LATTICES IN THE COHOMOLOGY OF SHIMURA CURVES

MATTHEW EMERTON, TOBY GEE, AND DAVID SAVITT

Abstract. We prove the main conjectures of [Bre12] (including a generalisation from the principal series to the cuspidal case) and [Dem], subject to a mild global hypothesis that we make in order to apply certain $R = \mathbb{T}$ theorems. More precisely, we prove a multiplicity one result for the mod p cohomology of a Shimura curve at Iwahori level, and we show that certain apparently globally defined lattices in the cohomology of Shimura curves are determined by the corresponding local p-adic Galois representations. We also indicate a new proof of the Buzzard–Diamond–Jarvis conjecture in generic cases. Our main tools are the geometric Breuil–Mézard philosophy developed in [EG14], and a new and more functorial perspective on the Taylor–Wiles–Kisin patching method. Along the way, we determine the tamely potentially Barsotti–Tate deformation rings of generic two-dimensional mod p representations, generalising a result of [BM12] in the principal series case.

Contents

The first author was supported in part by NSF grants DMS-1003339 and DMS-1249548, the second author was supported in part by a Marie Curie Career Integration Grant and an ERC starting grant, and the third author was supported in part by NSF grant DMS-0901049 and NSF CAREER grant DMS-1054032. The second and third authors would like to thank the mathematics department of Northwestern University for its hospitality.

1. Introduction

1.1. **Breuil's Conjectures.** The p-adic local Langlands correspondence for the group $\mathrm{GL}_2(\mathbb{Q}_p)$ (*cf.* [Col10, Paš13]) is one of the most important developments in the Langlands program in recent years, and the natural and definitive nature of the results is extremely suggestive of there being a similar correspondence for $\mathrm{GL}_n(K)$, where n is an arbitrary natural number and K an arbitrary finite extension of $\mathbb{Q}_p$. Unfortunately, all available evidence to date suggests that any more general such correspondence will be much more complicated to describe, even in the case of $\mathrm{GL}_2(K)$ for an unramified extension K of $\mathbb{Q}_p$. For example, one of the first steps in the construction of the correspondence for $\mathrm{GL}_2(\mathbb{Q}_p)$ was the explicit classification of irreducible admissible smooth representations of $\mathrm{GL}_2(\mathbb{Q}_p)$ in characteristic p [BL94, Bre03], from which it is straightforward to deduce a semisimple mod p local Langlands correspondence for $\mathrm{GL}_2(\mathbb{Q}_p)$ [Bre03]. To date, no analogous classification is known for any other $\mathrm{GL}_2(K)$, and it is known that the irreducible representations live in infinite families, described by parameters that have yet to be determined [BP12]. In particular, there are in a precise sense too many representations of $\mathrm{GL}_2(K)$ for there to be any simple one-to-one correspondence with Galois representations; rather, it seems that only certain particular representations of $\mathrm{GL}_2(K)$ will participate in the correspondence.

By the local-global compatibility result of [Eme], the completed cohomology of modular curves can be described in terms of the p-adic local Langlands correspondence for $\mathrm{GL}_2(\mathbb{Q}_p)$. Similarly, it is expected that the completed cohomology of Shimura curves can be described in terms of the anticipated p-adic local Langlands correspondence for $\mathrm{GL}_2(K)$. Accordingly, one way to get information about this hypothetical correspondence is to study the action of $\mathrm{GL}_2(K)$ on the cohomology of Shimura curves, for example with characteristic p coefficients. As one illustration of this approach, we recall that the $\mathrm{GL}_2(\mathcal{O}_K)$-socles of the irreducible $\mathrm{GL}_2(K)$-subrepresentations of this mod p cohomology are described by the Buzzard–Diamond–Jarvis conjecture ([BDJ10], proved in [BLGG13a], [GLS12], and [GK12]), and this information is an important guiding tool in the work of [BP12]. In particular, given $K/\mathbb{Q}_p$ unramified, and a generic Galois representation $\overline{\rho} : G_K \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$, [BP12] constructs an infinite family of irreducible $\overline{\mathbb{F}}_p$-representations of $\mathrm{GL}_2(K)$ whose $\mathrm{GL}_2(\mathcal{O}_K)$-socles are as predicted by [BDJ10], and conjectures that if $\bar{r}$ is a globalisation of $\overline{\rho}$ to a Galois representation over some totally real field, then the $\bar{r}$-part of the cohomology of an appropriate Shimura curve contains one of these representations.

Although the representations of [BP12] are constructed so as to have the correct socle (in the sense of [BDJ10]), proving that the cohomology contains such a representation lies deeper than the Buzzard–Diamond–Jarvis conjecture, and amounts to proving that certain subspaces of the mod p cohomology at pro-p Iwahori level are stable under the action of a Hecke operator. In [Bre12] Breuil develops two approaches to this problem, in the process making two conjectures of independent interest. The main results of this paper are proofs of these conjectures, as well as some generalisations. As a consequence, we demonstrate that the cohomology contains one of the representations constructed in [BP12].

The subspaces of the cohomology under consideration are those obtained by fixing the character through which the Iwahori subgroup acts on the cohomology at pro-p Iwahori level, and one approach to the Hecke-stability problem would be to show that these spaces are one-dimensional; that is, to prove a multiplicity one result at Iwahori level. In [Dem] Dembélé gives computational evidence that such a multiplicity one result should hold. One of our main results is the following (Theorem 10.2.1).

A. **Theorem.** *Suppose that $p \geq 5$ and the usual Taylor–Wiles condition holds. Then the multiplicity one conjecture of* [Dem] *holds.*

The other approach taken in [Bre12] is via a p-adic local-global compatibility conjecture for tamely potentially Barsotti–Tate representations. To be precise, let F be a totally real field in which p is inert (in fact, we only need p to be unramified, but we assume here that p is inert for simplicity of notation), and fix a finite extension $E/\mathbb{Q}_p$ (our field of coefficients) with ring of integers $\mathcal{O}$ and uniformiser ϖ_E. Write $f := [F : \mathbb{Q}]$, and let k be the residue field of F_p. Let $\overline{\rho} : G_F \to \mathrm{GL}_2(\mathbb{F})$ be an absolutely irreducible modular representation with $\overline{\rho}|_{G_{F_p}}$ generic (see Definition 2.1.1), and consider

$$M_{\overline{\rho}} := \varinjlim_{K_p} M(K_pK^p, \mathcal{O})_{\overline{\rho}}$$

where $M(K_pK^p, \mathcal{O})_{\overline{\rho}}$ is the $\overline{\rho}$-part of the cohomology of an appropriate Shimura curve at level K_pK^p, where K_p is a compact open subgroup of $\mathrm{GL}_2(F_p)$. This has a natural action of $\mathrm{GL}_2(F_p)$, and we consider some subrepresentation π of $M_{\overline{\rho}}[1/p]$ which is a tame principal series representation. Then $M_{\overline{\rho}}$ determines a lattice in π, and in particular, a lattice in the tame principal series type inside π (that is, the smallest $\mathrm{GL}_2(\mathcal{O}_{F_p})$-representation inside π). There are typically many homothety classes of $\mathrm{GL}_2(\mathcal{O}_{F_p})$-stable lattices inside this type, and there is no *a priori* reason that we are aware of to suppose that the lattice is a local invariant of the situation. Indeed, it is easy to check that classical local-global compatibility applied to π does not in general give enough information to determine the lattice.

However, the main conjecture of [Bre12] is that the lattice is determined by the local p-adic Galois representation in $M_{\overline{\rho}}$ corresponding to π in an explicit way. The local Galois representation is potentially Barsotti–Tate of tame principal series type, and various parameters can be read off from its Dieudonné module. Breuil classified the homothety classes of lattice in the tame principal series types in terms of similar parameters, and conjectured that the parameters for the lattice in π agree with those for the corresponding Galois representation. Our other main result confirms this conjecture (see Theorem 8.2.1).

B. **Theorem.** *Breuil's conjecture holds if $p \geq 5$ and the usual Taylor–Wiles condition holds; an analogous result holds for tame cuspidal types.*

We note that in the cuspidal case, we use an extension to (most) tame cuspidal types of Breuil's classification of lattices, which in turn uses a generalisation of some of the results of [BP12] to irregular weights. These results are presented in Section 5, and may be of independent interest.

Our main tools in proving Theorems A and B are the geometric Breuil–Mézard Conjecture, and its relation to the Taylor–Wiles–Kisin patching method. We have hopes that our arguments can be made to apply more generally, e.g. to show for arbitrary types that the lattices induced on them by embeddings into cohomology arising from global p-adic Galois representations are in fact completely determined by the corresponding local p-adic Galois representations (though perhaps without giving an explicit formula for the lattices). In the remainder of the introduction, we outline our methods in more detail.

1.2. **Lattices and their gauges.** Most of our efforts are directed towards proving Theorem B, while Theorem A will come as something of a byproduct of our method of proof via Taylor–Wiles–Kisin patching, much as Diamond in [Dia97] was able to deduce mod p multiplicity one results as a corollary of a patching argument.

Let σ be a tame type. The basic problem that we face is to determine the lattice σ° induced by integral cohomology when we are given an embedding $\sigma \hookrightarrow M_{\overline{\rho}}[1/p]$ arising from the local factor at p of the automorphic representation generated by some Hecke eigenform whose associated Galois representation lifts $\overline{\rho}$. It turns out that such a lattice can be encoded in a convenient way by what we call its *gauge*. Briefly, let $\overline{\sigma}$ be the reduction mod ϖ_E of σ (which is well-defined up to semisimplification). Then each Jordan–Hölder constituent of $\overline{\sigma}$ occurs with multiplicity one, and these consituents may be labelled by a certain collection $\mathcal{P}$ of subsets of $\{0, \dots, f-1\}$, which typically is equal to the full power set of $\{0, \dots, f-1\}$. For each $J \in \mathcal{P}$, we may find a lattice σ_J° in σ, uniquely determined up to homothety, whose cosocle equals the constituent of $\overline{\sigma}$ labelled by the subset J. For each $J \in \mathcal{P}$, we obtain a free rank one $\mathcal{O}$-module $\mathrm{Hom}_{\mathrm{GL}_2(k_v)}(\sigma_J^\circ, \sigma^\circ) \subset \mathrm{Hom}_{\mathrm{GL}_2(k_v)}(\sigma, \sigma) = E$, the collection of which we refer to as the gauge of σ (see Definition 4.1.3). The gauge of σ° determines σ° up to homothety (Proposition 4.1.4), and so our problem becomes that of determining the gauge of σ°.

If λ is the system of Hecke eigenvalues associated to our eigenform, and $M_{\overline{\rho}}[\lambda]$ denotes the λ-eigenspace in $M_{\overline{\rho}}$, then this amounts to understanding the collection of Hom-modules $\mathrm{Hom}_{\mathrm{GL}_2(\mathcal{O}_F)}\big(\sigma_J^\circ, M_{\overline{\rho}}[\lambda]\big)$. Rather than studying these for a single eigenform, we consider this collection simultaneously for a family of eigenforms whose associated Galois representations lift $\overline{\rho}$. More precisely, we employ the patching method of Taylor–Wiles [TW95], as further developed by Diamond [Dia97] and Kisin [Kis09b].

1.3. **Patching functors.** Our actual implementation of the Taylor–Wiles method is a further refinement of the viewpoint adopted in the papers [Kis09a, EG14], which patch not only spaces of modular forms, but certain filtrations on them. We take this idea to its natural conclusion, and patch together spaces of modular forms with all possible coefficient systems simultaneously (see Section 6). The patched spaces of modular forms obtained from the Taylor–Wiles method are naturally modules over certain universal Galois deformation rings, and we adopt a more geometric point of view, regarding them as coherent sheaves on deformation spaces of Galois representations.

In this way, we obtain what we call a *patching functor*, which is a functor from continuous representations of $\mathrm{GL}_2(\mathcal{O}_{F_p})$ on finite $\mathcal{O}$-modules to coherent sheaves on deformation spaces. These functors satisfy a number of natural properties (for example, certain of the sheaves obtained are maximal Cohen–Macaulay over their support, and this support is specified by a compatibility with the local Langlands correspondence and the Buzzard–Diamond–Jarvis conjecture), and we are able to define an abstract notion of a patching functor, with no reference to the Taylor–Wiles method. (One motivation for this is in order to clarify exactly which properties of these functors are used in our main theorems, and another is that it is sometimes possible to construct patching functors without using the Taylor–Wiles method, *cf.* [CEG$^+$a].)

In the particular case at hand, this construction realizes (the duals to) the above Hom-modules as the fibres of certain coherent sheaves (the patched modules) over the deformation space which parameterizes deformations of $\overline{\rho}|_{G_{F_p}}$ which are tamely potentially Barsotti–Tate of type σ. We can then reinterpret the problem of understanding the above Hom-modules as that of understanding these coherent sheaves, and in particular of understanding the natural morphisms between them induced by various inclusions between the σ_J°. We do this by exploiting an explicit description of the deformation space, and the geometric Breuil–Mézard philosophy of [EG14].

In Section 7 we give our explicit description of the deformation space, showing that it is the formal spectrum of a power series ring over the complete local ring

$\mathcal{O}[\![X_1, Y_1, \ldots, X_r, Y_r]\!]/(X_iY_i - p)_{i=1,\ldots,r}$, where r depends on $\overline{\rho}$ via the Buzzard–Diamond–Jarvis conjecture. (In the principal series case this is proved in [BM12], and we extend this computation to the cuspidal case by making use of a base change result coming from Taylor–Wiles patching via [EG14].) The X_i and Y_i have an interpretation in terms of the Dieudonné modules attached to the potentially Barsotti–Tate lifts of $\overline{\rho}|_{G_{F_p}}$ of type σ, and using this interpretation we can reduce Breuil's conjecture to understanding how the X_i and Y_i act on our coherent sheaves.

We do this in Section 8; the key is to show that the cokernels of various of the morphisms induced by the inclusions between the σ_J° are annihilated by certain products of the X_i and Y_i. Using the Cohen–Macaulay properties of the patching functors, this reduces to showing that the supports of certain cokernels are annihilated by these products, which in turn reduces to gaining a concrete understanding of the components of the special fibre of the deformation space in terms of the Serre weights specified by the Buzzard–Diamond–Jarvis conjecture. It follows from the results of [EG14] that these components are identified with the special fibres of certain Fontaine–Laffaille deformation spaces, and we need to make this correspondence completely explicit. We do this via an explicit p-adic Hodge theoretic computation, following [BM12] (and again reducing to a special case via a base change argument).

1.4. **Freeness.** In order to prove Conjecture A, we prove that certain of our patched coherent sheaves are free. By an argument originally due to Diamond [Dia97], this follows from the maximal Cohen–Macaulay property if our deformation space is regular. However, from the description above we see that this can only happen if $r = 1$. In order to prove freeness in the case that $r > 1$, we exploit the explicit description of the patching functor that we have obtained in the above work, and effectively proceed by induction on r. By Nakayama's lemma, it suffices to prove freeness in characteristic p. The functorial nature of our patching construction means that we can compare patched sheaves corresponding to the reductions modulo ϖ_E of varying tame types σ, and glue together the patched sheaves in general from those types which give $r = 1$, where we know freeness.

1.5. **The generic Buzzard–Diamond–Jarvis Conjecture.** As a byproduct of our constructions, we are able to give a new proof of the Buzzard–Diamond–Jarvis conjecture in the generic case (and indeed, we expect that this proof would allow a slight relaxing of the global hypotheses in the generic case when $p = 5$; see Remark B.1.1). This proof is related to the approach taken in [Gee11], but avoids both the combinatorial arguments of that paper and the dependence on the modularity lifting theorems proved in [Kis09b, Gee06]. The method is closely related to the arguments used in [Tay08] to avoid the use of Ihara's lemma in proving modularity lifting theorems. The idea is that our deformation spaces have irreducible generic fibres, but their special fibres have many irreducible components; from this and the properties of our patching functors, it follows that a patched module which is supported on the generic fibre of the deformation space is necessarily supported on each irreducible component of the special fibre.

As mentioned above, these components correspond to different Serre weights, and if we choose our tame type σ to maximise the number of components in the special fibre, we are able to prove modularity in every weight predicted by the Buzzard–Diamond–Jarvis conjecture. (In contrast, the arguments of [Gee11] choose tame types to ensure that the special fibres are irreducible, and use a base change argument and the results of [Kis09b, Gee06] to prove modularity in each tame type.) We explain this argument in Section 9, and in Appendix B.1 we explain how to remove the dependence on [Kis09b, Gee06] which is implicit in our use of the results of [EG14, GK12] in proving our explicit descriptions of the tamely Barsotti–Tate deformation spaces.

1.6. **A Morita-theoretic interpretation.** A Morita-theoretic interpretation of our arguments was suggested to us by David Helm. Namely, if we write

$$P := \oplus_{J \in \mathcal{P}} \sigma_J^{\circ},$$

then P is a projective generator in a certain abelian category of $\mathcal{O}[\mathrm{GL}_2(k_v)]$-modules, namely the full abelian subcategory of the category of all such modules generated by the collection of all lattices in σ. Our theory of gauges is thus a special case of general Morita theory, which shows that if $\mathcal{E} := \mathrm{End}_{\mathrm{GL}_2(k_v)}(P)$, then σ may be recovered from the $\mathcal{E}$-module $\mathrm{Hom}_{\mathrm{GL}_2(k_v)}(P, \sigma)$ via the formula

$$\sigma = \mathrm{Hom}_{\mathrm{GL}_2(k_v)}(P, \sigma) \otimes_{\mathcal{E}} P.$$

Furthermore, Theorem 5.2.4 may be regarded as an explicit description of the $\mathcal{O}$-algebra $\mathcal{E}$.

Since $\mathrm{Hom}_{\mathrm{GL}_2(k_v)}(P, \sigma)$ may be recovered as the fibre of the patched module for P at an appropriately chosen point, in order to describe this $\mathcal{E}$-module, it suffices to describe the $\mathcal{E}$-module structure on the corresponding patched module, and one interpretation of our results is that they give an explicit description of this patched module with its $\mathcal{E}$-action (and show, in particular, that it is purely local).

1.7. **Organisation of the paper.** For the convenience of the reader, we now briefly outline the structure of the paper. In Section 2 we recall the definition of a generic Galois representation, explain the Buzzard–Diamond–Jarvis conjecture for these representations, and give some definitions relating to our later base change arguments. In Section 3 we give a similar treatment of tame types. Section 4 contains some general results on the structure of lattices in irreducible representations whose reductions mod p have all Jordan–Hölder factors occurring with multiplicity one. In Section 5 we specialise to give an explicit description of the lattices in tame types, extending the results of [Bre12] for tame principal series types.

In Section 6 we define the notion of a patching functor, and use the Taylor–Wiles–Kisin method to construct examples from the cohomology of Shimura curves, as well as from spaces of algebraic modular forms on definite quaternion algebras and definite unitary groups. Our p-adic Hodge theoretic arguments appear in Section 7, which gives the explicit description of the tame Barsotti–Tate deformation spaces (and their special fibres) explained above. In Section 8 we prove Theorem B, and in Section 9 we explain our new proof of the Buzzard–Diamond–Jarvis conjecture in the generic case. Finally, Theorem A is proved in Section 10.

There are two appendices; in Appendix A we explain the connection between unipotent Fontaine–Laffaille theory and the theory of unipotent φ-modules, and in Appendix B we discuss the use of the geometric Breuil–Mézard conjecture in our arguments in greater detail. In particular, we explain how to prove our main theorems without making use of the results of [Kis09b, Gee06], and we give an argument that hints at a deeper relationship between the structures of the lattices in tame types and of the tamely potentially Barsotti–Tate deformation spaces.

1.8. **Notation and Conventions.** If K is any field let $G_K = \mathrm{Gal}(\overline{K}/K)$, where $\overline{K}$ is a fixed separable closure of K. Suppose that p is prime. When K is a finite extension of $\mathbb{Q}_p$, let $I_K \subset G_K$ denote the inertia subgroup.

Let F_v be a finite unramified extension of $\mathbb{Q}_p$. (Usually, F_v will be the completion of a number field F at a finite place v, but it will be convenient for us to use this notation more generally.) We write k_v to denote the residue field of v, and let $f = [k_v : \mathbb{F}_p]$. We write $\mathcal{O}_{F_v}$ for $W(k_v)$, so that F_v is the fraction field of $\mathcal{O}_{F_v}$. Write $f = [F_v : \mathbb{Q}_p]$, $q_v = p^f$, and $e = p^f - 1$.

Let L_v denote the quadratic unramified extension of F_v, with residue field l_v, the quadratic extension of k_v.

Let E be a finite extension of $\mathbb{Q}_p$ (which will be our field of coefficients) with ring of integers $\mathcal{O}$, uniformiser ϖ_E, and residue field $\mathbb{F}$. We will assume throughout the paper that E is sufficiently large in the sense that there exists an embedding of l_v into $\mathbb{F}$. We will sometimes make specific choices for E; in particular, at certain points in our arguments it will be important to take $E = W(\mathbb{F})[1/p]$.

Let $\varepsilon : G_{F_v} \to \mathcal{O}^\times$ denote the cyclotomic character.

If V is a de Rham representation of G_{F_v} over E and κ is an embedding $F_v \hookrightarrow E$ then the multiset $\mathrm{HT}_\kappa(V)$ of *κ-labeled Hodge–Tate weights* of V is defined to contain the integer i with multiplicity

$$\dim_E(V \otimes_{\kappa,K} \widehat{\overline{F}}_v(i))^{G_{F_v}},$$

with the usual notation for Tate twists. Thus for example $\mathrm{HT}_\kappa(\varepsilon) = \{-1\}$.

If L is a local or global field, we write Art_L for the Artin reciprocity map, normalised so that the global map is compatible with the local ones, and the local map takes a uniformiser to a geometric Frobenius element. If L is a number field and w is a finite place of L, we will write Frob_w for a geometric Frobenius element at w.

If V is a representation of some group which has a composition series, we write $\mathrm{JH}(V)$ for the Jordan–Hölder factors of V.

When discussing Galois deformations, we will use the terms "framed deformation" (which originated in [Kis09b]) and "lifting" (which originated in [CHT08]) interchangeably.

If σ is a finite-dimensional E-representation of a group Γ, then σ° (possibly with additional decorations) will denote a Γ-stable $\mathcal{O}$-lattice in σ (sometimes a general such lattice, and sometimes a particular lattice that has been defined in the context under consideration). We will then let $\overline{\sigma}^\circ$ (with the same decorations as σ°, if appropriate) denote the reduction of σ° modulo ϖ_E. We will write $\overline{\sigma}$ for some choice of $\overline{\sigma}^\circ$; but we will only use the notation $\overline{\sigma}$ in situations where the choice of a particular $\overline{\sigma}^\circ$ is unimportant, for example when discussing the Jordan–Hölder factors of $\overline{\sigma}$. We note that $\overline{\sigma}^\circ$ is finite-dimensional over $\mathbb{F}$, and so admits both a socle filtration and a cosocle filtration as a Γ-representation, each of finite length. The Γ-representation σ° itself is of infinite length, and does not admit a socle filtration. However, it does admit a cosocle filtration; furthermore, the cosocles of σ° and of $\overline{\sigma}^\circ$ coincide.

If A is a finitely generated $\mathcal{O}$-module, we write A^* for the Pontrjagin dual of A. If A is torsion, then we will sometimes write $A^\vee$ for A^*. If A is a finite free $\mathcal{O}$-module, we will write $A^\vee$ for its $\mathcal{O}$-dual. Note that with these conventions and in the notation of the previous paragraph, we have $(\overline{\sigma}^\circ)^\vee \cong \overline{(\sigma^\circ)^\vee}$.

Throughout the paper, we will number our socle and cosocle filtrations starting at 0, so that for example the 0-th layer of the cosocle filtration is the cosocle.

We let $\mathcal{S}$ denote the set $\{0, \ldots, f-1\}$, which we sometimes identify with the quotient $\mathbb{Z}/f\mathbb{Z}$ in the evident way. (Typically this identification will be invoked by the expression that certain algebraic expressions involving the elements of $\mathcal{S}$ are to be understood *cyclically*, which is to say, modulo f.) We draw the reader's attention to two involutions related to $\mathcal{S}$ that will play a role in the paper. Namely, in Subsection 5.2, and again in Section 8, we will fix a particular subset J_{base} of $\mathcal{S}$, and will let $\imath$ denote the involution on the power set of $\mathcal{S}$ defined by $J \mapsto J \triangle J_{\text{base}}$, while in Section 7 we will let ι denote the involution on $\mathcal{S}$ itself defined by $j \mapsto f-1-j$.

If F is a totally real field and π is a regular cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_F)$, then we fix an isomorphism $\imath : \overline{\mathbb{Q}}_p \to \mathbb{C}$ and associate to π the p-adic Galois representation $r_{p,\imath}(\pi) : G_F \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$, normalised as in [BLGGT14, Thm. 2.1.1]. (Note that we will use the symbol $\imath$ in other contexts in the body of the paper, but we will not have to mention the isomorphism $\overline{\mathbb{Q}}_p \to \mathbb{C}$ again, so we have denoted it by $\imath$ here for ease of reference to [BLGGT14].)

1.9. **The inertial local Langlands correspondence.** Let L be a local field of characteristic zero, so that L is a finite extension of $\mathbb{Q}_l$ for some prime l. An *inertial type* for L is a two-dimensional E-representation τ of I_L with open kernel, which may be extended to G_L. Henniart's appendix to [BM02] associates a finite dimensional irreducible E-representation $\sigma(\tau)$ of $\mathrm{GL}_2(\mathcal{O}_L)$ to τ; we refer to this association as the *inertial local Langlands correspondence*. To be precise, if $l = p$ then we normalise the correspondence as in Theorem 2.1.3 of [GK12], and if $l \neq p$ we use the version recalled in Section 5.2 of [GK12]. (The only difference between the two cases is that if $l \neq p$ we consider representations with non-trivial monodromy; the definition only differs in the case of scalar inertial types, which in the case $l = p$ correspond to twists of the trivial representation, and in the case $l \neq p$ to twists of the small Steinberg representation.)

If $l \neq p$, then let D denote the unique nonsplit quaternion algebra with centre L. We say that an inertial type τ is *discrete series* if it is either irreducible or scalar. In this case, a variant of Henniart's construction which is explained in Section 5.2.2 of [GK12] associates a finite dimensional irreducible E-representation $\sigma_D(\tau)$ of $\mathcal{O}_D^\times$ to τ.

2. Galois representations and Serre weights

In this section we recall various definitions and facts about the weight part of Serre's conjecture, largely following Section 4 of [Bre12]. We fix once and for all an embedding $\overline{\kappa}_0 : k_v \hookrightarrow \mathbb{F}$, and recursively define embeddings $\overline{\kappa}_i : k_v \hookrightarrow \mathbb{F}$ for $i \in \mathbb{Z}$ by $\overline{\kappa}_{i+1}^p = \overline{\kappa}_i$. Let ω_f denote the fundamental character of I_{F_v} of niveau f associated to $\overline{\kappa}_0$; that is, ω_f is the *reciprocal* of the character defined by the composite

$$I_{F_v} \xrightarrow{\mathrm{Art}_{F_v}^{-1}} \mathcal{O}_{F_v}^\times \longrightarrow k_v^\times \xrightarrow{\overline{\kappa}_0} \mathbb{F}^\times,$$

so that for example if $f = 1$ then $\omega_1 = \overline{\varepsilon}^{-1}$. Let ω_{2f} be a fundamental character of niveau $2f$ similarly associated to an embedding $\overline{\kappa}'_0 : l_v^\times \hookrightarrow \mathbb{F}$ extending $\overline{\kappa}_0$. Define $\overline{\kappa}'_i$ for $i \in \mathbb{Z}$ by analogy with the $\overline{\kappa}_i$.

2.1. **Basic notions.** We recall some basic notions related to mod p local Galois representations and their associated Serre weights, as well as the main result of [GLS12].

2.1.1. **Definition.** We say that a continuous representation $\overline{\rho}: G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$ is *generic* if and only if $\overline{\rho}|_{I_{F_v}}$ is isomorphic up to twist to a representation of one of the following two forms.

(1) $\begin{pmatrix} \omega_f^{\sum_{j=0}^{f-1}(r_j+1)p^j} & * \\ 0 & 1 \end{pmatrix}$ with $0 \le r_j \le p-3$ for each j, and the r_j not all equal to 0, and not all equal to $p-3$.

(2) $\begin{pmatrix} \omega_{2f}^{\sum_{j=0}^{f-1}(r_j+1)p^j} & 0 \\ 0 & \omega_{2f}^{q\sum_{j=0}^{f-1}(r_j+1)p^j} \end{pmatrix}$ with $1 \le r_0 \le p-2$, and $0 \le r_j \le p-3$ for $j > 0$.

It is easily checked that this definition is independent of the choice of $\overline{\kappa}_0$ and $\overline{\kappa}'_0$. The following lemma is also easily checked.

2.1.2. **Lemma.** *If $p > 3$ and $\overline{\rho}$ is generic, then $\overline{\rho}|_{G_{L_v}}$ is also generic.*

2.1.3. *Remark.* If $p = 3$ then there are no reducible generic representations, and in the above notation the only irreducible generic representations have $r_0 = 1$ and $r_j = 0$ for $j > 0$. The failure of Lemma 2.1.2 to hold in this case means that we will not consider the case $p = 3$ in our main results.

By definition, a *Serre weight* is an irreducible $\mathbb{F}$-representation of $\mathrm{GL}_2(k_v)$. Concretely, such a representation is of the form

$$\overline{\sigma}_{\vec{t},\vec{s}} := \otimes_{j=0}^{f-1}(\det{}^{t_j} \mathrm{Sym}^{s_j} k_v^2) \otimes_{k_v,\overline{\kappa}_{-j}} \mathbb{F},$$

where $0 \le s_j, t_j \le p-1$ and not all t_j are equal to $p-1$.

2.1.4. **Definition.** We say that a Serre weight $\overline{\sigma}_{\vec{t},\vec{s}}$ is *regular* if no s_j is equal to $p-1$.

2.1.5. *Remark.* This is the same definition of regular Serre weight as that in [GK12]; it is less restrictive than the definition used in [Gee11].

We let $\mathcal{D}(\overline{\rho})$ be the set of Serre weights associated to $\overline{\rho}$ by [BDJ10]; this definition is recalled in Section 4 of [Bre12]. While we will not need the full details of the definition of $\mathcal{D}(\overline{\rho})$, it will be convenient to recall certain of its properties, and in particular its definition in the case that $\overline{\rho}$ is semisimple; see [BDJ10] or Section 4 of [Bre12] for further details. We will also make use of the characterisation of $\mathcal{D}(\overline{\rho})$ of Theorem 2.1.8 below (at least implicitly, via its use in [GK12]).

We always have $\mathcal{D}(\overline{\rho}) \subset \mathcal{D}(\overline{\rho}^{\mathrm{ss}})$. If $\overline{\rho}$ is reducible and semisimple, then $\overline{\sigma}_{\vec{t},\vec{s}} \in \mathcal{D}(\overline{\rho})$ if and only if there is a subset $K \subseteq \mathcal{S} := \{0, \dots, f-1\}$ such that we can write

$$\overline{\rho}|_{I_{F_v}} \cong \omega_f^{\sum_{j=0}^{f-1} t_j p^j} \otimes \begin{pmatrix} \omega_f^{\sum_{j\in K}(s_j+1)p^j} & 0 \\ 0 & \omega_f^{\sum_{j\notin K}(s_j+1)p^j} \end{pmatrix}.$$

If $\overline{\rho}$ is irreducible, we set $\mathcal{S}' = \{0, \dots, 2f-1\}$, and we say that a subset $K \subseteq \mathcal{S}'$ is *symmetric* (respectively *antisymmetric*) if $i \in K$ if and only if $i + f \in K$

(respectively $i+f \notin K$) for all $i \in \mathcal{S}'$, numbered cyclically. Then $\overline{\sigma}_{\vec{t},\vec{s}} \in \mathcal{D}(\overline{\rho})$ if and only if there is an antisymmetric subset $K \subset \mathcal{S}'$ such that we can write

$$\overline{\rho}|_{I_{F_v}} \cong \omega_f^{\sum_{j=0}^{f-1} t_j p^j} \otimes \begin{pmatrix} \omega_{2f}^{\sum_{j\in K}(s_j+1)p^j} & 0 \\ 0 & \omega_{2f}^{\sum_{j\notin K}(s_j+1)p^j} \end{pmatrix}.$$

The following lemma is immediate from the above description of $\mathcal{D}(\overline{\rho})$.

2.1.6. **Lemma.** *If $\overline{\rho}$ is generic and $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$, then $\overline{\sigma}$ is regular; in addition, $\overline{\sigma}$ is not of the form $\overline{\sigma}_{\vec{t},\vec{0}}$.*

2.1.7. **Definition.** Let $\kappa_i : F_v \hookrightarrow E$ be the unique embedding that lifts $\overline{\kappa}_i$. If V is a two-dimensional de Rham representation of G_{F_v} over E and $\mathrm{HT}_{\kappa_{-i}}(V) = \{a_i, b_i\}$ with $a_i \le b_i$ for $0 \le i \le f-1$, we say that V has *Hodge type* $(\vec{a},\vec{b})$. If V has Hodge type $(\vec{0},\vec{1})$, we will also say that V has *Hodge type* 0.

If $\overline{\sigma} = \overline{\sigma}_{\vec{t},\vec{s}}$, we also say that V has *Hodge type* $\overline{\sigma}$ if V has Hodge type $(\vec{t},\vec{t}+\vec{s}+\vec{1})$.

The following theorem is the main result of [GLS12].

2.1.8. **Theorem.** *We have $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$ if and only if $\overline{\rho}$ has a crystalline lift of Hodge type $\overline{\sigma}$.*

2.2. **Base change for Galois representations and Serre weights.** In the sequel we will be studying base change from F_v to L_v, and it will be convenient to have standardized notation for this process.

In the case of Galois represenations, base change is simply restriction, and thus, if $\overline{\rho} : G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$, we write $\mathrm{BC}\,(\overline{\rho})$ for $\overline{\rho}|_{G_{L_v}}$.

For Serre weights the definition of base change is slightly more involved. Suppose given a Serre weight for $\mathrm{GL}_2(k_v)$, say $\overline{\sigma}_{\vec{t},\vec{s}} := \otimes_{j=0}^{f-1}(\det{}^{t_j}\mathrm{Sym}^{s_j} k_v^2) \otimes_{k_v,\overline{\kappa}_{-j}} \mathbb{F}$, and a Serre weight for $\mathrm{GL}_2(l_v)$, say $\overline{\Sigma}_{\vec{t'},\vec{s'}} := \otimes_{j=0}^{2f-1}(\det{}^{t'_j}\mathrm{Sym}^{s'_j} l_v^2) \otimes_{l_v,\overline{\kappa}'_{-j}} \mathbb{F}$. We say that $\overline{\Sigma} = \overline{\Sigma}_{\vec{t'},\vec{s'}}$ is the *base change* of $\overline{\sigma} = \overline{\sigma}_{\vec{t},\vec{s}}$, and write $\overline{\Sigma} = \mathrm{BC}\,(\overline{\sigma})$, if we have $t'_j = t_{\overline{j}}, s'_j = s_{\overline{j}}$ for all $0 \le j \le 2f-1$ (where $\overline{j}$ is $j \pmod f$). Note that not every Serre weight for $\mathrm{GL}_2(l_v)$ is the base change of a Serre weight for $\mathrm{GL}_2(k_v)$, and that every Serre weight for $\mathrm{GL}_2(k_v)$ has a unique base change to $\mathrm{GL}_2(l_v)$.

If $\overline{\rho}$ is generic, and $\overline{\sigma}$ is a Serre weight, then $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$ if and only if $\mathrm{BC}\,(\overline{\sigma}) \in \mathcal{D}(\overline{\rho}|_{G_{L_v}})$. The "only if" direction (which we will use in Section 3) is immediate from the definitions. It is presumably possible to prove the "if" direction from the definitions as well; we deduce it from the geometric Breuil–Mézard conjecture below (see Corollary 7.4.4).

3. Tame types, Serre weights, and base change

In this section we describe various results about tame types, and we explain how these results interact with base change for tame types, Serre weights and Galois representations. We restrict ourselves to quadratic unramified base change, which is all that we will need in this paper.

3.1. **Tame types.** In this section we write $K = \mathrm{GL}_2(\mathcal{O}_{F_v})$, and let I denote the standard Iwahori subgroup of K (i.e. the one whose reduction mod p consists of upper-triangular matrices). We consider tame types for K; that is, irreducible E-representations of K that arise by inflation from the irreducible E-representations of $\mathrm{GL}_2(k_v)$. By the results recalled in Section 1 of [Dia07], these representations are either principal series, cuspidal, one-dimensional, or twists of the Steinberg representation. In this paper we will only consider the principal series and cuspidal types, and all tame types occurring will be assumed to be of this kind without further comment. Under the inertial local Langlands correspondence of Section 1.9, this means that we will only consider non-scalar tame inertial types.

3.1.1. **Lemma.** *Any tame type σ is residually multiplicity free (in the sense that the Jordan–Hölder constituents of $\overline{\sigma}$ each appear with multiplicity one) and essentially self-dual (in the sense that there is an isomorphism $\sigma^\vee \otimes \chi \cong \sigma$ for some character $\chi : K \to \mathcal{O}^\times$). All of the Jordan–Hölder factors $\overline{\sigma}_i$ of $\overline{\sigma}$ also satisfy $\overline{\sigma}_i^\vee \otimes \overline{\chi} \cong \overline{\sigma}_i$. Furthermore, σ has a model over an unramified extension of $\mathbb{Q}_p$.*

Proof. That σ is residually multiplicity free follows at once from Propositions 1.1 and 1.3 of [Dia07], and the essential self-duality follows from the character tables for σ in Section 1 of *ibid.* These character tables also show that the character of σ is defined over an unramified extension of $\mathbb{Q}_p$, and that σ has dimension prime to p; thus the Schur index of σ is 1 by Theorem 2(a) of [Gow76], and σ is defined over the field of definition of its character. $\square$

We refer to these representations as tame types because they correspond to tame inertial types under the inertial local Langlands correspondence of Henniart's appendix to [BM02]. It is easy to see that a tame inertial type is reducible, and is either the sum of two copies of a single character which extends to G_{F_v} (the *scalar* case), the sum of two distinct characters which extend to G_{F_v} (the *principal series* case), or a sum $\psi \oplus \psi^{q_v}$ where ψ does not extend to G_{F_v} but does extend to G_{L_v} (the *cuspidal* case). In the principal series case, if $\tau = \eta_1 \oplus \eta_2$ then $\sigma(\tau)$ is the representation denoted $\sigma(\eta_1 \circ \mathrm{Art}_{F_v} \otimes \eta_2 \circ \mathrm{Art}_{F_v})$ below, while in the cuspidal case, if $\tau = \psi \oplus \psi^{q_v}$, then $\sigma(\tau)$ is the representation denoted $\Theta(\psi \circ \mathrm{Art}_{F_v})$ below. For typesetting reasons we will always write $\overline{\sigma}(\tau)$, $\overline{\Theta}(\psi)$, etc. in lieu of $\overline{\sigma(\tau)}$, $\overline{\Theta(\psi)}$, etc.

3.2. **Principal series types.** We first consider tame principal series types. We let $\eta, \eta' : k_v^\times \to \mathcal{O}^\times$ be two multiplicative characters, and we write $\chi = \eta \otimes \eta' : I \to \mathcal{O}^\times$ for the character

$$\begin{pmatrix} a & b \\ pc & d \end{pmatrix} \mapsto \eta(\overline{a})\eta'(\overline{d}).$$

We write $\chi^s := \eta' \otimes \eta$, and we write $\sigma(\chi) := \mathrm{Ind}_I^K \chi^s$, an E-vector space with an action of K.

We will assume from now on that $\eta \neq \eta'$, and we write $c = \sum_{j=0}^{f-1} c_j p^j$, $0 \leq c_j \leq p-1$, for the unique integer $0 < c_\chi < p^f - 1$ such that $\eta(\overline{x})\eta'^{-1}(\overline{x}) = [\overline{x}]^{c_\chi}$, where $\overline{x} \in k_v^\times$ and $[\overline{x}]$ is the Teichmüller lift of $\overline{x}$. We let $\mathcal{P}_{\sigma(\chi)}$ be the collection of subsets of $\mathcal{S}$ consisting of subsets J satisfying the conditions:

- if $j \in J$ and $j-1 \notin J$ then $c_j \neq p-1$, and
- if $j \notin J$ and $j-1 \in J$ then $c_j \neq 0$,

where $j-1$ is taken to mean $f-1$ if $j=0$.

The Jordan–Hölder factors of $\overline{\sigma}(\chi)$ are parameterised by $\mathcal{P}_{\sigma(\chi)}$ in the following fashion (see Lemma 2.2 of [BP12] or Proposition 1.1 of [Dia07]). For any $J \subseteq \mathcal{S}$, we let δ_J denote the characteristic function of J, and if $J \in \mathcal{P}_{\sigma(\chi)}$ we define $s_{J,i}$ by

$$s_{J,i} = \begin{cases} p-1-c_i-\delta_{J^c}(i-1) & \text{if } i \in J \\ c_i - \delta_J(i-1) & \text{if } i \notin J, \end{cases}$$

and we set $t_{J,i} = c_i + \delta_{J^c}(i-1)$ if $i \in J$ and 0 otherwise. Then we let $\overline{\sigma}(\chi)_J := \overline{\sigma}_{\vec{t},\vec{s}} \otimes \eta' \circ \det$; the $\overline{\sigma}(\chi)_J$ are precisely the Jordan–Hölder factors of $\overline{\sigma}(\chi)$. If $\tau = \eta \oplus \eta'$ corresponds to $\sigma(\chi)$ under the inertial local Langlands correspondence, we will often write $\overline{\sigma}(\tau)_J$ for $\overline{\sigma}(\chi)_J$. However, we caution that this notation depends implicitly on an ordering of the characters η, η', namely $\overline{\sigma}(\tau)_J = \overline{\sigma}(\eta \otimes \eta')_J = \overline{\sigma}(\eta' \otimes \eta)_{J^c}$.

3.3. **Cuspidal types.** Let $\psi : l_v^\times \to \mathcal{O}^\times$ be a multiplicative character which does not factor through the norm $l_v^\times \to k_v^\times$. We now recall the Jordan–Hölder factors of the reduction modulo p of the tame cuspidal type $\Theta(\psi)$, following [Dia07]. We can write ψ in the form $[\overline{x}]^{(q+1)b+1+c}$ where $0 \le b \le q-2$, $0 \le c \le q-1$. Write $c = \sum_{i=0}^{f-1} c_i p^i$ where $0 \le c_i \le p-1$. If $J \subseteq \mathcal{S}$ we set $J_0 = J \triangle \{f-1\}$, and we define $\mathcal{P}_{\Theta(\psi)}$ to be the collection of subsets of $\mathcal{S}$ consisting of those J satisfying the conditions:

- if $j \in J$ and $j-1 \notin J_0$ then $c_j \neq p-1$, and
- if $j \notin J$ and $j-1 \in J_0$ then $c_j \neq 0$.

The Jordan–Hölder factors of $\overline{\Theta}(\psi)$ are parameterised by $\mathcal{P}_{\Theta(\psi)}$ in the following fashion. (See Proposition 1.3 of [Dia07], but note that our parameterisation is slightly different: the Jordan–Hölder factor parameterised by J here is parameterised by J^c in [Dia07].) For any $J \in \mathcal{P}_{\Theta(\psi)}$ we define $s_{J,i}$ by

$$s_{J,i} = \begin{cases} p-1-c_i-\delta_{(J_0)^c}(i-1) & \text{if } i \in J \\ c_i - \delta_{J_0}(i-1) & \text{if } i \notin J, \end{cases}$$

and we set $t_{J,i} = c_i + \delta_{J^c}(i-1)$ if $i \in J$ and 0 otherwise. Then we let

$$\overline{\Theta}(\psi)_J := \overline{\sigma}_{\vec{t},\vec{s}} \otimes [\cdot]^{(q+1)b+\delta_J(0)\delta_J(f-1)+\delta_{J^c}(0)\delta_{J^c}(f-1)} \circ \det;$$

the $\overline{\Theta}(\psi)_J$ are precisely the Jordan–Hölder factors of $\overline{\Theta}(\psi)$.

We say that $\Theta(\psi)$ is *regular* if there is some i with $0 < c_i < p-1$, and *irregular* otherwise. If τ is a cuspidal tame inertial type, then we say that τ is regular if and only if $\sigma(\tau)$ is regular.

3.3.1. **Lemma.** *The cuspidal type $\Theta(\psi)$ is regular if and only if there is some $i \in \mathcal{S}$ such that $\{i, \dots, f-1\} \in \mathcal{P}_{\Theta(\psi)}$ and $\{0, \dots, i-1\} \in \mathcal{P}_{\Theta(\psi)}$. If $\Theta(\psi)$ is irregular, then there is a unique $J \in \mathcal{P}_{\Theta(\psi)}$ such that $\overline{\Theta}(\psi)_J$ is regular.*

Proof. If $\Theta(\psi)$ is regular, then it is easy to see that any i with $0 < c_i < p-1$ works. The converse is also straightforward. If $\Theta(\psi)$ is irregular and $\overline{\Theta}(\psi)_J$ is regular, we see that if $c_i = 0$ then we must have $i-1 \notin J_0$, and if $c_i = p-1$ then $i-1 \in J_0$, so J is uniquely determined. $\square$

3.4. **Base change for tame types.** Suppose first that $\sigma(\chi)$ is a tame principal series type for $\mathrm{GL}_2(\mathcal{O}_{F_v})$, where $\chi = \eta \otimes \eta'$ for multiplicative characters $\eta, \eta' : k_v^\times \to \mathcal{O}_{F_v}^\times$. Then we let $\mathrm{BC}(\sigma(\chi)) = \sigma(\chi')$, where $\chi' = \eta \circ N_{l_v/k_v} \otimes \eta' \circ N_{l_v/k_v}$, a tame principal series type for $\mathrm{GL}_2(\mathcal{O}_{L_v})$.

Suppose now that $\sigma(\tau)$ is a tame cuspidal type for $\mathrm{GL}_2(\mathcal{O}_{F_v})$, so that $\sigma(\tau) = \Theta(\psi)$, where $\psi : l_v^\times \to \mathcal{O}_{L_v}^\times$ is a multiplicative character which does not factor through N_{l_v/k_v} (see e.g. Section 1 of [Dia07]). Then we let $\mathrm{BC}\,(\sigma(\tau)) = \sigma(\chi')$, where $\chi' = \psi \otimes \psi^q$, a tame principal series type for $\mathrm{GL}_2(\mathcal{O}_{L_v})$. Note that this definition depends on the ψ, as opposed to only on $\sigma(\tau)$, so that the notation below for the Jordan–Hölder factors also depends on this choice. However, all of our main results are independent of this choice.

In either case we say that $\mathrm{BC}\,(\sigma(\tau))$ is the *base change* of $\sigma(\tau)$ to $\mathrm{GL}_2(\mathcal{O}_{L_v})$. Again, not every tame type for $\mathrm{GL}_2(\mathcal{O}_{L_v})$ is the base change of a tame type for $\mathrm{GL}_2(\mathcal{O}_{F_v})$.

By construction, this definition is compatible with the inertial local Langlands correspondence, in the sense that if τ is a tame inertial type for F_v, and if $\mathrm{BC}\,(\tau)$ denotes τ regarded as a tame intertial type for L_v, via the equality $I_{F_v} = I_{L_v}$, then $\mathrm{BC}\,(\sigma(\tau)) = \sigma(\mathrm{BC}\,(\tau))$.

If we write $\mathrm{BC}\,(\sigma(\tau)) = \sigma(\eta \otimes \eta')$, then an easy calculation shows that $\eta(\eta')^{-1} = [\cdot]^{\sum_{i=0}^{2f-1} a_i p^i}$ where in the principal series case $a_i = c_i$ if $0 \le i \le f-1$ and $a_i = c_{i-f}$ if $f \le i \le 2f-1$, while in the cuspidal case $a_i = c_i$ if $0 \le i \le f-1$ and $c_i = p-1-c_{i-f}$ if $f \le i \le 2f-1$.

Now, if $J \subseteq \mathcal{S}$ we write $\mathrm{BC}_{\mathrm{cusp}}(J)$ for the subset of $\{0, \dots, 2f-1\}$ defined by taking $i \in \mathrm{BC}_{\mathrm{cusp}}(J)$ if and only if either $0 \le i \le f-1$ and $i \in J$, or $f \le i \le 2f-1$ and $i - f \notin J$. Similarly, we write $\mathrm{BC}_{\mathrm{PS}}(J)$ for those $i \in \{0, \dots, 2f-1\}$ with $i \pmod f \in J$.

With these definitions, the following lemma is a tedious but elementary consequence of the above discussion in the cuspidal case, and is almost trivial in the principal series case.

3.4.1. **Lemma.** *Let $\sigma(\tau)$ be a tame type. If τ is a principal series (respectively cuspidal) type, then $\mathrm{BC}_{\mathrm{PS}}(J) \in \mathcal{P}_{\mathrm{BC}\,(\sigma(\tau))}$ and $\mathrm{BC}\,(\overline{\sigma}(\tau)_J) = \overline{\sigma}(\mathrm{BC}\,(\tau))_{\mathrm{BC}_{\mathrm{PS}}(J)}$ (respectively $\mathrm{BC}_{\mathrm{cusp}}(J) \in \mathcal{P}_{\mathrm{BC}\,(\tau)}$ and $\mathrm{BC}\,(\overline{\sigma}(\tau)_J) = \overline{\sigma}(\mathrm{BC}\,(\tau))_{\mathrm{BC}_{\mathrm{cusp}}(J)}$).*

3.5. **Tame types and Serre weights.** The following results will be useful in our later arguments.

3.5.1. **Proposition.** *Suppose that $\overline{\rho}$ is generic, and that $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$. Then there is a (non-scalar) tame inertial type τ such that $\mathrm{JH}(\overline{\sigma}(\tau)) \cap \mathcal{D}(\overline{\rho}) = \overline{\sigma}$.*

Proof. In the case that $\overline{\sigma} = \overline{\sigma}_{\vec{t},\vec{s}}$ with $1 \le s_i \le p-3$ for all i, then $\overline{\sigma}$ is regular in the sense of [Gee11], and the result is an immediate consequence of [Gee11, Prop. 3.10, Prop. 4.2]. In the more general setting that $0 \le s_i \le p-2$ for all i, we will see that the proof of [Gee11, Prop. 3.10] (the case where $\overline{\rho}$ reducible) goes through with only minor changes, while the irreducible case will follow from the reducible case by a base change argument. For ease of reference we will freely use the notation of [Gee11], with the minor changes that we will omit the letters τ and σ from the notation (so that for example we write c_i for c_{τ_i}) and we continue to write f where [Gee11] uses r. Our weight $\overline{\sigma}_{\vec{t},\vec{s}}$ is the weight $\overline{\sigma}_{\vec{a},\vec{b}}$ of [Gee11], where $a_i = t_{f-i}$ and $b_i = s_{f-i} + 1$.

Suppose first that $\overline{\rho}$ is reducible, and define c_i as in Section 3.5 of [Gee11]. Note firstly that the assumption of [Gee11] that not all c_i are equal to 0 and not all c_i are equal to $p-1$ is satisfied, as this could only fail to hold if $b_i = 1$ for all i, or equivalently $s_i = 0$ for all i, which would contradict Lemma 2.1.6. Then the proof of Proposition 3.10 of [Gee11] goes through unchanged, until the sentence beginning "By the assumption that $\overline{\sigma}_{\vec{a},\vec{b}}$ is regular". We may rewrite the displayed equation before this sentence as

$$\prod_{i\in S}\omega_i^{c_i+\delta_J(i+1)}=\prod_{i\in X}\omega_i^{c_i+\delta_{J'\cup(K'\cap J'^c)}(i+1)}\prod_{i\in X^c}\omega_i^{\delta_{(K'\cap J'^c)}(i+1)}, \tag{3.1}$$

where $X := (J'\cap K')\cup(J'^c\cap K'^c)$.

Let ψ be the character given by either side of (3.1). It is easy to check that the ratio of the two characters occurring in $\overline{\rho}$ is equal to $\psi\prod_{i\notin J}\omega_{i-1}^{-1}$. Since the genericity of $\overline{\rho}$ implies that this ratio (taken in either order) does not have the form $\prod_{i\in S}\omega_i^{d_i}$ with $d_i=0,1$ for all i, it follows in particular that ψ is neither trivial nor equal to $\prod_{i\in S}\omega_i$.

Now, we have $1\le b_i\le p-1$ for all i, and so we see from the definition of c_i that $0\le c_i\le p-1$ and $1\le c_i+\delta_J(i+1)\le p-1$. We claim moreover that no exponent on the right-hand side of (3.1) is equal to p, i.e., that no "carrying" can take place in (3.1). Indeed the only possibility for "carrying" is when $i\in X$, $c_i=p-1$, and $i+1\in J'\cup(K'\cap J'^c)$. But after carrying out all carries on the right hand side, the exponent of ω_i will be at most 1, whereas it is $p-1$ on the left hand side, a contradiction since ψ is non-trivial.

In particular (again using the fact that ψ is non-trivial) we can equate exponents on each side of (3.1). If $i\in X^c$ we must have $i+1\in K'\cap J'^c\subseteq X^c$, so that if X^c is nonempty then $X^c=S$, and every exponent on either side is equal to 1. This contradicts our observation that $\psi\ne\prod_{i\in S}\omega_i$. Thus $X=S$, so that $J'=K'$, and equating exponents shows that $J=J'$, as required.

Suppose next that $\overline{\rho}$ is irreducible (and let the characters ω_i now be fundamental characters of level $2f$). Let $I'_{\vec{a},\vec{b},J}$ be the cuspidal type constructed in [Gee11, §4.1]. One checks exactly as in the proof of [Gee11, Prop. 4.1] that $\overline{\sigma}_{\vec{a},\vec{b}}$ is a Jordan–Hölder factor of $\overline{I}'_{\vec{a},\vec{b},J}$. From the discussion of base change in Sections 2.2 and 3.4 we see that if $\overline{\sigma}\in\mathcal{D}(\overline{\rho})\cap\mathrm{JH}(\overline{I}'_{\vec{a},\vec{b},J})$ then $\mathrm{BC}\,(\overline{\sigma})$ lies in the intersection of $\mathcal{D}(\overline{\rho}|_{G_{L_v}})$ with the set of Jordan–Hölder factors of the reduction mod p of $\mathrm{BC}\,(I'_{\vec{a},\vec{b},J})$, and so it suffices to prove that this latter intersection has size 1.

Set $\vec{a}'=(a_0,\dots,a_{f-1},a_0,\dots,a_{f-1})$ and define $\vec{b}'$ similarly , so that $\mathrm{BC}\,(\overline{\sigma}_{\vec{a},\vec{b}})=\overline{\sigma}_{\vec{a}',\vec{b}'}$. We claim that $\mathrm{BC}\,(I'_{\vec{a},\vec{b},J})$ is the principal series type $I_{\vec{a}',\vec{b}',J}$ (in the notation of [Gee11, §3.5], but with $\vec{a}',\vec{b}'$ in place of $\vec{a},\vec{b}$) constructed in the reducible case for the representation $\overline{\rho}|_{G_{L_v}}$. Once we have checked this claim, the irreducible case will follow immediately from our proof of the reducible case.

To see the claim, one calculates directly from the definitions that the character $\tilde{\psi}_{\vec{a},\vec{b},J}\tilde{\omega}_f\prod_{i=1}^{f}\tilde{\omega}_i^{c_i}$ of [Gee11, §4.1] is the lift of $\chi:=\prod_{i=0}^{2f-1}\omega_i^{a_i}\prod_{i\in J}\omega_i^{b_i-p}$. Since J is antisymmetric we have $\chi^q=\prod_{i=0}^{2f-1}\omega_i^{a_i}\prod_{i\notin J}\omega_i^{b_i-p}$; in other words χ^q is equal to the character $\chi_{\vec{a}',\vec{b}',J}$ of [Gee11, §3.5]. Another straightforward calculation shows that $\chi_{\vec{a}',\vec{b}',J}\prod_{i=0}^{2f-1}\omega_i^{c_i}=\chi$ (with c_i as in [Gee11, §3.5] now), so that

$$I_{\vec{a}',\vec{b}',J}=I(\tilde{\chi}_{\vec{a}',\vec{b}',J},\tilde{\chi}_{\vec{a}',\vec{b}',J}\prod_i\tilde{\omega}_i^{c_i})=I(\tilde{\chi}^q,\tilde{\chi})\cong\mathrm{BC}\,(I'_{\vec{a},\vec{b},J})$$

as claimed. □

3.5.2. **Proposition.** *Suppose that $\overline{\rho}$ is generic. Then there is a (non-scalar) tame inertial type τ such that $\mathcal{D}(\overline{\rho})\subset\mathrm{JH}(\overline{\sigma}(\tau))$.*

Proof. Since $\mathcal{D}(\overline{\rho}) \subset \mathcal{D}(\overline{\rho}^{\mathrm{ss}})$ we may assume that $\overline{\rho}$ is semisimple, and the result now follows immediately from Theorems 2.1 and 3.2 of [Dia07] (since $\overline{\rho}$ is generic, Lemma 2.1.6 shows that there no exceptional weights in the terminology of [Dia07]). □

4. Lattices in residually multiplicity-free representations

In this section we record some general facts about lattices in (absolutely) irreducible representations whose reductions have all Jordan–Hölder factors occurring with multiplicity one.

Throughout this section we let Γ be a group, and we let σ be a finite-dimensional E-representation of Γ.

4.1. **Lattices and gauges.** We say that σ is *residually multiplicity free* if the Jordan–Hölder factors of $\overline{\sigma}$ occur with multiplicity one. The following lemma and its proof are well-known, but for lack of a convenient reference we give the details here.

4.1.1. **Lemma.** *Suppose that σ is irreducible and is residually multiplicity free. Let $\overline{\sigma}_i$ be a Jordan–Hölder factor of $\overline{\sigma}$. Then there is up to homothety a unique Γ-stable $\mathcal{O}$-lattice σ_i° in σ such that the socle of $\overline{\sigma}_i^\circ$ is precisely $\overline{\sigma}_i$. Similarly, there is up to homothety a unique Γ-stable $\mathcal{O}$-lattice σ_i° in σ such that the cosocle of $\overline{\sigma}_i^\circ$ is precisely $\overline{\sigma}_i$.*

Proof. We begin by proving the result for socles. For existence, let L be any Γ-stable $\mathcal{O}$-lattice in σ, and consider σ/L. Let M be the maximal $\mathcal{O}[\Gamma]$-submodule of σ/L such that none of the Jordan–Hölder factors of M is isomorphic to $\overline{\sigma}_i$. If M had infinite length, then the inverse image of M in σ would contain a non-zero Γ-stable E-subspace, so would be equal to σ, a contradiction. Thus M has finite length, and its inverse image in σ gives the required lattice.

For the uniqueness, suppose that L, L' are two Γ-stable lattices in σ with socle $\overline{\sigma}_i$. After scaling, we may, without loss of generality, suppose that $L' \subseteq L$ but that $L' \not\subseteq \varpi_E L$. Consider the induced map $L'/\varpi_E L' \to L/\varpi_E L$; it is non-zero by construction. If this is not an isomorphism, then (for length reasons) it is not surjective, and thus has non-trivial kernel, which necessarily contains the socle $\overline{\sigma}_i$ of $L'/\varpi_E L'$. Similarly, its image, being non-zero, contains the socle of $L/\varpi_E L$. Thus $\overline{\sigma}_i$ occurs at least twice in $L'/\varpi_E L'$, contradicting our assumption that σ is residually multiplicity free.

The statement for cosocles follows from the statement for socles via a consideration of contragredients. More precisely, the contragredient representation $\sigma^\vee$ is also residually multiplicity free, and the Jordan–Hölder factors of $\overline{\sigma}^\vee$ are the contragredients of the Jordan–Hölder factors of $\overline{\sigma}$. If L is a Γ-invariant lattice in σ, then its $\mathcal{O}$-dual $L^\vee$, equipped with the contragredient Γ-representation, is naturally a Γ-invariant lattice in $\sigma^\vee$. Furthermore, the cosocle of $L/\varpi_E L$ is contragredient to the socle of $L^\vee/\varpi_E L^\vee$. The case of the lemma for cosocles then follows from the case for socles, applied to $\sigma^\vee$. □

In fact, we will have more cause to apply the preceding lemma in the case of cosocles than in the case of socles. Note that the formation of cosocles is right exact, since the cosocle functor is the direct sum

$$L \mapsto \oplus_{\overline{\sigma}} \operatorname{Hom}_\Gamma(L, \overline{\sigma}) \otimes_{\mathbb{F}} \overline{\sigma},$$

where $\overline{\sigma}$ runs over all irreducible representations of Γ over $\mathbb{F}$. Then we have, for example, the following result.

4.1.2. **Lemma.** *Suppose that σ is irreducible and residually multiplicty free. Let σ° be a Γ-invariant lattice in σ, let $\{\overline{\sigma}_i\}$ denote the set of constituents of the cosocle of σ°, and for each $\overline{\sigma}_i$, choose a Γ-invariant lattice σ_i° whose cosocle is isomorphic to $\overline{\sigma}_i$, such that $\sigma_i^\circ \subseteq \sigma^\circ$, but such that $\sigma_i^\circ \not\subseteq \varpi_E\sigma^\circ$. (Note that σ_i° is uniquely determined, by Lemma 4.1.1.) Then $\sigma^\circ = \sum_i \sigma_i^\circ$.*

Proof. Write $\sigma^{\circ,\prime} := \sum_{i=1}^m \sigma_i^\circ \subseteq \sigma^\circ$. This inclusion induces an isomorphism on cosocles, from which we see that $\sigma^\circ/\sigma^{\circ,\prime}$ has trivial cosocle, and thus vanishes; so $\sigma^\circ = \sigma^{\circ,\prime}$, as required. □

Let us now label the Jordan–Hölder factors of $\overline{\sigma}$ as $\overline{\sigma}_1, \ldots, \overline{\sigma}_n$, and for each value of $i = 1, \ldots, n$, fix a choice of lattice σ_i° (unique up to homothety, by Lemma 4.1.1) such that the cosocle of $\overline{\sigma}_i^\circ$ equals $\overline{\sigma}_i$.

4.1.3. **Definition.** If σ° is any Γ-invariant $\mathcal{O}$-lattice in σ, define for each i the fractional ideal $I_i := \{a \in E \,|\, a\sigma_i^\circ \subseteq \sigma^\circ\}$, and define the *gauge* of σ° to be the n-tuple $(I_i)_{i=1,\ldots,n}$, thought of as an n-tuple of fractional ideals in E. We denote the gauge of σ° by $\mathcal{G}(\sigma^\circ)$.

4.1.4. **Proposition.** *Suppose that σ is irreducible and residually multiplicity free. A Γ-invariant $\mathcal{O}$-lattice σ° in σ is determined by its gauge $\mathcal{G}(\sigma^\circ)$; indeed, if we fix a generator ϕ_i for the fractional ideal I_i, then we have $\sigma^\circ = \sum_{i=1}^n \phi_i(\sigma_i^\circ)$.*

Proof. Write $\sigma^{\circ,\prime} := \sum_{i=1}^n \phi_i(\sigma_i^\circ) \subseteq \sigma^\circ$. Note that evidently $\sigma^{\circ,\prime}$ is well-defined independently of the choice of the generators ϕ_i, and hence depends only on the gauge of σ°. The proposition will follow if we prove that $\sigma^{\circ,\prime} = \sigma^\circ$; however, this follows immediately from Lemma 4.1.2. (Indeed, that lemma shows that we may restrict the sum defining $\sigma^{\circ,\prime}$ to those indices i for which $\overline{\sigma}_i$ is a constituent of the cosocle of σ°.) □

4.2. **Lattices in essentially self-dual representations.** In this subsection we assume that Γ is finite, and that Γ admits a character $\chi : \Gamma \to \mathcal{O}^\times$ for which there is a Γ-equivariant isomorphism $\sigma \cong \sigma^\vee \otimes \chi$. We let $\overline{\chi} : \Gamma \to k_E^\times$ denote the reduction mod ϖ_E of χ.

4.2.1. **Proposition.** *Suppose that σ is absolutely irreducible and that σ° is a Γ-invariant lattice in σ such that $\mathrm{cosoc}(\sigma^\circ)$ is absolutely irreducible and appears in $\mathrm{JH}(\overline{\sigma})$ with multiplicity one, and such that there is an isomorphism $\mathrm{cosoc}(\sigma^\circ) \cong \mathrm{cosoc}(\sigma^\circ)^\vee \otimes \overline{\chi}$. Then there is a Γ-equivariant embedding $(\sigma^\circ)^\vee \otimes \chi \hookrightarrow \sigma^\circ$ whose cokernel has exponent dividing $|\Gamma|$.*

Proof. Let P denote a projective envelope of σ° as an $\mathcal{O}[\Gamma]$-module. Then $P^\vee \otimes \chi$ is again a projective $\mathcal{O}[\Gamma]$-module, and there are Γ-equivariant isomorphisms

$$\begin{aligned}\mathrm{cosoc}(P^\vee \otimes \chi) = \mathrm{cosoc}(\overline{P}^\vee) \otimes \overline{\chi} \cong \mathrm{soc}(\overline{P})^\vee \otimes \overline{\chi} &\overset{(1)}{\cong} \mathrm{cosoc}(\overline{P})^\vee \otimes \overline{\chi} \\ = \mathrm{cosoc}(P)^\vee \otimes \overline{\chi} = \mathrm{cosoc}(\sigma^\circ)^\vee \otimes \overline{\chi} &\overset{(2)}{\cong} \mathrm{cosoc}(\sigma^\circ) = \mathrm{cosoc}(P),\end{aligned}$$

the isomorphism (1) following from the isomorphism $\mathrm{soc}(\overline{P}) \cong \mathrm{cosoc}(\overline{P})$ (a standard property of projective $k_E[\Gamma]$-modules), and the isomorphism (2) holding by assumption. The uniqueness up to isomorphism of projective envelopes thus implies that there is a Γ-equivariant isomorphism

$$P^\vee \otimes \chi \cong P. \tag{4.1}$$

The Γ-equivariant surjection $P \to \sigma^\circ$ (expressing the fact that P is a projective envelope of σ°) induces a saturated Γ-equivariant embedding $(\sigma^\circ)^\vee \otimes \chi \hookrightarrow P^\vee \otimes \chi$ (*saturated* meaning that the cokernel is ϖ_E-torsion free). Composing this with the isomorphism (4.1), we obtain a saturated Γ-equivariant embedding

$$(\sigma^\circ)^\vee \otimes \chi \hookrightarrow P. \tag{4.2}$$

Recall that there is an idempotent $e_\sigma \in \dfrac{1}{|\Gamma|}\mathcal{O}[\Gamma] \subseteq E[\Gamma]$ which, when applied to any Γ-representation V over E, projects onto the σ-isotypic part of V. We claim that $e_\sigma(E \otimes_{\mathcal{O}} P)$ consists of a single copy of σ. Indeed, since σ is absolutely irreducible, the number of such copies of τ is equal to the dimension of

$$\mathrm{Hom}_{E[\Gamma]}(E \otimes_{\mathcal{O}} P, \sigma) = E \otimes_{\mathcal{O}} \mathrm{Hom}_{\mathcal{O}[\Gamma]}(P, \sigma^\circ).$$

It thus suffices to show that $\mathrm{Hom}_{\mathcal{O}[\Gamma]}(P, \sigma^\circ)$, which is a finite rank free $\mathcal{O}$-module, is of rank one. To this end, we compute (using the projectivity of P) that

$$\begin{aligned}\mathrm{Hom}_{\mathcal{O}[\Gamma]}(P, \sigma^\circ)/\varpi_E \mathrm{Hom}_{\mathcal{O}[\Gamma]}(P, \sigma^\circ) &= \mathrm{Hom}_{k_E[\Gamma]}(\overline{P}, \overline{\sigma}^\circ) \\ &= \mathrm{Hom}_{k_E[\Gamma]}(\mathrm{cosoc}(P), \overline{\sigma}^{\mathrm{ss}}).\end{aligned}$$

The last of these spaces is one-dimensional, as by assumption $\mathrm{cosoc}(P) = \mathrm{cosoc}(\sigma^\circ)$ is absolutely irreducible and appears in $\overline{\sigma}^{\mathrm{ss}}$ with multiplicity one, and so our claim is proved.

Since (4.2) is a saturated embedding, it follows that its image is equal to $e_\sigma(E \otimes_{\mathcal{O}} P) \cap P$, and the composite of (4.2) and the projection $P \to \sigma^\circ$ is a Γ-equivariant embedding $(\sigma^\circ)^\vee \otimes \chi \hookrightarrow \sigma^\circ$. Since $|\Gamma| e_\sigma \in \mathcal{O}[\Gamma]$, we easily see that the image of this embedding contains $|\Gamma|\sigma^\circ$; indeed, $|\Gamma| e_\sigma P \subseteq e_\sigma(E \otimes_{\mathcal{O}} P) \cap P$, and the image of $|\Gamma| e_\sigma P$ under the projection to σ° is precisely $|\Gamma|\sigma^\circ$. This completes the proof of the proposition. ☐

It seems worthwhile to record the following result, which relates the conclusion of the preceding proposition to the length of the socle and cosocle filtrations of $\overline{\sigma}^\circ$, although in the particular context that we consider in the sequel, we will need a more precise statement (Theorem 5.2.4 below).

4.2.2. **Proposition.** *Continue to assume that $\sigma \cong \sigma^\vee \otimes \chi$. Suppose that σ is absolutely irreducible and residually multiplicity free, and that for each Jordan–Hölder factor $\overline{\sigma}_i$ of $\overline{\sigma}$, there is a Γ-equivariant isomorphism $\overline{\sigma}_i^\vee \otimes \overline{\chi} \cong \overline{\sigma}_i$. If σ° is any Γ-invariant lattice in σ for which there is a Γ-equivariant embedding*

$$(\sigma^\circ)^\vee \otimes \chi \hookrightarrow \sigma^\circ \tag{4.3}$$

whose cokernel is annihilated by ϖ_E^m, then the socle and cosocle filtrations of $\overline{\sigma}^\circ$ are both of length at most $m+1$.

Proof. If $m = 0$ then (4.3) is an isomorphism, and consequently there is a Γ-equivariant isomorphism $(\overline{\sigma}^\circ)^\vee \otimes \chi \cong \overline{\sigma}^\circ$. Passing to cosocles, we obtain an isomorphism $\mathrm{soc}(\overline{\sigma}^\circ)^\vee \otimes \overline{\chi} \cong \mathrm{cosoc}(\overline{\sigma}^\circ)$, which, by our hypotheses, may be rewritten as an isomorphism $\mathrm{soc}(\overline{\sigma}^\circ) \cong \mathrm{cosoc}(\overline{\sigma}^\circ)$. Since each Jordan–Hölder factor of $\overline{\sigma}$ appears with multiplicity one, this implies a literal equality of the socle and cosocle of $\overline{\sigma}^\circ$, i.e. it implies that $\overline{\sigma}^\circ$ is semisimple.

We treat the cases $m \geq 1$ by induction on m. Obviously it is no loss of generality to assume that m is minimal, i.e. that the image of the embedding (4.3) does not lie in $\varpi_E\sigma^\circ$. Since σ is absolutely irreducible by assumption, so that the only endomorphisms of σ are scalars, this condition then determines the embedding (4.3) up to multiplication by an element of $\mathcal{O}^\times$. Dualizing (4.3), and twisting by χ, we obtain another embedding $(\sigma^\circ)^\vee \otimes \chi \hookrightarrow \sigma^\circ$ with the same property (that its image is not contained in $\varpi_E\sigma^\circ$), which thus coincides with (4.3) up to multiplication by an element of $\mathcal{O}^\times$. Consequently we find that the cokernel of (4.3) is isomorphic to its own Pontrjagin dual, up to twisting by χ.

In particular, if $m = 1$, then the cokernel τ of (4.3) is a representation of Γ over k_E which satisfies $\tau^\vee \otimes \overline{\chi} \cong \tau$. Since τ is a quotient of $\overline{\sigma}^\circ$, we furthermore have that every Jordan–Hölder factor $\overline{\sigma}_i$ of τ appears with multiplicity one, and satifies $\overline{\sigma}_i^\vee \otimes \chi \cong \overline{\sigma}_i$. Applying the same argument to τ as we applied to $\overline{\sigma}^\circ$ in the case $m = 0$, we conclude that τ is semisimple. A similar argument shows that the cokernel of the induced embedding $\varpi_E\sigma^\circ \hookrightarrow (\sigma^\circ)^\vee \otimes \chi$ is semisimple. Thus $\overline{\sigma}^\circ$ admits a two-step filtration whose associated graded pieces are semisimple, proving the proposition in this case.

Suppose finally that $m > 1$. Regard $(\sigma^\circ)^\vee \otimes \chi$ as a sublattice of σ° via (4.3), and let $\sigma'^\circ := \varpi_E\sigma^\circ + (\sigma^\circ)^\vee \otimes \chi$. Then $(\sigma'^\circ)^\vee \otimes \chi = \big(\varpi_E^{-1}(\sigma^\circ)^\vee \otimes \chi\big) \cap \sigma^\circ$, and so multiplication by ϖ_E induces a Γ-equivariant embedding

$$(\sigma'^\circ)^\vee \otimes \chi \hookrightarrow \sigma'^\circ,$$

whose cokernel is easily checked to be annihilated by ϖ_E^{m-1}. By induction we conclude that the socle and cosocle filtrations of $\overline{\sigma}'^\circ$ are of length at most m, or equivalently (multiplying by ϖ_E^{-1}), that the socle and cosocle filtrations of $\varpi_E^{-1}\overline{\sigma}'^\circ/\overline{\sigma}'^\circ$ are of length at most m. Since $\sigma'^\circ \subseteq \sigma^\circ \subseteq \varpi_E^{-1}\sigma'^\circ$, we conclude that the socle and cosocle filtrations of $\sigma^\circ/\sigma'^\circ$ are also of length at most m.

Now $\sigma'^\circ/\varpi_E\sigma^\circ$ is a Γ-subrepresentation of $\overline{\sigma}^\circ$ whose cosocle is a quotient of

$$\operatorname{cosoc}\big((\sigma^\circ)^\vee \otimes \chi\big) \cong \operatorname{soc}(\overline{\sigma}^\circ)^\vee \otimes \overline{\chi} \cong \operatorname{soc}(\overline{\sigma}^\circ)$$

(the second isomorphism holding by our hypotheses that all the Jordan–Hölder factors of $\overline{\sigma}$ are self-dual up to twisting by χ). Since all the Jordan–Hölder factors of $\overline{\sigma}^\circ$ appear with multiplicity one, in fact the cosocle of $\sigma'^\circ/\varpi_E\sigma^\circ$ must be a subrepresentation of the socle of $\overline{\sigma}^\circ$, and consequently $\sigma'^\circ/\varpi_E\sigma^\circ$ must be semisimple. Combining this with the conclusion of the preceding paragraph, we find that the socle and cosocle filtrations of $\overline{\sigma}^\circ$ are of length at most $m+1$, as required. $\square$

5. Lattices in tame types

In this section we elaborate on the ideas of the preceding section in the context of tame types (in the sense of Section 3). In the specific case of tame principal series types, related results may be found in Section 2 of [Bre12], and several of our arguments are generalisations of those found in [Bre12].

5.1. **Socle and cosocle filtrations in tame types.** If τ is a tame inertial type, assumed to be non-scalar, we write $\overline{\sigma}_J(\tau)$ for the Jordan–Hölder factor $\overline{\sigma}(\tau)_J$ defined in 3.2 and 3.3, and $\mathcal{P}_\tau$ for $\mathcal{P}_{\sigma(\tau)}$.

For each $J \in \mathcal{P}_\tau$, we fix $\mathrm{GL}_2(\mathcal{O}_{F_v})$-invariant $\mathcal{O}$-lattices $\sigma^{\circ,J}(\tau)$ and $\sigma^\circ_J(\tau)$ in $\sigma(\tau)$, so that the socle of $\overline{\sigma}^{\circ,J}(\tau)$ (resp. the cosocle of $\overline{\sigma}^\circ_J(\tau)$) is equal to $\overline{\sigma}_J(\tau)$. (In the case of $\sigma^\circ_J(\tau)$, this can be expressed more simply by saying that the cosocle of $\sigma^\circ_J(\tau)$ itself is equal to $\overline{\sigma}_J(\tau)$.) Lemma 4.1.1 shows that these lattices are uniquely determined up to scaling. The following theorem describes the cosocle and socle filtrations of the mod ϖ_E reductions of these lattices. Recall that we number the layers of the cosocle and socle filtrations starting at 0.

5.1.1. **Theorem.** *The i-th layer of the cosocle filtration of $\overline{\sigma}^\circ_J(\tau)$ (respectively, the i-th layer of the socle filtration of $\overline{\sigma}^{\circ,J}(\tau)$) consists precisely of the Jordan–Hölder factors $\overline{\sigma}_{J'}(\tau)$ with $|J \triangle J'| = i$. Furthermore, if $J_1, J_2 \in \mathcal{P}_\tau$ with $|J_1 \triangle J_2| = 1$, so that $\overline{\sigma}_{J_1}(\tau)$ and $\overline{\sigma}_{J_2}(\tau)$ occur in adjacent layers of the cosocle filtration of $\overline{\sigma}^\circ_J(\tau)$ (respectively, the socle filtration of $\overline{\sigma}^{\circ,J}(\tau)$), then the extension between $\overline{\sigma}_{J_1}(\tau)$ and $\overline{\sigma}_{J_2}(\tau)$ in this filtration is nonsplit.*

Suppose that $J_1, J_2 \in \mathcal{P}_\tau$ with $|J_1 \triangle J_2| = 1$. It follows straightforwardly from Theorem 5.1.1, together with Lemma 3.1.1, that every subquotient of $\overline{\sigma}^\circ_J(\tau)$ (respectively $\overline{\sigma}^{\circ,J}(\tau)$) with Jordan–Hölder factors precisely $\overline{\sigma}_{J_1}(\tau)$ and $\overline{\sigma}_{J_2}(\tau)$ must be non-semisimple.

The remainder of this subsection is devoted to proving Theorem 5.1.1. The proof requires a generalization of several of the results from Sections 3 and 4 of [BP12], and we begin by establishing these generalizations. To this end, we fix an f-tuple of integers $(r_i)_{0 \le i \le f-1}$, such that $0 \le r_i \le p-1$ for all i. Theorem 5.1.1 is straightforward when $f = 1$, and we will assume that $f > 1$ in the below.

In [BP12, §3], Breuil and Paškūnas define a representation V_r of $\mathrm{GL}_2(k_v)$ over $\mathbb{F}$ for each integer $r \ge 0$, as the reduction modulo ϖ_E of a certain $\mathcal{O}$-lattice in $\mathrm{Sym}^r E^2$; here we recall that k_v is regarded as embedded in $\mathbb{F}$ via $\overline{\kappa}_0$. (Actually Breuil and Paškūnas work with coefficients in $\overline{\mathbb{Z}}_p$, $\overline{\mathbb{Q}}_p$, and $\overline{\mathbb{F}}_p$, rather than $\mathcal{O}$, E, and $\mathbb{F}$, but it is evident that all their constructions and results descend to these latter choices of coefficients.) In particular if $0 \le r \le p-1$ then V_r is isomorphic to $(\mathrm{Sym}^r k_v^2) \otimes_{k_v, \overline{\kappa}_0} \mathbb{F}$. By convention we let $V_{-1} = 0$.

Suppose that V is an $\mathbb{F}$-representation of $\mathrm{GL}_2(k_v)$. In this section we follow the notation of [BP12] and denote by V^{Fr^i} the representation of $\mathrm{GL}_2(k_v)$ obtained by letting $\mathrm{GL}_2(k_v)$ act on V via twists by powers of the arithmetic Frobenius; so for instance the Serre weight $\overline{\sigma}_{\vec{t},\vec{s}}$ of Section 2 can be written $\bigotimes_{i=0}^{f-1} (V_{s_i} \otimes \det^{t_i})^{\mathrm{Fr}^i}$.

The tensor product $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ is then naturally a representation of $\mathrm{GL}_2(k_v)$, whose socle is equal to $\bigotimes_{i=0}^{f-1} (V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i}$, unless all of the $r_i = 0$, in which case the socle also contains $\bigotimes_{i=0}^{f-1} V_{p-1}$. (See the discussion of [BP12, §3].) In the case when all $r_i \le p-2$, the submodule structure of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ is analyzed quite precisely in [BP12, §§3,4]; our goal is to make a similar analysis in the case when some of the r_i are equal to $p-1$. In the case that *all* of the r_i are equal to $p-1$, then the representation is irreducible, and we exclude this case from our analysis below.

Following Section 3 of [BP12], we can describe the Jordan–Hölder factors of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ in the following way. We let $x_0, \ldots, x_{f-1}$ be variables, and we define the set $\mathcal{I}(x_0, \ldots, x_{f-1})$ of all f-tuples $\lambda := (\lambda_0(x_0), \ldots, \lambda_{f-1}(x_{f-1}))$ which satisfy the following conditions for each i (numbered cyclically):

- $\lambda_i(x_i) \in \{x_i, x_i - 1, x_i + 1, p-2-x_i, p-3-x_i, p-1-x_i\}$,

- if $\lambda_i(x_i) \in \{x_i, x_i - 1, x_i + 1\}$, then $\lambda_{i+1}(x_{i+1}) \in \{x_{i+1}, p - 2 - x_{i+1}\}$, and
- if $\lambda_i(x_i) \in \{p - 2 - x_i, p - 3 - x_i, p - 1 - x_i\}$, then $\lambda_{i+1}(x_{i+1}) \in \{x_{i+1} - 1, x_{i+1} + 1, p - 3 - x_{i+1}, p - 1 - x_{i+1}\}$.

For $\lambda \in \mathcal{I}(x_0, \ldots, x_{f-1})$, we define $e(\lambda) := \frac{1}{2}\big(\sum_{i=0}^{f-1} p^i(x_i - \lambda_i(x_i))\big)$ if $\lambda_{f-1}(x_{f-1}) \in \{x_{f-1}, x_{f-1} - 1, x_{f-1} + 1\}$, and $e(\lambda) := \frac{1}{2}\big(p^f - 1 + \sum_{i=0}^{f-1} p^i(x_i - \lambda_i(x_i))\big)$ otherwise. For each $\lambda \in \mathcal{I}(x_0, \ldots, x_{f-1})$, we define

$$\overline{\sigma}_\lambda := \bigotimes_{i=0}^{f-1} (V_{\lambda_i(r_i)} \otimes \det{}^{p-1-r_i})^{\mathrm{Fr}^i} \otimes \det{}^{e(\lambda)(r_0,\ldots,r_{f-1})}$$

provided that $0 \le \lambda_i(r_i) \le p - 1$ for all i, and we leave $\overline{\sigma}_\lambda$ undefined otherwise.

5.1.2. **Proposition.** *Each Jordan–Hölder factor of* $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ *is isomorphic to* $\overline{\sigma}_\lambda$ *for a uniquely determined* $\lambda \in \mathcal{I}(x_0, \ldots, x_{f-1})$.

Proof. This follows immediately from Lemma 3.2 of [BP12], which shows that every Jordan–Hölder factor of an injective envelope of $(V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i}$ is isomorphic to $\overline{\sigma}_\lambda$ for a uniquely determined $\lambda \in \mathcal{I}(x_0, \ldots, x_{f-1})$, and Lemma 3.4 of [BP12] (and the discussion immediately preceding it), which shows that $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ is isomorphic to a subrepresentation of such an injective envelope. $\square$

For any Jordan–Hölder factor $\overline{\sigma}$ of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$, we define an integer $0 \le l(\overline{\sigma}) \le f$ as follows: write $\overline{\sigma} \cong \overline{\sigma}_\lambda$, and let $l(\overline{\sigma})$ be the number of $0 \le i \le f - 1$ with $\lambda_i(x_i) \in \{p - 2 - x_i, p - 3 - x_i, p - 1 - x_i\}$.

Let $\mathcal{E} := \{i \,|\, r_i = p - 1\}$, and write $\mathcal{S}' := \mathcal{S} \setminus \mathcal{E}$. (This is the set denoted $S_{\mathbf{r}}$ in [BP12, §3].) We begin our analysis of the submodule structure of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ by defining subrepresentations F'_J of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ indexed by subsets $J \subseteq \mathcal{S}'$. Note that by [BP12, Lem. 3.5], there is an embedding

$$V_{r_i} \otimes \det{}^{p-1-r_i} \hookrightarrow V_{2p-2-r_i}, \tag{5.1}$$

uniquely determined up to a non-zero scalar (of course, when $r_i = p - 1$, the embedding is an isomorphism). For each $i \in \mathcal{S}$, define $F'_J(i)$ to be the image of this map if $i \notin J$, and to be V_{2p-2-r_i} if $i \in J$, and then define

$$F'_J := \bigotimes_{i=0}^{f-1} F'_J(i)^{\mathrm{Fr}^i} \subseteq \bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}.$$

These subrepresentations are closely related to the filtration of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ considered in [BP12] in the discussion preceding Lemma 3.8. More precisely, one has that

$$\mathrm{Fil}^j \bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i} = \sum_{|J| = |\mathcal{S}'| - j} F'_J.$$

The association $J \mapsto F'_J$ is order-preserving, and, if we write

$$F'_{<J} := \sum_{i \in J} F'_{J \setminus \{i\}} = \sum_{J' \subsetneq J} F'_{J'},$$

then $F'_J / F'_{<J} = W_J$, where W_J is a certain subquotient of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ defined in [BP12, §3], whose precise description is as follows. It follows from [BP12, Lem. 3.5] that the cokernel of (5.1) is isomorphic to $V_{p-2-r_i} \otimes V_1^{\mathrm{Fr}}$. Thus $W_J :=$

$\otimes_{i=0}^{f-1} W_J(i)^{\mathrm{Fr}^i}$, where $W_J(i)$ equals $V_{r_i} \otimes \det^{p-1-r_i}$ if $i \notin J$, and equals $V_{p-2-r_i} \otimes V_1^{\mathrm{Fr}}$ if $i \in J$.

If $\mathcal{E} = \varnothing$, then each W_J is semi-simple, and the filtration Fil^i coincides with the cosocle filtration of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$. However, if $\mathcal{E} \neq \varnothing$, these statements are no longer true; it is this case that we wish to analyze further.

Before continuing our analysis, we need to consider extensions of Serre weights.

5.1.3. **Proposition.** *Let $\overline{\sigma}_{\vec{s},\vec{t}}$ and $\overline{\sigma}_{\vec{s}',\vec{t}'}$ be two Serre weights, and if $f = 1$ then suppose that $s_0, s_0' \neq p - 1$. Then there is a non-zero extension between $\overline{\sigma}_{\vec{s},\vec{t}}$ and $\overline{\sigma}_{\vec{s}',\vec{t}'}$ if and only if the following conditions hold:*

1. *If $f = 1$, then $s_0' = p - 2 - s_0 \pm 1$, and $t_0' - t_0 \equiv s_0 + 1 - p\frac{(1\pm1)}{2} \pmod{p-1}$.*
2. *If $f > 1$, then there is some k such that $s_j = s_j'$ if $j \neq k, k+1$, $s_k' = p-2-s_k$, $s_{k+1}' = s_{k+1} \pm 1$, and $\sum_{j=0}^{p-1} p^j(t_j' - t_j) \equiv p^k(s_k+1) - p^{k+1}\frac{(1\pm1)}{2} \pmod{p^f-1}$.*

In each of these cases both $\mathrm{Ext}^1_{\mathbb{F}[\mathrm{GL}_2(k_v)]}(\overline{\sigma}_{\vec{s},\vec{t}}, \overline{\sigma}_{\vec{s}',\vec{t}'})$ and $\mathrm{Ext}^1_{\mathbb{F}[\mathrm{GL}_2(k_v)]}(\overline{\sigma}_{\vec{s}',\vec{t}'}, \overline{\sigma}_{\vec{s},\vec{t}})$ are one-dimensional.

Proof. This is Corollary 5.6(i) of [BP12]. □

In what follows it will help to recall the explicit construction of the extensions that are described by the preceding proposition. First note that if we set $r_i = s_i$, then the Serre weight $\overline{\sigma}_{\vec{s},\vec{t}}$ is a twist of $\bigotimes_{i=0}^{f-1}(V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i}$; we thus may, and do, restrict ourselves to considering extensions of the latter such representations.

As we have already recalled, Lemma 3.5 of [BP12] shows that $V_{2p-2-r_i}^{\mathrm{Fr}^i}$ sits in a short exact sequence

$$0 \to (V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i} \to V_{2p-2-r_i}^{\mathrm{Fr}^i} \to V_{p-r_i-2}^{\mathrm{Fr}^i} \otimes V_1^{\mathrm{Fr}^{i+1}} \to 0. \tag{5.2}$$

If we fix some $i \in \{0, \dots, f-1\}$, and tensor this exact sequence with the representations $(V_{r_{i'}} \otimes \det^{p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$ for $i' \neq i$, we obtain a short exact sequence[1]

$$0 \to \bigotimes_{i'=0}^{f-1}(V_{r_{i'}} \otimes \det^{p-1-r_{i'}})^{\mathrm{Fr}^{i'}} \to V_{2p-2-r_i}^{\mathrm{Fr}^i} \otimes \bigotimes_{i' \neq i}(V_{r_{i'}} \otimes \det^{p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$$
$$\to V_{p-r_i-2}^{\mathrm{Fr}^i} \otimes (V_{r_{i+1}} \otimes V_1 \otimes \det^{p-1-r_{i+1}})^{\mathrm{Fr}^{i+1}} \otimes \bigotimes_{i' \neq i, i+1}(V_{r_{i'}} \otimes \det^{p-1-r_{i'}})^{\mathrm{Fr}^{i'}} \to 0.$$

The middle term of this short exact sequence is contained in $\bigotimes_{i'=0}^{f-1} V_{2p-2-r_{i'}}^{\mathrm{Fr}^{i'}}$, and since the socle of this latter representation is equal to $\bigotimes_{i'=0}^{f-1}(V_{r_{i'}} \otimes \det^{p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$ (unless all of the $r_{i'}$ are equal to 0, but the extra term $\bigotimes_{i'=0}^{f-1} V_{p-1}$ which is then present does not affect this argument), we conclude that the socle of the middle term of this short exact sequence is equal to the left-hand term.

[1]Here and below we have to consider many tensor products indexed by the elements of $\mathcal{S}$. It will be convenient, and save introducing additional notation, to not always write these tensor products with the indices taken in the standard order. This should cause no confusion, as long as it is understood that for any permutation π of $\mathcal{S}$, and any collection of representations A_i indexed by the elements $i \in \mathcal{S}$, we always identify $\bigotimes_i A_i$ and $\bigotimes_i A_{\pi(i)}$ via the isomorphism $a_0 \otimes \cdots \otimes a_{f-1} \mapsto a_{\pi(0)} \otimes \cdots \otimes a_{\pi(f-1)}$; i.e. we rearrange the order of the factors in the tensor product according to the permutation π^{-1}.

If $r_{i+1} \leq p-2$ then [BP12, Lem. 3.8(i)] shows that $V_{r_{i+1}} \otimes V_1 \cong V_{r_{i+1}+1} \oplus (V_{r_{i+1}-1} \otimes \det)$, and thus this short exact sequence gives rise to the non-split extensions of $V_{p-r_i-2}^{\mathrm{Fr}^i} \otimes (V_{r_{i+1}\pm 1} \otimes \det^{\,p-\frac{1\pm 1}{2}-r_{i+1}})^{\mathrm{Fr}^{i+1}} \otimes \bigotimes_{i'\neq i,i+1} (V_{r_{i'}} \otimes \det^{\,p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$ by $\bigotimes_{i'=0}^{f-1} (V_{r_{i'}} \otimes \det^{\,p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$ whose existence is asserted by Proposition 5.1.3. If $r_{i+1} = p-1$, then [BP12, Lems. 3.4, 3.5, and 3.8(ii)] show that $V_{r_{i+1}} \otimes V_1 = V_{p-1} \otimes V_1$ contains $V_{p-2} \otimes \det = V_{r_{i+1}-1} \otimes \det$ as a subrepresentation, and so the preceding short exact sequence gives rise to the non-split extension of $V_{p-r_i-2}^{\mathrm{Fr}^i} \otimes (V_{r_{i+1}-1} \otimes \det^{\,p-r_{i+1}})^{\mathrm{Fr}^{i+1}} \otimes \bigotimes_{i'\neq i,i+1} (V_{r_{i'}} \otimes \det^{\,p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$ by $\bigotimes_{i'=0}^{f-1} (V_{r_{i'}} \otimes \det^{\,p-1-r_{i'}})^{\mathrm{Fr}^{i'}}$ whose existence is asserted by Proposition 5.1.3.

From these explicit constructions, we can infer the following lemma.

5.1.4. **Lemma.** *Suppose that $f > 1$, and that $\overline{\sigma}_{\vec{s},\vec{t}}$ and $\overline{\sigma}_{\vec{s}',\vec{t}'}$ are two Serre weights related as in the statement of Proposition* 5.1.3 (2)*; namely there is some k such that $s_j = s'_j$ if $j \neq k, k+1$, $s'_k = p-2-s_k$, $s'_{k+1} = s_{k+1} \pm 1$, and $\sum_{j=0}^{p-1} p^j (t'_j - t_j) \equiv p^k(s_k+1) - p^{k+1}\frac{(1\pm 1)}{2} \pmod{p^f - 1}$. Suppose moreover that $s_i = s'_i = 0$ for some particular $i \neq k, k+1$. If $0 \to \overline{\sigma}_{\vec{s},\vec{t}} \to V \to \sigma_{\vec{s}',\vec{t}'} \to 0$ is a non-split extension of $\mathrm{GL}_2(k_v)$-representations over $\mathbb{F}$, and if $0 \leq r \leq p-1$, then the tensor product $0 \to \overline{\sigma}_{\vec{s},\vec{t}} \otimes V_r^{\mathrm{Fr}^i} \to V \otimes V_r^{\mathrm{Fr}^i} \to \sigma_{\vec{s}',\vec{t}'} \otimes V_r^{\mathrm{Fr}^i} \to 0$ is again a non-split extension of such representations.*

Proof. Define $\vec{r}$ by $r_j = s_j$ if $j \neq i$, and $r_i = r$, and similarly define $\vec{r'}$ by $r'_j = s'_j$ if $j \neq i$, and $r'_i = r$. Then $\overline{\sigma}_{\vec{r},\vec{t}}$ and $\overline{\sigma}_{\vec{r}',\vec{t}'}$ again satisfy condition (2) of Proposition 5.1.3, and so, by that proposition, there is a non-trivial extension of $\overline{\sigma}_{\vec{r}',\vec{t}'}$ by $\overline{\sigma}_{\vec{r},\vec{t}}$. The preceding discussion shows, though, that this extension is in fact obtained exactly by tensoring the extension of $\overline{\sigma}_{\vec{s}',\vec{t}'}$ by $\overline{\sigma}_{\vec{s},\vec{t}}$ by $V_r^{\mathrm{Fr}^i}$. This proves the lemma. $\square$

We need a further technical lemma related to extensions. Before stating it, we introduce a definition. If $0 \leq r \leq p-1$ and $0 \leq n < f-1$, then we define

$$V_r(n) := V_r \otimes V_1^{\mathrm{Fr}} \otimes \bigotimes_{i=1}^{n} V_{p-1}^{\mathrm{Fr}^i}.$$

5.1.5. **Lemma.** *The lattice of subrepresentations of $V_r(n)$ is totally ordered, and thus $V_r(n)$ admits a unique Jordan–Hölder filtration, which coincides with both its socle and cosocle filtration. The length of this filtration is $2n+1$, and the Jordan–Hölder factors of this filtration, taken in order are*

$$\begin{gathered}
V_r \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}} \otimes V_{p-1}^{\mathrm{Fr}^2} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^n}, \\
V_r \otimes V_0^{\mathrm{Fr}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^2} \otimes V_{p-1}^{\mathrm{Fr}^3} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^n}, \\
\cdots, \\
V_r \otimes V_0^{\mathrm{Fr}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^n}, \\
V_r \otimes V_0^{\mathrm{Fr}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^n} \otimes V_1^{\mathrm{Fr}^{n+1}}, \\
V_r \otimes V_0^{\mathrm{Fr}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^n}, \\
\cdots, \\
V_r \otimes V_0^{\mathrm{Fr}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^2} \otimes V_{p-1}^{\mathrm{Fr}^3} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^n},
\end{gathered}$$

$$V_r \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}} \otimes V_{p-1}^{\mathrm{Fr}^2} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^n}.$$

5.1.6. *Remark.* Of course we could omit the V_0 factors in the expressions, as V_0 is the trivial representation, but we have left them in place for clarity. Note also that the ith and $(2n-i)$th layers of the filtration are isomorphic.

Proof of Lemma 5.1.5. We proceed by induction on n, the case $n = 0$ being trivial. Suppose now that the result is proved for some value of n; writing $V_r(n+1) = V_r(n) \otimes V_{p-1}^{\mathrm{Fr}^{n+1}}$, we will deduce the claim of the lemma for $n+1$.

To this end, consider one of the extensions

$$0 \to \operatorname{soc}_i V_r(n)/\operatorname{soc}_{i-1} V_r(n) \to \operatorname{soc}_{i+1} V_r(n)/\operatorname{soc}_{i-1} V_r(n) \\ \to \operatorname{soc}_{i+1} V_r(n)/\operatorname{soc}_i V_r(n) \to 0 \tag{5.3}$$

arising from the socle filtration of $V_r(n)$, where we first suppose that either $i < n-1$ or $i > n$. Tensoring with $V_{p-1}^{\mathrm{Fr}^{n+1}}$, it follows from Lemma 5.1.4 that the resulting short exact sequence is again non-split.

If now $i = n-1$, then (applying our inductive hypothesis) the short exact sequence (5.3) becomes

$$\begin{aligned} 0 \to V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^n} \\ \to \operatorname{soc}_n V_r(n)/\operatorname{soc}_{n-2} V_r(n) \\ \to V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes V_0^{\mathrm{Fr}^n} \otimes V_1^{\mathrm{Fr}^{n+1}} \to 0, \end{aligned}$$

and tensoring with $V_{p-1}^{\mathrm{Fr}^{n+1}}$ yields the short exact sequence

$$\begin{aligned} 0 \to V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^n} \otimes V_{p-1}^{\mathrm{Fr}^{n+1}} \\ \to \big(\operatorname{soc}_n V_r(n)/\operatorname{soc}_{n-2} V_r(n)\big) \otimes V_{p-1}^{\mathrm{Fr}^{n+1}} \\ \to V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes V_0^{\mathrm{Fr}^n} \otimes (V_1 \otimes V_{p-1})^{\mathrm{Fr}^{n+1}} \to 0. \end{aligned}$$

Now, as a particular case of the above explicit description of extensions, we know that the socle of $V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes V_0^{\mathrm{Fr}^n} \otimes (V_1 \otimes V_{p-1})^{\mathrm{Fr}^{n+1}}$ is equal to $V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes V_0^{\mathrm{Fr}^n} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{n+1}}$, and that the extension of $V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes V_0^{\mathrm{Fr}^n} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{n+1}}$ by $V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^n} \otimes V_{p-1}^{\mathrm{Fr}^{n+1}}$ induced by the preceding short exact sequence is non-split.

Putting together everything shown so far, and taking into account the inductive hypothesis, we conclude that $\operatorname{soc}_i V_r(n+1)$ has the structure claimed in the statement of the lemma for $i \leq n$. A similar analysis applies if we tensor (5.3) with $V_{p-1}^{\mathrm{Fr}^{n+1}}$ in the case when $i = n$, and so we likewise conclude that $\operatorname{cosoc}_i V_r(n+1)$ has the claimed structure when $i \leq n$.

To complete the proof of the lemma, it suffices to show that the socle filtration of $\big(\operatorname{soc}_n V_r(n)/\operatorname{soc}_{n-1} V_r(n)\big) \otimes V_{p-1}^{\mathrm{Fr}^{n+1}}$ has length three, with Jordan–Hölder consituents (in order) equal to $V_r \otimes V_0^{\mathrm{Fr}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^n} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{n+1}}$, $V_r \otimes V_0^{\mathrm{Fr}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^{n+1}} \otimes V_1^{\mathrm{Fr}^{n+2}}$, and $V_r \otimes V_0^{\mathrm{Fr}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^n} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{n+1}}$. Now (again using our inductive hypothesis) we have that

$$\big(\operatorname{soc}_n V_r(n)/\operatorname{soc}_{n-1} V_r(n)\big) \otimes V_{p-1}^{\mathrm{Fr}^{n+1}}$$

$$= V_r \otimes V_0^{\mathrm{Fr}} \otimes \ldots \otimes V_0^{\mathrm{Fr}^{n-1}} \otimes V_0^{\mathrm{Fr}^n} \otimes (V_1 \otimes V_{p-1})^{\mathrm{Fr}^{n+1}}.$$

From [BP12, Lem. 3.8 (ii)] we see that $V_1 \otimes V_{p-1}$ is equal to the representation denoted there by R_{p-2}. From [BP12, Lems. 3.4, 3.5] and the surrounding discussion we see that R_{p-2} has socle equal to $V_{p-2} \otimes \det$, and that the quotient of R_{p-2} by its socle is equal to an extension of $V_{p-2} \otimes \det$ by V_1^{Fr}. Since V_{p-2} does not admit a non-trivial extension of itself as a $\mathrm{GL}_2(k_v)$-representation (by Proposition 5.1.3), that extension of $V_{p-2} \otimes \det$ by V_1^{Fr} must be non-split, and so in fact R_{p-2} has a socle filtration of length three, with constituents (in order) equal to $V_{p-2} \otimes \det$, V_1^{Fr}, and $V_{p-2} \otimes \det$. Our desired conclusion now follows from this, together with Lemma 5.1.4. □

Next, we introduce some additional notation. First, for any $i \in \mathcal{S}$, we let $l(i) := \max\{j \in \mathcal{S}' \mid j \leq i\}$. (Here the ordering on $\mathcal{S}$ is understood cyclically, i.e. $f-1 < 0$, and recall that $\mathcal{S}' \neq \varnothing$.) If $J \subseteq \mathcal{S}'$, then we write $\widetilde{J} := \{i \in \mathcal{S} \mid l(i) \in J\}$.

If $i \in \mathcal{S}'$, then define

$$X(i) := V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes V_1^{\mathrm{Fr}^{i+1}} \otimes \bigotimes_{j \in \mathcal{E} \text{ s.t. } l(j)=i} V_{p-1}^{\mathrm{Fr}^j}.$$

In terms of this definition (and taking into account the preceding description of the cokernel of (5.1)), we see that for any $J \subseteq \mathcal{S}'$, there is an isomorphism

$$W_J = \bigotimes_{i \in J} X(i) \otimes \bigotimes_{i \notin \widetilde{J}} (V_{r_i} \otimes \det{}^{p-1-r_i})^{\mathrm{Fr}^i}. \tag{5.4}$$

5.1.7. **Lemma.** *For any $i \in \mathcal{S}'$ the lattice of subrepresentations of $X(i)$ is totally ordered, and thus $X(i)$ admits a unique Jordan–Hölder filtration, which coincides with both its socle and cosocle filtration. If we let $i+n$ denote the maximal element of $\mathcal{S}$ for which $l(i+n) = i$, then the length of this filtration is $2n+1$, and the Jordan–Hölder factors of this filtration, taken in order are*

$$\begin{gathered}
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{i+1}} \otimes V_{p-1}^{\mathrm{Fr}^{i+2}} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^{i+n}}, \\
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes V_0^{\mathrm{Fr}^{i+1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{i+2}} \otimes V_{p-1}^{\mathrm{Fr}^{i+3}} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^{i+n}}, \\
\cdots, \\
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes V_0^{\mathrm{Fr}^{i+1}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^{i+n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{i+n}}, \\
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes V_0^{\mathrm{Fr}^{i+1}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^{i+n}} \otimes V_1^{\mathrm{Fr}^{i+n+1}}, \\
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes V_0^{\mathrm{Fr}^{i+1}} \otimes \cdots \otimes V_0^{\mathrm{Fr}^{i+n-1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{i+n}}, \\
\cdots, \\
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes V_0^{\mathrm{Fr}^{i+1}} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{i+2}} \otimes V_{p-1}^{\mathrm{Fr}^{i+3}} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^{i+n}}, \\
V_{p-2-r_i}^{\mathrm{Fr}^i} \otimes (V_{p-2} \otimes \det)^{\mathrm{Fr}^{i+1}} \otimes V_{p-1}^{\mathrm{Fr}^{i+2}} \otimes \cdots \otimes V_{p-1}^{\mathrm{Fr}^{i+n}}.
\end{gathered}$$

Proof. This follows immediately from Lemma 5.1.5. □

We let $\mathcal{P}$ denote the collection of subsets $J \subseteq \mathcal{S}$ for which $i \in J \cap \mathcal{E}$ implies that $i-1 \in J$. The following proposition constructs, for each element $J \in \mathcal{P}$, a subrepresentation F_J of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$. This construction refines the correspondence $J \mapsto F'_J$ (for $J \subseteq \mathcal{S}'$) introduced above.

5.1.8. **Proposition.** *There is a unique way to define an order-preserving map $J \mapsto F_J$ from $\mathcal{P}$ to the lattice of subrepresentations of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ such that:*

(1) *$F_J \subset F'_{J\cap\mathcal{S}'}$.*
(2) *The inclusion*
$$\sum_{\substack{J'\in\mathcal{P},\ J'\subsetneq J \text{ s.t.}\\ J'\cap\mathcal{S}'\subsetneq J\cap\mathcal{S}'}} F_{J'} \subset F_J \cap F'_{<J\cap\mathcal{S}'}$$
(which holds by (1) together with the fact that $J' \mapsto F_{J'}$ is order preserving) is an equality.
(3) *If for each $i \in J \cap \mathcal{S}'$ we let $n_i(J)$ denote the largest integer for which $\{i+1,\ldots,i+n_i(J)\} \subseteq J \cap \mathcal{E}$, then the image of $F_J/(F_J \cap F'_{<J\cap\mathcal{S}'}) = (F_J + F'_{<J\cap\mathcal{S}'})/F'_{<J\cap\mathcal{S}'}$ in $W_J = F'_{J\cap\mathcal{S}'}/F'_{<J\cap\mathcal{S}'}$ is identified under the isomorphism (5.4) with the tensor product*
$$\bigotimes_{i\in J\cap\mathcal{S}'} \operatorname{soc}_{n_i(J)} X(i) \otimes \bigotimes_{i\notin\widetilde{J}} (V_{r_i}\otimes\det{}^{p-1-r_i})^{\mathrm{Fr}^i}.$$
(Here $\operatorname{soc}_{n_i(J)} X(i)$ denotes the $n_i(J)$th stage of the socle filtration on $X(i)$.)

Proof. The given conditions allow us to recursively define F_J, beginning with $F_\varnothing = \bigotimes_{i=0}^{f-1} V_{r_i}^{\mathrm{Fr}^i}$. It is obvious that conditions (1)-(3) determine $F_J \subset F'_{J\cap\mathcal{S}'}$ uniquely. ∎

Note that $\mathcal{S}$ is the (unique) maximal element of $\mathcal{P}$.

5.1.9. **Lemma.** *If V is a subrepresentation of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ with socle equal to $\bigotimes_{i=0}^{f-1}(V_{r_i}\otimes\det^{p-1-r_i})^{\mathrm{Fr}^i}$, then the multiset $\mathrm{JH}(V)$ of Jordan–Hölder factors of V satisfies multiplicity one if and only if $V \subseteq F_{\mathcal{S}}$. Furthermore, any $\mathbb{F}$-representation W of $\mathrm{GL}_2(k_v)$ with socle $\bigotimes_{i=0}^{f-1} V_{r_i}^{\mathrm{Fr}^i}$ which satisfies multiplicity one is isomorphic to some such V.*

5.1.10. *Remark.* Provided that some $r_i \neq 0$, every subrepresentation of V has socle equal to $\bigotimes_{i=0}^{f-1}(V_{r_i}\otimes\det^{p-1-r_i})^{\mathrm{Fr}^i}$.

Proof of Lemma 5.1.9. Write $n_i := n_i(\mathcal{S})$. An examination of the statement of Lemma 5.1.7 shows that if $j > n_i$ then the multiset of Jordan–Hölder constituents of $\operatorname{soc}_j X(i)$ does *not* satisfy multiplicity one. From this a straightforward induction shows that if $\mathrm{JH}(V)$ does satisfy multiplicity one, then $F'_J \cap V = F_{\widetilde{J}} \cap V$ for each $J \subset \mathcal{S}'$. Taking $J = \mathcal{S}'$, so that $\widetilde{J} = \mathcal{S}$, we conclude that $V \subseteq F_{\mathcal{S}}$, as claimed.

To show that $F_{\mathcal{S}}$ satisfies multiplicity one, set $F_{<J} := \sum_{J'\subsetneq J} F_{J'}$. Then by Proposition 5.1.8(3), we see that
$$F_J/F_{<J} \cong \bigotimes_{i\in J\cap\mathcal{S}'} \left(\operatorname{soc}_{n_i(J)} X(i)/\operatorname{soc}_{n_i(J)-1} X(i)\right) \otimes \bigotimes_{i\notin\widetilde{J}} V_{r_i}^{\mathrm{Fr}^i}.$$

Suppose that $\overline{\sigma}_\lambda$ is a Jordan–Hölder factor of $F_J/F_{<J}$. From the explicit description of the socle filtration on $X(i)$, it is easy to check (using [BP12, Lem. 3.8(i)] whenever $n_i(J) = n_i(\mathcal{S})$) that we have $l(\overline{\sigma}_\lambda) = |J|$, and it is also clear that λ determines J. Therefore, in order to check that $F_{\mathcal{S}}$ satisfies multiplicity one, it suffices to show that each $F_J/F_{<J}$ satisfies multiplicity one, which is immediate from its explicit description.

Finally, note that any W with socle $\bigotimes_{i=0}^{f-1}(V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i}$ which satisfies multiplicity one is isomorphic to a subrepresentation of $\bigotimes_{i=0}^{f-1} V_{2p-2-r_i}^{\mathrm{Fr}^i}$ by Proposition 3.6 of [BP12]. $\square$

5.1.11. **Proposition.** *The socle filtration of $F_{\mathcal{S}}$ is given by*

$$\operatorname{soc}_n F_{\mathcal{S}} = \sum_{\substack{J \in \mathcal{P} \\ |J|=n}} F_J.$$

Furthermore, all extensions between constituents of successive layers of the socle filtration which can be non-zero according to Proposition 5.1.3*, are* non-zero.

Proof. Set $\operatorname{soc}'_n F_{\mathcal{S}} := \sum_{\substack{J\in\mathcal{P} \\ |J|=n}} F_J$. As in the proof of Lemma 5.1.9, any Jordan–Hölder factor $\overline{\sigma}$ of $\operatorname{soc}'_n F_{\mathcal{S}}/\operatorname{soc}'_{n-1} F_{\mathcal{S}}$ satisfies $l(\overline{\sigma}) = n$. It follows immediately from Proposition 5.1.3 that if $\overline{\sigma}'$ is another Jordan–Hölder factor of $F_{\mathcal{S}}$ which admits an extension with $\overline{\sigma}$, then $l(\overline{\sigma}') = l(\overline{\sigma}) \pm 1$, so in particular we see that $\operatorname{soc}'_n F_{\mathcal{S}}/\operatorname{soc}'_{n-1} F_{\mathcal{S}}$ is semisimple.

On the other hand, by Lemma 5.1.4 (together with the discussion preceding it) and the very definition of the F_J (concretely, the description of $F_J/F_{<J}$ given in the proof of Lemma 5.1.9), we see that each irreducible subrepresentation of the quotient $\operatorname{soc}'_n F_{\mathcal{S}}/\operatorname{soc}'_{n-1} F_{\mathcal{S}}$ extends some irreducible subrepresentation of the quotient $\operatorname{soc}'_{n-1} F_{\mathcal{S}}/\operatorname{soc}'_{n-2} F_{\mathcal{S}}$, and indeed that any possibly non-zero extension between constituents of these layers is in fact non-zero. Thus $\operatorname{soc}'_n F_{\mathcal{S}} = \operatorname{soc}_n F_{\mathcal{S}}$, and the result follows. $\square$

Proof of Theorem 5.1.1. The result for the cosocle filtrations of the $\overline{\sigma}^{\circ}_J(\tau)$ follows by duality from that for the socle filtrations of the $\overline{\sigma}^{\circ,J}(\tau)$. In this case, after twisting we may suppose that for some $0 \le r_i \le p-1$ not all equal to $p-1$, we have $\overline{\sigma}_J(\tau) \cong \bigotimes_{i=0}^{f-1}(V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i}$. By Lemma 3.1.1, $\overline{\sigma}^{\circ,J}(\tau)$ satisfies multiplicity one, and it has socle $\overline{\sigma}_J(\tau) \cong \bigotimes_{i=0}^{f-1}(V_{r_i} \otimes \det^{p-1-r_i})^{\mathrm{Fr}^i}$ by definition, so we see from Lemma 5.1.9 that it is isomorphic to a subrepresentation of $F_{\mathcal{S}}$.

The result is then immediate from Proposition 5.1.11 upon noting that for each Jordan–Hölder factor $\overline{\sigma}_{J'}(\tau)$ we have $l(\overline{\sigma}_{J'}(\tau)) = |J \triangle J'|$, and that (as already noted in the proof of Proposition 5.1.11), the irreducible constituents of the ith layer of the socle filtration of $F_{\mathcal{S}}$ are precisely those $\overline{\sigma}$ with $l(\overline{\sigma}) = i$. $\square$

5.2. **Specific lattices and their gauges.** Maintain the notation and assumptions of the previous sections, so that τ is a non-scalar tame inertial type, and for each $J \in \mathcal{P}_\tau$, we have $\mathrm{GL}_2(\mathcal{O}_{F_v})$-invariant $\mathcal{O}$-lattices $\sigma^{\circ,J}(\tau)$ and $\sigma^{\circ}_J(\tau)$ in $\sigma(\tau)$, which are uniquely determined up to scaling. In this section, we choose $E = \mathcal{O}[\frac{1}{p}]$ to be unramified over $\mathbb{Q}_p$; this is possible by Lemma 3.1.1. If τ is cuspidal, assume in addition that τ is regular. The arguments of this section are an abstraction and generalisation of the proof of Théorème 2.4 of [Bre12].

We can and do fix some $J_{\mathrm{base}} \in \mathcal{P}_\tau$ with $J^c_{\mathrm{base}} \in \mathcal{P}_\tau$; if τ is a principal series type, then we take $J_{\mathrm{base}} = \varnothing$, and if τ is cuspidal then by Lemma 3.3.1 we can and do set $J_{\mathrm{base}} = \{i, \dots, f-1\}$ for some i. We then have an involution $\imath$ on the subsets of $\mathcal{S}$ given by $\imath(J) := J \triangle J_{\mathrm{base}}$, so that $\imath(J_{\mathrm{base}}) = \varnothing$ and $\imath(J^c_{\mathrm{base}}) = \mathcal{S}$. Note that if τ is a principal series type, then $\imath(J) = J$ for all J. We will need the following simple combinatorial lemma.

5.2.1. **Lemma.** *For any $J \subseteq J' \subseteq \mathcal{S}$ such that $\imath(J), \imath(J') \in \mathcal{P}_\tau$, we can find a chain $J = J_0 \subsetneq J_1 \subsetneq \cdots \subsetneq J_{|J'|-|J|} = J'$ such that $|J_i| = i + |J|$ and $\imath(J_i) \in \mathcal{P}_\tau$ for each i.*

Proof. It suffices to consider the case $J' = \mathcal{S}$. (Observe that $\imath(J_i \cap J') \in \mathcal{P}_\tau$, and so replacing each J_i with $J_i \cap J'$ and discarding duplicates will give the desired chain.) The principal series case is part of [Bre12, Lem. 2.1(iii)], so we need only consider the (regular) cuspidal case. We must show that if $J \neq \mathcal{S}$ with $\imath(J) \in \mathcal{P}_\tau$, then there is some $j \notin J$ such that $\imath(J \cup \{j\}) \in \mathcal{P}_\tau$. We will prove this directly from the definition of $\mathcal{P}_\tau$. We use the notation of Section 3.3.

Continue to assume that τ is a regular cuspidal type, so that as above we have $J_{\text{base}} = \{i, \dots, f-1\}$ for some i with $0 < c_i < p-1$. For the sake of contradiction suppose that we have some $J \neq \mathcal{S}$ with $\imath(J) \in \mathcal{P}_\tau$ and for all $j \notin J$, $\imath(J \cup \{j\}) \notin \mathcal{P}_\tau$. Then we see that for each $j \notin J$ with $0 \le j \le i-1$, either

- $j - 1 \notin \imath(J)_0$ and $c_j = p-1$, or
- $j + 1 \notin \imath(J)$ and $c_{j+1} = 0$,

and for each $j \notin J$ with $i \le j \le f-1$, either

- $j - 1 \in \imath(J)_0$ and $c_j = 0$, or
- $j \neq f-1$, $j+1 \notin J$ and $c_{j+1} = p-1$, or
- $j = f-1$, $0 \notin J$, and $c_0 = 0$.

Note that in the final bullet point we have eliminated the possibility $j = f-1$, $i = 0$, $0 \in J$, and $c_0 = 0$ because this contradicts $c_i \neq 0$.

Suppose that $i \notin J$. Then since $c_i \neq 0$, we see that if $i < f-1$ then $c_{i+1} = p-1$ and $i+1 \notin J$. Continuing in this fashion, we see that $i+1, \dots, f-1 \notin J$, and that $c_{i+1} = \cdots = c_{f-1} = p-1$. Applying the same analysis again, we see that $0 \notin J$ and $c_0 = 0$. If $i = 0$ this is a contradiction; otherwise, continuing in the same way, we deduce that $c_1 = 0$ and $1 \notin \imath(J)$, and eventually we conclude that $c_i = 0$, a contradiction.

Suppose now that $i \in J$. Suppose that there is some $0 \le j \le i-1$ with $j \notin J$, and without loss of generality let j be maximal with this property. Then we see that $c_j = p-1$ and $j-1 \notin \imath(J)_0$ (using $c_i \neq 0$ in case $j = i-1$), so continuing in this fashion, we have in particular $0 \notin J$, $c_0 = p-1$, and $f-1 \notin J$. Then $c_{f-1} = 0$, $i \neq f-1$, and $f-2 \in \imath(J)_0$; so $f-2 \notin J$, and then $f-3 \notin J, \dots, i \notin J$, a contradiction. So we see that $j \in J$ for all $0 \le j \le i$, and thus we must have $j \notin J$ for some $i+1 \le j \le f-1$. Let j be maximal with this property. Then we see that $c_j = 0$ and $j-1 \in \imath(J)$ (recalling in case $j = f-1$ that we have $0 \in J$), so $j-1 \notin J$, and then $c_{j-1} = 0$, and continuing in this fashion we obtain $j-2 \notin J, \dots, i \notin J$, a contradiction. $\square$

If σ° is any $\mathrm{GL}_2(\mathcal{O}_{F_v})$-invariant $\mathcal{O}$-lattice in $\sigma(\tau)$ and $\imath(J) \in \mathcal{P}_\tau$, then we let $\varepsilon_{\imath(J)}(\sigma^\circ)$ denote the least integer such that $p^{\varepsilon_{\imath(J)}(\sigma^\circ)}\sigma^\circ_{\imath(J)}(\tau) \subseteq \sigma^\circ$. The tuple $(p^{\varepsilon_{\imath(J)}(\sigma^\circ)})_{\imath(J)\in\mathcal{P}_\tau}$ is the gauge of σ°. It is easy to see that the following fundamental property holds.

5.2.2. *Property.* For any lattice σ°, and for any J, J' with $\imath(J), \imath(J') \in \mathcal{P}_\tau$, we have

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) \ge \varepsilon_{\imath(J')}(\sigma^\circ) - \varepsilon_{\imath(J)}(\sigma^\circ),$$

with equality if and only if $\overline{\sigma}^\circ$ contains a subquotient with socle $\overline{\sigma}_{\imath(J')}(\tau)$ and cosocle $\overline{\sigma}_{\imath(J)}(\tau)$.

Note that if we take σ° to be $\sigma^{\circ,\imath(J')}(\tau)$, then for any $\overline{\sigma}_{\imath(J)}(\tau)$, the quotient $\overline{\sigma}^{\circ,\imath(J')}(\tau)$ contains a subobject with socle $\overline{\sigma}_{\imath(J')}(\tau)$ and cosocle $\overline{\sigma}_{\imath(J)}(\tau)$, and so in particular we deduce that

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) = \varepsilon_{\imath(J')}(\sigma^{\circ,\imath(J')}(\tau)) - \varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)). \tag{5.5}$$

We fix $\sigma^{\circ,\imath(\varnothing)}(\tau)$ once and for all. We then normalize the $\sigma^\circ_{\imath(J)}(\tau)$ by requiring that $\varepsilon_{\imath(J)}(\sigma^{\circ,\imath(\varnothing)}(\tau)) = 0$, and we normalize the $\sigma^{\circ,\imath(J)}(\tau)$ by requiring that $\varepsilon_{\imath(\varnothing)}(\sigma^{\circ,\imath(J)}(\tau)) = 0$. Note that, with these normalizations, (5.5) implies that $\varepsilon_{\imath(\varnothing)}(\sigma^\circ_{\imath(J)}(\tau)) = 0$, and Property 5.2.2 then implies that

$$0 = \varepsilon_{\imath(\varnothing)}(\sigma^\circ_{\imath(J)}(\tau)) \geq \varepsilon_{\imath(\varnothing)}(\sigma^{\circ,\imath(J')}(\tau)) - \varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)) = 0 - \varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)),$$

i.e. that

$$\varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)) \geq 0 \quad \text{for all } J, J' \text{ with } \imath(J), \imath(J') \in \mathcal{P}_\tau. \tag{5.6}$$

Our key input is the following.

5.2.3. **Proposition.** (1) *We have* $p^f\sigma^{\circ,\imath(\varnothing)}(\tau) \subseteq \sigma^\circ_{\imath(\varnothing)}(\tau) \subseteq \sigma^{\circ,\imath(\varnothing)}(\tau)$, *and* $\sigma^\circ_{\imath(\mathcal{S})}(\tau) = \sigma^{\circ,\imath(\varnothing)}(\tau)$.

(2) *The sets J with $\imath(J) \in \mathcal{P}_\tau$ such that $\overline{\sigma}_{\imath(J)}(\tau)$ appears in the ith layer of the socle filtration of $\overline{\sigma}^\circ_{\imath(\varnothing)}(\tau)$ are precisely those J with $|J| = f - i$.*

(3) *$\overline{\sigma}^{\circ,\imath(\varnothing)}(\tau)$ contains a subquotient with socle $\overline{\sigma}_{\imath(J')}(\tau)$ and cosocle $\overline{\sigma}_{\imath(J)}(\tau)$ if and only if $J' \subseteq J$, and $\overline{\sigma}^\circ_{\imath(\varnothing)}(\tau)$ contains a subquotient with socle $\overline{\sigma}_{\imath(J')}(\tau)$ and cosocle $\overline{\sigma}_{\imath(J)}(\tau)$ if and only if $J \subseteq J'$.*

Proof. (1) By Theorem 5.1.1, $\overline{\sigma}^{\circ,\imath(\varnothing)}(\tau)$ has cosocle $\overline{\sigma}_{\imath(\mathcal{S})}(\tau)$, so the assumption that $\varepsilon_{\imath(\mathcal{S})}(\sigma^{\circ,\imath(\varnothing)}(\tau)) = 0$ implies that $\sigma^\circ_{\imath(\mathcal{S})}(\tau) = \sigma^{\circ,\imath(\varnothing)}(\tau)$.

By Lemma 3.1.1, there is a character χ such that $\sigma(\tau)^\vee \otimes \chi \cong \sigma(\tau)$, and such that each $\overline{\sigma}_{\imath(J)}(\tau)$ satisfies $\overline{\sigma}_{\imath(J)}(\tau)^\vee \otimes \overline{\chi} \cong \overline{\sigma}_{\imath(J)}(\tau)$. There is thus an isomorphism $(\sigma^{\circ,\imath(\varnothing)}(\tau))^\vee \otimes \chi \cong \sigma^\circ_{\imath(\varnothing)}(\tau)$, and the result then follows from Proposition 4.2.1 and the assumption that $\varepsilon_{\imath(\varnothing)}(\sigma^{\circ,\imath(\varnothing)}(\tau)) = 0$.

(2) This is immediate from (1) and Theorem 5.1.1.

(3) Suppose $J' \subseteq J$. An easy induction on $|J \setminus J'|$ shows that any subrepresentation of $\overline{\sigma}^{\circ,\imath(\varnothing)}(\tau)$ with $\overline{\sigma}_{\imath(J)}(\tau)$ as a Jordan–Hölder factor also has $\overline{\sigma}_{\imath(J')}(\tau)$ as a Jordan–Hölder factor. (In the case $|J \setminus J'| = 1$ consider the quotient of $\overline{\sigma}^{\circ,\imath(\varnothing)}(\tau)$ by the maximal subrepresentation containing neither $\overline{\sigma}_{\imath(J)}(\tau)$ nor $\overline{\sigma}_{\imath(J')}(\tau)$ as a Jordan–Hölder factor, which must have socle $\overline{\sigma}_{\imath(J')}(\tau)$ by the paragraph following Theorem 5.1.1. For the induction step use Lemma 5.2.1.) The claim for $\overline{\sigma}^{\circ,\imath(\varnothing)}(\tau)$ follows easily from this, and the result for $\overline{\sigma}^\circ_{\imath(\varnothing)}(\tau)$ is similar. $\square$

From Property 5.2.2 and Proposition 5.2.3(3), we see that

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) = 0 \quad \text{if} \quad J' \subseteq J, \tag{5.7}$$

that

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) > 0 \quad \text{if} \quad J' \not\subseteq J, \tag{5.8}$$

and that

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) = \varepsilon_{\imath(J')}(\sigma^\circ_{\imath(\varnothing)}(\tau)) - \varepsilon_{\imath(J)}(\sigma^\circ_{\imath(\varnothing)}(\tau)) \quad \text{if} \quad J \subseteq J'. \tag{5.9}$$

(There is a fourth relation in the case $J \not\subseteq J'$, but we will not need this.)

Now combining relations (5.8) and (5.9), we see that if we partially order the subsets of $\imath(\mathcal{P}_\tau)$ by inclusion, then $\varepsilon_{\imath(J)}(\sigma^\circ_{\imath(\varnothing)}(\tau))$ is a strictly increasing function of J. By Proposition 5.2.3(1) we have $\varepsilon_{\imath(\mathcal{S})}(\sigma^\circ_{\imath(\varnothing)}(\tau)) \leq f$, so we see that $0 \leq \varepsilon_{\imath(J)}(\sigma^\circ_{\imath(\varnothing)}(\tau)) \leq f$ for each $J \in \imath(\mathcal{P}_\tau)$. By Lemma 5.2.1 (applied twice, to the pairs $\varnothing, J$ and $J, \mathcal{S}$), each $J \in \imath(\mathcal{P}_\tau)$ lies in a chain of elements of $\imath(\mathcal{P}_\tau)$ of length $f+1$, so we conclude that in fact

$$\varepsilon_{\imath(J)}(\sigma^\circ_{\imath(\varnothing)}(\tau)) = |J|,$$

and that

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) = |J' \setminus J| \quad \text{if} \quad J \subseteq J'. \tag{5.10}$$

As an application of this, note that if we take $J = \imath(\varnothing)$ in relation (5.5), we find that

$$\varepsilon_{\imath(J')}(\sigma^{\circ,\imath(J')}(\tau)) = |J'|. \tag{5.11}$$

As another application, combining (5.7) and (5.10) with Property 5.2.2, we find that for any $J, J' \in \imath(\mathcal{P}_\tau)$ we have

$$\begin{aligned} |J' \setminus J| &= \varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J\cap J')}(\tau)) \\ &\geq \varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) - \varepsilon_{J\cap J'}(\sigma^\circ_{\imath(J)}(\tau)) = \varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)). \end{aligned} \tag{5.12}$$

Fix some $J, J' \in \imath(\mathcal{P}_\tau)$, and consider a chain $J' = J_0, J_1, \ldots, J_n = J$ joining $\overline{\sigma}_{\imath(J')}(\tau)$ to $\overline{\sigma}_{\imath(J)}(\tau)$ in the socle filtration of $\overline{\sigma}^{\circ,\imath(J')}(\tau)$ (so that $\overline{\sigma}_{\imath(J_i)}(\tau)$ is in the i-th layer of the socle filtration of $\overline{\sigma}^{\circ,\imath(J')}(\tau)$, and is non-trivially extended by $\overline{\sigma}_{\imath(J_{i-1})}(\tau)$ in the socle filtration). By Theorem 5.1.1, J_{i+1} is obtained from J_i by either adding or removing an element. Property 5.2.2 gives that

$$\varepsilon_{\imath(J_i)}(\sigma^\circ_{\imath(J_{i+1})}(\tau)) = \varepsilon_{\imath(J_i)}(\sigma^{\circ,\imath(J')}(\tau)) - \varepsilon_{\imath(J_{i+1})}(\sigma^{\circ,\imath(J')}(\tau)).$$

From (5.7) we conclude that

$$\varepsilon_{\imath(J_i)}(\sigma^{\circ,\imath(J')}(\tau)) = \varepsilon_{\imath(J_{i+1})}(\sigma^{\circ,\imath(J')}(\tau)) \quad \text{if } J_i \subseteq J_{i+1},$$

while from (5.10) we conclude that

$$\varepsilon_{\imath(J_i)}(\sigma^{\circ,\imath(J')}(\tau)) = \varepsilon_{\imath(J_{i+1})}(\sigma^{\circ,\imath(J')}(\tau)) + 1 \quad \text{if } J_{i+1} \subseteq J_i.$$

Passing along the whole chain from J' to J, we find that

$$\varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)) \leq \varepsilon_{\imath(J')}(\sigma^{\circ,\imath(J')}(\tau)) - |J' \setminus J|. \tag{5.13}$$

Combining this with Property 5.2.2 again gives

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) = \varepsilon_{\imath(J')}(\sigma^{\circ,\imath(J')}(\tau)) - \varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)) \geq |J' \setminus J|. \tag{5.14}$$

Combining (5.12) and (5.14), we find that

$$\varepsilon_{\imath(J')}(\sigma^\circ_{\imath(J)}(\tau)) = |J' \setminus J|.$$

Since we have equality in (5.14), we must also have equality in (5.13), and combining this with (5.11), we find that

$$\varepsilon_{\imath(J)}(\sigma^{\circ,\imath(J')}(\tau)) = |J \cap J'|. \tag{5.15}$$

5.2.4. **Theorem.** *(1) We have* $\sigma^{\circ,\imath(J)}(\tau) = \sum_{J'} p^{|J\cap J'|}\sigma^{\circ}_{\imath(J')}(\tau)$.

(2) Suppose that $\imath(J) \in \mathcal{P}_\tau$*, that* $j \notin J$*, and that* $\imath(J \cup \{j\}) \in \mathcal{P}_\tau$*. Then we have inclusions*

$$p\sigma^{\circ,\imath(J)}(\tau) \subseteq \sigma^{\circ,\imath(J\cup\{j\})}(\tau) \subseteq \sigma^{\circ,\imath(J)}(\tau).$$

Furthermore the Serre weights occuring as Jordan–Hölder factors of the cokernel of the inclusion $\sigma^{\circ,\imath(J\cup\{j\})}(\tau) \subseteq \sigma^{\circ,\imath(J)}(\tau)$ *are the* $\overline{\sigma}_{\imath(J')}(\tau)$ *with* $j \in J'$*, and the Serre weights occuring in the cokernel of the inclusion* $p\sigma^{\circ,\imath(J)}(\tau) \subseteq \sigma^{\circ,\imath(J\cup\{j\})}(\tau)$ *are the* $\overline{\sigma}_{\imath(J')}(\tau)$ *with* $j \notin J'$.

(3) We have $\sigma^{\circ}_{\imath(J)}(\tau) = \sum_{J'} p^{|J^c\cap J'^c|}\sigma^{\circ,\imath(J')}(\tau)$.

(4) Suppose that $\imath(J) \in \mathcal{P}_\tau$*, that* $j \notin J$*, and that* $\imath(J \cup \{j\}) \in \mathcal{P}_\tau$*. Then we have inclusions*

$$p\sigma^{\circ}_{\imath(J\cup\{j\})}(\tau) \subseteq \sigma^{\circ}_{\imath(J)}(\tau) \subseteq \sigma^{\circ}_{\imath(J\cup\{j\})}(\tau).$$

Furthermore the Serre weights occuring as Jordan–Hölder factors of the cokernel of the inclusion $\sigma^{\circ}_{\imath(J)}(\tau) \subseteq \sigma^{\circ}_{\imath(J\cup\{j\})}(\tau)$ *are the* $\overline{\sigma}_{\imath(J')}(\tau)$ *with* $j \in J'$*, and the Serre weights occuring in the cokernel of the inclusion* $p\sigma^{\circ}_{\imath(J\cup\{j\})}(\tau) \subseteq \sigma^{\circ}_{\imath(J)}(\tau)$ *are the* $\overline{\sigma}_{\imath(J')}(\tau)$ *with* $j \notin J'$.

Proof. (1) This is immediate from (5.15) and Proposition 4.1.4.

(2) The inclusions are immediate from (1), and the claims on the Serre weights follow easily from considering the composite maps

$$p^{\varepsilon_{\imath(J')}(\sigma^{\circ,\imath(J\cup\{j\})}(\tau))}\sigma^{\circ}_{\imath(J')}(\tau) = p^{|J'\cap(J\cup\{j\})|}\sigma^{\circ}_{\imath(J')}(\tau) \subseteq \sigma^{\circ,\imath(J\cup\{j\})}(\tau) \subseteq \sigma^{\circ,\imath(J)}(\tau)$$

and comparing with $p^{\varepsilon_{\imath(J')}(\sigma^{\circ,\imath(J)}(\tau))}\sigma^{\circ}_{\imath(J')}(\tau) = p^{|J'\cap J|}\sigma^{\circ}_{\imath(J')}(\tau) \subseteq \sigma^{\circ,\imath(J)}(\tau)$.

(3) This follows by an argument analogous to the proof of (1).

(4) This follows by applying (2) to the contragredient of $\sigma(\tau)$. $\square$

6. Patching functors

In this section we will introduce the cohomology groups that we will consider, and we will apply the Taylor–Wiles–Kisin method to them. We will follow Sections 4 and 5 of [GK12], but we will take a more functorial approach, and we also follow [BD14] in laying the foundations for the multiplicity one results that we will prove in Section 10. Our functorial version of the Taylor–Wiles–Kisin method is a refinement of the approach taken in ([Kis09a], [EG14], [GK12]), where the patching argument is applied simultaneously for various lattices in types, and for certain filtrations on the reductions of these types modulo p; here we take this construction to its natural level of generality, and patch simultaneously over all finitely generated representations. (One could also presumably take a limit over all tame levels simultaneously, but as we have no need to do so for the applications we have in mind we have avoided setting up the necessary foundations.)

Of course, the resulting functor depends on the non-canonical choice of various auxiliary primes, but it seems reasonable to expect that it in fact enjoys various canonical properties. In fact, under appropriate genericity hypotheses, the results of Sections 8 and 10 below show that the restriction of the patching functor to the subquotients of lattices in a fixed tame type is independent of the choices of auxilliary primes, and of the particular global context; see Remark 10.1.15. See also [CEG$^+$a] for a further discussion of this in the case of $\mathrm{GL}_2/\mathbb{Q}_p$. With this in mind, and with an eye to future applications, we have chosen to set up a general abstract notion of a patching functor, and then to show that the various axioms for such a functor are satisfied in our setting. There are a variety of possible levels of generality in which one could work, and a variety of axioms one could require; for clarity, we have chosen to restrict ourselves to the case of two-dimensional representations, as this is what is used in the rest of the paper.

6.1. **Abstract patching functors.** Fix a prime p. Continue to fix a finite extension E of $\mathbb{Q}_p$ with ring of integers $\mathcal{O}$ and residue field $\mathbb{F}$. In this subsection, by a local field we mean a finite extension of some $\mathbb{Q}_l$, with l not necessarily equal to p.

Fix a finite set of local fields L_i, write l_i for the residue characteristic of L_i, and $\mathcal{O}_{L_i}$ for its ring of integers. If $l_i = p$, then we assume that $L_i/\mathbb{Q}_p$ is unramified. For each i we fix a continuous representation $\overline{\rho}_i : G_{L_i} \to \mathrm{GL}_2(\mathbb{F})$. (In the patching functors we construct, the L_i will all be localisations of some global field, and the $\overline{\rho}_i$ will be localisations of a global Galois representation, but with an eye to future applications we do not assume this in the definition of a patching functor.)

For each i, we will let K_i be either a compact open subgroup of $\mathrm{GL}_2(\mathcal{O}_{L_i})$, or a compact open subgroup of $\mathcal{O}_{D_i}^\times$, where $\mathcal{O}_{D_i}$ is a maximal order in the nonsplit quaternion algebra D_i with centre L_i. If $l_i = p$, we only allow K_i to be a compact open subgroup of $\mathrm{GL}_2(\mathcal{O}_{L_i})$. Write $K = \prod_i K_i$, and write $Z_K = Z(K)$. The reduced norm and determinant maps give a natural homomorphism $Z_K \to \prod_i \mathcal{O}_{L_i}^\times$.

For each i we have the universal framed deformation ring $R_i^\square$ for $\overline{\rho}_i$ over $\mathcal{O}$. Write $X_i = \mathrm{Spf}\, R_i^\square$. Write $R = \widehat{\otimes}_{\mathcal{O}} R_i^\square$, and write $X = \mathrm{Spf}\, R = \times_{\mathrm{Spf}\,\mathcal{O}} X_i$. Finally, fix some $h \geq 0$, and write $R_\infty = R[\![x_1, \dots, x_h]\!]$, $X_\infty = \mathrm{Spf}\, R_\infty$, where the x_i are formal variables.

Let $\mathcal{C}$ denote the category of finitely generated $\mathcal{O}$-modules with a continuous action of K, and let $\mathcal{C}'$ be a Serre subcategory of $\mathcal{C}$. For each i, let τ_i be an inertial type for I_{L_i}, assumed to be discrete series if K_i is a subgroup of $\mathcal{O}_{D_i}^\times$, and let $\sigma^\circ(\tau_i)$ denote some $\mathrm{GL}_2(\mathcal{O}_{L_i})$-stable $\mathcal{O}$-lattice in $\sigma(\tau_i)$ (in the case that K_i is a subgroup of $\mathrm{GL}_2(\mathcal{O}_{L_i})$), or an $\mathcal{O}_{D_i}^\times$-stable $\mathcal{O}$-lattice in $\sigma_D(\tau_i)$ (in the case that K_i is a subgroup of $\mathcal{O}_{D_i}^\times$). Write $\sigma^\circ(\tau) := \otimes_{\mathcal{O}} \sigma^\circ(\tau_i)$, an element of $\mathcal{C}$.

If $l_i \neq p$, we write $R_i^{\square,\tau}$ for the reduced, p-torsion free quotient of $R_i^\square$ corresponding to deformations of inertial type τ_i, while if $l_i = p$ we write $R_i^{\square,\tau}$ for the reduced, p-torsion free quotient of $R_i^\square$ corresponding to potentially crystalline deformations of inertial type τ_i and Hodge type 0. We write $R^\tau := \widehat{\otimes}_{\mathcal{O}} R_i^{\square,\tau}$, $R_\infty^\tau := R^\tau[\![x_1, \dots, x_h]\!]$, and $X_\infty(\tau) := \mathrm{Spf}(R_\infty^\tau)$.

Suppose that $\overline{\sigma} \in \mathcal{C}$ is killed by ϖ_E, and has the property that for each i with $l_i = p$, we have $K_i = \mathrm{GL}_2(\mathcal{O}_{L_i})$, and the restriction of $\overline{\sigma}$ to K_i is a direct sum of copies of an irreducible representation $\overline{\sigma}_i$ of K_i. In this case, for each i with $l_i = p$, we let $R_i^{\square,\overline{\sigma}}$ be the reduced, p-torsion free quotient of $R_i^\square$ corresponding to crystalline deformations of Hodge type $\overline{\sigma}_i$, and for each i with $l_i \neq p$ we set $R_i^{\square,\overline{\sigma}} = R_i^\square$. Then we write $R^{\overline{\sigma}} := \widehat{\otimes}_{\mathcal{O}} R_i^{\square,\overline{\sigma}}$, $R_\infty^{\overline{\sigma}} := R^{\overline{\sigma}}[\![x_1, \dots, x_h]\!]$, and $X_\infty(\overline{\sigma}) := \mathrm{Spf}(R_\infty^{\overline{\sigma}})$. We write $\overline{X}_\infty(\overline{\sigma})$ for the special fibre of the space $X_\infty(\overline{\sigma})$.

6.1.1. **Definition.** By a *patching functor* we will mean a covariant exact functor M_∞ from $\mathcal{C}'$ to the category of coherent sheaves on X_∞ which satisfies the following two properties for all $\sigma^\circ(\tau)$ and $\overline{\sigma}$ as above.

- $M_\infty(\sigma^\circ(\tau))$ is p-torsion free and supported on $X_\infty(\tau)$, and in fact is maximal Cohen–Macaulay over $X_\infty(\tau)$.

- $M_\infty(\overline{\sigma})$ is supported on $\overline{X}_\infty(\overline{\sigma})$, and in fact is maximal Cohen–Macaulay over $\overline{X}_\infty(\overline{\sigma})$.

6.1.2. *Remark.* Unless stated otherwise, all of our patching functors will be defined on all of $\mathcal{C}$, but it will be convenient in some arguments to have the additional flexibility of passing to a Serre subcategory.

We will now define the notion of a *fixed determinant patching functor.* For each i, in addition to the data above, we also fix a (lift of a) geometric Frobenius element $\mathrm{Frob}_i \in G_{L_i}$, and we fix an element $\alpha_i \in \mathcal{O}_{L_i}$ lifting $\det \overline{\rho}_i(\mathrm{Frob}_i)$. Let $\mathcal{C}_Z$ denote the subcategory of $\mathcal{C}$ consisting of representations which have a central character, so that the action of Z_K factors through the natural map to $\prod_i \mathcal{O}_{L_i}^\times$, and whose central character lifts the character $\prod_i(\overline{\varepsilon}\det\overline{\rho}_i|_{I_{L_i}})\circ \mathrm{Art}_{L_i}$. Let $\mathcal{C}'_Z$ be a Serre subcategory of $\mathcal{C}_Z$.

Take some $\sigma \in \mathcal{C}'_Z$. For each i, fix a character $\psi_{\sigma,i} : G_{L_i} \to \mathcal{O}^\times$ with the properties that $\psi_{\sigma,i}(\mathrm{Frob}_i) = \alpha_i$, and the composite of Art_{L_i} and the restriction of $\psi_{\sigma,i}$ to I_{L_i} gives the central character of σ. (If σ is not p-power torsion, then $\psi_{\sigma,i}$ is uniquely determined.) Write $R_i^{\square,\psi_\sigma}$ for the quotient of $R_i^\square$ corresponding to liftings with determinant $\psi_{\sigma,i}\varepsilon^{-1}$, and $X_i^{\psi_\sigma} = \mathrm{Spf}\, R_i^{\square,\psi_\sigma}$. Write $R^{\psi_\sigma} = \widehat{\otimes}_{\mathcal{O}} R_i^{\square,\psi_\sigma}$, and set $X^{\psi_\sigma} = \mathrm{Spf}\, R^{\psi_\sigma} = \times_{\mathrm{Spf}\,\mathcal{O}} X_i^{\psi_\sigma}$. Finally, fix some $h \geq 0$, and write $R_\infty^{\psi_\sigma} = R^{\psi_\sigma}[\![x_1,\dots,x_h]\!]$, $X_\infty^{\psi_\sigma} = \mathrm{Spf}\, R_\infty^{\psi_\sigma}$, where the x_i are formal variables.

Assuming that the τ_i above satisfy the condition that $\det\tau_i$ is a lift of $\overline{\varepsilon}\det\overline{\rho}_i|_{I_{L_i}}$, so that $\sigma^\circ(\tau) \in \mathcal{C}_Z$, we have the obvious quotient $R_i^{\square,\psi,\tau}$ of $R_i^{\square,\psi}$, and we write $R^{\psi,\tau} := \widehat{\otimes}_{\mathcal{O}} R_i^{\square,\psi,\tau}$, $R_\infty^{\psi,\tau} := R^{\psi,\tau}[\![x_1,\dots,x_h]\!]$, and $X^\psi(\tau) := \mathrm{Spf}(R_\infty^{\psi,\tau})$. If $\overline{\sigma} \in \mathcal{C}_Z$, then we have the quotient $R_i^{\square,\psi,\overline{\sigma}}$ of $R_i^{\square,\overline{\sigma}}$, and we write $R^{\psi,\overline{\sigma}} := \widehat{\otimes}_{\mathcal{O}} R_i^{\square,\psi,\overline{\sigma}}$, $R_\infty^{\psi,\overline{\sigma}} := R^{\psi,\overline{\sigma}}[\![x_1,\dots,x_h]\!]$, and $X^\psi(\overline{\sigma}) := \mathrm{Spf}(R_\infty^{\psi,\overline{\sigma}})$. We write $\overline{X}^\psi(\overline{\sigma})$ for the special fibre of the space $X^\psi(\overline{\sigma})$.

6.1.3. **Definition.** By a *fixed determinant patching functor* we will mean a covariant exact functor M_∞ from $\mathcal{C}'_Z$ to the category of coherent sheaves on X_∞ which satisfies the following properties.

- For all $\sigma \in \mathcal{C}'_Z$, $M_\infty(\sigma)$ is supported on $X_\infty^{\psi_\sigma}$.
- $M_\infty(\sigma^\circ(\tau))$ is p-torsion free and supported on $X_\infty^\psi(\tau)$, and in fact is maximal Cohen–Macaulay over $X_\infty^\psi(\tau)$.
- $M_\infty(\overline{\sigma})$ is supported on $\overline{X}_\infty^\psi(\overline{\sigma})$, and in fact is maximal Cohen–Macaulay over $\overline{X}_\infty^\psi(\overline{\sigma})$.

We have the following consequence of these definitions, an abstract version of the multiplicity one argument of [Dia97].

6.1.4. **Lemma.** *If M_∞ is a patching functor, and if the deformation space $X_\infty(\tau)$ is regular, then $M_\infty(\sigma^\circ(\tau))$ is free over $X_\infty(\tau)$. Similarly, if the generic fibre of $X_\infty(\tau)$ is regular, then the restriction of $M_\infty(\sigma^\circ(\tau))$ to the generic fibre is locally free. The analogous result also holds for fixed determinant patching functors.*

Proof. This follows immediately from the Auslander–Buchsbaum formula and the maximal Cohen–Macaulay property of $M_\infty(\sigma^\circ(\tau))$. □

We say that a patching functor (respectively a fixed determinant patching functor) M_∞ is *minimal* if for any inertial type τ for which $X_\infty(\tau)$ (respectively $X_\infty^\psi(\tau)$)

has regular generic fibre, $M_\infty(\sigma^\circ(\tau))$ is locally free of rank one over the generic fibre of $X_\infty(\tau)$ (respectively $X^\psi_\infty(\tau)$). Note that by Lemma 6.1.4, $M_\infty(\sigma^\circ(\tau))$ is locally free over the generic fibre of $X_\infty(\tau)$ (respectively $X^\psi_\infty(\tau)$), so this should be regarded as a multiplicity one hypothesis. (The reason we call these patching functors minimal is that we will construct minimal patching functors by considering the cohomology of Shimura curves at minimal level.)

6.1.5. **Definition.** We will sometimes say that M_∞ is a patching functor *indexed by* the $(L_i, \overline{\rho}_i)$, or that M_∞ is a patching functor for the $\overline{\rho}_i$. If the coefficient field E is unramified over $\mathbb{Q}_p$, we will say that M_∞ is a patching functor *with unramified coefficients*. We will also use the same terminology for fixed determinant patching functors.

In the rest of this section we will give examples of patching functors arising from the Taylor–Wiles–Kisin method. These patching functors will satisfy an additional property which will be important for us, but which we have not attempted to axiomatise, which is that certain fibres of the patching functors will give spaces of automorphic forms.

6.2. **Quaternion algebras.** In this subsection we will describe the various spaces of automorphic forms on quaternion algebras from which we will construct our patching functors. To this end, let F be a totally real field, and suppose that the prime p is odd and is unramified in F. Let $\overline{\rho}: G_F \to \mathrm{GL}_2(\mathbb{F})$ be a continuous representation. Assume that $\overline{\rho}$ is modular, and that $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible. If $p = 5$, assume further that the projective image of $\overline{\rho}|_{G_{F(\zeta_5)}}$ is not isomorphic to A_5. Choose a finite order character $\psi : G_F \to \mathcal{O}^\times$ such that $\det \overline{\rho} = \overline{\varepsilon}^{-1}\overline{\psi}$.

Let D be a quaternion algebra with centre F which is ramified at a set Σ of finite places of F. We assume that Σ does not contain any places lying over p. We also assume that either D is ramified at all infinite places (the *definite case*), or split at precisely one infinite place (the *indefinite case*). Let S be a set of finite places of F containing Σ, the places dividing p, and the places at which $\overline{\rho}$ or ψ is ramified. We will define a patching functor indexed by the $(F_w, \overline{\rho}|_{G_{F_w}})$ for $w \in S$, and from now on we will write the elements of our indexing set as w, rather than as i as in the previous section. We will write K_S for the group K in the definition of a patching functor.

By Lemma 4.11 of [DDT97] we can and do choose a finite place $w_1 \notin S$ with the properties that

- $\mathbf{N}w_1 \not\equiv 1 \pmod p$,
- the ratio of the eigenvalues of $\overline{\rho}(\mathrm{Frob}_{w_1})$ is not equal to $(\mathbf{N}w_1)^{\pm 1}$, and
- the residue characteristic of w_1 is sufficiently large that for any non-trivial root of unity ζ in a quadratic extension of F, w_1 does not divide $\zeta+\zeta^{-1}-2$.

We will consider compact open subgroups $K = \prod_w K_w$ of $(D \otimes_F \mathbb{A}_F^\infty)^\times$ for which $K_w \subset (\mathcal{O}_D)_w^\times$ for all w, $K_w = (\mathcal{O}_D)_w^\times$ if $w \notin S \cup \{w_1\}$, and K_{w_1} is the subgroup of $\mathrm{GL}_2(\mathcal{O}_{F_{w_1}})$ consisting of elements that are upper-triangular and unipotent modulo w_1. Our assumptions on w_1 imply that K is sufficiently small, in the sense that condition (2.1.2) of [Kis09a] holds.

We now define the spaces of modular forms with which we will work. In fact, the Taylor–Wiles–Kisin method naturally patches homology, rather than cohomology, so it will be convenient for us to work with Pontrjagin duals of spaces of modular forms. Consider some $\sigma \in \mathcal{C}_Z$. As in the definition of a fixed determinant patching functor, we fix characters $\psi_{\sigma,w} : G_{F_w} \to \mathcal{O}^\times$ for $w \in S$, determined (when σ is not p-power torsion) by the central character of σ. By the hypothesis on the central characters of elements of $\mathcal{C}_Z$, and the definition of ψ, we see that $\overline{\psi}_{\sigma,w}^{-1}\overline{\psi}|_{G_{F_w}} = 1$.

By Hensel's lemma (and the assumption that $p > 2$), for each $w \in S$ there is a unique character $\theta_w : G_{F_w} \to \mathcal{O}^\times$ with the properties that $\overline{\theta}_w = 1$ and $\theta_w^2 = \psi_{\sigma,w}^{-1}\psi|_{G_{F_w}}$. Write $\theta = \otimes_w \theta_w$, and write $\sigma(\theta)$ for the twist of σ by $\otimes_w(\theta_w \circ \mathrm{Art}_{F_w} \circ \det)$. Then $\sigma(\theta)$ has an action of K via the projection onto K_S, which we extend to an action of $K \cdot (\mathbb{A}_F^\infty)^\times$ by letting $(\mathbb{A}_F^\infty)^\times$ act via the composite $(\mathbb{A}_F^\infty)^\times \to (\mathbb{A}_F^\infty)^\times/F^\times \to \mathcal{O}^\times$, with the second map the one induced by $\psi \circ \mathrm{Art}_F$. (Note that the actions of K and $(\mathbb{A}_F^\infty)^\times$ agree on their intersection; since ψ is unramified outside of S, this is automatic at places away from S, and follows from the definition of the θ_w at the places in S.)

Suppose first that we are in the indefinite case. As in Section 3 of [BD14] there is a smooth projective algebraic curve X_K over F associated to K. Then there is a local system $\mathcal{F}_{\sigma(\theta)^*}$ on X_K corresponding to $\sigma(\theta)^*$ in the usual way, and we define

$$S(\sigma) := H^1((X_K)_{/\overline{F}}, \mathcal{F}_{\sigma(\theta)^*}).$$

Suppose now that we are in the definite case. Then we let $S(\sigma)$ be the space of continuous functions

$$f : D^\times \backslash (D \otimes_F \mathbb{A}_F^\infty)^\times \to \sigma(\theta)^*$$

such that we have $f(gu) = u^{-1}f(g)$ for all $g \in (D \otimes_F \mathbb{A}_F)^\times$, $u \in K(\mathbb{A}_F^\infty)^\times$.

We let $\mathbb{T}^{S,\mathrm{univ}}$ be the commutative polynomial algebra over $\mathcal{O}$ generated by formal variables T_w, S_w for each finite place $w \notin S \cup \{w_1\}$ of F. For each finite place w of F, fix a uniformiser ϖ_w of $\mathcal{O}_{F_w}$. Then $\mathbb{T}^{S,\mathrm{univ}}$ acts on $S(\sigma)$ in the usual way, with the corresponding double cosets being

$$T_w = \left[\mathrm{GL}_2(\mathcal{O}_{F_w}) \begin{pmatrix} \varpi_w & 0 \\ 0 & 1 \end{pmatrix} \mathrm{GL}_2(\mathcal{O}_{F_w})\right],$$

$$S_w = \left[\mathrm{GL}_2(\mathcal{O}_{F_w}) \begin{pmatrix} \varpi_w & 0 \\ 0 & \varpi_w \end{pmatrix} \mathrm{GL}_2(\mathcal{O}_{F_w})\right].$$

Let $\mathfrak{m}$ be the maximal ideal of $\mathbb{T}^{S,\mathrm{univ}}$ with residue field $\mathbb{F}$ with the property that for each finite place $w \notin S \cup \{w_1\}$ of F, the characteristic polynomial of $\overline{\rho}(\mathrm{Frob}_w)$ is equal to the image of $X^2 - T_w X + (\mathbf{N}w)S_w$ in $\mathbb{F}[X]$. Write $\mathbb{T}(\sigma)$ for the image of $\mathbb{T}^{S,\mathrm{univ}}$ in $\mathrm{End}_{\mathcal{O}}(S(\sigma))$. The assumption that $\overline{\rho}$ is modular means that we can (and do) assume that we have chosen S and ψ in such a way that for some σ, we have $S(\sigma)_\mathfrak{m} \neq 0$. In both the definite and indefinite cases the functor $\sigma \mapsto S(\sigma)_\mathfrak{m}$ is exact, because K is sufficiently small and $\mathfrak{m}$ is non-Eisenstein.

We will use our characters θ_w to twist the universal liftings we consider. The Taylor–Wiles–Kisin method applied to the spaces $S(\sigma)_\mathfrak{m}$ will naturally produce modules over $R^\psi = \widehat{\otimes}_{\mathcal{O}} R_w^{\square,\psi|_{G_{F_w}}}$, rather than over $R^{\psi_\sigma} = \widehat{\otimes}_{\mathcal{O}} R_w^{\square,\psi_\sigma}$. However, twisting the universal lifting to $R_w^{\square,\psi|_{G_{F_w}}}$ by $\theta_w^{-1} \circ \det$ gives a canonical isomorphism $R_w^{\square,\psi_\sigma} \xrightarrow{\sim} R_w^{\square,\psi|_{G_{F_w}}}$ for each w, and thus a canonical isomorphism $R^{\psi_\sigma} \xrightarrow{\sim} R^\psi$, so from now on we can and do think of any R^ψ-module as an R^{ψ_σ}-module.

Let R_S^{univ} be the universal deformation ring for deformations of $\overline{\rho}$ which are unramified outside of S, and let $\rho^{\mathrm{univ}} : G_{F,S} \to \mathrm{GL}_2(R_S^{\mathrm{univ}})$ denote a choice of the universal deformation. There is a natural homomorphism $R_S^{\mathrm{univ}} \to \mathbb{T}(\sigma)_\mathfrak{m}$ which for each finite place $w \notin S \cup \{w_1\}$ sends the characteristic polynomial of $\rho^{\mathrm{univ}}(\mathrm{Frob}_w)$ to $X^2 - T_w X + (\mathbf{N}w)S_w$. We write $\rho(\sigma)_{\mathfrak{m}}$ for the composite $G_{F,S} \xrightarrow{\rho^{\mathrm{univ}}} \mathrm{GL}_2(R_S^{\mathrm{univ}}) \to \mathrm{GL}_2(\mathbb{T}(\sigma)_{\mathfrak{m}})$.

In the definite case, write $M(\sigma)$ for $S(\sigma)_{\mathfrak{m}}^*$. In the indefinite case, we "factor out" the action of $G_{F,S}$ by defining

$$M(\sigma) := (\mathrm{Hom}_{\mathbb{T}(\sigma)_{\mathfrak{m}}[G_{F,S}]}(\rho(\sigma)_{\mathfrak{m}}, S(\sigma)_{\mathfrak{m}}))^*.$$

By a standard argument, the evaluation map

$$M(\sigma)^* \otimes_{\mathbb{T}(\sigma)_{\mathfrak{m}}} \rho(\sigma)_{\mathfrak{m}} \to S(\sigma)_{\mathfrak{m}}$$

is an isomorphism of $\mathbb{T}(\sigma)_{\mathfrak{m}}[G_{F,S}]$-modules. (See, for example, Théorème 4 of [Car94] or Proposition 5.5.3 of [Eme].)

6.3. **Infinite level.** With an eye ahead to Section 8, we also note a related construction, of spaces of modular forms of infinite p-power level. (Note that we will not patch these spaces via the Taylor–Wiles method, although it is possible to do so, *cf.* [CEG+b].) Let $v|p$ be a fixed place of F, and write the group K_S above as $K_S = K_v K^v$, with $K^v = \prod_{w \in S \setminus \{v\}} K_w$. Fix a continuous finitely generated representation σ^v of K^v, and let σ be the representation of K_S given by twisting the action of K^v on σ^v by the character $\psi \circ \mathrm{Art}_{F_v} \circ \det$ of K_v. We assume that $\sigma \in \mathcal{C}_Z$, and write $S(K_v, \sigma^v)_{\mathfrak{m}}$ for $S(\sigma)_{\mathfrak{m}}$.

Then we define

$$S^v(\sigma^v) := \varinjlim_{K_v} S(K_v, \sigma^v),$$

where the direct limit is taken over the directed system of compact open subgroups of $\mathrm{GL}_2(\mathcal{O}_{F_v})$. We let $\mathbb{T}(\sigma^v)$ denote the image of $\mathbb{T}^{S,\mathrm{univ}}$ in $\mathrm{End}_{\mathcal{O}}(S^v(\sigma^v))$, and we write $\rho(\sigma^v)_{\mathfrak{m}}$ for the composite $G_{F,S} \xrightarrow{\rho^{\mathrm{univ}}} \mathrm{GL}_2(R_S^{\mathrm{univ}}) \to \mathrm{GL}_2(\mathbb{T}(\sigma^v)_{\mathfrak{m}})$. If we are in the definite case then we set $M^v(\sigma^v) := S^v(\sigma^v)_{\mathfrak{m}}^*$, and if we are in the indefinite case we set

$$M^v(\sigma^v) := (\mathrm{Hom}_{\mathbb{T}(\sigma^v)_{\mathfrak{m}}[G_{F,S}]}(\rho(\sigma^v)_{\mathfrak{m}}, S^v(\sigma^v)_{\mathfrak{m}}))^*.$$

Again, we have an isomorphism

$$M^v(\sigma^v)^* \otimes_{\mathbb{T}(\sigma^v)_{\mathfrak{m}}} \rho(\sigma^v)_{\mathfrak{m}} \to S^v(\sigma^v)_{\mathfrak{m}}.$$

Since $\mathfrak{m}$ is non-Eisenstein, in either the define or indefinite cases if we fix some K_v and a representation σ_v of K_v such that $\sigma := \sigma_v \otimes_{\mathcal{O}} \sigma^v$ is an element of $\mathcal{C}_Z$, then we have a natural isomorphism

$$M(\sigma) \xrightarrow{\sim} \mathrm{Hom}_{K_v}(M^v(\sigma^v), \sigma_v^*)^*. \tag{6.1}$$

6.4. **Functorial Taylor–Wiles–Kisin patching.** Since there are only countably many isomorphism classes of finite length $\mathcal{O}$-modules with a smooth action of K_S, applying the usual Taylor–Wiles–Kisin patching argument to the $M(\sigma)$ as in Section 5 of [GK12] gives us the following data.

- positive integers g, q,
- $S_\infty = \mathcal{O}[[x_1, \dots, x_{4\#S+q-1}]]$, a power series ring in $4\#S + q - 1$ variables over $\mathcal{O}$,
- R_∞, a power series ring in g variables over R,
- an $\mathcal{O}$-algebra homomorphism $S_\infty \to R_\infty$,
- a covariant exact functor $\sigma \mapsto M_\infty(\sigma)$ from the subcategory of finite length objects of $\mathcal{C}_Z$ to the category of R_∞-modules.

We extend M_∞ to the whole of $\mathcal{C}_Z$ by defining $M_\infty(\sigma) = \varprojlim_n M_\infty(\sigma/\varpi_E^n\sigma)$. This data satisfies the following further properties.

- There is an isomorphism $M_\infty(\sigma)/(x_1, \ldots, x_{4\#S+q-1}) \cong M(\sigma)$, and a compatible isomorphism $R_\infty/(x_1, \ldots, x_{4\#S+q-1}) \cong R_S^{\mathrm{univ}}$.
- The action of R on $M_\infty(\sigma)$ factors through R^{ψ_σ}.
- If $R_\infty^{\psi,\tau} = R_\infty \otimes_R R^{\psi,\tau}$, then $\dim R_\infty^{\psi,\tau} = \dim S_\infty$.
- If σ is a finite free $\mathcal{O}$-module, then $M_\infty(\sigma)$ is a finite free S_∞-module (where the S_∞-module structure comes from the homomorphism $S_\infty \to R_\infty$).
- If σ is a finite free $\mathbb{F}$-module, then $M_\infty(\sigma)$ is a finite free $S_\infty \otimes_{\mathcal{O}} \mathbb{F}$-module.

Regard the R_∞-module $M_\infty(\sigma)$ as a coherent sheaf on $X_\infty = \operatorname{Spf} R_\infty$. We now verify the hypotheses of Definition 6.1.3; the first hypothesis is clear from the construction.

In order to check the hypothesis on $M_\infty(\sigma^\circ(\tau))$, we need the action of R on $M_\infty(\sigma^\circ(\tau))$ to factor through $R^{\psi,\tau}$, and we need $M_\infty(\sigma^\circ(\tau))$ to be maximal Cohen–Macaulay over $R_\infty^{\psi,\tau}$. The first property is an immediate consequence of local-global compatibility and the construction of M_∞, and the claim that $M_\infty(\sigma^\circ(\tau))$ is maximal Cohen–Macaulay over $R_\infty^{\psi,\tau}$ is immediate from the Auslander–Buchsbaum formula and the fact that $M_\infty(\sigma^\circ(\tau))$ is free over the regular local ring S_∞.

It remains to check the hypothesis on $M_\infty(\overline{\sigma})$. It is enough to check that the action of R on $M_\infty(\overline{\sigma})$ factors through $R^{\psi,\overline{\sigma}}$, as the claim that it is maximal Cohen–Macaulay then follows exactly as in the previous paragraph. This is immediate from Corollaries 5.6.4 and 4.5.7 of [GK12] and the construction of M_∞.

6.5. **Minimal level.** Under some additional hypotheses (which will be satisfied in the cases of ultimate interest to us in this paper), we now refine the constructions of the previous section to produce a minimal fixed determinant patching functor with unramified coefficients. We will use this functor in Section 10.

We continue to use the notation and assumptions introduced in the previous section, so that in particular p is unramified in F. For the remainder of this section, we make the following additional assumptions:

- $p > 3$,
- $\overline{\rho}|_{G_{F_w}}$ is generic for all places $w|p$, and
- if $w \in \Sigma$, then $\overline{\rho}|_{G_{F_w}}$ is not scalar.

Since $p > 3$ is unramified in F, it follows from the proof of Lemma 4.11 of [DT94] that we can choose the place w_1 so that $\overline{\rho}(\mathrm{Frob}_{w_1})$ has distinct eigenvalues, and we choose w_1 in this manner. If necessary, we replace $\mathbb{F}$ with a quadratic extension so that these eigenvalues are in $\mathbb{F}$. Following [BD14], we now explain a slight refinement of the above constructions in the case of minimal level.

Fix a place $v|p$. Set $E = W(\mathbb{F})[1/p]$, let ψ be the Teichmüller lift of $\overline{\varepsilon}\det\overline{\rho}$, and let S be the union of Σ, the places dividing p, and the places where $\overline{\rho}$ is ramified. We will ultimately construct a minimal patching functor for which the corresponding index set is just the place v, but we will begin by constructing a patching functor for which the index set is the set of places S.

Following Section 3.3 of [BD14] very closely (which defines, under slightly different hypotheses, the corresponding spaces in characteristic p), we begin by defining the spaces of modular forms that we will use. Set $K_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$. For each place $w \in S \setminus \{v\}$, we will now define a compact open subgroup K_w of $(\mathcal{O}_D)_w^\times$ and a continuous representation of K_w on a finite free $\mathcal{O}$-module L_w. Let S' be the union of the set of places $w|p, w \neq v$ for which $\overline{\rho}|_{G_{F_w}}$ is reducible and the set of places w for which $w \in \Sigma$ and $\overline{\rho}|_{G_{F_w}}$ is reducible, and the place w_1; then at the places $w \in S'$, we will also define Hecke operators T_w and scalars $\beta_w \in \mathbb{F}^\times$.

- If $w|p, w \neq v$, and $\overline{\rho}|_{G_{F_w}}$ is irreducible, then we let $K_w = \mathrm{GL}_2(\mathcal{O}_{F_w})$, we choose a Serre weight $\overline{\sigma}_w \in \mathcal{D}(\overline{\rho}|_{G_{F_w}})$, we let τ_w be the corresponding inertial type from Proposition 3.5.1, and we let L_w be a $\mathrm{GL}_2(\mathcal{O}_{F_w})$-stable $\mathcal{O}$-lattice in $\sigma(\tau_w)$.
- If $w|p, w \neq v$, and $\overline{\rho}|_{G_{F_w}}$ is reducible, then the assumption that $\overline{\rho}|_{G_{F_w}}$ is generic implies in particular that it is not scalar, so we are in the situation considered in Section 3.3 of [BD14]. Write $\overline{\rho}|_{G_{F_w}} \cong \begin{pmatrix} \overline{\xi}_w & * \\ 0 & \overline{\xi}'_w \omega^{-1} \end{pmatrix}$. Note that the assumption that $\overline{\rho}|_{G_{F_w}}$ is generic also implies that $\overline{\xi}_w|_{I_{F_w}} \neq \overline{\xi}'_w|_{I_{F_w}}$.

  Let K_w be the Iwahori subgroup of $\mathrm{GL}_2(\mathcal{O}_{F_w})$ consisting of matrices which are upper-triangular modulo w, and let $\overline{\gamma}_w : K_w \to \mathbb{F}^\times$ be the character which sends a matrix congruent to $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ modulo w to $(\overline{\xi} \circ \mathrm{Art}_{F_w})(a)(\overline{\xi}' \circ \mathrm{Art}_{F_w})(d)$. Let γ_w be the Teichmüller lift of $\overline{\gamma}_w$, and let $L_w = \mathcal{O}(\gamma_w)$.

  Choose a uniformiser ϖ_w of $\mathcal{O}_{F_w}$, let $V_w := \ker(\gamma_w)$, and let T_w be the Hecke operator defined by the double coset $V_w \begin{pmatrix} \varpi_w & 0 \\ 0 & 1 \end{pmatrix} V_w$. Let $\beta_w = (\overline{\xi}_w \circ \mathrm{Art}_{F_w})(\varpi_w)$.
- If $w \in \Sigma \cap S'$ (equivalently: if $w \in \Sigma$ and $\overline{\rho}|_{G_{F_w}}$ is reducible), then the assumption that some $S(\sigma)_\mathfrak{m} \neq 0$ implies that $\overline{\rho}|_{G_{F_w}}$ is a twist of an extension of the trivial character by the mod p cyclotomic character; furthermore, if $\mathbf{N}w \equiv 1 \pmod{p}$, then the assumption that $\overline{\rho}|_{G_{F_w}}$ is not scalar implies that this extension is non-trivial. We let $K_w = (\mathcal{O}_D)_w^\times$, and we let L_w be the rank one $\mathcal{O}$-module on which $(\mathcal{O}_D)_w^\times$ acts via $\gamma_w := \psi|_{I_{F_w}} \circ \mathrm{Art}_{F_w} \circ \det$. We choose a uniformiser Π_w of $(\mathcal{O}_D)_w$, we let $V_w := \ker(\gamma_w)$, we let T_w be the Hecke operator defined by the double coset $V_w \Pi_w V_w$, and we set $\beta_w := (\overline{\gamma}_w \circ \mathrm{Art}_{F_w})(\det \Pi_w)$.
- If $w \notin \Sigma$, $w \nmid p$ and $\overline{\rho}|_{G_{F_w}}$ is reducible, we choose a character $\overline{\xi}_w : G_{F_w} \to \mathbb{F}^\times$ such that $\overline{\rho}'_w := \overline{\xi}_w^{-1} \overline{\rho}|_{G_{F_w}}$ has minimal conductor among the twists of $\overline{\rho}|_{G_{F_w}}$ by characters. Write ξ_w for the Teichmüller lift of $\overline{\xi}_w$. We let n_w be the (exponent of the) conductor of $\overline{\rho}'_w$, we let μ_w be the Teichmüller lift of $\det \overline{\rho}'_w$, and we let K_w be the subgroup of $\mathrm{GL}_2(\mathcal{O}_{F_w})$ consisting of matrices which are upper-triangular modulo w^{n_w}.

  We let L_w be the rank one $\mathcal{O}$-module on which K_w acts via a character γ_w, which is defined as follows. If $n_w = 0$, then $\gamma_w = \xi_w \circ \mathrm{Art}_{F_w} \circ \det$. If $n_w > 1$, we let $\gamma_w(g) = (\xi_w \circ \mathrm{Art}_{F_w})(\det(g))(\mu_w \circ \mathrm{Art}_{F_w})(d)$, where $g \equiv \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \pmod{w^{n_w}}$.
- If $w \nmid p$ and $\overline{\rho}|_{G_{F_w}}$ is irreducible, then we let τ_w be any inertial type for which $\overline{\rho}|_{G_{F_w}}$ has a lift of type τ_w and determinant $\psi|_{G_{F_w}} \varepsilon^{-1}$, and we let L_w be an $(\mathcal{O}_D)_w^\times$-invariant lattice in $\sigma(\tau_w)$ (respectively $\sigma_D(\tau_w)$ if $w \in \Sigma$). (It is easy to check that we can choose τ_w and thus L_w to be defined over $\mathcal{O}$, either by a consideration of the rationality properties of the local Langlands

correspondence, or by explicitly choosing τ_w via the classification of the possible $\overline{\rho}|_{G_{F_w}}$.)

- If $w = w_1$, we let K_{w_1} be the subgroup of $\mathrm{GL}_2(\mathcal{O}_{F_{w_1}})$ consisting of those matrices which are upper-triangular modulo w_1. We also define a Hecke operator T_{w_1} by
$$T_{w_1} := \left[K_{w_1}\begin{pmatrix}\varpi_{w_1} & 0\\ 0 & 1\end{pmatrix}K_{w_1}\right].$$
Note that T_{w_1} depends on the choice of ϖ_{w_1}. Fix from now on a choice β_{w_1} of eigenvalue of $\overline{\rho}(\mathrm{Frob}_{w_1})$, where $\mathrm{Frob}_{w_1} = \mathrm{Art}_{F_{w_1}}(\varpi_{w_1})$.

Write $L := \otimes_{w\in S, w\neq v, \mathcal{O}} L_w$. Let $\mathcal{C}_Z^v$ be the category of finitely generated $\mathcal{O}$-modules σ_v with a continuous action of $K_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$, having the property that the central character of σ_v lifts the character $(\overline{\varepsilon}\det\overline{\rho}|_{I_{F_v}})\circ\mathrm{Art}_{F_v}$. Let $\mathcal{C}'_Z$ be the subcategory of $\mathcal{C}_Z$ consisting of modules of the form $L\otimes_{\mathcal{O}}\sigma_v$, where σ_v is an object of $\mathcal{C}_Z^v$.

For each object σ of $\mathcal{C}'_Z$, we define $S(\sigma)$ as in Section 6.4. We let $\mathbb{T}(\sigma)' = \mathbb{T}(\sigma)[T_w]_{w\in S'\cup\{w_1\}}$, and we denote the maximal ideal of $\mathbb{T}(\sigma)'$ generated by $\mathfrak{m}$ and the $T_w - \beta_w$ for $w \in S'$ by $\mathfrak{m}'$. By Lemme 3.3.1 of [BD14], there is a natural action of $\mathbb{T}(\sigma)'$ on $S(\sigma)$; we define $S^{\min}(\sigma)_{\mathfrak{m}} := S(\sigma)_{\mathfrak{m}'}$, and construct $M^{\min}(\sigma)$ from $S^{\min}(\sigma)_{\mathfrak{m}}$ in the same way that we constructed $M(\sigma)$ from $S(\sigma)_{\mathfrak{m}}$.

For each place $w \in S$, set $\alpha_w = \psi(\mathrm{Frob}_w)$. Then the patching arguments of Section 5 of [GK12] go through as in Section 6.4 to produce a fixed determinant patching functor $M_\infty^{\min}$ defined on $\mathcal{C}'_Z$, such that
$$M_\infty^{\min}(\sigma)/(x_1,\ldots,x_{4\#S+q-1}) \cong M^{\min}(\sigma).$$

We will now use the obvious functor from $\mathcal{C}_Z^v$ to $\mathcal{C}'_Z$ to construct a minimal fixed determinant patching functor on $\mathcal{C}_Z^v$ from $M_\infty^{\min}$. For each place $w \in S\setminus\{v\}$ we define a certain universal lifting ring $R_w^{\min}$ as follows.

- If $\overline{\rho}|_{G_{F_w}}$ is irreducible, then we let $R_w^{\min} = R_w^{\square,\psi|_{G_{F_w}},\tau_w}$ for the above choice of τ_w.
- If $w \nmid p$, and $\overline{\rho}|_{G_{F_w}}$ is reducible but is not a twist of an extension of the trivial character by the mod p cyclotomic character, then we let $R_w^{\min}$ be the complete local noetherian $\mathcal{O}$-algebra which prorepresents the functor which assigns to an Artinian $\mathcal{O}$-algebra A the set of lifts of $\overline{\rho}|_{G_{F_w}}$ to representations $\rho_A : G_{F_w} \to \mathrm{GL}_2(A)$ of determinant $\psi|_{G_{F_w}}\varepsilon^{-1}$, for which $\rho_A(I_{F_w}) \xrightarrow{\sim} \overline{\rho}(I_{F_w})$.
- If $w|p$ and $\overline{\rho}|_{G_{F_w}}$ is reducible, in which case we write $\overline{\rho}|_{G_{F_w}} \cong \begin{pmatrix}\overline{\eta}_w & *\\ 0 & \overline{\eta}'_w\omega^{-1}\end{pmatrix}$ for some characters $\overline{\eta}_w, \overline{\eta}'_w$, or if $w\nmid p$ and $\overline{\rho}|_{G_{F_w}} \cong \begin{pmatrix}\overline{\eta}_w & *\\ 0 & \overline{\eta}_w\omega^{-1}\end{pmatrix}$ for some character $\overline{\eta}_w$, then we let $R_w^{\min}$ be the complete local noetherian $\mathcal{O}$-algebra which prorepresents the functor which assigns to an Artinian $\mathcal{O}$-algebra A the set of pairs (ρ_A, L_A) as follows. The representation $\rho_A : G_{F_w} \to \mathrm{GL}_2(A)$ is a lift of $\overline{\rho}|_{G_{F_w}}$ of determinant $\psi\varepsilon^{-1}$, and L_A is a direct factor of A^2 (the module on which ρ_A acts), such that G_{F_w} acts on L_A by a character of the form η_w, where η_w is a lift of $\overline{\eta}_w$ such that $\eta_w(I_{F_w}) \xrightarrow{\sim} \overline{\eta}_w(I_{F_w})$.

In each case the ring $R_w^{\min}$ is formally smooth over $\mathcal{O}$; in the case that $w|p$ and $\overline{\rho}|_{G_{F_w}}$ is irreducible, this follows from Theorem 7.2.1 below, and in the other cases it follows from Lemma 3.4.1 of [BD14]. Write $R^{\min}_{S\setminus\{v\}} = \widehat{\otimes}_{w\in S, w\neq v,\mathcal{O}} R^{\min}_w$, and let $R^{\min,\psi_\sigma} = R^{\min}_{S\setminus\{v\}}\widehat{\otimes} R^{\square,\psi_\sigma}_v$. Then $R^{\min}_{S\setminus\{v\}}$ is formally smooth over $\mathcal{O}$, and it follows exactly as in Proposition 3.5.1 of [BD14] that the action of R on $M^{\min}_\infty(\sigma)$ factors through $R^{\min,\psi_\sigma}$.

Now, since $R^{\min}_{S\setminus\{v\}}$ is formally smooth over $\mathcal{O}$, the ring $R_\infty \otimes_{\mathcal{O}} R^{\min,\psi_\sigma}$ is formally smooth over $R^{\square,\psi_\sigma}_v$; so the functor which assigns $M^{\min}_\infty(\sigma_v \otimes_{\mathcal{O}} L)$ to each object σ_v of $\mathcal{C}^v_Z$ is a fixed determinant patching functor with unramified coefficients.

It only remains to check that $M^{\min}_\infty$ is minimal. In order to see this, it suffices to check that if τ_v is an inertial type for I_{F_v} with $\det \overline{\tau}_v = \overline{\varepsilon} \det \overline{\rho}|_{I_{F_v}}$, then $M(\sigma^\circ(\tau_v) \otimes_{\mathcal{O}} L)[1/p]$ is locally free of rank one over $\mathbb{T}(\sigma^\circ(\tau_v)\otimes_{\mathcal{O}} L)'[1/p]$; but this follows from the definition of L, *cf.* Proposition 3.5.1 of [BD14].

6.6. **Unitary groups.** We now give a concrete construction of a patching functor (without fixed determinant), by applying the Taylor–Wiles–Kisin method to forms of $U(2)$. The construction is extremely similar to the arguments above with quaternion algebras, and we content ourselves with sketching the details, following Section 4 of [GK12].

Let F be an imaginary CM field with maximal totally real subfield F^+, and suppose that the prime p is odd and is unramified in F, and that all places of F^+ lying over p split in F. Assume further that $\zeta_p \notin F$, that F/F^+ is unramified at all finite places, and that $[F^+ : \mathbb{Q}]$ is even. Then there is an algebraic group $G_{/\mathcal{O}_{F^+}}$ as in Section 3.1 of [GK12]. This group is a quasisplit form of $U(2)$, and is compact mod centre at the infinite places. If v is a finite place of F^+ which splits at ww^c in F, then there is an isomorphism $\imath_w : G(\mathcal{O}_{F^+_v}) \xrightarrow{\sim} \mathrm{GL}_2(\mathcal{O}_{F_w})$ which extends to an isomorphism $G(F^+_v) \xrightarrow{\sim} \mathrm{GL}_2(F_w)$.

Fix an absolutely irreducible representation $\bar{r} : G_F \to \mathrm{GL}_2(\mathbb{F})$ with $\bar{r}^c \cong \bar{r}^\vee \overline{\varepsilon}^{-1}$. We assume that $\bar{r}$ is automorphic in the sense of Section 3.2 of [GK12], that $\bar{r}(G_{F(\zeta_p)})$ is adequate in the sense of [Tho12], and that if w is a finite place of F for which $\bar{r}|_{G_{F_w}}$ is ramified, then $w|_{F^+}$ splits in F.

Let S be a set of finite places of F^+ which split in F, and assume that S contains the places dividing p and the places at which $\overline{\rho}$ is ramified. For each place $v \in S$, choose a place $\widetilde{v}$ of F lying over v, and let $\widetilde{S}$ denote the set of these places. The local fields L_i in the definition of our patching functor will be the $F_{\widetilde{v}}$ for $v \in S$, the representations will be the $\overline{\rho}|_{G_{F_{\widetilde{v}}}}$, and in this section for consistency with [GK12] we will write the various compact groups that occur as U rather than K, and in particular we will write U_S for the group K in the definition of a patching functor.

By Lemma 4.11 of [DDT97] we can and do choose a finite place $v_1 \notin S$ which splits as $w_1 w_1^c$ in F, with the properties that

- $\mathbf{N}w_1 \not\equiv 1 \pmod p$,
- the ratio of the eigenvalues of $\bar{r}(\mathrm{Frob}_{w_1})$ is not equal to $(\mathbf{N}w_1)^{\pm 1}$, and
- the residue characteristic of w_1 is sufficiently large that for any non-trivial root of unity ζ in a quadratic extension of F, w_1 does not divide $\zeta + \zeta^{-1} - 2$.

We consider compact open subgroups $U = \prod_v U_v$ of $G(\mathbb{A}^\infty_{F^+})$ with the properties that

- $U_v \subset G(\mathcal{O}_{F^+_v})$ for all v which split in F;
- U_v is a hyperspecial maximal compact subgroup of $G(F^+_v)$ if v is inert in F;
- $U_v = G(\mathcal{O}_{F^+_v})$ if $v|p$ or $v \notin S \cup \{v_1\}$;

- U_{v_1} is the preimage of the upper triangular unipotent matrices under

$$G(\mathcal{O}_{F^+_{v_1}}) \overset{\iota_{w_1}}{\underset{\sim}{\longrightarrow}} \mathrm{GL}_2(\mathcal{O}_{F_{w_1}}) \to \mathrm{GL}_2(k_{v_1}).$$

In particular, the assumptions on v_1 and U_{v_1} imply that U is sufficiently small. Set $U_S = \prod_{v|p} U_v$. Then for any $\sigma \in \mathcal{C}$, we have a space of algebraic modular forms $S(U,\sigma^*)^*$ defined as in Section 3.1 of [GK12]. We have a Hecke algebra $\mathbb{T}^{S\cup\{v_1\},\mathrm{univ}}$ defined as in Section 3.2 of [GK12], with a maximal ideal $\mathfrak{m}$ corresponding to $\bar{r}$, and we set $M(\sigma) := S(U,\sigma^*)^*_{\mathfrak{m}}$.

Then in the same way as above, the Taylor–Wiles–Kisin patching method (as explained in Section 4.3 of [GK12]) allows us to patch the $M(\sigma)$ to obtain a covariant exact functor $\sigma \to M_\infty(\sigma)$ from the category of finite length objects of $\mathcal{C}$ to the category of R_∞-modules, where R_∞ is a power series ring over R. We extend this to a functor on all of $\mathcal{C}$ by setting $M_\infty(\sigma) = \varprojlim_n M_\infty(\sigma/\varpi_E^n\sigma)$, and regard $M_\infty(\sigma)$ as a coherent sheaf on $X_\infty = \operatorname{Spf} R_\infty$.

As in Section 6.4, it is easy to verify that this is a patching functor; in particular, the hypothesis on $M_\infty(\sigma^\circ(\tau))$ follows from local-global compatibility in the same way as in Section 6.4, as does that on $M_\infty(\overline{\sigma})$ (*cf.* the proof of Lemma 4.1.3 of [BLGG13a]).

7. Tame deformation spaces

The goal of this section is to describe the spaces of tamely potentially Barsotti–Tate lifts of a given generic Galois representation $\overline{\rho} : G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$, where a lift ρ of $\overline{\rho}$ is called *tamely potentially Barsotti–Tate* if $\overline{\rho}$ becomes Barsotti–Tate (that is, crystalline of Hodge type 0) over a tame extension of F_v. For such a lift, the inertial type τ underlying the potentially crystalline Dieudonné module of ρ will correspond via inertial local Langlands to a tame type $\sigma(\tau)$, i.e. an irreducible representation of $\mathrm{GL}_2(\mathbb{F}_q)$, and we restrict our attention to the case when $\sigma(\tau)$ is principal series or cuspidal. (Thus we omit the case when $\sigma(\tau)$ is one-dimensional, which is the case when a tame twist of ρ is actually Barsotti–Tate over F_v itself.) For each such τ, our goal then is to describe the framed deformation space $X(\tau)$ of potentially Barsotti–Tate liftings of type τ.

In the case when τ is a principal series and $\overline{\rho}$ has only scalar automorphisms, a precise calculation of $X(\tau)$ is given in [BM12, Thm. 5.2.1]. We will extend this calculation to the case of arbitrary generic $\overline{\rho}$, and also to the case of cuspidal τ. Furthermore, we will give an identification of the various components of the special fibre $\overline{X}(\tau)$ with the mod p reductions of crystalline deformation rings (thus providing a concrete geometric interpretation of the matching between components and Serre weights of $\overline{\rho}$ established in [BM12, Corollaire 5.2.2]).

Our proof will be in several stages. We will make use of some of the geometric Breuil–Mézard results of [EG14] (encapsulated in Theorem 7.1.1 below), and we will use base-change arguments to deduce our result for cuspidal types from the corresponding result for principal series types with a minimum of additional calculation. A key ingredient of our computation (as in the computation of [BM12]) is the fact that a potentially Barsotti–Tate lift of $\overline{\rho}$ arises from a uniquely determined strongly divisible module, and in fact we will follow [Sav05] in working with deformations of strongly divisible modules rather than directly with deformations of Galois representations. Thus, in addition to considering framed deformations (i.e. liftings) on the Galois side, we also consider framings of strongly divisible modules;

it is not difficult to transfer information obtained in terms of framings of strongly divisible modules back to the usual lifting spaces of Galois representations.

7.1. **Deformation spaces of Galois representations.** We fix a generic continuous representation $\overline{\rho} : G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$ and a non-scalar tame inertial type τ. As in Section 6.1, we let $X\big(\tau\big)$ denote the deformation space parameterising liftings of $\overline{\rho}$ that are potentially Barsotti–Tate of type τ. Let R^τ be the universal lifting ring for such deformations, so that $X\big(\tau\big)$ is the formal spectrum of R^τ. Fix a character $\psi : G_{F_v} \to \mathcal{O}^\times$ lifting $\overline{\varepsilon}\det\overline{\rho}$, with $\psi|_{I_{F_v}} = \det\tau$, and consider the subspace $X^\psi\big(\tau\big) = \mathrm{Spf}\, R^{\psi,\tau}$ corresponding to lifts with determinant $\psi\varepsilon^{-1}$. For each Serre weight $\overline{\sigma}$ we also have the spaces $X\big(\overline{\sigma}\big)$ and $X^\psi\big(\overline{\sigma}\big)$ for crystalline lifts of Hodge type $\overline{\sigma}$.

7.1.1. **Theorem.** *Assume $p > 2$. Let $\overline{\rho}$ be generic, and let τ be a non-scalar tame inertial type. Then the mod ϖ_E fibre of the deformation space $X\big(\tau\big)$ is the union of the mod ϖ_E fibres $\overline{X}\big(\overline{\sigma}\big)$ where $\overline{\sigma}$ runs over the Jordan–Hölder factors of $\overline{\sigma}(\tau)$. Furthermore, $\overline{X}\big(\overline{\sigma}\big)$ is non-empty if and only if $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$.*

The analogous statements also hold for $X^\psi\big(\tau\big)$.

Proof. The first part is a special case of [EG14, Thm. 5.5.4] (note that the assumption that all predicted weights of $\overline{\rho}$ are regular follows from Lemma 2.1.6, and the multiplicities n_a occurring in [EG14] are all zero or one by Lemma 3.1.1). The analogous results for $X^\psi\big(\tau\big)$ follow from those for $X\big(\tau\big)$, because twisting by the universal unramified deformation of the trivial character gives an isomorphism $R^\tau \cong R^{\psi,\tau}[\![X]\!]$, by Lemma 4.3.1 of [EG14]. $\square$

The rest of this section is devoted to proving a refinement of this result (Theorem 7.2.1 below) which gives explicit equations for the deformation rings involved; Theorem 7.1.1 will be a key ingredient in the proof of this refinement. Before we state this refinement, we need to introduce some terminology and notation.

7.2. **The structure of potentially Barsotti–Tate deformation spaces.** Let L be an unramified extension of F_v, of degree f' over $\mathbb{Q}_p$. Suppose that L embeds into the coefficient field E. (In applications, L will be either F_v or its quadratic unramified extension L_v, so this supposition will hold; recall that our E is always assumed to contain an embedding of L_v.) Fix a root π of the polynomial $E(u) = u^{p^{f'}-1} + p$.

A *Dieudonné $\mathcal{O}$-module* is a free $\mathcal{O}_L \otimes_{\mathbb{Z}_p} \mathcal{O}$-module of finite rank M together with an injective endomorphism $\varphi : M \to M$ which is $\mathcal{O}$-linear and $\mathcal{O}_L$-semilinear, and which satisfies $pM \subset \varphi(M)$. We say that M is a *Dieudonné $\mathcal{O}$-module with descent data* if it is equipped with an $\mathcal{O}_L \otimes_{\mathbb{Z}_p} \mathcal{O}$-linear action of $\mathrm{Gal}(L(\pi)/L)$ which commutes with φ.

Given a p-divisible group $\mathcal{H}$ over $\mathcal{O}_L[\pi]$ with an action of $\mathcal{O}$ and descent data on the generic fibre to L, the contravariant Dieudonné module corresponding to $\mathcal{H} \times_{\mathcal{O}} \mathbb{F}$ is a Dieudonné $\mathcal{O}$-module with descent data. In particular, an $\mathcal{O}$-point of $X\big(\tau\big)$ gives a Dieudonné $\mathcal{O}$-module with descent data of type τ.

The usual isomorphism $\mathcal{O}_L \otimes_{\mathbb{Z}_p} \mathcal{O} \xrightarrow{\sim} \mathcal{O} \times \cdots \times \mathcal{O}$ induces a decomposition $M = M^0 \times \cdots \times M^{f'-1}$ such that $\varphi(M^i) \subseteq M^{i+1}$. (We remark that we have used the notation of [Bre12] instead of the notation of [BM12], so that the index set $\mathcal{S}$ is a set of integers rather than the set of embeddings of k_v into $\mathbb{F}$; for a Dieudonné

$\mathcal{O}$-module $M = M^0 \times \cdots \times M^{f'-1}$, the piece that we denote M^j here corresponds to the piece that would be denoted M^{κ_j} in [BM12].)

Since $\mathrm{Gal}(L(\pi)/L)$ has order prime to p, it follows that the $\mathcal{O}$-linear action of $\mathrm{Gal}(L(\pi)/L)$ on each M^j is via a sum of characters, which is independent of j. We refer to this sum of characters as the *type* of M, and indeed in the case that M arises from a p-divisible group $\mathcal{H}$ over $\mathcal{O}_L[\pi]$ with descent data on the generic fibre, it has type τ if and only if the potentially Barsotti–Tate representation associated to $\mathcal{H}$ has type τ (regarded as a representation of I_{F_v}).

Let ι denote the involution $\iota(j) = f'-1-j$ on $\{0,\dots,f'-1\}$. Given a Dieudonné $\mathcal{O}$-module M with descent data of type $\eta\oplus\eta'$ with $\eta\neq\eta'$, we define the *gauge* (more specifically, the η-gauge and the η'-gauge) of M as follows: fix a basis e_j, e'_j of M^j on which $\mathrm{Gal}(L(\pi)/L)$ acts by η,η' respectively (such a basis is well-defined up to $\mathcal{O}^\times$-scalars). Then we have $\varphi(e_j) = x_j e_{j+1}, \varphi(e'_j) = x'_j e'_{j+1}$ for some x_j, x'_j that are well-defined up to $\mathcal{O}^\times$, and we say that the j-part of the η-gauge (respectively the j-part of the η'-gauge) of M is $x_{\iota(j)}\mathcal{O}^\times$ (respectively $x'_{\iota(j)}\mathcal{O}^\times$); *cf.* [Bre12, Eq. (13)]. We will also refer to $\{x_{\iota(j)}\mathcal{O}^\times\}_j$ (respectively $\{x'_{\iota(j)}\mathcal{O}^\times\}_j$) as the η-gauge of M (respectively, the η'-gauge).

For the rest of this section we will freely identify characters of $\mathcal{O}_L^\times$ and I_L via Art_L. The main theorem of this section is the following.

7.2.1. **Theorem.** *Assume that $p>3$. Continue to assume that $\overline{\rho}$ is generic and let τ be a non-scalar tame inertial type. If $\sigma(\tau)$ is a principal series type, then we write $\sigma(\tau)=\sigma(\eta\otimes\eta')$, and let $L=F_v$ and $f'=f$ in the above discussion. If $\sigma(\tau)=\Theta(\eta)$ is a cuspidal type, then we write $\eta'=\eta^q$, so that $\mathrm{BC}\,(\sigma(\tau))=\sigma(\eta\otimes\eta')$, and we let $L=L_v$ (the quadratic unramified extension of F_v) and $f'=2f$ in the above discussion. Recall that $\mathcal{S}=\{0,\dots,f-1\}$ in either case.*

1. *The deformation space $X(\tau)$ is non-empty if and only if $\overline{\sigma}_J(\tau)\in\mathcal{D}(\overline{\rho})$ for at least one $J\in\mathcal{P}_\tau$. If this is the case, then there are subsets $J_{\min}\subseteq J_{\max}\subseteq\mathcal{S}$ with $J_{\min},J_{\max}\in\mathcal{P}_\tau$ such that $\overline{\sigma}_J(\tau)\in\mathcal{D}(\overline{\rho})$ if and only if $J_{\min}\subseteq J\subseteq J_{\max}$. Moreover, all $J\subseteq\mathcal{S}$ with $J_{\min}\subseteq J\subseteq J_{\max}$ are in $\mathcal{P}_\tau$.*
2. *Assume that $X(\tau)$ is non-empty. If τ is cuspidal, assume further that $\mathbb{F}$ is large enough for $X(\tau)$ to have a point over the ring of integers of some finite extension of E with residue field $\mathbb{F}$. Then the deformation space $X(\tau)$ is equal to the formal spectrum of a power series ring over*
$$\mathcal{O}[\![(X_j,Y_j)_{j\in J_{\max}\setminus J_{\min}}]\!]/(X_jY_j-p)_{j\in J_{\max}\setminus J_{\min}}.$$
3. *If $\lambda: R^\tau\to\overline{\mathbb{Z}}_p$ is a point on $X(\tau)$, then for $j\in J_{\max}\setminus J_{\min}$ the j-part of the η'- and η-gauges of the Dieudonné module of the potentially Barsotti–Tate representation associated to λ are represented by $\lambda(X_j)$ and $\lambda(Y_j)$ respectively.*
4. *In the context of (2), the mod ϖ_E fibre $\overline{X}(\tau)$ is the union of the mod ϖ_E fibres $\overline{X}(\overline{\sigma}_J)$, for each J such that $J_{\min}\subseteq J\subseteq J_{\max}$. Precisely, the component $\overline{X}(\overline{\sigma}_J)$ of $\overline{X}(\tau)$ is cut out by the equations $X_j=0$ (for $j\in J\setminus J_{\min}$) and $Y_j=0$ (for $j\in J_{\max}\setminus J$).*
5. *The analogous results also hold for $X^\psi(\tau)$.*

7.2.2. *Remark.* Since twisting by the universal unramified deformation of the trivial character gives an isomorphism $R^\tau\cong R^{\psi,\tau}[\![X]\!]$, part (5) of Theorem 7.2.1 follows from the previous four parts, and we will ignore it in the below.

7.2.3. *Remark.* We comment briefly on the extra hypothesis on the residue field $\mathbb{F}$ in part (2) of Theorem 7.2.1. In the principal series case, the explicit calculations of [Bre12] show that our running hypotheses on $\mathbb{F}$ already guarantee the existence of a point on $X(\tau)$ as in 7.2.1(2), and so no additional hypothesis on $\mathbb{F}$ is necessary in this case. We expect that a similar explicit calculation in the cusipdal case would eliminate the extra hypothesis here. We note that this extra hypothesis is unimportant in our applications of Theorem 7.2.1 in Sections 8 and 10, where we are free to choose $\mathbb{F}$ to be arbitrarily large.

7.3. **Strongly divisible modules in the principal series case.** We begin the proof of Theorem 7.2.1 by briefly recalling the notion of a strongly divisible module with coefficients and tame descent data; for the full details of this theory, see Sections 3 and 4 of [Sav05]. Set $e = p^{f'}-1$, and let $S_{\mathcal{O}_L,\mathbb{Z}_p}$ be the p-adic completion of $\mathcal{O}_L[u, u^{ie}/i!]_{i\geq 0}$. Let R be a complete local noetherian $\mathcal{O}_L$-algebra, and let $S_{\mathcal{O}_L,R}$ be the $\mathfrak{m}_R$-adic completion of $S_{\mathcal{O}_L}\otimes_{\mathbb{Z}_p} R$. When L is clear from context we will abbreviate $S_{\mathcal{O}_L,\mathbb{Z}_p}$ and $S_{\mathcal{O}_L,R}$ as S and S_R, respectively. Define $\operatorname{Fil}^1 S_R$ to be the completion of the ideal of S_R generated by the elements $E(u)^i/i!$ for all $i\geq 1$.

Let φ denote the endomorphism of S_R that acts R-linearly and $\mathcal{O}_L$-Frobenius-semilinearly, with $\varphi(u)=u^p$. Write $c = p^{-1}\varphi(E(u))$. For each $g\in \operatorname{Gal}(L(\pi)/L)$ we let $\widehat{g}$ be the linear endomorphism of S_R with $g(u) = (g(\pi)/\pi)u$.

7.3.1. **Definition.** A *strongly divisible module with tame descent data from $L(\pi)$ to L and R-coefficients* is a finitely generated free S_R-module $\mathcal{M}$, together with a sub-S_R-module $\operatorname{Fil}^1\mathcal{M}$, a φ-semilinear map $\varphi_1 : \operatorname{Fil}^1\mathcal{M}\to\mathcal{M}$, and additive bijections $\widehat{g}:\mathcal{M}\to\mathcal{M}$ for each $g\in\operatorname{Gal}(L(\pi)/L)$, satisfying the following conditions:

(1) $\operatorname{Fil}^1\mathcal{M}$ contains $(\operatorname{Fil}^1 S_R)\mathcal{M}$,
(2) $\operatorname{Fil}^1\mathcal{M}\cap I\mathcal{M} = I\operatorname{Fil}^1\mathcal{M}$ for all ideals I in R,
(3) $\varphi_1(sx)=\varphi_1(s)\varphi(x)$ for $s\in\operatorname{Fil}^1 S_R$ and $x\in\mathcal{M}$, with $\varphi(x) := c^{-1}\varphi_1(E(u)x)$,
(4) $\varphi_1(\operatorname{Fil}^1\mathcal{M})$ is contained in $\mathcal{M}$ and generates it over S_R,
(5) $\widehat{g}(sx)=\widehat{g}(s)\widehat{g}(x)$ for all $s\in S_R$, $x\in\mathcal{M}$, $g\in\operatorname{Gal}(L(\pi)/L)$,
(6) $\widehat{g}_1\circ\widehat{g}_2=\widehat{g_1\circ g_2}$ for all $g_1,g_2\in\operatorname{Gal}(L(\pi)/L)$,
(7) $\widehat{g}(\operatorname{Fil}^1\mathcal{M})\subset\operatorname{Fil}^1\mathcal{M}$ for all $g\in\operatorname{Gal}(L(\pi)/L)$, and
(8) φ commutes with $\widehat{g}$ for all $g\in\operatorname{Gal}(L(\pi)/L)$.

This differs from the definition of strongly divisible modules with descent data and coefficients in Definition 4.1 of [Sav05] in the following respects. First, we have set $k=2$ and $F=L(\pi)$, $F'=L$ (in the notation of [Sav05]), and we have not required R to be flat over $\mathcal{O}$. Second, we have equipped each strongly divisible module with a map $\varphi_1 : \operatorname{Fil}^1\mathcal{M}\to\mathcal{M}$ rather than a map $\varphi:\mathcal{M}\to\mathcal{M}$; when R is p-torsion-free (as in [Sav05]) these notions are equivalent. Finally, we have ignored the monodromy operator N: we can do this because the operator N will exist and be unique for all strongly divisible modules that we consider. Existence follows by the same argument as in [Bre00, Prop 5.1.3(1)] and [Bre00, Lem. 2.1.1.9] after noting that $\operatorname{Fil}^1\mathcal{M}/(\operatorname{Fil}^1 S_R)\mathcal{M}$ is free over $R\otimes_{\mathbb{Z}_p}\mathcal{O}_L[\pi]$ for the strongly divisible modules $\mathcal{M}$ that we consider; uniqueness is a consequence of the usual argument as in [Bre00, Prop 5.1.3(1)].

Proposition 4.13 of [Sav05] and the remarks immediately preceding it (together with the remarks in the previous paragraph) show that if $R=\mathcal{O}$, the category of strongly divisible modules with tame descent data and R-coefficients is equivalent to the category of G_L-stable $\mathcal{O}$-lattices in E-representations of G_L which become Barsotti–Tate over $L(\pi)$.

In particular, if $\mathcal{H}$ is a p-divisible group over $\mathcal{O}_L[\pi]$ with an action of $\mathcal{O}$ and descent data on the generic fibre to L, then there is a strongly divisible $\mathcal{O}$-module with descent data $\mathcal{M}$ corresponding to $\mathcal{H}$, and thus to the (descended) generic fibre ρ of $\mathcal{H}$. We recall that $M \cong \mathcal{M} \otimes_S \mathcal{O}$ where the $\mathcal{O}$-algebra map $S \to \mathcal{O}$ sends the variable u and its divided powers to 0, and M is the contravariant Dieudonné module corresponding to $\mathcal{H} \times_{\mathcal{O}} \mathbb{F}$ discussed above. The usual isomorphism $\mathcal{O}_L \otimes_{\mathbb{Z}_p} \mathcal{O} \xrightarrow{\sim} \mathcal{O} \times \cdots \times \mathcal{O}$ induces a decomposition $\mathcal{M} = \mathcal{M}^0 \times \cdots \times \mathcal{M}^{f'-1}$, compatible with the decomposition of M defined above, and since $\mathrm{Gal}(L(\pi)/L)$ has order prime to p, it follows that the $\mathcal{O}$-linear action of $\mathrm{Gal}(L(\pi)/L)$ on each $\mathcal{M}^j$ again is via a sum of characters, which is independent of j.

The strongly divisible modules that arise in the two-dimensional principal series case are studied in detail in [Bre12, §§5–8] and [BM12, §5], and we now recall some of these results. We note that although [Bre12] and [BM12] assume that E contains a splitting field of the polynomial $u^{p^{2f}-1} + p$, this assumption is never used in their study of strongly divisible modules, and so our weaker hypothesis on E is sufficient. (In particular it is possible to take E to be unramified.)

Let $\eta \neq \eta'$ be characters of $\mathrm{Gal}(L(\pi)/L)$, and let $\tau = \eta \oplus \eta'$, a tame inertial principal series type. As mentioned above, we identify η, η' with characters of $\mathcal{O}_L^\times$, and we use the notation of Section 3.2, so that in particular we have an integer $c = \sum_{i=0}^{f-1} c_i p^i$. Write $c^{(j)} := \sum_{i=0}^{f-1} (p-1-c_{i-j}) p^i$. (Recall that ω_f agrees with the *reciprocal* of the composition of $g \mapsto g(\pi)/\pi$ with $\overline{\kappa}_0$; this is the reason for writing $p-1-c_{i-j}$ rather than c_{i-j} in the preceding definition.) We say that a strongly divisible module $\mathcal{M}$ with tame descent data and R-coefficients has *shape* τ if there is a decomposition $\mathcal{S} = I_\eta \coprod I_{\eta'} \coprod II$ such that $\mathcal{M}$ can be written in the following form.

$$j \in I_\eta : \quad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}^j & = \langle e_\eta^j + a_j u^{c^{(j)}} e_{\eta'}^j, (u^e + p) e_{\eta'}^j \rangle \\ \varphi_1(e_\eta^j + a_j u^{c^{(j)}} e_{\eta'}^j) & = e_\eta^{j+1} \\ \varphi_1((u^e + p) e_{\eta'}^j) & = e_{\eta'}^{j+1} \end{cases}$$

$$j \in I_{\eta'} : \quad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}^j & = \langle (u^e + p) e_\eta^j, e_{\eta'}^j + a_j u^{e-c^{(j)}} e_\eta^j \rangle \\ \varphi_1((u^e + p) e_\eta^j) & = e_\eta^{j+1} \\ \varphi_1(e_{\eta'}^j + a_j u^{e-c^{(j)}} e_\eta^j) & = e_{\eta'}^{j+1} \end{cases}$$

$$j \in II : \quad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}^j & = \langle a_j e_\eta^j + u^{c^{(j)}} e_{\eta'}^j, -b_j e_{\eta'}^j + u^{e-c^{(j)}} e_\eta^j \rangle \\ \varphi_1(a_j e_\eta^j + u^{c^{(j)}} e_{\eta'}^j) & = e_\eta^{j+1} \\ \varphi_1(-b_j e_{\eta'}^j + u^{e-c^{(j)}} e_\eta^j) & = e_{\eta'}^{j+1} \end{cases}$$

with $a_j \in R$ if $j \in I_\eta, I_{\eta'}$, and $a_j, b_j \in \mathfrak{m}_R$ if $j \in II$, with $a_j b_j = p$. When $j = f-1$ the right-hand side of the defining expressions for φ_1 must be modified to αe_η^0 and $\alpha' e_{\eta'}^0$ for some $\alpha, \alpha' \in R^\times$. Here $\widetilde{\mathrm{Fil}}^1 \mathcal{M}^j := \mathrm{Fil}^1 \mathcal{M}^j / (\mathrm{Fil}^p S) \mathcal{M}^j$, and the descent data acts on e_η^j, $e_{\eta'}^j$ via η, η' respectively. It is not difficult to check that an object of this form is indeed a strongly divisible module in the sense of Definition 7.3.1 (in particular that the condition (2) is satisfied). We refer to the elements e_η^j and $e_{\eta'}^j$ in the above shape as a *gauge basis* of $\mathcal{M}^j$.

Fix a generic representation $\overline{\rho}$. If ρ is a potentially Barsotti–Tate lift of $\overline{\rho}$ of type τ, and $\mathcal{M}$ is the corresponding strongly divisible module with $\mathcal{O}$-coefficients and descent data, then by [Bre12, Prop. 5.2, Prop. 7.1] we see that $\mathcal{M}$ has shape τ.

7.3.2. *Remark.* Recall that our conventions for Hodge–Tate weights differ from those of [Bre12]. In particular where [Bre12] associates to $\mathcal{M}$ the Galois representation $T^L_{\mathrm{st},2}(\mathcal{M})$ (where $T^L_{\mathrm{st},2}$ is the functor defined between Remark 4.8 and Lemma 4.9 of [Sav05]), we instead have $\rho = T^L_{\mathrm{st},2}(\mathcal{M})(-1)$. That is, our representations ρ and $\overline{\rho}$ differ from those of [Bre12] by a twist by ε^{-1}.

Note that this twist preserves genericity, even recalling that our fundamental characters are the reciprocals of the ones in [Bre12]. In particular, items from *op. cit.* such as [Bre12, Prop. 7.1], which concern intrinsic statements about the shape of strongly divisible modules of type τ under the hypothesis that $\overline{\rho}$ is generic, will remain true in our setting, without any changes.

By [BM12, Thm. 5.1.1] the sets I_η, $I_{\eta'}$, II, and the indices j such that $a_j \in \mathcal{O}^\times$ are completely determined by $\overline{\rho}$ and the type τ; note that in light of Remark 7.3.2 this statement remains true in our conventions. Similarly we remark that [Bre12, Thm. 8.1] shows that we have $\overline{\rho}$ irreducible if and only if $|II|$ is odd and $\mathrm{val}_p(a_j) > 0$ for all j, non-split reducible if and only if some a_j is a unit, and split if and only if $|II|$ is even and $\mathrm{val}_p(a_j) > 0$ for all j.

7.4. **Deformation spaces of strongly divisible modules.** We continue to assume that τ is a principal series type. If $\mathcal{M}$ is a strongly divisible module with R-coefficients of shape τ, then the gauge bases $e^j_\eta, e^j_{\eta'}$ are determined by $\mathcal{M}$ up to scalar multiplication. (This is [Bre12, Rem. 5.5(iv)] and [Bre12, Prop. 5.4] in the case $R = \mathcal{O}$, but the argument goes over unchanged to the general case.) Once a choice of gauge basis $e^0_\eta, e^0_{\eta'}$ is fixed, then the gauge bases $e^j_\eta, e^j_{\eta'}$ and the a_j, b_j and α, α' are also uniquely determined.

We now consider a further three deformation problems. Fix $\overline{\rho}$ and τ such that $\overline{\rho}$ has a potentially Barsotti–Tate lift of type τ. There is a unique strongly divisible module $\overline{\mathcal{M}}$ (with $\mathbb{F}$-coefficients, of shape τ) that occurs as $\mathcal{M}/\mathfrak{m}\mathcal{M}$ as $\mathcal{M}$ varies over all strongly divisible modules with $\mathcal{O}$-coefficients of shape τ corresponding to potentially Barsotti–Tate lifts of $\overline{\rho}$ of type τ. (This is clear when $\overline{\rho}$ is reducible and split, because $\overline{a}_j = 0$ for all j, while the parameters $\overline{\alpha}, \overline{\alpha}'$ are determined by the unramified parts of the characters comprising $\overline{\rho}$ [Sav08, Ex. 3.7]; when $\overline{\rho}$ has scalar endomorphisms, the claim follows from the proof of [BM12, Prop. 5.1.2].) Fix a choice of the gauge basis elements $\overline{e}^0_\eta, \overline{e}^0_{\eta'}$ for $\overline{\mathcal{M}}^0$. We then let $R^\tau_{\overline{\mathcal{M}}}$ be the complete local noetherian $\mathcal{O}$-algebra representing the functor which assigns to a complete local noetherian $\mathcal{O}$-algebra R the set of isomorphism classes of tuples $(\mathcal{M}, e^0_\eta, e^0_{\eta'})$ where $\mathcal{M}$ is a strongly divisible module with R-coefficients of shape τ lifting $\overline{\mathcal{M}}$, and $e^0_\eta, e^0_{\eta'}$ are a choice of gauge basis lifting $\overline{e}^0_\eta, \overline{e}^0_{\eta'}$. We write $X_{\overline{\mathcal{M}}}(\tau) := \mathrm{Spf}\, R^\tau_{\overline{\mathcal{M}}}$.

Given a strongly divisible module $\mathcal{M}$ with R-coefficients of shape τ, we obtain a Galois representation $\rho_{\mathcal{M}} : G_L \to \mathrm{GL}_2(R)$ as follows. If R is Artinian, then we set $\rho_{\mathcal{M}} := T^L_{\mathrm{st},2}(\mathcal{M})(-1)$, with $T^L_{\mathrm{st},2}$ as in Remark 7.3.2. In general, we set $\rho_{\mathcal{M}} := \varprojlim_n \rho_{\mathcal{M}/\mathfrak{m}^n_R}$.

We let $R^{\square,\tau}_{\overline{\mathcal{M}}}$ denote the complete local noetherian $\mathcal{O}$-algebra representing the functor which assigns to a complete local noetherian $\mathcal{O}$-algebra R the set of isomorphism classes of tuples $(\mathcal{M}, e^0_\eta, e^0_{\eta'}, \rho)$, where $(\mathcal{M}, e^0_\eta, e^0_{\eta'})$ is as above, and ρ is a lifting of $\overline{\rho}$ such that $\rho \cong \rho_{\mathcal{M}}$ (i.e. the same data as that parameterised by $R^{\tau}_{\overline{\mathcal{M}}}$, together with a framing of $\rho_{\mathcal{M}}$).

We also consider $R^{\square}_{\overline{\mathcal{M}}}$, the complete local noetherian $\mathcal{O}$-algebra representing the functor which assigns to a complete local noetherian $\mathcal{O}$-algebra R the set of isomorphism classes of pairs $(\mathcal{M}, \rho)$, with $\mathcal{M}$ and ρ as in the previous paragraph.

Now, note that $R^{\square,\tau}_{\overline{\mathcal{M}}}$ is formally smooth over both $R^{\tau}_{\overline{\mathcal{M}}}$ and $R^{\square}_{\overline{\mathcal{M}}}$ (the additional data being a choice of framing of $\overline{\rho}_{\mathcal{M}}$ in the first case, and a choice of gauge basis $e^0_{\eta}, e^0_{\eta'}$ in the second case). In addition, we claim that there is a natural isomorphism $R^{\tau} \xrightarrow{\sim} R^{\square}_{\overline{\mathcal{M}}}$, induced by the forgetful functor $(\mathcal{M}, \rho) \mapsto \rho$. To see this, we argue as in the proof of Théorème 5.2.1 of [BM12]. This morphism certainly induces an isomorphism on $\mathcal{O}_{E'}$-points for each finite extension E'/E (this is just the statement recalled above, that the strongly divisible module associated to a potentially Barsotti–Tate lifting of $\overline{\rho}$ of type τ necessarily has shape τ), so it suffices to show that it induces a surjection on reduced tangent spaces. This reduces to a straightforward explicit calculation with filtered φ_1-modules, exactly as in the last paragraph of the proof of Théorème 5.2.1 of [BM12].

Thus parts (1) and (2) of Theorem 7.2.1 hold in the principal series case if and only if the analogous statements hold with $X(\tau)$ replaced with the space $X_{\overline{\mathcal{M}}}(\tau)$. Similarly we will be able to deduce parts (3) and (4) of Theorem 7.2.1 in the principal series case from calculations on $X_{\overline{\mathcal{M}}}(\tau)$ together with the description of the relationship between $X(\tau)$ and $X_{\overline{\mathcal{M}}}(\tau)$. For these reasons, in the principal series case we will work with $X_{\overline{\mathcal{M}}}(\tau)$ in place of $X(\tau)$ from now on.

We will prove Theorem 7.2.1 via a series of lemmas.

7.4.1. **Lemma.** *If $\sigma(\tau)$ is a principal series type, then parts (1), (2), and (3) of Theorem 7.2.1 hold, with $J_{\min} = \iota(I_{\eta'})$ and $J_{\max} \setminus J_{\min} = \iota(II)$ for the sets $I_{\eta'}, II$ determined by $\overline{\mathcal{M}}$.*

Proof. Parts (2) and (3) are straightforward from the form of a strongly divisible module of shape τ. Define $J_{\min} = \iota(I_{\eta'})$ and $J_{\max} \setminus J_{\min} = \iota(II)$, so we may set $X_j = a_{\iota(j)}, Y_j = b_{\iota(j)}$ for each $j \in J_{\max} \setminus J_{\min}$, and the unnamed formal variables correspond to the other values of a_j, and to α, α' (or more precisely to $a_j - [\overline{a}_j]$, where the brackets denote Teichmüller lift, and similarly for α, α'). Recall we have adopted the indexing from [Bre12] rather than the indexing from [BM12]; this explains the presence of the involution ι. The argument in part (3) is completed by noting that if we begin with a $\overline{\mathbb{Z}}_p$-point $\lambda \in X(\tau)$ and pass to a point $\lambda' \in X_{\overline{\mathcal{M}}}(\tau)$ by lifting to $\mathrm{Spf}(R^{\square,\tau}_{\overline{\mathcal{M}}})$ via formal smoothness and then projecting, then λ' corresponds to the strongly divisible module attached to the potentially Barsotti–Tate representation associated to λ.

Part (1) will follow from Proposition 4.3 of [Bre12] but requires a comparison of the conventions of our paper with the conventions of [Bre12]. Note that from [Bre12, Eq. (26)] the sets $J_{\min}, J_{\max}$ defined above are the *complements* of the sets $J_{\max}, J_{\min}$ from [Bre12]. By [Bre12, Thm. 8.1], the Serre weights of $\overline{\rho}(1)$ that are Jordan–Hölder factors of the representations that Breuil calls $\overline{\sigma}(\chi^s)$ (and using the definition of being modular of some weight from [Bre12]) are exactly the weights that Breuil calls $\overline{\sigma}_J$, for $J^c_{\max} \subseteq J \subseteq J^c_{\min}$ (in our conventions for these sets). We translate to our conventions. Note that because our normalization of the inertial local Langlands correspondence is opposite to that of [Bre12], the representation denoted $\overline{\sigma}(\chi^s)$ in *op. cit.* is our representation $\overline{\sigma}(\eta^{-1}\otimes(\eta')^{-1})$, or in other words our $\overline{\sigma}(\eta'\otimes\eta)\otimes\det(\tau)^{-1}$. In particular Breuil's $\overline{\sigma}_J$ is our $\overline{\sigma}(\tau)_{J^c}\otimes\det(\tau)^{-1}$.

On the other hand $\overline{\rho}(1)\cong\overline{\rho}^\vee\otimes\det(\tau)$. Now $\overline{\rho}$ has Serre weight σ (in our conventions for being modular of some weight) if and only if $\overline{\rho}^\vee$ has Serre weight σ (in Breuil's conventions for being modular of some weight), if and only if $\overline{\rho}(1)$ has Serre weight $\sigma\otimes\det(\tau)^{-1}$ (in Breuil's conventions for being modular of some weight, but our conventions for local Langlands). It follows that the Serre weights of $\overline{\rho}$ that are Jordan–Hölder factors of $\overline{\sigma}(\tau)$ are precisely the $\overline{\sigma}(\tau)_{J^c}$ for $J_{\max}^c\subseteq J\subseteq J_{\min}^c$, or in other words the $\overline{\sigma}(\tau)_J$ for $J_{\min}\subseteq J\subseteq J_{\max}$. ☐

Note that the proof of Lemma 7.4.1 gives a construction of the universal strongly divisible module on $X_{\overline{\mathcal{M}}}(\tau)$ in the same spirit as the construction (for $\overline{\rho}$ not split-reducible) of the universal strongly divisible module on $X(\tau)$ in the proof of [BM12, Thm. 5.2.1].

7.4.2. **Lemma.** *If $\sigma(\tau)$ is a principal series type, and $\overline{\rho}$ is reducible, then part (4) of Theorem 7.2.1 holds.*

Proof. Fix a set $J_{\min}\subseteq J\subseteq J_{\max}$. Write $\mathfrak{m}^\tau_{\overline{\mathcal{M}}}$ for the maximal ideal of $R^\tau_{\overline{\mathcal{M}}}$, and let $I_J\subseteq R^\tau_{\overline{\mathcal{M}}}$ be the ideal generated by ϖ_E and $(\mathfrak{m}^\tau_{\overline{\mathcal{M}}})^2$, together with the variables X_j for $j\in J\setminus J_{\min}$, Y_j for $j\in J_{\max}\setminus J$, and all of the unnamed extra power series variables of Theorem 7.2.1(2). Let $\mathbb{F}_J$ be the quotient $R^\tau_{\overline{\mathcal{M}}}/I_J$, so that $\mathbb{F}_J$ is an Artinian local ring, and write $\overline{\rho}_J$ for the deformation of $\overline{\rho}$ corresponding to the natural quotient map $R^\tau_{\overline{\mathcal{M}}}\to\mathbb{F}_J$. The deformation $\overline{\rho}_J$ lies on the component of $\overline{X}_{\overline{\mathcal{M}}}(\tau)$ cut out by the equations $X_j=0$ (for $j\in J\setminus J_{\min}$) and $Y_j=0$ (for $j\in J_{\max}\setminus J$), and only that component, and so by Theorem 7.1.1 (and from the relationship between $X(\tau)$ and $X_{\overline{\mathcal{M}}}(\tau)$) it suffices to show that $\overline{\rho}_J$ also lies on $\overline{X}(\overline{\sigma}_J)$.

We will carry this out by showing that a twist of $\overline{\rho}_J$ is the generic fibre of a suitable Fontaine–Laffaille module with $\mathbb{F}_J$-coefficients. We refer the reader to Appendix A for a discussion of what we need about (unipotent) Fontaine–Laffaille theory, and in particular for some of the notation used in this argument.

Fix a compatible system of p^nth roots of $-p$ in $\overline{F}_v$ and let $F_{v,\infty}$ be the extension of F_v that they generate. Since $\overline{\rho}$ is generic, the following lemma shows that both $\overline{\rho}_J$ and $\overline{\rho}$ are uniquely determined by their restrictions to $\mathrm{Gal}(\overline{F}_v/F_{v,\infty})$, and so in the remainder of the proof it suffices to consider these restrictions. (We thank Fred Diamond for simplifying our original proof of the following lemma.)

7.4.3. **Lemma.** *Suppose that A is a local Artin $\mathbb{F}_p$-algebra with maximal ideal $\mathfrak{m}_A$, and $\chi: G_{F_v}\to A^\times$ is a character such that $\chi \pmod{\mathfrak{m}_A}$ is neither trivial nor cyclotomic. Then the restriction map $\mathrm{res}: H^1(G_{F_v},\chi)\to H^1(G_{F_{v,\infty}},\chi)$ is injective.*

Proof. We note that our proof will work for any finite extension $F_v/\mathbb{Q}_p$ (not just unramified F_v). Write $\overline{\chi}$ for $\chi \pmod{\mathfrak{m}_A}$, and let $\widehat{F}_v$ be the Galois closure of $F_{v,\infty}$. By inflation-restriction, it suffices to show that $H^1(\mathrm{Gal}(\widehat{F}_v/F_v),\chi^{G_{\widehat{F}_v}})=0$. Since $\chi^{G_{\widehat{F}_v}}$ is a successive extension of copies of $\overline{\chi}^{G_{\widehat{F}_v}}$, by dévissage it suffices to show that $H^1(\mathrm{Gal}(\widehat{F}_v/F_v),\overline{\chi}^{G_{\widehat{F}_v}})=0$.

If $\overline{\chi}^{G_{\widehat{F}_v}}\neq 0$, then $\overline{\chi}$ is trivial on $G_{\widehat{F}_v}$ (so $\overline{\chi}^{G_{\widehat{F}_v}}=\overline{\chi}$), and so also on $G=\mathrm{Gal}(\overline{F}_v/F_v(\mu_{p^\infty}))$. Another application of inflation-restriction gives

$$1\to H^1(\mathrm{Gal}(F_v(\mu_{p^\infty})/F_v),\overline{\chi})\to H^1(\mathrm{Gal}(\widehat{F}_v/F_v),\overline{\chi})\to H^1(G,\overline{\chi})^{\mathrm{Gal}(F_v(\mu_{p^\infty})/F_v)}.$$

Since $G \cong \mathbb{Z}_p(1)$, the rightmost term is $\mathrm{Hom}_{\mathrm{Gal}(F_v(\mu_{p^\infty})/F_v)}(\mathbb{Z}_p(1), \overline{\chi})$, and this is nontrivial if and only if $\overline{\chi}$ is cyclotomic. As for the leftmost term, one more application of inflation-restriction shows that $H^1(\mathrm{Gal}(F_v(\mu_{p^\infty})/F_v), \overline{\chi})$ injects into $H^1(\mathrm{Gal}(F_v(\mu_{p^\infty})/F_v(\mu_p)), \overline{\chi})^{\mathrm{Gal}(F_v(\mu_p)/F_v)} \cong \mathrm{Hom}_{\mathrm{Gal}(F_v(\mu_p)/F_v)}(\mathbb{Z}_p, \overline{\chi})$, and this is non-zero if and only if $\overline{\chi}$ is trivial. The lemma follows. $\square$

We resume the proof of Lemma 7.4.2. From the construction of the universal strongly divisible module on $X_{\overline{\mathcal{M}}}(\tau)$ described above, the representation $\overline{\rho}_J$ is the generic fibre of a strongly divisible module $\mathcal{M}_J = \mathcal{M}_J^0 \times \cdots \times \mathcal{M}_J^{f-1}$ of shape τ with $\mathbb{F}_J$-coefficients and descent data. In particular the structure of $\mathcal{M}_J$ is described as follows.

$$j \in I_\eta : \qquad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}_J^j & = \langle e_\eta^j + a_j u^{c(j)} e_{\eta'}^j, u^e e_{\eta'}^j \rangle \\ \varphi_1(e_\eta^j + a_j u^{c(j)} e_{\eta'}^j) & = e_\eta^{j+1} \\ \varphi_1(u^e e_{\eta'}^j) & = e_{\eta'}^{j+1} \end{cases}$$

$$j \in I_{\eta'} : \qquad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}_J^j & = \langle u^e e_\eta^j, e_{\eta'}^j + a_j u^{e-c(j)} e_\eta^j \rangle \\ \varphi_1(u^e e_\eta^j) & = e_\eta^{j+1} \\ \varphi_1(e_{\eta'}^j + a_j u^{e-c(j)} e_\eta^j) & = e_{\eta'}^{j+1} \end{cases}$$

$$j \in II_X : \qquad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}_J^j & = \langle X_{\iota(j)} e_\eta^j + u^{c(j)} e_{\eta'}^j, u^{e-c(j)} e_\eta^j \rangle \\ \varphi_1(X_{\iota(j)} e_\eta^j + u^{c(j)} e_{\eta'}^j) & = e_\eta^{j+1} \\ \varphi_1(u^{e-c(j)} e_\eta^j) & = e_{\eta'}^{j+1} \end{cases}$$

$$j \in II_Y : \qquad \begin{cases} \widetilde{\mathrm{Fil}}^1 \mathcal{M}_J^j & = \langle u^{c(j)} e_{\eta'}^j, -Y_{\iota(j)} e_{\eta'}^j + u^{e-c(j)} e_\eta^j \rangle \\ \varphi_1(u^{c(j)} e_{\eta'}^j) & = e_\eta^{j+1} \\ \varphi_1(-Y_{\iota(j)} e_{\eta'}^j + u^{e-c(j)} e_\eta^j) & = e_{\eta'}^{j+1} \end{cases}$$

Here our notation is as in the definition of "shape τ" in 7.3.1, except that we have decomposed II as $II_X \coprod II_Y$ with $II_X = \iota(J_{\max} \setminus J)$ and $II_Y = \iota(J \setminus J_{\min})$. Recall also that by Lemma 7.4.1 we have $J_{\min} = \iota(I_{\eta'})$ and $\mathcal{S} \setminus J_{\max} = \iota(I_\eta)$. As usual, when $j = f-1$ the right-hand side of the defining expressions for φ_1 should be modified to αe_η^0 and $\alpha' e_{\eta'}^0$. For the remainder of this argument it is convenient for us to define $e_\eta^f := \alpha e_\eta^0$ and $e_{\eta'}^f := \alpha' e_{\eta'}^0$ so that from now on we can ignore this last complication.

The generic fibre of the strongly divisible module $\mathcal{M}_J$ is also the generic fibre of a certain étale φ-module; we show this following the method in the proof of Proposition 7.3 of [Bre12]. The first step is to note that by [Bre12, Prop. A.2(i)] the object $\mathcal{M}_J$ comes from a unique φ-module $\mathfrak{M}_J$ of type $\overline{\chi}$ over $k_v \otimes_{\mathbb{F}_p} \mathbb{F}[\![u]\!]$ (see the discussion preceding *loc. cit.*, which in particular gives the recipe for recovering $\mathcal{M}_J$ from $\mathfrak{M}$). By functoriality $\mathfrak{M}_J$ has an action of $\mathbb{F}_J$. Set $\mathfrak{D}_J = \mathfrak{M}_J[1/u]$; this is a φ-module of type χ in the sense of [Bre12, Déf. A1]. There is a decomposition $\mathfrak{D}_J = \mathfrak{D}_J^0 \times \cdots \times \mathfrak{D}_J^{f-1}$ where $\mathfrak{D}_J^j = \mathbb{F}_J((u))\mathfrak{e}_\eta^j \oplus \mathbb{F}_J((u))\mathfrak{e}_{\eta'}^j$ and $\mathrm{Gal}(F_{v,\infty}(\sqrt[e]{-p})/F_{v,\infty})$ acts on $\mathfrak{e}_\eta^j, \mathfrak{e}_{\eta'}^j$ via $\overline{\eta}, \overline{\eta}'$ respectively, and one calculates that

$$j \in I_\eta : \qquad \begin{cases} \varphi(\mathfrak{e}_\eta^{j-1}) & = u^e \mathfrak{e}_\eta^j - a_j u^{c(j)} \mathfrak{e}_{\eta'}^j \\ \varphi(\mathfrak{e}_{\eta'}^{j-1}) & = \mathfrak{e}_{\eta'}^j \end{cases}$$

$$j \in I_{\eta'}: \qquad \begin{cases} \varphi(\mathfrak{e}_{\eta}^{j-1}) &= \mathfrak{e}_{\eta}^{j} \\ \varphi(\mathfrak{e}_{\eta'}^{j-1}) &= u^{e}\mathfrak{e}_{\eta'}^{j} - a_j u^{e-c^{(j)}}\mathfrak{e}_{\eta}^{j} \end{cases}$$

$$j \in II_X: \qquad \begin{cases} \varphi(\mathfrak{e}_{\eta}^{j-1}) &= u^{c^{(j)}}\mathfrak{e}_{\eta'}^{j} \\ \varphi(\mathfrak{e}_{\eta'}^{j-1}) &= -X_j\mathfrak{e}_{\eta'}^{j} + u^{e-c^{(j)}}\mathfrak{e}_{\eta}^{j} \end{cases}$$

$$j \in II_Y: \qquad \begin{cases} \varphi(\mathfrak{e}_{\eta}^{j-1}) &= u^{c^{(j)}}\mathfrak{e}_{\eta'}^{j} + Y_j\mathfrak{e}_{\eta}^{j} \\ \varphi(\mathfrak{e}_{\eta'}^{j-1}) &= u^{e-c^{(j)}}\mathfrak{e}_{\eta}^{j}. \end{cases}$$

with a suitable modification when $j = f-1$. (To check this, it suffices to define $\mathfrak{M}_J$ by the same formulas, and then follow the recipe for recovering $\mathcal{M}_J$ from $\mathfrak{M}_J$.)

By [Bre12, Prop. A.2(ii)] the generic fibre of $\mathfrak{D}_J$ is $\overline{\rho}_J^{\vee}|_{\mathrm{Gal}(\overline{F}_v/F_{v,\infty})}$. Set $\mathfrak{e}^j = \mathfrak{e}_{\eta}^j$ and $\mathfrak{f}^j = u^{c^{(j)}}\mathfrak{e}_{\eta'}^j$; one checks that with respect to the basis $\mathfrak{e}^j, \mathfrak{f}^j$, the action of φ involves only powers of u^e; it follows from the isomorphism [Bre12, Eq. (47)] that replacing u^e with u gives a φ-module $\prod_j \mathbb{F}_J((u))\mathfrak{e}^j \oplus \mathbb{F}_J((u))\mathfrak{f}^j$ without descent data whose generic fibre is $(\overline{\rho}_J^{\vee} \otimes \overline{\eta})|_{\mathrm{Gal}(\overline{F}_v/F_{v,\infty})}$; the action of φ is given by

$$j \in I_{\eta}: \qquad \begin{cases} \varphi(\mathfrak{e}^{j-1}) &= u\mathfrak{e}^{j} - a_j\mathfrak{f}^{j} \\ \varphi(\mathfrak{f}^{j-1}) &= u^{p-1-c_{f-j}}\mathfrak{f}^{j} \end{cases}$$

$$j \in I_{\eta'}: \qquad \begin{cases} \varphi(\mathfrak{e}^{j-1}) &= \mathfrak{e}^{j} \\ \varphi(\mathfrak{f}^{j-1}) &= u^{p-c_{f-j}}(\mathfrak{f}^{j} - a_j\mathfrak{e}^{j}) \end{cases}$$

$$j \in II_X: \qquad \begin{cases} \varphi(\mathfrak{e}^{j-1}) &= \mathfrak{f}^{j} \\ \varphi(\mathfrak{f}^{j-1}) &= -X_j u^{p-1-c_{f-j}}\mathfrak{f}^{j} + u^{p-c_{f-j}}\mathfrak{e}^{j} \end{cases}$$

$$j \in II_Y: \qquad \begin{cases} \varphi(\mathfrak{e}^{j-1}) &= \mathfrak{f}^{j} + Y_j\mathfrak{e}^{j} \\ \varphi(\mathfrak{f}^{j-1}) &= u^{p-c_{f-j}}\mathfrak{e}^{j}. \end{cases}$$

Define

$$v_{f-j} = \begin{cases} p-1 & \text{if } f-j \in J \\ c_{f-j} & \text{if } f-j \notin J,\ f-j-1 \notin J \\ c_{f-j}-1 & \text{if } f-j \notin J,\ f-j-1 \in J. \end{cases}$$

Note that v_{f-j} has been chosen so that $v_i = s_{J,i} + t_{J,i}$ with $s_{J,i}, t_{J,i}$ as in Section 3.2. Write $\widetilde{\omega}_f$ for the extension of ω_f to $\mathrm{Gal}(\overline{F}_v/F_v)$ given by composing $g \mapsto (g(\pi)/\pi)^{-1}$ with $\overline{\kappa}_0$. We twist by $\widetilde{\omega}_f^{-\sum_{j=0}^{f-1} v_j p^j}$; as in the proof of [Bre12, Prop. 7.3], this has the effect of multiplying $\varphi(\mathfrak{e}^{j-1})$ and $\varphi(\mathfrak{f}^{j-1})$ by $u^{v_{f-j}}$ for all j. Make the change of basis $(\frac{1}{u}\mathfrak{e}^{j-1}, \frac{1}{u}\mathfrak{f}^{j-1})$ when $f-j \in J$ and $(\mathfrak{e}^{j-1}, \frac{1}{u}\mathfrak{f}^{j-1})$ when $f-j \notin J$, and let $\mathfrak{M}$ be the resulting φ-module.

Now define a Fontaine–Laffaille module $M = M^0 \times \cdots \times M^{f-1}$ with $M^j = \mathbb{F}_J e^j \oplus \mathbb{F}_J f^j$ and

$$j \in I_{\eta},\ f-j \in J: \qquad \begin{cases} \varphi(e^{j}) &= e^{j+1} - a_j f^{j+1} \\ \varphi_{p-c_{f-j}-1}(f^{j}) &= f^{j+1} \end{cases}$$

$$j \in I_{\eta},\ f-j \notin J: \qquad \begin{cases} \varphi_{c_{f-j}+1}(e^{j}) &= e^{j+1} - a_j f^{j+1} \\ \varphi(f^{j}) &= f^{j+1} \end{cases}$$

$$j \in I_{\eta'},\ f-j \in J: \quad \begin{cases} \varphi(e^j) & = e^{j+1} \\ \varphi_{p-c_{f-j}}(f^j) & = f^{j+1} - a_j e^{j+1} \end{cases}$$

$$j \in I_{\eta'},\ f-j \notin J: \quad \begin{cases} \varphi_{c_{f-j}}(e^j) & = e^{j+1} \\ \varphi(f^j) & = f^{j+1} - a_j e^{j+1} \end{cases}$$

$$j \in II_X,\ f-j \in J: \quad \begin{cases} \varphi(e^j) & = f^{j+1} \\ \varphi_{p-c_{f-j}-1}(f^j) & = -X_j f^{j+1} + e^{j+1} \end{cases}$$

$$j \in II_X,\ f-j \notin J: \quad \begin{cases} \varphi_{c_{f-j}+1}(e^j) & = f^{j+1} \\ \varphi(f^j) & = -X_j f^{j+1} + e^{j+1} \end{cases}$$

$$j \in II_Y,\ f-j \in J: \quad \begin{cases} \varphi(e^j) & = f^{j+1} + Y_j e^{j+1} \\ \varphi_{p-c_{f-j}}(f^j) & = e^{j+1} \end{cases}$$

$$j \in II_Y,\ f-j \notin J: \quad \begin{cases} \varphi_{c_{f-j}}(e^j) & = f^{j+1} + Y_j e^{j+1} \\ \varphi(f^j) & = e^{j+1}. \end{cases}$$

It is easy to check that M and $\mathfrak{M}$ are unipotent in the sense of Definitions A.1.3 and A.1.7 respectively, and that $\Theta_{p-1}(\mathfrak{M}) \cong \mathcal{F}_{p-1}(M)$ (see Appendix A for the functors $\Theta_{p-1}, \mathcal{F}_{p-1}$). By Theorem A.3.1 we see that $\Theta_{p-1}(\mathfrak{M}) \cong \mathcal{F}_{p-1}(M)$ is unipotent in the sense of Definition A.1.5, and so combining Propositions A.3.2 and A.3.3 shows that

$$T(M)|_{\mathrm{Gal}(\overline{F}_v/F_{v,\infty})} \cong T(\mathfrak{M}) \cong (\overline{\rho}_J^\vee \otimes \overline{\eta}\widetilde{\omega}_f^{-\sum_{j=0}^{f-1} v_j p^j})|_{\mathrm{Gal}(\overline{F}_v/F_{v,\infty})}.$$

Recalling that $f-j-1 \in J$ if and only if $j \in I_{\eta'} \cup II_Y$, we see that M corresponds to a point on the deformation space of crystalline liftings of $\overline{\rho}^\vee \otimes \overline{\eta}\widetilde{\omega}_f^{-\sum_{j=0}^{f-1} v_j p^j}$ of Hodge type $(-\vec{s}-\vec{1}, \vec{0})$ where

$$s_j = s_{J,j} = \begin{cases} p-1-c_j-\delta_{J^c}(j-1) & \text{if } j \in J \\ c_j - \delta_J(j-1) & \text{if } j \notin J. \end{cases}$$

Note that $\overline{\rho}^\vee \otimes \overline{\eta}\overline{\eta}'\overline{\varepsilon}^{-1} \simeq \overline{\rho}$. Similarly, since $\overline{\rho}_J$ lies on $X_{\overline{\mathcal{M}}}(\tau)$ we must have $\overline{\rho}_J^\vee \otimes \overline{\eta}\overline{\eta}'\overline{\varepsilon}^{-1}\nu \simeq \overline{\rho}_J$ for some unramified character ν. We conclude that

$$\overline{\rho}_J \simeq \overline{\rho}_J^\vee \otimes \overline{\eta}\overline{\eta}'\overline{\varepsilon}^{-1}\nu \simeq (\overline{\rho}_J^\vee \otimes \overline{\eta}\widetilde{\omega}_f^{-\sum_{j=0}^{f-1} v_j p^j}) \otimes \widetilde{\omega}_f^{\sum_{j=0}^{f-1} v_j p^j}\overline{\eta}'\overline{\varepsilon}^{-1}\nu$$

comes from a point on the deformation space of crystalline liftings of $\overline{\rho}$ of Hodge type $\overline{\sigma}_{\vec{v}-\vec{s},\vec{s}} \otimes (\overline{\eta}' \circ \det) = \overline{\sigma}_{\vec{t},\vec{s}} \otimes (\overline{\eta}' \circ \det) = \overline{\sigma}_J(\tau)$. □

Proof of Theorem 7.2.1. We use the theory of base change explained in Section 3. The first sentence of Theorem 7.2.1(1) is a consequence of Theorem 7.1.1, so we may assume throughout the argument that $X(\tau)$ is non-empty. Recall that L_v denotes the quadratic unramified extension of F_v. Then $\tau' := \mathrm{BC}(\tau)$ is a principal series representation, and $\overline{\rho}' := \overline{\rho}|_{G_{L_v}}$ is reducible, so Theorem 7.2.1 holds for $\overline{\rho}'$ and τ' by Lemmas 7.4.1 and 7.4.2.

We take $L = L_v$, $f' = 2f$ from now on. Given a strongly divisible module (with tame descent data and coefficients) $\mathcal{M}'$ such that $\mathcal{M}'$ has shape τ', we define another strongly divisible module $c(\mathcal{M}')$ of shape τ' in the following way. Let φ^f denote the Frobenius on L/F_v. Then we take $(c(\mathcal{M}'), \mathrm{Fil}^1 c(\mathcal{M}')) = (\mathcal{M}', \mathrm{Fil}^1 \mathcal{M}')$ as sets and as strongly divisible $S_{\mathcal{O}_{F_v},R}$-modules; we let $\mathcal{O}_L$ act on $c(\mathcal{M}')$ through φ^f; and the additive bijection $\widehat{g} : c(\mathcal{M}') \to c(\mathcal{M}')$ is given by $\widehat{\varphi^f g \varphi^f} : \mathcal{M}' \to \mathcal{M}'$ for each $g \in \mathrm{Gal}(L(\pi)/L)$. (In particular the roles of η, η' are exchanged if τ is cuspidal, and preserved if τ is principal series.) The identity map on underlying sets gives a φ^f-semilinear map $c_{\mathcal{M}'} : \mathcal{M}' \to c(\mathcal{M}')$ and a natural identification $(\mathcal{M}')^j \xrightarrow{\sim} c(\mathcal{M}')^{j+f}$. Furthermore, there is a natural isomorphism $\rho_{c(\mathcal{M}')} \cong \rho_{\mathcal{M}'}^{\varphi^f}$, where $\rho_{\mathcal{M}'}^{\varphi^f}$ is the conjugate of $\rho_{\mathcal{M}'}$ by the generator φ^f of the Galois group $\mathrm{Gal}(L_v/F_v)$. (To see this last statement, it suffices to consider the case that R is Artinian. In this case $\rho_{\mathcal{M}'} = T^L_{\mathrm{st},2}(\mathcal{M}')(-1) = \mathrm{Hom}_{\phi,\mathrm{Fil}^1}(\mathcal{M}', \hat{A}_{\mathrm{st},\infty})^\vee$, and it is easily checked that the map

$$\mathrm{Hom}_{\phi,\mathrm{Fil}^1}(\mathcal{M}', \hat{A}_{\mathrm{st},\infty}) \to \mathrm{Hom}_{\phi,\mathrm{Fil}^1}(c(\mathcal{M}'), \hat{A}_{\mathrm{st},\infty})$$

sending $h \mapsto \varphi^f \circ h \circ c_{\mathcal{M}'}^{-1}$ (where ϕ is the Frobenius on $\hat{A}_{\mathrm{st},\infty}$ as in Section 2.2.2 of [Bre99a]) is Galois equivariant after conjugating the left hand side by φ^f.)

Write $\overline{\mathcal{M}}'$ for the strongly divisible module (with $\mathbb{F}$-coefficients and descent data) of shape τ' corresponding to $\overline{\rho}'$. In the principal series case, note that $\overline{\mathcal{M}}' = l_v \otimes_{k_v} \overline{\mathcal{M}}$, with the action of $\mathrm{Gal}(F_v(\pi)/F_v)$ on $\overline{\mathcal{M}}$ extended l_v-linearly to give the action of $\mathrm{Gal}(L(\pi)/L)$ on $\overline{\mathcal{M}}'$. In particular there is a canonical $S_{\mathcal{O}_L,R}$-linear isomorphism $\gamma : \overline{\mathcal{M}}' \xrightarrow{\sim} c(\overline{\mathcal{M}}')$ given by the generator of $\mathrm{Gal}(l_v/k_v)$ on the first factor of $\overline{\mathcal{M}}' = l_v \otimes_{k_v} \overline{\mathcal{M}}$ and the identity on the second factor. In the cuspidal case we no longer *a priori* have an $\overline{\mathcal{M}}$, but there still exists a canonical isomorphism $\gamma : \overline{\mathcal{M}}' \xrightarrow{\sim} c(\overline{\mathcal{M}}')$ as above. By hypothesis there exists an $\mathcal{O}'$-point $x \in X(\tau)$ for the ring of integers $\mathcal{O}'$ in some finite extension E'/E with residue field $\mathbb{F}$. Let $x' \in X(\tau')$ be the point arising from x by restriction to G_{L_v}. If $\mathcal{M}, \mathcal{M}'$ are the strongly divisible modules corresponding to x, x', then $\mathcal{M}'$ is obtained from $\mathcal{M}$ by restricting the descent data. In particular there is an involutive isomorphism $\tilde{\gamma} : \mathcal{M}' \xrightarrow{\sim} c(\mathcal{M}')$ coming from the action on $\mathcal{M}$ of the element of $\mathrm{Gal}(L_v(\pi)/F_v)$ lifting φ^f and fixing π. Then γ is the reduction of this map modulo $\mathfrak{m}_{E'}$. In any case we see that $c_{\overline{\mathcal{M}}'}\gamma^{-1}c_{\overline{\mathcal{M}}'} = \gamma$. For the remainder of this argument we write c for $c_{\mathcal{M}'}$ and $\bar{c}$ for $c_{\overline{\mathcal{M}}'}$.

We have the universal deformation rings $R^{\tau'}_{\overline{\mathcal{M}}'}$ and $R^{\square,\tau'}_{\overline{\mathcal{M}}'}$ considered in Section 7.4, which depended on our choice of gauge basis elements $\bar{e}^0_\eta, \bar{e}^0_{\eta'}$ for $(\overline{\mathcal{M}}')^0$. Then we let $R^{\tau',f}_{\overline{\mathcal{M}}'}$ be the complete local noetherian $\mathcal{O}$-algebra representing the functor which assigns to a complete local noetherian $\mathcal{O}$-algebra R the set of isomorphism classes of tuples $(\mathcal{M}', e^0_\eta, e^0_{\eta'}, \epsilon^f_\eta, \epsilon^f_{\eta'})$, where $\mathcal{M}'$, $e^0_\eta, e^0_{\eta'}$ are as above, and $\epsilon^f_\eta, \epsilon^f_{\eta'}$ are a choice of gauge basis for $(\mathcal{M}')^f$ lifting $\{(\gamma^{-1}\bar{c})(\bar{e}^0_\eta), (\gamma^{-1}\bar{c})(\bar{e}^0_{\eta'})\}$. We write $X_{\overline{\mathcal{M}}',f}(\tau') := \mathrm{Spf}\, R^{\tau',f}_{\overline{\mathcal{M}}}$. Similarly, $R^{\square,\tau',f}_{\overline{\mathcal{M}}'}$ represents the functor assigning to R the set of isomorphism classes of tuples $(\mathcal{M}', e^0_\eta, e^0_{\eta'}, \epsilon^f_\eta, \epsilon^f_{\eta'}, \rho')$ with ρ' a lifting of $\overline{\rho}'$ such that $\rho' \cong \rho_{\mathcal{M}'}$.

There is a natural involutive action of φ^f on each of $R^{\tau',f}_{\overline{\mathcal{M}}'}$, $R^{\square,\tau',f}_{\overline{\mathcal{M}}'}$, $R^{\square}_{\overline{\mathcal{M}}'}$, and $R^{\tau'}$ by sending $(\mathcal{M}', \{e^0_\eta, e^0_{\eta'}\}, \{\epsilon^f_\eta, \epsilon^f_{\eta'}\})$ to $(c(\mathcal{M}'), \{c(\epsilon^f_\eta), c(\epsilon^f_{\eta'})\}, \{c(e^0_\eta), c(e^0_{\eta'})\})$ and conjugating ρ' by φ^f. (Note that the latter will indeed define a point of each of the deformation spaces we are considering, after identifying $(c(\overline{\mathcal{M}}'), \{\gamma(\bar{e}^0_\eta), \gamma(\bar{e}^0_{\eta'})\})$ with $(\overline{\mathcal{M}}', \{\bar{e}^0_\eta, \bar{e}^0_{\eta'}\})$ via γ^{-1}.)

Similarly if X, X' are the universal framed deformation spaces of $\overline{\rho}$ and $\overline{\rho}|_{G_{L_v}}$, then we have a natural map $X \to (X')^{\varphi^f}$ which is easily seen to be an isomorphism since $p > 2$. Since the extension L_v/F_v is unramified, we see that a lift ρ of $\overline{\rho}$ is potentially Barsotti–Tate of type τ if and only if $\rho' := \rho|_{G_{L_v}}$ is potentially Barsotti–Tate of type $\mathrm{BC}(\tau)$, so that furthermore $X(\tau) = X(\tau')^{\varphi^f}$. It is easy to see that $(\mathrm{Spf}\, R^{\square,\tau',f}_{\overline{\mathcal{M}'}})^{\varphi^f}$ is formally smooth over both $X(\tau')^{\varphi^f}$ and $X_{\overline{\mathcal{M}'},f}(\tau')^{\varphi^f}$. Furthermore, in the principal series case the Dieudonné module with descent data M' and strongly divisible module with descent data $\mathcal{M}'$ corresponding to ρ' are obtained from those corresponding to ρ by tensoring with $\mathcal{O}_{L_v}$ over $\mathcal{O}_{F_v}$ and extending φ semilinearly; in the cuspidal case, the Dieudonné module for ρ' is equal to the one for ρ, and the strongly divisible module is obtained by restricting the descent data from $\mathrm{Gal}(L_v(\pi)/F_v)$ to $\mathrm{Gal}(L_v(\pi)/L_v)$. Consequently, we see that in order to establish parts (1) and (2) of Theorem 7.2.1, it is enough to prove the analogous statements for $X_{\overline{\mathcal{M}'},f}(\tau')^{\varphi^f}$; parts (3) and (4) will also follow from considerations on this latter space.

Choose a gauge basis $\bar{e}^0_\eta, \bar{e}^0_{\eta'}$ for $(\overline{\mathcal{M}'})^0$, and recursively define $\bar{e}^j_\eta, \bar{e}^j_{\eta'}$ for $0 < j < 2f$ by the formulas of Section 7.3. Define $\bar{\epsilon}^{j+f}_\eta$ and $\bar{\epsilon}^{j+f}_{\eta'}$ to be $(\gamma^{-1}\bar{c})(\bar{e}^j_\eta)$ and $(\gamma^{-1}\bar{c})(\bar{e}^j_{\eta'})$ respectively if τ is principal series, and non-respectively if τ is cuspidal.

For $0 \le j < f$ we write $(\overline{\mathcal{M}'})^j$ in terms of the gauge basis $\bar{e}^j_\eta, \bar{e}^j_{\eta'}$ as in Section 7.3 (denoting the structure constants in the type I filtrations as $\bar{a}_j$; note that the structure constants in type II filtrations are 0 here). For $f \le j < 2f$ we instead write $(\overline{\mathcal{M}'})^j$ in terms of the gauge basis $\bar{\epsilon}^j_\eta, \bar{\epsilon}^j_{\eta'}$ (again denoting the structure constants as $\bar{a}_j$). In view of the fact that $\gamma : \overline{\mathcal{M}'} \to c(\overline{\mathcal{M}'})$ is an isomorphism, we see that the sets $I_\eta, I_{\eta'}, II$ for $\overline{\mathcal{M}'}$ are invariant under translation by f if τ is principal series, and II is similarly invariant under translation by f if τ is cuspidal while $I_\eta, I_{\eta'}$ are interchanged; that $\bar{a}_{j+f} = \bar{a}_j$ for all j; and that in this basis, if the right-hand side of the defining expressions for φ_1 for $j = 2f-1$ are written as $\bar{\beta}\bar{e}^0_\eta$ and $\bar{\beta}'\bar{e}^0_{\eta'}$, then the right-hand side of the defining expressions for φ_1 for $j = f-1$ can be written as $\bar{\beta}\bar{\epsilon}^f_\eta$ and $\bar{\beta}'\bar{\epsilon}^f_{\eta'}$.

Consider an arbitrary point $x = (\mathcal{M}', \{e^0_\eta, e^0_{\eta'}\}, \{\epsilon^f_\eta, \epsilon^f_{\eta'}\})$ of $X_{\overline{\mathcal{M}'},f}(\tau')$. Recursively define $e^j_\eta, e^j_{\eta'}$ for $0 \le j < f$ and $\epsilon^j_\eta, \epsilon^j_{\eta'}$ for $f \le j < 2f$ by the formulas of Section 7.3. As we did with $\overline{\mathcal{M}'}$, we write $\mathcal{M}'$ in terms of the gauge basis $e^j_\eta, e^j_{\eta'}$ for $0 \le j < f$ (denoting the structure constants in the filtration as a_j, b_j), and in terms of the gauge basis $\epsilon^j_\eta, \epsilon^j_{\eta'}$ for $f \le j < 2f$ (here denoting the structure constants as a'_j, b'_j). In this basis, let the right-hand side of the defining expressions for φ_1 for $j = f-1$ and $2f-1$ be $\beta\epsilon^f_\eta, \beta'\epsilon^f_{\eta'}$ and $\tilde{\beta}e^0_\eta, \tilde{\beta}'e^0_{\eta'}$ respectively. In particular $\beta, \tilde{\beta}$ are both lifts of $\bar{\beta}$, and similarly $\beta', \tilde{\beta}'$ are both lifts of $\bar{\beta}'$.

Then $\varphi^f(x) = x$ if and only if there is an isomorphism $c(\mathcal{M}') \xrightarrow{\sim} \mathcal{M}'$ lifting γ^{-1} and sending $\{c(\epsilon^f_\eta), c(\epsilon^f_{\eta'})\}$, $\{c(e^0_\eta), c(e^0_{\eta'})\}$ to $\{e^0_\eta, e^0_{\eta'}\}$, $\{\epsilon^f_\eta, \epsilon^f_{\eta'}\}$ respectively. It follows immediately from the above description that this is the case precisely when $\beta = \tilde{\beta}$, $\beta' = \tilde{\beta}'$, $a_j = a'_{j+f}$ for all $0 \le j < f$, and $b_j = b'_{j+f}$ for all $j \in II \cap \{0, \dots, f-1\}$, except that if τ is cuspidal we have $a_j = -b'_{j+f}$ and $b_j = -a'_{j+f}$ for all $j \in II \cap \{0, \dots, f-1\}$.

Now rewrite $\mathcal{M}'$ entirely in terms of the gauge bases $e_\eta^j, e_{\eta'}^j$ for $0 \le j < 2f$, as in Section 7.3. By the change-of-basis formulae in the proof of [Bre12, Prop. 5.4] we obtain $a_{j+f} = a_j(\beta/\beta')^{\gamma_j}$ for some $\gamma_j \in \{\pm 1\}$ for all $0 \le j < f$, except that again if τ is cuspidal and $j \in II \cap \{0, \dots, f-1\}$ we have $a_j = -b_{j+f}(\beta/\beta')^{\gamma_j}$ and $b_j = -a_{j+f}(\beta/\beta')^{-\gamma_j}$. (The exponents γ_j are determined by the sets $I_\eta, I_{\eta'}, II$, and so by $\overline{\rho}$ and τ.) Furthermore we have α, α' equal to $\beta^2, (\beta')^2$ if $II \cap \{0, \dots, f-1\}$ is even, and $\alpha = \alpha' = \beta\beta'$ if $II \cap \{0, \dots, f-1\}$ is odd.

For the remainder of this proof we will use ι for the involution $\iota(j) = f-1-j$ on $\{0, \dots, f-1\}$, and ι' for the analogous involution on $\{0, \dots, f'-1\}$. Write $J_0 = \iota(II \cap \{0, \dots, f-1\})$. It follows easily from the previous paragraph and from the description of $X_{\overline{\mathcal{M}'},f}(\tau')$ given by Lemma 7.4.1 that $X_{\overline{\mathcal{M}'},f}(\tau')^{\varphi^f}$ is the formal spectrum of a ring R equal to a power series ring over $\mathcal{O}[\![(X_j, Y_j)_{j \in J_0}]\!]/(X_jY_j - p)_{j \in J_0}$, where the variables X_j, Y_j for $j \in J_0$ are taken to be the restrictions to $X_{\overline{\mathcal{M}'},f}(\tau')^{\varphi^f}$ of the variables X_j, Y_j on $X_{\overline{\mathcal{M}'},f}(\tau')$. (Note that when $|J_0|$ is even we use the hypothesis that $p \ne 2$, so that $(1+Z)^{1/2} \in \mathcal{O}[\![Z]\!]$.) This gives part (2) of the Theorem.

Part (3) now follows by an argument similar to that in the proof of Lemma 7.4.1 (if a $\overline{\mathbb{Z}}_p$-point $\lambda \in X(\tau)$ leads to a point $\lambda' \in X_{\overline{\mathcal{M}'},f}(\tau')^{\varphi^f}$, and if λ corresponds to the potentially Barsotti–Tate representation ρ, then λ' corresponds to the strongly divisible module associated to $\rho|_{G_L}$) together with the behavior of η- and η'-gauges under base change; for the latter one considers the principal series and cuspidal cases separately.

Now let $J'_{\min}$ and $J'_{\max}$ be the sets from part (1) of the Theorem applied to $\overline{\rho}'$ and τ', so that $J'_{\min} = \iota'(I_{\eta'})$ and $J'_{\max} = \iota'(I_{\eta'} \cup II)$. Define $J_{\min} = J'_{\min} \cap \{0, \dots, f-1\}$ and $J_{\max} = J'_{\max} \cap \{0, \dots, f-1\}$. Observe that the set J_0 from the previous paragraph is precisely $J_{\max} \setminus J_{\min}$.

Write J' for $\mathrm{BC}_{\mathrm{cusp}}(J)$ or $\mathrm{BC}_{\mathrm{PS}}(J)$ depending on whether τ is a cuspidal or principal series inertial type; it is easy to verify that if $J \subseteq \mathcal{S}$ then $J_{\min} \subseteq J \subseteq J_{\max}$ if and only if $J'_{\min} \subseteq J' \subseteq J'_{\max}$.

Now, by Lemma 3.4.1 and the definition of $\mathcal{D}(\overline{\rho})$, we see that if $\overline{\sigma}_J(\tau) \in \mathcal{D}(\overline{\rho})$ then $\overline{\sigma}_{J'}(\tau') \in \mathcal{D}(\overline{\rho}')$. In particular, we have $J'_{\min} \subseteq J' \subseteq J'_{\max}$, or equivalently $J_{\min} \subseteq J \subseteq J_{\max}$. Conversely, we claim that if J is such that $J'_{\min} \subseteq J' \subseteq J'_{\max}$, then $\overline{\sigma}_J(\tau) \in \mathcal{D}(\overline{\rho})$. In order to see this, we need to check that $\overline{\rho}$ has a crystalline lift of Hodge type $\overline{\sigma}_J(\tau)$. In addition, we need to verify the assertion of Theorem 7.2.1(4), by identifying explicit equations for $\overline{X}(\overline{\sigma}_J(\tau))$.

In fact, we can verify these conditions simultaneously, as follows: let $\overline{X}_J$ be the component of $\overline{X}(\overline{\tau})$ cut out by the equations $X_j = 0$ (for $j \in J \setminus J_{\min}$) and $Y_j = 0$ (for $j \in J_{\max} \setminus J$). From Lemma 7.4.2, from the relationship between $X(\tau')$ and $X_{\overline{\mathcal{M}'}}(\tau')$, and by the discussion above, we see that $\overline{X}_J = \overline{X}(\overline{\sigma}_{J'}(\tau'))^{\varphi^f}$. On the other hand, by Theorem 7.1.1 we know that we must have $\overline{X}_J = \overline{X}(\overline{\sigma}(\tau)_K)$ for some K, and we then have $\overline{X}_J \subseteq \overline{X}(\mathrm{BC}\,(\overline{\sigma}(\tau)_K))^{\varphi^f}$ by definition. By Lemma 3.4.1, this means that $\overline{X}_J \subseteq \overline{X}(\overline{\sigma}(\tau')_{K'})^{\varphi^f}$, so that $K' = J'$, and $K = J$, as required. $\square$

We note the following useful corollary of the proof of Theorem 7.2.1.

7.4.4. **Corollary.** *Assume that $\overline{\rho}$ is generic, and that $\overline{\sigma}$ is a Serre weight. Then $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$ if and only if* $\mathrm{BC}\,(\overline{\sigma}) \in \mathcal{D}(\overline{\rho}|_{G_{L_v}})$.

7.4.5. *Remark.* Observe that in the principal series case we have given *two* constructions of the space $X(\tau)$: once in Lemma 7.4.1 (i.e. once from [BM12, Thm. 5.2.1]) and then again by the base change argument in the proof of Theorem 7.2.1 (i.e. by applying Lemma 7.4.1 to the case of principal series types over the quadratic extension). We caution that we have not checked that these two constructions give precisely the same formal variables X_j, Y_j (rather than differing by a unit).

8. Breuil's lattice conjecture

We now use the results of the preceding sections to prove the main conjecture of [Bre12], as well as its natural analogue for cuspidal representations.

8.1. **Patched modules of lattices.** Assume that $p > 3$. Let F_v be an unramified extension of $\mathbb{Q}_p$, let $\overline{\rho} : G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$ be a generic representation, and let M_∞ be a (possibly fixed determinant) patching functor indexed by a set of places and representations which includes $(F_v, \overline{\rho})$.

Let τ be a non-scalar tame inertial type for G_{F_v}, which we assume is either principal series or regular cuspidal. Recall from Section 5.2 that for each subset $J \in \mathcal{P}_\tau$ we have fixed a lattice $\sigma^\circ_J(\tau)$ in $\sigma(\tau)$ with the property that the cosocle of $\overline{\sigma}^\circ_J(\tau)$ is precisely $\overline{\sigma}_J(\tau)$. We have also fixed a subset $J_{\mathrm{base}} \subseteq \mathcal{S}$ and defined the involution $\imath(J) := J \triangle J_{\mathrm{base}}$ on the subsets of $\mathcal{S}$; if τ is principal series, then $J_{\mathrm{base}} = \varnothing$ and $\imath$ is just the identity.

We assume that we have fixed inertial types at all places in the indexing set for M_∞ other than v, and we drop these places from the notation from now on, so that, if L denotes the tensor product of the fixed inertial types at all places in the indexing set other than v, then we simply write (for example) $M_\infty(\sigma^\circ_J(\tau))$, rather than $M_\infty(\sigma^\circ_J(\tau) \otimes L)$.

It follows from Theorem 5.2.4 that if $\imath(J) \in \mathcal{P}_\tau$ then we have inclusions

$$p^{|J|}\sigma^\circ_{\imath(J)}(\tau) \subseteq \sigma^\circ_{\imath(\varnothing)}(\tau) \subseteq \sigma^\circ_{\imath(J)}(\tau),$$

and therefore inclusions

$$p^{|J|}M_\infty(\sigma^\circ_{\imath(J)}(\tau)) \subseteq M_\infty(\sigma^\circ_{\imath(\varnothing)}(\tau)) \subseteq M_\infty(\sigma^\circ_{\imath(J)}(\tau)).$$

Since $p > 3$ and $\overline{\rho}$ is assumed to be generic, we can apply Theorem 7.2.1, so we have subsets $J_{\min} \subseteq J_{\max} \subseteq \mathcal{S}$, and elements X_j, Y_j of $R_v^{\square,\tau}$ for each $j \in J_{\max} \setminus J_{\min}$, which we regard as elements of R^τ_∞ via the natural map $R_v^{\square,\tau} \to R^\tau_\infty$. It will be convenient for us to also consider the subsets

$$J'_{\min} := (J^c_{\mathrm{base}} \cap J_{\min}) \cup (J_{\mathrm{base}} \cap J^c_{\max}),\ J'_{\max} := (J^c_{\mathrm{base}} \cap J_{\max}) \cup (J_{\mathrm{base}} \cap J^c_{\min}).$$

Note that we have $J_{\min} \subseteq \imath(J) \subseteq J_{\max}$ if and only if $J'_{\min} \subseteq J \subseteq J'_{\max}$, and that if τ is principal series, then we have $J'_{\min} = J_{\min}, J'_{\max} = J_{\max}$. Observe also that in all cases we have $J'_{\max} \setminus J'_{\min} = J_{\max} \setminus J_{\min}$.

For each $j \in \mathcal{S}$, we define an element $\varpi_j \in R^\tau_\infty$ according to the prescription

$$\varpi_j := \begin{cases} p & \text{if } j \in J'_{\min} \\ Y_j & \text{if } j \in (J'_{\max} \setminus J'_{\min}) \cap J_{\mathrm{base}} \\ X_j & \text{if } j \in (J'_{\max} \setminus J'_{\min}) \cap J^c_{\mathrm{base}} \\ 1 & \text{if } j \notin J'_{\max}. \end{cases}$$

If $J \in \imath(\mathcal{P}_\tau)$, then we define $\varpi_J := \prod_{j \in J} \varpi_j$.

Similarly, for each $j \in \mathcal{S}$ we define $\varpi'_j \in R^\tau_\infty$ according to the prescription

$$\varpi'_j := \begin{cases} 1 & \text{if } j \in J'_{\mathrm{min}} \\ X_j & \text{if } j \in (J'_{\mathrm{max}} \setminus J'_{\mathrm{min}}) \cap J_{\mathrm{base}} \\ Y_j & \text{if } j \in (J'_{\mathrm{max}} \setminus J'_{\mathrm{min}}) \cap J^c_{\mathrm{base}} \\ p & \text{if } j \notin J'_{\mathrm{max}}, \end{cases}$$

and if $J \in \imath(\mathcal{P}_\tau)$, then we define $\varpi'_J := \prod_{j \in J} \varpi'_j$. Note that $\varpi_J \varpi'_J = p^{|J|}$.

8.1.1. **Proposition.** *Suppose that τ is either a principal series inertial type or a regular cuspidal inertial type. For each $J \in \imath(\mathcal{P}_\tau)$, we have equalities*

$$\varpi_J M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big) = M_\infty\big(\sigma^\circ_{\imath(\varnothing)}(\tau)\big)$$

and

$$\varpi'_J M_\infty\big(\sigma^\circ_{\imath(\varnothing)}(\tau)\big) = p^{|J|} M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big).$$

Proof. To begin with, fix J and $j \in J^c$ with the property that $\imath(J), \imath(J \cup \{j\}) \in \mathcal{P}_\tau$. By Theorem 5.2.4(4) there is an inclusion of lattices $\sigma^\circ_{\imath(J)}(\tau) \subseteq \sigma^\circ_{\imath(J\cup\{j\})}(\tau)$, and the cokernel of this inclusion is a successive extension of the weights $\overline{\sigma}_{\imath(J')}(\tau)$, where J' runs over subsets of $\mathcal{S}$ such that $j \in J'$ and $\imath(J') \in \mathcal{P}_\tau$. Note also that this cokernel is annihilated by p.

The cokernel $M_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)/\sigma^\circ_{\imath(J)}(\tau)\big)$ of

$$M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big) \hookrightarrow M_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)\big) \tag{8.1}$$

is thus a successive extension of the corresponding modules $M_\infty(\overline{\sigma}_{\imath(J')}(\tau))$, and is annihilated by p, and so its scheme-theoretic support is generically reduced, with underlying reduced subscheme equal to $\bigcup_{\substack{J'_{\mathrm{min}} \subseteq J' \subseteq J'_{\mathrm{max}} \\ j \in J'}} \overline{X}\big(\overline{\sigma}_{\imath(J')}(\tau)\big)$. Since the patched module $M_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)/\sigma^\circ_{\imath(J)}(\tau)\big)$ is Cohen–Macaulay, by the definition of a patching functor, it follows from [Mat89, Thm. 17.3(i)] that its scheme-theoretic support has no embedded associated primes, and so is in fact equal to

$$\bigcup_{\substack{J'_{\mathrm{min}} \subseteq J' \subseteq J'_{\mathrm{max}} \\ j \in J'}} \overline{X}\big(\overline{\sigma}_{\imath(J')}(\tau)\big).$$

If $j \notin J'_{\mathrm{max}}$, we see that this support is trivial, and hence this cokernel is trivial, i.e. is annihilated by multiplication by 1. If $j \in J'_{\mathrm{max}} \setminus J'_{\mathrm{min}}$, then by Theorem 7.2.1(4) this cokernel is annihilated by multiplication by Y_j if $j \in J_{\mathrm{base}}$ and by X_j if $j \in J^c_{\mathrm{base}}$. If $j \in J'_{\mathrm{min}}$, then this cokernel is annihilated by p. Thus in all cases, we find that the cokernel of (8.1) is annihilated by ϖ_j.

Similarly, we deduce that the cokernel of $pM_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)\big) \hookrightarrow M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big)$ is annihilated by ϖ'_j, since the cokernel of the inclusion $p\sigma^\circ_{\imath(J\cup\{j\})}(\tau) \subseteq \sigma^\circ_{\imath(J)}(\tau)$ is a successive extension of the weights $\overline{\sigma}_{\imath(J')}(\tau)$, where J' runs over the subsets of $\mathcal{S}$ such that $j \notin J'$.

We conclude that there are inclusions

$$\varpi_j M_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)\big) \subseteq M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big) \tag{8.2}$$

and

$$\varpi'_j M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big) \subseteq pM_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)\big). \tag{8.3}$$

Multiplying (8.2) by ϖ_j', we find that also

$$pM_\infty\big(\sigma^\circ_{\imath(J\cup\{j\})}(\tau)\big) \subseteq \varpi_j' M_\infty\big(\sigma^\circ_{\imath(J)}(\tau)\big),$$

and hence that (8.3) is in fact an equality. Multiplying this equality by ϖ_j (and recalling that $M_\infty\big(\sigma^\circ(\tau)\big)$ is p-torsion free), we deduce that (8.2) is also an equality. The statement of the Proposition now follows by induction on $|J|$, using Lemma 5.2.1. ☐

8.2. **Breuil's lattice conjecture.** We now apply Proposition 8.1.1 to the particular fixed determinant patching functors constructed in Section 6.2. (Of course, it could be applied in a similar way to the unitary group patching functors constructed in Section 6.6.) Continue to fix a prime $p > 3$, let F be a totally real field in which p is unramified, and fix a continuous representation $\overline{\rho} : G_F \to \mathrm{GL}_2(\mathbb{F})$ with the property that $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible. If $p = 5$, assume further that the projective image of $\overline{\rho}|_{G_{F(\zeta_5)}}$ is not isomorphic to A_5. Fix a place $v|p$, and assume that $\overline{\rho}|_{G_{F_v}}$ is generic. Let τ be some tame inertial type.

We now adopt the notation of Sections 6.2 and 6.3, fixing in particular a quaternion algebra D, a finite set of finite places S, and the maximal ideal $\mathfrak{m} \subset \mathbb{T}^{S,\mathrm{univ}}$ corresponding to $\overline{\rho}$. For each place $w \in S \setminus \{v\}$, choose an inertial type τ_w and a lattice $\sigma^\circ(\tau_w)$ in $\sigma(\tau_w)$, and let $\sigma^v := \otimes_{w\in S\setminus\{v\}}\sigma^\circ(\tau_w)$ be the corresponding representation of $K^v = \prod_{w\in S\setminus\{v\}} K_w$. Let $\lambda : \mathbb{T}_\mathfrak{m} \to E'$ be the system of Hecke eigenvalues corresponding to some minimal prime ideal of $\mathbb{T}_\mathfrak{m}$, where $\mathbb{T}_\mathfrak{m} := \mathbb{T}(\sigma^v\sigma^\circ(\tau))_\mathfrak{m}$ in the notation of Section 6.3 (note that this is independent of the choice of lattice $\sigma^\circ(\tau)$), and where E' is a finite extension of E. Set $K_v = \mathrm{GL}_2(\mathcal{O}_{F_v})$. Let $\pi := (M^v(\sigma^v)^*[1/p])[\lambda]$, which by strong multiplicity one for $D^\times$ is an irreducible tame representation of $\mathrm{GL}_2(F_v)$. Then the eigenspace $M^v(\sigma^v)^*[\lambda]$ is a lattice in π, and the intersection

$$\sigma^\circ(\pi) := (\sigma(\tau)\otimes_E E') \cap M^v(\sigma^v)^*[\lambda]$$

inside π defines a K_v-stable $\mathcal{O}_{E'}$-lattice in $\sigma(\tau)\otimes_E E'$.

Let $\rho : G_F \to \mathrm{GL}_2(\mathcal{O}_{E'})$ be the Galois representation associated to π. We now define a $\mathrm{GL}_2(\mathcal{O}_{F_v})$-stable $\mathcal{O}_{E'}$-lattice $\sigma^\circ(\rho)$ in $\sigma(\tau)\otimes_E E'$. Suppose first that τ is an irregular cuspidal type. Then by Lemmas 2.1.6 and 3.3.1, $\mathcal{D}(\overline{\rho}|_{G_{F_v}}) \cap \mathrm{JH}(\overline{\sigma}(\tau))$ contains a single weight $\overline{\sigma}_J(\tau)$, and we set $\sigma^\circ(\rho) := \sigma^{\circ,J}(\tau)\otimes_\mathcal{O}\mathcal{O}_{E'}$.

Now suppose that τ is a principal series or regular cuspidal type. For each $J \in \imath(\mathcal{P}_\tau)$, we let $\varpi_J(\lambda)$ denote the image of ϖ_J under the composite of λ and the homomorphism $R^\tau_\infty \to \mathbb{T}_\mathfrak{m}$ induced by the composite $R_\infty \to R^{\mathrm{univ}}_S \to \mathbb{T}_\mathfrak{m}$. Define

$$\sigma^\circ(\rho) := \sum_{J\in\imath(\mathcal{P}_\tau)} \varpi_J(\lambda)\sigma^\circ_{\imath(J)}(\tau).$$

8.2.1. **Theorem.** *Continue to maintain the above assumptions, so that in particular we have $p > 3$, and $\overline{\rho}|_{G_{F_v}}$ is generic. Then the lattices $\sigma^\circ(\pi)$ and $\sigma^\circ(\rho)$ are homothetic.*

Proof. Suppose first that τ is an irregular cuspidal type. By Lemma 4.1.1 and the definition of $\sigma^\circ(\rho)$, we must show that the only Serre weight occurring in the socle of $\overline{\sigma}^\circ(\pi)$ is $\overline{\sigma}_J(\tau)$; but this is an immediate consequence of Corollaries 5.6.4 and 4.5.7 of [GK12] (the Buzzard–Diamond–Jarvis conjecture).

Suppose now that τ is either a principal series type or a regular cuspidal type. By Proposition 4.1.4 we may write $\sigma^\circ(\pi) = \sum_{J\in\imath(\mathcal{P}_\tau)} p^{v_J}\sigma^\circ_{\imath(J)}(\tau)$, where p^{v_J} denotes an element of $\mathcal{O}_{E'}$ of some valuation v_J. Rescaling if necessary, we further assume that $v_{\imath(\varnothing)} = 0$.

We now apply the results of Subsection 8.1 to the fixed determinant patching functor M_∞ of Subsection 6.4. It follows from (6.1) that for any subset $J \in \imath(\mathcal{P}_\tau)$ there is a natural isomorphism

$$\mathrm{Hom}_{K_v}\big(\sigma^\circ_{\imath(J)}(\tau), \sigma^\circ(\pi)\big)^* \xrightarrow{\sim} M\big(\sigma^\circ_{\imath(J)}(\tau)\big)/\lambda,$$

and so (noting that $M\big(\sigma^\circ_{\imath(J)}(\tau)\big)$ is p-torsion free) Proposition 8.1.1 shows that

$$\mathrm{Hom}_{K_v}\big(\sigma^\circ_{\imath(J)}(\tau), \sigma^\circ(\pi)\big) = \varpi_J(\lambda)\,\mathrm{Hom}_{K_v}\big(\sigma^\circ_{\imath(\varnothing)}(\tau), \sigma^\circ(\pi)\big).$$

Proposition 4.1.4 then shows that $p^{v_J} = \varpi_J(\lambda)$ for each $J \in \imath(\mathcal{P}_\tau)$ (or more precisely, that they have the same valuation, which is all we need). □

8.2.2. *Remark.* In the case that τ is a principal series type and D is a definite quaternion algebra, this is precisely Conjecture 1.2 of [Bre12], under the mild additional assumptions that we have imposed in order to use the Taylor–Wiles–Kisin method. (See displayed equation (13) in Section 3 of [Bre12] together with [Bre12, Cor. 5.3], and note that our normalisations differ from those of [Bre12].) Note that while Breuil does not fix types at places other than v, the fact that Theorem 8.2.1 shows that the homothety class of the lattice is independent of the choices of types away from v means that Breuil's apparently more general conjecture is an immediate consequence of our result.

9. The generic Buzzard–Diamond–Jarvis conjecture

9.1. **The support of a patching functor.** We now explain how the results of the earlier sections can be applied to the weight part of Serre's conjecture, by an argument analogous to the arguments made to avoid the use of Ihara's lemma in [Tay08]. Fix an unramified extension $F_v/\mathbb{Q}_p$, and a continuous representation $\overline{\rho} : G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$. For simplicity, we consider (fixed determinant) patching functors indexed by just the single field F_v.

9.1.1. **Theorem.** *Suppose that $p > 3$ and that $\overline{\rho}$ is generic, let M_∞ be a fixed determinant patching functor for $\overline{\rho}$, and let $\overline{\sigma}$ be a Serre weight. Then $M_\infty(\overline{\sigma}) \neq 0$ if and only if $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$.*

Proof. First, suppose that $M_\infty(\overline{\sigma}) \neq 0$. Since $M_\infty(\overline{\sigma})$ is supported on $\overline{X}^\psi\big(\overline{\sigma}\big)$ by the definition of a fixed determinant patching functor, we have $\overline{X}^\psi\big(\overline{\sigma}\big) \neq 0$. By Theorem 7.1.1, this implies that $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$.

We now prove the converse. Let τ be an inertial type such that $\mathcal{D}(\overline{\rho}) \subset \mathrm{JH}(\overline{\sigma}(\tau))$, as in Proposition 3.5.2. We certainly have $M_\infty(\overline{\sigma}') \neq 0$ for some Serre weight $\overline{\sigma}'$, and this implies that $M_\infty(\sigma^\circ(\tau)) \neq 0$ (because we have shown that $\overline{\sigma}' \in \mathcal{D}(\overline{\rho})$, so that $\overline{\sigma}'$ is a Jordan–Hölder factor of $\overline{\sigma}^\circ(\tau)$). By Theorem 7.1.1 and the choice of τ, it suffices to show that $M_\infty(\overline{\sigma}^\circ(\tau))$ is supported on all of $\overline{X}^\psi\big(\tau\big)$.

In order to see this, note that by Theorem 7.2.1 the generic fibre of $X^\psi\big(\tau\big)$ is irreducible, so that $M_\infty(\sigma^\circ(\tau))$ is supported on the whole generic fibre of $X^\psi\big(\tau\big)$. Since $M_\infty(\sigma^\circ(\tau))$ is maximal Cohen–Macaulay over $X^\psi\big(\tau\big)$, it follows (*cf.* Lemmas 2.2(1) and 2.3 of [Tay08]) that $M_\infty(\overline{\sigma}^\circ(\tau))$ is supported on all of $\overline{X}^\psi\big(\tau\big)$, as required. □

9.1.2. *Remark.* Of course the analogous result also holds for patching functors without fixed determinant, and can be proved in the same way. The same analysis also goes through for patching functors indexed by a collection of fields, provided that at the fields of residue characteristic p, the fields are absolutely unramified over $\mathbb{Q}_p$, and the Galois representations are generic.

Applying Theorem 9.1.1 to the specific fixed determinant patching functors constructed in Section 6 recovers cases of the Buzzard–Diamond–Jarvis conjecture ([BDJ10]). Of course, as written this proof depends on the results of [GLS12], [GK12], etc. which were used in Section 6 to construct our patching functors, but it is possible to prove enough properties of the patching functors to recover Theorem 9.1.1 without using these results, and in particular without using the potential diagonalizability of potentially Barsotti–Tate representations proved in [Kis09b], [Gee06] — see Section B.1 below.

10. Freeness and multiplicity one

In this section we examine multiplicity one questions for cohomology with coefficients in the lattices under consideration.

10.1. **Freeness of patched modules.** Fix F_v, an unramified extension of $\mathbb{Q}_p$, and fix a continuous representation $\overline{\rho} : G_{F_v} \to \mathrm{GL}_2(\mathbb{F})$. Let $M_\infty^{\min}$ be a minimal fixed determinant patching functor with unramified coefficients, indexed by $(F_v, \overline{\rho})$. As noted in Subsection 5.2, the lattices $\sigma(\tau)_J^\circ$ may be defined over the ring of integers in an unramified extension of $\mathbb{Q}_p$, and we take our unramified coefficient field E to be large enough for all these lattices to be defined over its ring of integers $\mathcal{O}$. Our main theorem is then the following.

10.1.1. **Theorem.** *Assume that $p > 3$ and that $\overline{\rho}$ is generic, and let τ be a non-scalar tame inertial type for I_{F_v}. Fix some $J \in \mathcal{P}_\tau$. Then the coherent sheaf $M_\infty^{\min}(\sigma_J^\circ(\tau))$, regarded as a finitely generated module over $R_\infty^{\psi,\tau}$, is free of rank one.*

10.1.2. *Remark.* As usual, the analogous result also holds for minimal patching functors without fixed determinant, with an essentially identical proof.

In the case when the local deformation space $X^\psi(\tau)$ is regular, such a result follows from the method of Diamond [Dia97], and indeed we will use Diamond's method as an ingredient in our proof via Lemma 6.1.4. However, the deformation spaces $X^\psi(\tau)$ are typically *not* regular, and so more work is required than solely applying this method.

By Theorem 7.2.1, there are subsets $J_{\min} \subseteq J_{\max} \subseteq \mathcal{S}$ with $J_{\min}, J_{\max} \in \mathcal{P}_\tau$ such that $R_\infty^{\psi,\tau}$ is a formal power series ring over

$$\mathcal{O}[\![(X_j, Y_j)_{j \in J_{\max} \setminus J_{\min}}]\!]/(X_j Y_j - p)_{j \in J_{\max} \setminus J_{\min}}.$$

Since $M_\infty^{\min}$ is assumed to have unramified coefficients, this ring is regular if and only if $|J_{\max} \setminus J_{\min}| \leq 1$. In addition, the Jordan–Hölder factors of $\overline{\sigma}(\tau)$ which are elements of $\mathcal{D}(\overline{\rho})$ are precisely the $\overline{\sigma}_{J'}(\tau)$ with $J_{\min} \subseteq J' \subseteq J_{\max}$, and the equations $X_j = 0$ $(j \in J' \setminus J_{\min})$ and $Y_j = 0$ $(j \in J_{\max} \setminus J')$ cut out the component $\overline{X}^\psi\big(\overline{\sigma}_{J'}(\tau)\big)$ of $\overline{X}^\psi\big(\tau\big)$.

For the remainder of this section we fix J as in Theorem 10.1.1. It will be convenient for us to reindex the $\overline{\sigma}_{J'}(\tau)$ and the X_j, Y_j as follows. Set $\overline{\sigma}'_{J'}(\tau) := \overline{\sigma}_{J^c \triangle J'}(\tau)$, so that in particular $\overline{\sigma}'_{\mathcal{S}}(\tau) = \overline{\sigma}_J(\tau)$. Set

$$J'_{\min} := (J \cap J_{\min}) \cup (J^c \cap J^c_{\max}), \ J'_{\max} := (J \cap J_{\max}) \cup (J^c \cap J^c_{\min}).$$

As in Section 8, note that $J'_{\min} \subset J' \subset J'_{\max}$ if and only if $J_{\min} \subset J^c \triangle J' \subset J_{\max}$, and that $J'_{\max} \backslash J'_{\min} = J_{\max} \backslash J_{\min}$. For each $i \in J'_{\max} \backslash J'_{\min}$, we set $X'_i = X_i, Y'_i = Y_i$ if $i \in J$, and $X'_i = Y_i, Y'_i = X_i$ if $i \notin J$, so that the equations $X'_j = 0$ $(j \in J' \setminus J'_{\min})$ and $Y'_j = 0$ $(j \in J'_{\max} \setminus J')$ cut out the component $\overline{X}^{\psi}_{\infty}\big(\overline{\sigma}'_{J'}(\tau)\big)$ of $\overline{X}^{\psi}_{\infty}(\tau)$.

In fact, we will be more interested in the special fibre $\overline{R}^{\psi,\tau}_{\infty}$, where $\overline{R}^{\psi,\tau}_{\infty} = R^{\psi,\tau}_{\infty}/pR^{\psi,\tau}_{\infty}$ is a formal power series ring over

$$\mathbb{F}[\![(X'_j, Y'_j)_{j \in J'_{\max} \backslash J'_{\min}}]\!]/(X'_j Y'_j)_{j \in J'_{\max} \backslash J'_{\min}}.$$

We let $\mathcal{W} := \{J' \,|\, J'_{\min} \subseteq J' \subseteq J'_{\max}\}$. If $\mathcal{J} \subseteq \mathcal{W}$, then we let $I_{\mathcal{J}}$ denote the radical ideal in $\overline{R}^{\psi,\tau}_{\infty}$ which cuts out the reduced induced structure on the closed subset $\cup_{J' \in \mathcal{J}} \overline{X}^{\psi}_{\infty}\big(\overline{\sigma}'_{J'}(\tau)\big)$. If $\mathcal{J}_1, \mathcal{J}_2 \subseteq \mathcal{W}$, then evidently $I_{\mathcal{J}_1 \cup \mathcal{J}_2} = I_{\mathcal{J}_1} \cap I_{\mathcal{J}_2}$. On the other hand, while there is an inclusion $I_{\mathcal{J}_1} + I_{\mathcal{J}_2} \subseteq I_{\mathcal{J}_1 \cap \mathcal{J}_2}$, it is not generally an equality (although it is an equality generically on $\overline{X}^{\psi}_{\infty}(\tau)$).

10.1.3. *Example.* Suppose that $|J'_{\max} \setminus J'_{\min}| = 2$, and write $J'_{\max} \setminus J'_{\min} = \{j_1, j_2\}$.

(1) If we let $\mathcal{J}_1 := \{J'_{\max}\}, \mathcal{J}_2 := \{J'_{\min}\}$, then we have $I_{\mathcal{J}_1} = (X'_{j_1}, X'_{j_2}), I_{\mathcal{J}_2} = (Y'_{j_1}, Y'_{j_2})$, and $I_{\mathcal{J}_1} + I_{\mathcal{J}_2} = (X'_{j_1}, X'_{j_2}, Y'_{j_1}, Y'_{j_2})$, while $I_{\mathcal{J}_1 \cap \mathcal{J}_2} = I_{\varnothing}$ is the unit ideal.

(2) If we let $\mathcal{J}_1 := \{J'_{\max}, J'_{\max} \setminus \{j_2\}\}, \mathcal{J}_2 := \{J'_{\min} \cup \{j_2\}, J'_{\min}\}$, then $I_{\mathcal{J}_1} = (X'_{j_1}), I_{\mathcal{J}_2} = (Y'_{j_1})$, and $I_{\mathcal{J}_1} + I_{\mathcal{J}_2} = (X'_{j_1}, Y'_{j_1})$, while $I_{\mathcal{J}_1 \cap \mathcal{J}_2} = I_{\varnothing}$ is again the unit ideal.

(3) If we let $\mathcal{J}_1 := \{J'_{\max}, J'_{\max} \setminus \{j_2\}\}, \mathcal{J}_2 := \{J'_{\max}, J'_{\min}\}$, then we have $I_{\mathcal{J}_1} = (X'_{j_1}), I_{\mathcal{J}_2} = (X'_{j_1} Y'_{j_2}, X'_{j_2} Y'_{j_1})$, and $I_{\mathcal{J}_1} + I_{\mathcal{J}_2} = (X'_{j_1}, X'_{j_2} Y'_{j_1})$, while $I_{\mathcal{J}_1 \cap \mathcal{J}_2} = I_{\{J'_{\max}\}} = (X'_{j_1}, X'_{j_2})$.

10.1.4. **Definition.** We say that a subset $\mathcal{J} \subseteq \mathcal{W}$ is an *interval* if whenever $J_1 \subseteq J' \subseteq J_2$ with $J_1, J_2 \in \mathcal{J}$, then $J' \in \mathcal{J}$. We say that $\mathcal{J}$ is a *capped interval* if it contains a unique maximal element, which we then refer to as the *cap* of $\mathcal{J}$.

10.1.5. **Definition.** If $J_1 \subseteq J_2$ are elements of $\mathcal{W}$, then we define

$$\mathcal{F}(J_1, J_2) := \{J' \in \mathcal{W} \,|\, J_1 \subseteq J' \subseteq J_2\}$$

and

$$\mathcal{F}(J_1, J_2)^{\times} := \mathcal{F}(J_1, J_2) \setminus \{J_1\} = \{J' \in \mathcal{W} \,|\, J_1 \subsetneq J' \subseteq J_2\}.$$

10.1.6. **Lemma.** *The quotient $I_{\mathcal{F}(J_1,J_2)^{\times}}/I_{\mathcal{F}(J_1,J_2)}$ is isomorphic to $\overline{R}^{\psi,\tau}_{\infty}/I_{\{J_1\}}$ (and so in particular is cyclic), and is generated by the image of the element $\prod_{j \in J_2 \backslash J_1} X'_j$.*

Proof. The ideals $I_{\mathcal{F}(J_1,J_2)}$ and $I_{\mathcal{F}(J_1,J_2)^{\times}}$ admit simple descriptions, namely

$$I_{\mathcal{F}(J_1,J_2)} = (\{X'_j\}_{j \in J_1 \backslash J'_{\min}}, \{Y'_j\}_{j \in J'_{\max} \backslash J_2}),$$

while

$$I_{\mathcal{F}(J_1,J_2)^{\times}} = \Big(\{X'_j\}_{j \in J_1 \backslash J'_{\min}}, \{Y'_j\}_{j \in J'_{\max} \backslash J_2}, \prod_{j \in J_2 \backslash J_1} X'_j\Big).$$

Thus $I_{\mathcal{F}(J_1,J_2)^\times}/I_{\mathcal{F}(J_1,J_2)}$ is generated by $\prod_{j\in J_2\setminus J_1} X_j'$, as claimed. Furthermore, this quotient is supported precisely on the component $\overline{X}\big(\overline{\sigma}'_{J_1}(\tau)\big)$, and so its annihilator is $I_{\{J_1\}}$. □

10.1.7. *Remark.* The preceding lemma can be rephrased more geometrically, namely that the quotient $I_{\mathcal{F}(J_1,J_2)^\times}/(I_{\mathcal{F}(J_1,J_2)^\times}\cap I_{\{J_1\}}) \cong (I_{\mathcal{F}(J_1,J_2)^\times}+I_{\{J_1\}})/I_{\{J_1\}}$ is principal, which is to say that the component $\overline{X}^{\psi}_{\infty}\big(\overline{\sigma}'_{J_1}(\tau)\big)$ intersects the union of components

$$\bigcup_{J'\in\mathcal{F}(J_1,J_2)^\times} \overline{X}^{\psi}_{\infty}\big(\overline{\sigma}'_{J'}(\tau)\big)$$

in a Cartier divisor on $\overline{X}^{\psi}_{\infty}\big(\overline{\sigma}'_{J_1}(\tau)\big)$.

10.1.8. **Lemma.** *If $\mathcal{J}_1$ and $\mathcal{J}_2$ are two capped intervals in $\mathcal{W}$ that share a common cap, then $I_{\mathcal{J}_1}+I_{\mathcal{J}_2}=I_{\mathcal{J}_1\cap\mathcal{J}_2}$.*

Proof. Write $\mathcal{J} := \mathcal{J}_1\cap\mathcal{J}_2$, and let J_{cap} denote the common cap of $\mathcal{J}_1$ and $\mathcal{J}_2$. If $\mathcal{J}=\mathcal{J}_1$ then the statement of the lemma is trivial, and so we may assume that $\mathcal{J}\subsetneq\mathcal{J}_1$. We begin by assuming in addition that $|\mathcal{J}_1\setminus\mathcal{J}_2|=1$, and write $\mathcal{J}_1\setminus\mathcal{J}_2=\{J_1\}$.

The evident inclusion $I_{\mathcal{J}_2}\subseteq I_{\mathcal{J}}$ implies that

$$(I_{\mathcal{F}(J_1,J_{\mathrm{cap}})}+I_{\mathcal{J}_2})\cap I_{\mathcal{J}} = (I_{\mathcal{F}(J_1,J_{\mathrm{cap}})}\cap I_{\mathcal{J}})+I_{\mathcal{J}_2} = I_{\mathcal{F}(J_1,J_{\mathrm{cap}})\cup\mathcal{J}}+I_{\mathcal{J}_2}=I_{\mathcal{J}_1}+I_{\mathcal{J}_2}.$$

Since furthermore $I_{\mathcal{J}}\subseteq I_{\mathcal{F}(J_1,J_{\mathrm{cap}})^\times}$, we see that in order to deduce the lemma in the case we are considering, it suffices to prove that

$$I_{\mathcal{F}(J_1,J_{\mathrm{cap}})}+I_{\mathcal{J}_2}=I_{\mathcal{F}(J_1,J_{\mathrm{cap}})^\times}$$

(that is, it suffices to prove the result in the case that $\mathcal{J}_1=\mathcal{F}(J_1,J_{\mathrm{cap}})$), or equivalently, it suffices to prove that $I_{\mathcal{J}_2}$ surjects onto $I_{\mathcal{F}(J_1,J_{\mathrm{cap}})^\times}/I_{\mathcal{F}(J_1,J_{\mathrm{cap}})}$ under the quotient map $\overline{R}^{\psi,\tau}_{\infty}\to\overline{R}^{\psi,\tau}_{\infty}/I_{\mathcal{F}(J_1,J_{\mathrm{cap}})}$. Now Lemma 10.1.6 shows that the quotient $I_{\mathcal{F}(J_1,J_{\mathrm{cap}})^\times}/I_{\mathcal{F}(J_1,J_{\mathrm{cap}})}$ is cyclically generated by (the image of) $\prod_{j\in J_{\mathrm{cap}}\setminus J_1} X_j'$, and so it suffices to note that this element also lies in $I_{\mathcal{J}_2}$, since $J_1\notin\mathcal{J}_2$.

We now proceed by induction on $|\mathcal{J}_1\setminus\mathcal{J}_2|$, assuming that it is greater than 1. Let J_1 be a maximal element of $\mathcal{J}_1\setminus\mathcal{J}_2$, and write $\mathcal{J}'=\mathcal{J}\cup\{J_1\}$ and $\mathcal{J}_2'=\mathcal{J}_2\cup\{J_1\}$. The maximality of J_1 assures us that $\mathcal{J}'$ and $\mathcal{J}_2'$ are again both intervals. Since $\mathcal{J}_1\cap\mathcal{J}_2'=\mathcal{J}'$, we may assume by induction that

$$I_{\mathcal{J}_1}+I_{\mathcal{J}_2'}=I_{\mathcal{J}'}.$$

Adding $I_{\mathcal{J}_2}$ to both sides of this equality yields

$$I_{\mathcal{J}_1}+I_{\mathcal{J}_2}=I_{\mathcal{J}'}+I_{\mathcal{J}_2}.$$

Thus we are reduced to proving that $I_{\mathcal{J}'}+I_{\mathcal{J}_2}=I_{\mathcal{J}}$. But $|\mathcal{J}'\setminus\mathcal{J}_2|=|\mathcal{J}'\setminus\mathcal{J}|=1$, and so we have reduced ourselves to the situation already treated. This completes the proof of the lemma. □

10.1.9. *Remark.* We remind the reader that as the examples of 10.1.3 show, some sort of hypothesis on $\mathcal{J}_1$ and $\mathcal{J}_2$ as in Lemma 10.1.8 is necessary for that lemma to hold.

10.1.10. *Remark.* Note that Theorem 10.1.1 implies Lemma 10.1.8. To see this, apply Theorem 10.1.1 to the set $J'' \in \mathcal{P}_\tau$ such that $J^c \triangle J''$ is the common cap of $\mathcal{J}_1$ and $\mathcal{J}_2$. Then there are submodules M_1, M_2 of $\overline{\sigma}^\circ_{J''}(\tau)$ such that the Jordan–Hölder factors of $\overline{\sigma}^\circ_{J''}(\tau)/M_i$ are precisely the weights $\overline{\sigma}'_{J'}(\tau)$ with $J' \in \mathcal{J}_i$. Then $I_{\mathcal{J}_1}$ is the annihilator of $M_\infty^{\min}(\overline{\sigma}^\circ_{J''}(\tau)/M_1)$, $I_{\mathcal{J}_2}$ is the annihilator of $M_\infty^{\min}(\overline{\sigma}^\circ_{J''}(\tau)/M_2)$, and $I_{\mathcal{J}_1 \cap \mathcal{J}_2}$ is the annihilator of $M_\infty^{\min}(\overline{\sigma}^\circ_{J''}(\tau)/(M_1+M_2))$, and the result is immediate.

Before we begin the proof of Theorem 10.1.1, we check the following proposition.

10.1.11. **Proposition.** *Suppose that $\overline{\rho}$ is generic, and that for some $\overline{\sigma}, \overline{\sigma}' \in \mathcal{D}(\overline{\rho})$, there is a non-trivial extension $\overline{\theta}$ of $\overline{\sigma}$ by $\overline{\sigma}'$. Then there is a tame inertial type τ' and a K-stable $\mathcal{O}$-lattice $\sigma^\circ(\tau')$ in $\sigma(\tau')$ such that $\mathrm{JH}(\overline{\sigma}(\tau')) \cap \mathcal{D}(\overline{\rho}) = \{\overline{\sigma}, \overline{\sigma}'\}$, and the extension $\overline{\theta}$ is realised in the cosocle filtration of $\overline{\sigma}^\circ(\tau')$.*

Proof. First, note that since $\mathrm{Ext}^1_{\mathbb{F}[\mathrm{GL}_2(k_v)]}(\overline{\sigma}', \overline{\sigma})$ is one-dimensional (by Proposition 5.1.3), it suffices to find $\overline{\sigma}^\circ(\tau')$ which realises a non-split extension of $\overline{\sigma}'$ by $\overline{\sigma}$.

In fact, we claim that it suffices to find τ' such that $\mathrm{JH}(\overline{\sigma}(\tau')) \cap \mathcal{D}(\overline{\rho}) = \{\overline{\sigma}, \overline{\sigma}'\}$. To see this, suppose that such a τ' exists, and take $\sigma^\circ(\tau')$ to be the lattice $\sigma(\tau')_{\overline{\sigma}}$ provided by Lemma 4.1.1. It suffices to prove that $\overline{\sigma}'$ necessarily lies in the first layer of the cosocle filtration of $\sigma^\circ(\tau')$ (since the extension of $\overline{\sigma}$ by $\overline{\sigma}'$ must be nontrivial, as the socle of $\sigma^\circ(\tau')$ is just $\overline{\sigma}$); but this follows from Theorem 5.1.1 and Proposition 5.1.3.

We now find such a type τ'. Since $\mathcal{D}(\overline{\rho}) \subset \mathcal{D}(\overline{\rho}^{\mathrm{ss}})$, we may assume that $\overline{\rho}$ is semisimple. Without loss of generality we may assume that $\overline{\sigma} = \overline{\sigma}_{\vec{t},\vec{s}}$, $\overline{\sigma}' = \overline{\sigma}_{\vec{t}',\vec{s}'}$ as in the statement of Proposition 5.1.3, so that there is some k with $s_j = s'_j$ if $j \neq k, k+1$, $s'_k = p-2-s_k$, and $s'_{k+1} = s_{k+1} \pm 1$. Let $\overline{\rho}' := \mathrm{BC}\,(\overline{\rho})$. By Corollary 7.4.4, it is enough to find a tame inertial type τ' such that if $\tau'' := \mathrm{BC}\,(\tau')$, then $\mathrm{BC}\,(\overline{\sigma}), \mathrm{BC}\,(\overline{\sigma}')$ are Jordan–Hölder factors of $\overline{\sigma}(\tau'')$, and no other elements of $\mathrm{JH}(\overline{\sigma}(\tau'')) \cap \mathcal{D}(\overline{\rho}')$ are in the image of BC. We will construct τ'' directly as a principal series representation, and it will be immediate by the construction that τ'' is in the image of BC.

The construction of such a τ'' (and thus of τ') is an easy exercise; we sketch the details. Suppose first that $\overline{\rho}$ is reducible, so that we may write

$$\overline{\rho}|_{I_{F_v}} \cong \omega_f^{\sum_{j=0}^{f-1} t_j p^j} \begin{pmatrix} \omega_f^{\sum_{j\in K}(s_j+1)p^j} & 0 \\ 0 & \omega_f^{\sum_{j\notin K}(s_j+1)p^j} \end{pmatrix}$$

for some subset $K \subseteq \mathcal{S}$; then we let $K' = \mathrm{BC}_{\mathrm{PS}}(K) \subseteq \mathcal{S}'$. If on the other hand $\overline{\rho}$ is irreducible, then we may write

$$\overline{\rho}|_{I_{F_v}} \cong \omega_f^{\sum_{j=0}^{f-1} t_j p^j} \begin{pmatrix} \omega_{2f}^{\sum_{j\in K}(s_j+1)p^j} & 0 \\ 0 & \omega_{2f}^{\sum_{j\notin K}(s_j+1)p^j} \end{pmatrix},$$

and we let $K' = K \subseteq \mathcal{S}'$.

In either case we let $J' = K' \triangle \{k+1, \ldots, k+f\}$, so that J' is antisymmetric if $\overline{\rho}$ is reducible, and symmetric if $\overline{\rho}$ is irreducible. Define $\tau'' := I(\eta \otimes \eta')$, where $\eta' = [\cdot]^{\sum_{j=0}^{2f-1} x_j p^j}$, $\eta(\eta')^{-1} = [\cdot]^{\sum_{j=0}^{2f-1} c_j p^j}$, with

$$x_i = \begin{cases} t_i + s_i + 1 - p & \text{if } i \in J' \\ t_i & \text{if } i \notin J', \end{cases}$$

$$c_i = \begin{cases} p-1-s_i-\delta_{(J')^c}(i-1) & \text{if } i \in J' \\ s_i+\delta_{J'}(i-1) & \text{if } i \notin J'. \end{cases}.$$

Note by Lemma 2.1.6 that each c_i is in the range $[0, p-1]$; observe also that if $\overline{\rho}$ is reducible (respectively irreducible) then there is a cuspidal (respectively principal series) inertial type τ' with $\mathrm{BC}\,(\tau') = \tau''$.

Now, by definition we have $\overline{\sigma}_{J'}(\tau'') = \mathrm{BC}\,(\overline{\sigma})$, and $\overline{\sigma}(\tau'')_{J' \triangle \{k,k+f\}} = \mathrm{BC}\,(\overline{\sigma}')$. It is elementary to check (as an application of Proposition 4.3 of [Bre12], and especially equation (19) from its proof) that there are precisely two other Jordan–Hölder factors of $\overline{\sigma}(\tau'')$ which are contained in $\mathcal{D}(\overline{\rho}')$, namely $\overline{\sigma}(\tau'')_{J' \triangle \{k\}}$ and $\overline{\sigma}(\tau'')_{J' \triangle \{k+f\}}$, neither of which is in the image of BC. The result follows. $\square$

Proof of Theorem 10.1.1. Since $M_\infty^{\min}(\sigma_J^\circ(\tau))$ is a faithful $R_\infty^{\psi,\tau}$-module (as by Lemma 6.1.4 the module $M_\infty^{\min}(\sigma_J^\circ(\tau))[1/p]$ is locally free over $R_\infty^{\psi,\tau}[1/p]$), the module $M_\infty^{\min}(\sigma_J^\circ(\tau))$ is free of rank one if and only if it is cyclic. By Nakayama's lemma, $M_\infty^{\min}(\sigma_J^\circ(\tau))$ is cyclic if and only if $M_\infty^{\min}(\overline{\sigma}_J^\circ(\tau))$ is cyclic.

It follows easily from Theorem 5.1.1 that for each interval $\mathcal{J} \subseteq \mathcal{W}$, there is a subquotient $\overline{\sigma}^{\mathcal{J}}$ of $\overline{\sigma}_J^\circ(\tau)$ uniquely characterized by the property that its set of Jordan–Hölder factors is precisely the set $\{\overline{\sigma}'_{J'}(\tau)\}_{J' \in \mathcal{J}}$.

We will prove, by induction on $|\mathcal{J}|$, that for each capped interval $\mathcal{J} \subseteq \mathcal{W}$, the patched module $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ is cyclic, with annihilator equal to $I_{\mathcal{J}}$. (The assumption that $\mathcal{J}$ be *capped* is crucial; see Remark 10.1.14 below.) The theorem will follow by taking $\mathcal{J} = \mathcal{W}$; indeed, although $M_\infty^{\min}(\overline{\sigma}^{\mathcal{W}})$ is *a priori* a subquotient of $M_\infty^{\min}(\overline{\sigma}_J^\circ(\tau))$, in fact the two are canonically isomorphic, because $M_\infty^{\min}(\overline{\sigma}'_{J'}(\tau))$ vanishes if $J' \notin \mathcal{W}$, by Theorem 7.1.1.

The induction begins with the following lemmas.

10.1.12. **Lemma.** *If $\mathcal{J}$ is an interval with $|\mathcal{J}| \le 2$, then $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ is cyclic, with annihilator $I_{\mathcal{J}}$.*

Proof. Since $\mathcal{J}$ is an interval with at most two elements, it is either empty (in which case $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}}) = 0$ and the lemma is trivially true), consists of a single element, or consists of two adjacent sets J_1 and J_2. Suppose that $\mathcal{J}$ is non-empty.

Since $|\mathcal{J}| \le 2$, by Propositions 3.5.1 and 10.1.11 we may find an inertial type τ' such that $\mathrm{JH}(\overline{\sigma}(\tau')) \cap \mathcal{D}(\overline{\rho}) = \{\overline{\sigma}'_{J'}(\tau)\}_{J' \in \mathcal{J}}$. Since $X_\infty^\psi(\tau')$ is regular, $M_\infty^{\min}(\sigma^\circ(\tau'))$ is free of rank one over $X_\infty^\psi(\tau')$ for any choice of lattice $\sigma^\circ(\tau')$, by Lemma 6.1.4. Note that by Theorem 7.1.1, $I_{\mathcal{J}}$ cuts out the reduced induced structure on the subset $\overline{X}_\infty^\psi(\tau')$ of $\overline{X}_\infty^\psi(\tau)$.

We claim that we can choose $\sigma^\circ(\tau')$ so that $M_\infty^{\min}(\overline{\sigma}^\circ(\tau')) = M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$, from which the lemma will follow. In the case that $\mathcal{J}$ is a singleton, this is trivial. In the case that $\mathcal{J} = \{J_1, J_2\}$, it follows from Proposition 10.1.11 and Theorem 5.2.4. $\square$

We will also require the following lemma.

10.1.13. **Lemma.** *Let R be a local ring, and let $M'' \subsetneq M' \subseteq M$ be R-modules such that M' and M/M'' are both cyclic. Then M is cyclic.*

Proof. Let m be a lift to M of a cyclic generator of M/M''. We claim that $M = Rm$. In order to see this, it suffices to show that $M'' \subseteq Rm$, and thus it suffices to show that $M' \subseteq Rm$.

Choose $r \in R$ such that $m' := rm$ generates the cyclic submodule M'/M'' of M/M'', so that $Rm' \subseteq M'$. Since M'/M'' is non-zero, we see by Nakayama's lemma that m' is in fact a generator of M', so $M' = Rm'$. Thus $M' = Rm' \subseteq Rm$, as required. $\square$

We now complete the induction, using Lemma 10.1.12 to provide the base case $|\mathcal{J}| \leq 2$. Suppose then that $|\mathcal{J}| > 2$. We consider two cases: when $\mathcal{J}$ has a unique minimal element, and when $\mathcal{J}$ has at least two minimal elements.

Suppose first that $\mathcal{J}$ contains a unique minimal element J_0, and let J' be a minimal element of $\mathcal{J} \setminus \{J_0\}$. Then we have (by Theorem 5.1.1) inclusions $\overline{\sigma}^{\{J_0\}} \subsetneq \overline{\sigma}^{\{J_0,J'\}} \subseteq \overline{\sigma}^{\mathcal{J}}$, and an equality $\overline{\sigma}^{\mathcal{J}}/\overline{\sigma}^{\{J_0\}} = \overline{\sigma}^{\mathcal{J}\setminus\{J_0\}}$, which induce (since $M_\infty^{\min}$ is an exact functor) inclusions $M_\infty^{\min}(\overline{\sigma}^{\{J_0\}}) \subsetneq M_\infty^{\min}(\overline{\sigma}^{\{J_0,J'\}}) \subseteq M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ (the first inclusion being strict since $M_\infty^{\min}(\overline{\sigma}^{\{J'\}}) \neq 0$, as $J' \in \mathcal{W}$), and an equality $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})/M_\infty^{\min}(\overline{\sigma}^{\{J_0\}}) = M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}\setminus\{J_0\}})$. By induction, we know that each of $M_\infty^{\min}(\overline{\sigma}^{\{J_0,J'\}})$ and $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}\setminus\{J_0\}})$ is cyclic, and hence by Lemma 10.1.13 we deduce that $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ is cyclic.

It remains to determine the annihilator of $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$. If we denote this annihilator by I, then certainly $I \subseteq I_{\mathcal{J}}$, and $I_{\mathcal{J}\setminus\{J_0\}}/I \xrightarrow{\sim} M_\infty^{\min}(\overline{\sigma}'_{J_0}(\tau)) \xrightarrow{\sim} \overline{R}_\infty^{\psi,\tau}/I_{\{J_0\}}$. Thus to show that $I = I_{\mathcal{J}}$, it suffices to prove that $I_{\mathcal{J}\setminus\{J_0\}}/I_{\mathcal{J}} \xrightarrow{\sim} \overline{R}_\infty^{\psi,\tau}/I_{\{J_0\}}$. For this, note that (letting J_{cap} denote the cap of $\mathcal{J}$)

$$\begin{aligned} I_{\mathcal{J}\setminus\{J_0\}}/I_{\mathcal{J}} = I_{\mathcal{J}\setminus\{J_0\}}/I_{\mathcal{J}\setminus\{J_0\}} \cap I_{\mathcal{F}(J_0,J_{\mathrm{cap}})} &\xrightarrow{\sim} (I_{\mathcal{J}\setminus\{J_0\}} + I_{\mathcal{F}(J_0,J_{\mathrm{cap}})})/I_{\mathcal{F}(J_0,J_{\mathrm{cap}})} \\ &= I_{\mathcal{F}(J_0,J_{\mathrm{cap}})^\times}/I_{\mathcal{F}(J_0,J_{\mathrm{cap}})} \xrightarrow{\sim} \overline{R}_\infty^{\psi,\tau}/I_{\{J_0\}} \end{aligned}$$

(the first and second equalities being evident, the third being an application of Lemma 10.1.8, and the isomorphism being provided by Lemma 10.1.6), as required.

Suppose now that $\mathcal{J}$ contains at (at least) two distinct minimal elements, say J_1 and J_2. Write $\mathcal{J}_1 := \mathcal{J} \setminus \{J_1\}$, $\mathcal{J}_2 := \mathcal{J} \setminus \{J_2\}$. Then $\overline{\sigma}^{\mathcal{J}}$ is naturally identified with the fibre product of $\overline{\sigma}^{\mathcal{J}_1}$ and $\overline{\sigma}^{\mathcal{J}_2}$ over $\overline{\sigma}^{\mathcal{J}_1\cap\mathcal{J}_2}$, and so $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ is naturally identified with the fibre product of $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}_1})$ and $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}_2})$ over $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}_1\cap\mathcal{J}_2})$ (since $M_\infty^{\min}$ is an exact functor). By induction, there are isomorphisms $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}_1}) \xrightarrow{\sim} \overline{R}_\infty^{\psi,\tau}/I_{\mathcal{J}_1}$, $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}_2}) \xrightarrow{\sim} \overline{R}_\infty^{\psi,\tau}/I_{\mathcal{J}_2}$, and $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}_1\cap\mathcal{J}_2}) \xrightarrow{\sim} \overline{R}_\infty^{\psi,\tau}/I_{\mathcal{J}_1\cap\mathcal{J}_2}$. From Lemma 10.1.8 we know that $I_{\mathcal{J}_1\cap\mathcal{J}_2} = I_{\mathcal{J}_1} + I_{\mathcal{J}_2}$, while clearly $I_{\mathcal{J}_1} \cap I_{\mathcal{J}_2} = I_{\mathcal{J}_1\cup\mathcal{J}_2} = I_{\mathcal{J}}$. Since the fibre product of $\overline{R}_\infty^{\psi,\tau}/I_{\mathcal{J}_1}$ and $\overline{R}_\infty^{\psi,\tau}/I_{\mathcal{J}_2}$ over $\overline{R}_\infty^{\psi,\tau}/(I_{\mathcal{J}_1} + I_{\mathcal{J}_2})$ is $\overline{R}_\infty^{\psi,\tau}/(I_{\mathcal{J}_1} \cap I_{\mathcal{J}_2})$, we conclude that indeed $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ is isomorphic to $\overline{R}_\infty^{\psi,\tau}/I_{\mathcal{J}}$, completing the inductive step, and so completing the proof of the theorem. $\square$

10.1.14. *Remark.* The assumption in the preceding proof that $\mathcal{J}$ be a *capped* interval is crucial for deducing that $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ is cyclic. Indeed, if $\mathcal{J}$ admits two distinct maximal elements J_1 and J_2, then $\overline{\sigma}^{\mathcal{J}}$ admits a surjection onto $\overline{\sigma}'_{J_1}(\tau) \oplus \overline{\sigma}'_{J_2}(\tau)$, and hence $M_\infty^{\min}(\overline{\sigma}^{\mathcal{J}})$ admits a surjection onto $M_\infty^{\min}(\overline{\sigma}'_{J_1}(\tau)) \oplus M_\infty^{\min}(\overline{\sigma}'_{J_2}(\tau))$; thus it cannot be cyclic.

10.1.15. *Remark.* By Proposition 8.1.1 and Theorem 10.1.1, we see that if the tame inertial type τ is either principal series or regular cuspidal, then the restriction of $M_\infty^{\min}$ to the category of subquotients of lattices in $\sigma(\tau)$ is independent of the particular choice of $M_\infty^{\min}$. Indeed, we can explicitly define a particular choice of $M_\infty^{\min}$ with no formal variables in the following way: we fix some $J \in \mathcal{P}_\tau$, and we let $M_\infty^{\min}(\sigma_J^\circ(\tau))$ be the structure sheaf of $R^{\psi,\tau}$. Then the other $M_\infty^{\min}(\sigma_{J'}^\circ(\tau))$ are uniquely determined by Proposition 8.1.1, and since any lattice in $\sigma(\tau)$ is in the span of the $\sigma^\circ_{J'}(\tau)$ by Proposition 4.1.4, we see that this determines $M_\infty^{\min}$ completely. Any other choice of $M_\infty^{\min}$ is then equivalent to one obtained from this one by base extension to $R_\infty^{\psi,\tau}$.

10.2. **Application to a conjecture of Dembélé.** We now give a proof of a conjecture of Dembélé ([Dem]; see also the introduction to [Bre12]), by applying Theorem 10.1.1 with $M_\infty^{\min}$ the minimal fixed determinant patching functor of Section 6.5, and τ a principal series type. (Using the construction of Section 6.6, a similar result could also be proved for unitary groups.)

We put ourselves in the setting of Section 6.2. In particular, we have a prime $p\geq 5$, a totally real field F in which p is unramified, and a modular representation $\overline{\rho}:G_F\to \mathrm{GL}_2(\mathbb{F})$ such that $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible, and if $p=5$ then we assume further that the projective image of $\overline{\rho}(G_{F(\zeta_5)})$ is not isomorphic to A_5. We also have a quaternion algebra D with centre F which is ramified at all but at most one of the infinite places of F, and which is ramified at a set Σ of finite places of F that does not contain any places lying over p. We assume further that

- $\overline{\rho}|_{G_{F_w}}$ is generic for all places $w|p$, and
- if $w\in\Sigma$, then $\overline{\rho}|_{G_{F_w}}$ is not scalar.

Let S be the union of Σ, the set of places dividing p, and the places where $\overline{\rho}$ is ramified. Fix a place $v|p$ of F; then for each place $w\in S\setminus\{v\}$, we defined a compact open subgroup K_w of $(\mathcal{O}_D)_w^\times$ in Section 6.5, and a finite free $\mathcal{O}$-module L_w with an action of K_w. At some places $w\in S\setminus\{v\}$ we also defined Hecke operators T_w and scalars $\beta_w\in\mathbb{F}^\times$. We then set $K_v=\mathrm{GL}_2(\mathcal{O}_{F_v})$, $K=\prod_w K_w$, and defined a space $S^{\min}(\sigma_v)_\mathfrak{m}:=S(\sigma_v\otimes(\otimes_{w\in S,w\neq v}L_w))_{\mathfrak{m}'}$ for each σ_v a finitely generated $\mathcal{O}$-module with a continuous action of $\mathrm{GL}_2(\mathcal{O}_{F_v})$, having a central character which lifts $(\overline{\varepsilon}\det\overline{\rho}|_{I_{F_v}})\circ\mathrm{Art}_{F_v}$. By definition, $S^{\min}(\sigma_v)_\mathfrak{m}$ is a space of modular forms (strictly speaking, it is a localisation of a space of modular forms). We then constructed $M^{\min}(\sigma_v)$ by taking a dual, and additionally by factoring out the Galois action in the indefinite case.

We extend this definition slightly in the following way. We write I_v for the Iwahori subgroup of $\mathrm{GL}_2(\mathcal{O}_{F_v})$ consisting of matrices which are upper triangular modulo p, and for each character $\chi:I_v\to\mathbb{F}^\times$ such that $\det\chi=\overline{\varepsilon}\det\overline{\rho}|_{I_{F_v}}$, we define a space $S^{\min}(K^vI_v,\chi):=S(K^vI_v,\chi(\otimes_{w\in S,w\neq v}L_w))_{\mathfrak{m}'}$ exactly as above. By factoring out the Galois action in the indefinite case, and then taking duals, we also define a module $M^{\min}(K^vI_v,\chi)$.

10.2.1. **Theorem.** *With the above notation and assumptions (so in particular $p\geq 5$, $\overline{\rho}|_{G_{F_w}}$ is generic for all places $w|p$, and if $w\in\Sigma$, then $\overline{\rho}|_{G_{F_w}}$ is not scalar), we have $\dim_\mathbb{F} M^{\min}(K^vI_v,\chi)^*[\mathfrak{m}']\leq 1$.*

Proof. If $M^{\min}(K^vI_v,\chi)=0$ then there is nothing to prove. Otherwise, we must show that $\dim_\mathbb{F} M^{\min}(K^vI_v,\chi)^*[\mathfrak{m}']=1$. By Frobenius reciprocity, we have

$$M^{\min}(K^vI_v,\chi)^*[\mathfrak{m}']=M^{\min}(\mathrm{Ind}_{I_v}^{K_v}\chi)^*[\mathfrak{m}']=M^{\min}(\overline{\sigma}^\circ_\mathcal{S}(\chi))^*[\mathfrak{m}'].$$

Suppose first that χ is not scalar (that is, we can write $\chi=\eta\otimes\eta'$ with $\eta\neq\eta'$). Then, applying Theorem 10.1.1 with $\sigma(\tau)=\sigma(\chi)$, $J=\mathcal{S}$, and $M_\infty^{\min}$ the patching functor of Section 6.5, we see that $M_\infty^{\min}(\sigma^\circ_\mathcal{S}(\chi))$ is a cyclic $R_\infty^{\psi,\tau}$-module. By construction, $M^{\min}(\overline{\sigma}^\circ_\mathcal{S}(\chi))^*[\mathfrak{m}']$ is dual to the cyclic k-module

$$M_\infty^{\min}(\sigma^\circ_\mathcal{S}(\chi))/\mathfrak{m}_{R_\infty^{\psi,\tau}}M_\infty^{\min}(\sigma^\circ_\mathcal{S}(\chi)),$$

and we are done in this case.

It remains to deal with the case that χ is scalar. Twisting, we may without loss of generality assume that χ is trivial. Then the Jordan–Hölder factors of $\overline{\sigma}^{\circ}_{\mathcal{S}}(\chi)$ are $\overline{\sigma}_{\vec{0},\vec{0}}$ and $\overline{\sigma}_{\vec{0},\vec{p}-\vec{1}}$, so we have $M^{\min}(\overline{\sigma}^{\circ}_{\mathcal{S}}(\chi)) = 0$ by Lemma 2.1.6 and Theorem 7.1.1, and we are done. $\square$

10.2.2. *Remark.* The relationship of this result to Conjecture B.1 of [Dem] is as follows. Our condition that there are no places $w \in \Sigma$ with $\overline{\rho}|_{G_{F_w}}$ scalar follows from Dembélé's assumption that there are no places $w \in \Sigma$ with $\mathbf{N}w \equiv 1 \pmod p$, and our only additional restrictions are that $p \geq 5$ and $\overline{\rho}|_{G_{F(\zeta_p)}}$ is absolutely irreducible. Conjecture B.1 of [Dem] predicts the dimension of a certain space of mod p modular forms at pro-p Iwahori level (note that while the statement of the conjecture does not specify this, it is intended that the mod p modular forms considered are new at places away from p). There is a natural variant of this conjecture in our setting (with our slightly different choice of tame level), and this conjecture follows from Theorem 10.2.1 and [BP12, Proposition 14.7] (which shows that the number of characters χ with $M^{\min}(K^v I_v, \chi)^*[\mathfrak{m}'] \neq 0$ is precisely the dimension conjectured in [Dem]).

10.2.3. **Corollary.** *Maintaining the assumptions of Theorem* 10.2.1*, the* $\mathrm{GL}_2(F_v)$*-representation* $\varinjlim_{K_v} M^{\min}(K_v, \mathbb{F})^*_{\mathfrak{m}'}$ *contains one of the representations constructed in* [BP12]*.*

Proof. This follows immediately from Theorem 10.2.1, because if I_v^1 denotes the pro-p Sylow subgroup of I_v, then $M^{\min}(K^v I_v^1, \chi)^*[\mathfrak{m}']$ is stabilised by $\begin{pmatrix} 0 & 1 \\ p & 0 \end{pmatrix}$ (*cf.* the discussion between Théorèmes 1.4 and 1.5 of [Bre12]). $\square$

10.2.4. *Remark.* Breuil has shown that Corollary 10.2.3 can also be deduced from Theorem 8.2.1, without using the results of this section; see the proof of Théorème 10.1 of [Bre12].

Acknowledgements. We would like to thank Christophe Breuil and Fred Diamond for sending us a preliminary version of their paper [BD14], which was very helpful in writing Section 6.5. We would like to thank Christophe Breuil, Fred Diamond, Hui Gao, David Helm, Florian Herzig and Vytautas Paškūnas for helpful conversations. The debt that this paper owes to the work of Christophe Breuil will be obvious to the reader, and it is a pleasure to acknowledge this.

Appendix A. Unipotent Fontaine–Laffaille modules and ϕ-modules

A.1. **Unipotent objects.** Recall that results in p-adic Hodge theory that are valid in the Fontaine–Laffaille range can often be extended slightly to the so-called *unipotent* case (see below for specific examples of what we mean by this). In the proof of Lemma 7.4.2, we need to compare the Galois representations associated to certain unipotent φ-modules with the Galois representations associated to unipotent Fontaine–Laffaille modules. Breuil makes a similar comparison (without allowing the unipotent case) in the proofs of [Bre12, Prop. 7.3] and [Bre12, Prop. A.3], making use of certain results from [Bre99b]. In this appendix we will extend those results to the unipotent case, so that we can carry out the same argument. Since these extensions are minor, we will generally indicate where the proofs of [Bre99b] need to be changed, rather than repeating any proofs in their entirety. We attempt to keep our notation as consistent as possible with that of [Bre99b]; for instance we write ϕ in this appendix where we wrote φ in the body of the paper.

A.1.1. *Remark.* As in Appendix A of [Bre12], there exist $\mathbb{F}$-coefficient versions of all of the following results (rather than just $\mathbb{F}_p$-coefficients); in fact we can take the coefficients to be any Artinian local $\mathbb{F}_p$-algebra (such as the the ring $\mathbb{F}_J$ in the proof of Lemma 7.4.2). These more general results follow entirely formally from their $\mathbb{F}_p$-coefficient versions, and so to simplify our notation we will omit coefficients in what follows.

Throughout this appendix we let k be a perfect field of characteristic $p > 0$, and let K_0 denote the fraction field of $W(k)$. Fix a uniformizer π of K_0. We begin by recalling the definitions of several categories of objects that we wish to consider.

A.1.2. **Definition.** Suppose that $0 \le h \le p-1$. We define the category $\underline{\mathrm{MF}}_k^{f,h}$ (the category of mod p Fontaine–Laffaille modules of height h) to be the category whose objects consist of:

- a finite-dimensional k-vector space M,
- a filtration $(\mathrm{Fil}^i M)_{i\in\mathbb{Z}}$ such that $\mathrm{Fil}^i M = M$ for $i \le 0$ and $\mathrm{Fil}^i M = 0$ for $i \ge h+1$, and
- for each integer i, a semi-linear map $\phi_i : \mathrm{Fil}^i M \to M$ such that $\phi_i|_{\mathrm{Fil}^{i+1} M} = 0$ and $\sum \mathrm{Im}\, \phi_i = M$.

Morphisms in $\underline{\mathrm{MF}}_k^{f,h}$ are k-linear maps that are compatible with filtrations and commute with the ϕ_i.

When $h \le p-2$, there is an exact and fully faithful functor from the category $\underline{\mathrm{MF}}_k^{f,h}$ to the category of $\mathbb{F}_p$-representations of the absolute Galois group of K_0; this is no longer the case when $h = p-1$, but as we will see, it remains true if we restrict to a certain subcategory of $\underline{\mathrm{MF}}_k^{f,p-1}$.

A.1.3. **Definition.** We say that an object M of $\underline{\mathrm{MF}}_k^{f,h}$ is *multiplicative* if $\mathrm{Fil}^h M = M$; we say that M is *unipotent* if it has no nontrivial multiplicative quotients. Let $\underline{\mathrm{MF}}_k^{u,p-1}$ denote the full subcategory of $\underline{\mathrm{MF}}_k^{f,p-1}$ consisting of the unipotent objects.

We turn next to strongly divisible modules. Let $\overline{S} = k\langle u\rangle$ be the divided power polynomial algebra in the variable u, and let $\mathrm{Fil}^h \overline{S}$ be the ideal generated by $\gamma_i(u)$ for $i \ge h$. Write c for the element $-\gamma_p(u) - (\pi/p) \in \overline{S}^{\times}$, and for $0 \le h \le p-1$ define $\phi_h : \mathrm{Fil}^h \overline{S} \to \overline{S}$ by setting $\phi_h(\gamma_h(u)) = c^h/h!$ and $\phi_h(\gamma_i(u)) = 0$ for $i > h$, except that when $h = p-1$ we set $\phi_{p-1}(\gamma_p(u)) = c^p/(p-1)! \in \overline{S}^{\times}$. Set $\phi = \phi_0$.

A.1.4. **Definition.** Suppose that $0 \le h \le p-1$. We define the category $\underline{\mathcal{M}}_{0,k}^h$ (the category of mod p strongly divisible modules of height h) to be the category whose objects consist of:

- a free, finite rank $\overline{S}$-module $\mathcal{M}$,
- a sub-$\overline{S}$-module $\mathrm{Fil}^h \mathcal{M}$ of $\mathcal{M}$ containing $(\mathrm{Fil}^h \overline{S})\mathcal{M}$
- a ϕ-semi-linear map $\phi_h : \mathrm{Fil}^h \mathcal{M} \to \mathcal{M}$ such that for each $s \in \mathrm{Fil}^h \overline{S}$ and $x \in \mathcal{M}$ we have $\phi_h(sx) = c^{-h}\phi_h(s)\phi_h(u^h x)$ and such that $\phi_h(\mathrm{Fil}^h \mathcal{M})$ generates $\mathcal{M}$ over $\overline{S}$.

Morphisms in $\underline{\mathcal{M}}_{0,k}^h$ are $\overline{S}$-linear maps that are compatible with filtrations and commute with ϕ_h.

A.1.5. **Definition.** An object of $\underline{\mathcal{M}}_{0,k}^h$ is called *multiplicative* if $\mathrm{Fil}^h \mathcal{M} = \mathcal{M}$, and *unipotent* if it has no non-zero multiplicative quotients. Let $\underline{\mathcal{M}}_{0,k}^{u,p-1}$ denote the full subcategory of $\underline{\mathcal{M}}_{0,k}^{p-1}$ consisting of the unipotent objects.

Finally we define several categories of *étale ϕ-modules*.

A.1.6. **Definition.** Let $\underline{\mathfrak{D}}$ be the category whose objects are finite-dimensional $k(\!(u)\!)$-vector spaces $\mathfrak{D}$ equipped with an injective map $\mathfrak{D} \to \mathfrak{D}$ that is semilinear with respect to the pth power map on $k(\!(u)\!)$; morphisms in $\underline{\mathfrak{D}}$ are $k(\!(u)\!)$-linear maps that commute with ϕ.

Let $\underline{\mathfrak{M}}$ be the category whose objects are finite-dimensional $k[\![u]\!]$-modules $\mathfrak{M}$ equipped with an injective map $\mathfrak{M} \to \mathfrak{M}$ that is semilinear with respect to the pth power map on $k[\![u]\!]$; morphisms in $\underline{\mathfrak{M}}$ are $k[\![u]\!]$-linear maps that commute with ϕ.

Write ϕ^* for the $k[\![u]\!]$-linear map $\mathrm{Id} \otimes \phi : k[\![u]\!] \otimes_\phi \mathfrak{M} \to \mathfrak{M}$. We say that an object $\mathfrak{M}$ of $\underline{\mathfrak{M}}$ has height h if $\mathfrak{M}/\mathrm{Im}(\phi^*)$ is killed by u^h. Write $\underline{\mathfrak{M}}^h$ for the full subcategory of $\underline{\mathfrak{M}}$ of objects of height h, and $\underline{\mathfrak{D}}^h$ for the full subcategory of $\underline{\mathfrak{D}}$ consisting of objects isomorphic to $\mathfrak{M}[\frac{1}{u}]$ for some $\mathfrak{M} \in \underline{\mathfrak{M}}^h$.

A.1.7. **Definition.** We say that an object $\mathfrak{M}$ of $\underline{\mathfrak{M}}^h$ is *unipotent* if it has no non-zero quotients $\mathfrak{M}'$ such that $\phi(\mathfrak{M}') \subset u^h \mathfrak{M}'$. Let $\underline{\mathfrak{M}}^{u,p-1}$ denote the full subcategory of $\underline{\mathfrak{M}}^{p-1}$ consisting of the unipotent objects; let $\underline{\mathfrak{D}}^{u,p-1}$ denote the full subcategory of $\underline{\mathfrak{D}}^{p-1}$ consisting of objects isomorphic to $\mathfrak{M}[\frac{1}{u}]$ for $\mathfrak{M} \in \underline{\mathfrak{M}}^{u,p-1}$.

We recall the following fundamental result of Fontaine.

A.1.8. **Theorem.** ([Fon90, B.1.7.1]) *The functor $\mathfrak{M} \rightsquigarrow \mathfrak{M}[\frac{1}{u}]$ is an equivalence of categories from $\underline{\mathfrak{M}}^h$ to $\underline{\mathfrak{D}}^h$ for $0 \le h \le p-2$, as well as from $\underline{\mathfrak{M}}^{u,p-1}$ to $\underline{\mathfrak{D}}^{u,p-1}$.*

A.2. **Generic fibres.** We fix an algebraic closure $\overline{K}_0$ of K_0, and write $G_{K_0} = \mathrm{Gal}(\overline{K}_0/K_0)$. We also fix a system of elements π_n in $\overline{K}_0$, for $n \ge 0$, such that $\pi_0 = \pi$ and $\pi_n^p = \pi_{n-1}$ for $n \ge 1$. We write $K_\infty = \cup_n K_0(\pi_n)$, and $G_\infty = \mathrm{Gal}(\overline{K}_0/K_\infty)$.

For each of the categories in Definitions A.1.2–A.1.7 we will now define a "generic fibre" functor to k-representations of G_{K_0} (for Fontaine–Laffaille modules) or G_∞ (for strongly divisible modules and ϕ-modules).

Let A_{cris} be the ring (of the same name) defined in [Bre98, §3.1.1]. The ring A_{cris} has a filtration $\mathrm{Fil}^h A_{\mathrm{cris}}$ for $h \ge 0$ and an endomorphism ϕ with the property that $\phi(\mathrm{Fil}^h A_{\mathrm{cris}}) \subset p^h A_{\mathrm{cris}}$ for $0 \le h \le p-1$; for such h we may define $\phi_h : \mathrm{Fil}^h A_{\mathrm{cris}} \to A_{\mathrm{cris}}$ to be ϕ/p^h. By [Bre98, Lem. 3.1.2.2] the ring $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ is the ring denoted R^{DP} in [Bre99b]. We have an induced filtration on $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ as well as induced maps $\phi_h : \mathrm{Fil}^h(A_{\mathrm{cris}}/pA_{\mathrm{cris}}) \to A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ for $0 \le h \le p-1$. The ring A_{cris} (and so also $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$) has a natural action of G_{K_0} that commutes with all of the above structures. We regard $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ as an $\overline{S}$-algebra by sending the divided power $\gamma_i(u)$ to $\gamma_i(\underline{\pi})$, where $\underline{\pi} \in A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ is the element defined in [Bre99b, §3.2]; then the group G_{K_∞} acts trivially on the image of $\overline{S}$ in $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$.

A.2.1. **Definition.** We define the following functors.

(1) If $M \in \underline{\mathrm{MF}}_k^{f,h}$ for $0 \le h \le p-2$ or $M \in \underline{\mathrm{MF}}_k^{u,p-1}$, we set
$$T(M) = \mathrm{Hom}_{\mathrm{Fil}^\bullet,\phi_\bullet}(M, A_{\mathrm{cris}}/pA_{\mathrm{cris}}),$$
the k-linear morphisms that preserve filtrations and commute with the ϕ_i's. This is a G_{K_0}-representation via $g(f)(x) = g(f(x))$ for $g \in G_{K_0}$, $f \in T(M)$, and $x \in M$.

(2) If $\mathcal{M} \in \underline{\mathcal{M}}_{0,k}^{h}$ for $0 \leq h \leq p-2$ or $M \in \underline{\mathcal{M}}_{0,k}^{u,h}$ with $h = p-1$, we set
$$T(\mathcal{M}) = \mathrm{Hom}_{\mathrm{Fil}^h,\phi_h}(\mathcal{M}, A_{\mathrm{cris}}/pA_{\mathrm{cris}}),$$
the $\overline{S}$-linear morphisms that preserve filtrations Fil^h and commute with ϕ_h. This is a G_∞-representation via $g(f)(x) = g(f(x))$ for $g \in G_\infty$, $f \in T(\mathcal{M})$, and $x \in \mathcal{M}$.

(3) If $\mathfrak{D} \in \underline{\mathfrak{D}}$, we set
$$T(\mathfrak{D}) = \mathrm{Hom}_{\phi}(\mathfrak{D}, k(\!(u)\!)^{\mathrm{sep}}),$$
the $k[[u]]$-linear morphisms that commute with ϕ. This is naturally a G_∞-representation by the theory of *corps des normes* (which gives an isomorphism between $\mathrm{Gal}(k(\!(u)\!)^{\mathrm{sep}}/k(\!(u)\!))$ and G_∞; see [Win83, 3.2.3]).

(4) If $\mathfrak{M} \in \underline{\mathfrak{M}}^h$ for $0 \leq h \leq p-2$ or $\mathfrak{M} \in \underline{\mathfrak{M}}^{u,p-1}$, we set $T(\mathfrak{M}) := T(\mathfrak{M}[\frac{1}{u}])$.

We refer to $T(M)$ (respectively $T(\mathcal{M})$, $T(\mathfrak{D})$, $T(\mathfrak{M})$) as the generic fibre of M (respectively $\mathcal{M}$, $\mathfrak{D}$, $\mathfrak{M}$).

A.2.2. **Proposition.** *We have the following.*

(1) *If $M \in \underline{\mathrm{MF}}_k^{f,h}$ for $0 \leq h \leq p-2$ or $M \in \underline{\mathrm{MF}}_k^{u,p-1}$, then*
$$\dim_k T(M) = \dim_k M.$$

(2) *If $\mathcal{M} \in \underline{\mathcal{M}}_{0,k}^{h}$ for $0 \leq h \leq p-2$ or $\mathcal{M} \in \underline{\mathcal{M}}_{0,k}^{u,p-1}$, then*
$$\dim_k T(\mathcal{M}) = \mathrm{rank}_{\overline{S}}\, \mathcal{M}.$$

(3) *If $\mathfrak{D} \in \underline{\mathfrak{D}}$, then*
$$\dim_k T(\mathfrak{D}) = \dim_{k(\!(u)\!)} \mathfrak{D}.$$

Proof. Parts (1) and (3) are classical: for instance, (1) is [FL82, Thm. 6.1], and (3) is [Fon90, Prop. 1.2.6]. (In (3) the functor T is even an anti-equivalence of categories.)

As for (2), the case $0 \leq h \leq p-2$ is [Bre99b, Lem. 3.2.1]. The proof for $\underline{\mathcal{M}}_{0,k}^{u,p-1}$ is exactly the same, except that the reference to [Wac97, 2.3.2.2] must be supplemented with a reference to [Wac97, Lem. 2.3.3.1]. □

A.3. **Equivalences.** Regard $k[\![u]\!]$ naturally as a subring of $\overline{S}$. Following [Bre99b, §4], if $0 \leq h \leq p-1$ we define a functor $\Theta_h : \underline{\mathfrak{M}}^h \to \underline{\mathcal{M}}_{0,k}^h$ as follows. If $\mathfrak{M}$ is an object of $\underline{\mathfrak{M}}^h$, we define $\Theta_h(\underline{\mathfrak{M}})$ to be the object $\mathcal{M}$ constructed as follows:

- $\mathcal{M} = \overline{S} \otimes_\phi \mathfrak{M}$,
- $\mathrm{Fil}^h \mathcal{M} = \{y \in \mathcal{M} : (\mathrm{Id} \otimes \phi)(y) \in (\mathrm{Fil}^h \overline{S}) \otimes_\phi \mathfrak{M}\}$,
- $\phi_h : \mathrm{Fil}^h \mathcal{M} \to \mathcal{M}$ is the composition of $\mathrm{Id} \otimes \phi : \mathrm{Fil}^h \mathcal{M} \to (\mathrm{Fil}^h \overline{S}) \otimes_\phi \mathfrak{M}$ with $\phi_h \otimes \mathrm{Id} : (\mathrm{Fil}^h \overline{S}) \otimes_\phi \mathfrak{M} \to \overline{S} \otimes_\phi \mathfrak{M} \cong \mathcal{M}$.

A.3.1. **Theorem.** *The functor Θ_h induces an equivalence of categories from $\underline{\mathfrak{M}}^h$ to $\underline{\mathcal{M}}_{0,k}^h$ for $0 \leq h \leq p-2$, as well as an equivalence of categories from $\underline{\mathfrak{M}}^{u,p-1}$ to $\underline{\mathcal{M}}_{0,k}^{u,p-1}$.*

Proof. For $0 \leq h \leq p-2$ this is [Bre99b, Thm. 4.1.1], while the unipotent case is explained in the proof of Theorem 2.5.3 of [Gao]. (Note, however, that some of our terminology is dual to that of [Gao]: what we call a multiplicative object of $\underline{\mathcal{M}}_{0,k}^h$ or $\underline{\mathrm{MF}}_k^{f,h}$, Gao calls étale.) We briefly recall the argument from [Gao]. If A is the matrix of ϕ on $\mathfrak{M}$ with respect to some fixed basis, the condition that $\mathfrak{M}$ is unipotent is precisely the condition that the product $\prod_{n=1}^{\infty} \phi^n(u^{p-1}A^{-1})$ converges to 0, and this is exactly what is required for the proof of full faithfulness in [Bre99b, Thm. 4.1.1] to go through (as well as to show that $\Theta_h(\mathfrak{M})$ is actually unipotent); the essential surjectivity in [Bre99b, Thm. 4.1.1] is true when $h = p-1$ without any unipotence condition. $\square$

A.3.2. **Proposition.** *If $\mathfrak{M} \in \underline{\mathfrak{M}}^{u,p-1}$, we have a canonical isomorphism of G_∞-representations $T(\mathfrak{M}) \cong T(\Theta_{p-1}(\mathfrak{M}))$.*

Proof. The same statement for $\Theta_h : \underline{\mathfrak{M}}^h \to \underline{\mathcal{M}}^h_{0,k}$ with $0 \le h \le p-2$ is [Bre99b, Prop. 4.2.1]. The proof in the unipotent case is identical, noting that the condition $\prod_{n=1}^{\infty} \phi^n(u^{p-1}A^{-1}) = 0$ is precisely what is needed for the last two sentences of the argument in [Bre99b] to go through. $\square$

Following [Bre99b, §5], if $0 \le h \le p-1$ we define a functor $\mathcal{F}_h : \underline{\mathrm{MF}}^{f,h}_k \to \underline{\mathcal{M}}^h_{0,k}$ as follows. If M is an object of $\underline{\mathrm{MF}}^{f,h}_k$, we define $\mathcal{F}_h(M)$ to be the object $\mathcal{M}$ constructed as follows:

- $\mathcal{M} = \overline{S} \otimes_k M$,
- $\mathrm{Fil}^h \mathcal{M} = \sum_{i=0}^h \mathrm{Fil}^i \overline{S} \otimes_k \mathrm{Fil}^{h-i} M$,
- $\phi_h = \sum_{i=0}^h \phi_i \otimes \phi_{h-i}$.

While we expect it to be true that if M is unipotent then so is $\mathcal{F}_{p-1}(M)$, and that the resulting functor $\mathcal{F}^u_{p-1} : \underline{\mathrm{MF}}^{u,p-1}_k \to \underline{\mathcal{M}}^{u,p-1}_{0,k}$ is fully faithful, we will not need these assertions, and so we do not prove them. Instead, we note the following.

A.3.3. **Proposition.** *If $M \in \underline{\mathrm{MF}}^{u,p-1}_k$ and $\mathcal{F}_{p-1}(M)$ is unipotent, then we have a canonical isomorphism of G_∞-representations $T(M)|_{G_\infty} \cong T(\mathcal{F}_{p-1}(M))$.*

Proof. The same argument as in the paragraph before Lemme 5.1 of [Bre99b] (i.e., the proof of [Bre98, 3.2.1.1]) shows that the natural map $T(M) \to T(\mathcal{F}_{p-1}(M))$ is injective. Now the result follows by parts (1) and (2) of Proposition A.2.2 together with the assumption that $\mathcal{F}_{p-1}(M)$ is unipotent. $\square$

Appendix B. Remarks on the geometric Breuil–Mézard philosophy

In this appendix we will explain some variants on the arguments of the main part of the paper, which are either more conceptual, or which avoid the use of results from papers such as [GK12], [BLGG13a], and [GLS12]. Since nothing in the main body of the paper depends on this appendix, and since a full presentation of the arguments would be rather long, we only sketch the proofs.

B.1. **Avoiding the use of results on the weight part of Serre's conjecture.** In this section we will indicate how the main results of the paper (namely, the application of the results of Sections 8–10 to the specific examples of patching functors constructed from spaces of modular forms in Section 6) could be proved without relying on the main results of [GK12], and without using the potential diagonalizability of potentially Barsotti–Tate representations proved in [Kis09b] and [Gee06], but rather just the explicit computations of potentially Barsotti–Tate deformation rings in Section 7. Of course, the results of Section 7 depend on [GK12] (via Theorem 7.1.1, which depends on [GK12] via [EG14]), so we will need to explain how to prove the results of both Sections 6 and 7 (or at least enough of them to prove the main results of Sections 8–10) without using these results. We will still make use of [GLS12] in order to prove Theorem 7.1.1, but we remark that it should be possible to establish all of the results of Section 7 by directly extending the computations of Section 5 of [BM12] to the tame cuspidal case, so it should ultimately be possible to remove this dependence as well.

In fact, if we examine Sections 6 and 7, we see that there are two things that need to be established: Theorem 7.2.1 (which only depends on the above items via Theorem 7.1.1), and the claim that if $\overline{\sigma}$ is a Serre weight and M_∞ is constructed from the spaces of modular forms for a quaternion algebra using the Taylor–Wiles–Kisin method as in Section 6.2, then the action of the universal lifting ring R on $M_\infty(\overline{\sigma})$ factors through $R^{\psi,\overline{\sigma}}$ (or in fact through its reduction mod ϖ_E). (We recall that this is not immediately deducible from local-global compatibility because of parity issues: if σ denotes the algebraic representation of $\mathrm{GL}_2(O_{F_v})$ lifting $\overline{\sigma}$, then the weights of σ may not satisfy the parity condition necessary for constructing a local system on the congruence quotients associated to our given quaternion algebra.) Note that we only need to prove these claims under the assumption that $\overline{\rho}$ is generic (in the local case), or that $\overline{\rho}|_{G_{F_v}}$ is generic for all $v|p$ (in the global case), and we will make this assumption from now on.

The claim about $M_\infty(\overline{\sigma})$ will follow from the results of Section 7, in the following fashion. Suppose first that $\overline{\sigma} \notin \mathcal{D}(\overline{\rho})$; we will show that $M_\infty(\overline{\sigma}) = 0$ by an argument similar to those used in [GK12]. To do this, note that by Lemma 4.5.2 of [GK12] (and its proof), in the Grothendieck group of mod p representations of $\mathrm{GL}_2(k_v)$ we can write $\overline{\sigma}$ as a linear combination $\sum_\tau n_\tau \overline{\sigma}(\tau)$ where τ runs over the tame inertial types (allowing for the moment τ to be the trivial or small Steinberg type). Then if e denotes the Hilbert–Samuel multiplicity of a coherent sheaf on X_∞^ψ, we have

$$e(M_\infty(\overline{\sigma})) = \sum_\tau n_\tau e(M_\infty(\overline{\sigma}(\tau))).$$

By the argument of Lemma 4.3.9 of [GK12], and the irreducibility of the generic fibres of the deformation spaces (see the discussion below), $\sum_\tau n_\tau e(M_\infty(\overline{\sigma}(\tau)))$ is a multiple of $\sum_\tau n_\tau e(R^{\psi,\tau})$ (and the constant of proportionality is independent of $\overline{\sigma}$), so it suffices to check that this last quantity is zero.

To see this, note that the $e(R^{\psi,\tau})$ are completely determined by Theorem 7.2.1 and the observation that since $\overline{\rho}$ is assumed generic, $R^{\psi,\tau} = 0$ if τ is the trivial or Steinberg type (this reduces to checking that the reduction mod p of a semistable (possibly crystalline) Galois representation of Hodge type 0 is not generic, which is clear). In fact, we see that $e(R^{\psi,\tau}) = |\operatorname{JH}(\overline{\sigma}(\tau)) \cap \mathcal{D}(\overline{\rho})|$. Applying the argument of the previous paragraph to *all* $\overline{\sigma}$ simultaneously, and noting that again by Lemma 4.5.2 of [GK12] (and its proof) the $e(M_\infty(\overline{\sigma}(\tau)))$ determine the $e(M_\infty(\overline{\sigma}))$, we see that there must be some constant c such that for any $\overline{\sigma}$, we have $e(M_\infty(\overline{\sigma})) = c$ if $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$, and 0 otherwise. In particular, if $\overline{\sigma} \notin \mathcal{D}(\overline{\rho})$, then $M_\infty(\overline{\sigma}) = 0$, as required.

Now suppose that $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$. Then by Proposition 3.5.1 there is a tame inertial type τ such that $\operatorname{JH}(\overline{\sigma}(\tau)) \cap \mathcal{D}(\overline{\rho}) = \{\overline{\sigma}\}$, so by the result of the previous paragraph we have $M_\infty(\overline{\sigma}^\circ(\tau)) = M_\infty(\overline{\sigma})$ for any lattice $\sigma^\circ(\tau)$ in $\sigma(\tau)$. Also, by Theorem 7.1.1, the reduction mod ϖ_E of $R^{\psi,\overline{\sigma}}$ is equal to the reduction mod ϖ_E of $R^{\psi,\tau}$, so it suffices to note that the action of R on $M_\infty(\overline{\sigma}^\circ(\tau))$ factors through $R^{\psi,\tau}$.

It remains to prove Theorem 7.1.1. Examining the proof of [EG14, Thm. 5.5.4], we see that we have to establish the existence of a suitable globalization in the sense of Section 5.1 of [EG14]), and we have to avoid the use of Lemma 4.4.1

of [GK12] (the potential diagonalizability of potentially Barsotti–Tate representations.). (Note that we still appeal to Lemma 4.4.2 of [GK12], and that this is where we make use of [GLS12].) However, Lemma 4.4.1 of [GK12] is only used to establish that certain patched spaces of modular forms are supported on every component of the generic fibre of each local deformation ring which we consider, and this will be automatic provided we know that these generic fibres are domains. By a standard base change argument, it suffices to know this after a quadratic base change, so we can reduce to the case of tame principal series types, which follows from Lemma 7.4.1 (the proof of which makes no use of Theorem 7.1.1).

It remains to check the existence of a suitable globalization, which amounts to checking Conjecture A.3 of [EG14], the existence of a potentially diagonalizable lift of $\overline{\rho}$. In the proof of [EG14, Thm. 5.5.4], this is done by appealing to results of [GK12], which use the potential diagonalizability of potentially Barsotti–Tate representations. However, as we are assuming that $\overline{\rho}$ is generic, if we choose some weight $\overline{\sigma} \in \mathcal{D}(\overline{\rho})$ then a crystalline lift of $\overline{\rho}$ of Hodge type $\overline{\sigma}$ is Fontaine–Laffaille and thus potentially diagonalizable, as required.

B.1.1. *Remark.* We suspect that the above analysis would make it possible to remove the assumption that when $p = 5$ the projective image of $\overline{\rho}(G_{F(\zeta_5)})$ is not isomorphic to A_5 from our main global theorems, but we have not attempted to check the details.

B.1.2. *Remark.* In particular, if we apply the above analysis to Theorem 9.1.1, we see that we can reprove the main result of [Gee11] without making use of the results on the components of Barsotti–Tate deformation rings proved in [Kis09b], [Gee06], and without needing to make an ordinarity assumption. (Of course, this ordinarity assumption was already removed by [BLGG13b].)

B.2. **Deformation rings and the structure of lattices.** The main theme of this paper is the close connection between the structure of the lattices in tame types for $\mathrm{GL}_2(F_v)$, and the structures of the corresponding potentially Barsotti–Tate deformation rings for generic residual representations. We believe that this connection is even tighter than we have been able to show; in particular, we suspect that it should be possible to use explicit descriptions of the deformation rings to recover Theorem 5.1.1 without making use of the results of [BP12].

Unfortunately, we have not been able to prove this, and our partial results are fragmentary. One difficulty is that in general there will be irregular weights occurring as Jordan–Hölder factors of the reductions of the lattices, and it would be necessary to have an explicit description of the deformation rings for non-generic residual representations. However, even in cases where all the Jordan–Hölder factors are regular, we were still unable to prove complete results. As an illustration of what we were able to prove, we have the following modest result, whose proof will illustrate the kind of arguments that we have in mind.

Let $p > 3$ be prime, let F_v be an unramified extension of $\mathbb{Q}_p$, let τ be a tame type, and consider the lattice $\sigma^\circ_J(\tau)$ for some $J \in \mathcal{P}_\tau$.

B.2.1. **Proposition.** *Suppose that $\overline{\sigma}_{J_1}(\tau), \overline{\sigma}_{J_2}(\tau)$ are two weights in adjacent layers of the cosocle filtration of $\overline{\sigma}^\circ_J(\tau)$, and that there exists a nonsplit extension between $\overline{\sigma}_{J_1}(\tau), \overline{\sigma}_{J_2}(\tau)$ as $\mathrm{GL}_2(F_v)$-representations. Then the extension induced by the cosocle filtration of $\overline{\sigma}^\circ_J(\tau)$ is also nonsplit.*

Proof. By an explicit (but tedious) computation "dual" to that used in the proof of Proposition 10.1.11, the existence of a nonsplit extension between $\overline{\sigma}_{J_1}(\tau), \overline{\sigma}_{J_2}(\tau)$ implies that there is a semisimple generic representation $\overline{\rho}$ such that $\mathrm{JH}(\overline{\sigma}(\tau)) \cap \mathcal{D}(\overline{\rho}) = \{\overline{\sigma}_{J_1}(\tau), \overline{\sigma}_{J_2}(\tau)\}$. Since $\overline{\rho}$ is semisimple, it is easy to globalise it as the local mod p representation corresponding to some CM Hilbert modular form, and the constructions of Section 6.5 then give a minimal fixed determinant patching functor $M_\infty^{\min}$ with unramified coefficients $\mathcal{O}$, indexed by $(F_v, \overline{\rho})$.

As was already noted in Subsection 5.2, it follows from Lemmas 4.1.1 and 3.1.1 that $\sigma_J^\circ(\tau)$ is defined over the ring of integers of an unramified extension of $\mathbb{Q}_p$, so extending scalars if necessary, we can assume that it is defined over $\mathcal{O}$. By Theorem 7.2.1, and the assumptions that $|\,\mathrm{JH}(\overline{\sigma}(\tau)) \cap \mathcal{D}(\overline{\rho})| = 2$ and $\mathcal{O}$ is unramified, we see that the deformation space $X^\psi\big(\tau\big)$ is regular, so by Lemma 6.1.4 we see that $M_\infty^{\min}(\sigma_J^\circ(\tau))$ is free of rank one over $X_\infty^\psi\big(\tau\big)$, and thus in particular $M_\infty^{\min}(\overline{\sigma}_J^\circ(\tau))$ is free of rank one over $\overline{X}_\infty^\psi\big(\tau\big)$.

Now, if the extension between $\overline{\sigma}_{J_1}(\tau), \overline{\sigma}_{J_2}(\tau)$ induced by the cosocle filtration on $\overline{\sigma}_J^\circ(\tau)$ were split, we would have

$$M_\infty^{\min}(\overline{\sigma}_J^\circ(\tau)) = M_\infty^{\min}(\overline{\sigma}_{J_1}(\tau)) \oplus M_\infty^{\min}(\overline{\sigma}_{J_1}(\tau))$$

which would not be free of rank one, which is a contradiction; so the extension must be nonsplit, as claimed. $\square$

References

[BD14] Christophe Breuil and Fred Diamond, *Formes modulaires de Hilbert modulo p et valeurs d'extensions Galoisiennes*, to appear in Ann. Scient. de l'E.N.S. Volume 47, fascicule 4, 2014.

[BDJ10] Kevin Buzzard, Fred Diamond, and Frazer Jarvis, *On Serre's conjecture for mod l Galois representations over totally real fields*, Duke Math. J. **155** (2010), no. 1, 105–161.

[BL94] L. Barthel and R. Livné, *Irreducible modular representations of* GL_2 *of a local field*, Duke Math. J. **75** (1994), no. 2, 261–292.

[BLGG13a] Thomas Barnet-Lamb, Toby Gee, and David Geraghty, *Serre weights for rank two unitary groups*, Math. Ann. **356** (2013), no. 4, 1551–1598.

[BLGG13b] Tom Barnet-Lamb, Toby Gee, and David Geraghty, *Congruences between Hilbert modular forms: constructing ordinary lifts, II*, Math. Research Letters (to appear) (2013).

[BLGGT14] Tom Barnet-Lamb, Toby Gee, David Geraghty, and Richard Taylor, *Potential automorphy and change of weight*, Annals of Math. (2) **179** (2014), no. 2, 501–609.

[BM02] Christophe Breuil and Ariane Mézard, *Multiplicités modulaires et représentations de* $\mathrm{GL}_2(\mathbf{Z}_p)$ *et de* $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ *en* $l = p$, Duke Math. J. **115** (2002), no. 2, 205–310, With an appendix by Guy Henniart.

[BM12] Christophe Breuil and Ariane Mézard, *Mulitplicités modulaires raffinées*, Bull. Soc. Math. de France. (to appear) (2012).

[BP12] Christophe Breuil and Vytautas Paškūnas, *Towards a modulo p Langlands correspondence for* GL_2, Memoirs of the American Mathematical Society **216** (2012), no. 1016.

[Bre98] Christophe Breuil, *Construction de représentations p-adiques semi-stables*, Ann. Sci. École Norm. Sup. (4) **31** (1998), no. 3, 281–327.

[Bre99a] ______, *Représentations semi-stables et modules fortement divisibles*, Invent. Math. **136** (1999), no. 1, 89–122.

[Bre99b] ______, *Une application de corps des normes*, Compositio Math. **117** (1999), no. 2, 189–203.

[Bre00] ______, *Groupes p-divisibles, groupes finis et modules filtrés*, Ann. of Math. (2) **152** (2000), no. 2, 489–549.

[Bre03] ———, *Sur quelques représentations modulaires et p-adiques de* $\mathrm{GL}_2(\mathbf{Q}_p)$*. I*, Compositio Math. **138** (2003), no. 2, 165–188.

[Bre12] ———, *Sur un problème de compatibilité local-global modulo p pour* GL_2, J. Reine Angew. Math. (to appear) (2012).

[Car94] Henri Carayol, *Formes modulaires et représentations galoisiennes à valeurs dans un anneau local complet*, p-adic monodromy and the Birch and Swinnerton-Dyer conjecture (Boston, MA, 1991), Contemp. Math., vol. 165, Amer. Math. Soc., Providence, RI, 1994, pp. 213–237.

[CEG$^+$a] Ana Caraiani, Matthew Emerton, Toby Gee, David Geraghty, Vytautas Paškūnas, and Sug Woo Shin, *Local and global aspects of the p-adic Langlands program*, in preparation.

[CEG$^+$b] ———, *Patching and the p-adic local Langlands correspondence*, in preparation.

[CHT08] Laurent Clozel, Michael Harris, and Richard Taylor, *Automorphy for some l-adic lifts of automorphic mod l Galois representations*, Pub. Math. IHES **108** (2008), 1–181.

[Col10] Pierre Colmez, *Représentations de* $\mathrm{GL}_2(\mathbf{Q}_p)$ *et* (ϕ,Γ)*-modules*, Astérisque (2010), no. 330, 281–509.

[DDT97] Henri Darmon, Fred Diamond, and Richard Taylor, *Fermat's last theorem*, Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993), Int. Press, Cambridge, MA, 1997, pp. 2–140.

[Dem] Lassina Dembélé, Appendix B to [Bre12].

[Dia97] Fred Diamond, *The Taylor-Wiles construction and multiplicity one*, Invent. Math. **128** (1997), no. 2, 379–391.

[Dia07] ———, *A correspondence between representations of local Galois groups and Lie-type groups*, L-functions and Galois representations, London Math. Soc. Lecture Note Ser., vol. 320, Cambridge Univ. Press, Cambridge, 2007, pp. 187–206.

[DT94] Fred Diamond and Richard Taylor, *Nonoptimal levels of mod l modular representations*, Invent. Math. **115** (1994), no. 3, 435–462.

[EG14] Matthew Emerton and Toby Gee, *A geometric perspective on the Breuil-Mézard conjecture*, J. Inst. Math. Jussieu **13** (2014), no. 1, 183–223.

[Eme] Matthew Emerton, *Local-global compatibility in the p-adic Langlands programme for* $\mathrm{GL}_2/\mathbb{Q}$.

[FL82] Jean-Marc Fontaine and Guy Laffaille, *Construction de représentations p-adiques*, Ann. Sci. École Norm. Sup. (4) **15** (1982), no. 4, 547–608 (1983).

[Fon90] Jean-Marc Fontaine, *Représentations p-adiques des corps locaux. I*, The Grothendieck Festschrift, Vol. II, Progr. Math., vol. 87, Birkhäuser Boston, Boston, MA, 1990, pp. 249–309.

[Gao] Hui Gao, *Galois lattices and strongly divisible lattices in the unipotent case*, Preprint, 2012.

[Gee06] Toby Gee, *A modularity lifting theorem for weight two Hilbert modular forms*, Math. Res. Lett. **13** (2006), no. 5-6, 805–811.

[Gee11] ———, *On the weights of mod p Hilbert modular forms*, Invent. Math. **184** (2011), no. 1, 1–46.

[GK12] Toby Gee and Mark Kisin, *The Breuil–Mézard conjecture for potentially Barsotti–Tate representations*, 2012.

[GLS12] Toby Gee, Tong Liu, and David Savitt, *The Buzzard–Diamond–Jarvis conjecture for unitary groups*, 2012.

[Gow76] R. Gow, *Schur indices of some groups of Lie type*, J. Algebra **42** (1976), no. 1, 102–120.

[Kis09a] Mark Kisin, *The Fontaine-Mazur conjecture for* GL_2, J. Amer. Math. Soc. **22** (2009), no. 3, 641–690.

[Kis09b] ———, *Moduli of finite flat group schemes, and modularity*, Annals of Math.(2) **170** (2009), no. 3, 1085–1180.

[Mat89] Hideyuki Matsumura, *Commutative ring theory*, second ed., Cambridge Studies in Advanced Mathematics, vol. 8, Cambridge University Press, Cambridge, 1989, Translated from the Japanese by M. Reid.

[Paš13] Vytautas Paškūnas, *The image of Colmez's Montreal functor*, Publ. Math. Inst. Hautes Études Sci. **118** (2013), 1–191.

[Sav05] David Savitt, *On a conjecture of Conrad, Diamond, and Taylor*, Duke Math. J. **128** (2005), no. 1, 141–197.

[Sav08] ______, *Breuil modules for Raynaud schemes*, J. Number Theory **128** (2008), no. 11, 2939–2950.

[Tay08] Richard Taylor, *Automorphy for some l-adic lifts of automorphic mod l Galois representations. II*, Pub. Math. IHES **108** (2008), 183–239.

[Tho12] Jack Thorne, *On the automorphy of l-adic Galois representations with small residual image*, J. Inst. Math. Jussieu **11** (2012), no. 4, 855–920, With an appendix by Robert Guralnick, Florian Herzig, Richard Taylor and Thorne.

[TW95] Richard Taylor and Andrew Wiles, *Ring-theoretic properties of certain Hecke algebras*, Ann. of Math. (1995), no. 141, 553–572.

[Wac97] Nathalie Wach, *Représentations cristallines de torsion*, Compositio Math. **108** (1997), no. 2, 185–240.

[Win83] Jean-Pierre Wintenberger, *Le corps des normes de certaines extensions infinies de corps locaux; applications*, Ann. Sci. École Norm. Sup. (4) **16** (1983), no. 1, 59–89.

E-mail address: emerton@math.uchicago.edu

Department of Mathematics, University of Chicago, 5734 S. University Ave., Chicago, IL 60637

E-mail address: toby.gee@imperial.ac.uk

Department of Mathematics, Imperial College London

E-mail address: savitt@math.arizona.edu

Department of Mathematics, University of Arizona