LEVEL LOWERING FOR GSP(4) AND VANISHING CYCLES ON SIEGEL THREEFOLDS

HAINING WANG

Abstract. In this article we prove some level lowering results for Siegel modular forms of degree 2 with paramodular level structure by adapting a method of Ribet in his proof of the epsilon conjecture. The proof relies on the description of the supersingular locus of a quaternionic Siegel threefold which is obtained by the author in a previous work. The heart of the proof is the so called $p,q$ switch trick introduced by Ribet where $p,q$ are two distinct prime numbers and it relies on the comparison between the vanishing cycle at $q$ on the paramodular Siegel threefold and the vanishing cycle at $p$ on the quaternionic Siegel threefold.

Contents

1. Introduction 1
2. Review of nearby and vanishing cycles 4
3. Automorphic and Galois representations for $GSp_4$ 7
4. Siegel threefold of paramodular level 12
5. Mazur’s principle for paramodular Siegel modular forms 14
6. Quaternionic Siegel threefold of paramodular level 17
7. Ribet’s principle for paramodular Siegel modular forms 20
References 24

1. Introduction

1.1. Motivations. In this article we provide two level lowering principles for Galois representations attached to a cuspidal automorphic representation of the symplectic similitude group $GSp_4$ of degree 4 that has paramodular fixed vectors. We refer to these as the Mazur’s principle and Ribet’s principle. These terminologies come from the celebrated work of Ribet on Serre’s epsilon conjecture [Rib90], [Rib91] and Mazur’s previous work. To motivate our results, we first recall their results in the $GL_2$ setting. Let

$$\rho : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(k)$$

be an irreducible representation of the absolute Galois group over $\mathbb{Q}$ valued in a finite field $k$ of characteristic $l \geq 3$. Suppose that $\rho$ is modular of level $N$ which means $\rho$ comes as the reduction of...
an $l$-adic representation
\[ \rho_{\pi,l} : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{GL}_2(E_\lambda) \]
attached to a weight 2 newform of level $\Gamma_0(N)$ whose associated automorphic representation which we denote by $\pi$. Here $E_\lambda$ is a suitable finite extension over $\mathbb{Q}$. Let $p$ be a prime which exactly divide $N$ and suppose that $\rho$ is unramified at $p$. Then Serre [Ser87] conjectured that $\rho$ is modular of level $N/p$. This conjecture is known as the Serre’s epsilon conjecture. The following theorem is the main result of [Rib90].

**Theorem** (Epsilon conjecture). Assume $\rho$ is modular of level $N$ and unramified at $p$ with $p \mid N$. Then $\rho$ is modular of level $N/p$ if $N$ is prime to $l$.

We remark that Ribet is able to prove Fermat’s last theorem using this result under the assumption that all the elliptic curves over $\mathbb{Q}$ are modular using Frey’s construction. The proof of the above theorem in fact breaks into two parts. First one proves the following theorem which is known as the Mazur’s principle.

**Theorem** (Mazur’s principle). Assume $\rho$ is modular of level $N$ and unramified at $p$ with $p \mid N$. Then $\rho$ is modular of level $N/p$ if $N$ is prime to $l$ and $p \not\equiv 1 \mod l$.

Note that by a theorem of Langlands [Lan73], $\rho_{G_0(p)}$ is of the form
\[ \begin{pmatrix} \chi & * \\ 0 & 1 \end{pmatrix} \]
where $\chi$ is the mod $l$ cyclotomic character. The assumption that $p \not\equiv 1 \mod l$ signifies that the two Frobenius eigenvalues of $\rho$ at $p$ are distinct even modulo $l$. One then proves next the following theorem which we call the Ribet’s principle.

**Theorem** (Ribet’s principle). Let $p, q$ be two distinct primes which are coprime to $lN$. Assume $\rho$ is modular of level $pqN$. Assume that $\rho$ is unramified at $p$ but ramified at $q$. Then $\rho$ is modular of level $qN$.

This result is not quite explicitly proved in [Rib90] but rather it is formulated and proved in [Rib91]. Combining Mazur’s principle and Ribet’s principle with level raising results, one can then prove the epsilon conjecture. The existence of the auxiliary prime $q$ allows Ribet to use the so-called $p, q$ switch trick. Here the crucial observation is that the singular locus of a Shimura curve associated to a suitable indefinite quaternion algebra and the singular locus of a suitable modular curve can be parametrized by a same discrete Shimura set. We make this more precise now. Let $B = B_{pq}$ be an indefinite quaternion algebra with discriminant $pq$. Let $X_{\Gamma_0(N)}^B$ be the Shimura curve associated to $B$ and level $\Gamma_0(N)$ and let $X_{\Gamma_0(pq)}^B$ be the usual modular curve with $\Gamma_0(pqN)$ level structure. Consider the special fiber $X_{\Gamma_0(N), \mathfrak{F}_p}^B$ of $X_{\Gamma_0(N)}^B$ at $p$. The singular locus of $X_{\Gamma_0(N), \mathfrak{F}_p}^B$ can be identified with the double coset
\[ B' \setminus \mathbb{A}_f^\times /K'((pq))K_0(N) \]
where $B' = B_{q\infty}$ is the definite quaternion algebra with discriminant $q$, $K_0(N) = \Gamma_0(N)$ is the completion of $\Gamma_0(N)$ and $K'((pq)) = \text{Iw}(p)\mathcal{O}_{B_q}^\times$ with $\text{Iw}(p)$ the Iwahori subgroup of $\text{GL}_2(\mathbb{Q}_p)$ and $\mathcal{O}_{B_q}$ the maximal order of $B'$. On the other hand, the singular locus of $X_{\Gamma_0(pq), \mathfrak{F}_q}^B$ at the prime $q$ can be identified with the same double coset. Therefore the vanishing cycles on $X_{\Gamma_0(N), \mathfrak{F}_p}^B$ and on $X_{\Gamma_0(pq), \mathfrak{F}_q}^B$ can be identified. The starting point of this article is the observation that the same phenomenon happens for the Siegel modular threefold with paramodular level structure and its quaternionic analogue.
1.2. Main results. Let $\pi$ be an automorphic representation of $\text{GSp}_4$ over $\mathbb{Q}$ which we assume throughout this article that is non-endoscopic and non-CAP. Suppose $\pi_{\infty}$ is in the holomorphic discrete series and cohomological of weights $a \geq b \geq 0$. Then one can attach a Galois representation

$$\rho_{\pi,l} : \text{Gal}((\overline{\mathbb{Q}}/\mathbb{Q})) \to \text{GSp}_4(E_{\lambda})$$

for a finite extension $E_{\lambda}$ over $\mathbb{Q}_l$. We will denote by

$$\overline{\rho}_{\pi,l} : \text{Gal}((\overline{\mathbb{Q}}/\mathbb{Q})) \to \text{GSp}_4(k)$$

its associated residual representation valued in the residue field $k$ of $E_{\lambda}$. We will assume that the residual characteristic $l > a + b + 4$. In fact this Galois representation is realized in the cohomology of the Siegel threefold in degree three. That is it is realized in $\text{Hom}_{\text{GSp}_4(\mathbb{A}_f)}(\pi_f, H^3(X_{\mathbb{Q}}, \mathbb{V}))$ where

$$H^3(X_{\mathbb{Q}}, \mathbb{V}) = \lim_{\substack{\longrightarrow \ U}} H^3(X_U, \mathbb{V})$$

is the interior cohomology of the Siegel threefolds indexed by the open compact subgroups $U$ of $\text{GSp}_4(\mathbb{A}_f)$ and $\mathbb{V}$ is a suitable $l$-adic étale automorphic sheaf which we will make it more precise in the main body of this article. Recall that there are two maximal parahoric subgroups in $\text{GSp}_4(\mathbb{Q}_p)$ up to isomorphism, one is known as the hyperspecial subgroup which we will denote by $H$ and the other is known as the paramodular subgroup which we denote by $K(p)$. Suppose that the local component of $\pi$ at $p$ is para-spherical which means that $\pi^K(p)$ is non-zero. In this case, $\pi$ is ramified at $p$. We will let $U = K(p)U^p \subset \text{GSp}_4(\mathbb{Q}_p)\text{GSp}_4(\mathbb{A}_f^p)$ be a neat open compact subgroup such that $\pi^U \neq 0$. Let $X_{K(p)} = X_U$ be the Siegel threefold with paramodular level at $p$. Let $T_p$ be a suitable Hecke algebra which we will make it more explicit in the main body of this article and let $\mathfrak{m}$ be the maximal ideal in $T_p$ corresponding to $\overline{\rho}_{\pi,l}$. Then our first result is the analogue of Mazur’s principle for holomorphic Siegel modular forms of paramodular level.

Theorem 1. Let $\pi$ be a cuspidal automorphic representation of $\text{GSp}_4$ that is cohomological which is non-CAP and non-endoscopic whose component at infinity $\pi_{\infty}$ is a holomorphic discrete series of Harish Chandra parameter $(a, b, -a - b + 3)$ with $a \geq b \geq 0$ and such that $l > a + b + 4$. Let $p$ be a prime distinct from $l$ and such that $p \not\equiv 1 \pmod{l}$. Let $U = K(p)U^p \subset \text{GSp}_4(\mathbb{Q}_p)\text{GSp}_4(\mathbb{A}_f^p)$ be a neat open compact subgroup such that $\pi^U \neq 0$. We assume that $\pi_p$ is ramified. Suppose the residual Galois representation $\overline{\rho}_{\pi,l}$ satisfies the following assumptions

- $\overline{\rho}_{\pi,l}$ is unramified at $p$;
- $\overline{\rho}_{\pi,l}$ is irreducible;
- $H^3(X_{K(p),\overline{\mathbb{Q}}}, \mathbb{V}) \otimes k$ is semisimple as a Galois module.

Then there exists a cuspidal automorphic representation $\pi'$ with the same type as $\pi$ at $\infty$ such that $\overline{\rho}_{\pi,l} \cong \overline{\rho}_{\pi',l}$ and $\pi'_p$ is unramified.

We remark that in a previous work [VH19], Van Hoften was able to prove another version of Mazur’s principle in this setting. Our assumptions are quite different from his. In particular we are not relying on the vanishing results of Lan-Suh [LS13] and thus we allow more general weights for the étale automorphic sheaf $\mathbb{V}$. In addition, we do not require that the Frobenius eigenvalues of $\overline{\rho}_{\pi,l}$ at $p$ to be distinct instead we only assume that $p \not\equiv 1 \pmod{l}$. However we introduce the additional assumption that $H^3(X_{K(p),\overline{\mathbb{Q}}}, \mathbb{V}) \otimes k$ is semisimple. This assumption is the analogue of the main result of [BLR91]. However the proof of [BLR91] relies on the fact the representation of interest is 2 dimensional and therefore their methods do not generalize directly to our setting. It is nevertheless an interesting question to study the semi-simplicity of $H^3(X_{K(p),\overline{\mathbb{Q}}}, \mathbb{V}) \otimes k$ which we hope to treat in another occasion.
Now let $p, q$ be two distinct primes which are different from $l$. We fix $U^{pq}$ an open compact subgroup of $\text{GSp}_4(K_f^{(pq)})$ which is sufficiently small. Suppose now that for $U = K(p)K(q)U^{pq}$, $\pi^U \neq 0$. We consider the Siegel threefold $X_{K(pq)} = X_U$ with paramodular level at $p$ and $q$. Our second result is the analogue of Ribet’s principle for holomorphic Siegel modular forms of paramodular level.

**Theorem 2.** Let $\pi$ be a cuspidal automorphic representation of $\text{GSp}_4$ that is cohomological which is non-CAP and non-endoscopic whose component at infinity $\pi_{\infty}$ is a holomorphic discrete series of Harish Chandra parameter $(a, b, -a - b + 3)$ with $a \geq b \geq 0$ such that $l \geq a + b + 4$. Let $p, q$ be two distinct primes different from $l$. Let $U = K(p)K(q)U^{pq} \subset \text{GSp}_4(\mathbb{Q}_p)\text{GSp}_4(\mathbb{Q}_q)\text{GSp}_4(\mathbb{A}^{(p)})$ be a neat open compact subgroup such that $\pi^U \neq 0$. We assume that $\pi_p$ and $\pi_q$ are both ramified. Suppose the residual Galois representation $\bar{\rho}_{\pi, l}$ satisfies the following assumptions

- $\bar{\rho}_{\pi, l}$ is unramified at $p$ and is ramified at $q$;
- $\bar{\rho}_{\pi, l}$ is irreducible;
- $H^1_f(X_{K(pq)}, \mathbb{V})_m \otimes k$ is a semisimple Galois module.

Then there exists a cuspidal automorphic representation $\pi'$ with the same type as $\pi$ at $\infty$ and such that $\bar{\rho}_{\pi, l} \cong \bar{\rho}_{\pi', l}$ and $\pi'_p$ is unramified.

The proof of this theorem relies heavily on the interplay between the geometry of the Siegel threefold $X_{K(pq)}$ and its quaternionic analogue $X_B$. Here the quaternion algebra $B$ is again the indefinite quaternion algebra over $\mathbb{Q}$ with discriminant $pq$ and $X_B$ is the Shimura variety associated to the quaternionic unitary group $\text{GU}_2(B)$. We will rely on the description of the supersingular locus of $X_B$ in [Wang19a]. In particular the starting point of this work is the observation that the singular locus of $X_{K(pq)}$ and the singular locus $X_B^{\mathbb{F}_p}$ are parametrized by the same double coset and therefore one can identify the vanishing cycles on $X_{K(pq)}$ and the vanishing cycles on $X_B^{\mathbb{F}_p}$. Also the role of $p, q$ is symmetric in the study of these two Shimura varieties and this give us the flexibility to permute the two primes. It is natural to ask if one can upgrade this result to remove the assumption that there exits a prime $q$ such that $\pi_q$ is ramified and $\bar{\rho}_{\pi, l}$ is ramified at $q$. This seems to rely on suitable level raising results which are unknown to the author.

Our results seem to be the first to generalize Ribet’s techniques to higher dimensional Shimura varieties. In the case of Hilbert modular forms, there are works of Jarvis [Jar99] and Rajaei [Raj01] which work with Shimura curves over totally real field. Note that there are also other techniques to prove level lowering results, notably the work of Skinner-Wiles [SW01]. In the setting of $\text{GSp}_4$, Sorensen [Sor09a] is able to prove a potential level lowering result using the techniques of [SW01].

1.3. **Acknowledgement.** This work is completed when the author is a postdoctoral fellow at McGill university and he would like to thank Henri Darmon and Pengfei Guan for their generous support. We would like to thank Marc-Hubert Nicole for informing him the work of Van Hoften [VH19] and discussions about paramodular Siegel threefold.

2. **Review of nearby and vanishing cycles**

2.1. **Nearby and vanishing cycles.** Let $S = \text{Spec}(R)$ be the spectrum of a henselian DVR. We choose a uniformizer $\pi$ of $R$. We denote by $s$ the closed point of $S$ and by $\eta$ the generic point of $S$. We assume the residue field $k(s)$ at $s$ is of characteristic $p$. Let $\bar{s}$ be a geometric point of $S$ over $s$ and $\tilde{S}$ be the localization $S_{(\bar{s})}$ of $S$ at $\bar{s}$. We will denote by $\bar{s}$ and $\bar{\eta}$ the closed and generic point of $\tilde{S}$. Let
\( \tilde{\eta} \) be a separable closure of \( \eta \). We summarize the above notations in the following diagram.

\[
\begin{array}{ccc}
\tilde{s} & \to & \tilde{S} \\
\downarrow & & \downarrow \\
\tilde{s} & \to & \tilde{S} \\
\downarrow & & \downarrow \\
s & \to & S
\end{array}
\]

For a morphism \( f : X \to S \), we obtain by base change the following maps

\[
\begin{array}{ccc}
X_{\tilde{s}} & \to & X_{\tilde{S}} \\
\downarrow & & \downarrow \\
X_{\tilde{s}} & \to & X_{\tilde{S}} \\
\downarrow & & \downarrow \\
X_s & \to & X
\end{array}
\]

Let \( K \in D^+(X_{\tilde{\eta}}, \Lambda) \) with coefficient in \( \Lambda = \mathbb{Z}/l^n\mathbb{Z} \) or \( \mathbb{Z}_l \), then we define the nearby cycle complex \( R\Psi(K) \in D^+(X_s, \Lambda) \) by

\[
R\Psi(K) = \tilde{i}^* R^j_{s*}(K|X_{\tilde{\eta}}).
\]

For \( K \in D^+(X, \Lambda) \), the adjunction map defines the following distinguished triangle

\[
K|X_s \to R\Psi(K|X_{\tilde{\eta}}) \to R\Phi(K) \to
\]

The complex \( R\Phi(K) \) is known as the vanishing cycle complex and is the cone of the previous map. If \( f \) is proper, then we have

\[
R\Gamma(X_{\tilde{\eta}}, K|X_{\tilde{\eta}}) = R\Gamma(X_s, R\Psi(K|X_{\tilde{\eta}}))
\]

and in general we always have the following long exact sequence

\[
\cdots \to H^i(X_{\tilde{s}}, K|X_{\tilde{s}}) \xrightarrow{sp} H^i(X_{\tilde{s}}, R\Psi(K|X_{\tilde{\eta}})) \to H^i(X_s, R\Phi(K)) \to \cdots.
\]

The first map \( sp \) is called the specialization map and the cohomology of vanishing cycles \( H^i(X_{\tilde{s}}, R\Phi(K)) \) measures the defect of \( sp \) from being an isomorphism. If we assume that \( f \) is smooth, then \( H^i(X_s, R\Phi(K)) \) vanishes and the specialization map is an isomorphism.

2.2. Isolated singularities. Suppose that \( f : X \to S \) be a regular, flat, finite type morphism of relative dimension \( n \) which is smooth outside a finite collection \( \Sigma \) of closed points in \( X_s \). In this subsection, we only consider the trivial coefficient \( \Lambda \) for simplicity and all the results recalled here can be extended to more general constructible coefficients. In this case we have

\[
R\Phi(\Lambda)|X_s - \Sigma = 0
\]

and moreover

\[
R\Phi(\Lambda) = \bigoplus_{x \in \Sigma} R^n\Phi(\Lambda)_x
\]

is concentrated at degree \( n \). Therefore we have the following exact sequence

\[
0 \to H^n(X_s, \Lambda) \xrightarrow{sp} H^n(X_{\tilde{s}}, R\Psi(\Lambda)) \xrightarrow{\alpha} \bigoplus_{x \in \Sigma} R^n\Phi(\Lambda)_x \to
\]

\[
H^{n+1}(X_s, \Lambda) \to H^{n+1}(X_{\tilde{s}}, R\Psi(\Lambda)) \to 0.
\]

Assume next that for every \( x \in \Sigma \) is an ordinary quadratic singularity and furthermore \( H^n(X_{\tilde{s}}, R\Phi(\Lambda)) = H^n(X_{\tilde{\eta}}, \Lambda) \). This means \( X \) is étale locally near \( x \) isomorphic to

- \( V(\Sigma_{1 \leq i \leq m} x_i x_{i+m} + \pi) \subset \mathbb{A}_S^{2m} \) if \( n = 2m - 1 \);
- \( V(\Sigma_{1 \leq i \leq m} x_i x_{i+m} + x_{2m+1}^2 + \pi) \subset \mathbb{A}_S^{2m+1} \) if \( n = 2m \).
Then in this case
\[ R^n\Phi(\Lambda)_x = \Lambda \]
noncanonically. Let \( I \subset \text{Gal}(\bar{\eta}/\eta) \) be the inertia group and let \( \sigma \in I \). Then we have the variation map \( \text{Var}(\sigma) : (R^n\Phi\Lambda)_x \to H^n_x(X_{\bar{\eta}}, R\Psi(\Lambda)) \) and the action of \( \sigma - 1 \) on \( H^n(X_{\bar{\eta}}, \Lambda) \) can be factored as
\[
\begin{align*}
H^n(X_{\bar{\eta}}, \Lambda) &\xrightarrow{\sigma - 1} \bigoplus_{x \in \Sigma} R^n\Phi(\Lambda)_x \\
&\xrightarrow{\text{Var}(\sigma)} H^n(X_{\bar{\eta}}, \Lambda)
\end{align*}
\]
Here \( \beta \) is the composite
\[
\bigoplus_{x \in \Sigma} H^n_x(X_{\bar{s}}, R\Psi(\Lambda)) \rightarrow H^n(X_{\bar{s}}, \Lambda) \rightarrow H^n(X_{\bar{\eta}}, \Lambda)
\]
where the first map is the Gysin map under the identification
\[
\bigoplus_{x \in \Sigma} H^n_x(X_{\bar{s}}, R\Psi(\Lambda)) = \bigoplus_{x \in \Sigma} H^n_x(X_{\bar{s}}, \Lambda)
\]
and the second map is the specialization map. One has also a Frobenius equivariant version
\[
\begin{align*}
H^n(X_{\bar{\eta}}, \Lambda)(1) &\xrightarrow{\sigma} \bigoplus_{x \in \Sigma} R^n\Phi(\Lambda)_x(1) \\
&\xrightarrow{\text{Var}(\sigma)} H^n(X_{\bar{\eta}}, \Lambda)
\end{align*}
\]
where \( N \) and \( N_x \) are the monodromy and local monodromy operator. We will loosely refer to this diagram as the Picard-Lefschetz formula for \( X_{\bar{\eta}} \).

2.3. **Nearby cycles of automorphic étale sheaves.** Suppose \( f : X \to S \) is proper. Then the proper base change theorem implies that we have an isomorphism
\[
H^i(X_{\bar{\eta}}, \Lambda) \cong H^i(X_{\bar{s}}, R\Psi(\Lambda)).
\]
However if \( f : X \to S \) is not proper, then it is not always true that
\[
H^i(X_{\bar{\eta}}, \Lambda) \cong H^i(X_{\bar{s}}, R\Psi(\Lambda)).
\]
In the setting of Shimura varieties with good compactification, we do have such isomorphism. In particular all the Shimura varieties we will consider in this article have good compactifications. In fact all the Shimura varieties belong to the case of
- \((Nm)\): A flat integral model defined by taking normalization of a characteristic 0 PEL type moduli problem over a product of good reduction integral models of smooth PEL type moduli problem

in the classification of \([LS18a]\) and \([LS18b]\). The following theorem summarizes the results we need.

**Theorem 2.2** ([LS18b Corollary 4.6]). Suppose \( X \) is a Shimura variety that is in the case of \((Nm)\) and let \( \mathcal{V} \) be an automorphic étale sheaf defined as in \([LS18b] \S 3\). Then the canonical adjunction morphisms
\[
H^i(X_{\bar{\eta}}, \mathcal{V}) \rightarrow H^i(X_{\bar{s}}, R\Psi(\mathcal{V})).
\]
and
\[
H^i_c(X_{\bar{s}}, R\Psi(\mathcal{V})) \rightarrow H^i_c(X_{\bar{\eta}}, \mathcal{V})
\]
are isomorphisms for all \( i \).
3. Automorphic and Galois representations for $\text{GSp}_4$

3.1. The group $\text{GSp}_4$. Let $\text{GSp}_4$ be the symplectic similitude group defined by the set of matrices $g$ in $\text{GL}_4$ that satisfy $g^t J g = c(g) J$ for

\[
J = \begin{pmatrix}
1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 1
\end{pmatrix}
\]

and some $c(g) \in \mathbb{G}_m$. We define the map $c : \text{GSp}_4 \to \mathbb{G}_m$ by sending $g \in \text{GSp}_4$ to $c(g)$ and refer to it as the similitude map. The kernel of this map is by definition the symplectic group $\text{Sp}(4)$. There are two conjugacy classes of maximal parabolic subgroups given by

- the Siegel parabolic subgroup $P$, whose Levi factor is
  \[
  M_P = \left\{ \begin{pmatrix} A & uA' \\ u & A' \end{pmatrix} : u \in \text{GL}_1, A \in \text{GL}_2 \right\} \cong \text{GL}_1 \times \text{GL}_2
  \]
  where $A' = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} (A^{-1})^t \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$

- the Klingen parabolic subgroup $Q$, whose Levi factor is
  \[
  M_Q = \left\{ \begin{pmatrix} u & A \\ u^{-1} \det(A) & A \end{pmatrix} : u \in \text{GL}_1, A \in \text{GL}_2 \right\} \cong \text{GL}_1 \times \text{GL}_2.
  \]

Let $T$ be the diagonal torus in $\text{GSp}_4$ and $X^*(T)$ be its character group which we identify with subset \{(a, b, c) \in \mathbb{Z}^3 : a + b \equiv c \mod 2\} of $\mathbb{Z}^3$ by associating a triple $(a, b; c)$ the character

\[
(3.1) \quad \left( \begin{array}{c}
t_1 \\
t_2 \\
v t_2^{-1} \\
v t_1^{-1}
\end{array} \right) \to t_1^a t_2^b v^c.
\]

Let $B$ be the Borel subgroup of upper triangular matrices in $\text{GSp}_4$. Then the set of dominant weights $X^*(T)^+$ with respect to $B$ is given explicitly by \{(a, b; c) \in X^*(T) : a \geq b \geq 0\}.

3.2. Parahoric subgroups of $\text{GSp}_4(\mathbb{Q}_p)$. The affine Dynkin diagram of type $\check{C}_2$ is given by

\[
\bullet \quad \overset{0}{\longrightarrow} \quad \overset{1}{\bullet} \quad \overset{2}{\longleftarrow} \quad \overset{0}{\bullet}
\]

and correspondingly we have the parahoric subgroups given by

- The Iwahori subgroup $I$;
- The Siegel parahoric subgroup $K_{\{0,2\}}$, this is obtained from the Iwahori subgroup by adding the affine root group corresponding to 1; Its reduction is the Siegel parabolic $P$.
- The Klingen parahoric subgroup $K_{\{0,1\}}$, this is obtained from the Iwahori subgroup by adding the affine root group corresponding to 2; Its reduction is the Klingen parabolic $Q$.
- The paramodular parahoric subgroup $K_{\{1\}}$, this is obtained from the Iwahori subgroup by adding the affine root groups corresponding to 0 and 2;
- The hyperspecial parahoric subgroup $K_{\{0\}}$ and $K_{\{2\}}$, this is obtained from the Iwahori subgroup by adding the affine root groups corresponding to 1, 2 or 0. Note that they are conjugate to each other.
In this article we will be particularly interested to representations and automorphic forms related to the paramodular subgroup and therefore we denote it by $K(p)$ instead of $K_{(1)}$. We will also denote by $H$ the hyperspecial parahoric subgroup. Note that the hyperspecial subgroup and the paramodular subgroup are the maximal parahoric subgroups in $\text{GSp}_4(\mathbb{Q}_p)$, therefore there is no map between Siegel threefold of paramodular level to the Siegel threefold of hyperspecial level. This is one of the difference between our case and the case of modular curves or Shimura curves.

3.3. Hecke algebra. Let $p$ be a prime number. We let $\mathcal{H}_p$ be the spherical Hecke algebra over $\mathbb{Z}$. This is a commutative algebra isomorphic to $\mathbb{Z}[T_{p,0}, T_{p,1}, T_{p,2}]$ where

$$T_{p,0} = \text{char}(\text{GSp}_4(\mathbb{Z}_p)) \begin{pmatrix} p & p \\ p & p \end{pmatrix} \text{GSp}_4(\mathbb{Z}_p),$$

$$T_{p,1} = \text{char}(\text{GSp}_4(\mathbb{Z}_p)) \begin{pmatrix} 1 & 1 \\ p & p \end{pmatrix} \text{GSp}_4(\mathbb{Z}_p),$$

$$T_{p,2} = \text{char}(\text{GSp}_4(\mathbb{Z}_p)) \begin{pmatrix} 1 & p \\ p & p^2 \end{pmatrix} \text{GSp}_4(\mathbb{Z}_p).$$

(3.2)

where $\text{char}(\cdot)$ is the characteristic function. The Hecke polynomial is by definition given by

$$Q_p(X) = 1 - T_{p,2}X + p(T_{p,1} + (p^2 + 1)T_{p,0})X^2 - p^3T_{p,2}T_{p,0}X^3 + p^6T_{p,0}^2X^6.$$  

To any $\pi_p$ irreducible unramified admissible representation of $\text{GSp}_4(\mathbb{Q}_p)$, we associate a character

$$\Theta_{\pi_p} : \mathcal{H}_p \to \text{End}(\pi_p^{\text{GSp}_4(\mathbb{Z}_p)}) = \mathbb{C}$$

and a Langlands parameter

$$\left( \begin{array}{c} \alpha \\ \beta \\ \gamma \\ \delta \end{array} \right)$$

considered as an element in $\text{GSp}_4(\mathbb{C}) = \ell^4 \text{GSp}_4$. Then we have the following identity

$$\Theta_{\pi_p}(Q_p(X)) = (1 - p^{2/3}\alpha X)(1 - p^{2/3}\beta X)(1 - p^{2/3}\gamma X)(1 - p^{2/3}\delta X).$$

(3.3)

Given an automorphic representation $\pi$, we let $S$ be the product of primes at which that $\pi$ is ramified. Then we define the unramified Hecke algebra to be the restricted tensor product

$$T^S = \bigotimes_{p \nmid S} \mathcal{H}_p.$$  

3.4. Representations with paramodular fixed vector. Let $\pi_p$ be an irreducible smooth representation of $\text{GSp}_4(\mathbb{Q}_p)$ such that $\pi_p^{K(p)} \neq 0$. One can find the classification of these representations in [Schm05] for example. The classification shows that there are five types (IIa, IVc, Vb, Vc, VIc) and only IIa is generic. Let $| \cdot | : \mathbb{Q}_p^+ \to \mathbb{Q}_l$ be the normalized absolute value such that $|p| = \frac{1}{p}$ and let $\sigma$
be an unramified character \( \mathbb{Q}_p^\times \) and \( \chi \) be character of \( \mathbb{Q}_p^\times \) such that \( \chi \neq |\cdot|^{\pm}, |\cdot|^{\pm 3/2} \). Then the type IIa irreducible representation \( \chi_{St_{GL_2}} \otimes \sigma \) corresponds to the Weil-Deligne representation given by

\[
\begin{pmatrix}
\chi^2 \sigma & |\cdot|^{1/2} \chi \sigma \\
|\cdot|^{-1/2} \chi \sigma & |\cdot|^{-1/2} \chi \sigma
\end{pmatrix},
\]

with monodromy operator of rank one given by

\[
N = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{pmatrix}
\]

under the local Langlands correspondence for GSp\(4\) established in [GT11]. In other words, the \( L \)-parameter is given by the \( \chi^2 \sigma \oplus |\cdot|^{-1/2} \chi \sigma \otimes \text{Sp}_2 \oplus \sigma \). We will later refer to \( |\cdot|^{-1/2} \chi \sigma \otimes \text{Sp}_2 \) as the twisted Steinberg part of the Weil-Deligne representation.

We define the following elements in the paramodular Hecke algebra

\[
U_{p,1} = \text{char}(K(p)) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} K(p),
\]

\[
U_{p,2} = \text{char}(K(p)) \begin{pmatrix} 0 & 1 \\ 1 & p \end{pmatrix} K(p).
\]

The element which we call the Atkin-Lehner operator

\[
u_p = \begin{pmatrix} 1 & -1 \\ p & p \end{pmatrix}
\]

normalizes the paramodular subgroup \( K(p) \) and thus will acts on the space of \( K(p) \) fixed vectors.

**Lemma 3.6.** Let \( \pi_p = \chi_{St_{GL_2}} \otimes \sigma \) be a representation of type IIa with \( \chi \) and \( \sigma \) as above. Then \( \pi_{K(p)} \) is one dimensional on which \( u_p \) acts by the scalar \( \chi \sigma(p) \).

**Proof.** This is [VH19, Lemma 2.2.3]. \( \square \)

Let \( p \) and \( q \) be two distinct odd primes. Then we define the following Hecke algebras which we will also use in addition to \( T^S \)

\[
T_p = T^S[U_{p,0}, U_{p,1}, u_p].
\]

\[
T_{pq} = T^S[U_{p,0}, U_{p,1}, u_p, U_{q,0}, U_{q,1}, u_q].
\]
3.5. Cohomology of Siegel threefolds. Let $h : \text{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m = \mathbb{C}^\times \to \text{GSp}_4(\mathbb{R})$ sending $x + iy$ to 
\[
\begin{pmatrix}
  xI_2 & yS \\
 -yS & xI_2
\end{pmatrix}
\]
where $I_2$ is the identity matrix of size 2 and $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Let $K^h$ be the centralizer of $h$ in $\text{GSp}_4(\mathbb{R})$. Then $K^h = K_\infty \mathbb{R}^\times$ where $K_\infty$ is the maximal compact subgroup of $\text{GSp}_4(\mathbb{R})$. Let $U \subset \text{GSp}_4(\mathbb{A}^f)$ be an open compact subgroup and we have the Siegel threefold with level $U$ denoted by $X_U$. This is a quasi-projective variety whose $\mathbb{C}$-points are given by 
\[
X_U(\mathbb{C}) = \text{GSp}_4(\mathbb{Q}) \backslash (\text{GSp}_4(\mathbb{R})/K^h \times \text{GSp}_4(\mathbb{A}^f))/U.
\]
Let $\mu = (a, b; c) \in X^*(T)$ and $V_\mu$ be the corresponding irreducible representation of highest weight $\mu$. We fix an isomorphism $\mathbb{C} \cong \mathbb{Q}_l$ and let $E_\lambda/\mathbb{Q}_l$ be a large enough coefficient field with ring of integral elements $O = \mathcal{O}_{E_\lambda}$ and uniformizer $\varpi$. We can consider $V_\mu$ as a representation of $\text{GSp}_4$ over $E_\lambda$ and we fix a lattice $V_\mu^{\text{int}}$ in $V_\mu$. We review the construction of an automorphic étale sheaf $\mathcal{V}_\mu$ on $X_U$. For $m$ sufficiently large, let $U(l^m)$ be the open compact subgroup in $U = \text{GSp}_4(\mathbb{Z})$ that acts trivially on $V_\mu^{\text{int}}/\varpi^m$. Then $\mathcal{V}_\mu^{\text{int}}$ is defined to be 
\[
(\varprojlim X_U(l^m)/U \times V_\mu^{\text{int}}/\varpi^m)/U.
\]
We that $l > a + b + 4$ and this guarantees that the étale automorphic sheaf is independent of the choice of lattice. We denote its associated $\mathbb{Q}_l$-local system $\mathcal{V}_\mu^{\text{int}} \otimes \mathbb{Q}_l$ by $V_\mu$ and its associated $\mathbb{C}$-local system $\mathcal{V}_{\mu, \mathbb{C}} = V_\mu^{\text{int}} \otimes \mathbb{C}$.

- The intersection cohomology $\mathcal{IH}^*(X_U(\mathbb{C}), V_\mu)$;
- The $L^2$-cohomology $H^*_2(X_U(\mathbb{C}), V_{\mu, \mathbb{C}})$ which can be computed by the Matsushima formula 
\[
H^*_2(X_U(\mathbb{C}), V_{\mu, \mathbb{C}}) = \bigoplus_{\pi} \pi_U \otimes H^*(\mathfrak{h}, K_\infty, V_{\mu, \mathbb{C}} \otimes \pi_\infty)
\]
where $\pi$ runs through all the automorphic representations $\pi = \pi_f \otimes \pi_\infty$ that occurs in the discrete spectrum of $L^2(\text{GSp}_4(\mathbb{Q})/\text{GSp}_4(\mathbb{A}))$ with multiplicity $m(\pi)$. There is a natural direct summand $H^*_\text{cusp}(X_U(\mathbb{C}), V_{\mu, \mathbb{C}})$ of cuspidal cohomology which is the direct sum over cuspidal automorphic representations $\pi$.

- The interior cohomology 
\[
H^*_i(X_U(\mathbb{C}), V_\mu) = \text{Im}(H^*_c(X_U(\mathbb{C}), V_\mu) \to H^*(X_U(\mathbb{C}), V_\mu)).
\]
Moreover we have an isomorphism 
\[
\mathcal{IH}^*(X_U(\mathbb{C}), V_{\mu, \mathbb{C}}) = H^*_2(X_U(\mathbb{C}), V_{\mu, \mathbb{C}}).
\]

3.6. Galois representations. Let $\pi$ be a cuspidal automorphic representation of $\text{GSp}_4$ which is not CAP, whose $\pi_\infty$ is a holomorphic discrete series of Harish Chandra parameter $(a, b, -a - b + 3)$ with $a \geq b \geq 0$. Let $S$ be the product of primes $p$ such that $\pi_p$ is ramified. We review the construction of the Galois representation attached to $\pi$. We define 
\[
H^*_i(X_{\mathbb{Q}_l}, V_\mu) = \varprojlim_{U'} H^*_i(X_{U'\mathbb{Q}_l}, V_\mu).
\]
In this case, it is known that $\pi$ is concentrated in degree 3 in $H^*_3(X_{\mathbb{Q}_l}, V_\mu)$. Then we define the Galois module $W_{\pi, l}$ by 
\[
W_{\pi, l} = \text{Hom}_{\text{GSp}_4(\mathbb{A}^f)}(\pi_f, H^*_3(X_{\mathbb{Q}_l}, V_\mu)).
\]
In fact by [Weis09, Theorem 1.1], we have

\[ W_{\pi,l} = \text{Hom}_{GSp_4(K)}(\pi_f, H^3(X_{\bar{Q}}, V_\mu)) \]

(3.13)

\[ = \text{Hom}_{GSp_4(K)}(\pi_f, H^3(X_{\bar{Q}}, V_\mu)) \]

\[ = \text{Hom}_{GSp_4(K)}(\pi_f, H^3(X_{\bar{Q}}, V_\mu)). \]

This will allow us pass from the usual cohomology \( H^3(X_{\bar{Q}}, V_\mu) \) or the cohomology with compact support \( H^3_c(X_{\bar{Q}}, V_\mu) \) to the interior cohomology \( H^3_! (X_{\bar{Q}}, V_\mu) \) after localizing at a suitable maximal ideal in the Hecke algebra.

The following theorem summarizes the properties of \( W_{\pi,l} \) and the proof of it can be found in [Tay93], [Laum05], [Weis09], [Sor10] and [Mo14].

**Theorem 3.14.** Let \( \pi \) be as above and \( \Theta_\pi : \mathbb{T}^S \to \mathbb{C} \) be the character giving the action of \( \mathbb{T}^S \) on \( \pi \). Then \( W_{\pi,l} \) gives rise to a Galois representation

\[ \rho_{\pi,l} : \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to GSp_4(\mathbb{Q}_l) \]

satisfying

- The representation \( \rho_{\pi,l} \) is unramified at a prime \( p \) for \( p \neq l \) and \( \pi_p \) is unramified. Moreover we have

\[ \det(1 - \rho_{\pi,l}(\text{Frob}_p)X) = \Theta_\pi(Q_p(X)) \]

- At all places \( p \neq l \) we have the local-global compatibility of Langlands correspondence, this means

\[ \text{WD}(\rho_{\pi,l}(G_{Q_p}))^{F-ss} \cong \text{rec}_p(\pi_p \otimes \cdot |^{-3/2}) \]

where \( \text{rec}_p \) is the local Langlands reciprocity map of [GT11] and \( F-ss \) denotes the Frobenius semi-simplification of the Weil-Deligne representation.

3.7. **Some inner forms of \( GSp_4 \).** Let \( B \) be a quaternion algebra over \( \mathbb{Q} \). If \( B \) split at \( \infty \), then we call \( B \) indefinite and otherwise we call \( B \) definite. For an element \( b \in B \), we denote by \( b \) the image of \( b \) under the main involution of \( B \). We choose an element \( \tau \in B \) such that \( \bar{\tau} = -\tau \) if \( B \) is indefinite and \( \tau = 1 \) if \( B \) is definite. We define a new involution \( \ast \) on \( B \) by putting \( b^\ast = \tau b \tau^{-1} \). Then consider the form \( (\cdot, \cdot) \) on \( V = B \oplus B \) defined by

\[ (x, y) = \text{tr}(\tau^{-1}(x_1 y_1^\ast + x_2 y_2^\ast)) \]

(3.15)

where \( x = (x_1, x_2) \) and \( y = (y_1, y_2) \) are elements in \( V \). Then we define the quaternionic unitary similitude group of degree 2 by

\[ \text{GU}_2(B)(\mathbb{Q}) = \{ g \in \text{GL}_2(B)(\mathbb{Q}) : (gx, gy) = c(g)(x, y), c(g) \in \mathbb{Q}^\times \} \]

We write \( D \) the quaternion division algebra over \( \mathbb{Q}_p \). If \( B_p = D \) at place \( p \), then \( \text{GU}_2(B)(\mathbb{Q}_p) = \text{GU}_2(D) \) where \( \text{GU}_2(D) \) is defined similarly as in (3.16) for \( V = D \oplus D \). If If \( B_p = M_2(\mathbb{Q}_p) \), then \( \text{GU}_2(B)(\mathbb{Q}_p) = \text{GSp}_4(\mathbb{Q}_p) \). Since \( D \) splits over \( \mathbb{Q}_p \), we also have an identification

\[ \text{GU}_2(D)(\mathbb{Q}_{p^2}) \cong \text{GSp}_4(\mathbb{Q}_{p^2}). \]

There are two kinds of parahoric subgroups of \( \text{GU}_2(D) \). This can be explained using the affine Dynkin diagram of \( \text{GU}_2(D) \). The affine Dynkin diagram is still of type \( \tilde{C}_2 \) but with Frobenius acting non-trivially on the diagram by switching the nodes 0 and 2 while fixing 1.

```
0 1 2
```

Therefore we have the following parahoric subgroups of $\text{GU}_2(D)$.

- The Iwahori subgroup $I'$ which corresponds to the Iwahori subgroup $I$ in $\text{GSp}_4$ when base change to $\mathbb{Z}_p^2$;
- The Siegel parahoric $H' = K'(02)$ which corresponds to the Siegel parahoric $K(02)$ in $\text{GSp}_4$ when base changed to $\mathbb{Z}_p^2$;
- The paramodular parahoric $K'(p) = K'_{(1)}$ which corresponds to the paramodular parahoric $K(p) = K_{(1)}$ in $\text{GSp}_4$ when base changed to $\mathbb{Z}_p^2$.

4. Siegel threefold of paramodular level

4.1. Integral models. Let $U^p$ be a fixed prime to $p$ open compact subgroup of $\text{GSp}_4(\mathbb{A}_f)$. Let $H = \text{GSp}_4(\mathbb{Z}_p)$ be the hyperspecial subgroup of $\text{GSp}_4(\mathbb{Q}_p)$ and recall that $K(p)$ is the paramodular subgroup. In this section, we will omit $U^p$ from all the notations. For example we have $X_H = X_{\text{HU}^p}$ and $X_{K_p} = X_{K(p)U^p}$. These Shimura varieties admits natural integral models $X_H$ and $X_{K(p)}$ over $\mathbb{Z}_p$ which are well documented in the literature and we will only recall their moduli interpretations.

- $X = X_H$ classifies triples of the form $(A, \lambda, \eta)$ up to isomorphism where $A$ is an abelian scheme of relative dimension $2$ over a test scheme $S$ over $\mathbb{Z}_p$ equipped with a principal polarization $\lambda$ and $U^p$-level structure $\eta$.
- $X_{K(p)}$ classifies triples of the form $(A, \lambda, \eta)$ up to isomorphism where $A$ is an abelian scheme of relative dimension $2$ and $\lambda$ is a polarization on $A$ such that $\ker(\lambda)[p]$ has rank $2$. Again $\eta$ is a prime to $p$ level structure.

Next we explain the structure of the supersingular locus and the singular locus of the special fiber of $X_{K(p)}$. First we recall a result of Yu [Yu11a], let $X_{K(p), \mathbb{F}_p}$ be the special fiber of $X_{K(p)}$ base changed to $\mathbb{F}_p$ where $\mathbb{F}_p$ is an algebraically closed field containing $\mathbb{F}_p$.

**Theorem 4.1.** The scheme $X_{K(p), \mathbb{F}_p}$ has isolated quadratic singularities. The singular locus $\Sigma_p(K(p))$ consists of those $(A, \lambda, \mu)$ such that $\ker(\lambda) \subset A[p]$ and $A$ is superspecial.

**Proof.** The first assertion is proved by a local model computation in [Yu11a]. For the second assertion, the singular locus consists of those $(A, \lambda, \mu)$ such that $\ker(\lambda) \cong \alpha_p \times \alpha_p$ by [Yu11a, Theorem 4.7]. Thus $\ker(\lambda) \subset A[p]$ and moreover it implies that $A$ is superspecial. □

Next we recall the description of the supersingular locus $S_{K(p)}$ of $X_{K(p), \mathbb{F}_p}$ using the Bruhat-Tits stratification in [Wang19a]. We fix a $p$-divisible group $X$ of dimension $2$ which is isoclinic of slope $1/2$ equipped with a polarization of height $2$. We denote by $N$ its associated isocrystal. The polarization induces an alternating form $(\cdot, \cdot)$ on $N$. Let $b \in B(\text{GSp}_4)$ be the element corresponds to $N$ in the Kottwitz set $B(\text{GSp}_4)$ of $\text{GSp}_4$. Let $J_b$ be the group functor over $\mathbb{Q}_p$ defined by assigning each $\mathbb{Q}_p$-algebra $R$ the group

$$J_b(R) = \{g \in \text{GSp}_4(R \otimes \mathbb{Q}_p, K_0) : g(b \sigma) = (b \sigma) g\}.$$  

It is well known that in this case $J_b \cong \text{GU}_2(D)$ in our case. In fact if we denote by $C = N^{\tau=1}$ with $\tau = F^{-1}V$, then $C$ can be viewed as an $D$-module equipped with an alternating form $(\cdot, \cdot)$ which is the restriction of $(\cdot, \cdot)$ on $N$. Then we have $J_b \cong \text{GU}_2(C)$. We consider the Rapoport-Zink space $N_{K(p)}$ for $\text{GSp}_4(\mathbb{Q}_p)$ with paramodular level structure. Let $W_0 = W(\mathbb{F}_p)$ and denote by $(\text{Nilp})$ the category of $W_0$-schemes $S$ on which $p$ is locally nilpotent. This is the set valued functor on $(\text{Nilp})$ that classifies the data $(X, \lambda_X, \rho_X)$ where

- $X$ is a two dimensional $p$-divisible group over $S$ of height $4$;
• \( \lambda_X : X \to X^\vee \) is a polarization of height 2;
• \( \rho_X : X \times_{S_p} \bar{S} \to \bar{X} \times_S \bar{S} \) is a quasi-isogeny.

Here \( \bar{S} \) is the special fiber of \( S \) at \( p \). This functor is representable by a formal scheme \( \mathcal{N}_{K(p)} \) over \( \text{Spf}(W_0) \) and \( J_0(\mathbb{Q}_p) \) acts on \( \mathcal{N}_{K(p)} \) naturally. The formal scheme \( \mathcal{N}_{K(p)} \) can be decomposed into connected components

\[
\mathcal{N}_{K(p)} = \bigsqcup_{i \in \mathbb{Z}} \mathcal{N}_{K(p)}^{(i)}
\]

where \( \mathcal{N}_{K(p)}^{(i)} \) classifies those \((X, \lambda_X, \rho_X)\) such that \( \rho_X \) has height \( i \). We denote by \( \mathcal{M}_{K(p)} \) the underlying reduced scheme of \( \mathcal{N}_{K(p)}^{(0)} \).

We recall the notion of vertex lattices in \( C \). A vertex lattice \( L \) is by definition a \( \mathbb{Z}_p \)-lattice \( L \subset C \) such that \( pL^\perp \subset L \subset L^\perp \) where \( L^\perp \) is the integral dual of \( L \) in \( C \). They are naturally divided into three kinds according to the indices of the inclusions:

• \( L \) is a vertex lattice of type 0 if \( pL^\perp \subset L^0 \subset L^\perp \);
• \( L \) is a vertex lattice of type 1 if \( pL^\perp \subset L^1 \subset L^\perp \);
• \( L \) is a vertex lattice of type 2 if \( pL^\perp \subset L^2 \subset L^\perp \).

We also define a vertex lattice of type 02 by:

• A pair \((L_0, L_2)\) is a vertex lattice of type 02 if \( L_0 \) is a vertex lattice of type 0 and \( L_2 \) is a vertex lattice of type 2.

The vertex lattices of type 02 and type 1 are called vertex lattices of \( \text{GU}_2(C) \).

**Remark 4.4.** Each vertex lattice \( L_i \) for \( i = 0, 1, 2 \) gives rise to a maximal parahoric subgroup in \( \text{GSp}_4(\mathbb{Q}_p) \) by taking the stabilizer. The vertex lattices of type 0 and 2 correspond to the hyperspecial parahorics \( H \) and the vertex lattice of type 1 corresponds to the paramodular parahoric \( K(p) \).

On the other hand, vertex lattices of type 02 and type 1 can be linked to the maximal parahoric subgroups of \( \text{GU}_2(D) \). Here vertex lattices of type 02 will correspond to the Siegel parahoric \( H' \) and vertex lattice of type 1 will correspond to the paramodular parahoric \( K'(p) \).

For each vertex lattice \( L_{02} = (L_0, L_2) \) of type 02 and each vertex lattice \( L_1 \) of type 1, we have defined in [Wang19a] projective subschemes \( \mathcal{M}_{K(p), L_{02}} \) and \( \mathcal{M}_{K(p), L_1} \) of \( \mathcal{M}_{K(p)} \). In fact, each \( \mathcal{M}_{K(p), L_{02}} \) is isomorphic to a projective line \( \mathbb{P}^1 \) and \( \mathcal{M}_{K(p), L_1} \) is isomorphic to a reduced point which can be defined over \( \mathbb{F}_p^2 \). We will refer to these projective schemes as the lattice strata. We also define the open lattice strata by

\[
\begin{align*}
\mathcal{M}_{K(p), L_{02}}^0 & = \mathcal{M}_{L_{02}} - \bigcup L_1 \mathcal{M}_{L_1}, \\
\mathcal{M}_{K(p), L_1}^0 & = \mathcal{M}_{L_1} - \bigcup L_0 \mathcal{M}_{L_{02}},
\end{align*}
\]

where the index \( L_1 \) runs through all the vertex lattices of type 1. This allows us to define the Bruhat-Tits strata of \( \mathcal{M}_{K(p)} \) by

\[
\begin{align*}
\mathcal{M}_{K(p), \{02\}}^0 & = \bigcup_{L_{02}} \mathcal{M}_{K(p), L_{02}}^0, \\
\mathcal{M}_{K(p), \{1\}}^0 & = \bigcup_{L_1} \mathcal{M}_{K(p), L_1}^0.
\end{align*}
\]

Then by [Wang19a] Theorem 7.3] we have the following description of the scheme \( \mathcal{M}_{K(p)} \).

**Proposition 4.5.** The scheme \( \mathcal{M}_{K(p)} \) admits the Bruhat-Tits stratification

\[
\mathcal{M}_{K(p)} = \mathcal{M}_{K(p), \{02\}}^0 \cup \mathcal{M}_{K(p), \{1\}}^0.
\]
The irreducible components of $\mathcal{M}_{K(p)}$ are given by the lattice strata of the form $\mathcal{M}_{K(p), L_02}$ for a vertex lattice $L_{02}$ of type 02. The singular locus of $\mathcal{M}_{K(p)}$ is given by $\mathcal{M}_{K(p), 1}$. 

The Rapoport-Zink uniformization theorem \cite[Theorem 6.30]{RZ96} furnishes an isomorphism

$$S_{K(p)} \cong \mathbb{I}(\mathbb{Q}) \backslash \mathbb{N}_{K(p)} \times \text{GSp}_4(\mathbb{A}_f^p)/U_p.$$ 

Here $\mathbb{I}(\mathbb{Q}) \cong \text{GU}_{B'}(\mathbb{Q})$ and $B' = B_{p\infty}$ is the quaternion algebra over $\mathbb{Q}$ which ramifies at $p$ and infinity.

**Proposition 4.7.**

- The irreducible components of $S_{K(p)}$ can be parametrized by the double coset
  
  $$(4.8) \quad \text{GU}_{B'}(\mathbb{Q}) \backslash \text{GU}_{B'}(\mathbb{A}_f)/H' U_p.$$ 

- The singular locus of $S_{K(p)}$ can be parametrized by the double coset
  
  $$(4.9) \quad \text{GU}_{B'}(\mathbb{Q}) \backslash \text{GU}_{B'}(\mathbb{A}_f)/K(p) U_p.$$ 

**Proof.** By Proposition 4.5, the irreducible components of the scheme $\mathcal{M}_{K(p)}$ corresponds to vertex lattices of type 02. The derived group of $\text{GU}_2(D)$ acts transitively on these lattices and the stabilizer corresponds to the Siegel parahoric $H'$. Therefore the irreducible components of $\mathcal{N}_{K(p)}$ is parametrized by $\text{GU}_2(D)/H'$. Now we apply (4.6) to conclude the proof of (4.8).

By Proposition 4.5, the singular locus of the scheme $\mathcal{M}_{K(p)}$ corresponds to vertex lattices of type 1. Then we proceed similarly as above to prove (4.9). \hfill \Box

## 5. Mazur’s Principle for Paramodular Siegel Modular Forms

**5.1. Mazur’s principle.** We recall that we are concerned with a cuspidal automorphic representation $\pi$ for $\text{GSp}_4$ which is non-CAP and non-endoscopic whose component at infinity $\pi_\infty$ is a holomorphic discrete series of Harish Chandra parameter $(a, b, -a - b + 3)$ with $a \geq b \geq 0$. We also assume that the representation $\pi$ is para-spherical and its local component $\pi_p$ at $p$ is of type IIa under Schmidt’s classification. In this case, its associated Galois representation $\rho_{\pi, l}$ is realized in the Galois module $W_{\pi, l}$ introduced in (3.12). The integral cohomology $H^3_!(\bar{X}_{\mathbb{Q}^p}, \mathbb{V}_{\mu}^\text{int})$ gives a Galois stable lattice $\Lambda_{\pi, l} \subset W_{\pi, l}$ such the natural reduction of $\Lambda_{\pi, l}$ gives rise to a representation of $\text{Gal}(\bar{k}/k)$ valued in $k$ for some finite field of characteristic $l$

$$\tilde{\rho}_{\pi, l} : \text{Gal}(\bar{k}/k) \rightarrow \text{GSp}_4(k)$$

which we will refer to as the residual representation of $\rho_{\pi, l}$. We will make the following assumption throughout the rest of this article that

(Irr) \quad $\tilde{\rho}_{\pi, l}$ is irreducible.

The residual Galois representation $\tilde{\rho}_{\pi, l}$ corresponds to a maximal ideal $m \subset \mathfrak{p}_p$ with residue field $k$. We make the following assumption

(Semi) \quad $H^3_!(\bar{X}_{K(p), \mathbb{Q}}, \mathbb{V}_{\mu}^\text{int})_m \otimes k$ is semisimple as a Galois module.

This implies that

$$H^3_!(\bar{X}_{K(p), \mathbb{Q}}, \mathbb{V}_{\mu}^\text{int})_m \otimes k = \tilde{\rho}^{\otimes s}$$

for some positive number $s$.

In this section, we will prove the following theorem which we will refer to as the Mazur’s principle.
Theorem 5.1. Let $\pi$ be a cuspidal automorphic representation of $\text{GSp}_4$ that is cohomological and is non-CAP and non-endoscopic whose component at infinity $\pi_{\infty}$ is a holomorphic discrete series of Harish Chandra parameter $(a, b, -a - b + 3)$ with $a \geq b \geq 0$ and $l > a + b + 4$. Let $p$ be a prime distinct from $l$ and such that $p \not\equiv 1 \mod l$. Let $U = K(p)U^p \subset \text{GSp}_4(\mathbb{Q}_p)\text{GSp}_4(\mathbb{A}_f^p)$ be a neat open compact subgroup such that $\pi^U \neq 0$. We assume that $\pi_p$ is ramified of type IIa in the classification of Schmidt. Suppose the residual Galois representation $\rho_{\pi,l}$ satisfies the following assumptions:

- $\bar{\rho}_{\pi,l}$ is unramified at $p$;
- $\rho_{\pi,l}$ satisfies $(\text{Irr})$;
- $\rho_{\pi,l}$ satisfies $(\text{Semi})$;

Then there exists a cuspidal automorphic representation $\pi'$ with the same type as $\pi$ at $\infty$ such that $\bar{\rho}_{\pi,l} \equiv \bar{\rho}_{\pi',l}$ and $\pi'_p$ is unramified.

5.2. Picard-Lefschetz formula. We consider $X = X_{K(p)}$ in the setting of $\mathbb{Q}$ Here $\bar{\eta} = \text{Spec}(\mathbb{Q}_p)$ and $\bar{s} = \text{Spec}(\overline{\mathbb{F}}_p)$, we will set $X_{K(p),\bar{\eta}_p} = X_{K(p),\bar{\eta}}$ $X_{K(p),\bar{\Sigma}_p} = X_{K(p),\bar{s}}$ and $\Sigma = \Sigma_p(K(p))$. Recall also that $S_{K(p)}$ is the supersingular locus of $X_{K(p),\bar{\Sigma}_p}$. Furthermore, we will let $V = V^{\text{int}}_{\mu}$ to further light up the notation. In this subsection, we are mainly concerned with analyzing the following diagram which we refer to as the Picard-Lefschetz formula for $X_{K(p),\bar{\Sigma}_p}$

$$
\begin{align*}
H^3(X_{K(p),\bar{\Sigma}_p}, V)(1) &\xrightarrow{\alpha} \bigoplus_{x \in \Sigma} R^3\Phi(V)_x(1) \\
&\quad \downarrow N \\
H^3(X_{K(p),\bar{\Sigma}_p}, V) &\xleftarrow{\beta} \bigoplus_{x \in \Sigma} H^3_x(X_{K(p),\bar{\Sigma}_p}, V).
\end{align*}
$$

By general theory, the map $\beta$ and $\alpha$ are dual to each other and therefore $\beta$ is injective if and only if $\alpha$ is surjective. We prove first the following important proposition

Proposition 5.3. The map $\beta$ localized at $m$ is injective

$$
\beta : \bigoplus_{x \in \Sigma} H^3_x(X_{K(p),\bar{\Sigma}_p}, V)_m \to H^3_c(X_{K(p),\bar{\Sigma}_p}, V)_m
$$

and therefore

$$
\alpha : H^3_c(X_{K(p),\bar{\Sigma}_p}, V)_m(1) \to \bigoplus_{x \in \Sigma} R^3\Phi(V)_x,m(1)
$$

is surjective.

Proof. Since $H^3_c(X_{K(p),\bar{\Sigma}_p}, V) \to H^3_c(X_{K(p),\bar{\Sigma}_p}, V)$ is injective, we need to show the Gysin map

$$
\bigoplus_{x \in \Sigma} H^3_x(X_{K(p),\bar{\Sigma}_p}, V) \to H^3_c(X_{K(p),\bar{\Sigma}_p}, V)
$$

is injective after localizing at $m$. Note the kernel of this morphism is the image of the connecting homomorphism

$$
H^2_c(U, V) \to \bigoplus_{x \in \Sigma} H^2_x(X_{K(p),\bar{\Sigma}_p}, V)
$$

where $U = X_{K(p),\bar{\Sigma}_p} - \Sigma$. We claim that $H^2_c(U, V) = 0$ after localizing at $m$. To prove this, we use the following excision exact sequence

$$
H^2_{S_{K(p)}}(X_{K(p),\bar{\Sigma}_p} - \Sigma, V) \to H^2_c(U, V) \to H^2_c(X_{K(p),\bar{\Sigma}_p} - S_{K(p)}, V).
$$
Note that $X_{K(p),\bar{\mathcal{E}}_p} - \Sigma$ is now smooth and $S_{K(p)} - \Sigma$ is of codimension 2 in $X_{K(p),\bar{\mathcal{E}}_p} - \Sigma$. Then by purity, $H^2_{S_{K(p)} - \Sigma}(X_{K(p),\bar{\mathcal{E}}_p} - \Sigma, \mathcal{V})$ vanishes. Also notice that $X_{K(p),\bar{\mathcal{E}}_p} - S_{K(p)}$ is the ordinary locus. Let $X'_{K(p),\bar{\mathcal{E}}_p}$ be the minimal compactification of $X_{K(p),\bar{\mathcal{E}}_p}$, then $X'_{K(p),\bar{\mathcal{E}}_p} - S_{K(p)}$ is affine and therefore

$$H^2_c(X'_{K(p),\bar{\mathcal{E}}_p} - S_{K(p)}, \mathcal{V}) = 0$$

by Artin vanishing. Since the difference between

$$H^2_c(X_{K(p),\bar{\mathcal{E}}_p} - S_{K(p)}, \mathcal{V}) = 0$$

and

$$H^2_c(X'_{K(p),\bar{\mathcal{E}}_p} - S_{K(p)}, \mathcal{V}) = 0$$

is supported on the boundary of the minimal compactification, the difference vanishes after localizing at $\mathfrak{m}$ by our assumption (Irr). Therefore $H^2_c(X_{K(p),\bar{\mathcal{E}}_p} - S_{K(p)}, \mathcal{V})$ vanishes after localizing at $\mathfrak{m}$. □

By Proposition 5.3 we have the following fundamental diagram

$$\bigoplus_{x \in \Sigma} H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})_m \otimes E \xrightarrow{\alpha} H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})_m \otimes E \longrightarrow H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})_m \otimes E \longrightarrow \bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m} \otimes E.$$

We have

$$H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})_m \otimes E = \bigoplus_{\pi'} \otimes m(\pi')$$

where $\pi'$ runs through all the automorphic representation $\pi'$ whose Galois representation $\rho_{\pi',l}$ is congruent to $\rho_{\pi,l}$ and $m(\pi')$ is the multiplicity of $\pi'$ which could be zero. Let $\pi'$ be an automorphic representation such that $\pi'_p$ is ramified. Then we have the following description of $\rho_{\pi',l}$ in terms of the fundamental diagram (5.6).

**Proposition 5.8.** Let $\pi'$ be a representation appearing in (5.6) such that $\pi'_p$ is ramified. Then

- $\bigoplus_{x \in \Sigma} H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})_m \otimes E$ contributes to $\rho_{\pi',l|G_{\mathbb{Q}_p}}$ as a 1-dimensional subspace;
- $\bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m} \otimes E$ contributes to $\rho_{\pi',l|G_{\mathbb{Q}_p}}$ as a 1-dimensional quotient;
- $H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})_m \otimes E/\text{im}(\beta)$ contributes to $\rho_{\pi',l|G_{\mathbb{Q}_p}}$ as a 2-dimensional subspace.

**Proof.** The local-global compatibility of Langlands correspondence as in Theorem 3.14 implies that the monodromy operator $N$ has one dimensional image. Thus the Picard-Lefschetz formula in (7.19) implies that $\bigoplus_{x \in \Sigma} H^3_c(X_{K(p),\bar{\mathcal{E}}_p}, \mathcal{V})$ has one dimensional contribution to $\rho_{\pi',l}$. Since the local monodromy map $\beta_x$ are isomorphisms, $\bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m} \otimes E$ contributes to $\rho_{\pi',l|G_{\mathbb{Q}_p}}$ as a 1-dimensional quotient. Then the third claim is also clear from what we have proved. □

**5.3. Proof of the Mazur’s principle.** Now we come to the proof of Theorem 5.1. We will prove the theorem by contradiction. Assume that we can not find such a $\pi'$ in the statement of the theorem. This means that all the $\pi'$ appears in (5.7) are ramified at $p$. We first prove a congruence relation on $\bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m}$ under this assumption.

**Lemma 5.9.** Suppose that all the $\pi'$ appears in (5.7) are ramified at $p$. The Frobenius action $\text{Frob}_p$ on the Galois module

$$\bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m}$$

is given by

$$\bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m} \otimes E \longrightarrow \bigoplus_{x \in \Sigma} R^3\Phi(\mathcal{V})_{x,m} \otimes E.$$
is given by the Atkin-Lehner operator $u_p$.

**Proof.** Examine the Weil-Deligne representation corresponding to $\rho_{\pi',l}$ in (5.10) and the shape of the monodromy operator (3.5). Let $x \in \Sigma$ such that $H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes E$ contributes to the Galois representation $\rho_{\pi',l|G_{\infty p}}$ as the character $\chi \tau \cdot \cdot$. Since the local monodromy operator $N_x$ induces an isomorphism between $\bigoplus_{x \in \Sigma} R^3\Phi(V)(1) \otimes E$ and $\bigoplus_{x \in \Sigma} H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes E$. It follows that Then the result follows from Lemma 5.6. □

We continue with the proof of the Mazur’s principle. The reduction modulo $l$ of (5.7) gives

$$\text{(5.10)} \quad H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes k = \bar{\rho}_{\pi,l|G_{\infty p}}^*$$

which is unramified by our assumption. Therefore the monodromy operator $N \otimes k$ degenerates to 0.

We calculate this monodromy operator using again the Picard-Lefschetz formula

$$\text{(5.11)} \quad H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes k \xrightarrow{\alpha \otimes k} \bigoplus_{x \in \Sigma} R^3\Phi(V)(1) \otimes k \xrightarrow{\beta \otimes k} \bigoplus_{x \in \Sigma} H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes k.$$

By Proposition (5.3), $\alpha \otimes k$ is surjective and $N_x \otimes k$ is an isomorphism. Therefore $\beta \otimes k$ has to be 0.

Now we examine the mod $l$ version of the fundamental diagram (5.6)

$$\text{(5.12)} \quad \bigoplus_{x \in \Sigma} H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes k \xrightarrow{\gamma} H^3(X_{K(p)},\mathcal{F}_p,\mathbb{V})_m \otimes k \xrightarrow{\alpha \otimes k} \bigoplus_{x \in \Sigma} R^3\Phi(V)(1) \otimes k.$$

The map $\gamma$ here is still injective. This can be proved by snake lemma and the fact that $\bigoplus_{x \in \Sigma} R^3\Phi(V)(1) \otimes k$ is obviously torsion free. Therefore $\bigoplus_{x \in \Sigma} R^3\Phi(V)(1) \otimes k$ contributes to $\bar{\rho}_{\pi,l|G_{\infty p}}^*$ as the two dimensional quotient which is isomorphic to the reduction modulo $l$ of the twisted Steinberg $\chi St_{GL_2}$ part of Weil-Deligne representation in (3.4). In particular the determinant of the element $\text{Frob}_p$ on this two dimensional quotient is of the form $\chi_2^2 \sigma^2 \cdot \cdot$ modulo $l$. On the other hand, by the congruence relation in Lemma 5.9 $\text{Frob}_p$ acts by the character $\chi \sigma$ modulo $l$ on this two dimensional quotient. This implies that its determinant is of the form $\chi_2^2 \sigma^2$. It then follows that $p \equiv 1 \mod l$ which is a contradiction. This finishes the proof of Theorem 5.1.

6. Quaternionic Siegel threefold of paramodular level

6.1. Integral model. In this section we review some results obtained in [Wang19a] about the supersingular locus of the quaternionic Siegel threefold. In light of the forthcoming application, we let $B = B_{pq}$ be the indefinite quaternion algebra with discriminant $pq$ over $\mathbb{Q}$. Here $p$ and $q$ are two distinct primes that are distinct from $l$. Let $*$ be a $\text{neben-involution}$ on $B$ defined as in [KR00, A.4] and $O_B$ be a maximal order fixed by $*$. We set $V = B \oplus B$ and let

$$(\cdot, \cdot) : V \times V \to \mathbb{Q}$$

be the alternating form defined in (3.1.15). This form satisfies the following equation

$$(av_1, v_2) = (v_1, a^*v_2)$$

for all $v_1, v_2 \in V$ and $a \in B$. Let $G = GU_2(B)$ be the quaternionic unitary group of degree 2 defined as in (3.10). Since $B$ splits over $\mathbb{R}$. We have the identification $G(\mathbb{R}) \cong \text{GSp}_4(\mathbb{R})$. Let $\mathbb{S} = \text{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_m$
be the Deligne torus and let $h : S \to G_{\mathbb{R}}$ be the Hodge cocharacter that it induces the minuscule cocharacter
\[
\mu : \mathbb{G}_{m, \mathbb{C}} \to G_{\mathbb{C}} \cong \text{GSp}(4)_{\mathbb{C}}
\]
that sends $z \in \mathbb{G}_{m, \mathbb{C}}$ to $\text{diag}(z, z, 1, 1) \in \text{GSp}(4)_{\mathbb{C}}$. Let $U^{pq} \subset \text{GSp}_{4}(A_{f}^{pq})$ be a sufficiently small subgroup which we will fix throughout the rest of this article. Let $L_{p} = \mathcal{O}_{B_{p}} \oplus \mathcal{O}_{B_{q}}$ and let $L_{q} = \mathcal{O}_{B_{q}} \oplus \mathcal{O}_{B_{q}}$. Then the stabilizer of this lattice in $G(\mathbb{Q}_{p})$ and $G(\mathbb{Q}_{q})$ are the paramodular subgroups denoted by $K'(p)$ and $K'(q)$ in [3.7]. Then the tuple
\[
(B, \mathcal{O}_{B}, V, (\cdot, \cdot), \mu, L_{p}, U^{p} = K'(q)U^{pq})
\]
gives a PEL-data over $\mathbb{Z}_{p}$ and the tuple
\[
(B, \mathcal{O}_{B}, V, (\cdot, \cdot), \mu, L_{q}, U^{q} = K'(p)U^{pq})
\]
gives a PEL-data over $\mathbb{Z}_{q}$. They define PEL-type moduli problems $X_{K'(p)}^{B}$ over $\mathbb{Z}_{p}$ and $X_{K'(q)}^{B}$ over $\mathbb{Z}_{q}$. As the role of $p$ and $q$ are completely symmetric in this section, we will only state the moduli problem $X_{K'(p)}^{B}$ over $\mathbb{Z}_{p}$. For each $\mathbb{Z}_{p}$ scheme $S$, a point of the functor $X_{K'(p)}^{B}$ with value in $S$ is given by the following data up to isomorphism.

- $A$ is an abelian scheme over $S$ of dimension $4$;
- $\iota : \mathcal{O}_{B} \to \text{End}_{S}(A)$ is a ring homomorphism;
- $\lambda : A \to A^{\vee}$ is a prime to $p$ polarization;
- $\eta : H_{1}(A, k_{p}^{\vee}) \cong V \otimes k_{p}^{\vee}$ mod $U^{p}$ is a $U^{p}$-level structure that respects the bilinear forms on both sides up to a constant in $(k_{p}^{\vee})^{\times}$.

The quadruple $(A, \iota, \lambda, \eta)$ is required to satisfy the following additional conditions:

- The Kottwitz condition
\[
\det(T - \iota(a); \text{Lie}(A)) = (T^{2} - \text{Trd}(a)T + \text{Nrd}(a))^{2}
\]
for all $a \in \mathcal{O}_{B}$;

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

This functor is representable by a quasi-projective scheme $X_{K'(p)}^{B}$ over $\mathbb{Z}_{p}$ of relative dimension $3$. We remark that $\text{GU}_{2}(B)$ has semi-simple $\mathbb{Q}$-rank one and therefore $X_{K'(p)}^{B}$ is not proper. The boundary of the minimal compactification of $X_{K'(p)}^{B}$ consists of cusps. The scheme $X_{K'(p)}^{B}$ is regular but not smooth over $\text{Spec}(\mathbb{Z}_{p})$. Since the group $G$ can be identified with $\text{GSp}_{4}$ over $\mathbb{Z}_{p^{2}}$ and the parahoric subgroup $K'(p)$ can be identified with $K$ over the same extension, the local model for $X_{K'(p)}^{B}$ agrees with that of $X_{K'(p)}$ after the base change to $\mathbb{Z}_{p^{2}}$. Therefore $X_{K'(p)}^{B}$ also has isolated singularities which are ordinary and quadratic. We denote by $X_{K'(p)}^{B, \text{qs}}$ the generic fiber of $X_{K'(p)}^{B}$ base changed to $\overline{\mathbb{Q}}_{p}$.

The set of singular points $\Sigma_{p}(B)$ again agree with the superspecial points on $X_{K'(p)}^{B, \text{qs}}$.

**Lemma 6.4.** The singular locus $\Sigma_{p}(B)$ consists of those $(A, \iota, \lambda, \eta) \in X_{K'(p)}^{B, \text{qs}}$ such that $A$ is superspecial.

**Proof.** This follows from [Wang19a, Lemma 2.1] and the discussion on local model following it. \(\square\)

Next we recall the description of the supersingular locus of $X_{K'(p)}^{B, \text{qs}}$ using the Bruhat-Tits stratification in [Wang19a]. Let $S_{K'(p)}$ be the supersingular locus considered as a reduced closed subscheme of $X_{K'(p)}^{B, \text{qs}}$. We fix a $p$-divisible group $\mathcal{X}$ over $\mathbb{F}_{p}$ with an action $\iota : \mathcal{O}_{\mathcal{D}} \to \text{End}(\mathcal{X})$ and a polarization $\lambda : \mathcal{X} \to \mathcal{X}^{\vee}$. We assume that $\mathcal{X}$ is isoclinic of slope $\frac{1}{2}$ and we denote by $N$ its associated isocrystal. Let $b \in B(G)$ be the element associated to $N$ in the Kottwitz set $B(G)$. Then the group $J_{b}$ defined as
in \([4,2]\) can be identified with the split \(\text{GSp}_4/\mathbb{Q}_p\) in this case. In fact let \(C = N^{\tau=1}\) with \(\tau = \Pi^{-1}F\) equipped with the restriction of the alternating form \((\cdot, \cdot)\) to \(C\). Then we have \(J_b = \text{GSp}_4(C)\). We consider the Rapoport-Zink space \(\mathcal{N}_{K'(p)}\) for the group \(G\). This is the set valued functor on \((\text{Nilp})\) which assigns each \(S \in (\text{Nilp})\) the set of isomorphism classes of quadruples \((X, \iota_X, \lambda_X, \rho_X)\) where

\[
\mathcal{N}_{K'(p)}(S) = \bigsqcup_{i \in \mathbb{Z}} \mathcal{N}_{K'(p)}^{(i)}
\]

where \(\mathcal{N}_{K'(p)}^{(i)}\) classifies those \((X, \iota_X, \lambda_X, \rho_X)\) such that \(\rho_X\) has height \(i\). We denote by \(\mathcal{M}_{K'(p)}\) the reduced scheme of the formal scheme \(\mathcal{N}_{K'(p)}^{(0)}\). Then we define vertex lattices in \(C\) similarly as before. Recall that they are naturally divided into three kinds

- \(L\) is a vertex lattice of type 0 if \(pL \perp_c 1 L \subset 0 L \perp_c 1\);
- \(L\) is a vertex lattice of type 1 if \(pL \perp_c 2 L \subset 4 L \perp_c 1\);
- \(L\) is a vertex lattice of type 2 if \(pL \perp_c 3 L \subset 4 L \perp_c 3\).

Each vertex lattice gives rise to a lattice stratum which is a projective sub scheme of \(\mathcal{M}_{K'(p)}\), see [Wang19a] Theorem 5.1.

- For a vertex lattice \(L_0\) of type 0, the scheme \(\mathcal{M}_{K'(p), L_0}\) is a projective surface of the form
\[
x_3^p x_0 - x_0^p x_3 + x_2^p x_1 - x_1^p x_2 = 0
\]
for a projective coordinate \([x_0 : x_1 : x_2 : x_3]\) of \(\mathbb{P}^3\);

- For a vertex lattice \(L_2\) of type 2, the scheme \(\mathcal{M}_{K'(p), L_2}\) is a projective surface of the form
\[
x_3^p x_0 - x_0^p x_3 + x_2^p x_1 - x_1^p x_2 = 0
\]
for a projective coordinate \([x_0 : x_1 : x_2 : x_3]\) of \(\mathbb{P}^3\);

- For a vertex lattice \(L_1\) of type 1, the scheme \(\mathcal{M}_{K'(p), L_1}\) consists of a superspecial point.

There are two types of irreducible components of \(\mathcal{M}_{K'(p)}\). They are given by lattice strata \(\mathcal{M}_{K'(p), L_0}\) of type 0 and lattice strata \(\mathcal{M}_{K'(p), L_2}\) of type 2. The Rapoport-Zink uniformization theorem [RZ96] Theorem 6.30] then yields

\[
S_{K'(p)} \cong I(\mathbb{Q}) \backslash \mathcal{N}_{K'(p)} \times \text{GU}_2(B)((\mathbb{A}_f)^p)/U^p.
\]

Here \(I(\mathbb{Q}) \cong \text{GU}_{B''}(\mathbb{Q})\) and \(B'' = B_{\text{qu}}\) is the quaternion algebra over \(\mathbb{Q}\) which ramifies at \(q\) and infinity.

**Proposition 6.7.**

- The irreducible components of \(S_{K'(p)}\) can be parametrized by the union of two double cosets
\[
\text{GU}_2(B'')(\mathbb{Q}) \backslash \text{GU}_2(B'')(\mathbb{A}_f)/K[0]U^p \bigsqcup \text{GU}_2(B'')(\mathbb{Q}) \backslash \text{GU}_2(B'')((\mathbb{A}_f)^p)/K(2)U^p.
\]

- The singular locus \(\Sigma_p(B)\) can be parametrized by the following double coset
\[
\text{GU}_2(B'')(\mathbb{Q}) \backslash \text{GU}_2(B'')(\mathbb{A}_f)/K(p)U^p.
\]
Proof. Recall there are two kinds of irreducible components of \( \mathcal{M}_{K(p)} \) and they correspond to vertex lattices of type 0 and vertex lattice of type 2. Therefore by (6.6) the irreducible components of type 0 can be parametrized by

\[
I(\mathbb{Q})\backslash J_6(\mathbb{Q}_p)/K(\mathfrak{q}) \times GU_2(B)(A_f^{(p)})/U^p.
\]

One then sees that this is exactly

\[
GU_2(B'')(\mathbb{Q})\backslash GU_2(B''')(A_f)/K(0)^p.
\]

Similarly, the irreducible components of type 2 can be parametrized by

\[
I(\mathbb{Q})\backslash J_6(\mathbb{Q}_p)/K(\mathfrak{q}) \times GU_2(B)(A_f^{(p)})/U^p.
\]

One then sees that this is exactly

\[
GU_2(B''')(\mathbb{Q})\backslash GU_2(B''')(A_f)/K(0)^p.
\]

By Lemma 6.4, the singular locus is exactly the superspecial locus which correspond to vertex lattices of type 1. Therefore by (6.6), it is parametrized by the double coset

\[
I(\mathbb{Q})\backslash J_6(\mathbb{Q}_p)/K(p) \times GU_2(B)(A_f^{(p)})/U^p.
\]

One then sees that this is exactly

\[
GU_2(B''')(\mathbb{Q})\backslash GU_2(B''')(A_f)/K(p)^p.
\]

Recall our notations for the quaternion algebras: \( B = B_{pq}, B' = B_{p\infty} \) and \( B'' = B_{q\infty} \). The following observation will be important in the subsequent discussions.

Corollary 6.14. The singular locus \( \Sigma_p(K(pq)) \) of \( X_{K(pq),\tilde{\mathfrak{q}}_p} \) is parametrized by the same set as the singular locus \( \Sigma_q(B) \) of \( X_{K^{(q)},\tilde{\mathfrak{q}}_q} \). More precisely, we have

\[
\Sigma_p(K(pq)) = GU_2(B')(\mathbb{Q})\backslash GU_2(B')(A_f)/K'(p)K(q)U^{pq} = \Sigma_q(B).
\]

Symmetrically, The singular locus \( \Sigma_q(K(pq)) \) of \( X_{K(pq),\tilde{\mathfrak{q}}_q} \) is parametrized by the same set as the singular locus \( \Sigma_p(B) \) of \( X_{K^{(p)},\tilde{\mathfrak{q}}_p} \). More precisely, we have

\[
\Sigma_q(K(pq)) = GU_2(B'')(\mathbb{Q})\backslash GU_2(B'')(A_f)/K(p)K'(q)U^{pq} = \Sigma_p(B).
\]

Proof. This follows easily from Proposition 4.7 and Proposition 6.7.

\[\square\]

7. Ribet’s principle for paramodular Siegel modular forms

7.1. Ribet’s principle. Now we come to state the second main theorem of this article which we call Ribet’s principle for the paramodular Siegel modular forms. As in the proof of the Mazur’s principle we will make the following semi-simplicity assumption

\[
\text{(Semi-pq)} \quad H^3(X_{K(pq),\tilde{\mathfrak{q}}_p,\mathfrak{V}^{\text{int}}}_t)(m) \otimes k \text{ is semisimple as a Galois module.}
\]

Theorem 7.1. Let \( \pi \) be a cuspidal automorphic representation of \( \text{GSp}_4 \) that is cohomological which is non-CAP and non-endoscopic whose component at infinity \( \pi_\infty \) is a holomorphic discrete series of Harish Chandra parameter \( (a, b, -a - b + 3) \) with \( a \geq b \geq 0 \) such that \( l \geq a + b + 4 \). Let \( p, q \) be two distinct primes different from \( l \). Let \( U = K(p)K(q)U^{pq} \subset \text{GSp}_4(\mathbb{Q}_p)\text{GSp}_4(\mathbb{Q}_q)\text{GSp}_4(A_f^{(pq)}) \) be a neat open compact subgroup such that \( \pi^U \neq 0 \). We assume that \( \pi_p \) and \( \pi_q \) are both ramified of type IIa. Suppose the residual Galois representation \( \bar{\rho}_{\pi,l} \) satisfies the following assumptions

\[\square\]
• $\bar{\rho}_{\pi,l}$ is unramified at $p$ and is ramified at $q$;
• $\rho_{\pi,l}$ satisfies $[\text{Tr}]$;
• $\rho_{\pi,l}$ satisfies $[\text{Semi-pq}]$;

Then there exists a cuspidal automorphic representation $\pi'$ with the same type as $\pi$ at $\infty$ such that $\bar{\rho}_{\pi,l} \cong \bar{\rho}_{\pi',l}$ and $\pi'_q$ is unramified.

7.2. Proof of Ribet’s principle. Recall that we have defined the Hecke algebra $T_{pq}$ in (3.8). Let $\mathfrak{m}$ be the maximal ideal corresponding to $\bar{\rho}_{\pi,l}$ in $T_{pq}$. Then by our assumption $[\text{Semi-pq}]$, we have
\[(7.2) \quad H^3_l(X_{K(pq),\mathfrak{q}},\mathcal{V})_\mathfrak{m} \otimes k = \bar{\rho}^{\mathfrak{m}_{\pi,l}}.
\]
for some positive integer $s$.

Let $U$ be an open compact subgroup of $GU_2(B)(\mathbb{A}_f)$. Consider $X^B_U$ the quaternionic Siegel threefold whose $\mathbb{C}$-points is given by the symmetric space
\[(7.3) \quad X^B_U(\mathbb{C}) = GU_2(B)(\mathbb{Q})/(\text{GSp}_4(\mathbb{R})/K^h \times GU_2(B)(\mathbb{A}_f))/U.
\]
Then $X^B_U$ has a canonical model over $\mathbb{Q}$. Since $GU_2(B)(L) = \text{GSp}_4(L)$ for a suitable quadratic extension $L/\mathbb{Q}$, we can define a $\mathbb{Q}_l$-local system $\mathcal{V}_\mu$ exactly the same way as in (3.8) for each $\mu = (a,b,c) \in X^*(T)$ on $X^B_U$. As before we impose the assumption that $l > a + b + 4$. Then under this assumption, we define $\mathcal{V} = \mathcal{V}_\mu^{\mathfrak{m}_{\pi}}$ an integral structure of $\mathcal{V}_\mu$ on $X^B_U$.

Now let $U^{pq}$ be the small open compact subgroup that we have fixed before. Let $U = K(p)K(q)U^{pq}$ and $U' = K'(p)K'(q)U^{pq}$. We set $X^B = X^B_U$ and $X^B_{K(pq)} = X_U$. By the Jacquet-Langlands correspondence for $\text{GSp}_4$ [RW19], we have an injection
\[H^3_l(X^B_{K(pq),\mathfrak{q}},\mathcal{V})_\mathfrak{m} \hookrightarrow H^3_l(X^B_{K(pq),\mathfrak{q}},\mathcal{V}_\mu).
\]
The subspace $H^3_l(X^B_{\mathfrak{q}},\mathcal{V}_\mu)$ is preserved by the Hecke algebra $T_{pq}$. Therefore
\[(7.4) \quad H^3_l(X^B_{\mathfrak{q}},\mathcal{V})_\mathfrak{m} \otimes k = \bar{\rho}^{\mathfrak{m}_{\pi,l}}
\]
for some non-negative integer $t$ under the assumption $[\text{Semi-pq}]$. Note that $t$ is positive in our case, since in the decomposition
\[(7.5) \quad H^3_l(X^B_{\mathfrak{q}},\mathcal{V})_\mathfrak{m} \otimes E = \bigoplus_{\pi'} \rho^{\mathfrak{m}_{\pi',l}}
\]
where $\pi'$ runs through all the representations that is congruent to $\pi$, there is at least one $\pi'$ such that $\pi'_p$ and $\pi'_q$ are both ramified. Our strategy is to prove the theorem by contradiction and therefore we assume that all the representation $\pi'$ appear in the following equations
\[(7.6) \quad H^3_l(X^B_{\mathfrak{q}},\mathcal{V})_\mathfrak{m} \otimes E = \bigoplus_{\pi'} \rho^{\mathfrak{m}_{\pi',l}}
\]
are ramified at $p$. As we have assumed that $\bar{\rho}_{\pi',l}$ is ramified at $q$, $\pi_q$ is necessarily ramified. Therefore it follows that in particular $t$ in (7.4) is positive. Our analysis relies on the Picard-Lefschetz formulas for $X_{K(pq),\mathfrak{q}}$ and $X^B_{\mathfrak{q}_p}$, that is the following diagrams
\[(7.7) \quad H^3_l(X_{K(pq),\mathfrak{q}},\mathcal{V})(1) \overset{\alpha}{\rightarrow} \bigoplus_{\pi \in \Sigma_q(K(pq))} R^3\Phi(\mathcal{V})_{\pi}(1) \overset{N}{\rightarrow} H^2_l(X_{K(pq),\mathfrak{q}},\mathcal{V}) \overset{\beta}{\leftarrow} \bigoplus_{\pi \in \Sigma_q(K(pq))} H^3_l(X_{K(pq),\mathfrak{q}_p},\mathcal{V}).
\]
Since $K$ presentions that are unramified at $\Sigma$.

Note that we also have a natural injection

$$H^3_c(X_{\mathbb{Q}_p}, \mathcal{V}) \cong \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(x)(1)$$

(7.8)

$$H^3_c(X_{\mathbb{Q}_p}, \mathcal{V}) \overset{\beta}{\longrightarrow} \bigoplus_{x \in \Sigma_p(B)} H^3_c(X_{\mathbb{Q}_p}, \mathcal{V}).$$

**Proposition 7.9.** Suppose that all the $\pi'$ are ramified in (7.9). The map $\beta$ localized at $m$ in (7.8) is injective

$$\beta : \bigoplus_{x \in \Sigma_p(B)} H^3_c(X_{\mathbb{Q}_p}, \mathcal{V})_m \to H^3_c(X_{\mathbb{Q}_p}, \mathcal{V})_m$$

and therefore

$$\alpha : H^3_c(X_{\mathbb{Q}_p}, \mathcal{V})_m(1) \to \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(x)_x(1).$$

is surjective.

**Proof.** Since $H^3_c(X_{\mathbb{F}_p}, \mathcal{V}) \to H^3_c(X_{\mathbb{Q}_p}, \mathcal{V})$ is injective, we need to show the Gysin map

$$\bigoplus_{x \in \Sigma_p(B)} H^3_c(X_{\mathbb{Q}_p}, \mathcal{V}) \to H^3_c(X_{\mathbb{F}_p}, \mathcal{V})$$

is injective after localizing at $m$. Note the kernel of this morphism is the image of the connecting homomorphism

$$\Pi^3_c(U, \mathcal{V}) \to \bigoplus_{x \in \Sigma_p(B)} H^3_c(X_{\mathbb{F}_p}, \mathcal{V})$$

where $U = X_{\mathbb{F}_p}^B - \Sigma_p(B)$. We claim that $H^3_c(U, \mathcal{V}) = 0$ after localizing at $m$. To prove this, we use the following excision exact sequence

$$H^2_{S_{K'(p)} - \Sigma_p(B)}(X_{\mathbb{F}_p}^B - \Sigma_p(B), \mathcal{V}) \to H^2_c(U, \mathcal{V}) \to H^2_c(X_{\mathbb{F}_p}^B - S_{K'(p)}, \mathcal{V}).$$

Note that $X_{\mathbb{F}_p}^B - \Sigma_p(B)$ is now smooth and $S_{K'(p)} - \Sigma_p(B)$ is of codimension 1 in $X_{\mathbb{F}_p}^B - \Sigma_p(B)$. Then by purity, $H^2_{S_{K'(p)} - \Sigma_p(B)}(X_{\mathbb{F}_p}^B - \Sigma_p(B), \mathcal{V}) = H^0(S_{K'(p)} - \Sigma_p(B), \mathcal{V})$. Note the set of components of $S_{K'(p)} - \Sigma_p(B)$ can be identified with the union of discrete Shimura varieties

$$X_{K'(0)}^B \sqcup X_{K'(2)}^B$$

where

$$X_{K'(0)}^B = GU_2(B^0)(\mathbb{Q})\backslash GU_2(B^0)(\mathbb{A}_f)/K_{(0)}U^p$$

$$X_{K'(2)}^B = GU_2(B^0)(\mathbb{Q})\backslash GU_2(B^0)(\mathbb{A}_f)/K_{(2)}U^p.$$

Note that we also have a natural injection

$$H^0(S_{K'(p)} - \Sigma_p(B), \mathcal{V}) \to H^0(X_{K'(0)}^B, \mathcal{V}) \oplus H^0(X_{K'(2)}^B, \mathcal{V}).$$

Since $K_{(0)}$ and $K_{(2)}$ are both hyperspecial, the set $H^0(S_{K'(p)} - \Sigma_p(B), \mathcal{V})$ only supports those representations that are unramified at $p$. Therefore $H^0(S_{K'(p)} - \Sigma_p(B), \mathcal{V})$ vanishes when localized at $m$ by our assumption. Also notice that $X_{\mathbb{F}_p}^B - S_{K'(p)}$ is the ordinary locus. Let $X_{\mathbb{F}_p}^{B'}$ be the minimal compactification of $X_{\mathbb{F}_p}^B$, then $X_{\mathbb{F}_p}^{B'} - S_{K'(p)}$ is affine and therefore

$$H^3_c(X_{\mathbb{F}_p}^{B'}, S_{K'(p)}, \mathcal{V}) = 0$$
by Artin vanishing. Since the difference between
\[ H_c^2(X_{\mathbb{F}_p}^{B*} - S_{K^\prime(p)}, V) = 0 \]
and
\[ H_c^2(X_{\mathbb{F}_p}^{B} - S_{K^\prime(p)}, V) = 0 \]
is supported on the boundary of the minimal compactification, the difference vanishes after localizing at \( m \) by our assumption \((7.1)\). Therefore \( H_c^2(X_{\mathbb{F}_p}^{B} - S_{K^\prime(p)}, V) \) vanishes after localizing at \( m \). Then we conclude that \( H_c^2(U, V) = 0 \) vanishes after localizing at \( m \).

The above Proposition gives the following fundamental diagram
\[ \bigoplus_{x \in \Sigma_p(B)} H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m \]
\[ \beta \]
\[ H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m \]
\[ \gamma \]
\[ H_3^3(x_{\mathbb{F}_p}^B, V)_m \]
\[ \alpha \]
\[ \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(V)_{x,m} \]
\[ (7.12) \]

**Proposition 7.13.** Suppose that those representation \( \pi' \) appearing in \((7.5)\) are all ramified at \( p \). Then
\[ \dim_k \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(V)_{x,m} \otimes k = 2t. \]

**Proof.** Take the fundamental diagram modulo \( l \) gives
\[ \bigoplus_{x \in \Sigma_p(B)} H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m \otimes k \]
\[ \beta \otimes k \]
\[ H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m \otimes k \]
\[ \gamma \otimes k \]
\[ H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m \otimes k \]
\[ \alpha \otimes k \]
\[ \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(V)_{x,m} \otimes k. \]

Since \( H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m \otimes k = \bar{\rho}_{\pi,l|G_{\mathbb{Q}_p}}^p \) and \( \bar{\rho}_{\pi,l|G_{\mathbb{Q}_p}} \) is unramified at \( p \), the rank one monodromy \( N \otimes k \) degenerates and has rank 0. By the mod \( l \) version of the Picard-Lefschetz formula
\[ H^3_\Sigma(x_{\mathbb{F}_p}^B, V)_m(1) \otimes k \]
\[ \alpha \otimes k \]
\[ \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(V)_{x,m}(1) \otimes k \]
\[ (7.15) \]
and the fact that \( \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(V)_{x,m} \) is obviously torsion free, we obtain that \( \beta \otimes k \) is zero and \( \gamma \otimes k \) is injective exactly the same way as in the proof of the Mazur’s principle. Therefore \( \bigoplus_{x \in \Sigma_p(B)} R^3\Phi(V)_{x,m} \otimes k \) contributes to \( \bar{\rho}_{\pi,l|G_{\mathbb{Q}_p}} \) as the two dimensional quotient which is isomorphic to the reduction modulo \( l \) of the twisted Steinberg part of the Weil-Deligne representation in \((3.4)\). Then it follows that
\[ \dim_k \bigoplus_{x \in \Sigma_p(K(pq))} R^3\Phi(V)_{x,m} \otimes k = 2t. \]

**Proposition 7.17.** We have
\[ \dim_k \bigoplus_{x \in \Sigma_p(K(pq))} R^3\Phi(V)_{x,m} \otimes k = s. \]

\[ (7.16) \]
Proof. Consider the mod $l$ fundamental diagram in this case
\[
\bigoplus_{x \in \Sigma_q(K(pq))} H^3(X_{K(pq),\tilde{q}_q}, \mathcal{V})_m \otimes k \xrightarrow{\beta \otimes k} H^3(X_{K(pq),\tilde{q}_q}, \mathcal{V})_m \otimes k \xrightarrow{\gamma \otimes k} H^3(X_{K(pq),\tilde{q}_q}, \mathcal{V})_m \otimes k \xrightarrow{\alpha \otimes k} \bigoplus_{x \in \Sigma_q(K(pq))} R^3\Phi(\mathcal{V})_{x,m} \otimes k
\]
and the mod $l$ version of the Picard-Lefschetz formula
\[
H^3(X_{K(pq),\tilde{q}_q}, \mathcal{V})_m(1) \otimes k \xrightarrow{N \otimes k} \bigoplus_{x \in \Sigma_q(K(pq))} R^3\Phi(\mathcal{V})_{x,m}(1) \otimes k \xrightarrow{N_x \otimes k} \bigoplus_{x \in \Sigma_q(K(pq))} H^3(X_{K(pq),\tilde{q}_q}, \mathcal{V})_m \otimes k.
\]
Since $\rho_{\pi,l|G_{\tilde{q}_q}}$ is ramified, $N \otimes k$ does not degenerate and still has rank one. Therefore $\beta \otimes k$ is in fact injective and contributes to $\rho_{\pi,l|G_{\tilde{q}_q}}$ as a one dimensional subspace. It follows then that
\[
\bigoplus_{x \in \Sigma_q(K(pq))} R^3\Phi(\mathcal{V})_{x,m}(1) \otimes k
\]
contributes to $\rho_{\pi,l|G_{\tilde{q}_q}}$ as a one dimensional quotient since $N_x \otimes k$ is an isomorphism. Then
\[
\dim_k \bigoplus_{x \in \Sigma_q(K(pq))} R^3\Phi(\mathcal{V})_{x,m} \otimes k = s.
\]
Recall that in Corollary 6.14, we have proved that $\Sigma_q(B)$ and $\Sigma_q(K(pq))$ can be identified. It follows immediately that $s = 2t$. Note also that if we switch the role of $pq$ in the previous arguments, we will obtain that
\[
\dim_k \bigoplus_{x \in \Sigma_q(K(pq))} R^3\Phi(\mathcal{V})_{x,m} \otimes k = 2s
\]
\[
\dim_k \bigoplus_{x \in \Sigma_q(B)} R^3\Phi(\mathcal{V})_{x,m} \otimes k = t.
\]
By Corollary 6.14 again, we have $t = 2s$. Therefore we obtain that $s = t = 0$ which is obviously impossible. Hence Ribet’s principle is proved.

References
[BLR91] N. Boston, H. Lenstra and K. Ribet, Quotients of group rings arising from two-dimensional representations, C. R. Acad. Sci. Paris Sér. I Math. 312 (1991), no. 4, 323–328.
[GH15] U. Görtz and X-H. He, Basic loci of Coxeter type in Shimura varieties, Camb. J. Math. 3 (2015), no. 3, 323–353.
[GH18] U. Görtz and X-H. He, Erratum to: Basic loci in Shimura varieties of Coxeter type, Camb. J. Math. 6 (2018), no. 3, 323–353
[GHN16] U. Görtz, X-H. He and S-A. Nie, Fully Hodge-Newton decomposable Shimura varieties, (2016), Preprint, arxiv:1610.05381.
[GT11] W-T. Gan and S. Takeda, The local Langlands conjecture for GSp(4), Ann. of Math. 3 (2011), no. 3, 1841–1882.
[Jar99] F. Jarvis, Mazur’s principle for totally real fields of odd degree, Compositio Math. 116 (1999), no. 1, 39–79.
[Illu00] L. Illusie, Sur la formule de Picard-Lefschetz, Algebraic geometry 2000, Azumino (Hotaka) Adv. Stud. Pure Math. Math. Soc. Japan, Tokyo (2002), no. 36, 249–268.
[KR00] S. Kudla and M. Rapoport, Cycles on Siegel threefolds and derivatives of Eisenstein series, Ann. Sci. École Norm. Sup. (4) 33 (2000), no. 5, 695–756.
[Kot85] R. Kottwitz, Isocrystals with additional structure, Compositio Math. 56 (1985), no. 2, 201–220.
[Lan73] R Langlands, Modular forms and $\ell$-adic representations Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Lecture Notes in Math. (1973), no. 349, 1–66.

[Laum05] G Laumon, Fonctions zêtas des variétés de Siegel de dimension trois, Formes automorphes. II. Le cas du groupe $\text{GSp}(4)$, Astérisque (2005), no. 302, 1–66.

[LS13] K-W. Lan and J. Suh, Vanishing theorems for torsion automorphic sheaves on general PEL-type Shimura varieties Adv. Math. 242 (2013), 228–286.

[LS18b] K-W. Lan and B. Stroh, Nearby cycles of automorphic étale sheaves, Compos. Math. 154 (2018), no. 1, 80–119.

[LS18a] K-W. Lan and B. Stroh, Nearby cycles of automorphic étale sheaves II, Cohomology of arithmetic groups, Springer Proc. Math. Stat. 245 (2018), Springer, Cham, 83–106.

[Laum05] G Laumon, Fonctions zêtas des variétés de Siegel de dimension trois, Formes automorphes. II. Le cas du groupe $\text{GSp}(4)$, Astérisque (2005), no. 302, 1–66.

[MP16] K. Madapusi Pera, Integral canonical models for spin Shimura varieties, Compositio Math.

[Mo14] C-P Mok, Galois representations attached to automorphic forms on $GL_2$ over CM fields, Compositio Math. 150 (2014), no. 4, 523–567.

[Raj01] A. Rajaei, On the levels of mod $F$ Hilbert modular forms, J. Reine Angew. Math. 537 (2001), 33–65.

[Rap05] M Rapoport, A guide to the reduction modulo $p$ of Shimura varieties, Automorphic forms. I Astérisque (2005), vol. 298, 271–318.

[Rib90] K. Ribet, On modular representations of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms, Invent. Math. 100 (1991), no. 2, 431–476.

[Rib91] K. Ribet, Lowering the levels of modular representations without multiplicity one, Internat. Math. Res. Notices (1991), no. 2, 15–19

[RT96] M. Rapoport and T. Zink, Period spaces for $p$-divisible groups, Annals of Mathematics Studies 141, Princeton University Press, Princeton, NJ (1996), xxii+324.

[RTW14] M. Rapoport, U. Terstiege and S. Wilson, The supersingular locus of the Shimura variety for $GU(1, n − 1)$ over a ramified prime, Math. Z. 276 (2014), no. 3-4, 1165–1188.

[RTZ13] M. Rapoport, U. Terstiege and W. Zhang, On the arithmetic fundamental lemma in the minuscule case, Compos. Math. 149 (2013), no. 10, 1631–1666.

[RW99] M. Rösner and R. Weissauer, Global liftings between inner forms of $\text{GSp}(4)$, upcoming work.

[Ser87] J-P. Serre, Sur les représentations modulaires de degré 2 de $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Duke Math. J. 54 (1987), no. 1, 179–230.

[Schm95] R. Schmidt, Iwahori-spherical representations of $\text{GSp}(4)$ and Siegel modular forms of degree 2 with square-free level $J$, Math. Soc. Japan 57 (2005), no. 1, 259–293.

[SGA7II] P. Deligne and N. Katz, Groupes de monodromie en géométrie algébrique. II, Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7, II), Lecture Notes in Mathematics. Springer-Verlag vol 340 (1973), x+438.

[Sor09a] C. Sorensen, Level-raising for Saito-Kurokawa forms Compositio Math. 15 (2010), 623–670.

[Sor09b] C. Sorensen, Potential level-lowering for $\text{GSp}(4)$ J. Inst. Math. Jussieu 8 (2009), 595–622.

[Sor10] C. Sorensen, Galois representations attached to Hilbert-Siegel modular forms Doc. Math. 15 (2010), 623–670.

[SW01] C. Skinner and A. Wiles, Base change and a problem of Serre Duke Math. J.

[Sor09a] C. Sorensen, Level-raising for Saito-Kurokawa forms Compositio Math. 15 (2010), 623–670.

[SW01] C. Skinner and A. Wiles, Base change and a problem of Serre Duke Math. J. 54 (2001), no. 1, 15–25.

[Tay93] R. Taylor, On the $\ell$-adic cohomology of Siegel threefolds, Invent. Math. 114 (1993), no. 2, 289–310.

[Tits77] J. Tits, Reductive groups over local fields, Automorphic forms, representations and $L$-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Amer. Math. Soc., Providence, R.I. (1979).

[VH19] P. Van Hoften, A geometric Jacquet-Langlands correspondence for paramodular Siegel threefolds, Preprint (2019), arXiv:1906.04008.

[Voll10] I. Voll, The supersingular locus of the Shimura variety for $GU(1, s)$, Canad. J. Math. 62 (2010), no. 3, 668–720.

[VW10] I. Voll, T. Wedhorn, The supersingular locus of the Shimura variety of $GU(1, n − 1)$ II, Invent. Math.. 184 (2011), no. 3, 591–627.

[Wang19a] H-N. Wang, On the Bruhat-Tits stratification for a quaternionic unitary Rapoport-Zink space, Preprint, arXiv:1907.05999.

[Wang19b] H-N. Wang, On quaternionic unitary Rapoport-Zink spaces with parahoric level structures, Preprint, arXiv:1909.12263.

[Weis05] R. Weissauer, Four dimensional Galois representations, Formes automorphes. II. Le cas du groupe $\text{GSp}(4)$, Astérisque no.302 (2005), 67–150.

[Weis09] R. Weissauer, Endoscopy for $\text{GSp}(4)$ and the cohomology of Siegel modular threefolds. Lecture Notes in Mathematics 1986, Springer-Verlag, Berlin, 2009.

[Yu06] C-F. Yu, The supersingular loci and mass formulas on Siegel modular varieties, Doc. Math. 11 (2006), 449–468.

[Yu11a] C-F. Yu, Geometry of the Siegel modular threefold with paramodular level structure, Proc. Amer. Math. Soc. 139 (2011), no. 9, 3181–3190.
