

## ON THE COHOMOLOGY OF COMPACT UNITARY GROUP SHIMURA VARIETIES AT RAMIFIED SPLIT PLACES

PETER SCHOLZE AND SUG WOO SHIN

### CONTENTS

1.	Introduction	261
2.	Kottwitz triples	267
3.	Endoscopic transfer of test functions	268
4.	Stabilization - proof of Theorem 1.5	271
5.	Pseudostabilization	273
6.	The stable Bernstein center	276
7.	Character identities satisfied by $f_{\tau,h}^{\mathbf{H}}$ : Conjectures	278
8.	Character identities satisfied by $f_{\tau,h}^{\mathbf{H}}$ : Proofs	281
9.	Cohomology of Shimura varieties: Nonendoscopic case	286
10.	Cohomology of Shimura varieties: Endoscopic case	287
11.	Construction of Galois representations	291
	Acknowledgments	292
	References	292

### 1. INTRODUCTION

This paper deals with the problem of determining the Galois action on the cohomology of Shimura varieties, specifically at places of bad reduction. Let us first briefly recall the expected description, due to Langlands.

Let  $\mathrm{Sh}_K$  be some Shimura variety associated to a reductive group  $\mathbf{G}$  over  $\mathbb{Q}$  and a compact open subgroup  $K \subset \mathbf{G}(\mathbb{A}_f)$ , and some additional data. It is defined over a number field  $E$ , canonically embedded into  $\mathbb{C}$ . For definiteness, let us assume that the Shimura variety is compact, as we will only deal with this case; it is equivalent to requiring that  $\mathbf{G}$  is anisotropic modulo its center.<sup>1</sup> Then one considers the étale cohomology groups

$$H^i = \varinjlim_K H_{\mathrm{et}}^i(\mathrm{Sh}_K \otimes_E \bar{\mathbb{Q}}, \bar{\mathbb{Q}}_\ell),$$

---

Received by the editors November 9, 2011 and, in revised form, July 23, 2012.

2010 *Mathematics Subject Classification*. Primary 11F70, 11F80, 11G18, 11R39, 11S37; Secondary 14G35, 11F72, 22E50, 22E55.

This work was written while the first author was a Clay Research Fellow.

The second author's work was supported by the National Science Foundation during his stay at the Institute for Advanced Study under agreement No. DMS-0635607. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

<sup>1</sup>In the general case, one considers the intersection cohomology groups of the Baily-Borel compactification, and similar results are expected; cf. [35].

which carry an action of  $\mathbf{G}(\mathbb{A}_f)$  (via Hecke correspondences) and of  $\text{Gal}(\overline{\mathbb{Q}}/E)$ . In the main body of the paper we consider the cohomology groups with coefficients in a local system associated to an algebraic representation  $\xi$  of  $\mathbf{G}$ , where everything works without essential change. Let  $H^*$  be the alternating sum of the  $H^i$  in a suitable Grothendieck group; then one can write

$$H^* = \sum_{\pi_f} \pi_f \otimes \sigma(\pi_f) ,$$

where  $\pi_f$  runs through irreducible admissible representations of  $\mathbf{G}(\mathbb{A}_f)$ , and  $\sigma(\pi_f)$  is some virtual finite-dimensional representation of  $\text{Gal}(\overline{\mathbb{Q}}/E)$ .

If one forgets about the action of  $\text{Gal}(\overline{\mathbb{Q}}/E)$  and is only interested in describing the  $\mathbf{G}(\mathbb{A}_f)$ -action, then one can base-change to  $\mathbb{C}$  and use Matsushima’s formula, which computes the cohomology as the real-analytic de Rham cohomology of  $\text{Sh}_K(\mathbb{C})$ . As  $\text{Sh}_K(\mathbb{C})$  is a (finite disjoint union of) locally symmetric varieties for  $\mathbf{G}(\mathbb{R})$ , automorphic forms for  $\mathbf{G}$  enter the stage, and in particular one sees that if  $\pi_f$  appears in  $H^*$ , then there is some  $\pi_\infty$  such that  $\pi_f \otimes \pi_\infty$  is an irreducible automorphic representation of  $\mathbf{G}(\mathbb{A})$ . Using relative Lie algebra cohomology of  $\pi_\infty$ , one expresses the contribution of any automorphic representation  $\pi_f \otimes \pi_\infty$  of  $\mathbf{G}(\mathbb{A})$  to  $H^*$ .

The situation becomes more complicated when one is trying to understand the Galois action. Let us recall the general recipe given in Kottwitz’s paper [28]. Let us restrict to the contribution of tempered representations for simplicity. (The general case is similar, using  $A$ -parameters instead.) Let us remark that tempered representations should only show up in the middle degree cohomology, so that this really only gives part of the whole picture.

First, note that any irreducible automorphic representation  $\pi$  of  $\mathbf{G}(\mathbb{A})$  should give rise to a global Langlands parameter, which should be some map

$$\varphi_\pi : \mathcal{L}_{\mathbb{Q}} \rightarrow {}^L\mathbf{G} ,$$

where  $\mathcal{L}_{\mathbb{Q}}$  is the conjectural global Langlands group, and  ${}^L\mathbf{G}$  is the  $L$ -group of  $\mathbf{G}$ , which is the semidirect product of the dual group  $\widehat{\mathbf{G}}$  over  $\mathbb{C}$  and  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ . The parameter  $\varphi_\pi$  should be discrete due to our assumption on  $\mathbf{G}$ . Fix an isomorphism  $\overline{\mathbb{Q}}_\ell \cong \mathbb{C}$ . Henceforth the subscript  $\ell$  for a complex group will designate the base change from  $\mathbb{C}$  to  $\overline{\mathbb{Q}}_\ell$ . In the cases of interest here, such a parameter  $\varphi_\pi$  should be closely related to an  $\ell$ -adic Galois representation

$$\varphi_{\pi,\ell} : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \rightarrow {}^L\mathbf{G}_\ell .^2$$

A discrete tempered global Langlands parameter

$$\varphi : \mathcal{L}_{\mathbb{Q}} \rightarrow {}^L\mathbf{G} ,$$

upon restriction to each place  $v$  of  $\mathbb{Q}$ , should give rise to local  $L$ -packets  $\Pi(\varphi_v)$ , which are finite sets of isomorphism classes of representations of  $\mathbf{G}(\mathbb{Q}_v)$ . We let  $\Pi_f(\varphi)$  be the set of isomorphism classes of irreducible admissible representations  $\pi_f$  of  $\mathbf{G}(\mathbb{A}_f)$  such that  $\pi_p \in \Pi(\varphi_p)$  for all  $p$ .

On the other hand, the data defining the Shimura variety give rise to a natural representation

$$r : {}^L(\mathbf{G} \otimes_{\mathbb{Q}} E) \rightarrow \text{GL}_N ,$$

---

<sup>2</sup>We refer to [9] for a detailed discussion of this point. There is some subtle twisting issue which we ignore here.

for some integer  $N$ . In particular, one gets an  $\ell$ -adic Galois representation  $r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)}$ :

$$\text{Gal}(\bar{\mathbb{Q}}/E) \xrightarrow{\varphi_\ell} {}^L(\mathbf{G} \otimes_{\mathbb{Q}} E)_\ell \xrightarrow{r} \text{GL}_N(\bar{\mathbb{Q}}_\ell) .$$

A first approximation would be that the contribution of tempered representations to  $H^*$  is given by (up to twists)

$$(1) \quad \sum_{\varphi} \sum_{\pi_f \in \Pi_f(\varphi)} m(\pi_f) \pi_f \otimes (r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)}) .$$

Here  $m(\pi_f)$  are certain (possibly negative) integers related to multiplicities of automorphic representations.

This approximation turns out to be correct in the situations where one can ignore endoscopy. A part of this paper will be restricted to such situations, but the construction of Galois representations requires working in a more general context.

The additional ingredient needed is the group  $S_\varphi \subset \widehat{\mathbf{G}}$  of self-equivalences of  $\varphi$ , i.e. the subgroup of  $\widehat{\mathbf{G}}$  which centralizes the image of  $\varphi$  up to a locally trivial 1-cocycle of  $\mathcal{L}_{\mathbb{Q}}$  into  $Z(\widehat{\mathbf{G}})$ ; cf. [22, §11]. Let  $\mathfrak{S}_\varphi = S_\varphi/Z(\widehat{\mathbf{G}})$ , which is a finite abelian group in the cases of interest.<sup>3</sup> If  $\mathfrak{S}_\varphi$  is trivial, then the contribution of  $\varphi$  to (1) should be correct. Such  $\varphi$  are called stable.

To give the general recipe, let us look at the corresponding  $\ell$ -adic version. We get an action of  $S_{\varphi,\ell}$  on the Galois representation  $r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)}$ . One checks that the subgroup  $Z(\widehat{\mathbf{G}})_\ell$  acts by a certain fixed character  $\mu_1$ . Let  $\nu$  be a character of  $S_{\varphi,\ell}$  whose restriction to  $Z(\widehat{\mathbf{G}})_\ell$  is  $\mu_1$ , and denote by  $(r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)})_\nu$  the direct summand of  $r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)}$  on which  $S_{\varphi,\ell}$  acts through the character  $\nu$ . Note that the set of such  $\nu$  is a principal homogeneous space under the dual group of  $\mathfrak{S}_\varphi$ .

Now, the general contribution of tempered representations to  $H^*$  should be (up to twists)

$$(2) \quad \sum_{\varphi} \sum_{\pi_f \in \Pi_f(\varphi)} \sum_{\nu} m(\pi_f, \nu) \pi_f \otimes (r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)})_\nu .$$

Again,  $m(\pi_f, \nu)$  are certain integers, whose general definition is somewhat subtle.

The object of this paper is to verify such a description in as many cases as possible. As even the ingredients in this formula remain largely conjectural, one has to be more specific about the precise statement one wants to prove. We fix a prime  $\mathfrak{p}$  of  $E$  above a rational prime  $p \neq \ell$ , and restrict the representation of  $\text{Gal}(\bar{\mathbb{Q}}/E)$  to the Weil group  $W_{E_{\mathfrak{p}}}$  of  $E_{\mathfrak{p}}$ . Then  $\pi_p$  has a local  $L$ -parameter, which gives a map

$$\varphi_{\pi_p} : W_{\mathbb{Q}_p} \rightarrow {}^L\mathbf{G}_\ell .$$

One expects that if  $\pi_f \in \Pi(\varphi)$ , then  $\varphi_{\pi_p}$  is the restriction of  $\varphi_\ell$  to the local Weil group; this is known as the local-global compatibility. In particular, we have

$$(r \circ \varphi_\ell|_{\text{Gal}(\bar{\mathbb{Q}}/E)})|_{W_{E_{\mathfrak{p}}}} = r \circ \varphi_{\pi_p}|_{W_{E_{\mathfrak{p}}}} ,$$

so that the local Weil group representations are determined by the local components of the representation  $\pi_f$ . We want to prove that the contribution of tempered

---

<sup>3</sup>In fact, in most cases it is just a product of some copies of  $\mathbb{Z}/2\mathbb{Z}$ .

representations to  $H^*$  is

$$(3) \quad \sum_{\varphi} \sum_{\pi_f \in \Pi_f(\varphi)} \sum_{\nu} m(\pi_f, \nu) \pi_f \otimes (r \circ \varphi_{\pi_p}|_{W_{E_p}})_{\nu}$$

as a virtual  $\mathbf{G}(\mathbb{A}_f) \times W_{E_p}$ -representation.

This still requires the existence of the local Langlands correspondence for the group  $\mathbf{G}(\mathbb{Q}_p)$ . By [17] and [19], it is known for general linear groups. It is in this setup that we prove our results. In other words, we start with a Shimura variety associated to a reductive group  $\mathbf{G}$  over  $\mathbb{Q}$  which is anisotropic modulo its center, such that locally at  $p$ ,  $\mathbf{G}$  is a product of Weil restrictions of general linear groups. As  $\mathbf{G}$  will be a unitary (similitude) group in this case, we use the term compact unitary group Shimura varieties, and the prime  $p$  is called a split place.<sup>4</sup> We note that we allow both the representation  $\pi_p$  and the implicit CM field defining the unitary group  $\mathbf{G}$  to be ramified at  $p$ .

Our first theorem is the following. Here, we allow coefficients given by a general irreducible algebraic representation  $\xi$  of  $\mathbf{G}$ . The corresponding cohomology group is denoted by  $H_{\xi}^*$ .

**Theorem 1.1.** *Assume that  $\text{Sh}$  is a compact unitary group Shimura variety,  $p$  is a split place, and endoscopy is trivial, as in [29]. Then*

$$H_{\xi}^* = \sum_{\pi_f} a(\pi_f) \pi_f \otimes (r \circ \varphi_{\pi_p}|_{W_{E_p}}) \cdot | \cdot |^{-\dim \text{Sh}/2}$$

as a virtual  $\mathbf{G}(\mathbb{A}_f^p) \times \mathbf{G}(\mathbb{Z}_p) \times W_{E_p}$ -representation. Here  $\pi_f$  runs through irreducible admissible representations of  $\mathbf{G}(\mathbb{A}_f)$ , the integer  $a(\pi_f)$  is as in [29, p. 657], and  $\varphi_{\pi_p}$  is the local Langlands parameter associated to  $\pi_p$ . Also  $\mathbf{G}(\mathbb{Z}_p)$  is a certain maximal compact subgroup of  $\mathbf{G}(\mathbb{Q}_p)$ .

The restriction of the  $\mathbf{G}(\mathbb{Q}_p)$ -action to  $\mathbf{G}(\mathbb{Z}_p)$  is due to our method of proof. Let us remark that in the situation of the theorem, one expects that the  $p$ -component  $\pi_p$  is determined by  $\pi_f^p$ . Granting this, the identity would extend to an identity of  $\mathbf{G}(\mathbb{A}_f) \times W_{E_p}$ -representations. Unconditionally, the theorem gives a description of the semisimple Hasse-Weil zeta function.

**Corollary 1.2.** *In the situation of the theorem, let  $K \subset \mathbf{G}(\mathbb{A}_f)$  be any sufficiently small compact open subgroup. Then the semisimple local Hasse-Weil zeta function of  $\text{Sh}_K$  at the place  $\mathfrak{p}|p$  of  $E$  is given by*

$$\zeta_{\mathfrak{p}}^{\text{ss}}(\text{Sh}_K, s) = \prod_{\pi_f} L^{\text{ss}}(s - \dim \text{Sh}/2, \pi_p, r)^{a(\pi_f) \dim \pi_f^K}.$$

We remark that in these nonendoscopic cases, no form of the fundamental lemma or related work is needed, except for the existence of base-change transfers for  $\text{GL}_n$ , as proved by Arthur-Clozel, [3].

The second theorem is a description of the cohomology of compact unitary group Shimura varieties at split primes in endoscopic cases. In this case, the full machinery of endoscopy (for unitary groups) is used, and we have to make some additional assumptions concerning our data to establish unconditional results. Specifically,

---

<sup>4</sup>This is a slightly unfortunate terminology, as it only implies that  $\mathbf{G}_{\mathbb{Q}_p}$  is quasisplit, but not that it is split. A justification is that the primes above  $p$  should be split between the totally real field and its CM quadratic extension defining the Hermitian form for  $\mathbf{G}$ .

consider a unitary group  $\mathbf{G}$  relative to a CM quadratic extension  $F/F_0$  such that  $F$  contains an imaginary quadratic field  $\mathcal{K}$ . We assume that  $F_0 \neq \mathbb{Q}$ . Fix a rational prime  $p$  split in  $\mathcal{K}$  and a prime  $\mathfrak{p}$  of the reflex field  $E$  above  $p$ . We assume that  $\mathbf{G}$  is quasisplit locally at all finite places but we do not fix the signature at infinity, except that  $\mathbf{G}$  has a compact factor at one infinite place. The group  $\mathbf{G}(\mathbb{A}_{\mathcal{K}})$  admits an order 2 automorphism  $\theta$  induced by the complex conjugation on  $\mathcal{K}$ . Let  $\Pi$  be a  $\theta$ -stable automorphic representation of  $\mathbf{G}(\mathbb{A}_{\mathcal{K}}) \cong \mathrm{GL}_1(\mathbb{A}_{\mathcal{K}}) \times \mathrm{GL}_n(\mathbb{A}_F)$  which has infinite component determined by  $\xi$ . We define a virtual representation of  $\mathrm{Gal}(\overline{\mathbb{Q}}/E)$ ,

$$(4) \quad H_{\xi}^*(\pi_f) = \mathrm{Hom}_{\mathbf{G}(\mathbb{A}_f)}(\pi_f, H_{\xi}^*),$$

and also the  $\Pi$ -part  $H_{\xi}^*(\Pi)$  of  $H_{\xi}^*$ . The latter is roughly the sum of  $H_{\xi}^*(\pi_f)$  over the set of  $\pi_f$  whose base-change is  $\Pi_f$  (cf. (25)). We consider two cases, which are loosely described as:

- (Case ST)  $\Pi$  is cuspidal.
- (Case END)  $\Pi$  is induced from a  $\theta$ -stable cuspidal automorphic representation on a maximal proper Levi subgroup of  $\mathbf{G}(\mathbb{A}_{\mathcal{K}})$ .

**Theorem 1.3.** *The following hold up to an explicit constant (cf. (26)). In (Case ST),*

$$H_{\xi}^*(\Pi)|_{W_{E_{\mathfrak{p}}}} = (r \circ \varphi_{\pi_p}|_{W_{E_{\mathfrak{p}}}})| \cdot | \cdot |^{-\dim \mathrm{Sh}/2}.$$

*In (Case END), there exists a natural decomposition*

$$r \circ \varphi_{\pi_p}|_{W_{E_{\mathfrak{p}}}} = (r \circ \varphi_{\pi_p}|_{W_{E_{\mathfrak{p}}}})_1 + (r \circ \varphi_{\pi_p}|_{W_{E_{\mathfrak{p}}}})_2$$

*into two representations, and some  $i \in \{1, 2\}$  depending only on  $\Pi$  such that*

$$H_{\xi}^*(\Pi)|_{W_{E_{\mathfrak{p}}}} = (r \circ \varphi_{\pi_p}|_{W_{E_{\mathfrak{p}}}})_i | \cdot |^{-\dim \mathrm{Sh}/2}.$$

As a corollary of this theorem, one can reprove the existence of Galois representations associated to regular algebraic conjugate self-dual cuspidal automorphic representations of  $\mathrm{GL}_n$  over CM fields, as previously constructed by one of us, [44]; cf. Theorem 11.1. Of course the construction of the Galois representations with the Langlands-Kottwitz method goes back to more than two decades ago and there have been a number of important predecessors. Among them are [29], [13], [8], [17] and [35] for the case  $n \geq 3$ . Also note that [14] gave a construction around the same time as [44] with a weaker local-global compatibility result, namely only at unramified places. (The original Langlands-Kottwitz method was used in [14] whereas a somewhat different approach [17] was a basis for [44].) The fundamental lemma used to be a main obstacle blocking further progress in the past despite the presence of key ideas, but its proof by Ngô (as well as Waldspurger and others) opened the door for recent developments such as [35], [14] and [44].

We can also reprove the Ramanujan-Petersson conjecture at unramified places; cf. [12]:

**Corollary 1.4.** *Let  $F$  be a CM field, and let  $\Pi$  be a regular algebraic conjugate self-dual cuspidal automorphic representation for  $\mathrm{GL}_n(\mathbb{A}_F)$ . Then for all finite places  $v$  of  $F$  lying above a rational prime  $p$  above which  $F$  and  $\Pi$  are unramified, the component  $\Pi_v$  is tempered.*

Although our theorem gives information in the ramified case as well, we do not see a direct way of proving temperedness at ramified places, as done by Caraiani, [10].

Let us explain the strategy of the proof. This paper is a sequel to [40]. We recall that under certain circumstances, the main theorem of [40] gives an equality of the form

$$(5) \quad \text{tr}(\tau \times hf^p|H_\xi^*) = \sum_{\substack{(\gamma_0; \gamma, \delta) \\ \alpha(\gamma_0; \gamma, \delta)=1}} c(\gamma_0; \gamma, \delta) O_\gamma(f^p) TO_{\delta\sigma}(\phi_{\tau, h}) \text{tr} \xi(\gamma_0)$$

for the trace of a Hecke operator  $hf^p$  twisted with an element of the local Weil group  $\tau$  at a prime  $\mathfrak{p}|p$  of the reflex field  $E$ . We refer to Section 4, especially Theorem 4.1, for a precise statement. Here,  $h \in C_c^\infty(\mathbf{G}(\mathbb{Z}_p))$  for a certain extension of  $\mathbf{G}$  to a group scheme over  $\mathbb{Z}_{(p)}$ . There is no naive generalization of this theorem to general  $h$ , and for this reason our method only gives information about the  $\mathbf{G}(\mathbb{A}_f^p) \times \mathbf{G}(\mathbb{Z}_p) \times W_{E_p}$ -action.

We give a general stabilization of this expression, following the arguments of Kottwitz in [28]. Write  $\mathcal{E}_{\text{ell}}(\mathbf{G})$  for the set of isomorphism classes of elliptic endoscopic triples for  $\mathbf{G}$ . For  $(\mathbf{H}, s, \eta) \in \mathcal{E}_{\text{ell}}(\mathbf{G})$ , let  $f^{\mathbf{H}} = f^{\mathbf{H}, p} f_{\tau, h}^{\mathbf{H}} f_\xi^{\mathbf{H}}$  be the function on  $\mathbf{H}(\mathbb{A})$  obtained from  $f^p$ ,  $\phi_{\tau, h}$  and  $\xi$ . Here  $f^{\mathbf{H}, p}$  is the usual transfer of  $f^p$ , whereas  $f_{\tau, h}^{\mathbf{H}}$  is a twisted transfer of  $\phi_{\tau, h}$  relative to a suitable  $L$ -morphism. The function  $f_\xi^{\mathbf{H}}$  at infinity is an explicit linear combination of pseudo-coefficients for discrete series. Let  $ST_{\text{ell}}^{\mathbf{H}}$  denote the stable distribution on  $\mathbf{H}(\mathbb{A})$  given as the sum of stable orbital integrals on elliptic semisimple elements of  $\mathbf{H}(\mathbb{Q})$ .

**Theorem 1.5.** *The formula (5) is stabilized as follows, where the sum runs over  $\mathcal{E}_{\text{ell}}(\mathbf{G})$ :*

$$\text{tr}(\tau \times hf^p|H_\xi^*) = \sum_{(\mathbf{H}, s, \eta)} \iota(\mathbf{G}, \mathbf{H}) ST_{\text{ell}}^{\mathbf{H}}(f^{\mathbf{H}}).$$

To apply Theorem 1.5, it is essential to understand character identities satisfied by  $f_{\tau, h}^{\mathbf{H}}$ . We are able to formulate a precise conjecture (cf. Conjecture 7.1 and Conjecture 7.2), which is purely local in nature and may be thought of as a common generalization of the identities [41, Thm. 1.2.(a)] and [28, (9.7)]. As explained in the introduction of [40], it is also related to a conjecture of Kottwitz about similar functions arising via the bad reduction of Shimura varieties of parahoric type, and related ideas were expressed recently by Haines.

The conjecture relates (the endoscopic transfers of) the functions  $\phi_{\tau, h}$  arising from deformation spaces of  $p$ -divisible groups with the local Langlands Correspondence. In this sense it is similar in spirit to the conjectures that describe the cohomology of Rapoport-Zink spaces in terms of the local Langlands Correspondence; the precise relation between the two points of view remains to be clarified. Let us only mention here that our conjecture has the advantage that all Langlands parameters of  $\mathbf{H}$  occur in the same way, as was the case in [41], Theorem 1.2.

As evidence we verify the conjecture when all data are unramified; cf. Lemma 7.3. More importantly, we prove the conjecture in all cases of EL type; cf. Corollary 8.6. This extends the result of [41, Theorem 1.2], to arbitrary signature. We stress here that we allow ramified EL data.

Granting this conjecture, we follow the arguments of [28] and verify that for any compact PEL type Shimura variety of type A or C one gets the expected

description of the cohomology, assuming some additional hypotheses on the global  $A$ -packet classification and the stable trace formula, as in [28]. This justifies the general form of the conjecture.

Knowing the general character identities satisfied by (the endoscopic transfers of) the test function  $\phi_{\tau,h}$ , our main theorems follow from Theorem 1.5 and applications of suitable trace formulas.

We remark on one more application of our arguments in the nonendoscopic case. Namely, we prove that for certain compact unitary group Shimura varieties without endoscopy, all Newton strata are nonempty at “split” places in the above sense; cf. Corollary 8.3 and Corollary 8.4. In particular, this proves that the deformation spaces considered in [40] are algebraizable in all cases of EL type; cf. Corollary 8.5.

Let us summarize the content of the different sections. In the first sections, we give the stabilization leading to Theorem 1.5. We also explain the process of pseudostabilization in the nonendoscopic case; cf. Section 5. Afterwards, we state the conjecture about the character identities satisfied by the endoscopic transfers  $f_{\tau,h}^{\mathbf{H}}$  of  $\phi_{\tau,h}$  in Section 7. This will involve the theory of the stable Bernstein center, which we recall in Section 6. We go on by proving the conjecture in cases of EL type in Section 8. The last sections apply this result to prove our main theorems. We note that a reader only interested in our results in the nonendoscopic case finds that the argument in Section 5, the first half of Section 8, and Section 9.

## 2. KOTTWITZ TRIPLES

The definitions of this section are group-theoretic and do not require any reference to Shimura varieties even though that is where the motivation comes from. Let  $\mathbf{G}$  be a connected reductive group over  $\mathbb{Q}$  with  $A_{\mathbf{G}}$  the maximal split component of its center such that

- (1)  $(\mathbf{G}/A_{\mathbf{G}})(\mathbb{R})$  contains a compact maximal torus,
- (2)  $\mathbf{G}^{\text{der}}$  is simply connected,
- (3)  $\mathbf{G}_{\mathbb{Q}_p}$  is quasisplit.

Condition (1) is indispensable. Both (2) and (3) are not essential, but are assumed to avoid complication.

Let  $\mathbf{h} : \text{Res}_{\mathbb{C}/\mathbb{R}}\mathbb{G}_m \rightarrow \mathbf{G}_{\mathbb{R}}$  be a group morphism. Taking the base-change to  $\mathbb{C}$ , this gives rise to a morphism  $\mathbb{G}_m \times \mathbb{G}_m \rightarrow \mathbf{G}_{\mathbb{C}}$ , and we denote the first factor by  $\mu : \mathbb{G}_m \rightarrow \mathbf{G}_{\mathbb{C}}$ . Choosing a maximal torus  $\mathbf{T}$  of  $\mathbf{G}$  over  $\mathbb{C}$ , one can replace  $\mu$  by a conjugate which factors over  $\mathbf{T}$ ; this induces by duality a cocharacter of  $\widehat{\mathbf{T}}$ . Restricting this cocharacter to  $Z(\widehat{\mathbf{G}})$  gives an element  $\mu_1 \in X^*(Z(\widehat{\mathbf{G}}))$  independent of all choices.<sup>5</sup>

The  $\mathbf{G}(\mathbb{C})$ -conjugacy class  $\bar{\mu}$  of  $\mu$  has its field of definition  $E \subset \mathbb{C}$ , which is finite over  $\mathbb{Q}$ . We fix a prime  $\mathfrak{p}$  of  $E$  above  $p$ . Let  $f(\mathfrak{p})$  denote the inertia degree of  $E_{\mathfrak{p}}$  over  $\mathbb{Q}_p$ .

Set  $L$  to be the fraction field of  $W(\overline{\mathbb{F}}_p)$ . Write  $B(\mathbf{G}_{\mathbb{Q}_p})$  for the set of  $\sigma$ -conjugacy classes in  $\mathbf{G}(L)$ . Let

$$\kappa_{\mathbf{G},p} : B(\mathbf{G}_{\mathbb{Q}_p}) \rightarrow X^*(Z(\widehat{\mathbf{G}})^{\Gamma(p)})$$

denote the Kottwitz map; cf. [28], Section 6. Here  $\Gamma(p)$  denotes the absolute Galois group of  $\mathbb{Q}_p$ . Use  $\mathbf{G}(\mathbb{Q})_{\text{ss}}$  to denote the set of semisimple elements in  $\mathbf{G}(\mathbb{Q})$ .

<sup>5</sup>In particular, we revert back to the normalization used by Kottwitz, which is different from the one used in [40].

**Definition 2.1.** Let  $j \geq 1$  be an integer and set  $r = jf(\mathfrak{p})$ . A degree  $j$  Kottwitz triple  $(\gamma_0; \gamma, \delta)$  consists of  $\gamma_0 \in \mathbf{G}(\mathbb{Q})_{\text{ss}}$ ,  $\gamma \in \mathbf{G}(\mathbb{A}_f^p)$  and  $\delta \in \mathbf{G}(\mathbb{Q}_{p^r})$  such that

- $\gamma_0$  is elliptic in  $\mathbf{G}(\mathbb{R})$ ,
- $\gamma_0$  and  $\gamma$  are stably conjugate in  $\mathbf{G}(\mathbb{Q}_v)$  for all  $v \neq p, \infty$ ,
- $\gamma_0$  and  $N\delta = \delta\delta^\sigma \cdots \delta^{\sigma^{r-1}}$  are stably conjugate, i.e. conjugate in  $\mathbf{G}(\overline{\mathbb{Q}}_p)$ ,
- $\kappa_{\mathbf{G},p}(\delta) = -\mu_1$  in  $X^*(Z(\widehat{\mathbf{G}})^{\Gamma(p)})$ .

Two triples  $(\gamma_0; \gamma, \delta)$  and  $(\gamma'_0; \gamma', \delta')$  are considered equivalent if  $\gamma_0 \sim_{\text{st}} \gamma'_0$ ,  $\delta \sim_\sigma \delta'$  and  $\gamma_v \sim \gamma'_v$  for all  $v \neq p, \infty$ . Denote by  $\text{KT}_j$  (a set of representatives for) equivalence classes of degree  $j$  Kottwitz triples.

Let  $I_0$  be the centralizer of  $\gamma_0$  in  $\mathbf{G}$ . Kottwitz defines an invariant  $\beta_v(\gamma_0; \gamma, \delta) \in X^*(Z(\widehat{I_0})^{\Gamma(v)}Z(\widehat{\mathbf{G}}))$  for each place  $v$  of  $\mathbb{Q}$  in [28, Section 2]. Set  $\widetilde{\mathfrak{K}}(I_0/\mathbb{Q}) = \bigcap_v Z(\widehat{I_0})^{\Gamma(v)}Z(\widehat{\mathbf{G}})$ . The product

$$\beta(\gamma_0; \gamma, \delta) = \prod_v \beta_v(\gamma_0; \gamma, \delta)|_{\widetilde{\mathfrak{K}}(I_0/\mathbb{Q})} \in X^*(\widetilde{\mathfrak{K}}(I_0/\mathbb{Q}))$$

is trivial on  $Z(\widehat{\mathbf{G}})$  by construction, hence defines a character  $\alpha(\gamma_0; \gamma, \delta)$  of  $\mathfrak{K}(I_0/\mathbb{Q}) = (\bigcap_v Z(\widehat{I_0})^{\Gamma(v)}Z(\widehat{\mathbf{G}}))/Z(\widehat{\mathbf{G}})$ .

### 3. ENDOSCOPIC TRANSFER OF TEST FUNCTIONS

In this section we define the test functions needed to stabilize the formula given in Theorem 4.1 below, following Kottwitz.

First, we recollect some general facts about endoscopy. Let  $F$  be a local or global field of characteristic 0. Let  $\mathbf{G}$  be a connected reductive group over  $F$ . As in the last section, we assume for convenience that

- $\mathbf{G}^{\text{der}}$  is simply connected.

Let  $(\mathbf{H}, s, \eta) \in \mathcal{E}(\mathbf{G})$  be an endoscopic triple consisting of a quasisplit connected reductive group  $\mathbf{H}$  over  $F$ , an element  $s \in Z(\widehat{\mathbf{H}})$  and an embedding of complex Lie groups  $\eta : \widehat{\mathbf{H}} \rightarrow \widehat{\mathbf{G}}$  satisfying the conditions of [22, 7.4]. Such a triple is elliptic if  $(Z(\widehat{\mathbf{H}})^{\Gamma})^0 \subset Z(\widehat{\mathbf{G}})$ . The notion of isomorphism for endoscopic triples is defined in [22, §7]. We use the Weil form of the  $L$ -group, and fix an  $L$ -morphism  ${}^L\mathbf{H} \rightarrow {}^L\mathbf{G}$  extending  $\eta$ , which exists since  $\mathbf{G}^{\text{der}}$  is simply connected ([33, Prop. 1]). The latter  $L$ -morphism is also denoted  $\eta$  by abuse of notation. Write  $\mathcal{E}(\mathbf{G})$  (resp.  $\mathcal{E}_{\text{ell}}(\mathbf{G})$ ) for the set of isomorphism classes of endoscopic (resp. elliptic endoscopic) triples for  $\mathbf{G}$ .

**3.1. Places  $v \neq p, \infty$ : Untwisted endoscopy.** Now assume that  $F$  is a local field. Langlands and Shelstad ([34]) define a transfer factor

$$\Delta : \mathbf{H}(F)_{\text{ss},(\mathbf{G},\mathbf{H})\text{-reg}} \times \mathbf{G}(F)_{\text{ss}} \rightarrow \mathbb{C},$$

which is canonical up to a nonzero constant. Some basic properties are that  $\Delta(\gamma_{\mathbf{H}}, \gamma)$  depends only on the stable conjugacy class of  $\gamma_{\mathbf{H}}$  and the  $\mathbf{G}(F)$ -conjugacy class of  $\gamma$  and that  $\Delta(\gamma_{\mathbf{H}}, \gamma) = 0$  unless  $\gamma_{\mathbf{H}}$  is  $(\mathbf{G}, \mathbf{H})$ -regular and associated to  $\gamma$ .



The fundamental lemma and the transfer conjecture, which are now theorems due to Ngô, Waldspurger and others (see [48], [36] and the references therein), assert

**Theorem 3.1.** *For each  $f \in C_c^\infty(\mathbf{G}(F))$ , there exists an  $f^{\mathbf{H}} \in C_c^\infty(\mathbf{H}(F))$  such that for every  $\gamma_{\mathbf{H}} \in \mathbf{H}(F)_{\text{ss},(\mathbf{G},\mathbf{H})\text{-reg}}$ ,*

$$SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(F)}(f^{\mathbf{H}}) = \sum_{\gamma \in \mathbf{G}(F)_{\text{ss}}/\sim} \Delta(\gamma_{\mathbf{H}}, \gamma) e(\mathbf{G}_\gamma) O_\gamma^{\mathbf{G}(F)}(f),$$

where  $\gamma$  runs over the set of representatives for all conjugacy classes in  $\mathbf{G}(F)_{\text{ss}}$  and  $\mathbf{G}_\gamma$  denotes the centralizer of  $\gamma$  in  $\mathbf{G}$ . Now assume that  $\mathbf{H}$ ,  $\mathbf{G}$  and  $\eta$  are unramified and that  $f \in \mathcal{H}^{\text{ur}}(\mathbf{G}(F))$ . Then  $\Delta$  can be normalized such that  $f^{\mathbf{H}} = \eta^*(f)$  satisfies the above formula, where  $\eta^* : \mathcal{H}^{\text{ur}}(\mathbf{G}(F)) \rightarrow \mathcal{H}^{\text{ur}}(\mathbf{H}(F))$  is the natural induced morphism of unramified Hecke algebras. In particular, if  $f$  is the idempotent associated to a hyperspecial maximal compact subgroup, then one may take  $f^{\mathbf{H}}$  as the idempotent of a hyperspecial maximal compact subgroup as well.

We will say that  $f^{\mathbf{H}}$  as above is an  $(\eta)$ -transfer of  $f$ .

*Remark 3.2.* Our transfer factor follows the convention of [34] and [26], which are opposite to the one of [28]. (See [28, p. 178] and [35, p. 201].) This introduces the sign change  $s \mapsto s^{-1}$  when citing results from the latter.

**3.2. Place  $p$ : Twisted endoscopy.** Let us recall the setup in which functions  $\phi_{\tau,h}$  were defined in [40]. In particular, we have a quasisplit reductive group  $\mathbf{G}$  over  $\mathbb{Q}_p$  and a conjugacy class  $\bar{\mu}$  of cocharacters  $\mu : \mathbb{G}_m \rightarrow \mathbf{G}_{\bar{\mathbb{Q}}_p}$  with field of definition  $E \subset \bar{\mathbb{Q}}_p$ .<sup>6</sup> We get  $\mu_1 \in X^*(Z(\widehat{\mathbf{G}}))$ . Let  $f$  be the inertial degree of the extension  $E/\mathbb{Q}_p$ , and let  $r = jf$  for some  $j \geq 1$ . In [40], certain functions  $\phi_{\tau,h} \in C_c^\infty(\mathbf{G}(\mathbb{Q}_{p^r}))$  are defined, depending on  $\tau$  in the Weil group of  $E$ , acting as the  $j$ -th power of geometric Frobenius on the residue field, and  $h \in C_c^\infty(\mathbf{G}(\mathbb{Z}_p))$  for a suitable integral model of  $\mathbf{G}$ . One important property is that

- $\phi_{\tau,h}(\delta) = 0$  unless  $\kappa_{\mathbf{G}}(\delta) = -\mu_1$  in  $X^*(Z(\widehat{\mathbf{G}}))^{\Gamma(p)}$ .

Here  $\kappa_{\mathbf{G}} : B(\mathbf{G}) \rightarrow X^*(Z(\widehat{\mathbf{G}}))^{\Gamma(p)}$  is the Kottwitz morphism as above.

We want to define the twisted endoscopic transfers of  $\phi_{\tau,h}$ . Let  $(\mathbf{H}, s, \eta) \in \mathcal{E}(\mathbf{G}_{\mathbb{Q}_p})$  be any endoscopic triple. First, we are going to give the construction of the endoscopic transfer  $f_{\tau,h}^{\mathbf{H}}$  of the function  $\phi_{\tau,h}$  when  $s \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}$ . It will be explained at the end of this subsection what to do in the general case where  $s \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}Z(\widehat{\mathbf{G}})$ . Let  $R_r = \text{Res}_{\mathbb{Q}_{p^r}/\mathbb{Q}_p} \mathbf{G}_{\mathbb{Q}_{p^r}}$ , equipped with the obvious automorphism  $\theta$  such that  $R_r^\theta = \mathbf{G}$ . We have an identification  $\widehat{R}_r = \widehat{\mathbf{G}} \times \cdots \times \widehat{\mathbf{G}}$  ( $r$  copies) on which any lift of the arithmetic Frobenius  $\sigma$  acts by  $(g_1, \dots, g_r) \mapsto (\sigma(g_2), \dots, \sigma(g_r), \sigma(g_1))$ . In particular, there is a natural  $L$ -morphism  $\xi : {}^L\mathbf{G} \rightarrow {}^L R_r$  sending  $g \rtimes w \in {}^L\mathbf{G}$  to  $(g, \dots, g) \rtimes w \in {}^L R_r$ . Choose elements  $t_i \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}$  for  $i = 1, \dots, r$  such that  $t_1 t_2 \cdots t_r = s$  and write  $t = (t_1, \dots, t_r)$ . Consider the  $L$ -morphism

$$(6) \quad \tilde{\eta} : {}^L\mathbf{H} \rightarrow {}^L R_r$$

defined by  $\tilde{\eta}(h) = \xi(\eta(h))$  for  $h \in \widehat{\mathbf{H}}$  and  $\tilde{\eta}(\tilde{\sigma}) = t\xi(\eta(\tilde{\sigma}))$  for any lift  $\tilde{\sigma} \in {}^L\mathbf{H}$  of  $\sigma$ . Then  $(\mathbf{H}, t, \tilde{\eta})$  is a twisted endoscopic group for  $(R_r, \theta)$ .

<sup>6</sup>Here, we switch to Kottwitz’s normalization of  $\mu$ : Let  $\mu_{RZ}$  denote the  $\mu$  considered in [40]; then the product  $\mu\mu_{RZ}$  is to be the central morphism  $\mathbb{G}_m \rightarrow \mathbf{G}$  that sends  $t \in \mathbb{G}_m$  to multiplication by  $t$  on  $V$ .

Ngô and Waldspurger (see [49]) showed that the twisted endoscopic transfer exists. The implication in our situation is that there exists a function  $f_{\tau,h}^{\mathbf{H}} \in C_c^\infty(\mathbf{H}(\mathbb{Q}_p))$  such that for all  $(\mathbf{G}, \mathbf{H})$ -regular semisimple  $\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{Q}_p)$ ,

$$(7) \quad SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{Q}_p)}(f_{\tau,h}^{\mathbf{H}}) = \sum_{\delta \in \mathbf{G}(\mathbb{Q}_p)/\sim_\sigma} \Delta_p(\gamma_{\mathbf{H}}, \delta) e(\mathbf{G}_{\delta\sigma}) TO_{\delta\sigma}(\phi_{\tau,h}),$$

where  $\Delta_p$  denotes the twisted transfer factor, which is nonzero exactly when the twisted conjugacy class of  $\delta$  transfers to the stable conjugacy class of  $\gamma_{\mathbf{H}}$ , and  $\mathbf{G}_{\delta\sigma}$  is the  $\sigma$ -centralizer of  $\delta$  in  $\mathbf{G}$ .

As in our situation,  $\mathbf{G}$  is quasisplit over  $\mathbb{Q}_p$ ; more can be said about the transfer factors. Kottwitz showed ([35, Appendix]) that if canonical transfer factors ([35, A.1.5]) are used, then

$$(8) \quad \Delta_p(\gamma_{\mathbf{H}}, \delta) = \langle \alpha(\gamma_0; \delta), s \rangle^{-1} \Delta_p(\gamma_{\mathbf{H}}, \gamma_0),$$

provided that  $\gamma_0 \in \mathbf{G}(\mathbb{Q}_p)_{\text{ss}}$  and the  $\gamma_{\mathbf{H}}$  have matching stable conjugacy classes. Such a  $\gamma_0$  exists since  $\mathbf{G}$  is quasisplit over  $\mathbb{Q}_p$ . Henceforth we will always choose  $\Delta_p(\gamma_{\mathbf{H}}, \delta)$  and  $\Delta_p(\gamma_{\mathbf{H}}, \gamma_0)$  to be canonical transfer factors.

For later use, we recall the definition of  $\alpha(\gamma_0; \delta)$  here; cf. [28, p. 167]. Let us write  $I_0$  for the centralizer of  $\gamma_0$  in  $\mathbf{G}$  over  $\mathbb{Q}_p$ . We have  $I_0 \subset \mathbf{H}$  since  $\gamma_0$  is  $(\mathbf{G}, \mathbf{H})$ -regular. Recall that the set of  $\sigma$ -conjugacy classes in  $I_0(L)$  is denoted  $B(I_0)$ . Using Steinberg’s theorem, choose any  $d \in \mathbf{G}(L)$  such that  $\delta\delta^\sigma \cdots \delta^{\sigma^{r-1}} = d\gamma_0 d^{-1}$ . Then it turns out that  $d^{-1}\delta d^\sigma$  belongs to  $I_0(L)$  and yields a well-defined element of  $B(I_0)$ . The image of the latter element under the Kottwitz map  $\kappa_{I_0} : B(I_0) \rightarrow X^*(Z(\widehat{I_0})^{\Gamma(p)})$  is called  $\alpha(\gamma_0; \delta)$ .

Putting this together, we get

$$(9) \quad SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{Q}_p)}(f_{\tau,h}^{\mathbf{H}}) = \sum_{\delta \in \mathbf{G}(\mathbb{Q}_p)/\sim_\sigma} \langle \alpha(\gamma_0; \delta), s \rangle^{-1} \Delta_p(\gamma_{\mathbf{H}}, \gamma_0) e(\mathbf{G}_{\delta\sigma}) TO_{\delta\sigma}(\phi_{\tau,h}).$$

So far we treated the case where  $s \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}$ . In general, we may write  $s = s_0 z$  for  $s_0 \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}$  and  $z \in Z(\widehat{\mathbf{G}})$ . At this point, we remark that the twisted orbital integrals  $TO_{\delta\sigma}(\phi_{\tau,h})$  vanish unless  $\kappa_{\mathbf{G}}(\delta) = -\mu_1$  in  $X^*(Z(\widehat{\mathbf{G}})^{\Gamma(p)})$ . This allows us to define  $\beta(\gamma_0; \delta) \in X^*(Z(\widehat{I_0})^{\Gamma(p)} Z(\widehat{\mathbf{G}}))$  as the extension of  $\alpha(\gamma_0; \delta)$  which is  $-\mu_1$  on  $Z(\widehat{\mathbf{G}})$ . Then  $\langle \beta(\gamma_0; \delta), s \rangle^{-1} = \mu_1(z) \langle \beta(\gamma_0; \delta), s_0 \rangle^{-1}$ . We define  $f_{\tau,h}^{\mathbf{H}}$  relative to  $(\mathbf{H}, s, \eta)$  as  $\mu_1(z)$  times the function  $f_{\tau,h}^{\mathbf{H}}$  relative to  $(\mathbf{H}, s_0, \eta)$  as constructed above, so that (9) still holds if one restricts the sum to  $\delta$  satisfying  $\kappa_{\mathbf{G}}(\delta) = -\mu_1$  and replaces  $\alpha(\gamma_0; \delta)$  by  $\beta(\gamma_0; \delta)$ .

**3.3. Place  $\infty$ : Pseudocoefficients.** In this subsection, we assume that  $F = \mathbb{R}$ . Moreover, we fix an algebraic representation  $\xi$  of  $\mathbf{G}$  over  $\mathbb{C}$ . Assume that  $(\mathbf{H}, s, \eta)$  is an elliptic endoscopic triple of  $\mathbf{G}$ . We recall a construction of Kottwitz based on Shelstad’s theory of real endoscopy.

For a discrete series representation  $\pi_{\mathbf{H}}$  of  $\mathbf{H}(\mathbb{R})$ , write  $\phi_{\pi_{\mathbf{H}}}$  for its pseudo-coefficient. For any discrete  $L$ -parameter  $\varphi_{\mathbf{H}} : W_{\mathbb{R}} \rightarrow {}^L\mathbf{H}$ , write  $\Pi^2(\varphi_{\mathbf{H}})$  for its associated  $L$ -packet. Set  $\phi_{\varphi_{\mathbf{H}}} = |\Pi^2(\varphi_{\mathbf{H}})|^{-1} \sum_{\pi_{\mathbf{H}} \in \Pi^2(\varphi_{\mathbf{H}})} \phi_{\pi_{\mathbf{H}}}$ . Write  $\varphi_{\xi} : W_{\mathbb{R}} \rightarrow {}^L\mathbf{G}$  for the discrete  $L$ -parameter such that  $\Pi^2(\varphi_{\xi})$  consists of discrete series with the same central character and infinitesimal character as  $\xi^\vee$ . Define  $\Phi_{\mathbf{H}}(\varphi_{\xi})$  to be the

set of discrete  $\varphi_{\mathbf{H}}$  such that  $\eta\varphi_{\mathbf{H}}$  and  $\varphi_{\xi}$  are equivalent. Set

$$f_{\xi}^{\mathbf{H}} = \langle \mu_{\mathbf{h}}, s \rangle^{-1} \sum_{\varphi_{\mathbf{H}} \in \Phi_{\mathbf{H}}(\varphi_{\xi})} (-1)^{q(\mathbf{G})} \det(\omega_*(\varphi_{\mathbf{H}})) \phi_{\varphi_{\mathbf{H}}},$$

where  $\det(\omega_*(\varphi_{\mathbf{H}}))$  is defined on pp. 184-185 and  $\langle \mu_{\mathbf{h}}, s \rangle^{-1}$  is defined on p. 185 of [28]. Kottwitz shows that for every  $\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{R})_{\text{ss},(\mathbf{G},\mathbf{H})\text{-reg}}$ ,

$$(10) \quad SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{R})}(f_{\xi}^{\mathbf{H}}) = \langle \beta_{\infty}(\gamma_0), s \rangle^{-1} \Delta_{\infty}(\gamma_{\mathbf{H}}, \gamma_0) e(I) \text{tr} \xi(\gamma_0) \text{vol}(A_{\mathbf{G}}(\mathbb{R})^{\circ} \backslash I(\mathbb{R}))^{-1}$$

if  $\gamma_{\mathbf{H}}$  is elliptic and  $SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{R})}(f_{\infty}^{\mathbf{H}}) = 0$  otherwise; cf. [28, (7.4), page 186]. Here  $\Delta_{\infty}(\gamma_{\mathbf{H}}, \gamma_0)$  is as on page 184 of [28], and  $I$  is the inner form of the centralizer  $I_0$  of  $\gamma_0$  in  $\mathbf{G}$  that is anisotropic modulo its center. Moreover Proposition 3.3.4 (cf. Remark 6.2.2) of [35] implies that

$$(11) \quad SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{R})}(f_{\xi}^{\mathbf{H}}) = 0$$

if  $\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{R})_{\text{ss}}$  is not  $(\mathbf{G}, \mathbf{H})$ -regular.

#### 4. STABILIZATION - PROOF OF THEOREM 1.5

Let us first briefly recall the main theorem from [40].

Let  $B$  be a simple  $\mathbb{Q}$ -algebra with center  $F$  and maximal  $\mathbb{Z}_{(p)}$ -order  $\mathcal{O}_B$  that is stable under a positive involution  $*$  on  $B$ . Let  $V$  be a finitely generated left  $B$ -module with a nondegenerate  $*$ -Hermitian form  $(\ , \ )$ . We assume that there is an  $\mathcal{O}_B$ -stable self-dual  $\mathbb{Z}_{(p)}$ -lattice  $\Lambda \subset V$ , which we fix. Moreover, we let  $F_0 = F^{*-1}$ , which is a totally real field. We assume that at all places above  $p$ , the extension  $F/F_0$  is unramified and the  $F$ -algebra  $B$  is split.

We let  $C = \text{End}_B(V)$  and  $\mathcal{O}_C = \text{End}_{\mathcal{O}_B}(\Lambda)$ ; both carry an involution  $*$  induced from  $(\ , \ )$ . We recall (cf. [30, p. 375]) that over  $\bar{\mathbb{Q}}$ , the algebra  $C$  together with the involution  $*$  is of one of the following types:

- (A)  $M_n \times M_n^{\text{opp}}$  with  $(x, y)^* = (y, x)$ ,
- (C)  $M_{2n}$  with  $x^*$  being the adjoint of  $x$  with respect to a nondegenerate alternating form in  $2n$  variables,
- (D)  $M_{2n}$  with  $x^*$  being the adjoint of  $x$  with respect to a nondegenerate symmetric form in  $2n$  variables.

We assume that case (A) or (C) occurs. We get the reductive group  $\mathbf{G}/\mathbb{Q}$  of  $B$ -linear similitudes of  $V$ ; in fact, we can extend it to an algebraic group over  $\mathbb{Z}_{(p)}$  as the group representing the functor

$$\mathbf{G}(R) = \{g \in (\mathcal{O}_C \otimes_{\mathbb{Z}_p} R)^{\times} \mid gg^* \in R^{\times}\}.$$

Finally, we fix a homomorphism  $\mathbf{h}_0 : \mathbb{C} \rightarrow C \otimes \mathbb{R}$  such that  $\mathbf{h}_0(\bar{z}) = \mathbf{h}_0(z)^*$  for all  $z \in \mathbb{C}$ , and such that the symmetric real-valued bilinear form  $(v, \mathbf{h}_0(i)w)$  on  $V \otimes \mathbb{R}$  is positive definite.

We write  $\mathbf{h}$  for the map  $\mathbb{S} \rightarrow \mathbf{G} \otimes \mathbb{R}$  from Deligne’s torus  $\mathbb{S}$  (i.e., the algebraic torus over  $\mathbb{R}$  with  $\mathbb{S}(\mathbb{R}) = \mathbb{C}^{\times}$ ) that is given on  $\mathbb{R}$ -valued points by  $\mathbf{h}(z) = \mathbf{h}_0(z)$ ,  $z \in \mathbb{C}^{\times}$ . Then one gets a tower  $\text{Sh}_K$ ,  $K \subset \mathbf{G}(\mathbb{A}_f)$  running through compact open subgroups of the adelic points of  $\mathbf{G}$ , of Shimura varieties associated to the pair  $(\mathbf{G}, \mathbf{h}^{-1})$ .

Of course, we get  $\mu$  and the reflex field  $E$  as before, and  $\text{Sh}_K$  has a canonical model defined over  $E$ . We fix a prime  $\mathfrak{p}$  of  $E$  above  $p$ .

Finally, fix an algebraic representation  $\xi$  of  $\mathbf{G}$ , defined over a number field  $L$ , and let  $\lambda$  be a place of  $L$  lying over the rational prime  $\ell \neq p$ . This gives  $\ell$ -adic local systems  $\mathcal{F}_{\xi,K}$  on  $\text{Sh}_K$  to which the action of the Hecke operators extends. Now we define

$$H_{\xi}^* = \varinjlim_K H^*(\text{Sh}_K \otimes \bar{\mathbb{Q}}, \mathcal{F}_{K,\xi})$$

as a virtual representation of  $\text{Gal}(\bar{\mathbb{Q}}/E) \times \mathbf{G}(\mathbb{A}_f)$ .

We make the assumption that the integral models defined in [40] are proper; in particular, the Shimura varieties themselves need to be proper, i.e.  $\mathbf{G}$  needs to be anisotropic modulo its center.

We choose Haar measures on various groups exactly as in [35, §5.1]; cf. also [28, §3]. We merely remark that adelic groups are given the Tamagawa measure and that the local measures at non-Archimedean places assign rational numbers to open compact subgroups. Moreover, in our situation we have defined a model of our group over  $\mathbb{Z}_p$ , and we require that  $\mathbf{G}(\mathbb{Z}_p)$  and  $\mathbf{G}(\mathbb{Z}_p^r)$  get volume 1.

Now we have the following theorem from [40].

**Theorem 4.1.** *Let  $\tau \in W_{E_p}$  project to the  $j$ -th power of the geometric Frobenius, for some  $j \geq 1$ . Let  $h \in C_c^\infty(\mathbf{G}(\mathbb{Z}_p))$  and  $f^p \in C_c^\infty(\mathbf{G}(\mathbb{A}_f^p))$ . Then*

$$\text{tr}(\tau \times hf^p | H_{\xi}^*) = \sum_{\substack{(\gamma_0; \gamma, \delta) \in \text{KT}_j \\ \alpha(\gamma_0; \gamma, \delta) = 1}} c(\gamma_0; \gamma, \delta) O_{\gamma}(f^p) T O_{\delta\sigma}(\phi_{\tau,h}) \text{tr} \xi(\gamma_0),$$

where  $c(\gamma_0; \gamma, \delta)$  is a volume constant defined as in [28, p. 172].

We remark that a priori these expressions are numbers in an algebraic closure of  $L_{\lambda}$ ; we choose an isomorphism  $\bar{L}_{\lambda} \cong \mathbb{C}$  and consider everything as  $\mathbb{C}$ -valued from now on.

Just as (4.2) was obtained from (3.1) in [28], we derive from Theorem 4.1 that  $\text{tr}(\tau \times hf^p | H_{\xi}^*)$  is equal to

$$(12) \quad \tau(\mathbf{G}) \sum_{\gamma_0} \sum_{\kappa} \sum_{(\gamma, \delta)} \langle \alpha(\gamma_0; \gamma, \delta), \kappa \rangle^{-1} e(\gamma, \delta) O_{\gamma}(f^p) T O_{\delta\sigma}(\phi_{\tau,h}) \cdot \text{tr} \xi(\gamma_0) \text{vol}(A_{\mathbf{G}}(\mathbb{R})^\circ \backslash I(\infty)(\mathbb{R}))^{-1},$$

where the inner sum runs over  $(\gamma, \delta)$  such that  $(\gamma_0; \gamma, \delta) \in \text{KT}_j$ . Again,  $I(\infty)/\mathbb{R}$  is the compact modulo center inner form of the centralizer  $I_0$  of  $\gamma_0$  in  $\mathbf{G}$ . For every  $\kappa$  in the sum, fix a lift  $\tilde{\kappa}$  of  $\kappa$  to  $\tilde{\mathfrak{K}}(I_0/\mathbb{Q})$ . Note  $\langle \alpha(\gamma_0; \gamma, \delta), \kappa \rangle = \langle \beta(\gamma_0; \gamma, \delta), \tilde{\kappa} \rangle$ .

Lemma 9.7 of [23] and the global hypothesis for transfer factors ([34, §6.4]) may be summarized as follows.

**Lemma 4.2.** *There exists a canonical map*

$$\begin{aligned} \mathcal{T} : & \coprod_{(\mathbf{H}, s, \eta) \in \mathcal{E}_{\text{ell}}(\mathbf{G})} \mathbf{H}(\mathbb{Q})_{(\mathbf{G}, \mathbf{H})\text{-reg, ss}} / \sim_{\text{st}} \\ & \longrightarrow \{(\gamma_0, \kappa) : \gamma_0 \in \mathbf{G}(\mathbb{Q})_{\text{ss}} / \sim_{\text{st}}, \kappa \in \mathfrak{K}(I_0/\mathbb{Q})\} \cup \{\emptyset\} \end{aligned}$$

such that:

- (1)  $(\mathbf{H}, s, \eta, \gamma_{\mathbf{H}}) \mapsto (\gamma_0, \kappa)$  if  $\gamma_0$  and  $\gamma_{\mathbf{H}}$  have matching stable conjugacy classes,  $s$  lands in  $\tilde{\mathfrak{K}}(I_0/\mathbb{Q}) \subset Z(\hat{I}_0)$  via  $Z(\hat{I}_{\gamma_{\mathbf{H}}}) \simeq Z(\hat{I}_0)$  and the image of  $s$  maps to  $\kappa \in \mathfrak{K}(I_0/\mathbb{Q})$  via projection.
- (2)  $(\mathbf{H}, s, \eta, \gamma_{\mathbf{H}}) \mapsto \emptyset$  if  $\gamma_{\mathbf{H}}$  does not transfer to  $\mathbf{G}(\mathbb{Q})$ .

For any  $(\gamma_0, \kappa)$ , there is a unique elliptic endoscopic triple  $(\mathbf{H}, s, \eta) \in \mathcal{E}_{\text{ell}}(\mathbf{G})$  containing a preimage of  $(\gamma_0, \kappa)$ ; in that case, we have

$$|\mathcal{T}^{-1}(\gamma_0, \kappa)| = |\text{Out}((\mathbf{H}, s, \eta))|.$$

Moreover, for each  $(\mathbf{H}, s, \eta) \in \mathcal{E}_{\text{ell}}(\mathbf{G})$ , the transfer factors  $\Delta_v$  may be normalized globally such that the following holds: for all  $\gamma_0 \in \mathbf{G}(\mathbb{Q})_{\text{ss}}$  and  $\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{Q})_{(\mathbf{G}, \mathbf{H})\text{-reg, ss}}$ ,

$$\prod_v \Delta_v(\gamma_{\mathbf{H}}, \gamma_0) = \begin{cases} 1, & \text{if } \gamma_{\mathbf{H}} \mapsto \gamma_0, \\ 0, & \text{otherwise.} \end{cases}$$

Thanks to the fundamental lemma (Theorem 3.1), we can find an  $\eta$ -transfer  $f^{\mathbf{H}, p}$  of  $f^p$  such that for every  $\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{A}_f^p)_{\text{ss}, (\mathbf{G}, \mathbf{H})\text{-reg}}$ ,

$$(13) \quad SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{A}_f^p)}(f^{\mathbf{H}, p}) = \sum_{\gamma \in \mathbf{G}(\mathbb{A}_f^p)_{\text{ss}}/\sim} \Delta^p(\gamma_{\mathbf{H}}, \gamma) e^p(\gamma) O_{\gamma}(f^p).$$

Kottwitz’s construction of  $\beta_v$  at  $v \neq p, \infty$  implies that (see the second to last equality of [28, p. 169], keeping Remark 3.2 in mind)

$$(14) \quad \Delta^p(\gamma_{\mathbf{H}}, \gamma) = \Delta^p(\gamma_{\mathbf{H}}, \gamma_0) \frac{\Delta^p(\gamma_{\mathbf{H}}, \gamma)}{\Delta^p(\gamma_{\mathbf{H}}, \gamma_0)} = \Delta^p(\gamma_{\mathbf{H}}, \gamma_0) \langle \beta^p(\gamma_0; \gamma, \delta), s \rangle^{-1}.$$

We have defined  $f_{\tau, h}^{\mathbf{H}}$  and  $f_{\xi}^{\mathbf{H}}$  in the last section. Put  $f^{\mathbf{H}} = f^{\mathbf{H}, p} f_{\tau, h}^{\mathbf{H}} f_{\xi}^{\mathbf{H}}$ . Lemma 4.2, formulas (9), (8), (11), (13), (14) and (10) imply (cf. argument in [28, pp. 188-189]) that the expression (12) is equal to

$$\begin{aligned} & \sum_{(\mathbf{H}, s, \eta) \in \mathcal{E}_{\text{ell}}(\mathbf{G})} \tau(\mathbf{G}) |\text{Out}((\mathbf{H}, s, \eta))|^{-1} \sum_{\substack{\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{Q})_{\text{ss}}/\sim_{\text{st}} \\ \text{elliptic in } \mathbf{H}(\mathbb{R})}} SO_{\gamma_{\mathbf{H}}}(f^{\mathbf{H}}) \\ &= \sum_{(\mathbf{H}, s, \eta)} \iota(\mathbf{G}, \mathbf{H}) \tau(\mathbf{H}) \sum_{\substack{\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{Q})_{\text{ss}}/\sim_{\text{st}} \\ \text{elliptic in } \mathbf{H}(\mathbb{Q})}} SO_{\gamma_{\mathbf{H}}}(f^{\mathbf{H}}) = \sum_{(\mathbf{H}, s, \eta)} \iota(\mathbf{G}, \mathbf{H}) ST_{\text{ell}}^{\mathbf{H}}(f^{\mathbf{H}}). \end{aligned}$$

Here  $\iota(\mathbf{G}, \mathbf{H})$  is by definition  $\tau(\mathbf{G}) \tau(\mathbf{H})^{-1} |\text{Out}((\mathbf{H}, s, \eta))|^{-1}$ . The proof of Theorem 1.5 is complete.

*Remark 4.3.* The stable orbital integrals of the function  $f^{\mathbf{H}}$  depend only on the isomorphism class of  $(\mathbf{H}, s, \eta)$ . For instance suppose that  $(\mathbf{H}, s, \eta)$  is replaced with an isomorphic triple  $(\mathbf{H}, sz, \eta)$  with  $z \in Z(\widehat{\mathbf{G}})$ . Then  $f^{\mathbf{H}, p}$  does not change whereas  $f_{\tau, h}^{\mathbf{H}}$  and  $f_{\xi}^{\mathbf{H}}$  are multiplied by  $\mu_1(z)$  and  $\mu_1^{-1}(z)$ , respectively.

### 5. PSEUDOSTABILIZATION

In this section, we employ the easier process of pseudostabilization as in [29] that works for certain compact unitary group Shimura varieties with trivial endoscopy. It has the advantage that one can avoid mention of endoscopy almost completely.<sup>7</sup>

We recall the setup; cf. [29] and Section 8.1.1.3 in [15]. We start with a CM field  $F$  with totally real subfield  $F_0$ . Let  $D/F$  be a central division algebra of dimension  $n^2$  with an involution  $*$  that restricts to complex conjugation on  $F$ . We want to construct PEL data with  $C = D$ , compatible with  $*$ . First, we set  $B = D^{\text{op}}$ , acting

<sup>7</sup>In particular, all of our results in the nonendoscopic case are independent of recent work on the fundamental lemma. Only the base-change fundamental lemma by Kottwitz, Clozel and Labesse is needed.

on  $V = B$  from the left. In order to define the pairing on  $V$ , we choose some  $b \in B^\times$  with  $b^* = b$  such that the pairing

$$(x, y) = \text{tr}_B(xby^*)$$

gives a positive pairing on  $B \otimes_{\mathbb{Q}} \mathbb{R}$ ; this is possible because the set of all  $b \in B^{*=1} \otimes \mathbb{R}$  with this property is an open cone in  $B^{*=1}$  by Lemma 2.8 in [30]. We endow  $B$  with the involution  $x \mapsto x^\sharp = bx^*b^{-1}$ ; the previous remarks show that this is a positive involution, as required. Now take some  $\beta \in F$  with  $\beta^\sharp = -\beta$  and endow  $V$  with the inner product

$$\langle x, y \rangle = \text{tr}_B(xb^{-1}\beta y^\sharp).$$

One checks that starting with  $(B, \sharp)$ , this makes  $V$  a left  $B$ -module with a non-degenerate  $\sharp$ -Hermitian form, and  $C = \text{End}_B(V) = D$ , inducing the involution  $*$  on  $C = D$ . Finally, we take some  $*$ -homomorphism  $\mathbf{h}_0 : \mathbb{C} \rightarrow D \otimes \mathbb{R}$  such that  $\langle x, \mathbf{h}_0(i)y \rangle$  is a positive pairing on  $V$ .

This defines global PEL data. Let  $E$  be the associated reflex field,  $p$  a prime satisfying our local assumptions; in particular, we have additional integral data at  $p$ , and the group  $\mathbf{G}$ , defined over  $\mathbb{Q}$  by

$$\mathbf{G}(R) = \{x \in D \otimes_{\mathbb{Q}} R \mid xx^* \in R^\times\},$$

has a model over  $\mathbb{Z}_{(p)}$ . Also, let  $\mathfrak{p}$  be a place of  $E$  above  $p$ , and let  $\xi$  be an algebraic representation of  $\mathbf{G}$  over a number field  $L$ ; let  $\lambda$  be a place of  $L$  lying above  $\ell \neq p$ .

**Theorem 5.1.** *Let  $\tau, h$  and  $f^p$  be as usual. Then*

$$N \text{tr}(\tau \times hf^p | H_\xi^*) = \text{tr}(f_{\tau,h}^{\mathbf{G}} f^p | H_\xi^*).$$

Here  $N = |\Pi^2(\xi)| \cdot |\pi_0(G(\mathbb{R})/Z(G)(\mathbb{R}))|$  (cf. [29, p. 659]) and  $f_{\tau,h}^{\mathbf{G}}$  is the endoscopic transfer of  $\phi_{\tau,h}$  for the trivial endoscopic triple  $(\mathbf{H}, s, \eta) = (\mathbf{G}, 1, \text{id})$ .

*Proof.* This follows from following the arguments in [29]. The crucial point in comparison to the general case is that  $\mathfrak{K}(I_0/\mathbb{Q})$  is trivial for all Kottwitz triples  $(\gamma_0; \gamma, \delta)$ . This makes the sum over  $\kappa$  in (12) superfluous, and one can factor everything into stable orbital integrals directly. The outcome is (cf. [29, (5.2)]) that the left-hand side equals

$$N\tau(\mathbf{G}) \sum_{\gamma_0} SO_{\gamma_0}(f_{\tau,h}^{\mathbf{G}} f^p f_\infty),$$

where  $\gamma_0$  runs through stable conjugacy classes in  $\mathbf{G}(\mathbb{Q})$ . Applying the trace formula (and Lemma 4.1 of [29]), this can be rewritten as

$$N \sum_{\pi} m(\pi) \text{tr} \pi(f_{\tau,h}^{\mathbf{G}} f^p f_\infty).$$

Now one applies Lemma 4.2 of [29] to rewrite this as

$$\sum_{\pi_f} \text{tr} \pi_f(f_{\tau,h}^{\mathbf{G}} f^p) \sum_{\pi_\infty} m(\pi_f \otimes \pi_\infty) \text{ep}(\pi_\infty \otimes \xi),$$

with the notation  $\text{ep}$  denoting the Euler-Poincaré characteristics of the Lie algebra cohomology as in [29, p. 660]. Finally, one uses Matsushima’s formula (cf. [29, p. 655]) to rewrite this as

$$\text{tr}(f_{\tau,h}^{\mathbf{G}} f^p | H_\xi^*),$$

as desired. □

In general one would expect that the following statement is true at places  $p$ , such that all places of  $F_0$  above  $p$  are split in  $F$ : If  $\pi_f, \pi'_f$  are two irreducible admissible representations of  $\mathbf{G}(\mathbb{A}_f)$  such that the  $\pi_f$ -isotypic and  $\pi'_f$ -isotypic components of  $H_\xi^*$  are nonzero and the representations are isomorphic away from  $p$ , i.e.  $\pi_f^p \cong \pi'_f{}^p$ , then also  $\pi_p \cong \pi'_p$ .

Under some extra conditions, this can be deduced from results of Harris and Labesse, [16], which show the existence of a stable base-change lift of automorphic representations  $\pi$  of  $\mathbf{G}$  to automorphic representations of  $\mathrm{GL}_n$  over  $F$ ; cf. [15, Section A.3].

In particular, assume that  $F$  is the composite of  $F_0$  and an imaginary quadratic field  $\mathcal{K}$ . Note that if  $q$  is a rational prime that splits in  $\mathcal{K}$  as  $w_0 w_0^c$ , there is a decomposition

$$\mathbf{G}(\mathbb{Q}_p) = \prod_{v|w_0} \mathrm{GL}_n(F_v) \times \mathbb{Q}_p^\times,$$

which induces a decomposition

$$\pi_q = \bigotimes_{v|w_0} \pi_v \otimes \chi_{w_0}$$

for every irreducible admissible representation  $\pi_q$  of  $\mathbf{G}(\mathbb{Q}_p)$ .

**Theorem 5.2.** *Assume that  $F = F_0\mathcal{K}$  is the composite of the totally real field  $F_0$  and an imaginary quadratic field  $\mathcal{K}$  in which  $p$  splits. Assume that the division algebra  $D$  is split at all places of  $F_0$  which do not split in  $F$ , and at all non-Archimedean places either split or a division algebra. Let  $\pi_f$  be an irreducible admissible representation of  $\mathbf{G}(\mathbb{A}_f)$  that occurs in  $H_\xi^*$ , and such that there is some rational prime  $q = w_0 w_0^c$  split in  $\mathcal{K}$  such that in the decomposition*

$$\pi_q = \bigotimes_{v|w_0} \pi_v \otimes \chi_{w_0},$$

one component  $\pi_v$  is supercuspidal.

Then there exists a cuspidal automorphic representation  $\Pi$  of  $\mathrm{GL}_n(\mathbb{A}_F)$  and a Hecke character  $\chi$  of  $\mathbb{A}_\mathcal{K}^\times$  such that  $\Pi_\infty$  is regular algebraic,  $\Pi^\vee \cong \Pi \circ c$ , where  $c : F \rightarrow F$  is complex conjugation, and for all places  $p = u_0 u_0^c$  split in  $\mathcal{K}$ , the decomposition of  $\pi_p$  from above is given by

$$\pi_p = \bigotimes_{v|u_0} \Pi_v \otimes \chi_{u_0}.$$

A similar description holds true at unramified inert places.

*Proof.* This follows from Theorem A.3.1 in [15]. □

**Corollary 5.3.** *In the situation of the theorem, let  $p \neq q$  be split in  $\mathcal{K}$  and let  $\pi'_f$  be a second irreducible admissible representation of  $\mathbf{G}(\mathbb{A}_f)$  which occurs in  $H_\xi^*$  and for which  $\pi_f^p \cong \pi'_f{}^p$ . Then  $\pi_p \cong \pi'_p$ .*

*Proof.* Use strong multiplicity 1 for  $\mathrm{GL}_n$  and the quadratic base-change for  $\mathbf{G}$  in the nonendoscopic case; cf. Corollary VI.2.3 of [17]. □

Moreover, the results of Harris-Taylor, [18, Theorem C], prove the existence of  $\ell$ -adic Galois representations associated to  $\Pi$  that satisfy a local-global compatibility

at all finite places not dividing  $\ell$ . We remark that for our purposes, the slightly weaker statement proven in [41] would suffice.

**Corollary 5.4.** *In the situation of the theorem, there is an  $\ell$ -adic Galois representation*

$$R_\ell(\Pi) : \text{Gal}(\bar{F}/F) \rightarrow \text{GL}_n(\bar{\mathbb{Q}}_\ell)$$

such that for all primes  $v$  of  $F$  not dividing  $\ell$ , the restriction  $R_\ell(\Pi)|_{W_{F_v}}$  is given by  $\sigma_\ell(\Pi_v^\vee) \otimes |\cdot|^{(1-n)/2}$ , where  $\sigma_\ell$  denotes the local Langlands correspondence for  $\text{GL}_n(F_v)$ .  $\square$

### 6. THE STABLE BERNSTEIN CENTER

The most convenient way to state our conjecture is to use the conjectural theory of the stable Bernstein center; cf. [47]. The importance of the stable Bernstein center in relation to test functions appearing in generalizations of the Langlands-Kottwitz method has also been emphasized recently by Haines.

Let us recall the main definitions. Let  $G$  be a connected reductive group over a  $p$ -adic field  $F$ . In our convention varieties are not required to be of finite type or connected.

**Definition/Proposition 6.1** ([7]). A supercuspidal pair for  $G$  is a pair  $(M, \sigma)$ , where  $M \subset G$  is a Levi subgroup of  $G$ , and  $\sigma$  is a supercuspidal representation of  $M(F)$ . An irreducible smooth representation  $\pi$  of  $G(F)$  is said to have supercuspidal support  $(M, \sigma)$ , if it appears as a subquotient of the normalized parabolic induction of  $\sigma$  to  $G(F)$ ; for any  $\pi$ , there is a unique such  $(M, \sigma)$  up to conjugation by  $G(F)$ .

The variety  $\Omega(G)$  of all supercuspidal pairs  $(M, \sigma)$  up to conjugation by  $G(F)$  has a natural structure as the  $\mathbb{C}$ -valued points of an infinite disjoint union of algebraic varieties over  $\mathbb{C}$ , which are quotients of tori by finite group actions. Let  $\mathcal{Z}(G)$  be the ring of regular functions on  $\Omega(G)$ , the Bernstein center of  $G$ . Then there is a natural  $\mathcal{Z}(G)$ -algebra structure on  $C_c^\infty(G(F))$ , given by a convolution map

$$\mathcal{Z}(G) \times C_c^\infty(G(F)) \rightarrow C_c^\infty(G(F)) : (z, f) \mapsto z * f ,$$

such that for all  $f \in C_c^\infty(G(F))$ , all  $z \in \mathcal{Z}(G)$  and any irreducible  $\pi$  with supercuspidal support  $(M, \sigma)$ , we have

$$\text{tr}(z * f | \pi) = z(M, \sigma) \text{tr}(f | \pi) .$$

**Definition/Proposition 6.2.** A semisimple  $L$ -parameter<sup>8</sup> for  $G$  is a homomorphism  $\lambda : W_F \rightarrow {}^L G$  from the Weil group  $W_F$  of  $F$  into the  $L$ -group of  $G$  that commutes with the projection maps to the Galois group of  $F$ , and such that  $\lambda(W_F)$  consists of semisimple elements.

The variety  $\Omega^{\text{st}}(G)$  of all semisimple  $L$ -parameters for  $G$  up to  $\widehat{G}$ -conjugation has a natural structure as an infinite disjoint union of algebraic varieties over  $\mathbb{C}$ , which are quotients of tori by finite group actions. Let  $\mathcal{Z}^{\text{st}}(G)$  be the ring of regular functions on  $\Omega^{\text{st}}(G)$ , the stable Bernstein center of  $G$ .

---

<sup>8</sup>Vogan named it an infinitesimal character in [47], which unfortunately conflicts with the common usage of the term as a character of  $\mathcal{Z}(G)$ , which Vogan refers to as a classical infinitesimal character.



Now assume for motivation that the local Langlands conjecture is known. In that case, any irreducible  $\pi$  has an associated  $L$ -parameter  $\varphi_\pi : W_F \times \mathrm{SL}_2(\mathbb{C}) \rightarrow {}^L G$ , and the restriction  $\varphi_\pi|_{W_F}$  defines a semisimple  $L$ -parameter  $\lambda_\pi$ . Here, we embed  $W_F$  into  $W_F \times \mathrm{SL}_2(\mathbb{C})$  by using the identity map on the first factor, and the map

$$w \mapsto \begin{pmatrix} |w|_F^{1/2} & 0 \\ 0 & |w|_F^{-1/2} \end{pmatrix}$$

from  $W_F$  to  $\mathrm{SL}_2(\mathbb{C})$ , using the norm character  $|\cdot|_F : W_F \rightarrow \mathbb{R}_{>0}^\times$ . Conjecturally (cf. Conjecture 7.18 in [47]), this induces a finite map of algebraic varieties

$$\Omega(G) \rightarrow \Omega^{\mathrm{st}}(G) : (M, \sigma) \mapsto \lambda(M, \sigma)$$

such that whenever  $\pi$  has supercuspidal support  $(M, \sigma)$ , then  $\lambda_\pi = \lambda(M, \sigma)$ .

We will use the following form of the local Langlands conjecture. Cf. Remark 6.4 below.

**Conjecture 6.3.** *There is a natural finite map of algebraic varieties*

$$\Omega(G) \rightarrow \Omega^{\mathrm{st}}(G) .$$

*The induced map  $\mathcal{Z}^{\mathrm{st}}(G) \rightarrow \mathcal{Z}(G)$  has the following property: If  $f \in C_c^\infty(G(F))$  has vanishing stable orbital integrals, then for any  $z^{\mathrm{st}} \in \mathcal{Z}^{\mathrm{st}}(G)$  with image  $z \in \mathcal{Z}(G)$ , the convolution  $z * f \in C_c^\infty(G(F))$  has vanishing stable orbital integrals.*

The conjecture is true for (products of Weil restrictions of)  $\mathrm{GL}_n$ , by [17], [19]. In the case of  $GL_n$  one of us essentially constructed the map  $\mathcal{Z}^{\mathrm{st}}(G) \rightarrow \mathcal{Z}(G)$  directly from Lubin-Tate towers in [39].

Moreover, let us recall some statements from [47] about stable characters. Here we only explain the case where  $G$  is quasisplit and refer the reader to [47] or [1] for general  $G$ . The easiest case is the case of tempered  $L$ -parameters; so let

$$\varphi : W_F \times \mathrm{SL}_2(\mathbb{C}) \rightarrow {}^L G$$

be a tempered  $L$ -parameter. Let  $C_\varphi$  be the group of  $g \in \widehat{G}$  centralizing the image of  $\varphi$ . Define  $\mathfrak{S}_\varphi = \pi_0(C_\varphi/Z(\widehat{G})^{\mathrm{Gal}(\overline{F}/F)})$ .<sup>9</sup> It is expected that there exists a bijection  $\pi \mapsto r_\pi$  from the  $L$ -packet  $\Pi(\varphi)$  to the set of irreducible finite-dimensional representations of  $\mathfrak{S}_\varphi$ . Moreover, one expects that the distribution

$$S\Theta_\varphi = \sum_{\pi \in \Pi(\varphi)} \dim r_\pi \mathrm{tr} \pi$$

is stable. The bijection  $\pi \mapsto r_\pi$  is supposed to be canonical up to twist by a character, so that the stable character is independent of all choices.

In particular, in this case, a suitable linear combination of the characters of the representations in the  $L$ -packet induces a stable character. In general, this will not be the case, and one has to include some representations from other  $L$ -packets, as in the example of  $A$ -packets. However, all representations appearing will have the same *semisimple*  $L$ -parameter. Indeed, Vogan conjectures that there are many stable characters of the form

$$(15) \quad S\Theta = \sum_{\pi} a(\pi) \mathrm{tr} \pi ,$$

---

<sup>9</sup>Some authors use  $S_\varphi$  to denote our  $C_\varphi$  but Kottwitz uses  $S_\varphi$  differently; cf. [22, §10]. Our choice of notation is consistent with Kottwitz's.

where  $a(\pi)$  is nonzero for only finitely many  $\pi$ , and for all such  $\pi$ , the semisimple  $L$ -parameter  $\lambda = \lambda_\pi$  is the same. Let  $z^{\text{st}} \in \mathcal{Z}^{\text{st}}(G)$ , with associated  $z \in \mathcal{Z}(G)$ . Then for all  $f \in C_c^\infty(G(F))$ , we have

$$(16) \quad S\Theta(z * f) = z^{\text{st}}(\lambda)S\Theta(f).$$

Indeed, this is an immediate consequence of Conjecture 6.3 and the formula for the action of the Bernstein center on the Hecke algebra.

*Remark 6.4.* The work of Kazhdan and Varshavsky [20] suggests a harmonic analytic definition of the stable Bernstein center without reference to semisimple  $L$ -parameters.<sup>10</sup> Recall that  $\mathcal{Z}(G)$  can be viewed as the algebra of  $G(F)$ -invariant distributions  $\alpha$  on  $G(F)$  which preserve  $C_c^\infty(G(F))$  under the convolution  $*$ . Define  $\mathcal{Y}^{\text{st}}(G)$  to be the subspace of  $\mathcal{Z}(G)$  consisting of  $\delta$  such that  $\delta * f$  has vanishing stable orbital integrals for every  $f \in C_c^\infty(G(F))$  whose stable orbital integrals vanish. It is clear that  $\mathcal{Y}^{\text{st}}(G)$  is a subalgebra of  $\mathcal{Z}(G)$ . In this optic, Conjecture 6.3 would hold true if the following are verified:  $\mathcal{Z}(G)$  is a finitely generated  $\mathcal{Y}^{\text{st}}(G)$ -module and there is a natural surjection  $\mathcal{Z}^{\text{st}}(G) \rightarrow \mathcal{Y}^{\text{st}}(G)$ , which should be induced by an infinitesimal version of the local Langlands correspondence. If  $G$  is quasi-split, then the map  $\mathcal{Z}^{\text{st}}(G) \rightarrow \mathcal{Y}^{\text{st}}(G)$  should be an isomorphism. This anticipated isomorphism justifies our terminology for  $\mathcal{Z}^{\text{st}}(G)$ .

### 7. CHARACTER IDENTITIES SATISFIED BY $f_{\tau,h}^{\mathbf{H}}$ : CONJECTURES

The contents of this section are purely local in nature, and we go back to the local setting of [40] and Subsection 3.2. Let  $(\mathbf{H}, s, \eta)$  be an endoscopic triple for  $\mathbf{G}$ .

We will state two versions of our conjecture. One relies on the local Langlands correspondence for  $\mathbf{H}$ , at least for tempered  $L$ -parameters. The other relies on Conjecture 6.3 about the stable Bernstein center, and is stronger.

Recall that we have a conjugacy class  $\bar{\mu}$  of cocharacters  $\mu : \mathbb{G}_m \rightarrow \mathbf{G}_{\overline{\mathbb{Q}}_p}$  as defined in Subsection 3.2. Fix a maximal torus  $\mathbf{T}$  of  $\mathbf{G}$  over  $\overline{\mathbb{Q}}_p$  with a set of positive roots. After conjugation we may assume that  $\mu$  factors through  $\mathbf{T}$  and is a dominant coweight. Write  $\rho$  for the half sum of all positive roots. Up to isomorphism there exists a unique complex representation  $r_{-\mu}$  of  ${}^L\mathbf{G}$  characterized by [21, Lem. 2.1.2]. In particular,  $r_{-\mu}|_{\widehat{\mathbf{G}}}$  is irreducible, with highest weight given by (the dominant representative of)  $-\mu$ . Let  $\text{Frob}$  denote a geometric Frobenius element in  $W_E$ .

**Conjecture 7.1.** *Let  $\tau \in \text{Frob}^j I_E$  for  $j \geq 1$  and  $h \in C_c^\infty(\mathbf{G}(\mathbb{Z}_p))$ . Denote by  $h^{\mathbf{H}}$  an  $\eta$ -transfer of  $h$  from  $\mathbf{G}$  to  $\mathbf{H}$ . For every tempered  $L$ -parameter  $\varphi_{\mathbf{H}}$  of  $\mathbf{H}$ , with associated semisimple  $L$ -parameter  $\lambda_{\mathbf{H}} = \varphi_{\mathbf{H}}|_{W_{\mathbb{Q}_p}}$ , we have*

$$(17) \quad S\Theta_{\varphi_{\mathbf{H}}}(f_{\tau,h}^{\mathbf{H}}) = \text{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda_{\mathbf{H}}|_{W_E})| \cdot |_E^{-(\rho,\mu)})S\Theta_{\varphi_{\mathbf{H}}}(h^{\mathbf{H}}).$$

Here is a simple consistency check. Let  $z \in Z(\widehat{\mathbf{G}})$ . Suppose that  $(\mathbf{H}, s, \eta)$  is replaced with  $(\mathbf{H}, sz, \eta)$  and that  $f_{\tau,h}^{\mathbf{H}}$  is constructed from the latter. Then both sides of (17) are multiplied by  $\mu_1(z)$  (cf. Remark 4.3).

It is expected that stable tempered characters are dense in the space of stable distributions; cf. [42, Conjecture 9.2]. Namely for any stable distribution  $\mathcal{D}$  on

<sup>10</sup>In fact our definition is not the same as the one in [20]. Though the two definitions are supposed to be equivalent, and this is the case for  $G = GL_n$ , it is difficult to check in general. Our definition has the advantage that  $\mathcal{Y}^{\text{st}}(G)$  is easily seen to be a subalgebra of  $\mathcal{Z}(G)$ .

$\mathbf{H}(\mathbb{Q}_p)$ , the value of  $\mathcal{D}(f_{\tau,h}^{\mathbf{H}})$  should be determined by (17) for tempered  $\varphi_{\mathbf{H}}$ . In particular, there should be an analogue of (17) for nontempered  $\varphi_{\mathbf{H}}$ . We feel that the best way to formulate this conjecture is to use the stable Bernstein center.

Let  $z_{\tau}^{\mathbf{H},\text{st}} \in \mathcal{Z}^{\text{st}}(\mathbf{H})$  be the function in the stable Bernstein center of  $\mathbf{H}$  that takes each semisimple  $L$ -parameter  $\lambda : W_{\mathbb{Q}_p} \rightarrow {}^L\mathbf{H}$  to

$$\text{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda|_{W_E})| \cdot |_E^{-\langle \rho, \mu \rangle}).$$

General requirements on the local Langlands correspondence ensure that this is a regular function on  $\Omega^{\text{st}}(\mathbf{H})$ . Let  $z_{\tau}^{\mathbf{H}} \in \mathcal{Z}(\mathbf{H})$  be its image.

**Conjecture 7.2.** *Let  $\tau$  and  $h$  be as above. Then for any  $\eta$ -endoscopic transfer  $h^{\mathbf{H}}$  of  $h$ , the function  $z_{\tau}^{\mathbf{H}} * h^{\mathbf{H}}$  is a twisted endoscopic  $\tilde{\eta}$ -transfer of  $\phi_{\tau,h}$ .*

This conjecture says that the twisted endoscopic transfers of  $\phi_{\tau,h}$  factor as a product of two factors, one depending only on  $\tau$ , the other depending only on  $h$ . This property by itself would be of great interest.

As a consequence of this conjecture, assume that  $S\Theta$  is a stable distribution as in (15). Let  $f_{\tau,h}^{\mathbf{H}}$  denote any transfer of  $\phi_{\tau,h}$ . Then

$$(18) \quad S\Theta(f_{\tau,h}^{\mathbf{H}}) = \text{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda|_{W_E})| \cdot |_E^{-\langle \rho, \mu \rangle})S\Theta(h^{\mathbf{H}}).$$

This generalizes (17), and should apply in particular to the stable characters associated to  $A$ -packets.

If  $\mathbf{G}(\mathbb{Q}_p)$  is a product of general linear groups over finite extensions of  $\mathbb{Q}_p$ , as happens in all cases of EL or quasi-EL type (cf. Remark 8.2 below), then  $\mathbf{H}(\mathbb{Q}_p)$  is of the same type, and the relevant local Langlands conjecture is known by [17] and [19]. In this case the statements of Conjecture 7.1 and Conjecture 7.2 do not rely on any unverified conjecture and simplify as stability issues do not arise. It is an easy exercise to verify that the two conjectures are equivalent. The conjecture was shown to be true by one of us, [39], in the Lubin-Tate case. In §8 we will settle Conjecture 7.1 in all EL cases.

A version of this conjecture also makes sense in the unramified case without any assumption thanks to the unramified Langlands correspondence. Hence assume that we are in the case of unramified EL or PEL data. Given our assumptions, this amounts to the only additional requirement that  $F$  is unramified. In particular,  $\mathbf{G}$  is unramified over  $\mathbb{Q}_p$ . Suppose moreover that  $\mathbf{H}, \eta : {}^L\mathbf{H} \rightarrow {}^L\mathbf{G}$ , and  $h$  are unramified; in particular (up to a scalar)  $h = \mathbf{1}_{\mathbf{G}(\mathbb{Z}_p)}$ . If  $\varphi_{\mathbf{H}}$  is unramified, then we write  $\pi_{\varphi_{\mathbf{H}}}$  for the unramified representation corresponding to  $\varphi_{\mathbf{H}}$ . We define a character  $S\Theta_{\varphi_{\mathbf{H}}}$  on  $\mathcal{H}^{\text{ur}}(\mathbf{H}(\mathbb{Q}_p))$ , by setting  $S\Theta_{\varphi_{\mathbf{H}}}$  equal to  $\text{tr } \pi_{\varphi_{\mathbf{H}}}$  if  $\varphi_{\mathbf{H}}$  is unramified and 0 otherwise.

**Lemma 7.3.** *Assume that  $F$  is unramified, and that  $p \neq 2$  in the PEL case. One can choose  $f_{\tau,h}^{\mathbf{H}}, h^{\mathbf{H}} \in \mathcal{H}^{\text{ur}}(\mathbf{H}(\mathbb{Q}_p))$ , and for all  $L$ -parameters  $\varphi_{\mathbf{H}}$  with semisimple  $L$ -parameter  $\lambda_{\mathbf{H}}$ , we have*

$$(19) \quad S\Theta_{\varphi_{\mathbf{H}}}(f_{\tau,h}^{\mathbf{H}}) = \text{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda_{\mathbf{H}}|_{W_E})| \cdot |_E^{-\langle \rho, \mu \rangle})S\Theta_{\varphi_{\mathbf{H}}}(h^{\mathbf{H}}).$$

*Remark 7.4.* The lemma shows that (17) naturally extends the identity in the unramified case treated in [28].

*Proof.* From Proposition 4.7 of [40], we know that  $\phi_{\tau,h}$  has the same twisted orbital integrals as Kottwitz’s  $\phi_r$  in [28, p. 173] when  $h = \mathbf{1}_{\mathbf{G}(\mathbb{Z}_p)}$ . We may thus work with  $\phi_r$  instead. We note that  $\phi_r$  lies in the unramified Hecke algebra  $\mathcal{H}^{\text{ur}}(\mathbf{G}(\mathbb{Q}_{p^r}))$ .

The  $L$ -morphism (6) gives rise to the transfer  $\tilde{\eta}^* : \mathcal{H}^{\text{ur}}(\mathbf{G}(\mathbb{Q}_p)) \rightarrow \mathcal{H}^{\text{ur}}(\mathbf{H}(\mathbb{Q}_p))$  such that  $\phi \mapsto \tilde{\eta}^*(\phi)$  is an  $\tilde{\eta}$ -transfer with respect to the canonical transfer factor of §3.2 (cf. [35, A.1.5]).

In particular, we can choose  $f_{\tau,h}^{\mathbf{H}}, h^{\mathbf{H}} \in \mathcal{H}^{\text{ur}}(\mathbf{H}(\mathbb{Q}_p))$ , as desired. Moreover, both sides of (19) vanish unless  $\varphi_{\mathbf{H}}$  is unramified.

Now assume that  $\varphi_{\mathbf{H}}$  is unramified. Then the argument for (9.7) of [28], which is based on the Satake transform (thus unconditional), shows that

$$S\Theta_{\varphi_{\mathbf{H}}}(f_{\tau,h}^{\mathbf{H}}) = p^{r(\rho,\mu)} \text{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda_{\mathbf{H}}|_{W_E})),$$

which is equivalent to (19). Note that  $\Delta_p(\psi_{\mathbf{H}}, \pi_{\eta\varphi_{\mathbf{H}}}) = 1$  in (9.7) of [28] since the above transfer factor is compatible with the algebra map of unramified Hecke algebras. □

As the final topic of this section, and as further supporting evidence for the conjecture, we verify that it leads to the expected description of the cohomology. We switch back to global notation and assume that  $\mathbf{G}$  is anisotropic over  $\mathbb{Q}$  modulo its center so that our Shimura varieties are compact.<sup>11</sup>

The following are accepted on faith:

- (1) the  $A$ -packet classification for discrete automorphic representations of  $\mathbf{G}$  and its elliptic endoscopic groups over  $\mathbb{Q}$ , encompassing the multiplicity formula [28, (8.2)] and a product formula for spectral transfer factors [28, (9.6)],<sup>12</sup>
- (2)  $ST_{\text{ell}}^{\mathbf{H}}(f^{\mathbf{H}}) = ST_{\text{disc}}^{\mathbf{H}}(f^{\mathbf{H}})$  for all  $(\mathbf{H}, s, \eta) \in \mathcal{E}_{\text{ell}}(\mathbf{G})$  (cf. [28, (10.1)]).

Our argument copies the one in [28, §§9-10] with changes occurring only at  $p$ . Therefore we will only indicate the changes without repeating the details. All references are made to Kottwitz’s article.

- Use  $f^{\mathbf{H},p}, f_{\tau,h}^{\mathbf{H}}$  and  $f_{\xi}^{\mathbf{H}}$  in place of  $h^p, h_p$  and  $h_{\infty}$ .
- Drop the assumption that  $\psi_p$  is unramified. Now (9.3) is equal to (9.7) if the latter is replaced by

$$\sum_{\pi_p \in \Pi(\eta\psi_{\mathbf{H},p})} \Delta_p(\psi_{\mathbf{H}}, \pi_p) \text{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\varphi_{\mathbf{H}}|_{W_{E_p}})| \cdot |_{\mathfrak{p}}^{-\dim \text{Sh}/2}) \text{tr} \pi_p(h).$$

This key equality is justified by Conjecture 7.1, or more precisely its version for  $A$ -parameters.

- On page 198, drop conditions (a) and (c).
- (9.9) is valid if  $\text{tr} \pi_f(f)$  is changed to  $\text{tr} \pi_f^p(f^p)$  and the line below reads

$$A(x, \pi_p) = (-1)^{q(\mathbf{G})} \text{tr}(s^{-1}\tau|(r_{-\mu} \circ \varphi_{\pi_p}|_{W_{E_p}})| \cdot |_{\mathfrak{p}}^{-\dim \text{Sh}/2}) \text{tr} \pi_p(h).$$

- (10.4) holds if  $\text{tr} \pi_f(f)$  is replaced with  $\text{tr} \pi_f^p(f^p)$  and we set in the line below

$$A(\pi_p, \nu) = (-1)^{q(\mathbf{G})} \nu(s_{\psi}) \text{tr}(\tau|V_{\nu}| \cdot |_{\mathfrak{p}}^{-\dim \text{Sh}/2}) \text{tr} \pi_p(h).$$

---

<sup>11</sup>We also assume that the flat closure of the generic fibre in the integral models defined in [40] is proper.

<sup>12</sup>In particular, we assume that there is the global Langlands group. This should be avoidable in all PEL cases by using alternative global parameters as in [2].

- From (10.4) we obtain (cf. (10.5)) that  $\text{tr}(\tau \times hf^p | H_\xi^*)$  equals

$$\sum_{[\psi]} \sum_{\pi_f} \text{tr} \pi_f(hf^p) \sum_{\nu} (-1)^{q(\mathbf{G})} m(\pi_f, \nu) \nu(s_\psi) \text{tr}(\tau | V(\psi, \nu)).$$

- The above equality establishes that

$$H_\xi^* = \sum_{[\psi]} \sum_{\pi_f} \sum_{\nu} (-1)^{q(\mathbf{G})} m(\pi_f, \nu) \nu(s_\psi) (\pi_f \otimes V(\psi_p, \nu))$$

as virtual  $\mathbf{G}(\mathbb{A}_f^p) \times \mathbf{G}(\mathbb{Z}_p) \times W_{E_p}$ -representations, as desired.

Of course, one would expect that this identity stays true as  $\mathbf{G}(\mathbb{A}_f) \times W_{E_p}$ -representations. If  $\mathbf{G}_{\mathbb{Q}_p}$  is a product of Weil restrictions of general linear groups, then  $\pi_p$  should be determined by  $\pi_f^p$ , which easily implies that this identity extends to  $\mathbf{G}(\mathbb{A}_f) \times W_{E_p}$ . In the general case, the first obstruction is that if  $\pi_f$  and  $\pi'_f$  are two representations appearing in  $H_\xi^*$  such that  $\pi_f^p \cong \pi'^p_f$ , then it is not clear that  $\pi_p$  and  $\pi'_p$  belong to the same  $A$ -packet. Ignoring this obstacle, the question of whether one can isolate representations within a local  $A$ -packet by Hecke operators in  $\mathbf{G}(\mathbb{Z}_p)$  appears:

**Question 7.5.** Let  $G$  be a connected reductive group over a  $p$ -adic field  $F$ , and let  $K \subset G(F)$  be a special maximal compact subgroup. Let  $\psi$  be an  $A$ -parameter for  $G$ , with associated  $A$ -packet  $\Pi(\psi) = \{\pi_1, \dots, \pi_k\}$ . Are  $\pi_1|_K, \dots, \pi_k|_K$  linearly independent in the Grothendieck group of admissible representations of  $K$ ?

### 8. CHARACTER IDENTITIES SATISFIED BY $f_{\tau,h}^{\mathbf{H}}$ : PROOFS

Let us keep the local notation of the last section (except that global notation will appear in the proof of Theorem 8.1). Recall that the local Langlands correspondence for (a product of) general linear groups over a  $p$ -adic field is known thanks to Harris-Taylor and Henniart ([17], [19]). In what follows  $\varphi_\pi$  denotes the  $L$ -parameter attached to an irreducible representation  $\pi$  on such a group.

**Theorem 8.1.** *Conjecture 7.1 is true in all EL and quasi-EL cases for the trivial endoscopic triple  $(\mathbf{H}, s, \eta) = (\mathbf{G}, 1, \text{id})$ . In fact, for every irreducible representation  $\pi$  of  $\mathbf{G}(\mathbb{Q}_p)$  with semisimple  $L$ -parameter  $\lambda_\pi$ ,*

$$(20) \quad \text{tr}(f_{\tau,h}^{\mathbf{G}} | \pi) = \text{tr}(\tau | (r_{-\mu} \circ \lambda_\pi |_{W_E}) \cdot |_{E}^{-\langle \rho, \mu \rangle}) \text{tr}(h | \pi).$$

*Remark 8.2.* We recall that in [40], quasi-EL data were defined as those PEL data where the field  $F$  decomposes as  $F_0 \times F_0$  with the involution  $*$  acting by  $(x, y)^* = (y, x)$ . We note that there is a bijection between EL data and quasi-EL data, and that quasi-EL data come about by localization of global PEL data of type A at a split place.

*Proof.* The conjectures in an EL case and the corresponding quasi-EL case are equivalent by Proposition 4.10 of [40]. Moreover, the conjecture in the EL case reduces to the simple EL case by Proposition 4.9 of [40]. By Morita equivalence, we may moreover assume that  $V = B$  as left  $B$ -module, where  $B$  is a matrix algebra over  $F$ .

Summarizing, we may assume that we are in the quasi-EL case corresponding to a simple EL case, where  $B = B_0 \times B_0$  is a matrix algebra over  $F = F_0 \times F_0$ , following the notation for the quasi-EL case in [40]. The involution on  $B$  is given by  $(x, y)^* = (y, x)$ , and  $V = B$  as left  $B$ -module.

For the proof we may fix  $\tau$  and  $h$ . Also fix an algebraic representation  $\xi$  of  $\mathbf{G}$  of regular highest weight. As our proof will be global in nature, we will adopt global notation as needed. Let us globalize the local datum over  $\mathbb{Q}_p$  to a global datum  $(B, \mathcal{O}_B, F, V, \Lambda, C, \mathcal{O}_C, \mathbf{h})$  and a  $*$ -Hermitian form on  $V$  for a compact unitary Shimura variety with trivial endoscopy as in Section 5, satisfying the assumptions of Theorem 5.2 as done in [15, Prop. 8.1.3]. In particular  $F$  is a CM field containing an imaginary quadratic field  $\mathcal{K}$ ,  $p$  splits in  $\mathcal{K}$ ,  $F$  decomposes as the product of its totally real subfield  $F_0$  and  $\mathcal{K}$ , and  $p$  is inert in  $F_0$ , with its local field  $F_{0,p}$  being isomorphic to what was previously called  $F_0$ . After base-change from  $\mathbb{Q}$  to  $\mathbb{Q}_p$ , we get the previous local PEL data and there exists a prime  $\mathfrak{p}$  of the reflex field  $E$  of the datum such that the completion  $E_{\mathfrak{p}}$  is isomorphic to the local reflex field  $E$  occurring in (20), compatible with the cocharacter  $\mu$ . Let  $\mathbf{G}$  denote the reductive group over  $\mathbb{Z}_{(p)}$  for the global datum, and  $\pi_p$  denote  $\pi$  of (20).

The first step is to verify (20) assuming

- $\pi_p$  appears as the  $p$ -component of an irreducible automorphic representation  $\pi$  of  $\mathbf{G}(\mathbb{A})$  such that  $\pi_{\infty} \in \Pi^2(\varphi_{\xi})$  and  $\pi_v$  is supercuspidal for some prime  $q \neq p$  split in  $\mathcal{K}$  and some place  $v|q$  of  $F$ .

Let

$$H_{\xi}^*(\pi_f) = \text{Hom}_{\mathbf{G}(\mathbb{A}_f)}(\pi_f, H_{\xi}^*).$$

Using Corollary 5.3, it follows from Theorem 5.1 that

$$(21) \quad N \text{tr}(\tau|H_{\xi}^*(\pi_f)) \text{tr}(h|\pi_p) = \dim H_{\xi}^*(\pi_f) \text{tr}(f_{\tau,h}^{\mathbf{G}}|\pi_p).$$

If  $\pi_p$  is unramified, one may take  $h = \mathbf{1}_{\mathbf{G}(\mathbb{Z}_p)}$  and use Lemma 7.3 to determine the restriction  $H_{\xi}^*(\pi_f)|_{W_{E_{\mathfrak{p}}}}$ ; cf. Theorem 1 in [29]. This gives

$$(22) \quad H_{\xi}^*(\pi_f)|_{W_{E_{\mathfrak{p}}}} = a(\pi_f)(r_{-\mu} \circ \lambda_{\pi_p}|_{W_{E_{\mathfrak{p}}}}) \cdot |_{\mathfrak{p}}^{-\dim \text{Sh}/2}$$

for some integer  $a(\pi_f) \in \mathbb{Z}$ . This equation implies in particular that  $N = \dim r_{-\mu}$  and  $\dim H_{\xi}^*(\pi_f) = a(\pi_f)N$ . We also recall that  $\dim \text{Sh} = \langle 2\rho, \mu \rangle$ . This argument works for any unramified split place  $p$  (not only the one fixed in advance). Using Corollary 5.4, one can build a representation of  $\text{Gal}(\bar{\mathbb{Q}}/EK)$  that agrees with  $H_{\xi}^*(\pi_f)|_{\text{Gal}(\bar{\mathbb{Q}}/EK)}$  at all unramified split places; we refer to the proof of [15, Thm. A.7.2] for the precise construction. From the Chebotarev density theorem, it follows that they agree everywhere, which shows that (22) continues to hold at ramified split places, in particular at our given  $\mathfrak{p}$ .

Now one may combine (21) with (22) to deduce (20).

It remains to get rid of the hypothesis on  $\pi_p$ . Consider the two linear forms  $\mathcal{F}_1$  (resp.  $\mathcal{F}_2$ ) on the Grothendieck group of admissible representations of  $\mathbf{G}(\mathbb{Q}_p)$  mapping each irreducible  $\pi_p$  to the value of the left-hand (resp. right-hand) side of (20). It is easy to see that both  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are trace functions in the context of the trace Paley-Wiener theorem of [6]. (This is a tautology for  $\mathcal{F}_1$ .) Since [43, Thm. 4.4] shows that a dense subset of the unitary dual of  $\mathbf{G}(\mathbb{Q}_p)$  satisfies the above hypothesis, we know that  $\mathcal{F}_1 = \mathcal{F}_2$  on that dense subset. Since  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are regular functions on (an infinite disjoint union of) algebraic varieties, we deduce that  $\mathcal{F}_1(\pi_p) = \mathcal{F}_2(\pi_p)$  for all  $\pi_p$ .  $\square$

At this point, let us give a corollary of the proof. We note that the identities of Conjecture 7.1 determine the orbital integrals of  $f_{\tau,h}^{\mathbf{G}}$  uniquely. On the other hand, the orbital integrals of  $f_{\tau,h}^{\mathbf{G}}$  determine the twisted orbital integrals of  $\phi_{\tau,h}$  uniquely.

But (cf. the remarks after Theorem 5.7 of [40]), there was some freedom in the choice of  $\phi_{\tau,h}$ . This gives the following corollary.

**Corollary 8.3.** *Let  $\text{Sh}$  be a compact unitary group Shimura variety without endoscopy as in Section 5, satisfying the hypothesis of Theorem 5.2. Let  $p$  be a place split in  $\mathcal{K}$ , let  $\mathfrak{p}|p$  be a place of  $E$  and let  $\mathcal{M}_{K^p}/\mathcal{O}_{E_{\mathfrak{p}}}$  be the integral model of (a finite disjoint union of copies of)  $\text{Sh}_{\mathbf{G}(\mathbb{Z}_p)K^p}$  over  $\mathcal{O}_{E_{\mathfrak{p}}}$  defined in [40].*

*Let  $\mathcal{D}$  be the corresponding local data of quasi-EL type, associated to data  $\mathcal{D}_0$  of EL type. Then any  $p$ -divisible group  $\overline{H}$  with  $\mathcal{D}$ -structure over  $\overline{\mathbb{F}}_p$  which admits a deformation to characteristic 0 occurs as the  $p$ -divisible group with  $\mathcal{D}$ -structure associated to some point over  $\overline{\mathbb{F}}_p$  on  $\mathcal{M}_{K^p}$ , for some  $K^p$ .*

*Proof.* Note that a  $p$ -divisible group  $\overline{H}$  with  $\mathcal{D}$ -structure which admits a deformation to characteristic 0, i.e. such that  $X_{\overline{H}} \neq \emptyset$ , is parametrized by some  $\delta \in \mathbf{G}(W(\overline{\mathbb{F}}_p)[\frac{1}{p}])$ ; cf. Proposition 3.10 of [40].

We can define two functions  $\phi_{\tau,h}^{(0)}$  and  $\phi_{\tau,h}^{(1)}$ , both satisfying the conditions after Theorem 5.7 of [40], but such that  $\phi_{\tau,h}^{(0)}$ , resp.  $\phi_{\tau,h}^{(1)}$ , is 0, resp. 1, on all  $\delta$  for which nothing is imposed by these conditions. We want to prove that  $\phi_{\tau,h}^{(0)} = \phi_{\tau,h}^{(1)}$ .

Now by what we have shown, necessarily all twisted orbital integrals of  $\phi_{\tau,h}^{(1)} - \phi_{\tau,h}^{(0)}$  vanish. But this is the characteristic function of some open bounded subset, and it follows that this subset has to be trivial.  $\square$

In the case that the local PEL data are unramified, this implies nonemptiness of Newton strata. Recall that the Newton stratification is a stratification of  $\mathcal{M}_{K^p} \otimes \overline{\mathbb{F}}_p$  into locally closed subvarieties  $\mathcal{M}_{K^p,b}$ ,  $b \in B(\mathbf{G}_{\mathbb{Q}_p}, \mu)$ , parametrized by a certain subset of  $\sigma$ -conjugacy classes  $B(\mathbf{G}_{\mathbb{Q}_p}, \mu) \subset B(\mathbf{G}_{\mathbb{Q}_p})$ ; cf. the work of Rapoport and Richartz, [38]. The question of nonemptiness of Newton strata is a deep question in the theory of the reduction of Shimura varieties. We refer to the survey article of Rapoport, [37], for an extensive discussion of related work. Recently, Viehmann and Wedhorn, [46], have used local geometric methods to prove nonemptiness of Newton strata in many cases, including those that we handle here. Our method is entirely different, in that it uses global and automorphic arguments, similar to earlier work of Fargues, [15], showing that the basic locus is nonempty.

**Corollary 8.4.** *Assume that in the situation of Corollary 8.3, the field  $F$  is unramified at  $p$ , so that  $\mathbf{G}$  is unramified at  $p$  and  $\mathbf{G}(\mathbb{Z}_p)$  is a hyperspecial maximal compact subgroup. Then for all  $b \in B(\mathbf{G}_{\mathbb{Q}_p}, \mu)$ , the Newton stratum  $\mathcal{M}_{K^p,b}$  is nonempty.*

*Proof.* We only have to show that there is some  $p$ -divisible group with  $\mathcal{D}$ -structure that has the correct Newton polygon. In the notation of [37], this question is equivalent to the question  $X(b, \mu)_K \neq \emptyset$  by Corollary 3.12 of [37]. This is resolved by a theorem of Kottwitz and Rapoport in the case considered here, [25]; cf. Theorem 5.5 in [37].  $\square$

Another interesting application of Corollary 8.3 is the following algebraization result that works in the ramified case and seems to be hard to prove by other methods.

**Corollary 8.5.** *Let  $\mathcal{D}$  be data of EL type, and let  $\overline{\mathbf{H}}$  be any  $p$ -divisible group with  $\mathcal{D}$ -structure over  $\overline{\mathbb{F}}_p$ . Then the deformation spaces  $X_{\overline{\mathbf{H}},K}/X_{\overline{\mathbf{H}}}$  defined in [40] are algebraizable as in Theorem 2.9 of [40].*

*Proof.* It suffices to consider the case of simple EL data; as in the general case, everything decomposes into a product. By Morita equivalence, we can assume that  $V = B$  again. Changing to quasi-EL data, we may use Corollary 8.3 to embed the situation into a Shimura variety, which provides the desired algebraization.  $\square$

Now we continue with the proof of Conjecture 7.1 in EL cases.

**Corollary 8.6.** *Conjecture 7.1 is true in all EL and quasi-EL cases for general endoscopic triples  $(\mathbf{H}, s, \eta)$  of  $\mathbf{G}$ ; i.e., for all irreducible representations  $\pi_{\mathbf{H}}$  of  $\mathbf{H}(\mathbb{Q}_p)$  with semisimple  $L$ -parameter  $\lambda_{\pi_{\mathbf{H}}}$ , we have*

$$\mathrm{tr}(f_{\tau,h}^{\mathbf{H}}|\pi_{\mathbf{H}}) = \mathrm{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda_{\pi_{\mathbf{H}}}|_{W_E})| \cdot |E|^{-\langle \rho, \mu \rangle}) \mathrm{tr}(h^{\mathbf{H}}|\pi_{\mathbf{H}}) .$$

*Proof.* Let  $K$  be a finite extension of  $\mathbb{Q}_p$ . It suffices to consider the case of simple EL data with  $B = K$ ,  $V = K^n$ , where  $\mathbf{G} \simeq \mathrm{Res}_{K/\mathbb{Q}_p} \mathrm{GL}_n$ . In this case, any endoscopic triple  $(\mathbf{H}, s, \eta)$  of  $\mathbf{G}$  is of the form  $\mathbf{H} \simeq \mathrm{Res}_{K/\mathbb{Q}_p} (\prod_i \mathrm{GL}_{n_i})$  for some  $n_i \geq 1$  such that  $\sum_i n_i = n$  and

$$\eta : \left( \prod_i \mathrm{GL}_{n_i}(\mathbb{C}) \right)^{\mathrm{Hom}(K, \overline{\mathbb{Q}}_p)} \rtimes W_{\mathbb{Q}_p} \rightarrow \mathrm{GL}_n(\mathbb{C})^{\mathrm{Hom}(K, \overline{\mathbb{Q}}_p)} \rtimes W_{\mathbb{Q}_p}$$

is the canonical Levi embedding; i.e., we identify  $\mathbf{H}$  and  $\widehat{\mathbf{H}}$  with Levi subgroups of  $\mathbf{G}$  and  $\widehat{\mathbf{G}}$  via block diagonal embeddings. Moreover we may assume that  $s \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}$  as the general case is reduced to this case (cf. the last paragraph of §3.2). Write  $s = (s_i)_i \in Z(\widehat{\mathbf{H}})^{\Gamma(p)} = \prod_i \mathbb{C}^\times$  (which embeds in the center of  $(\prod_i \mathrm{GL}_{n_i}(\mathbb{C}))^{\mathrm{Hom}(K, \overline{\mathbb{Q}}_p)}$  diagonally). Choose  $t_0 = (t_{0,i})_i \in Z(\widehat{\mathbf{H}})^{\Gamma(p)}$  such that  $t_0^\circ = s$ . Let  $c : W_{\mathbb{Q}_p} \rightarrow Z(\widehat{\mathbf{H}})^{\Gamma(p)}$  be the cocycle  $w \mapsto t_0^{-\mathrm{ord}(w)}$ , where we normalize  $\mathrm{ord} : W_E \rightarrow \mathbb{Z}$  by requiring that it sends a geometric Frobenius element to 1. Define  $\eta' : {}^L\mathbf{H} \rightarrow {}^L\mathbf{G}$  by  $\eta'(g \rtimes w) = c(w)\eta(g \rtimes w)$ . Recall that  $R_r$  and  $\tilde{\eta}$  were defined in §3.2. Write  $\xi : {}^L\mathbf{G} \rightarrow {}^L R_r$  for the base-change  $L$ -morphism sending  $g \rtimes w$  to  $(g, g, \dots, g) \rtimes w$ . Then it is straightforward to verify that

$$(23) \quad \tilde{\eta} = \xi \circ \eta' .$$

The function  $f_{\tau,h}^{\mathbf{H}}$  is a twisted transfer of  $\phi_{\tau,h}$  relative to  $\tilde{\eta}$ . We claim that it is also an  $\eta'$ -transfer of  $f_{\tau,h}^{\mathbf{G}}$  (denoted  $(\eta')^*(f_{\tau,h}^{\mathbf{G}})$  below).

Let us accept the claim for now and finish the proof. Note that an  $\eta$ -transfer of  $f_{\tau,h}^{\mathbf{G}}$ , denoted  $\eta^*(f_{\tau,h}^{\mathbf{G}})$ , is given by the constant term along a parabolic subgroup with  $\mathbf{H}$  as a Levi component. It satisfies a character identity

$$\mathrm{tr}(\eta^*(f_{\tau,h}^{\mathbf{G}})|\pi_{\mathbf{H}}) = \mathrm{tr} \left( f_{\tau,h}^{\mathbf{G}} \Big| \mathrm{n}\text{-ind}_{\mathbf{H}(\mathbb{Q}_p)}^{\mathbf{G}(\mathbb{Q}_p)}(\pi_{\mathbf{H}}) \right)$$

for every irreducible admissible representation  $\pi_{\mathbf{H}}$  of  $\mathbf{H}(\mathbb{Q}_p)$ , where  $\mathrm{n}\text{-ind}$  denotes the normalized parabolic induction (independent of the choice of the parabolic). Let  $\chi_c : H(\mathbb{Q}_p) \rightarrow \mathbb{C}^\times$  be the character corresponding to  $c \in H^1(W_{\mathbb{Q}_p}, Z(\widehat{\mathbf{H}})^{\Gamma(p)})$ .

Since  $\eta' = \eta \cdot c$ , one can show that  $(\eta')^* f_{\tau,h}^{\mathbf{G}}$  may be given as  $\chi_c^{-1} \cdot \eta^*(f_{\tau,h}^{\mathbf{G}})$ . (The reason is that the Langlands-Shelstad transfer factor gets multiplied by  $\chi_c^{-1}$  when  $\eta$  is multiplied by  $c$ .) As we can take  $f_{\tau,h}^{\mathbf{H}} = (\eta')^* f_{\tau,h}^{\mathbf{G}}$  by the claim above,

$$f_{\tau,h}^{\mathbf{H}}(\gamma) = \eta^*(f_{\tau,h}^{\mathbf{G}})(\gamma)\chi_c^{-1}(\gamma), \quad \forall \gamma \in \mathbf{H}(\mathbb{Q}_p) .$$



To put it in an invariant way,

$$\mathrm{tr}(f_{\tau,h}^{\mathbf{H}}|\pi_{\mathbf{H}}) = \mathrm{tr}(\eta^*(f_{\tau,h}^{\mathbf{G}})|\pi_{\mathbf{H}} \otimes \chi_c) = \mathrm{tr}\left(f_{\tau,h}^{\mathbf{G}} \Big| \mathrm{n}\text{-ind}_{\mathbf{H}(\mathbb{Q}_p)}^{\mathbf{G}(\mathbb{Q}_p)}(\pi_{\mathbf{H}} \otimes \chi_c)\right).$$

Note that all subquotients  $\pi$  of  $\mathrm{n}\text{-ind}_{\mathbf{H}(\mathbb{Q}_p)}^{\mathbf{G}(\mathbb{Q}_p)}(\pi_{\mathbf{H}} \otimes \chi_c)$  give rise to the same semisimple  $L$ -parameter  $\lambda_{\pi} = \eta\lambda_{\pi_{\mathbf{H}} \otimes \chi_c}$ . Hence, by (20) the right-hand side equals

$$\begin{aligned} & \mathrm{tr}(\tau|(r_{-\mu} \circ \eta\lambda_{\pi_{\mathbf{H}} \otimes \chi_c}|_{W_E})) \mathrm{tr}\left(h \Big| \mathrm{n}\text{-ind}_{\mathbf{H}(\mathbb{Q}_p)}^{\mathbf{G}(\mathbb{Q}_p)}(\pi_{\mathbf{H}} \otimes \chi_c)\right) \\ &= \mathrm{tr}(\tau|(r_{-\mu} \circ \eta\lambda_{\pi_{\mathbf{H}} \otimes \chi_c}|_{W_E})) \mathrm{tr}(\eta^*(h)|\pi_{\mathbf{H}} \otimes \chi_c). \end{aligned}$$

We recall that  $c(\tau) = t_0^{-\mathrm{ord}(\tau)} = t_0^{-r} = s^{-1}$ , from which it is not difficult to deduce that the twist by  $\chi_c$  has the same effect as replacing  $\tau$  by  $s^{-1}\tau$ :

$$\mathrm{tr}(\tau|(r_{-\mu} \circ \eta\lambda_{\pi_{\mathbf{H}} \otimes \chi_c}|_{W_E})) = \mathrm{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda_{\pi_{\mathbf{H}}}|_{W_E})).$$

On the other hand, note that

$$\mathrm{tr}(\eta^*(h)|\pi_{\mathbf{H}} \otimes \chi_c) = \mathrm{tr}(\eta^*(h)|\pi_{\mathbf{H}}).$$

Indeed, the constant term  $\eta^*(h)$  is supported only on the conjugacy classes of  $\mathbf{H}(\mathbb{Q}_p)$  meeting  $\mathbf{H}(\mathbb{Z}_p)$ , since  $h$  is supported on  $\mathbf{G}(\mathbb{Z}_p)$ . Since the character  $\chi_c$  is trivial on such conjugacy classes (because  $c$  is an unramified cocycle), the identity follows.

Putting the above identities together, we conclude that

$$\mathrm{tr}(f_{\tau,h}^{\mathbf{H}}|\pi_{\mathbf{H}}) = \mathrm{tr}(s^{-1}\tau|(r_{-\mu} \circ \eta\lambda_{\pi_{\mathbf{H}}}|_{W_E})) \mathrm{tr}(\eta^*(h)|\pi_{\mathbf{H}}),$$

which is the equality (17) to be proved.

It remains to justify the unproven claim that  $f_{\tau,h}^{\mathbf{H}}$  is an  $\eta'$ -transfer of  $f_{\tau,h}^{\mathbf{G}}$ . For any  $(\mathbf{G}, \mathbf{H})$ -regular semisimple  $\gamma_{\mathbf{H}} \in \mathbf{H}(\mathbb{Q}_p)$ , if  $\gamma_{\mathbf{H}}$  is not a transfer of any twisted conjugacy class of  $\delta \in \mathbf{G}(\mathbb{Q}_{p^r})$ , then  $SO_{\gamma_{\mathbf{H}}}(f_{\tau,h}^{\mathbf{H}}) = SO_{\gamma_{\mathbf{H}}}((\eta')^* f_{\tau,h}^{\mathbf{G}}) = 0$ . From now on suppose that  $\gamma_{\mathbf{H}}$  is a transfer of some  $\delta$ . We note that in the cases that occur here, taking a stable (twisted) orbital integral amounts to multiplying the usual (twisted) orbital integral by a sign, given as the Kottwitz sign attached to the (twisted) centralizer of the element considered. By (9), (8),

$$SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{Q}_p)}(f_{\tau,h}^{\mathbf{H}}) = \Delta_p^{\eta}(\gamma_{\mathbf{H}}, \gamma) \langle \alpha(\gamma; \delta), s \rangle^{-1} STO_{\delta\sigma}(\phi_{\tau,h}).$$

The superscript  $\eta$  indicates that the transfer factor is relative to  $\eta$ . On the other hand, Theorem 3.1 and its analogue for the base change relative to  $\xi$  tell us that

$$SO_{\gamma_{\mathbf{H}}}^{\mathbf{H}(\mathbb{Q}_p)}((\eta')^* f_{\tau,h}^{\mathbf{G}}) = \Delta_p^{\eta'}(\gamma_{\mathbf{H}}, \gamma) SO_{\gamma}^{\mathbf{G}(\mathbb{Q}_p)}(f_{\tau,h}^{\mathbf{G}}) = \Delta_p^{\eta'}(\gamma_{\mathbf{H}}, \gamma) STO_{\delta\sigma}(\phi_{\tau,h}).$$

As we have noted above,  $\Delta_p^{\eta'}(\gamma_{\mathbf{H}}, \gamma) = \chi_c^{-1}(\gamma_{\mathbf{H}}) \Delta_p^{\eta}(\gamma_{\mathbf{H}}, \gamma)$ . Hence the proof boils down to showing that

$$(24) \quad \chi_c(\gamma_{\mathbf{H}}) = \langle \alpha(\gamma; \delta), s \rangle.$$

Define  $R_r^{\mathbf{H}} = \mathrm{Res}_{\mathbb{Q}_{p^r}/\mathbb{Q}_p} \mathbf{H}$ . We can find a  $\delta_{\mathbf{H}} \in R_r^{\mathbf{H}}(\mathbb{Q}_p) = \mathbf{H}(\mathbb{Q}_{p^r})$  which is  $\sigma$ -conjugate to  $\delta$  and transfers to  $\gamma_{\mathbf{H}}$  via base change for  $\mathbf{H}$ . We write  $\delta_{\mathbf{H}} = (\delta_{\mathbf{H},i})_i$  and  $\gamma_{\mathbf{H}} = (\gamma_{\mathbf{H},i})_i$  with respect to  $\mathbf{H} \simeq \mathrm{Res}_{K/\mathbb{Q}_p}(\prod_i GL_{n_i})$ . As usual  $v_p$  denotes the additive valuation on a  $p$ -adic field such that  $v_p(p) = 1$ . It can be seen from the definition of  $\chi_c$  that

$$\chi_c(\gamma_{\mathbf{H}}) = \prod_i t_{0,i}^{v_p(N_{K/\mathbb{Q}_p} \det \gamma_{\mathbf{H},i})}.$$

As  $\gamma_{\mathbf{H},i}$  is a norm of  $\delta_{\mathbf{H},i}$ , we get

$$\chi_c(\gamma_{\mathbf{H}}) = \prod_i s_i^{v_p(N_{K/\mathbb{Q}_p} \det \delta_{\mathbf{H},i})}.$$

In order to evaluate the right-hand side of (24), we begin by recalling the definition of  $\alpha(\gamma; \delta)$ . We are free to replace  $\delta$  by a  $\sigma$ -conjugate, and hence we assume that  $\delta = \delta_{\mathbf{H}}$ . Let  $L$  be the fraction field of  $W(\widehat{\mathbb{F}}_p)$ . Using Steinberg’s theorem, one finds  $d \in \mathbf{G}(L)$  such that  $N\delta = d\gamma_{\mathbf{H}}d^{-1}$ . In fact, it is possible to take  $d \in \mathbf{H}(L)$ , by choice of  $\delta$ , and as  $\gamma_{\mathbf{H}}$  is a base-change transfer of  $\delta_{\mathbf{H}}$ . In general,  $d^{-1}\delta d^\sigma \in I_0(L)$ , where  $I_0$  is the centralizer of  $\gamma$  in  $\mathbf{G}$ , and  $\alpha(\gamma; \delta)$  is defined as the image of  $d^{-1}\delta d^\sigma$  under Kottwitz’s map  $\kappa_{I_0} : B(I_0) \rightarrow X^*(Z(\widehat{I_0})^\Gamma)$ .

The Kottwitz maps  $\kappa_{I_0}$  and  $\kappa_{\mathbf{H}}$  fit into a commutative diagram

$$\begin{CD} B(I_0) @>>> B(\mathbf{H}) @= \prod_i B(\text{Res}_{K/\mathbb{Q}_p} \text{GL}_{n_i}) \\ @V \kappa_{I_0} VV @V \kappa_{\mathbf{H}} VV @V (\kappa_{\text{Res}_{K/\mathbb{Q}_p} \text{GL}_{n_i}})_i VV \\ X^*(Z(\widehat{I_0})^\Gamma) @>>> X^*(Z(\widehat{\mathbf{H}})^\Gamma) @= \prod_i X^*(\mathbb{C}^\times). \end{CD}$$

The image of  $\alpha(\gamma; \delta)$  in  $X^*(Z(\widehat{\mathbf{H}})^\Gamma)$  is the image under  $\kappa_{\mathbf{H}}$  of the  $\sigma$ -conjugacy class of  $d^{-1}\delta_{\mathbf{H}}d^\sigma$ , i.e. the  $\sigma$ -conjugacy class of  $\delta_{\mathbf{H}}$ , as  $d \in \mathbf{H}(L)$ . Then

$$\langle \alpha(\gamma; \delta), s \rangle = \langle \kappa_{\mathbf{H}}(\delta_{\mathbf{H}}), s \rangle = \prod_i \langle \kappa_{\text{Res}_{K/\mathbb{Q}_p} \text{GL}_{n_i}}(\delta_{\mathbf{H},i}), s_i \rangle.$$

The right-hand side is equal to  $\prod_i s_i^{v_p(N_{K/\mathbb{Q}_p} \det \delta_{\mathbf{H},i})}$  thanks to Lemma 8.7 below. We have finished the proof of (24). □

The following well-known fact was used in the proof.

**Lemma 8.7.** *Let  $G = \text{Res}_{K/\mathbb{Q}_p} \text{GL}_n$  for a finite extension  $K$  of  $\mathbb{Q}_p$ . Then the Kottwitz map*

$$B(G) \rightarrow X^*(Z(\widehat{G})^{\Gamma(p)}) = X^*(\mathbb{C}^\times) \xrightarrow{\text{canon.}} \mathbb{Z}$$

*is induced by the map  $g \mapsto v_p(N_{K/\mathbb{Q}_p} \det g)$  from  $G(L) = \text{GL}_n(K \otimes_{\mathbb{Q}_p} L)$  to  $\mathbb{Z}$ .*

*Proof.* The functoriality of the Kottwitz map ([31, 4.9]) with respect to

$$\det : \text{Res}_{K/\mathbb{Q}_p} \text{GL}_n \rightarrow \text{Res}_{K/\mathbb{Q}_p} \text{GL}_1$$

reduces the proof to the case of  $n = 1$ , where the assertion is standard; cf. [27, 2.5]. □

### 9. COHOMOLOGY OF SHIMURA VARIETIES: NONENDOSCOPIC CASE

Consider a compact unitary group Shimura variety with trivial endoscopy as in Section 5.

**Theorem 9.1.** *Assume that all places of  $F_0$  above  $p$  are split in  $F$ . Then we have an identity*

$$H_\xi^* = \sum_{\pi_f} a(\pi_f) \pi_f \otimes (r_{-\mu} \circ \varphi_{\pi_p} |_{W_{E_p}}) \cdot |^{-\dim \text{Sh}/2}$$

*as virtual  $\mathbf{G}(\mathbb{A}_f^p) \times \mathbf{G}(\mathbb{Z}_p) \times W_{E_p}$ -representations. Here  $\pi_f$  runs through irreducible admissible representations of  $\mathbf{G}(\mathbb{A}_f)$ , the integer  $a(\pi_f)$  is as in [29, p. 657], and  $\varphi_{\pi_p}$  is the local Langlands parameter associated to  $\pi_p$ .*

*Proof.* First, we note that the assumption implies that Conjecture 7.1 is true in this case, by Theorem 8.1. Matsushima’s formula can be reformulated as the identity

$$H_\xi^* = N \sum_{\pi_f} a(\pi_f) \pi_f$$

as a  $\mathbf{G}(\mathbb{A}_f)$ -representation; cf. Lemma 4.2 of [24]. Now Theorem 5.1 and the identity from Conjecture 7.1 show that

$$\begin{aligned} \mathrm{tr}(\tau \times hf^p | H_\xi^*) &= N^{-1} \mathrm{tr}(f_{\tau,h}^{\mathbf{G}} f^p | H_\xi^*) \\ &= \sum_{\pi_f} a(\pi_f) \mathrm{tr}(\tau | (r_{-\mu} \circ \varphi_{\pi_p} |_{W_{E_p}}) | \cdot |^{-\dim \mathrm{Sh}/2}) \mathrm{tr}(hf^p | \pi_f) . \end{aligned}$$

This gives the desired identity. □

**Corollary 9.2.** *In the situation of the theorem, let  $K \subset \mathbf{G}(\mathbb{A}_f)$  be any sufficiently small compact open subgroup. Then the semisimple local Hasse-Weil zeta function of  $\mathrm{Sh}_K$  at the place  $\mathfrak{p}$  of  $E$  is given by*

$$\zeta_{\mathfrak{p}}^{\mathrm{ss}}(\mathrm{Sh}_K, s) = \prod_{\pi_f} L^{\mathrm{ss}}(s - \dim \mathrm{Sh}/2, \pi_p, r_{-\mu})^{a(\pi_f) \dim \pi_f^K} .$$

*Proof.* One can assume that  $K \subset \mathbf{G}(\mathbb{A}_f^p) \times \mathbf{G}(\mathbb{Z}_p)$ . Now the corollary follows directly from the previous theorem and the definitions. □

### 10. COHOMOLOGY OF SHIMURA VARIETIES: ENDOSCOPIC CASE

In this section we proceed to study compact unitary group Shimura varieties at split places when endoscopy is nontrivial. We will work in the setup of [44, §5.1] to be recalled now. Suppose that

- $B = F$  is a CM field with complex conjugation as the involution  $*$ , and
- $F$  contains an imaginary quadratic field  $\mathcal{K}$ .

Fix a prime  $p$  split in  $\mathcal{K}$  and let  $\mathfrak{p}$  be a prime of  $E$  above  $p$ . We also fix a field isomorphism  $\iota_\ell : \overline{\mathbb{Q}}_\ell \simeq \mathbb{C}$  throughout. Set  $n = \dim_F V \geq 1$  and  $F_0 = F^{*=1}$ . Let  $\mathrm{Spl}_{F/F_0}$  denote the set of all rational primes  $v$  such that every prime of  $F_0$  above  $v$  splits in  $F$ . We further assume

- (1)  $F_0 \neq \mathbb{Q}$ ,
- (2) if a prime  $v$  is ramified in  $F$ , then  $v \in \mathrm{Spl}_{F/F_0}$ ,
- (3)  $\mathbf{G}_{\mathbb{Q}_v}$  is quasisplit at all finite places  $v$ ,
- (4) the flat closure of the generic fibre in the integral models of [40] is proper.

Condition (4) is necessary for our machinery. A sufficient condition for (4) is that the pairing  $(\ , \ )$  on  $V$  is positive or negative definite with respect to some complex embedding  $F \hookrightarrow \mathbb{C}$ , as proved by Kai-Wen Lan, [32, Theorem 5.3.3.1, Remark 5.3.3.2]; cf. also [40, Theorem 5.8], for an explanation.

Conditions (1) and (2) are imposed only to avoid issues with  $L$ -packets for unitary groups and will become unnecessary in due course. We remark that this situation is somewhat more general than in [44], as we allow general signatures at  $\infty$ .

Condition (3) contrasts the case considered in the previous section, and means that endoscopy is seen in its full strength. It implies that  $\mathbf{G}(\mathbb{A})$  admits an automorphic representation whose base-change is  $\Pi$ , for any  $\Pi$  as below. (If  $\mathbf{G}_{\mathbb{Q}_v}$  is not quasi-split at a finite  $v$ , then there would be a local obstruction at  $v$ .)

Let  $\Pi = \psi \otimes \Pi^1$  be an automorphic representation of  $\mathbf{G}(\mathbb{A}_{\mathcal{K}}) \simeq \mathrm{GL}_1(\mathbb{A}_{\mathcal{K}}) \times \mathrm{GL}_n(\mathbb{A}_F)$ . Let  $\theta$  denote the automorphism of  $\mathbf{G}(\mathbb{A}_{\mathcal{K}})$  induced by the nontrivial element  $c$  of  $\mathrm{Gal}(\mathcal{K}/\mathbb{Q})$ . Suppose that

- $\Pi \simeq \Pi \circ \theta$  (such a  $\Pi$  is called  $\theta$ -stable),
- the  $L$ -parameter for  $\Pi_{\infty}$  is the base-change of  $\varphi_{\xi}$ ,
- $\Pi$  is ramified only at the places of  $F$  above  $\mathrm{Spl}_{F/F_0}$ .

We will put ourselves in either (Case ST) or (Case END) of [44, §6.1], to be briefly recalled here. The first case refers to the case when  $\Pi$  is cuspidal. In the second,  $\Pi$  is required to be parabolically induced from a  $\theta$ -stable cuspidal automorphic representation on a maximal proper  $\theta$ -stable Levi subgroup which is isomorphic to a quadratic base-change of an elliptic endoscopic group of  $\mathbf{G}$ . Such a Levi subgroup is isomorphic to  $\mathrm{GL}_1(\mathbb{A}_{\mathcal{K}}) \times \mathrm{GL}_{m_1}(\mathbb{A}_F) \times \mathrm{GL}_{m_2}(\mathbb{A}_F)$  for some  $m_1 \geq m_2 \geq 1$  such that  $m_1 + m_2 = n$ . It is isomorphic to a quadratic base-change of a (necessarily unique up to isomorphism) elliptic endoscopic group  $(\mathbf{H}, s, \eta)$  of  $\mathbf{G}$  in all cases except when  $m_1, m_2$  and  $[F^+ : \mathbb{Q}]$  are all odd. We may and do assume that  $s^2 = 1$ . Let  $\pi_p$  be the irreducible smooth representation of  $\mathbf{G}(\mathbb{Q}_p)$  whose base-change is isomorphic to  $\Pi_p$ . Similarly, in (Case END), we have a representation  $\pi_{\mathbf{H},p}$  of  $\mathbf{H}(\mathbb{Q}_p)$ ; its  $\eta$ -transfer is  $\pi_p$ .

Fix a set of representatives for  $\mathcal{E}_{\mathrm{ell}}(\mathbf{G})$  once and for all as in [44, §3.2], by choosing a Hecke character  $\varpi : \mathbb{A}_{\mathcal{K}}^{\times}/\mathcal{K}^{\times} \rightarrow \mathbb{C}^{\times}$  such that  $\varpi|_{\mathbb{A}^{\times}/\mathbb{Q}^{\times}}$  corresponds to the quadratic character associated with  $\mathcal{K}/\mathbb{Q}$  via class field theory, using Lemma 7.1 of [44]. Define  $S$  to be the finite set of all primes which are either  $p$  or a prime where any of  $F, \Pi$  or  $\varpi$  is ramified; our assumptions imply that  $S \subset \mathrm{Spl}_{F/F_0}$ . Define a virtual representation  $H_{\xi}^*(\Pi)$  of  $\mathrm{Gal}(\overline{E}/E)$  by

$$(25) \quad H_{\xi}^*(\Pi) = \sum_{\pi_f} H_{\xi}^*(\pi_f),$$

where the sum runs over  $\pi_f$  such that  $H_{\xi}^*(\pi_f) \neq 0$ ,  $\pi_f^S$  is unramified,  $BC(\iota_{\ell}\pi_f^S) \simeq \Pi^S$  and  $BC(\iota_{\ell}\pi_S) \simeq \Pi_S$ ; this is a finite sum. Here  $BC$  denotes the local base change map, which makes sense as  $\pi_f^S$  is unramified and  $S \subset \mathrm{Spl}_{F/F_0}$  ([44, §4.2]).

We have the following proposition, which is not explicitly stated in [44].

**Proposition 10.1.** *Let  $\pi_f$  be such that  $H_{\xi}^*(\pi_f) \neq 0$ ,  $\pi_f^S$  is unramified and  $BC(\iota_{\ell}\pi_f^S) \simeq \Pi^S$ . Then  $BC(\iota_{\ell}\pi_S) \simeq \Pi_S$ .*

*Proof.* One can argue with the trace formula as in the proof of Corollary 6.5 (iv) of [44], which is similar to the proof of Theorem 6.1 of [44], but relies on the usual trace formula rather than the trace formula for Igusa varieties. One derives formulas just as (6.29) and (6.30) of [44]. Then the proposition follows from the strong multiplicity one theorem for general linear groups and the injectivity of local base change on  $S$  (since  $S \subset \mathrm{Spl}_{F/F_0}$ ; see [44, §4.2]). (In (Case END),  $\Pi$  is not cuspidal and it is implicitly used in the argument that the induced representation of an irreducible unitary representation is irreducible for general linear groups (cf. [44, (6.2)]) to ensure that not only the supercuspidal support of  $\Pi_S$  but  $\Pi_S$  itself is uniquely determined.) □

Set

$$(26) \quad C_{\mathbf{G}} = (-1)^{\dim \mathrm{Sh}_{\tau}(\mathbf{G})}.$$

The following generalizes [44, Thm. 6.4] to other compact unitary Shimura varieties.

**Theorem 10.2.** *In (Case ST),*

$$H_{\xi}^*(\Pi)|_{W_{E_p}} = C_{\mathbf{G}} \cdot (r_{-\mu} \circ \varphi_{\pi_p}|_{W_{E_p}})| \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2}.$$

*In (Case END) there exists  $e_{\Pi} \in \{\pm 1\}$  (independent of  $p$  and  $\mathfrak{p}$ ) such that*

$$H_{\xi}^*(\Pi)|_{W_{E_p}} = C_{\mathbf{G}} \cdot \frac{1}{2} \left( (r_{-\mu} \circ \varphi_{\pi_p}|_{W_{E_p}}) + e_{\Pi} (r_{-\mu} \circ \eta\varphi_{\pi_{\mathbf{H},p}}|_{W_{E_p}})^s \right) | \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2}.$$

*Here  $(r_{-\mu} \circ \eta\varphi_{\pi_{\mathbf{H},p}}|_{W_{E_p}})^s$  is the virtual representation  $V_{-\mu}^+ - V_{-\mu}^-$  of  $W_{E_p}$ , where  $V_{-\mu} = V_{-\mu}^+ \oplus V_{-\mu}^-$  is the decomposition into  $+1$  and  $-1$  eigenspaces for the  $s$ -action on  $r_{-\mu} \circ \eta\varphi_{\pi_{\mathbf{H},p}}$ .*

*Remark 10.3.* Even though  $\eta\varphi_{\pi_{\mathbf{H},p}}$  is equivalent to  $\varphi_{\pi_p}$ , we keep the expression  $\eta\varphi_{\pi_{\mathbf{H},p}}$  to remember how  $s$  acts on the underlying vector space. Note that the assertion in (Case END) can be restated as

$$(27) \quad H_{\xi}^*(\Pi)|_{W_{E_p}} = \begin{cases} C_{\mathbf{G}} V_{-\mu}^+ | \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2}, & \text{if } e_{\Pi} = +1, \\ C_{\mathbf{G}} V_{-\mu}^- | \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2}, & \text{if } e_{\Pi} = -1. \end{cases}$$

*Proof.* We will first only prove this theorem up to sign. The sign will be determined later in Corollary 10.6. We adopt the strategy of proof as well as notation and convention (e.g., the choice of Haar measures, transfer factors and intertwining operators) from [44, Thm. 6.1]. Once we begin with Theorem 1.5 (playing the role of Proposition 5.5 in that paper), basically the same argument works if suitable changes are made at  $p$ . As for the choice of test functions, explained at the start of the proof of *loc. cit.*, it suffices to remark that  $f_{\tau,h}^{\mathbf{H}}$  (replacing  $\phi_{\text{Ig},p}^{\tilde{n}}$ ) is in the image of the base-change transfer for each  $\mathbf{H}$  since  $p$  splits in  $\mathcal{K}$  and that  $f_{\xi}^{\mathbf{H}}$  is the same function as  $\phi_{\text{Ig},\infty}^{\tilde{n}}$  by construction (when  $\mathbf{H} = G_{\tilde{n}}$ ). Let us write  $\tilde{f}^{\mathbf{H}}$  (with subscript or superscript) for the function on  $\mathbf{G} \otimes_{\mathbb{Q}} \mathcal{K}$  whose base-change transfer is  $f^{\mathbf{H}}$ . Hence our notation  $f$  and  $\tilde{f}$  correspond to  $\phi$  and  $f$  in [44], respectively.

Let  $\mathbf{G}^*$  be a quasi-split inner form of  $\mathbf{G}$ , which is a principal endoscopic group. In (Case ST), (6.14) of [44] is supplanted by

$$\begin{aligned} \text{tr}(\Pi_p(\tilde{f}_p^{\mathbf{G}^*})A_{\Pi_p}^0) &= \text{tr } \iota_l \pi_p(f_{\tau,h}^{\mathbf{G}^*}) \\ &= \text{tr} \left( \tau \left| (r_{-\mu} \circ \varphi_{\pi_p}|_{W_{E_p}}) \right| \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2} \right) \text{tr}(h|\pi_p), \end{aligned}$$

where the last equality is justified by Theorem 8.1. In (Case END), the following is put in place of (6.19) of [44]:<sup>13</sup>

$$\begin{aligned} \text{tr}(\Pi_{\mathbf{H},p}(\tilde{f}_p^{\mathbf{H}})A_{\Pi_{\mathbf{H},p}}^0) &= \text{tr } \iota_l \pi_{\mathbf{H},p}(f_{\tau,h}^{\mathbf{H}}) \\ &= \text{tr} \left( s\tau \left| (r_{-\mu} \circ \eta\varphi_{\pi_{\mathbf{H},p}}|_{W_{E_p}}) \right| \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2} \right) \text{tr}(h|\pi_p). \end{aligned}$$

This time the last equality comes from Corollary 8.6. The right-hand side is equal to  $\text{tr}(\tau|(r_{-\mu} \circ \eta\varphi_{\pi_{\mathbf{H},p}}|_{W_{E_p}})^s| \cdot |\mathfrak{p}|^{-\dim \text{Sh}/2})$ . Set  $e_{\Pi} = e_1^{-1}e_2$ . With these changes, the argument of [44] proves Theorem 10.2 up to sign.  $\square$

The theorem implies the Ramanujan-Petersson conjecture for regular algebraic conjugate self-dual cuspidal automorphic representations of  $\text{GL}_n$ . This was proved previously by Clozel and Caraiani, [12], [10]. We note that Caraiani’s result is stronger in that it works for all finite places.

<sup>13</sup>There are typos in (6.19) and (6.20) of [44]:  $\Pi_p$  and  $\Pi_{\infty}$  should read  $\Pi_{\mathbf{H},p}$  and  $\Pi_{\mathbf{H},\infty}$ , resp.

**Corollary 10.4.** *Let  $M$  be a CM field, and let  $\Pi$  be a cuspidal automorphic representation of  $\mathrm{GL}_n(\mathbb{A}_M)$  such that  $\Pi^\vee \cong \Pi \circ c$  and  $\Pi_\infty$  is regular algebraic, i.e. has the same infinitesimal character as some algebraic representation of  $R_{M/\mathbb{Q}}\mathrm{GL}_n$ . Then  $\Pi_v$  is tempered for all finite places  $v$  of  $M$  above rational primes  $p$  above which  $M$  and  $\Pi$  are unramified.*

*Remark 10.5.* In fact the proof shows that  $\Pi_v$  is tempered at all ramified places as well unless  $n$  is even and condition (3) of Theorem 11.1 fails.

*Proof.* Fix the place  $v$  of  $M$ , and let  $p$  be the rational prime below  $v$ . We adjoin several quadratic extensions of  $\mathbb{Q}$  which are split at  $p$  to  $M$  to make the following assumptions true:

- $F' = M$  is the composite of the totally real subfield  $F'_0$  and an imaginary quadratic field  $\mathcal{K}$  in which  $p$  splits,
- $\Pi$  and  $F'$  are only ramified at places above  $\mathrm{Spl}_{F'/F'_0}$ ,
- $F'_0 \neq \mathbb{Q}$ .

Note that the quadratic base-change of  $\Pi$  stays conjugate self-dual and  $\Pi_\infty$  stays regular algebraic. If  $\Pi$  is unramified at some place where the quadratic field ramifies, then it also stays cuspidal; cf. [3]. Finally, we adjoin one more real quadratic extension of  $\mathbb{Q}$ , split at  $p$ , to  $F'_0$  and  $F'$ , to get  $F_0$  and  $F$ . One checks that all assumptions remain true for  $F$  in place of  $F'$ . Let  $\Pi'$  be the base-change of the original  $\Pi$  to  $F'$ , and let  $\Pi$  be the further base-change to  $F$ .

Using Lemma 7.2 of [44], we arrive at a situation where the assumptions of (Case ST) are satisfied. Consider a PEL datum whose associated group  $\mathbf{G}$  is quasi-split at all finite places and has as signature at infinity either

- (i)  $(1, n - 1), (0, n), \dots, (0, n)$  or
- (ii)  $(1, n - 1), (1, n - 1), (0, n), \dots, (0, n)$ .

In case (ii), we require that the two infinite places of  $F_0$  with signature  $(1, n - 1)$  lie above one place of  $F'_0$ . If  $n$  is odd, then there is no obstruction for choosing the signature. In general there is a parity obstruction but one can still find such a PEL datum with one of the two signatures. Note that all assumptions of this section are satisfied.

Let  $\mathrm{Sh}$  denote the corresponding Shimura variety. Note that the reflex field  $E$  is given by  $F$  in case (i), and by  $F'$  in case (ii). In case (i), the proof of [44, Cor. 6.5] (following the proof presented on page 207 of [17]), can be combined with the Weil conjecture to imply that  $\Pi$  is tempered at  $v$ . We note that this argument works even at ramified places. Let us explain the argument in the more subtle case (ii). Note that we have a place  $v$  of  $F'$  and want to prove that  $\Pi'_v$  is tempered. We use the previous theorem at the place  $v$  of  $E = F'$ . The representation  $r_{-\mu} \circ \varphi_{\pi_p}|_{W_{F'_v}}$  is given by a unitary twist of  $\varphi_{\Pi'_v}^{\otimes 2}$ . The classification of unitary generic representations shows that  $\Pi'_v$  is induced from a representation

$$\bigotimes_i (\pi_i | \cdot |^{a_i} \otimes \pi_i | \cdot |^{-a_i}) \otimes \bigotimes_j \pi_j$$

of a Levi subgroup, where  $\pi_i$  are tempered representations of smaller  $\mathrm{GL}_k$ 's, and  $0 < a_i < \frac{1}{2}$  are real numbers. We want to prove that the first factor is trivial. Our assumptions ensure that this representation already occurs at an unramified level, for which there exists a smooth model of our Shimura variety. Thus, it follows from the Weil conjectures and the fact that  $\varphi_{\Pi'_v}^{\otimes 2}$  occurs in the cohomology of the

Shimura variety that  $4a_i \in \mathbb{Z}$ , i.e.  $a_i = \frac{1}{4}$  for all  $i$ . If the first factor is nonzero, then it follows that  $\varphi_{\Pi'_v}^{\otimes 2}$  contributes nontrivially to three consecutive degrees. On the other hand, the cohomology is an alternating sum, with consecutive degrees contributing different signs (which cannot cancel each other as they have different weights). This contradiction finishes the proof.  $\square$

**Corollary 10.6.** *In both (Case ST) and (Case END), for any  $\pi_f$  such that  $H_\xi^*(\pi_f) \neq 0$ ,  $\pi_f^S$  is unramified and  $BC(\iota_\ell \pi_f^S) \simeq \Pi^S$ , the  $\pi_f$ -isotypic component  $H_\xi^i(\pi_f)$  is nonzero only for  $i = \dim \text{Sh}$ . In particular, the sign in Theorem 10.2 is correct.*

*Proof.* This is a direct application of the Weil conjectures.  $\square$

*Remark 10.7.* We restricted ourselves to (Case ST) and (Case END) above for simplicity, as this is enough for the application of the next section. It should be possible to treat the case of general  $\Pi$  once finer information becomes available about the terms in the twisted trace formula for  $\mathbf{G} \otimes_{\mathbb{Q}} \mathcal{K}$ . Such information would be obtained by carrying over the results of [50] and [2] to the case of unitary similitude groups.

### 11. CONSTRUCTION OF GALOIS REPRESENTATIONS

In this final section we remark that the approach used in this paper provides a shorter proof of the main result of [44], which was based on a generalization of the method of [17]. In particular one can bypass the use of Igusa varieties, the Newton stratification, and the analogues of the first and second basic identities ([17, Thm. VI.2.9, V.5.4]; cf. [44, Prop. 5.2, Thm. 6.1]) of Harris and Taylor.

**Theorem 11.1** ([44, Thm. 1.2]). *Let  $F$  be any CM field. Let  $\Pi$  be a cuspidal automorphic representation of  $\text{GL}_n(\mathbb{A}_F)$  such that*

- (1)  $\Pi^\vee \simeq \Pi \circ c$ ,
- (2)  $\Pi_\infty$  has the same infinitesimal character as some irreducible algebraic representation  $\Xi^\vee$  of the restriction of scalars  $R_{F/\mathbb{Q}}\text{GL}_n$ , and
- (3)  $\Xi$  is slightly regular if  $n$  is even.

*Then for each prime  $\ell$  and an isomorphism  $\iota_\ell : \overline{\mathbb{Q}}_\ell \simeq \mathbb{C}$ , there exists a continuous semisimple representation  $R_{\ell, \iota_\ell}(\Pi) : \text{Gal}(\overline{F}/F) \rightarrow \text{GL}_n(\overline{\mathbb{Q}}_\ell)$  such that for any place  $y$  of  $F$  not dividing  $\ell$ , the Frobenius semisimple Weil-Deligne representation associated to  $R_{\ell, \iota_\ell}(\Pi)|_{\text{Gal}(\overline{F}_y/F_y)}$  corresponds to  $\iota_\ell^{-1}\Pi_y$  via (a suitably normalized) local Langlands correspondence, except possibly for the monodromy operator  $N$ .*

See [44] for the meaning of (3) and a complete list of assertions regarding  $R_{\ell, \iota_\ell}(\Pi)$ . It is not difficult to extend the above theorem to the case where  $\Pi^\vee \simeq \Pi \circ c$  holds up to a character twist or where  $\Pi^\vee \simeq \Pi$  up to a character and  $F$  is a totally real field. Moreover using the above theorem as a starting point, several mathematicians have strengthened it ([11], [10], [4], [5]). For instance (3) has become unnecessary and a similar conclusion holds at  $y$  dividing  $\ell$ .

*Proof.* In [44], the main theorems of section 7 follow from Theorem 6.4 (and its corollaries). Since our Theorem 10.2 reproves Theorem 6.4 (by a different method) when applied to the Shimura varieties of [44], the main results of that article can be derived from Theorem 10.2 by the same reasoning. (Our proof of Theorem 10.2 shows that the sign  $e_\Pi$  is equal to  $e_1^{-1}e_2$  of [44]. Hence it can be controlled by

the choice of Archimedean parameters as in Lemma 7.3 of that paper.) Note that as in [44] it is enough to consider Shimura varieties attached to unitary similitude groups in an odd number of variables with signature  $(1, m-1)$ ,  $(0, m)$ ,  $\dots$ ,  $(0, m)$  at infinity where  $m = n$  if  $n$  is odd and  $m = n + 1$  if  $n$  is even, even though Theorem 10.2 covers more general cases.  $\square$

Finally, let us add one remark about the determination of the monodromy operator  $N$ , which is the most delicate extra assertion in [44, Thm. 1.2]. One easily proves temperedness of  $\Pi_v$  for all  $v$ , so it remains to verify the weight-monodromy conjecture for  $R_{\ell, \ell}(\Pi)$ , as in the paper of Taylor and Yoshida, [45]; cf. [44, Section 7]. The input that one needs is a description of the cohomology of each Newton stratum. This can be achieved by using the Langlands-Kottwitz method. First, one has to check that the arguments of [40] compute the cohomology of the Newton stratum corresponding to the  $\sigma$ -conjugacy class  $b \in B(\mathbf{G})$  once  $\phi_{\tau, h}$  is replaced by  $\phi_{\tau, h} \chi_b$ , where  $\chi_b$  is the characteristic function of the  $\sigma$ -conjugacy class  $b$ , at least if  $j$  is sufficiently large. Now one has to follow the arguments of this paper with  $\phi_{\tau, h} \chi_b$  in place of  $\phi_{\tau, h}$ . Note that we have already determined the twisted orbital integrals of  $\phi_{\tau, h}$ , which in turn determine those of  $\phi_{\tau, h} \chi_b$ . This gives the desired description of the cohomology of each Newton stratum, which may then be used to determine  $N$ .

#### ACKNOWLEDGMENTS

This work was started as a result of a discussion at the IAS in Princeton during the special year on Galois representations and automorphic forms. The authors want to thank this institution for the hospitality and also thank T. Haines for related discussions. The first author wants to thank R. Taylor for the invitation to speak there. Moreover, he thanks M. Rapoport for his constant encouragement and help. The second author wishes to thank D. Kazhdan and D. Vogan for answering his questions.

#### REFERENCES

- [1] J. Arthur. A note on  $L$ -packets. *Pure Appl. Math. Q.*, 2(1, Special Issue: In honor of John H. Coates. Part 1):199–217, 2006. MR2217572 (2006k:22014)
- [2] J. Arthur. *The endoscopic classification of representations: orthogonal and symplectic groups*. 2011.
- [3] J. Arthur and L. Clozel. *Simple algebras, base change, and the advanced theory of the trace formula*, volume 120 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1989. MR1007299 (90m:22041)
- [4] T. Barnet-Lamb, T. Gee, D. Geraghty, and R. Taylor. Local-global compatibility for  $l = p$ . I. *preprint*, <http://math.ias.edu/~rtaylor>.
- [5] T. Barnet-Lamb, T. Gee, D. Geraghty, and R. Taylor. Local-global compatibility for  $l = p$ . II. *preprint*, <http://math.ias.edu/~rtaylor>.
- [6] J. Bernstein, P. Deligne, and D. Kazhdan. Trace Paley-Wiener theorem for reductive  $p$ -adic groups. *J. Analyse Math.*, 47:180–192, 1986. MR874050 (88g:22016)
- [7] J. N. Bernstein. Le “centre” de Bernstein. In *Representations of reductive groups over a local field*, Travaux en Cours, pages 1–32. Hermann, Paris, 1984. Edited by P. Deligne. MR771671 (86e:22028)
- [8] D. Blasius and J. D. Rogawski. Zeta functions of Shimura varieties. In *Motives (Seattle, WA, 1991)*, volume 55 of *Proc. Sympos. Pure Math.*, pages 525–571. Amer. Math. Soc., Providence, RI, 1994. MR1265563 (95e:11051)
- [9] K. Buzzard and T. Gee. The conjectural connections between automorphic representations and Galois representations. arXiv:1009.0785.



- [10] A. Caraiani. Local-global compatibility and the action of monodromy on nearby cycles. *preprint*, arXiv:1010:2188.
- [11] G. Chenevier and M. Harris. Construction of automorphic Galois representations, II. *preprint*, <http://fa.institut.math.jussieu.fr/node/45>.
- [12] L. Clozel. Purity reigns supreme. <http://www.institut.math.jussieu.fr/projets/fa/bpFiles/Clozel2.pdf>.
- [13] L. Clozel. Représentations galoisiennes associées aux représentations automorphes autoduales de  $GL(n)$ . *Inst. Hautes Études Sci. Publ. Math.*, (73):97–145, 1991. MR1114211 (92i:11055)
- [14] L. Clozel, M. Harris, and J.-P. Labesse. Construction of automorphic Galois representations, I. In *On the stabilization of the trace formula*, volume 1 of *Stab. Trace Formula Shimura Var. Arith. Appl.*, pages 497–527. Int. Press, Somerville, MA, 2011. MR2742611
- [15] L. Fargues. Cohomologie des espaces de modules de groupes  $p$ -divisibles et correspondances de Langlands locales. *Astérisque*, 291:1–200, 2004. MR2074714 (2005g:11110b)
- [16] M. Harris and J.-P. Labesse. Conditional base change for unitary groups. *Asian J. Math.*, 8(4):653–683, 2004. MR2127943 (2006g:11098)
- [17] M. Harris and R. Taylor. *The geometry and cohomology of some simple Shimura varieties*, volume 151 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2001. With an appendix by Vladimir G. Berkovich. MR1876802 (2002m:11050)
- [18] M. Harris and R. Taylor. *The geometry and cohomology of some simple Shimura varieties*, volume 151 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2001. With an appendix by Vladimir G. Berkovich. MR1876802 (2002m:11050)
- [19] G. Henniart. Une preuve simple des conjectures de Langlands pour  $GL(n)$  sur un corps  $p$ -adiques. *Invent. Math.*, 139:439–455, 2000. MR1738446 (2001e:11052)
- [20] D. Kazhdan and Y. Varshavsky. Endoscopic decomposition of certain depth zero representations. In *Studies in Lie theory*, volume 243 of *Progr. Math.*, pages 223–301. Birkhäuser Boston, Boston, MA, 2006. MR2214251 (2007e:22008)
- [21] R. Kottwitz. Shimura varieties and twisted orbital integrals. *Math. Ann.*, 269:287–300, 1984. MR761308 (87b:11047)
- [22] R. Kottwitz. Stable trace formula: Cuspidal tempered terms. *Duke Math.*, 51:611–650, 1984. MR757954 (85m:11080)
- [23] R. Kottwitz. Stable trace formula: Elliptic singular terms. *Math. Ann.*, 275:365–399, 1986. MR858284 (88d:22027)
- [24] R. Kottwitz. On the  $\lambda$ -adic representations associated to some simple Shimura varieties. *Invent. Math.*, 108:653–665, 1992. MR1163241 (93f:11046)
- [25] R. Kottwitz and M. Rapoport. On the existence of  $F$ -crystals. *Comment. Math. Helv.*, 78(1):153–184, 2003. MR1966756 (2004c:11221)
- [26] R. Kottwitz and D. Shelstad. Foundations of twisted endoscopy. *Astérisque*, no. 255, 1999, vi + 190 pp. MR1687096 (2000k:22024)
- [27] R. E. Kottwitz. Isocrystals with additional structure. *Compositio Math.*, 56(2):201–220, 1985. MR809866 (87i:14040)
- [28] R. E. Kottwitz. Shimura varieties and  $\lambda$ -adic representations. In *Automorphic forms, Shimura varieties, and  $L$ -functions, Vol. I (Ann Arbor, MI, 1988)*, volume 10 of *Perspect. Math.*, pages 161–209. Academic Press, Boston, MA, 1990. MR1044820 (92b:11038)
- [29] R. E. Kottwitz. On the  $\lambda$ -adic representations associated to some simple Shimura varieties. *Invent. Math.*, 108(3):653–665, 1992. MR1163241 (93f:11046)
- [30] R. E. Kottwitz. Points on some Shimura varieties over finite fields. *J. Amer. Math. Soc.*, 5(2):373–444, 1992. MR1124982 (93a:11053)
- [31] R. E. Kottwitz. Isocrystals with additional structure. II. *Compositio Math.*, 109(3):255–339, 1997. MR1485921 (99e:20061)
- [32] K.-W. Lan. *Arithmetic compactifications of PEL-type Shimura varieties*. ProQuest LLC, Ann Arbor, MI, 2008. Thesis (Ph.D.)—Harvard University. MR2711676
- [33] R. Langlands. Stable conjugacy: definitions and lemmas. *Can. J. Math.*, 31:700–725, 1979. MR540901 (82j:10054)
- [34] R. Langlands and D. Shelstad. On the definition of transfer factors. *Math. Ann.*, 278:219–271, 1987. MR909227 (89c:11172)
- [35] S. Morel. *On the cohomology of certain noncompact Shimura varieties*. Number 173 in Princeton University Press. Annals of Mathematics Studies, Princeton, New Jersey, 2010. MR2567740 (2011b:11073)

- [36] B. C. Ngô. Le lemme fondamental pour les algèbres de Lie. *Publ. Math. Inst. Hautes Études Sci.*, (111):1–169, 2010. MR2653248 (2011h:22011)
- [37] M. Rapoport. A guide to the reduction modulo  $p$  of Shimura varieties. *Astérisque*, (298):271–318, 2005. Automorphic forms. I. MR2141705 (2006c:11071)
- [38] M. Rapoport and M. Richartz. On the classification and specialization of  $F$ -isocrystals with additional structure. *Compositio Math.*, 103(2):153–181, 1996. MR1411570 (98c:14015)
- [39] P. Scholze. The Langlands-Kottwitz approach for some simple Shimura varieties. arXiv:1003.2451, to appear in *Invent. Math.*
- [40] P. Scholze. The Langlands-Kottwitz method and deformation spaces of  $p$ -divisible groups. arXiv:1110.0230.
- [41] P. Scholze. A new approach to the local Langlands correspondence for  $GL_n$  over  $p$ -adic fields. arXiv:1010.1540, to appear in *Invent. Math.*
- [42] F. Shahidi. A proof of Langlands’ conjecture on Plancherel measures; complementary series for  $p$ -adic groups. *Ann. of Math. (2)*, 132(2):273–330, 1990. MR1070599 (91m:11095)
- [43] S. W. Shin. Automorphic Plancherel density theorem. *to appear in Israel J. of Math.*
- [44] S. W. Shin. Galois representations arising from some compact Shimura varieties. *Annals of Math.*, 173:1645–1741, 2011. MR2800722
- [45] R. Taylor and T. Yoshida. Compatibility of local and global Langlands correspondences. *J. Amer. Math. Soc.*, 20(2):467–493, 2007. MR2276777 (2007k:11193)
- [46] E. Viehmann and T. Wedhorn. Ekedahl-Oort and Newton strata for Shimura varieties of PEL type. 2010.
- [47] D. A. Vogan, Jr. The local Langlands conjecture. In *Representation theory of groups and algebras*, volume 145 of *Contemp. Math.*, pages 305–379. Amer. Math. Soc., Providence, RI, 1993. MR1216197 (94e:22031)
- [48] J.-L. Waldspurger. Le lemme fondamental implique le transfert. *Comp. Math.*, 105(2):153–236, 1997. MR1440722 (98h:22023)
- [49] J.-L. Waldspurger. L’endoscopie tordue n’est pas si tordue. *Mem. Amer. Math. Soc.*, 194(908):x+261, 2008. MR2418405 (2011d:22020)
- [50] P.-J. White. Tempered automorphic representations of the unitary group. arXiv: 1106.1127v1 [math.NT].

MATHEMATISCHES INSTITUT DER UNIVERSITÄT BONN, ENDENICHER ALLEE 60, 53115 BONN, GERMANY

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MASSACHUSETTS 02139 – AND – KOREA INSTITUTE FOR ADVANCED STUDY, 85 HOEGIRO, DONGDAEMUN-GU, SEOUL 130-722, REPUBLIC OF KOREA