A GEOMETRIC JACQUET-LANGLANDS CORRESPONDENCE FOR PARAMODULAR SIEGEL THREEFOLDS

POL VAN HOFTEN

ABSTRACT. We study the Picard-Lefschetz formula for Siegel modular threefolds of paramodular level and prove the weight-monodromy conjecture for its middle degree inner cohomology. We give some applications to the Langlands programme: Using Rapoport-Zink uniformisation of the supersingular locus of the special fiber, we construct a geometric Jacquet-Langlands correspondence between GSp_4 and a definite inner form, proving a conjecture of Ibukiyama [21]. We also prove an integral version of the weight-monodromy conjecture and use it to deduce a level lowering result for cohomological cuspidal automorphic representations of GSp_4 .

1. Introduction

In this paper we will study the cohomology of paramodular Siegel threefolds, using geometric results of [62] and the Picard-Lefschetz formula of [6]. Our first main result (Theorem 2) is a geometric Jacquet-Langlands correspondence for such paramodular Siegel threefolds, which proves a conjecture of Ibukiyama [21] that we will explain below. Our second main result (Theorem 3) is a level lowering result for cohomological cuspidal automorphic representations of GSp_4 , in the spirit of Mazur's level lowering results for modular forms. Our main technical result (Theorem 1) is the weight-monodromy conjecture for the inner cohomology of Siegel threefolds over \mathbb{Q}_p with paramodular level structure at p.

1.1. **Ibukiyama's conjecture.** Let D be the quaternion algebra over \mathbb{Q} that is non-split over \mathbb{R} and \mathbb{Q}_p for a prime p and split at all other places. Let $S_k[\Gamma_0(p)]^{\text{new}}$ be the space of modular forms of weight $k \geq 2$ and level $\Gamma_0(p)$ that are p-new. A classical result of Eichler gives a Hecke equivariant injection (surjective for $k \geq 3$)

$$S_k[\Gamma_0(p)]^{\mathrm{new}} \hookrightarrow \mathcal{A}_k^{D^{\times}}[\mathcal{O}_D^{\times}],$$

where $\mathcal{A}_k^{D^{\times}}[\mathcal{O}_D^{\times}]$ denotes the space of algebraic modular forms for D^{\times} of weight k and level \mathcal{O}_D^{\times} . These can be described explicitly as 'functions' on the (finite) class set of D which are easier to understand computationally than modular forms. In fact this description is used in practice to compute bases of newforms for spaces of modular forms. This correspondence is a special case of the Jacquet-Langlands correspondence between the algebraic group $\operatorname{GL}_2/\mathbb{Q}$ and its inner form D^{\times}/\mathbb{Q} , proven in [24] using the trace formula.

The quaternionic similitude group $G = GU_2(D)$ associated with D is an inner form of GSp_4 , such that $G(\mathbb{R})$ is compact modulo centre. General conjectures of Langlands predict a transfer from automorphic representations of G to automorphic representations of GSp_4 . A particular instance of

1

Key words and phrases. Shimura varieties, Siegel modular forms, Weight-monodromy.

This work was supported by the Engineering and Physical Sciences Research Council [EP/L015234/1], The EPSRC Centre for Doctoral Training in Geometry and Number Theory (The London School of Geometry and Number Theory), University College London and King's College London.

this transfer was first conjectured by Ihara and Ibukiyama [19,23] and later extended by Ibukiyama. We will state a slight reformulation of Conjecture 5.1 of [21] below. Let $S_{k,j}[K(p)]$ denote the space of Siegel modular forms of weight k, j and level K(p), where K(p) is the paramodular group (c.f. Section 3). Let $\mathcal{A}_{k,j}^G[K_2(p)]$ denote the space of algebraic modular forms for G of weight k, j and level $K_2(p)$, where $K_2(p)$ is an analogue of the paramodular group (c.f. Section 4).

Conjecture 1 (Ibukiyama). For $k \geq 0$ and for $j \geq 3$ there is an injective map

$$S_{k,j}[K(p)]^{new} \hookrightarrow \mathcal{A}_{k,j}^G[K_2(p)],$$

which is Hecke-equivariant for the prime-to-p Hecke operators.

Remark 1.1.1. This conjecture can be used to do explicit computations with Siegel modular forms, we give some examples: In [8] the author constructs an algorithm computing Hecke eigenvalues of Siegel modular forms, assuming a form of Conjecture 1; in [31] the authors do extensive computations with algebraic modular forms for G; in [10] the author uses the conjecture to find computational evidence for Harder's conjecture on congruences between Siegel modular forms and elliptic modular forms.

It is an old idea of Serre that the classical Jacquet-Langlands correspondence between modular forms and quaternion modular forms can be realised geometrically by restricting mod p modular forms (which are sections of a line bundle on the modular curve $Y_0(N)$) to the supersingular locus. This mod p Jacquet-Langlands correspondence can be upgraded to a characteristic zero correspondence by considering the Picard-Lefschetz formula in étale cohomology [6] for the modular curve $Y_0(Np)$. Ghitza [15] generalised Serre's ideas to produce a mod p Jacquet-Langlands correspondence for Siegel modular forms (by restricting to the superspecial locus). Our approach to Conjecture 1 is then to use the Picard-Lefschetz formula for the Siegel threefold with paramodular level at p, which is possible because Yu [62] computed its singularities. Using results of [27] and [61] we can identify the finite set of singular points of $X \otimes \mathbb{F}_p$ with an adelic double quotient of the form

$$G(\mathbb{Q})\setminus G(\mathbb{A}_f)/(U^pK_2(p)).$$

Algebraic modular forms for G are basically functions on this finite set and the Picard-Lefschetz formula gives us a map from the space of these algebraic modular forms to the middle cohomology of $X_{\overline{\mathbb{Q}}_p}$. This map has an interpretation in terms of the action of the inertia group and our first result concerns this action.

Theorem 1. Let X/\mathbb{Q}_p be the Siegel threefold of neat level $U = U^pK(p)$ and let \mathbb{V} be an automorphic local system of L-vector spaces where L/\mathbb{Q}_ℓ is a finite extension $(\ell \neq p)$ of sufficiently regular weight. Then the weight-monodromy conjecture holds for $H^3_!(X_{\overline{\mathbb{Q}}_p}, \mathbb{V})$.

Remark 1.1.2. The result, without restrictions on the weight or level, also follows from Arthur's classification of automorphic forms on GSp_4 together with results on weight-monodromy for GL_4 by Caraiani [4]. The point of the theorem is to give a geometric proof, using the cohomological vanishing theorems of [30]. In fact we will need an integral refinement of the theorem for our level-lowering result, which does not follow from automorphic considerations.

Remark 1.1.3. When X is a non-compact Shimura variety of Hodge type, with minimal compactification $j: X \hookrightarrow X^*$, one expects the weight-monodromy conjecture to hold for $H^{\bullet}(X^*, j_{!*}\mathbb{V})$. Indeed, it follows from the usual weight-monodromy conjecture for smooth projective varieties, see Appendix A. In general the inner cohomology $H_1^{\bullet}(X, \mathbb{V})$ is only a sub-quotient of $H^{\bullet}(X^*, j_{!*}\mathbb{V})$ and

it is unclear to the author if one should expect that the weight-monodromy conjecture holds for it. However when X is a Siegel threefold, the inner cohomology $H_!^{\bullet}(X, \mathbb{V})$ is a direct summand of $H^{\bullet}(X^*, j_{!*}\mathbb{V})$, because it is equal to the cuspidal cohomology.

At this point we have all the ingredients to prove Conjecture 1, but we will actually prove more. We can work with arbitrary level away from p to transfer certain cuspidal automorphic representations π of GSp_4 to G, which is a geometric incarnation of the Jacquet-Langlands correspondence between G and GSp_4 . We will state a short version of our main theorem below, see Theorem 8.2.1 for a more general statement.

- **Theorem 2.** (1) Let π be a non-endoscopic cohomological cuspidal automorphic representation of GSp_4 such that π_∞ is in the discrete series and such that π_p is ramified and K(p)-spherical. Then there is a cuspidal automorphic representation σ of G such that $\pi_v \cong \sigma_v$ for finite places $v \neq p$, such that σ_p is $K_2(p)$ -spherical and with σ_∞ determined by π_∞ . Moreover, σ occurs with multiplicity one in the cuspidal spectrum of G.
 - (2) Let $k \geq 0, j \geq 3$ and let N be a squarefree integer with $p \mid N$, then there is an injective map $\varphi: S_{k,j}[K(N)]^{p-new} \hookrightarrow \mathcal{A}_{k,j}^G[K_2(N)]$

equivariant for the prime-to-p Hecke operators, which proves Conjecture 1.

- Remark 1.1.4. Sorensen constructs a Jacquet-Langlands transfer in [48] using the stable trace formula. However he has to assume that there is a certain auxiliary prime $q \neq p$ where σ_q is ramified enough, which is never satisfied for paramodular forms of squarefree level. Moreover, while I was writing up this paper, Weissauer and Rösner informed me that they can construct a general Jacquet-Langlands correspondence between GSp_4 and its inner forms, also using trace formula methods, see their forthcoming article [42].
- Remark 1.1.5. Ibukiyama gives a conjectural characterisation of the image of φ when N=p. We will give a different characterisation of its image in Theorem 8.2.1, which is probably easier to verify in practice. We will also discuss the image of the map $\pi \mapsto \sigma$ in Theorem 8.2.1.
- 1.2. Mazur's principle. In this section we describe a level lowering result for cohomological cuspidal automorphic representations of GSp_4 . We start by recalling a classical result of Mazur concerning levels of modular forms, which answers the following question: Given a normalised eigenform $f \in S_2[\Gamma_0(Np)]$ that is p-new, is there a congruence $f \equiv g \mod \ell$, for some $\ell \neq p$, with $g \in S_2[\Gamma_0(N)]$? We can translate this congruence into an isomorphism of mod ℓ Galois representations $\overline{\rho}_{f,\ell} \cong \overline{\rho}_{g,\ell}$. Since the Galois representation $\rho_{f,\ell}$ is unramified at p, a necessary condition for such a congruence to exist is that the Galois representation $\overline{\rho}_{f,\ell}$ is unramified at p, and Serre conjectured in [46] that the converse should be true. Ribet proved this conjecture in [39], which was famously used to show that modularity of semi-stable elliptic curves implies Fermat's last theorem. Below we state a slightly weaker version of Ribet's theorem, due to Mazur:

Theorem. (Mazur's principle, Theorem 6.1 in [39]) Assume that $\overline{\rho_{f,\ell}}$ is irreducible, unramified at p and that it has two distinct Frob_p eigenvalues (equivalently, $p \neq 1 \mod \ell$). Then there is a normalised eigenform $g \in S_2[\Gamma_0(N)]$ with $\overline{\rho_{g,\ell}} \simeq \overline{\rho_{f,\ell}}$.

Our second main result is the following analogue of Mazur's principle for cohomological cuspidal automorphic representations of GSp_4 .

Theorem 3. Let π be a cuspidal automorphic representation of GSp_4 that is cohomological of weight a > b > 0 such that π_p is ramified and K(p)-spherical. Let $U = U^p \cdot K(p) \subset \mathrm{GSp}_4(\mathbb{A}_f)$ be a neat

compact open subgroup such that $\pi^U \neq 0$. Let $\ell \neq p$ be a prime such that $\overline{\rho_{\pi,\ell}}$ is irreducible, such that the group U_ℓ is hyperspecial at ℓ and such that $\ell > a + b + 4$. Then if $\overline{\rho_{\pi,\ell}}$ is unramified at p and has four distinct Frob_p eigenvalues, there exists a cuspidal automorphic representation π' , of the same weight and level U^p away from p, such that $\overline{\rho_{\pi,\ell}} = \overline{\rho_{\pi',\ell}}$ and such that π'_p is unramified.

Remark 1.2.1. We need the assumption that a > b > 0, that $\ell > a + b + 4$ and that K_{ℓ} is hyperspecial to apply Theorem 10.1 of [30], which is a torsion vanishing result. Our proof only needs this torsion vanishing result after localising at a 'nice' maximal ideal of the Hecke algebra, so it should be possible to relax these assumptions. After a first version of this paper appeared on the arxiv, Haining Wang proved a version of Proposition 7.1.2 localised at such a maximal ideal (see Proposition 5.3 of [57]). This means that the assumption that a > b > 0 and that K_{ℓ} hyperspecial is hyperspecial is no longer necessary. The assumption that $\overline{\rho_{\pi,\ell}}$ is irreducible and has four distinct Frobenius eigenvalues is crucial for the actual level lowering argument. We note that Haining Wang proves a version of Theorem 3 under different assumptions, see Theorem 5.1 of [57]. To be precise he has weaker assumptions on the Frobenius eigenvalues, but has to assume semi-simplicity of the middle degree étale cohomology of the Siegel threefold.

An essential ingredient in the proof of Mazur's principle is the fact that the component group of the Jacobian of the modular curve $X_0(Np)$ is Eisenstein. Following Jarvis and Rajaei [25, 38] we define a cohomological analogue of the component group and show that it is zero if the aforementioned torsion-vanishing result holds. This statement can be seen as an integral refinement of the weight-monodromy conjecture because the component group is defined as the cokernel of a certain monodromy operator.

Remark 1.2.2. Gee and Geraghty [13] prove more general level lowering results for cuspidal automorphic representations of GSp₄, under the assumption that $\overline{\rho_{\pi,\ell}}$ has large image and is 'ordinary', see Theorem 7.5.2 of op. cit. for a precise statement. It might be possible to generalise their results using recent work of Yamauchi [60]. Moreover Sorensen proves a potential level lowering result for GSp₄ in [49].

Remark 1.2.3. The paramodular group K(Np) is not neat, so a level lowering result from $S_{k,j}[K(Np)]$ to $S_{k,j}[K(N)]$ does not follow. However, it should be possible to prove such a result by cleverly choosing some auxiliary level structure (but we would need some extra conditions on $\overline{\rho_{\pi,\ell}}$, c.f. the main theorem of [25]).

1.3. Overview of the paper. We start by giving an overview of the theory of Siegel modular forms and automorphic representations of $\mathrm{GSp}_4(\mathbb{A})$ in Section 2. In Section 3 we discuss the Vogan-Zuckerman classification and Arthur's classification and use these to describe the cohomology of Siegel threefolds. We define algebraic modular forms for G in Section 4. In Section 5 we survey some results on integral models of Siegel threefolds and describe the combinatorics of their supersingular loci explicitly in terms of Shimura sets for G. Section 7 is the technical heart of the paper where we prove the vanishing of the component group. We deduce the main theorems from this in Section 8.

2. Automorphic forms

In this section we will discuss (cuspidal) automorphic representations of $GSp_4(\mathbb{A})$. We will first describe classical Siegel modular forms of vector-valued weight for GSp_4 , which give rise to such automorphic representations. We then carry out some local computations needed in our proof of Mazur's principle. We end by recalling the Galois representations associated with cohomological cuspidal automorphic representations of $GSp_4(\mathbb{A})$.

2.1. Siegel modular forms of genus 2. We define the group GSp_4 as the group scheme over \mathbb{Z} defined by the functor sending a commutative ring R to

$$\{(g,\lambda) \in \operatorname{GL}_4(R) \times R^{\times} \mid J = \lambda g J g^t \},$$

where $J = \begin{pmatrix} 0 & 1_2 \\ -1_2 & 0 \end{pmatrix}$ and 1_2 is the 2×2 identity matrix. The projection map $(g, \lambda) \mapsto \lambda$ is called the similitude character and is denoted by

$$sim : GSp_4 \to \mathbb{G}_m$$
.

The group $\mathrm{GSp}_4(\mathbb{R})$ acts transitively on the Siegel upper half space \mathbb{H}_2^{\pm} defined as

$$\{z = x + iy \in \operatorname{Mat}_{2 \times 2}(\mathbb{C}) \mid z \text{ symmetric and } y \text{ positive- or negative definite} \}$$

by $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := (az+b)(cz+d)^{-1}$. Let $\Gamma \subset \mathrm{GSp}_4(\mathbb{Q})$ be a congruence subgroup and for $k \geq 0, j \geq 0$ consider the irreducible representation of $\mathrm{GL}_2(\mathbb{C})$ given by

$$W_{k,j} := \operatorname{Sym}^k V \otimes \det^j V,$$

where V is the standard representation of $GL_2(\mathbb{C})$.

Definition 2.1.1. A Siegel modular form of weight k, j and level Γ is a holomorphic function $f: \mathbb{H}_2^{\pm} \to W_{k,j}$ such that such that for all $z \in \mathbb{H}_2^{\pm}$ and all $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ we have

$$f\left(\frac{az+b}{cz+d}\right) = \rho_{k,j}(cz+d)f(z).$$

A Siegel modular form is called a cusp-form if $\Phi(f) = 0$, where $\Phi(f)$ is the Siegel operator (c.f. [55, pp. 12]), we write $S_{k,j}[\Gamma]$ for the \mathbb{C} -vector space of cusp forms of weight k, j and level Γ .

We will be particularly interested in the case that Γ is given by the paramodular group K(N) for some squarefree N. Here

$$K(N) = \left\{g \in \mathrm{GSp}_4(\mathbb{Q}) \cap \left(\begin{smallmatrix} \mathbb{Z} & N\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ \mathbb{Z} & \mathbb{Z} & \mathbb{Z} & N^{-1}\mathbb{Z} \\ \mathbb{Z} & N\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \end{smallmatrix} \right) \right\}.$$

We usually like to think of our levels as compact open subgroups in $GSp_4(\mathbb{A}_f)$ when translating Siegel modular forms into automorphic representations. For a prime p we define

$$H = H(p) := \operatorname{GSp}_{4}(\mathbb{Z}_{p})$$

$$K = K(p) := \left\{ g \in \operatorname{GSp}_{4}(\mathbb{Q}_{p}) \cap \begin{pmatrix} \mathbb{Z}_{p} & p\mathbb{Z}_{p} & \mathbb{Z}_{p} & \mathbb{Z}_{p} \\ \mathbb{Z}_{p} & \mathbb{Z}_{p} & \mathbb{Z}_{p} & p^{-1}\mathbb{Z}_{p} \\ \mathbb{Z}_{p} & p\mathbb{Z}_{p} & \mathbb{Z}_{p} & \mathbb{Z}_{p} \end{pmatrix} : \operatorname{sim}(g) \in \mathbb{Z}_{p}^{\times} \right\}$$

$$Q = Q(p) := K(p) \cap H(p),$$

we call K the paramodular group and Q the Klingen parahoric. The paramodular group K(N) now corresponds to the compact open subgroup

$$\prod_{p \mid N} K(p) \prod_{p \nmid N} \mathrm{GSp}_4(\mathbb{Z}_p) \subset \mathrm{GSp}_4(\mathbb{A}_f).$$

We will implicitly use the following result throughout the paper (c.f. [2, Theorem 2]):

Proposition 2.1.2. Given a normalised eigenform $f \in S_{k,j}(K(N))$ with N squarefree then there is an associated cuspidal automorphic representation π of $GSp_4(\mathbb{A})$ such that π_{∞} is in the holomorphic discrete series and such that π_f is unramified for $p \nmid N$ and $\pi_f^{K(p)} \neq 0$ for $p \mid N$.

2.2. The local Langlands correspondence. This section is about irreducible smooth representations π of $\mathrm{GSp}_4(\mathbb{Q}_p)$, that are ramified and K(p)-spherical, and their associated Weil-Deligne representations under the local Langlands correspondence of [12]. Table A.13 of [41] describes all of these representations and their associated Weil-Deligne representations can be found in Chapter 2.4 of op. cit. These representations break up into 5 types (IIa, IVc, Vb, Vc, VIc) of which only one (type IIa) is generic (c.f. Table A1 of op. cit.).

Let v be the normalised absolute value $v: \mathbb{Q}_p^{\times} \to \overline{\mathbb{Q}_\ell}$, let σ be an unramified character of \mathbb{Q}_p^{\times} and let χ be a character of \mathbb{Q}_p^{\times} such that $\chi^2 \neq v^{\pm 1}$ and $\chi \neq v^{\pm 3/2}$. We consider all of these as character of the Weil group via local class field theory (uniformisers go to geometric Frobenius elements). The representation $\chi \operatorname{St}_{\operatorname{GL}_2} \rtimes \sigma$ of type IIa is generic and its associated Weil-Deligne representation is given by:

$$\begin{pmatrix} \chi^2 \sigma & 0 & 0 & 0 \\ 0 & v^{1/2} \chi \sigma & 0 & 0 \\ 0 & 0 & v^{-1/2} \chi \sigma & 0 \\ 0 & 0 & 0 & \sigma \end{pmatrix}, \qquad N := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We note that the central character of this representation is $\chi^2 \sigma^2$ which is also equal to the similitude character of the Weil-Deligne representation. Furthermore, the Weil-Deligne representation satisfies the weight-monodromy conjecture, i.e., it is pure in the sense of Taylor-Yoshida [52, pp.6]. More concretely, this means that the monodromy operator has rank 'as large as possible'.

Remark 2.2.1. The other four types (IVc, Vb, Vc, VIc) are not generic, and their associated Weil-Deligne representations do not satisfy the weight-monodromy conjecture. After we prove Theorem 1 (weight-monodromy), we can use this observation to deduce that certain representations must be of type IIa, which will be used in the proof of Theorem 3.

2.2.2. Atkin-Lehner eigenvalues. The element

$$u := \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ p & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \end{pmatrix}$$

normalises the paramodular group K and so it acts on the space of invariants π^K . Moreover since $u^2 = pI$ is central it must act on π^K via the central character of π .

Lemma 2.2.3. Let $\pi = \chi \operatorname{St}_{\operatorname{GL}_2} \rtimes \sigma$ be a representation of type IIa with χ, σ unramified and central character $\chi^2 \sigma^2$. Then π^K is one-dimensional and u acts via the scalar $(\chi \sigma)(p)$

Proof. When $\chi \sigma$ is trivial, Roberts and Schmidt compute that u acts by (c.f. Table A.12 in [41])

$$(\chi \sigma)(p) = 1,$$

one can reduce to this case by twisting and using the fact that χ, σ are unramified.

2.3. Galois representations. Galois representations associated with cohomological cuspidal automorphic representations of $GSp_4(\mathbb{A})$ were constructed in [51], [32], [58]. They satisfy local-global compatibility with the local Langlands correspondence of [12] by work of Mok [35, Theorem 3.5]. To summarise we state the following version of Theorem 3.5 of [35]

Theorem 2.3.1. Let π be a cohomological cuspidal automorphic representation of $GSp_4(\mathbb{A})$ that is not CAP. Then for every prime number ℓ there exists a continuous Galois representation

$$\rho_{\pi,\ell}: \operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \operatorname{GSp}_4(\overline{\mathbb{Q}_\ell})$$

such that:

- (1) The representation is unramified at primes $p \neq \ell$ where π_p is unramified, de Rham at ℓ and crystalline at ℓ if π_{ℓ} is unramified.
- (2) For all primes $p \neq \ell$ we have local-global compatibility with the local Langlands correspondence of [12]. To be precise we have

$$WD\left(\rho_{\pi,\ell}\big|_{G_{\mathbb{Q}_p}}\right)^{F\text{-}ss} \cong rec_p(\pi_p \otimes sim(v^{3/2})),$$

where rec_p is the Langlands reciprocity map of [12] and F-ss denotes Frobenius semi-simplification of a Weil-Deligne representation.

(3) The representation satisfies

$$\rho_{\pi,\ell}^{\vee} \cong \rho_{\pi,\ell} \otimes \operatorname{sim}(\rho_{\pi,\ell})$$

where $sim(\rho_{\pi,\ell})$ is the similar character of $\rho_{\pi,\ell}$.

Remark 2.3.2. If π is CAP then the above theorem also holds except that local global compatibility is only true up to semi-simplification (this is Proposition 3.4 of [35]).

3. Cohomology of Siegel Threefolds

In this section we will describe the inner cohomology $H_!^3$ of Siegel threefolds in terms of cuspidal automorphic representations of GSp_4 using the Vogan-Zuckerman classification (c.f. [37,51]). We then make use of Arthur's classification [1,14] of these cuspidal automorphic representations to make our discussion explicit.

3.1. Automorphic local systems. Let Y_U be the Shimura variety for $(GSp_4, \mathbb{H}_2^{\pm})$ of level $U = U^{\ell}U_{\ell} \subset GSp_4(\mathbb{A}_f)$. Irreducible representations V of $GSp_4(\mathbb{C})$ give rise to local systems \mathbb{V} on Y_U , called automorphic local systems, as in [53, pp.24], [29, section 3]):

Recall that irreducible representations V of $\mathrm{GSp}_4(\mathbb{C})$ are parametrized by tuples $(a \geq b \geq 0, c)$, where $a \geq b \geq 0$ determines the restriction of V to Sp_4 and c determines the action of the similitude character. If $\ell > a+b+4$ then the lattice $V_{\mathcal{O}}$ is unique up to homothety because $V_{\mathcal{O}}/\varpi$ is an irreducible representation of H, which makes the local system $\mathbb{V}_{\mathcal{O}}$ independent on the choice of lattice. Changing c amounts to Tate-twisting the local system, which has no interesting effect on the cohomology. Therefore we will set c = a + b, which normalises things such that $\mathbb{V}_{1,0}$ is the local system corresponding to the (dual of the) relative ℓ -adic Tate module of the universal abelian variety. In general the local system $\mathbb{V}_{a,b} \otimes L$ occurs as a direct summand of

$$\mathbb{V}_{1.0}^{\otimes (a+b)}$$

and these local systems are self-dual up to a Tate twist.

3.2. Cohomology theories. Before studying the Galois representations that occur in the étale cohomology of Y_U , we will first study the Hodge structure on the singular cohomology. Write $\mathbb{V}_{\mathbb{C}}$ for the \mathbb{C} -linear local system on $Y_U(\mathbb{C})$ of weight $a \geq b \geq 0$. One is usually interested in the intersection cohomology

$$H^{\bullet}(\overline{Y}_U(\mathbb{C}), j_{!*}\mathbb{V}_{\mathbb{C}}),$$

where $j: Y_U \to \overline{Y}_U$ is the inclusion of Y_U into its Baily-Borel compactification \overline{Y}_U . By the Zucker conjecture (proven in [33] and [44]) this is equal to the L^2 cohomology

$$H_{(2)}^{\bullet}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}),$$

which can be described in terms of automorphic representations as follows:

$$H_{(2)}^{\bullet}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}) = \bigoplus_{\pi} m(\pi) \pi_{\text{fin}}^U \otimes H^{\bullet}(\mathfrak{g}, K_{\infty}; \mathbb{V}_{\mathbb{C}} \otimes \pi_{\infty}),$$

where the sum runs over automorphic representations $\pi = \pi_{\infty} \otimes \pi_{\text{fin}}$ occurring in the discrete spectrum of $L^2(\text{GSp}_4(\mathbb{Q}) \setminus \text{GSp}_4(\mathbb{A}))$ with multiplicity $m(\pi)$. There is a natural direct summand $H^{\bullet}_{\text{cusp}}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}})$ of cuspidal cohomology which is the direct sum over cuspidal automorphic representations π . Moreover it follows from the results announced in [1] and proven in [14] that $m(\pi) \leq 1$ for all π . In this article we will work with the inner cohomology which is defined to be the image of the natural map

$$H_!^{ullet}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}) := \operatorname{Im} \left(H_c^{ullet}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}) \to H^{ullet}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}) \right)$$

where subscript c denotes compactly supported cohomology. In middle degree there is an isomorphism (c.f. [51, pp. 294])

$$H^3_!(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}) \cong H^3_{\text{cusp}}(Y_U(\mathbb{C}), \mathbb{V}_{\mathbb{C}}).$$

3.3. Vogan-Zuckerman classification. The Vogan-Zuckerman classification (c.f. [56]) describes the possible π_{∞} for which

$$H^i(\mathfrak{g}, K_{\infty}; \mathbb{V}_{\mathbb{C}} \otimes \pi_{\infty})$$

is nonzero, where $\mathbb{V}_{\mathbb{C}}$ is a local system of weight $a \geq b \geq 0$. If π_{∞} is in the discrete series, then it is always cohomological. It turns out that the discrete series representations are the only ones that contribute to middle degree cuspidal cohomology. For a fixed choice of weight a, b, there are precisely two choices of π_{∞} in the discrete series, which we will denote by π^H and π^W , where the H stands for 'holomorphic' and the W for 'Whittaker' (the latter is generic). The $(\mathfrak{g}, K_{\infty})$ -cohomology of these representations is two-dimensional and occurs in degree 3 with Hodge numbers (3,0), (0,3) and (2,1), (1,2) respectively. We note that the representations π with $\pi_{\infty} = \pi^H$ correspond to holomorphic cuspidal Siegel modular forms of weight a - b, b + 3 (c.f. Proposition 2.1.2).

- 3.4. **Arthur's classification.** Arthur has classified the cuspidal automorphic representations π in [1] into six different types. Only three of these contribute to the inner cohomology in middle degree:
 - The cuspidal automorphic representations π of general type, which occur with multiplicity one. These are stable which means that both $\pi_{\text{fin}} \otimes \pi^H$ and $\pi_{\text{fin}} \otimes \pi^W$ are cuspidal automorphic. This implies that the four-dimensional Galois representation associated with π occurs in the inner cohomology of the Siegel threefold.
 - The cuspidal automorphic representations π of Yoshida type which occur with multiplicity $m(\pi) = 1$. These are unstable which means that precisely one of $\pi_{\text{fin}} \otimes \pi^H$ and $\pi_{\text{fin}} \otimes \pi^W$ is cuspidal automorphic. These are related to pairs (π_1, π_2) of cuspidal automorphic representations of GL₂, which we will make precise in Section 3.5.
 - The cuspidal automorphic representations π of Saito-Kurokawa type which are CAP and occur with multiplicity one in the cuspidal spectrum. They only occur in scalar weight (so a = b) and have π_{∞} isomorphic to either $\pi^H, \pi^W, \pi^+, \pi^-$ (the latter two don't contribute to middle degree cohomology). These lifts come from cuspidal automorphic representations σ of GL₂ of weight a + b + 4 and the associated Galois representation is

$$\rho_{\pi \ell} = \rho_{\sigma \ell} \oplus \mathbb{Q}_{\ell}(-b-2) \oplus \mathbb{Q}_{\ell}(-b-1).$$

This Galois representation cannot occur in the middle degree cohomology by purity (outside a finite set of places the inner cohomology will be pure of weight a + b + 4). In fact we will only see the two-dimensional Galois representation $\rho_{\sigma,\ell}$ in the middle cohomology.

3.5. Yoshida lifts. We will now describe Yoshida lifts in terms of cuspidal automorphic representations of GL_2 . Let π_1, π_2 be cuspidal automorphic representations of $GL_2(\mathbb{A})$ with the same central character and suppose that $\pi_{1,\infty}, \pi_{2,\infty}$ are in the holomorphic discrete series with Harish-Chandra parameters $r_1 - 1, r_2 - 2$ satisfying $r_1 > r_2 \ge 2$ with $r_1 = a + b + 4$ and $r_2 = a - b + 2$.

Definition 3.5.1. Let π be a cuspidal automorphic representation of $GSp_4(\mathbb{A})$ that is not CAP and such that π_{∞} is in the discrete series. We say that π is a Yoshida lift of (π_1, π_2) if for almost all places v the Satake parameter π_v agrees with the Satake parameter of (π_1, π_2) via the inclusion of dual groups $GL_2 \times_{\mathbb{G}_m} GL_2 \to GSp_4$.

Alternatively, we can explicitly construct an L-packet $L(\pi_1, \pi_2)$ as the restricted tensor product of local L-packets $L(\pi_{1,v}, \pi_{2,v})$ for all places v. These local L-packets are constructed using the θ -correspondence between GSp_4 and $(\mathrm{GL}_2 \times \mathrm{GL}_2/\mathbb{G}_m)$ and have size two if v is infinite or if v is finite and both $\pi_{1,v}$ and $\pi_{2,v}$ are essentially square-integrable. In all other cases the local L-packet has size one and moreover always contains a generic member. However not all members of this L-packet are cuspidal, which is quantified in the following result of Weissauer.

Theorem 3.5.2. (c.f. Theorem 5.2 of [59]) A cuspidal automorphic representation π is a Yoshida lift of (π_1, π_2) if and only if it is a cuspidal member of the L-packet $L(\pi_1, \pi_2)$. Moreover, elements π of this L-packet occur in the cuspidal spectrum with multiplicity

$$m(\pi) = \frac{1 + (-1)^{(c(\pi))}}{2}$$

where $c(\pi)$ is the number of places where π_v is non-generic. Moreover π_{∞} is cohomological with respect to the local system of weight a, b.

The following result describes what Galois representation occur in the inner cohomology:

Proposition 3.5.3. (c.f. Corollary 4.2 of [59]) Let π be a Yoshida lift of (π_1, π_2) , then π is in the discrete series which means that the $(\mathfrak{g}, K_{\infty})$ -cohomology of π_{∞} is two-dimensional. The Galois action on this two-dimensional piece is given by

$$\begin{cases} \rho_{\pi_1,\ell} & \text{if } \pi_{\infty} = \pi^H \\ \rho_{\pi_2,\ell}(-b-1) & \text{if } \pi_{\infty} = \pi^W \end{cases}.$$

Definition 3.5.4. We call π an irrelevant Yoshida lift if π_p is ramified but the associated two-dimensional Galois representation occurring in cohomology is unramified.

Corollary 3.5.5. If $\pi \in L(\pi_1, \pi_2)$ has nonzero invariant vectors for K(N) for squarefree N, then $L(\pi_1, \pi_2)$ has precisely one cuspidal member π , with π_v generic for all places v. In particular there are no holomorphic Yoshida lifts of level K(N) and so the Galois representation that we see in the cohomology is

$$\rho_{\pi_{2,\ell}}(-b-1).$$

Proof. This is just Remark 3.5 in Section 3.3 of [43].

3.6. Vanishing theorems. In this section we discuss vanishing theorems for

$$H_c^{\bullet}(Y_U, \mathbb{V}_{a,b}),$$

where we will now take $V_{a,b}$ to be a \mathcal{O} -linear local system, note that these are only well defined if $\ell > a+b+4$.

Theorem 3.6.1. (Theorem 10.1, Corollary 10.2 of [30]) Let Y_U be the Siegel threefold of neat level U and let $\mathbb{V} = \mathbb{V}_{a,b}$ be an automorphic local system of \mathcal{O} -modules. Assume that a > b > 0, that U_ℓ is hyperspecial and that $\ell > a + b + 4$. Then

$$H^3_!(Y_{U,\overline{\mathbb{O}}}, \mathbb{V}_{\mathcal{O}})$$

is a free O-module of finite rank and the natural map (with $\mathbb{F} = \mathcal{O}/\varpi$)

$$H^3_!(Y_{U,\overline{\mathbb{O}}}, \mathbb{V}_{\mathcal{O}}) \otimes_{\mathcal{O}} \mathbb{F} \to H^3_!(Y_{U,\overline{\mathbb{O}}}, \mathbb{V}_{\mathcal{O}} \otimes_{\mathcal{O}} \mathbb{F})$$

is surjective. Moreover

$$H_c^i(Y_{U,\overline{\mathbb{Q}}}, \mathbb{V}_{\mathcal{O}})) = 0 \text{ for } i > 3$$

$$H^i(Y_{U\overline{\mathbb{O}}}, \mathbb{V}_{\mathcal{O}})) = 0 \text{ for } i < 3.$$

Choosing ℓ sufficiently large and using comparison results between singular cohomology and étale cohomology we get results for local systems with coefficients in L, without restrictions on ℓ .

Corollary 3.6.2. If we now use $V_{a,b}$ to denote an L-linear local system of weight a > b > 0 then we have

$$H_c^i(Y_{U\overline{\mathbb{O}}}, \mathbb{V}_{a,b}) = 0 \text{ for } i > 3$$

$$H^i(Y_{U,\overline{\mathbb{Q}}}, \mathbb{V}_{a,b}) = 0 \text{ for } i < 3.$$

Remark 3.6.3. Theorem 10.1 of op. cit. has the assumptions that $\mu = (a, b, c) \in X_{\mathrm{GSp}_4}^{++,<\ell}$, that $|\mu|_{\mathrm{re},+} < \ell$ and that $|\mu|'_{\mathrm{comp}} \le \ell - 2$. Working through the definitions we find the first condition is equivalent to a > b > 0 and $\ell \ge a + b + 5$, the second condition is equivalent to $\ell \ge a + b + 2$ and the third condition is equivalent to $\ell \ge a + b + 4$. We conclude that the conditions a > b > 0 and $\ell > a + b + 4$ imply all the hypotheses of Theorem 10.1 of [30] in the case of Siegel modular threefolds.

4. Algebraic modular forms

In this section we will discuss algebraic modular forms on the group $G = GU_2(D)$. The general reference for this section is [16], we have also drawn from chapter 3 of [11].

Let G/\mathbb{Q} be a connected reductive group such that $G(\mathbb{R})$ is connected and compact modulo centre, in this paper we will use $G = D^{\times}$ and $G = GU_2(D)$ with D/\mathbb{Q} a definite quaternion algebra. Let L/\mathbb{Q}_{ℓ} be a finite extension as before and let V be an irreducible algebraic L-linear representation of $G(\mathbb{Q}_{\ell})$. Let $U = U^{\ell}U_{\ell}$ be a compact open subgroup of $G(\mathbb{A}_f)$, then we define the L-vector space of algebraic modular forms of level U with coefficients in V

$$A_V^G[U] := \left\{ F: G(\mathbb{Q}) \setminus G(\mathbb{A}_f) / U^\ell \to V \mid F(g \cdot u_\ell) = u_\ell^{-1} F(g) \text{ for } u_\ell \in U_\ell \right\}.$$

This vector space is finite dimensional because the double coset spaces

$$G(\mathbb{Q}) \setminus G(\mathbb{A}_f)/U$$

are finite (Proposition 4.3 of [16]). The following result is well known and can be proven as in [17, Section 2] or as in [9, Section 2]:

Proposition 4.0.1. Choose an embedding $L \hookrightarrow \mathbb{C}$, then there is a Hecke-equivariant isomorphism

$$A_V^G[U] \otimes_E \mathbb{C} \cong \bigoplus_{\sigma} m(\sigma) \sigma^U$$

where σ runs over cuspidal automorphic representations of G such that $\sigma_{\infty} \cong V \otimes_E \mathbb{C}$ and $m(\sigma)$ is the multiplicity with which σ occurs in the discrete spectrum of G.

4.1. The inner form. Let D be the definite quaternion algebra over \mathbb{Q} ramified at a single prime p with canonical involution $d \mapsto \overline{d}$ and consider the algebraic group G/\mathbb{Q} whose R points are given by

$$G(R) := \{(g, \lambda) \in \operatorname{GL}_2(D \otimes_{\mathbb{Q}} R) \times R^{\times} \mid {}^t \overline{g} g = \lambda I \},$$

where \overline{g} is entry-wise application of $d \mapsto \overline{d}$ and I is the identity matrix. One can show that $G(\mathbb{R})/Z(G(\mathbb{R})) \cong \mathrm{USp}(4)/\{\pm I\}$ where $\mathrm{USp}(4)$ is the compact \mathbb{R} -form of Sp_4 and that for primes $q \neq p$ there are isomorphisms $G(\mathbb{Q}_q) \cong \mathrm{GSp}_4(\mathbb{Q}_q)$. Both of these results are worked out in detail in Section 3.5 of [11]. The statement of Conjecture 1 concerns algebraic modular forms of weight k, j and level $K_2(p)$ and we will discuss these now.

4.2. Weights. For $\ell \neq p$ choose an isomorphism $G_{\mathbb{Q}_{\ell}} \simeq \mathrm{GSp}_{4,\mathbb{Q}_{\ell}}$. For every irreducible representation V of GSp_4 we get an irreducible representation V of G defined over \mathbb{Q}_{ℓ} , as before these are parametrised parametrized by integers $(a \geq b \geq 0, c)$, just like the local systems on Siegel threefolds. Siegel modular forms of weight k, j correspond to the local system of weight k+j-3, j-3, k+2j-6 and we define

$$\mathcal{A}_{k,j}^G[U] := \mathcal{A}_V^G[U],$$

where V denotes the irreducible representation of highest weight (k+j-3, j-3, k+2j-6). We will later see that cohomology of $\mathbb{V}_{a,b}$ with support in the supersingular locus of the Siegel modular threefold can be identified with a space of algebraic modular forms for G of weight a-b, b+3.

4.3. **Levels.** Our level subgroups $U \subset G(\mathbb{A}_f)$ will be of the form U^pU_p . The level U^p will come from GSp_4 under the isomorphism $G(\mathbb{A}_f^p) \cong \mathrm{GSp}_4(\mathbb{A}_f^p)$, the level U_p will be a maximal compact subgroup of $G(\mathbb{Q}_p)$ which we will describe now. Let B be the non-split quaternion algebra over \mathbb{Q}_p with standard involution $b \mapsto \overline{b}$ and uniformiser ϖ_B , then $G(\mathbb{Q}_p)$ can also be described as (because all Hermitian forms on $B^{\oplus 2}$ are equivalent)

$$G(\mathbb{Q}_p) = \left\{ (g, \lambda) \in \operatorname{GL}_2(B) \times \mathbb{Q}_p^{\times} : \overline{g} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} g = \lambda \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

We will be interested in the maximal compact subgroup

$$K_2 = \left\{ (g, \lambda) \in G(\mathbb{Q}_p) : g \in \left(\begin{smallmatrix} \mathcal{O}_B & \varpi_B \mathcal{O}_B \\ \varpi_B^{-1} \mathcal{O}_B & \mathcal{O}_B \end{smallmatrix} \right), \ \lambda \in \mathbb{Z}_p^{\times} \right\},$$

which is the stabiliser $\mathcal{O}_B \oplus \varpi_B \mathcal{O}_B$. Now let N be a squarefree integer with $p \mid N$, then we define

$$K_2(N) := K_2(p) \times \prod_{v \mid \frac{N}{p}} K(v) \prod_{v \nmid N} \mathrm{GSp}_4(\mathbb{Z}_v).$$

4.4. **Theta lifts.** In section 3 of [20] Ibukiyama defines a notion of ' θ -lift' and suggests that these are precisely the algebraic modular forms that do *not* correspond to holomorphic Siegel modular forms. He predicts (Conjecture 5.3 of [21]) that the θ -lifting gives injective maps to $\mathcal{A}_{k,j}^G[K_2(p)]$ from

(1)
$$\begin{cases} S_{2j-2+k}[\Gamma_0(1)] \times S_{k+2}[\Gamma_0(p)]^{\text{new}} & \text{if } k > 0 \text{ or } j \text{ odd} \\ S_{2j-2}[\Gamma_0(1)] \times M_2[\Gamma_0(p)] & \text{if } k = 0 \text{ and } j \text{ even} \end{cases}.$$

We remark that dim $M_2[\Gamma_0(p)] = 1 + \dim S_2[\Gamma_0(p)]$ so that Ibukiyama's formula essentially predicts a compact version of the Saito-Kurokawa lifting

$$S_{2j-2}[\Gamma_0(1)] \to \mathcal{A}_{0,j}^G[K_2(N)]$$

for odd j. We will actually prove that there are injective lifts from (1) to cuspidal automorphic representations of G, but it is unclear to us that these are θ -lifts in the sense of Ibukiyama (which is why we do not prove Conjecture 5.1 of [21] but something slightly different). It will follow from our arguments that most of these lifts (all if k > 0, j > 3) will be weakly endoscopic in the following sense:

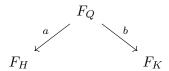
Definition 4.4.1. A cuspidal automorphic representation π of G is called weakly endoscopic if there is a cuspidal automorphic representation $\pi_1 \boxtimes \pi_2$ of $GL_2(\mathbb{A}) \times GL_2(\mathbb{A})$ such that the Satake parameters of π and $\pi_1 \boxtimes \pi_2$ agree for almost all primes via the inclusion of dual groups $GL_2 \times_{\mathbb{G}_m} GL_2 \to GSp_4$.

5. Integral models of Siegel modular threefolds

In this section we will discuss Siegel threefolds with various parahoric level structures at a fixed prime p, define integral models and describe their singularities, following [54] and [62]. We will then study the natural maps between these models and compute the fibers in characteristic p. The supersingular locus of the Siegel threefold with good reduction is well understood by classical work of Katsura-Oort [26] (c.f. [27]). Combining this description with our understanding of the fibers of the natural maps allows us to describe the supersingular loci of the other Siegel threefolds (c.f. [61]).

- 5.1. **Moduli functors.** Let $U^p \subset \mathrm{GSp}_4(\mathbb{A}^{p,\infty})$ be a fixed compact open subgroup away from p. Following section 2 of [54], we define moduli problems for three different parahorics levels at p, which are defined over \mathbb{Z}_p .
 - The moduli functor F_H of hyperspecial level H which parametrizes prime-to-p isogeny classes of triples (A, λ, η) where A is an abelian scheme of relative dimension two, λ is a prime-to-p polarisation and η is a U^p level structure (c.f. [28, Definition 1.3.7.1.]).
 - The moduli functor F_K of paramodular level K which parametrizes prime-to-p isogeny classes of triples (A, λ, η) where A is an abelian scheme of relative dimension two, λ is a polarisation such that $\ker \lambda \subset A[p]$ has rank p^2 , and η is a U^p level structure.
 - The moduli functor F_Q of Klingen level Q which parametrizes prime-to-p isogeny classes of quadruples (A, λ, η, H) where A is an abelian scheme of relative dimension two, λ is a prime-to-p polarisation, η is a U^p level structure and $H \subset A[p]$ is a finite locally free group scheme of rank p.

There are natural maps



where a is the forgetful map and b is the map that takes A to A/H with the induced polarisation and level structure. To elaborate, a sub-group scheme $H \subset A[p]$ of rank p is automatically isotropic for the Weil pairing induced by λ and hence the polarisation $p \cdot \lambda$ descends uniquely to a polarisation on A/H (c.f. Proposition 11.25 of [36]). The following result is standard (c.f. [28, Theorem 1.4.1.11])

Proposition 5.1.1. If U is neat (c.f. [28, Definition 1.4.1.8]), then the functors F_{\star} are representable by quasi-projective \mathbb{Z}_p schemes A_H , A_Q , A_K . The first one is smooth while the latter two have smooth generic fibres. Moreover the maps a and b are proper, finite étale on the generic fibre and induce bijections on geometrically connected components.

5.2. **Singularities.** In this section we will describe the bad reduction of A_K and A_Q following [54] and [62].

Proposition 5.2.1. (Theorem 3 of [54]) The geometric special fiber of the Siegel modular threefold of Klingen level decomposes as $A_{Q,\overline{s}} = X^e \cup X^m$, where X^e, X^m are smooth and intersect in a smooth surface S. Both X^e and X^m have a unique connected component lying over each connected component of $A_{H,\overline{s}}$. The locus $X^e \setminus X^m$ parametrizes those quadruples where H is étale, the locus $X^m \setminus X^e$ parametrizes those quadruples where H is multiplicative and S parametrizes the quadruples where H is étale locally isomorphic to α_p .

Proposition 5.2.2. (Theorem 1.3 of [62]) The scheme A_K/\mathbb{Z}_p is regular, and it is smooth over \mathbb{Z}_p away from a finite set of closed points Σ in the special fiber. At the singular points in the geometric special fibre $A_{K,\overline{s}}$ the completed local ring is isomorphic to

$$\overline{\mathbb{F}}_p[\![X,Y,Z,W]\!]/(XY-ZW).$$

The singular points Σ correspond precisely to the polarized abelian surfaces (A, λ) where $\ker \lambda$ is etale locally isomorphic to $\alpha_p \times \alpha_p$. Moreover for every connected component of $A_{K,\overline{\eta}}$ there is a unique connected component of $A_{K,\overline{s}}$.

5.3. **Geometry of Hecke correspondences.** The following two lemmata can be proven using the moduli descriptions and Section 4 of [61], see also Section 6.3 of [47].

Lemma 5.3.1. Let (A, λ, η) be an $\overline{\mathbb{F}}_p$ point of $A_{H,\overline{s}}$, let $a^{-1}(A, \lambda, \eta)$ denote the scheme theoretic fiber. The underlying reduced scheme of $a^{-1}(A, \lambda, \eta)$ has: cardinality 2(p+1) if A is ordinary, cardinality 3 if A has p-rank one, cardinality 1 when A is supersingular but not superspecial. If A is superspecial then $a^{-1}(A, \lambda, \eta)$ is isomorphic to \mathbb{P}^1 .

Lemma 5.3.2. Let (A, λ, η) be an $\overline{\mathbb{F}}_p$ point of $A_{K,\overline{s}}$ and let $b^{-1}(A, \lambda, \eta)$ denote the scheme-theoretic fiber. We will describe the underlying reduced scheme of $b^{-1}(A, \lambda, \eta)$ for various (A, λ, η) . It consists of two points when A has p-rank two, one point when A has p-rank one and $\ker \lambda$ is étale-multiplicative, two points when A has p-rank one and $\ker \lambda$ is local-local, and it consists of a single point when A is supersingular but $(A, \lambda, \eta) \notin \Sigma$ (recall that Σ is the singular locus of $A_{K,\overline{s}}$). For $(A, \lambda, \eta) \in \Sigma$ the fiber $b^{-1}(A, \lambda, \eta)$ is isomorphic to \mathbb{P}^1 .

Proof. Let (A, λ, η) as in the lemma. We want to find all principally polarised abelian varieties (B, μ) with a degree p isogenies $\phi : B \to A$ such that $\ker \phi \subset B[p]$ has degree p and such that the following diagram commutes

$$\begin{array}{ccc}
A & \xrightarrow{\lambda} & A^t \\
\phi \uparrow & & \downarrow \phi^t \\
B & \xrightarrow{p\mu} & B^t.
\end{array}$$

There is a unique isogeny $\lambda^t: A^t \to A$ such that $\lambda^t \circ \lambda = [p]$ and the choice of (B, μ) is equivalent to the choice of degree p subgroup of ker λ^t . The lemma now follows from a case-by-case analysis; for the sake of brevity we only highlight two cases (c.f. Theorem 4.7 of [61] for the supersingular cases).

- If A has p-rank one and ker λ is local-local then ker $\lambda^t \simeq \mu_p \times Z/p\mathbb{Z}$ and there are two degree p subgroups.
- If A has p-rank one and ker λ is étale-multiplicative then ker $\lambda^t \simeq E[p]$ (the p-torsion of a supersingular elliptic curve), and so there is only one degree p subgroup.

5.4. Supersingular loci. In this section we will give a description of the supersingular loci of $A_{Q,\overline{s}}$, $A_{H,\overline{s}}$ and $A_{K,\overline{s}}$, following [27] and [61]. It is a classical result due to Katsura-Oort [26] that all the irreducible components of the supersingular locus of the (coarse) moduli space of principally polarized abelian surfaces are isomorphic to \mathbb{P}^1 . In Chapter 4 of [27], it is shown that the supersingular locus of $A_{H,\overline{s}}$ is a union of projective lines and that the irreducible components are in bijection with

$$N = G(\mathbb{Q}) \setminus G(\mathbb{A}_f)/U^p K_2$$

equivariant from the prime-to-p Hecke operators. Our computation of the fibers of a now tells us that the irreducible components of the supersingular locus $A_{Q,\bar{s}}$ are parametrised by

$$N\coprod M$$
,

where M is the set of superspecial points in $A_{H,\overline{s}}(\overline{\mathbb{F}}_p)$. We will later write

$$A_{Q,\overline{s}}^{\mathrm{ss}} = E \cup F$$

where E corresponds to N and F corresponds to M. In [61], Yu proves the following result, which will be important for us in the proof of our main theorem.

Proposition 5.4.1. (Theorem 1.2 of [61]) The irreducible components of E are pairwise disjoint, and so are the irreducible components of F. The map $b: A_{Q,\overline{s}} \to A_{K,\overline{s}}$ contracts the irreducible components of E to points, identifying the singular points of $A_{K,\overline{s}}$ with N.

Remark 5.4.2. The bijection $\Sigma \simeq N$ can also be proven directly, using the fact that all superspecial abelian surfaces are isomorphic; then Σ corresponds to choices of a polarisation λ and a level structure η on a fixed superspecial abelian surface A.

6. The Picard-Lefschetz formula

In this section we discuss the Picard-Lefschetz formula for A_K , following [6] Exposé XV and using results of [29]. Let $\ell \neq p$, let \mathcal{O} denote \mathbb{Z}_{ℓ} and let \mathbb{V} be an automorphic local system of \mathcal{O} -modules on A_K , and recall that there is a distinguished triangle on the special fiber $A_{K,\overline{s}}$

$$(2) \mathbb{V} \to R\Psi \mathbb{V} \to R\Phi \mathbb{V}$$

where $R\Psi\mathbb{V}$ is the complex of nearby cycles and $R\Psi\mathbb{V}$ is the complex of vanishing cycles. It follows from Corollary 4.6 of [29] that

(3)
$$H^{\bullet}(X_{\overline{s}}, R\Psi \mathbb{V}) = H^{\bullet}(X_{\overline{\eta}}, \mathbb{V})$$
$$H^{\bullet}_{c}(X_{\overline{s}}, R\Psi \mathbb{V}) = H^{\bullet}_{c}(X_{\overline{\eta}}, \mathbb{V}).$$

6.1. The exact sequence. Proposition 3.1.2. of Exposé XV of [6] then tells us that the sheaves $R^i\Phi\mathbb{V}$ are zero for $i\neq 3$ and supported on the singular locus Σ for i=3. This means in particular that

$$R^{3}\Phi \mathbb{V} = \bigoplus_{x \in \Sigma} i_{x,*} \left(R^{3}\Phi \mathbb{V} \right)_{x}$$

and so the cohomology of $R^3\Phi\mathbb{V}$ is concentrated in degree zero where we find

How the conomology of
$$H$$
 we is concentrated in degree zero where we find
$$H^3_c(A_{K,\overline{s}},R\Phi\mathbb{V})=H^3(A_{K,\overline{s}},R\Phi\mathbb{V})=H^0_c(A_{K,\overline{s}},R^3\Phi\mathbb{V})=H^0(A_{K,\overline{s}},R^3\Phi\mathbb{V})=\bigoplus_{x\in\Sigma}\left(R^3\Phi\mathbb{V}\right)_x.$$

It follows from (3) and (2)) that

$$\begin{split} &H^i_c(A_{K,\overline{s}},\mathbb{V}) = H^i_c(A_{K,\overline{\eta}},\mathbb{V}) \\ &H^i(A_{K,\overline{s}},\mathbb{V}) = H^i(A_{K,\overline{\eta}},\mathbb{V}) \end{split}$$

for $i \neq 3, 4$ and also that we get the following diagram of exact sequences (4)

$$0 \longrightarrow H^3_c(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^3_c(A_{K,\overline{\eta}}, \mathbb{V}) \stackrel{\alpha}{\longrightarrow} H^3_c(A_{K,\overline{s}}, R\Phi\mathbb{V}) \longrightarrow H^4_c(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^4_c(A_{K,\overline{\eta}}, \mathbb{V}) \longrightarrow 0$$

$$\downarrow \qquad \qquad \qquad \downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow$$

$$0 \longrightarrow H^3(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^3(A_{K,\overline{\eta}}, \mathbb{V}) \stackrel{\alpha}{\longrightarrow} H^3(A_{K,\overline{s}}, R\Phi\mathbb{V}) \longrightarrow H^4(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^4(A_{K,\overline{\eta}}, \mathbb{V}) \longrightarrow 0.$$

6.2. The action of inertia. Let $\mathbb{Z}_{\ell}(1)$ with generator σ be the maximal pro- ℓ quotient of the inertia group I_p of $G_{\mathbb{Q}_p}$, then the general theory tells us that the action of the inertia group factors through $\mathbb{Z}_{\ell}(1)$ and moreover can be computed as follows (c.f. page 33 of Exposé XIII of [6], note that the variation map goes from ordinary cohomology to compactly supported cohomology)

$$H^{3}(A_{K,\overline{\eta}}, \mathbb{V}) \xrightarrow{\alpha} \bigoplus_{x \in \Sigma} \left(R^{3}\Phi\mathbb{V}\right)_{x}$$

$$\downarrow^{\sigma-1} \qquad \qquad \downarrow^{\oplus \operatorname{Var}_{x}(\sigma)}$$

$$H^{3}_{c}(A_{K,\overline{\eta}}, \mathbb{V}) \xleftarrow{\beta} \bigoplus_{x \in \Sigma} H^{3}_{x}(A_{K,\overline{s}}, R\Psi\mathbb{V})$$

Here $\operatorname{Var}_x(\sigma)$ is the local variation map, which is an isomorphism by 3.3.5 of Exposé XV of [6], and $H_x(A_{K,\bar{s}}, -)$ denotes cohomology with support in the closed subscheme $\{x\}$. The map $\sigma - 1$ is

a priori not equivariant for the action of Frobenius, but we can fix this by defining a monodromy operator N as the logarithm of σ . Since $(\sigma - 1)^2 = 0$ as an endomorphism of $H^3_!(A_{K,\overline{\eta}}, \mathbb{V})$, we have $N = \sigma - 1$ and so the following diagram is Frob_p -equivariant:

$$H_{!}^{3}(A_{K,\overline{\eta}}, \mathbb{V})(1) \xrightarrow{\alpha} \bigoplus_{x \in \Sigma} (R^{3}\Phi\mathbb{V})_{x}(1)$$

$$\downarrow^{N} \qquad \qquad \downarrow^{\oplus N_{x}}$$

$$H_{!}^{3}(A_{K,\overline{\eta}}, \mathbb{V}) \leftarrow_{\beta} \bigoplus_{x \in \Sigma} H_{x}^{3}(A_{K,\overline{s}}, R\Psi\mathbb{V}).$$

Additionally, the general theory tells us that β is the Poincare dual of α up to a Tate twist, since \mathbb{V} is self dual up to a Tate twist. Furthermore, the natural map (induced by (2))

$$H_x^3(A_{K,\overline{s}}, \mathbb{V}) \to H_x^3(A_{K,\overline{s}}, R\Psi\mathbb{V})$$

is an isomorphism by 2.2.5.8 of Exposé XV of [6]. Taking the direct sum over $x \in \Sigma$ we get an isomorphism

$$H^3_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \cong H^3_x(A_{K,\overline{s}}, R\Psi\mathbb{V})$$

and under this identification β factors as

$$H^{3}_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^{3}_{c}(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^{3}_{c}(A_{K,\overline{\eta}}, \mathbb{V})$$

$$\downarrow \qquad \qquad \downarrow \qquad \qquad \downarrow$$

$$H^{3}(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^{3}(A_{K,\overline{\eta}}, \mathbb{V})$$

$$\beta$$

since cohomology with support in a proper subscheme factors through compactly supported cohomology.

6.3. The weight filtration. Now let \mathbb{V} be an automorphic local system of $L = \mathbb{Q}_{\ell}$ -vector spaces of pure weight k. We will show that the weight filtration on $H^3_!(A_{K,\overline{\eta}},\mathbb{V})$ has nonzero graded pieces only of weights k+2,k+3,k+4 by explicitly writing down the weight filtration. Here the weight filtration is defined as usual in terms of the action of Frob_p on $H^3_!(A_{K,\overline{\eta}},\mathbb{V})$.

Lemma 6.3.1. The weight filtration on $H^3(A_{K,\overline{\eta}}, \mathbb{V})$ is given by

$$\operatorname{Im} \beta \subset H^3_!(A_{K,\overline{s}}, \mathbb{V}) \subset H^3_!(A_{K,\overline{\eta}}, \mathbb{V}).$$

Proof. The quotient $H^3_!(A_{K,\overline{\eta}}, \mathbb{V})/H^3_!(A_{K,\overline{s}}, \mathbb{V})$ is a subspace of $H^3(A_{K,\overline{s}}, R\Phi\mathbb{V})$ by (4). The derived projection formula tells us that (with tensor products derived)

$$R\Psi \mathbb{V} = \mathbb{V} \otimes_L R\Psi L$$
,

where L/\mathbb{Q}_{ℓ} is our coefficient field. This implies that $R\Phi\mathbb{V} = \mathbb{V} \otimes_L R\Phi L$ and the latter is given by

$$R\Phi L \cong \bigoplus_{x \in \Sigma} i_{x,*} L(-2)[-3],$$

by the discussion preceding 2.2.5.7 of Exposé XV of [6](m=1 in our case). This shows that

$$R\Phi \mathbb{V} = \bigoplus_{x \in \Sigma} i_{x,*} \mathbb{V}_x(-2)[-3],$$

and so $H^3(A_{K,\overline{s}}, R\Phi \mathbb{V}) = \bigoplus_{x \in \Sigma} \mathbb{V}_x(-2)$ is pure of weight k+4. We deduce from this that Im β is pure of weight k+2 using the isomorphism

$$H^3(A_{K,\overline{s}}, R\Phi \mathbb{V})(1) \cong H^3_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}),$$

coming from the local monodromy maps. We are left to show that $H_!^3(A_{K,\overline{s}}, \mathbb{V})/\operatorname{Im}\beta$ is pure of weight k+3. We know it has weight at most k+3 by Theorem 3.3.1 of [7], so it suffices to show the weight is at least k+3. Let $U=A_{K,\overline{s}}\setminus\Sigma$ be the smooth locus of $A_{K,\overline{s}}$, then there is a long exact sequence

$$\cdots \longrightarrow H^3_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \stackrel{\beta}{\longrightarrow} H^3(A_{K,\overline{s}}, \mathbb{V}) \longrightarrow H^3(U, \mathbb{V}) \longrightarrow \cdots$$

This implies that $H^3_!(A_{K,\overline{s}}, \mathbb{V})/\operatorname{Im}\beta$ is contained in $H^3(U, \mathbb{V})$, which has weight at least k+3 by smoothness of U.

6.4. The cokernel of monodromy. Now let $\mathbb{V} = \mathbb{V}_{\mathcal{O}}$ be an automorphic local system of \mathcal{O} -modules, then we can compute the monodromy operator using the following diagram :

(5)
$$H_{!}^{3}(A_{K,\overline{\eta}}, \mathbb{V})(1) \xrightarrow{\alpha} \operatorname{Im} \alpha \longleftrightarrow \bigoplus_{x \in \Sigma} \left(R^{3}\Phi\mathbb{V}\right)(1)_{x} \\ \downarrow^{N} \qquad \qquad \downarrow^{\gamma} \qquad \qquad \downarrow_{\oplus N_{x}} \\ H_{!}^{3}(A_{K,\overline{\eta}}, \mathbb{V}) \longleftrightarrow \operatorname{Im} \beta \overset{\alpha}{\longleftarrow} \bigoplus_{x \in \Sigma} H_{x}^{3}(A_{K,\overline{s}}, R\Psi\mathbb{V}).$$

Definition 6.4.1. Let Θ be the cokernel of γ (which depends on the weight, the level and ℓ), we call it the component group.

If we invert ℓ , then Im α is the top graded piece of the monodromy filtration and Im β is the bottom graded piece. The weight monodromy conjecture predicts that γ is an isomorphism after inverting ℓ , i.e., that Θ is torsion. We call Θ the component group because it is a direct generalisation of the component group of the Jacobian of a modular curve, in its cohomological incarnation (c.f. [38]). We will show that $\Theta = 0$ under the assumptions of Theorem 3.6.1. The same proof will show that γ is an isomorphism after inverting ℓ , under the assumption that the weight is sufficiently regular (now using Corollary 3.6.2). Showing that γ is an isomorphism after inverting ℓ in singular weight requires automorphic input.

7. The component group vanishes

7.1. Statement of the main result. In this section we will prove that the component group vanishes under the assumptions of Theorem 3.6.1, it will be useful to give these assumptions a label.

Assumption 7.1.1. Assume that a > b > 0, that U_{ℓ} is hyperspecial and that $\ell > a + b + 4$.

Proposition 7.1.2. If Assumption 7.1.1 holds then

$$\Theta = 0$$

The strategy of the proof is to show that the map

(6)
$$\alpha: H_c^3(A_{K,\overline{\eta}}, \mathbb{V}) \to H_c^3(A_{K,\overline{s}}, R\Phi \mathbb{V})$$

is surjective, which then shows that β is injective by duality. Proving this implies the same statement on inner cohomology, because the map from compactly supported cohomology to ordinary cohomology is self dual and because α factors through inner cohomology. The cokernel of (6) is given by the image of the map

(7)
$$H_c^3(A_{K,\overline{s}}, R\Phi \mathbb{V}) \to H_c^4(A_{K,\overline{s}}, \mathbb{V}),$$

which has a geometric interpretation in terms of the cycle classes of irreducible components of the supersingular locus of $A_{H,\bar{s}}$. To be precise we will identify the image of (7) with the image of the map

$$H_Z^4(A_{H,\overline{s}}, \mathbb{V}) \to H_c^4(A_{H,\overline{s}}, \mathbb{V}),$$

where Z is the supersingular locus of $A_{H,\bar{s}}$. Now since $A_{H,\bar{s}}$ is smooth we know that

$$H_c^4(A_{H,\overline{s}}, \mathbb{V}) = H_c^4(A_{H,\overline{\eta}}, \mathbb{V})$$

and we can now apply our vanishing theorems to the latter. The proof proceeds in two steps, which we formulate as two lemmas: We first relate the image of (7) to the cycle class of the supersingular locus on A_Q . More precisely we will relate it to the image of the cycle class map $H^0(E, \mathbb{V})(-2) \to H^4(X^e, \mathbb{V})$, where we recall that $A_{Q,\bar{s}}^{ss} = E \cup F \subset X^e$. Since E and X^e are smooth we can instead rewrite this using cohomology with support in E, which will appear in the following lemma.

Lemma 7.1.3. There is an isomorphism making the following diagram commute

$$H_{E}^{4}(X^{e}, \mathbb{V}) \longrightarrow H_{c}^{4}(X^{e}, \mathbb{V})$$

$$\uparrow \qquad \qquad \qquad \downarrow j \uparrow$$

$$H^{3}(A_{K,\overline{s}}, R\Phi \mathbb{V}) \longrightarrow H_{c}^{4}(A_{K,\overline{s}}, \mathbb{V})$$

and moreover the natural map j is an isomorphism.

Next we relate this to the cycle class map for the supersingular locus Z of $A_{H,\overline{s}}$. Because $a_1^{-1}(Z) = E \cup F$ (where $a_1 : X^e \to A_{H,\overline{s}}$ is the map induced by a) there is no natural map $H^4_Z(A_{H,\overline{s}}, \mathbb{V}) \to H^4_E(X^e, \mathbb{V})$, however we can produce one as follows:

Lemma 7.1.4. The natural map $H_Z^4(A_{H,\overline{s}}, \mathbb{V}) \to H_{E \cup F}^4(X^e, \mathbb{V})$ factors through $H_E^4(X^e, \mathbb{V})$ via an isomorphism. To be precise, there is an isomorphism making the following diagram commute:

Here Z is the supersingular locus of $A_{H,\overline{s}}$ and $E \cup F$ is the supersingular locus of $A_{Q,\overline{s}}$.

Proof of Proposition 7.1.2. Lemma 7.1.3 and Lemma 7.1.4 give us a commutative diagram

$$H^{3}(A_{K,\overline{s}}, R\Phi \mathbb{V}) \longrightarrow H^{4}_{c}(A_{K,\overline{s}}, \mathbb{V})$$

$$\downarrow \sim \qquad \qquad \downarrow \sim$$

$$H^{4}_{E}(X^{e}, \mathbb{V}) \longrightarrow H^{4}_{c}(X^{e}, \mathbb{V})$$

$$\uparrow \qquad \qquad \uparrow$$

$$H^{4}_{Z}(A_{H,\overline{s}}, \mathbb{V}) \longrightarrow H^{4}_{c}(A_{H,\overline{s}}, \mathbb{V})$$

To show that the top map is zero it suffices to show that the bottom map is zero. Now we note that

$$H_c^4(A_{H,\overline{s}}, \mathbb{V}) = H_c^4(A_{H,\overline{\eta}}, \mathbb{V}),$$

by smoothness of A_H . Moreover by Assumption 7.1.1 we can apply Theorem 3.6.1 to deduce that $H_c^4(A_{H,\overline{\eta}}, \mathbb{V}) = 0$. We have shown that the map ζ in the following exact sequence is zero

$$H^3_c(A_{K,\overline{\eta}},\mathbb{V}) \stackrel{\alpha}{\longrightarrow} H^3(A_{K,\overline{s}},R\Phi\mathbb{V}) \stackrel{\zeta}{\longrightarrow} H^4_c(A_{K,\overline{s}},\mathbb{V})$$

which shows that α is surjective and by Poincaré duality we conclude that β is injective. This shows that the component group is zero, because the local monodromy operators N_x in (5) are isomorphisms.

Corollary 7.1.5. If Assumption 7.1.1 holds then the map

$$\beta \otimes \mathbb{F} : H^3_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \otimes_{\mathcal{O}} \mathbb{F} \to H^3_{\Gamma}(A_{K,\overline{n}}, \mathbb{V}) \otimes_{\mathcal{O}} \mathbb{F}$$

is injective.

Proof. Theorem 3.6.1 says that

$$H^3_!(A_{K,\overline{\eta}}, \mathbb{V}) \otimes_{\mathcal{O}} \mathbb{F} \to H^3_!(A_{K,\overline{\eta}}, \mathbb{V} \otimes_{\mathcal{O}} \mathbb{F})$$

is surjective. Moreover since $R\Phi \mathbb{V}$ only has cohomology in degree 3 we find that

$$H^3(A_{K,\overline{s}}, R\Phi \mathbb{V}) \otimes_{\mathcal{O}} \mathbb{F} \to H^3(A_{K,\overline{s}}, R\Phi \mathbb{V} \otimes_{\mathcal{O}} \mathbb{F})$$

is an isomorphism. Now these natural maps fit into a commutative cube (which we won't draw) with the maps

$$\overline{\alpha}: H^3_!(A_{K,\overline{\eta}}, \mathbb{V} \otimes_{\mathcal{O}} \mathbb{F}) \to H^3(A_{K,\overline{s}}, R\Phi \mathbb{V} \otimes_{\mathcal{O}} \mathbb{F})$$
$$\overline{\beta}: H^3_{\Sigma}(A_{K,\overline{s}}, \mathbb{V} \otimes_{\mathcal{O}} \mathbb{F}) \to H^3_!(A_{K,\overline{\eta}}, \mathbb{V} \otimes_{\mathcal{O}} \mathbb{F}).$$

The fact that $\alpha \otimes \mathbb{F}$ is surjective implies that $\overline{\alpha}$ is surjective. By duality we find that $\overline{\beta}$ is injective which shows that $\beta \otimes \mathbb{F}$ is injective by chasing the cube.

Corollary 7.1.6. Let $\mathbb{V} = \mathbb{V}_{a,b}$ be a local system of L-vector spaces and suppose that the weight is sufficiently regular, i.e., that a > b > 0. Then the weight-monodromy conjecture holds for

$$\alpha: H^3_!(A_{K,\overline{\eta}}, \mathbb{V}),$$

in other words Theorem 1 holds.

Proof. The proof is the same as the proof of Proposition 7.1.2, except that we note that the vanishing of $H_c^4(A_{H,\overline{\eta}}, \mathbb{V})$ holds without restrictions on ℓ (this is Corollary 3.6.2).

7.2. **Proof of Lemma 7.1.3.**

Proof of Lemma 7.1.3. We apply the functors $R\Gamma_c$ and $R\Gamma_{\Sigma}$ to the triangle $\mathbb{V} \to R\Psi\mathbb{V} \to R\Phi\mathbb{V}$ to get a commutative diagram with exact rows:

The map h is zero by 2.2.5.8 of Exposé of [6] showing that g is injective. The group $H^4_{\Sigma}(A_{K,\overline{s}}, R\psi \mathbb{V})$ vanishes by the discussion preceding op. cit. 2.2.5.1, showing that g is an isomorphism. Since the sheaf $R\Phi\mathbb{V}$ is supported on Σ we conclude that f is an isomorphism using the long exact sequence

$$\cdots \to H^2(U, R\Phi \mathbb{V}) \to H^3_{\Sigma}(A_{K,\overline{s}}, R\Phi \mathbb{V}) \to H^3(A_{K,\overline{s}}, R\Phi \mathbb{V}) \to H^3(U, R\Phi \mathbb{V}) \to \cdots$$

Recall that $b_1: X^e \to A_{K,\overline{s}}$ satisfies $b_1^{-1}(\Sigma) = E$ by Proposition 5.4.1. This means that we get a morphism of cohomology with supports $H^4_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \to H^4_E(X^e, \mathbb{V})$ which fits into the following diagram:

$$H^{3}(A_{K,\overline{s}}, R\Phi \mathbb{V}) \xrightarrow{\cong} H^{3}_{\Sigma}(A_{K,\overline{s}}, R\Phi \mathbb{V}) \xrightarrow{\cong} H^{4}_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \xrightarrow{H^{4}_{c}(A_{K,\overline{s}}, \mathbb{V})} \downarrow j$$

$$H^{4}_{E}(X^{e}, \mathbb{V}) \xrightarrow{H^{4}_{c}(X^{e}, \mathbb{V})} H^{4}_{c}(X^{e}, \mathbb{V})$$

The map $\pi = b_1$ induces a map $\mathbb{V} \to R\pi_*\pi^*\mathbb{V}$ which we extend to a distinguished triangle

$$(8) \mathbb{V} \to R\pi_*\pi^*\mathbb{V} \to Q.$$

Recall that $b_1: X^e \to A_{K,\overline{s}}$ has fibers isomorphic to singletons over all points, except that over the singular points of $A_{K,\overline{s}}$ the fibers are isomorphic to \mathbb{P}^1 (Lemma 5.3.2). Using the proper base change theorem and the fact that the pullback of \mathbb{V} to $E_x \cong \mathbb{P}^1$ is constant and isomorphic to \mathbb{V}_x we find that

$$Q \cong \bigoplus_{x \in \Sigma} i_{x,*} \mathbb{V}_x(-1)[-2]$$

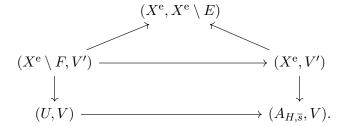
where $i_x:\{x\}\to X$ is the inclusion of the point x. This means in particular that $H^{\bullet}_{\Sigma}(A_{K,\overline{s}},Q)=H^{\bullet}(A_{K,\overline{s}},Q)$ since the restriction of Q to U is zero. If we now apply $R\Gamma_{\Sigma}$ to (8) we get a long exact sequence (using the fact that $R\Gamma_{\Sigma}R\pi_*\cong R\Gamma_{\pi^{-1}(\Sigma)}$)

$$\cdots \to H^i_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \to H^i_{E}(X^e, \mathbb{V}) \to H^i_{\Sigma}(A_{K,\overline{s}}, Q) \to \cdots$$

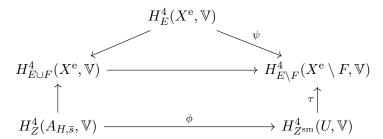
Because the cohomology of Q is concentrated in degree 2 we see that the map $H^4_{\Sigma}(A_{K,\overline{s}}, \mathbb{V}) \to H^4_E(X^e, \mathbb{V})$ is an isomorphism. By a similar long exact sequence argument (applying $R\Gamma$ to (8)) it follows that j is an isomorphism.

7.3. **Proof of Lemma 7.1.4.** The lemma will follow from the fact that $a_1: E \to Z$ induces a bijection on irreducible components and the fact that the sheaf $\mathbb V$ is constant on E and Z (because the components are projective lines). However, since there is no natural map $H_Z^4(A_{H,\overline{s}},\mathbb V) \to H_E^4(X^e,\mathbb V)$ we have to pass to the smooth locus of Z which makes the proof look more involved than it is.

Proof of Lemma 7.1.4. Let $V = A_{H,\overline{s}} \setminus Z$ be the complement of Z, let U be the complement of the singular locus of Z and let $V' = X^e \setminus (E \cup F)$. Then we have a commutative diagram of maps of pairs (recall that a(F) is precisely the set of singular points of Z)



There are induced (contravariant) maps in cohomology with support



and we are going to show that ϕ , ψ and τ are all isomorphisms, providing the required factorisation. The map ϕ fits into a long exact sequence for $A_{H,\overline{s}} \supset U \supset V$ (c.f. Chapter 23 of [34])

$$\cdots \to H^4_{Z^{\operatorname{sing}}}(A_{H,\overline{s}}, \mathbb{V}) \to H^4_Z(A_{H,\overline{s}}, \mathbb{V}) \to H^4_{Z^{\operatorname{sm}}}(U, \mathbb{V}) \to H^5_{Z^{\operatorname{sing}}}(A_{H,\overline{s}}, \mathbb{V}) \to \cdots$$

Since Z^{sing} is smooth of codimension 3 in $A_{H,\bar{s}}$, cohomological purity (Theorem 16.1 of [34]) tells us that

$$H_{Z^{\operatorname{sing}}}^{i}(A_{H,\overline{s}}, \mathbb{V}) = H^{i-6}(Z^{\operatorname{sing}}, \mathbb{V})(-3)$$

which is zero for i=4,5, so ϕ is an isomorphism. Now we note that both Z^{sm} and $E \setminus F$ are smooth of codimension 2 in U and $X^{\mathrm{e}} \setminus F$ respectively so we can use cohomological purity twice to produce a commutative diagram

$$\begin{array}{ccc} H^4_{E\backslash F}(X^{\mathrm{e}} \setminus F, \mathbb{V}) & \stackrel{\sim}{\longrightarrow} & H^0(E \setminus F, \mathbb{V})(-2) \\ & & & & & & & & \\ \tau & & & & & & & \\ f & & & & & & & & \\ H^4_{Z^{\mathrm{sm}}}(U, \mathbb{V}) & \stackrel{\sim}{\longrightarrow} & H^0(Z^{\mathrm{sm}}, \mathbb{V})(-2). \end{array}$$

The map σ is an isomorphism because the map $E \setminus F \to Z^{\mathrm{sm}}$ induces a bijection on irreducible components (Section 5.4) and the sheaf \mathbb{V} is constant on each of those components (because the

components are isomorphic to \mathbb{P}^1). Similarly, we get a commutative diagram

$$\begin{array}{ccc} H_E^4(X^e,\mathbb{V}) & \stackrel{\sim}{-\!\!\!\!-\!\!\!\!-\!\!\!\!-} & H^0(E,\mathbb{V})(-2) \\ & & & & & & \downarrow \xi \\ \\ H_{E\backslash F}^4(X^e \backslash F,\mathbb{V}) & \stackrel{\sim}{-\!\!\!\!-\!\!\!\!-} & H^0(E \backslash F,\mathbb{V})(-2), \end{array}$$

and ξ is an isomorphism because $(E \setminus F) \hookrightarrow E$ induces a bijection on irreducible components and \mathbb{V} is constant on E because $E \simeq \mathbb{P}^1$.

8. Proofs of the main results

8.1. Mazur's principle. Let us recall the theorem that we are trying to prove:

Theorem 3. Let π be a cuspidal automorphic representation of GSp_4 that is cohomological of weight a > b > 0 such that π_p is ramified and K(p)-spherical. Let $U = U^p \cdot K(p) \subset \mathrm{GSp}_4(\mathbb{A}_f)$ be a neat compact open subgroup such that $\pi^U \neq 0$. Let $\ell \neq p$ be a prime such that $\overline{\rho_{\pi,\ell}}$ is irreducible, such that the group U_ℓ is hyperspecial at ℓ and such that $\ell > a + b + 4$. Then if $\overline{\rho_{\pi,\ell}}$ is unramified at p and has four distinct Frob_p eigenvalues, there exists a cuspidal automorphic representation π' , of the same weight and level U^p away from p, such that $\overline{\rho_{\pi,\ell}} = \overline{\rho_{\pi',\ell}}$ and such that π'_p is unramified.

Proof. Assume that there are no congruences to unramified cuspidal automorphic representations π' , in other words, that the conclusion of the theorem does not hold. We will prove that $\overline{\rho_{\pi,\ell}}$ has at most three Frobenius eigenvalues, showing that the assumptions of the theorem are not satisfied. Let $\mathbb T$ be the Hecke algebra containing the unramified Hecke operators and the Hecke operators at p and let $\mathfrak m$ be the maximal ideal of the Hecke algebra corresponding to $\overline{\rho_{\pi,\ell}}$, then we have a decomposition

(9)
$$H^3_!(A_{K,\overline{\eta}}, \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} L = \bigoplus \rho_{\pi',\ell},$$

where $\rho_{\pi',\ell}$ is congruent modulo ℓ to $\rho_{\pi,\ell}$ and might appear multiple times in the direct sum. Note that Yoshida lifts and Saito-Kurokawa lifts don't contribute to the direct sum because the associated Galois representations are reducible and so cannot be congruent to $\rho_{\pi,\ell}$. This means that (9) is really a direct sum of irreducible four-dimensional Galois representations. By assumption all the π'_p over which the sum is indexed are ramified at p and so by weight-monodromy (Corollary 7.1.6) and local-global compatibility we know that π'_p is of type IIa. This means that the monodromy operator $N: \rho_{\pi',\ell}(1) \to \rho_{\pi',\ell}$ has a one-dimensional image, which shows that

$$\alpha(\rho_{\pi',\ell})$$

is a one-dimensional subspace of $H^3(A_{K,\overline{s}}, R\Phi \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} L$. Since α is moreover surjective we find that the dimension of $H^3(A_{K,\overline{s}}, R\Phi \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} L$ is equal to n, the number of $\rho_{\pi',\ell}$'s appearing in the direct sum (9).

By Lemma 2.2.3 we know that Frobenius acts on $\alpha(\rho_{\pi',\ell})$ by the scalar $p\lambda_{\pi'}$, where $\lambda_{\pi'}$ is the eigenvalue of the u-operator on $(\pi'_p)^{K(p)}$, introduced in Section 2.2.2. Localising at \mathfrak{m} means fixing a mod ℓ eigenvalue of u and so Frobenius acts with a single eigenvalue on

$$H^3(A_{K,\overline{s}}, R\Phi \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F}.$$

We also know that $H^3(A_{K,\overline{\eta}}, \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F}$ is isomorphic (up to semi-simplification) to

$$\overline{\rho_{\pi,\ell}}^{\oplus r}$$

for the same n as before (by vanishing of torsion). Since $\overline{\rho_{\pi,\ell}}$ has four distinct Frobenius eigenvalues, we know that the space $H^3_!(A_{K,\overline{\eta}}, \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F}$ decomposes into four generalised Frobenius eigenspaces, each of dimension n. Now $\alpha \otimes \mathbb{F}$ is surjective because α is surjective hence the kernel of $\alpha \otimes \mathbb{F}$ only contains three Frobenius eigenvalues. Because $\overline{\rho_{\pi,\ell}}$ is irreducible we can find a copy

$$\overline{\rho_{\pi,\ell}} \subset H^3_!(A_{K,\overline{\eta}}, \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F},$$

which is contained in the kernel of $N \otimes \mathbb{F}$ since $\overline{\rho_{\pi,\ell}}$ is unramified. As usual we compute the monodromy operator using the following diagram:

$$H^{3}_{!}(A_{K,\overline{\eta}}, \mathbb{V})(1)_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F} \xrightarrow{\alpha \otimes \mathbb{F}} H^{3}(A_{K,\overline{s}}, R\Phi \mathbb{V})(1)_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F}$$

$$\downarrow^{N \otimes \mathbb{F}} \qquad \qquad \downarrow$$

$$H^{3}_{!}(A_{K,\overline{\eta}}, \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F} \xleftarrow{\beta \otimes \mathbb{F}} H^{3}_{\Sigma}(A_{K,\overline{s}}, \mathbb{V})_{\mathfrak{m}} \otimes_{\mathcal{O}} \mathbb{F}$$

Corollary 7.1.5 tells us that $\beta \otimes \mathbb{F}$ is injective which means that $\ker(N \otimes \mathbb{F}) = \ker(\alpha \otimes \mathbb{F})$. Therefore our copy of $\overline{\rho_{\pi,\ell}}$ must be contained in the kernel of $\alpha \otimes \mathbb{F}$, but then $\overline{\rho_{\pi,\ell}}$ only has three distinct Frobenius eigenvalues, which gives a contradiction.

8.2. A geometric Jacquet-Langlands correspondence. Let us start by stating a precise version of Theorem 2:

- **Theorem 8.2.1.** (1) Let π be a cohomological cuspidal automorphic representation of GSp_4 that is not an irrelevant Yoshida lift and such that π_p is ramified and K(p)-spherical. Then there is a cuspidal automorphic representation σ of G such that $\pi_v \cong \sigma_v$ for finite places $v \neq p$, such that σ_p is $K_2(p)$ -spherical and with σ_∞ determined by π_∞ . Moreover, σ occurs with multiplicity one in the cuspidal spectrum of G. Conversely, a given cuspidal automorphic representation σ of G comes from such a π if σ_∞ has weight k > 0, j > 3.
 - (2) Let $k \geq 0, j \geq 3$ and let N be a squarefree integer such that $p \mid N$, then there is an injective map

(10)
$$S_{k,j}[K(N)]^{p-new} \hookrightarrow \mathcal{A}_{k,j}^G[K_2(N)],$$

equivariant for the prime-to-p Hecke operators, which proves Conjecture 1. For k > 0, j > 3 the image consists precisely of algebraic modular forms that are not weakly-endoscopic.

(3) Let $k \geq 0, j \geq 3$ and N = p be prime, then there is an injective lift

$$S_{2j-2+k}[\Gamma_0(1)] \times S_{k+2}[\Gamma_0(p)]^{new} \to \mathcal{A}_{k,j}^G[K_2(p)]$$

and for k = 0 an injective lift

$$S_{2j-2}[\Gamma_0(1)] \to \mathcal{A}_{0,j}^G[K_2(p))].$$

The image of (10) is a complementary subspace to the space generated by the images of these lifts (and the constant algebraic modular forms if k=0, j=3). Moreover for $k \geq 0, j \geq 3$ we have an equality

$$\dim \mathcal{A}_{k,j}^G[K_2(p)] = \dim S_{2j-2+k}[\Gamma_0(1)] \cdot \dim S_{k+2}[\Gamma_0(p)]^{new} + \dim S_{k,j}[\Gamma_0(p)] + \\ - 2\dim S_{k,j}[\Gamma_0(1)] + \delta_{k,0}\dim S_{2j-2}[\Gamma_0(1)] + \delta_{k,0} \cdot \delta_{j,0}.$$

Remark 8.2.2. The dimension formula in (iii) is proven by Ibukiyama for $j \neq 3, 4$ in [21] and for k = 0, j = 3 in [22].

The main idea of the proof is the incarnation of the singular locus Σ of $A_{K,\overline{s}}$ as a Shimura set for G. We then get a map α from $H^3_!(A_{K,\overline{\eta}},\mathbb{V})$, which we can explicitly describe in terms of automorphic forms and Galois representations, to the space of algebraic modular forms. Theorem 1 combined with local-global compatibility allows us to control the image of α in a Galois-theoretic way. For parts (2) and (3) the main ingredients are the multiplicity one result of [1, 14] and the main Theorem of [37], which computes the cohomology of $A_{H,\overline{\eta}}$ if $U^p = \mathrm{GSp}(\hat{\mathbb{Z}}^p)$. We start by relating spaces of algebraic modular forms to cohomology groups on $A_{K,\overline{s}}$, using our description of the singular locus:

Lemma 8.2.3. There is an isomorphism

$$H^0(\Sigma, \mathbb{V}_{a,b}) \cong \mathcal{A}_{a-b,b+3}^G[U^p K_2(p)],$$

equivariant for the prime-to-p Hecke operators.

Proof. This is standard, see the proof of Proposition 6.4 of [53].

Proof of Theorem 8.2.1 (1). We can identify the following cohomology groups

$$H^{3}(A_{K,\overline{s}}, R\Phi \mathbb{V}) = \bigoplus_{x \in \Sigma} (R^{3}\Phi \mathbb{V})_{x} = \bigoplus_{x \in \Sigma} \mathbb{V}_{x}(-2) = H^{0}(\Sigma, \mathbb{V})(-2),$$

and the key player of the proof will be the Hecke equivariant map

$$\alpha: H^3_!(A_{K,\overline{\eta}}, \mathbb{V}) \to H^3(A_{K,\overline{s}}, R\Phi \mathbb{V}) = A^G_{a-b,b+3}[U^p K_2(p)].$$

We can write the domain of α in terms of Hecke modules and Galois representations as follows (using the results of Section 3.2 and the multiplicity one result of [1,14])

$$H^3_!(A_{K,\overline{\eta}}, \mathbb{V}) = \bigoplus_{\pi} \pi^U_{\mathrm{fin}} \otimes \rho_{\pi,\ell}.$$

We are going to hide some things in the notation for bookkeeping purposes: If π is a Yoshida lift or CAP then by $\rho_{\pi,\ell}$ we mean the two-dimensional Galois representation that occurs in the cohomology. If π is of general type with π_{∞} holomorphic then $\rho_{\pi,\ell}$ is the four-dimensional Galois representation associated with π and if π_{∞} is generic then $\rho_{\pi,\ell} = 0$. The reason for doing this is that there is only one four-dimensional Galois representation $\rho_{\pi,\ell}$ for the two automorphic representations $\pi_{\text{fin}} \otimes \pi^W$ and $\pi_{\text{fin}} \otimes \pi^H$ as they both contribute a two-dimensional piece to cohomology.

Because α has an interpretation in terms of the action of inertia, we know that $\rho_{\pi,\ell}/(\ker \alpha \cap \rho_{\pi,\ell})$ is one dimensional for all ramified $\rho_{\pi,\ell}$ occurring in $H^3_!(A_{K,\overline{\eta}}, \mathbb{V})$. To be precise this follows from the conjunction of weight-monodromy and local-global compatibility (for weight-monodromy we need to use the results of Arthur to transfer generic cuspidal automorphic representations to GL_4 , where weight-monodromy holds by [4], c.f. Theorem 2.1.1.(2) of [3]). Therefore we will write

$$\operatorname{Im}\alpha\cong\bigoplus_{\pi}\pi_{\mathrm{f}}^{U}$$

where π runs over certain cuspidal automorphic representations of GSp_4 . Since $\pi_p^{K(p)}$ is one-dimensional we can rewrite this as

$$\operatorname{Im} \alpha \cong \bigoplus_{\pi} (\pi_{\mathrm{f}}^p)^{U^p}.$$

Similarly we can write

$$A_{a-b,b+3}^G[U^pK_2(p)] = \bigoplus_{\sigma} m(\sigma)\sigma_{\mathbf{f}}^{U'}$$

where σ runs over certain cuspidal automorphic representations of G and $U' = U^p K_2(p)$. Since $\sigma_p^{K_2(p)}$ is one-dimensional we can also write this as

$$A_{a-b,b+3}^G[U^pK_2(p)] = \bigoplus_{\sigma} m(\sigma)(\sigma_{\mathrm{f}}^p)^{U^p}.$$

Let \mathbb{T} be the Hecke-algebra away from p for both G and GSp_4 , which makes sense as soon as we fix an isomorphism $\mathrm{GSp}_4(\mathbb{A}_f^p) \cong G(\mathbb{A}_f^p)$. Then the spaces $(\pi_{\mathrm{f}}^p)^{U^p}$ and $(\sigma_f^p)^{U^p}$ are simple \mathbb{T} -modules and moreover the map α is equivariant for the action of \mathbb{T} .

Now fix a cohomological automorphic representation π of GSp_4 that is not an irrelevant Yoshida lift such that π_{∞} is in the discrete series and such that π_p is ramified and K(p)-spherical. Then we can choose U^p sufficiently small such that $\rho_{\pi,\ell}$ occurs in $H^3_!(A_{K,\overline{\eta}},\mathbb{V})$. Because π is not an irrelevant Yoshida lift we know that the summand

$$F_{\pi} := \pi_{\mathrm{f}}^U \subset \operatorname{Im} \alpha$$

is nonzero (because the Galois representation occurring in cohomology is ramified). There we see that

$$F_{\pi} \subset A_{a-b,b+3}^G[U^p K_2(p)] = \bigoplus_{\sigma} m(\sigma) \sigma_{\mathbf{f}}^{U'}$$

and so there is a cuspidal automorphic representation σ of G such that $(\sigma^p)^{U^p} \cong (\pi^p)^{U^p}$ which implies that $\sigma_v \cong \pi_v$ for all finite places $v \neq p$. Moreover σ_∞ is determined by π_∞ because it is determined by the weights a-b,b+3. When a>b>0 or equivalently k>0,j>3 then α is surjective by Corollary 7.1.6 so every cuspidal automorphic representation of G arises in this way. To prove multiplicity one for σ 'in the image' of this transfer we first prove a claim:

Claim 8.2.4. The summand $P_{\sigma} := m(\sigma)\sigma_f^{U'} \subset \mathcal{A}_{a-b,b+3}^G$ is in the image of α (we have only shown so far that Im $\alpha \cap P_{\sigma} \neq \emptyset$).

Proof of Claim. Let π be a cuspidal automorphic representation of GSp_4 that 'maps to σ '. If P_{σ} maps nontrivially to coker α then the Hecke eigenvalues of σ occur in $H^4_c(A_{H,\overline{\eta}}, \mathbb{V})$ and so the Hecke eigenvalues associated with π occur there. Therefore π has the same prime-to-p Hecke eigenvalues as an automorphic form π' whose Hecke eigenvalues occur in $H^4_c(A_{H,\overline{s}}, \mathbb{V})$. By the Chebotarev density theorem this means that the semi-simplification of $\rho_{\pi,\ell}$ is equal to the semi-simplification of $\rho_{\pi',\ell}$. But $\rho_{\pi,\ell}$ is already semi-simple and since $\rho_{\pi',\ell}$ is unramified at p (because π'_p is unramified), we conclude that $\rho_{\pi,\ell}$ is unramified at p, a contradiction.

From the claim we get cuspidal automorphic representations $\pi_1, \dots, \pi_{m(\sigma)}$ such that $\alpha(\rho_{\pi_i,\ell})$ maps to the summand P_{σ} . Therefore we have $\pi_{i,v} \cong \pi_{j,v}$ for all finite places $v \neq p$ and so all the π_i are in the same L-packet. When the L-packet is not CAP, then the $\pi_{i,p}$ are generic (by weight-monodromy) and this means that $\pi_{i,p} \cong \pi_{j,p}$ because local L-packets have unique generic constituents. When the L-packet is CAP, then the π_i are Saito-Kurokawa lifts and the fact that $\pi_{i,p}$ is K(p)-spherical also means that $\pi_{i,p} \cong \pi_{j,p}$ (c.f. Table 2 of [45]). We conclude that all the π_i are isomorphic and therefore by multiplicity one for GSp₄ we deduce that $m(\sigma) = 1$.

Remark 8.2.5. In the proof of parts (2),(3) of Theorem 8.2.1 we will work with the open compact subgroup K(N), which is not neat. This means that the moduli functors we defined are not representable and so we cannot, strictly speaking, take the étale cohomology of the 'Shimura variety of level K(N)'. To fix this we choose a neat compact open subgroup $U \subset K(N)$ that is normal in K(N) and then define (with similar definitions for compactly supported and inner cohomology)

$$H^{\bullet}(Y_{K(N),\overline{\mathbb{Q}}}, \mathbb{V}) := H^{\bullet}(Y_{U,\overline{\mathbb{Q}}}, \mathbb{V})^{H},$$

where H is the finite group H = K(N)/U. Note that because H is finite and \mathbb{V} is a local system of L-vector spaces, taking H invariants is exact and so all the exact sequences from the previous sections carry over to this setting. Furthermore, this definition does not depend on the choice of U. Moreover the corresponding space of algebraic modular forms $\mathcal{A}^G[U']$ satisfies

$$\mathcal{A}_{k,j}^G[U']^H = \mathcal{A}_{k,j}^G[K_2(N)]$$

for all k, j, which means that there is an induced map

$$\alpha: H^3_!(A_{K,\overline{\eta}}, \mathbb{V}) \to \mathcal{A}^G_{k,j}[K_2(N)].$$

Proof of Theorem 8.2.1(2). Write

$$H^3_!(A_{Q,\overline{\eta}}, \mathbb{V}) := \bigoplus_{\pi} \pi^U_{\text{fin}} \otimes \rho_{\pi,\ell},$$

with the notation as in the proof of part (1) of Theorem 8.2.1 (so $\rho_{\pi,\ell}$ is either two or four-dimensional). Let S be the subspace of $H^3_!(A_{K,\overline{\eta}},\mathbb{V})$ spanned by the summands $\pi^U_{\mathrm{fin}}\otimes\rho_{\pi,\ell}$ for π such that π_{∞} is holomorphic and such that π_p is ramified. Then $\alpha(S)$ has dimension equal to the dimension of $S_{k,j}[K(N)]^{p-\mathrm{new}}$ which gives us an injective map (after choosing a basis of normalised eigenforms)

(11)
$$S_{k,j}[K(N)]^{p-\text{new}} \to \mathcal{A}_{k,j}^G[K_2(N)] \otimes \mathbb{C}.$$

Since there are no holomorphic Yoshida lifts, it is clear that the algebraic modular forms in the image are not weakly endoscopic. If a > b > 0 then α is surjective by Corollary 7.1.6 and the image of (11) is complementary to the subspace of weakly endoscopic algebraic modular forms. Indeed, the only cohomological cuspidal automorphic representations of this level that don't come from $S_{k,i}[K(N)]$ are non-holomorphic Yoshida lifts.

Proof of Theorem 8.2.1(3). Now suppose that N=p, then Theorem 2.1 of [37] tells us that $H_c^4(A_{H,\overline{\mathbb{Q}}}, \mathbb{V}_{a,b})=0$ unless a=b is even. This means that α will be surjective unless a=b is even and so we can prove dimension formulas if we understand the dimension of the image of α .

Claim 8.2.6. The dimension of the image of α is equal to:

(12)
$$\dim \operatorname{Im} \alpha = \dim S_{2j-2+k}[\Gamma_0(1)] \times \dim S_{k+2}[\Gamma_0(p)]^{\text{new}}$$

$$+ \dim S_{k,j}[K(p)] - 2 \dim S_{k,j}[K(1)] + \delta_{k,0} \frac{1 + (-1)^j}{2} \dim S_{2j-2}[\Gamma_0(1)].$$

Proof of Claim 8.2.6. The cohomological cuspidal automorphic representations π with nonzero invariants under K(p) are either non-holomorphic Yoshida lifts or correspond to holomorphic Siegel cusp forms. We only care about holomorphic Siegel cusp forms that are new at p and similarly about Yoshida lifts with π_p ramified that are not irrelevant. Since $\pi^{K(p)}$ is one-dimensional for all these π it suffices to simply count them. The results of [40] tell us that every eigenform $f \in S_{k,j}[K(1)]$

produces two oldforms in $S_{k,j}[K(p)]$, which are distinct unless k=0, j is even and f is in the image of the (injective) Saito-Kurokawa lift $S_{2j-2}[\Gamma_0(1)] \to S_{0,j}[K(1)]$. This means that

$$\dim S_{k,j}[K(p)]^{\text{new}} = \dim S_{k,j}[K(p)] - 2\dim S_{k,j}[K(1)] + \delta_{k,0} \frac{1 + (-1)^j}{2} \dim S_{2j-2}[\Gamma_0(1)].$$

Section 3.5.5 tells us that the number of relevant Yoshida lifts is equal to dim $S_{2j-2+k}[\Gamma_0(1)] \times \dim S_{k+2}[\Gamma_0(p)]^{\text{new}}$ and the formula follows.

If a > b or if b is odd (equivalently k > 0 or $j \ge 3$ even) then α is surjective which proves that

$$\dim \mathcal{A}_{k,j}^G[K_2(p)] = \dim \operatorname{Im} \alpha = (\star),$$

where (\star) is given by (12). When a=b is even and b>0 (equivalently k=0 and $j\geq 3$ odd) then the dimension formula to be proven follows from Theorem 5.2 of [21]. The case a=b=0 (equivalently k=0, j=3) is proven by Ibukiyama in [22] and can also be deduced from the fact that $H_c^4(A_{H,\overline{\eta}}, \mathbb{Q}_\ell) = \mathbb{Q}_\ell(-2)$.

Now let us return to the case that a=b is even. Theorem 2.1 of [37] tells us that

$$H^4_c(A_{H,\overline{\mathbb{Q}}}, \mathbb{V}_{a,b}) = L(b-2)^{\oplus s_{a+b+4}}$$

where $s_{a+b+4} = \dim S_{a+b+4}[\Gamma_0(1)]$. Moreover, the Hecke eigenvalues occurring in $H_c^4(A_{H,\overline{\mathbb{Q}}}, \mathbb{V}_{a,b}) = H_c^4(A_{H,\overline{\mathbb{q}}}, \mathbb{V}_{a,b})$ are weakly equivalent to those of automorphic representations parabolically induced from the Siegel parabolic (they are conjecturally non-cuspidal). Theorem 5.2 of [21] combined with Claim 8.2.6 tells us the dimension of the cokernel of α is also equal to s_{a+b+4} , and so

$$H_Z^4(A_{H,\overline{\eta}}, \mathbb{V}_{a,b}) \to H_c^4(A_{H,\overline{\eta}}, \mathbb{V}_{a,b})$$

is surjective (recall that coker $\alpha = \operatorname{Im} \gamma$). In any case this means that there is an injective lift $S_{a+b+4}[\Gamma_0(1)] \to \mathcal{A}_{0,b+3}^G[K_2(p)]$ as claimed in the theorem. We conclude that the image of

$$S_{k,j}[K(p)]^{\text{new}} \to \mathcal{A}_{k,j}^G[K_2(p)]$$

is a complementary subspace to the subspace generated by the lift from

$$S_{2j-2+k}[\Gamma_0(1)] \times S_{k+2}[\Gamma_0(p)]^{\text{new}},$$

the lift from $S_{2j-2}[\Gamma_0(1)]$ and the constant algebraic modular forms (if k=0, j=3).

8.3. Acknowledgements. I am very grateful to James Newton for his guidance and for carefully reading preliminary versions of this work. I would like to thank Ana Caraiani for her encouragement and her helpful suggestions. I would also like to thank Andrew Graham and Dougal Davis for many useful discussions. Furthermore I would like to thank the anonymous referee for their helpful comments and suggestions.

APPENDIX A. ON THE WEIGHT-MONODROMY CONJECTURE

In this appendix we prove the following result (which is presumably well known to experts):

Theorem A.1. Let F/\mathbb{Q}_p be a finite extension, let X/F be a Shimura variety of Hodge type, let $j: X \hookrightarrow X^*$ be the inclusion of X into its minimal compactification and let \mathbb{V} be an automorphic local system on X. Then the weight-monodromy conjecture for

$$H^i(X_{\overline{F}}^*, j_{!*}\mathbb{V}),$$

where $j_{!*}$ denotes the intermediate extension of perverse sheaves (up to shift), follows from the weight-monodromy conjecture for smooth and proper varieties over F

Proof of Theorem A.1. Let $\pi: A \to X$ be the universal family of abelian varieties over X (here we use that X is of Hodge type) and let $\pi^n: A^n \to X$ be the n-fold self fiber product of $\pi: A \to X$. Then Proposition 3.2 of [29] tells us that every automorphic local system \mathbb{V} is a direct summand of $R\pi_*^n\mathbb{Q}_\ell$, up to shift and Tate twist. So it suffices to prove weight monodromy for

$$H^{\bullet}(X_{\overline{F}}^*, j_{!*}R\pi_*^n\mathbb{Q}_{\ell}).$$

Let $\sigma: A^* \to X^*$ be a projective morphism extending π (which exists by [5]) and let σ^n be its *n*-fold self fiber product, which sits in the following Cartesian diagram over \overline{F} .

$$A_{\overline{F}}^{n} \xrightarrow{k} A_{\overline{F}}^{*,n}$$

$$\downarrow^{\pi^{n}} \qquad \downarrow^{\sigma^{n}}$$

$$X_{\overline{F}} \xrightarrow{j} X_{\overline{F}}^{*}$$

Let us note that $k_{!*}\mathbb{Q}_{\ell}[d]$, where d is the dimension of A^n , is the intersection cohomology complex $IC_{A^{\star,n}}$ of $A^{\star,n}$. Theorem 1.8 of [50] (a refinement of the decomposition theorem) tells us that there is a Galois equivariant decomposition

$$R\sigma_*^n \operatorname{IC}_{A^{\star,n}} \simeq \bigoplus_i {}^{\mathfrak{p}} \mathcal{H}^i (R\sigma_*^n \operatorname{IC}_{A^{\star,n}})[-i].$$

and similarly

$$R\pi^n_*\mathbb{Q}_\ell[d] \simeq \bigoplus_i R^i \pi^n \mathbb{Q}_\ell[d-i].$$

Claim A.1.1. For each i there is a Galois-equivariant splitting

$$j_{!*}R^i\pi^n_*\mathbb{Q}_\ell[d-i] \subset^{\oplus} {}^{\mathfrak{p}}\mathcal{H}^i(R\sigma^n_*\operatorname{IC}_{A^{\star,n}})[-i].$$

Proof. The proper base change theorem tells us that the left hand side and the right hand side have the same restriction to $X_{\overline{F}} \subset X_{\overline{F}}^*$. Then Lemma 2.2.8 of [50] gives us the result (rather its Galois equivariance).

Lemma 1.4 of [52] tells us that validity of weight-monodromy passes to direct summands. We compute

$$H^{\bullet}(X_{\overline{F}}^{*}, j_{!*}R\pi_{*}^{n}\mathbb{Q}_{\ell}) \subset^{\oplus} H^{\bullet}(X_{\overline{F}}^{*}, R\sigma_{*}^{n}k_{!*}\mathbb{Q}_{\ell})$$

$$= H^{\bullet}(A_{\overline{F}}^{*,n}, k_{!*}\mathbb{Q}_{\ell})$$

$$= IH^{\bullet}(A_{\overline{F}}^{*,n}, \mathbb{Q}_{\ell}),$$

and deduce that weight monodromy for the cohomology of $j_{!*}\mathbb{V}$ follows from weight-monodromy for the intersection cohomology of $A^{*,n}$. Now let $\rho: Y \to A^{*,n}$ be a smooth projective morphism with Y smooth projective.

Lemma A.1.2 (Corollary 1 of [18]). The intersection cohomology complex of $A^{*,n}$ is a Galois-equivariant direct summand of

$$R\rho_*\mathbb{Q}_\ell$$
.

Proof. We know that ${}^{\mathfrak{p}}\mathcal{H}^{0}(R\rho_{*}\mathbb{Q}_{\ell})$ is a Galois-equivariant direct summand of $R\rho_{*}\mathbb{Q}_{\ell}$ by Theorem 1.8 of [50]. Let $a:U\to A^{*,n}$ be a dense open subset over which ρ is smooth and such that U is contained in A^{n} . Then there is a Galois-equivariant decomposition

$$a_{!*}a^{*\mathfrak{p}}\mathcal{H}^0(R\rho_*\mathbb{Q}_\ell) \subset^{\oplus \mathfrak{p}}\mathcal{H}^0(R\rho_*\mathbb{Q}_\ell)$$

by Lemma 2.2.8 of [50]. We know that the intersection cohomology complex of $A^{*,n}$ is equal to $a_{!*}a^*\mathbb{Q}_{\ell}[d]$ and since intermediate extension is fully faithful it suffices to show that \mathbb{Q}_{ℓ} is a Galois-equivariant direct summand of $a^{*p}\mathcal{H}^0(R\rho_*\mathbb{Q}_{\ell})$. But the latter is equal to

$$R^0\tau_*\mathbb{O}_\ell$$
.

where τ is the restriction of ρ to U. The natural map

$$\mathbb{O}_{\ell} \to R^0 \tau_* \mathbb{O}_{\ell}$$

has a section coming from the trace map.

This means that weight-monodromy for Y implies weight-monodromy for the intersection cohomology of $A^{*,n}$, proving the theorem.

References

- [1] James Arthur, Automorphic representations of GSp(4), Contributions to automorphic forms, geometry, and number theory, 2004, pp. 65–81. MR2058604
- Mahdi Asgari and Ralf Schmidt, Siegel modular forms and representations, Manuscripta Math. 104 (2001), no. 2, 173-200. MR1821182
- [3] Thomas Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor, *Potential automorphy and change of weight*, Ann. of Math. (2) **179** (2014), no. 2, 501–609. MR3152941
- [4] Ana Caraiani, Local-global compatibility and the action of monodromy on nearby cycles, Duke Math. J. 161 (2012), no. 12, 2311–2413. MR2972460
- [5] Brian Conrad, Deligne's notes on Nagata compactifications, J. Ramanujan Math. Soc. 22 (2007), no. 3, 205–257. MR2356346
- [6] P. Deligne and N. Katz (eds.), Groupes de monodromie en géométrie algébrique. II, Lecture Notes in Mathematics, Vol. 340, Springer-Verlag, Berlin-New York, 1973. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II). MR0354657
- [7] Pierre Deligne, La conjecture de Weil. II, Inst. Hautes Études Sci. Publ. Math. 52 (1980), 137–252. MR601520
- [8] Lassina Dembélé, On the computation of algebraic modular forms on compact inner forms of GSp₄, Math. Comp.
 83 (2014), no. 288, 1931–1950. MR3194136
- [9] Fred Diamond and Richard Taylor, Nonoptimal levels of mod l modular representations, Invent. Math. 115 (1994), no. 3, 435–462. MR1262939
- [10] Dan Fretwell, Genus 2 paramodular Eisenstein congruences, Ramanujan J. 46 (2018), no. 2, 447–473. MR3803970
- [11] Daniel Fretwell, Level p paramodular congruences of Harder type, 2015.
- [12] Wee Teck Gan and Shuichiro Takeda, The local Langlands conjecture for GSp(4), Ann. of Math. (2) 173 (2011), no. 3, 1841–1882. MR2800725
- [13] Toby Gee and David Geraghty, Companion forms for unitary and symplectic groups, Duke Math. J. 161 (2012), no. 2, 247–303. MR2876931
- [14] Toby Gee and Olivier Taïbi, Arthur's multiplicity formula for GSp\$_4\$ and restriction to Sp\$_4\$, arXiv e-prints (2018Jul), arXiv:1807.03988, available at 1807.03988.
- [15] Alexandru Ghitza, Hecke eigenvalues of Siegel modular forms (mod p) and of algebraic modular forms, J. Number Theory 106 (2004), no. 2, 345–384. MR2059479
- [16] Benedict H. Gross, Algebraic modular forms, Israel J. Math. 113 (1999), 61–93. MR1729443
- [17] Lucio Guerberoff, Modularity lifting theorems for Galois representations of unitary type, Compos. Math. 147 (2011), no. 4, 1022–1058. MR2822860
- [18] David Hansen, A trick and the decomposition theorem, 2020.

- [19] Tomoyoshi Ibukiyama, On symplectic Euler factors of genus two, Proc. Japan Acad. Ser. A Math. Sci. 57 (1981), no. 5, 271–275. MR624040
- [20] _____, On relations of dimensions of automorphic forms of Sp(2, **R**) and its compact twist Sp(2). I, Automorphic forms and number theory (Sendai, 1983), 1985, pp. 7–30. MR876100
- [21] ______, Paramodular forms and compact twist, Automorphic forms on GSp(4): Proceedings of the 9th Autumn Workshop on Number Theory, November 6 to 10, 2006, Hakuba, Japan, 2006, pp. 37–48.
- [22] ______, Dimension formulas of siegel modular forms of weight 3 and supersingular abelian varieties, Proceedings of the 4-th Spring Conference on modular forms and related topics, "Siegel Modular Forms and Abelian Varieties", 2007
- [23] Yasutaka Ihara, On certain arithmetical Dirichlet series, J. Math. Soc. Japan 16 (1964), 214–225. MR0179142
- [24] H. Jacquet and R. P. Langlands, Automorphic forms on GL(2), Lecture Notes in Mathematics, Vol. 114, Springer-Verlag, Berlin-New York, 1970. MR0401654
- [25] Frazer Jarvis, Mazur's principle for totally real fields of odd degree, Compositio Math. 116 (1999), no. 1, 39–79. MR1669444
- [26] Toshiyuki Katsura and Frans Oort, Families of supersingular abelian surfaces, Compositio Math. 62 (1987), no. 2, 107–167. MR898731
- [27] Stephen S. Kudla and Michael Rapoport, Cycles on Siegel threefolds and derivatives of Eisenstein series, Ann. Sci. École Norm. Sup. (4) 33 (2000), no. 5, 695–756. MR1834500
- [28] Kai-Wen Lan, Arithmetic compactifications of PEL-type Shimura varieties, London Mathematical Society Monographs Series, vol. 36, Princeton University Press, Princeton, NJ, 2013. MR3186092
- [29] Kai-Wen Lan and Benoît Stroh, Nearby cycles of automorphic étale sheaves, II, Cohomology of arithmetic groups, 2018, pp. 83–106. MR3848816
- [30] Kai-Wen Lan and Junecue Suh, Vanishing theorems for torsion automorphic sheaves on general PEL-type Shimura varieties, Adv. Math. 242 (2013), 228–286. MR3055995
- [31] Joshua Lansky and David Pollack, Hecke algebras and automorphic forms, Compositio Math. 130 (2002), no. 1, 21–48. MR1883690
- [32] Gérard Laumon, Fonctions zêtas des variétés de Siegel de dimension trois, Astérisque **302** (2005), 1–66. Formes automorphes. II. Le cas du groupe GSp(4). MR2234859
- [33] Eduard Looijenga, L^2 -cohomology of locally symmetric varieties, Compositio Math. **67** (1988), no. 1, 3–20. MR949269
- [34] James S. Milne, Lectures on etale cohomology (v2.21), 2013. 202 pages, Available at www.jmilne.org/math/.
- [35] Chung Pang Mok, Galois representations attached to automorphic forms on GL₂ over CM fields, Compos. Math. **150** (2014), no. 4, 523–567. MR3200667
- [36] Moonen, Ben, van der Geer, Gerard, and Edixhoven, Bas, Abelian varieties, 2014.
- [37] Dan Petersen, Cohomology of local systems on the moduli of principally polarized abelian surfaces, Pacific J. Math. 275 (2015), no. 1, 39–61. MR3336928
- [38] Ali Rajaei, On the levels of mod l Hilbert modular forms, J. Reine Angew. Math. 537 (2001), 33-65. MR1856257
- [39] K. A. Ribet, On modular representations of $Gal(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms, Invent. Math. **100** (1990), no. 2, 431–476. MR1047143
- [40] Brooks Roberts and Ralf Schmidt, On modular forms for the paramodular groups, Automorphic forms and zeta functions, 2006, pp. 334–364. MR2208781
- [41] _____, Local newforms for GSp(4), Lecture Notes in Mathematics, vol. 1918, Springer, Berlin, 2007. MR2344630
- [42] Mirko Rösner and Rainer Weissauer, Global liftings between inner forms of GSp(4), 2020. in preparation.
- [43] Abhishek Saha and Ralf Schmidt, Yoshida lifts and simultaneous non-vanishing of dihedral twists of modular L-functions, J. Lond. Math. Soc. (2) 88 (2013), no. 1, 251–270. MR3092267
- [44] Leslie Saper and Mark Stern, L₂-cohomology of arithmetic varieties, Ann. of Math. (2) 132 (1990), no. 1, 1–69. MR1059935
- [45] Ralf Schmidt, Paramodular forms in CAP representations of GSp(4), 2018. preprint, available at http://www.math.ou.edu/~rschmidt/papers/CAP_paramodular.pdf.
- [46] J.-P. Serre, Lettre à J.-F. Mestre, Current trends in arithmetical algebraic geometry (Arcata, Calif., 1985), 1987, pp. 263–268. MR902597
- [47] Xu Shen, Chia-Fu Yu, and Chao Zhang, EKOR strata for Shimura varieties with parahoric level structure, arXiv e-prints (October 2019), available at 1910.07785.

- [48] Claus M. Sorensen, Level-raising for Saito-Kurokawa forms, Compos. Math. 145 (2009), no. 4, 915–953. MR2521249
- [49] _____, Potential level-lowering for GSp(4), J. Inst. Math. Jussieu 8 (2009), no. 3, 595–622. MR2516307
- [50] Shenghao Sun and Weizhe Zheng, Parity and symmetry in intersection and ordinary cohomology, Algebra Number Theory 10 (2016), no. 2, 235–307. MR3477743
- [51] Richard Taylor, On the l-adic cohomology of Siegel threefolds, Invent. Math. 114 (1993), no. 2, 289–310. MR1240640
- [52] Richard Taylor and Teruyoshi Yoshida, Compatibility of local and global Langlands correspondences, J. Amer. Math. Soc. 20 (2007), no. 2, 467–493. MR2276777
- [53] Jack A. Thorne, Raising the level for GL_n, Forum Math. Sigma 2 (2014), e16, 35. MR3264255
- [54] J. Tilouine, Siegel varieties and p-adic Siegel modular forms, Doc. Math. Extra Vol. (2006), 781–817. MR2290605
- [55] Gerard van der Geer, Siegel Modular Forms, arXiv Mathematics e-prints (2006May), math/0605346, available at math/0605346.
- [56] David A. Vogan Jr. and Gregg J. Zuckerman, Unitary representations with nonzero cohomology, Compositio Math. 53 (1984), no. 1, 51–90. MR762307
- [57] Haining Wang, Level lowering for GSp(4) and vanishing cycles on Siegel threefolds, arXiv e-prints (October 2019), arXiv:1910.07569, available at 1910.07569.
- [58] Rainer Weissauer, Existence of Whittaker models related to four dimensional symplectic Galois representations, Modular forms on Schiermonnikoog, 2008, pp. 285–310. MR2530981
- [59] _____, Endoscopy for GSp(4) and the cohomology of Siegel modular threefolds, Lecture Notes in Mathematics, vol. 1968, Springer-Verlag, Berlin, 2009. MR2498783
- [60] Takuya Yamauchi, Serre weights for GSp₄ over totally real fields, arXiv e-prints (June 2020), arXiv:2006.07824, available at 2006.07824.
- [61] Chia-Fu Yu, The supersingular loci and mass formulas on Siegel modular varieties, Doc. Math. 11 (2006), 449–468. MR2288077
- [62] _____, Geometry of the Siegel modular threefold with paramodular level structure, Proc. Amer. Math. Soc. 139 (2011), no. 9, 3181–3190. MR2811273

Email address: pol.van_hoften@kcl.ac.uk

King's College London