MODULI OF HIGH RANK VECTOR BUNDLES OVER SURFACES

DAVID GIESEKER AND JUN LI

0. Introduction

The purpose of this work is to apply the degeneration theory developed in [GL] to study the moduli space of stable vector bundles of arbitrary rank on any smooth algebraic surface (over $\mathbb{C}$). We will show that most of the recent progress in understanding moduli of rank two vector bundles can be carried over to high rank cases.

After introducing the notion of stable vector bundles, the first author constructed the moduli schemes of vector bundles on surfaces. He showed that for any smooth algebraic surface $X$ with ample divisor $H$ and line bundle $I$ on $X$, there is a coarse moduli scheme $\mathcal{M}_{X}^{r,d}(I,H)$ parameterizing (modulo equivalence relation) the set of all $H$-semistable rank $r$ torsion free sheaves $E$ on $X$ with $\det E = I$ and $c_2(E) = d$. Since then, many mathematicians have studied the geometry of this moduli space, especially for the rank two case. To cite a few, Maruyama, Taubes and the first author showed that the moduli space $\mathcal{M}_{X}^{2,d}(= \mathcal{M}_{X}^{r,d}(I,H))$ is non-empty when $d$ is large. Moduli spaces of vector bundles of some special surfaces have been studied also.

The deep understanding of $\mathcal{M}_{X}^{r,d}$ for arbitrary $X$ and $r = 2$ begins with Donaldson’s generic smoothness result. Roughly speaking, Donaldson [Do] (later generalized by Friedman [Fr] and K. Zhu [Zh]) showed that when $d$ is large enough, then the singular locus $\text{Sing}(\mathcal{M}_{X}^{2,d})$ of $\mathcal{M}_{X}^{2,d}$ is a proper subset of $\mathcal{M}_{X}^{2,d}$ and its codimension in $\mathcal{M}_{X}^{2,d}$ increases linearly in $d$. This theorem indicates that the moduli $\mathcal{M}_{X}^{2,d}$ behaves as expected when the second Chern class $d$ is large. Later, using general deformation theory, the second author proved that $\mathcal{M}_{X}^{2,d}$ is normal, and has local complete intersection (l.c.i.) singularities at stable sheaves provided $d$ is large [L2]. He also showed that when $X$ is a surface of general type satisfying some mild technical conditions, then $\mathcal{M}_{X}^{2,d}$ is of general type for $d \gg 0$ [L2]. In our paper [GL], we also proved that $\mathcal{M}_{X}^{2,d}$ is irreducible if $d$ is large.

In this and subsequent papers, we shall show that the geometry of $\mathcal{M}_{X}^{2,d}$ and the geometry of $\mathcal{M}_{X}^{r,d}$, $r \geq 3$, is rather similar. The main obstacle in doing so is the lack of an analogy of the generic smoothness result in high rank case. In this paper, we will use the degeneration of moduli developed in [GL] to establish the following main technical theorem.

Received by the editors August 25, 1993 and, in revised form, October 1, 1994.

1991 Mathematics Subject Classification. Primary 14D20, 14D22, 14D25, 14J60.

The first author was partially supported by NSF grant DMS-9305657 and the second author was partially supported by NSF grant DMS-9307892.
Theorem 0.1. Let $X$ be a smooth algebraic surface, $H$ an ample line bundle and $I$ a line bundle on $X$. Let $r \geq 2$ be any integer. Then for any constant $C_1$ and any divisor $D \subseteq X$, there is an $N$ such that whenever $d \geq N$, then we have

$$\dim \{ E \in \mathcal{M}^{r,d}_X \mid \text{Ext}^0(E, E(D))^0 \neq \{ 0 \} \} \leq \eta(r, d, I) - C_1,$$

where $\eta(r, d, I) = 2rd - (r - 1)I^2 - (r^2 - 1)\chi(\mathcal{O}_X)$ is the expected dimension of $\mathcal{M}^{r,d}_X$ ($= \mathcal{M}^{r,d}_X(I, H)$) and the superscript 0 stands for the traceless part of $\text{Ext}^0(\cdot, \cdot)$.

According to [Ar], [Mu], $\mathcal{M}^{r,d}_X$ is regular at $E$ if $E$ is stable and $\text{Ext}^2(E, E)^0 = \{ 0 \}$. As to the subset of strictly semistable sheaves in $\mathcal{M}^{r,d}_X$, it is easy to show that its dimension is much less than $\eta(r, d, I) - C_1$ when $d$ is large. After applying Theorem 0.1 to the divisor $D = K_X$ and using the Serre duality, we conclude that for $d$ sufficiently large,

$$\dim \text{Sing}(\mathcal{M}^{r,d}_X) \leq \eta(r, d, I) - C_1.$$

On the other hand, based on deformation theory, each component of $\mathcal{M}^{r,d}_X$ has dimension at least $\eta(r, d, I)$. Thus, we have proved the following theorem.

Theorem 0.2. Let $X$ be a smooth algebraic surface, $H$ an ample line bundle and $I$ a line bundle on $X$. Let $r \geq 2$ be any integer. Then for any constant $C_1$, there is an $N$ such that whenever $d \geq N$, then $\mathcal{M}^{r,d}_X$ has pure dimension $\eta(r, d, I)$ and further,

$$\text{codim}(\text{Sing}(\mathcal{M}^{r,d}_X), \mathcal{M}^{r,d}_X) \geq C_1.$$

Once we have settled the generic smoothness result, we can generalize some other properties of $\mathcal{M}^{r,d}_X$ to high rank case. In this paper, we will prove

Theorem 0.3. With the notation as in Theorem 0.2, there is an $N$ such that whenever $d \geq N$, then:

1. $\mathcal{M}^{r,d}_X$ is normal. Further, if $s \in \mathcal{M}^{r,d}_X$ is a closed point corresponding to a stable sheaf, then $\mathcal{M}^{r,d}_X$ is a local complete intersection at $s$.
2. The set of locally free $\mu$-stable sheaves $(\mathcal{M}^{r,d}_X)^{\mu} \subseteq \mathcal{M}^{r,d}_X$ is dense in $\mathcal{M}^{r,d}_X$.
3. For any polarizations $H_1$ and $H_2$ of $X$, the moduli $\mathcal{M}^{r,d}_X(I, H_1)$ is birational to $\mathcal{M}^{r,d}_X(I, H_2)$. (In this case, $N$ depends on both $H_1$ and $H_2$.)

To illustrate the idea of the proof of our main theorem (Theorem 0.1), let us first recall the degeneration of moduli $\mathcal{M}^{r,d}_X$ constructed in [GL]. Let $0 \in C \subseteq \text{Spec} \mathbb{C}[t]$ be a smooth curve that functions as a parameter space and let $Z \to C$ be a family of surfaces that is the result of blowing-up $X \times C$ along $\Sigma \times \{ 0 \}$, where $\Sigma \in |H|$ is a smooth very ample divisor. Clearly, $Z_t = \pi^{-1}(t)$ is $X$ and $Z_0 = X \cup \Delta$, where $\Delta$ is a ruled surface over $\Sigma$. Over $C^* = C \setminus \{ 0 \}$, we have a constant family $\mathcal{M}^{r,d}_X \times C^*$. In [GL], we have constructed completions of $\mathcal{M}^{r,d}_X \times C^*$ over $C$. These completions depend on the choice of ample divisors on $Z$. The ample divisor which we will use is a multiple of the $\mathbb{Q}$-divisor $p^*_X H(- (1 - \varepsilon)\Delta)$ that depends on the rational $\varepsilon \in (0, \frac{1}{2})$. We denote this completion by $\mathcal{M}^{r,d}_{\varepsilon}$. There is a nice description of closed points of the special fiber $\mathcal{M}^{r,d}_{\varepsilon}_0$: Any point of $\mathcal{M}^{r,d}_{\varepsilon}_0$ corresponds uniquely to an equivalence class of semistable sheaves on $Z_0$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Now let $D \subseteq X$ be any divisor and $N \subseteq M^{r,d}_X$ be the set of sheaves $E$ such that
$$(0.1) \quad \text{Hom}(E, E(D))^0 \neq \{0\}.$$ 
Put $M^{d,\varepsilon} \subseteq M^{d,\varepsilon}_X$ as the closure of $N \times C^*$ in $M^{d,\varepsilon}_X$. To show that for any constant $C_1$ and large $d$ we have
$$\dim N \leq \eta(r, d, I) - C_1,$$
it suffices to show
$$\dim (0.2) \quad \dim M^{d,\varepsilon}_0 \leq \eta(r, d, I) - C_1.$$ 

Now let $E \in M^{d,\varepsilon}_0$ be any sheaf. Note that $E$ is a limit of sheaves in $N$ and that sheaves in $N$ satisfy (0.1). So by the semicontinuity theorem, for any invertible sheaf $L$ on $Z$ such that $L|_Z \sim O_X(D)$, we have
$$\text{Hom}(Z, 0) \neq \{0\}.$$ 
In particular, if we choose $L$ to be $p^*X O_X(D)^{-k\Delta}$, where $p_X : Z \to X$ is the projection, we get
$$(0.3) \quad \text{Ext}^0_Z(E, E \otimes p^*X O_X(D)^{-k\Delta})^0 \neq \{0\}, \quad \forall k \in \mathbb{Z}, \ E \in M^{d,\varepsilon}_0.$$ 
Since $E$ is semistable, $E|_X$ and $E|_\Delta$ as sheaves on $X$ and $\Delta$ respectively will satisfy some weak stability conditions. (For simplicity, here we assume $E$ is locally free.) On the other hand, for large $k$, the non-vanishing of
$$\text{Ext}^0_X(E|_X, E|_X(D - k\Sigma))^0$$
will force $E|_X$ to be very unstable. Therefore, we can choose a $k > 0$ (independent of $d$, $\varepsilon$ and $N$) such that (0.4) is always trivial. Thus (0.3) will force
$$(0.5) \quad \text{Ext}^0_\Delta(E|_\Delta, E|_\Delta \otimes p^*X O_X(D)(k\Sigma))^0 \neq \{0\}.$$ 
(0.5) certainly is possible for sheaves over $\Delta$. However, if we can show that the number of moduli of the set of sheaves $F$ (over $\Delta$) satisfying (0.5) is strictly less than
the number of moduli of $\{E_\Delta | E \in M^{d,\varepsilon}_0\} - C_1$,
then $\text{codim}(M^{d,\varepsilon}_0, M^{d,\varepsilon}_0) \geq C_1$, which is exactly what we need. Therefore, the proof of Theorem 0.1 is reduced to the proof of the following theorem.

**Theorem 0.4.** Let $X$ be any ruled surface and let $H$ and $I$ be as in Theorem 0.1. Then for any integer $r$, any divisor $D \subseteq X$ and any constant $C_1$, there is a constant $N$ such that for $d \geq N$,
$$\dim \{E \in M^{r,d}_X | \text{Ext}^0(E, E(D))^0 \neq \{0\}\} \leq \eta(r, d, I) - C_1.$$ 

The advantage of working with a ruled surface lies in the fact that every vector bundle on a ruled surface can be constructed explicitly as follows: Let $X = \Delta$ and let $E$ be a vector bundle on $\Delta$. For simplicity, we assume for the general fiber $P_\xi$ of
\( \pi : \Delta \to \Sigma \), the restriction sheaf \( E|_{P_i} \cong \mathcal{O}_{P_i}^{\oplus r} \). Then there is a unique rank \( r \) vector bundle \( V \) on \( \Sigma \) and a sheaf \( F \) supported on a finite number of fibers of \( \pi \) such that

\[
0 \longrightarrow E \longrightarrow \pi^*V \overset{\varphi}{\longrightarrow} F \longrightarrow 0
\]

is exact. When \( E \) is general, \( F \) is of the form \( \bigoplus \mathcal{O}_{P_i}(1) \), where \( P_i \) are fibers of \( \pi \). Thus the condition under which \( E \) admits a traceless homomorphism \( E \to E(D) \) can be interpreted in terms of the location of \( P_i \)'s and the choice of homomorphism \( \varphi \). The argument to carry out this approach is rather straightforward though quite technical and will occupy the first section of this paper. In \( \S \)2, we will review the degeneration construction and use it to prove Theorem 0.1. Theorems 0.2–0.4 will be proved in \( \S \)3. We remark that after the completion of the initial version of this work, O’Grady has improved our results in his paper [OG].

**Conventions and preliminaries.** All schemes are defined over the field of complex numbers \( \mathbb{C} \) and are of finite type. All points are closed points unless otherwise mentioned. We shall always identify a vector bundle with its sheaf of sections. If \( I \) and \( J \) are two line bundles on a surface, then we denote by \( I \cdot J \) the intersection \( c_1(I) \cdot c_1(J) \) and by \( I^2 \) the self-intersection \( c_1(I) \cdot c_1(I) \). We will use \( \sim \) to denote the numerical equivalence of divisors (line bundles). For the coherent sheaf \( F \), we denote by \( \text{rk}(F) \) the rank of \( F \). In case \( F \) is supported on a finite number of points on \( X \), we denote by \( \ell(F) \) the length of \( F \). If \( p \) and \( q \) are two polynomials with real coefficients, we say \( p \succeq q \) (resp. \( p \preceq q \)) if \( p(n) > q(n) \) (resp. \( p(n) \geq q(n) \)) for all \( n \gg 0 \).

In the following, \( X \) will always denote a smooth projective surface. Let \( H \) be a very ample line bundle on \( X \). For any sheaf \( E \) on \( X \), we denote by \( \chi_E \) the Poincaré polynomial of \( E \), namely, \( \chi_E(n) = \chi(E(n)) \), \( E(n) = E \otimes H^\otimes n \), and denote by \( p_E \) the polynomial \( \frac{1}{\text{rk}(E)} \chi_E \) when \( \text{rk}(E) \neq 0 \). Unless the contrary is mentioned, the degree of a sheaf \( E \) is \( c_1(E) \cdot H \). We recall the notion of stability:

**Definition 0.5.** A sheaf \( E \) on \( X \) is said to be stable (resp. semistable) with respect to \( H \) if \( E \) is coherent, torsion free and if one of the following two equivalent conditions hold:

1. Whenever \( F \subset E \) is a proper subsheaf, then \( p_F \prec p_E \) (resp. \( p_F \preceq p_E \)).
2. Whenever \( E \to Q \) is a quotient sheaf, \( \text{rk}(Q) > 0 \), then \( p_E \prec p_Q \) (resp. \( p_E \preceq p_Q \)).

When \( E \) is a torsion free coherent sheaf on \( X \), we define the slope \( \mu(E) = \frac{1}{\text{rk}(E)} \deg E \).

**Definition 0.6.** Let \( e \) be a constant. The sheaf \( E \) is said to be \( e \)-stable if one of the following two equivalent conditions holds:

1. Whenever \( F \subset E \) is a subsheaf with \( 0 < \text{rk}(F) < \text{rk}(E) \), then \( \mu(F) \prec \mu(E) + \frac{\sqrt{H^2 \cdot e}}{\text{rk}(F)} \).
2. Whenever \( E \to Q \) is a quotient sheaf with \( 0 < \text{rk}(Q) < \text{rk}(E) \), then \( \mu(Q) \prec \mu(E) + \frac{\sqrt{H^2 \cdot e}}{\text{rk}(Q)} \).

We call \( E \) \( \mu \)-stable if \( E \) is \( e \)-stable with \( e = 0 \). When the strict inequality is replaced by \( \preceq \), then we call \( E \) \( e \)-semistable.

Let \( W \to S \) be a flat morphism and let \( E \to W \) be any sheaf on \( W \). For any closed \( s \in S \), we will use \( W_s \) to denote the fiber of \( W \) over \( s \) and use \( E_s \).
to denote the restriction of $E$ to $W_s$. For any subscheme $T \subseteq W$, we denote by $E_T$ the restriction of $E$ to $T$. We shall adopt the following convention: If $\mathcal{R}$ is a set of sheaves on $X$, then the number of moduli of $\mathcal{R}$ is the smallest integer $m$ so that there are countably many schemes (of finite types) of dimension at most $m$, say $S_1, S_2, \ldots$, and flat families of sheaves $E_{S_1}, E_{S_2}, \ldots$ on $X \times S_1, X \times S_2, \ldots$ respectively of which the following holds: For any $F \in \mathcal{R}$, there is a closed $s \in S_k$ for some $k$ such that $F \cong E_{S_k,s}$. We will denote by $\#_{\text{mod}}(\mathcal{R})$ the number of moduli of $\mathcal{R}$. In case $\mathcal{R}$ is a scheme parameterizing a family of sheaves and $t \in R$, then we denote by $\#_{\text{loc}}(\mathcal{R},[t])$ the number of moduli of sheaves parameterized by the germ of $\mathcal{R}$ at $r$. In particular, we write $\#_{\text{mod}}(E)$, where $E$ is any sheaf, for $\#_{\text{loc}}(\mathcal{Q}_E([E]) (= \text{the number of moduli of the set of all "small" deformations of } E)$, where $\mathcal{Q}$ is Grothendieck’s Quot-scheme $[\text{Gr}]$ that contains all deformations of $E$ as quotient sheaves of some appropriate locally free sheaf. Another notion we use frequently is $\#_{\text{aut}}(E) = \dim \text{Aut}(E)$, where $\text{Aut}(E)$ is the group of automorphisms of $E$. Note $\#_{\text{aut}}(E) = h^0(\mathcal{E}nd(E))$. When $\mathcal{R}$ is a set of sheaves, then $\#_{\text{aut}}(\mathcal{R}) = \max\{\#_{\text{aut}}(E) \mid E \in \mathcal{R}\}$.

1. Vector bundles on a ruled surface

The purpose of this section is to prove an analogy of Theorem 0.1 for a ruled surface $\Delta$. Before giving the precise statement of the theorem, we first introduce some notation. Let $\Sigma$ be a smooth curve and let $\pi: \Delta \to \Sigma$ be a ruled surface. For simplicity, we assume $\Delta$ is the projective bundle of a direct sum of a trivial line bundle with a very ample line bundle (over $\Sigma$). Hence $\pi: \Delta \to \Sigma$ has a unique section $\Sigma^-$ with $\Sigma^- \cdot \Sigma^- < 0$ and has many sections with positive self-intersection. We choose one such section and denote it by $\Sigma^+$. By assumption, $|\Sigma^+|$ is base point free. Let $H$ be an ample line bundle on $\Delta$ that is numerically equivalent to $(\sim) a\Sigma^+ + bP_\xi$, where $P_\xi$ is a general fiber of $\pi$. Let $e$ be a constant, let $I$ be a line bundle on $\Delta$ and let $D$ be any divisor on $\Delta$. In this section, we will study the set $\mathcal{A}_{r,d,H}^{e,1}$ of all $e$-semistable (with respect to $H$) rank $r$ locally free sheaves $E$ with $\det E = I$ and $c_2(E) = d$ and the set

$$\mathcal{A}_{r,d,H}^{e,1}(D) = \{E \in \mathcal{A}_{r,d,H}^{e,1} \mid \text{Hom}(E,E(D))^0 \neq \{0\}\}.$$ 

Here and in the following, the superscript 0 always stands for the traceless part of the group or sheaf. For technical reasons, we will choose $H$ to be very close to $\Sigma^+$ in the sense that $b/a$ is very small. With the choice of $H$ understood, we will not build $H$ into the notation and will write $\mathcal{A}_{r,d}^{e,1}$ (resp. $\mathcal{A}_{r,d,H}^{e,1}(D)$) for $\mathcal{A}_{r,d,H}^{e,1}$ (resp. $\mathcal{A}_{r,d,H}^{e,1}(D)$). We will also use $\eta_{\Delta}(E) = \eta_{\Delta}(\text{rk}(E),c_2(E),c_1(E))$ to denote the number

$$\eta_{\Delta}(r,d,I) = 2rd - (r - 1)I^2 - (r^2 - 1)\chi(\mathcal{O}_\Delta).$$ 

$\eta_{\Delta}(r,d,I)$ is the expected dimension of $\mathcal{A}_{r,d,H}^{e,1}$. Because in this section we work solely with the surface $\Delta$, we will simply write $\eta$ for $\eta_{\Delta}$. The theorem we will prove in this section is the following.

**Theorem 1.1.** Given $r$ and $\Delta$, there is an $\varepsilon_0 > 0$ depending on $r$ and $\Delta$ for which the following holds: For any ample divisor $H \sim a\Sigma^+ + bP_\xi$ with $b/a < \varepsilon_0$ and for
any choice of constants $e$, $C$ and divisor $D \subseteq \Delta$, $I \in \text{Pic}(\Delta)$, there is an integer $N$ such that whenever $d \geq N$, then we have

\begin{equation}
\#_{\text{mod}}(\mathcal{A}_{e,I}^{r,d}(D)) \leq \eta(r,d,I) - C.
\end{equation}

The advantage of working with ruled surfaces lies in having a powerful structure theorem of torsion free sheaves on $\Delta$. Let $E$ be any torsion free sheaf of rank $r$. By Grothendieck’s splitting theorem, its restriction to a generic fiber $P_\xi$ has the form

\begin{equation}
E|_{P_\xi} \cong \bigoplus_{i=1}^{n} O_{P_\xi}(\alpha_i)^{\otimes r_i}, \quad \alpha_1 > \cdots > \alpha_n.
\end{equation}

In the following, we call $\alpha = (\alpha_1^{\otimes r_1}, \cdots, \alpha_n^{\otimes r_n})$ the generic fiber type of $E$. (The integer sequence $\{\alpha_i\}$ is always assumed to be strictly decreasing.) We let $\ell(\alpha) = \sum_{i=1}^{n} r_i \alpha_i$. Clearly, $r = \sum_{i=1}^{n} r_i$ and further, when $\det E = I$ and $\deg I|_{P_\xi} = m$, then $m = \ell(\alpha)$. $\mathcal{A}_{e,I}^{r,d}$ can be divided into strata according to the generic fiber types of individual vector bundles. Let $r \in \mathbb{N}$ and $I \in \text{Pic}(\Delta)$ be fixed. Without loss of generality, we can assume $0 \leq \deg I|_{P_\xi} \leq r - 1$. Let $m = \deg I|_{P_\xi}$ and let $1_m$ be the fiber type $(1^{\otimes m}, 0^{\otimes (r-m)})$. For any fiber type $\alpha$ with $\ell(\alpha) = m$, we let

\begin{equation}
\mathcal{A}_{e,I}^{r,d}(\alpha) = \{ E \in \mathcal{A}_{e,I}^{r,d} \mid E \text{ has generic fiber type } \alpha \}.
\end{equation}

The first observation we have is that except for $\alpha = 1_m$, none of $\#_{\text{mod}}(\mathcal{A}_{e,I}^{r,d}(\alpha))$ are close to $\eta(r,d,I)$. More precisely, we have

**Theorem 1.2.** Let $m = \deg I|_{P_\xi}$. There are constants $C_1$ and $\varepsilon_0$ depending on $(r, \Delta)$ such that for any ample divisor $H \sim a\Sigma^+ + bP_\xi$ with $b/a < \varepsilon_0$ and any fiber type $\alpha \neq 1_m$, we have

\begin{equation}
\#_{\text{mod}}(\mathcal{A}_{e,I}^{r,d}(\alpha)) \leq (2r - 1)d + C_1.
\end{equation}

The proof of Theorem 1.2 goes as follows: Let $\alpha = (\alpha_1^{\otimes r_1}, \cdots, \alpha_n^{\otimes r_n})$ be any fiber type. Then each $E \in \mathcal{A}_{e,I}^{r,d}(\alpha)$ admits a relative Hardar-Narasimhan filtration

\begin{equation}
0 = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_n = E
\end{equation}

of which the quotient sheaves $F_i = E_i/E_{i-1}$ are torsion free with generic fiber types $(\alpha_i^{\otimes r_i})$ respectively. Clearly, the deformation of $E$ within $\mathcal{A}_{e,I}^{r,d}(\alpha)$ depends on deformation of individual $F_i$ and the extension $E_i \to E_{i+1} \to F_{i+1}$. The contribution of these data to the number of moduli of $\mathcal{A}_{e,I}^{r,d}(\alpha)$ can be estimated by using Riemann-Roch. The details of the proof will be provided shortly.

In light of Theorem 1.2, to prove Theorem 1.1 we only need to study the stratum $\mathcal{A}_{e,I}^{r,d}(1_m)$ and

\begin{equation}
\mathcal{A}_{e,I}^{r,d}(1_m, D) = \{ E \in \mathcal{A}_{e,I}^{r,d}(1_m) \mid \text{Hom}(E, E(D))^0 \neq 0 \}.
\end{equation}

In this section, we will first establish Theorem 1.1 for the stratum $\mathcal{A}_{e,I}^{r,d}(1_0, D)$ and derive the remainder by induction on $m$. 
Let $E$ be any vector bundle of generic fiber type $1_0=(0^{r_1})$. Let $x \in \Sigma$ be any point, let $P_x$ be the fiber of $\pi$ over $x \in \Sigma$ and let $\beta_x(E)=(\beta_x^1,\cdots,\beta_x^{r_n})$ be the fiber type of $E_{|P_x}$. In case $\beta_x(E) \neq 1_0$, we call $P_x$ a jumping line of $E$. Let $P_x$ be a jumping line of $E$. We then perform semistable reduction on $E$ along $P_x$ by taking $F$ to be the kernel of the (unique surjective) homomorphism $E \rightarrow O_{P_x}(\beta_n)^{\oplus r_n}$. For convenience, we will use $\Upsilon_x$ to denote this operation and denote $F=\Upsilon_x(E)$ and $\omega_x(E)=\beta_x^{\oplus r_n}$. Clearly, $F$ belongs to the exact sequence

$$0 \rightarrow F \rightarrow E \rightarrow \mathcal{O}_{P_x}(\beta_n)^{\oplus r_n} \rightarrow 0.$$  

(1.6)

An easy calculation based on Riemann-Roch yields

Lemma 1.3. Let $F=\Upsilon_x(E)$ with $\omega_x(E)=t^{\oplus s}$. Then $c_1(F)=c_1(E)-s[P_x]$ and $c_2(F)=c_2(E)+s\cdot t$. In particular, $\eta(r,c_2(F),c_1(F))=\eta(r,c_2(E),c_1(E))=2rs\cdot t$.

Proof. See [Br, p. 166].

In case $F$ still has a jumping line, say $P_y$ of type $(\cdots,\gamma_l^{\oplus s_l})$, then we can further perform semistable reduction on $F$ to get $F_2=\Upsilon_y(F)$. We can iterate this process as long as the resulting vector bundle $F_k$ still admits jumping lines. In general, if $F_k$ is derived by successively performing this type of elementary transformations, namely, $F_0=E$ and $F_{i+1}=\Upsilon_x(F_i)$ with $\omega_x(F_i)=t^{i\oplus s_i}$, for $i=0,\cdots,k-1$, then we will write

$F_k=\Upsilon_\Lambda(E), \ \Lambda=(x_1,\cdots,x_k)$

and define $\omega_\Lambda(E)=(t_1^{\oplus s_1},\cdots,t_k^{\oplus s_k})$. We call $k$ the length of $\Lambda$.

Lemma 1.4. For any vector bundle $E$ of generic fiber type $1_0$, there is a finite length $\Lambda=(x_1,\cdots,x_k)$ such that $\Upsilon_\Lambda(E)$ has no jumping lines.

Proof. By Lemma 1.3, the second Chern class of $\Upsilon_x(E)$ is strictly less than $c_2(E)$ because $\beta_n<0$ when $\beta_x(E) \neq 1_0$. Thus Lemma 1.4 follows if we can show that any vector bundle of generic fiber type $1_0$ has non-negative second Chern class.

Indeed, let $E$ be any vector bundle of generic fiber type $1_0$. We choose a divisor $D$ supported on fibers of $\pi$ such that $O_D$ is a subsheaf of $E$ with $E/O(D)$ torsion free. Since $E/O(D)$ has generic fiber type $1_0$, we can assume $c_2(E/O(D)) \geq 0$ by the induction hypothesis on the rank of $E$. Hence,

$$c_2(E) = c_2(E/O(D)) + D \cdot (c_1(E) - D) = c_2(E/O(D)) \geq 0.$$ 

This completes the proof of Lemma 1.4.

Let $E$ be a vector bundle of generic fiber type $1_0$ and let $\Lambda=(x_1,\cdots,x_k)$ be such that $F=\Upsilon_\Lambda(E)$ has no jumping lines. Then $F$ is a pull-back vector bundle $\pi^*V$ whose dual belongs to the exact sequence

$$0 \rightarrow E^\vee \rightarrow \pi^*V^\vee \rightarrow J \rightarrow 0,$$  

(1.7)

where $J$ is a torsion sheaf supported on the union of fibers $P_{x_i}$. Usually, the sheaf $J$ near some fiber $P_{x_i}$ can be very complicated. The case that is easy to understand and will be dealt with extensively in the subsequent discussion is when $J \cong O_{P_{x_1}}(1)$ near $P_{x_1}$. The following theorem says that when the number of moduli of $\mathfrak{A}^{r,d}_c(1_0)$ is close to $\eta(r,d,I)$, then for general $E \in \mathfrak{A}^{r,d}_c(1_0)$ with the exact sequence (1.7), $J \cong O_{P_{x_1}}(1)$ near $P_{x_1}$ for all $x_i \in \{x_1,\cdots,x_k\}$ except for a bounded number of fibers.
Theorem 1.5. For any constant $e$, there is a constant $C_2$ such that

\begin{equation}
\#_{\text{mod}} \mathfrak{A}_{e,I}^{r,d} \leq \eta(r, d, I) + C_2.
\end{equation}

Further, for any constant $C$, there are integers $l, l_1, l_2$ and $N_1$ of which the following holds: Assume $d \geq N_1$ and that $S$ is a variety parameterizing a subset of $\mathfrak{A}_{e,I}^{r,d}(1_0)$ satisfying $\#_{\text{mod}}(S) \geq \eta(r, d, I) - C$. Then there is a line bundle $L$ on $\Sigma$ of degree $[(d-c)/r] + l_1$, where $c = I \cdot \Sigma^+$, so that for general $E \in S$, there are

1. $d - l$ distinct points $x_1, \ldots, x_{d-l} \in \Sigma$ in general position, a surjective homomorphism $\tau_1 : \pi^*L^{\oplus r} \to \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{x_i}}(1)$ and

2. a zero-dimensional scheme (divisor) $z_0 \subseteq \Sigma$ away from $\{x_1, \ldots, x_{d-l}\}$ with $\ell(z_0) \leq l_2$ and a sheaf of $\mathcal{O}_{\pi^{-1}(z_0)}$-modules $J$ with a quotient homomorphism $\tau_0 : \pi^*L^{\oplus r} \to J$ so that $E'$ belongs to the exact sequence

\begin{equation}
0 \longrightarrow E' \longrightarrow \pi^*(L^{\oplus r})^{\tau_0,J} \oplus \left( \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{x_i}}(1) \right) \longrightarrow 0.
\end{equation}

This theorem holds for a very simple reason: To maximize the number of moduli of the set of those $E$ in (1.9), we need to maximize the number of moduli of the set of homomorphisms $\tau_1$ and the quotient sheaves in (1.9). This can only be achieved by letting $J = \{0\}$ and $x_i$ general. Hence, if $\#_{\text{mod}} \mathfrak{A}_{e,I}^{r,d}(1_0)$ is close to the expected dimension $\eta(r, d, I)$, then the number of fibers in $\text{supp}(J)$ cannot be too large.

Now we sketch how this structure theorem of $\mathfrak{A}_{e,I}^{r,d}(1_0)$ leads to the proof of Theorem 1.1. We first prove the case $m = 0$ by contradiction. Assume $\#_{\text{mod}} \mathfrak{A}_{e,I}^{r,d}(1_0, D) \geq \eta(r, d, I) - C$. Then by Theorem 1.5, the general element $E \in \mathfrak{A}_{e,I}^{r,d}(1_0, D)$ fits into the exact sequence (1.9) with $\{x_1, \ldots, x_{d-l}\}$ and $\tau_1 : \pi^*L^{\oplus r} \to \bigoplus \mathcal{O}_{P_{x_i}}(1)$ general. Now let $F = \ker(\pi^*L^{\oplus r} \to J)$ and let $f : E \to E(D)$ be a non-trivial traceless homomorphism. Then $f$ induces a non-trivial traceless homomorphism

$$\tilde{f} : F' \longrightarrow F(D + \pi^{-1}(z_0)),$$

where $z_0$ is a divisor of $\Sigma$ as in Theorem 1.5(2). Because the position of $x_1, \ldots, x_{d-l}$ and the homomorphism $\tau_1$ are general, we will see by degeneration theoretic methods that $\text{Hom}(F', F'(D + \pi^{-1}(z_0))) \neq 0$ for the torsion free sheaf

$$F' = \ker(\pi^*L^{\oplus r} \to \bigoplus \mathcal{O}_{P_{x_i}}(1)),$$

where $p_i \in P_{x_i}$. Because of the special choice of $F'$, the non-vanishing of the previous group amounts to saying that for any choice of $p_i \in \Delta$, there are sections of $H^0(\mathcal{O}(D + \pi^{-1}(z_0)))$ that vanish on $[(d-I)/r]$ of $p_1, \ldots, p_{d-I}$. On the other hand, since $D$ is fixed and $\pi^{-1}(z_0)$ is bounded, this is impossible if $\{p_i\}$ are generic and $d$ is sufficiently large. This leads to a contradiction which ensures that $\eta(r, d, I) - \#_{\text{mod}} \mathfrak{A}_{e,I}^{r,d}(1_0, D)$ can be arbitrarily large.

For the general case, we use induction on $m$ (with $r \geq m$ fixed). Assume the theorem holds for $m-1 \geq 0$ and assume $\#_{\text{mod}} \mathfrak{A}_{e,I}^{r,d}(1_m, D) \geq \eta(r, d, I) - C$. Then for general $E \in \mathfrak{A}_{e,I}^{r,d}(1_m, D)$, we can perform an elementary transformation on $E$ along a section $\Sigma^+$ to get a new vector bundle $\tilde{E} \in \mathfrak{A}_{e,I}^{r,d'}(1_{m-1}, D)$. By carefully
studying this correspondence, we will get the desired estimate of \( \#_\text{mod} \mathfrak{A}^{r,d}_{e,I}(1_m, D) \) from the known estimate of \( \#_\text{mod} \mathfrak{A}^{r,d}_{e,I}(1_{m-1}, \hat{D}) \), thus establishing Theorem 1.1.

In the following, we will fill in the details of the above sketch. We continue to use the notation introduced before Lemma 1.4. We begin with the estimate of the number of moduli of vector bundles of generic fiber type 1. Let \( E_0 \in \mathfrak{A}^{r,d}_{e,I}(1_0) \) be any vector bundle of generic fiber type 1. Let \( \Lambda = \langle x_1, \ldots, x_\ell \rangle \) be such that \( F = \Upsilon_\Lambda(E_0) \) has no jumping line. Then \( F \) is a pull-back vector bundle \( \pi^*V \). Let \( \omega \) be \( \omega_\Lambda(E_0) = (t_1^{\otimes s_1}, \ldots, t_k^{\otimes s_k}) \) and let

\[
\mathcal{S}_{\Lambda, \omega}(F) = \{ E \in \mathfrak{A}^{r,d}_{e,I}(1_0) \mid \omega_\Lambda(E) = \omega \ \text{and} \ \Upsilon_\Lambda(E) = F \}.
\]

In the following, we will estimate the number of moduli of this set. We first study the case where \( \Lambda = \langle x \rangle \) and \( \omega = (t^{\otimes s}) \). Let \( \beta_x(F) = (\cdot, \cdot^{\otimes r_1}) \). Because of the following lemma, either \( t < \beta_t \) or \( t = \beta_t \).

**Lemma 1.6.** Suppose \( E_P \) has fiber type \((\cdot, \cdot, \gamma_1^{\otimes s_n})\) and that \( \Upsilon_x(E) \) has fiber type \((\cdot^{\otimes s_1})\) at \( x \). Then either \( \gamma_n < \beta_t \) or \( \gamma_n = \beta_t \) and \( r_1 \leq s_n \).

**Proof.** Since \( F = \Upsilon_x(E) \) is the kernel of \( E \to \mathcal{O}_P(\gamma_n)^{\otimes s_n} \), \( F \) belongs to the exact sequence

\[
0 \to \mathcal{O}_P(\gamma_n)^{\otimes s_n} \to F_1 | P \to \bigoplus_{i=1}^{n-1} \mathcal{O}_P(\gamma_i)^{\otimes s_i} \to 0.
\]

Then the lemma follows because \( \gamma_n < \gamma_{n-1} < \cdots < \gamma_1 \).

Let \( E \in \mathcal{S}_{\omega}(F) \). By dualizing the sequence (1.6), we get

\[
(1.10) \quad 0 \to E^\vee \to F^\vee \to \mathcal{O}_P(-t)^{\otimes s} \to 0.
\]

Clearly, all possible \( E^\vee \) that fit into (1.10) are parameterized by a subset of \( \Xi \) that is the total space of \( \text{Hom}(E^\vee, \mathcal{O}_P(-t)^{\otimes s}) \). Now let \( \Theta \subseteq \Xi \) be the subset consisting of \( \gamma : \gamma^\vee : \mathcal{O}_P(-t)^{\otimes s} \) such that \( \ker(\gamma)^\vee \in \mathcal{S}_{\omega}(F) \). \( \Theta \) admits a left \( GL(s, \mathbb{C}) \) action and a right \( \text{Aut}(F^\vee) \) action as follows: Let \( \varphi_1 \in \text{Aut}(F^\vee) \) and let \( \varphi_2 \in GL(s) = \text{Aut}(\mathcal{O}_P(-t)^{\otimes s}) \). Then

\[
\varphi_2 \cdot \gamma \cdot \varphi_1 = \varphi_2 \circ \gamma \circ \varphi_1 \in \text{Hom}(F^\vee, \mathcal{O}_P(-t)^{\otimes s}).
\]

Geometrically, \( \varphi \cdot \gamma \cdot \varphi_1 \) corresponds to a locally free sheaf \( E' \) defined by

\[
0 \to E^\vee \to F^\vee \to \mathcal{O}_P(-t)^{\otimes s} \to 0.
\]

Clearly, \( E' \) is isomorphic to \( E = \ker(\gamma)^\vee \). Conversely, suppose \( E_1 \) and \( E_2 \) are two isomorphic locally free sheaves associated to \( \gamma_1, \gamma_2 \in \Theta \). Then the isomorphism \( \varphi : E_1 \to E_2 \) induces an isomorphism between \( \Upsilon_x(E_1) \) and \( \Upsilon_x(E_2) \). Hence, there is an automorphism \( \varphi_1 : F^\vee \to F^\vee \) fitting into the (commutative) diagram

\[
\begin{array}{ccc}
0 & \to & E^\vee_2 \to F^\vee \to \mathcal{O}_P(-t)^{\otimes s} \to 0 \\
\downarrow \varphi^\vee & & \downarrow \varphi_1 \\
0 & \to & E^\vee_1 \to F^\vee \to \mathcal{O}_P(-t)^{\otimes s} \to 0
\end{array}
\]

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
In particular, there is a $\varphi_2:O_{P_2}(-t)^{\oplus s} \rightarrow O_{P_2}(-t)^{\oplus s}$ such that $\varphi_2 \circ \gamma_1 = \gamma_2 \circ \varphi_1$. Therefore those points in $\Theta$ that give rise to isomorphic sheaves form an $\text{Aut}(F') \times GL(s)$ orbit. Next, we will determine the size of the stabilizer in $\text{Aut}(F') \times GL(s)$ of any $\gamma \in \Theta$. Suppose $\varphi_1 \in \text{Aut}(F')$ and $\varphi_2 \in GL(s)$ are such that the right rectangle below is commutative:

$$
\begin{array}{ccc}
0 & \longrightarrow & E' \\
& \downarrow & \\
0 & \longrightarrow & F' \\
& & \downarrow \varphi_1 \\
& & \varphi_2^{-1} \\
\end{array}
\longrightarrow
\begin{array}{ccc}
0 & \longrightarrow & O_{P_2}(-t)^{\oplus s} \\
& \downarrow & \\
0 & \longrightarrow & O_{P_2}(-t)^{\oplus s} \\
& & \downarrow \\
\end{array}
$$

Then it induces a $\varphi \in \text{Aut}(E')$. One sees that such a map $\text{Stab}_{\gamma} \rightarrow \text{Aut}(E')$ is injective. Thus, if we let $\Theta_k \subseteq \Theta$ be the set of $\gamma$'s such that $\#_{\text{aut}}(\ker(\gamma)) = k$ (for a set $R$ of sheaves, we define $\#_{\text{aut}}(R) = \max_{E \in R} \{\dim \text{Hom}(E, E)\}$), then

\begin{align}
(1.11) \quad \dim \left( GL(s) \setminus \Theta_k / \text{Aut}(E') \right) \leq \dim \text{Hom}(F', O_{P_2}(-t)^{\oplus s}) - (s^2 + \#_{\text{aut}}(F')) + k.
\end{align}

Finally, let $(\beta_1^{t_1}, \cdots, \beta_n^{t_n})$ be the fiber type of $F|_{P_2}$. Then by Lemma 1.6, $t \leq \beta_i$. Because $\sum r_i \beta_i = 0$, we have

\begin{align}
(1.12) \quad \dim \text{Hom}(F', O_{P_2}(-t)) = \sum_{i=1}^n r_i \dim H^0(O_{P_2}(\beta_i - t)) = r(-t + 1).
\end{align}

Returning to the general case $\Lambda = (x_1, \cdots, x_k)$ and $\omega = (t_1^{\oplus s_1}, \cdots, t_n^{\oplus s_n})$, we will prove:

**Lemma 1.7.** With the notation as above and let $E \in S_{\Lambda, \omega}(F)$, then

\begin{align}
(\#_{\text{mod}} - \#_{\text{aut}})(S_{\Lambda, \omega}(F)) &
\leq \eta(E) - \left( r \sum_{i=1}^n s_i(-t_i - 1) + \sum_{i=1}^n s_i^2 \right) - \#_{\text{aut}}(F) - (r^2 - 1)(g - 1).
\end{align}

**Proof.** We only need to prove the inequality

\begin{align}
(1.14) \quad (\#_{\text{mod}} - \#_{\text{aut}})(S_{\Lambda, \omega}(F)) \leq \sum_{i=1}^n (rs_i(-t_i + 1) - s_i^2) - \#_{\text{aut}}(F)
\end{align}

because then (1.13) follows from $c_2(E) = -\sum_{i=1}^n s_i t_i$ and $\eta(E) = -2r \sum_{i=1}^n s_i t_i + (r^2 - 1)(g - 1)$. We prove (1.14) by induction on $n$. When $n = 1$, (1.14) follows from (1.11) and (1.12) because $\#_{\text{aut}}(S_{\Lambda, \omega}(F)) = \sup \{ k | \Theta_k \neq \emptyset \}$. Now assume (1.14) is true for $n - 1$. We divide $S_{x_1, \omega_1}(F)$, $\omega_1 = (t_1^{\oplus s_1})$, into subsets $W_k$ such that $F' \in W_k$ if $\#_{\text{aut}}(F') = k$. Let $\Lambda_2 = (x_2, \cdots, x_k)$ and $\omega_2 = (t_2^{\oplus s_2}, \cdots, t_n^{\oplus s_n})$. Then by the induction hypothesis, for $F' \in W_k$,

\begin{align}
(\#_{\text{mod}} - \#_{\text{aut}})(S_{\Lambda_2, \omega_2}(F')) \leq \sum_{i=2}^n (rs_i(-t_i + 1) - s_i^2) - k
\end{align}

and therefore,

\[ (#_{\text{mod}} - #_{\text{aut}})(S_{\Delta,\omega}(F)) \leq \sup_k \left\{ \sum_{i=2}^{n} (rs_i(-t_i + 1) - s_i^2) - k + #_{\text{mod}}(W_k) \right\} \]

\[ \leq \sum_{i=2}^{n} (rs_i(-t_i + 1) - s_i^2) + (rs_1(-t_1 + 1) - s_1^2 - #_{\text{aut}}(F)) \]

\[ = \sum_{i=1}^{n} (rs_i(-t_i + 1) - s_i^2) - #_{\text{aut}}(F). \]

Now we are ready to prove our structure theorem for subsets of \( \mathcal{A}^{r,d}_{e,I}(0_c). \)

Proof of Theorem 1.5. Inequality (1.8) follows directly from Riemann-Roch and the fact that there is a constant \( C_2 \) depending on \( (\Delta, H, e) \) such that for any \( E \in \mathcal{A}^{r,d}_{e,I}, \)
\[ #_{\text{aut}}(E) \leq C_2. \]

We now prove the second part of the theorem. Let \( S \subseteq \mathcal{A}^{r,d}_{e,I}(1_0) \) be any (irreducible) algebraic set and let \( E \in S \) be a generic element. By Lemma 1.4, after performing a sequence of semistable reduction at \( y_1, \ldots, y_n, \) we get a vector bundle with no jumping line, say \( \pi^*F \) with \( F \) a vector bundle over \( \Sigma. \) Clearly, \( n = n(E) \) depends on \( E. \) We let \( S_0 \subseteq S \) be the open set of \( E' \in S \) with \( n(E') = E \) and let \( n_0 \) be the integer so that when \( E \) varies in \( S_0, \) the number of moduli of the (unordered) set \( y_1, \ldots, y_n \) is \( n_0. \) In other words, \( n_0 \) of \( (y_1, \ldots, y_n) \) are in generic position. We know that the number of moduli of rank \( r \) vector bundles on \( \Sigma \) is \( r^2(g - 1) + 1. \) Also, since \( E \) is \( e \)-stable, \( #_{\text{aut}}(E) \) is bounded by a constant \( C_3 \) independent of \( d \) and \( I \) (see Lemma 1.10). Combining these with (1.13), we get

\[ (1.15) \quad #_{\text{mod}}(S) \leq \eta(E) - r \sum_{i=1}^{n} s_i(-t_i - 1) - \sum_{i=1}^{n} s_i^2 + n_0 - #_{\text{aut}}(F) + g + C_3. \]

Since we have assumed \( #_{\text{mod}}(S) \geq \eta(r, d, I) - C, \) for \( C_3 = C + C_3' + g, \) we get

\[ (1.16) \quad C_3 \geq r \sum_{i=1}^{n} s_i(-t_i - 1) + \left( \sum_{i=1}^{n} s_i^2 - n_0 \right) + #_{\text{aut}}(F). \]

Because \( t_i < 0, \) all terms in (1.16) are non-negative. This immediately gives us \( n - n_0 \leq C_3. \) Next, we define the multiplicity \( m(y_i) \) of \( y_i \) to be the number of appearances of the point \( y_i \) in \( (y_1, \ldots, y_n). \) Then by (1.16),

\[ \frac{1}{2} \#\{ y_i \mid m(y_i) \geq 2 \} \leq \sum_{i=1}^{n} s_i^2 - n_0 \leq C_3. \]

So the total multiplicity of multiple points is bounded. Without loss of generality, we can assume \( y_1, \ldots, y_{n_0} \) are in general position for general \( E \in S. \) For convenience, we call \( y_i \in (y_1, \ldots, y_n) \) a simple point if \( m(y_i) = 1 \) and \( \omega_{y_i}(E) = \omega^{(s_i)} \) is \((-1)^{s_i}. \) We claim that then

\[ (1.17) \quad \sum_{y_i \neq \text{simple}} (-s_i t_i) \leq 2 \sum_{i=1}^{n} s_i(-t_i - 1) + \left( \sum_{i=1}^{n} s_i^2 - n_0 \right) + \#\{ y_i \mid m(y_i) \geq 2 \} \leq \left( \frac{2}{r} + 2 \right) C_3. \]
Indeed, when \( t_i \leq -2 \), then the term \( s_i(-t_i) \) is bounded from above by the term \( 2s_i(t_i - 1) \) in the middle of (1.17), and when \( t_i = -1 \) and \( s_i \geq 2 \), then we have \( -s_i t_i \leq s_i^2 - 1 \). The only remaining situation is when \( m(y_i) \geq 2 \), \( t_i = 1 \) and \( s_i = 1 \). But in this case, \( (-t_i)s_i = 1 \) can be absorbed by the term \( \#\{y_i \mid m(y_i) \geq 2\} \).

Hence, (1.17) holds. Finally, since \( d = \sum_{i=1}^{n} (-t_i)s_i \),

\[
\#\{y_i \mid y_i \text{ simple}\} = d + \sum_{y_i \neq \text{simple}} s_i t_i \geq d - 4C_3.
\]

Therefore, combined with \( n - n_0 \leq C_3 \), we get

\[
(1.18) \quad d \geq n \geq n_1 = \#\{y_i \mid y_i \text{ simple}, 1 \leq i \leq n_0\} \geq d - 5C_3.
\]

Now we let \( l = [5C_3] + 1 \). Without loss of generality, we can assume \( \{y_1, \ldots, y_{d-1}\} \) are simple points in \( \{y_1, \ldots, y_{n_0}\} \). Then the sheaf \( E \) must belong to the exact sequence

\[
0 \rightarrow \pi^*F \rightarrow E \rightarrow \left( \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{y_i}}(-1) \right) \oplus J' \rightarrow 0.
\]

To prove the proposition, we need to have an estimate on \( F \) and \( J' \). By definition, \( J' \) admits a filtration

\[
0 = J_{d-l} \subseteq J_{d-l+1} \subseteq \cdots \subseteq J_n = J'
\]

such that \( J_{i+1}/J_i \cong \mathcal{O}_{P_{y_i}}(t_i)^{\oplus s_i} \). Thus there is a zero scheme \( z' \subseteq \Sigma \) supported on \( \{y_{d-l+1}, \ldots, y_n\} \) of length \( l(z') \leq n - (d - l) \leq 5C_3 \) (because of (1.18)) such that \( J' \) is an \( \mathcal{O}_{\Sigma_{d-l}}(z') \)-module and further,

\[
0 \leq c_1(E) \cdot \Sigma^+ - (\deg F + d - l) = c_1(J') \cdot \Sigma^+
\]

\[
= \sum_{i=d-l+1}^{n} s_i \leq \sum_{i=d-l+1}^{n} (-t_i)s_i \leq 5C_3.
\]

(1.20)

Here, the last inequality holds because of (1.17) and \( n - n_0 \leq C_3 \). Also, since \( \#_{\text{aut}}(F) \leq C_3 \) (from (1.16)), there is a constant \( C_4 \) such that \( F \) is \( C_4 \)-stable.

It remains to show that we can find an integer \( l_1 \) (independent of \( d \)) and find a single line bundle \( L \) of degree \( [(d - c)/r] + l_1 \) \( (c = I \cdot \Sigma^+) \) so that for any \( E \in S_0 \), \( E \) belongs to the exact sequence

\[
0 \rightarrow E^\vee \rightarrow \pi^*(L^\oplus r) \rightarrow \left( \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{y_i}}(1) \right) \oplus J \rightarrow 0
\]

specified in Theorem 1.5. First, there is a constant \( l_1 \) and a line bundle \( L \) of degree \( [(d - c)/r] + l_1 \) such that for any \( C_4 \)-stable rank \( r \) vector bundle \( F \) on \( \Sigma \) satisfying (1.20), \( L \otimes F \) is generated by \( H^0(L \otimes F) \). Now for any \( E \in S_0 \) with the data given by (1.19), we choose \( \pi^*F^\vee \rightarrow \pi^*L^\oplus r \) so that the support of \( \pi^*(L^\oplus r)/\pi^*F^\vee \) is disjoint
from \( \bigcup_{i=1}^{d-l} P_{y_i} \). Then by dualizing (1.19) and coupled with \( \pi^*F^\vee \to \pi^*(L^{\otimes r}) \), we get

\[
0 \to E^\vee \to \pi^*(L^{\otimes r}) \to J \oplus \left( \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{y_i}}(1) \right) \to 0.
\]

Finally, it is easy to see that there is an integer \( l_2 \) depending only on \( C_3 \) and \( l_1 \) such that for some subscheme \( z \subseteq \Sigma \) of length \( \ell(z) \leq l_2 \), \( J \) is a sheaf of \( \mathcal{O}_{\pi^{-1}(z)} \)-modules. This completes the proof of the theorem.

Now we prove Theorem 1.2.

**Proof of Theorem 1.2.** We begin with a general vector bundle \( E \in \mathcal{A}_{e,l}(\alpha), \alpha = (\alpha_1^{\otimes r_1}, \cdots, \alpha_n^{\otimes r_n}) \neq 1_m \). Let

\[
(1.22) \quad 0 = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_n = E
\]

be the relative Harder-Narasimhan filtration such that \( F_i = E_i/E_{i-1} \) are torsion free of generic fiber types \( (\alpha_i^{\otimes r_i}) \) respectively. We call this the relative filtration of \( E \). ((1.22) can be derived by using the usual Harder-Narasimhan filtration of \( E \) with respect to the divisor \( kP_x + \Sigma^+ \) with \( k \gg 0 \).)

We fix \( F_i = E_i/E_{i-1} \) and let \( W(\{F_i\}_1^n) \) be the set of all vector bundles \( V \) whose relative filtrations \( 0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V \) satisfy \( V_i/V_{i-1} \cong F_i \). Our first step is to estimate the number of local moduli \( \#_{\text{loc}}(W(\{F_i\}_1^n)) \) at \( [E] \). Let \( A_i = \frac{1}{r_i} c_i(F_i) \) and \( d_i = c_2(F_i) - (\frac{1}{2})^2 A_i^2 \). Note that by the proof of Lemma 1.4, \( d_i \geq 0 \). Now an easy calculation shows that

\[
(1.23) \quad d = c_2(E) = \frac{1}{2} \sum_{i=1}^{n} r_i(I - A_i). A_i + \sum_{i=1}^{n} d_i = \frac{1}{2} I^2 - \frac{1}{2} \sum_{i=1}^{n} r_i A_i^2 + \sum_{i=1}^{n} d_i.
\]

From the exact sequence

\[
0 \to E_{n-1} \to E \to F_n \to 0
\]

and the argument similar to (1.11), we have

\[
\#_{\text{loc}}(W(\{F_i\}_1^n)) \leq \#_{\text{mod}}(W(\{F_i\}_1^{n-1})) + \dim \text{Ext}^1(F_n, E_{n-1}) - \#_{\text{aut}}(F_n) - \#_{\text{aut}}(E) - \dim \text{Hom}(F_n, E_{n-1}).
\]

Further, because \( E_{n-1}^\vee \otimes F_n \) has generic fiber type \( ((\alpha_n - \alpha_1)^{\otimes r_1}, \cdots, (\alpha_n - \alpha_{n-1}^\otimes)^{\otimes r_{n-1}}) \) and \( \alpha_{i+1} < \alpha_i \), \( \text{Ext}^2(F_n, E_{n-1}) = 0 \) by Serre duality. Hence

\[
(1.25) \quad \dim \text{Hom}(F_n, E_{n-1}) - \dim \text{Ext}^1(F_n, E_{n-1}) = \chi(F_n, E_{n-1}),
\]

where the right-hand side of (1.25) is the abbreviation of \( \chi(\text{Ext}^1(F_n, E_{n-1})) \). Finally, by using the filtration (1.22), we have

\[
\#_{\text{loc}}(W(\{F_i\}_1^n)) - \#_{\text{aut}}(E) \leq \#_{\text{mod}}(W(\{F_i\}_1^{n-1})) + \sum_{i=1}^{n-1} \chi(F_n, E_{n-1}) - \#_{\text{aut}}(E) - \#_{\text{aut}}(F_n)
\]

\[
(1.26) \quad \leq \sum_{i=1}^{n} \chi(F_i, F_j) - \sum_{i=1}^{n} \#_{\text{aut}}(F_i).
\]
The last inequality is derived by iterating the first part of (1.26). Therefore,

\[
\#_{\text{mod}} \mathcal{U}_{e,i}^d(\alpha) \leq \sup \left\{ \sum_{i<j} \chi(F_i, F_j) + \sum_{i=1}^n \left( \#_{\text{mod}}(F_i) - \#_{\text{aut}}(F_i) \right) \right\} \\
+ \max \{ \#_{\text{aut}}(E) \mid E \in \mathcal{U}_{e,i}^d(\alpha) \},
\]

where the sup is taken over all possible relative filtrations (1.22) of \( E \)'s in \( \mathbb{A}_{e,i}^r(\alpha) \).

We now calculate the right-hand side of (1.27) by Riemann-Roch. First,

\[
\chi(F_i, F_j) = r_ir_j \left( \frac{1}{2}(A_j - A_i)^2 - \frac{1}{2}(A_j - A_i) \cdot K_{\Delta} + (1 - g) \right) - r_id_j - r_jd_i.
\]

For simplicity, in the following we will group all terms that are bounded independently of \( r_i, d_i, A_i \) and \( \alpha_m \neq 1_m \) into \( O(1) \). We have

\[
\sum_{i>j} \chi(F_i, F_j)
= - \sum_{i>j} r_ir_j \left( \frac{1}{2}(A_j - A_i)^2 - \frac{1}{2}(A_j - A_i) \cdot K_{\Delta} + \frac{d_i}{r_i} + \frac{d_j}{r_j} \right) + O(1).
\]

Further, one calculates

\[
\eta(F_i) = 2r_i (d_i + \left( \frac{r_i}{2} \right) A_i^2) - (r_i - 1)r_i^2 A_i^2 - (r_i^2 - 1)(1 - g)
= 2r_i d_i - (r_i^2 - 1)(1 - g).
\]

Thus by combining (1.23), (1.27)–(1.29) and the fact that \( \#_{\text{mod}}(F_i) - \#_{\text{aut}}(F_i) \leq \chi(F_i, F_i) \), we obtain

\[
(\#_{\text{mod}}(E) - \#_{\text{aut}}(E)) (\mathcal{U}_{e,i}^r(\alpha)) - (2r - 1)d
\leq - \sum_{i>j} r_ir_j \left( \frac{1}{2}(A_j - A_i)^2 - \frac{1}{2}(A_j - A_i) \cdot K_{\Delta} + \frac{d_i}{r_i} + \frac{d_j}{r_j} \right)
+ \sum_{i=1}^n 2r_id_i - (2r - 1) \left( \sum_{i=1}^n d_i - \frac{1}{2} \sum_{i=1}^n r_i A_i^2 \right) + O(1).
\]

To analyze (1.30), we first note that

\[
\sum_{i>j} r_ir_j (A_j - A_i)^2 = r \sum_{i=1}^n A_i^2 - r_i^2; \quad \sum_{i>j} r_ir_j \left( \frac{d_i}{r_i} + \frac{d_j}{r_j} \right) = \sum_{i=1}^n (r - r_i) d_i.
\]

Now if we let \( A_i \sim \alpha_i \Sigma^- + c_i P_\xi \) and let \( \delta = \Sigma^+ \cdot \Sigma^+ \), then the right-hand side of (1.30) is equal to

\[
\sum_{i=1}^n \left( \frac{1}{2} (r - 1)r_i A_i^2 - (r - r_i - 1)d_i \right) + \frac{1}{2} \sum_{i>j} r_ir_j (A_j - A_i) \cdot K_{\Delta} + O(1)
= \sum_{i=1}^n \frac{1}{2} (r - 1) r_i (-\delta \alpha_i^2 + 2\alpha_i c_i) - \sum_{i=1}^n (r - r_i - 1)d_i
+ \frac{1}{2} \sum_{i>j} r_ir_j (\alpha_j - \alpha_i) \Sigma^- \cdot K_{\Delta} + \frac{1}{2} \sum_{i>j} r_ir_j (c_j - c_i) P_\xi \cdot K_{\Delta} + O(1).
\]
which is bounded from above by (note \( d_i \geq 0 \))

\[
- \frac{1}{4} \delta \sum_{i=1}^{n} \alpha_i^2 + (r - 1) \sum_{i=1}^{n} r_i \alpha_i c_i - \sum_{i>j} r_i r_j (c_j - c_i) + O(1) \\
(1.31)
\]

\[
= - \frac{1}{4} \delta \sum_{i=1}^{n} \alpha_i^2 + \sum_{k=1}^{n} \left( r_k c_k (\alpha_k + \sum_{i=1}^{k-1} r_i - \sum_{i=k+1}^{n} r_i) \right) + O(1).
\]

Let \( p_k = (r - 1) \alpha_k + \sum_{i=1}^{k-1} r_i - \sum_{i=k+1}^{n} r_i \). Then when \( \alpha \neq 1_m \), or equivalently when \( n \geq 2 \) or \( n = 2 \) and \( \alpha_1 - \alpha_2 \geq 2 \), we will always have

\[
p_k - p_{k+1} = (r - 1)(\alpha_k - \alpha_{k+1}) - (r_k + r_{k+1}) \geq 0, \quad k \leq n - 1.
\]

We rewrite

\[
\sum_{k=1}^{n} \left( r_k c_k (\alpha_k + \sum_{i=1}^{k-1} r_i - \sum_{i=k+1}^{n} r_i) \right) = \sum_{k=1}^{n-1} \left( (p_k - p_{k+1}) (\sum_{i=1}^{k} r_i c_i) \right) + \sum_{k=1}^{n} r_i c_i.
\]

Finally, we shall make use of the fact that \( E \) is \( e \)-stable. If \( H \sim a \Sigma + b P \), then for any \( k \),

\[
\deg(E_k) = a \sum_{i=1}^{k} r_i c_i + b \sum_{i=1}^{k} r_i \alpha_i \leq \frac{r_k(E_k)}{r} H \cdot I + e \sqrt{H}.
\]

Therefore, for \( k \leq n - 1 \),

\[
\sum_{i=1}^{k} r_i c_i \leq 1 + \frac{b}{a} + e \sqrt{\delta} + 2 \frac{b}{a} - \frac{b}{a} \sum_{i=1}^{k} r_i \alpha_i.
\]

Thus we get

\[
\sum_{k=1}^{n} \left( (p_k - p_{k+1}) (\sum_{i=1}^{k} r_i c_i) \right) \\
(1.32) \leq \sum_{k=1}^{n-1} (p_k - p_{k+1}) \left( \frac{a + b + e \sqrt{2} \delta + 2ab}{a} - \frac{b}{a} \sum_{i=1}^{k} r_i \alpha_i \right) + p_m m \\
\leq \frac{b}{a} \cdot r \left( \sum_{i=1}^{m} |\alpha_i| \right)^2 + 4r^2(2 + e\delta)(1 + \frac{b}{a}) \left( \sum_{i=1}^{m} |\alpha_i| \right) + O(1).
\]

Here we have used the fact that \( p_k - p_{k+1} \leq \sum_{i=1}^{m} |\alpha_i| + r \) and \( p_m \leq 0 \) because \( \alpha \neq 1_m \) and \( \sum_{i=1}^{m} r_i c_i = m \geq 0 \). Now if we assume

\[
\frac{b}{a} \cdot r^2 < \frac{1}{16},
\]

then everything in (1.32) can be absorbed by the quadratic term \(-\frac{1}{4} \delta \sum_{i=1}^{n} \alpha_i^2 \) (in (1.31)) with the help of some constant \( C_1 \). Thus combined with (1.31), we have proved

\[
\#_{\text{mod}} \mathcal{A}_{e,I}^{d} (a) \leq (2r - 1)d + C_1 + \max\{ \#_{\text{aut}}(E) \mid E \in \mathcal{A}_{e,I}^{d} (a) \}.
\]

Theorem 1.2 will be proved if we can bound \( \text{Hom}(E, E) \) for \( E \in \mathcal{A}_{e,I}^{d} \). Since \( E \) is \( e \)-stable, \( E^\vee \otimes E \) must be \((2|e| + 1)\)-stable. (This can be proved by using the fact that the Harder-Narasimhan filtration of \( E \) will induce the Harder-Narasimhan filtration of \( E^\vee \otimes E \).) Thus \( \#_{\text{aut}}(E) \) is bounded independently of \( d \) by the following lemma.
Lemma 1.8. For constants $e_1$, $e_2$ and integer $r$, there is a constant $C'$ such that whenever $V$ is a rank $r$ $e_1$-stable vector bundle on $\Delta$ such that $|\deg(V)| \leq e_2$, then we have $\dim H^0(V) \leq C'$.

Proof. We prove the lemma by induction on $r$. The case $r = 1$ is obvious. Assume the lemma is true for vector bundles of rank $\leq r-1$ and assume $V$ has $H^0(V) \neq \{0\}$. Then there is a line bundle $L$, $\deg L \geq 0$, such that $V$ belongs to the exact sequence

$$0 \longrightarrow L \longrightarrow V \longrightarrow V/L \longrightarrow 0$$

with $V/L$ torsion free. Since $V$ is $e_1$-stable and $|\deg(V)| \leq e_2$, there are constants $e'_1$ and $e'_2$ such that $|\deg L|, |\deg V/L| \leq e'_2$ and $V/L$ is $e'_1$-stable. Thus by the induction hypothesis, there is a constant $C'$ such that $h^0(L) \leq C'$ and $h^0(V/L) \leq C'$. The lemma then follows.

We now prove Theorem 1.1 by induction on $m$. We first establish the case $m = 0$. Let $e$ and $C$ be any constants, $r \geq 2$ be an integer and $D \subseteq \Delta$ be any divisor. We assume $H$ is an ample divisor satisfying the condition of Theorem 1.2. To prove the theorem, we need to show that there is a constant $N$ depending only on $(X, H, r, I, e, D)$ so that if for some $d$ we have

$$(1.33) \quad \#_{\text{mod}} \mathfrak{A}_{e, l}^d(D) \geq \eta(r, d, I) - C,$$

then $d \leq N$. Now assume (1.33) does hold. Thanks to Theorem 1.2, there is an $N_1 \geq 0$ such that if $d \geq N_1$, then the set $\mathfrak{A}_{e, l}^d(D, 1_0)$ satisfies

$$(1.34) \quad \#_{\text{mod}} \mathfrak{A}_{e, l}^d(D, 1_0) = \#_{\text{mod}} \mathfrak{A}_{e, l}^d(D) \geq \eta(r, d, I) - C.$$ 

Of course, $\mathfrak{A}_{e, l}^d(D, 1_0)$ is a constructible set. Let $S$ be an irreducible variety parameterizing a subset of $\mathfrak{A}_{e, l}^d(D, 1_0)$ such that $\#_{\text{mod}} S \geq \eta(r, d, I) - C$. By Theorem 1.5, there are constants $t$, $l_1$, $l_2$ (independent of $d$) and a line bundle $L$ of degree $[(d - c)/r] + l_1$ such that associated to a general $E \in S$, there are $x_1, \cdots, x_{d-l} \in \Sigma$ in general position and a quotient sheaf $J_E$ of $\pi^*(L^\oplus r)$ such that $E$ belongs to the exact sequence

$$(1.35) \quad 0 \longrightarrow E' \longrightarrow \pi^*(L^\oplus r) \longrightarrow J_E \oplus \left( \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{x_i}}(1) \right) \longrightarrow 0.$$ 

Clearly, $E$ is determined by the surjective homomorphisms

$$\pi^*(L^\oplus r) \overset{\tau_0}{\longrightarrow} J_E \quad \text{and} \quad \pi^*(L^\oplus r) \overset{\tau_1}{\longrightarrow} \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{x_i}}(1).$$

Hence the combined number of moduli of the sets of these quotient sheaves that come from $E \in S$ is no less than $\eta(r, d, I) - C$. Let

$$\Xi_0 = \{ \tau_0 : \pi^*(L^\oplus r) \to J_E \mid E \in S \},$$

$$\Xi_1 = \{ \tau_1 : \pi^*(L^\oplus r) \to \bigoplus_{i=1}^{d-l} \mathcal{O}_{P_{x_i}}(1) \mid E \in S \}.$$ 

Because of the following lemma, the information contained in $\Xi_0$ is minimal.
Lemma 1.9. There is a constant $C_5$ independent of $d$ such that $\#_{\text{mod}}(\Xi_0) \leq C_5$.

Proof. We first calculate the Hilbert polynomials of the sheaves $J_E$. Let $J_E(n) = J_E \otimes H^\otimes n$. Then

$$\chi(J_E(n)) = \chi(\pi^*L^{\otimes r}(n)) - \chi(E(n)) - (d-1)\chi(O_{P_z}(1) \otimes H^\otimes n) = a_1(d)n + a_0(d),$$

where $a_1(d) = (r[(d-c)/r] - d + rl_1 + l) \cdot (H \cdot P_k) - I \cdot H$ and $a_0(d) = (r[(d-c)/r] - d) + rl_1 - \frac{1}{2}r^2 + \frac{1}{2}r \cdot K + 2l$. Since $r[(d-c)/r] - d$ can only attain integer values between $-c - r$ and $-c$ for integers $d$, the function $a_1(d)$ (resp. $a_0(d)$) can only attain $r$ values. Hence, $\{\chi(J_E(\cdot)) \mid E \in S\}$ is a finite set (independent of $d$) and by [Gr, p. 12], the set $\Xi_0$ is bounded. Thus, there is a constant $C_5$ such that $\#_{\text{mod}}\Xi_0 \leq C_5$.

Since $\#_{\text{mod}}S \leq \#_{\text{mod}}\Xi_0 + \#_{\text{mod}}\Xi_1$, we have

$$\#_{\text{mod}}\Xi_1 \geq \eta(r, d, I) - (C + C_5).$$

Let $\tau_1 \in \Xi_1$ and $F = \ker\{\tau_1\}$. In the following, we seek to relate the non-vanishing of $\Hom(E, E(D))^0$ to the non-vanishing of $\Hom(F, F(D'))^0$ for some divisor $D'$. First of all, by (2) of Theorem 1.5, there is a divisor $z \in \Sigma$ (of degree $\leq l_2$) such that the composition

$$F(-\pi^{-1}(z)) \hookrightarrow \pi^*L(-\pi^{-1}(z))^{\otimes r} \longrightarrow \pi^*L^{\otimes r} \longrightarrow J_E$$

is trivial. Because of (1.35), $F(-\pi^{-1}(z))$ is a subsheaf of $E^\vee$. Therefore, any non-trivial traceless homomorphism $\varphi: E \rightarrow E(D)$ will provide us a non-trivial traceless homomorphism

$$F(-\pi^{-1}(z)) \longrightarrow E^\vee \xrightarrow{\varphi} E^\vee(D) \longrightarrow F(D).$$

Further, let $\bar{\xi}$ be a fixed divisor on $\Sigma$ of degree $l_2 + 2g$. Since

$$h^0(\Sigma, O_\Sigma(\bar{\xi} - z)) \neq 0,$$

$\Hom(F, F(D + \pi^{-1}(z))) \neq 0$ implies $\Hom(F, F(D + \pi^{-1}(\bar{\xi}))) \neq 0$. Thus we have proved:

Lemma 1.10. With the notation as before, there is a divisor $z \in \Sigma$ independent of $d$ and $D$ such that for any sheaf $F = \ker\{\tau_1\}$, where $\tau_1 \in \Xi_1$, and for $D_1 = D + \pi^{-1}(z)$, we have $\Hom(F, F(D_1)) \neq \{0\}$. \hfill $\square$

Our next step is to investigate the set $\Xi_1$ by utilizing this non-vanishing property. We first fix $d-l$ general points $x_1, \ldots, x_{d-l} \in \Sigma$ and let $U$ be the set of all quotient homomorphisms

$$\sigma: \pi^*(L^{\otimes r}) \longrightarrow \bigoplus_{i=1}^{d-l} O_{P_{x_i}}(1).$$

$U$ is (canonically) parameterized by an open subset of the product of $d-l$ copies of projective space $\mathbb{P}^{2r-1}$ after fixing the basis of each $H^0(O_{P_{x_i}}(1))$. In the following, for any $u \in \Pi^{d-l}\mathbb{P}^{2r-1}$ we denote by $\sigma_u$ the associated homomorphism $\sigma_u: \pi^*(L^{\otimes r}) \rightarrow \bigoplus_{i=1}^{d-l} O_{P_{x_i}}(1)$. Let

$$\Xi_1(\Phi) = \{u \in \Pi^{d-l}\mathbb{P}^{2r-1} \mid \sigma_u \in \Xi_1\}.$$
Since the points of \( y = (x_1, \cdots, x_{d-l}) \) are general,

\[
\dim \Xi_1(y) \geq \# \text{mod } \Xi_1 - (d-l)
\geq \eta(r, d, l) - (C + C_6) - (d-l) \geq (2r-1)d - C_6
\]

for some integer \( C_6 \). Now let \( l_3 = l + C_6 + 1 \). Possibly after rearranging the order of \( (x_1, \cdots, x_{d-l}) \), we can further assume that the restriction to \( \Xi_1(y) \) of the projection from \( \Pi^{d-1} \mathbf{P}^{2r-1} \) to the first \( d - l_3 \) factors,

\[
\Xi_1(y) \subseteq \Pi^{d-l_3} \mathbf{P}^{2r-1} \rightarrow \Pi^{d-l_3} \mathbf{P}^{2r-1},
\]

is dominant. That is, for general \( v \in \Pi^{d-l_3} \mathbf{P}^{2r-1} \) with the associated homomorphism \( \sigma'_v : \pi^*L^{\oplus r} \rightarrow \bigoplus_{1=1}^{d-l_3} \mathcal{O}_{P_{x_i}}(1) \), there is at least one

\[
\xi : \pi^*L^{\oplus r} \rightarrow \bigoplus_{i=d-l_3+1}^{d-l} \mathcal{O}_{P_{x_i}}(1)
\]

such that \( \sigma'_v \oplus \xi \) considered as a quotient sheaf belongs to \( \Xi_1(y) \). Thus if we let \( V = \ker \{ \sigma'_v \oplus \xi \} \) and let \( V_v = \ker \{ \sigma'_v \} \), \( V \) and \( V_v \) fit into the exact sequence

\[
0 \rightarrow V \rightarrow V_v \rightarrow \bigoplus_{i=d-l_3+1}^{d-l} \mathcal{O}_{P_{x_i}}(1) \rightarrow 0.
\]

Let \( A = \bigcup_{i=d-l_3+1}^{d-l} P_{x_i} \) be a divisor in \( \Delta \). Following the argument in Lemma 1.10, the non-trivial homomorphism \( \phi \) in Lemma 1.10 induces a non-trivial homomorphism

\[
\phi' : V_v \rightarrow V_v(D_1 + A).
\]

Therefore for general \( v \in \Pi^{d-l_3} \mathbf{P}^{2r-1} \), \( \text{Hom}_{\Delta}(V_v, V_v(D_1 + A))^0 \neq \{0\} \). Finally, as in Lemma 1.9, for any fixed divisor \( A_0 \subseteq \Delta \) that consists of \( l_3 + 2g \) fibers of \( \Delta \), we must have \( \text{Hom}(V_v, V_v(D_1 + A_0))^0 \neq \{0\} \) as well. Therefore, Theorem 1.1 (when \( m = 0 \)) follows from

**Proposition 1.11.** For any divisor \( D \subseteq \Delta \) and any integer \( l_0 \), there is a constant \( N \) for which the following holds: Assume \( d \geq N \), that \( L \) is a line bundle on \( \Sigma \) of degree \( L = [d/r] + l_0 \) and that \( x_1, \cdots, x_d \) are general points in \( \Sigma \). Then for general \( v \in \Pi^{d} \mathbf{P}^{2r-1} \), the sheaf \( E_v = \ker \{ \sigma_v \} \), where \( \sigma_v \) is the associated homomorphism \( \pi^*L^{\oplus r} \rightarrow \bigoplus_{i=1}^{d} \mathcal{O}_{P_{x_i}}(1) \), satisfies \( \text{Hom}(E_v, E_v(D))^0 = \{0\} \).

**Proof.** We prove it by contradiction. The trick is to first prove the vanishing of this homomorphism group for a special quotient sheaf and then apply the semicontinuity theorem to derive the general case. Let \( p_i \in P_{x_i} \) be a general closed point and let \( U \) be a small disk containing 0. There is a torsion free sheaf \( J_t \) on \( P_{x_i} \times U \) flat over \( U \) such that \( J_t|_{P_{x_i} \times \{0\}} \cong \mathcal{O}_{P_{x_i}} \oplus \mathbb{C}_{p_i} \) (\( \mathbb{C}_{p_i} \) is the skyscraper sheaf supported on \( p_i \)) and for \( t \neq 0 \), \( J_t|_{P_{x_i} \times \{t\}} \cong \mathcal{O}_{P_{x_i}}(1) \). It is easy to see that any surjective homomorphism

\[
f_i : (\mathcal{O}_{P_{x_i}})^{\oplus r} \rightarrow \mathcal{O}_{P_{x_i}} \oplus \mathbb{C}_{p_i}
\]
can be extended to a (surjective) homomorphism

$$F_i : \left( \mathcal{O}_{P_n} \right)_{\otimes r} \longrightarrow J_i.$$ 

In general, we can extend any surjective homomorphism

$$(1.40) \quad f : \pi^* (L \otimes r) \longrightarrow \bigoplus_{i=1}^{d} \pi^* (L \otimes r)_{|P_{x_i}} \longrightarrow \bigoplus_{i=1}^{d} (\mathcal{O}_{P_{x_i}} \oplus C_{p_i})$$

to a (surjective) homomorphism

$$F : \pi^* (L \otimes r) \otimes \mathcal{O} \Delta \longrightarrow \bigoplus_{i=1}^{d} J_i.$$ 

Let

$$V_t = \ker \{ F[\Delta \times (t) : 1] \}.$$

Then $V_t$ is a flat family of torsion free sheaves on $\Delta$ parameterized by $U$. Assuming, for general $\sigma_v : \pi^* (L \otimes r) \rightarrow \bigoplus_{i=1}^{d} \mathcal{O}_{P_{x_i}} (1)$, $\text{Hom}(E_v, E_v(D)) \neq \{0\}$, then by the semicontinuity theorem, $\text{Hom}(V_t, V_t(D)) \neq \{0\}$ for $t \neq 0$ and consequently, $\text{Hom}(V_0, V_0(D)) \neq \{0\}$.

Now we seek to find a contradiction by choosing $V_0$ (i.e. $f$ in (1.40)) carefully. We first divide the set $\{x_1, \cdots, x_d\}$ into $2r$ subsets, say $\Lambda_1, \cdots, \Lambda_{2r}$, such that each contains either $[d/2r]$ or $[d/2r] + 1$ points. We write $f_i = f_i^1 \oplus f_i^2$ according to (1.39). For $x_i \in \Lambda_{2k-1}$, we define $f_i^1$ to be the composition

$$(1.41) \quad f_i^1 : \pi^* (L \otimes r) \stackrel{\text{rest}}{\longrightarrow} \pi^* (L \otimes r)_{|P_{x_i}} \stackrel{pr_k}{\longrightarrow} \pi^* (L)_{|P_{x_i}},$$

where $pr_k$ is the projection onto the $k$th component, and define $f_i^2$ to be the composition

$$f_i^2 : \pi^* (L \otimes r) \stackrel{\text{rest}}{\longrightarrow} \pi^* (L \otimes r)_{|P_{x_i}} \stackrel{pr_k+1 \oplus pr_k}{{\longrightarrow}^\pi} \pi^* (L)_{|P_{x_i}} \stackrel{ev}{\longrightarrow} C_{p_i},$$

where $ev : \pi^* (L)_{|P_{x_i}} \rightarrow C_{p_i}$ is the evaluation map. (Here we agree $pr_{r+1} = pr_1$.)

For $i \in \Lambda_{2k}$, we define $f_i^1$ as in (1.41) while we let $f_i^2$ be

$$f_i^2 : \pi^* (L \otimes r) \stackrel{\text{rest}}{\longrightarrow} \pi^* (L \otimes r)_{|P_{x_i}} \stackrel{pr_k+1 \oplus pr_k+2}{{\longrightarrow}^\pi} \pi^* (L)_{|P_{x_i}} \oplus \pi^* (L)_{|P_{x_i}} \stackrel{ev+ev}{\longrightarrow} C_{x_i}.$$ 

($pr_{r+2} = pr_2$.) We claim that when $d$ is sufficiently large, the sheaf $E \subseteq \pi^* (L \otimes r)$ that is the kernel of $\bigoplus_{i=1}^{d} (f_i^1 \oplus f_i^2)$ has $\text{Hom}(E, E(D)) \neq \{0\}$. Indeed, let

$$L_k = L(- \sum_{i \in \Lambda_{2k-1} \cup \Lambda_{2k}} x_i)$$

be the line bundle on $\Sigma$ of degree between $l_0 - 2$ and $l_0 + 1$ and let $L_k = \pi^* L_k$. Then, $E$ is a subsheaf of $\bigoplus_{i=1}^{r} L_k$ with cokernel $\bigoplus_{i=1}^{r} C_{p_i}$. Let $s \in \text{Hom}(E, E(D))$. Then $s$ induces a homomorphism

$$(s_{ij})_{r \times r} : \bigoplus_{h=1}^{r} L_h \longrightarrow \bigoplus_{h=1}^{r} L_h(D)$$
with $s_{ij} \in H^0(L_j^{-1} \otimes L_j(D))$. Since $L_j^{-1} \otimes L_j$ is a pull back of the line bundle on $\Sigma$ that has degree $-1$, 0 or 1, $h^0(L_j^{-1} \otimes L_j(D))$ is bounded by a constant $C_6$ independent of $d$. On the other hand, by our construction of $E$, when $i \in \Lambda_{2k-1}$, the composition

$$
\bigoplus_{h=1}^r L_h \xrightarrow{(s_{**})} \bigoplus_{h=1}^r L_h \xrightarrow{pr_{k+1}} L_{k+1} \xrightarrow{ev} C_{p_i}
$$

is trivial. Hence for $j \neq k+1$, $s_{jk+1}$ vanishes on $p_i$ for all $i \in \Lambda_{2k-1}$. Now we let $N = 2r([C_6] + 3)$ and assume $d \geq N$. Because $p_i$ are general and

$$\#(\Lambda_{2k-1}) \geq \lfloor d/2r \rfloor > C_6 + 2 > h^0(L_j^{-1} \otimes L_{k+1}) + 1,$$

$s_{jk+1}$ must be 0 for $j \neq k+1$.

It remains to show that $s = g_0 \cdot \text{id}$ for some $g_0 \in H^0(O(D))$. Let $g_j \in H^0(O(D))$ be sections so that $s_{jj} = g_j \cdot \text{id} : L_j \rightarrow L_j(D)$. Let $i \in \Lambda_{2k}$ and let

$$v_i \neq 0 \in \ker \{ (L_{k+1} + L_{k+2}) | p_i \}.$$

Then because $(s_{ij}) = \text{diag}(s_{11}, \cdots, s_{rr})$ is induced from $s \in \text{Hom}(E, E(D))$, we must have

$$(ev, ev) \circ (pr_{k+1} \oplus pr_{k+2}) \circ (s_{**})v_i = 0.$$

It is straightforward to check that this is equivalent to $(g_{k+1} - g_{k+2})(p_i) = 0$. Hence, because $p_i$ are general and $\#(\Lambda_{2k}) > h^0(O(D)) + 1$, we must have $g_{k+1} = g_{k+2}$. Therefore, Hom$(E, E(D))^0 = 0$. This completes the proof of the theorem for $m = 0$.

Now we use induction on $m$ to establish the remaining cases. The strategy is as follows: We first fix a section $\Sigma^+ \subseteq \Delta$ of $\pi : \Delta \rightarrow \Sigma$ of positive self-intersection $\delta$. Let $E \in \mathfrak{A}^{r,d}(D, 1_m)$ be any sheaf. We choose a quotient sheaf $E_{|\Sigma^+} \rightarrow L_E$ with $L_E$ a locally free sheaf of $O_{\Sigma^+}$-modules and define $\tilde{E} = \ker \{ E \rightarrow L_E \}$. $\tilde{E}$ is locally free with Chern classes

$$I' = \det(\tilde{E}) = I(-\Sigma^+), \quad r_0 = \text{rank} L;$$

$$d' = c_2(\tilde{E}) = d + \deg L_E + \frac{1}{2} r_0 (r_0 - 1)\delta - r_0(I \cdot \Sigma^+).$$

Moreover, $\tilde{E} \in \mathfrak{A}^{r',d'}(D + \Sigma^+)$ for a constant $e'$ independent of $L$ and $d$. Hence by applying the induction hypothesis to $\mathfrak{A}^{r',d'}_{e',l}(D + \Sigma^+)$, we get an upper bound of $\#_{\text{mod}} \{ E \mid E \in \mathfrak{A}^{r',d'}_{e',l}(D) \}$. Thus if we understand the correspondence $E \rightarrow \tilde{E}$ well, we can translate the estimate of $\#_{\text{mod}} \{ E \mid E \in \mathfrak{A}^{r,d}_{e,l}(D) \}$ to the estimate of $\#_{\text{mod}} \mathfrak{A}^{r,d}_{e,l}(D)$. We now give the details of this argument.

First, we choose $e_0 > 0$ so that $h^1(\Sigma, F) = 0$ holds for all semistable vector bundles $F$ on $\Sigma$ with $\text{rk}(F) \leq r^2$ and $\mu(F) \geq e_0$. Put $e_1 = r(e_0 + \delta)$. There is a decomposition of $\mathfrak{A}^{r,d}_{e,l}(D, 1_m)$ according to whether the restriction of an element $E \in \mathfrak{A}^{r,d}_{e,l}(D, 1_m)$ to $\Sigma^+$ is $e_1$-stable or not. We denote these sets by $W^+$ and $W^-$ respectively. Let $L_0$ be a line bundle on $\Sigma^+$ such that $H^0_{\Sigma^+}(F^\vee \otimes L_0)$ generates $F^\vee \otimes L_0$ for any $e_1$-stable rank $r$ vector bundle $F$ on $\Sigma^+$ of degree $I \cdot \Sigma^+$. Then for
any $E \in W^+$, we let $L_E = L_0$ and fix a surjective homomorphism $\sigma : E \to L_E$. In case $E \in W^-$, we let

$$0 = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_k = E_{|\Sigma^+}$$

be the Harder-Narasimhan filtration of $E_{|\Sigma^+}$. That is, $F_{i+1}/F_i$ are semistable and $\mu(F_i/F_{i-1}) > \mu(F_{i+1}/F_i)$. We let $i_0$ be the largest integer so that

$$\mu(F_{i_0}/F_{i_0-1}) \geq \mu(F_{i_0+1}/F_{i_0}) + e_0.$$ 

Such $i_0$ exists because $E_{|\Sigma^+}$ is not $e_1$-stable and $e_1 > re_0$. Then by our choice of $e_0$, $E_{|\Sigma^+} \cong M_E \oplus L_E$, $M_E = F_{i_0}$ and $L_E = E_{|\Sigma^+}/F_{i_0}$.

We choose our quotient sheaf to be $\sigma : E \to L_E$. Note that $L_E$ is $re_0$-stable and has degree $\leq \frac{r}{\varphi} I \cdot \Sigma^+$.

Now let $\tilde{E}$ be the kernel of $E \to L_E$. Then $\tilde{E}$ is locally free whose first and second Chern classes are given in (1.43) and (1.44). It can easily be checked that $\tilde{E}$ is $e'$-stable, $e' = e_1 + r(H \cdot \Sigma^+)$, and $\text{Hom}_\Delta(\tilde{E}, \tilde{E}(D + \Sigma^+)) \neq \emptyset$. Therefore, we have obtained a map

$$(1.45) \quad \Psi : \mathcal{A}_{e,I}^{r,d}(D,1_m) \to \bigcup_{d',I'} \mathcal{A}_{e,I'}^{r,d'}(D + \Sigma^+),$$

where $d'$ can be any integer and $I'$ can possibly be $I(-\Sigma^+), \cdots, I(-(r-1)\Sigma^+)$. We wish to find an upper bound on

$$\#_{\text{mod}} \Psi^{-1}(\mathcal{A}_{e,I'}^{r,d'}(D + \Sigma^+))$$

that is independent of $(d',I')$. We begin with an estimate of $\#_{\text{mod}} \Psi^{-1}(\Psi(E))$ for any $E \in \mathcal{A}_{e,I}^{r,d}(D,1_m)$. Because $E$ belongs to the exact sequence

$$(1.46) \quad 0 \to F \to E \to L \to 0$$

($L = L_E$ as before) for $M = E_{|\Sigma^+}/L$, $F_{|\Sigma^+}$ fits into the exact sequence

$$(1.47) \quad 0 \to L(-\Sigma^+) \to F_{|\Sigma^+} \to M \to 0.$$ 

On the other hand, elements of $\Psi^{-1}(F)$ are parameterized by $\text{Ext}_\Delta^1(L,F)$, a subset of $\mathcal{P}$. Since $F$ is locally free,

$$\dim \text{Ext}_\Delta^1(L,F) = \dim \text{Ext}_\Delta^1(F, L \otimes K_\Delta) = \dim H^1_{\Sigma^+}(\text{Hom}(F, L) \otimes K_\Delta)$$

$$= \dim H^0_{\Sigma^+}(L \otimes F_{|\Sigma^+}(\Sigma^+)) \leq h^0(L \otimes L) + h^0(L \otimes M(\Sigma^+)).$$

Here the last inequality follows from (1.47). Since $L$ is $re_0$-stable, $h^0(L \otimes L)$ is bounded from above by a constant. In case $E \in W^+$, because $E_{|\Sigma^+}$ is $e_1$-stable, $h^0(L \otimes M(\Sigma^+))$ is also bounded from above. Hence for some constant $C_3$ depending only on $e_1$, $r$ and $I$, we have

$$(1.48) \quad \#_{\text{mod}} \Psi^{-1}(\Psi(E)) \leq C_3, \quad \forall E \in W^+.$$
When $E \in W^-$, $h^1(L^\vee \otimes M(\Sigma^+)) = 0$ because of our choice of $\epsilon_0$. Thus
\[
h^0(L^\vee \otimes M(\Sigma^+)) = \chi(L^\vee \otimes M(\Sigma^+)) = -\deg L \cdot \text{rk}(M) + \deg M \cdot \text{rk}(L) + \text{rk}(M) \cdot \text{rk}(L)(\delta + 1 - g).
\]
Combined with $\deg L + \deg M = I \cdot \Sigma^+$, we get
\[
#_{\text{mod}} \Psi^{-1}(\Psi(E)) \leq -r \deg L + C_4
\]
for some constant $C_4 \geq 0$ independent of $E \in W^-$. Now we use the induction hypothesis. Because $\text{rk}L = r_0 < r$, for any $E \in \mathfrak{A}_{\alpha_1^r, \alpha_2^r}(D, 1_m)$, either the generic fiber type of $E$ is $1_{m-r_0}$ or it is not of the form $(\alpha_1^{r_0}, \alpha_2^{r_0})$. Hence, with $I' = I(-r_0 \Sigma^+)$, $1 \leq r_0 \leq r$, and
\[
C_5 = C + C_4 + 2r^2 \delta + 2rI \cdot \Sigma^+,
\]
we can use Theorem 1.2 and the induction hypothesis to conclude that there is an $N_1$ and a constant $C_6$ so that when $d' \geq N_1$, we have
\[
#_{\text{mod}} \mathfrak{A}_{\alpha_1^{r}, \alpha_2^{r}}(D + \Sigma^+) \leq \eta(r, d', I') - C_5
\]
and when $d' \leq N_1$, we have
\[
#_{\text{mod}} \mathfrak{A}_{\alpha_1^{r}, \alpha_2^{r}}(D + \Sigma^+) \leq \eta(r, d', I') + C_6.
\]
We claim that when
\[
d \geq N = N_1 + r \delta + (2 + r)|I \cdot \Sigma^+| + C + C_4 + C_6,
\]
then
\[
#_{\text{mod}} W^- \leq \eta(r, d, I) - C.
\]
We break the estimate into two cases. In the case $d' = c_2(\tilde{E}) \geq N_1$, by (1.49) and (1.50),
\[
#_{\text{mod}} \Psi^{-1}(\mathfrak{A}_{\alpha_1^{r}, \alpha_2^{r}}(D + \Sigma^+)) \leq (\eta(r, d', I') - C_5) + (-r \deg L + C_4)
\]
\[
= \eta(r, d, I) - 2r_0 I \cdot \Sigma^+ + (-r_0^2(r - 1) + \frac{1}{2}r_0(r_0 - 1)) + r \deg L - C_5 + C_4
\]
\[
\leq \eta(r, d, I) - C.
\]
The last inequality holds because $\deg L \leq \frac{d}{r} I \cdot \Sigma^+$. Now assume $d' = c_2(\tilde{E}) < N_1$. Then
\[
#_{\text{mod}} \Psi^{-1}(\mathfrak{A}_{\alpha_1^{r}, \alpha_2^{r}}(D + \Sigma^+)) \leq (\eta(r, d', I') + C_0) + (-r \deg L + C_4)
\]
\[
\leq \eta(r, d, I) + r \deg L + 2r I \cdot \Sigma^+ + r^2 \delta + C_6 + C_4 \leq \eta(r, d, I) - C.
\]
Here we have used the fact that $\deg L \leq -2|I \cdot \Sigma^+| - r \delta - C_6 - C_4 - C$ which follows from (1.44), (1.52) and $d' < N_1$. Now we consider $E \in W^+$. Since $#_{\text{mod}} \Psi^{-1}(\Psi(E)) \leq C_3$ from (1.48) and $c_2(\tilde{E}) = d + \eta$ with $\eta$ a fixed integer independent of $d$, an argument similar to that of $W^-$ shows that there is an $N'$ such that for $d \geq N'$, we have $#_{\text{mod}} W^- \leq \eta(r, d, I) - C$. This establishes Theorem 1.1.
2. Degeneration of moduli space

We now recall briefly the construction of degeneration of moduli and refer the details of this construction to [GL]. We first fix a very ample line bundle \( H \) and a line bundle \( I \) on \( X \). Let \( C \) be a Zariski neighborhood of \( 0 \in \text{Spec } \mathbb{C} \). By choosing a smooth divisor \( \Sigma \in |H| \) we can form a threefold \( Z \) over \( C \) by blowing up \( X \times C \) along \( \Sigma \times \{0\} \). Clearly, \( Z_t \cong X, \ t \neq 0 \), and \( Z_0 \) consists of two smooth components \( X \) and a ruled surface \( \Delta \) that intersect normally along \( \Sigma \subseteq X \) and \( \Sigma^- \subseteq \Delta \). For any line bundle \( I \) on \( X \) and integers \( r \) and \( d \), let \( \mathcal{M}_{X}^{r,d} \) be the moduli space of rank \( r \) \( H \)-semistable sheaves over \( X \) of \( \det E = I \) and \( c_2(E) = d \). Let \( \mathcal{M}_{X}^{r,d} \times C^* \rightarrow C^* \), \( C^* = C \setminus \{0\} \), be the constant family over \( C^* \). The degeneration we construct will be a flat family \( \mathcal{M}_d \) (over \( C \)) extending the family \( \mathcal{M}_{X}^{r,d} \times C^* \) such that the closed points of the special fiber \( \mathcal{M}_d^0 = \mathcal{M}_d \times C \text{ Spec } \mathbb{C}[0] \) are in one-one correspondence with the semistable sheaves on \( Z_0 \) that will be defined shortly.

We first introduce the notion of torsion free sheaves on the surface \( Z_0 \):

**Definition 2.1.** A sheaf \( E \) on \( Z_0 \) is said to be torsion free at \( z \in Z_0 \) if whenever \( f \in \mathcal{O}_{Z_0,z} \) is a zero divisor of the \( \mathcal{O}_{Z_0,z} \)-modules \( E_z \), then \( f \) is a zero divisor of the \( \mathcal{O}_{Z_0,z} \)-modules \( \mathcal{O}_{Z_0,z} \). The sheaf \( E \) is said to be torsion free if \( E \) is torsion free everywhere.

Let \( E \) be any coherent sheaf on \( Z_0 \). We denote by \( E^{(1)} \) (resp. \( E^{(2)} \)) the torsion free part of \( E |_X \) (resp. \( E |_\Delta \)). We define the rank of \( E \) to be a pair of integers, \( \text{rk}(E) = (\text{rk}(E^{(1)}), \text{rk}(E^{(2)})) \). When \( \text{rk}(E) = (r, r) \), we simply call \( E \) a rank \( r \) sheaf.

Let \( \varepsilon \in (0, \frac{1}{2}) \) be a rational number. We define a \( \mathbb{Q} \)-ample divisor \( H(\varepsilon) \) on \( Z \) as follows: Let \( p_X : Z \rightarrow X \) be the projection and put

\[
H(\varepsilon) = p_X^* H(-(1-\varepsilon)\Delta).
\]

Clearly, for integer \( n_0 \) so that \( n_0 \cdot \varepsilon \in \mathbb{Z} \),

\[
H(\varepsilon)^{\otimes n_0} = p_X^* H^{\otimes n_0}(-(n_0 - n_0\varepsilon)\Delta)
\]

is an ample divisor. In the sequel, we will constantly use the tensor power \( H(\varepsilon)^{\otimes n} \). We agree without further mentioning that in such cases, \( n \) is always divisible by \( n_0 \).

Let \( \alpha = (\alpha_1, \alpha_2) \) be a pair of rational numbers:

\[
\alpha_1 = (H(\varepsilon)|_X \cdot H(\varepsilon)|_\Delta)/(H \cdot H), \quad \alpha_2 = (H(\varepsilon)|_\Delta \cdot H(\varepsilon)|_\Delta)/(H \cdot H).
\]

Note that \( \alpha_1 + \alpha_2 = 1 \). For any sheaf \( E \) on \( Z_0 \) with \( \text{rk}(E) \neq (0, 0) \), we define \( p_E \) to be the polynomial \( p_E = \frac{1}{\text{rk}(E) \cdot \alpha} \chi_E \).

We remark that since \( \chi_E(n) = \chi(E \otimes H^{\otimes n}) \) is well defined for those \( n \) divisible by \( n_0 \) and is a restriction of a polynomial in \( n \), we can define \( \chi_E \) to be that polynomial. Once we have the polynomial \( p_E \), we can define the \( H(\varepsilon) \)-stability (or \( H(\varepsilon) \)-semistability) of \( E \) by mimicking Definition 0.5 word for word.
Definition 2.2. A torsion free sheaf $E$ on $Z_0$ is said to be $H(\varepsilon)$-stable (resp. $H(\varepsilon)$-semistable) if whenever $F \subseteq E$ is a proper subsheaf, then $p_F < p_E$ (resp. $\leq$).

We fix a line bundle $I$ on $X$ and an integer $r \geq 2$. We let $\chi(n)$ be the polynomial that depends on $(r, d, I, H, X)$:
\[
\chi(n) = \frac{r}{2}n^2(H \cdot H) + n((H \cdot I) - \frac{r}{2}(H \cdot K_X)) + (r - 1)\chi(O_X) + \chi(I) - d.
\]
\[
\chi(\cdot)\text{ is the Hilbert polynomial of a rank } r \text{ sheaf of } c_1 = I \text{ and } c_2 = d.
\]
We also fix a rational $\varepsilon \in (0, \frac{1}{2})$ momentarily and the $\mathbb{Q}$-ample line bundle $H(\varepsilon)$. For convenience, we will denote by $E(n)$ the sheaf $E \otimes p_Z^*H(\varepsilon)^{\otimes n}$ for any sheaf $E$ over $Z_S$. (Here $S$ is any scheme over $C$ and $Z_S = Z \times_C S$.)

We now construct the degeneration $\mathcal{M}^d$ promised at the beginning of this section. Recall that the moduli space $\mathcal{M}_X^{d, r}$ was constructed as a GIT quotient of Grothendieck’s Quot-scheme. Here, we shall adopt the same approach to construct $\mathcal{M}^d$. We first fix a sufficiently large $n$ and let $\rho = \chi(n)$. Following A. Grothendieck [Gr], we define $\text{Quot}_{Z/C}^{\chi, O^n}$ to be the functor sending any scheme $S$ of finite type over $C$ to the set of all quotient sheaves $E(n)$ of $O_{Z_S}^{\varepsilon, \rho}$ on $Z_S$ flat over $S$ so that $\chi_E(m) = \chi(n + m)$ for any closed $s \in S$. $\text{Quot}_{Z/C}^{\chi, O^n}$ is represented by a scheme $\text{Quot}_{Z/C}^{\chi, O^n}$, that is projective over $C$, called Grothendieck’s Quot-scheme. Similarly, we have Grothendieck’s Quot-scheme $\text{Quot}_{Z/C}^{\chi, O^n}$ parameterizing all quotient sheaves $O_{Z_S}^{\varepsilon, \rho} \to E(n)$ on $X$ with $\chi_E \equiv \chi(\cdot + n)$. Let $U_{X}^{d, l} \subseteq \text{Quot}_{Z/C}^{\chi, O^n}$ be the subset of all $H$-semistable quotient sheaves $E(n)$ obeying one further restriction: det $E = I$. $U_{X}^{d, l}$ is locally closed.

We define $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$ to be the closure of $U_{X}^{l, d} \times C^* \subseteq \text{Quot}_{Z/C}^{\chi, O^n}$ endowed with the reduced scheme structure and denote by $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$ the normalization of $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$. $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$ has the property that it is normal, projective and flat over $C$. Finally, we define $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l} \subseteq \text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$ to be the subset of all closed points whose associated quotient sheaves are $H(\varepsilon)$-semistable.

Clearly, $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$ depends on the choice of $(r, d, n, I, H, \varepsilon)$. In the sequel, $r$, $I$ and $H$ will be fixed once and for all. Of course, $d$ should be viewed as a variable. For technical reasons, the choice of $\varepsilon$ will depend on $d$. After this, we will choose $n$ sufficiently large (the exact value of $n$ is irrelevant to our discussion as long as it meets the requirement of [Gi, Corollary 1.3], [GL, Corollary 1.11]). If all of these are understood, then we will abbreviate $\text{Quot}_{Z/C, H(n)}^{\chi, O^n, l}$ to $U_{d, \varepsilon}$. By abuse of notation, we will call $E$ the universal family of $U_{d, \varepsilon}$, where $E(n)$ is the pullback of the universal quotient family on $Z \times_C \text{Quot}_{Z/C}^{\chi, O^n}$.

Let $SL_C = SL(\rho, C) \otimes_C C$ be the special linear group scheme over $C$. Clearly, $\text{Quot}_{Z/C}^{\chi, O^n}$ is an $SL_C$-scheme. By our construction, this action lifts to $U_{d, \varepsilon}$. Further, we have

Theorem 2.3 ([GL, Theorems 2.10 and 2.11]). The good quotient $\mathcal{M}^{d, \varepsilon} = U_{d, \varepsilon}/\mathcal{SL}_C$ exists. $\mathcal{M}^{d, \varepsilon}$ is normal, projective and flat over $C$. Further, for any closed $t \neq 0$, $\mathcal{M}^{d, \varepsilon}_t$ is isomorphic to the normalization of the moduli scheme $\mathcal{M}^{d, \varepsilon}_X$.

To make use of this degeneration, we need to analyze the closed points of $\mathcal{M}^{d, \varepsilon}_0$. Since $\mathcal{M}^{d, \varepsilon}_0$ is a GIT quotient of $U_{d, \varepsilon}^0$, each point of $\mathcal{M}^{d, \varepsilon}_0$ associates to an equivalent
class of sheaves $E(n)$ in $U_0^{d,\varepsilon}$. In the following, we will find bounds on $c_1(E^{(1)})$, $c_1(E^{(2)})$ and $c_2(E^{(2)})$ that are independent of the choice of $\varepsilon$ and $E(n) \in U_0^{d,\varepsilon}$.

First we study $c_1(E^{(1)})$ and $c_1(E^{(2)})$. Following [GL, Lemma 1.6], there is a sheaf of $O_{\Sigma}$-modules $E^{(0)}$ (of rank $r_0$) such that $E$ belongs to the exact sequence

$$0 \longrightarrow E \longrightarrow E^{(1)} \oplus E^{(2)} \longrightarrow \iota_* E^{(0)} \longrightarrow 0,$$

where $\iota: \Sigma \hookrightarrow X$. Because $U_0^{d,\varepsilon}$ is flat over $C$, there are integers $a_1$, $a_2$ with

$$a_1 + a_2 = r_0 - r$$

such that $\det E^{(1)} = I(a_1 H)$, $\det E^{(2)} = I_0 \Delta(a_2 \Sigma^{-})$, where $I_0 = p_X^* I_{\Sigma_0}$ [GL, §4].

Then since both $E^{(1)}$ and $E^{(2)}$ are quotient sheaves of $E$, by the $H(\varepsilon)$-stability of $E$, we have

$$a_1 \geq (1 - \varepsilon) \frac{H \cdot (rK_X - 2I)}{2(H \cdot H)}; \quad a_2 \geq -(1 - \varepsilon) \frac{H \cdot (rK_X - 2I)}{2(H \cdot H)} - r.$$

Since $H$ is a very ample divisor on $X$, we may and will assume that $r$ divides $H \cdot I$ and

$$H \cdot H \geq 18|K_X \cdot H| + 18|I \cdot H|.$$

Therefore, $r \geq a_1 \geq 0$ and $0 \geq a_2 \geq -r$.

The bound of $c_2(E^{(2)})$ is achieved by applying Bogomolov’s argument which shows that when $V$ is an $H$-stable vector bundle on $X$, then the restriction of $V$ to a high degree hyperplane curve is semistable. We follow the argument in [GL, §4] and indicate the necessary changes needed in the higher rank case.

**Lemma 2.4** (cf. [GL, Lemma 4.3]). There is a constant $c_1$ independent of $d$ and $\varepsilon$ such that the sheaf $E^{(1)}$ (on $X$) is $c_1$-stable.

**Lemma 2.5.** For any constant $c_1$ and integer $r$, there is a constant $C_1$ such that whenever $V$ is an $c_1$-stable torsion free sheaf of rank $\leq r$, then

$$c_2(V) - \frac{\text{rk}(V) - 1}{2 \text{rk}(V)} c_1(V)^2 \geq C_1.$$

**Proof.** Lemma 2.4 is true because any quotient sheaf $Q$ of $E^{(1)}$ is also a quotient sheaf of $E$. Hence the degree of $Q$ has to satisfy an inequality, which combined with (2.4) gives us the desired inequality. The details of the argument can be found in [GL, Lemma 4.3]. Now we prove Lemma 2.5 following the suggestion of the referee. By Riemann-Roch,

$$\chi(V, V) = \chi(\text{Ext}^1(V, V)) = -2r(c_2(V) - \frac{r - 1}{2r} c_1(V)^2) + r^2 \chi(O_X).$$

Since $V$ is $c_1$-stable, by Lemma 1.8, there is a constant $C_1$ such that $\text{dim} \text{Hom}(V, V)$ and $\text{dim} \text{Hom}(E, E \otimes K_X)$ are bounded from above by $C_1$. Hence

$$c_2(V) - \frac{r - 1}{2r} c_1(V)^2 = \frac{1}{2r} \chi(O_X) - \frac{1}{2r} \chi(V, V) \geq \frac{1}{2r} \chi(O_X) - 2C_1. \quad \square$$
Lemma 2.6 (cf. [GL, Lemma 4.4]). For any constant $c_1$, there is a constant $C_2$ such that whenever $V$ is an $e_1$-stable, rank $r$ vector bundle on $X$ with $\det V = I(aH)$, $|a| \leq r$, and that $Q$ is an $O_{\Sigma}$-module that is a quotient sheaf of $V_{|\Sigma}$, then we have $\chi(Q) \geq -c_2(V) + C_2$.

Proof. Let $W$ be the kernel of $V \to Q$. By Riemann-Roch, $c_1(W) = I + (a - c)|\Sigma|$ and

$$c_2(W) = c_2(V) + \chi(Q) + \frac{1}{2} c(K_X + cH) \cdot H - c(I + aH) \cdot H.$$ 

Thus

$$\chi(Q) \geq -c_2(V) + c_2(W) - 2r^2H^2.$$ 

On the other hand, since $V$ is $e_1$-stable, $W$ is $(e_1 + 1)$-stable. So by Lemma 2.5, there is a constant $C_1$ so that $c_2(W) \geq C_1$. This completes the proof of Lemma 2.6.

Proposition 2.7. There is a constant $C_3$ independent of $\epsilon$ and $d$ such that for any $E(n) \in U_{0}^d, ε$,

$$c_2(E^{(2)}) \leq d + C_3.$$ 

Proof. By (2.3), we have

$$\chi_{E^{(2)}}(\cdot) = \chi_{E}(\cdot) + (\chi_{E^{(0)}}(\cdot) - \chi_{E^{(1)}}(\cdot)).$$ 

Hence the proposition follows if we can show that the constant term of $\chi_{E^{(0)}}(\cdot) - \chi_{E^{(1)}}(\cdot)$ is bounded from below but this follows from Lemma 2.6. The details of the proof are given in [GL, Proposition 4.6].

Our next goal is to construct Donaldson’s line bundle $\mathcal{L}$ on $\mathfrak{M}_{d,\epsilon}$ and to establish the following key property of $\mathcal{L}$: Whenever $W_0 \subseteq \mathfrak{M}_{0}^{d,\epsilon}$ is a dimension $c$ subvariety such that $|\mathcal{E}|^c(W_0) > 0$, then

$$\#_{\text{mod}} \{E^{(2)} \mid E(n) \in W_0\} = c.$$ 

We first sketch the construction of $\mathcal{L}$. The full account of this construction appeared in [GL, §5]. For any integer $h \geq 1$, let $D^h \subseteq Z$ be a smooth divisor such that $\pi: D^h \to C$ is smooth, that $D^h = \pi^{-1}(t) \in [hH]$ for $t \neq 0$ and that $D^h_0 \subseteq \Delta \setminus \Sigma$. We call such $D^h$ good divisors in $[hH_C(-h\Delta)]$, where $H_C = p^*_X H$. Since $H$ is very ample, the set of good divisors in $[hH_C(-h\Delta)]$ is base point free. Associated to each $D^h$ we can find an étale covering $\tilde{C} \to C$ such that on $D^h_{\tilde{C}} = D^h \times_C \tilde{C}$ there is a line bundle $\tilde{\theta}^h$ satisfying $(\tilde{\theta}^h) \otimes 2r = K_{D^h_{\tilde{C}}} \otimes p^*_X I_{D^h_{\tilde{C}}} \otimes 2r$ for all closed $\tilde{C} \in \tilde{C}$, where $K_{D^h_{\tilde{C}}}$ is the canonical divisor of $D^h_{\tilde{C}}$. We remark that such $\tilde{\theta}^h$ exist because $[D^h_{\tilde{C}}] \cdot I = H \cdot I$ is divisible by $r$.

We first construct a line bundle on $U_{\tilde{C}}^{d,\epsilon} = U^{d,\epsilon} \times_C \tilde{C}$ as follows: Let $\mathcal{E}(n)$ be the universal quotient family on $Z \times_C U^{d,\epsilon}$. Since $\mathcal{E}$ is a family of torsion free sheaves flat over $U^{d,\epsilon}$, $\mathcal{E}$ admits length two locally free resolution near $D^h$. Thus the restriction of $\mathcal{E}$ to $D^h \times_C U^{d,\epsilon}$ (denoted by $\mathcal{E}_{|D^h}$) has a length two locally free resolution also
(see [L1]). Let \( p_{12} \) (resp. \( p_{13} \); resp. \( p_{23} \)) be the projection from \( D^h \times_C U^d_{\tilde{C}} \) to \( D^h \) (resp. to \( D^h \times U^d_{\tilde{C}} \); resp. to \( U^d_{\tilde{C}} \)). Note that \( p_{23} \) is smooth. Hence

\[
(2.8) \quad R^1p_{23*}(p_{13*}(\mathcal{E}_{\mathcal{D}^h}) \otimes p_{12}^*\theta^h)
\]

is a perfect complex on \( U^d_{\tilde{C}} \) [KM]. Following [KM], we can define a determinant line bundle

\[
(2.9) \quad \det\left(R^1p_{23*}(p_{13*}(\mathcal{E}_{\mathcal{D}^h}) \otimes p_{12}^*\theta^h)\right)
\]

on \( U^d_{\tilde{C}} \) whose inverse we call \( \mathcal{L}_\mathcal{D}(D^h) \). If we choose another good divisor \( D^{h'} \in |hH_C(-h\Delta)| \) and form the corresponding line bundle \( \mathcal{L}_\mathcal{D}(D^{h'}) \) on \( U^d_{\tilde{C}} \), then since the set of good divisors in \( |hH_C(-h\Delta)| \) is an irreducible set, for any \( v \in \tilde{C} \) and \( v' \in \tilde{C}' \) which lie over the same closed point \( t \in C \), the line bundles \( \mathcal{L}_\mathcal{D}(D^h)|U^d_{\tilde{C}} \) and \( \mathcal{L}_\mathcal{D}(D^{h'})|U^d_{\tilde{C}} \) are algebraically equivalent.

Remark. Indeed, more is true. There is a single line bundle \( \mathcal{L}_\mathcal{D}(h) \) on \( U^d_{\tilde{C}} \) such that the line bundles \( \mathcal{L}_\mathcal{D}(D^h) \) on \( U^d_{\tilde{C}} \) are pullback of \( \mathcal{L}_\mathcal{D}(h) \) via \( U^d_{\tilde{C}} \to U^d_{\tilde{C}} \).

Our next task is to show that under favorable conditions, these line bundles descend to line bundles on \( \mathfrak{M}^d \cdot \mathcal{C} \). We need the following result of Kempf:

**Lemma 2.8** (Descent lemma [DN, Theorem 2.3]). Let \( \mathcal{L} \) be an \( \mathcal{S}\mathcal{L}_{\mathcal{C}} \) line bundle on \( U^d_{\tilde{C}} \). \( \mathcal{L} \) descends to \( \mathfrak{M}^d \cdot \mathcal{C} \) if and only if for every closed point \( w \in U^d_{\tilde{C}} \) with closed orbit \( \mathcal{S}\mathcal{L}_{\mathcal{C}} \cdot \{w\} \), the stabilizer \( \text{stab}(w) \subseteq \mathcal{S}\mathcal{L}_{\mathcal{C}} \) acts trivially on \( \mathcal{L}_w = \mathcal{L} \otimes k(w) \).

We have

**Proposition 2.9.** There is a function \( \kappa: \mathbb{Z}^+ \to (0, \frac{1}{2}) \) for which the following holds:

For any \( d \), there is a large \( h \) such that when \( \varepsilon \in (0, \kappa(d)) \cap \mathbb{Q} \) and \( D^h \in |hH_C(-h\Delta)| \) is a good divisor, then the line bundle \( \mathcal{L}_\mathcal{D}(D^h) \) (on \( U^d_{\tilde{C}} \)) descends to a line bundle on \( \mathfrak{M}^d \cdot \mathcal{C} \). We denote the descent by \( \mathcal{L}_\mathfrak{M}(D^h) \).

**Proof.** It is straightforward to check that \( w = E(n) \in U^d_{\tilde{C}} \) (over \( t \in C \)) has closed orbit if and only if \( E \) splits into a direct sum of stable sheaves \( F_1, \ldots, F_k \). Then following [L1, p. 426], the stabilizer \( \text{stab}(w) \) acts trivially on \( \mathcal{L}_\mathcal{D}(D^h)_w \) if and only if

\[
\frac{1}{\text{rk}(F_1)}c_1(F_1) \cdot D^h_1 = \cdots = \frac{1}{\text{rk}(F_k)}c_1(F_k) \cdot D^h_k.
\]

These identities follow if we can prove

**Proposition 2.10.** There is a function \( \kappa: \mathbb{Z}^+ \to (0, \frac{1}{2}) \) and a constant \( N \) for which the following holds: Given \( d_0 \), there is an \( h \geq 1 \) such that for any \( \varepsilon \in (0, \kappa(d_0)) \), whenever \( d \leq d_0 \), and that \( E(n) \in U^d_{\mathcal{C}} \) is an \( H(\varepsilon) \)-semistable sheaf over \( t \in C \), then for a generic good divisor \( D^h \in |hH_C(-h\Delta)| \), \( E|_{D^h_t} \) is semistable.

**Completion of the proof of Proposition 2.9.** Assume \( E = F_1 \oplus \cdots \oplus F_k \). By Proposition 2.10, there is a good divisor \( D^{h'} \in |hH_C(-h\Delta)| \) such that \( E|_{D^{h'}_t} \) is semistable. Then the value \( \frac{1}{\text{rk}(F_i)}c_1(F_i) \cdot D^{h'}_i = \frac{1}{\text{rk}(F_i)}c_1(F_i) \cdot D^h_i \) is identical for all \( i \).

Proposition 2.10 will be proved shortly.
Remark. Let \( t \neq 0 \in C \) be any closed point. Then the line bundle \( \mathcal{L}_U(D^h)_t \) on \( \mathcal{U}_D^d \) descends regardless of the choice of \( d \) and \( \varepsilon \) [L1, p. 426]. In particular, \( \mathcal{L}_M(D^h)_t \) always exists on \( \mathfrak{M}_t^{d,\varepsilon} \).

Now we explain how to construct global sections of \( \mathcal{L}_M(D^h)^\otimes_m \) on \( \mathfrak{M}_t^{d,\varepsilon}, v \in \check{C} \). All we need to know about the line bundle \( \mathcal{L}_M(D^h)_v \) is how to calculate its intersection numbers on various subvarieties of \( \mathfrak{M}_t^{d,\varepsilon} \). So in the following, we will not distinguish between the line bundles \( \mathcal{L}_M(D^h)_v \) and \( \mathcal{L}_M(D^h)_{v'} \) (resp. \( \mathcal{L}_U(D^h)_v \) and \( \mathcal{L}_U(D^h)_{v'} \)) when \( v, v' \in \check{C} \) lie over the same closed point \( t \in C \). By abuse of notation, we will denote both of them by \( \mathcal{L}_M(D^h)_t \) (resp. \( \mathcal{L}_U(D^h)_t \)).

For any good \( D^h \in |hH_C(-h\Delta)| \) and any closed \( t \in C \), let \( \mathcal{U}_t^{d,\varepsilon}[D^h_t] \subseteq \mathfrak{U}_t^{d,\varepsilon} \) be the open set of all \( s \in \mathfrak{U}_t^{d,\varepsilon} \) such that \( \mathcal{E}_{s[D^h]} \) is semistable. In the following, we abbreviate \( D = D^h_t \). By restricting \( E(n) \in \mathfrak{U}_t^{d,\varepsilon}[D] \) to \( D \), we obtain a morphism

\[
\Phi_D : \mathcal{U}_t^{d,\varepsilon}[D] \longrightarrow \mathfrak{M}_t^{r,I}(D),
\]

where \( \mathfrak{M}_t^{r,I}(D) \) is the moduli scheme of rank \( r \) semistable vector bundles \( V \) on \( D \) with \( \det V = \rho_X I_D \). If we view \( \mathfrak{M}_t^{r,I}(D) \) as an \( SL(\rho, \mathbb{C}) \) scheme with trivial group action, the morphism \( \Phi_D \) is \( SL(\rho, \mathbb{C}) \)-equivalent.

**Proposition 2.11 [Donaldson].** There is an ample line bundle \( \mathcal{L}_D \) on \( \mathfrak{M}_t^{r,I}(D) \) so that its pull back under \( \Phi_D \) is canonically isomorphic to the restriction to \( \mathcal{U}_t^{d,\varepsilon}[D] \) of \( \mathcal{L}_U \). Further, this isomorphism is \( SL(\rho, \mathbb{C}) \)-equivariant.

**Proof.** For the details of the proof, the readers are advised to look at [L2, p. 31]. Though the author only treated the case \( r = 2 \) in the proof, the proof of the higher rank case is similar.

Now let \( m \) be a large positive integer. Since the isomorphism

\[
\Phi^*_D(\mathcal{L}_D) \cong \mathcal{L}_U|_{\mathcal{U}_t^{d,\varepsilon}[D]}
\]

is \( SL(\rho) \)-equivalent, for any \( \xi \in H^0(\mathfrak{M}_t^{r,I}(D), \mathcal{L}_D^\otimes_m) \), \( \Phi^*_D(\xi) \) is an \( SL(\rho) \)-invariant section of \( \mathcal{L}_U|_{\mathcal{U}_t^{d,\varepsilon}[D]} \).

**Lemma 2.12.** Let \( D^h \in |hH_C(-h\Delta)| \) be any good divisor and for any \( t \in C \) with \( D = D^h_t \), let \( \xi \in H^0(\mathfrak{M}_t^{r,I}(D), \mathcal{L}_D^\otimes_m) \) be any section. Then the pullback section \( \Phi^*_D(\xi) \) (on \( \mathcal{U}_t^{d,\varepsilon}[D] \)) extends canonically over \( \mathfrak{U}_t^{d,\varepsilon} \) to an \( SL(\rho, \mathbb{C}) \)-invariant section.

We shall denote this extension (and its descent to \( \mathfrak{M}_t^{d,\varepsilon} \) if no confusion is possible) by \( \Phi^*_D(\xi)_{ex} \). Furthermore,

\[
\Phi^*_D(\xi)_{ex}^{-1}(0) = \left( \mathfrak{U}_t^{d,\varepsilon} \setminus \mathfrak{U}_t^{d,\varepsilon}[D] \right) \cup \left\{ F(n) \in \mathfrak{U}_t^{d,\varepsilon}[D] \mid \xi(F_D) = 0 \right\}.
\]

**Proof.** In case \( \mathfrak{U}_t^{d,\varepsilon} \) is normal, we can apply [GL, Lemma 5.6], [GL, Proposition 5.7] and [L2, Lemma 4.10] to our situation. In general, we need to use GIT to prove this lemma [L1, p. 435].

In the following, we seek to estimate the self-intersection numbers of \( \mathcal{L}_M(D^h) \) on subvarieties \( W \subseteq \mathfrak{M}_t^{d,\varepsilon} \) and to relate the non-vanishing of such numbers to the estimate of the numbers (2.7). Our immediate goal is to prove
Proposition 2.13. Let \( t \neq 0 \in \mathcal{C} \) be any closed point and let \( W_t \subseteq \mathfrak{M}_{t,c}^d \) be an irreducible variety of dimension \( c \). Then for sufficiently large \( h \) and for any good \( D^h \in |hH_C(-h\Delta)| \),

\[
(2.13) \quad [\mathcal{L}_M(D^h)]^c(W_t) \geq 0.
\]

Further, if we assume that the general points of \( W_t \) are locally free \( H,\mu \)-stable sheaves, then the strict inequality holds.

Proof. To prove (2.13), it suffices to find divisors \( D_1, \cdots, D_c \in |hH| \) and sections \( \varphi_1, \cdots, \varphi_c \) of \( \mathcal{L}_M(D^h)_t \) such that \( \cap_{i=1}^c \varphi_i^{-1}(0) \) is a finite set. But this is obvious because for sufficiently large \( h \), the restriction of each \( E \in \mathfrak{M}_{t,c}^d \) to general \( D \in |hH| \) is semistable (Proposition 2.10). Now we prove the second part of the proposition. Let \( h \) be large so that for any locally free \( E \in \mathfrak{M}_{t,c}^d \), \( H^1(\mathcal{M}^D_0(E)(-hH)) = 0 \). Then for any \( D \in |hH| \) and any locally free stable \( E_1, E_2 \in W_t \), \( E_1|_D = E_2|_D \) implies \( E_1 \cong E_2 \). We can also assume that the restriction of any \( E \in \mathfrak{M}_{t,c}^d \) to a general \( D \in |hH| \) is semistable.

Choose \( D \in |hH| \) so that \( U^d_{t,c}[D] \cap W_t \) is non-empty. Then because the line bundle \( \mathcal{L}_D \) is ample on \( \mathfrak{M}_{t,c}^d(D) \) and because

\[
\Psi_D : \mathfrak{M}_{t,c}^d[D] \cap W_t \rightarrow \mathfrak{M}_{t,c}^d(D)
\]

\((\mathfrak{M}_{t,c}^d[D] \text{~is~the~image~of~}\mathcal{L}_D^{d,c}[D] \text{~under~the~projection})\) is generically one-to-one, there is a section \( \xi \in H^0(\mathfrak{M}_{t,c}^d(D), \mathcal{L}_D^{d,m}) \), \( m \) large, such that the extension of the pullback section \( \Phi_D^*(\xi)_{ex} \) (over \( \mathfrak{M}_{t,c}^d \)) is non-trivial over \( W_t \) and

\[
\dim \left( \Phi^*_D(\xi)_{ex}^{-1}(0) \cap W_t \right) = \dim W_t - 1.
\]

Since being locally free and stable are open conditions, we can assume that general points of at least one irreducible component of \( \Phi_D^*(\xi)_{ex}^{-1}(0) \cap W_t \) are still locally free and \( H,\mu \)-stable. Therefore, we can use induction on \( \dim W_t \) to conclude that \( [\mathcal{L}_M(D^h)]^{c-1}(W_t) \geq 0 \) for any irreducible component \( W'_t \) of \( \Phi_D^*(\xi)_{ex}^{-1}(0) \cap W_t \), and for at least one of these component, this number is positive. Therefore, the strict inequality (2.13) holds.

The converse to the proposition is that if a set \( W_t \subseteq \mathfrak{M}_{t,c}^d \) with \( \dim W_t = c \) has the property that

\[
[\mathcal{L}_M(D^h)]^c(W_t) > 0,
\]

then \( \#_{\mathfrak{mod}}(W) \geq c \). But this is a tautology since \( \mathfrak{M}_{t,c}^d \) is the normalization of the moduli scheme. What we need is a similar result in \( t = 0 \). We will prove

Proposition 2.14. Let \( W_0 \subseteq \mathfrak{M}_{0,c}^d \) be any (complete) subvariety of dimension \( c \). Assume for some large \( h \) (given by Proposition 2.10) and good \( D^h \in |hH_C(-h\Delta)| \) we have

\[
[\mathcal{L}_M(D^h)]^c(W_0) > 0.
\]

Then \( \#_{\mathfrak{mod}}\{ E^{(2)} \mid E(n) \in W_0 \} = c \).

Proof. We prove it by contradiction. Assume \( \#_{\mathfrak{mod}}\{ E^{(2)} \mid E(n) \in W_0 \} < c \). Then \( \{ E^{(2)} \mid E(n) \in W_0 \} \) can be parameterized by finite irreducible varieties
of dimension at most $c - 1$. Let them be $S_1, \cdots, S_k$ and let $E_1, \cdots, E_k$ be the corresponding families. Thanks to Proposition 2.10, there is a large $h$ such that for any $F \in \{E^{(2)} | E(n) \in W_0\}$, $F|_{D^h_D}$ is semistable for generic $D^h \in [hH_C(-h\Delta)]$. We fix such an $h$. We choose a $D^h \in [hH_C(-h\Delta)]$ so that $E_i,s_i|D^h_D$ are semistable for some closed $s_i \in S_i, i = 1, \cdots, k$. Since $D^h_D$ is ample, we can further choose $\xi \in H^0(G\mathcal{M}^{-1}(D^h_D), L^{\otimes m}_D), m \geq 1$, so that $\xi(E_i,s_i|D^h_D) \neq 0$ for all $i$.

Let $\Psi^*_D(\xi)_{c\varepsilon}$ be the extension of the pullback of $\xi$ in $H^0(G\mathcal{M}(D^h_D), L^{\otimes m}_D)$. Put $W' = W_0 \cap \Psi^*_D(\xi)_{c\varepsilon}^{-1}(0)$. By our construction, $\dim W' \leq \dim W_0 - 1$ and

$$\#_{\text{mod}}\{E^{(2)} | E(n) \in W'_0\} \leq \max_{i=1, \cdots, k} \{\dim S_i - 1\} \leq \#_{\text{mod}}\{E^{(2)} | E(n) \in W_0\} - 1.$$

Note that $[\mathcal{L}_M(D^h)]_{c-1}(W'_0) = m[\mathcal{L}_M(D^h)]_c(W_0) > 0$. So by the induction hypothesis, we have $\#_{\text{mod}}\{E^{(2)} | E(n) \in W'_0\} \geq c - 1$. Therefore,

$$\#_{\text{mod}}\{E^{(2)} | E(n) \in W_0\} \geq \#_{\text{mod}}\{E^{(2)} | E(n) \in W'_0\} + 1 \geq c.$$

The proposition follows because $\#_{\text{mod}}\{E^{(2)} | E(n) \in W_0\} \leq \dim W_0 = c$.

In the remainder of this section, we will give the proof of Proposition 2.10 that is parallel to the treatment for the rank two situation given in [GL, 5.13]. Let $E(n) \in U^{d,e}_t$ be any $H(\varepsilon)$-semistable sheaf over $t \in C$. When $t \neq 0$, then $E$ is an $H$-semistable sheaf over $X$ and [MR] tells us that for large $h$ and generic $D \in [hH]$, $E|_D$ is semistable. In case $t = 0$, namely when $E$ is an $H(\varepsilon)$-semistable sheaf on $Z_0$, the situation is quite tricky because $Z_0$ is reducible and the divisorial ray $\mathbb{R}[D^h_D]$ is different from $\mathbb{R}[H(\varepsilon)\Delta]$. However, it is essential that $\mathbb{R}[D^h_D]$ and $\mathbb{R}[H(\varepsilon)\Delta]$ become very close when $\varepsilon$ becomes small. Before going into the details of the proof, let us state the following stability criterion of $E^{(2)}$.

**Lemma 2.15.** There is a constant $c_2$ such that for any $d, \varepsilon$ and any $E(n) \in U^{d,e}_0$, $E^{(2)}$ is $c_2$-stable with respect to $H(\varepsilon)|\Delta$.

**Proof.** See [GL, 5.14].

**Proof of Proposition 2.10.** Let $V$ be the double dual of $E^{(2)}$. By (2.4) and Proposition 2.7, det $V = I_0(a_2 \Sigma^-), -r \leq a_2 \leq 0$, and $c_2(V) \leq c_2(E) + C_3$, where $C_3$ is a constant independent of $E$ and $d$. Since $I_0 \cdot \Sigma^-$ is divisible by $r$, by tensoring $V$ with some line bundle, we can assume $c_1(V) \sim a_2[\Sigma^-]$. Note that $c_2(V)$ is still bounded by $d_0 + C_3$ possibly with a new constant $C_3$. Clearly, the proposition will be established if we can show that there is an $\varepsilon_0$ and an integer $h$ such that whenever $\varepsilon < \varepsilon_0$ and $V$ is $c_2\varepsilon$-stable with respect to $H(\varepsilon)|\Delta$ as before, then for generic $D \in [h\Sigma^+]$, $V_D$ is semistable.

The argument we adopt is a direct generalization of Bogomolov’s theorem showing that the restriction of any $\mu$-stable rank two vector bundle $E$ to any smooth hyperplane section of degree $\geq 2c_2(E) + 1$ is stable. We prove it by contradiction. Assume otherwise. Then there is a rank $s$ $(1 \leq s \leq r - 1)$ quotient vector bundle $Q$ of $V_D$ such that $0 = \mu(V_D) > \mu(Q)$. Let $W$ be the kernel of $V \rightarrow Q$. Then $W$ is a locally free sheaf on $\Delta$ with $c_1(W) \sim a_2[\Sigma^-] - sh[\Sigma^+]$ and

$$c_2(W) = c_2(V) + \frac{1}{2}s(s - 1)h^2H^2 + \deg Q < c_2(V) + \frac{1}{2}s(s - 1)h^2H^2.$$
Thus a simple calculation gives us
\begin{equation}
2rc_2(W) - (r - 1)c_1(W)^2 < 2rc_2(V) - (s(r - s)h^2 - (r - 1)a_2^2)H^2.
\end{equation}
Because \( c_2(V) \leq d_0 + C_3 \), when
\[
h^2 \geq r^2 + \frac{2r}{r - 1} \frac{d_0 + C_3}{H^2}
\]
the right-hand side of (2.14) is negative. Therefore, Bogomolov’s inequality shows that \( W \) is unstable. Let
\begin{equation}
0 = W_0 \subset W_1 \subset \cdots \subset W_n = W
\end{equation}
be the Harder-Narasimhan filtration of \( W \) such that the sheaves \( F_i = W_i/W_{i-1} \) are \( \mu \)-semistable and \( \mu(F_i) > \mu(F_{i+1}) \). Let \( r_i = \text{rk}(F_i) \) and let \( \Gamma_i \) be the \( \mathbb{Q} \)-divisor supported on fibers of \( \Delta \rightarrow \Sigma \) such that
\[
c_1(F_i) \sim r_i(b_i\Sigma^- + \Gamma_i).
\]
We let \( c_i = \Gamma_i \cdot \Sigma^+/H^2 \). Then \( b_i \) and \( c_i \) satisfy the following inequalities:
\begin{equation}
(e_2 + \frac{a_2}{r})\varepsilon \geq \varepsilon b_1 + (1 - \varepsilon)c_1 + \cdots + \varepsilon b_n + (1 - \varepsilon)c_n.
\end{equation}
The first inequality holds because \( E^{(2)} \) is \( e_2\varepsilon \)-stable and the remainder inequalities come from \( \mu(F_i) > \mu(F_{i+1}) \). On the other hand, we have \( \sum_{i=1}^n c_1(F_i) = c_1(W) \). So
\begin{equation}
\sum_{i=1}^n r_i b_i = a_2 - sh, \quad \sum_{i=1}^n r_i c_i = -sh.
\end{equation}
Finally, we calculate
\begin{equation}
c_2(W) = \sum_{i < j} c_1(F_i) \cdot c_1(F_j) + \sum_{i=1}^n c_2(F_i)
\end{equation}
\begin{equation}
\geq \frac{1}{2} \left( \sum_{i=1}^n c_1(F_i) \right)^2 - \sum_{i=1}^n c_1(F_i)^2 + \sum_{i=1}^n \frac{r_i - 1}{2r_i} c_1(F_i)^2
\end{equation}
\[= \frac{1}{2}(s^2h^2 - a_2^2)H^2 + \sum_{i=1}^n \frac{r_i}{2}(b_i^2 - 2b_i c_i)H^2.
\]
Here we have used Bogomolov’s inequalities \( 2r_i c_2(F_i) - (r_i - 1)c_1(F_i)^2 \geq 0 \). Combining (2.18) with \( c_2(W) \leq c_2(V) \), we have
\begin{equation}
(s^2h^2 - a_2^2) + \sum_{i=1}^n \left( r_i(b_i - c_i)^2 - r_i c_i^2 \right) \leq \frac{2(d_0 + C_3)}{H^2}.
\end{equation}
In the following, we will argue that there are \( h \) and \( \varepsilon_0 \) so that whenever \( 0 < \varepsilon < \varepsilon_0 \), then the only tuples \((b_i, c_i)\) that satisfy (2.16)–(2.19) must have \( c_i = 0 \) for \( i = \)
1, \ldots , n - 1. First of all, let \( \Lambda \) be the set of indices \( i \) so that \( c_i > 0 \). Then for those \( i \in \Lambda \), \( c_i \geq 1/rH^2 \) and by (2.16), for small \( \varepsilon \), we have

\[
(2.20) 
\quad b_i - c_i \leq \left( c_2 + \frac{a_2}{r} \right) + \frac{1}{\varepsilon} (-c_i) < \frac{1}{2\varepsilon} (-c_i).
\]

Thus

\[
\sum_{i=1}^{n} (r_i(b_i - c_i)^2 - r_ic_i^2) \geq \sum_{i \in \Lambda} r_i \left( \frac{1}{4\varepsilon^2} - 1 \right) c_i^2 - \sum_{i \notin \Lambda} r_i c_i^2.
\]

On the other hand, since \( \sum_{i \notin \Lambda} r_i c_i = -(sh + \sum_{i \in \Lambda} r_i c_i) \) and \( c_i \leq 0 \) for \( i \notin \Lambda \), \( \sum_{i \notin \Lambda} r_i c_i^2 \) is bounded from above by \( (sh + \sum_{i \in \Lambda} r_i c_i)^2 \) which in turn is no more than \( 2s^2h^2 + 2(\sum_{i \in \Lambda} r_i c_i)^2 \). Combined with (2.19), we must have

\[
(2.21) 
\quad \sum_{i \in \Lambda} r_i \left( \frac{1}{4\varepsilon^2} - 1 \right) c_i^2 - \left( 2s^2h^2 + 2(\sum_{i \in \Lambda} r_i c_i)^2 \right) + (s^2h^2 - a_2^2) \leq \frac{2(d_0 + C_3)}{H^2}.
\]

(2.21) is impossible if we assume

\[
(2.22) 
\quad \frac{1}{4\varepsilon^2} \geq r^2(p^2h^2 + r^2) \cdot H^2 + 2r(d_0 + C_3) + 4.
\]

Thus under the assumption (2.22), we must have \( c_i \leq 0 \) for all \( i \).

It remains to show that we can choose \( h \) large enough so that \( c_1 = \ldots = c_{n-1} = 0 \). Suppose there are \( c_{i_0} < c_i < 0 \). Then \( \sum_{i \neq i_0} r_i c_i = -sh - r_{i_0} c_{i_0} \). Again since \( c_i \leq 0 \), we have

\[
\sum_{i=1}^{n} r_i c_i^2 = r_{i_0} c_{i_0}^2 + \sum_{i \neq i_0} r_i c_i^2 \leq r_{i_0} c_{i_0}^2 + (sh + r_{i_0} c_{i_0})^2.
\]

Therefore, from (2.19), we have

\[
\frac{2(d_0 + C_3)}{H^2} \geq \left( s^2h^2 - a_2^2 \right) + \sum_{i=1}^{n} r_i(b_i - c_i)^2 - \sum_{i=1}^{n} r_i c_i^2 
\geq \left( s^2h^2 - a_2^2 \right) - \left( (sh + r_{i_0} c_{i_0})^2 + r_{i_0} c_{i_0}^2 \right) 
\geq \left( s^2h^2 - a_2^2 \right) - \left( (sh - \frac{1}{rH^2})^2 + \left( \frac{1}{rH^2} \right)^2 \right) 
= \frac{2sh}{rH^2} - \left( a_2^2 + \frac{2}{r^2H^2} \right).
\]

Clearly (2.23) is impossible if we choose

\[
(2.24) 
\quad h \geq 2r(d_0 + C_3 + r^2H^2) + 5.
\]

Now, we can choose \( h \) large according to (2.24) and then choose \( \varepsilon_0 \) small so that \( \varepsilon_0 \leq 1/2r \) and (2.22) holds with \( \varepsilon \) replaced by \( \varepsilon_0 \). Thus by our previous argument, if \( V_{_D} \) is not semistable, then in the filtration (2.15) all but one \( c_1(W_i/W_{i-1}) \cdot [\Sigma^+] = 0 \). We claim that \( c_1(W_i/W_{i-1}) \cdot [\Sigma^+] \neq 0 \). Indeed, assume \( c_j \neq 0 \), \( j < n \). Then \( c_j = -sh/r \) and then by (2.19),

\[
(2.25) 
\quad \left( \sum_{i \neq j} r_i b_i^2 \right) + r_j(b_j - c_j)^2 + \left( 1 - \frac{1}{r_j} \right) s^2h^2 - a_2^2 \leq \frac{2(d_0 + C_3)}{H^2}.
\]
Thus $|b_i| \leq 2\sqrt{d_0 + C_3}/\sqrt{H^2}$ for $i \neq j$ and $|b_j - c_j| < 2\sqrt{d_0 + C_3}/\sqrt{H^2}$. In particular, we will have

$$
\mu(F_j) = (b_j \varepsilon - \frac{(1 - \varepsilon)sh_j}{r_j})H^2 \leq b_n \varepsilon H^2 = \mu(F_n).
$$

This contradicts $\mu(F_i) > \mu(F_n)$. Thus we have proved the claim.

The next step is to reconstruct $V$ from the filtration $\{W_i\}$. We first construct a filtration of $V$ out of the filtration $\{W_i\}$ by letting $V_i \supseteq W_i$ be the subsheaf of $V$ so that $V/V_i$ is torsion free and $\text{rk}(W_i) = \text{rk}(V_i)$. We claim that $W_i = V_i$ for all $i \leq n - 1$. Indeed, let $V_i$ be the first among which $V_i \neq W_i$. Since $V_i = W_i$ on $\Delta \setminus D$, we must have

$$
c_1(V_i) = c_1(W_i) + \alpha[D], \quad \alpha \geq 1.
$$

On the other hand, $c_1(W_i) = (\sum_{j=1}^i r_j b_j) \Sigma^{-}$ and $|b_j| \leq 2\sqrt{d_0 + C_3}/\sqrt{H^2}$ because of (2.25). Thus

$$
\mu(V_i) = \frac{1}{\text{rk}(V_i)} c_1(V_i) \cdot H(\varepsilon)|_{\Delta}

= \frac{1}{\text{rk}(V_i)} \left( \sum_{j=1}^i r_j b_j \varepsilon + \alpha h \right)H^2 > \mu(V) + \frac{1}{\text{rk}(V_i)} \varepsilon H^2,
$$

which violates the $\varepsilon$-stability of $V$. Therefore, $V_i = W_i$ for all $i \leq n - 1$. In particular, the filtration

$$
0 = V_0 \subset V_1 \subset \cdots \subset V_n = V
$$

has the property that for $i \leq n - 1$, $V_i/V_{i-1}$ are $\mu$-semistable and $c_1(V_i/V_{i-1}) \sim r_i b_i [\Sigma^{-}]$. Let $F_i = V_i/V_{i-1}$. We intend to use induction on the rank $r$ to complete the proof of the proposition. In order to do this, we need to show that $F_n$ is $\mu$-semistable and

$$
c_2(V_i/V_{i-1}) - \frac{r_i - 1}{2r_i} c_1(V_i/V_{i-1})^2 \leq d_0 + C_3
$$

(2.26)

for all $i \leq n$. We show $F_n$ is $\mu$-semistable by showing that $r_n = 1$. Indeed, a combination of (2.25) (with $j = n$) and (2.24) guarantees $r_n = 1$. Thus $F_n$ is stable. Next, we have

$$
c_2(V) - \frac{r - 1}{2r} c_1(V)^2 = \sum_{i=1}^n c_2(F_i) + \sum_{i<j} c_1(F_i) \cdot c_1(F_j) + (-\frac{1}{2} + \frac{1}{2r})(\sum_{i=1}^n c_1(F_i))^2

= \sum_{i=1}^n \left( c_2(F_i) - \frac{r_i - 1}{2r_i} c_1(F_i)^2 - \frac{1}{2r_i} c_1(F_i)^2 \right) + \frac{1}{2r}(\sum_{i=1}^n c_1(F_i))^2

= \sum_{i=1}^n \left( c_2(F_i) - \frac{1}{2r_i} c_1(F_i)^2 \right) + \frac{1}{2r} c_1(V)^2 + \frac{1}{2} \sum_{i=1}^n \frac{1}{r_i} b_i^2 H^2.
$$

Because each $c_2(F_i) - \frac{1}{2r_i} c_1(F_i)^2$ is non-negative, (2.26) must be true. Therefore, we can apply the induction argument to $V_i/V_{i-1}$ to conclude that we can find large $h$ and small $c_0$ so that for any $\varepsilon < \varepsilon_0$, we must have $(V_i/V_{i-1})|_D$ semistable for generic $D \in |h\Sigma^+|$. Since $\deg(V_i/V_{i-1})|_D = 0$, $V_D$ must be semistable also. This completes the proof of Proposition 2.10.
3. Main theorems

In this section, we will prove our main theorems. We will show that when the second Chern class $d$ is large enough, then the moduli scheme $\mathcal{M}_X^{r,d}$ ($= \mathcal{M}_X^{r,d}(I, H)$) is smooth at a dense open subset. We shall further show that $\mathcal{M}_X^{r,d}$ is normal and for any constant $C$, there is an $N$ depending on $(X, I, H, r, C)$ such that whenever $d \geq N$, then

$$\text{codim}(\text{Sing } \mathcal{M}_X^{r,d}, \mathcal{M}_X^{r,d}) \geq C.$$ 

Finally, we will investigate the dependence of the moduli scheme $\mathcal{M}_X^{r,d}(I, H)$ on the polarization $H$. In case $r = 2$, Qin’s work [Qi] shows that for any two polarizations $H_1$ and $H_2$, the corresponding moduli spaces $\mathcal{M}_X^{r,d}(I, H_1)$ and $\mathcal{M}_X^{r,d}(I, H_2)$ are birational when $d$ is sufficiently large. Here, we shall demonstrate that similar phenomena also occur in high rank cases. But first, we shall continue our discussion of the degeneration to finish the proof our main technical theorem, Theorem 0.1.

For the moment, we shall keep the notation developed in §2. For any divisor $D \subseteq X$, we define $S_{e,I}^{r,d}$ be the set of all $e$-stable (with respect to the fixed $H$) rank $r$ sheaves $E$ of $\det E = I$ and $c_2(E) = d$ and define

$$S_{e,I}^{r,d}(D) = \{ E \in S_{e,I}^{r,d} \mid \text{Hom}(E, E(D)) \neq 0 \}.$$ 

Similarly, we define $\mathcal{V}_{e,I}^{r,d}$ and $\mathcal{V}_{e,I}^{r,d}(D)$ to be the subsets of locally free sheaves in $S_{e,I}^{r,d}$ and $S_{e,I}^{r,d}(D)$ respectively. For technical reasons, we will first attack the set $\mathcal{V}_{\mu,I}^{r,d}(D)$ which is the set of $\mu$-stable locally free sheaves $E$ with the mentioned constraint on $c_1, c_2$ and $h^0$. Namely, $\mathcal{V}_{\mu,I}^{r,d}(D) = \mathcal{V}_{0,I}^{r,d}(D)$. We shall prove

**Theorem 3.1.** For any choice of $r$, $I$ and $D$, and any constant $C_1$, there is a constant $N$ such that whenever $d \geq N$, we have

$$\#_{\text{mod}} \mathcal{V}_{\mu,I}^{r,d}(D) \leq \eta_X(r, d, I) - C_1.$$ 

**Proof.** Clearly, $\mathcal{V}_{\mu,I}^{r,d}(D)$ is a subset of $\mathcal{M}_X^{r,d}$. Since being locally free and $\mu$-stable are open conditions and having non-vanishing $\text{Hom}(E, E(D)) \neq 0$ is a closed condition, $\mathcal{V}_{\mu,I}^{r,d}(D)$ is a locally closed subset of $\mathcal{M}_X^{r,d}$. Let $A \subseteq \mathcal{M}_X^{r,d}$ be the closure of any irreducible component of $\mathcal{V}_{\mu,I}^{r,d}(D)$.

In the following, we seek to utilize the degeneration $\mathcal{M}_X^{r,d,\varepsilon} \rightarrow C$ (of the normalization of $\mathcal{M}_X^{r,d}$) constructed in Theorem 2.3. When $t \neq 0$, $\mathcal{M}_X^{r,d,\varepsilon}$ is just the normalization of $\mathcal{M}_X^{r,d}$. For such $t$, we let $W_i \subseteq \mathcal{M}_X^{r,d,\varepsilon}$ be the preimage of $A \subseteq \mathcal{M}_X^{r,d}$. $\bigcup_{t \neq 0} W_i$ is a constant family over $C^*$. We then let $W$ be the closure of $\bigcup W_i$ in $\mathcal{M}_X^{r,d,\varepsilon}$ and let $W_0$ be the special fiber of $W$ over $0 \in C$.

Here is our strategy: Take a large $h$ and a good $D^h \in |hH_C(h\Delta)|$. By Proposition 2.13, for any $t \neq 0$ and $c = \dim W_i$, the top self-intersection number $[L_{D^h}]^c(W_i) > 0$. Then since $W$ is flat and proper over $C$, $[L_{D^h}]^c(W_0) > 0$. Therefore, according to Proposition 2.14,

\begin{equation}
\#_{\text{mod}} \{ E^{(2)} \mid E \in W_0 \} = c.
\end{equation}
On the other hand, since every sheaf $E \in W_t$, $t \neq 0$, has non-vanishing $\text{Hom}(E, E(D))$, the upper-semicontinuity theorem tells us that there is a divisor $D' \subseteq \Delta$ (we can make it independent of the choice of $W_0$) such that for any $E \in W_0$,

$$\text{Hom}_\Delta(E^{(2)}, E^{(2)}(D')) \neq \{0\}.$$  

Thus, by applying Theorem 1.1, we get an upper bound of (3.1) and hence an upper bound of $c$. We now fill in the details of this approach.

To establish (3.2) for some $D' \subseteq \Delta$, we argue as follows: First of all, let $E \in W_0$ be any point. Since $W$ is flat over $C$, there is a smooth affine curve $S$ over $C$ and a flat family of torsion free sheaves $E_S$ on $Z_S = Z \times_S S$ such that for any closed $s \in S$ over $t \neq 0 \in C$, $E_s \in W_t$ and further, there is a closed $s_0 \in S$ over $0 \in C$ so that $E_{s_0} = E$. For any integer $k$, we consider the divisor $D_C - k\Delta$ on $Z$, where $D_C = p_X^*(D)$, and the pullback divisor (of $D_C - k\Delta$) on $Z_S$ which we denote by $D_k$. Clearly, the restriction of $D_k \otimes k(s_0)$ to $X \subseteq Z_0$ is $D - kH$. Now consider the vector space $\text{Hom}_{Z_0}(E_S, E_S(D_k))$. By assumption, for general $s \in S$, $\text{Hom}_{Z_0}(E_s, E_s(D_k)) \neq 0$. Thus $\text{Hom}_{Z_S}(E_S, E_S(D_k)) \neq \{0\}$. Let $w \in \text{Hom}_{Z_S}(E_S, E_S(D_k))$ be a non-trivial section and let $\xi$ be the uniformizing parameter of $S$ at $s_0$. Then because $E_S$ is flat over $S$, there is an $n \geq 0$ such that the restriction of $w/\xi^n$ to $Z_{s_0}$ gives rise to a non-trivial homomorphism $\varphi : E_{s_0} \rightarrow E_{s_0}(D_k)$.

Next, because $E^{(1)}$ is a quotient sheaf of $E$, $\varphi$ induces a homomorphism $E \rightarrow E^{(1)}(D_k)$ and further because $E^{(1)}$ is torsion free, it comes from $\varphi_1 : E^{(1)} \rightarrow E^{(1)}(D_k)$. Similarly, we have $\varphi_2 : E^{(2)} \rightarrow E^{(2)}(D_k)$. Because $E$ is torsion free, at least one $\varphi_i$ is non-trivial. Now we claim that we can choose a $k$ (independent of $d$ and $\varepsilon$) so that $\varphi_1$ is always trivial. Indeed, we first choose $k$ so that $H \cdot (D - kH) < 0$. Then since $\text{det} \varphi_1 \in H^0(O_X(rD - rkH)) = \{0\}$, $\text{det} \varphi_1$ is trivial. If we let $A \subset E^{(1)}$ be the kernel of $\varphi_1$, then $E^{(1)}/A$ is torsion free and further, there is a $\psi$ making the following diagram commutative

$$\begin{array}{ccc}
A & \longrightarrow & E^{(1)} \\
\downarrow \varphi_1 & & \downarrow \psi \\
E^{(1)}(D_k) & \longrightarrow & E^{(1)}(D_k) \\
\end{array}$$

(3.3)

On the other hand, by Lemma 2.4, there is a constant $e_1$ independent of $d$ and $\varepsilon$ such that $E^{(1)}$ is $e_1$-stable. Thus if $\varphi_2 \neq 0$, then $0 < \text{rk}(E^{(1)}/A) < \text{rk}(E^{(1)})$ and $E^{(1)}/A$ is both a subsheaf of $E^{(1)}(D_k)$ and a quotient sheaf of $E^{(1)}$. Therefore,

$$\mu(E^{(1)}(D_k)) + \frac{1}{\text{rk}(E^{(1)}/A)} \sqrt{H^2 \cdot e_1} > \mu(E^{(1)}) > \mu(E^{(1)}) - \frac{1}{\text{rk}(E^{(1)}/A)} \sqrt{H^2 \cdot e_1}.$$  

A straightforward calculation shows that this is impossible if we let

$$k > \frac{1}{H^2}(D \cdot H + 2\sqrt{H^2 \cdot e_1}).$$  

(3.4)

Hence $\varphi_1$ must be trivial.
From now on, we fix such a $k$. Then our previous argument shows that all $V \in \Theta = \{E^{(2)} \mid E \in W_0\}$ have non-vanishing $\text{Hom}_\Delta(V', V(D_k))_0$. As we explained, our intention is to apply Theorem 1.1 to the set $\Theta$ to get the bound:

\begin{equation}
\#_{\text{mod}} \Theta \leq \eta_X(r, d, I) - C_1, \quad d \gg 0.
\end{equation}

First of all, all $V \in \Theta$ are $e_2$-stable by Lemma 2.15 and have $\text{det} V = I_2$, $I_2 \in \Lambda = \{I_0(-r\Sigma^+), \ldots, I_0\}$. Next, for each $I_2 \in \Lambda$, there is an $\varepsilon_0(I_2) > 0$ specified by Theorem 1.1. We let $\varepsilon_0 = \min_{I_2 \in \Lambda}\{\varepsilon_0(I_2)\}$. Then for any $\varepsilon$ smaller than the $\varepsilon_0$, the ample divisor $H(\varepsilon)|_\Delta$ on $\Delta$ satisfies the condition of Theorem 1.1. In order to apply Theorem 1.1, we need to know that the general element of $\Theta$ is locally free, which certainly is quite delicate in general. The solution we propose is to use the double dual operation to relate any sheaf $F \in \Theta$ to its double dual $F^\vee = F^\vee \vee$. $F^\vee \vee$ is always locally free because $\Delta$ is a smooth surface. Assume $d_2 = c_3(F(F))$; then $d_2 \leq c_2(F)$ and the equality holds if and only if $F$ is locally free. Following the notation introduced at the beginning of §1, we have

\begin{equation}
F : \Theta \longrightarrow \bigcup_{d_2 \in \mathbb{Z}; I_2 \in \Lambda} \mathfrak{A}^{r, d_2}_{e_2, I_2}(D_k|_\Delta).
\end{equation}

(We use $\mathfrak{A}$ to denote sets related to $\Delta$ and use $\mathfrak{V}$ to denote sets related to $X$.) Here we have used the fact that $\text{Hom}(F, F(D_k|_\Delta))_0 \neq 0$ implies $\text{Hom}(F^\vee, F^\vee(D_k|_\Delta))_0 \neq 0$. Next, we divide $\Theta$ into subsets $\Theta_{d_1}$ according to the value of the second Chern class of $F \in \Theta$. Then, $\Theta = \bigcup \Theta_{d_1}$. We have the following estimate which will be proved shortly.

**Lemma 3.2.** For any $V \in \mathfrak{A}^{r, d_2}_{e_2, I_2}$, $\#_{\text{mod}}(F^{-1}(V) \cap \Theta_{d_1}) \leq (r + 1)(d_1 - d_2)$.

Now we are ready to complete the proof of the theorem. First of all, by applying Theorem 1.1 to the set $\mathfrak{A}^{r, d_2}_{e_2, I_2}(D_k|_\Delta)$, we know that for any constant $C_2$, there is an $N_2$ such that whenever $d_2 \geq N_2$, we have

\begin{equation}
\#_{\text{mod}} \mathfrak{A}^{r, d_2}_{e_2, I_2}(D_k|_\Delta) \leq \eta_\Delta(r, d_2, I_2) - C_2, \quad I_2 \in \Lambda.
\end{equation}

To control the left-hand side of (3.6) for small $d_2$, we invoke Theorem 1.5 to get

\begin{equation}
\#_{\text{mod}} \mathfrak{A}^{r, d_2}_{e_2, I_2} \leq \eta_\Delta(r, d_2, I_2) + C_3, \quad I_2 \in \Lambda,
\end{equation}

where $C_3$ is a constant. Another estimate we need was established in Proposition 2.7,

\begin{equation}
c_2(F) \leq d + C_4, \quad \forall F \in \Theta.
\end{equation}

The proof of (3.5) then goes as follows: For any constant $C_1$, we let $C_2$ be such that

\begin{equation}
C_2 \geq C + \eta_\Delta(r, C_4, I_2) - \eta_X(r, 0, I), \quad \forall I_2 \in \Lambda,
\end{equation}

and let $N_2$ be the constant that makes (3.6) hold. We then let $N$ be so that

\begin{equation}
(r + 1)(N - N_2) + \eta_\Delta(r, N_2, I_2) + C_3 \leq \eta_X(r, N, I) - C_1.
\end{equation}
We claim that when \( d \geq N \) and \( \varepsilon < \varepsilon_0 \), then (3.5) holds. Indeed, let \( d_1 \leq d + C_4 \) be any integer. Then for \( d_2 \leq N_2 \), by (3.7) and (3.10),

\[
\#_{\mathrm{mod}} \left( \Theta_{d_1} \cap \mathcal{F}^{-1}(\mathfrak{A}^{r,d_2}_{e_2,t_2}(D_{k|\Delta})) \right) \\
\leq (r + 1)(d_1 - d_2) + \#_{\mathrm{mod}} \left( \mathcal{F}(\Theta_{d_1}) \cap \mathfrak{A}^{r,d_2}_{e_2,t_2}(D_{k|\Delta}) \right) \\
\leq (r + 1)(d_1 - d_2) + \eta_{\Delta}(r, d_2, I) + C_3 \leq \eta_X(r, d, I) - C_1.
\]

Assume \( d_2 \geq N_2 \). By (3.6) and (3.9), we have

\[
\#_{\mathrm{mod}} \left( \Theta_{d_1} \cap \mathcal{F}^{-1}(\mathfrak{A}^{r,d_2}_{e_2,t_2}(D_{k|\Delta})) \right) \leq \#_{\mathrm{mod}} \mathfrak{A}^{r,d_2}_{e_2,t_2}(D_{k|\Delta}) + (r + 1)(d_1 - d_2) \\
\leq \eta_{\Delta}(r, d_2, I) - C + (r + 1)(d_1 - d_2) \leq \eta_X(r, d, I) - C_1.
\]

Thus we have established (3.5).

To finish the proof of the theorem, it suffices to show that \( \#_{\mathrm{mod}} \Theta = \#_{\mathrm{mod}} W_t \). For this, we will use Donaldson’s line bundle \( \mathcal{L}_{D^h} \). First of all, for any \( d \), we choose \( \varepsilon < \min(\varepsilon_0, \kappa(d)) \). (\( \kappa(d) \) was specified in Proposition 2.10.) We then apply Proposition 2.13 to the set \( W_t \), \( t \neq 0 \). Proposition 2.13 asserts that with \( c = \dim W_t \), \( [\mathcal{L}_{D^h}]^e(W_t) > 0 \). Since \( W \) is flat and proper over \( C \), we have

\[
[\mathcal{L}_{D^h}]^e(W_0) = [\mathcal{L}_{D^h}]^e(W_t) > 0.
\]

In particular, Proposition 2.14 tells us that then

\[
\#_{\mathrm{mod}} \{ E^{(2)} \mid E \in W_0 \} = c.
\]

Therefore, combined with inequality (3.5), we have that for \( d \geq N \),

\[
\dim \{ E \in \mathfrak{M}^{r,d}_X \mid \text{Hom}_X(E, E(D)) \neq 0 \} \leq \eta_X(r, d, I) - C.
\]

This completes the proof of Theorem 3.1.

Before we go any further, let us finish the proof of Lemma 3.2.

**Proof of Lemma 3.2.** The situation when \( r = 2 \) was proved in [L1, p. 461]. In general, let \( E \) be any rank \( r \) torsion free sheaf and let \( V = \mathcal{F}(E) \). Then \( E \) is uniquely determined by the quotient sheaf \( V \to V/E \), where \( V/E \) is supported on a discrete set of length \( \ell(V/E) = c_2(E) - c_2(V) \). Therefore, \( \mathcal{F}^{-1}(V) \cap \{ \text{sheaves of } c_2 = c_2(V) + c \} \) is exactly the set of all quotient sheaves \( V \to A \) such that \( A \) is supported on a discrete set and \( \ell(A) = c \). Let \( \text{Quot}^e_V \) be Grothendieck’s Quot-scheme of all quotient sheaves \( A \) of \( V \) with \( \ell(A) = c \). \( \text{Quot}^e_V \) is projective by [Gr, p. 13]. Observe also that when \( A \) is supported on \( c \) distinct points, then by [Gr, p. 21],

\[
(3.11) \quad \dim T_A \text{Quot}^e_V = (r + 1)c.
\]

Thus the lemma will be established if we can show that for any quotient sheaf \( A_0 \in \text{Quot}^e_V \), there is a deformation \( A_t \) of \( A_0 \) such that for generic \( t \), \( A_t \) is supported
on c distinct points [L1, p. 461]. In the following, we will demonstrate how to construct such a deformation.

Clearly, this is a local problem. Let $U$ be a classical neighborhood of $0 \in \mathbb{C}^2$ with coordinate $z = (z_1, z_2)$. Assume $A_0$ is a quotient sheaf of $\mathcal{O}_U^{gr}$ of length $c$ supported at the origin $0$. Let $E = \ker \{ \mathcal{O}_U^{gr} \to A_0 \}$. Along the lines of the argument given in [L1, p. 462], we can show that there are $f_1, \cdots, f_n \in \mathcal{O}_U^{gr}$ such that $\{ f_i \}_{i=r+1}^n$ are divisible by $z_1$ and $\{ f_i \}_{1}^n$ generate the submodule $E$.

Next, we define

$$f_i(z, t) = \begin{cases} f_i(z), & 1 \leq i \leq r; \\ (z_1 - t) \frac{f_i(z_1)}{z_1}, & r + 1 \leq i \leq n. \end{cases}$$

We then define a submodule $E_D \subset \mathcal{O}_{U \times D}^{gr}$, where $D$ is a small disk with parameter $t$, by

$$E_D = (f_1(z, t), \cdots, f_n(z, t)) \cdot \mathcal{O}_{U \times D} \subset \mathcal{O}_{U \times D}^{gr}.$$ 

Let $A_D = \mathcal{O}_{U \times D}^{gr} / E_D$. $E_D$ and $A_D$ can be viewed as families of sheaves parameterized by $D$. It is easy to see that when $E_D \otimes k(0)$ is torsion free, then $A_D \otimes k(0) = A_0$ and for $t$ small, $A_D$ is a (flat) deformation of $A_0$. Now we check that $E_D \otimes k(0)$ is torsion free. Suppose there are $h \in E_D$ and $f \in \mathcal{O}_U$ such that $f \cdot h = th'$ for some $h' \in E_D$. Let

$$h = \sum_{i=1}^n g_i(z, t) \cdot f_i(z, t).$$

Then the fact that $f(z) \cdot h \equiv 0 \mod(t)$ in $\mathcal{O}_{U \times D}^{gr}$ and that $f_1(z)/(z_1), \cdots, f_r(z)/(z_1)$ generate a rank $r$ $\mathcal{O}_U/(z_1)$-module implies

$$z_1 \mid g_i(z, 0), \quad i = 1, \cdots, r;$$

$$\sum_{i=1}^r \frac{g_i(z, 0)}{z_1} \cdot f_i(z) + \sum_{i=r+1}^n g_i(z, 0) \cdot \frac{f_i(z)}{z_1} \equiv 0. \tag{3.15}$$

Further, if we write $g_i(z, t) = \alpha_i(z) + t \beta_i(z, t)$, then the following identities hold in $\mathcal{O}_{U \times D}^{gr}$:

$$h = \sum_{i=1}^r \left( \alpha_i(z) f_i(z) + t \beta_i(z, t) f_i(z) \right)$$

$$+ \sum_{i=r+1}^n \left( \alpha_i(z) \frac{f_i(z)}{z_1} + t \beta_i(z, t) \frac{f_i(z)}{z_1} \right)$$

$$= (z_1 - t) \left( \sum_{i=1}^r \frac{\alpha_i(z)}{z_1} f_i(z) + \sum_{i=r+1}^n \frac{\alpha_i(z)}{z_1} f_i(z) \right)$$

$$+ t \left( \sum_{i=1}^r \frac{\alpha_i(z)}{z_1} + \beta_i(z, t) \right) f_i(z) + \sum_{i=r+1}^n \beta_i(z, t) (z_1 - t) f_i(z)$$

$$= th'',$$

$$= \cdots.$$
where $h''$ obviously belongs to $E_D$. Since $E_D$ is a submodule of $O_{U \times D}^{\oplus r}$, $h$ must be equal to $th''$ in $E_D$. Therefore, $f \cdot h_{|t=0} = 0$ implies $h_{|t=0} = 0$ in $E_D \otimes k(0)$ or that $F_D \otimes k(0)$ is torsion free.

In general, $A_t$ is not supported on $c$ distinct points. But at least we expect that $A_t$ is simpler than $A_0$, say $\text{Supp}(A_t)$ has at least two distinct points. In the following, we will show that this is indeed the case. Without loss of generality, we can assume that $f_t(z)$ all vanish at the origin. (Since otherwise, $A_0$ is essentially a quotient sheaf of $O_U^{\oplus (r-1)}$ and we can use induction on $r$ to take care of this situation.) For small $t$, the equation

\[(3.16) \quad f_1(t, z_2) \land \cdots \land f_r(t, z_2) = 0\]

has solutions, say $z_2 = w_2$, because $f_1(0) \land \cdots \land f_r(0) = 0$ and $f_1(0, z_2) \land \cdots \land f_r(0, z_2) \neq 0$ for generic $z_2$. Note that $(t, w_2) \in \text{Supp}(A_t)$. If $\text{Supp}(A_t)$ is a single point, then $f_{r+1}(z)/z_1, \cdots, f_n(z)/z_1$ must generate $O_U^{\oplus r}$ at the origin. Thus by discarding some extra terms, we will have $n = 2r$ and further, by eliminating terms in $f_1, \cdots, f_r$ that involve $z_1$ by using combinations of $f_{r+1}, \cdots, f_n$, we can assume $z_2|f_1(z), \cdots, z_2|f_r(z)$. Therefore, we can consider the deformation of $A_0$ derived from

\[E'_D = \left( (z_2 - t) \frac{f_1(z)}{z_2}, \cdots, (z_2 - t) \frac{f_r(z)}{z_2}, f_{r+1}(z), \cdots, f_n(z) \right).\]

In case $\text{Supp}(A'_t)$ is still a single point for generic $t$, then $\left( \frac{f_1(z)}{z_2}, \cdots, \frac{f_r(z)}{z_2} \right)$ will generate $O_U^{\oplus r}$ at 0 also. In particular, $A_0 = \bigoplus \mathbb{C}$ and then the desired deformation can be written by hand.

In the remainder of this section, we will complete the proof of the theorems stated at the beginning of this paper. We first investigate the sets $S_{e, I}^{r, d}$ and $S_{e, I}^{r, d}(D)$ introduced at the beginning of this section. We shall prove

**Theorem 3.3.** For any choice of $r$, $I$ and $D$ and any choice of constants $e$ and $C$, there is an integer $N$ such that whenever $d \geq N$, then we have

\[(3.17) \quad \#_{\mod S_{e, I}^{r, d}} = \eta(r, d, I),\]

\[(3.18) \quad \#_{\mod S_{e, I}^{r, d}(D)} \leq \eta_X(r, d, I) - C.\]

**Proof of (3.17).** Let $\mathcal{V}_{e, I}^{r, d} = S_{e, I}^{r, d} \cap \{\text{locally free sheaves}\}$ and let $\mathcal{V}_{e, I}^{r, d}(D) = \mathcal{V}_{e, I}^{r, d} \cap S_{e, I}^{r, d}(D)$. Clearly, (3.17) is a stronger statement than

\[(3.19) \quad \#_{\mod \mathcal{V}_{e, I}^{r, d}} = \eta_X(r, d, I),\]

which in turn is stronger (in case $e > 0$) than

\[(3.20) \quad \#_{\mod \mathcal{V}_{\mu, I}^{r, d}} = \eta_X(r, d, I).\]

Our strategy is first to prove statement (3.20) and then prove (3.19) and (3.17). We proceed by induction on the rank $r$. (3.17) and (3.20) are trivial when $r = 1$. 
For $r \geq 2$ and $E \in \mathcal{V}_{r,d}^*$, the Kodaira-Spencer-Kuranishi deformation theory tells us that there is a holomorphic map

$$f : U \subset H^1(X, \mathcal{E}nd^0(E)) \longrightarrow H^2(X, \mathcal{E}nd^0(E)),$$

where $U$ is an (analytic) neighborhood of the origin, such that $f^{-1}(0)$ is the versal deformation space of $E$. Since $h^0(\mathcal{E}nd^0(E)) = 0$ (since $E$ is $\mu$-stable),

$$\#_{\text{mod}}(\mathcal{V}_{r,d}^*, [E]) \geq h^1(\mathcal{E}nd^0(E)) - h^2(\mathcal{E}nd^0(E)),$$

and when $h^2(\mathcal{E}nd^0(E)) = 0$, $\#_{\text{mod}}(\mathcal{V}_{r,d}^*, [E]) = h^1(\mathcal{E}nd^0(E))$. Next, by Riemann-Roch, one calculates $\chi(\mathcal{E}nd^0(E)) = \eta_X(r, d, I)$. Thus one gets

$$\#_{\text{mod}}(\mathcal{V}_{r,d}^*, [E]) \geq \eta_X(r, d, I). \tag{3.21}$$

On the other hand, since $h^2(\mathcal{E}nd^0(E)) = h^0(\mathcal{E}nd^0(E) \otimes K_X)$, by Theorem 3.1, there is an $N$ such that whenever $d \geq N$, we have

$$\#_{\text{mod}}\{E \in \mathcal{V}_{r,d}^* \mid h^0(\mathcal{E}nd^0(E) \otimes K_X) > 0\} \leq \eta_X(r, d, I) - 1.$$

Therefore, for generic $E \in \mathcal{V}_{r,d}^*$, $\#_{\text{mod}}(\mathcal{V}_{r,d}^*, [E]) = \eta_X(r, d, I)$. Thus we have proved (3.20) provided $d \geq N$. To further attack (3.19) and (3.17), we need the following estimate which is interesting in its own right.

**Theorem 3.4.** For any choice of $r$, $I$ and two constants $c_1 > c_2$, there is a constant $C'$ such that

$$\#_{\text{mod}}(\mathcal{S}_{c_1, i} \setminus \mathcal{S}_{c_2, i}) \leq (2r - 1)d + C'. \tag{3.22}$$

**Proof.** Let $E$ be any torsion free sheaf in $\mathcal{S}_{c_1, i} \setminus \mathcal{S}_{c_2, i}$. Since $E$ is not $c_2$-stable, there is a torsion free subsheaf $F_1 \subseteq E$ such that $E/F_1$ is torsion free and that $\mu(F_1) \geq \mu(E) + e_1\sqrt{H^2}/\text{rk}(F_1)$. Because $E$ is $c_1$-stable, $\mu(F_1)$ is bounded from above by $\mu(E) + e_1\sqrt{H^2}/\text{rk}(F_1)$. Combined, we get

$$\frac{1}{r_1} I \cdot H + \frac{1}{r_1} e_1 \sqrt{H^2} \leq \frac{1}{r_1} I \cdot H \leq \frac{1}{r_1} I \cdot H + \frac{1}{r_1} e_1 \sqrt{H^2}, \tag{3.23}$$

where $r_1 = \text{rk} F_1$, $d_i = c_2(F_i)$ and $I_i = \text{det} F_i$ with $F_2 = E/F_1$. Note that $E$ belongs to the exact sequence

$$0 \longrightarrow F_1 \longrightarrow E \longrightarrow F_2 \longrightarrow 0. \tag{3.24}$$

We call $(r_i, d_i, I_i)$ admissible if they do come from (3.24) with $E \in \mathcal{S}_{c_1, i} \setminus \mathcal{S}_{c_2, i}$. We claim that $F_i$ are $e_3$-stable with $e_3 = e_1 + |e_2|$. Indeed, let $L \subset F_1$ be any subsheaf. Because $L$ is also a subsheaf of $E$,

$$\mu(L) < \mu(E) + \frac{1}{\text{rk}(L)} e_1 \sqrt{H^2} \leq \mu(F_1) - \frac{1}{r_1} e_2 \sqrt{H^2} + \frac{1}{\text{rk}(L)} e_1 \sqrt{H^2}.$$

Thus, $F_1$ is $e_3$-stable. $F_2$ is $e_3$-stable for the same reason. Therefore, $F_i \in \mathcal{S}_{e_3, i}$.

Finally, because of (3.24),

$$\#_{\text{mod}}(\mathcal{S}_{c_1, i} \setminus \mathcal{S}_{c_2, i}) \leq \sup_{(r_i, d_i, I_i)} \{\#_{\text{mod}}(\mathcal{S}_{e_1, i}) + \#_{\text{mod}}(\mathcal{S}_{e_2, i}) + \#_{\text{mod}}(\mathcal{S}_{e_3, i})\} + \sup\{\dim \text{Ext}^1(F_2, F_1) \mid F_1 \in \mathcal{S}_{e_3, i}\}, \tag{3.25}$$

where the supremum is taken over all admissible tuples $(r_i, d_i, I_i)$. Note that we only have numerical restriction on $I_i$ (cf. (3.23)) and $d_i$ can be small, thus we cannot expect an estimate of type (3.17) to hold for all $\mathcal{S}_{e_3, i}$. Nevertheless, we have
Lemma 3.5. There is a constant $C_1$ depending only on $r$, $e_3 > 0$ and $\deg I'$, $I' \in \text{Pic}(X)$, such that for $r' \leq r$, we have

$$\#_{\text{mod}} S_{e_3, I'}^{r', d'} \leq \eta_X(r', d', I') + C_1.$$ 

Proof. It suffices to show that there is a constant $C_1$ such that for any $E \in S_{e_3, I'}^{r', d'}$,

$$\dim \text{Ext}^1(E, E)^0 \leq 2r'd' - (r' - 1)I'^2 + C_1.$$ 

First, since $E$ is $e_3$-stable and $e_3 > 0$, $E$ is $2r_3$-stable. Hence by Lemma 1.8, both $h^0(E^d(E))$ and $h^0(E^d(E) \otimes K_X)$ are bounded from above by a constant, say $C_1$. By Serre duality, $\dim \text{Ext}^2(E, E)^0 = h^0(E^d(E) \otimes K_X)$. Therefore,

$$\dim \text{Ext}^1(E, E)^0 = 2r'd' - (r' - 1)I'^2 - (r' - 1)^2\chi(O_X)$$
$$+ \dim \text{Ext}^0(E, E)^0 + \dim \text{Ext}^2(E, E)^0$$
$$\leq 2r'd' - (r' - 1)I'^2 + (2C_1 - (r' - 1)^2\chi(O_X)).$$

This completes the proof of the lemma.

Returning to the proof of Theorem 3.4, we need to estimate the term $\dim \text{Ext}^1(F_1, F_2)$ in (3.25). First of all, by Riemann-Roch, for $F_i \in S_{e_3, I_i}^{r_i, d_i}$,

$$\dim \text{Ext}^1(F_1, F_2)$$
$$= \dim \text{Ext}^0(F_1, F_2) + \dim \text{Ext}^2(F_1, F_2)$$
$$- \left(\frac{r_1}{2}I_1^2 + \frac{r_1}{2}I_2^2 - \left(\frac{r_1}{2}I_2 - \frac{r_2}{2}I_1\right) \cdot K_X - I_1 \cdot I_2 + r_1 r_2 \chi(O_X) - r_1 d_2 - r_2 d_1\right).$$

Because $F_1$ and $F_2$ are $e_3$-stable, $F_1^\vee \otimes F_2$ and $F_2^\vee \otimes F_1$ are $2r_3$-stable. Also, the degree of $F_1^\vee \otimes F_2$ and $F_2^\vee \otimes F_1$ are bounded (from both sides) by constants depending on $r$, $e$ and $I \cdot H$. Thus, there is a constant $C_2$ depending on these parameters only so that

$$\dim \text{Ext}^0(F_1, F_2), \quad \dim \text{Ext}^2(F_1, F_2) \leq C_2.$$ 

Therefore, for any admissible $(r_i, d_i, I_i)$,

$$\#_{\text{mod}} S_{e_3, I_i}^{r_i, d_i} + \#_{\text{mod}} S_{e_3, I_i}^{r_2, d_2} + \sup \{ \dim \text{Ext}^1(F_1, F_2) \mid F_i \in S_{e_3, I_i}^{r_i, d_i} \}$$
$$\leq 2r_1 d_1 - (r_1 - 1)I_1^2 + 2r_2 d_2 - (r_2 - 1)I_2^2 + 2C_1$$
$$+ \left(\frac{r_2 I_2^2 - r_1 I_2^2}{2} + \frac{r_1 I_2 - r_2 I_1}{2} \cdot K_Xight)$$
$$+ I_1 \cdot I_2 - r_1 r_2 \chi(O_X) + r_1 d_2 + r_2 d_1 + 2C_2$$
$$\leq (2r - 1)d + (1 - r_2)d_1 + (1 - r_1)d_2 - (r_1 + \frac{r_2}{2} - 1)I_1^2 - (r_2 + \frac{r_1}{2} - 1)I_2^2$$
$$+ I_1 \cdot I_2 + \left(\frac{r_1}{2} - \frac{r_2}{2} I_1\right) \cdot K_X - r_1 r_2 \chi(O_X) + 2C_1 + 2C_2.$$
Thanks to Lemma 2.5, there is a constant $C_3 \leq 0$ depending on $r$, $e$ and $I \cdot H$ only such that $d_i - \frac{c_0}{2r_i} I_i^2 \geq C_3$. Thus combined with $d = c_2(E) = I_1 \cdot I_2 + d_1 + d_2$ and $I_2 = I - I_1$, the right-hand side of the above inequality is

\[(3.26)\]

\[
\leq (2r - 1)d - (r + \frac{1 - r_2}{2r_1} - \frac{3}{2}) I_1^2 - (r + \frac{1 - r_1}{2r_2} - \frac{3}{2}) I_2^2
\]

\[- (2r - 2) I_1 \cdot I_2 + (\frac{r_1}{2} I_2 - \frac{r_2}{2} I_1) \cdot K_X + r^2 |\chi(O_X)| + C_1 + 2C_2 - rC_3
\]

\[
= (2r - 1)d + \left( (1 + \frac{r_1 - 1}{2r_2} + \frac{r_2 - 1}{2r_1}) I_1^2 + \frac{r_2 + 1 - r_1}{2r_2} I_1 \cdot I_2 - \frac{r}{2} I_1 \cdot K_X \right) + C_4.
\]

Finally, because $|I_1 \cdot H| \leq |I \cdot H| + e \sqrt{H^2}$, the Hodge index theorem tells us that the sum of three middle terms in the last line of (3.26) is bounded from above by a constant $C_5$. Therefore, combined with (3.25), we have

\[
\#_{mod}(S_{r,d}^I \setminus S_{\mu,I}^r) \leq (2r - 1)d + C'.
\]

**Proof of (3.18).** We shall only consider the case where $e \geq 0$. The case $e < 0$ can be proved similarly. First of all, by letting $e_1 = e$ and $e_2 = 0$ in Theorem 3.4, we know that there is a constant $C_1$ such that $\#_{mod}(S_{r,d}^I \setminus S_{\mu,I}^r) \leq (2r - 1)d + C_1$. Then by choosing $N$ large, we have $\#_{mod}V_{r,d}^e = \eta_X(r, d, I)$ and $(2r - 1)d + C_1 \leq \eta_X(r, d, I)$ whenever $d \geq N$. Thus

\[
\#_{mod}V_{r,d}^e \leq \max \left\{ \#_{mod}V_{r,d}^e, \#_{mod} \left( V_{r,d}^e \setminus V_{\mu,I}^r \right) \right\} = \eta_X(r, d, I).
\]

To prove (3.17), we will use the double dual operation $\mathcal{F}$. Let

\[(3.27)\]

\[
\mathcal{F} : S_{r,d}^I \rightarrow \bigcup_{d' \leq d} V_{r,d}^{e,I}
\]

be the map sending $E$ to $E^{\vee \vee}$. Thanks to Lemma 3.2, we have

\[
\#_{mod}S_{r,d}^I \leq \sup_{d' \leq d} \left\{ \#_{mod}V_{r,d}^{e,I} + (r + 1)(d - d') \right\}.
\]

Further, let $C_1 \geq 0$ be a constant such that

\[
\#_{mod}V_{r,d}^{e,I} \leq \eta(r, d', I) + C_1.
\]

Then, for $d \geq N + C_1$, $\#_{mod}S_{r,d}^I$ is no more than either

\[
\sup_{d' \leq N} \left\{ \eta(r, d', I) + C_1 + (r + 1)(d - d') \right\} \leq \eta(r, N, I) + C_1 \leq \eta_X(r, d, I)
\]

or

\[
\sup_{N \leq d' \leq d} \left\{ \eta_X(r, d', I) + (r + 1)(d - d') \right\} \leq \eta_X(r, d, I).
\]

This establishes (3.17). (3.18) can be proved similarly based on Theorem 3.1. We shall omit it.
Proposition 3.6. With the notation as before and for any constant $N$ constant and when the codimension is at least 2, prove Theorem 0.2, it suffices to classify the singular locus of $U \rho + (\text{all quotient sheaves} E)$ is an ideal generated by at most $q$.

Proof. Let $\text{SL} \rho, \text{SL}$ is a principal bundle, the singular locus $\text{Sing}(U)$ is contained in $X r,d$ of all quotient sheaves, then $\text{U} \rho$ is a principal bundle, the singular locus $\text{Sing}(U)$ is contained in $X r,d$ of all quotient sheaves, then $\pi: U \rightarrow \pi(U) \subset M^r_d(I, H)$ is a principal $\text{SL} \rho, \text{SL}$-bundle. With this set-up in mind, one sees that in order to prove Theorem 0.2, it suffices to classify the singular locus of $U$.

Further, when the codimension is at least 1, then $U$ locally is a complete intersection, and when the codimension is at least 2, then $U$ is normal.

Proof. Let $E \in U$ be any quotient sheaf, let $q_2 = h^0(\text{End}^E(E) \otimes K_X)$ and let $q_1 = \eta_X(r, d, I) + (\rho^2 - 1) + q_2$. Then the argument in [L2, p. 8] demonstrates that the completion of the local ring of $U$ at $E$ is of the form $k[[t_1, \cdots, t_{q_1}]]/J$, where $J$ is an ideal generated by at most $q_2$ elements. In particular, for each component $U \subseteq U$, we always have

\[ \dim U \geq \eta_X(r, d, I) + (\rho^2 - 1). \]  

Next, by [Ar], [Mu], [Ma, p. 594], the singular locus $\text{Sing}(U)$ is exactly the set of all quotient sheaves $E$ with $\text{Ext}^2(E, E)^0 \neq 0$. By Theorem 3.3, for any constant $C$, there is an $N$ such that whenever $d \geq N$, the set

\[ U \cap S^r_d(K_X) = \{ E \in U \mid h^0(\text{End}^E(E) \otimes K_X) \neq 0 \} \]

obeys $\#_{\text{mod}}(U \cap S^r_d(K_X)) \leq \eta_X(r, d, I) - C$. Therefore,

\[ \dim \text{Sing}(U) \leq \#_{\text{mod}}(U \cap S^r_d(K_X)) + \dim SL(\rho) \leq \eta_X(r, d, I) + (\rho^2 - 1) - C. \]

When $C \geq 1$, this inequality and (3.28) imply that $U$ has pure dimension $\eta_X(r, d, I) + (\rho^2 - 1)$ and $\text{codim}(\text{Sing}(U), U) \geq C$. Because the completion of the local rings of $U$ are of the form $k[[t_1, \cdots, t_{q_1}]]/J$ with $J = (f_1, \cdots, f_{q_2})$, $U$ is a local complete intersection. $U$ will be normal if we further assume $\text{codim}(\text{Sing}(U), U) \geq 2$.

Corollary 3.7. Let $X$ be a smooth algebraic surface, $H$ an ample divisor and $I$ a line bundle on $X$. Let $r \geq 2$ be an integer. Then for any constant $C$, there is an $N$ such that whenever $d \geq N$, then $M^r_d(I, H)$ has pure dimension $\eta_X(r, d, I)$ and further, $\text{codim}(\text{Sing}(M^r_d), M^r_d) \geq C$.

Proof. Since $\pi: U^* \rightarrow \pi(U^*) \subset M^r_d$ is a principal bundle, the singular locus $\text{Sing}(M^r_d)$ is contained in $\pi(\text{Sing}(U^*)) \cap \pi(U \setminus U^*)$. 

By Proposition 3.6, we know that for $d$ large, we can arrange
\[
\text{codim}(\pi(\text{Sing}(U^s)), \mathcal{M}^{r,d}_X) \geq C.
\]
Therefore, to prove the corollary, we only need to find an upper bound of the dimension of $\pi(U \setminus U^s)$.

Let $E \in U \setminus U^s$. Then $E$ admits a filtration $0 = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_k = E$ such that $F_i = E_i/E_{i-1}$ are strictly stable. According to [Gi], $E$ and $gr(E) = \bigoplus F_i$ have the same image in $\mathcal{M}^{r,d}_X$ under $\pi$. Thus $\dim \pi(U \setminus U^s)$ can be bounded easily in terms of the dimension of moduli of lower rank stable sheaves. Similar to the proof of Theorem 3.4, we can show that there is a constant $C_1$ such that
\[
(3.29) \quad \dim \pi(U \setminus U^s) \leq (2r - 1)d + C_1.
\]
(If we let $\mathcal{S}^{r,d}_{s,1}$ be the set introduced in §2, $\pi(U \setminus U^s) \subseteq \mathcal{S}^{r,d}_{s,1} \setminus \mathcal{S}^{r,d}_{s,1}$ and then (3.29) follows directly from Theorem 3.4.) Thus for large $N$, we will have for $d \geq N$, $\dim \pi(U \setminus U^s) \leq \eta_X(r, d, I) - C$. This completes the proof of the corollary and Theorem 0.2.

**Corollary 3.8.** With the notation as before, there exists $N$ such that whenever $d \geq N$, then
1. $\mathcal{M}^{r,d}_X$ is normal. Further, if $s \in \mathcal{M}^{r,d}_X$ is any closed point corresponding to a stable sheaf, then $\mathcal{M}^{r,d}_X$ is a local complete intersection at $s$.
2. The set of locally free $\mu$-stable sheaves $(\mathcal{M}^{r,d}_X)^{vb} \subseteq \mathcal{M}^{r,d}_X$ is dense in $\mathcal{M}^{r,d}_X$.

**Proof.** Let $N$ be given by Proposition 3.6 so that whenever $d \geq N$, $U$ has pure dimension $\eta_X(r, d, I) + (g^2 - 1)$ and $\text{codim}(\text{Sing}(U), U) \geq 2$. Then since $U$ is normal, $\mathcal{M}^{r,d}_X$ must be normal and since $U^s$ is a local complete intersection, $\pi(U^s)$ is a good quotient and $U^s \rightarrow \pi(U^s)$ is a principal bundle. The last statement can be proved easily similar to that of Theorem 3.1. We shall omit it here.

The last subject we will study is the dependence of the moduli spaces on the choice of the polarizations. We prove

**Theorem 3.9.** For any choice $(r, I)$ and polarizations $H_1$ and $H_2$, there is a constant $N$ so that whenever $d \geq N$, then $\mathcal{M}^{r,d}_X(I, H_1)$ and $\mathcal{M}^{r,d}_X(I, H_2)$ are birational to each other.

**Proof.** Let $W \subseteq \mathcal{M}^{r,d}_X(I, H_1)$ be the set of quotient sheaves $E$ such that $E$ are not $H_2$-stable. Then every $E \in W$ belongs to the exact sequence
\[
0 \rightarrow F_1 \rightarrow E \rightarrow F_2 \rightarrow 0
\]
such that $\mu(F_1) \geq \mu(E)$ (with respect to $H_2$). Then by repeating the argument in Theorem 3.4, we can find a constant $C_1$ (depending on $H_1$ and $H_2$) such that $\text{dim} W \leq (2r - 1)d + C_1$. Therefore, by letting $N$ be large, we will have
\[
\text{dim} \mathcal{M}^{r,d}_X(I, H_1) = \text{dim} \mathcal{M}^{r,d}_X(I, H_2) = \eta_X(r, d, I)
\]
and $\text{dim} W \leq \eta_X(r, d, I) - 1$ provided $d \geq N$. Therefore, by the universality of the moduli scheme, there is a morphism
\[
\Phi : \mathcal{M}^{r,d}_X(I, H_1) \setminus W \rightarrow \mathcal{M}^{r,d}_X(I, H_2)
\]
which is generically one-one and onto. Thus $\mathcal{M}^{r,d}_X(I, H_1)$ is birational to $\mathcal{M}^{r,d}_X(I, H_2)$. 

REFERENCES


Department of Mathematics, University of California, Los Angeles, California 90024
E-mail address, D. Gieseker: dag@math.ucla.edu
E-mail address, J. Li: jli@math.ucla.edu