\frac{j}{2}+1} \geq 1$, whence by (4.5),

$$jM_j(t) \geq 1 + t > m$$

which gives (4.2) in this case.

(3) To establish $M_m(-t) > 0$ for $t > 2$, note that for $m = 1$ one has $-M_1(-t) = t > 0$ for $t > 0$. For $m \geq 2$ we have for $t > 2$ that

$$m\,M_m(-t) \geq t^m - \left(\sum_{j=1}^{\lfloor m/2 \rfloor} t^j\right) \geq t^m - t^{\frac{m}{2}} + t > 0. \tag{4.6}$$

For the second part, it suffices to show for $1 \leq j \leq m/2$ that

$$(2j)M_{2j}(-t) > m - 2j \quad \left(\geq 2j\left(\left\lfloor \frac{m}{2j} \right\rfloor - 1\right)\right).$$

For $2j = 2$ we have by hypothesis

$$2M_2(-t) = t(t+1) > m-2.$$

For $2j \geq 4$ and $m \geq 6$, the condition $t(t+1) > m-2$ implies $t > 2$. Then (4.6) applies and we obtain.

$$2jM_{2j}(-t) \geq t^{2j} - t^j + t \geq t(t+1) > m-2 > m-2j.$$

as required. The remaining case is $m = 4$ and $2j = 4$, where $m - 2j = 0$, the condition $t(t+1) > 2$ implies $t > 1$, whence

$$4M_4(-t) = t^4 - t^2 > 0,$$

as required. □

4.2 Cycle polynomials

To any partition $\mu \vdash n$ we associate the *cycle polynomial*

$$N_\mu(X) := \prod_{i=1}^{n} \binom{M_i(X)}{c_i(\mu)} \tag{4.7}$$

Here $N_\mu(X) \in \mathbb{Q}[X]$ is a polynomial of degree n (since $\sum_{i=1}^{n} ic_i = n$).

The values $N_\mu(X)$ for prime powers $X = q = p^f$ count the number of square-free polynomial factorizations of type μ in $\mathbb{F}_q[X]$, as shown in Section 4.4.

Lemma 4.3. *(Properties of Cycle Polynomials) Let $n \geq 2$. For any partition $\mu \vdash n$ the cycle polynomial $N_\mu(X)$ has the following properties.*

(1) The polynomial $N_\mu(X) \in \frac{1}{n!}\mathbb{Z}[X]$ is integer-valued.

(2) The polynomial $N_\mu(X)$ has lead term

$$\left(\prod_{i=1}^{n} \frac{1}{i^{c_i(\mu)} c_i(\mu)!}\right) X^n = \frac{|C_\mu|}{n!} X^n.$$

(3) The polynomial $N_\mu(X)$ is divisible by X^m, where $m \geq 1$ counts the number of distinct cycle lengths appearing in μ.

(4) There holds

$$\sum_{\mu \vdash n} N_\mu(X) = X^{n-1}(X-1). \tag{4.8}$$

Proof. (1) The definition (4.7) implies that $N_\mu(X) \in \frac{1}{d(\mu)}\mathbb{Z}[X]$ with

$$d(\mu) = \prod_{i=1}^{n} i^{c_i(\mu)} c_i(\mu)!$$

By comparison with equation (2.2) we have $d(\mu) = \frac{n!}{|C_\mu|}$, which shows that $d(\mu)$ divides $n!$, with equality when $\mu = \langle 1^n \rangle$. The integrality of $N_\mu(k)$ for $k \in \mathbb{Z}$ follows from the definition using the integrality of all $M_i(k)$ (Lemma 4.1(2)).

(2) The property follows by direct calculation of the top degree term in (4.7).

(3) The divisibility property is immediate from the definition (4.7) since X divides $\binom{M_i(X)}{c_i(\mu)}$ whenever $c_i(\mu) > 0$.

(4) Both sides of the identity (4.8) are polynomials of degree n, so it suffices to verify that the identity holds at $n+1$ distinct values of X. To this end, we make use of a combinatorial interpretation of $N_\mu(X)$ for $X = q = p^k$ a prime power, given in Proposition 4.5 below. The sum on the left evaluated at $X = p^k$ counts all possible degree n monic polynomials over $\mathbb{F}_q[X]$ for $q = p^k$ that have a square-free factorization, i.e. nonvanishing discriminant over $\mathbb{F}_q$. The resulting polynomial $F(X)$ satisfies $F(q) = q^n - q^{n-1}$, according to Proposition 4.5 (1), verifying the identity at $X = q$. □

4.3 Splitting measures

For each $n \geq 2$ we define the parametric family of (necklace) *splitting measures* $\nu^*_{n,z}$ on the symmetric group S_n, with family parameter $z \in \mathbb{C}$, by means of their values on conjugacy classes

$$\nu^*_{n,z}(C_\mu) := \frac{1}{z^{n-1}(z-1)} N_\mu(z) = \frac{1}{z^{n-1}(z-1)} \prod_{i=1}^{n} \binom{M_i(z)}{c_i(\mu)}. \tag{4.9}$$

For any element $g \in C_\mu$ we set

$$\nu^*_{n,z}(g) := \frac{1}{|C_\mu|} \nu^*_{n,z}(C_\mu). \tag{4.10}$$

The latter formula coincides with the definition (2.1) for $\nu^*_{n,z}(g)$. Since this formula is a rational function of z for each μ, with possible poles only at $z = 0, 1$, this defines a complex-valued function on S_n constant on conjugacy classes, for all z on the Riemann sphere $\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ except possibly at $z = 0, 1$. The measure at $z = \infty$ is the uniform measure, $\nu_{n,\infty}(g) = \frac{1}{n!}$, a result that follows from Lemma 4.2 (2)–see also the proof of Theorem 2.3 (3) below. The measure at $z = 1$ also turns out to be well-defined but is now a signed measure, which is studied by the first author in [28].

We next show that these measures have total (complex-valued) mass one.

Proposition 4.4. *For $n \geq 1$, for all $z \in \hat{\mathbb{C}} \smallsetminus \{0\}$ and denoting conjugacy classes in S_n by C_μ with $\mu \vdash n$,*

$$\sum_{\mu \vdash n} \nu^*_{n,z}(C_\mu) = 1.$$

Equivalently, for all $g \in S_n$,

$$\sum_{g \in S_n} \nu^*_{n,z}(g) = 1.$$

Proof. For $z \in \hat{\mathbb{C}} \smallsetminus \{0, 1, \infty\}$ the lemma follows from the normalization identity (4.8) for the cycle polynomials. It extends by analytic continuation in z to the values $z = 1, \infty$. □

4.4 Splitting measures and finite field factorizations

A main rationale for the study of z-splitting measures is that when $z = q = p^k$ is a prime power these measures occur in the statistics of factorization of monic polynomials of degree n in $\mathbb{F}_q[X]$, drawn from a uniform distribution, conditioned on being square-free.

Recall that a monic polynomial $f(x) \in \mathbb{F}_q[x]$ factors uniquely as $f(x) = \prod_{i=1}^{k} g_i(x)^{e_i}$, where the e_i are positive integers and the $g_i(x)$ are distinct, monic, irreducible, and non-constant. We have the following basic facts about square-free factorizations.

Proposition 4.5. *Fix a prime $p \geq 2$, and let $q = p^f$. Consider the set $\mathcal{F}_{n,q}$ of all monic polynomials in $\mathbb{F}_q[X]$ of degree n, so that $|\mathcal{F}_{n,q}| = q^n$.*

(1) Exactly q^{n-1} of these polynomials have discriminant $Disc(f) = 0$ in $\mathbb{F}_q$. Equivalently, exactly q^{n-1} of these polynomials are not square-free when factored into irreducible factors over $\mathbb{F}_q[X]$.

(2) The number $N(\mu;q)$ of $f(x) \in \mathcal{F}_{n,q}$ whose factorization over $\mathbb{F}_q$ into irreducible factors is square-free of degree type $\mu := (\mu_1, \ldots, \mu_r)$, with $\mu_1 \geq \mu_2 \cdots \geq \mu_r$ having $c_i = c_i(\mu)$ factors of degree i satisfies

$$N(\mu;q) = \prod_{i=1}^{n} \binom{M_i(q)}{c_i(\mu)} = N_\mu(q), \tag{4.11}$$

in which $N_\mu(X)$ denotes the cycle polynomial for μ.

Proof. (1) This result can be found in ([39], Prop. 2.3). Another proof, due to M. Zieve, is given in ([45], Lemma 4.1).

(2) This result is well known, see for example S. D. Cohen ([10], p. 256). It follows from counting all unique factorizations of the given type. □

This proposition has the following consequence.

Proposition 4.6. *Consider a random monic polynomial $g(X)$ of degree n drawn from $\mathbb{F}_q[x]$ with the uniform distribution, where $q = p^f$. Then the probability of $g(x)$ having a factorization into irreducible factors of splitting type $\mu \in T_n^*$, conditioned on $g(x)$ having a square-free factorization, is exactly $\nu_{n,q}^*(C_\mu)$. That is,*

$$\nu_{n,q}^*(C_\mu) = \mathit{Prob}[\, g(x) \text{ has splitting type } \mu \mid g(x) \text{ is square-free}\,].$$

Proof. Proposition 4.5 (1), and (2) together evaluate the conditional probability

$$\text{Prob}[\, g(x) \text{ has splitting type } C_\mu \mid g(x) \text{ is square-free}] = \frac{1}{q^n - q^{n-1}} \prod_{i=1}^{n} \binom{M_i(q)}{c_i(\mu)}.$$

Comparing the right side with the definition (2.3) of the splitting measure shows that it equals $\nu_{n,q}^*(C_\mu)$. □

4.5 Nonnegativity conditions for splitting measures: Positive real *z*

This paper is concerned with the case that $z = t$ is a real number ($t \neq 0, 1$). In this case the measure is real-valued, and is a signed measure, of total (signed) mass one by Proposition 4.4.

We now treat ranges of positive real z and prove Theorem 2.3, which specifies several real parameter ranges where these measures are nonnegative, and so define probability measures; these parameter values include all integer values $z = m \geq 2$.

Proof. To decide on nonnegativity or positivity of $\nu_{n,t}^*(C_\mu)$, it suffices to study the individual terms $\binom{M_i(t)}{c_i(\mu)}$ for $1 \leq i \leq n$ and to show nonnegativity (resp. positivity) of each of the numerators

$$(M_i(t))_{c_i(\mu)} = M_i(t)(M_i(t) - 1) \cdots (M_i(t) - c_i(\mu) + 1). \tag{4.12}$$

(1) We verify that for $t > n - 1$ and all μ, all terms in the definition (4.9) of $\nu_{n,t}^*(C_\mu)$ for $1 \leq i \leq n$ have $\binom{M_i(t)}{c_i(\mu)} > 0$. The positivity of the terms in (4.12) is immediate for $n = 1$ so suppose $n \geq 2$. Using Lemma 4.2 (2) for $t > n - 1$ we have

$$M_i(t) > \left\lfloor \frac{n}{i} \right\rfloor - 1 \geq c_i(\mu) - 1,$$

whence all factors in the product (4.12) are positive, as asserted.

(2) For each integer $2 \le k \le n-1$, the normalizing factor $\frac{1}{k^{n-1}(k-1)}$ in the definition is positive.

Since $M_i(X)$ is an integer-valued polynomial for all $i \ge 1$, each term in the product definition of $\nu^*_{n,k}(C_\mu)$ is a binomial coefficient, hence is nonnegative. This proves nonnegativity of the k-splitting measure. Finally, for the identity conjugacy class $C_{\langle 1^n \rangle} = \{e\}$, for $2 \le k \le n-1$ we have that $\nu^*_{n,k}(C_{\langle 1^n \rangle}) = 0$, since in these case the $i = 1$ factor in (4.12) has $(M_1(k))_n = 0$.

(3) The limit as $t \to \infty$ is driven by the lead term asymptotics of the polynomial $M_i(t)$. Using $\sum_i i c_i(\mu) = n$ and $M_i(t) = \frac{1}{i}t^i + O(t^{i-1})$ we obtain

$$\begin{aligned}\lim_{t\to\infty} \nu^*_{n,t}(C_\mu) &= \lim_{t\to\infty} \prod_{i=1}^{n} \frac{M_i(t)^{c_i(\mu)}}{c_i(\mu)!\ t^{ic_i(\mu)}} \\ &= \prod_{i=1}^{n} \frac{1}{c_i(\mu)!\, i^{c_i(\mu)}} = \frac{|C_\mu|}{n!},\end{aligned}$$

as asserted. □

4.6 Nonnegativity conditions for splitting measures: negative real *z*

We prove complementary results specifying some negative real parameter values $z = -t < 0$ where $\mu_{n,-t}(C_\mu)$ is nonnegative and so defines a probability measure.

Theorem 4.7. *Let $n \ge 2$. Then for real $z = -t < 0$ the signed measures $\nu^*_{n,t}$ on S_n have the following properties:*

(1) For all real values $t > 0$ having $t(t+1) > n-2$, the measure $\nu^*_{n,-t}$ on S_n is strictly positive, so that it defines a probability measure on S_n with full support.

(2) For all integers $k \ge 1$ having $k(k+1) \le n-2$ the measure $\nu_{n,-k}$ is nonnegative and defines a probability measure on S_n. This measure does not have full support. it is zero on the conjugacy class C_μ with $\mu = \langle 2^{n/2} \rangle$ if n is even, and on the conjugacy class with $\mu = \langle 1, 2^{(n-1)/2} \rangle$ if n is odd.

(3) There holds for all $g \in S_n$,

$$\lim_{t\to\infty} \nu^*_{n,-t}(g) = \frac{1}{n!}.$$

Proof. (1) To show positivity of the measure we keep track of the signs of all the factors in the definition (4.9). Since $t > 0$ the prefactor has sign

$$\mathrm{Sign}\left(\frac{1}{(-t)^{n-1}(-t-1)}\right) = (-1)^n.$$

Lemma 4.2 (3) then gives for $t(t+1) > n-2$ that

$$M_{2j}(-t) > \left\lfloor \frac{t}{2j} \right\rfloor - 1.$$

Since $c_j(\mu) \le \left\lfloor \frac{n}{j} \right\rfloor$, we obtain the positivity of all even degree terms, as

$$(M_{2j}(-t))_{c_{2j}(\mu)} = M_{2j}(-t)(M_{2j}(-t)-1)\cdots(M_{2j}(-t)-c_{2j}(\mu)+1) > 0.$$

We assert that all odd degree terms have $M_{2j+1}(-t) < 0$. Assuming this is proved, we obtain $\mathrm{Sign}((M_{2j+1}(-t))_{c_{2j+1}(\mu)}) = (-1)^{c_{2j+1}(\mu)} = (-1)^{(2j+1)c_{2j+1}(\mu)}$. It follows that

$$\mathrm{Sign}(\nu^*_{n,t}(C_\mu)) = (-1)^n(-1)^{\sum_i j c_i(\mu)} = (-1)^{2n} = 1,$$

showing the required positivity.

It remains to show that all $M_{2j+1}(-t) < 0$. This holds for $t \geq 2$ by Lemma 4.2 (3), and $t \geq 2$ whenever $n \geq 8$. For the remaining cases we check $M_1(-t) = -t < 0$ for $t > 0$, and that for $2j+1 = 3, 5, 7$ we have $M_{2j+1}(-t) = -t^{2j+1} + t < 0$ for $t > 1$,

(2) To show nonnegativity of the measure $\nu^*_{n,-k}$ for those positive integer k with $k(k-1) \leq n-2$, the argument of (1) still applies with the following changes. For even indices $2j$, we use the fact that $M_{2j}(-k)$ is a positive integer, so either the descending factorial remains positive or else is zero if a zero is encountered. So the sign of this term may be treated as positive. For the odd indices $2j+1$, either the initial value $M_{2j+1}(-k) = 0$, in which case the measure is 0, or else $M_{2j+1}(-k) < 0$ and the sign argument above applies. One has $M_{2j+1}(-k) < 0$ if $k \leq 2$ so the only problematic value is $M_{2j+1}(-1)$. This value is always 0, as may be checked. Thus nonnegativity of the measure follows.

It remains to show that the measure does not have full support. We verify for $n = 2\ell$ that for $\mu = \langle 2^\ell \rangle$, one has $\nu^*_{n,-k}(C_\mu) = 0$ for all positive integers k with $k(k+1) \leq n-2$. Here $m_2(\mu) = \ell$ and the integer

$$1 \leq M_2(-k) = \frac{1}{2}k(k+1) \leq \frac{n-2}{2} = \ell - 1,$$

so that the descending factorial $(M_2(-m))_\ell = 0$. One verifies similarly that for $n = 2\ell + 1 \geq 3$ one has $\nu_{n,-k}(C_\mu) = 0$ for $\mu = \langle 1, 2^\ell \rangle$, where again $m_2(\mu) = \ell$ and $\frac{n-2}{2} = \ell - 1$.

(3) This limit behavior follows similarly to the case of Theorem 2.3 (3). □

5 Counting S_n-polynomials with specified splitting types

5.1 Counting monic S_n-polynomials with coefficients in a box

It is well-known that, in a suitable sense, almost all monic polynomials with $\mathbb{Z}$ coefficients have a splitting field that is an S_n-extension of $\mathbb{Q}$. This was proved in 1936 by van der Waerden [44], who showed that the fraction of all monic degree n polynomials in $\mathbb{Z}[x]$ having all coefficients in a box $|a_i| \leq B$ that have a splitting field with Galois group S_n approaches 1 as $B \to \infty$. An improved quantitative form of this assertion was given in 1973 by Gallagher [20], which we formulate as follows.

Theorem 5.1 (Gallagher)**.** *For integer $B \geq 1$ let $\mathcal{F}_n(B)$ be the set of monic, degree n polynomials in $\mathbb{Z}[x]$ with all coefficients in the box $[-B+1, B]$; there are $(2B)^n$ such polynomials. Let $E_n(B)$ denote the proportion of polynomials in $\mathcal{F}_n(B)$ which do not have splitting field with Galois group S_n. Then there exists a positive constant α_n, depending only on n, such that for all $B > 2$,*

$$\frac{E_n(B)}{(2B)^n} \leq \alpha_n \frac{\log B}{\sqrt{B}}. \tag{5.1}$$

We remark that all polynomials with coefficients in the box $\mathcal{F}_n(B)$ satisfy

$$|\mathrm{Disc}(f)| \leq (4B)^{n(n-1)}. \tag{5.2}$$

Indeed, we have $\mathrm{Disc}(f) = \prod_{1 \leq i < j \leq n} (\theta_i - \theta_j)^2$, so it suffices to show that $f(x) \in \mathcal{F}_n(B)$ have all roots $|\theta_i| < 2B$. This holds because if some root $|\theta| \geq 2B$ then $|\theta^{n-j}| \leq |\theta|^n/(2B)^j$, whence

$$|a_{n-1}\theta^{n-1} + \cdots + a_1\theta + a_0| \leq B\left(\sum_{j=1}^{n} \frac{|\theta|^n}{(2B)^j}\right) = |\theta|^n \left(\sum_{j=1}^{n} \frac{1}{2^j}\right) < |\theta^n|.$$

which contradicts θ being a root of $f(x)$.

The error estimate in Gallagher's estimate was recently improved by Dietmann [15] to

$$\frac{E_n(B)}{(2B)^n} = O_\epsilon\left(B^{-(2-\sqrt{2})+\epsilon}\right). \tag{5.3}$$

Improvements of Gallagher's results in some other directions are given in Zywina [49].

5.2 Density of S_n-polynomials with specified splitting types

Our object is to refine the result above by counting the number of such polynomials generating an S_n-extension that have a given splitting type at a finite set of primes. As above, for integer B let $\mathcal{F}_n(B)$ denote the set of monic polynomials of degree n with coefficients $-B < a_i \le B$, so that $\#\{f(x) \in \mathcal{F}_n(B)\} = (2B)^n$. Theorem 2.4 is the special case $r = 1$ of the following result.

Theorem 5.2. *Let $n \ge 2$ be given, and let $\mathcal{S} = \{p_1, \dots, p_r\}$ be a finite set of primes and let $\mathfrak{U} = \{\mu_1, \dots, \mu_r\}$ be a corresponding set of splitting symbols.*

(1) Let $\mathcal{F}_n(B;\mathcal{S})$ denote the set of all polynomials $f(x)$ in $\mathcal{F}_n(B)$ such that $\gcd(Disc(f), \prod_{i=1}^r p_i) = 1$. *Then*

$$\lim_{B\to\infty} \frac{\#\{f(x) \in \mathcal{F}_n(B;\mathcal{S})\}}{\#\{f(x) \in \mathcal{F}_n(B)\}} = \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right). \tag{5.4}$$

(2) Let $\mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\})$ denote the set of all $f(x)$ in $\mathcal{F}_n(B,\mathcal{S})$ such that:

(i) $f(x)$ has splitting field K_f that is an S_n-extension of $\mathbb{Q}$.

(ii) The splitting type of $f(x) \pmod{p_i}$ is C_{μ_i} for $1 \le i \le r$.

Then

$$\lim_{B\to\infty} \frac{\#\{f(x) \in \mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\})\}}{\#\{f(x) \in \mathcal{F}_n(B;\mathcal{S})\}} = \prod_{i=1}^r \nu^*_{n,p_i}(C_{\mu_i}). \tag{5.5}$$

We note that the condition $\gcd(\mathrm{Disc}(f), \prod_{i=1}^r p_i) = 1$ on a monic irreducible polynomial guarantees that the field $K = \mathbb{Q}(\theta)$ generated by a single root of $f(x)$ is unramified over all the primes in $\mathcal{S}$. In that case, the discriminant $Disc(f)$ detects the discriminant of the ring $\mathcal{O}_f = \mathbb{Z}[1, \theta, \dots, \theta^{n-1}]$, which is a subring of the full ring of integers $\mathcal{O}(K)$ of the field $K = \mathbb{Q}(\theta)$ generated by a root of the polynomial. We have

$$\mathrm{Disc}(f) = \mathrm{Disc}(K)[\,\mathcal{O}(K) : \mathcal{O}_f]^2,$$

so that $p \nmid \mathrm{Disc}(f)$ implies $p \nmid \mathrm{Disc}(K)$.

We will derive Theorem 5.2 from two quantitative estimates given below. We begin with an estimate for the event $\gcd\left(\mathrm{Disc}(f), \prod_{i=1}^r p_i\right) = 1$.

Lemma 5.3. *Let $n \ge 2$. Let $\mathcal{S} = \{p_1, p_2, \dots, p_r\}$ and $M = \prod_{i=1}^r p_i$. Then for $B \ge 2nM$,*

$$\left|\frac{\#\{f(x) \in \mathcal{F}_n(B;\mathcal{S})\}}{\#\{f(x) \in \mathcal{F}_n(B)\}} - \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)\right| \le \frac{2nM}{B}.$$

Proof. For each prime p the behavior of $Disc(f) \pmod{p}$ is determined by $(a_0, a_1, \dots, a_{n-1}) \pmod{p}$. Thus if M divides B then Proposition 4.5 (1) shows that exactly a fraction of $\frac{1}{p}$ of these polynomials have $\mathrm{Disc}(f) \equiv 0 \pmod{p}$. The polynomials are

labelled by lattice points in the closed box $[-B+1, B]^n$, and we call a lattice point *admissible* if it corresponds to a polynomial in $\mathcal{F}_n(B, \mathcal{S})$. For a general B we first round down to a box of side $B' = M\left\lfloor \frac{B}{M} \right\rfloor$, and using there the Chinese remainder theorem we find exactly $(2B')^n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)$ admissible polynomials in the smaller box belong to $\mathcal{F}_n(B, \mathcal{S})$. This number undercounts $(2B)^n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)$ by amount $\prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)((2B)^n - (2B')^n)$. Similarly we may round up to a box of side $B'' = M\lceil \frac{B}{M} \rceil$ and using there a similar argument we find exactly $\prod_{i=1}^r (1 - \frac{1}{p_i})(2B'')^n$ admissible polynomials in the larger box. Thus

$$(2B')^n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right) \le |\mathcal{F}_n(B; \mathcal{S})| \le (2B'')^n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)$$

We now use the inequality, valid for real $|x| \le \frac{1}{2n}$,

$$1 + 2n|x| \ge (1 + x)^n \ge 1 - 2n|x|.$$

Since $B'' - B' \le M$, the inequality gives for $B \ge 2nM$,

$$(2B'')^n - (2B')^n \le (2B)^n \left(\left(1 + 2n\frac{B'' - B}{B}\right) - \left(1 - 2n\frac{B - B'}{B}\right)\right) \le (2B)^n \left(\frac{2nM}{B}\right). \tag{5.6}$$

This yields the estimate

$$\#\{f(x) \in \mathcal{F}_n(B; \mathcal{S})\} = (1 + \epsilon_n(B; \mathcal{S}))(2B)^n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right), \tag{5.7}$$

with

$$|\epsilon_n(B; \mathcal{S})| \le \frac{2nM}{B}.$$

Dividing both sides by $\#\{f(x) \in \mathcal{F}_n(B)\} = (2B)^n$ yields the desired bound. □

Now we derive the main estimate from which Theorem 5.2 will follow.

Theorem 5.4. *Let $n \ge 2$. Let $\mathcal{S} := \{p_1, p_2, \ldots, p_r\}$ be a finite set of primes and let $\mathfrak{U} := \{\mu_1, \ldots, \mu_r\}$ be a set of splitting types. Let $\mathcal{F}_n(B; \{S, \mathfrak{U}\})$ denote the set of all polynomials $f(x)$ in $\mathcal{F}_n(B)$ such that:*

- *(i)* $\gcd(Disc(f), \prod_{i=1}^r p_i) = 1$;
- *(ii)* *The splitting type of $f(x) \pmod{p_j}$ is C_{μ_j}, for $1 \le j \le r$;*
- *(iii)* *$f(x)$ has splitting field K_f that is an S_n-extension of $\mathbb{Q}$.*

Then, setting $M = \prod_i p_i$, for $B \ge 4nM$ there holds

$$\left| \frac{\#\{f(x) \in \mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})\}}{\#\{f(x) \in \mathcal{F}_n(B; \mathcal{S})\}} - \prod_{i=1}^r \nu^*_{n,p_i}(C_{\mu_i}) \right| \le 2 \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)^{-1} \alpha_n \frac{\log B}{\sqrt{B}} + \frac{4nM}{B}.$$

Proof. Let $\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})^+$ denote the set of all polynomials $f(x)$ in $\mathcal{F}_n(B)$ that satisfy conditions (i) and (ii) above. Theorem 5.1 then gives

$$0 \le |\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})^+| - |\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})| \le (2B)^n \left(\alpha_n \frac{\log B}{\sqrt{B}}\right).$$

For splitting types on box of side $B' = M\left\lfloor \frac{B}{M} \right\rfloor$ by reduction (mod M) together with Proposition 4.5 (2) and the Chinese remainder theorem we get a product distribution of all splitting types (mod p_i) for $1 \le i \le r$,

$$0 \le |\mathcal{F}_n(B'; \{\mathcal{S}, \mathfrak{U}\})^+| = (2B')^n \prod_{i=1}^{n} \frac{1}{p_i^n} N_{\mu_i}(p_i),$$

where $N_{\mu_i}(\cdot)$ is a cycle polynomial. We have a similar formula for an enclosing box of side $B'' = M\lceil \frac{B}{M} \rceil$, with $(2B'')^n$ replacing $(2B')^n$. Assuming $B \ge 2nM$ we obtain by an application of (5.6) that

$$0 \le |\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})^+| = (1 + \epsilon_n(B; \{\mathcal{S}, \mathfrak{U}\}))(2B)^n \prod_{i=1}^{n} \frac{1}{p_i^n} N_{\mu_i}(p_i),$$

with the error estimate

$$|\epsilon_n(B; \{\mathcal{S}, \mathfrak{U}\})| \le \frac{2nM}{B}.$$

Next we note that $\frac{1}{q^n} N_\mu(q) = \left(1 - \frac{1}{q}\right) \nu^*_{n,q}(C_\mu)$. Substituting this for each p_i in the formula above and using our original bound for $|\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})|$ yields

$$\left| |\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})| - (2B)^n \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right) \nu^*_{n,p_i}(C_{\mu_i}) \right| \le (2B)^n \left(\frac{2nM}{B} \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right) + \alpha_n \frac{\log B}{\sqrt{B}} \right).$$

For $B \ge 2nM$, we replace $(2B)^n \prod_i (1 - \frac{1}{p_i})$ with $|\mathcal{F}_n(B; \mathcal{S})|$ using (5.7) we obtain

$$\left| |\mathcal{F}_n(B; \{\mathcal{S}, \mathfrak{U}\})| - |\mathcal{F}_n(B; \mathcal{S})| \prod_{i=1}^{r} \nu^*_{n,p_i}(C_{\mu_i}) \right| \le (2B)^n \left(\frac{4nM}{B} \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right) + \alpha_n \frac{\log B}{\sqrt{B}} \right)$$

The result follows on dividing both sides by

$$|\mathcal{F}_n(B; \mathcal{S})| = (1 - \epsilon_n(B; \mathcal{S}))(2B)^n \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right),$$

noting for $B \ge 4nM$ that (5.7) implies $|\epsilon_n(B; \mathcal{S})| \le \frac{1}{2}$. □

Proof of Theorem 5.2. This follows directly from Lemma 5.3 and Theorem 5.4 on letting $B \to \infty$. □

Remark. The conclusion in Theorem 5.2 is insensitive to the shape of the box bounding the coefficients as long as the box increases homothetically as $B \to \infty$, e.g. one can use $-c_{n,j}B < a_j < c_{n,j}B$, where $c_{n,j}$ are positive constants independent of B, and derive exactly the same limiting formula. For example, $c_{n,j} = \binom{n}{j}$ is another natural choice. □

5.3 Existence of S_n-number fields with specified splitting types: Proof of Theorem 2.5

We first remark on a special property of the symmetric group S_n as a Galois group, represented as a permutation group acting transitively on the roots of a polynomial, that distinguishes it from some of its subgroups. Let G be a permutation group $G \subset S_n$ (i.e. a permutation representation of the abstract group G). The elements in a conjugacy class in G necessarily have the same cycle type as permutations, but the converse need not hold. That is, the cycle type of a conjugacy class in G need not determine it uniquely. For example, the group $G = \{(12)(34), (13)(24), (14)(23), (1)(2)(3)(4)\} \subset S_4$ is abelian so all conjugacy classes have size 1 but three of these classes have identical cycle structures.

This uniqueness property does hold for cycle types for the full symmetric group S_n, which has the consequence that the cycle type of an S_n polynomial having a square-free factorization (mod p) uniquely determines the Artin symbol for an S_n-number field obtained by adjoining one root of it.

Proof of Theorem 2.5. (1) $\Rightarrow$ (3). By hypothesis we are given $\mathcal{S} = \{p_1, \ldots, p_r\}$ and splitting types $\mathfrak{U} = \{\mu_1, \ldots, \mu_r\}$ with the property that all $\nu^*_{n,p_i}(C_{\mu_i}) > 0$. We will show the number of S_n-number fields K whose Galois closure K' has the given Artin symbols

$$\left[\frac{K'/\mathbb{Q}}{(p_i)}\right] = C_{\mu_i}, \qquad 1 \le i \le r \tag{5.8}$$

is infinite, by showing it is arbitrarily large. Since the splitting type of a polynomial $f(X)$ generating an S_n-number field modulo p determines the corresponding Artin symbol, it suffices to specify factorizations of polynomials (mod p_i) which we can do using Theorem 5.2.

Given $k \ge 1$ we choose $\mathcal{S}_k := \mathcal{S} \bigcup \mathcal{S}^*_k$ with $\mathcal{S}^*_k = \{p_{r+1}, \cdots, p_{r+k}\}$ being a set of k auxiliary primes that satisfy $n \le p_{r+1} < p_{r+2} < \ldots < p_{r+k}$ and disjoint from the primes in $\mathcal{S}$. In that case we may choose splitting symbols $\mathfrak{U}_k := \{\mu_{r+1}, \ldots, \mu_{r+k}\}$ arbitrarily in S_n for the auxiliary primes and the condition $\nu^*_{n,p_{r+j}}(C_{\mu_{r+j}}) > 0$ will automatically hold by Theorem 2.3 (2). The square-free condition on the polynomial modulo each p_i guarantees that the polynomial discriminant is relatively prime to $p_1 p_2 \cdots p_{r+k}$ and this property guarantees that (P1) holds. Theorem 5.2 now implies the existence of infinitely many S_n-polynomials having the given splitting behavior at all $r+k$ primes; thus (P2) holds for such fields. In particular there exists at least one such S_n-number field K exhibiting the given splitting behavior. Since each S_n for $n \ge 2$ has at least two distinct conjugacy classes, we obtain in this way at least 2^k different S_n-number fields, all of which match the splitting types C_{μ_i} for $1 \le i \le r$ in (5.8) and which are distinguishable among themselves by how the auxiliary primes p_{r+j}, $1 \le j \le k$ split. Since k can be arbitrarily large, the result follows.

(3) $\Rightarrow$ (2). Immediate.

(2) $\Rightarrow$ (1). By hypothesis the given field K possesses is a monogenic order $\mathbb{Z}[1, \theta, \ldots, \theta^{n-1}]$ satisfying (P1). The minimal polynomial for θ is then a monic polynomial $f(x) \in \mathbb{Z}[X]$ which satisfies $\gcd(Disc(f), p_1 \cdots p_r) = 1$. This polynomial then has square-free factorization (modp_i) yielding the splitting types C_{μ_i} for $1 \le i \le r$, see Lang ([30], I. §8, Proposition 25). (The hypothesis of Lang's Proposition 25 requires $\mathbb{Z}[1, \theta, \cdots, \theta^{n-1}]$ to be integrally closed, i.e. the full ring of integers O_k. As he notes, the argument can be done by localizing over each prime ideal (p_i), and here $(Disc(f), p_i) = 1$ implies that the integral closure condition holds locally.) We next observe that the splitting type conditions are congruence conditions (mod$p_1 \cdots p_n$) on the coefficients of f, and they enforce the condition $\gcd(Disc(f), p_1 p_2 \cdots p_r) = 1$. These congruence conditions are satisfied for a positive proportion of polynomials in the box, so in Theorem 5.2 the left side of (5.5) is positive, which certifies that each $\nu^*_{n,p_i}(C_{\mu_i}) > 0$. □

5.4 Vanishing values of splitting measures: Proof of Theorem 2.6

We characterize pairs (n, p, μ) where $\nu^*_{n,p}(C_\mu) = 0$.

Proof. $(C1) \Leftrightarrow (C2)$. Since $p \ne 0, 1$ the condition $\nu^*_{n,p}(C_\mu) = 0$ holds if and only if $N_\mu(p) = 0$. By Proposition 4.5 (2) the latter condition holds if any only if no degree n

monic polynomial in $\mathbb{F}_p[X]$ with $Disc(f) \neq 0 \in \mathbb{F}_p$ has a square-free factorization of splitting type μ. The latter condition is exactly (C2).

$(C1) \Leftrightarrow (C3)$. We establish the contrapositive. Suppose (C3) does not hold. This says that there exists a S_n-number field K which at (p) is unramified and has splitting type μ. Now the equivalence of Theorem 2.6 applied for a single prime $p_1 = p$ shows that $\nu^*_{n,p}(\mu) > 0$, which is equivalent to the condition that $(C1)$ does not hold.

We remark that this argument does not establish whether or not there exist any S_n extensions K which satisfy condition (C3) for given splitting data μ. □

6 Number of S_n-Polynomials with specified splitting types over number fields

We consider polynomials with coefficients drawn from an algebraic number field k, not necessarily Galois over $\mathbb{Q}$. We set $[k : \mathbb{Q}] = d$, and say that an extension L/k with $[L : k] = n$ is a *relative S_n-number field* if the Galois closure L' of L over k has $Gal(L'/k) \simeq S_n$. We let D_k denote the absolute discriminant of k over $\mathbb{Q}$.

Let $\mathcal{O}_k$ denote the ring of algebraic integers in k. We consider monic polynomials

$$f(x) = x^n + \sum_{j=0}^{n-1} \alpha_j x^j,$$

with all $\alpha_j \in \mathcal{O}_k$. Choose an integral basis $\mathcal{O}_k = \mathbb{Z}[\omega_1, \omega_2, \ldots, \omega_d]$, and let $\Omega = (\omega_1, \ldots, \omega_d)$ denote this (ordered) integral basis. We now have

$$\alpha_j = \sum_{k=1}^{d} m_{j,k}\omega_k, \quad 1 \le j \le n,$$

for unique $m_{i,j} \in \mathbb{Z}$. We define $\mathcal{F}_n(B;\Omega)$ to be the set of all monic degree n polynomials over $\mathcal{O}_k$ whose coefficients have all $m_{i,j}$ satisfying $-B+1 \le m_{i,j} \le B$, so there are $(2B)^{nd}$ polynomials in the box.

Next we let $\mathcal{S} = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_r\}$ denote a finite ordered set of (distinct) prime ideals in $\mathcal{O}_k$. We allow different ideals in the list to have residue class fields of the same characteristic, i.e. to lie over the same rational prime. We set $N_{k/\mathbb{Q}}\mathfrak{p}_j = p_j^{f_j}$. We let $\mathfrak{U} = \{\mu_1, \ldots, \mu_r\}$ denote a finite ordered set of splitting types of S_n (the different μ_j need not be distinct).

Theorem 6.1. *Suppose that $k/\mathbb{Q}$ is a number field, not necessarily Galois over $\mathbb{Q}$, Let $\mathcal{S} = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_r\}$ be an ordered finite set of distinct prime ideals in $\mathcal{O}_k$ and let $\mathbb{F}_{q_i}$ denote the residue class field for $\mathfrak{p}_i$, with $q_i = N\mathfrak{p}_i = p_i^{f_i}$. Suppose $\mathfrak{U} = \{\mu_1, \ldots, \mu_r\}$ is a given ordered set of splitting symbols. Then for fixed $n \ge 2$, the following hold.*

(1) Let $\mathcal{F}_n(B;\mathcal{S},\Omega)$ denote the set of all degree n polynomials $f(x)$ in $\mathcal{F}_n(B;\Omega)$ such that $\gcd(Disc(f), \prod_{i=1}^{r} \mathfrak{p}_i) = (1)$, *viewed as ideals in $\mathcal{O}_k$. Then*

$$\lim_{B\to\infty} \frac{\#\{f(x) \in \mathcal{F}_n(B;\mathcal{S},\Omega)\}}{\#\{f(x) \in \mathcal{F}_n(B;\Omega)\}} = \prod_{i=1}^{r}\left(1 - \frac{1}{q_i}\right). \tag{6.1}$$

(2) Let $\mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)$ denote the set of all $f(x)$ in $\mathcal{F}_n(B;\mathcal{S},\Omega)$ such that:

(i) The splitting type of $f(x) \pmod{\mathfrak{p}_i}$ is C_{μ_i} for $1 \le i \le r$.

(ii) $f(x)$ has relative splitting field K_f over k that is an S_n-extension of k.

Then

$$\lim_{B\to\infty} \frac{\#\{f(x) \in \mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)\}}{\#\{f(x) \in \mathcal{F}_n(B;\mathcal{S},\Omega)\}} = \prod_{i=1}^{r} \nu^*_{n,q_i}(C_{\mu_i}). \tag{6.2}$$

Proof. This result parallels the proof of Theorem 5.2. We only sketch the details, indicating the main changes needed. Suppose $[k:\mathbb{Q}] = d$.

Firstly, we have

$$\#\{f(x) \in \mathcal{F}_n(B;\Omega)\} = (2B)^{nd}.$$

The condition for the polynomial discriminant $\gcd\left(Disc(f), \prod_{i=1}^{r} \mathfrak{p}_i\right) = (1)$ is exactly that the polynomial $f(x)$ has square-free factorization (mod $\mathfrak{p}_i$) for $1 \le i \le r$. Set $M = \prod_{i=1}^{r} q_i = \prod_{i=1}^{r} (p_i)^{f_i}$. For the limit in (1) we obtain an exact count when going through boxes having all sides $B = Mm$ for some integer $m \ge 1$, which is

$$|\mathcal{F}_n(B;\mathcal{S},\Omega)| = (2B)^{nd} \prod_{i=1}^{r} \left(1 - \frac{1}{N\mathfrak{p}_i}\right) = (2B)^{nd} \prod_{i=1}^{r} \left(1 - \frac{1}{q_i}\right)$$

For each prime ideal $\mathfrak{p}_i$ this holds using Proposition 4.5 (1) since we have an integral multiple of complete residue systems (mod $\mathfrak{p}_i$) in the box, and it holds for all $\mathfrak{p}_i$ simultaneously using the Chinese remainder theorem for the box. Allowing a general B adjusts this formula by a multiplicative amount $1 + O(\frac{ndM}{B})$, and letting $B \to \infty$ yields (6.1).

Secondly, we introduce $\mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)^+$ to be those elements of $\mathcal{F}_n(B;\mathcal{S},\Omega)$ that satisfy condition (i) only. We then have a bound for the number of these $f(x)$ that do not give S_n-extensions of k, which is

$$0 \le \mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)^+ - \mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega) \le \alpha_n(k)(2B)^{nd} \frac{d\log B}{\sqrt{B^d}}.$$

This result follows using an upper bound of Cohen ([11], Theorem 2.1), in his result specifying that $F_{\mathbf{t}}(x) = X^n + \sum_{i=0}^{n-1} \mathbf{t}_i X^i$, that $K = k$, and noting the Galois group $G = S_n$ for $F_{\mathbf{t}}(X)$ over the function field $k(\mathbf{t}_1, \cdots, \mathbf{t}_n)$.

Thirdly, on restricting the box size to the special form $B = Mm$ with $m \ge 1$, one gets an exact count

$$|\mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)| = (2B)^{nd} \prod_{i=1}^{r} \frac{1}{(N\mathfrak{p}_i)^n} N_{\mu_i}(N\mathfrak{p}_i).$$

This formula is equivalent to

$$|\mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)| = (2B)^{nd} \prod_{i=1}^{r} \left(1 - \frac{1}{q_i}\right) \nu^*_{n,q_i}(C_{\mu_i}).$$

Changing the box size to an arbitrary integer B introduces at most a multiplicative roundoff error of $1 + O\left(\frac{ndM}{B}\right)$.

Fourthly, we combine the above estimates to obtain an analogue of Theorem 5.4, stating that

$$\left| \frac{|\mathcal{F}_n(B;\{\mathcal{S},\mathfrak{U}\},\Omega)|}{|\mathcal{F}_n(B;\mathcal{S},\Omega)|} - \prod_{i=1}^{r} \nu^*_{n,q_i}(C_{\mu_i}) \right| \le 2 \prod_{i=1}^{r} \left(1 - \frac{1}{q_i}\right)^{-1} \alpha_n(k) \frac{d\log B}{\sqrt{B^d}} + \frac{4ndM}{B}.$$

The formula (6.2) follows on letting $B \to \infty$. □

Remark. The conclusion in Theorem 6.1 is insensitive to the shape of the box bounding the coefficients as long as it is increased homothetically as $B \to \infty$, e.g. $-c_{n,j}B < a_j < c_{n,j}B$, where $c_{n,j}$ are positive constants independent of B. □

We next obtain a result parallel to Theorem 2.5 on the existence of relative S_n-number fields K over k having prescribed splitting above a given finite set of prime ideals $\mathcal{S} = \{\mathfrak{p}_i : 1 \le i \le r\}$, and setting $N_{k/\mathbb{Q}}\mathfrak{p}_i = (q_i)$, provided that all the quantities $\nu^*_{n,q_i}(C_{\mu_i}) > 0$. We follow the convention that a *relative S_n-number field K over k* is a degree n extension of k whose Galois closure over k has Galois group S_n. We recall that the *(relative) discriminant* $Disc(O_K \mid O_k)$ of any order O of K that contains O_k is that ideal of O_k that is generated by the discriminants $(\alpha_1, \dots, \alpha_n)$ of all the bases of K/k which are contained in O. ([38], III (2.8)). The prime ideal powers dividing the relative discriminant can be computed locally ([37], Prop. 5.7, p. 219).

Theorem 6.2. *Let $k/\mathbb{Q}$ be a number field, not necessarily Galois over $\mathbb{Q}$. Let $\mathcal{S} = \{\mathfrak{p}_1, \dots, \mathfrak{p}_r\}$ denote a finite set of prime ideals of k. and let $\mathfrak{U} = \{\mu_1, \dots, \mu_r\}$ with $\mu_j \vdash n$ be a prescribed set of splitting symbols for these prime ideals. Set $q_i = N_{k/\mathbb{Q}}\mathfrak{p}_i$. Then the following conditions are equivalent.*

(1) *The positive measure condition*

$$\nu^*_{n,q_i}(C_{\mu_i}) > 0 \text{ for } 1 \le i \le r$$

holds.

(2) *There exists a relative S_n-number field K/k having the following two properties:*

- *(P1-k)* *The field K contains a monogenic order $O = O_k[1, \theta, \dots \theta^{n-1}]$ whose relative discriminant $Disc(O|O_k)$ is relatively prime to $\mathfrak{p}_1\mathfrak{p}_2 \cdots \mathfrak{p}_r$.*
- *(P2-k)* *The Galois closure K^{spl} of K over k is unramified at all prime ideals above those in $\mathcal{S}$ and the primes in $\mathcal{S}$ have prescribed Artin symbols*

$$\left[\frac{K^{spl}/k}{(\mathfrak{p}_i)}\right] = C_{\mu_i}, \qquad 1 \le i \le r.$$

(3) *There exist infinitely many relative S_n-number fields K over k having properties* (P1-*k*) *and (P2-k).*

Proof. The proof parallels that of Theorem 2.5, using Theorem 6.1 in place of Theorem 5.2. For (1) $\Rightarrow$ (3) we use the fact that for a monic polynomial $f(x) \in O_k[x]$ that is irreducible over O_k one has the equality of polynomial discriminants and relative discriminants of the associated monogenic order in $K = k(\theta)$, for θ a root of $f(x)$. That is, setting $O_f := O_k[1, \theta, \cdots, \theta^{n-1}]$.

one has the equality

$$(Disc(f))O_k = Disc[O_f \mid O_k]. \tag{6.3}$$

of O_k-ideals; here $(Disc(f))$ is a principal ideal. We use this fact to show that (P1-k) is satisfied, and apply Theorem 6.1 to show (P2-k) is satisfied.

For (2) $\Rightarrow$ (1) the hypothesis (P1-k) with the identity (6.3) implies $\mathfrak{p}_i \nmid (Disc(f))$ as an O_k-ideal and the square-free factorization of $f(x) \pmod{\mathfrak{p}_i}$ for each of the $\mathfrak{p}_i$. This fact

gives the required Artin symbols C_{μ_i}, and positive density follows by Theorem 6.1 since all conditions imposed are congruence conditions. □

To conclude the paper we formulate a generalization of Theorem 2.6. For a relative extension K/k of degree n we say that a prime ideal $\mathfrak{p}$ of O_k is called an *essential relative discriminant divisor* if it divides the relative discriminants $Disc(O|O_k)$ all monogenic orders $O := O_k[1, \theta, \cdots, \theta^{n-1}]$ of the field K over k.

Theorem 6.3. *Let a number field k together with a prime ideal $\mathfrak{p}$ be given. Let $\mathfrak{p}$ have ideal norm $N_{k/\mathbb{Q}}(\mathfrak{p}) = (q) = (p^k)$. For a set of splitting types $\mu \vdash n$, with $n \geq 2$, the following three conditions are equivalent.*

(C1-k) *The splitting measure at $z = q = p^k$ has*
$$\nu^*_{n,q}(C_\mu) = 0.$$

(C2-k) *For all degree n monic integer polynomials $f(x)$ with coefficients in O_k whose* $(\mathrm{mod}\ \mathfrak{p})$ *factorization has splitting type μ, the relative discriminant $Disc(O_f|O_k)$ is divisible by $\mathfrak{p}$.*

(C3-k) *All relative S_n-extensions K of k in which $\mathfrak{p}$ is unramified and has splitting type μ necessarily have $\mathfrak{p}$ as an essential relative discriminant divisor.*

Proof. The proof parallels that of Theorem 2.6. We note only that to establish the equivalence (C1-k) ⇔ (C2-k), one uses (6.3). □

In cases where (C1-k) holds this proof does not establish that there exist any fields satisfying (C3-k).

7 Generalizations

7.1 Characteristic polynomials of random integer matrices

The problem studied in this paper can be viewed as a special case of study of characteristic polynomials of random matrices. One may consider random matrices drawn from a group like $GL(n, \mathbb{Z})$ with constraints on the size of the matrix $\mathbf{A} = [a_{i,j}]$ (measured in some matrix norm), and also putting side conditions on the allowed elements. The problem for degree n polynomials above can be encoded as such random $n \times n$ matrices (having entries $|a_{i,j}| \leq B$) by mapping the polynomial $f(x)$ to the companion matrix having characteristic polynomial $f(x)$. After reduction $(\mathrm{mod}\, p)$ from $GL(n, \mathbb{Z})$ one obtains a particular distribution of random matrices having entries over the finite field $\mathbb{F}_p$ with a side condition forcing many matrix entries to be zero. Our imposed restriction on factorization of polynomials being squarefree corresponds requiring that the associated matrices in $GL(n; \mathbb{F}_p)$ have distinct eigenvalues, i.e. they belong to semisimple conjugacy classes. One can ask whether there are further interesting generalizations of the model of this paper results in the random matrix context.

There are many results known considering random integer matrices in more general models. In 2008 Kowalski ([27], Chap. 7) showed that the characteristic polynomial of a random matrix in $SL(n, \mathbb{Z})$ drawn using a random walk is an S_n-polynomial with probability approaching 1 as the number of steps increases. For splitting fields of characteristic polynomials of random elements drawn from more general split reductive arithmetic

groups G see work of Gorodnik and Nevo [22], Jouve, Kowalski and Zywina [25]. In their framework the Galois group S_n is replaced by the Weyl group $W(G)$ of the underlying algebraic group G; the case $W(G) = S_n$ corresponds to $G = SL_n$. Lubotzky and Rosenzweig [32] give a further generalization to a wider class of groups with coefficients in a wider class of fields, where the "generic" Galois group of a random element may have a more complicated behavior.

There are also many results known on the distribution of characteristic polynomials of random matrices over finite fields $\mathbb{F}_q$; this subject is surveyed in Fulman [18]. His paper puts emphasis on $Mat(n, \mathbb{F}_q)$ and $GL(n, \mathbb{F}_q)$, and includes results on factorization type of characteristic polynomials (see also [19]). Example 2 in ([18], Section 2.2) observes that the factorization type for a uniformly drawn matrix in $Mat(n, \mathbb{F}_q)$ has a distribution depending on n and q that approaches that of a random degree n monic polynomial in $\mathbb{F}_q[X]$ as $q \to \infty$. Fulman ([18], Section 3.1) also introduces a family of probability measures $M_{GL,u,q}$ on conjugacy classes of $GL(n, \mathbb{F}_q)$, which when conditioned on fixed n do not depend on the parameter u and have the rational function interpolation property in the parameter q. They therefore extend to a complex parameter z, defining complex-valued measures $M_{GL,z,q}$. He remarks ([18], Section 3.3) that this distribution coincides with the distribution on partitions describing the Jordan block structure of a random unipotent element of $GL(n, q)$. It would be interesting to determine whether the measures $M_{GL,u,q}$ have any relation to the splitting measures studied in this paper.

7.2 Square-free polynomials and homological stability

The splitting measures $\nu^*_{n,q}(C_\mu)$ count the relative fraction of monic square-free polynomials (mod p) that have a given factorization type in $\mathbb{F}_q[x]$. Recently, as a special case of a general theory, Church, Ellenberg and Farb [6] observed that the monic square-free polynomials in $\mathbb{F}_q[x]$ for $q = p^k$ label points in an interesting moduli space $Y_n(\mathbb{F}_q)$ defined over $\mathbb{F}_q$, the complement of the discriminant locus, which carries an S_n-action. They relate point counts on the space $Y_n(\mathbb{F}_q)$ specified by factorizations of square-free polynomials in $\mathbb{F}_q[x]$ to the topology of the *configuration space*

$$X_n(\mathbb{C}) = PConf_n(\mathbb{C}) := \{(z_1, z_2, \cdots, z_n) : z_i \in \mathbb{C}, z_i \neq z_j\},$$

which itself carries an S_n-action. The configuration space $\mathrm{PConf}_n(\mathbb{C})$ is an affine variety which is the complement of a set of hyperplanes. (It is a special case of a discriminant variety, see Lehrer [31].) Church, Ellenberg and Farb study the S_n-representations produced by the S_n-action on the homology of this space and show certain homological stability properties of these representations hold as $n \to \infty$. They then study limiting behaviors of polynomial statistics of these points attached to a fixed multivariate polynomial $P(x_1, \ldots, x_m) \in \mathbb{Q}[x_1, \ldots, x_m]$ and relate this behavior to homological stability.

The statistics they study over $Y_n(\mathbb{F}_q)$ can be expressed in terms of the q-splitting measures $\nu^*_{n,q}(\cdot)$, which may permit an alternative way to view some of their results. We hope to consider this topic further elsewhere.

For general results on homological stability properties under S_n-actions see Church et al. [4,5,7].

Endnote

[a] Gauss found this formula on August 25, 1797, according to his *Tagebuch*, see Frei [17].

Acknowledgments
The authors thank Dani Neftin for remarks on characteristic polynomials of random matrices and for bringing relevant references to our attention. They thank the reviewer for helpful comments. Some of the work of the second author was done at the University of Michigan and at the Technion, whom he thanks for support.
The first author was partially supported by NSF grants DMS-1101373 and DMS-1401224.

Author details
[1]Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, USA. [2]Department of Mathematics, University of Maine, Orono, ME 04469, USA.

Received: 13 January 2015 Accepted: 24 January 2015
Published online: 09 July 2015

References

1. Bhargava, M: The density of discriminants of quartic rings and fields. Ann. Math. **162**(2), 1031–1063 (2005)
2. Bhargava, M: Mass formulae for extensions of local fields, and conjectures on the density of number field discriminants. Int. Math. Res. Not. IMRN. **2007**(17). 20. Art. ID rnm052
3. Bhargava, M: The density of discriminants of quintic rings and fields. Ann. Math. **172**(3), 1559–1591 (2010)
4. Church, T: Homological stability for configuration spaces of manifolds. Inventiones Math. **188**(2), 465–504 (2012)
5. Church, T, Ellenberg, JS, Farb, B: FI-modules and stability for representations of symmetric groups. Duke Math. J. to appear. (arXiv:1204.4533)
6. Church, T, Ellenberg, JS, Farb, B: Representation stability in cohomology and asymptotics for families of varieties over finite fields. Algebraic topology: applications and new directions, 1–54. Contemp. Math. 620, Amer. Math. Soc., Providence, RI, 2014
7. Church, T, Farb, B: Representation theory and homological stability. Adv. Math. **245**, 250–314 (2013)
8. Cohen, H: A Course in Computational Algebraic Number Theory. Springer-Verlag, New York (1993)
9. Cohen, H, Diaz y Diaz, F, Olivier, M: Counting discriminants of number fields. J. Théor. Nombres Bordeaux. **18**(3), 573–593 (2006)
10. Cohen, SD: The distribution of polynomials over finite fields. Acta Arithmetica. **17**, 255–271 (1970)
11. Cohen, SD: The distribution of Galois groups and Hilbert's irreducibility theorem. Proc. London Math. Soc., Series 3. 43, 227–250 (1981)
12. Cohn, H: A Classical Invitation to Algebraic Numbers and Class Fields. Springer-Verlag, New York (1978)
13. Davenport, H, Heilbronn, H: On the density of discriminants of cubic fields II. Proc. Roy. Soc. London Ser. A. **322**(1551), 405–420 (1971)
14. Dedekind, R: Über Zusammenhang zwischen der Theorie der Ideale und der Theorie der höhere Kongruenzen, Abh. König. Ges. der Wissen. zu Göttingen. **23**, 1–23 (1878)
15. Dietmann, RLProbabilisticGaloistheory. Bull. London Math. Soc. **45**(3), 453–462 (2013)
16. Evertse, J-H: A survey on monogenic orders. Publ. Math. Debrecen. **79**(3-4), 411–422 (2011)
17. Frei, G: The Unpublished Section Eight: On the Way to Function Fields over a Finite Field, Sect. II.4. In: Goldstein, C, Shappacher, N, Schwermer, J (eds.) The Shaping of Arithmetic after C. F. Gauss' Disquisitiones Arithmeticae, pp. 159–198. Springer, New York, (2007)
18. Fulman, J: Random matrix theory over finite fields. Bull. Am. Math. Soc. (N.S.) **39**(1), 51–85 (2002)
19. Fulman, J, Neumann, PM, Praeger, CE: A generating function approach to the enumeration of matrices in the classical groups over finite fields. Memoirs Amer. Math. Soc. **176**(830), 90 (2005)
20. Gallagher, PX: The large sieve and probabilistic Galois theory. Analytic number theory Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972, pp. 91–101. Amer. Math. Soc. Providence, RI, (1973)
21. Gauss, CF: Werke, Vol. II Höhere Arithmetik, ed. Königliche Geschellshaft der Wissenschaften zu Göttingen. Universitäts-Druckerei, Göttingen. 2nd Ed. (1876)
22. Gorodnik, A, Nevo, A: Splitting fields of elements in arithmetic groups. Math. Res. Lett. **18**(6), 1281–1288 (2011)
23. Hazewinkel, M: Witt vectors. I. Handbook of Algebra, Vol. 6, pp. 319–472. Elsevier/North-Holland, Amsterdam, (2009)
24. Hall, M: A basis for free Lie rings and higher commutators in free groups. Proc. Amer. Math. Soc. **1**, 575–581 (1951)
25. Jouve, F, Kowalski, E, Zywina, D: Splitting fields of characteristic polynomials of random elements in arithmetic groups, Israel. J. Math. **193**(1), 263–307 (2013)
26. Kedlaya, KS: Mass formulas for local Galois representations. With an appendix by Daniel Gulotta. Int. Math. Res. Not. IMRN. **2007**(17), 26. Art. ID rnm021 (2007)
27. Kowalski, E: The large sieve and its applications. Arithmetic geometry, random walks and discrete groups. Cambridge Tracts in Math., 175. Cambridge Univ. Press, Cambridge (2008)
28. Lagarias, JC: A family of measures on symmetric groups and the field with one element, arXiv:1505.05745, v1
29. Lagarias, JC, Odlyzko, AM: Effective versions of the Chebotarev density theorem. In: Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), pp. 409–464. Academic Press, London, (1977)
30. Lang, S: Algebraic Number Theory, Second Edition, Vol. 110. Springer, New York (1994)
31. Lehrer, GI: Rational points and cohomology of discriminant varieties, Adv. Math. **186**(1), 229–250 (2004)
32. Lubotzky, A, Rosenzweig, L: The Galois group of random elements of linear groups. Amer. J. Math. **136**(5), 1347–1383 (2014)
33. Macdonald, IG: Symmetric Functions and Hall Polynomials. Second Edition. Oxford University Press, Oxford (1995)
34. Maser, H: Carl Friedrich Gauss' Untersuchungen der höhere arithmetik. Julius Springer, Berlin (1889)
35. Metropolis, N, Rota, G-C: Witt vectors and the algebra of necklaces. Adv. Math. **50**, 95–125 (1983)
36. Moreau, C: Sur les permutations circulaires distinctes, Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale. Sér. 2. **11**, 309–314 (1872)
37. Narkiewicz, W: Elemantary and Analytic Theory of Algebraic Numbers. Second Edition. Springer-Verlag, Berlin (1990)

38. Neukirch, J: Algebraic number theory, volume 322 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin (1999). Translated by Norbert Schappacher
39. Rosen, M: Number theory in function fields, volume 210 of Graduate Texts in Mathematics. Springer-Verlag, New York (2002)
40. Śliwa, J: On the nonessential discriminant divisor of an algebraic number field. Acta Arithmetica. **42**(1), 57–72 (1982)
41. Stanley, RP: Enumerative combinatorics. Vol. 1, Second Edition. volume 49 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (2012)
42. Tornheim, L: Minimal basis and inessential discriminant divisors, Pacific. J. Math. **5**, 621–631 (1955)
43. Venkatesh, A, Ellenberg, JS: Statistics of Number Fields and Function Fields. In: Proceedings of the ICM, Hyderabad, India 2010 (Volume II), pp. 383–402. Hindustan Book Agency, New Delhi, (2010)
44. van der Waerden, BL: Die Seltenheit der reduziblen Gleichungen und der Gleichungen mit Affekt. Monatsh. Math. Phys. **43**(1), 133–147 (1936)
45. Weiss, BL: Probabilistic Galois theory over p-adic fields. J. Number Theory. **133**, 1537–1563 (2013)
46. Witt, E: Treue Darstellung Lieschen Ringe. J. reine Angew. Math. 177, 152–160 (1937)
47. Wood, MM: Mass formulas for local Galois representations of wreath products and cross products. Algebra and Number Theory. **2**(4), 391–405 (2008)
48. Wood, MM: On the probabilities of local behaviors in abelian field extensions. Compositio Math. **146**(1), 102–128 (2010)
49. Zywina, D: Hilbert's Irreducibility Theorem and the Larger Sieve. eprint: arXiv:1011.6465 v1