Modularity theorems for abelian surfaces

Toby Gee

October 3, 2025

Abstract. This is a brief account of my results with George Boxer, Frank Calegari and Vincent Pilloni on the (potential) modularity of abelian surfaces.

1 Introduction This article gives an overview of the proofs of the main results of [7, 8]. I have attempted to complement the introductions to those papers, and to concentrate on aspects of the proofs that are not already covered in other survey articles. Consequently, I devote little space to innovations in the Taylor–Wiles method (in particular, the Calegari–Geraghty method), for which I refer the reader to Calegari's excellent surveys [16, 15].

We begin by recalling the main theorems of the papers. Our first paper [7] proves the Hasse–Weil conjecture for abelian surfaces (and genus 2 curves) over totally real fields. To recall what this means, let X be a smooth, projective variety over a number field F with good reduction outside a finite set of primes S. Associated to X, one may write down a global Hasse–Weil zeta function:

$$\zeta_X(s) = \prod \frac{1}{1 - N(x)^{-s}},$$

where the product runs over all the closed points x of some (any) smooth proper integral model $\mathcal{X}/\mathcal{O}_F[1/S]$ for X. (Different choices of S only change $\zeta_X(s)$ by a finite number of Euler factors. For curves and abelian varieties there is a natural definition of the Euler factors at all places, and our modularity results are compatible with these factors, but we ignore this point from now on.) The function $\zeta_X(s)$ is absolutely convergent for $\operatorname{Re}(s) > 1 + \dim X$. We have the following:

CONJECTURE 1.1 (Hasse–Weil Conjecture, cf. [43], in particular Conj. C9). *The function $\zeta_X(s)$ extends to a meromorphic function of $\mathbf{C}$. There exists a positive real number $A \in \mathbf{R}^{>0}$, non-zero rational functions $P_v(T)$ for $v|S$, and infinite Gamma factors $\Gamma_v(s)$ for $v|\infty$ such that:*

$$\xi(s) = \zeta_X(s) \cdot A^{s/2} \cdot \prod_{v|\infty} \Gamma_v(s) \cdot \prod_{v|S} P_v(N(v)^{-s})$$

satisfies the functional equation $\xi(s) = w \cdot \xi(\dim X + 1 - s)$ with $w = \pm 1$.

(In Serre's formulation of the conjecture, the Gamma factors are also given explicitly in terms of the Archimedean Hodge structures of X.)

If $F = \mathbf{Q}$ and X is a point, then $\zeta_X(s)$ is the Riemann zeta function, and Conjecture 1.1 follows from Riemann's functional equation. By work of Hecke and Brauer, the conjecture is known if X is zero-dimensional, or if more generally the Galois representations associated to the l-adic cohomology of X are *potentially* abelian (e.g. an abelian variety with CM). All subsequent progress on Conjecture 1.1 has been via the Langlands program. Write $\operatorname{Gal}_{\mathbf{Q}}$ for the absolute Galois group of $\mathbf{Q}$; more generally, for any field K, we write $\overline{K}$ for a separable closure, and $\operatorname{Gal}_K := \operatorname{Gal}(\overline{K}/K)$ for the absolute Galois group. If K is a local field, then we write I_K for the inertia subgroup of Gal_K. Via the Grothendieck–Lefschetz trace formula, one writes for each prime ℓ (at least up to a finite number of Euler factors)

$$\zeta_X(s) = \prod_{n=0}^{2 \dim X} L\big(H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell), s\big)^{(-1)^n}$$

where $L\big(H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell), s\big)$ is the L-function of the ℓ-adic representation $H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell)$ of $\operatorname{Gal}_{\mathbf{Q}}$. The Langlands conjectures predict that each $L\big(H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell), s\big)$ is a product of automorphic L-functions; more precisely, we

have $L(H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell), s) = \prod_j L(\pi_j, s)$, where each π_j is an automorphic representation of $\mathrm{GL}_{n_j}(F)$ for some integers n_j with $\sum_j n_j = \dim_{\overline{\mathbf{Q}}_\ell} H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell)$. Since (completed) automorphic L-functions have meromorphic (usually holomorphic) continuations and functional equations, this prediction implies the Hasse–Weil conjecture.

DEFINITION 1.2. *If X/F is a curve or an abelian variety, we write $L(X, s) := L(H^1(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell), s)$, and we say that X is* modular *if $L(X, s)$ is a product of automorphic L-functions. We say that X is* potentially modular *if there is a finite extension F'/F such that $X_{F'}$ is modular. Similarly, we say that a representation $\rho : \mathrm{Gal}_F \to \mathrm{GL}_n(\overline{\mathbf{Q}}_\ell)$ is* modular *if $L(\rho, s)$ is a product of automorphic L-functions, and we say that a representation $\overline{\rho} : \mathrm{Gal}_F \to \mathrm{GL}_n(\overline{\mathbf{F}}_l)$ is* modular *if it is the reduction modulo ℓ of a modular Galois representation.*

Remark 1.3. The notion of potential modularity was introduced by Taylor [46], who observed that Brauer's methods apply in this setting, so that if X is potentially modular, then $L(X, s)$ has the expected meromorphic continuation and functional equations. However one cannot in general say anything about the poles (or lack thereof) of this meromorphic continuation.

If X is a curve of genus zero, then (up to bad Euler factors) $\zeta_X(s) = \zeta_F(s)\zeta_F(s-1)$, and Conjecture 1.1 follows immediately. The fundamental work of Wiles [49, 47] and the subsequent work of Breuil, Conrad, Diamond, and Taylor [20, 12] proved Conjecture 1.1 for curves $X/\mathbf{Q}$ of genus one, since if we write $E = \mathrm{Jac}(X)$, then $\zeta_X(s) = \zeta(s)\zeta(s-1)/L(E, s)$ so the modularity of E implies the holomorphy and functional equation for $L(E, s)$. More generally, Taylor's potential modularity results [46] prove Conjecture 1.1 for curves X/F of genus one over any totally real field.

The methods used in these papers have been vastly generalized over the past 30 years due to the enormous efforts of many people, and as a consequence one knows for example that if F is totally real and X is such that the Hodge numbers $h^{p,q} = \dim H^{p,q}_{\mathrm{dR}}(X) = \dim H^q(X, \Omega^p)$ of X are at most 1 for all p and q with $p + q = n$, then $L(H^n(X_{\overline{F}}, \overline{\mathbf{Q}}_\ell), s)$ has the expected meromorphic continuation and functional equation (see [31, Cor. B]).

Unfortunately, this multiplicity one condition on Hodge numbers is fundamental to the original Taylor–Wiles method, and there is a paucity of natural geometric examples satisfying this condition. In particular, it fails for curves of genus $g > 1$ and for abelian varieties of dimension $g > 1$, where $h^{1,0} = h^{0,1} = g$. The main theorem of [7] is the following.

THEOREM 1.4. *Let X be either a genus two curve or an abelian surface over a totally real field F. Then X is potentially modular, and Conjecture 1.1 holds for X.*

(The deduction of Conjecture 1.1 from potential modularity is straightforward in this case, using that the cohomology of X is given by the wedge powers of the cohomology in degree 1, and known Langlands functoriality results for wedge powers.)

While Theorem 1.4 resolves the Hasse–Weil conjecture for abelian surfaces $A/\mathbf{Q}$, for many purposes (e.g. applications to the Birch–Swinnerton-Dyer conjecture) one wishes to know modularity.

Modularity is known in some cases, using the results recalled above (in particular, the modularity of elliptic curves). More precisely, by [8, Thm. 10.2.1], it is known unless $A/\mathbf{Q}$ is "challenging" in the sense of [7, §9.2], which means that either

(1) $\mathrm{End}(A_{\overline{\mathbf{Q}}}) = \mathbf{Z}$, or

(2) there exists a quadratic field $K/\mathbf{Q}$ so that $\mathrm{End}(A) = \mathbf{Z}$ but $\mathrm{End}(A_K) \otimes \mathbf{Q}$ is either $\mathbf{Q} \oplus \mathbf{Q}$ or a real quadratic field.

A natural source of abelian surfaces of type (2) are those of the form $\mathrm{Res}_{K/\mathbf{Q}}(E)$ for a non-CM elliptic curve E which is not isogenous to its $\mathrm{Gal}(K/\mathbf{Q})$-conjugate. In this case the modularity of A would follow from the modularity of E. If $K/\mathbf{Q}$ is real quadratic, then E is modular by Freitas–Le Hung–Siksek [23], while if K is imaginary quadratic, then the modularity of E is known in many cases by work of Caraiani–Newton [18]. On the other hand, the endomorphism algebra $\mathrm{End}(A_K) \otimes \mathbf{Q}$ could also be a real quadratic field E rather than $\mathbf{Q} \times \mathbf{Q}$, in which case A/K will be a simple abelian surface of GL_2-type, and the modularity of such abelian surfaces remains open in general even for real quadratic fields K.

In view of this, we concentrate from now on the "typical" case that $\mathrm{End}_{\overline{\mathbf{Q}}}(A) = \mathbf{Z}$, where one has the more precise expectation that $L(A, s) = L(\pi, s)$ for some cuspidal automorphic representation π of $\mathrm{GSp}_4 /\mathbf{Q}$, and Brumer and Kramer [13] formulated the *paramodular conjecture*, which gives a precise prescription for the

"optimal" level structure for an automorphic form corresponding to a given abelian surface; in particular, this in principle reduces the conjecture for a given A to an explicit computation of a (finite-dimensional) space of Siegel modular forms. Using the Faltings–Serre method, and elaborate explicit computations of low weight Siegel modular forms, developed in part by Poor and Yuen [36, 35, 11], the modularity of (finitely many, up to twist) abelian surfaces A with $\mathrm{End}_{\overline{\mathbf{Q}}}(A) = \mathbf{Z}$ was established in the papers [36, 14, 3].

If A/F is an abelian surface, we write $\rho_{A,p}$ for the Galois representation associated to $H^1(A_{\overline{F}}, \mathbf{Z}_p)$, which we often think of as a representation

$$\rho_{A,p} : \mathrm{Gal}_F \to \mathrm{GSp}_4(\mathbf{Q}_p)$$

with multiplier given by the inverse cyclotomic character ε^{-1}. By definition, A is modular if and only if $\rho_{A,p}$ is modular for some p (equivalently, for all p). We also let $\overline{\rho}_{A,p}$ denote the Galois representation associated to $H^1(A_{\overline{F}}, \mathbf{F}_p)$. If A admits a principal polarization of degree prime to p, then we can and do think of $\overline{\rho}_{A,p}$ as a representation

$$\overline{\rho}_{A,p} : \mathrm{Gal}_F \to \mathrm{GSp}_4(\overline{\mathbf{F}}_p).$$

The main theorem of [8] is as follows.

THEOREM 1.5. *Let $A/\mathbf{Q}$ be an abelian surface with a polarization of degree prime to 3. Suppose the following hold:*

(1) *The mod 3 representation*

$$\overline{\rho}_{A,3} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\mathbf{F}_3)$$

   is surjective.

(2) *The representation $\overline{\rho}_{A,3}|_{\mathrm{Gal}_{\mathbf{Q}_2}}$ is unramified, and the characteristic polynomial of $\overline{\rho}_{A,3}(\mathrm{Frob}_2)$ is not $(x^2 \pm x + 2)^2$.*

(3) *A has good ordinary reduction at 3, and the characteristic polynomial of Frob_3 does not have repeated roots.*

Then A is modular.

More precisely, there exists a cuspidal automorphic representation π of $\mathrm{GL}_4/\mathbf{Q}$ (the transfer of a cuspidal automorphic representation of $\mathrm{GSp}_4/\mathbf{Q}$ of weight 2) such that $L(s, H^1(A)) = L(s, \pi)$. Consequently, $L(s, H^1(A))$ has a holomorphic continuation to $\mathbf{C}$ and satisfies the expected functional equation.

Remark 1.6. We claim that Theorem 1.5 applies to a positive proportion of abelian surfaces over $\mathbf{Q}$, counted in any reasonable sense. As one justification of this, suppose that one samples genus two curves

$$X : y^2 + h(x)y = f(x)$$

with $h(x), f(x) \in \mathbf{Z}[x]$ of degrees ≤ 3 and ≤ 6 in any way in which the distributions modulo 2 and 3 are equidistributed, and considers those curves X with the following properties:

(1) $\overline{\rho}_{\mathrm{Jac}\,X,3} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\mathbf{F}_3)$ is surjective,

(2) X has good reduction at 2,

(3) X has good ordinary reduction at 3,

(4) $\overline{\rho}_{\mathrm{Jac}(X),3}(\mathrm{Frob}_2)$ does not have characteristic polynomial $(x^2 \pm x + 2)^2$,

(5) The characteristic polynomial of Frob_3 has distinct eigenvalues.

Theorem 1.5 proves the modularity of $\mathrm{Jac}(X)$ for any such X, and one can check [8, §10.1] that these X form a subset of density $\frac{5551}{46656} = 0.1189\ldots$ Another point of comparison is with the curves in the database [48] (see also [6]). This contains 63107 genus two curves $X/\mathbf{Q}$ with $\mathrm{End}\,\mathrm{Jac}(X)_{\overline{\mathbf{Q}}} = \mathbf{Z}$, and Theorem 1.5 applies to 11384 of them.

2   The 2–3 switch   The proof of Theorem 1.5 follows Wiles's strategy for proving the modularity of semistable elliptic curves, and in particular, we make use of an analogue of the 3-5 switch used by Wiles [49] to prove residual modularity. That switch exploited the rationality of certain twists of the modular curve $X(5)/\mathbf{Q}$. In our case, we use a rational moduli space of abelian surfaces to carry out a 2-3 switch.

In outline, the 2-3 switch proving Theorem 1.5 divides into three steps as follows.

Step 1 Show that $\overline{\rho}_{B,2}$ is modular for many abelian surfaces $B/\mathbf{Q}$. (See Lemma 2.3.)

Step 2 Show that for any abelian surface A as in Theorem 1.5, there exists B as in (Step 1) with $\overline{\rho}_{B,3} \cong \overline{\rho}_{A,3}$. (See Lemma 2.4.)

Step 3 Prove the following (imprecisely stated) modularity lifting theorem, which applies in particular to the representations $\rho_{B,2}$ and $\rho_{A,3}$ for A, B as in the previous two steps:

> THEOREM 2.1. *Suppose that $\rho : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\overline{\mathbf{Z}}_p)$ is unramified at all but finitely many primes and de Rham at p, and:*
>
>   (i) *$\overline{\rho} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\overline{\mathbf{F}}_p)$ is modular.*
>   (ii) *$\overline{\rho}(\mathrm{Gal}_{\mathbf{Q}})$ is large.*
>   (iii) *ρ is pure.*
>   (iv) *$\rho|_{\mathrm{Gal}_{\mathbf{Q}_p}}$ is ordinary, p-distinguished, and has Hodge–Tate weights $0, 0, 1, 1$.*
>
> *Then ρ is modular.*

Proof of Theorem 1.5, given these steps. Suppose that A satisfies the hypotheses of Theorem 1.5, and let B be as in Step 2. Then $\overline{\rho}_{B,2}$ is modular by Step 1, so that $\rho_{B,2}$ is modular by Theorem 2.1. Equivalently, $\rho_{B,3}$ is modular, so that $\overline{\rho}_{B,3}$ is modular. Since $\overline{\rho}_{A,3} \cong \overline{\rho}_{B,3}$ by assumption, we can apply Theorem 2.1 to deduce that $\rho_{A,3}$ is modular, as required.   □

Remark 2.2. Hypothesis (ii) of Theorem 2.1 is responsible for assumption (1) in Theorem 1.5, while the more serious hypothesis (iv) corresponds to assumption (3) there (and is also responsible for (2)).

Our supply of abelian surfaces $B/\mathbf{Q}$ for Step 1 will be certain Jacobians $B = \mathrm{Jac}(X)$, where $X/\mathbf{Q}$ is a genus two curve. Let r_i for $i = 1, \ldots, 6$ be the Weierstrass points of X over $\overline{\mathbf{Q}}$; then the non-zero elements of $B[2]$ are given by the $r_i - r_j$ for $i < j$. Considering the action of $\mathrm{Gal}_{\mathbf{Q}}$ on the r_i, one has an identification $S_6 \xrightarrow{\sim} \mathrm{Sp}_4(\mathbf{F}_2) = \mathrm{GSp}_4(\mathbf{F}_2)$. There are two conjugacy classes of subgroup $S_5 \subset S_6$; we denote by $S_5(b)$ the standard copy of S_5 in S_6 (and below we write $A_5(b)$ for the copy of A_5 in $S_5(b)$). Thus X has a rational Weierstrass point (so that X can be written in the form $y^2 = f(x)$ with f quintic) if and only if $\overline{\rho}_{A,2}$ factors through a conjugate of $S_5(b)$.

The following lemma, which exploits some coincidences in the representation theory of A_5, allows us to find many X for which we know that $\overline{\rho}_{B,2}$ is modular.

LEMMA 2.3. *Suppose that $X/\mathbf{Q}$ is a genus two curve with a rational Weierstrass point, and that $B := \mathrm{Jac}(X)$ has semistable ordinary or good ordinary reduction at 2. Suppose also that*

$$\overline{\rho}_{B,2} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\mathbf{F}_2) \simeq S_6$$

has image $S_5(b)$, and that the image of complex conjugation has conjugacy class $(**)(**)$. Then $\overline{\rho}_{B,2}$ is modular, arising from an ordinary weight 3 Siegel modular form.

Proof. If F^+ is the quadratic field given by the kernel of the composite $\mathrm{Gal}_{\mathbf{Q}} \to S_5(b) \to \mathbf{Z}/2\mathbf{Z}$, then $\overline{\rho}(\mathrm{Gal}_{F^+}) = A_5(b)$, and F^+ is real by the assumption on complex conjugation. Let

$$\overline{\varrho} : \mathrm{Gal}_{F^+} \to \mathrm{SL}_2(\mathbf{F}_4) \simeq A_5$$

denote the residual 2-dimensional Galois representation associated to this A_5-extension. (There are two such representations which are permuted by the outer automorphism; choose either.) Either by an easy Brauer character computation, or as a consequence of the Steinberg tensor product theorem for $\mathrm{SL}_2(\mathbf{F}_4)$, we have

$$\overline{\rho}_{B,2} \cong \mathrm{Sym}^3 \overline{\varrho}.$$

By a theorem of Tate the composite $\overline{\varrho} : \mathrm{Gal}_{F^+} \to A_5 \hookrightarrow \mathrm{PGL}_2(\mathbf{C})$ lifts to an odd representation $\varrho : \mathrm{Gal}_{F^+} \to \mathrm{GL}_2(\mathbf{C})$ with finite image (which will be some central extension of A_5). By the odd Artin conjecture for GL_2 (i.e. by the main results of [34] or [40]), ϱ is modular. By Hida theory, it follows that $\overline{\varrho}$ is modular, coming from an ordinary Hilbert modular eigenform of parallel weight 2. By symmetric cube functoriality [27], $\overline{\rho}_{B,2}|_{\mathrm{Gal}_{F^+}}$ is modular, arising from an ordinary weight 3 Hilbert–Siegel modular form. By solvable base change and a standard use of the Khare–Wintenberger method [26], $\overline{\rho}_{B,2}$ itself is modular, arising from an ordinary weight 3 Siegel modular form, as required. □

Step 2 is provided by the following lemma (see [8, Lem. 9.4.2]).

LEMMA 2.4. *Let $A/\mathbf{Q}$ be an abelian surface with a polarization of degree prime to 3. Assume that $\overline{\rho}_{A,3}|_{\mathrm{Gal}_{\mathbf{Q}_2}}$ is unramified, and the characteristic polynomial of $\overline{\rho}_{A,3}(\mathrm{Frob}_2)$ is not $(x^2 \pm x + 2)^2$.*

Then there exists a genus two curve $X/\mathbf{Q}$ with a rational Weierstrass point, with $B = \mathrm{Jac}(X)$ having the following properties:

(1) $\overline{\rho}_{B,3} \cong \overline{\rho}_{A,3}$.

(2) B has good ordinary or semistable ordinary reduction at 2, and is 2-distinguished.

(3) B has good ordinary reduction at 3.

(4) $\mathrm{End}(B_{\overline{\mathbf{Q}}}) = \mathbf{Z}$.

(5) The representation

$$\overline{\rho}_{B,2} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\mathbf{F}_2)$$

has image $S_5(b)$, and the image of complex conjugation has conjugacy class $(**)(**)$.

Proof. We consider the moduli space $M(\overline{\rho}_{A,3})/\mathbf{Q}$ of genus-two curves X equipped with a fixed rational Weierstrass point and a symplectic isomorphism $\overline{\rho}_{\mathrm{Jac}(X),3} \cong \overline{\rho}_{A,3}$. By definition, any $\mathbf{Q}$-point of $M(\overline{\rho}_{A,3})$ gives $B = \mathrm{Jac}(X)$ satisfying ((1)), and $\overline{\rho}_{B,2}(\mathrm{Gal}_{\mathbf{Q}}) \subseteq S_5(b)$ (due to the rational point).

By [7, Lem. 10.2.4], the variety $M(\overline{\rho}_{A,3})/\mathbf{Q}$ is rational. (The proof of this lemma is very similar to that of [45, Lem. 1.1]: namely, a Galois cohomology argument reduces to checking that the corresponding moduli space for the representation $1 \oplus 1 \oplus \overline{\varepsilon}^{-1} \oplus \overline{\varepsilon}^{-1}$ is $\mathrm{PGL}_4(\mathbf{F}_3)$-equivariantly rational over $\mathbf{Q}$, which is known by [5, Theorem 0.0.1].) In particular, by weak approximation (combined with Hilbert irreducibility as in [44, §3.4, §3.5]), for any finite set of places S and any nonempty open subsets $\Omega_\ell \subset M(\overline{\rho}_{A,3})(\mathbf{Q}_l)$ (in the ℓ-adic topology) for each $\ell \in S$, we can find a $\mathbf{Q}$-point of $M(\overline{\rho}_{A,3})$ lying in Ω_ℓ for each $\ell \in S$, for which $\overline{\rho}_{B,2}(\mathrm{Gal}_{\mathbf{Q}}) = S_5(b)$ (which in turn implies that $\mathrm{End}(B_{\overline{\mathbf{Q}}}) = \mathbf{Z}$). We then take $S = \{2, 3, \infty\}$; the hypothesis that the characteristic polynomial of $\overline{\rho}_{A,3}(\mathrm{Frob}_2)$ is not $(x^2 \pm x + 2)^2$ is used to guarantee the existence of suitable points over $\mathbf{Q}_2$, by writing down appropriate abelian surfaces over $\mathbf{Q}_2$ for each of the other possibilities for this characteristic polynomial. □

For Theorem 1.4, we use a variant of the above strategy, where we choose primes p, q splitting completely in our totally real field F, with the further properties that A admits a polarization of degree prime to pq, the representations $\overline{\rho}_{A,p}$ and $\overline{\rho}_{A,q}$ have large image, and A has good ordinary reduction at all places dividing pq. Then we have:

LEMMA 2.5. *Let A/F be a challenging abelian surface over a totally real field. Then there is a finite Galois extension of totally real fields F'/F such that p and q split completely in F', and an abelian surface B/F' with good ordinary reduction at all places dividing pq, such that $\overline{\rho}_{B,p} \cong \overline{\rho}_{A,p}|_{\mathrm{Gal}_{F'}}$, while $\overline{\rho}_{B,q}$ is modular of parallel weight 2, and has large image.*

Proof. We consider the moduli space Y of abelian surfaces B equipped with isomorphisms $B[p] \xrightarrow{\sim} A[p]$ and $B[q] \cong \overline{\rho}_q$, for any choice of $\overline{\rho}_q : \mathrm{Gal}_F \to \mathrm{GSp}_4(\mathbf{F}_q)$. There is no reason that $Y(F)$ should be nonempty, but a theorem of Moret-Bailly [28] guarantees that we can find F' as above so that $Y(F')$ is nonempty, and satisfies the required local conditions at primes dividing pq. Choosing $\overline{\rho}_q$ to be induced from a 2-dimensional representation, we are able to arrange (using the (known) potential modularity of elliptic curves) that after a further extension of F', $\overline{\rho}_{B,q}$ is modular. □

Remark 2.6. As well as the (important!) difference between modularity and potential modularity, another significant difference between the Lemma 2.5 and Lemmas 2.3 and 2.4 is that Lemma 2.5 shows residual potential modularity in weight 2, while Lemma 2.3 shows residual modularity in weight 3. (which is regular).

The Galois representations associated to Siegel modular eigenforms of weight $k \geq 2$ have Hodge–Tate weights $0, k-2, k-1, 2k-3$ (see Section 3.1 below for the definition of Hodge–Tate weights), while the Galois representations associated to abelian surfaces have Hodge–Tate weights $0, 0, 1, 1$. It follows that if an abelian surface is modular, corresponding to a Siegel modular form, that modular form must be of weight 2; so when we have residual modularity in weight 3, we have to "change weight" in some way (see Remark 2.7 below for more discussion of this). On the other hand, if $k > 2$ then the Hodge–Tate weights $0, k-2, k-1, 2k-3$ are pairwise distinct, and we say that k is a "regular weight", while $k = 2$ is an "irregular weight". The irregular weight cases behave quite differently (and are in general much more complicated than the regular weight cases) on both the automorphic and Galois sides of the Langlands correspondences, as we will see below.

The remainder of this survey is devoted to Step 3 (and its analogue in [7]). This step divides into two parts. Firstly, we use the Taylor–Wiles method to show that ρ is p-adically modular, in the sense that it contributes to a Hida family of Siegel modular forms. This is relatively standard, although the need to consider the primes $p = 2, 3$ causes some pain. Secondly, we prove a classicality criterion for weight 2 ordinary Siegel p-adic modular forms. It is in this step that the difference mentioned in Remark 2.6 becomes significant. In [7], the relevant classicality theorem is a "small slope implies classical" theorem in the style of Coleman, but this approach does not suffice in the situation considered in [8]. In this case the proof, building upon work of Rodríguez Camargo [38], is a generalisation from GL_2 to GSp_4 of a part of Lue Pan's remarkable work [30]. In the remainder of this survey, we concentrate on the classicality theorems, before briefly returning to the Taylor–Wiles method in Section 5.

Remark 2.7. It was well-known for many years that Theorem 1.4 could be deduced from strong enough modularity lifting theorems; we refer the reader to [2, §6] and [16, §11.2] for some of the history. The paper [7] was posted online at the end of 2018, and we found the strategy outlined above for proving Theorem 1.5 in March 2019. At that time, we imagined that we would need to prove a "low weight mod p companion form" result to show that the weight 3 Siegel modular form in Lemma 2.3 is congruent to a weight 2 Siegel modular form, and then apply the modularity lifting theorems of [7]. We still do not know how to prove such companion form theorems, but the situation changed in 2020 with Lue Pan's paper [30]. By the spring of 2022, we were confident that it was possible to use Pan's techniques to prove an appropriate classicality result, but we had not yet proved (regular weight) modularity lifting theorems for GSp_4 which could applied to the representations $\overline{\rho}_{B,2}$. After some false starts, we managed to do this in spring 2023, and we wrote [8] over the following two years.

3 Galois representations

3.1 Sen theory Let $\mathbf{Q}_p(\zeta_{p^\infty}) \coloneqq \cup_n \mathbf{Q}_p(\zeta_{p^n})$ be the cyclotomic extension of $\mathbf{Q}_p$. Then $\mathrm{Gal}_{\mathbf{Q}_p(\zeta_{p^\infty})}$ is the kernel of the cyclotomic character $\varepsilon : \mathrm{Gal}_{\mathbf{Q}_p} \to \mathbf{Z}_p^\times$. By the Ax–Sen–Tate theorem, we have $\mathbf{C}_p^{\mathrm{Gal}_{\mathbf{Q}_p(\zeta_{p^\infty})}} = \widehat{\mathbf{Q}_p(\zeta_{p^\infty})}$, the completion of $\mathbf{Q}_p(\zeta_{p^\infty})$. Let $V/\mathbf{Q}_p$ be a finite-dimensional $\mathbf{Q}_p$-vector space equipped with a continuous action of $\mathrm{Gal}_{\mathbf{Q}_p}$, so that $\rho : \mathrm{Gal}_{\mathbf{Q}_p} \to \mathrm{GL}(V)$ is a Galois representation. Then $V \otimes_{\mathbf{Q}_p} \mathbf{C}_p$ has a semilinear action of $\mathrm{Gal}_{\mathbf{Q}_p}$, and Sen showed that it descends to $\widehat{\mathbf{Q}_p(\zeta_{p^\infty})}$, in the sense that

$$(V \otimes_{\mathbf{Q}_p} \mathbf{C}_p)^{\mathrm{Gal}_{\mathbf{Q}_p(\zeta_{p^\infty})}} \otimes_{\widehat{\mathbf{Q}_p(\zeta_{p^\infty})}} \mathbf{C}_p \cong V \otimes_{\mathbf{Q}_p} \mathbf{C}_p.$$

Furthermore, it descends to $\mathbf{Q}_p(\zeta_{p^\infty})$: there is a unique $\mathrm{Gal}(\mathbf{Q}_p(\zeta_{p^\infty})/\mathbf{Q}_p)$-stable $\mathbf{Q}_p(\zeta_{p^\infty})$-subspace $\mathbf{D}_{\mathrm{Sen}}(V)$ of $V \otimes_{\mathbf{Q}_p} \mathbf{C}_p$ such that

$$\mathbf{D}_{\mathrm{Sen}}(V) \otimes_{\mathbf{Q}_p(\zeta_{p^\infty})} \mathbf{C}_p \cong V \otimes_{\mathbf{Q}_p} \mathbf{C}_p.$$

In fact, $\mathbf{D}_{\mathrm{Sen}}(V)$ is the union of the finite-dimensional $\mathrm{Gal}(\mathbf{Q}_p(\zeta_{p^\infty})/\mathbf{Q}_p)$-stable $\mathbf{Q}_p(\zeta_{p^\infty})$-subspaces of $V \otimes_{\mathbf{Q}_p} \mathbf{C}_p$.

The $\mathbf{Q}_p(\zeta_{p^\infty})$-vector space $\mathbf{D}_{\mathrm{Sen}}(V)$ has an action of $\mathrm{Gal}(\mathbf{Q}_p(\zeta_{p^\infty})/\mathbf{Q}_p)$, and thus a linear action of $\mathrm{Lie}\,(\mathrm{Gal}(\mathbf{Q}_p(\zeta_{p^\infty})/\mathbf{Q}_p))$. More explicitly, we have the Sen operator Θ_V, which is the $\mathbf{Q}_p(\zeta_{p^\infty})$-linear map given by $\Theta_V \coloneqq \log(\gamma)/\log_p(\varepsilon(\gamma))$ for any $\gamma \in \mathrm{Gal}(\mathbf{Q}_p(\zeta_{p^\infty})/\mathbf{Q}_p)$ sufficiently close to 1.

By definition, V is Hodge–Tate, with Hodge–Tate weights $h_1, \ldots, h_n \in \mathbf{Z}$, if there exists an isomorphism of $\mathbf{C}_p[\mathrm{Gal}_{\mathbf{Q}_p}]$-modules

$$V \otimes_{\mathbf{Q}_p} \mathbf{C}_p \cong \oplus_{i=1}^n \mathbf{C}_p(-h_i),$$

where $\mathbf{C}_p(n)$ is the nth Tate twist; thus V is Hodge–Tate if and only if the Sen operator Θ_V is semisimple and has eigenvalues in $\mathbf{Z}$, in which case the Hodge–Tate weights are the negatives of the eigenvalues of Θ_V.

3.2   Ordinary Galois representations There are various definitions in the literature of ordinary Galois representations; the following will be convenient for us.

DEFINITION 3.1. *We say that $\rho : \mathrm{Gal}_{\mathbf{Q}_p} \to \mathrm{GL}(V)$ is ordinary if there are integers $h_1 \leq h_2 \cdots \leq h_n$ such that ρ is conjugate to an upper-triangular representation*

$$\begin{pmatrix} \chi_1 & * & \cdots & * \\ 0 & \chi_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \chi_n \end{pmatrix}$$

where each $\chi_i : \mathrm{Gal}_{\mathbf{Q}_p} \to \mathbf{Q}_p^\times$ is a character with $\chi_i|_{I_{\mathbf{Q}_p}} = \varepsilon^{-h_i}$.

It is straightforward to check that an ordinary Galois representation is de Rham if and only if it is Hodge–Tate, i.e. if and only if the Sen operator Θ_V is semisimple. If $h_1 < \cdots < h_n$, this is automatic (e.g. because the eigenvalues of the Sen operator are distinct), but in general it need not hold. Indeed, a standard example of a representation which is not Hodge–Tate but whose Sen operator has integral eigenvalues is the ordinary representation

$$\begin{pmatrix} 1 & \log_p \varepsilon \\ 0 & 1 \end{pmatrix},$$

which by definition has Sen operator $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

4 p-adic modular forms and classicality We now review the approach to p-adic modular forms and classicality theorems taken in [8], which relies on the work of Pan [30] and its generalisations to other Shimura varieties by Rodríguez Camargo [38], as well as the higher Coleman and Hida theories of Boxer–Pilloni [9, 10]. There is a long history of such classicality theorems, going back in the case of Coleman theory to [19]; we highlight in particular Kassaei's paper [24], and its generalisation in [4]. Another important ingredient is the families of p-adic automorphic forms introduced by Andreatta–Iovita–Pilloni [1] (see [2] for a survey).

For simplicity we say almost nothing below about compactifications of Shimura varieties, or about the distinction between cuspidal and usual cohomology. These both play an important role in the foundations of the theory, but disappear in our main results, as we always localise at a non-Eisenstein maximal ideal of a Hecke algebra. Similarly, we will sometimes elide the difference between functors on the abelian and derived level where it makes no difference for our final statements. We will also only work with p-adic modular forms of integral weight, as this is all that is needed for our main theorems and allows us to simplify the exposition in places. We say very little about families of p-adically varying weight, although these are an important ingredient in the proofs of some results that we state.

We will also be extremely informal in our treatment of p-adic functional analysis and condensed mathematics. A justification for this is that at the time of writing, the "correct" foundations for the constructions we discuss are not yet available. In particular, the arguments below use a p-adic version of Beilinson–Bernstein localisation; ideally, this should be defined in the framework of the analytic de Rham stack of Rodríguez Camargo [39], similarly to Scholze's treatment of classical Beilinson–Bernstein localisation in [42]. This formalism is expected to be available soon, but in the meantime [8] proceeds in a somewhat ad hoc fashion.

While the only Shimura varieties considered here are the Siegel threefolds associated to the group GSp_4, many of the results that we explain below are proved in [8] for Hodge-type Shimura varieties, or (in the case of results requiring a non-Eisenstein localisation) are expected to hold in this generality. Accordingly, where possible we phrase our results without reference to specific features of GSp_4, although we caution the reader that they should turn to [8] to see the precise hypotheses under which each result is proved.

Accordingly, from now on we write $G = \mathrm{GSp}_4$; when we eventually need to be concrete, we will realise G as the subgroup of GL_4 acting on the free $\mathbf{Z}$-modules of rank 4, with basis $e_1, \cdots, e_4$ and preserving up to a similitude factor the symplectic form with matrix

$$J = \begin{pmatrix} 0 & S \\ -S & 0 \end{pmatrix}$$

where S is the 2×2 anti-diagonal matrix with only 1's on the anti-diagonal. We take P to be the ("block lower-triangular") Siegel parabolic stabilising e_3, e_4, and B the Borel inside it which is upper-triangular in each of the diagonal 2×2 blocks. We let M be the Levi quotient of P, and let U, U_P be the unipotent radicals of B, P respectively. We let T be the diagonal torus. The sets of M-dominant and G-dominant characters are respectively denoted by $X^*(T)^{M,+}$ and $X^*(T)^+$. We let $\mu \in X_*(T)$ be the minuscule dominant cocharacter $t \mapsto \operatorname{diag}(1, 1, t, t)$.

We let W be the Weyl group of G, with length function $\ell : W \to \mathbf{Z}_{\geq 0}$, and write w_0 for the longest element of W. Let Φ^+ be the set of positive roots of G, and let ρ be half the sum of the positive roots. Write W_M for the Weyl group of M, and let ${}^M W \subseteq W$ be the set of Kostant representatives of $W_M \backslash W$ (i.e. those $w \in W$ with $wX^*(T)^+ \subseteq X^*(T)^{M,+}$; this is a set of coset representatives of minimal length).

We denote by $\mathfrak{g}$, $\mathfrak{b}$, $\mathfrak{h}, \mathfrak{p}, \mathfrak{m}, \mathfrak{u}_\mathfrak{p}$ the Lie algebras of G, B, T, P, M, U_P respectively. For each $w \in W$, we let $P_w := w^{-1}Pw$, with Lie algebra $\mathfrak{p}_w := w^{-1}\mathfrak{p}w$, and similarly we define $\mathfrak{u}_{\mathfrak{p}_w}$, $\mathfrak{m}_w$, and so on. For $w \in {}^M W$, $\mathfrak{b}_{\mathfrak{m}_w} := \mathfrak{b} \cap \mathfrak{m}_w$ is a Borel in $\mathfrak{m}_w$.

4.1 Modular forms and the Hodge–Tate period map We fix throughout a tame level $K^p \subset \operatorname{GSp}(\mathbf{A}^{\infty,p})$ (which we ultimately choose in order to guarantee that various spaces of p-adic modular forms are 1-dimensional, using the theory of newforms developed in [37]), and for each open compact subgroup $K_p \subset \operatorname{GSp}_4(\mathbf{Q}_p)$, we let $\operatorname{Sh}_{K_p}/\mathbf{C}_p$ be the analytic adic space attached to (a toroidal compactification of) the Siegel Shimura variety of level $K_p K^p$.

For each $\kappa \in X^*(T)^{M,+}$ and each finite level Shimura variety Sh_{K_p}, we have the usual sheaf ω^κ of modular forms of weight κ on Sh_{K_p}. The coherent cohomology $\operatorname{R\Gamma}(\operatorname{Sh}_{K_p}, \omega^\kappa)$ has an action of a Hecke algebra $\mathbf{T}$ (generated by the Hecke operators at the places where K^p is hyperspecial), and the corresponding eigenclasses can be computed in terms of automorphic forms on G. More precisely, cuspidal automorphic representations contribute according to their Archimedean components π_∞, and the essentially tempered π_∞ which contribute to coherent components are the so-called non-degenerate limits of discrete series. The upshot for us is that in order to prove that a Galois representation is modular, it suffices to show that its corresponding system of Hecke eigenvalues contributes to some $\operatorname{R\Gamma}(\operatorname{Sh}_{K_p}, \omega^\kappa)$.

We let $\operatorname{Sh}_\infty = \lim_{K_p} \operatorname{Sh}_{K_p}$, which is a perfectoid space with an action of $\operatorname{Gal}_{\mathbf{Q}_p}$, admitting a $\operatorname{Gal}_{\mathbf{Q}_p}$-equivariant Hodge–Tate period map

$$\pi_{\mathrm{HT}} : \operatorname{Sh}_\infty \to \mathcal{FL},$$

where $\mathcal{FL}$ is the base change to $\mathbf{C}_p$ of the (partial) flag variety $P \backslash G$. This was introduced by Scholze [41], and has revolutionised the study of p-adic modular forms. As a first illustration of this, let $\mathcal{L}_\kappa$ be the G-equivariant sheaf on $\mathcal{FL}$ whose fibre at e is the inflation from M to P of the irreducible representation of M of highest weight κ, and we set $\omega^{\kappa,\mathrm{sm}} := (\pi_{\mathrm{HT}}^* \mathcal{L}_\kappa)^{\mathrm{sm}}$. By the definition of π_{HT}, one finds that the sheaf $\omega^{\kappa,\mathrm{sm}}$ descends to ω^κ on each Sh_{K_p}. Thus $\operatorname{R\Gamma}(\operatorname{Sh}_\infty, \omega^{\kappa,\mathrm{sm}})$ is a complex of smooth admissible $G(\mathbf{Q}_p)$-representations, equal to $\operatorname{colim}_{K_p} \operatorname{R\Gamma}(\operatorname{Sh}_{K_p}, \omega^\kappa)$. This suggests the possibility of proving results about modular forms by working on the flag variety $\mathcal{FL}$, an idea exploited to great effect by Scholze in [41].

Returning to the problem of the modularity of Galois representations, a basic difficulty now presents itself: there is no Galois action on $\operatorname{R\Gamma}(\operatorname{Sh}_\infty, \omega^{\kappa,\mathrm{sm}})$, thus no direct connection to Galois representations. This difficulty was resolved in the case of the modular curve by Pan [30], who invented geometric Sen theory and combined it with Scholze's ideas to prove remarkable new "p-adic Eichler–Shimura" results, relating the $\operatorname{R\Gamma}(\operatorname{Sh}_\infty, \omega^{\kappa,\mathrm{sm}})$ to étale cohomology groups, which naturally have an action of $\operatorname{Gal}_\mathbf{Q}$. Furthermore, as we will see below, Pan's theory often allows one to reduce to questions on the flag variety, and thus to explicit computations.

In fact, the spaces of classical modular forms $\operatorname{R\Gamma}(\operatorname{Sh}_\infty, \omega^{\kappa,\mathrm{sm}})$ do not directly show up in p-adic Eichler–Shimura theory; rather, one sees spaces of p-adic modular forms which are defined using the flag variety, and are closely related to the higher Coleman theory of Boxer–Pilloni [9]. From our point of view, this is a feature rather than a bug: it is these spaces of p-adic modular forms to which we can apply the Taylor–Wiles method in order to prove our modularity lifting theorems, and we then consider separately the problem of proving the classicality of a p-adic modular form. This classicality problem is solved by a generalisation of Pan's ideas, in combination with results of Boxer–Pilloni.

4.2 Completed cohomology By Scholze's primitive comparison theorem, the cohomology of the structure sheaf $\operatorname{R\Gamma}_{\mathrm{an}}(\operatorname{Sh}_\infty, \mathcal{O}_{\operatorname{Sh}_\infty})$ is naturally identified with $\widetilde{\operatorname{R\Gamma}}(\operatorname{Sh}_\infty, \mathbf{Q}_p) \otimes_{\mathbf{Q}_p} \mathbf{C}_p$, where

$$\widetilde{\operatorname{R\Gamma}}(\operatorname{Sh}_\infty, \mathbf{Q}_p) := \lim_n \operatorname{colim}_{K_p} \operatorname{R\Gamma}(\operatorname{Sh}_{K_p}, \mathbf{Z}_p/p^n) \otimes_{\mathbf{Z}_p} \mathbf{Q}_p$$

denotes Emerton's completed étale cohomology (with $\mathbf{Q}_p$ coefficients). As well as the action of $\mathrm{Gal}_{\mathbf{Q}_p}$ and of the Hecke algebra $\mathbf{T}$, this has an action of $\mathrm{GSp}_4(\mathbf{Q}_p)$, and we can consider the (derived) locally analytic vectors $\widetilde{\mathrm{R\Gamma}}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}} \otimes_{\mathbf{Q}_p} \mathbf{C}_p$, which have an action of the Lie algebra $\mathfrak{g}$. We can define subsheaves of smooth and locally analytic vectors for the $\mathrm{GSp}_4(\mathbf{Q}_p)$-action:

$$\mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{sm}} \subseteq \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}} \subseteq \mathcal{O}_{\mathrm{Sh}_\infty},$$

and by results of Rodríguez Camargo (following Pan), we have a natural identification

$$(4.1) \qquad \widetilde{\mathrm{R\Gamma}}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}} \otimes_{\mathbf{Q}_p} \mathbf{C}_p = \mathrm{R\Gamma}_{\mathrm{an}}(\mathrm{Sh}_\infty, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}}).$$

Fix $\lambda \in X^*(T)$, and write M_λ for the Verma module for $\mathfrak{g}$ with highest weight λ. By (4.1) we have

$$(4.2) \qquad \mathrm{RHom}_{\mathfrak{b}}(\lambda, \widetilde{\mathrm{R\Gamma}}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}}) \otimes_{\mathbf{Q}_p} \mathbf{C}_p = \mathrm{R\Gamma}_{\mathrm{an}}(\mathrm{Sh}_\infty, \mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}})).$$

We think of the left hand side of (4.2) as a space of p-adic modular forms of weight λ; the goal of p-adic Eichler–Shimura theory is to understand it in terms of the $\mathrm{R\Gamma}(\mathrm{Sh}_\infty, \omega^{\kappa, \mathrm{sm}})$.

The formula (4.2) motivates the following definitions. Write $\mathcal{O}(\mathfrak{g}, \mathfrak{b})$ for the usual category $\mathcal{O}$ of finitely generated left $U(\mathfrak{g})$-modules with locally finite $\mathfrak{b}$-action and semisimple $\mathfrak{h}$-action, and $\mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}}$ for the full subcategory of objects all of whose weights are in $X^*(T)$; this category contains the Verma modules of integral highest weight. The action of $\mathfrak{b}$ on an object of $\mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}}$ can naturally be integrated to an action of B, so from now on we regard these objects as $(\mathfrak{g}, B)$-modules (and we we regard the left hand side of 4.2 as a smooth $B(\mathbf{Q}_p)$-representation). Writing $\mathrm{Mod}_{B(\mathbf{Q}_p)}^{\mathrm{sm}}(\mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{sm}})$ for the derived category of solid $\mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{sm}}$-modules with a smooth action of $B(\mathbf{Q}_p)$, the right hand side of (4.2) suggests that we should consider the functor

$$(4.3) \qquad \begin{aligned} \mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}} &\to \mathrm{Mod}_{B(\mathbf{Q}_p)}^{\mathrm{sm}}(\mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{sm}}), \\ M &\mapsto \mathrm{RHom}_{\mathfrak{g}}(M, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}}). \end{aligned}$$

4.3 Localisation to the partial flag variety A key point is that the functor $\mathrm{RHom}_{\mathfrak{g}}(-, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}})$ factors through a natural analogue of Beilinson–Bernstein localisation for the partial flag variety $\mathcal{FL}$, as we now explain (see Theorem 4.1 below). From now on we regard G an an affinoid analytic group over $\mathbf{C}_p$. There is a natural action of G on $\mathcal{FL}$ by right multiplication, which induces an action of $\mathfrak{g}$ by derivations on $\mathcal{O}_{\mathcal{FL}}$. We have a filtration of G-equivariant coherent sheaves: $\mathfrak{u}_{\mathfrak{p}}^0 \subseteq \mathfrak{p}^0 \subseteq \mathfrak{g}^0 = \mathcal{O}_{\mathcal{FL}} \otimes \mathfrak{g}$, whose fibres at a point $x \in \mathcal{FL}$ are $\mathfrak{u}_{\mathfrak{p}_x} = x^{-1}\mathfrak{u}_{\mathfrak{p}}x \subseteq \mathfrak{p}_x = x^{-1}\mathfrak{p}x \subseteq \mathfrak{g}$. We let G_r be the analytic subgroup of G consisting of elements reducing to the identity e modulo p^r, and we set $\mathcal{O}_{G,e} := \mathrm{colim}_r \mathcal{O}_{G_r}$. Then we define $\hat{U}(\mathfrak{g}) := \mathcal{O}_{G,e}^\vee$; this is a completion of the universal enveloping algebra $U(\mathfrak{g})$. We also write $G_e = \lim_r G_r$, the limit being taken in the category of locally ringed spaces; this has a single point e (the identity element of G), with structure sheaf $\mathcal{O}_{G,e}$.

We now define the ring of universal twisted differential operators

$$\tilde{\mathcal{D}}^{\mathrm{la}} = (\mathcal{O}_{\mathcal{FL}} \otimes \hat{U}(\mathfrak{g}))/\mathfrak{u}_{\mathfrak{p}}^0(\mathcal{O}_{\mathcal{FL}} \otimes \hat{U}(\mathfrak{g})),$$

and we let $\mathcal{C}^{\mathrm{la}} = (\mathcal{O}_{G,e} \otimes \mathcal{O}_{\mathcal{FL}})^{\mathfrak{u}_{\mathfrak{p}}^0}$ where the invariants are for the diagonal action on the two factors, where the action on $\mathcal{O}_{G,e}$ is by left translation. Then $\mathcal{C}^{\mathrm{la}}$ is a $\tilde{\mathcal{D}}^{\mathrm{la}}$-module, and it carries an action of $\mathfrak{g}$ (by right translation on $\mathcal{O}_{G,e}$) which commutes with the $\tilde{\mathcal{D}}^{\mathrm{la}}$-module structure. We define a localisation functor from the (derived) category of solid $U(\mathfrak{g})$-modules to the (derived) category of solid $\tilde{\mathcal{D}}^{\mathrm{la}}$-modules (i.e. twisted D-modules):

$$\begin{aligned} \mathrm{Loc} : \mathrm{Mod}(U(\mathfrak{g})) &\to \mathrm{Mod}(\tilde{\mathcal{D}}^{\mathrm{la}}) \\ M &\mapsto \mathrm{RHom}_{\mathfrak{g}}(M, \mathcal{C}^{\mathrm{la}}). \end{aligned}$$

For any M, the sheaf $\mathrm{Loc}(M)$ admits a natural "horizontal" action of the centre $Z(\mathfrak{m})$ of the universal enveloping algebra $U(\mathfrak{m})$, via the natural injection $\mathcal{O}_{\mathcal{FL}} \otimes Z(\mathfrak{m}) \hookrightarrow U(\mathfrak{m}^0)$ (with $\mathfrak{m}^0 = \mathfrak{g}^0/\mathfrak{u}_{\mathfrak{p}}^0$), and in particular it admits an action of $\mu \in Z(\mathfrak{m})$. If $M \in \mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}}$ then we regard $\mathrm{Loc}(M)$ as a $(\mathfrak{g}, B)$-equivariant sheaf on $\mathcal{FL}$; the B-equivariant structure comes from the action of B on M, while the $\mathfrak{g}$-action comes from the $\tilde{\mathcal{D}}^{\mathrm{la}}$-module structure.

The following theorem is a consequence of the geometric version of Sen theory introduced by Pan and generalized by Rodríguez Camargo. Its formulation is a generalisation of Pilloni's interpretation [33] of Pan's work for the modular curve.

THEOREM 4.1. *The functor* (4.3) *factors through the functor* Loc; *more precisely, for* $M \in \mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}}$ *we have*

$$\mathrm{RHom}_{\mathfrak{g}}\big(M, \mathcal{O}^{\mathrm{la}}_{\mathrm{Sh}_\infty}\big) = \big(\pi^*_{\mathrm{HT}} \mathrm{Loc}(M)\big)^{\mathrm{sm}},$$

and

$$(4.4) \qquad \mathrm{RHom}_{\mathfrak{g}}(M, \widetilde{\mathrm{R\Gamma}}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}}) \otimes_{\mathbf{Q}_p} \mathbf{C}_p = \mathrm{R\Gamma}_{\mathrm{an}}(\mathrm{Sh}_\infty, \mathrm{RHom}_{\mathfrak{g}}\big(M, \mathcal{O}^{\mathrm{la}}_{\mathrm{Sh}_\infty}\big)).$$

Furthermore, the action of $\mu \in Z(\mathfrak{m})$ *on* $\mathrm{Loc}(M)$ *via the horizontal action induces the Sen operator on the left hand side of* (4.4).

In view of Theorem 4.1, we will sometimes refer to the horizontal action of μ as "the Sen operator" in the following.

4.4 p-adic Eichler–Shimura To go further, we use excision with respect to the Bruhat stratification of $\mathcal{FL}$, i.e. the decomposition into B-orbits $\mathcal{FL} = P\backslash G = \coprod_{w \in {}^M W} P\backslash PwB$, indexed by the Kostant representatives ${}^M W$; the dimension of $P\backslash PwB$ is $\ell(w)$. We write $j_w : C_w = P\backslash PwB \hookrightarrow \mathcal{FL}$ for the locally closed immersion of the Bruhat cell corresponding to w, and $C^\dagger_w$ for the dagger neighbourhood of C_w in $\mathcal{FL}$. We write $j_{w, \mathrm{Sh}_\infty} : \pi^{-1}_{\mathrm{HT}}(C_w) \hookrightarrow \mathrm{Sh}_\infty$ for the morphism induced by j_w.

For each $w \in {}^M W$, we can consider the composite of Loc with restriction to $C^\dagger_w$. Let M be an object of $\mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}}$, so that $\mathrm{Loc}(M)$ is a $(\mathfrak{g}, B)$-equivariant sheaf; since C_w is a B-orbit, and $\mathrm{Loc}(M)$ is B-equivariant, the sheaf $\mathrm{Loc}(M)|_{C^\dagger_w}$ is determined by its fibre at w, which is a representation of the stabiliser of w for the $(\mathfrak{g}, B)$-action. This stabiliser admits an explicit description as follows. The action of $(\mathfrak{g}, B)$ can be upgraded to an action of the semi-direct product $G_e \rtimes B$, which acts on $C^\dagger_w$ via $(g, b) \mapsto wgb$. Write $\mathrm{Stab}(w)$ for the stabiliser of w for this action; then it follows from the definitions that the map $(g, b) \mapsto (gb, b)$ is an injection

$$\mathrm{Stab}(w) \hookrightarrow P_w \times B.$$

It is then elementary to check that $\mathrm{Stab}(w)$ is generated by its subgroups $\mathrm{Stab}(w)_e = P_{w,e} \times B_e$ and $P_w \cap B \subseteq B$, and in fact

$$(4.5) \qquad (P_w \cap B)_e \backslash (P_{w,e} \times B_e) \rtimes (P_w \cap B).$$

The explicit description (4.5) allows us to construct a contravariant functor

$$HCS_w : \mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})_{\mathrm{alg}} \to \mathrm{Mod}_{(\mathfrak{g}, B)}(C^\dagger_w)^{\mathfrak{u}^0_\mathfrak{p}}$$

where $\mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})_{\mathrm{alg}}$ is the algebraic category $\mathcal{O}$ for the pair $(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})$, as defined above, and the target category has the obvious meaning. (Here "HCS" stands for "higher Coleman sheaf".) This functor is defined as follows: the action of $\mathfrak{b}_{\mathfrak{m}_w}$ on $V \in \mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})_{\mathrm{alg}}$ integrates to an action of $B_{M_w} = M_w \cap B$, and thus determines an action of $P_w \cap B$. The admissible $\hat{U}(\mathfrak{g})$-module $(V \otimes_{U(\mathfrak{g})} \hat{U}(\mathfrak{g}))^\vee$ therefore has actions of $P_w \cap B$ and of $P_{w,e}$ (the latter action factoring through the action of $M_{w,e}$, which comes from the action of $\mathfrak{m}_w$), and one checks that by allowing B_e to act trivially, we obtain an action of $\mathrm{Stab}(w)$. The corresponding $(\mathfrak{g}, B)$-equivariant sheaf is $HCS_w(V)$.

The reason for introducing this functor is that there is a commutative diagram

$$(4.6) \qquad \begin{array}{ccc} \mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}} & \xrightarrow{j^{-1}_w \mathrm{Loc}} & \mathrm{Mod}_{(\mathfrak{g}, B)}(C^\dagger_w)^{\mathfrak{u}^0_\mathfrak{p}} \\ {}_{\mathbf{C}_p \otimes^L_{\mathfrak{u}_{\mathfrak{p}_w}} -}\Big\downarrow & \nearrow_{HCS_w} & \\ \mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})_{\mathrm{alg}} & & \end{array}$$

where the functor $\mathbf{C}_p \otimes^L_{\mathfrak{u}_{\mathfrak{p}_w}} - : \mathrm{Mod}(U(\mathfrak{u}_{\mathfrak{p}_w})) \to \mathrm{Mod}(\mathbf{C}_p)$ is the functor of "Lie algebra homology", with $H_i(\mathfrak{u}_{\mathfrak{p}_w}, M) = H^{-i}(\mathbf{C}_p \otimes_{U(\mathfrak{g})} M)$. To see that (4.6) commutes, one only has to check that the fibres at w of the $(\mathfrak{g}, B)$-equivariant sheaves are isomorphic representations of $\mathrm{Stab}(w)$. This boils down to noting that the fibre $\mathcal{C}^{\mathrm{la}}_w$ is $\mathcal{O}_{U_{P_w}\backslash G, e}$, with the action of $\mathrm{Stab}(w)$ being determined by the action of $P_{w,e}$ by left translation, the action of B_e by right translation, and the action of $P_w \cap B$ by conjugation.

We now return to the cohomology of Shimura varieties. For any sheaf $\mathcal{F}$ on a subset of Sh_∞ containing $\pi^{-1}_{\mathrm{HT}}(\mathcal{C}_w)$, we write

$$\mathrm{R\Gamma}_w(\mathrm{Sh}_\infty, \mathcal{F}) := \mathrm{R\Gamma}(\mathrm{Sh}_\infty, j_{w,\mathrm{Sh},!}\mathcal{F}|_{\pi^{-1}_{\mathrm{HT}}(\mathcal{C}_w)}),$$

where $j_{w,\mathrm{Sh},!}$ is the extension by zero functor on abelian sheaves of solid abelian groups. Then we define the functor

$$(4.7) \qquad \begin{aligned} HC_w : \mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})_{\mathrm{alg}} &\to \mathrm{Mod}^{\mathrm{sm}}_{B(\mathbf{Q}_p)}(\mathbf{C}_p), \\ M &\mapsto \mathrm{R}\Gamma_w(\mathrm{Sh}_\infty, (\pi_{\mathrm{HT}}^* HCS_w(M))^{\mathrm{sm}}). \end{aligned}$$

Combining (4.6) with Theorem 4.1 and the spectral sequence for a filtered complex, we obtain the following theorem, a basic form of the p-adic Eichler–Shimura decomposition.

THEOREM 4.2. *For any $M \in \mathcal{O}(\mathfrak{g}, \mathfrak{b})_{\mathrm{alg}}$, we have a spectral sequence:*

$$E_1^{p,q} = \oplus_{w \in {}^M W, \ell(w)=p} H^{p+q}(HC_w(\mathbf{C}_p \otimes^L_{\mathfrak{u}_{\mathfrak{p}_w}} M))$$

converging to $H^{p+q}(\mathrm{RHom}_{\mathfrak{g}}(M, \widetilde{\mathrm{R}\Gamma}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}})) \otimes_{\mathbf{Q}_p} \mathbf{C}_p$. The Sen operator is induced by the action of $w\mu \in Z(\mathfrak{m}_w)$ on $H_*(\mathfrak{u}_{\mathfrak{p}_w}, M)$.

4.5 The ordinary part Let $T^+(\mathbf{Q}_p) := \{t \in T(\mathbf{Q}_p), \ \forall \alpha \in \Phi^+, \ v_p(\alpha(t)) \geq 0\}$. Given a smooth representation V of $B(\mathbf{Q}_p)$, there is a Hecke action of $T^+(\mathbf{Q}_p)$ on $V^{U(\mathbf{Z}_p)}$; as usual, the action of $t \in T^+(\mathbf{Q}_p)$ is defined via the (normalised) trace $V^{tU(\mathbf{Z}_p)t^{-1}} \to V^{U(\mathbf{Z}_p)}$. The finite slope part V^{fs} of V is by definition the subspace on which $T^+(\mathbf{Q}_p)$ acts invertibly. For any $\lambda \in X^*(T)_{\mathbf{R}}$, we say that the slopes of V^{fs} are at least λ (respectively, are equal to λ) if for every $t \in T^+(\mathbf{Q}_p)$ and every eigenvalue α of t acting on V^{fs}, we have $v_p(\alpha) \geq v_p(\lambda(t))$ (respectively, $v_p(\alpha) = v_p(\lambda(t))$). The spaces $HC_w(M)^{fs}$ are very closely related to the higher Coleman theories of Boxer–Pilloni (see [8, Thm. 4.6.56] for a precise statement). The following key slope bound was proved by Boxer–Pilloni (it is essentially [10, Cor. 6.2.16]). It is proved by a careful analysis of integral models of Hecke correspondences (and in particular, it is not deduced from a statement on the flag variety). We will see below that it immediately implies classicality theorems and p-adic Eichler–Shimura decompositions for the ordinary part of the cohomology (and more generally for the "small slope" part, although we do not discuss that here).

THEOREM 4.3. *Let $w \in {}^M W$, and let $M \in \mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})_{\mathrm{alg}}$ be a module generated by a highest weight vector of weight ν. Then the slopes of $HC_w(M)^{fs}$ are at least $-\nu + w^{-1} w_{0,M} \rho + \rho$.*

We fix from now on a character $\lambda \in X^*(T)^{M,+}$, and assume that V is a smooth representation of $B(\mathbf{Q}_p)$ such that the slopes of V^{fs} are at least $-\lambda$. Then the *ordinary part* V^{ord} of V is by definition the subspace of V^{fs} whose slopes are equal to $-\lambda$.

In particular, by Theorem 4.3, the slopes of $HC_w(M_\lambda)^{fs}$ are all at least $-\lambda$, so we may consider the ordinary part $HC_w(M_\lambda)^{\mathrm{ord}}$. Write $d = \dim \mathfrak{u}_{\mathfrak{p}} = \dim \mathrm{Sh}_{K_p}$. The homology groups $H_i(\mathfrak{u}_{\mathfrak{p}_w}, M_\lambda)$ all belong to $\mathcal{O}(\mathfrak{m}_w, \mathfrak{b}_{\mathfrak{m}_w})$, and they can (at least in principle) be computed by a Chevalley–Eilenberg complex. Write $L(\mathfrak{m}_w)_\nu$ for the simple quotient of the Verma module for $\mathfrak{m}_w$ of highest weight ν. Then an analysis of this complex shows that $H_{d-\ell(w)}(\mathfrak{u}_{\mathfrak{p}_w}, M_\lambda)$ has a unique subquotient isomorphic to $L(\mathfrak{m}_w)_{\lambda + w^{-1} w_{0,M} \rho + \rho}$, and that every other Jordan–Hölder factor of any $H_i(\mathfrak{u}_{\mathfrak{p}_w}, M_\lambda)$ is generated by a highest weight vector of the form

$$(4.8) \qquad \lambda + w^{-1} w_{0,M} \rho + \rho - \sum_{\alpha \in \Phi^+} n_\alpha \alpha,$$

where $n_\alpha \in \mathbf{Z}_{\geq 0}$ and $n_\alpha > 0$ for some α. Applying Theorem 4.3 to (4.8), it follows that for any such Jordan–Hölder factor, say X, the slopes of $HC_w(X)^{fs}$ are at least $-\lambda + \sum_{\alpha \in \Phi^+} n_\alpha \alpha > -\lambda$, so that $HC_w(X)^{\mathrm{ord}} = 0$. Going back to Theorem 4.2, we see that the slopes of $H^{p+q}(\mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \widetilde{\mathrm{R}\Gamma}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}})) \otimes_{\mathbf{Q}_p} \mathbf{C}_p$ are all at least $-\lambda$, and that there is a spectral sequence

$$(4.9) \qquad E_1^{p,q} = \oplus_{w \in {}^M W, \ell(w)=p} H^{2p+q-d}(HC_w(L(\mathfrak{m}_w)_{\lambda + w^{-1} w_{0,M} \rho + \rho}))^{\mathrm{ord}}$$

converging to $H^{p+q}(\mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \widetilde{\mathrm{R}\Gamma}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}}))^{\mathrm{ord}} \otimes_{\mathbf{Q}_p} \mathbf{C}_p$.

Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T}$ with corresponding Galois representation $\overline{\rho}_{\mathfrak{m}} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\overline{\mathbf{F}}_p)$. We say that $\mathfrak{m}$ is *non-Eisenstein* if $\overline{\rho}_{\mathfrak{m}}$ is irreducible. Then a comparison of the cuspidal and non-cuspidal versions of HC_w, and an analysis of the cohomology of the boundary of the toroidal compactification (which we have only carried

out for $G = \mathrm{GSp}_4$, although we expect the analogous results to hold more generally) together show that after localising at a non-Eisenstein $\mathfrak{m}$, each HC_w only has cohomology in degree $\ell(w)$, and $H^i(\mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}}))_{\mathfrak{m}}^{\mathrm{ord}}$ vanishes outside of degree $i = 3$ (i.e. outside of middle degree). Thus the spectral sequence (4.9) degenerates, proving the following theorem.

THEOREM 4.4. *Suppose that* $\lambda \in X^*(T)$, *and that* $\mathfrak{m}$ *is non-Eisenstein. Then*

$$\mathrm{RHom}_{\mathfrak{b}}(\lambda, \widetilde{\mathrm{R}\Gamma}(\mathrm{Sh}_\infty, \mathbf{Q}_p)^{\mathrm{la}})_{\mathfrak{m}}^{\mathrm{ord}} \otimes_{\mathbf{Q}_p} \mathbf{C}_p = \mathrm{R}\Gamma_{\mathrm{an}}(\mathrm{Sh}_\infty, \mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}}))_{\mathfrak{m}}^{\mathrm{ord}}$$

is concentrated in degree 3, *and there is a decreasing filtration on* $H^3(\mathrm{Sh}_\infty, \mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \mathcal{O}_{\mathrm{Sh}_\infty}^{\mathrm{la}}))_{\mathfrak{m}}^{\mathrm{ord}}$ *with ith graded piece given by*

$$(4.10) \qquad \oplus_{w \in {}^M W, \ell(w) = i} H^i(HC_w(L(\mathfrak{m}_w)_{\lambda + w^{-1} w_{0,M} \rho + \rho}))_{\mathfrak{m}}^{\mathrm{ord}}.$$

Under appropriate dominance hypotheses, the summands in (4.10) can be described in terms of the sheaves $\omega^{\kappa, \mathrm{sm}}$. Indeed, recall that $\mathcal{L}_\kappa$ is the G-equivariant sheaf on $\mathcal{FL}$ corresponding to the irreducible representation of M of highest weight κ. Unwinding the definitions, one finds that for each $w \in {}^M W$, we have

$$(4.11) \qquad \mathcal{L}_\kappa|_{\mathcal{C}_w^\dagger} = HCS_w(L(\mathfrak{m}_{\mathfrak{w}})_{-w^{-1} w_{0,M} \kappa}).$$

Set

$$\kappa_w := -w_{0,M} w(\lambda + \rho) - \rho,$$

so that

$$-w^{-1} w_{0,M} \kappa_w = \lambda + w^{-1} w_{0,M} \rho + \rho.$$

Provided that $\kappa_w \in X^*(T)^{M,+}$, it follows from (4.11) that

$$\omega^{\kappa_w, \mathrm{sm}}|_{\pi_{HT}^{-1} C_w^\dagger} = (\pi_{HT}^* HCS_w(L(\mathfrak{m}_w)_{\lambda + w^{-1} w_{0,M} \rho + \rho}))^{\mathrm{sm}},$$

and (by (4.7), i.e. by the definition of HC_w) the contribution from w to (4.10) is

$$(4.12) \qquad H^i(HC_w(L(\mathfrak{m}_w)_{\lambda + w^{-1} w_{0,M} \rho + \rho}))_{\mathfrak{m}}^{\mathrm{ord}} = H_w^i(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}}.$$

Additionally, considering the horizontal action of μ, we see that the Sen operator acts on $\omega^{\kappa_w, \mathrm{sm}}|_{\pi_{HT}^{-1} C_w^\dagger}$ via $\langle \mu, \kappa_w \rangle$.

We now consider the difference between the spaces of ordinary p-adic modular forms $H_w^i(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}}$ and of ordinary classical modular forms $H^i(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}}$. Bearing in mind (4.12), the slope bound of Theorem 4.3 shows that for each pair $v, w \in {}^M W$, the slopes of $H_v^i(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})^{fs}$ are at least

$$(4.13) \qquad -\lambda + ((\lambda + \rho) - v^{-1} w(\lambda + \rho)).$$

If we assume that $\lambda + \rho \in X^*(T)^+$, we see that these slopes are at least $-\lambda$, and that $H_v^i(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}}$ vanishes unless $v^{-1} w(\lambda + \rho) = (\lambda + \rho)$; equivalently, unless $\kappa_v = \kappa_w$. Accordingly, we now consider the subgroup W_λ of W consisting of those w' with $w'(\lambda + \rho) = (\lambda + \rho)$. This subgroup corresponds to a standard parabolic subgroup Q of G; for example, if $\lambda \in X^*(T)^+$, then this subgroup is trivial and $Q = B$. Then $P \backslash PwQ$ is the union of those Bruhat cells C_v for which $H_v^i(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}}$ could be nonzero, and we let $j_{w,Q} : P \backslash PwQ \hookrightarrow \mathcal{FL}$, $j_{w,Q,\mathrm{Sh}_\infty} : \pi_{HT}^{-1}(P \backslash PwQ) \hookrightarrow \mathrm{Sh}_\infty$ be the corresponding locally closed immersions. Writing

$$\mathrm{R}\Gamma_{w,Q}(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}}) := \mathrm{R}\Gamma(\mathrm{Sh}_\infty, j_{w,Q,\mathrm{Sh},!} \omega^{\kappa_w, \mathrm{sm}}|_{\pi_{HT}^{-1}(P \backslash PwQ)}),$$

we obtain the following classicality result.

THEOREM 4.5. *Suppose that* $\kappa_w \in X^*(T)^{M,+}$, *and* $\lambda + \rho \in X^*(T)^+$. *Then*

$$\mathrm{R}\Gamma_{w,Q}(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}} = \mathrm{R}\Gamma(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}}.$$

Proof. It suffices to note that for each $v \notin wW_\lambda$, we have $\mathrm{R}\Gamma_v(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})_{\mathfrak{m}}^{\mathrm{ord}} = 0$ by (4.13). $\qquad \square$

Remark 4.6. The cohomology $R\Gamma_{w,Q}(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$ is supported in degrees $[\ell(w_{\min}), \ell(w_{\max})]$, where $w_{\min}, w_{\max}$ are respectively the minimal and maximal length representatives in $^M W$ for the double coset $W_M w W_\lambda$.

Remark 4.7. As noted above, if $\lambda \in X^*(T)^+$ then $Q = B$, so that $H^{\ell(w)}_{w,Q}(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}}) = H^{\ell(w)}_w(\mathrm{Sh}_\infty, \omega^{\kappa_w, \mathrm{sm}})$, and the Eichler–Shimura decomposition in Theorem 4.4 takes a particularly simple form. However, a stronger result (without taking ordinary parts or making a non-Eisenstein localisation) was already known in this case (going back to Faltings–Chai [22] in the Siegel case); from our perspective, it can be obtained by replacing M_λ above with V_λ, the algebraic representation of G of highest weight λ. Then $\mathrm{Loc}(V_\lambda)$ is G-equivariant, rather than merely B-equivariant, and one can compute directly on the whole of $\mathcal{FL}$, without needing to restrict to Bruhat strata.

4.6 Sen equals Cousin We now make things more explicit in the case of interest to us, where λ is not regular, i.e. $\lambda \notin X^*(T)^+$. We now introduce some more explicit notation for $G = \mathrm{GSp}_4$. We label the elements of T by $t = \mathrm{diag}(zt_1, zt_2, zt_2^{-1}, zt_1^{-1})$, and the characters $X^*(T)$ of T by tuples $\kappa = (k_1, k_2; w) \in \mathbf{Z}^2 \times \mathbf{Z}$ with $w \equiv k_1 + k_2 \pmod 2$, where $\kappa(t) = z^w t_1^{k_1} t_2^{k_2}$. Thus a character $(k_1, k_2; w)$ is M-dominant if $k_1 \geq k_2$, and G-dominant if $0 \geq k_1 \geq k_2$. We have $\rho = (-1, -2; 0)$.

The Weyl group W is generated by s_α and s_β where $s_\alpha(k_1, k_2; w) = (k_2, k_1; w)$ and $s_\beta(k_1, k_2; w) = (-k_1, k_2; w)$, so that $w_0 = s_\alpha s_\beta s_\alpha s_\beta$, and $w_0(k_1, k_2; w) = (-k_1, -k_2; w)$. We have $W_M = \{\mathrm{Id}, w_{0,M} = s_\alpha\}$. The elements of $^M W$ are $^0 w = \mathrm{Id}, {}^1 w = s_\beta, {}^2 w = s_\beta s_\alpha, {}^3 w = s_\beta s_\alpha s_\beta$, where $\ell({}^i w) = i$. From now on we take $\lambda = (1, 1; -2)$, so that $\lambda + \rho = (0, -1; -2)$. Then $W_\lambda = \{1, s_\beta\}$, and Q is the Klingen parabolic. We have two Q-orbits on $\mathcal{FL}$, namely $C_{^0 w} \cup C_{^1 w}$ and $C_{^2 w} \cup C_{^3 w}$, with the corresponding κ_w being $(2, 2; 2)$ and $(1, 1; 2)$ respectively. This is illustrated in Figure 4.1 (which draws a character $(k_1, k_2; w)$ at (k_1, k_2)).

In our earlier paper [7], the input to our modularity lifting theorem is a class in $H^0(\mathrm{Sh}_\infty, \omega^{(2,2;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$, i.e. an ordinary Siegel modular form of weight 2 (see Remark 2.6). Accordingly in that paper we work with $R\Gamma_{\mathrm{id}, Q}(\mathrm{Sh}_\infty, \omega^{(2,2;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$, which has cohomology in degrees $0, 1$, and by Theorem 4.5 agrees with the classical ordinary cohomology $R\Gamma(\mathrm{Sh}_\infty, \omega^{(2,2;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$.

However in [8] our input is a class in weight 3, rather than weight 2. Using Hida theory, we can produce a congruence to a class in $H^0_{\mathrm{id}}(\mathrm{Sh}_\infty, \omega^{(2,2;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$, and ultimately the modularity lifting machinery produces another class in the same cohomology group. We then need to understand when such a class extends to $H^0_{\mathrm{id}, Q}(\mathrm{Sh}_\infty, \omega^{(2,2;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$. We now briefly explain how we do this; in fact for technical reasons we found it more convenient in [8] to study the problem of extending from $H^2_{^3 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$ to $H^2_{^3 w, Q}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$, so we do the same here.

From now on we work on $C_{^3 w, Q} := C_{^3 w} \cup C_{^2 w}$. We have a short exact sequence of sheaves over $\pi_{HT}^{-1}(C_{^3 w, Q})$, corresponding to the stratification of $C_{^3 w, Q}$ into B-orbits, with $j_{^3 w, \mathrm{Sh}_\infty} : \pi_{HT}^{-1}(C_{^3 w}) \hookrightarrow \pi_{HT}^{-1}(C_{^3 w, Q})$:

$$(4.14) \qquad 0 \to (j_{^3 w, \mathrm{Sh}_\infty})_! \omega^{(1,1;2), \mathrm{sm}}\big|_{\pi_{HT}^{-1}(C_{^3 w})} \to \omega^{(1,1;2), \mathrm{sm}}\big|_{\pi_{HT}^{-1}(C_{^3 w, Q})} \to \omega^{(1,1;2), \mathrm{sm}}\big|_{\pi_{HT}^{-1}(C_{^2 w})} \to 0$$

Thus $R\Gamma_{^3 w, Q}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$ is computed by the following complex in degrees $2, 3$, where *Cous* is induced by the class of the extension (4.14):

$$(4.15) \qquad [H^2_{^2 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}} \xrightarrow{\mathrm{Cous}} H^3_{^3 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}].$$

On the other hand, as noted just below (4.12), the Sen operator Θ on $H^3(\mathrm{RHom}_{\mathfrak{g}}(M_\lambda, \mathcal{O}^{\mathrm{la}}_{\mathrm{Sh}_\infty}))^{\mathrm{ord}}_{\mathfrak{m}}$ acts by $\langle \mu, \kappa_{^2 w} \rangle = \langle \mu, \kappa_{^3 w} \rangle = 0$ on each of $H^2_{^2 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$ and $H^3_{^3 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$, so it induces a map:

$$(4.16) \qquad H^2_{^2 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}} \xrightarrow{\mathrm{Sen}} H^3_{^3 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}.$$
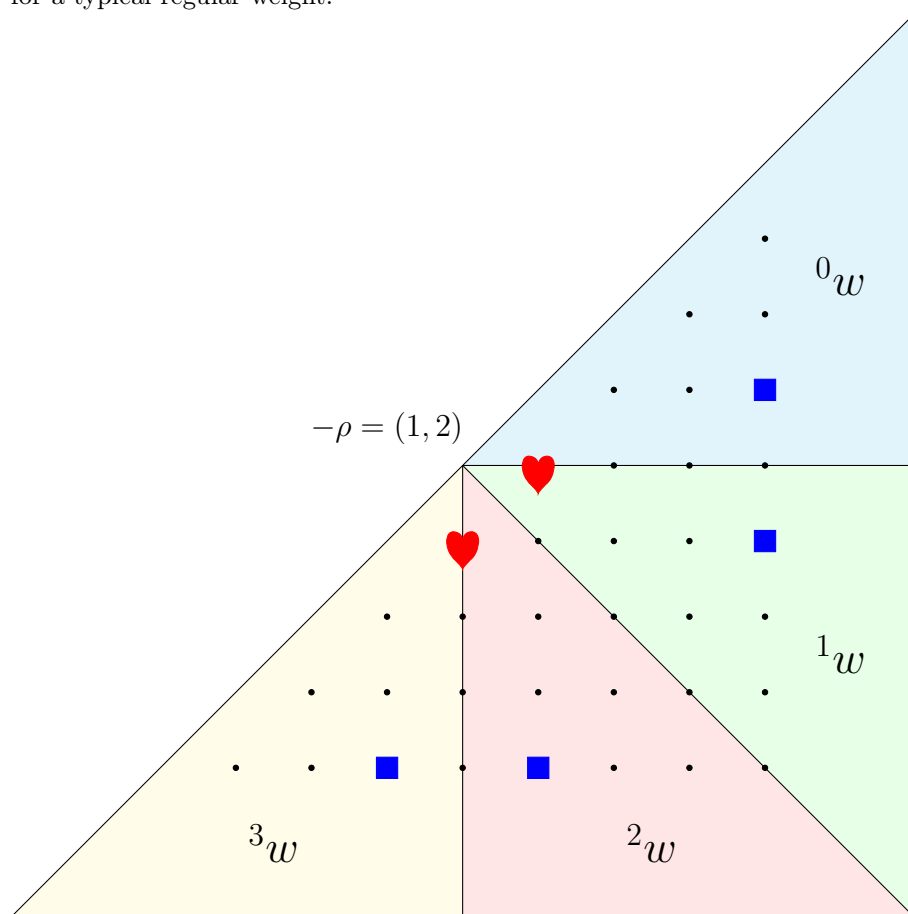
The key result is the following theorem, showing that "Sen equals Cousin". It is proved by an explicit calculation on the flag variety $\mathcal{FL}$.

THEOREM 4.8. *The two maps*

$$\mathrm{Cous}, \mathrm{Sen} : H^2_{^2 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}} \to H^3_{^3 w}(\mathrm{Sh}_\infty, \omega^{(1,1;2), \mathrm{sm}})^{\mathrm{ord}}_{\mathfrak{m}}$$

agree up to a non-zero scalar.

Figure 4.1: (Shifted) M-dominant Weyl chambers of weights (k_1, k_2). The G-dominant Weyl chamber is labelled by ${}^3w$. The red hearts are at $\kappa_{{}^3w} = \kappa_{{}^2w} = (1,1)$ and $\kappa_{{}^1w} = \kappa_{{}^0w} = (2,2)$ for our $\lambda = (1,1)$, while the blue squares represent the κ_w for a typical regular weight.

Combining Theorem 4.8 and (4.15), we see that a class $c \in H^2_{2_w}(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}_\mathfrak{m}$ extends to $H^2_{3_w,Q}(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}_\mathfrak{m}$ (equivalently, to $H^2(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}_\mathfrak{m}$) if and only if $\mathrm{Sen}(c) = 0$. It only remains to relate this condition to the Sen operator on the Galois representation associated to a p-adic modular form. We do this using the Eichler–Shimura relations; more precisely, we use the following lemma, which is deduced from results of Nekovář, [29].

LEMMA 4.9. *Let* $f \in H^0_{0_w}(\mathrm{Sh}_\infty, \omega^{(2,2;2),\mathrm{sm}})^{\mathrm{ord}}$ *be an ordinary overconvergent modular eigenform with corresponding Galois representation* $\rho_f : \mathrm{Gal}_\mathbf{Q} \to \mathrm{GSp}_4(\overline{\mathbf{Q}}_p)$. *Let* $\mathfrak{m}_f$ *be the corresponding maximal ideal of the Hecke algebra* $\mathbf{T}[1/p]$. *We assume that* $\overline{\rho}_f$ *is irreducible and the Zariski closure of* $\rho_f(\mathrm{Gal}_\mathbf{Q})$ *contains* Sp_4. *Then* $H^3(\mathrm{RHom}_\mathfrak{b}(\lambda, \mathrm{R\Gamma}(\mathrm{Sh}_\infty, \overline{\mathbf{Q}}_p)^{\mathrm{la}}))^{\mathrm{ord}}[\mathfrak{m}_f] = \rho_f \otimes_{\overline{\mathbf{Q}}_p} W$ *for some finite-dimensional vector space* $W \neq 0$.

Finally, we deduce from this the following classicality theorem.

THEOREM 4.10. *Let* $f \in H^0_{0_w}(\mathrm{Sh}_{K^p}, \omega^{(2,2;2),\mathrm{sm}})^{\mathrm{ord}}$ *be an ordinary overconvergent modular eigenform with Galois representation* $\rho_f : \mathrm{Gal}_\mathbf{Q} \to \mathrm{GSp}_4(\overline{\mathbf{Q}}_p)$. *Let* $\mathfrak{m}_f$ *be the corresponding maximal ideal. We assume that:*

(1) *the Zariski closure of* $\rho_f(\mathrm{Gal}_\mathbf{Q})$ *contains* Sp_4.

(2) *The representation* $\rho_f|_{\mathrm{Gal}_{\mathbf{Q}_p}}$ *is de Rham.*

(3) *We have* $\dim_{\mathbf{C}_p} H^i_{i_w}(\mathrm{Sh}_\infty, \omega^{(2,2;2),\mathrm{sm}})^{\mathrm{ord}}[\mathfrak{m}_f] = 1$ *for* $i = 0, 1$ *and* $\dim_{\mathbf{C}_p} H^i_{i_w}(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}[\mathfrak{m}_f] = 1$ *for* $i = 2, 3$.

(4) *The representation* $\overline{\rho}_f$ *is irreducible.*

Then f *is a classical Siegel modular form.*

Proof. Assumption (3) guarantees that the passage to the $\mathfrak{m}_f$ torsion is exact, so that we have induced maps

$$(4.17) \qquad \mathrm{Cous}, \mathrm{Sen} : H^2_{2_w}(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}[\mathfrak{m}_f] \to H^3_{3_w}(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}[\mathfrak{m}_f].$$

These two maps agree by Theorem 4.8. Since $\rho_f|_{\mathrm{Gal}_{\mathbf{Q}_p}}$ is de Rham, the corresponding Sen operator is semisimple (see Section 3.2), which by Lemma 4.9 implies that the maps (4.17) vanish. As noted above, this implies that $H^2(\mathrm{Sh}_\infty, \omega^{(1,1;2),\mathrm{sm}})^{\mathrm{ord}}[\mathfrak{m}_f] \neq 0$, which by Arthur's multiplicity formula implies that $H^0(\mathrm{Sh}_\infty, \omega^{(2,2;2),\mathrm{sm}})^{\mathrm{ord}}[\mathfrak{m}_f] \neq 0$, whence it is a one-dimensional $\mathbf{C}_p$-vector space spanned by f, as required. □

5 Modularity lifting Finally, we very briefly explain the use of modularity lifting theorems in our two papers. In [8], we needed to prove Theorem 2.1, which we deduce from Theorem 4.10. We work with the ordinary higher Coleman families $\mathrm{R\Gamma}_{i_w}(\mathrm{Sh}_\infty, \omega^{\kappa,\mathrm{sm}})^{\mathrm{ord}}$, or rather their integral versions, the higher Hida theories of [10]. Since for each i, the complex $\mathrm{R\Gamma}_{i_w}(\mathrm{Sh}_\infty, \omega^{\kappa,\mathrm{sm}})^{\mathrm{ord}}$ has cohomology only in degree i, a standard application of the usual Taylor–Wiles method is able to prove modularity lifting theorems for each of these higher Hida theories. (In practice we find it useful to work with Hida families in which the weight varies, for example in proving local-global compatibility, but this is again standard.) Furthermore, following Diamond [21], the Taylor–Wiles method is able to prove multiplicity one theorems for spaces of modular forms, and we use it to verify hypothesis (3) in Theorem 4.10. (It is here that we use the hypothesis of p-distinguishedness in assumption (iv) of Theorem 2.1.) The only significant difficulty that we have to overcome in applying the Taylor–Wiles method is that we have to consider residual Galois representations with rather small image, namely A_5 (when $p = 2$) and $\mathrm{GSp}_4(\mathbf{F}_3)$ (when $p = 3$). This causes us considerable pain, but does not involve any significant innovations.

As we explained above, in our earlier paper [7] we worked with the complexes $\mathrm{R\Gamma}_{\mathrm{id},Q}(\mathrm{Sh}_\infty, \omega^{(2,2;2),\mathrm{sm}})^{\mathrm{ord}}_\mathfrak{m}$, which have cohomology in degrees $0, 1$. Again, there is an integral version, the higher Hida theory introduced by Pilloni in [32]. Since there is cohomology in multiple degrees, the usual Taylor–Wiles method does not work in this case, and instead we use the version of the Taylor–Wiles method introduced by Calegari–Geraghty [17]. Our main difficulty now is that rather than working over $\mathbf{Q}$, we are over an arbitrary totally real field F of degree d, say, and the analogous complexes (constructed from the cohomology of Hilbert–Siegel Shimura varieties) now have cohomology in degrees $0, \ldots, d$; but we do not know how to prove the local-global compatibility results needed for the Calegari–Geraghty method unless $d = 1$. We get around this difficulty by working only with primes p which split completely in F, and for each $v | p$, considering complexes which at v behave like $\mathrm{R\Gamma}_{\mathrm{id},Q}$, and at the other places above p behave like $\mathrm{R\Gamma}_{\mathrm{id}}$. We are able to compare these complexes, and we show that an appropriately compatible set of cohomology classes combine to give a classical cohomology class.

6   Future directions It is hopefully clear that there is considerable scope for improving on Theorem 1.5. Firstly, by working with Hilbert–Siegel Shimura varieties, it should be possible to extend Theorem 1.5 to abelian surfaces over arbitrary totally real fields. By making a solvable base change to a totally real field with suitable behaviour at the primes above 2, such an improvement would allow us to remove Hypothesis (2) in that theorem, which arose due to the interaction between the local conditions at 2 and 3 in Lemma 2.4. Secondly, our theorems should not require the full strength of the ordinarity condition, but only a suitable "small slope" condition, which would allow us to relax Hypothesis (3) in Theorem 1.5. Finally, many of the arguments in [8] are written for more general Shimura data, and it would be interesting to prove classicality theorems beyond the case of GSp_4. We intend to pursue all of these ideas in future sequels to [8]. Finally, one other possible route to the modularity of abelian surfaces would be to deduce it from an appropriate version of Serre's conjecture for GSp_4. Such a deduction would be analogous to Khare's theorem [25], which showed that Serre's conjecture implies the Artin conjecture for odd 2-dimensional representations of $\mathrm{Gal}_{\mathbf{Q}}$. With this in mind, we use our theorems to prove the following implication (see [8, Lem. 10.4.1]).

LEMMA 6.1. *Suppose that for every residual representation:*

$$\overline{\rho} : \mathrm{Gal}_{\mathbf{Q}} \to \mathrm{GSp}_4(\mathbf{F}_p)$$

satisfying the following conditions:

(1) $\overline{\rho}$ has multiplier $\overline{\varepsilon}^{-1}$,

(2) $\overline{\rho}$ is absolutely irreducible, and

(3) the semi-simplification of $\overline{\rho}|_{\mathrm{Gal}_{\mathbf{Q}_p}}$ is a direct sum of characters,

there exists an ordinary classical Siegel modular form f for $\mathrm{GSp}_4/\mathbf{Q}$ of weight at least 3 such that

$$\overline{\rho}_{f,p} \cong \overline{\rho}.$$

Then all abelian surfaces $A/\mathbf{Q}$ are modular.

References

[1] F. ANDREATTA, A. IOVITA, AND V. PILLONI, *p-adic families of Siegel modular cuspforms*, Ann. of Math. (2), 181 (2015), pp. 623–697, https://doi.org/10.4007/annals.2015.181.2.5.

[2] F. ANDREATTA, A. IOVITA, AND V. PILLONI, *p-adic variation of automorphic sheaves*, in Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. II. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 249–276.

[3] T. BERGER AND K. KLOSIN, *Deformations of Saito–Kurokawa type and the paramodular conjecture*, Amer. J. Math., 142 (2020), pp. 1821–1875, https://doi.org/10.1353/ajm.2020.0052. With an appendix by Chris Poor, Jerry Shurman, and David S. Yuen.

[4] S. BIJAKOWSKI, V. PILLONI, AND B. STROH, *Classicité de formes modulaires surconvergentes*, Ann. of Math. (2), 183 (2016), pp. 975–1014, https://doi.org/10.4007/annals.2016.183.3.5.

[5] M. BOLOGNESI, *On Weddle surfaces and their moduli*, Adv. Geom., 7 (2007), pp. 113–144, https://doi.org/10.1515/ADVGEOM.2007.008.

[6] A. R. BOOKER, J. SIJSLING, A. V. SUTHERLAND, J. VOIGHT, AND D. YASAKI, *A database of genus-2 curves over the rational numbers*, LMS J. Comput. Math., 19 (2016), pp. 235–254, https://doi.org/10.1112/S146115701600019X.

[7] G. BOXER, F. CALEGARI, T. GEE, AND V. PILLONI, *Abelian surfaces over totally real fields are potentially modular*, Publ. Math. Inst. Hautes Études Sci., 134 (2021), pp. 153–501, https://doi.org/10.1007/s10240-021-00128-2.

[8] G. BOXER, F. CALEGARI, T. GEE, AND V. PILLONI, *Modularity theorems for abelian surfaces*, 2025, https://arxiv.org/abs/2502.20645.

[9] G. BOXER AND V. PILLONI, *Higher Coleman theory*. arXiv 2110.10251, 2021.

[10] G. BOXER AND V. PILLONI, *Higher Hida theory for Siegel modular forms*, 2023. preprint.

[11] J. BREEDING, II, C. POOR, AND D. S. YUEN, *Computations of spaces of paramodular forms of general level*, J. Korean Math. Soc., 53 (2016), pp. 645–689, https://doi.org/10.4134/JKMS.j150219.

[12] C. BREUIL, B. CONRAD, F. DIAMOND, AND R. TAYLOR, *On the modularity of elliptic curves over* **Q**: *wild 3-adic exercises*, J. Amer. Math. Soc., 14 (2001), pp. 843–939 (electronic), https://doi.org/10.1090/S0894-0347-01-00370-8.

[13] A. BRUMER AND K. KRAMER, *Paramodular abelian varieties of odd conductor*, Trans. Amer. Math. Soc., 366 (2014), pp. 2463–2516, https://doi.org/10.1090/S0002-9947-2013-05909-0.

[14] A. BRUMER, A. PACETTI, C. POOR, G. TORNARÍA, J. VOIGHT, AND D. S. YUEN, *On the paramodularity of typical abelian surfaces*, Algebra Number Theory, 13 (2019), pp. 1145–1195, https://doi.org/10.2140/ant.2019.13.1145.

[15] F. CALEGARI, *Motives and L-functions*, in Current developments in mathematics 2018, Int. Press, Somerville, MA, [2020], pp. 57–123.

[16] F. CALEGARI, *Reciprocity in the Langlands program since Fermat's last theorem*, in ICM—International Congress of Mathematicians. Vol. 2. Plenary lectures, EMS Press, Berlin, [2023], pp. 610–651.

[17] F. CALEGARI AND D. GERAGHTY, *Modularity lifting beyond the Taylor–Wiles method*, Inventiones mathematicae, 211 (2018), pp. 297–433, https://doi.org/10.1007/s00222-017-0749-x.

[18] A. CARAIANI AND J. NEWTON, *On the modularity of elliptic curves over imaginary quadratic fields*, 2023, https://arxiv.org/abs/2301.10509.

[19] R. F. COLEMAN, *Classical and overconvergent modular forms*, Invent. Math., 124 (1996), pp. 215–241, https://doi.org/10.1007/s002220050051.

[20] B. CONRAD, F. DIAMOND, AND R. TAYLOR, *Modularity of certain potentially Barsotti–Tate Galois representations*, J. Amer. Math. Soc., 12 (1999), pp. 521–567.

[21] F. DIAMOND, *The Taylor–Wiles construction and multiplicity one*, Invent. Math., 128 (1997), pp. 379–391, https://doi.org/10.1007/s002220050144.

[22] G. Faltings and C.-L. Chai, *Degeneration of abelian varieties*, vol. 22 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], Springer-Verlag, Berlin, 1990, https://doi.org/10.1007/978-3-662-02632-8. With an appendix by David Mumford.

[23] N. Freitas, B. V. Le Hung, and S. Siksek, *Elliptic curves over real quadratic fields are modular*, Invent. Math., 201 (2015), pp. 159–206, https://doi.org/10.1007/s00222-014-0550-z.

[24] P. L. Kassaei, *A gluing lemma and overconvergent modular forms*, Duke Math. J., 132 (2006), pp. 509–529, https://doi.org/10.1215/S0012-7094-06-13234-9.

[25] C. Khare, *Remarks on mod p forms of weight one*, Internat. Math. Res. Notices, (1997), pp. 127–133, https://doi.org/10.1155/S1073792897000093.

[26] C. Khare and J.-P. Wintenberger, *On Serre's conjecture for 2-dimensional mod p representations of* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Ann. of Math. (2), 169 (2009), pp. 229–253, https://doi.org/10.4007/annals.2009.169.229.

[27] H. H. Kim and F. Shahidi, *Functorial products for* $\mathrm{GL}_2 \times \mathrm{GL}_3$ *and the symmetric cube for* GL_2, Ann. of Math. (2), 155 (2002), pp. 837–893, https://doi.org/10.2307/3062134. With an appendix by Colin J. Bushnell and Guy Henniart.

[28] L. Moret-Bailly, *Groupes de Picard et problèmes de Skolem II*, Ann. Scient. Ec. Norm. Sup., 22 (1989), pp. 181–194.

[29] J. Nekovář, *Eichler–Shimura relations and semisimplicity of étale cohomology of quaternionic Shimura varieties*, Ann. Sci. Éc. Norm. Supér. (4), 51 (2018), pp. 1179–1252, https://doi.org/10.24033/asens.2374.

[30] L. Pan, *On locally analytic vectors of the completed cohomology of modular curves*, Forum Math. Pi, 10 (2022), pp. Paper No. e7, 82, https://doi.org/10.1017/fmp.2022.1.

[31] S. Patrikis and R. Taylor, *Automorphy and irreducibility of some l-adic representations*, Compos. Math., 151 (2015), pp. 207–229, https://doi.org/10.1112/S0010437X14007519.

[32] V. Pilloni, *Higher coherent cohomology and p-adic modular forms of singular weights*, Duke Math. J., 169 (2020), pp. 1647–1807, https://doi.org/10.1215/00127094-2019-0075.

[33] V. Pilloni, *Faisceaux equivariants sur* $\mathbb{P}^1$ *et faisceaux automorphes*, Annales de la Faculté des sciences de Toulouse : Mathématiques, 6e série, 33 (2024), pp. 1155–1213, https://doi.org/10.5802/afst.1794.

[34] V. Pilloni and B. Stroh, *Surconvergence, ramification et modularité*, Astérisque, (2016), pp. 195–266.

[35] C. Poor, J. Shurman, and D. S. Yuen, *Siegel paramodular forms of weight 2 and squarefree level*, Int. J. Number Theory, 13 (2017), pp. 2627–2652, https://doi.org/10.1142/S1793042117501469.

[36] C. Poor and D. S. Yuen, *Paramodular cusp forms*, Math. Comp., 84 (2015), pp. 1401–1438, https://doi.org/10.1090/S0025-5718-2014-02870-6.

[37] B. Roberts and R. Schmidt, *Local newforms for GSp(4)*, vol. 1918 of Lecture Notes in Mathematics, Springer, Berlin, 2007, https://doi.org/10.1007/978-3-540-73324-9.

[38] J. E. Rodríguez Camargo, *Locally analytic completed cohomology*, 2022, https://arxiv.org/abs/2209.01057.

[39] J. E. Rodríguez Camargo, *The analytic de Rham stack in rigid geometry*, 2024, https://arxiv.org/abs/2401.07738.

[40] S. Sasaki, *Integral models of Hilbert modular varieties in the ramified case, deformations of modular Galois representations, and weight one forms*, Invent. Math., 215 (2019), pp. 171–264, https://doi.org/10.1007/s00222-018-0825-x.

[41] P. Scholze, *On torsion in the cohomology of locally symmetric varieties*, Ann. of Math. (2), 182 (2015), pp. 945–1066, https://doi.org/10.4007/annals.2015.182.3.3.

[42] P. Scholze, *Geometrization of the real local Langlands correspondence (draft version, used for ARGOS seminar)*, 2024, https://people.mpim-bonn.mpg.de/scholze/RealLocalLanglands.pdf.

[43] J.-P. Serre, *Facteurs locaux des fonctions zêta des variétés algébriques (définitions et conjectures)*, Séminaire Delange-Pisot-Poitou. Théorie des nombres, 11 (1969-1970), pp. 1–15, http://eudml.org/doc/110758.

[44] J.-P. Serre, *Topics in Galois theory*, vol. 1 of Research Notes in Mathematics, A K Peters, Ltd., Wellesley, MA, second ed., 2008. With notes by Henri Darmon.

[45] N. I. Shepherd-Barron and R. Taylor, mod 2 *and* mod 5 *icosahedral representations*, J. Amer. Math. Soc., 10 (1997), pp. 283–298, https://doi.org/10.1090/S0894-0347-97-00226-9.

[46] R. Taylor, *Remarks on a conjecture of Fontaine and Mazur*, J. Inst. Math. Jussieu, 1 (2002), pp. 125–143, https://doi.org/10.1017/S1474748002000038.

[47] R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (2), 141 (1995), pp. 553–572, https://doi.org/10.2307/2118560.

[48] The LMFDB Collaboration, *The L-functions and modular forms database.* https://www.lmfdb.org, 2024. [Online; accessed 1 June 2024].

[49] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2), 141 (1995), pp. 443–551, https://doi.org/10.2307/2118559.