Inventiones mathematicae (2023) 235:441–487
https://doi.org/10.1007/s00222-023-01226-0

On the distribution of the Hodge locus

Gregorio Baldi[1] **· Bruno Klingler**[2] **· Emmanuel Ullmo**[1]

Received: 3 February 2023 / Accepted: 20 October 2023 / Published online: 3 November 2023

Abstract
Given a polarizable $\mathbb{Z}$-variation of Hodge structures $\mathbb{V}$ over a complex smooth quasi-projective base S, a classical result of Cattani, Deligne and Kaplan says that its Hodge locus (i.e. the locus where exceptional Hodge tensors appear) is a *countable* union of irreducible algebraic subvarieties of S, called the special subvarieties for $\mathbb{V}$. Our main result in this paper is that, if the level of $\mathbb{V}$ is at least 3, this Hodge locus is in fact a *finite* union of such special subvarieties (hence is algebraic), at least if we restrict ourselves to the Hodge locus factorwise of positive period dimension (Theorem 1.5). For instance the Hodge locus of positive period dimension of the universal family of degree d smooth hypersurfaces in $\mathbf{P}^{n+1}_{\mathbb{C}}$, $n \geq 3$, $d \geq 5$ and $(n, d) \neq (4, 5)$, is algebraic. On the other hand we prove that in level 1 or 2, the Hodge locus is analytically dense in S^{an} as soon as it contains one typical special subvariety. These results follow from a complete elucidation of the distribution in S of the special subvarieties in terms of typical/atypical intersections, with the exception of the atypical special subvarieties of zero period dimension.

Keywords Hodge theory and Mumford–Tate domains · Functional transcendence · Unlikely and likely intersections

Mathematics Subject Classification (2020) 14D07 · 14C30 · 14G35 · 22F30 · 03C64

Contents

✉ B. Klingler
bruno.klingler@hu-berlin.de

G. Baldi
baldi@ihes.fr

E. Ullmo
ullmo@ihes.fr

[1] I.H.E.S., Université Paris-Saclay, CNRS, Laboratoire Alexandre Grothendieck, 35 Route de Chartres, 91440 Bures-sur-Yvette, France

[2] Humboldt Universität zu Berlin, Berlin, Germany

1 Motivation and first main result

1.1 Hodge locus and special subvarieties

Let $f : X \to S$ be a smooth projective morphism of smooth irreducible complex quasi-projective varieties. Motivated by the study of the Hodge conjecture for the fibres of f, one defines the Hodge locus $\mathrm{HL}(S, f)$ as the locus of points $s \in S^{\mathrm{an}} = S(\mathbb{C})$ for which the Hodge structure $H^\bullet(X_s^{\mathrm{an}}, \mathbb{Z})_{\mathrm{prim}}/(\mathrm{torsion})$ admits more *Hodge tensors* than the primitive cohomology of the very general fibre. Here a Hodge class of a pure $\mathbb{Z}$-Hodge structure $V = (V_\mathbb{Z}, F^\bullet)$ is a class in $V_\mathbb{Q}$ whose image in $V_\mathbb{C}$ lies in the zeroth piece $F^0 V_\mathbb{C}$ of the Hodge filtration, or equivalently a morphism of Hodge structures $\mathbb{Q}(0) \to V_\mathbb{Q}$; and a Hodge tensor for V is a Hodge class in $V^\otimes := \bigoplus_{a,b\in\mathbb{N}} V^{\otimes a} \otimes (V^\vee)^{\otimes b}$, where $V^\vee$ denotes the Hodge structure dual to V. In this geometric case Weil [63] asked whether $\mathrm{HL}(S, f)$ is a countable union of closed algebraic subvarieties of S (he noticed that a positive answer follows easily from the rational Hodge conjecture).

More generally, let $\mathbb{V} := (\mathbb{V}_\mathbb{Z}, \mathcal{V}, F^\bullet, \nabla)$ be a polarizable variation of $\mathbb{Z}$-Hodge structures ($\mathbb{Z}$VHS) on S. Thus $\mathbb{V}_\mathbb{Z}$ is a finite rank locally free $\mathbb{Z}_{S^{\mathrm{an}}}$-local system on the complex manifold S^{an} and $(\mathcal{V}, F^\bullet, \nabla)$ is a filtered regular algebraic flat connection on S such that the analytification $(\mathcal{V}^{\mathrm{an}}, \nabla^{\mathrm{an}})$ is isomorphic to $\mathbb{V}_\mathbb{Z} \otimes_{\mathbb{Z}_{S^{\mathrm{an}}}} \mathcal{O}_{S^{\mathrm{an}}}$ endowed with the holomorphic flat connection defined by $\mathbb{V}_\mathbb{Z}$ (the filtration $F^\bullet$ is called the Hodge filtration). The Hodge locus $\mathrm{HL}(S, \mathbb{V}^\otimes)$ is the subset of points $s \in S^{\mathrm{an}}$ for which the Hodge structure $\mathbb{V}_s$ admits more Hodge tensors than the very general fiber $\mathbb{V}_{s'}$. We recover the geometric situation by considering the polarizable $\mathbb{Z}$VHS "of geometric origin"

$$\mathbb{V} = \left(\bigoplus_{k\in\mathbb{N}} (R^k f_*^{\mathrm{an}}\mathbb{Z})_{\mathrm{prim}}/(\mathrm{torsion}), \mathcal{V} = \bigoplus_{k\in\mathbb{N}} R^k f_* \Omega^\bullet_{X/S}, F^\bullet, \nabla\right).$$

In this case the Hodge filtration $F^\bullet$ on $\mathcal{V}$ is induced by the stupid filtration on the algebraic de Rham complex $\Omega^\bullet_{X/S}$ and ∇ is the Gauß-Manin connection.

Cattani, Deligne and Kaplan [15, Theorem 1.1] proved the following celebrated result, which in particular answers positively Weil's question (we also refer to [5, Theorem 1.6] for an alternative proof, using o-minimality):

Theorem 1.1 (Cattani-Deligne-Kaplan) *Let S be a smooth connected complex quasi-projective algebraic variety and $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS over S. Then $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is a countable union of closed irreducible algebraic subvarieties of S: the strict special subvarieties of S for $\mathbb{V}$.*

We refer to Definition 4.4 for a detailed description of the *special* subvarieties of S. Accordingly to such a definition, S itself is a special subvariety. We are therefore concerned with the distribution of the strict special subvarieties of S (i.e. the ones different from S).

1.2 Distribution of special subvarieties

After Theorem 1.1 one would like to understand the distribution in S of the (strict) special subvarieties for $\mathbb{V}$. For instance (see [34, Question 1.2]): are there any geometric constraints on the Zariski closure of $\mathrm{HL}(S, \mathbb{V}^{\otimes})$?

The fundamental tool for studying these questions is the holomorphic period map

$$\Phi : S^{\mathrm{an}} \to \Gamma\backslash D, \tag{1.1}$$

completely describing the $\mathbb{Z}$VHS $\mathbb{V}$ (see Sect. 4 for more details on what follows). Here $(\mathbf{G}, D)$ denotes the generic Hodge datum of $\mathbb{V}$ and $\Gamma\backslash D$ is the associated Hodge variety. The Mumford-Tate domain D decomposes as a product $D_1 \times \cdots \times D_k$, according to the decomposition of the adjoint group $\mathbf{G}^{\mathrm{ad}}$ into a product $\mathbf{G}_1 \times \cdots \times \mathbf{G}_k$ of simple factors (notice that some factors $\mathbf{G}_i$ may be $\mathbb{R}$-anisotropic). Replacing S by a finite étale cover and reordering the factors if necessary, the lattice $\Gamma \subset \mathbf{G}^{\mathrm{ad}}(\mathbb{R})^+$ decomposes as a direct product $\Gamma \cong \Gamma_1 \times \cdots \times \Gamma_r$, $r \leq k$, where $\Gamma_i \subset \mathbf{G}_i(\mathbb{R})^+$, $1 \leq i \leq r$, is an arithmetic lattice, Zariski-dense in $\mathbf{G}_i$. Writing $D' = D_{r+1} \times \cdots \times D_k$ for the product of factors where the monodromy is trivial (it contains in particular all the factors D_i for which $\mathbf{G}_i$ is $\mathbb{R}$-anisotropic), the period map is written

$$\Phi : S^{\mathrm{an}} \to \Gamma\backslash D \cong \Gamma_1\backslash D_1 \times \cdots \times \Gamma_r\backslash D_r \times D' \ , \tag{1.2}$$

and the projection of $\Phi(S^{\mathrm{an}})$ on D' is a point. For more details, see also *the structure theorem for VHS* from [28, (III.A.2)].

By enlarging S if necessary, one can assume without loss of generality that Φ is proper (see [30, Theorem 9.5]). The image $\Phi(Z^{\mathrm{an}})$ for any closed algebraic subvariety $Z \subset S$ is then a closed analytic subvariety of $\Gamma\backslash D$, by Remmert's proper mapping theorem. It is thus natural to distinguish between special subvarieties Z of *zero period dimension* (i.e. $\Phi(Z^{\mathrm{an}})$ is a point of $\Gamma\backslash D$), which are geometrically elusive; and those of *positive period dimension* (i.e. $\dim_{\mathbb{C}} \Phi(Z^{\mathrm{an}}) > 0$), that are susceptible of a variational study. Taking into account product phenomena, we are lead to:

Definition 1.2

(1) A subvariety Z of S is said of *positive period dimension for* $\mathbb{V}$ if $\Phi(Z^{\mathrm{an}})$ has positive dimension. If moreover the projection of $\Phi(Z^{\mathrm{an}})$ on each factor $\Gamma_i\backslash D_i$, $1 \leq i \leq r$, has positive dimension, then Z is said to be *factorwise of positive period dimension*.

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

Thus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})$ and the first inclusion is an equality if $\mathbf{G}^{\mathrm{ad}}$ is simple.

Recently Otwinowska and the second author [35] proved the following theorem (they work in the case where $\mathbf{G}^{\mathrm{ad}}$ is simple, but their proof adapts immediately to the general case):

Theorem 1.3 ([35, Theorem 1.5]) *Let* $\mathbb{V}$ *be a polarizable* $\mathbb{Z}$*VHS on a smooth connected complex quasi-projective variety* S. *Then either* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ *is an algebraic subvariety of* S (*not necessarily irreducible*) *or it is Zariski-dense in* S.

Remark 1.4 Saying that $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ is algebraic is equivalent to saying that the set of strict special subvarieties of S for $\mathbb{V}$ factorwise of positive period dimension has only finitely many maximal elements for the inclusion.

1.3 An algebraicity result

Theorem 1.3 has an important flaw: it does not provide any criterion for deciding which branch of the alternative holds true. Our most striking result in this paper proves that $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ is algebraic in most cases. A simple measure of the complexity of $\mathbb{V}$ is its level: roughly, the length of the Hodge filtration on the holomorphic tangent space of D. See Definition 4.15 for the precise definition in terms of the algebraic monodromy group of $\mathbb{V}$. While special subvarieties usually abound for $\mathbb{Z}$VHSs of level one (e.g. families of abelian varieties or families of K3 surfaces) and for some $\mathbb{Z}$VHS of level two (e.g Green's famous example of the Noether-Lefschetz locus for degree d ($d > 3$) surfaces in $\mathbf{P}^3_{\mathbb{C}}$, see [62, Proposition 5.20]), we show:

Theorem 1.5 *Let* $\mathbb{V}$ *be a polarizable* $\mathbb{Z}$*VHS on a smooth connected complex quasi-projective variety* S. *If* $\mathbb{V}$ *is of level at least* 3 *then* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ *is a finite union of maximal atypical special subvarieties* (*hence is algebraic*). *In particular, if moreover* $\mathbf{G}^{\mathrm{ad}}$ *is simple*, *then* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}}$ *is algebraic in* S.

As a simple geometric illustration of Theorem 1.5:

Corollary 1.6 *Let* $\mathbf{P}^{N(n,d)}_{\mathbb{C}}$ *be the projective space parametrising the hypersurfaces* X *of* $\mathbf{P}^{n+1}_{\mathbb{C}}$ *of degree* d (*where* $N(n,d) = \binom{n+d+1}{d} - 1$). *Let* $U_{n,d} \subset \mathbf{P}^{N(n,d)}_{\mathbb{C}}$ *be the*

Zariski-open subset parametrising the smooth hypersurfaces X and let $\mathbb{V} \to U_{n,d}$ be the $\mathbb{Z}$VHS corresponding to the primitive cohomology $H^n(X, \mathbb{Z})_{\mathrm{prim}}$.

If $n = 3$ and $d \geq 5$; $n = 4$ and $d \geq 6$; $n = 5, 6, 8$ and $d \geq 4$; and $n = 7$ or ≥ 9 and $d \geq 3$, then the level of $\mathbb{V} \to U_{n,d}$ is at least 3, and therefore $\mathrm{HL}(U_{n,d}, \mathbb{V}^{\otimes})_{\mathrm{pos}} \subset U_{n,d}$ is algebraic.

Let us also mention that Theorem 1.5, combined with the main result of [39], provides a nice first result in the direction of the so called *refined Bombieri-Lang conjecture*. See our note [8].

2 Special subvarieties as (a)typical intersection loci: conjectures

The heuristic behind the proof of Theorem 1.5 is based on a remarkable feature of Hodge theory: thanks to the existence of period maps, special subvarieties can also be defined as *intersection loci*. Indeed, a closed irreducible subvariety $Z \subset S$ is special for $\mathbb{V}$ (we will equivalently say that it is special for Φ) precisely if Z^{an} coincides with an analytic irreducible component $\Phi^{-1}(\Gamma'\backslash D')^0$ of $\Phi^{-1}(\Gamma'\backslash D')$, for $(\mathbf{G}', D') \subset (\mathbf{G}, D)$ the generic Hodge subdatum of Z and $\Gamma'\backslash D' \subset \Gamma\backslash D$ the associated Hodge subvariety.

Remark 2.1 The description of special subvarieties as intersection loci makes the study of the Hodge locus simpler than the study of the Tate locus (its analogue obtained when replacing $\mathbb{V}$ by a lisse ℓ-adic sheaf over a smooth variety S over a field of finite type and the Hodge tensors by the Tate tensors).

This suggests a fundamental dichotomy between *typical* and *atypical* intersections, which should govern whether or not the Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is algebraic, and in particular which branch of the alternative is satisfied in Theorem 1.3. Such a dichotomy was proposed for the first time by the second author in [34]:

Definition 2.2 Let $Z = \Phi^{-1}(\Gamma'\backslash D')^0 \subset S$ be a special subvariety for $\mathbb{V}$, with generic Hodge datum $(\mathbf{G}', D')$. It is said to be *atypical* if $\Phi(S^{\mathrm{an}})$ and $\Gamma'\backslash D'$ do not intersect generically along $\Phi(Z)$:

$$\mathrm{codim}_{\Gamma\backslash D}\, \Phi(Z^{\mathrm{an}}) < \mathrm{codim}_{\Gamma\backslash D}\, \Phi(S^{\mathrm{an}}) + \mathrm{codim}_{\Gamma\backslash D}\, \Gamma'\backslash D' \ , \tag{2.1}$$

or if Z is singular for $\mathbb{V}$ (meaning that $\Phi(Z^{\mathrm{an}})$ is contained in the singular locus of $\Phi(S^{\mathrm{an}})$). Otherwise it is said to be *typical*.

Remark 2.3 In Definition 2.2, deciding that if Z is singular for $\mathbb{V}$ then it is atypical for $\mathbb{V}$ hides the fact that the numerical condition (2.1) is too naive when Z is singular. This is the right convention for hoping Conjecture 2.7 below to be true. Notice that if we define the *singular locus of S for* $\mathbb{V}$ as the preimage $S_{\mathbb{V}}^{\mathrm{sing}}$ under Φ of the singular locus of the complex analytic variety $\Phi(S^{\mathrm{an}})$ (in particular any $Z \subset S$ special and singular for $\mathbb{V}$ is contained in $S_{\mathbb{V}}^{\mathrm{sing}}$), it follows from the definability of Φ in the o-minimal structure $\mathbb{R}_{\mathrm{an,exp}}$ [5, Theorem 1.3] that $S_{\mathbb{V}}^{\mathrm{sing}}$ is actually a closed (strict) algebraic subvariety of S (not necessarily irreducible), see [35, Sect. 8] for a proof. The precise structure of the singular special locus is described in Lemma 6.7.

Definition 2.4 The *atypical Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})$ (resp. the *typical Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})$) is the union of the atypical (resp. strict typical) special subvarieties of S for $\mathbb{V}$.

2.1 Conjectures

We expect the Hodge locus

$$\mathrm{HL}(S, \mathbb{V}^{\otimes}) = \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}} \bigcup \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$$

to satisfy the following two conjectures:

Conjecture 2.5 (Zilber–Pink conjecture for the atypical Hodge locus, strong version) *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on an irreducible smooth quasi-projective variety S. The atypical Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}}$ is a finite union of atypical special subvarieties of S for $\mathbb{V}$. Equivalently: the set of atypical special subvarieties of S for $\mathbb{V}$ has finitely many maximal elements for the inclusion.*

Remark 2.6 Conjecture 2.5 is stronger than the Zilber-Pink conjecture for $\mathbb{Z}$VHS originally proposed in [34, Conjecture 1.9]. Indeed, our Definition 2.2 of atypical special subvarieties is more general (and simpler) than the one considered in *op. cit.*. We refer to Sect. 5.3 for a detailed comparison with [34].

Conjecture 2.7 (Density of the typical Hodge locus) *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. If $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is not empty then it is dense (for the analytic topology) in S.*

Conjecture 2.5 and Conjecture 2.7 imply immediately the following, which clarifies the possible alternatives in Theorem 1.3:

Conjecture 2.8 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on an irreducible smooth quasi-projective variety S. If $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is empty then $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is algebraic; otherwise $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is analytically dense in S.*

In view of Conjecture 2.7 and Conjecture 2.8, we are led to the:

Question 2.9 Is there a simple combinatorial criterion on $(\mathbf{G}, D)$ for deciding whether $\mathrm{HL}(S, \mathbb{V})_{\mathrm{typ}}$ is empty (which would imply, following Conjecture 2.8, that $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is algebraic)?

3 Special subvarieties as (a)typical intersection loci: results and applications

In this paper we prove Conjecture 2.5 and Conjecture 2.7 (and thus Conjecture 2.8) except for the description of the atypical locus of zero period dimension, see Theorem 3.1 and Theorem 3.9 respectively. We also provide a powerful criterion for $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ to be empty, thus answering Question 2.9, see Theorem 3.3. Our Theorem 1.5 that "in most cases" $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ is algebraic then follows from Theorem 3.1 and Theorem 3.3. Let us notice that all these results are independent of Theorem 1.3, which we do not use and of which we prove a variant, see Corollary 3.12.

3.1 On the atypical Hodge locus

Our first main result establishes the *geometric part* of Conjecture 2.5: we prove that the maximal atypical special subvarieties of positive period dimension arise in a finite number of families whose geometry we control. We cannot say anything on the atypical locus of zero period dimension, for which different ideas are certainly needed.

Theorem 3.1 (Geometric Zilber–Pink) *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. Let Z be an irreducible component of the Zariski closure of* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}} := \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}} \cap \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}}$ *in S. Then*:

(a) *Either Z is a maximal atypical special subvariety*;
(b) *Or the generic adjoint Hodge datum $(\mathbf{G}_Z^{\mathrm{ad}}, D_{G_Z})$ decomposes as a non-trivial product $(\mathbf{G}', D') \times (\mathbf{G}'', D'')$, inducing (after replacing S by a finite étale cover if necessary)*

$$\Phi_{|Z^{\mathrm{an}}} = (\Phi', \Phi'') : Z^{\mathrm{an}} \to \Gamma_{\mathbf{G}_Z} \backslash D_{G_Z} = \Gamma' \backslash D' \times \Gamma'' \backslash D'' \subset \Gamma \backslash D \ ,$$

such that Z contains a Zariski-dense set of atypical special subvarieties for Φ'' of zero period dimension. Moreover Z is Hodge generic in a special subvariety

$$\Phi^{-1}(\Gamma_{\mathbf{G}_Z} \backslash D_{G_Z})^0$$

of S for Φ which is typical.

Remark 3.2 Conjecture 2.5, which also takes into account the atypical special subvarieties of zero period dimension, predicts that the branch (b) of the alternative in the conclusion of Theorem 3.1 never occurs.

In Sect. 3.5.1 and Sect. 3.5.2 we discuss special examples of Theorem 3.1 of particular interest.

3.2 A criterion for the typical Hodge locus to be empty

The following result answers Question 2.9:

Theorem 3.3 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and algebraic monodromy group $\mathbf{H}$. If $\mathbb{V}$ is of level at least 3 then $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}} = \emptyset$ (and thus $\mathrm{HL}(S, \mathbb{V}^{\otimes}) = \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{atyp}}$).*

In level 2 we prove that strict typical special subvarieties do satisfy some geometric constraints:

Proposition 3.4 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S of level 2. Suppose that the Lie algebra $\mathfrak{g}^{\mathrm{ad}}$ of its adjoint generic Mumford-Tate group is simple. If $Z \subset S$ is a typical special subvariety then its adjoint generic Mumford-Tate group $\mathbf{G}_Z^{\mathrm{ad}}$ is simple.*

3.3 On the algebraicity of the Hodge locus

As we will show in Sect. 8, Theorem 1.5 easily follows from Theorem 3.1 and Theorem 3.3. Notice that the full Conjecture 2.5, allied with Theorem 3.3, would imply in the same way:

Conjecture 3.5 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. If $\mathbb{V}$ is of level at least 3 then $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is algebraic.*

Remark 3.6 Notice that Conjecture 3.5 implies, if S is defined over $\overline{\mathbb{Q}}$, that S contains a Hodge generic $\overline{\mathbb{Q}}$-point, as predicted, at least if $\mathbb{V}$ is of geometric origin (in the sense of Sect. 1.1), by the conjecture that Hodge classes are absolute Hodge classes.

In level two, the typical locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is not necessarily empty, but, thanks to Proposition 3.4, we can still prove:

Theorem 3.7 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S with generic Mumford-Tate datum $(\mathbf{G}, D)$. Suppose that $\mathbb{V}$ has level 2 and that $\mathbf{G}^{\mathrm{ad}}$ is simple. Then $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}}$ is algebraic.*

Corollary 3.8 *Let $\mathbf{P}^N_{\mathbb{C}}$ be the projective space parametrising the hypersurfaces X of $\mathbf{P}^3_{\mathbb{C}}$ of degree d (with $N = \frac{(d+3)(d+2)(d+1)}{6} - 1$). Let $U_{2,d} \subset \mathbf{P}^N_{\mathbb{C}}$ be the Zariski-open subset parametrising the smooth hypersurfaces X and let $\mathbb{V} \to U_{2,d}$ be the $\mathbb{Z}$VHS corresponding to the primitive cohomology $H^2(X, \mathbb{Z})_{\mathrm{prim}}$. If $d \geq 5$ then $\mathrm{HL}(U_{2,d}, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}} \subset U_{2,d}$ is algebraic.*

3.4 On the typical Hodge locus in level one and two

Let us now turn to the typical Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$.

3.4.1 The "all or nothing" principle

In view of Theorem 3.3, $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is non-empty only if $\mathbb{V}$ is of level one or two. In that case its behaviour is predicted by Conjecture 2.7. In this direction we obtain our last main result:

Theorem 3.9 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. If the typical Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}}$ is non-empty then $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is analytically (hence Zariski) dense in S.*

Here and elsewhere, by analytically dense, we mean dense in every non-empty open subset of S (for the usual Euclidean topology).

Remark 3.10 Notice that, in Theorem 3.9, we also treat the typical Hodge locus of zero period dimension.

Remark 3.11 Theorem 3.9 is new even for S a subvariety of a Shimura variety. Its proof is inspired by the arguments of Colombo-Pirola [18], Izadi [32] and Chai [16] in that case.

In the same way that Conjecture 2.5 and Conjecture 2.7 imply Conjecture 2.8, we deduce from Theorem 3.1, Theorem 3.3 and Theorem 3.9 the following:

Corollary 3.12 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. Then either $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$ is a finite union of maximal atypical special subvarieties of S for $\mathbb{V}$, hence is algebraic; or the level of $\mathbb{V}$ is one or two, $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{typ}} \neq \emptyset$, and $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is analytically dense in S.*

Remark 3.13 Notice that Corollary 3.12 clarifies the alternative of Theorem 1.3, and strengthens it, as density is now in the analytic topology rather than in the (weaker) Zariski topology. However it does not recover exactly Theorem 1.3, as in Corollary 3.12 the density could come from typical varieties which are not factorwise of positive dimension.

3.4.2 Examples of dense typical Hodge locus in level 1

Given a $\mathbb{Z}$VHS on S of level one or two, can we deduce from the generic Mumford-Tate datum $(\mathbf{G}, D)$ whether or not the typical Hodge locus is dense in S? Here we discuss a simple example (related to Remark 3.11) where density holds and, in the next paragraph, use it to present an arithmetic application of independent interest.

Theorem 3.14 (Chai + ϵ) *Let $(\mathbf{G}, X)$ be a Shimura datum containing a Shimura sub-datum $(\mathbf{H}, X_H)$ such that X_H is one dimensional and the normaliser of H in G is H. Assume moreover that $\mathbf{G}$ is absolutely simple. Let $S \subset \Gamma\backslash X$ be an irreducible subvariety of codimension one. Then the Hodge locus of S is dense.*

Remark 3.15 Regarding the Shimura variety parametrising g-dimensional principally polarised abelian varieties $\mathcal{A}_g = \mathrm{Sh}_{\mathbf{Sp}_{2g}(\mathbb{Z})}(\mathbf{GSp}_{2g}, \mathbb{H}_g)$, we recall that its dimension is $\frac{g(g+1)}{2}$ and the biggest dimension of the (strict) special subvarieties is $\frac{g(g-1)}{2}+1$, realised for instance by $\mathcal{A}_{g-1} \times \mathcal{A}_1 \subset \mathcal{A}_g$. It follows that the typical locus of any subvariety of $\mathcal{A}_g$ of dimension smaller than $g-1$ is empty. Let $S \subset \mathcal{A}_g$, be a (closed, not necessarily smooth) subvariety of dimension q, one would expect that the typical Hodge locus is (analytically) dense in S if and only if $q \geq g-1$.

Remark 3.16 We do not address here the question of whether the typical Hodge locus is equidistributed in level 1 or 2. This interesting question has been investigated in more details for the typical Hodge locus of zero period dimension by Tayou and Tholozan in [55, 56]. In particular Theorem 3.14 is also a corollary of their work. See also the work of Koziarz and Maubon [38] and [41, Proposition 3.1].

3.4.3 An arithmetic application

In this section we present an arithmetic application of Theorem 3.14, that was our motivation for studying such a problem. Mumford [46] shows that there exist principally polarized abelian varieties X of dimension 4 having trivial endomorphism ring, that are not Hodge generic in $\mathcal{A}_4$ (they have an exceptional Hodge class in $H^4(X^2, \mathbb{Z})$). A question often attributed to Serre is to describe “as explicitly as possible” such abelian varieties *of Mumford’s type*. The most satisfying way would be to show the existence of a smooth projective curve over $\overline{\mathbb{Q}}$ of genus 4, whose Jacobian is of Mumford’s type. Gross [24, Problem 1] asked the weaker question over $\mathbb{C}$. See also a related question of Oort [47, Sect. 7] (which we will address at the very end of the paper, see indeed Remark 11.6). We also notice that, at the end of August 2021, F. Calegari has proposed a polymath project to find an explicit example of a Mumford 4-fold over $\mathbb{Q}$.

Theorem 3.17 *There exists a smooth projective curve $C/\overline{\mathbb{Q}}$ of genus* 4 *whose Jacobian has Mumford-Tate group isogenous to a $\mathbb{Q}$-form of the complex group* $\mathbb{G}_m \times \mathrm{SL}_2 \times \mathrm{SL}_2 \times \mathrm{SL}_2$.

Remark 3.18 Actually our proof shows the existence of infinitely many such curves. Finding explicit equations for such a curve remains an open problem.

Remark 3.19 A crucial input in the proof is the André–Oort conjecture for $\mathcal{A}_4$, as established in [59, Theorem 1.3]. For $g = 4$, [60, Theorem 5.1] and [49, Theorem 1.3] are actually enough, see Remark 11.5 for more details.

3.5 Complements and applications

We conclude this section by discussing two applications of Theorem 3.1 of particular interest.

3.5.1 The Shimura locus

Recall that a $\mathbb{Z}$VHS $\mathbb{V}$ is said *of Shimura type* if its generic Hodge datum $(\mathbf{G}, D)$ is a Shimura datum, see [34, Definition 3.7]. The $\mathbb{Z}$VHSs of Shimura type form the simplest class of $\mathbb{Z}$VHSs: the target $\Gamma\backslash D$ of the period map (1.1) is a Shimura variety (a quasi projective algebraic variety by [3]) and the period map $\Phi : S^{\mathrm{an}} \to \Gamma\backslash D$ is algebraic, see [11, Theorem 3.10].

Given $\mathbb{V}$ a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, a special subvariety $Z \subset S$ for $\mathbb{V}$ is said to be *of Shimura type* if $\mathbb{V}_{|Z}$ is of Shimura type, i.e. its generic Hodge datum $(\mathbf{G}', D')$ is a Shimura datum. In that case, it is said to have *dominant period map* if the algebraic period map $\Phi_{|Z^{\mathrm{an}}} : Z^{\mathrm{an}} \to \Gamma'\backslash D'$ is dominant. The André-Oort conjecture for $\mathbb{Z}$VHS, as formulated in [28, page 275] and [34, Conjecture 5.2], states that if the union of the special subvarieties of S for $\mathbb{V}$ which are of Shimura type with dominant period maps is Zariski-dense in S, then $(S, \mathbb{V})$ itself is of Shimura type with dominant period map. This conjecture easily follows from Conjecture 2.5, see [34, Sect. 5.2].

More generally let us define the *Shimura locus* of S for $\mathbb{V}$ as the union of the special subvarieties of S for $\mathbb{V}$ which are of Shimura type (but not necessarily with dominant period maps). In [34, Question 5.8], the second author asked the following question, which generalizes the André–Oort conjecture for $\mathbb{Z}$VHSs:

Question 3.20 Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. Suppose that the Shimura locus of S for $\mathbb{V}$ is Zariski-dense in S. Is it true that necessarily $\mathbb{V}$ is of Shimura type?

Thanks to Theorem 3.1, we answer affirmatively the geometric part of Question 3.20:

Corollary 3.21 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$, and assume that $\mathbf{G}^{\mathrm{der}}$ is the generic Mumford-Tate group of $\mathbb{V}$. Suppose that the Shimura locus of S for $\mathbb{V}$ of positive period dimension is Zariski-dense in S.*

If the adjoint group of the generic Mumford-Tate group $\mathbf{G}$ of $\mathbb{V}$ is simple then $(\mathbf{G}, D)$ is a Shimura datum.

In general, either $(\mathbf{G}, D)$ is a Shimura datum, or there exists a decomposition $\widehat{S}^{\mathrm{an}} \xrightarrow{(\Phi', \Phi'')} \Gamma'\backslash D' \times \Gamma''\backslash D'' \to \Gamma\backslash D$, where $\widehat{S} \to S$ is a finite surjective morphism; $\Gamma'\backslash D'$ is a Shimura variety; Φ' is dominant, and $\Gamma'\backslash D' \times \Gamma''\backslash D'' \to \Gamma\backslash D$, is a finite Hodge morphism; $\Phi(S^{\mathrm{an}})$ is dense in the image of $(\Gamma'\backslash D' \times \Phi''(S^{\mathrm{an}}))$ in $\Gamma\backslash D$; and the infinitely many Shimura subvarieties of $\Gamma\backslash D$ intersecting $\Phi(S^{\mathrm{an}})$ are coming from the image of

$$\text{Shimura sub-variety} \times \text{CM-point} \subset \Gamma'\backslash D' \times \Gamma''\backslash D''$$

in $\Gamma\backslash D$.

Remark 3.22 Notice that Corollary 3.21 implies the geometric part of the André–Oort conjecture for $\mathbb{Z}$VHS, which has the same conclusion but requires the sub-Shimura varieties to be fully contained in $\Phi(S^{an})$ rather than intersecting it in positive dimension. Such a statement has been recently proven independently of this paper and with different techniques by Chen, Richard and the third author [51, Theorem A.7].

Remark 3.23 Since the horizontal (in the sense of Sect. 5.3) tangent subbundle of $\Gamma\backslash D$ is locally homogeneous, any Hodge variety $\Gamma\backslash D$ containing one positive dimensional Shimura subvariety (induced by a Hodge-subdatum) contains infinitely many of such (compare with Remark 3.27 and [52, Lemma 3.2]). For a discussion regarding the horizontal tangent bundle we refer to the beginning of Sect. 5.3. The same applies to *CM-points*, that is points whose Mumford–Tate group is commutative, which are always analytically dense in $\Gamma\backslash D$.

3.5.2 The modular locus

Any $\mathbb{Z}$VHS $\mathbb{V}$ of Shimura type with dominant (algebraic) period map $\Phi : S \to \Gamma\backslash D$ endows S with a large collection of algebraic self-correspondences, coming from the Hecke correspondences on the Shimura variety $\Gamma\backslash D$ associated to elements $g \in \mathbf{G}(\mathbb{Q})_+$. These correspondences are examples of *special correspondences* in the following sense (generalising [7, Definition 6.2.1]):

Definition 3.24 Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. A *special correspondence* for $(S, \mathbb{V})$ is an irreducible subvariety $W \subset S \times S$ such that:

(1) Both projection maps $W \rightrightarrows S$ are surjective finite morphisms;
(2) W is a special subvariety for $(S \times S, \mathbb{V} \times \mathbb{V})$.

The *modular locus* for $(S, \mathbb{V})$ is the union in $S \times S$ of the special correspondences.

Thus the modular locus is a subset of the Hodge locus of $(S \times S, \mathbb{V} \times \mathbb{V})$. If $(S, \mathbb{V})$ is of Shimura type with dominant period map, this modular locus is well-known to be Zariski-dense in $S \times S$. As a second corollary to Theorem 3.1, we prove the converse:

Corollary 3.25 (Hodge commensurability criterion) *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. Then $(S, \mathbb{V})$ is of Shimura type with dominant period map if and only if the modular locus for $(S, \mathbb{V})$ is Zariski-dense in $S \times S$.*

Remark 3.26 A special case of Corollary 3.25 has been recently proven by the first and third authors. They used it to give a new proof of the Margulis commensurability criterion for arithmeticity (for complex hyperbolic lattices), see [7, Theorem 6.2.2] and the discussion in Sect. 6.2 in *op. cit.*.

Remark 3.27 Recall that a Hodge variety $\Gamma\backslash D$ can contain no positive dimensional Hodge subvariety at all, that is the only sub-Hodge datum $(\mathbf{H}, D_H) \subsetneq (\mathbf{G}, D)$ are associated to CM-points. This is of course the case if $\Gamma\backslash D$ has dimension one, but there are also higher dimensional Shimura varieties with the same property. The most famous example is perhaps given by $D = \mathbb{B}^n$ (the Hermitian symmetric space associated to $\mathrm{PU}(1, n)$), with $n + 1$ a prime number and Γ an arithmetic lattice associated to a simple division algebra (see indeed [2, Example 9.2] for all details).

However $\Gamma\backslash D$ always admits infinitely many special correspondences. Corollary 3.25 says that these special correspondences do very rarely pullback to finite correspondences on S.

3.6 Related work

A key ingredient in the proof of Theorem 3.1 is the Ax-Schanuel conjectured by the second author [34, Conjecture 7.5] and recently proved by Bakker and Tsimerman [4] (which is recalled in Sect. 4.7). Starting from Pila and Zannier [50], a link between functional transcendence and the Zilber–Pink conjecture has been observed in [19, 31, 60], at least for subvarieties of abelian and pure Shimura varieties. For the more general case of $\mathbb{Z}$VHS it has appeared for the first time in the proof of [7, Theorem 1.2.1]. In retrospect even the very first proof of the so called *geometric Manin-Mumford* used some functional transcendence (in the form of Bloch-Ochiai Theorem). See indeed [33, Theorem 4]. We remark here that a new Ax-Schanuel result for foliated principal bundles has recently been obtained by Blázquez-Sanz, Casale, Freitag, and Nagloo [10]. Such an Ax-Schanuel implies the one we are using, it provides a new proof which does not rely on o-minimality (even if our proof of Theorem 3.1 still uses some light o-minimality).

While this paper was being completed it seems that Pila and Scanlon also noticed a difference between the two Zilber–Pink conjectures for $\mathbb{Z}$VHS, namely Conjecture 2.5 and [34, Conjecture 1.9] (a point which is addressed in more details in Sect. 5.3). See indeed [48, Theorem 2.15 and Remark 2.15]. In the same paper they also speculate about the Zariski density of the typical Hodge locus (see Conjecture 5.10 in *op. cit.*). It could be that the strategy presented here applies also in their function field versions.

While this paper was being completed, we learned that de Jong [20] had also wondered about Conjecture 2.5. It is remarkable that he had considered such a problem even before the Zilber–Pink conjecture for Shimura varieties happened to be in print. In the same document he discusses the difference between Conjecture 2.5 and the Zilber–Pink conjecture as stated in [34, Conjecture 1.9], and asks if Theorem 3.3 could be true. We thank him for sharing with us his insightful notes.

3.7 Outline of the paper

Section 4 contains some preliminaries on variational Hodge theory, most notably discussing the notion of special and weakly special subvariety, the level of a $\mathbb{Z}$VHS and functional transcendence properties of period maps. Section 5 describes in details the notions of atypicality and our variant of the Zilber–Pink conjecture for $\mathbb{Z}$VHS, presenting equivalent formulations and applications. Section 6 is devoted to atypical intersections, proving the geometric part of the Zilber–Pink conjecture Theorem 3.1, and the fact that families of maximal atypical weakly special subvarieties lie in typical intersections. Section 7 proves Theorem 3.3: in level at least 3, every special subvariety is atypical. We combine Theorem 3.1 and Theorem 3.3 to obtain Theorem 1.5

in Sect. 8. In Sect. 9, we prove the results announced in Sect. 3.3 and Sect. 3.5. Section 10 proves the "all or nothing" principle Theorem 3.9 for typical intersections. Finally Sect. 11 discusses the density of the typical Hodge locus in level one, and Sect. 11.2 concludes the paper presenting our application to the existence of curves of genus four with a given Mumford–Tate group (Serre-Gross' question).

4 Preliminaries

In this section we recall the definitions and results from Hodge theory we will need, as well as adequate references for more details.

4.1 Some notation

- In this paper an algebraic variety S is a reduced scheme of finite type over the field of complex numbers, not necessarily irreducible;
- If S is an algebraic (resp. analytic) variety, by a subvariety $Y \subset S$ we always mean a *closed* algebraic (resp. analytic) subvariety. Its smooth locus is denoted by Y^{sm};
- Given $\mathbf{G}$ an algebraic group, we denote by $\mathbf{G}^{\mathrm{ad}}$ its adjoint group and by $\mathbf{G}^{\mathrm{der}}$ its derived group. Similarly if $\mathfrak{g}$ is a Lie algebra we denote by $\mathfrak{g}^{\mathrm{ad}}$ its adjoint Lie algebra and by $\mathfrak{g}^{\mathrm{der}} = [\mathfrak{g}, \mathfrak{g}]$ its derived Lie algebra. When $\mathbf{G}$ (resp. $\mathfrak{g}$) is reductive the natural morphism $\mathbf{G}^{\mathrm{der}} \to \mathbf{G}^{\mathrm{ad}}$ (resp. $\mathfrak{g}^{\mathrm{der}} \to \mathfrak{g}^{\mathrm{ad}}$) is an isogeny (resp. an isomorphism). If $\mathbf{G}$ is a $\mathbb{Q}$-algebraic group we denote by G the connected component of the identity $\mathbf{G}(\mathbb{R})^+$ of the real Lie group $\mathbf{G}(\mathbb{R})$. We use the index $_+$ to denote the inverse image of $\mathbf{G}^{\mathrm{ad}}(\mathbb{R})^+$ in $\mathbf{G}(\mathbb{R})$. Finally we set $\mathbf{G}(\mathbb{Q})^+ = G \cap \mathbf{G}(\mathbb{Q})$ and $\mathbf{G}(\mathbb{Q})_+ = \mathbf{G}(\mathbb{R})_+ \cap \mathbf{G}(\mathbb{Q})$. If $\mathbf{H} \subset \mathbf{G}$ is an algebraic subgroup we denote by $\mathbf{Z}_{\mathbf{G}}(\mathbf{H})$ its centraliser in $\mathbf{G}$ and by $\mathbf{N}_{\mathbf{G}}(\mathbf{H})$ its normaliser;
- We refer to [62] for standard definitions on Hodge structures and variations of Hodge structures. We denote by $\mathbf{S}$ the Deligne torus $\mathrm{Res}_{\mathbb{C}/\mathbb{R}} \mathbb{G}_{\mathrm{m}}$. Given $R = \mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$ we recall that an R-Hodge structure on a finitely generated R-module V is equivalent to the datum of a morphism of $\mathbb{R}$-algebraic group $h : \mathbf{S} \to \mathbf{GL}(V_{\mathbb{R}})$, where $V_{\mathbb{R}} := V \otimes_R \mathbb{R}$. For $R = \mathbb{Z}$ or $\mathbb{Q}$ a Hodge class for V is a $\mathbf{S}$-invariant vector $v \in V_{\mathbb{Q}}$. We remark here that we do not need to assume that h restricted to $\mathbb{G}_{\mathrm{m}} \subset \mathbf{S}$ is defined over $\mathbb{Q}$. A Hodge tensor for V is a Hodge class for $\bigoplus_{a,b \in \mathbb{Z}} V^{\otimes a} \otimes (V^{\vee})^{\otimes b}$ where $V^{\vee}$ denotes the Hodge structure dual to V. In this paper all Hodge structures and variations of Hodge structures are polarizable ones;
- The category of $\mathbb{Q}$-Hodge-structures is Tannakian. Recall that the Mumford–Tate group $\mathbf{MT}(V) \subset \mathbf{GL}(V)$ of a $\mathbb{Q}$-Hodge structure V is the Tannakian group of the Tannakian subcategory $\langle V \rangle^{\otimes}$ of $\mathbb{Q}$-Hodge structures generated by V. Equivalently, it is the smallest $\mathbb{Q}$-algebraic subgroup of $\mathbf{GL}(V)$ whose base-change to $\mathbb{R}$ contains the image of $h : \mathbf{S} \to \mathbf{GL}(V_{\mathbb{R}})$. It is also the fixator in $\mathbf{GL}(V)$ of the Hodge tensors for V. As V is polarised, this is a reductive group. We recall here that $\mathrm{int}(h(i))$ is a Cartan involution of $\mathbf{MT}(V)_{\mathbb{R}}/h(\mathbf{G}_{\mathrm{m},\mathbb{R}})$. See [46], and [42, Sect. 2.1] for more details;
- We remark here that the main results of the paper could also apply to the case of admissible, graded-polarizable variation of mixed Hodge structures.

4.2 Generic Mumford–Tate group and algebraic monodromy group

Let $\mathbb{V}$ be a $\mathbb{Z}$VHS on a smooth quasi-projective variety S and $Y \subset S$ a closed irreducible algebraic subvariety (possibly singular).

A point s of Y^{an} is said to be Hodge-generic in Y for $\mathbb{V}$ if $\mathbf{MT}(\mathbb{V}_{s,\mathbb{Q}})$ has maximal dimension when s ranges through Y^{an}. Two Hodge-generic points in Y^{an} for $\mathbb{V}$ have isomorphic Mumford-Tate group, where an isomorphism is given by horizontal transport along a path between the two points, called *generic Mumford-Tate group* $\mathbf{G}_Y = \mathbf{MT}(Y, \mathbb{V}_{|Y})$ of Y for $\mathbb{V}$. The Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is also the subset of points of S which are not Hodge-generic in S for $\mathbb{V}$.

The algebraic monodromy group $\mathbf{H}_Y$ of Y for $\mathbb{V}$ is the identity component of the Zariski-closure in $\mathbf{GL}(V_{\mathbb{Q}})$ of the monodromy of the restriction to Y of the local system $\mathbb{V}_{\mathbb{Z}}$. It follows from Deligne's "Theorem of the fixed part" and "Semisimplicity Theorem" [21, Sect. 4] that $\mathbf{H}_Y$ is a normal subgroup of the derived group $\mathbf{G}_Y^{\mathrm{der}}$, see [1, Theorem 1].

A simple example where the inclusion $\mathbf{H} := \mathbf{H}_S \subset \mathbf{G}^{\mathrm{der}}$ is strict is provided by a $\mathbb{Z}$VHS of the form $\mathbb{V} = \mathbb{V}_1 \otimes V_2$, where $\mathbb{V}_1$ is a $\mathbb{Z}$VHS on S with infinite monodromy and V_2 is a non-trivial constant $\mathbb{Z}$VHS (identified with a single non-trivial $\mathbb{Z}$-Hodge structure).

4.3 Period map and Hodge varieties

Let $\mathbf{G} := \mathbf{G}_S$ be the generic Mumford-Tate group of S for $\mathbb{V}$. Any point $\tilde{s}$ of the universal cover $\widetilde{S^{\mathrm{an}}}$ of S^{an} defines a morphism of real algebraic groups $h_{\tilde{s}} : \mathbf{S} \to \mathbf{G}_{\mathbb{R}}$. All such morphisms belong to the same connected component $D = D_S$ of a $\mathbf{G}(\mathbb{R})$-conjugacy class in $\mathrm{Hom}(\mathbf{S}, \mathbf{G}_{\mathbb{R}})$, which has a natural structure of complex analytic space (see [34, Proposition 3.1]). The space D is a so-called *Mumford-Tate domain*, a refinement of the classical period domain for $\mathbb{V}$ defined by Griffiths. The pair $(\mathbf{G}, D)$ is a connected Hodge datum in the sense of [34, Sect. 3.1], called the generic Hodge datum of $\mathbb{V}$. The $\mathbb{Z}$VHS $\mathbb{V}$ is entirely described by its holomorphic period map

$$\Phi : S^{\mathrm{an}} \to \Gamma \backslash D \ .$$

Here $\Gamma \subset \mathbf{G}(\mathbb{Z})$ is a finite index subgroup and $\Gamma \backslash D$ is the associated connected *Hodge variety* (see [34, Definition 3.18 and below]). We denote by $\tilde{\Phi} : \widetilde{S^{\mathrm{an}}} \to D$ the lift of Φ at the level of universal coverings.

Enlarging S if necessary by adding the components of a normal crossing divisor at infinity around which the monodromy for $\mathbb{V}$ is finite, we obtain a period map $\Phi' : S' \to \Gamma \backslash D$ extending Φ which is proper (see [30, Theorem 9.5]). Replacing first S by a finite étale covering if necessary, we can moreover assume that the arithmetic group Γ is neat (in particular torsion-free, thus $\Gamma \backslash D$ is a smooth complex analytic variety); that the monodromy at infinity of $\mathbb{V}$ is unipotent [53, Lemma (4.5)]. We leave it to the reader to verify that the results of the introduction are true for the original variety if and only if they are true for the modified variety. As a result in the following we will always assume that these conditions are satisfied.

4.4 Special and weakly special subvarieties

For this section we refer to [35, Sect. 4] for more details.

Any connected Hodge variety $\Gamma\backslash D$ is naturally endowed with a countable collection of irreducible complex analytic subvarieties: its special subvarieties $\Gamma_{\mathbf{G}'}\backslash D' \subset \Gamma\backslash D$ for $(\mathbf{G}', D') \subset (\mathbf{G}, D)$ a Hodge subdatum, and $\Gamma_{\mathbf{G}'} = \mathbf{G}'(\mathbb{Q})_+ \cap \Gamma$. More generally one defines the notion of weakly special subvariety of a Hodge variety. Let $\Gamma\backslash D$ be a Hodge variety, with associated connected Hodge datum $(\mathbf{G}, D)$. A weakly special subvariety of the Hodge variety $\Gamma\backslash D$ is either a special subvariety or a subvariety image of

$$\Gamma_{\mathbf{H}}\backslash D_H \times \{t\} \subset \Gamma_{\mathbf{H}}\backslash D_H \times \Gamma_{\mathbf{L}}\backslash D_L \xrightarrow{f} \Gamma\backslash D \ ,$$

where $(\mathbf{H} \times \mathbf{L}, D_H \times D_L)$ is a Hodge subdatum of $(\mathbf{G}^{\mathrm{ad}}, D)$, $\{t\}$ is a Hodge generic point in $\Gamma_{\mathbf{L}}\backslash D_L$ and f is a finite morphism of Hodge varieties. The datum $((\mathbf{H}, D_H), t)$ is called a *weak Hodge subdatum* of $(\mathbf{G}, D)$.

Remark 4.1 Any special subvariety of $\Gamma\backslash D$ is weakly special but the converse does not hold: the simplest example of a weakly special subvariety which is not special is provided by a Hodge generic point in a Shimura variety. Notice however that a weakly special subvariety of $\Gamma\backslash D$ containing a special subvariety $\Gamma_{\mathbf{G}'}\backslash D' \subset \Gamma\backslash D$ is special: with the notation above, the existence of the morphism of Hodge data $(\mathbf{G}', D') \to (\mathbf{H} \times \mathbf{L}, D_H \times D_L)$ sending $\Gamma_{\mathbf{G}'}\backslash D'$ to $\Gamma_{\mathbf{H}}\backslash D_H \times \{t\}$ forces $\mathbf{L}$ to be a torus and t to be CM-point.

Let $Y \subset S$ be a closed irreducible algebraic subvariety, with generic Mumford-Tate group $\mathbf{G}_Y$ and algebraic monodromy group $\mathbf{H}_Y$ for $\mathbb{V}_{|Y}$. Suppose that $\mathbf{H}_Y$ is a strict normal subgroup of $\mathbf{G}_Y^{\mathrm{der}}$. The adjoint group $\mathbf{G}_Y^{\mathrm{ad}}$ decomposes as a non-trivial product $\mathbf{H}_Y^{\mathrm{ad}} \times \mathbf{L}_Y$. Let $\widetilde{Y^{\mathrm{an}}} \subset \widetilde{S^{\mathrm{an}}}$ be an irreducible complex analytic component of the preimage of Y^{an} in $\widetilde{S^{\mathrm{an}}}$. The image $\tilde{\Phi}(\widetilde{Y^{\mathrm{an}}})$ is contained in a unique closed $\mathbf{G}_Y(\mathbb{R})^+$-orbit $D_{G_Y} = D_{H_Y} \times D_{L_Y}$ (a Mumford-Tate domain for $\mathbf{G}_Y$) in D, and in a unique closed $\mathbf{H}_Y(\mathbb{R})^+$-orbit $D_{H_Y} \times \{\tilde{t}_Y\}$ (a weak Mumford-Tate domain) in D. This gives rise (replacing Y by a finite étale cover if necessary) to a factorization

$$\Phi_{|Y^{\mathrm{an}}} : Y^{\mathrm{an}} \to \Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \times \{t_Y\} \hookrightarrow \Gamma_{\mathbf{G}_Y}\backslash D_{G_Y} = \Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \times \Gamma_{\mathbf{L}_Y}\backslash D_{L_Y} \subset \Gamma\backslash D\,,$$

where $\Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \times \{t_Y\}$ (resp. $\Gamma_{\mathbf{G}_Y}\backslash D_{\mathbf{G}_Y}$) is the smallest weakly special (resp. special) subvariety of $\Gamma\backslash D$ containing $\Phi(Y)$, called the weakly special closure (resp. the special closure) of Y in $\Gamma\backslash D$. In the case where $\mathbf{H}_Y = \mathbf{G}_Y^{\mathrm{der}}$ the weakly special closure and the special closure of Y in $\Gamma\backslash D$ coincide.

Notation 4.2 To simplify the notation we will often write hereafter simply

$$\Phi_{|Y^{\mathrm{an}}} : Y^{\mathrm{an}} \to \Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \subset \Gamma\backslash D$$

for the period map for Y, calling $(\mathbf{H}_Y, D_{H_Y})$ the weak Hodge datum associated to Y and specifying t_Y and $\mathbf{L}_Y$ only when we need them.

Remark 4.3 In particular, with the notation of (1.2), $D_{\mathbf{H}} = D_1 \times \cdots \times D_r$, $\Gamma_{\mathbf{H}_S} = \Gamma_1 \times \cdots \times \Gamma_r$, and $\{t_S\}$ is the unique point of D' determined by the projection of $\Phi(S^{\mathrm{an}})$.

Theorem 1.1 can be rephrased by saying that the preimage under Φ of any special subvariety of $\Gamma\backslash D$ is an algebraic subvariety of S. More generally the preimage under Φ of any weakly special subvariety of $\Gamma\backslash D$ is an algebraic subvariety of S, see [35, Corollary 4.8].

Definition 4.4 An irreducible algebraic subvariety $Y \subset S$ is called *special*, resp. *weakly special* for $\mathbb{V}$ or Φ if it is an irreducible component of the Φ-preimage of its special closure (resp. its weakly special closure) in $\Gamma\backslash D$.

The Hodge locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is then the countable union of the strict special subvarieties of S for $\mathbb{V}$. Notice, on the other hand, that any point of S is weakly special for $\mathbb{V}$. We thus define:

Definition 4.5 The *weakly special locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{ws}}$ of S for $\mathbb{V}$ is the union of the strict weakly special subvarieties of S for $\mathbb{V}$ of positive period dimension.

In particular: $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}} \subset \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{ws}}$.

Remark 4.6 For studying $\mathrm{HL}(S, \mathbb{V}^{\otimes})$, $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{ws}}$, $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}}$, or $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}}$, we can without loss of generality assume that the algebraic monodromy group $\mathbf{H} := \mathbf{H}_S$ of S for $\mathbb{V}$ coincide with $\mathbf{G}^{\mathrm{der}}$, hence $\Gamma_{\mathbf{H}}\backslash D_H = \Gamma\backslash D$. Indeed if $\mathbf{H}$ is a strict normal subgroup of $\mathbf{G}^{\mathrm{der}}$ the morphism of Hodge datum

$$(\mathbf{G}, D) \to (\mathbf{G}^{\mathrm{ad}} = \mathbf{H}^{\mathrm{ad}} \times \mathbf{L}, D_H \times D_L) \xrightarrow{p_1} (\mathbf{H}^{\mathrm{ad}}, D_H)$$

induces a commutative diagram of period maps

$$\begin{array}{ccccc} S^{\mathrm{an}} & \xrightarrow{\Phi} & \Gamma_{\mathbf{H}}\backslash D_H \times \{t_S\} & \hookrightarrow & \Gamma\backslash D \\ & \searrow^{\Phi'} & & & \downarrow p_1 \\ & & & & \Gamma_{\mathbf{H}}\backslash D_H \end{array}$$

One easily checks that the special (resp. weakly special) subvarieties of S for Φ coincide with the special (resp. weakly special) subvarieties of S for Φ'.

As proven in [35, Corollary 4.14] one obtains the following equivalent definitions of special and weakly special subvarieties of S for $\mathbb{V}$:

Lemma 4.7 *The special subvarieties of S for $\mathbb{V}$ are the closed irreducible algebraic subvarieties $Y \subset S$ maximal among the closed irreducible algebraic subvarieties Z of S such that the generic Mumford-Tate group $\mathbf{G}_Z$ of Z for $\mathbb{V}$ equals $\mathbf{G}_Y$.*

Lemma 4.8 *The weakly special subvarieties of S for $\mathbb{V}$ are the closed irreducible algebraic subvarieties $Y \subset S$ maximal among the closed irreducible algebraic subvarieties Z of S whose algebraic monodromy group $\mathbf{H}_Z$ with respect to $\mathbb{V}$ equals $\mathbf{H}_Y$.*

4.5 Hodge-Lie algebras

Definition 4.9 Let $K = \mathbb{Q}$ or $\mathbb{R}$, or a totally real number field. If $K = \mathbb{Q}$ or $\mathbb{R}$, a *K-Hodge-Lie algebra* is a reductive (finite dimensional) Lie algebra $\mathfrak{g}$ over K endowed with a K-Hodge structure of weight zero

$$\mathfrak{g}_{\mathbb{C}} = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}^i, \qquad \text{and} \quad \mathfrak{g}^{-i} = \overline{\mathfrak{g}^i} \quad \forall i \in \mathbb{Z}\ ,$$

such that the Lie bracket $[\cdot,\cdot] : \bigwedge^2 \mathfrak{g} \to \mathfrak{g}$ is a morphism of K-Hodge structures, and the negative of the Killing form $B_{\mathfrak{g}} : \mathfrak{g}^{\mathrm{ad}} \otimes \mathfrak{g}^{\mathrm{ad}} \to K$ is a polarisation of the K-Hodge substructure $\mathfrak{g}^{\mathrm{ad}}$ (where we identify the adjoint Lie algebra $\mathfrak{g}^{\mathrm{ad}}$ with the derived one $\mathfrak{g}^{\mathrm{der}} := [\mathfrak{g}, \mathfrak{g}]$).

If K is a totally real number field a *K-Hodge-Lie algebra* is the datum of a reductive Lie algebra $\mathfrak{g}$ over K and of a $\mathbb{Q}$-Hodge-Lie algebra structure on $\mathrm{Res}_{K/\mathbb{Q}}\, \mathfrak{g}$.

Remark 4.10 In Definition 4.9 the notation $\mathfrak{g}^i$ abbreviates the classical $\mathfrak{g}^{i,-i}$ of Hodge theory.

Remark 4.11 If $\mathfrak{g}$ is a K-Hodge-Lie algebra then the derived Lie algebra $\mathfrak{g}^{\mathrm{der}}$ is a K-Hodge-Lie subalgebra of $\mathfrak{g}$, which is equal to the adjoint K-Hodge-Lie algebra $\mathfrak{g}^{\mathrm{ad}}$ quotient of $\mathfrak{g}$ by its center; if $\mathfrak{g}$ is semi-simple as a Lie algebra, then its simple factors are naturally K-Hodge-Lie algebras; and being simple as a K-Lie algebra is equivalent to being simple as a K-Hodge-Lie algebra.

Remark 4.12 Given $h : \mathbf{S} \to \mathbf{G}_{\mathbb{R}}$ a pure Hodge structure on a real algebraic group $\mathbf{G}_{\mathbb{R}}$ in the sense of [54, (page 46)], the adjoint action of h on the Lie algebra $\mathfrak{g}_{\mathbb{R}}$ endows $\mathfrak{g}_{\mathbb{R}}$ with the structure of a real Hodge-Lie algebra. Conversely one easily checks that a real Hodge-Lie algebra structure on $\mathfrak{g}_{\mathbb{R}}$ integrates into a Hodge structure on some connected real algebraic group $\mathbf{G}_{\mathbb{R}}$ with Lie algebra $\mathfrak{g}_{\mathbb{R}}$. In particular if a simple real Lie algebra $\mathfrak{g}_{\mathbb{R}}$ admits a Hodge-Lie structure its complexification $\mathfrak{g}_{\mathbb{C}}$ is still simple, see [54, 4.4.10].

4.6 Level

Definition 4.13 Given a simple $\mathbb{R}$-Hodge-Lie algebra $\mathfrak{g}_{\mathbb{R}}$, its *level* is the largest integer k such that $\mathfrak{g}^k \neq 0$. The *level* of a simple $\mathbb{Q}$-Hodge-Lie algebra $\mathfrak{g}$ is the *maximum* of the level of the irreducible factors of $\mathfrak{g}_{\mathbb{R}} := \mathfrak{g} \otimes_{\mathbb{Q}} \mathbb{R}$. The *level* of a semi-simple $\mathbb{Q}$-Hodge-Lie algebra is the *minimum* of the levels of its simple factors. If $K = \mathbb{Q}$ or $\mathbb{R}$, the *level* of a K-Hodge structure V is the level of its adjoint K-Hodge-Lie Mumford-Tate algebra $\mathfrak{g}^{\mathrm{ad}}$.

Remark 4.14 Notice that our Definition 4.13 is *not* the standard one. Usually, the level of any pure real Hodge structure V is defined as the maximum of $k-l$ for $V^{k,l} \neq 0$. We believe that our definition, which takes into account only the adjoint Mumford-Tate Lie algebra of V, is more fundamental. For instance, the level of the weight 2 Hodge structure $H^2(S,\mathbb{Q})$, for S a $K3$-surface, is one for Definition 4.13, reflecting the fact that the motive of S is of abelian type, while it would be two with the usual definition.

It follows from Remark 4.12 that if $\mathbb{V}$ is a polarizable $\mathbb{Z}$VHS on a smooth connected quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and algebraic monodromy group $\mathbf{H}$, then any Hodge generic point $x \in \tilde{\Phi}(\widetilde{S^{\mathrm{an}}})$ defines a $\mathbb{Q}$-Hodge-Lie algebra structure $\mathfrak{g}_x^{\mathrm{ad}}$ on $\mathfrak{g}^{\mathrm{ad}}$, with $\mathbb{Q}$-Hodge-Lie subalgebra $\mathfrak{h}_x$. One immediately checks that the levels of these two structures are independent of the choice of the Hodge generic point x in D.

Definition 4.15 Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with algebraic monodromy group $\mathbf{H}$. The *level* of $\mathbb{V}$ is the level of the $\mathbb{Q}$-Hodge-Lie algebra $\mathfrak{h}_x$ for x any Hodge generic point of $\tilde{\Phi}(\widetilde{S^{\mathrm{an}}})$.

4.7 Functional transcendence

Let S be a smooth complex quasi-projective variety supporting a polarizable $\mathbb{Z}$VHS $\mathbb{V}$, with generic Hodge datum $(\mathbf{G}, D)$, algebraic monodromy group $\mathbf{H}$, and period map $\Phi : S \to \Gamma_{\mathbf{H}}\backslash D_H \subset \Gamma\backslash D$. Without loss of generality we can assume, replacing if necessary S by a finite étale cover, that Γ is torsion-free. Let $\tilde{\Phi} : \widetilde{S^{\mathrm{an}}} \to D_H$ be the lift of Φ. The domain D_H is canonically embedded as an open complex analytic real semi-algebraic subset in a flag variety $D_H^\vee$ (called its *compact dual*). Following [34] we define an irreducible algebraic subvariety of D_H (resp. $S \times D_H$) as a complex analytic irreducible component of the intersection of an algebraic subvariety of $D_H^\vee$ (resp. $S \times D_H^\vee$) with D_H (resp. $S \times D_H$).

The following result is the so called Ax-Schanuel Theorem. It was conjectured by the second author [34, Conjecture 7.5] and later proved by Bakker and Tsimerman [4, Theorem 1.1], generalising the work of Mok, Pila and Tsimerman [40, Theorem 1.1] from level one to arbitrary levels.

Theorem 4.16 (Bakker–Tsimerman) *Let $W \subset S \times D_H$ be an algebraic subvariety. Let U be an irreducible complex analytic component of $W \cap S \times_{\Gamma_{\mathbf{H}}\backslash D_H} D_H$ such that*

$$\mathrm{codim}_{S\times D_H} U < \mathrm{codim}_{S\times D_H} W + \mathrm{codim}_{S\times D_H}(S \times_{\Gamma_{\mathbf{H}}\backslash D_H} D_H) \ .$$

Then the projection of U to S is contained in a strict weakly special subvariety of S for $\mathbb{V}$.

Notice that $S \times_{\Gamma_{\mathbf{H}}\backslash D_H} D_H$ is simply the image of the graph of $\tilde{\Phi} : \widetilde{S^{\mathrm{an}}} \to D_H$ under $\widetilde{S^{\mathrm{an}}} \times D_H \to S \times D_H$.

Remark 4.17 Following the notation of Sect. 4.3, the intersection $W \cap S \times_{\Gamma_{\mathbf{H}} \backslash D_H} D_H$ can be identified with the intersection in $S \times D_H$ between W and the image of $\widetilde{S}^{\mathrm{an}}$ in $S \times D_H$ along the map $(\pi, \tilde{\Phi})$.

For the applications presented in our paper, the (simpler) case of $W = W' \times W''$ is enough, where W' is an algebraic subvariety of S, and W'' an algebraic subvariety of D_H. Notice also that in the conclusion of Theorem 4.16 the projection of U to S can be zero dimensional.

Remark 4.18 A version of the Ax-Schanuel conjecture for graded-polarized admissible variations of mixed Hodge structures has recently been established by Chiu [17, Theorem 1.2] and, independently, in joint work of Gao and the second author [26, Theorem 1.1]. On a somewhat different direction, the first and third author [7, Theorem 1.2.2] proved an Ax-Schanuel conjecture for quotients of Hermitian symmetric spaces by irreducible *non-arithmetic* lattices.

5 Typicality versus atypicality and a strong Zilber–Pink conjecture

Even if our paper is focused on the case of *pure* $\mathbb{Z}$-VHS, we remark that everything in this section can be translated to the more general case of graded-polarized admissible mixed $\mathbb{Z}$VHS (and probably all the results hold true, using Remark 4.18).

5.1 Two notions of typicality and atypicality

The choice to consider either the generic Mumford-Tate group of a $\mathbb{Z}$VHS or its algebraic monodromy group, which leads to distinguish special subvarieties from the more general weakly special subvarieties, also leads to distinguish two notions of atypicality:

Definition 5.1 Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth complex quasi-projective variety S and $\Phi : S^{\mathrm{an}} \to \Gamma_{\mathbf{H}} \backslash D_H \subset \Gamma \backslash D$ its period map. Let $Y \subset S$ be a closed irreducible algebraic subvariety, with weakly special and special closures $\Gamma_{\mathbf{H}_Y} \backslash D_{H_Y} \subset \Gamma_{\mathbf{G}_Y} \backslash D_{\mathbf{G}_Y}$ in $\Gamma \backslash D$. We define the *Hodge codimension* of Y for $\mathbb{V}$ as

$$\text{H-cd}(Y, \mathbb{V}) := \dim D_{\mathbf{G}_Y} - \dim \Phi(Y^{\mathrm{an}}) \ ,$$

and the *monodromic codimension* of Y for $\mathbb{V}$ as

$$\text{M-cd}(Y, \mathbb{V}) := \dim D_{H_Y} - \dim \Phi(Y^{\mathrm{an}}) \ .$$

Similarly:

Definition 5.2 The subvariety $Y \subset S$ is said to be *atypical* for $\mathbb{V}$ if either Y is singular for $\mathbb{V}$, or if $\Phi(S^{\mathrm{an}})$ has an excess intersection in $\Gamma \backslash D$ with the special subvariety $\Gamma_{\mathbf{G}_Y} \backslash D_{\mathbf{G}_Y}$:

$$\text{H-cd}(Y, \mathbb{V}) < \text{H-cd}(S, \mathbb{V}) \ ,$$

that is

$$\operatorname{codim}_{\Gamma\backslash D} \Phi(Y^{\mathrm{an}}) < \operatorname{codim}_{\Gamma\backslash D} \Phi(S^{\mathrm{an}}) + \operatorname{codim}_{\Gamma\backslash D} \Gamma_{\mathbf{G}_Y}\backslash D_{\mathbf{G}_Y} \ .$$

Otherwise Y is said to be *typical*.

Definition 5.3 The subvariety Y is said to be *monodromically atypical* for $\mathbb{V}$ if either Y is singular for $\mathbb{V}$, or if $\Phi(S^{\mathrm{an}})$ has an excess intersection in $\Gamma_{\mathbf{H}}\backslash D_H$ with the weakly special subvariety $\Gamma_{\mathbf{H}_Y}\backslash D_{H_Y}$:

$$\text{M-cd}(Y, \mathbb{V}) < \text{M-cd}(S, \mathbb{V}) \ ,$$

that is:

$$\operatorname{codim}_{\Gamma_{\mathbf{H}}\backslash D_H} \Phi(Y^{\mathrm{an}}) < \operatorname{codim}_{\Gamma_{\mathbf{H}}\backslash D_H} \Phi(S^{\mathrm{an}}) + \operatorname{codim}_{\Gamma_{\mathbf{H}}\backslash D_H} \Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \ .$$

Otherwise Y is said to be *monodromically typical*.

Let us generalize Definition 2.4 to the weakly special locus:

Definition 5.4 The *monodromically atypical weakly special Hodge locus* $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{ws,w\text{-}atyp}}$ of S for $\mathbb{V}$ is the union of the monodromically atypical weakly special subvarieties of S for $\mathbb{V}$ of positive period dimension.

Although we won't use it in this paper, the notion of atypicality can be refined as in [34, Definition 1.8]:

Definition 5.5 An irreducible subvariety $Y \subset S$ is *optimal* (resp. *weakly optimal*) for $\mathbb{V}$ if for any irreducible subvariety $Y \subset S$ containing Y' strictly, the following inequality holds true

$$\text{H-cd}(Y, \mathbb{V}) < \text{H-cd}(Y', \mathbb{V}) \qquad (\text{resp. M-cd}(Y, \mathbb{V}) < \text{M-cd}(Y', \mathbb{V}) \ ) \ .$$

5.2 Atypicality versus weak atypicality

Let us now compare atypicality with weak atypicality.

Lemma 5.6 *Any atypical subvariety is monodromically atypical, and therefore any weakly typical subvariety is typical.*

Proof Writing as above $\mathbf{G}_S^{\mathrm{ad}} = \mathbf{H}_S^{\mathrm{ad}} \times \mathbf{L}_S$ and $\mathbf{G}_Y^{\mathrm{ad}} = \mathbf{H}_Y^{\mathrm{ad}} \times \mathbf{L}_Y$ (the groups $\mathbf{L}_S$ and $\mathbf{L}_Y$ being possibly trivial) the inequality

$$\text{H-cd}(Y, \mathbb{V}) < \text{H-cd}(S, \mathbb{V})$$

of Hodge codimensions can be rewritten in terms of monodromic codimensions as

$$\text{M-cd}(Y, \mathbb{V}) < \text{M-cd}(S, \mathbb{V}) + \dim D_{\mathbf{L}_S} - \dim D_{\mathbf{L}_Y} \ . \tag{5.1}$$

If follows immediately that M-cd$(Y, \mathbb{V}) <$ M-cd$(S, \mathbb{V})$ holds true in the case where $\dim D_{\mathbf{L}_S} = 0$, that is if $\mathbf{H}_S = \mathbf{G}_S^{\mathrm{der}}$. In the general case notice that the morphism of Hodge data $(\mathbf{G}_Y, D_Y) \to (\mathbf{G}_S, D_S)$ induces by projection a morphism of Hodge data $(\mathbf{G}_Y, D_Y) \to (\mathbf{L}_S, D_{\mathbf{L}_S})$. As $\tilde{\Phi}(\widetilde{S^{\mathrm{an}}})$ is contained in the weak Mumford-Tate subdomain $D_{\mathbf{H}_S} \times \{t\} \subset D_{\mathbf{H}_S} \times D_{\mathbf{L}_S}$, this morphism sends $D_{\mathbf{L}_Y}$ to a Mumford-Tate subdomain of $D_{\mathbf{L}_S}$ containing the point t. As t is Hodge generic in $D_{\mathbf{L}_S}$ it follows that $D_{\mathbf{L}_Y}$ surjects onto $D_{\mathbf{L}_S}$. Thus one always has $\dim D_{\mathbf{L}_S} \leq \dim D_{\mathbf{L}_Y}$, hence M-cd$(Y, \mathbb{V}) <$ M-cd$(S, \mathbb{V})$ from (5.1). □

The converse of Lemma 5.6 is not true:

Example 5.7 There exists a $\mathbb{Z}$VHS $\mathbb{V}$ on a smooth quasi-projective surface S and a special curve $Y \subset S$ for $\mathbb{V}$ which is typical but monodromically atypical.

Proof Let us denote by $Y_1 = \Gamma_1(7)\backslash\mathbb{H}$ the classical modular curve of level $\Gamma_1(7)$, where $\mathbb{H}$ denotes the Poincaré half-plane (the level plays no particular role in the sequel, but it is needed to have a torsion free lattice). We will construct S as a surface contained in the Shimura variety $\Gamma\backslash D = Y_1^3$. To do so let us fix a CM-point $x \in Y_1$, that is a point with Mumford-Tate group a torus $\mathbf{T}$. Let $t \in Y_1$ be a Hodge generic point. We define

$$Y := Y_1 \times \{t\} \times \{x\} \subset Y_1^3 \ .$$

Let us choose an irreducible algebraic curve $C \subset Y_1^2$ such that the intersection $C \cap (Y_1 \times \{x\})$ is a finite set of points, containing the point $\{t\} \times \{x\} \in Y_1^2$. Without loss of generality we can assume that (the compactification of) C in (the compactification X_1^2 of) Y_1^2 is of degree high enough that the algebraic monodromy group of C is $\mathbf{SL}_2 \times \mathbf{SL}_2$. Finally define

$$S := Y_1 \times C \subset Y_1^3 \ .$$

The surface S is Hodge-generic in the Shimura variety Y_1^3, its generic Mumford-Tate group $\mathbf{G}$ is $\mathbf{GL}_2 \times \mathbf{GL}_2 \times \mathbf{GL}_2$, and its algebraic monodromy group is $\mathbf{SL}_2 \times \mathbf{SL}_2 \times \mathbf{SL}_2$. Its Hodge codimension is 1, its monodromic codimension is 1.

The curve $Y \subset S$ is an irreducible component of the intersection of S with the special surface $Y_1^2 \times \{x\}$ of Y_1^3, hence is a special curve in S (although it is only a weakly special curve in Y_1^3). Its generic Mumford-Tate group is $\mathbf{GL}_2 \times \mathbf{GL}_2 \times \mathbf{T}$, its algebraic monodromy group is $\mathbf{SL}_2 \times \{1\} \times \{1\}$. Its Hodge codimension is 1, its monodromic codimension is zero.

Hence the special curve $Y \subset S$ is typical but monodromically atypical. □

Remark 5.8 Although Example 5.7 shows one has to be careful when using the notion of (a)typicality, one can without loss of generality assume that $\mathbf{H} = \mathbf{G}^{\mathrm{der}}$ when studying $\mathrm{HL}(S, \mathbb{V})_{\mathrm{atyp}}$ and $\mathrm{HL}(S, \mathbb{V})_{\mathrm{typ}}$. More precisely, with the notations of Remark 4.6, we claim that a special (resp. weakly special) subvariety of S for Φ is atypical (resp. monodromically atypical) if and only if it is for Φ'. Indeed let us write as above $\mathbf{G}_S^{\mathrm{ad}} = \mathbf{H}_S^{\mathrm{ad}} \times \mathbf{L}_S$ and $\mathbf{G}_Y^{\mathrm{ad}} = \mathbf{H}_Y^{\mathrm{ad}} \times \mathbf{L}_Y$ (the groups $\mathbf{L}_S$ and $\mathbf{L}_Y$ being possibly trivial) the adjoint groups of the Mumford-Tate groups for Φ. Notice that by definition the group $\mathbf{L}_Y$ surjects onto $\mathbf{L}_S$. The corresponding Mumford-Tate groups for Φ' are then $\mathbf{H}_S^{\mathrm{ad}}$ for S, and $\mathbf{H}_Y^{\mathrm{ad}} \times \ker(\mathbf{L}_Y \twoheadrightarrow \mathbf{L}_S)$ for Y. In particular, although the Hodge codimensions H-cd(S, Φ) and H-cd(S, Φ') (resp. H-cd(Y, Φ) and H-cd(Y, Φ')) differ in general, one still has the equality:

$$\text{H-cd}(S, \Phi) - \text{H-cd}(Y, \Phi) = \text{H-cd}(S, \Phi') - \text{H-cd}(Y, \Phi') \ .$$

Hence the result.

5.3 Comparison with [34]

In [34] the second author defined a smaller Hodge codimension, which will be called in this paper the *horizontal Hodge codimension*. Similarly we can also define a *horizontal monodromic codimension*.

Let $(\mathbf{G}, D)$ be a Hodge datum. Each point $x \in D$ defines a $\mathbb{Q}$-Hodge-Lie algebra structure on the adjoint Lie algebra $\mathfrak{g}^{\mathrm{ad}}$ of $\mathbf{G}$, hence a Hodge filtration $F_x^p \mathfrak{g}_{\mathbb{C}}^{\mathrm{ad}}$ on $\mathfrak{g}_{\mathbb{C}}^{\mathrm{ad}}$ of $\mathbf{G}(\mathbb{C})$. This Hodge filtration induces a decreasing filtration on the tangent space $T_x D^\vee = \mathfrak{g}_{\mathbb{C}}^{\mathrm{ad}} / F_x^0 \mathfrak{g}_{\mathbb{C}}^{\mathrm{ad}}$. These filtrations glue together to induces a $\mathbf{G}(\mathbb{C})$- equivariant holomorphic decreasing filtration $F^\bullet T D^\vee$, $\bullet < 0$, on the $\mathbf{G}(\mathbb{C})$-homogeneous tangent bundle $T D^\vee$. The horizontal tangent bundle $T_h D$ of D is (the restriction to D of) the $\mathbf{G}(\mathbb{C})$-equivariant subbundle $F^{-1} T D^\vee$. Griffiths' transversality for VHS, i.e. the fact that $\nabla F^\bullet \mathcal{V} \subset F^{\bullet - 1} \mathcal{V} \otimes \Omega_S^1$, translates into the differential constraint $(d\Phi)(TS) \subset T_h(\Gamma \backslash D)$. Notice that the rank of the holomorphic vector bundle $T_h(\Gamma \backslash D)$ equals the dimension of $\mathfrak{g}_Y^{-1} = \mathfrak{g}_Y^{-1,1}$.

Definition 5.9 ([34, Definition 1.5 and 1.7]) Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S and let $Y \subset S$ be a closed irreducible algebraic subvariety, with generic Mumford-Tate datum $(\mathbf{G}_Y, D_Y)$ for $\mathbb{V}$. The *horizontal Hodge codimension* of Y for $\mathbb{V}$ is

$$\text{hH-cd}(Y, \mathbb{V}) := \dim_{\mathbb{C}} \left(\mathfrak{g}_Y^{-1} \right) - \dim \Phi(Y^{\mathrm{an}}) \ .$$

Similarly the *horizontal monodromy codimension* of Y for $\mathbb{V}$ is

$$\text{hM-cd}(Y, \mathbb{V}) := \dim_{\mathbb{C}} \left(\mathfrak{h}_Y^{-1} \right) - \dim \Phi(Y^{\mathrm{an}}) \ .$$

Definition 5.10 A closed irreducible algebraic subvariety $Y \subset S$ is said to be *horizontally atypical* for $\mathbb{V}$ if

$$\text{hH-cd}(Y, \mathbb{V}) < \text{hH-cd}(S, \mathbb{V}).$$

It is said to be *horizontally monodromically atypical* for $\mathbb{V}$ if

$$\text{hM-cd}(Y, \mathbb{V}) < \text{hM-cd}(S, \mathbb{V}).$$

Lemma 5.11 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth complex quasi-projective variety S. Then*:

(a) *For any closed irreducible algebraic subvariety* $Y \subset S$:

$$\text{hH-cd}(Y, \mathbb{V}) \leq \text{H-cd}(Y, \mathbb{V}) \quad \text{and} \quad \text{hM-cd}(Y, \mathbb{V}) \leq \text{M-cd}(Y, \mathbb{V}) \ ;$$

(b) $\text{hH-cd}(Y, \mathbb{V}) = \text{H-cd}(Y, \mathbb{V})$ *if and only if* $(\mathbf{G}_Y, D_Y)$ *is of Shimura type*;
(c) *A subvariety which is horizontally atypical* (*resp. horizontally monodromically atypical*) *is atypical* (*resp. monodromically atypical*).

Proof For (a): The inequality $\text{hH-cd}(Y, \mathbb{V}) \leq \text{H-cd}(Y, \mathbb{V})$ follows immediately from the obvious inequality $\dim T_h D_{\mathbf{G}_Y} \leq \dim T D_{\mathbf{G}_Y}$. Similarly $\text{hM-cd}(Y, \mathbb{V}) \leq \text{M-cd}(Y, \mathbb{V})$ follows immediately from $\dim T_h D_{H_Y} \leq \dim T D_{H_Y}$.

For (b): The equality $\text{hH-cd}(Y, \mathbb{V}) = \text{H-cd}(Y, \mathbb{V})$ implies that $F^1\mathfrak{g}_{Y,\mathbb{C}} = \mathfrak{g}_{Y,\mathbb{C}}$, i.e. $(\mathbf{G}_Y, D_Y)$ is of Shimura type.

For (c): Suppose hat $Y \subset S$ is a closed irreducible algebraic subvariety which is horizontally atypical for $\mathbb{V}$, that is $\text{hH-cd}(Y, \mathbb{V}) < \text{hH-cd}(S, \mathbb{V})$. As

$$\text{H-cd}(Y, \mathbb{V}) = \text{hH-cd}(Y, \mathbb{V}) + \sum_{k \geq 2} \mathfrak{g}_Y^{-k}$$

and $\text{H-cd}(S, \mathbb{V}) = \text{hH-cd}(S, \mathbb{V}) + \sum_{k \geq 2} \mathfrak{g}_S^{-k}$, the inequality $\text{H-cd}(Y, \mathbb{V}) < \text{H-cd}(S, \mathbb{V})$ follows immediately from the inclusions $\mathfrak{g}_Y^{-k} \subset \mathfrak{g}_S^{-k}$ for all $k \geq 2$. Thus Y is atypical for $\mathbb{V}$. The assertion concerning monodromically atypicality is proved in the same way, replacing the Lie algebra of the generic Mumford-Tate group with the Lie algebra of the algebraic monodromy group. □

5.4 Zilber–Pink conjecture (strong form)

Let us now state the Zilber-Pink conjecture for $\mathbb{Z}$VHS. It is an enhanced version (using atypicality rather than horizontal atypicality) of the Main Conjecture proposed in [34] (see Conjectures 1.9, 1.10, 1.11 and 1.11 in *op. cit.*). In view of Lemma 5.11 it implies the Main Conjecture in *loc. cit.*

Conjecture 5.12 *For any irreducible smooth quasi-projective variety S endowed with a polarizable variation of Hodge structures $\mathbb{V} \to S$, the following equivalent conditions hold true*:

(a) *The subset* $\text{HL}(S, \mathbb{V}^{\otimes})_{\text{atyp}}$ *is a finite union of maximal atypical special subvarieties of S*;
(b) *The subset* $\text{HL}(S, \mathbb{V}^{\otimes})_{\text{atyp}}$ *is a strict algebraic subvariety of S*;
(c) *The subset* $\text{HL}(S, \mathbb{V}^{\otimes})_{\text{atyp}}$ *is not Zariski-dense in S*;
(d) *The variety S contains only finitely many irreducible subvarieties optimal for $\mathbb{V}$.*

The equivalence between the four conditions presented above is obtained by arguing as in [34, Proposition 5.1].

Remark 5.13 Strictly speaking a Zilber–Pink type conjecture predicts only the behaviour of atypical intersections. It should be implicitly understood that every other intersection (namely a typical intersection) has a more predictable behaviour and satisfies an all or nothing principle: this is the content of Conjecture 2.7. Once this is understood the difference between Conjecture 5.12 and the Main Conjecture of [34] becomes crucial, since they would provide two different definitions of typicality.

5.5 An example: two-attractors

In this section we describe a concrete example which is atypical, and hence covered by Conjecture 5.12, but not horizontally atypical hence not covered by the weaker [34, Conjecture 1.9]. This example already attracted some attention in the literature, and the link with the Zilber–Pink conjecture could give a strategy for attacking it.

Consider a Calabi-Yau Hodge structure V of weight 3 with Hodge numbers $h^{3,0} = h^{2,1} = 1$ (for instance the Hodge structure given by the mirror dual quintic). Its universal deformation space is a quasi-projective curve S, since $h^{2,1} = 1$ (the projective line minus three points in the case of the mirror dual quintic), which carries a $\mathbb{Z}$VHS $\mathbb{V}$ of the same type. This gives a non-trivial period map $\Phi : S^{\mathrm{an}} \to \Gamma\backslash D$, where $D = \mathbf{Sp}(4, \mathbb{R})/U(1) \times U(1)$. The Hodge variety $\Gamma\backslash D$ is a complex manifold of dimension 4, while its horizontal tangent bundle has rank 2 (see for example [28, page 257-258]). The analytic curve $\Phi(S^{\mathrm{an}})$ is known to be Hodge generic in $\Gamma\backslash D$ (in fact its algebraic monodromy is $\mathbf{Sp}_4$). Thus $\mathrm{hH\text{-}cd}(S, \mathbb{V}) = 2 - 1 = 1$ while $\mathrm{H\text{-}cd}(S, \mathbb{V}) = 4 - 1 = 3$.

Consider the special points $s \in S$ where the weight 3 Hodge structure

$$\mathbb{V}_{s,\mathbb{C}} = V_s^{3,0} \oplus V_s^{2,1} \oplus V_s^{1,2} \oplus V_s^{0,3}$$

splits as a sum of two twisted weight one Hodge structures: $(V_s^{2,1} \oplus V_s^{1,2})$ and its orthogonal for the Hodge metric $(V_s^{3,0} \oplus V_s^{0,3})$. Following Moore [44, 45], such points $s \in S$ are referred to as *two-attractors*. They are irreducible components of the intersection of $\Phi(S^{\mathrm{an}})$ with a product $Y_n \subset \Gamma\backslash D$ of two modular curves (one of them non-horizontally embedded).

If the two-attractor $s \in S$ is moreover a CM-point then $\mathrm{hH\text{-}cd}(s, \mathbb{V}) = 0 - 0 = 0$, therefore s is horizontally atypical (hence also atypical). On the other hand if the two-attractor $s \in S$ is not CM then $\mathrm{hH\text{-}cd}(s, \mathbb{V}) = 1 - 0 = 1$ and $\mathrm{H\text{-}cd}(s, \mathbb{V}) = 2 - 0 = 2$. Hence s is horizontally-typical, but atypical.

We remark here that [34, Conjecture 1.9] predicts that the subset of CM two-attractor points in S is finite, but says nothing for the full two-attractor locus. On the other hand Conjecture 5.12 predicts that the full two-attractor locus in S is finite. It agrees with the expectation of [12, (page 44)], where the authors propose some evidence for the finiteness of the two-attractors on S.

6 Atypical locus – the geometric Zilber–Pink conjecture

In this section we first prove a version of Theorem 3.1 for the weakly atypical weakly special locus $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{ws,w\text{-}atyp}}$: the maximal monodromically atypical weakly special subvarieties of positive period dimension arise in a finite number of families whose geometry we control. The proof is mainly inspired by the arguments appearing in [60, Theorem 4.1], [7, Theorem 1.2.1] and the work of Daw and Ren [19]. In Sect. 6.4 we explain how the proof adapts to the case of $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}}$, thus proving Theorem 3.1.

Theorem 6.1 (Geometric Zilber–Pink conjecture for the weakly special atypical locus) *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$. Let Z be an irreducible component of the Zariski closure $\overline{\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{ws},w\text{-}atyp}}^{\mathrm{Zar}}$ in S. Then:*

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

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

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

Remark 6.2 Theorem 6.1 is more satisfying than Theorem 3.1: while the second branch of the alternative in Theorem 3.1 should not occur due to Conjecture 2.5, Theorem 6.1 is a complete description of varieties containing a Zariski-dense set of monodromically atypical weakly special subvarieties.

6.1 Preliminaries for the proof of Theorem 6.1

6.1.1 Definable fundamental sets

Let $\Phi : S \to \Gamma_{\mathbf{H}} \backslash D_H$ be the period map for $\mathbb{V}$. We refer the reader to [61] for the notion of o-minimal structure and to [5] for its use in Hodge theory. As recalled in Sect. 4.7, the Mumford-Tate domain D_H is a real semi-algebraic open subset of its compact dual $D_H^{\vee}$ and that its *algebraic subvarieties* are the irreducible complex analytic components of the intersections with D_H or complex algebraic subvarieties of $D_H^{\vee}$. Recall also that S being an algebraic variety, it is naturally definable in any extension of the o-minimal structure $\mathbb{R}_{\mathrm{alg}}$. *From now on, definable will be always understood in the o-minimal structure* $\mathbb{R}_{\mathrm{an,exp}}$.

Let us introduce a notion of "definable fundamental set" of S for Φ, arguing as at the beginning of [4, Sect. 3]. Let $(\overline{S}, E)$ be a log-smooth compactification of S, and choose a definable atlas of S by finitely many polydisks $\Delta^k \times (\Delta^*)^{\ell}$. Let

$$\exp : \Delta^k \times \mathbb{H}^{\ell} \to \Delta^k \times (\Delta^*)^{\ell}$$

be the standard universal cover, and choose

$$\Sigma = [-b, b] \times [1, +\infty[\subset \mathbb{H}$$

such that $\Delta^k \times \Sigma^\ell$ is a fundamental set for the $\mathbb{Z}^\ell$-action by covering transformation. Let $\mathcal{F}$ be the disjoint union of $\Delta^k \times \Sigma^\ell$ over all charts and choose lifts $\Delta^k \times \mathbb{H}^\ell \to D$ of the period map restricted to each chart to obtain a lift $\tilde{\Phi} = \tilde{\Phi}_{\mathcal{F}} : \mathcal{F} \to D_H$. From the Nilpotent Orbit Theorem [53, (4.12)] we obtain the following diagram in the category of definable complex manifolds:

$$\begin{array}{ccc} \mathcal{F} & \xrightarrow{\tilde{\Phi}} & D_H \\ \Big\downarrow{\scriptstyle \exp} & & \\ S^{\mathrm{an}} & & \end{array}$$

6.1.2 Notation for the proof of Theorem 6.1

Let $(\mathbf{M}, D_M)$ be a weak Hodge subdatum of $(\mathbf{H}, D_H)$, and set $M = \mathbf{M}(\mathbb{R})^+ \subset H = \mathbf{H}(\mathbb{R})^+$. We recall here that M and H are semialgebraic, that is definable in $\mathbb{R}_{\mathrm{alg}}$. We introduce the set

$$\Pi(\mathbf{M}, D_M) := \{(x, g) \in \mathcal{F} \times H \mid \tilde{\Phi}(x) \in g \cdot D_M\} \ , \tag{6.1}$$

where $\mathcal{F}$ denotes the definable fundamental set constructed in Sect. 6.1.1. Notice that if $(x, g) \in \Pi(\mathbf{M}, D_M)$ then $g \cdot Mg^{-1}\tilde{\Phi}(x) = g \cdot D_M$ is, following the conventions from Sect. 4.7, an algebraic subvariety of D_H. Denote by

$$\pi : D_H \to \Gamma_{\mathbf{H}} \backslash D_H$$

the projection map. We have that $\pi(gMg^{-1} \cdot \tilde{\Phi}(x)) \subset \Gamma_{\mathbf{H}} \backslash D_H$ is closed in $\Gamma_{\mathbf{H}} \backslash D_H$ only if it defines a weakly special subvariety of $\Gamma \backslash D$.

By "dimension" we will always mean the complex dimension (and possibly the local dimension at some point). Consider the function

$$d : \Pi(\mathbf{M}, D_M) \to \mathbb{R}, \qquad (x, g) \mapsto d(x, g) := \dim_{\tilde{\Phi}(x)} \left(g \cdot D_M \cap \tilde{\Phi}(\mathcal{F})\right).$$

It defines a natural decreasing filtration (for $0 \le j < n := \dim \Phi(S^{\mathrm{an}})$)

$$\Pi^j(\mathbf{M}, D_M) := \{(x, g) \in \Pi(\mathbf{M}, D_M) : d(x, g) \ge j\} \subset \Pi(\mathbf{M}, D_M). \tag{6.2}$$

Finally we define

$$\Sigma^j(\mathbf{M}, D_M) := \{(g\mathbf{M}g^{-1}, g \cdot D_M) \mid \exists (x, g) \in \Pi^j(\mathbf{M}, D_M)\} \ . \tag{6.3}$$

Notice that if we write

$$\tau : \Pi^j(\mathbf{M}, D_M) \to H/N_H(M), \quad (x, g) \mapsto g \cdot N_H(M), \tag{6.4}$$

where $N_H(M)$ denotes the normaliser of $M = \mathbf{M}(\mathbb{R})^+$ in H, then one also has

$$\Sigma^j(\mathbf{M}, D_M) = \tau(\Pi^j(\mathbf{M}, D_M)) \ .$$

Lemma 6.3 *Let $(\mathbf{M}, D_M) \subset (\mathbf{H}, D_H)$ be a weak Hodge subdatum. Then the subsets $\Pi(\mathbf{M}, D_M) \subset \mathcal{F} \times H$, $\Pi^j(\mathbf{M}, D_M) \subset \mathcal{F} \times H$, and $\Sigma^j(\mathbf{M}, D_M) \subset H/N_H(M)$ are definable subsets (where $\mathcal{F} \times H$ and $H/N_H(M)$ are endowed with their natural definable structures).*

Proof That $\Pi(\mathbf{M}, D_M) \subset \mathcal{F} \times H$ is a definable subset follows from (6.1), and from the facts that $\tilde{\Phi} : \mathcal{F} \to D_H$ is definable (see Sect. 6.1.1), that $D_M \subset D_H$ is definable and that the H-action on D_H is definable.

Observe that, since the dimension of a definable set at a point is a definable function, the set $\Pi^j(\mathbf{M}, D_M)$ is a definable subset of $\mathcal{F} \times H$.

Finally, as the projection $H \to H/N_H(M)$ is semi-algebraic, the map τ introduced in (6.4) is definable. Hence its image $\Sigma^j(\mathbf{M}, D_{\mathbf{M}}) \subset H/N_H(M)$ is a definable subset. □

6.2 Main Proposition and Ax-Schanuel

From now on, we may and do assume that the period map Φ is immersive. This is possible thanks to [6, Theorem 1.1], which asserts that the image of the period map is algebraic. We do this because the atypicality condition appearing in the Ax-Schanuel theorem concerns the dimension of S, rather than its period dimension (as it appears in our notion of atypical special subvarieties Sect. 5.1).

Definition 6.4 Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and generic weak Hodge datum $(\mathbf{H}, D_H)$. A weak Hodge subdatum $(\mathbf{M}, D_M)$ of $(\mathbf{G}, D)$ is said to be *monodromically atypical* for $\mathbb{V}$ if there exists $Z \subset S$ a monodromically atypical weakly special subvariety of S for $\mathbb{V}$, such that $(\mathbf{H}_Z, D_{H_Z}) = (\mathbf{M}, D_M)$.

Remark 6.5 Let $Z \subset S$ be as in Definition 6.4. Thus $Z = \Phi^{-1}(\Gamma_{\mathbf{M}} \backslash D_M)^0$ and either Z is singular for $\mathbb{V}$ or

$$\dim \Phi(S^{\mathrm{an}}) - \dim \Phi(Z^{\mathrm{an}}) < \dim D_H - \dim D_M \ . \tag{6.5}$$

Recall that, there are only finitely many conjugacy classes of semisimple subgroups of H. In particular, there are only finitely many weak Hodge subdata $(\mathbf{H}_i, D_{H_i})$, $i \in \{1, \dots, w\}$, of $(\mathbf{G}, D)$ such that every monodromically atypical weakly special subvariety of S for $\mathbb{V}$ is associated to a G-conjugate of some $(\mathbf{H}_i, D_{H_i})$. Let m be the maximal period dimension $\dim \Phi(Z^{\mathrm{an}})$, for $Z \subset S$ ranging through the monodromically atypical weakly special subvarieties of S for $\mathbb{V}$, which are of positive period dimension and non-singular for $\mathbb{V}$ (by assumption $m > 0$).

The main ingredient in the proof of Theorem 6.1 is the following proposition:

Proposition 6.6 *Let $(\mathbf{M}, D_M) \subset (\mathbf{G}, D)$ be a weak Hodge subdatum, monodromically atypical for $\mathbb{V}$ giving rise to a m-dimensional monodromically atypical weakly special subvariety of S. Then the set $\Sigma^m(\mathbf{M}, D_M)$ is finite.*

The proof uses (multiple times) the Ax–Schanuel Theorem recalled in Sect. 4.7.

Proof As $\Sigma^m(\mathbf{M}, D_M)$ is a definable subset of $H/N_H(M)$, it is enough to show that $\Sigma^m(\mathbf{M}, D_M)$ is countable to conclude that it is finite.

To show that $\Sigma^m(\mathbf{M}, D_M)$ is countable it is enough to show that for each $(x, g) \in \Pi^m(\mathbf{M}, D_M)$ the subset

$$\pi(g \cdot D_M) \subset \Gamma_{\mathbf{H}} \backslash D_H \subset \Gamma \backslash D$$

is a weakly special subvariety (as the Hodge variety $\Gamma \backslash D$ admits only countably many families of weakly special subvarieties).

For $(x, g) \in \Pi^m(\mathbf{M}, D_M)$ consider the algebraic subset of $S \times D_H$

$$W_{(x,g)} := S \times g \cdot D_M \ .$$

Choose $U_{(x,g)}$ an irreducible complex analytic component of maximal dimension at $(\exp(x), \tilde{\Phi}(x))$ of

$$W_{(x,g)} \cap \left(S \times_{\Gamma_{\mathbf{H}} \backslash D_H} D_H\right) \subset S \times D_H.$$

We claim that $U_{(x,g)}$ is an atypical intersection, that is

$$\operatorname{codim}_{S \times D_H} U_{(x,g)} < \operatorname{codim}_{S \times D_H} W_{(x,g)} + \operatorname{codim}_{S \times D_H} \left(S \times_{\Gamma_{\mathbf{H}} \backslash D_H} D_H\right).$$

Indeed this last inequality can be rewritten as

$$\dim \Phi(S^{\mathrm{an}}) - \dim U_{(x,g)} < \dim D_H - \dim g \cdot D_M \ . \tag{6.6}$$

As $(x, g) \in \Pi^m(\mathbf{M}, D_M)$, one has $\dim U_{(x,g)} = m = \dim \Phi(Z^{\mathrm{an}})$. On the other hand

$$\dim g \cdot D_M = \dim D_M \ .$$

Hence the required inequality (6.6) is exactly the inequality (6.5), thus $U_{(x,g)}$ is an atypical intersection.

The Ax–Schanuel conjecture for $\mathbb{Z}$VHS, Theorem 4.16, thus asserts that the projection $S_{(x,g)}$ of $U_{(x,g)}$ to S is contained in some strict weakly special subvariety of S. Let $S' = \Phi^{-1}(\Gamma_{\mathbf{H}'} \backslash D_{H'})$ be such a strict weakly special subvariety of S containing $S_{(x,g)}$, of smallest possible period dimension:

$$S_{(x,g)} \subset S' \subsetneq S.$$

In particular $\dim \Phi(S'^{\mathrm{an}}) \geq \dim S_{(x,g)} = m$ with equality if and only if $S' = S_{(x,g)}$.

Suppose that S' is monodromically typical for $\mathbb{V}$. Thus

$$\dim \Phi(S^{\mathrm{an}}) - \dim \Phi(S'^{\mathrm{an}}) = \dim D_H - \dim D_{H'} \ . \tag{6.7}$$

Consider the algebraic subvariety

$$W'_{(x,g)} := S' \times g \cdot D_M \subset S' \times D_{H'}$$

and let $U'_{(x,g)}$ be an irreducible complex analytic component at $(\exp(x), \tilde{\Phi}(x))$ of

$$W'_{(x,g)} \cap (S' \times_{\Gamma_{\mathbf{H}'} \backslash D_{H'}} D_{H'}) \subset S' \times D_{H'}$$

containing $U_{(x,g)}$. The typicality condition (6.7) ensures that $U'_{(x,g)}$ is still an atypical intersection. Applying again the Ax–Schanuel conjecture, Theorem 4.16 produces a strict weakly special subvariety $S'' \subsetneq S'$ containing $S_{(x,g)}$. This contradicts the minimality of S'.

Thus S' has to be monodromically atypical (and non-singular) for $\mathbb{V}$. As $\dim \Phi(S') \geq m$ it follows from the maximality of m that $\dim \Phi(S') = m$, thus $S' = S_{(x,g)}$. This is to say that $\pi(g \cdot D_M)$ is a weakly special subvariety of $\Gamma\backslash D$.

This concludes the proof of Proposition 6.6. □

6.3 Proof of Theorem 6.1

We have all the elements to finally prove Theorem 6.1. For some similar arguments we refer to the proofs of [60, Theorem 4.1] and [7, Theorem 1.2.1. and Sect. 6.1.3].

Proof of Theorem 6.1 We need the following lemma, describing the weakly special subvarieties of S singular for $\mathbb{V}$:

Lemma 6.7 *Let $Y \subset S$ be a weakly special subvariety of S for $\mathbb{V}$, singular for $\mathbb{V}$, and maximal for these properties. Let S' be an irreducible component of $S_{\mathbb{V}}^{\text{sing}}$ containing Y. If Y satisfies the equality*

$$\dim \Phi(S^{\text{an}}) - \dim \Phi(Y^{\text{an}}) = \dim \Gamma_{\mathbf{H}}\backslash D_H - \dim \Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \ , \tag{6.8}$$

then either $Y \subset S'$ is monodromically atypical for $\mathbb{V}_{|S'}$ or $Y = S'$.

Proof Suppose that $Y \subset S'$ is monodromically typical for $\mathbb{V}_{|S'}$. In particular:

$$\dim \Phi(S'^{\text{an}}) - \dim \Phi(Y^{\text{an}}) = \dim \Gamma_{\mathbf{H}_{S'}}\backslash D_{H_{S'}} - \dim \Gamma_{\mathbf{H}_Y}\backslash D_{H_Y} \ .$$

Subtracting from (6.8) and as $\dim \Phi(S'^{\text{an}}) < \dim \Phi(S^{\text{an}})$ by definition of the singular locus, it follows that S' is contained in a strict weakly special subvariety of S. As S' contains Y which is maximal among the weakly special subvarieties of S, we deduce by irreducibility of Y and S' that $S' = Y$. □

Let Z be as in Theorem 6.1. If Z contains a Zariski-dense set of weakly special subvarieties of S singular for $\mathbb{V}$, each satisfying the equality (6.8), and maximal for these properties, it then follows from Lemma 6.7 that:

- Either Z coincides with an irreducible component S' of $S_{\mathbb{V}}^{\text{sing}}$ which is a weakly special subvariety of S for $\mathbb{V}$. But then Z satisfies (a) of Theorem 6.1;
- Or Z is an irreducible component of $\overline{\text{HL}(S', \mathbb{V}^{\otimes}_{|S'})_{\text{ws,w-atyp}}}^{\text{Zar}}$ for such an irreducible component S'.

Without loss of generality (replacing S by S' if necessary, and iterating) we are reduced to the case where Z contains a Zariski-dense set of monodromically atypical weakly special subvarieties of S for $\mathbb{V}$, non-singular for $\mathbb{V}$ and of positive period dimension.

Recall that m is the maximum of the period dimensions of the weakly atypical weakly special subvarieties of S. For any $j \leq m$ consider the set

$$\mathcal{E}_j := \{\text{maximal monodromically atypical weakly special subvarieties of } S \\ \text{non-singular for } \mathbb{V} \text{ and of period dimension } j\}\ , \quad (6.9)$$

and, following the notation introduced in Sect. 6.2, write

$$\mathcal{E}_j = \coprod_{1\leq i\leq w} \coprod_{(\mathbf{M}, D_M)\in\Sigma^j(\mathbf{H}_i, D_{H_i})} \mathcal{E}_j(\mathbf{M}, D_M)\ , \quad (6.10)$$

where $\mathcal{E}_j(\mathbf{M}, D_M)$ denotes the set of atypical weakly special subvarieties of S of period dimension j with generic weak Hodge datum $(\mathbf{M}, D_M)$.

Let Z be as in Theorem 6.1. There exists a larger $j \in \{1, \dots, m\}$ and an $i \in \{1, \dots, w\}$ such that the union of the special subvarieties of $\mathcal{E}_j(\mathbf{M}, D_M)$ contained in Z, for $(\mathbf{M}, D_M)$ ranging through $\Sigma^j(\mathbf{H}_i, D_{H_i})$, is Zariski-dense in Z.

Suppose first that $j = m$. As $\Sigma^m(\mathbf{H}_i, D_{H_i})$ is a finite set by Proposition 6.6, there exists $(\mathbf{M}, D_M) \in \Sigma^m(\mathbf{H}_i, D_{H_i})$ such that the union of the special subvarieties of $\mathcal{E}_j(\mathbf{M}, D_M)$ contained in Z is Zariski-dense in Z. Let $\mathbf{L} := \mathbf{Z}_{\mathbf{G}^{\mathrm{ad}}}(\mathbf{M}^{\mathrm{ad}})$.

Either L is compact, in which case, similarly to [60, Proof of Theorem 4.1], we observe that $\mathcal{E}_m(\mathbf{M}, D_M)$ is a finite set as $\Sigma^m(\mathbf{M}, D_M)$ is finite, and that each element in $\mathcal{E}_m(\mathbf{M}, D_M)$ is a special subvariety. It follows that Z satisfies (a) in Theorem 6.1.

Or L is not compact. In that case the pair

$$(\mathbf{M}', D_{M'}) := (\mathbf{M}^{\mathrm{ad}} \cdot \mathbf{L}, Z_G(M) \cdot D_M)$$

is a Hodge subdatum of $(\mathbf{G}^{\mathrm{ad}}, D)$, satisfying the inclusions of weak Hodge data

$$(\mathbf{M}, D_M) \subsetneq (\mathbf{M}', D_{M'}) \subset (\mathbf{G}^{\mathrm{ad}}, D)\ .$$

Let $Z' \subset S$ be an irreducible component of $\Phi^{-1}(\Gamma_{\mathbf{M}'}\backslash D_{M'})$ containing Z. It is a special subvariety of S containing Z. By maximality of Z among the monodromically atypical weakly special subvarieties, we observe that Z' is weakly typical and of product type. Thus Z satisfies (b) in Theorem 6.1.

This finishes the proof of Theorem 6.1 in the case where $j = m$.

We now argue by induction (downward) on j. Consider the set

$$\Sigma_2^j(\mathbf{H}_i, D_{H_i}) := \{(g\mathbf{H}_i g^{-1}, g \cdot D_{H_i}) \mid \exists (x, g) \in \Sigma^j(\mathbf{H}_i, D_{H_i}),\ \forall j' > j\ \ S_{x,g} \notin \mathcal{E}_{j'}\},$$

where $\Sigma^j(\mathbf{H}_i, D_{H_i})$ is defined as in (6.3). The inductive assumption implies that the second condition is definable. The proof of Proposition 6.6 shows that $\Sigma_2^j(\mathbf{H}_i, D_{H_i})$ is finite. Arguing as in the case $j = m$ we conclude again that Z satisfies either (a) or (b) of Theorem 6.1.

This concludes the proof of the Geometric Zilber–Pink conjecture for monodromically atypical weakly special subvarieties. □

6.4 Proof of Theorem 3.1

To prove Theorem 3.1 one repeats the proof of Theorem 6.1, replacing the inputs "weakly special" and "monodromically atypical" by "special" and "atypical" respectively; and, in the proof, "weak Hodge subdatum" by "Hodge subdatum" and all the inequalities concerning the monodromic codimension by one with the Hodge codimension. As atypical subvarieties are monodromically atypical by Lemma 5.6, we can still apply the Ax-Schanuel conjecture in the proof of Proposition 6.6 to produce a weakly special subvariety S', which is not only monodromically atypical but atypical. The argument then gives rise to the alternative (a) or (b) in Theorem 3.1 in the same way it gave rise to the alternative (a) or (b) for Theorem 6.1 (for the branch (a) we use Remark 4.1 to deduce that Z, which is an atypical weakly special subvariety containing a special subvariety, is actually an atypical special subvariety).

7 A criterion for the typical Hodge locus to be empty

In this section we prove Theorem 3.3 and Proposition 3.4. Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and algebraic monodromy group $\mathbf{H}$. Theorem 3.3 and Proposition 3.4 are variations of the following results (see Sect. 7.2 for this deduction):

Theorem 7.1 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and algebraic monodromy group $\mathbf{H}$. If the level of $\mathbb{V}$ is at least 3, then $\mathbb{V}$ does not admit any strict monodromically typical weakly special subvariety.*

Proposition 7.2 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected quasi-projective variety S, with generic Hodge datum $(\mathbf{G}, D)$ and algebraic monodromy group $\mathbf{H}$.*

If the level of $\mathbb{V}$ is 2 and $\mathfrak{h}$ is $\mathbb{Q}$-simple, then any strict monodromically typical weakly special subvariety is associated to a sub-datum $(\mathbf{G}', D_{\mathbf{G}'})$, such that $\mathbf{G}'^{\mathrm{ad}}$ is simple.

7.1 Proof of Theorem 7.1

The proof uses the fact that the $\mathbb{Q}$-Hodge-Lie algebra $\mathfrak{h}$ of the algebraic monodromy group $\mathbf{H}$ is generated in level 1, see below; and the analysis by Kostant [37] of root systems of Levi factors for complex semi-simple Lie algebras.

Definition 7.3 Let $K = \mathbb{Q}$ or $\mathbb{R}$. A K-Hodge-Lie algebra $\mathfrak{g}$ is said to be *generated in level* 1 if the smallest K-Hodge-Lie subalgebra of $\mathfrak{g}$ whose complexification contains $\mathfrak{g}^{-1}$ and $\mathfrak{g}^{1}$ coincides with $\mathfrak{g}$.

Proposition 7.4 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, with algebraic monodromy group $\mathbf{H}$. Then the $\mathbb{Q}$-Hodge-Lie subalgebra structure $\mathfrak{h}_s$ defined by any point $s \in S$ on $\mathfrak{h}$ is generated in level* 1.

More precisely, the $\mathbb{R}$-Hodge-Lie subalgebra structure defined by any point $s \in S$ on the non-compact part $\mathfrak{h}_{\mathbb{R}}^{\mathrm{nc}}$ of $\mathfrak{h}_{\mathbb{R}} := \mathfrak{h} \otimes_{\mathbb{Q}} \mathbb{R}$ is generated in degree 1, *and the compact part $\mathfrak{h}_{\mathbb{R}}^{\mathrm{c}}$ is of Hodge type* $(0, 0)$.

Proposition 7.5 *Suppose $\mathfrak{g}_{\mathbb{R}}$ is a simple $\mathbb{R}$-Hodge-Lie algebra generated in level* 1, *and of level at least* 3. *If $\mathfrak{g}'_{\mathbb{R}} \subset \mathfrak{g}_{\mathbb{R}}$ is an $\mathbb{R}$-Hodge-Lie subalgebra satisfying ${\mathfrak{g}'}^{i} = \mathfrak{g}^i$ for all $|i| \geq 2$ then $\mathfrak{g}' = \mathfrak{g}$.*

Assuming for the moment Proposition 7.4 and Proposition 7.5, let us finish the proof of Theorem 7.1. Let $Y \subset S$ be a monodromically typical weakly special subvariety for $\mathbb{V}$, with algebraic monodromy group $\mathbf{H}_Y$. In the following we fix a point $x \in D_{H_Y} \cap \widetilde{\Phi}(\widetilde{S^{\mathrm{an}}})$ and consider the $\mathbb{Q}$-Hodge-Lie algebra pair $\mathfrak{h}_Y \subset \mathfrak{h}$ defined by this point. The typicality condition writes:

$$\dim \left(\bigoplus_{i \geq 1} \mathfrak{h}_Y^{-i} \right) - \dim \Phi(Y^{\mathrm{an}}) = \dim \left(\bigoplus_{i \geq 1} \mathfrak{h}^{-i} \right) - \dim \Phi(S^{\mathrm{an}}) \ . \tag{7.1}$$

As Y is monodromically typical, it is also horizontally monodromically typical by Lemma 5.11(c):

$$\dim \mathfrak{h}_Y^{-1} - \dim \Phi(Y^{\mathrm{an}}) = \dim \mathfrak{h}^{-1} - \dim \Phi(S^{\mathrm{an}}) \ . \tag{7.2}$$

We thus deduce from (7.1) and (7.2) that

$$\dim \left(\bigoplus_{i \geq 2} \mathfrak{h}_Y^{-i} \right) = \dim \left(\bigoplus_{i \geq 2} \mathfrak{h}^{-i} \right) \ ,$$

thus $\mathfrak{h}_Y^{-i} = \mathfrak{h}^{-i}$ as $\mathfrak{h}_Y^{-i} \subset \mathfrak{h}^{-i}$, for all $i \geq 2$. As $\mathfrak{h}_Y^{i} = \overline{\mathfrak{h}_Y^{-i}}$ and $\mathfrak{h}^{i} = \overline{\mathfrak{h}^{-i}}$ we finally deduce:

$$\mathfrak{h}_Y^{i} = \mathfrak{h}^{i} \quad \text{for all } |i| \geq 2. \tag{7.3}$$

The assumption that $\mathfrak{h}$ is of level at least 3 says that each simple $\mathbb{Q}$-factor of $\mathfrak{h}$ is at level at least 3. Moreover the equality (7.3) remains true if we replace $\mathfrak{h}$ by any of its $\mathbb{Q}$-simple factors and $\mathfrak{h}_Y$ by its intersection with this simple factor. To deduce that $\mathfrak{h}_Y = \mathfrak{h}$, we can thus without loss of generality assume that $\mathfrak{h}$ is $\mathbb{Q}$-simple.

By Proposition 7.4 it follows that one simple $\mathbb{R}$-factor $\mathfrak{l}_{\mathbb{R}}$ of $\mathfrak{h}_{\mathbb{R}}^{\mathrm{nc}}$ is generated in level 1 and of level at least 3. It then follows from Proposition 7.5 that the intersection $\mathfrak{h}_{Y,\mathbb{R}} \cap \mathfrak{l}_{\mathbb{R}}$ equals $\mathfrak{l}_{\mathbb{R}}$. As both $\mathfrak{h}_Y$ and $\mathfrak{h}$ are defined over $\mathbb{Q}$ and $\mathfrak{h}$ is the smallest $\mathbb{Q}$-algebra whose $\mathbb{R}$-extension contains $\mathfrak{l}_{\mathbb{R}}$, we conclude that $\mathfrak{h}_Y = \mathfrak{h}$. Thus $Y = S$, which finishes the proof of Theorem 7.1, assuming Proposition 7.4 and Proposition 7.5. □

Proof of Proposition 7.4 Passing to a finite étale cover of S if necessary, we can and will assume without loss of generality that the monodromy of the local system $\mathbb{V}$ is contained in $\mathbf{H}(\mathbb{R})$. Let $\mathbb{V}_{\mathfrak{h}}$ be the $\mathbb{Q}$VHS associated to the adjoint representation of the algebraic monodromy group $\mathbf{H}$. For each $s \in S$, let $\mathfrak{h}_s$ be the fiber at s of $\mathbb{V}_{\mathfrak{h}}$. This is a semi-simple $\mathbb{Q}$-Hodge-Lie algebra. We write $\mathfrak{h}_{\mathbb{R},s} = \mathfrak{h}^{\mathrm{nc}}_{\mathbb{R},s} \oplus \mathfrak{h}^{\mathrm{c}}_{\mathbb{R},s}$ for the decomposition of the $\mathbb{R}$-Hodge-Lie-algebra $\mathfrak{h}_{\mathbb{R},s}$ into its non-compact and compact part.

As $\mathbf{H}$ has no $\mathbb{Q}$-anisotropic factor, $\mathfrak{h}_{\mathbb{R},s}$ is the smallest subalgebra of $\mathfrak{h}_{\mathbb{R},s}$ defined over $\mathbb{Q}$ and containing $\mathfrak{h}^{\mathrm{nc}}_{\mathbb{R},s}$. Proving that $\mathfrak{h}^{\mathrm{nc}}_{\mathbb{R},s}$ is generated in level 1 thus implies that $\mathfrak{h}_s$ is generated in level 1. Let us now turn to the proof of this last statement.

Let $\mathfrak{h}'_{\mathbb{C},s} \subset \mathfrak{h}_{\mathbb{C},s}$ be the complex Lie subalgebra generated by $\mathfrak{h}_s^{-1}$ and $\mathfrak{h}_s^1$. It is naturally defined over $\mathbb{R}$: $\mathfrak{h}'_{\mathbb{C},s} = \mathfrak{h}'_{\mathbb{R},s} \otimes_{\mathbb{R}} \mathbb{C}$ and $\mathfrak{h}'_{\mathbb{R},s} \subset \mathfrak{h}_{\mathbb{R},s}$ is a real Lie subalgebra. We claim that the collection $\left(\mathfrak{h}'_{\mathbb{R},s}\right)_{s\in S}$ defines a local subsystem of $\mathbb{V}_{\mathfrak{h}_{\mathbb{R}}} := \mathbb{V}_{\mathfrak{h}} \otimes_{\mathbb{Q}} \mathbb{R}$. To prove the claim, it is enough to show that the corresponding holomorphic subbundle of $\mathcal{V}_{\mathfrak{h}} := \mathbb{V}_{\mathfrak{h}_{\mathbb{C}}} \otimes_{\mathbb{C}} \mathcal{O}_S$ is stable under the flat connection ∇ defined by $\mathbb{V}_{\mathfrak{h}_{\mathbb{C}}}$. But a local holomorphic vector field X on S identifies under the period map with a local section u of $F^{-1}\mathcal{V}_{\mathfrak{h}}$ over $\Phi(S^{\mathrm{an}})$; and, under this identification, the derivation ∇_X at a point s is nothing else than $\mathrm{ad}\, u$, which preserves $\mathfrak{h}'_{\mathbb{C},s}$ at each point according to the very definition of the latter. Hence the claim.

It follows that $\mathfrak{h}'_{\mathbb{R},s}$ is an $\mathbf{H}_s(\mathbb{R})$-submodule of the Lie algebra $\mathfrak{h}_{\mathbb{R},s}$ under the adjoint action, i.e. an ideal of $\mathfrak{h}_{\mathbb{R},s}$. As $\mathfrak{h}_{\mathbb{R},s}$ is semi-simple we obtain a decomposition of real Lie algebras, hence of $\mathbb{R}$-Hodge-Lie algebras (see Remark 4.11):

$$\mathfrak{h}_{\mathbb{R},s} = \mathfrak{h}'_{\mathbb{R},s} \oplus \mathfrak{l}_s \ ;$$

and of the associated flag domains $D_H = D_{H'} \times D_L$. The subvariety $\widetilde{\Phi}(\widetilde{S^{\mathrm{an}}}) \subset D_H$ is horizontal, thus tangent to $D_{H'}$ at every point. Hence there exists $d_L \in D_L$ such that $\widetilde{\Phi}(\widetilde{S^{\mathrm{an}}}) \subset D_{H'} \times \{d_L\}$. It follows that the monodromy $(\mathrm{Ad}\,\rho)(\pi_1(S,s))$ of the real local system $\mathbb{V}_{\mathfrak{h}_{\mathbb{R}}}$, hence also its Zariski-closure $\mathbf{H}_{\mathbb{R}}$, is contained in $\mathbf{H}' \times \mathbf{M}_L$, where $\mathbf{M}_L$ is the $\mathbb{R}$-anisotropic stabilizer of d_L in $\mathbf{L}$. Thus $\mathbf{H}_{\mathbb{R}} = \mathbf{H}'_{\mathbb{R}} \times \mathbf{M}_L$; $\mathbf{H}'_{\mathbb{R}}$ is the non-compact part of $\mathbf{H}_{\mathbb{R}}$, which is thus generated in degree 1; and $\mathbf{L} = \mathbf{M}_L$ is the compact part of $\mathbf{H}_{\mathbb{R}}$, of pure type $(0, 0)$. Hence the result. □

Remark 7.6 Proposition 7.4 should be compared with [52, Prop. 3.4]. There Robles obtains a weaker result for the more general situation of an horizontal subvariety Z of D_H not necessarily coming from an $\mathbb{R}$VHS on a quasi-projective base (her result is stated for $(\mathbf{H}, D_H) = (\mathbf{G}, D)$ but the proof adapts immediately to the general case; moreover, as indicated to us by C. Robles the $\mathbb{Q}$s appearing in [52, Proposition 3.4] have to be replaced by $\mathbb{R}$s). In her case the group $\mathbf{H}'_{\mathbb{R}}$ is not necessarily a factor of $\mathbf{H}_{\mathbb{R}}$. Our stronger conclusion (and easier proof) comes from the ubiquitous use of Deligne's semisimplicity theorem.

Proof of Proposition 7.5 The proof follows from the results of Kostant in [37]. We will use his notation in our setting to help the reader. Thus let us write $\mathfrak{m} := \mathfrak{g}^0$, a Levi factor of the proper parabolic subalgebra $\mathfrak{q} := F^0\mathfrak{g}_{\mathbb{C}}$. Let $\mathfrak{t} := \mathrm{cent}\,\mathfrak{m}$ be the center of $\mathfrak{m}$ and let $\mathfrak{s} := [\mathfrak{m}, \mathfrak{m}]$. Thus $\mathfrak{m} = \mathfrak{t} \oplus \mathfrak{s}$. Let $\mathfrak{r}$ be the orthogonal complement for the Killing form of $\mathfrak{m}$ in $\mathfrak{g}$ so that $\mathfrak{g}_{\mathbb{C}} = \mathfrak{m} \oplus \mathfrak{r}$ and $[\mathfrak{m}, \mathfrak{r}] \subset \mathfrak{r}$. A nonzero element $\nu \in \mathfrak{t}^*$ is called a $\mathfrak{t}$-root if $\mathfrak{g}_\nu \neq 0$, where

$$\mathfrak{g}_\nu = \{z \in \mathfrak{g}_{\mathbb{C}}, \quad \mathrm{ad}\, x(z) = \nu(x)z,\ \forall x \in \mathfrak{t}\} \ .$$

We denote by $R \subset \mathfrak{t}^*$ the set of all $\mathfrak{t}$-roots. Following [37, Theorem 0.1], the root space $\mathfrak{g}_\nu$ is an irreducible $\mathrm{ad}\,\mathfrak{m}$-module for any $\nu \in R$, and any irreducible $\mathfrak{m}$-submodule of $\mathfrak{r}$ is of this form. Moreover if $\nu, \mu \in R$ and $\nu + \mu \in R$ then $[\mathfrak{g}_\nu, \mathfrak{g}_\mu] = \mathfrak{g}_{\nu+\mu}$, see indeed Theorem 2.3 in *op. cit.*.

Let $R^+ \subset R$ be the set of positive $\mathfrak{t}$-roots defined in [37, (page 139)]. Thus the unipotent radical $\mathfrak{n} = F^1\mathfrak{g}_\mathbb{C}$ of $\mathfrak{q} = F^0\mathfrak{g}_\mathbb{C}$ coincides with $\bigoplus_{\nu \in R^+} \mathfrak{g}_\nu$. Let $T \in \mathfrak{t}$ be the grading element defining the Hodge graduation $\mathfrak{g}_\mathbb{C} = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}^i$, see [52, Sect. 2.2]. Thus $\mathfrak{g}_\nu \subset \mathfrak{g}^{\nu(T)}$ and $\nu \in R^+$ if and only if $\nu(T) > 0$.

A $\mathfrak{t}$-root in R^+ is called simple if it cannot be written as a sum of two elements of R^+. Let $R_{\text{simple}} \subset R^+$ denote the set of simple roots. In [37, Theorem 2.7] Kostant proves that the elements of R_{simple} form a basis of $\mathfrak{t}^*$. If $R_{\text{simple}} = \{\beta_1, \ldots, \beta_l\}$ (where $\dim \mathfrak{t} = l$) then moreover the $\mathfrak{g}_{\beta_i}$, $i \in I := \{1, \ldots, l\}$, generate $\mathfrak{n}$ under bracket.

As $\mathfrak{g}_\mathbb{R}$ is a simple real algebra, it follows from Remark 4.12 that $\mathfrak{g}_\mathbb{C}$ is a simple complex Lie algebra. Since $\mathfrak{g}_\mathbb{R}$ is assumed to be generated in level 1, it follows that for all $i \in I$, $\mathfrak{g}_{\beta_i} \subset \mathfrak{g}^1$ (otherwise β_i would not be a simple root); and $\mathfrak{g}^1 = \bigoplus_{i \in I} \mathfrak{g}_{\beta_i}$.

Let $\mathfrak{h}_\mathbb{C} \subset \mathfrak{g}_\mathbb{C}$ be the complex Lie subalgebra generated by $\bigoplus_{i \geq 2} \mathfrak{g}^i$ and $\bigoplus_{i \geq 2} \mathfrak{g}^{-i}$. It is naturally an $\mathfrak{m}$-submodule of $\mathfrak{g}_\mathbb{C}$, hence $\mathfrak{h}_\mathbb{C} \cap \mathfrak{g}^1$ decomposes as a direct sum $\bigoplus_{j \in J \subset I} \mathfrak{g}_{\beta_j}$. As $\mathfrak{g}_\mathbb{R}$ is generated in degree one and the level k of $\mathfrak{g}$ is at least 3, it follows from [37, Theorem 2.3] that the subspace $[\mathfrak{g}^{-2}, \mathfrak{g}^3] \subset \mathfrak{h}_\mathbb{C} \cap \mathfrak{g}^1$ is not 0, thus J is not empty. Notice that, for any $i \in I - J$ and $j \in J$, one has $[\mathfrak{g}_{\beta_i}, \mathfrak{g}_{\beta_j}] = 0$. Indeed the Killing form is $\mathrm{ad}(\mathfrak{m})$-invariant, and $\mathfrak{h}$ is closed under $\mathrm{ad}(\mathfrak{m})$; hence $\mathfrak{h}^\perp$ is closed under $\mathrm{ad}(\mathfrak{m})$, *a fortiori* under $\mathrm{ad}(\mathfrak{t})$, whence $(\mathfrak{h}^\perp)^1 = \bigoplus_{i \in I - J} \mathfrak{g}_{\beta_i}$. Since $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^\perp$ as a Lie algebra, the desired conclusion follows.

If $J \neq I$ this contradicts the fact that the restriction to $\mathfrak{t}$ of the highest root of $\mathfrak{g}_\mathbb{C}$ is of the form $\sum_{i \in I} n_{\beta_i} \beta_i$, with all $n_{\beta_i} > 0$ for all $i \in I$, see [37, (2.39) and (2.44)]. Thus $I = J$ and $\mathfrak{h}_\mathbb{C} = \mathfrak{g}_\mathbb{C}$. □

Remark 7.7 (Proposition 7.5 does not hold in level two) Let $(\mathbf{G}, D)$ be the Hodge datum parametrising polarized Hodge structures of weight 2 with Hodge numbers $(h^{2,0}, h^{1,1}, h^{0,2} = h^{2,0})$ for some $h^{2,0} > 0$, $h^{1,1} > 0$. Let $(\mathbf{H}, D_H)$ be a sub-Hodge datum of $(\mathbf{G}, D)$. Recall that $\mathfrak{g}^{-2}$ is non zero, as long as $h^{2,0} > 1$. Then the fact that the canonical inclusion

$$\mathfrak{h}^{-2} \subseteq \mathfrak{g}^{-2}$$

is an equality implies that D_H is parametrising Hodge structures of weight 2 with Hodge numbers $(h^{2,0}, a, h^{2,0})$ with $0 < a < h^{1,1}$. That is, in level two, typical intersections may (and do) happen, as recalled in Sect. 3.4.2. In Sect. 7.3, we prove that the period domain $D_\mathbf{H}$ cannot be a non-trivial product, indeed in such an example $\mathbf{H} = \mathbf{SO}(2h^{2,0}, a)$ and $\mathbf{G} = \mathbf{SO}(2h^{2,0}, h^{1,1})$.

7.2 Proof of Theorem 3.3

Without loss of generality (see Remark 5.8) we can assume that $\mathbf{H} = \mathbf{G}^{\text{der}}$. In the proof of Theorem 7.1, we may then replace the weakly typical subvarieties by the typical ones, and $\mathfrak{h}_Y$ by $\mathfrak{g}_Y^{\text{der}}$. Everything works the same. It is important that $\mathbf{H} = \mathbf{G}^{\text{der}}$ so that Proposition 7.4 applies to $\mathfrak{g}^{\text{der}}$. With this formulation, Theorem 7.1 implies immediately Theorem 3.3. □

7.3 Proof of Proposition 7.2

To prove Proposition 7.2, and therefore Proposition 3.4, it is enough to prove the following:

Proposition 7.8 *Suppose* $\mathfrak{g}$ *is a simple* $\mathbb{Q}$*-Hodge-Lie algebra generated in level* 1, *and of level* 2. *If* $\mathfrak{g}' \subset \mathfrak{g}$ *is a* $\mathbb{Q}$*-Hodge-Lie subalgebra satisfying* $\mathfrak{g}'^{-2} = \mathfrak{g}^{-2}$, *then* $\mathfrak{g}'$ *is simple.*

Proof Arguing as in the previous section, we can suppose that $\mathbf{G}$ is $\mathbb{R}$-simple. This implies that $\mathfrak{g}^{-2}$ is a simple (non-zero) $\mathfrak{g}^0$-module. Notice also that $\mathfrak{g}^0$ is a product $[\mathfrak{g}^{-2}, \mathfrak{g}^2] \times \mathfrak{u}$ (where $\mathfrak{u}$ is possibly trivial). This implies that $\mathfrak{g}^{-2}$ is of the form $R \otimes S$, in the case where $\mathfrak{u}$ is non trivial, or simply R if $\mathfrak{u}$ is trivial, where R is a simple $[\mathfrak{g}^{-2}, \mathfrak{g}^2]$-module and S a simple $\mathfrak{u}$-module.

Let $\mathfrak{g}' \subset \mathfrak{g}$ be a Hodge-Lie subalgebra such that $\mathfrak{g}'^{-2} = \mathfrak{g}^{-2}$. We want to show that $\mathfrak{g}'$ is simple. Suppose by contradiction it is not simple. Thus

$$\mathfrak{g}' = \mathfrak{g}'_1 \oplus \mathfrak{g}'_2 \ , \tag{7.4}$$

for two non-trivial Hodge-Lie subalgebras $\mathfrak{g}'_i$, $i = 1, 2$. The decomposition (7.4) gives rise to a decomposition

$$\mathfrak{g}^{\pm 2} = \mathfrak{g}'^{\pm 2} = \mathfrak{g}_1'^{\pm 2} \oplus \mathfrak{g}_2'^{\pm 2} \ . \tag{7.5}$$

Suppose first that both $\mathfrak{g}_1'^{-2}$ and $\mathfrak{g}_2'^{-2}$ are non-zero. In that case

$$[\mathfrak{g}^{-2}, \mathfrak{g}^2] = [\mathfrak{g}_1'^{-2}, \mathfrak{g}_1'^{2}] \oplus [\mathfrak{g}_2'^{-2}, \mathfrak{g}_2'^{-2}]$$

is a non-trivial decomposition, and the decomposition (7.5) is a non-trivial decomposition of the $\mathfrak{g}^0$-module $\mathfrak{g}^{-2}$, where $\mathfrak{g}^0$ acts on $\mathfrak{g}_i'^{-2}$, $i = 1, 2$, through its quotient $[\mathfrak{g}_i'^{-2}, \mathfrak{g}_i'^{2}] \times \mathfrak{u}$. This contradicts the simplicity of $\mathfrak{g}^{-2}$ as a $\mathfrak{g}^0$-module.

Thus, exchanging factors if necessary, we can assume without loss of generality that $\mathfrak{g}_1'^{-2} = \mathfrak{g}^{-2}$ and $\mathfrak{g}_2'^{-2} = 0$. It follows in particular that:

$$\mathfrak{g}^2 \oplus [\mathfrak{g}^{-2}, \mathfrak{g}^2] \oplus \mathfrak{g}^{-2} \subset \mathfrak{g}'_{1\mathbb{C}} \ . \tag{7.6}$$

As $\mathfrak{g}'_2$ centralizes $\mathfrak{g}'_1$, it centralizes, and hence in particular normalizes, $\mathfrak{g}^2 \oplus [\mathfrak{g}^{-2}, \mathfrak{g}^2] \oplus \mathfrak{g}^{-2}$. The normaliser $\mathfrak{n}$ of $\mathfrak{g}^2 \oplus [\mathfrak{g}^{-2}, \mathfrak{g}^2] \oplus \mathfrak{g}^{-2}$ in $\mathfrak{g}$ is therefore a Hodge-Lie subalgebra containing $\mathfrak{g}'_2$. As $\mathfrak{g}^2 \oplus [\mathfrak{g}^{-2}, \mathfrak{g}^2] \oplus \mathfrak{g}^{-2}$ is a $\mathfrak{g}^0$-module, its normaliser $\mathfrak{n}$ is a $\mathfrak{g}^0$-module too. Thus $\mathfrak{n}^{-1}$ decomposes into a sum of irreducible $\mathfrak{g}^0$-modules. As already explained in Sect. 7.1, for any non-trivial irreducible $\mathfrak{g}^0$-submodule $E \subset \mathfrak{g}^{-1}$, we have that $[E, \mathfrak{g}^2]$ is a non-zero subspace of $\mathfrak{g}^1$. In particular such an E cannot normalize $\mathfrak{g}^2 \oplus [\mathfrak{g}^{-2}, \mathfrak{g}^2] \oplus \mathfrak{g}^{-2}$.

Therefore we have proven that $\mathfrak{n}^{-1} = 0$, which implies $\mathfrak{g}_2'^{-1} = 0$. Finally $\mathfrak{g}'_2 = \mathfrak{g}_2'^{0}$ is a compact Lie algebra, establishing the contradiction we were aiming for. □

8 Proof of Theorem 1.5

Let $\mathbb{V}$ be, as in Theorem 1.5, of level at least 3. Using Remark 4.6 we can without loss of generality (in particular without changing the level) assume that $\mathbf{H} = \mathbf{G}^{\mathrm{der}}$. It then follows from Theorem 3.3 (proven as Theorem 7.1) that $\mathrm{HL}(S, \mathbb{V}) = \mathrm{HL}(S, \mathbb{V})_{\mathrm{atyp}}$. Let $Z \subset S$ be an irreducible component of the Zariski-closure of $\mathrm{HL}(S, \mathbb{V})_{\text{f-pos}} = \mathrm{HL}(S, \mathbb{V})_{\text{f-pos,atyp}}$. Let us apply Theorem 3.1 to Z. If we are in case (b) of Theorem 3.1, then necessarily $Z = S$, as S does not admit any *strict* typical special subvariety of positive dimension. With the notations of the proof of Theorem 3.1, we thus obtain a decomposition $\mathbf{G}^{\mathrm{ad}} = \mathbf{M} \times \mathbf{L}$, where $\mathbf{L}$ is such that D_L is not a point, and $\mathbf{M}$ is the algebraic monodromy group of special subvarieties Y_i, $i \in \mathbb{N}$, contained in $\mathrm{HL}(S, \mathbb{V})_{\text{f-pos}}$. But then the projection of $\Phi({Y_i}^{\mathrm{an}})$ on $\Gamma_{\mathbf{L}} \backslash D_L$ is a point: contradiction to the fact that Y_i is factorwise of positive dimension. Thus we are necessarily in case (a) of Theorem 3.1: Z is a maximal strict (atypical) special subvariety of S for $\mathbb{V}$. Thus $\mathrm{HL}(S, \mathbb{V})_{\text{f-pos}}$ is algebraic. Finally, If $\mathbf{G}^{\mathrm{ad}}$ is simple, then

$$\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\text{f-pos}} = \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos}} = \mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}}$$

is algebraic in S. Theorem 1.5 is therefore proven. □

The proof of Theorem 3.7 is essentially the same, replacing Theorem 3.3 by Proposition 7.2.

Proof of Theorem 3.7 Assume that $\mathbb{V}$ is of level 2 and that $\mathbf{G}^{\mathrm{ad}}$ is simple. Thanks to Proposition 7.2, and we obtain that $\mathrm{HL}(S, \mathbb{V}^{\otimes})_{\mathrm{pos,atyp}}$ is algebraic, as claimed in Theorem 3.7. □

9 Applications in higher level

In this section we combine the Geometric Zilber–Pink conjecture (Theorem 3.1) with the fact that, in level > 1, the typical Hodge locus is constrained (as established in Sect. 7). In particular we prove the results announced in Sect. 3.3 and then the applications presented in Sect. 3.5.

9.1 Proof of Corollary 1.6 and Corollary 3.8

Let $n \geq 2$ and $d \geq 3$. Let $\mathbf{P}_{\mathbb{C}}^{N(n,d)}$ be the projective space parametrising the hypersurfaces X of $\mathbf{P}_{\mathbb{C}}^{n+1}$ of degree d, with

$$N(n, d) = \binom{n+d+1}{d} - 1.$$

Let $U_{n,d} \subset \mathbf{P}_{\mathbb{C}}^{N(n,d)}$ be the Zariski-open subset parametrising the smooth hypersurfaces X (its complement, the so called *discriminant locus*, is irreducible and of codimension one). Let $\mathbf{G}_{n,d}$ be the group of automorphisms of $H^n(X, \mathbb{Q})_{\mathrm{prim}}$ preserving the cup-product. When n is odd the primitive cohomology is the same as the cohomology. When n is even it is the orthogonal complement of $h^{n/2}$, where h is some fixed hyperplane class. Thus $\mathbf{G}_{n,d}$ is either a symplectic or an orthogonal group depending on the parity of n, and it is a simple $\mathbb{Q}$-algebraic group. One knows that the monodromy group $\mathbf{H}$ of $\mathbb{V}$ coincides with the simple group $\mathbf{G}_{n,d}$, see indeed the remark below. As $\mathbf{H} \subset \mathbf{G}^{\mathrm{ad}} \subset \mathbf{G}_{n,d}$ we deduce $\mathbf{G}^{\mathrm{ad}} = \mathbf{G}_{n,d}$, hence $\mathbf{G}^{\mathrm{ad}}$ is simple.

If $n = 3$ and $d \geq 5$; $n = 4$ and $d \geq 6$; $n = 5, 6, 8$ and $d \geq 4$; and $n = 7$ or ≥ 9 and $d \geq 3$, one checks that the level of $\mathbf{G}^{\mathrm{ad}} = \mathbf{G}_{n,d}$ is at least 3, giving Corollary 1.6. This follows from Griffiths' residue theory [29], and the computation of the Hodge diamond of $H^n(X, \mathbb{Z})_{\mathrm{prim}}$. See in particular [27, Lecture 4], and also [14], for similar computations. For example the level is n, as soon as $d \geq n + 2$, and the level is at least three also if $n = 3$, and $d = 5$. If $n = 2$, one sees that the level is 2, as soon as $d \geq 5$ (it is a well known fact that the level is one if $d = 4$).

The results then follow from Theorem 1.5 and Theorem 3.7, proved in Sect. 8.

Remark 9.1 Beauville [9, Theorem 2 and Theorem 4], building on the work of Ebeling and Janssen, computes exactly the image of the monodromy representation mentioned above. In our argument we just need to know $\mathbf{H}$, the Zariski closure of the image of such a monodromy representation. The easier fact that $\mathbf{H} = \mathbf{G}_{n,d}$ follows from the *Picard-Lefschetz formulas* and it is due to Deligne [22, Proposition 5.3 and Theorem 5.4], see also [23, Sect. 4.4].

9.2 Proof of Corollary 3.21

Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S, and $\Gamma\backslash D$ be the associated Mumford–Tate domain. We first argue assuming that the adjoint group of the generic Mumford–Tate $\mathbf{G}$ of $\mathbb{V}$ is simple.

Assume that S contains a Zariski-dense set of special subvarieties Z_n which are of Shimura type (but not necessarily with dominant period map), and of period dimension > 0. Heading for a contradiction let us assume that they are atypical intersections. Theorem 3.1 then implies that S is an atypical intersection, that is that

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

Which is a contradiction. That means that the Z_n are typical subvarieties for $(S, \mathbb{V})$. As explained in Sect. 7.1, to have a typical intersection between $\Phi(S^{\mathrm{an}})$ and the special closure of Z_n, which we denote by $\Gamma_{\mathbf{H}_n}\backslash D_{\mathbf{H}_n}$ we must have the following equality at the level of holomorphic tangent spaces (at some point $P \in Z_n$):

$$T(S) + T(\Gamma_{\mathbf{H}_n}\backslash D_{\mathbf{H}_n}) = T(\Gamma\backslash D). \tag{9.1}$$

By definition, however, the left hand side lies in the horizontal tangent bundle $T_h(\Gamma\backslash D)$, which was introduced at the beginning of Sect. 5.3. It follows that $\Gamma\backslash D$ must be a Shimura variety, as explained in (b) of Lemma 5.11, since we must have

$$T(D) = T_h(D).$$

Moreover we can observe that S is Hodge generic, and its Hodge locus is Zariski-dense. Corollary 3.21 is thereby proven.

For the case where $\mathbf{G}^{\mathrm{ad}}$ is not assumed to be simple we argue as follows, using the full power of the geometric Zilber–Pink. Assume, as above, that S contains a Zariski-dense set of special subvarieties Z_n which are of Shimura type, of period dimension > 0, and atypical intersections. Then, by Theorem 3.1, we have a decomposition $\mathbf{G}^{\mathrm{ad}} = \mathbf{G}' \times \mathbf{G}''$, in such a way that the Z_n (or a covering thereof) have trivial period dimension for Φ''. Moreover we can assume that there are no further decompositions of $(\mathbf{G}', D_{G'})$ in such a way that the Z_n have trivial period dimension for one of such factors. But then S is atypical for Φ', which is a contradiction. That means that the Z_n have to be typical intersection for Φ', and so, as explained previously, $\Gamma' \backslash D'$ is a Shimura variety.

Remark 9.2 (Shimura varieties inside a period domain) Let D be the classifying space for Hodge structures of weight two on a fixed integral lattice, and $\Gamma\backslash D$ be the associated space of isomorphism classes. To fix the notation, as in Remark 7.7, let

$$h^{2,0} - h^{1,1} - h^{0,2}$$

be the Hodge diamond we are parametrising. To be more explicit we notice here that $\mathbf{G} = \mathbf{SO}(2h^{2,0}, h^{1,1})$.

Assume that $h^{1,1}$ is even, and write:

$$h^{2,0} = p > 1, \quad h^{1,1} = 2q > 0.$$

Shimura-Hodge subvarieties of D appear for example from the trivial map *from complex to real*

$$\mathrm{SU}(p,q) \to \mathrm{SO}(2p, 2q),$$

which indeed induces a geodesic immersion (from the Hermitian space associated to $\mathrm{SU}(p,q)$ to D), as one can read in [13, Construction 1.4] and [52].

9.3 Proof of Corollary 3.25 – special correspondences are atypical intersections

Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. Here we describe the geometry of a portion of the Hodge locus of $(S \times S, \mathbb{V} \times \mathbb{V})$, which we refer to as the *modular locus*, as introduced in Definition 3.24.

Let $\Phi : S^{\mathrm{an}} \to \Gamma\backslash D$ be the period map associated to $\mathbb{V}$, and, as usual assume that $(\mathbf{G}, D) = (\mathbf{G}_S, D_S)$, that is that $\Phi(S^{\mathrm{an}})$ is Hodge generic in $\Gamma\backslash D$. A special correspondence is a $\dim S$-subvariety of $S \times S$ which come from sub-Hodge datum of $(\mathbf{G} \times \mathbf{G}, D \times D)$ of the form $T_g \subset \Gamma\backslash D \times \Gamma\backslash D$, for some g in the commensurator of Γ. That is

$$(\Phi \times \Phi)(W^{\mathrm{an}}) = \left((\Phi \times \Phi)(S^{\mathrm{an}} \times S^{\mathrm{an}}) \cap T_g\right)^0.$$

To prove Corollary 3.25, via Theorem 3.1, we show that the above is an atypical (maximal[1]) intersection in $\Gamma\backslash D \times \Gamma\backslash D$, unless Φ is dominant, which implies that $\Gamma\backslash D$ is a Shimura variety (see for example [34, Lemma 4.11]). Notice that

$$2\dim\Gamma\backslash D - \dim\Phi(S^{\mathrm{an}}) = \operatorname{codim}\Phi(W^{\mathrm{an}})$$

is strictly smaller than

$$\operatorname{codim}((\Phi\times\Phi)(S^{\mathrm{an}}\times S^{\mathrm{an}})) + \operatorname{codim} T_g = 2(\dim\Gamma\backslash D - \dim\Phi(S^{\mathrm{an}})) + \dim\Gamma\backslash D,$$

(where all codimensions are computed in $\Gamma\backslash D \times \Gamma\backslash D$), if and only if

$$\dim\Phi(S^{\mathrm{an}}) < \dim\Gamma\backslash D.$$

That is, if and only if Φ is not dominant. Corollary 3.25 is therefore a simple consequence of Theorem 3.1.

10 Typical locus – all or nothing

In this section we prove Theorem 3.9, which, for the reader's convenience, is recalled below. It immediately implies Corollary 3.12, providing the elucidation of Theorem 1.3 we were looking for.

Theorem 10.1 *Let $\mathbb{V}$ be a polarizable $\mathbb{Z}$VHS on a smooth connected complex quasi-projective variety S. If the typical Hodge locus $\mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{typ}}$ is non-empty then $\mathrm{HL}(S,\mathbb{V}^{\otimes})$ is analytically (hence Zariski) dense in S.*

As usual, we let

$$\Phi : S^{\mathrm{an}} \to \Gamma\backslash D$$

be the period map associated to $\mathbb{V}$, where $(\mathbf{G}, D) = (\mathbf{G}_S, D_S)$ is the generic Hodge datum associated to $(S, \mathbb{V})$. The proof of the above result builds on a local computation (at some smooth point).

10.1 Proof of Theorem 10.1

Let $\Gamma'\backslash D'$ be a period subdomain of $\Gamma\backslash D$ such that

$$0 \le \dim((\Phi(S^{\mathrm{an}}) \cap \Gamma'\backslash D')^0) = \dim\Phi(S^{\mathrm{an}}) + \dim\Gamma'\backslash D' - \dim\Gamma\backslash D, \qquad (10.1)$$

that is we have one typical intersection $Z = \Phi^{-1}(\Gamma'\backslash D')^0$ (accordingly to Definition 2.2). By definition we also have that Z is not singular for $\mathbb{V}$. Let $\mathbf{H} \subset \mathbf{G}$ be the generic Mumford–Tate of D', and, as usual, write $H = \mathbf{H}(\mathbb{R})^+ \subset G = \mathbf{G}(\mathbb{R})^+$ and

[1] Maximality follows directly form the fact that T_g is not contained in any strict period sub-domains of $\Gamma\backslash D \times \Gamma\backslash D$.

$\mathfrak{h}$ for its Lie algebra, which is an $\mathbb{R}$-Hodge substructure of $\mathfrak{g} = \mathrm{Lie}(G)$. Finally we denote by $N_G(H)$ the normaliser of H in G. Let

$$\mathcal{C}_H := G/N_G(H) = \{H' = gHg^{-1} : g \in G\}$$

be the set of all subgroups of G that are conjugated to H (under G) with its natural structure of real-analytic manifold (a manifold which, unsurprisingly, appeared also in the proof of Theorem 3.1, see indeed Proposition 6.6). Set

$$\Pi_H := \{(x, H') \in D \times \mathcal{C}_H : x(\mathbf{S}) \subset H'\} \subset D \times \mathcal{C}_H,$$

and let π_1 (resp. π_2) be the natural projection to D (resp. to $\mathcal{C}_H$). Notice that π_i are real-analytic G-equivariant maps (where G acts diagonally on Π_H).

Let $\bar{S}$ be the preimage of $\Phi(S^{\mathrm{an}})$ in D, along the natural projection map $D \to \Gamma\backslash D$, and $\tilde{S}$ be the preimage of $\overline{S}$ in Π_H, along π_1. By restricting π_2 we have a real-analytic map

$$f : \tilde{S} \to \mathcal{C}_H.$$

By a simple topological argument, as explained for example in [16, Proposition 1], to prove that $\mathrm{HL}(S, \mathbb{V}^{\otimes})$ is dense in S, it is enough to prove that f is generically a submersion (that is a submersion outside a nowhere-dense real analytic subset B of $\tilde{S}$).[2] As being submersive is an open condition (for the real analytic topology), it is enough to find a smooth point in $\tilde{S}$ at which f is submersive.

Let us analyse this condition. Let $y = (P, H') \in D \times \mathcal{C}_H$ be a point of $\tilde{S}$. The real tangent space of $\mathcal{C}_H$ at $f(y)$ is canonically isomorphic to

$$\mathfrak{g}/N_G(\mathrm{Lie}(H')) = \mathfrak{g}/\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}').$$

The image of df at $y = (\overline{P}, H')$ is equal to

$$(\mathfrak{m} + T_P(\overline{S})_{\mathbb{R}} + \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}'))/\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}') \tag{10.2}$$

where $\mathfrak{m}$ is the Lie algebra of M (from the identification $D = G/M$). Thus f is a submersion at $y = (P, H')$ if and only if (10.2) is equal to $\mathfrak{g}/\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')$. Now we work on the *complex* tangent spaces (by tensoring all our $\mathbb{R}$-Lie algebras with $\mathbb{C}$), and use the fact that $\mathfrak{h}$, $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')$, and $\mathfrak{m}$ are naturally endowed with an $\mathbb{R}$-Hodge structure. Thus

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}^{-w} \oplus \cdots \oplus \mathfrak{g}^{-1} \oplus \mathfrak{g}^0 \oplus \mathfrak{g}^1 \oplus \cdots \oplus \mathfrak{g}^w,$$

for some $w \geq 1$. To ease the notation set

$$U := T_P(\overline{S})_{\mathbb{C}} \subset \mathfrak{g}^{-1}.$$

[2] Because of its simplicity, for the convenience of the reader, we recall here Chai's argument. Let $\Omega \subset S^{\mathrm{an}}$ be an open subset. Since f is open, the image of $\pi^{-1}(\Omega) - B$ in $\mathcal{C}_H$ is open, hence meets in a dense set the $\mathbf{G}(\mathbb{Q})_+$-conjugates of H, as $\mathbf{G}(\mathbb{Q})_+$ is topologically dense in G.

It follows that f is a submersion at $y=(P,H')$ if and only if the following two conditions are satisfied (compare with [16, Proposition 2]):

$$\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')^{-k}=\mathfrak{g}^{-k} \tag{10.3}$$

for any $k=w,\dots,2$ (recall that $\mathfrak{m}$ is pure of type $(0,0)$, so $\mathfrak{m}^{-k}=0$ in this range), and

$$U+\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h}')^{-1}=\mathfrak{g}^{-1}. \tag{10.4}$$

Let us now exhibit a point $y=(P,H')\in\tilde{S}$ satisfying both (10.3) and (10.4). Choose P one lift in D of a smooth Hodge-generic point of $\Phi(Z^{\mathrm{an}})$ and $H'=\mathbf{H}'(\mathbb{R})^+$ where $\mathbf{H}'$ is the Mumford-Tate group at that point. Then (10.1) implies the following decomposition of the holomorphic tangent bundle of D at P:

$$U+T_P(D')_{\mathbb{C}}=T_P(D)_{\mathbb{C}}=\mathfrak{g}^{-w}\oplus\cdots\oplus\mathfrak{g}^{-1}. \tag{10.5}$$

In particular we have $(\mathfrak{h}')^{-k}=\mathfrak{g}^{-k}$, for all $k>1$, which implies (10.3) (since $(\mathfrak{h}')^{-k}\subset\mathfrak{h}'$) and $(\mathfrak{h}')^{-1}+U=\mathfrak{g}^{-1}$, which implies that (10.4) is satisfied at y.

Theorem 10.1 is proven.

Remark 10.2 If $\mathbb{V}$ is of level two and $\mathbf{G}^{\mathrm{ad}}$ is simple, then the density established above has to come either from the typical Hodge locus, or from the atypical Hodge locus of zero period dimension (as $\mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{pos,atyp}}$ is algebraic, as established in Theorem 3.7). That is, if $\mathrm{HL}(S,\mathbb{V}^{\otimes})_{\mathrm{pos,typ}}$ is non-empty, then it is dense in S.

11 On a question of Serre and Gross

To conclude the paper we are left to prove the results announced in Sect. 3.4.2 and Sect. 3.4.3.

11.1 Proof of Theorem 3.14, after Chai

We first recall the main theorem of Chai [16], whose proof was of inspiration for the argument of Sect. 10.

Let $(\mathbf{G},X)$ be a connected Shimura datum, and Γ a torsion free finite index subgroup of $\mathbf{G}(\mathbb{Z})$. The quotient $\Gamma\backslash X$ is a smooth quasi-projective variety [3], let n be its complex dimension, which we assume to be >0. Recall that a *special* subvariety of $\Gamma\backslash X$ is, by definition, a component of the Hodge locus $\mathrm{HL}(\Gamma\backslash X,\mathbb{V}^{\otimes})$, where $\mathbb{V}$ denotes any $\mathbb{Z}$VHS on $\Gamma\backslash X$ defined by a faithful algebraic representation of $\mathbf{G}$ (this Hodge locus is independent of the choice of such a representation).

Let $(\mathbf{H},X_H)$ be a sub-Shimura datum of $(\mathbf{G},X)$, and let $S\subset\Gamma\backslash X$ be an irreducible closed subvariety. As usual we set $H=\mathbf{H}(\mathbb{R})^+\subset G=\mathbf{G}(\mathbb{R})^+$. Consider the following subset of $\mathrm{HL}(S,\mathbb{V}^{\otimes})$:

$$\mathrm{HL}(S,\mathbf{H}):=\{x\in S:\mathbf{MT}(x)\subset g\mathbf{H}g^{-1}\text{ for some }g\in\mathbf{G}(\mathbb{Q})_+\}. \tag{11.1}$$

Theorem 11.1 ([16]) *There exists a constant* $c = c(G, X, H) \in \mathbb{N}$ *such that if* S *has codimension at most* c *in* $\Gamma\backslash X$, *then* $\mathrm{HL}(S, \mathbf{H})$ *is (analytically) dense in* S.

To establish Theorem 3.14 we just need to prove the following, which builds on an explicit computation of $c = c(G, X, H)$ given by Chai:

Lemma 11.2 *Assume that* $\mathbf{G}$ *is absolutely simple*, $\Gamma_H\backslash X_H \subset \Gamma\backslash X$ *has dimension one, and* $N_G(H) = H$. *Then* $c(G, X, H) > 0$.

Proof Let $h : \mathbf{S} \to H$ be a Hodge generic point of X_H. Let K_h be the centraliser of h in G. And recall that, as explained in Sect. 7.1, $\mathfrak{g}^{-1}$ is an irreducible $K_h \otimes \mathbb{C}$-module.

Chai [16, beginning of page 409] proves[3] that $c(G, h, H)$ is the largest non-negative integer such that for every $\mathbb{C}$-vector subspace W of $\mathfrak{g}^{-1}$ of codimension at most $c(G, h, H)$, there exists an element $k \in K_h$ with

$$\mathfrak{g}^{-1} = W + \mathrm{Ad}\, k(\mathfrak{h}^{-1}) \ .$$

Notice that the sub-$\mathbb{C}$-vector space of $\mathfrak{g}^{-1}$ generated by $\bigcup_{k \in K_h} \mathrm{Ad}\, k(\mathfrak{h}^{-1})$ is stable under the action of K_h, and therefore, by the irreducibility of $\mathfrak{g}^{-1}$ as K_h-module, it is equal to $\mathfrak{g}^{-1}$. As $\mathfrak{h}^{-1}$ is one dimensional, there must exist a $k_0 \in K_h$ for which

$$\mathrm{Ad}\, k_0(\mathfrak{h}^{-1}) \nsubseteq W.$$

That is, if W has codimension one in $\mathfrak{g}^{-1}$, then $\mathfrak{g}^{-1} = W + \mathrm{Ad}\, k_0(\mathfrak{h}^{-1})$. Which is saying that $c(G, h, H) \geq 1$, as desired. □

Combining the above lemma and Theorem 11.1 we obtain the following, which implies Theorem 3.14:

Theorem 11.3 (Chai $+ \epsilon$) *Assume that* $\mathbf{G}$ *is absolutely simple*, $X_H \subset X$ *has dimension one, and that* $N_G(H) = H$. *If* S *has codimension one in* $\Gamma\backslash X$, *then* $\mathrm{HL}(S, \mathbf{H})$ *is dense in* S. *That is*

$$\bigcup_{g \in \mathbf{G}(\mathbb{Q})_+} S \cap T_g(\Gamma_H\backslash X_H) \qquad (11.2)$$

is dense in S.

Remark 11.4 Under the André–Oort conjecture for Shimura varieties, Theorem 11.3 could be upgraded saying that the *typical* Hodge locus is in fact dense in S (as we will do in the next section in a special case), as long as S is not a special subvariety of $\Gamma\backslash X$. That is to say that the intersections appearing in (11.2) are, up to a finite number, zero-dimensional and with Mumford-Tate group equal to $g\mathbf{H}g^{-1}$, for some $g \in \mathbf{G}(\mathbb{Q})_+$.

[3] That is to say that, in the notation of Definition 2.2 in *op. cit.*, $c(G, h, H) = d(K_h, \mathfrak{g}^{-1}, \mathfrak{h}^{-1})$.

11.2 Proof of Theorem 3.17 – typical and atypical intersections altogether

Let $\mathcal{M}_g$ be the moduli space of curves of genus g, and

$$j : \mathcal{M}_g \hookrightarrow \mathcal{A}_g$$

be the Torelli morphism. Recall that the left hand side is irreducible and has dimension $3g-3$, while the right hand side has dimension $\frac{g(g+1)}{2}$. We denote the image of j by

$$\mathcal{T}_g^0 = j(\mathcal{M}_g) \subset \mathcal{A}_g,$$

which will be referred to as the *open Torelli locus*, and by $\mathcal{T}_g$ its Zariski closure (the so called *Torelli locus*). To be more precise a n-level structure, for some $n>5$, should be fixed. If $g>3$, it is well known that $\mathcal{T}_g$ is Hodge generic in $\mathcal{A}_g$. This can indeed be observed by using the fact that $\mathcal{T}_g$ is not special and that the fundamental group of $\mathcal{M}_g$ surjects onto the one of $\mathcal{A}_g$, see for example [43, Remark 4.5].

From now on, we set $g=4$. This is important because it is the only g for which

$$\operatorname{codim}_{\mathcal{A}_g}(\mathcal{T}_g) = 1 = 10-9 = \operatorname{codim}_{\mathcal{A}_4}(\mathcal{T}_4).$$

Recall that $\mathcal{A}_4$ contains a special curve Y whose generic Mumford-Tate group $\mathbf{H}$ is isogenous to a $\mathbb{Q}$-form of $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3/\mathbb{C}$, as proven by Mumford [46]. Mumford's construction is such that the normaliser of $\mathbf{H}$ in $\mathbf{G}$ is indeed $\mathbf{H}$. Theorem 11.3 shows that $\mathrm{HL}(\mathcal{T}_4, \mathbf{H})$ is analytically dense[4] in $\mathcal{T}_4$. This means that $\mathcal{T}_4$ cuts many Hecke translates Y_n of Y. In particular, upon extracting a sequence of Y_n, we have

$$Y_n \cap \mathcal{T}_4^0 \neq \emptyset, \tag{11.3}$$

since it is not possible that all intersections happen on $\mathcal{T}_4 - \mathcal{T}_4^0$.

Since the Y_n are one dimensional, for each $P \in Y_n$, there are only two possibilities:

- P is a special point, i.e. $\mathbf{MT}(P)$ is a torus (in which case it corresponds to a principally polarised abelian 4-fold with CM);
- $\mathbf{MT}(P) = \mathbf{MT}(Y_n)$, which is isogenous to some $\mathbb{Q}$-form of $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$.

Therefore, to conclude, we have to find a n and a non-special $P \in Y_n \cap \mathcal{T}_4^0$. Heading for a contradiction, suppose that, for all n, all points of $Y_n \cap \mathcal{T}_4^0$ are special. By density of $\mathrm{HL}(\mathcal{T}_4, \mathbf{H})$ in $\mathcal{T}_4$, this means that $\mathcal{T}_4$ contains a dense set of special points. André–Oort now implies that $\mathcal{T}_4$ is special, which is the contradiction we were looking for. This shows the existence of a point of $\mathcal{T}_4^0(\mathbb{C})$, corresponding to a curve $C/\mathbb{C}$ of genus 4 whose Jacobian has Mumford-Tate group isogenous to a $\mathbb{Q}$-form of the $\mathbb{C}$-group $\mathbb{G}_{\mathrm{m}} \times (\mathrm{SL}_2)^3$. To conclude the proof of Theorem 3.17 we just have to observe that $\mathcal{M}_4$, $\mathcal{A}_4$, j and all the Y_n can be defined over $\overline{\mathbb{Q}} \subset \mathbb{C}$ (see for example [25] or think about the moduli problem they are describing), and therefore all intersections we considered during the proof are defined over $\overline{\mathbb{Q}}$. The result follows.

[4] What makes Theorem 3.17 interesting is that Y is a special subvariety of $\mathcal{A}_4$ which is not of PEL type (while every special subvariety of $\mathcal{A}_2$ and $\mathcal{A}_3$ is necessary of PEL type).

Remark 11.5 The André–Oort conjecture of $\mathcal{A}_g$ is a deep theorem whose proof appears in [59, Theorem 1.3], and builds on the work of several people. We refer to [36] for the history of the proof. We note here that, for $1 < g \leq 6$, it is known by combining the Pila-Zannier strategy with the Ax–Lindemann conjecture and the Galois bound appearing in [58] (which is easier than the Galois bound needed in $\mathcal{A}_g$, for g arbitrary). This is explained in [60, Theorem 5.1] and in [49, Theorem 1.3].

Remark 11.6 (Oort's question) In [47, Sect. 7] it is asked whether one of Mumford's special curves $Y \subset \mathcal{A}_4$ can actually lie in $\mathcal{T}_4^0$. As a simple application of a result of Toledo [57] we show that there are no such Ys. In *op. cit.* it is proven that if a compact geodesic curve $Z \subset \mathcal{A}_g$ is contained in $\mathcal{T}_4^0$ then it has curvature $1/l$ with $1 < l \leq (g-1)/3$. If $g = 4$ it means that l, which is a priori an integer in the interval $(1, g)$, is both ≤ 1 and > 1. Since Mumford's curves are compact, this shows the desired claim.

Acknowledgements The authors would like to thank S. Tayou and N. Tholozan for pointing out a mistake related to Remark 3.15 in a previous version; S. Tayou and T. Kreutz for their careful reading and comments on a first version of this paper; C. Robles for pointing out a mistake in a previous version and sharing her thoughts on Sect. 7.1; B. Farb for sharing the reference [57]; and M. Green, P. Griffiths and C. Robles for their interest in this work. Finally we thank the referees for their careful reading, which improved the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. André, Y.: Mumford-Tate groups of mixed Hodge structures and the theorem of the fixed part. Compos. Math. **82**(1), 1–24 (1992)
2. Bader, U., Fisher, D., Miller, N., Stover, M.: Arithmeticity, superrigidity and totally geodesic submanifolds of complex hyperbolic manifolds. Invent. Math. **233**(1), 169–222 (2023)
3. Baily, W.L. Jr., Borel, A.: Compactification of arithmetic quotients of bounded symmetric domains. Ann. Math. (2) **84**, 442–528 (1966)
4. Bakker, B., Tsimerman, J.: The Ax–Schanuel conjecture for variations of Hodge structures. Invent. Math. **217**(1), 77–94 (2019)
5. Bakker, B., Klingler, B., Tsimerman, J.: Tame topology of arithmetic quotients and algebraicity of Hodge loci. J. Am. Math. Soc. **33**(4), 917–939 (2020)
6. Bakker, B., Brunebarbe, Y., Tsimerman, J.: o-minimal GAGA and a conjecture of Griffiths. Invent. Math. **232**(1), 163–228 (2023)
7. Baldi, G., Ullmo, E.: Special subvarieties of non-arithmetic ball quotients and Hodge theory. Ann. Math. **197**(1), 159–220 (2023)
8. Baldi, G., Klingler, B., Ullmo, E.: On the geometric Zilber–Pink theorem and the Lawrence–Venkatesh method. Expo. Math. **41**(3), 718–722 (2023)

9. Beauville, A.: Le groupe de monodromie des familles universelles d'hypersurfaces et d'intersections complètes. In: Complex Analysis and Algebraic Geometry (Göttingen, 1985). Lecture Notes in Math., vol. 1194, pp. 8–18. Springer, Berlin (1986)
10. Blázquez-Sanz, D., Casale, G., Freitag, J., Nagloo, J.: A differential approach to the Ax-Schanuel, I. arXiv e-prints (2021). arXiv:2102.03384
11. Borel, A.: Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem. J. Differ. Geom. **6**, 543–560 (1972)
12. Candelas, P., de la Ossa, X., Elmi, M., van Straten, D.: A one parameter family of Calabi-Yau manifolds with attractor points of rank two. J. High Energy Phys. **2020**(10), 202 (2020)
13. Carlson, J.A.: Bounds of the dimension of variations of Hodge structure. Trans. Am. Math. Soc. **294**, 45–64 (1986)
14. Carlson, J.A., Donagi, R.: Hypersurface variations are maximal. I. Invent. Math. **89**, 371–374 (1987)
15. Cattani, E., Deligne, P., Kaplan, A.: On the locus of Hodge classes. J. Am. Math. Soc. **8**(2), 483–506 (1995)
16. Chai, C.-L.: Density of members with extra Hodge cycles in a family of Hodge structures. Asian J. Math. **2**(3), 405–418 (1998)
17. Chiu, K.C.T.: Ax-Schanuel for variations of mixed Hodge structures. arXiv e-prints (2021). arXiv:2101.10968
18. Colombo, E., Pirola, G.P.: Some density results for curves with non-simple Jacobians. Math. Ann. **288**(1), 161–178 (1990)
19. Daw, C., Ren, J.: Applications of the hyperbolic Ax-Schanuel conjecture. Compos. Math. **154**(9), 1843–1888 (2018)
20. de Jong, A.J.: Beyond the André-Oort conjecture. 6 pages note, communicated to the second author in June 2021 (2004)
21. Deligne, P.: Théorie de Hodge. II. Publ. Math. Inst. Hautes Études Sci. **40**, 5–57 (1971)
22. Deligne, P.: La conjecture de Weil. I. Publ. Math. Inst. Hautes Études Sci. **43**, 273–307 (1973)
23. Deligne, P.: La conjecture de Weil. II. Publ. Math. Inst. Hautes Études Sci. **52**, 137–252 (1980)
24. Edixhoven, S.J., Moonen, B.J.J., Oort, F.: Open problems in algebraic geometry. Bull. Sci. Math. **125**(1), 1–22 (2001)
25. Faltings, G.: Arithmetic varieties and rigidity. In: Seminar on Number Theory, Paris 1982–83 (Paris, 1982/1983). Progr. Math., vol. 51, pp. 63–77. Birkhäuser, Boston (1984)
26. Gao, Z., Klingler, B.: The Ax-Schanuel conjecture for variations of mixed Hodge structures. Math. Ann. (2021). arXiv:2101.10938. https://doi.org/10.1007/s00208-023-02614-w
27. Green, M.L.: Infinitesimal methods in Hodge theory. In: Algebraic Cycles and Hodge Theory. Lectures Given at the 2nd Session of the Centro Internazionale Matematico Estivo (C.I.M.E.), Torino, Italy, June 21–29, 1993, pp. 1–92. Springer, Berlin (1994)
28. Green, M., Griffiths, P., Kerr, M.: Mumford-Tate Groups and Domains: Their Geometry and Arithmetic. Annals of Mathematics Studies, vol. 183. Princeton University Press, Princeton (2012).
29. Griffiths, P.A.: On the periods of certain rational integrals. I, II. Ann. Math. (2) **90**, 460–495, 496–541 (1969)
30. Griffiths, P.A.: Periods of integrals on algebraic manifolds. III: some global differential-geometric properties of the period mapping. Publ. Math. Inst. Hautes Études Sci. **38**, 125–180 (1970)
31. Habegger, P., Pila, J.: O-minimality and certain atypical intersections. Ann. Sci. Éc. Norm. Supér. (4) **49**(4), 813–858 (2016)
32. Izadi, E.: Density and completeness of subvarieties of moduli spaces of curves or Abelian varieties. Math. Ann. **310**(2), 221–233 (1998)
33. Kawamata, Y.: On Bloch's conjecture. Invent. Math. **57**, 97–100 (1980)
34. Klingler, B.: Hodge loci and atypical intersections: conjectures. In: Motives and Complex Multiplication (2017, to appear)
35. Klingler, B., Otwinowska, A.: On the closure of the Hodge locus of positive period dimension. Invent. Math. **225**(3), 857–883 (2021)
36. Klingler, B., Ullmo, E., Yafaev, A.: Bi-algebraic geometry and the André-Oort conjecture. In: Algebraic Geometry: Salt Lake City 2015. Proc. Sympos. Pure Math., vol. 97, pp. 319–359. Am. Math. Soc., Providence (2018)
37. Kostant, B.: Root systems for Levi factors and Borel-de Siebenthal theory. In: Symmetry and Spaces. In Honor of Gerry Schwarz on the Occasion of His 60th Birthday, pp. 129–152. Birkhäuser, Basel (2010)
38. Koziarz, V., Maubon, J.: Finiteness of totally geodesic exceptional divisors in Hermitian locally symmetric spaces. Bull. Soc. Math. Fr. **146**(4), 613–631 (2018)

39. Lawrence, B., Venkatesh, A.: Diophantine problems and p-adic period mappings. Invent. Math. **221**(3), 893–999 (2020)
40. Mok, N., Pila, J., Tsimerman, J.: Ax-Schanuel for Shimura varieties. Ann. Math. (2) **189**(3), 945–978 (2019)
41. Möller, M., Toledo, D.: Bounded negativity of self-intersection numbers of Shimura curves in Shimura surfaces. Algebra Number Theory **9**(4), 897–912 (2015)
42. Moonen, B.: Families of motives and the Mumford-Tate conjecture. Milan J. Math. **85**, 257–307 (2017)
43. Moonen, B., Oort, F.: The Torelli locus and special subvarieties. In: Handbook of Moduli, vol. II, pp. 549–594. Higher Education Press, Somerville; International Press, Beijing (2013)
44. Moore, G.: Arithmetic and attractors. arXiv e-prints (1998). arXiv:hep-th/9807087
45. Moore, G.W.: Les Houches lectures on strings and arithmetic. arXiv e-prints (2004). arXiv:hep-th/0401049
46. Mumford, D.: A note of Shimura's paper: discontinuous groups and Abelian varieties. Math. Ann. **181**, 345–351 (1969)
47. Oort, F.: Some questions in algebraic geometry. In: Open Problems in Arithmetic Algebraic Geometry. Based on the Presentations at the Conference 'Arithmetic and Geometry of Abelian Varieties'. With appendices by Frans Oort, 2019, pp. 263–283. Higher Education Press, Somerville; International Press, Beijing (1995)
48. Pila, J., Scanlon, T.: Effective transcendental Zilber-Pink for variations of Hodge structures. arXiv e-prints (2021). arXiv:2105.05845
49. Pila, J., Tsimerman, J.: Ax-Lindemann for $\mathcal{A}_g$. Ann. Math. (2) **179**(2), 659–681 (2014)
50. Pila, J., Zannier, U.: Rational points in periodic analytic sets and the Manin-Mumford conjecture. Atti Accad. Naz. Lincei, Rend. Lincei, Mat. Appl. **19**(2), 149–162 (2008)
51. Richard, R., Ullmo, E.: Equidistribution de sous-variétés spéciales et o-minimalité: André-Oort géométrique, with an appendix with J. Chen. Ann. Inst. Fourier (2021). arXiv:2104.04439
52. Robles, C.: Schubert varieties as variations of Hodge structure. Sel. Math. New Ser. **20**(3), 719–768 (2014)
53. Schmid, W.: Variation of Hodge structure: the singularities of the period mapping. Invent. Math. **22**, 211–319 (1973)
54. Simpson, C.T.: Higgs bundles and local systems. Publ. Math. Inst. Hautes Études Sci. **75**, 5–95 (1992)
55. Tayou, S.: On the equidistribution of some Hodge loci. J. Reine Angew. Math. **762**, 167–194 (2020)
56. Tayou, S., Tholozan, N.: Equidistribution of Hodge loci. II. Compos. Math. **159**(1), 1–52 (2023)
57. Toledo, D.: Non-existence of certain closed complex geodesics in the moduli space of curves. Pac. J. Math. **129**(1), 187–192 (1987)
58. Tsimerman, J.: Brauer-Siegel for arithmetic tori and lower bounds for Galois orbits of special points. J. Am. Math. Soc. **25**(4), 1091–1117 (2012)
59. Tsimerman, J.: The André-Oort conjecture for $\mathcal{A}_g$. Ann. Math. (2) **187**(2), 379–390 (2018)
60. Ullmo, E.: Applications du théorème d'Ax-Lindemann hyperbolique. Compos. Math. **150**(2), 175–190 (2014)
61. van den Dries, L.: Tame Topology and O-Minimal Structures. London Mathematical Society Lecture Note Series, vol. 248. Cambridge University Press, Cambridge (1998)
62. Voisin, C.: Théorie de Hodge et Géométrie Algébrique Complexe. Cours Spécialisés [Specialized Courses], vol. 10. Société Mathématique de France, Paris (2002)
63. Weil, A.: Abelian varieties and the Hodge ring. Collect. Pap. **III**, 421–429 (1979)