GENERAL SERRE WEIGHT CONJECTURES

TOBY GEE, FLORIAN HERZIG, AND DAVID SAVITT

ABSTRACT. We formulate a number of related generalisations of the weight part of Serre's conjecture to the case of GL_n over an arbitrary number field, motivated by the formalism of the Breuil–Mézard conjecture. We give evidence for these conjectures, and discuss their relationship to previous work. We generalise one of these conjectures to the case of connected reductive groups which are unramified over $\mathbb{Q}_p$, and we also generalise the second author's previous conjecture for $GL_n/\mathbb{Q}$ to this setting, and show that the two conjectures are generically in agreement.

CONTENTS

1. INTRODUCTION

The goal of this paper is to formulate a number of related generalisations of the weight part of Serre's conjecture to the case of GL_n over an arbitrary number field. Since this is a problem with a long and involved history and since we work in significant generality in this paper, we begin with an extended introduction, in which we try to summarize this history (Sections 1.1 to 1.5) and give a detailed overview of the approach that we have taken (Sections 1.6 and 1.7).

1.1. **Serre's conjecture for GL_2 over $\mathbb{Q}$.** Let p be a prime. Serre's conjecture, as originally formulated in 1973 (see [SD73, p. 9] and [Ser75, §3]), predicted that every odd irreducible continuous representation $\bar{r} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ which is unramified outside p has a twist by a power of the mod p cyclotomic character which arises from a cuspidal modular Hecke eigenform of level 1 and weight at most $p+1$. The theory of the θ-operator then implies that $\bar{r}$ itself is modular of weight at most p^2-1. This was a bold conjecture, for at the time there was little evidence outside of the cases $p = 2, 3$. In those cases, since there are no cusp forms of level 1 and weight less than 12 the conjecture simply predicts that there are no such representations. This can be established via discriminant bounds, as in [Tat94] and [Ser86, p. 710].

Serre later formulated ([Ser87]) a version of the conjecture with no restriction on the ramification of $\bar{r}$, which included a precise recipe for both the weight and the level of a modular eigenform giving rise to $\bar{r}$. In this way the conjecture became computationally verifiable, and was tested in a number of cases in which $\bar{r}$ has small image.

At least as far back as [Ser75], it had been known that in the theory of mod p modular forms, one can trade off the weight and level (and Nebentypus) at p. For this reason Serre restricted his attention to modular forms of level prime to p. He conjectured that the minimal possible level of the candidate eigenform giving rise to $\bar{r}$ could be taken to be the prime-to-p Artin conductor of $\bar{r}$, while his conjectural recipe for the minimal possible weight of the eigenform (in prime-to-p-level) was more intricate, and depended on the ramification behaviour of $\bar{r}$ at p.

The part of Serre's conjecture which predicts that every odd irreducible continuous representation $\bar{r} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ arises from some modular eigenform is often referred to as "the weak form of Serre's conjecture", while the form of the conjecture that includes the precise recipes for the minimal weight and level is called "the strong form of Serre's conjecture". Much of the early work concerning Serre's conjecture was focussed on proving that the weak form implies the strong form, and it is natural to expect that work on generalisations of Serre's conjecture will follow the same pattern. (Indeed, the eventual proof of Serre's conjecture [KW09a, KW09b, Kis09b] relied on the work that had been done to prove the equivalence of the weak and strong forms.)

Serre's conjectural recipe for the minimal weight of an eigenform of prime-to-p level giving rise to $\bar{r}$ was more subtle than the recipe for the level, but essentially amounted to providing the minimal weight k that was consistent with the known properties of the restriction to a decomposition group at p of the Galois representations associated to eigenforms. To make this precise one nowadays uses the language of p-adic Hodge theory. Given a modular eigenform f of weight $k \geq 2$ and prime-to-p level, the associated p-adic Galois representation $r_f : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ has the property that the restriction $r_f|_{G_{\mathbb{Q}_p}}$ to a decomposition group at p is crystalline with Hodge–Tate weights $k-1$ and 0. Therefore any results on the reduction mod p of crystalline representations of $G_{\mathbb{Q}_p}$ with Hodge–Tate weights $k-1$ and 0, such as the early results of Deligne and Fontaine–Serre when $k \leq p$, give information (purely in terms of $\bar{r}|_{G_{\mathbb{Q}_p}}$) about the possible weights k of the modular eigenforms giving rise to $\bar{r}$.

To give a concrete example, let ε denote the cyclotomic character, and $\bar{\varepsilon}$ its reduction mod p. Suppose that

$$(1.1.1) \qquad \bar{r}|_{I_{\mathbb{Q}_p}} \cong \begin{pmatrix} \bar{\varepsilon}^{k-1} & * \\ 0 & 1 \end{pmatrix}$$

where $I_{\mathbb{Q}_p}$ is the inertia group at p, and $2 < k < p + 1$. Then the minimal weight predicted by Serre's recipe is k. Indeed, it is known that any crystalline representation $\rho : G_{\mathbb{Q}_p} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$ with Hodge–Tate weights $k - 1$ and 0 (with k in the given range) and whose reduction mod p is reducible must be an extension of an unramified character by an unramified twist of ε^{k-1}, and therefore the shape of $\bar{\rho}|_{I_{\mathbb{Q}_p}}$ must be as on the right-hand side of (1.1.1).

We make one further remark about the above example. Suppose that the extension class $*$ vanishes, and assume for simplicity that $k < p - 1$. Serre observed that

$$(1.1.2) \qquad (\bar{r} \otimes \bar{\varepsilon}^{1-k})|_{I_{\mathbb{Q}_p}} \cong \begin{pmatrix} \bar{\varepsilon}^{p-k} & 0 \\ 0 & 1 \end{pmatrix}$$

and therefore has minimal weight $p + 1 - k$. Thus, although Serre's conjecture predicts that any $\bar{r}$ has a twist which is modular with weight at most $p + 1$, in this split case there are actually two such twists. This is the so-called "companion forms" phenomenon.

1.2. **Serre weights.** We now explain a representation-theoretic reformulation of the weight k in Serre's conjecture. This optic first appears in the work of Ash–Stevens [AS86], and both simplifies the original weight recipe for GL_2 over $\mathbb{Q}$ and has proved crucial for formulating the weight part of Serre's conjecture for other groups and over other fields.

The Eichler–Shimura isomorphism allows one to reinterpret Serre's conjecture in terms of the cohomology of arithmetic groups. If V is an $\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_2(\mathbb{F}_p)$ and N is prime to p, then we have a natural action of the Hecke algebra of $\Gamma_1(N)$ on $H^1(\Gamma_1(N), V)$, and so it makes sense to speak of a continuous representation $\bar{r} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ being associated to an eigenclass in that cohomology group. If $\bar{r}$ is odd and irreducible, then the Eichler–Shimura isomorphism implies that $\bar{r}$ is modular of weight k and prime-to-p level N if and only if $\bar{r}$ is associated to an eigenclass in $H^1(\Gamma_1(N), \mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2)$, where $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$ is the $(k-2)$th symmetric power of the standard representation of $\mathrm{GL}_2(\mathbb{F}_p)$ on $\overline{\mathbb{F}}_p^2$. By dévissage one deduces that $\bar{r}$ is modular of weight k and prime-to-p level N if and only if $\bar{r}$ is associated to an eigenclass in $H^1(\Gamma_1(N), V)$ for some Jordan–Hölder factor V of $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$. (Recall that the representation $\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2$ is reducible as soon as $k > p + 1$.)

It is then natural to associate to $\bar{r}$ the set $W(\bar{r})$ of irreducible $\overline{\mathbb{F}}_p$-representations V of $\mathrm{GL}_2(\mathbb{F}_p)$ such that $\bar{r}$ is associated to an eigenclass in $H^1(\Gamma_1(N), V)$ for some prime-to-p level N. Thanks to the argument in the previous paragraph, the (finite) set $W(\bar{r})$ determines all weights in which $\bar{r}$ occurs in prime-to-p level, and not just the minimal such weight. For this reason such representations of $\mathrm{GL}_2(\mathbb{F}_p)$ are now often referred to as *Serre weights*, or even simply *weights*.

To illustrate, suppose once again that $\bar{r}$ is as in (1.1.1), with $2 < k < p - 1$. If the extension class $*$ is non-split, then we have $W(\bar{r}) = \{\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2\}$. However, in the companion forms case where the extension class $*$ is split, we have

$$W(\bar{r}) = \{\mathrm{Sym}^{k-2}\overline{\mathbb{F}}_p^2, \det^{k-1} \otimes \mathrm{Sym}^{p-1-k}\overline{\mathbb{F}}_p^2\}$$

instead. Here the second weight comes from observing via (1.1.2) that the weight $\text{Sym}^{p-1-k}\,\overline{\mathbb{F}}_p^2$ should lie in $W(\bar{r} \otimes \bar{\varepsilon}^{1-k})$, and then undoing the twist.

Serre in fact asked [Ser87, §3.4] whether a "mod-p Langlands philosophy" exists which would give a more natural definition of the weight, and which would allow for generalisations of the conjecture to other groups and number fields. This is now known to be true for GL_2 over $\mathbb{Q}$ ([Col10, Eme10]) and the set $W(\bar{r})$ intervenes naturally from this point of view (see for example [Bre10]). There is considerable evidence that such a philosophy remains true in more general settings, although it is far from completely developed at this point. Indeed the results to date on generalisations of the weight part of Serre's conjecture have been a major guiding influence on the development of the mod p Langlands program, rather than a consequence of it.

1.3. **Early generalisations.** Formulations of very general versions of the weak conjecture have been known to the experts for many years; the main issue is to define the correct generalisation of "odd", for which see for example [Gro07] and [BV13, §6]. (If one wishes to consider automorphic forms or cohomology classes for groups which are not quasi-split, it is also necessary to impose conditions on the ramification of $\bar{r}$ at places at which the underlying group is ramified; see [GK14, Def. 4.5.3] for the case of quaternion algebras.) Moreover, granting an understanding of classical local Langlands and its relationship to local-global compatibility, it is reasonably straightforward to generalise the definition of the (prime-to-p) level in terms of the prime-to-p ramification of $\bar{r}$. For example, for generalisations to GL_n over arbitrary number fields, one again expects to take the level to be the prime-to-p Artin conductor of $\bar{r}$; see e.g. [ADP02, §2.2] for the case that the number field is $\mathbb{Q}$.

However, formulating the weight part of the conjecture in any generality has proved difficult. We stress at the outset that, in keeping with the mod p Langlands philosophy, one conjectures that the set of Serre weights associated to $\bar{r}$ depends only on the restrictions of $\bar{r}$ to decomposition groups at places dividing p. For this reason all of the weight predictions that we discuss in this paper are formulated in terms of local Galois representations.

For Hilbert modular forms over a totally real field F in which p is unramified, a precise conjecture was formulated by Buzzard–Diamond–Jarvis in [BDJ10]. It was essential for [BDJ10] to use the "Serre weight" point of view, since weights of Hilbert modular forms are $[F : \mathbb{Q}]$-tuples of integers and so there isn't a natural notion of minimal weight of $\bar{r}$. In this context a Serre weight is an irreducible $\overline{\mathbb{F}}_p$-representation of $\prod_{v|p} \text{GL}_2(k_v)$, where k_v is the residue field of the completion F_v. The recipe of [BDJ10] predicts the set of weights $W(\bar{r})$ in terms of the Hodge–Tate weights of crystalline lifts of $\bar{r}|_{G_{F_v}}$ for $v \mid p$, in line with the discussion at the end of Section 1.1. The prediction of [BDJ10] is now known to be correct [GLS14, GK14].

In another direction, the study of the weight part of Serre's conjecture for GL_n over $\mathbb{Q}$ was initiated by Ash and his collaborators [AS00, ADP02], with a particular focus on GL_3. They gave a combinatorial recipe for a predicted set of weights, in the spirit of Serre's original recipe but using the language of Serre weights. The combinatorial recipe takes as input the tame inertia weights of $\bar{r}|_{I_{\mathbb{Q}_p}}$ (the base p

"digits" of the exponents when $\bar{r}|_{I_{\mathbb{Q}_p}}$ is written as a successive extension of powers of fundamental characters), much as in the examples (1.1.1), (1.1.2) and their reformulations in Section 1.2.

In the case where $\bar{r}|_{G_{\mathbb{Q}_p}}$ is semisimple, the thesis [Her09] of the second-named author gave a representation-theoretic recipe for a predicted set of weights, which for generic $\bar{r}|_{G_{\mathbb{Q}_p}}$ should be the full set of weights. The prediction is made in terms of the reduction mod p of Deligne–Lusztig representations, and involves a mysterious involution $\mathcal{R}$ on the set of Serre weights. In particular [Her09] predicts some weights for GL_3 that are not predicted by [ADP02], and that were subsequently computationally confirmed (in some concrete cases) by Doud and Pollack. (We stress that [AS00, ADP02] did not claim to predict the full set of weights for $\bar{r}$.)

1.4. **The Breuil–Mézard conjecture.** We now turn to the Breuil–Mézard conjecture, which gives a new way of looking at the weight part of Serre's conjecture.

Originally the Breuil–Mézard conjecture arose in the context of attempts to generalise the Taylor–Wiles method [TW95], and was also one of the starting points of the p-adic Langlands program. It was clear early on that understanding the geometry of deformation spaces of local mod p Galois representations with prescribed p-adic Hodge-theoretic conditions was essential for proving automorphy lifting theorems; the earliest automorphy lifting theorems required the smoothness of such deformation spaces. The Breuil–Mézard conjecture gives a measure of the complexity of these deformation spaces, in terms of the modular representation theory of GL_2.

We state a version of this conjecture for GL_n over $\mathbb{Q}_p$, following [EG14]. We need the following data and terminology:

- a continuous representation $\bar{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$,
- a *Hodge type* λ, which in this setting is an n-tuple $\lambda = (\lambda_1, \ldots, \lambda_n)$ of integers with $\lambda_1 \geq \cdots \geq \lambda_n$, and
- an *inertial type* τ, i.e. a representation $I_{\mathbb{Q}_p} \to \mathrm{GL}_n(\overline{\mathbb{Q}}_p)$ with open kernel and that can be extended to a representation of $G_{\mathbb{Q}_p}$.

Kisin [Kis08] associates to this data a lifting ring $R_{\bar{\rho}}^{\lambda,\tau}$ whose characteristic 0 points parameterise the lifts of $\bar{\rho}$ that are potentially crystalline with type τ and Hodge–Tate weights

$$(1.4.1) \qquad\qquad \lambda_1 + n - 1, \ldots, \lambda_{n-1} + 1, \lambda_n.$$

The Breuil–Mézard conjecture predicts the Hilbert–Samuel multiplicity $e(R_{\bar{\rho}}^{\lambda,\tau} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p)$ of $R_{\bar{\rho}}^{\lambda,\tau} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$, as follows.

The inertial local Langlands correspondence (cf. Henniart's appendix to [BM02]) associates to τ a finite-dimensional smooth $\overline{\mathbb{Q}}_p$-representation $\sigma(\tau)$ of $\mathrm{GL}_n(\mathbb{Z}_p)$. On the other hand associated to λ is the irreducible algebraic representation $W(\lambda)$ of $\mathrm{GL}_n(\mathbb{Q}_p)$ of highest weight λ.

Conjecture 1.4.2 (The generalised Breuil–Mézard conjecture)**.** *There exist nonnegative integers $\mu_V(\bar{\rho})$, indexed by Serre weights V, such that for all Hodge types λ and inertial types τ we have*

$$e(R_{\bar{\rho}}^{\lambda,\tau} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) = \sum_V n_{\lambda,\tau}(V)\mu_V(\bar{\rho})$$

where $n_{\lambda,\tau}(V)$ is the multiplicity of V in the reduction modulo p of $W(\lambda) \otimes_{\overline{\mathbb{Q}}_p} \sigma(\tau)$ (as a $\mathrm{GL}_n(\mathbb{Z}_p)$-representation).

This conjecture was first formulated by Breuil–Mézard [BM02] for GL_2 with certain restrictions on λ and τ. In the special case where the Serre weight V is actually a Weyl module, and therefore lifts to some $W(\lambda)$ in characteristic zero, taking τ trivial in Conjecture 1.4.2 gives an equality $\mu_V(\overline{\rho}) = e(R_{\overline{\rho}}^{\lambda,\mathrm{triv}} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p)$. For this reason we typically refer to the integers $\mu_V(\overline{\rho})$ as multiplicities.

Based on some explicit calculations, Breuil and Mézard furthermore gave predictions for the multiplicities $\mu_V(\overline{\rho})$, and observed a close connection between these multiplicities and the weight part of Serre's original modularity conjecture: namely, that one appeared to have $\mu_V(\overline{r}|_{G_{\mathbb{Q}_p}}) > 0$ if and only if V was a predicted Serre weight for $\overline{r}$.

More recently, the first-named author and Kisin [GK14] suggested that one could turn this around and use the Breuil–Mézard conjecture to define the set of Serre weights $W_{\mathrm{BM}}(\overline{\rho}) = \{V : \mu_V(\overline{\rho}) > 0\}$ associated to a local Galois representation $\overline{\rho}$. One would then conjecture that $W(\overline{r}) = W_{\mathrm{BM}}(\overline{r}|_{G_{\mathbb{Q}_p}})$.

Note that this prediction for the set of Serre weights associated to $\overline{r}$, while very general, is contingent on the truth of the Breuil–Mézard conjecture. In fact, what [GK14] actually do is prove the Breuil–Mézard conjecture for GL_2 and $\lambda = 0$ (for arbitrary $K/\mathbb{Q}_p$), which allows them unconditionally to define a set of weights $W_{\mathrm{BT}}(\overline{\rho})$. Here BT stands for Barsotti–Tate. This description of the set of weights turns out to be extremely useful, and was an important part of the resolution in [GK14] and [GLS15] of the conjectures of [BDJ10] and their generalisations to arbitrary totally real fields.

The key technique used by [GK14] is the method of Taylor–Wiles–Kisin patching. One first constructs a globalisation $\overline{r}$ of $\overline{\rho}$. Write $X_\infty = \mathrm{Spf}\, R_{\overline{\rho}}[[x_1, \ldots, x_h]]$, with $R_{\overline{\rho}}$ the universal lifting ring of $\overline{\rho}$ and $h \geq 0$ a certain integer. Similarly write $X_\infty^\tau = \mathrm{Spf}\, R_{\overline{\rho}}^{0,\tau}[[x_1, \ldots, x_h]]$, which if non-empty is of dimension $d + 1$ for some d independent of τ. In the context of [GK14] a *patching functor* is a non-zero covariant exact functor M_∞ from the category of finitely generated $\overline{\mathbb{Z}}_p$-modules with a continuous action of $\mathrm{GL}_2(\mathcal{O}_K)$, to the category of coherent sheaves on X_∞, with the properties that:

- for all inertial types τ the sheaf $M_\infty(\sigma(\tau))$ is p-torsion free and has scheme-theoretic support X_∞^τ, and in fact is maximal Cohen–Macaulay over X_∞^τ;
- the (maximal Cohen–Macaulay over a regular scheme, so) locally free sheaf $M_\infty(\sigma(\tau))[1/p]$ has rank one over the generic fibre of X_∞^τ, and
- for all Serre weights V, the support of the sheaf $M_\infty(V)$ either has dimension d or is empty.

This is an abstraction of the output of the Taylor–Wiles–Kisin patching method applied to spaces of automorphic forms. The existence of a patching functor can be shown to imply that the Breuil–Mézard conjecture holds (in the cases under consideration in [GK14]), and moreover that $W_{\mathrm{BT}}(\overline{\rho})$ is precisely the set of weights V for which $M_\infty(V) \neq 0$. On the other hand, [GK14] construct such a functor, and the construction implies that $M_\infty(V) \neq 0$ if and only if $\overline{r}$ is automorphic of weight V. Putting these together, [GK14] conclude that the set $W_{\mathrm{BT}}(\overline{\rho})$ is indeed the correct weight set for $\overline{r}$.

1.5. **Shadow weights and the crystalline lifts conjecture.** One of the features of the weight part of Serre's conjecture for GL_n ($n \geq 3$) that distinguishes it from the GL_2 case is that there exist Serre weights that do not lift to characteristic 0. For example, for GL_3 over $\mathbb{Q}$ roughly half the Serre weights are so-called "upper alcove weights". These have the property that if W is the irreducible representation in characteristic 0 with the same highest weight as an upper alcove weight U, then the reduction mod p of W is not irreducible but rather has two Jordan–Hölder factors, one of which is U and another which we denote by L (for "lower alcove").

It was observed in the conjecture of [Her09] (in the semisimple case) as well as in the computations of [ADP02] (including some non-semisimple examples) that whenever L was a predicted Serre weight for some $\bar{r}$, so also was U. For this reason one began to refer to the weight U as a *shadow* of the weight L. The conjecture that U occurs in the set of Serre weights of $\bar{r}$ whenever L does (as well as its natural generalisation to the GL_n setting) became known as the shadow weight conjecture.

In the optic of the Breuil–Mézard conjecture, the shadow weight conjecture says that if $\mu_L(\bar{\rho}) > 0$ for some local Galois representation $\bar{\rho}$ and if U is a shadow of L, then also $\mu_U(\bar{\rho}) > 0$. The Breuil–Mézard conjecture itself implies that if $\bar{\rho}$ has Serre weight U then $\bar{\rho}$ has a crystalline lift with Hodge–Tate weights corresponding to the highest weight of U, or equivalently to the highest weight of W; and conversely that if $\bar{\rho}$ has such a crystalline lift, then $\bar{\rho}$ has at least one Serre weight that occurs in the reduction of W. In combination with the shadow weight conjecture, this is elevated to an if-and-only-if: that $\bar{\rho}$ has Serre weight U if and only if a crystalline lift of $\bar{\rho}$ with Hodge–Tate weights as above exists.

This attractive picture (as well as its generalisation to GL_n over more general number fields) was known as the crystalline lifts version of the weight part of Serre's conjecture, and was widely believed for a number of years. For the sake of historical accuracy, we should remark that the crystalline lifts version of the weight part of Serre's conjecture emerged [Gee11, §4] before the Breuil–Mézard optic, motivated by its evident parallels with the GL_2 case (both Serre's original conjecture and the conjecture of [BDJ10]) and its compatibility with the conjectures of [Her09].

The crystalline lifts version of the weight part of Serre's conjecture was contained in drafts of the present paper as recently as 2014. We had a narrow escape, then, when (prior to the completion of this paper) Le, Le Hung, Levin, and Morra [LLHLM15] produced counterexamples to the shadow weight conjecture for GL_3 over $\mathbb{Q}$ (in the non-semisimple case), thus also disproving the crystalline lifts version of the weight part of Serre's conjecture for GL_3.

The geometric explanation seems to be as follows. The papers [EG15, EG] construct a finite type equidimensional Artin stack $\overline{\mathcal{X}}$ over $\mathbb{F}_p$ whose $\overline{\mathbb{F}}_p$-points naturally correspond to the isomorphism classes of representations $\bar{\rho} : G_{\mathbb{Q}_p} \to GL_3(\overline{\mathbb{F}}_p)$. The stack $\overline{\mathcal{X}}$ should have among its irreducible components $\overline{\mathcal{X}}(U)$ and $\overline{\mathcal{X}}(L)$, whose $\overline{\mathbb{F}}_p$-points are precisely the representations $\bar{\rho}$ for which U and L respectively are Serre weights, and these components appear to intersect in codimension one. Since [ADP02] make computations for representations $\bar{r}$ which by construction have small image, it is not surprising in hindsight that those representations might lie in special loci of $\overline{\mathcal{X}}$.

1.6. **This paper.** In this paper, we explain a general formulation of the weight part of Serre's conjecture (Conjecture 3.2.7) in terms of the Breuil–Mézard conjecture,

based on the philosophy outlined in Section 1.4. Moreover, there are compelling reasons (coming from the Fontaine–Mazur conjecture and the Taylor–Wiles method) to believe that this recipe gives the correct weights in full generality. In particular, in Proposition 4.2.1 we prove that the existence of a suitable patching functor would on the one hand imply the Breuil–Mézard conjecture, and would on the other hand imply that the set $W_{\mathrm{BM}}(\overline{\rho})$ is the set of Serre weights of globalisations of $\overline{\rho}$.

Although we believe this description of the weights is the "correct" one, and it seems likely that any proof of the weight part of Serre's conjecture in general situations will need to make use of this formulation, it is of interest to have more explicit descriptions of the set of weights. For a variety of reasons (which we discuss in the body of the paper), it seems unlikely that in general there will be explicit and complete descriptions of the sort that one finds for GL_2 in [GLS14, GLS15, DDR16, CEGM17], but it does seem reasonable to hope for something more concrete in the case that $\overline{\rho}$ is semisimple and suitably generic.

Indeed, it remains plausible that the crystalline lifts version of the weight part of Serre's conjecture, despite being false in general, is nevertheless true in the case where $\bar{r}$ is semisimple locally at places above p. For instance, there is considerable evidence in the 3-dimensional case over $\mathbb{Q}$: when $\bar{r}|_{G_{\mathbb{Q}_p}}$ is suitably generic many cases of the conjecture are proved in [EGH13] and [LLHLM15], and for some non-generic $\bar{r}|_{G_{\mathbb{Q}_p}}$ there is computational evidence due to [ADP02]. The more recent papers [LLHLM16, LLHL16] extend the results of [LLHLM15] to the case of totally real fields in which p is unramified, and establish *weight elimination* (that the set of modular weights is a subset of the set of predicted weights) in arbitrary dimension in this setting, again with a genericity hypothesis on $\bar{r}$ locally above p.

In Conjecture 5.1.7 we formulate the crystalline lifts version of the weight part of Serre's conjecture for Galois representations that are semisimple locally at primes above p. We remark that when the extension $K/\mathbb{Q}_p$ is ramified, the definition of the weight set in terms of crystalline lifts involves a choice of lifting $K \hookrightarrow \overline{\mathbb{Q}}_p$ for each embedding $k \hookrightarrow \overline{\mathbb{F}}_p$ of the residue field k of K. (The former embeddings index Hodge–Tate weights, the latter are used to parameterise Serre weights.) This leads us to define two weight sets, $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ and $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$. The former is the set of weights obtained by taking the union over all such choices of liftings, and the latter is the set of weights obtained by taking the intersection. We conjecture that these two sets are in fact equal.

Section 6 contains a brief and informal discussion of some intuition for Serre weight conjectures that is suggested to us by the Galois moduli stacks of [EG15, EG].

We next explore the possibility of making the conjectures of Section 5 explicit. Our basic idea is that in the case when $\overline{\rho}$ is semisimple, we can explicitly construct many crystalline lifts of $\overline{\rho}$ by lifting each irreducible factor of $\overline{\rho}$ separately (and this comes down to constructing crystalline lifts of characters, since each irreducible mod p representation of G_K is induced from a character of an unramified extension of K). We call a crystalline lift obtained in this way an *obvious crystalline lift*, and correspondingly we obtain a set of weights $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$. (In fact this is not quite accurate: the set $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ also takes into account our expectation that the set of Serre weights of $\overline{\rho}$ should depend only on $\overline{\rho}|_{I_K}$; see Definition 7.1.3.) To illustrate,

the representation

$$\overline{\rho} \cong \begin{pmatrix} \overline{\varepsilon}^{k-1}\overline{\chi}_1 & 0 \\ 0 & \overline{\chi}_2 \end{pmatrix}$$

of $G_{\mathbb{Q}_p}$ with $\overline{\chi}_i$ unramified has a crystalline lift ρ of the form

$$\rho \cong \begin{pmatrix} \varepsilon^{k-1}\chi_1 & 0 \\ 0 & \chi_2 \end{pmatrix}$$

where each χ_i is unramified and lifts $\overline{\chi}_i$ ($i = 1, 2$). The representation ρ has Hodge–Tate weights $k - 1$ and 0; taking into account the shift by $(1, 0)$ as in (1.4.1), we predict that the Serre weight $\mathrm{Sym}^{k-2}\,\overline{\mathbb{F}}_p^2$, which is described by the highest weight $(k - 2, 0)$, is contained in $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$. Similarly, $\overline{\rho}$ has a crystalline lift with Hodge–Tate weights $p - 1$, $k - 1$ obtained by lifting $\overline{\chi}_2$ instead to $\varepsilon^{p-1}\chi_2$, leading to the inclusion $\det^{k-1} \otimes \mathrm{Sym}^{p-1-k}\,\overline{\mathbb{F}}_p^2 \in \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$, in accordance with Section 1.2.

As will be discussed in Remark 5.1.11, under the assumption of the generalised Breuil–Mézard conjecture, the shadow weight conjecture and the crystalline lifts conjecture are equivalent. Therefore, our explicit weight set for $\overline{\rho}$ needs to be closed under the consequences of the shadow weight conjecture. We denote the smallest set of Serre weights that satisfies this requirement and that contains $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ by $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$ (see Section 7.2). We call the weights that lie in the complement $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho})) \setminus \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ *shadow weights*. The simplest example occurs for GL_3 over $\mathbb{Q}_p$, as explained in Section 1.5.

For a period of time, we hoped that the set $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$ might explain the full set of weights of $\overline{\rho}$ arising from crystalline lifts; unfortunately, this cannot always be the case. Again this phenomenon first occurs for GL_3 over $\mathbb{Q}_p$. In some cases we can inductively construct further crystalline lifts of $\overline{\rho}$ coming from Levi subgroups. The idea is that we write $\overline{\rho} = \oplus_i \overline{\rho}_i$ and take the direct sum of certain crystalline lifts ρ_i of $\overline{\rho}_i$ whose existence would be implied by the generalised Breuil–Mézard conjecture in combination with the explicitly constructed weight set for $\overline{\rho}_i$. The weight set resulting from these (hypothetical) crystalline lifts is denoted by $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$. It contains in fact all shadow weights. We call the weights in the complement $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \setminus \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$ *obscure weights*. See Example 7.2.5 for examples of such weights in the case of GL_3 over $\mathbb{Q}_p$. In Section 7.4 we ask furthermore whether the weight set of $\overline{\rho}$ should be closed under certain "weight shifts", and we give a limited amount of evidence for a positive answer.

In general we do not know how close our explicit weight set $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ is to the actual set of weights of $\overline{\rho}$. In Section 8 we compare this predicted weight set to all existing conjectures and computational evidence that we are aware of. Then, in the final part of our paper, we give strong evidence that in case $K/\mathbb{Q}_p$ is unramified and $\overline{\rho}$ is sufficiently generic (a genericity condition on the tame inertia weights of $\overline{\rho}|_{I_{\mathbb{Q}_p}}$) we are not missing any weights. It turns out that it is most natural to work in the setting of unramified groups G over $\mathbb{Q}_p$, considering GL_n over K as the restriction of scalars $\mathrm{Res}_{K/\mathbb{Q}_p}\mathrm{GL}_n$ to $\mathbb{Q}_p$. (For our precise conditions on G, see Hypothesis 9.1.1.) We extend both our explicit weight set $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ as well as the weight set $W^?(\overline{\rho})$ of [Her09] to this general setting and then prove the following theorem.

Theorem 1.6.1 (Theorem 10.2.11). *If $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ is semisimple and sufficiently generic, then $W^?(\overline{\rho}) = \mathrm{W}_{\mathrm{expl}}(\overline{\rho}) = \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$.*

In particular, there are no obscure weights in this generic setting. The proof is not immediate but requires some subtle modular representation theory. We thus see this result as an encouraging sign that $W_{\mathrm{expl}}(\bar{\rho})$ is correct in the generic unramified case.

This paper has two appendices. Appendix A contains the proof of a theorem of J. C. Ye and J. P. Wang on alcove geometry that is needed in Section 10; as far as we know the only published proofs are in Chinese. In Appendix B we prove by combinatorial arguments that the explicit set of Serre weights defined in Section 7 is always non-empty.

The only part of the paper that we have not yet mentioned is Section 2, in which we describe a global framework for formulations of generalisations of Serre's conjecture, in terms of the cohomology of arithmetic quotients of adèle groups. Although it is of course necessary to have chosen such a framework before one can begin to speak about Serre's conjecture (e.g. in order to define what one means when one says that $\bar{r}$ to be modular of a given weight!), this discussion is in some sense secondary to the rest of the paper, which is entirely local except for parts of Section 4 on patching.

1.7. **Index to the weight sets defined in this paper.**

Associated to a global Galois representation $\bar{r} : G_F \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$:

- $W(\bar{r})$, the Serre weights of $\bar{r}$: Definition 2.1.3;
- $W_v(\bar{r})$, the local Serre weights of $\bar{r}$ at a place $v \mid p$: Conjecture 2.1.5.

Associated to a local Galois representation $\bar{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$:

- $W_{\mathrm{BM}}(\bar{\rho})$, the Breuil–Mézard predicted weights for $\bar{\rho}$: Definition 3.2.6;
- $W_{\mathcal{S}}(\bar{\rho})$, the $\mathcal{S}$-Breuil–Mézard predicted weights for $\bar{\rho}$: Definition 3.3.7;
- $W_{\mathrm{cris}}^{\exists}(\bar{\rho})$ and $W_{\mathrm{cris}}^{\forall}(\bar{\rho})$, the crystalline weights for $\bar{\rho}$: Definition 5.1.5.

Associated to $\bar{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ such that $\bar{\rho}|_{I_K}$ is semisimple:

- $W_{\mathrm{obv}}(\bar{\rho})$, the obvious weights for $\bar{\rho}$: Definition 7.1.3;
- $\mathcal{C}(W_{\mathrm{obv}}(\bar{\rho}))$, the obvious and shadow weights for $\bar{\rho}$: Definition 7.2.1;
- $W_{\mathrm{expl}}(\bar{\rho})$, the explicit predicted weights for $\bar{\rho}$: Definition 7.2.3;
- $W^{?}(\bar{\rho})$ ($n = 3$), the weights predicted by [Her09]: Proposition 8.2.8.

See Section 7.5 for a summary of our conjectures about these weight sets.

Associated to a tame inertial L-parameter $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}})$, for a group G as in Hypothesis 9.1.1 (generalising the corresponding definitions for $\bar{\rho}$ such that $\bar{\rho}|_{I_K}$ is semisimple):

- $W_{\mathrm{obv}}(\tau)$, the obvious weights for τ: Definition 9.3.6;
- $\mathcal{C}(W_{\mathrm{obv}}(\tau))$, the obvious and shadow weights for τ: Definition 9.3.9;
- $W_{\mathrm{expl}}(\tau)$, the explicit predicted weights for τ: Definition 9.3.10;
- $W^{?}(\tau)$, the weights for τ predicted in the manner of [Her09]: Definition 9.2.5.

The latter three of these sets coincide for sufficiently generic τ (Theorem 10.2.11).

1.8. **Acknowledgments.**

1.9. **Notation and conventions.** We fix a prime p. If K is any field, we let $\overline{K}$ be a separable closure of K, and let $G_K = \mathrm{Gal}(\overline{K}/K)$; nothing we do will depend on the choice of $\overline{K}$, and in particular we will sometimes consider G_L to be a subgroup of G_K when $K \subset L$. In Section 9 we will instead denote $\mathrm{Gal}(\overline{K}/K)$ by Γ_K to avoid a conflict of notation. All Galois representations are assumed to be continuous with respect to the profinite topology on the Galois group and the natural topology on the coefficients (which will usually be either the p-adic topology or the discrete topology).

If K is a finite extension of $\mathbb{Q}_p$, we write $\mathcal{O}_K$ and k respectively for the ring of integers and residue field of K, I_K for the inertia subgroup of G_K, and Frob_K for a geometric Frobenius element of G_K. If F is a number field and v is a finite place of F then we let Frob_v denote a geometric Frobenius element of G_{F_v} and we write k_v for the residue field of the ring of integers of F_v.

We will use E to denote our coefficient field, a finite extension of $\mathbb{Q}_p$ contained in $\overline{\mathbb{Q}}_p$. We write $\mathcal{O} = \mathcal{O}_E$ for the ring of integers of E and $\mathbb{F}$ for its residue field. When we are working with representations of the absolute Galois group of a finite extension $K/\mathbb{Q}_p$, we will often assume that E is *sufficiently large*, by which we mean that the images of all embeddings $K \hookrightarrow \overline{\mathbb{Q}}_p$ are contained in E. We also let $\overline{\mathbb{Z}}_p$ denote the ring of integers of $\overline{\mathbb{Q}}_p$ and $\overline{\mathbb{F}}_p$ its residue field (it is thus our fixed choice of algebraic closure of $\mathbb{F}_p$).

Let K be a finite extension of $\mathbb{Q}_p$, and let $\mathrm{Art}_K : K^\times \to W_K^{\mathrm{ab}}$ be the isomorphism provided by local class field theory, which we normalise so that uniformisers correspond to geometric Frobenius elements. Let rec denote the local Langlands correspondence from isomorphism classes of irreducible smooth representations of $\mathrm{GL}_n(K)$ over $\mathbb{C}$ to isomorphism classes of n-dimensional Frobenius semisimple Weil–Deligne representations of W_K as in the introduction to [HT01], so that when $n = 1$ we have $\mathrm{rec}(\pi) = \pi \circ \mathrm{Art}_K^{-1}$. We fix an isomorphism $\imath : \overline{\mathbb{Q}}_p \to \mathbb{C}$ and define the local Langlands correspondence rec_p over $\overline{\mathbb{Q}}_p$ by $\imath \circ \mathrm{rec}_p = \mathrm{rec} \circ \imath$. This depends only on $\imath^{-1}(\sqrt{p})$ (and the only ambiguity is a quadratic unramified twist, so that in particular $\mathrm{rec}_p|_{I_K}$ does not depend on any choices).

Assume for the rest of this section that E is sufficiently large. For the purposes of defining the notation below, we also allow $E = \overline{\mathbb{Q}}_p$, $\mathcal{O} = \overline{\mathbb{Z}}_p$, $\mathbb{F} = \overline{\mathbb{F}}_p$ in what follows in this section. Define $S_k = \{\sigma : k \hookrightarrow \mathbb{F}\}$ and $S_K = \{\kappa : K \hookrightarrow E\}$. If $\kappa \in S_K$, we let $\overline{\kappa}$ be the induced element of S_k. Let ε denote the p-adic cyclotomic character, and $\overline{\varepsilon}$ the mod p cyclotomic character. For each $\sigma \in S_k$ we define the fundamental character ω_σ corresponding to σ to be the composite

$$I_K \longrightarrow\!\!\!\!\rightarrow \mathcal{O}_K^\times \longrightarrow k^\times \stackrel{\sigma}{\longrightarrow} \mathbb{F}^\times,$$

where the first map is induced by Art_K^{-1}. In particular $\left(\prod_{\sigma \in S_k} \omega_\sigma\right)^{e(K/\mathbb{Q}_p)} = \overline{\varepsilon}|_{I_K}$. When $k = \mathbb{F}_p$ and $\sigma : k \to \mathbb{F}$ is the unique embedding, we will often write ω in place of ω_σ. If χ is a character of G_K or I_K, we denote its reduction mod p by $\overline{\chi}$.

If W is a de Rham representation of G_K over E, then for each $\kappa \in S_K$ we will write $\mathrm{HT}_\kappa(W)$ for the multiset of Hodge–Tate weights of W with respect to κ. By definition this set contains the integer $-i$ with multiplicity $\dim_E(W \otimes_{\kappa, K} \widehat{\overline{K}}(i))^{G_K}$. Thus for example $\mathrm{HT}_\kappa(\varepsilon) = \{1\}$. The set $\mathrm{HT}_\kappa(W)$ is invariant under extensions of the coefficient field, and so also makes sense for de Rham representations over $\overline{\mathbb{Q}}_p$ (and embeddings $\kappa : K \hookrightarrow \overline{\mathbb{Q}}_p$).

We say that W has *regular* Hodge–Tate weights if for each κ, the elements of $\mathrm{HT}_\kappa(W)$ are pairwise distinct. Let $\mathbb{Z}_+^n$ denote the set of tuples $(\lambda_1, \ldots, \lambda_n)$ of integers with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. A *Hodge type* is an element of $(\mathbb{Z}_+^n)^{S_K}$. Then if W has regular Hodge–Tate weights, there is a Hodge type λ such that for each $\kappa \in S_K$ we have

$$\mathrm{HT}_\kappa(\rho) = \{\lambda_{\kappa,1} + n - 1, \lambda_{\kappa,2} + n - 2, \ldots, \lambda_{\kappa,n}\},$$

and we say that W is regular of weight λ (or Hodge type λ).

An *inertial type* is a representation $\tau : I_K \to \mathrm{GL}_n(E)$ with open kernel and which extends to the Weil group W_K. Then we say that a de Rham representation $\rho : G_K \to \mathrm{GL}_n(E)$ has inertial type τ and Hodge type λ, or more briefly that ρ has type (λ, τ), if ρ is regular of weight λ, and the restriction to I_K of the Weil–Deligne representation $\mathrm{WD}(\rho)$ associated to ρ is equivalent to τ.

For any $\lambda \in \mathbb{Z}_+^n$, view λ as a dominant weight of the algebraic group $\mathrm{GL}_{n/\mathcal{O}_K}$ in the usual way, and let M_λ' be the algebraic $\mathcal{O}_K$-representation of GL_n given by

$$M_\lambda' := \mathrm{Ind}_{B_n}^{\mathrm{GL}_n}(w_0 \lambda)_{/\mathcal{O}_K}$$

where B_n is the Borel subgroup of upper-triangular matrices of GL_n, and w_0 is the longest element of the Weyl group. (This representation is denoted by $H^0_{\mathcal{O}_K}(\lambda)$ in [Jan03, §II.8]. Note that its generic fibre is irreducible with highest weight λ by [Jan03, II.5.6].) Write M_λ for the $\mathcal{O}_K$-representation of $\mathrm{GL}_n(\mathcal{O}_K)$ obtained by evaluating M_λ' on $\mathcal{O}_K$. For any $\lambda \in (\mathbb{Z}_+^n)^{S_K}$ we write $L_{\lambda, \mathcal{O}}$ for the $\mathcal{O}$-representation of $\mathrm{GL}_n(\mathcal{O}_K)$ defined by

$$\otimes_{\kappa \in S_K}(M_{\lambda_\kappa} \otimes_{\mathcal{O}_K, \kappa} \mathcal{O}),$$

although when $\mathcal{O}$ is clear from the context we will suppress it and write simply L_λ.

We remark that the sets S_K for varying (sufficiently large) coefficient fields E can be naturally identified, and we will freely do so; and similarly for the sets S_k, $(\mathbb{Z}_+^n)^{S_K}$, and $(\mathbb{Z}_+^n)^{S_k}$.

If A is a Noetherian local ring with maximal ideal $\mathfrak{m}$ of dimension d, and M is a finite A-module, then there is polynomial $P_M^A(X)$ of degree at most d (the Hilbert–Samuel polynomial of M), uniquely determined by the requirement that for $n \gg 0$, the value $P_M^A(n)$ is equal to the length of $M/\mathfrak{m}^{n+1}M$ as an A-module. Then the Hilbert–Samuel multiplicity $e(M, A)$ is defined to be $d!$ times the coefficient of X^d in $P_M^A(X)$, and we write $e(A)$ for $e(A, A)$.

2. A global setting

2.1. GL_n **over a number field.**
In this section we briefly explain a possible global setup in which we can formulate the weight part of Serre's conjecture for GL_n over a number field F. It is presumably possible to formulate such conjectures for a general connected reductive group over a number field (by a characteristic p analogue of the conjectures of [BG15]), but this would entail developing a great deal of material of no relevance to the bulk of this paper.

The point of this paper is to formulate and study only the weight part of Serre's conjecture, and this is expected to be a purely local question (see Conjecture 2.1.5 below). Indeed, essentially everything in this paper (other than various comparisons to results in the literature, giving evidence for our conjectures) after the present section is purely local. On the other hand, to make a general global formulation requires a careful discussion of various technical issues (such as: the association of mod p Satake parameters to characteristic polynomials of Frobenii; characteristic p analogues of the various considerations of [BG15], such as the C-group; variants using the Galois action on the étale cohomology of Shimura varieties, rather than the Hecke action on the Betti cohomology; and so on).

As all of these points are orthogonal to our goals in the remainder of the paper, we have restricted ourselves to a brief description of the case of GL_n, as this suffices for the bulk of the paper, and for much of the computational evidence to date. (The remaining computations concern forms of GL_2.)

Let $\mathbb{A}_F$ denote the adèles of F, and let $\mathbb{A}_F^\infty$ denote the finite adèles of F. Let $U = U^p U_p$ be a compact open subgroup of $\mathrm{GL}_n(\mathbb{A}_F^\infty)$, where U^p is a compact open subgroup of $\mathrm{GL}_n(\mathbb{A}_F^{\infty,p})$, assumed to be sufficiently small, and $U_p = \mathrm{GL}_n(\mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Z}_p)$. Let $A_\infty^\circ = \mathbb{R}_{>0}^\times$, embedded diagonally in $\prod_{v|\infty} \mathrm{GL}_n(F_v)$, and write $U_\infty^\circ = \prod_{v|\infty} U_v^\circ \subset \prod_{v|\infty} \mathrm{GL}_n(F_v)$, where $U_v^\circ = \mathrm{SO}_n(\mathbb{R})$ if v is real and $U_v^\circ = \mathrm{U}_n(\mathbb{R})$ if v is complex. Set

$$Y(U) := \mathrm{GL}_n(F) \backslash \mathrm{GL}_n(\mathbb{A}_F) / U A_\infty^\circ U_\infty^\circ.$$

Let W be an irreducible smooth $\overline{\mathbb{F}}_p$-representation of U_p; the action of U_p on W necessarily factors through $\prod_{v|p} \mathrm{GL}_n(k_v)$, and we write $W \cong \otimes_{v|p} W_v$, where W_v is an irreducible $\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_n(k_v)$. We can define a local system $\mathcal{W}$ of $\overline{\mathbb{F}}_p$-vector spaces on $Y(U)$ via

$$(2.1.1) \qquad \mathcal{W} := \big((\mathrm{GL}_n(F) \backslash \mathrm{GL}_n(\mathbb{A}_F) / U^p A_\infty^\circ U_\infty^\circ) \times W\big)/U_p.$$

By shrinking U^p we are free to assume that it is a product $U^p = \prod_{v \nmid p} U_v$. There is a finite set Σ_0 of finite places of F (dependent on U) which contains all places dividing p, and which has the property that if $v \notin \Sigma_0$ is a finite place of F, then $U_v = \mathrm{GL}_n(\mathcal{O}_{F_v})$. For each $v \notin \Sigma_0$, the spherical Hecke algebra $\mathcal{H}_v := \mathcal{H}(\mathrm{GL}_n(\mathcal{O}_{F_v}) \backslash \mathrm{GL}_n(F_v) / \mathrm{GL}_n(\mathcal{O}_{F_v}), \overline{\mathbb{Z}}_p)$ (cf. [Gro98b]) with coefficients in $\overline{\mathbb{Z}}_p$ acts naturally on each cohomology group $H^i(Y(U), \mathcal{W})$. Indeed, $\mathcal{H}_v$ is identified with the subalgebra $\mathcal{H}(U \backslash U \mathrm{GL}_n(F_v) U / U, \overline{\mathbb{Z}}_p)$ of the usual adelic Hecke algebra $\mathcal{H}(U \backslash \mathrm{GL}_n(\mathbb{A}_F^\infty) / U, \overline{\mathbb{Z}}_p)$, and the bigger subalgebra $\mathcal{H}(U \backslash U \mathrm{GL}_n(\mathbb{A}_F^{\infty,p}) U / U, \overline{\mathbb{Z}}_p)$ acts naturally on each cohomology group $H^i(Y(U), \mathcal{W})$ (the prime-to-p condition being relevant, as W may be non-trivial).

Let $\bar{r} : G_F \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ be an irreducible representation. For any U as above, and for any finite set Σ of places of F containing Σ_0, all the finite places at which $\bar{r}$ is ramified, and all the infinite places of F, we may define a maximal ideal $\mathfrak{m} = \mathfrak{m}(\bar{r}, U, \Sigma)$ of $\mathbb{T}_\Sigma := \otimes'_{v \notin \Sigma} \mathcal{H}_v$ with residue field $\overline{\mathbb{F}}_p$ by demanding that for all places $v \notin \Sigma$, the semisimple part of $\bar{r}(\mathrm{Frob}_v^{-1})$ is conjugate to the class defined by the $\mathcal{H}_v$-eigenvalues determined by $\mathfrak{m}$ under the (suitably twisted) Satake isomorphism (cf. [Gro99, §17]). (Of course, since we are working with GL_n, this just amounts to specifying the characteristic polynomial of $\bar{r}(\mathrm{Frob}_v)$, as in [CHT08, Prop. 3.4.4(2)], but the formulation we have used here generalises more easily to more general groups.)

Definition 2.1.2. We say that $\bar{r}$ is *automorphic* if there is some W, U, Σ as above such that $H^i(Y(U), \mathcal{W})_{\mathfrak{m}} \neq 0$ for some $i \geq 0$.

Definition 2.1.3. Suppose that $\bar{r}$ is automorphic. Let $\mathrm{W}(\bar{r})$ denote the set of isomorphism classes of irreducible representations W of $\prod_{v|p} \mathrm{GL}_n(k_v)$ for which $H^i(Y(U), \mathcal{W})_{\mathfrak{m}} \neq 0$ for some $i \geq 0$. We refer to $\mathrm{W}(\bar{r})$ as the set of *Serre weights* of $\bar{r}$.

Remark 2.1.4. Let $U_p(1)$ be the kernel of the homomorphism $U_p \to \prod_{v|p} \mathrm{GL}_n(k_v)$. A natural variant of the definition of the Serre weights of $\bar{r}$ would be to ask that $\mathrm{Hom}_{U_p}(W^\vee, H^i(Y(U_p(1)), \overline{\mathbb{F}}_p)_{\mathfrak{m}}) \neq 0$ for some $i \geq 0$. We do not know how to show unconditionally that the two definitions always give the same set of weights, but this would follow from conjectures in the literature, as we now explain.

Let r_1 be the number of real places of F, and let r_2 be the number of complex-conjugate pairs of complex places. Set $q_0 = r_1 \lfloor n^2/4 \rfloor + r_2 n(n-1)/2$; this is the minimal degree of cohomology to which tempered cohomological automorphic representations of $\mathrm{GL}_n(\mathbb{A}_F)$ will contribute. According to the conjectures of [CE12], as expanded upon in [Eme14, §3.1.1], it is expected that if some $H^i(Y(U), \overline{\mathbb{F}}_p)_{\mathfrak{m}}$ is non-zero (where now U_p can be arbitrarily small), then in fact $H^{q_0}(Y(U), \overline{\mathbb{F}}_p)_{\mathfrak{m}} \neq 0$, while $H^i(Y(U), \overline{\mathbb{F}}_p)_{\mathfrak{m}} = 0$ for $i < q_0$; there is a similar expectation for the $H^i(Y(U), \mathcal{W})_{\mathfrak{m}}$. If this conjecture holds, then it follows easily from the Hochschild–Serre spectral sequence that this variant definition gives the same set of Serre weights as Definition 2.1.3.

We now have the following general formulation of a weak version of the weight part of Serre's conjecture.

Conjecture 2.1.5. *Suppose that $\bar{r}$ is automorphic. Then we may write $\mathrm{W}(\bar{r}) = \bigotimes_{v|p} \mathrm{W}_v(\bar{r})$, where $\mathrm{W}_v(\bar{r})$ is a set of isomorphism classes of irreducible representations of $\mathrm{GL}_n(k_v)$, which depends only on $\bar{r}|_{G_{F_v}}$.*

In fact one expects something more, namely that the set $\mathrm{W}_v(\bar{r})$ should depend only on $\bar{r}|_{I_{F_v}}$, a point that will be important for making explicit Serre weight conjectures later in the paper (see especially Section 7.1).

Much of the rest of the paper will be occupied with the question of making Conjecture 2.1.5 more precise (and for giving evidence for the more precise conjectures) in the sense of giving conjectural descriptions of the sets $\mathrm{W}_v(\bar{r})$ in terms of $\bar{r}|_{G_{F_v}}$.

2.2. **Groups which are compact modulo centre at infinity.** While it is natural to work with the group GL_n/F, just as in the characteristic 0 Langlands program it is often advantageous to work with other choices of group, in particular those that admit discrete series. From the point of view of Serre's conjecture, it is particularly advantageous to work with groups which are compact mod centre at infinity; the associated arithmetic quotients only admit cohomology in degree 0, which facilitates an easy exchange between information in characteristic 0 and characteristic p. (In the more general context of groups that admit discrete series, there is an expectation that after localising at a maximal ideal $\mathfrak{m}$ as above which is "non-Eisenstein" in the sense that it corresponds to an irreducible Galois representation, cohomology should only occur in a single degree; however there are at present only fragmentary results in this direction, beyond the case of groups of semisimple rank 1.)

In particular, in the papers [Gro99], [Gro98a] and [Gro07] Gross considers questions relating to the weak form of Serre's conjecture for certain groups over $\mathbb{Q}$ which are compact mod centre at infinity. While he does not consider the weight part of Serre's conjecture in his setting (although the discussion of [Gro99, §4] could be viewed as a starting point in this direction), the conjectures we will make in this paper, especially those for more general reductive groups, could be used to make explicit Serre weight conjectures for "algebraic modular forms" (in Gross's terminology). It seems likely that computations with these automorphic forms would be a good way to investigate our general Serre weight conjectures.

A great deal of progress has been made on these and related questions for a particular class of such groups, namely unitary groups or quaternion algebras over totally real and CM fields, which are compact mod centre at infinity. (In the case of quaternion algebras over a totally real field, it is also possible to allow the quaternion algebra to split at a single infinite place: in that case the semisimple rank is 1, and it is easy to show that the cohomology of the associated Shimura curves vanishes outside of degree 1 after localising at a non-Eisenstein maximal ideal, for example via [BDJ10, Lem. 2.2]).

In particular, for these groups the association of Galois representations (valued in $\mathrm{GL}_n(\overline{\mathbb{Q}}_p)$) to automorphic representations is well-understood (see [Shi11] and the references therein), and the Taylor–Wiles machinery is also well-developed ([CHT08]) and has been successfully applied to the problem of the weight part of Serre's conjecture (see for example [GK14]). The relevance of these results (which address characteristic 0 Galois representations and how characteristic p Galois representations deform to characteristic 0) to the weight part of Serre's conjecture is the following simple principle, which underlies the proofs of most of what is known about the weight part of Serre's conjecture to date, and also motivates much of the material in the following sections.

Let us abusively adopt the notation of Section 2.1 above, although the groups we are considering are now (say) unitary groups which are compact mod centre at infinity. Let V be a finite free $\overline{\mathbb{Z}}_p$-module with a continuous action of U_p, and let $\overline{V} = V \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$. Then (recalling that U is sufficiently small) we can define a local system of $\overline{\mathbb{Z}}_p$-modules $\mathcal{V}$ on $Y(U)$ as in (2.1.1), and since $Y(U)$ only has cohomology in degree 0, we see that $H^0(Y(U), \mathcal{V})_{\mathfrak{m}} \neq 0$ if and only if $H^0(Y(U), \overline{\mathcal{V}})_{\mathfrak{m}} \neq 0$ if and only if $H^0(Y(U), \mathcal{W})_{\mathfrak{m}} \neq 0$ for some Jordan–Hölder factor W of $\overline{V}$.

Now, if $H^0(Y(U), \mathcal{V})_{\mathfrak{m}} \neq 0$ then we may consider the p-adic Galois representations attached to the automorphic representations contributing to $H^0(Y(U), \mathcal{V})_{\mathfrak{m}}$; these will lift our representation $\bar{r}$, and in particular for places $v|p$ the restrictions to G_{F_v} of these representations will lift $\bar{r}|_{G_{F_v}}$. The known p-adic Hodge-theoretic properties of the Galois representations associated to these automorphic representations then prescribe non-trivial relationships between the $\bar{r}|_{G_{F_v}}$ (for $v|p$) and V, and thus between the $\bar{r}|_{G_{F_v}}$ and $\overline{V}$. In particular, by considering the Jordan–Hölder factors W of $\overline{V}$, we obtain necessary conditions in terms of the $\bar{r}|_{G_{F_v}}$ for $\bar{r}$ to be automorphic of Serre weight W. The basic perspective of this paper (which was perhaps first considered in Section 4 of [Gee11], and was refined in [GK14]) is that these necessary conditions are often also sufficient.

Example 2.2.1. As a specific example of these considerations, consider the case of a definite quaternion algebra over $\mathbb{Q}$ that is split at p. Up to twist, an irreducible

$\overline{\mathbb{F}}_p$-representation of $\mathrm{GL}_2(\mathbb{F}_p)$ is of the form $W = \mathrm{Sym}^{k-2}\,\overline{\mathbb{F}}_p^2$ for some $2 \le k \le p+1$. Taking $V = \mathrm{Sym}^{k-2}\,\overline{\mathbb{Z}}_p^2$ in the above discussion, and using the Jacquet–Langlands correspondence, we find that if $\bar{r}$ is automorphic of Serre weight W, then $\bar{r}$ can be lifted to the Galois representation attached to a newform of weight k and level prime to p. By local-global compatibility, this means that $\bar{r}|_{G_{\mathbb{Q}_p}}$ has a lift to a crystalline representation with Hodge–Tate weights $\{k-1, 0\}$. If one assumes that conversely the only obstruction to $\bar{r}$ being automorphic of Serre weight W is this property of $\bar{r}|_{G_{\mathbb{Q}_p}}$ having a crystalline lift with Hodge–Tate weights $\{k-1, 0\}$, then an examination of the possible reductions modulo p of such crystalline representations recovers Serre's original conjecture [Ser87] (or rather, the specialisation of the conjecture of [BDJ10] to the case of modular forms over $\mathbb{Q}$, which implies Serre's original conjecture by, for example, an explicit comparison of Serre's original recipe for a minimal weight with the explicit list of Serre weights; see the proof of [BDJ10, Thm. 3.17]).

This example generalises in an obvious fashion to the case of forms of U(2) over totally real fields which are compact at infinity, and allows one to recover the Serre weight conjecture of [BDJ10]. (In the case of quaternion algebras over totally real fields there is a parity obstruction to finding lifts to characteristic zero, coming from the global units. However, in line with Remark 2.2.2 below, the weight part of Serre's conjecture is known for both quaternion algebras and compact forms of U(2) over totally real fields, and the sets of Serre weights are the same in both cases. We will elaborate on this point, and in particular say a few words about its proof, in Remark 4.2.5 below.)

More generally one can work over a totally real field with a form of U(n) which is compact at infinity, and employ similar considerations; the general theory of "change of weight" for Galois representations developed in [BLGGT14] (which generalises an argument of Khare–Wintenberger) shows that it is reasonable to expect that the only obstructions to producing automorphic lifts of particular weights will be the local ones prescribed by p-adic Hodge theory. However, for most choices of W it is no longer possible to find a representation V for which $\overline{V} \cong W$, and it is far from clear how to extract complete information in characteristic p from information in characteristic zero, and accordingly far from clear how to generalise the description of the weight part of Serre's conjecture for GL_2. However, we do still obtain information (for example, that being automorphic of some Serre weight implies the existence of a crystalline lift of some specific Hodge–Tate weights), and much of this paper is devoted to exploring the relationship between the weight part of Serre's conjecture and p-adic Hodge theory. In particular, a consequence of the philosophy of the paper [GK14] is that information about potentially semistable lifts is sufficient to determine the set of Serre weights in general; we explain this in Sections 3 and 4 below.

Remark 2.2.2. It is generally expected that there is a mod p Langlands correspondence satisfying local-global compatibility at places dividing p; this is known for $\mathrm{GL}_2/\mathbb{Q}$ by the results of [Eme10]. A consequence of such a compatibility would be that the sets $\mathrm{W}_v(\bar{r})$ would only depend on the reductive group over F_v. It therefore seems reasonable to use considerations from groups which are compact modulo centre at infinity to make conjectures about $\mathrm{W}_v(\bar{r})$ for more general groups; in particular, one can use considerations about unitary groups which split at places

above p (as in Example 2.2.1) to make predictions about the weight part of Serre's conjecture for GL_n.

Remark 2.2.3. One could consider the question of the relationship of the ramification of the Galois representation away from p to the tame level ("the level in Serre's conjecture"), and the question of sufficient conditions for a mod p Galois representation to correspond to a Hecke eigenclass in the first place ("the weak form of Serre's conjecture" which should correspond to an oddness condition at infinite places, see for example [Gro07] and [BV13, §6]). Again, these questions lie in a rather different direction to the concerns of this paper, and we will not address them here.

3. THE BREUIL–MÉZARD FORMALISM FOR GL_n AND SERRE WEIGHTS

In this section we will recall the formalism of the general Breuil–Mézard conjecture for GL_n, following [EG14], and then explain how the formalism leads to a Serre weight conjecture. As in [EG14], we will only formulate the potentially crystalline (as opposed to potentially semistable) version of the conjecture, as this is all that we will need. We expect an analogous conjecture to hold in the potentially semistable case, and we refer the reader to Section 1.1.4 of [Kis09a] for a discussion of the differences between the potentially crystalline and potentially semistable versions of the conjecture in the case of $\mathrm{GL}_2/\mathbb{Q}_p$. (See also Lemma 5.2 of [GG15], which shows for GL_2 that the potentially crystalline and potentially semistable conjectures predict the same multiplicities; we anticipate that the proof will extend to GL_n.)

3.1. **Serre weights.** Let $K/\mathbb{Q}_p$ be a finite extension, and assume throughout this section that the field $E/\mathbb{Q}_p$ (our field of coefficients) is sufficiently large. Recall that k and $\mathbb{F}$ denote the residue fields of K and E respectively. Fix a representation $\bar{\rho} : G_K \to \mathrm{GL}_n(\mathbb{F})$. (We will use $\bar{\rho}$ to denote a local Galois representation, typically of the group G_K, in contrast to $\bar{r}$ which we reserve for a global Galois representation, typically of the group G_F.)

Definition 3.1.1. A *Serre weight* is an isomorphism class of irreducible $\mathbb{F}$-representations of $\mathrm{GL}_n(k)$. (This definition will be extended to more general reductive groups in Definition 9.1.4.)

We will sometimes (slightly abusively) refer to an individual irreducible representation as a Serre weight.

Remark 3.1.2. From the results recalled below, it follows that all Serre weights can be defined over k (note that our running assumptions imply in particular that $\mathbb{F}$ contains the images of all embeddings $k \hookrightarrow \overline{\mathbb{F}}_p$). Hence the choice of coefficient field is irrelevant, will occasionally be elided below, and will be taken to be $\overline{\mathbb{F}}_p$ from Section 5 onwards.

Let $\mathrm{W}(k, n)$ denote the set of Serre weights for our fixed k and n. In the following paragraphs we give an explicit description of this set.

Write $X_1^{(n)}$ for the subset of $\mathbb{Z}_+^n$ consisting of tuples (a_i) such that $p-1 \geq a_i - a_{i+1}$ for all $1 \leq i \leq n-1$. If $a = (a_{\sigma,i}) \in (\mathbb{Z}_+^n)^{S_k}$, write a_σ for the component of a indexed by $\sigma \in S_k$. Set $f = [k : \mathbb{F}_p]$, and let $\sim$ denote the equivalence relation on $(\mathbb{Z}_+^n)^{S_k}$ in which $a \sim a'$ if and only if there exist integers x_σ such that $a_{\sigma,i} - a'_{\sigma,i} = x_\sigma$ for

all σ, i and for any labeling σ_j of the elements of S_k such that $\sigma_j^p = \sigma_{j+1}$ we have $\sum_{j=0}^{f-1} p^j x_{\sigma_j} \equiv 0 \pmod{p^f - 1}$. When $k = \mathbb{F}_p$ we can omit the subscript σ, and the above equivalence relation amounts to $a_i - a_i' = (p-1)y$ for some integer y, independent of i.

Given any $a \in X_1^{(n)}$, we define the k-representation P_a of $\mathrm{GL}_n(k)$ to be the representation obtained by evaluating $\mathrm{Ind}_{B_n}^{\mathrm{GL}_n}(w_0 a)_{/\mathcal{O}_K}$ on k (so we have a natural $\mathrm{GL}_n(\mathcal{O}_K)$-equivariant isomorphism $M_a \otimes_{\mathcal{O}_K} k \xrightarrow{\sim} P_a$), and let N_a be the irreducible sub-k-representation of P_a generated by the highest weight vector (that this is indeed irreducible follows from the analogous result for the algebraic group GL_n, cf. II.2.2–II.2.6 in [Jan03], and the appendix to [Her09]).

If $a \in (X_1^{(n)})^{S_k}$ then we define an irreducible $\mathbb{F}$-representation F_a of $\mathrm{GL}_n(k)$ by

$$F_a := \otimes_{\sigma \in S_k} (N_{a_\sigma} \otimes_{k, \sigma} \mathbb{F}).$$

The representations F_a are absolutely irreducible, and every irreducible $\mathbb{F}$-representation of $\mathrm{GL}_n(k)$ is of the form F_a for some a (see for example the appendix to [Her09]). Furthermore $F_a \cong F_{a'}$ if and only if $a \sim a'$, and so the map $a \mapsto F_a$ gives a bijection from $(X_1^{(n)})^{S_k}/\sim$ to the set of Serre weights. We identify the two sets $\mathrm{W}(k, n)$ and $(X_1^{(n)})^{S_k}/\sim$ under this bijection, and we refer to the elements of $(X_1^{(n)})^{S_k}/\sim$ as Serre weights. (We will abuse this terminology in two specific ways: if $a \in (X_1^{(n)})^{S_k}$ and W is a set of weights, we may write $a \in \mathrm{W}$ when literally we mean $F_a \in \mathrm{W}$, and we may write "the weight a" when literally we mean "the Serre weight represented by a.")

If $k = \mathbb{F}_p$ we will also write $F(a_{\sigma,1}, \ldots, a_{\sigma,n})$ for F_a, where $S_k = \{\sigma\}$.

3.2. **The Breuil–Mézard conjecture.** By the main results of [Kis08], for each Hodge type λ and inertial type τ there is a unique reduced and p-torsion free quotient $R_{\overline{\rho}, \mathcal{O}}^{\lambda, \tau}$ of the universal lifting $\mathcal{O}$-algebra $R_{\overline{\rho}, \mathcal{O}}$ which is characterised by the property that if E'/E is a finite extension of fields, then an $\mathcal{O}$-algebra homomorphism $R_{\overline{\rho}, \mathcal{O}} \to E'$ factors through $R_{\overline{\rho}, \mathcal{O}}^{\lambda, \tau}$ if and only if the corresponding representation $G_K \to \mathrm{GL}_n(E')$ is potentially crystalline of Hodge type λ and inertial type τ. The ring $R_{\overline{\rho}, \mathcal{O}}^{\lambda, \tau}[1/p]$ is regular by [Kis08, Thm. 3.3.8]. When $\mathcal{O}$ is clear from the context, we will suppress it and write simply $R_{\overline{\rho}}^{\lambda, \tau}$. If τ is trivial we will write $R_{\overline{\rho}}^{\lambda}$ for $R_{\overline{\rho}}^{\lambda, \tau}$.

Given an inertial type τ, there is a finite-dimensional smooth irreducible $\overline{\mathbb{Q}}_p$-representation $\sigma(\tau)$ of $\mathrm{GL}_n(\mathcal{O}_K)$ associated to τ by the "inertial local Langlands correspondence", as in the following consequence of the results of [SZ99], which is Theorem 3.7 of [CEG+16].

Theorem 3.2.1. *If τ is an inertial type, then there is a finite-dimensional smooth irreducible $\overline{\mathbb{Q}}_p$-representation $\sigma(\tau)$ of $\mathrm{GL}_n(\mathcal{O}_K)$ such that if π is any irreducible smooth $\overline{\mathbb{Q}}_p$-representation of G, then the restriction of π to $\mathrm{GL}_n(\mathcal{O}_K)$ contains (an isomorphic copy of) $\sigma(\tau)$ as a subrepresentation if and only if $\mathrm{rec}_p(\pi)|_{I_K} \sim \tau$ and $N = 0$ on $\mathrm{rec}_p(\pi)$. Furthermore, in this case the restriction of π to $\mathrm{GL}_n(\mathcal{O}_K)$ contains a unique copy of $\sigma(\tau)$.*

Remark 3.2.2. In particular, if τ is the trivial inertial type, then $\sigma(\tau) \cong \overline{\mathbb{Q}}_p$ is the trivial one-dimensional representation of $\mathrm{GL}_n(\mathcal{O}_K)$.

Remark 3.2.3. In general the type $\sigma(\tau)$ need not be unique, although it is a folklore conjecture (which is known for $n = 2$, see Henniart's appendix to [BM02]) that $\sigma(\tau)$ is unique if $p > n$. The Breuil–Mézard conjecture, as formulated below, should hold for any choice of $\sigma(\tau)$; indeed it seems plausible that the semisimplification of the reduction mod p of $\sigma(\tau)$ does not depend on any choices (this is the case when $n = 2$ by Proposition 4.2 of [BD14]).

Enlarging E if necessary, we may assume that $\sigma(\tau)$ is defined over E. Since it is a finite-dimensional representation of the compact group $\mathrm{GL}_n(\mathcal{O}_K)$, it contains a $\mathrm{GL}_n(\mathcal{O}_K)$-stable $\mathcal{O}$-lattice L_τ. Set $L_{\lambda,\tau} := L_\tau \otimes_{\mathcal{O}} L_\lambda$, a finite free $\mathcal{O}$-module with an action of $\mathrm{GL}_n(\mathcal{O}_K)$. Then we may write

$$(L_{\lambda,\tau} \otimes_{\mathcal{O}} \mathbb{F})^{\mathrm{ss}} \overset{\sim}{\longrightarrow} \oplus_a F_a^{n_{\lambda,\tau}(a)},$$

where the sum runs over Serre weights $a \in \mathrm{W}(k,n)$, and the $n_{\lambda,\tau}(a)$ are non-negative integers. Then we have the following conjecture.

Conjecture 3.2.4 (The generalised Breuil–Mézard conjecture)**.** *There exist non-negative integers $\mu_a(\overline{\rho})$ depending only on $\overline{\rho}$ and a such that for all Hodge types λ and inertial types τ we have $e(R_{\overline{\rho}}^{\lambda,\tau}/\varpi) = \sum_a n_{\lambda,\tau}(a)\mu_a(\overline{\rho})$.*

Here ϖ is a uniformiser of $\mathcal{O}$. The finitely many integers $\mu_a(\overline{\rho})$ are in fact hugely overdetermined by the infinitely many equations $e(R_{\overline{\rho}}^{\lambda,\tau}/\varpi) = \sum_a n_{\lambda,\tau}(a)\mu_a(\overline{\rho})$. We will return to this point in the next subsection.

Remark 3.2.5. The multiplicities $\mu_a(\overline{\rho})$ in Conjecture 3.2.4 will be independent of the coefficient field E, in the following sense. Let E' be a finite extension of E, with ring of integers $\mathcal{O}'$ and residue field $\mathbb{F}'$. Write $\overline{\rho}' = \overline{\rho} \otimes_{\mathbb{F}} \mathbb{F}'$ and $\tau' = \tau \otimes_E E'$. One knows ([BLGGT14, Lem. 1.2.1 and §1.4]) that there is an isomorphism $R_{\overline{\rho}',\mathcal{O}'}^{\lambda',\tau'} \cong R_{\overline{\rho},\mathcal{O}}^{\lambda,\tau} \otimes_{\mathcal{O}} \mathcal{O}'$. It follows that if Conjecture 3.2.4 holds then $\mu_a(\overline{\rho}') = \mu_a(\overline{\rho})$ for all $a \in \mathrm{W}(k,n)$.

The generalised Breuil–Mézard conjecture is almost completely understood when $n = 2$ and $K = \mathbb{Q}_p$ i.e. in the setting originally studied and conjectured by Breuil and Mézard [BM02][1]. In fact, it is completely understood in this setting when $p > 3$ [Kis09a, Paš15, HT15]; when $p = 2, 3$ it is known in all cases except when the representation $\overline{\rho}$ is reducible and the characters on the diagonal of $\overline{\rho}$ have ratio $\overline{\varepsilon}$ ($= \overline{\varepsilon}^{-1}$ when $p \leq 3$) [Paš16, Kis09a, San16]. The multiplicities $\mu_a(\overline{\rho})$ are described in most cases (those for which $\overline{\rho}$ has only scalar endomorphisms) in [BM02, §2.1.2], and in general in [Kis09a, §1.1] together with [San14].

Assuming Conjecture 3.2.4, we make the following definition and conjecture.

Definition 3.2.6. We define $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$, the *Breuil–Mézard predicted weights for $\overline{\rho}$*, to be the set of Serre weights a such that $\mu_a(\overline{\rho}) > 0$.

Conjecture 3.2.7. *In the weight part of Serre's conjecture (Conj. 2.1.5), we may take $\mathrm{W}_v(\overline{r}) = \mathrm{W}_{\mathrm{BM}}(\overline{r}|_{G_{F_v}})$.*

---

[1]Breuil and Mézard restricted their original conjecture to the case of Hodge types $\lambda = (r, 0)$ with $0 \leq r \leq p - 3$, due to the lack of a suitable integral p-adic Hodge theory at the time, and considered potentially semistable deformation rings. The conjecture was later extended to arbitrary Hodge types and adapted to the potentially crystalline setting by Kisin [Kis10].

In Section 4 below we will explain how the formalism of the Taylor–Wiles–Kisin patching method shows that this is a natural definition for the set of predicted Serre weights.

3.3. **Breuil–Mézard systems.** We now describe a family of variants of Conjectures 3.2.4 and 3.2.7. Let $\mathcal{S}$ be a set of pairs (λ, τ) such that λ is a Hodge type and τ is an inertial type (both for our fixed K and n). We say that $\mathcal{S}$ is a *Breuil–Mézard system* if the map $\mathbb{Z}^{\mathrm{W}(k,n)} \to \mathbb{Z}^{\mathcal{S}}$ given by the formula

$$(x_a)_{a \in \mathrm{W}(k,n)} \mapsto \left(\sum_a n_{\lambda,\tau}(a) x_a \right)_{(\lambda,\tau) \in \mathcal{S}}$$

is injective; in particular, if $\mathcal{S}$ is a Breuil–Mézard system then for each representation $\overline{\rho}$ the equations $e(R_{\overline{\rho}}^{\lambda,\tau}/\varpi) = \sum_a n_{\lambda,\tau}(a)\mu_a(\overline{\rho})$, regarded as a system of linear equations in the variables $\mu_a(\overline{\rho})$, can have at most one solution.

We remark that $\mathcal{S}$ is a Breuil–Mézard system if and only if the image of the map $\mathbb{Z}[\mathcal{S}] \to K_0(\mathrm{Rep}_{\mathbb{F}}(\mathrm{GL}_n(k)))$ sending $(\lambda, \tau) \mapsto [L_{\lambda,\tau} \otimes_{\mathcal{O}} \mathbb{F}]$ has finite index. (Here we write $K_0(\mathrm{Rep}_{\mathbb{F}}(\mathrm{GL}_n(k)))$ for the Grothendieck group of finite-dimensional $\mathbb{F}[\mathrm{GL}_n(k)]$-modules.) Indeed, if $\mathcal{S}$ is finite then this is precisely the dual of the definition in the previous paragraph; in general, for the 'only if' direction one simply notes that any Breuil–Mézard system contains a finite Breuil–Mézard system, and similarly for the 'if' direction.

Example 3.3.1. Take $n = 2$ and let BT be the set of pairs $(0, \tau)$, so that BT is the set of potentially Barsotti–Tate types. Then [GK14, Lem. 3.5.2] shows that BT is a Breuil–Mézard system, and indeed that this is true even if we restrict to types τ such that $\det \tau$ is tame.

To give another example, we make the following definition.

Definition 3.3.2. We say that an element $\lambda \in (\mathbb{Z}_+^n)^{S_K}$ is a *lift* of an element $a \in (\mathbb{Z}_+^n)^{S_k}$ if for each $\sigma \in S_k$ there exists $\kappa_\sigma \in S_K$ lifting σ such that $\lambda_{\kappa_\sigma} = a_\sigma$, and $\lambda_{\kappa'} = 0$ for all other $\kappa' \neq \kappa_\sigma$ in S_K lifting σ. In that case we may say that the lift λ is taken with respect to the choice of embeddings (κ_σ). When $a \in (X_1^{(n)})^{S_k}$ we will also say that λ is a *lift* (with respect to the choice of embeddings (κ_σ)) of the Serre weight represented by a.

Example 3.3.3. Fix a lift λ_b for each Serre weight b, and let $\widetilde{\mathrm{cr}}$ be the set of pairs $(\lambda_b, \mathrm{triv})$, where triv denotes the trivial type. Then $\widetilde{\mathrm{cr}}$ is a Breuil–Mézard system, because Lemma 3.3.5 below shows (inductively) that the natural map $\mathbb{Z}[\widetilde{\mathrm{cr}}] \to K_0(\mathrm{Rep}_{\mathbb{F}}(\mathrm{GL}_n(k)))$ is surjective.

Definition 3.3.4. For $a \in (\mathbb{Z}_+^n)^{S_k}$, let $\|a\| := \sum_{i,\sigma}(n + 1 - 2i)a_{\sigma,i} \in \mathbb{Z}_{\geq 0}$.

Lemma 3.3.5. *If λ is a lift of $a \in (X_1^{(n)})^{S_k}$, then $L_\lambda \otimes_{\mathcal{O}} \mathbb{F}$ has socle F_a, and every other Jordan–Hölder factor of $L_\lambda \otimes_{\mathcal{O}} \mathbb{F}$ is of the form F_b with $b \in (X_1^{(n)})^{S_k}$ and $\|b\| < \|a\|$.*

To prove Lemma 3.3.5 it is best to work not with the group GL_n over k but rather its restriction of scalars to $\mathbb{F}_p$. For this reason we defer the proof until Section 10.3. However, we stress that Lemma 3.3.5 will only be used in our discussion of the Breuil–Mézard system $\widetilde{\mathrm{cr}}$.

In the following conjectures and definition, we let $\mathcal{S}$ be a Breuil–Mézard system.

Conjecture 3.3.6 (The Breuil–Mézard conjecture for representations of type $\mathcal{S}$). *There exist non-negative integers $\mu_a(\overline{\rho})$ depending only on $\overline{\rho}$ and a such that for all $(\lambda, \tau) \in \mathcal{S}$ we have $e(R_{\overline{\rho}}^{\lambda, \tau}/\varpi) = \sum_a n_{\lambda, \tau}(a)\mu_a(\overline{\rho})$.*

Definition 3.3.7. Suppose that the Breuil–Mézard conjecture for representations of type $\mathcal{S}$ is true for $\overline{\rho}$. We define $W_{\mathcal{S}}(\overline{\rho})$ to be the set of Serre weights a such that $\mu_a(\overline{\rho}) > 0$.

Conjecture 3.3.8 (The $\mathcal{S}$-weight part of Serre's conjecture). *Suppose that the Breuil–Mézard conjecture for representations of type $\mathcal{S}$ (Conj. 3.3.6) is true. Then the weight part of Serre's conjecture (Conj. 2.1.5) holds with $W_v(\overline{r}) = W_{\mathcal{S}}(\overline{r}|_{G_{F_v}})$.*

Of course if the generalised Breuil–Mézard conjecture (Conj. 3.2.4) holds, then so does the Breuil–Mézard conjecture for representations of any type $\mathcal{S}$, and in that case we must always have $W_{BM}(\overline{\rho}) = W_{\mathcal{S}}(\overline{\rho})$. In particular, if we believe Conjecture 3.2.4 (and as we explain in Section 4 below, we certainly *should* believe Conjecture 3.2.4!), then the Breuil–Mézard predicted weights for $\overline{\rho}$ are completely determined by information about the crystalline lifts of $\overline{\rho}$ of bounded Hodge–Tate weights.

Example 3.3.9. Assume that $p > 2$. Gee and Kisin [GK14, Cor. 3.5.6] have established the Breuil–Mézard conjecture for potentially Barsotti–Tate representations; that is, they have shown that Conjecture 3.3.6 holds for the system BT of Example 3.3.1. In fact they also prove (subject to a Taylor–Wiles-type hypothesis) that the BT-weight part of Serre's conjecture holds in this setting, i.e. that the analogue of Conjecture 2.1.5 for quaternion algebras or forms of U(2) over totally real fields holds with $W_v(\overline{r}) = W_{BT}(\overline{r}|_{G_{F_v}})$ ([GK14, Cor. 4.5.4]); see also the discussion in Section 4.2 of this paper.

Example 3.3.10. Let $\widetilde{\mathrm{cr}}$ be one of the Breuil–Mézard systems of Example 3.3.3. Then a weak version of the Breuil–Mézard conjecture for representations of type $\widetilde{\mathrm{cr}}$ is trivially true; namely, there are uniquely determined integers $\mu_a(\overline{\rho})$ satisfying the required equations, but it is not clear that these integers are non-negative. (Since the system $\widetilde{\mathrm{cr}}$ is in bijection with the set of Serre weights, it is immediate that there are uniquely determined rational numbers $\mu_a(\overline{\rho})$ satisfying the required equations, and that they are in fact integers follows easily from Lemma 3.3.5.)

If $n = 2$, it follows trivially that the $\mu_a(\overline{\rho})$ are indeed non-negative integers (so that the Breuil–Mézard conjecture holds for representations of type $\widetilde{\mathrm{cr}}$), and that $a \in W_{\widetilde{\mathrm{cr}}}(\overline{\rho})$ if and only if $\overline{\rho}$ has a crystalline lift of Hodge type λ_a (for our chosen lift λ_a of a).

The set $W_{\widetilde{\mathrm{cr}}}(\overline{\rho})$ *a priori* could depend on the choice of lifts in the construction of $\widetilde{\mathrm{cr}}$. However, if $n = 2$ and $p > 2$, it is proved in [GLS15, Thm. 6.1.8] that the set $W_{\widetilde{\mathrm{cr}}}(\overline{\rho})$ is independent of these choices, and indeed is equal to $W_{BT}(\overline{\rho})$ and thus (as explained in Example 3.3.9) under a mild Taylor–Wiles hypothesis the analogue of Conjecture 2.1.5 for quaternion algebras or forms of U(2) over totally real fields holds with $W_v(\overline{r}) = W_{\widetilde{\mathrm{cr}}}(\overline{r}|_{G_{F_v}})$.

4. Patching functors and the Breuil–Mézard formalism

4.1. **Patching functors.** The most general results available to date on the weight part of Serre's conjecture have been based on the method of Taylor–Wiles patching

(see, for example, [BLGG13] and [GK14]). In this section, we give a general formalism for these arguments, and we explain how the resolution of the weight part of Serre's conjecture for Hilbert modular forms in [GK14, BLGG13, GLS15] fits into this framework.

The formalism we have in mind is a generalisation of the one employed in [GK14], which in turn is based on Kisin's work on the Breuil–Mézard conjecture [Kis09a]. Since our aim in this paper is not to prove new global theorems, but rather to explain what we believe should be true, we avoid making specific Taylor–Wiles patching arguments, and instead use the abstract language of patching functors, originally introduced for GL_2 in [EGS15]. Our patching functors will be for GL_n, and will satisfy slightly different axioms from those in [EGS15], but are motivated by the same idea, which is to abstract the objects produced by Taylor–Wiles patching. In practice one often wants to consider all places above p at once, but for simplicity of notation we will work at a single place in this section.

Continue to work in the context of Section 3, so that we have a fixed representation $\overline{\rho} : G_K \to \mathrm{GL}_n(\mathbb{F})$. Fix some $h \geq 0$, and write $R_\infty := R_{\overline{\rho}}[[x_1, \ldots, x_h]]$, $X_\infty :=$ Spf R_∞. (In applications, the x_i will be the auxiliary variables that arise in the Taylor–Wiles method; they will be unimportant in our discussion, and the reader unfamiliar with the details of the Taylor–Wiles method will lose nothing by assuming that $h = 0$.) We write $R_\infty^{\lambda,\tau} := R_{\overline{\rho}}^{\lambda,\tau}[[x_1, \ldots, x_h]]$ and $X_\infty(\lambda, \tau) := $ Spf $R_\infty^{\lambda,\tau}$. Write $\overline{X}_\infty$ and $\overline{X}_\infty(\lambda, \tau)$ for the special fibres of X_∞ and $X_\infty(\lambda, \tau)$ respectively. Write $d + 1$ for the dimension of the non-zero $X_\infty(\lambda, \tau)$ (which is independent of the choice of λ, τ).

Let $\mathcal{C}$ denote the category of finitely generated $\mathcal{O}$-modules with a continuous action of $\mathrm{GL}_n(\mathcal{O}_K)$; in particular, we have $L_{\lambda,\tau} \in \mathcal{C}$ for any λ, τ. Fix a Breuil–Mézard system $\mathcal{S}$ in the sense of Section 3.3.

Definition 4.1.1. A *patching functor* for $\mathcal{S}$ is a non-zero covariant exact functor M_∞ from $\mathcal{C}$ to the category of coherent sheaves on X_∞, with the properties that:
  ○ for all pairs $(\lambda, \tau) \in \mathcal{S}$, the sheaf $M_\infty(L_{\lambda,\tau})$ is p-torsion free and has scheme-theoretic support $X_\infty(\lambda, \tau)$, and in fact is maximal Cohen–Macaulay over $X_\infty(\lambda, \tau)$;
  ○ for all Serre weights F_a, the support $\overline{X}_\infty(F_a)$ of the sheaf $M_\infty(F_a)$ either has dimension d or is empty;
  ○ the (maximal Cohen–Macaulay over a regular scheme, so) locally free sheaf $M_\infty(L_{\lambda,\tau})[1/p]$ has rank one over the generic fibre of $X_\infty(\lambda, \tau)$.

Remark 4.1.2. In practice, examples of patching functors M_∞ come from the Taylor–Wiles–Kisin patching method applied to spaces of automorphic forms, localised at a maximal ideal of a Hecke algebra which corresponds to a global Galois representation $\overline{r}$ which locally at some place above p restricts to give $\overline{\rho}$. For example, the functor $\sigma^\circ \mapsto M_\infty(\sigma^\circ)$ defined in [CEG$^+$16, §4] is conjecturally a patching functor; the only difficulty in verifying this is that the usual Auslander–Buchsbaum argument only shows that $M_\infty(L_{\lambda,\tau})$ is maximal Cohen–Macaulay over its support, which is a union of irreducible components of the generic fibre of $X_\infty(\lambda, \tau)$.

Showing that this support is in fact the whole of $X_\infty(\lambda, \tau)$ is one of the major open problems in the field; it is closely related to the Fontaine–Mazur conjecture, and is therefore strongly believed to hold in general. By the main results of [BLGGT14], this is known whenever all potentially crystalline representations of

Hodge type λ and inertial type τ are potentially diagonalisable, but this condition seems to be hard to verify in practice.

Remark 4.1.3. The assumption that $M_\infty(L_{\lambda,\tau})[1/p]$ has rank one corresponds to the notion of a minimal patching functor in [EGS15]. The following arguments go through straightforwardly if one allows the rank to be higher, and in applications coming from Taylor–Wiles–Kisin patching, it is occasionally necessary to allow this (due to the need to ensure that the tame level is sufficiently small when the image of the global Galois representation is also small), but it makes no essential difference to the discussion below. However, these cases are rare, and in particular the patching constructions of [CEG+16] give examples where the rank is one.

4.2. **The relationship to the Breuil–Mézard conjecture.** The connection between patching functors and Serre weights is the following result, which is an abstraction of one of the main ideas of [GK14].

Proposition 4.2.1. *If a patching functor for $\mathcal{S}$ exists, then the Breuil–Mézard conjecture for representations of type $\mathcal{S}$ (Conj. 3.3.6) holds, and the set $\mathrm{W}_\mathcal{S}(\overline{\rho})$ is precisely the set of weights $\overline{\sigma}$ for which $M_\infty(\overline{\sigma}) \neq 0$.*

Proof. Let M_∞ be a patching functor for $\mathcal{S}$. The $\overline{X}_\infty(\lambda, \tau)$ are all equidimensional of dimension d by [BM14, Lem. 2.1]. (Strictly speaking the context of *loc. cit.* has $n = 2$, but its proof is completely general.) By [Mat89, Thm. 14.6] we have

$$(4.2.2) \qquad e(M_\infty(L_{\lambda,\tau} \otimes_\mathcal{O} \mathbb{F}), \overline{X}_\infty(\lambda, \tau)) = \sum_a n_{\lambda,\tau}(a) e(M_\infty(F_a), \overline{X}_\infty(\lambda, \tau))$$

(noting that $M_\infty(F_a)$ is supported on $\overline{X}_\infty(\lambda, \tau)$ whenever $n_{\lambda,\tau}(a) > 0$). Now from [Mat89, Thm. 14.7] it follows that if $A \twoheadrightarrow B$ is a surjection of Noetherian local rings of the same dimension and M is a finitely generated B-module, then $e(B, M) = e(A, M)$, where on the right-hand side M is regarded as an A-module via the given map. If $\overline{X}_\infty(F_a)$ is non-empty, then from the definition of a patching functor it has dimension d, and it follows that

$$(4.2.3) \qquad n_{\lambda,\tau}(a) e(M_\infty(F_a), \overline{X}_\infty(\lambda, \tau)) = n_{\lambda,\tau}(a) e(M_\infty(F_a), \overline{X}_\infty(F_a)).$$

If we make the convention that $e(0, \varnothing) = 0$, then (4.2.3) holds in general, since the left-hand side is 0 when $\overline{X}_\infty(F_a)$ is empty.

From the third bullet point in the definition of a patching functor, we know that $M_\infty(L_{\lambda,\tau})_\mathfrak{p}$ is free of rank 1 for any minimal prime $\mathfrak{p}$ of $R_\infty^{\lambda,\tau}$ (note that the latter ring has no p-torsion). Let $S := R_\infty^{\lambda,\tau} \setminus \bigcup_\mathfrak{p} \mathfrak{p}$, the union being taken over all minimal primes of $R_\infty^{\lambda,\tau}$. We have $S^{-1}M_\infty(L_{\lambda,\tau}) \cong \prod_\mathfrak{p} M_\infty(L_{\lambda,\tau})_\mathfrak{p}$ as an $S^{-1}R_\infty^{\lambda,\tau} \cong \prod_\mathfrak{p}(R_\infty^{\lambda,\tau})_\mathfrak{p}$-module, the products again being taken over all minimal primes of $R_\infty^{\lambda,\tau}$. Hence we can find find $m' \in M_\infty(L_{\lambda,\tau})$ such that for any such $\mathfrak{p}$, the image of m' in $M_\infty(L_{\lambda,\tau})_\mathfrak{p}$ is a basis as a $(R_\infty^{\lambda,\tau})_\mathfrak{p}$-module. It follows that [Kis09a, Proposition 1.3.4](2) applies with $A = M = R_\infty^{\lambda,\tau}$, $M' = M_\infty(L_{\lambda,\tau})$, $G = 1$, $x = \varpi$, and $f : M \to M'$ the map sending $1 \mapsto m'$, from which we find that

$$(4.2.4) \qquad e(M_\infty(L_{\lambda,\tau} \otimes_\mathcal{O} \mathbb{F}), \overline{X}_\infty(\lambda, \tau)) = e(R_{\overline{\rho}}^{\lambda,\tau}/\varpi).$$

(Note that $e(\overline{X}_\infty(\lambda, \tau)) = e(R_\infty^{\lambda,\tau}/\varpi) = e(R_{\overline{\rho}}^{\lambda,\tau}/\varpi)$.)

Putting together equations (4.2.2), (4.2.3), and (4.2.4), we find that Conjecture 3.3.6 holds with

$$\mu_a(\overline{\rho}) := e(M_\infty(F_a), \overline{X}_\infty(F_a)).$$

By the definition of a Breuil–Mézard system, the $\mu_a(\overline{\rho})$ are uniquely determined. Finally, it follows from [Mat89, Formula 14.2] that $e(M_\infty(F_a), \overline{X}_\infty(F_a)) > 0$ if and only if $M_\infty(F_a) \neq 0$, and the result follows. $\square$

Remark 4.2.5. In the cases that M_∞ arises from the Taylor–Wiles–Kisin patching construction, $M_\infty(F_a)$ corresponds to (patched) spaces of mod p automorphic forms of weight a, and it is immediate from the definition that $M_\infty(F_a) \neq 0$ if and only if $\overline{r}$ is automorphic of Serre weight a. Thus in cases where it can be shown that the Taylor–Wiles–Kisin method gives a patching functor for $\mathcal{S}$ (which, as explained in Remark 4.1.2, amounts to showing that the support of the $M_\infty(L_{\lambda,\tau})$ is as large as possible), the $\mathcal{S}$-weight part of Serre's conjecture (Conj. 3.3.8) follows from Proposition 4.2.1.

As explained in Remark 4.1.2, it is not in general known that potentially crystalline representations are potentially diagonalisable, which limits the supply of patching functors for general Breuil–Mézard systems.

The situation is better when $n = 2$, and indeed as a result of the papers [GK14], [BLGG13] and [GLS15], it is now known that if $p > 2$, then $\mathrm{W}_{\mathrm{BT}}(\overline{\rho}) = \mathrm{W}_{\widetilde{\mathrm{cr}}}(\overline{\rho})$, where BT is the Breuil–Mézard system of Example 3.3.1, and $\widetilde{\mathrm{cr}}$ is any of the Breuil–Mézard systems of Example 3.3.3; and it is known that the analogue of Conjecture 2.1.5 for quaternion algebras or forms of U(2) over totally real fields holds for this set of weights.

We briefly recall the argument. By the results of [Kis09c, Gee06] potential diagonalisability is known for the system BT, and Proposition 4.2.1 (applied to the Taylor–Wiles–Kisin patching method for automorphic forms on suitable quaternion algebras or forms of U(2)) then implies the result of [GK14] discussed in Example 3.3.9. Indeed, as we have already explained, Proposition 4.2.1 is an abstraction of the arguments of [GK14]. It remains to show that $\mathrm{W}_{\mathrm{BT}}(\overline{\rho}) = \mathrm{W}_{\widetilde{\mathrm{cr}}}(\overline{\rho})$.

Since this question is purely local, it suffices to work in the U(2) setting, where it is essentially immediate (by the considerations explained in Example 2.2.1) that $\mathrm{W}_{\mathrm{BT}}(\overline{\rho}) \subset \mathrm{W}_{\widetilde{\mathrm{cr}}}(\overline{\rho})$ (note that by the previous paragraph, the left hand side is known at this point in the argument to be the set of weights that occur globally). The purely local results of [GLS15], coming from a detailed study of the underlying integral p-adic Hodge theory, show that if $F_a \in \mathrm{W}_{\widetilde{\mathrm{cr}}}(\overline{\rho})$, then $\overline{\rho}$ necessarily has a potentially diagonalisable crystalline lift of Hodge type $\lambda_{\widetilde{a}}$. The above machinery then shows that $\mathrm{W}_{\widetilde{\mathrm{cr}}}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{BT}}(\overline{\rho})$ (again using that the right hand side is the set of weights that occur globally; this part of the argument is carried out in [BLGG13]), as required.

5. Crystalline lifts and Serre weights

The Breuil–Mézard version of the weight part of Serre's conjecture (Conjecture 3.2.7) has the obvious drawback that even the definition of the conjectural set of weights $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ is contingent on Conjecture 3.2.4. (Of course, in theory it is possible to determine the conjectural values of the $\mu_a(\overline{\rho})$'s without proving the generalised Breuil–Mézard conjecture first, by computing $e(R_{\overline{\rho}}^{\lambda,\tau}/\varpi)$ for enough choices of λ and τ, but in practise this seems to be very difficult.) In this section we will, under the assumption that $\overline{\rho}|_{I_K}$ is semisimple, define another conjectural set of Serre weights in terms of crystalline lifts. Although this set of weights may not be any more computable than $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$, its definition will not depend on any

unproven conjectures, and perhaps more importantly it will provide a bridge between the Breuil–Mézard description of the set of Serre weights and a much more explicit set of Serre weights to be defined in Section 7.1.

It is perhaps also worth recalling that, although we have emphasized the Breuil–Mézard perspective in this article, the crystalline lifts perspective historically came first. Indeed, the original explicit description of weights given in [Ser87] can in retrospect be understood as the most optimistic conjecture that one could make given the constraints provided by known results on the reduction mod p of the crystalline representations associated to modular forms, and similarly the conjecture of [BDJ10] arose from the consideration of crystalline lifts via Fontaine–Laffaille theory. Unfortunately, when $n > 2$ it seems that (contrary to the conjectures made in [Gee11]) the obvious extension of these conjectures to the general case that $\overline{\rho}|_{I_K}$ is not semisimple is false, and it now seems likely that a precise description of the sets of weights in general will be extremely complicated; see Sections 6.2 and 7.3 below.

5.1. **Crystalline lifts.** We fix a finite extension $K/\mathbb{Q}_p$ and a representation $\overline{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$.

Remark 5.1.1. Note that we have now switched (for the remainder of the paper) to working with $\overline{\rho}$ whose coefficients are algebraically closed. By Remark 3.2.5 it still makes sense to speak of $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$: choose any sufficiently large finite extension $\mathbb{F}/\mathbb{F}_p$ such that $\overline{\rho}$ has a model $\overline{\rho}_{\mathbb{F}}$ over $\mathbb{F}$, set $\mu_a(\overline{\rho}) = \mu_a(\overline{\rho}_{\mathbb{F}})$, and take $\mathrm{W}_{\mathrm{BM}}(\overline{\rho}) = \{a : \mu_a(\overline{\rho}) > 0\}$ as usual. (Recall that Serre weights can equally well be taken to be defined over $\overline{\mathbb{F}}_p$; cf. Remark 3.1.2.) Similarly for any Hodge type λ we can write $R_{\overline{\rho}}^{\lambda} = R_{\overline{\rho},\mathcal{O}}^{\lambda} \otimes_{\mathcal{O}} \overline{\mathbb{Z}}_p$ for any sufficiently large $\mathcal{O}$, and Remark 3.2.5 again shows that this is well-defined. Correspondingly, in this section L_λ will mean $L_{\lambda,\mathcal{O}} \otimes_{\mathcal{O}} \overline{\mathbb{Z}}_p$ for any sufficiently large $\mathcal{O}$.

Definition 5.1.2. Suppose that $\lambda \in (\mathbb{Z}_+^n)^{S_K}$. A *crystalline lift* of $\overline{\rho}$ of Hodge type λ is a representation $\rho : G_K \to \mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ such that

- $\rho \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p \cong \overline{\rho}$ and
- $\rho \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{Q}}_p$ is crystalline and regular of weight λ.

To motivate our reformulation of the weight part of Serre's conjecture in terms of crystalline lifts, we consider the following lemma.

Lemma 5.1.3. *Assume that the generalised Breuil–Mézard conjecture (Conj. 3.2.4) holds. Then $\overline{\rho}$ has a crystalline lift of Hodge type λ if and only if $\mathrm{W}_{\mathrm{BM}}(\overline{\rho}) \cap \mathrm{JH}_{\mathrm{GL}_n(k)}(L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) \neq \varnothing$.*

Proof. Since a representation $\overline{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ has image contained in $\mathrm{GL}_n(\mathcal{O}')$ for some finite $\mathcal{O}'/\mathbb{Z}_p$, it follows that $\overline{\rho}$ has a crystalline lift of Hodge type λ if and only if $R_{\overline{\rho}}^{\lambda} \neq 0$. Under the assumption of Conjecture 3.2.4, this is equivalent to there being a Jordan–Hölder factor F_a of $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ such that $\mu_a(\overline{\rho}) > 0$, which by definition is equivalent to $a \in \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$. $\square$

Corollary 5.1.4. *Assume the generalised Breuil–Mézard conjecture (Conj. 3.2.4) holds, and let λ be a lift of the Serre weight a. If $a \in \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$, then $\overline{\rho}$ has a crystalline lift of Hodge type λ.*

Proof. Suppose that λ is a lift of a with respect to the lift (κ_σ) of S_k. From Definition 3.3.2 we see that $L_\lambda = \otimes_{\sigma \in S_k} M_{a_\sigma} \otimes_{\mathcal{O}_K, \kappa_\sigma} \overline{\mathbb{Z}}_p$, and so $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p \cong \otimes_{\sigma \in S_k} P_{a_\sigma} \otimes_{k,\sigma} \overline{\mathbb{F}}_p$. In particular $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ has F_a as a Jordan–Hölder factor and Lemma 5.1.3 applies. $\square$

We are thus led to make the following definition.

Definition 5.1.5. We define $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$, the *crystalline weights for $\overline{\rho}$*, to be the set of Serre weights a such that the representation $\overline{\rho}$ has a crystalline lift of Hodge type λ for some lift λ of a. We further define $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$ to be the set of Serre weights a such that $\overline{\rho}$ has a crystalline lift of Hodge type λ for every lift λ of a.

It is not difficult to check that this definition is reasonable in the following sense: let $a \in (X_1^{(n)})^{S_k}/\sim$ be a Serre weight, and suppose that $\lambda = (a_{\sigma,i})$ and $\lambda' = (a'_{\sigma,i})$ are two lifts of a to $(\mathbb{Z}_+^n)^{S_K}$, each taken with respect to the same choice of embeddings (κ_σ); then $\overline{\rho}$ has a crystalline lift of Hodge type λ if and only if it has a crystalline lift of Hodge type λ'. To see this, we first recall the following basic fact about crystalline characters and their reductions modulo p.

Lemma 5.1.6. *Let $\Lambda = \{\lambda_\kappa\}_{\kappa \in S_K}$ be a collection of integers.*
  (i) *There is a crystalline character $\psi_\Lambda^K : G_K \to \overline{\mathbb{Z}}_p^\times$ such that for each $\kappa \in S_K$ we have $\mathrm{HT}_\kappa(\psi_\Lambda^K) = \lambda_\kappa$; this character is uniquely determined up to unramified twists.*
  (ii) *We have $\overline{\psi}_\Lambda^K|_{I_K} = \prod_{\sigma \in S_k} \omega_\sigma^{b_\sigma}$, where $b_\sigma = \sum_{\kappa \in S_K : \overline{\kappa} = \sigma} \lambda_\kappa$.*

Proof. Existence in (i) is well-known; see for instance [Ser79, §2.3, Cor. 2] or [Con11, Prop. B.4]. If ψ and ψ' are crystalline characters of G_K with the same labeled Hodge–Tate weights, then $\psi^{-1}\psi'$ is a crystalline representation all of whose Hodge–Tate weights are zero, and so is unramified. This proves (i), while (ii) is a consequence of [Con11, Prop. B.3] (see also the proof of [GLS14, Prop. 6.7]). $\square$

Now, to justify the claim preceding Lemma 5.1.6, write $a'_{\sigma,i} - a_{\sigma,i} = x_\sigma$. Then the lift of Hodge type λ' can be obtained by twisting the lift of type λ by a crystalline character with κ_σ-labeled Hodge–Tate weight x_σ for each $\sigma \in S_k$, κ'-labeled Hodge–Tate weights 0 for all other $\kappa' \in S_K$, and trivial reduction; such a character exists by Lemma 5.1.6.

In general, we obviously have $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$, and assuming the generalised Breuil–Mézard conjecture we even have $\mathrm{W}_{\mathrm{BM}}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$. (The equality follows from Lemma 5.1.3, noting that $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p \cong L_{\lambda'} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ for any two lifts λ, λ' of the same Serre weight.) If $\overline{\rho}|_{I_K}$ is semisimple, then as in [Gee11, Conj. 4.2.1], we make the following conjecture. As we have already remarked, we do not believe that [Gee11, Conj. 4.2.1] is true without the semisimplicity hypothesis that we impose here; even in the semisimple case, where there is (as we will see below) considerable evidence in favour of the conjecture, we do not have a fully satisfying reason to believe that it holds in complete generality, in the sense that for instance we do not know how to see that it would follow from other widely-believed conjectures.

Conjecture 5.1.7 (The weight part of Serre's conjecture in terms of crystalline lifts)**.**
  (i) *We have $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$.*

(ii) If $\bar\rho|_{I_K}$ is semisimple, then in the context of the generalised Breuil–Mézard conjecture (Conj. 3.2.4), one has $\mathrm{W}_{\mathrm{BM}}(\bar\rho) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\bar\rho) = \mathrm{W}^{\forall}_{\mathrm{cris}}(\bar\rho)$.

(iii) If $\bar r|_{I_{F_v}}$ is semisimple for all places $v|p$, then the weight part of Serre's conjecture (Conj. 2.1.5) holds with $\mathrm{W}_v(\bar r) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\bar r|_{G_{F_v}})$.

If one believes the $\mathcal{S}$-weight part of Serre's conjecture (Conj. 3.3.8) — and as explained in Section 4 above, the Taylor–Wiles–Kisin method strongly suggests that we *should* believe Conjecture 3.3.8 — then the mysterious part of Conjecture 5.1.7 is the assertion that $\mathrm{W}^{\exists}_{\mathrm{cris}}(\bar\rho)$ is no larger than $\mathrm{W}_{\mathcal{S}}(\bar\rho)$. The evidence for this conjecture is for the most part limited to the case $n \le 2$ (but see Remark 5.1.9 below) and the case of $\mathrm{GL}_3(\mathbb{Q}_p)$, and from a theoretical point of view the conjecture is rather mysterious; however, the evidence for the case of $\mathrm{GL}_3(\mathbb{Q}_p)$ is striking (see Section 8 for a detailed discussion of the theoretical and computational evidence in this case), and makes the conjecture seem plausible in general.

Remark 5.1.8. Considerable progress has been made on Conjecture 5.1.7 in the case where $\bar\rho$ is at most two-dimensional.

If $n = 1$ then Conjecture 5.1.7 is a consequence of class field theory together with an analysis of crystalline characters and their reductions modulo p. (For example, part (i) of the conjecture when $n = 1$ follows from Lemma 5.1.6.)

If $n = 2$ and $p > 2$ then, as explained in Remark 4.2.5, part (i) of Conjecture 5.1.7 is known, and the analogue of part (iii) for quaternion algebras and forms of $\mathrm{U}(2)$ over totally real fields is also known. If $n = 2$ and $K = \mathbb{Q}_p$ then part (ii) is known whenever the Breuil–Mézard conjecture is known; that is, it is known unless $p = 2, 3$, the representation $\bar\rho$ is reducible, and the characters on the diagonal of $\bar\rho$ have ratio $\bar\varepsilon$ ($= \bar\varepsilon^{-1}$ when $p \le 3$) [Paš15, HT15, Paš16, San16]. Indeed, all of these results hold without the assumption of semisimplicity.

Remark 5.1.9. Again assuming the generalised Breuil–Mézard conjecture, we note that the weights in $\mathrm{W}_{\mathrm{BM}}(\bar\rho)$ and $\mathrm{W}^{\exists}_{\mathrm{cris}}(\bar\rho)$ which are in the closure of the lowest alcove (i.e. the weights a such that $a_{\sigma,1} - a_{\sigma,n} + (n-1) \le p$ for each σ) must always coincide: this follows by considerations similar to those in the proofs of Lemma 5.1.3 and Corollary 5.1.4, noting that if λ is a lift of such a weight, the representation $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ is irreducible. In particular, when $n \le 2$ all Serre weights are in the closure of the lowest alcove, so that the progress towards Conjecture 5.1.7 in the case $n \le 2$ should be regarded as relatively weak evidence for the general case.

It is worth mentioning that while it is an open problem to prove that $\mathrm{W}^{\exists}_{\mathrm{cris}}(\bar\rho)$ is non-empty in general, we strongly believe that this is the case. Indeed, if $\bar\rho$ arises as the local mod p representation associated to an automorphic representation of some unitary group which is compact at infinity, then this is automatic from the considerations explained in Example 2.2.1 (in brief: the corresponding system of Hecke eigenvalues will show up in the cohomology associated to some Serre weight W, and lifting to characteristic 0 gives a global Galois representation which is crystalline of the appropriate Hodge–Tate weights).

While it might seem that this is a rather restrictive requirement on $\bar\rho$, it is expected that such an automorphic representation exists for every choice of $\bar\rho$ (of course, one has to allow unitary groups associated to arbitrary CM fields). Indeed, as explained in [EG14, App. A], the methods of [Cal12] allow one to globalise $\bar\rho$ to a representation which should (under the assumption of a weak version of Serre's conjecture for unitary groups) correspond to an automorphic representation on

some unitary group. Furthermore, even without knowing weak Serre, under the assumptions that $p \nmid 2n$ and that $\overline{\rho}$ admits a potentially diagonalisable lift with regular Hodge–Tate weights, the potential automorphy results of [BLGGT14] imply that $\overline{\rho}$ can indeed be globalised to an automorphic Galois representation [EG14, Cor. A.7], so that $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$ is provably non-empty for such representations. It is widely expected that every $\overline{\rho}$ admits such a potentially diagonalisable lift, and this is known if $\overline{\rho}$ is semisimple by [CEG^{+}16, Lem. 2.2]. (These considerations are expanded upon in [GHLS15, §3].)

We close this section with the observation that the generalised Breuil–Mézard conjecture and the crystalline lifts version of the weight part of Serre's conjecture (Conjectures 3.2.4 and 5.1.7) together with Lemma 5.1.3 entail the following conjecture.

Conjecture 5.1.10. *Suppose that $\overline{\rho}|_{I_K}$ is semisimple. If $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho}) \cap \mathrm{JH}_{\mathrm{GL}_n(k)}(L_\lambda \otimes_{\overline{\mathbb{Z}}_p}$ $\overline{\mathbb{F}}_p) \neq \varnothing$ for some lift λ of the Serre weight a, then $a \in \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$.*

It is possible to use (global) potential automorphy techniques to prove Conjecture 5.1.10 in certain special cases; see [GHLS15, §3] for details.

Remark 5.1.11. Assume that the generalised Breuil–Mézard conjecture holds so that, as we have already observed, $\mathrm{W}_{\mathrm{BM}}(\overline{\rho}) \subset \mathrm{W}^{\vee}_{\mathrm{cris}}(\overline{\rho}) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$. Then Conjecture 5.1.7(ii) is equivalent to the variant of Conjecture 5.1.10 where $\mathrm{W}^{\exists}_{\mathrm{cris}}$ is replaced with W_{BM} (both times). Indeed, this variant is equivalent to $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ by Lemma 5.1.3.

6. THE PICTURE

6.1. **A geometric perspective.** We now explain a geometric perspective ("the picture") on the weight part of Serre's conjecture. Full details will appear in the papers [EG15, EG, CEGS]. Continue to fix a finite extension $K/\mathbb{Q}_p$ and an integer $n \geq 1$. Assume that p is odd. Then the papers [EG15, EG, CEGS] construct a finite type equidimensional Artin stack $\overline{\mathcal{X}}$ over $\mathbb{F}_p$ (of dimension $[K : \mathbb{Q}_p]\binom{n}{2}$), whose $\overline{\mathbb{F}}_p$-points naturally correspond to the isomorphism classes of those representations $\overline{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ that admit a de Rham lift to $\mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ (of course, these are conjecturally all the $\overline{\rho}$, but as far as we are aware this is only known if $n \leq 3$; the case $n = 3$ is due to Muller [Mul13]).

Furthermore, for each pair (λ, τ) consisting of a Hodge type λ and an inertial type τ, there is a finite type formal Artin stack $\mathcal{X}_{\lambda,\tau}$ over $\mathrm{Spf}\,\mathbb{Z}_p$, whose $\overline{\mathbb{Z}}_p$-points are in natural bijection with the isomorphism classes of de Rham representations $\rho : G_K \to \mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ of type (λ, τ). There is a specialisation morphism $\pi : \mathcal{X}_{\lambda,\tau} \to \overline{\mathcal{X}}$, which on points just sends ρ to its reduction modulo p. The underlying reduced substack of $\pi(\mathcal{X}_{\lambda,\tau})$ is a union of irreducible components of $\overline{\mathcal{X}}$.

Each irreducible component of $\overline{\mathcal{X}}$ has a dense open subset of closed points that lie only on that component, and which correspond to certain maximally non-split upper-triangular representations with characters $\chi_1, \ldots, \chi_n$ on the diagonal such that the characters $\chi_i|_{I_K}$ are fixed. We refer to these points as the *generic $\overline{\mathbb{F}}_p$-points* of the component.

Suppose for example that $n = 2$, and fix characters $\psi_i : I_K \to \overline{\mathbb{F}}_p^{\times}$ for $i = 1, 2$ that extend to G_K. Then whenever $\psi_1 \psi_2^{-1} \neq \overline{\varepsilon}$, there is a unique component whose generic $\overline{\mathbb{F}}_p$-points correspond to extensions of χ_2 by χ_1 with $\chi_i|_{I_K} \cong \psi_i$, and these

representations have a unique Serre weight. We label the irreducible component by the corresponding Serre weight. Note that this Serre weight can be read off directly from an expression of the χ_i in terms of fundamental characters (that is, from the tame inertial weights).

To illustrate what happens when $\psi_1 \psi_2^{-1} = \bar{\varepsilon}$, suppose further that $K = \mathbb{Q}_p$. There is one component of $\overline{\mathcal{X}}$ whose generic $\overline{\mathbb{F}}_p$-points are très ramifiée extensions of χ by $\chi \bar{\varepsilon}$, where χ is any unramified character, and also another component whose generic $\overline{\mathbb{F}}_p$-points are extensions of χ_2 by $\chi_1 \bar{\varepsilon}$, where $\chi_1 \neq \chi_2$ are any unramified characters. The peu ramifiée extensions of χ by $\chi \bar{\varepsilon}$ lie on both components (and so are not generic $\overline{\mathbb{F}}_p$-points on either of them). We label the first component by the Serre weight $\mathrm{Sym}^{p-1} \overline{\mathbb{F}}_p^2$, while the second is labeled by both 1 and $\mathrm{Sym}^{p-1} \overline{\mathbb{F}}_p^2$, the two Serre weights of a generic $\overline{\mathbb{F}}_p$-point on that component. In particular every component of $\overline{\mathcal{X}}$ labeled by 1 is also labeled by $\mathrm{Sym}^{p-1} \overline{\mathbb{F}}_p^2$. All other components of $\overline{\mathcal{X}}$ are labeled by a single Serre weight, as in the previous paragraph, and in fact each other Serre weight is the label for a unique irreducible component of $\overline{\mathcal{X}}$.

More generally, we expect that when $n > 2$ there will be a set of weights associated to each component, and the Serre weights of any $\bar{\rho}$ will be precisely the union of the sets of weights associated to the components that it lies on. In particular the labels of a component must therefore be the Serre weights of its generic $\overline{\mathbb{F}}_p$-points. This structure, with the set of Serre weights for $\bar{\rho}$ being the set $\mathrm{W}_{\mathcal{S}}(\bar{\rho})$ for a Breuil–Mézard system $\mathcal{S}$, should be a consequence of the Breuil–Mézard conjecture for representations of type $\mathcal{S}$. Indeed for $n = 2$ (with K arbitrary) and $\mathcal{S} = \mathrm{BT}$ this can be proved, as a consequence of the results of [GK14] (see [CEGS]).

Accordingly, understanding the weight part of Serre's conjecture should reduce to understanding the components of $\overline{\mathcal{X}}$ on which a given representation lies, and understanding what the Serre weights are for maximally non-split upper-triangular representations (that are generic enough to lie on a single component).

While this structure is already (at least to us) very attractive, we expect that the picture is both simpler and more structured than what is entailed by the Breuil–Mézard conjecture. Specifically, we expect that most components are labeled by a single weight, and that in the cases where there are multiple weights labeling a component, they are frequently related in a simple way (see Section 7.4). For example, if $K = \mathbb{Q}_p$ and a component has $F(a_1, \ldots, a_n)$ as a label, then the generic representations on the component are of the form

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

where $\chi_i|_{I_{\mathbb{Q}_p}} = \omega^{a_i + n - i}$. Furthermore, if none of the $a_i - a_{i+1}$ are equal to 0 or $p - 1$, then we expect there to be a unique component labeled by this weight, and this component should be labeled only by $F(a_1, \ldots, a_n)$. We will discuss the case where some $a_i - a_{i+1}$ are equal to 0 or $p - 1$ in Section 7.4.

6.2. **Crystalline lifts.** We briefly explain what light the geometric perspective of Section 6.1 sheds on the crystalline lifts conjectures of Section 5, and on their expected failure to extend to the case of non-semisimple representations.

Let a be a Serre weight, and let λ be a lift of a. As explained in Section 6.1, there is a specialisation morphism $\pi : \mathcal{X}_{\lambda,\text{triv}} \to \overline{\mathcal{X}}$, which on points just sends a crystalline representation ρ of weight λ to its reduction modulo p. The underlying reduced substack of $\pi(\mathcal{X}_{\lambda,\text{triv}})$ is a union of irreducible components of $\overline{\mathcal{X}}$, and the geometrisation of the Breuil–Mézard conjecture of [BM14, EG14] strongly suggests that these irreducible components should be precisely the ones that have some Jordan–Hölder factor of $L_\lambda \otimes \overline{\mathbb{F}}_p$ among their labels.

If the conjectures of Section 5 held for arbitrary (not necessarily semisimple) $\overline{\rho}$, then we would be forced to conclude that the Serre weight F_a is a label of each of the above components. However, work of Lê Hùng Viêt Bao, Brandon Levin, Dan Le and Stefano Morra [LLHLM15] contradicts this conclusion; instead, their calculations indicate that already for $n = 3$ and $K = \mathbb{Q}_p$, if a is in the upper alcove and is suitably generic, then the two Jordan–Hölder factors F_a, F_b of $L_\lambda \otimes \overline{\mathbb{F}}_p$ correspond to two components of $\overline{\mathcal{X}}$, labeled by the single weight F_a (resp. F_b), which meet in a codimension one substack. Thus the generic $\overline{\rho}$ on the component labeled by the weight in the lower alcove do not satisfy the conjectures of Section 5; it is only those $\overline{\rho}$ which lie in a special position which do so. (The limited evidence available in the cases $n = 2, 3$ suggests that it is possible that two components labeled by suitably generic weights F, F' meet in codimension i, where i is minimal such that $\text{Ext}^i_{\text{GL}_n(\mathbb{F}_p)}(F, F') \neq 0$, but we do not know if it is reasonable to expect this to be true in general.)

Of course, the most special position is that occupied by semisimple $\overline{\rho}$, which agrees (in the case that $\overline{\rho}$ is a sum of characters) with the conjectures of Section 5. Note also that in general it seems reasonable to expect that any component containing $\overline{\rho}$ also contains $\overline{\rho}^{\text{ss}}$, which is consistent with the folklore belief that the set of Serre weights for $\overline{\rho}$ should be a subset of those for $\overline{\rho}^{\text{ss}}$.

7. Explicit weight conjectures in the semisimple case

Once again assume that $\overline{\rho}|_{I_K}$ is semisimple. The set $\text{W}^{\exists}_{\text{cris}}(\overline{\rho})$ is, in general, very badly understood: for instance at the time of writing we do not know how to prove, in general, that it is non-empty! (Though we do when $\overline{\rho}$ is semisimple; see Appendix B.) We would therefore like to have a version of the weight part of Serre's conjecture that is more explicit than conjectures we have already described, such as Conjecture 5.1.7 in terms of crystalline lifts.

In Sections 7.1 and 7.2 we construct various sets of weights that we have good reason (e.g. as a consequence of the generalised Breuil–Mézard conjecture) to believe are contained in $\text{W}^{\exists}_{\text{cris}}(\overline{\rho})$. On the other hand there is no reason to think that in general any of these sets are actually equal to $\text{W}^{\exists}_{\text{cris}}(\overline{\rho})$; to the contrary, we explain in Example 7.2.9 and Section 7.4 why we believe that this should *not* be the case. These examples illustrate the difficulty in making a general explicit conjecture.

However, we do expect that at least for generic $\overline{\rho}$ and unramified $K/\mathbb{Q}_p$, the set $\mathcal{C}(\text{W}_{\text{obv}}(\overline{\rho}))$ defined below in Section 7.2 *is* equal to $\text{W}^{\exists}_{\text{cris}}(\overline{\rho})$, motivated by a comparison with the conjectures of [Her09]; this will be explained in Section 10.

7.1. **Obvious lifts.** Recall from Section 2 that if $\overline{r} : G_F \to \text{GL}_n(\overline{\mathbb{F}}_p)$ is automorphic, one may hope that the set $\text{W}_v(\overline{r})$ depends only on $\overline{r}|_{I_{F_v}}$. We do not understand this as well as we would like; for instance, it appears to be somewhat more than can be deduced easily from the Breuil–Mézard formalism, because even

in the case of $\mathrm{GL}_2(\mathbb{Q}_p)$, the quantities $\mu_a(\overline{\rho})$ do not depend only on $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ (see [San14, Thm. 1]). However, since p-adic Hodge theoretic conditions are fundamentally conditions about ramification, it is not unreasonable to imagine that the set $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{r}|_{G_{F_v}})$ depends exclusively on $\overline{r}|_{I_{F_v}}$ and not on the image of Frob_v under $\overline{r}$. For instance this is known to be true when $n = 2$ and $p > 2$ by [GLS15, Prop. 6.3.1].

To make Conjecture 5.1.7 explicit, one can imagine trying to exhibit specific elements of $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$ by constructing crystalline lifts of $\overline{\rho}$ of various Hodge types, for instance by taking sums ρ' of inductions of crystalline characters. This is essentially what we will do; however, an immediate flaw with this plan is that for such ρ' one has limited control over the image of Frob_K under $\overline{\rho}'$, and in particular one may not be able to match the image of Frob_K under $\overline{\rho}$. (This will however be possible in generic situations.) Guided by the expectation that $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$ should depend only on $\overline{\rho}|_{I_K}$, we will be satisfied with constructing certain crystalline representations ρ' (that we call *obvious lifts* of $\overline{\rho}$) with the property that $\overline{\rho}'|_{I_K} \cong \overline{\rho}|_{I_K}$. In particular we caution that an obvious lift of $\overline{\rho}$ need not literally be a lift of $\overline{\rho}$. When an obvious lift ρ' of $\overline{\rho}$ has Hodge type λ_a, with λ_a a lift of a Serre weight a, we will call a an *obvious weight* of $\overline{\rho}$.

We now set up some basic results about crystalline characters. For each integer $n \geq 1$, let K_n be the unique extension of K inside $\overline{K}$ which is unramified of degree n. We denote the residue field of K_n by k_n. Given a character $\chi : G_{K_n} \to \overline{\mathbb{Z}}_p^\times$, we define a character $\chi^{(r)}$ by

$$\chi^{(r)}(g) = \chi(\mathrm{Frob}_K^r \cdot g \cdot \mathrm{Frob}_K^{-r}).$$

Note that this character does not depend on the choice of Frob_K. The following lemma is elementary (see also Lemma 9.3.2(iii) for a generalisation).

Lemma 7.1.1. *If $\chi : G_{K_n} \to \overline{\mathbb{Z}}_p^\times$ is crystalline, then so is $\chi^{(r)}$ and for any $\kappa' \in S_{K_n}$ we have $\mathrm{HT}_{\kappa'}(\chi^{(r)}) = \mathrm{HT}_{\kappa' \circ \mathrm{Frob}_K^{-r}}(\chi)$.*

Recall from Lemma 5.1.6 that for any collection of integers $\Lambda = \{\lambda_\kappa\}_{\kappa \in S_K}$ there exists a crystalline character $\psi_\Lambda^K : G_K \to \overline{\mathbb{Z}}_p^\times$ such that for each $\kappa \in S_K$ we have $\mathrm{HT}_\kappa(\psi_\Lambda^K) = \lambda_\kappa$, and that this character is uniquely determined up to unramified twists.

Corollary 7.1.2. *Let $\Lambda = \{\lambda_{\kappa'}\}_{\kappa' \in S_{K_n}}$ be a collection of integers. The representation $\rho_\Lambda^K := \mathrm{Ind}_{G_{K_n}}^{G_K} \overline{\mathbb{Z}}_p(\psi_\Lambda^{K_n})$ is crystalline, and for each $\kappa \in S_K$ we have*

$$\mathrm{HT}_\kappa(\rho_\Lambda^K) = \{\lambda_{\kappa'} \,:\, \kappa' \in S_{K_n} \text{ such that } \kappa'|_K = \kappa\}.$$

Moreover we have

$$\overline{\rho}_\Lambda^K|_{I_K} \cong \bigoplus_{i=0}^{n-1} \left(\prod_{\sigma \in S_{k_n}} \omega_\sigma^{b_\sigma} \right)^{q^i}$$

where $b_\sigma = \sum_{\kappa' \in S_{K_n} : \overline{\kappa}' = \sigma} \lambda_{\kappa'}$ and $q = \#k$.

Proof. If ρ is a Hodge–Tate representation of G_K and L is a finite extension of K, then $\mathrm{gr}^{-i}(D_{\mathrm{HT}}(\rho|_{G_L})) = \mathrm{gr}^{-i}(D_{\mathrm{HT}}(\rho)) \otimes_K L$. From this we deduce that if $\kappa' \in S_L$ then $\mathrm{HT}_{\kappa'}(\rho|_{G_L}) = \mathrm{HT}_{\kappa'|_K}(\rho)$. Applying this statement for $L = K_n$, the Corollary now follows from Lemma 7.1.1, the fact that $\mathrm{Ind}_{G_{K_n}}^{G_K}(\psi_\Lambda^{K_n})|_{G_{K_n}} \cong \oplus_{r=0}^{n-1}(\psi_\Lambda^{K_n})^{(r)}$, and (for the first part of the statement) the fact that the property of

being crystalline only depends on the restriction to inertia. The formula for $\overline{\rho}_\Lambda^K|_{I_K}$ follows from Lemma 5.1.6(ii). □

Definition 7.1.3. Suppose that $\overline{\rho}|_{I_K}$ is semisimple. We define an *obvious lift* of $\overline{\rho}$ to be a representation of the form $\rho' = \rho_{\Lambda_1}^K \oplus \cdots \oplus \rho_{\Lambda_d}^K$ (for some partition $n_1 + \cdots + n_d = n$ of n) such that $\overline{\rho}'|_{I_K} \cong \overline{\rho}|_{I_K}$.

We define $W_{\mathrm{obv}}(\overline{\rho})$ to be the set of Serre weights a such that $\overline{\rho}$ has an obvious lift ρ' of Hodge type λ for some lift λ of a. (In this case we say that the lift ρ' *witnesses* the obvious weight a.)

It is essential in this definition that we have required $\overline{\rho}'|_{I_K} \cong \overline{\rho}|_{I_K}$ rather than $\overline{\rho}' \cong \overline{\rho}$: as we will see in Example 7.1.9 making the latter definition would sometimes have produced a different (too small) set of weights. We note that if $\overline{\rho}$ has an obvious lift of some Hodge type lifting the Serre weight a, then it has a lift of any Hodge type lifting a: this follows from Corollary 7.1.2 (specifically, the fact that $\overline{\rho}_\Lambda^K|_{I_K}$ only depends on the multisets $\{\lambda_{\kappa'} : \kappa' \in S_{K_n} \text{ lifting } \sigma \in S_{k_n}\}$) and an argument as in the paragraph following Definition 5.1.5.

Remark 7.1.4. The set $W_{\mathrm{obv}}(\overline{\rho})$ is always non-empty. This is not at all immediate from the definitions, and unfortunately the only proof we have been able to find proceeds via a direct and somewhat painful combinatorial argument; for this reason we have deferred the proof to Appendix B.

Since we expect that the possible Hodge types of the crystalline lifts of $\overline{\rho}$ depend only on $\overline{\rho}|_{I_K}$, we make the following conjecture.

Conjecture 7.1.5. *We have* $W_{\mathrm{obv}}(\overline{\rho}) \subset W_{\mathrm{cris}}^\forall(\overline{\rho})$.

We consider several illustrative examples.

Example 7.1.6. When $n = 1$, any obvious lift of $\overline{\rho}$ is an unramified twist of a genuine crystalline lift of $\overline{\rho}$, from which it follows that $W_{\mathrm{obv}}(\overline{\rho}) = W_{\mathrm{cris}}^\forall(\overline{\rho}) = W_{\mathrm{cris}}^\exists(\overline{\rho})$.

Example 7.1.7 (Comparison with Schein's conjecture). We determine the obvious weights of a representation $\overline{\rho} : G_K \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ such that $\overline{\rho}|_{I_K}$ is semisimple. Let e be the absolute ramification index of K. The weight part of Serre's conjecture has been formulated in this context by Schein [Sch08].

Suppose first that $\overline{\rho}$ is irreducible. Consider a Serre weight a represented by $(x_\sigma, y_\sigma)_{\sigma \in S_k} \in (\mathbb{Z}_+^2)^{S_k}$, and let $\lambda \in (\mathbb{Z}_+^2)^{S_K}$ be a lift of (x_σ, y_σ). An obvious lift ρ' of $\overline{\rho}$ must be of the form $\mathrm{Ind}_{G_{K_2}}^{G_K} \psi_\Lambda^{K_2}$. Suppose that ρ' witnesses a. We may take ρ' to have Hodge type λ, so that the Hodge–Tate weights $\lambda_{\kappa'}$ of $\psi_\Lambda^{K_2}$ are as follows. For each $\sigma \in S_k$, there is a pair $(\kappa'_{\sigma,1}, \kappa'_{\sigma,2})$ of K_2/K-conjugate embeddings $K_2 \hookrightarrow \overline{\mathbb{Q}}_p$ such that $\overline{\kappa}'_{\sigma,1}, \overline{\kappa}'_{\sigma,2} : k_2 \hookrightarrow \overline{\mathbb{F}}_p$ extend σ, and $\lambda_{\kappa'_{\sigma,1}} = x_\sigma + 1$, $\lambda_{\kappa'_{\sigma,2}} = y_\sigma$. For the remaining $e - 1$ pairs (κ'_1, κ'_2) of K_2/K-conjugate embeddings $K_2 \hookrightarrow \overline{\mathbb{Q}}_p$ such that $\overline{\kappa}'_1, \overline{\kappa}'_2 : k_2 \hookrightarrow \overline{\mathbb{F}}_p$ extend σ, we have $\{\lambda_{\kappa'_1}, \lambda_{\kappa'_2}\} = \{1, 0\}$. Write σ_1 for $\overline{\kappa}'_{\sigma,1} \in S_{k_2}$ and σ_2 for its k_2/k-conjugate. Let $0 \leq m_\sigma \leq e - 1$ be the number of embeddings $\kappa'_1 \neq \kappa'_{\sigma,1}$ with $\overline{\kappa}'_1 = \sigma_1$ and $\lambda_{\kappa'_1} = 1$. Then we see from Corollary 7.1.2 that

$$\overline{\rho}|_{I_K} \cong \begin{pmatrix} \prod_{\sigma \in S_k} \omega_{\sigma_1}^{x_\sigma+1+m_\sigma} \omega_{\sigma_2}^{y_\sigma+e-1-m_\sigma} & 0 \\ 0 & \prod_{\sigma \in S_k} \omega_{\sigma_2}^{x_\sigma+1+m_\sigma} \omega_{\sigma_1}^{y_\sigma+e-1-m_\sigma} \end{pmatrix}.$$

In other words, we have $a \in W_{\mathrm{obv}}(\overline{\rho})$ if and only if for each $\sigma \in S_k$ we can write the elements of S_{k_2} extending σ as σ_1, σ_2 so that the above formula holds for some

choice of integers $0 \leq m_\sigma \leq e-1$. Observe that these are precisely the Serre weights predicted for $\overline{\rho}$ in [Sch08, Thm. 2.4].

Next suppose that $\overline{\rho}$ is reducible, and let a be a Serre weight as in the previous paragraph. If $\overline{\rho}|_{I_K}$ is non-scalar then every obvious lift of $\overline{\rho}$ must be a sum of two characters, but if $\overline{\rho}|_{I_K}$ is scalar then $\overline{\rho}$ may also have obvious lifts whose generic fibres are irreducible. Consider first the obvious lifts ρ' of $\overline{\rho}$ that have Hodge type λ (hence witness a) and that are sums of two characters. Say $\rho' = \psi_\Lambda^K \oplus \psi_{\Lambda'}^K$ with $\Lambda = \{\lambda_\kappa\}$ and $\Lambda' = \{\lambda'_\kappa\}$. For each $\sigma \in S_k$, there is an embedding $\kappa_\sigma \in S_K$ lifting σ such that $\{\lambda_{\kappa_\sigma}, \lambda'_{\kappa_\sigma}\} = \{x_\sigma + 1, y_\sigma\}$. For the remaining $e-1$ embeddings $\kappa \in S_K$ extending σ, we have $\{\lambda_\kappa, \lambda'_\kappa\} = \{1, 0\}$. Define $J = \{\sigma \in S_k : \lambda_{\kappa_\sigma} = x_\sigma + 1\}$. If $\sigma \in J$ we let $0 \leq m_\sigma \leq e-1$ be the number of embeddings $\kappa \neq \kappa_\sigma$ extending σ such that $\lambda_\kappa = 1$, while if $\sigma \notin J$ we let $0 \leq m_\sigma \leq e-1$ be the number of embeddings $\kappa \neq \kappa_\sigma$ extending σ such that $\lambda'_\kappa = 1$. Then we see from Corollary 7.1.2 that $\overline{\rho}|_{I_K}$ is isomorphic to

$$
\begin{pmatrix}
\prod_{\sigma \in J} \omega_\sigma^{x_\sigma + 1 + m_\sigma} \prod_{\sigma \notin J} \omega_\sigma^{y_\sigma + e - 1 - m_\sigma} & 0 \\
0 & \prod_{\sigma \notin J} \omega_\sigma^{x_\sigma + 1 + m_\sigma} \prod_{\sigma \in J} \omega_\sigma^{y_\sigma + e - 1 - m_\sigma}
\end{pmatrix}.
$$

In other words the weight $a \in \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ is witnessed by an obvious lift whose generic fibre is reducible if and only if the above formula holds for some subset $J \subset S_k$ and a choice of integers $0 \leq m_\sigma \leq e-1$.

In fact if $\overline{\rho}|_{I_K}$ is scalar, then it turns out that every weight $a \in \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ that is witnessed by an obvious lift whose generic fibre is irreducible is also witnessed by an obvious lift whose generic fibre is reducible, so that the previous paragraph still describes the whole set $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ in this case. This observation is an elementary but not necessarily straightforward exercise that we leave to the reader. (One first reduces to the case $e \leq p-1$ by noting that if $e \geq p$ then *every* weight a whose central character is compatible with $\det(\overline{\rho})|_{I_K}$ lies in $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ and is witnessed by an obvious lift whose generic fibre is reducible. Alternately, if $p \geq 3$ the observation can be deduced from the local results in [GLS15], specifically Theorem 4.1.6 of *loc. cit.*, while for $p = 2$ one reduces to the case $e = 1$ as above. But the case $e = 1$ is straightforward: after twisting one may suppose that $y_\sigma = 0$ for all σ; then $x_\sigma + 1 \in \{1, p-1, p\}$ for all σ (see the last paragraph of the proof of [GLS14, Thm. 10.1] for a more precise statement), and one checks that $\overline{\rho}|_{I_K}$ has the above shape with $J = \{\sigma : x_\sigma = 0\}$.) Observe that these are precisely the Serre weights predicted for $\overline{\rho}$ in [Sch08, Thm. 2.5].

Example 7.1.8. Consider $(a, b, c) \in \mathbb{Z}_+^3$ with $a - b, b - c > 1$ and $a - c < p - 2$. We determine the obvious weights of a representation $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ such that $\overline{\rho}|_{I_{\mathbb{Q}_p}} = \omega^a \oplus \omega^b \oplus \omega^c$. Any obvious lift ρ' must be a sum of characters. In particular ρ' has the form $\psi_{\{x\}}^{\mathbb{Q}_p} \oplus \psi_{\{y\}}^{\mathbb{Q}_p} \oplus \psi_{\{z\}}^{\mathbb{Q}_p}$ where $(x-2, y-1, z) \in X_1^{(3)}$, and $\{x, y, z\}$ and $\{a, b, c\}$ reduce to the same subset of $\mathbb{Z}/(p-1)$. It follows that the only possibilities for (x, y, z) (up to translation by $\mathbb{Z}(p-1, p-1, p-1)$) are

$$(a, b, c),\ (b, c, a - p + 1),\ (c + p - 1, a, b),$$
$$(c + p - 1, b, a - p + 1),\ (a, c, b - p + 1),\ (b + p - 1, a, c)$$

and therefore

$$\mathrm{W}_{\mathrm{obv}}(\overline{\rho}) = \{F(a - 2, b - 1, c),\ F(b - 2, c - 1, a - p + 1),\ F(c + p - 3, a - 1, b),$$
$$F(c + p - 3, b - 1, a - p + 1),\ F(a - 2, c - 1, b - p + 1),\ F(b + p - 3, a - 1, c)\}.$$

We see from this example that we *cannot* expect to have $\mathrm{W}_{\mathrm{obv}}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$: this is because (at least if $\overline{\rho}$ is semisimple) the weights $F(c + p - 2, b - 1, a - p)$, $F(a-1, c-1, b-p)$, and $F(b+p-2, a-1, c-1)$ also belong to $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$. We explain this for $F(c+p-2, b-1, a-p)$; the others are similar. We need to exhibit a lift ρ' of $\overline{\rho}$ with Hodge–Tate weights $\{c+p, b, a-p\}$. Since $p+1 < 2p-(a-c) < 2p$, for example by [Ber10, Thm. 3.2.1(3)], we can take ρ' to be the sum of an unramified twist of ε^b and an unramified twist of $\varepsilon^{a-p} \otimes V$ where V is a suitable crystalline representation with irreducible generic fibre and Hodge–Tate weights $\{2p - (a - c), 0\}$ (one of the representations $V_{2p+1-(a-c), a_p}$ considered in [Ber10]).

Example 7.1.9. Next, we determine the obvious weights of an unramified representation $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$. The reader can verify that the family of obvious lifts $\psi_{\{p-1\}}^{\mathbb{Q}_p} \oplus \psi_{\{0\}}^{\mathbb{Q}_p} \oplus \psi_{\{-p+1\}}^{\mathbb{Q}_p}$ witness the weight $F(p - 3, -1, -p + 1)$; the obvious lifts $\rho_{\{-1,p\}}^{\mathbb{Q}_p} \oplus \psi_{\{0\}}^{\mathbb{Q}_p}$ witness the weight $F(p - 2, -1, -1)$; the obvious lifts $\rho_{\{-1,p\}}^{\mathbb{Q}_p} \oplus \psi_{\{p-1\}}^{\mathbb{Q}_p}$ witness the weight $F(p - 2, p - 2, -1)$; and that these are the only weights in $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ when $p > 2$. When $p = 2$, it is easy to check that we have $\mathrm{W}_{\mathrm{obv}}(\overline{\rho}) = W(\mathbb{F}_2, 3)$ (so there are four weights in this case). This example illustrates two points. First, although $\overline{\rho}$ is a sum of characters, there are obvious weights of $\overline{\rho}$ that cannot be witnessed by sums of characters. Second, we remark that many unramified representations $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ do not have literal lifts of the form $\psi_{\{p-1\}}^{\mathbb{Q}_p} \oplus \psi_{\{0\}}^{\mathbb{Q}_p} \oplus \psi_{\{-p+1\}}^{\mathbb{Q}_p}$ (or of the other two shapes above). For instance if $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ has a lift of the form $\psi_{\{p-1\}}^{\mathbb{Q}_p} \oplus \psi_{\{0\}}^{\mathbb{Q}_p} \oplus \psi_{\{-p+1\}}^{\mathbb{Q}_p}$ then $\overline{\rho}(\mathrm{Frob}_{\mathbb{Q}_p})$ will be semisimple. Similarly, it may be the case that $\overline{\rho}^{\mathrm{ss}}$ may not have a literal lift of the form $\rho_{\{-1,p\}}^{\mathbb{Q}_p} \oplus \psi_{\{0\}}^{\mathbb{Q}_p}$ or $\rho_{\{-1,p\}}^{\mathbb{Q}_p} \oplus \psi_{\{p-1\}}^{\mathbb{Q}_p}$, since possessing such a lift imposes restrictions on the eigenvalues of $\overline{\rho}(\mathrm{Frob}_{\mathbb{Q}_p})$.

7.2. **Shadow and obscure weights.** Now we would like to address the observation (from Example 7.1.8) that in general we need not have $\mathrm{W}_{\mathrm{obv}}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$. To begin to account for this, Conjecture 5.1.10 motivates the following definition.

Definition 7.2.1. If W is a set of Serre weights, we define $\mathcal{C}(\mathrm{W})$ to be the smallest set of weights with the properties:

- $\mathrm{W} \subset \mathcal{C}(\mathrm{W})$, and
- if $\mathcal{C}(\mathrm{W}) \cap \mathrm{JH}_{\mathrm{GL}_n(k)}(L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) \neq \varnothing$ for some lift λ of the Serre weight a, then $a \in \mathcal{C}(\mathrm{W})$.

For instance, Conjecture 5.1.10 asserts that we should have $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}) = \mathcal{C}(\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}))$.

Example 7.2.2. Return to the case of GL_3 over $\mathbb{Q}_p$. If $F(x, y, z)$ is a Serre weight such that $\mathcal{C}(\{F(x, y, z)\}) \supsetneq \{F(x, y, z)\}$, then $x - z < p - 2$ and $\mathcal{C}(\{F(x, y, z)\}) = \{F(x, y, z), F(z + p - 2, y, x - p + 2)\}$. Indeed, if $x - z < p - 2$ then by Proposition 3.18 of [Her09] there is a short exact sequence

$$0 \to F(\lambda) \to L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p \to F(x, y, z) \to 0$$

where $\lambda = (z + p - 2, y, x - p + 2)$, so that $F(\lambda) \in \mathcal{C}(\{F(x, y, z)\})$, and these give all the instances of reducible $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ with $\lambda \in X_1^{(3)}$.

For instance, in the setting of Example 7.1.8 we see that $F(c+p-2, b-1, a-p) \in \mathcal{C}(\{F(a - 2, b - 1, c)\})$, that $F(a - 1, c - 1, b - p) \in \mathcal{C}(\{F(b - 2, c - 1, a - p + 1)\})$, and that $F(b + p - 2, a - 1, c - 1) \in \mathcal{C}(\{F(c + p - 3, a - 1, b)\})$. In fact one can

check in this setting that $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$ is precisely $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ together with these three extra weights. (We note that this same prediction can be found in the discussion immediately following [ADP02, Def. 3.5].)

If one believes Conjecture 5.1.10, then one might hope that also $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho})) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$, and indeed we will show in Section 10 that this is a reasonable expectation when $K/\mathbb{Q}_p$ is unramified and $\overline{\rho}$ is sufficiently generic in a precise sense. However, the following generalisation of the principle behind Conjecture 5.1.10 will show that this cannot be true in all cases.

Suppose that $\overline{\rho}|_{I_K} \cong (\oplus_{j=1}^r \overline{\rho}^{(j)})|_{I_K}$ with $\overline{\rho}^{(j)} : G_K \to \mathrm{GL}_{n_j}(\overline{\mathbb{F}}_p)$ not necessarily irreducible. Write $\eta_m = (m-1, \ldots, 1, 0)$ for any $m \geq 1$. Let a be a Serre weight, and suppose that λ is some lift of a. Suppose that $\lambda^{(j)}$ (for each $1 \leq j \leq r$) are Hodge types in $(\mathbb{Z}^{n_j}_+)^{S_K}$ such that the $\lambda^{(j)}_\kappa + \eta_{n_j}$ for each κ are obtained by partitioning $\lambda_\kappa + \eta_n$ into r decreasing subsequences of length n_j. (We will say that the $\lambda^{(j)}$ are an η-*partition* of a.) If $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho}^{(j)}) \cap \mathrm{JH}_{\mathrm{GL}_{n_j}(k)}(L_{\lambda^{(j)}} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) \neq \varnothing$, then Conjecture 5.1.10 entails that $\overline{\rho}^{(j)}$ has a crystalline lift of Hodge type $\lambda^{(j)}$. The direct sum of these lifts would be a crystalline lift of $\oplus_j \overline{\rho}^{(j)}$ of Hodge type λ, in which case $a \in \mathrm{W}^{\exists}_{\mathrm{cris}}(\oplus_j \overline{\rho}^{(j)})$. Since we expect that $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$ depends only on $\overline{\rho}|_{I_K}$, we then also expect to have $a \in \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$.

We are thus led to the following definition.

Definition 7.2.3. Suppose that $\overline{\rho}|_{I_K}$ is semisimple. We recursively define $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$, the *explicit predicted weights for* $\overline{\rho}$, to be the smallest set containing $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ and satisfying the expectation described in the previous paragraph: that is, $a \in \mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ for any Serre weight a such that there exists a decomposition $\overline{\rho}|_{I_K} \cong \oplus_{j=1}^r \overline{\rho}^{(j)}|_{I_K}$ and an η-partition $\lambda^{(j)}$ of a such that $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}^{(j)}) \cap \mathrm{JH}_{\mathrm{GL}_{n_j}(k)}(L_{\lambda^{(j)}} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) \neq \varnothing$ for each j.

Taking $r = 1$ in this definition we see that $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho})) \subset \mathrm{W}_{\mathrm{expl}}(\overline{\rho})$. We say that an element of $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho})) \setminus \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ is a *shadow weight*, while an element of $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \setminus \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$ is an *obscure weight*.

Example 7.2.4. If $n \leq 2$ it is easily checked that $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) = \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$. It is shown in [GLS15, Thm. 4.1.6] that when $n = 2$ and $p > 2$ we have $\mathrm{W}^{\forall}_{\mathrm{cris}}(\overline{\rho}) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$, and that these sets agree with the prediction of Schein [Sch08]. It follows that if $n = 1$, or $n = 2$ with $p > 2$, then $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) = \mathrm{W}_{\mathrm{obv}}(\overline{\rho}) = \mathrm{W}^{\forall}_{\mathrm{cris}}(\overline{\rho}) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$.

Explicitly, if $n = 1$, Lemma 5.1.6 implies that $a \in \mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ for a Serre weight a if and only if $\overline{\rho}|_{I_K} = \prod_{\sigma \in S_k} \omega^{a_\sigma}_\sigma$. If $n = 2$, Example 7.1.7 shows that $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ coincides with the set of weights predicted by Schein [Sch08]. Since $\mathrm{W}_{\mathrm{obv}}(\overline{\rho}) = \mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ in this setting, the claim follows from the above results of [GLS15].

Example 7.2.5. The existence of shadow weights in the case of GL_3 over $\mathbb{Q}_p$ was discussed in Example 7.2.2. We now classify the obscure weights in this case (showing, in particular, that they sometimes exist). We will repeatedly make use of our knowledge of $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ for $n \leq 2$, see Example 7.2.4. Suppose that $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ is a representation such that $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ is semisimple.

Since $\mathrm{Sym}^r \overline{\mathbb{F}}^2_p$ is irreducible as a $\mathrm{GL}_2(\mathbb{F}_p)$-representation for $r \leq p - 1$, it is straightforward to see that the weight $F(x, y, z)$ can be obscure for $\overline{\rho}$ only if we have:

    ∘ $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong (\overline{\rho}^{(1)} \oplus \overline{\rho}^{(2)})|_{I_{\mathbb{Q}_p}}$ with $\dim \overline{\rho}^{(i)} = i$,

    ◦ $F(y+1) \in \mathrm{W}_{\mathrm{expl}}(\overline{\rho}^{(1)})$, i.e. $\overline{\rho}^{(1)}|_{I_{\mathbb{Q}_p}} \cong \omega^{y+1}$, and

    ◦ $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}^{(2)}) \cap \mathrm{JH}_{\mathrm{GL}_2(\mathbb{F}_p)}(L_{(x+1,z)} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p) \neq \varnothing.$

Moreover, as $F(x, y, z)$ is not obvious, the restriction $\overline{\rho}^{(2)}|_{I_{\mathbb{Q}_p}}$ does not have the form $\omega^{x+2} \oplus \omega^z$ or $\omega_{\sigma_1}^{(x+2)+pz} \oplus \omega_{\sigma_2}^{(x+2)+pz}$, where $S_{\mathbb{F}_{p^2}} = \{\sigma_1, \sigma_2\}$. Hence $x - z \geq p - 1$. A calculation shows that the irreducible constituents of $L_{(x+1,z)} \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ are $F(x-p+2, z)$, $F(x-p+1, z+1)$, $F(z+p-1, x-p+2)$ if $p - 1 \leq x - z < 2p - 2$ (where the second weight is omitted if $x - z = p - 1$) and $F(z + p - 1, z + 1)$ (twice), $F(z + 1, z)$ if $x - z = 2p - 2$. Hence $\overline{\rho}^{(2)}|_{I_{\mathbb{Q}_p}}$ is either $\omega^{x+1} \oplus \omega^{z+1}$ or $\omega_{\sigma_1}^{(x+3)+p(z-1)} \oplus \omega_{\sigma_2}^{(x+3)+p(z-1)}$ (the latter only if $x - z \neq 2p - 2$). (This is of course compatible with [Ber10, Thm. 3.2.1] computing the reduction of crystalline representations with Hodge–Tate weights $\{x + 2, z\}$.)

    In the first case, one finds that $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \omega^{x+1} \oplus \omega^{y+1} \oplus \omega^{z+1}$. If $x - y, y - z < p - 1$, then $F(z + p - 2, y, x - p + 2)$ is an obvious weight for $\overline{\rho}$, and $F(x, y, z)$ is its shadow (so in particular is not obscure). Suppose on the other hand that $x - y = p - 1$ or $y - z = p - 1$. Once again $F(x, y, z)$ cannot be a shadow weight (as a shadow weight $F(x, y, z)$ always has $x - y, y - z < p - 1$) but sometimes it is an obvious weight.

    The weight $F(y + p - 1, y, z)$ is straightforwardly checked to be obvious precisely when $p = 2$, or else $p > 2$ and $y - z \in \{0, p - 2\}$. (When $y - z = p - 2$ the obvious lift is a sum of characters, while when $y = z$ the obvious lift has the shape $\rho_{\{y+p+1,z\}}^{\mathbb{Q}_p} \oplus \psi_{\{y+1\}}^{\mathbb{Q}_p}$.) Thus the weight $F(y + p - 1, y, z)$ is an obscure weight for $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \omega^{y+1} \oplus \omega^{y+1} \oplus \omega^{z+1}$ exactly when $p > 2$ and $y - z \notin \{0, p - 2\}$. By a similar analysis the weight $F(x, y, y - p + 1)$ is an obscure weight for $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \omega^{x+1} \oplus \omega^{y+1} \oplus \omega^{y+1}$ exactly when $p > 2$ and $x - y \notin \{0, p - 2\}$.

    Now suppose instead that $x - z \neq 2p - 2$ and $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \omega^{y+1} \oplus \omega_{\sigma_1}^{(x+3)+p(z-1)} \oplus \omega_{\sigma_2}^{(x+3)+p(z-1)}$. If $x - y, y - z < p - 1$ then again the weight $F(z + p - 2, y, x - p + 2)$ is obvious (the obvious lift has the shape $\psi_{\{y+1\}}^{\mathbb{Q}_p} \oplus \rho_{\{x-p+2,z+p\}}^{\mathbb{Q}_p}$) and $F(x, y, z)$ is its shadow. Suppose on the other hand that $x - y = p - 1$ or $y - z = p - 1$. Then once again $F(x, y, z)$ cannot be a shadow weight, while sometimes it is an obvious weight.

    The weight $F(y + p - 1, y, z)$ with $y - z \neq p - 1$ can be checked to be obvious precisely when $y - z = p - 2$; in this case the obvious lift has the shape $\psi_{\{z\}}^{\mathbb{Q}_p} \oplus \rho_{\{y+p+1,y+1\}}^{\mathbb{Q}_p}$. (Note that $y - z = p - 1$ is excluded because $x - z \neq 2p - 2$.) Thus the weight $F(y + p - 1, y, z)$ is an obscure weight for $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \omega^{y+1} \oplus \omega_{\sigma_1}^{(y+2)+pz} \oplus \omega_{\sigma_2}^{(y+2)+pz}$ precisely when $y - z \notin \{p - 2, p - 1\}$. By a similar argument the weight $F(x, y, y - p + 1)$ is an obscure weight for $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \omega^{y+1} \oplus \omega_{\sigma_1}^{(x+2)+py} \oplus \omega_{\sigma_2}^{(x+2)+py}$ exactly when $x - y \notin \{p - 2, p - 1\}$. This completes our analysis of obscure weights for $\mathrm{GL}_3(\mathbb{Q}_p)$.

    One might optimistically hope that there is an equality $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$; for example this is known to be the case when $n \leq 2$ (except for $n = 2$ and $p = 2$) thanks to [GLS15]. Unfortunately we do not expect this to be true in general; for example, in Example 7.4.5 we give some explicit examples in the case of $\mathrm{GL}_3(\mathbb{Q}_p)$ of weights which are not in $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ but which we suspect are in $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$. Furthermore

we remark that the sets $W_{\mathrm{cris}}^{\exists}(\overline{\rho})$ must also be compatible with other functorial operations, such as suitable tensor products and inductions, and it is far from clear whether or not the sets $W_{\mathrm{expl}}(\overline{\rho})$ satisfy these compatibilities.

On the other hand, in the unramified setting we are prepared to conjecture that these two weight sets are equal at least for sufficiently generic $\overline{\rho}$.

Conjecture 7.2.6. *Suppose that $K/\mathbb{Q}_p$ is unramified and that $\overline{\rho}|_{I_K}$ is semisimple and sufficiently generic. Then $W_{\mathrm{expl}}(\overline{\rho}) = W_{\mathrm{cris}}^{\exists}(\overline{\rho})$.*

Conjecture 7.2.7. *Suppose that for each $v|p$, the extension $F_v/\mathbb{Q}_p$ is unramified and $\overline{r}|_{I_{F_v}}$ is semisimple and sufficiently generic. Then the weight part of Serre's conjecture (Conj. 2.1.5) holds with $W_v(\overline{r}) = W_{\mathrm{expl}}(\overline{r}|_{G_{F_v}})$.*

The general definition of "sufficiently generic" will be given in Definition 10.1.12, but to give the reader a sense of the meaning of this term, we spell it out in the case where $\overline{\rho}$ is a direct sum of characters.

Example 7.2.8. Suppose that K is an unramified extension of $\mathbb{Q}_p$ and that $\overline{\rho}$ is a sum of characters, so that $\overline{\rho}|_{I_K} \cong \oplus_{i=1}^n \prod_{\sigma \in S_k} \omega_\sigma^{\mu_{\sigma,i}}$ for integers $\mu_{\sigma,i}$ (very much not uniquely defined). Fix $\delta > 0$. We say that $\overline{\rho}$ is δ-generic if it is possible to choose the integers $\mu_{\sigma,i}$ such that $\mu_{\sigma,i} - \mu_{\sigma,i+1} \geq \delta$ for all $1 \leq i < n$ and all σ, and furthermore $\mu_{\sigma,1} - \mu_{\sigma,n} \leq p - n - \delta$. We say that a statement is true for sufficiently generic $\overline{\rho}$ if there exists $\delta > 0$ (independent of p) such that the statement is true for all δ-generic $\overline{\rho}$.

We will prove in Theorem 10.2.11 that Conjecture 7.2.7 agrees with all other conjectures in the literature, in particular that of [Her09] (hence our willingness to make the conjecture, even though it is stronger than what is entailed by the generalised Breuil–Mézard formalism and by Conjecture 5.1.7). In fact we will show in Theorem 10.2.11 that for sufficiently generic $\overline{\rho}|_{I_K}$ and $K/\mathbb{Q}_p$ unramified we have $W_{\mathrm{expl}}(\overline{\rho}) = \mathcal{C}(W_{\mathrm{obv}}(\overline{\rho}))$ (that is, there are no obscure weights), so that in the context of Conjecture 7.2.7 the construction of the set $W_{\mathrm{expl}}(\overline{r}|_{G_{F_v}})$ is somewhat simplified.

We stress that for any fixed $\overline{\rho}$ the set of weights $W_{\mathrm{expl}}(\overline{\rho})$ is quite explicit in principle, at least for p large: the calculation of $W_{\mathrm{obv}}(\overline{\rho})$ is a combinatorial exercise (as in Examples 7.1.8 and 7.1.9), and then the shadow and obscure weights are determined by the Jordan–Hölder decompositions of the representations $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$. As for the computability of those decompositions, consider first the case $k = \mathbb{F}_p$. One needs to decompose GL_n-modules $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ with λ dominant and $\|\lambda\| < Np$ (for some N independent of p, and with $\|\cdot\|$ as in Definition 3.3.4) into simple $\mathrm{GL}_n(\mathbb{F}_p)$-modules. For $p \gg 0$, Lusztig's conjecture allows one to recursively decompose $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ into simple GL_n-modules when λ is p-regular ([Jan03, II.8.22], [Fie12]). For the remaining λ one uses [Jan03, II.7.17(b)]. For decomposing simple GL_n-modules as representations of $\mathrm{GL}_n(\mathbb{F}_p)$, see for example [Jan87, §1.5]. For general k one follows the same strategy, replacing GL_n with the algebraic group $G = \mathrm{Res}_{W(k)/\mathbb{Z}_p} \mathrm{GL}_n$ and $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p$ with the dual Weyl module $W(\lambda)$ as defined in Sections 9–10.

Example 7.2.9. As remarked above, we will show in Theorem 10.2.11 that for sufficiently generic $\overline{\rho}|_{I_K}$ and $K/\mathbb{Q}_p$ unramified we have $W_{\mathrm{expl}}(\overline{\rho}) = \mathcal{C}(W_{\mathrm{obv}}(\overline{\rho}))$. In this example, we show that this statement does not extend to the case where $K/\mathbb{Q}_p$ is ramified.

Suppose that $K/\mathbb{Q}_p$ is ramified quadratic and $\overline{\rho} : G_K \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ is such that $\overline{\rho}|_{I_K} \cong \omega^{a+3} \oplus \omega^{b+2} \oplus \omega^{c+1}$, where $a > b > c$ and $a - c < p - 4$. We claim that $F(a, b, c)$ is an obscure weight of $\overline{\rho}$.

If we had $F(a, b, c) \in \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$, then $F(a, b, c) \in \mathrm{W}_{\mathrm{obv}}(\overline{\rho})$, as (a, b, c) lies in the lowest alcove. As $\overline{\rho}|_{I_K}$ is a sum of distinct characters, any obvious crystalline lift of $\overline{\rho}|_{I_K}$ is a sum of characters. From Lemma 5.1.6 we would get that $\overline{\rho}|_{I_K} \cong \omega^r \oplus \omega^s \oplus \omega^t$ with $(r, s, t) = (a + 2, b + 1, c) + w(2, 1, 0)$ for some permutation $w \in S_3$. By the bounds on (a, b, c) we get a contradiction.

To show that in fact $F(a, b, c) \in \mathrm{W}_{\mathrm{expl}}(\overline{\rho})$, note that we can find $\overline{\rho}^{(i)} : G_K \to \mathrm{GL}_i(\overline{\mathbb{F}}_p)$ $(i = 1, 2)$ with $\overline{\rho}^{(1)}|_{I_K} \cong \omega^{b+2}$ and $\overline{\rho}^{(2)}|_{I_K} \cong \omega^{a+3} \oplus \omega^{c+1}$. By Lemma 5.1.6 we have $F(b + 2) \in \mathrm{W}_{\mathrm{obv}}(\overline{\rho}^{(1)})$ and $F(a + 2, c) \in \mathrm{W}_{\mathrm{obv}}(\overline{\rho}^{(2)})$. Let $S_K = \{\sigma_1, \sigma_2\}$. We define an η-partition of $F(a, b, c)$ as follows:

$$\lambda_{\sigma_1} = (a, b, c), \qquad\qquad\qquad \lambda_{\sigma_2} = 0,$$
$$\lambda_{\sigma_1}^{(1)} = (b + 1), \qquad\qquad\qquad \lambda_{\sigma_2}^{(1)} = (1),$$
$$\lambda_{\sigma_1}^{(2)} = (a + 1, c), \qquad\qquad\qquad \lambda_{\sigma_2}^{(2)} = (1, 0).$$

Then $L_{\lambda^{(1)}} \otimes \overline{\mathbb{F}}_p \cong F(b + 2)$ and $L_{\lambda^{(2)}} \otimes \overline{\mathbb{F}}_p \cong \mathrm{Sym}^{a-c+1} \overline{\mathbb{F}}_p^2 \otimes \mathrm{Sym}^1 \overline{\mathbb{F}}_p^2 \otimes \det^c$. We see that $F(a + 2, c)$ is a Jordan–Hölder factor of $L_{\lambda^{(2)}} \otimes \overline{\mathbb{F}}_p$, for example by Brauer's formula [Jan03, II.5.8(b)]. (The only other factor is $F(a+1, c+1)$.) From Definition 7.2.3 we see that indeed $F(a, b, c) \in \mathrm{W}_{\mathrm{expl}}(\overline{\rho})$.

7.3. Remarks on the general (non-semisimple) case.

Now let us drop our assumption that $\overline{\rho}|_{I_K}$ is semisimple, and consider what we might say about explicit weights for $\overline{\rho}$. As mentioned in Section 6.2, one expects that the Serre weights of $\overline{\rho}$ should be a subset of the Serre weights of $\overline{\rho}^{\mathrm{ss}}$. However, we hesitate to make any sort of precise conjecture: evidence is scant beyond the two-dimensional case, and the limited information that we do possess suggests that there are serious complications that arise already in the three-dimensional case.

We begin with a brief review of the two-dimensional case (for $p > 2$ and general $K/\mathbb{Q}_p$) as studied in [GLS15]. It is shown that $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ depends only on $\overline{\rho}|_{I_K}$ ([GLS15, Prop. 6.3.1]), and that $\mathrm{W}_{\mathrm{cris}}^{\vee}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}^{\mathrm{ss}})$. Suppose now that $\overline{\rho} : G_K \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ is an extension of characters χ_1 by χ_2. By Example 7.1.7 any weight $a \in \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}^{\mathrm{ss}})$ is witnessed by a sum of characters. Let $L(\chi_1, \chi_2, a)$ be the subset of $H^1(G_K, \chi_2\chi_1^{-1})$ obtained by taking the union, over all literal lifts ψ_1, ψ_2 of χ_1, χ_2 such that $\psi_1 \oplus \psi_2$ witnesses $a \in \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}^{\mathrm{ss}})$, of the image in $H^1(G_K, \chi_2\chi_1^{-1})$ of $H^1_f(G_K, \overline{\mathbb{Z}}_p(\psi_2\psi_1^{-1}))$. Then $a \in \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ if and only if the extension class corresponding to $\overline{\rho}$ lies in $L(\chi_1, \chi_2, a)$.

In fact it is almost always true that if ψ_1, ψ_2 as above are chosen so that the dimension of $H^1_f(G_K, \overline{\mathbb{Z}}_p(\psi_2\psi_1^{-1}))$ is as large as possible, then the image of that space in $H^1(G_K, \chi_2\chi_1^{-1})$ is actually equal to $L(\chi_1, \chi_2, a)$. The lone exception occurs when $\chi_2\chi_1^{-1}$ is the cyclotomic character and a is represented by $(x_\sigma, y_\sigma)_{\sigma \in S_k}$ with $x_\sigma - y_\sigma = p - 1$ for all $\sigma \in S_k$. In that case, if the ψ_i as above are chosen so that the dimension of $H^1_f(G_K, \overline{\mathbb{Z}}_p(\psi_2\psi_1^{-1}))$ is as large as possible, then the images of the spaces $H^1_f(G_K, \overline{\mathbb{Z}}_p(\lambda\psi_2\psi_1^{-1}))$ cover $L(\chi_1, \chi_2, a)$ as λ varies over all unramified characters with trivial reduction mod p (cf. [GLS15, Thm. 5.4.1, Thm. 6.1.8] and their proofs).

In three dimensions, the situation appears to be considerably more complicated. In addition to the discussion of Section 6.2, we have the following example.

Example 7.3.1. Suppose that $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ is such that

$$\overline{\rho} \sim \begin{pmatrix} \chi_1 & * & * \\ & \chi_2 & * \\ & & \chi_3 \end{pmatrix}.$$

Suppose moreover that $\chi_1|_{I_{\mathbb{Q}_p}} = \omega^{a+1}$, $\chi_2|_{I_{\mathbb{Q}_p}} = \omega^{b+1}$, $\chi_3|_{I_{\mathbb{Q}_p}} = \omega^{c+1}$ with integers $a > b > c > a - (p-1)$, where all gaps in the inequalities are at least 3, and that $\overline{\rho}$ is maximally non-split (i.e. χ_1 is the unique subrepresentation and χ_3 the unique quotient representation). When the χ_i are fixed, the isomorphism class of $\overline{\rho}$ is determined by an invariant $\mathrm{FL}(\overline{\rho}) \in \mathbb{P}^1(\overline{\mathbb{F}}_p) \setminus \{\chi_2(p)\}$. In the global setting of a suitable compact unitary group the Serre weights of $\overline{\rho}$ are almost completely determined in [HLM17]: with the possible addition of the shadow weight $F(c+p-1, b, a-p+1)$, the set of Serre weights equals

$$\begin{cases} \{F(a-1, b, c+1)\} & \text{if } \mathrm{FL}(\overline{\rho}) \notin \{0, \infty\}, \\ \{F(a-1, b, c+1), F(b+p-1, a, c)\} & \text{if } \mathrm{FL}(\overline{\rho}) = 0, \\ \{F(a-1, b, c+1), F(a, c, b-p+1)\} & \text{if } \mathrm{FL}(\overline{\rho}) = \infty. \end{cases}$$

That is, the set of Serre weights consists of one element of $\mathrm{W}_{\mathrm{obv}}(\overline{\rho}^{\mathrm{ss}})$, namely the obvious weight coming from the diagonal characters of $\overline{\rho}$ in their given order, together with a set of shadow weight(s) depending on the parameter $\mathrm{FL}(\overline{\rho})$. The occasional presence of the weights $F(b+p-1, a, c)$ and $F(a, c, b-p+1)$ suggests that there is no naive explicit conjecture for non-semisimple $\overline{\rho}$. We make two further remarks. First, [HLM17] verify that in this setting there exists an ordinary crystalline lift of $\overline{\rho}$ that witnesses the containment $F(a-1, b, c+1) \in \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$. Second, when the maximal non-splitness assumption above is dropped, an upper bound on the set of Serre weights of $\overline{\rho}$ was obtained by Morra–Park [MP17].

7.4. **Shifted weights.** We continue to consider $\overline{\rho}$ such that $\overline{\rho}|_{I_K}$ may not be semisimple. Recall from Section 6.1 that when $n = 2$ and $K = \mathbb{Q}_p$, every component of $\overline{\mathcal{X}}$ labeled by 1 is also labeled by $\mathrm{Sym}^{p-1} \overline{\mathbb{F}}_p^2$; equivalently, every $\overline{\rho}$ with 1 as a Serre weight also has $\mathrm{Sym}^{p-1} \overline{\mathbb{F}}_p^2$ as a Serre weight. This can be viewed as the first instance of the following more general question: for which pairs of Serre weights F, F' does $F \in \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ imply that one must have $F' \in \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ as well? In this case we say that the weight F *entails* the weight F'.

The geometric perspective explained in Section 6.1 (combined with the Breuil–Mézard conjecture) allows a significant reduction to this question. The weight F will entail the weight F' if and only if every component of $\overline{\mathcal{X}}$ labeled by F is also labeled by F'; to check the latter it suffices to check that every generic $\overline{\mathbb{F}}_p$-point (of some component) that has F as a Serre weight also has F' as a Serre weight. In particular, if one believes that the Breuil–Mézard conjecture holds, then one should believe that F entails F' for arbitrary $\overline{\rho}$ as long as the same holds for maximally non-split upper-triangular $\overline{\rho}$ (or even those that are sufficiently generic to lie on just one component of $\overline{\mathcal{X}}$).

In the remainder of this section we will discuss the following specific instance of the weight entailment question.

Definition 7.4.1. If a, b are Serre weights, we say that b is a *shift* of a if there exists $1 \leq i_0 < n$ such that

$$b_{\sigma,i} - a_{\sigma,i} = \begin{cases} p-1 & \text{if } i \leq i_0 \\ 0 & \text{if } i > i_0 \end{cases}$$

for all $\sigma \in S_k$.

Note that this definition only depends on i_0 but not on the choice of representative $a \in (X_1^{(n)})^{S_k}/\sim$, and that we must have $a_{\sigma,i_0} = a_{\sigma,i_0+1}$ for all $\sigma \in S_k$ in order for any shift of a to exist.

Question 7.4.2. *If the weight b is a shift of the weight a, does $a \in \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ entail $b \in \mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ for representations $\overline{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$?*

We equally well ask the same question with $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ replaced by any set that is conjecturally the same as it, such as $\mathrm{W}^{\forall}_{\mathrm{cris}}(\overline{\rho})$, $\mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho})$, or $\mathrm{W}_{\mathcal{S}}(\overline{\rho})$ for any Breuil–Mézard system $\mathcal{S}$.

Remark 7.4.3. This question was suggested to us by the work of Ash–Pollack–Soares [APS04] and Doud [Dou07]: the weight sets conjectured for $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_2)$ in [APS04, §2], resp. for irreducible $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ in [Dou07, Conj. 2.10] are by definition closed under shifts (cf. also [Dou07, Def. 2.7]).

Example 7.4.4. Suppose that $n = 2$ and $p > 2$. Twisting by a suitable character, Question 7.4.2 when $n = 2$ can be reduced to the case where $a = 0$ and $b_\sigma = (p-1, 0)$ for all $\sigma \in S_k$. Since one knows (even if $\overline{\rho}|_{I_K}$ is not semisimple) that $\mathrm{W}_{\mathrm{BT}}(\overline{\rho}) = \mathrm{W}^{\exists}_{\mathrm{cris}}(\overline{\rho}) = \mathrm{W}^{\forall}_{\mathrm{cris}}(\overline{\rho})$ in this setting by the work of [GLS15], an affirmative answer to Question 7.4.2 for any of these sets is equivalent to the statement that if $\overline{\rho} : G_K \to \mathrm{GL}_2(\overline{\mathbb{F}}_p)$ has a regular Barsotti–Tate lift then it also has a crystalline lift with Hodge type some lift of b, which is well known (and can be proved for example via the techniques of [GLS15], or by using the corresponding fact for automorphic forms and the potential modularity techniques of [GK14, App. A]).

Example 7.4.5. We now give an extended discussion of the case $\mathrm{GL}_3/\mathbb{Q}_p$ which suggests to us that Question 7.4.2 may have an affirmative answer in this setting as well. Computational evidence for this (due to [ADP02, APS04, Dou07]) will be reviewed in Section 8.7. Our discussion will be heuristic; in particular we will assume the Breuil–Mézard conjecture, and will extrapolate the labelling of the irreducible components of $\overline{\mathcal{X}}$ from the case $n = 2$ in a speculative fashion. In particular, note that for $n = 2$, the labeling of the irreducible components of $\overline{\mathcal{X}}$ is dictated by the restrictions to inertia of the characters of generic reducible $\overline{\rho}$ on those components, with the subtlety that in the ambiguous case that these weights could either be one-dimensional or twists of the Steinberg representation, we always predict the twist of the Steinberg representation, and only predict the one-dimensional representation in the case that (twists of) these generic $\overline{\rho}$ admit a crystalline lift of Hodge type 0.

In particular, every component labeled by a one-dimensional weight is also labeled by the corresponding twist of the Steinberg representation, and this fact is reflected by the fact that a generic reducible representation admitting a crystalline lift of Hodge type 0 also necessarily admits one of Hodge type corresponding to the Steinberg representation. We will now assume that similar considerations apply for $n = 3$, and see what is implied.

We first suppose that $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ is semisimple and observe that the set $W_{\mathrm{expl}}(\overline{\rho})$ described in Section 7.2 is *not* necessarily closed under shifts, so that a positive answer to Question 7.4.2 means that $W_{\mathrm{expl}}(\overline{\rho})$ is at best a proper subset of $W_{\mathrm{cris}}^{\exists}(\overline{\rho})$. We leave it as an exercise to the reader to check the following. If $\overline{\rho}$ is reducible, then $W_{\mathrm{expl}}(\overline{\rho})$ is closed under shifts. (Use that weights $F(x, y, y)$ or $F(y, y, z)$ are either obvious or obscure.) On the other hand if

$$\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \oplus_{\sigma \in S_{\mathbb{F}_{p^3}}} \omega_\sigma^{(y+2)+p(y+1)+p^2 z}$$

with $0 \le y - z \le p - 2$ then $F(y, y, z) \in W_{\mathrm{expl}}(\overline{\rho})$ but $F(y + p - 1, y, z) \notin W_{\mathrm{expl}}(\overline{\rho})$ (this can be checked by hand, or seen from the tables in Proposition 8.2.14), and dually if

$$\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \oplus_{\sigma \in S_{\mathbb{F}_{p^3}}} \omega_\sigma^{y+p(y+1)+p^2(x+2)}$$

with $0 \le x - y \le p - 2$ then $F(x, y, y) \in W_{\mathrm{expl}}(\overline{\rho})$ but $F(x, y, y - p + 1) \notin W_{\mathrm{expl}}(\overline{\rho})$; and moreover these are the only shifts missing from $W_{\mathrm{expl}}(\overline{\rho})$ for irreducible $\overline{\rho}$. (It is perhaps worth pointing out that shifts do not account for all of the obscure weights of Example 7.2.5, so that neither shifts nor obscure weights alone can account for the difference between $\mathcal{C}(W_{\mathrm{obv}}(\overline{\rho}))$ and the full set of weights.)

Let us now consider the weight entailment problem for weights of the form $F = F(y, y, z)$ and $F' = F(y+p-1, y, z)$; the case of $F(x, y, y)$ and $F(x, y, y-p+1)$ will be dual. Recall (e.g. from Example 7.2.2) that $L_\lambda \otimes_{\overline{\mathbb{Z}}_p} \overline{\mathbb{F}}_p = F(\lambda)$ for both $\lambda = (y, y, z)$ and $\lambda = (y + p - 1, y, z)$, so that we expect that $F(y, y, z)$ (resp. $F(y + p - 1, y, z)$) is a weight for $\overline{\rho}$ if and only if $\overline{\rho}$ has a crystalline lift of Hodge type (y, y, z) (resp. $(y+p-1, y, z)$). Suppose that a component $\mathcal{Z}$ of $\overline{\mathcal{X}}$ has $F(y, y, z)$ among its labels, so that a generic $\overline{\mathbb{F}}_p$-point on $\mathcal{Z}$ corresponds to $\overline{\rho}$ that has a crystalline lift of Hodge type (y, y, z). We wish to know whether $\overline{\rho}$ also has a crystalline lift of Hodge type $(y + p - 1, y, z)$.

If $y - z \le p - 3$, Fontaine–Laffaille theory implies that a generic $\overline{\mathbb{F}}_p$-point on $\mathcal{Z}$ corresponds to $\overline{\rho}$ having the shape

$$(7.4.6) \qquad \overline{\rho}|_{I_{\mathbb{Q}_p}} \sim \begin{pmatrix} \omega^{y+2} & * & * \\ & \omega^{y+1} & * \\ & & \omega^z \end{pmatrix}.$$

The same conclusion seems likely to hold if $y - z = p - 2$: an argument as in [GLS14, Prop. 7.8] shows at least that $\overline{\rho}$ has the same semisimplification as the representation (7.4.6), and it seems quite plausible that the order of the characters on the diagonal will be correct. Suppose this is so.

Let χ_1, χ_2, χ_3 be the characters on the diagonal of $\overline{\rho}$ (in the same order as given in (7.4.6)). Then as long as none of χ_i / χ_j with $i < j$ are cyclotomic it is straightforward to show that a crystalline lift of $\overline{\rho}$ with Hodge type $(y + p - 1, y, z)$ exists. One can even take this lift to be upper-triangular; see for example [GG12, Lem. 3.1.5]. Even if some χ_i / χ_j is cyclotomic, it is reasonable to imagine that the same conclusion holds; e.g. when $y - z \le p - 3$ this is immediate from [GHLS15, Cor. 2.3.5].

Alternately, it is plausible that any $\overline{\rho}$ having the shape (7.4.6) and having a crystalline lift of Hodge type (y, y, z) has an *ordinary* such lift, with characters down the diagonal having Hodge–Tate weights $y + 2$, $y + 1$, z (in that order); cf. the first remark at the end of Example 7.3.1, as well as the discussion of the case

$n = 2$ and $K/\mathbb{Q}_p$ arbitrary in Section 7.3. Write this lift as an extension of a two-dimensional crystalline representation V (with Hodge–Tate weights $\{y + 1, z\}$) by a character W. One may then hope to produce the desired lift of $\overline{\rho}$ of Hodge type $(y + p - 1, y, z)$ by considering extensions of V by unramified twists of $W \otimes \varepsilon^{p-1}$.

We remark that the above arguments are agnostic regarding the case $y - z = p - 1$. However, it is at least the case for $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ semisimple that the set $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ contains $F(y + p - 1, y, y - p + 1)$ whenever it contains $F(y + p - 1, y, y)$.

When $n > 3$ the heuristic arguments in Example 7.4.5 at least make it plausible that Question 7.4.2 has an affirmative answer for shifts of weights $F = F(a_1, \ldots, a_n)$ with $a_1 - a_n$ small (e.g. when $a_1 - a_n \le p - n$, so that Fontaine–Laffaille theory still determines the shape of $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ for $\overline{\rho}$ corresponding to a generic $\overline{\mathbb{F}}_p$-point on a component of $\overline{\mathcal{X}}$ labeled by F).

7.5. Summary.
We briefly summarize the Serre weight conjectures that we have explained in this section.

Definition 7.5.1. Let $\overline{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ be a representation.
- If the generalised Breuil–Mézard conjecture holds, we define $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ to be the set of Serre weights a such that $\mu_a(\overline{\rho}) > 0$.
- We define $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ to be the set of Serre weights a such that $\overline{\rho}$ has a crystalline lift of Hodge type λ_a for some lift λ_a of a.
- We define $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$ to be the set of Serre weights a such that $\overline{\rho}$ has a crystalline lift of Hodge type λ_a for every lift λ_a of a.
- If $\overline{\rho}|_{I_K}$ is semisimple, we define a non-empty set of obvious weights $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ in Definition 7.1.3, and a set of explicit weights $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \supset \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$ in Definition 7.2.3.

Conjecture 7.5.2. *Let $\overline{\rho} : G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ be a representation. Assume that $\overline{\rho}|_{I_K}$ is semisimple.*
- (i) *We have $\mathcal{C}(\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$.*
- (ii) *The sets $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ and $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$ depend only on $\overline{\rho}|_{I_K}$, as does $\mathrm{W}_{\mathrm{BM}}(\overline{\rho})$ if it is defined (i.e. if the generalised Breuil–Mézard conjecture holds).*
- (iii) *We have $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$.*
- (iv) *If the generalised Breuil–Mézard conjecture holds then $\mathrm{W}_{\mathrm{BM}}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$.*
- (v) *If $K/\mathbb{Q}_p$ is unramified, and $\overline{\rho}|_{I_K}$ is sufficiently generic, then $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho}) = \mathrm{W}_{\mathrm{expl}}(\overline{\rho}) = \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\overline{\rho}))$.*

Conjecture 7.5.3. *If each $\overline{r}|_{I_{F_v}}$ is semisimple, then the weight part of Serre's conjecture (Conj. 2.1.5) holds with $\mathrm{W}_v(\overline{r}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{r}|_{G_{F_v}})$.*

Finally (assuming again that $\overline{\rho}|_{I_K}$ is semisimple), we recall that $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ by definition; that if the generalised Breuil–Mézard conjecture holds then we have $\mathrm{W}_{\mathrm{BM}}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho}) = \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ (cf. Lemma 5.1.3); and that if the set $\mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ depends only on $\overline{\rho}|_{I_K}$, then $\mathrm{W}_{\mathrm{obv}}(\overline{\rho}) \subset \mathrm{W}_{\mathrm{cris}}^{\exists}(\overline{\rho})$ (and similarly for $\mathrm{W}_{\mathrm{cris}}^{\forall}(\overline{\rho})$).

8. Existing conjectures in the literature

In this section we review the theoretical and computational evidence for our conjectures, beyond the case $n = 2$ which was discussed in detail above. We also make comparisons with other conjectures in the literature.

8.1. **The case of** $\mathrm{GL}_3(\mathbb{Q}_p)$. Take $K = \mathbb{Q}_p$ and fix an odd and irreducible representation $\bar{r} : G_\mathbb{Q} \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ such that $\bar{r}|_{I_{\mathbb{Q}_p}}$ is semisimple. The first Serre weight conjectures in this context were made by Ash, Doud, Pollack, and Sinnott [AS00], [ADP02]. We will discuss their work in Section 8.4 below. Later in the paper (Section 10), we will show that Conjecture 7.2.7 agrees with the Serre weight conjecture made by the second author in [Her09]. (In fact we will ultimately work in a somewhat more general context than this.) Recall, however, that in Conjecture 7.2.7 the representation $\bar{r}|_{I_{\mathbb{Q}_p}}$ is assumed to be sufficiently generic. In this next section we will check that in the 3-dimensional case the conjecture of [Her09] agrees *completely* with the explicit set of weights described in the previous section.

8.2. **The conjecture of** [Her09] **for** GL_3 **over** $\mathbb{Q}_p$. Recall that [Her09, Conj. 6.9] predicts the set of *regular* Serre weights for which a given irreducible, odd representation $\bar{r} : G_\mathbb{Q} \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ is automorphic. Regular Serre weights are defined as follows.[2]

Definition 8.2.1. A weight $a \in \mathrm{W}(k, n)$ is said to be *regular* if $a_{\sigma,i} - a_{\sigma,i+1} < p - 1$ for all σ, i, and *irregular* otherwise. Let $\mathrm{W}_{\mathrm{reg}} \subset \mathrm{W}(k, n)$ be the set of regular weights.

The set of Serre weights predicted in [Her09] is denoted $\mathrm{W}^?(\bar{r}|_{I_{\mathbb{Q}_p}})$, so we want to check that $\mathrm{W}^?(\bar{\rho}|_{I_{\mathbb{Q}_p}}) = \mathrm{W}_{\mathrm{expl}}(\bar{\rho}) \cap \mathrm{W}_{\mathrm{reg}}$ for a local representation $\bar{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ such that $\bar{\rho}|_{I_{\mathbb{Q}_p}}$ is semisimple. To describe the set $\mathrm{W}^?(\bar{\rho}|_{I_{\mathbb{Q}_p}})$, we begin with the following definitions.

Definition 8.2.2. Suppose that $(w, \mu) \in S_n \times \mathbb{Z}^n$. Let $w = w_1 \cdots w_m$ be the unique decomposition of the permutation w into disjoint cycles (including trivial cycles), and write $\mu = (\mu_1, \ldots, \mu_n)$.

   (i) If $w_i = (c_0 \cdots c_{d_i-1})$ we set $N_i = \sum_{j=0}^{d_i-1} p^j \mu_{c_j}$, write $k_i = \mathbb{F}_{p^{d_i}}$, and define $\tau_{d_i}(w_i, \mu)$ to be the isomorphism class of the inertial Galois representation $\oplus_{\sigma \in S_{k_i}} \omega_\sigma^{N_i}$ of dimension d_i.

   (ii) We define $\tau(w, \mu)$ to be the isomorphism class of the inertial Galois representation $\oplus_{i=1}^m \tau_{d_i}(w_i, \mu)$ of dimension n.

   (iii) We say that the pair (w, μ) is *good* if for all $1 \le i \le m$ and for all $d \mid d_i$, $d \ne d_i$ we have $(p^d - 1)N_i \not\equiv 0 \pmod{p^{d_i} - 1}$.

It is straightforward to verify that these definitions do not depend on any of the choices involved.

Remark 8.2.3. The above definitions are concrete instances of the more general and more canonical [Her09, (6.15)] and [Her09, Def. 6.19]. We will recall this more canonical definition, and extend it to other groups, in Proposition 9.2.3.

If $\bar{\rho}|_{I_{\mathbb{Q}_p}} \cong \tau(w, \mu)$, the condition that the pair (w, μ) is good means, concretely, that the dimensions of the Jordan–Hölder factors of $\bar{\rho}$ correspond to the cycle type of w.

Example 8.2.4. Suppose that w is the transposition swapping i and j. Then the pair (w, μ) is good if and only if $p + 1 \nmid \mu_i + p\mu_j$, or equivalently $p + 1 \nmid \mu_i - \mu_j$. In

---

particular this is always the case if $|j - i| = 1$ and $\mu = \lambda + (n - 1, \ldots, 1, 0)$ where λ is the lift of a Serre weight.

Example 8.2.5. Suppose that $n = 3$ and $w = (i\,j\,k)$ is a 3-cycle. If $\mu = \lambda + (2, 1, 0)$ where λ represents a regular Serre weight, it is a straightforward exercise to verify that the pair (w, μ) is always good.

Suppose for the remainder of this section that $n = 3$, so that W_{reg} refers to the regular weights in $W(\mathbb{F}_p, 3)$, and write $\eta = (2, 1, 0)$.

Definition 8.2.6. We define $X_{\mathrm{reg}}^{(3)} \subset X_1^{(3)}$ to be the set of triples such that $a \geq b \geq c$ and $a - b, b - c \leq p - 2$. Note that $X_{\mathrm{reg}}^{(3)}/\!\sim \,\cong W_{\mathrm{reg}}$.

If $(x, y, z) \in \mathbb{Z}^3$, let $\mathrm{reg}(x, y, z)$ be the unique element of W_{reg} represented by some $(x', y', z') \in X_{\mathrm{reg}}^{(3)}$ with $(x', y', z') \equiv (x, y, z)$ modulo $(p - 1)\mathbb{Z}^3$.

Definition 8.2.7. Following [Her09, Prop. 3.18] (see also Example 7.2.2), we define the function (from weights to sets of weights)

$$r\big(F(x, y, z)\big) = \begin{cases} \{F(x, y, z)\} & \text{if } x - z \geq p - 2, \\ \{F(x, y, z), \ F(z + p - 2, y, x - p + 2)\} & \text{if } x - z < p - 2. \end{cases}$$

Following [Her09, Def. 7.3] we set $\mathcal{A}(\mu) = r(\mathrm{reg}(\mu - \eta))$ for each $\mu \in X_1^{(3)}$. Note that if $F \in W_{\mathrm{reg}}$ then $r(F) \subset W_{\mathrm{reg}}$, and so $\mathcal{A}(\mu) \subset W_{\mathrm{reg}}$ for any $\mu \in X_1^{(3)}$.

The set $W^?(\overline{\rho}|_{I_{\mathbb{Q}_p}})$ in [Her09] is defined in terms of a Deligne–Lusztig representation associated to $\overline{\rho}|_{I_{\mathbb{Q}_p}}$. In the three-dimensional case we have the following explicit description of this set, which is all we will need for the purposes of this section.

Proposition 8.2.8. ([Her09, Prop. 7.4]) *Let* $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ *be a representation such that* $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ *is semisimple. Set*

$$\mathcal{C}(\overline{\rho}|_{I_{\mathbb{Q}_p}}) = \{\mu \in X_1^{(3)} : \text{there exists } w \in S_3 \text{ with } (w, \mu) \text{ good and } \overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \tau(w, \mu)\}.$$

Then

$$W^?(\overline{\rho}|_{I_{\mathbb{Q}_p}}) = \bigcup_{\mu \in \mathcal{C}(\overline{\rho}|_{I_{\mathbb{Q}_p}})} \mathcal{A}(\mu).$$

Example 8.2.9. Suppose that $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ is unramified. Then $W_{\mathrm{expl}}(\overline{\rho})$ consists of the four weights $F(p - 3, -1, -p + 1)$, $F(p - 2, -1, -1)$, $F(p - 2, p - 2, -1)$, $F(p - 2, -1, -p)$ by Examples 7.1.9 and 7.2.5. On the other hand, $\mathcal{C}(\overline{\rho}|_{I_{\mathbb{Q}_p}})$ consists of all $\mu \in X_1^{(3)} \cap (p - 1)\mathbb{Z}^3$, and $\mathrm{reg}(\mu - \eta) = F(p - 3, -1, -p + 1)$ for all $\mu \in \mathcal{C}(\overline{\rho}|_{I_{\mathbb{Q}_p}})$. We therefore have $W^?(\overline{\rho}|_{I_{\mathbb{Q}_p}}) = \{F(p - 3, -1, -p + 1)\}$. Hence we confirm that $W^?(\overline{\rho}|_{I_{\mathbb{Q}_p}}) = W_{\mathrm{expl}}(\overline{\rho}) \cap W_{\mathrm{reg}}$, since the other three weights in $W_{\mathrm{expl}}(\overline{\rho})$ are irregular.

We can now prove the main result of this section.

Proposition 8.2.10. *Let* $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ *be a representation such that* $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ *is semisimple. Then we have* $W_{\mathrm{expl}}(\overline{\rho}) \cap W_{\mathrm{reg}} = W^?(\overline{\rho}|_{I_{\mathbb{Q}_p}})$.

Remark 8.2.11. In Proposition 8.2.14 we will describe the irregular weights in $W_{\mathrm{expl}}(\overline{\rho})$.

Proof. Note that obscure weights for $\mathrm{GL}_3/\mathbb{Q}_p$ were analysed completely in Example 7.2.5, and were all found to be irregular, so that $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \cap \mathrm{W}_{\mathrm{reg}}$ consists entirely of obvious and shadow weights. It is then easy to see from the definition of $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$ (or alternatively from Proposition 9.3.7 in the next section), together with the discussion of Examples 7.2.2, that

$$\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \cap \mathrm{W}_{\mathrm{reg}} = \bigcup_{\mu \in \mathcal{C}'(\overline{\rho}|_{I_{\mathbb{Q}_p}})} \mathcal{A}(\mu)$$

where

$$\mathcal{C}'(\overline{\rho}|_{I_{\mathbb{Q}_p}}) = \{\mu \in X^{(3)}_{\mathrm{reg}} + \eta : \text{there exists } w \in S_3 \text{ with } \overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \tau(w, \mu)\}.$$

To show the inclusion "$\subset$" in the Proposition, we have to consider $\mu \in X^{(3)}_{\mathrm{reg}} + \eta$ such that there exists $w \in S_3$ with (w, μ) not good and $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \tau(w, \mu)$. By Examples 8.2.4 and 8.2.5 this only happens when $w = (1\,3)$, $\mu = (a, b, a - p - 1)$. The condition $\mu \in X^{(3)}_{\mathrm{reg}} + \eta$ forces $2 \le a - b \le p - 1$. Then $\tau(w, \mu) \cong \omega^{a-1} \oplus \omega^{a-1} \oplus \omega^b$ and so $\tau(w, \mu) \cong \tau(1, \mu')$ with $\mu' = (a - 1, b, a - p)$. Since $(1, \mu')$ is good and $\mu' \in X^{(3)}_1$, we have $\mu' \in \mathcal{C}(\overline{\rho}|_{I_{\mathbb{Q}_p}})$. Directly from Definition 8.2.7 one calculates that $\mathcal{A}(\mu) = \{F(a - 2, b - 1, a - p - 1)\}$ and $\mathcal{A}(\mu') = \{F(a - 3, b - 1, a - p), F(a - 2, b - 1, a - p - 1)\}$. In particular $\mathcal{A}(\mu) \subset \mathcal{A}(\mu') \subset \mathrm{W}^?(\overline{\rho}|_{I_{\mathbb{Q}_p}})$, as required.

For the reverse inclusion, one must consider $\mu \in X^{(3)}_1 \setminus (X^{(3)}_{\mathrm{reg}} + \eta)$ such that there exists $w \in S_3$ with (w, μ) good and $\overline{\rho}|_{I_{\mathbb{Q}_p}} \cong \tau(w, \mu)$. There are three cases. First, if $\mu = (a, a, a)$, then $w = 1$ and $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ is a sum of three copies of ω^a. After a twist we can reduce to the unramified case, which we have already considered in Example 8.2.9. (Alternately, just note that $\mu' = (a + p - 1, a, a - p + 1) \in X^{(3)}_{\mathrm{reg}} + \eta$ with $\tau(1, \mu') \cong \tau(1, \mu)$ and $\mathcal{A}(\mu') = \mathcal{A}(\mu)$.)

Second, suppose $\mu = (a, a, c)$ with $0 < a - c \le p - 1$, so that $\mathcal{A}(\mu) = \{F(a + p - 3, a - 1, c)\}$. Without loss of generality we may assume $w \in \{1, (1\,3), (1\,2\,3)\}$ (note that the pair $((1\,2), \mu)$ is not good). For each of these three possibilities it is easy to check that there exists $w' \in S_3$ such that $\tau(w, \mu) \cong \tau(w', \mu')$ and $\mu' \in \{(a + p - 1, a, c), (c + p, a, a - 1)\} \subset X^{(3)}_{\mathrm{reg}} + \eta$; for instance if $w = (1\,2\,3)$ we take $w' = (1\,3\,2)$ and $\mu' = (c + p, a, a - 1)$. In particular $\mu' \in \mathcal{C}'(\overline{\rho}|_{I_{\mathbb{Q}_p}})$ and either $\mathcal{A}(\mu) = \mathcal{A}(\mu')$ (if $\mu' = (a + p - 1, a, c)$) or $\mathcal{A}(\mu) \subset \mathcal{A}(\mu')$ (if $\mu' = (c + p, a, a - 1)$), as required.

Finally, if $\mu = (a, c, c)$ with $0 < a - c \le p - 1$ then one can argue as in the previous case; alternately we can reduce to the previous case by duality using the following lemma, valid for GL_n (cf. [Her09, Prop. 6.23(ii)]), whose proof is straightforward. □

Lemma 8.2.12. *We have* $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}^\vee) = \{F^\vee \otimes \det^{1-n} : F \in \mathrm{W}_{\mathrm{expl}}(\overline{\rho})\}.$

We now describe the irregular weights in $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$. As a preliminary, we observe that the possibilities for $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ are given by the following alternatives.

Lemma 8.2.13. *Suppose that* $\tau : I_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ *is semisimple and extends to a representation of* $G_{\mathbb{Q}_p}$. *Then precisely one of the following alternatives holds:*

    (i) $\tau \cong \tau(1, (a, b, c))$ *where* $a \ge b \ge c$ *and* $a - c \le p - 1$,

    (ii) $\tau \cong \tau((2\,3), (a, b, c))$ *where* $a \ge b > c$ *and* $a - c \le p - 1$,

    (iii) $\tau \cong \tau((1\,2\,3), (a, b, c))$ *where* $a > b \ge c$ *and* $a - c \le p$,

(iv) $\tau^\vee \cong \tau((1\,2\,3),(a,b,c))$ *where* $a > b \geq c$ *and* $a - c \leq p$.

Moreover, in (i) the triple (a,b,c) *is unique up to the equivalence relation generated by* $(a,b,c) \sim (c+p-1,a,b)$*; in (ii) the triple* (a,b,c) *is unique up to translation by* $(p-1,p-1,p-1)\mathbb{Z}$*; in (iii) and (iv), the triple* (a,b,c) *is unique up to the equivalence relation generated by* $(a,b,c) \sim (c+p,a-1,b)$*.*

Proof. Parts (iii) and (iv) follow from [EGH13, Lem. 5.2.2]. The rest of the proof is left to the reader. $\square$

Proposition 8.2.14. *Let* $\overline{\rho} : G_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ *be a representation such that* $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ *is semisimple. Then the weights in* $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \setminus \mathrm{W}_{\mathrm{reg}}$ *are described as follows.*

(i) *Suppose that* $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ *is as in Lemma 8.2.13(i). The set* $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \setminus \mathrm{W}_{\mathrm{reg}}$ *consists of the weights* $F(\mu - \eta)$ *for triples* μ *as in the second column of the following table, under the conditions as in the first column.*

condition	μ
$a - b = 1,\ b - c \neq 0$	$(b+p, b, c)$
$a - b = 1,\ b - c \leq 1$	$(b+p, b, c-p+1)$
$b - c = 1,\ a - c \neq p - 1$	$(c+p, c, a-p+1)$
$b - c = 1,\ a - c \geq p - 2$	$(c+p, c, a-2p+2)$
$a - c = p - 2,\ a - b \neq 0$	$(a+p, a, b)$
$a - c = p - 2,\ a - b \leq 1$	$(a+p, a, b-p+1)$
$b - c = 1,\ a - b \neq 0$	$(a, b, b-p)$
$b - c = 1,\ a - b \leq 1$	$(a+p-1, b, b-p)$
$a - b = 1,\ a - c \neq p - 1$	$(c+p-1, a, a-p)$
$a - b = 1,\ a - c \geq p - 2$	$(c+2p-2, a, a-p)$
$a - c = p - 2,\ b - c \neq 0$	$(b, c, c-p)$
$a - c = p - 2,\ b - c \leq 1$	$(b+p-1, c, c-p)$
$a - b = 0$	$(b+p, b, c-1),\ (c+p, a, a-p)$
$b - c = 0$	$(c+p, c, a-p),\ (a+1, b, b-p)$
$a - c = p - 1$	$(a+p, a, b-1),\ (b+1, c, c-p)$

(ii) *Suppose that* $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ *is as in Lemma 8.2.13(ii). The set* $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \setminus \mathrm{W}_{\mathrm{reg}}$ *consists of the weights* $F(\mu - \eta)$ *for triples* μ *as in the second column of the following table, under the conditions as in the first column.*

condition	μ
$b - c = 1,\ a - c \neq p - 1$	$(c+p, c, a-p+1)$
$a - b = 1$	$(b+p, b, c)$
$a - c = 2$	$(c+p+1, c+1, b-p)$
$a - b = 0$	$(b+p, b, c-1),\ (c+p, a, a-p)$
$b - c = 1,\ a - b \neq 0$	$(a, b, b-p)$
$a - c = p - 2$	$(b, c, c-p)$
$a - b = p - 3$	$(c+p, b-1, b-p-1)$
$a - c = p - 1$	$(b+1, c, c-p),\ (a+p, a, b-1)$

(iii) *Suppose that* $\overline{\rho}|_{I_{\mathbb{Q}_p}}$ *is as in Lemma 8.2.13(iii). The set* $\mathrm{W}_{\mathrm{expl}}(\overline{\rho}) \setminus \mathrm{W}_{\mathrm{reg}}$ *consists of the weights* $F(\mu - \eta)$ *for triples* μ *as in the second column of the following table, under the conditions as in the first column.*

condition	μ
$a - b = 2$	$(c + p, a - 1, a - p - 1)$
$a - b = 1,\ a - c \neq 1$	$(c + p + 1, a - 1, a - p - 1)$
$a - c = p - 1$	$(b + 1, c, c - p)$
$a - c = p,\ b - c \neq p - 1$	$(b + 2, c, c - p)$
$b - c = 1$	$(a, b, b - p)$
$b - c = 0,\ a - b \neq p$	$(a + 1, b, b - p)$

Remark 8.2.15. Suppose that $\bar{\rho}|_{I_{\mathbb{Q}_p}}$ is as in Lemma 8.2.13(iv). The description of set $W_{\mathrm{expl}}(\bar{\rho}) \setminus W_{\mathrm{reg}}$ can be extracted from Proposition 8.2.14(iii) by duality using Lemma 8.2.12.

Also, we remind the reader that in part (iii), the list of weights given here does not include the shifted weights described at the beginning of Example 7.4.5.

Proof. Note that if $F(x, y, z)$ is a weight with $x - z < p - 2$, then the weight $F(z + p - 2, y, x - p + 2)$ as in Example 7.2.2 is regular, so that all the weights in these tables must either be obvious or obscure. The obscure weights are analyzed in Example 7.2.5, and are listed in the final three rows of the table in (i), and the second half of rows 4 and 8 of the table in (ii). (These rows also contain some obvious weights.)

Suppose, then, that the irregular weight $F\big((b + p, b, c) - \eta\big)$ lies in $W_{\mathrm{obv}}(\bar{\rho})$, with $0 < b - c \leq p$. It follows from the definitions that $\bar{\rho}|_{I_{\mathbb{Q}_p}} \cong \tau(w, (b + p, b, c))$ for some $w \in S_3$. For each w one then expresses $\bar{\rho}|_{I_{\mathbb{Q}_p}}$ in the form of Lemma 8.2.13 (in all possible ways) to generate the lines of the above tables containing triples of the form $\mu = (\mu_1, \mu_2, \mu_3)$ in the second column with $\mu_1 - \mu_2 = p$ (relabelling as necessary, as well as keeping in mind the equivalence relations in Lemma 8.2.13). For example, if $p > 2$ then rewriting $\tau(1, (b + p, b, c))$ as $\tau(1, (b + 1, b, c))$ when $0 < b - c \leq p - 2$ and as $\tau(1, (b + 1, b, c + p - 1))$ when $p - 1 \leq b - c \leq p$ gives the first two lines of the table in (i); then the next four lines come from the first two lines via the equivalence relation of Lemma 8.2.13. We leave the rest of the details as an exercise for the reader (for which we suggest considering the case $p = 2$ separately, at least in the cases when $\bar{\rho}$ is reducible). Dualising using Lemma 8.2.12 we obtain a similar list for $\bar{\rho}|_{I_{\mathbb{Q}_p}}$ for the weight $F\big((a, b, b - p) - \eta\big)$. □

8.3. **The results of** [EGH13]. The paper [EGH13] considers the weight part of Serre's conjecture for Galois representations $\bar{r} : G_F \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$, where F is a totally real field in which p splits completely, and $\bar{r}|_{G_{F_v}}$ is irreducible for each $v | p$. The main results are proved with respect to some abstract axioms, which are in particular satisfied for the cohomology of forms of $\mathrm{U}(3)$ which are compact at infinity, and show that if each $\bar{r}|_{G_{F_v}}$ satisfies a mild genericity condition, then the set of weights in which $\bar{r}$ is automorphic contains the set of weights predicted by the conjecture of [Her09], and that any other weight for which $\bar{r}$ is automorphic is non-generic. (Here a weight is generic if it is sufficiently far away from the walls of any alcove. As with the definition of genericity for a Galois representation, this will be made precise in Section 10 below.)

In the light of the discussion of Section 8.1, these results are completely consistent with our conjectures.

8.4. **The conjecture of Ash, Doud, Pollack, and Sinnott for** GL_n **over** $\mathbb{Q}_p$. Let $\bar{r} : G_{\mathbb{Q}} \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ be odd and irreducible. The first Serre weight conjectures

for such $\bar{r}$ with $n > 2$ were made by Ash, Doud, Pollack, and Sinnott [AS00], [ADP02]. When $n = 3$ a detailed comparison between their conjecture and the conjecture of [Her09] can be found in *ibid.*, §7. The purpose of this section is to note the following result.

Proposition 8.4.1. *Let $\bar{r}$ be as above and suppose that $\bar{r}|_{I_{\mathbb{Q}_p}}$ is semisimple and sufficiently generic. Then the Serre weights predicted in* [ADP02, Conj. 3.1] *are a subset of* $W_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}})$.

For the term "sufficiently generic" we once again refer the reader to Definition 10.1.12 (but see also Example 7.2.8 for the case when $\bar{r}|_{G_{\mathbb{Q}_p}}$ is a sum of characters).

Remark 8.4.2. It turns out that the subset of Serre weights predicted in [ADP02, Conj. 3.1] consists of a mix of some (but not all) obvious weights for $\bar{r}|_{G_{\mathbb{Q}_p}}$ and some (but very far from all) shadow weights. In any case we stress that [ADP02] do not claim to predict the full set of weights for $\bar{r}$.

Since the proof of Proposition 8.4.1 will make use of terminology and results from Sections 9–10 we defer the proof until Section 10.4. (We may safely do so because nothing in the paper depends logically on Proposition 8.4.1.)

We make two further remarks. First, [ADP02] still give a conjectural set of Serre weights even when $\bar{r}|_{I_{\mathbb{Q}_p}}$ is not semisimple, a context in which we do not make an explicit prediction; we have no heuristic by which to predict whether or not all the Serre weights conjectured by [ADP02] in this context are indeed weights of $\bar{r}$. Second, we note that when $\bar{r}|_{G_{\mathbb{Q}_p}}$ is irreducible, Doud [Dou07, Conj. 2.10] predicts precisely the set $W_{\mathrm{obv}}(\bar{r}|_{G_{\mathbb{Q}_p}})$ together with all weight shifts as described in Section 7.4.

8.5. **The results of** [BLGG14]**.** The article [BLGG14] applies the machinery of the paper [BLGGT14] to the problem of the weight part of Serre's conjecture for unitary groups over CM fields. As explained in Section 4, a lack of general results on the potential diagonalisability of crystalline representations limits the scope for proving general comprehensive results. However, under mild conditions the paper shows that when $n = 3$, p splits completely in an imaginary CM field F, and $\bar{r} : G_F \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ is such that $\bar{r}|_{G_{F_v}}$ is semisimple for each $v|p$, then $\bar{r}$ is automorphic for every obvious predicted weight in the sense of Definition 7.1.3 above. This is, of course, consistent with our conjectures.

8.6. GSp_4**.** The paper [HT13] formulates a version of the weight part of Serre's conjecture for irreducible representations $\bar{r} : G_{\mathbb{Q}} \to \mathrm{GSp}_4(\overline{\mathbb{F}}_p)$, under the assumption that $\bar{r}|_{G_{\mathbb{Q}_p}}$ is a sum of characters, and under a mild regularity condition on the weights. The formulation follows that of [Her09], and is a special case of the more general conjectures that we formulate in Sections 9 and 10, which show that these conjectures are also consistent with the philosophy of this paper.

8.7. **Computational evidence.** The paper [Tor12] carried out computations for the weight part of Serre's conjecture (Conj. 2.1.5) in the case that $F = \mathbb{Q}(i)$ and $n = 2$. In this setting the Taylor–Wiles method is not available in anything like the generality required to make arguments along the lines of those explained in Remark 4.2.5 for totally real fields, and so there are no theoretical results on the

weight part of Serre's conjecture. However, the computations of [Tor12] are all consistent with the expectation that the weight part of Serre's conjecture will behave identically in this case, and are thus completely consistent with the conjectures of this paper. (It is explained in Section 7 of [Tor12] that there was one example where two expected weights were not found; however Mehmet Haluk Şengün has independently reproduced the calculations of [Tor12] in unpublished work, and has found complete agreement, *except* that the two "missing" weights were also obtained.)

The paper [ADP02] explicitly carries out calculations for the weight part of Serre's conjecture (Conj. 2.1.5) for the case $F = \mathbb{Q}$ and $n = 3$. These calculations, and some additional calculations that Doud and Pollack carried out at the request of the second author, are all consistent with the conjecture of [Her09], and thus with our conjectures; see Section 8 of [Her09] for more details of this.

Remark 8.7.1. In fact, the calculations in [ADP02] would have been consistent with Conjectures 5.1.7 and 5.1.10 even without the hypothesis that $\overline{\rho}|_{I_K}$ is semisimple. To be precise, set $F = F(x, y, z)$ with $x - z < p - 2$, and set $F' = F(z + p - 2, y, x - p + 2)$; then in the calculations of [ADP02], one finds that whenever F is a Serre weight for some $\overline{\rho}$, so is F'.

On the other hand, as we have explained in Section 6.2, Conjectures 5.1.7 and 5.1.10 are now known to be false if one omits the hypothesis that $\overline{\rho}|_{I_K}$ is semisimple. One thus expects that every $\overline{\rho}$ with Serre weight F considered by [ADP02] happens to lie on the (codimension one) intersection between the two components of the stack $\overline{\mathcal{X}}$ labeled by the Serre weights F and F' (cf. the discussion in Section 6); while at the time of writing we do not know for certain that this is the case, an examination of the explicit representations considered by [ADP02] suggests that they are indeed in a rather special position.

8.8. **Computational evidence for irregular Serre weights.** Consider a representation $\overline{r} : G_{\mathbb{Q}} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ that is odd and irreducible, and such that $\overline{r}|_{I_p}$ is semisimple. Since our conjectures cover more weights than those of [Her09] (namely, the weights which are irregular), we now give computational evidence for such weights.

8.8.1. *Examples from* [ADP02]. The weight predictions in [ADP02, Conj. 3.1] are ambiguous for irregular weights: if $x \equiv y$ or $y \equiv z \pmod{p-1}$, then their weight prediction of $F(x, y, z)'$ means that $\overline{r}$ occurs in *at least one* weight $F(x', y', z')$ such that $(x', y', z') \equiv (x, y, z) \pmod{(p-1)\mathbb{Z}^3}$ (so there are either two or four such weights, the latter precisely when $x \equiv y \equiv z \pmod{p-1}$).

In [ADP02] there are six examples with $\overline{r}|_{G_{\mathbb{Q}_p}}$ of length two (see [ADP02, Table 10]) and two examples with $\overline{r}|_{G_{\mathbb{Q}_p}}$ irreducible (see [ADP02, §7.2]) where ambiguous weight predictions occur. In each case, except for the second entry of [ADP02, Table 10] which lies outside the scope of his program, Doud has checked for us that $\overline{r}$ appears in *both* weights implied by the ambiguous notation (testing Hecke eigenvalues for all $l \leq 47$, as in [ADP02]). This is consistent with our conjecture, as all weights in question are obvious.

8.8.2. *Examples from* [Dou07]. The paper [Dou07] provided computational evidence for several $\overline{r}$ with $\overline{r}|_{G_{\mathbb{Q}_p}}$ occurring in irregular weights. Recall that for such $\overline{r}$ his predicted weight set is obtained by adjoining all weight shifts to $\mathrm{W}_{\mathrm{expl}}(\overline{r}|_{G_{\mathbb{Q}_p}})$.

In most of his examples the irregular weights are obvious for $\bar{r}|_{G_{\mathbb{Q}_p}}$; see Section 8.8.4 below for the remaining cases.

8.8.3. *Obscure weights.* Consider the irreducible polynomial $f(x) = x^4 - x^3 + 5x^2 - 4x + 3$ over $\mathbb{Q}$ with Galois group A_4, as in [ADP02, Ex. 5.4]. By taking the unique 3-dimensional irreducible representation of A_4 over $\overline{\mathbb{F}}_{13}$, we obtain a Galois representation $\bar{r}$ as above with $\bar{r}|_{I_{13}} \cong \tau(1, (6, 6, 0))$. We have $\mathrm{W}_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}}) = \{F(16, 5, 0), F(16, 11, 6), F(22, 17, 6), F(17, 11, 5), F(29, 17, 11), F(23, 17, 5)\}$, where the last two weights are obscure. Doud could provide for us computational evidence that $\bar{r}$ is automorphic in each of these weights, and showed that it doesn't occur in any other irregular weight $F(a, b, c)$ with $a - c \notin \{21, 24\}$. (Note that the central character forces $a - c \equiv 0 \pmod{3}$ for irregular weights.)

8.8.4. *Weight shifts.* In the literature we found evidence for shifted weights that are not contained in $\mathrm{W}_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}})$ in the following cases. First, when $p = 2$, [APS04, Table 3] contains three examples where $\bar{r}|_{I_2} \cong \tau((1\,2\,3), (1, 0, 0))$ (or its dual). We have that $\mathrm{W}_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}}) = \{F(0, 0, 0), F(1, 1, 0), F(2, 1, 0)\}$ (all are obvious), and $F(1, 0, 0)$ is a shift of $F(0, 0, 0)$. In each case [APS04] gives computational evidence that $\bar{r}$ occurs in all four weights.

Second, when $p = 3$, [Dou07, §5.3] considers an example in which $\bar{r}|_{I_3} \cong \tau((1\,2\,3), (2, 0, 0))$ (or its dual; these are denoted by $m = 2$, $m = 8$ in [Dou07, Table 2]). In this case

$$\mathrm{W}_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}}) = \{F(1, 1, 1), F(1, 0, 0), F(3, 2, 0), F(3, 3, 1), F(5, 3, 1), F(2, 1, 0)\}$$

(all but the last weight being obvious), and $F(3, 1, 1)$ is a shift of $F(1, 1, 1)$. Doud [Dou07] gives computational evidence for all seven weights. (Note that $F(3, 1, 1)$ is missing from [Dou07, Table 2], but Doud confirmed to us that this is just a typo.)

9. Unramified groups

We now explain how to extend the definition of the set of weights $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$, as well as the set of weights $\mathrm{W}^?(\overline{\rho})$ defined in [Her09], to the more general setting of unramified groups over $\mathbb{Q}_p$. In this section and the next, we will use Γ_K instead of G_K to denote the absolute Galois group of K, to avoid confusion with our notation for algebraic groups.

9.1. *L-groups and L-parameters.* Let G be a connected reductive group over $\mathbb{Z}_p$, i.e. a smooth affine algebraic group whose geometric fibres are connected reductive. Then $G \times \mathbb{Q}_p$ is unramified (i.e. quasisplit and split over an unramified extension of $\mathbb{Q}_p$), and conversely every unramified group over $\mathbb{Q}_p$ arises in this way (by choosing a hyperspecial point in the building). Let B be a Borel subgroup of G with Levi subgroup $T \subset B$, so T is a maximal torus of G. Note that we have a canonical identification of character groups $X(T \times \overline{\mathbb{Q}}_p) \cong X(T \times \overline{\mathbb{F}}_p)$, which is compatible with the Galois action of $\Gamma_{\mathbb{Q}_p} \twoheadrightarrow \Gamma_{\mathbb{F}_p}$. We sometimes write just $X(T)$ for this Galois module and similarly $Y(T)$ for the co-character group $Y(T \times \overline{\mathbb{Q}}_p) \cong Y(T \times \overline{\mathbb{F}}_p)$. Let $W := (N(T)/T)(\overline{\mathbb{Q}}_p) \cong (N(T)/T)(\overline{\mathbb{F}}_p)$ denote the Weyl group. Let $\Delta = \Delta(B, T) \subset X(T)$, respectively $\Delta^\vee = \Delta^\vee(B, T) \subset Y(T)$, denote the simple roots (respectively coroots) defined by B. Then $\Gamma_{\mathbb{Q}_p}$ naturally acts on the based root datum $\Psi_0(G, B, T) := (X(T), \Delta, Y(T), \Delta^\vee)$. Let $L \subset \overline{\mathbb{Q}}_p$ denote the splitting

field of G, i.e. the finite unramified extension of $\mathbb{Q}_p$ cut out by the $\Gamma_{\mathbb{Q}_p}$-action on $\Psi_0(G, B, T)$.

A *dual group* of G is a quadruple $(\widehat{G}, \widehat{B}, \widehat{T}, \{x_\alpha\}_{\alpha \in \Delta(\widehat{B}, \widehat{T})})$, where $\widehat{G}$ is a split connected reductive group over $\mathbb{Z}_p$, $\widehat{B}$ a Borel of $\widehat{G}$, $\widehat{T} \subset \widehat{B}$ a Levi subgroup, $x_\alpha : \mathbb{G}_a \xrightarrow{\sim} \widehat{U}_\alpha$ isomorphisms of algebraic groups (where $\Delta(\widehat{B}, \widehat{T})$ is the set of simple roots determined by $\widehat{B}$, and $\widehat{U}_\alpha$ is the root subgroup of α), together with an isomorphism $\phi : \Psi_0(\widehat{G}, \widehat{B}, \widehat{T}) \xrightarrow{\sim} \Psi_0(G, B, T)^\vee$. This isomorphism induces an action of $\Gamma_{\mathbb{Q}_p}$ on $(\widehat{G}, \widehat{B}, \widehat{T}, \{x_\alpha\})$ that factors through $\mathrm{Gal}(L/\mathbb{Q}_p)$, and we define the L-group ${}^L G := \widehat{G} \rtimes \mathrm{Gal}(L/\mathbb{Q}_p)$, a reductive group over $\mathbb{Z}_p$. The Weyl group of $\widehat{T}$ is naturally identified with W via the duality isomorphism. We remark that any two pinnings $(\widehat{G}, \widehat{B}, \widehat{T}, \{x_\alpha\})$ of $\widehat{G}$ are $\widehat{G}(\mathbb{Z}_p)$-conjugate provided that $Z(\widehat{G})$ is connected. (This is equivalent to G^{der} being simply connected, which we will assume in a moment.) We also remark that our definition of the L-group is compatible with that of [Kot84] and [BG15] who work with canonical based root data. (The reason is that (B, T) and $(\widehat{B}, \widehat{T})$ are defined over $\mathbb{Z}_p$.)

From now on we suppose the following.

Hypothesis 9.1.1. Assume that the group G^{der} is simply connected, that $Z(G)$ is connected, and that G has a local twisting element η, which by definition means that $\eta \in X(T)^{\Gamma_{\mathbb{Q}_p}}$ and $\langle \eta, \alpha^\vee \rangle = 1$ for all $\alpha \in \Delta$. (Twisting elements are defined in the same way for groups over number fields in [BG15, §5.2]; they are a key part of the general conjectures made in [BG15] on the association of Galois representations to automorphic representations.)

In the following definitions, A is a topological $\mathbb{Z}_p$-algebra, i.e. a $\mathbb{Z}_p$-algebra that is also a topological ring.

Definition 9.1.2. An *L-parameter* is a continuous homomorphism $\Gamma_{\mathbb{Q}_p} \to {}^L G(A)$ that is compatible with the projections to $\mathrm{Gal}(L/\mathbb{Q}_p)$.

Definition 9.1.3. An *inertial L-parameter* is a continuous homomorphism $I_{\mathbb{Q}_p} \to \widehat{G}(A)$ that admits an extension to an L-parameter $\Gamma_{\mathbb{Q}_p} \to {}^L G(A)$.

We say that (inertial) L-parameters ρ_1, ρ_2 are *equivalent* if they are $\widehat{G}(A)$-conjugate, and we write $\rho_1 \cong \rho_2$.

Definition 9.1.4. A *Serre weight* is an isomorphism class of irreducible $\overline{\mathbb{F}}_p$-representations of $G(\mathbb{F}_p)$. (Just as for GL_n, we will sometimes abuse terminology and refer to an individual irreducible representation as a Serre weight.)

Given a tamely ramified inertial L-parameter $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ we will define below sets of Serre weights $\mathrm{W}^?(\tau)$ and $\mathrm{W}_{\mathrm{expl}}(\tau)$. These generalise respectively the construction in [Her09] and the construction in Section 7. (To be precise, in the latter case we will only generalise the case of GL_n over unramified extensions of $\mathbb{Q}_p$.) Our main result, Theorem 10.2.11, will establish that the two sets are equal for generic τ.

9.2. **Definition of** $\mathrm{W}^?(\tau)$. In this section we generalise [Her09, §§6.3–6.4]. To simplify notation, let $(\underline{G}, \underline{B}, \underline{T}) := (G, B, T) \times \overline{\mathbb{F}}_p$ and $(\underline{G}^*, \underline{B}^*, \underline{T}^*) := (\widehat{G}, \widehat{B}, \widehat{T}) \times \overline{\mathbb{F}}_p$. Let $F : \underline{G} \to \underline{G}$ denote the relative Frobenius, so $\underline{G}^F = G(\mathbb{F}_p)$. Let $F^* : \underline{G}^* \to \underline{G}^*$ denote the composite $\mathrm{Fr} \circ \varphi = \varphi \circ \mathrm{Fr}$, where Fr denotes the relative Frobenius on

$\underline{G}^*$ and $\varphi \in \Gamma_{\mathbb{Q}_p}$ denotes from now on a geometric Frobenius element. Then F^* is the relative Frobenius for a different $\mathbb{F}_p$-structure on $\underline{G}^*$, as φ has finite order on $\underline{G}^*$.

Recall that we fixed an isomorphism $\phi : \Psi_0(\underline{G}^*, \underline{B}^*, \underline{T}^*) \xrightarrow{\sim} \Psi_0(\underline{G}, \underline{B}, \underline{T})^\vee$ above that is by definition $\Gamma_{\mathbb{Q}_p}$-equivariant. Our conventions are that $\Gamma_{\mathbb{Q}_p}$ and the Weyl group act on the left on $X(\underline{T})$ and $Y(\underline{T})$; so $\gamma(\mu) = \mu \circ \gamma^{-1}$ and $w(\mu) = \mu \circ w^{-1}$ for $\mu \in X(\underline{T})$, $\gamma \in \Gamma_{\mathbb{Q}_p}$, $w \in W$. However, F is not invertible, so we set $F(\mu) = \mu \circ F$. Similar comments apply to $\underline{G}^*$. With these conventions we have $F \circ \phi = \phi \circ F^*$ on $Y(\underline{T}^*)$, as[3] $F = p\varphi$ on $X(\underline{T})$ and $F^* = p\varphi$ on $Y(\underline{T}^*)$. Thus ϕ is a duality between $(\underline{G}, F)$ and $(\underline{G}^*, F^*)$ in the sense of Deligne–Lusztig [DL76]. (Note once again that $\Psi_0(\underline{G}, \underline{B}, \underline{T})$, $\Psi_0(\underline{G}^*, \underline{B}^*, \underline{T}^*)$ are canonically isomorphic to the canonical based data, as $(\underline{B}, \underline{T})$ is F-stable and $(\underline{B}^*, \underline{T}^*)$ is F^*-stable.)

Fix from now on a generator $(\zeta_{p^i-1}) \in \varprojlim_{i \geq 1} \mathbb{F}_{p^i}^*$. Recall the following facts from [DL76, §5].

(i) The (canonical) Weyl group $W = N(\underline{T})/\underline{T}$ is canonically identified with $N(\underline{T}^*)/\underline{T}^*$ such that $w \circ \phi = \phi \circ w$ for all $w \in W$. The actions of F and F^* on W are inverse to each other.

(ii) There is a canonical bijection between $\underline{G}^F$-conjugacy classes of pairs $(\mathbb{T}, \theta)$ consisting of an F-stable maximal torus $\mathbb{T} \subset \underline{G}$ and a character $\theta : \mathbb{T}^F \to \overline{\mathbb{Q}}_p^\times$ and $\underline{G}^{*F^*}$-conjugacy classes of pairs $(\mathbb{T}^*, s)$ consisting of an F^*-stable maximal torus $\mathbb{T}^* \subset \underline{G}^*$ and a semisimple element $s \in \mathbb{T}^{*F^*}$.

(iii) If the classes of $(\mathbb{T}^*, s)$, $(\mathbb{T}, \theta)$ are in bijection in (ii), then they are both said to be *maximally split* if $\mathbb{T}^* \subset Z_{\underline{G}^*}(s)$ is a maximally split torus (i.e. contained in an F^*-stable Borel subgroup of $Z_{\underline{G}^*}(s)$).

Proposition 9.2.1. *We have the following commutative diagram*

$$
\begin{array}{ccc}
\left\{ \begin{array}{c} \text{maximally split} \\ (\mathbb{T}^*, s) \end{array} \right\}_{/\underline{G}^{*F^*}} & \xleftarrow{\text{duality}} & \left\{ \begin{array}{c} \text{maximally split} \\ (\mathbb{T}, \theta) \end{array} \right\}_{/\underline{G}^F} \\
\uparrow & & \uparrow \\
\left\{ \begin{array}{c} \text{tame inertial } L\text{-parameters} \\ I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p) \end{array} \right\}_{/\cong} & \xdashrightarrow{V_\phi} & \left\{ \begin{array}{c} \text{representations of} \\ G(\mathbb{F}_p) \text{ over } \overline{\mathbb{Q}}_p \end{array} \right\}_{/\cong}
\end{array}
$$

defining the map V_ϕ.

Proof. We proceed as in the proof of [Her09, Prop. 6.14]. Recall that (ζ_{p^i-1}) gives rise to a generator $g_{\mathrm{can}} \in I_{\mathbb{Q}_p}/I_{\mathbb{Q}_p}^w$, where $I_{\mathbb{Q}_p}^w$ is the wild ramification subgroup. Since $\Gamma_{\mathbb{Q}_p}/I_{\mathbb{Q}_p}^w = I_{\mathbb{Q}_p}/I_{\mathbb{Q}_p}^w \rtimes \overline{\langle \varphi \rangle}$, where $\varphi^{-1} g \varphi = g^p$ for $g \in I_{\mathbb{Q}_p}/I_{\mathbb{Q}_p}^w$, the bottom left-hand corner of the diagram is in bijection with $\widehat{G}(\overline{\mathbb{F}}_p)$-conjugacy classes of semisimple elements $s' \in \widehat{G}(\overline{\mathbb{F}}_p) = \underline{G}^*(\overline{\mathbb{F}}_p)$ satisfying that $\varphi^{-1}(s')$ is $\underline{G}^*(\overline{\mathbb{F}}_p)$-conjugate to $(s')^p$. Equivalently, by conjugating s' to $\underline{T}^*(\overline{\mathbb{F}}_p)$ and using that $F^* = p\varphi$ on $\underline{T}^*$, these are F^*-stable $\underline{G}^*(\overline{\mathbb{F}}_p)$-conjugacy classes of semisimple elements of $\underline{G}^*(\overline{\mathbb{F}}_p)$. Using that $Z(\underline{G})$ is connected, every such conjugacy class has a representative in $\underline{G}^{*F^*}$, unique up to $\underline{G}^{*F^*}$-conjugacy. We then get the bijection on the left as in [Her09]. The map on the right is given by $(\mathbb{T}, \theta) \mapsto \varepsilon_{\underline{G}} \varepsilon_{\mathbb{T}} R_{\mathbb{T}}^\theta$, with notation as in

---

[3]In order not to get confused about actions on $X(\underline{T})$, it helps to think in terms of the actions on $\underline{T}$. For example, $F = p\varphi^{-1}$ on $\underline{T}$.

[Her09, §4.1] and [Her09, Lem. 4.2]. It is a genuine representation of $\underline{G}^F = G(\mathbb{F}_p)$ by [DL76, Prop. 10.10]. □

The explicit description of V_ϕ in [Her09] generalises, as we now explain. Recall that for $w \in W$ we choose $g_w \in \underline{G}(\overline{\mathbb{F}}_p)$ such that $g_w^{-1} F(g_w) \in N(\underline{T})(\overline{\mathbb{F}}_p)$ represents w and define $\underline{T}_w := g_w \underline{T} g_w^{-1}$. Then $\underline{T}_w$ is an F-stable maximal torus. Define $\theta_{w,\mu} : \underline{T}_w^F \to \overline{\mathbb{Q}}_p^\times$ for $\mu \in X(\underline{T})$ by $\theta_{w,\mu}(t) := \tilde{\mu}(g_w^{-1} t g_w)$, where tilde denotes the Teichmüller lift. Any pair $(\mathbb{T}, \theta)$ consisting of an F-stable maximal torus $\mathbb{T}$ and a character $\theta : \mathbb{T}^F \to \overline{\mathbb{Q}}_p^\times$ is $\underline{G}^F$-conjugate to $(\underline{T}_w, \theta_{w,\mu})$ for some (w, μ).

Definition 9.2.2. Let $R(w, \mu) := \varepsilon_{\underline{G}} \varepsilon_{\underline{T}_w} R_{\underline{T}_w}^{\theta_{w,\mu}}$ be defined as in the proof of Proposition 9.2.1. (It may be virtual if (w, μ) is not maximally split.)

For $d \geq 1$ let $\omega_d : I_{\mathbb{Q}_p} \to \overline{\mathbb{F}}_p^\times$ be the character ω_σ, where $\sigma : \mathbb{F}_{p^d} \to \overline{\mathbb{F}}_p$ denotes the inclusion of the unique subfield of $\overline{\mathbb{F}}_p$ of degree d over $\mathbb{F}_p$. Let $\tau(w, \mu) : I_{\mathbb{Q}_p} \to \widehat{T}(\overline{\mathbb{F}}_p)$ denote the tame representation

$$\tau(w, \mu) := N_{(F^* \circ w^{-1})^d / F^* \circ w^{-1}}\big(\mu(\omega_d)\big),$$

where $d \geq 1$ is chosen such that $(F^* \circ w^{-1})^d = p^d$ on $Y(\underline{T}^*)$, μ is considered as element of $Y(\underline{T}^*)$ via ϕ, and $N_{A^d / A} = \prod_{i=0}^{d-1} A^i$.

Proposition 9.2.3. *The representation $\tau(w, \mu)$ is an inertial L-parameter. If $(\underline{T}_w, \theta_{w,\mu})$ is maximally split then it corresponds to $\tau(w, \mu)$ under the bijections of Proposition 9.2.1 and we have $V_\phi\big(\tau(w, \mu)\big) \cong R(w, \mu)$. In particular, V_ϕ is independent of the choice of $(\zeta_{p^i-1})_i$.*

Proof. This is the same as in [Her09, Prop. 6.14]. □

Let $X(\underline{T})_+$ denote the subset of $X(\underline{T})$ consisting of dominant weights, and let $X_1(\underline{T}) := \{\mu \in X(\underline{T}) : 0 \leq \langle \mu, \alpha^\vee \rangle \leq p - 1 \text{ for all } \alpha \in \Delta\}$, $X^0(\underline{T}) := \{\mu \in X(\underline{T}) : \langle \mu, \alpha^\vee \rangle = 0 \text{ for all } \alpha \in \Delta\}$. For $\mu \in X(\underline{T})_+$ let $F(\mu)$ denote the irreducible algebraic $\underline{G}$-representation of highest weight μ. (As the referee points out, this notation is ambiguous, since F is also the Frobenius. However, below $F(\nu)$ for $\nu \in X(\underline{T})$ will always mean the algebraic representation and never the weight $\nu \circ F$.)

Lemma 9.2.4. *The map*

$$\frac{X_1(\underline{T})}{(F-1)X^0(\underline{T})} \to \big\{\text{Serre weights of } G(\mathbb{F}_p) = \underline{G}^F\big\}_{/\cong}$$

$$\mu \mapsto F(\mu)|_{\underline{G}^F}$$

is a (well-defined) bijection.

Proof. We claim first that there exists a finite order automorphism π of $(\underline{G}, \underline{B}, \underline{T})$ that commutes with F and that induces the action φ^{-1} on $\Psi_0(\underline{G}, \underline{B}, \underline{T})$. To see this, note first that we can choose a pinning $x_\alpha : \mathbb{G}_a \xrightarrow{\sim} U_\alpha$ for $\alpha \in \Delta$ such that $F \circ x_{\varphi\alpha} = x_\alpha \circ F_a$ for all $\alpha \in \Delta$, where F_a is the relative Frobenius on $\mathbb{G}_a$ and U_α is the root subgroup of α. Then let π be the unique automorphism of $(\underline{G}, \underline{B}, \underline{T}, \{x_\alpha\}_\Delta)$ inducing φ^{-1} on $\Psi_0(\underline{G}, \underline{B}, \underline{T})$. To see that F and π commute, note that both maps send U_α onto $U_{\varphi^{-1}\alpha}$.

Then $F \circ \pi^{-1}$ is the relative Frobenius of a *split* $\mathbb{F}_p$-structure on $\underline{G}$, since $F \circ \pi^{-1} = F\varphi = p$ on $\underline{T}$. The lemma now follows from Proposition 1.3 in the appendix to [Her09], noting that $F = p\pi^{-1}$ on $X(\underline{T})$. □

Let $X_{\mathrm{reg}}(\underline{T}) := \{\mu \in X(\underline{T}) : 0 \le \langle \mu, \alpha^\vee \rangle < p - 1 \text{ for all } \alpha \in \Delta\} \subset X_1(\underline{T})$. Then $\mu \mapsto w_0 \cdot (\mu - p\eta)$ defines a self-bijection of $X_{\mathrm{reg}}(\underline{T})$ (where $w_0 \in W$ is the longest element), which passes to the quotient $\frac{X_{\mathrm{reg}}(\underline{T})}{(F-1)X^0(\underline{T})}$. Via Lemma 9.2.4 this quotient is identified with a subset $\mathrm{W}_{\mathrm{reg}}$ of all (isomorphism classes of) Serre weights of $G(\mathbb{F}_p)$. We write

$$\mathcal{R}: \quad \begin{aligned} \mathrm{W}_{\mathrm{reg}} &\to & \mathrm{W}_{\mathrm{reg}} \\ F(\mu) &\mapsto & F\big(w_0 \cdot (\mu - p\eta)\big) \end{aligned}$$

for the induced bijection.

Definition 9.2.5. For a tame inertial L-parameter $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ let

$$\mathrm{W}^?(\tau) := \big\{ \mathcal{R}(F) : \ F \in \mathrm{W}_{\mathrm{reg}} \text{ an irreducible constituent of } \overline{V_\phi(\tau)} \big\}.$$

Remark 9.2.6. Note that this set does not agree completely with the set $W^?(\tau)$ defined in [Her09, §6] when $G = \mathrm{GL}_n$. The only discrepancy occurs for weights $F(\lambda)$ with $\langle \lambda, \alpha^\vee \rangle = p - 2$ for some $\alpha \in \Delta$, which is irrelevant for our main result (Theorem 10.2.11).

9.3. Definition of $\mathrm{W}_{\mathrm{expl}}(\tau)$. We will now define $\mathrm{W}_{\mathrm{expl}}(\tau)$ for a tame inertial L-parameter $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$. When $G = \mathrm{Res}_{K/\mathbb{Q}_p} \mathrm{GL}_n$ with $K/\mathbb{Q}_p$ finite unramified we will recover the set of weights given in Definition 7.2.3 (this will be Proposition 9.5.3).

9.3.1. *Hodge–Tate co-characters.* Suppose for the moment that $K/\mathbb{Q}_p$ is finite, or more generally that $K/\mathbb{Q}_p$ is algebraic with finite ramification index, and that H is a (not necessarily connected) algebraic group over $\overline{\mathbb{Q}}_p$. Recall that a continuous homomorphism $\rho : \Gamma_K \to H(\overline{\mathbb{Q}}_p)$ is said to be Hodge–Tate (resp. crystalline) if for some faithful (and hence any) representation $H \to \mathrm{GL}_N$ over $\overline{\mathbb{Q}}_p$, the resulting N-dimensional Galois representation is Hodge–Tate (resp. crystalline). Given any $\rho : \Gamma_K \to H(\overline{\mathbb{Q}}_p)$ that is Hodge–Tate, and any homomorphism $j : \overline{K} \to \overline{\mathbb{Q}}_p$, it is explained in [BG15, §2.4] that there is an $H^\circ(\overline{\mathbb{Q}}_p)$-conjugacy class $\mathrm{HT}_j(\rho)$ of co-characters $\mathbb{G}_m \to H$ over $\overline{\mathbb{Q}}_p$ (or equivalently, an element of $Y(\widehat{T})/W$, where $\widehat{T}$ is a fixed maximal torus and W its Weyl group in H°). These classes satisfy the relation

(9.3.1) $$\mathrm{HT}_{j \circ \gamma^{-1}}(\rho) = \rho(\gamma) \, \mathrm{HT}_j(\rho) \rho(\gamma)^{-1} \qquad \text{for all } \gamma \in \Gamma_K.$$

Our normalisation is such that $\mathrm{HT}_j(\varepsilon) = 1$ for all j. If $\mu : \mathbb{G}_m \to H$ is a co-character we let $[\mu]$ denote its class in $Y(\widehat{T})/W$. The following lemma is elementary.

Lemma 9.3.2. *Suppose $\rho : \Gamma_K \to H(\overline{\mathbb{Q}}_p)$ is Hodge–Tate.*

(i) *If $f : H \to H'$ is a map of algebraic groups over $\overline{\mathbb{Q}}_p$, then for $j : \overline{K} \to \overline{\mathbb{Q}}_p$,*

$$f \circ \mathrm{HT}_j(\rho) = \mathrm{HT}_j(f \circ \rho).$$

(ii) *If $K' \subset \overline{K}$ contains K, then for $j : \overline{K} \to \overline{\mathbb{Q}}_p$,*

$$\mathrm{HT}_j(\rho|_{\Gamma_{K'}}) = \mathrm{HT}_j(\rho).$$

(iii) *Suppose that L is another field and that $\gamma : \overline{K} \xrightarrow{\sim} \overline{L}$ is an isomorphism sending K onto L. Then for $j : \overline{L} \to \overline{\mathbb{Q}}_p$,*

$$\mathrm{HT}_{j \circ \gamma}(\rho) = \mathrm{HT}_j \big(\rho \circ (\gamma^{-1}(-)\gamma) \big).$$

Now specialize to an L-parameter $\rho : \Gamma_{\mathbb{Q}_p} \to {}^L G(\overline{\mathbb{Q}}_p)$ that is crystalline. Then $\mathrm{HT}_1(\rho) \in Y(\widehat{T})/W = X(T)/W$ associated to $\mathrm{id} : \overline{\mathbb{Q}}_p \to \overline{\mathbb{Q}}_p$ determines all Hodge–Tate co-characters by (9.3.1). Also, $\mathrm{HT}_j(\rho)$ depends only on $j|_L$.

Suppose that $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ is a tame inertial L-parameter.

Definition 9.3.3. We say that an L-parameter $\rho : \Gamma_{\mathbb{Q}_p} \to {}^L G(\overline{\mathbb{Z}}_p)$ is an *obvious crystalline lift of* τ if

    (i) $\overline{\rho|_{I_{\mathbb{Q}_p}}}$ is $\widehat{G}(\overline{\mathbb{F}}_p)$-conjugate to τ;
    (ii) ρ is crystalline;
    (iii) there is a maximal torus $T^* \subset {}^L G_{/\overline{\mathbb{Z}}_p}$ such that

$$\rho(\Gamma_{\mathbb{Q}_p}) \subset N_{{}^L G_{/\overline{\mathbb{Z}}_p}}(T^*)(\overline{\mathbb{Z}}_p) \text{ and } \rho(I_{\mathbb{Q}_p}) \subset T^*(\overline{\mathbb{Z}}_p).$$

Remark 9.3.4. As any two maximal tori of ${}^L G_{/\overline{\mathbb{Z}}_p}$ are $\widehat{G}(\overline{\mathbb{Z}}_p)$-conjugate, we may assume without loss of generality that $T^* = \widehat{T}_{/\overline{\mathbb{Z}}_p}$ in (iii). We also remark that $N_{{}^L G}(\widehat{T}) = N_{\widehat{G}}(\widehat{T}) \rtimes \mathrm{Gal}(L/\mathbb{Q}_p)$.

Remark 9.3.5. Note that (iii) implies that there exists $K \subset \overline{\mathbb{Q}}_p$ with $K/\mathbb{Q}_p$ finite unramified such that $\rho(\Gamma_K) \subset T^*(\overline{\mathbb{Z}}_p)$.

Definition 9.3.6. We define

$$\mathrm{W}_{\mathrm{obv}}(\tau) := \big\{ F(\mu) : \mu \in X_1(T), \ \tau \text{ admits an obvious crystalline lift } \rho$$
$$\text{with } \mathrm{HT}_1(\rho) = [\mu + \eta] \in X(T)/W \big\}.$$

Proposition 9.3.7. *There holds the equality* $\mathrm{W}_{\mathrm{obv}}(\tau) = \{ F(\mu) : \mu \in X_1(T), \ \tau \cong \tau(w, \mu + \eta) \text{ for some } w \in W \}.$

Proof. More generally we will show that

$$\{\mathrm{HT}_1(\rho) : \rho \text{ an obvious crystalline lift of } \tau\}$$
$$= \{[\mu] : \mu \in X(T) \text{ such that } \tau \cong \tau(w, \mu) \text{ some } w \in W\}.$$

(Note that $\tau(\sigma w F(\sigma)^{-1}, \sigma\mu) \cong \tau(w, \mu)$ for $\sigma \in W$.)

First suppose that $\rho : \Gamma_{\mathbb{Q}_p} \to N_{{}^L G}(\widehat{T})(\overline{\mathbb{Z}}_p)$ is an obvious crystalline lift of τ. Fix $K \subset \overline{\mathbb{Q}}_p$ with $K/\mathbb{Q}_p$ finite unramified such that $\rho(\Gamma_K) \subset \widehat{T}(\overline{\mathbb{Z}}_p)$. Then for all $j : \overline{\mathbb{Q}}_p \to \overline{\mathbb{Q}}_p$, the co-character $\mu_j := \mathrm{HT}_j(\rho|_{\Gamma_K}) \in Y(\widehat{T})$ is a lift of $\mathrm{HT}_j(\rho) \in Y(\widehat{T})/W$ (by Lemma 9.3.2). Note that μ_j depends only on $j|_K$. Also, $\mathrm{conj}(\rho(\varphi)) \circ \rho|_{\Gamma_K} = \rho|_{\Gamma_K} \circ (\varphi(-)\varphi^{-1})$ (where $\mathrm{conj}(g)$ denotes conjugation by g), so by Lemma 9.3.2, $\mathrm{conj}(\rho(\varphi)) \circ \mu_j = \mu_{j \circ \varphi^{-1}}$. Writing $\rho(\varphi) = \varphi \dot{w}^{-1} \in N_{{}^L G}(\widehat{T})(\overline{\mathbb{Z}}_p)$, for some $\dot{w} \in N_{\widehat{G}}(\widehat{T})(\overline{\mathbb{Z}}_p)$ lifting $w \in W$, we get $\mu_{j \circ \varphi^{-1}} = \varphi w^{-1} \mu_j$. For $\nu \in X(\widehat{T})$, $\nu \circ \rho|_{\Gamma_K} : \Gamma_K \to \overline{\mathbb{Z}}_p^\times$ is crystalline with Hodge–Tate co-characters $\mathrm{HT}_j(\nu \circ \rho|_{\Gamma_K}) = \langle \nu, \mu_j \rangle \in \mathbb{Z}$, hence by Lemma 5.1.6,

$$\overline{\nu \circ \rho|_{I_{\mathbb{Q}_p}}} = \prod_{j \in \Gamma_{\mathbb{Q}_p}/\Gamma_K} \bar{j}(\omega_d)^{\langle \nu, \mu_j \rangle},$$

where $d := [K : \mathbb{Q}_p]$. It follows that

$$(9.3.8) \qquad \overline{\rho|_{I_{\mathbb{Q}_p}}} = \prod_{s=0}^{d-1} \mu_{\varphi^{-s}}(\omega_d^{p^s}) = \omega_d^{\sum p^s (\varphi w^{-1})^s \mu_1} = \tau(w, \mu_1),$$

as $F^* = p\varphi$ on $Y(\widehat{T}) = X(T)$.

Conversely, given $(w, \mu) \in W \times X(T)$, choose $d \geq 1$ such that $(\varphi w^{-1})^d$ acts trivially on μ. Let $K \subset \overline{\mathbb{Q}}_p$ be unramified over $\mathbb{Q}_p$ of degree d. Let $\rho' : \Gamma_K \to \widehat{T}(\overline{\mathbb{Z}}_p)$ be crystalline such that $\mathrm{HT}_{\varphi^{-s}}(\rho') = (\varphi w^{-1})^s \mu \in Y(\widehat{T})$ for all $s \in \mathbb{Z}$. (Note that $\mathrm{HT}_j(\rho')$ only depends on $j|_K$. To construct ρ' write $\widehat{T} \cong \mathbb{G}_m^r$ and use Lemma 5.1.6.) It follows by Lemma 5.1.6 that $\varphi w^{-1} \circ \rho' = \rho' \circ (\varphi(-)\varphi^{-1})$ holds on $I_{\mathbb{Q}_p}$. Therefore we may define an L-parameter $\rho : \Gamma_{\mathbb{Q}_p} \to {}^L G(\overline{\mathbb{Z}}_p)$ by (i) $\rho|_{I_{\mathbb{Q}_p}} = \rho'|_{I_{\mathbb{Q}_p}}$ and (ii) $\rho(\varphi) = \varphi \dot{w}^{-1}$ where $\dot{w} \in N_{\widehat{G}}(\widehat{T})(\overline{\mathbb{Z}}_p)$ is any fixed lift of w. Then ρ is crystalline, as $\rho|_{I_{\mathbb{Q}_p}}$ is crystalline, and $\mathrm{HT}_1(\rho) = [\mu]$ by Lemma 9.3.2. Also, $\overline{\rho|_{I_{\mathbb{Q}_p}}} = \tau(w, \mu)$ by (9.3.8). □

For $\nu \in X(\underline{T})_+$ let $W(\nu)$ denote the $\underline{G}$-module $\mathrm{Ind}_{\underline{B}}^{\underline{G}}(w_0\nu)$ defined in [Jan03, II.2]. It has unique highest weight ν and $\underline{G}$-socle $F(\nu)$.

Definition 9.3.9. If W is a set of Serre weights of $G(\mathbb{F}_p)$, we define $\mathcal{C}(\mathrm{W})$ to be the smallest set of Serre weights with the properties:

     ∘ $\mathrm{W} \subset \mathcal{C}(\mathrm{W})$, and
     ∘ if $\mathrm{W} \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\nu) \neq \varnothing$, where $\nu \in X_1(T)$, then $F(\nu) \in \mathcal{C}(\mathrm{W})$.

9.3.2. *Levi predictions.* The Levi subgroups $M \subset G$ that contain T are in bijection with the $\Gamma_{\mathbb{Q}_p}$-stable subsets of Δ, by sending M to Δ_M. Note that each M satisfies Hypothesis 9.1.1 with the same twisting element η as G. For each M fix a dual group $(\widehat{M}, \widehat{B}_M, \widehat{T}_M, \{x_{\alpha,M}\}_{\alpha \in \widehat{\Delta}_M})$. Then there is a unique $\Gamma_{\mathbb{Q}_p}$-equivariant homomorphism $i : (\widehat{M}, \widehat{B}_M, \widehat{T}_M) \to (\widehat{G}, \widehat{B}, \widehat{T})$ such that $i^* : X(\widehat{T}) \to X(\widehat{T}_M)$ corresponds to $\mathrm{id}_{Y(T)}$ and $i \circ x_{\alpha,M} = x_\alpha$ for all $\alpha \in \widehat{\Delta}_M$. In fact, i is a closed immersion, so we can and will think of $\widehat{M}$ as the Levi subgroup of $\widehat{G}$ containing $\widehat{T}$ defined by $\Delta_M^\vee$, with induced structures.

We have variants of the definitions of Sections 9.1–9.3 with M replacing G, and we will indicate these by decorating notation with an M: in particular $X_1^M(\underline{T})$, $W^M(\mu)$ for $\mu \in X(\underline{T})_{+,M}$, $[\cdot]_M$, and in Section 10 also Φ_M, Φ_M^+, $\|\cdot\|_M$, $\uparrow_M$, $\tau^M(w, \mu)$ for $w \in W_M, \mu \in X(\underline{T})$.

Definition 9.3.10. We recursively define $\mathrm{W}_{\mathrm{expl}}(\tau)$ to be the smallest set containing $\mathrm{W}_{\mathrm{obv}}(\tau)$ that is closed under the following operation: whenever $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ factors (perhaps after conjugation) through an inertial L-parameter $\tau^M : I_{\mathbb{Q}_p} \to \widehat{M}(\overline{\mathbb{F}}_p)$ with M as above, and $\mathrm{W}_{\mathrm{expl}}(\tau^M) \cap \mathrm{JH}_{M(\mathbb{F}_p)} W^M(w \cdot \nu) \neq \varnothing$ (where $\nu \in X_1(\underline{T})$, $w \in W$ such that $w \cdot \nu \in X(\underline{T})_{+,M}$), then $F(\nu) \in \mathrm{W}_{\mathrm{expl}}(\tau)$.

Remark 9.3.11. If we take $M = G$ in the recursive step of this definition, we obtain that $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau)) \subset \mathrm{W}_{\mathrm{expl}}(\tau)$.

The motivation for Definition 9.3.10 is as in Section 7.2: we expect that τ^M has a crystalline lift of Hodge–Tate co-character $[w \cdot \nu + \eta]_M$, hence that τ has a crystalline lift of Hodge–Tate co-character $[\nu + \eta]$.

Lemma 9.3.12. *If $G = T$ is a torus and τ a tame inertial L-parameter, then $\mathrm{W}_{\mathrm{expl}}(\tau) = \mathrm{W}_{\mathrm{obv}}(\tau) = \{F(\mu - \eta)\}$ for any $\mu \in X(\underline{T})$ such that $\tau \cong \tau(1, \mu)$.*

Proof. The equality $\mathrm{W}_{\mathrm{expl}}(\tau) = \mathrm{W}_{\mathrm{obv}}(\tau)$ is clear by Remark 9.3.11. By Proposition 9.2.3, there exists $\mu \in X(\underline{T})$ such that $\tau \cong \tau(1, \mu)$ (as $W = 1$), so $F(\mu - \eta) \in$

$W_{\mathrm{obv}}(\tau)$ by Proposition 9.3.7. By the same result, if we take any weight $F(\mu' - \eta) \in W_{\mathrm{obv}}(\tau)$, then $\tau \cong \tau(1, \mu')$. It follows that $(1 + F^* + \cdots + (F^*)^{d-1})(\mu - \mu') \equiv 0 \pmod{(p^d - 1)Y(\underline{T}^*)}$, where d is chosen as in §9.2. Equivalently, $(1 + F + \cdots + F^{d-1})(\mu - \mu') \equiv 0 \pmod{(p^d - 1)X(\underline{T})}$. As $F^d - 1 = p^d - 1$ is injective on $X(\underline{T})$ we see that $\mu \equiv \mu' \pmod{(F - 1)X(\underline{T})}$, i.e. $F(\mu - \eta) \cong F(\mu' - \eta)$. □

Remark 9.3.13. For general G it follows from Lemma 9.3.12 that if $M = T$ in the recursive step of Definition 9.3.10, then the non-emptiness of the intersection implies that $F(\nu) \in W_{\mathrm{obv}}(\tau)$, so no new weights are obtained. (To see this, note that the non-emptiness implies that $\tau \cong \tau(1, w \cdot \nu + \eta)$ for some $w \in W$, by applying Proposition 9.3.7 to τ^T. Hence $\tau \cong \tau(w^{-1}F(w), \nu + \eta)$, which implies the claim.)

9.4. **Restriction of scalars.** Suppose for the rest of this section that $K \subset \overline{\mathbb{Q}}_p$ with $K/\mathbb{Q}_p$ finite unramified. In the following, if X is a set (resp. group, resp. group scheme) with a smooth left action of Γ_K, then we denote by $\mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} X$ the induced set (resp. group, resp. group scheme) consisting of functions $\Gamma_{\mathbb{Q}_p} \to X$ that are Γ_K-equivariant. For $\gamma \in \Gamma_{\mathbb{Q}_p}$ let $\mathrm{ev}_\gamma : \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} X \to X$ denote the evaluation map at γ. If Y is a set of representatives of $\Gamma_K \backslash \Gamma_{\mathbb{Q}_p}$, the $(\mathrm{ev}_y)_{y \in Y}$ provide a non-canonical isomorphism $\mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} X \xrightarrow{\sim} X^{[K:\mathbb{Q}_p]}$ of sets (resp. groups, resp. group schemes).

Suppose that H is a connected reductive group over $\mathcal{O}_K$ with Borel B_H, Levi $T_H \subset B_H$ and simple roots $\Delta_H \subset X(T_H)$. Suppose that H satisfies Hypothesis 9.1.1 (or rather its analogue over K) with local twisting element $\eta_H \in X(T_H)^{\Gamma_K}$. We may then obtain a group G as in Section 9.1 by restriction of scalars:

$$(G, B, T) := \mathrm{Res}_{\mathcal{O}_K/\mathbb{Z}_p}(H, B_H, T_H).$$

Note that $G \times \overline{\mathbb{Q}}_p \cong \prod_{\kappa: K \to \overline{\mathbb{Q}}_p} H \times_{K,\kappa} \overline{\mathbb{Q}}_p$, so G^{der} is simply connected and $Z(G)$ is connected. In particular,

$$X(T) \cong \bigoplus_{\kappa: K \to \overline{\mathbb{Q}}_p} X(T_H \times_\kappa \overline{\mathbb{Q}}_p) \cong \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} X(T_H).$$

It follows that $\Psi_0(G, B, T) \cong \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} \Psi_0(H, B_H, T_H)$ (where strictly speaking Δ consists of those functions in $\mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} \Delta_H$ that are supported on a single coset of Γ_K, and similarly for $\Delta^\vee$) and that $\eta \in X(T)^{\Gamma_{\mathbb{Q}_p}} \cong X(T_H)^{\Gamma_K}$ defined by η_H is a local twisting element of G. Hence G satisfies Hypothesis 9.1.1. Let $L \subset \overline{\mathbb{Q}}_p$ denote the splitting field of H, so $L/\mathbb{Q}_p$ is finite unramified, $L \supset K$, and $G \times L$ is also split.

Let $(\widehat{H}, \widehat{B}_H, \widehat{T}_H, \{x_{\alpha'}\})$ be a dual group of H as in Section 9.1, and define

$$(\widehat{G}, \widehat{B}, \widehat{T}) := \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}}(\widehat{H}, \widehat{B}_H, \widehat{T}_H).$$

Then $\Gamma_{\mathbb{Q}_p}$ preserves a pinning $\{x_\alpha\}$ of $\widehat{G}$ that is naturally induced from $\{x_{\alpha'}\}$. We also see that $\Psi_0(\widehat{G}, \widehat{B}, \widehat{T}) \cong \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} \Psi_0(\widehat{H}, \widehat{B}_H, \widehat{T}_H)$ (with the same proviso as above), so via the induced isomorphism $\Psi_0(\widehat{G}, \widehat{B}, \widehat{T}) \xrightarrow{\sim} \Psi_0(G, B, T)^\vee$ we can consider $(\widehat{G}, \widehat{B}, \widehat{T}, \{x_\alpha\})$ as a dual group of G. We let $^L H := \widehat{H} \rtimes \mathrm{Gal}(L/K)$.

In the following, note that the notions of (inertial) L-parameter and obvious crystalline lift carry over to representations of Γ_K (repectively I_K). Also note that $\mathrm{ev}_1 : \widehat{G}(A) \to \widehat{H}(A)$ is Γ_K-equivariant hence extends to a homomorphism $\mathrm{ev}_1 : \widehat{G}(A) \rtimes \mathrm{Gal}(L/K) \to {}^L H(A)$.

Lemma 9.4.1. *Suppose that A is a topological $\mathbb{Z}_p$-algebra. Then we have a bijection*

$$(9.4.2) \qquad \left\{ \begin{array}{c} L\text{-parameters} \\ \Gamma_{\mathbb{Q}_p} \xrightarrow{\rho} {}^L G(A) \end{array} \right\}_{/\widehat{G}(A)} \xrightarrow{\sim} \left\{ \begin{array}{c} L\text{-parameters} \\ \Gamma_K \xrightarrow{\rho_K} {}^L H(A) \end{array} \right\}_{/\widehat{H}(A)}.$$

sending ρ to $\rho_K = \mathrm{ev}_1(\rho|_{\Gamma_K})$. If $A = \overline{\mathbb{Q}}_p$ then ρ is crystalline if and only if ρ_K is crystalline.

Proof. For the moment let us consider A with the discrete topology. By writing $\rho(g) = \rho^0(g) \rtimes g$ we see that ρ^0 defines a 1-cocycle $\Gamma_{\mathbb{Q}_p} \to \widehat{G}(A)$, and in this way we get a bijection between $\widehat{G}(A)$-conjugacy classes of L-parameters ρ and the pointed set $H^1(\Gamma_{\mathbb{Q}_p}, \widehat{G}(A))$. As $\widehat{G}(A) \cong \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} \widehat{H}(A)$, the non-abelian Shapiro lemma ([Sti10, Prop. 8]) shows that $H^1(\Gamma_{\mathbb{Q}_p}, \widehat{G}(A)) \cong H^1(\Gamma_K, \widehat{H}(A))$ where ρ^0 is sent to $\mathrm{ev}_1(\rho^0|_{\Gamma_K})$. This proves (9.4.2) if A is discrete. From the description $\rho_K = \mathrm{ev}_1(\rho|_{\Gamma_K})$ and (9.4.4) it follows in general that ρ is continuous iff $\rho|_{\Gamma_L}$ is continuous iff ρ_K is continuous, and similarly for the crystalline condition when $A = \overline{\mathbb{Q}}_p$. For later reference we recall from [Sti10] a description of a representative ρ in the inverse image of ρ_K. Let Y be a set of representatives of $\Gamma_K \backslash \Gamma_{\mathbb{Q}_p}$ with $1 \in Y$. Then ρ is defined by

$$(9.4.3) \qquad \mathrm{ev}_\gamma \rho^0(\gamma') = \rho_K^0(\delta)^{-1} \cdot \rho_K^0(\delta') \in \widehat{H}(A),$$

where $\gamma = \delta y$, $\gamma\gamma' = \delta'y'$ with $\delta, \delta' \in \Gamma_K$, $y, y' \in Y$. $\qquad\square$

Note that in the context of Lemma 9.4.1 for any $\gamma \in \Gamma_{\mathbb{Q}_p}$ we have $\gamma \circ \rho|_{\Gamma_L} \cong \rho|_{\Gamma_L} \circ (\gamma(-)\gamma^{-1})$, so

$$(9.4.4) \qquad \mathrm{ev}_\gamma(\rho|_{\Gamma_L}) \cong \rho_K|_{\Gamma_L} \circ (\gamma(-)\gamma^{-1}) \; : \; \Gamma_L \to \widehat{H}(A).$$

Lemma 9.4.5. *Suppose that A is a topological $\mathbb{Z}_p$-algebra. Then we have a bijection*

$$\left\{ \begin{array}{c} \text{inertial } L\text{-parameters} \\ I_{\mathbb{Q}_p} \xrightarrow{\tau} \widehat{G}(A) \end{array} \right\}_{/\widehat{G}(A)} \xrightarrow{\sim} \left\{ \begin{array}{c} \text{inertial } L\text{-parameters} \\ I_K \xrightarrow{\tau_K} \widehat{H}(A) \end{array} \right\}_{/\widehat{H}(A)}$$

sending τ to $\tau_K = \mathrm{ev}_1(\tau)$.

Remark 9.4.6. Note that $I_K = I_{\mathbb{Q}_p}$.

Proof. By Lemma 9.4.1 the map is well defined and surjective. Suppose that τ_1, τ_2 are inertial L-parameters such that $\tau_{1,K} \cong \tau_{2,K}$. As the τ_i extend to L-parameters, from (9.4.4) we get that $\mathrm{ev}_\gamma(\tau_1) \cong \mathrm{ev}_\gamma(\tau_2)$ for any $\gamma \in \Gamma_{\mathbb{Q}_p}$. Let Y be a set of representatives of $\Gamma_K \backslash \Gamma_{\mathbb{Q}_p}$, and choose $h_y \in \widehat{H}(A)$ ($y \in Y$) such that

$$\mathrm{ev}_y(\tau_1) = h_y \cdot \mathrm{ev}_y(\tau_2) \cdot h_y^{-1}, \qquad \text{for all } y \in Y.$$

Then $\tau_1 = g \cdot \tau_2 \cdot g^{-1}$, where $g \in \widehat{G}(A)$ is defined by $g(y) = h_y$ for $y \in Y$. $\qquad\square$

Lemma 9.4.7. *Suppose that $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ is a tame inertial L-parameter and that $\rho : \Gamma_{\mathbb{Q}_p} \to {}^L G(\overline{\mathbb{Z}}_p)$ is an L-parameter. Then ρ is an obvious crystalline lift of τ iff ρ_K is an obvious crystalline lift of τ_K.*

Proof. The "only if" implication is immediate from Lemma 9.4.1. Conversely, if ρ_K is an obvious crystalline lift of τ_K, by Lemma 9.4.5, $\overline{\rho|_{I_{\mathbb{Q}_p}}} \cong \tau$, as it is true after evaluating at 1. Also, ρ is crystalline by Lemma 9.4.1. Now assume without loss of generality that ρ_K takes values in $N_{{}^L H}(\widehat{T}_H)(\overline{\mathbb{Z}}_p)$. Then ρ (or rather a representative

of ρ in its conjugacy class) is obtained from ρ_K by formula (9.4.3), which shows that ρ takes values in $N_{^LG}(\widehat{T})(\overline{\mathbb{Z}}_p)$. Now if $\gamma' \in I_{\mathbb{Q}_p}$, then $\rho(\gamma') \in \widehat{T}(\overline{\mathbb{Z}}_p)$ follows from (9.4.3), as $y = y'$ and $\delta' \in \delta I_K$. $\qquad\square$

9.5. The case of GL_n. Suppose that $H = \mathrm{GL}_n$, B_H the upper-triangular Borel subgroup and T_H the diagonal torus, all over $\mathcal{O}_K$. Define $(\widehat{H}, \widehat{B}_H, \widehat{T}_H)$ likewise but over $\mathbb{Z}_p$. Identify $X(T_H) \xrightarrow{\sim} Y(\widehat{T}_H)$ by sending $\mathrm{diag}(x_1, \ldots, x_n) \mapsto \prod x_i^{a_i}$ to $x \mapsto \mathrm{diag}(x^{a_1}, \ldots, x^{a_n})$. Note that $L = K$ and $^LH = \widehat{H}$. Let $\eta_H \in X(T_H)$ be the local twisting element $\mathrm{diag}(x_1, \ldots, x_n) \mapsto \prod x_i^{n-i}$.

Lemma 9.5.1. *The representation $\rho_K : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ is an obvious crystalline lift of $\tau_K : I_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ in the sense of the current section if and only if it is an obvious crystalline lift in the sense of Definition 7.1.3.*

Proof. It suffices to show that $\rho_K : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ satisfies condition (iii) in Definition 9.3.3 iff it is isomorphic to $\bigoplus \mathrm{Ind}_{\Gamma_{K_i}}^{\Gamma_K} \overline{\mathbb{Z}}_p(\chi_i)$ for some $K_i \subset \overline{\mathbb{Q}}_p$ with K_i/K finite unramified and characters $\chi_i : \Gamma_{K_i} \to \overline{\mathbb{Z}}_p^\times$. This is clear, as either condition is equivalent to $\overline{\mathbb{Z}}_p^n$ being a direct sum of n rank 1 free $\overline{\mathbb{Z}}_p$-submodules that are permuted by Γ_K, with I_K preserving each summand. $\qquad\square$

Combining the previous lemma with Lemma 9.4.7, we obtain the following.

Corollary 9.5.2. *$\rho : \Gamma_{\mathbb{Q}_p} \to {}^LG(\overline{\mathbb{Z}}_p)$ is an obvious crystalline lift of $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ if and only if $\rho_K : \Gamma_K \to \mathrm{GL}_n(\overline{\mathbb{Z}}_p)$ is an obvious crystalline lift of $\tau_K : I_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ in the sense of Definition 7.1.3.*

We can now show that when $G = \mathrm{Res}_{K/\mathbb{Q}_p} \mathrm{GL}_n$, the set $\mathrm{W}_{\mathrm{expl}}(\tau)$ recovers the collection of Serre weights given by Definition 7.2.3. More precisely, we have the following.

Proposition 9.5.3. *For $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ a tame inertial L-parameter, $\mathrm{W}_{\mathrm{expl}}(\tau) = \mathrm{W}_{\mathrm{expl}}(\tau_K)$, where the latter set is computed according to Definition 7.2.3.*

Remark 9.5.4. Note that we have canonical isomorphisms $G(\mathbb{F}_p) \cong H(k) = \mathrm{GL}_n(k)$, where k is the residue field of K.

Proof. Let $r : \mathcal{O}_K \to k$ be the reduction map and $\sigma_0 : k \to \overline{\mathbb{F}}_p$ the inclusion. We have canonical isomorphisms

$$(9.5.5) \qquad G \times \overline{\mathbb{F}}_p \cong \prod_{\sigma:k\to\overline{\mathbb{F}}_p} H \times_{\sigma r} \overline{\mathbb{F}}_p,$$

$$(9.5.6) \qquad X(T \times \overline{\mathbb{F}}_p) \cong \bigoplus_{\sigma:k\to\overline{\mathbb{F}}_p} X(T_H \times_{\sigma r} \overline{\mathbb{F}}_p) \cong \mathrm{Ind}_{\Gamma_K}^{\Gamma_{\mathbb{Q}_p}} X(T_H \times_{\sigma_0 r} \overline{\mathbb{F}}_p).$$

For $\mu = (\mu_\sigma) \in X(T \times \overline{\mathbb{F}}_p)$ we can write $\mu_\sigma = \nu_\sigma \times_\sigma \overline{\mathbb{F}}_p$ with $\nu_\sigma \in \mathrm{Hom}_k(T_H \times_r k, \mathbb{G}_m)$, as T_H is split. Then in (9.5.6), for $\gamma \in \Gamma_{\mathbb{Q}_p}$,

$$\mathrm{ev}_{\gamma^{-1}}(\mu) = \mathrm{ev}_1(\gamma^{-1}\mu) = (\gamma^{-1}\mu)_{\sigma_0} = \gamma^{-1}\mu_{\overline{\gamma}\circ\sigma_0} = \nu_{\overline{\gamma}\circ\sigma_0} \times \overline{\mathbb{F}}_p.$$

In particular, as $\eta \in X(T)^{\Gamma_{\mathbb{Q}_p}}$, $\mathrm{ev}_{\gamma^{-1}}(\eta) = \eta_H = \eta_{H,0} \times \overline{\mathbb{F}}_p$, where $\eta_{H,0} = (n-1, n-2, \ldots, 0) \in \mathrm{Hom}_k(T_H \times_r k, \mathbb{G}_m)$.

Suppose that ρ is an obvious crystalline lift of τ with $\mathrm{HT}_1(\rho) = [\mu + \eta]$, where $\mu \in X_1(T)$. Then from Lemma 9.3.2 and (9.4.4) we get for any $\gamma \in \Gamma_{\mathbb{Q}_p}$,

$$(9.5.7) \qquad \begin{aligned} \mathrm{HT}_\gamma(\rho_K) &= \mathrm{HT}_1\left(\rho_K \circ (\gamma^{-1}(-)\gamma)\right) = \mathrm{ev}_{\gamma^{-1}}\left(\mathrm{HT}_1(\rho)\right) \\ &= [\mathrm{ev}_{\gamma^{-1}}(\mu + \eta)] = \left[(\nu_{\overline{\gamma} \circ \sigma_0} + \eta_{H,0}) \times \overline{\mathbb{F}}_p\right]. \end{aligned}$$

With respect to the decomposition (9.5.5), the Serre weight $F(\mu) \in \mathrm{W}_{\mathrm{obv}}(\tau)$ is isomorphic to $\bigotimes_\sigma F(\mu_\sigma)$ as representations of $G(\mathbb{F}_p) \cong H(k)$, so $F(\mu) \cong \bigotimes F_k(\nu_\sigma) \otimes_\sigma \overline{\mathbb{F}}_p$, where $F_k(\nu_\sigma)$ denotes the irreducible algebraic $H_{/k}$-representation with highest weight ν_σ. Then $F_k(\nu_\sigma) \cong N_{\nu_\sigma}$ (cf. Section 3.1) as representations of $G(\mathbb{F}_p) \cong H(k)$. By (9.5.7), $F(\mu)$ is the Serre weight associated to the obvious crystalline lift ρ_K by Definition 7.1.3. Hence $\mathrm{W}_{\mathrm{obv}}(\tau) = \mathrm{W}_{\mathrm{obv}}(\tau_K)$.

Similarly, with the above notation we get an isomorphism $W(\mu) \cong \bigotimes W_k(\nu_\sigma) \otimes_\sigma \overline{\mathbb{F}}_p$, where $W_k(\nu_\sigma)$ denotes the $H_{/k}$-module $\mathrm{Ind}_{B_{H_{/k}}}^{H_{/k}}(w_0 \nu_\sigma)$, so $W_k(\nu_\sigma) \cong M'_{\nu_\sigma} \otimes_{\mathcal{O}_K} k$ (cf. Section 1.9). We deduce that $\mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau)) = \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau_K))$, where the latter is computed according to Definitions 7.1.3 and 7.2.1.

To compare explicit predicted weights, note first that any Levi M of G containing T is of the form $M = \mathrm{Res}_{\mathcal{O}_K/\mathbb{Z}_p} M_H$ with M_H a Levi of H containing T_H, so that $M_H \cong \prod_{j=1}^r \mathrm{GL}_{n_j}$ for some r and n_j's; then τ factors through τ^M if and only if there is an isomorphism $\tau_K \cong \oplus_j \tau_K^{(j)}$ with $\dim \tau_K^{(j)} = n_j$ for all j. In general, whenever $(G, B, T) \cong \prod_j (G_j, B_j, T_j)$ factors as a product of pinned groups, then $\eta = \sum_j \eta_j$, where η_j is a local twisting element of G_j, and $\mathrm{W}_{\mathrm{expl}, \eta}(\tau_1 \times \cdots \times \tau_r) = \{\boxtimes_j F_j : F_j \in \mathrm{W}_{\mathrm{expl}, \eta_j}(\tau_j)\}$ where the subscripts η and η_j indicate the dependence on the local twisting element. Moreover, if $\mu = \sum \mu_j$ with $\mu_j \in X(T_j)_+$, then $F(\mu) \cong \boxtimes_j F^{G_j}(\mu_j)$ and $W(\mu) \cong \boxtimes_j W^{G_j}(\mu_j)$. From Proposition 9.3.7 we get that $\mathrm{W}_{\mathrm{expl}, \eta}(\tau) = \mathrm{W}_{\mathrm{expl}, \eta'}(\tau) \otimes F(\eta' - \eta)$ whenever η' is another local twisting element. Putting these observations together, we see that $\mathrm{W}_{\mathrm{expl}}(\tau) = \mathrm{W}_{\mathrm{expl}}(\tau_K)$, where the latter is computed according to Definition 7.2.3. $\qquad\square$

9.6. A unitary group example.

We work out another example of the constructions in this section. Suppose that $n \geq 1$ and let $J = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}$. For any $\mathbb{Z}_p$-algebra A define $G(A) = \{g \in \mathrm{GL}_n(A \otimes_{\mathbb{Z}_p} \mathbb{Z}_{p^2}) : {}^t g \cdot J \cdot \overline{g} = J\}$, where conjugation $g \mapsto \overline{g}$ is trivial on A and the non-trivial Galois automorphism on $\mathbb{Z}_{p^2}$. Then G is a connected reductive group over $\mathbb{Z}_p$ with generic fibre the unramified unitary group over $\mathbb{Q}_p$ of absolute rank n.

We consider the upper-triangular Borel B and diagonal maximal torus T. Then the splitting field is $\mathbb{Q}_{p^2}$ and $\varphi \in \mathrm{Gal}(\mathbb{Q}_{p^2}/\mathbb{Q}_p)$ acts as $(a_1, \ldots, a_n) \mapsto -(a_n, \ldots, a_1)$ on $X(T) \cong \mathbb{Z}^n$, i.e. as $-w_0$. We assume that n is odd so that G has a local twisting element, namely $\eta = (\frac{n-1}{2}, \frac{n-3}{2}, \ldots, -\frac{n-1}{2})$. As dual group we take $\widehat{G} = \mathrm{GL}_n$ over $\mathbb{Z}_p$ with upper-triangular Borel $\widehat{B}$ and diagonal maximal torus $\widehat{T}$, with pinning given by the isomorphisms sending $a \in \mathbb{G}_a$ to the upper-triangular unipotent matrix having unique off-diagonal element a in the $(i, i+1)$-entry (for $1 \leq i < n$) and obvious identification ϕ (as in §9.5). With these choices, ${}^L G = \mathrm{GL}_n \rtimes \mathrm{Gal}(\mathbb{Q}_{p^2}/\mathbb{Q}_p)$ with φ acting as $g \mapsto J' \cdot {}^t g^{-1} \cdot (J')^{-1}$, where $J' = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ -1 & & \end{pmatrix}$ (with alternating signs along the diagonal).

We can identify $\underline{G}$ with $\mathrm{GL}_{n/\overline{\mathbb{F}}_p}$ via the inclusion $\mathbb{F}_{p^2} \hookrightarrow \overline{\mathbb{F}}_p$, and then $F = -pw_0$ on $X(\underline{T})$. Serre weights are identified with equivalence classes $X_1^{(n)}/\sim$, where $a \sim a'$ if and only if $a - a' \in (p+1, \ldots, p+1)\mathbb{Z}$.

To finish, here is an explicit example with $n = 3$. Consider $\mu = (a, b, c)$ sufficiently generic in the lowest alcove and suppose that the inertial L-parameter $\tau : I_{\mathbb{Q}_p} \to \mathrm{GL}_3(\overline{\mathbb{F}}_p)$ is given by $\tau \cong \tau(1, \mu)$, i.e. $\tau \cong \begin{pmatrix} \omega_2^{a-pc} & & \\ & \omega_2^{b(1-p)} & \\ & & \omega_2^{c-pa} \end{pmatrix}$. Then

$$\mathrm{W}_{\mathrm{expl}}(\tau) = \{F((a, b, c) - \eta), F((b, c - 1, a - p) - \eta), F((c + p, a + 1, b) - \eta),$$
$$F((c + p - 1, b, a - p + 1) - \eta), F((a, c, b - p - 1) - \eta), F((b + p + 1, a, c) - \eta),$$
$$F((c + p, b, a - p) - \eta), F((a, c - 1, b - p) - \eta), F((b + p, a + 1, c) - \eta)\},$$

where the first six weights are obvious and the last three are shadows. For example, the second weight is obvious by Proposition 9.3.7 since $\tau \cong \tau((1\,2\,3), (b, c-1, a-p))$. Note that there are no obscure weights by Remarks 9.3.11 and 9.3.13, as G and T are the only Levi subgroups of G that contain T.

10. COMPARISON WITH [Her09]

In this section we will prove that for L-parameters $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ that are sufficiently generic, the sets $\mathrm{W}^?(\tau)$ and $\mathrm{W}_{\mathrm{expl}}(\tau)$ are equal. This establishes in particular that Conjecture 7.2.7 is in agreement with the Serre weight conjecture of [Her09].

10.1. **The weight set $\mathrm{W}^?(\tau)$ in the generic case.** We begin by giving an alternate characterization of the set $\mathrm{W}^?(\tau)$ (for τ sufficiently generic in a sense to be made precise below) in terms of the $\uparrow$ relation on alcoves. We refer the reader to [Jan03, II.6.5] for the definition of the $\uparrow$ relation; see also [Her09, Def. 3.15].

We use the same notation as in Section 9.2. Recall from the proof of Lemma 9.2.4 that there is a finite order automorphism π of $(\underline{G}, \underline{B}, \underline{T})$ that induces the action of φ^{-1} on $\Psi_0(\underline{G}, \underline{B}, \underline{T})$. In particular, $F = p\pi^{-1}$ on $X(\underline{T})$.

Let $\Phi \subset X(\underline{T})$ denote the set of roots, Φ^+ the subset of positive roots, $W_p := p\mathbb{Z}\Phi \rtimes W$ the affine Weyl group, and $\widetilde{W}_p := pX(\underline{T}) \rtimes W$ the extended affine Weyl group. We refer to [Jan03, II.6] for the definition of alcoves and for the basic facts about them. We denote by C_0 the lowest (or fundamental) alcove. We say that a weight $\lambda \in X(\underline{T})$ is p-regular if it does not lie on any alcove walls; equivalently, $\mathrm{Stab}_{W_p}(\lambda) = 1$.

Lemma 10.1.1. Suppose $\lambda \in X(\underline{T})$ is p-regular. Then $\mathrm{Stab}_{\widetilde{W}_p}(\lambda) = 1$.

Proof. The proof of [Her09, Lemma 5.6] applies, as $Z(\underline{G})$ is connected. $\square$

Recall from [Her09, §5.2] the definition of $\mu \in X(\underline{T})$ lying δ-*deep* in an alcove. We say that a statement is true for μ lying *sufficiently deep* in some alcove C if there is a $\delta > 0$ depending only on the based root datum $\Psi_0(\underline{G}, \underline{B}, \underline{T})$ together with its automorphism π (and in particular not on p) such that the statement holds for all μ which are δ-deep in C.

Recall from [Jan03, II.9] the definitions of $\underline{G}_1\underline{T}$-modules $\widehat{Z}_1(\lambda)$ and $\widehat{L}_1(\lambda)$ for $\lambda \in X(\underline{T})$, where the group scheme $\underline{G}_1$ is the kernel of $F : \underline{G} \to \underline{G}$. Supposing there

exists $\mu \in C_0 \cap X(\underline{T})$ (equivalently, $p > \langle \eta, \alpha^\vee \rangle \ \forall \alpha \in \Phi^+$), let

$$D_1 := \{u \in \widetilde{W}_p : u \cdot \mu \in X_1(\underline{T})\}.$$

This set is independent of the choice of μ, and it is a finite union of $pX^0(\underline{T})$-cosets.

Proposition 10.1.2. *For μ lying sufficiently deep in the alcove C_0, we have*

$$\overline{R(w, \mu + \eta)} = \sum_{u \in D_1/pX^0(\underline{T})} \sum_{\nu \in X(\underline{T})} [\widehat{Z}_1(\mu + p\eta) : \widehat{L}_1(p\nu + u \cdot \mu)] \, F(u \cdot (\mu + w\pi\nu))$$

in the Grothendieck group of finite-dimensional $\overline{\mathbb{F}}_p[G(\mathbb{F}_p)]$-modules.

Remark 10.1.3. Recall that the Deligne–Lusztig representation $R(w, \mu)$ was defined in Definition 9.2.2. The notation $[\widehat{Z}_1(\lambda) : \widehat{L}_1(\mu)]$ signifies the multiplicity of the simple $\underline{G}_1\underline{T}$-module $\widehat{L}_1(\mu)$ as Jordan–Hölder factor of $\widehat{Z}_1(\lambda)$.

Remark 10.1.4. By Lemma 9.2.4 the inner sum depends only on the coset of u in $D_1/pX^0(\underline{T})$.

Proof. This proposition is a generalisation of Jantzen's generic decomposition formula for Deligne–Lusztig representations [Jan81, Satz 4.3]. For a generalisation of [Jan81, §1–3] to reductive groups with simply connected derived subgroups, see the appendix to [Her09]. We now explain how [Jan81, §4] generalises to the same context. We only leave aside the part of [Jan81, §4.1] that follows equation 4.1(2). Without further comment, any reference in the remainder of this paragraph will be to [Jan81], and we keep the same notation and conventions as in the appendix of [Her09]. For example, any occurrence of $\rho, \rho_w, \varepsilon_w, \gamma_{w_1,w_2}$ should be replaced by $\rho', \rho'_w, \varepsilon'_w, \gamma'_{w_1,w_2}$. In addition, in §4 any occurence of the term "$\mu + \rho$" should be replaced by "$\mu + \pi\rho'$". We let $h := \max\{\langle \rho', \alpha^\vee \rangle + 1 : \alpha \in R^+\}$. Furthermore, any occurence of D_n as the index of a sum should be replaced by (a fixed set of representatives of) $D_n/p^nX^0(T)$, which is finite. In particular, Satz 4.3 says that

$$\widetilde{R}_w(n, \mu + \pi\rho') = \Psi \sum_{\substack{u \in D_n/p^nX^0(T) \\ \nu \in X(T)}} [\widehat{Z}(n, \mu + p^n\rho') : \widehat{L}(n, p^n\nu + u \cdot \mu)] \, \chi_p(u \cdot (\mu + w\pi\nu)).$$

In §4.3 and §4.4, $\alpha_0^\vee$ denotes any choice of highest coroot. The inequality in line -3 of page 472 is no longer true in general, but the following line still holds. In line -1 of page 472 the second occurrence of ν_1 should be ν_2 (a typo). In the proof of Lemma 4.4, $-w'\varepsilon'_{w_0w'} = \rho' - \rho'_{w'}$ only holds modulo $X^0(T)$, but this is sufficient. The diagonal elements of the upper-triangular matrix now lie in $X^0(T) \subset \mathbb{Z}[X(T)]^W$, and so the terms χ likewise need to be multiplied by elements of $X^0(T)$. Similar comments apply to the following two displayed equations.

To deduce our proposition, we choose Jantzen's split $G/\mathbb{F}_p$ such that $G \times \overline{\mathbb{F}}_p \cong \underline{G}$ with relative Frobenius $F \circ \pi^{-1}$ (see the proof of Lemma 9.2.4). We then choose $\rho' = \eta$ (noting that $\pi\eta = \eta$), take $n = 1$, and use Lemma 10.1.1. $\qquad \square$

Lemma 10.1.5. *Suppose $p \geq 2\max\{\langle \eta, \alpha^\vee \rangle : \alpha \in \Phi^+\}$. For weights $\lambda \in X(\underline{T})$, $\mu \in X_1(\underline{T})$ we have*

$$(10.1.6) \qquad [\widehat{Z}_1(\lambda) : \widehat{L}_1(\mu)] \neq 0 \iff \sigma \cdot (\lambda - p\eta) \uparrow w_0 \cdot (\mu - p\eta) \text{ for all } \sigma \in W.$$

Remark 10.1.7. Here we do not need to assume that $Z(\underline{G})$ is connected or that η is $\mathrm{Gal}(L/\mathbb{Q}_p)$-invariant. We remark that [DS87, Cor. 2.7] relies on Corollary A.1.2(ii).

Proof. First suppose that $\underline{G} = \underline{G}^{\mathrm{der}}$. Then [DS87, Cor. 2.7] shows that $[\widehat{Z}_1(\lambda) : \widehat{L}_1(\mu)] \neq 0$ if and only if $\mu \in \cap_{y \in W_v} I_{y,1}^{-1} \cdot \mathrm{SL}(y \cdot \lambda)$ in their notation, where $v \in -\eta + pX(\underline{T})$ is arbitrary. Taking $v = -\eta$ and $y \in W_{-\eta} = W$, we find that this is equivalent to $\mu + p(y\eta - \eta) \uparrow y \cdot \lambda$ for all $y \in W$, or equivalently $w_0 y \cdot (\lambda - p\eta) \uparrow w_0 \cdot (\mu - p\eta)$ for all $y \in W$.

In the general case, note that both sides of (10.1.6) imply that $\mu \in W_p \cdot \lambda$. For $\mu \in W_p \cdot \lambda$ we have $[\widehat{Z}_1(\lambda) : \widehat{L}_1(\mu)]_{\underline{G}} = [\widehat{Z}_1(\lambda) : \widehat{L}_1(\mu)]_{\underline{G}^{\mathrm{der}}}$ and $\mu \uparrow \lambda$ if and only if $\mu|_{\underline{T} \cap \underline{G}^{\mathrm{der}}} \uparrow \lambda|_{\underline{T} \cap \underline{G}^{\mathrm{der}}}$. (Note that $W_p \cdot \lambda \subset \lambda + \mathbb{Z}\Phi$. The restriction map $X(\underline{T}) \twoheadrightarrow X(\underline{T} \cap \underline{G}^{\mathrm{der}})$ induces a bijection $\lambda + \mathbb{Z}\Phi \xrightarrow{\sim} \lambda|_{\underline{T} \cap \underline{G}^{\mathrm{der}}} + \mathbb{Z}\Phi$, which identifies $\leq$, W_p-actions, and hence $\uparrow$, on both sides. Also, $\widehat{Z}_1(\lambda)$, $\widehat{L}_1(\mu)$ restrict to corresponding objects for $\underline{G}^{\mathrm{der}}$, see the proof of [Her09, Prop. 5.7].) This reduces the claim to the case $\underline{G} = \underline{G}^{\mathrm{der}}$. $\square$

Proposition 10.1.8. *For μ lying sufficiently deep in the alcove C_0, and for $\lambda \in X_1(\underline{T})$, we have that $F(\lambda)$ is a Jordan–Hölder constituent of $\overline{R(w, \mu + \eta)}$ if and only if there exists $\nu \in X(\underline{T})$ such that*

$$\sigma \cdot (\mu + (w\pi - p)\nu) \uparrow w_0 \cdot (\lambda - p\eta) \qquad \text{for all } \sigma \in W.$$

Proof. For μ lying sufficiently deep in C_0 we have $p \geq 2\max\{\langle \eta, \alpha^\vee \rangle : \alpha \in \Phi^+\}$, so we can (and do) assume this inequality. From Proposition 10.1.2 we know that

$$\overline{R(w, \mu + \eta)} = \sum_{u \in D_1/pX^0(\underline{T})} \sum_{\nu \in X(\underline{T})} [\widehat{Z}_1(\mu + p(\eta - \nu)) : \widehat{L}_1(u \cdot \mu)] F(u \cdot (\mu + w\pi\nu)).$$

By Lemma 10.1.5, the (u, ν) term of the double sum is non-zero if and only if

$$(10.1.9) \qquad \sigma \cdot (\mu - p\nu) \uparrow w_0 \cdot (u \cdot \mu - p\eta) \qquad \text{for all } \sigma \in W.$$

As in the proof of [Her09, Prop. 5.7], for μ sufficiently deep in C_0 we obtain $\sigma \cdot (\mu + (w\pi - p)\nu) \uparrow w_0 \cdot (u \cdot (\mu + w\pi\nu) - p\eta)$ for all $\sigma \in W$, which proves the "only if" part of the proposition. (Note that in (10.1.9) there are only finitely many possibilities for ν modulo $X^0(\underline{T})$, independent of μ.)

Conversely, if $\sigma \cdot (\mu + (w\pi - p)\nu) \uparrow w_0 \cdot (\lambda - p\eta)$ for all $\sigma \in W$, we may reverse the above argument, as explained in the proof of [Her09, Prop. 5.7]. $\square$

Lemma 10.1.10. *For μ lying sufficiently deep in the alcove C_0, we have that $(\underline{T}_w, \theta_{w,\mu})$ is maximally split for all $w \in W$.*

Proof. The dual pair is $(\mathbb{T}^*, s)$ with $\mathbb{T}^* = \underline{T}^*_{F^*(w^{-1})}$ and $s = g^*_{F^*(w^{-1})} s'(g^*_{F^*(w^{-1})})^{-1}$, and where $s' := N_{(F^* \circ w^{-1})^d / F^* \circ w^{-1}} \mu(\zeta_{p^d-1})$ with $d > 0$ chosen such that $(F^* \circ w^{-1})^d = p^d$. We can define $\hat{s} : X(\mathbb{T}^*) \to \overline{\mathbb{Q}}_p^\times$ by $\hat{s}(\mu) := \widetilde{\mu(s)}$. Then $\widetilde{w(s)} = w(\hat{s})$ for $w \in N(\mathbb{T}^*)/\mathbb{T}^*$, so $\mathrm{Stab}_{N(\mathbb{T}^*)/\mathbb{T}^*}(\hat{s}) \cong \mathrm{Stab}_{N(\mathbb{T}^*)/\mathbb{T}^*}(s) \cong \mathrm{Stab}_W(s')$. By [DL76, Thm. 5.13], this group is generated by reflections, as $Z(\underline{G})$ is connected. A reflection

$s_\alpha \in W$ fixes $s' \in \underline{T}^*$ if and only if

$$(1 - s_\alpha) \left(\sum_{i=0}^{d-1} (F^* \circ w^{-1})^i \mu \right) (\zeta_{p^d - 1}) = 1$$

$$\iff (1 - s_\alpha) \left(\sum_{i=0}^{d-1} (F^* \circ w^{-1})^i \mu \right) \equiv 0 \pmod{(p^d - 1)X(\underline{T})}$$

$$\iff \left\langle \sum_{i=0}^{d-1} (p\pi^{-1} w^{-1})^i \mu, \alpha^\vee \right\rangle \equiv 0 \pmod{p^d - 1},$$

where we used that $\langle \alpha, Y(\underline{T}) \rangle = \mathbb{Z}$, as $Z(\underline{G})$ is connected. Equivalently,

$$\sum_{i=0}^{d-1} p^i \langle \mu, (w\pi)^i \alpha^\vee \rangle \equiv 0 \pmod{p^d - 1}.$$

If $\mu \in C_0$, the left-hand side has to be zero, so $\langle \mu, \alpha^\vee \rangle \equiv 0 \pmod p$, and this is impossible if μ lies $(h-1)$-deep in C_0, where $h = \max\{\langle \eta, \beta^\vee \rangle + 1 : \beta \in \Phi^+\}$. Thus for μ lying sufficiently deep in C_0, the Weyl group of $\mathbb{T}^*$ in the connected reductive group $Z_{\underline{G}^*}(s)$ is trivial, so $Z_{\underline{G}^*}(s) = \mathbb{T}^*$, which implies that $(\mathbb{T}^*, s)$ is maximally split. $\qquad\square$

The group $X(\underline{T}) \rtimes W$ acts on the set $W \times X(\underline{T})$ by

$$(10.1.11) \qquad {}^{(\nu, \sigma)}(w, \mu) = (\sigma w \pi \sigma^{-1} \pi^{-1}, \sigma\mu + (p - \sigma w \pi \sigma^{-1})\nu),$$

see [Jan81, §3.1]. This action has the same orbits as the action considered in [Her09, §4.1], as $F = p\pi^{-1}$ on $X(\underline{T})$ and $F(\sigma) = \pi\sigma\pi^{-1}$ in $\mathrm{Aut}(\underline{T})$ for $\sigma \in W$. In particular, [Her09, Lem. 4.2] still applies. Note also that $\tau(w, \mu)$ depends only on the orbit of (w, μ).

Definition 10.1.12. We say that a tame inertial L-parameter τ is δ-*generic* if $\tau \cong \tau(w, \mu)$ for some $w \in W$ and μ lying δ-deep in C_0. As in [Her09, §6.5] we say that a statement is true for all *sufficiently generic* tame inertial L-parameters τ if it holds for all $\tau \cong \tau(w, \mu)$ with $w \in W$ and for μ lying sufficiently deep in C_0; in other words, if there exists $\delta > 0$ depending only in the based root datum $\Psi_0(\underline{G}, \underline{B}, \underline{T})$ together with its automorphism π (and in particular not on p) such that the statement holds for all τ that are δ-generic.

We remark that this definition of δ-generic differs slightly from the one given at [Her09, Def. 6.27], as we do not require that the pair (w, μ) occurring in the definition be "good" (cf. [Her09, Def. 6.19]). On the other hand, this change does not affect what it means for a statement to be true for all sufficiently generic τ: indeed, [Her09, Lem. 6.24] (whose analogue in this paper is Lemma 10.1.10) shows that the pair (w, μ) is automatically good for μ lying sufficiently deep in C_0.

Proposition 10.1.13. *For all sufficiently generic tame inertial L-parameters τ :* $I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ *and for all $\lambda \in X_1(\underline{T})$, we have that $F(\lambda) \in \mathrm{W}^?(\tau)$ if and only if $\tau \cong \tau(w, \lambda' + \eta)$ for some dominant $\lambda' \uparrow \lambda$ and some $w \in W$.*

Remark 10.1.14. Alternatively the equivalence holds for λ lying sufficiently deep in a restricted alcove. (See [Her09, Prop. 6.28].)

Proof. Write $\tau \cong \tau(w, \mu + \eta)$. For μ lying sufficiently deep in C_0, Proposition 10.1.8 gives that $W^?(\tau)$ consists of the Serre weights $F(\lambda)$ for $\lambda \in X_{\mathrm{reg}}(\underline{T})$ such that there exists $\nu \in X(\underline{T})$ with

$$\sigma \cdot (\mu + (w\pi - p)\nu) \uparrow \lambda \qquad \text{for all } \sigma \in W.$$

Equivalently, this relation holds for the unique σ making the left-hand side dominant. The proof concludes as in [Her09, Prop. 6.28], using Lemma 10.1.10 and the formula (10.1.11). $\qquad\square$

10.2. **The main result.** The main result of this section is Theorem 10.2.11, which shows that for all sufficiently generic tame inertial L-parameters τ the sets $W^?(\tau)$ and $W_{\mathrm{expl}}(\tau)$ coincide, and moreover that the Levi predictions of Definition 9.3.10 do not produce any new weights beyond those already in $\mathcal{C}(W_{\mathrm{obv}}(\tau))$.

Define $\|\lambda\| := \sum_{\alpha > 0} \langle \lambda, \alpha^\vee \rangle$ for $\lambda \in X(\underline{T})$. Then we have:

   (i) $\lambda < \mu \implies \|\lambda\| < \|\mu\|$.
   (ii) $\lambda \in X(\underline{T})_+ \implies \|\lambda\| \geq 0$, with equality if and only if $\lambda \in X^0(\underline{T})$.
   (iii) $\|\lambda\| = \|\pi(\lambda)\|$ for all $\lambda \in X(\underline{T})$.

Lemma 10.2.1. *Fix $N \in \mathbb{Z}_{>0}$. Suppose that $\lambda \in X(\underline{T})_+$ with $\|\lambda\| < Np$.*

   (i) *For all sufficiently generic τ, if $\tau \cong \tau(w, \lambda)$ then λ is as deep as we like in its alcove.*
   (ii) *For $\lambda' \in X_1(\underline{T})$ lying sufficiently deep in a restricted alcove, if $F(\lambda') \in \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ then λ is as deep as we like in its alcove.*

For instance, to be precise, the statement in (i) means that for each fixed $\delta > 0$ and for all sufficiently generic τ, if $\tau \cong \tau(w, \lambda)$ then λ is δ-deep in its alcove; the meaning of (ii) is similar.

Proof. (i) There is a finite collection of alcoves (independent of p) such that any λ allowed by $\|\lambda\| < Np$ lies in the closure of one of them. Therefore, as explained after [Her09, Def. 6.27], modulo $(p-1)X^0(\underline{T})$ there are only finitely many possible λ (independent of p) and, for τ sufficiently generic, each one is δ-deep in its alcove. (In the paragraph before [Her09, Prop. 6.28] note that $p - \pi$ is injective on the free abelian group $X(\underline{T})/\mathbb{Z}\Phi$, as π has finite order.)

(ii) In the argument that follows, if ν is an element of $X(\underline{T})_+$ we will often write $\nu = \nu_0 + p\nu_1$ with $\nu_0 \in X_1(\underline{T})$ and $\nu_1 \in X(\underline{T})_+$ (so that ν_1 is unique modulo $X^0(\underline{T})$).

Choose $\mu \in X(\underline{T})_+$ such that $F(\lambda') \in \mathrm{JH}_{G(\mathbb{F}_p)} F(\mu)$ and $F(\mu) \in \mathrm{JH}_{\underline{G}} W(\lambda)$. Then $\mu \uparrow \lambda$, so $\|\mu\| < Np$ and μ lies as deep in its alcove as λ.

If $\mu \in X_1(\underline{T})$, then $\mu \equiv \lambda' \pmod{(p - \pi)X^0(\underline{T})}$, and we are done. Otherwise, $F(\mu) \cong F(\mu_0) \otimes F(\mu_1)^{(\pi)}$ as $G(\mathbb{F}_p)$-representations, so there exists $\mu^{(1)} \in X(\underline{T})_+$ such that $F(\lambda') \in \mathrm{JH}_{G(\mathbb{F}_p)} F(\mu^{(1)})$ and $F(\mu^{(1)}) \in \mathrm{JH}_{\underline{G}}(F(\mu_0) \otimes F(\mu_1)^{(\pi)})$. In particular, $\mu^{(1)} \leq \mu_0 + \pi\mu_1$, so as $\mu_1 \notin X^0(\underline{T})$ we have

$$(10.2.2) \qquad \|\mu^{(1)}\| \leq \|\mu_0\| + \|\mu_1\| < \|\mu\| - (p-1) < \|\mu\| - p/2.$$

Iterating, we can find a sequence of dominant weights $\mu = \mu^{(0)}, \mu^{(1)}, \ldots, \mu^{(r)}$ with

   $\circ$ $F(\mu^{(i+1)}) \in \mathrm{JH}_{\underline{G}}(F(\mu_0^{(i)}) \otimes F(\mu_1^{(i)})^{(\pi)})$ for all $0 \leq i < r$,
   $\circ$ $\mu^{(i)} \notin X_1(\underline{T})$ for all $0 \leq i < r$, but $\mu^{(r)} \in X_1(\underline{T})$,
   $\circ$ $F(\mu^{(r)}) \cong F(\lambda')$ as $G(\mathbb{F}_p)$-representations.

Moreover, by (10.2.2), we know that $r < 2N$.

On the other hand, as in the proof of [Her09, Prop. 9.1] we can write $F(\mu_0) \otimes F(\mu_1)^{(\pi)} = \sum a_\varepsilon b_{\mu'_0} W(\mu'_0 + \pi\varepsilon)$, where the sum runs over $\varepsilon \in X(\underline{T})$ such that $w\varepsilon \leq \mu_1$ for all $w \in W$, and dominant $\mu'_0 \uparrow \mu_0$. Hence $\mu^{(1)} \uparrow \sigma \cdot (\mu'_0 + \pi\varepsilon)$ for some such ε, μ'_0 and some $\sigma \in W$. It follows that if $\mu^{(1)}$ is $(\delta + N)$-deep in its alcove, then μ'_0 (and hence μ) is δ-deep in its alcove. Therefore, as $r < 2N$, if λ' is $(\delta + 2N^2)$-deep in its alcove, then λ is δ-deep in its alcove. $\qquad\square$

Lemma 10.2.3. *Suppose $M \subset G$ is a Levi subgroup containing T, and that $\tau^M : I_{\mathbb{Q}_p} \to \widehat{M}(\overline{\mathbb{F}}_p)$ denotes a tame inertial L-parameter. Let $\tau : I_{\mathbb{Q}_p} \to \widehat{G}(\overline{\mathbb{F}}_p)$ denote the composite of τ^M and the inclusion $\widehat{M} \subset \widehat{G}$. Fix $\delta > 0$. For all τ sufficiently generic, τ^M is δ-generic.*

Proof. Write $\tau^M \cong \tau^M(w, \lambda)$ with $w \in W_M$, $\lambda \in X(\underline{T})$. Write $\lambda = \lambda_0 + p\lambda_1$ with $\lambda_0 \in X_1(\underline{T})$, $\lambda_1 \in X(\underline{T})_+$. Then $\tau^M(w, \lambda) \cong \tau^M(w, \lambda')$, where $\lambda' = \lambda_0 + w\pi\lambda_1$. For $\nu \in X(\underline{T})$ let $|\nu| := \sum_{\alpha > 0} |\langle v, \alpha^\vee \rangle|$. Then

$$|\lambda'| \leq |\lambda_0| + |\lambda_1| = \frac{p-1}{p}|\lambda_0| + \frac{1}{p}|\lambda| \leq \frac{(p-1)^2}{p} \sum_\Delta n_\beta + \frac{|\lambda|}{p},$$

where we write $\sum_{\alpha > 0} \alpha^\vee = \sum_{\beta \in \Delta} n_\beta \beta^\vee$.

Iterating, we deduce that $\tau^M \cong \tau^M(w, \mu)$ with $|\mu| < p\sum_\Delta n_\beta$, so μ lies in the closure of a finite union of alcoves (for $\underline{G}$, hence also for $\underline{M}$). A fortiori, $\tau \cong \tau(w, \mu)$. As in the proof of Lemma 10.2.1(i), for τ sufficiently generic we have that μ lies as deep as we like in its alcove (for $\underline{G}$, hence also for $\underline{M}$). By reversing the argument we deduce that τ^M is as generic as we like. $\qquad\square$

Lemma 10.2.4. *Suppose that $\lambda, \mu \in X(\underline{T})_{+,M} - \eta$. Choose $w, w' \in W$ such that $w \cdot \lambda, w' \cdot \mu$ are in $X(\underline{T})_+ - \eta$. Then*

$$\lambda \uparrow_M \mu \qquad \Longrightarrow \qquad w \cdot \lambda \uparrow w' \cdot \mu.$$

Proof. Let $\mu' := w' \cdot \mu$, the unique element in $(X(\underline{T})_+ - \eta) \cap W \cdot \mu$. We may assume that $w' \in W$ has least possible length, i.e. w' is a Kostant representative for $\mathrm{Stab}_W(\mu' + \eta) \backslash W$ (noting that the stabiliser is generated by simple reflections).

First we claim that $w'(\Phi_M^+) \subset \Phi^+$, or equivalently $w'(\Delta_M) \subset \Phi^+$. Suppose that $\alpha \in \Delta_M$. As $\mu \in X(\underline{T})_{+,M} - \eta$ we know that $\langle \mu' + \eta, w'(\alpha)^\vee \rangle \geq 0$. Hence if $w'(\alpha) \in \Phi^-$, then equality holds, i.e. $s_{w'(\alpha)} \in \mathrm{Stab}_W(\mu' + \eta)$. By our choice of w' it follows that $w's_\alpha = s_{w'(\alpha)}w' > w'$, hence $w'(\alpha) \in \Phi^+$. This proves the claim.

By Corollary A.1.2 and induction, we may assume that $\lambda = s_{\alpha, np} \cdot \mu$ for some $\alpha \in \Phi_M^+$, $n \in \mathbb{Z}$ and that $\lambda \neq \mu$. As $\lambda \in X(\underline{T})_{+,M} - \eta$, we deduce that $\langle \mu + \eta, \alpha^\vee \rangle > np > 0$. Hence $w' \cdot \lambda = s_{w'(\alpha), np}w' \cdot \mu$ with $\langle w' \cdot \mu + \eta, w'(\alpha)^\vee \rangle > np > 0$ and $w'(\alpha) \in \Phi^+$ by the above. Then [Jan03, II.6.9] shows that

$$w''w' \cdot \lambda = w''s_{w'(\alpha), np}w' \cdot \mu \uparrow w' \cdot \mu$$

for any $w'' \in W$ making $w''w'(\lambda + \eta)$ dominant. $\qquad\square$

Lemma 10.2.5. *Suppose that $\mu \in X(\underline{T})_+ - \eta$ and that $\nu \in X(\underline{T})_+$. Then for $\lambda \in X(\underline{T})_+ - \eta$, we have $\lambda \uparrow \mu + p\nu$ if and only if $\lambda = \sigma \cdot (\mu' + p\varepsilon)$ for some $\sigma \in W$, some $\mu' \uparrow \mu$ with $\mu' \in X(\underline{T})_+ - \eta$ and some $\varepsilon \in X(\underline{T})$ such that $w\varepsilon \leq \nu$ for all $w \in W$.*

Proof. Let $X(\mu, \nu)$ denote the subset of $\lambda \in X(\underline{T})_+ - \eta$ defined by the right-hand side of the claimed equivalence.

For the "if" direction of the lemma note that $w \cdot (\mu' + p\varepsilon) \uparrow \mu' + p\nu$ for all $w \in W$ by [Her09, Lem. 9.4] and that $\mu' + p\nu \uparrow \mu + p\nu$ by [Jan03, II.6.4(4)]. (We note that the proof of [Her09, Lem. 9.4] holds in our more general context. The only necessary modifications are that in the statement of that lemma the weights μ, ν are to be taken in $X(\underline{T})_+ - \eta$ and in the proof of reduction step (R1) we may assume $i > 0$ and then the first displayed inequality becomes $0 < pi \le \langle \lambda' + \eta, w\alpha^\vee \rangle$.)

Conversely, suppose $\lambda \uparrow \mu + p\nu$ with λ, μ, ν as in the statement of the lemma. By Corollary A.1.2 there is a sequence $\lambda = \lambda_r \uparrow \lambda_{r-1} \uparrow \ldots \uparrow \lambda_0 = \mu + p\nu$, where $\lambda_i \in X(\underline{T})_+ - \eta$ and there exist affine reflections $s_{\alpha_i, n_i p} = s_{\alpha_i} + n_i p \alpha_i \in W_p$ ($\alpha_i \in \Phi^+, n_i \in \mathbb{Z}$) such that $\lambda_{i+1} = s_{\alpha_i, n_i p} \cdot \lambda_i$. Without loss of generality, $\lambda_{i+1} < \lambda_i$ for all i.

We now show that $\lambda_i \in X(\mu, \nu)$ by induction on i. This is obvious when $i = 0$. For the induction step we are reduced to the following statement. Given λ, λ' in $X(\underline{T})_+ - \eta$ such that $\lambda = s_{\alpha, np} \cdot \lambda'$ with $\langle \lambda' + \eta, \alpha^\vee \rangle > np$ and $\alpha \in \Phi^+$, then $\lambda' \in X(\mu, \nu)$ implies $\lambda \in X(\mu, \nu)$. (Note that here λ no longer denotes the element λ_r above.) Note that $n > 0$, as $np > \langle \lambda + \eta, \alpha^\vee \rangle \ge 0$. As $\lambda' \in X(\mu, \nu)$ we can write $\lambda' = \sigma \cdot (\mu' + p\varepsilon)$ as in the statement of the lemma. Then

$$(10.2.6) \qquad \lambda = s_{\alpha, np} \cdot \lambda' = s_\alpha \sigma \cdot (\mu' + p(\varepsilon - n\sigma^{-1}\alpha)).$$

Case 1: Assume that $\langle \varepsilon, \sigma^{-1}\alpha^\vee \rangle \ge n$. To see that $\lambda \in X(\mu, \nu)$, by (10.2.6) it suffices to show that $w(\varepsilon - n\sigma^{-1}\alpha) \le \nu$ for all $w \in W$. Let $\varepsilon' := s_{\sigma^{-1}\alpha}\varepsilon = \varepsilon - \langle \varepsilon, \sigma^{-1}\alpha^\vee \rangle \sigma^{-1}\alpha$. As $\langle \varepsilon, \sigma^{-1}\alpha^\vee \rangle \ge n$, the sequence $w\varepsilon$, $w(\varepsilon - n\sigma^{-1}\alpha)$, $w\varepsilon'$ is monotonic with respect to $\le$ (i.e. either increasing or decreasing). As $w\varepsilon \le \nu$ and $w\varepsilon' \le \nu$ by our assumption on ε, we conclude that $w(\varepsilon - n\sigma^{-1}\alpha) \le \nu$.

Case 2: Assume that $\langle \varepsilon, \sigma^{-1}\alpha^\vee \rangle = n - r$ for some $r > 0$. As $\langle \mu' + \eta, \sigma^{-1}\alpha^\vee \rangle = \langle \lambda' + \eta, \alpha^\vee \rangle - p \langle \varepsilon, \sigma^{-1}\alpha^\vee \rangle$, we see that $\langle \mu' + \eta, \sigma^{-1}\alpha^\vee \rangle > rp$. As $\mu' \in X(\underline{T})_+ - \eta$ and $r > 0$, we get $\sigma^{-1}\alpha \in \Phi^+$. Let $w \in W$ be such that $\mu'' := ws_{\sigma^{-1}\alpha, rp} \cdot \mu' \in X(\underline{T})_+ - \eta$. Then $\mu'' \uparrow \mu'$ by [Jan03, II.6.9] and

$$\sigma w^{-1} \cdot (\mu'' + pw\varepsilon) = \sigma \cdot (s_{\sigma^{-1}\alpha} \cdot \mu' + rp\sigma^{-1}\alpha + p\varepsilon)$$
$$= s_\alpha \sigma \cdot (\mu' - rp\sigma^{-1}\alpha + p\varepsilon - p\langle \varepsilon, \sigma^{-1}\alpha^\vee \rangle \sigma^{-1}\alpha)$$
$$= s_\alpha \sigma \cdot (\mu' + p(\varepsilon - n\sigma^{-1}\alpha)),$$

which equals λ by (10.2.6). Hence $\lambda \in X(\mu, \nu)$. $\qquad\square$

Recall the definition of $d(C) \in \mathbb{Z}$ for an alcove C ([Jan03], II.6.6). For all $\alpha \in \Phi^+$ there is a unique $n_\alpha \in \mathbb{Z}$ such that

$$(10.2.7) \qquad n_\alpha p < \langle \lambda + \eta, \alpha^\vee \rangle < (n_\alpha + 1)p$$

for all $\lambda \in C$. Then $d(C) = \sum_{\Phi^+} n_\alpha$. If C is dominant, then $d(C)$ is the number of affine root hyperplanes separating C and the lowest alcove. If $\lambda \in C$, then we set $d(\lambda) := d(C)$. Note that if $\lambda, \mu \in X(\underline{T})$ are p-regular, then $d(\mu) \le d(\lambda)$ for $\mu \uparrow \lambda$ and $d(\mu) \le d(\mu + p\nu)$ for $\nu \in X(\underline{T})_+$, where equality holds only if $\mu = \lambda$, respectively $\nu \in X^0(\underline{T})$ ([Jan03], II.6.6).

Proposition 10.2.8. *Fix* $N \in \mathbb{Z}_{>0}$. *Then for τ sufficiently generic and any* $\lambda \in X(\underline{T})_+$ *with* $\|\lambda\| < Np$, *the following are equivalent.*

(i) $W^?(\tau) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda) \ne \varnothing$.

(ii) $\tau \cong \tau(w, \lambda' + \eta)$ *for some dominant $\lambda' \uparrow \lambda$ and some $w \in W$.*

*Moreover, if (ii) holds then $\lambda' = \lambda \in X_1(\underline{T})$ or there exists $F(\nu)$ in $W^?(\tau) \cap$
$\mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ with $d(\nu) < d(\lambda)$.*

Proof. (i) $\Rightarrow$ (ii): Suppose $F(\lambda') \in W^?(\tau) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ for some $\lambda' \in X_1(\underline{T})$. By
Proposition 10.1.13 we have $\tau \cong \tau(w, \lambda'' + \eta)$ for some dominant $\lambda'' \uparrow \lambda'$ and some
$w \in W$. By Lemma 10.2.1 we see that λ'', λ', and λ are as deep in their respective
alcoves as we like. To show (ii) we can now follow the proof of [Her09, Prop. 9.1],
noting that it never uses that λ is restricted (as is assumed there) and making the
following modifications: $F(\mu_1)$ should be replaced by its π-twist $F(\mu_1)^{(\pi)} \cong F(\pi\mu_1)$
and ρ by η. In the expressions $\mu_0 + \varepsilon, \mu_0' + \varepsilon, \mu_0 + w'\varepsilon$, ε should be replaced by
$\pi\varepsilon$. Starting with [Her09, (9.2)], the expression $pw^{-1}w'\varepsilon$ should be replaced by
$p\pi^{-1}w^{-1}w'\pi\varepsilon$, as well as $\sigma w \sigma^{-1}$ by $\sigma w \pi \sigma^{-1}\pi^{-1}$. (Note also that $\pi \in W$ in *loc.
cit.* is now a bad choice of letter.)

(ii) $\Rightarrow$ (i): Suppose that

(10.2.9) $\tau \cong \tau(w, \lambda' + \eta)$ for some dominant $\lambda' \uparrow \lambda$, some $w \in W$.

By Lemma 10.2.1, we see that λ' (and hence λ) lie as deep in their respective
alcoves as we like. If $\lambda' = \lambda \in X_1(\underline{T})$, then $F(\lambda) \in W^?(\tau) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ by
Proposition 10.1.13, as required. Thus from now on we may assume $\lambda' = \lambda \notin X_1(\underline{T})$
or $\lambda' \neq \lambda$.

We will first find $\lambda'' \in X(\underline{T})_+$ such that $F(\lambda'')$ is a $\underline{G}$-constituent of $W(\lambda)$, and
such that $\lambda'' \neq \lambda$ if $\lambda' \neq \lambda$. If $\lambda' = \lambda \notin X_1(\underline{T})$, we take $\lambda'' := \lambda$. If however
$\lambda' \neq \lambda$, choose $\lambda'' \neq \lambda$ maximal such that λ'' is dominant and $\lambda' \uparrow \lambda'' \uparrow \lambda$. By
Corollary A.1.2 there exists an affine reflection $s_{\beta,np} \in W_p$ ($\beta \in \Phi^+, n \in \mathbb{Z}$) such
that $s_{\beta,np} \cdot \lambda = \lambda''$. As λ'' is dominant, $\langle \lambda + \eta, \beta^\vee \rangle > np > 0$. Jantzen's sum
formula [Jan03, II.8.19] says that for a certain descending filtration $(V(\lambda)^i)_{i \geq 0}$ on
the Weyl module $V(\lambda)$ we have

$$\sum_{i>0} \mathrm{ch}\, V(\lambda)^i = \sum_{\alpha \in \Phi^+} \sum_{0 < mp < \langle \lambda+\eta, \alpha^\vee \rangle} \nu_p(mp) \, \mathrm{sgn}(w_{\alpha,m}) \, \mathrm{ch}\, W(w_{\alpha,m} s_{\alpha,mp} \cdot \lambda),$$

where $w_{\alpha,m} \in W$ is chosen such that $w_{\alpha,m} s_{\alpha,mp} \cdot \lambda$ is dominant. Note that the
p-adic valuation $\nu_p(mp)$ is positive, as $m > 0$. By [Jan03, II.6.8], for each term in
the sum, $w_{\alpha,m} s_{\alpha,mp} \cdot \lambda \uparrow \lambda$ and equality does not hold. Also, as λ is p-regular, all
$w_{\alpha,m} s_{\alpha,mp} \cdot \lambda$ that occur in this sum are distinct. (See also [Jan03, II.8.19, Rk. 3].)
Note that $w_{\beta,n} = 1$ by the previous paragraph. Therefore, by the maximality of λ''
and by the strong linkage principle, $F(\lambda'')$ is a $\underline{G}$-constituent of $W(\lambda)$, as claimed.
(It occurs once in $W(w_{\beta,n} s_{\beta,np} \cdot \lambda)$, but cannot occur in any other term.)

Suppose first that $\lambda'' \in X_1(\underline{T})$, so $\lambda'' \neq \lambda$. Then $F(\lambda'') \in W^?(\tau)$ by Proposi-
tion 10.1.13 and (10.2.9), so $F(\lambda'') \in W^?(\tau) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ and $d(\lambda'') < d(\lambda)$, as
required.

Alternatively, if $\lambda'' \notin X_1(\underline{T})$, then $\lambda'' = \lambda_0'' + p\lambda_1''$, where $\lambda_0'' \in X_1(\underline{T})$ and
$\lambda_1'' \in X(\underline{T})_+ - X^0(\underline{T})$. By Lemma 10.2.5, as $\lambda' \uparrow \lambda''$ we can write $\lambda' = \sigma \cdot (\mu + p\varepsilon)$
for some $\sigma \in W$, some dominant $\mu \uparrow \lambda_0''$, and some $\varepsilon \in X(\underline{T})$ such that $w'\varepsilon \leq \lambda_1''$
for all $w' \in W$. As $(w, \lambda' + \eta) = (w, \sigma(\mu + p\varepsilon + \eta))$ is in the same $X(\underline{T}) \rtimes W$-orbit
as $(w', \mu + \pi\varepsilon' + \eta)$, where $w' := \sigma^{-1}w\pi\sigma\pi^{-1}$ and $\varepsilon' := \pi^{-1}w'\pi\varepsilon$, we have

(10.2.10) $\tau \cong \tau(w', \mu + \pi\varepsilon' + \eta).$

By genericity, we may assume that p is large enough such that $\lambda_1'' \in C_0$ and that $\mu + \pi\varepsilon'$ lies in the same alcove as μ for any possible λ_1'' and ε'. Then $\mu + \pi\varepsilon' \uparrow \lambda_0'' + \pi\varepsilon''$ for some $\varepsilon'' \in W\varepsilon' = W\varepsilon$. Note that ε'' is a weight of $F(\lambda_1'') = W(\lambda_1'')$. Hence (as in the proof of [Her09, Prop. 9.1]), $F(\lambda_0'' + \pi\varepsilon'')$ is a $G(\mathbb{F}_p)$-constituent of $F(\lambda'') \cong F(\lambda_0'') \otimes F(\lambda_1'')^{(\pi)}$, hence by the above a $G(\mathbb{F}_p)$-constituent of $W(\lambda)$. By Proposition 10.1.13 and (10.2.10), $F(\lambda_0'' + \pi\varepsilon'') \in W^?(\tau)$, hence $F(\lambda_0'' + \pi\varepsilon'') \in W^?(\tau) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ and $d(\lambda_0'' + \pi\varepsilon'') = d(\lambda_0'') < d(\lambda)$, as required. $\qquad\square$

Theorem 10.2.11. *For sufficiently generic tame inertial L-parameters τ we have $W^?(\tau) = \mathrm{W}_{\mathrm{expl}}(\tau) = \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau))$.*

Proof. It suffices to show $\mathrm{W}_{\mathrm{expl}}(\tau) \subset W^?(\tau) \subset \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau))$. To see that $\mathrm{W}_{\mathrm{expl}}(\tau) \subset W^?(\tau)$, first note that $\mathrm{W}_{\mathrm{obv}}(\tau) \subset W^?(\tau)$ by Propositions 9.3.7 and 10.1.13. Suppose there is a Levi $M \subset G$ containing T such that τ factors via $\tau^M : I_{\mathbb{Q}_p} \to \widehat{M}(\overline{\mathbb{F}}_p)$. Then τ^M is as generic as we like by Lemma 10.2.3. It remains to check that if $\nu \in X_1(\underline{T})$ and $w \in W$ are such that $w \cdot \nu \in X(\underline{T})_{+,M}$, then $W^?(\tau^M) \cap \mathrm{JH}_{M(\mathbb{F}_p)}(W^M(w \cdot \nu)) \neq \varnothing$ implies $F(\nu) \in W^?(\tau)$. Noting that $\|w \cdot \nu\|_M \leq \|\nu\|$, we get from Proposition 10.2.8 that $\tau^M \cong \tau^M(w', \lambda' + \eta)$ for some M-dominant $\lambda' \uparrow_M w \cdot \nu$ and some $w' \in W_M$. By Lemma 10.2.4 we have $\sigma \cdot \lambda' \uparrow \nu$, where $\sigma \in W$ such that $\sigma(\lambda' + \eta) \in X(\underline{T})_+$. Hence $\tau \cong \tau(w', \lambda' + \eta) \cong \tau(\sigma w' \pi \sigma^{-1} \pi^{-1}, \sigma \cdot \lambda' + \eta)$. From this we deduce as in Lemma 10.2.1(i) that $\sigma \cdot \lambda' \in X(\underline{T})_+$, as it is as deep as we like in its alcove, hence that $F(\nu) \in W^?(\tau)$ by Proposition 10.1.13.

To show $W^?(\tau) \subset \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau))$, we show

$$F(\lambda) \in W^?(\tau) \implies F(\lambda) \in \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau)) \qquad \text{for all } \lambda \in X_1(\underline{T})$$

by induction on $d(\lambda)$. (Note that $d(\lambda)$ is bounded, independent of p.) As $F(\lambda) \in W^?(\tau)$, Proposition 10.1.13 implies that $\tau \cong \tau(w, \lambda' + \eta)$ for some dominant $\lambda' \uparrow \lambda$ and some $w \in W$. If $\lambda' = \lambda$, then $F(\lambda) \in \mathrm{W}_{\mathrm{obv}}(\tau)$ by Proposition 9.3.7. Otherwise, by Proposition 10.2.8 there exists $\nu \in X_1(\underline{T})$ such that $F(\nu) \in W^?(\tau) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$ and $d(\nu) < d(\lambda)$. By induction, we have $F(\nu) \in \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau)) \cap \mathrm{JH}_{G(\mathbb{F}_p)} W(\lambda)$, hence by definition of $\mathcal{C}$ we get $F(\lambda) \in \mathcal{C}(\mathrm{W}_{\mathrm{obv}}(\tau))$. $\qquad\square$

Remark 10.2.12. In principle the implied constant in this theorem (as well as in all other results in Section 10) can be made explicit. We also remark that none of the results we use depend on Lusztig's conjecture.

10.3. **The proof of Lemma 3.3.5.** In this section we prove Lemma 3.3.5, which we restate here (using once again the notation of Section 3).

Lemma 10.3.1. *If λ is a lift of $a \in (X_1^{(n)})^{S_k}$, then $L_\lambda \otimes_{\mathcal{O}} \mathbb{F}$ has socle F_a, and every other Jordan–Hölder factor of $L_\lambda \otimes_{\mathcal{O}} \mathbb{F}$ is of the form F_b with $b \in (X_1^{(n)})^{S_k}$ and $\|b\| < \|a\|$.*

Proof. It suffices to prove the analogous claim over $\overline{\overline{\mathbb{F}}}_p$. For this we work in the following more general setting: let G denote a connected reductive group over $\mathbb{F}_p$ such that G^{der} is simply connected[4]. We also let B be a Borel subgroup of G with Levi subgroup T, $(\underline{G}, \underline{B}, \underline{T}) := (G, B, T) \times \overline{\overline{\mathbb{F}}}_p$, and let $F : \underline{G} \to \underline{G}$ denote the relative Frobenius. Let π be the finite order automorphism of $(\underline{G}, \underline{B}, \underline{T})$ as in the proof of Lemma 9.2.4; in particular, $F = p\pi^{-1}$ on $X(\underline{T})$. For the moment we work with

---

[4]That is, G is the special fibre of one of the groups that we considered in Section 9, except we don't assume that $Z(G)$ is connected or that G has a local twisting element.

the definition of $\|.\|$ given in Section 10.2, and check at the end of the proof that it agrees with Definition 3.3.4.

We will show that for $a \in X_1(\underline{T})$, $W(a)$ has $\underline{G}^F$-socle $F(a)$, and that every other Jordan–Hölder factor is of the form $F(b)$, $b \in X_1(\underline{T})$, $\|b\| < \|a\|$.

We first leave aside the socle and show by induction on $\|a\|$ that if V is a $\underline{G}$-module with unique highest weight $a \in X_1(\underline{T})$, and $\dim V_a = 1$, then $[V : F(a)]_{\underline{G}^F} = 1$, and every other Jordan–Hölder factor of the $\underline{G}^F$-representation V is of the form $F(b)$, $b \in X_1(\underline{T})$, $\|b\| < \|a\|$.

Any irreducible $\underline{G}$-constituent of V is of the form $F(b)$ with $b \leq a$. Hence it is enough to show that if $[F(b) : F(c)]_{\underline{G}^F} > 0$ $(c \in X_1(\underline{T}))$ then $\|c\| \leq \|b\|$ and that $[F(b) : F(a)]_{\underline{G}^F} = \delta_{ab}$.

If $b \in X_1(\underline{T})$, then $c \equiv b \pmod{X^0(\underline{T})}$ by Lemma 9.2.4 and we are done. Otherwise, $b = b_0 + pb_1$ with $b_0 \in X_1(\underline{T})$ and $b_1 \in X(\underline{T})_+ - X^0(\underline{T})$. Then

$$F(b) \cong F(b_0) \otimes F(pb_1) \cong F(b_0) \otimes F(\pi(b_1))$$

as $\underline{G}^F$-representations, and the latter $\underline{G}$-module has unique highest weight $b_0 + \pi(b_1)$. As $\|b_0 + \pi(b_1)\| = \|b\| - (p-1)\|b_1\| < \|b\|$, we get by induction that $\|c\| \leq \|b_0 + \pi(b_1)\| < \|b\|$.

The claim about the socle follows by dualising the statement of [Hum06, Thm. 5.9]. (In the proof replace σ by any element of $X(\underline{T})$ that pairs to $p-1$ with any simple coroot and $\leq_{\mathbb{Q}}$ by $\|\cdot\| \leq \|\cdot\|$, keeping in mind the above result about Jordan–Hölder factors.)

To deduce the lemma, apply the above with $G = \operatorname{Res}_{k/\mathbb{F}_p}\operatorname{GL}_n$ as in Section 9.5. We have canonical identifications $\underline{G}^F \cong \operatorname{GL}_n(k)$ and $X(\underline{T}) \cong (\mathbb{Z}^n)^{S_k}$. In the notation of Section 3.1 and the proof of Proposition 9.5.3 we get:

$$L_\lambda \otimes \overline{\mathbb{F}}_p \cong \prod_{\sigma \in S_k} \left(M_{a_\sigma} \otimes_{k,\sigma} \overline{\mathbb{F}}_p \right) \cong W(a),$$

$$F_a \otimes \overline{\mathbb{F}}_p \cong \prod_{\sigma \in S_k} \left(N_{a_\sigma} \otimes_{k,\sigma} \overline{\mathbb{F}}_p \right) \cong F(a).$$

To recover Definition 3.3.4, note that $(\sum \alpha^\vee)_\sigma = (n-1, n-3, \ldots, -n+1) \in \mathbb{Z}_+^n$ for any $\sigma \in S_k$.   □

10.4. **Comparison with [ADP02].** Let $\bar{r} : G_{\mathbb{Q}} \to \operatorname{GL}_n(\mathbb{F}_p)$ be odd and irreducible. In this section we prove Proposition 8.4.1, i.e. we show that if $\bar{r}|_{I_{\mathbb{Q}_p}}$ is semisimple and sufficiently generic then the Serre weights predicted in [ADP02] are a subset of $\mathrm{W}_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}})$.

Suppose that $F(\lambda)$, with $\lambda \in X_1^{(n)}$ sufficiently deep in its alcove, is predicted for $\bar{r}$ by [ADP02, Conj. 3.1]. Then according to Definition 2.23 of *loc. cit.*, but using our terminology, there exist integers n_i, an η-partition $(\lambda^{(i)})$ of λ with $\lambda^{(i)} \in \mathbb{Z}_+^{n_i}$, weights $\mu^{(i)} \in X_1^{(n_i)}$, and n_i-cycles $w_i \in S_{n_i}$ such that:

○ $\lambda^{(i)} \equiv \mu^{(i)} \pmod{(p-1)\mathbb{Z}^{n_i}}$ for all i,
○ $\bar{r}|_{I_{\mathbb{Q}_p}} \cong \oplus_i \tau(w_i, \mu^{(i)} + \eta_{n_i})$, where each summand is irreducible, and
○ $\mu_1^{(i)} - \mu_{n_i}^{(i)} \leq p-1$ for all i.

(In fact, in what follows we make no use of the final condition in the above list, nor of the irreducibility of the summands, nor of the fact that w_i is an n_i-cycle.)

Write $\lambda^{(i)} = \mu^{(i)} + (p-1)\nu^{(i)}$ with $\nu^{(i)} \in \mathbb{Z}_+^{n_i}$. Then by (10.1.11) we have

$$\tau(w_i, \mu^{(i)} + \eta_{n_i}) \cong \tau(\sigma w_i \sigma^{-1}, \sigma \cdot ((\lambda^{(i)} - p\nu^{(i)}) + pw_i^{-1}\nu^{(i)}) + \eta_{n_i})$$

for all $\sigma \in S_{n_i}$. By Lemma 10.2.5 we have

$$\sigma_i \cdot ((\lambda^{(i)} - p\nu^{(i)}) + pw_i^{-1}\nu^{(i)}) \uparrow (\lambda^{(i)} - p\nu^{(i)}) + p\nu^{(i)} = \lambda^{(i)},$$

where σ_i is chosen so that the left-hand side of this equation is dominant. Proposition 10.2.8 then gives

$$W^?(\tau(w_i, \mu^{(i)} + \eta_{n_i})) \cap \mathrm{JH}_{\mathrm{GL}_{n_i}(\mathbb{F}_p)} W(\lambda^{(i)}) \neq \varnothing,$$

and so by Proposition 9.5.3, Theorem 10.2.11, and Definition 7.2.3 we deduce that $F(\lambda) \in W_{\mathrm{expl}}(\bar{r}|_{G_{\mathbb{Q}_p}})$.

10.5. **Beyond unramified groups.** It is at present unclear how to formulate versions of the various conjectures of this paper for general ramified groups, where crystalline representations are not available. It seems reasonable to expect that at least for inner forms of GL_n, it should be possible to use the Breuil–Mézard formalism; indeed, this is carried out for quaternion algebras in the papers [GS11, GG15]. For more general groups the absence of a local Langlands correspondence and a mature theory of types at present mean that it is unclear whether to expect the Breuil–Mézard formalism to extend in the necessary fashion.

Appendix A. Wang's result on the $\uparrow$-ordering of alcoves

We give Wang's proof of the following theorem on the geometry of alcoves (see [Ye86], [Wan87]). The following treatment is based on Chuangxun (Allen) Cheng's translation of parts of [Wan87].

A.1. **Ye and Wang's result.** Let $\underline{G}$ denote a connected reductive group over $\overline{\mathbb{F}}_p$, and let $\underline{B}$ be a Borel subgroup of $\underline{G}$ with Levi subgroup $\underline{T}$. We then keep the same notation as in Sections 9–10, for example we have Φ, Φ^+, W_p, $\uparrow$. However, for convenience, in this section C_0 and w_0 do not have their usual meaning.

Theorem A.1.1 (Ye, Wang). *Suppose C, C' are dominant alcoves such that $C \uparrow C'$. Then there exists a sequence of dominant alcoves $C = C_0 \uparrow C_1 \uparrow \cdots \uparrow C_k = C'$ such that $d(C_i) - d(C_{i-1}) = 1$ for all i.*

The proof of this theorem will be discussed below. We first deduce a corollary. Let $\rho := \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha$. We say that $\lambda \in X(\underline{T})$ is ρ-dominant if $\langle \lambda + \rho, \alpha^\vee \rangle \geq 0$ for all $\alpha \in \Delta$.

Corollary A.1.2.
  (i) *Suppose C, C' are dominant alcoves such that $C \uparrow C'$. Then there exists a sequence of dominant alcoves $C = C_0 \uparrow C_1 \uparrow \cdots \uparrow C_k = C'$ and reflections $s_i \in W_p$ such that $s_i \cdot C_{i-1} = C_i$ for all i.*
  (ii) *Suppose λ, $\lambda' \in X(\underline{T})$ are ρ-dominant such that $\lambda \uparrow \lambda'$. Then there exists a sequence of ρ-dominant weights $\lambda = \lambda_0 \uparrow \lambda_1 \uparrow \cdots \uparrow \lambda_k = \lambda'$ and reflections $s_i \in W_p$ such that $s_i \cdot \lambda_{i-1} = \lambda_i$ for all i.*

Proof. Part (i) follows from Theorem A.1.1 and the definition of $\uparrow$, as $d(C) < d(C')$ whenever $C \uparrow C'$ with $C \neq C'$ (and this in fact implies part (ii) in case λ and λ' are p-regular).

For part (ii), let F (resp. F') be the facet containing λ (resp. λ'). Let C be the unique maximal alcove with respect to $\uparrow$ that contains λ, or equivalently F, in its closure. It exists by [Jan03, II.6.11(5)], taking $C = C^+(F)$ in the notation used there. Similarly we let C' be the unique maximal alcove such that $\lambda' \in \overline{C'}$. As λ, λ' are ρ-dominant, we see from [Jan03, II.6.11] that C, C' are dominant alcoves.

We claim that $C \uparrow C'$. An argument exactly as in [Jan03, II.6.11(4)] (reflecting up from C rather than down from $w \cdot C^-$) shows that $C \uparrow C''$ for some alcove C'' such that $\lambda' \in \overline{C''}$, i.e. $F' \subset \overline{C''}$. By the maximality of C' we deduce that $C \uparrow C'' \uparrow C'$.

Applying part (i) we get a sequence of dominant alcoves $C = C_0 \uparrow C_1 \uparrow \cdots \uparrow C_k = C'$ and reflections $s_i \in W_p$ such that $s_i \cdot C_{i-1} = C_i$ for all i. For each i let λ_i be the unique W_p-translate of λ in $\overline{C_i}$. Then $s_i \cdot \lambda_{i-1} = \lambda_i$ for all i, $\lambda_k = \lambda'$, and λ_i is ρ-dominant as C_i is dominant. $\qquad\square$

In the following, let $\mathcal{A}$ denote the set of alcoves and $\mathcal{A}^+$ the subset of dominant alcoves. Let $\mathcal{H}$ denote the set of all hyperplanes

(A.1.3) $$H_{\alpha,np} = \{\lambda \in X(\underline{T}) \otimes \mathbb{R} : \langle \lambda + \rho, \alpha^\vee \rangle = np\}$$

for $\alpha \in \Phi^+$, $n \in \mathbb{Z}$. For each hyperplane $H = H_{\alpha,np} \in \mathcal{H}$, let $s_H \in W_p$ be the reflection in H. It is denoted by $s_{\alpha,np}$ in [Jan03], II.6.1. We will loosely say H is a wall of an alcove C if H contains a facet of C of codimension one.

Given a hyperplane $H = H_{\alpha,np} \in \mathcal{H}$, we let H^- (resp., H^+) denote the half-space obtained by replacing "$=$" by "$<$" (resp., "$>$") in (A.1.3). Recall that $C \uparrow C'$ if there exists a sequence of alcoves $C = C_0, C_1, \ldots, C_k = C'$ such that $C_i = s_{H_i} \cdot C_{i-1}$ and $C_i \subset H_i^+$ for all i ([Jan03], II.6.5). We say that $C \Uparrow C'$ if there is such a sequence satisfying moreover that $-\rho \in \overline{H_i^-}$ for all i. (This was considered, for example, in [And80].) We will see in Corollary A.1.15 that the two partial orders agree on the set of dominant alcoves.

Let C^+ denote the lowest alcove and $D^+ = \{\lambda : \langle \lambda + \rho, \alpha^\vee \rangle > 0$ for all $\alpha \in \Phi^+\}$ the (ρ-shifted) dominant Weyl chamber.

Lemma A.1.4. *If $C \in \mathcal{A}$ and $H \in \mathcal{H}$ then $d(C) \neq d(s_H \cdot C)$, and $C \uparrow s_H \cdot C$ if and only if $d(C) < d(s_H \cdot C)$. In particular if $d(s_H \cdot C) = d(C) + 1$, then $H = H_{\beta,(n_\beta+1)p}$ for some $\beta \in \Phi^+$ and n_β as in (10.2.7) (with ρ replacing η). Therefore there are only finitely many alcoves C' such that $C \uparrow C'$ and $d(C') - d(C) = 1$.*

Proof. This follows easily from Lemma II.6.6 in [Jan03] and its proof. $\qquad\square$

Lemma A.1.5. *Suppose $C \in \mathcal{A}$ and that H is a wall of C. Let $s = s_H$. Suppose that r, $w \in W_p$ and that r is a reflection. If $w \cdot C \uparrow rw \cdot C$ and $rws \cdot C \uparrow ws \cdot C$, then $rw = ws$.*

Proof. Let $H_1 \in \mathcal{H}$ be the hyperplane fixed by r. Then $w \cdot C \subset H_1^-$ and $ws \cdot C \subset H_1^+$. But $w \cdot H$ is the *unique* hyperplane separating alcoves $w \cdot C$ and $ws \cdot C$. Thus $H_1 = w \cdot H$ and $r = s_{w \cdot H} = wsw^{-1}$. $\qquad\square$

Proposition A.1.6. *Suppose that C, H, s are as in Lemma A.1.5. If there exists a sequence $w_0, \ldots, w_h \in W_p$ such that $w_i w_{i-1}^{-1}$ is a reflection for all i and $w_0 \cdot C \uparrow \cdots \uparrow w_h \cdot C$, then one of the following is true:*

(i) $w_0 s \cdot C \uparrow \cdots \uparrow w_h s \cdot C$.

(ii) *There are integers j, k such that $1 \le j \le k \le h$ such that $w_0 s \cdot C \uparrow \cdots \uparrow$ $w_{j-1} s \cdot C = w_j \cdot C \uparrow \cdots \uparrow w_h \cdot C$ and $w_0 \cdot C \uparrow \cdots \uparrow w_{k-1} \cdot C = w_k s \cdot C \uparrow$ $\cdots \uparrow w_h s \cdot C$.*

Proof. Suppose (i) does not hold. Letting j be the minimum and k be the maximum of the non-empty set $\{i : w_i s \cdot C \uparrow w_{i-1} s \cdot C\}$, the proposition easily follows from Lemma A.1.5. $\qquad\square$

Corollary A.1.7. *Suppose that C, H, s are as in Lemma A.1.5. If w, $w' \in W_p$ such that $w \cdot C \uparrow w' \cdot C$ and $w' s \cdot C \uparrow w' \cdot C$, then $w s \cdot C \uparrow w' \cdot C$.*

Proof. We can find $w_0, \ldots, w_h \in W_p$ such that $w_i w_{i-1}^{-1}$ is a reflection for all i and $w \cdot C = w_0 \cdot C \uparrow \cdots \uparrow w_h \cdot C = w' \cdot C$. By the proposition, $w \cdot C \uparrow w' \cdot C$ implies $ws \cdot C \uparrow w's \cdot C$ or $ws \cdot C \uparrow w' \cdot C$. We are done in the second case. In the first case use $w's \cdot C \uparrow w' \cdot C$ to conclude. $\qquad\square$

Lemma A.1.8. *Suppose $C \in \mathcal{A}^+$ or that C has a wall H such that $s_H \cdot C \in \mathcal{A}^+$. Let $r \in W_p$ be a reflection. Then $C \uparrow r \cdot C \iff C \Uparrow r \cdot C$.*

Proof. Clearly if $C \Uparrow r \cdot C$ then $C \uparrow r \cdot C$. Conversely, suppose that $C \uparrow r \cdot C$. Say $r = s_{H_1}$, where $H_1 = H_{\alpha,mp}$ with $\alpha \in \Phi^+$. We have $C \subset H_1^-$ and we want to show that $-\rho \in \overline{H_1^-}$. If $D^+ \cap H_1^- \ne \varnothing$, then for any point x in the intersection, $0 < \langle x + \rho, \alpha^\vee \rangle < mp$, so $-\rho \in H_1^-$ and we are done. If $C \in \mathcal{A}^+$, then $C \subset D^+ \cap H_1^-$ and we are done.

If $C \notin \mathcal{A}^+$, then $s_H \cdot C \in \mathcal{A}^+$ for some wall H of C. If $H \ne H_1$, then C and $s_H \cdot C$ lie on the same side of H_1, so $s_H \cdot C \subset D^+ \cap H_1^-$ and we are done. If $H = H_1$, then $r \cdot C = s_H \cdot C \in \mathcal{A}^+$, so H_1 is a wall of D^+ and thus $-\rho \in H_1 \subset \overline{H_1^-}$. $\qquad\square$

Proposition A.1.9. *Suppose C, $C' \in \mathcal{A}$ with $C' \Uparrow C$. If w, $w' \in W$ such that $w' \cdot C'$ and $w \cdot C$ are dominant, then there exists a sequence of dominant alcoves $w' \cdot C' = C_0 \Uparrow \cdots \Uparrow C_k = w \cdot C$ such that $d(C_i) - d(C_{i-1}) = 1$ for all i.*

Proof. By the definition of $\Uparrow$ we can reduce to the case when $C' = s_{\alpha,np} \cdot C$ for some $\alpha \in \Phi^+$, $-\rho \in \overline{H_{\alpha,np}^-}$, and $C \subset H_{\alpha,np}^+$. Thus $\langle x + \rho, \alpha^\vee \rangle > np \ge 0$ for all $x \in C$. If $\beta = w\alpha$, we see that $\langle y + \rho, \beta^\vee \rangle > np \ge 0$ for all $y \in w \cdot C$. Since $w \cdot C$ is dominant, we have $\beta \in \Phi^+$. Thus $-\rho \in \overline{H_{\beta,np}^-}$ and $w \cdot C \subset H_{\beta,np}^+$. Now we apply [Jan03], II.6.8 to the dominant alcove $w \cdot C$ and the reflection $s_{\beta,np}$ to obtain a sequence of dominant alcoves

$$\text{(A.1.10)} \qquad\qquad w'' s_{\beta,np} w \cdot C = C_0 \uparrow \cdots \uparrow C_k = w \cdot C$$

such that $d(C_i) - d(C_{i-1}) = 1$ for all i, for some $w'' \in W$. Finally notice that $s_{\beta,np} w = w s_{\alpha,np}$ and that $\uparrow$ can be replaced with $\Uparrow$ in (A.1.10) by Lemma A.1.8. $\quad\square$

Note that the translations of $X(\underline{T}) \otimes \mathbb{R}$ that stabilise $\mathcal{H}$ are precisely given by $pX(\underline{T})$. The following lemma is obvious.

Lemma A.1.11. *Suppose $t \in pX(\underline{T})$ and that C, C' are alcoves.*
(i) *$d(C') - d(C) = d(t \cdot C') - d(t \cdot C)$.*
(ii) *$C' \uparrow C \iff t \cdot C' \uparrow t \cdot C$.*

Proposition A.1.12. *Suppose C is an alcove and $n \in \mathbb{Z}_{\ge 0}$. Then*

$$\mathcal{A}(C, n) = \{C' : C \uparrow C' \text{ and } d(C') - d(C) \le n\}$$

is finite. If $t \in pX(\underline{T})$, then $\mathcal{A}(t \cdot C, n) = t \cdot \mathcal{A}(C, n)$.

Proof. By [Jan03], II.6.10 the first claim is reduced to the case $n = 1$, which is covered by Lemma A.1.4. The second claim is immediate from Lemma A.1.11. $\square$

Given $n \in \mathbb{Z}_{\geq 0}$ and an alcove C, we say that C is in *general n-position* if $\mathcal{A}(C, n) \subset \mathcal{A}^+$. The finiteness of $\mathcal{A}(C, n)$ guarantees the existence of alcoves in general n-position.

Lemma A.1.13. *If $n \in \mathbb{Z}_{\geq 0}$ and $C \in \mathcal{A}^+$, then there exists a sequence of dominant alcoves $C = \widetilde{C}_0 \uparrow \cdots \uparrow \widetilde{C}_h$ with $h \in \mathbb{Z}_{\geq 0}$ such that $\widetilde{C}_{i-1}$ and $\widetilde{C}_i$ are adjacent for all i (i.e. there is only one hyperplane between them) and such that $\widetilde{C}_h$ is in general n-position.*

Proof. First consider the case when $C = C^+$. Take $C' = w \cdot C$ in general n-position, where $w \in W_p$. Let S be the set of reflections in the walls of C^+. Pick a reduced expression $w = s_1 \cdots s_r$ $(s_i \in S)$ in the Coxeter group (W_p, S). Letting $C_i := s_1 \cdots s_i \cdot C^+$, it is clear that C_{i-1}, C_i are adjacent for all i. We claim that the C_i are dominant and that $C^+ = C_0 \uparrow \cdots \uparrow C_r = C'$. If $w := s_1 \cdots s_{r-1}$ and $H \in \mathcal{H}$ denotes the hyperplane fixed by $w s_r w^{-1}$, then $\ell(s_H w) > \ell(w)$ implies by [Bou02, Thm. V.3.2.1] that $C_0 = C^+$ and $C_{r-1} = w \cdot C^+$ lie on the same side of H. As H is the common wall of C_{r-1}, C_r, we see that C_0, C_r lie on opposite sides of H. Since $C_0 = C^+$ and C_r are dominant, it follows that $C_0 \subset H^-$, so $C_r \subset H^+$ and $C_{r-1} \uparrow C_r$. As H is not a wall of D^+, we deduce that C_{r-1} is dominant. The claim follows by induction.

If C is general, write $C = w \cdot C^+$ for some $w \in W_p$ and write $w = \sigma + p\nu$ with $\sigma \in W$, $\nu \in X(\underline{T})$. Then it is easy to see that $\langle \nu, \alpha^\vee \rangle > -1$ for all $\alpha \in \Phi^+$, i.e. $\nu \in X(\underline{T})_+$. Hence $t := p\nu$ maps dominant alcoves to dominant alcoves and alcoves in general n-position to alcoves in general n-position. So the lemma is true if $C = t \cdot C^+$.

If $C = w \cdot C^+ \neq t \cdot C^+$, we only have to find a sequence of dominant alcoves $C = \widetilde{C}_0 \uparrow \cdots \uparrow \widetilde{C}_h = t \cdot C^+$ such that $\widetilde{C}_{i-1}$ and $\widetilde{C}_i$ are adjacent for all i. We use an induction on the number of hyperplanes between C and $t \cdot C^+$. Let H be a wall of C that lies between C and $t \cdot C^+$, so H cannot be a wall of D^+. Let $\widetilde{C}_1 = s_H \cdot C = s_H w \cdot C^+$. Then $\widetilde{C}_1 \in \mathcal{A}^+$ and we have $t \cdot (-\rho) = w \cdot (-\rho) \in H$, so $t \cdot (-\rho) = s_H w \cdot (-\rho)$. Since $t \cdot C^+ \subset H^+$, it follows that $C \uparrow \widetilde{C}_1$. Moreover, the number of hyperplanes between $\widetilde{C}_1$ and $t \cdot C^+$ is one less than the number of hyperplanes between C and $t \cdot C^+$. $\square$

Given an alcove C and $n \in \mathbb{Z}_{\geq 0}$, we let $h(C, n)$ be the minimum possible value h occurring as the length of the sequence in Lemma A.1.13.

Proof of Theorem A.1.1. We prove this by induction on $d := d(C') - d(C) \geq 0$. When d is fixed, we induct on $h(C, d)$. The cases $d \leq 1$ are trivial. For any d, the case $h(C, d) = 0$ is trivial. Now for fixed C and d we have a sequence of dominant alcoves

$$C = \widetilde{C}_0 \uparrow \cdots \uparrow \widetilde{C}_h$$

as in Lemma A.1.13 and such that $h = h(C, d)$. If $\widetilde{C}_1 \uparrow C'$ then $d(C') - d(\widetilde{C}_1) = d(C') - d(C) - 1$ and we are done by the induction hypothesis. So we can assume from now on that $\widetilde{C}_1 \not\uparrow C'$. We can write $\widetilde{C}_1 = s_H \cdot C$ and $C' = w \cdot C$ for some wall H of C and some $w \in W_p$. Let $\widetilde{C}'_1 = w \cdot \widetilde{C}_1 = w s_H \cdot C$. We claim that $C' \uparrow \widetilde{C}'_1$.

Otherwise $\widetilde{C}_1' \uparrow C'$. So $C \uparrow w \cdot C$ and $w s_H \cdot C \uparrow w \cdot C$. By Cor. A.1.7 this implies that $\widetilde{C}_1 = s_H \cdot C \uparrow w \cdot C = C'$, a contradiction.

Thus $C' \uparrow \widetilde{C}_1'$, in particular $C \uparrow \widetilde{C}_1'$. We apply Cor. A.1.7 again to $C \uparrow w s_H \cdot C$, $w \cdot C \uparrow w s_H \cdot C$ and get that $\widetilde{C}_1 = s_H \cdot C \uparrow w s_H \cdot C = \widetilde{C}_1'$. Now note that $d(\widetilde{C}_1') - d(\widetilde{C}_1) = d(C') - d(C) = d$, but $h(\widetilde{C}_1, d) = h(C, d) - 1$. By induction hypothesis we have a sequence of dominant alcoves

$$\widetilde{C}_1 = w_0 \cdot \widetilde{C}_1 \uparrow w_1 \cdot \widetilde{C}_1 \uparrow \cdots \uparrow w_d \cdot \widetilde{C}_1 = \widetilde{C}_1'$$

such that $d(w_i \cdot \widetilde{C}_1) - d(w_{i-1} \cdot \widetilde{C}_1) = 1$ and so $w_i w_{i-1}^{-1}$ is a reflection in W_p for all i. Note that $w_0 = 1$ and $w_d = w$. Since $\widetilde{C}_1 \not\uparrow C' = w s_H \cdot \widetilde{C}_1$, by Prop. A.1.6 we have

(A.1.14) $$C = w_0 \cdot C \uparrow w_1 \cdot C \uparrow \cdots \uparrow w_d \cdot C = C'.$$

Since $d = d(C') - d(C)$, we have $d(w_i \cdot C) - d(w_{i-1} \cdot C) = 1$ for all i. As $w_i \cdot C$ and the dominant alcove $w_i \cdot \widetilde{C}_1$ are adjacent, we may replace $\uparrow$ by $\Uparrow$ in (A.1.14) (by Lemma A.1.8). In particular, we have $w_i \cdot C \Uparrow C'$ for all i. If some $w_i \cdot C$ is not dominant, then $w_i \cdot H$ is a wall of D^+ and $w_i s_H w_i^{-1} \in W$. By Prop. A.1.9, $w_i \cdot \widetilde{C}_1 = w_i s_H \cdot C = (w_i s_H w_i^{-1}) w_i \cdot C \Uparrow C'$, so $\widetilde{C}_1 \uparrow w_i \cdot \widetilde{C}_1 \uparrow C'$, contradiction. Thus $w_i \cdot C \in \mathcal{A}^+$ for all i and (A.1.14) satisfies the condition in the theorem. $\square$

Corollary A.1.15. *If C, C' are dominant alcoves, then $C \uparrow C'$ if and only if $C \Uparrow C'$.*

Proof. This follows from Corollary A.1.2 and Lemma A.1.8. $\square$

APPENDIX B. $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ IS NON-EMPTY

The purpose of this appendix is to give a proof of the following result, which was promised in Remark 7.1.4.

Theorem B.1.1. *Suppose $K/\mathbb{Q}_p$ is a finite extension, and let $\overline{\rho}: G_K \to \mathrm{GL}_n(\overline{\mathbb{F}}_p)$ be a representation such that $\overline{\rho}|_{I_K}$ is semisimple. Then the set $\mathrm{W}_{\mathrm{obv}}(\overline{\rho})$ of obvious weights for $\overline{\rho}$ is non-empty.*

A fortiori the same is true for $\mathrm{W}_{\mathrm{expl}}(\overline{\rho})$. Moreover, the proof shows that $\mathrm{W}_{\mathrm{cris}}^\vee(\overline{\rho})$ is non-empty when $\overline{\rho}$ is semisimple.

Proof. For each $\sigma \in S_k$ we fix an element $\kappa_\sigma \in S_K$ lifting σ. Throughout this proof, if we refer to *the* lift of some Serre weight F, we mean any lift λ of F for which $\lambda_\kappa = 0$ if $\kappa \notin \{\kappa_\sigma\}_{\sigma \in S_k}$ (cf. Definition 3.3.2). We will prove that $\overline{\rho}$ has an obvious lift ρ of Hodge type λ, where λ is the lift of some Serre weight.

We may without loss of generality assume that $\overline{\rho}$ itself is semisimple. We begin by explaining how to reduce to the case where $\overline{\rho}$ is irreducible, by induction on the number of Jordan–Hölder factors of $\overline{\rho}$.

Indeed, suppose that $\overline{\rho} = \overline{\rho}' \oplus \overline{\rho}''$, where $\overline{\rho}'$ has dimension $d' > 0$ and $\overline{\rho}''$ is irreducible. By induction $\overline{\rho}'$ has an obvious lift ρ' of Hodge type λ', the lift of some Serre weight. Similarly $\overline{\rho}'' \otimes \overline{\varepsilon}^{-d'}$ has an obvious lift ρ'' of Hodge type λ'', the lift of some Serre weight.

For each $\sigma \in S_k$, let $H_\sigma = \max \mathrm{HT}_{\kappa_\sigma}(\rho')$ and $h_\sigma = \min \mathrm{HT}_{\kappa_\sigma}(\rho'') + d'$. Also let $\Lambda \subset \mathbb{Z}^{S_k}$ be the sublattice consisting of tuples (x_σ) such that $\prod \omega_\sigma^{x_\sigma} = 1$. It is elementary to see that there exists $x = (x_\sigma) \in \Lambda$ such that $h_\sigma + x_\sigma \in [H_\sigma + 1, H_\sigma + p]$ for all $\sigma \in S_k$. (This comes down to the fact that $\mathbb{Z}^{S_k}/\Lambda \cong \mathbb{Z}/(p^f - 1)\mathbb{Z}$, along with

the fact that integers have base p representations.) Let χ be a crystalline character whose Hodge type is the lift of x, and such that $\overline{\chi}$ is trivial; such a character exists by Lemma 5.1.6(i) and (ii). Define $\rho := \rho' \oplus (\rho'' \otimes \varepsilon^{d'} \otimes \chi)$. Then one checks (considering separately the sets $\mathrm{HT}_\kappa(\rho)$ where $\kappa = \kappa_\sigma$ for some $\sigma \in S_k$, and the sets $\mathrm{HT}_\kappa(\rho)$ where $\kappa \notin \{\kappa_\sigma\}_{\sigma \in S_k}$) that ρ is an obvious lift of $\overline{\rho}$ whose Hodge type is the lift of some Serre weight.

It remains to consider the case where $\overline{\rho}$ is irreducible. Let $d = \dim \overline{\rho}$, and write $\overline{\rho} \cong \mathrm{Ind}_{K_d}^K \overline{\psi}$ where K_d/K is the unramified extension of degree d and $\overline{\psi} : G_{K_d} \to \overline{\mathbb{F}}_p^\times$ is a character. We wish to prove the existence of d-tuples of integers $\{(h_{\sigma,0}, \ldots, h_{\sigma,d-1})\}_{\sigma \in S_k}$ such that $0 < h_{\sigma,i} - h_{\sigma,i+1} \leq p$ for all σ and i, and a crystalline character ψ lifting $\overline{\psi}$ such that

$$\bigcup_{\substack{\kappa' \in S_{K_d} \\ \kappa'|_K = \kappa_\sigma}} \mathrm{HT}_{\kappa'}(\psi) = \{h_{\sigma,0}, \ldots, h_{\sigma,d-1}\}$$

for each σ, and such that if $\kappa \in S_K$ but $\kappa \notin \{\kappa_\sigma\}_{\sigma \in S_k}$ then

$$\bigcup_{\substack{\kappa' \in S_{K_d} \\ \kappa'|_K = \kappa}} \mathrm{HT}_{\kappa'}(\psi) = \{0, 1, \ldots, d-1\}.$$

Let χ be any crystalline character of G_{K_d} such that $\mathrm{HT}_{\kappa'}(\chi) = \{0\}$ whenever $\kappa'|_K \in \{\kappa_\sigma\}_{\sigma \in S_k}$, and such that if $\kappa \in S_K$ but $\kappa \notin \{\kappa_\sigma\}_{\sigma \in S_k}$ then

$$\bigcup_{\substack{\kappa' \in S_{K_d} \\ \kappa'|_K = \kappa}} \mathrm{HT}_{\kappa'}(\chi) = \{0, 1, \ldots, d-1\}.$$

Then the theorem comes down to the existence of integers $h_{\sigma,i}$ as above and a crystalline character χ' of G_{K_d} such that

$$\bigcup_{\substack{\kappa' \in S_{K_d} \\ \kappa'|_K = \kappa_\sigma}} \mathrm{HT}_{\kappa'}(\chi') = \{h_{\sigma,0}, \ldots, h_{\sigma,d-1}\}$$

for each $\sigma \in S_k$, such that $\mathrm{HT}_{\kappa'}(\chi') = \{0\}$ if $\kappa'|_K \notin \{\kappa_\sigma\}_{\sigma \in S_k}$, and such that $\overline{\chi'} = \overline{\psi}\overline{\chi}^{-1}$ (for then one can take $\psi = \chi'\chi$).

Unless $(d, f) = (2, 1)$, where $p^f = \#k$, the existence of χ' is an immediate consequence of Proposition B.1.2 below (in combination with both parts of Lemma 5.1.6). When $(d, f) = (2, 1)$, the existence of χ' will follow in the same way provided that the character $\overline{\psi}\overline{\chi}^{-1}$ of G_{K_2} does not extend to G_K. If also $e(K/\mathbb{Q}_p) = 1$ then $\overline{\chi}$ is unramified, and since $\overline{\psi}$ does not extend to G_K, the same is true of $\overline{\psi}\overline{\chi}^{-1}$. If instead $e(K/\mathbb{Q}_p) > 1$, it is possible that $\overline{\psi}\overline{\chi}^{-1}$ extends to G_K. In that case choose any $\kappa' \in S_{K_2}$ such that $\mathrm{HT}_{\kappa'}(\chi) = \{1\}$. Let $\kappa'' \in S_{K_2}$ be the other embedding such that $\kappa''|_K = \kappa'|_K$ (so that $\mathrm{HT}_{\kappa''}(\chi) = \{0\}$). Let χ_0 be a crystalline character of G_{K_2} with the same labeled Hodge–Tate weights as χ, except that $\mathrm{HT}_{\kappa'}(\chi_0) = \{0\}$ and $\mathrm{HT}_{\kappa''}(\chi_0) = \{1\}$. We note that $\overline{\chi\chi_0^{-1}}|_{I_{K_2}} = \omega_{\overline{\kappa}''}^{p-1}$. Hence $\overline{\chi\chi_0^{-1}}$ does not extend to G_K (since $\omega_{\overline{\kappa}''}$ is a fundamental character of niveau 2 and its exponent is not a multiple of $p+1$); so neither does the character $\overline{\psi}\overline{\chi}_0^{-1}$, and the result follows from Proposition B.1.2 using $\overline{\psi}\overline{\chi}_0^{-1}$ in place of $\overline{\psi}\overline{\chi}^{-1}$. $\qquad\square$

Proposition B.1.2. *Given positive integers d and f, any residue class modulo $p^{df} - 1$ (with the exception of the residue classes congruent to 0 modulo $p+1$ when*

$d = 2$, $f = 1$) *is of the form* $\sum_{i=0}^{df-1} x_i p^i$, *where for any* $i_0 \in \mathbb{Z}$ *the set* $\{x_i : i \equiv i_0 \pmod{f}\}$ *is of the form* $\{h_0, \ldots, h_{d-1}\}$ *with* $0 < h_i - h_{i+1} \le p$ *for all* i.

Proof. Let N be the representative in the interval $[0, p^{df} - 1)$ of our given residue class modulo $p^{df} - 1$, and let $x_0, \ldots, x_{df-1}$ be the digits in the base p expansion of N, so that certainly $N \equiv \sum_{i=0}^{df-1} x_i p^i \pmod{p^{df} - 1}$. We will argue by altering the x_i's, preserving this congruence, until the condition on the sets $\{x_i : i \equiv i_0 \pmod{f}\}$ is met. The typical alteration will be to add δp to x_i and $-\delta$ to x_{i+1} (with x_{df} taken to mean x_0). We break into cases depending on the value of f.

(1) We consider first the case where f is even. Restrict our attention to the x_j's with $j \equiv 0, 1 \pmod{f}$. Relabel the pairs $\{(x_{if}, x_{if+1}) : i = 0, \ldots, d-1\}$ as pairs $(a_0, b_0), \ldots, (a_{d-1}, b_{d-1})$, but *not necessarily in the same order*; instead, we choose the labeling so that

    (i) $b_0 \ge \cdots \ge b_{d-1}$, and
    (ii) if $b_i = b_{i+1}$ then $a_{i+1} \le a_i$.

There exist integers $\delta_i \in \mathbb{Z}$ such that

(B.1.3) $$(a_{i+1} + \delta_{i+1} p) - (a_i + \delta_i p) \in (0, p]$$

for all $i \in [0, d-1)$. As $a_{i+1} - a_i \in (-p, p)$, we have $\delta_{i+1} - \delta_i \in [0, 1]$. For each i we define $(a_i', b_i') = (a_i + \delta_i p, b_i - \delta_i)$, thereby also altering the corresponding x's. We claim that $a_{i+1}' - a_i'$ and $b_i' - b_{i+1}'$ both lie in $(0, p]$ for all i. The first of these claims is precisely (B.1.3). For the second claim, write

$$b_i' - b_{i+1}' = (b_i - b_{i+1}) + (\delta_{i+1} - \delta_i),$$

and observe that $b_i - b_{i+1} \in [0, p-1]$ by (i), while $\delta_{i+1} - \delta_i \in [0, 1]$. It remains to note that if $b_i = b_{i+1}$ then $a_{i+1} \le a_i$ by (ii), implying $\delta_{i+1} - \delta_i = 1$ by (B.1.3), and so $b_i' - b_{i+1}' > 0$ in all cases.

The two claims together show that after making these alterations, the sets $\{x_j : j \equiv j_0 \pmod{f}\}$ for $j_0 = 0, 1$ are both of the desired form $\{h_0, \ldots, h_{d-1}\}$ with $0 < h_i - h_{i+1} \le p$ for all i. Iterating the above procedure for the x_j's with $j \equiv 2j_0, 2j_0 + 1 \pmod{f}\}$ for each $j_0 \in [1, f/2)$ in turn, the proposition follows in this case.

(2) Next we suppose that f is odd and $f \ge 3$. It is enough to explain how to alter the triples $\{(x_{if}, x_{if+1}, x_{if+2}) : i = 0, \ldots, d-1\}$, for then we can deal with the remaining consecutive pairs of residue classes as in the case where f was even. The truth of the proposition is certainly unchanged under multiplication of the given residue class by a power of p, or equivalently, under cyclic permutation of the x_j's. We observe (trivially) that it is possible to cyclically permute the x_j's so that it is *not* the case that the pairs (x_{if+1}, x_{if+2}) are all of the form $(p-1, p-1)$ or $(0, 0)$, with both occurring, and we make such a cyclic permutation.

Now rewrite the triples $(x_{if}, x_{if+1}, x_{if+2})$ as $(a_0, b_0, c_0), \ldots, (a_{d-1}, b_{d-1}, c_{d-1})$ with the labeling so that

    (i) $c_0 \ge \cdots \ge c_{d-1}$, and
    (ii) if $c_i = c_{i+1}$ then $b_{i+1} \le b_i$.

Conditions (i) and (ii), together with the condition on the pairs (x_{if+1}, x_{if+2}) from the previous paragraph, imply

(B.1.4)   there is no value of i such that $(b_i - b_{i+1}, c_i - c_{i+1}) = (p-1, p-1)$.

There exist $\delta_i \in \mathbb{Z}$ such that

(B.1.5) $$(a_{i+1} + \delta_{i+1}p) - (a_i + \delta_i p) \in (0, p]$$

for all $i \in [0, d-1)$, and also $\epsilon_i \in \mathbb{Z}$ such that

(B.1.6) $$\lambda_i := (b_{i+1} + \epsilon_{i+1}p) - (b_i + \epsilon_i p) \in (0, p+1]$$

for all $i \in [0, d-1)$, with

(B.1.7) $$\begin{cases} \lambda_i = p+1 & \implies a_{i+1} \leq a_i, \\ \lambda_i = 1 & \implies a_{i+1} > a_i. \end{cases}$$

As in case (1) we have $\delta_{i+1} - \delta_i \in [0, 1]$ for all i; similarly we have $\epsilon_{i+1} - \epsilon_i \in [0, 2]$, with

(B.1.8) $$\epsilon_{i+1} - \epsilon_i = 2 \quad \text{iff} \quad (b_i, b_{i+1}) = (p-1, 0) \text{ and } a_{i+1} \leq a_i.$$

For each i we define

$$(a_i', b_i', c_i') = (a_i + \delta_i p, b_i + \epsilon_i p - \delta_i, c_i - \epsilon_i),$$

thereby altering the corresponding x's, and claim that we then have $a_{i+1}' - a_i'$, $b_{i+1}' - b_i'$, $c_i' - c_{i+1}'$ in $(0, p]$ for all i. Case (2) will be complete once we have proved this claim.

That $a_{i+1}' - a_i' \in (0, p]$ is immediate from (B.1.5). Next, we have

$$b_{i+1}' - b_i' = \lambda_i - (\delta_{i+1} - \delta_i)$$

with the first term on the right-hand side in $(0, p+1]$ and the second term in $[0, 1]$. If $\lambda_i = p+1$ then $a_{i+1} \leq a_i$ by (B.1.7), which implies $\delta_{i+1} - \delta_i = 1$ by (B.1.5); similarly if $\lambda_i = 1$ then $a_{i+1} > a_i$ and $\delta_{i+1} - \delta_i = 0$. Thus in all cases we have $b_{i+1}' - b_i' \in (0, p]$, as desired.

Finally,

$$c_i' - c_{i+1}' = (c_i - c_{i+1}) + (\epsilon_{i+1} - \epsilon_i)$$

with the first term on the right-hand side in $[0, p-1]$ and the second term in $[0, 2]$. If $c_i = c_{i+1}$ then $b_{i+1} \leq b_i$, implying $\epsilon_{i+1} - \epsilon_i > 0$; thus $c_i' - c_{i+1}' > 0$ in all cases. Suppose on the other hand that $c_i - c_{i+1} = p - 1$. Then by (B.1.4) we have $b_i - b_{i+1} \neq p - 1$, and so $\epsilon_{i+1} - \epsilon_i \neq 2$ by (B.1.8); thus $c_i' - c_{i+1}' \leq p$ in all cases, and case (2) is complete.

(3) Finally we turn to the case $f = 1$. As usual, we take the x_i's at the outset to be the digits in the base p expansion of N. As in part (2) we will make use of the fact that the truth of the proposition is unchanged when multiplying the given residue class by a power of p (i.e. under cyclic permutation of the x_i's), as well as when adding any multiple of $(p^d - 1)/(p - 1)$ to the residue class (i.e. adding the same constant to each x_i).

We first dispense with the case where d is even and N is divisible by $(p^d - 1)/(p - 1)$. By hypothesis we have $d \geq 4$ (recall that in the case $d = 2$, $f = 1$ the residue classes divisible by $p+1$ are excluded from the statement of the proposition). Subtracting the appropriate multiple of $(p^d-1)/(p-1)$ we may suppose that $N = 0$. Then writing $d = 2m + 2$ with $m \geq 1$, we alter the x_i's by replacing them with

$$(x_0', \ldots, x_{d-1}') := (p, 2p-1, p-2, -1, 2p, -2, \ldots, mp, -m).$$

In the remaining cases, we can reduce to one of the following three situations.

(I) d is odd, each x_i lies in $[0, p-1]$, and $x_{d-1} = \max_i x_i$.

(II) d is even, each x_i lies in $[0, p-1]$, $x_1 = \max_i x_i > 0$, and $x_i = 0$ for all even i.

(III) d is even, each x_i lies in $[0, p-1]$ except $x_1 = p$, and x_i is non-zero for some odd $i > 1$.

To see this, argue as follows. If d is odd, cyclically permute to assume that x_{d-1} is maximal to put ourselves in case (I). Now suppose d is even, so that $(p^d-1)/(p-1) \nmid N$ and not all the x_i's are equal. Subtracting $\min_i x_i$ from each x_i, we can further suppose that some x_i is 0. Since not all x_i's are zero, we can suppose (after cyclically permuting if necessary) that $x_1 = 0$ and $x_2 > 0$. If x_i is non-zero for some odd $i > 1$, then we add p to x_1 and -1 to x_2 to put ourselves in case (III). Otherwise $x_i = 0$ for all odd i but not for all even i. After cyclically permuting so that $x_1 = \max_i x_i$, we have $x_{2j} = 0$ for all j and are in case (II). This completes the reduction.

Write $d = 2m$ if d is even and $d = 2m+1$ if d is odd. We relabel the pairs of variables $(x_0, x_1), \ldots, (x_{2m-2}, x_{2m-1})$ as $(a_0, b_0), \ldots, (a_{m-1}, b_{m-1})$, ordered as usual so that

(i) $b_0 \geq \cdots \geq b_{m-1}$ and
(ii) if $b_i = b_{i+1}$ then $a_{i+1} \leq a_i$.

Note that when d is odd, x_{d-1} is not relabeled. When d is even, we can (and do) take $(a_0, b_0) = (x_0, x_1)$: in case (II) this is a consequence of the fact that $a_i = 0$ for all i, whereas in case (III) it is automatic.

There exist unique integers $\delta_i \in \mathbb{Z}$ such that $\delta_0 = 1$ and

(B.1.9) $$(a_{i+1} + \delta_{i+1} p) - (a_i + \delta_i p) \in (0, p]$$

for all $i \in [0, m-1)$. For each i we define $(a_i', b_i') = (a_i + \delta_i p, b_i - \delta_i)$, thereby altering the corresponding x's. (Note when d is odd that x_{d-1} is unchanged.) It follows almost exactly as in (1) that we have $a_{i+1}' - a_i'$, $b_i' - b_{i+1}' \in (0, p]$ for all i; the only modification required is to note that in case (III), although $b_0 = p$ we still have $b_0 - b_1 \in [0, p-1]$ because of the condition that x_i is non-zero for some odd $i > 1$.

To complete the proof, it will suffice to show that

$$\begin{cases} a_0' - b_0' \in (0, p] & \text{if } d \text{ is even,} \\ a_0' - x_{d-1}, \ x_{d-1} - b_0' \in (0, p] & \text{if } d \text{ is odd.} \end{cases}$$

First suppose that d is even. Since $\delta_0 = 1$ we have

$$a_0' - b_0' = (x_0 - x_1) + (p+1).$$

In case (II) we have $x_0 - x_1 \in (-p, 0)$, while in case (III) we have $x_0 - x_1 \in [-p, 0)$; in either case $a_0' - b_0' \in (0, p]$.

Finally suppose that d is odd. We have $a_0' = a_0 + p$ and $b_0' = b_0 - 1$. Since $a_0 \leq x_{d-1}$ and both are in the range $[0, p-1]$, we have $a_0' - x_{d-1} = p - (x_{d-1} - a_0) \in (0, p]$. Similarly $b_0 \leq x_{d-1}$ and both are in the range $[0, p-1]$, so that $x_{d-1} - b_0' = (x_{d-1} - b_0) + 1 \in (0, p]$. This completes the proof. $\square$

References

[ADP02] Avner Ash, Darrin Doud, and David Pollack, *Galois representations with conjectural connections to arithmetic cohomology*, Duke Math. J. **112** (2002), no. 3, 521–579.

[And80] Henning Haahr Andersen, *The strong linkage principle*, J. Reine Angew. Math. **315** (1980), 53–59.

[APS04]   Avner Ash, David Pollack, and Dayna Soares, $SL_3(\mathbb{F}_2)$-*extensions of* $\mathbb{Q}$ *and arithmetic cohomology modulo 2*, Experiment. Math. **13** (2004), no. 3, 298–307.

[AS86]   Avner Ash and Glenn Stevens, *Modular forms in characteristic l and special values of their L-functions*, Duke Math. J. **53** (1986), no. 3, 849–868.

[AS00]   Avner Ash and Warren Sinnott, *An analogue of Serre's conjecture for Galois representations and Hecke eigenclasses in the mod p cohomology of* $GL(n, \mathbf{Z})$, Duke Math. J. **105** (2000), no. 1, 1–24.

[BD14]   Christophe Breuil and Fred Diamond, *Formes modulaires de Hilbert modulo p et valeurs d'extensions entre caractères galoisiens*, Ann. Sci. Éc. Norm. Supér. (4) **47** (2014), no. 5, 905–974.

[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.

[Ber10]   Laurent Berger, *Représentations modulaires de* $GL_2(\mathbf{Q}_p)$ *et représentations galoisiennes de dimension 2*, Astérisque (2010), no. 330, 263–279.

[BG15]   Kevin Buzzard and Toby Gee, *The conjectural connections between automorphic representations and Galois representations*, Automorphic Forms and Galois Representations, London Math. Soc. Lecture Note Ser. 414 (2015), 135–187.

[BLGG13]   Thomas Barnet-Lamb, Toby Gee, and David Geraghty, *Serre weights for rank two unitary groups*, Math. Ann. **356** (2013), no. 4, 1551–1598.

[BLGG14]   _____, *Serre weights for* $U(n)$, J. Reine Angew. Math. (to appear), 2014.

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

[BM02]   Christophe Breuil and Ariane Mézard, *Multiplicités modulaires et représentations de* $GL_2(\mathbf{Z}_p)$ *et de* $Gal(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ *en* $l = p$, Duke Math. J. **115** (2002), no. 2, 205–310, With an appendix by Guy Henniart.

[BM14]   _____, *Multiplicités modulaires raffinées*, Bull. Soc. Math. France **142** (2014), no. 1, 127–175.

[Bou02]   Nicolas Bourbaki, *Lie groups and Lie algebras. Chapters 4–6*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 2002, Translated from the 1968 French original by Andrew Pressley.

[Bre10]   Christophe Breuil, *The emerging p-adic Langlands programme*, Proceedings of the International Congress of Mathematicians. Volume II, Hindustan Book Agency, New Delhi, 2010, pp. 203–230.

[BV13]   Nicolas Bergeron and Akshay Venkatesh, *The asymptotic growth of torsion homology for arithmetic groups*, J. Inst. Math. Jussieu **12** (2013), no. 2, 391–447.

[Cal12]   Frank Calegari, *Even Galois Representations and the Fontaine-Mazur conjecture II*, J. Amer. Math. Soc. **25** (2012), no. 2, 533–554.

[CE12]   Frank Calegari and Matthew Emerton, *Completed cohomology—a survey*, Non-abelian fundamental groups and Iwasawa theory, London Math. Soc. Lecture Note Ser., vol. 393, Cambridge Univ. Press, Cambridge, 2012, pp. 239–257.

[CEG+16]   Ana Caraiani, Matthew Emerton, Toby Gee, David Geraghty, Vytautas Paˇskūnas, and Sug Woo Shin, *Patching and the p-adic local Langlands correspondence*, Camb. J. Math. **4** (2016), no. 2, 197–287.

[CEGM17]   Frank Calegari, Matthew Emerton, Toby Gee, and Lambros Mavrides, *Explicit Serre weights for two-dimensional Galois representations*, Compositio Mathematica **153** (2017), no. 9, 1893–1907.

[CEGS]   Ana Caraiani, Matthew Emerton, Toby Gee, and David Savitt, *Moduli spaces of Kisin modules with descent data*, in preparation.

[CHT08]   Laurent Clozel, Michael Harris, and Richard Taylor, *Automorphy for some l-adic lifts of automorphic mod l Galois representations*, Publ. Math. Inst. Hautes Études Sci. (2008), no. 108, 1–181, With Appendix A, summarizing unpublished work of Russ Mann, and Appendix B by Marie-France Vignéras.

[Col10]   Pierre Colmez, *Représentations de* $GL_2(\mathbf{Q}_p)$ *et* (ϕ, Γ)-*modules*, Astérisque (2010), no. 330, 281–509.

[Con11]   Brian Conrad, *Lifting global representations with local properties*, preprint, 2011.

[DDR16]   Lassina Dembélé, Fred Diamond, and David P. Roberts, *Serre weights and wild ramification in two-dimensional Galois representations*, Forum Math. Sigma **4** (2016), e33, 49.

[DL76]   P. Deligne and G. Lusztig, *Representations of reductive groups over finite fields*, Ann. of Math. (2) **103** (1976), no. 1, 103–161.

[Dou07]   Darrin Doud, *Supersingular Galois representations and a generalization of a conjecture of Serre*, Experiment. Math. **16** (2007), no. 1, 119–128.

[DS87]   Stephen R. Doty and John B. Sullivan, *Filtration patterns for representations of algebraic groups and their Frobenius kernels*, Math. Z. **195** (1987), no. 3, 391–407.

[EG]   Matthew Emerton and Toby Gee, *Moduli spaces of p-adic and mod p representations of Galois groups of p-adic local fields*, in preparation.

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

[EG15]   M. Emerton and T. Gee, *"Scheme-theoretic images" of morphisms of stacks*, ArXiv e-prints (2015), 116 pages.

[EGH13]   Matthew Emerton, Toby Gee, and Florian Herzig, *Weight cycling and Serre-type conjectures for unitary groups*, Duke Math. J. **162** (2013), no. 9, 1649–1722.

[EGS15]   Matthew Emerton, Toby Gee, and David Savitt, *Lattices in the cohomology of Shimura curves*, Invent. Math. **200** (2015), no. 1, 1–96.

[Eme10]   Matthew Emerton, *Local-global compatibility in the p-adic Langlands programme for* $GL_2 / \mathbb{Q}$, preprint, 2010.

[Eme14]   _____, *Completed cohomology and the p-adic Langlands program*, Proceedings of the 2014 ICM, Volume II (2014), 319–342.

[Fie12]   Peter Fiebig, *An upper bound on the exceptional characteristics for Lusztig's character formula*, J. Reine Angew. Math. **673** (2012), 1–31.

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

[Gee11]   _____, *Automorphic lifts of prescribed types*, Math. Ann. **350** (2011), no. 1, 107–144.

[GG12]   Toby Gee and David Geraghty, *Companion forms for unitary and symplectic groups*, Duke Math. J. **161** (2012), no. 2, 247–303.

[GG15]   _____, *The Breuil-Mézard conjecture for quaternion algebras*, Ann. Inst. Fourier (Grenoble) **65** (2015), no. 4, 1557–1575.

[GHLS15]   T. Gee, F. Herzig, T. Liu, and D. Savitt, *Potentially crystalline lifts of certain prescribed types*, ArXiv e-prints (2015), 22 pages.

[GK14]   Toby Gee and Mark Kisin, *The Breuil–Mézard conjecture for potentially Barsotti–Tate representations*, Forum Math. Pi **2** (2014), e1 (56 pages).

[GLS14]   Toby Gee, Tong Liu, and David Savitt, *The Buzzard-Diamond-Jarvis conjecture for unitary groups*, J. Amer. Math. Soc. **27** (2014), no. 2, 389–435.

[GLS15]   _____, *The weight part of Serre's conjecture for* GL(2), Forum Math. Pi **3** (2015), e2, 52.

[Gro98a]   Benedict H. Gross, *Modular forms* (mod p) *and Galois representations*, Internat. Math. Res. Notices (1998), no. 16, 865–875.

[Gro98b]   _____, *On the Satake isomorphism*, Galois representations in arithmetic algebraic geometry (Durham, 1996), London Math. Soc. Lecture Note Ser., vol. 254, Cambridge Univ. Press, Cambridge, 1998, pp. 223–237.

[Gro99]   _____, *Algebraic modular forms*, Israel J. Math. **113** (1999), 61–93.

[Gro07]   _____, *Odd Galois representations*, preprint, 2007.

[GS11]   Toby Gee and David Savitt, *Serre weights for quaternion algebras*, Compos. Math. **147** (2011), no. 4, 1059–1086.

[Her09]   Florian Herzig, *The weight in a Serre-type conjecture for tame n-dimensional Galois representations*, Duke Math. J. **149** (2009), no. 1, 37–116.

[HLM17]   Florian Herzig, Daniel Le, and Stefano Morra, *On mod p local-global compatibility for* GL_3 *in the ordinary case*, Compos. Math. (to appear), 2017.

[HT01]   Michael Harris and Richard Taylor, *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies, vol. 151, Princeton University Press, Princeton, NJ, 2001, With an appendix by Vladimir G. Berkovich.

[HT13]   Florian Herzig and Jacques Tilouine, *Conjecture de type de Serre et formes compagnons pour* GSp_4, J. Reine Angew. Math. **676** (2013), 1–32.

[HT15]   Yongquan Hu and Fucheng Tan, *The Breuil-Mézard conjecture for non-scalar split residual representations*, Ann. Sci. Éc. Norm. Supér. (4) **48** (2015), no. 6, 1383–1421.

[Hum06]   James E. Humphreys, *Modular representations of finite groups of Lie type*, London Mathematical Society Lecture Note Series, vol. 326, Cambridge University Press, Cambridge, 2006.

[Jan81]   Jens Carsten Jantzen, *Zur Reduktion modulo p der Charaktere von Deligne und Lusztig*, J. Algebra **70** (1981), no. 2, 452–474.

[Jan87]   J. C. Jantzen, *Representations of Chevalley groups in their own characteristic*, The Arcata Conference on Representations of Finite Groups (Arcata, Calif., 1986), Proc. Sympos. Pure Math., vol. 47, Amer. Math. Soc., Providence, RI, 1987, pp. 127–146.

[Jan03]   Jens Carsten Jantzen, *Representations of algebraic groups*, second ed., Mathematical Surveys and Monographs, vol. 107, American Mathematical Society, Providence, RI, 2003.

[Kis08]   Mark Kisin, *Potentially semi-stable deformation rings*, J. Amer. Math. Soc. **21** (2008), no. 2, 513–546.

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

[Kis09b]   _____, *Modularity of 2-adic Barsotti-Tate representations*, Invent. Math. **178** (2009), no. 3, 587–634.

[Kis09c]   _____, *Moduli of finite flat group schemes, and modularity*, Annals of Math.(2) **170** (2009), no. 3, 1085–1180.

[Kis10]   _____, *The structure of potentially semi-stable deformation rings*, Proceedings of the International Congress of Mathematicians. Volume II (New Delhi), Hindustan Book Agency, 2010, pp. 294–311.

[Kot84]   Robert E. Kottwitz, *Stable trace formula: cuspidal tempered terms*, Duke Math. J. **51** (1984), no. 3, 611–650.

[KW09a]   Chandrashekhar Khare and Jean-Pierre Wintenberger, *Serre's modularity conjecture. I*, Invent. Math. **178** (2009), no. 3, 485–504.

[KW09b]   _____, *Serre's modularity conjecture. II*, Invent. Math. **178** (2009), no. 3, 505–586.

[LLHL16]   Daniel Le, Bao V. Le Hung, and Brandon Levin, *Weight elimination in Serre type conjectures*, preprint, 2016.

[LLHLM15]   Daniel Le, Bao V. Le Hung, Brandon Levin, and Stefano Morra, *Potentially crystalline deformation rings and Serre weight conjectures: shapes and shadows*, preprint, 2015.

[LLHLM16]   _____, *Serre weight conjectures and Breuil's lattice conjecture in dimension three*, preprint, 2016.

[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.

[MP17]   Stefano Morra and Chol Park, *Serre weights for three dimensional ordinary Galois representations*, J. Lond. Math. Soc. (to appear), 2017.

[Mul13]   Alain Muller, *Relèvements cristallins de représentations galoisiennes*, Université de Strasbourg Ph.D. thesis, 2013.

[Paš15]   Vytautas Paškūnas, *On the Breuil–Mézard conjecture*, Duke Math. J. **164** (2015), no. 2, 297–359.

[Paš16]   _____, *On 2-dimensional 2-adic Galois representations of local and global fields*, Algebra Number Theory **10** (2016), no. 6, 1301–1358.

[San14]   Fabian Sander, *Hilbert-Samuel multiplicities of certain deformation rings*, Math. Res. Lett. **21** (2014), no. 3, 605–615.

[San16]   _____, *A local proof of the Breuil-Mézard conjecture in the scalar semi-simplification case*, J. Lond. Math. Soc. (2) **94** (2016), no. 2, 447–461.

[Sch08]   Michael M. Schein, *Weights in Serre's conjecture for Hilbert modular forms: the ramified case*, Israel J. Math. **166** (2008), 369–391.

[SD73]   H. P. F. Swinnerton-Dyer, *On l-adic representations and congruences for coefficients of modular forms*, Modular functions of one variable, III (Proc. Internat. Summer

School, Univ. Antwerp, 1972), Springer, Berlin, 1973, pp. 1–55. Lecture Notes in Math., Vol. 350.

[Ser75]   Jean-Pierre Serre, *Valeurs propres des opérateurs de Hecke modulo l*, Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, 1974), Soc. Math. France, Paris, 1975, pp. 109–117. Astérisque, Nos. 24–25.

[Ser79]   ———, *Groupes algébriques associés aux modules de Hodge-Tate*, Journées de Géométrie Algébrique de Rennes. (Rennes, 1978), Vol. III, Astérisque, vol. 65, Soc. Math. France, Paris, 1979, pp. 155–188.

[Ser86]   ———, *Œuvres. Vol. III*, Springer-Verlag, Berlin, 1986, 1972–1984.

[Ser87]   ———, *Sur les représentations modulaires de degré 2 de* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. **54** (1987), no. 1, 179–230.

[Shi11]   Sug Woo Shin, *Galois representations arising from some compact Shimura varieties*, Ann. of Math. (2) **173** (2011), no. 3, 1645–1741.

[Sti10]   Jakob Stix, *Trading degree for dimension in the section conjecture: the non-abelian Shapiro lemma*, Math. J. Okayama Univ. **52** (2010), 29–43.

[SZ99]   P. Schneider and E.-W. Zink, *K-types for the tempered components of a p-adic general linear group*, J. Reine Angew. Math. **517** (1999), 161–208, With an appendix by Schneider and U. Stuhler.

[Tat94]   John Tate, *The non-existence of certain Galois extensions of* **Q** *unramified outside* 2, Arithmetic geometry (Tempe, AZ, 1993), Contemp. Math., vol. 174, Amer. Math. Soc., Providence, RI, 1994, pp. 153–156.

[Tor12]   Rebecca Torrey, *On Serre's conjecture over imaginary quadratic fields*, J. Number Theory **132** (2012), no. 4, 637–656.

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

[Wan87]   Jian Pan Wang, *Partial orderings on affine Weyl groups*, J. East China Norm. Univ. Natur. Sci. Ed. (1987), no. 4, 15–25.

[Ye86]   Jia Chen Ye, *A theorem on the geometry of alcoves*, Tongji Daxue Xuebao **14** (1986), no. 1, 57–64.

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON
E-mail address: `toby.gee@imperial.ac.uk`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO
E-mail address: `herzig@math.toronto.edu`

DEPARTMENT OF MATHEMATICS, JOHNS HOPKINS UNIVERSITY
E-mail address: `savitt@math.jhu.edu`