The Cayley hyperbolic space and volume entropy rigidity

Yuping Ruan

Received: 27 April 2022 / Accepted: 9 October 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract
Let $M$ be a Riemannian manifold with dimension greater than or equal to 3 which admits a complete, finite-volume Riemannian metric $g_0$ locally isometric to a rank one symmetric space of non-compact type. The volume entropy rigidity theorem (Besson et al. in Geom Funct Anal 5(5), 731–799, 1995, Theorémè principal) asserts that $g_0$ minimizes a normalized volume growth entropy among all complete, finite-volume, Riemannian metric on $M$. We will repair a gap in the proof when $g_0$ is locally isometric to the Cayley hyperbolic space.

Keywords Volume growth entropy · The Cayley hyperbolic space · Rigidity · Busemann functions

Contents
1 Introduction ................................................
2 Octonions and the Cayley hyperbolic space ................................
3 An overview of Besson–Courtois–Gallot’s approach to volume entropy rigidity in rank one symmetric spaces of non-compact type ........................................
4 The gap in the proof of Theorem 1.1 from [3] and its repair .......................
  4.1 The gap in the proof of Theorem 1.1 from [3] regarding the Cayley hyperbolic space .......
  4.2 Repairing the gap: reducing a weaker version of (8) into a linear algebra problem ........
  4.3 Repairing the gap: a linear algebra lemma ..............................
  4.4 Repairing the gap: proof of (8) ....................................
References ...................................................

1 Introduction
Let $h(g)$ be the volume growth entropy of a closed $n$-dimensional Riemannian manifold $(M, g)$, i.e.

$$h(g) = \lim_{r \to \infty} \frac{1}{r} \ln(\text{Vol}(B(x, r))).$$
where $B(x, r)$ denotes the ball of radius $r$ centered at some point $x$ in the universal cover of $(M, g)$. The definition of $h(g)$ is independent of the choice of $x$. When $M$ is not compact but has finite volume, we define $h(g)$ instead by

$$ h(g) = \inf \left\{ s \geq 0 \mid \exists K > 0 \text{ s.t. } \int_{\tilde{M}} e^{-sd(x, z)} dg(z) < K \right\}, $$

where $dg$ denotes the volume form on the universal cover $\tilde{M}$ of $M$ induced by the metric $g$. We introduce the normalized entropy

$$ \text{ent}(g) = \frac{h(g)}{\text{Vol}(M, g)} $$

so that it remains invariant under scaling. In [3] Besson–Courtois–Gallot proved the following theorem.

**Theorem 1.1** (Besson–Courtois–Gallot, Théorème principal, [3]) Let $M$ be a Riemannian manifold of dimension $n$ which admits a complete, finite-volume Riemannian metric $g_0$ locally isometric to a rank one symmetric space of non-compact type with dimension at least 3. Let $g$ be another complete, finite-volume Riemannian metric on $M$. Assuming that $(\tilde{M}, g)$, the universal cover of $(M, g)$, has bounded geometry (see the remark below), then

$$ \text{ent}(g) \geq \text{ent}(g_0) \quad (1) $$

with equality achieved if and only if the map $\text{Id} : (M, g) \to (M, g_0)$ is homotopic to an isometry.

More generally, if $(N, g)$ is any $n$-dimensional finite-volume Riemannian manifold of bounded geometry, and if $f : N \to M$ is any proper, smooth map with non-zero degree, then

$$ \text{ent}(N, g) \geq |\text{deg } f| \text{ent}(M, g_0) $$

with equality achieved if and only if $f$ is homotopic to a Riemannian covering.

**Remark** (1). A Riemannian manifold $X$ has bounded geometry if for any $\epsilon > 0$, there exists some constant $C > 0$ such that any geodesic ball of radius $\epsilon$ has volume greater or equal to $C$;

(2). The above theorem is slightly different from [3, Théorème principal] because Besson–Courtois–Gallot assumed that $M$ and $N$ are compact. The non-compact case requires a few extra arguments and was proved by Boland–Connell–Souto in [5, Theorem 1.3]. Also, Storm removed the bounded geometry condition in [13, Theorem 1.1].

We noticed a technical gap in the proof of the above theorem when $(M, g_0)$ is locally isometric to the Cayley hyperbolic space $\mathbb{O}H^2$. The gap appears in the assertion (see [3, page 751]) which claims that there exist fiber-wise linear maps $J_k : T\tilde{M} \to T\tilde{M}$, $k = 1, \ldots, 7$ such that the Hessian of the Busemann function satisfies the following formula

$$ \text{Hess} B_\theta |_x (v, w) = g_0(v, w) - dB_\theta(v)dB_\theta(w) + \sum_{k=1}^{7} dB_\theta(J_k v)dB_\theta(J_k w), \quad x \in M, v, w \in T_x\tilde{M}, $$

where $\tilde{M}$ denotes the universal cover of $(M, g_0)$ and $B_\theta$ is the Busemann function with respect to the geodesic ray starting at a fixed point towards a point $\theta$ in the boundary at infinity $\partial_\infty \tilde{M}$ of $\tilde{M}$. We observe that such maps $J_k$ do not exist in the Cayley hyperbolic case due to the non-associativity of octonionic multiplication, which makes this case different from other

\[ \text{Springer} \]
rank one symmetric spaces of non-compact type. The goal of this paper is to repair this gap so that Theorem 1.1 still holds in the Cayley hyperbolic space, thus holds for all rank one symmetric spaces.

In Sect. 2, we will introduce the “matrix model” and the “vector model” for the Cayley hyperbolic space. The “vector model” has appeared in multiple resources (See e.g. [9]). It is an analogue of the Minkowski model for the real hyperbolic space. The “matrix model” is a variant of the model in [8, §19. Spaces of $\mathbb{R}$-rank one] which is convenient for us to define the distance function. Towards the end of Sect. 2, we will introduce Cayley lines as analogues of the complex lines and explain their relations with the curvature data in the “vector model” for the Cayley hyperbolic space. These computations lead to a gap in [3, page 751], i.e. the aforementioned maps $J_k$ do not exist. This gap is related to some kind of incompatibility between the sectional curvature data and multiplying vectors by purely imaginary octonions due to the non-associativity of octonionic multiplication. To be specific, we compare the complex hyperbolic space $\mathbb{C}H^n$ with the Cayley hyperbolic space $\mathbb{O}H^2$ (assuming their sectional curvatures take values in $[-4, -1]$). For any unit vector $v \in T\mathbb{C}H^n$, the sectional curvature between $v$ and $iv$ is always $-4$ because they lie on the same complex line. However, for any unit vector $v \in T\mathbb{O}H^2$ and any purely imaginary unit octonion $e$, the sectional curvature between $v$ and $ev$ can vary because they are not necessarily on the same Cayley line. Nevertheless, the computations in this section show that there are still a lot of similarities between the Cayley structure on $\mathbb{O}H^2$ and the complex structure on $\mathbb{C}H^n$.

Section 3 is an outline of Besson–Courtois–Gallot’s original proof of Theorem 1.1 and Sect. 4.1 points out the gap in their proof in the case of Cayley hyperbolic space. This gap is located in the technical part of Besson–Courtois–Gallot’s argument regarding the following inequality

$$\det \left( \int_{\partial_\infty \tilde{M}} (dB_\theta)^2(x) d\mu(\theta) \right)^{1/2} \leq \left( \frac{\sqrt{16}}{16 + 8 - 2} \right)^{16}.$$ 

Here $\partial_\infty \tilde{M}$ is the boundary at infinity (diffeomorphic to a sphere of dimension $\dim M - 1$) of $\tilde{M}$ and $\mu$ is a probability measure in the Lebesgue measure class of $\partial_\infty \tilde{M}$. We will repair this gap in Sects. 4.2–4.4.

The main idea to repair the gap is to control the “bad outcome” caused by the loss of the aforementioned $J_k$ maps. Section 4.2 reduces the above inequality to a pure linear algebra problem. To be specific, we first notice that we can define $J_k$ maps on each Cayley line (but not the entire tangent space!) which behave nicely with respect to curvature. This enables us to find a suitable basis on $T_x \tilde{M}$ such that the symmetric bilinear forms in both the numerator and the denominator have matrix forms almost as nice as the complex (or quaternionic) hyperbolic setting. Therefore we can translate the geometric information into certain properties on matrices and use these properties instead for the rest of the proof. In Sect. 4.3, we prove Lemma 4.4 which deals with the linear algebra problem from Sect. 4.2. This is the main part and the most technical part of this paper. We first observe that we can make small adjustments to Besson–Courtois–Gallot’s proof in [3, Appendice B] so that the inequality holds when eigenvalues of these matrices are relatively close to their averages respectively. This corresponds to Case 1 in Lemma 4.4 considered as the “good” case. The complementary case is much more unpleasant to work with and hence considered as the “bad” case (see Case 2 in Lemma 4.4). In this case, some eigenvalues of the matrices are extremely large compared to other eigenvalues. Although the proof in this case is very technical, the overall idea behind is much simpler. By [3, B.4 Lemme], the extreme eigenvalues will actually help us in the
proof. Therefore the general idea in this case is to show that these extreme eigenvalues help us more than the possible trouble created by the loss of the aforementioned $J_k$ maps. Section 4.4 proves a slightly stronger inequality which is also used in Besson–Courtois–Gallot’s proof (see [3, B.5 Proposition]).

2 Octonions and the Cayley hyperbolic space

This section is a very brief review on octonions and the Cayley hyperbolic space. A more detailed version can be found in [10, Appendix A].

The set of octonions $O$ is an 8-dimensional non-associative, non-commutative division algebra over $\mathbb{R}$. Let $(\cdot, \cdot)$ be the Euclidean inner product on $O$ and $| \cdot |$ the induced norm. Then we have the following properties. (See [12])

1. $|ab| = |a||b|$ for any $a, b \in O$;
2. $\langle ab, ac \rangle = \langle ba, ca \rangle = |a|^2 \langle b, c \rangle$ for any $a, b, c \in O$;
3. $\langle ac, bd \rangle + (\langle ad, bc \rangle = 2\langle a, b \rangle \langle c, d \rangle$ for any $a, b, c, d \in O$;
4. $\overline{a} = 2\langle a, 1 \rangle - a$ for any $a \in O$. We call $\overline{a}$ the conjugate of $a$;
5. $2\langle a, b \rangle = 2\langle \overline{a}, \overline{b} \rangle = \overline{ab} + \overline{ba} = \overline{a}b + b\overline{a}$ for any $a, b \in O$;
6. $(ba)\overline{a} = \overline{a}(ab) = |a|^2 b$ for any $a, b \in O$;
7. $a(\overline{bc}) + b(\overline{ac}) = (\overline{ca})b + (\overline{cb})a = \langle a, b \rangle c$ for any $a, b, c \in O$;
8. (Moufang Identities)
   1. $(ab)(ca) = a((bc)a)$ for any $a, b, c \in O$;
   2. $a(b(ac)) = (a(ba))c$ for any $a, b, c \in O$;
   3. $b(a(ca)) = ((ba)c)a$ for any $a, b, c \in O$;
9. Multiplications involving only two octonions are associative.

Let

$$I_{1,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and

$$\mathcal{J}(1, 2, O) = \left\{ X \in \text{Mat}_{3 \times 3}(O) : I_{1,2}X^*I_{1,2} = X \right\},$$

where $X^*$ denotes the conjugate transpose of $X$. Any element $X \in \mathcal{J}(1, 2, O)$ can be written in the following form

$$X(\theta, a) = \begin{pmatrix} \theta_1 & a_3 & \overline{a}_2 \\ -\overline{a}_3 & -\theta_2 & -a_1 \\ -a_2 & -\overline{a}_1 & -\theta_3 \end{pmatrix}, \quad \theta_j \in \mathbb{R}, a_j \in \mathbb{O}, j = 1, 2, 3,$$

where $\theta = (\theta_1, \theta_2, \theta_3)$ and $a = (a_1, a_2, a_3)$.

**Definition 2.1** (“Matrix model”) We define the Cayley hyperbolic space $OH^2$ as

$$OH^2 = \left\{ X \in \mathcal{J}(1, 2, O) : X^2 = X, \text{tr}(X) = 1, X_{11} > 0 \right\}.$$

**Remark** The way we define the Cayley hyperbolic space using matrices is slightly different from that in [8, §19 Spaces of $\mathbb{R}$-rank one], but one can prove that they are equivalent. See [10, Proposition A.4] for a detailed argument.
Proposition 2.2 For any trace 1 idempotent \( \mathcal{X} \in \mathcal{J}(1, 2, \mathbb{O}) \) with \( \mathcal{X}_{11} \neq 0 \), there exists a unique vector \( (\theta, a, b) \in \mathbb{R}_+ \times \mathbb{O}^2 \) such that
\[
\mathcal{X} = \text{sgn}(\mathcal{X}_{11}) I_{1,2}(\theta, a, b)^*(\theta, a, b),
\]
where \( \text{sgn}(t) = t/|t| \) when \( t \in \mathbb{R} \setminus \{0\} \). The set
\[
\mathcal{J}_{1,0} := \{ \mathcal{X} \in \mathcal{J}(1, 2, \mathbb{O}) : \mathcal{X}^2 = \mathcal{X}, \text{tr}(\mathcal{X}) = 1, \mathcal{X}_{11} = 0 \}
\]
is isomorphic to \( \mathbb{O}\mathbb{P}^1 \cong S^8 \).

Proof See [10, Proposition A.2].

Therefore we have the following alternative definition for the Cayley hyperbolic space (also see [9, page 87]).

Definition 2.3 (“Vector model”) The Cayley hyperbolic space can be alternatively defined as
\[
\mathcal{O}\mathbb{H}^2 = \{(\theta, a, b) \in \mathbb{R}_+ \times \mathbb{O}^2 : \theta^2 - |a|^2 - |b|^2 = 1 \}.
\]

Remark One can easily check that for any \( \mathcal{X} \in \mathcal{O}\mathbb{H}^2, \mathcal{X}_{11} \geq 1 \).

We refer to [8, §19. Spaces of \( \mathbb{R} \)-rank one], [10, Appendix A] and [2] for an overview of the related subjects. A very detailed and general theory can be found in [11, 12] by Springer and Veldkamp for further reference.

Denote by \( g_0 \) the symmetric metric on \( \mathcal{O}\mathbb{H}^2 \) and \( M_0 = (\mathcal{O}\mathbb{H}^2, g_0) \) such that the distance function on \( \mathcal{O}\mathbb{H}^2 \) is given by
\[
\cosh(2d(\mathcal{X}, \mathcal{Y})) = \text{tr}(\mathcal{X}\mathcal{Y} + \mathcal{Y}\mathcal{X}) - 1, \quad \forall \mathcal{X}, \mathcal{Y} \in \mathcal{O}\mathbb{H}^2
\]
as in [8]. \( (\mathcal{X}\mathcal{Y} + \mathcal{Y}\mathcal{X})/2 =: \mathcal{X} \circ \mathcal{Y} \) is also known as the Jordan multiplication of \( \mathcal{X} \) and \( \mathcal{Y} \).

Let \( x_0 = (1, 0, 0) \in M_0 \) using the “vector model”. Then we can identify \( T_{x_0}M_0 \) with \( \mathbb{O}^2 \) such that each unit vector \( v = (a, b) \in \mathbb{O}^2 \) corresponds to the initial vector of
\[
\gamma_v(t) = \gamma_{(a,b)}(t) := (\cosh(t), a \sinh(t), b \sinh(t))
\]
(2)

We will compute the Riemannian curvature data at \( x_0 \) by understanding the geodesic hinge \( \angle X_{x_0}Y \) for any \( X, Y \in M \setminus \{x_0\} \), where \( \angle X_{x_0}Y \) consists of two geodesic segments \( X_{x_0}, Y_{x_0} \) and the angle \( \angle X_{x_0}Y \). A comparison hinge of \( \angle X_{x_0}Y \) in some space form \( M' \) is a geodesic hinge \( \angle X'_{x_0}Y' \) in \( M' \) with the same angle such that the lengths of geodesic segments \( X_{x_0}, Y_{x_0} \) and \( X'_{x_0}, Y'_{x_0} \) are equal respectively.

Proposition 2.4 The following hold for the Cayley hyperbolic space.

(1) For any unit vector \( v = (a, b) \in \mathbb{O}^2 \), \( \gamma_v(t) \) gives a unit speed geodesic starting at \( x_0 \) with initial vector \( v \);

(2) The Riemannian metric at \( x_0 \) is given by the Euclidean inner product on \( \mathbb{O}^2 \). Hence the map \( \chi : \mathbb{O}^2 \to M \) such that
\[
\chi(a, b) = \left( \cosh((|a, b|)), \frac{a}{|a, b|} \sinh(|(a, b)|), \frac{b}{|a, b|} \sinh(|(a, b)|) \right)
\]
gives the geodesic normal coordinates centered at \( x_0 \) (hence \( d\chi := (d\chi)|_{(0,0)} : \mathbb{O}^2 \to T_{x_0}M_0 \) is the isometric correspondence from \( \mathbb{O}^2 \) with Euclidean inner product to \( T_{x_0}M_0 \) mentioned above)
Corollary 2.5

For any non-parallel pair of non-zero unit vectors \( (a, b), (c, d) \) contained in the same Cayley line and any pair of points \( X \in \gamma(a, b)(\mathbb{R}_+) \) and \( Y \in \gamma(c, d)(\mathbb{R}_+) \), (where \( \gamma(a, b) \) and \( \gamma(c, d) \) are defined in (2),) the comparison hinge \( \triangle X'Y_0Z' \) of \( \triangle X_0Y_0Z_0 \) in a space form of constant sectional curvature \(-4\) satisfies \( d(X', Y') = d(X, Y) \);

(4) For any \( a, b \in \mathbb{O} \setminus \{0\} \), any pair of points \( X \in \gamma(a, 0)(\mathbb{R}_+) \) and \( Y \in \gamma(0, b)(\mathbb{R}_+) \), the comparison hinge \( \triangle X'Y_0Z' \) of \( \triangle X_0Y_0Z_0 \) in a space form of constant sectional curvature \(-1\) satisfies \( d(X', Y') = d(X, Y) \).

**Proof**

The proof can be found in [10, Proposition A.5].

A direct corollary of the above proposition is the following.

**Corollary 2.5**

For any non-zero \( v, w \in \mathbb{O}^2 \) the following hold.

1. If \( v, w \) belong to the same Cayley line and \( v \neq Rw \), then the sectional curvature of the 2-dimensional plane spanned by \( dX(v), dX(w) \) is \(-4\);
2. If \( \text{Cay}(v) \perp \text{Cay}(w) \), then the sectional curvature of the 2-dimensional plane spanned by \( dX(v), dX(w) \) is \(-1\).
3. Recall that from classical results that \( \mathbb{O}^2 = F_4^{-20} / \text{Spin}(9) \) with \( \text{Spin}(9) \) the stabilizer of \( x_0 \). Since \( F_4^{-20} \) acts by isometries on \( M_0 = (\mathbb{O}^2, g_0) \), identifying \( TX_0M \) with \( \mathbb{O}^2 \), one can view \( \text{Spin}(9) \) as a subgroup of \( \text{SO}(\mathbb{O}^2) \cong \text{SO}(16) \). In particular, \( \text{Spin}(9) \) maps Cayley lines to Cayley lines and acts transitively on the set of all Cayley lines.

**Remark**

(1) By (3), one can easily check that \( \text{Cay}(1, a) \perp = \text{Cay}(1, -a/|a|^2) \) for any \( a \in \mathbb{O} \setminus \{0\} \) and \( \text{Cay}(1, 0) \perp = \text{Cay}(0, 1) \in \mathbb{O}^2 \). Therefore the orthogonal complement of any Cayley line in \( \mathbb{O}^2 \) is a Cayley line in \( \mathbb{O}^2 \). Hence Corollary 2.5 determines all possible sectional curvature data for 2-dimensional planes in \( TX_0M = (dX)^{-1} \mathbb{O}^2 \). In particular, for any pair of non-zero vectors \( v, w \in \mathbb{O}^2 \) satisfying \( v \neq Rw \), the sectional curvature of the 2-dimensional plane spanned by \( dX(v), dX(w) \) is \(-4\) if and only if \( \text{Cay}(v) = \text{Cay}(w) \);

(2) Let \( p \) be a fixed point in \( \mathbb{CH}^n \). Then there exists a map \( J : T_p\mathbb{CH}^n \rightarrow T_p\mathbb{CH}^n \) satisfying the following two conditions:

- \( J \) is \( \mathbb{R} \)-linear and invertible;
- For any unit vector \( v \in T_p^1\mathbb{CH}^n \) with respect to the symmetric metric, the sectional curvature between \( v \) and \( J(v) \) is \(-4\).

In fact, we can simply let \( J(v) = vi \) for any \( v \in T_p\mathbb{CH}^n \). The same argument holds true in quaternionic hyperbolic spaces but **NOT** in the Cayley hyperbolic space. This is due to the non-associativity of octonionic multiplication. In fact, if such a map \( J \) exists, without loss of generality we can assume that \( p = x_0 \) and identify the tangent space at \( x_0 \) with \( \mathbb{O}^2 \) via \( dX \) as in Proposition 2.4.

Since we assumed that the sectional curvature between \( v \) and \( J(v) \) is \(-4\), by the previous remark and (3), \( J(1, 0) \in \text{Cay}(1, 0) = \mathbb{O} \cdot (1, 0) \) and \( J(v) \) is not parallel to \( (1, 0) \). Hence \( J(1, 0) = (a, 0) \) for some \( a \in \mathbb{O} \setminus \mathbb{R} \). Similarly, \( J(0, b) \in \text{Cay}(0, b) \).
null

\[ \bigcirc \cdot (0, 1) \text{ for any } b \in \bigcirc \setminus \{0\}. \]

As a corollary of the above discussions and \( \mathbb{R} \)-linearity of \( J \), for any octonion \( b \) and any \( \theta \in [0, \pi/2) \),

\[ J(\cos \theta, b \sin \theta) - (a \cos \theta, 0) = J(0, b \sin \theta) \in \bigcirc \cdot (0, 1). \]

On the other hand, by (3), \( J(\cos \theta, b \sin \theta) \in \text{Cay}(\cos \theta, b \sin \theta) = \bigcirc \cdot (1, b \tan \theta) = \bigcirc \cdot (\cos \theta, b \sin \theta) \). Therefore, \( J(\cos \theta, b \sin \theta) = a(\cos \theta, b \sin \theta) \). In particular, \( J(0, 1) = (0, a) \).

One can apply a similar argument to \( J(b \sin \theta, \cos \theta) \) and obtain \( J(b \sin \theta, \cos \theta) = a(b \sin \theta, \cos \theta) \) for any octonion \( b \) and any \( \theta \in [0, \pi/2) \), which implies that \( J(v) = av \) for any \( v \in \bigcirc^2 \). Therefore by the previous remark and (3), for any octonions \( b \neq 0 \) and \( c \), we have \( J(b, bc) = (ab, a(bc)) \in \text{Cay}(b, bc) = \text{Cay}(1, c) \). This implies that \( (ab)c = a(bc) \). Notice that if \( (ab)c = a(bc) \) for any \( b, c \in \bigcirc \), then \( a \) must be real. This contradicts the assumption that \( a \in \bigcirc \setminus \mathbb{R} \).

Let \( \mathcal{A} \) be a maximal abelian subgroup of \( F_{-20}^4 \) and \( F_{-20}^4 = \mathcal{KAN} \) be a corresponding Iwasawa decomposition of \( F_{-20}^4 \). Denote by \( v_{l,t} = (\delta_{1l}e_r, \delta_{2l}e_r) \) with \( l = 1, 2, \delta_{lm} \) the Kronecker delta, \( 0 \leq t \leq 7 \) and \( \{e_l\}_{0 \leq l \leq 7} \) the standard orthonormal basis for \( \bigcirc \). Then we can construct \( \mathcal{A}N \)-invariant vector fields \( \tilde{\xi}_l,t \) such that

\[ \tilde{\xi}_l,t(x_0) = \Psi(v_{l,t}) \quad \forall l = 1, 2 \text{ and } 0 \leq t \leq 7, \]

where \( \Psi \) is the same as \( d_X \) in Proposition 2.4.

Define \( \Psi_{\xi} : \mathbb{O}^2 \rightarrow T_x M_0 \) such that

\[ \Psi_{\xi}(v_{l,t}) = \tilde{\xi}_l,t(x), \quad \forall l = 1, 2 \text{ and } 0 \leq t \leq 7. \]

For simplicity we denote by

\[ \text{Cay}(\xi) = \Psi_{\xi} \left( \text{Cay}(\Psi^{-1}_x(\xi)) \right), \quad \forall \xi \in T_x M_0 \setminus \{0\} \]

the Cayley line containing \( \xi \). Since the isometry group acts transitively on the Cayley hyperbolic space, the same arguments as in Corollary 2.5 can be applied to all points in \( M_0 \). Namely,

**Corollary 2.6** For any \( x \in M_0 = (\bigcirc \mathbb{H}^2, g_0) \) and any non-zero \( \tilde{\xi}, \tilde{\xi}' \in T_x M_0 \),

1. If \( \tilde{\xi}, \tilde{\xi}' \) belong to the same Cayley line and \( \tilde{\xi} \notin \mathbb{R}\tilde{\xi}' \), then the sectional curvature \( K_{M_0}(\tilde{\xi}, \tilde{\xi}') = -4 \);
2. If \( \text{Cay}(\tilde{\xi}) \perp \text{Cay}(\tilde{\xi}') \), then the sectional curvature \( K_{M_0}(\tilde{\xi}, \tilde{\xi}') = -1 \).

### 3 An overview of Besson–Courtois–Gallot’s approach to volume entropy rigidity in rank one symmetric spaces of non-compact type

Following the notations in Theorem 1.1, we denote by \( (\widetilde{N}, g) \) and \( (\widetilde{M}, g_0) \) the universal (Riemannian) covers of \( (N, g) \) and \( (M, g_0) \) respectively. Let \( \Gamma \) be the fundamental group of \( M \). For any \( x \in \widetilde{M} \), denote by \( \mu_x \) the Patterson–Sullivan measure (see [1]) at \( x \) for \( \Gamma \).

In this particular case, \( \mu_x \) is also the probability measure on the boundary at infinity \( \partial_{\infty} \widetilde{M} \) invariant under \( \text{Stab}_{\text{Isom}(\widetilde{M})}(x) \). Notice that in rank one symmetric spaces, \( \text{Stab}_{\text{Isom}(\widetilde{M})}(x) \) can be viewed as a subgroup of \( \text{SO}(T_x \widetilde{M}) \) and acts transitively on \( T_x^1 \widetilde{M} \), one can view \( \mu_x \) as the push forward of the Haar measure on \( T_x^1 \widetilde{M} \) to the boundary at infinity via opposite geodesic rays centered at \( x \). For any \( s > h(g) \) we define a family of probability measures \( \{\sigma^x_y\}_{y \in \widetilde{N}} \) on \( \partial_{\infty} \widetilde{M} \) as

\[ \sigma^x_y(U) = \frac{\int_{\widetilde{N}} e^{-sd_x(y,z)} \mu_\tilde{g}(z) dg(z)}{\int_{\widetilde{N}} e^{-sd_x(y,z')} dg(z')}, \quad \forall y \in \widetilde{N}, U \subset \partial_{\infty} \widetilde{M}, \]
where \( \tilde{f} : \tilde{N} \to \tilde{M} \) is a lift of \( f \) to universal covers with respect to a pair of base points \( p \in N \) and \( f(p) \in M \).

Fix a point \( x_0 \in \tilde{M} \), we define the Busemann function as

\[
B_\theta(x) = B_\theta(x, x_0) = \lim_{t \to \infty} (d(x, \gamma_\theta(t)) - t), \quad \forall x \in M, \theta \in \partial_\infty M.
\]

where \( \gamma_\theta(t) \) is the geodesic ray starting from \( x_0 \) towards \( \theta \). Then we define a \( C^1 \)-map (due to implicit function theorem. See [4, pages 635-636].) \( \tilde{F}_s : \tilde{N} \to \tilde{M} \) such that \( \tilde{F}_s(y) \) is the unique critical point of the following function

\[
B_{s,y}(x) := \int_{\partial_\infty \tilde{M}} B_\theta(x)d\sigma_\gamma^s(\theta),
\]

where uniqueness comes from the convexity of Busemann functions and that \( \sigma_\gamma^s \) is in the same measure class as the Haar measure. It is not hard to see that the definition of \( B_{s,y} \) is independent of the choice of \( x_0 \) up to an additive constant and the map \( \tilde{F}_s \) naturally descends to a map \( F_s : N \to M \). Also when \( s \) goes to \( \infty \), \( F_s \) goes to \( f \), which implies that \( F_s \) is homotopic to \( f \). Direct computation shows that for any \( v \in T_{\tilde{F}_s(y)}\tilde{M} \) and any \( w \in T_y\tilde{N} \), we have

\[
0 = D_yD_x=\tilde{F}_s(y)B_{s,y}(x)(v, w) = \int_{\partial_\infty \tilde{M}} \text{Hess}B_\theta|_{\tilde{F}_s(y)}(D_y\tilde{F}_s(w), v)d\sigma_\gamma^s(\theta)
- s \int_{\tilde{N}} \int_{\partial_\infty M} dB_\theta|_{\tilde{F}_s(y)}(v)\langle \nabla_y d_\theta(y, z), w \rangle \frac{e^{-sd_\theta(y, z)}}{\int_{\tilde{N}} e^{-sd_\theta(y, z)}dg(z)} d\mu_{\tilde{F}_s(y)}(\theta)dg(z).
\]

Hence as matrices

\[
D_y\tilde{F}_s = sW_y^{-1} \cdot H_y,
\]

where \( W_y \) and \( H_y \) are matrices defined by bilinear forms on \( T_{\tilde{F}_s(y)}\tilde{M} \times T_{\tilde{F}_s(y)}\tilde{M} \) and \( T_y\tilde{N} \times T_{\tilde{F}_s(y)}\tilde{M} \) respectively such that

\[
W_y(v', v) = \int_{\partial_\infty \tilde{M}} \text{Hess}B_\theta|_{\tilde{F}_s(y)}(v', v)d\sigma_\gamma^s(\theta)
\]

and

\[
H_y(w, v) = \int_{\tilde{N}} \int_{\partial_\infty M} \frac{e^{-sd_\theta(y, z)}dB_\theta|_{\tilde{F}_s(y)}(v)\langle \nabla_y d_\theta(y, z), w \rangle}{\int_{\tilde{N}} e^{-sd_\theta(y, z)}dg(z')}d\mu_{\tilde{F}_s(y)}(\theta)dg(z)
\]

for any \( w \in T_y\tilde{N} \) and \( v, v' \in T_{\tilde{F}_s(y)}\tilde{M} \).

In particular, by Hölder’s inequality we have

\[
|\text{tr}H_y| \
\leq \left[ \text{tr} \left( \int_{\partial_\infty \tilde{M}} dB^2_\theta|_{\tilde{F}_s(y)}d\sigma_\gamma^s(\theta) \right) \text{tr} \left( \int_{\tilde{N}} \frac{\langle \nabla_y d_\theta(y, z), w \rangle^2 e^{-sd_\theta(y, z)}}{\int_{\tilde{N}} e^{-sd_\theta(y, z)}dg(z')}d\mu_{\tilde{F}_s(y)}(\theta)dg(z) \right) \right]^{\frac{1}{2}}
\leq 1
\]

(5)
and
\[
| \det H_y | \\
\leq \det \left( \int_{\partial M} B_{\theta}^2 |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta) \right)^{1/2} \det \left( \int \frac{\langle \nabla_y d g(y, z), w \rangle^2 e^{-sd_{g}(y, z)}}{\int_{N} e^{-sd_{g}(y, z')} d g(z')} d g(z) \right)^{1/2} \\
\leq \left( \frac{1}{\sqrt{n}} \right)^n \det \left( \int_{\partial M} B_{\theta}^2 |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta) \right)^{1/2}.
\]
Therefore
\[
| \text{Jac} F_\theta(y) | \leq \left( \frac{s}{\sqrt{n}} \right)^n \frac{\det \left( \int_{\partial M} B_{\theta}^2 |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta) \right)^{1/2}}{\det \left( \int_{\partial M} \text{Hess} B_{\theta} |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta) \right)}.
\]
In order to prove the inequality (1), it suffices to show that the right hand side of (7) is smaller than or equal to \((s/h(g_0))^n\). Note that in the non-compact case, one also needs to verify that \(F_\theta\) is proper. Moreover
\[
\left( \frac{s}{\sqrt{n}} \right)^n \frac{\det \left( \int_{\partial M} B_{\theta}^2 |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta) \right)^{1/2}}{\det \left( \int_{\partial M} \text{Hess} B_{\theta} |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta) \right)} \leq \left( \frac{s}{h(g_0)} \right)^n \left[ 1 - K \sum_{j=1}^{n} \left( v_j - \frac{1}{n} \right)^2 \right]
\]
for some positive constant \(K\), where \(v_1, \ldots, v_n\) are eigenvalues of \(\int_{\partial M} B_{\theta}^2 |F_\theta(y) dy | d\sigma^\gamma_{\theta}(\theta)\). This is also the part where the proof in [3] has a gap. We will explain this in Sect. 4.1.

When equality holds in (1), assuming \(h(g) = h(g_0)\) after scaling, [3, 7.6 Lemme] and [3, 7.8 Lemme] imply that there exists a sequence \(s_k \to h(g)^+\) such that \(F_{s_k}\) converge to a 1-Lipschitz map \(F : N \to M\) uniformly on compact subsets. Note that in the non-compact case, \(F\) is also a proper map following [5, Proposition 5.1] or [6, Lemma 6.7]. Then it follows from [3, C.1 Proposition] that \(F\) is a local isometry. Since we can still prove (8) in the Cayley hyperbolic setting (see Corollary 4.9), the whole argument for the equality case in (1) still works for the Cayley hyperbolic case without any changes.

### 4 The gap in the proof of Theorem 1.1 from [3] and its repair

#### 4.1 The gap in the proof of Theorem 1.1 from [3] regarding the Cayley hyperbolic space

Following the notations from the previous sections, we recall that the Busemann function satisfies
\[
B_{\theta}(x) = b_{\theta}(x, x_0) = \lim_{t \to \infty} (d(x, y_{\theta}(t)) - t), \quad \forall x \in \tilde{M}, \theta \in \partial \tilde{M}.
\]
where \(y_{\theta}(t)\) is the geodesic ray starting from \(x_0\) towards \(\theta\). In [7, page 36] we have
\[
\text{Hess} B_{\theta}(x) = \sqrt{-R(\text{grad} B_{\theta}, \cdot, \text{grad} B_{\theta}, \cdot)},
\]
where \(R\) denotes the Riemannian curvature tensor.

In [3, page 751], Besson–Courtois–Gallot claimed that there exist fiberwise linear maps \(J_l : TM \to T\tilde{M}\) with \(1 \leq l \leq d - 1\), \((d = \dim_{\mathbb{R}} \odot = 8\) in the Cayley hyperbolic case.)
such that

\[
\text{Hess}B_\theta(x) = g_0 - dB_\theta^2 + \sum_{l=1}^{d-1} (dB_\theta \circ J_l)^2.
\] (10)

By Corollary 2.6, \( J_l(\text{grad} B_\theta(x)) \) is contained in \( \text{Cay}(\text{grad} B_\theta(x)) \). However, the remark after Corollary 2.5 proves that such linear maps do not exist. Hence we believe there is a gap here. Since this calculation only affects (8), Theorem 1.1 still works in the Cayley hyperbolic space following [3] provided we can prove (8).

### 4.2 Repairing the gap: reducing a weaker version of (8) into a linear algebra problem

Following the notations from Sect. 4.1, we will first discuss a weaker version of (8) when \( M \) is locally isometric to the Cayley hyperbolic space. Namely the following

\[
\left( \frac{s}{\sqrt{n}} \right)^n \frac{\det \left( \int_{\partial_{\infty} \tilde{M}} dB_\theta^2 | \tilde{F}_\theta(x) \right) \left| d\sigma_\gamma^x(\theta) \right|^2}{\det \left( \int_{\partial_{\infty} \tilde{M}} \text{Hess} B_\theta | \tilde{F}_\theta(x) \right) \left| d\sigma_\gamma^x(\theta) \right|^2} \leq \left( \frac{s}{h(g_0)} \right)^n.
\] (11)

We will leave the proof of (8) in Sect. 4.4. Notice that \( n = 16 \), \( h(g_0) = 16 + 8 - 2 = 22 \) in the Cayley hyperbolic setting, (11) follows from the following proposition.

**Proposition 4.1** Let \( M \) be a locally symmetric space with finite volume whose universal cover \( \tilde{M} \) is the Cayley hyperbolic space, Assume that \( x \in \tilde{M} \) is an arbitrary point. Denote by \( \mu \) a probability measure on \( \partial_{\infty} \tilde{M} \) in the Lebesgue measure class. Then we have

\[
\frac{\det \left( \int_{\partial_{\infty} \tilde{M}} (dB_\theta)^2(x) \right) \left| d\mu(\theta) \right|^2}{\det \left( \int_{\partial_{\infty} \tilde{M}} \text{Hess} B_\theta(x) \right) \left| d\mu(\theta) \right|^2} \leq \left( \frac{\sqrt{16}}{16 + 8 - 2} \right)^{16}
\] (12)

whenever the LHS of (4.1) is well-defined.

We want to reduce Proposition 4.1 to a pure linear algebra problem. To this end, we first introduce some reductions.

**Proof** (Reductions of Proposition 4.1) For any \( w \in T_x \tilde{M} \setminus \{0\} \), we denote by \( \pi_{\text{Cay}(w)} : T_x \tilde{M} \to T_x \tilde{M} \) the orthogonal projection onto \( \text{Cay}(w) \). For simplicity we write \( \pi_{x,\theta} = \pi_{\text{Cay}(\text{grad} B_\theta(x))} \). In the Cayley hyperbolic space, by (9) we can explicitly compute the Hessian of Busemann functions as

\[
\text{Hess} B_\theta(x)(\cdot, \cdot) = g_0(\cdot, \cdot) - 2dB_\theta(\cdot)dB_\theta(\cdot) + g_0(\cdot, \pi_{x,\theta}(\cdot)).
\]

We will be using this formula in place of (10) in the Cayley hyperbolic setting because (10) does not hold in the Cayley hyperbolic case (see Sect. 4.1). We also note that a corresponding version of this formula is equivalent to (10) in the complex (or quaternionic) hyperbolic setting.

To simplify notations, we let

\[
H = \int_{\partial_{\infty} \tilde{M}} (dB_\theta)^2(x) \left| d\mu(\theta) \right|^2.
\] (13)
and
\[ \hat{H} = \int_{\bar{\partial}_\infty M} \langle \cdot , \pi_x, \theta (\cdot) \rangle d\mu(\theta), \]  
(14)

where \( \langle \cdot , \cdot \rangle \) denotes the inner product induced by \( g_0 \). If we identify \( H \) and \( \hat{H} \) with their induced self-adjoint linear operators, \( H \) is the \( \mu \)-weighted average of orthogonal projections onto 1-dimensional subspaces spanned by unit vectors and \( \hat{H} \) is the \( \mu \)-weighted average of orthogonal projections onto Cayley lines of unit vectors. Then the inequality (12) is equivalent to
\[ \frac{(\det H)^{1/2}}{\det (\text{Id} - H + (\hat{H} - H))} \leq \left( \frac{\sqrt{n}}{n + d - 2} \right)^n = \left( \frac{\sqrt{16}}{16 + 8 - 2} \right)^{16} \]  
(15)

where \( n = \dim M \) and \( d = \dim \mathbb{R} \circ \mathbb{R} = 8 \). Since we assumed that the LHS of (4.1) is well-defined, \( \text{Id} - H + (\hat{H} - H) \) is invertible. Notice that (15) holds trivially if \( \det(H) = 0 \), from now on we assume that \( H \) is invertible and hence is positive definite.

A remark on (15) in the complex (or quaternionic) hyperbolic setting: One can similarly define \( H \) and \( \hat{H} \) in the complex (or quaternionic) hyperbolic setting as the \( \mu \)-weighted averages of orthogonal projections onto unit vectors and their complex (or quaternionic) lines respectively. In those cases which are already well-understood in [3, Appendice B],
\[ \hat{H} = \sum_{l=0}^{d-1} J_l H J_l^{-1}, \]
where \( J_0 = \text{Id} \) and \( J_1, \ldots, J_{d-1}, d = 2, 4 \) come from the complex (or quaternionic) structure(s). Here we want to focus on the following two properties of \( \hat{H} \) and \( H \) in the complex (or quaternionic) hyperbolic cases:

(P1). Let \( v, w \) be unit vectors. If \( v, w \) lie in the same complex (or quaternionic) line, then \( w^T \hat{H} w = v^T \hat{H} v \);

(P2). \( \hat{H} - H \) is the sum of \( d - 1 \) matrices which are all conjugate to \( H \) via orthogonal matrices. Moreover, these orthogonal matrices are given by the complex (or quaternionic) structure on \( M \). In particular, these orthogonal matrices map every complex (or quaternionic) line to itself.

Going back to the Cayley hyperbolic case. Unfortunately, due to the lack of \( J_l \) maps compatible with curvature data (see the remark after Corollary 2.5), the second property described above does not hold for the Cayley hyperbolic setting. Nonetheless, (P1) and a weaker version of (P2) still hold for the Cayley hyperbolic space. To be specific, we have the following lemma.

### Lemma 4.2

Let \( v \) be any unit vector in \( T_x^1 M \). We denote by \( \hat{H}|_{\text{Cay}(v)} \) the symmetric bilinear form on \( \text{Cay}(v) \) induced by \( \hat{H} \). Then we have the following properties of \( \hat{H} \) and \( H \).

(P1'). \( \hat{H}|_{\text{Cay}(v)} = \lambda_v \cdot g_0|_{\text{Cay}(v)} \), where \( \lambda_v \) is a real number related to \( v \). In particular, if \( v \) is an eigenvector of \( \hat{H} \), all non-zero vectors in \( \text{Cay}(v) \) are eigenvectors of \( \hat{H} \) with the same eigenvalue. This is the Cayley hyperbolic version of (P1).

(P2'). There exists orthogonal linear maps \( J_{t,v} \in \text{O(Cay}(v)) \), \( t = 0, 1, \ldots, 7 \) such that \( J_{0,v} = \text{Id} \) and
\[ \hat{H}|_{\text{Cay}(v)} = \sum_{t=0}^7 H|_{\text{Cay}(v)} \circ (J_{t,v}(\cdot), J_{t,v}(\cdot)). \]

This is weaker than an analogue of (P2) in the Cayley hyperbolic setting.
**Proof of Lemma 4.2** (P1’). Since \( \hat{H} \) is a weighted average of 2-forms \( \langle \cdot, \pi_{Cay(w)}(\cdot) \rangle \), where \( w \in T^1_x M \), it suffices to prove the conclusion for \( \langle \cdot, \pi_{Cay(w)}(\cdot) \rangle \). It is natural to consider the octonionic coordinates \( \Psi_x : \mathbb{O}^2 \to T_x M \) introduced at the end of Sect. 2. Since \( Stab_{\text{Isom}(\mathbb{H})}(x) \) acts transitively on \( T^1_x M_0 \), we can assume without loss of generality that \( v = \Psi_x(1, 0) \) and \( w = \Psi_x(b \cos \theta, ba \sin \theta) \), where \( \theta \in \mathbb{R} \) and \( a, b \) are unit octonions.

Then for any unit octonion \( c \) and any \( \nu_x = \Psi_x(c, 0) \), we have
\[
\pi_{Cay(w)}(\nu_x) = \Psi_x(c \cos^2 \theta, ca \sin \theta \cos \theta).
\]
Hence
\[
\langle \nu_x, \pi_{Cay(w)}(\nu_x) \rangle = \langle (c, 0), (c \cos^2 \theta, ca \sin \theta \cos \theta) \rangle = \cos^2 \theta,
\]

The first assertion follows since the above inner product is independent of \( c \).

(P2’). Let \( w \in T^1_x M_0 \). Similar to the proof of the first assertion, we can assume without loss of generality that \( v = \Psi_x(1, 0) \) and \( w = \Psi_x(b \cos \theta, ba \sin \theta) \), where \( \theta \in \mathbb{R} \) and \( a, b \) are unit octonions. Also it suffices to prove the conclusion for \( \langle \cdot, \pi_{Cay(w)}(\cdot) \rangle \). Let \( e_0 = 1, e_1, \ldots, e_7 \) be the standard orthonormal basis for \( \mathbb{O} \). Then
\[
\{w_\ell := \Psi_x(\overline{e}_\ell b \cos \theta, (\overline{e}_\ell b) a \sin \theta) : 0 \leq \ell \leq 7 \}
\]
is an orthonormal basis for \( Cay(w) \). Let \( \pi_H : T_x M \to T_x M \) be the orthogonal projection onto \( \mathbb{R} u \) for any non-zero vector \( u \in T_x M \). Then
\[
\langle v', \pi_{Cay(w)}(v') \rangle = \sum_{t=0}^{7} \langle v', \pi_{w_t} v' \rangle, \forall v' \in Cay(v).
\]

Notice that any \( v' \) has the form \( \Psi_x(c, 0) \) for some unit octonion \( c \) and
\[
\langle v', \pi_{w_t} v' \rangle = \langle \Psi_x(c, 0), \Psi_x(\overline{e}_\ell b \cos \theta, (\overline{e}_\ell b) a \sin \theta) \rangle^2
\]
\[
= \langle c, \overline{e}_\ell b \cos \theta \rangle^2
\]
\[
= \langle e_\ell c, b \cos \theta \rangle^2
\]
\[
= \langle \Psi_x(e_\ell c, 0), \nu_w \pi_{\Psi_x}^{-1}(v') \rangle^2 = \langle \Psi_x(e_\ell \Psi_x^{-1}(v')), \pi_w \Psi_x(e_\ell \Psi_x^{-1}(v')) \rangle, \quad 0 \leq \ell \leq 7.
\]

Choose \( J_{t,v}(\cdot) = \Psi_x(e_\ell \Psi_x^{-1}(\cdot)) \) and the second assertion follows.

\[\square\]

Back to the proof of Proposition 4.1, (P1) suggests that \( \hat{H} \) has at most 2 eigenvalues \( \lambda, \mu \geq 0 \) with \( \lambda + \mu = 1 \). When \( \lambda \neq \mu \), the two eigenspaces are a pair of Cayley lines perpendicular to each other. This is because the largest and the smallest eigenvalue of \( \hat{H} \) can be defined as
\[
\lambda_{\text{max}} = \max_{v \in T^1_x M} \hat{H}(v, v) \text{ and } \lambda_{\text{min}} = \min_{v \in T^1_x M} \hat{H}(v, v).
\]
A unit vector \( v \in T^1_x M \) is an eigenvector of \( \hat{H} \) corresponding to the largest (smallest resp.) eigenvalue if and only if \( \hat{H}(v, v) \) equals to the largest (smallest resp.) eigenvalue. Hence (P1) implies that the eigenspace of \( \hat{H} \) corresponding to the largest (smallest resp.) eigenvalue must contain all Cayley lines which have non-trivial intersections with it. Therefore under a suitable orthonormal basis \( H \) and \( \hat{H} \) have matrices of the form
\[
H = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix} \quad \text{and} \quad \hat{H} = \begin{pmatrix} \lambda \cdot \text{Id} & 0 \\ 0 & \mu \cdot \text{Id} \end{pmatrix}, \quad (16)
\]
where $A, B$ are $8 \times 8$ diagonal matrices with trace equal to $\lambda, \mu$ respectively. The trace data of $A$ and $B$ follows from (P2). By the definitions of $H$ and $\hat{H}$ in (13) and (14), we have $\hat{H} \geq H$. Therefore

$$\hat{H} - H = \begin{pmatrix} \lambda \cdot \text{Id} - A & -C \\ -C^* & \mu \cdot \text{Id} - B \end{pmatrix}$$

(17)

is positive semi-definite. Recall that $H$ is assumed to be positive definite, (see the comments after (15).) in particular, $A$ and $B$ are positive definite. Hence any eigenvalue of $A$ (or $B$) is strictly smaller than $\lambda = \text{tr}A$ (or $\mu = \text{tr}(B)$). This implies that $\lambda \cdot \text{Id} - A$ and $\mu \cdot \text{Id} - B$ are positive definite.

Before we outline the main idea of the proof for (15) in the Cayley hyperbolic setting, we first recall Besson–Courtois–Gallot’s method in proving the complex (or quaternionic) hyperbolic version of (15). (See [3, B.3 Lemme].)

**Besson–Courtois–Gallot’s proof for (15) in the complex (or quaternionic) hyperbolic setting:** We first apply the log-concavity of the determinant function for positive semi-definite matrices (see [3, B.2 Lemme]) and obtain

$$\det(\text{Id} - H + (\hat{H} - H)) \geq (n + d - 2)^n \det\left(\frac{\text{Id} - H}{n - 1}\right)^{\frac{n-1}{n+d-2}} \det\left(\frac{\hat{H} - H}{d - 1}\right)^{\frac{d-1}{n+d-2}},$$

(18)

where $n$ is the dimension of a complex (or quaternionic) hyperbolic space and $d = 2$ (or 4). It follows from [3, B.4 Lemme] that

$$\det((\text{Id} - H)/(n - 1)) \geq n^{-\frac{n(n+d-2)}{2(n-1)}} \det(H)^{\frac{n-d}{2(n-1)}}$$

(19)

for any positive semi-definite matrix $H$ with $\text{tr}H = 1$. (This follows from using $\alpha = 2(n-1)/(n-d)$ in [3, B.4 Lemme].) By Property (P2) for the complex (or quaternionic) hyperbolic setting and again the log-concavity of determinant function for positive semi-definite matrices, we have

$$\det((\hat{H} - H)/(d - 1)) \geq \det(H).$$

(20)

Therefore the complex (or quaternionic) hyperbolic version of (15) follows from (18), (19) and (20).

**Going back to the Cayley hyperbolic setting.** The argument above does not work directly for the Cayley hyperbolic case due to the lack of Property (P2). Namely, it is not obvious that $\det((\hat{H} - H)/(d - 1)) = \det((\hat{H} - H)/7) \geq \det H$ because we cannot write $\hat{H} - H$ as the sum of 7 conjugates of $H$. (See the remark after Corollary 2.5 or Sect. 4.1.) However, we can make some adjustments to the above argument so that it still works (with a lot more technical details) in the Cayley hyperbolic case. Set

$$U = \begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}$$

and $U_\alpha = \alpha U$.

for some $0 \leq \alpha \leq 1$ to be determined, where $C$ is the same as in (16). If we assume that $0 < \lambda_1, \ldots, \lambda_8 < \lambda$ and $0 < \mu_1, \ldots, \mu_8 < \mu$ are eigenvalues of $A$ and $B$ respectively (due to the comments after (17)), then by the matrix expressions (16) and (17) for $H$ and $\hat{H} - H$, we can apply the arithmetic mean-geometric mean inequality and obtain

$$\det\left[\frac{\hat{H} - (H - U)}{7}\right] = \det\left(\frac{1}{7} \begin{pmatrix} \lambda \cdot \text{Id} - A & 0 \\ 0 & \mu \cdot \text{Id} - B \end{pmatrix}\right)$$

Springer
\[
\prod_{i=1}^{8} \left( \frac{\lambda - \lambda_i}{7} \right) \left( \frac{\mu - \mu_i}{7} \right) \\
= \prod_{i=1}^{8} \left( \frac{\sum_{1 \leq j \leq 8} \lambda_j}{7} \right) \left( \frac{\sum_{1 \leq j \leq 8} \mu_j}{7} \right) \\
\geq \prod_{i=1}^{8} \left( \prod_{1 \leq j \leq 8, j \neq i} \frac{\lambda_j}{7} \right) \left( \prod_{1 \leq j \leq 8, j \neq i} \frac{\mu_j}{7} \right) \\
= \prod_{i=1}^{8} \lambda_i \mu_i = \det(A) \det(B) = \det \left( \begin{array}{cc} A & 0 \\ 0 & B \end{array} \right). \tag{21}
\]

In order to compare the RHS of (21) with \( \det(H) \), we first introduce a linear algebra fact which will be used extensively in the rest of this paper.

**Fact 4.3** Let \( Q_1, Q_2 \) be \( n \times n \) symmetric matrices and \( P \) an \( n \times n \) matrix. Assume that \( Q_1 \) is invertible. Then

\[
\begin{pmatrix} Q_1 & P \\ P^* & Q_2 \end{pmatrix} = \begin{pmatrix} \text{Id} & \begin{pmatrix} 0 \\ P^* \end{pmatrix} Q_1^{-1} & \end{pmatrix} \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 - P^* Q_1^{-1} P \end{pmatrix} \begin{pmatrix} \text{Id} & \begin{pmatrix} 0 \\ Q_1^{-1} P \end{pmatrix} \end{pmatrix}.
\]

Hence \( \begin{pmatrix} Q_1 & P \\ P^* & Q_2 \end{pmatrix} \) is positive (semi-)definite if and only if both \( Q_1 \) and \( Q_2 - P^* Q_1^{-1} P \) are positive (semi-)definite. Also we have the determinant equality

\[
\det \begin{pmatrix} Q_1 & P \\ P^* & Q_2 \end{pmatrix} = \det(Q_1) \det(Q_2 - P^* Q_1^{-1} P).
\]

Back to the comparison between the RHS of (21) and \( \det(H) \), since \( H, A \) and \( B \) are positive definite, by Fact 4.3, \( B - C^* A^{-1} C \) is positive definite. Therefore

\[
\det \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det(A) \det(B) \geq \det(A) \det(B - C^* A^{-1} C) = \det \left( \begin{array}{cc} A & C \\ C^* & B \end{array} \right) = \det(H). \tag{22}
\]

In other words, the main reason (20) is not clear in the Cayley hyperbolic setting comes from the “troublesome” \( U \). Moreover, since both \( \lambda \cdot \text{Id} - A \) and \( \mu \cdot \text{Id} - B \) are positive definite (see the comments after (17)), and \( \hat{H} - H \) is positive semi-definite, Fact 4.3 and (17) imply that \( \mu \cdot \text{Id} - B - C^* (\lambda \cdot \text{Id} - A)^{-1} C \) is positive semi-definite. Therefore by Fact 4.3 and (17) again, \( \hat{H} - H + U_\alpha \) is positive definite and

\[
\det(\hat{H} - H + U_\alpha) = \det(\lambda \cdot \text{Id} - A) \det[\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^* (\lambda \cdot \text{Id} - A)^{-1} C] \tag{23}
\]

is an increasing function for \( 0 \leq \alpha \leq 1 \).

From (21), (22) and (23) one can expect that for some \( 0 \leq \alpha \leq 1 \) which is likely close to 1, we have

\[
\det[\hat{H} - H + U_\alpha] \geq \det(H). \tag{24}
\]
Recall that it is not clear whether (20) holds in the Cayley hyperbolic setting. We will use (24) instead of (20) for our proof of (15).

We will also use the log-concavity of the determinant function for positive semi-definite matrices similar to (18). Namely,

\[
\det(\Id + \hat{H} - 2H) \geq 22^{16} \det\left(\frac{\Id - H - U_\alpha}{15}\right)^{\frac{15}{7}} \det\left(\frac{\hat{H} - H + U_\alpha}{7}\right)^{\frac{7}{22}}
\]  

(25)

when \(\Id - H - U_\alpha\) is positive semi-definite. (We have already proved that \(\hat{H} - H + U_\alpha\) is positive definite right before (23).) This inequality (25) will play a similar role in the proof of Proposition 4.1 as (18) does in Besson–Courtois–Gallot’s proof for (15) in the complex (or quaternionic) hyperbolic setting. (See (19), (20) and (18).) The last remaining building-block which corresponds to (19) is to prove

\[
\det(\Id - H - U_\alpha) \geq \frac{15^{16}}{16^{16}(1 - 4/15)} \det(H)^{4/15}
\]  

(26)

for the same \(\alpha\) in (24). Given (19) which is the same as (26) in the Cayley hyperbolic setting when \(\alpha = 0\), one might guess that this \(\alpha\) should be close to 0. We summarize all the reductions discussed above as the following idea of the proof for (15) and hence for Proposition 4.1.

\[\square\]

Idea of the proof for Proposition 4.1 Recall that proving (15) is equivalent to proving Proposition 4.1. Notice that (25) holds for any \(\alpha \in [0, 1]\). If there exists some \(\alpha \in [0, 1]\) such that (26) and (24) hold, then (15) is a direct corollary of (25), (26) and (24). Therefore it remains for us to solve the following linear algebra problem: Finding an \(\alpha \in [0, 1]\) such that both (26) and (24) hold.

Finding such an \(\alpha\) is also the main difficulty in this proof since naively we expect that (26) holds when \(\alpha\) is “close” to 0 and (24) holds when \(\alpha\) is “close” to 1. Therefore we need to verify the existence of an \(\alpha\) which is “close” to both 0 and 1 in the above sense.

Lemma 4.4 in Sect. 4.3 confirms the existence of such an \(\alpha\) and hence completes the proof of Proposition 4.1.

4.3 Repairing the gap: a linear algebra lemma

Recall that from the end of Sect. 4.2, it remains for us to find an \(\alpha \in [0, 1]\) such that (26) and (24) hold. This section gives a positive answer to the existence of such an \(\alpha\). This can be summarized as the following linear algebra lemma.

Lemma 4.4 Following the notations from Sect. 4.2, there exists some \(0 \leq \alpha \leq 1\) such that

1. \(\det(\hat{H} - H + U_\alpha) \geq 7^{16} \det(H)\) (see also (24));
2. \(\det(\Id - H - U_\alpha) \geq \frac{15^{16}}{16^{16}(1 - 4/15)} \det(H)^{4/15}\) (see also (26)).

Proof Let \(\lambda_1 \geq \cdots \geq \lambda_8 \geq 0\) and \(\mu_1 \geq \cdots \geq \mu_8\) be eigenvalues (and therefore the diagonal elements) of \(A\) and \(B\) respectively. By the comments after (17), \(0 < \lambda_j < \lambda\) and \(0 < \mu_j < \mu\) for any \(1 \leq j \leq 8\).

Without loss of generality we assume that

\[
A = \begin{pmatrix}
\lambda_1 & & \\
& \ddots & \\
& & \lambda_8
\end{pmatrix}
\quad \text{and} \quad
B = \begin{pmatrix}
\mu_1 & & \\
& \ddots & \\
& & \mu_8
\end{pmatrix}.
\]

\[\square\]
Write $C = (c_{ij})_{1 \leq i,j \leq 8}$. Denote by $C_k = (c_{ki}c_{kj})_{1 \leq i,j \leq 8}$ with $1 \leq k \leq 8$. Then for any diagonal matrix $D = \text{diag}(d_1, \ldots, d_8)$, we have

$$
C^*DC = \left( \sum_{k=1}^{8} c_{ki}d_kc_{kj} \right)_{1 \leq i,j \leq 8} = \sum_{k=1}^{8} d_kC_k.
$$

(27)

When $C_k \neq 0$, we can think of $C_k/\text{tr}C_k$ as the matrix for the orthogonal projection onto the $k$-th row vector of $C$, where $(\text{tr}C_k)^{1/2}$ equal to the norm of this vector. It is clear that $C_k$ has rank at most 1. Then (27) states the linear algebra fact that any matrix of the form $C^*DC$ with $D$ diagonal is a linear combination of orthogonal projections onto the row vectors of $C$. Applying Fact 4.3 and (27) to $H$ and $\widehat{H} - H$ yields

$$
B - C^*A^{-1}C = B - \left( \sum_{k=1}^{8} C_k/\lambda_k \right) \geq 0
$$

(28)

and

$$
\mu \cdot \text{Id} - B - C^*(\lambda \cdot \text{Id} - A)^{-1}C = \mu \cdot \text{Id} - B - \left( \sum_{k=1}^{8} C_k/(\lambda - \lambda_k) \right) \geq 0.
$$

(29)

The rest of the proof is divided into two cases. We will first discuss the “good case”.

Motivations for the “good case”: Find conditions on $A$, $B$ when $\alpha = 0$ works.

Recall that in Besson–Courtois–Gallot’s proof for (15) in the complex (or quaternionic) hyperbolic setting, (see (19), (20) and (18)), we can choose $\alpha = 0$ and the corresponding version of Lemma 4.4 holds. We want to first discuss to what extent $\alpha = 0$ works in the Cayley hyperbolic setting. Notice that (19) (see [3, B.4 Lemme]) is true for any positive semi-definite matrix $H$ with trace 1, the second inequality in Lemma 4.4 always holds when $\alpha = 0$. Therefore it suffices to find the conditions when the first inequality holds and $\alpha = 0$ simultaneously. By Fact 4.3 we have

$$
det(H) = det(A) det(B - C^*A^{-1}C)
$$

(30)

and

$$
det(\widehat{H} - H) = det(\lambda \cdot \text{Id} - A) det(\mu \cdot \text{Id} - B - C^*(\lambda \cdot \text{Id} - A)^{-1}C).
$$

(31)

Since $A$ and $B$ are both positive definite and have trace $\lambda$ and $\mu$ respectively, one can apply the arithmetic mean-geometric mean inequality (in a similar way compare to (22)) and obtain

$$
det(\lambda \cdot \text{Id} - A) \geq 7^8 det(A) \text{ and } det(\mu \cdot \text{Id} - B) \geq 7^8 det(B).
$$

(32)

Therefore if we can prove that

$$
\frac{det(\widehat{H} - H)}{det(\lambda \cdot \text{Id} - A) det(\mu \cdot \text{Id} - B)} \geq \frac{det(H)}{det(A) det(B)}.
$$

then the first inequality of Lemma 4.4 holds when $\alpha = 0$. This is equivalent to proving

$$
\frac{det(\mu \cdot \text{Id} - B - C^*(\lambda \cdot \text{Id} - A)^{-1}C)}{det(\mu \cdot \text{Id} - B)} \geq \frac{det(B - C^*A^{-1}C)}{det(B)}.
$$

(33)

One can naively expect (33) to be true if $\mu \cdot \text{Id} - B$ is more positive than $B$ and $C^*(\lambda \cdot \text{Id} - A)^{-1}C$ is less positive than $C^*A^{-1}C$. This comes from the observation that if all the matrices in (33) are of size $1 \times 1$, then it reduces to the elementary fact that $(a - d)/a \geq (b - c)/c$ for any positive numbers $a \geq b \geq c \geq d > 0$. Therefore we discuss the following “good” case.
Case 1: If $2\lambda_1 \leq \lambda$ and $2\mu_1 \leq \mu$, then we have $\mu \cdot \text{Id} - B \geq B$ and

$$C^* A^{-1} C = \left( \sum_{k=1}^{8} C_k / \lambda_k \right) \geq \left( \sum_{k=1}^{8} C_k / (\lambda - \lambda_k) \right) = C^* (\lambda \cdot \text{Id} - A)^{-1} C,$$

where two symmetric matrices $Q_1 \geq Q_2$ if and only if $Q_1 - Q_2$ is positive semi-definite. Therefore

$$0 \leq \mu \cdot \text{Id} - B - \left( \sum_{k=1}^{8} C_k / \lambda_k \right) \leq \mu \cdot \text{Id} - B - \left( \sum_{k=1}^{8} C_k / (\lambda - \lambda_k) \right).$$

(34)

Hence

$$\frac{\text{det}(B - C^* A^{-1} C)}{\text{det}(\mu \cdot \text{Id} - B - C^* (\lambda \cdot \text{Id} - A)^{-1} C)} \leq \frac{\text{det}(B - C^* A^{-1} C)}{\text{det}(\mu \cdot \text{Id} - B - C^* A^{-1} C)}$$

(35)

Before we finish the estimate, we introduce a linear algebra lemma.

**Lemma 4.5** Let $H_1 \geq H_2 \geq W \geq 0$ be $n \times n$ positive semi-definite matrices. Then

$$\text{det}(H_2 - W) \text{det}(H_1) \leq \text{det}(H_1 - W) \text{det}(H_2).$$

In particular, if $n = 1$, then we have the naive inequality $(H_2 - W)H_1 \leq (H_1 - W)H_2$ for any non-negative real numbers $H_1 \geq H_2 \geq W \geq 0$.

**Proof of Lemma 4.5** If $H_2 - W$ is singular then the result is trivial. Without loss of generality we assume that $H_2 - W$ is invertible. Since $H_1 \geq H_2 \geq H_2 - W \geq 0$, $H_1$ and $H_2$ are also invertible and we can assume WLOG that $H_1 = \text{Id}$. We first prove the special case when $W$ is a rank one matrix. After conjugating $H_1, H_2, W$ simultaneously by a suitable orthogonal matrix, we can assume WLOG that $W = \text{diag}(c, 0, \ldots, 0)$. Write $H_2 = (h_{ij})_{1 \leq i, j \leq n}$. Define

$$\hat{H}_2 = \left( \begin{array}{cccc} h_{11} & 0 & \cdots & 0 \\ 0 & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & h_{n2} & \cdots & h_{nn} \end{array} \right).$$

Then $\text{det} H_2 \leq \text{det} \hat{H}_2$ and

$$\text{det}(\hat{H}_2) - \text{det}(\hat{H}_2 - W) = \text{det}(H_2) - \text{det}(H_2 - W)$$

Hence

$$\frac{\text{det}(H_2 - W)}{\text{det}(H_2)} \leq \frac{\text{det}(\hat{H}_2 - W)}{\text{det}(\hat{H}_2)} = \frac{\text{det}(H_1 - W)}{\text{det}(H_1)}.$$

This finishes the proof of Lemma 4.5 when $W$ is a rank one matrix. For general $W$, we can write $W = W_1 + \cdots + W_k$ for some $0 \leq k \leq n$, where $W_1, \ldots, W_k$ are positive semi-definite rank one matrices. (This is because we can assume WLOG that $W$ is diagonal after conjugating a suitable orthogonal matrix.) Let

$$H_1^{(0)} := H_1 \text{ and } H_1^{(j)} := H_1 - \sum_{i=1}^{j} W_j, \quad \forall j = 1, 2 \text{ and } 1 \leq j \leq k.$$
It follows directly from the above definitions that $H_1^{(j)} - H_2^{(j)} = H_1 - H_2 \geq 0$ for any $0 \leq j \leq k$ and

$$H_1^{(j+1)} = H_1^{(j)} - W_{j+1} \geq H_2^{(j)} - W_{j+1} = H_2^{(j+1)} = H_2 - W + \sum_{j+2 \leq i \leq k} W_i \geq H_2 - W \geq 0, \forall 0 \leq j \leq k - 1.$$ 

Hence $H_1^{(j)} \geq H_2^{(j)} \geq W_{j+1}$ for any $0 \leq j \leq k - 1$. Since we also assumed that $H_2 - W$ is invertible, $H_1^{(j)} - W_{j+1}$ and $H_2^{(j)} - W_{j+1}$ are invertible. By the above the special case of Lemma 4.5 applied to $H_1^{(j)} \geq H_2^{(j)} \geq W_{j+1}$, we have

$$\frac{\det(H_2 - W)}{\det(H_2)} = \frac{\det(H_2^{(k)})}{\det(H_2^{(0)})} = \prod_{j=0}^{k-1} \frac{\det(H_2^{(j+1)})}{\det(H_2^{(j)})} = \prod_{j=0}^{k-1} \frac{\det(H_2^{(j)} - W_{j+1})}{\det(H_2^{(j)})} \leq \prod_{j=0}^{k-1} \frac{\det(H_1^{(j)} - W_{j+1})}{\det(H_1^{(j)})} = \prod_{j=0}^{k-1} \frac{\det(H_1^{(j+1)})}{\det(H_1^{(j)})} = \frac{\det(H_1 - W)}{\det(H_1)}.$$

Choose $H_1 = \mu \cdot \text{Id} - B$, $H_2 = B$ and $W = C^* A^{-1} C$ in Lemma 4.5, following (32) and (35) we have

$$\frac{\det(B - C^* A^{-1} C)}{\det(\mu \cdot \text{Id} - B - C^* (\lambda \cdot \text{Id} - A)^{-1} C)} \leq \frac{\det(B)}{\det(\mu \cdot \text{Id} - B)} \leq 7^{-8}.$$ 

Hence by (32),

$$\frac{\det(H)}{\det(H - H)} \leq \frac{\det(A) \det(B)}{\det(\lambda \cdot \text{Id} - A) \det(\mu \cdot \text{Id} - B)} \leq 7^{-16}. \quad (36)$$

Choose $\alpha = 0$ and the proof of Lemma 4.4 in this case follows from [3, B.3–B.4 Lemmes].

It remains for us to prove Lemma 4.4 in the following “bad” case.

Case 2 : If $2\lambda_1 > \lambda$ or $2\mu_1 > \mu$, without loss of generality we assume that $2\mu_1 > \mu$. This case is trickier since $\mu \cdot \text{Id} - B$ is no longer more positive than $B$ and $C^* (\lambda \cdot \text{Id} - A)^{-1} C$ is no longer less positive than $C^* A^{-1} C$. Therefore the arguments in the previous case which are slight adjustments to Besson–Courtois–Gallot’s approach (see also [3, B.3–B.4 Lemmes]) do not work in this case. Hence choosing $\alpha = 0$ may not work for (25) and Lemma 4.4. ($\alpha = 0$ could still work in this case but it is not obvious how to verify it.) Another reason why this case is “bad” (or actually “good” depending on one’s perspective) comes from the following observation: Equality for the first inequality in Lemma 4.4 holds when $\lambda_1 = \cdots = \lambda_8$, $\mu_1 = \cdots = \mu_8$ and $C = 0$ for every $\alpha$. The assumptions on eigenvalues $\lambda_j$’s and $\mu_j$’s for $A$ and
in this case are very far from the above mentioned equality conditions. Therefore one can expect that too much difference between eigenvalues of $A$ and $B$ will be the main reason why the inequalities in Lemma 4.4 hold in this case. In addition to that, one can expect that equalities will not be achieved here.

Since we need to find an $\alpha$ satisfying 2 inequalities in Lemma 4.4, we split the rest of the proof into 2 steps.

**Step 1:** Find the conditions for $\alpha$ such that the first inequality in Lemma 4.4 holds in Case 2.

By Fact 4.3, we have
\[
\det(\tilde{H} - H + U_\alpha) = \det(\lambda \cdot \text{Id} - A) \det(\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1} C).
\]
Recall that by (32) we have $\det(\lambda \cdot \text{Id} - A) \geq 7^8 \det(A)$. In order to satisfy the first requirement, it suffices to find $\alpha$ such that
\[
\det(\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1} C) \geq \beta^8 \det(B - C^* A^{-1} C). \tag{37}
\]

**Main idea to find a lower bound for the left hand side of (37):** Usually there is no obvious way to give a non-trivial lower bound on the determinant of a positive semi-definite matrix (e.g. the left hand side of (37)). However we notice that the matrix $\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1} C$ has lower bounds on its trace and its smallest eigenvalue. If we view $\mu$ as the main reason why the inequalities in Lemma 4.4 hold in this case. In addition to that, one can expect that equalities will not be achieved here.

Details for finding a lower bound for the left hand side of (37): Since $\lambda - \lambda_k \geq \lambda_k$ for any $2 \leq k \leq 8$, (28) implies that
\[
\text{tr} \left( \sum_{k=2}^{8} C_k / (\lambda - \lambda_k) \right) \leq \text{tr} \left( \sum_{k=2}^{8} C_k / \lambda_k \right) \leq \text{tr}(C^* A^{-1} C) \leq \text{tr} B = \mu. \tag{38}
\]
Hence we set
\[
\text{tr} \left( \sum_{k=2}^{8} C_k / (\lambda - \lambda_k) \right) = \beta \mu \tag{39}
\]
for some $0 \leq \beta \leq 1$. On the other hand, since $\mu \cdot \text{Id} - B$ has norm bounded above by $\mu$ and $C_1$ has rank 1, by (29) we have
\[
\text{tr}(C_1 / (\lambda - \lambda_1)) = \|C_1 / (\lambda - \lambda_1)\| \leq \|C^* (\lambda \cdot \text{Id} - A)^{-1} C\| \leq \|\mu \cdot \text{Id} - B\| \leq \mu.
\]
Therefore the matrix $\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^* (\lambda \cdot \text{Id} - A)^{-1} C$ (introduced in (29)) has trace at least $[7 - (1 + \beta)(1 - \alpha)^2] \mu$. Notice that (29) implies
\[
\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^* (\lambda \cdot \text{Id} - A)^{-1} C \geq \mu \cdot \text{Id} - B - (1 - \alpha)^2 (\mu \cdot \text{Id} - B)
\]
\[
= [1 - (1 - \alpha)^2] (\mu \cdot \text{Id} - B).
\]
Hence the smallest eigenvalue of $\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^* (\lambda \cdot \text{Id} - A)^{-1} C$ is greater than or equal to $[1 - (1 - \alpha)^2] (\mu - \mu_1)$, where $\mu - \mu_1$ is the smallest eigenvalue of $\mu \cdot \text{Id} - B$. We summarize these estimates on trace and the smallest eigenvalue for reader’s convenience.

**(E1)** Lower bound for trace: $\text{tr}[\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^* (\lambda \cdot \text{Id} - A)^{-1} C] \geq [7 - (1 + \beta)(1 - \alpha)^2] \mu$, where $\beta \in [0, 1]$ is given in (39);
(E2) Lower bound for the smallest eigenvalue: The smallest eigenvalue of $[\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1}C]$ is greater or equal to $[1 - (1 - \alpha)^2](\mu - \mu_1)$.

Since $\mu \cdot \text{Id} \geq \mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1}C$, the norm of $\text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1}C$ is bounded above by $\mu$. If the smallest eigenvalue of $\text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1}C$ equals to $[1 - (1 - \alpha)^2](\mu - \mu_1)$ and has multiplicity at least 3, then by (E2), we have

$$\text{tr}[[\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1}C]\begin{array}{c}
\leq 5\mu + 3[1 - (1 - \alpha)^2](\mu - \mu_1) \\
< 5\mu + 1.5[1 - (1 - \alpha)^2]\mu \\
= [6.5 - 1.5(1 - \alpha)^2]\mu \leq [7 - 2(1 - \alpha)^2]\mu \leq [7 - (1 + \beta)(1 - \alpha)^2]\mu.
\end{array}$$

This contradicts (E1). Therefore if $[1 - (1 - \alpha)^2](\mu - \mu_1)$ is an eigenvalue of $\mu \cdot \text{Id} - B - (1 - \alpha)^2 C^*(\lambda \cdot \text{Id} - A)^{-1}C$, its multiplicity is at most 2. The following linear algebra lemma explains how to obtain a lower bound on its determinant given the above observation.

Lemma 4.6 Let $Q$ be an $8 \times 8$ symmetric matrix with real entries such that its eigenvalues $q_1, \ldots, q_8$ lie in $[m, M]$ for some $0 < m \leq M$. Assume that there exists some constant $K$ such that $\text{tr}(Q) \geq K > 3m + 5M$, then

$$\text{det}(Q) \geq m^2(K - 2m - 5M)M^5.$$ 

Proof of Lemma 4.6 WLOG we assume $m \leq q_1 \leq \cdots \leq q_8 \leq M$. We define a function $P(x_1, \ldots, x_8) = x_1x_2\ldots x_8$ with domain defined by $x_1 + \cdots + x_8 = \text{tr}(Q)$ and $m \leq x_1 \leq \cdots \leq x_8 \leq M$. Assume that $P$ achieves global minimum at $(y_1, \ldots, y_8)$ in its domain. Then there exists at most one $y_i \notin \{m, M\}$. (This is because if there exists $1 \leq i < j \leq 8$ such that $y_i, y_j \notin \{m, M\}$, WLOG we can assume that $y_i = m$ for $l < i$ and $y_k = M$ for $k > j$. Then we can choose $0 < \epsilon \ll 1$ such that $m < y_i - \epsilon < y_j + \epsilon < M$. It is not hard to verify that $(y_1, \ldots, y_i - \epsilon, \ldots, y_j + \epsilon, \ldots, y_8)$ is still in the domain of $P(x_1, \ldots, x_8)$ and $P(y_1, \ldots, y_i - \epsilon, \ldots, y_j + \epsilon, \ldots, y_8) < P(y_1, \ldots, y_8)$, which contradicts with the assumption on $(y_1, \ldots, y_8)$.) Since $\text{tr}(Q) \geq K > 3m + 5M$, we have $y_4 = \cdots = y_8 = M$.

Case (1): If $y_1 = y_2 = m$, then

$$\text{det}(Q) = P(q_1, \ldots, q_8) = \text{det}(Q), \quad y_1, \ldots, y_8, \quad = m^2(\text{tr}(Q) - 2m - 5M)M^5 \geq m^2(K - 2m - 5M)M^5.$$

Case (2): If $y_2 > m$ and $y_1 = m$, then $m < \text{tr}(Q) - m - 6M \leq M$ and

$$\text{det}(Q) \geq P(y_1, \ldots, y_8) = m\text{tr}(Q) - m - 6M)M^6 \geq m^2(\text{tr}(Q) - 2m - 5M)M^5 \geq m^2(K - 2m - 5M)M^5.$$ 

Case (3): If $y_1 > m$, then $m < \text{tr}(Q) - 7M \leq M$ and

$$\text{det}(Q) \geq P(y_1, \ldots, y_8) = (\text{tr}(Q) - 7M)M^7 \geq m^2(\text{tr}(Q) - 2m - 5M)M^5 \geq m^2(K - 2m - 5M)M^5.$$ 

Here the strict inequality in Case (3) follows from the fact that $\text{tr}(Q) - m - 6M \geq M$ and that

$$\text{tr}(Q) - 7M)^M^2 > m\text{tr}(Q) - m - 6M)M > m^2(\text{tr}(Q) - 2m - 5M).$$ 

\[\square\]
Applying Lemma 4.6 to the case when $Q = \mu \cdot \text{Id} - B - (1 - \alpha)^2C^*(\lambda \cdot \text{Id} - A)^{-1}C$, $m = [1 - (1 - \alpha)^2](\mu - \mu_1)$, $M = \mu$ and $K = [7 - (1 + \beta)(1 - \alpha)^2]\mu$, we have

$$\det(\mu \cdot \text{Id} - B - (1 - \alpha)^2C^*(\lambda \cdot \text{Id} - A)^{-1}C) \geq [(1 - (1 - \alpha)^2)(\mu - \mu_1)]^2 \cdot [2\mu_1[1 - (1 - \alpha)^2] + (1 - \beta)(1 - \alpha)^2\mu] \cdot \mu^5$$

$$= [1 - (1 - \alpha)^2]^2(\mu - \mu_1)^2[2\mu_1[1 - (1 - \alpha)^2] + (1 - \alpha)^2\mu - \beta(1 - \alpha)^2\mu]\mu^5. \ (40)$$

It remains for us to prove that the lower bound in (40) is greater than or equal to $\det(B - C^*A^{-1}C)$. We will do so by finding a suitable upper bound for $\det(B - C^*A^{-1}C)$ (equivalently for the right hand side of (37)). Similar to the arguments on finding a lower bound for the left hand side of (37), we first summarize the idea behind finding this upper bound estimate which is less complicated than the actual arguments.

**Main idea to find an upper bound for the right hand side of (37):** Equivalently we want to find an upper bound for $\det(B - C^*A^{-1}C)$. Since we have assumed $\mu_1 \geq \mu/2$ in this case, the possibilities of $B - C^*A^{-1}C$ having equal eigenvalues are very low when $C$ is close to 0. Therefore directly applying the arithmetic mean-geometric mean inequality to the eigenvalues of $B - C^*A^{-1}C$ may overkill. We want to use the idea of the arithmetic mean-geometric mean inequality in a more delicate way. Namely, if trace of $B - C^*A^{-1}C$ is fixed, $\det(B - C^*A^{-1}C)$ becomes larger when its eigenvalues are closer to each other, even in the case when it is impossible for all eigenvalues to be the same. In particular, we noticed that if $C$ is close to 0, the largest possible value of $\det(B - C^*A^{-1}C)$ as a function on its eigenvalues is achieved when $B$ and $B - C^*A^{-1}C$ have the same eigenvalues except the largest one. The main reason is that $\mu_1$ is way larger than other eigenvalues $\mu_j$ of $B$ in this case. The technical details of the above idea is presented below.

**Details for finding an upper bound for the right hand side of (37):** By (28), (38) and (39) we have $B - C^*A^{-1}C$ is positive semi-definite with trace smaller than or equal to $(1 - \beta)\mu$. Since $B \geq B - C^*A^{-1}C$, we can let $\mu_j - b_j$ be the $j$-th diagonal entry of $B - C^*A^{-1}C$. By Fact 4.3 we know that the determinant of any positive semi-definite matrix is smaller than or equal to the product of its diagonal elements. In particular, $\det(B - C^*A^{-1}C) \leq \prod_{j=1}^{8}(\mu_j - b_j)$.

Define

$$E_\beta(z_1, \ldots, z_8) = \prod_{j=1}^{8} \left(1 - \frac{z_j}{\mu_j}\right), \quad 0 \leq z_j \leq \mu_j, \sum_{j=1}^{8} z_j \geq \beta \mu.$$  

Then we have $0 \leq \mu_j - b_j \leq \mu_j$ and $b_1 + \cdots + b_8 \geq \beta \mu$ by (28), (38) and (39). Hence $(b_1, \ldots, b_8)$ is in the domain of $E_\beta$ and

$$\det(B - C^*A^{-1}C) \leq \det(B)E_\beta(b_1, \ldots, b_8) \leq \det(B) \cdot \sup E_\beta \leq \mu_1 \left(\frac{\mu - \mu_1}{7}\right)^7 \sup E_\beta.$$  

The following lemma provides a more delicate arithmetic mean-geometric mean estimate regarding $E_\beta$.

**Lemma 4.7** Under the above assumptions, $E_\beta$ achieves maximum at a point $(\xi_1, \ldots, \xi_8)$ satisfying the following additional conditions.

1. $\xi_1 + \cdots + \xi_8 = \beta \mu$;
2. $\mu_1 - \xi_1 \geq \cdots \geq \mu_8 - \xi_8$;
3. If $\xi_j \neq 0$, then $\mu_1 - \xi_1 = \mu_j - \xi_j$.  

[Springer]
Remark One way to understand these conditions is to vary $\beta$ from 0 to 1. When $\beta$ is small, $\xi_j = 0$ for $j \geq 2$ and hence $\xi_1 = \beta \mu$. $\xi_2$ will not become positive until $\beta$ become too large that $\mu_1 - \beta \mu \leq \mu_2$. Naively, if we think of $\mu_j$ as the “pre-tax income of the $j$-th person” and $\xi_j$ as “the tax he/she has to pay”, the process of finding $\sup E_\beta$ shows an example of “the one who earns more pays more taxes” with “the least possible inequality in after-tax income”. This explains its similarity to the arithmetic mean-geometric mean inequality since we are maximizing $E_\beta$ by minimizing the differences among $\mu_j - \xi_j$ under the given conditions.

Proof of Lemma 4.7 Since the domain of $E_\beta$ is compact, the global maximum can be achieved at some point $(\xi_1, \ldots, \xi_8)$. We verify all three conditions by contradictions.

1. If $\xi_1 + \cdots + \xi_8 =: (1 + \epsilon) \beta \mu > \beta \mu$ for some $\epsilon > 0$, then $E_\beta((1 + \epsilon) \beta \mu) > E_\beta(\xi_1, \ldots, \xi_8)$. This contradicts our assumption on $(\xi_1, \ldots, \xi_8)$. Therefore the first condition must hold;

2. If $\mu_i - \xi_i < \mu_j - \xi_j$ for some $i < j$, we define

$$\xi'_k = \begin{cases} 
\xi_k, & k \neq i, j; \\
\mu_k - \xi_i + \mu_j - \xi_j, & k = i \text{ or } j.
\end{cases}$$

Then we have $\mu_i \geq \mu_j \geq \mu_j - \xi_j > \mu_i - \xi'_j = (\mu_i - \xi_i + \mu_j - \xi_j)/2 > 0$. Hence $(\xi'_1, \ldots, \xi'_8)$ is in the domain of $E_\beta$ and $E_\beta(\xi'_1, \ldots, \xi'_8) > E_\beta(\xi_1, \ldots, \xi_8)$ by the arithmetic mean-geometric mean inequality applied to $(\mu_i - \xi_i)(\mu_j - \xi_j)$. This contradicts our assumption on $(\xi_1, \ldots, \xi_8)$. Therefore the second condition must hold;

3. If $\xi_j > 0$ and $\mu_1 - \xi_1 > \mu_j - \xi_j$ for some $j \geq 2$, WLOG we can assume that $\mu_1 - \xi_1 = \mu_k - \xi_k$ for any $k < j$. Then we can choose some $0 < \epsilon < \xi_j$ such that $\mu_{j-1} - (\xi_{j-1} + \epsilon) \geq \mu_j - (\xi_j - \epsilon)$. Notice that $[\mu_{j-1} - (\xi_{j-1} + \epsilon)][\mu_j - (\xi_j - \epsilon)] > (\mu_{j-1} - \xi_{j-1})(\mu_j - \xi_j)$, we have

$$E_\beta(\xi_1, \ldots, \xi_{j-1} + \epsilon, \xi_j - \epsilon, \ldots, \xi_8) > E_\beta(\xi_1, \ldots, \xi_8).$$

This contradicts our assumption on $(\xi_1, \ldots, \xi_8)$. Therefore the third condition must hold.

Now we will give an upper bound for $\sup E_\beta$ given the above lemma. Notice that $\mu_1 - \xi_1 \leq \sum_{j=2}^{8} (\mu_j - \xi_j)$ if $(\xi_2, \ldots, \xi_8) \neq 0$ and that $\mu_1 \geq \sum_{j=2}^{8} \mu_j$. Therefore $\xi_1 \geq \beta \mu/2 \geq \sum_{j=2}^{8} \xi_j$. Notice that $\xi_1 \geq \beta \mu/2 \geq \sum_{j=2}^{8} \xi_j$ is trivially true when $(\xi_2, \ldots, \xi_8) = 0$, we then have

$$\sup E_\beta = \prod_{j=1}^{8} \frac{\mu_j - \xi_j}{\mu_j} \leq \frac{\mu_1 - \xi_1}{\mu_1} \leq \frac{\mu_1 - \beta \mu}{\mu_1}.$$  

Hence

$$\det(B - CA^{-1}C) \leq \mu_1 \left(\frac{\mu - \mu_1}{7}\right)^7 \sup E_\beta \leq \left(\mu_1 - \frac{\beta \mu}{2}\right) \left(\frac{\mu - \mu_1}{7}\right)^7. \quad (41)$$

We are now in a position to find conditions on $\alpha$ such that (37) holds. Given (40) and (41), it suffices to find $\alpha$ such that

$$[1 - (1 - \alpha)^2]^2(\mu - \mu_1)^2(2\mu_1[1 - (1 - \alpha)^2] + (1 - \alpha)^2(\mu - \beta(1 - \alpha)^2)) \mu^5$$
Moreover, to make \((43)\) hold in order to have a nicer expression to work with. Then this leads to the assumption that \(\alpha \geq 1 - \sqrt{2/3}\). Then \((1-\alpha)^2 \leq 2/3\) and \[1-(1-\alpha)^2\] \(\geq 1/3\). Therefore
\[
2\mu_1[1-(1-\alpha)^2] + (1-\alpha)^2\mu = 2\mu_1 - (1-\alpha)^2(2\mu_1 - \mu).
\]
This implies that
\[
\frac{2\mu_1[1-(1-\alpha)^2] + (1-\alpha)^2\mu - \beta (1-\alpha)^2\mu}{\mu_1 - \frac{\beta\mu}{2}} \geq 2(1-\alpha)^2, \quad \forall \ 0 \leq \beta \leq 1.
\] (43)

The idea behind assuming \(\alpha \geq 1 - \sqrt{2/3}\) comes from some “backward thinking”. We want to make (43) hold in order to have a nicer expression to work with. Then this leads to the assumption that \(\alpha \geq 1 - \sqrt{2/3}\). Let \(\tau = (\mu - \mu_1)/\mu \in [0, 1/2]\), then by (43) applied to (42) it suffices to find an \(\alpha \geq 1 - \sqrt{2/3}\) such that
\[
[1-(1-\alpha)^2]^2 \geq \frac{7\tau^5}{2(1-\alpha)^2}, \quad \forall \tau \in [0, 1/2).
\]
Equivalently we want to find an \(\alpha \geq 1 - \sqrt{2/3}\) such that
\[
[1-(1-\alpha)^2]^2 (1-\alpha)^2 = [1-(1-\alpha)^2]^2 - [1-(1-\alpha)^2]^3 \geq \frac{7}{26}.
\]
Notice that the function \(x(t) := t^2 - t^3\) is increasing on \([0, 2/3]\) and decreasing on \([2/3, 1]\). Moreover, \(x(4/9) = 80/729 > 7/2^6\) (since \((80/729) \cdot (2^6/7) = 512/5103 > 1\) and \(x(3/4) = 9/2^6 > 7/2^6\). Therefore we can choose \(\alpha\) such that
\[
\frac{4}{9} \leq [1-(1-\alpha)^2] \leq \frac{3}{4}.
\]
This is equivalent to
\[
\frac{1}{2} \geq \alpha \geq 1 - \frac{\sqrt{5}}{3} \in [0.254, 0.255].\] (44)
Hence any \(\alpha\) satisfying (44) also satisfies the first requirement of Lemma 4.4 in Case 2.

It remains for us to find conditions for \(\alpha\) satisfying the second requirement of Lemma 4.4.

**Step 2: Find the conditions for \(\alpha\) such that the second inequality in Lemma 4.4 holds in Case 2.**

Recall that in [3, B.4 Lemme] Besson–Courtois–Gallot proved that for any \(1/(n-1) \leq s \leq 1\) the function
\[
H \rightarrow \frac{\det(H)^s}{\det(\text{Id} - H)}, \quad \text{tr}(H) = 1 \text{ and } H \geq 0
\]
defined on \(n \times n\) trace 1 positive semi-definite matrices achieves global maximum when \(H = \text{Id}/n\). This proves (19) as a corollary. In particular, the second inequality in Lemma 4.4 (see also (26)) holds when \(\alpha = 0\). However, \(\alpha = 0\) does not satisfy (44). For non-zero \(\alpha\), the second inequality in Lemma 4.4 may still hold in Case 2 if \(\alpha\) satisfies the following properties.
(1) $\det(\text{Id} - H - U_\alpha)$ is “not too small” compared with $\det(\text{Id} - H)$. The lower bound between their ratio is dependent on $\alpha$;

(2) $\det(H)^{4/15}/\det(\text{Id} - H)$ in this case is “much smaller” than $16^{16(1-4/15)}. 15^{16} = \det(\text{Id}/16)^{4/15}/\det(15\text{Id}/16)$ since $H$ differs a lot from $\text{Id}/16$ by the assumption of Case 2. In particular, the upper bound between their ratio should be smaller than the ratio in the first property.

To be more specific, we first rewrite the second inequality in Lemma 4.4 as

\[
\frac{\det(H)^{4/15}}{\det(\text{Id} - H - U_\alpha)} = \frac{\det(H)^{4/15}}{\det(\text{Id} - H)} \frac{\det(\text{Id} - H)}{\det(\text{Id} - H - U_\alpha)} \leq \frac{16^{16(1-4/15)}}{15^{16}}.
\]

(45)

Based on (45), the main idea to find conditions on $\alpha$ is to find upper bounds $K_1(\alpha), K_2$ for $\det(\text{Id} - H)/\det(\text{Id} - H - U_\alpha)$ and $\det(H)^{4/15}/\det(\text{Id} - H)$ respectively. Then the conditions on $\alpha$ will be given in the form of

\[
K_1(\alpha) K_2 \leq \frac{16^{16(1-4/15)}}{15^{16}} = \frac{\det(\text{Id}/16)^{4/15}}{\det(15\text{Id}/16)}.
\]

(46)

**Finding an upper bound** $K_1(\alpha)$ for $\det(\text{Id} - H)/\det(\text{Id} - H - U_\alpha)$: Since both $H$ and $\hat{H} - H$ are positive semi-definite, by their matrix forms in (16) and (17) we have

\[
\begin{pmatrix}
A & \pm C \\
\pm C^* & B
\end{pmatrix},
\begin{pmatrix}
\lambda \cdot \text{Id} - A & \pm C \\
\pm C^* & \mu \cdot \text{Id} - B
\end{pmatrix}
\]

are positive semi-definite. Therefore

\[
\begin{pmatrix}
\lambda \cdot \text{Id} & 2C \\
2C^* & \mu \cdot \text{Id}
\end{pmatrix} \geq \begin{pmatrix}
A & C \\
C^* & B
\end{pmatrix} \geq 0,
\]

Notice from Fact 4.3 that

\[
\begin{pmatrix}
\lambda \cdot \text{Id} & 2C \\
2C^* & \mu \cdot \text{Id}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\text{Id} & 0 \\
2C^*/\lambda & \text{Id}
\end{pmatrix} \begin{pmatrix}
\lambda \cdot \text{Id} & 0 \\
0 & \mu \cdot \text{Id} - 4\lambda^{-1}C^*C
\end{pmatrix} \begin{pmatrix}
\text{Id} & 2C/\lambda \\
0 & \text{Id}
\end{pmatrix},
\]

we have $\lambda \mu \cdot \text{Id} \geq 4C^*C$, which implies that

\[
\begin{pmatrix}
\mu \cdot \text{Id} & -2C \\
-2C^* & \lambda \cdot \text{Id}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\text{Id} & 0 \\
-2C^*/\mu & \text{Id}
\end{pmatrix} \begin{pmatrix}
\mu \cdot \text{Id} & 0 \\
0 & \lambda \cdot \text{Id} - 4\mu^{-1}C^*C
\end{pmatrix} \begin{pmatrix}
\text{Id} & -2C/\mu \\
0 & \text{Id}
\end{pmatrix}
\]

is positive semi-definite. This implies that

\[
\text{Id} - H - 2U = \begin{pmatrix}
\text{Id} - A & -3C \\
-3C^* & \text{Id} - B
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\mu \cdot \text{Id} & -2C \\
-2C^* & \lambda \cdot \text{Id}
\end{pmatrix} + \begin{pmatrix}
\lambda \cdot \text{Id} - A & -C \\
-C^* & \mu \cdot \text{Id} - B
\end{pmatrix}
\]

is positive semi-definite. In particular,

\[
\text{Id} - B \geq 9C^*(\text{Id} - A)^{-1}C.
\]

(47)
By Fact 4.3 and (47), we have \( \text{Id} - H - U_\alpha \) is positive semi-definite (which is a prerequisite for using (25)) and
\[
\det(\text{Id} - H - U_\alpha) \\
= \det(\text{Id} - A) \det(\text{Id} - B - (1 + \alpha)^2 C^* (\text{Id} - A)^{-1} C) \\
\geq \det(\text{Id} - A) \det \left[ \frac{9 - (1 + \alpha)^2}{8} \cdot (\text{Id} - B - C^* (\text{Id} - A)^{-1} C) \right] \\
\geq \left[ \frac{9 - (1 + \alpha)^2}{8} \right]^8 \det(\text{Id} - A) \det(\text{Id} - B - C^* (\text{Id} - A)^{-1} C) \\
= \left[ \frac{9 - (1 + \alpha)^2}{8} \right]^8 \det(\text{Id} - H).
\]

Hence
\[
\frac{\det(\text{Id} - H)}{\det(\text{Id} - H - U_\alpha)} \leq K_1(\alpha) := \left[ \frac{9 - (1 + \alpha)^2}{8} \right]^{-8}. \tag{48}
\]

**Finding an upper bound** \( K_2 \) for \( \det(H)^{4/15} / \det(\text{Id} - H) \): Denote by \( v_1 \geq \cdots \geq v_{16} \geq 0 \) the eigenvalues of \( H \). By the assumptions of Case 2, we have
\[
v_1 \geq \mu_1 \geq \mu - \mu_1 = \sum_{j=2}^{8} \mu_j \geq \sum_{j=1}^{7} v_{16-j+1}. \tag{49}
\]

Define
\[
F(x_1, \ldots, x_{16}) = \prod_{j=1}^{16} \frac{x_j^{4/15}}{1 - x_j}, \quad \sum_{j=1}^{16} x_j = 1 \text{ and } x_j \geq 0.
\]

Then we have
\[
F(v_1, \ldots, v_{16}) = \frac{\det(H)^{4/15}}{\det(\text{Id} - H)}. \tag{50}
\]

Define \( f(t) = \ln(1 - t) + 4 \ln(t) / 15 \) with \( t \in (0, 1) \). Then \( \ln F(x_1, \ldots, x_{16}) = \sum_{j=1}^{16} f(x_j) \). Direct computations yield
\[
f'(t) = \frac{4}{15 t} + \frac{1}{1 - t}, \quad f''(t) = -\frac{4}{15 t^2} + \frac{1}{(1 - t)^2}.
\]

Therefore \( f(t) \) is concave down on \((0, \sqrt[3]{4/15}/(1 + \sqrt[3]{4/15}))\) and concave up on \((\sqrt[3]{4/15}/(1 + \sqrt[3]{4/15}), 1)\). Since \( \sqrt[3]{4/15}/(1 + \sqrt[3]{4/15}) > 1/3 \) and \( v_1 + \cdots + v_{16} = \text{tr}(H) = 1 \), we have \( v_j < \sqrt[3]{4/15}/(1 + \sqrt[3]{4/15}) \) for any \( 3 \leq j \leq 16 \). Therefore we can use concavity arguments to find an upper bound for \( F(v_1, \ldots, v_{16}) \).

**Lemma 4.8** Under the above assumptions on \( F \) and \( f \), there exist some \( v'_1 \geq \cdots \geq v'_{16} \) such that
\[
F(v_1, \ldots, v_{16}) \leq F(v'_1, \ldots, v'_{16}).
\]

Moreover, \( (v'_1, \ldots, v'_{16}) \) satisfies the following conditions
\[
(1) \ v'_2 = \cdots = v'_9 \text{ and } v'_{10} = \cdots = v'_{16}.
\]
(2) $\nu'_1 = \max \left\{ v_1, v_1 + v_2 - \frac{\sqrt{4/15}}{1+\sqrt{4/15}} \right\} \geq \sum_{j=1}^{7} v'_{16-j+1} \geq \sum_{j=1}^{7} v_{16-j+1}$;

(3) If the first inequality in the second condition is strict, then $v'_2 = \cdots = v'_{16}$;

(4) If $v'_2 > v'_{16}$, then $v'_1 = \sum_{j=1}^{7} v'_{16-j+1}$;

(5) $\sum_{j=1}^{16} v'_j = 1$.

**Proof of Lemma 4.8** The main idea behind this inequality is to apply a few adjustments to the original $\nu$’s so that the value of $F$ goes up. We first use the fact that $f(x)$ is concave up on $(\sqrt{4/15}/(1 + \sqrt{4/15}), 1)$ to obtain

$$f(v_1) + f(v_2) \leq f(v'_1) + f(v_1 + v_2 - v'_1),$$

where $v'_1 = \max \left\{ v_1, v_1 + v_2 - \frac{\sqrt{4/15}}{1+\sqrt{4/15}} \right\}$ and $v_1 + v_2 - v'_1 = \min \left\{ v_2, \frac{\sqrt{4/15}}{1+\sqrt{4/15}} \right\}$. Notice that

$$\ln F(v_1, \ldots, v_{16}) = \sum_{j=1}^{16} f(v_j).$$

Therefore

$$F(v_1, \ldots, v_{16}) \leq F(v'_1, v_1 + v_2 - v'_1, v_3, \ldots, v_{16}). \quad (51)$$

Let $x' = (v_{10} + \cdots + v_{16})/7 \leq \sqrt{4/15}/(1 + \sqrt{4/15})$ and $y' = [(v_1 + v_2 - v'_1) + v_3 + \cdots + v_{16}] / 8 \leq \sqrt{4/15}/(1 + \sqrt{4/15})$. It is clear from the ordering of $v_j$ that $y' \geq x'$. Since $\sqrt{4/15}/(1 + \sqrt{4/15}) \geq (v_1 + v_2 - v'_1) \geq v_3 \geq \cdots \geq v_{16}$, by the fact that $f(x)$ is concave down on $(0, \sqrt{4/15}/(1 + \sqrt{4/15}))$, we have

$$f(v_1 + v_2 - v'_1) + f(v_3) + \cdots + f(v_9) \leq 8 f(y')$$

and

$$f(v_{10}) + \cdots + f(v_{16}) \leq 7 f(x').$$

These two inequalities imply that

$$F(v'_1, v_1 + v_2 - v'_1, v_3, \ldots, v_{16}) \leq F(v'_1, y', \ldots, y', x', \ldots, x'), \quad (52)$$

where $y'$ shows up 8 times and $x'$ shows up 7 times.

**Case (1):** If $v'_1 \geq 7(8y' + 7x')/15$, then by the fact that $f(x)$ is concave down on $(0, \sqrt{4/15}/(1 + \sqrt{4/15}))$, we have

$$8 f(y') + 7 f(x') \leq 15 f((8y' + 7x')/15).$$

Choose $v'_2 = \cdots = v'_{16} = (8y' + 7x')/15$. Then

$$F(v'_1, y', \ldots, y', x', \ldots, x') \leq F(v'_1, \ldots, v'_{16}). \quad (53)$$

Lemma 4.8 then follows from (51), (52) and (53).

**Case (2):** If $v'_1 < 7(8y' + 7x')/15$, we let $x = v'_1/7$ and $y = (1 - 2v'_1)/8 = (8y' + 7x' - 7x)/8 > x$. Clearly $8y' + 7x' = 8y + 7x$. Since $v'_1 \geq v_1 \geq v_{10} + \cdots + v_{16} = 7x'$, we have $y' \geq y > x \geq x'$. By the fact that $f(x)$ is concave down on $(0, \sqrt{4/15}/(1 + \sqrt{4/15}))$, we have

$$8 f(y') + 7 f(x') \leq 8 f(y) + 7 f(x).$$
Choose \( v'_2 = \cdots = v'_9 = y \) and \( v'_{10} = \cdots = v'_{16} = x \), we have
\[
F(v'_1, y', \ldots, y', x', \ldots, x') \leq F(v'_1, \ldots, v'_{16}).
\] (54)

Lemma 4.8 then follows from (51), (52) and (54).

Back to finding an upper bound \( K_2 \) for \( \det(H)^{4/15} / \det(\text{Id} - H) = F(v_1, \ldots, v_{16}) \). Let
\( v'_1 \geq v'_2 = \cdots = v'_9 = y \geq v'_{10} = \cdots = v'_{16} = x \) be the same as in Lemma 4.8. Then we have \( v'_1 \geq 7x \) and \( y \geq x \). Moreover, \( v'_1 > 7x \) implies that \( y = x \).

**Case 2a:** \( v'_1 = 7x \). Define
\[
L(t) = \frac{(7t^8)^{4/15}}{(1-t)^7(1-7t)} \cdot \left[ \frac{(14t/8)^{32/15}}{(1-14t/8)^8} \right]^{-1}, \quad 0 \leq t \leq \frac{1}{22}.
\]
Then by [3, B.4 Lemma] we have
\[
F(7x, y, \ldots, y, x, \ldots, x) = L(x)F(14x/8, y, \ldots, y, 14x/8, \ldots, 14x/8) \leq L(x)F(\text{Id}/16),
\]
where \( y \) shows up 8 times. Notice that
\[
\left( \ln L(t) \right)' = \frac{7}{1-7t} + \frac{7}{1-t} - \frac{14}{(1-14t/8)} \geq 0.
\]
Therefore by Lemma 4.8 and (50) we have
\[
\frac{\det(H)^{4/15}}{\det(\text{Id} - H)} = F(v_1, \ldots, v_{16}) \leq F(v'_1, \ldots, v'_{16}) \leq L(1/22)F(1/16, \ldots, 1/16) = L(1/22) \cdot \frac{16^{16(1-4/15)}}{15^{16}} =: K_2. \quad (55)
\]

**Case 2b:** \( v_1 > 7x \) and therefore \( y = x \). Let
\[
R(t) = F(1-15t, t, \ldots, t) = \frac{(1-15t)^{4/15}t^4}{15t(1-t)^{15}}, \quad 0 \leq t \leq \frac{1}{22}.
\]
Then
\[
\left( \ln R(t) \right)' = -\frac{4}{1-15t} + \frac{3}{t} + \frac{15}{1-t} > 0.
\]
Therefore \( R(t) \) is increasing and
\[
F(1-15t, t, \ldots, t) \leq F(7/22, 1/22, \ldots, 1/22) \leq L(1/22)F(1/16, \ldots, 1/16).
\]
Hence this case reduces to Case 2a.

By (46), (48) and (55) it suffices to find an \( \alpha \) satisfying (44) and
\[
\left[ \frac{9 - (1 + \alpha)^2}{8} \right]^8 \cdot L(1/22) \leq 1.
\]
The above inequality is equivalent to
\[
\alpha \leq -1 + \sqrt{9 - 8L(1/22)^{1/8}} \in [0.266, 0.267], \quad (56)
\]
which gives a sufficient condition on \( \alpha \) such that the second requirement in Lemma 4.4 holds in Case 2.

Notice that the lower bound in (44) is smaller than 0.26 and the smallest upper bound for \( \alpha \) is greater than 0.26. We can choose \( \alpha = 0.26 \) and Lemma 4.4 follows in Case 2. \( \square \)
Remark (1). In fact, the equality of the second inequality in Lemma 4.4 can not be achieved in Case 2 since (56) is strict when $\alpha = 0.26$. To be specific, (48), (55) and (56) imply that $0 < K_1(0.26)K_2 < \frac{16^{16(1-4/15)}}{15^{16}}$. Therefore by (45) and (46) we have

$$\det(\text{Id} - H - U_\alpha) \geq \frac{\det(H)^{4/15}}{K_1(0.26)K_2} > \frac{15^{16}\det(H)^{4/15}}{16^{16(1-4/15)}}.$$ 

Since the first inequality in Lemma 4.4 holds in Case 2 when $\alpha = 0.26$, it follows from (25) and the above inequality that

$$\det(\text{Id} + \hat{H} - 2H)$$

$$\geq 22^{16}\det((\text{Id} - H - U_\alpha)/15)^{15/22}\det((\hat{H} - H + U_\alpha)/7)^{7/22}$$

$$\geq 22^{16}15^{-(15/22)-16}K_1(0.26)^{-15/22}K_2^{-15/22}\det(H)^{4/22+7/22}$$

$$> 22^{16}15^{-(15/22)-16}(16^{16(1-4/15)}15^{-16})^{-15/22}\det(H)^{1/2}$$

$$= 22^{16}6^{-8}\det(H)^{1/2}. \quad (57)$$

This shows that the equality in (15) (and equivalently Proposition 4.1) cannot be achieved under the assumptions of Case 2.

(2). The original proof of [3] actually proved a stronger inequality in the sense that the power of $\det(H)$ in the numerator can be reduced to $d/(n + d - 2)$, where $d = 1, 2, 4$ in the real, complex and quaternionic setting respectively and $n$ denotes the dimension of the symmetric space. The estimates in our proof are clearly not optimal since we can choose $\alpha$ from the interval $[1 - \sqrt{5}/3, 0.266]$. Using the same method we can reduce the power of $\det(H)$ in the numerator to $-\epsilon + 1/2$ for some small $\epsilon > 0$, which is sufficient to prove entropy rigidity in the Cayley hyperbolic setting but weaker than the corresponding inequality in [3]. We suspect that in order to prove the stronger inequality as in [3], we need some stronger estimates on $\det(\hat{H} - H)$ or some new ideas.

4.4 Repairing the gap: proof of (8)

Following the notations in previous subsections, we have the following analogue of [3, B.5 Proposition], which proves (8) in the Cayley hyperbolic setting.

Corollary 4.9 Let $H$ and $\hat{H}$ be as in Proposition 4.1. Then there exists a constant $K$ such that

$$\frac{(\det H)^{1/2}}{\det(\text{Id} - 2H + \hat{H})} \leq \left(\frac{\sqrt{16}}{16 + 8 - 2}\right)^{16}\left(1 - K\sum_{j=1}^{16}(\nu_j - \frac{1}{16})^2\right), \quad (58)$$

where $\nu_1, \ldots, \nu_{16}$ denote the eigenvalues of $H$.

Proof Recall that in Proposition 4.1 we proved that under a suitable basis one can write

$$H = \begin{pmatrix} A & C \\ C^* & B \end{pmatrix}$$

and

$$\hat{H} = \begin{pmatrix} \lambda \cdot \text{Id} & 0 \\ 0 & \mu \cdot \text{Id} \end{pmatrix},$$
where $A$, $B$ are $8 \times 8$ diagonal matrices with trace equal to $\lambda$, $\mu \geq 0$ respectively such that $\lambda + \mu = 1$. Moreover,

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_8 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_8 \end{pmatrix},$$

where $\lambda_1 \geq \cdots \geq \lambda_8 \geq 0$ and $\mu_1 \geq \cdots \geq \mu_8 \geq 0$ are eigenvalues of $A$ and $B$ respectively. Similar to the proof of Proposition 4.1, we consider two separate cases.

Case 1. If $2\lambda_1 \leq \lambda$ and $2\mu_1 \leq \mu$, then (18) and (36) imply that it suffices to prove

$$\frac{\det(H)^{1/2 - 7/22}}{\det(I - H)^{15/22}} \leq \frac{16^8}{(1516)^{15/22}} \left( 1 - K \sum_{j=1}^{16} \left( \nu_i - \frac{1}{16} \right)^2 \right)$$

for some constant $K$. This follows directly from the proof of [3, Proposition B.5].

Case 2. Under the assumptions of Case 2 in the proof for Lemma 4.4, the left hand side of (15) has a strictly smaller upper bound comparing to its right hand side. This follows from (57) in the remark of Lemma 4.4. If $2\lambda_1 > \lambda$ or $2\mu_1 > \mu$, then the corollary directly follows from the fact that $\sum_{j=1}^{16} \left( \nu_i - \frac{1}{16} \right)^2 < 16$.

\[ \square \]

**Remark** Equivalently, Sects. 4.2–4.4 guarantee that all results in [3, Appendix B] are correct in the Cayley hyperbolic setting.

**Acknowledgements** I would like to heartily thank my advisor Ralf Spatzier for his support during the entire work. I sincerely thank Gérald Besson, Gilles Courtois and Sylvain Gallot for patiently proofreading this paper and providing valuable suggestions, especially for pointing out a mistake in an earlier version of the manuscript. I am also very grateful to Chris Connell for helpful and thorough discussions on this subject.

**Data availability** No datasets were generated or analysed during the current study.

**References**

1. Albuquerque, P.: Patterson–Sullivan theory in higher rank symmetric spaces. Geom. Funct. Anal. 9(1), 1–28 (1999)
2. Baez, J.C.: The octonions. Bull. Am. Math. Soc. (N.S.) 39(2), 145–205 (2002)
3. Besson, G., Courtois, G., Gallot, S.: Entropies et rigidités des espaces localement symétriques de courbure strictement négative. (French) [Entropy and rigidity of locally symmetric spaces with strictly negative curvature]. Geom. Funct. Anal. 5(5), 731–799 (1995)
4. Besson, G., Courtois, G., Gallot, S.: Minimal entropy and Mostow’s rigidity theorems. Ergodic Theory Dyn. Syst. 16(4), 623–649 (1996)
5. Boland, J., Connell, C., Souto, J.: Volume rigidity for finite volume manifolds. Am. J. Math. 127(3), 535–550 (2005)
6. Connell, C., Farb, B.: Minimal entropy rigidity for lattices in products of rank one symmetric spaces. Commun. Anal. Geom. 11(5), 1001–1026 (2003)
7. Connell, C., Farb, B.: The degree theorem in higher rank. J. Differ. Geom. 65(1), 19–59 (2003). (Erratum for The degree theorem in higher rank. J. Differential Geom. 105(1), 21-32 (2017))
8. Mostow, G.D.: Strong Rigidity of Locally Symmetric Spaces. Annals of Mathematics Studies, vol. 78. Princeton University Press, University of Tokyo Press, Princeton (1973)
9. Parker, J.R.: Hyperbolic Spaces. Jyväskylä Lectures in Mathematics 2 (2008)
10. Ruan, Y.: Filling volume minimality and boundary rigidity of metrics close to a negatively curved symmetric metric. arXiv:2201.09175 [math.DG]
11. Springer, T.A., Veldkamp, F.D.: Elliptic and hyperbolic octave planes I, II, III, vol. 66, pp. 413–451. Springer, Berlin (1963)
12. Springer, T.A., Veldkamp, F.D.: Octonions, Jordan Algebras and Exceptional Groups. Springer Monographs in Mathematics. Springer, Berlin (2000). (ISBN: 3-540-66337-1)
13. Storm, P.A.: The minimal entropy conjecture for nonuniform rank one lattices. Geom. Funct. Anal. 16(4), 959–980 (2006)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.