Ergod. Th. & Dynam. Sys., (2025), **45**, 2249–2272 © The Author(s), 2025. Published by Cambridge University Press.
doi:10.1017/etds.2024.136

Rich representations and superrigidity

GREGORIO BALDI†, NICHOLAS MILLER‡, MATTHEW STOVER§ and EMMANUEL ULLMO¶

† *CNRS, IMJ-PRG, Sorbonne Université, Paris 75005, France*
(e-mail: baldi@imj-prg.fr)
‡ *Department of Mathematics, University of Oklahoma, Norman, OK 73019, USA*
(e-mail: nickmbmiller@ou.edu)
§ *Department of Mathematics, Temple University, Philadelphia, PA 19147, USA*
(e-mail: mstover@temple.edu)
¶ *I.H.E.S., Bures-sur-Yvette 91440, France*
(e-mail: ullmo@ihes.fr)

(*Received* 11 *March* 2024 *and accepted in revised form* 28 *November* 2024)

Abstract. We investigate and compare applications of the Zilber–Pink conjecture and dynamical methods to rigidity problems for arithmetic real and complex hyperbolic lattices. Along the way, we obtain new general results about reconstructing a variation of Hodge structure from its typical Hodge locus that may be of independent interest. Applications to Siu's immersion problem are also discussed, the most general of which only requires the hypothesis that infinitely many closed geodesics map to proper totally geodesic subvarieties under the immersion.

Key words: complex hyperbolic lattices, superrigidity, Hodge theory and Mumford–Tate domains, functional transcendence, unlikely intersections
2020 Mathematics Subject Classification: 14G35, 22E40, 03C64, 14P10, 32H02, 32M15 (Primary)

1. *Introduction and motivation*

A cornerstone in the representation theory of lattices in Lie groups is the Margulis superrigidity theorem [31, p. 2]. Indeed, the implications of superrigidity are profound, including arithmeticity of irreducible lattices in all non-compact semisimple Lie groups except $\mathrm{PO}(1,n)$ and $\mathrm{PU}(1,n)$, that is, for all but real and complex hyperbolic lattices. Specifically, Margulis proved superrigidity of irreducible lattices in semisimple Lie groups of real rank at least two. Using the theory of harmonic maps, this was later extended to the rank-one groups $\mathrm{Sp}(1,n)$ and $\mathrm{F}_4^{(-20)}$ by Corlette [14] and Gromov and Schoen [22]. While the harmonic maps method was extended to give a new proof in higher rank [32], it is open (and quite relevant in the context of the blend of techniques used in this paper) whether the dynamical methods of Margulis can give an independent proof in the rank-one cases.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

To be precise, let Γ be a lattice in a Lie group G. A representation ρ from Γ to a topological group H is said to *extend* if there is a continuous homomorphism $\widetilde{\rho} : G \to H$ so that ρ is the restriction of $\widetilde{\rho}$ to Γ under its lattice embedding in G. Then, Γ is *superrigid* if every unbounded, Zariski dense representation of Γ to a connected, adjoint simple algebraic group H over a local field k of characteristic zero extends. It is known that there are lattices in $\mathrm{PO}(1, n)$ and $\mathrm{PU}(1, n)$ that are not superrigid; see [23, §5] for real hyperbolic examples and [34, §22] or [27, §1] for some results in the complex hyperbolic setting. The study of representations of real and complex hyperbolic lattices is very difficult and rich in open questions, and the extent to which superrigidity fails has yet to be made precise. For example, one motivation for this paper is the first part of a question asked by David Fisher [20, Question 3.15].

Question 1.1. (D. Fisher) Let $\Gamma \subset G$ be a lattice where $G = \mathrm{SO}(1, n)$ or $\mathrm{SU}(1, n)$. What conditions on a representation $\rho : \Gamma \to \mathrm{GL}_m(k)$ imply that ρ extends or almost extends? What conditions on Γ imply that Γ is arithmetic?

In particular, this paper builds on the techniques from [1, 2, 8] to explore *geometric* conditions on a representation that allow one to use tools from dynamics and/or Hodge theory to prove new rigidity phenomena. The conditions we provide, which address Question 1.1 through the behavior of totally geodesic submanifolds under maps related to the representation, have also been considered in studying the infinite volume setting; see for example [25, 26]. There is some history of success along these lines. For just one family of examples, some classes of representations studied in the literature have *maximal* geometric behavior, where two distinct instances are:

- maximal for the *Toledo invariant* (Koziarz and Maubon [29] and many others);
- maximal for their *limit set* (Besson, Courtois and Gallot [10] and Shalom [39]).

Here, we study a class of representations generalizing those shown to be superrigid in [1, 2], namely those that are *geodesically rich*.

Before delving into technical definitions and stating our main results, we discuss a primary application of our methods that can be stated much more simply.

Since it arises several times, we briefly recall what it means for $\Gamma \subset G$ to be an arithmetic lattice of the simplest kind. These are well studied in the literature, so we are terse. In the real hyperbolic case, this means that there is a totally real number field F with integer ring $\mathcal{O}_F$ and a quadratic form q on F^{n+1} so that Γ is commensurable with $\mathcal{G}(\mathcal{O}_F)$, where $\mathcal{G}$ is the F-algebraic group $\mathrm{PO}(q)$. In the complex hyperbolic case, there is a totally imaginary extension E of the totally real field F so that Γ is commensurable with $\mathcal{G}(\mathcal{O}_F)$, where $\mathcal{G}$ is now $\mathrm{PU}(h)$ for a Hermitian form h on E^{n+1} (which is still an F-algebraic group). In each case, it is necessarily the case that the form has signature $(1, n)$ at one place of F and is definite at all other places. The associated locally symmetric spaces have an abundance of totally geodesic submanifolds, which are the main object of study in this paper.

1.1. *Siu's immersion problem.* An original motivation for studying the more general rigidity questions considered in this paper is Siu's *immersion problem* [41, Problem (b),

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

pp. 182]. Let $\mathbb{B}^n$ denote the unit ball in $\mathbb{C}^n$ with its Bergman metric. If Γ is a lattice in the holomorphic isometry group $\mathrm{PU}(1, n)$ of $\mathbb{B}^n$, then S_Γ will denote the *ball quotient* $\Gamma\backslash\mathbb{B}^n$.

Question 1.2. (Siu, 1985) Is it true that every holomorphic embedding between compact ball quotients of dimension at least 2 must have totally geodesic image?

To our knowledge the only previous result is by Cao and Mok [**13**], which answered Question 1.2 in the affirmative for $f : \Gamma\backslash\mathbb{B}^n \to \Lambda\backslash\mathbb{B}^m$ when $m < 2n$. See [**2**, Theorem 1.7] and the surrounding references for more on Siu's analogous *submersion problem*. The following is our main contribution toward a positive answer to Siu's immersion problem.

THEOREM 1.3. *Let $f : S_{\Gamma_1} \to S_{\Gamma_2}$ be an immersive holomorphic map between arithmetic ball quotients such that $f(S_{\Gamma_1})$ does not lie in any smaller totally geodesic subvariety of S_{Γ_2}. Suppose that S_{Γ_1} contains a proper totally geodesic subvariety of positive dimension. Then, the following are equivalent.*

(1) *The map f is covering.*
(2) *There is a generic sequence of positive-dimensional proper totally geodesic subvarieties $Y_\ell \subset S_{\Gamma_1}$ such that each $f(Y_\ell)$ lies is a proper totally geodesic subvariety $W_\ell \subset S_{\Gamma_2}$.*

The positive-dimensional hypothesis on the subvarieties of S_{Γ_1} is to rule out the trivial counterexample of points. We offer two proofs of Theorem 1.3, one using homogeneous dynamics which gives a slightly weaker statement in §3, and the second using Hodge theory and functional transcendence in §4, as an application of Theorem 1.18 below.

For arithmetic lattices of simplest kind, Theorem 1.3 reduces Question 1.2 to checking a single dimension.

COROLLARY 1.4. *If Question 1.2 has a positive answer when S_{Γ_1} is a two-dimensional arithmetic ball quotient of simplest type, then it has a positive answer for S_{Γ_1} an arithmetic ball quotient of simplest type of any dimension $n \geq 2$.*

Remark 1.5. By [**13**], the first open case is of a holomorphic embedding $f : S_{\Gamma_1} \to S_{\Gamma_2}$ with $\dim S_{\Gamma_1} = 2$, $\dim S_{\Gamma_2} = 4$. Here, Theorem 1.3 implies that it suffices to prove that there is an infinite collection of distinct totally geodesic curves $C_i \subset S_{\Gamma_1}$ such that each $f(C_i)$ lies in some strict totally geodesic subvariety $Y_i \subset S_{\Gamma_2}$.

There are various similar generalizations that follow from Theorem 1.17. For example, a period map $\psi : S_{\Gamma_1} \to \Gamma\backslash D$ such that the $f^{-1}(Y_i)$ are totally geodesic subvarieties of S_{Γ_1} for a generic sequence of Mumford–Tate subdomains $Y_i \subset \Gamma\backslash D$ is necessarily an isomorphism. Here, 'generic' means that they are not all contained in a finite union of maximal totally geodesic subvarieties of S.

From a dynamical point of view, the dimension at least two hypothesis in the previous results is natural and critical in that it allows one to access tools from unipotent dynamics. From a Hodge theoretic point of view, one typically works with a complex subvariety. However, there is still something one can say for closed geodesics.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

THEOREM 1.6. *Let $S = \Gamma_n\backslash\mathbb{B}^n$ be a ball quotient with an immersive holomorphic map*

$$f : S \longrightarrow \Gamma_m\backslash\mathbb{B}^m$$

such that $f(S)$ does not lie in any smaller totally geodesic subvariety of $\Gamma_m\backslash\mathbb{B}^m$. Let $\{C_\ell\}$ be a generic sequence of distinct closed geodesics in S in the sense that they are not all contained in a finite union of maximal totally geodesic subvarieties of S. Assume that for each ℓ, there exits a strict complex totally geodesic subvariety $Y_\ell \subset \Gamma_m\backslash\mathbb{B}^m$ such that $f(C_\ell) \subset Y_\ell$. Then, f is a totally geodesic immersion.

See Corollary 4.4 for the proof. Note that Theorem 1.6 applies for non-arithmetic Γ, since all finite volume ball quotients contain infinitely many free homotopy classes of closed geodesics. That f must take certain closed geodesics into a proper totally geodesic subspace is a very strong hypothesis, but it is not entirely unprecedented. Indeed, one can show that some of the maps associated with surjections between complex hyperbolic lattices take a certain (necessarily finite) collection of totally geodesic subspaces of S_{Γ_n} to totally geodesic subspaces of S_{Γ_m}. We close this section with a final question that would generalize Theorem 1.6.

Question 1.7. Let Γ be any lattice in $\mathrm{PU}(1, n)$, $n > 1$, and $f : S_\Gamma \to \Gamma_D\backslash D$ be any period map (e.g., $\Gamma_D\backslash D$ is a Hermitian symmetric space and f is holomorphic). If f sends a generic sequence of closed geodesics $\{C_\ell\}$ of S_Γ to closed geodesics of $\Gamma_D\backslash D$, is then $f(S_\Gamma)$ a totally geodesic subvariety?

1.2. *Rich representations and the main theorem.* Let $\Gamma \subset G$ be a torsion-free arithmetic lattice, where G is $\mathrm{SU}(1, n)$ or $\mathrm{SO}(1, n)$, and let X denote the symmetric space associated with G. Then, S_Γ will denote the manifold $\Gamma\backslash X$, which is a smooth manifold that is also a smooth variety in the complex hyperbolic case. Since both contexts are relevant to this paper, we use the term submanifold when considering S_Γ as a differentiable manifold, and subvariety when S_Γ is complex hyperbolic and the submanifold is also a subvariety for the natural algebraic structure. We assume that S_Γ contains one, and therefore infinitely many, proper totally geodesic submanifolds of real dimension at least 2. This implication follows easily from the fact that the associated algebraic group commensurates the lattice. In the other direction, if S_Γ contains infinitely many maximal totally geodesic submanifolds of dimension at least 2, then Γ is arithmetic; see [**1**, **2**, **8**]. When this is the case, we say that S_Γ (or Γ) is *geodesically rich*. Let $\{S_{\Gamma_i}\}_{i\in\mathbb{N}}$ be a sequence of distinct immersed maximal totally geodesic submanifolds, where maximal means that the only totally geodesic submanifold strictly containing S_{Γ_i} is S_Γ itself.

The group Γ_i is the stabilizer in Γ of a totally geodesic subspace Y_i of X, and the restriction of Γ_i to Y_i induces a surjective homomorphism from Γ_i onto the fundamental group of S_{Γ_i}. The kernel of this homomorphism encodes the action of Γ_i on the normal bundle to Y_i in X. The main purpose of this paper is to introduce the notion of geodesically rich representations, which are morally those that 'preserve' a significant amount of the rich structure of Γ.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Definition 1.8. Let k be a local field and $\mathbf{H}$ be a semisimple k-algebraic group without compact factors. Let $\rho : \Gamma \to \mathbf{H}(k)$ be a representation with unbounded, Zariski dense image. We say that ρ is *geodesically rich* (just *rich* when it is clear from context) if the following are satisfied:

- Γ is geodesically rich; and
- there are infinitely many i so that $\dim \mathbf{H}_i < \dim \mathbf{H}$, where $\mathbf{H}_i$ is defined by

$$\mathbf{H}_i := \overline{\rho(\Gamma_i)} \subset \mathbf{H} \tag{1.1}$$

for each i, where closure is with respect to the Zariski topology.

Special consideration will be given in this paper to the case where S_Γ and the subspaces S_{Γ_i} are all complex hyperbolic.

Definition 1.9. Let $\Gamma \subset G = \mathrm{PU}(1, n)$ be a complex hyperbolic lattice and $\rho : \Gamma \to \mathbf{H}(k)$ be a rich representation. Then, ρ is *complex geodesically rich* if infinitely many of the S_{Γ_i} for which $\dim \mathbf{H}_i < \dim \mathbf{H}$ are complex hyperbolic.

Broadly, this paper is concerned with identifying situations where the following question has a positive answer.

Question 1.10. Do rich representations extend? Equivalently, are there non-trivial rich representations (that is, other than the case where $\mathbf{H}(k) = G$ and ρ is given by conjugation)?

Before stating one of our primary results toward a positive answer, we recall a definition that appears in [3] that will be used to give a generalization/abstract version of the method of [1, 2].

Definition 1.11. Fix a parabolic subgroup $P \subset G$ with unipotent radical U. A pair $(k, \mathbf{H})$ as above is *strongly compatible* with G if for any subgroup $\{1\} \neq \mathbf{H}' \subset \mathbf{H}$ and every continuous group homomorphism $\tau : P \to (N_{\mathbf{H}}(\mathbf{H}')/\mathbf{H}')(k)$, we have that $U \leq \mathrm{Ker}(\tau)$.

Remark 1.12. If k is non-Archimedean, then any pair $(k, \mathbf{H})$ is strongly compatible with G.

Our main theorem establishes Conjecture 1.10 in two cases.

THEOREM 1.13. *Let Γ be an arithmetic lattice in G which is either* $\mathrm{PO}(1, n)$ *or* $\mathrm{PU}(1, n)$, *$\mathbf{H}$ be a connected, adjoint algebraic group over a local field k of characteristic zero, and $\rho : \Gamma \to \mathbf{H}(k)$ be a rich representation. If either:*

(1) *$\mathbf{H}$ is strongly compatible (in the sense of Definition 1.11) and k-simple; or*
(2) *Γ is complex hyperbolic, ρ is complex geodesically rich, and ρ is cohomologically rigid (that is, $H^1(\Gamma, \mathrm{Ad}_{\mathbf{H}} \circ \rho) = 0$),*

then ρ extends.

Notice that the former condition is just on the pair $(k, \mathbf{H})$, whereas the latter depends only on ρ. The proof of part (1), which follows from the methods developed in [1, 2], is in §2. The proof of part (2), contained in Theorem 7, is related to the study of variations of Hodge structures and, in fact, it will be related to the so-called *Zilber–Pink conjecture*

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

[7]. Regarding part (2), it might be possible to study some non-rigid representations via the so-called *factorization theorem*, see for example [40, Theorem 10] and various more recent generalizations. The fact that $(k, \mathbf{H})$ is always compatible when k is non-Archimedean has the following fairly direct corollary.

COROLLARY 1.14. *Suppose that* $\Gamma \subset \mathrm{PU}(1, n)$ *is an arithmetic lattice and that* $\rho : \Gamma \to \mathbf{H}(k)$ *is a rich representation to a k-simple algebraic group* $\mathbf{H}$ *defined over a number field k. Then,* ρ *is integral.*

Corollary 1.14 should be viewed as a complement to the fundamental recent work of Esnault and Groechenig [18] that proves integrality of certain cohomologically rigid local systems. Note that Corollary 1.14 does not assume cohomological rigidity, only that the representation is defined over a number field. Moreover, note that there are homomorphisms from complex hyperbolic lattices onto surface groups [43], and hence not all representations of complex hyperbolic lattices are definable over a number field. As noticed in [2], the version of Corollary 1.14 that implicitly follows from the results there makes the main theorem of that paper independent of [18].

Remark 1.15. In contrast with Theorem 1.6, we now give a simple cautionary example regarding generalizing richness to closed geodesics. Let G be $\mathrm{PO}(1, n)$ for $n \geq 3$ or $\mathrm{PU}(1, n)$ for $n \geq 2$, and suppose that Γ_1, Γ_2 are arithmetic lattices in G of simplest type for which there is a surjection $\rho : \Gamma_1 \to \Gamma_2$ with infinite kernel; such representations are known to exist in both settings. Choose a maximal subgroup $\Delta \subset \Gamma_2$ associated with a proper totally geodesic submanifold of S_{Γ_2}. Then, $\rho^{-1}(\Delta)$ contains the kernel of ρ, and therefore is Zariski dense in G. Consequently, $\rho^{-1}(\Delta)$ contains a profinitely open collection of regular semisimple elements [37], and hence they preserve no totally geodesic subspace of X except their axis of translation, and thus the associated closed geodesics on S_Γ are maximal. In the complex hyperbolic setting, the associated closed geodesics are *generic* in the sense that their union is Zariski dense in S_Γ, and hence this provides a fairly rich collection of one-dimensional maximal totally geodesic subspaces.

Representations as in Remark 1.15 are unbounded and Zariski dense with connected, adjoint target, but they cannot extend since they have infinite kernel. They have a well-distributed sequence of maximal closed geodesics for which the associated elements in $\rho(\Gamma_1)$ all preserve a totally geodesic subspace of the target, and hence the representation could be considered rich in analogy with the higher-dimensional setting. While one might reasonably interpret this as saying the one-dimensional case is hopeless, Theorem 1.6 indicates that stronger geometric hypotheses can still lead to a rigidity theorem.

We now comment on how richness may differentiate between real and complex hyperbolic lattices. For example, we do not know the answer to the following.

Question 1.16. Is every faithful linear representation of an arithmetic lattice in $\mathrm{PU}(1, n)$ of simplest kind rich?

Bending deformations of lattices in $\mathrm{PO}(1, n)$ [23, §5] easily imply that Question 1.16 is false for real hyperbolic lattices of simplest kind. However, Goldman–Millson rigidity

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

[21] implies that there are no analogous deformations in the complex hyperbolic setting. It seems that examples giving a negative answer to Question 1.16 would come from representations of a new kind. However, a positive answer would (for example) prove superrigidity of unbounded, Zariski dense, and faithful representations to connected, adjoint targets.

Since Theorem 1.13 proves that all rich representations to connected, adjoint k-simple groups over non-Archimedean fields of characteristic zero are bounded, all other results in this paper consider the cases where k is $\mathbb{R}$ or $\mathbb{C}$. For the remainder of this introduction, we consider rich representations of complex hyperbolic lattices with coefficients in $\mathbb{R}$.

1.3. *Holomorphic and variation of Hodge structure results.* To enter in the world of variations of Hodge structures (VHSs), we focus on *complex hyperbolic lattices* $\Gamma \subset G = \mathrm{PU}(1, n)$ with $n > 1$. See §5 for the various notions from Hodge theory used in this subsection. In this setting, we can restate Definition 1.9 in a more geometric way. Suppose for example that H is the automorphism group of a Hermitian symmetric domain X_H, $\rho : \Gamma \to H$ takes values in a lattice $\Gamma_H \subset H$, and is induced by taking the induced map on fundamental groups for a holomorphic map

$$f : S_\Gamma \longrightarrow \Gamma_H \backslash X_H. \tag{1.2}$$

Then, the complex richness of ρ translates into the following: $f(S_{\Gamma_i})$ lies in a strict complex totally geodesic submanifold $Y_i \subset \Gamma_H \backslash X_H$ for infinitely many i.

We remark here that the setting described in equation (1.2) is a special case of the variation of Hodge structure (VHS, from now on) formalism. The next theorem shows that rich representations underlying an integral pure polarized VHS extend.

THEOREM 1.17. *If ρ is a complex rich representation underlying a direct $\mathbb{C}$-factor of an irreducible $\mathbb{Q}$VHS, then ρ extends.*

The proof of Theorem 1.17 builds on the main results of [7], namely the *geometric part of the Zilber–Pink conjecture*. In particular, the proof crucially uses the following result proved in §6, which may be of independent interest.

THEOREM 1.18. *Let S be a smooth quasi-projective complex variety supporting $\mathbb{Z}$VHSs $\mathbb{V}_1, \mathbb{V}_2$ whose algebraic monodromy groups are $\mathbb{Q}$-simple and non-trivial. If*

$$\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}} = \mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset,$$

then $\mathbb{V}_1, \mathbb{V}_2$ are isogenous.

Given a smooth quasi-projective complex variety S and $\mathbb{V} \to S$ a variation of Hodge structures, we write $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,typ}}$ for the so-called *typical Hodge locus of positive period dimension*, as introduced in [7]. See Definition 6.1 for the precise definition of an isogeny in this context.

1.4. *Outline of the paper.* In §2, we prove the statement about compatibility. The rest of the paper is devoted to the complex hyperbolic case. In §3, we give the argument based on various forms of Ratner's equidistribution theorem. In §4, we give a simple self-contained

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

proof of the results about Siu's immersion problem. The rest of the paper uses more ideas and techniques from Hodge theory. In particular, §5 is devoted to all the required preliminaries from Hodge theory and the Zilber–Pink paradigm. In §6, we prove the most general statement on isogenies of VHSs, namely Theorem 1.18. Finally, we prove the main results about the extension of rich representations underlying a VHS in Theorem 7.

2. *Dynamics background and compatibility*

This section contains some background on the dynamical side, along with the proof of the first part of Theorem 1.13. Let G be either $\mathrm{PO}(1, n)$ or $\mathrm{PU}(1, m)$. We first state some basic facts and establish some notation that will be used throughout this paper. What is described immediately below recaps facts that are given a detailed treatment in [**1**, **2**].

Fixing a maximal compact subgroup K in G, the associated locally symmetric space is $X = G/K$. Fix a representative $Y \subset X$ for each type (that is, G-orbit) of totally geodesic subspace of X with dimension at least two and call such a representative a *standard* subspace. If W is the subgroup of the stabilizer of Y in G generated by unipotent elements, then the full stabilizer of Y is the normalizer $N_G(W)$, which is a compact extension of W.

The Haar measure on G will be denoted μ_G, and (after appropriate scaling) given a lattice Γ in G, this induces a probability measure $\mu_{\Gamma\backslash G}$ on the quotient $\Gamma\backslash G$. Suppose that the translate gY of the standard subspace Y has the property that $gLg^{-1} \cap \Gamma$ defines a lattice Δ in $T = gLg^{-1}$ for some $W \subseteq L \subseteq N_G(W)$. One obtains an immersion $S_\Delta \looparrowright S_\Gamma$ induced by the embedding

$$\Gamma\backslash \Gamma gL \hookrightarrow \Gamma\backslash G$$

and an associated push-forward of (the g-translate of) $\mu_{\Delta\backslash T}$ to a probability measure on $\Gamma\backslash G$. Thus, one obtains an homogeneous W-invariant measure on $\Gamma\backslash G$ with proper support. Upon taking an ergodic component of this measure, one can moreover choose L satisfying the above properties for which the corresponding measure is W-ergodic. If $\{S_{\Gamma_i}\}$ is an infinite sequence of distinct maximal totally geodesic subspaces of S_Γ, then one can extract a subsequence so that they are all associated with G-translates of the same Y, so Γ_i is a lattice in $T_i = g_i L_i g_i^{-1}$ with $W \subseteq L_i \subseteq N_G(W)$ for all i. The maximality assumption ensures that $\mu_{\Gamma\backslash G}$ is a weak-$*$ limit of (the g_i translates of) the measures $\mu_{\Gamma_i\backslash T_i}$. All limits of measures in this paper will be in the weak-$*$ sense.

THEOREM 2.1. *Suppose that $\Gamma \subset G$ is a geodesically rich lattice. Let k be a local field of characteristic zero and $\mathbf{H}$ be a connected, adjoint k-algebraic group so that the pair $(k, \mathbf{H})$ is compatible with G. If $\rho : \Gamma \to \mathbf{H}(k)$ is a rich representation, then ρ extends.*

Proof. By Definition 1.8, ρ is a homomorphism with unbounded, Zariski dense image. Following [**1**, Theorem 1.6 and Remark 1.7], to prove that there exists a continuous extension $\widetilde{\rho} : G \to \mathbf{H}(k)$ of ρ, it suffices to find:

(1) a non-compact almost-simple, proper subgroup $W \subset G$;
(2) a k-rational faithful, irreducible representation $\mathbf{H} \to \mathbf{SL}(V)$ on a finite-dimensional k-vector space V; and
(3) a W-invariant measure ν on $\Gamma\backslash(G \times \mathbb{P}(V))$ that projects to the Haar measure $\mu_{\Gamma\backslash G}$ on $\Gamma\backslash G$.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

We now explain how richness implies the existence of items (1)–(3), which follows the argument given in [1, §3] in a similar context.

Recall that by hypothesis, S_Γ contains infinitely many distinct maximal totally geodesic subspaces S_{Γ_i}. Passing to a subsequence, we assume they are of the same type (in the sense described in this section). This defines a sequence of homogeneous, W-ergodic measures $\{\mu_i\}$ with $\mathrm{supp}(\mu_i) \neq \Gamma\backslash G$ for which $\mu_i \to \mu_{\Gamma\backslash G}$, where $W \subset G$ is a fixed proper subgroup generated by unipotent elements. This provides the W in item (1) above.

The condition that μ_i is W-ergodic and homogeneous is precisely the condition that there exists $L_i \subset G$ with $W \leq L_i \leq N_G(W)$ and $g_i \in G$ for which $\Gamma\backslash\Gamma g_i L_i \subset \Gamma\backslash G$ is closed, and the measure obtained by push-forward is ergodic for the W-action. If $T_i = g_i L_i g_i^{-1}$, then $\Gamma\backslash\Gamma g_i L_i$ being closed in $\Gamma\backslash G$ is equivalent to $\Gamma_i = T_i \cap \Gamma$ being a lattice in T_i. The condition that μ_i is W-ergodic is equivalent to the condition that T_i is the identity component of the Zariski closure of $g_i N_G(W) g_i^{-1} \cap \Gamma$.

Returning to the proof at hand, recall from Definition 1.8 that $\mathbf{H}_i = \overline{\rho(\Gamma_i)}$ is a k-defined non-compact proper subgroup of $\mathbf{H}$ with $\dim(\mathbf{H}_i) < \dim(\mathbf{H})$ for infinitely many i. Passing to a subsequence, we can assume that $\dim(\mathbf{H}_i)$ is constant, and we denote this constant by d. Consider the dth exterior power of the adjoint representation

$$\wedge^d(\mathrm{Ad}) : \mathbf{H}(k) \longrightarrow \mathbf{GL}\Big(\bigwedge^d \mathfrak{h}\Big),$$

where $\mathfrak{h}$ is the Lie algebra of $\mathbf{H}(k)$. The Lie algebra $\mathfrak{h}_i = \mathrm{Lie}(\mathbf{H}_i(k))$ determines a line $\ell_i \subset \bigwedge^d \mathfrak{h}$. The representation $\wedge^d(\mathrm{Ad})$ need not be irreducible; however, there exists a non-trivial summand $\theta : \mathbf{H} \to \mathbf{GL}(V)$, where V is a k-defined vector space, for which infinitely many ℓ_i project non-trivially to V. This representation satisfies item (2) above.

Finally, for item (3), we again pass to a subsequence to assume that all ℓ_i project non-trivially to V and, therefore, $\theta(\rho(\Gamma_i))$ is contained in the stabilizer of ℓ_i for infinitely many i. Let x_i be the point corresponding to the image of ℓ_i in the projectivization $\mathbb{P}(V)$ and we denote the composition of $\theta \circ \rho$ with the projectivization by $\overline{\theta}$.

The *suspension space*, $\Gamma\backslash(G \times \mathbb{P}(V))$, is formed via the identification

$$(g, [v]) \sim (\gamma g, \overline{\theta}(\gamma)[v]),$$

which is naturally a fiber bundle over $\Gamma\backslash G$. Since the right G-action on the first factor of $G \times \mathbb{P}(V)$ commutes with this Γ-action, it descends to a right G-action on the suspension space. Recalling the notation from the beginning of the proof, the fact that $\Gamma\backslash\Gamma T_i$ is closed in $\Gamma\backslash G$ is equivalent to the natural map $\Gamma_i\backslash T_i \to \Gamma\backslash\Gamma T_i$ being a homeomorphism. As $\overline{\theta}(\Gamma_i)$ fixes $\{x_i\}$, it is straightforward that $\Gamma_i\backslash T_i$ is isomorphic to $\Gamma_i\backslash(T_i \times \{x_i\})$ under the same action as above. Recalling that $g_i L_i = T_i g_i$, commutativity with the right G-action yields the following commutative diagram:

$$\begin{array}{ccccc}
\Gamma_i\backslash(L_i \times \{x_i\})g_i & \xrightarrow{\cong} & \Gamma_i\backslash(T_i \times \{x_i\}) & \hookrightarrow & \Gamma\backslash(G \times \mathbb{P}(V)) \\
\downarrow \cong & & \downarrow \cong & & \downarrow \\
\Gamma\backslash\Gamma g_i L_i = \Gamma\backslash\Gamma T_i g_i & \xrightarrow{\cong} & \Gamma\backslash\Gamma T_i \cong \Gamma_i\backslash T_i & \hookrightarrow & \Gamma\backslash G
\end{array}$$

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

If σ_i denotes the induced map $\sigma_i : \Gamma\backslash\Gamma g_i L_i \hookrightarrow \Gamma\backslash(G \times \mathbb{P}(V))$, then $\nu_i = (\sigma_i)_*\mu_i$ provides a W-invariant, ergodic measure on $\Gamma\backslash(G \times \mathbb{P}(V))$. Again passing to a subsequence, let ν_∞ denote any weak-$*$ limit of the ν_i. Then, any ergodic component of ν_∞ projects to $\mu_{\Gamma\backslash G}$ and hence satisfies item (3) since projection and weak-$*$ limit commute. This completes the proof. □

3. *Dynamical proof for a variant of Theorem 1.3*

The purpose of this section is to give a dynamical proof of the following variant of Theorem 1.3. We continue with the notation established in §2.

THEOREM 3.1. *Let $S_n = \Gamma_n\backslash\mathbb{B}^n$ be a smooth complex ball quotient that is arithmetic of simplest kind, and suppose that there is immersive holomorphic map*

$$f : S_n \longrightarrow S_m \coloneqq \Gamma_m\backslash\mathbb{B}^m$$

to another arithmetic ball quotient of simplest kind. Suppose that Y_ℓ is a sequence of distinct irreducible totally geodesic divisors on S_m for which the top dimensional irreducible components of $f^{-1}(Y_\ell)$ are totally geodesic subvarieties of S_n. Then, f is a totally geodesic immersion.

Proof. Throughout this proof, let $G = \mathrm{SU}(1, n)$. Let ω be the Kähler form on S_m descending from the G-invariant form on $\mathbb{B}^m$ and α be the analogous form on S_n. We will prove that $f^*\omega = \lambda\alpha$ for some constant λ, and hence the induced metric on S_n is a multiple of the locally symmetric metric. This implies that f must be a totally geodesic immersion and thus, under the standard normalizations, $f^*\omega = \alpha$.

A simple transversality argument implies that $f^{-1}(Y_\ell)$ has top-dimensional irreducible components with codimension one for all but finitely many ℓ. Let β be a positive $(n-1, n-1)$ form on S_m with compact support. The Y_ℓ equidistribute in S_m, so by applying [30, Corollary 1.5] (also see the related paper of Möller and Toledo [33] and the more general work of Tayou and Tholozan [42]) to both β and ω^{n-1}, we deduce that

$$\frac{\int_{f(S_n)\cap Y_\ell} \beta}{\int_{f(S_n)\cap Y_\ell} \omega^{n-1}} \longrightarrow \frac{\int_{f(S_n)} \beta \wedge \omega}{\int_{f(S_n)} \omega^n} \tag{3.1}$$

as ℓ goes to infinity. By definition of the pullback, the left-hand side of equation (3.1) is also equal to

$$\frac{\int_{f^{-1}(Y_\ell)} f^*\beta}{\int_{f^{-1}(Y_\ell)} f^*\omega^{n-1}} = \frac{\int_{f^{-1}(Y_\ell)} \alpha^{n-1}}{\int_{f^{-1}(Y_\ell)} f^*\omega^{n-1}} \frac{\int_{f^{-1}(Y_\ell)} f^*\beta}{\int_{f^{-1}(Y_\ell)} \alpha^{n-1}}.$$

We want to compute the limits of $(\int_{f^{-1}(Y_\ell)} f^*\beta / \int_{f^{-1}(Y_\ell)} \alpha^{n-1})$ and $(\int_{f^{-1}(Y_\ell)} f^*\omega^{n-1} / \int_{f^{-1}(Y_\ell)} \alpha^{n-1})$.

Since there are possibly several irreducible components, we write

$$f^{-1}(Y_\ell) = Z_{\ell,1} \cup \cdots \cup Z_{\ell,n_\ell}.$$

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Now we use Ratner's theorem in S_n, using the fact that $f^{-1}(Y_\ell)$ consists of totally geodesic divisors. Until this point, we only used that we had some typical intersections, now we redo the same computations using what we know about S_n and $f^{-1}(Y_\ell)$.

LEMMA 3.2. *Let* $\{\mu_{f^{-1}(Y_\ell)}\}$ *be the sequence*

$$\mu_{f^{-1}(Y_\ell)} \coloneqq \frac{1}{n_\ell}\sum_{i=1}^{n_\ell} \mu_{Z_{\ell,i}}$$

of probability measures on $\Gamma_n\backslash G$. *Then, the* $\mu_{f^{-1}(Y_\ell)}$ *weak-*$*$ *converge to* μ_{S_n}.

Proof. We need to recall the description of the totally geodesic divisors on $S_n = \Gamma_n\backslash\mathbb{B}^n$. We refer to [2, Lemma 8.2] for a more detailed description and for a proof of the following facts. We have an isomorphism $\mathbb{B}^n = G/K$, where $K = \mathrm{S}(\mathrm{U}(1)\times\mathrm{U}(n))$ is a maximal compact subgroup of G, and a fixed embedding $W \coloneqq \mathrm{SU}(1, n-1) \subset G$. Let $\pi : \Gamma_n\backslash G \longrightarrow S_n$ be the natural projection associated with the right quotient by K. Then, any totally geodesic divisor Z on S_n is of the form

$$Z = \pi(\Gamma_n\backslash\Gamma_n g W)$$

for some $g \in \mathrm{SU}(1, n)$ such that $gWg^{-1}\cap\Gamma_n$ is a lattice in gWg^{-1} (see §2 and Remark 3.3 at the end of the proof).

Using this description, we can write $Z_{\ell,i} = \pi(\Gamma_n\backslash\Gamma_n g_{\ell,i} W)$ for some $g_{\ell,i} \in \mathrm{SU}(1, n)$, where $1 \le i \le n_\ell$ for each ℓ. Let $\mu_{\ell,i}$ be the W-invariant probability measure on $\Gamma_n\backslash G$ with support $\Gamma_n\backslash\Gamma_n g_{\ell,i} W$ and define

$$\mu_\ell \coloneqq \frac{1}{n_\ell}\sum_{i=1}^{n_\ell}\mu_{\ell,i}.$$

Note that $\pi_*\mu_\ell = \mu_{f^{-1}(Y_\ell)}$ by [1, Lemma 8.2(1)]. As each $\mu_{\ell,i}$ is W-invariant, so is μ_ℓ and consequently any weak-$*$ limit μ of a subsequence of the μ_ℓ is a W-invariant measure. Moreover, the similar argument to [1, Lemma 8.3] using quantitative non-divergence shows that there is no escape of mass and thus μ is also a probability measure.

Let ν be an ergodic component of μ. By Ratner's theorem [38], ν is an homogeneous, W-invariant, W-ergodic measure and therefore is either a constant multiple of the Haar measure on $\Gamma_n\backslash G$ or a constant multiple of the Haar measure on a proper closed W-orbit $V \subsetneq \Gamma_n\backslash G$. However, note that for any such closed W-orbit, $\mu_{\ell,i}(V) = 0$ for sufficiently large ℓ and all $1 \le i \le n_\ell$, and hence the latter cannot occur. Therefore, only the former possibility occurs and hence $\nu = \mu_{\Gamma_n\backslash G}$. Since pushforward and weak-$*$ limits commute, we conclude that $\mu_{f^{-1}(Y_\ell)} \to \mu_{S_n}$, as required. □

As a consequence, again by upgrading Ratner to a statement about forms as in [30, Corollary 1.5], but this time in S_n rather than in $\mathbb{B}^m$, we get

$$\frac{\int_{f^{-1}(Y_\ell)} f^*\beta}{\int_{f^{-1}(Y_\ell)}\alpha^{n-1}} \longrightarrow \frac{\int_{S_n} f^*\beta\wedge\alpha}{\int_{S_n}\alpha^n}.$$

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

When we apply this result to the case $\beta = \omega^{n-1}$, we find that

$$\frac{\int_{f^{-1}(Y_\ell)} f^*\omega^{n-1}}{\int_{f^{-1}(Y_\ell)} \alpha^{n-1}} \longrightarrow \frac{\int_{S_n} f^*\omega^{n-1} \wedge \alpha}{\int_{S_n} \alpha^n}.$$

Putting everything together, we obtain

$$\frac{\int_{S_n} f^*\beta \wedge \alpha}{\int_{S_n} f^*\omega^{n-1} \wedge \alpha} = \frac{\int_{S_n} f^*\beta \wedge f^*\omega}{\int_{S_n} f^*\omega^n}.$$

Since the above is true for all positive $(n-1, n-1)$ forms β with compact support on S_m and $f^* : \Omega^p(S_m) \to \Omega^p(S_n)$ is locally surjective because f is an immersion, we deduce that $f^*\omega = \lambda\alpha$ for

$$\lambda := \frac{\int_{S_n} f^*\omega^n}{\int_{S_n} f^*\omega^{n-1} \wedge \alpha} > 0.$$

As explained at the beginning of the proof, this implies that f is indeed a totally geodesic immersion, as required. □

The 'double equidistribution computation' employed above is reminiscent of an argument appearing in [5, 12], which was again motivated by the Zilber–Pink philosophy.

Remark 3.3. Technically speaking, using [2, Lemma 8.2], one only obtains the fact that any totally geodesic divisor Z is of the form $Z = \pi(\Gamma_n \backslash \Gamma_n g L)$ for some $g \in \mathrm{SU}(1, n)$, where L is an intermediate subgroup $W \subseteq L \subseteq N_G(W)$ and $gLg^{-1} \cap \Gamma_n$ is a lattice in gLg^{-1}. Moreover, there are many instances where it is necessary to take L strictly larger than W for $gLg^{-1} \cap \Gamma_n$ to be a lattice in gLg^{-1}. However, in the case where Γ_n is a lattice of simplest kind, one can always take $L = W$ which is what we are implicitly using in Lemma 3.2.

4. *The proof of Theorem 1.3*

We now prove Theorem 1.3 using some Hodge theoretic features behind the theory of Shimura varieties. This can serve as a warm-up case for later arguments that use the full power of the theory of variations of Hodge structures, though it only requires the language used in previous sections, allowing us to defer the introduction of more technical definitions and results.

Suppose that $S_n = \Gamma_n \backslash \mathbb{B}^n$ is an arithmetic ball quotient and that

$$f : S_n \longrightarrow S_m := \Gamma_m \backslash \mathbb{B}^m$$

is an immersive holomorphic map to another arithmetic ball quotient. From now on, we always assume that $f(S_n)$ is *monodromy generic*, that is, is not contained in any strict complex totally geodesic subvariety S' of S_m. Let Γ_f be the graph of f in $S_n \times S_m$. We now prove the following equivalent formulation of Theorem 1.3.

THEOREM 4.1. *Let $\{Z_\ell\}$ (respectively $\{Y_\ell\}$) be an infinite sequence of distinct complex totally geodesic subvarieties of S_n (respectively S_m). If $f(Z_\ell) \subset Y_\ell$ for all ℓ, then f is a totally geodesic immersion.*

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

We first state the *geometric Zilber–Pink conjecture*, which is a theorem in our special case. In fact, in the more general setting of arbitrary Shimura varieties, this is a theorem of Daw and Ren [**15**] (see also the more general Theorem 5.8, which will be used later). First, recall that an intersection of subvarieties is *atypical* if its dimension is smaller than the expected dimension coming from a differential topology codimension count.

THEOREM 4.2. (Special case of [**7**, **15**]) *Let Y be a closed irreducible subvariety of $S_n \times S_m$. If Y has a Zariski dense set of atypical intersections with totally geodesic subvarieties of positive dimension, then Y is contained in a strict complex totally geodesic subvariety $S_0 \subset S_n \times S_m$.*

Proof of Theorem 4.1. Define $U_\ell \coloneqq \Gamma_f \cap (Z_\ell \times Y_\ell)$. The key calculation in the proof is the following.

LEMMA 4.3. *The intersection U_ℓ between Γ_f and $Z_\ell \times Y_\ell$ is atypical for infinitely many ℓ, that is,*

$$\operatorname{codim}_{S_n \times S_m}(U_\ell) < \operatorname{codim}_{S_n \times S_m}(\Gamma_f) + \operatorname{codim}_{S_n \times S_m}(Z_\ell \times Y_\ell). \tag{4.1}$$

Proof. We can pass to a subsequence and assume that the dimensions of Z_ℓ and Y_ℓ are independent of ℓ. Set $d = \dim Z_\ell$ and $\dim Y_\ell = m - e$. We then have

$$\begin{aligned} \operatorname{codim}_{S_n \times S_m}(U_\ell) &= n + m - d, \\ \operatorname{codim}_{S_n \times S_m}(\Gamma_f) + \operatorname{codim}_{S_n \times S_m}(Z_\ell \times Y_\ell) &= m + (n + m - d - m + e) \\ &= n + m - d + e. \end{aligned}$$

Since $e > 0$, equation (4.1) holds. □

Since we assumed that $f(S_n)$ is monodromy generic, we must prove that $f(S_n) = S_m$. Applying Theorem 4.2 to $\Gamma_f \subset S_n \times S_m$, we conclude that Γ_f is contained in a **strict** totally geodesic subvariety S_0 of $S_n \times S_m$. As $f(S_n)$ is monodromy generic, the second projection of S_0 on S_m is surjective. Therefore, S_0 is a strict totally geodesic subvariety of $S_n \times S_m$ dominating both factors. This is possible only when S_n is commensurable with S_m and S_0 is a modular correspondence. Since $\Gamma_f \subset S_0$ have the same dimension and are irreducible, it follows that $\Gamma_f = S_0$, that is, f is a totally geodesic embedding. □

Using a similar argument, we can also prove the following, where a sequence of geodesics is *generic* when they are not contained in a finite union of maximal totally geodesic subvarieties.

COROLLARY 4.4. *Let $S_n \coloneqq \Gamma_n \backslash \mathbb{B}^n$ be a finite-volume ball quotient with an immersive holomorphic map*

$$f : S_n \longrightarrow S_m \coloneqq \Gamma_m \backslash \mathbb{B}^m$$

to another finite-volume ball quotient. Let $\{C_\ell\}$ be a generic sequence of distinct closed geodesics in S_n for which there exists a strict complex totally geodesic subvariety $Y_\ell \subset S_m$ such that $f(C_\ell) \subset Y_\ell$ for all ℓ. Then, f is a totally geodesic immersion.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

The proof of Corollary 4.4 uses the so-called *Baby Ax–Lindemann theorem* [**44**, Proposition 2.6], which we recall below in the special case we need.

PROPOSITION 4.5. *Let $Z \subset \mathbb{B}^n$ be a real totally geodesic subvariety (of any real dimension). Let $\Gamma \subset \mathrm{PU}(1,n)$ be an arithmetic lattice and $\pi : \mathbb{B}^n \to S_\Gamma = \Gamma\backslash\mathbb{B}^n$ be the associated locally Hermitian space. Then, the Zariski closure of $\pi(Z)$ is a complex totally geodesic subvariety of S_Γ.*

Proof of Corollary 4.4. Note that we may reduce to the case that S_m is arithmetic. Indeed, if S_m were non-arithmetic, the fact that $\{C_\ell\}$ is generic implies that, after passing to maximal proper totally geodesic subvarieties, $f(S_n)$ is contained in a finite union of totally geodesic subvarieties. Since the map is a smooth immersion, the image is contained in a single such subvariety S'. Replacing S_m with S', we see that S' must have infinitely many distinct maximal totally geodesic subvarieties (see [**2**] or [**8**]), and hence is arithmetic.

By assumption, $C_\ell \subset f^{-1}(Y_\ell)$. Since f is in fact an algebraic morphism by Borel's extension theorem [**11**], the right-hand side is Zariski closed and

$$Z_\ell := \overline{C_\ell} \subset f^{-1}(Y_\ell) \subsetneq S_n.$$

Thanks to Proposition 4.5, $\overline{C_\ell}$ is a complex totally geodesic subvariety of S_n, which we denote by Y_ℓ. Since the sequence $\{C_\ell\}$ is generic, we can apply Theorem 4.1 to conclude the proof. □

5. *Hodge theoretic preliminaries and a quick introduction to Zilber–Pink*

We begin this section by recalling the André–Oort and Zilber–Pink conjectures and some formalism from Hodge theory. Loosely speaking, our results on (closed) geodesics are inspired by the analogy between closed geodesics and CM points and the following. A point of a Shimura variety is called *CM* if its associated Mumford–Tate group is a torus.

THEOREM 5.1. (Special case of André–Oort in $S_1 \times S_2$, [**36**]) *A holomorphic map $f : S_1 \to S_2$ between Shimura varieties sending a Zariski dense set of CM points of S_1 to CM points of S_2 is a morphism of Shimura varieties. In particular, it is totally geodesic.*

For this, we refer the reader to [**7**, **28**].

Remark 5.2. It may be interesting to compare this to what happens with abelian varieties. If A, B are complex abelian varieties, any morphism $f : A \to B$ has totally geodesic image. This follows from a simple *rigidity lemma* as in [**35**, Corollary 1, p. 43]. If f sends a torsion point of A to a torsion point of B, then the aforementioned corollary shows that f is in fact a homomorphism. This latter statement is the analogue of Theorem 5.1 and shows that a version of Siu's immersion problem (Question 1.2) holds for abelian varieties.

5.1. *Hodge theory and period domains.*

We begin with some general notation and conventions. In what follows, an algebraic variety S is a reduced scheme of finite type over the field of complex numbers, but may be reducible. If S is an algebraic (respectively analytic) variety, by a subvariety $Y \subset S$, we always mean a *closed* algebraic (respectively analytic) subvariety.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

A $\mathbb{Q}$-Hodge structure of weight n on a finite-dimensional $\mathbb{Q}$-vector space V is a decreasing filtration $F^\bullet$ on the complexification $V_\mathbb{C}$ such that

$$V_\mathbb{C} = \bigoplus_{p\in\mathbb{Z}} F^p \oplus \overline{F^{n-p}}.$$

The category of pure $\mathbb{Q}$-Hodge structures is Tannakian and moreover semisimple when considering polarizable Hodge structures, as we frequently will. The Mumford–Tate group $\mathbf{MT}(V) \subset \mathbf{GL}(V)$ of a $\mathbb{Q}$-Hodge structure V is the Tannakian group of the Tannakian subcategory $\langle V\rangle^\otimes$ of $\mathbb{Q}$-Hodge structures generated by V. Equivalently, $\mathbf{MT}(V)$ is the smallest $\mathbb{Q}$-algebraic subgroup of $\mathbf{GL}(V)$ whose base-change to $\mathbb{R}$ contains the image of $h : \mathbb{S} \to \mathbf{GL}(V_\mathbb{R})$. It is also the stabilizer in $\mathbf{GL}(V)$ of the Hodge tensors for V. When V is polarized, this is a reductive algebraic group.

In the two sections below, we recall the setting and the main results from [7]. We refer also to the companion [6] (especially §2 in *op. cit.*) for a quick introduction to this circle of ideas.

5.2. *Typical and atypical intersections.* To understand a VHS $\mathbb{V} \to S$, we consider the associated holomorphic period map

$$\Phi : S^{\mathrm{an}} \longrightarrow \Gamma\backslash D, \tag{5.1}$$

which completely describes $\mathbb{V}$. We let $(\mathbf{G}, D)$ denote the generic Hodge datum of $\mathbb{V}$ and $\Gamma\backslash D$ the associated Hodge variety. The Mumford–Tate domain D decomposes as a product $D_1 \times \cdots \times D_k$ according to the decomposition of the adjoint group $\mathbf{G}^{\mathrm{ad}}$ into a product $\mathbf{G}_1 \times \cdots \times \mathbf{G}_k$ of simple factors, where some $\mathbf{G}_i$ may be $\mathbb{R}$-anisotropic. Replacing S by a finite étale cover and reordering the factors if necessary, the lattice $\Gamma \subset \mathbf{G}^{\mathrm{ad}}(\mathbb{R})^+$ decomposes as a direct product $\Gamma \cong \Gamma_1 \times \cdots \times \Gamma_r$ with $r \le k$, where $\Gamma_i \subset \mathbf{G}_i(\mathbb{R})^+$ is an arithmetic lattice for each i. Writing $D_{\mathrm{triv}} = D_{r+1} \times \cdots \times D_k$ for the product of factors where the monodromy is trivial, which includes all of the factors D_i for which $\mathbf{G}_i$ is $\mathbb{R}$-anisotropic, the period map can be written as

$$\Phi : S^{\mathrm{an}} \longrightarrow \Gamma\backslash D \cong \Gamma_1\backslash D_1 \times \cdots \times \Gamma_r\backslash D_r \times D_{\mathrm{triv}}, \tag{5.2}$$

where the projection of $\Phi(S^{\mathrm{an}})$ to D_{triv} is a point.

We distinguish between special subvarieties Z of *zero period dimension*, which are geometrically elusive, and those of *positive period dimension*.

Definition 5.3.

(1) A subvariety Z of S is said to be of *positive period dimension for* $\mathbb{V}$ if $\Phi(Z^{\mathrm{an}})$ has positive dimension, that is, if $\dim_\mathbb{C} \Phi(Z^{\mathrm{an}}) > 0$.

(2) The *Hodge locus of positive period dimension*, $\mathrm{HL}(S, \mathbb{V}^\otimes)_{\mathrm{pos}}$, is the union of the special subvarieties of S for which $\mathbb{V}$ has positive period dimension.

Using period maps, special subvarieties can also be defined as *intersection loci*. Indeed, a closed irreducible subvariety $Z \subset S$ is special for $\mathbb{V}$ precisely when Z^{an} coincides with an analytic irreducible component $\Phi^{-1}(\Gamma'\backslash D')^0$ of $\Phi^{-1}(\Gamma'\backslash D')$ for $(\mathbf{G}', D') \subset (\mathbf{G}, D)$

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

the generic Hodge subdatum of Z and $\Gamma'\backslash D' \subset \Gamma\backslash D$ the associated Hodge subvariety. We will equivalently say that Z is special for Φ.

Definition 5.4. Let $Z = \Phi^{-1}(\Gamma'\backslash D')^0 \subset S$ be a special subvariety for $\mathbb{V}$ with generic Hodge datum $(\mathbf{G}', D')$. Then, Z is said to be *atypical* if $\Phi(S^{\mathrm{an}})$ and $\Gamma'\backslash D'$ do not intersect generically along $\Phi(Z^{\mathrm{an}})$. That is, Z is atypical when

$$\operatorname{codim}_{\Gamma\backslash D} \Phi(Z^{\mathrm{an}}) < \operatorname{codim}_{\Gamma\backslash D} \Phi(S^{\mathrm{an}}) + \operatorname{codim}_{\Gamma\backslash D} \Gamma'\backslash D'. \tag{5.3}$$

Otherwise, Z is said to be *typical*. The *atypical Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})$ (respectively the *typical Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})$) is the union of the atypical (respectively strict typical) special subvarieties of S for $\mathbb{V}$.

Remark 5.5. Since it will be important in the following, we remark here that the notion of typicality and atypicality is always computed with respect to the smallest special subvariety of $\Gamma\backslash D$, usually called the special closure. It can, and does, happen that

$$Z = \Phi^{-1}(\Gamma'\backslash D')^0 = \Phi^{-1}(\Gamma''\backslash D'')^0$$

for some bigger $\Gamma''\backslash D''$, but computing codimensions with the latter would give the wrong conclusions.

Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on an irreducible smooth quasi-projective variety S, from [7], we expect the following conjecture.

Conjecture 5.6. (Zilber–Pink conjecture for the atypical Hodge locus, strong version) The atypical Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}}$ is a finite union of atypical special subvarieties of S for $\mathbb{V}$.

Conjecture 5.7. (Density of the typical Hodge locus) If $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is not empty, then it is analytically dense in S.

5.3. *Some results on the distribution of the Hodge locus.* We now recall some results from [7, Theorems 6.1, 10.1, and Remark 10.2]. To simplify the statements, the reader can first assume that the generic Mumford–Tate group is simple; however, later it will be used in a product situation.

THEOREM 5.8. (Geometric Zilber–Pink) *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and Z be an irreducible component of the Zariski closure of the union of the atypical special subvarieties of positive period dimension in S. Then, either:*

(a) *Z is a maximal atypical special subvariety; or*
(b) *the adjoint Mumford–Tate group $\mathbf{G}_Z^{\mathrm{ad}}$ decomposes as a non-trivial product $\mathbf{H}_Z^{\mathrm{ad}} \times \mathbf{L}_Z$, Z contains a Zariski-dense set of fibers of $\Phi_{\mathbf{L}_Z}$ which are atypical weakly special subvarieties of S for Φ, where (possibly up to an étale covering)*

$$\Phi_{|Z^{\mathrm{an}}} = (\Phi_{\mathbf{H}_Z}, \Phi_{\mathbf{L}_Z}) : Z^{\mathrm{an}} \longrightarrow \Gamma_{\mathbf{G}_Z}\backslash D_{\mathbf{G}_Z} = \Gamma_{\mathbf{H}_Z}\backslash D_{\mathbf{H}_Z} \times \Gamma_{\mathbf{L}_Z}\backslash D_{\mathbf{L}_Z} \subset \Gamma\backslash D,$$

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

and Z is Hodge generic in a special subvariety $\Phi^{-1}(\Gamma_{\mathbf{G}_Z}\backslash D_{\mathbf{G}_Z})^0$ of S for Φ which is monodromically typical and therefore typical.

We refer also to [9, Theorem 7.1] for an effective proof of a more general statement (proved using differential geometry rather than o-minimality).

THEOREM 5.9. *If the typical Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,typ}}$ is non-empty, then $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,typ}}$ is analytically dense in S.*

The above statement already hides an application of Theorem 5.8, as explained in [7, Remark 10.2]. Indeed, it is proved by first showing that $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}}$ is dense, then invoking Theorem 5.8 to say that $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}}$ is algebraic, and therefore is exactly $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,typ}}$, which is dense. In future applications, we will actually use this finer version. See also [24, Theorem 1.6 and Remark 1.7] for a related discussion.

Remark 5.10. Hodge theory actually gives a simple combinatorial criterion to decide whether $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is empty or not. Indeed, see the recent work [19, 24].

6. *General Hodge theoretic statement: from typical to atypical intersections*

6.1. *Isogenies between VHSs.* For an introduction to Tannakian categories, we refer for example to the article of Deligne and Milne [17]. The category of $\mathbb{Q}$VHS on a smooth quasi-projective base S, which we denote by $\mathbb{Q}\mathrm{VHS}/S$, is Tannakian (we can fix a fiber functor ω_s corresponding to some base point $s \in S$). Given a $\mathbb{Z}$VHS $\mathbb{V}$, we denote the associated $\mathbb{Q}$VHS by $\mathbb{V}_{\mathbb{Q}}$.

Definition 6.1. Let $\mathbb{V}_1, \mathbb{V}_2 \in \mathbb{Z}\ \mathrm{VHS}/S$. We say that $\mathbb{V}_1$ and $\mathbb{V}_2$ are *isogenous* if there is an equivalence of tensor categories $\langle \mathbb{V}_{1,\mathbb{Q}} \rangle^{\otimes} \cong \langle \mathbb{V}_{2,\mathbb{Q}} \rangle^{\otimes}$, where $\langle \mathbb{V}_{i,\mathbb{Q}} \rangle^{\otimes}$ denotes the smallest Tannakian subcategory of $\mathbb{Q}$VHSs containing $\mathbb{V}_{i,\mathbb{Q}}$.

Of course, the Tannakian categories $\langle \mathbb{V}_{i,\mathbb{Q}} \rangle^{\otimes}$ appearing above are equivalent (as tensor categories) to the category of finite-dimensional representations of their generic Mumford–Tate group. (The equivalence is realized by the functor of $\otimes$-automorphisms of the fiber functor). It can happen that two VHSs have isomorphic Mumford–Tate groups, but the isomorphism does not induce an equivalence of tenor categories, cf. the article of Deligne and Milne [17, Definition 1.10 and Theorem 2.11].

Remark 6.2. If two complex principally polarized abelian varieties A, B are isogenous in the usual sense (either via a polarized or an unpolarized isogeny), then the Hodge structures $H^1(A, \mathbb{Z})$, $H^1(B, \mathbb{Z})$ are also isogenous in the sense of Definition 6.1 (here, we take as base S the spectrum of $\mathbb{C}$). However, the converse is in general not true, for example, $H^1(A, \mathbb{Z})$ is isogenous to $H^1(A \times A, \mathbb{Z})$. However, for two principally polarized g-dimensional abelian varieties with Mumford–Tate group $\mathbf{GSp}_{2g}$, the two notions agree, since this is the case when the Mumford–Tate group is as big as possible.

Note that isogeny is an equivalence relation which we denote by $\mathbb{V}_1 \simeq \mathbb{V}_2$. Let $\Gamma_1\backslash D_1$ and $\Gamma_2\backslash D_2$ be two Mumford–Tate domains. A *modular correspondence* between $\Gamma_1\backslash D_1$

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

and $\Gamma_2\backslash D_2$ is a subvariety

$$V \subset \Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$$

such that the two projections $p_1 : V \to \Gamma_1\backslash D_1$ and $p_2 : V \to \Gamma_2\backslash D_2$ are finite étale surjective maps (see also [7, Definition 3.23] for more details). In this situation, V is a Mumford–Tate domain, that is, $V = \Gamma\backslash D$, and we have that $D \simeq D_1 \simeq D_2$ and that Γ_1 is commensurable with Γ_2. The following proposition translates the isogeny relation from a Tannakian statement to a more useful one on period maps.

PROPOSITION 6.3. *Let $\mathbb{V}_1, \mathbb{V}_2$ be $\mathbb{Z}$VHS on S with associated Hodge data $(\mathbf{G}_i, D_i)$ and period map ψ_i. The following are equivalent:*

(1) *there is an isogeny between $\mathbb{V}_1$ and $\mathbb{V}_2$;*
(2) *the Hodge data $(\mathbf{G}_1, D_1)$, $(\mathbf{G}_2, D_2)$ are isomorphic. Using $(\mathbf{G}, D)$ for the corresponding Hodge datum and $\psi_1 : S \to \Gamma_1\backslash D$, $\psi_2 : S \to \Gamma_2\backslash D$ for the period maps associated with $\mathbb{V}_1$, $\mathbb{V}_2$ (respectively), then there exists a modular correspondence*

$$\Gamma\backslash D \subset \Gamma_1\backslash D \times \Gamma_2\backslash D$$

and a period map $\psi_0 : S \to \Gamma\backslash D$ such that if $p_i : \Gamma\backslash D \to \Gamma_i\backslash D$ are the natural projections, then $\psi_i = p_i \circ \psi_0$ for $i \in \{1, 2\}$.

Proof. Let $\mathbb{V}$ be a $\mathbb{Z}$VHS on S with generic Hodge datum $(\mathbf{G}, D)$. Then, the Tannakian category generated by $\mathbb{V}_\mathbb{Q}$ is determined by $(\mathbf{G}, D)$, as any $\mathbb{W}$ in $\langle \mathbb{V}_\mathbb{Q} \rangle^\otimes$ is, by definition of a Tannakian category, given by a representation of $\mathbf{G}$ on the fiber $W_\mathbb{Q}$ of $\mathbb{W}_\mathbb{Q}$ at a point $s \in S$ and the Hodge structure on $W_\mathbb{Q}$ is determined by some Hodge morphism $\alpha : \mathbb{S} \to \mathbf{G}$ belonging to D. Therefore, if $(\mathbf{G}_1, D_1) \cong (\mathbf{G}_2, D_2)$, then $\mathbb{V}_1, \mathbb{V}_2$ are isogenous.

If $\mathbb{V}_1, \mathbb{V}_2$ are isogenous, they have isomorphic generic Mumford–Tate groups, denoted by $\mathbf{G}$, under the interpretation of the Mumford–Tate group in terms of Tannakian categories recalled above. Thus,

$$\langle \mathbb{V}_{1,\mathbb{Q}} \rangle^\otimes \cong \langle \mathbb{V}_{2,\mathbb{Q}} \rangle^\otimes \cong \langle \mathbb{V}_{1,\mathbb{Q}} \oplus \mathbb{V}_{2,\mathbb{Q}} \rangle^\otimes$$

and therefore the generic Mumford–Tate group of $\mathbb{V}_1 \oplus \mathbb{V}_2$ is also $\mathbf{G}$.

The $\mathbb{Z}$VHS $\mathbb{V}_1 \oplus \mathbb{V}_2$ corresponds to a period map

$$\psi_0 : S \longrightarrow \Gamma\backslash D \subset \Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2,$$

where $(\mathbf{G}, D)$ is the generic Hodge subdatum associated with $\mathbb{V}_1 \oplus \mathbb{V}_2$. Moreover, if the natural projections are $p_i : \Gamma\backslash D \to \Gamma_i\backslash D_i$, we have $\psi_i = p_i \circ \psi_0$. Let $\alpha_i : \mathbb{S} \to \mathbf{G}$ be Hodge morphisms such that $D_i = \mathbf{G}(\mathbb{R})^+\alpha_i$. Then, there exists an embedding of $\mathbf{G}$ in $\mathbf{G} \times \mathbf{G}$ such that D is the $\mathbf{G}(\mathbb{R})^+$-orbit of $\alpha = (\alpha_1, \alpha_2)$. Therefore, $D \simeq D_1 \simeq D_2$. □

6.2. *The main statement.*

THEOREM 6.4. *Let S be a smooth quasi-projective variety and $\mathbb{V}_1, \mathbb{V}_2$ two pure polarized $\mathbb{Z}$VHSs on S. Assume that the generic Mumford–Tate groups of $\mathbb{V}_1$ and $\mathbb{V}_2$ are $\mathbb{Q}$-simple. If*

$$\mathrm{HL}(S, \mathbb{V}_1^\otimes)_{\mathrm{pos,typ}} = \mathrm{HL}(S, \mathbb{V}_2^\otimes)_{\mathrm{pos,typ}} \neq \emptyset,$$

then $\mathbb{V}_1$ is isogenous to $\mathbb{V}_2$. As a consequence, $\mathrm{HL}(S, \mathbb{V}_1^\otimes) = \mathrm{HL}(S, \mathbb{V}_2^\otimes)$.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Some stronger variants with the *weakly special Hodge locus* replacing the Hodge locus can be given. Also, the same proof works if there exists a Zariski dense subset of components of $\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$ that is contained in $\mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}}$.

Remark 6.5. By Theorem 5.9, the condition $\mathrm{HL}(S, \mathbb{V}_i^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset$ implies that $\mathrm{HL}(S, \mathbb{V}_i^{\otimes})_{\mathrm{pos,typ}}$ is analytically dense in S. Conjecture 5.7 predicts that $\mathrm{HL}(S, \mathbb{V}_i^{\otimes})_{\mathrm{typ}}$ should be analytically dense whenever it is non-empty. The condition

$$\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}} = \mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset$$

is therefore quite restrictive. Moreover, by [7, Theorem 1.5], the levels of the $\mathbb{Z}$VHS $\mathbb{V}_1$ and $\mathbb{V}_2$ must be 1 (the Shimura case) or 2.

To better explain Theorem 6.4, we give a simple and more concrete corollary about two principally polarized families of abelian varieties $h_i : \mathcal{A}_i \to S$.

COROLLARY 6.6. *Fix $g \geq 2$. Let S be a smooth quasi projective variety and $h_i : \mathcal{A}_i \to S$ be two principally polarized g-dimensional families of abelian varieties whose monodromy group is $\mathbf{Sp}_{2g}$. Set $\mathbb{V}_i := R^1 h_{i,*}\mathbb{Z}$ for the associated VHS. If*

$$\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}} = \mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset, \tag{6.1}$$

then $\mathcal{A}_1$ is isogenous to $\mathcal{A}_2$ as principally polarized S-abelian schemes.

Proof. We will use Remark 6.2 and the following fact: given two principally polarized families of abelian varieties $h_i : \mathcal{A}_i \to S$, one has

$$\mathrm{Hom}_S(\mathcal{A}_1, \mathcal{A}_2) = \mathrm{Hom}(R_1 h_{1,*}\mathbb{Z}, R_1 h_{2,*}\mathbb{Z}),$$

where the second Hom is in the category of $\mathbb{Z}$VHS; see [16, Corollary 4.4.15]. Indeed, Theorem 6.4 implies that the two $\mathbb{Z}$VHSs $\mathbb{V}_i$ for $i \in \{1, 2\}$ are isogenous in the sense of Definition 6.1, and the above facts imply that this notion of isogeny translates precisely to the traditional one. □

Remark 6.7. The above statement bears some similarities with work of Baker and DeMarco [4] in complex dynamics, which shows that for any fixed $a, b \in \mathbb{C}$ and any integer $d \geq 0$, the set of $c \in \mathbb{C}$ for which both a and b are preperiodic for $z^d + c$ is infinite if and only if $a^d = b^d$. Also see generalizations like in the work of Yuan and Zhang [45]. Both proofs mentioned above use equidistribution, but here we work with positive dimensional subvarieties. The statement without *pos* is deeper and would follow from Conjecture 5.6.

Proof of Theorem 6.4. The main step in the proof is the following lemma, where we do not need to assume that the generic Mumford–Tate group of $\mathbb{V}_i$ is $\mathbb{Q}$-simple.

LEMMA 6.8. *Assuming the hypothesis of Theorem 6.4, let $f_i : S \to \Gamma_i \backslash D_i$ be the period map associated with $\mathbb{V}_i$, $i = 1, 2$, and let*

$$f_1 \times f_2 : S \longrightarrow \Gamma_1 \backslash D_1 \times \Gamma_2 \backslash D_2$$

be the period map associated with $\mathbb{V}_1 \oplus \mathbb{V}_2$. Then, $\mathrm{HL}(S, (\mathbb{V}_1 \oplus \mathbb{V}_2)^{\otimes})_{\mathrm{pos,atyp}}$ is analytically dense in S.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Proof. As $\mathrm{HL}(S, \mathbb{V}_i^{\otimes})_{\mathrm{pos,typ}} \neq \emptyset$, by Theorem 5.9, there exists sequences of special subvarieties $\{Z_{i,\ell}\}_{\ell\in\mathbb{N}}$ in $\Gamma_i\backslash D_i$ and a sequence $\{W_\ell\}_{\ell\in\mathbb{N}}$ of components of both $\mathrm{HL}(S, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$ and $\mathrm{HL}(S, \mathbb{V}_2^{\otimes})_{\mathrm{pos,typ}}$, such that

$$W_\ell = f_1^{-1}(Z_{1,\ell})^0 = f_2^{-1}(Z_{2,\ell})^0$$

for some components $f_i^{-1}(Z_{i,\ell})^0$ of $f_i^{-1}(Z_{i,\ell})$ and for which $\bigcup_\ell W_\ell$ is Zariski dense in S. The fact that the W_ℓ are again typical intersections, guaranteed by Theorem 5.9, implies that W_ℓ is Hodge generic for $\mathbb{V}_i$ in $Z_{i,\ell}$ (see also the discussion in Remark 5.5). By passing to a subsequence, we assume that $Z_{i,\ell}$ has fixed dimension z_i.

Since W_ℓ is realized in two ways as a typical intersection, we have

$$\mathrm{codim}_{\Gamma_i\backslash D_i}\ f_i(W_\ell) = \mathrm{codim}_{\Gamma_i\backslash D_i}\ Z_{i,\ell} + \mathrm{codim}_{\Gamma_i\backslash D_i}(f_i(S))$$

for each $i \in \{1, 2\}$. That is, if $w_i \coloneqq \dim\ f_i(W)$, $d_i \coloneqq \dim\ \Gamma_i\backslash D_i$, and $s_i \coloneqq \dim\ f_i(S)$, then

$$d_i - w_i = d_i - z_i + d_i - s_i. \tag{6.2}$$

Note that we do not assume that our period maps are immersive. Notice that

$$W_\ell \subset (f_1 \times f_2)^{-1}(Z_{1,\ell} \times Z_{2,\ell}) \subsetneq S.$$

Let W'_ℓ be an irreducible component of $(f_1 \times f_2)^{-1}(Z_{1,\ell} \times Z_{2,\ell})$ containing W_ℓ. We claim that W'_ℓ is an atypical intersection of positive period dimension in the Hodge locus $\mathrm{HL}(S, (\mathbb{V}_1 \oplus \mathbb{V}_2)^{\otimes})$ for all ℓ. For this, we need to show that

$$\begin{aligned}\mathrm{codim}_{\Gamma_1\backslash D_1\times\Gamma_2\backslash D_2}((f_1 \times f_2)(W'_\ell)) < \mathrm{codim}_{\Gamma_1\backslash D_1\times\Gamma_2\backslash D_2}(Z_{1,\ell} \times Z_{2,\ell})\\ + \mathrm{codim}_{\Gamma_1\backslash D_1\times\Gamma_2\backslash D_2}((f_1 \times f_2)(S)).\end{aligned}$$

Summing equation (6.2) for $i = 1, 2$, we have

$$d_1 + d_2 - w_1 - w_2 = d_1 + d_2 - z_1 - z_2 + d_1 + d_2 - s_1 - s_2. \tag{6.3}$$

This implies the requisite equation by noticing that

$$\max\{\dim(f_1(Y)), \dim(f_2(Y))\} \leq \dim((f_1 \times f_2)(Y)) \leq \dim(f_1(Y)) + \dim(f_2(Y))$$

for any Y in S. □

LEMMA 6.9. *Each of the following hold.*

(1) *If the monodromy of $\mathbb{V}_i$ is equal to the derived subgroup of its Mumford–Tate group $\mathbf{G}_i$, then $(f_1 \times f_2)(S)$ is not Hodge generic in $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$.*

(2) *If the Mumford–Tate group $\mathbf{G}_i$ of $\mathbb{V}_i$ is $\mathbb{Q}$-simple, then $\mathbf{G}_1^{\mathrm{ad}} \cong \mathbf{G}_2^{\mathrm{ad}}$, $D_1 \simeq D_2$. Denoting the corresponding Hodge datum by $(\mathbf{G}, D)$, we have that $(f_1 \times f_2)(S)$ is Hodge generic in a modular correspondence*

$$\Gamma\backslash D \subset \Gamma_1\backslash D \times \Gamma_2\backslash D.$$

Proof. By Lemma 6.8, S has a Zariski dense set of positive dimensional atypical components of the Hodge locus for the period map $f_1 \times f_2$. Therefore, S satisfies conclusion (a) or (b) of Theorem 5.8. If conclusion (a) holds, then $(f_1 \times f_2)(S)$ is not

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Hodge generic in $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$, as desired. If the monodromy of $\mathbb{V}_i$ is equal to the derived subgroup of its Mumford–Tate group $\mathbf{G}_i$, then the projection of $f_i(S)$ on every factor of $\Gamma_i\backslash D_i$ is positive dimensional and Hodge generic. The same is true for the period map $f_1 \times f_2$. Therefore, conclusion (b) is not satisfied. This finishes the proof of the first part of the lemma.

If the Mumford–Tate group $\mathbf{G}_i$ of $\mathbb{V}_i$ is $\mathbb{Q}$-simple, then the monodromy of $\mathbb{V}_i$ is equal to the derived subgroup of its Mumford–Tate group $\mathbf{G}_i$ and by the first part, $(f_1 \times f_2)(S)$ is not Hodge generic in $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$. Moreover, in this situation, the only strict special subvarieties of $\Gamma_1\backslash D_1 \times \Gamma_2\backslash D_2$ whose projections on both factors are surjective are the modular correspondences. This forces $D_1 \simeq D_2$ and $\mathbf{G}_1^{\mathrm{ad}} \cong \mathbf{G}_2^{\mathrm{ad}}$. This finishes the proof of the second part of the lemma. □

Theorem 6.4 is now a consequence of the characterization of isogenous $\mathbb{Z}$VHSs given in Proposition 6.3. □

Remark 6.10. Theorem 6.4 (or, more precisely, a slight generalization of it) implies Theorem 4.1 for ball quotients, which we now explain. On the smaller ball quotient S_n, which we assume to be arithmetic for simplicity, we have a $\mathbb{Z}$VHS $\mathbb{V}_1$ induced by any faithful linear representation of $\mathrm{GU}(1, n)$. The Hodge locus for such a VHS does not depend on the linear representation by a theorem of Deligne and it is equal to the union of all sub-Shimura varieties. With the terminology introduced in Definition 5.4, a sub-Shimura variety is always in the typical Hodge locus

$$\mathrm{HL}(S_n, \mathbb{V}_1^{\otimes})_{\mathrm{pos}} = \mathrm{HL}(S_n, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$$

and this locus is equal to the union of all totally geodesic subvarieties. Now, a map $f : S_n \to S_m$ gives another VHS $\mathbb{V}_2$ by pulling back the natural one on S_m. Once this translation from immersions to Hodge theory is made, the proof of Theorem 4.1 is just a special case of the one given in Theorem 6.4.

7. *Proofs of Theorems 1.13 and 1.17*

In this section, we prove the second part of Theorem 1.13 and then Theorem 1.17 following a similar argument. Let $\Gamma \subset \mathrm{PU}(1, n)$ for $n > 1$ be a lattice of the simplest kind, which we can assume to be neat (that is, no element of Γ has a non-trivial root of unity as an eigenvalue under the adjoint representation). Consider the smooth quasi-projective variety S_Γ and let

$$\rho : \Gamma \longrightarrow \mathbf{H}(k)$$

be a cohomologically rigid representation, that is, $H^1(\Gamma, \mathrm{Ad}_{\mathbf{H}}\, \rho) = 0$, where

$$\mathrm{Ad}_{\mathbf{H}} : \mathbf{H}(k) \longrightarrow \mathrm{Aut}(\mathfrak{h})$$

is the adjoint representation over some local field k. Assuming that ρ is geodesically rich in the sense of Definition 1.8, we would like to show that ρ extends. The proof is a generalization of the arguments from [**8**, §§4 and 5], improved by combining them with §6.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Proof of Theorem 1.13(2). The proof proceeds in two steps. The first is to build a rich $\mathbb{Z}$VHS, using the rigidity assumption, and the second is to use richness to finish the proof.

From the vanishing $H^1(\Gamma, \mathrm{Ad}_{\mathbf{H}}\,\rho) = 0$, we have that ρ underlies a $\mathbb{C}$VHS. This is essentially due to Simpson and is explained in detail in [8, §4]. Moreover,

$$\mathbb{Z}(\rho) := \mathbb{Z}[\mathrm{tr}\,\mathrm{Ad}_{\mathbf{H}}(\rho(\gamma)) : \gamma \in \Gamma]$$

has a natural embedding in $\mathcal{O}_K$, the ring of integers of a totally real number field K. There are two ways to show that $\mathbb{Z}(\rho)$ lies in $\mathcal{O}_K$, as opposed to the ring of S-integers for some finite set of finite places S of K. In [8, §4], it is done by invoking the recent work of Esnault and Groechenig [18]. However, here we can give a self-contained proof by applying Corollary 1.14.

From here forward, we assume that $\rho(\Gamma)$ is Zariski dense in the k-group $\mathbf{H}$ and will then prove that $\mathbf{H}(K) \cong G$. In particular, from the above, we obtain a natural K-form of $\mathbf{H}$ (that we denote by the same symbol $\mathbf{H}$). Let $\sigma_i : K \to \mathbb{R}$ be the places of K, ordered in such a way that σ_1 is the identity for the given lattice embedding. Let $\widehat{\mathbf{H}} \subseteq \mathrm{Res}_{K/\mathbb{Q}}\,\mathbf{H}$ and consider the $\mathbb{Z}$-local system

$$\begin{aligned} \hat{\rho} : \Gamma &\longrightarrow \widehat{\mathbf{H}}(\mathbb{R}) \\ \gamma &\longmapsto \hat{\rho}(\gamma) = (\sigma_i(\rho(\gamma))). \end{aligned}$$

Again, by cohomological rigidity of ρ, we have that $\hat{\rho}$ naturally underlies a $\mathbb{Z}$VHS $\widehat{\mathbb{V}}$ (compare with [8, Theorem 4.2.4]). As recalled in §5, $\widehat{\mathbb{V}}$ corresponds to a period map

$$\psi : S_\Gamma^{\mathrm{an}} \longrightarrow \widehat{\mathbf{H}}(\mathbb{Z})\backslash D.$$

Now that we have a $\mathbb{Z}$VHS, as in Remark 6.10, we use the results of §6 to prove the theorem. Let $S_\Gamma = \Gamma\backslash\mathbb{B}^n$ be the associated ball quotient, and $\mathbb{V}_1$ the $\mathbb{Z}$VHS on S_Γ associated with some faithful linear rational representation of $\mathbf{G}$. By construction, $\mathrm{HL}(S_\Gamma, \mathbb{V}_1^{\otimes})_{\mathrm{pos}}$ is the union of the complex totally geodesic subvarieties of S_Γ, and it is equal to $\mathrm{HL}(S_\Gamma, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$; in particular, it is non-empty.

It is straightforward to see that the fact that ρ is complex rich implies that $\hat{\rho}$ is complex rich as well. Indeed, let $S_{\Gamma_i} \subset S_\Gamma$ be a totally geodesic subvariety such that $\mathbf{H}_i = \overline{\rho(\Gamma_i)}$ has dimension smaller than $\dim \mathbf{H}$. Then, $\dim \widehat{\mathbf{H}}_i := \overline{\hat{\rho}(\Gamma_i)} < \dim \widehat{\mathbf{H}}$. Moreover, S_{Γ_i} lies in the $\mathrm{HL}(S_\Gamma, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$, as well as $\mathrm{HL}(S_\Gamma, \widehat{\mathbb{V}}^{\otimes})_{\mathrm{pos,typ}}$. Since this is true for a sequence of i, (the proof of) Theorem 6.4 implies that $\mathbb{V}_1, \widehat{\mathbb{V}}$ are isogenous, which means that the period domain $\widehat{\mathbf{H}}(\mathbb{Z})\backslash D$ is isomorphic to S_Γ and that ψ is isogenous to the identity (as in Proposition 6.3). Therefore, $\hat{\rho}$ extends, which is only possible if the initial representation ρ extends. That is, ψ is an isomorphism. This concludes the proof of the desired statement. □

Proof of Theorem 1.17. The proof follows the same argument given above. The only difference is we are given a $\mathbb{Q}$VHS $\mathbb{V}$ on the algebraic variety S_Γ. By applying Corollary 1.14, $\mathbb{V}$ must be a $\mathbb{Z}$VHS, and the richness assumption implies that there is a generic sequence of subvarieties $S_{\Gamma_i} \in \mathrm{HL}(S_\Gamma, \mathbb{V}^{\otimes})_{\mathrm{pos,typ}} \cap \mathrm{HL}(S_\Gamma, \mathbb{V}_1^{\otimes})_{\mathrm{pos,typ}}$. Therefore, the results follow once again from Theorem 6.4. □

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

Acknowledgements. G.B. and E.U. were partially supported by the NSF grant DMS-1928930, while they were in residence at the MSRI in Berkeley, during the Spring 2023 semester. N.M. was partially supported by Grant Number DMS-2005438/2300370 and M.S. was partially supported by Grant Number DMS-2203555 from the National Science Foundation. G.B. also thanks the IHES for the excellent working conditions.

References

[1] U. Bader, D. Fisher, N. Miller and M. Stover. Arithmeticity, superrigidity, and totally geodesic submanifolds. *Ann. of Math. (2)* **193**(3) (2021), 837–861.
[2] U. Bader, D. Fisher, N. Miller and M. Stover. Arithmeticity, superrigidity and totally geodesic submanifolds of complex hyperbolic manifolds. *Invent. Math.* **233**(1) (2023), 169–222.
[3] U. Bader and A. Furman. An extension of Margulis's superrigidity theorem. *Dynamics, Geometry, Number Theory: The Impact of Margulis on Modern Mathematics*. Ed. D. Fisher, D. Kleinbock and G. Soifer. The University of Chicago Press, Chicago, IL, 2022, pp. 47–65.
[4] M. Baker and L. DeMarco. Preperiodic points and unlikely intersections. *Duke Math. J.* **159**(1) (2011), 1–29.
[5] G. Baldi. On a conjecture of Buium and Poonen. *Ann. Inst. Fourier (Grenoble)* **70**(2) (2020), 457–477.
[6] G. Baldi, B. Klingler and E. Ullmo. Non-density of the exceptional components of the Noether–Lefschetz locus. *Int. Math. Res. Not. IMRN* **2024**(21) (2024), 13642–13650.
[7] G. Baldi, B. Klingler and E. Ullmo. On the distribution of the Hodge locus. *Invent. Math.* **235**(2) (2024), 441–487.
[8] G. Baldi and E. Ullmo. Special subvarieties of non-arithmetic ball quotients and Hodge theory. *Ann. of Math. (2)* **197**(1) (2023), 159–220.
[9] G. Baldi and D. Urbanik. Effective atypical intersections and applications to orbit closures. *Preprint*, 2024, arXiv:2406.16628.
[10] G. Besson, G. Courtois and S. Gallot. Lemme de Schwarz réel et applications géométriques. *Acta Math.* **183**(2) (1999), 145–169.
[11] A. Borel. Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem. *J. Differential Geom.* **6** (1972), 543–560.
[12] A. Buium and B. Poonen. Independence of points on elliptic curves arising from special points on modular and Shimura curves. I. Global results. *Duke Math. J.* **147**(1) (2009), 181–191.
[13] H.-D. Cao and N. Mok. Holomorphic immersions between compact hyperbolic space forms. *Invent. Math.* **100**(1) (1990), 49–61.
[14] K. Corlette. Flat G-bundles with canonical metrics. *J. Differential Geom.* **28**(3) (1988), 361–382.
[15] C. Daw and J. Ren. Applications of the hyperbolic Ax-Schanuel conjecture. *Compos. Math.* **154**(9) (2018), 1843–1888.
[16] P. Deligne. Théorie de Hodge. II. *Publ. Math. Inst. Hautes Études Sci.* **40** (1971), 5–57.
[17] P. Deligne, J. S. Milne, A. Ogus and K.-Y. Shih. *Hodge Cycles, Motives, and Shimura Varieties (Lecture Notes in Mathematics, 900)*. Springer, Cham, 1982.
[18] H. Esnault and M. Groechenig. Cohomologically rigid local systems and integrality. *Selecta Math. (N.S.)* **24**(5) (2018), 4279–4292.
[19] S. Eterović and T. Scanlon. Likely intersections. *Preprint*, 2023, arXiv:2211.10592.
[20] D. Fisher. Superrigidity, arithmeticity, normal subgroups: results, ramifications, and directions. *Dynamics, Geometry, Number Theory: The Impact of Margulis on Modern Mathematics*. Ed. D. Fisher, D. Kleinbock and G. Soifer. University of Chicago Press, Chicago, IL, 2022, pp. 9–46.
[21] W. M. Goldman and J. J. Millson. Local rigidity of discrete groups acting on complex hyperbolic space. *Invent. Math.* **88**(3) (1987), 495–520.
[22] M. Gromov and R. Schoen. Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one. *Publ. Math. Inst. Hautes Études Sci.* **76** (1992), 165–246.
[23] D. Johnson and J. J. Millson. Deformation spaces associated to compact hyperbolic manifolds. *Discrete Groups in Geometry and Analysis: Papers in Honor of G. D. Mostow 60th Birthday*. Ed. R. Howe. Birkhäuser Boston, Inc., Boston, MA, 1987, pp. 48–106.
[24] N. Khelifa and D. Urbanik. Existence and density of typical Hodge loci. *Preprint*, 2024, arXiv:2303.16179.
[25] D. Kim and H. Oh. Rigidity of Kleinian groups via self-joinings: measure theoretic criterion. *Preprint*, 2023, arXiv:2302.03552.
[26] D. Kim and H. Oh. Rigidity of Kleinian groups via self-joinings. *Invent. Math.* **234**(3) (2023), 937–948.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press

[27] B. Klingler. Local rigidity for complex hyperbolic lattices and Hodge theory. *Invent. Math.* **184**(3) (2011), 455–498.
[28] B. Klingler, E. Ullmo and A. Yafaev. Bi-algebraic geometry and the André–Oort conjecture. *Algebraic Geometry: Salt Lake City 2015* (*Proceedings of Symposia in Pure Mathematics, 97)*. Ed. T. de Fernex, B. Hassett, M. Mustaţă, M. Olsson, M. Popa and R. Thomas. American Mathematical Society, Providence, RI, 2018, pp. 319–359.
[29] V. Koziarz and J. Maubon. Maximal representations of uniform complex hyperbolic lattices. *Ann. of Math. (2)* **185**(2) (2017), 493–540.
[30] V. Koziarz and J. Maubon. Finiteness of totally geodesic exceptional divisors in Hermitian locally symmetric spaces. *Bull. Soc. Math. France* **146**(4) (2018), 613–631.
[31] G. A. Margulis. *Discrete Subgroups of Semisimple Lie Groups (Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 17)*. Springer-Verlag, Berlin, 1991.
[32] N. Mok, Y. T. Siu and S.-K. Yeung. Geometric superrigidity. *Invent. Math.* **113**(1) (1993), 57–83.
[33] M. Möller and D. Toledo. Bounded negativity of self-intersection numbers of Shimura curves in Shimura surfaces. *Algebra Number Theory* **9**(4) (2015), 897–912.
[34] G. D. Mostow. On a remarkable class of polyhedra in complex hyperbolic space. *Pacific J. Math.* **86**(1) (1980), 171–276.
[35] D. Mumford. *Abelian Varieties (Tata Institute of Fundamental Research, Studia Mathematica, 5)*. Oxford University Press, London, 1970.
[36] J. Pila, A. N. Shankar, J. Tsimerman, H. Esnault and M. Groechenig. Canonical heights on Shimura varieties and the André–Oort conjecture. *Preprint*, 2022, arXiv:2109.08788.
[37] G. Prasad and A. S. Rapinchuk. Generic elements of a Zariski-dense subgroup form an open subset. *Trans. Moscow Math. Soc.* **78** (2017), 299–314.
[38] M. Ratner. On Raghunathan's measure conjecture. *Ann. of Math. (2)* **134**(3) (1991), 545–607.
[39] Y. Shalom. Rigidity, unitary representations of semisimple groups, and fundamental groups of manifolds with rank one transformation group. *Ann. of Math. (2)* **152**(1) (2000), 113–182.
[40] C. T. Simpson. The ubiquity of variations of Hodge structure. *Complex Geometry and Lie Theory. Proceedings of a Symposium (Sundance, UT, USA, May 26–30, 1989)*. Ed. J. Carlson, C. H. Clemens and D. Morrison. American Mathematical Society, Providence, RI, 1991, pp. 329–348.
[41] Y.-T. Siu. Some recent results in complex manifold theory related to vanishing theorems for the semipositive case. *Arbeitstagung Bonn, 1984: Proceedings of the Meeting held by the Max-Planck-Institut für Mathematik, Bonn, June 15–22, 1984* (*Lecture Notes in Mathematics, 111)*. Ed. F. Hirzebruch, J. Schwermer and S. Suter. Springer, Berlin, 1985, pp. 169–192.
[42] S. Tayou and N. Tholozan. Equidistribution of Hodge loci. II. *Compos. Math.* **159**(1) (2023), 1–52.
[43] D. Toledo. Maps between complex hyperbolic surfaces. *Geom. Dedicata* **97** (2003), 115–128.
[44] E. Ullmo and A. Yafaev. Algebraic flows on Shimura varieties. *Manuscripta Math.* **155**(3–4) (2018), 355–367.
[45] S. Yuan and S. Zhang. Calabi–Yau theorem and algebraic dynamics. *Preprint*, 2010.

https://doi.org/10.1017/etds.2024.136 Published online by Cambridge University Press