# RESEARCH STATEMENT: YIFENG HUANG

My primary interests are algebraic geometry and number theory, and I am interested in various questions on the interface of arithmetic statistics, combinatorics and algebraic topology. I have studied problems involving moduli spaces, Diophantine equations, random matrices, elliptic curves, singular curves, zeta functions, partitions and $q$-series. Among other problems, I have initiated the investigation of "matrix points on varieties", which amounts to the study of certain matrix equations naturally arising from algebraic geometry. This topic has turned out to have rich connections with a broad territory of mathematics, ranging from modular forms, hypergeometric functions to knot theory and mathematical physics. I will explain matrix points in context and some of my results along this line, among my other works. Later, I will offer a glimpse of my potential future work.

## 1. Counting matrix points on varieties

For millenia, mathematicians aspire to understand polynomial equations and their solutions. An important aspect of this immortal subject is counting solutions in finite fields. The investigation of it has led to a number of deep findings in the last century. One celebrated example is the Sato–Tate conjecture in 1960s, which concerns $\mathbb{F}_p$-point counts on elliptic curves and their distribution as the prime $p$ varies. Half a century later, it was famously proved by a series of works [5, 19, 27, 47] which helped Richard Taylor win the Breakthrough Prize in Mathematics in 2015.

I am interested in the matrix solutions of polynomial equations. A natural question in this aspect is to count tuples of commuting $n \times n$-matrices over finite fields that satisfy a polynomial equation. For example, over a finite field $\mathbb{F}_q$ with $q = p^r$ elements, consider the elliptic curve defined as

$$E_\lambda : y^2 = x(x-1)(x-\lambda), \qquad \lambda \in \mathbb{F}_q \setminus \{0,1\}. \tag{1.1}$$

Then the set of $n \times n$-matrix points on $E_\lambda$ is defined as

$$E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q)) := \{(A,B) \in C_{n,2}(\mathbb{F}_q) : B^2 = A(A - I_n)(A - \lambda I_n)\}, \tag{1.2}$$

where $C_{n,2}(\mathbb{F}_q)$ denotes the set of commuting pairs of $n \times n$-matrices over $\mathbb{F}_q$. It is natural to ask whether there is an exact formula for $\#E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$ in terms of $\lambda$, and how $\#E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$ distributes as $\lambda$ varies. More generally, what can we say about matrix point counts for other varieties? What role do the ordinary point counts play in answering the above question?

It turns out that all the questions above have interesting and promising answers. In joint work with Ono and Saad [35], we found a formula for $\#E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$ in terms of certain combinatorial polynomials depending on $n$, and certain arithmetic functions [26] depending on $\lambda$. Moreover, we proved a matrix analog for the Sato–Tate distribution concerning the ordinary point counts $\#E_\lambda(\mathbb{F}_q)$. To be specific, for any $n \geq 1$, let $a_n^*(\lambda; q)$ denote a suitable normalization of $\#E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$, then we have the following result.

**Theorem 1.1** (H.–Ono–Saad [35])**.** *The histogram of $\{a_n^*(\lambda; q) : \lambda \in \mathbb{F}_q\}$ approaches the semicircular distribution that has famously appeared in the Sato–Tate conjecture. More precisely, for $-2 \leq a < b \leq 2$ and $n, r \geq 1$, we have*

$$\lim_{p \to \infty} \frac{\#\{\lambda \in \mathbb{F}_{p^r} : a_n^*(\lambda; p^r) \in [a,b]\}}{p^r} = \frac{1}{2\pi} \int_a^b \sqrt{4 - t^2}\, dt. \tag{1.3}$$

Our proof involves a combinatorial analysis of our exact formula for $\#E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$, drawing from distributional results about the ordinary point counts $\#E_\lambda(\mathbb{F}_q)$ by Ono, Saad and Saikia [44].

I now move on to explain my results about matrix points on other varieties, with the guiding questions in mind.

1.1. **Smooth curves and surfaces in general.** Before continuing the discussion, I now define matrix points in greater generality. Let $X$ be an affine variety over a field $k$ defined by the polynomial equation $f_1(T_1, \dots, T_m) = \cdots = f_r(T_1, \dots, T_m) = 0$, and let $C_{n,m}(k)$ denote the set of $m$-tuples of pairwise commuting $n \times n$-matrices over $k$. Then the set of $n \times n$-**matrix points** on $X$ is defined as

$$X(\mathrm{Mat}_n(k)) := \{(A_1, \dots, A_m) \in C_{n,m}(k) : [A_i, A_j] = 0, f_i(A_1, \dots, A_m) = 0\}. \tag{1.4}$$

For example, when $X$ is the affine plane, $X(\mathrm{Mat}_n(\mathbb{F}_q))$ corresponds to the set of commuting pairs of $n \times n$-matrices over $\mathbb{F}_q$. Feit and Fine established a remarkable theorem [24] in 1960 that provides the count of these matrix points. Their formula is combinatorial and involves partitions.

More generally, if $X$ is any smooth curve or smooth surface over $\mathbb{F}_q$, I gave explicit formulas [31] for $\#X(\mathrm{Mat}_n(\mathbb{F}_q))$ in terms of the local zeta function of $X$, a function that records the point counts of $X$ over finite extensions of $\mathbb{F}_q$.

As a consequence, $\mathrm{Mat}_n(\mathbb{F}_q)$-point counts on smooth curves and surfaces are determined by ordinary point counts over $\mathbb{F}_{q^r}$ for $1 \le r \le n$. To prove these formulas, I first establish a local-to-global formula inspired by a pivotal work of Cohen and Lenstra [21], and combine it with Feit–Fine's theorem.

1.2. **A family of $K3$ surfaces.** Theorem 1.1 exemplifies that deep theorems about ordinary point counts can sometimes lead to theorems about $n \times n$-matrix point counts for $n > 1$. In the same vein, I proved analogous results for a family of $K3$ surfaces [1] with remarkable arithmetic properties, in joint work [35] with Ono and Saad. More precisely, consider the $K3$ surface with function field

$$X_\lambda : s^2 = xy(x+1)(y+1)(x+\lambda y), \qquad \lambda \in \mathbb{F}_q \setminus \{0,-1\} \tag{1.5}$$

and its set of matrix points

$$X_\lambda(\mathrm{Mat}_n(\mathbb{F}_q)) := \{(A,B,C) \in C_{n,3}(\mathbb{F}_q) : C^2 = AB(A+I_n)(B+I_n)(A+\lambda B), C \in \mathrm{GL}_n(\mathbb{F}_q)\}. \tag{1.6}$$

We found an exact formula for $\#X_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$ that is analogous but more complicated than our formula for $\#E_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$. Moreover, we proved that as $\lambda$ varies, the limiting distribution of a suitable normalization of $\#X_\lambda(\mathrm{Mat}_n(\mathbb{F}_q))$ has a density function featuring two verticle asymptotes resembling "Batman's ears". This distribution was previously identified by Ono, Saad and Saikia [44] as the limiting distribution of suitably normalized $\#X_\lambda(\mathbb{F}_q)$, and it is very different from the Sato–Tate distribution.

1.3. **Singular curves.** So far, we have only discussed matrix points on smooth curves or smooth surfaces. Further generalization from this point turns out to be quite challenging. As the first successful attempt [31], I explicitly determined $\#X(\mathrm{Mat}_n(\mathbb{F}_q))$ whenever $X/\mathbb{F}_q$ is a curve whose singularities are nodes. More precisely, I determined the following generating function for $\#X(\mathrm{Mat}_n(\mathbb{F}_q))$ featuring a standard normalizing factor $\#\,\mathrm{GL}_n(\mathbb{F}_q)$:

$$\widehat{Z}_X(t) := 1 + \sum_{n=1}^\infty \frac{\#X(\mathrm{Mat}_n(\mathbb{F}_q))}{\#\,\mathrm{GL}_n(\mathbb{F}_q)} t^n. \tag{1.7}$$

To state the result, let $X/\mathbb{F}_q$ be a reduced curve with exactly $r$ singular points. Assume every singular point is the $\{xy=0\}$ singularity, namely, a point on $X$ at which the completed local ring is isomorphic to $\mathbb{F}_q[[x,y]]/(xy)$. Let $\widetilde{X} \to X$ be the desingularization of $X$. Then $\widehat{Z}_X(t)$ is given in terms of the local zeta function $Z_{\widetilde{X}}(t)$ of $\widetilde{X}$ and a certain basic hypergeometric function, which we expand explicitly using the $q$-Pochhammer symbol $(a;q)_n := \prod_{i=0}^{n-1}(1-aq^i)$, $(a;q)_\infty := \prod_{i=0}^\infty (1-aq^i)$.

**Theorem 1.2** (H. [31])**.** *Assume the setting above. Then we have*

$$\widehat{Z}_X(t) = H(t;q^{-1})^r \prod_{j=1}^\infty Z_{\widetilde{X}}(tq^{-j}), \tag{1.8}$$

*where*

$$H(t;q) := (tq;q)_\infty \cdot {}_0\phi_1(-;tq;q,t^2q) = \sum_{k=0}^\infty \frac{t^{2k}q^{k^2}}{(q;q)_k}(tq^{k+1};q)_\infty. \tag{1.9}$$

The proof involves counting matrix points on the "union of axes" $\{xy=0\}$, which amounts to counting pairs of matrices $A, B \in \mathrm{Mat}_n(\mathbb{F}_q)$ with $AB = BA = 0$. The expression of $H(t;q)$ is then derived from these counts using partition identities.

The availability of an explicit formula is striking. By analyzing the series $H(t;q)$, Theorem 1.2 immediately leads to several surprising consequences. I will list two. First, we have the following analytic phenomenon that results from the fact that $H(t;q)$ converges for all $t \in \mathbb{C}$ and $|q| < 1$.

**Corollary 1.3** (H. [31])**.** *Assume the setting above. Then $\widehat{Z}_X(t)$ has a meromorphic continuation to all of $\mathbb{C}$. Moreover, as meromorphic functions on $\mathbb{C}$, $\widehat{Z}_X(t)$ and $\widehat{Z}_{\widetilde{X}}(t)$ have the same set of poles, counting multiplicities.*

Second, investigating the special values $H(\pm 1;q)$ reveals exact formulas about module statistics over $R := \mathbb{F}_q[[x,y]]/(xy)$. These formulas are module analogues of famous "partition statistics = modular form" identities, such as a number of Euler's partition identities and the celebrated Rogers–Ramanujan identities. In our case, the existence of such analogues is surprising because a combinatorial classification of modules over the singular ring $R$ is out of reach.

As is standard, the dimension of an $R$-module $M$ refers to $\dim_{\mathbb{F}_q} M$, and we count each $R$-module $M$ with weight $1/\#\mathrm{Aut}\, M$, where $\mathrm{Aut}\, M$ is the automorphism group of $M$ as an $R$-module.

**Corollary 1.4** (H. [31]). *Let $m(q), m_e(q)$ and $m_o(q)$ denote the total weighted count of isomorphism classes of finite-dimensional, even-dimensional and odd-dimensional $\mathbb{F}_q[[x,y]]/(xy)$-modules, respectively. Then*

$$m(q) = (q^{-1}; q^{-1})_\infty^{-2}, \text{ and } m_e(q) - m_o(q) = (-q^{-2}; q^{-2})_\infty^{-1}. \tag{1.10}$$

It is natural to ask whether these (and potentially more) consequences of Theorem 1.2 generalize to other curve singularities, but it would not take one too long to realize that any such speculation is quite challenging in the absence of an exact formula. The following conjecture, however, has shown promising signs.

**Conjecture 1.5** (H. [31]). Let $\widetilde{X} \to X$ be the desingularization of *any* singular curve $X/\mathbb{F}_q$. Then the conclusion of Corollary 1.3 still holds.

The conjectured phenomenon is an instance of approximation by desingularization, a recurring theme in algebraic geometry. My joint work [34] with Jiang gave positive evidences for the conjecture in the case where $X$ has only the cusp singularity $\{y^2 = x^3\}$, and our current work [33] proves the conjecture if $X$ has only singularities of the form $\{y^2 = x^n\}$ with $n \geq 2$.

## 2. Enumerative questions in moduli spaces

The research field of enumerative geometry aims to count "how many" of certain geometric configurations are there. The general approach to this goal is to consider the moduli space parametrizing all possible configurations. Then the original question can often be answered by determining a certain measure of the moduli space. For example, one may ask how many parameters are needed to specify a smooth projective curve over $\mathbb{C}$ of genus $g > 1$. It is answered by the fact that the moduli of curves $\mathcal{M}_g$ has dimension $3g - 3$, which is discovered by Riemann. For another example, one may ask how many 2-dimensional subspaces of $(\mathbb{F}_q)^4$ are there. It is answered by the fact that the Grassmannian $\mathrm{Gr}(2,4)$ has $q^4 + q^3 + 2q^2 + q + 1$ points over $\mathbb{F}_q$, which is a gateway example in the vast subject of Schubert calculus.

I am generally interested in counting "$n$-point configurations" on a variety, a question that can be made precise in many fundamentally different ways. I have studied several moduli spaces towards this question.

### 2.1. Configuration spaces.

A fundamental question in topology is to find the Betti numbers of the configuration space $\mathrm{Conf}_n(X)$, which parametrizes $n$ unordered *distinct* points on a manifold $X$. In 1969, Arnol'd [3] famously determined all the Betti numbers of $\mathrm{Conf}_n(\mathbb{C})$. In 2014, Church, Ellenberg and Farb [17] pointed out a remarkable numerical correspondence between the theorem of Arnol'd and the classical problem of counting square-free polynomials in $\mathbb{F}_q[t]$. To make it precise, we first note that $n$ unordered distinct points on $\mathbb{C}$ can be parametrized by (roots of) monic square-free polynomials of degree $n$. In this spirit, $\#\,\mathrm{Conf}_n(\mathbb{F}_q)$ is the number of monic square-free polynomials of degree $n$ in $\mathbb{F}_q[T]$. Church, Ellenberg and Farb pointed out that

$$\begin{array}{lll} n = 0,1, \quad \#\,\mathrm{Conf}_n(\mathbb{F}_q) = q^n & \longleftrightarrow & H^0(\mathrm{Conf}_n(\mathbb{C})) = \mathbb{Q}; \\ n \geq 2, \quad \#\,\mathrm{Conf}_n(\mathbb{F}_q) = q^n - q^{n-1} & \longleftrightarrow & H^0(\mathrm{Conf}_n(\mathbb{C})) = \mathbb{Q}, H^1(\mathrm{Conf}_n(\mathbb{C})) = \mathbb{Q}. \end{array}$$

They then gave a beautiful explanation of it using purity of the $\ell$-adic cohomology and the mixed Hodge structure.

The correspondence above relates the $\mathbb{C}$-points and the $\mathbb{F}_q$-points of $\mathrm{Conf}_n(X)$, where $X = \mathbb{A}^1$ is the affine line. It is natural to seek its analogues where $X$ is another variety. Any analogue, if found, would immediately lead to a clean formula for the Betti numbers of $\mathrm{Conf}_n(X)$, because the $\mathbb{F}_q$-point count has one thanks to the classical Macdonald's formula [40] and a theorem of Vakil and Wood [49].

For $g, r \geq 0$, let $\Sigma_{g,r}$ be an $r$-punctured genus-$g$ algebraic curve; note that $\mathbb{A}^1 = \Sigma_{0,1}$ fits into this family. The Betti numbers of $\mathrm{Conf}_n(\Sigma_{g,r})$ were explicitly computed by Drummond-Cole and Knudsen [22] in 2017. However, Church–Ellenberg–Farb type correspondences remained unclear for $\Sigma_{g,r}$ when $g \geq 1$.

In joint work with Cheong [12], I found a satisfying correspondence for the one-punctured elliptic curve $\Sigma_{1,1}$. More precisely, we proved that an "irregularly shifted" generating function for the Betti numbers of $\mathrm{Conf}_n(\Sigma_{1,1})$ coincides with an explicit rational function recording the $\mathbb{F}_q$-point count of $\mathrm{Conf}_n(\Sigma_{1,1})$.

**Theorem 2.1** (Cheong–H. [12]). *Let $h^i(\mathrm{Conf}_n(\Sigma_{1,1}))$ denote the $i$-th Betti number of $\mathrm{Conf}^n(\Sigma_{1,1})$. Then we have*

$$\sum_{n,i \geq 0} (-1)^i h^i(\mathrm{Conf}_n(\Sigma_{1,1})) u^{2n - w(i)} t^n = \frac{(1-ut)^2(1-u^2t^2)}{(1-u^2t)(1-ut^2)^2}, \text{ where } w(i) = \lfloor \frac{3i}{2} \rfloor. \tag{2.1}$$

The rational function at the right-hand side is the generating function for $\#\,\mathrm{Conf}_n(\Sigma_{1,1})(\mathbb{F}_q)$ with a change of variables. To establish Theorem 2.1, we proved a geometric statement involving Deligne's mixed Hodge theory, which naturally explains the irregular shift pattern $w(i)$.

**Theorem 2.2** (Cheong–H. [12]). *For $n, i \geq 1$, the mixed Hodge structure of the $i$-th cohomology of $\mathrm{Conf}_n(\Sigma_{1,1})$ is pure of weight $w(i)$.*

**Remark.** From Theorem 2.2, we also determined the mixed Hodge numbers of $\mathrm{Conf}_n(\Sigma_{1,1})$, refining Theorem 2.1. Moreover, we proved that their generating function is related to the $\mathbb{F}_q$-point count of $\mathrm{Conf}_n(\Sigma_{1,1})$ in a way analogous to Theorem 2.1.

So far, the mixed Hodge structure have played an essential role to connect the Betti numbers to $\mathbb{F}_q$-point counts. Towards the more general study of $\mathrm{Conf}_n(\Sigma_{g,r})$ beyond $g = r = 1$, a natural problem is to determine the mixed Hodge numbers of $\mathrm{Conf}_n(\Sigma_{g,r})$. However, we pointed out in [12, §1.2] an immediate challenge that the analogue of Theorem 2.2 no longer holds if $g, r \geq 1$, unless $g = r = 1$. Despite the absence of purity, in [29], I determined the mixed Hodge numbers of $\mathrm{Conf}_n(\Sigma_{1,r})$ for all $r \geq 1$ by making a *different* connection to the $\mathbb{F}_q$-point counts. More precisely, for every $r \geq 1$, the generating function for the mixed Hodge numbers of $\mathrm{Conf}_n(\Sigma_{g,r})$ has a simple expression in terms of its $r = 1$ case, which is known thanks to Theorem 2.2.

**Theorem 2.3** (H. [29]). *For any complex variety $X$, let $h^{p,q;i}(X)$ denote the $(p, q)$-th mixed Hodge number of the $i$-th cohomology of $X$. Then for $g \geq 0$ and $r \geq 1$, we have*

$$\sum_{p,q,i,n\geq 0} h^{p,q;i}(\mathrm{Conf}_n(\Sigma_{g,r}))x^p y^q (-u)^i t^n = \frac{1}{(1+xyut)^{r-1}} \sum_{p,q,i,n\geq 0} h^{p,q;i}(\mathrm{Conf}_n(\Sigma_{g,1}))x^p y^q (-u)^i t^n. \tag{2.2}$$

Up to a change of variables, the relation (2.2) is the same as the relation between the generating function for $\#\,\mathrm{Conf}_n(\Sigma_{g,r})(\mathbb{F}_q)$ and the generating function for $\#\,\mathrm{Conf}_n(\Sigma_{g,1})(\mathbb{F}_q)$.

Such relations are in fact ubiquitous, at least for Betti numbers. Analogoues of Theorem 2.3 have appeared since 1981 and were referred to as "splittings" [20, 25, 41, 42]. These results relate the Betti numbers of the configuration spaces of a topological surface and its punctured version. In 2008, Kallel [36] generalize some of these results to higher dimensions. It is then natural to ask if splittings occur in higher dimensions on the level of mixed Hodge numbers. I gave an affirmative answer.

**Theorem 2.4** (H. [29]). *Let $V$ be a smooth compact complex variety of dimension $d$, and let $V_r$ denote the $r$-punctured version of $V$. Then for $r \geq 1$, we have*

$$\sum_{p,q,i,n\geq 0} h^{p,q;i}(\mathrm{Conf}_n(V_r))x^p y^q (-u)^i t^n = \frac{1}{(1+x^d y^d u^{2d-1} t)^{r-1}} \sum_{p,q,i,n\geq 0} h^{p,q;i}(\mathrm{Conf}_n(V_1))x^p y^q (-u)^i t^n. \tag{2.3}$$

Theorem 2.4 also gives a refinement and alternative proof of Kallel's Betti number result. My proof relies on an equivariant strengthening of Theorem 2.4, which I established by examining the spectral sequence [48] for a certain hyperplane-like arrangement, and using mixed-Hodge purity of $V_1$.

2.2. **Hilbert and Quot schemes of points.** We now shift our attention to $n$-point configurations where the points can collide. A natural moduli space to consider is the symmetric product $\mathrm{Sym}_n(X)$ that parametrizes multisets of size $n$ on $X$. However, we also consider moduli spaces that keep track of extra data on the points of collision, and they are sometimes preferred for geometric reasons. For example, for an algebraic variety $X$, the Hilbert scheme (of points) $\mathrm{Hilb}_n(X)$ parametrizes 0-dimensional subschemes of $X$ of degree $n$. If $X$ is a smooth surface and $n \geq 2$, it is famously known that $\mathrm{Sym}_n(X)$ is not smooth and $\mathrm{Hilb}_n(X)$ is a resolution of singularities of $X$. In this and the next section, I will discuss several moduli spaces of this flavor and some of my results on their enumerative aspects.

A common theme about these moduli spaces is that their associated extra data at a collision point $p$ reflect the local geometry of $X$ at $p$. A rich source of local geometries is curve singularities, where concrete examples abound. It is thus natural to ask how the enumerative aspects of such a moduli space on a singular curve $X$ depends on the nature of singularities of $X$. This is the guiding question for all of the moduli spaces I will discuss in this section and the next. The enumerative invariant I focus on is the **motive** in the Grothendieck ring of varieties. Roughly speaking, the motive measures how a variety is built from simpler varieties (such as affine spaces) using locally closed decompositions. If a variety has a cell decomposition, then its motive is essentially a polynomial whose degree-$n$ coefficient counts the number of cells of dimension $n$. We denote the motive of a variety $X$ by $[X]$. The motive is a strong enumerative invariant that always determines the Euler characteristics over $\mathbb{C}$ and the point count over $\mathbb{F}_q$. In the presence of a cell decomposition, it also determines the Betti numbers and the mixed Hodge numbers.

Let us revisit $\mathrm{Hilb}_n(X)$, but with $X$ being a reduced singular curve over $\mathbb{C}$. When $X$ has only planar singularities, a recent conjecture of Oblomkov, Rassmussen and Shende [43] states an exact relation between the motive of $\mathrm{Hilb}_n(X)$ and knot-theoretic invariants of the associated links of singularities of $X$. In particular, the motive of $\mathrm{Hilb}_n(X)$ is expected to have a formula with regular combinatorial patterns, when the singularities of $X$ are planar. For a general reduced singular curve $X$, in the absence of an explicit formula, Bejleri, Ranganathan and Vakil [8] proved that the Hilbert zeta function

$$\sum_{n=0}^{\infty}[\mathrm{Hilb}_n(X)]\, t^n \tag{2.4}$$

is rational in $t$, with denominator governed by the desingularization of $X$.

Joint with Jiang [33, 34], I extended some of the general results as well as explicit formulas to an infinite family of natural generalizations of the Hilbert scheme, namely, the **Quot scheme of points**. To define it, we recall that $\mathrm{Hilb}_n(X)$ equivalently parametrizes quotient sheaves of $\mathcal{O}_X$ that are 0-dimensional and of degree $n$. For $d \geq 0$, replacing $\mathcal{O}_X$ by the vector bundle $\mathcal{O}_X^d = \mathcal{O}_X^{\oplus d}$ gives rise to the following definition

$$\mathrm{Quot}_{d,n}(X) = \mathrm{Quot}_{\mathcal{O}_X^d,n}(X) := \{\mathcal{O}_X^d \twoheadrightarrow M : \dim M = 0, \deg M = n\}. \tag{2.5}$$

When $X$ is a smooth curve, the motive of $\mathrm{Quot}_{d,n}(X)$ is known to Bifet [9]. The following result is a common generalization of theorems of Bifet and Bejleri–Ranganathan–Vakil.

**Theorem 2.5** (H.–Jiang [33]). *Let $X$ be a reduced curve over an algebraically closed field. Then for $d \geq 0$, the Quot zeta function*

$$\mathcal{Z}_{\mathcal{O}_X^d}(t) := \sum_{n=0}^{\infty}[\mathrm{Quot}_{d,n}(X)]\, t^n \tag{2.6}$$

*is rational in $t$. Moreover, if $\widetilde{X}$ is the desingularization of $X$, then $\mathcal{Z}_{\mathcal{O}_X^d}(t)/\mathcal{Z}_{\mathcal{O}_{\widetilde{X}}^d}(t)$ is a polynomial in $t$.*

Furthermore, we gave explicit formulas for the Quot zeta functions of reduced curves with certain planar singularities. To distinguish the contribution of the singularity from that of the global geometry, for a complete local ring $R$, we define the polynomial

$$\mathcal{NZ}_{R^d}(t) := \mathcal{Z}_{\mathcal{O}_X^d}(t)/\mathcal{Z}_{\mathcal{O}_{\widetilde{X}}^d}(t), \tag{2.7}$$

where $X$ is any reduced curve whose only singularity is $p$ and the completed local ring of $X$ at $p$ is isomorphic to $R$. We explicitly computed $\mathcal{NZ}_{R^d}(t)$ when $R = \mathbb{C}[[x,y]]/(y^2 - x^n)$ for $n \geq 2$. The general formulas involve summations over partitions. For simplicity, I only state the special case $n = 3$.

**Theorem 2.6** (H.–Jiang [33]). *Let $R = \mathbb{C}[[x,y]]/(y^2 - x^3)$ and $d \geq 0$. Then we have*

$$\mathcal{NZ}_{R^d}(t) = \sum_{r=0}^{d}[\mathrm{Gr}(r,d) \times \mathbb{A}^{dr}]\, t^{2r}. \tag{2.8}$$

Both Theorem 2.5 and Theorem 2.6 were conjectured in our previous work [34], where we proved Theorem 2.5 for curves with only the $\{y^2 = x^3\}$ singularity and arbitrary $d \geq 0$, and Theorem 2.6 for $d \leq 3$. The techniques involved showing an $n \to \infty$ stabilization for a Gröbner stratification using the combinatorial construction we introduced in [32].

Another motivation to consider the special cases for $\mathcal{NZ}_{R^d}(t)$ is to shed light on a possible compactification of $\mathrm{Bun}_{\mathrm{GL}_d}(X)$ (the moduli of vector bundles on $X$) that interacts well with the Serre duality, when $X$ is a projective plane curve. When $d = 1$, the compactified Jacobian provides a satisfactory compactification, which leads to a functional equation for the Hilbert zeta function $\mathcal{Z}_{\mathcal{O}_X}(t)$ (see for instance [45]). When $d > 1$, in the absence of analogous constructions, we proved the function equation for special singularities using our explicit formulas. To state it, define the Lefschetz motive $\mathbb{L} := [\mathbb{A}^1]$.

**Theorem 2.7** (H.–Jiang [33]). *Let $d \geq 0$ and $m \geq 1$. If $R = \mathbb{C}[[x,y]]/(y^2 - x^{2m+1})$, then $\mathcal{NZ}_{R^d}(t)$ satisfies the functional equation*

$$\mathcal{NZ}_{R^d}(t) = (\mathbb{L}^d t^2)^{dm} \mathcal{NZ}_{R^d}(\mathbb{L}^{-d} t^{-1}). \tag{2.9}$$

*On the other hand, if $R = \mathbb{C}[[x,y]]/(y^2 - x^{2m})$, then (2.9) for $\mathcal{NZ}_{R^d}(t)$ reduces to the assertion that an explicit identity involving Hall–Littlewood symmetric functions holds.*

2.3. **Moduli stack of finite-length coherent sheaves.** We have seen that $\mathrm{Hilb}_n(X)$ models a notion of $n$-point configurations using an coherent sheaf $M$ of dimension 0 and degree $n$, "framed" by a surjection $\mathcal{O}_X \to M$. We may drop the framing and consider $\mathrm{Coh}_n(X)$, which parametrizes length-$n$ coherent sheaves on $X$ up to isomorphism. While its definition seems simpler, $\mathrm{Coh}_n(X)$ actually has wilder geometry, in that $\mathrm{Coh}_n(X)$ is a stack, not a scheme. One can still make sense of its motive using [7]. The stack $\mathrm{Coh}_n(X)$ is closely related to mathematical physics. For example, when $X = \mathbb{C}^2$ is the affine plane over $\mathbb{C}$, the stack $\mathrm{Coh}_n(X)$ can be recognized as (a quotient of) a Nakajima quiver variety and its motive has played a crucial role in the study of the motivic Donaldson–Thomas theory of $\mathrm{Hilb}_n(\mathbb{C}^3)$ by Behrend, Bryan and Szendrői [6].

I now present some of my explicit results on $[\mathrm{Coh}_n(X)]$, where $X$ is a reduced curve with special singularities. They use different approaches. The first and classical approach is to recognize $\mathrm{Coh}_n(X)$ as the quotient stack $[C_n(X)/\mathrm{GL}_n]$, where $C_n(X)$ is the variety that parametrizes $n \times n$-matrix points on $X$. This yields the following. To state it, let $\mathcal{Z}_X(t)$ be the motivic zeta function of $X$ due to Kapranov [37], and recall the two-variable power series $H(t;q)$ from (1.9).

**Theorem 2.8** (H. [31])**.** *Let $X/\mathbb{C}$ be a reduced curve with exactly $r$ singularities, all of which being the $\{xy = 0\}$ singularity. Let $\widetilde{X}$ be the desingularization of $X$. Then we have*

$$\widehat{\mathcal{Z}}_X(t) := \sum_{n=0}^{\infty} [\mathrm{Coh}_n(X)]\, t^n = H(t;\mathbb{L}^{-1})^r \prod_{j=1}^{\infty} \mathcal{Z}_{\widetilde{X}}(\mathbb{L}^{-j}t) \tag{2.10}$$

*as power series in $\mathbb{L}^{-1}$ and $t$.*

The proof is exactly the same as Theorem 1.2.

The second and new approach is based on a formula that computes $\widehat{\mathcal{Z}}_X(t)$ in terms of $\mathcal{Z}_{\mathcal{O}_X^d}(t)$ for all $d \geq 0$, proved in my joint work [34] with Jiang. The high-level reason is that any $M \in \mathrm{Coh}_n(X)$ can be framed by $\mathcal{O}_X^d$ for sufficiently large $d$. Based on this, we explicitly computed $\widehat{\mathcal{Z}}_X(t)$ for reduced curves with only $\{y^2 = x^n\}$ singularities. Again, to single out the contribution of the singularity, for a complete local ring $R$, define

$$\mathcal{N}\widehat{\mathcal{Z}}_R(t) := \widehat{\mathcal{Z}}_X(t)/\widehat{\mathcal{Z}}_{\widetilde{X}}(t), \tag{2.11}$$

where $X$ is any reduced curve whose only singularity is $p$ and the completed local ring of $X$ at $p$ is isomorphic to $R$. I just state the case $n = 3$.

**Theorem 2.9** (H.–Jiang [33])**.** *Let $R = \mathbb{C}[[x,y]]/(y^2 - x^3)$. Then we have*

$$\mathcal{N}\widehat{\mathcal{Z}}_R(t) = \sum_{n=0}^{\infty} \frac{t^{2n}}{[\mathrm{GL}_n]} = \sum_{n=0}^{\infty} \frac{\mathbb{L}^{-n^2}}{(\mathbb{L}^{-1};\mathbb{L}^{-1})_n}\, t^{2n}. \tag{2.12}$$

The formula (2.12) was conjectured in our previous work [34], where we gave a proof conditional on Theorem 2.6, which was then only known for $d \leq 3$. When $t = \pm 1$, the right-hand side of (2.12) can be recognized as $1/((\mathbb{L}^{-1};\mathbb{L}^{-5})_\infty(\mathbb{L}^{-4};\mathbb{L}^{-5})_\infty)$ by the Rogers–Ramanujan identity.

Having access to explicit formulas for many special $\mathcal{N}\widehat{\mathcal{Z}}_R(t)$ unveils that the modularity phenomenon in Corollary 1.4 is not a coincidence. To motivate the results, we state a conjecture first.

**Conjecture 2.10.** Let $R/\mathbb{C}$ be the completed local ring of a curve singularity such that $\mathcal{N}\widehat{\mathcal{Z}}_R(t)$ is polynomial-count in the sense of Katz [28]. Say $N\widehat{Z}_R(t;q) \in \mathbb{Q}(q)[[t]]$ is the point-count version of $\mathcal{N}\widehat{\mathcal{Z}}_R(t)$ over $\mathbb{F}_q$. Then $N\widehat{Z}_R(\pm 1; q^{-1})$ converge for $|q| < 1$, and are weakly holomorphic modular forms of weight 0 in $\tau$, where $q = e^{2\pi i \tau}$.

**Theorem 2.11** (H.–Jiang [33])**.** *Let $n \geq 2$ and $R = \mathbb{C}[[x,y]]/(y^2 - x^n)$. If $n = 2m+1$, then*

$$N\widehat{Z}_R(\pm 1; q^{-1}) = \prod_{\substack{i=1 \\ i \not\equiv 0, \pm(m+1) \bmod 2m+3}}^{\infty} (1 - q^i)^{-1}. \tag{2.13}$$

*On the other hand, if $n = 2m$, then $N\widehat{Z}_R(1;q^{-1}) = 1$ and*

$$N\widehat{Z}_R(-1;q^{-1}) = \frac{(q^2;q^2)_\infty (q^{m+1};q^{m+1})_\infty^2}{(q;q)_\infty^2 (q^{2m+2};q^{2m+2})_\infty}, \tag{2.14}$$

*except that when $m \geq 2$, the truth of the last equality is conditional on the assertion that an explicit identity involving Hall–Littlewood symmetric functions holds.*

The explicit formulas above directly verify Conjecture 2.10 for $R = \mathbb{C}[[x,y]]/(y^2 - x^n)$ for $n \geq 2$.

## 3. Noncommutative Diophantine equations

A central question in number theory and arithmetic geometry is to find all integer or rational solutions of a polynomial equation. An important situation where the goal can be achieved (in finite time) is when one shows that there are only finitely many solutions, and moreover, their "heights" are bounded by an effectively computable number. We call it an effective finiteness theorem for this equation.

The unit equations have played an important role in the theory of Diophantine equations in general. A unit equation, in a very generic language, is an equation $x + y = 1$, where $x, y$ are in certain sets multiplicatively generated by finitely many elements. For example, we may require $x, y$ to be in certain finitely generated subgroups of $\mathbb{C}^\times$. A toy example is $2^n - 3^m = 1$, where $n, m \in \mathbb{Z}$. A celebrated theorem [38] of Lang in 1960 states that any unit equation on $\mathbb{C}^\times$ has only finitely many solutions. As an application of the unit equations to other Diophantine equations, Lang used this result to show that a curve over $\mathbb{Q}$ of genus $\geq 1$ cannot have infinitely many integral points. Since the birth of Baker's method [4], both finiteness theorems above have been known to be effective due to work of Evertse and Győry; see [23].

Since the finiteness theorem on unit equations is merely a statement that "addition and multiplication should not be compatible", nothing stops us from expecting analogous theorems in a noncommutative setting. However, my result below is the first finiteness theorem in this setting. Moreover, the finiteness is effective. To state it, let $\mathbb{H}$ be the quaternion algebra over $\mathbb{R}$ defined by $i^2 = j^2 = k^2 = -1, ij = k$. Let $\mathbb{H}_a$ be the algebra of algebraic quaternions, namely, $a + bi + cj + dk$ with $a, b, c, d \in \mathbb{R} \cap \overline{\mathbb{Q}}$. The norm of a quaternion $a + bi + cj + dk$ is $\sqrt{a^2 + b^2 + c^2 + d^2}$.

**Theorem 3.1** (H. [30]). *Let $\Gamma_1, \Gamma_2$ be semigroups of $\mathbb{H}_a^\times$ generated by finitely many elements of norms greater than 1, and fix $a, a', b, b' \in \mathbb{H}_a^\times$. If $\Gamma_1$ is commutative, then the equation*

$$afa' + bgb' = 1 \tag{3.1}$$

*has only finitely many solutions with $f \in \Gamma_1$ and $g \in \Gamma_2$. Moreover, there are effective bounds for the solutions $(f, g)$.*

Along the way, I proved that a commutative subsemigroup $\Gamma$ of $\mathbb{H}_a^\times$ generated by finitely many elements of norms $> 1$ cannot infinitely intersect an $(\mathbb{R}\cap\overline{\mathbb{Q}})$-hyperplane of $\mathbb{H}_a$ not passing through the origin. This is a noncommutative analogue of the Mordell–Lang theorem even though $\Gamma$ is commutative, because the real algebraic group $\mathbb{H}^\times$ is not. The Mordell–Lang theorem essentially says that in a semiabelian variety, such as $(\mathbb{C}^\times)^n$, subgroups and subvarieties are not compatible. Semiabelian varieties are all commutative as algebraic groups, but there are many noncommutative groups where the main obstruction to the Mordell–Lang theorem (namely, containing the additive group) is not present; for example, $\mathbb{H}^\times$, or the multiplicative group of any division algebra. As such, unit equations in division algebras and the Mordell–Lang theorem for noncommutative algebraic groups are natural follow-up directions of research after my results.

## 4. Discrete random matrix theory

Random matrix theory, originated in 1950s from modeling distributions in physics, is nowadays a vast and ever-growing field that is deeply connected with many areas of mathematics, such as number theory and representation theory. Classical random matrix theory concerns the distribution of spectral invariants (eigenvalues, singular values, etc.) of a random $n \times n$ real or complex matrices, as $n \to \infty$. A result of Wigner (see [2, Chap. 2]) that marks the birth of this subject reveals the ubiquity of the "universality" phenomenon: it states that a random $n \times n$ Hermitian matrix with independent entries (subject to the Hermitian constraint) whose distributions are mean 0 and variance 1 has eigenvalue distribution approaching the semicircular distribution as $n \to \infty$, after a suitable normalization. Here, the limiting distribution of eigenvalues does not depend on the precise distributions of the entries.

Discrete random matrix theory, on the other hand, concerns integer matrices or $p$-adic integer matrices. An important spectral invariant is the Smith normal form, or equivalently, the structure of the cokernel as an abelian group. Since the proposal of the Cohen–Lenstra heuristics [21], the cokernels of random ($p$-adic) integer matrices have been used to model the distributions of many finite abelian groups arising in arithmetic statistics. The universality phenomenon is also ubiquitous in discrete random matrix theory. An exemplary result of this phenomenon by Wood [51] states that if a random $n \times n$ matrix $X$ over $\mathbb{Z}_p$ have independent entries whose residue classes mod $p$ are not too concentrated, then the limiting distribution of $\operatorname{cok}(X)$ always converges to the "Cohen–Lenstra distribution" as $n \to \infty$. The Cohen–Lenstra distribution assigns to a finite abelian $p$-group a probability inversely proportional to the number of its automorphisms. In other models, such as random symmetric matrices [50], the universal distributions may be variants of the Cohen–Lenstra distribution, but they all feature $1/\#\operatorname{Aut}$.

Widely studied random matrix models generally have independent entries, possibly subject to symmetry constraint. With Cheong [11], I discovered the presence of the Cohen–Lenstra distribution in a model where entries are intricately dependent. More precisely, we fix a monic polynomial $P(t) \in \mathbb{Z}_p[t]$, and consider the random matrix $P(X)$, where $X \in \operatorname{Mat}_n(\mathbb{Z}_p)$ is uniformly random (with respect to the Haar measure). More generally, we fix monic polynomials

$P_1(t), \dots, P_l(t) \in \mathbb{Z}_p[t]$ that are pairwise coprime mod $p$, and consider the joint distribution of $\operatorname{cok}(P_j(X))$ for $1 \le j \le l$. It is expected that as $n \to \infty$, the limiting distributions of $\operatorname{cok}(P_j(X))$ should be independent and each should approach a variant of the Cohen–Lenstra distribution depending on how $P_j(t)$ factorizes modulo $p$. We proved the first few nontrivial cases in [11].

**Theorem 4.1** (Cheong–H. [11])**.** *Let $P_1(t), \dots, P_l(t) \in \mathbb{Z}_p[t]$ be monic polynomials whose reductions mod $p$ are distinct irreducible polynomials in $\mathbb{F}_p[t]$. Suppose $\deg P_l(t) = 1$, then given a finite abelian $p$-group $G$, we have*

$$\lim_{n\to\infty} \operatorname*{Prob}_{X \in \mathrm{Mat}_n(\mathbb{Z}_p)^{\mathrm{Haar}}} \begin{pmatrix} \operatorname{cok} P_1(X) = \cdots = \operatorname{cok} P_{l-1}(X) = 0 \\ \operatorname{cok} P_l(X) \cong G \end{pmatrix} = \frac{1}{\#\operatorname{Aut}(G)} \prod_{j=1}^{l} \prod_{i=1}^{\infty} (1 - p^{-i \deg P_j}). \tag{4.1}$$

After the appearance of this work, more general theorems were proved in [14–16, 39]. In my recent joint work [13] with Cheong, I obtained results about the distribution of $\operatorname{cok}(P(X))$ where $X$ has a concentrated residue class mod $p$. They are the first results where $X$ violates Wood's "$\varepsilon$-balanced" assumption that the entries are not too concentrated mod $p$. Even the simplest case $P(t) = t$ gives a new distribution. To provide a glimpse, I state this case in an imprecise language.

**Theorem 4.2** (Cheong–H. [13])**.** *Fix a matrix $A \in \mathrm{Mat}_n(\mathbb{F}_p)$, and denote by $\mathrm{Mat}_n(\mathbb{Z}_p)_A$ the set of matrices in $\mathrm{Mat}_n(\mathbb{Z}_p)$ that reduces to $A$ mod $p$. Let $X$ be a random matrix in $\mathrm{Mat}_n(\mathbb{Z}_p)_A$. Assume $n$ is large, and the entries of $X$ are independent and not too concentrated mod $p^2$. Let $r := \dim_{\mathbb{F}_p} \operatorname{cok}(A)$ be the corank of $A$. Then for a finite abelian $p$-group $G$, we have*

$$\operatorname*{Prob}_{X \in \mathrm{Mat}_n(\mathbb{Z}_p)_A} (\operatorname{cok}(X) \cong G) \approx \frac{p^{r^2}}{\#\operatorname{Aut}(G)} \prod_{i=1}^{r} (1 - p^{-i})^2 \tag{4.2}$$

*if $\dim_{\mathbb{F}_p} G/pG = r$, and the probability is zero otherwise.*

## 5. Miscellany and future research

I realize that my statement is already quite long. I hope I have reasonably illustrated the breadth and depth of my interests and expertise in number theory and algebraic geometry, and my potential to continue conducting novel research independently or collaboratively. However, I have left out some of my work in combinatorics and commutative algebra. In my joint work [18], for a random permutation of $m$ copies of $\{1, \dots, n\}$, we investigated the statistics of its longest continuously increasing subsequences, and gave exact and asymptotic formulas for the expected value. When discussing my joint works [32–34] before, I have not mentioned that they fit in a rich context of lattice zeta functions introduced by Solomon [46]. As a necessary step to prove Theorem 2.5, we showed that the motivic lattice zeta function of any lattice over the completed local ring of any reduced curve singularity is rational. This can be viewed as a function field analogue of a result of Bushnell and Reiner [10]. In [32], for the purpose of classifying lattices of $k[[T]]^d$, we found an analogue of the Hermite normal form that directly recovers the Smith normal form, which is what the Hermite normal form does not do. Along the way to prove some results in my works [13, 31], I proved statistical formulas about Cohen–Lenstra-random modules over singular rings in the absence of a combinatorial classification of these modules.

I will continue the investigation of topics I have discussed in the previous sections. Every finding, even in special cases, has led to new insights, connections and conjectures. I am excited by the plethora of prospects that my current works have presented me. For example, in current work with Ono, I proved the Sato–Tate distribution for matrix points of small sizes on certain Calabi–Yau 3-folds. The crux is finding asymptotic counts of commuting triples of $n \times n$ matrices, which is out of reach when $n$ is large. I am investigating whether weaker assumptions on commuting triple counts would imply the Sato–Tate distribution. For another example, many explicit formulas about the Quot scheme of points appear to give polynomials with positive coefficients. I am very interested in its proof, both combinatorial (is there a cancellation-free formula?) and geometrical (does the relevant moduli space have a cell decomposition?). I am investigating whether careful analysis of Hall–Littlewood symmetric functions or the torus action on the moduli space would achieve this goal. Finally, I wish to branch out and begin studying problems in neighboring fields of mathematics, such as intersection theory and $p$-adic representation theory, partially for their potential relevance in Theorem 2.7 and Conjecture 2.10.

## References


[1] S. Ahlgren, K. Ono, and D. Penniston. Zeta functions of an infinite family of $K3$ surfaces. *Amer. J. Math.*, 124(2):353–368, 2002.

[2] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[3] V. I. Arnol'd. The cohomology ring of the group of dyed braids. *Mat. Zametki*, 5:227–231, 1969.

[4] A. Baker. Linear forms in the logarithms of algebraic numbers. I. *Mathematika*, 13:204–216, 1966.

[5] T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor. A family of Calabi–Yau varieties and potential automorphy II. *Publ. Res. Inst. Math. Sci.*, 47(1):29–98, 2011.

[6] K. Behrend, J. Bryan, and B. Szendrői. Motivic degree zero Donaldson-Thomas invariants. *Invent. Math.*, 192(1):111–160, 2013.

[7] K. Behrend and A. Dhillon. On the motivic class of the stack of bundles. *Adv. Math.*, 212(2):617–644, 2007.

[8] D. Bejleri, D. Ranganathan, and R. Vakil. Motivic Hilbert zeta functions of curves are rational. *J. Inst. Math. Jussieu*, 19(3):947–964, 2020.

[9] E. Bifet. Sur les points fixes du schéma $\mathrm{Quot}_{\mathscr{O}_X^r/X/k}$ sous l'action du tore $\mathbf{G}_{m,k}^r$. *C. R. Acad. Sci. Paris Sér. I Math.*, 309(9):609–612, 1989.

[10] C. J. Bushnell and I. Reiner. Zeta functions of arithmetic orders and Solomon's conjectures. *Math. Z.*, 173(2):135–161, 1980.

[11] G. Cheong and Y. Huang. Cohen–Lenstra distributions via random matrices over complete discrete valuation rings with finite residue fields. *Illinois Journal of Mathematics*, 65(2):385–415, 2021.

[12] G. Cheong and Y. Huang. Betti and Hodge numbers of configuration spaces of a punctured elliptic curve from its zeta functions. *Trans. Amer. Math. Soc.*, 375(9):6363–6383, 2022.

[13] G. Cheong and Y. Huang. Universality of the cokernel of a random integral matrix with concentrated residue. In preparation, 2023++.

[14] G. Cheong and N. Kaplan. Generalizations of results of Friedman and Washington on cokernels of random $p$-adic matrices. *J. Algebra*, 604:636–663, 2022.

[15] G. Cheong, Y. Liang, and M. Strand. Polynomial equations for matrices over integers modulo a prime power and the cokernel of a random matrix. *Linear Algebra Appl.*, 677:1–30, 2023.

[16] G. Cheong and M. Yu. The distribution of the cokernel of a polynomial evaluated at a random integral matrix. Preprint `https://arxiv.org/abs/2303.09125`.

[17] T. Church, J. S. Ellenberg, and B. Farb. Representation stability in cohomology and asymptotics for families of varieties over finite fields. In *Algebraic topology: applications and new directions*, volume 620 of *Contemp. Math.*, pages 1–54. Amer. Math. Soc., Providence, RI, 2014.

[18] A. Clifton, B. Deb, Y. Huang, S. Spiro, and S. Yoo. Continuously increasing subsequences of random multiset permutations. *European J. Combin.*, 110:103708, 2023.

[19] L. Clozel, M. Harris, and R. Taylor. Automorphy for some $\ell$-adic lifts of automorphic mod $\ell$ Galois representations. *Publ. Math. Inst. Hautes Études Sci.*, 108:1–181, 2008. With Appendix A, summarizing unpublished work of Russ Mann, and Appendix B by Marie-France Vignéras.

[20] F. R. Cohen. On the mapping class groups for punctured spheres, the hyperelliptic mapping class groups, SO(3), and $\mathrm{Spin}^c(3)$. *Amer. J. Math.*, 115(2):389–434, 1993.

[21] H. Cohen and H. W. Lenstra, Jr. Heuristics on class groups of number fields. In *Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983)*, volume 1068 of *Lecture Notes in Math.*, pages 33–62. Springer, Berlin, 1984.

[22] G. C. Drummond-Cole and B. Knudsen. Betti numbers of configuration spaces of surfaces. *J. Lond. Math. Soc. (2)*, 96(2):367–393, 2017.

[23] J.-H. Evertse and K. Győry. *Unit Equations in Diophantine Number Theory*. Cambridge University Press, Cambridge, 2015.

[24] W. Feit and N. J. Fine. Pairs of commuting matrices over a finite field. *Duke Math. J.*, 27:91–94, 1960.

[25] V. V. Gorjunov. Cohomology of braid groups of series $C$ and $D$. *Trudy Moskov. Mat. Obshch.*, 42:234–242, 1981.

[26] J. Greene. Hypergeometric functions over finite fields. *Trans. Amer. Math. Soc.*, 301(1):77–101, 1987.

[27] M. Harris, N. Shepherd-Barron, and R. Taylor. A family of Calabi–Yau varieties and potential automorphy. *Ann. of Math. (2)*, 171(2):779–813, 2010.

[28] T. Hausel and F. Rodriguez-Villegas. Mixed Hodge polynomials of character varieties. *Invent. Math.*, 174(3):555–624, 2008. With an appendix by Nicholas M. Katz.

[29] Y. Huang. Cohomology of configuration spaces on punctured varieties. Preprint `https://arxiv.org/abs/2011.07153`, 2020+.

[30] Y. Huang. Unit equations on quaternions. *Q. J. Math.*, 71(4):1521–1534, 2020.

[31] Y. Huang. Mutually annihilating matrices, and a Cohen–Lenstra series for the nodal singularity. *J. Algebra*, 619:26–50, 2023.

[32] Y. Huang and R. Jiang. Spiral shifting operators from the enumeration of finite-index submodules of $\mathbb{F}_q[[T]]^d$. Preprint `https://arxiv.org/abs/2210.10215`, 2022+.

[33] Y. Huang and R. Jiang. Generating series for torsion-free bundles over singular curves: rationality, duality and modularity. In preparation, 2023++.

[34] Y. Huang and R. Jiang. Punctual Quot schemes and Cohen–Lenstra series of the cusp singularity. Preprint `https://arxiv.org/abs/2305.06411`, 2023+.

[35] Y. Huang, K. Ono, and H. Saad. Counting matrix points on certain varieties over finite fields. *Contemp. Math., Amer. Math. Soc.*, accepted for publication, 2023. `https://arxiv.org/abs/2302.04830`.

[36] S. Kallel. Symmetric products, duality and homological dimension of configuration spaces. In *Groups, homotopy and configuration spaces*, volume 13 of *Geom. Topol. Monogr.*, pages 499–527. Geom. Topol. Publ., Coventry, 2008.

[37] M. Kapranov. The elliptic curve in the $S$-duality theory and Eisenstein series for Kac–Moody groups. Preprint `https://arxiv.org/abs/math/0001005`.

[38] S. Lang. Integral points on curves. *Publications Mathématiques de l'IHÉS*, 6:27–43, 1960.

[39] J. Lee. Joint distribution of the cokernels of random $p$-adic matrices. *Forum Math.*, 35(4):1005–1020, 2023.

[40] I. G. Macdonald. The Poincaré polynomial of a symmetric product. *Proc. Cambridge Philos. Soc.*, 58:563–568, 1962.

[41] F. Napolitano. Configuration spaces on surfaces. *C. R. Acad. Sci. Paris Sér. I Math.*, 327(10):887–892, 1998.

[42] F. Napolitano. On the cohomology of configuration spaces on surfaces. *J. London Math. Soc. (2)*, 68(2):477–492, 2003.

[43] A. Oblomkov, J. Rasmussen, and V. Shende. The Hilbert scheme of a plane curve singularity and the HOMFLY homology of its link. *Geom. Topol.*, 22(2):645–691, 2018. With an appendix by Eugene Gorsky.

[44] K. Ono, H. Saad, and N. Saikia. Distribution of values of Gaussian hypergeometric functions. *Pure Appl. Math. Q.*, 19(1):371–407, 2023.

[45] R. Pandharipande and R. P. Thomas. Stable pairs and BPS invariants. *J. Amer. Math. Soc.*, 23(1):267–297, 2010.

[46] L. Solomon. Zeta functions and integral representation theory. *Advances in Math.*, 26(3):306–326, 1977.
[47] R. Taylor. Automorphy for some $\ell$-adic lifts of automorphic mod $\ell$ Galois representations. II. *Publ. Math. Inst. Hautes Études Sci.*, 108:183–239, 2008.
[48] P. Tosteson. Lattice spectral sequences and cohomology of configuration spaces. 2016. arXiv: 1612.06034.
[49] R. Vakil and M. M. Wood. Discriminants in the Grothendieck ring. *Duke Math. J.*, 164(6):1139–1185, 2015.
[50] M. M. Wood. The distribution of sandpile groups of random graphs. *J. Amer. Math. Soc.*, 30(4):915–958, 2017.
[51] M. M. Wood. Random integral matrices and the Cohen-Lenstra heuristics. *Amer. J. Math.*, 141(2):383–398, 2019.