

MOTIVIC COHOMOLOGY OF QUATERNIONIC SHIMURA VARIETIES AND  
LEVEL RAISING  
COHOMOLOGIE MOTIVIQUE DES VARIÉTÉS DE SHIMURA  
QUATERNIONIQUE ET AUGMENTATION DU NIVEAU

RONG ZHOU

ABSTRACT. We study the motivic cohomology of the special fiber of quaternionic Shimura varieties at a prime of good reduction. We exhibit classes in these motivic cohomology groups and use this to give an explicit geometric realization of level raising between Hilbert modular forms. The main ingredient for our construction is a form of Ihara’s Lemma for compact quaternionic Shimura surfaces which we prove by generalizing a method of Diamond–Taylor. Along the way we also verify the Hecke orbit conjecture for these quaternionic Shimura varieties which is a key input for our proof of Ihara’s Lemma.

RÉSUMÉ. Nous étudions la cohomologie motivique de la fibre spéciale des variétés de Shimura quaternioniques à un nombre premier de bonne réduction. Nous explicitons des classes dans ces groupes de cohomologie motivique et utilisons cela pour donner une réalisation géométrique explicite de l’augmentation de niveau entre les formes modulaires de Hilbert. L’ingrédient principal de notre construction est une forme du lemme d’Ihara pour les surfaces de Shimura compactes quaternioniques, que nous prouvons en généralisant une méthode de Diamond–Taylor. En cours de route, nous vérifions également la conjecture de l’orbite de Hecke pour ces variétés de Shimura quaternioniques, qui est un point-clé de notre preuve du lemme d’Ihara.

1. INTRODUCTION

1.1. **Main Theorem.** The aim of this paper is to study the motivic cohomology of the special fiber of certain quaternionic Shimura varieties. For a scheme of finite type over a field, its motivic cohomology groups are a generalization of the usual Chow groups, and the main new observation of this paper is that for certain Shimura varieties, these groups can encode very rich arithmetic information. More precisely, we will show that the cycle class map from motivic cohomology to étale cohomology gives a geometric realization of level raising between Hilbert modular forms. The construction is also related to a geometric realization of the mod  $\ell$  Jacquet–Langlands correspondence.

We now state our main result. Let  $F$  be a totally real field of even degree  $[F : \mathbb{Q}] = g$  and  $p > 2$  a prime which is *inert* in  $F$ . Let  $B$  be a totally indefinite quaternion algebra over  $F$  which is unramified at the unique prime  $\mathfrak{p}$  above  $p$  and  $G$  the associated reductive group over  $\mathbb{Q}$ . Let  $K$  be a sufficiently small compact open subgroup of  $G(\mathbb{A}_f)$  such that  $K = K_{\mathfrak{p}}K^p$  where  $K_{\mathfrak{p}} \subset G(\mathbb{Q}_{\mathfrak{p}}) = \mathrm{GL}_2(F_{\mathfrak{p}})$  is the standard hyperspecial maximal compact  $\mathrm{GL}_2(\mathcal{O}_{F_{\mathfrak{p}}})$  and  $K^p \subset G(\mathbb{A}_f^p)$ . Then there is a Shimura variety  $\mathrm{Sh}_K(G)$  defined over  $\mathbb{Q}$ ; it extends to a smooth integral model  $\underline{\mathrm{Sh}}_K(G)$  over  $\mathbb{Z}_{(p)}$ . We let  $\mathcal{S}_K(G)$  denote its special fiber over  $\mathbb{F}_p$  and  $\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}$  its base change to  $\mathbb{F}_{p^g}$ .

Fix an irreducible cuspidal automorphic representation  $\Pi$  of  $\mathrm{GL}_2(F)$  of parallel weight 2 defined over a number field  $\mathbf{E}$ . Let  $R$  be a finite set of places of  $F$  not containing  $\mathfrak{p}$  and away from which  $\Pi$  is unramified and  $K$  is hyperspecial. We also choose a prime  $\lambda$  of  $\mathcal{O}_{\mathbf{E}}$  whose residue characteristic

---

*Date:* October 5, 2021.

*2010 Mathematics Subject Classification.* 11F33, 11F41, 14F42,

*Key words and phrases.* Shimura varieties, motivic cohomology, level raising, Ihara’s Lemma.

is coprime to  $p$  and write  $k_\lambda = \mathcal{O}_{\mathbf{E}}/\lambda$ . We write  $\mathbf{H}_{\mathcal{M}}^i(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, k_\lambda(j))$  for the motivic cohomology group with  $k_\lambda$  coefficients defined in [SV00]. By [Voe02], we may identify this with the higher Chow group  $\mathrm{Ch}^j(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 2j - i, k_\lambda)$  defined in [Blo86]. When  $2j = i$ , this group is just the usual Chow group of codimension  $j$  cycles modulo rational equivalence (with coefficients in  $k_\lambda$ ). The group  $\mathrm{Ch}^j(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 2j - i, k_\lambda)$  is equipped with the following cycle class map to the absolute étale cohomology:

$$(1.1.1) \quad \mathrm{Ch}^j(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 2j - i, k_\lambda) \rightarrow \mathbf{H}_{\text{ét}}^i(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, k_\lambda(j)).$$

We let  $\mathbf{T}_R$  denote the abstract Hecke algebra of  $\mathrm{GL}_2(F)$  away from  $R$ ; it is the  $\mathbb{Z}$ -algebra generated by elements  $T_{\mathfrak{q}}, S_{\mathfrak{q}}$  where  $\mathfrak{q}$  runs over primes of  $F$  away from  $R$ . Then the Hecke eigenvalues of  $\Pi$  induce a map

$$\phi_\lambda^\Pi : \mathbf{T}_R \rightarrow \mathcal{O}_{\mathbf{E}} \rightarrow k_\lambda.$$

We write  $\mathfrak{m}_R := \ker(\phi_\lambda^\Pi)$  a maximal ideal of  $\mathbf{T}_R$  and we let  $\mathfrak{m} \subset \mathbf{T}_{R \cup \{\mathfrak{p}\}}$  be the preimage of  $\mathfrak{m}_R$  in  $\mathbf{T}_{R \cup \{\mathfrak{p}\}}$  under the natural inclusion  $\mathbf{T}_{R \cup \{\mathfrak{p}\}} \rightarrow \mathbf{T}_R$ .

The Hecke algebra  $\mathbf{T}_{R \cup \{\mathfrak{p}\}}$  acts on the étale cohomology  $\mathbf{H}_{\text{ét}}^\bullet(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(-))$  and higher Chow groups  $\mathrm{Ch}^j(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 2j - i, k_\lambda)$  of  $\mathcal{S}_K(G)$ . Upon making a large image assumption on the  $\bmod \lambda$  Galois representation associated to  $\Pi$  (see Assumption 4.1.1) and localizing at the maximal ideal  $\mathfrak{m}$ , there is an isomorphism

$$\mathbf{H}_{\text{ét}}^{g+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, k_\lambda(g/2 + 1))_{\mathfrak{m}} \cong \mathbf{H}^1(\mathbb{F}_{p^g}, \mathbf{H}_{\text{ét}}^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(g/2 + 1))_{\mathfrak{m}}).$$

The cycle class map then induces the *Abel–Jacobi map*:

$$(1.1.2) \quad \mathrm{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}} \rightarrow \mathbf{H}^1(\mathbb{F}_{p^g}, \mathbf{H}_{\text{ét}}^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(g/2 + 1))_{\mathfrak{m}}).$$

In §5.3, we will define a subgroup  $\mathrm{Ch}_{\mathrm{Ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}}$  of  $\mathrm{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}}$  using the geometry of Goren–Oort cycles (in fact the cycles we consider arise from the supersingular locus) on  $\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}$  as studied in [TX16], [TX19] and [LT20]. As the notation suggests, this subgroup is related to level raising. The main theorem of the paper is the following; we refer to §5.3 for the precise statement.

**Theorem 1.1.1.** *Suppose that  $p$  is a  $\lambda$ -level raising prime in the sense of Definition 5.3.2 and that Assumptions 4.1.1 and 5.3.5 are satisfied; in particular  $T_{\mathfrak{p}} \equiv p^g + 1 \pmod{\mathfrak{m}_R}$  and  $S_{\mathfrak{p}} \equiv 1 \pmod{\mathfrak{m}_R}$ . Then the map*

$$\mathrm{Ch}_{\mathrm{Ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}} \rightarrow \mathbf{H}^1(\mathbb{F}_{p^g}, \mathbf{H}^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(g/2 + 1))_{\mathfrak{m}})$$

*induced by (1.1.2) is surjective.*

We note that as in [LT20, Remark 4.2, 4.6], if there exist rational primes inert in  $F$ , and  $\Pi$  is not dihedral and not isomorphic to a twist by a character of any of its internal conjugates, then for all but finitely many  $\lambda$ , the set of primes  $p$  which are  $\lambda$ -level raising primes has positive density.

In general it is a difficult problem to produce non-zero classes in motivic cohomology. The key input to proving the surjectivity in Theorem 1.1.1 is a form of Ihara’s Lemma which we prove by generalizing a method of Diamond–Taylor [DT94]; see the next subsection for more details.

We now give an example of the construction of  $\mathrm{Ch}_{\mathrm{Ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$  which makes clear the relationship with level raising. We assume  $g = 2$  so that  $\dim \mathcal{S}_K(G) = 2$ .

We write  $B'$  for the totally definite quaternion algebra which agrees with  $B$  at all finite places. We fix an isomorphism

$$B' \otimes_{\mathbb{Q}} \mathbb{A}_f \cong B \otimes_{\mathbb{Q}} \mathbb{A}_f$$

which allows us to consider  $K$  as a compact open subgroup of  $B' \otimes_{\mathbb{Q}} \mathbb{A}_f$ . We let  $\mathcal{X}'$  and  $\mathcal{X}'_0(\mathfrak{p})$  denote the discrete Shimura sets

$$\mathcal{X}' := B' \backslash B' \otimes_{\mathbb{Q}} \mathbb{A}_f / K, \quad \mathcal{X}'_0(\mathfrak{p}) := B' \backslash B' \otimes_{\mathbb{Q}} \mathbb{A}_f / K_0(\mathfrak{p})$$

where the compact open subgroup  $K_0(\mathfrak{p}) \subset B' \otimes_{\mathbb{Q}} \mathbb{A}_f$  agrees with  $K$  away from  $\mathfrak{p}$  and is the standard Iwahori subgroup of  $\mathrm{GL}_2(F_{\mathfrak{p}})$  at  $\mathfrak{p}$ . We let

$$\pi_1, \pi_2 : \mathcal{X}'_0(\mathfrak{p}) \rightarrow \mathcal{X}'$$

denote the natural degeneracy maps so that the diagram

$$\mathcal{X}' \xleftarrow{\pi_1} \mathcal{X}'_0(\mathfrak{p}) \xrightarrow{\pi_2} \mathcal{X}'$$

is the usual Hecke correspondence for  $\mathcal{X}'$ . For any finite set  $S$ , we write  $\Gamma(S, k_\lambda)$  for the abelian group of  $k_\lambda$ -valued functions on  $S$ .

We may think of  $\mathcal{S}_K(G)$  as a moduli space of abelian varieties with multiplication by some maximal order  $\mathcal{O}_B$  in  $B$ . We let  $\mathcal{S}_K(G)^{\mathrm{ss}}$  be the locus where the underlying abelian variety is supersingular. Using the geometry of  $\mathcal{S}_K(G)^{\mathrm{ss}}$  one can show that under the assumptions of Theorem 1.1.1,  $\mathrm{Ch}^2(\mathcal{S}_K(G)_{\mathbb{F}_{p^2}}, 1, k_\lambda)_{\mathfrak{m}}$  admits a map from

$$K_{\mathfrak{m}} := \ker((\pi_{1*}, \pi_{2*}) : \Gamma(\mathcal{X}'_0(\mathfrak{p}), k_\lambda) \rightarrow \Gamma(\mathcal{X}', k_\lambda))_{\mathfrak{m}}.$$

The construction uses an interpretation of classes in  $\mathrm{Ch}^2(\mathcal{S}_K(G)_{\mathbb{F}_{p^2}}, 1, k_\lambda)_{\mathfrak{m}}$  as codimension 1 cycles together with a rational function on the cycle; see §5.3 for the details. Then  $\mathrm{Ch}_{\mathrm{Ir}}^2(\mathcal{S}_K(G)_{\mathbb{F}_{p^2}}, 1, k_\lambda)_{\mathfrak{m}}$  is defined to be the image of  $K_{\mathfrak{m}}$ . Theorem 1.1.1 in this case follows from the following stronger result:

**Theorem 1.1.2.** *Let  $g = 2$ . Suppose that  $p$  is a  $\lambda$ -level raising prime and that Assumption 4.1.1 is satisfied. Then the map*

$$(1.1.3) \quad K_{\mathfrak{m}} \rightarrow \mathrm{H}^1(\mathbb{F}_{p^2}, \mathrm{H}_{\mathrm{ét}}^2(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}})$$

is surjective.

The relationship with level raising should now be clear. Indeed under the Jacquet–Langlands correspondence,  $\mathbf{T}_{R \cup \{\mathfrak{p}\}}$  acts on left hand side of (1.1.3) via the  $\mathfrak{p}^{\mathrm{new}}$  quotient in the sense of [Rib88], whereas it is well known that it acts via the  $\mathfrak{p}^{\mathrm{old}}$  quotient on the right hand side. In this sense, the Abel–Jacobi map gives an explicit realization of the congruence between old and new forms.

*Relation with motivic Bloch–Kato conjecture:* It is known by the work of many authors that the motivic cohomology groups satisfy many of the formal properties of a cohomology theory. However there is much that is still not understood, we refer to [Gei05] for a brief survey. We may use Theorem 1.1.2 to show that in certain cases of Shimura surfaces, motivic cohomology coincides with étale cohomology, upon localizing at  $\mathfrak{m}$ .

**Theorem 1.1.3.** *Let  $g = 2$ . Suppose that  $p$  is a  $\lambda$ -level raising prime and that Assumption 4.1.1 is satisfied. Then the cycle class map induces an isomorphism*

$$\mathrm{H}_{\mathcal{M}}^3(\mathcal{S}_K(G)_{\mathbb{F}_{p^2}}, k_\lambda(2))_{\mathfrak{m}} \xrightarrow{\sim} \mathrm{H}_{\mathrm{ét}}^3(\mathcal{S}_K(G)_{\mathbb{F}_{p^2}}, k_\lambda(2))_{\mathfrak{m}}.$$

When  $i \leq j$ , Voevodsky’s proof of the motivic Bloch–Kato conjecture [Voe11] implies that  $\mathrm{Ch}^j(X, 2j - i, k_\lambda)$  is isomorphic to  $\mathrm{H}_{\mathrm{ét}}^i(X, k_\lambda(j))$ , for  $X$  proper smooth over any base field. For  $i > j$ , not much seems to be known.

*Remark 1.1.4.* When  $g$  is odd, there is an Abel–Jacobi map

$$(1.1.4) \quad \mathrm{Ch}^{\frac{g-1}{2}}(\mathcal{S}_K(G)_{\mathbb{F}_{p^{2g}}}, k_\lambda)_{\mathfrak{m}} \rightarrow \mathrm{H}^1(\mathbb{F}_{p^{2g}}, \mathrm{H}_{\mathrm{ét}}^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(\lfloor g/2 \rfloor + 1))_{\mathfrak{m}}).$$

In this case the supersingular locus  $\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}^{\mathrm{ss}}$  is equidimensional of dimension  $\frac{g-1}{2}$  and we may consider the subgroup  $\mathrm{Ch}_{\mathrm{Ir}}^{\frac{g-1}{2}}(\mathcal{S}_K(G)_{\mathbb{F}_{p^{2g}}}, k_\lambda)_{\mathfrak{m}}$  generated by the irreducible components in  $\mathcal{S}_K(G)_{\mathbb{F}_{p^{2g}}}^{\mathrm{ss}}$ . Then [LT20, Theorem 1.3] have shown the surjectivity of (1.1.4) modulo  $\mathfrak{m}$  restricted to this subgroup. Thus our Theorem 1.1.1 can be thought of as the even-dimensional analogue of the theorem of Liu–Tian. The main new observation of this work is that we are able to produce certain classes in motivic cohomology, or higher Chow groups, as opposed to ordinary Chow groups. Its conceptual

importance lies in the fact that we are able to obtain a geometric interpretation of *even*-dimensional Galois cohomology.

*Remark 1.1.5.* In [LT20], the geometric realization of level raising was a key ingredient in their proof of certain cases of the Bloch–Kato conjecture<sup>1</sup> on special values of  $L$ -functions. Our work should have applications to cases of this conjecture for non-central  $L$ -values; we aim to carry this out in a future work.

**1.2. Proof of main result and Ihara’s Lemma.** We now explain the proof of Theorem 1.1.1. Our approach follows that of [LT20], but there are many new difficulties in the even dimensional case.

Firstly, using the construction of the group  $\mathrm{Ch}_{\mathrm{Ih}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$  and the intersection pairing between certain Goren–Oort strata proved in [TX19], we reduce to proving the surjectivity statement in the case of quaternionic Shimura surfaces (cf. Proposition 5.4.3). The statement in this case follows from the following form of Ihara’s Lemma. These type of results first appeared in Ribet’s ICM article [Rib84] for the case of modular curves and over the last thirty years they have seen many important arithmetic applications. Therefore our result in the case of surfaces should be of independent interest.

For simplicity, we only state the result in the totally indefinite case; we refer to Theorem 4.1.4 for the more general statement. Thus we assume  $g = 2$  as in the example of the previous subsection.

**Theorem 1.2.1** (Ihara’s Lemma). *Under Assumption 4.1.3, the map*

$$\pi_1^* + \pi_2^* : \mathrm{H}_{\mathrm{ét}}^2(\mathrm{Sh}_K(G)_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}^2 \rightarrow \mathrm{H}_{\mathrm{ét}}^2(\mathrm{Sh}_{K_0(p)}(G)_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}$$

*is injective.*

Here  $\mathrm{Sh}_{K_0(p)}(G)$  denotes the quaternionic Shimura surface with Iwahori level structure at  $p$  and  $\pi_1, \pi_2$  are the natural degeneracy maps. In fact the appropriate Abel–Jacobi map in this case can be related to the map  $\pi_1^* + \pi_2^*$  in the statement of Theorem 1.2.1; they are essentially dual to one another. To show the existence of this duality requires a careful analysis of the global geometry of the mod  $p$  fiber of the quaternionic Shimura surface *with Iwahori level structure* at  $p$ . We note that in this case the Shimura surface has bad reduction at  $p$ . The main result which is Corollary A.5.4 is proved in an appendix and is analogous to the results of [Sta97] in the case of Hilbert modular surfaces.

We now describe our approach to Theorem 1.2.1. The result in the case of Hilbert modular varieties has been proved by Dimitrov [Dim09]. However his proof relied crucially on the existence of a  $q$ -expansion. Note that when  $g > 2$ , even if one is interested in Theorem 1.1.1 for Hilbert modular varieties, the reduction to the case of surfaces will require us to consider compact Shimura surfaces where a  $q$ -expansion is not available. We therefore take another approach by generalizing a method of Diamond–Taylor who proved the result for Shimura curves [DT94].

We first apply a crystalline comparison isomorphism to reduce the problem to proving injectivity of a certain map between global sections of line bundles over the mod  $\ell$  reduction of the Shimura surface (Proposition 4.4.1), where  $\ell$  is the rational prime underlying  $\lambda$ . The property that a non-zero section lies in the kernel implies that the divisor  $D$  corresponding to this section is invariant under  $p$ -power Hecke operators. In the case of Shimura curves, it is known that the image of an ordinary point under  $p$ -power Hecke operators is infinite; this constrains  $D$  to be supported on the supersingular locus. Since  $p$ -power Hecke operators act transitively on supersingular points, the support contains the supersingular locus and this is enough to deduce a contradiction for degree reasons.

In the case of surfaces, we need a stronger result to constrain the support of the divisor  $D$ . In section §3, we prove a version of the Hecke orbit conjecture of Chai–Oort [CO19] for the ordinary locus on quaternionic Shimura varieties. We assume  $\ell$  is a prime where the compact open  $K$  is hyperspecial and we write  $\mathcal{S}_K^\ell(G)$  for the mod  $\ell$  reduction of the integral model  $\underline{\mathrm{Sh}}_K^\ell(G)$  at a prime of the reflex field above  $\ell$ . We write  $\mathcal{S}_K^\ell(G)^{\mathrm{ord}}$  for the locus where the universal abelian variety is ordinary.

<sup>1</sup>This is a different conjecture to the motivic Bloch–Kato conjecture mentioned above.

**Theorem 1.2.2** (Hecke orbit Conjecture). *Let  $x \in \mathcal{S}_K^\ell(G)^{\text{ord}}(\overline{\mathbb{F}}_\ell)$ . Then the prime-to- $\ell$  Hecke orbit is Zariski dense in  $\mathcal{S}_K^\ell(G)$ .*

In fact we prove this result in a more general situation; we refer to §3.1 for the statement. Using the strong approximation theorem, we deduce that the  $p$ -power Hecke orbit of  $x \in \mathcal{S}_K^\ell(G)^{\text{ord}}(\overline{\mathbb{F}}_\ell)$  is Zariski dense in the connected component of  $\mathcal{S}_K^\ell(G)_{\overline{\mathbb{F}}_\ell}$  containing it. This allows one to show that  $D$  is supported on the complement of  $\mathcal{S}_K^\ell(G)^{\text{ord}}$ . A computation involving intersection numbers of  $D$  with certain cycles on  $\mathcal{S}_K^\ell(G)$  then gives the desired contradiction.

*Remark 1.2.3.* In [LT20], the authors reduce their surjectivity result to a form of Ihara’s Lemma for Shimura curves, which was proved by Manning–Shotton [MS21] using a different method.

We note that many of the quaternionic Shimura varieties we consider do not admit good moduli interpretations. Thus in order to obtain the geometric results we need, we study the geometry of certain auxiliary unitary Shimura varieties which are of PEL type. Using [TX16, §2], the results for unitary Shimura varieties transfer easily to the quaternionic side. The moduli interpretation for the unitary Shimura varieties allow us to adapt many proofs in the case of Hilbert modular varieties to the quaternionic case.

**1.3. Outline of the paper.** In §2 we begin with some basics on Shimura varieties and define the quaternionic Shimura varieties of interest. We recall the construction of the auxiliary unitary Shimura varieties of PEL type as in [TX16, §3], and recall the description of Goren–Oort cycles obtained in [TX19]. In §3 we prove the Hecke orbit conjecture for quaternionic Shimura varieties. We deduce our results from the corresponding statement for the auxiliary unitary Shimura varieties. Using the moduli interpretation, the proof in the unitary case follows the strategy of [Cha95] who proved the result for Hilbert modular varieties. A key input here is Moonen’s generalization of Serre–Tate theory for ordinary abelian varieties [Moo04]. In §4 we study the intersection pairing of cycles on the mod  $\ell$  reduction of Shimura surfaces and use this to prove Theorem 1.2.1. Finally in §5, we recall the definition of motivic cohomology groups and higher Chow groups, paying extra attention in the most pertinent case of surfaces, and we construct the level raising subgroup  $\text{Ch}_{\text{ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_m$ . We then prove Theorem 1.1.1 using the strategy outlined above. In the appendix we describe the bad reduction of quaternionic Shimura surfaces with Iwahori level structure.

*Acknowledgments:* The author would like to thank Tony Feng, Bao Le Hung, Chao Li, Ananth Shankar, Richard Taylor, Liang Xiao and Xinwen Zhu for useful discussions and comments. This paper started out as a joint work between the author and Akshay Venkatesh, who initially observed that the motivic cohomology of Shimura varieties should encode interesting arithmetic information. He generously declined to be included on the final version of the paper despite making a substantial contribution. The author would like to thank him heartily for his contribution and for many hours of interesting and enjoyable discussion concerning the material in this paper. Finally the author would like to thank the anonymous referees for a careful reading of the manuscript and for many useful comments and suggestions. The author was partially supported by NSF grant No. DMS-1638352 through membership at the Institute for Advanced Study.

#### 1.4. Notation.

- If  $F$  is a number field we write  $\mathcal{O}_F$  for its ring of integers. If  $v$  is a place of  $F$ , we write  $F_v$  for the completion of  $F$  at  $v$  and if  $v$  is finite, we write  $k_v$  for its residue field at  $v$ .
- If  $F$  is a local field we write  $\mathcal{O}_F$  for its ring of integers.
- For any field  $F$ , we write  $\overline{F}$  for a fixed algebraic closure of  $F$ .
- We write  $\mathbb{A}$  for the ring of adèles and  $\mathbb{A}_f$  the ring of finite adèles. If  $p$  is a prime,  $\mathbb{A}_f^p$  denotes the finite adèles with trivial  $p$ -component.
- If  $R \rightarrow S$  is a map of algebras and  $X$  is an  $R$ -scheme, we write  $X_S$  for the base change of  $X$  to  $S$ .

- If  $X$  is a scheme, we write  $\mathcal{O}_X$  for its structure sheaf. We write  $\mathbf{H}^\bullet(X, -)$  for the étale cohomology of  $X$ . For any closed subscheme  $Y \subset X$ , we write  $\mathbf{H}_Y^\bullet(X, -)$  for the étale cohomology supported on  $Y$ .

## 2. GEOMETRY OF QUATERNIONIC SHIMURA VARIETIES AND GOREN–OORT STRATA

In this section we recall the results concerning the geometry of quaternionic Shimura varieties and Goren–Oort cycles following [TX16] and [TX19] that we will need.

**2.1. Basics on quaternionic Shimura varieties.** Let  $F$  be a totally real field with  $[F : \mathbb{Q}] = g$  and let  $p > 2$  be a prime which is *unramified* in  $F$ . We are mainly interested in the case when  $p$  is inert in  $F$ ; this is the case considered in [TX16] and [LT20]. However, we will sometimes need to consider the reduction mod  $\ell$  of these Shimura varieties, so we will keep the more general assumption for now. We write  $\Sigma_p$  (resp.  $\Sigma_\infty$ ) for the set of  $p$ -adic places (resp. infinite places). We fix once and for all an isomorphism  $\iota_p : \mathbb{C} \cong \overline{\mathbb{Q}}_p$ , which we will use to identify  $\Sigma_\infty$  with the set of  $p$ -adic embeddings of  $F$ . For  $\mathfrak{p} \in \Sigma_p$  we let  $g_{\mathfrak{p}} := [F_{\mathfrak{p}} : \mathbb{Q}_p]$  and  $\Sigma_{\infty/\mathfrak{p}}$  the set of  $p$ -adic embeddings  $\tau \in \Sigma_\infty$  which induce  $\mathfrak{p}$ . As  $p$  is unramified in  $F$ , the  $p$ -Frobenius  $\sigma$  induces an action on  $\Sigma_{\infty/\mathfrak{p}}$ .

We fix a totally indefinite quaternion algebra  $B$  over  $F$  which is split at all the places above  $p$ . Let  $S \subset \Sigma_\infty \cup \Sigma_p$  be a subset of even cardinality. We set  $S_\infty = S \cap \Sigma_\infty$  and for each  $\mathfrak{p}$  we set  $S_{\infty/\mathfrak{p}} = S \cap \Sigma_{\infty/\mathfrak{p}}$ . We will make the assumption that  $\mathfrak{p} \in S$  only if  $S_{\infty/\mathfrak{p}} = \Sigma_{\infty/\mathfrak{p}}$ .

We write  $B_S$  for the quaternion algebra over  $F$  whose ramification set is precisely the union of  $S$  and the places in  $F$  over which  $B$  ramifies. For each  $\mathfrak{l}$  a place of  $F$  away from the ramification set for  $B_S$ , we fix an isomorphism  $B_S \otimes_F F_{\mathfrak{l}} \cong \mathrm{GL}_2(F_{\mathfrak{l}})$ . We define  $G_S$  to be the reductive group over  $\mathbb{Q}$  such that for any  $\mathbb{Q}$ -algebra  $R$  we have

$$G_S(R) = (B_S \otimes_{\mathbb{Q}} R)^\times.$$

When  $S = \emptyset$ , we simply write  $G$  for the above group. For  $\ell \neq p$  a prime, we have an isomorphism  $G(\mathbb{Q}_\ell) \cong G_S(\mathbb{Q}_\ell)$ . Hence we may fix an isomorphism

$$G(\mathbb{A}_f^p) \cong G_S(\mathbb{A}_f^p).$$

Let  $T \subset S_\infty$  and  $T_{\mathfrak{p}} := S_{\infty/\mathfrak{p}} \cap T$ . We consider the following homomorphism:

$$\begin{aligned} h_{S,T} : \mathcal{S}(\mathbb{R}) \cong \mathbb{C}^\times &\longrightarrow B_S(\mathbb{R}) \cong \mathrm{GL}_2(\mathbb{R})^{\Sigma_\infty - S_\infty} \times \mathbb{H}^T \times \mathbb{H}^{S_\infty - T} \\ x + yi &\longrightarrow ((x + yi)^{\Sigma_\infty - S_\infty}, (x^2 + y^2)^T, 1^{S_\infty - T}). \end{aligned}$$

Then  $G_S$  and the conjugacy class of  $h_{S,T}$  forms a (weak) Shimura datum in the sense of [TX16, §2.2]. We let  $\mathbf{E}_{S,T}$  denote the reflex field which is the subfield of the Galois closure  $\tilde{F}$  of  $F$  in  $\mathbb{C}$  fixed by the subgroup of  $\mathrm{Gal}(\tilde{F}/\mathbb{Q})$  stabilizing  $S_\infty$  and  $T$ . We let  $v$  be the  $p$ -adic place of  $\mathbf{E}_{S,T}$  induced by the embedding  $\mathbf{E}_{S,T} \hookrightarrow \mathbb{C} \cong \overline{\mathbb{Q}}_p$ . We define the compact open subgroup  $K_p := \prod_{\mathfrak{p} \in \Sigma_p} K_{\mathfrak{p}} \subset G_S(\mathbb{Q}_p)$ , where

- $K_{\mathfrak{p}} = \mathrm{GL}_2(\mathcal{O}_{F_{\mathfrak{p}}})$  if  $\mathfrak{p} \notin S$ .
- $K_{\mathfrak{p}} = \mathcal{O}_{B_{\mathfrak{p}}}$  the unique maximal compact of  $B_S \otimes_F F_{\mathfrak{p}}$  if  $\mathfrak{p} \in S$ .

For a sufficiently small compact open subgroup  $K^p \subset G(\mathbb{A}_f^p)$ , we write  $K = K_p K^p$  and let  $\mathrm{Sh}_K(G_{S,T})$  denote the Shimura variety associated to the above data. We use the notation of [TX16] so that  $T$  determines the Deligne homomorphism. It is an algebraic variety over  $\mathbf{E}_{S,T}$  whose complex points are given by

$$\mathrm{Sh}_K(G_{S,T})(\mathbb{C}) = G_S(\mathbb{Q}) \backslash (\mathfrak{h}^\pm)^{\Sigma_\infty - S_\infty} \times G_S(\mathbb{A}_f) / K,$$

where  $\mathfrak{h}^\pm$  is the union of the complex upper and lower half planes. We note that the algebraic variety  $\mathrm{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}$  is independent of  $T$ . However, different choices of  $T$  will give rise to different  $\mathbf{E}_{S,T}$  varieties, see for example [TX19, p1563]. We also point out the abuse of notation here, where the compact open subgroup  $K$  implicitly depends on the choice of  $S$ . When  $S_\infty = \Sigma_\infty$ ,  $\mathrm{Sh}_K(G_{S,T})(\overline{\mathbb{Q}})$  is a discrete set and the action of  $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbf{E}_{S,T})$  can be described explicitly as in [LT20, §2.1].

We also set

$$\mathrm{Sh}_{K_p}(G_{S,T}) := \lim_{\leftarrow K^p} \mathrm{Sh}_K(G_{S,T})$$

and we write  $\mathrm{Sh}_{K_p}^\circ(G_{S,T})$  for the neutral connected component of  $\mathrm{Sh}_{K_p}(G_{S,T})_{\overline{\mathbb{Q}}}$ . This is the component containing the image of the point

$$(i^{\Sigma_\infty - S_\infty}, 1) \in (\mathfrak{h}^\pm)^{\Sigma_\infty - S_\infty} \times G_S(\mathbb{A}_f).$$

**2.2. Unitary Shimura varieties.** In this section we define an auxiliary unitary Shimura variety which is of PEL type in order to define integral models for  $\mathrm{Sh}_K(G_{S,T})$ .

Let  $E/F$  be a CM-extension such that the following two conditions are satisfied:

- (1)  $E/F$  is inert at every place that  $B_S$  is ramified.
- (2) For  $\mathfrak{p} \in \Sigma_p$ ,  $E/F$  is split at  $\mathfrak{p}$  if  $\mathfrak{p} \notin S$  and  $\Sigma_\infty - S_{\infty/\mathfrak{p}}$  is even, and is inert if  $\mathfrak{p} \notin S$  and  $\Sigma_\infty - S_{\infty/\mathfrak{p}}$  is odd.

Let  $\Sigma_{E,\infty}$  denote the set of complex embeddings of  $E$ , which we identify with the set of  $p$ -adic embeddings via  $\iota_p$ . For  $\tilde{\tau} \in \Sigma_{E,\infty}$ , we let  $\tilde{\tau}^c$  denote its complex conjugate. For  $\mathfrak{p} \in \Sigma_p$ , let  $\Sigma_{E,\infty/\mathfrak{p}}$  denote the subset of  $p$ -adic embeddings of  $E$  inducing  $\mathfrak{p}$ . Similarly for  $\mathfrak{q}$  a  $p$ -adic place of  $E$ , we let  $\Sigma_{E,\infty/\mathfrak{q}}$  denote the set of  $p$ -adic embeddings of  $E$  inducing  $\mathfrak{q}$ .

We choose a subset  $\tilde{S}_\infty \subset \Sigma_{E,\infty}$  satisfying the property that for each  $\mathfrak{p} \in \Sigma_p$ , the natural restriction map  $\Sigma_{E,\infty/\mathfrak{p}} \rightarrow \Sigma_{\infty/\mathfrak{p}}$  induces a bijection  $\tilde{S}_{\infty/\mathfrak{p}} \xrightarrow{\sim} S_{\infty/\mathfrak{p}}$ , where  $\tilde{S}_{\infty/\mathfrak{p}} = \tilde{S}_\infty \cap \Sigma_{E,\infty/\mathfrak{p}}$

For each  $\tilde{\tau} \in \Sigma_{E,\infty}$ , the choice of  $\tilde{S}_\infty$  determines a collection of numbers  $s_{\tilde{\tau}} \in \{0, 1, 2\}$  given by:

$$(2.2.1) \quad s_{\tilde{\tau}} = \begin{cases} 0 & \text{if } \tilde{\tau} \in \tilde{S}_\infty \\ 2 & \text{if } \tilde{\tau}^c \in \tilde{S}_\infty \\ 1 & \text{otherwise} \end{cases}.$$

Let  $\tilde{S} = (S, \tilde{S}_\infty)$  and  $T_E := \mathrm{Res}_{E/\mathbb{Q}}(\mathbb{G}_m)$ . We let  $K_{E,p} \subset T_E(\mathbb{Q}_p)$  denote the compact open  $(\mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_p)^\times$ . We define the homomorphism

$$h_{E,\tilde{S},T} : \mathbb{S}(\mathbb{R}) \cong \mathbb{C}^\times \longrightarrow T_E(\mathbb{R}) \cong \prod_{\tau \in \Sigma_\infty} (E \otimes_{F,\tau} \mathbb{R})^\times \cong (\mathbb{C}^\times)^{S_\infty - T} \times (\mathbb{C}^\times)^T \times (\mathbb{C}^\times)^{\Sigma_\infty - S_\infty}$$

$$z = x + yi \longrightarrow ((\bar{z}, \dots, \bar{z}), (z^{-1}, \dots, z^{-1}), (1, \dots, 1)).$$

Here for  $\tau \in S_\infty$ , we identify  $(E \otimes_{F,\tau} \mathbb{R})^\times$  with  $\mathbb{C}^\times$  via the embedding  $\tilde{\tau} : E \rightarrow \mathbb{C}$ , where  $\tilde{\tau} \in \tilde{S}_\infty$  is the unique lift of  $\tau$ . We use the above data to define a Shimura datum for a unitary similitude group which will give rise to a moduli interpretation of the unitary Shimura variety.

Let  $D_S := B_S \otimes_F E$ , which is isomorphic to  $\mathrm{Mat}_2(E)$  by our assumptions on  $E$ . We let  $b \mapsto \bar{b}$  denote the involution on  $D_S$  defined by the product of the canonical involution on  $B_S$  and complex conjugation on  $E/F$ . Fix a totally imaginary element  $\alpha \in E$  such that  $\alpha$  is a  $p$ -adic unit for every place above  $p$ . Choose an element  $\delta \in D_S$  such that  $\bar{\delta} = \delta$  as in [TX16, Lemma 3.8]. We define a new involution of  $D_S$  by  $b \mapsto b^* = \delta^{-1} \bar{b} \delta$ .

We consider  $W := D_S$  as a left  $D_S$ -module of rank 1. It is equipped with the following pairing

$$(2.2.2) \quad \psi : W \times W \rightarrow \mathbb{Q}, \quad \psi(x, y) = \mathrm{Tr}_{E/\mathbb{Q}}(\mathrm{Tr}_{D_S/E}^\circ(\alpha x \delta y^*)),$$

where  $\mathrm{Tr}_{D_S/E}^\circ$  is the reduced trace. It is easy to see this pairing is alternating and non-degenerate. Moreover it satisfies the following property

$$\psi(bx, y) = \psi(x, b^*y), \quad b \in D_S^\times.$$

The unitary similitude group is defined to be

$$G_S'(\mathbb{Q}) := \{g \in \mathrm{GL}_{D_S}(W) \mid \psi(xg, yg) = c(g)\psi(x, y), \text{ for some } c(g) \in \mathbb{Q}^\times\}$$

which arises as the  $\mathbb{Q}$ -points of a reductive group  $G'_S$  over  $\mathbb{Q}$ . We may also describe this group in the following way. Since  $\mathrm{GL}_{D_S}(W) \cong D_S^\times$ , we have  $g \in D_S^\times$  lies in  $G'_S(\mathbb{Q})$  if and only if

$$\mathrm{Tr}_{E/\mathbb{Q}}(\mathrm{Tr}_{D_S/E}^\circ(\alpha x g \delta g^* y^*)) = c(g) \mathrm{Tr}_{E/\mathbb{Q}}(\mathrm{Tr}_{D_S/E}^\circ(\alpha x \delta y^*)), \quad \forall x, y \in D_S.$$

This identity is equivalent to  $g \delta g^* = c(g) \delta$ , i.e.  $g \bar{g} = c(g) \in \mathbb{Q}^\times$ . Thus

$$G'_S(\mathbb{Q}) = \{g = (b, t) \in B_S^\times \times^{F^\times} E^\times \mid \nu_S(b) \mathrm{Nm}_{E/F}(t) \in \mathbb{Q}^\times\}$$

where  $\nu_S : B_S \rightarrow F$  is the reduced norm. Here  $B_S^\times \times^{F^\times} E^\times$  denotes the quotient of  $B_S^\times \times E^\times$  by the central embedding  $F^\times \rightarrow B_S^\times \times E^\times$  given by  $x \mapsto (x, x^{-1})$ .

Let  $T_F$  denote the torus  $\mathrm{Res}_{F/\mathbb{Q}} \mathbb{G}_m$ . We define the group

$$G''_S := G'_S \times^{T_F} T_E.$$

Applying the above to points valued in a  $\mathbb{Q}$ -algebra, we see that  $G'_S$  is identified with the subgroup of  $G''_S$  corresponding to the preimage of  $\mathbb{G}_m \subset T_F$  under the map

$$N : G''_S \rightarrow T_F, \quad (b, t) \mapsto \nu_S(b) \mathrm{Nm}_{E/F}(t).$$

We now let  $h''_S : \mathbb{S} \rightarrow G''_S$  denote the morphism induced by  $(h_{S,T}, h_{E,\mathbb{S},T})$ ; it is independent of  $T$ . The image of  $h''_S$  lies in  $G'_S$  and we let  $h'_S$  denote the induced map. Let  $K''_p \subset G''_S(\mathbb{Q}_p)$  be the compact open subgroup given by the image of  $K_p \times K_{E,p}$  and  $K'_p = G'_S(\mathbb{Q}_p) \cap K''_p$ . For sufficiently small compact open subgroups  $K''^p \subset G''_S(\mathbb{A}_f^p)$  and  $K'^p \subset G'_S(\mathbb{A}_f^p)$ , we obtain Shimura varieties  $\mathrm{Sh}_{K''}(G''_S), \mathrm{Sh}_{K'}(G'_S)$  where  $K'' = K''_p K''^p$  and  $K' = K'_p K'^p$ . We also set

$$\mathrm{Sh}_{K'_p}(G'_S) := \lim_{\leftarrow K'^p} \mathrm{Sh}_{K'}(G'_S), \quad \mathrm{Sh}_{K''_p}(G''_S) := \lim_{\leftarrow K''^p} \mathrm{Sh}_{K''}(G''_S).$$

Let  $\mathbf{E}_S$  denote the common reflex field of these Shimura varieties, which is a subfield of  $\mathbb{C}$ . The isomorphism  $\iota_p : \overline{\mathbb{Q}}_p \cong \mathbb{C}$  induces a  $p$ -adic place  $\tilde{v}$  of  $\mathbf{E}_S$ . We let  $\mathrm{Sh}_{K''_p}(G''_S)^\circ$  (resp.  $\mathrm{Sh}_{K'_p}(G'_S)^\circ$ ) denote the neutral geometric connected component of  $\mathrm{Sh}_{K''_p}(G''_S)_{\overline{\mathbb{Q}}_p}$  (resp.  $\mathrm{Sh}_{K'_p}(G'_S)_{\overline{\mathbb{Q}}_p}$ ). Then both  $\mathrm{Sh}_{K''_p}(G''_S)_{\overline{\mathbb{Q}}_p}^\circ$  and  $\mathrm{Sh}_{K'_p}(G'_S)_{\overline{\mathbb{Q}}_p}^\circ$  can be descended to  $\mathbb{Q}_p^{\mathrm{ur}}$ .

We have the following diagram of groups

$$G_S \leftarrow G_S \times T_E \rightarrow G''_S \leftarrow G'_S$$

compatible with Deligne homomorphisms and such that the induced maps on the derived and adjoint groups are isomorphisms. By [TX16, Corollary 2.17] this induces isomorphisms

$$\mathrm{Sh}_{K_p}(G_S)_{\overline{\mathbb{Q}}_p}^\circ \xleftarrow{\sim} \mathrm{Sh}_{K''_p}(G''_S)_{\overline{\mathbb{Q}}_p}^\circ \xrightarrow{\sim} \mathrm{Sh}_{K'_p}(G'_S)_{\overline{\mathbb{Q}}_p}^\circ$$

of neutral geometric connected components. Since Shimura varieties may be constructed from its neutral connected component and the action of Hecke and Galois, we may transfer integral models from one group to the other. See [TX16, §2.11] for the details.

**2.3. Moduli interpretation for unitary Shimura varieties and integral models.** The Shimura variety  $\mathrm{Sh}_{K'}(G'_S)$  is a Shimura variety of PEL type and thus admits an integral model as a moduli space. Recall the  $D_S$ -module  $W$  together with the non-degenerate alternating form  $\psi$  introduced in the last subsection. We also fix some integral PEL data. Let  $\mathcal{O}_{D_S} \subset D_S$  be an order which is maximal at  $p$  and  $\Lambda \subset W$  an  $\mathcal{O}_{D_S}$ -lattice such that  $\psi(\Lambda, \Lambda) \subset \mathbb{Z}$  and  $\Lambda \otimes_{\mathbb{Z}} \mathbb{Z}_p$  is self-dual. Let  $K'^p \subset G'_S(\mathbb{A}_f^p)$  be a sufficiently small compact open subgroup stabilizing  $\Lambda \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^p$ . We consider the moduli functor  $\underline{\mathrm{Sh}}_{K'}(G'_S)$  that associates to each  $\mathcal{O}_{\mathbf{E}_{\mathbb{S}, \tilde{v}}}$ -scheme  $S$  the set of isomorphism classes of triples  $(A, \iota, \lambda, \varepsilon_{K'^p})$  where:

- $A$  is an abelian scheme over  $S$  of dimension  $4[F : \mathbb{Q}]$ .



•  $\iota : \mathcal{O}_{D_S} \hookrightarrow \text{End}_S(A)$  is an embedding such that the induced action of  $a \in \mathcal{O}_{D_S}$  on  $\text{Lie}(A/S)$  satisfies

$$(2.3.1) \quad \det(T - \iota(a) | \text{Lie}(A/S)) = \prod_{\tilde{\tau} \in \Sigma_{E, \infty}} (T - \tilde{\tau}(a))^{2s_{\tilde{\tau}}}.$$

•  $\lambda : A \rightarrow A^\vee$  is a polarization such that:

-The Rosati involution on  $\text{End}_S(A)$  defined by  $\lambda$  induces the involution  $b \mapsto b^*$  on  $\mathcal{O}_{D_S}$ .

-If  $\mathfrak{p} \notin S$ ,  $\lambda$  induces an isomorphism of  $p$ -divisible groups  $A[\mathfrak{p}^\infty] \cong A^\vee[\mathfrak{p}^\infty]$ .

-If  $\mathfrak{p} \in S$ ,  $(\ker \lambda)[\mathfrak{p}^\infty]$  is a finite flat group scheme contained in  $A[\mathfrak{p}]$  of rank  $p^{4g_p}$  and the cokernel of the map  $\lambda_* : H_1^{\text{dR}}(A/S) \rightarrow H_1^{\text{dR}}(A^\vee/S)$  is a locally free module of rank 2 over  $\mathcal{O}_S \otimes_{\mathbb{Z}_p} \mathcal{O}_E/\mathfrak{p}$ . Here  $H_1^{\text{dR}}(A/S)$  denotes the relative de Rham homology.

•  $\varepsilon_{K'^p}$  is a  $K'^p$ -level structure, i.e. a  $K'^p$ -orbit of isomorphisms

$$\varepsilon_{K'^p} : \Lambda \otimes_{\mathbb{Z}} \widehat{\mathbb{Z}}^p \cong \widehat{T}^p A$$

which respects the action of  $\mathcal{O}_{D_S}$  on both sides and such that the pairing  $\psi$  is compatible with the  $\lambda$ -Weil pairing for some choice of isomorphism  $\widehat{\mathbb{Z}}^p \cong \widehat{\mathbb{Z}}^p(1)$ . Here  $\widehat{T}^p A := \varprojlim_{\leftarrow p^n} A[n]$  denotes the prime-to- $p$  Tate module of  $A$ .

By [TX16, Theorem 3.14], the moduli problem  $\underline{\text{Sh}}_{K'}(G'_S)$  is representable by a smooth quasi-projective scheme over  $\mathcal{O}_{\mathbb{E}_{\tilde{S}, \tilde{v}}}$ , and its generic fiber is identified with  $\text{Sh}_{K'}(G'_S)_{\mathbb{E}_{\tilde{S}, \tilde{v}}}$ . Moreover it is an integral canonical model for  $\text{Sh}_{K'}(G'_S)$  in the sense of [TX16, §2.4]; see also [Mil92]. We write

$$\underline{\text{Sh}}_{K'_p}(G'_S) := \varprojlim_{\leftarrow K'^p} \underline{\text{Sh}}_{K'_p K'^p}(G'_S).$$

Taking the closure of  $\text{Sh}_{K'_p}(G'_S)^\circ$  in  $\underline{\text{Sh}}_{K'_p}(G'_S) \otimes_{\mathcal{O}_{\mathbb{E}_{\tilde{S}, \tilde{v}}}} \mathbb{Z}_p^{\text{ur}}$  and using Deligne's recipe to transfer across to  $G_S$  we obtain an integral canonical model  $\underline{\text{Sh}}_K(G_{S, \text{T}})$  for  $\text{Sh}_K(G_{S, \text{T}})$ ; see [TX16, §2.11]. We write  $\mathcal{S}_K(G_{S, \text{T}})$  (resp.  $\mathcal{S}_{K'}(G'_S)$ ) for the special fiber of  $\underline{\text{Sh}}_K(G_{S, \text{T}})$  (resp.  $\underline{\text{Sh}}_{K'}(G'_S)$ ) over  $k_v$  (resp.  $k_{\tilde{v}}$ ). We let  $k_0$  be any finite field containing all the residue fields of the  $p$ -adic places of  $E$ .

We fix an isomorphism  $\mathcal{O}_{D_{S, p}} := \mathcal{O}_{D_S} \otimes_{\mathbb{Z}} \mathbb{Z}_p \cong M_{2 \times 2}(\mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_p)$  and let  $e$  denote the idempotent of  $M_2(\mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_p)$  given by

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Suppose  $M$  is a module with an action of  $\mathcal{O}_{D_{S, p}}$ . We write  $M^\circ$  for the sub  $\mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_p$ -module  $eM \subset M$ .

Now suppose  $k$  is a perfect field of characteristic  $p$  containing all residue fields at the  $p$ -adic places of  $E$ . Let  $(A, \iota, \lambda, \varepsilon_{K'^p})$  be the data corresponding to an  $S$ -point of  $\text{Sh}_{K'}(G'_S)$  where  $S$  is a  $k_{\tilde{v}}$ -scheme. Then we have an exact sequence:

$$0 \longrightarrow \omega_{A^\vee}^\circ \longrightarrow H_1^{\text{dR}}(A/S)^\circ \longrightarrow \text{Lie}(A)^\circ \longrightarrow 0.$$

Here  $\omega_{A^\vee}$  is the module of invariant differential forms for the dual abelian variety  $A^\vee$  and  $H_1^{\text{dR}}(A/S)$  is the first relative de Rham homology; see [TX16, §4.1]. For  $\tilde{\tau} \in \Sigma_{E, \infty}$ , we use a subscript  $\tilde{\tau}$  to denote the subspace over which  $\mathcal{O}_{E, p} := \mathcal{O}_E \otimes_{\mathbb{Z}} \mathbb{Z}_p$  acts via  $\tilde{\tau}$ . Then the above induces an exact sequence

$$(2.3.2) \quad 0 \longrightarrow \omega_{A^\vee, \tilde{\tau}}^\circ \longrightarrow H_1^{\text{dR}}(A/S)_{\tilde{\tau}}^\circ \longrightarrow \text{Lie}(A)_{\tilde{\tau}}^\circ \longrightarrow 0$$

and the dimensions of these three factors are  $2 - s_{\tilde{\tau}}$ , 2 and  $s_{\tilde{\tau}}$  respectively. In particular, for  $\tau \in \Sigma_\infty - S_\infty$ , and  $\tilde{\tau}$  a lift of  $\tau$ ,  $\omega_{A^\vee, \tilde{\tau}}^\circ$  is a line bundle over  $S$ .

For simplicity, we will write  $X'$  for  $\mathcal{S}_{K'}(G'_S)_{k_0}$  and  $(\mathcal{A}', \iota, \lambda, \varepsilon_{K'^p})$  the universal abelian scheme over  $X'$  together with polarization and level structure. Recall the Kodaira–Spencer map

$$\text{KS}_{(\mathcal{A}', \iota, \lambda, \varepsilon_{K'^p})} : \omega_{\mathcal{A}'} \otimes_{\mathcal{O}_{X'}} \omega_{\mathcal{A}'^\vee} \rightarrow \Omega_{X'/k_0}^1$$

which induces a map

$$\text{KS}_{(\mathcal{A}', \iota, \lambda, \varepsilon_{K'^p})}^\circ : \omega_{\mathcal{A}'}^\circ \otimes_{\mathcal{O}_{X'}} \omega_{\mathcal{A}'^\vee}^\circ \rightarrow \Omega_{X'/k_0}^1.$$

We write  $\mathcal{M} \subset \omega_{\mathcal{A}'}^\circ \otimes_{\mathcal{O}_{X'} \otimes \mathcal{O}_{E,p}} \omega_{\mathcal{A}'\vee}^\circ$  for the  $\mathcal{O}_{X'}$ -submodule given by the sheafification of the presheaf whose sections are generated by  $\lambda_*(y) \otimes z - \lambda_*(z) \otimes y$ , where  $y, z$  are sections of  $\omega_{\mathcal{A}'\vee}^\circ$ . We write  $\omega_{\mathcal{A}'}^\circ \otimes_{(\mathcal{O}_{X'} \otimes \mathcal{O}_{E,p}, \lambda)} \omega_{\mathcal{A}'\vee}^\circ$  for the quotient of  $\omega_{\mathcal{A}'}^\circ \otimes_{\mathcal{O}_{X'} \otimes \mathcal{O}_{E,p}} \omega_{\mathcal{A}'\vee}^\circ$  by  $\mathcal{M}$ .

**Proposition 2.3.1.** *The map  $\mathrm{KS}_{(\mathcal{A}', \iota, \lambda, \varepsilon_{K^p})}^\circ$  factors as*

$$\omega_{\mathcal{A}'}^\circ \otimes_{\mathcal{O}_{X'}} \omega_{\mathcal{A}'\vee}^\circ \rightarrow \omega_{\mathcal{A}'}^\circ \otimes_{(\mathcal{O}_{X'} \otimes \mathcal{O}_{E,p}, \lambda)} \omega_{\mathcal{A}'\vee}^\circ \rightarrow \Omega_{X'/k_0}^1$$

and induces an isomorphism  $\omega_{\mathcal{A}'}^\circ \otimes_{(\mathcal{O}_{X'} \otimes \mathcal{O}_{E,p}, \lambda)} \omega_{\mathcal{A}'\vee}^\circ \cong \Omega_{X'/k_0}^1$ .

*Proof.* This follows from [Lan13, Proposition 2.3.5.2] applied to the universal abelian scheme over  $X'$ .  $\square$

**Corollary 2.3.2.** *The map  $\mathrm{KS}_{(\mathcal{A}', \iota, \lambda, \varepsilon_{K^p})}^\circ$  induces an isomorphism of line bundles*

$$\bigotimes_{\tau \in \Sigma_\infty - S_\infty} \omega_{\mathcal{A}'\vee, \tilde{\tau}}^\circ \otimes_{\mathcal{O}_{X'}} \omega_{\mathcal{A}'\vee, \tilde{\tau}^c}^\circ \cong \Omega_{X'/k_0}^{|\Sigma_\infty - S_\infty|}$$

where  $\tilde{\tau}$  is a lift of  $\tau$ .

*Proof.* We have an isomorphism

$$\omega_{\mathcal{A}'}^\circ \otimes_{\mathcal{O}_{X'} \otimes \mathcal{O}_{E,p}} \omega_{\mathcal{A}'\vee}^\circ \cong \bigoplus_{\tilde{\tau} \in \Sigma_{E,\infty}} \omega_{\mathcal{A}', \tilde{\tau}}^\circ \otimes_{\mathcal{O}_{X'}} \omega_{\mathcal{A}'\vee, \tilde{\tau}}^\circ.$$

Now for  $\tau \in S_\infty$  and  $\tilde{\tau} \in \Sigma_{E,\infty}$  a lift of  $\tau$ , either  $\omega_{\mathcal{A}', \tilde{\tau}}^\circ$  or  $\omega_{\mathcal{A}'\vee, \tilde{\tau}}^\circ$  has rank 0.

For  $\tilde{\tau} \in \Sigma_{E,\infty}$  a lift of  $\tau \in \Sigma_\infty - S_\infty$ , the polarization  $\lambda$  induces an isomorphism  $\omega_{\mathcal{A}', \tilde{\tau}}^\circ \cong \omega_{\mathcal{A}'\vee, \tilde{\tau}^c}^\circ$  compatible with the action of  $\mathcal{O}_{D_S}$  and hence induces an isomorphism  $\omega_{\mathcal{A}', \tilde{\tau}}^\circ \cong \omega_{\mathcal{A}'\vee, \tilde{\tau}^c}^\circ$ . Thus Proposition 2.3.1 implies we have an isomorphism

$$\bigoplus_{\tau \in \Sigma_\infty - S_\infty} \omega_{\mathcal{A}'\vee, \tilde{\tau}}^\circ \otimes_{\mathcal{O}_{X'}} \omega_{\mathcal{A}'\vee, \tilde{\tau}^c}^\circ \cong \Omega_{X'/k_0}^1,$$

where  $\tilde{\tau}$  is a lift of  $\tau$ . Taking the top exterior power gives the result.  $\square$

Now let  $X$  denote the special fiber  $\mathcal{S}_K(G_{S,T})_{k_0}$ . Using [TX16, Corollary 2.13], we may transfer the vector bundles  $\omega_{\mathcal{A}'\vee, \tilde{\tau}}^\circ$  from  $X'$  to  $X$ . For each  $\tau \in \Sigma_\infty - S_\infty$ , we write  $\omega_\tau$  to be the line bundle on  $X$  coming from  $\omega_{\mathcal{A}'\vee, \tilde{\tau}}^\circ$  on  $X'$  for some choice of lift  $\tilde{\tau}$  of  $\tau$  which we now fix. Then  $\omega_\tau$  is independent of the choice of lift  $\tilde{\tau}$  up to a torsion element in the Picard group of  $X$  by [TX16, Lemma 6.2]. For any line bundle  $\mathcal{L}$  on  $X$ , we write  $[\mathcal{L}]$  for the image of  $\mathcal{L}$  in the rational Picard group. The next corollary follows immediately from Corollary 2.3.2.

**Corollary 2.3.3.** *There exists an isomorphism*

$$\left[ \bigotimes_{\tau \in \Sigma_\infty - S_\infty} \omega_\tau^2 \right] \cong \left[ \Omega_{X/k_0}^{|\Sigma_\infty - S_\infty|} \right]$$

in the rational Picard group of  $X$ .  $\square$

**2.4. Goren–Oort divisors and Goren–Oort cycles.** We now let  $k_0$  be the smallest subfield of  $\overline{\mathbb{F}}_p$  containing the residue fields of the  $p$ -adic places of  $E$ . Then  $k_0 = \mathbb{F}_{p^h}$ , where  $h$  is the least common multiple of  $\{(1 + g_p - 2[g_p/2])g_p \mid \mathfrak{p} \in \Sigma_\infty\}$ . In this section we will recall the description of Goren–Oort divisors and Goren–Oort cycles in  $\mathcal{S}_K(G_{S,T})_{k_0}$  obtained in [TX16] and [TX19]. As before  $S \subset \Sigma_p \cup \Sigma_\infty$  is a set of even cardinality. For each  $\tau \in \Sigma_\infty - S_\infty$ , [TX16] defines a Goren–Oort divisor  $\mathcal{S}_K(G_{S,T})_{k_0, \tau} \subset \mathcal{S}_K(G_{S,T})_{k_0}$  by transferring over a certain divisor on an auxiliary unitary Shimura variety. We briefly recall its construction.

Let  $(A, \iota, \lambda, \varepsilon_{K^p})$  be an  $S$ -point of  $\mathcal{S}_{K'}(G'_S)$  for  $S$  a  $k_0$ -scheme. For  $\tilde{\tau} \in \Sigma_{E,\infty}$ , we define the *essential Verschiebung*

$$V_{\mathrm{es}, \tilde{\tau}} : H_1^{\mathrm{dR}}(A/S)_{\tilde{\tau}}^\circ \rightarrow H_1^{\mathrm{dR}}(A^{(p)}/S)_{\tilde{\tau}}^\circ \cong H_1^{\mathrm{dR}}(A/S)_{\sigma^{-1}\tilde{\tau}}^{\circ, (p)}$$

to be the usual Verschiebung if  $s_{\sigma^{-1}\tilde{\tau}} = 0$  or 1, and to be the inverse of the usual Frobenius if  $s_{\sigma^{-1}\tilde{\tau}} = 2$ . For every integer  $n \geq 1$  we define

$$V_{\text{es}}^n : H_1^{\text{dR}}(A/S)_{\tilde{\tau}}^{\circ} \rightarrow H_1^{\text{dR}}(A^{(p^n)}/S)_{\tilde{\tau}}^{\circ} \cong H_1^{\text{dR}}(A/S)_{\sigma^{-n}\tilde{\tau}}^{\circ, (p^n)}$$

to be the  $n^{\text{th}}$ -iteration of the essential Verschiebung.

Now let  $\tau \in \Sigma_{\infty} - S_{\infty}$  and  $\tilde{\tau} \in \Sigma_{E, \infty}$  a lift of  $\tau$ . We define the integer  $n_{\tau}$  to be the smallest integer such that  $\sigma^{-n_{\tau}}\tau \in \Sigma_{\infty} - S_{\infty}$ . Then the restriction of  $V_{\text{es}}^{n_{\tau}}$  to  $\omega_{A^{\vee}, \tilde{\tau}}^{\circ}$  defines a map

$$h_{\tilde{\tau}}(A) : \omega_{A^{\vee}, \tilde{\tau}}^{\circ} \rightarrow \omega_{A^{\vee}, \sigma^{-n_{\tau}}\tilde{\tau}}^{\circ, (p^{n_{\tau}})} \cong (\omega_{A^{\vee}, \sigma^{-n_{\tau}}\tilde{\tau}}^{\circ})^{\otimes p^{n_{\tau}}}.$$

Applying this to the universal abelian variety  $\mathcal{A}'$  over  $\mathcal{S}_{K'}(G'_{\mathbb{S}})_{k_0}$ , we obtain a global section

$$h_{\tilde{\tau}} \in \Gamma(\mathcal{S}_{K'}(G'_{\mathbb{S}})_{k_0}, (\omega_{\mathcal{A}'^{\vee}, \sigma^{-n_{\tau}}\tilde{\tau}}^{\circ})^{\otimes p^{n_{\tau}}} \otimes (\omega_{\mathcal{A}'^{\vee}, \tilde{\tau}}^{\circ})^{\otimes -1}).$$

We call this the  $\tau^{\text{th}}$  partial Hasse-invariant. We let  $\mathcal{S}_{K'}(G'_{\mathbb{S}})_{k_0, \tilde{\tau}} \subset \mathcal{S}_{K'}(G'_{\mathbb{S}})_{k_0}$  denote the vanishing locus of  $h_{\tilde{\tau}}$ , and we let  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, \tilde{\tau}}$  be the corresponding divisor on  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0}$ . By [TX16, Lemma 4.5],  $\mathcal{S}_{K'}(G'_{\mathbb{S}})_{k_0, \tilde{\tau}}$ , and hence  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, \tilde{\tau}}$ , is independent of the lifting  $\tilde{\tau}$  of  $\tau$ . We may thus write  $\mathcal{S}_{K'}(G'_{\mathbb{S}})_{k_0, \tau}$  and  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, \tau}$  for these divisors respectively. The divisor  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, \tau}$  is known as the *Goren–Oort divisor corresponding to  $\tau$* . It can also be described as the vanishing locus of a certain section of the line bundle of  $\omega_{\sigma^{-n_{\tau}}\tau}^{\otimes p^{n_{\tau}}} \otimes \omega_{\tau}^{-1}$  on  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0}$ .

For  $J \subset \Sigma_{\infty} - S_{\infty}$ , we set

$$\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, J} := \bigcap_{\tau \in J} \mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, \tau}.$$

The closed subvarieties  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, J}$  as  $J$  runs over subsets of  $\Sigma_{\infty} - S_{\infty}$  give the *Goren–Oort stratification* of  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})$ .

To state the main structure theorem of [TX16], we introduce the following notation. For  $\tau \in \Sigma_{\infty} - S_{\infty}$ , we write  $\tau^{-}$  for  $\sigma^{-n_{\tau}}\tau$  and we let  $\mathfrak{p}$  denote the  $p$ -adic place induced by  $\tau$ . We write  $\mathbb{T}_{\tau} = \mathbb{T} \cup \{\tau\}$  and define

$$S_{\tau} = \begin{cases} S_{\infty/\mathfrak{p}} \cap \{\tau, \tau^{-}\} & \text{if } S_{\infty/\mathfrak{p}} \cup \{\tau\} \neq \Sigma_{\infty/\mathfrak{p}} \\ S \cup \{\tau, \mathfrak{p}\} & \text{if } S_{\infty/\mathfrak{p}} \cup \{\tau\} = \Sigma_{\infty/\mathfrak{p}} \end{cases}.$$

**Theorem 2.4.1** ([TX16, Theorem 5.2], [LT20, Proposition 3.9]). *Under the above notations, there is a morphism of  $\mathbb{F}_p$ -varieties*

$$\pi_{\tau} : \mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{\mathbb{F}_p, \tau} \rightarrow \mathcal{S}_K(G_{S_{\tau}, \mathbb{T}_{\tau}})_{\mathbb{F}_p}$$

*equivariant for the prime-to- $p$  Hecke action, such that*

- (1) *if  $S_{\infty/\mathfrak{p}} \cup \{\tau\} \neq \Sigma_{\infty/\mathfrak{p}}$ ,  $\pi_{\tau}$  is a  $\mathbb{P}^1$ -fibration which descends to a morphism  $\mathcal{S}_K(G_{\mathbb{S}, \mathbb{T}})_{k_0, \tau} \rightarrow \mathcal{S}_K(G_{S_{\tau}, \mathbb{T}_{\tau}})_{k_0}$  of  $k_0$ -varieties.*
- (2) *if  $S_{\infty/\mathfrak{p}} \cup \{\tau\} = \Sigma_{\infty/\mathfrak{p}}$ ,  $\pi_{\tau}$  is an isomorphism.*

Note that in case (2), the level structure at  $\mathfrak{p}$  needs to be modified.

We now recall the framework for the construction of Goren–Oort cycles. These cycles are parameterized by certain combinatorial objects called periodic semi-meanders whose definition we now recall. We refer to [LT20, §3.4] for more details.

For  $\mathfrak{p}$  a place of  $F$  above  $p$ , let  $d_{\mathfrak{p}}(S) := g_{\mathfrak{p}} - |S_{\infty/\mathfrak{p}}|$ . We consider the cylinder  $C := \{x^2 + y^2 = 1\}$  in 3-dimensional space. Let  $\Sigma_{\infty/\mathfrak{p}} = \{\tau_0, \dots, \tau_{g_{\mathfrak{p}}-1}\}$  where  $\sigma(\tau_j) = \tau_{j+1}$  for  $j \in \mathbb{Z}/g_{\mathfrak{p}}\mathbb{Z}$ . We use the  $\tau_j$  to label the points on the  $xy$ -plane by identifying  $\tau_j$  with the point  $(\cos \frac{2\pi j}{g_{\mathfrak{p}}}, \sin \frac{2\pi j}{g_{\mathfrak{p}}})$ . If  $\tau_j \in S_{\infty/\mathfrak{p}}$  we put a  $+$  at the point  $\tau_j$ , otherwise we put a node. We call the series of nodes and  $+$ 's the *band* associated to  $S_{\infty/\mathfrak{p}}$ .

A periodic semi-meander for  $S_{\infty/\mathfrak{p}}$  is a series of arcs and semi-lines on  $C$  (an arc connects two nodes and a semi-line connects a node and  $+\infty$ ) satisfying the following properties:

- (1) All arcs and semi-lines lie above the band.

- (2) Every node is the end-point of an arc or a semi-line.
- (3) There are no intersection points among the arcs and semi-lines.

We identify two periodic semi-meanders if one can be continuously deformed into the other. We write  $r$  for the number of arcs in a periodic semi-meander and  $d_{\mathfrak{p}}(S) - r$  is the defect. We write  $\mathfrak{B}(S_{\infty/\mathfrak{p}}, r)$  for the set of periodic semi-meanders with  $r$  arcs.

For any  $\mathfrak{a} \in \mathfrak{B}(S_{\infty/\mathfrak{p}}, r)$  we define the sets  $S_{\mathfrak{a}}$  and  $T_{\mathfrak{a}}$  to be

$$(2.4.1) \quad S_{\mathfrak{a}} := S \cup \{\tau \in \Sigma_{\infty} \mid \tau \text{ is an end-point of an arc in } \mathfrak{a}\}$$

$$(2.4.2) \quad T_{\mathfrak{a}} := T \cup \{\tau \in \Sigma_{\infty} \mid \tau \text{ is the right end-point of an arc in } \mathfrak{a}\}.$$

To any periodic semi-meander  $\mathfrak{a} \in \mathfrak{B}(S_{\infty/\mathfrak{p}}, r)$ , [LT20, §3.5] constructs the *Goren–Oort* cycle

$$Z_{S,T}(\mathfrak{a}) \hookrightarrow \mathcal{S}_K(G_{S,T})_{k_0}.$$

corresponding to  $\mathfrak{a}$  using the method of [TX19, §3.7]. The construction is by induction on  $r$  and the resulting cycle is an  $r^{\text{th}}$ -iterated  $\mathbb{P}^1$ -bundle over  $\mathcal{S}_K(G_{S_{\mathfrak{a}}, T_{\mathfrak{a}}})_{k_0}$ . When  $r = 0$ ,  $Z_{S,T}(\mathfrak{a})$  is defined to be  $\mathcal{S}_K(G_{S,T})_{k_0}$ . For  $r \geq 1$ , we say an arc  $\delta$  in  $\mathfrak{a}$  is basic if it does not lie above any other arc. Choose such a basic arc  $\delta$  and write  $\tau$  and  $\tau^-$  for its right and left endpoints respectively. Consider the Goren–Oort divisor  $\mathcal{S}_K(G_{S,T})_{k_0, \tau}$  together with the fibration  $\pi_{\tau} : \mathcal{S}_K(G_{S,T})_{k_0, \tau} \rightarrow \mathcal{S}_K(G_{S_{\tau}, T_{\tau}})_{k_0}$ . Let  $\mathfrak{a}_{\delta} \in \mathfrak{B}(S_{\tau, \infty/\mathfrak{p}}, r - 1)$  denote the periodic semi-meander given by removing the arc  $\delta$  from  $\mathfrak{a}$  and replacing the endpoints with  $+$  signs. By induction hypothesis, we have the cycle  $Z_{S_{\tau}, T_{\tau}}(\mathfrak{a}_{\delta}) \subset \mathcal{S}_K(G_{S_{\tau}, T_{\tau}})_{k_0}$  which is an  $(r - 1)^{\text{th}}$ -iterated  $\mathbb{P}^1$ -bundle over  $\mathcal{S}_K(G_{S_{\mathfrak{a}}, T_{\mathfrak{a}}})_{k_0}$ . The Goren–Oort cycle  $Z_{S,T}(\mathfrak{a})$  is defined to be the preimage of  $Z_{S_{\tau}, T_{\tau}}(\mathfrak{a}_{\delta})$  in  $\mathcal{S}_K(G_{S,T})_{k_0}$  under the projection  $\pi_{\tau}$ . We write

$$\pi_{\mathfrak{a}} : Z_{S,T}(\mathfrak{a}) \rightarrow \mathcal{S}_K(G_{S_{\mathfrak{a}}, T_{\mathfrak{a}}})_{k_0}$$

for the projection map.

The following proposition is clear from the construction and Theorem 2.4.1.

**Proposition 2.4.2.**  *$Z_{S,T}(\mathfrak{a})$  is an  $r^{\text{th}}$ -iterated  $\mathbb{P}^1$ -fibration over the Shimura variety  $\mathcal{S}_K(G_{S_{\mathfrak{a}}, T_{\mathfrak{a}}})_{k_0}$ . Moreover the inclusion  $Z_{S,T}(\mathfrak{a}) \rightarrow \mathcal{S}_K(G_{S,T})_{k_0}$  is equivariant for the prime-to- $p$  Hecke correspondences.*  $\square$

**2.5. Goren–Oort cycles and Shimura surfaces.** We now give a more detailed description of the Goren–Oort cycles which will be used in the construction of the level raising subgroup in motivic cohomology.

For this we will impose the extra assumptions that  $[F : \mathbb{Q}] = g$  is even and that  $p$  is inert in  $F$ ; then  $\mathfrak{p}$  will denote the unique prime above  $p$ . In this case we may take  $k_0 = \mathbb{F}_{p^g}$ . We consider the set  $\mathfrak{B}(\emptyset, g/2 - 1)$  of periodic semi-meanders which is easily seen to have  $\binom{g}{g/2-1}$  elements.

Fix  $\mathfrak{a} \in \mathfrak{B}(\emptyset, g/2 - 1)$ . Then we have constructed the Goren–Oort cycle  $Z_{\emptyset}(\mathfrak{a}) \subset \mathcal{S}_K(G)_{\mathbb{F}_{p^g}}$  and the projection

$$\pi_{\mathfrak{a}} : Z_{\emptyset}(\mathfrak{a}) \rightarrow \mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, T_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}$$

which is  $(g/2 - 1)$ -iterated  $\mathbb{P}^1$ -fibration; here  $T_{\mathfrak{a}}$  is defined as in (2.4.2) with  $T = \emptyset$ . Since  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, T_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}$  is the special fiber of a Shimura surface, the Goren–Oort cycle  $Z_{\emptyset}(\mathfrak{a})$  has dimension  $g/2 + 1$ .

We write  $\Sigma_{\infty} \setminus \emptyset_{\mathfrak{a}} = \{\tau_i, \tau_j\}$ . Then we may consider the Goren–Oort divisors  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, T_{\mathfrak{a}}})_{\mathbb{F}_{p^g}, \tau_i}$  and  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, T_{\mathfrak{a}}})_{\mathbb{F}_{p^g}, \tau_j}$  of the Shimura surface  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, T_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}$  which are  $\mathbb{P}^1$ -bundles over the discrete Shimura sets  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, \tau_i, T_{\mathfrak{a}}, \tau_i})_{\mathbb{F}_{p^g}}$  and  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, \tau_j, T_{\mathfrak{a}}, \tau_j})_{\mathbb{F}_{p^g}}$  respectively. The  $\overline{\mathbb{F}}_p$ -points of these Shimura sets may be identified (upon fixing a base-point) with

$$G_{\Sigma_{\infty}}(\mathbb{Q}) \backslash G_{\Sigma_{\infty}}(\mathbb{A}_f) / K.$$

We let  $K_0(\mathfrak{p}) \subset G_{\Sigma_{\infty}}(\mathbb{A}_f)$  denote the compact open subgroup  $K_{0, \mathfrak{p}} K^{\mathfrak{p}}$  where  $K_{0, \mathfrak{p}}$  is the standard Iwahori subgroup of  $\text{GL}_2(F_{\mathfrak{p}})$  consisting of matrices in  $\text{GL}_2(\mathcal{O}_{F_{\mathfrak{p}}})$  which reduce to the upper triangular matrices mod  $\mathfrak{p}$ . In [TX16, 3.15], there is a construction of the integral model for the

discrete Shimura set  $\text{Sh}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})$ . We can explicitly describe the  $\overline{\mathbb{F}}_p$ -points of the special fiber  $\mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}}$  and the action of  $\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})$  as follows. We equip the discrete set  $G_{\Sigma_\infty}(\mathbb{Q})\backslash G_{\Sigma_\infty}(\mathbb{A}_f)/K_0(\mathfrak{p})$  with the  $\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})$ -action where the arithmetic  $p^g$ -Frobenius  $\sigma_{p^g}$  acts via multiplication by the central element  $\underline{p}^{-g/2} \in F \otimes_{\mathbb{Q}} \mathbb{A}_f^\times \subset G_{\Sigma_\infty}(\mathbb{A}_f)$ . Here  $\underline{p}^{-g/2}$  is the idele which is  $p^{-g/2}$  at the  $p$ -adic place  $\mathfrak{p}$  and 1 elsewhere. Then there is a bijection

$$\mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p) \cong G_{\Sigma_\infty}(\mathbb{Q})\backslash G_{\Sigma_\infty}(\mathbb{A}_f)/K_0(\mathfrak{p})$$

compatible with the Galois action.

The action of  $\sigma_{p^g}$  on  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p)$  is defined in the same way. We write  $n_{K_\infty}$  for the order of  $\sigma_{p^g}$  acting on  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p)$ . It is easily seen to be independent of the choice of  $\mathfrak{a}$  and  $\tau_i$ .

The following proposition is contained in [TX19, Proposition 2.32].

**Proposition 2.5.1.** (1) *There is an isomorphism*

$$\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}},\mathbb{T}_{\mathfrak{a}}})_{\mathbb{F}_{p^g},\{\tau_i,\tau_j\}} \cong \mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}}$$

of  $\mathbb{F}_{p^g}$ -varieties.

(2) *There is an isomorphism*

$$\eta_{\tau_j} : \mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_j},\mathbb{T}_{\mathfrak{a},\tau_j}})_{\mathbb{F}_{p^g}} \xrightarrow{\sim} \mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}}$$

of  $\mathbb{F}_{p^g}$ -varieties such that the induced diagram

$$\begin{array}{ccc} & \mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}} & \\ \swarrow \pi_{\tau_i} \circ \iota_{\tau_i} & & \searrow \pi_{\tau_j} \circ \iota_{\tau_j} \\ \mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}} & & \mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_j},\mathbb{T}_{\mathfrak{a},\tau_j}})_{\mathbb{F}_{p^g}} \xrightarrow{\eta_{\tau_j}} \mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}} \end{array}$$

is identified with the base change to  $\mathbb{F}_{p^g}$  of the Hecke correspondence for  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})$ .

Here the maps

$$\iota_{\tau_i} : \mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}} \rightarrow \mathcal{S}_K(G_{\emptyset_{\mathfrak{a}},\mathbb{T}_{\mathfrak{a}}})_{\mathbb{F}_{p^g},\tau_i},$$

$$\iota_{\tau_j} : \mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset_{\mathfrak{a},\tau_i},\mathbb{T}_{\mathfrak{a},\tau_i}})_{\mathbb{F}_{p^g}} \rightarrow \mathcal{S}_K(G_{\emptyset_{\mathfrak{a}},\mathbb{T}_{\mathfrak{a}}})_{\mathbb{F}_{p^g},\tau_j}$$

are the natural embeddings coming from (1). □

### 3. HECKE ORBIT CONJECTURE

In this section we prove a version of the Hecke orbit conjecture for the ordinary locus of quaternionic Shimura varieties. The desired result follows from the corresponding statement for the auxiliary unitary Shimura varieties of PEL type constructed in the previous section. The result for these Shimura varieties can then be deduced using the method of [Cha95].

**3.1. Statement of Hecke orbit conjecture.** We keep the notation introduced in §2. We let  $\underline{\text{Sh}}_K(G_{\mathbb{S},\mathbb{T}})$  be the integral model for the quaternionic Shimura variety  $\text{Sh}_K(G_{\mathbb{S},\mathbb{T}})$ . We will assume in this section that  $\mathbb{S} \subsetneq \Sigma_\infty$  so that the compact open subgroup  $K = K_p K^p$  is hyperspecial at  $p$ . Since the case of Hilbert modular varieties has been proved in [Cha95], we also assume  $\text{Sh}_K(G_{\mathbb{S},\mathbb{T}})$ , and hence  $\underline{\text{Sh}}_K(G_{\mathbb{S},\mathbb{T}})$ , is compact, or in other words that the quaternion algebra  $B_{\mathbb{S}}$  is not totally split. By our assumption on  $\mathbb{S}$ ,  $\underline{\text{Sh}}_K(G_{\mathbb{S},\mathbb{T}})$  is not discrete, in other words it arises from a genuine Shimura datum (as opposed to a weak Shimura datum). For ease of notation we will write  $\mathcal{X} := \mathcal{S}_K(G_{\mathbb{S},\mathbb{T}})_{\overline{\mathbb{F}}_p}$  for the geometric special fiber of  $\underline{\text{Sh}}_K(G_{\mathbb{S},\mathbb{T}})$ .

We also write  $\text{Sh}_{K'}(G'_{\mathbb{S}})$  for the unitary Shimura variety of PEL type associated to  $\text{Sh}_K(G_{\mathbb{S},\mathbb{T}})$  by the choice of an imaginary quadratic field  $E/F$  and a subset  $\tilde{\mathbb{S}}_\infty \subset \Sigma_{E,\infty}$  satisfying the conditions

in §2.2. We similarly write  $\mathcal{X}' = \mathcal{S}_{K'}(G'_S)_{\overline{\mathbb{F}}_p}$  for the geometric special fiber of the integral model  $\underline{\text{Sh}}_{K'}(G'_S)$ .

Our formulation of the Hecke orbit conjecture for  $\mathcal{X}'$  differs from the statement in [Cha95]; we need a slightly stronger statement in order to transfer the result to the quaternionic side. More precisely we will consider the orbit under Hecke correspondences coming from the derived group. We let  $G'_{\text{der}}$  denote the derived group of  $G'_S$ . Since  $G'_S$  and  $G_S$  have the same derived group,  $G'_S$  is the reductive group corresponding to  $(B_S^\times)^{\nu_S=1}$ ; in particular  $G'_{\text{der}}$  is simply connected. We write  $T'_{\text{ab}}$  for the quotient of  $G'_S$  by its derived group; it is isomorphic to the subtorus  $\text{Nm}_{E/F}^{-1}(\mathbb{G}_m) \subset T_E$ . We write  $T'_{\text{ab}}(\mathbb{R})^\dagger = \text{Im}(Z'(\mathbb{R}) \rightarrow T'_{\text{ab}}(\mathbb{R}))$ , where  $Z'$  is the center of  $G'_S$ . We let  $T'_{\text{ab}}(\mathbb{Q})^\dagger = T'_{\text{ab}}(\mathbb{R})^\dagger \cap T'_{\text{ab}}(\mathbb{Q})$ .

We write  $\nu' : G'_S \rightarrow T'_{\text{ab}}$  for the quotient map. This induces a bijection

$$(3.1.1) \quad \pi_0(\text{Sh}_{K'}(G'_S)) \cong T'_{\text{ab}}(\mathbb{Q})^\dagger \backslash T'_{\text{ab}}(\mathbb{A}_f) / \nu'(K')$$

where the left hand side is the set of geometric connected components. Since  $\underline{\text{Sh}}_{K'}(G'_S)$  is smooth, there is a bijection

$$\pi_0(\mathcal{X}') \xrightarrow{\sim} T'_{\text{ab}}(\mathbb{Q})^\dagger \backslash T'_{\text{ab}}(\mathbb{A}_f) / \nu'(K')$$

compatible with specialization from (3.1.1). For an element  $\mathfrak{c}'$  on the right hand side, we write  $\mathcal{X}'^{\mathfrak{c}'}$  for the corresponding connected component of  $\mathcal{X}'$ .

We consider the inverse limit scheme

$$\underline{\text{Sh}}_{K'_p}(G'_S) := \varprojlim_{K'^p} \underline{\text{Sh}}_{K'_p K'^p}(G'_S)$$

which has an action by the group  $G'_S(\mathbb{A}_f^p)$ . Let  $x \in \underline{\text{Sh}}_{K'}(G'_S)(\overline{\mathbb{F}}_p) = \mathcal{X}'(\overline{\mathbb{F}}_p)$  and  $\tilde{x} \in \underline{\text{Sh}}_{K'_p}(G'_S)(\overline{\mathbb{F}}_p)$  a lift of  $x$ . We write  $\tilde{Y}'^p(\tilde{x}) \subset \underline{\text{Sh}}_{K'_p}(G'_S)(\overline{\mathbb{F}}_p)$  for the orbit of  $\tilde{x}$  under the group  $G'_{\text{der}}(\mathbb{A}_f^p)$ . Similarly, for  $\ell$  a prime coprime to  $p$ , we write  $\tilde{Y}'_\ell(\tilde{x}) \subset \underline{\text{Sh}}_{K'_p}(G'_S)(\overline{\mathbb{F}}_p)$  for the orbit of  $\tilde{x}$  under the group  $G'_{\text{der}}(\mathbb{Q}_\ell)$ .

**Definition 3.1.1.** The reduced prime-to- $p$  Hecke orbit  $Y'^p(x)$  of  $x$  is defined to be the image of  $\tilde{Y}'^p(\tilde{x})$  in  $\mathcal{X}'(\overline{\mathbb{F}}_p)$ . Similarly the reduced  $\ell$ -power Hecke orbit  $Y'_\ell(x)$  is defined to be the image of  $\tilde{Y}'_\ell(\tilde{x})$  in  $\mathcal{X}'(\overline{\mathbb{F}}_p)$ . We write  $Z'^p(x)$  and  $Z'_\ell(x)$  for the Zariski closure of  $Y'^p(x)$  and  $Y'_\ell(x)$  respectively.

It is clear from the definition that  $Y'^p(x)$  and  $Y'_\ell(x)$  are independent of the choice of  $\tilde{x}$  and that  $Y'_\ell(x) \subset Y'^p(x)$ . When we consider  $Y'_\ell(x)$  we will also assume  $K'$  factors as  $K'_p K'_\ell K'^{p,\ell}$ , where  $K'_\ell \in G'_S(\mathbb{Q}_\ell)$  and  $K'^{p,\ell} \subset G'_S(\mathbb{A}_f^{p,\ell})$  are compact open subgroups; here  $\mathbb{A}_f^{p,\ell}$  denotes the adèles with trivial component at  $p$  and  $\ell$ .

By the main result of [Ham15], there is a stratification of  $\mathcal{X}'$  parametrized by the set  $B(G'_{S, \mathbb{Q}_p}, \{\mu\})$  where  $\{\mu\}$  is the geometric conjugacy class of cocharacters of  $G'_S$  containing the inverse of the Hodge cocharacter associated to the Shimura datum. This is a certain subset of the set of  $\sigma$ -conjugacy classes of  $G'_S(L)$ , where  $L$  is the completion of the maximal unramified extension of  $\mathbb{Q}_p$ ; we refer to §A.2 for the definition. The set  $B(G'_{S, \mathbb{Q}_p}, \{\mu\})$  is equipped with a partial order. It contains a unique maximal element  $[b]^{\text{ord}}$  and a unique minimal element  $[b]^{\text{ss}}$ . We write  $\mathcal{X}'^{\text{ord}}$  and  $\mathcal{X}'^{\text{ss}}$  for the corresponding strata.

*Remark 3.1.2.* In the literature,  $\mathcal{X}'^{\text{ord}}$  is usually known as the  $\mu$ -ordinary locus. If  $\tilde{S}_\infty$  satisfies [LT20, Assumption 2.2], the stratum  $\mathcal{X}'^{\text{ss}}$  coincides with the locus where the universal abelian variety is supersingular. In what follows, we will abuse notation and refer to  $\mathcal{X}'^{\text{ss}}$  as the supersingular locus even when this assumption is not satisfied.

The main result of this section is the following.

**Theorem 3.1.3.** *Let  $\mathfrak{c}' \in T'_{\text{ab}}(\mathbb{Q})^\dagger \backslash T'_{\text{ab}}(\mathbb{A}_f) / \nu'(K')$ . Then for any  $x \in \mathcal{X}'^{\mathfrak{c}'} \cap \mathcal{X}'^{\text{ord}}(\overline{\mathbb{F}}_p)$  we have*

$$Z'^p(x) = \mathcal{X}'^{\mathfrak{c}'}$$

*Remark 3.1.4.* (1) The inclusion  $Z'^p(x) \subset \mathcal{X}'^{c'}$  is clear since  $G'_{\text{der}} = \ker \nu'$ .

(2) Since  $G'_S(\mathbb{A}_f^p)$  acts transitively on the set of connected components of  $\text{Sh}_{K'_p}(G'_S)$  (see for example [Kis10, Lemma 2.2.5]), this theorem implies the prime-to- $p$  Hecke orbit of any  $x \in \mathcal{X}'^{\text{ord}}(\overline{\mathbb{F}}_p)$  is Zariski dense in  $\mathcal{X}'$ . Here the prime-to- $p$  Hecke orbit is defined by replacing the  $G'_{\text{der}}(\mathbb{A}_f^p)$ -orbit with the  $G'_S(\mathbb{A}_f^p)$ -orbit in the definition of  $\tilde{Y}'^p(\tilde{x})$ .

The rest of this section is devoted to the proof of Theorem 3.1.3. We briefly outline the strategy following [Cha95]. The first step is to analyze the formal completion  $\widehat{Z}'^p(x)$  of  $Z'^p(x)$  at a smooth ordinary point. The stability of  $Z'^p(x)$  under Hecke correspondences constrains the possibilities for this formal completion. Indeed, using Moonen's generalization of the Serre–Tate coordinates [Moo04] in this context, one can show that the formal completion of  $\mathcal{X}'$  at an ordinary point has a formal group structure; then  $\widehat{Z}'^p(x)$  is actually a formal subgroup. The second step is to show that  $Z'^p(x)$  contains a supersingular point; this uses the quasi-affineness of the Ekedahl–Oort stratification on  $\mathcal{X}'$  proved in [VW13]. Finally we analyze the formal completion of  $Z'^p(x)$  at a supersingular point  $s \in \mathcal{X}'^{\text{ss}} \cap Z'^p(x)$ ; we use the fact that  $s$  is the specialization of an ordinary point  $x'$  and that  $\mathcal{A}_s$  has a large endomorphism group to show that the formal completion of  $Z'^p(x)$  at  $x'$  is the whole formal neighborhood in  $\mathcal{X}'$ . This implies the result.

A key ingredient of the proof of the Hecke orbit conjecture for Hilbert modular varieties in [Cha95] was an explicit description of the structure of isogeny classes associated to an abelian variety with extra structure. This required an understanding of the endomorphism algebra of these abelian varieties, see [Cha95, Lemma 6]. We will instead use a group theoretic approach using results of [Kis17] which avoids some of the case-by-case analysis of loc. cit.

Let  $k$  be an algebraically closed field of characteristic  $p$ . Let  $x \in \mathcal{X}'(k)$  and  $(\mathcal{A}_x, \iota, \lambda, \varepsilon_{K'^p})$  the associated quadruple. We write  $I_x$  for the reductive group over  $\mathbb{Q}$  such that  $I_x(\mathbb{Q}) := \text{Aut}_{(\mathcal{O}_{D_S}, \lambda)}^0(\mathcal{A}_x)$  the group of quasi-isogenies of  $\mathcal{A}_x$  respecting the  $\mathcal{O}_{D_S}$ -structure and preserving the polarization  $\lambda$  up to a scalar in  $\mathbb{Q}^\times$ .

We would like to apply the results of [Kis17, §2] to our situation. Suppose that  $x \in \mathcal{X}'(\overline{\mathbb{F}}_p)$ . We write  $I_x^{\text{Kis}}$  for the group denoted  $I$  in [Kis17, §2.1]. Then it is easy to see that  $I_x^{\text{Kis}} \cong I_x$ ; indeed in [Kis17] we may take the tensors  $s_\alpha \in \Lambda_{\mathbb{Z}(p)}^\otimes$  to be the classes corresponding to the endomorphisms and polarization. Then the condition defining  $I_x^{\text{Kis}}$  in [Kis17, §2.1] precisely corresponds to the condition defining  $I_x$ . By [Kis17, Corollary 2.3.5], there is a subgroup  $I_0 \subset G'_S$  equipped with an inner twisting  $I_0 \otimes_{\mathbb{Q}} \overline{\mathbb{Q}} \cong I_x \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}$ . The map  $\nu' : G'_S \rightarrow T'_{\text{ab}}$  determines a map  $\nu'_x : I_x \rightarrow T'_{\text{ab}}$  and we let  $I_x^1 \subset I_x$  denote the kernel of  $\nu'_x$ .

For all primes  $\ell \neq p$ , upon fixing a trivialization of  $T_\ell \mathcal{A}_x$ , there is a map  $I_x(\mathbb{Q}) \rightarrow G'_S(\mathbb{Q}_\ell)$  given by the action on the  $\ell$ -adic Tate module. We write  $\mathcal{G}_x := \mathcal{A}_x[p^\infty]$  for the  $p$ -divisible group associated to  $\mathcal{A}_x$  and  $\mathbb{D}(\mathcal{G}_x)$  its Dieudonné module over  $\mathcal{O}_L$ . Then upon fixing an  $\mathcal{O}_L$ -linear bijection

$$(3.1.2) \quad \mathbb{D}(\mathcal{G}_x) \cong \Lambda \otimes_{\mathbb{Z}} \mathcal{O}_L$$

preserving the action of  $\mathcal{O}_{D_S}$  and the pairings on both sides, we obtain an embedding  $I_x(\mathbb{Q}_p) \hookrightarrow G'_S(L)$ . We write  $H_p \subset I_x(\mathbb{Q}_p)$  for the subgroup of elements which stabilize  $\Lambda \otimes_{\mathbb{Z}} \mathcal{O}_L$ ; it is a compact open subgroup of  $I_x(\mathbb{Q}_p)$  which does not depend on the choice of bijection (3.1.2). The following proposition can be deduced in a standard way as in [Cha95, p448].

**Proposition 3.1.5.** *Suppose  $x \in \mathcal{X}'(\overline{\mathbb{F}}_p)$ . There are bijections*

$$(3.1.3) \quad Y'^p(x) \cong I_x^1(\mathbb{Q}) \cap H_p \backslash G'_{\text{der}}(\mathbb{A}_f^p) K'^p / K'^p$$

$$(3.1.4) \quad Y'_\ell(x) \cong I_x^1(\mathbb{Q}) \cap H_p K'^{p,\ell} \backslash G'_{\text{der}}(\mathbb{Q}_\ell) K'_\ell / K'_\ell$$

such that (3.1.3) (resp. (3.1.4)) is equivariant for reduced prime-to- $p$  (resp. reduced  $\ell$ -power) Hecke operators.

*Proof.* Let  $\tilde{x} \in \underline{\text{Sh}}_{K'_p}(G'_S)(\overline{\mathbb{F}}_p)$  be a lift of  $x$ . Then by the argument of [Cha95, p448], the stabilizer of  $\tilde{x}$  in  $G'_S(\mathbb{A}_f^p)$  is identified with  $I_x(\mathbb{Q}) \cap H_p$  and hence the  $G'_S(\mathbb{A}_f^p)$ -orbit of  $\tilde{x}$  can be identified with  $I_x(\mathbb{Q}) \cap H_p \backslash G'_S(\mathbb{A}_f^p)$ , cf. also [Kis17, Proposition 2.1.3]. Since  $G'_{\text{der}}(\mathbb{A}_f^p) \cap I_x(\mathbb{Q}) = I_x^1(\mathbb{Q})$ , the  $G'_{\text{der}}(\mathbb{A}_f^p)$ -orbit of  $\tilde{x}$  is identified with  $I_x^1(\mathbb{Q}) \cap H_p \backslash G'_{\text{der}}(\mathbb{A}_f^p)$  and the description of  $Y'^p(x)$  follows by taking the image in  $\mathcal{X}'$ . The description of  $Y'_\ell(x)$  is proved in a similar way.  $\square$

**3.2. Serre–Tate theory following [Moo04] and analysis at ordinary points.** In this subsection we study the formal neighbourhood of a point in  $\mathcal{X}'^{\text{ord}}$ . We begin with some preliminaries concerning  $p$ -divisible groups with an action of the ring of integers  $\mathcal{O}$  of a finite unramified extension of  $\mathbb{Q}_p$ . Recall we have the integral PEL datum  $(\mathcal{O}_{D_S} \subset D_S, *, \Lambda \subset W, \psi)$ ; we write  $\mathcal{D}$  for the base change of this datum to  $\mathbb{Z}_p$ .

**Definition 3.2.1.** Let  $S$  be an  $\mathcal{O}_{\mathbb{E}_{S,\bar{v}}}$ -scheme. A  $p$ -divisible group with  $\mathcal{D}$ -structure over  $S$  is a triple  $(\mathcal{G}, \lambda, \iota)$  where:

- $\mathcal{G}$  is a  $p$ -divisible group over  $S$  of height  $\dim W$ .
- $\iota : \mathcal{O}_{D_S,p} \rightarrow \text{End}(\mathcal{G})$  is a homomorphism satisfying

$$(3.2.1) \quad \det(T - \iota(a) \mid \text{Lie}(\mathcal{G}/S)) = \prod_{\tilde{\tau} \in \Sigma_{E,\infty}} (T - \tilde{\tau}(a))^{2s_{\tilde{\tau}}}.$$

- $\lambda : \mathcal{G} \rightarrow \mathcal{G}^\vee$  is a polarization such that

$$\iota(a) = \lambda^{-1} \circ \iota(a^*)^\vee \circ \lambda.$$

It is easy to see that for any  $x \in \underline{\text{Sh}}'_K(G'_S)(S)$  with associated quadruple  $(\mathcal{A}_x, \iota, \lambda, \varepsilon_{K^p})$ , the  $p$ -divisible group  $\mathcal{G}_x := \mathcal{A}_x[p^\infty]$  is a  $p$ -divisible group with  $\mathcal{D}$ -structure.

Suppose  $p$  factors as  $\mathfrak{p}_1 \dots \mathfrak{p}_r$  in  $F$ ; we have a decomposition  $F \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong F_{\mathfrak{p}_1} \times \dots \times F_{\mathfrak{p}_r}$ . This induces a decomposition

$$D_{S,p} \cong D_{S,\mathfrak{p}_1} \times \dots \times D_{S,\mathfrak{p}_r}.$$

The datum  $\mathcal{D}$  decomposes as a product of data  $\mathcal{D}_i := (\mathcal{O}_{D_{S,\mathfrak{p}_i}} \subset D_{S,\mathfrak{p}_i}, *, \Lambda_i \subset W_i, \psi_i)$ , where  $W_i$  is the subspace of  $W \otimes_{\mathbb{Q}} \mathbb{Q}_p$  on which  $D_{S,p}$  acts via  $D_{S,\mathfrak{p}_i}$  and  $\psi_i$  and  $*$  are the restrictions of  $\psi \otimes_{\mathbb{Q}} \mathbb{Q}_p$  and  $*$  to  $W_i$ . Similarly to Definition 3.2.1, we may define the notion of  $p$ -divisible group with  $\mathcal{D}_i$ -structure. By [Ham15, Corollary 4.5 (1)], there is an equivalence of categories

$$(3.2.2) \quad \{p\text{-divisible groups with } \mathcal{D}\text{-structure}\} \cong \prod_i \{p\text{-divisible groups with } \mathcal{D}_i\text{-structure}\}$$

which takes isogenies to isogenies. Moreover this equivalence gives a decomposition

$$\mathcal{G} \cong \prod_{i=1}^r \mathcal{G}_i$$

for  $\mathcal{G}$  a  $p$ -divisible group with  $\mathcal{D}$ -structure. The  $\mathcal{G}_i$  are equipped with polarizations  $\lambda_i$  and  $\mathcal{O}_{D_{S,\mathfrak{p}_i}}$ -actions and the isomorphism identifies  $\lambda$  with the product polarization and  $\iota$  with the product of the actions of  $\mathcal{O}_{D_{S,\mathfrak{p}_i}}$ .

For any  $x \in \mathcal{X}'(\overline{\mathbb{F}}_p)$ , we write  $J_x$  for the group of automorphism of  $\mathcal{G}_x$  in the isogeny category preserving the  $\mathcal{O}_{D_{S,p}}$ -action and the polarization up to a  $\mathbb{Q}_p^\times$  scalar. Then  $J_x$  is a reductive group over  $\mathbb{Q}_p$  which is an inner form of a Levi subgroup of  $G'_{S,\mathbb{Q}_p}$  and there is a natural embedding  $I_x \otimes_{\mathbb{Q}} \mathbb{Q}_p \hookrightarrow J_x$ . Similarly to the definition of  $I_x^1$ , the map  $G'_S \rightarrow T'_{\text{ab}}$  induces a natural map  $J_x \rightarrow T'_{\text{ab},\mathbb{Q}_p}$  and we write  $J_x^1$  for its kernel. Then  $J_x^1$  breaks up into a product  $J_{x,i}^1$  corresponding to the decomposition  $\mathcal{G}_x \cong \prod_{i=1}^r \mathcal{G}_{x,i}$ .

Now suppose  $k$  is an algebraically closed field of characteristic  $p$  and  $x \in \mathcal{X}'^{\text{ord}}(k)$ . We write  $\widehat{U}_x$  for the formal neighborhood of  $\mathcal{X}'$  at the point  $x$ . By the Serre–Tate theorem,  $\widehat{U}_x$  is identified with the characteristic  $p$  deformation space of  $\mathcal{G}_x$ ; see [Ham15, Proposition 2.9] for example. In other words,



let  $\text{Def}_{\mathcal{G}_x, \lambda}$  be the functor on local Artinian rings  $R/k$  with residue field  $k$  satisfying:

$$\text{Def}_{\mathcal{G}_x, \lambda}(R) = \left\{ (\tilde{\mathcal{G}}, \tilde{\lambda}, \iota) / R \text{ a } p\text{-divisible group with } \mathcal{D}\text{-structure, } \theta : \tilde{\mathcal{G}} \otimes_R k \xrightarrow{\sim} \mathcal{G}_x \right\} / \sim.$$

Here  $\theta$  is an isomorphism of  $p$ -divisible groups with  $\mathcal{D}$  structure (i.e. preserves actions and polarizations) and  $\sim$  is the equivalence relation identifying isomorphic  $p$ -divisible groups with  $\mathcal{D}$ -structure. Then  $\text{Def}_{\mathcal{G}_x, \lambda}$  is representable by a formal scheme  $\text{Spf } R$ , and  $\text{Spf } R \cong \widehat{U}_x$ . The equivalence of categories (3.2.2) implies that there is an isomorphism

$$\text{Def}_{\mathcal{G}_x, \lambda} \cong \prod_{i=1}^r \text{Def}_{\mathcal{G}_{x,i}, \lambda_i}$$

where  $\text{Def}_{\mathcal{G}_{x,i}, \lambda_i}$  is the deformation space of the  $p$ -divisible group  $\mathcal{G}_{x,i}$  with  $\mathcal{D}_i$ -structure

(3.2.3)

$$\text{Def}_{\mathcal{G}_{x,i}, \lambda_i}(R) = \left\{ (\tilde{\mathcal{G}}_i, \tilde{\lambda}_i, \iota_i) / R \text{ a } p\text{-divisible group with } \mathcal{D}_i\text{-structure, } \theta_i : \tilde{\mathcal{G}}_i \otimes_R k \xrightarrow{\sim} \mathcal{G}_{x,i} \right\} / \sim.$$

**Theorem 3.2.2** ([Moo04, Theorem p226, Example 3.3.2]). *For  $i = 1, \dots, r$ ,  $\text{Def}_{\mathcal{G}_{x,i}, \lambda_i}$  is a formal  $p$ -divisible group. Hence  $\text{Def}_{\mathcal{G}_x, \lambda}$  is a formal  $p$ -divisible group.*

The results in [Moo04] in fact give an explicit description of the group structure of these deformation spaces which we now recall. To do this we introduce some notations. Let  $\mathcal{O}$  be the ring of integers in an unramified extension of  $\mathbb{Q}_p$  of degree  $g$ . Recall a  $p$ -divisible group with  $\mathcal{O}$ -structure over a scheme  $S$  is a  $p$ -divisible group  $\mathcal{G}/S$  together with a homomorphism of  $\mathbb{Z}_p$ -algebras  $\iota : \mathcal{O} \rightarrow \text{End}_S(\mathcal{G})$ ; see §A.3. The  $\mathcal{O}$ -height  $\text{ht}_{\mathcal{O}}(\mathcal{G})$  of  $\mathcal{G}$  satisfies

$$\text{ght}_{\mathcal{O}}(\mathcal{G}) = \text{ht}(\mathcal{G}).$$

We identify the set of embeddings  $\mathcal{O} \rightarrow W(k)$  with  $\mathcal{I} := \{1, \dots, g\}$  so that  $\sigma(m) \equiv m + 1 \pmod{g}$ . Let  $d$  be an integer and  $\mathfrak{f} : \mathcal{I} \rightarrow \{0, \dots, d\}$ . For  $m \in \mathcal{I}$  we let  $M_m$  be a  $W(k)$ -module of rank  $d$  with basis  $(e_{m,n})_{n=1, \dots, d}$ . We equip  $M := \bigoplus_{m=1}^g M_m$  with a Frobenius  $\varphi$  given by

$$(3.2.4) \quad \varphi(e_{m,n}) = \begin{cases} e_{m+1,n} & \text{if } n \leq d - \mathfrak{f}(m) \\ pe_{m+1,n} & \text{if } n > d - \mathfrak{f}(m) \end{cases}.$$

This gives  $M$  the structure of a Dieudonné module and hence corresponds to a  $p$ -divisible group  $\mathcal{G}(d, \mathfrak{f})$  over  $k$ . The action of  $\mathcal{O}$  on  $M$  given by letting  $\mathcal{O}$  act on  $M_m$  via the embedding  $m$  induces an action of  $\mathcal{O}$  on  $\mathcal{G}(d, \mathfrak{f})$ . Then  $\mathcal{G}(d, \mathfrak{f})$  is a  $p$ -divisible group with  $\mathcal{O}$ -structure of  $\mathcal{O}$ -height  $d$ . The above discussion extends in the obvious way to the case where  $\mathcal{O} = \prod_{i=1}^r \mathcal{O}_i$  and  $\mathcal{O}_i$  is the ring of integers in a finite unramified extension of  $\mathbb{Q}_p$ .

We now return to the previous notations so that  $x \in \mathcal{X}'^{\text{ord}}(k)$ . Let  $i \in \{1, \dots, r\}$  and  $\mathcal{O}' := \mathcal{O}_{E_{p_i}}$ ; we write  $g := [F_{p_i} : \mathbb{Q}_p]$ . By Morita equivalence, there exists a polarized  $p$ -divisible group  $(\mathcal{G}^\circ, \lambda^\circ)$  with  $\mathcal{O}'$ -structure of  $\mathcal{O}'$ -height 2 such that  $(\mathcal{G}^\circ)^2 = \mathcal{G}_{x,i}$ . The polarization  $\lambda^\circ$  induces an involution  $*$  on  $\mathcal{O}'$ . Then  $(\mathcal{O}', *)$  is one of the following two forms, see [Moo04, §3.1.2]:

Case (AL):  $\mathfrak{p}_i$  splits as  $\mathfrak{q}_i \bar{\mathfrak{q}}_i$  in  $E$  and  $\mathcal{O}' = W(\mathbb{F}_{p^g}) \times W(\mathbb{F}_{p^g})$ , where the involution  $*$  is given by  $(a, b) \mapsto (b, a)$ .

Case (AU):  $\mathfrak{p}_i$  is inert in  $E$  and  $\mathcal{O}' = W(\mathbb{F}_{p^{2g}})$ . The involution  $*$  is given by the  $p^g$ -Frobenius.

In the case (AL), we write  $\mathcal{O} = \mathcal{O}_{E_{\mathfrak{q}_i}} \cong W(\mathbb{F}_{p^g})$ . Then by [Ham15, Corollary 4.5 (2) I], there is a  $p$ -divisible group  $\mathcal{G}'$  with  $\mathcal{O}$ -structure of  $\mathcal{O}$ -height 2 such that  $\mathcal{G}^\circ \cong \mathcal{G}' \times \mathcal{G}'^V$ . We identify the set of embeddings  $\mathcal{O} \rightarrow W(k)$  with  $\mathcal{I} := \{1, \dots, g\}$  as above, and we also identify this with  $\Sigma_{E, \infty / \mathfrak{q}_i}$ . Then  $\mathcal{G}' \cong \mathcal{G}(2, \mathfrak{s})$  where  $\mathfrak{s} : \mathcal{I} \rightarrow \{0, 1, 2\}$  is defined by  $\mathfrak{s}(m) = s_m$ ; here  $s_m$  is the number defined in (2.2.1).

We define a function  $\mathfrak{d} : \mathcal{S} \rightarrow \{0, 1\}$  by setting

$$\mathfrak{d}(m) = \begin{cases} 1 & \text{if } s_m = 1 \\ 0 & \text{otherwise} \end{cases}.$$

Then by [Moo04, Theorem p226] there is an isomorphism

$$\text{Def}_{\mathcal{G}_{x,i,\lambda_i}} \cong \text{Def}_{\mathcal{G}'} \cong \mathcal{G}(1, \mathfrak{d})^{\text{for}},$$

where we write  $\mathcal{G}(1, \mathfrak{d})^{\text{for}}$  for the formal group associated to  $\mathcal{G}(1, \mathfrak{d})$  and the space  $\text{Def}_{\mathcal{G}'}$  is the deformation space of the  $p$ -divisible group  $\mathcal{G}'$  with  $\mathcal{O}$ -structure (without polarization). The first isomorphism here is via [Ham15, Corollary 4.5 (2) I]. Then  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  is (the formal group associated to) a  $p$ -divisible group with  $\mathcal{O}$ -structure of  $\mathcal{O}$ -height 1.

In the case (AU), we write  $\mathcal{O} := \mathcal{O}_{F_{p_i}} \cong W(\mathbb{F}_{p^g})$ ; recall  $\mathcal{O}' := \mathcal{O}_{E_{p_i}} \cong W(\mathbb{F}_{p^{2g}})$ . Then  $\mathcal{G}^\circ \cong \mathcal{G}(2, \mathfrak{s})$  where  $\mathfrak{s} : \{1, \dots, 2g\} \rightarrow \{0, 1, 2\}$  is defined by  $\mathfrak{s}(m) = s_m$  as above. Here we have identified  $\{1, \dots, 2g\}$  with  $\Sigma_{E, \infty/p_i}$ . There exists a  $p$ -divisible group with  $\mathcal{O}'$ -structure  $\mathcal{G}'$  of  $\mathcal{O}'$ -height 1 and an isomorphism  $\mathcal{G}^\circ \cong \mathcal{G}' \times \mathcal{G}'^{\vee}$ . The polarization  $\lambda^\circ$  on  $\mathcal{G}^\circ$  switches the two factors.

We define a  $p$ -divisible group  $\tilde{\mathcal{G}} := \mathcal{G}(1, \mathfrak{d})$  with  $\mathcal{O}'$ -structure where  $\mathfrak{d}$  is defined by

$$\mathfrak{d}(m) = \begin{cases} 1 & \text{if } s_m = 1 \\ 0 & \text{otherwise} \end{cases}$$

as above. In this case, the polarization  $\lambda^\circ$  induces an automorphism of  $\tilde{\mathcal{G}}^{\text{for}}$ ; then there is an isomorphism

$$\text{Def}_{\mathcal{G}_{x,i,\lambda_i}} \cong \text{Def}_{\mathcal{G}^\circ, \lambda^\circ} \cong (\tilde{\mathcal{G}}^{\text{for}})^{\lambda^\circ},$$

where  $(\tilde{\mathcal{G}}^{\text{for}})^{\lambda^\circ}$  denotes the formal subgroup of  $\tilde{\mathcal{G}}^{\text{for}}$  fixed by the automorphism induced by  $\lambda^\circ$ . In this case the deformation space  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  is no longer stable under the action of  $\mathcal{O}'$ . However it is stable under  $\mathcal{O}$  and indeed it is (the formal group associated to) a  $p$ -divisible group with  $\mathcal{O}$ -structure of  $\mathcal{O}$ -height 1.

Using the definition of slopes for a  $p$ -divisible group with  $\mathcal{O}$ -structure as in [Moo04, §1.2.5], it is easily checked that  $\mathcal{G}'$  in case (AL) and  $\mathcal{G}^\circ$  in case (AU) has one slope if  $\Sigma_{\infty/p_i} - S_{\infty/p_i} = \emptyset$  and two slopes otherwise. Moreover we have that  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  has dimension  $|\Sigma_{\infty/p_i} - S_{\infty/p_i}|$  in both cases (AL) and (AU); in particular  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  is trivial if  $\Sigma_{\infty/p_i} - S_{\infty/p_i} = \emptyset$ .

We also need a description of the endomorphism algebra of  $\mathcal{G}_{x,i}$  and its action on the deformation space. For  $i \in \{1, \dots, r\}$ , we write  $J_{x,i} := \text{Aut}_{(\mathcal{O}', \lambda_i^\circ)}(\mathcal{G}_{x,i}^\circ)$  for the group of automorphisms of  $\mathcal{G}_{x,i}^\circ$  respecting the  $\mathcal{O}'$  action and polarization. The group  $J_{x,i}$  acts on  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  via modification of the isomorphism  $\theta_i$  in (3.2.3).

In case (AL),  $J_{x,i}$  can be identified with the group of automorphisms of  $\mathcal{G}'$  respecting the  $\mathcal{O}_{F_{p_i}}$ -structure. If  $\mathcal{G}'$  has one slope, in other words if  $s_m \neq 1$  for any  $m \in \Sigma_{E, \infty/q_i}$ , then  $J_{x,i} \cong M_2(\mathcal{O}_{p_i})$ . In this case the deformation space is trivial and  $J_{x,i}$  acts trivially on  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$ . If  $\mathcal{G}'$  has more than one slope, which occurs if  $S_{\infty/p_i} \neq \Sigma_{\infty/p_i}$ , then  $J_{x,i} \cong (\mathcal{O}_{F_{p_i}}^\times)^2$ . In this case  $J_{x,i}$  acts on  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  via  $(a, b) \mapsto \iota(ab^{-1})$ ; here  $\iota$  gives the  $\mathcal{O}_{F_{p_i}}$ -structure on  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$ .

In case (AU), if  $\mathcal{G}^\circ$  has one slope, then  $J_{x,i}$  is identified with the subgroup in  $\text{GL}_2(\mathcal{O}_{E_{p_i}})$  preserving the Hermitian form given by the matrix  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ . In this case the deformation space is trivial and  $J_{x,i}$  acts trivially on  $\text{Def}_{\mathcal{G}_{x,i}}$ . If  $\mathcal{G}^\circ$  has two slopes,  $J_{x,i}$  is identified with  $\mathcal{O}_{E_{p_i}}^\times$ . Then  $J_{x,i}$  acts on  $\text{Def}_{\mathcal{G}_{x,i}}$  via  $a \mapsto \iota(\text{Nm}_{E_{p_i}/F_{p_i}}(a))$ .

It follows from these descriptions that  $\text{Def}_{\mathcal{G}_{x,\lambda}}$  is equipped with an action of  $(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p)^\times$ .

Now assume  $x' \in \mathcal{X}'^{\text{ord}}(\overline{\mathbb{F}}_p)$ . We let  $x \in Z'^p(x')(\overline{\mathbb{F}}_p)$  be a  $\mu$ -ordinary point which lies in the smooth locus of  $Z'^p(x')$  and we write  $\widehat{Z}'^p(x') \subset \text{Def}_{\mathcal{G}_{x,\lambda}}$  for the formal completion of  $Z'^p(x')$  at  $x$ . The analysis of the actions of  $J_{x,i}$  on  $\text{Def}_{\mathcal{G}_{x,i,\lambda_i}}$  leads to the following proposition.

**Proposition 3.2.3.**  $\widehat{Z}_x^{lp}(x') \subset \text{Def}_{\mathcal{G}_x}$  is stable under the action of an open subgroup of  $(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p)^\times$ .

*Proof.* Recall we have the groups  $J_x$  and  $J_x^1$  defined above. Let  $T \subset I_x$  denote a maximal torus and  $T^1$  its intersection with  $I_x^1$ . By [Kis17, Corollary 2.1.7],  $I_x$  and  $J_x$  have the same rank, and hence  $I_x^1$  and  $J_x^1$  have the same rank. It follows that  $T_{\mathbb{Q}_p}^1$  is a maximal torus of  $J_x^1$ .

Now  $T^1(\mathbb{Q}_p)$  breaks up into a product  $\prod_{i=1}^r T_i(\mathbb{Q}_p)$  where  $T_i(\mathbb{Q}_p) \subset J_{x,i}^1$ . Let  $i \in \{1, \dots, r\}$  be an element such that  $\mathcal{G}'_{x,i}$  in case (AL) (resp.  $\mathcal{G}^\circ_{x,i}$  in case (AU)) has more than one slope. It can be checked in the two cases there is the following description of  $T_i^1(\mathbb{Q}_p)$ :

Case (AL):  $T_i^1(\mathbb{Q}_p) \subset \text{Aut}_{\mathcal{O}}^0(\mathcal{G}'_{x,i}) \cong (F_{\mathfrak{p}_i}^\times)^2$  is the subgroup of elements of the form  $(a, b)$  with  $ab = 1$ . Here  $\mathcal{G}'_{x,i}$  is as in the discussion above and  $\text{Aut}_{\mathcal{O}}^0(\mathcal{G}'_{x,i})$  is the group of automorphisms in the isogeny category preserving the  $\mathcal{O}$ -action.

Case (AU):  $T_i^1(\mathbb{Q}_p) \subset \text{Aut}_{\mathcal{O}', \lambda^\circ}^0(\mathcal{G}^\circ_{x,i}) \cong E_{\mathfrak{p}_i}^\times$  is the subgroup consisting of elements  $a \in E_{\mathfrak{p}_i}^\times$  with  $a\bar{a}^{-1} = 1$ . Here  $\text{Aut}_{\mathcal{O}', \lambda^\circ}^0(\mathcal{G}^\circ_{x,i})$  denotes the group of automorphisms in the isogeny category preserving the  $\mathcal{O}'$ -actions and the polarization.

By weak approximation for  $T^1$ , see [San81, Corollary 3.5],  $T^1(\mathbb{Q}) \cap J_x$  is dense in  $T^1(\mathbb{Q}_p) \cap J_x$ . Since the action of  $T^1(\mathbb{Q}) \cap J_x$  on  $\text{Def}_{\mathcal{G}_x, \lambda}$  preserves  $\widehat{Z}^{lp}(x')$  so does the action of  $T^1(\mathbb{Q}_p) \cap J_x$ . By the description of  $T^1(\mathbb{Q}_p)$  in the previous paragraph, the action of  $T^1(\mathbb{Q}_p) \cap J_x^1$  on the deformation space can be described in the following way. We assume  $i \in \{1, \dots, r\}$  is such that  $\mathcal{G}^\circ_{x,i}$  has two slopes. Then in either case (AU) or (AL),  $T_i^1(\mathbb{Q}_p) \cap J_{x,i} \cong \mathcal{O}_{F_{\mathfrak{p}_i}}^\times$  acts on  $\text{Def}_{\mathcal{G}_{x,i}, \lambda_i}$  via  $a \mapsto \iota(a^2)$ . Since the image of  $a \mapsto a^2$  in  $\mathcal{O}_{F_{\mathfrak{p}_i}}^\times$  is open, this proves the proposition.  $\square$

We may relabel the primes  $\mathfrak{p}_i$  so that  $\mathcal{G}'_{x,i}$  in case (AL) (resp.  $\mathcal{G}^\circ_{x,i}$  in case (AU)) has two slopes for  $i = 1, \dots, a$  and one slope for  $i = a + 1, \dots, r$ . Then  $\text{Def}_{\mathcal{G}_{x,i}, \lambda_i}$  is non-trivial for  $i = 1, \dots, a$  and trivial for  $i = a + 1, \dots, r$ , and we have an isomorphism:

$$\text{Def}_{\mathcal{G}_x, \lambda} \cong \prod_{i=1}^a \text{Def}_{\mathcal{G}_{x,i}, \lambda_i}.$$

**Corollary 3.2.4.** *There exists a subset  $\omega \subset \{1, \dots, a\}$  such that*

$$\widehat{Z}_x^{lp}(x') = \prod_{i \in \omega} \text{Def}_{\mathcal{G}_{x,i}, \lambda_i}.$$

*Proof.* By [Cha08, Proposition 4.2],  $\widehat{Z}_x^{lp}(x')$  is a formal  $p$ -divisible subgroup of  $\text{Def}_{\mathcal{G}_{x,i}, \lambda_i}$ . Indeed it can be checked from the explicit description of the action of  $T^1(\mathbb{Q}_p) \cap J_x$  on  $\text{Def}_{\mathcal{G}_x}$  that the conditions stated in [Cha08, Proposition 4.2] are satisfied. By Proposition 3.2.3,  $\widehat{Z}_x^{lp}(x')$  is a formal  $p$ -divisible subgroup stable under the action of an open subgroup of  $(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p)^\times$ . The only such subgroups are the ones described.  $\square$

This proposition can be globalized to the following. Recall the tangent sheaf  $T_{\mathcal{X}'}$  is equipped with an action of  $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$  by Proposition 2.3.1. We write  $T_{\mathcal{X}'}(\mathfrak{p}_i)$  for the subspace of  $T_{\mathcal{X}'}$  on which  $\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p$  acts via the projection to  $\mathcal{O}_{F_{\mathfrak{p}_i}}$ . The next corollary follows immediately from 3.2.4 and faithfully flat descent.

**Corollary 3.2.5.** *Let  $W$  be an irreducible component of the smooth locus  $Z^{lp}(x')^{\text{sm}}$  of  $Z^{lp}(x')$ . Then there exists a subset  $\omega \subset \{1, \dots, a\}$  such that the tangent sheaf  $T_W$  of  $W$  is given by*

$$\bigoplus_{i \in \omega} T_{\mathcal{X}'}(\mathfrak{p}_i) \otimes_{\mathcal{O}_{\mathcal{X}'}} \mathcal{O}_W \subset T_{\mathcal{X}'} \otimes_{\mathcal{O}_{\mathcal{X}'}} \mathcal{O}_W.$$

$\square$

**3.3. Proof of the main theorem.** Recall for  $x \in \mathcal{X}'(\overline{\mathbb{F}}_p)$ ,  $Y^p(x)$  and  $Y_\ell(x)$  are the reduced prime-to- $p$  and  $\ell$ -power Hecke orbits respectively, and  $Z'^p(x)$  and  $Z'_\ell(x)$  are their respective closures.

**Proposition 3.3.1.**  $Z'_\ell(x) \cap \mathcal{X}'^{\text{ss}} \neq \emptyset$ .

*Proof.* Since  $\mathcal{X}'$  is proper over  $\overline{\mathbb{F}}_p$ ,  $Z'_\ell(x)$  is proper. By the main theorem of [VW13], there is a stratification on  $\mathcal{X}'$  whose strata are quasi-affine and stable under prime-to- $p$  Hecke correspondences. The strata  $\mathcal{X}'_w$  are parametrized by a certain subset  ${}^JW$  of the Weyl group  $W$  of  $G'_S$ . The closure relations are given by a partial ordering on  ${}^JW$ .

We let  $w \in {}^JW$  be a minimal element for this partial order such that  $\mathcal{X}'_w \cap Z'_\ell(x) \neq \emptyset$ . Then  $\mathcal{X}'_w \cap Z'_\ell(x)$  is closed in  $\mathcal{X}'$  by the minimality of  $w$  and the properness of  $Z'_\ell(x)$ , hence  $\mathcal{X}'_w \cap Z'_\ell(x)$  is proper. Since  $\mathcal{X}'_w \cap Z'_\ell(x)$  is also quasi-affine, it follows that  $\mathcal{X}'_w \cap Z'_\ell(x)$  is 0-dimensional. By Lemma 3.3.2 below,  $\mathcal{X}'_w \cap Z'_\ell(x) \subset \mathcal{X}'^{\text{ss}} \neq \emptyset$ . □

**Lemma 3.3.2.** *Let  $Y \subset \mathcal{X}'$  be a 0-dimensional subscheme of  $\mathcal{X}'$  which is stable under reduced  $\ell$ -power Hecke correspondences. Then  $Y \subset \mathcal{X}'^{\text{ss}}$ .*

*Conversely if  $x \in \mathcal{X}'^{\text{ss}}(\overline{\mathbb{F}}_p)$ , then  $Y^p(x)$  is finite and hence  $Z'^p(x)$  is 0-dimensional.*

*Proof.* First assume  $x \in \mathcal{X}'^{\text{ss}}(\overline{\mathbb{F}}_p)$ . Then  $I_x \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \cong G'_S \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$  for all  $\ell \neq p$  by [XZ, 7.2.14] and hence  $I_x^1 \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \cong G'_{\text{der}} \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ . Therefore  $Y^p(x) \cong I_x^1(\mathbb{Q}) \cap H_p \backslash G'_{\text{der}}(\mathbb{A}_f^p) K'^p / K'^p$  is finite by the finiteness of class groups.

Now we assume  $Y$  is finite and  $x \in Y(\overline{\mathbb{F}}_p)$ ; then  $Y_\ell(x)$  is finite. It follows that  $I_x^1(\mathbb{Q}_\ell) \backslash G'_{\text{der}}(\mathbb{Q}_\ell)$  is compact. The same argument as in [Kis17, Corollary 2.1.7] shows that

$$I_x^1 \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \cong G'_{\text{der}} \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$$

and hence

$$I_x \otimes_{\mathbb{Q}} \mathbb{Q}_\ell \cong G'_S \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$$

since  $I_x$  and  $G'_S$  have the same ranks by [Kis17, Corollary 2.1.7]. But  $I_x \otimes_{\mathbb{Q}} \mathbb{Q}_p$  is a subgroup of  $J_x$  which is an inner form of a Levi subgroup of  $G'_S \otimes_{\mathbb{Q}} \mathbb{Q}_p$ . Since  $I_x$  and  $G'_S$  have the same dimension,  $J_x$  is an inner form of  $G'_S$ . Therefore the Newton cocharacter defined as in §A.2 is central and hence  $x$  is a supersingular point. □

We will need the following two results on the structure of  $I_x^1$  for supersingular points.

**Proposition 3.3.3.** *Suppose  $x \in \mathcal{X}'^{\text{ss}}(\overline{\mathbb{F}}_p)$ .*

- (1) *The natural map  $I_x^1 \otimes_{\mathbb{Q}} \mathbb{Q}_p \rightarrow J_x^1$  is an isomorphism.*
- (2)  *$I_x^1$  is an inner form of  $G'_{\text{der}}$ . In particular  $I_x^1$  is simply connected.*

*Proof.* By [XZ, Lemma 7.2.14]  $I_x$  is an inner form of  $G$  and  $I_x \otimes_{\mathbb{Q}} \mathbb{Q}_p \rightarrow J_x$  is an isomorphism. Therefore  $I_x^1$  is an inner form of  $G'_{\text{der}}$  and  $I_x^1 \otimes_{\mathbb{Q}} \mathbb{Q}_p \rightarrow J_x^1$  is an isomorphism. □

We can now complete the proof of the main theorem.

*Proof of Theorem 3.1.3.* We proceed as in [Cha95, Proposition 7]; we will assume  $Z'^p(x)$  is not equal to  $\mathcal{X}'^{\text{cl}}$  and deduce a contradiction. By Proposition 3.3.1,  $Z'^p(x)$  contains a supersingular point  $s \in \mathcal{X}'^{\text{ss}}(\overline{\mathbb{F}}_p)$ . We write  $\widehat{Z}'_s{}^p(x)$  for the completion of  $Z'^p(x)$  at the point  $s$ .

Let  $\mathcal{A}_s$  be the abelian variety with  $\mathcal{O}_{D_S}$ -multiplication associated to  $s$ . The corresponding  $p$ -divisible group  $(\mathcal{G}_s, \lambda, \iota)$  with  $\mathcal{D}$ -structure breaks up into a product  $\mathcal{G}_{s,1} \times \cdots \times \mathcal{G}_{s,r}$  of  $p$ -divisible groups with  $\mathcal{D}_i$ -structure and there is a decomposition of the deformation space

$$\text{Def}_{\mathcal{G}_s, \lambda} \cong \text{Def}_{\mathcal{G}_{s,1}, \lambda_1} \times \cdots \times \text{Def}_{\mathcal{G}_{s,r}, \lambda_r}.$$

We write  $J_s$  for the group of automorphisms of  $\mathcal{G}_s$  respecting the  $\mathcal{O}_{D_s}$ -action and polarization, and  $J_s^1$  its intersection with  $J_s^1$ . Then  $J_s^1$  is a compact open subgroup of  $J_s^1$  and there is a product decomposition

$$J_s^1 \cong \prod_{i=1}^r J_{s,i}^1.$$

Since  $I_s^1$  is simply connected and  $I_s^1 \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong J_s^1$  by Proposition 3.3.3, the weak approximation theorem implies that  $I_s^1(\mathbb{Q}) \cap J_s^1$  is dense in  $J_s^1$ . Since  $I_s^1(\mathbb{Q}) \cap J_s^1$  preserves the subspace  $\widehat{Z}'^p(x) \subset \text{Def}_{\mathcal{G}_s, \lambda}$ , it is preserved by the group  $J_s^1$ .

There exists an irreducible component  $W$  of the smooth locus of  $Z'^p(x)$  such that  $s$  lies in the closure of  $W$ . By Corollary 3.2.5, the tangent sheaf  $T_W$  is equal to a product

$$\bigoplus_{i \in \omega} T_{\mathcal{X}'(\mathfrak{p}_i)} \otimes_{\mathcal{O}_{\mathcal{X}'}} \mathcal{O}_W$$

where  $\omega \subset \{1, \dots, a\}$  is a proper subset since we assumed  $Z'^p(x) \neq \mathcal{X}'^{c'}$ . We fix a choice of  $j \in \{1, \dots, a\} - \omega$ .

Let  $\xi : \text{Spec } k[[t]] \rightarrow \text{Def}_{\mathcal{G}_s, \lambda}$  be a formal curve which generically lies in  $W$  and with special fiber  $s$ , and we write  $\xi_i : \text{Spec } k[[t]] \rightarrow \text{Def}_{\mathcal{G}_s, i, \lambda_i}$  for the  $\mathfrak{p}_i$ -component of the map  $\xi$ . We write  $\eta$  for the generic fiber of  $\xi$ ; then  $\eta$  is an ordinary point. It follows that  $J_{\eta, j}^1$  is a torus by the description given in 3.2. Therefore, there exist elements  $u_j^{(n)} \in J_{s, j}$  such that  $u_j^{(n)} \xi \neq \xi$  but  $(u_j^{(n)} \xi)$  is congruent to  $\xi$  modulo arbitrarily high powers of  $t$  as  $n \rightarrow \infty$ . For  $i \neq j$ , the  $\mathfrak{p}_i$ -component of  $u_j^{(n)} \xi$  is equal to  $\xi_i$ .

Let  $\widehat{W}_s$  denote the formal neighbourhood at  $s$  of the closure of  $W$  and we write  $\widehat{W}_{s, \mathfrak{p}_j}$  for the image of  $\widehat{W}_s$  in the  $\mathfrak{p}_j$ -component  $\text{Def}_{\mathcal{G}_s, j, \lambda_j}$ . Then upon replacing  $u_j^{(n)}$  by a subsequence, there is an irreducible component  $\widehat{U} \subset \widehat{W}_{s, \mathfrak{p}_j}$  containing the  $\mathfrak{p}_j$ -component of  $u_j^{(n)} \xi$  for all  $n$ . Let  $\widehat{W}_{s, \mathfrak{p}_j} = \text{Spf } R_j$  and let  $I \subset R_j$  be the defining ideal for  $\widehat{U}$ . Then  $f(\xi_j) = 0$  for any  $f \in I$  by continuity, and hence  $\widehat{U}$  contains  $\xi_j$ . By considering the tangent space at the  $k((t))$ -valued point  $R_j/I \xrightarrow{\xi_j} k[[t]] \rightarrow k((t))$ , one sees that the tangent space to  $W$  at the point  $\eta : \text{Spec } k((t)) \rightarrow W$  contains a non-zero  $\mathfrak{p}_j$ -component, which is a contradiction.  $\square$

We may use this to deduce the following corollary

**Corollary 3.3.4.** *Let  $x \in \mathcal{X}'^{c'} \cap \mathcal{X}'^{\text{ord}}(\overline{\mathbb{F}}_p)$  and  $\ell$  a prime such that  $G'_{\mathbb{S}}$  is unramified at  $\ell$ . Then  $Z'_{\ell}(x) = \mathcal{X}'^{c'}$ .*

*Proof.* By Theorem 3.1.3, it suffices to show that  $Z'_{\ell}(x)$  is stable under reduced prime-to- $p$  Hecke correspondences; we use the same argument as in [Cha95, Theorem 1].

By Proposition 3.3.1,  $Z'_{\ell}(x)$  contains a supersingular point  $s \in \mathcal{X}'^{\text{ss}}(\overline{\mathbb{F}}_p)$ . As in the proof of Lemma 3.3.2, we have an isomorphism  $I_s^1 \otimes_{\mathbb{Q}} \mathbb{A}_f^p \cong G'_{\text{der}} \otimes_{\mathbb{Q}} \mathbb{A}_f^p$ . Since  $I_s^1$  is simply connected, the strong approximation theorem implies  $I_s^1(\mathbb{Q}) I_s^1(\mathbb{Q}_{\ell})$  is dense in  $I_s^1(\mathbb{A}_f)$ . It follows that  $I_s^1(\mathbb{Q}) \cap H_p K'^{p, \ell}$  is dense in  $I_s^1(\mathbb{Q}_p) \cap H_p$ . Since  $I_s^1(\mathbb{Q}) \cap H_p K'^{p, \ell}$  is contained in the stabilizer of  $Z'_{\ell}(x)$  at  $s$ , the completion  $\widehat{Z}'_{\ell, s}(x)$  at  $s$  is stable under the action of  $I_s^1(\mathbb{Q}_p) \cap H_p$ .

Let  $\tilde{\gamma} \in G'_{\text{der}}(\mathbb{A}_f^p)$  and we write  $\gamma$  for the corresponding reduced prime-to- $p$  Hecke correspondence. We write  $\gamma(Z'_{\ell}(x))$  for the image of  $Z'_{\ell}(x)$  under  $\gamma$ ; we need to show that  $\gamma(Z'_{\ell}(x)) = Z'_{\ell}(x)$ . The strong approximation theorem also implies that  $Y'^p(s) = Y'_{\ell}(s)$  by the description in Proposition 3.1.5. It follows that upon right multiplying  $\tilde{\gamma} \in G'_{\text{der}}(\mathbb{A}_f^p)$  by an element of  $G'_{\text{der}}(\mathbb{Q}_{\ell})$ , which does not alter  $\gamma(Z'_{\ell}(x))$ , we may assume  $\tilde{\gamma} = ik$  for some  $i \in I_s^1(\mathbb{Q}) \cap H_p$  and  $k \in K'^p$ . Modifying again by  $k^{-1}$ , we may assume  $\tilde{\gamma} \in I_s^1(\mathbb{Q}) \cap H_p$ . By the previous paragraph,  $\tilde{\gamma}$  preserves  $\widehat{Z}'_{\ell, s}(x)$  and hence the formal completion of  $Z'_{\ell}(x) \cup \gamma(Z'_{\ell}(x))$  at  $s$  is equal to  $\widehat{Z}'_{\ell, s}(x)$ . Applying the argument to all supersingular points in  $Z'_{\ell}(x)$ , we find that  $Z'_{\ell}(x) \cup \gamma(Z'_{\ell}(x))$  is equal to  $Z'_{\ell}(x) \cup W$  where  $W$  is a closed subscheme

of  $\mathcal{X}'$  stable under reduced  $\ell$ -power Hecke correspondences and which does not contain supersingular points. Then Proposition 3.3.1 implies that  $W$  is empty. Since  $\gamma$  was arbitrary, it follows that  $Z'_\ell(x)$  is stable under reduced prime-to- $p$  Hecke correspondences as desired.  $\square$

**3.4. The result in the quaternionic case.** We now explain how Theorem 3.1.3 may be used to deduce the Hecke-orbit conjecture for quaternionic Shimura varieties. The open subscheme  $\mathcal{X}'^{\text{ord}}$  transfers to an open subscheme  $\mathcal{X}^{\text{ord}}$  of  $\mathcal{X}$ , see for example [SZ]<sup>2</sup>. We write  $G_{\text{der}}$  for the derived group of  $G_{\text{S}}$ . Then  $G_{\text{der}} \cong G'_{\text{der}}$ .

For  $x \in \mathcal{X}(\overline{\mathbb{F}}_p)$ , we may define the reduced prime-to- $p$  Hecke orbit  $Y^p(x)$  and reduced  $\ell$ -power Hecke orbit  $Y_\ell(x)$  in the same way as for  $\mathcal{X}'$ . Namely, we let  $\tilde{x}$  denote a lift of  $x$  to  $\text{Sh}_{K_p}(G_{\text{S,T}})(\overline{\mathbb{F}}_p)$  and we write  $\tilde{Y}^p(x)$  (resp.  $\tilde{Y}_\ell(x)$ ) for the  $G_{\text{der}}(\mathbb{A}_f^p)$ -orbit (resp.  $G_{\text{der}}(\mathbb{Q}_\ell)$ -orbit) of  $\tilde{x}$ . Then  $Y^p(x)$  (resp.  $Y_\ell(x)$ ) is defined to be the image of  $\tilde{Y}^p(x)$  (resp.  $\tilde{Y}_\ell(x)$ ) in  $\mathcal{X}$ . We write  $Z^p(x)$  and  $Z_\ell(x)$  for the closures of  $Y^p(x)$  and  $Y_\ell(x)$  respectively.

Recall  $\nu_{\text{S}} : G_{\text{S}} \rightarrow T_F$  denotes the reduced norm. Then  $\nu_{\text{S}}$  induces an isomorphism

$$\pi_0(\mathcal{X}) \cong T_F(\mathbb{Q})^\dagger \backslash T_F(\mathbb{A}_f) / \nu_{\text{S}}(K)$$

where  $T_F(\mathbb{Q})^\dagger$  is defined as in 3.1. For  $\mathfrak{c}$  in the right hand side we write  $\mathcal{X}^\mathfrak{c}$  for the corresponding connected component

Let  $\ell$  be a prime coprime to  $p$ . The following result follows immediately from Theorem 3.1.3 by using the fact that the unitary Shimura variety and quaternionic Shimura variety have isomorphic geometric connected components (upon taking the inverse limit over the prime-to- $p$  level) and the fact that  $G'_{\text{der}} \cong G_{\text{der}}$ .

**Theorem 3.4.1.** *Let  $\mathfrak{c} \in T_F(\mathbb{Q})^\dagger \backslash T_F(\mathbb{A}_f) / \nu_{\text{S}}(K)$  and  $x \in \mathcal{X}^\mathfrak{c} \cap \mathcal{X}^{\text{ord}}(\overline{\mathbb{F}}_p)$ . Then*

- (1)  $Z^p(x) = \mathcal{X}^\mathfrak{c}$ .
- (2)  $Z_\ell(x) = \mathcal{X}^\mathfrak{c}$ .  $\square$

*Remark 3.4.2.* As in Remark 3.1.4, one can use this to deduce that the prime-to- $p$  Hecke orbit of an ordinary point is dense in  $\mathcal{X}$ .

#### 4. IHARA'S LEMMA FOR SHIMURA SURFACES

In this section we prove a version of Ihara's lemma for certain quaternionic Shimura surfaces. The argument combines the arguments of [Dim09] and [DT94].

**4.1. Statement of Ihara's Lemma.** We keep the previous notations, so that  $B$  is a totally indefinite quaternion algebra over a totally real field  $F$  of degree  $g := [F : \mathbb{Q}]$ , and we let  $K$  be a sufficiently small compact open subgroup of  $(B \otimes_{\mathbb{Q}} \mathbb{A}_f)^\times$ .

Let  $\Pi$  be an irreducible cuspidal automorphic representation of  $\text{GL}_2(\mathbb{A}_F)$  of parallel weight 2 defined over a number field  $\mathbf{E}$ . We write  $\mathcal{O}_{\mathbf{E}}$  for the ring of integers of  $\mathbf{E}$  and  $k_\lambda$  for the residue field of  $\mathcal{O}_{\mathbf{E}}$  at a prime  $\lambda$ . Let  $\mathbf{R}$  be a finite set of places of  $F$  away from which  $\Pi$  is unramified and  $K$  is hyperspecial; in particular  $\mathbf{R}$  contains the ramification set of  $B$ . Let  $\mathbf{T}_{\mathbf{R}}$  denote the (abstract) Hecke algebra away from  $\mathbf{R}$ , i.e. the polynomial ring over  $\mathbb{Z}$  generated by  $T_{\mathfrak{q}}, S_{\mathfrak{q}}$  where  $\mathfrak{q}$  is a prime away from  $\mathbf{R}$ . The representation  $\Pi$  determines a homomorphism

$$\phi_{\mathbf{R}}^{\Pi} : \mathbf{T}_{\mathbf{R}} \rightarrow \mathcal{O}_{\mathbf{E}}$$

via the Hecke eigenvalues of  $\Pi$ . For every prime  $\lambda$  of  $\mathbf{E}$  there is an attached Galois representation

$$\rho_{\Pi, \lambda} : \text{Gal}(\overline{F}/F) \rightarrow \text{GL}_2(\mathcal{O}_{\mathbf{E}, \lambda})$$

<sup>2</sup>It is not a priori clear that the definitions of ordinary locus is independent of the choice of auxiliary PEL data. For the applications, it is not necessary to know this.

unramified outside of  $R \cup R_\lambda$  where  $R_\lambda$  is the set of primes of  $F$  having the same zdue characteristic as  $\lambda$ . It is characterized by the property that for  $\mathfrak{q} \notin R \cup R_\lambda$ , the characteristic polynomial of the arithmetic Frobenius  $\text{Frob}_\mathfrak{q}$  at  $\mathfrak{q}$  is given by

$$X^2 - \phi_{\mathbb{R}}^{\Pi}(T_{\mathfrak{q}})X + \text{Nm}_{F/\mathbb{Q}}(\mathfrak{q})\phi_{\mathbb{R}}^{\Pi}(S_{\mathfrak{q}}).$$

We write  $\bar{\rho}_{\Pi,\lambda}$  for the reduction of  $\rho_{\Pi,\lambda} \pmod{\lambda}$ .

We now assume  $p$  is a prime which is inert in  $F$  and such that  $K = K_p K^p$  where  $K_p \subset G_S(\mathbb{Q}_p)$  is a hyperspecial subgroup. We let  $\mathfrak{p}$  denote the unique prime of  $F$  above  $p$ . Let  $\mathfrak{m} := \mathfrak{m}_{\Pi,\lambda} \subset \mathbf{T}_{R \cup \{\mathfrak{p}\}}$  denote the preimage of the ideal  $(\lambda) \subset \mathcal{O}_{\mathbf{E}}$  under the map  $\phi_{R \cup \{\mathfrak{p}\}}^{\Pi}$ ; it is a maximal ideal of  $\mathbf{T}_{R \cup \{\mathfrak{p}\}}$ . For any  $\mathbf{T}_{R \cup \{\mathfrak{p}\}}$ -module  $M$ , we write  $M_{\mathfrak{m}}$  for the localization of  $M$  at the ideal  $\mathfrak{m}$ . We introduce the following assumptions:

**Assumption 4.1.1.** Let  $\ell$  be the underlying prime of  $\lambda$ .

- (1)  $\ell$  is coprime to  $R$ ,  $\text{disc}(F)$ , and the cardinality of  $F^\times \backslash \mathbb{A}_{F,f}^\times / \mathbb{A}_{F,f}^\times \cap K$ , where we consider  $\mathbb{A}_{F,f}^\times := (F \otimes_{\mathbb{Q}} \mathbb{A}_f)^\times$  as the center of  $(B \otimes_{\mathbb{Q}} \mathbb{A}_f)^\times$ .
- (2)  $\ell \geq g + 2$ .
- (3) The representation  $\bar{\rho}_{\Pi,\lambda}$  satisfies the condition  $\mathbf{LI}_{\text{Ind}\bar{\rho}_{\Pi,\lambda}}$  in [Dim05, Proposition 0.1].
- (4)  $B$  splits at all places above  $\ell$  and the compact open subgroup  $K$  factors as  $K = K_\ell K^\ell$  where  $K_\ell \subset \text{GL}_2(\mathbb{Q}_\ell)$  is a hyperspecial subgroup.

Let  $S \subset \Sigma_\infty \cup \Sigma_p$  and  $T \subset S_\infty$  as in §2.1. Then we have the associated Shimura variety  $\text{Sh}_K(G_{S,T})$ .

**Proposition 4.1.2.** *Suppose Assumption 4.1.1 is satisfied. Then*

- (1)  $H^j(\text{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, \mathcal{O}_{\mathbf{E}_\lambda})_{\mathfrak{m}} = 0$  unless  $j = g_S := |\Sigma_\infty - S|$ .
- (2)  $H^{g_S}(\text{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, \mathcal{O}_{\mathbf{E}_\lambda})_{\mathfrak{m}}$  is a free  $\mathcal{O}_{\mathbf{E}_\lambda}$ -module.

*Proof.* The case of Hilbert modular varieties is proved in [Dim09, Theorem 6.6]. We thus assume  $\text{Sh}_K(G_{S,T})$  is compact.

For (1), it suffices to show that  $H^j(\text{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}} = 0$  for  $j = 0, \dots, g_S - 1$  by Nakayama's Lemma and Poincaré duality. We let  $\tilde{F}$  denote the Galois closure of  $F$  and  $\mathcal{G}_{\tilde{F}}$  its absolute Galois group. For any irreducible representation  $\bar{\rho}'$  of  $\mathcal{G}_{\tilde{F}}$  which appears as a subquotient of  $H^j(\text{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}$ , the same argument as in [Liu19, Theorem 3.21] shows that  $g_S$  appears as a Fontaine–Laffaille weight of  $\bar{\rho}'$ ; note here we need to use Assumption 4.1.1 (2) in order to apply this theory. By Assumption 4.1.1 (4),  $\text{Sh}_K(G_{S,T})_{\mathbb{F}_{\ell^h}}$  admits a proper smooth model over  $W(\mathbb{F}_{\ell^h})$ , where  $h$  is the residue degree of a prime in  $\tilde{F}$  above  $\ell$ . Therefore, by Faltings' Comparison Theorem [Fal89], we know that  $g_S$  cannot be a Fontaine–Laffaille weight for  $H^j(\text{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}$ ,  $j < g_S$ . This proves (1).

For (2) the proof is the same as [Dim05, Theorem 6.6].  $\square$

For the rest of this section we make the following assumption.

- Assumption 4.1.3.** (1)  $\mathfrak{p} \notin S$  and  $|\Sigma_\infty - S| = 2$ .  
(2)  $B_S^\times \neq \text{GL}_2(F)$ .

In particular this condition implies that the associated Shimura varieties are compact. We write  $\Sigma_\infty - S = \{\tau_1, \tau_2\}$  (note we do not use the convention in §2.4 so that  $\sigma(\tau_1)$  is not necessarily equal to  $\tau_2$ ). As in §A.1, we fix an isomorphism  $(B_S \otimes_{\mathbb{Q}} \mathbb{Q}_p)^\times \cong \text{GL}_2(F_p)$  such that  $K_p$  corresponds to the standard hyperspecial subgroup of  $\text{GL}_2(F_p)$ , and we let  $K_{0,p}$  correspond to the standard Iwahori subgroup of  $\text{GL}_2(F_p)$ . We let  $K_0(\mathfrak{p}) = K_{0,p} K^p$ . We then have the Shimura variety  $\text{Sh}_{K_0(\mathfrak{p})}(G_{S,T})$  with Iwahori level structure at  $p$  which is equipped with two finite étale maps  $\pi_1, \pi_2 : \text{Sh}_{K_0(\mathfrak{p})}(G_{S,T}) \rightarrow \text{Sh}_K(G_{S,T})$ .

The main theorem of this section is the following.

**Theorem 4.1.4.** *Under Assumptions 4.1.1 and 4.1.3, the map*

$$\pi_1^* + \pi_2^* : H^2(\text{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}^2 \rightarrow H^2(\text{Sh}_{K_0(\mathfrak{p})}(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}$$

is injective.

*Remark 4.1.5.* In the case of Hilbert modular varieties, this theorem was proved by Dimitrov [Dim09]. For Shimura curves, the case  $F = \mathbb{Q}$  is due to Diamond–Taylor [DT94] and the case of general totally real  $F$  is due to Manning–Shotton [MS21].

We outline the strategy for the proof of this theorem. Following the idea of Diamond–Taylor, we use Faltings’ comparison theorem [Fal89] to reduce to proving the injectivity of a certain map between global sections of line bundles over the  $\text{mod } \ell$  reduction of the Shimura surface (Proposition 4.4.1). The property that such a global section lies in the kernel of the map (4.4.1) implies that the divisor corresponding to this section is stable under certain Hecke correspondences. The Hecke orbit conjecture proved in the previous section shows that this divisor (if non-trivial) must be supported on the complement of the ordinary locus. We then compute the intersection pairing of this divisor with certain Goren–Oort divisors to deduce a contradiction.

**4.2. Intersection numbers.** In this section we compute the intersection numbers of certain cycles on the special fiber of quaternionic Shimura surfaces. The corresponding calculations for Hilbert modular surfaces were carried out in [AG04].

Recall we have assumed  $\ell$  is a prime which is unramified in  $F$  and such that  $K$  is hyperspecial at  $\ell$ ; in other words that  $K = K_\ell K^\ell$  where  $K_\ell \subset G_S(\mathbb{Q}_\ell) \cong \text{GL}_2(F \otimes_{\mathbb{Q}} \mathbb{Q}_\ell)$  is hyperspecial. In this subsection we need to consider the  $\text{mod } \ell$  reductions of the quaternionic Shimura varieties constructed in section 2. Fix an isomorphism  $\iota_\ell : \overline{\mathbb{Q}}_\ell \xrightarrow{\sim} \mathbb{C}$  through which we identify  $\Sigma_\infty$  with  $\Sigma_\ell$ . We let  $X_\ell$  denote the integral model of  $\text{Sh}_K(G_{S,T})$  over  $\mathcal{O}_{\mathbf{E}_{\tilde{S},\tilde{\ell}}}$  determined by a choice of imaginary quadratic extension  $E/F$  and a subset  $\tilde{S}_\infty \subset \Sigma_{E,\infty}$  as in §2.2; here  $\tilde{\ell}$  is the prime of  $\mathbf{E}_{\tilde{S}}$  induced by  $\iota_\ell$ . We may apply the construction in §2 upon replacing  $p$  by  $\ell$ . In this section we will only need to consider geometric special fibers, so we do not need to keep track of the fields of definitions as in §2. In particular, the subset  $T$  will not play a role. We write  $\mathcal{X}_\ell$  for the special fiber of  $X_\ell$  over  $\overline{\mathbb{F}}_\ell$ .

For a choice of auxiliary PEL data, the Newton stratification on the corresponding unitary Shimura variety defines a stratification on  $\mathcal{X}_\ell$ , see §3.1. We write  $\mathcal{X}_\ell^{\text{ord}}$  for the  $\mu$ -ordinary locus of  $\mathcal{X}_\ell$  and we write  $\mathcal{X}_\ell^{\text{n-ord}}$  for the complement of  $\mathcal{X}_\ell^{\text{ord}}$  in  $\mathcal{X}_\ell$ . We would like to understand the intersection numbers of certain cycles supported on  $\mathcal{X}_\ell^{\text{n-ord}}$ . This is possible since in the case of surfaces, we may give an explicit description of the non-ordinary locus in terms of Goren–Oort cycles. The relationship with Goren–Oort strata follows easily from an examination of the Dieudonné modules of the universal  $p$ -divisible group over the unitary Shimura variety. Since the calculations are completely analogous to the case of Hilbert modular varieties (see for example [Sta97] and [LT20] for some cases of totally indefinite quaternion algebras), we omit the computations and just give the statements. We consider the following two cases separately:

Case (1): There exists a prime  $\mathfrak{l}$  above  $\ell$  such that  $\tau_1, \tau_2 \in \Sigma_{\mathfrak{l}/\infty}$ .

Case (2): There exists distinct primes  $\mathfrak{l}_1, \mathfrak{l}_2$  above  $\ell$  such that  $\tau_1 \in \Sigma_{\mathfrak{l}_1/\infty}$  and  $\tau_2 \in \Sigma_{\mathfrak{l}_2/\infty}$ .

We first consider Case (1). Then we have

$$\mathcal{X}_\ell^{\text{n-ord}} \cong \mathcal{X}_\ell^{\text{ss}} = \mathcal{X}_{\ell,\tau_1} \cup \mathcal{X}_{\ell,\tau_2}$$

where  $\mathcal{X}_{\ell,\tau_1}$  and  $\mathcal{X}_{\ell,\tau_2}$  are the Goren–Oort divisors over  $\overline{\mathbb{F}}_\ell$  corresponding to  $\tau_1, \tau_2 \in \Sigma_\infty - S_\infty$ . Here  $\mathcal{X}_\ell^{\text{ss}}$  is defined as in §3.1.

By Proposition 2.4.1,  $\mathcal{X}_{\ell,\tau_1}$  and  $\mathcal{X}_{\ell,\tau_2}$  are  $\mathbb{P}^1$ -bundles over the discrete Shimura set

$$G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K$$

and by [TX19, Proposition 2.32 (3)] there is an isomorphism

$$\mathcal{X}_{\ell,\tau_1} \cap \mathcal{X}_{\ell,\tau_2} \cong G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K_0(\mathfrak{l})$$



where  $K_0(\mathfrak{l}) \subset K$  denotes the compact open which agrees with  $K$  away from  $\mathfrak{l}$  and with Iwahori level structure at  $\mathfrak{l}$ . Here we consider these finite sets as discrete schemes over  $\overline{\mathbb{F}}_\ell$ . We write  $s_\ell(K)$  for the cardinality of  $G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K$  and  $s_\ell(K_0(\mathfrak{l}))$  for the cardinality of  $G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K_0(\mathfrak{l})$ . Then  $s_\ell(K)$  and  $s_\ell(K_0(\mathfrak{l}))$  are related by the equality

$$s_\ell(K_0(\mathfrak{l})) = (\ell^{g_\mathfrak{l}} + 1) s_\ell(K)$$

where  $g_\mathfrak{l} = [F_\mathfrak{l} : \mathbb{Q}_p]$ .

By [TX19, Theorem 4.3]<sup>3</sup> the intersection matrix of  $\mathcal{X}_{\ell, \tau_1}$  and  $\mathcal{X}_{\ell, \tau_2}$  considered as divisors on  $\mathcal{X}_\ell$  is given by

$$\begin{pmatrix} -2\ell^{n_{\tau_1}} s_\ell(K) & s_\ell(K_0(\mathfrak{l})) \\ s_\ell(K_0(\mathfrak{l})) & -2\ell^{n_{\tau_2}} s_\ell(K) \end{pmatrix} = s_\ell(K) \begin{pmatrix} -2\ell^{n_{\tau_1}} & \ell^{g_\mathfrak{l}} + 1 \\ \ell^{g_\mathfrak{l}} + 1 & -2\ell^{n_{\tau_2}} \end{pmatrix}$$

We write  $\omega_{\tau_1}$  (resp.  $\omega_{\tau_2}$ ) for the line bundle on  $\mathcal{X}_\ell$  defined in §2.3 corresponding to a choice of lift  $\tilde{\tau}_1$  (resp.  $\tilde{\tau}_2$ ) in  $\tilde{S}_\infty$ . Since  $\mathcal{X}_{\ell, \tau_1}$  (resp.  $\mathcal{X}_{\ell, \tau_2}$ ) is the vanishing locus of a section of  $(\omega_{\tau_2})^{\ell^{n_{\tau_1}}} \otimes (\omega_{\tau_1})^{-1}$  (resp.  $(\omega_{\tau_1})^{\ell^{n_{\tau_2}}} \otimes (\omega_{\tau_2})^{-1}$ ), applying a change of basis matrix we compute the intersection matrix of  $\omega_{\tau_1}$  and  $\omega_{\tau_2}$  to be

$$\begin{pmatrix} 0 & s_\ell(K) \\ s_\ell(K) & 0 \end{pmatrix}.$$

Now we consider Case (2). In this case we have the two Goren–Oort divisors  $\mathcal{X}_{\ell, \tau_1}$  and  $\mathcal{X}_{\ell, \tau_2}$  and

$$\mathcal{X}_\ell^{\text{n-ord}} = \mathcal{X}_{\ell, \tau_1} \cup \mathcal{X}_{\ell, \tau_2}.$$

By Theorem 2.4.1,  $\mathcal{X}_{\ell, \tau_1}$  and  $\mathcal{X}_{\ell, \tau_2}$  are isomorphic to  $\mathcal{S}_K(G_{S_{\tau_1, T_{\tau_1}}})_{\overline{\mathbb{F}}_\ell}$  and  $\mathcal{S}_K(G_{S_{\tau_2, T_{\tau_2}}})_{\overline{\mathbb{F}}_\ell}$  respectively; these are the special fibers of certain Shimura curves. Moreover, by [TX16, Theorem 5.2],  $\mathcal{X}_{\ell, \tau_1}$  and  $\mathcal{X}_{\ell, \tau_2}$  intersect transversally and there is an identification of  $\mathcal{X}_{\ell, \tau_1} \cap \mathcal{X}_{\ell, \tau_2}$  with the discrete Shimura set

$$G_{S \cup \{\tau_1, \tau_2, \mathfrak{l}_1, \mathfrak{l}_2\}}(\mathbb{Q}) \backslash G_{S \cup \{\tau_1, \tau_2, \mathfrak{l}_1, \mathfrak{l}_2\}}(\mathbb{A}_f) / K_{\mathfrak{l}_1, \mathfrak{l}_2} = G_{\Sigma_\infty \cup \{\mathfrak{l}_1, \mathfrak{l}_2\}}(\mathbb{Q}) \backslash G_{\Sigma_\infty \cup \{\mathfrak{l}_1, \mathfrak{l}_2\}}(\mathbb{A}_f) / K_{\mathfrak{l}_1, \mathfrak{l}_2}$$

where  $K_{\mathfrak{l}_1, \mathfrak{l}_2}$  is the compact open which agrees with  $K$  away from  $\mathfrak{l}_1$  and  $\mathfrak{l}_2$  and is the unique maximal compact at the places  $\mathfrak{l}_1$  and  $\mathfrak{l}_2$ . We write  $s_\ell(K_{\mathfrak{l}_1, \mathfrak{l}_2})$  for the cardinality of this finite set. It follows that the intersection number  $\mathcal{X}_{\ell, \tau_1} \cdot \mathcal{X}_{\ell, \tau_2}$  is equal to  $s_\ell(K_{\mathfrak{l}_1, \mathfrak{l}_2})$ .

For a projective scheme  $S$  of dimension  $d$  over a field  $k$  and  $\mathcal{F}$  a coherent sheaf on  $S$ , we write

$$\chi_S(\mathcal{F}) := \sum_{i=0}^d (-1)^i \dim_k H^i(S, \mathcal{F})$$

for the Euler characteristic of  $\mathcal{F}$ . We write  $\chi_S(0)$  for the Euler characteristic of the structure sheaf.

By [LT20, Proposition 3.2] and Corollary 2.3.3, we have  $\omega_{\tau_2}^2|_{\mathcal{X}_{\ell, \tau_1}} \cong \Omega_{\mathcal{X}_{\ell, \tau_1}}^1$  in the rational Picard group of  $\mathcal{X}_{\ell, \tau_1}$ . Thus

$$\begin{aligned} 2\chi_{\mathcal{X}_{\ell, \tau_1}}(0) &= 2 \deg(\omega_{\tau_2}|_{\mathcal{X}_{\ell, \tau_1}}) \\ &= 2(\ell^{n_{\tau_1}} - 1)\omega_{\tau_1} \cdot \omega_{\tau_2}. \end{aligned}$$

On the other hand, the adjunction formula implies that

$$\begin{aligned} 2\chi_{\mathcal{X}_{\ell, \tau_1}}(0) &= \mathcal{X}_{\ell, \tau_1} \cdot (\mathcal{X}_{\ell, \tau_1} + K_{\text{can}}) \\ &= (\omega_{\tau_1})^{\ell^{n_{\tau_1}} - 1} \cdot \left( (\omega_{\tau_1})^{\ell^{n_{\tau_1}} + 1} + (\omega_{\tau_2})^2 \right) \\ &= (\ell^{2n_{\tau_1}} - 1)\omega_{\tau_1} \cdot \omega_{\tau_1} + 2(\ell^{n_{\tau_1}} - 1)\omega_{\tau_1} \cdot \omega_{\tau_2}. \end{aligned}$$

<sup>3</sup>In [TX19], there is a running assumption that the prime  $p$  ( $\ell$  in our setting) is inert in  $F$ . This is required for the cohomological results to hold. However it is easily checked that the computation of the intersection pairing, which is purely geometric, holds if the prime is only assumed unramified.

where  $K_{\text{can}}$  is the canonical bundle on  $\mathcal{X}_\ell$  and for the second equality we have used Corollary 2.3.3. Thus since  $\ell > 4$ , it follows that  $\omega_{\tau_1} \cdot \omega_{\tau_1} = 0$ , and hence  $\mathcal{X}_{\ell, \tau_1} \cdot \mathcal{X}_{\ell, \tau_1} = 0$ . Similarly we have  $\mathcal{X}_{\ell, \tau_2} \cdot \mathcal{X}_{\ell, \tau_2} = 0$ .

In summary, the intersection matrix of  $\mathcal{X}_{\ell, \tau_1}, \mathcal{X}_{\ell, \tau_2}$  is given by

$$\begin{pmatrix} 0 & s_\ell(K_{I_1, I_2}) \\ s_\ell(K_{I_1, I_2}) & 0 \end{pmatrix}$$

and the intersection matrix of  $\omega_1, \omega_{\tau_2}$  is

$$\frac{1}{(\ell^{n_{\tau_1}} - 1)(\ell^{n_{\tau_2}} - 1)} \begin{pmatrix} 0 & s_\ell(K_{I_1, I_2}) \\ s_\ell(K_{I_1, I_2}) & 0 \end{pmatrix}.$$

**4.3. Connected components.** We will need a variant of the above computations for the connected components of  $\mathcal{X}_\ell$ ; this is due to the fact that we will need to apply the strong approximation theorem in the proof of Proposition 4.3.4 and this only holds for the derived group of  $G_S$ . We keep the notations and assumptions of §4.2.

Recall  $\nu_S : G_S \rightarrow T_F$  denotes the reduced norm and this induces an isomorphism

$$(4.3.1) \quad \pi_0(\mathcal{X}_\ell) \cong T_F(\mathbb{Q})^\dagger \backslash T_F(\mathbb{A}_f) / \nu_S(K).$$

We write  $\text{Cl}_F(K)$  for the right hand side. For an element  $\mathfrak{c} \in \text{Cl}_F(K)$ , we write  $\mathcal{X}_\ell^\mathfrak{c}$  for the connected component of  $\mathcal{X}_\ell$  corresponding to  $\mathfrak{c}$ . We also write  $\omega_{\tau_1}^\mathfrak{c}, \omega_{\tau_2}^\mathfrak{c}$  for the restriction of the lines bundles  $\omega_{\tau_1}, \omega_{\tau_2}$  to  $\mathcal{X}_\ell^\mathfrak{c}$ . We would like to obtain the intersection matrix of the line bundles  $\omega_{\tau_1}^\mathfrak{c}, \omega_{\tau_2}^\mathfrak{c}$  on  $\mathcal{X}_\ell^\mathfrak{c}$ .

There are maps between the Shimura sets

$$(4.3.2) \quad G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K \rightarrow T_F(\mathbb{Q})^\dagger \backslash T_F(\mathbb{A}_f) / \nu_{\Sigma_\infty}(K), \quad \text{in Case (1).}$$

(4.3.3)

$$G_{\Sigma_\infty \cup \{I_1, I_2\}}(\mathbb{Q}) \backslash G_{\Sigma_\infty \cup \{I_1, I_2\}}(\mathbb{A}_f) / K_{I_1, I_2} \rightarrow T_F(\mathbb{Q})^\dagger \backslash T_F(\mathbb{A}_f) / \nu_{\Sigma_\infty \cup \{I_1, I_2\}}(K_{I_1, I_2}), \quad \text{in Case (2).}$$

In each of the equations (4.3.1), (4.3.2) and (4.3.3),  $T_F(\mathbb{Q})^\dagger$  is identified with the set of totally positive elements in  $F$  and the images of  $K$  and  $K_{I_1, I_2}$  under the various reduced norms are all identified, therefore we may identify the right hand sides of (4.3.2) and (4.3.3) with  $\text{Cl}_F(K)$ . We thus write  $(G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K)^\mathfrak{c}$  (resp.  $(G_{\Sigma_\infty \cup \{I_1, I_2\}}(\mathbb{Q}) \backslash G_{\Sigma_\infty \cup \{I_1, I_2\}}(\mathbb{A}_f) / K_{I_1, I_2})^\mathfrak{c}$ ) for the preimage of  $\mathfrak{c}$  under (4.3.2) (resp. (4.3.3)), and we write  $s_\ell(K)^\mathfrak{c}$  (resp.  $s_\ell(K_{I_1, I_2})^\mathfrak{c}$ ) for its cardinality.

Suppose we are in Case (1) of §4.2. For  $i = 1, 2$ , we have an identification of the irreducible components of  $\mathcal{X}_{\ell, \tau_i}$  with  $G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K$ . We may choose this identification compatibly with (4.3.1). In other words, if we denote by  $\mathcal{X}_{\ell, \tau_i}^\mathfrak{c}$  the union of irreducible components of  $\mathcal{X}_{\ell, \tau_i}$  corresponding to  $(G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K)^\mathfrak{c}$ , then  $\mathcal{X}_{\ell, \tau_i}^\mathfrak{c} = \mathcal{X}_{\ell, \tau_i} \cap \mathcal{X}_\ell^\mathfrak{c}$  for  $i = 1, 2$ .

**Proposition 4.3.1.** *In Case (1), the intersection matrix of  $\omega_{\tau_1}^\mathfrak{c}, \omega_{\tau_2}^\mathfrak{c}$  is given by*

$$\begin{pmatrix} 0 & s_\ell(K)^\mathfrak{c} \\ s_\ell(K)^\mathfrak{c} & 0 \end{pmatrix}.$$

*Proof.* Using the fact that  $\mathcal{X}_{\ell, \tau_i} \cap \mathcal{X}_\ell^\mathfrak{c} \cong \mathcal{X}_{\ell, \tau_i}^\mathfrak{c}$  and [TX19, Theorem 4.3], we find that in Case (1) the intersection matrix of  $\mathcal{X}_{\ell, \tau_1}^\mathfrak{c}, \mathcal{X}_{\ell, \tau_2}^\mathfrak{c}$  is given by

$$s_\ell(K)^\mathfrak{c} \begin{pmatrix} -2\ell^{n_{\tau_1}} & \ell^{g_1} + 1 \\ \ell^{g_1} + 1 & -2\ell^{n_{\tau_2}} \end{pmatrix}.$$

Applying a change of basis matrix gives the desired result.  $\square$

Now suppose we are in Case (2). For  $i = 1, 2$  we may identify

$$\pi_0(\mathcal{X}_{\ell, \tau_i}) \cong \text{Cl}_F(K)$$

compatibly with the identification  $\pi_0(\mathcal{X}_\ell) \cong \text{Cl}_F(K)$ . In other words if we write  $\mathcal{X}_{\ell, \tau_i}^\mathfrak{c}$  for the component of  $\mathcal{X}_{\ell, \tau_i}$  corresponding to  $\mathfrak{c}$ , we have  $\mathcal{X}_{\ell, \tau_i}^\mathfrak{c} = \mathcal{X}_\ell^\mathfrak{c} \cap \mathcal{X}_{\ell, \tau_i}$  for  $i = 1, 2$ .

**Proposition 4.3.2.** *In Case (2), the intersection matrix of  $\omega_{\tau_1}^c, \omega_{\tau_2}^c$  is given by*

$$\frac{1}{(\ell^{n_{\tau_1}} - 1)(\ell^{n_{\tau_2}} - 1)} \begin{pmatrix} 0 & s_\ell(K_{I_1, I_2})^c \\ s_\ell(K_{I_1, I_2})^c & 0 \end{pmatrix}.$$

*Proof.* Using the fact that  $\mathcal{X}_{\ell, \tau_1}^c$  and  $\mathcal{X}_{\ell, \tau_2}^c$  intersect transversally and that there are  $s_\ell(K_{I_1, I_2})^c$  intersection points, we see that  $\mathcal{X}_{\ell, \tau_1}^c \cdot \mathcal{X}_{\ell, \tau_2}^c = s_\ell(K_{I_1, I_2})^c$ .

As in §4.2, we have  $\omega_{\tau_2}^2|_{\mathcal{X}_{\ell, \tau_1}^c} \cong \Omega_{\mathcal{X}_{\ell, \tau_1}^c}^1$  and  $\omega_{\tau_1}^2|_{\mathcal{X}_{\ell, \tau_2}^c} \cong \Omega_{\mathcal{X}_{\ell, \tau_2}^c}^1$  in the rational Picard group. Applying the adjunction formula to each connected component and arguing as in §4.2, we find that  $\mathcal{X}_{\ell, \tau_i}^c \cdot \mathcal{X}_{\ell, \tau_i}^c = 0$ , and hence  $\omega_{\tau_i}^c \cdot \omega_{\tau_i}^c = 0$  for  $i = 1, 2$ .  $\square$

We let  $\mathcal{X}_0(\mathfrak{p})_\ell$  denote the special fiber of the integral model for  $\text{Sh}_{K_0(\mathfrak{p})}(G_{S, T})$  over  $\overline{\mathbb{F}}_\ell$ . We note that, since we have only changed level away from  $\ell$ ,  $\mathcal{X}_0(\mathfrak{p})_\ell$  is still smooth.

Recall there are two degeneracy maps

$$\pi_1, \pi_2 : \text{Sh}_{K_0(\mathfrak{p})}(G_{S, T}) \rightarrow \text{Sh}_K(G_{S, T}).$$

We also write  $\pi_1, \pi_2 : \mathcal{X}_0(\mathfrak{p})_\ell \rightarrow \mathcal{X}_\ell$  for the corresponding maps on the special fiber; both of these maps are finite étale.

**Definition 4.3.3.** Let  $x, x' \in \mathcal{X}_\ell(\overline{\mathbb{F}}_\ell)$ . We write  $x \sim_{\mathfrak{p}} x'$  if there exists  $\tilde{x}, \tilde{x}' \in \mathcal{X}_0(\mathfrak{p})_\ell(\overline{\mathbb{F}}_\ell)$  such that

$$\pi_1(\tilde{x}) = x, \pi_1(\tilde{x}') = x', \text{ and } \pi_2(\tilde{x}) = \pi_2(\tilde{x}').$$

We write  $\sim$  for the equivalence relation on  $\mathcal{X}_\ell(\overline{\mathbb{F}}_\ell)$  generated by  $\sim_{\mathfrak{p}}$ .

It is easy to see that for  $x \in \mathcal{X}_\ell(\overline{\mathbb{F}}_\ell)$ , the  $\sim$  equivalence class of  $x$  may be identified with  $Y_p(x)$  in the notation of §3.4. Note that the roles of  $p$  and  $\ell$  have been switched.

Recall we have the discrete Shimura sets  $G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K$ ,  $G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K_0(\mathfrak{p})$  and natural degeneracy maps

$$\pi_1, \pi_2 : G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K_0(\mathfrak{p}) \rightarrow G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K.$$

Similarly to the above we may define an equivalence relation on this set by specifying  $x \sim_{\mathfrak{p}} x'$  for  $x, x' \in G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K$  if there exists  $\tilde{x}, \tilde{x}' \in G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K_0(\mathfrak{p})$  such that

$$\pi_1(\tilde{x}) = x, \pi_1(\tilde{x}') = x', \text{ and } \pi_2(\tilde{x}) = \pi_2(\tilde{x}').$$

We write  $\sim$  for the equivalence relation generated by  $\sim_{\mathfrak{p}}$ , and we write  $Y_p(x)$  for the equivalence class of  $x$ .

**Proposition 4.3.4.** *For any  $x \in (G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K)^c$ , we have*

$$Y_p(x) = (G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K)^c.$$

*Proof.* Let  $G_{\Sigma_\infty}^1$  denote the kernel of  $\nu_{\Sigma_\infty} : G_{\Sigma_\infty} \rightarrow T_F$ . Then  $G_{\Sigma_\infty}^1$  is the derived group of  $G_{\Sigma_\infty}$  and is simply connected. Let  $g \in G_{\Sigma_\infty}(\mathbb{A}_f)$  be a representative of  $(G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K)^c$ . It is easy to see that

$$\begin{aligned} (G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K)^c &= G_{\Sigma_\infty}^1(\mathbb{Q}) \backslash G_{\Sigma_\infty}^1(\mathbb{A}_f) g K / K \\ &\cong G_{\Sigma_\infty}^1(\mathbb{Q}) \backslash G_{\Sigma_\infty}^1(\mathbb{A}_f) g K g^{-1} / g K g^{-1}; \end{aligned}$$

the isomorphism is given by right multiplication by  $g^{-1}$ .

Then  $Y_p(x)$  is identified with

$$G_{\Sigma_\infty}^1(\mathbb{Q}) \cap K^p \backslash G_{\Sigma_\infty}^1(\mathbb{Q}_p) g K g^{-1} / g K g^{-1} \subset G_{\Sigma_\infty}^1(\mathbb{Q}) \backslash G_{\Sigma_\infty}^1(\mathbb{A}_f) g K g^{-1} / g K g^{-1}.$$

By the strong approximation theorem,  $G_{\Sigma_\infty}^1(\mathbb{Q}) G_{\Sigma_\infty}^1(\mathbb{Q}_p)$  is dense in  $G_{\Sigma_\infty}^1(\mathbb{A}_f)$  and hence

$$G_{\Sigma_\infty}^1(\mathbb{Q}) \cap K^p \backslash G_{\Sigma_\infty}^1(\mathbb{Q}_p) g K g^{-1} / g K g^{-1} = G_{\Sigma_\infty}^1(\mathbb{Q}) \backslash G_{\Sigma_\infty}^1(\mathbb{A}_f) g K g^{-1} / g K g^{-1}.$$

$\square$

**4.4. Ihara's Lemma.** We now prove the main theorem of the section. We keep the notations of the previous subsection.

**Proposition 4.4.1.** *The map*

$$(4.4.1) \quad \pi_1^* + \pi_2^* : H^0(\mathcal{X}_\ell, \Omega_{\mathcal{X}_\ell/\overline{\mathbb{F}}_\ell}^2)^2 \rightarrow H^0(\mathcal{X}_0(\mathfrak{p})_\ell, \Omega_{\mathcal{X}_0(\mathfrak{p})_\ell/\overline{\mathbb{F}}_\ell}^2)$$

*is injective.*

*Proof.* Let  $(f_1, f_2)$  be an element of the kernel. We show  $f_1 = 0$ ; in fact we will show  $f_1|_{\mathcal{X}_\ell^c} = 0$  for all  $\mathfrak{c} \in \text{Cl}_F(K)$ . Assume for contradiction that  $f_1|_{\mathcal{X}_\ell^c} \neq 0$ . Suppose  $f_1$  has a zero at a point  $x \in \mathcal{X}_\ell^c(\overline{\mathbb{F}}_p)$ . Let  $y, z \in \mathcal{X}_\ell(\overline{\mathbb{F}}_p)$  be such that

$$\pi_1^{-1}(x) \cap \pi_2^{-1}(y) \neq \emptyset, \quad \pi_1^{-1}(z) \cap \pi_2^{-1}(y) \neq \emptyset.$$

Then since  $\pi_1^*(f_1) = -\pi_2^*(f_2)$ ,  $f_1$  has a zero at  $z$ . This implies  $f_1$  vanishes at every point of  $Y_p(x)$ . If  $f_1$  vanishes at an ordinary point of  $\mathcal{X}_\ell^c$ , then by Theorem 3.4.1,  $f_1|_{\mathcal{X}_\ell^c} = 0$ .

Therefore  $\text{div}(f_1|_{\mathcal{X}_\ell^c})$  is supported on the complement of the ordinary locus.

We first consider Case (1). Then  $\text{div}(f_1)$  is supported on  $\mathcal{X}_\ell^{\text{ss}}$ .

Let us consider  $D^c := \text{div}(f_1)|_{\mathcal{X}_\ell^c}$  for some  $\mathfrak{c} \in \text{Cl}_F(K)$ . Then  $D^c$  is supported on  $\mathcal{X}_\ell^c \cap \mathcal{X}_\ell^{\text{ss}}$ . Recall  $\mathcal{X}_\ell^c \cap \mathcal{X}_\ell^{\text{ss}} = \mathcal{X}_{\ell, \tau_1}^c \cup \mathcal{X}_{\ell, \tau_2}^c$  and each  $\mathcal{X}_{\ell, \tau_i}^c$  is a  $\mathbb{P}^1$ -fibration over  $(G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f)/K)^c$ . We let  $D_{\tau_1}^c$  (resp.  $D_{\tau_2}^c$ ) be the divisor corresponding to the sum of irreducible components in  $\mathcal{X}_{\ell, \tau_1}^c = \mathcal{X}_{\ell, \tau_1} \cap \mathcal{X}_\ell^c$  (resp.  $\mathcal{X}_{\ell, \tau_2}^c = \mathcal{X}_{\ell, \tau_2} \cap \mathcal{X}_\ell^c$ ). We claim there are non-negative integers  $a, b$  such that

$$D^c = aD_{\tau_1}^c + bD_{\tau_2}^c.$$

Indeed if  $f_1$  vanishes on some irreducible component of  $\mathcal{X}_{\ell, \tau_1}$  then by Proposition 4.3.4,  $f_1$  vanishes on every irreducible component of  $\mathcal{X}_{\ell, \tau_1}$ . We let  $h_{\tau_1} \in H^0(\mathcal{X}_\ell^c, (\omega_{\tau_2}^c)^{\ell^{n_{\tau_1}}} \otimes (\omega_{\tau_1}^c)^{-1})$  be a section with divisor  $D_{\tau_1}^c$ . The section  $h_{\tau_1}$  can be chosen to be compatible with prime-to- $\ell$  level structure and so  $\pi_1^*(h_{\tau_1}) = \pi_2^*(h_{\tau_1})$ . Then  $(f_1 h_{\tau_1}^{-1}, f_2 h_{\tau_1}^{-1}) \in \ker(\pi_1^* + \pi_2^*)$ . Repeating the argument we see that there exists  $a$  such that  $f_1 h_{\tau_1}^{-a}|_{\mathcal{X}_\ell^c}$  is non-vanishing on  $\mathcal{X}_{\ell, \tau_1}^c$ . We may use the same argument for  $\mathcal{X}_{\ell, \tau_2}^c$  to obtain  $b$  such that  $f_1 h_{\tau_2}^{-b}|_{\mathcal{X}_\ell^c}$  is non-vanishing on  $\mathcal{X}_{\ell, \tau_2}^c$ . It follows that the zero-locus of  $f_1 h_{\tau_1}^{-a} h_{\tau_2}^{-b}|_{\mathcal{X}_\ell^c}$  has codimension 2 and hence is empty; this proves the claim.

Now  $f_1$  corresponds to a section of the line bundle  $\Omega_{\mathcal{X}_\ell/\overline{\mathbb{F}}_\ell}^2$ , which is isomorphic to  $(\omega_{\tau_1})^2 \otimes (\omega_{\tau_2})^2$  up to a torsion element of the Picard group by Corollary 2.3.3. Furthermore  $D_{\tau_1}^c$  (resp.  $D_{\tau_2}^c$ ) is the divisor corresponding to a global section of  $(\omega_{\tau_2}^c)^{\ell^{n_{\tau_1}}} \otimes (\omega_{\tau_1}^c)^{-1}$  (resp.  $(\omega_{\tau_1}^c)^{\ell^{n_{\tau_2}}} \otimes (\omega_{\tau_2}^c)^{-1}$ ). Therefore the line bundle

$$((\omega_{\tau_2}^c)^{\ell^{n_{\tau_1}}} \otimes (\omega_{\tau_1}^c)^{-1})^{-a} \otimes ((\omega_{\tau_1}^c)^{\ell^{n_{\tau_2}}} \otimes (\omega_{\tau_2}^c)^{-1})^{-b} \otimes \Omega_{\mathcal{X}_\ell/\overline{\mathbb{F}}_\ell}^2$$

has a non-vanishing section; hence it must be the trivial bundle. It follows that

$$((\omega_{\tau_2}^c)^{\ell^{n_{\tau_1}}} \otimes (\omega_{\tau_1}^c)^{-1})^a \otimes ((\omega_{\tau_1}^c)^{\ell^{n_{\tau_2}}} \otimes (\omega_{\tau_2}^c)^{-1})^b \otimes (\omega_{\tau_1}^c)^{-2} \otimes (\omega_{\tau_2}^c)^{-2} \cong (\omega_{\tau_2}^c)^{a\ell^{n_{\tau_1}} - b - 2} \otimes (\omega_{\tau_1}^c)^{b\ell^{n_{\tau_2}} - a - 2}$$

is torsion in the Picard group of  $\mathcal{X}_\ell$ , hence pairs with any divisor to be 0. Since  $n_{\tau_1}, n_{\tau_2} > 0$  and  $\ell > 4$ , at least one of  $a\ell^{n_{\tau_1}} - b - 2, b\ell^{n_{\tau_2}} - a - 2$  is non-zero.

Recall the intersection matrix for the line bundles  $\omega_{\tau_1}^c$  and  $\omega_{\tau_2}^c$  on  $\mathcal{X}_\ell^c$  is given by

$$\begin{pmatrix} 0 & s_\ell(K)^c \\ s_\ell(K)^c & 0 \end{pmatrix}.$$

Thus  $(\omega_{\tau_2}^c)^{a\ell^{n_{\tau_1}} - b - 2} \otimes (\omega_{\tau_1}^c)^{b\ell^{n_{\tau_2}} - a - 2}$  pairs with either  $\omega_{\tau_1}^c$  or  $\omega_{\tau_2}^c$  to give a non-zero number. This is a contradiction.

We now consider Case (2). Then  $\text{div}(f_1)$  is supported on  $\mathcal{X}_{\ell, \tau_1}^c \cup \mathcal{X}_{\ell, \tau_2}^c$ . Let  $D^c := \text{div}(f_1)|_{\mathcal{X}_\ell^c}$ , and we write  $D_{\tau_1}^c$  (resp.  $D_{\tau_2}^c$ ) for the divisor  $\mathcal{X}_{\ell, \tau_1} \cap \mathcal{X}_\ell^c$  (resp.  $\mathcal{X}_{\ell, \tau_2} \cap \mathcal{X}_\ell^c$ ). Then

$$D^c = aD_{\tau_1}^c + bD_{\tau_2}^c$$

for some integers  $a, b$  since  $D_{\tau_1}^c$  and  $D_{\tau_2}^c$  are primitive divisors. Since  $D_{\tau_1}^c$  (resp.  $D_{\tau_2}^c$ ) is the vanishing locus of a section of  $(\omega_{\tau_1}^c)^{\ell^{n\tau_1}-1}$  (resp.  $(\omega_{\tau_2}^c)^{\ell^{n\tau_2}-1}$ ), the same argument as in case (1) shows that

$$(\omega_{\tau_1}^c)^{a\ell^{n\tau_1}-a-2} \otimes (\omega_{\tau_2}^c)^{b\ell^{n\tau_2}-b-2}$$

is a torsion element in the Picard group of  $\mathcal{X}_\ell$ . Since  $\ell > 4$ ,  $a\ell^{n\tau_1} - a - 2 \neq 0$  for any value of  $a$ . Thus intersecting with  $D_{\tau_2}^c$  gives a non-zero number. This is a contradiction.  $\square$

*Proof of Theorem 4.1.4.* Let

$$W := H^2(\mathrm{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}, \quad W_0(\mathfrak{p}) := H^2(\mathrm{Sh}_{K_0(\mathfrak{p})}(G_{S,T})_{\overline{\mathbb{Q}}}, k_\lambda)_{\mathfrak{m}}.$$

The image  $\mathbf{T}_{\mathrm{RU}\{\mathfrak{p}\}}^W$  of  $\mathbf{T}_{\mathrm{RU}\{\mathfrak{p}\}}$  in  $\mathrm{End}_{k_\lambda}(W)$  is a local Artinian ring and  $\mathfrak{m}^i W$  is a finite decreasing filtration of  $W$ . By the freeness result in Proposition 4.1.2 (2), each graded piece  $\mathfrak{m}^i W / \mathfrak{m}^{i+1} W$  is a quotient of two  $\mathbf{T}_{\mathrm{RU}\{\mathfrak{p}\}}[\mathcal{G}_{\tilde{F}}]$ -lattices in  $H^2(\mathrm{Sh}_K(G_{S,T})_{\overline{\mathbb{Q}}}, \mathcal{O}_{\mathbf{E}_\lambda})_{\mathfrak{m}}$ . By Lemma [Dim05, Lemma 6.5] and the Eichler–Shimura congruence relation proved by Nekovar [Nek18, A.6], the irreducible subquotients of  $W$  are all isomorphic as Galois representations of  $\mathcal{G}_{\tilde{F}}$ . The same statement holds for the irreducible subquotients of  $W_0(\mathfrak{p})$ . Using Faltings’ comparison theorem [Fal89], we may therefore check the injectivity on the last graded pieces of the Fontaine–Laffaille filtration. By the degeneracy of the Hodge–de Rham spectral sequence proved in [DI87], this follows from the assertion of Proposition 4.4.1.  $\square$

*Remark 4.4.2.* In Case (1), there is an alternative argument to prove Theorem 4.1.4. We are grateful to the anonymous referee for explaining this to us. In this case, we assume the two Frobenius eigenvalues at  $\ell$  are distinct. Then using the Tate conjecture for  $\mathcal{S}_K(G_{S,T})$ , which can be proved as in [TX19], one can deduce that  $(\pi_{1*}, \pi_{2*}) : W_0(\mathfrak{p})^* \rightarrow (W^*)^2$  is surjective by reducing it to the surjectivity in the case of discrete Shimura sets which is well known.

## 5. ABEL–JACOBI MAP AND GEOMETRIC LEVEL RAISING

In this section we will use the results from the previous sections to construct classes in the motivic cohomology of quaternionic Shimura varieties. We begin by recalling the definition of higher Chow groups and the associated cycle class maps.

**5.1. Higher Chow groups and  $\ell$ -adic cycle class maps.** Let  $X$  be a smooth variety over a field  $k$  and let  $\Delta^n$  denote the standard  $n$ -simplex  $\mathrm{Spec} k[x_0, \dots, x_n] / (\sum_{i=0}^n x_i - 1)$ . For integers  $n, r$ , we define  $z^r(X, n)$  to be the free abelian group generated by the integral closed subvarieties  $Z$  of  $X \times \Delta^n$  such that for any face  $F \subset \Delta^n$  we have

$$\mathrm{codim}_{X \times F}(Z \cap (X \times F)) \geq r.$$

The groups  $z^r(X, n)$  fit into a complex

$$(5.1.1) \quad \dots \rightarrow z^r(X, n) \rightarrow z^r(X, n-1) \rightarrow \dots \rightarrow z^r(X, 0) \rightarrow 0$$

where the differential is given by taking the alternating sum of the induced face maps. The *higher Chow group*  $\mathrm{Ch}^r(X, n)$  is defined to be the  $n^{\mathrm{th}}$  homology of the above complex. It is easy to see that  $\mathrm{Ch}^r(X, 0)$  is the standard Chow group of codimension  $r$  cycles on  $X$ . By [Voe02], we have an isomorphism

$$\mathrm{Ch}^j(X, 2j-i) \cong H_{\mathcal{M}}^i(X, \mathbb{Z}(j))$$

where  $H_{\mathcal{M}}^i(X, \mathbb{Z}(j))$  is the motivic cohomology group of [SV00].

We may also consider a variant of this construction by introducing coefficients. Let  $R$  be any ring and we let  $\mathrm{Ch}^r(X, n, R)$  denote the  $n^{\mathrm{th}}$  homology of the sequence (5.1.1) tensored with  $R$ . As before there is an isomorphism  $\mathrm{Ch}^j(X, 2j-i, R) \cong H_{\mathcal{M}}^i(X, R(j))$ ; see [Voe02, Corollary 2]. From now on, we will work with higher Chow groups.

Let  $\ell$  be a prime which is invertible in  $k$ , then there is an  $\ell$ -adic cycle class map

$$\mathrm{Ch}^j(X, 2j - i) \rightarrow \mathrm{H}_{\mathrm{cont}}^i(X, \mathbb{Z}_\ell(j))$$

where  $\mathrm{H}_{\mathrm{cont}}^i(X, \mathbb{Z}_\ell(j))$  is the continuous étale cohomology defined by [Jan88]. This coincides with the usual cycle class map when  $i = 2j$ .

Similarly if  $k_\lambda$  is a finite extension of  $\mathbb{F}_\ell$ , there is a cycle class map

$$\mathrm{Ch}^j(X, 2j - i, k_\lambda) \rightarrow \mathrm{H}^i(X, k_\lambda(j)).$$

Let  $Y \rightarrow X$  denote a fibration with connected smooth fibers of dimension  $s$  and  $R$  any coefficient ring. Then taking preimages of cycles under the projection  $Y \times \Delta^n \rightarrow X \times \Delta^n$  induces a map

$$(5.1.2) \quad \mathrm{Ch}^r(X, n, R) \rightarrow \mathrm{Ch}^r(Y, n, R).$$

Finally let  $Z \hookrightarrow X$  be a closed immersion where  $Z$  is smooth and of codimension  $t$  in  $X$ . Then pushforward of cycles along  $Z \times \Delta^n \rightarrow X \times \Delta^n$  induces a map

$$(5.1.3) \quad \mathrm{Ch}^r(Z, n, R) \rightarrow \mathrm{Ch}^{r+t}(X, n, R).$$

**5.2. Motivic cohomology of surfaces and dual graphs.** We are particularly interested in the case of surfaces. In this case, the motivic cohomology  $\mathrm{H}_{\mathcal{M}}^3(X, \mathbb{Z}(2)) \cong \mathrm{Ch}^2(X, 1)$  is given by the homology of the following sequence (see [Sch00] for example):

$$(5.2.1) \quad K_2(k(X)) \xrightarrow{\partial} \bigoplus_{S \subset X} k(S)^\times \xrightarrow{\mathrm{div}} \bigoplus_{x \in X} \mathbb{Z}.$$

Here the middle sum runs over the set of irreducible curves  $S \subset X$  and  $k(X)$  (resp.  $k(S)$ ) denotes the field of rational functions on  $X$  (resp.  $S$ ). The term  $K_2(k(X))$  denotes the second Milnor  $K$ -group of the rational function field  $k(X)$  and the  $S$ -component of the map  $\partial$  is the tame symbol associated to the valuation  $\mathrm{ord}_S$ . The map  $\mathrm{div}$  sends a rational function  $f$  on  $S$  to its divisor  $\mathrm{div}(f)$ .

There is a special case where we can understand a part of  $\mathrm{Ch}^2(X, 1)$  in a purely combinatorial way. If  $k'/k$  is a finite extension, we write  $\mathbb{P}_{k'/k}^1$  for the projective line  $\mathbb{P}_{k'}^1$  over  $k'$  considered as a  $k$ -scheme. Therefore  $\mathbb{P}_{k'/k}^1 \otimes_k \bar{k}$  can be identified with  $[k' : k]$  copies of  $\mathbb{P}_{\bar{k}}^1$  corresponding to the embeddings  $k' \rightarrow \bar{k}$ . Let  $Y \subset X$  be a codimension 1 subvariety satisfying the following conditions:

- (1) Each irreducible component  $S$  of  $Y$  is isomorphic to  $\mathbb{P}_{k_S/k}^1$  where  $k_S$  is a finite extension of  $k$ .
- (2) Any two irreducible components of  $Y_{\bar{k}}$  intersect transversally and no three components have a common intersection point.

The non-smooth points of  $Y_{\bar{k}}$  are the intersection points of the components in  $Y_{\bar{k}}$  and it is naturally a closed subscheme of  $X$  defined over  $k$ . We note that by (1), the irreducible components of  $Y_{\bar{k}}$  are isomorphic to  $\mathbb{P}_{\bar{k}}^1$ .

**Definition 5.2.1.** We define  $\mathrm{Ch}_Y^2(X, 1)$  to be the subgroup of  $\mathrm{Ch}^2(X, 1)$  supported on  $Y$ . In other words, it is generated by elements of the form  $\bigoplus_{S \subset X} f_S$  where  $f_S \in k(S)^\times$  is trivial unless  $S \subset Y$ .

We now describe how  $\mathrm{Ch}_Y^2(X, 1)$  can be interpreted in terms of the combinatorics of the configuration of  $\mathbb{P}_{\bar{k}}^1$ 's on  $Y_{\bar{k}}$ .

**Definition 5.2.2.** The dual graph  $\mathfrak{G}$  associated to  $Y_{\bar{k}}$  is the unoriented graph defined by the following:

- (1) The set of vertices  $\mathcal{V}$  is identified with the set of irreducible components of  $Y_{\bar{k}}$ . For  $i \in \mathcal{V}$ , we let  $S_i$  denote the corresponding irreducible component.
- (2) The set of edges  $\mathcal{E}$  is identified with the set of intersection points of two irreducible components in  $Y_{\bar{k}}$ , where an edge  $e$  connects  $i, j \in \mathcal{V}$  if  $e \in S_i \cap S_j$ .

The Galois group  $\mathrm{Gal}(\bar{k}/k)$  naturally acts on  $\mathfrak{G}$  and hence on the homology  $\mathrm{H}_1(\mathfrak{G}, \mathbb{Z})$  of the graph  $\mathfrak{G}$ . We will define a map

$$\Theta : \mathrm{H}_1(\mathfrak{G}, \mathbb{Z})^{\mathrm{Gal}(\bar{k}/k)} \rightarrow \mathrm{Ch}_Y^2(X, 1).$$

To do this, we first calculate  $H_1(\mathfrak{G}, \mathbb{Z})$ . Consider the bouquet of circles  $\tilde{\mathfrak{G}}$  given by contracting all the elements of  $\mathcal{Y}$ . For each  $e \in \mathcal{E}$  we choose an orientation of  $e$ , i.e. an ordering  $(v^1(e), v^2(e))$  of the two vertices adjacent to  $e$ . This choice determines an isomorphism

$$H_1(\tilde{\mathfrak{G}}, \mathbb{Z}) \cong \mathbb{Z}^{\mathcal{E}}.$$

We may then identify  $H_1(\mathfrak{G}, \mathbb{Z})$  as the subgroup of  $\mathbb{Z}^{\mathcal{E}}$  corresponding to the kernel of the map

$$d : \mathbb{Z}^{\mathcal{E}} \rightarrow \mathbb{Z}^{\mathcal{Y}}$$

given by

$$e \mapsto v_1(e) - v_2(e).$$

The map  $\Theta$  can then be defined as follows. Fix a basis of the free  $\mathbb{Z}$ -module  $H_1(\mathfrak{G}, \mathbb{Z})^{\text{Gal}(\bar{k}/k)}$  and let  $m := (m_e)_{e \in \mathcal{E}} \in H_1(\mathfrak{G}, \mathbb{Z})^{\text{Gal}(\bar{k}/k)}$  be an element of this basis. Let  $i \in \mathcal{Y}$ . We let  $Z(i), P(i) \subset \mathcal{E}$  be the two subsets of  $\mathcal{E}$  defined by

$$Z(i) = \{e \in \mathcal{E} \mid v_1(e) = i\}$$

$$P(i) = \{e \in \mathcal{E} \mid v_2(e) = i\}.$$

By definition,  $Z(i)$  and  $P(i)$  are identified with subsets of points on  $S_i$ .

**Lemma 5.2.3.** *There exists an element  $(f_S^m)_{S \subset Y} \in \bigoplus_{S \subset Y} k(S)^\times$  such that its image  $(f_i^m)_{i \in \mathcal{Y}} \in \bigoplus_{i \in \mathcal{Y}} \bar{k}(S_i)^\times$  satisfies*

$$\text{div}(f_i^m) = \sum_{e \in Z(i)} m_e(e) - \sum_{e \in P(i)} m_e(e).$$

*Proof.* Let  $S \cong \mathbb{P}_{k_S/k}^1$  be an irreducible component of  $Y$ . Let  $\mathcal{Y}_S \subset \mathcal{Y}$  denote the set of irreducible components of  $S_{\bar{k}}$ . We consider the divisor

$$\sum_{i \in \mathcal{Y}_S} \sum_{e \in Z(i)} m_e(e) - \sum_{i \in \mathcal{Y}_S} \sum_{e \in P(i)} m_e(e)$$

on  $S_{\bar{k}}$ . Since  $m \in \ker(d)$ , this divisor has degree 0 on every component  $S_i$ . Moreover, it is Galois invariant, hence there exists a function  $f_S^m \in k(S)^\times$  with this divisor (we note that  $k(S)$  is isomorphic to the field of rational functions in one variable  $k_S(x)$  over  $k_S$ ). Repeating the argument for all components  $S \subset Y$  proves the lemma.  $\square$

Let  $(f_S^m)_{S \subset Y}$  be an element as in the lemma above. Then we have

$$\begin{aligned} \sum_{S \subset Y} \text{div}(f_S^m) &= \sum_{i \in \mathcal{Y}} \text{div}(f_i^m) \\ &= \sum_{i \in \mathcal{Y}} \sum_{e \in Z(i)} m_e(e) - \sum_{i \in \mathcal{Y}} \sum_{e \in P(i)} m_e(e) \\ &= \sum_{e \in \mathcal{E}} m_e(e) - \sum_{e \in \mathcal{E}} m_e(e) \\ &= 0. \end{aligned}$$

Thus by (5.2.1), the tuple  $(f_S^m)_{S \subset Y}$  gives an element of  $\text{Ch}_Y^2(X, 1)$  and we define  $\Theta(m)$  to be this element. Repeating for the other basis elements and extending linearly gives the definition of  $\Theta$ .

In fact the map  $\Theta$  is “essentially” surjective in the following sense. Note that the choice of  $f_S^m$  is well-defined up to multiplication by an element of  $k(S)^\times$  whose divisor is 0, i.e. an element of  $k_S^\times$ . It follows that  $\Theta$  induces a map

$$(5.2.2) \quad \tilde{\Theta} : \left( \bigoplus_{S \subset Y} k_S^\times \right) \oplus H_1(\mathfrak{G}, \mathbb{Z})^{\text{Gal}(\bar{k}/k)} \rightarrow \text{Ch}_Y^2(X, 1).$$

**Proposition 5.2.4.** *The map  $\tilde{\Theta}$  is surjective.*

*Proof.* Let  $(f_S)_{S \subset Y} \in \bigoplus_{S \subset Y} k(S)^\times$  be a representative of an element of  $\text{Ch}_Y^2(X, 1)$ . We write  $(f_i)_{i \in \mathcal{Y}}$  for its image in  $\bigoplus_{i \in \mathcal{Y}} \bar{k}(S_i)$ . Since  $\sum_{S \subset Y} \text{div}(f_S) = \sum_{i \in \mathcal{Y}} \text{div}(f_i) = 0$ , the divisor  $\text{div}(f_i)$  can only be supported on the intersection of  $S_i$  with some other  $S_j$ , i.e. on points corresponding to elements of  $\mathcal{E}$ .

For each  $e \in \mathcal{E}$ , let  $S_i$  denote the divisor corresponding to  $v_1(e)$  and define  $m_e \in \mathbb{Z}$  to be the order of the zero of  $f_1$  at  $e$ . Reversing the above argument, we see that  $m = (m_e)_{e \in \mathcal{E}}$  defines an element of  $H_1(\mathfrak{G}, \mathbb{Z})$  which is Galois invariant since  $\text{div}(f_S)$  is.

By definition the  $S^{\text{th}}$ -component of  $\Theta(m)$  has the same divisor as  $f_S$ , thus we may modify  $\Theta(m)$  by an element of  $\bigoplus_{S \subset Y} k_S^\times$  to get the element  $(f_S)_{S \subset Y}$ .  $\square$

We will consider an analogue of this construction with torsion coefficients. Indeed upon tensoring with  $k_\lambda$  we obtain a map

$$(5.2.3) \quad \tilde{\Theta}_{k_\lambda} : \left( \bigoplus_{S \subset Y} k_S^\times \otimes_{\mathbb{Z}} k_\lambda \right) \oplus H_1(\mathfrak{G}, k_\lambda)^{\text{Gal}(\bar{k}/k)} \rightarrow \text{Ch}_Y^2(X, 1, k_\lambda)$$

and same proof as Proposition 5.2.4 shows that  $\tilde{\Theta}_{k_\lambda}$  is surjective. Here we consider the abelian group  $\bigoplus_{S \subset Y} k_S^\times$  as a  $\mathbb{Z}$ -module in taking the tensor product, and in defining  $\text{Ch}^2(X, 1, k_\lambda)$  we may take the homology of the sequence (5.2.1) tensored with  $k_\lambda$ ; the subgroup  $\text{Ch}_Y^2(X, 1, k_\lambda)$  is then defined in the same way as Definition 5.2.1.

**5.3. Motivic cohomology and level raising.** We keep the notations of §2.1, but now  $F$  will be a totally real field with  $[F : \mathbb{Q}] = g$  even and  $p$  will be a prime which is *inert* in  $F$ ; we let  $\mathfrak{p}$  denote the unique prime of  $F$  above  $p$ . We let  $B$  be a quaternion algebra over  $F$  whose ramification set does not intersect with  $\Sigma_p \cup \Sigma_\infty$ . We will apply the construction of the previous subsections for the case  $\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}$ . Specifically we are interested in the cycle class map:

$$(5.3.1) \quad \text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda) \rightarrow H^{g+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, k_\lambda(g/2 + 1)).$$

Let  $\Pi$  be an irreducible cuspidal automorphic representation of  $\text{GL}_2(F)$  defined over  $\mathbf{E}$  as in §4.1. Let  $\lambda$  be a prime of  $\mathbf{E}$  with underlying rational prime  $\ell$  satisfying Assumption 4.1.1. Recall for  $R$  a finite set of places of  $F$  away from which  $K$  is hyperspecial and  $\Pi$  is unramified, we have the prime-to- $R$  Hecke algebra  $\mathbf{T}_R$  and a homomorphism  $\phi_R^\Pi : \mathbf{T}_R \rightarrow \mathcal{O}_{\mathbf{E}}$  induced by the Hecke eigenvalues of  $\Pi$ . We assume  $\mathfrak{p} \notin R$  and we write  $\mathfrak{m}_R \subset \mathbf{T}_R$  (resp.  $\mathfrak{m} \subset \mathbf{T}_{R \cup \{\mathfrak{p}\}}$ ) for the maximal ideal given by  $(\phi_R^\Pi)^{-1}(\lambda)$  (resp.  $(\phi_{R \cup \{\mathfrak{p}\}}^\Pi)^{-1}(\lambda)$ ). Then  $\mathbf{T}_{R \cup \{\mathfrak{p}\}}$  acts on  $\text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$  and the map (5.3.1) is equivariant for this action. Localizing at  $\mathfrak{m}$ , we obtain

$$\text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}} \rightarrow H^{g+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, k_\lambda(g/2 + 1))_{\mathfrak{m}}.$$

By the Hochschild–Serre spectral sequence and Proposition 4.1.2, there is an isomorphism

$$H^1(\mathbb{F}_{p^g}, H^g(\mathcal{S}_K(G)_{\bar{\mathbb{F}}_p}, k_\lambda(g/2 + 1))_{\mathfrak{m}}) \cong H^{g+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, k_\lambda(g/2 + 1))_{\mathfrak{m}},$$

where  $H^1(\mathbb{F}_{p^g}, -)$  denotes the Galois cohomology. We let

$$\text{AJ}_{\mathfrak{m}} : \text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}} \rightarrow H^1(\mathbb{F}_{p^g}, H^g(\mathcal{S}_K(G)_{\bar{\mathbb{F}}_p}, k_\lambda(g/2 + 1))_{\mathfrak{m}})$$

denote the induced map which we call the *Abel–Jacobi map*.

We would like to show that  $\text{AJ}_{\mathfrak{m}}$  is surjective modulo  $\mathfrak{m}$ . In fact we will identify a certain subgroup

$$\text{Ch}_{\text{Ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda) \subset \text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$$

such that  $\text{AJ}_{\mathfrak{m}}$  restricted to this subgroup already surjects. Moreover as suggested by the notation, this subgroup is related to level raising. The group  $\text{Ch}_{\text{Ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$  is defined using the geometry of Goren–Oort cycles as follows. We note that we may take  $k_0 = \mathbb{F}_{p^g}$  in this section.



Let  $\mathfrak{B} := \mathfrak{B}(\emptyset, g/2 - 1)$  as in §2.5; thus  $|\mathfrak{B}| = \binom{g}{g/2-1}$ . Let  $\mathfrak{a} \in \mathfrak{B}$ , then by the discussion in §2.5, the Goren–Oort cycle  $Z_\emptyset(\mathfrak{a})$  is a  $(g/2 - 1)$ -iterated  $\mathbb{P}^1$ -bundle over the Shimura surface  $\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}}$  where  $\mathfrak{T} = \emptyset$ . We have  $|\Sigma_\infty - \emptyset_\mathfrak{a}| = 2$  and we write  $\Sigma_\infty - \emptyset_\mathfrak{a} = \{\tau_i, \tau_j\}$ .

The Goren–Oort divisors  $\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}, \tau_i}$ ,  $\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}, \tau_j}$  are  $\mathbb{P}^1$ -bundles over the discrete Shimura sets  $\mathcal{S}_K(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}$  and  $\mathcal{S}_K(G_{\emptyset, \tau_j, \mathfrak{T}_\mathfrak{a}, \tau_j})_{\mathbb{F}_{p^g}}$  via maps  $\pi_{\tau_i}$ ,  $\pi_{\tau_j}$  respectively. These Shimura sets are isomorphic over  $\mathbb{F}_{p^g}$  and their  $\overline{\mathbb{F}}_p$ -points are isomorphic to

$$G_{\Sigma_\infty}(\mathbb{Q}) \backslash G_{\Sigma_\infty}(\mathbb{A}_f) / K.$$

By Proposition 2.5.1, the intersection of these two divisors  $\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}, \{\tau_i, \tau_j\}}$  is identified with the discrete Shimura set  $\mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}$  and the induced diagram

$$\mathcal{S}_K(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}} \xleftarrow{\pi_{\tau_i}} \mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}} \xrightarrow{\eta_{\tau_j} \circ \pi_{\tau_j}} \mathcal{S}_K(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}$$

is identified with the Hecke correspondence for  $\mathcal{S}_K(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}$ .

We may apply the construction of §5.2 in this setting to obtain a map

$$H_1(\mathfrak{G}, k_\lambda)^{\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})} \rightarrow \text{Ch}^2(\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}}, 1, k_\lambda),$$

where  $\mathfrak{G}$  is the dual graph associated to the configuration of projective lines in  $\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\overline{\mathbb{F}}_p, \tau_i}$  and  $\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\overline{\mathbb{F}}_p, \tau_j}$ . We may describe  $H_1(\mathfrak{G}, k_\lambda)$  more explicitly as follows. For any finite set  $S$ , we write  $\Gamma(S, k_\lambda)$  for the abelian group of  $k_\lambda$ -valued functions on  $S$ . The maps  $\pi_{\tau_i}$  and  $\eta_{\tau_j} \circ \pi_{\tau_j}$  induce maps

$$\pi_{\tau_i*}, (\eta_{\tau_j} \circ \pi_{\tau_j})_* : \Gamma(\mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p), k_\lambda) \rightarrow \Gamma(\mathcal{S}_K(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p), k_\lambda)$$

Then  $H_1(\mathfrak{G}, k_\lambda)$  is identified with

$$(5.3.2) \quad \ker \left( (\pi_{\tau_i*}, (\eta_{\tau_j} \circ \pi_{\tau_j})_*) : \Gamma(\mathcal{S}_{K_0(\mathfrak{p})}(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p), k_\lambda) \rightarrow \Gamma(\mathcal{S}_K(G_{\emptyset, \tau_i, \mathfrak{T}_\mathfrak{a}, \tau_i})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p), k_\lambda)^2 \right).$$

We write  $K(\mathfrak{a})$  for this kernel, and we write  $\Theta(\mathfrak{a})$  for the map

$$(5.3.3) \quad \Theta(\mathfrak{a}) : K(\mathfrak{a})^{\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})} \rightarrow \text{Ch}^2(\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}}, 1, k_\lambda)$$

which is equivariant for the action of  $\mathbf{T}_{\text{RU}\{\mathfrak{p}\}}$ .

Using (5.1.2) and (5.1.3) we obtain a map

$$(5.3.4) \quad \text{Ch}^2(\mathcal{S}_K(G_{\emptyset, \mathfrak{T}_\mathfrak{a}})_{\mathbb{F}_{p^g}}, 1, k_\lambda) \xrightarrow{(5.1.2)} \text{Ch}^2(Z_\emptyset(\mathfrak{a}), 1, k_\lambda) \xrightarrow{(5.1.3)} \text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$$

Composing (5.3.4) and (5.3.3) and taking the direct sum over  $\mathfrak{a} \in \mathfrak{B}$  we obtain a map:

$$(5.3.5) \quad \bigoplus_{\mathfrak{a} \in \mathfrak{B}} K(\mathfrak{a})^{\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})} \rightarrow \text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$$

**Definition 5.3.1.** We define the level raising subgroup

$$\text{Ch}_{\text{lr}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda) \subset \text{Ch}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)$$

to be the image of (5.3.5).

Localizing (5.3.5) at  $\mathfrak{m}$  and composing with the map  $\text{AJ}_\mathfrak{m}$  we obtain a map:

$$\Psi_\mathfrak{m} : \bigoplus_{\mathfrak{a} \in \mathfrak{B}} K(\mathfrak{a})_\mathfrak{m}^{\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})} \rightarrow H^1(\mathbb{F}_{p^g}, H^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(g/2 + 1))_\mathfrak{m})$$

**Definition 5.3.2.** A prime  $p$  is a  $\lambda$ -level raising prime (with respect to  $F, B, \Pi, K$ ) if the following four conditions are satisfied:

- (1)  $p$  is inert in  $F$  and coprime to  $\text{R} \cup \text{R}_\lambda$ .
- (2)  $\ell \nmid \prod_{i=1}^g (p^{2gi} - 1)$ .

- (3)  $\phi_{\mathbb{R}}^{\Pi}(T_{\mathfrak{p}})^2 \equiv (p^g + 1)^2 \pmod{\lambda}$  and  $\phi_{\mathbb{R}}^{\Pi}(S_{\mathfrak{p}}) \equiv 1 \pmod{\lambda}$ .  
(4) For any  $\mathfrak{a} \in \mathfrak{B}$  and  $\tau_i \in \Sigma_{\infty} - \emptyset_{\mathfrak{a}}$ , we have  $\ell \nmid (p^{g n_{K_{\infty}}} - 1)$  where  $n_{K_{\infty}}$  is the order of  $\sigma_p^g$  acting on  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, \tau_i, \Gamma_{\mathfrak{a}, \tau_i}})_{\mathbb{F}_{p^g}}(\overline{\mathbb{F}}_p)$ .

*Remark 5.3.3.* As in [LT20, Remark 4.6], it can be shown that for  $\ell$  sufficiently large, there are infinitely many  $\lambda$ -level raising primes with positive density as long as there are rational primes inert in  $F$  and  $\lambda$  satisfies Assumption 4.1.1.

**Lemma 5.3.4.** *Assume  $p$  is a  $\lambda$ -level raising prime. Then  $\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})$  acts trivially on  $K(\mathfrak{a})_{\mathfrak{m}}$  for any  $\mathfrak{a} \in \mathfrak{B}$ .*

*Proof.* By the definition of the Galois action in §2.5 and Definition 5.3.2 (3),  $\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})$  acts trivially on  $K(\mathfrak{a})/\mathfrak{m}$ . By Assumption 4.1.1 (1), the trivial action on  $K(\mathfrak{a})_{\mathfrak{m}}$  is the only lifting of the trivial action on  $K(\mathfrak{a})/\mathfrak{m}$ ; the lemma follows.  $\square$

We will need to make the following additional assumption:

**Assumption 5.3.5.**  $H^g(\text{Sh}_K(G)_{\overline{\mathbb{Q}}}, k_{\lambda})/\mathfrak{m}$  has dimension  $2^g \dim(\Pi_B)^K$  over  $k_{\lambda}$ , where  $\Pi_B$  denotes the Jacquet–Langlands transfer of  $\Pi$  to  $B \otimes_{\mathbb{Q}} \mathbb{A}$ .

**Theorem 5.3.6.** *Let  $p$  be a  $\lambda$ -level raising prime and suppose that Assumptions 4.1.1 and 5.3.5 are satisfied. Then the induced map*

$$(5.3.6) \quad \bigoplus_{\mathfrak{a} \in \mathfrak{B}} K(\mathfrak{a})/\mathfrak{m} \rightarrow H^1(\mathbb{F}_{p^g}, H^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_{\lambda}(g/2 + 1))/\mathfrak{m})$$

*is surjective.*

*Remark 5.3.7.* In the case when  $[F : \mathbb{Q}] = 2$  and  $B = \text{GL}_2(F)$ , the surjectivity of  $\Psi_{\mathfrak{m}}$  implies level raising for Hilbert modular forms. Indeed in this case there is a unique  $\mathfrak{a} \in \mathfrak{B}$  and the description of  $K(\mathfrak{a})$  in (5.3.2) shows that it is identified under the Jacquet–Langlands correspondence with the space of  $\text{mod } \ell$  Hilbert modular forms of parallel weight 2 and level  $K_0(\mathfrak{p})$  which are new at  $\mathfrak{p}$ . In particular  $\mathbf{T}_{\text{RU}\{\mathfrak{p}\}}$  acts on  $K(\mathfrak{a})$  via the  $\mathfrak{p}$ -new quotient as in [Rib88], whereas it is well known  $\mathbf{T}_{\text{RU}\{\mathfrak{p}\}}$  acts on  $H^1(\mathbb{F}_{p^g}, H^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_{\lambda}(g/2 + 1)))$  via the  $\mathfrak{p}$ -old quotient.

Before embarking on the proof, we state an immediate corollary of Theorem 5.3.6, which is Theorem 1.1.1 of the introduction.

**Corollary 5.3.8.** *The map  $\text{AJ}_{\mathfrak{m}} \pmod{\mathfrak{m}}$  restricted to  $\text{Ch}_{\text{Ir}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_{\lambda})/\mathfrak{m}$  is surjective.*

$\square$

**5.4. Proof of Theorem 5.3.6.** The rest of this section will be devoted to the proof of Theorem 5.3.6. For notational convenience, we will write  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}$  for what was denoted  $\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}, \Gamma_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}$  in the previous subsection.

Recall for  $\mathfrak{a} \in \mathfrak{B}(\emptyset, g/2 - 1)$  we have the correspondence:

$$\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}})_{\mathbb{F}_{p^g}} \xleftarrow{\pi_{\mathfrak{a}}} Z_{\emptyset}(\mathfrak{a}) \hookrightarrow \mathcal{S}_K(G)_{\mathbb{F}_{p^g}}$$

where  $\pi_{\mathfrak{a}}$  is a  $(g/2 - 1)$ -iterated  $\mathbb{P}^1$ -bundle. This induces a Gysin map:

$$\text{Gys}(\mathfrak{a}) : H^*(\mathcal{S}_K(G_{\emptyset_{\mathfrak{a}}})_{\overline{\mathbb{F}}_p}, k_{\lambda}(2)) \rightarrow H^{*+g-2}(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_{\lambda}(g/2 + 1)).$$

Then the map  $\bigoplus_{\mathfrak{a} \in \mathfrak{B}} K(\mathfrak{a})_{\mathfrak{m}} \rightarrow H^1(\mathbb{F}_{p^g}, H^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_{\lambda}(g/2 + 1))_{\mathfrak{m}})$  factors as

$$\begin{array}{ccc}
 \mathrm{Ch}_{\mathrm{lr}}^{g/2+1}(\mathcal{S}_K(G)_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}} & \longrightarrow & \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^g(\mathcal{S}_K(G)_{\overline{\mathbb{F}}_p}, k_\lambda(g/2+1))_{\mathfrak{m}}) \\
 \uparrow & & \uparrow \sum_{\mathfrak{a} \in \mathfrak{B}} \mathrm{Gys}(\mathfrak{a})_{\mathfrak{m}} \\
 \bigoplus_{\mathfrak{a} \in \mathfrak{B}} \mathrm{K}(\mathfrak{a})_{\mathfrak{m}} & \longrightarrow & \bigoplus_{\mathfrak{a} \in \mathfrak{B}} \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(\mathcal{S}_K(G_{\theta_{\mathfrak{a}}})_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}})
 \end{array}$$

where the bottom map is the composition of the direct sums of the maps  $\Theta(\mathfrak{a})_{\mathfrak{m}}$  and the Abel–Jacobi maps

$$\mathrm{AJ}(\mathfrak{a})_{\mathfrak{m}} : \mathrm{Ch}^2(\mathcal{S}_K(G_{\theta_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}, 1, k_\lambda)_{\mathfrak{m}} \rightarrow \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(\mathcal{S}_K(G_{\theta_{\mathfrak{a}}})_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}).$$

We write  $\Psi(\mathfrak{a})_{\mathfrak{m}} : \mathrm{K}(\mathfrak{a})_{\mathfrak{m}} \rightarrow \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(\mathcal{S}_K(G_{\theta_{\mathfrak{a}}}), k_\lambda(2))_{\mathfrak{m}})$  for this composition.

**Proposition 5.4.1.** *The map  $\sum_{\mathfrak{a} \in \mathfrak{B}} \mathrm{Gys}(\mathfrak{a}) \pmod{\mathfrak{m}}$  is surjective.*

*Proof.* The proof is the same as in [LT20, Proof of Theorem 4.7]. The idea is to show that  $\sum_{\mathfrak{a} \in \mathfrak{B}} \mathrm{Gys}(\mathfrak{a}) \pmod{\mathfrak{m}}$  is injective; the surjectivity then follows by a dimension count.  $\square$

*Remark 5.4.2.* The proof of Proposition 5.4.1 uses Definition 5.3.2 (2), Assumption 5.3.5, and the freeness result in Proposition 4.1.2 in an essential way.

Thus in order to prove Theorem 5.3.6, it suffices to prove the following proposition.

**Proposition 5.4.3.** *For each  $\mathfrak{a} \in \mathfrak{B}$ , the map*

$$\Psi(\mathfrak{a})_{\mathfrak{m}} : \mathrm{K}(\mathfrak{a})_{\mathfrak{m}} \rightarrow \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(\mathcal{S}_K(G_{\theta_{\mathfrak{a}}}), k_\lambda(2))_{\mathfrak{m}})$$

*is surjective.*

Fix  $\mathfrak{a} \in \mathfrak{B}$ . We let  $\Sigma_\infty - \mathrm{S}_\infty = \{\tau_i, \tau_j\}$ . To ease notation, we write  $X = \mathcal{S}_K(G_{\theta_{\mathfrak{a}}})_{\mathbb{F}_{p^g}}$  and we write  $Z_1 = \mathcal{S}_K(G_{\theta_{\mathfrak{a}}})_{\mathbb{F}_{p^g}, \tau_i}$ ,  $Z_2 = \mathcal{S}_K(G_{\theta_{\mathfrak{a}}})_{\mathbb{F}_{p^g}, \tau_j}$  for the Goren–Oort divisors. We let  $Z_{12}$  denote the intersection  $Z_1 \cap Z_2$  and we write  $X^{\mathrm{ord}}$  for the complement of  $Z_1 \cup Z_2$  in  $X$ .

The first observation is that the map  $\Psi(\mathfrak{a})_{\mathfrak{m}}$  factors through a certain cohomology group with supports.

**Proposition 5.4.4.** *The map  $\Psi(\mathfrak{a})_{\mathfrak{m}}$  factors through  $\mathrm{H}_{Z_1 \cup Z_2}^3(X, k_\lambda(2))_{\mathfrak{m}}$ . Moreover the induced map*

$$\mathrm{K}(\mathfrak{a})_{\mathfrak{m}} \rightarrow \mathrm{H}_{Z_1 \cup Z_2}^3(X, k_\lambda(2))_{\mathfrak{m}}$$

*is surjective.*

*Proof.* The factoring property follows from the definition of the cycle class map. Indeed we have the following diagram with exact columns and rows

$$\begin{array}{ccccccc}
 & & & & 0 & & 0 \\
 & & & & \uparrow & & \uparrow \\
 0 & \longrightarrow & \mathbf{L} & \longrightarrow & \mathrm{H}^1(Z_1 \cap Z_2 - Z_{12}, k_\lambda(1)) & \xrightarrow{c} & \mathrm{H}^0(Z_{12}, k_\lambda) \\
 & & \uparrow a & & \uparrow b & & \parallel \\
 0 & \longrightarrow & \mathbf{K} & \longrightarrow & \mathrm{H}^0(Z_1 \cap Z_2 - Z_{12}, \mathbb{G}_m) \otimes_{\mathbb{Z}} k_\lambda^{\mathrm{div} \otimes_{\mathbb{Z}} k_\lambda} & \longrightarrow & \mathrm{H}^0(Z_{12}, k_\lambda).
 \end{array}$$

Here  $\mathrm{H}^0(Z_1 \cap Z_2 - Z_{12}, \mathbb{G}_m)$  is identified with the group of rational functions on  $Z_1 \cap Z_2$  with zeros and poles only at the points in  $Z_{12}$ , and the map  $\mathrm{div} \otimes_{\mathbb{Z}} k_\lambda$  is induced from the map taking such a function to its divisor. The groups  $\mathbf{L}$  and  $\mathbf{K}$  are by definition the kernels of the maps on the right. The map  $b$  comes from the Kummer sequence and its surjectivity follows from the fact that  $Z_1 \cup Z_2 - Z_{12}$  is a union of open subsets in  $\mathbb{A}_{\mathbb{F}_{p^g}/\mathbb{F}_{p^g}}^1$  for  $g \mid h$  hence all have trivial Picard group.

Here we write  $\mathbb{A}_{\mathbb{F}_p^h/\mathbb{F}_{p^g}}^1$  for the affine line  $\mathbb{A}^1$  over  $\mathbb{F}_{p^h}$  considered as a  $\mathbb{F}_{p^g}$  scheme. The map  $\text{div} \otimes_{\mathbb{Z}} k_\lambda$  factors through  $b$ , hence we obtain the map  $c$ .

By purity, we may identify the top row of the exact sequence with the exact sequence of the triple  $(X, Z_{12}, Z_1 \cup Z_2)$ :

$$\begin{aligned} \mathrm{H}_{Z_{12}}^3(X, k_\lambda(2)) = 0 &\rightarrow \mathrm{H}_{Z_1 \cup Z_2}^3(X, k_\lambda(2)) \\ &\rightarrow \mathrm{H}_{Z_1 \cup Z_2 - Z_{12}}^3(X - Z_{12}, k_\lambda(2)) \rightarrow \mathrm{H}_{Z_{12}}^4(X, k_\lambda(2)). \end{aligned}$$

It is easy to see from the definition that the group  $\mathrm{Ch}_{Z_1 \cup Z_2}^2(X, 1, k_\lambda)$  is a quotient of the group  $\mathbf{K}$ . Then the cycle class map is induced by

$$\mathbf{K} \rightarrow \mathbf{L} \cong \mathrm{H}_{Z_1 \cup Z_2}^3(X, k_\lambda(2)) \rightarrow \mathrm{H}^3(X, k_\lambda(2)).$$

To deduce the moreover part, we note that the map

$$\mathbf{K}(\mathfrak{a})_{\mathfrak{m}} \rightarrow \mathrm{Ch}_{Z_1 \cup Z_2}^2(X, 1, k_\lambda)_{\mathfrak{m}}$$

is surjective. Indeed by the discussion after Proposition 5.2.4, the map

$$\left( \bigoplus_{S \subset Z_1 \cup Z_2} k_S^\times \otimes_{\mathbb{Z}} k_\lambda \right) \oplus \mathbf{K}(\mathfrak{a})^{\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})} \rightarrow \mathrm{Ch}_{Z_1 \cup Z_2}^2(X, 1, k_\lambda)$$

is surjective. But  $k_S^\times \otimes_{\mathbb{Z}} k_\lambda = 0$  for all  $S$  by Definition 5.3.2 (4). It therefore suffices to show that the map  $a$  is surjective. We write

$$\begin{aligned} \mathbf{A} &:= \mathrm{Im}(\mathrm{H}^0(Z_1 \cap Z_2 - Z_{12}, \mathbb{G}_m) \rightarrow \mathrm{H}^0(Z_{12}, k_\lambda)) \\ \mathbf{B} &:= \mathrm{Im}(\mathrm{H}^1(Z_1 \cap Z_2 - Z_{12}, k_\lambda(1)) \rightarrow \mathrm{H}^0(Z_{12}, k_\lambda)). \end{aligned}$$

The map  $\mathbf{A} \rightarrow \mathbf{B}$  is injective, hence by the snake lemma,  $a$  is surjective.  $\square$

Thus in order to prove Proposition 5.4.3, and hence Theorem 5.3.6, it suffices to show that the map

$$\Phi'_{\mathfrak{m}} : \mathrm{H}_{Z_1 \cup Z_2}^3(X, k_\lambda(2))_{\mathfrak{m}} \rightarrow \mathrm{H}^3(X, k_\lambda(2))_{\mathfrak{m}}$$

is surjective.

By the Hochschild–Serre spectral sequence, this map fits into the following diagram with exact rows:

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}_{Z_1 \cup Z_2}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) & \longrightarrow & \mathrm{H}_{Z_1 \cup Z_2}^3(X, k_\lambda(2))_{\mathfrak{m}} & \longrightarrow & \mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}_{Z_1 \cup Z_2}^3(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) \longrightarrow 0 \\ & & \downarrow & & \Phi'_{\mathfrak{m}} \downarrow & & \downarrow \\ 0 & \longrightarrow & \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) & \longrightarrow & \mathrm{H}^3(X, k_\lambda(2))_{\mathfrak{m}} & \longrightarrow & \mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}^3(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) \longrightarrow 0. \end{array}$$

Since  $\mathrm{H}^3(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}} = 0$  by Proposition 4.1.2, we have an isomorphism

$$\mathrm{H}^3(X, k_\lambda(2))_{\mathfrak{m}} \cong \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}).$$

The diagram above then induces a map

$$\Phi_{\mathfrak{m}} : \mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}_{Z_1 \cup Z_2}^3(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) \rightarrow \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) / \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}_{Z_1 \cup Z_2}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}})$$

and surjectivity of  $\Phi'_{\mathfrak{m}}$  is equivalent to the surjectivity of  $\Phi_{\mathfrak{m}}$ . It therefore suffices to prove  $\Phi_{\mathfrak{m}}$  is surjective.

We may dualize the above to obtain a map

$$\Phi_{\mathfrak{m}}^* : \left( \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) / \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}_{Z_1 \cup Z_2}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}}) \right)^* \rightarrow \mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}_{Z_1 \cup Z_2}^3(X_{\overline{\mathbb{F}}_p}, k_\lambda(2))_{\mathfrak{m}})^*.$$

Consider the exact sequence of cohomology

$$\mathrm{H}^1(X_{\overline{\mathbb{F}}_p}, k_\lambda) \rightarrow \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda) \rightarrow \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\mathrm{ord}}, k_\lambda) \rightarrow \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda) \rightarrow \mathrm{H}^2(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)$$

arising from the triple  $Z_1 \cup Z_2 \xrightarrow{i} X \xleftarrow{j} X^{\text{ord}}$ . When  $X$  is the special fiber of the Hilbert modular surface (hence non-compact), we abuse notation and write  $H_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)$  for  $H^2(X_{\overline{\mathbb{F}}_p}, j_! k_\lambda)$ . Upon localizing at  $\mathfrak{m}$ ,  $H^1(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} = 0$ , and we obtain the following boundary map for the long exact sequence of Galois cohomology:

$$(5.4.1) \quad \ker \left( H^0(\mathbb{F}_{p^g}, H^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}) \rightarrow H^0(\mathbb{F}_{p^g}, H^2(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}) \right) \rightarrow H^1(\mathbb{F}_{p^g}, H^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}).$$

By Poincaré duality and the duality of Galois cohomology over finite fields, (5.4.1) is identified with  $\Phi_{\mathfrak{m}}^*$ . We mention that in the case of Hilbert modular surfaces, we need to use the canonical isomorphism

$$H^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \cong H_c^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}},$$

which follows from [LS18, Corollary 5.20] and [Dim09, Theorem 2.3].

Let  $X_0(\mathfrak{p})$  be the special fiber  $\mathcal{S}_{K_0(\mathfrak{p})}(G_{\theta_a})_{\mathbb{F}_{p^g}}$  for the Shimura variety  $\text{Sh}_{K_0(\mathfrak{p})}(G_{\theta_a})$  with Iwahori level structure at  $p$  constructed in A.1. We claim  $\Phi_{\mathfrak{m}}^*$  can be related to a certain map relating the geometry of  $X$  and  $X_0(\mathfrak{p})$ .

By Corollary A.5.4 there is a decomposition

$$X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} = \mathcal{X}_{1\overline{\mathbb{F}}_p} \cup \mathcal{X}_{2\overline{\mathbb{F}}_p} \cup \mathcal{X}_{3\overline{\mathbb{F}}_p}$$

where  $\mathcal{X}_{1\overline{\mathbb{F}}_p}$  and  $\mathcal{X}_{2\overline{\mathbb{F}}_p}$  are the two copies of  $X_{\overline{\mathbb{F}}_p}$  corresponding to the essential Frobenius and Verschiebung isogenies respectively and  $\mathcal{X}_{3\overline{\mathbb{F}}_p}$  is the ‘‘supersingular locus’’. For  $k, l, m \in \{1, 2, 3\}$  distinct, we write

$$\begin{aligned} i_k &: \mathcal{X}_{k\overline{\mathbb{F}}_p} \rightarrow X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} \\ i_{kl} &: \mathcal{X}_{k\overline{\mathbb{F}}_p} \cap \mathcal{X}_{l\overline{\mathbb{F}}_p} \rightarrow X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} \\ i_{klm} &: \mathcal{X}_{k\overline{\mathbb{F}}_p} \cap \mathcal{X}_{l\overline{\mathbb{F}}_p} \cap \mathcal{X}_{m\overline{\mathbb{F}}_p} \rightarrow X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} \end{aligned}$$

for the closed immersions. Then we have an exact sequence of sheaves on  $X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$ :

$$0 \rightarrow k_\lambda \xrightarrow{f} i_{1*}k_\lambda \oplus i_{2*}k_\lambda \oplus i_{3*}k_\lambda \rightarrow i_{23*}k_\lambda \oplus i_{13*}k_\lambda \oplus i_{23*}k_\lambda \xrightarrow{g} i_{123*}k_\lambda \rightarrow 0.$$

Here the maps are induced by (pushing forward) the usual unit maps of the adjunction. We define

$$C := \text{coker}(f) = \ker(g).$$

Then the above exact sequence breaks up into two short exact sequences of sheaves on  $X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$ .

$$(5.4.2) \quad 0 \rightarrow k_\lambda \xrightarrow{f} i_{1*}k_\lambda \oplus i_{2*}k_\lambda \oplus i_{3*}k_\lambda \rightarrow C \rightarrow 0$$

$$(5.4.3) \quad 0 \rightarrow C \rightarrow i_{23*}k_\lambda \oplus i_{13*}k_\lambda \oplus i_{23*}k_\lambda \xrightarrow{g} i_{123*}k_\lambda \rightarrow 0$$

Taking cohomology of the sequence (5.4.2), we obtain:

$$0 \longrightarrow \mathbf{R} \longrightarrow H^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \xrightarrow{(i_1^*, i_2^*, i_3^*)} H^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \oplus H^2(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}$$

where  $\mathbf{R} \cong H^1(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, C)_{\mathfrak{m}}/H^1(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}$ .

Consider the map

$$\pi_1^* + \pi_2^* : H^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \rightarrow H^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}$$

induced by the degeneracy maps  $\pi_1, \pi_2 : X_0(\mathfrak{p}) \rightarrow X$ . Let  $R\psi_{k_\lambda}$  denote the nearby cycles functor applied to the sheaf  $k_\lambda$  on  $X_{\overline{\mathbb{Q}}_p}$ . Then we have canonical isomorphisms

$$(5.4.4) \quad H^2(\text{Sh}_K(G_{\theta_a})_{\overline{\mathbb{Q}}_p}, k_\lambda)_{\mathfrak{m}} \cong H^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}$$

$$(5.4.5) \quad H^2(\text{Sh}_{K_0(\mathfrak{p})}(G_{\theta_a})_{\overline{\mathbb{Q}}_p}, k_\lambda)_{\mathfrak{m}} \cong H^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, R\psi_{k_\lambda})_{\mathfrak{m}},$$

where these isomorphisms follow from proper base change if  $\mathrm{Sh}_K(G_{\theta_a})$  and  $\mathrm{Sh}_{K_0(\mathfrak{p})}(G_{\theta_a})$  are proper, and [LS18, Corollary 5.20] if they are not proper.

We write

$$\pi_{1,\bar{\eta}}^* + \pi_{2,\bar{\eta}}^* : \mathrm{H}^2(\mathrm{Sh}_K(G_{\theta_a})_{\overline{\mathbb{Q}}_p}, k_\lambda)_{\mathfrak{m}}^2 \rightarrow \mathrm{H}^2(\mathrm{Sh}_{K_0(\mathfrak{p})}(G_{\theta_a})_{\overline{\mathbb{Q}}_p}, k_\lambda)_{\mathfrak{m}}$$

for the map induced by the degeneracy maps on the generic fiber which is injective by Theorem 4.1.4. Then under the the identifications (5.4.4) and (5.4.5),  $\pi_{1,\bar{\eta}}^* + \pi_{2,\bar{\eta}}^*$  factors as

$$\mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \xrightarrow{\pi_1^* + \pi_2^*} \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \rightarrow \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, R\psi_{k_\lambda})_{\mathfrak{m}}.$$

It follows that  $\pi_1^* + \pi_2^* : \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \rightarrow \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}$  is injective.

We consider the kernel  $\mathbf{S}$  of the composition

$$\mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \xrightarrow{\pi_1^* + \pi_2^*} \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \xrightarrow{(i_1^*, i_2^*, i_3^*)} \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \oplus \mathrm{H}^2(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}.$$

We obtain a map  $\Psi_{\mathfrak{m}} : \mathbf{S} \rightarrow \mathbf{R}$  which fits into the diagram:

$$\begin{array}{ccccc} 0 & \longrightarrow & \mathbf{R} & \longrightarrow & \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} & \xrightarrow{(i_1^*, i_2^*, i_3^*)} & \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \oplus \mathrm{H}^2(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \\ & & \uparrow \Psi_{\mathfrak{m}} & & \uparrow \pi_1^* + \pi_2^* & & \\ & & \mathbf{S} & \longrightarrow & \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 & & \end{array}$$

Since  $\pi_1^* + \pi_2^*$  is injective, it follows that  $\Psi_{\mathfrak{m}}$  is injective. Therefore in order to prove Proposition 5.4.3, it suffices to prove the following proposition.

**Proposition 5.4.5.** *The map  $\Psi_{\mathfrak{m}}$  can be identified with the map  $\Phi_{\mathfrak{m}}^*$ .*

*Proof.* We first identify the groups  $\mathbf{S}$  and  $\mathbf{R}$  with the corresponding domain and codomain for the map  $\Phi_{\mathfrak{m}}^*$ .

First note that the composition

$$\mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2 \xrightarrow{\pi_1^* + \pi_2^*} \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \xrightarrow{(i_1^*, i_2^*)} \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2$$

induces the endomorphism

$$(5.4.6) \quad \begin{pmatrix} 1 & \mathrm{Fr}' \\ \mathrm{Fr}' S_{\mathfrak{p}}^{-1} & 1 \end{pmatrix},$$

where  $\mathrm{Fr}'$  is the essential Frobenius as defined in A.4. By Definition 5.3.2 (3) and Assumption 4.1.1 (1), this is the same as the endomorphism

$$\begin{pmatrix} 1 & \mathrm{Fr}' \\ \mathrm{Fr}' & 1 \end{pmatrix}.$$

By Proposition A.4.2 and Definition 5.3.2 (3),  $\mathrm{Fr}'^2 = \mathrm{Fr}_{\mathfrak{p}}$  the  $p^g$ -Frobenius, thus we may identify the kernel of (5.4.6) with the image of the injective map

$$\mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda))_{\mathfrak{m}} \cong \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^{\mathrm{Fr}_{\mathfrak{p}}=1} \xrightarrow{(-\mathrm{Fr}', \mathrm{id})} \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}^2.$$

Therefore we have an identification

$$\mathbf{S} \cong \ker \left( \mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda))_{\mathfrak{m}} \rightarrow \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \xrightarrow{i_3^*} \mathrm{H}^2(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \right).$$

The map in the above equation factors as

$$\mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda))_{\mathfrak{m}} \rightarrow \mathrm{H}^2(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}} \hookrightarrow \mathrm{H}^2(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_{\mathfrak{m}}$$

Here the second map is an injection by [Liu19, Lemma 3.26 (2)] since  $\mathcal{X}_{3\overline{\mathbb{F}}_p} \cong \mathcal{Z}_{1\overline{\mathbb{F}}_p} \cup \mathcal{Z}_{2\overline{\mathbb{F}}_p}$  where  $\mathcal{Z}_{1\overline{\mathbb{F}}_p}$  and  $\mathcal{Z}_{2\overline{\mathbb{F}}_p}$  are  $\mathbb{P}^1$ -fibrations over  $Z_{1\overline{\mathbb{F}}_p}$  and  $Z_{2\overline{\mathbb{F}}_p}$  respectively and the intersection  $\mathcal{Z}_{1\overline{\mathbb{F}}_p} \cap \mathcal{Z}_{2\overline{\mathbb{F}}_p}$  is 0-dimensional. Therefore we may identify  $\mathbf{S}$  with the domain in the definition of  $\Phi_m^*$ .

Consider the following commutative diagram of  $\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})$ -modules with exact rows:

$$\begin{array}{ccccccc}
& & & \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m & \longrightarrow & \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_m & \\
& & & \downarrow \pi_2^* - \pi_1^* \text{Fr}' & & \downarrow \pi_2^* - \pi_1^* \text{Fr}' & \\
0 & \longrightarrow & \mathrm{H}^1(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_m & \xrightarrow{\alpha} & \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m^2 & \longrightarrow & \mathrm{H}^2(X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}, k_\lambda)_m \xrightarrow{i_3^*} \mathrm{H}^2(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_m \\
& & \downarrow \Delta & & \parallel & & \downarrow (i_1^*, i_2^*) \\
0 & \longrightarrow & \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m^2 & \xrightarrow{\gamma \oplus \gamma} & \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m^2 & \longrightarrow & \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_m^2 \xrightarrow{\delta \oplus \delta} \mathrm{H}^2(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m^2. \\
& & & & & & \downarrow \Delta'
\end{array}$$

Here the middle row is the exact sequence of cohomology arising from the triple

$$\mathcal{X}_{3\overline{\mathbb{F}}_p} \xrightarrow{i_3} X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} \leftarrow X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} - \mathcal{X}_{3\overline{\mathbb{F}}_p} \cong X^{\text{ord}} \sqcup X^{\text{ord}}.$$

The maps  $\Delta$  and  $\Delta'$  are induced by the inclusion maps  $i_{12}, i_{13}$ .

Now recall  $\Phi_m^*$  is identified with the coboundary map after taking cohomology of the short exact sequence:

$$(5.4.7) \quad 0 \rightarrow \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m \xrightarrow{\gamma} \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m \rightarrow \text{coker}(\gamma) \rightarrow 0.$$

To identify  $\mathbf{R}$  with the codomain  $\mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m)$ , note that by definition we have a canonical identification

$$\mathbf{R} \cong \ker \left( \text{coker}(\alpha) \xrightarrow{(i_{1*}, i_{2*})} \text{coker}(\gamma \oplus \gamma) \right).$$

We then use the snake lemma applied to the diagram:

$$\begin{array}{ccccccc}
0 & \longrightarrow & \mathrm{H}^1(\mathcal{X}_{3\overline{\mathbb{F}}_p}, k_\lambda)_m & \xrightarrow{\alpha} & \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m^2 & \longrightarrow & \text{coker}(\alpha) \longrightarrow 0 \\
& & \downarrow \Delta & & \parallel & & \downarrow \\
0 & \longrightarrow & \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m^2 & \xrightarrow{\gamma \oplus \gamma} & \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m^2 & \longrightarrow & \text{coker}(\gamma \oplus \gamma) \longrightarrow 0
\end{array}$$

to deduce an isomorphism

$$\mathbf{R} \cong \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m \cong \mathrm{H}^1(\mathbb{F}_{p^g}, \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m)$$

where we have identified  $\text{coker}(\Delta)$  with  $\mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m$  via projection onto the second factor and the second isomorphism follows from Definition 5.3.2 (3) and Assumption 4.1.1, which implies  $\text{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_{p^g})$  acts trivially on  $\mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m$ . Using these identifications, we may now compare the maps  $\Phi_m^*$  and  $\Psi_m$ .

Let  $x \in \mathbf{S} \subset \mathrm{H}^0(\mathbb{F}_{p^g}, \mathrm{H}^2(X_{\overline{\mathbb{F}}_p}, k_\lambda)_m)$ . We will compute  $\Psi_m(x) \in \mathbf{R}$  considered as an element of  $\text{coker}(\Delta) \cong \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m$  and show that it coincides with the image under the coboundary map of (5.4.7).

To compute  $\Psi_m(x)$  we let  $\tilde{x} \in \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m$  be a lift of  $x$ , which exists since  $x$  maps to 0 in  $\mathrm{H}^2(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m$ . Moreover we may assume  $\tilde{x}$  is fixed by  $S_{\mathfrak{p}}$  since  $S_{\mathfrak{p}}$  acts semi-simply on  $\mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m$ . Then  $\pi_2^*(\tilde{x}) - \pi_1^* \text{Fr}'(\tilde{x}) \in \mathrm{H}_c^2(X_{\overline{\mathbb{F}}_p}^{\text{ord}}, k_\lambda)_m^2$  is an element lifting  $\pi_2^*(x) - \pi_1^* \text{Fr}'(x) \in \ker(i_1^*, i_2^*)$ , hence comes from an element

$$c(\tilde{x}) \in \mathrm{H}^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, k_\lambda)_m^2.$$

By definition,  $\Psi_{\mathfrak{m}}(x)$  is the image of  $\pi_2^*(x) - \pi_1^* \text{Fr}'(x) \in \ker(i_1^*, i_2^*)$  in  $\text{coker}(\Delta)$ . Therefore the image of  $c(\tilde{x})$  is identified with  $\Psi_{\mathfrak{m}}(x)$ .

However we also have

$$\pi_2^*(\tilde{x}) - \pi_1^* \text{Fr}'(\tilde{x}) = (S_{\mathfrak{p}}^{-1} \text{Fr}'(\tilde{x}), \tilde{x}) - (\text{Fr}'(\tilde{x}), \text{Fr}_{\mathfrak{p}}(\tilde{x})) \cong (0, (1 - \text{Fr}_{\mathfrak{p}})(\tilde{x})).$$

By definition of the coboundary map, the image of this element in  $\text{coker}(\Delta) \cong H^1(Z_{1\overline{\mathbb{F}}_p} \cup Z_{2\overline{\mathbb{F}}_p}, \Lambda)_{\mathfrak{m}}$  is equal to  $\Phi_{\mathfrak{m}}^*(x)$ .  $\square$

We finish by stating the following corollary.

**Corollary 5.4.6.** *Let  $p$  be a  $\lambda$ -level raising prime and suppose that Assumption 4.1.1 is satisfied. Then the cycle class map*

$$H_{\mathcal{M}}^3(\mathcal{S}_K(G_{\theta_a})_{\mathbb{F}_{p^g}}, k_{\lambda}(2))_{\mathfrak{m}} \rightarrow H^3(\mathcal{S}_K(G_{\theta_a})_{\mathbb{F}_{p^g}}, k_{\lambda}(2))_{\mathfrak{m}}$$

is an isomorphism.

*Proof.* The surjectivity follows from Proposition 5.4.3. The injectivity is [Voe11, Theorem 6.17].  $\square$

## APPENDIX A. BAD REDUCTION OF QUATERNIONIC SHIMURA SURFACES

In this appendix we give a global description of the special fiber of certain quaternionic Shimura surfaces with *Iwahori level structure at  $p$* . These are associated to the group  $G_S$  with  $|\Sigma_{\infty} - S_{\infty}| = 2$ . The idea is that only the two places  $\Sigma_{\infty} - S_{\infty}$  should contribute to the geometry, and thus the structural results we obtain are completely analogous to the results of Stamm [Sta97] in the case of the Hilbert modular surface. Indeed the proofs are also completely analogous, with some extra technical difficulties since we must transfer the results from unitary Shimura varieties. The key difference is that we must replace the notion of the usual notion of Frobenius and Verschiebung in the case of Hilbert modular surfaces with the notion of essential Frobenius and Verschiebung. For this reason, we will refer to [Sta97] for some of the computations, because they are exactly the same.

**A.1. Moduli interpretation and local models.** We keep the notation of §2.1. In particular  $B/F$  is a totally indefinite quaternion algebra and  $S \subset \Sigma_p \cup \Sigma_{\infty}$  a set of even cardinality. In fact in this section we will make the following further assumptions.

**Assumption A.1.1.** (1)  $g := [F : \mathbb{Q}]$  is even and the prime  $p$  is inert in  $F$ .  
 (2)  $\mathfrak{p} \notin S$  and  $|\Sigma_{\infty} - S_{\infty}| = 2$ .

We write  $\mathfrak{p}$  for the unique prime of  $F$  above  $p$ . We will also exclude the case  $B = \text{GL}_2(F)$  and  $g = 2$ , i.e. the case of the Hilbert modular surfaces as this case is already covered in [Sta97]. In particular this assumption implies the Shimura varieties we consider are compact. We fix an identification  $\Sigma_{\infty} \cong \{\tau_1, \dots, \tau_g\}$  such that  $\sigma(\tau_i) = \tau_{i+1}$  for  $i = 1, \dots, g-1$  and  $\sigma(\tau_g) = 1$ , and  $\Sigma_{\infty} - S_{\infty} = \{\tau_1, \tau_c\}$  for some  $c \in \{2, \dots, g\}$ .

Fix a subset  $T \subset S_{\infty}$  such that  $2|T| = |S|$  and let  $K_{\mathfrak{p}}$  be the standard hyperspecial subgroup of  $(B_S \otimes_F F_{\mathfrak{p}})^{\times} \cong \text{GL}_2(F_{\mathfrak{p}})$ . We also define  $K_{0,\mathfrak{p}}$  to denote the following compact open subgroup

$$K_{0,\mathfrak{p}} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathcal{O}_{F_{\mathfrak{p}}}) \mid \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{\mathfrak{p}} \right\};$$

it is an Iwahori subgroup of  $\text{GL}_2(F_{\mathfrak{p}})$ .

For  $K^p$  sufficiently small, we define  $K = K_{\mathfrak{p}} K^p$  and  $K_0(\mathfrak{p}) = K_{0,\mathfrak{p}} K^p$ . We then have the Shimura varieties  $\text{Sh}_{K_0(\mathfrak{p})}(G_{S,T})$ ,  $\text{Sh}_K(G_{S,T})$  defined over the common reflex field  $\mathbf{E}_{S,T}$ . These are equipped with finite étale maps

$$\pi_1, \pi_2 : \text{Sh}_{K_0(\mathfrak{p})}(G_{S,T}) \rightarrow \text{Sh}_K(G_{S,T}).$$

Using a similar procedure as in §2.2, we may define an integral model for  $\text{Sh}_{K_0(\mathfrak{p})}(G_{S,T})$  as follows.



We fix a choice  $E/F$  of CM-extension and a pair  $\tilde{S} = (S, \tilde{S}_\infty)$  as in §2.2; in particular  $E$  is split over the prime  $\mathfrak{p}$  and we write  $\mathfrak{q}$  and  $\bar{\mathfrak{q}}$  for the primes above  $\mathfrak{p}$ . We make the following assumption on  $\tilde{S}_\infty$ .

**Assumption A.1.2.** Let  $\tilde{\tau} \in \Sigma_{E, \infty/\mathfrak{q}}$  be a lift of  $\tau \in \Sigma_{\infty/\mathfrak{q}}$ . Then  $\tilde{\tau} \in \tilde{S}_\infty$  if and only if  $\tau \in T$ .

We let  $D_S := B_S \otimes_F E$  and  $W := D_S$  considered as a left  $D_S$ -module of rank 1. Then  $W$  is equipped with a pairing  $\psi$  of (2.2.2). Recall the associated groups  $G'_S$  and  $G''_S$  which fit in the diagram

$$G_S \leftarrow G_S \times T_E \rightarrow G''_S \leftarrow G'_S$$

We let  $K''_{0,\mathfrak{p}}$  denote the image of  $K_{0,\mathfrak{p}} \times K_{E,\mathfrak{p}}$  in  $G''_S(\mathbb{Q}_p)$ .

To define the level structures for  $G'_S$ , note that an element  $g_p \in G'_S(\mathbb{Q}_p)$  corresponds to an element of  $\text{End}(W \otimes_{\mathbb{Q}} \mathbb{Q}_p)$  such that

$$\psi(vg_p, wg_p) = c(g_p)\psi(v, w), \quad \forall v, w \in W \otimes_{\mathbb{Q}} \mathbb{Q}_p.$$

Here we abuse notation so that  $\psi$  also denotes the base change of the pairing  $\psi$  to  $\mathbb{Q}_p$ . We fix an isomorphism  $\mathcal{O}_{B_S} \otimes_{\mathcal{O}_F} \mathcal{O}_{F_{\mathfrak{p}}} \cong \text{Mat}_2(\mathcal{O}_{F_{\mathfrak{p}}})$  which induces an identification

$$\mathcal{O}_{D_S} \otimes_F F_{\mathfrak{p}} \cong \text{Mat}_2(\mathcal{O}_{E_{\mathfrak{q}}}) \times \text{Mat}_2(\mathcal{O}_{E_{\bar{\mathfrak{q}}}}).$$

We then define the following lattice chain  $\Lambda_{1,p} \subset \Lambda_{2,p}$  of  $\mathcal{O}_{D_S} \otimes_F F_{\mathfrak{p}}$ -modules where

$$\Lambda_{1,p} = \begin{pmatrix} \mathcal{O}_{\mathfrak{q}} & \mathfrak{q} \\ \mathcal{O}_{\bar{\mathfrak{q}}} & \mathfrak{q} \end{pmatrix} \oplus \begin{pmatrix} \mathcal{O}_{\bar{\mathfrak{q}}} & \mathcal{O}_{\bar{\mathfrak{q}}} \\ \mathcal{O}_{\mathfrak{q}} & \mathcal{O}_{\mathfrak{q}} \end{pmatrix}, \quad \Lambda_{2,p} = \begin{pmatrix} \mathcal{O}_{\mathfrak{q}} & \mathcal{O}_{\mathfrak{q}} \\ \mathcal{O}_{\bar{\mathfrak{q}}} & \mathcal{O}_{\bar{\mathfrak{q}}} \end{pmatrix} \oplus \begin{pmatrix} \bar{\mathfrak{q}}^{-1} & \mathcal{O}_{\bar{\mathfrak{q}}} \\ \mathcal{O}_{\mathfrak{q}} & \bar{\mathfrak{q}}^{-1} \end{pmatrix}.$$

Let  $K'_{0,\mathfrak{p}} \subset G'_S(\mathbb{Q}_p)$  denote the compact open subgroup stabilizing the lattice chain  $\Lambda_{2,p} \subset \Lambda_{1,p}$ . We fix an  $\mathcal{O}_{D_S}$  lattice  $\Lambda_1$  of  $W$  such that  $\Lambda_1 \otimes_{\mathbb{Z}} \mathbb{Z}_p = \Lambda_{1,p}$ . For a sufficiently small compact open subgroup of  $K'^p \subset G'_S(\mathbb{A}_f^p)$  (resp.  $K''^p \subset G''_S(\mathbb{A}_f^p)$ ) which stabilizes  $\Lambda_1 \otimes_{\mathbb{Z}} \mathbb{Z}^{(p)}$ , we write  $K'_0(\mathfrak{p})$  (resp.  $K''_0(\mathfrak{p})$ ) for the compact open subgroup  $K'_{0,\mathfrak{p}}K'^p \subset G'_S(\mathbb{A}_f)$  (resp.  $K''_{0,\mathfrak{p}}K''^p \subset G''_S(\mathbb{A}_f)$ ). We then have the associated Shimura varieties  $\text{Sh}_{K'_0(\mathfrak{p})}(G'_S)$ ,  $\text{Sh}_{K''_0(\mathfrak{p})}(G''_S)$  over the common reflex field  $\mathbf{E}_{\tilde{S}}$ , and the inverse limit schemes:

$$\text{Sh}_{K'_{0,\mathfrak{p}}}(G'_S) := \lim_{\leftarrow K'^p} \text{Sh}_{K'_0(\mathfrak{p})}(G'_S), \quad \text{Sh}_{K''_{0,\mathfrak{p}}}(G''_S) := \lim_{\leftarrow K''^p} \text{Sh}_{K''_0(\mathfrak{p})}(G''_S).$$

These Shimura varieties are equipped with finite étale maps

$$\begin{aligned} \pi'_1, \pi'_2 &: \text{Sh}_{K'_0(\mathfrak{p})}(G'_S) \rightarrow \text{Sh}_{K'}(G'_S) \\ \pi''_1, \pi''_2 &: \text{Sh}_{K''_0(\mathfrak{p})}(G''_S) \rightarrow \text{Sh}_{K''}(G''_S). \end{aligned}$$

As before we also have the identification of neutral connected components:

$$\text{Sh}_{K'_{0,\mathfrak{p}}}(G'_S)_{\mathbb{Q}_p}^{\circ} \xleftarrow{\sim} \text{Sh}_{K''_{0,\mathfrak{p}}}(G''_S)_{\mathbb{Q}_p}^{\circ} \xrightarrow{\sim} \text{Sh}_{K'_{0,\mathfrak{p}}}(G'_S)_{\mathbb{Q}_p}^{\circ}.$$

In fact these isomorphisms descend to  $\mathbb{Q}_p^{\text{ur}}$ , and this will allow us to transfer the integral models that we will construct.

In order to define integral models, we follow the procedure in [RZ96] adapted to our situation.

We first make the following definition.

**Definition A.1.3.** Let  $A/S$  be the abelian variety associated to some point of  $\text{Sh}_{K'}(G'_S)$ . A *cyclic isogeny*

$$\rho : A \rightarrow A'$$

is an isogeny of abelian varieties over  $S$  such that the kernel of  $\rho$  is an  $\mathcal{O}_{D_S}$ -stable subgroup  $H_{\mathfrak{q}} \oplus H_{\bar{\mathfrak{q}}}$  of  $A[\mathfrak{q}] \oplus A[\bar{\mathfrak{q}}]$  of order  $|k_{\mathfrak{p}}|^4$  such that  $H_{\mathfrak{q}}$  and  $H_{\bar{\mathfrak{q}}}$  are dual to one another under the pairing

$$H_{\mathfrak{q}} \times H_{\bar{\mathfrak{q}}} \rightarrow \mu_p$$

induced by the polarization. Moreover we require that the induced action of  $\mathcal{O}_{D_S}$  on  $A'$  to satisfy the condition (2.3.1).

We now consider the moduli problem  $\underline{\text{Sh}}_{K'_0(\mathfrak{p})}(G'_S)$  that associates to an  $\mathcal{O}_{\mathbf{E}_{\mathfrak{S},\bar{v}}}$ -scheme  $S$  the set of isomorphism classes of tuples  $(\rho : A \rightarrow A', \iota, \lambda, \varepsilon_{K'^p})$  where :

- $(A, \iota, \lambda, \varepsilon_{K'^p})$  is an  $S$ -point of  $\underline{\text{Sh}}_{K'}(G'_S)$ .
- $\rho : A \rightarrow A'$  is a cyclic isogeny.

Then  $\underline{\text{Sh}}_{K'_0(\mathfrak{p})}(G'_S)$  is representable by a quasi-projective variety over  $\mathcal{O}_{\mathbf{E}_{\mathfrak{S},\bar{v}}}$  and it is an integral model for  $\text{Sh}_{K'_0(\mathfrak{p})}(G'_S)$ . The two degeneracy maps extend to maps also denoted  $\pi_1, \pi_2$  of  $\mathcal{O}_{\mathbf{E}_{\mathfrak{S},\bar{v}}}$ -schemes:

$$\pi_1, \pi_2 : \underline{\text{Sh}}_{K'_0(\mathfrak{p})}(G'_S) \rightarrow \underline{\text{Sh}}_{K'}(G'_S).$$

These maps can be described explicitly in terms of the moduli interpretation as follows.  $\pi_1$  sends the tuple  $(\rho : A \rightarrow A', \iota, \lambda, \varepsilon_{K'^p})$  to  $(A, \iota, \lambda, \varepsilon_{K'^p})$ . To define  $\pi_2$ , note that given a tuple  $(\rho : A \rightarrow A', \iota, \lambda, \varepsilon_{K'^p})$ , since  $H_{\mathfrak{q}} \oplus H_{\bar{\mathfrak{q}}} := \ker(\rho)$  is  $\mathcal{O}_{D_S}$ -stable, the action of  $\mathcal{O}_{D_S}$  on  $A$  extends to an action of  $A'$ . Moreover the polarization  $\lambda$  and level structure  $\varepsilon_{K'^p}$  induce a polarization  $\lambda'$  and level structure  $\varepsilon'_{K'^p}$  on  $A'$ ; here  $\lambda'$  satisfies  $p\lambda = \rho^\vee \circ \lambda' \circ \rho$ . It is easy to check the tuple  $(A', \iota', \lambda', \varepsilon'_{K'^p})$  satisfies the conditions defining a point of  $\underline{\text{Sh}}_{K'}(G'_S)$ . This defines the map  $\pi_2$ .

In order to transfer the integral model to the quaternionic side, we can use the construction in [TX16, §2]. Note however that since the models we construct do not satisfy the correct extension property, there is a subtlety in defining the  $G^{\text{ad}}(\mathbb{Q})^+$ -action on the integral model. We may instead use a direct description of this action as in [KP18, §4.4] by twisting abelian varieties; the rest of the argument then goes through and we obtain an integral model  $\underline{\text{Sh}}_{K_0(\mathfrak{p})}(G_{S,T})$  for  $\text{Sh}_{K_0(\mathfrak{p})}(G_{S,T})$  over  $\mathcal{O}_{\mathbf{E}_{\mathfrak{S},\bar{v}}}$ .

*Remark A.1.4.* Alternatively, we may use [TX16, Corollary 2.13] to define a universal  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure over  $\underline{\text{Sh}}_K(G_{S,T})_{\mathcal{O}_{\mathbf{E}_{\mathfrak{S},\bar{v}}}}$  in the sense of Definition A.3.2. We may then define  $\text{Sh}_{K_0(\mathfrak{p})}(G_{S,T})$  as classifying cyclic isogenies of this universal  $p$ -divisible group. We refer to [Car86] for the details in the case of Shimura curves. We prefer to transfer the results from the unitary Shimura variety since in this case we may directly apply theorems which are known for certain PEL type Shimura varieties.

The integral model  $\underline{\text{Sh}}_{K_0(\mathfrak{p})}(G_{S,T})$  is equipped with degeneracy maps

$$\pi_1, \pi_2 : \underline{\text{Sh}}_{K_0(\mathfrak{p})}(G_{S,T}) \rightarrow \underline{\text{Sh}}_K(G_{S,T}).$$

Note also that we are not interested in any canonicity properties of these models. The only thing we use is the existence of such a model with the necessary geometric properties that we will describe in this section. The integral model  $\underline{\text{Sh}}_{K_0(\mathfrak{p})}(G_{S,T})$  is constructed from the connected component

$$\underline{\text{Sh}}_{K_0,\mathfrak{p}}(G_{S,T})_{\mathbb{Z}_p^{\text{ur}}}^\circ \xrightarrow{\sim} \underline{\text{Sh}}_{K'_0,\mathfrak{p}}(G'_S)_{\mathbb{Z}_p^{\text{ur}}}^\circ$$

and the action of a certain group  $\mathcal{E}_{G_{S,T},p}$ , cf. [TX16, 2.11].

For ease of notation, we shall write  $X'$  for the special fiber of  $\underline{\text{Sh}}_{K'}(G'_S)$  and  $X'_0(\mathfrak{p})$  for the special fiber of  $\underline{\text{Sh}}_{K'_0(\mathfrak{p})}(G'_S)$ . We will also write  $X$  and  $X_0(\mathfrak{p})$  for the special fibers of  $\underline{\text{Sh}}_K(G_{S,T})$  and  $\underline{\text{Sh}}_{K_0(\mathfrak{p})}(G_{S,T})$  respectively.

We begin with a few basic facts concerning these moduli spaces.

**Proposition A.1.5.** (1)  $\underline{\text{Sh}}_{K'_0(\mathfrak{p})}(G'_S)$  is a flat normal scheme over  $\mathcal{O}_{\mathbf{E}_{\mathfrak{S},\bar{v}}}$  with reduced special fiber.  
 (2) Each irreducible component of  $X'_0(\mathfrak{p})$  is smooth of dimension 2.

*Proof.* Both of these properties follow from the corresponding properties of the local models; see [Gör01].  $\square$

**A.2. Group theoretic preliminaries.** Let  $G$  be a reductive group over  $\mathbb{Z}_p$ ; in particular its generic fiber is quasi-split and splits over an unramified extension of  $\mathbb{Q}_p$ . Let  $k$  be an algebraically closed field of characteristic  $p$ ,  $L := W(k)[\frac{1}{p}]$  and  $\mathcal{O}_L = W(k)$ . We write  $\sigma$  for the Frobenius element of

$\text{Aut}(L/\mathbb{Q}_p)$ . We fix a maximal  $\mathbb{Q}_p^{\text{ur}}$ -split torus  $T$  and a Borel subgroup  $B$  containing  $G$ . We let  $X_*(T)$  (resp.  $X^*(T)$ ) denote the group of cocharacters (resp. characters) of  $T$ , and  $X_*(T)_+$  (resp.  $X^*(T)_+$ ) the submonoid of dominant cocharacters (resp. characters) with respect to the choice of Borel subgroup  $B$ .

For  $b \in G(L)$  we let  $[b] = \{g^{-1}b\sigma(g) \in G(L) | g \in G(L)\}$  denote its  $\sigma$ -conjugacy class in  $G(L)$  and we write  $B(G)$  to denote the set of all  $\sigma$ -conjugacy classes. The set  $B(G)$  has been classified by Kottwitz in [Kot97].

For  $b \in G(L)$  we let  $\bar{\nu}_b \in X_*(T)_{\mathbb{Q},+}^\sigma$  denote its dominant Newton cocharacter; it depends only on the image of  $b$  in  $B(G)$ . We let

$$\kappa_G : G(L) \rightarrow \pi_1(G)_\Gamma$$

denote the Kottwitz homomorphism, where  $\Gamma := \text{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$  and  $\pi_1(G)_\Gamma$  denotes the  $\Gamma$ -coinvariants of  $\pi_1(G)$ . This induces a map, also denoted  $\kappa_G$  from  $B(G)$  to  $\pi_1(G)_\Gamma$ .

By [Kot97, §4.13], the map

$$B(G) \rightarrow X_*(T)_{\mathbb{Q},+}^\sigma \times \pi_1(G)_\Gamma, \quad [b] \mapsto (\bar{\nu}_b, \kappa_G(b))$$

is injective.

We define a partial order on the set  $X_*(T)_{\mathbb{Q},+}^\sigma \times \pi_1(G)_\Gamma$  by setting  $(\nu_1, \kappa_1) \leq (\nu_2, \kappa_2)$  if  $\kappa_1 = \kappa_2$  and  $\nu_2 - \nu_1$  is a non-negative rational linear combination of positive coroots.

*Example A.2.1.* (1) Let  $G = \text{GL}_n$ . Then we have a bijection

$$B(G) \leftrightarrow \{\text{isocrystals over } k \text{ of height } n\}$$

given by taking  $[b] \in B(G)$  to the isocrystal  $(L^n, b\sigma)$ . We may take  $T$  to be the diagonal matrices and  $B$  the upper triangular matrices. There is an identification  $X_*(T) \cong \mathbb{Q}^n$  and we have  $\pi_1(G) \cong \mathbb{Z}$ . The first isomorphism identifies

$$X_*(T)_{\mathbb{Q},+} \leftrightarrow \{(\nu_1, \dots, \nu_n) \in \mathbb{Q}^n | \nu_1 \geq \dots \geq \nu_n\} \in \mathbb{Q}^n$$

and  $\kappa_G$  takes  $b \in G(L)$  to the valuation of its determinant. For  $[b] \in B(G)$ , the element  $\bar{\nu}_b$  corresponds under the above identification to the Newton slopes of the associated isocrystal. In this case  $\kappa_G(b)$  is determined by the  $\bar{\nu}_b$ , and Kottwitz's classification recovers the Dieudonné–Manin classification of isocrystals by their Newton slopes.

(2) Let  $F$  denote a finite unramified extension of  $\mathbb{Q}_p$  of degree  $d$  and let  $G = \text{Res}_{\mathcal{O}_F/\mathbb{Z}_p} \text{GL}_n$ . As in the previous example, the association  $[b] \mapsto (F \otimes_{\mathbb{Q}_p} L, b\sigma)$  defines a bijection between  $B(G)$  and the set of isocrystals of height  $dn$  with an action of  $F$ . Given such an isocrystal  $N$ , we have a decomposition

$$N \cong \prod_{\tau: F \rightarrow L} N_\tau$$

where  $N_\tau$  is the subspace of  $N$  over which  $F$  acts via the embedding  $\tau : F \rightarrow L$ . As  $N_\tau$  is fixed by  $b \in G(L)$  and  $\sigma$  induces a bijection between  $N_\tau$  and  $N_{\sigma\tau}$ , it follows that  $(b\sigma)^d$  takes  $N_\tau$  to itself. This gives  $N_\tau$  the structure of a  $\sigma^d$ -isocrystal which is easily seen to be independent of the choice of  $\tau$ . One checks that the association  $N \mapsto N_\tau$  induces a bijection

$$\{\text{isocrystals of height } dn \text{ with an action of } F\} \leftrightarrow \{\sigma^d\text{-isocrystals of height } n\}.$$

We let  $T$  denote the diagonal torus and  $B$  the Borel subgroup of upper triangular matrices. We have an identification

$$X_*(T)_{\mathbb{Q}}^\sigma \cong \mathbb{Q}^n$$

identifying

$$X_*(T)_{\mathbb{Q},+}^\sigma \leftrightarrow \{(\nu_1, \dots, \nu_n) | \nu_1 \geq \dots \geq \nu_n\}.$$

For  $[b] \in B(G)$ , the Newton cocharacter  $\bar{\nu}_b$  corresponds under the above identification to the slopes of the  $\sigma^d$ -isocrystal  $(N_\tau, (b\sigma)^d)$ .

By the Cartan decomposition we have an identification

$$G(L) \cong \coprod_{\mu \in X_*(T)_+} G(\mathcal{O}_L)\mu(p)G(\mathcal{O}_L).$$

Now fix a cocharacter  $\mu \in X_*(T)_+$ . We define  $\bar{\mu} \in X_*(T)$  to be the Galois average of  $\mu$ . More precisely we take a Galois extension  $E/\mathbb{Q}_p$  over which  $\mu$  is defined and we define

$$\bar{\mu} = \frac{1}{[E:\mathbb{Q}_p]} \sum_{\tau \in \text{Gal}(E/\mathbb{Q}_p)} \tau(\mu).$$

We also define  $\mu^\natural$  to be the image of  $\mu$  in  $\pi_1(G)_\Gamma$ .

**Proposition A.2.2** ([RR96, Thm. 4.2]). *Let  $b \in G(\mathcal{O}_L)\mu(p)G(\mathcal{O}_L)$ . Then we have*

$$(\bar{\nu}_b, \kappa_G(b)) \leq (\bar{\mu}, \mu^\natural).$$

**A.3.  $p$ -divisible groups with  $\mathcal{O}$ -structure and Dieudonné theory.** In this subsection we recall the notion of  $p$ -divisible groups with an action of the ring of integers of a finite unramified extension of  $\mathbb{Q}_p$ .

Let  $F$  be a finite unramified extension of  $\mathbb{Q}_p$  of degree  $d$  with ring of integers  $\mathcal{O}$ . We let  $q = p^d$  denote the cardinality of its residue field.

**Definition A.3.1.** Let  $S$  be a scheme. A  $p$ -divisible group with  $\mathcal{O}$ -structure over  $S$  is a pair  $(\mathcal{G}, \iota)$  where  $\mathcal{G}$  is a  $p$ -divisible group over  $S$  and  $\iota: \mathcal{O} \rightarrow \text{End}(\mathcal{G})$  is a homomorphism.

For any  $p$ -divisible group with  $\mathcal{O}$ -structure  $(\mathcal{G}, \iota)$ , we write  $\mathcal{G}[p^n]$  for the kernel of multiplication by  $p^n$ . Then there exists an integer  $h := \text{ht}_{\mathcal{O}}\mathcal{G}$ , the  $\mathcal{O}$ -height of  $\mathcal{G}$ , such that  $\mathcal{G}[p^n]$  has rank  $q^{nh}$ . It is easily verified that we have the equality

$$\text{ht}\mathcal{G} = [F:\mathbb{Q}_p]\text{ht}_{\mathcal{O}}\mathcal{G}$$

where  $\text{ht}\mathcal{G}$  is the usual height of  $\mathcal{G}$  as a  $p$ -divisible group.

Let  $S$  be a scheme in which  $p$  is locally nilpotent. For  $\mathcal{G}$  a  $p$ -divisible group over  $S$ , we write  $\mathbb{D}(\mathcal{G})$  for the contravariant Dieudonné crystal of  $\mathcal{G}$ . This is a locally free crystal on the crystalline site of  $S$ , equipped with a map  $\sigma^*\mathbb{D}(\mathcal{G}) \rightarrow \mathbb{D}(\mathcal{G})$ .

Now suppose  $\mathcal{G}$  is a  $p$ -divisible group with  $\mathcal{O}$ -structure over an algebraically closed field  $k$  of characteristic  $p$  of  $\mathcal{O}$ -height  $n$ . Then we identify  $\mathbb{D}(\mathcal{G})$  with its Dieudonné module (i.e.  $\mathbb{D}(\mathcal{G})$  evaluated at  $\mathcal{O}_L := W(k)$ ) which is a finite free  $\mathcal{O}_L$ -module of rank  $dn$  equipped with an injective  $\sigma$  semi-linear map  $\varphi: \mathbb{D}(\mathcal{G}) \rightarrow \mathbb{D}(\mathcal{G})$  and an action of  $\mathcal{O}$ . Fixing an  $\mathcal{O} \otimes_{\mathbb{Z}_p} \mathcal{O}_L$ -basis of  $\mathbb{D}(\mathcal{G})$ , we obtain an element  $b \in G(L)$  where  $G = \text{Res}_{\mathcal{O}/\mathbb{Z}_p} \text{GL}_n$ , such that  $\varphi = b\sigma$  under the identification  $\mathbb{D}(\mathcal{G}) \cong (\mathcal{O}_F \otimes_{\mathbb{Z}_p} \mathcal{O}_L)^n$ . The element  $b$  is well-defined up to  $\sigma$ -conjugation by  $G(\mathcal{O}_L)$ .

We define the *Hodge polygon* of  $\mathcal{G}$  to be the element  $\mu \in X_*(T)^+$  such that

$$b \in G(\mathcal{O}_L)\sigma(\mu(p))G(\mathcal{O}_L).$$

By Proposition A.2.2, it follows that  $(\bar{\nu}_b, \kappa_G(b)) \leq (\bar{\mu}, \mu^\natural)$ . If  $S$  is a scheme of characteristic  $p$ , its Hodge polygon (resp. Newton polygon) is the function assigning to any geometric point  $\bar{s}$  of  $S$  the Hodge polygon (resp. the Newton polygon) of the base change of  $\mathcal{G}$  to  $\bar{s}$ .

We now describe how the study of  $p$ -divisible groups which arise naturally from the moduli problem in the last subsection can be reduced to the case of  $p$ -divisible groups with  $\mathcal{O}$ -structure. Recall we have the integral PEL datum  $(\mathcal{O}_{D_S} \subset D_S, *, W, \psi, \Lambda_1)$  and we write  $\mathcal{D}$  for the base change of this datum to  $\mathbb{Z}_p$ . Recall  $g = [F:\mathbb{Q}]$ . Let  $D_{S,p}$  denote the completion of  $D_S$  at the place  $p$  and  $F_p$  (resp.  $E_p$ ) the completion of  $F$  (resp.  $E$ ) at  $p$ . Then  $F_p \cong F_{\mathfrak{p}}$  and  $E_p \cong E_{\mathfrak{q}} \times E_{\bar{\mathfrak{q}}}$ ; we write  $\mathcal{O}$  for the ring of integers  $\mathcal{O}_{F_p}$ . Fixing isomorphisms  $E_{\mathfrak{q}} \cong F_{\mathfrak{p}}$  and  $E_{\bar{\mathfrak{q}}} \cong F_{\mathfrak{p}}$ , we obtain an isomorphism

$$D_{S,p} \cong \text{Mat}_2(E_{\mathfrak{q}}) \times \text{Mat}_2(E_{\bar{\mathfrak{q}}}) \cong \text{Mat}_2(F_{\mathfrak{p}}) \times \text{Mat}_2(F_{\mathfrak{p}}).$$

Let  $\mathcal{O}_{D_{S,p}}$  denote the maximal order  $\text{Mat}_2(\mathcal{O}) \times \text{Mat}_2(\mathcal{O})$ . Then the involution  $*$  on  $D_{S,p}$  can be identified with

$$(a, b) \mapsto (b^t, a^t).$$

Recall we have defined the notion of  $p$ -divisible group with  $\mathcal{D}$ -structure in Definition 3.2.1. It follows easily from the definitions that if  $(A, \iota, \lambda, \varepsilon_{K'^p})$  is an  $S$ -point of  $\underline{\text{Sh}}_{K'}(G'_S)$ , then the associated  $p$ -divisible group  $A[p^\infty]$  together with the induced  $\mathcal{O}_{D_{S,p}}$ -action and polarization is a  $p$ -divisible group with  $\mathcal{D}$ -structure.

We write  $W^1$  for the sub  $F_p$  vector space over which  $D_{S,p}$  acts via the first factor, and we write  $W^\circ$  for subspace  $eW^1$  where  $e$  is the idempotent  $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ . We define an integer  $s_\tau \in \{0, 1, 2\}$  for  $\tau \in S_\infty$  by

$$s_\tau = \begin{cases} 0 & \text{if } \tau \in \mathbf{T} \\ 2 & \text{if } \tau \in S_\infty - \mathbf{T} \\ 1 & \text{otherwise} \end{cases}.$$

**Definition A.3.2.** Let  $S$  be scheme over  $\mathcal{O}_{\mathbf{E}_{S,v}}$ . A  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure is a  $p$ -divisible group  $\mathcal{G}$  with  $\mathcal{O}$ -structure such that

- (1)  $\text{ht}_{\mathcal{O}} \mathcal{G} = 2$ .
- (2) For  $a \in \mathcal{O}$ , we have an equality:

$$(A.3.1) \quad \text{char}(\iota(a)|\text{Lie} \mathcal{G}) = \prod_{\tau \in \Sigma_\infty} (T - \tau(a))^{s_\tau}.$$

Let  $\mathcal{G}/S$  be a  $p$ -divisible with  $\mathcal{D}$ -structure. Then we have a decomposition

$$\mathcal{G} = \mathcal{G}_q \times \mathcal{G}_{\bar{q}}$$

where  $\mathcal{O}_{D_{S,p}}$  acts on  $\mathcal{G}_q$  via the projection to  $\text{Mat}_2(E_q)$  and on  $\mathcal{G}_{\bar{q}}$  via the projection to  $\text{Mat}_2(E_{\bar{q}})$ . Moreover by [Ham15, Lemma 4.1], there exists a  $p$ -divisible group  $\mathcal{G}'$  with  $\mathcal{O}$ -structure such that  $\mathcal{G}_q \cong \mathcal{G}'^2$ . The condition (3.2.1) and the Assumption A.1.2 implies that  $\mathcal{G}'$  is a  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure.

The following proposition follows from the discussion above and [Ham15, Corollary 4.5 (2)].

**Proposition A.3.3.** *The association  $\mathcal{G} \mapsto \mathcal{G}'$  induces an equivalence of categories*

$$(A.3.2) \quad \{p\text{-divisible groups with } \mathcal{D}\text{-structure}\} \xrightarrow{\sim} \{p\text{-divisible groups with } \mathcal{D}^\circ\text{-structure}\}$$

*preserving isogenies.* □

Let  $\mathcal{G}'$  be a  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure over an algebraically closed field  $k$  of characteristic  $p$ . Fixing a trivialization of  $\mathbb{D}(\mathcal{G}')(W(k))$  respecting the  $\mathcal{O}$ -structure, we obtain an element  $b \in G(L)$  where  $G = \text{Res}_{\mathcal{O}/\mathbb{Z}_p} \text{GL}_2$ . Let  $T$  be the diagonal maximal torus of  $G$ , then we may identify  $X_*(T)$  (resp.  $X_*(T)_{\mathbb{Q}}$ ) with  $g$  copies of  $\mathbb{Z}^2$  (resp.  $g$  copies of  $\mathbb{Q}^2$ ) and  $X_*(T)_+$  (resp.  $X_*(T)_{\mathbb{Q},+}$ ) with the subset such that for each factor of  $\mathbb{Z}^2$  (resp.  $\mathbb{Q}^2$ ), the terms  $(a, b)$  are decreasing. The condition A.3.1 implies that the Hodge polygon  $\mu$  of  $\mathcal{G}'$  corresponds to  $[(a_i, b_i)]_{i=1, \dots, g}$  where

$$a_i = \begin{cases} 1 & \text{if } i = 1, c \text{ or } \tau_i \in S_\infty - \mathbf{T} \\ 0 & \text{if } \tau_i \in \mathbf{T} \end{cases}$$

$$b_i = \begin{cases} 1 & \text{if } \tau_i \in S_\infty - \mathbf{T} \\ 0 & \text{if } i = 1, c \text{ or } \tau_i \in \mathbf{T}. \end{cases}$$

We write  $\mu \in X_*(T)$  for this cocharacter.

**Proposition A.3.4.** (1) *There exists exactly two elements  $[b] \in B(G, \mu)$ .*

(2) For the maximal element  $[\mu(p)]$ , there exists a unique  $p$ -divisible group with  $\mathcal{O}$ -structure with these Newton slopes.

*Proof.* (1) It is easy to check using the explicit description of  $\mu$  that under the identification  $X_*(T)^\sigma \cong \mathbb{Q}^2$ , that the only two elements of  $B(G, \{\mu\})$  correspond to  $\nu^{\text{ord}} = (\frac{g+1}{2}, \frac{g-1}{2})$  and  $\nu^{\text{ss}} = (\frac{g}{2}, \frac{g}{2})$ . Here we use the fact that  $2|T| = |S|$ .

(2) This is [Moo04, Theorem 3.2.7]; note in this case the Newton cocharacter is  $\nu^{\text{ord}}$ .  $\square$

We say a  $p$ -divisible group  $\mathcal{G}/k$  with  $\mathcal{D}^\circ$ -structure is ordinary (resp. supersingular) if the corresponding Newton vector is equal to  $\nu^{\text{ord}}$  (resp.  $\nu^{\text{ss}}$ ). Similarly a  $p$ -divisible group  $\mathcal{G}$  with  $\mathcal{D}$ -structure is ordinary (resp. supersingular) if the corresponding  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure is. For  $k$  an algebraically closed field of characteristic  $p$ , we write  $\mathcal{G}^{\text{ord}}$  for the unique isomorphism class of ordinary  $p$ -divisible groups from part (2) of Proposition A.3.4.

If  $A_{X'}$  denotes the universal abelian variety over  $X'$ , we write  $X'^{\text{ss}}$  (resp.  $X'^{\text{ord}}$ ) for the locus where  $A_{X'}[p^\infty]$  is supersingular (resp. ordinary). Similarly if  $A \rightarrow A'$  denotes the universal cyclic isogeny over  $X'_0(\mathfrak{p})$ , we define  $X'_0(\mathfrak{p})^{\text{ss}}$  (resp.  $X'_0(\mathfrak{p})^{\text{ord}}$ ) as the locus where  $A[p^\infty]$  (equivalently  $A'[p^\infty]$ ) is supersingular (resp. ordinary). Then  $X'^{\text{ss}}$  is the union of the Goren–Oort divisors corresponding to  $\tau_1$  and  $\tau_c$ ; see [LT20, §3] for example.

**Proposition A.3.5.** *Let  $x \in X'_0(\mathfrak{p})^{\text{ord}}(\overline{\mathbb{F}}_p)$ . Then  $x$  is a smooth point of  $X'_0(\mathfrak{p})$ .*

*Proof.* The local model in this case has a stratification by the  $\mu$ -admissible set  $\text{Adm}(\{\mu\})$  and this induces a stratification of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$ , cf. [HZ20, §9]. The strata corresponding to translation elements are all smooth. By [HN17, Theorem 2.6],  $X'_0(\mathfrak{p})^{\text{ord}}$  is contained in these strata, hence  $x$  is a smooth point of  $X'_0(\mathfrak{p})$ . Note that the Axioms of [HR17] for these Shimura varieties have been verified in [HZ20] so that [HN17, Theorem 2.6] is applicable.  $\square$

**Definition A.3.6.** Let  $(\mathcal{G}, \lambda, \iota)$  and  $(\mathcal{G}', \lambda', \iota')$  be  $p$ -divisible groups with  $\mathcal{D}$ -structure over  $S$ . A *cyclic isogeny* between  $\mathcal{G}$  and  $\mathcal{G}'$  over  $S$  is an isogeny  $f : \mathcal{G} \rightarrow \mathcal{G}'$  such that

- $f$  is compatible with the actions  $\iota, \iota'$  and we have  $p\lambda = f^\vee \circ \lambda' \circ f$ .
- $\ker(f) = K_{\mathfrak{q}} \oplus K_{\overline{\mathfrak{q}}} \subset \mathcal{G}_{\mathfrak{q}}[\mathfrak{q}] \oplus \mathcal{G}_{\overline{\mathfrak{q}}}[\mathfrak{q}]$  is of order  $|k_{\mathfrak{p}}|^4$  and  $K_{\mathfrak{q}}$  is dual to  $K_{\overline{\mathfrak{q}}}$  under the pairing induced by  $\lambda$ .

Similarly, for  $p$ -divisible groups with  $\mathcal{D}^\circ$ -structure  $\mathcal{G}$  and  $\mathcal{G}'$  over  $S$ , a *cyclic isogeny* between  $\mathcal{G}$  and  $\mathcal{G}'$  is an isogeny  $f : \mathcal{G} \rightarrow \mathcal{G}'$  such that:

- $f$  is compatible with the action of  $\mathcal{O}$  on  $\mathcal{G}$  and  $\mathcal{G}'$ .
- $\ker(f) \subset \mathcal{G}[p]$  is of order  $|k_{\mathfrak{p}}|$ .

One can check that under the equivalence of categories in Proposition A.3.3, the cyclic isogenies correspond to one another. Moreover if  $(A, \iota, \lambda, \varepsilon_{K'^p})$  is an  $S$ -point of  $\text{Sh}_{K'}(G'_S)$ , then a cyclic isogeny  $A \rightarrow A'$  corresponds precisely to a cyclic isogeny of the associated  $p$ -divisible group.

**A.4. Essential Frobenius and Verschiebung isogenies.** In this section we define two canonical cyclic isogenies associated to a point in  $X'$ . These isogenies will define sections to the projections  $\pi_1, \pi_2 : X'_0(\mathfrak{p}) \rightarrow X'$  analogous to the Frobenius and Verschiebung isogenies in the case of the Hilbert modular surface (see [Sta97, §4]). For this reason we will call these isogenies the essential Frobenius and essential Verschiebung isogenies respectively.

We let  $q = p^g$  and let  $(A, \iota, \lambda, \varepsilon_{K'^p})$  correspond to an  $S$ -point of  $X'$  where  $S$  is a smooth  $\mathbb{F}_q$ -scheme. We first define the essential Frobenius isogeny  $A \rightarrow A'$ . Note that in order to define a cyclic isogeny it suffices to define a cyclic isogeny of the associated  $p$ -divisible group  $\mathcal{G} := A[p^\infty]$ . Let  $\mathcal{G} = (\mathcal{G}')^2 \times (\mathcal{G}'^\vee)^2$  denote the decomposition of  $\mathcal{G}$  coming from Proposition A.3.3; then it suffices to define a cyclic isogeny of  $\mathcal{G}'$ . In order to do this we introduce some notation.

Let  $k$  be a perfect field of characteristic  $p$  and  $R$  a smooth  $k$ -algebra. By a frame for  $R$  we mean a  $p$ -adically complete and separated flat  $W(k)$ -algebra  $\mathcal{R}$  lifting  $R$  together with a lift of Frobenius  $\sigma : \mathcal{R} \rightarrow \mathcal{R}$ .

A Dieudonné  $F$ -crystal over  $R$  is a quadruple  $(M, F, V, \nabla)$  where

- $M$  is a finite locally free  $\mathcal{R}$  module.
- $F : \sigma^* M \rightarrow M$  and  $V : M \rightarrow \sigma^* M$  are injective  $\mathcal{R}$ -linear maps such that  $FV = p$  and  $VF = p$ .
- $\nabla$  is a topologically nilpotent integrable connection such that  $F$  is parallel for  $\nabla$ .

If  $\mathcal{G}$  is a  $p$ -divisible group over  $R$ , then by [dJ95]  $\mathbb{D}(\mathcal{G})(\mathcal{R})$  is a Dieudonné  $F$ -crystal and the association  $\mathcal{G} \mapsto \mathbb{D}(\mathcal{G})(\mathcal{R})$  induces an anti-equivalence of categories between  $p$ -divisible groups over  $R$  and Dieudonné  $F$ -crystals. Similarly the association induces an anti-equivalence of categories between  $p$ -divisible groups with  $\mathcal{O}$ -structure and Dieudonné crystals with an action of  $\mathcal{O}$ .

It will follow from the canonicity of the construction that we may assume  $S$  is affine, since we may glue the construction over an affine cover. Thus let  $S = \text{Spec } R$  be a smooth  $\mathbb{F}_q$ -scheme and  $\mathcal{R}$  a frame for  $R$  as above. Then in order to define a cyclic isogeny, it suffices to find an  $\mathcal{R}$ -lattice  $M \subset \mathbb{D}(\mathcal{G})(\mathcal{R})$  satisfying the following conditions:

- (1)  $M$  is stable for the action of  $\mathcal{O}$
- (2)  $M$  is stable under  $F, V, \nabla$ .
- (3) We have the inclusions  $p\mathbb{D}(\mathcal{G})(\mathcal{R}) \subset M \subset \mathbb{D}(\mathcal{G})(\mathcal{R})$ .
- (4) (A.3.1) holds for the module  $M/VM$ .

We now construct such an  $\mathcal{R}$ -lattice. For  $\tau \in \Sigma_\infty$ , we identify this with an embedding  $\tau : \mathcal{O} \rightarrow W(\mathbb{F}_q)$  and we let  $\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau$  the submodule where  $\mathcal{O}$  acts via  $\tau$ . Similarly to §2.4 we define the essential Verschiebung to be

$$V_{\text{es}, \tau} : \mathbb{D}(\mathcal{G})(\mathcal{R})_\tau \rightarrow \mathbb{D}(\mathcal{G})(\mathcal{R})_{\sigma^{-1}(\tau)}$$

to be the usual Verschiebung if  $\sigma^{-1}(\tau) \notin S_\infty$  or  $\sigma^{-1}(\tau) \in S_\infty - T$  and the inverse of Frobenius if  $\sigma^{-1}(\tau) \in T$ . For  $\tau \in \Sigma_\infty$ , let  $n_\tau$  be the smallest positive integer such that  $\sigma^{-n_\tau}(\tau) \notin S_\infty$ . We define  $M_\tau \subset \mathbb{D}(\mathcal{G})(\mathcal{R})_\tau$  to be the preimage of  $p\mathbb{D}(\mathcal{G})(\mathcal{R})_{\sigma^{-n_\tau}(\tau)}$  under the map  $V_{\text{es}}^{n_\tau}$ . Then we define  $M := \bigoplus_{\tau \in \Sigma_\infty} M_\tau$ . Then  $M$  is a locally free  $\mathcal{R}$ -submodule of  $\mathbb{D}(\mathcal{G})(\mathcal{R})$  of full rank.

**Proposition A.4.1.**  *$M \subset \mathbb{D}(\mathcal{G})(\mathcal{R})$  satisfies the properties (1), (2), (3), (4), above.*

*Proof.* (1) is clear since  $M$  is a direct sum of locally free submodules of  $\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau$ . To show  $M$  is stable under  $F$ , we must show  $F(M_\tau) \subset M_{\sigma(\tau)}$ . We consider the separate cases  $\tau \in S_\infty$  and  $\tau \notin S_\infty$ . If  $\tau \in S_\infty$ , we have

$$V_{\text{es}}^{n_{\sigma(\tau)}}|_{\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau} = V_{\text{es}}^{n_\tau} \circ V_{\text{es}, \sigma(\tau)}$$

and  $V_{\text{es}, \sigma(\tau)}$  is the usual Verschiebung or the inverse of Frobenius. Since  $VF = p$ , we have  $V_{\text{es}, \sigma(\tau)}FM_\tau \subset M_\tau$  and hence  $FM_\tau \subset V_{\text{es}, \sigma(\tau)}^{-1}M_\tau = M_{\sigma(\tau)}$ .

In the second case we have

$$M_{\sigma(\tau)} = V_{\text{es}, \sigma(\tau)}^{-1}(p\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau) = F(\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau).$$

Since

$$M_\tau = (V_{\text{es}}^{n_\tau})^{-1}(p\mathbb{D}(\mathcal{G})(\mathcal{R})_{\sigma^{n_\tau}(\tau)}) \subset \mathbb{D}(\mathcal{G})(\mathcal{R})_\tau,$$

we have  $F(M_\tau) \subset M_{\sigma(\tau)}$ . The verification of the stability for  $V$  follows similarly.

The stability under  $\nabla$  follows from the stability of  $p\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau$  under  $\nabla$  and the fact that  $F$  and  $V$  are horizontal for  $\nabla$ . It follows that (2) is satisfied.

(3) follows from the inclusions

$$pV\mathbb{D}(\mathcal{G})(\mathcal{R})_{\sigma(\tau)} \subset p\mathbb{D}(\mathcal{G})(\mathcal{R})_\tau \subset V\mathbb{D}(\mathcal{G})(\mathcal{R})_{\sigma(\tau)}$$

for  $\tau \in \Sigma_\infty - S_\infty$ .

(4) is equivalent to the condition

$$\dim(M/VM)_\tau = \begin{cases} 1 & \text{if } \tau \in \Sigma_\infty - S_\infty \\ 0 & \text{if } \tau \in S_\infty - T \\ 2 & \text{if } \tau \in T \end{cases}$$

which follows similarly.  $\square$

It follows that  $M$  corresponds to a  $p$ -divisible group  $\mathcal{G}^{(p')}$  equipped with a cyclic isogeny  $\text{Fr}' : \mathcal{G} \rightarrow \mathcal{G}^{(p')}$ . We define  $\mathcal{G}^{(p')}$  to be the *essential Frobenius twist* of  $\mathcal{G}$  and  $\text{Fr}'$  to be the *essential Frobenius isogeny*. Similarly we may define the *essential Verschiebung*

$$\text{Ver}' : \mathcal{G}^{(p')} \rightarrow \mathcal{G}$$

to be defined by the submodule

$$p\mathbb{D}(\mathcal{G})(\mathcal{R}) \subset M.$$

It follows from the definition that  $\text{Fr}' \circ \text{Ver}' = p$  and  $\text{Ver}' \circ \text{Fr}' = p$ .

Let  $A$  be the universal abelian scheme over  $X'$ . Taking a cover of  $X'$  by smooth affine opens and gluing the construction we obtain isogenies

$$\text{Fr}'_A : A \rightarrow A^{(p')}, \quad \text{Ver}'_A : A^{(p')} \rightarrow A.$$

**Proposition A.4.2.** *We have*

$$(A^{(p')})^{(p')} = A^{(p^g)}$$

where  $A^{(p^g)}$  is the usual  $p^g$ -Frobenius twist. Moreover the composition  $\text{Fr}' \circ \text{Fr}'$  considered as a map  $X' \rightarrow X'$  corresponds to  $S_p^{-(g/2-1)}\text{Fr}_p$  where  $S_p$  is the standard Hecke operator at  $p$  and  $\text{Fr}_p$  is the  $p^g$ -Frobenius.

*Proof.* We may reduce to the case  $S = \text{Spec } R$  a smooth affine scheme. Moreover, it suffices to prove this for the  $p$ -divisible  $\mathcal{G}'$  associated to  $A[p^\infty]$  by Proposition A.3.3.

The submodule  $F^g\mathbb{D}(\mathcal{G}')(\mathcal{R}) \subset \mathbb{D}(\mathcal{G}')(\mathcal{R})$  corresponds to the  $p^g$ -Frobenius isogeny. We let  $M \subset \mathbb{D}(\mathcal{G}')(\mathcal{R})$  be the submodule corresponding to  $\mathcal{G}'^{(p')}$  and  $M' \subset M$  the submodule corresponding to  $(\mathcal{G}'^{(p')})^{(p')}$ . Then by the definition of  $M'$ , we have that  $p^{\frac{g}{2}-1}M' = F^g\mathbb{D}(\mathcal{G}')(\mathcal{R})$ . In particular,  $F^g\mathbb{D}(\mathcal{G}')(\mathcal{R}) \subset M'$  corresponds to the isogeny

$$(A.4.1) \quad p^{\frac{g}{2}-1} : (\mathcal{G}'^{(p')})^{(p')} \rightarrow (\mathcal{G}'^{(p')})^{(p')}.$$

Therefore  $(\mathcal{G}'^{(p')})^{(p')}$  and  $\mathcal{G}^{(p)}$  are isomorphic and the moreover part follows.  $\square$

**Proposition A.4.3.** *Let  $\mathcal{G}$  be an ordinary  $p$ -divisible with  $\mathcal{O}$ -structure over a reduced and irreducible  $\mathbb{F}_q$ -scheme  $S$ . Let  $f : \mathcal{G} \rightarrow \mathcal{G}'$  be a cyclic isogeny, then  $f = \text{Fr}'$  or  $f = \text{Ver}'$ .*

*Proof.* First assume  $S = \text{Spec } k$  where  $k$  is an algebraically closed field of characteristic  $p$ . Then  $\mathcal{G} \cong \mathcal{G}^{\text{ord}}$  by Proposition A.3.4. The same proof as in [Sta97, Proposition 4.3] shows that  $f = \text{Fr}'$  or  $f = \text{Ver}'$ .

Let  $S^{\text{Fr}'}$  (resp  $S^{\text{Ver}'}$ ) denote the subscheme of  $S$  where  $f$  coincides with  $\text{Fr}'$  (resp.  $\text{Ver}'$ ). Then  $S^{\text{Fr}'}$  and  $S^{\text{Ver}'}$  are closed subschemes of  $S$ , since the locus where two subgroups of  $\mathcal{G}$  coincide is closed. For each closed  $s$  point of  $S$ , we have by the case above that  $f_s = \text{Fr}'$  or  $f_s = \text{Ver}'$ . Therefore  $S^{\text{Fr}'} \cup S^{\text{Ver}'} = S$ , hence by irreducibility  $S = S^{\text{Ver}'}$  or  $S = S^{\text{Fr}'}$ .  $\square$

The association  $(A, \iota, \lambda, \varepsilon_{K'/p}) \mapsto (A^{(p')}, \iota', \lambda', \varepsilon'_{K'/p})$  induces a map

$$\text{Fr}' : X' \rightarrow X'$$

By Proposition A.4.2, we have  $\text{Fr}' \circ \text{Fr}' = S_p^{-(g/2-1)}\text{Fr}_p$ . These isogenies induce maps

$$\mathcal{F}, \mathcal{V} : X' \rightarrow X'_0(\mathfrak{p})$$



which are defined by

$$\begin{aligned}\mathcal{F}(A, \iota, \lambda, \varepsilon_{K'^p}) &= (\mathrm{Fr}'_A : A \rightarrow A^{(p')}, \iota, \lambda, \varepsilon_{K'^p}) \\ \mathcal{V}(A, \iota, \lambda, \varepsilon_{K'^p}) &= (\mathrm{Ver}'_A : A^{(p')} \rightarrow A, \iota', \lambda', \varepsilon'_{K'^p}).\end{aligned}$$

By definition we have the following properties

$$\begin{aligned}\pi_1 \circ \mathcal{F} &= \mathrm{id}_{X'}, \quad \pi_2 \circ \mathcal{F} = \mathrm{Fr}' \\ \pi_1 \circ \mathcal{V} &= S_{\mathfrak{p}}^{-1} \mathrm{Fr}', \quad \pi_2 \circ \mathcal{V} = \mathrm{id}_{X'}.\end{aligned}$$

We have the following proposition regarding these maps.

**Proposition A.4.4.** (1)  $\mathcal{F}$  and  $\mathcal{V}$  are closed immersions.

(2) Let  $(\rho : A \rightarrow A', \iota, \lambda, \varepsilon_{K'^p})$  be an  $S$ -point of  $X_0(\mathfrak{p})^{\mathrm{ord}}$ , where  $S$  is a reduced scheme, then  $x \in \mathcal{F}(X) \cup \mathcal{V}(X)$ .

*Proof.* (1) We prove the result for  $\mathcal{F}$ , the case of  $\mathcal{V}$  is analogous. First note that  $\mathcal{F}$  is injective on points since  $\pi_1 \circ \mathcal{F} = \mathrm{id}_X$ . Therefore it suffices to show  $\mathcal{F}$  is proper; the valuative criterion in this case follows from standard properties of Neron models.

(2) Follows directly from Proposition A.4.3.  $\square$

**A.5. Global structure of quaternionic Shimura surfaces with Iwahori level structure.** In this subsection we prove the main theorem of the appendix. We first need to study the fibers of the map  $\pi_1$ .

Let  $\mathcal{G}/\overline{\mathbb{F}}_p$  be a supersingular  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure. We will define a universal cyclic isogeny of  $\mathcal{G}$ ; this will parameterize the fiber of the projection  $\pi_1 : X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p} \rightarrow X'_{\overline{\mathbb{F}}_p}$ . As in §2.4, we may define a version of essential Verschiebung for any  $p$ -divisible group with  $\mathcal{D}^\circ$  structure  $\mathcal{G}$  over an  $\overline{\mathbb{F}}_p$ -scheme  $S$ . We have the exact sequence of sheaves over  $S$

$$(A.5.1) \quad 0 \rightarrow \omega_{\mathcal{G}, \tau_i} \rightarrow \mathbb{D}(\mathcal{G})(S)_{\tau_i} \rightarrow \omega_{\mathcal{G}^\vee, \tau_i}^* \rightarrow 0$$

for each  $i = 1, \dots, g$ , where  $\mathcal{G}^\vee$  denotes the Cartier dual and  $\omega_{\mathcal{G}}$  (resp.  $\omega_{\mathcal{G}^\vee}$ ) is the sheaf of invariant differential forms on  $\mathcal{G}$  (resp.  $\mathcal{G}^\vee$ ). If  $\mathcal{G}$  arises from an abelian variety with  $\mathcal{D}$ -structure  $A$ , then  $\omega_{\mathcal{G}^\vee, \tau_i} \cong \omega_{A^\vee, \tilde{\tau}_i}^\circ$  as in §2.4, where  $\tilde{\tau}_i \in \Sigma_{E, \infty/q}$  lifts  $\tau_i$ . We obtain sections

$$h_{\tau_i} \in \Gamma(S, \omega_{\mathcal{G}^\vee, \sigma^{-n_{\tau_i} \tau_i}}^{\otimes p^{n_{\tau_i}}} \otimes \omega_{\mathcal{G}^\vee, \tau_i}^{\otimes -1})$$

for  $i = 1, c$ , using the same construction from §2.4; here we need to use the essential *Frobenius* to define these sections since we are using contravariant Dieudonné theory. Applying this to the universal  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure on  $X'$ , these may be identified with the partial Hasse invariants defined in §2.4.

Let  $\mathcal{G}/\overline{\mathbb{F}}_p$  be a supersingular  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure. We separate the following two cases:

Case (1): Both  $h_{\tau_1}$  and  $h_{\tau_c}$  vanish.

Case (2): Exactly one of  $h_{\tau_1}$  or  $h_{\tau_c}$  vanish.

The following proposition can be proved in the same way as [Sta97, §5] where the analogous calculations are carried out for the Hilbert modular surface.

**Proposition A.5.1.** Let  $\mathcal{O}_L := W(\overline{\mathbb{F}}_p)$ .

In Case (1), there exists an  $\mathcal{O}_L$ -basis  $\{e_i, f_i\}$  of  $\mathbb{D}(\mathcal{G})(\mathcal{O}_L)_{\tau_i}$  for  $i = 1, \dots, g$  such that

$$\begin{aligned}\varphi(e_i) &= f_{i+1}, \varphi(f_i) = pe_{i+1}, \text{ for } i = 1, c \\ \varphi(e_i) &= pe_{i+1}, \varphi(f_i) = pf_{i+1}, \text{ for } \tau_i \in S_\infty - \mathbb{T} \\ \varphi(e_i) &= e_{i+1}, \varphi(f_i) = f_{i+1}, \text{ for } \tau_i \in \mathbb{T}\end{aligned}$$

In Case (2); we assume without loss of generality  $h_{\tau_c} \neq 0$ . Then there exists an element  $u \in \overline{\mathbb{F}} - \mathbb{F}_{p^g}$  and an  $\mathcal{O}_L$ -basis  $\{e_i, f_i\}$  of  $\mathbb{D}(\mathcal{G})(\mathcal{O}_L)_{\tau_i}$  such that

$$\begin{aligned} \varphi(e_i) &= f_{i+1}, \quad \varphi(f_i) = pe_{i+1} - [u^{p^{g-c+1}}]f_{i+1}, \quad \text{for } i = 1 \\ \varphi(e_i) &= f_{i+1} + [u^p]e_{i+1}, \quad \varphi(f_i) = pe_{i+1}, \quad \text{for } i = c \\ \varphi(e_i) &= pe_{i+1}, \quad \varphi(f_i) = pe_{i+1}, \quad \text{for } \tau_i \in S_\infty - T \\ \varphi(e_i) &= e_{i+1}, \quad \varphi(f_i) = f_{i+1}, \quad \text{for } \tau_i \in T \end{aligned}$$

where  $[u] \in \mathcal{O}_L$  denotes the Teichmüller lifting.  $\square$

We now define the universal cyclic isogeny of  $\mathcal{G}$ . In the two cases above we fix a basis of  $\mathbb{D}(\mathcal{G})(\mathcal{O}_L)_{\tau_i}$  as in the previous proposition.

Case (1): In this case we will define a cyclic isogeny of  $\mathcal{G} \times T$  where  $T$  is the scheme consisting of two copies of  $\mathbb{P}^1$  intersecting transversally at an  $\overline{\mathbb{F}}_p$ -point. To do this we first introduce some notation.

Let  $S = \text{Spec } R$  be a smooth  $\overline{\mathbb{F}}_p$ -scheme.

Let  $\mathcal{R} = \mathcal{O}_L\langle x \rangle$  denote the ring of restricted power series over  $\mathcal{O}_L$  equipped with the lift of Frobenius given by the usual Frobenius on  $\mathcal{O}_L$  and  $x \mapsto x^p$ . Then  $\mathcal{R}$  is a frame for  $R = \overline{\mathbb{F}}_p[x]$ .

We let  $M$  be the  $\mathcal{R}$ -module  $\mathbb{D}(\mathcal{G})(\mathcal{O}_L) \otimes_{\mathcal{O}_L} \mathcal{R}$ , equipped with the induced Frobenius, Verschiebung and the trivial connection

$$\nabla := 1 \otimes d,$$

where  $d : \mathcal{R} \rightarrow \Omega_{\mathcal{R}/\mathcal{O}_L}^1$  is the universal derivation. Then  $(M, F, V, \nabla)$  is the Dieudonné  $F$ -crystal associated to  $\mathcal{G} \otimes_{\overline{\mathbb{F}}_p} R$ . We now define the submodule  $\tilde{M}_0 \subset M$  to be the submodule generated by generated by  $pM$  and

$$\langle e_i + x^{p^{i-c}} f_i \mid i = c+1, \dots, g, 1 \rangle \cup \langle f_i \mid i = 2, \dots, c \rangle.$$

One checks that  $\tilde{M}_0$  is stable under  $F, V$  and  $\nabla$  hence corresponds to a cyclic isogeny

$$\tilde{\rho}_0 : \mathcal{G} \times \text{Spec } R \rightarrow \tilde{\mathcal{G}}_0$$

where  $\tilde{\mathcal{G}}_0$  is a  $p$ -divisible group with  $\mathcal{O}$ -structure over  $R$ .

Similarly letting  $\mathcal{R}' = \mathcal{O}_L\langle y \rangle$ , we define the submodule  $\tilde{M}'_0$  of  $M' := \mathbb{D}(\mathcal{G})(\mathcal{O}_L) \otimes_{\mathcal{O}_L} \mathcal{R}'$  to be generated by  $pM'$  and

$$\langle y^{p^{i-c}} e_i + f_i \mid i = c+1, \dots, g, 1 \rangle \cup \langle f_i \mid i = 2, \dots, c \rangle.$$

As before this corresponds to a cyclic isogeny

$$\tilde{\rho}'_0 : \mathcal{G} \times \text{Spec } R \rightarrow \tilde{\mathcal{G}}'_0.$$

Using the identification  $x \leftrightarrow \frac{1}{y}$ , we see that  $\tilde{\rho}_0$  and  $\tilde{\rho}'_0$  glue to give an isogeny

$$\rho_0 : \mathcal{G} \times \mathbb{P}^1 \rightarrow \mathcal{G}_0.$$

Similarly we may define an isogeny

$$\rho_1 : \mathcal{G} \times \mathbb{P}^1 \rightarrow \mathcal{G}_1$$

by gluing the isogenies corresponding to submodules  $\tilde{M}_1 \subset M$  and  $\tilde{M}'_1 \subset M'$  where

$$\langle f_i \mid i = c+1, \dots, g, 1 \rangle \cup \langle e_i + x^{p^{i-1}} f_i \mid i = 2, \dots, c \rangle.$$

and

$$\langle f_i \mid i = c+1, \dots, g, 1 \rangle \cup \langle y^{p^{i-1}} e_i + f_i \mid i = 2, \dots, c \rangle.$$

Let  $z_0 = (1 : 0)$  and  $z_1 = (0 : 1)$ . Then  $\rho_0|_{z_0} : \mathcal{G} \rightarrow \mathcal{G}_{0, z_0}$  and  $\rho_1|_{z_1} : \mathcal{G} \rightarrow \mathcal{G}_{1, z_1}$  agree and hence we may glue  $\rho_0$  and  $\rho_1$  to give an isogeny

$$\rho : \mathcal{G} \times T \rightarrow \mathcal{G}'.$$

Case (2): Assume  $h_{\tau_c} \neq 0$ .

As in case (1), we let  $\mathcal{R} = \mathcal{O}_L\langle x \rangle$  and  $M = \mathbb{D}(\mathcal{G})(\mathcal{R})$ . We define the submodule  $\tilde{M}$  of  $M$  to be generated by  $pM$  and

$$\langle e_i + x^{p^{i-c}} f_i \mid i = c+1, \dots, g, 1 \rangle \cup \langle f_i \mid i = 2, \dots, c \rangle.$$

Similarly  $\tilde{M}'$  is the submodule generated by  $pM$  and

$$\langle y^{p^{i-c}} e_i + f_i \mid i = c+1, \dots, g, 1 \rangle \cup \langle f_i \mid i = 2, \dots, c \rangle.$$

As in Case (1), these submodules correspond to isogenies which glue along  $x \leftrightarrow \frac{1}{y}$  and define an isogeny  $\mathcal{G} \times \mathbb{P}^1 \rightarrow \mathcal{G}'$ .

If  $h_{\tau_1} \neq 0$ , we may switch the roles of 1 and  $c$  to obtain an isogeny  $\mathcal{G} \times \mathbb{P}^1 \rightarrow \mathcal{G}'$ .

Let  $x \in X'(\overline{\mathbb{F}}_p)$  lie in the intersection of the two Goren–Oort divisors  $X'_{\tau_1, \overline{\mathbb{F}}_p}$ ,  $X'_{\tau_c, \overline{\mathbb{F}}_p}$  corresponding to  $\tau_1$  and  $\tau_c$ ; we write  $A_x$  for the associated abelian variety and  $\mathcal{G}'_x$  for the  $p$ -divisible group with  $\mathcal{D}^\circ$ -structure associated to  $A_x[p^\infty]$  as in Proposition A.3.3. Then we are in Case (1) and have defined a cyclic isogeny

$$\rho : \mathcal{G}'_x \times T \rightarrow \mathcal{G}'$$

where  $T$  is the union of two copies of  $\mathbb{P}^1$  intersecting transversally at point. This induces a cyclic isogeny  $A_x \times T \rightarrow A'$ . We therefore obtain a map

$$\beta : T \rightarrow X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$$

whose image lies in the fiber of  $x$  under the degeneracy map  $\pi_1$ .

Similarly if  $x \in X'(\overline{\mathbb{F}}_p)$  lies in a unique Goren–Oort strata for  $\tau_1$  or  $\tau_c$ , then we are in Case (2) and we obtain a cyclic isogeny  $A_x \times \mathbb{P}^1 \rightarrow A'$ . This induces a map

$$\alpha : \mathbb{P}^1 \rightarrow X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$$

whose image lies in the fiber of  $x$  under  $\pi_1$ .

**Proposition A.5.2.** *In Case (1), the map  $\beta : T \rightarrow \pi_1^{-1}(x)$  is a bijection on  $\overline{\mathbb{F}}_p$ -points.*

*In Case (2), the map  $\alpha : \mathbb{P}^1 \rightarrow \pi_1^{-1}(x)$  is a bijection on  $\overline{\mathbb{F}}_p$ -points.*

*Proof.* This follows from the same calculation as in [Sta97, Proposition 6.5].  $\square$

**Theorem A.5.3.** (1)  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  can be decomposed as

$$\mathcal{V}(X'_{\overline{\mathbb{F}}_p}) \cup \mathcal{F}(X'_{\overline{\mathbb{F}}_p}) \cup X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}$$

where each term is a union of irreducible components of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  of dimension 2. Any irreducible component of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  is contained in exactly one of the terms and  $\mathcal{V}(X'_{\overline{\mathbb{F}}_p}) \cap \mathcal{F}(X'_{\overline{\mathbb{F}}_p})$  is 0-dimensional.

(2) For each irreducible component  $C$  of  $X'_{\overline{\mathbb{F}}_p}$ , there is a unique irreducible component  $R$  of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  such that  $\pi_1(R) = C$ . Moreover the projection  $\pi_1 : R \rightarrow C$  exhibits  $R$  as a  $\mathbb{P}^1$ -bundle over  $C \cong \mathbb{P}^1$  and for any other irreducible component  $C'$  of  $X'_{\overline{\mathbb{F}}_p}$ , we have:

$$R \cap R' = \begin{cases} \{pt\} & \text{if } C \cap C' \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}.$$

Similarly, the result holds if we replace  $\pi_1$  by  $\pi_2$ .

The rest of this subsection will be devoted to the proof. We refer to [Sta97, p409] for a pictorial representation of the geometry.

*Proof.* (1) By the theory of local models every irreducible component of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  has dimension 2. Now since  $\mathcal{F}$  and  $\mathcal{V}$  are closed immersions,  $\mathcal{V}(X'_{\overline{\mathbb{F}}_p})$  and  $\mathcal{F}(X'_{\overline{\mathbb{F}}_p})$  are unions of irreducible components of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$ . Let  $Z$  be an irreducible component of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  not contained in  $\mathcal{F}(X'_{\overline{\mathbb{F}}_p}) \cup \mathcal{V}(X'_{\overline{\mathbb{F}}_p})$  and let

$x \in Z(\overline{\mathbb{F}}_p)$ . If  $x$  corresponds to an ordinary abelian variety, then  $x \in \mathcal{F}(X'_{\overline{\mathbb{F}}_p}) \cup \mathcal{V}(X'_{\overline{\mathbb{F}}_p})$  by Proposition A.4.4, hence it lies in more than one irreducible component of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$ . In particular  $x$  is not a smooth point of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  contradicting Proposition A.3.5, therefore  $x$  is supersingular.

To show  $\mathcal{V}(X'_{\overline{\mathbb{F}}_p}) \cap \mathcal{F}(X'_{\overline{\mathbb{F}}_p})$  is 0-dimensional, we note that by the closure relations of the Kottwitz-Rapoport stratification on the local model, the intersection is contained in the minimal stratum which is 0-dimensional. Part (1) follows.

(2) Let  $R$  be an irreducible component of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}$ . Then since  $\pi_1|_{X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}}$  is surjective with one dimensional fibers, there exists a unique irreducible component  $C$  of  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$  such that  $\pi_1(R) = C$ . Now Proposition A.5.2 implies the fibers of  $\pi_1|_{X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}}$  are irreducible over a dense subset of  $C$ , hence  $R$  is the unique irreducible component of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}$  mapping to  $C$ .

To show  $R$  is a  $\mathbb{P}^1$ -bundle over  $C$ , we first show the morphism  $\pi_1|_R$  is smooth. Since  $R$  and  $C$  are both smooth and  $\overline{\mathbb{F}}_p$  is algebraically closed, it suffices to show the map is surjective on tangent spaces. This follows in the same way as [Sta97, Proposition 6.7].

For  $x$  a smooth point of  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$ , we have a map  $\mathbb{P}^1 \rightarrow \pi_1^{-1}(x)$  which is bijective on  $\overline{\mathbb{F}}_p$ -points by Proposition A.5.2. Since  $\pi_1^{-1}(x)$  is smooth, it follows that  $\mathbb{P}^1 \cong \pi_1^{-1}(x)$ . Thus it suffices to show that for  $x$  a non-smooth point of  $C \subset X^{\text{ss}}$ , we have  $\pi_1^{-1}(x) \cap R \cong \mathbb{P}^1$ . In this case there is a map  $T \rightarrow \pi^{-1}(x)$  which is bijective on  $\overline{\mathbb{F}}_p$ -points by Proposition A.5.2, where  $T$  is the transversal intersection of two copies of  $\mathbb{P}^1$ . Since each component of  $\pi^{-1}(x)$  is smooth, the same argument as above shows that  $T \cong \pi^{-1}(x)$ . Since  $\pi_1^{-1}(x) \cap R$  is smooth and one-dimensional, it follows that  $\pi_1^{-1}(x) \cap R \cong \mathbb{P}^1$ .

Now let  $R$  and  $R'$  be irreducible components of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}$  which map to irreducible components  $C$  and  $C'$  respectively in  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$  under  $\pi_1$ . If  $C \cap C' \neq \emptyset$ , then they intersect at a single point  $x \in X_{\overline{\mathbb{F}}_p}^{\text{ss}}(\overline{\mathbb{F}}_p)$ , and these are the only irreducible components of  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$  containing  $x$ . Then  $R \cap R' \subset \pi^{-1}(x)$ . If  $R \cap R'$  is one-dimensional, then there is another component  $R''$  of  $X'_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}$  which intersects  $\pi^{-1}(x)$ . Thus its image  $C''$  in  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$  also contains  $x$  which is a contradiction. It follows that  $R \cap R'$  is a point.  $\square$

Transferring to the quaternionic side, we obtain the analogous results for quaternionic Shimura varieties.

**Corollary A.5.4.** (1)  $X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  can be decomposed as

$$\mathcal{V}(X_{\overline{\mathbb{F}}_p}) \cup \mathcal{F}(X_{\overline{\mathbb{F}}_p}) \cup X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}^{\text{ss}}$$

where each term is a union of irreducible components of  $X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  of dimension 2. Any irreducible component of  $X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  is contained in exactly one of the terms and  $\mathcal{V}(X_{\overline{\mathbb{F}}_p}) \cap \mathcal{F}(X_{\overline{\mathbb{F}}_p})$  is 0-dimensional.

(2) For each irreducible component  $C$  of  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$ , there is a unique irreducible component  $R$  of  $X_0(\mathfrak{p})_{\overline{\mathbb{F}}_p}$  such that  $\pi_1(R) = C$ . Moreover the projection  $\pi_1 : R \rightarrow C$  exhibits  $R$  as a  $\mathbb{P}^1$ -bundle over  $C \cong \mathbb{P}^1$  and for any other irreducible component  $C'$  of  $X_{\overline{\mathbb{F}}_p}^{\text{ss}}$ , we have:

$$R \cap R' = \begin{cases} \{pt\} & \text{if } C \cap C' \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}.$$

Similarly, the result holds if we replace  $\pi_1$  by  $\pi_2$ .  $\square$

## REFERENCES

- [AG04] F. Andreatta and E. Goren, *Hilbert modular varieties of low dimension*, ch. Geometric aspects of Dwork theory, pp. 113–175, De Gruyter, 2004.
- [Blo86] S. Bloch, *Algebraic cycles and higher K-theory*, Adv. Math **61** (1986), 267–304.
- [Car86] H. Carayol, *Sur la mauvaise réduction des courbes de Shimura*, Compos. Math. **59** (1986), 151–230.

- [Cha95] C. Chai, *Every ordinary symplectic isogeny class in positive characteristic is dense in the moduli*, Invent. Math. **121** (1995), no. 1, 439–479.
- [Cha08] C-L. Chai, *A rigidity result for  $p$ -divisible formal groups*, Asian J. Math. **12** (2008), no. 2, 193–202.
- [CO19] C. Chai and F. Oort, *The Hecke orbit conjecture: a survey and outlook*, ch. Open Problems in Arithmetic Algebraic Geometry, International Press, 2019.
- [DI87] P. Deligne and L. Illusie, *Relèvements modulo  $p^2$  et décomposition du complexe de de Rham*, Invent. Math. **89** (1987), 247–270.
- [Dim05] M. Dimitrov, *Galois representations modulo  $p$  and cohomology and cohomology of Hilbert modular varieties*, Ann. Sci. École Norm. Sup. **38** (2005), no. 4, 505–551.
- [Dim09] ———, *On Ihara’s lemma for Hilbert modular varieties*, Compos. Math. **145** (2009), no. 5, 1114–1146.
- [dJ95] A. J. de Jong, *Crystalline Dieudonné module theory via formal and rigid geometry*, Publ. Math. Inst. Hautes Études Sci. **82** (1995), 5–96.
- [DT94] F. Diamond and R. Taylor, *Non-optimal levels of mod  $l$  modular representations*, Invent. Math. **115** (1994), no. 3, 435–462.
- [Fal89] G. Faltings, *Crystalline cohomology and  $p$ -adic Galois representations*, Algebraic Analysis, Geometry and Number Theory, Proc. JAMI Inaugural Conference, Johns-Hopkins University Press, 1989, pp. 25–79.
- [Gei05] T. Geisser, *Motivic cohomology,  $K$ -theory and topological cyclic homology*, Handbook of K-theory (Berlin), vol. 1, Springer, 2005, pp. 193–234.
- [Gör01] U. Görtz, *On the flatness of models of certain Shimura varieties of PEL-type*, Math. Ann. **321** (2001), no. 3, 689–727.
- [Ham15] P. Hamacher, *The geometry of Newton strata in the reduction modulo  $p$  of Shimura varieties of PEL type*, Duke Math. J. **164** (2015), no. 15, 2809–2895.
- [HN17] X. He and S. Nie, *On the  $\mu$ -ordinary locus of a Shimura variety*, Adv. Math. **321** (2017), 513–528.
- [HR17] X. He and M. Rapoport, *Stratifications in the reduction of Shimura varieties*, Manuscripta Math. **152** (2017), 317–343.
- [HZ20] X. He and R. Zhou, *On the connected components of affine Deligne-Lusztig varieties*, Duke Math. J. **169** (2020), no. 14, 2697–2765.
- [Jan88] U. Jannsen, *Continuous étale cohomology*, Math. Ann. **280** (1988), no. 2, 207–246.
- [Kis10] M. Kisin, *Integral models for Shimura varieties of abelian type*, J. Amer. Math. Soc. **23** (2010), no. 4, 967–1012.
- [Kis17] ———, *Mod  $p$  points on Shimura varieties of abelian type*, J. Amer. Math. Soc. **30** (2017), no. 3, 819–914.
- [Kot97] R. Kottwitz, *Isocrystals with additional structure. II*, Compositio Math. **109** (1997), no. 3, 255–339.
- [KP18] M. Kisin and G. Pappas, *Integral models of Shimura varieties with parahoric level structure*, Publ. Math. Inst. Hautes Études Sci. **128** (2018), no. 1, 121–218.
- [Lan13] Kai-Wen Lan, *Arithmetic compactifications of PEL-type Shimura varieties*, London Mathematical Society Monographs, vol. 36, Princeton University Press, 2013.
- [Liu19] Y. Liu, *Bounding cubic-triple product Selmer groups of elliptic curves*, J. Eur. Math. Soc. **21** (2019), no. 5, 1411–1508.
- [LS18] K. W. Lan and B. Stroh, *Nearby cycles of automorphic étale sheaves II*, Cohomology of Arithmetic Groups: On the Occasion of Joachim Schwermer’s 66th birthday, Springer Proceedings in Mathematics and Statistics, vol. 245, Springer International Publishing, 2018, Preprint, pp. 83–106.
- [LT20] Y. Liu and Y. Tian, *Supersingular locus of Hilbert modular varieties, arithmetic level raising, and Selmer groups*, Algebra Number Theory **14** (2020), no. 8, 2059–2119.
- [Mil92] J. Milne, *The points on a Shimura variety modulo a prime of good reduction*, ch. The zeta functions of Picard modular surfaces, pp. 151–253, 1992.
- [Moo04] B. Moonen, *Serre-Tate theory for moduli spaces of PEL type*, Ann. Sci. École Norm. Sup. (4) **37** (2004), no. 2, 223–269. MR 2061781
- [MS21] J. Manning and J. Shotton, *Ihara’s Lemma for Shimura curves over totally real fields via patching*, Math. Ann. **379** (2021), 187–234.
- [Nek18] J. Nekovář, *Eichler-Shimura relations and semisimplicity of étale cohomology of quaternionic Shimura varieties*, Ann. Sci. École Norm. Sup. **51** (2018), 1179–1252.
- [Rib84] K. Ribet, *Congruence relations between modular forms*, Proceedings of the International Congress of Mathematicians (Warsaw), vol. 1,2, PWN, 1984, pp. 503–514.
- [Rib88] ———, *Séminaire de théorie des nombres*, ch. Raising the levels of modular representations, pp. 259–271, Birkhäuser Boston, 1987–88.
- [RR96] M. Rapoport and M. Richartz, *On the classification and specialization of  $F$ -isocrystals with additional structure*, Compositio Math. **103** (1996), no. 2, 153–181.
- [RZ96] M. Rapoport and Th. Zink, *Period spaces for  $p$ -divisible groups*, Annals of Mathematics Studies, vol. 141, Princeton University Press, Princeton, NJ, 1996.

- [San81] J.-J. Sansuc, *Groupe de Brauer et arithmétique des groupes algébriques linéaires sur un corps de nombres*, J. reine angew. Math. **327** (1981), 12–80.
- [Sch00] A. Scholl, *Integral elements in K-theory and products of modular curves*, The Arithmetic and Geometry of Algebraic Cycles, NATO Sciences Series, Series C: Mathematical and Physical Sciences, vol. 548, Kluwer, 2000, pp. 467–489.
- [Sta97] H. Stamm, *On the reduction of the Hilbert-Blumenthal moduli scheme with  $\Gamma_0(p)$ -level structure*, Forum mathematicum **9** (1997), no. 4, 405–456.
- [SV00] A. Suslin and V. Voevodsky, *Bloch-Kato conjecture and motivic cohomology with finite coefficients*, The Arithmetic and Geometry of Algebraic Cycles, NATO Sciences Series, Series C: Mathematical and Physical Sciences, vol. 548, Kluwer, 2000, pp. 117–189.
- [SZ] X. Shen and C. Zhang, *Stratifications in good reductions of Shimura varieties of abelian type*, arxiv:1707.00439.
- [TX16] Y. Tian and L. Xiao, *On Goren-Oort stratification for quaternionic Shimura varieties*, Compos. Math. **152** (2016), 2134–2220.
- [TX19] ———, *Tate cycles on quaternionic Shimura varieties over finite fields*, Duke Math. J. **168** (2019), no. 9, 1551–1639.
- [Voe02] V. Voevodsky, *Motivic cohomology groups are isomorphic to higher Chow groups in any characteristic*, Int. Math. Res. Not. (2002), no. 351–355.
- [Voe11] ———, *Motivic cohomology with  $\mathbb{Z}/l$  coefficients*, Ann. of Math. **174** (2011), no. 1, 401–438.
- [VW13] E. Viehmann and T. Wedhorn, *Ekedahl-Oort and Newton strata for Shimura varieties of PEL type*, Math. Ann. **356** (2013), 1493–1550.
- [XZ] L. Xiao and X. Zhu, *Cycles on Shimura varieties via geometric Satake*, arxiv:1707.05700.

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS, WILBERFORCE ROAD, CAMBRIDGE, UNITED KINGDOM, CB3 0WB.

*Email address:* rz240@dpms.cam.ac.uk