Cohomology of configuration spaces on punctured varieties

Yifeng Huang[*]

February 5, 2022

Abstract

We describe how the rational cohomology groups of unordered configuration spaces are affected by puncturing a point from the base space, by giving an identity involving Betti numbers that holds whenever the base space is a smooth noncompact complex variety satisfying a flexible mixed-Hodge-theoretic assumption. This extends certain results of Napolitano and Kallel. We also obtain a similar identity involving mixed Hodge numbers, which refines the identity about Betti numbers. Furthermore, we refine the above results to an equivariant theorem about the cohomology of ordered configuration spaces. Our approach involves explicit algebraic computations based on an Orlik–Solomon spectral sequence.

1 Introduction

1.1 History and motivation

Let X be a topological space. Define the n-th *ordered configuration space* of X as the complement

$$F(X,n) := X^n - \bigcup_{i<j}\{(x_1,...,x_n)\in X^n : x_i = x_j\} \tag{1.1}$$

of the big diagonal in the product space, and define the n-th *unordered configuration space* as a topological quotient

$$\mathrm{Conf}^n(X) := F(X,n)/S_n \tag{1.2}$$

by the symmetric group S_n acting by permuting coordinates. If X is a quasiprojective variety over a field, then $F(X,n)$ is defined by the same formula, understood as an open subvariety of X^n, and $\mathrm{Conf}^n(X)$ is the scheme-theoretic quotient of $F(X,n)$ by the natural action of S_n. Both $F(X,n)$ and $\mathrm{Conf}^n(X)$ are quasiprojective varieties as well.

[*]Dept. of Mathematics, University of Michigan. Email Address: huangyf@umich.edu.

Configuration spaces form a playground where topology and number theory often interplay. As a notable example, Church, Ellenberg and Farb [CEF14] pointed out and explained the connection between the Betti numbers of configuration spaces of the complex plane $\mathbb{C}$ and the counting of square-free polynomials over a finite field. Let $\mathbb{A}^1_{\mathbb{F}_q}$ denote the affine line over the finite field $\mathbb{F}_q$ with q elements. Church, Ellenberg and Farb realized that the set of $\mathbb{F}_q$-points of the $\mathbb{F}_q$-variety $\mathrm{Conf}^n(\mathbb{A}^1_{\mathbb{F}_q})$ can be identified with the set of degree-n monic square-free $\mathbb{F}_q$-polynomials, and used this connection to show the identity

$$|\mathrm{Conf}^n(\mathbb{A}^1_{\mathbb{F}_q})(\mathbb{F}_q)| = \sum_{i=0}^{\infty}(-1)^i h^i(\mathrm{Conf}^n(\mathbb{C}))q^{n-i}, \tag{1.3}$$

where the left-hand side is the number of monic square-free $\mathbb{F}_q$-polynomials of degree n, and the right-hand side involves $h^i(\mathrm{Conf}^n(\mathbb{C}))$, the i-th (rational) Betti number of the configuration space $\mathrm{Conf}^n(\mathbb{C})$.

The identity (1.3) compares the following two aspects about the affine line $\mathbb{A}^1$: the cohomology of its configuration spaces over $\mathbb{C}$, and the point counts of its configuration spaces over $\mathbb{F}_q$. While such a connection does not exist for a variety in general, other phenomena do exist that occur in parallel in these two aspects. The focus of this paper is a "splitting" formula that compares the cohomology of configuration spaces of a topological space X with the configuration spaces of X minus one point. The study of such a splitting dates back to Fred Cohen [Coh93] and Gorjunov [Gor81].

Suppose X is a connected noncompact topological surface without boundary and P is a point of X. Napolitano [Nap03] showed that

$$H^i(\mathrm{Conf}^n(X-P);\mathbb{Z}) \cong \bigoplus_{t=0}^{\infty} H^{i-t}(\mathrm{Conf}^{n-t}(X);\mathbb{Z}). \tag{1.4}$$

In particular, this formula determines the Betti numbers of $\mathrm{Conf}^n(X-P)$ from the Betti numbers of $\mathrm{Conf}^m(X-P), m \leq n$.

He noted that (1.4) does not hold if X is compact, even in field coefficients. In fact, if X is a connected closed orientable surface of genus g and P is a point of X, then a different formula holds:

$$H^i(\mathrm{Conf}^n(X);\mathbb{Z}/2\mathbb{Z}) = H^i(\mathrm{Conf}^n(X-P);\mathbb{Z}/2\mathbb{Z}) \oplus H^{i-2}(\mathrm{Conf}^{n-1}(X-P),\mathbb{Z}/2\mathbb{Z}). \tag{1.5}$$

Kallel [Kal08] extended the splitting (1.4) in field coefficients to higher dimensions. Let X be a connected noncompact manifold of even dimension $2d$, and let P be a point of X. Kallel showed that the following formula holds for any coefficient field $\mathbb{F}$ if X can be obtained from removing $r \geq 1$ points from a connected closed orientable manifold:

$$H^i(\mathrm{Conf}^n(X-P);\mathbb{F}) \cong \bigoplus_{t=0}^{\infty} H^{i-(2d-1)t}(\mathrm{Conf}^{n-t}(X);\mathbb{F}). \tag{1.6}$$

If X is a several-punctured orientable surface, then $d=1$, so (1.6) recovers the field-coefficient version of (1.4). If X is a puncturing of a connected smooth compact variety over $\mathbb{C}$ of (complex) dimension d, then the formula (1.6) holds, and the special case $\mathbb{F} = \mathbb{Q}$ implies the following identity of power series in u and t:

$$\sum_{i,n\geq 0} h^i(\mathrm{Conf}^n(X-P))(-u)^i t^n = \frac{1}{1+u^{2d-1}t} \sum_{i,n\geq 0} h^i(\mathrm{Conf}^n(X))(-u)^i t^n. \tag{1.7}$$

Notably, an analogous formula also holds in terms of point counting over finite fields. Let X be any variety over $\mathbb{F}_q$ and let P be an $\mathbb{F}_q$-point of X. As a consequence of a result due to Vakil and Wood [VW15, Proposition 5.9], we have

$$\sum_{n=0}^{\infty} |\mathrm{Conf}^n(X-P)(\mathbb{F}_q)| t^n = \frac{1}{1+t} \sum_{n=0}^{\infty} |\mathrm{Conf}^n(X)(\mathbb{F}_q)| t^n. \tag{1.8}$$

The main contribution of this paper is to provide a version of (1.7) and (1.8) in terms of mixed Hodge numbers. Our formula refines both (1.7) and a consequence of (1.8) in a uniform way, which explains the analogy between (1.7) and (1.8). We also extend Kallel's formula (1.7) to more general families of smooth noncompact complex varieties.

1.2 Main result

To state the main theorem and explain how it connects (1.7) and (1.8), we recall some necessary concepts in the mixed Hodge theory (for detailed references, see Deligne's [Del75] or [Del71]). For any complex variety X and every $p, q, i \geq 0$, Deligne defined a complex vector space $H^{p,q;i}(X)$ that is a subquotient of the singular cohomology $H^i(X;\mathbb{C})$ in complex coefficients. The dimension $h^{p,q;i}(X)$ of $H^{p,q;i}(X)$ is called the *mixed Hodge number* of X of *Hodge type* (p, q). We have

$$h^i(X) = \sum_{p,q\geq 0} h^{p,q;i}(X), \tag{1.9}$$

so the data of mixed Hodge numbers refine the data of Betti numbers.

Theorem 1.1 (Main Result)**.** *Let X be a connected compact smooth complex variety of dimension d with $r \geq 1$ points punctured (in particular, X is never compact). Let P be a point of X, then the mixed Hodge numbers of the further-punctured variety $X - P$ are given by*

$$h^{p,q;i}(\mathrm{Conf}^n(X-P)) = \sum_{t\geq 0} h^{p-dt,q-dt;i-(2d-1)t}(\mathrm{Conf}^{n-t}(X)). \tag{1.10}$$

Equivalently,

$$\begin{aligned} &\sum_{p,q,i,n\geq 0} h^{p,q;i}(\mathrm{Conf}^n(X-P)) x^p y^q (-u)^i t^n \\ &= \frac{1}{1+(xy)^d u^{2d-1} t} \sum_{p,q,i,n\geq 0} h^{p,q;i}(\mathrm{Conf}^n(X)) x^p y^q (-u)^i t^n. \end{aligned} \tag{1.11}$$

The formula (1.11) describes the relationship between mixed Hodge numbers of $\mathrm{Conf}^n(X-P)$ for all n and mixed Hodge numbers of $\mathrm{Conf}^n(X)$ for all n. If we substitute $x = y = 1$ in (1.11), we recover Kallel's formula (1.7) about Betti numbers.

To explain the connection between Theorem 1.1 and the point counting formula (1.8), we introduce an analogue of point counting for a complex variety. For a connected smooth complex variety X of dimension d, define the *E-polynomial* (following the notation of [HRV08]) as

$$E(X; x, y) = \sum_{p,q,i \geq 0} (-1)^i h^{p,q;i}(X) x^{d-p} y^{d-q} \in \mathbb{Z}[x, y]. \tag{1.12}$$

The notion of E-polynomial can be uniquely extended to all complex varieties, such that the followings hold:

(a) For any complex varieties X and Y, we have $E(X \times Y; x, y) = E(X; x, y)E(Y; x, y)$;

(b) If Z is a closed subvariety of a complex variety X, and $U = X - Z$ is the complement of Z, then $E(X; x, y) = E(U; x, y) + E(Z; x, y)$.

In other words, the E-polynomial behaves well with the cartesian product and "cut-and-paste", just like the point count for varieties over a finite field. Like the E-polynomial, an invariant that satisfies properties (a) and (b) above is said to be *motivic*. We remark that Betti numbers and mixed Hodge numbers are not motivic.

Because the E-polynomial is motivic, the formula [VW15, Theorem 5.9] of Vakil and Wood from which the point count formula (1.8) is derived also implies

$$\sum_{n=0}^{\infty} E(\mathrm{Conf}^n(X - P); x, y)t^n = \frac{1}{1+t} \sum_{n=0}^{\infty} E(\mathrm{Conf}^n(X); x, y)t^n, \tag{1.13}$$

where X is any complex variety and P is a point of X.

Given (1.11), if we substitute $x \mapsto x^{-1}, y \mapsto y^{-1}, u \mapsto 1$ and $t \mapsto (xy)^d t$, then we recover (1.13) about E-polynomials. We point out that the notion of E-polynomial of a complex variety is indeed a reasonable analogue of the point count of a variety over $\mathbb{F}_q$: in fact, due to a result of Nicholas Katz [HRV08, §6], the E-polynomial of a complex variety X is uniquely determined by the point count of spreadings out of X to "enough" finite fields.

In conclusion, for a complex variety satisfying the assumption of Theorem 1.1, the formula (1.11) is a common refinement of Kallel's formula (1.7) about Betti numbers and Vakil and Wood's formula (1.13) about E-polynomials, which can be viewed as an analogue of the point count formula (1.8) but for complex varieties.

1.3 Applications to algebraic curves

Let $\Sigma_{g,r}$ be a smooth projective algebraic curve of genus g minus r points. Drummond-Cole and Knudsen [DCK17] computed the Betti numbers $h^i(\mathrm{Conf}^n(\Sigma_{g,r}))$. As an application of Theorem 1.1, we describe the picture about the mixed Hodge numbers $h^{p,q;i}(\mathrm{Conf}^n(\Sigma_{g,r}))$. The nicest case is when $g = 0, r \geq 1$, where the rational mixed Hodge structure on $H^i(\mathrm{Conf}^n(\Sigma_{0,r}))$ was shown to be pure of weight $2i$ by Kim [Kim94]. When $g = r = 1$, Cheong and the author [CH] proved that the rational mixed Hodge structure on $H^i(\mathrm{Conf}^n(\Sigma_{1,1}))$ is pure of weight $w(i) = \lfloor 3i/2 \rfloor$.

In both cases above ($g = 0, r \geq 1$ or $g = r = 1$), thanks to the purity statement, the motivic formula of Vakil and Wood gives rise to a rational generating function in four variables that encodes the mixed Hodge numbers $h^{p,q;i}(\mathrm{Conf}^n(\Sigma_{g,r}))$ with p, q, i, n varying, as is explained in [CH]. The methods of [CH] can also be applied to the case $g \geq 2, r = 1$ to give an explicit basis for $H^{p,q;i}(\mathrm{Conf}^n(\Sigma_{g,1}))$, thus giving the mixed Hodge numbers as a counting formula, except that the mixed Hodge structure on $H^i(\mathrm{Conf}^n(\Sigma_{g,1}))$ is no longer pure, so that the "motivic trick" does not work to give a rational generating function.

Applying Theorem 1.1 to the formula in [CH] for $g = r = 1$, we obtain the mixed Hodge numbers $h^{p,q;i}(\mathrm{Conf}^n(\Sigma_{g,r}))$ for $g = 1, r = 2$. Similarly we can obtain an analogous formula for $g = 1$ and any $r \geq 2$ by applying Theorem 1.1 repetitively. In particular, we may generate an explicit table of mixed Hodge numbers of $\mathrm{Conf}^n(\Sigma_{1,r})$ for $r \geq 2$. We emphasize that this very table would show that the mixed Hodge structure on $H^i(\mathrm{Conf}^n(\Sigma_{1,r}))$ is not pure for all $r \geq 2$ (except possibly for small i and n), so the computation of $h^{p,q;i}(\mathrm{Conf}^n(\Sigma_{1,r}))$ for $r \geq 2$ could not be a consequence of the "motivic trick". For $g \geq 2, r \geq 1$, it turns out that the mixed Hodge structure on $H^i(\mathrm{Conf}^n(\Sigma_{g,r}))$ is not pure, either, but Theorem 1.1 asserts that the mixed Hodge numbers of $\mathrm{Conf}^n(\Sigma_{g,r})$ (for all n) are determined by the mixed Hodge numbers of $\mathrm{Conf}^n(\Sigma_{g,1})$ (for all n).

Theorem 1.1 does not work to relate $\mathrm{Conf}^n(\Sigma_{g,0})$ and $\mathrm{Conf}^n(\Sigma_{g,1})$. Fortunately, when $r = 0$, Pagaria [Pag] obtained a (not necessarily rational) generating function that computes all mixed Hodge numbers $h^{p,q;i}(\mathrm{Conf}^n(\Sigma_{g,0}))$. In conclusion, the formulas for $h^{p,q;i}(\mathrm{Conf}^n(\Sigma_{g,r}))$ can be organized into three essentially different cases: the compact case $r = 0$, following [Pag]; the one-punctured case $r = 1$, following [CH]; and the multi-punctured case $r \geq 2$, which can be directly obtained from the $r = 1$ case via Theorem 1.1.

1.4 Generalizations and refinements

The conclusion of Theorem 1.1 and Kallel's Betti number formula (1.7) generalize to more examples of X. For the purpose of this paper, we only consider mixed Hodge structures in rational coefficients. Given a rational number $\lambda \geq 0$, we say a complex variety X to be *pure of slope* λ if the mixed Hodge structure of $H^i(X, \mathbb{Q})$ is pure of weight $\lambda \cdot i$ for any integer $i \geq 0$, namely, the mixed Hodge number $h^{p,q;i}(X)$ is zero unless $p + q = \lambda \cdot i$. We point out that part of the requirement is that $H^i(X, \mathbb{Q}) = 0$ for all i such that $\lambda \cdot i$ is not an integer. By [Del75, §7, p. 82], the slope of a smooth complex variety (if exists) must be a rational number $1 \leq \lambda \leq 2$.

Remark 1.2. Technically, the slope may not be unique. If $h^i(X) = 0$ for all $i > 0$, then a slope exists and is arbitrary; in any other cases, the slope is unique (if exists). For the purpose of convenience only, we say *the* slope of X is 1 if $h^i(X) = 0$ for all $i > 0$.

Theorem 1.3. *Let $\overline{X}$ be a connected smooth noncompact complex variety that is pure of slope λ for some rational number $1 \leq \lambda \leq 2$. Let X be an r-puncture of $\overline{X}$ for some $r \geq 0$; in other words, X is of the form $X = \overline{X} - \{P_1, \ldots, P_r\}$ where P_i are points of $\overline{X}$. Then for any point P of X, the conclusion of Theorem* 1.1 *holds as well. As a consequence, Kallel's* (1.7) *holds for X and P.*

Theorem 1.1 is a special case of Theorem 1.3 because for any multi-punctured variety X in the setting of Theorem 1.1, we can choose $\overline{X}$ in Theorem 1.3 to be a one-punctured variety (a connected smooth compact variety minus one point). It was noted by Dupont [Dup16, Theorem 2.10] that a one-punctured variety is pure of slope 1, so Theorem 1.3 applies. Therefore, the proof of Theorem 1.1 is complete after we prove Theorem 1.3.

To list a few other examples of varieties that satisfy the assumption for $\overline{X}$ (see Section 5 for details),

- $\overline{X}$ is an affine space $\mathbb{C}^d$, or the complement of a hyperplane arrangement therein. In this case, $\overline{X}$ is pure of slope 2.
- $\overline{X}$ is the torus $(\mathbb{C}^*)^d$, or the complement of a union of 1-codimensional subtori therein. In this case, $\overline{X}$ is pure of slope 2.
- $\overline{X}$ is the complement of a smooth plane curve in the projective plane $\mathbb{P}^2$. In this case, X is pure of slope 3/2.
- If $\overline{X}$ is pure of slope λ, then so is a smooth quotient of $\overline{X}$ by a finite group. For example, we could take $\overline{X} = \mathrm{Conf}^m(\mathbb{C})$, in which case $\overline{X}$ is pure of slope 2.

By taking $\overline{X}$ from the list above, and possibly puncturing finitely many points, we get abundant examples of even-dimensional noncompact manifolds where Kallel's formula (1.7) holds. We remark that our current proof of the purely topological statement (1.7) for these manifolds requires the structure of complex variety and mixed Hodge theory.

The proof of Theorem 1.3 depends on an equivariant result, which is itself a strong refinement of Theorem 1.3. The isomorphism in the next theorem is in the category of split mixed Hodge structures, a category that encodes the complex vector spaces $H^{p,q}(M)$ associated to a mixed Hodge structure M (and thus the mixed Hodge numbers $h^{p,q}(M) := \dim_{\mathbb{C}} H^{p,q}(M)$) but forgets the extra structures encoded in the weight filtration and the Hodge filtration. A *split mixed Hodge structure* (over $\mathbb{Q}$) is a direct sum of pure Hodge structures. An *S_n-split mixed Hodge structure* is a split mixed Hodge structure together with an action by S_n. For a mixed Hodge structure M, its associated graded $\mathrm{gr}\, M$ with respect to the weight filtration is a split mixed Hodge structure.

Theorem 1.4. *Let X be a complex variety as in Theorem 1.3, and P be a point of X. Recall that $F(X, n)$ denotes the n-th ordered configuration space of X. Then we have a noncanonical isomorphism of S_n-representations*

$$H^i(F(X-P,n),\mathbb{Q}) \cong \bigoplus_{t=0}^{\infty} \mathrm{Ind}_{S_{n-t}}^{S_n} H^{i-(2d-1)t}(F(X,n-t),\mathbb{Q}). \tag{1.14}$$

Moreover, we have a noncanonical isomorphism in the (semi-simple) category of S_n-split mixed Hodge structures

$$\mathrm{gr}\, H^i(F(X-P,n),\mathbb{Q}) \cong \bigoplus_{t=0}^{\infty} \mathrm{Ind}_{S_{n-t}}^{S_n} \mathrm{gr}\, H^{i-(2d-1)t}(F(X,n-t),\mathbb{Q})(-d\cdot t), \tag{1.15}$$

where $(-d \cdot t)$ denotes the Tate twist.

Here, the operator $\mathrm{Ind}_{S_{n-t}}^{S_n}$ denotes the induction of an S_{n-t}-representation to an S_n-representation, where S_{n-t} is the subgroup of S_n consisting of permutations that permute the first $n-t$ elements and fix the last t elements. The Tate twist shifts a mixed Hodge structure according to the rule

$$H^{p,q}(M(n)) = H^{p+n,q+n}(M). \tag{1.16}$$

As a special case of Shapiro's lemma [Wei94, p. 172] (or alternatively, the Frobenius reciprocity), we have the following standard fact: for any subgroup H of a finite group G, and a representation V of H, then the G-invariant of the induction of V is isomorphic to the H-invariant of V as a vector space:

$$(\mathrm{Ind}_H^G V)^G \cong V^H. \tag{1.17}$$

As a result, by taking the S_n-invariants of both sides of (1.15) and extracting the mixed Hodge numbers of Hodge type (p,q), we get

$$h^{p,q;i}(\mathrm{Conf}^n(X-P)) = \sum_{t=0}^{\infty} h^{p-dt,q-dt;i-(2d-1)t}(\mathrm{Conf}^{n-t}(X)), \tag{1.18}$$

which is equivalent to our main result (1.11).

We conclude this section with an alternative form of (1.15) that notably does not involve the dimension d. Let $H_c^i(X;\mathbb{Q})$ denote the compactly supported cohomology of X, which is also equipped with a mixed Hodge structure. Using the Poincaré duality, the formula (1.15) is equivalent to the following isomorphism of S_n-split mixed Hodge structures:

$$\mathrm{gr}\, H_c^i(F(X-P,n),\mathbb{Q}) \cong \bigoplus_{t=0}^{\infty} \mathrm{Ind}_{S_{n-t}}^{S_n} \mathrm{gr}\, H_c^{i-t}(F(X,n-t),\mathbb{Q}). \tag{1.19}$$

1.5 Overview of methods

Our approach to Theorem 1.1 is to prove an equivariant theorem (Theorem 1.4) about the cohomology of ordered configuration spaces $F(X,n)$. To prove Theorem 1.4, we use a spectral sequence similar to a Leray spectral sequence described by Totaro [Tot96]. The spectral sequence we use has its E_1 page as a differential graded algebra with explicit generators and relations, and the spectral sequence degenerates on the E_2 page; these allow explicit computation of cohomology groups of $F(X,n)$. We point out that in order to have degeneration on E_2, we must use a spectral sequence different from the one described in [Tot96].

Like the Leray spectral sequence in [Tot96], our spectral sequence also remembers mixed-Hodge-theoretic information of $F(X,n)$ and respects the action of S_n on $F(X,n)$. Looking at a similar spectral sequence for $F(X-P,n)$ and comparing it with the spectral sequence for $F(X,n)$, we are able to construct (artificially using the generators and relations on the E_1 page) an equivariant isomorphism that expresses the cohomology of $F(X-P,n)$ in terms of the cohomology of $F(X,m)$ for all $m \le n$; this is Theorem 1.4. Taking the S_n-invariant parts of both sides, we get an identity that compares the mixed Hodge numbers of $\mathrm{Conf}^n(X)$ (for all n) and $\mathrm{Conf}^n(X-P)$ (for all n), which is the identity required in Theorem 1.1.

We now state the desired properties of the spectral sequence we use.

Theorem 1.5. *Let $\overline{X}$ be a connected smooth complex variety and X be an r-puncture of $\overline{X}$, where $r \geq 0$. Then for any point P of X, the following statements hold:*

(a) *There exist spectral sequences of S_n-mixed Hodge structures $E_1^{i,j}(X-P,n) \implies H^{i-j}(F(X-P,n);\mathbb{Q})$ and $E_1^{i,j}(X,n) \implies H^{i-j}(F(X,n);\mathbb{Q})$ with explicit description of the first page (see Section 3) for all n, such that there exists an isomorphism (see Lemma 3.1) of S_n-mixed Hodge structures*

$$\Phi : E_1^{i,j}(X,n) \oplus \operatorname{Ind}_{S_{n-1}}^{S_n} E_1^{i-2d,j-1}(X-P,n-1)(-d) \to E_1^{i,j}(X-P,n), \tag{1.20}$$

where $(-d)$ denotes the Tate twist. The construction of both sequences $E(X-P,n)$ and $E(X,n)$ depend on $\overline{X}$ despite the notation.

(b) *If $\overline{X}$ is noncompact, then Φ commutes with the first-page differential map $d_1^{i,j} : E_1^{i,j} \to E_1^{i,j-1}$.*

(c) *If $\overline{X}$ is pure of slope λ, then both spectral sequences $E(X-P,n)$ and $E(X,n)$ degenerate on E_2, namely, all higher-page differentials d_r ($r \geq 2$) vanish.*

We point out that the important assumptions of Theorem 1.3 about $\overline{X}$ result from the assumptions of parts (b) and (c) of Theorem 1.5.

1.6 Organization of the paper

The rest of the paper will be devoted to the proof of Theorem 1.4. In Section 2, we describe a general spectral sequence (Proposition 2.3) used in the construction of $E(X-P,n)$ and $E(X,n)$ in Theorem 1.5. In Section 3, we construct the isomorphism Φ in Theorem 1.5 explicitly, and prove each part of Theorem 1.5 using explicit computations. In Section 4, we prove Theorem 1.4 from Theorem 1.5, and thus conclude the proof of all of the main results. In Section 5, we discuss the possibilities and difficulties of further generalizing Theorem 1.3.

Contents

2 Description of the spectral sequences

2.1 Arrangements and Orlik–Solomon algebra

In order to construct the spectral sequences in Theorem 1.5, we give an introduction of necessary preliminaries about hyperplane arrangements and the Orlik–Solomon algebra. For a reference, see for instance [OT92].

A *hyperplane arrangement* is a set $\mathcal{A} := \{Y_1, \ldots, Y_h\}$ of complex hyperplanes in $\mathbb{C}^n$. A *stratum* F of $\mathcal{A}$ is the intersection of zero or more hyperplanes in $\mathcal{A}$ (so $\mathbb{C}^n$ is also a stratum). The *lattice* $L(\mathcal{A})$ of $\mathcal{A}$ is the partially ordered set (poset) of strata of $\mathcal{A}$, ordered by inclusion. It has a top element $\hat{1} := \mathbb{C}^n$. In the rest of this section, all of the constructions associated to $\mathcal{A}$ are determined by the poset $L(\mathcal{A})$ alone.

The *rank* of a stratum F is its complex codimension in $\mathbb{C}^n$, or equivalently, the length of any maximal chain from F to $\hat{1}$. A subset S of $\mathcal{A}$ is called (1) *independent*, if $\bigcap S := \bigcap_{F\in S} F \neq \varnothing$ and $\mathrm{rk}(\bigcap S) = |S|$; (2) *dependent*, if $\bigcap S \neq \varnothing$ and $\mathrm{rk}(\bigcap S) < |S|$; (3) *vanishing*, if $\bigcap S = \varnothing$.

Define the algebra $B(\mathcal{A})$ to be the free graded-commutative $\mathbb{Z}$-algebra (i.e. the exterior algebra) generated by degree-one elements $e_Y, Y \in \mathcal{A}$. Let $\partial : B(\mathcal{A}) \to B(\mathcal{A})$ be the unique $\mathbb{Z}$-linear map on $B(\mathcal{A})$ that satisfies the graded Leibniz rule for derivations and $\partial e_Y = 1$ for $Y \in \mathcal{A}$. It has the property that $\partial^2 = 0$. For an ordered subset S of $\mathcal{A}$, we denote $e_S := \prod_{Y\in S} e_Y$, with $e_\varnothing = 1$ and $\bigcap \varnothing = X$ by convention. There is an important observation ([OT92, Lemma 3.7]) for any nonempty S that e_S is always divisible by $\partial(e_S)$ in $A(\mathcal{A})$. Define $I(\mathcal{A})$ to be the ideal of $B(\mathcal{A})$ generated by e_S for S vanishing and $\partial(e_S)$ for S dependent. It turns out (as a consequence of the important observation above) that $I(\mathcal{A})$ is generated by e_S for minimal vanishing sets S and $\partial(e_S)$ for minimal dependent sets S. Define the *Orlik–Solomon algebra* of

$\mathcal{A}$ (or of $L(\mathcal{A})$) as $A(\mathcal{A}) := B(\mathcal{A})/I(\mathcal{A})$. Denote by g_Y the image of $e_Y \in B(\mathcal{A})$ in $A(\mathcal{A})$ under the quotient map. The differential ∂ on $B(\mathcal{A})$ descends to a differential $\partial : A(\mathcal{A}) \to A(\mathcal{A})$, making $A(\mathcal{A})$ a differential graded algebra.

More concretely, $A(\mathcal{A})$ is the graded-commutative $\mathbb{Z}$-algebra generated by degree-one elements g_Y, $Y \in \mathcal{A}$ with the following relations:

(a) $g_S = 0$ if S is vanishing;

(b) If $S = \{Y_{i_1}, \ldots, Y_{i_k}\}$ is dependent, then

$$\sum_{j=1}^{k} (-1)^{j+1} g_{Y_{i_1}} \ldots \widehat{g}_{Y_{i_j}} \ldots g_{Y_{i_k}} = 0 \tag{2.1}$$

where the notation $\widehat{g}_{Y_{i_j}}$ means skipping j-th factor in the product. Note that relation (b) implies that $g_S = 0$ for S dependent (again due to the observation [OT92, Lemma 3.7]).

The *complement* of $\mathcal{A}$ is the complex variety $M(\mathcal{A}) := \mathbb{A}^n - \bigcup_{Y \in \mathcal{A}} Y$. The cohomology ring $H^*(M(\mathcal{A})) := \bigoplus_{i=0}^{\infty} H^i(M(\mathcal{A}), \mathbb{Z})$ is isomorphic to the graded algebra $A(\mathcal{A})$.

Let F be a stratum of $\mathcal{A}$. Consider the subarrangement $\mathcal{A}_F := \{Y \in \mathcal{A} : Y \supseteq F\}$, then we have $M(\mathcal{A}) \subseteq M(\mathcal{A}_F)$, and the pullback of the inclusion map gives a morphism of graded algebras $H^*(M(\mathcal{A}_F)) \to H^*(M(\mathcal{A}))$. Via the identification above, this is the ring homomorphism $A(\mathcal{A}_F) \to A(\mathcal{A})$ that sends g_Y to g_Y. If $F' \subseteq F$ are two strata of $\mathcal{A}$, then we have a natural map $A(\mathcal{A}_F) \to A(\mathcal{A}_{F'})$.

Let $A_F(\mathcal{A})$ be the abelian subgroup of $A(\mathcal{A})$ generated by g_S with $\bigcap S = F$. Then $A(\mathcal{A}) = \bigoplus_F A_F(\mathcal{A})$, where F ranges over all strata of $\mathcal{A}$. Also, $A_F(\mathcal{A})$ is the image of $A(\mathcal{A}_F)_{\operatorname{rk} F} \to A(\mathcal{A})_{\operatorname{rk} F}$, where $(\cdot)_d$ is taking the degree-d part.

2.2 Lattice spectral sequence

Consider a smooth complex variety V and a collection of smooth d-codimensional closed subvarieties $\mathcal{A} = \{Y_1, \ldots, Y_h\}$. We say that $Y_1, \ldots, Y_h$ intersect *like a hyperplane arrangement* if

(a) For any $S \subseteq \mathcal{A}$, the intersection $\bigcap S$ (if nonempty) is smooth and connected. Call such an intersection a *stratum* of $\mathcal{A}$.

(b) Every stratum has codimension a multiple of d.

(c) The poset of strata of $\mathcal{A}$ is isomorphic to the lattice of some hyperplane arrangement of a complex vector space. Denote this poset by $L(\mathcal{A})$, also called the *lattice* of $\mathcal{A}$.

(d) The set $\mathcal{A}$ is precisely the set of rank one strata in $L(\mathcal{A})$, and every stratum $F \in L(\mathcal{A})$ satisfies $\operatorname{codim}(F) = d \cdot \operatorname{rk}(F)$. (Recall the *rank* of a stratum F is the length of any maximal chain from F to $\hat{1} = V$.)

(The last two statements are actually redundant as they are consequences of the first two.)

We define $A(\mathcal{A})$ to be the Orlik–Solomon algebra associated to the lattice $L(\mathcal{A})$. Denote $M(\mathcal{A}) := V - \bigcup_{i=1}^{h} Y_i$.

Tosteson [Tos16] describes a spectral sequence converging to $H^*(M(\mathcal{A}), \mathbb{Z})$ that works in a general setting. In the special case we describe above, the spectral sequence can be vastly simplified into a form similar to [Dup16, Theorem 3.1].

Proposition 2.1. *Let V and $\mathcal{A}$ be as above. Then there is a spectral sequence of integral-coefficient mixed Hodge structures $E_1^{i,j}(\mathcal{A}) \implies H^{i-j}(M(\mathcal{A}), \mathbb{Z})$ such that the abelian group $E_1(\mathcal{A}) := \bigoplus_{i,j\geq 0} E_1^{i,j}(\mathcal{A})$ bigraded by (i, j) is given by*

$$E_1(\mathcal{A}) = \bigoplus_{F\in L(\mathcal{A})} H^*(F, \mathbb{Z}) \otimes A_F(\mathcal{A}) \tag{2.2}$$

where $H^i(F, \mathbb{Z})$ has bidegree $(i, 0)$ and $A_F(\mathcal{A})$ has bidegree $(2\operatorname{codim}_{\mathbb{C}} F, \operatorname{rk} F)$ and Hodge type $(\operatorname{codim}_{\mathbb{C}} F, \operatorname{codim}_{\mathbb{C}} F)$.

The group $E_1(\mathcal{A})$ has a structure of graded-commutative algebra[1] *induced from the algebra structure of $A(\mathcal{A}) = \bigoplus_F A_F(\mathcal{A})$ and the cup product on $H^*(F, \mathbb{Z})$.*

The first-page differential map $d_1^{i,j} : E_1^{i,j} \to E_1^{i,j-1}$ is given by the Gysin map on the cohomology and the differential ∂ on $A(\mathcal{A})$. It makes $E_1(\mathcal{A})$ a differential graded algebra[2]*.*

The spectral sequence is functorial in automorphisms of V that preserve $\mathcal{A}$.

Remark 2.2. In the case where $\mathcal{A}$ is the big diagonal arrangement that gives the ordered configuration space $F(X, n)$ of a d-dimensional variety X, the spectral sequence here is the same as [Tot96] but with a degree shifting, so that $E_1^{2d,1}$ here corresponds to Totaro's $E_{2d}^{0,2d-1}$. The spectral sequence in Proposition 2.1 can also be viewed as a high-codimensional analogue of the Orlik–Solomon spectral sequence in [Loo93, §2] and [Dup15].

Proof. One can reuse Totaro's argument in [Tot96] based on the result [GM88, pp. 237–239] of Goresky and MacPherson about arrangements of k-codimensional subspaces of $\mathbb{R}^n$ whose all strata have codimension a multiple of k (see the Remark after the proof of Lemma 3 in [Tot96] for a discussion). This recognizes $E_1(\mathcal{A})$ as the E_{2d} page of the Leray spectral sequence of $F(X_r, n) \hookrightarrow X^n$, where a dga structure is present. See also [MP20, Lemma 3.1].

Alternatively, one can use the spectral sequence described in [Tos16, Theorem 1.8]:

$$E_1^{i,j}(\mathcal{A}) = \bigoplus_{F\in L(\mathcal{A})} \widetilde{H}_{j-2}((F, \hat{1}); H^i(V, V - F; \mathbb{Z})) \tag{2.3}$$

where $\widetilde{H}_{j-2}((F, \hat{1}))$ is the reduced homology of the order complex of the poset $(F, \hat{1})$, with a special convention when $F = \hat{1}$. We refer the reader to [Wac07] to a detailed account for these concepts.

[1] The direct summand $E_1^{i,j}$ is assigned the degree $i - j$.

[2] A *differential graded algebra* (dga) is a graded-commutative algebra equipped with a linear map d of degree 1 (namely, sending a degree-k element to a degree-$(k+1)$ element), called the differential map, such that $d \circ d = 0$ and the graded Leibniz rule is satisfied.

Since $L(\mathcal{A})$ is isomorphic to the lattice of a hyperplane arrangement, we have

$$\widetilde{H}_{j-2}((F,\hat{1});\mathbb{Z}) = \begin{cases} \mathbb{Z}^{|\mu(F,\hat{1})|} = A_F(\mathcal{A}), & j = \operatorname{rk} F \\ 0 & j \neq \operatorname{rk} F \end{cases} \tag{2.4}$$

for all $F \in L(\mathcal{A})$ (including $F = \hat{1}$), where μ is the Möbius function of the lattice $L(\mathcal{A})$ (see [Fol66, Theorem 4.1] and [OT92, §4.5]).

By the tubular neighborhood theorem and the excision theorem, we have $H^i(V, V - F;\mathbb{Z}) \cong H^i(\mathcal{N}_F, \mathcal{N}_F - F;\mathbb{Z})$, where $\mathcal{N}_F$ is the normal bundle of F in V. Since we are considering complex manifolds, the normal bundle has a canonical orientation, which gives a canonical Thom isomorphism

$$H^i(V, V - F;\mathbb{Z}) \cong H^{i-2\operatorname{codim}_{\mathbb{C}} F}(F,\mathbb{Z}). \tag{2.5}$$

Combining the above, we get $E_1^{i,j}(\mathcal{A}) = \bigoplus_{\operatorname{rk} F = j} H^{i-2\operatorname{codim}_{\mathbb{C}} F}(F,\mathbb{Z}) \otimes A_F(\mathcal{A})$ as required. For the compatibility with the mixed Hodge structure, see [Pet17, §3.2]. The structure of differential graded algebra on $E_1(\mathcal{A})$ can be constructed using the functoriality of Tosteson's spectral sequence along the diagonal map $V \to V \times V$. □

2.3 Arrangements arising from punctured varieties

In this section, we provide a simplified description of the spectral sequence in the following special setting. Consider a connected smooth complex variety X of dimension d and distinct points $P^1, \ldots, P^r$ ($r \geq 1$) of X. (We use superscripts for the points for future convenience.) Fix $n \geq 0$, and consider the arrangement $\mathcal{A}$ of X^n consisting of the following d-codimensional closed subvarieties:

$$\Delta_{ij} := \{(x_1, \ldots, x_n) \in X^n : x_i = x_j\}$$
$$\Delta_i^s := \{(x_1, \ldots, x_n) \in X^n : x_i = P^s\}$$

for $1 \leq i \neq j \leq n$ and $1 \leq s \leq r$. Then the complement $M(\mathcal{A})$ of the arrangement $\mathcal{A}$ is the ordered configuration space $F(X_r, n)$, where $X_r := X - \{P^1, \ldots, P^r\}$.

The goal of this section is Proposition 2.3, an explicit description of the differential graded algebra $E_1(\mathcal{A}) \otimes \mathbb{Q}$ in terms of generators and relations. In the rest of this section, *we assume every cohomology is in rational coefficients*.

We denote by $p_i : X^n \to X$ the projection map onto the i-th coordinate, and denote $p_i^*(\alpha)$ by α_i. Since we are working in rational coefficients, by Künneth's formula $H^*(X^n) \cong H^*(X)^{\otimes n}$, the cohomology ring $H^*(X^n)$ is generated by elements of the form α_i where $\alpha \in H^*(X)$ and $1 \leq i \leq n$. For $i \neq j$, let $p_{ij} : X^n \to X^2$ be the projection map $(x_1, ..., x_n) \mapsto (x_i, x_j)$. Let Δ be the diagonal of X^2. In the notation of above, we have

$$p_{ij}^{-1}(\Delta) = \Delta_{ij} \tag{2.6}$$
$$p_i^{-1}(P^s) = \Delta_i^s \tag{2.7}$$

To a smooth irreducible d-codimensional closed subvariety Z of any smooth variety Y, we associate a cohomology class $[Z] \in H^{2d}(Y)$ given by the image of the canonical generator

$1 \in H^0(Z)$ under the Gysin map $H^*(Z) \to H^{*+2d}(Y)$. One way to define the Gysin map is using the Poincaré dual $H^i(Z) = H_c^{2(n-d)-i}(Z)^\vee \to H_c^{2n-(i+2d)}(Y)^\vee = H^{i+2d}(Y)$ of the pullback map $H^{2n-(i+2d)}(Y) \to H^{2(n-d)-i}(Z)$. The class of a closed subvariety satisfies the following property: if $f : X \to Y$ is a flat map of constant relative dimension, and D is an algebraic cycle of Y, then $f^*[D] = [f^{-1}(D)]$, where $f^{-1}(D)$ is the pullback.

Define $E_1(X_r, n) := E_1(\mathcal{A}) \otimes \mathbb{Q}$. We point out that the definition of $E_1(X_r, n)$ depends on X together with $\{P^1, \ldots, P^r\}$ despite the notation.

Proposition 2.3. *Let $\mathcal{A}$ be the arrangement above. Then the differential graded algebra $E_1(X_r, n)$ is given by*

$$E_1(X_r, n) := \frac{H^*(X^n; \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{Q}[g_{ij}, g_i^s : 1 \le i \ne j \le n, 1 \le s \le r]}{\text{(relations)}}$$

with $H^k(X^n)$ having bidegree $(k, 0)$, and g_{ij} and g_i^s having bidegree $(2d, 1)$ and Hodge type (d, d), subject to relations given by

$$g_{ij} = g_{ji} \tag{2.8}$$

$$g_{ij}g_{jk} + g_{jk}g_{ki} + g_{ki}g_{ij} = 0 \text{ for } i, j, k \text{ distinct} \tag{2.9}$$

$$g_{ij}\alpha_i = g_{ij}\alpha_j \text{ for } \alpha \in H^*(X) \tag{2.10}$$

$$g_i^s\alpha_i = 0 \text{ for } \alpha \in H^{\ge 1}(X) := \bigoplus_{p \ge 1} H^p(X) \tag{2.11}$$

$$g_i^s g_j^s - g_{ij}g_i^s + g_{ij}g_j^s = 0 \tag{2.12}$$

$$g_i^s g_i^t = 0 \text{ for } s \ne t \tag{2.13}$$

The differential $d := d_1$ on the E_1 page is determined by

$$dg_{ij} = p_{ij}^*[\Delta] \tag{2.14}$$

$$dg_i^s = [\Delta_i^s] \tag{2.15}$$

$$d|_{H^*(X^n)} = 0 \tag{2.16}$$

The symmetric group S_n acts on $E_1(X_r, n)$ by permuting subscripts.

We note that the notation $\mathbb{Q}[g_{ij}, g_i^s]$ above does not denote a polynomial ring, but an exterior algebra, since the generators have degree 1.

2.4 Proof of Proposition 2.3

The statement and the proof idea of Proposition 2.3 is similar to Bibby's [Bib16, Theorem 4.1], but we present a proof in full detail here due to the lack of a direct reference.

A stratum F of $\mathcal{A}$ can be uniquely indexed by a pair $(\chi, \sim)$ of a *coloring* function $\chi : \{1, \ldots, n\} \to \{0, \ldots, r\}$ and an equivalence relation $\sim$ on $\chi^{-1}(0)$, according to the rule

$$F_{(\chi,\sim)} = \{(x_1, \ldots, x_n) \in X^n : x_i = P^{\chi(i)} \text{ if } \chi(i) \ne 0, \tag{2.17}$$

$$\text{and } x_i = x_j \text{ if } \chi(i) = \chi(j) = 0 \text{ and } i \sim j. \} \tag{2.18}$$

In other words, a coordinate that is colored s ($1 \le s \le r$) is required to take P^s as value, and the coordinates colored 0 have no such a requirement, but they must agree if they belong to the same block of the partition given by $\sim$. It is helpful to think of color 0 as "uncolored".

The arrangement $\mathcal{A}$ satisfies the following key property. Recall that every cohomology is in rational coefficients unless otherwise notated.

Lemma 2.4. *Let $F = F_{(\chi,\sim)}$ be a stratum of $\mathcal{A}$ described above. Then the pullback map $H^*(X^n) \to H^*(F)$ is surjective. Moreover, its kernel is the ideal generated by $p_i^*\alpha$ for $\alpha \in H^{\ge 1}(X), \chi(i) \neq 0$, and $p_i^*\alpha - p_j^*\alpha$ for $\alpha \in H^*(X), \chi(i) = \chi(j) = 0, i \sim j$.*

Proof. By Künneth's formula, we can deal with each i with $\chi(i) \neq 0$ and each equivalence class of $\chi^{-1}(0)$ separately. It suffices to prove the lemma for the following cases:

(a) $n = 1, F = \Delta_1^s$, which is the point P^s. Then since $H^*(F) = H^0(F)$ and both X and F are connected, the kernel of $H^*(X) \to H^*(F)$ is $H^{\ge 1}(X)$.

(b) $n \ge 2, F = \{(x_1, \dots, x_n) : x_1 = \cdots = x_n\}$. Then $F \hookrightarrow X^n$ is isomorphic to the diagonal map $X \hookrightarrow X^n, x \mapsto (x, \dots, x)$, so the kernel of $H^*(X^n) \to H^*(F)$ is the same as the kernel of the multiplication map $H^*(X)^{\otimes n} \to H^*(X)$.

We recall a standard fact in commutative algebra, but now we state and prove an extension in the following setting. Let R be a commutative ring and A a graded-commutative R-algebra. The R-module $A^{\otimes n} := A \otimes_R \cdots \otimes_R A$ has a structure of graded-commutative algebra that satisfies

- For $1 \le i \le n$, we have a degree-preserving algebra homomorphism $\theta_i : A \to A^{\otimes n}$ sending $a \in A$ to $1 \otimes \cdots \otimes a \otimes \cdots \otimes 1$ where a appears at the i-th factor;
- $a_1 \otimes \cdots \otimes a_n = \theta_1(a_1) \dots \theta_n(a_n)$.

We claim that the kernel of the multiplication map $\mu_n : A^{\otimes n} \to A, a_1 \otimes \cdots \otimes a_n \mapsto a_1 \dots a_n$ is generated by $\theta_i(a) - \theta_j(a)$ for all $1 \le i, j \le n$ and $a \in A$. Lemma 2.4 then follows from the claim by letting $A = H^*(X)$ and $R = \mathbb{Q}$.

It is obvious that $\theta_i(a) - \theta_j(a)$ is in the kernel. We shall prove the reverse inclusion by induction on n.

(i) Case $n = 2$. An element of $\ker \mu_2$ is of the form

$$\sum_{j=1}^{h} a_j \otimes b_j = \sum_{j=1}^{h} \theta_1(a_j)\theta_2(b_j)$$

such that $\sum a_j b_j = 0$ in A.
Then

$$\begin{aligned} \sum a_j \otimes b_j &= \sum \theta_1(a_j)\theta_2(b_j) - \theta_1(\sum a_j b_j) && (2.19)\\ &= \sum (\theta_1(a_j)\theta_2(b_j) - \theta_1(a_j)\theta_1(b_j)) && (2.20)\\ &= \sum \theta_1(a_j)(\theta_2(b_j) - \theta_1(b_j)) && (2.21) \end{aligned}$$

is in the ideal generated by $\theta_2(b_j) - \theta_1(b_j)$.

(ii) Case $n > 2$. Decompose μ_n into two maps:

$$A^{\otimes(n-1)} \otimes_R A \xrightarrow{\mu_{n-1}\otimes 1} A \otimes_R A \xrightarrow{\mu_2} A \tag{2.22}$$

If $x \in \ker(\mu_n)$, then $(\mu_{n-1} \otimes 1)(x)$ must be in the kernel of μ_2. By the $n = 2$ case above, the kernel of μ_n is generated by preimages of $1 \otimes a - a \otimes 1, a \in A$ under $\mu_{n-1} \otimes 1$. One of its preimages is $\theta_n(a) - \theta_{n-1}(a)$, and all other preimages must differ from this one by an element of $\ker(\mu_{n-1} \otimes 1) = \ker(\mu_{n-1})\theta_n(A)$. By the induction hypothesis, the kernel of μ_n is contained in the ideal generated by $\theta_n(a) - \theta_{n-1}(a)$ and $\theta_i(a) - \theta_j(a), 1 \leq i, j \leq n-1, a \in A$.

□

Lemma 2.4 allows a simplication of Proposition 2.1.

Lemma 2.5. *Let $\mathcal{A}$ be an arrangement of d-codimensional subvarieties of a smooth variety V that intersect like a hyperplane arrangement. Assume in addition that for any stratum F of $\mathcal{A}$, the pullback map $H^*(V) \to H^*(F)$ is surjective with kernel I_F. Then*

$$E_1(\mathcal{A}) \otimes \mathbb{Q} \cong \frac{H^*(V) \otimes A(\mathcal{A})}{(I_F \cdot A_F(\mathcal{A}) : F \in L(\mathcal{A}))} \tag{2.23}$$

Proof. The graded algebra $E_1(\mathcal{A}) \otimes \mathbb{Q}$ in Proposition 2.1 is given by

$$\begin{aligned}
\bigoplus_F H^*(F) \otimes A_F(\mathcal{A}) &= \bigoplus_F \frac{H^*(V)}{I_F} \otimes A_F(\mathcal{A}) \\
&= \frac{\bigoplus_F H^*(V) \otimes A_F(\mathcal{A})}{\sum_F I_F \cdot A_F(\mathcal{A})} \\
&= \frac{H^*(V) \otimes \bigoplus_F A_F(\mathcal{A})}{\sum_F I_F \cdot A_F(\mathcal{A})} \\
&= \frac{H^*(V) \otimes A(\mathcal{A})}{\sum_F I_F \cdot A_F(\mathcal{A})}
\end{aligned}$$

□

We are now ready to compute $E_1(X_r, n)$. We need the following technical lemma.

Lemma 2.6. *Let $\mathcal{A}$ be a hyperplane arrangement. Let $S_1, S_2 \subseteq \mathcal{A}$ be disjoint subsets and $Y \in \mathcal{A}$ be an element not in $S_1 \sqcup S_2$. Then inside the algebra $B(\mathcal{A})$, we have*

$$\partial e_{S_1 \sqcup S_2} \in (\partial e_{S_1 \sqcup \{Y\}}, \partial e_{S_2 \sqcup \{Y\}}) B(\mathcal{A})$$

Proof. Write $A = e_{S_1}$, $B = e_{S_2}$ and $e = e_Y$. We have

$$\begin{aligned}
\partial(e\partial(AB)) &= (\partial e)(\partial(AB)) - e\partial^2(AB) && (2.24)\\
&= 1 \cdot \partial(AB) - 0 && (2.25)\\
&= \partial(AB) && (2.26)
\end{aligned}$$

Thus

$$\partial(AB) = \partial(e\partial(AB)) \tag{2.27}$$
$$= \partial(e((\partial A)B \pm A\partial B)) \tag{2.28}$$
$$= \pm\partial(eB\partial A) \pm \partial(eA\partial B) \tag{2.29}$$
$$= \pm\partial(eB)\partial(A) \pm \partial(eA)\partial B \text{ (since } \partial^2 = 0) \tag{2.30}$$
$$\in (\partial(eA), \partial(eB)) \tag{2.31}$$

□

Proof of Proposition 2.3. By Lemma 2.5, for the arrangement $\mathcal{A} := \{\Delta_{ij}, \Delta_i^s : 1 \le i \ne j \le n, 1 \le s \le r\}$, the algebra $E_1(\mathcal{A})_{\mathbb{Q}}$ is given by

$$E_1(X_r, n) = \frac{H^*(X^n) \otimes A(\mathcal{A})}{(I_F \cdot A_F(\mathcal{A}) : F)}$$
$$= \frac{H^*(X^n) \otimes A(\mathcal{A})}{(I_F g_S : \bigcap S = F \text{ and } S \text{ is independent})}$$

We denote the generators of $A(\mathcal{A})$ by $g_{ij} = g_{\Delta_{ij}}, g_i^s = g_{\Delta_i^s}$ (so that $g_{ij} = g_{ji}$ are the same generator). We work out the relation ideal $(I_F g_S : \bigcap S = F$ and S is independent$)$ first. We claim that it is enough to use F of rank one. In other words, the relation ideal is equal to the ideal J generated by

(a) $g_{ij}(\alpha_i - \alpha_j)$;

(b) $g_i^s \alpha_i, \alpha \in H^{\ge 1}(X)$.

Let $F = F_{\chi,\sim}$ be a stratum of $\mathcal{A}$. We need to show that for any independent $S \subset \mathcal{A}$ such that $\bigcap S = F$, the ideal $g_S I_F$ is in the ideal J.

Such an S is classified by the following indirected graph, consisting of

(a) An (unrooted) spanning tree on each equivalence class of $\sim$ on $\chi^{-1}(0)$;

(b) A forest of rooted trees on $\chi^{-1}(s)$, for each $1 \le s \le r$.

The set S then consists of $\Delta_i^{\chi(i)}$ for each i that appears as a root and Δ_{ij} for each (i,j) that appears as an edge.

We observe that to generate $I_F = (\alpha_i - \alpha_j\ (i \sim j), \alpha_i\ (\chi(i) \ne 0))$, a part of the generators suffices: $\alpha_i - \alpha_j$ for $\Delta_{ij} \in S$ and α_i for $\Delta_i^s \in S, \alpha \in H^{\ge 1}(X)$. Indeed, this can be done by joining a path from i to j in the tree (if $i \sim j \in \chi^{-1}(0)$) or by joining a path from i to the root of the tree where i belongs (if $\chi(i) \ne 0$). But g_S multiplied by each of these special generators lies in J. This proves the claim and finishes the computation of $(I_F \cdot A_F(\mathcal{A}) : F \in L(\mathcal{A}))$.

It remains to compute a presentation of $A(\mathcal{A})$. Let $J(\mathcal{A})$ be the ideal of $B(\mathcal{A})$ generated by relations (2.9), (2.12) and (2.13) of Proposition 2.3. Claim that $J(\mathcal{A}) = I(\mathcal{A})$, the defining ideal for $A(\mathcal{A})$.

The ideal $I(\mathcal{A})$ is generated by e_S for minimal vanishing set e_S and ∂e_S for minimal dependent set S. These include

(1) $e_i^s e_j^t e_\gamma$, $s \neq t$;

(2) $\partial(e_{i_1 i_2} e_{i_2 i_3} \dots e_{i_{h-1} i_h} e_{i_h i_1})$, with $h \geq 3$ and $i_1, \dots, i_h$ are distinct;

(3) $\partial(e_i^s e_j^s e_\gamma)$,

where in both (1) and (3), $\gamma = (i = i_0 \to i_1 \to \cdots \to i_h = j)$ is a path joining i and j, and e_γ means $e_{i_0 i_1} e_{i_1 i_2} \dots e_{i_{h-1} i_h}$. Here h is allowed to be 0, in which case $i = j$ and $e_\gamma = 1$.

We need to show that these generators are in $J(\mathcal{A})$.

First, we prove that (2) is in $J(\mathcal{A})$ by induction on h. If $h = 3$, then (2) is just (2.9). If $h > 3$, we set $S_1 = \{\Delta_{i_1 i_2}, \Delta_{i_2 i_3}, \dots, \Delta_{i_{h-2} i_{h-1}}\}, S_2 = \{\Delta_{i_{h-1} i_h}, \Delta_{i_h i_1}\}$ and $Y = \Delta_{i_1 i_{h-1}}$, then $\partial e_{S_1 \sqcup Y} \in J(\mathcal{A})$ by induction hypothesis, $\partial e_{S_2 \sqcup Y} \in J(\mathcal{A})$ by $h = 3$ case. Applying Lemma 2.6, we get $\partial e_{S_1 \sqcup S_2} \in J(\mathcal{A})$.

Next, we prove that (3) is in $J(\mathcal{A})$ by induction on h, the length of $\gamma = (i = i_0 \to i_1 \to \cdots \to i_h = j)$. The base case $h = 1$ is (2.12), and the induction step is proved similarly with $S_1 = \{\Delta_i^2, \Delta_{i_0 i_1}, \Delta_{i_1 i_2}, \dots, \Delta_{i_{h-2} i_{h-1}}\}, S_2 = \{\Delta_{i_{h-1} i_h}, \Delta_j^s\}$ and $Y = \Delta_{i_{h-1}}^s$.

Finally, for (1), to prove that $e_i^s e_j^t e_\gamma \in J(\mathcal{A})$ for $s \neq t$, we note that

$$\partial(e_i^t e_j^t e_\gamma) = e_j^t e_\gamma - e_i^t \partial(e_i^t e_\gamma) \tag{2.32}$$

is just proved to be in $J(\mathcal{A})$ by case (3). Hence

$$e_j^t e_\gamma \equiv e_i^t \partial(e_i^t e_\gamma) \mod J(\mathcal{A}), \tag{2.33}$$

so

$$e_i^s e_j^t e_\gamma \equiv e_i^s e_i^t \partial(e_i^t e_\gamma) \mod J(\mathcal{A}). \tag{2.34}$$

But $e_i^s e_i^t$ is just (2.13), so $e_i^s e_j^t e_\gamma \in J(\mathcal{A})$.

In summary, we have proved that

$$\begin{aligned} E_1(X_r, n) &= \frac{H^*(X^n) \otimes A(\mathcal{A})}{(I_F \cdot A_F(\mathcal{A}) : F)} && (2.35) \\ &= \frac{H^*(X^n) \otimes \mathbb{Q}[g_{ij}, g_i^s]/((2.8), (2.9), (2.12), (2.13))}{((2.10), (2.11))} && (2.36) \\ &= \frac{H^*(X^n)[g_{ij}, g_i^s]}{((2.8) \text{ through } (2.13))} && (2.37) \end{aligned}$$

This proves the description of $E_1(X_r, n)$ in Proposition 2.3. $\square$

3 Proof of Theorem 1.5

Let $\overline{X}$ be a connected smooth complex variety, and let $P^1, \dots, P^r$ be distinct points of $\overline{X}$, where $r \geq 1$. Let $X = \overline{X} - \{P^1, \dots, P^{r-1}\}$ and $P = P^r$. Noting that $X = \overline{X}_{r-1}$ and $X - P = \overline{X}_r$ in the notation of Proposition 2.3, we construct spectral sequences

$$E_1(X, n) := E_1(\overline{X}_{r-1}, n) \implies H^{i-j}(F(X, n); \mathbb{Q}) \tag{3.1}$$

and

$$E_1(X - P, n) := E_1(\overline{X}_r, n) \implies H^{i-j}(F(X - P, n); \mathbb{Q}) \tag{3.2}$$

based on Proposition 2.3. We shall prove that these two spectral sequences satisfy the statements of Theorem 1.5.

3.1 Proof of Theorem 1.5(a)

Using the description of $E_1(\overline{X}_r, n)$ and $E_1(\overline{X}_{r-1}, n)$ in Proposition 2.3 as differential graded algebras, we can express $E_1(X - P, n)$ as an algebra over $E_1(X, n)$ according to

$$E_1(X - P, n) = E_1(X, n)[g_1^r, ..., g_n^r]/(\text{new relations}) \tag{3.3}$$

where the new relations consist of

$$g_i^r g_j^r - g_{ij} g_i^r + g_{ij} g_j^r = 0 \text{ for } 1 \leq i, j \leq n \tag{3.4}$$

$$g_i^s g_i^r = 0 \text{ for } 1 \leq s \leq r - 1 \tag{3.5}$$

$$g_i^r \alpha_i = 0 \text{ for } \alpha \in H^p(\overline{X}), p \geq 1 \tag{3.6}$$

We will repetitively use the following elementary fact about quadratic algebras. Let R be a graded-commutative ring with identity, and let $x_1, \dots, x_m$ be indeterminates of degree one. Consider the graded-commutative R-algebra A generated by $x_1, \dots, x_m$ with relations $x_i x_j = L_{ij}(x_1, \dots, x_m)$ for all $1 \leq i < j \leq m$, where $L_{ij}(x_1, \dots, x_m)$ is a left R-linear combination of $x_1, \dots, x_m$. Then A is isomorphic to $R\langle 1, x_1, \dots, x_m \rangle$ as a left R-module, where $R\langle 1, x_1, \dots, x_m \rangle$ denotes the free left R-module with basis $1, x_1, \dots, x_m$.

We denote by $[n]$ the finite set $\{1, \dots, n\}$, and by $[n] - i$ the set $\{j \in [n] : j \neq i\}$. For any finite set I of integers and for $Y = X$ or $X - P$, we denote by $E_1(Y, I)$ a copy of $E_1(Y, |I|)$, but with lower indices of the generators taken from I instead of $\{1, \dots, |I|\}$. Note that $E_1(Y, n)$ and $E_1(Y, [n])$ are precisely the same.

We now express $E_1(X - P, n)$ as a left module over $E_1(X, n)$. Note that the relation (3.4) is equivalent to $g_i^r g_j^r = g_{ij} g_i^r - g_{ij} g_j^r$, the right-hand side being a linear combination of g_i^r, g_j^r with coefficients g_{ij} in the ring $E_1(X, n)$. Thus $E_1(X, n)[g_i^r : i \in [n]]/(3.4)$ is a quadratic algebra over $E_1(X, n)$, and we have

$$\frac{E_1(X, n)[g_i^r : i \in [n]]}{(3.4)} \cong E_1(X, n)\langle 1, g_i^r : i \in [n] \rangle \tag{3.7}$$

as a left $E_1(X,n)$-module, while its ring structure is given by the multiplication table

$$g_i^r g_j^r = g_{ij} g_i^r - g_{ij} g_j^r. \tag{3.8}$$

Therefore,

$$\begin{aligned} E_1(X-P,n) &= \frac{E_1(X,n)[g_i^r : i \in [n]]/(3.4)}{((3.5),(3.6))} && (3.9)\\ &= \frac{E_1(X,n)\langle 1, g_i^r : i\in[n]\rangle}{(g_i^s g_i^r, \alpha_i g_i^r : i \in [n], s \neq r, \alpha \in H^{\geq 1}(X))} && (3.10)\\ &= E_1(X,n) \oplus \bigoplus_{i=1}^n E_1(X,n)\langle g_i^r\rangle/(g_i^s g_i^r, \alpha_i g_i^r : s\neq r) && (3.11)\end{aligned}$$

Our next goal is to further decompose each summand of (3.11). Fix $i \in [n]$, and we shall compute the presentation $E_1(X,n)$ as an algebra over $E_1(X,[n]-i)$. Comparing the presentations of $E_1(X,n)$ and $E_1(X,[n]-i)$ in Proposition 2.3, we get the following presentation, where the subscripts j,k always range over $[n]-i$, the superscripts s,t always range over $\{1,\dots,r-1\}$, and α is taken from $H^{\geq 1}(\overline{X})$.

$$E_1(X,n) = \frac{E_1(X,[n]-i)[g_{ij}, g_i^s] \otimes_{\mathbb{Q}} H^*(X)}{\begin{pmatrix} g_{ij}g_{ik} &=& -g_{jk}g_{ij} + g_{jk}g_{ik}, & j\neq k\\ g_{ij}g_i^s &=& -g_j^s g_i^s - g_j^s g_{ij}, & j,s\\ g_i^s g_i^t &=& 0, & s\neq t\\ \alpha_i g_{ij} &=& \alpha_j g_{ij}, & j,\alpha\\ g_i^s \alpha_i &=& 0, & s,\alpha \end{pmatrix}} \tag{3.12}$$

where $H^*(X)$ in the tensor factor contributes to the i-th coordinate, namely, $\{\alpha_i : \alpha \in H^*(\overline{X})\}$. From now on, every tensor product is over $\mathbb{Q}$.

We note that if we only take the first three relations in (3.12), we get a quadratic algebra over $E_1(X,[n]-i) \otimes H^*(X)$. Thus

$$E_1(X,n) = \frac{\big(E_1(X,[n]-i)\otimes H^*(\overline{X})\big)\langle 1, g_{ij}, g_i^s\rangle}{((\alpha_i-\alpha_j)g_{ij}, \alpha_i g_i^s)}. \tag{3.13}$$

Recall that for any α in $H^{\geq 1}(\overline{X})$, the element α_i of $E_1(X,n)$ is understood as the tensor $1\otimes\alpha$ in $E_1(X,[n]-i)\otimes H^*(\overline{X})$, and the element α_j is understood as the tensor $\alpha_j \otimes 1$ in $E_1(X,[n]-i)\otimes H^*(\overline{X})$. Recalling that $H^*(\overline{X}) = \mathbb{Q}\oplus H^{\geq 1}(\overline{X})$, we may decompose the numerator of (3.13) as

$$\begin{aligned}&\big(E_1(X,[n]-i)\otimes H^*(\overline{X})\big)\langle 1, g_{ij}, g_i^s\rangle = E_1(X,[n]-i)\otimes\\ &\left(H^*(\overline{X}) \oplus \bigoplus_j \mathbb{Q}g_{ij} \oplus \bigoplus_j H^{\geq 1}(\overline{X}) g_{ij} \oplus \bigoplus_s \mathbb{Q} g_i^s \oplus \bigoplus_s H^{\geq 1}(\overline{X}) g_i^s\right).\end{aligned} \tag{3.14}$$

The relation $(\alpha_i - \alpha_j)g_{ij} = 0$ in (3.13) identifies a general element $\alpha_i g_{ij}$ of the summand $H^{\geq 1}(\overline{X})g_{ij}$ with the element $\alpha_j g_{ij}$ of $E_1(X, n-[i])\otimes\mathbb{Q}g_{ij}$. As a result, the relation $(\alpha_i-\alpha_j)g_{ij}$

kills the summand $H^{\geq 1}(\overline{X})g_{ij}$ without introducing other identifications. The same argument can be applied to the relation $\alpha_i g_i^s = 0$. It follows that

$$\begin{aligned}
E_1(X,n) &= \frac{\big(E_1(X,[n]-i)\otimes H^*(\overline{X})\big)\langle 1, g_{ij}, g_i^s\rangle}{((\alpha_i-\alpha_j)g_{ij}, \alpha_i g_i^s)} && (3.15)\\
&= E_1(X,[n]-i)\otimes (H^*(\overline{X})\oplus \mathbb{Q}\langle g_{ij}, g_i^s\rangle) && (3.16)\\
&= E_1(X,[n]-i)\otimes (H^{\geq 1}(\overline{X})\oplus \mathbb{Q}\langle 1, g_{ij}, g_i^s\rangle). && (3.17)
\end{aligned}$$

as a module over $E_1(X,[n]-i)$, and the ring structure of $E_1(X,n)$ can be read from this representation with a multiplication table given by the relations in (3.12).

We are now ready to give a presentation of each summand of (3.11). Fix i. In the computation below, the convention for α, j, s, t is as before, and β ranges over $H^{\geq 1}(X)$.

$$\begin{aligned}
\frac{E_1(X,n)\langle g_i^r\rangle}{E_1(X,n)(g_i^s g_i^r, \alpha_i g_i^r : s,\alpha)} &= \frac{E_1(X,[n]-i)\otimes (H^{\geq 1}(\overline{X})\oplus \mathbb{Q}\langle 1, g_{ij}, g_i^s\rangle)}{E_1(X,[n]-i)(\beta_i, 1, g_{ij}, g_i^t)(g_i^s, \alpha_i)}\langle g_i^r\rangle && (3.18)\\
&= \frac{E_1(X,[n]-i)\otimes (H^{\geq 1}(\overline{X})\oplus \mathbb{Q}\langle 1, g_{ij}, g_i^s\rangle)}{E_1(X,[n]-i)\begin{pmatrix} g_i^s\beta_i, & g_i^s, & g_i^s g_{ij}, & g_i^s g_i^t \\ \alpha_i\beta_i, & \alpha_i, & \alpha_i g_{ij}, & \alpha_i g_i^t\end{pmatrix}}\langle g_i^r\rangle && (3.19)\\
&\overset{(3.12)}{=} \frac{E_1(X,[n]-i)\otimes (H^{\geq 1}(\overline{X})\oplus \mathbb{Q}\langle 1, g_{ij}, g_i^s\rangle)}{E_1(X,[n]-i)\begin{pmatrix} 0, & g_i^s, & g_j^s g_i^s + g_j^s g_{ij}, & 0 \\ (\alpha\beta)_i, & \alpha_i, & \alpha_j g_{ij}, & 0\end{pmatrix}}\langle g_i^r\rangle && (3.20)\\
&= \frac{E_1(X,[n]-i)_{\mathbb{Q}}\langle 1, g_{ij} : j\in [n]-i\rangle}{(g_j^s g_{ij}, \alpha_j g_{ij} : j\in[n]-i, s\neq r, \alpha\in H^{\geq 1}(\overline{X}))}\langle g_i^r\rangle && (3.21)\\
&= \frac{E_1(X,[n]-i)\langle g_r, g_{ij}g_r : j\in [n]-i\rangle}{(g_j^s g_{ij}g_r, \alpha_j g_{ij}g_r : j\in[n]-i, s\neq r, \alpha\in H^{\geq 1}(\overline{X}))} && (3.22)
\end{aligned}$$

where we note that the generators g_i^s and the summand $H^{\geq 1}(X)$ are eliminated by the relations. In line (3.22), we view g_r and $g_{ij}g_r$ as two separate formal generators.

We now prove Theorem 1.5(a). The important observation is that the presentation (3.22) is isomorphic to the presentation of $E_1(X-P,[n]-i)$ given by (3.10) applied to the index set $[n]-i$:

$$E_1(X-P,[n]-i) = \frac{E_1(X,[n]-i)\langle 1, g_j^r : j\in[n]-i\rangle}{(g_j^s g_j^r, \alpha_j g_j^r : j\in[n]-i, s\neq r, \alpha\in H^{\geq 1}(\overline{X}))}, \tag{3.23}$$

where the correspondence is given by $g_r \mapsto 1$ and $g_{ij}g_r\mapsto g_j^r$.

Putting the summands of (3.11) together, we finish construction of the required isomorphism Φ, as is stated below.

Lemma 3.1. *There exists a $\mathbb{Q}$-linear isomorphism*

$$\Phi : E_1(X,n)\oplus \bigoplus_{i=1}^n E_1(X-P,[n]-i)\to E_1(X-P,n), \tag{3.24}$$

such that $\Phi|_{E_1(X,n)} : E_1(X,n) \to E_1(X-P,n)$ *is the natural ring homomorphism, and* $\Phi|_{E_1(X-P,[n]-i)}$ *is the* $E_1(X,[n]-i)$*-linear map that sends* 1 *to* g_i^r *and sends* g_j^r *to* $g_{ij}g_i^r$ *for all* $j \in [n]-i$*.*

We note that $\bigoplus_{i=1}^n E_1(X-P,[n]-i)$ is isomorphic to $\mathrm{Ind}_{S_{n-1}}^{S_n} E_1(X-P,n-1)$ as an S_n-module. Indeed, the induction of $E_1(X-P,n-1)$ from S_{n-1} to S_n is the direct sum of images of $E_1(X-P,n-1)$ under cosets of S_{n-1} in S_n. There are n cosets, and for each $i \in [n]$, there is precisely one coset such that every element in it maps $E_1(X-P,n-1)$ onto $E_1(X-P,[n]-i)$. Hence the induction of $E_1(X-P,n-1)$ is the direct sum of $E_1(X-P,[n]-i)$ for $i \in [n]$.

Note that the elements g_i^r and g_{ij} have bidegree $(2d,1)$ and Hodge type (d,d) in the dga $E_1(X-P,n)$. Keeping track of the bidegree shifting and the Tate twist, and noting that S_n acts by permuting the subscripts, we obtain an isomorphism

$$\Phi : E_1^{i,j}(X,n) \oplus \mathrm{Ind}_{S_{n-1}}^{S_n} E_1^{i-2d,j-1}(X-P,n-1)(-d) \to E_1^{i,j}(X-P,n) \tag{3.25}$$

as S_n-mixed Hodge structures, which finishes the proof of Theorem 1.5(a).

3.2 Proof of Theorem 1.5(b)

Now assume that $\overline{X}$ is a d-dimensional smooth complex variety that is not compact. We need to show that Φ constructed above commutes with the differential.

Recall that the restriction of Φ on the i-th summand is given by an $E_1(X,[n]-i)$-linear map

$$\Phi_i : E_1(X-P,[n]-i) \to E_1(X-P,n) \tag{3.26}$$

such that $\Phi_i(1) = g_i^r$ and $\Phi_i(g_j^r) = g_{ij}g_i^r$ for $j \in [n]-i$. It suffices to show that $d(\Phi_i(1)) = \Phi_i(d(1))$ and $d(\Phi_i(g_j^r)) = \Phi_i(d(g_j^r))$.

We claim that as elements of $E_1(X-P,n)$, we have $dg_i^r = 0$ and $d(g_{ij}g_i^r) = 0$ for $j \in [n]-i$.

Since $\overline{X}$ is not compact, the top cohomology $H^{2d}(]bbarX)$ vanishes. We have $dg_i^r = [\Delta_i^r] = p_i^*([P^r])$ in the notation of Section 2.3. But $[P^r]$ lies in $H^{2d}(\overline{X})$ and $H^{2d}(\overline{X}) = 0$, so $dg_i^r = 0$ for all $i \in [n]$.

To compute $d(g_{ij}g_i^r)$, we use the graded Leibniz rule, noting that $dg_i^r = 0$ just obtained above.

$$d(g_{ij}g_i^r) = d(g_{ij})g_i^r - g_{ij}dg_i^r = d(g_{ij})g_i^r = p_{ij}^*[\Delta]g_i^r. \tag{3.27}$$

Note that $[\Delta] \in H^{2d}(\overline{X} \times \overline{X})$, but $H^{2d}(\overline{X}) = 0$, so Künneth's formula implies that $[\Delta] \in \bigoplus_{p=1}^{2d-1} H^p(\overline{X}) \otimes H^{2d-p}(\overline{X})$. In particular, $p_{ij}^*[\Delta]$ can be expressed as a $\mathbb{Q}$-linear combination of terms of the form $\alpha_i\beta_j$, where $\alpha, \beta \in H^{\geq 1}(\overline{X})$. Since $\alpha_i g_i^r = 0$ in $E_1(X-P,n)$ for every $\alpha \in H^{\geq 1}(\overline{X})$ (Relation (2.11)), we see that $d(g_{ij}g_i^r) = 0$ in $E_1(X-P,n)$.

We now verify that Φ commutes with the differential. Since $dg_i^r = 0$, the left-hand side of $d(\Phi_i(1)) = \Phi_i(d(1))$ is zero, verifying the equality. Note that we have proved $dg_i^r = 0$ in $E_1(X-P,[n])$ for *all* $i \in [n]$. Applying this fact to the index set $[n]-i$, we have $dg_j^r = 0$ for $j \in [n]-i$. Hence, the right-hand side of $d(\Phi_i(g_j^r)) = \Phi_i(d(g_j^r))$ is zero. But the left-hand side is $d(g_{ij}g_i^r)$, which we have computed to be zero as well. This finishes the verification of the equalities, and hence the proof of Theorem 1.5(b).

3.3 Proof of Theorem 1.5(c)

Recall that the spectral sequences $E(X-P,n)$ and $E(X,n)$ are constructed as $E(\overline{X}_r,n)$ and $E(\overline{X}_{r-1},n)$, respectively, in Proposition 2.3. Therefore, it suffices to show that for any connected smooth complex variety X that is pure of slope λ and any integer $r \geq 0$, the spectral sequence $E(X_r,n)$ degenerates on E_2. (For convenience purposes, we use X in place of $\overline{X}$.)

For any $k,l \geq 0$, the bidegree-(k,l) part $E_1^{k,l}(X_r,n)$ of $E_1(X_r,n)$ is generated by the product of l Orlik–Solomon generators (denoted g_{ij} and g_i^s in Proposition 2.3, each with bidegree $(2d,1)$ and Hodge type (d,d)) and an element of $H^{k-2l}(X^n)$. By Kúnneth's formula, X^n is pure of slope λ as well, so $H^{k-2l}(X^n)$ is pure of weight $\lambda(k-2l)$. Each of the l Orlik–Solomon generators has a weight of $d+d=2d$, so that

$$E_1^{k,l}(X_r,n) \text{ is pure of weight } \lambda(k-2l)+l(2d). \tag{3.28}$$

The same purity holds for any higher page $E_h^{k,l}(X_r,n)$ (where $h \geq 2$) as well, since E_h is a subquotient of E_1 from the general definition of spectral sequences.

We now show that the spectral sequence $E(X_r,n)$ degenerates on E_2 for weight reason. Suppose the differential map

$$d_h : E_h^{k,l}(X_r,n) \to E_h^{k-h+1,l-h}(X_r,n) \tag{3.29}$$

is nonzero for some $h \geq 2$ and $k,l \geq 0$. Then $E_h^{k,l}(X_r,n)$ and $E_h^{k-h+1,l-h}(X_r,n)$ must be nonzero pure Hodge structures with equal weight, since the differentials are strictly compatible with the weight filtration. From (3.28), we get

$$\lambda(k-2l)+2dl = \lambda((k-h+1)-2(l-h))+2d(l-h), \tag{3.30}$$

which simplifies to

$$\lambda = \frac{2dh}{1+2dh-h}. \tag{3.31}$$

Since $h>1$, we have

$$\lambda = \frac{2dh}{1+2dh-h} > \frac{2dh}{1+2dh-1} = 1. \tag{3.32}$$

Since X is noncompact, the cohomology of X is concentrated in degrees $0 \leq i \leq 2d-1$, so the denominator of λ as a rational number is at most $2d-1$. It follows that $\lambda \geq \frac{2d}{2d-1}$. However,

$$\lambda = \frac{2dh}{1+2dh-h} < \frac{2dh}{2dh-h} = \frac{2d}{2d-1}, \tag{3.33}$$

a contradiction. This finishes the proof of Theorem 1.5(c).

Remark 3.2. This weight argument is well-known in the case $\lambda = 1$; see for example [Tot96] and [Bib16].

4 Proof of Theorem 1.4

Let $\overline{X}$ and X be as in the assumption of Theorem 1.3. By Theorem 1.5, we have spectral sequences of S_n-mixed Hodge structures $E_1^{i,j}(X-P,n) \implies H^{i-j}(F(X-P,n))$ and $E_1^{i,j}(X,n) \implies H^{i-j}(F(X,n))$ that degenerate on the E_2 page. We have also an isomorphism of S_n-mixed Hodge structures

$$\Phi : E_1^{i,j}(X,n) \oplus \operatorname{Ind}_{S_{n-1}}^{S_n} E_1^{i-2d,j-1}(X-P,n-1)(-d) \to E_1^{i,j}(X-P,n) \tag{4.1}$$

that commutes with the differential map. Because of its commutativity with the differential, the isomorphism Φ descends to an isomorphism on the E_2 page, which is the same as the E_∞ page due to degeneracy. We get an isomorphism of S_n-mixed Hodge structures

$$\Phi_\infty^{i,j} : E_\infty^{i,j}(X,n) \oplus \operatorname{Ind}_{S_{n-1}}^{S_n} E_\infty^{i-2d,j-1}(X-P,n-1)(-d) \to E_\infty^{i,j}(X-P,n). \tag{4.2}$$

Fix $k \in \mathbb{Z}$. The convergence of the spectral sequence $E_1^{i,j}(X,n) \implies H^{i-j}(F(X,n))$ means that there is a filtration of $H^k(F(X,n))$ whose successive quotients consist of $E_\infty^{k+t,t}(X,n)$ for all $t \in \mathbb{Z}$. In particular, if we apply the exact functor gr from the category of (S_n-)mixed Hodge structures to the category of (S_n-)split mixed Hodge structures, defined by taking the associated graded according to the weight filtration, we get a noncanonical isomorphism

$$\operatorname{gr} H^k(F(X,n)) \cong \bigoplus_{t\in\mathbb{Z}} \operatorname{gr} E_\infty^{k+t,t}(X,n), \tag{4.3}$$

since the category of (S_n-)split mixed Hodge structures is semisimple. Similarly for $F(X-P,n)$.

Summing up the isomorphisms $\Phi_\infty^{k+t,t}$ for all $t \in \mathbb{Z}$ and applying the gr functor, we get (working from the right-hand side)

$$\begin{aligned}
\operatorname{gr} H^k(F(X-P,n)) &\cong \bigoplus_{t\in\mathbb{Z}} \operatorname{gr} E_\infty^{k+t,t}(X-P,n) && (4.4)\\
&\cong \bigoplus_{t\in\mathbb{Z}} \Big(\operatorname{gr} E_\infty^{k+t,t}(X,n) \oplus \operatorname{Ind}_{S_{n-1}}^{S_n} \operatorname{gr} E_\infty^{k+t-2d,k+t-1}(X-P,n-1)(-d)\Big) && (4.5)\\
&\cong \operatorname{gr} H^k(F(X,n)) \oplus \operatorname{Ind}_{S_{n-1}}^{S_n} \operatorname{gr} H^{k-2d,k-1}(F(X-P,n-1))(-d). && (4.6)
\end{aligned}$$

By repetitively applying the above isomorphism, or by induction on n, we get

$$\operatorname{gr} H^k(F(X-P,n),\mathbb{Q}) \cong \bigoplus_{t=0}^{\infty} \operatorname{Ind}_{S_{n-t}}^{S_n} \operatorname{gr} H^{k-(2d-1)t}(F(X,n-t),\mathbb{Q})(-d\cdot t), \tag{4.7}$$

finishing the proof of (1.15). By forgetting the mixed Hodge structures and noting that the category of finite-dimensional S_n-representations over $\mathbb{Q}$ is semisimple, we complete the proof of(1.14) and thus Theorem 1.4. As is explained in Section 1.4, Theorem 1.4 implies Theorem 1.1 and Theorem 1.3, thus the proof of all results is complete.

5 Examples and discussions

We discuss several examples of connected smooth noncompact complex varieties that are pure of a certain slope, and therefore satisfy the assumptions for $\overline{X}$ in Theorem 1.3 and 1.4. Recall that a complex variety $\overline{X}$ is pure of slope λ if the mixed Hodge structure on $H^i(\overline{X};\mathbb{Q})$ is pure of weight $\lambda \cdot i$ whenever $\lambda \cdot i$ is an integer, and $H^i(\overline{X};\mathbb{Q})$ is zero whenever $\lambda \cdot i$ is not an integer.

(a) Let $\overline{X}$ be the complement of a hyperplane arrangement in $\mathbb{C}^d$ or a toric arrangement in $(\mathbb{C}^*)^d$. Then $H^i(\overline{X};\mathbb{Q})$ is pure of weight $2i$ by [Dup16, Theorems 3.7, 3.8]. In other words, $\overline{X}$ is pure of slope 2.

(b) Let C be a smooth projective curve of genus g embedded in the projective plane $\mathbb{P}^2$, and let $\overline{X} = \mathbb{P}^2 - C$. We have

$$H^i(X,\mathbb{Q}) = \begin{cases} \mathbb{Q}, \text{pure of weight } 0, & i = 0 \\ \mathbb{Q}^{2g}, \text{pure of weight } 3, & i = 2 \\ 0, & i \neq 0, 2. \end{cases} \tag{5.1}$$

Hence X is pure of weight $3/2$.

(c) Suppose X is smooth and pure of slope λ, and G is a finite group that acts on X freely such that the scheme-theoretic quotient X/G is also smooth. Then $H^i(X;\mathbb{Q})$ is pure of weight $\lambda \cdot i$ for all i. Since $H^i(X/G;\mathbb{Q})$ is the G-invariant of $H^i(X;\mathbb{Q})$, we have that $H^i(X/G;\mathbb{Q})$ is also pure of weight $\lambda \cdot i$. Thus X/G is also pure of slope λ. As an application, the generalized configuration space $F(\mathbb{C}, m)/G$ for a subgroup G of S_m is pure of slope 2.

We formulate the following questions in an attempt to generalize the formula (1.11), and consequently Kallel's (1.7), beyond the already flexible family of examples of X described in Theorem 1.3. We emphasize that both formulas are known to be false if X is compact.

Question 5.1.

(a) Does the conclusion of Theorem 1.1 hold for any connected noncompact smooth complex variety?

(b) If yes, how about the refined statements in Theorem 1.4?

The difficulty of further generalizing the main theorems lies in the Hodge-theoretic hypothesis of Theorem 1.5(c). The isomorphism Φ in Theorem 1.5(a) is constructed "artifitially" based on matching explicit generators and relations. It is unlikely that Φ has an equivalent construction that is functorial, because Φ does not necessarily commute with the differential maps if $\overline{X}$ is compact, as is clear from the proof of Theorem 1.5(b). Therefore, we do not have control of Φ on higher pages of the spectral sequences, so the degeneracy statement in Theorem 1.5(c) is necessary to obtain an isomorphism on the E_∞ page.

We conclude the paper by giving an example of $\overline{X}$ where the hypothesis of Theorem 1.5(c) is false, but it is unknown whether the conclusion of Theorem 1.5(c) or Theorem 1.3 is true.

Example 5.2. Let E be an elliptic curve, and O be a point of E. Consider the smooth surface $\overline{X} = \mathrm{Conf}^2(E - O)$. As is computed in [CH], we have

$$H^i(\overline{X}, \mathbb{Q}) = \begin{cases} \mathbb{Q}, \text{pure of weight } 0, & i = 0 \\ \mathbb{Q}^2, \text{pure of weight } 1, & i = 1 \\ \mathbb{Q}^2, \text{pure of weight } 3, & i = 2 \\ 0, & i \geq 3 \end{cases} \tag{5.2}$$

We observe that when $n \geq 6$, the weight-8 part of $H^7(\overline{X}^n, \mathbb{Q})$, namely, the successive quotient of the weight filtration that is pure of weight 8, is nonzero, since the summand $H^1(\overline{X})^{\otimes 5} \otimes H^2(\overline{X}) \otimes H^0(\overline{X})^{\otimes(n-6)}$ in the Kunneth's formula is in $H^7(\overline{X}^n, \mathbb{Q})$ and is pure of weight $1+1+1+1+1+3 = 8$. If $\overline{X}_r$ is $\overline{X}$ with r punctures ($r \geq 0$), then the second-page differential

$$d_2^{8,2} : E_2^{8,2}(\overline{X}_r, n) \to E_2^{7,0}(\overline{X}_r, n) \tag{5.3}$$

is not yet known to be zero, because the left-hand side is nonzero and pure of weight 8 and the right-hand side is $H^7(X^n, \mathbb{Q})$, which has a nonzero weight-8 part. Regarding the left-hand side, recall that $E_1(\overline{X}_r, n)$ is generated by elements of bidegrees $(k, 0)$ and $(4, 1)$ (noting that $d = 2$) according to Proposition 2.3, so $E_2^{8,2}(\overline{X}_r, n)$ is generated (as a vector space) by products of two Orlik–Solomon generators, each being of bidegree $(4, 1)$ and weight 4.

Acknowledgements

We thank Gilyoung Cheong, Jeffery Lagarias, Mircea Mustaţă, Roberto Pagaria, Dan Petersen, Andrew Snowden, John Stembridge, Burt Totaro and Craig Westerland for helpful conversations and useful comments. We thank Christin Bibby and Philip Tosteson for informing the author about numerous crucial results and references. This work was done with the support of National Science Foundation grants DMS-1840234 and DMS-1701576.

References

[Bib16] Christin Bibby. Cohomology of abelian arrangements. *Proc. Amer. Math. Soc.*, 144(7):3093–3104, 2016.

[CEF14] Thomas Church, Jordan S. Ellenberg, and Benson 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.

[CH] Gilyoung Cheong and Yifeng Huang. Rationality for the Betti numbers of the unordered configuration spaces of a punctured torus. Submitted. arXiv: 2009.07976.

[Coh93] F. R. Cohen. On the mapping class groups for punctured spheres, the hyperelliptic mapping class groups, SO(3), and Spin^c(3). *Amer. J. Math.*, 115(2):389–434, 1993.

[DCK17] Gabriel C. Drummond-Cole and Ben Knudsen. Betti numbers of configuration spaces of surfaces. *J. Lond. Math. Soc. (2)*, 96(2):367–393, 2017.

[Del71] Pierre Deligne. Théorie de Hodge. II. *Inst. Hautes Études Sci. Publ. Math.*, (40):5–57, 1971.

[Del75] Pierre Deligne. Poids dans la cohomologie des variétés algébriques. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 1*, pages 79–85, 1975.

[Dup15] Clément Dupont. The Orlik-Solomon model for hypersurface arrangements. *Ann. Inst. Fourier (Grenoble)*, 65(6):2507–2545, 2015.

[Dup16] Clément Dupont. Purity, formality, and arrangement complements. *Int. Math. Res. Not. IMRN*, (13):4132–4144, 2016.

[Fol66] Jon Folkman. The homology groups of a lattice. *J. Math. Mech.*, 15:631–636, 1966.

[GM88] Mark Goresky and Robert MacPherson. *Stratified Morse theory*, volume 14 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1988.

[Gor81] V. V. Gorjunov. Cohomology of braid groups of series C and D. *Trudy Moskov. Mat. Obshch.*, 42:234–242, 1981.

[HRV08] Tamás Hausel and Fernando Rodriguez-Villegas. Mixed Hodge polynomials of character varieties. *Invent. Math.*, 174(3):555–624, 2008. With an appendix by Nicholas M. Katz.

[Kal08] Sadok 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.

[Kim94] Minhyong Kim. Weights in cohomology groups arising from hyperplane arrangements. *Proc. Amer. Math. Soc.*, 120(3):697–703, 1994.

[Loo93] Eduard Looijenga. Cohomology of $\mathcal{M}_3$ and $\mathcal{M}_3^1$. *Mapping class groups and moduli spaces of Riemann surfaces*, pages 205–228, 1993.

[MP20] L. Moci and R. Pagaria. On the cohomology of arrangements of subtori. 2020. arXiv: 2001.05180.

[Nap03] Fabien Napolitano. On the cohomology of configuration spaces on surfaces. *J. London Math. Soc. (2)*, 68(2):477–492, 2003.

[OT92] Peter Orlik and Hiroaki Terao. *Arrangements of hyperplanes*, volume 300 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1992.

[Pag] Roberto Pagaria. Cohomology of configuration spaces of surfaces. arXiv: 1905.04724.

[Pet17] Dan Petersen. A spectral sequence for stratified spaces and configuration spaces of points. *Geom. Topol.*, 21(4):2527–2555, 2017.

[Tos16] P. Tosteson. Lattice spectral sequences and cohomology of configuration spaces. 2016. arXiv: 1612.06034.

[Tot96] Burt Totaro. Configuration spaces of algebraic varieties. *Topology*, 35(4):1057–1067, 1996.

[VW15] Ravi Vakil and Melanie Matchett Wood. Discriminants in the Grothendieck ring. *Duke Math. J.*, 164(6):1139–1185, 2015.

[Wac07] Michelle L. Wachs. Poset topology: tools and applications. In *Geometric combinatorics*, volume 13 of *IAS/Park City Math. Ser.*, pages 497–615. Amer. Math. Soc., Providence, RI, 2007.

[Wei94] Charles A. Weibel. *An introduction to homological algebra*, volume 38 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1994.