Received: 11 December 2025 | Revised: 19 May 2026 | Accepted: 18 June 2026

DOI: 10.1112/plms.70182

RESEARCH ARTICLE

Proceedings of the London Mathematical Society

An elementary approach to Wehrl-type entropy bounds in quantitative form

Fabio Nicola | **Federico Riccardi** | **Paolo Tilli**

Dipartimento di Scienze Matematiche, Politecnico di Torino, Torino, Italy

Correspondence
Federico Riccardi, Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy.
Email: federico.riccardi@polito.it

Abstract
We consider the problem of the stability (with sharp exponent) of the Lieb–Solovej inequality for symmetric $SU(N)$ coherent states, which was obtained only recently by the authors. Here, we propose an elementary proof of this result, based on reformulating the Wehrl-type entropy as a function defined on the unit sphere in $\mathbb{C}^d$, for some suitable d, and on some explicit (and somewhat surprising) computations.

MSC 2020
81R30, 22E70, 30H20, 32A10, 47N50, 49J40

1 | INTRODUCTION

Recent years have witnessed an intense and fruitful interest in various questions related to the *Wehrl-type entropy conjecture*. The origins of this problem date back to the late 1970s, when Wehrl [23] introduced a notion of classical entropy for quantum-mechanical density matrices on $L^2(\mathbb{R})$ and conjectured that Glauber coherent states, which are coherent states in representations of the Heisenberg group, are the only optimizers for this entropy. Within a short period, Lieb [16] proved that coherent states are indeed the optimizers not only of Wehrl's original entropy, but also of other kinds of entropies (where the function $x \ln x$ in Wehrl's definition is replaced with convex powers x^p). The uniqueness of the optimizers was later established by Carlen [2]. In his paper, Lieb also suggested that an analogous result should hold for Bloch coherent spin states, corresponding to the irreducible representations of $SU(2)$. This new conjecture was proved in full generality (i.e. with an arbitrary convex function in the definition of the entropy) several years later by Lieb and Solovej [18], who further generalized their proof [19] obtaining a similar result for the symmetric $SU(N)$ coherent states ($N \geqslant 2$) (see also [5, 22] for the full conjecture on Glauber coherent states and [12] for $SU(1,1)$ coherent states). However, the question of uniqueness of the optimizers for the $SU(2)$ and the $SU(N)$ Wehrl-type entropy remained open until recently. For the $SU(2)$ case, it was proved independently and simultaneously in [5] and in [13], while the $SU(N)$ case was finally settled in [21]. Along with the question of uniqueness, the issue of stability was also addressed in recent years. Roughly speaking, the goal of a stability estimate is to prove and quantify the idea that states with almost minimal entropy should be almost coherent states. The stability of the Lieb–Solovej inequality for $SU(2)$ coherent states was proved in [8], while for symmetric $SU(N)$ coherent states it was proved in [21] (see also [6, 9, 10, 20] for the stability of the estimate in the case of Glauber and $SU(1,1)$ coherent states). In all these cases, the proof relies on a careful and highly nontrivial study of the measure of the super-level sets of the Husimi function (see below) using a technique first introduced in [9, Lemma 2.1].

The main goal of this paper is to prove the stability of the estimate for the Wehrl-type entropy conjecture for symmetric $SU(N)$ coherent states using only elementary tools and avoiding any study of the measure of the super-level sets of the Husimi function. In fact, to simplify matters and avoid additional technicalities, we will focus on the case of pure states, although this method could likely be adapted to handle density matrices as well. The main idea of the proof is to rephrase the problem in order to express the Wehrl-type entropy as a function on the unit sphere in a high-dimensional Euclidean space. We compute the second differential of the entropy on the submanifold where the entropy achieves its minimum value, and we can show that it is uniformly positive definite on the normal bundle to the submanifold (in fact, we will instead consider an equivalent formulation with a convex function rather than a concave one, which corresponds to a maximization problem).

Although this approach is elementary, it requires some lengthy—and somewhat surprising—computations. In order to make things clearer, we have organized the paper as follows.

In Section 2, we fix the notation and rephrase the problem in order to write the Wehrl-type entropy as a function on the unit sphere in $\mathbb{C}^d$, for a suitable dimension d. This geometric setting requires some preliminary notions from differential geometry, which are gathered at the end of the section. In Section 3, we prove the stability result under an additional smoothness assumption (which is satisfied in the special cases of interest; see Remark 3.3 below). This enables us to present the core concept of the proof without encountering further technical complications. The proof of the result in full generality is then given in Section 4 and will require some technical (and more cumbersome) lemmas that are collected in Appendix.

2 | NOTATION AND REPHRASING OF THE PROBLEM IN TERMS OF HOLOMORPHIC POLYNOMIALS

2.1 | Notation

For $z = (z_1, \dots, z_N)$, $w = (w_1, \dots, w_N) \in \mathbb{C}^N$ and a multi-index $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$ we adopt the standard notation

$$z^\alpha = z_1^{\alpha_1} \cdots z_N^{\alpha_N}, \quad |z|^2 = |z_1|^2 + \cdots + |z_N|^2, \quad z \cdot w = z_1 w_1 + \cdots + z_N w_N,$$

$$|\alpha| = \alpha_1 + \cdots + \alpha_N, \quad \alpha! = \alpha_1! \cdots \alpha_N!.$$

Given $M \in \mathbb{N}$ and a multi-index $\alpha \in \mathbb{N}^N$ of length at most M we have

$$\binom{M}{\alpha} = \frac{M!}{\alpha_1! \cdots \alpha_N!(M-|\alpha|)!}.$$

Dealing with sums over a multi-index, we denote by

$$\sum_{|\alpha|\leqslant M}$$

the sum over all multi-indices $\alpha \in \mathbb{N}^N$ of total length at most M, and by

$$\sum_{\alpha\neq 0}$$

the sum over all multi-indices of total length at most M except for the null multi-index, the condition $|\alpha| \leqslant M$ being understood (this should cause no confusion, since M is fixed throughout the paper). Similarly, we denote by

$$\sum_{\alpha\neq\beta}$$

the double sum over all multi-indices α and β (of length at most M) that are not equal (i.e. they differ in at least one component).

2.2 | From the original setting to the unit sphere

The original theorem by Lieb and Solovej [19] is stated in terms of symmetric representations of $SU(N+1)$ as follows. Fix $N \geqslant 1$, $M \geqslant 1$ and consider the Hilbert space $\mathcal{H}_M = (\bigotimes^M \mathbb{C}^{N+1})_{\mathrm{sym}}$, that is the image of $\bigotimes^M \mathbb{C}^{N+1}$ under the projection

$$P_M(v_1 \otimes \cdots \otimes v_M) = \frac{1}{M!}\sum_{\sigma\in S_M} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(M)},$$

where S_M denotes the permutation group over $\{1,\dots,M\}$. We denote by $\langle\psi|\phi\rangle$ the inner product of $\psi, \phi \in \mathcal{H}_M$, with the agreement that it is linear in the second argument. On $\mathcal{H}_M$, we consider the representation of the group $SU(N+1)$ given by

$$R(v_1 \otimes \cdots \otimes v_M) = (Rv_1) \otimes \cdots \otimes (Rv_M), \qquad R \in SU(N+1)$$

and extended on the whole $\mathcal{H}_M$ by linearity. Fix a normalized vector $v \in \mathbb{C}^{N+1}$. Given a density matrix ρ (i.e. a nonnegative self-adjoint operator on $\mathcal{H}_M$ with trace one), its *Husimi function* u_ρ : $SU(N+1) \to \mathbb{R}$ is defined as

$$u_\rho(R) = \langle \otimes^M Rv|\rho| \otimes^M Rv\rangle, \quad R \in SU(N+1).$$

It is easy to see [19] that u_ρ is continuous (in fact real-analytic), and satisfies $0 \leqslant u_\rho \leqslant 1$, as well as

$$\dim(\mathcal{H}_M) \int_{SU(N+1)} u_\rho(R)\, dR = 1, \tag{2.1}$$

where dR denotes the Haar probability measure on $SU(N+1)$.

Then, the Lieb–Solovej inequality [19, Theorem 2.1] states that, for any convex function $\Phi :\ [0,1] \to \mathbb{R}$, the Wehrl-type entropy

$$G(\rho) = \int_{SU(N+1)} \Phi(u_\rho(R))\, dR$$

is maximized by (symmetric) coherent states, that is, when

$$\rho = | \otimes^M v\rangle\langle \otimes^M v|$$

for some $v \in \mathbb{C}^{N+1}$ of norm 1. In fact, with a slight abuse of language, we will call a coherent state any vector of the type $\otimes^M v \in \mathcal{H}_M$, with $v \in \mathbb{C}^{N+1}$, $|v| = 1$.

Remark 2.1. Note that any coherent state can be written as $\otimes^M R v_0$ for some $R \in SU(N+1)$, where $v_0 = (1, 0, \dots, 0) \in \mathbb{C}^{N+1}$. In other words, the subset of coherent states is the orbit in $\mathcal{H}_M$ of the coherent state associated with the vector v_0 under the action of $SU(N+1)$.

Note that, when Φ is affine, G is constant (as easily seen from the normalization (2.1)), whereas if Φ is strictly convex in some interval $(\ell, 1)$ with $\ell \in (0,1)$, then coherent states are the only maximizers (see [21, Theorem 1.2 and Remark 4.3]). Moreover, the estimate is stable, meaning that a density matrix ρ that almost maximizes the functional G is almost a coherent state. To make this precise, we introduce the distance, in the trace norm, between ρ and the subset of coherent states, that is,

$$D[\rho] = \inf_{v \in \mathbb{C}^{N+1}, |v|=1} \|\rho - | \otimes^M v\rangle\langle \otimes^M v|\|_1 .$$

Theorem 2.2 [21]. *For every convex function $\Phi :\ [0,1] \to \mathbb{R}$ that is strictly convex in some interval $(\ell, 1)$ with $\ell \in (0,1)$, there exists a constant $c > 0$ such that for every density matrix ρ on $\mathcal{H}_M$ we have*

$$\int_{SU(N+1)} \Phi(u_0(R))\, dR - \int_{SU(N+1)} \Phi(u_\rho(R))\, dR \geqslant c D[\rho]^2,$$

where u_0 is the Husimi function of any coherent state.

Remark 2.3. Actually, this theorem holds as long as Φ is not affine in $(\ell, 1)$ for every $\ell \in (0,1)$ (see [21, Remark 4.4]). However, for the sake of clarity, we will work under the slightly stronger assumption that Φ is strictly convex in some interval $(\ell, 1)$, for some $\ell \in (0,1)$.

As already mentioned in the introduction, the goal of this paper is to give an elementary proof of this stability estimate for pure states (i.e. when $\rho = |\psi\rangle\langle\psi|$ for some $\psi \in \mathcal{H}_M$). In order to pursue this goal, we need to rephrase the problem in a more geometrical setting. Following the argument presented in [21, section 3.1] one can show that, for pure states, the Wehrl-type entropy can be written as

$$\int_{\mathbb{C}^N} \Phi\left(\frac{|F(z)|^2}{(1+|z|^2)^M}\right) d\nu(z),$$

where

$$F(z) = \langle \otimes^M (1, \overline{z}) | \psi \rangle, \quad z \in \mathbb{C}^N \tag{2.2}$$

is a holomorphic polynomial (since the inner product is linear in the second argument) on $\mathbb{C}^N$ of degree at most M and $d\nu$ is a probability measure (see below for the explicit expression). Since the correspondence between unit pure state vectors and suitably normalized holomorphic polynomials of this kind is one-to-one, we are naturally led to consider the reproducing kernel Hilbert space $\mathcal{P}_M$ of holomorphic polynomials on $\mathbb{C}^N$ of degree at most M, endowed with the norm

$$\|F\|^2_{\mathcal{P}_M} = \dim(\mathcal{P}_M) \int_{\mathbb{C}^N} \frac{|F(z)|^2}{(1+|z|^2)^M}\, d\nu(z), \tag{2.3}$$

where

$$\dim(\mathcal{P}_M) = \binom{M+N}{N} =: d \tag{2.4}$$

and

$$d\nu(z) = c_N (1+|z|^2)^{-N-1} dA(z), \quad c_N = N!/\pi^N \tag{2.5}$$

is a probability measure on $\mathbb{C}^N$; here $dA(z)$ denotes the Lebesgue measure on $\mathbb{C}^N$. The normalized monomials

$$e_\alpha(z) = c_\alpha z^\alpha, \quad z \in \mathbb{C}^N,\ \alpha \in \mathbb{N}^N,\ |\alpha| \leqslant M,$$

where $c_\alpha = \sqrt{\binom{M}{\alpha}}$, form an orthonormal basis of $\mathcal{P}_M$, therefore, we immediately have the expression of the reproducing kernel

$$K(z,w) := \sum_{|\alpha| \leqslant M} e_\alpha(z) \overline{e_\alpha(w)} = (1 + z \cdot \overline{w})^M,$$

whereas the normalized reproducing kernel at a point $w \in \mathbb{C}^N$ is

$$k_w(z) := \frac{K(z,w)}{\|K(\cdot, w)\|_{\mathcal{P}_M}} = \frac{(1+z\cdot\overline{w})^M}{(1+|w|^2)^{M/2}}, \quad z \in \mathbb{C}^N.$$

With a slight abuse of notation, we define the Wehrl-type entropy of a normalized polynomial $F \in \mathcal{P}_M$ as

$$G(F) = \int_{\mathbb{C}^N} \Phi\left(\frac{|F(z)|^2}{(1+|z|^2)^M}\right) d\nu(z). \tag{2.6}$$

Among all normalized polynomials in $\mathcal{P}_M$ this entropy is maximized by the normalized (multiples of the) reproducing kernels, that is, on the subset

$$V := \{e^{i\theta} k_w \;:\; w \in \mathbb{C}^N,\ \theta \in \mathbb{R}\} \subset \mathcal{P}_M,$$

since the normalized (multiples of the) reproducing kernels are exactly the polynomials that arise from coherent states through the formula (2.2). Thus, the supremum of G over all possible normalized polynomials in $\mathcal{P}_M$ is achieved at any normalized reproducing kernel, and choosing the reproducing kernel at $w = 0$ (i.e. the constant function 1) we obtain

$$\sup G = \int_{\mathbb{C}^N} \Phi\left(\frac{1}{(1+|z|^2)^M}\right) d\nu(z).$$

Given a normalized polynomial $F \in \mathcal{P}_M$, the distance from F to V is, by definition,

$$D[F] = \inf\{\|F - e^{i\theta} k_w\|_{\mathcal{P}_M} \;:\; w \in \mathbb{C}^N,\ \theta \in \mathbb{R}\}.$$

Then, Theorem 2.2 (for pure states) can be rephrased as follows.

Theorem 2.4. *For every convex function $\Phi : [0,1] \to \mathbb{R}$ that is strictly convex in some interval $(\ell, 1)$ with $\ell \in (0,1)$, there exists a constant $c > 0$ such that*

$$\sup G - G(F) \geqslant c D[F]^2$$

for every normalized polynomial $F \in \mathcal{P}_M$.

Remark 2.5. Let $\rho = |\psi\rangle\langle\psi|$ and let $F \in \mathcal{P}_M$ be the polynomial associated with the pure state $\psi \in \mathcal{H}_M$. Since unit pure state vectors are in one-to-one correspondence with normalized polynomials in $\mathcal{P}_M$, and since we are dealing with finite-dimensional spaces, the distances $D[\rho]$ and $D[F]$ are equivalent, meaning that there is a constant $C > 0$ such that

$$C^{-1} D[\rho] \leqslant D[F] \leqslant C D[\rho].$$

For this reason, Theorem 2.2 (for pure states) and Theorem 2.4 are equivalent.

To prove Theorem 2.4, we will not work directly with polynomials, but rather with their coefficients in the monomial basis, expressing a normalized polynomial $F \in \mathcal{P}_M$ as

$$F(z) = \sum_{|\alpha| \leqslant M} X_\alpha e_\alpha(z), \quad z \in \mathbb{C}^N,$$

where the vector $X = (X_\alpha)_{|\alpha|\leqslant M}$, due to the orthonormality of the e_α's, belongs to the $2d-1$ dimensional sphere $S := S^{2d-1} \subset \mathbb{R}^{2d} \simeq \mathbb{C}^d$. Hence, with a slight abuse of notation, for $X = (X_\alpha) \in S$ we can write

$$G(X) = \int_{\mathbb{C}^N} \Phi\left(\frac{|\sum_{|\alpha|\leqslant M} X_\alpha e_\alpha(z)|^2}{(1+|z|^2)^M} \right) d\nu(z). \tag{2.7}$$

In this way, we can regard the Wehrl-type entropy as a function $G : S \to \mathbb{R}$, which achieves its maximum at every point of the set

$$V := \left\{ X \in S : \sum_{|\alpha|\leqslant M} X_\alpha e_\alpha = e^{i\theta} k_w \text{ for some } \theta \in \mathbb{R}, w \in \mathbb{C}^N \right\}.$$

In the following, we will see that V is a smooth submanifold of S, of dimension $2N+1$.

Endowing S with the Riemannian metric induced by $\mathbb{R}^{2d} \simeq \mathbb{C}^d$, and denoting by $\operatorname{dist}(X, V)$ the geodesic distance from $X \in S$ to V, we can restate Theorem 2.4 as follows.

Theorem 2.6. *For every convex function $\Phi : [0,1] \to \mathbb{R}$ that is strictly convex in some interval $(\ell, 1)$, with $\ell \in (0,1)$, there exists a constant $c > 0$ such that*

$$\sup_S G - G(X) \geqslant c \operatorname{dist}(X, V)^2 \tag{2.8}$$

for every $X \in S$.

This theorem is equivalent to Theorem 2.4 because $D[F]$ is the Euclidean distance from X to V, and the geodesic distance on S is equivalent to the Euclidean distance, when S is seen as a subset of $\mathbb{C}^d$. Indeed, if $X_1, X_2 \in S$,

$$|X_1 - X_2| = 2\sin(\operatorname{dist}(X_1, X_2)/2).$$

Remark 2.7. So far we have seen that there is a unitary correspondence between $\mathcal{H}_M$ and the space of polynomials $\mathcal{P}_M$, and now we have another unitary map $\mathcal{P}_M \to \mathbb{C}^d \supset S$. Due to these correspondences, the action of $SU(N+1)$ on $\mathcal{H}_M$ is automatically transferred on $\mathcal{P}_M$, $\mathbb{C}^d$ and S. Moreover, since the set of coherent states in $\mathcal{H}_M$ is an orbit of $SU(N+1)$ and since the coherent states are in one-to-one correspondence with the points of V (whether viewed as a subset of the normalized polynomials in $\mathcal{P}_M$ or as a subset of S), the isometric action of $SU(N+1)$ on S is, in fact, transitive on V. This will be of great importance in the following, as it will allow us to transfer any result proved for a single point in V to every other point in V.

We point out that this unitary correspondence between $\mathcal{H}_M$ and $\mathbb{C}^d$ can be made explicit considering a particular orthonormal basis of $\mathcal{H}_M$. In fact, denoting with $\{\iota_n\}_{n=0}^N$ the canonical basis of $\mathbb{C}^{N+1}$, it is easy to see that the vectors

$$\psi_\alpha = \sqrt{\binom{M}{\alpha}} P_M(\underbrace{\iota_0 \otimes \cdots \iota_0}_{M-|\alpha| \text{ times}} \otimes \underbrace{\iota_1 \otimes \cdots \iota_1}_{\alpha_1 \text{ times}} \otimes \cdots \otimes \underbrace{\iota_N \otimes \cdots \otimes \iota_N}_{\alpha_N \text{ times}}),$$

with $\alpha \in \mathbb{N}^N$ of length at most M, form an orthonormal basis of $\mathcal{H}_M$ and are such that

$$\langle \otimes^M(1, \overline{z}) | \psi_\alpha \rangle = e_\alpha(z).$$

Therefore, the unitary correspondence that we obtained through $\mathcal{P}_M$ is simply the map

$$\mathcal{H}_M \ni \psi = \sum_{|\alpha| \leqslant M} X_\alpha \psi_\alpha \mapsto (X_\alpha)_{|\alpha| \leqslant M} \in \mathbb{C}^d.$$

2.3 | Preliminaries of differential geometry

We conclude this section with some preliminaries of differential geometry. First of all, we notice that the subset V where G achieves its maximum is a smooth submanifold of S of dimension $2N + 1$ since it is the preimage of the Veronese variety under the canonical projection $\pi : S \to \mathbb{C}P^{d-1}$ (see [7, section 11.3] and [21, Remark 2.1(ii)]), which is smooth (see [11, p. 184]), and the above projection is a submersion (see [14, Theorem 6.30]). For more information on the connection between coherent states and rational curves, see, for example, [1].

For any point $X \in V \subset S$ we have the decomposition

$$T_X S = T_X V \oplus (T_X V)^\perp.$$

In the following, we are going to consider the ε-neighborhood of V, that is the set

$$U_\varepsilon = \{X \in S : \operatorname{dist}(X, V) < \varepsilon\},$$

where the distance is intended in the geodesic sense. Then, it is well known (see [15, Theorem 6.40]) that U_ε is also a ε-tubular neighborhood of V, which means that every point in U_ε is of the form $X(t)$ where $X : (-\varepsilon, \varepsilon) \to S$ is a geodesic with unit speed and such that $X(0) \in V$, $X'(0) \in (T_{X(0)}V)^\perp \subset T_{X(0)}S$. In other words, U_ε is the image of the exponential map restricted to the normal bundle of V.

Rather than at a generic point $X \in S$, we will perform the computations at $X_0 = (1, 0, \dots, 0)$, where the expressions of the tangent and normal spaces take the simple forms

$$T_{X_0} S = \{Y = (Y_\alpha)_{|\alpha| \leqslant M} \in \mathbb{C}^d : \operatorname{Re} Y_0 = 0\},$$

$$T_{X_0} V = \{Y = (Y_\alpha)_{|\alpha| \leqslant M} \in T_{X_0} S : Y_\alpha = 0 \text{ for } |\alpha| \geqslant 2\},$$

$$(T_{X_0} V)^\perp = \{Y = (Y_\alpha)_{|\alpha| \leqslant M} \in T_{X_0} S : Y_\alpha = 0 \text{ for } |\alpha| \leqslant 1\}.$$

To obtain the expression for $T_{X_0}V$, consider a smooth curve $v : (-\varepsilon, \varepsilon) \to S^{2N+1} \subset \mathbb{C}^{N+1}$ such that $v(0) = v_0 = (1, 0, \dots, 0) \in \mathbb{C}^{N+1}$, and note that $\otimes^M v(t)$ is a curve on the submanifold of coherent states passing through $\otimes^M v_0$, whose corresponding polynomials are given by

$$\langle \otimes^M(1, \overline{z}) | \otimes^M v(t) \rangle = \langle (1, \overline{z}) | v(t) \rangle^M = \sum_{|\alpha| \leqslant M} X_\alpha(t) e_\alpha(z), \quad z \in \mathbb{C}^N,$$

where $X(t)$ is a smooth curve in $V \subset S$ such that $X(0) = X_0$. Differentiation at $t = 0$ gives

$$M\left(v_0'(0)+\sum_{j=1}^{N} v_j'(0)z_j\right)=\sum_{|\alpha|\leqslant M} X_\alpha'(0)e_\alpha(z),$$

which proves the inclusion $T_{X_0}V \subseteq \{Y \in T_{X_0}S : Y_\alpha = 0 \text{ for } |\alpha| \geqslant 2\}$. The reverse inclusion follows by observing that the two spaces have the same dimension, since V is a submanifold of S of dimension $2N+1$.

3 | PROOF OF THEOREM 2.6 WHEN Φ IS OF CLASS C^2

To illustrate the main idea of the proof, in this section we prove Theorem 2.6 under the additional assumption that $\Phi \in C^2([0,1])$. In this case, also the Wehrl-type entropy function G in (2.7) is easily seen to be of class C^2, which simplifies the proof. The key points of the proof are:

- the value of G at points far from V is uniformly distant from the maximal value of G;
- the second differential of G at every point $X \in V$ is negative semidefinite on $T_X S$ and negative definite on the orthogonal space $(T_X V)^\perp$.

These ideas are formalized by the following lemmas, the first of which does not actually depend on the regularity of G and, in fact, will also be used in the proof for general Φ in Section 4. We also observe that the constant c in this lemma is explicit, in the sense that it is not obtained by a compactness argument.

Lemma 3.1. *Let $\Phi : [0,1] \to \mathbb{R}$ be a convex function that is strictly convex in some interval $(\ell, 1)$ with $\ell \in (0,1)$. Then, for every $\varepsilon > 0$ small enough there exists $c > 0$ such that, for every $X \in S$, if* $\operatorname{dist}(X,V) \geqslant \varepsilon$ *then*

$$\sup_S G - G(X) \geqslant c.$$

In particular, up to reducing the constant c, we have

$$\sup_S G - G(X) \geqslant c\operatorname{dist}(X,V)^2 \quad \textit{whenever} \quad \operatorname{dist}(X,V) \geqslant \varepsilon.$$

Proof. We use the notation of Section 2. Let $F = \sum_{|\alpha|\leqslant M} X_\alpha e_\alpha \in \mathcal{P}_M$ be the polynomial associated with $X = (X_\alpha) \in S$ and let

$$T := \sup_{z\in\mathbb{C}^N} \frac{|F(z)|^2}{(1+|z|^2)^M} \in [0,1].$$

We start by proving that if $\operatorname{dist}(X,V) \geqslant \varepsilon$, then $1 - T \geqslant \tau$ for some positive constant τ that depends only on ε. We have

$$\begin{aligned} D[F]^2 &= \inf_{\theta\in\mathbb{R}, w\in\mathbb{C}^N} \|F - e^{i\theta}k_w\|^2_{\mathcal{P}_M} = \inf_{\theta\in\mathbb{R}, w\in\mathbb{C}^N}\left(2 - 2\operatorname{Re} e^{-i\theta}\frac{F(w)}{(1+|w|^2)^{M/2}}\right) \\ &= 2 - 2\sup_{\theta\in\mathbb{R}, w\in\mathbb{C}^N}\operatorname{Re} e^{-i\theta}\frac{F(w)}{(1+|w|^2)^{M/2}} = 2 - 2\sup_{w\in\mathbb{C}^N}\frac{|F(w)|}{(1+|w|^2)^{M/2}} = 2 - 2\sqrt{T}. \end{aligned}$$

As observed in Section 2, $D[F]$ is just the Euclidean distance from X to V in $\mathbb{C}^d$, while for the geodesic distance in the unit sphere we have $\mathrm{dist}(X, V) \leqslant CD[F]$ for some positive constant C. Therefore,

$$\varepsilon \leqslant \mathrm{dist}(X, V) \leqslant CD[F] = C\sqrt{2 - 2\sqrt{T}},$$

and rearranging this expression we obtain that $1 - T \geqslant \tau$ for every ε small enough and for some $\tau > 0$ depending on ε.

Let F, X and T be as above, with $\mathrm{dist}(X, V) \geqslant \varepsilon$ so that $1 - T \geqslant \tau$, and consider the decomposition $\Phi = \Phi_1 + \Phi_2$, where

$$\Phi_2(t) = \begin{cases} \Phi(t), & t \in [0, T] \\ \Phi'_-(T)(t - T) + \Phi(T), & t \in (T, 1] \end{cases}$$

(Φ'_- denotes the left derivative of Φ). Denoting by G_1 and G_2 the corresponding Wehrl-type entropies, since both Φ_1 and Φ_2 are convex, we have

$$\sup_S G - G(F) = \sup_S G_1 - G_1(F) + \sup_S G_2 - G_2(F) \geqslant \sup_S G_1 - G_1(F) = \sup_S G_1,$$

where the first equality holds since G_1 and G_2 both achieve their maximum on V (by the Lieb–Solovej inequality), while the last equality follows from the fact that $\Phi_1(t) = 0$ for $t \leqslant T$.

To give an estimate from below of $\sup G_1$, we consider the distribution function of $1/(1 + |z|^2)^M$, that is,

$$\mu_0(t) = \nu(\{z \in \mathbb{C}^N \,:\, 1/(1 + |z|^2)^M > t\}).$$

Its explicit expression $\mu_0(t) := (1 - t^{1/M})^N$ for $t \in [0, 1]$ (see [21, eq. (4.3)]) is not needed here, where we just observe that $\mu_0(t) > 0$ for $t \in [0, 1)$.

Recalling the expression of the supremum of G_1 (2.6) and using the layer cake representation [17, Theorem 1.13], since $\lim_{t \to 0^+} \Phi_1(t) = 0$ we have

$$\begin{aligned} \sup G_1 &= \int_{\mathbb{C}^N} \Phi_1\left(\frac{1}{(1 + |z|^2)^M}\right) d\nu(z) \\ &= \int_0^1 \Phi_1'(t)\mu_0(t)\, dt \\ &= \int_T^1 (\Phi'(t) - \Phi'_-(T))\, \mu_0(t)\, dt \\ &\geqslant \int_{1-\tau}^1 (\Phi'(t) - \Phi'_-(1 - \tau))\, \mu_0(t)\, dt \\ &= c \geqslant \frac{c}{2} \cdot 2(1 - \sqrt{T}) = \frac{c}{2} D[F]^2 \geqslant c' \,\mathrm{dist}(X, V)^2, \end{aligned}$$

where $c, c' > 0$ because by assumption Φ is strictly convex in an interval of positive length contained in $(1 - \tau, 1)$. This concludes the proof. ∎

We now establish the desired stability estimate near V, for which the following lemma plays a crucial role (we alert the reader that some of the computations are deferred to the Appendix to maintain continuity of reading).

Lemma 3.2. *Let $\Phi : [0,1] \to \mathbb{R}$ be a convex function of class $C^2([0,1])$ that is strictly convex in some interval $(\ell, 1)$ with $\ell \in (0,1)$. Then, the second differential of G at $X_0 \in V$ is negative semidefinite on $T_{X_0}S$ and negative definite on the orthogonal space $(T_{X_0}V)^\perp$, that is,*

- $\mathrm{d}^2_{X_0} G(Y) \leqslant 0$ *for every* $Y \in T_{X_0}S$;
- $\mathrm{d}^2_{X_0} G(Y) \leqslant -c|Y|^2$ *for every* $Y \in (T_{X_0}V)^\perp$,

where $c > 0$.

Proof. By $SU(N+1)$ invariance (see Remark 2.7), it suffices to consider the point $X_0 = (1, 0, \dots, 0) \in V$. Consider a smooth curve $X : (-\varepsilon, \varepsilon) \to S$ such that $X(0) = X_0$ and let $Y = (Y_\alpha)_{|\alpha| \leqslant M} = X'(0) \in T_{X_0}S$. Since $X(0) = (1, 0, \dots, 0)$ and $\mathrm{Re} Y_0 = 0$ (see Section 2.3),

$$\mathrm{d}^2_{X_0} G(Y) := \frac{d^2}{dt^2} G(X(t))|_{t=0} = \int_{\mathbb{C}^N} \Phi''\left(\frac{1}{(1+|z|^2)^M}\right) \frac{(2\mathrm{Re}\sum_{\alpha\neq 0} Y_\alpha e_\alpha(z))^2}{(1+|z|^2)^{2M}} d\nu(z)$$
$$+ \int_{\mathbb{C}^N} \Phi'\left(\frac{1}{(1+|z|^2)^M}\right) \frac{2\mathrm{Re}\sum_{|\alpha|\leqslant M} X''_\alpha(0) e_\alpha(z) + 2|\sum_{|\alpha|\leqslant M} Y_\alpha e_\alpha(z)|^2}{(1+|z|^2)^M} d\nu(z).$$

We point out that the term $\mathrm{Re}\sum_{|\alpha|\leqslant M} X''_\alpha(0) e_\alpha(z)$ can be replaced with just $\mathrm{Re} X''_0(0)$ (the subscript 0 stands for the multi-index in $\mathbb{N}^N$ with 0 in all components). For, passing to polar coordinates and integrating in the angular component first, one sees that the integrals of $\mathrm{Re}\, e_\alpha(z)$ vanish for $\alpha \neq 0$, since $e_\alpha(0) = 0$ and $e_\alpha(z)$ (being the real part of a holomorphic function) is harmonic. We also observe that

$$\mathrm{Re} X''_0(0) = -\sum_{|\alpha|\leqslant M} |Y_\alpha|^2,$$

as can be seen by differentiating twice the identity $|X(t)|^2 = 1$ and evaluating at $t = 0$. Concerning the other terms, we have

$$(2\mathrm{Re}\sum_{\alpha\neq 0} Y_\alpha e_\alpha(z))^2 = 2\sum_{\alpha\neq 0} |Y_\alpha|^2 |e_\alpha(z)|^2 + 2\mathrm{Re}\left[\sum_{\alpha,\beta\neq 0} Y_\alpha Y_\beta e_\alpha(z) e_\beta(z)\right.$$
$$\left. + \sum_{\alpha\neq\beta;\alpha,\beta\neq 0} Y_\alpha \overline{Y_\beta} e_\alpha(z) \overline{e_\beta(z)}\right].$$

Again, using polar coordinates one sees that all the terms on the right-hand side, with the exception of the first, give a zero contribution in the integral (the term $\mathrm{Re}\, Y_\alpha \overline{Y_\beta} e_\alpha(z) \overline{e_\beta(z)}$ is not harmonic in general, yet its integral on every shell is zero since $\alpha \neq \beta$). Similarly,

$$\left|\sum_{|\alpha|\leqslant M} Y_\alpha e_\alpha(z)\right|^2 = \sum_{|\alpha|\leqslant M} |Y_\alpha|^2 |e_\alpha(z)|^2 + \sum_{\alpha\neq\beta} Y_\alpha \overline{Y_\beta} e_\alpha(z) \overline{e_\beta(z)}$$

and, again, the integral of the last sum is zero.

Thus, since $|e_0(z)|^2 = 1$, we are left with

$$\frac{1}{2}\mathrm{d}^2_{X_0} G(Y) = \sum_{\alpha\neq 0} b_\alpha |Y_\alpha|^2, \tag{3.1}$$

where

$$b_\alpha = \int_{\mathbb{C}^N} \Phi''\left(\frac{1}{(1+|z|^2)^M}\right)\frac{|e_\alpha(z)|^2}{(1+|z|^2)^{2M}}\, d\nu(z) + \int_{\mathbb{C}^N} \Phi'\left(\frac{1}{(1+|z|^2)^M}\right)\frac{|e_\alpha(z)|^2 - 1}{(1+|z|^2)^M}\, d\nu(z).$$

So, to prove the lemma, we just need to prove that every coefficient b_α is nonpositive for $|\alpha| = 1$ and strictly negative for $2 \leqslant |\alpha| \leqslant M$. To do so, we pass to the polar coordinates $z = \rho\omega$, where $\rho \geqslant 0$ and $\omega \in S^{2N-1} \subset \mathbb{C}^N$ (unit sphere in $\mathbb{C}^N$), so that

$$d\nu = \frac{c_N \rho^{2N-1}}{(1+\rho^2)^{N+1}}\, d\rho\, d\omega,$$

where c_N is the normalization constant for the measure $d\nu$ in (2.5). Observe that $1/c_N$ is the volume of the unit ball in $\mathbb{C}^N$ (i.e. the real $2N$-ball), therefore $c_N/(2N)d\omega$ is the normalized surface measure on S^{2N-1}. For ease of notation, we let

$$\tilde{c}^2_\alpha := c^2_\alpha \int_{S^{2N-1}} |\omega^\alpha|^2 \frac{c_N}{2N}\, d\omega, \tag{3.2}$$

whose explicit expression can be found below in Lemma A.2.

Hence, passing to polar coordinates in the integrals defining b_α, after the further change of variable $\rho^2 = s$ we have to study the sign of

$$N\int_0^\infty \Phi''\left(\frac{1}{(1+s)^M}\right)\frac{\tilde{c}^2_\alpha s^{|\alpha|+N-1}}{(1+s)^{2M+N+1}} ds + N\int_0^\infty \Phi'\left(\frac{1}{(1+s)^M}\right)\frac{\tilde{c}^2_\alpha s^{|\alpha|+N-1} - s^{N-1}}{(1+s)^{M+N+1}}\, ds.$$

In view of integrating by parts the second term, using Lemma A.1, we can preliminarily compute

$$\begin{aligned}\int_0^s \frac{\tilde{c}^2_\alpha \sigma^{|\alpha|+N-1} - \sigma^{N-1}}{(1+\sigma)^{M+N+1}}\, d\sigma &= \frac{\tilde{c}^2_\alpha (1+s)^{-(M+N)}}{A_{M,N,|\alpha|}} \sum_{j=|\alpha|+N}^{M+N} \binom{M+N}{j} s^j \\ &\quad - \frac{(1+s)^{-(M+N)}}{A_{M,N,0}} \sum_{j=N}^{M+N} \binom{M+N}{j} s^j \\ &= -\frac{\tilde{c}^2_\alpha (1+s)^{-(M+N)}}{A_{M,N,|\alpha|}} \sum_{j=N}^{|\alpha|+N-1} \binom{M+N}{j} s^j,\end{aligned}$$

where in the last step we used Lemma A.2 (see below for the explicit expression of $A_{M,N,|\alpha|}$). From this and from the fact that $\Phi'(t)$ has finite limits at 0^+ and at 1^-, it is immediate to see that the boundary term in the integration by parts vanishes and therefore we end up with

$$b_\alpha = N \int_0^\infty \Phi''\left(\frac{1}{(1+s)^M}\right) \frac{\tilde{c}_\alpha^2}{(1+s)^{2M+N+1}} \left[s^{|\alpha|+N-1} - \frac{M}{A_{M,N,|\alpha|}} \sum_{j=N}^{|\alpha|+N-1} \binom{M+N}{j} s^j \right] ds. \tag{3.3}$$

Now, we have

$$\frac{M}{A_{M,N,1}} \binom{M+N}{N} = 1,$$

so for $|\alpha| = 1$ the integral is 0 (which is clear since $Y \in T_{X_0}V$ and G is constant on V), while for $|\alpha| > 1$, we have

$$\frac{M}{A_{M,N,|\alpha|}} \binom{M+N}{|\alpha|+N-1} = \frac{M}{M-|\alpha|+1} > 1,$$

that is, the coefficient of the monomial of degree $|\alpha| + N - 1$ is strictly negative. Hence, the expression in square brackets is strictly negative for every $s > 0$. Finally, by assumption, $\Phi'' \geqslant 0$ and Φ is strictly convex in an interval of the type $(\ell, 1)$, for some $\ell \in (0,1)$, and the desired conclusion follows. □

Proof of Theorem 2.6 for Φ of class C^2. Assume, in addition to the hypothesis in the statement, that $\Phi \in C^2([0,1])$. Let $\varepsilon > 0$ to be fixed later. As already recalled in Section 2.3, every point in the ε-neighborhood of V can be reached by a geodesic $X : (-\varepsilon, \varepsilon) \to S$ parameterized with unit speed such that $X(0) = X_0 \in V$, with an initial velocity $Y = X'(0) \in (T_{X_0}V)^\perp$ such that $|Y| = 1$. In this case, for ε sufficiently small and $t \in (-\varepsilon, \varepsilon)$ we have, for some $\xi \in [-|t|, |t|]$ and $c > 0$,

$$\begin{aligned} \sup_S G - G(X(t)) &= G(X(0)) - G(X(t)) \\ &= -\frac{1}{2}(G \circ X)''(\xi) t^2 \\ &\geqslant ct^2 = c\, \mathrm{dist}(X(t), V)^2, \end{aligned}$$

where we applied Lemma 3.2 and the fact that the function $(G \circ X)''(t)$ is continuous, uniformly with respect to $Y \in (T_{X_0}V)^\perp$, with $|Y| = 1$. This fixes the value of ε and gives the desired estimate where $\mathrm{dist}(X, V) < \varepsilon$. Then, for this value of ε we apply Lemma 3.1 and obtain the estimate also where $\mathrm{dist}(X, V) \geqslant \varepsilon$. □

Remark 3.3. Wehrl's original entropy is defined with $\Phi(t) = t \log t$, which is not in $C^2([0,1])$. Nevertheless, one can easily adapt the above proof to this case, using the decomposition $\Phi = \Phi_1 + \Phi_2$ where

$$\Phi_1(t) = \begin{cases} \frac{1}{2}\Phi''(a)(t-a)^2 + \Phi'(a)(t-a) + \Phi(a), & t \in [0,a] \\ \Phi(t), & t \in (a,1] \end{cases},$$

for some $a \in (0,1)$. Indeed, since $\Phi_2 := \Phi - \Phi_1$ is convex, reasoning as in the proof of Lemma 3.1 one sees that it is enough to prove the desired estimate for the Wehrl-type entropy associated with Φ_1, which belongs to $C^2([0,1])$ and satisfies the assumptions of Theorem 2.6.

4 | PROOF OF THEOREM 2.6 FOR AN ARBITRARY CONVEX FUNCTION

When Φ is not C^2 the previous argument breaks down and some additional work is required. In fact, we will approximate Φ by smooth functions and compute the second differential of the smoothed G's at points that are not in V, in which case the computation is more involved. The proof of the following lemma is postponed to Appendix.

Lemma 4.1. *Let $X_0 = (1, 0, \dots, 0) \in \mathbb{C}^d$. For every $0 < a < b < 1$ there exist $\varepsilon > 0$ and a continuous function $h(\tau, t, Y)$, defined for $\tau \in (a, b)$, $|t| < \varepsilon$ and $Y \in (T_{X_0}V)^\perp$ with $|Y| = 1$, such that the following representation formula holds.*

If $X(t,Y)$ denotes the unit-speed geodesic in $S \subset \mathbb{C}^d$, with $X(0,Y) = X_0$ and initial velocity $\partial_t X(0,Y) = Y$, for every convex function $\Phi \in C^2([0,1])$ such that Φ'' is supported in (a,b) it holds

$$\partial^2_{tt} G(X(t,Y)) = \int_0^1 \Phi''(\tau) h(\tau, t, Y)\, d\tau, \quad |t| < \varepsilon,\ Y \in (T_{X_0}V)^\perp,\ |Y| = 1. \tag{4.1}$$

Moreover, $h(\tau, 0, Y) < 0$ for every $\tau \in (a,b)$ and $Y \in (T_{X_0}V)^\perp$ with $|Y| = 1$.

Proof of Theorem 2.6 in full generality. Let $\varepsilon > 0$ be sufficiently small, to be fixed later. In view of Lemma 3.1 it is sufficient to prove the estimate

$$\sup_S G - G(X) \geqslant c\, \mathrm{dist}(X, V)^2 \qquad \text{if } \mathrm{dist}(X, V) < \varepsilon. \tag{4.2}$$

To this end, let $0 < a < b < 1$ and decompose $\Phi = \Phi_1 + \Phi_2$, where Φ_1 is the affine continuation of $\Phi|_{[a,b]}$, that is,

$$\Phi_1(t) = \begin{cases} \Phi'_+(a)(t-a) + \Phi(a), & t \in [0,a) \\ \Phi(t), & t \in [a,b] \\ \Phi'_-(b)(t-b) + \Phi(b), & t \in (b,1]. \end{cases}$$

Then, since both Φ_1 and Φ_2 are convex, denoting with G_1 and G_2 the corresponding entropies, we have

$$\sup_S G - G(X) = \sup_S G_1 - G_1(X) + \sup_S G_2 - G_2(X) \geqslant \sup_S G_1 - G_1(X),$$

where in the first equality we used the fact that both G_1 and G_2 achieve their supremum at every point of V (by Lieb–Solovej inequality). Therefore, it suffices to prove the estimate (4.2) for Φ_1. Also, by the assumption on Φ, if b is sufficiently close to 1 (in particular, if $b > \ell$), Φ_1 will be strictly convex in some small interval contained in $[a, b]$.

Hence, we can assume that $\Phi : [0, 1] \to \mathbb{R}$ is convex, affine in $[0, a]$ and in $[b, 1]$, with Φ'', regarded as a positive Radon measure in (0,1), nonzero and supported in $[a, b]$.

Again we recall that every point of the ε-neighborhood of V in S is of the form $X(t)$, where $X : (-\varepsilon, \varepsilon) \to S$ is a geodesic, parameterized with unit speed, starting from a point $X_0 \in V$ and such that $Y = X'(0) \in (T_{X_0}V)^{\perp}$. With this in mind, the inequality (4.2) can be rephrased as

$$G(X(0)) - G(X(t)) \geqslant ct^2 \quad \forall t \in (-\varepsilon, \varepsilon) \tag{4.3}$$

for some $c > 0$, since $|t|$ is exactly the geodesic distance from $X(t)$ to V.

We now approximate Φ with more regular convex functions. First, we extend Φ to a function defined on $\mathbb{R}$, affine on $(-\infty, a]$ and on $[b, +\infty)$. Then, by convolution with a standard mollifier (see [4, section 6.3]) we construct a sequence of convex functions $\Phi_n : [0, 1] \to \mathbb{R}$ such that, possibly for a slightly larger interval $[a, b] \subset (0, 1)$,

- $\Phi_n \in C^2([0, 1])$;
- $\Phi_n \to \Phi$ uniformly in [0,1];
- Φ_n'' has compact support in (a, b);
- $\Phi_n'' \to \Phi''$ as (positive, finite) Radon measures on (0,1), that is, in the distribution sense.

We denote with G_n the corresponding Wehrl-type entropies. Using dominated convergence we have

$$G(X(0)) - G(X(t)) = \lim_{n\to\infty} (G_n(X(0)) - G_n(X(t)))$$

and applying Taylor's formula to the G_ns we see that

$$G_n(X(0)) - G_n(X(t)) = -\frac{1}{2}(G_n \circ X)''(\xi_n)t^2$$

for some ξ_n such that $|\xi_n| \leqslant |t|$. Therefore, to prove the stability estimate (4.3), we just need to prove that $(G_n \circ X)''(\xi) < -c < 0$ for some constant c independent of n and for every $\xi \in (-\varepsilon, \varepsilon)$. By $SU(N + 1)$ invariance (see Remark 2.7) it is sufficient to prove this when $X(0) = X_0 = (1, 0, \dots, 0) \in \mathbb{C}^d$.

From Lemma 4.1 (applied to a slightly larger interval (a, b)) we have

$$(G_n \circ X)''(\xi) = \int_0^1 \Phi_n''(\tau)h(\tau, \xi, Y)d\tau,$$

(recall, $Y = X'(0)$) where $h(\tau, 0, Y) < -c < 0$ for $\tau \in [a, b]$ and $Y \in (T_{X_0}V)^{\perp}$ with $|Y| = 1$. However, since h is continuous, possibly with a smaller constant c we have, for ε sufficiently small, that $h(\tau, \xi, Y) < -c < 0$ for every $\tau \in (a, b)$, $\xi \in (-\varepsilon, \varepsilon)$ and $Y \in (T_{X_0}V)^{\perp}$ with $|Y| = 1$. Hence, for $t \in (-\varepsilon, \varepsilon)$, we have

$$\begin{aligned} G(X(0)) - G(X(t)) &= \lim_{n\to\infty} G_n(X(0)) - G_n(X(t)) \\ &= \lim_{n\to\infty} -\frac{1}{2}(G_n \circ X)''(\xi_n)t^2 \\ &= \lim_{n\to\infty} -\frac{t^2}{2}\int_a^b \Phi_n''(\tau)h(\tau,\xi_n,Y)d\tau, \end{aligned}$$

and therefore

$$\begin{aligned} G(X(0)) - G(X(t)) &\geqslant ct^2 \lim_{n\to\infty}\int_a^b \Phi_n''(\tau)d\tau \\ &= ct^2 \lim_{n\to\infty}\int_0^1 \Phi_n''(\tau)d\tau \\ &= c\Phi''((0,1))t^2 = c't^2 \end{aligned}$$

where $\Phi''((0,1)) > 0$ denotes the total mass of the measure Φ''. The stability estimate (4.3) is therefore proved and this concludes the proof. □

APPENDIX
TECHNICAL LEMMAS

In this appendix, we have gathered several technical lemmas.

Lemma A.1. *Let $N \geqslant 1$, $M \geqslant 1$ and $0 \leqslant K \leqslant M$. Then, for $s \geqslant 0$,*

$$\int_0^s \frac{\sigma^{K+N-1}}{(1+\sigma)^{M+N+1}}\,d\sigma = \frac{(1+s)^{-(M+N)}}{A_{M,N,K}}\sum_{j=K+N}^{M+N}\binom{M+N}{j}s^j \tag{A.1}$$

where

$$A_{M,N,K} = (M-K+1)\binom{M+N}{K+N-1}. \tag{A.2}$$

Proof. After changing the variable $\tau = \sigma/(1+\sigma)$ one recognizes the expression of the incomplete beta function (up to the normalization constant), for which the expression (A.1) has been known for a long time (see [3, Formula 2.5]). □

Lemma A.2. *The coefficients $\tilde{c}_\alpha^2$ defined in (3.2) are given by*

$$\tilde{c}_\alpha^2 = \frac{(N-1)!M!}{(N+|\alpha|-1)!(M-|\alpha|)!} = \frac{A_{M,N,|\alpha|}}{A_{M,N,0}},$$

where the numbers $A_{M,N,|\alpha|}$ are defined in (A.2).

Proof. Recalling the expression of the norm (2.3) in $\mathcal{P}_M$ we have

$$
\begin{aligned}
1 &= d\int_{\mathbb{C}^N} \frac{c_\alpha^2 |z^\alpha|^2}{(1+|z|^2)^{M+N+1}} c_N \, dA(z) \\
&= dN \int_{S^{2N-1}} c_\alpha^2 |\omega^\alpha|^2 \frac{c_N}{2N} d\omega \int_0^\infty \frac{s^{|\alpha|+N-1}}{(1+s)^{M+N+1}} ds \\
&= \frac{dN}{A_{M,N,|\alpha|}} \tilde{c}_\alpha^2,
\end{aligned}
$$

where in the last equality we used (A.1). Then, recalling the explicit expression of d (2.4) it is immediately seen that $dN = A_{M,N,0}$. □

Proof of Lemma 4.1. For simplicity of notation, let us set $X(t) = X(t, Y)$. Similarly, we will omit the dependence on Y of various functions below.

Let $\varepsilon > 0$, to be chosen later on. Let Φ be as in the statement, in particular with Φ'' compactly supported in (a, b). We have to compute the second derivative of

$$
(G \circ X)(t) = \int_{\mathbb{C}^N} \Phi(u(z,t)) \, d\nu(z)
$$

for $|t| < \varepsilon$, where

$$
u(z,t) = u(z,t,Y) = \frac{|\sum_{|\alpha|\leqslant M} X_\alpha(t) e_\alpha(z)|^2}{(1+|z|^2)^M}
$$

(recall that $X_\alpha(t) = X_\alpha(t, Y)$ will also depend smoothly on Y).

Since $\int_{\mathbb{C}^N} u(z,t) d\nu(z) = 1/d$ for all t and Y, we can suppose, without loss of generality, that $\Phi(0) = \Phi'(0) = 0$. Therefore, we have

$$
\Phi(v) = \int_0^v \int_0^\sigma \Phi''(\tau) \, d\tau \, d\sigma,
$$

which allows us to rewrite

$$
(G \circ X)(t) = \int_{\mathbb{C}^N} \int_0^{u(z,t)} \int_0^\sigma \Phi''(\tau) \, d\tau \, d\sigma \, d\nu(z).
$$

Differentiating twice with respect to t leads to

$$
\begin{aligned}
(G \circ X)''(t) &= \int_{\mathbb{C}^N} \Phi''(u(z,t)) (\partial_t u(z,t))^2 \, d\nu(z) \\
&\quad + \int_{\mathbb{C}^N} \left(\int_0^{u(z,t)} \Phi''(\tau) \, d\tau \right) \partial_{tt}^2 u(z,t) \, d\nu(z) \\
&= \int_{\mathbb{C}^N} \Phi''(u(z,t)) (\partial_t u(z,t))^2 \, d\nu(z) \\
&\quad + \int_0^1 \Phi''(\tau) \left(\int_{\{u(\cdot,t)>\tau\}} \partial_{tt}^2 u(z,t) \, d\nu(z) \right) d\tau.
\end{aligned}
$$

Our goal is to rewrite the first integral using the coarea formula in order to obtain the representation formula (4.1). Before doing so, we notice that there exist $\varepsilon > 0$ and a compact subset $K \subset \mathbb{C}^N \setminus \{0\}$ such that, for all $|t| < \varepsilon$ and $Y \in (T_{X_0}V)^{\perp}$, with $|Y| = 1$,

$$a \leqslant u(z,t) \leqslant b \Rightarrow z \in K.$$

Indeed, since for $t = 0$ we have $u(z,0) = (1+|z|^2)^{-M}$ and since $X(t)$ has unit speed, we have $|u(z,t) - (1+|z|^2)^{-M}| \leqslant C\varepsilon$ for $|t| < \varepsilon$ and $Y \in (T_{X_0}V)^{\perp}$ with $|Y| = 1$, if ε is small enough, for some constant $C > 0$. Hence, one can take

$$K = \{z \in \mathbb{C}^N \ : \ a - C\varepsilon \leqslant (1+|z|^2)^{-M} \leqslant b + C\varepsilon\}.$$

Moreover, since $\nabla u(z,0) \neq 0$ for $z \in K$, up to reducing ε if necessary, we can suppose that $|\nabla u(z,t)| > c > 0$ for every $z \in K$, $|t| < \varepsilon$ and $Y \in (T_{X_0}V)^{\perp}$, with $|Y| = 1$. This allows us to use the coarea formula [4, Theorem 3.13(ii)] and obtain

$$(G \circ X)''(t) = \int_0^1 \Phi''(\tau) h(\tau,t)\, d\tau, \tag{A.3}$$

where (cf. (2.5))

$$h(\tau,t) = h(\tau,t,Y) = \int_{\{u(\cdot,t)=\tau\}} \frac{c_N(\partial_t u(z,t))^2}{|\nabla u(z,t)|^2 (1+|z|^2)^{N+1}}\, d\mathcal{H}^{2N-1}(z) + \int_{\{u(\cdot,t)>\tau\}} \partial^2_{tt} u(z,t)\, d\nu(z)$$

which is the desired representation—here $d\mathcal{H}^{2N-1}$ denotes the $(2N-1)$-dimensional Hausdorff measure. Moreover, since for $\tau \in (a,b)$, $|t| < \varepsilon$ and $Y \in (T_{X_0}V)^{\perp}$, with $|Y| = 1$, the equality $u(z,t) = \tau$ implies $z \in K$ and $|\nabla u(z,t)| > c > 0$, it is easy to see that h is continuous, as claimed.

To conclude the proof, we need to show that $h(\tau,0)$ is strictly negative for every $\tau \in (a,b)$. We observe that setting $t = 0$ in (A.3) we obtain

$$(G \circ X)''(0) = \int_0^1 \Phi''(\tau) h(\tau,0)\, d\tau.$$

However, we have already computed $(G \circ X)''(0)$ in the proof of Lemma 3.2. In fact, combining Equations (3.1) and (3.3), with $Y = (Y_\alpha)_{|\alpha| \leqslant M} = X'(0)$ and with the change of variable $\tau = (1+s)^{-M}$ we have obtained that

$$(G \circ X)''(0) = \int_0^1 \Phi''(\tau) \tilde{h}(\tau) d\tau,$$

for a continuous function $\tilde{h}(\tau) = \tilde{h}(\tau, Y)$[†], which was found to be strictly negative for $\tau \in (0,1)$ whenever $Y \in (T_{X_0}V)^{\perp}$. Therefore, we just need to prove that $\tilde{h}(\tau) = h(\tau,0)$ for every $\tau \in (a,b)$.

[†] Specifically, with

$$\tilde{h}(\tau, Y) = \sum_{\alpha \neq 0} 2|Y_\alpha|^2 N \tilde{c}_\alpha^2 \tau^{1+\frac{N}{M}} \left[\frac{(\tau^{-1/M} - 1)^{|\alpha|+N-1}}{M} - \frac{1}{A_{M,N,|\alpha|}} \sum_{j=N}^{|\alpha|+N-1} \binom{M+N}{j} (\tau^{-1/M} - 1)^j \right].$$

This is clear, because the equality

$$\int_0^1 \Phi''(\tau)\tilde{h}(\tau)\,d\tau = \int_0^1 \Phi''(\tau)h(\tau,0)\,d\tau$$

holds for every convex function $\Phi \in C^2([0,1])$ whose second derivative is compactly supported in (a,b) and, therefore, holds for every positive Radon measure in (a,b), which implies $h(\tau,0) = \tilde{h}(\tau)$ for $\tau \in (a,b)$ and $Y \in (T_{X_0}V)^{\perp}$, with $|Y| = 1$. $\square$

ACKNOWLEDGMENTS

F. N. is a Fellow of the Accademia delle Scienze di Torino and a member of the Società Italiana di Scienze e Tecnologie Quantistiche (SISTEQ).

Open access publishing facilitated by Politecnico di Torino, as part of the Wiley - CRUI-CARE agreement.

JOURNAL INFORMATION

The *Proceedings of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

ORCID

Federico Riccardi https://orcid.org/0009-0005-8534-7795

REFERENCES

1. D. C. Brody and E.-M. Graefe, *Coherent states and rational surfaces*, J. Phys. A **43** (2010), no. 25, 255205, 14.
2. E. A. Carlen, *Some integral identities and inequalities for entire functions and their application to the coherent state transform*, J. Funct. Anal. **97** (1991), no. 1, 231–249.
3. J. Dutka, *The incomplete beta function—a historical profile*, Arch. Hist. Exact Sci. **24** (1981), no. 1, 11–29.
4. L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Textbooks in Mathematics, revised edition, CRC Press, Boca Raton, FL, 2015.
5. R. L. Frank, *Sharp inequalities for coherent states and their optimizers*, Adv. Nonlinear Stud. **23** (2023), no. 1, Paper No. 20220050, 28.
6. R. L. Frank, F. Nicola, and P. Tilli, *The generalized Wehrl entropy bound in quantitative form*, J. Eur. Math. Soc. (2025), published online first.
7. W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.
8. M. A. García-Ferrero and J. Ortega-Cerdà, *Stability of the concentration inequality on polynomials*, Comm. Math. Phys. **406** (2025), no. 5, Paper No. 112, 35.
9. J. Gómez, A. Guerra, J. P. G. Ramos, and P. Tilli, *Stability of the Faber-Krahn inequality for the short-time Fourier transform*, Invent. Math. **236** (2024), no. 2, 779–836.
10. J. Gómez, D. Kalaj, P. Melentijević, and J. P. G. Ramos, *Uniform stability of concentration inequalities and applications*, Proc. Lond. Math. Soc. (3) **131** (2025), no. 6, Paper No. e70114, 53.
11. J. Harris, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1992. A first course.
12. A. Kulikov, *Functionals with extrema at reproducing kernels*, Geom. Funct. Anal. **32** (2022), no. 4, 938–949.
13. A. Kulikov, F. Nicola, J. Ortega-Cerdà, and P. Tilli, *A monotonicity theorem for subharmonic functions on manifolds*, Adv. Math. **479** (2025), Paper No. 110423, 18.

14. J. M. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics, vol. 218, 2nd ed., Springer, New York, 2013.
15. J. M. Lee, *Introduction to Riemannian manifolds*, Graduate Texts in Mathematics, vol. 176, 2nd ed., Springer, Cham, 2018.
16. E. H. Lieb, *Proof of an entropy conjecture of Wehrl*, Comm. Math. Phys. **62** (1978), no. 1, 35–41.
17. E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics, vol. 14, 2nd ed., American Mathematical Society, Providence, RI, 2001.
18. E. H. Lieb and J. P. Solovej, *Proof of an entropy conjecture for Bloch coherent spin states and its generalizations*, Acta Math. **212** (2014), no. 2, 379–398.
19. E. H. Lieb and J. P. Solovej, *Proof of the Wehrl-type entropy conjecture for symmetric $SU(N)$ coherent states*, Comm. Math. Phys. **348** (2016), no. 2, 567–578.
20. P. Melentijević, *Stability of convex functionals on weighted Bergman spaces with applications*, arXiv:2508.01939, 2025.
21. F. Nicola, F. Riccardi, and P. Tilli, *The Wehrl-type entropy conjecture for symmetric $SU(N)$ coherent states: cases of equality and stability*, Duke Math. J. to appear (arXiv:2412.10940).
22. F. Nicola and P. Tilli, *The Faber-Krahn inequality for the short-time Fourier transform*, Invent. Math. **230** (2022), no. 1, 1–30.
23. A. Wehrl, *On the relation between classical and quantum-mechanical entropy*, Rep. Math. Phys. **16** (1979), no. 3, 353–358.