Math. Res. Lett.
Volume 25, Number 3, 759–782, 2018

On multiplicities of Galois representations in cohomology groups of Shimura curves

Chuangxun Cheng and Ji Fu

In this paper, we construct examples where the multiplicities of rank two Galois representations in the localized first cohomology groups of Shimura curves over totally real fields are greater than one. These examples generalize the result of Ribet for Jacobians of Shimura curves over $\mathbb{Q}$ and are closely related to the BDJ conjecture.

1. Introduction

In [18], starting with a multiplicity one result on modular curves, Ribet constructed examples where the multiplicities of rank two Galois representations of $G_{\mathbb{Q}}$ in the Jacobians of Shimura curves over $\mathbb{Q}$ are greater than one (cf. [18, Theorem 3]). In this paper, we generalize Ribet's construction and show that similar result holds for cohomology groups of Shimura curves over general totally real fields.

To state the main result precisely, we first introduce some notation. Let F be a totally real field with $[F:\mathbb{Q}] = d$. Fix an infinite place $\tau_1 : F \hookrightarrow \mathbb{R}$. Let D be a quaternion algebra over F which is ramified at all infinite places except τ_1 and ramified at the finite primes in S_D. For simplicity, we also denote by S_D the product of finite primes in the set. Fix an odd prime p that is unramified in F and $(p, S_D) = 1$.

Fix a maximal order $\mathcal{O}_D$ of D and isomorphisms $\mathcal{O}_D \otimes_{\mathcal{O}_F} \mathcal{O}_{F_v} \cong M_2(\mathcal{O}_{F_v})$ for finite primes $v \nmid S_D$. Here $\mathcal{O}_F$ and $\mathcal{O}_{F_v}$ denote the ring of integers of F and F_v respectively. Let $K_0 \subset (D \otimes_F \mathbb{A}_F^{\infty})^{\times}$ be the open compact subgroup such that the v-component $(K_0)_v = (\mathcal{O}_D \otimes_{\mathcal{O}_F} \mathcal{O}_{F_v})^{\times}$. In particular, $(K_0)_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for finite primes $v \nmid S_D$. Define $K_0(N)$, for an ideal N prime to S_D, to be the subgroup of K_0 consisting of those u for which u_v congruent to

2010 Mathematics Subject Classification: 11F80, 11G18.
Key words and phrases: Shimura curves; Galois representations; p-adic uniformization.

$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{v^{\mathrm{ord}_v(N)}}$ for every $v \mid N$. Let $M_{K_0(N)}$ be the Shimura curve attached to D with $K_0(N)$-level structure (cf. Section 2.1). In the following, we assume that N is square-free, N is prime to pS_D, and there are at least two distinct finite primes that divide N, say $\mathfrak{pq} \mid N$.

Let $\bar{\rho} : G_F \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ be an irreducible rank two mod p Galois representation of $G_F = \mathrm{Gal}(\bar{F}/F)$. Assume that $\bar{\rho}$ appears in the space $H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{F}, \mathcal{F})$. Here $\mathcal{F}$ is an $\bar{\mathbb{F}}_p$-sheaf on the curve associated with a Serre weight (cf. Section 2.1). Let $\mathbb{T}_N$ be the Hecke algebra attached to $H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{F}, \mathcal{F})$ (cf. Section 2.2). Let $\mathfrak{m}$ be the maximal ideal of $\mathbb{T}_N$ generated by $\{p,\ T_v - \mathrm{Tr}\,\bar{\rho}(\mathrm{Frob}_v),\ \mathrm{Norm}_{F/\mathbb{Q}}(v)S_v - \det\bar{\rho}(\mathrm{Frob}_v) :\ v \nmid pNS_D\}$, where Frob_v is a lift of the arithmetic Frobenius at a finite place v. Fix an embedding $\mathbb{T}_N/\mathfrak{m} \hookrightarrow \bar{\mathbb{F}}_p$. By Eichler-Shimura relation, we have $H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{F}, \mathcal{F}) \otimes_{\mathbb{T}_N} \bar{\mathbb{F}}_p \cong \bar{\rho}^a$ ([12, Lemma 18.4] and [1]). In particular, we have

$$\dim_{\bar{\mathbb{F}}_p} \mathrm{Hom}_{G_F}(\bar{\rho}, H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{F}, \mathcal{F})_{\mathfrak{m}}) = a.$$

Let D' be another quaternion algebra over F such that D' is ramified at $\mathfrak{p}$, $\mathfrak{q}$ and at the primes where D is ramified. Fix a maximal order $\mathcal{O}_{D'}$ of D' and isomorphisms $\mathcal{O}_{D'} \otimes_{\mathcal{O}_F} \mathcal{O}_{F_v} = M_2(\mathcal{O}_{F_v})$ for finite primes $v \nmid \mathfrak{pq}S_D$. Let M'_O be the Shimura curve attached to D' with level O, where $O_v = K_0(N)_v$ if $v \nmid \mathfrak{pq}$, $O_v = (\mathcal{O}_{D'} \otimes_{\mathcal{O}_F} \mathcal{O}_{F_v})^\times$ if $v \mid \mathfrak{pq}$.

Theorem 1.1. *With the notation as above, if* $\bar{\rho} : G_F \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ *is ramified at* $\mathfrak{q}$ *and* $\dim_{\bar{\mathbb{F}}_p} \mathrm{Hom}_{G_F}(\bar{\rho}, H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{F}, \mathcal{F})_{\mathfrak{m}}) = 1$, *then*

$$\dim_{\bar{\mathbb{F}}_p} \mathrm{Hom}_{G_F}(\bar{\rho}, H^1_{\mathrm{et}}(M'_O \otimes \bar{F}, \mathcal{F})_{\mathfrak{m}}) \leq 2.$$

Moreover, assume that $\bar{\rho}$ *is not induced from a character if* $F \supset \mathbb{Q}(\mu_p)^+$. *Then the dimension is two if and only if* $\bar{\rho}$ *is unramified at* $\mathfrak{p}$ *and* $\bar{\rho}(\mathrm{Frob}_{\mathfrak{p}})$ *is* ± 1.

Remark 1.2. In order to prove the claim that the dimension is two if $\bar{\rho}$ is unramified at $\mathfrak{p}$ and $\bar{\rho}(\mathrm{Frob}_{\mathfrak{p}})$ is ± 1, we need a level-lowering result (cf. [16, Main Theorem 1] and [14, Theorem 0.1]). Hence in the statement of the theorem, we have the assumption that $\bar{\rho}$ is not induced from a character if $F \supset \mathbb{Q}(\mu_p)^+$ (cf. [16, Pages 58-59]).

Remark 1.3. If $F = \mathbb{Q}$, $S_D = \emptyset$, $V \cong \bar{\mathbb{F}}_p$ is the trivial Serre weight, Theorem 1.1 is exactly [18, Theorem 3]. Note that here one needs the compact modular curves. In Section 3.4.1, we explain that this generalization is not vacuous. Specifically, we prove a *level-raising* result Proposition 3.10, hence the construction after Theorem 3 of [18] can be generalized.

Remark 1.4. If $F=\mathbb{Q}$, $S_D=\emptyset$, $V=\mathrm{Symm}^b\,\bar{\mathbb{F}}_p^2$ with $0\le b\le p-3$, we show that under some technical conditions, we always have

$$\dim_{\bar{\mathbb{F}}_p}\mathrm{Hom}_{G_F}(\bar{\rho},H^1_{\mathrm{et}}(M_{K_0(N)}\otimes\bar{F},\mathcal{F})_{\mathfrak{m}})=1\quad\text{(cf. Section 3.4.2).}$$

2. Backgrounds on Shimura curves

2.1. The curves

Let G be the algebraic group over $\mathbb{Q}$ attached to $D^\times$. Let X be the $G(\mathbb{R})$-conjugacy class of the map

$$h:\mathbb{C}^\times\to G(\mathbb{R})\simeq \mathrm{GL}_2(\mathbb{R})\times\mathbb{H}^\times\times\cdots\times\mathbb{H}^\times,$$

which maps $a+ib$ to $\left(\begin{pmatrix}a & b\\ -b & a\end{pmatrix}^{-1},1,\dots,1\right)$. The conjugacy class X is naturally identified with the union of the upper and lower half plane by the map $g^{-1}hg\mapsto g(i)$, where $g(i)=\frac{a+ib}{c+id}$ if $g=\left(\begin{pmatrix}a & b\\ c & d\end{pmatrix},\dots\right)$.

Let $M=M(G,X)=(M_H)_H$ be the canonical model defined over F of the Shimura variety defined by G and X. Here H runs through the open compact subgroups of $G(\mathbb{A}^\infty)$. Each M_H is proper and smooth but not necessarily geometrically connected over F, and

$$M_H(\mathbb{C})\simeq G(\mathbb{Q})\backslash X\times G(\mathbb{A}^\infty)/H.$$

For each H and H' and $g\in G(\mathbb{A}^\infty)$ with $g^{-1}H'g\subset H$ and H sufficiently small, by [3, Lemma 1.4.1.1], there is an etale map $\varrho_g: M_{H'}\to M_H$ which on complex points coincides with the one induced by right multiplication by g in $G(\mathbb{A}^\infty)$. For a normal subgroup H' of H, the etale cover $\varrho_1:M_{H'}\to M_H$ is Galois, and mapping $g^{-1}\mapsto\varrho_g$ defines an isomorphism of H/H' with a group of covering maps.

Remark 2.1. The weight homomorphism $w:\mathbb{G}_{m,\mathbb{R}}\to G_{\mathbb{R}}$ attached to h is given by

$$w(r)=\left(\begin{pmatrix}r & 0\\ 0 & 1\end{pmatrix}^{-1},1,\dots,1\right)\quad r\in\mathbb{R}.$$

This is not defined over $\mathbb{Q}$ if $n > 1$. Thus we cannot expect a description of $M_H(\mathbb{C})$ as a moduli space of abelian varieties. Indeed, if $M_H(\mathbb{C})$ is a moduli space, we could describe h as given by the Hodge structure of an abelian variety, which is rational.

Following [2, Section 2], we construct some mod p sheaves on the curves. Denote by k_v the residue field of F at a finite prime v. Let V be an irreducible $\bar{\mathbb{F}}_p$-representation of $\mathrm{GL}_2(\mathcal{O}_F/p) = \prod_{v|p} \mathrm{GL}_2(k_v)$. It is called a *Serre weight.* We say a Serre weight is *regular* if each b_τ is less or equal to $p-2$ when we write the Serre weight as in [2, Section 2]. Let $H \subset G(\mathbb{A}^\infty)$ be open compact such that $H_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for all $v \mid p$. We attach to V a locally constant sheaf

$$\begin{array}{c} \mathcal{F}_V = G(\mathbb{Q})\backslash X \times G(\mathbb{A}^\infty) \times V/H \\ \Big\downarrow{\scriptstyle \pi} \\ M_H \end{array} \tag{2.1}$$

on M_H.

2.2. The Hecke algebras

Suppose that H and H' are sufficiently small open compact subgroups of $G(\mathbb{A}^\infty)$. Let $\mathcal{F}$ be a sheaf on the Shimura curves attached to a certain module (see [7, Section 6.1]). Let $g \in G(\mathbb{A}^\infty)$ such that g acts trivially on $\mathcal{F}$. There is a natural identification of sheaves on $M_{H\cap gH'g^{-1}}$: $\mathcal{F}^{gH'g^{-1}} |_{M_{H\cap gH'g^{-1}}} = \mathcal{F}^{H\cap gH'g^{-1}}$. Here $\mathcal{F}^R$ means that we consider $\mathcal{F}$ as a sheaf on the curve M_R. Then define

$$\begin{aligned}[HgH'] : H^j(M_{H'} \otimes \bar{F}, \mathcal{F}^{H'}) &\to H^j(M_{H'\cap g^{-1}Hg} \otimes \bar{F}, \mathcal{F}^{H'} |_{M_{H'\cap g^{-1}Hg}}) \\ &\to H^j(M_{gH'g^{-1}\cap H} \otimes \bar{F}, \mathcal{F}^{gH'g^{-1}} |_{M_{gH'g^{-1}\cap H}}) \\ &= H^j(M_{gH'g^{-1}\cap H} \otimes \bar{F}, \mathcal{F}^{H\cap gH'g^{-1}}) \\ &\to H^j(M_H \otimes \bar{F}, \mathcal{F}^H),\end{aligned}$$

where the first arrow is the restriction map, the second arrow is induced from $\varrho_g : M_{H'\cap g^{-1}Hg} \to M_{gH'g^{-1}\cap H}$, and the last arrow is the trace map. See [12, Section 15] for more details. In the following, we consider the case where $H_v = H'_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$ for all $v \mid p$ and $\mathcal{F}$ is associated with a Serre weight. Hence if $g \in G(\mathbb{A}^\infty)$ with $g_v = 1$ for all $v \mid p$, the operator $[HgH']$ exists.

Let $H = H'$. If $\mathfrak{q}$ is a prime of $\mathcal{O}_F$ which does not divide pS_D, let $\omega_\mathfrak{q} \in \mathbb{A}_F^\infty$ be such that $\omega_\mathfrak{q}$ is a uniformizer at $\mathfrak{q}$ and is 1 at every other place. Then write

$$T_\mathfrak{q} = \left[H \begin{pmatrix} 1 & 0 \\ 0 & \omega_\mathfrak{q} \end{pmatrix} H \right].$$

If also $H_\mathfrak{q} = \mathrm{GL}_2(\mathcal{O}_\mathfrak{q})$, define

$$S_\mathfrak{q} = \left[H \begin{pmatrix} \omega_\mathfrak{q} & 0 \\ 0 & \omega_\mathfrak{q} \end{pmatrix} H \right].$$

Denote by $\mathbb{T}(H, \mathcal{F})$ the sub-algebra of $\mathrm{End}(H^1_{\mathrm{et}}(M_H \otimes_F \bar{F}, \mathcal{F}))$ generated by $T_\mathfrak{q}$ and $S_\mathfrak{q}$ for those $\mathfrak{q}$ with $\mathfrak{q} \nmid pS_D$ and $H_\mathfrak{q} = \mathrm{GL}_2(\mathcal{O}_\mathfrak{q})$. If $H = K_0(N)$, write $U_{\omega_\mathfrak{q}} = T_\mathfrak{q}$ if $\mathfrak{q} \mid N$.

Definition 2.2. (Cf. [13, Definition 4.1]) A maximal ideal of $\mathbb{T}(H, \mathcal{F})$ is *Eisenstein* if it contains $T_v - 2$ and $S_v - 1$ for all but finitely many primes v of F which split completely in some finite abelian extension of F.

Let $U \subset K_0$ be a sufficiently small open compact subgroup such that $U_v = (\mathcal{O}_D \otimes \mathcal{O}_{F_v})^\times$ for almost all v and for all $v \mid pS_D$. Let $\mathfrak{q}$ be a finite prime of F such that $\mathfrak{q} \nmid pS_D$ and $U_\mathfrak{q} = \mathrm{GL}_2(\mathcal{O}_\mathfrak{q})$. Let $U_0(\mathfrak{q})$ be the subgroup of U defined by $U_0(\mathfrak{q})_v = U_v$ if $v \neq \mathfrak{q}$ and

$$U_0(\mathfrak{q})_\mathfrak{q} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in U_\mathfrak{q} \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{\mathfrak{q}} \right\}.$$

Let $\eta_\mathfrak{q} \in (D \otimes_F \mathbb{A}_F^\infty)^\times$ be the element $\begin{pmatrix} 1 & 0 \\ 0 & \omega_\mathfrak{q} \end{pmatrix}$. Then we have $\eta_\mathfrak{q}^{-1} U_0(\mathfrak{q}) \eta_\mathfrak{q} \subset U$ and two maps $\varrho_1, \varrho_{\eta_\mathfrak{q}} : M_{U_0(\mathfrak{q})} \to M_U$. We have the following conjecture, which is usually referred as Ihara's Lemma.

Conjecture 2.3. *Assume that the genus of M_U is greater than 1, V is a regular Serre weight. Then the kernel of the following map*

$$A : H^1_{\mathrm{et}}(M_U \otimes \bar{F}, \mathcal{F}_V) \oplus H^1_{\mathrm{et}}(M_U \otimes \bar{F}, \mathcal{F}_V) \to H^1_{\mathrm{et}}(M_{U_0(\mathfrak{q})} \otimes \bar{F}, \mathcal{F}_V)$$
$$(f_1, f_2) \mapsto (\varrho_1)^* f_1 + (\varrho_{\eta_\mathfrak{q}})^* f_2.$$

is Eisenstein, i.e., $(\mathrm{Ker}(A))_\mathfrak{m}$ *is trivial if* $\mathfrak{m}$ *is a non-Eisenstein maximal ideal of the Hecke algebra* $\mathbb{T}(U, \mathcal{F}_V)$.

The following result is [10, Theorem 4]. See also [10, Theorem 2] and [20, Lemma 3.1] for similar results where D is definite.

Theorem 2.4. *If $F = \mathbb{Q}$ and $V = \mathrm{Symm}^b \bar{\mathbb{F}}_p^2$ with $b \leq p-3$, then Conjecture 2.3 holds.*

2.3. Supersingular and ordinary points on the special fibre of M_K

Before we prove the main theorem, we review some properties of integral models of Shimura curves.

Let $\wp$ be a finite prime of F, K be an open compact subgroup of $G(\mathbb{A}^\infty)$. We suppose that K factors as $K_\wp H$. Then we have the following results. The first one is proved in [3, Sections 6, 9]. The second one is indicated in [3] and a detail proof is given in [12, Section 10].

Theorem 2.5. 1) *Suppose that $K_\wp$ is the subgroup $K_\wp^0 = \mathrm{GL}_2(\mathcal{O}_{F_\wp})$. If H is sufficiently small, there exists a model $\mathbf{M}_{0,H}$ of M_K defined over $\mathcal{O}_{F,(\wp)}$. This model is proper and smooth.*

2) *Suppose that $K_\wp$ is the subgroup $K_\wp^n$ of matrices congruent to $\mathrm{I}_{2\times 2}$ modular $\wp^n$. If H is sufficiently small, there exists a regular model $\mathbf{M}_{n,H}$ of M_K with a map to $\mathbf{M}_{0,H}$. The morphism $\mathbf{M}_{n,H} \to \mathbf{M}_{0,H}$ is finite and flat.*

Theorem 2.6. 1) *Suppose that $K_\wp$ is the group*

$$\Gamma_0(\wp) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_{F_\wp}) \mid c \in \wp \right\}.$$

If H is sufficiently small, there exists a regular model $\mathbf{M}_{\Gamma_0(\wp)H}$ of M_K defined over $\mathbf{M}_{0,H}$. The morphism $\mathbf{M}_{\Gamma_0(\wp)H} \to \mathbf{M}_{0,H}$ is finite and flat.

2) *The special fibre $\mathbf{M}_{\Gamma_0(\wp)H} \otimes \bar{k}_\wp$ is isomorphic to a union of two copies of $\mathbf{M}_{0,H} \otimes \bar{k}_\wp$ intersecting transversally above a finite set of points Σ_H.*

The points in Σ_H are the supersingular points of $\mathbf{M}_{0,H} \otimes \bar{k}_\wp$. We can describe this set in another way (cf. [12, Theorem 10.2], [13, Section 2], [3, Section 11]). Let $\bar{D}$ be another quaternion algebra over F ramified at primes in $S_D \cup \{\wp\} \cup \{\tau \mid \tau|\infty\}$. So it is totally definite. Let $\bar{G} = \mathrm{Res}_{F/\mathbb{Q}} \bar{D}^\times$ be the corresponding algebraic group over $\mathbb{Q}$, then there is a bijection

$$\begin{aligned} \Sigma_H &\cong \bar{G}(\mathbb{Q}) \backslash \bar{G}(\mathbb{A}^\infty)/H \times \mathcal{O}_{\bar{D}_\wp}^\times \\ &\cong \bar{G}(\mathbb{Q}) \backslash \bar{G}(\mathbb{A}^\infty)^\wp \times F_\wp^\times / H \times \mathcal{O}_{F_\wp}^\times, \end{aligned} \tag{2.2}$$

where the second bijection is induced by the reduced norm $\bar{D}_\wp^\times \to F_\wp^\times$. Since $\bar{D}$ is totally definite, Σ_H is a finite set.

According to [3, Section 11.1.1], the action of $\mathrm{GL}_2(F_\wp)$ on the inverse system of Shimura curves descends to an action on the set Σ_H of supersingular points, and the action factors through $\det : \mathrm{GL}_2(F_\wp) \to F_\wp^\times$. Further, if we normalize the reciprocity map of class field theory so that the arithmetic Frobenius elements correspond to uniformizers, then an element $\sigma \in W(F_\wp^{\mathrm{ab}}/F_\wp)$ acts on the set Σ_H in exactly the same way as the element $[\sigma] \in F_\wp^\times$ corresponding to σ by class field theory (cf. [3, Section 11.2(2)]).

Let $K = \mathrm{GL}_2(\mathcal{O}_\wp)H$ and $\sharp(\Sigma_H)$ be the number of supersingular points on $\mathbf{M}_{0,H} \times \bar{k}_\wp$. Let g_R be the genus of any geometric fibre of $\mathbf{M}_R$. Then we have the following lemma.

Lemma 2.7. $\sharp(\Sigma_H) = (\mathrm{Norm}(\wp) - 1)(g_{\mathrm{GL}_2(\mathcal{O}_\wp)H} - 1)$.

Proof. The argument is similar to the proof of [10, Lemma 6]. We compute $g_{\Gamma_0(\wp)H}$ in two ways. On one hand, in characteristic 0, we have a map $M_{\Gamma_0(\wp)H} \to M_{\mathrm{GL}_2(\mathcal{O}_\wp)H}$ which is of degree $\mathrm{Norm}(\wp) + 1$. Therefore, $g_{\Gamma_0(\wp)H} - 1 = (\mathrm{Norm}(\wp) + 1)(g_{\mathrm{GL}_2(\mathcal{O}_\wp)H} - 1)$. On the other hand, modulo $\wp$, $\mathbf{M}_{\Gamma_0(\wp)H}$ is singular. One can still define its arithmetic genus and it is equal to $g_{\Gamma_0(\wp)H}$ since $\mathbf{M}_{\Gamma_0(\wp)H}$ is flat. Applying Riemann-Hurwitz formula to the map $(\mathbf{M}_{\Gamma_0(\wp)H})_{/\wp} \to (\mathbf{M}_{\mathrm{GL}_2(\mathcal{O}_\wp)H})_{/\wp}$, we obtain the following equation

$$2g_{\Gamma_0(\wp)H} - 2 = 2(2g_{\mathrm{GL}_2(\mathcal{O}_\wp)H} - 2) + 2\sharp(\Sigma_H).$$

Comparing the two equations, the lemma follows. □

By [12, Theorem 10.2], we have a normalization map $r : \mathbf{M}_{0,H} \otimes \bar{k}_\wp \sqcup \mathbf{M}_{0,H} \otimes \bar{k}_\wp \to \mathbf{M}_{\Gamma_0(\wp)H} \otimes \bar{k}_\wp$. We describe the map r using the moduli description of the curves (cf. [12, Sections 8-10]). Denote by σ the arithmetic Frobenius over $k_\wp$. For any scheme Z over $\bar{k}_\wp$, define $Z^{(\sigma)} := Z \otimes_{\bar{k}_\wp,\sigma} \bar{k}_\wp$. For any scheme S over $k_\wp$, denote by $\mathrm{Fr} : S \to S$ the absolute Frobenius morphism, which on the affine rings is given by the map $s \mapsto s^{|k_\wp|}$.

Let x be a geometric point of $\mathbf{M}_{0,H} \otimes k_\wp$. Attached to x, there is a p-divisible group $\mathbf{E}_\infty|_x$. By [12, Proposition 9.7], up to isomorphism, there exist exactly two $\wp$-cyclic isogenies (cf. [12, Definition 9.2]) from $\mathbf{E}_\infty|_x$ given by Frobenius and Verschiebung respectively:

$$F : \mathbf{E}_\infty|_x \to (\mathbf{E}_\infty|_x)^{(\sigma)}, \quad V : (\mathbf{E}_\infty|_y)^{(\sigma)} \to \mathbf{E}_\infty|_y.$$

Here $\mathrm{Fr}\, y = x$ and we use the identity $(\mathbf{E}_\infty|_x)^{(\sigma)} = \mathbf{E}_\infty|_{\mathrm{Fr}\, x}$ from [12, Corollary 9.6]. By [12, Theorem 10.2], we obtain two maps ϕ_1 and $\phi_2 : \mathbf{M}_{0,H} \otimes$

$k_\wp \to \mathbf{M}_{\Gamma_0(\wp)H} \otimes k_\wp$, given by $x \mapsto (x, F : \mathbf{E}_\infty|_x \to (\mathbf{E}_\infty|_x)^{(\sigma)})$ and $x \mapsto (\mathrm{Fr}\, x, V : \mathbf{E}_\infty|_{\mathrm{Fr}\, x} \to \mathbf{E}_\infty|_x)$ respectively. These two maps induce the normalization map r.

Define $w : \mathbf{M}_{\Gamma_0(\wp)H} \otimes k_\wp \to \mathbf{M}_{\Gamma_0(\wp)H} \otimes k_\wp$ by

$$(x, F : \mathbf{E}_\infty|_x \to (\mathbf{E}_\infty|_x)^{(\sigma)}) \mapsto (\mathrm{Fr}\, x, V : \mathbf{E}_\infty|_{\mathrm{Fr}\, x} \to \mathbf{E}_\infty|_x).$$

Then $\phi_2 = w \circ \phi_1$.

Lemma 2.8. *The morphism w is an involution on $\mathbf{M}_{\Gamma_0(\wp)H} \otimes k_\wp$.*

Proof. To prove the lemma, we need to understand $\mathrm{Fr}\, x$. In order to do so, we go back to the unitary Shimura curves $\mathbf{M}'_{K'}$ as in [3, Section 7.5]. More precisely, with the same notation as in [3], let $(A, \iota, \theta, k^\wp, \varphi)$ be the data defining point x and let $(A', \iota', \theta', (k')^\wp, \varphi')$ be the data defining $\mathrm{Fr}\, x$. Then $A' = A/\operatorname{Ker}(F)$ as the p-divisible group is just $(A_{p^\infty})_1^{2,1}$ (cf. [3, Section 5.4]). One then writes down other data explicitly and obtain $w^2 = 1$. □

The two maps ϱ_1 and $\varrho_{\eta_\wp} : M_{\Gamma_0(\wp)H} \to M_{\mathrm{GL}_2(\mathcal{O}_\wp)H}$ induce two maps $\mathbf{M}_{\Gamma_0(\wp)H} \otimes k_\wp \to \mathbf{M}_{0,H} \otimes k_\wp$, which we still denote by ϱ_1 and $\varrho_{\eta_\wp}$. We have the following result (cf. [8, Chap. 5, Section 1.15]).

Lemma 2.9. *With the notation as above, $\varrho_1 \circ \phi_1 = \mathrm{id}$, $\varrho_1 \circ \phi_2 = \mathrm{Fr}$, $\varrho_{\eta_\wp} \circ \phi_1 = \mathrm{Fr}$, $\varrho_{\eta_\wp} \circ \phi_2 = \mathrm{id}$.*

Proof. The first two identities are obvious. For the last two identities, it suffices to show that $\varrho_{\eta_\wp} = \varrho_1 \circ w$. This follows from [3, Section 10.3]. Note that in [3, Section 10.3], Frobenius means the geometric Frobenius. □

2.4. p-adic uniformization of Shimura curves

Let v be a finite place of F at which D' is ramified. Here D' is the quaternion algebra defined in Section 1. Let P be an open compact subgroup of $(D' \otimes_F \mathbb{A}_F)^\times$ such that $P_v = (\mathcal{O}_{D'} \otimes_{\mathcal{O}_F} \mathcal{O}_{F_v})^\times$. Then we have a Shimura curve M'_P which is defined over F.

Let F_v^{ur} be the maximal unramified extension of F_v, let Ω_{F_v} be the Drinfeld's upper half plane over F_v and $\Omega = \Omega_{F_v} \otimes_{F_v} F_v^{\mathrm{ur}}$. Let $g \in \mathrm{GL}_2(F_v)$ act on Ω via the natural (left) action on Ω_{F_v} and the action of $\mathrm{Frob}_v^{\mathrm{val}(\det g)}$ on F_v^{ur}. Let $n \in \mathbb{Z}$ act on Ω through the action of Frob_v^{-n} on F_v^{ur}. This gives an F_v-rational action of $\mathrm{GL}_2(F_v) \times \mathbb{Z}$ on Ω. Moreover, the F_v-analytic space $\mathrm{GL}_2(F_v)\backslash(\Omega \times (P^v\backslash(D' \otimes_F \mathbb{A}_F^v)^\times/G'(\mathbb{Q})))$ algebraizes canonically to a scheme $\mathfrak{X}_S$ over F_v. Let M' and $\mathfrak{X}$ be the inverse limits of M'_P and $\mathfrak{X}_P$ over all P. Then a special case of [23, Theorem 5.3] gives the following theorem, which generalizes the result of Cherednik and Drinfeld [11, Section 4]. See also [15, Section 1].

Theorem 2.10. *There exists a $(D' \otimes_F \mathbb{A}_F^v)^\times \times \mathbb{Z}$-equivariant, F_v-rational isomorphism*

$$M' \otimes_F F_v \cong \mathfrak{X}.$$

In particular, we have an F_v-rational isomorphism

$$M'_P \otimes_F F_v \cong \mathfrak{X}_P.$$

Furthermore, there exists an integral model $\mathbf{M}'_P$ for M'_P over $\mathcal{O}_{F_v}$, and the above isomorphism can be extended to schemes over $\mathcal{O}_{F_v}$.

We apply the above theorem to the case $P = O$ and $v = \mathfrak{q}$ to study the singular points of the mod $\mathfrak{q}$ reduction of $\mathbf{M}'_O$ (cf. [17, Section 4], [16, Section 3.2]).

The special fibre of $\mathbf{M}'_O$ has non-degenerate quadratic singular points. The dual graph $\mathfrak{G}$ attached to the special fibre of $\mathbf{M}'_O \otimes k_\mathfrak{q}$ is the quotient

$$\mathrm{GL}_2(F_\mathfrak{q})^+ \backslash (\Delta \times (O^\mathfrak{q} \backslash G'(\mathbb{A}^\infty)^\mathfrak{q} / G'(\mathbb{Q})),$$

where Δ is the well-known tree attached to $\mathrm{SL}_2(F_\mathfrak{q})$, $\mathrm{GL}_2(F_\mathfrak{q})^+$ is the kernel of the map

$$\mu : \mathrm{GL}_2(F_\mathfrak{q}) \to \mathbb{Z}/2\mathbb{Z}$$

defined by the formula

$$\mu(\gamma) = \mathrm{ord}_\mathfrak{q}(\det \gamma) \pmod 2.$$

The singular points of $\mathbf{M}'_O \otimes k_\mathfrak{q}$ correspond to the edges of $\mathfrak{G}$. Let $\bar{D}'$ be the definite quaternion algebra over F ramified at places in $S_D \cup \{\mathfrak{p}\} \cup \{v \mid v|\infty\}$. Note that D' is ramified at $\mathfrak{q}$ and unramified at τ_1, and $\bar{D}'$ is unramified at $\mathfrak{q}$ and ramified at τ_1. Let $\bar{G}' = \mathrm{Res}_\mathbb{Q}^F(\bar{D}')^\times$ be the associated algebraic group. Let S be an open compact subgroup of $\bar{G}'(\mathbb{A}^\infty)$ such that $S_v = O_v$ if $v \neq \mathfrak{q}$, and $S_\mathfrak{q} = \Gamma_0(\mathfrak{q})$ where

$$\Gamma_0(\mathfrak{q}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathcal{O}_{F_\mathfrak{q}}) \mid c \in \mathfrak{q} \right\}.$$

Then the edges of Δ are in one-to-one correspondence with $\mathrm{GL}_2(F_\mathfrak{q})^+/(S_\mathfrak{q} F^\times)$. Therefore, the edges of $\mathfrak{G}$ are bijective to $\bar{G}'(\mathbb{Q}) \backslash \bar{G}'(\mathbb{A})/S$.

Remark 2.11. Consider the curves $\mathbf{M}_{K_0(N)}$ over $\mathcal{O}_{\mathfrak{p}}$ and $\mathbf{M}'_O$ over $\mathcal{O}_{\mathfrak{q}}$ and their special fibres. By the definitions of the quaternion algebras D and D', we see that there is a bijection between the set of singular points of $\mathbf{M}_{K_0(N)}$ (mod $\mathfrak{p}$) and the set of singular points of $\mathbf{M}'_O$ (mod $\mathfrak{q}$).

3. Proof of the main result

In this section, we study certain exact sequences following Carayol [4] and Jarvis [12] and prove Theorem 1.1. The sequences are also used in [16] to prove a level-lowering result. Note that when we take a reduction of a Shimura curve at a prime, we use the integral model of that curve as given in Sections 2.3 and 2.4.

3.1. Exact sequences for $M_{K_0(N)}$

First, we consider the Shimura curve $M_{K_0(N)}$ and the sheaf $\mathcal{F} = \mathcal{F}_V$. Let $H = K_0(N)^{\mathfrak{p}}$, then $K_0(N) = \Gamma_0(\mathfrak{p})H$. The exact sequence of vanishing cycles for the proper morphism $\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \mathcal{O}_{F_{\mathfrak{p}}} \to \operatorname{Spec} \mathcal{O}_{F_{\mathfrak{p}}}$ and the sheaf $\mathcal{F}$ is

$$(3.1)\quad 0 \to H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_{\mathfrak{p}}, \mathcal{F}) \to H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{F}_{\mathfrak{p}}, \mathcal{F}) \to \bigoplus_{x \in \Sigma_H} (R^1\Phi_{\bar{\eta}}\mathcal{F})_x \to H^2_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_{\mathfrak{p}}, \mathcal{F}) \to H^2_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{F}_{\mathfrak{p}}, \mathcal{F}) \to \cdots,$$

where Σ_H denotes the set of singular points of the $\bar{k}_{\mathfrak{p}}$-scheme $\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_{\mathfrak{p}}$. Note that Σ_H consists of a finite number of non-degenerate quadratic points.

Write

- $L(H) = \operatorname{Ker}(H^2_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_{\mathfrak{p}}, \mathcal{F}) \to H^2_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{F}_{\mathfrak{p}}, \mathcal{F}))$,
- $Z(H) = H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_{\mathfrak{p}}, \mathcal{F})$,
- $M(H) = H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{F}_{\mathfrak{p}}, \mathcal{F})$,
- $X(H) = \bigoplus_{x \in \Sigma_H} (R^1\Phi_{\bar{\eta}}\mathcal{F})_x$,
- $\tilde{X}(H) = \operatorname{Ker}(X(H) \to L(H))$.

Then we have a short exact sequence

$$(3.2)\qquad 0 \to Z(H) \to M(H) \to \tilde{X}(H) \to 0.$$

We construct a second exact sequence, based on the comparison between the cohomology of the special fibre and the cohomology of the normalization of the special fibre. Recall we have a map $r : \mathbf{M}_{0,H} \otimes \bar{k}_\mathfrak{p} \sqcup \mathbf{M}_{0,H} \otimes \bar{k}_\mathfrak{p} \to \mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}$, and $\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}$ can be regarded as two copies of $\mathbf{M}_{0,H} \otimes \bar{k}_\mathfrak{p}$ glued together transversally above each supersingular point of $\mathbf{M}_{0,H} \otimes \bar{k}_\mathfrak{p}$. As r is an isomorphism away from supersingular points, there is an exact sequence of sheaves

$$0 \to \mathcal{F} \to r_* r^* \mathcal{F} \to \mathcal{G} \to 0,$$

where $\mathcal{G}$ is a skyscraper sheaf supported on Σ_H. Taking the long exact sequence, we have

$$\begin{aligned} (3.3) \qquad 0 &\to H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{F}) \to H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, r_* r^* \mathcal{F}) \\ &\to H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{G}) \to H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{F}) \\ &\to H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, r_* r^* \mathcal{F}) \to 0. \end{aligned}$$

Write

- $\tilde{L}(H) = \mathrm{Im}(H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, r_* r^* \mathcal{F}) \to H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{G}))$,
- $Y(H) = H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{G}) = \bigoplus_{x \in \Sigma_H} \mathcal{G}_x$,
- $R(H) = H^1_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, r_* r^* \mathcal{F})$,
- $\tilde{Y}(H) = Y(H)/\tilde{L}(H)$.

Then we have another short exact sequence

$$(3.4) \qquad 0 \to \tilde{Y}(H) \to Z(H) \to R(H) \to 0.$$

The following lemma collects several properties of the exact sequences (3.2) and (3.4).

Lemma 3.1. 1) *The sequences* (3.2) *and* (3.4) *are equivariant for the Hecke action and the Galois action.*

2) $H^i_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, r_* r^* \mathcal{F}) \cong H^i_{\text{et}}(\mathbf{M}_{0,H} \otimes \bar{k}_\mathfrak{p}, \mathcal{F})^2$.

3) *There is an isomorphism*

$$H^1_{\Sigma_H}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{F}) \tilde{\to} \bigoplus_{x \in \Sigma_H} H^1_{\{x\}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, R\Psi_{\bar{\eta}} \mathcal{F}).$$

4) *There is an isomorphism*

$$H^0_{\text{et}}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{G}) \cong H^1_{\Sigma_H}(\mathbf{M}_{\Gamma_0(\mathfrak{p})H} \otimes \bar{k}_\mathfrak{p}, \mathcal{F}).$$

5) *There is an isomorphism*

$$N : X(H)(1) \to Y(H).$$

6) $\tilde{L}(H)_{\mathfrak{m}} = 0$, $L(H)_{\mathfrak{m}} = 0$, *if* $\mathfrak{m}$ *is a non-Eisenstein maximal ideal of the Hecke algebra.*

Proof. These results are proved in [12, Sections 16-18]. □

Remark 3.2. 1) We have an isomorphism

$$H^1_{\text{et}}(\mathbf{M}_{0,H} \otimes \bar{k}_{\mathfrak{p}}, \mathcal{F}) \cong H^1_{\text{et}}(\mathbf{M}_{0,H} \otimes \bar{F}_{\mathfrak{p}}, \mathcal{F}),$$

because $\mathbf{M}_{0,H}$ has good reduction at $\mathfrak{p}$. Therefore, we have

$$R(H) \cong H^1_{\text{et}}(\mathbf{M}_{0,H} \otimes \bar{F}_{\mathfrak{p}}, \mathcal{F})^2.$$

2) Jarvis [12] proved more than those listed in the lemma. In particular, [12, Proposition 17.4] gives the following diagram which describes the action of the inertia group and is very important in our application:

$$\begin{array}{ccc} M(H) & \xrightarrow{\tau-1} & M(H) \\ {\scriptstyle\beta}\downarrow & & \uparrow{\scriptstyle\alpha} \\ X(H) & \xrightarrow[\text{Var}(\tau)]{} & Y(H). \end{array} \tag{3.5}$$

Here τ is an element in the inertia group $I_{F_{\mathfrak{p}}}$, and $\text{Var}(\tau)$ is the variation map.

By Lemma 3.1, $\tilde{X}(H)_{\mathfrak{m}} = X(H)_{\mathfrak{m}}$ and $\tilde{Y}(H)_{\mathfrak{m}} = Y(H)_{\mathfrak{m}}$. Then combining exact sequences (3.2) and (3.4), we have the following diagram

(3.6)
$$\begin{array}{ccccccccc} & & 0 & & & & & & \\ & & \downarrow & & & & & & \\ & & Y(H)_{\mathfrak{m}} & & & & & & \\ & & \downarrow & & & & & & \\ 0 & \longrightarrow & Z(H)_{\mathfrak{m}} & \longrightarrow & M(H)_{\mathfrak{m}} & \longrightarrow & X(H)_{\mathfrak{m}} & \longrightarrow & 0 \\ & & \downarrow & & & & & & \\ & & R(H)_{\mathfrak{m}} & & & & & & \\ & & \downarrow & & & & & & \\ & & 0 & & & & & & \end{array}$$

together with an isomorphism $N : (X(H)(1))_{\mathfrak{m}} \tilde{\to} Y(H)_{\mathfrak{m}}$.

Fix an embedding $\mathbb{F} = \mathbb{T}(K_0(N), \mathcal{F})_{\mathfrak{m}}/\mathfrak{m} \to \bar{\mathbb{F}}_p$. Tensoring the above diagram with $\bar{\mathbb{F}}_p$, we obtain

(3.7)
$$\begin{array}{ccccccc} Y(H) \otimes \bar{\mathbb{F}}_p & & & & & & \\ \downarrow \gamma & \overset{\alpha}{\searrow} & & & & & \\ Z(H) \otimes \bar{\mathbb{F}}_p & \xrightarrow{\delta} & M(H) \otimes \bar{\mathbb{F}}_p & \xrightarrow{\beta} & X(H) \otimes \bar{\mathbb{F}}_p & \longrightarrow & 0 \\ \downarrow & & & & & & \\ R(H) \otimes \bar{\mathbb{F}}_p & & & & & & \\ \downarrow & & & & & & \\ 0 & & & & & & \end{array}$$

together with an isomorphism $N : X(H)(1) \otimes \bar{\mathbb{F}}_p \tilde{\to} Y(H) \otimes \bar{\mathbb{F}}_p$.

Lemma 3.3. *If $\beta : M(H) \otimes \bar{\mathbb{F}}_p \to X(H) \otimes \bar{\mathbb{F}}_p$ is an isomorphism, then $\bar{\rho}$ is unramified at $\mathfrak{p}$.*

Proof. We compute the action of the inertia group $I_{F_{\mathfrak{p}}}$. An element $\tau \in I_{F_{\mathfrak{p}}}$ acts via the diagram

$$
\begin{array}{ccc}
M(H)\otimes\bar{\mathbb{F}}_p & \xrightarrow{\tau-1} & M(H)\otimes\bar{\mathbb{F}}_p \\
{\scriptstyle\beta}\big\downarrow & & \big\uparrow{\scriptstyle\alpha} \\
X(H)\otimes\bar{\mathbb{F}}_p & \xrightarrow[\mathrm{Var}(\tau)]{} & Y(H)\otimes\bar{\mathbb{F}}_p.
\end{array}
\tag{3.8}
$$

If β is an isomorphism, then from diagram (3.7), we see that α is the zero map. Therefore, $\tau-1=0$, i.e., the action is unramified. □

We have the following key lemma, which corresponds to [18, Proposition 1].

Lemma 3.4. *With the assumptions in Theorem 1.1,* $\dim_{\bar{\mathbb{F}}_p} X(H)\otimes\bar{\mathbb{F}}_p\le 2$. *The dimension is* 2 *if and only if* $\bar\rho$ *is unramified at* $\mathfrak{p}$ *and* $\mathrm{Frob}_{\mathfrak{p}}$ *acts as* ± 1.

Proof. We have a surjection $\beta: M(H)\otimes\bar{\mathbb{F}}_p\twoheadrightarrow X(H)\otimes\bar{\mathbb{F}}_p$ with $\dim M(H)\otimes\bar{\mathbb{F}}_p=2$. Hence $\dim_{\bar{\mathbb{F}}_p} X(H)\otimes\bar{\mathbb{F}}_p\le 2$. If $\dim X(H)\otimes\bar{\mathbb{F}}_p=2$, then $\beta: M(H)\otimes\bar{\mathbb{F}}_p\to X(H)\otimes\bar{\mathbb{F}}_p$ is an isomorphism. By Lemma 3.3, the $G_{F_{\mathfrak{p}}}$-action on $M(H)\otimes\bar{\mathbb{F}}_p$ is unramified. Note that $\mathrm{Frob}_{\mathfrak{p}}$ acts on $Y(H)\otimes\bar{\mathbb{F}}_p$ as $\pm\,\mathrm{Norm}(\mathfrak{p})$ by the discussion in Section 2.3, where the action of $\mathrm{Frob}_{\mathfrak{p}}$ on Σ_H is given explicitly. Hence $\mathrm{Frob}_{\mathfrak{p}}$ acts on $X(H)\otimes\bar{\mathbb{F}}_p$ as $\pm\,\mathrm{Norm}(\mathfrak{p})$. Taking the determinant of $\mathrm{Frob}_{\mathfrak{p}}$, on $X(H)\otimes\bar{\mathbb{F}}_p$ we obtain $\mathrm{Norm}(\mathfrak{p})^2$, and on $M(H)\otimes\bar{\mathbb{F}}_p$ we obtain $\mathrm{Norm}(\mathfrak{p})$. Thus $\mathrm{Norm}(\mathfrak{p})^2\equiv\mathrm{Norm}(\mathfrak{p}) \pmod{\mathfrak{m}}$ and $\mathrm{Frob}_{\mathfrak{p}}$ acts as ± 1.

Conversely, assume that $\bar\rho$ is not induced from a character if $F\supset\mathbb{Q}(\mu_p)^+$, $\bar\rho$ is unramified at $\mathfrak{p}$, and $\mathrm{Frob}_{\mathfrak{p}}$ acts as ± 1. Taking determinant, we see that $\mathrm{Norm}(\mathfrak{p})\equiv 1 \pmod{\mathfrak{m}}$. Because $\bar\rho$ is unramified at $\mathfrak{p}$ and is irreducible, by the level-lowering results [16, Main Theorem 1] and [14, Theorem 0.1], the map

$$A_0 := \varrho_1^* + \varrho_{\eta_{\mathfrak{p}}}^* : H^1_{\mathrm{et}}(M_{K_0(N/\mathfrak{p})}\otimes\bar F,\mathcal{F})^2_{\mathfrak{m}} \to H^1_{\mathrm{et}}(M_{K_0(N)}\otimes\bar F,\mathcal{F})_{\mathfrak{m}}$$

is surjective. Note that here $\mathfrak{m}$ in the first term contains the element $T_{\mathfrak{p}}-\mathrm{Tr}\,\bar\rho(\mathrm{Frob}_{\mathfrak{p}})$. Let C be the diagonal map (resp. anti-diagonal map)

$$C: H^1_{\mathrm{et}}(M_{K_0(N/\mathfrak{p})}\otimes\bar F,\mathcal{F})_{\mathfrak{m}} \to H^1_{\mathrm{et}}(M_{K_0(N/\mathfrak{p})}\otimes\bar F,\mathcal{F})^2_{\mathfrak{m}}$$

if $\mathrm{Frob}_{\mathfrak{p}}$ acts as -1 (resp. as $+1$). We claim that $(A_0\otimes\bar{\mathbb{F}}_p)\circ(C\otimes\bar{\mathbb{F}}_p)$ is surjective.

We prove the claim by adapting an idea from Taylor-Wiles system construction (cf. [20, Lemma 2.2]). Indeed, the map A_0 commutes with the Hecke operator T_v for $v \nmid pNS_D$. For the prime $\mathfrak{p}$, we have

$$U_{\omega_\mathfrak{p}} \circ A_0 = A_0 \circ \begin{pmatrix} T_\mathfrak{p} & \mathrm{Norm}(\mathfrak{p})S_\mathfrak{p} \\ -1 & 0 \end{pmatrix}.$$

To verify this identity, it suffices to verify it in the complex case, which follows from the double coset decompositions of

$$K_0(N)\eta_\mathfrak{p}K_0(N) \quad \text{and} \quad K_0(N/\mathfrak{p})\eta_\mathfrak{p}K_0(N/\mathfrak{p}) \quad \text{(cf. [20, Section 1]).}$$

Hence as operators in $\mathrm{End}(H^1_{\mathrm{et}}(M_{K_0(N/\mathfrak{p})} \otimes \bar{F}, \mathcal{F})^{\oplus 2})$, we have

$$U_{\omega_\mathfrak{p}}^2 - T_\mathfrak{p}U_{\omega_\mathfrak{p}} + \mathrm{Norm}(\mathfrak{p})S_\mathfrak{p} = 0.$$

Note that $\mathrm{Frob}_\mathfrak{p}$ also satisfies the above equation by Eichler-Shimura. After tersoring with $\bar{\mathbb{F}}_p$, the equation has one root with multiplicity two. Hence $U_{\omega_\mathfrak{p}}$ acts the same way as $\mathrm{Frob}_\mathfrak{p}$ on $(H^1_{\mathrm{et}}(M_{K_0(N/\mathfrak{p})} \otimes \bar{F}, \mathcal{F}) \otimes \bar{\mathbb{F}}_p)^{\oplus 2}$. The claim then follows since $\mathrm{Image}((A_0 \otimes \bar{\mathbb{F}}_p) \circ (C \otimes \bar{\mathbb{F}}_p))$ is exactly the eigenspace of $U_{\omega_\mathfrak{p}}$. Thus the composition

$$\begin{aligned}(3.9)\quad H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N/\mathfrak{p})} \otimes \bar{k}_\mathfrak{p}, \mathcal{F})_\mathfrak{m} \otimes \bar{\mathbb{F}}_p &\xrightarrow{C'} (H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N/\mathfrak{p})} \otimes \bar{k}_\mathfrak{p}, \mathcal{F})_\mathfrak{m} \otimes \bar{\mathbb{F}}_p)^2 \\ &= (H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N/\mathfrak{p})} \otimes \bar{F}_\mathfrak{p}, \mathcal{F})_\mathfrak{m} \otimes \bar{\mathbb{F}}_p)^2 \\ &\xrightarrow{A_0'} H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N)} \otimes \bar{F}_\mathfrak{p}, \mathcal{F})_\mathfrak{m} \otimes \bar{\mathbb{F}}_p \\ &= M(H) \otimes \bar{\mathbb{F}}_p\end{aligned}$$

is surjective, where C' is the diagonal map (resp. anti-diagonal map) if C is the diagonal map (resp. anti-diagonal map), A_0' is defined by the same formula as for A_0.

Define the map A by

$$\begin{aligned} A : H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N/\mathfrak{p})} \otimes \bar{k}_\mathfrak{p}, \mathcal{F})^2 &\to H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N)} \otimes \bar{k}_\mathfrak{p}, \mathcal{F}) \\ (f, g) &\mapsto (\varrho_1)^* f + (\varrho_{\eta_\mathfrak{p}})^* g \end{aligned}$$

and the map B to be induced by the normalization map

$$B : H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N)} \otimes \bar{k}_\mathfrak{p}, \mathcal{F}) \to H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N/\mathfrak{p})} \otimes \bar{k}_\mathfrak{p}, \mathcal{F})^2.$$

By Lemma 2.9 (cf. [18, Lemma 1]), $B \circ A = \begin{pmatrix} \mathrm{id} & \mathrm{Frob}_\mathfrak{p} \\ \mathrm{Frob}_\mathfrak{p} & \mathrm{id} \end{pmatrix}$. By our definition of C', we have $(B \otimes \bar{\mathbb{F}}_p) \circ (A \otimes \bar{\mathbb{F}}_p) \circ C' = 0$. Let $W \subset H^1_{\mathrm{et}}(\mathbf{M}_{K_0(N/\mathfrak{p})} \otimes$

$\bar{k}_{\mathfrak{p}}, \mathcal{F})_{\mathfrak{m}} \otimes \bar{\mathbb{F}}_p$ be a lift of $M(H) \otimes \bar{\mathbb{F}}_p$ via the surjection (3.9). Since $\delta \circ (A \otimes \bar{\mathbb{F}}_p) \circ C'(W) = M(H) \otimes \bar{\mathbb{F}}_p$,

$$\begin{aligned} \dim M(H) \otimes \bar{\mathbb{F}}_p &\leq \dim((A \otimes \bar{\mathbb{F}}_p) \circ C'(W)) \\ &\leq \dim \operatorname{Ker}(B \otimes \bar{\mathbb{F}}_p) \leq \dim Y(H) \otimes \bar{\mathbb{F}}_p \\ &= \dim X(H) \otimes \bar{\mathbb{F}}_p. \end{aligned} \tag{3.10}$$

Hence the above inequalities are all equalities and the lemma follows. □

3.2. Exact sequences for M'_O

We now describe the corresponding picture for the curve M'_O. Note that here we consider the reduction mod $\mathfrak{q}$. The exact sequence of vanishing cycles for the proper morphism $\mathbf{M}'_O \otimes \mathcal{O}_{F_\mathfrak{q}} \to \operatorname{Spec} \mathcal{O}_{F_\mathfrak{q}}$ and the sheaf $\mathcal{F}$ is

$$\begin{aligned} 0 &\to H^1_{\text{et}}(\mathbf{M}'_O \otimes \bar{k}_\mathfrak{q}, \mathcal{F}) \to H^1_{\text{et}}(\mathbf{M}'_O \otimes \bar{F}_\mathfrak{q}, \mathcal{F}) \to \bigoplus_{x \in \Sigma_O} (R^1\Phi_{\bar{\eta}}\mathcal{F})_x \\ &\to H^2_{\text{et}}(\mathbf{M}'_O \otimes \bar{k}_\mathfrak{q}, \mathcal{F}) \to H^2_{\text{et}}(\mathbf{M}'_O \otimes \bar{F}_\mathfrak{q}, \mathcal{F}) \to \cdots, \end{aligned} \tag{3.11}$$

where Σ_O denotes the set of singular points for the $\bar{k}_\mathfrak{q}$-scheme $\mathbf{M}'_O \otimes \bar{k}_\mathfrak{q}$. Note that Σ_O consists of a finite number of non-degenerate quadratic points.

Write

- $L(O) = \operatorname{Ker}(H^2_{\text{et}}(\mathbf{M}'_O \otimes \bar{k}_\mathfrak{q}, \mathcal{F}) \to H^2_{\text{et}}(\mathbf{M}'_O \otimes \bar{F}_\mathfrak{q}, \mathcal{F}))$,
- $Z(O) = H^1_{\text{et}}(\mathbf{M}'_O \otimes \bar{k}_\mathfrak{q}, \mathcal{F})$,
- $M(O) = H^1_{\text{et}}(\mathbf{M}'_O \otimes \bar{F}_\mathfrak{q}, \mathcal{F})$,
- $X(O) = \bigoplus_{x \in \Sigma_O} (R^1\Phi_{\bar{\eta}}\mathcal{F})_x$,
- $\tilde{X}(O) = \operatorname{Ker}(X(O) \to L(O))$.

Then we have a short exact sequence

$$0 \to Z(O) \to M(O) \to \tilde{X}(O) \to 0. \tag{3.12}$$

We construct a second exact sequence. Let r be the normalization map of the special fibre of $\mathbf{M}'_O$. We define $\mathcal{G}'$ via the following exact sequence of sheaves

$$0 \to \mathcal{F} \to r_* r^* \mathcal{F} \to \mathcal{G}' \to 0.$$

Then $\mathcal{G}'$ is a skyscraper sheaf supported on Σ_O. Taking the long exact sequence, we obtain

$$(3.13)\quad \begin{aligned} 0 &\to H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{F}) \to H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, r_* r^* \mathcal{F}) \to H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{G}') \\ &\to H^1_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{F}) \to H^1_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, r_* r^* \mathcal{F}) \to 0. \end{aligned}$$

Write

- $\tilde{L}(O) = \mathrm{Im}(\alpha : H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, r_* r^* \mathcal{F}) \to H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{G}'))$,
- $Y(O) = H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{G}') = \bigoplus_{x \in \Sigma_O} \mathcal{G}'_x$,
- $R(O) = H^1_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, r_* r^* \mathcal{F})$,
- $\tilde{Y}(O) = Y(O)/\tilde{L}(O)$.

Then we have another short exact sequence

$$(3.14)\qquad 0 \to \tilde{Y}(O) \to Z(O) \to R(O) \to 0.$$

Lemma 3.5. 1) *$H^i_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, r_* r^* \mathcal{F}) \cong H^i_{\mathrm{et}}(\mathbb{P}^1 \otimes \bar{k}_{\mathfrak{q}}, \mathcal{F})^c$, for some $c \in \mathbb{Z}_{\geq 1}$.*

2) *There is an isomorphism*

$$H^1_{\Sigma_O}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{F}) \tilde{\to} \bigoplus_{x \in \Sigma_O} H^1_{\{x\}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, R\Psi_{\bar{\eta}} \mathcal{F}).$$

3) *There is an isomorphism*

$$H^0_{\mathrm{et}}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{G}') \cong H^1_{\Sigma_O}(\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}, \mathcal{F}).$$

4) *There is an isomorphism*

$$N : X(O)(1) \to Y(O).$$

5) *$\tilde{L}(O)_{\mathfrak{m}} = 0$, $L(O)_{\mathfrak{m}} = 0$, $R(O)_{\mathfrak{m}} = 0$, if $\mathfrak{m}$ is a non-Eisenstein maximal ideal of the Hecke algebra.*

Proof. We know that the reduction Ω (mod $\mathfrak{p}$) is a union of $\mathbb{P}^1$. By Theorem 2.10, the special fibre $\mathbf{M}'_O \otimes \bar{k}_{\mathfrak{q}}$ is several copies of $\mathbb{P}^1$ glued together transversally above the singular points. The normalization of the special fibre is a disjoint union of finitely many $\mathbb{P}^1$. The first statement follows.

For the last statement, the first two terms vanish because they come from the 0th and 2nd cohomology groups and $\mathfrak{m}$ is a non-Eisenstein ideal

([13, Section 4] and [2, Lemma 2.2]). The third term vanishes because the 1st cohomology group of $\mathbb{P}^1$ vanishes.

The arguments for other claims are the same as the arguments for Lemma 3.1. □

Combining exact sequences (3.12) and (3.14), we obtain the following short exact sequence

$$0 \to Y(O)_{\mathfrak{m}} \to M(O)_{\mathfrak{m}} \to X(O)_{\mathfrak{m}} \to 0,$$

together with an isomorphism $N : (X(O)(1))_{\mathfrak{m}} \tilde{\to} Y(O)_{\mathfrak{m}}$. Then we obtain the following exact sequence

$$Y(O) \otimes \bar{\mathbb{F}}_p \xrightarrow{\alpha'} M(O) \otimes \bar{\mathbb{F}}_p \xrightarrow{\beta'} X(O) \otimes \bar{\mathbb{F}}_p \to 0, \tag{3.15}$$

together with an isomorphism $N : X(O)(1) \otimes \bar{\mathbb{F}}_p \tilde{\to} Y(O) \otimes \bar{\mathbb{F}}_p$.

Lemma 3.6. *There is an isomorphism of Hecke modules*

$$Y(H)_{\mathfrak{m}} \cong Y(O)_{\mathfrak{m}}.$$

Proof. We deduce this from the exact sequence (3.14) of [16]. Rewriting that sequence in our setting, we have the exact sequence

$$0 \to Y(H')^2_{\mathfrak{m}} \to Y(H)_{\mathfrak{m}} \to Y(O)_{\mathfrak{m}} \to 0.$$

Here $Y(H')$ corresponds to the group $Y(H)$ for the curve $M_{K'}$ where $K' = K_0(N/\mathfrak{q})$. Note that $\bar{\rho}$ is ramified at $\mathfrak{q}$, hence $\mathfrak{m}$ is not $\mathfrak{q}$-old. On the other hand, $M_{K'}$ has full level structure at $\mathfrak{q}$ and $Y(H')$ is $\mathfrak{q}$-old. Therefore, $Y(H')_{\mathfrak{m}} = 0$ and the claim follows. □

3.3. Proof of Theorem 1.1

We combine the results in the above sections to prove Theorem 1.1.

Lemma 3.7. *$\bar{\rho}$ is unramified at $\mathfrak{q}$ if and only if*

$$\beta' : M(O) \otimes \bar{\mathbb{F}}_p \to X(O) \otimes \bar{\mathbb{F}}_p \tag{3.16}$$

is an isomorphism.

Proof. By the diagram in Remark 3.2, we can compute the action of the inertia group $I_{F_{\mathfrak{q}}}$. An element $\tau \in I_{F_{\mathfrak{q}}}$ acts via the diagram

$$\begin{array}{ccc} M(O)\otimes\bar{\mathbb{F}}_p & \xrightarrow{\tau-1} & M(O)\otimes\bar{\mathbb{F}}_p \\ {\scriptstyle\beta'}\big\downarrow & & \big\uparrow{\scriptstyle\alpha'} \\ X(O)\otimes\bar{\mathbb{F}}_p & \xrightarrow[\mathrm{Var}(\tau)]{} & Y(O)\otimes\bar{\mathbb{F}}_p. \end{array} \tag{3.17}$$

If $\bar{\rho}$ is unramified at $\mathfrak{q}$, then $M(O)\otimes\bar{\mathbb{F}}_p$ is unramified at $\mathfrak{q}$. Thus $\alpha'\,\mathrm{Var}(\tau)\beta' = \tau - 1 = 0$ for all $\tau \in I_{F_{\mathfrak{q}}}$. This is the same as $\alpha' N\beta' = 0$ since $\mathrm{Var}(\tau) = -\epsilon(\tau)N$ (cf. [12, Proposition 17.4]). Now β' is a surjection and N is an isomorphism, so $\alpha' = 0$. Thus by the exact sequence (3.15), β' is an isomorphism.

On the other hand, if β' is an isomorphism, then by the exact sequence (3.15), we see that α' is the zero map. Therefore, $\tau - 1 = 0$, i.e., the action is unramified. □

We have assumed that $\bar{\rho}$ is ramified at $\mathfrak{q}$, then we have the following lemma.

Lemma 3.8. $\dim_{\bar{\mathbb{F}}_p} M(O)\otimes\bar{\mathbb{F}}_p = 2\dim_{\bar{\mathbb{F}}_p} X(O)\otimes\bar{\mathbb{F}}_p$.

Proof. By exact sequence (3.15), we have

$$\begin{aligned} \dim_{\bar{\mathbb{F}}_p} M(O)\otimes\bar{\mathbb{F}}_p &\le \dim_{\bar{\mathbb{F}}_p} X(O)\otimes\bar{\mathbb{F}}_p + \dim_{\bar{\mathbb{F}}_p} Y(O)\otimes\bar{\mathbb{F}}_p \\ &= 2\dim_{\bar{\mathbb{F}}_p} Y(O)\otimes\bar{\mathbb{F}}_p = 2\dim_{\bar{\mathbb{F}}_p} Y(H)\otimes\bar{\mathbb{F}}_p \\ &= 2\dim_{\bar{\mathbb{F}}_p} X(H)\otimes\bar{\mathbb{F}}_p \le 2\dim_{\bar{\mathbb{F}}_p} M(H)\otimes\bar{\mathbb{F}}_p \le 4. \end{aligned} \tag{3.18}$$

By Lemma 3.7, we see that $\dim_{\bar{\mathbb{F}}_p} M(O)\otimes\bar{\mathbb{F}}_p = \dim_{\bar{\mathbb{F}}_p} X(O)\otimes\bar{\mathbb{F}}_p = 2$ cannot happen since $\bar{\rho}$ is ramified at $\mathfrak{q}$. The lemma follows. □

Proof of Theorem 1.1. $\dim_{\bar{\mathbb{F}}_p} \mathrm{Hom}(\bar{\rho}, H^1_{\mathrm{et}}(M'_O\otimes\bar{F},\mathcal{F})_{\mathfrak{m}}) \le 2$ follows from equation (3.18). Since $\dim_{\bar{\mathbb{F}}_p} Y(H)\otimes\bar{\mathbb{F}}_p = \dim_{\bar{\mathbb{F}}_p} Y(O)\otimes\bar{\mathbb{F}}_p$, we have

$$\dim_{\bar{\mathbb{F}}_p} X(H)\otimes\bar{\mathbb{F}}_p = \dim_{\bar{\mathbb{F}}_p} X(O)\otimes\bar{\mathbb{F}}_p.$$

In particular, $\dim_{\bar{\mathbb{F}}_p} X(O)\otimes\bar{\mathbb{F}}_p \le 2$. The dimension is 2 if and only if the $\mathrm{Gal}(\bar{F_{\mathfrak{p}}}/F_{\mathfrak{p}})$-action is unramified and $\mathrm{Frob}_{\mathfrak{p}}$ acts as ± 1 by Lemma 3.4. Theorem 1.1 is now clear. □

Remark 3.9. Now we have a multiplicity two result. If we start with a Shimura curve such that the cohomology group has multiplicity two, we can do the same thing as above (as long as we have enough primes in the level): construct a new Shimura curve by changing local invariants, compare the corresponding cohomology groups, etc. Then we may get a multiplicity 4 result similar to Theorem 1.1. In this way we may construct Shimura curves with multiplicity 2^n. This is closely related to [2, Conjecture 4.9].

3.4. Some remarks

In order to show that Theorem 1.1 is not vacuous, we show that the construction after Theorem 3 of [18] carries over in our setting. For this, we prove a level-raising result and prove a multiplicity one result.

3.4.1. Level raising. We have the following level-raising result corresponding to [19, Theorem 1]. Note that for our purpose, being *new* or *old* is not essential, hence the argument is simpler than those in [17, Theorem 7.5] and [16, Theorem 5]. We sketch the proof for completeness.

Proposition 3.10. *Let $\bar{\rho} : G_F \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ be a rank two modular representation appears in $H^1_{\mathrm{et}}(M_{K_0(\mathfrak{n})} \otimes_F \bar{F}, \mathcal{F})_{\mathfrak{m}}$, where $\mathfrak{n}$ is square-free, $\mathcal{F}$ is an $\bar{\mathbb{F}}_p$-sheaf attached to a Serre weight, $\mathfrak{m}$ is a non-Eisenstein maximal ideal of the Hecke algebra $\mathbb{T}(K_0(\mathfrak{n}), \mathcal{F})$. Let v be a prime of F such that $v \nmid p\mathfrak{n}S_D$ and satisfies one or both the identities*

$$\operatorname{Tr} \bar{\rho}(\mathrm{Frob}_v) \equiv \pm(\mathrm{Norm}(v) + 1) \pmod{p}.$$

Then there is a maximal ideal $\mathcal{M}$ of $\mathbb{T}(K_0(v\mathfrak{n}), \mathcal{F})$, such that $\bar{\rho}$ appears in the cohomology group $H^1_{\mathrm{et}}(M_{K_0(v\mathfrak{n})} \otimes_F \bar{F}, \mathcal{F})_{\mathcal{M}}$.

Proof. As explained in [19, Section 3], it suffices to construct a maximal ideal $\mathcal{M}$ of $\mathbb{T}(K_0(v\mathfrak{n}), \mathcal{F})$ such that $\mathbb{T}(K_0(v\mathfrak{n}), \mathcal{F})/\mathcal{M} \cong \mathbb{T}(K_0(\mathfrak{n}), \mathcal{F})/\mathfrak{m}$. Here $\mathfrak{m} = \mathfrak{m}_{\bar{\rho}}$ is the maximal ideal of $\mathbb{T}(K_0(\mathfrak{n}), \mathcal{F})$ generated by $\{p,\ T_w - \operatorname{Tr} \bar{\rho}(\mathrm{Frob}_w),\ \mathrm{Norm}_{F/\mathbb{Q}}(w) S_w - \det \bar{\rho}(\mathrm{Frob}_w) :\ w \nmid p\mathfrak{n}S_D\}$. As in the proof of Lemma 3.4, consider the map

$$\begin{aligned} A : H^1_{\mathrm{et}}(M_{K_0(\mathfrak{n})} \otimes \bar{F}, \mathcal{F}) \oplus H^1_{\mathrm{et}}(M_{K_0(\mathfrak{n})} \otimes \bar{F}, \mathcal{F}) &\to H^1_{\mathrm{et}}(M_{K_0(v\mathfrak{n})} \otimes \bar{F}, \mathcal{F}) \\ (f, g) &\mapsto (\varrho_1)^* f + (\varrho_{\eta_v})^* g. \end{aligned}$$

This map commutes with the Hecke operator T_w for $w \nmid p\mathfrak{n}vS_D$. For the prime v, we have

$$U_{\omega_v} \circ A = A \circ \begin{pmatrix} T_v & \mathrm{Norm}(v)S_v \\ -1 & 0 \end{pmatrix}.$$

Hence as operators in $\mathrm{End}(H^1_{\mathrm{et}}(M_{K_0(\mathfrak{n})} \otimes \bar{F}, \mathcal{F})^{\oplus 2})$, we have

$$U_{\omega_v}^2 - T_v U_{\omega_v} + \mathrm{Norm}(v)S_v = 0.$$

If $\mathrm{Tr}\,\bar{\rho}(\mathrm{Frob}_v) \equiv -(\mathrm{Norm}(v)+1) \pmod p$, we take

$$\mathcal{M} = \{p,\ U_{\omega_v}+1,\ T_w - \mathrm{Tr}\,\bar{\rho}(\mathrm{Frob}_w),\\ \mathrm{Norm}_{F/\mathbb{Q}}(w)S_w - \det\bar{\rho}(\mathrm{Frob}_w) : w \nmid p\mathfrak{n}S_D\}.$$

If $\mathrm{Tr}\,\bar{\rho}(\mathrm{Frob}_v) \equiv (\mathrm{Norm}(v)+1) \pmod p$, we take

$$\mathcal{M} = \{p,\ U_{\omega_v}-1,\ T_w - \mathrm{Tr}\,\bar{\rho}(\mathrm{Frob}_w),\\ \mathrm{Norm}_{F/\mathbb{Q}}(w)S_w - \det\bar{\rho}(\mathrm{Frob}_w) : w \nmid p\mathfrak{n}S_D\}.$$

The $\mathcal{M}$ constructed as above then satisfies the properties as explained after Lemma 1 of [19]. Hence the proposition follows. □

3.4.2. More on the case $F = \mathbb{Q}$. As indicated in [19, Section 1] and explained in [9], the results of [19] lead to results for cusp forms of weight $k \geq 2$. In the following, we show that the results hold for general quaternion algebras over $\mathbb{Q}$. In particular, we have a multiplicity one result to start our construction (cf. [18, Theorem 1]).

Use the same notation as in Section 1 and take $F = \mathbb{Q}$. Note that we allow $S_D = \emptyset$, in which case we use the compact modular curves and it is the case considered in [9, 18].

Let $M_{K_0(N)}$ be the Shimura curve over $\mathbb{Q}$ with level $K_0(N)$, $\mathfrak{m}$ be a maximal non-Eisenstein ideal of the corresponding Hecke algebra. Then we have the following result.

Theorem 3.11. *Let $\bar{\rho} : G_{\mathbb{Q}} \to GL_2(\bar{\mathbb{F}}_p)$ be a modular Galois representation with conductor dividing pNS_D. Let b be an integer such that $0 \leq b \leq p-3$. Let V be the Serre weight* $\mathrm{Symm}^b\,\bar{\mathbb{F}}_p^2$. *Assume that $\bar{\rho}$ satisfies the following conditions*

1) *the restriction $\bar{\rho}|_{G_{F'}}$ is irreducible, where $F' = \mathbb{Q}(\sqrt{(-1)^{(p-1)/2}p})$;*

2) *if* $v = p$, *then* $\mathrm{End}_{\bar{\mathbb{F}}_p[G_{\mathbb{Q}_v}]}(\bar{\rho}|_{G_{\mathbb{Q}_v}}) = \bar{\mathbb{F}}_p$*;*

3) *if* $v \mid S_D$*, and* $v^2 \equiv 1 \pmod{p}$*, then* $\bar{\rho}$ *is ramified at* v*.*

Then $\dim_{\bar{\mathbb{F}}_p} \mathrm{Hom}_{G_{\mathbb{Q}}}(\bar{\rho}, H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{\mathbb{Q}}, \mathcal{F}_V)_{\mathfrak{m}}) \leq 1$.

Proof. We show that if the dimension is not zero, then it is one. First, we assume that if $v \mid N$, then $\bar{\rho}$ is ramified at v. This is the so called *minimal case* in Taylor-Wiles method (cf. [21, 24]). In this case, we define a deformation condition for $\bar{\rho}$ as in [22, Section 2], except at $v = p$, where we require the deformations to be Fontaine-Laffaille with weights 0 and $(b+1)$ (cf. [6, Section 2.4.1]). Constructing a Taylor-Wiles system as in [22, Section 3] and applying [9, Theorem 2.1], we obtain that $H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{\mathbb{Q}}, \mathcal{F}_V)_{\mathfrak{m}}$ is free of rank two over $\mathbb{T}(K_0(N), \mathcal{F}_V)_{\mathfrak{m}}$. Hence $\dim_{\bar{\mathbb{F}}_p} \mathrm{Hom}_{G_{\mathbb{Q}}}(\bar{\rho}, H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{\mathbb{Q}}, \mathcal{F}_V)_{\mathfrak{m}}) = 1$.

Combining Theorem 2.4 and the minimal case above, by a standard argument in Taylor-Wiles method (see for example [9, Section 3], [20, Section 3]), we can pass from the minimal case to non-minimal case and prove that $H^1_{\mathrm{et}}(M_{K_0(N)} \otimes \bar{\mathbb{Q}}, \mathcal{F}_V)_{\mathfrak{m}}$ is free of rank two over $\mathbb{T}(K_0(N), \mathcal{F}_V)_{\mathfrak{m}}$. The theorem then follows. □

Remark 3.12. For general totally real fields F, [5, Theorem 5.3] gives similar multiplicity one result in the minimal case. Since we do not have Ihara's Lemma in general, the above argument no longer applies and we only consider $F = \mathbb{Q}$ in this section.

Acknowledgements

The authors would like to thank Matthew Emerton for many enlightening discussions. They would like to thank Brian Conrad, Fred Diamond, and Mark Kisin for helpful conversations and correspondences, Panagiotis Tsaknias for pointing out a mistake in an early version of this paper. Finally, the authors would like to thank the referee for numerous comments and suggestions.

References

[1] Nigel Boston, Hendrik W. Lenstrai Jr., and Kenneth Ribet, *Quotients of group rings arising from two-dimensional representations*, C. R. Acad. Sci. Paris Sér. I Math. **312** (1991), no. 4, 323–328.

[2] Kevin Buzzard, Fred Diamond, and Frazer Jarvis, *On Serre's conjecture for mod ℓ Galois representations over totally real fields*, Duke Math. J. **155** (2010), no. 1, 105–161.

[3] Henri Carayol, *Sur la mauvaise réduction des courbes de Shimura*, Compositio Math. **59** (1986), no. 2, 151–230.

[4] Henri Carayol, *Sur les représentations l-adiques associées aux formes modulaires de Hilbert*, Ann. Sci. École Norm. Sup. (4) **19** (1986), no. 3, 409–468.

[5] Chuangxun Cheng, *Multiplicity one of regular Serre weights*, Israel J. Math. **198** (2013), no. 1, 419–466.

[6] Laurant Clozel, Michael Harris, and Richard Taylor, *Automorphy for some l-adic lifts of automorphic mod l representations*, Pub. Math. IHES **108** (2008), 1–181.

[7] Brian Conrad, Fred Diamond, and Richard Taylor, *Modularity of certain potentially Barsotti-Tate Galois representations*, J. Amer. Math. Soc. **12** (1999), no. 2, 521–567.

[8] Pierre Deligne and Michael Rapoport, *Les schémas de modules de courbes elliptiques*, Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 143–316. Lecture Notes in Math., Vol. 349, Springer, Berlin, 1973.

[9] Fred Diamond, *The Taylor-Wiles construction and multiplicity one*, Invent. Math. **128** (1997), no. 2, 379–391.

[10] Fred Diamond and Richard Taylor, *Non-optimal levels of mod l modular representations*, Invent. Math. **115** (1994), no. 3, 435–462.

[11] Vladimir Gershonovich Drinfeld, *Coverings of p-adic symmetric regions*, Funct. Anal. Appl. **10** (1976), 29–40.

[12] Frazer Jarvis, *Mazur's principle for totally real fields of odd degree*, Compositio Math. **116** (1999), no. 1, 39–79.

[13] Frazer Jarvis, *Correspondences on Shimura curves and Mazur's principle at p*, Pacific J. Math. **213** (2004), no. 2, 267–280.

[14] Frazer Jarvis, *Level lowering for modular mod l representations over totally real fields*, Mathematische Annalen **313** (1999), no. 1, 141–160.

[15] Bruce Jordan, Ron Livne, and Yakov Varshavsky, *Local points on p-adically uniformazed Shimura varieties*, `arXiv:math/0204361`.

[16] Ali Rajaei, *On levels of mod l Hilbert modular forms*, J. reine angew. Math. **537** (2001), 33–65.

[17] Kenneth A. Ribet, *On modular representations of* $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), no. 2, 431–476.

[18] Kenneth A. Ribet, *Multiplicities of Galois representations in Jacobians of Shimura curves*, Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, Part II (Ramat Aviv, 1989), pp. 221–236, Israel Math Conf. Proc. **3**, Weizmann, Jerusalem, 1990.

[19] Kenneth A. Ribet, *Raising the levels of modular representations*, Séminaire de Théorie des Nombres, Paris (1987-88), pp. 259–271, Progr. Math. **81**, Birkhäuser Boston, Boston, MA, 1990.

[20] Richard Taylor, *On the meromorphic continuation of degree two L-functions*, Doc. Math. (2006), Extra Vol., 729–779.

[21] Richard Taylor and Andrew Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (2) **141** (1995), no. 3, 553–572.

[22] Lea Terracini, *A Taylor-Wiles system for quaternionic Hecke algebras*, Compositio Math. **137** (2003), no. 1, 23–47.

[23] Yakov Varshavsky, *p-adic uniformization of unitary Shimura varieties*, J. Differential Geometry **49** (1998), 75–113.

[24] Andrew Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), no. 3, 443–551.

Department of Mathematics, Nanjing University
Nanjing 210093, China
E-mail address: cxcheng@nju.edu.cn

Department of Mathematics, Nanzhou College
Nanxian 413221, China
E-mail address: zmtheorem@qq.com

Received June 26, 2016