Invent. math. (2023) 233:169–222
https://doi.org/10.1007/s00222-023-01186-5

Arithmeticity, superrigidity and totally geodesic submanifolds of complex hyperbolic manifolds

Uri Bader[1] · David Fisher[2] · Nicholas Miller[3] · Matthew Stover[4]

To Domingo Toledo with admiration on the occasion of his 75th birthday

Accepted: 2 February 2023 / Published online: 1 March 2023

Abstract For $n \geq 2$, we prove that a finite volume complex hyperbolic n-manifold containing infinitely many maximal properly immersed totally geodesic submanifolds of real dimension at least two is arithmetic, paralleling our previous work for real hyperbolic manifolds. As in the real hyperbolic case, our primary result is a superrigidity theorem for certain representations of complex hyperbolic lattices. The proof requires developing new general tools not needed in the real hyperbolic case. Our main results also have a number of other applications. For example, we prove nonexistence of certain maps between complex hyperbolic manifolds, which is related to a question of Siu, that certain hyperbolic 3-manifolds cannot be totally geodesic submanifolds

✉ David Fisher
davidfisher@rice.edu

Uri Bader
bader@weizmann.ac.il

Nicholas Miller
nickmbmiller@ou.edu

Matthew Stover
mstover@temple.edu

1 Faculty of Mathematics and Computer Science, Weizmann Institute of Science, 234 Herzl Street, Rehovot 7610001, Israel

2 Department of Mathematics, Rice University, Houston, TX 77005, USA

3 Department of Mathematics, University of Oklahoma, Norman, OK 73019, USA

4 Department of Mathematics, Temple University, Philadelphia, PA 19122, USA

of complex hyperbolic manifolds, and that arithmeticity of complex hyperbolic manifolds is detected purely by the topology of the underlying complex variety, which is related to a question of Margulis. Our results also provide some evidence for a conjecture of Klingler that is a broad generalization of the Zilber–Pink conjecture.

1 Introduction

Throughout this paper, a *geodesic submanifold* will always mean a properly immersed totally geodesic subspace, and a geodesic submanifold is called *maximal* if it is not contained in a proper geodesic submanifold of larger dimension. Here and throughout a finite volume complex hyperbolic n-manifold will always have complex dimension n. In this paper, we will prove the following.

Theorem 1.1 (Arithmeticity) *Suppose that $n \geq 2$ and M is a finite volume complex hyperbolic n-manifold containing infinitely many maximal totally geodesic submanifolds of real dimension at least* 2. *Then M is arithmetic.*

In previous work, we proved arithmeticity of *real* hyperbolic manifolds containing infinitely many maximal geodesic subspaces of dimension at least two [2, Thm. 1.1], which answered a question independently due to Alan Reid and Curtis McMullen [19]. As in that case, Theorem 1.1 can be restated purely in terms of homogeneous dynamics, where one obtains the exact same statement as [2, Thm. 1.5]. We also note that, after learning of our result, Baldi and Ullmo recently gave a very different proof of Theorem 1.1 in the special case of totally geodesic complex subvarieties [5].

To answer Reid and McMullen's question we introduced the notion of *compatibility* of an algebraic group $\mathbf{H}$ over a local field k with a semisimple Lie group G and used this with $G = \mathrm{SO}_0(n, 1)$ to prove a *superrigidity theorem* for certain representations of real hyperbolic lattices [2, Thm. 1.6]. However, for $G = \mathrm{SU}(n, 1)$ superrigidity with compatible targets is not enough to prove arithmeticity. Thus, in this paper we introduce new tools for proving superrigidity theorems in rank one with target groups that are not compatible with G. While these new tools can be used for $\mathrm{SO}_0(n, 1)$, see for instance Remark 6.8, our focus will be to apply these ideas in the case $G = \mathrm{SU}(n, 1)$ to prove Theorem 1.1 and the other applications described later in this introduction.

It is well-known that if G is a connected adjoint semisimple Lie group with no compact factors and $\Gamma < G$ is an irreducible nonarithmetic lattice, then G is necessarily isomorphic to either $\mathrm{PO}_0(n, 1)$ or $\mathrm{PU}(n, 1)$ for some $n \geq 2$. Therefore, combining [2, Thm. 1.1] and Theorem 1.1 with the fact that geodesic subspaces of arithmetic manifolds are arithmetic yields:

Corollary 1.2 (Finiteness) *Let M be a nonarithmetic finite volume irreducible locally symmetric space of noncompact type. Then*

(1) *M contains only finitely many maximal geodesic subspaces of real dimension at least* 2*, and*
(2) *M contains only finitely many nonarithmetic geodesic subspaces of real dimension at least* 2.

Note that a complex hyperbolic n-manifold can contain geodesic submanifolds of real dimension k that are real hyperbolic ($2 \leq k \leq n$) or complex hyperbolic ($2 \leq k \leq 2n-2$ even), see Sect. 9.1 for examples. We also remark that there are currently only 24 known commensurability classes of nonarithmetic complex hyperbolic manifolds of finite volume, twenty-two in complex dimension 2 and two in complex dimension 3 [20]. Finding more examples or any examples in complex dimension 4 and higher is an important open problem, e.g., see [40, Prob. 9] or [30, Conj. 2.6].

We also note that the current technology for building nonarithmetic examples is consistent with the strategies in the real hyperbolic setting in that it produces examples with special geodesic subspaces. In particular, all known nonarithmetic complex hyperbolic lattices are commensurable with complex reflection groups; see [21] for a comprehensive discussion for PU(2, 1). There are also some attempts to use a hybrid construction inspired by the work of Gromov and Piatetski-Shapiro [25], for example [30, Conj. 2.7] or [51]. In this direction Theorem 1.1 can be viewed as a criterion for (non)arithmeticity that can be applied when one has intimate knowledge of the collection of geodesic submanifolds, and we will interpret this criterion in terms of the topology of M as an algebraic variety in Theorem 1.6 below.

We now state our general superrigidity result. Part (1) mirrors [2, Thm. 1.6]. However, the tools developed in [2] cannot be used to prove part (2). The heart of this paper develops new tools for proving superrigidity and applies them to prove part (2).

Theorem 1.3 (Superrigidity) *Let G be* SU(n, 1) *for $n \geq 2$, $W < G$ be a noncompact connected almost simple subgroup, and $\Gamma < G$ be a lattice. Suppose that k is a local field, $\mathbf{H}$ is a connected adjoint k-algebraic group, and $\rho : \Gamma \to \mathbf{H}(k)$ is a homomorphism with unbounded, Zariski dense image. Moreover, suppose that there is a faithful irreducible representation of $\mathbf{H}(k)$ on a k-vector space V of dimension at least two and a W-invariant, ergodic measure ν on $(G \times \mathbb{P}(V))/\Gamma$ that projects to Haar measure on G/Γ. If either*

(1) *the pair consisting of k and $\mathbf{H}$ is compatible with G, or*
(2) *$k = \mathbb{R}$ and $\mathbf{H}(\mathbb{R}) \cong$* PU($n$, 1),

then ρ extends to a continuous homomorphism from G to $\mathbf{H}(k)$.

We briefly recall the definition of compatibility from [2]. Let P be a minimal parabolic subgroup of G and U its unipotent radical (see Sect. 2.2). A pair consisting of a local field k and a k-algebraic group $\mathbf{H}$ is said to be *compatible*

with G if for every nontrivial k-subgroup $\mathbf{J} < \mathbf{H}$ and any continuous homomorphism $\tau : P \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$, where $N_{\mathbf{H}}(\mathbf{J})$ is the normalizer of $\mathbf{J}$ in $\mathbf{H}$, we have that the Zariski closure of $\tau(U')$ coincides with the Zariski closure of $\tau(U)$ for every nontrivial subgroup $U' < U$. Note that minimal parabolic subgroups are unique up to conjugation, so this is a property of G, not just P. We will extend this definition and elaborate on it in Sect. 3.

Remark 1.4 Margulis asked when superrigidity holds for arithmetic complex hyperbolic lattices [40, Prob. 9], and our results provide a partial answer to his question. Indeed, Theorem 1.3 proves superrigidity of certain representations of arithmetic complex hyperbolic lattices, and we will describe applications of this later in the introduction. We also note that there are a number of previous superrigidity results for particular representations of complex hyperbolic lattices. For example, see the famous work of Toledo [61] and Corlette [17] on what are now known as *maximal representations*, and see more recent work of Burger–Iozzi [11], Pozzetti [52], and Koziarz–Maubon [36] for more on results of this kind and historical remarks. Another famous example is the proof by Klingler of superrigidity of low-dimensional representations of fundamental groups of fake projective planes, which he then used to deduce their arithmeticity [33]; see [34] for a more general result along these lines.

We now describe the general ideas behind the proof of Theorem 1.3(2). The failure of $\mathrm{PU}(n, 1)$ to be compatible with $\mathrm{SU}(n, 1)$ can be measured precisely, and we describe this in Sect. 3. Specifically, we define an *incompatibility datum* for a group G with respect to an algebraic group $\mathbf{H}$ over a local field k as a measure of the failure of compatibility. When G is a real semisimple group, Proposition 3.2 shows that the only relevant case is when k is $\mathbb{R}$ or $\mathbb{C}$ and Lemma 3.3 proves that an incompatibility datum for the pair $(k, \mathbf{H})$ is associated with a particular parabolic subgroup Q of $\mathbf{H}(k)$.

When $G = \mathrm{SU}(n, 1)$, $k = \mathbb{R}$, and $\mathbf{H}(k) = \mathrm{PU}(n, 1)$, we use this general setup to produce a measurable self-map of the boundary of complex hyperbolic space having certain equivariance properties. A delicate fiber product argument translates the original map into an incidence geometry problem about *chains* on the boundary of complex hyperbolic space. This problem was solved by Pozzetti [52, Thm. 1.6] (see Theorem 6.5 for a statement in our language), and this allows us to complete the proof of Theorem 1.3(2). More generally, the tools developed in Sect. 3 can be used to turn some superrigidity questions into problems about incidence geometry on boundaries. This is reminiscent of both Mostow's proof of strong rigidity in higher rank and the proof by Margulis and Mohammadi of the analogue of Theorem 1.1 for compact hyperbolic 3-manifolds [44] (see Remark 6.8).

The basic strategy for proving Theorem 1.1 is to employ the general outline for Margulis's proof of arithmeticity of higher rank lattices [41,42]. This is

the same strategy used in [2] for the real hyperbolic case; see Sect. 8.2 for discussion. However, we need more than Theorem 1.3 to prove Theorem 1.1 and we now describe the additional input we require.

Suppose that M is a finite volume complex hyperbolic n-manifold, $n \geq 2$, and that $\Gamma < \mathrm{SU}(n, 1)$ is a lattice so that M is biholomorphic to $\Gamma\backslash\mathbb{B}^n$. Canonically associated with Γ is its *adjoint trace field* ℓ, and the Zariski closure of Γ under the adjoint representation is an adjoint simple ℓ-algebraic group $\overline{\mathbf{G}}$. The projection $\Gamma \to \mathrm{PU}(n, 1)$ is determined by a real place v_0 of ℓ so that $\overline{\mathbf{G}}(\ell_{v_0}) \cong \mathrm{PU}(n, 1)$, where ℓ_v denotes the completion of ℓ at a place v. Then Γ is arithmetic if and only if Γ is precompact in $\overline{\mathbf{G}}(\ell_v)$ for all places $v \neq v_0$ of ℓ.

In proving the analogue of Theorem 1.1 in the real hyperbolic setting, the pairs ℓ_v and $\overline{\mathbf{G}}$ one encounters are always compatible with $\mathrm{SO}_0(n, 1)$ and so the analogue of Theorem 1.3(1) is all we needed. Certain pairs of interest to us in the proof of Theorem 1.1, particularly $(\mathbb{R}, \mathrm{PU}(r, s))$ with $2 \leq s \leq r < n$ and $r + s = n + 1$, are compatible with $\mathrm{SU}(n, 1)$. Theorem 1.3(2) takes care of the additional case where $\ell_v = \mathbb{R}$ and $\overline{\mathbf{G}}(\ell_v) \cong \mathrm{PU}(n, 1)$. However, the above is not sufficient to handle the pair $(\mathbb{C}, \mathrm{SL}_{n+1}(\mathbb{C}))$.

To finish the proof of Theorem 1.1, we apply Simpson's celebrated results on linear representations of Kähler groups [58] and their generalization to the quasi-projective case. These methods actually lead to a more general theorem about the possibilities for the adjoint trace field ℓ of a lattice Γ in $\mathrm{SU}(n, 1)$ and the associated ℓ-algebraic group, namely the ℓ-Zariski closure of Γ under the adjoint representation (see Sect. 7 for the precise definition). In Sect. 7 we will prove the following, which is known to experts but is not in the literature.

Theorem 1.5 (Hodge type and integral) *Let $\Gamma < \mathrm{SU}(n, 1)$ be a lattice, ℓ be its adjoint trace field, and $\mathbf{G}$ be the absolutely almost simple simply connected ℓ-algebraic group canonically associated with Γ. Then:*

(1) *ℓ is totally real and the quadratic extension ℓ'/ℓ associated with $\mathbf{G}$ as a group of type ${}^2\mathrm{A}_n$ is totally complex;*
(2) *for each real place v of ℓ, $\mathbf{G}(\ell \otimes_v \mathbb{R})$ is isomorphic to $\mathrm{SU}(r_v, s_v)$ for some $r_v, s_v \geq 0$ with $r_v + s_v = n + 1$;*
(3) *Γ is integral, i.e., there is an action of $\mathbf{G}(\ell)$ on an ℓ-vector space V and an $\mathcal{O}_\ell$-lattice $\mathcal{L} \subset V$ so that a finite index subgroup of Γ stabilizes $\mathcal{L}$.*

In particular, if v_0 is the place of ℓ associated with the lattice embedding of Γ in $\mathrm{SU}(n, 1)$, then Γ is arithmetic if and only if $\mathbf{G}(\ell_v) \cong \mathrm{SU}(n + 1)$ for all archimedean places $v \neq v_0$ of ℓ.

The first two statements and their proofs were described to us by Domingo Toledo, and the third follows from work of Esnault–Groechenig [22]. To prove Theorem 1.1, we only need parts (1) and (2) of Theorem 1.5, and in the cocompact case these are immediate from Simpson's result that rigid representations

of Kähler groups are of Hodge type [58, Lem. 4.5]. There is in fact considerable overlap between the cases in Theorem 1.1 covered by Theorems 1.3(1) and 1.5. Theorem 1.5(3) rules out the cases where ℓ_v is nonarchimedean, but these cases are perhaps more easily handled by Theorem 1.3(1), since any simple algebraic group over a nonarchimedean local field is compatible with G by Proposition 3.2.

We now describe some other applications and interpretations of our results in the language of algebraic and complex geometry.

One application of Theorem 1.1 is to provide evidence for a conjecture of Klingler [32, Conj. 1.12]. Let $M = \mathbb{B}^n/\Gamma$ be a finite volume complex hyperbolic manifold of complex dimension $n \geq 2$. As discussed above, there exists a totally real number field with fixed embedding $v_0 : \ell \to \mathbb{R}$ and an almost simple ℓ-algebraic group $\mathbf{G}$ with $\mathbf{G}(\ell \otimes_{v_0} \mathbb{R}) \cong \mathrm{SU}(n, 1)$ such that, up to passing to a subgroup of finite index, $\Gamma \subseteq \mathbf{G}(\mathcal{O}_\ell)$, where $\mathcal{O}_\ell$ denotes the ring of integers of ℓ. Let $\mathbf{H}$ denote the Weil restriction of scalars from ℓ to $\mathbb{Q}$ of $\mathbf{G}$, which is a semisimple $\mathbb{Q}$-group with $\Gamma \subseteq \mathbf{H}(\mathbb{Z})$. Any faithful representation $\rho : \mathbf{H} \to \mathrm{GL}(V)$ defined over $\mathbb{Z}$ induces a polarizable $\mathbb{Z}$-variation of Hodge structure $\mathbb{V}$ on M.

As soon as Γ is nonarithmetic, the group $\mathbf{H}(\ell \otimes_{\mathbb{Q}} \mathbb{R})$ admits at least two noncompact factors and one easily checks that any totally geodesic subvariety of M is atypical for $(M, \mathbb{V})$ in the sense of [32, Def. 1.7]. One can in fact check that maximal totally geodesic subvarieties of M are optimal in the sense of [32, Def. 1.8]. Then [32, Conj. 1.12] implies that M contains at most finitely many maximal totally geodesic subvarieties. Therefore, our results confirm this consequence of [32, Conj. 1.12]. See [5] for more about these connections.

An application in another direction is to Margulis's question as to whether arithmeticity can be detected at the topological level. We explain in Sect. 9.2 how Theorem 1.1 can be interpreted as saying that arithmeticity is detected by the topology of the complex variety underlying a complex hyperbolic manifold. In fact, arithmeticity is recognized by the structure of the intersection product on cohomology. For simplicity, we state our result for complex dimension 2n here and refer to Sect. 9.2 for more on the higher-dimensional case.

Suppose X is a smooth complex projective variety with canonical divisor K_X and $D \subset X$ is a (possibly empty) smooth divisor such that $M = X \smallsetminus D$ is a complex hyperbolic 2-manifold for which X is a smooth toroidal compactification of M. We recall that any complex projective curve C on X satisfies

$$3\, C \cdot C + 3 \deg(D \cap C) \geq -K_X \cdot C + 2\, D \cdot C, \tag{1}$$

with respect to the intersection pairing on $H^2(X)$, and equality holds if and only if C is totally geodesic. Note that D is empty if and only if M is compact, where Equation (1) reduces to

$$3\,C \cdot C \geq -K_X \cdot C.$$

We also note that every complex hyperbolic manifold admits a finite covering with a compactification as assumed in Theorem 1.6; see the proof of Theorem 1.5 for a precise discussion. Therefore, up to finite covers there is no loss of generality in making the assumptions we make in the following result.

Theorem 1.6 (Arithmeticity and the intersection pairing) *Suppose that X is the smooth toroidal compactification of a complex hyperbolic 2-manifold with (possibly empty) compactification divisor D. If X contains infinitely many complex projective curves C where equality holds in (1) then M is arithmetic. Equivalently, if M is nonarithmetic, then there are only finitely many curves C on X for which equality holds in (1).*

Another application of Theorem 1.3 is to ruling out existence of certain mappings between complex hyperbolic manifolds. For example, Siu asked whether or not there are surjective holomorphic mappings $M_1 \to M_2$ between ball quotients with $2 \leq \dim_{\mathbb{C}}(M_2) < \dim_{\mathbb{C}}(M_1)$ [59]. The only progress thus far is by Koziarz and Mok, who ruled out the case of holomorphic submersions [38]. Using Theorem 1.3 we prove the following in Sect. 9.3.

Theorem 1.7 (Nonexistence of certain maps) *Let M be a finite volume complex hyperbolic manifold of complex dimension $n \geq 2$ containing a family $\{Z_i\}$ of geodesic submanifolds, all of real dimension at least two, that are equidistributed in the sense of Proposition 8.1. Suppose that N is a finite volume complex hyperbolic manifold N with $\dim_{\mathbb{C}}(N) = d \leq n$ and $f : M \to N$ is a continuous map such that $f(Z_i)$ is contained in a proper geodesic submanifold of N for all i and*

$$f_*(\pi_1(M)) \leq \pi_1(N) < PU(d, 1)$$

is Zariski dense. Then $d = n$ and f is homotopic to a finite cover.

An example application of Theorem 1.7 is to the maps between Deligne–Mostow orbifolds, described in detail in work of Toledo [62, Thm. 1]. There are surjective holomorphic maps $f : M \to N$ between complex hyperbolic 2-orbifolds that are not homotopic to a finite cover where M is arithmetic with a family $\{Z_i\}$ as in Theorem 1.7. The maps are additionally surjective on the level of orbifold fundamental groups, hence Theorem 1.7 implies that $f_*(\pi_1(Z_i))$ must be Zariski dense in PU(2, 1) for all but finitely many i. However, note that it is possible that some finite number of Z_i are contracted by the map, hence $f_*(\pi_1(Z_i))$ is trivial for those i.

Our last application is the following. There are lattices $\Gamma < \mathrm{SL}_2(\mathbb{C})$ such that there are $\gamma \in \Gamma$ so that $\mathrm{Tr}(\mathrm{Ad}(\gamma))$ is not an algebraic integer. See [39,

Section 5.2.2] and [15] for many examples. In Sect. 9.4 we will show that part (3) of Theorem 1.5 has the following consequence pointed out to us by Ian Agol.

Theorem 1.8 (Restricting geodesic submanifolds) *There are both closed and noncompact finite volume hyperbolic* 3*-manifolds that are not isometric to an immersed totally geodesic submanifold of a complex hyperbolic* n*-manifold for any* n.

We now discuss the organization of this paper. Section 2 starts by discussing $\mathrm{SU}(n,1)$ and some of its subgroups. In Sect. 3 we introduce the notion of an incompatibility datum, define compatibility, and prove some general results. We also prove compatibility of $\mathrm{SU}(n,1)$ with certain algebraic groups over local fields and study the failure of $\mathrm{PU}(n,1)$ to be compatible with $\mathrm{SU}(n,1)$. In Sect. 4 we discuss existence of equivariant maps and algebraic representations. In Sect. 5 we make the necessary modifications to the arguments in [2] to prove Theorem 1.3(1). We continue in Sect. 6 with the proof of our main technical result, Theorem 1.3(2). In Sect. 7, we prove Theorem 1.5 and give some additional algebraic setup for the proof of Theorem 1.1, which is contained in Sect. 8. Then Sect. 9 contains examples that exhibit the possible behavior of geodesic submanifolds of arithmetic complex hyperbolic manifolds, the proofs Theorems 1.6, 1.7, and 1.8, and some final questions.

2 Preliminaries on SU(n, 1)

Throughout this section and the remainder of the paper we fix the group $G \cong \mathrm{SU}(n,1)$ with $n \geq 2$. We will consider G as being the group of real points of a real algebraic group.

2.1 The group G and its standard subgroups.

In this subsection we fix some notation and give convenient coordinates for certain subgroups of G that will be important in what follows.

We start by giving a convenient matrix representation for G. The group $U(n,1)$ is often described as the automorphism group of the hermitian form

$$h_0(x_1,\ldots,x_{n+1}) = \sum_{i=1}^{n} |x_i|^2 - |x_{n+1}|^2,$$

on $\mathbb{C}^{n+1}$. Under the linear change of variables

$$y_1 = \frac{\sqrt{2}}{2}(x_1 + x_{n+1}),\ y_2 = x_2,\ \ldots,\ y_n = x_n,\ y_{n+1} = \frac{\sqrt{2}}{2}(x_1 - x_{n+1})$$

this form becomes

$$h(y_1, \ldots, y_{n+1}) = y_1\overline{y}_{n+1} + y_{n+1}\overline{y}_1 + \sum_{i=2}^{n} |y_i|^2. \tag{2}$$

Hereafter, we will view G as the subgroup of $\mathrm{SL}_{n+1}(\mathbb{C})$ preserving h.

Remark 2.1 For $n = 1$, the stabilizer of h in $\mathrm{SL}_2(\mathbb{C})$ is the image of the homomorphism:

$$\begin{aligned} \mathrm{SL}_2(\mathbb{R}) &\to \mathrm{SL}_2(\mathbb{C}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \begin{pmatrix} a & ib \\ -ic & d \end{pmatrix} \end{aligned}$$

This explicitly realizes the well-known isomorphism $\mathrm{SU}(1, 1) \cong \mathrm{SL}_2(\mathbb{R})$. We will tacitly assume from here forward that $n \geq 2$.

In what follows, let e_i denote the ith standard basis vector in $\mathbb{C}^{n+1}$ given by $y_i = 1$ and $y_j = 0$ for $j \neq i$. Note that the restriction of h to the complex line spanned by the vector $e_1 - e_{n+1}$ is negative definite. We denote by $K \leq G$ the stabilizer in G of this line. Then K also stabilizes the h-orthogonal complement of this line, namely the complex hyperplane spanned by the vectors $e_2, \ldots, e_n$ and $e_1 + e_{n+1}$, and note that h is positive definite on this hyperplane. One sees easily that

$$K \cong \mathrm{S}(U(n) \times U(1)) \cong U(n),$$

and that it is a maximal compact subgroup of G. In particular, every compact subgroup of G is conjugate to a subgroup of K.

For $1 \leq m \leq n$ let $W_m^c \leq G$ be the subgroup of G fixing each of the $n - m$ standard basis vectors $e_{m+1}, \ldots, e_n \in \mathbb{C}^{n+1}$, and note that $W_m^c \cong \mathrm{SU}(m, 1)$. We also consider the subgroup

$$W_m^c \cap \mathrm{GL}_{n+1}(\mathbb{R}) < \mathrm{GL}_{n+1}(\mathbb{C}),$$

which is isomorphic to $\mathrm{SO}(m, 1)$ and we denote its identity component by W_m^r, thus $W_m^r \cong \mathrm{SO}_0(m, 1)$.

Definition 2.2 The subgroups $W_1^c, \ldots, W_n^c$ and $W_2^r, \ldots, W_n^r$ of G defined above are said to be *the standard almost simple subgroups*, or for short just *the standard subgroups* of G.

Identifying $K\backslash G$ with complex hyperbolic n-space $\mathbb{B}^n$, the standard subgroups of G have a special relationship with its totally geodesic subspaces. It

is shown in §3.1 of Goldman's book [24] (see in particular §3.1.11) that every totally geodesic subspace of complex hyperbolic n-space with real dimension at least two is isometric to either real hyperbolic m-space for some $2 \leq m \leq n$ or complex hyperbolic m-space for some $1 \leq m \leq n$.

Moreover, the group G acts transitively on the collection of totally geodesic subspaces of any given *type*, where the type of a totally geodesic subspace describes whether it is isometric to real or complex hyperbolic space of a given fixed dimension. We note that a real hyperbolic 2-plane has a different type from a complex hyperbolic line. Indeed, under the restriction of the complex hyperbolic metric the former has a constant sectional curvature $-1/4$ while the latter has curvature -1.

It is evident that for every standard subgroup $W \leq G$, $W \cap K$ is a maximal compact subgroup of W and that

$$(W \cap K)\backslash W \xrightarrow{\sim} K\backslash KW \subseteq K\backslash G,$$

is a totally geodesic embedding of a real or complex hyperbolic space of the corresponding type in $K\backslash G$. The following proposition summarizes the above discussion.

Proposition 2.3 ([24, Section 3.1]) *The totally geodesic subspaces of (real) dimension at least two in the symmetric space $K\backslash G$ are exactly the subsets of the form $K\backslash KWg$ for an element $g \in G$ and a standard subgroup $W \leq G$ corresponding to the type of the given totally geodesic subspace.*

Proposition 2.3 will play a prominent role in Sect. 8.1, along with Lemma 2.7 below in which we will record some basic facts about standard subgroups. First we state an important corollary of Proposition 2.3 regarding the classification of almost simple subgroups of G, which follows from the fact that every noncompact, almost simple subgroup stabilizes a totally geodesic subspace constructed by inclusion of maximal compact subgroups.

Proposition 2.4 *The standard subgroups of G are noncompact, connected, almost simple, closed subgroups generated by unipotent elements, and every noncompact, connected, almost simple, closed subgroup of G generated by unipotent elements is conjugate to a unique standard subgroup of G.*

Definition 2.5 Given a noncompact, connected, almost simple, closed subgroup of G generated by unipotent elements, we say that its *type* is the type of the unique standard subgroup of G to which it is conjugate.

Remark 2.6 The analogue of Proposition 2.4 for $\mathrm{SO}_0(n, 1)$ holds as well: every noncompact, connected, almost simple, closed subgroup of $\mathrm{SO}_0(n, 1)$ is conjugate to a unique standard subgroup of the form $\mathrm{SO}_0(m, 1)$ for some

$2 \leq m \leq n$. The proof of this fact is similar to the proof of Proposition 2.4 to be presented below, only it is easier and well-known, so we will omit it.

The proof of Proposition 2.4 will be derived simultaneously with the proof of Lemma 2.7 given below. For a subgroup $S \leq G$, we use the notation S^+ to denote the closed subgroup of G generated by all the one-dimensional unipotent subgroups of S. We note that S^+ is necessarily connected, and it is either a unipotent subgroup or a noncompact, almost simple subgroup of G. This follows from the fact that G has rank 1. Indeed, if S^+ is contained in a parabolic subgroup P then it is contained in P^+, which is the unipotent radical of P, hence S^+ is unipotent. Otherwise, S^+ has a trivial unipotent radical, hence it is semisimple, and since it has no compact factor, it follows that it is almost simple and noncompact.

Lemma 2.7 *Fix a standard subgroup $W \leq G$ and let $N \leq G$ be its normalizer. Then the following results hold.*

(1) *If $S \leq G$ is a connected, almost simple, closed subgroup generated by unipotent elements that preserves $K\backslash KW$ and acts transitively on it, then $S = W$.*
(2) *The stabilizer in G of the totally geodesic subspace $K\backslash KW$ is N. Moreover, $N \subseteq KW$ thus N/W is compact and $N^+ = W$.*
(3) *Assume that $S \leq G$ is a closed intermediate subgroup $W \leq S \leq N$. Then $S^+ = W$ and $K\backslash KW = K\backslash KS = K\backslash KN$.*
(4) *Let $S \leq G$ be a closed subgroup containing W. Then there exists $k \in K$ and a standard subgroup $W_0 \leq G$ such that $kS^+k^{-1} = W_0$ and*

$$K\backslash KS = K\backslash KS^+ = K\backslash KW_0k.$$

Moreover, $K\backslash KS$ is a totally geodesic subspace of $K\backslash G$ whose volume measure coincides, up to normalization, with the push-forward of Haar measure on S to $(S \cap K)\backslash S \simeq K\backslash KS$.

Proof of Proposition 2.4 and Lemma 2.7 We first prove part (1) of the lemma. Let N_0 be the stabilizer in G of $K\backslash KW$ and observe that S is a subgroup of N_0 and $S = S^+ \leq N_0^+$. Since it contains the almost simple subgroup S, N_0^+ is not unipotent, thus it is connected and almost simple. We will show that in fact $S = N_0^+$.

Since N_0 acts by isometries on the symmetric space $Z = W \cap K\backslash W \simeq K\backslash KW$ we have a natural continuous homomorphism from N_0 to the group of isometries, $\mathrm{Isom}(Z)$, endowed with the compact open topology. It goes back to Elie Cartan that the image of S coincides with the identity component of $\mathrm{Isom}(Z)$, as S is semisimple connected group of isometries of the symmetric space Z [26, Thm. V.4.1]. The same holds for N_0^+. Since

both S and N_0^+ are almost simple, the restriction of the above homomorphism to each has a finite kernel. It follows that S has finite index in N_0^+, hence indeed $S = N_0^+$, since N_0^+ is connected.

In the above discussion S was arbitrary, thus applying it in the special case $S = W$ we conclude that $W = N_0^+$. Therefore $S = W$ holds *a priori* and this proves part (1).

Next we prove part (2) of the lemma. By the previous discussion we have that $N_0 \leq N$, since $W = N_0^+$ is normal in N_0. Observe that $K\backslash KW$ is the unique W-invariant totally geodesic subspace of its type. Indeed, given any W-invariant totally geodesic subspace of the same type $Z \subseteq K\backslash G$ and $z \in Z$, the function $d(\cdot, zW)$ measuring distance to the W-orbit of z is constant on $K\backslash KW$. This implies that zW and $K\backslash KW$ have identical boundaries in the visual compactification of $K\backslash G$. Since $K\backslash G$ has negative curvature, this implies that $zW \subset K\backslash KW$, hence $Z = K\backslash KW$.

It also follows that $N \leq N_0$, so $N = N_0$ and N is indeed the stabilizer in G of the totally geodesic subspace $K\backslash KW$. We furthermore see that $N^+ = N_0^+ = W$. For every $n \in N$, it also follows that $K\backslash Kn = K\backslash Kw$ for some $w \in W$, thus $N \subseteq KW$. This proves part (2).

Part (3) of the lemma follows immediately from part (2). Indeed, notice that $W = W^+ \leq S^+ \leq N^+ = W$ implies that $S^+ = W$ and the sequence of inclusions

$$K\backslash KW \subseteq K\backslash KS \subseteq K\backslash KN \subseteq K\backslash K(KW) = K\backslash KW$$

then has equality everywhere.

We now turn to the proof of Proposition 2.4. The fact that the standard subgroups are pairwise nonconjugate, noncompact, connected, almost simple, closed subgroups of G generated by unipotent elements is obvious, so we only need to show that any other such group is conjugate to a standard one. Let $S \leq G$ be a noncompact, connected, almost simple, closed subgroup generated by unipotent elements. By the Karpelevich–Mostow Theorem [31, 47] for some $h \in G$ the S-orbit $K\backslash KhS \subseteq K\backslash G$ is totally geodesic. Thus, by Proposition 2.3, there exist an element $g \in G$ and a standard subgroup $W_0 \leq G$ such that $K\backslash KhS = K\backslash KW_0g$. Rewriting we have $K\backslash Khg^{-1}S^g = K\backslash KW_0$ and we conclude that S^g preserves $K\backslash KW_0$ and acts transitively on it. By part (1) we get that $S^g = W_0$ and this proves Proposition 2.4.

We are now in a position to prove part (4) of the lemma. Note that $W \leq S$ implies that $W = W^+ \leq S^+$. It follows that S^+ is not unipotent, thus it is connected and almost simple. Using Proposition 2.4 there exist an element $g \in G$ and a standard subgroup $W_0 \leq G$ such that $(S^+)^g = W_0$. We conclude that $W^g \leq W_0$. We claim that the group W^g, which is a noncompact, connected, almost simple, closed subgroup of W_0, is conjugate in W_0 to a standard

subgroup of W_0. Indeed, when $W_0 = W_m^c$ for some $1 \leq m \leq n$ this follows from Proposition 2.4, replacing the role of G by $W_0 \cong \mathrm{SU}(m, 1)$, and when $W_0 = W_m^r$ for some $2 \leq m \leq n$ this follows from Remark 2.6 applied to $W_0 \cong \mathrm{SO}_0(m, 1)$. In any case, this standard subgroup of W_0 must be of the same type as W^g, so it must be W itself.

Therefore, there is an $h \in W_0$ such that $W^{hg} = (W^g)^h = W$. It follows that $hg \in N$, thus by part (2), $hg = kw$ for some $k \in K$ and $w \in W$. The sequence of equations

$$kS^+k^{-1} = kwS^+w^{-1}k^{-1} = hgS^+g^{-1}h^{-1} = hW_0h^{-1} = W_0,$$

implies that $(kSk^{-1})^+ = kS^+k^{-1} = W_0$, and therefore

$$W_0 \leq kSk^{-1} \leq N_0,$$

where N_0 is the normalizer of W_0 in G. By part (3) we get

$$K\backslash KS^+k^{-1} = K\backslash K(S^+)^k = K\backslash KW_0 = K\backslash KS^k = K\backslash KSk^{-1},$$

and upon applying k on the right we conclude that indeed

$$K\backslash KS = K\backslash KS^+ = K\backslash KW_0k.$$

This is a totally geodesic subspace of $K\backslash G$ by Proposition 2.3. The final statement follows from the essential uniqueness of an S-invariant measure on $(S \cap K)\backslash S \simeq K\backslash KS$. This proves part (4) and thus completes the proof. □

2.2 Parabolic subgroups of G and some of their subgroups

In this subsection we continue the discussion of special subgroups of G begun in the previous one. See [35, Section VII.7] for definitions and properties of the structure theory for simple Lie groups and their parabolic subgroups. Now we focus mostly on parabolic subgroups of G and some of their subgroups. Recall that $e_1, \ldots, e_{n+1}$ denotes the standard basis of $\mathbb{C}^{n+1}$ and that G is the subgroup of $\mathrm{SL}_{n+1}(\mathbb{C})$ preserving the form h given in Equation (2).

Let $P < G$ be the stabilizer of the isotropic line spanned by e_1. This is a parabolic subgroup of G and all proper parabolic subgroups of G are conjugate to P, since G has real rank 1. Some important subgroups of P are

$$A = \left\{ \left(\begin{array}{c|c|c} \lambda & 0 & 0 \\ \hline 0 & I_{n-1} & 0 \\ \hline 0 & 0 & \lambda^{-1} \end{array} \right) : \lambda \in \mathbb{R}^* \right\},$$

$$M = \left\{ \left(\begin{array}{c|c|c} \theta & 0 & 0 \\ \hline 0 & T & 0 \\ \hline 0 & 0 & \theta \end{array} \right) : \begin{array}{c} \theta \in U(1), \\ T \in U(n-1) \\ \theta^2 = \det(T)^{-1} \end{array} \right\},$$

$$U = \left\{ \left(\begin{array}{c|c|c} 1 & -v^* & ib - \frac{1}{2}\|v\|^2 \\ \hline 0 & I_{n-1} & v \\ \hline 0 & 0 & 1 \end{array} \right) : v \in \mathbb{C}^{n-1},\ b \in \mathbb{R} \right\}.$$

Here v^* denotes the complex conjugate transpose of v. Note that U is isomorphic to the real $(2n-1)$-dimensional Heisenberg group $\mathbb{H}_{2n-1}(\mathbb{R})$ and that $A \cong \mathbb{R}^*$.

One checks that P is generated by M, A, and U. Note that A is a maximal $\mathbb{R}$-split torus of G, U is the unipotent radical of P, and M is a compact reductive group that commutes with A. Thus the Langlands decomposition of P is given by $P = MAU$.

Every element of P can be represented in an obvious way by parameters $(\lambda, \theta, T, v, b)$ using the coordinates introduced above. Such a representation is unique up to the order two intersection group $M \cap A$, that is, $(\lambda, \theta, T, v, b)$ represents the same element as $(-\lambda, -\theta, T, v, b)$. In the remainder of this subsection we will use this representation often. We will also represent elements of M, A, and U by (θ, T), λ, and (v, b) in the obvious way.

Let $C < M$ be the subgroup of scalar matrices in G. These scalars are the $(n+1)^{st}$ roots of unity, thus C is a cyclic group of order $n+1$. Note that C is the center of G, and it is easy to see that it is also the center of P.

The group U is a two step nilpotent group with center

$$Z = \left\{ \left(\begin{array}{c|c|c} 1 & 0 & ib \\ \hline 0 & I_n & 0 \\ \hline 0 & 0 & 1 \end{array} \right) b \in \mathbb{R} \right\}.$$

This is a characteristic subgroup of U, hence it is a normal subgroup of P. Under the identification $Z \cong \mathbb{R}$, the conjugation action of P on Z is given by

the homomorphism:

$$\begin{aligned} P &\to \mathbb{R}_+^* < \mathbb{R}^* \cong \mathrm{GL}_1(\mathbb{R}) \\ (\lambda, \theta, T, v, b) &\mapsto \lambda^2 \end{aligned}$$

Since $\mathbb{R}_+^*$ has no nontrivial compact subgroups, we obtain the following, which we record for future reference.

Lemma 2.8 *Every compact subgroup of P commutes with Z.*

The quotient group U/Z is naturally identified with $\mathbb{C}^{n-1}$ by the map $(v, b) \mapsto v$. Under this identification, the conjugation action of P on U/Z is given by the homomorphism:

$$\begin{aligned} P &\to \mathbb{R}^* \cdot U(n-1) < \mathrm{GL}_{n-1}(\mathbb{C}) \\ (\lambda, \theta, T, v, b) &\mapsto \lambda\theta^{-1}T \end{aligned}$$

Here $\mathbb{R}^* < \mathrm{GL}_{n-1}(\mathbb{C})$ is considered as the group of real scalar matrices. Note that this homomorphism induces isomorphisms $M/C \cong U(n-1)$ and $MA/C \cong \mathbb{R}^* \cdot U(n-1)$. In particular, the conjugation action of MA/C on $(U/Z) \smallsetminus \{0\}$ is transitive and faithful. This is crucial in proving the following lemma, which describes the normal subgroups of P.

Lemma 2.9 *If $N \trianglelefteq P$ is a normal subgroup, either $U \leq N$ or $N \leq CZ$.*

Proof Consider the quotient map $\sigma : P \to P/CZ$. Then $\sigma(N) \cap \sigma(U)$ is normal in P/CZ, hence transitivity of the MA/C conjugation action on $(U/Z) \smallsetminus \{0\}$ implies that this intersection is either trivial or all of $\sigma(U)$. Assume the latter case. We get that $UC \leq NCZ$, thus $UC \leq (UC \cap N)CZ$. Taking commutators and using the centrality of CZ in CU, we get

$$\begin{aligned} Z = [U, U] &= [UC, UC] \\ &\leq [(UC \cap N)CZ, (UC \cap N)CZ] = [CU \cap N, CU \cap N] \end{aligned}$$

and we deduce that $Z \leq N$. It follows that $UC \leq CN$, but U is connected and CN/N is totally disconnected, hence $U \leq N$ and we are done.

Therefore, it remains to consider the case where $\sigma(N) \cap \sigma(U)$ is trivial. It follows that $\sigma(N)$ commutes with $\sigma(U)$, as they are both normal subgroups of P/CZ. Since the action of $\sigma(MA) \cong MA/C$ on $\sigma(U) \cong U/Z$ is faithful, we deduce that N is in the kernel of the natural map $P = MAU \to MA/C$, that is $N \leq UC$. By triviality of $\sigma(N) \cap \sigma(U) = \sigma(N) \cap \sigma(UC)$, we see that N is in the kernel of σ, hence $N \leq CZ$ as desired. □

Let $D < G$ be the subgroup stabilizing the plane spanned by e_1 and e_{n+1}. Thus D also stabilizes the h-orthogonal complement of this plane, the subspace spanned by $e_2, \ldots, e_n$, and we have:

$$D = \left\{ \begin{pmatrix} \theta a & 0 & i\theta b \\ 0 & T & 0 \\ -i\theta c & 0 & \theta d \end{pmatrix} \quad \begin{array}{c} a, b, c, d \in \mathbb{R},\ ad - bc = 1, \\ \theta \in U(1),\ T \in U(n-1), \\ \theta^2 = \det(T)^{-1} \end{array} \right\}.$$

We will be particularly interested in the group $P \cap D$ stabilizing both the line spanned by e_1 and the plane spanned by e_1 and e_{n+1}.

Lemma 2.10 *We have $P \cap D = MAZ$.*

Proof This is easily deduced from the matrix representations of P and D. □

The following lemma is easy, nevertheless we give a detailed proof due to its importance to our considerations.

Lemma 2.11 *Fix $d \geq 1$ and consider the group*

$$\mathbb{R}^* \cdot U(d) \ltimes \mathbb{C}^d \leq \mathrm{GL}_d(\mathbb{C}) \ltimes \mathbb{C}^d.$$

Its group of outer automorphisms, $\mathrm{Out}(\mathbb{R}^* \cdot U(d) \ltimes \mathbb{C}^d)$*, is of order 2 and its nontrivial element corresponds to complex conjugation of* $\mathrm{GL}_d(\mathbb{C}) \ltimes \mathbb{C}^d$.

Proof For brevity, we set $S = \mathbb{R}^* \cdot U(d)$ and $V = \mathbb{C}^d$. We consider an automorphism $\tau : S \ltimes V \to S \ltimes V$ and will show that, up to an inner automorphism, τ is either trivial or complex conjugation.

Identifying V with $\mathbb{R}^{2d}$ in the standard way, we identify its group of continuous automorphisms with the the real general linear group $\mathrm{GL}_{2d}(\mathbb{R})$. For an operator $t \in \mathrm{GL}_{2d}(\mathbb{R})$ and $v \in V$ we denote the corresponding action by $t \cdot v$. We identify the conjugation action of S on V as the linear action of $S < \mathrm{GL}_{2d}(\mathbb{R})$ and for $s \in S$ and $v \in V$, define $s \cdot v = svs^{-1} = v^s$.

We have that $\tau(S) < S \ltimes V$ is a Levi subgroup. See [10, Section 6.8] for details about Levi decompositions (at the level of Lie algebras). As all Levi subgroups are conjugate, we assume as we may that $\tau(S) = S$. Let α denote the automorphism of S induced by τ. Since V is a characteristic subgroup of $S \ltimes V$, we have $\tau(V) = V$. Thus τ induces a continuous automorphism of V corresponding to a fixed element $t \in \mathrm{GL}_{2d}(\mathbb{R})$. For $s \in S$ and $v \in V$ we have

$$ts \cdot v = t \cdot v^s = \tau(v^s) = \tau(v)^{\tau(s)} = (t \cdot v)^{\alpha(s)} = \alpha(s)t \cdot v.$$

Thus $\alpha(s) = s^t$. We conclude that t is contained in N, where N is the normalizer of S in $\mathrm{GL}_{2d}(\mathbb{R})$, and α is the automorphism of S obtained by conjugating by t.

We claim that $S < N$ is of index two, where the nontrivial coset is generated by complex conjugation. We fix $g \in N$ and argue that, up to complex conjugation, g is in S. Note that N commutes with the group of real scalars $\mathbb{R}^* < S$ and the group $U(1) < S$ consisting of modulus 1 complex scalars is characteristic. Indeed, $U(d) < S$ is characteristic, as it is the unique maximal compact subgroup of S, and $U(1)$ is its center. Thus g normalizes $U(1)$.

Since the unique nontrivial automorphism of $U(1)$ is complex conjugation, we can assume that g acts trivially on $U(1)$. We conclude that g commutes with the group of complex scalars $\mathbb{C}^*$, hence g is $\mathbb{C}$-linear, i.e., $g \in \mathrm{GL}_d(\mathbb{C})$. If $d = 1$ then $g \in \mathrm{GL}_d(\mathbb{C}) = S$, and we are done. We thus assume $d \geq 2$. Since the derived subgroup $\mathrm{SU}(d) < U(d)$ is also characteristic, we get that g normalizes it as well. However, $\mathrm{SU}(d)$ is a maximal subgroup of $\mathrm{SL}_d(\mathbb{C})$, hence it equals its own normalizer in $\mathrm{SL}_d(\mathbb{C})$. It follows that the normalizer of $\mathrm{SU}(d)$ in $\mathrm{GL}_d(\mathbb{C})$ is $\mathbb{C}^* \cdot \mathrm{SU}(d) = S$, thus $g \in S$. This proves the claim.

We conclude that indeed, up to an inner automorphism, τ is either trivial or complex conjugation. □

Proposition 2.12 *Consider $\overline{P} = P/CZ \cong \mathbb{R}^* \cdot U(n-1) \ltimes \mathbb{C}^{n-1}$ and let $\tau : P \to \overline{P}$ be a continuous homomorphism with one-dimensional kernel. Then there exists an inner automorphism $i : P \to P$ such that*

$$\textit{either} \quad \tau = \theta \circ i \quad \textit{or} \quad \tau = c \circ \theta \circ i,$$

where $\theta : P \to \overline{P}$ is the obvious quotient map and $c : \overline{P} \to \overline{P}$ corresponds to complex conjugation on $\mathbb{R}^* \cdot U(n-1) \ltimes \mathbb{C}^{n-1}$ as discussed in Lemma 2.11.

In particular, τ is surjective, $\ker(\tau) = \ker(\theta) = CZ$ and, up to precomposing τ by an inner automorphism of P, $\tau(P \cap D) = \theta(P \cap D)$.

Proof We first note that Lemma 2.9 implies that $Z \leq \ker(\tau) \leq CZ$, since $\dim(U) = 2n - 1 \geq 3$ and $Z \cong \mathbb{R}$ is the identity component of CZ. By dimension considerations, $\tau(P)$ is Zariski dense in $\overline{P}$, as $\overline{P}$ is Zariski connected [29, Section 23.1, Cor. B]. Thus $C \leq \ker(\tau)$, as $\overline{P}$ has trivial center, hence $\ker(\tau) = CZ$ and τ induces a continuous injection $\overline{\tau} : \overline{P} \to \overline{P}$ such that $\tau = \overline{\tau} \circ \theta$.

A continuous injective endomorphism of a Lie group with finitely many connected components is surjective, so $\overline{\tau}$ is necessarily an automorphism. By Lemma 2.11 we get that either $\bar{\tau} = c \circ \mathrm{inn}(\bar{g})$ or $\bar{\tau} = \mathrm{inn}(\bar{g})$ for some $\bar{g} \in \overline{P}$, where c denotes complex conjugation. Fixing $g \in P$ such that $\bar{g} = \theta(g)$ we get that either $\tau = c \circ \theta \circ \mathrm{inn}(g)$ or $\tau = \theta \circ \mathrm{inn}(g)$, proving the first part of the proposition.

It remains to show that $\tau(P \cap D) = \theta(P \cap D)$ up to an inner automorphism. This follows from Lemma 2.10, since $\theta(P \cap D) = \theta(MA)$ corresponds to the subgroup $\mathbb{C}^* \cdot \mathrm{SU}(n-1) \subset \mathbb{C}^* \cdot \mathrm{SU}(n-1) \ltimes \mathbb{C}^{n-1}$, which is invariant under complex conjugation. □

3 Compatibility and measuring incompatibility

While much of this section applies more broadly, in what follows G will always denote $\mathrm{SU}(n, 1)$. We fix its matrix presentation as discussed in Sect. 2 and freely use the notation introduced there. In this section, we recall the definition of compatibility from [2, Section 3.4] and study compatibility of certain groups over local fields with G. In this paper, we will also need to measure the extent to which compatibility fails. This leads us to begin with the following sequence of definitions.

Definition 3.1 Let P be a minimal parabolic subgroup of G and U its unipotent radical. An *incompatibility datum* for G is a tuple $(k, \mathbf{H}, \mathbf{J}, U', \tau)$, where k is a local field, $\mathbf{H}$ is a k-algebraic group, $\mathbf{J} < \mathbf{H}$ is a nontrivial k-subgroup, $U' < U$ is a nontrivial proper subgroup, and $\tau : P \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$ is a continuous homomorphism such that the Zariski closure of $\tau(U')$ is not equal to the Zariski closure of $\tau(U)$. Here $N_{\mathbf{H}}(\mathbf{J})$ denotes the normalizer of $\mathbf{J}$ in $\mathbf{H}$. We then have:

- When k, $\mathbf{H}$, and $\mathbf{J}$ are fixed, τ is an *incompatible homomorphism* for $\mathbf{J}$ if there exists $U' < U$ as above so the corresponding tuple forms an incompatibility datum for G. If no such U' exists, then τ is called a *compatible homomorphism*.
- When k and $\mathbf{H}$ are fixed, $\mathbf{J}$ is an *incompatible subgroup* if there exists an incompatible homomorphism for $\mathbf{J}$. Otherwise $\mathbf{J}$ is a *compatible subgroup*.
- When k is fixed, $\mathbf{H}$ is an *incompatible k-group* for G if there is a nontrivial k-subgroup $\mathbf{J} < \mathbf{H}$ that is an incompatible subgroup for the pair $(k, \mathbf{H})$. Otherwise we call $\mathbf{H}$ a *compatible k-group*.
- A local field k is *incompatible* with G if there exists an incompatible k-group. Otherwise k is called *compatible*.

We will see an example of an incompatibility datum in Proposition 3.4. First we prove some general results. To stay as notationally consistent with [4] as possible, if $\mathbf{A}$, $\mathbf{B}$ are k-algebraic groups we use the notation $\mathbf{A}/\mathbf{B}(k)$ to mean $(\mathbf{A}/\mathbf{B})(k)$, that is, the k-points of $\mathbf{A}/\mathbf{B}$.

Proposition 3.2 *Every nonarchimedean local field is compatible with G.*

Proof If k is a nonarchimedean local field, then for any k-group $\mathbf{H}$ and k-subgroup $\mathbf{J} < \mathbf{H}$, $N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$ is totally disconnected. Therefore every con-

tinuous homomorphism $\tau : P \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$ must be trivial on the connected subgroup $U < P$. Compatibility clearly follows. □

In view of Proposition 3.2 we will be concerned in the rest of this subsection with archimedean local fields, i.e., $k = \mathbb{R}$ or $k = \mathbb{C}$. We then have the following general lemma.

Lemma 3.3 *Let $(k, \mathbf{H}, \mathbf{J}, U', \tau)$ be an incompatibility datum for G where k is $\mathbb{R}$ or $\mathbb{C}$. Suppose that $\mathbf{S}$ is the Zariski closure in $\mathbf{H}$ of the preimage of $\tau(P)$ under the map $N_{\mathbf{H}}(\mathbf{J})(k) \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$. Then $\mathbf{S}$ is not reductive. In particular, if $\mathbf{H}$ is reductive then $\mathbf{S}$ is contained in a proper parabolic subgroup of $\mathbf{H}$.*

Proof We first observe that if $\mathbf{H}$ is reductive and $\mathbf{S}$ is not reductive then $\mathbf{S}$ is contained in a proper parabolic subgroup of $\mathbf{H}$ by [9, Section 3]. It then suffices to show that $\mathbf{S}$ is not reductive.

By extending scalars if necessary, we can assume that $k = \mathbb{C}$. We identify algebraic groups with their $\mathbb{C}$-points, writing H for $\mathbf{H}(\mathbb{C})$ with similar notation for the other groups. We assume that S is reductive and will prove that its Lie algebra has a noncentral nilpotent ideal, which is a contradiction.

Incompatibility of τ implies that $\tau(U)$ is nontrivial, since $\tau(U')$ must be a proper subgroup of $\tau(U)$. Note that $[P, U] = U$, thus $\tau(U)$ is a noncentral normal nilpotent subgroup of $\tau(P)$. We now consider G, H, and their subgroups as Lie groups, τ as a morphism of Lie groups, and denote their Lie algebras by the corresponding Gothic letters. Since $\mathfrak{s}$ is reductive, the ideal $\mathfrak{j} \trianglelefteq \mathfrak{s}$ has a direct complement $\mathfrak{i} \trianglelefteq \mathfrak{s}$, i.e., $\mathfrak{s} \cong \mathfrak{i} \oplus \mathfrak{j}$. The image of the composition

$$\mathfrak{u} \to \mathfrak{s}/\mathfrak{j} \to \mathfrak{i} \to \mathfrak{s}$$

defines a nontrivial noncentral nilpotent ideal in $\mathfrak{s}$, giving the desired contradiction and thus proving the lemma. □

To prove Theorem 1.5(1) we will be concerned exclusively with the case $k = \mathbb{R}$ and $\mathbf{H}(\mathbb{R}) = \mathrm{PU}(r, s)$ with $r + s = n + 1$. For the proof of Theorem 1.7, we will care about the case $r + s < n + 1$. In particular, we will need the following technical result.

Proposition 3.4 *Let the tuple $(k, \mathbf{H}, \mathbf{J}, U', \tau)$ be an incompatibility datum for G with $k = \mathbb{R}$ and $\mathbf{H} = \mathrm{PU}(r, s)$ where $r + s \leq n + 1$, $r \geq s \geq 1$. Then*

(1) *$r = n$ and $s = 1$, i.e., $\mathbf{H} = \mathrm{PU}(n, 1)$, and*
(2) *$\mathbf{J}$ is the center of the unipotent radical of a proper parabolic subgroup of $\mathbf{H}$ and τ has one-dimensional kernel.*

In particular, $\mathbf{H} = \mathrm{PU}(r, s)$ is compatible with G for either $r + s < n + 1$ or $s > 1$.

Proof We work directly with real points of real algebraic groups and abandon the bold face notation. We again let S denote the Zariski closure of the preimage of $\tau(P)$ under the map $N_H(J) \to N_H(J)/J$. Lemma 3.3 implies that $S \leq N_H(J)$ is contained in some maximal proper parabolic subgroup $Q < H$. Set $N = N_Q(J)$ and note that $S \leq N \leq Q$. Restricting the codomain, we consider τ as a map from P to $N/J \leq N_H(J)/J$.

The group $U(r, s)$ acts on $\mathbb{C}^{r+s}$ preserving the hermitian form

$$h_{r,s}(x_1, \ldots, x_{n+1}) = \sum_{i=1}^{r} |x_i|^2 - \sum_{i=1}^{s} |x_{r+i}|^2,$$

and Q is the projectivization of the stabilizer in $U(r, s)$ of a totally isotropic complex subspace $V \subset \mathbb{C}^{r+s}$. Let $d = \dim_{\mathbb{C}}(V)$ and note that $d \leq s$, as s is the dimension of a maximal $h_{r,s}$-isotropic subspace $\mathbb{C}^{r+s}$. A parabolic subgroup of H is conjugate to Q if it is the stabilizer of a d-dimensional totally isotropic subspace. A parabolic subgroup is opposite to Q if it is the stabilizer of a maximal totally isotropic subspace V' disjoint from V such that the restriction of $h_{r,s}$ to $V \oplus V'$ is nondegenerate.

A Levi factor of Q is obtained by intersecting Q with an opposite parabolic subgroup Q'. If V' is the totally isotropic subspace associated with Q', then $L = Q \cap Q'$ stabilizes the subspaces V, V' and $(V \oplus V')^{\perp}$, and these subspaces form a direct sum decomposition of $\mathbb{C}^{r+s}$. Taking $V \oplus V'$ and $(V \oplus V')^{\perp}$ with the induced forms, we identify $\mathrm{GL}(V)$ with the stabilizer of V in $U(V \oplus V')$ under the action $T(v, v') = (Tv, (\overline{T}^{-1})^* v')$. This leads to the identification

$$L \cong \mathrm{P}(\mathrm{GL}(V) \times U((V \oplus V')^{\perp})) \cong \mathrm{P}(\mathrm{GL}_d(\mathbb{C}) \times U(r-d, s-d)),$$

and we see that $K \cong \mathrm{P}(U(d) \times U(r-d) \times U(s-d))$ is a maximal compact subgroup of L.

We use the notation introduced in Sect. 2.2 and recall that $M < P$ is locally isomorphic to $U(n-1)$. The map $S \to S/J$ is real algebraic, thus the compact subgroup $\tau(M) < S/J$ has a compact lift $\widetilde{M} < S$ such that the map $S \to S/J$ restricts to a local isomorphism $\widetilde{M} \to \tau(M)$. Incompatibility of τ means that U cannot be in the kernel of τ, and it follows from Lemma 2.9 that $\ker(\tau) \leq CZ$. In particular, τ is almost injective on M.

We therefore get that $\widetilde{M}$ is locally isomorphic to $\tau(M)$, which is locally isomorphic to M, therefore $\widetilde{M}$ is locally isomorphic to $U(n-1)$. Let L be a Levi factor of Q that contains $\widetilde{M}$ and $K_0 < L$ be a maximal compact subgroup containing $\widetilde{M}$. We conclude that d, s, r and n are parameters satisfying $1 \leq d \leq s \leq r \leq n$ and $r + s \leq n + 1$ such that the group $\mathrm{P}(U(d) \times U(r-d) \times U(s-d))$ contains a subgroup locally isomorphic to $U(n-1)$.

We claim that $r = n$, which will imply $s = 1$ and will prove (1). Assume for a contradiction that $r \leq n-1$. Note that the possibility $d = s = r$ is excluded. Indeed, in this case $\mathrm{P}(U(d) \times U(r-d) \times U(s-d)) = \mathrm{PU}(d)$ does not contain a subgroup locally isomorphic to $U(n-1)$ by dimension considerations, as $d \leq r \leq n-1$. If $n = 2$ we necessarily have $1 = d = s = r$, which gives a contradiction. We thus have $n \geq 3$. Thus the commutator subgroup of the subgroup of $\mathrm{P}(U(d) \times U(r-d) \times U(s-d))$, which is locally isomorphic to $U(n-1)$, is almost simple, hence it projects almost injectively to one of the groups $\mathrm{PU}(d)$, $\mathrm{PU}(r-d)$, or $\mathrm{PU}(s-d)$. Since this commutator subgroup is locally isomorphic to $\mathrm{SU}(n-1)$, it follows by dimension considerations that $n-1$ is less than or equal to one of the three numbers d, $r-d$, or $s-d$. Since $s-d \leq r-d \leq n-2$ we must have that $n-1 \leq d$ and we conclude that $d = s = r = n-1$. This gives a contradiction. Hence we must have $r = n$, completing the proof of (1).

We now have that $H = \mathrm{PU}(n, 1)$. In particular, we have a natural surjection

$$G = \mathrm{SU}(n, 1) \to G/C \cong \mathrm{PU}(n, 1) = H,$$

with finite kernel. Let P' be the preimage of Q under this surjection. Then $P' < G$ is a proper parabolic subgroup, and hence is conjugate to P. Precomposing the above map with the corresponding conjugation we get a surjection $\theta : G \to H$ such that $P = \theta^{-1}(Q)$ and $C = \theta^{-1}(e)$.

Therefore P is locally isomorphic to Q and $\dim(Q) = \dim(P)$. Note that τ factors through an injection $P/\ker(\tau) \to N/J = N_Q(J)/J$. Recalling that $\ker(\tau) \leq CZ$, we have that $\dim(\ker(\tau)) \leq 1$. We then have the following chain of inequalities

$$\begin{aligned}
\dim(P) - 1 &\leq \dim(P) - \dim(\ker(\tau)) = \dim(P/\ker(\tau)) \\
&\leq \dim(N) - \dim(J) = \dim(N/J) \\
&\leq \dim(Q) - \dim(J) \\
&= \dim(P) - \dim(J)
\end{aligned} \tag{3}$$

that all follow easily from the above observations.

We claim that $\dim(N) = \dim(Q)$. If not, then $\dim(N) < \dim(Q)$ and we see from (3) that $\dim(P) - 1 < \dim(P) - \dim(J)$ so $\dim(J) = 0$. In other words, $J < Q$ is finite, and so $\theta^{-1}(J) < P$ is also finite, hence compact. Then $\theta^{-1}(J)$ commutes with $Z < P$ by Lemma 2.8. Setting $\overline{Z} = \theta(Z)$, J commutes with $\overline{Z} < Q$, and in particular $\overline{Z} \leq N$.

Since $Z \lhd P$ is normal, $\overline{Z} \lhd Q$ is normal and hence $J\overline{Z} \lhd N$ is normal. Consider the natural map $\pi : N/J \to N/J\overline{Z}$. Since Z is one-dimensional, $\overline{Z}$ is one-dimensional, and so $\ker(\pi) = J\overline{Z}/J$ is one-dimensional. Recall that $\ker(\tau) \leq CZ$, so it is at most one-dimensional. We conclude that $\ker(\pi \circ \tau)$

is at most two-dimensional. However, $\dim(U) \geq 3$, thus $\ker(\pi \circ \tau) \leq CZ$ by Lemma 2.9, and in particular $\ker(\pi \circ \tau)$ is at most one-dimensional.

Considering the injective map $P/\ker(\pi \circ \tau) \to N/J\overline{Z}$, we obtain the contradictory chain of inequalities:

$$\begin{aligned}\dim(P) - 1 &\leq \dim(P) - \dim(\ker(\pi \circ \tau)) = \dim(P/\ker(\pi \circ \tau)) \\ &\leq \dim(N) - \dim(J\overline{Z}) = \dim(N/J\overline{Z}) \\ &= \dim(N) - 1 \\ &< \dim(Q) - 1 \\ &= \dim(P) - 1\end{aligned}$$

Therefore $\dim(N) = \dim(Q)$ as claimed.

We now have that $N \leq Q$ and $\dim(N) = \dim(Q)$. Since Q is Zariski connected (again, see [29, Section 23.1]), we conclude that $N = Q$ and $J \lhd Q$ is normal. From Equation (3) we get that $\dim(J) \leq 1$. If J was finite, it would be central in Q, since $J \lhd Q$ is normal, Q is Zariski connected, and any discrete normal subgroup of a connected group is central. However Q has trivial center and J is nontrivial by the incompatibility assumption, thus we have that $\dim(J) = 1$.

Then $\theta^{-1}(J)$ is normal in P and one-dimensional. Lemma 2.9 implies that $\theta^{-1}(J) \leq CZ$ and we conclude that $J \leq \overline{Z} = \theta(CZ)$. Since $\dim(J) = 1$ and $\overline{Z} \cong Z \cong \mathbb{R}$ is connected, we must have $J = \overline{Z}$. Noting that $Q < H$ is a proper parabolic subgroup and $\overline{Z} < Q$ is the center of its unipotent radical, we have shown that J is indeed the center of the unipotent radical of a proper parabolic subgroup of H. Finally, using the fact that $\dim(J) = 1$ we get from Equation (3) that $\dim(\ker(\tau)) = 1$. This completes the proof of (2). □

4 Preliminaries for the proof of Theorem 1.3

4.1 Existence of equivariant maps

Assume that k is a local field and $\mathbf{H}$ is a connected adjoint k-algebraic group. Let $\Gamma < G$ be a lattice and consider a homomorphism $\rho : \Gamma \to \mathbf{H}(k)$ with unbounded, Zariski dense image. Assume that $W < G$ is a closed noncompact subgroup for which there exists a faithful irreducible representation of $\mathbf{H}(k)$ on a k-vector space V of dimension at least two and a W-invariant ergodic measure ν on the bundle $(G \times \mathbb{P}(V))/\Gamma$ that projects to Haar measure on G/Γ, where the action of W is induced by the left action of W on G. The proof of the following result is exactly the same as the proof of [2, Prop. 4.1].

Proposition 4.1 *Under the assumptions above, there is a proper noncompact k-algebraic subgroup $\mathbf{L} < \mathbf{H}$ and a measurable W-invariant and*

Γ-equivariant map $\phi : G \to \mathbf{H}/\mathbf{L}(k)$. We can also view ϕ as a measurable Γ-map from $W\backslash G$ to $\mathbf{H}/\mathbf{L}(k)$.

4.2 Algebraic representations

Here we recall the work of Bader and Furman [4] that will be used in the proof of Theorem 1.3. We also refer to [2, Section 4.2] for further discussion. We again assume that k is a local field and $\mathbf{H}$ is a connected adjoint k-algebraic group.

We fix a lattice $\Gamma < G$ and a Zariski dense representation $\rho : \Gamma \to \mathbf{H}(k)$. For a closed subgroup $T < G$, a *T-algebraic representation of G* consists of:

- a k-algebraic group $\mathbf{I}$,
- a k-($\mathbf{H} \times \mathbf{I}$)-algebraic variety $\mathbf{V}$ that is a left $\mathbf{H}$-space and a right $\mathbf{I}$-space on which the $\mathbf{I}$-action is faithful,
- a Zariski dense homomorphism $\tau : T \to \mathbf{I}(k)$,
- an *algebraic representation* of G on $\mathbf{V}$, by which we mean an almost-everywhere defined measurable map $\phi : G \to \mathbf{V}(k)$ such that

$$\phi(tg\gamma^{-1}) = \rho(\gamma)\phi(g)\tau(t)^{-1}$$

  for every $\gamma \in \Gamma$, every $t \in T$, and almost every $g \in G$.

The data for a T-algebraic representation is denoted by $\mathbf{I}_{\mathbf{V}}$, $\tau_{\mathbf{V}}$, and $\phi_{\mathbf{V}}$.

A T-algebraic representation is *coset T-algebraic representation* when $\mathbf{V}$ is the coset space $\mathbf{H}/\mathbf{J}$ for some k-algebraic subgroup $\mathbf{J} < \mathbf{H}$ and $\mathbf{I}$ is a k-subgroup of $N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}$, where $N_{\mathbf{H}}(\mathbf{J})$ is the normalizer of $\mathbf{J}$ in $\mathbf{H}$ and $\mathbf{H}/\mathbf{J}$ is endowed with the standard right action of $N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}$. The collection of T-algebraic representations of G forms a category. The Howe–Moore theorem implies that if T is noncompact then the T-action on G/Γ is mixing, hence weakly mixing. In this case, by [4, Thm. 4.3], the above category has an initial object which is a coset T-algebraic representation.

Definition 4.2 Suppose that G, Γ, and T are as above. An initial object in the category of T-algebraic representations of G is called the *gate* or, when wishing to emphasize its dependence on T, the *T-gate*.

Remark 4.3 As remarked in [3, Def. 4.4] there is not a unique T-gate in the category of T-algebraic representations. Indeed any conjugate of an initial coset T-algebraic representation will yield another T-gate (see [4, Rem. 4.5]). The gate is easily seen to be unique up to unique isomorphism and for this reason, and to remain consistent with the discussion in [3], we continue to refer to such a choice as *the T-gate*.

We recall one last definition from [2, Section 4.2]. Given closed noncompact subgroups $S, T < G$, consider their gates $\phi_S, \mathbf{V}_S$ and $\phi_T, \mathbf{V}_T$. We say that these gates *have the same map* if $\mathbf{V}_S = \mathbf{V}_T$ and ϕ_T, ϕ_S agree almost everywhere. We then have the following.

Lemma 4.4 (cf. [2, Lem. 4.4]) *Assume that $T < G$ is a closed noncompact subgroup. Assume $T = T_0 \lhd T_1 \lhd \cdots \lhd T_n = T'$ is a sequence of subgroups of G such that T_{i-1} is normal in T_i for each $i = 1, \ldots, n$. Then the gates for T and T' can be chosen to have the same map.*

Proof By induction we can assume that $n = 1$, that is $T < T' \leq N_G(T)$. Then [2, Lem. 4.4] implies we can assume that the gates for T and $N_G(T)$ have the same map $\phi : G \to \mathbf{H}/\mathbf{J}(k)$. Since ϕ is an $N_G(T)$-algebraic representation, it is also a T'-algebraic representation. Let $\phi' : G \to \mathbf{H}/\mathbf{J}'(k)$ be the gate in the category of T'-algebraic representations. It follows that we can find an $\mathbf{H}$-equivariant k-map $\mathbf{H}/\mathbf{J}' \to \mathbf{H}/\mathbf{J}$. Viewing ϕ' as a T-algebraic representation, we can also find an $\mathbf{H}$-equivariant k-map $\mathbf{H}/\mathbf{J} \to \mathbf{H}/\mathbf{J}'$. By [4, Cor. 4.7] we see that the gates for T and T' can indeed be chosen to have the same map. □

5 The proof of Theorem 1.3(1)

In this section we state and prove Proposition 5.1 and then use it to prove Theorem 1.3(1). Proposition 5.1 will also be used in the proof of Theorem 1.3(2) in the next section. Throughout this section we rely on the notation, definitions, and results of the previous sections, particularly the notion of incompatibility given in Definition 3.1 and the notion of a gate from Definition 4.2.

Proposition 5.1 *Let G be $\mathrm{SU}(n, 1)$ for some $n \geq 2$, $\Gamma < G$ be a lattice, and $W < G$ be a noncompact connected almost simple subgroup. Suppose that k is a local field, $\mathbf{H}$ is a connected adjoint k-algebraic group, and that $\rho : \Gamma \to \mathbf{H}(k)$ is a homomorphism with unbounded, Zariski dense image. Assume moreover that there is a faithful irreducible representation of $\mathbf{H}(k)$ on a k-vector space V of dimension at least two and a W-invariant, ergodic measure ν on $(G \times \mathbb{P}(V))/\Gamma$ that projects to Haar measure on G/Γ.*

Suppose that $U' < W$ is a nontrivial unipotent subgroup, consider the category of U'-algebraic representations of G, and let $\Psi : G \to \mathbf{H}/\mathbf{J}(k)$ be the corresponding gate, where $\mathbf{J} < \mathbf{H}$ is a k-algebraic subgroup and Ψ is a measurable map that is $(U' \times \Gamma)$-equivariant with respect to a homomorphism $\tau : U' \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$. Then τ extends to a continuous homomorphism $\overline{\tau} : P \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$ with $\overline{\tau}|_{U'} = \tau$ such that Ψ is $(P \times \Gamma)$-equivariant with respect to $\overline{\tau}$. Furthermore:

(1) *If $\mathbf{J}$ is trivial, then ρ extends to a continuous homomorphism from G to $\mathbf{H}(k)$.*

(2) *If* $\mathbf{J}$ *is nontrivial, then it is an incompatible subgroup of* $\mathbf{H}$ *and* $\overline{\tau}$ *is an incompatible homomorphism.*

We note that Proposition 5.1 is essentially proved in [2, Section 4.3]. The main difference is that here we emphasize (in)compatibility of $\mathbf{J}$ and $\overline{\tau}$, rather than compatibility of the group $\mathbf{H}$. Apart from that and minor differences due to the fact that $G = \mathrm{SU}(n, 1)$ here rather than $\mathrm{SO}(n, 1)$, the proofs are the same. Nevertheless, we include the full proof for the reader's convenience.

Proof of Proposition 5.1 By [9, Section 3] we can find a proper parabolic subgroup $P < G$ containing U'. Let M, A, U, and Z be as in Sect. 2.2. Since G has real rank 1, P is minimal parabolic and all its unipotent elements are contained in U, hence $U' < U$.

Now consider the group $U'Z < U$, where Z is the center of U. We then have $U' \lhd U'Z \lhd U \lhd P$, and applying Lemma 4.4 we see that the gates for U' and P can be chosen to have the same map. Therefore, τ extends to a continuous homomorphism $\overline{\tau} : P \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$ with respect to which Ψ is $(P \times \Gamma)$-equivariant.

Assume $\mathbf{J}$ is trivial. As $A < P$, Ψ is A-equivariant through $\overline{\tau}|_A$. Triviality of $\mathbf{J}$ implies that this must be the gate in the category of A-algebraic representations. Then Lemma 4.4 implies that $\overline{\tau}|_A$ extends to a homomorphism $N_G(A) \to \mathbf{H}(k)$ for which Ψ is $N_G(A)$-equivariant, where $N_G(A)$ is the normalizer of A in G.

Since $N_G(A)$ contains a Weyl element for A (i.e., a generator for the Weyl group) we have that the group $\langle P, N_G(A)\rangle$ generated by P and $N_G(A)$ is G. Since Ψ is equivariant for both P and $N_G(A)$, using [4, Prop. 5.1] and following the end of the proof of [4, Thm. 1.3], we conclude that $\rho : \Gamma \to \mathbf{H}(k)$ extends to a continuous homomorphism $\widehat{\rho} : G \to \mathbf{H}(k)$. This proves the proposition when $\mathbf{J}$ is trivial.

We now assume that $\mathbf{J}$ is nontrivial but $\overline{\tau}$ is compatible and derive a contradiction. Compatibility of $\overline{\tau}$ implies that the Zariski closures of $\overline{\tau}(U')$ and $\overline{\tau}(U)$ coincide. Proposition 4.1 implies that there exists a proper noncompact k-algebraic subgroup $\mathbf{L} \lneq \mathbf{H}$ and a measurable W-invariant, Γ-equivariant map $\phi : G \to \mathbf{H}/\mathbf{L}(k)$.

The W-invariant map ϕ is also U'-invariant, as $U' < W$. Thus it factors through $\Psi : G \to \mathbf{H}/\mathbf{J}(k)$ and

$$G \to \mathbf{H}/\mathbf{J}(k) \to (\mathbf{H}/\mathbf{J})/\overline{\overline{\tau}(U')}(k) = (\mathbf{H}/\mathbf{J})/\overline{\overline{\tau}(U)}(k)$$

by U'-invariance, where $\overline{\overline{\tau}(U')} = \overline{\overline{\tau}(U)}$ are the corresponding Zariski closures. Since Ψ is U-equivariant, the above composition is U-invariant, and it follows that ϕ is also U-invariant. Since ϕ is also W-invariant and $\langle U, W\rangle = G$, we obtain that $\phi : G \to \mathbf{H}/\mathbf{L}(k)$ is an essentially constant Γ-equivariant map,

hence $\rho(\Gamma)$ has a fixed point on $\mathbf{H}/\mathbf{L}(k)$. This is impossible since $\rho(\Gamma)$ is Zariski dense in $\mathbf{H}$ and $\mathbf{L}$ is a proper algebraic subgroup of the connected adjoint group $\mathbf{H}$. This is a contradiction, which completes the proof. □

We are now prepared to prove part (1) of Theorem 1.3.

Proof of Theorem 1.3(1) Suppose $U' < W$ is a nontrivial unipotent subgroup and consider the category of U'-algebraic representations with gate $\Psi : G \to \mathbf{H}/\mathbf{J}(k)$, where $\mathbf{J} < \mathbf{H}$ is a k-algebraic subgroup and Ψ is a measurable map that is $(U' \times \Gamma)$-equivariant with respect to a homomorphism $\tau : U' \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$.

We claim that $\mathbf{J}$ is trivial. Indeed, if it was nontrivial then it would be incompatible by Proposition 5.1, contradicting the compatibility of $\mathbf{H}$ with $\mathrm{SU}(n, 1)$. Applying Proposition 5.1 again, we see that ρ extends to a continuous homomorphism from G to $\mathbf{H}(k)$. This completes the proof. □

6 The proof of Theorem 1.3(2)

We begin by stating Proposition 6.1 and then use it and Proposition 5.1 to prove Theorem 1.3(2). The rest of this section is then devoted to the proof of Proposition 6.1.

Proposition 6.1 *Suppose that G is $\mathrm{SU}(n, 1)$ for $n \geq 2$ and $\Gamma < G$ is a lattice. Let $H = \mathrm{PU}(n, 1)$, $\overline{Z} < H$ be the center of the unipotent radical of a proper parabolic subgroup $Q < H$, and $\rho : \Gamma \to H$ be a homomorphism with unbounded, Zariski dense image. Assume that there exists a continuous homomorphism $\tau : P \to Q/\overline{Z}$ with one-dimensional kernel and a measurable map $\Phi : G \to H/\overline{Z}$ that is $(P \times \Gamma)$-equivariant with respect to the left Γ-action and the right P-action on $H/\overline{Z}$ via τ. Then ρ extends to a continuous homomorphism from G to H.*

Using this, we prove part (2) of Theorem 1.3.

Proof of Theorem 1.3(2) given Proposition 6.1 Here $k = \mathbb{R}$ and $\mathbf{H}$ is the k-algebraic group corresponding to H. Let $U' < W$ be a nontrivial unipotent subgroup and consider the category of U'-algebraic representations of G and the corresponding gate $\Psi : G \to \mathbf{H}/\mathbf{J}(k)$, where $\mathbf{J} < \mathbf{H}$ is a k-algebraic subgroup and Ψ is a measurable map that is $(U' \times \Gamma)$-equivariant with respect to a homomorphism $\tau : U' \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$. By Proposition 5.1, τ extends to a continuous homomorphism $\overline{\tau} : P \to N_{\mathbf{H}}(\mathbf{J})/\mathbf{J}(k)$ such that Ψ is $(P \times \Gamma)$-equivariant with respect to $\overline{\tau}$.

If $\mathbf{J}$ is trivial then we are done by Proposition 5.1. We therefore assume that $\mathbf{J}$ is nontrivial. It follows, again by Proposition 5.1, that $\overline{\tau}$ is incompatible. Proposition 3.4 implies that $\mathbf{J}$ is the center of the unipotent radical of a proper

parabolic subgroup of $\mathbf{H}$ and $\ker(\overline{\tau})$ is one-dimensional. By Proposition 6.1 we conclude that ρ indeed extends to a continuous homomorphism from G to $\mathbf{H}(k) = \mathrm{PU}(n, 1)$. This completes the proof. □

The proof of Proposition 6.1 will be given in Sect. 6.5. Before giving the proof, we require preliminary subsections discussing the notion of *fiber products* in the measured category and the incidence geometry of *chains* on the ideal boundary of complex hyperbolic space.

6.1 Fiber products

We begin by recalling some standard definitions. We will consider the category of Lebesgue spaces and their morphisms, where a Lebesgue space is a standard Borel space endowed with a measure class. An almost everywhere defined measurable map between Lebesgue spaces is said to be measure class preserving if the preimage of every null set is null. A morphism of Lebesgue spaces is an equivalence class of almost everywhere defined measure class preserving measurable maps, where two such maps are equivalent if they agree almost everywhere.

We now recall the definition of the fiber product in this category. Suppose that X, Y, Z are Lebesgue spaces endowed with probability measures μ_X, μ_Y, and μ_Z, respectively. Let $\phi : X \to Z$ and $\psi : Y \to Z$ be maps such that the measures are compatible in the sense that $\phi_*\mu_X$ and $\psi_*\mu_Y$ are in the same measure class as μ_Z. The corresponding set theoretic fiber product is

$$X \times_Z Y = \{(x, y) \in X \times Y \;\; \phi(x) = \psi(y)\} \subseteq X \times Y.$$

For μ_Z almost every $z \in Z$, disintegration of ϕ and ψ give measures $\nu_{X,z}$ and $\nu_{Y,z}$ on the corresponding fibers $\phi^{-1}(z) \subset X$ and $\psi^{-1}(z) \subset Y$ such that

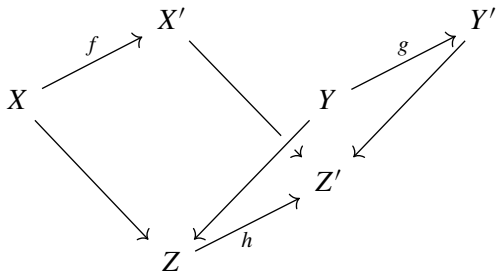
$$\mu_X = \int_Z \nu_{X,z}\, d\mu_Z \qquad \mu_Y = \int_Z \nu_{Y,z}\, d\mu_Z.$$

We then define a measure

$$\mu_X \times_Z \mu_Y = \int_Z (\nu_{X,z} \times \nu_{Y,z})\, d\mu_Z,$$

on $X \times Y$ whose equivalence class is supported on the set theoretic fibered product $X \times_Z Y$. Further, note that this measure class is independent of the choices of representatives for ϕ and ψ and the choices of representatives for μ_X, μ_Y, and μ_Z. We thus view this measure as a measure class on $X \times_Z Y$ and call it the *fiber product measure*. For details, see for instance [55, p. 265]. We will need the following lemma.

Lemma 6.2 *Consider Lebesgue spaces X, Y, and Z endowed with morphisms $X \to Z$ and $Y \to Z$. Assume X', Y', and Z' are standard Borel spaces endowed with Borel maps $X' \to Z'$ and $Y' \to Z'$, and moreover that there are almost everywhere defined measurable maps $f : X \to X'$, $g : Y \to Y'$, and $h : Z \to Z'$ such that the following diagram commutes:*



Consider the product map $f \times g : X \times Y \to X' \times Y'$ and the fiber product measure class $\mu_X \times_Z \mu_Y$ on $X \times_Z Y \subset X \times Y$. Then for $\mu_X \times_Z \mu_Y$ almost every (x, y), one has that $(f \times g)(x, y) \in X' \times_{Z'} Y'$.

Proof Using the choices and notation introduced above, we have

$$(f \times g)_*(\mu_X \times_Z \mu_Y) = \int_{Z'} (f_*\nu_{X,z} \times g_*\nu_{Y,z})\, dh_*\mu_Z,$$

and conclude that $(f \times g)_*(\mu_X \times_Z \mu_Y) = f_*\mu_X \times_{Z'} g_*\mu_Y$. That is, the push forward of the fiber product measure is the fiber product of the push forward measures. It follows that this measure is indeed supported on the set theoretical fiber product $X' \times_{Z'} Y'$. We then get that:

$$\begin{aligned}
&(\mu_X \times_Z \mu_Y)\left((f \times g)^{-1}((X' \times Y') \smallsetminus (X' \times_{Z'} Y'))\right) \\
&\quad = (f \times g)_*(\mu_X \times_Z \mu_Y)((X' \times Y') \smallsetminus (X' \times_{Z'} Y')) \\
&\quad = (f_*\mu_X \times_{Z'} g_*\mu_Y)((X' \times Y') \smallsetminus (X' \times_{Z'} Y')) \\
&\quad = 0
\end{aligned}$$

This proves the lemma. □

6.2 Real algebraic varieties as Lebesgue spaces

Assume V is a real algebraic variety. Any choice of smooth volume form on the nonsingular locus of V gives rise to a same measure class on V which we will call the *volume measure class* and denote μ_V. Thus V has a natural structure as a Lebesgue space. We will use the following fact.

Lemma 6.3 *Let X, Y, and Z be real algebraic varieties and let $\phi : X \to Z$, $\psi : Y \to Z$ be real rational maps that are surjective on real points such that $\phi_*\mu_X$ and $\psi_*\mu_Y$ are in the same measure class as μ_Z. Consider the corresponding fiber product $X \times_Z Y$ in the category of real algebraic varieties. Then the volume measure class on $X \times_Z Y$ coincides with the corresponding fiber product measure class of the volume measure classes on X, Y, and Z.*

6.3 The incidence geometry of chains

Let $\mathbb{B}^n$ denote complex hyperbolic n-space, and recall the notation from Sect. 2.2. Then $G = \mathrm{SU}(n, 1)$ acts transitively on $\mathbb{B}^n$ with point stabilizers conjugates of $K = \mathrm{S}(U(n) \times U(1))$, the maximal compact subgroup of G. We can also identify $\mathbb{B}^n$ with the unit ball in $\mathbb{C}^n$ with its Bergman metric. Note that $\mathbb{B}^1$ is naturally identified with the Poincaré disk model of hyperbolic 2-space.

Let $\partial\mathbb{B}^n$ denote the ideal boundary of $\mathbb{B}^n$, which is homeomorphic to the $(2n-1)$-sphere S^{2n-1}. Identify $\partial\mathbb{B}^n$ with the space of isotropic lines in $\mathbb{C}^{n+1}$ with respect to a hermitian form h of signature $(n, 1)$, thus with the real algebraic variety $P\backslash G$ for $P < G$ a minimal parabolic subgroup. The space of pairs of points $\partial\mathbb{B}^n \times \partial\mathbb{B}^n$ is then identified with $P\backslash G \times P\backslash G$ and the Zariski open subset $(\partial\mathbb{B}^n)^{(2)}$ consisting of pairs of distinct points is identified with $MA\backslash G$, i.e., the space of oriented geodesics in $\mathbb{B}^n$.

A totally geodesic holomorphic embedding $f : \mathbb{B}^1 \hookrightarrow \mathbb{B}^n$ induces an embedding

$$f_\infty : \partial\mathbb{B}^1 \simeq S^1 \hookrightarrow \partial\mathbb{B}^n,$$

and $f_\infty(\partial\mathbb{B}^1)$ is called a *chain* on $\partial\mathbb{B}^n$. Denote the space of chains on $\partial\mathbb{B}^n$ by $\mathcal{C}$. Chains were originally studied by Cartan [13] and we refer to [24] for basic facts about them. Chains are in one-to-one correspondence with two-dimensional subspaces of $\mathbb{C}^{n+1}$ on which h is nondegenerate of signature $(1, 1)$. In particular, the G-action on $\mathcal{C}$ is transitive and the stabilizers are conjugate in G to $D = \mathrm{S}(U(1, 1) \times U(n-1))$. Therefore, we have the identification $\mathcal{C} \simeq D\backslash G$.

Given a natural number k, we denote by $\mathcal{C}_k$ the space whose elements are a chain and k points lying on that chain. That is,

$$\mathcal{C}_k = \{(c, x_1, \ldots, x_k) \;\; c \in \mathcal{C}, \;\; x_1, \ldots, x_k \in c\}.$$

Let $\mathcal{C}_{(k)}$ denote the subset of $\mathcal{C}_k$ where the points are distinct, i.e.,

$$\mathcal{C}_{(k)} = \left\{(c, x_1, \ldots, x_k) \in \mathcal{C}_k \;\; x_i \neq x_j \text{ for } i \neq j\right\}.$$

Note that $\mathcal{C}_0 = \mathcal{C}$ and $\mathcal{C}_1 = \mathcal{C}_{(1)}$ is the set of pairs consisting of a chain and a point on that chain. Thus $\mathcal{C}_0 \simeq D\backslash G$ and $\mathcal{C}_1 \simeq (P \cap D)\backslash G = MAZ\backslash G$. In particular, these spaces have real algebraic variety structures and we endow them with the corresponding volume measure classes, which are also the unique G-invariant measure classes. We have the projection $\mathcal{C}_1 \to \mathcal{C}$ forgetting the point, which corresponds to the projection $(P \cap D)\backslash G \to D\backslash G$. This map is real algebraic and G-equivariant. The space $\mathcal{C}_k$ can be seen as the kth fiber power of the above projection and is thus also endowed with a natural real algebraic variety structure and a corresponding measure class. The space $\mathcal{C}_{(k)}$ is a dense Zariski open subset of $\mathcal{C}_k$ and then is also endowed with a real algebraic variety structure and a corresponding measure class. In particular, the inclusion $\mathcal{C}_{(k)} \subset \mathcal{C}_k$ is an isomorphism of Lebesgue spaces.

In this paper we will primarily be concerned with $\mathcal{C}_{(2)}$ and $\mathcal{C}_{(3)}$. A distinct pair of points $x, y \in \partial\mathbb{B}^n$ determines a unique chain, which is easily seen by considering x and y as isotropic lines in $\mathbb{C}^{n+1}$ and taking their span. Thus there are G-equivariant isomorphisms $MA\backslash G \simeq (\partial\mathbb{B}^n)^{(2)} \simeq \mathcal{C}_{(2)}$.

We now consider the subset

$$\mathcal{C}_k^0 = \left\{(c, x_1, \ldots, x_k) \in \mathcal{C}_k \;\; x_i \neq x_j \text{ for some } i \neq j\right\} \subset \mathcal{C}_k$$

consisting of k-tuples of points containing at least two distinct points and the subset

$$\mathcal{C}_k^1 = \{(c, x_1, \ldots, x_k) \in \mathcal{C}_k \;\; x_i \neq x_1 \text{ for } i \neq 1\} \subset \mathcal{C}_k^0$$

where the first point in each chain is different from the others. We note that

$$\mathcal{C}_{(k)} \subset \mathcal{C}_k^1 \subset \mathcal{C}_k^0 \subset \mathcal{C}_k,$$

and these are all dense Zariski open subsets of $\mathcal{C}_k$. In particular, they are all isomorphic in the category of Lebesgue spaces.

Remark 6.4 The space $\mathcal{C}_k^0$ can be alternatively described as:

$$\left\{(x_1, \ldots, x_k) \;\; x_i \subset \mathbb{C}^{n+1} \text{ an isotropic line, } \dim_{\mathbb{C}} \operatorname{span}\{x_1, \ldots, x_n\} = 2\right\}.$$

It is this space that is considered in Pozzetti's chain rigidity theorem, which is Theorem 6.5 below.

We now make an observation that will be useful later. For each k we define $\alpha_k : \mathcal{C}_k^1 \to \mathcal{C}_1$ by $(c, x_1, \ldots, x_k) \mapsto (c, x_1)$. In particular we have the map $\alpha_2 : \mathcal{C}_{(2)} = \mathcal{C}_2^1 \to \mathcal{C}_1$ that forgets the second point of each chain. We observe that $\alpha_k : \mathcal{C}_k^1 \to \mathcal{C}_1$ can be identified in the category of algebraic varieties with

the $(k-1)^{st}$ fibered power of α_2. It follows by Lemma 6.3 that $\mathcal{C}_k^1$ is isomorphic to the $(k-1)^{st}$ fibered power of α_2 also in the category of Lebesgue spaces.

Note that $\mathcal{C}_{(2)} \simeq MA\backslash G$, but G is not transitive on $\mathcal{C}_{(k)}$ already for $k = 3$. Indeed, $\mathcal{C}_{(3)}$ has two G-orbits distinguished by the cyclic orientations of the given triple of points on each chain. The stabilizer in G of each element of $\mathcal{C}_{(3)}$ is a compact subgroup, namely a conjugate of $M < G$. It follows that the measure class on $\mathcal{C}_{(3)}$ is the sum of the unique G-invariant measure classes on the two G-orbits. Each of $\mathcal{C}_3^0$, $\mathcal{C}_3^1$, and $\mathcal{C}_3$ are obtained by adding some lower-dimensional manifolds, which are null sets, to $\mathcal{C}_{(3)}$.

6.4 Pozzetti's chain rigidity theorem

In [13], Cartan studied the incidence geometry of chains in $\partial\mathbb{B}^n$ and showed that a map preserving this geometry must come from an isometry of $\mathbb{B}^n$. His work was later generalized by Burger–Iozzi to almost everywhere defined measurable maps under the added assumption that the map preserves orientation [11]. This assumption was later removed by Pozzetti [52], and this is the result we will need.

In the formulation of the theorem below we use the notation introduced in Sect. 6.3. Recall that $\partial\mathbb{B}^n$ represents the variety of isotropic lines in $\mathbb{C}^{n+1}$, and under this identification, $\mathcal{C}_3^0$ represents triples of isotropic lines that span a two-dimensional subspace. Briefly, Pozzetti's chain rigidity theorem is that a measurable map from $\partial\mathbb{B}^n$ to $\partial\mathbb{B}^n$ that takes triples of points on a chain to triples of points on a chain is necessarily rational; for more on rationality in this general context see [69, Ch. 3]. Restated precisely in our language, this is the following.

Theorem 6.5 (Thm. 1.6 [52]) *For $n \geq 2$, let $\phi : \partial\mathbb{B}^n \to \partial\mathbb{B}^n$ be a measurable map whose essential image is Zariski dense. Endow $(\partial\mathbb{B}^n)^3$ with the measure class $[\mu]$ associated with the volume form on $\mathcal{C}_3^0$ and consider the induced map $\phi^3 : (\partial\mathbb{B}^n)^3 \to (\partial\mathbb{B}^n)^3$. If the $[\mu]$-essential image of ϕ^3 is contained in $\mathcal{C}_3^0$, then ϕ agrees almost everywhere with a rational map.*

On the proof of Theorem 6.5 The statement of [52, Thm. 1.6] assumes that the target of ϕ is the Shilov boundary $\mathcal{S}_{m_1,m_2}$ associated with SU(m_1, m_2) when $1 < m_1 < m_2$. However, nowhere in the proof of the theorem does Pozzetti use the assumption that $1 < m_1$, and hence her result also holds for $\mathcal{S}_{1,m}$, $m > 1$, and in particular for $\mathcal{S}_{1,n}$. The Shilov boundary is the space of maximal isotropic subspaces for the relevant hermitian form, hence $\mathcal{S}_{1,n} \simeq \partial\mathbb{B}^n$. In other words, Pozzetti's proof applies without alteration to give Theorem 6.5 as stated. The reason for Pozzetti's assumption that $m_1 > 1$ is that the analogue of her main result was previously proved by Burger–Iozzi [11], however Burger and

Iozzi did not need a statement as strong as Theorem 6.5 to prove their main result. □

We recall from Sect. 6.3 that $\mathcal{C}_3^1$ can be identified with the fibered square of the forgetful map $\alpha_2 : \mathcal{C}_{(2)} \to \mathcal{C}_1$ and that $\mathcal{C}_{(2)} \simeq (\partial\mathbb{B}^n)^{(2)}$. Upon making this identification, we consider the map $\alpha : (\partial\mathbb{B}^n)^{(2)} \to \mathcal{C}_1$ and regard $\mathcal{C}_3^1$ as the fibered square of α. We then obtain the following corollary of Theorem 6.5.

Corollary 6.6 *Let G be* $\mathrm{SU}(n,1)$ *for $n \geq 2$, $\Gamma < G$ be a lattice, H be* $\mathrm{PU}(n,1)$*, and $\rho : \Gamma \to H$ be a homomorphism with unbounded, Zariski dense image. Assume that $\phi : \partial\mathbb{B}^n \to \partial\mathbb{B}^n$ and $\psi : \mathcal{C}_1 \to \mathcal{C}_1$ are measurable, Γ-equivariant maps, where Γ acts on the domain through its inclusion into G and on the target via ρ. Then the following assertions hold.*

(1) *The essential image of $\phi^2 : (\partial\mathbb{B}^n)^2 \to (\partial\mathbb{B}^n)^2$ is contained in the set $(\partial\mathbb{B}^n)^{(2)}$ of distinct points.*
(2) *Considering the restricted map $\phi^{(2)} : (\partial\mathbb{B}^n)^{(2)} \to (\partial\mathbb{B}^n)^{(2)}$, if $\alpha \circ \phi^{(2)}$ and $\psi \circ \alpha$ agree almost everywhere as maps from $(\partial\mathbb{B}^n)^{(2)}$ to $\mathcal{C}_1$, then ρ extends to a continuous homomorphism from G to H.*

The content of (2) in Corollary 6.6 is that the boundary map ϕ sends chains to chains and that the induced map on chains is ψ.

Proof We first prove (1). Note that $(\partial\mathbb{B}^n)^{(2)} \subset (\partial\mathbb{B}^n)^2$ is open and Zariski dense, hence of full measure. Moreover $(\partial\mathbb{B}^n)^{(2)}$ is isomorphic to $MA\backslash G$ as a G-space and so it is Γ-ergodic by Howe–Moore. It follows that the essential image of ϕ^2 is either contained in $(\partial\mathbb{B}^n)^{(2)}$ or in its complement, which is the diagonal in $(\partial\mathbb{B}^n)^2$. If the latter were true then ϕ would be essentially constant, and its essential image is invariant under Γ. This would imply that $\rho(\Gamma)$ is contained in a proper parabolic subgroup of H, contradicting Zariski density of $\rho(\Gamma)$. It therefore follows that the essential image of ϕ^2 is contained in the set of pairs of distinct points, $(\partial\mathbb{B}^n)^{(2)}$. This proves (1).

To prove (2), assume that $\alpha \circ \phi^{(2)}$ and $\psi \circ \alpha$ agree almost everywhere. By a well-known lemma of Margulis [69, Lem. 5.1.3], to show that ρ extends it suffices to show that ϕ is rational. Therefore it suffices to show that Theorem 6.5 applies.

Endow $(\partial\mathbb{B}^n)^3$ with the measure class $[\mu]$ associated with the volume form on $\mathcal{C}_3^0$. We will show that the $[\mu]$-essential image of ϕ^3 is contained in $\mathcal{C}_3^0$. In fact, we will show that it is contained in the subset $\mathcal{C}_3^1$ of $\mathcal{C}_3^0$. Since $\mathcal{C}_3^1$ is conull in $\mathcal{C}_3^0$, we view $[\mu]$ as the measure class associated with the volume form on $\mathcal{C}_3^1$.

As indicated above, we identify $\mathcal{C}_3^1$ with the fibered square of the map α from $(\partial\mathbb{B}^n)^{(2)}$ to $\mathcal{C}_1$, which is naturally a subset of $(\partial\mathbb{B}^n)^{(2)} \times (\partial\mathbb{B}^n)^{(2)}$, even

though $\mathcal{C}_3^1$ is a subset of $(\partial\mathbb{B}^n)^3$ by its original definition. This is clarified by the commutative diagram

$$\begin{array}{ccc} (\partial\mathbb{B}^n)^{(2)} \times_{\mathcal{C}_1} (\partial\mathbb{B}^n)^{(2)} & \hookrightarrow & \left\{(x,y),\ (x,z) \in (\partial\mathbb{B}^n)^{(2)}\ \ y,z \neq x\right\} \\ \updownarrow\simeq & & \updownarrow\simeq \\ \mathcal{C}_3^1 & \hookrightarrow & \left\{(x,y,z) \in (\partial\mathbb{B}^n)^3\ \ y,z \neq x\right\} \end{array}$$

in which the top two spaces are subsets of $(\partial\mathbb{B}^n)^{(2)} \times (\partial\mathbb{B}^n)^{(2)}$ and the bottom two spaces are subsets of $(\partial\mathbb{B}^n)^3$ containing the support of $[\mu]$. Notice that this diagram also commutes with applying $\phi^{(2)} \times \phi^{(2)}$ to the top line and ϕ^3 to the bottom line.

Further, the assumption that $\alpha \circ \phi^{(2)}$ and $\psi \circ \alpha$ agree almost everywhere allows us to apply Lemma 6.2 with:

$$\begin{array}{ll} X = Y = X' = Y' = (\partial\mathbb{B}^n)^{(2)} & f = g = \phi^{(2)} \\ Z = Z' = \mathcal{C}_1 & h = \psi \end{array}$$

This allows us to conclude that the essential image of $\phi^{(2)} \times \phi^{(2)}$ is contained in $(\partial\mathbb{B}^n)^{(2)} \times_{\mathcal{C}_1} (\partial\mathbb{B}^n)^{(2)}$ with respect to the fibered measure class of $(\partial\mathbb{B}^n)^{(2)} \times_{\mathcal{C}_1} (\partial\mathbb{B}^n)^{(2)}$ on $(\partial\mathbb{B}^n)^{(2)} \times (\partial\mathbb{B}^n)^{(2)}$. In view of the above discussion, it follows that, with respect to the measure class associated with $\mathcal{C}_3^1$ on $(\partial\mathbb{B}^n)^3$, the essential image of ϕ^3 is contained in $\mathcal{C}_3^1$. This completes the proof. □

6.5 Proof of Proposition 6.1

Let G be SU$(n, 1)$ for $n \geq 2$ and $\Gamma < G$ be a lattice. Suppose that $H = \mathrm{PU}(n, 1)$, $\overline{Z} < H$ is the center of the unipotent radical of a proper parabolic subgroup $Q < H$, and $\rho : \Gamma \to H$ is a homomorphism with unbounded, Zariski dense image. Assume that there exists a continuous homomorphism $\tau : P \to Q/\overline{Z}$ with one-dimensional kernel and a measurable map $\Phi : G \to H/\overline{Z}$ that is $(P \times \Gamma)$-equivariant with respect to the left Γ-action via ρ and the right P-action through τ. We must show that ρ extends to a continuous homomorphism from G to H.

We will freely use the notation introduced in Sect. 2.2. Identifying Q with P/C we consider the natural map $\theta : P \to P/CZ \cong Q/\overline{Z}$. For a fixed $p_0 \in P$, the map $G \to H/\overline{Z}$ given by $g \mapsto \Phi(gp_0^{-1})$ is $(P \times \Gamma)$-equivariant with respect to the left Γ-action and the right P-action via $\tau \circ \mathrm{inn}(p_0)$. We can then replace τ with $\tau \circ \mathrm{inn}(p_0)$, hence Proposition 2.12 allows us to assume that τ is surjective, $\ker(\tau) = CZ$, and $\tau(P \cap D) = \theta(P \cap D)$.

The composition of $\Phi : G \to H/\overline{Z}$ with $H/\overline{Z} \to H/Q$ is right P-invariant, and hence gives rise to a Γ-map $P\backslash G \to H/Q$. Upon identifying $P\backslash G$ with $H/Q \simeq \partial\mathbb{B}^n$, we consider the above composition as a Γ-equivariant measurable map $\phi : \partial\mathbb{B}^n \to \partial\mathbb{B}^n$. By Corollary 6.6(1), the map $\phi^2 : (\partial\mathbb{B}^n)^2 \to (\partial\mathbb{B}^n)^2$ restricts to a map $\phi^{(2)} : (\partial\mathbb{B}^n)^{(2)} \to (\partial\mathbb{B}^n)^{(2)}$.

Next we consider the map α obtained by composing the natural identification $(\partial\mathbb{B}^n)^{(2)} \simeq \mathcal{C}_{(2)}$ with the forgetful map $\alpha_2 : \mathcal{C}_{(2)} \to \mathcal{C}_1$. In other words, α maps the pair of distinct points (x, y) to the pair consisting of the unique chain through x and y and the point x. Identify $MA\backslash G$ with $(\partial\mathbb{B}^n)^{(2)}$ and $\mathcal{C}_1$ with $(H/\overline{Z})/\theta(P \cap D)$, where the latter identification comes from the fact that $\theta(P \cap D) = \tau(P \cap D)$.

Our goal is now to prove the existence of the dashed arrows in the following diagram of $(P \times \Gamma)$-equivariant measurable maps:

$$\begin{tikzcd}
G \arrow[d] & & & \\
MA\backslash G \arrow[r, "\simeq"] \arrow[d] & (\partial\mathbb{B}^n)^{(2)} \arrow[r, "\phi^{(2)}"] \arrow[d, "\alpha"] & (\partial\mathbb{B}^n)^{(2)} \arrow[d, "\alpha"] & \\
(P \cap D)\backslash G \arrow[r, "\simeq"] \arrow[rrr, dashed, bend right, "\psi'"'] & \mathcal{C}_1 \arrow[r, dashed, "\psi"] & \mathcal{C}_1 \arrow[r, "\simeq"] & (H/\overline{Z})/\theta(P \cap D)
\end{tikzcd}$$

For this, the composed map $G \to (H/\overline{Z})/\theta(P \cap D)$ is $(P \cap D)$-invariant, since $\theta(P \cap D) = \tau(P \cap D)$, thus it descends to a map from $(P \cap D)\backslash G$ to $(H/\overline{Z})/\theta(P \cap D)$, which proves the existence of ψ'. The existence of ψ is then obtained by pre- and post-composing with the corresponding isomorphisms. We thus have that $\alpha \circ \phi^{(2)}$ and $\psi \circ \alpha$ agree almost everywhere, and therefore we can apply Corollary 6.6(2) and conclude that ρ extends to a continuous homomorphism from G to H. This proves the proposition. □

Remark 6.7 The use of incidence geometry to prove rigidity theorems goes back at least to Mostow's use of work of Tits in his proof of Mostow Rigidity [48]. Rigidity of chain preserving maps is older, going back to Cartan's 1932 paper [13]. An important application of rigidity of chain preserving maps in the study of representations of discrete groups was by Burger and Iozzi in [11]. The idea of exploiting triples of points on a chain goes back to earlier work of Toledo on rigidity of certain surface group representations into $\mathrm{PU}(n, 1)$ [61], and Toledo attributes this general idea to Thurston. As mentioned above, we cannot use Cartan's result because our boundary map is only measurable and we cannot use Burger–Iozzi's because they assume orientability of the map on chains, but Pozzetti chain rigidity theorem suffices for our purposes.

Remark 6.8 Margulis and Mohammadi proved a version of Theorem 1.1 for cocompact lattices in SO(3, 1) using incidence geometry. It is possible to use the methods of this section combined with results in [2] to produce the input for their incidence geometry result, i.e., to prove the existence of a *circle preserving map* between boundaries. However, the proof of Theorem 1.1 in the real hyperbolic case is easier to complete using compatibility as we did in [2]. More generally though, it is possible for one to apply Lemma 3.3 above and use the methods of this paper to turn some problems about homomorphisms from lattices in $\mathrm{SO}_0(n, 1)$ to incompatible targets into incidence geometry problems for boundary maps.

7 Algebraic groups associated with lattices in SU(n, 1) and Theorem 1.5

We begin by establishing some notation we need for the proof of Theorem 1.5. Let $\Gamma < \mathrm{SU}(n, 1)$ be a lattice, $n \geq 2$. Following the discussion of [2, Section 3], local rigidity (due to Calabi–Vesentini [12] in the cocompact case and Raghunathan [53] otherwise) along with work of Vinberg [66] implies that the field $\ell = \mathrm{Tr}(\mathrm{Ad}(\Gamma))$ is both a number field and a minimal field of definition for Γ. Let $\overline{\mathbf{G}}$ denote the connected adjoint ℓ-algebraic group defined by the Zariski closure of Γ under the adjoint representation.

Then $\overline{\mathbf{G}}$ is the adjoint form of a unique simply connected ℓ-group $\mathbf{G}$ of type ${}^2\mathrm{A}_n$ [60]. In other words, there is a quadratic extension ℓ'/ℓ and a central division algebra D over ℓ' of degree $d \mid (n+1)$ with involution σ of the second kind so that $\mathbf{G}$ is the special unitary group $\mathrm{SU}_{(n+1)/d}(D, h)$ for some nondegenerate σ-hermitian form h on $D^{(n+1)/d}$. Recall that σ is of the second kind if its restriction to the center of D is the nontrivial Galois involution of ℓ'/ℓ. The image of Γ under the adjoint representation lies in $\overline{\mathbf{G}}(\ell)$ and the kernel of $\mathrm{SU}(n, 1) \to \mathrm{PU}(n, 1)$ is cyclic of order $(n+1)$, hence we can replace Γ with a subgroup of finite index and assume that $\Gamma < \mathbf{G}(\ell)$.
With this setup, we now prove Theorem 1.5.

Proof of Theorem 1.5 Let $\Gamma < \mathrm{SU}(n, 1)$ be a lattice, and retain all previous definitions and notation from this section. As described above, we can assume that $\Gamma < \mathbf{G}(\ell)$. For each place v of ℓ, we obtain a homomorphism

$$\rho_v : \Gamma \hookrightarrow G_v = \mathbf{G}(\ell_v),$$

where ℓ_v is the completion of ℓ associated with v. Let v_0 be the place associated with the lattice embedding of Γ into $\mathrm{SU}(n, 1)$.

Our first goal is to prove that ρ_v is locally rigid for all v. Recall that local rigidity of ρ_v is the vanishing of the cohomology group $H^1(\rho_v, \mathfrak{g}_v)$, where $\mathfrak{g}_v$ is the Lie algebra of G_v. The assumption that $\Gamma < \mathbf{G}(\ell)$ implies that there is

an ℓ-form $\mathfrak{g}_\ell$ of $\mathfrak{su}(n,1)$ so that $\mathfrak{g}_v \cong \mathfrak{g}_\ell \otimes_\ell \ell_v$ for all v. Note that vanishing of $H^1(\rho_v, \mathfrak{g}_v)$ is equivalent to vanishing of $H^1(\Gamma, \mathfrak{g}_\ell)$ since

$$H^1(\rho_v, \mathfrak{g}_v) \cong H^1(\Gamma, \mathfrak{g}_\ell) \otimes_\ell \ell_v.$$

However, $H^1(\Gamma, \mathfrak{g}_\ell)$ is trivial since $H^1(\rho_{v_0}, \mathfrak{g}_{v_0})$ is trivial by local rigidity of the lattice embedding of Γ into $\mathrm{SU}(n,1)$. This proves local rigidity of ρ_v for all v.

We now make some technical reductions needed to apply some results as they are stated in the literature. First, without loss of generality we can pass to a finite index torsion-free subgroup of Γ, since ℓ, ℓ', and $\mathbf{G}$ are commensurability invariants by [66]. When Γ is cocompact, we can therefore assume that $\mathbb{B}^n/\Gamma$ is a compact Kähler manifold. When Γ is not cocompact, we can assume that $\mathbb{B}^n/\Gamma$ is a smooth quasiprojective variety admitting a smooth toroidal compactification by a smooth divisor D (e.g., see [1,45]).

More precisely in the noncompact case, we can assume by passing to a finite index subgroup that the cusps of $\mathbb{B}^n/\Gamma$ are diffeomorphic to bundles over the punctured disk with fiber an abelian variety. The associated peripheral subgroup of Γ is a two-step nilpotent group with infinite cyclic center naturally realized as a torsion-free lattice in the unipotent radical of the Borel subgroup of $\mathrm{PU}(n,1)$ (cf. the structure of U in Sect. 2.2). The center of this nilpotent group is generated by the free homotopy class of a loop projecting to a loop around the puncture of the disk in the base of the bundle. One can then smoothly compactify $\mathbb{B}^n/\Gamma$ to obtain a smooth projective variety by adding a certain abelian variety above the puncture in the disk. See Fig. 1 and see [28, Section 4.2] for further details in this specific case.

Suppose that v is archimedean, i.e., that ℓ_v is $\mathbb{R}$ or $\mathbb{C}$. Rigidity of ρ_v then implies that the real Zariski closure of $\rho_v(\Gamma)$, namely G_v, is a group of *Hodge type*. See [58, Lem. 4.5] when Γ is cocompact. In general see [43, Section 10.2.3], which handles representations of fundamental groups of smooth quasiprojective varieties admitting a smooth compactification, and hence by the above assumptions applies to $\mathbb{B}^n/\Gamma$. Since $\mathbf{G}$ is of absolute type A_n, considering [58, p. 50–51], we conclude that v is real and G_v must be $\mathrm{SU}(r_v, s_v)$ for an appropriate pair r_v, s_v, completing the proof of the first two parts of the theorem.

It remains to prove that Γ is integral. Indeed, the last statement of Theorem 1.5 is an immediate consequence of the previous statements along with the definition of arithmeticity. Integrality will follow from a theorem of Esnault–Groechenig [22, Thm. 1.1] once we verify that their assumptions hold in our setting.

Since ρ_v is locally rigid and has determinant one, it remains to verify that ρ_v has quasi-unipotent monodromy at infinity. See [22, Section 2] for the pre-

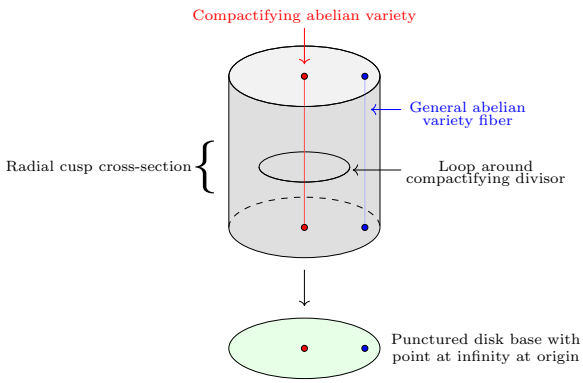


Fig. 1 Smooth compactification of a cusp

cise definition. Indeed, loops around the compactifying divisor of $\mathbb{B}^n/\Gamma$ are associated with central elements of peripheral subgroups of Γ, which are naturally unipotent subgroups of $\mathbf{G}(\ell)$. These elements clearly remain unipotent under ρ_v, hence ρ_v has quasi-unipotent monodromy at infinity. We then conclude from [22, Thm. 1.1] that Γ is integral, which completes the proof of the theorem. □

Remark 7.1 With notation as in Theorem 1.5, if $\Gamma < \mathrm{PU}(n, 1)$ is a nonarithmetic lattice, there is another real place v so that $\mathbf{G}(\ell_v) \cong \mathrm{PU}(p, q)$ for some $p+q = n+1$ and $p \geq q$. When $n = 2$, the only possibility is $(p, q) = (2, 1)$. There are only two known commensurability classes of nonarithmetic lattices, and in each case $(p, q) = (3, 1)$ for every noncompact $\mathbf{G}(\ell_v)$ [20]. We do not know if it is possible to have a nonarithmetic lattice with $\mathbf{G}(\ell_v) \cong \mathrm{PU}(p, q)$ with $q \geq 2$.

8 The proof of Theorem 1.1

8.1 Equidistribution on G/Γ

This subsection describes the equidistribution results needed to show that Theorem 1.3 applies to prove Theorem 1.1. Each result in this section has a direct analogue in [2], and here we provide the necessary modifications for the SU(n, 1) setting.

Throughout this section we fix a lattice $\Gamma < G$. Recall that a measure μ on G/Γ is called *homogeneous* if there exists a closed subgroup $S \leq G$ and a closed S-orbit in G/Γ such that μ is the push-forward of Haar measure on S along this orbit. If $W \leq S$ is a closed subgroup with respect to which μ is ergodic, we call the measure *W-ergodic*. Given a homogeneous, W-ergodic measure μ, we will refer to its support, $\mathrm{supp}(\mu)$, as a *homogeneous, W-ergodic subspace* of G/Γ.

We ask the reader to recall the notation introduced in Sect. 2.1. In particular, we have a fixed maximal compact subgroup $K \leq G$ and we consider the symmetric space $K\backslash G$. The locally symmetric space $K\backslash G/\Gamma$ will be denoted by M. In Definition 2.2 we introduced the standard (almost simple) subgroups of G, namely certain copies in G of the group $\mathrm{SU}(m,1)$ for $1 \leq m \leq n$ and $\mathrm{SO}_0(m,1)$ for $2 \leq m \leq n$ which, by Proposition 2.4, are representatives of conjugacy classes of all noncompact, connected, almost simple, closed subgroups of G.

In what follows "geodesic subspace" will always mean a properly immersed totally geodesic subspace of either M or its universal cover $K\backslash G$. Our goal in this subsection is to prove the following proposition, which translates the existence of infinitely many maximal geodesic subspaces of M into a statement about measures on G/Γ that are invariant under a standard subgroup of G.

Proposition 8.1 (Cf. Prop. 3.1 in [2]) *The following are equivalent:*

(1) *The complex hyperbolic space $M = K\backslash G/\Gamma$ contains infinitely many maximal geodesic subspaces of real dimension at least* 2.
(2) *There exists a standard subgroup $W < G$ and an infinite sequence $\{\mu_i\}$ of W-invariant, W-ergodic measures on G/Γ with proper support for which Haar measure on G/Γ is a weak-$*$ limit of the μ_i.*
(3) *There exists a standard subgroup $W < G$ and an infinite sequence $\{\mu_i\}$ of homogeneous, W-ergodic measures on G/Γ for which Haar measure on G/Γ is a weak-$*$ limit of the μ_i.*

As noted in [2], work of Ratner [56] shows that (2) and (3) are equivalent, and therefore it remains to prove that (1) and (3) are equivalent. For this, we closely follow the strategy in [2, Section 3.1], adapting the statements therein to the current setting.

For the remainder of this section, fix the notation $\pi : G/\Gamma \to K\backslash G/\Gamma = M$ for the natural projection map. Note that π is a proper map. We begin with the following lemma.

Lemma 8.2 *The following hold:*

(1) *Let S be a closed subgroup of G such that $W \leq S$ for some standard subgroup*
$W < G$, and suppose there exists $h \in G$ for which $Sh\Gamma/\Gamma$ is a closed S-orbit. Then $Z = \pi(Sh\Gamma/\Gamma)$ is a closed totally geodesic subspace of M of real dimension at least 2. *Up to normalization, the* $\dim(Z)$*-volume of Z is the push-forward of Haar measure on $Sh\Gamma/\Gamma$ via the projection map π.*
(2) *Under the assumptions of part (1), $Z = M$ if and only if $S = G$ and $Z = \pi(Wg\Gamma/\Gamma)$ for some $g \in G$ if and only if $W \leq S \leq N$, where N is the normalizer of W in G. In the latter case, $Nh\Gamma/\Gamma$ is also closed with projection $\pi(Nh\Gamma/\Gamma) = Z$.*

(3) *Conversely, every totally geodesic subspace Z in M of real dimension at least 2 has finite measure. Moreover, for any such Z there is a standard subgroup W of G with normalizer N, an intermediate subgroup $W \leq S \leq N$, and an element $h \in G$ such that $\Gamma \cap S^{h^{-1}}$ is a lattice in $S^{h^{-1}}$ and $Sh\Gamma/\Gamma$ is a homogeneous, W-ergodic subspace of G/Γ for which $Z = \pi(Wh\Gamma/\Gamma) = \pi(Sh\Gamma/\Gamma)$.*

Proof The fact that Z is closed follows from the fact that π is a proper map and the assumption that $Sh\Gamma/\Gamma$ is closed. By Lemma 2.7(4), $K\backslash KSh$ is a totally geodesic subspace of $K\backslash G$, thus Z, which is the image of $K\backslash KS$ under the covering map $K\backslash G \to M$, is a totally geodesic subspace of M. The statement about the measure also follows from Lemma 2.7(4), and this proves part (1).

Next we consider part (2). Clearly $S = G$ implies $Z = M$ and the fact that $Z = M$ implies $S = G$ follows by applying Lemma 2.7(1) in the special case $W = G$. If $W \leq S \leq N$ then

$$Z = \pi(Sh\Gamma/\Gamma) = \pi(Wh\Gamma/\Gamma),$$

by Lemma 2.7(3). Conversely, assume $Z = \pi(Wg\Gamma/\Gamma)$ for some $g \in G$. Then both $K\backslash KSh$ and $K\backslash KWg$ cover Z in $K\backslash G$, thus they are totally geodesic subspaces of the same type. Therefore $K\backslash KSh = K\backslash KWg_0$ for some $g_0 \in G$ by Proposition 2.3, and Lemma 2.7(4) implies that $K\backslash KSh = K\backslash KS^+h$. Hence we obtain the sequence of equalities

$$K\backslash Khg_0^{-1}(S^+)^{g_0h^{-1}} = K\backslash KS^+hg_0^{-1} = K\backslash KShg_0^{-1} = K\backslash KW,$$

and conclude that $(S^+)^{g_0h^{-1}}$ preserves $K\backslash KW$ and acts transitively on it. Thus by Lemma 2.7(1) we see that $(S^+)^{g_0h^{-1}} = W$.

Since $W \leq S$, we have $(S^+)^{g_0h^{-1}} = W = W^+ \leq S^+$ and deduce that $S^+ = (S^+)^{g_0h^{-1}} = W$. Since S normalizes S^+ it follows that $S \leq N$. Then W is cocompact in N, so S is cocompact in N by Lemma 2.7(2). It follows that $Nh\Gamma/\Gamma$ is closed in G/Γ, since $Sh\Gamma/\Gamma$ is closed by hypothesis. That $\pi(Nh\Gamma/\Gamma) = Z$ follows from Lemma 2.7(3), and this completes the proof of part (2).

We now prove part (3). Fix a totally geodesic subspace Z in M of real dimension at least 2. The fact that Z has a finite volume is a well-known consequence of the existence of a thick-thin decomposition. See [23] for detailed argument in real hyperbolic space that is easily adapted to any rank one symmetric space.

The preimage of Z under the Γ-invariant map $K\backslash G \to M$ is a collection of mutually disjoint totally geodesic subspaces on which Γ acts. We fix one of these, which by Proposition 2.3 is of the form $K\backslash KWg$ for some $g \in G$ and a standard subgroup $W \leq G$. By Lemma 2.7(2), the subgroup of G that stabilizes

this totally geodesic subspace is $N^{g^{-1}}$, thus the subgroup of Γ stabilizing it is $\Gamma \cap N^{g^{-1}}$. By Lemma 2.7(4) we have that $K\backslash KWg = K\backslash KNg = K\backslash KgN^{g^{-1}}$ which is isomorphic as an $N^{g^{-1}}$-space to $(K^{g^{-1}} \cap N^{g^{-1}})\backslash N^{g^{-1}}$. It follows that

$$Z = K\backslash KgN^{g^{-1}}\Gamma/\Gamma \simeq (K^{g^{-1}} \cap N^{g^{-1}})\backslash N^{g^{-1}}/(\Gamma \cap N^{g^{-1}}).$$

Since Z has a finite volume and $K^{g^{-1}} \cap N^{g^{-1}}$ is a compact subgroup of $N^{g^{-1}}$, we conclude that $\Gamma \cap N^{g^{-1}}$ is a lattice in $N^{g^{-1}}$. By [54, Thm. 1.13] we obtain that $Ng\Gamma/\Gamma$ is closed in G/Γ and conclude that it is a closed homogeneous N-orbit of finite volume for which the associated Haar measure on this orbit is a homogeneous measure.

However $Ng\Gamma/\Gamma$ may not be W-ergodic, even though it is W-invariant. To complete the proof, it remains to show that there exists an intermediate subgroup $W \leq S \leq N$ and an $h \in G$ such that $Sh\Gamma/\Gamma$ is a homogeneous, W-ergodic subspace of G/Γ. Let μ be a W-ergodic measure in the ergodic decomposition of Haar measure on $Ng\Gamma/\Gamma$, and let S denote the stabilizer of μ in N. Then $W \leq S \leq N$ by W-invariance, and Ratner's theorem implies that μ is S-homogeneous. Write the corresponding closed homogeneous space as $Sh\Gamma/\Gamma$ for some $h \in G$. Since $Sh\Gamma/\Gamma \simeq S/(\Gamma \cap S^{h^{-1}})$ we see that $\Gamma \cap S^{h^{-1}}$ is a lattice in $S^{h^{-1}}$. Since h is in the N-homogeneous space $Ng\Gamma/\Gamma$, we have that $Nh\Gamma/\Gamma = Ng\Gamma/\Gamma$. Using Lemma 2.7(4) twice, we see that

$$\begin{aligned}\pi(Wh\Gamma/\Gamma) &= \pi(Sh\Gamma/\Gamma) = \pi(Nh\Gamma/\Gamma)\\ &= \pi(Ng\Gamma/\Gamma) = \pi(Wg\Gamma/\Gamma) = Z,\end{aligned}$$

which completes the proof. □

We now collect some useful facts about limits of homogeneous, W-ergodic measures that we need in the proof of Proposition 8.1. Despite the difference in presentation, the next three results are direct analogues of results in the real hyperbolic case [2, Thm 3.3] and use an essentially identical argument. All of these results are relatively straightforward consequences of work of Dani–Margulis [18, Thm. 6.1] and Mozes–Shah [49]. We first show that there is no escape of mass for the sequence of measures under consideration.

Lemma 8.3 *Let $W < G$ be a closed, connected, almost simple subgroup of G that is generated by unipotent elements. If $\{\mu_i\}$ is a sequence of homogeneous, W-ergodic probability measures on G/Γ that weak-$*$ converges to a measure μ in the space of all finite Radon measures, then μ is not the zero measure.*

Proof Since the claim is invariant under conjugation, we can assume by Proposition 2.4 that W is a standard subgroup of G.

Let C be any fixed compact set in M whose interior contains the compact core of M. See [6, Thm. 10.5] for the existence of a compact core in this setting. Then C has the property that $C \cap Z$ contains a nonempty open subset of Z for any closed geodesic subspace Z of M. Set $F = \pi^{-1}(C)$.

Applying [18, Thm. 6.1] for $\epsilon = 1/2$, there exists a compact set $F' \subseteq G/\Gamma$ such that

$$\frac{1}{2} \leq \frac{1}{T}\lambda\left(\left\{t \in [0, T] \mid u_t x \in F'\right\}\right) = \frac{1}{T}\int_0^T \chi_{F'}(u_t x)dt, \tag{4}$$

for every one-parameter unipotent subgroup $\{u_t\}$ of G, every $x \in F$, and every $T \geq 0$, where $\chi_{F'}$ is the characteristic function of F' and λ is Lebesgue measure on $\mathbb{R}$. We claim that $\mu(F') \geq 1/2$, which proves that μ is not the zero measure.

Fix a one-parameter unipotent subgroup $U = \{u_t\}$ of W. Then W-ergodicity and the Howe–Moore theorem imply that each μ_i is U-ergodic. The measures μ_i are homogeneous and W-ergodic, so Lemma 8.2(1) implies that for each i there exists a closed totally geodesic subspace Z_i of M with real dimension at least 2 such that $\pi_*\mu_i$ is a constant multiple of the volume measure on Z_i. Since $C \cap Z_i$ contains an open subset of Z_i, we see that

$$\pi_*\mu_i(C) = \mu_i(\pi^{-1}(C)) > 0.$$

In particular, there exists a U-generic point $x_i \in F$ for each μ_i. The Birkhoff ergodic theorem applied to $\chi_{F'}$ combined with Equation (4) then implies that $\mu_i(F') \geq 1/2$. Therefore, $\mu(F') \geq 1/2$, which completes the proof. □

Corollary 8.4 *Under the assumptions of Lemma 8.3, μ is a homogeneous, W-ergodic probability measure on G/Γ. Moreover, μ is ergodic with respect to any nontrivial subgroup generated by unipotent elements of its stabilizer in G.*

Proof Let S denote the stabilizer of μ in G. Then μ is not the zero measure by Lemma 8.3, thus [49, Cor. 1.3] shows that μ is a homogeneous, S^+-ergodic probability measure on G/Γ. Since any nontrivial subgroup generated by unipotents in S is a noncompact subgroup of S^+, μ is ergodic with respect to any such subgroup by the Howe–Moore theorem. In particular, this applies to W, as $W \leq S$ because the set of W-invariant measures is weak-$*$ closed. □

We now apply [49, Thm. 1.1] to understand the relationship between the support of the μ_i and the support of μ.

Theorem 8.5 *Let $W < G$ be a closed, connected, almost simple subgroup generated by unipotent elements and $\{\mu_i\}$ be a sequence of homogeneous,*

W-ergodic measures on G/Γ that weak-$*$ converges to μ. Then there exist a sequence of elements $\{g_i\}$ in G and a natural number i_0 such that for all $i \geq i_0$, the measures $g_i\mu$ are homogeneous, W-ergodic probability measures on G/Γ with $\mathrm{supp}(\mu_i) \subseteq g_i\,\mathrm{supp}(\mu)$.

Proof We fix unipotent subgroups $U_1, \ldots, U_s$ that generate W. By the Howe–Moore theorem, μ_i is U_j-ergodic for every $i \in \mathbb{N}$ and every $1 \leq j \leq s$. Therefore, for each i, the set of points for which μ_i is U_j-generic for all $1 \leq j \leq s$ is of full measure and hence is dense in $\mathrm{supp}(\mu_i)$.

By Corollary 8.4, μ is a nonzero homogeneous W-ergodic probability measure and $\mathrm{supp}(\mu)$ is a nonempty, closed homogeneous subspace of G/Γ. Fix a point x_∞ in $\mathrm{supp}(\mu)$. Then by the above we can find a sequence of points $\{x_i\}$ converging to x_∞ such that, for each i, $x_i \in \mathrm{supp}(\mu_i)$ and x_i is a U_j-generic point for every $1 \leq j \leq s$. We also fix a sequence $\{g_i\}$ of elements of G converging to the identity such that $g_i x_\infty = x_i$ for each i.

Applying [49, Thm. 1.1] to each unipotent subgroup U_j, we find a natural number i_j such that for all $i \geq i_j$ the measure μ is $U_j^{g_i^{-1}}$-invariant and $\mathrm{supp}(\mu_i) \subseteq g_i\,\mathrm{supp}(\mu)$. Let $i_0 = \max\{i_1, \ldots, i_s\}$. We now have that for every $i \geq i_0$, $\mathrm{supp}(\mu_i) \subseteq g_i\,\mathrm{supp}(\mu)$ and we are left to show that for every such i, the measure $g_i\mu$, which is clearly a homogeneous probability measure, is in fact W-ergodic. Equivalently, we are left to show that for every $i \geq i_0$, the measure μ is $W^{g_i^{-1}}$-ergodic.

Fix $i \geq i_0$. We have that μ is $U_j^{g_i^{-1}}$-invariant for every $1 \leq j \leq s$. We conclude that μ is $W^{g_i^{-1}}$-invariant, as $W^{g_i^{-1}}$ is the group generated by the unipotent subgroups $U_j^{g_i^{-1}}$, $1 \leq j \leq s$. It follows from Corollary 8.4 that indeed μ is $W^{g_i^{-1}}$-ergodic, which completes the proof of the theorem. □

We finally have all of the necessary ingredients to prove Proposition 8.1, which has an essentially identical proof to that of [2, Prop. 3.1]. As remarked immediately following its statement, it suffices to show that (1) and (3) are equivalent.

Proof of Proposition 8.1 We first prove that (3) implies (1). Recall that $\{\mu_i\}$ is a sequence of homogeneous, W-ergodic measures with weak-$*$ limit μ, which is Haar measure on G/Γ. Then $\pi_*\mu$ is the volume measure on M and hence $\mathrm{supp}(\pi_*\mu) = M$. As the measures μ_i are homogeneous, $\pi_*\mu_i$ is a constant multiple of the volume form on some closed totally geodesic subspace of M by Lemma 8.2(1), therefore $\mathrm{supp}(\pi_*\mu_i)$ is contained in some closed maximal geodesic subspace Z_i of M. Then $M = \mathrm{supp}(\pi_*\mu)$ is contained in closure of $\cup Z_i$, and we conclude that there must be infinitely many distinct Z_i in this union. This implies (1).

Finally, we show that (1) implies (3). Assume there are infinitely many distinct closed maximal totally geodesic subspaces $\{Z_i\}$ of M. By Lemma 8.2(3), for every i there exists a standard subgroup $W_i \leq G$, a homogeneous W_i-ergodic measure μ_i on G/Γ, and an element $h_i \in G$ such that $Z_i = \pi(W_i h_i \Gamma/\Gamma) = \pi(X_i)$, where $X_i = \mathrm{supp}(\mu_i)$. Since the collection of standard subgroups of G is finite, we may and do pass to a subsequence for which the subgroups W_i all coincide. We denote this common standard subgroup by W. Upon passing to a further subsequence, which we still denote by $\{\mu_i\}$, we assume that the μ_i weak-$*$ converge to a probability measure μ that is homogeneous and W-ergodic by Corollary 8.4. The proof will be complete once we show that μ is Haar measure on G/Γ.

For a contradiction, assume μ is not Haar measure. Since μ is homogeneous, $X_\infty = \mathrm{supp}(\mu)$ is a closed homogeneous subspace $Sh\Gamma/\Gamma$ of G/Γ, where $h \in G$ and S is the stabilizer of μ. By the assumption that μ is not Haar measure on G/Γ, $S < G$ is a proper subgroup. By Theorem 8.5, there exists a sequence of elements $\{g_i\}$ in G and a natural number i_0 such that for all $i \geq i_0$, the measures $g_i\mu$ are homogeneous, W-ergodic probability measures on G/Γ with $\mathrm{supp}(\mu_i) \subseteq g_i\,\mathrm{supp}(\mu)$. Upon passing again to a subsequence, we assume that $i_0 = 1$, thus for every i, $X_i \subset g_i X_\infty$ and $g_i\mu$ is a homogeneous, W-ergodic probability measure on G/Γ.

We will find a contradiction by showing that the spaces Z_i all coincide with $\pi(X_\infty)$, contradicting the assumption that they are all distinct. From now on we fix a natural number i and will argue that $Z_i = \pi(X_\infty)$.

By Lemma 8.2(1), $\pi(g_i X_\infty)$ is a totally geodesic subspace of M, and by Lemma 8.2(2) it is a proper subspace of M, since the stabilizer of $g_i\mu$, namely S^{g_i}, is a proper subgroup of G. From $X_i \subseteq g_i X_\infty$ we have that $Z_i = \pi(X_i) \subseteq \pi(g_i X_\infty)$ and, by the maximality assumption on the totally geodesic subspace Z_i, we conclude that $Z_i = \pi(g_i X_\infty)$. Since $g_i\mu$ is W-invariant and its stabilizer is S^{g_i}, we have that $W \leq S^{g_i}$. We also have $\pi(S^{g_i} g_i h\Gamma/\Gamma) = \pi(W h_i \Gamma/\Gamma)$, by the sequence of equations

$$\pi(S^{g_i} g_i h\Gamma/\Gamma) = \pi(g_i Sh\Gamma/\Gamma) = \pi(g_i X_\infty) = Z_i = \pi(W h_i \Gamma/\Gamma).$$

Lemma 8.2(2) implies that $W \leq S^{g_i} \leq N$, where N is the normalizer of W.

It follows from Lemma 2.7(3) that $W = (S^{g_i})^+$. As $W \leq S$, we also have that $W^{g_i} \leq (S^{g_i})^+ = W$, hence $W^{g_i} = W$ and in particular, $g_i \in N$. Using Lemma 8.2(2) again, we conclude that $N g_i h\Gamma/\Gamma$ is a closed homogeneous subspace of G/Γ such that $\pi(S^{g_i} g_i h\Gamma/\Gamma) = \pi(N g_i h\Gamma/\Gamma)$. Conjugating by $g^{-1} \in N$, the inclusion $W \leq S^{g_i} \leq N$ implies that $W \leq S \leq N$ and another application of Lemma 8.2(2) gives $\pi(X_\infty) = \pi(Sh\Gamma/\Gamma) = \pi(Nh\Gamma/\Gamma)$.

Therefore,

$$\begin{aligned} Z_i &= \pi(g_i X_\infty) = \pi(S^{g_i} g_i h\Gamma/\Gamma) \\ &= \pi(N g_i h\Gamma/\Gamma) = \pi(Nh\Gamma/\Gamma) = \pi(X_\infty). \end{aligned}$$

Thus $Z_i = \pi(X_\infty)$, and this completes the proof that μ is Haar measure, hence we have shown that (1) implies (3). □

Remark 8.6 It is also worthy of mention here that there has been considerable previous work on equidistribution in the context of Shimura varieties and special subvarieties. For example, see deep work of Clozel–Ullmo on equidistribution for strongly special subvarieties [16]. Also see [37,46,64,65,67] for subsequent results in this direction.

8.2 Setup for the proof of Theorem 1.1

The purpose of this section is to collect the necessary final setup so that we can apply Theorem 1.3 in the proof of Theorem 1.1, which is given in §8.3.

Let $G = \mathrm{SU}(n, 1)$, and suppose that $\Gamma < G$ is a lattice such that the complex hyperbolic orbifold $K\backslash G/\Gamma$ contains infinitely many maximal geodesic subspaces. Pass to a subsequence so that they are all either real or complex hyperbolic subspaces of the same type. By Lemma 8.2(3), there is a fixed standard subgroup $W < G$ with normalizer N and elements $g_i \in G$ so that

$$\Delta_i = N^{g_i^{-1}} \cap \Gamma,$$

is a lattice in $N^{g_i^{-1}}$ associated with the ith maximal geodesic subspace. Specifically, Δ_i acts on the totally geodesic subspace $K\backslash KNg_i^{-1}$ of $K\backslash G$ with finite covolume. We will need the following general result.

Proposition 8.7 (Cf. Prop. 3.3 [2]) *In addition to the above assumptions, suppose that k is a local field and $\mathbf{H}$ is a connected adjoint k-algebraic group. If $\rho : \Gamma \to \mathbf{H}(k)$ is a representation so that $\rho(\Delta_i)$ has proper Zariski closure in $\mathbf{H}(k)$ for infinitely many i, then there is a k-vector space V of dimension at least two, an irreducible representation of $\mathbf{H}$ on V, and a W-invariant, ergodic measure ν on the bundle $(G \times \mathbb{P}(V))/\Gamma$ that projects to Haar measure on G/Γ.*

Proof The proof is almost exactly the same as [2, Prop. 3.3], so we only sketch the proof. The assumption on $\rho(\Delta_i)$ implies that we can pass to a subsequence to assume that the Zariski closures of the $\rho(\Delta_i)$ are all contained in proper, nontrivial, k-algebraic subgroups $\mathbf{J}_i < \mathbf{H}$ all of which have the same dimension $d > 0$. Let $\mathfrak{h}$ be the Lie algebra of H over k and $\mathfrak{j}_i \subset \mathfrak{h}$ the subalgebra associated

with $\mathbf{J}_i$. We then take V to be the dth exterior power of $\mathfrak{h}$, so each $\mathfrak{j}_i$ defines a point in $\mathbb{P}(V)$.

Let S_i be the closed subgroup of $N^{g_i^{-1}}$ determined by Lemma 8.2(3). Then Δ_i is a lattice in S_i. Consequently, for each i we obtain a measurable section $S_i/\Delta_i \to (G \times \mathbb{P}(V))/\Gamma$. Let ν_i be the push-forward of Haar measure on S_i/Δ_i and ν be an ergodic component of the weak-$*$ limit of the ν_i. Ratner's theorem, the assumption that the geodesic submanifolds are distinct and maximal, and Proposition 8.1 implies that ν is a W-invariant, W-ergodic measure that projects to Haar measure on G/Γ. □

8.3 The proof of Theorem 1.1

Suppose that $\Gamma < \mathrm{SU}(n, 1)$ is a lattice so that $\mathbb{B}^n/\Gamma$ contains infinitely many maximal geodesic subspaces. Let ℓ be the adjoint trace field of Γ and $\mathbf{H}$ be the connected adjoint algebraic group associated with Γ as in Sect. 7. Given a place v of ℓ, let ℓ_v be the completion of ℓ at this place and $\rho_v : \Gamma \to \mathbf{H}(\ell_v) = H_v$ be the natural inclusion. We have a place v_0 so that $H_{v_0} \cong \mathrm{PU}(n, 1)$ and $\rho_{v_0}(\Gamma)$ is the lattice embedding. To prove that Γ is arithmetic we must show that $\rho_v(\Gamma)$ is precompact for all $v \neq v_0$.

Assume $\rho_v(\Gamma)$ is not precompact for some $v \neq v_0$. Let $k = \ell_v$, and W and $\{\Delta_i\}$ be as in Sect. 8.2. We clearly have that $\rho_v(\Delta_i)$ has proper Zariski closure in H_v. Then Proposition 8.7 applies to produce a vector space V as in the conclusion of the proposition and a W-invariant measure ν on $(G \times \mathbb{P}(V))/\Gamma$ that projects to Haar measure on G/Γ.

Then Theorem 1.5 along with Propositions 3.2 and 3.4 imply that either the pair (k, H_v) is compatible with G, or $k = \mathbb{R}$ and $H_v \cong \mathrm{PU}(n, 1)$. Since $\rho_v(\Gamma)$ is unbounded, Theorem 1.3 applies and ρ_v extends to a continuous homomorphism $\widehat{\rho}_v : G \to H_v$, and this is a contradiction as explained in [2, Section 3.2]. Therefore $\rho_v(\Gamma)$ must be precompact for all $v \neq v_0$, which proves that Γ is arithmetic. □

9 Examples, other results, and final comments

9.1 Submanifolds of arithmetic manifolds

In this section, we briefly give three examples exhibiting some of the possibilities for geodesic submanifolds of arithmetic quotients of $\mathbb{B}^n$.

The general construction is as follows. See [60]. Let ℓ be a totally real number field and ℓ' a totally imaginary quadratic extension of ℓ. Suppose that D is a central simple division algebra over ℓ' of degree d admitting an involution τ of *second kind*, i.e., so that the restriction of τ to ℓ' is the nontrivial Galois

involution of ℓ'/ℓ. For $r \geq 1$, τ extends to an anti-involution of the matrix algebra $\mathrm{M}_r(D)$ by τ-conjugate transposition, denoted $x \mapsto x^*$. Note that under any embedding of ℓ' in $\mathbb{C}$, this involution extends to complex conjugate transposition on $\mathrm{M}_r(D) \otimes_{\ell'} \mathbb{C} \cong \mathrm{M}_{rd}(\mathbb{C})$.

An element $h \in \mathrm{M}_r(D)$ is called *τ-hermitian* if $h^* = h$, and then we can define the ℓ-algebraic unitary group $\mathbf{G}$ with ℓ-points

$$\mathbf{G}(\ell) = \{x \in \mathrm{SL}_r(D) \;\; x^*hx = h\}.$$

Choosing a maximal order $\mathcal{O}$ of D, we obtain an arithmetic group

$$\Gamma^1_{\mathcal{O}} = \{x \in \mathrm{SL}_r(\mathcal{O}) \;\; x^*hx = h\} < \mathbf{G}(\ell).$$

Let $n = rd - 1$. Choosing D and h so that

$$\mathrm{Res}_{\ell/\mathbb{Q}}(\mathbf{G})(\mathbb{R}) \cong \mathrm{SU}(n, 1) \times \mathrm{SU}(n+1)^{[\ell:\mathbb{Q}]-1},$$

one has that the projection of $\Gamma^1_{\mathcal{O}}$ to $\mathrm{SU}(n, 1)$ is an arithmetic lattice. For any $n \geq 2$, all arithmetic subgroups of $\mathrm{SU}(n, 1)$ are commensurable with some such $\Gamma^1_{\mathcal{O}}$.

Example 9.1 When $d = 1$, we have $D = \ell$. Then h is a τ-hermitian form on an ℓ'-vector space V of dimension $n+1$, $\mathcal{O}$ is the ring of integers of ℓ', and $\Gamma^1_{\mathcal{O}}$ is sometimes called an arithmetic lattice of *simplest type*. For concreteness, we take $\ell = \mathbb{Q}$, $\ell' = \mathbb{Q}(\alpha)$ with $\alpha^2 = -1$, and h to be the hermitian form fixed in Sect. 2.1. Then

$$\Gamma^1_{\mathcal{O}} = \{g \in \mathrm{SL}_{n+1}(\mathbb{Z}[\alpha]) \;\; g^*hg = h\},$$

is the subgroup of $\mathrm{SL}_{n+1}(\mathbb{Z}[\alpha])$ preserving the form h and $\mathbf{G}$ is a $\mathbb{Q}$-algebraic group with $\mathbf{G}(\mathbb{R}) \cong \mathrm{SU}(n, 1)$.

We claim that $\mathbb{B}^n/\Gamma^1_{\mathcal{O}}$ contains all possible types of geodesic submanifolds of a complex hyperbolic n-manifold. Indeed, let $\{e_i\}$ be the standard basis for our vector space $V \cong \mathbb{C}^{n+1}$. Restricting the form to the span of $\{e_1, \ldots, e_m, e_{n+1}\}$ visibly gives an arithmetic subgroup $\Lambda^1_{\mathcal{O}} < \mathrm{SU}(m, 1)$ contained in $\Gamma^1_{\mathcal{O}}$, where $\mathrm{SU}(m, 1)$ denotes the standard subgroup in the notation of Sect. 2.1. This leads to properly immersed totally geodesic complex hyperbolic m-submanifolds for all m. Considering the real span of this subspace instead, the restriction of h now defines a quadratic form on $\mathbb{R}^{n+1}$ stabilized by the standard $\mathrm{SO}_0(m, 1)$ subgroup, hence one similarly finds real hyperbolic submanifolds of every real dimension between 2 and n.

Example 9.2 At the other extreme, assume that $d = n + 1$ is prime, so $\mathbf{G}(\ell)$ is a subgroup of the group $\mathrm{SL}_1(D)$ of units of D with reduced norm 1. Fix

a maximal order $\mathcal{O}$ of D and consider the arithmetic group $\Gamma^1_{\mathcal{O}} < \mathrm{SU}(n,1)$. For example, in the case $n = 2$ all fake projective planes arise from this construction [14].

Were $\mathbb{B}^n/\Gamma^1_{\mathcal{O}}$ to contain a proper geodesic subspace that is complex hyperbolic, then we would obtain an injection $\mathrm{M}_r(D') \hookrightarrow D$, where D' is a central simple ℓ'_0-division algebra of degree d' for some totally complex subfield ℓ'_0 of ℓ' whose intersection ℓ_0 with ℓ is totally real. Moreover, rd' divides $n+1$.

Since $n+1$ is prime, we claim that $r = 1$ and $d' = n+1$. Indeed, the case $r = d$ and $d' = 1$ is impossible since the algebra $\mathrm{M}_{n+1}(\ell'_0)$ contains subalgebras of degree $1 < e < n+1$. After taking the tensor product with ℓ', this contradicts the fact that D has prime degree and hence contains no such subalgebras. Now we rule out the case that ℓ'_0 is a proper subfield of ℓ'. To see this, note that in this case the ℓ-algebraic unitary group $\mathbf{G}$ associated with D now contains $\mathbf{H} \otimes_{\ell_0} \ell$, where $\mathbf{H}$ is the ℓ_0-algebraic unitary group associated with D'. However, $\mathbf{H}$ is noncompact at exactly one real place of ℓ_0 and $\mathbf{G}$ is consequently noncompact at exactly $[\ell : \ell_0]$ real places of ℓ. Since $\mathbf{G}$ is noncompact at exactly one real place of ℓ, we have that $\ell = \ell_0$ and $D = D'$.

It follows that $\mathbb{B}^n/\Gamma^1_{\mathcal{O}}$ contains no complex hyperbolic geodesic subspaces. A similar argument shows that $\mathbb{B}^n/\Gamma^1_{\mathcal{O}}$ also contains no real hyperbolic subspaces of any dimension $m \geq 2$. In other words, the only properly immersed geodesic subspaces of $\mathbb{B}^n/\Gamma^1_{\mathcal{O}}$ are the closed geodesics.

Remark 9.3 We leave it to the reader to show that the case $\mathrm{M}_r(D)$ for $r > 1$ and D of possibly composite degree $d > 1$ gives intermediate examples living between the previous two.

9.2 Geodesic submanifolds and the intersection pairing

In this section, we give some background on the algebraic and complex geometry of complex hyperbolic manifolds and explain the proof of Theorem 1.6. For simplicity we restrict to the case complex dimension two, then give references for the analogous results in higher dimension at the end of the section.

If M is a closed complex hyperbolic 2-manifold, then the associated complex hyperbolic manifold M is a smooth projective surface of general type whose Chern numbers satisfy $c_1^2(M) = 3\,c_2(M)$ [27]. We recall that $c_1^2(M)$ is the self-intersection of the canonical divisor $K_M \in H_2(M)$ and $c_2(M)$ is the Euler number of M. Moreover, Yau's famous solution to the Calabi conjecture says that this equality of Chern numbers holds for surfaces of general type if and only if $M = \mathbb{B}^2/\Gamma$ for some torsion-free cocompact lattice Γ in $\mathrm{PU}(2,1)$, where $\mathbb{B}^2$ denotes the unit ball in $\mathbb{C}^2$ with its complex hyperbolic metric [68]. It is a well-known consequence of Selberg's lemma that any cocompact lattice

$\Gamma < \mathrm{PU}(2,1)$ contains a finite index subgroup Γ' so that $\mathbb{B}^2/\Gamma'$ is a manifold, hence there is no loss of generality in assuming this is the case.

In the noncompact setting, as described in detail in the proof of Theorem 1.5, we can replace a lattice $\Gamma < \mathrm{PU}(2,1)$ with a subgroup of finite index so that $M = \mathbb{B}^2/\Gamma$ is a manifold of the form $X \smallsetminus D$, where X is a smooth projective surface and D is a certain smooth divisor on X. In this setting, one has the *logarithmic* canonical divisor $K_X + D \in H_2(X)$, where K_X is the canonical divisor on X. To unify the two cases, when M is compact we consider D to be the empty divisor and so $X = M$.

Proof of Theorem 1.6 Applying a very general theorem due to Borel [8, Thm. 3.7] to the compactification [1,45], an immersed totally geodesic complex hyperbolic submanifold of M determines an irreducible complex curve C_0 on M that compactifies to a curve C on X. Hirzebruch–Höfer relative proportionality [7, Section B.3] is precisely the statement that a projective curve C on X determines a totally geodesic subspace of M if and only if Equation (1) holds with respect to the divisor $K_X + D$. See [50, Thm. 0.1] for our statement, which is slightly more general than the original. In particular, the theorem is an immediate consequence of Theorem 1.1 and Corollary 1.2. □

Remark 9.4 Margulis asked whether arithmeticity is detected purely by the topology of the locally symmetric space. Theorem 1.6 implies that if M contains a totally geodesic curve, then arithmeticity of M is completely determined by the restriction of the intersection pairing on $H_2(M)$ to the curves on M. In other words, arithmeticity is detected by the topology of the underlying variety.

We now give references for where one can give a precise version of our statements in higher dimensions. For the equality of (logarithmic) Chern numbers that characterizes higher-dimensional complex hyperbolic manifolds, see [63]. A general version of relative proportionality that uniquely determines complex hyperbolic totally geodesic submanifolds was given by Müller-Stach, Viehweg, and Zuo. See Theorem 2.3 and Addendum 2.4 in [50].

This interpretation of our main results leads to the following question.

Question 9.5 *Can one classify the totally geodesic curves on nonarithmetic Deligne–Mostow orbifolds?*

This seems particularly approachable in dimension two in the sense that the underlying spaces for these orbifolds are closely related to blowups of the complex projective plane. In particular, one can connect geodesic curves to classical plane curves, where immersed geodesic curves will have singularities arising from self-intersections. One can then use relative proportionality to detect which curves are totally geodesic.

9.3 The proof of Theorem 1.7

Proof of Theorem 1.7 With notation as in the statement of the theorem, suppose that $f : M \to N$ is a surjective mapping so that $f(Z_i)$ is contained in a proper geodesic subspace of N for each i. By hypothesis, the induced map on fundamental groups induces a Zariski dense homomorphism $\rho : \Gamma \to \mathrm{PU}(m, 1)$, where Γ is the lattice in $\mathrm{SU}(n, 1)$ associated with M and $d = \dim(N)$. Let $\Lambda < \mathrm{PU}(d, 1)$ be the fundamental group of N.

We now proceed exactly as in Proposition 8.7. Let $\Delta_i < \Gamma$ be the subgroup associated with Z_i. The assumption on f implies that $\rho(\Delta_i)$ is not a Zariski dense subgroup of $\mathrm{PU}(d, 1)$. In particular, we pass to an infinite subsequence so that each each Z_i is a geodesic submanifold of M of the same type with associated with the standard subgroup W of G and $\rho(\Delta_i)$ has Zariski closure contained in some conjugate of a fixed proper, nontrivial, positive-dimensional $J < \mathrm{PU}(d, 1)$.

Let S_i be the closed subgroup of G associated with Δ_i by Lemma 8.2(3). The appropriate exterior power of the Lie algebra of H defines a vector space V such that for each i we can construct a measurable section from S_i/Δ_i to $(G \times \mathbb{P}(V))/\Gamma$ in order to build a W-invariant, W-ergodic measure on the bundle that projects to Haar measure on G/Γ. This and Proposition 3.4 allow us to apply Theorem 1.3(1) to conclude that ρ extends to a continuous homomorphism from $\mathrm{SU}(n, 1)$ to $\mathrm{PU}(d, 1)$. Therefore, $d = n$ and $\rho(\Gamma)$ is a lattice in $\mathrm{PU}(n, 1)$. Since Λ is also a lattice and $\rho(\Gamma) \leq \Lambda$, we see that $\rho(\Gamma)$ is a finite index subgroup. It follows that f is homotopic to a cover. This proves the theorem. □

9.4 The proof of Theorem 1.8

Proof of Theorem 1.8 To prove the theorem we must take care to differentiate between Tr Ad for $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{PO}_0(3, 1)$, since lattices in $\mathrm{PU}(3, 1)$ naturally contain lattices in $\mathrm{PO}_0(3, 1)$, but results on nonintegral traces for Kleinian groups typically are stated for $\mathrm{SL}_2(\mathbb{C})$. We denote the two adjoint representations by $\mathrm{Ad}_{\mathbb{C}}$ and $\mathrm{Ad}_{\mathbb{R}}$, respectively, as they denote the traces of the adjoint representation for $\mathfrak{sl}_2(\mathbb{C})$ considered as a complex (resp. real) vector space.

There are many known hyperbolic 3-manifolds M with nonintegral traces. See [39, Section 5.2.2] for a closed example and [15] for a hyperbolic link complement. However, in the literature this means that for a given lift of $\Gamma = \pi_1(M)$ from $\mathrm{PSL}_2(\mathbb{C})$ to $\mathrm{SL}_2(\mathbb{C})$ there is a $\gamma \in \Gamma$ with a nonintegral trace for the associated 2×2 matrix, i.e., that $\mathrm{Tr}(\gamma)$ is not an algebraic integer. The embedding of Γ in $\mathrm{PSL}_2(\mathbb{C})$ is the lattice embedding associated with the complete hyperbolic structure, and we can choose any lift to $\mathrm{SL}_2(\mathbb{C})$, as γ will

have nonintegral trace for any chosen lift. In what follows, we identify γ with a matrix in $\mathrm{SL}_2(\mathbb{C})$.

Let $\lambda^{\pm 1}$ be the eigenvalues of γ. Then $\lambda^{\pm 1}$ both have minimal polynomial $t^2 - \mathrm{Tr}(\gamma)t + 1$, and it follows that λ and λ^{-1} also are not algebraic integers. A direct calculation gives:

$$\begin{aligned}\mathrm{Tr}(\mathrm{Ad}_{\mathbb{C}}(\gamma)) &= \lambda^2 + \lambda^{-2} + 1\\ &= \mathrm{Tr}(\gamma)^2 - 1\end{aligned}$$

It follows easily from the assumption that $\mathrm{Tr}(\gamma)$ is not an algebraic integer that $\mathrm{Tr}(\mathrm{Ad}_{\mathbb{C}}(\gamma))$ is also not an algebraic integer.

Now, suppose that M is a geodesic submanifold of a complex hyperbolic n-manifold. When $n = 3$, the inclusion $\Gamma < \pi_1(M)$ gives

$$\Gamma < \mathrm{PO}_0(3, 1) < \mathrm{PU}(3, 1).$$

When $n > 3$, we instead have

$$\Gamma < \mathrm{SO}_0(3, 1) < \mathrm{PO}_0(n, 1) < \mathrm{PU}(n, 1),$$

where we have chosen a lift of Γ from $\mathrm{PO}_0(3, 1)$ to $\mathrm{SO}_0(3, 1)$. In each case, we want to show that $\mathrm{Tr}(\mathrm{Ad}_{\mathfrak{su}(n,1)}(\gamma))$ is not an algebraic integer. This will contradict Theorem 1.5(3) and complete the proof of the theorem.

For $n = 3$, we can conjugate and choose our hermitian form so that γ maps to the image in $\mathrm{PO}_0(3, 1) < \mathrm{PU}(3, 1)$ of the diagonal matrix with entries $\{\lambda^2, 1, 1, \lambda^{-2}\}$. A direct calculation from the root space decomposition for $\mathfrak{su}(3, 1)$ gives:

$$\begin{aligned}\mathrm{Tr}(\mathrm{Ad}_{\mathfrak{su}(3,1)}(\gamma)) &= 4\lambda^2 + 4\lambda^{-2} + \lambda^4 + \lambda^{-4} + 5\\ &= (\lambda^2 + \lambda^{-2})^2 + 4(\lambda^2 + \lambda^{-2}) + 3\end{aligned}$$

Since the above expression is a monic integral polynomial in $\lambda^2 + \lambda^{-2}$, which is not an algebraic integer, we conclude that $\mathrm{Tr}(\mathrm{Ad}_{\mathfrak{su}(3,1)}(\gamma))$ is also not an algebraic integer. We leave it to the reader to verify that the same proof works for higher n where 3 is now a sum of roots of unity and 4 is changed to $2(n-1)$. □

We note that Theorem 1.8 does not apply for large families of hyperbolic 3-manifolds. For example, if M is a finite-volume hyperbolic 3-manifold commensurable with a non-Haken hyperbolic 3-manifold, then $\pi_1(M)$ has all traces integral [39, Section 5.2]. More concretely, if $K \subset S^3$ is a *small knot*, then $\mathrm{Tr}(\mathrm{Ad}(\gamma))$ is an algebraic integer for all $\gamma \in \pi_1(S^3 \smallsetminus K)$. The figure-eight

knot is the only knot whose complement is arithmetic [57], and one can show that it is commensurable with a totally geodesic submanifold of a complex hyperbolic manifold arising from the construction in Example 9.1. It follows that if K is any other hyperbolic knot and M is a complex hyperbolic manifold containing $S^3 \smallsetminus K$ as a geodesic submanifold, then M is necessarily nonarithmetic. We know of no such example.

Question 9.6 *Which knot (or link) complements are isometric to a totally geodesic submanifold of a complex hyperbolic n-manifold for some $n \geq 3$? Are there combinatorial/diagrammatic obstructions to a knot complement being a geodesic submanifold of a complex hyperbolic manifold?*

There are only two known commensurability classes of nonarithmetic lattices in PU(3, 1) [20]. The known geometric constructions of orbifolds in each commensurability class might allow one to at least find some examples. Of course, Theorem 1.1 implies that only finitely many commensurability classes of knot and link complements could be associated with each commensurability class of nonarithmetic lattices in PU(3, 1).

Acknowledgements The authors thank Domingo Toledo for describing how Simpson's work leads to Theorem 1.5 and Olivier Biquard for correspondence about the nonuniform case. We thank Bruno Klingler for explaining the relationship to his conjecture discussed in the introduction, Ian Agol for pointing out Theorem 1.8 and Venkataramana, Nicolas Bergeron, and Emmanuel Ullmo for insightful conversations. We also thank Jean Lécureux and Beatrice Pozzetti for conversations related to the proof of Theorem 1.3(2). It is a pleasure to thank Baldi and Ullmo for having explained their approach to us and for various conversations around Hodge Theory. We are grateful to the referees for the constructive comments and recommendations which improved the quality of our exposition.

Funding Bader was supported by the ISF Moked 713510 grant number 2919/19. Fisher was supported by NSF DMS-1906107. Miller was supported by NSF DMS-2005438/2300370. Stover was supported by Grant Number 523197 from the Simons Foundation/SFARI and Grants DMS-1906088, DMS-2203555 from the National Science Foundation.

References

1. Ash, A., Mumford, D., Rapoport, M., Tai, Y.-S.: Smooth Compactifications of Locally Symmetric Varieties. Cambridge Mathematical Library, 2nd edn. Cambridge University Press, Cambridge (2010)
2. Bader, U., Fisher, D., Miller, N., Stover, M.: Arithmeticity, superrigidity, and totally geodesic submanifolds. Ann. Math. (2) **193**(3), 837–861 (2021)
3. Bader, U., Furman, A.: Super-rigidity and non-linearity for lattices in products. Compos. Math. **156**(1), 158–178 (2020)
4. Bader, U., Furman, A.: An extension of Margulis's superrigidity theorem. In: Dynamics. Geometry, Number Theory–The Impact of Margulis on Modern Mathematics, pp. 47–65. University Chicago Press, Chicago, IL (2022)
5. Baldi, G., Ullmo, E.: Special subvarieties of non-arithmetic ball quotients and Hodge Theory. Ann. of Math. (2) **197**(1), 159–220 (2023)

6. Ballmann, W., Gromov, M., Schroeder, V.: Manifolds of Nonpositive Curvature, Volume 61 of Progress in Mathematics. Birkhäuser, Basel (1985)
7. Barthel, G., Hirzebruch, F., Höfer, T.: Geradenkonfigurationen und Algebraische Flächen. Aspects of Mathematics, D4. Friedr Vieweg & Sohn, Braunschweig (1987)
8. Borel, A.: Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem. J. Differ. Geom. **6**, 543–560 (1972)
9. Borel, A., Tits, J.: Éléments unipotents et sous-groupes paraboliques de groupes réductifs. I. Invent. Math. **12**, 95–104 (1971)
10. Bourbaki, N.: Éléments de mathématique. XXVI. Groupes et algèbres de Lie Chapitre 1: Algèbres de Lie., vol. 1285. Actualités Sci. Indust., Hermann (1960)
11. Burger, M., Iozzi, A.: A measurable Cartan theorem and applications to deformation rigidity in complex hyperbolic geometry. Pure Appl. Math. Q. **4**(1, Special Issue: In honor of Grigory Margulis. Part 2), 181–202 (2008)
12. Calabi, E., Vesentini, E.: On compact, locally symmetric Kähler manifolds. Ann. Math. **2**(71), 472–507 (1960)
13. Cartan, E.: Sur le groupe de la géométrie hypersphérique. Comment. Math. Helv. **4**(1), 158–171 (1932)
14. Cartwright, D., Steger, T.: Enumeration of the 50 fake projective planes. C. R. Math. Acad. Sci. Paris **348**(1–2), 11–13 (2010)
15. Chesebro, E., DeBlois, J.: Algebraic invariants, mutation, and commensurability of link complements. Pacific J. Math. **267**(2), 341–398 (2014)
16. Clozel, L., Ullmo, E.: Équidistribution de sous-variétés spéciales. Ann. Math. (2) **161**(3), 1571–1588 (2005)
17. Corlette, K.: Flat G-bundles with canonical metrics. J. Differ. Geom. **28**(3), 361–382 (1988)
18. Dani, S.G., Margulis, G.A.: Limit distributions of orbits of unipotent flows and values of quadratic forms. In: Gel'fand, I.M. (ed.) Advances in Soviet Mathematics, vol. 16, pp. 91–137. American Mathematical Society, Providence (1993)
19. Delp, K., Hoffoss, D., Manning, J.: Problems in Groups, Geometry, and Three-Manifolds. Preprint: arXiv:1512.04620
20. Deraux, M.: A new nonarithmetic lattice in PU(3, 1). Algebr. Geom. Topol. **20**(2), 925–963 (2020)
21. Deraux, M., Parker, J.R., Paupert, J.: New nonarithmetic complex hyperbolic lattices II. Michigan Math. J. **70**(1), 135–205 (2021)
22. Esnault, H., Groechenig, M.: Cohomologically rigid local systems and integrality. Selecta Math. (N.S.) **24**(5), 4279–4292 (2018)
23. Gelander, T., Levit, A.: Counting commensurability classes of hyperbolic manifolds. Geom. Funct. Anal. **24**(5), 1431–1447 (2014)
24. Goldman, W.M.: Complex Hyperbolic Geometry. Oxford Mathematical Monographs, Oxford University Press, Oxford (1999)
25. Gromov, M., Piatetski-Shapiro, I.: Nonarithmetic groups in lobachevsky spaces. Inst. Hautes Études Sci. Publ. Math. **66**, 93–103 (1988)
26. Helgason, S.: Differential Geometry, Lie Groups, and Symmetric Spaces, Volume 80 of Pure and Applied Mathematics. Academic Press, Cambridge (1978)
27. Hirzebruch, F.: Automorphe Formen und der Satz von Riemann-Roch. In: Symposium internacional de topología algebraica International symposium on algebraic topology, pp 129–144. Universidad Nacional Autónoma de México and UNESCO, (1958)
28. Holzapfel, R.-P.: Ball and Surface Arithmetics. Aspects of Mathematics, vol. E29. Friedr. Vieweg & Sohn, Braunschweig (1998)
29. Humphreys, J.E.: Linear Algebraic Groups. Graduate Texts in Mathematics, vol. 21. Springer-Verlag, New York (1975)
30. Kapovich, M.: A survey of complex hyperbolic Kleinian groups. In: In The Tradition of Thurston II. Geometry and Groups, pp. 7–51. Springer, Cham (2022)

31. Karpelevič, F.I.: Surfaces of transitivity of a semisimple subgroup of the group of motions of a symmetric space. Doklady Akad. Nauk SSSR (N.S.) **93**, 401–404 (1953)
32. Klingler, B.: Hodge loci and atypical intersections: conjectures. Preprint: arXiv:1711.09387
33. Klingler, B.: Sur la rigidité de certains groupes fondamentaux, l'arithméticité des réseaux hyperboliques complexes, et les "faux plans projectifs". Invent. Math. **153**(1), 105–143 (2003)
34. Klingler, B.: Symmetric differentials, Kähler groups and ball quotients. Invent. Math. **192**(2), 257–286 (2013)
35. Knapp, A.W.: Lie Groups Beyond an Introduction, Volume 140 of Progress in Mathematics, 2nd edn. Birkhäuser, Boston (2002)
36. Koziarz, V., Maubon, J.: Maximal representations of uniform complex hyperbolic lattices. Ann. Math. (2) **185**(2), 493–540 (2017)
37. Koziarz, V., Maubon, J.: Finiteness of totally geodesic exceptional divisors in Hermitian locally symmetric spaces. Bull. Soc. Math. France **146**(4), 613–631 (2018)
38. Koziarz, V., Mok, N.: Nonexistence of holomorphic submersions between complex unit balls equivariant with respect to a lattice and their generalizations. Am. J. Math. **132**(5), 1347–1363 (2010)
39. Maclachlan, C., Reid, A.W.: The Arithmetic of Hyperbolic 3-Manifolds, Volume 219 of Graduate Texts in Mathematics. Springer-Verlag, New York (2003)
40. Margulis, G.: Problems and conjectures in rigidity theory. In: Mathematics: Frontiers and Perspectives, pp. 161–174. American Mathematical Society, Providence RI, (2000)
41. Margulis, G.A.: Arithmeticity of the irreducible lattices in the semisimple groups of rank greater than 1. Invent. Math. **76**(1), 93–120 (1984)
42. Margulis, G.A.: Discrete Subgroups of Semisimple Lie Groups, volume 17 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3). Springer-Verlag, New York (1991)
43. Mochizuki, T.: Kobayashi-Hitchin correspondence for tame harmonic bundles and an application. Astérisque, (309) (2006)
44. Mohammadi, A., Margulis, G.: Arithmeticity of hyperbolic 3-manifolds containing infinitely many totally geodesic surfaces. Ergod. Theory Dyn. Syst. **42**(3), 1188–1219 (2022)
45. Mok, N.: Projective algebraicity of minimal compactifications of complex-hyperbolic space forms of finite volume. In: Perspectives in Analysis, Geometry, and Topology, volume 296 of Progr. Math., pp. 331–354. Birkhäuser/Springer, New York, (2012)
46. Möller, M., Toledo, D.: Bounded negativity of self-intersection numbers of Shimura curves in Shimura surfaces. Algebra Number Theory **9**(4), 897–912 (2015)
47. Mostow, G.D.: Some new decomposition theorems for semi-simple groups. Mem. Am. Math. Soc. **14**, 31–54 (1955)
48. Mostow, G.D.: Strong Rigidity of Locally Symmetric Spaces. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, (1973). Annals of Mathematics Studies, No. 78
49. Mozes, S., Shah, N.: On the space of ergodic invariant measures of unipotent flows. Ergod. Theory Dyn. Syst. **15**(1), 149–159 (1995)
50. Müller-Stach, S., Viehweg, E., Zuo, K.: Relative proportionality for subvarieties of moduli spaces of $K3$ and abelian surfaces. Pure Appl. Math. Q. **5**(3, Special Issue: In honor of Friedrich Hirzebruch. Part 2), 1161–1199 (2009)
51. Paupert, J., Wells, J.: Hybrid lattices and thin subgroups of Picard modular groups. Topol. Appl. **269**, 106918 (2020)
52. Pozzetti, M.B.: Maximal representations of complex hyperbolic lattices into SU(M, N). Geom. Funct. Anal. **25**(4), 1290–1332 (2015)
53. Raghunathan, M.S.: Cohomology of arithmetic subgroups of algebraic groups. II. Ann. Math. **2**(87), 279–304 (1967)

54. Raghunathan, M.S.: Discrete Subgroups of Lie Groups. Springer-Verlag, New York-Heidelberg, (1972). Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68
55. Ramsay, A.: Virtual groups and group actions. Adv. Math. **6**(253–322), 1971 (1971)
56. Ratner, M.: Raghunathan's topological conjecture and distributions of unipotent flows. Duke Math. J. **63**(1), 235–280 (1991)
57. Reid, A.W.: Arithmeticity of knot complements. J. Lond. Math. Soc. (2) **43**(1), 171–184 (1991)
58. Simpson, C.: Higgs bundles and local systems. Inst. Hautes Études Sci. Publ. Math. **75**, 5–95 (1992)
59. Siu, Y.-T.: Some recent results in complex manifold theory related to vanishing theorems for the semipositive case. In: Workshop Bonn 1984 (Bonn, 1984), volume 1111 of Lecture Notes in Math., pp. 169–192. Springer (1985)
60. Tits, J.: Classification of algebraic semisimple groups. In: Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965), pp. 33–62. Amer. Math. Soc., (1966)
61. Toledo, D.: Representations of surface groups in complex hyperbolic space. J. Differ. Geom. **29**(1), 125–133 (1989)
62. Toledo, D.: Maps between complex hyperbolic surfaces. vol. 97, pp. 115–128. (2003). Special volume dedicated to the memory of Hanna Miriam Sandler (1960–1999)
63. Tsuji, H.: A characterization of ball quotients with smooth boundary. Duke Math. J. **57**(2), 537–553 (1988)
64. Ullmo, E.: Equidistribution de sous-variétés spéciales. II. J. Reine Angew. Math. **606**, 193–216 (2007)
65. Ullmo, E., Yafaev, A.: Galois orbits and equidistribution of special subvarieties: towards the André-Oort conjecture. Ann. Math. (2) **180**(3), 823–865 (2014)
66. Vinberg, E.B.: Rings of definition of dense subgroups of semisimple linear groups. Izv. Akad. Nauk SSSR Ser. Mat. **35**, 45–55 (1971)
67. Yafaev, A.: Galois orbits and equidistribution: Manin-Mumford and André-Oort. J. Théor. Nombres Bordeaux **21**(2), 493–502 (2009)
68. Yau, S.-T.: Calabi's conjecture and some new results in algebraic geometry. Proc. Nat. Acad. Sci. U.S.A. **74**(5), 1798–1799 (1977)
69. Zimmer, R.J.: Ergodic Theory and Semisimple Groups. Monographs in Mathematics, vol. 81. Birkhäuser Verlag, Basel (1984)