

Hamilton-Jacobi approach to potential functions in information geometry

Cite as: J. Math. Phys. 58, 063506 (2017); <https://doi.org/10.1063/1.4984941>

Submitted: 31 August 2016 . Accepted: 22 May 2017 . Published Online: 08 June 2017

Florio M. Ciaglia , Fabio Di Cosmo , Domenico Felice, Stefano Mancini, Giuseppe Marmo, and Juan M. Pérez-Pardo



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Pedal coordinates, dark Kepler, and other force problems](#)

Journal of Mathematical Physics **58**, 063505 (2017); <https://doi.org/10.1063/1.4984905>

[Dirac oscillator and minimal length](#)

Journal of Mathematical Physics **58**, 063504 (2017); <https://doi.org/10.1063/1.4984312>

[On a class of integrable systems of Monge-Ampère type](#)

Journal of Mathematical Physics **58**, 063508 (2017); <https://doi.org/10.1063/1.4984982>



Don't let your writing keep you from getting published!

AIP | Author Services

Learn more today!



Hamilton-Jacobi approach to potential functions in information geometry

Florio M. Ciaglia,¹ Fabio Di Cosmo,¹ Domenico Felice,² Stefano Mancini,² Giuseppe Marmo,¹ and Juan M. Pérez-Pardo³

¹*Dipartimento di Fisica, Università di Napoli "Federico II," Via Cinthia Edificio 6, I-80126 Napoli, Italy and INFN-Sezione di Napoli, Via Cinthia Edificio 6, I-80126 Napoli, Italy*

²*School of Science and Technology, University of Camerino Via Madonna delle Carceri 7, I-62032 Camerino, Italy and INFN-Sezione di Perugia, Via A. Pascoli, I-06123 Perugia, Italy*

³*INFN-Sezione di Napoli, Via Cinthia Edificio 6, I-80126 Napoli, Italy and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. de la Universidad 30, Leganés, Madrid, Spain*

(Received 31 August 2016; accepted 22 May 2017; published online 8 June 2017)

The search for a potential function S allowing us to reconstruct a given metric tensor g and a given symmetric covariant tensor T on a manifold \mathcal{M} is formulated as the Hamilton-Jacobi problem associated with a canonically defined Lagrangian on $T\mathcal{M}$. The connection between this problem, the geometric structure of the space of pure states of quantum mechanics, and the theory of contrast functions of classical information geometry are outlined. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4984941>]

I. INTRODUCTION

The recent development of quantum information theory has led to a growing interest in the geometrical description of the space of quantum states (Ref. 1). In this direction, the geometrical approach to quantum mechanics developed, for example, in Refs. 2 and 3, allows us to reformulate the information encoded in the scalar product $\langle \cdot, \cdot \rangle$ of the Hilbert space \mathcal{H} of the system in terms of two tensors on the space of pure states $\mathcal{P}(\mathcal{H})$, namely, a metric tensor g called the Fubini-Study metric, and a symplectic form ω . The manifold $\mathcal{P}(\mathcal{H})$ together with these tensors forms what is known as a Kähler manifold, i.e., $\mathcal{P}(\mathcal{H})$ admits a complex structure J such that the metric tensor g and the symplectic form ω are mutually related according to the following compatibility condition:

$$g(X, J(Y)) = \omega(X, Y), \quad (1)$$

where X and Y are the arbitrary vector fields on $\mathcal{P}(\mathcal{H})$. These two tensors together define a Hermitian tensor $h = g + i\omega$ on $\mathcal{P}(\mathcal{H})$ (Refs. 2–4). The geometric information of a Kähler manifold is completely encoded in the so-called (local) Kähler potential K . This is a (local) function on the manifold which allows us to recover the explicit expression of h , and thus of g and ω , as follows:

$$h_{j\bar{k}} = \frac{\partial^2 K}{\partial z^j \partial \bar{z}^k}, \quad (2)$$

where z^j, \bar{z}^k are the holomorphic coordinates on the Kähler manifold. As it is clear, the way in which h is extracted from the Kähler potential K highly depends on the complex structure J on $\mathcal{P}(\mathcal{H})$ by means of the derivatives with respect to the holomorphic coordinates.

Following Ref. 3, the pullback of h to the Hilbert space \mathcal{H} reads

$$\tilde{h} = \pi^* h = \frac{\langle d\psi | d\psi \rangle}{\langle \psi | \psi \rangle} - \frac{\langle d\psi | \psi \rangle \langle \psi | d\psi \rangle}{\langle \psi | \psi \rangle^2}, \quad (3)$$

where the projection map $\pi : \mathcal{H} \rightarrow \mathcal{P}(\mathcal{H})$ is given by

$$\pi(|\psi\rangle) = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}. \quad (4)$$

To get a more concrete feeling of \tilde{h} , let us consider a finite-dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^n$ and write a vector in \mathcal{H} as $|\psi\rangle = \sqrt{p_j} e^{i\varphi_j} |e_j\rangle$, where (p_1, \dots, p_n) is a probability vector, that is, $\sum p_j = 1$, $e^{i\varphi_j}$ is a phase factor, and the summation on j is understood. Then we have

$$\begin{aligned} \tilde{h} = & \frac{1}{4} \left[\langle d(\ln \vec{p}) \otimes d(\ln \vec{p}) \rangle_{\vec{p}} - \langle d(\ln \vec{p}) \rangle_{\vec{p}} \otimes \langle d(\ln \vec{p}) \rangle_{\vec{p}} \right] + \\ & + \langle d\vec{\varphi} \otimes d\vec{\varphi} \rangle_{\vec{p}} - \langle d\vec{\varphi} \rangle_{\vec{p}} \otimes \langle d\vec{\varphi} \rangle_{\vec{p}} + \frac{i}{2} \left[\langle d(\ln \vec{p}) \wedge d\vec{\varphi} \rangle_{\vec{p}} - \langle d(\ln \vec{p}) \rangle_{\vec{p}} \wedge \langle d\vec{\varphi} \rangle_{\vec{p}} \right], \end{aligned} \quad (5)$$

where $\langle \cdot \rangle_{\vec{p}}$ denotes the expectation value with respect to the probability vector \vec{p} . The real part of this tensor is symmetric and defines the pullback to \mathcal{H} of the Fubini-Study metric g , while the imaginary part is antisymmetric and defines the pullback to \mathcal{H} of the symplectic form ω . Note that (the pullback of) g is composed of two terms, the first one is equivalent to the Fisher-Rao metric on the space of probability vectors $(p_1 \dots p_n)$, while the second term can be interpreted as a quantum contribution to the Fisher-Rao metric due to the phase of the state. In Ref. 5, it is shown that, once a particular submanifold of pure states $\mathcal{M} \subseteq \mathcal{P}(\mathcal{H})$ is chosen, the Hermitian tensor h induces a tensor $h_{\mathcal{M}}$ on \mathcal{M} , which can be degenerate. The real part of $h_{\mathcal{M}}$ defines a metric tensor $g_{\mathcal{M}}$ which, again, can be interpreted as a quantum analog of the Fisher-Rao metric on \mathcal{M} . However, \mathcal{M} is not necessarily a Kähler submanifold of the space of pure states; hence, the existence of a Kähler potential there is not guaranteed.

Since the idea of using a potential function to describe geometrical structures naturally fits into the conceptual framework of the geometrical formulation of quantum mechanics, it makes sense to ask whether there is a “potential-like function” for the quantum Fisher-Rao metric on \mathcal{M} . Because we cannot rely on complex coordinates, a possible potential function should be expressed in terms of real coordinates on the real manifold \mathcal{M} . Interestingly, a similar problem is deeply investigated in the geometrical approach to classical information theory pioneered by Amari,⁶ where the Fisher-Rao metric of a statistical manifold is recovered by means of the so-called contrast function S . This is a two-point function, i.e., a function on $\mathcal{M} \times \mathcal{M}$, that, analogous to the Kähler potential K , contains all the geometrical information of the statistical manifold. Furthermore, in Ref. 7, the possibility of using the contrast function S of information geometry in order to define a Kähler structure on $\mathcal{M} \times \mathcal{M}$ is analyzed.

It is then natural to look at this classical case with a more “quantum” attitude in order to unveil differences and analogies that could lead to a bidirectional flow of ideas and mathematical tools between the two settings. For instance, since there is no complex structure J in the classical setting, we think that a more thorough investigation of this situation could help to understand and point out what the role of J in the quantum setting is. Such a programme is highly non-trivial and cannot be accomplished in a single work.

In this contribution, we want to analyze the problem of finding a potential function for a statistical manifold in the framework of classical information geometry from a different point of view. We will formulate the problem in a geometric and dynamical framework common to both the classical and quantum setting, and we will show that a solution of the Hamilton-Jacobi problem for a canonically defined Lagrangian \mathcal{Q} is actually a potential function for a given statistical manifold \mathcal{M} . We point out that this formulation of the problem does not depend on the fact that \mathcal{M} is a statistical manifold in the sense of information geometry, that is, our formulation can be naturally applied to cases in which the metric tensor on \mathcal{M} is not the classical Fisher-Rao metric. The results presented here suggest that it could be possible to think of well-known divergence functions, such as the Kullback-Leibler divergence, as the Hamilton principal function of some suitable Lagrangian.

II. HAMILTON-JACOBI THEORY AND POTENTIAL FUNCTIONS

Following Lauritzen (Ref. 8), a statistical manifold is a triple (\mathcal{M}, g, T) , where \mathcal{M} is a differential manifold whose points parameterize a family of probability distributions, g is a metric tensor on \mathcal{M} , and T is a symmetric covariant tensor of order 3 on \mathcal{M} , called the skewness tensor. Starting with the tensor T and the Christoffel symbols ${}_{\circ}\Gamma$ of the Levi-Civita connection of g , it is possible to define a

family of affine torsionless connections ∇^α whose Christoffel symbols are

$${}^\alpha\Gamma_{jkl} := {}_g\Gamma_{jkl} - \frac{\alpha}{2}T_{jkl}, \tag{6}$$

where T_{jkl} and ${}_g\Gamma_{jkl}$ are, respectively, the components of T and ${}_g\Gamma$.

In the context of classical information geometry, we have a measure space \mathcal{X} , the space $\mathcal{P}(\mathcal{X})$ of probability distributions on \mathcal{X} , and a statistical manifold \mathcal{M} is identified with a submanifold of $\mathcal{P}(\mathcal{X})$ by means of an injective map

$$\mathcal{M} \ni m \mapsto p(\mathbf{x}, m) \, dx \in \mathcal{P}(\mathcal{X}), \tag{7}$$

where dx is a reference measure on \mathcal{X} . Then, introducing a coordinate system $\{\xi^j\}$ on \mathcal{M} , the Fisher-Rao metric g and the skewness tensor T can be obtained by

$$g_{jk} = \int_{\mathcal{X}} p(\mathbf{x}, \xi) \left(\frac{\partial \log(p)}{\partial \xi^j} \right) \left(\frac{\partial \log(p)}{\partial \xi^k} \right) dx, \tag{8}$$

$$T_{jkl} = \int_{\mathcal{X}} p(\mathbf{x}, \xi) \left(\frac{\partial \log(p)}{\partial \xi^j} \right) \left(\frac{\partial \log(p)}{\partial \xi^k} \right) \left(\frac{\partial \log(p)}{\partial \xi^l} \right) dx. \tag{9}$$

For every pair $(\alpha, -\alpha)$, the connections ∇^α and $\nabla^{-\alpha}$ are dual with respect to the metric g in the sense that the following equation holds:

$$Z(g(X, Y)) = g(\nabla_Z^\alpha X, Y) + g(X, \nabla_Z^{-\alpha} Y) \tag{10}$$

for all vector fields X, Y, Z on \mathcal{M} . The Levi-Civita connection $\nabla_g = \nabla^0$ is the only self-dual torsionless connection with respect to the metric g , and a statistical manifold for which $T = 0$ is called self-dual.

Remark 1. A statistical structure on \mathcal{M} can be alternatively defined using g and the pair of torsionless dual connections $\nabla \equiv \nabla^1, \nabla^* \equiv \nabla^{-1}$ (Refs. 6 and 8). In this framework, if the connections ∇ and ∇^* are flat connections, the statistical manifold $(\mathcal{M}, g, \nabla, \nabla^*)$ is called dually flat. This obviously poses some topological obstructions on \mathcal{M} , for instance, it turns out that the space of pure states of a finite-level quantum system does not admit a dually flat structure (see Refs. 9 and 10). According to Eq. (6), it is clear that the statistical manifolds (\mathcal{M}, g, T) and $(\mathcal{M}, g, \nabla, \nabla^*)$ are completely equivalent.

It is possible to prove^{6,8,11} that the geometrical structure of every statistical manifold can be completely encoded in a two-point function $S: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ called the contrast function. This is a distance-like function such that

$$S(m_1, m_2) \geq 0 \quad \forall m_1, m_2, \tag{11}$$

$$S(m_1, m_2) = 0 \quad \text{iff } m_1 = m_2. \tag{12}$$

Here, the first \mathcal{M} is thought of as the manifold of initial points whose coordinates we denote by q_{in} , and the second \mathcal{M} is the manifold of final points whose coordinates we denote by q_{fin} . If S is at least C^3 ,¹¹ it follows that

$$\left. \frac{\partial S}{\partial q_{in}^j} \right|_{q_{in}=q_{fin}} = \left. \frac{\partial S}{\partial q_{fin}^j} \right|_{q_{in}=q_{fin}} = 0. \tag{13}$$

The metric g and the tensor T are recovered from it as follows:

$$\left. \frac{\partial^2 S}{\partial q_{in}^j \partial q_{in}^k} \right|_{q_{in}=q_{fin}} = \left. \frac{\partial^2 S}{\partial q_{fin}^j \partial q_{fin}^k} \right|_{q_{in}=q_{fin}} = - \left. \frac{\partial^2 S}{\partial q_{fin}^j \partial q_{in}^k} \right|_{q_{in}=q_{fin}} = g_{jk}, \tag{14}$$

$$\left. \frac{\partial^3 S}{\partial q_{in}^l \partial q_{fin}^k \partial q_{fin}^j} \right|_{q_{in}=q_{fin}} - \left. \frac{\partial^3 S}{\partial q_{fin}^l \partial q_{in}^k \partial q_{in}^j} \right|_{q_{in}=q_{fin}} = T_{jkl}. \tag{15}$$

Note that the restriction of the second and third derivatives of S to the diagonal defines the tensor fields because of Eq. (13). It is important to note that S is never unique, and this leads to the need for the definition of a contrast function which is canonical in some suitable sense.¹²

What we propose is to interpret the task of finding a canonical potential function for the statistical manifold (\mathcal{M}, g, T) in the context of Hamilton-Jacobi theory associated with a particular Lagrangian built directly from the metric g and the symmetric tensor T . Note that the dynamical approach to potential functions presented here is purely geometric in the sense that it relies only on the geometrical structure of \mathcal{M} . This means that \mathcal{M} need not to be a statistical manifold endowed with the Fisher-Rao metric and the canonical skewness tensor of information geometry, but it could be a generic manifold endowed with a generic metric tensor g and a generic skewness tensor T . This is particularly useful with respect to quantum mechanics, where quantum states are probability amplitudes and not genuine probability distributions, and where, for invertible mixed states, there is an infinite number of possible generalizations of the Fisher-Rao metric²⁶ (Refs. 13 and 14).

To keep the article as self-contained as possible, we briefly recall the main points of Hamilton-Jacobi theory (for a more detailed formulation of the problem we refer to Refs. 15–17). In the variational formulation of dynamics,¹⁸ the solutions of the equations of motion are expressed as the critical points of the action functional

$$I(\gamma) = \int_{t_{\text{in}}}^{t_{\text{fin}}} \mathfrak{L}(\gamma, \dot{\gamma}) dt, \quad (16)$$

where γ are curves with fixed extreme points $q(t_{\text{in}}) = q_{\text{in}}$ and $q(t_{\text{fin}}) = q_{\text{fin}}$, and \mathfrak{L} is the Lagrangian function of the system. The evaluation of the action functional on a critical point γ_c gives a two-point function

$$S(q_{\text{in}}, q_{\text{fin}}) = I(\gamma_c), \quad (17)$$

which is known in the literature as the Hamilton characteristic function. It is a solution of the Hamilton-Jacobi equation for the dynamics

$$H\left(q, \frac{\partial S}{\partial q}\right) = E, \quad (18)$$

where H is the Hamiltonian function (Ref. 19) associated with the Lagrangian \mathfrak{L} and $E \in \mathbb{R}$ is a constant. In particular, $S(q_{\text{in}}, q_{\text{fin}})$ is called a complete solution when

$$\det \left\| \frac{\partial^2 S}{\partial q_{\text{in}}^j \partial q_{\text{fin}}^k} \right\| \neq 0. \quad (19)$$

Therefore, S is the generating function of a canonical transformation on the phase space of the system. Specifically, we have the following relations:

$$p_j^{\text{in}} = - \frac{\partial S}{\partial q_{\text{in}}^j}, \quad (20)$$

$$p_j^{\text{fin}} = \frac{\partial S}{\partial q_{\text{fin}}^j}, \quad (21)$$

where $\{p_j^{\text{in}}\}$ (respectively, $\{p_j^{\text{fin}}\}$) are the canonical momenta associated with q_{in}^j 's (respectively, q_{fin}^j 's).

The fact that S is a two-point function allows us to read the problem of finding a canonical contrast function on a statistical manifold as the Hamilton-Jacobi problem associated with suitable Lagrangian and Hamiltonian functions. Indeed, consider a statistical manifold (\mathcal{M}, g, T) , let $\alpha \neq 0$ be a real number, and let us define the following Lagrangian function:

$$\mathfrak{L}_\alpha(q, v) = \frac{1}{2} g_{jk}(q) v^j v^k + \frac{\alpha}{6} T_{jkl}(q) v^j v^k v^l. \quad (22)$$

Our claim is that a complete solution S_α of the Hamilton-Jacobi equation associated with this Lagrangian is a potential function for our statistical manifold in the sense that it allows us to recover the geometric structure of the manifold as follows:

$$\left. \frac{\partial^2 S_\alpha}{\partial q_{\text{fin}}^j \partial q_{\text{in}}^k} \right|_{q_{\text{in}}=q_{\text{fin}}} = -g_{jk}, \quad (23)$$

$$\left. \frac{\partial^3 S_\alpha}{\partial q_{\text{in}}^l \partial q_{\text{in}}^k \partial q_{\text{fin}}^j} \right|_{q_{\text{in}}=q_{\text{fin}}} - \left. \frac{\partial^3 S_\alpha}{\partial q_{\text{fin}}^l \partial q_{\text{fin}}^k \partial q_{\text{in}}^j} \right|_{q_{\text{in}}=q_{\text{fin}}} = 2\alpha T_{jkl}. \tag{24}$$

Equation (24) is slightly different from Eq. (15), consequently, we have chosen the name potential function instead of contrast function for S_α because, as we shall see, S_α allows us to recover the geometrical structures of the statistical manifold. Notice further that S_α does not need to be positive semidefinite, while a contrast function must be so.

Remark 2. In the following, we will use the term “contrast function” for a two-point function satisfying the conditions given in Eqs. (11) and (12), see paragraph after Remark 1. We will use the term “potential function” for a generic two-point function by means of which it is possible to recover the geometrical structure of a statistical manifold (\mathcal{M}, g, T) using a suitably defined procedure, such as that given by Eqs. (23) and (24).

Note that it is possible to write the Lagrangian (22) in intrinsic form as follows:

$$\mathfrak{L}_\alpha = \mathfrak{L}_g + \frac{1}{3} L_{\alpha\Gamma} \mathfrak{L}_g, \tag{25}$$

where \mathfrak{L}_g is the metric Lagrangian associated with g and $\alpha\Gamma$ is the second order vector field²⁰ associated with the affine connection ∇^α (6). By looking at this expression we can notice that this Lagrangian can be considered as a sort of first-order approximation of a more complete function $\widetilde{\mathfrak{L}}_\alpha$ which also includes all successive Lie derivatives with respect to the vector field Γ , shortly

$$\widetilde{\mathfrak{L}}_\alpha = e^{\tau L_{\alpha\Gamma}} \mathfrak{L}_g = (\Phi_\tau^\alpha)^* \mathfrak{L}_g, \tag{26}$$

where Φ_τ^α is the flow of the second order vector field $\alpha\Gamma$ on the tangent bundle $T\mathcal{M}$ of the statistical manifold \mathcal{M} . However, only (25) contributes to the determination of metric and skewness tensors, as we will prove in the following.

According to Eq. (20), we have $\frac{\partial S_\alpha}{\partial q_{\text{in}}^j} = -p_j^{\text{in}}$. Furthermore, the momenta p_j can be expressed in terms of the Lagrangian function as

$$p_j = \frac{\partial \mathfrak{L}}{\partial v^j}. \tag{27}$$

In particular, for our Lagrangian we get

$$p_j = g_{jk}(q)v^k + \frac{\alpha}{2} T_{jkl}(q)v^k v^l. \tag{28}$$

From this, it follows that

$$\frac{\partial S_\alpha}{\partial q_{\text{in}}^j} = -p_j^{\text{in}} = -g_{jk}(q_{\text{in}})v_{\text{in}}^k - \frac{\alpha}{2} T_{jkl}(q_{\text{in}})v_{\text{in}}^k v_{\text{in}}^l, \tag{29}$$

where v_{in}^j 's must be expressed in terms of the initial and final positions. The link between initial and final positions $(q_{\text{in}}, q_{\text{fin}})$ and the initial velocity (v_{in}) is provided by the dynamical trajectories γ_c associated with the Lagrangian \mathfrak{L} . The Euler-Lagrange equations associated with \mathfrak{L} are

$$(g_{jk}(q) + \alpha T_{jkl} v^l) \dot{v}^k = -g_{\Gamma jkl} v^l v^k - \frac{\alpha}{6} \left(\frac{\partial T_{jkl}}{\partial q^m} + \frac{\partial T_{jlm}}{\partial q^k} + \frac{\partial T_{jkm}}{\partial q^l} - \frac{\partial T_{klm}}{\partial q^j} \right) v^k v^l v^m, \tag{30}$$

where $v^j = \frac{dq^j}{dt}$ and $\dot{v}^j = \frac{dv^j}{dt}$. A series expansion of $\gamma_c(t) = (q^1(t), \dots, q^n(t))$ around $t = 0$ yields

$$q^j(t) = q_{\text{in}}^j + t \left. \frac{dq^j}{dt} \right|_{t=0} + \frac{t^2}{2} \left. \frac{d^2 q^j}{dt^2} \right|_{t=0} + \mathcal{O}(t^3). \tag{31}$$

We know that $v_{\text{in}}^j = \left. \frac{dq^j}{dt} \right|_{t=0}$, so that naming $q_{\text{fin}}^j := q^j(1)$, we can write

$$v_{\text{in}}^j = q_{\text{fin}}^j - q_{\text{in}}^j - \frac{1}{2} \left. \frac{dv_{\text{in}}^j}{dt} \right|_{t=0}, \tag{32}$$

where higher order terms in the expansion have been neglected. Since v_{in} is a function of q_{in} , q_{fin} , we can express the derivatives with respect to q_{fin} in terms of the derivatives of v_{in} and vice versa. Indeed²⁷

$$\frac{\partial}{\partial q_{fin}^j} = \frac{\partial v_{in}^k}{\partial q_{fin}^j} \frac{\partial}{\partial v_{in}^k}, \tag{33}$$

$$\frac{\partial^2}{\partial q_{fin}^k \partial q_{fin}^j} = \frac{\partial^2 v_{in}^l}{\partial q_{fin}^k \partial q_{fin}^j} \frac{\partial}{\partial v_{in}^l} + \frac{\partial v_{in}^l}{\partial q_{fin}^j} \frac{\partial v_{in}^r}{\partial q_{fin}^k} \frac{\partial}{\partial v_{in}^r} \frac{\partial}{\partial v_{in}^l}, \tag{34}$$

and we need to evaluate these expressions on the diagonal $q_{fin} = q_{in}$. Note that the condition $q_{fin} = q_{in}$ is equivalent to the fact that the dynamical trajectory is $q^j(t) = q_{in}^j$, and thus, according to the equations of motion, this corresponds to $v_{in} = 0$.

Equation (30) can be written as follows:

$$\dot{v}^l = -\alpha T_{jk}^l v^k \dot{v}^j - g \Gamma_{kj}^l v^j v^k - \frac{\alpha}{6} g^{lj} \left(\frac{\partial T_{jkr}}{\partial q^m} + \frac{\partial T_{jrm}}{\partial q^k} + \frac{\partial T_{jkm}}{\partial q^r} - \frac{\partial T_{krm}}{\partial q^j} \right) v^k v^r v^m. \tag{35}$$

If we suppose that \dot{v}^j is an analytic function of $\{v^j\}$ in a neighbourhood of $v^j = 0$, we can write

$$\dot{v}^k = \sum_{m=0}^{\infty} \sum_{j_1, \dots, j_m=1}^n a_{j_1 \dots j_m}^k v^{j_1} \dots v^{j_m}. \tag{36}$$

By inserting this expression into Eq. (30), we get the coefficients

$$a_0^k = 0, \tag{37}$$

$$a_{j_1}^k = 0, \tag{38}$$

$$a_{j_1 j_2}^k = -g \Gamma_{j_1 j_2}^k \tag{39}$$

and so on.

Therefore \dot{v}^j is a function of order $\mathcal{O}(|v|^2)$ and higher order derivatives $\{\ddot{v}^j, \dots\}$ will be at least of order $\mathcal{O}(|v|^3)$. We can now put these results into Eq. (32) to obtain

$$v_{in}^j = q_{fin}^j - q_{in}^j + \frac{1}{2} g \Gamma_{kl}^j v_{in}^k v_{in}^l + \mathcal{O}(|v|^3). \tag{40}$$

Deriving this expression with respect to q_{fin} and then evaluating it at $v_i = 0$, we have

$$\left. \frac{\partial v_i^k}{\partial q_{fin}^j} \right|_{q_{in}=q_{fin}} = \delta_j^k, \tag{41}$$

$$\left. \frac{\partial^2 v_i^l}{\partial q_{fin}^k \partial q_{fin}^j} \right|_{q_{in}=q_{fin}} = g \Gamma_{jk}^l. \tag{42}$$

Eventually we get

$$\left. \frac{\partial^2 S_\alpha}{\partial q_{fin}^k \partial q_{in}^j} \right|_{q_{in}=q_{fin}} = - \left. \frac{\partial p_j^{in}}{\partial v_{in}^k} \right|_{v_{in}=0} = -g_{jk}, \tag{43}$$

$$\left. \frac{\partial^3 S_\alpha}{\partial q_{fin}^l \partial q_{fin}^k \partial q_{in}^j} \right|_{q_{in}=q_{fin}} = -g \Gamma_{kl}^r \left. \frac{\partial p_j^{in}}{\partial v_{in}^r} \right|_{v_{in}=0} - \left. \frac{\partial^2 p_j^{in}}{\partial v_{in}^l \partial v_{in}^k} \right|_{v_{in}=0} = -g \Gamma_{jkl} - \alpha T_{jkl}. \tag{44}$$

From them it follows immediately that the metric tensor is derived from the metric Lagrangian only, whereas information about the connection depends on the ‘‘interaction term,’’ as it should be since quadratic terms alone cannot contribute to third order derivatives. In particular, when $\alpha = 0$, we get the Christoffel symbols of the Levi-Civita connection associated with the metric g .

In order to extract the symmetric tensor from our potential function, we need to take derivatives in a different order, according to Eq. (24). The main difference with respect to our previous description is the fact that

$$\frac{\partial S_\alpha}{\partial q_{\text{fin}}^j} = p_j^{\text{fin}}, \quad (45)$$

which is the canonical momentum at the extreme $\gamma(t=1) = q_{\text{fin}}$.

Following the procedure just outlined, we have to express the dependence of v_{fin} on the variables $(q_{\text{in}}, q_{\text{fin}})$. This relation is provided by the dynamics, which is reversible. Then it follows that

$$v_{\text{fin}}^j = q_{\text{fin}}^j - q_{\text{in}}^j - \frac{1}{2} g \Gamma_{kl}^j v_{\text{fin}}^k v_{\text{fin}}^l + \mathcal{O}(|v|^3). \quad (46)$$

Eventually we get

$$\left. \frac{\partial^2 S_\alpha}{\partial q_{\text{in}}^k \partial q_{\text{fin}}^j} \right|_{q_{\text{in}}=q_{\text{fin}}} = - \left. \frac{\partial p_j^{\text{fin}}}{\partial v_{\text{fin}}^k} \right|_{v_{\text{fin}}=0} = -g_{jk}, \quad (47)$$

$$\left. \frac{\partial^3 S_\alpha}{\partial q_{\text{in}}^l \partial q_{\text{in}}^k \partial q_{\text{fin}}^j} \right|_{q_{\text{in}}=q_{\text{fin}}} = -g \Gamma_{kl}^r \left. \frac{\partial p_j^{\text{fin}}}{\partial v_{\text{fin}}^r} \right|_{v_{\text{fin}}=0} + \left. \frac{\partial^2 p_j^{\text{fin}}}{\partial v_{\text{fin}}^l \partial v_{\text{fin}}^k} \right|_{v_{\text{fin}}=0} = -g \Gamma_{jkl} + \alpha T_{jkl}. \quad (48)$$

From Eqs. (23) and (24), it follows that S_α is actually a potential function for the statistical manifold (M, g, T) . Note that S_α need not be positive as a contrast function would be; however, this is not an obstruction in determining metric and skewness tensors. Indeed, it is sufficient that S_α has a local extreme on the diagonal $q_{\text{in}} = q_{\text{fin}}$ (or $v = 0$), and it is true in our case as shown above, by calculating the Hessian matrices (43) and (47). Furthermore, as already noticed, one could also think to add other interaction terms to the basic Lagrangian (25), and by a suitable choice of the coupling constants, it is possible to make the contrast function positive definite.

Furthermore, we want to stress that, from the purely mathematical point of view, the information contained in the metric tensor g and the skewness tensor T is completely uncorrelated with the fact that \mathcal{M} is actually a manifold of probability distributions as it is the case in classical information geometry. The Hamilton-Jacobi approach to potential functions introduced here makes reference only to the geometrical properties of \mathcal{M} encoded in g and T and no reference to the fact that \mathcal{M} is a manifold of probability distributions is needed. Therefore, this approach works perfectly well when \mathcal{M} is not a manifold of probability distributions, as it is the case for quantum information geometry, where the focus is on quantum states, i.e., probability amplitudes, and the relevant statistical manifolds are manifolds of quantum states.

III. FEATURES OF THE HAMILTON-JACOBI APPROACH

Here we will discuss some features of the Hamilton-Jacobi approach to potential functions described in Sec. II.

Let us start pointing out the connection between the potential function S_α defined here and the canonical contrast function defined on self-dual statistical manifolds (Ref. 6). We recall that these are statistical manifolds for which the symmetric tensor T identically vanishes so that the only connection at our disposal is the self-dual Levi-Civita connection ∇_g associated with the metric g .

For self-dual manifolds, a canonical contrast function S_d exists which is given by

$$S_d(\mathbf{q}_{\text{in}}, \mathbf{q}_{\text{fin}}) = \frac{1}{2} d^2(\mathbf{q}_{\text{in}}, \mathbf{q}_{\text{fin}}), \quad (49)$$

where $d^2(q_{\text{in}}, q_{\text{fin}})$ is the square of the Riemannian geodesic distance associated with the metric g on \mathcal{M} .

Applying our procedure to the case of self-dual manifolds, it is clear that the potential function we obtain is precisely the canonical contrast function S_d defined above. To see this, recall that the metric Lagrangian \mathcal{Q}_g associated with the metric tensor g , and all of its functions $F(\mathcal{Q}_g)$ with F analytic, gives rise to the same dynamical trajectories (Ref. 20). Furthermore, \mathcal{Q}_g and $F(\mathcal{Q}_g)$ are all constants of

the motion for the dynamics. Denoting with γ_c the geodesic connecting $\mathbf{q}_{\text{in}} = \gamma_c(0)$ and $\mathbf{q}_{\text{fin}} = \gamma_c(1)$, this implies that

$$I_{F(\mathcal{Q}_g)}(\mathbf{q}_{\text{in}}, \mathbf{q}_{\text{fin}}) := \int_{t_{\text{in}}}^{t_{\text{fin}}} F(\mathcal{Q}_g) dt = F(\mathcal{Q}_g) \Big|_{\gamma_c(t)}. \tag{50}$$

Now, the Riemannian geodesic distance $d^2(q_{\text{in}}, q_{\text{fin}})$ is given by

$$d(\mathbf{q}_{\text{in}}, \mathbf{q}_{\text{fin}}) = \int_{t_{\text{in}}}^{t_{\text{fin}}} \sqrt{2\mathcal{Q}_g} dt, \tag{51}$$

and thus

$$\begin{aligned} d^2(\mathbf{q}_{\text{in}}, \mathbf{q}_{\text{fin}}) &= \left(\int_{t_{\text{in}}}^{t_{\text{fin}}} \sqrt{2\mathcal{Q}_g} dt \right)^2 = \left(\sqrt{2\mathcal{Q}_g} \right)^2 \Big|_{\gamma_c(t)} = \\ &= 2\mathcal{Q}_g \Big|_{\gamma_c(t)} = 2 \int_{t_{\text{in}}}^{t_{\text{fin}}} \mathcal{Q}_g dt = 2 I_{\mathcal{Q}_g}(\mathbf{q}_{\text{in}}, \mathbf{q}_{\text{fin}}). \end{aligned} \tag{52}$$

Equations (16), (17), and (22), with $T = 0$, allow us to immediately realize that the Hamilton characteristic function associated with the Lagrangian $\mathcal{Q}_\alpha = \mathcal{Q}_g$ is precisely the canonical contrast function of Eq. (49).

A physically interesting example of the self-dual statistical manifold is given by the manifold $\mathcal{P}(\mathcal{H})$ of pure states of quantum mechanics. As it is shown in Ref. 21, a meaningful notion of the statistical distance between pure states can be defined by means of the concepts of distinguishability and statistical fluctuations in the outcomes of measurements. It turns out that this two-point function on pure states coincides with the Riemannian geodesic distance associated with the Fubini-Study metric so that the statistical structure determined by this two-point function makes $\mathcal{P}(\mathcal{H})$ a self-dual manifold. Consequently, we can apply our procedure and conclude that the statistical distance introduced by Wootters coincides with the Hamilton principal function associated with the metric Lagrangian of the Fubini-Study metric on $\mathcal{P}(\mathcal{H})$. The relevance of the statistical structure on $\mathcal{P}(\mathcal{H})$ emerging from Wootters statistical distance is enforced by the results of Refs. 9 and 10, where it is shown that the set of pure states of quantum mechanics does not admit a dually flat statistical structure.

Now, consider the statistical manifold (\mathcal{M}, g, T) , where $\mathcal{M} = \mathbb{R}^+$, $g = \frac{1}{\xi^2}$, ${}_g\Gamma = -\frac{1}{\xi}$ is the Christoffel symbol of the Levi-Civita connection, and $T = -\frac{2}{\xi^3}$. This manifold arises as the statistical manifold associated with the exponential distributions

$$p(x, \xi) = \xi e^{-x\xi} \quad \xi, x > 0. \tag{53}$$

The metric g and the tensor T are then obtained by

$$g = \int_0^{+\infty} p(x, \xi) \left(\frac{d \log(p)}{d\xi} \right)^2 dx, \tag{54}$$

$$T = \int_0^{+\infty} p(x, \xi) \left(\frac{d \log(p)}{d\xi} \right)^3 dx. \tag{55}$$

The ‘‘deformed’’ Lagrangian function \mathcal{Q} with respect to the connection ∇ reads

$$\mathcal{Q}_\alpha = \frac{v^2}{2\xi^2} - \frac{\alpha}{3} \frac{v^3}{\xi^3}, \tag{56}$$

where $\mathcal{Q}_g = \frac{v^2}{2\xi^2}$ is the metric Lagrangian. It is clear that \mathcal{Q}_α is a function of \mathcal{Q}_g , specifically, it is $\mathcal{Q}_\alpha = \mathcal{Q}_g + \frac{2\sqrt{2}\alpha}{3} \left(\mathcal{Q}_g \right)^{\frac{3}{2}}$. Consequently, the solutions of the Euler-Lagrange equations associated with the metric Lagrangian \mathcal{Q}_g , i.e., the geodesics of g , are solutions of the Euler-Lagrange equations associated with the Lagrangian \mathcal{Q}_α , and the explicit expression of the dynamical trajectories $\gamma_c(t) = \xi(t)$ is

$$\xi(t) = \xi_{\text{in}} e^{\frac{v_{\text{in}}}{\xi_{\text{in}}} t}. \tag{57}$$

A complete solution of the Hamilton-Jacobi problem for \mathcal{Q} is given by

$$I_\alpha(\gamma_c) = \int_{t_{\text{in}}}^{t_{\text{fin}}} \mathcal{Q}_\alpha(\gamma_c(t), \dot{\gamma}_c(t)) dt, \quad (58)$$

where the curve γ_c has fixed extreme points $\xi_{\text{in}} = \gamma_c(t_{\text{in}})$ and $\xi_{\text{fin}} = \gamma_c(t_{\text{fin}})$, and integration is performed between $t_{\text{in}} = 0$ and $t_{\text{fin}} = 1$. In our case, since the Lagrangian \mathcal{Q}_α is a constant of the motion, we have

$$I_\alpha(\gamma_c) = \frac{v_{\text{in}}^2}{2\xi_{\text{in}}^2} - \frac{\alpha}{3} \frac{v_{\text{in}}^3}{\xi_{\text{in}}^3}. \quad (59)$$

The link between ξ_{in} , ξ_{fin} and v_{in} can easily be extracted from the explicit expression of $\gamma_c(t)$, indeed

$$v_{\text{in}} = \xi_{\text{in}} \ln \left(\frac{\xi_{\text{fin}}}{\xi_{\text{in}}} \right), \quad (60)$$

and thus, the contrast function S reads

$$S_\alpha(\xi_{\text{in}}, \xi_{\text{fin}}) = \frac{\ln^2 \left(\frac{\xi_{\text{fin}}}{\xi_{\text{in}}} \right)}{2} - \frac{\alpha}{3} \ln^3 \left(\frac{\xi_{\text{fin}}}{\xi_{\text{in}}} \right). \quad (61)$$

An explicit calculation gives

$$\left. \frac{\partial^2 S_\alpha}{\partial \xi_{\text{fin}} \partial \xi_{\text{in}}} \right|_{\xi_{\text{in}} = \xi_{\text{fin}} = \xi} = -\frac{1}{\xi^2}, \quad (62)$$

$$\left. \frac{\partial^3 S_\alpha}{\partial \xi_{\text{fin}} \partial \xi_{\text{fin}} \partial \xi_{\text{in}}} \right|_{\xi_{\text{in}} = \xi_{\text{fin}} = \xi} = \frac{2\alpha + 1}{\xi^3} = -{}_g\Gamma - \alpha T, \quad (63)$$

$$\left. \frac{\partial^3 S_\alpha}{\partial \xi_{\text{fin}} \partial \xi_{\text{in}} \partial \xi_{\text{in}}} \right|_{\xi_{\text{in}} = \xi_{\text{fin}} = \xi} = \frac{1 - 2\alpha}{\xi^3} = -{}_g\Gamma + \alpha T, \quad (64)$$

showing that S_α is a potential function for the statistical manifold (\mathcal{M}, g, T) of exponential distributions.

The statistical structure of \mathcal{M} can be alternatively derived starting with the Kullback-Leibler divergence function S_{KL} ,

$$S_{KL}(\xi_{\text{in}}, \xi_{\text{fin}}) = \int_0^{+\infty} p(x, \xi_{\text{in}}) \ln \left(\frac{p(x, \xi_{\text{in}})}{p(x, \xi_{\text{fin}})} \right) dx = \ln \left(\frac{\xi_{\text{in}}}{\xi_{\text{fin}}} \right) + \frac{\xi_{\text{fin}}}{\xi_{\text{in}}} - 1. \quad (65)$$

As it is clear, the potential function S_α in Eq. (61) does not coincide with the Kullback-Leibler divergence. This is not surprising since, for a given statistical manifold, there are infinite many potential (contrast) functions generating the same statistical structure. However, we will now show that it is possible to read the Kullback-Leibler divergence S_{KL} as the Hamilton principal function I_{KL} associated with a suitably defined Lagrangian. At this purpose, let us perform the following diffeomorphism between $\mathcal{M} = \mathbb{R}_+$ and $\mathcal{N} = \mathbb{R}$:

$$\xi \mapsto y = \ln(\xi). \quad (66)$$

This diffeomorphism gives rise to a diffeomorphism between $\mathcal{M} \times \mathcal{M}$ and $\mathcal{N} \times \mathcal{N}$,

$$(\xi_{\text{in}}, \xi_{\text{fin}}) \mapsto (y_{\text{in}} = \ln(\xi_{\text{in}}), y_{\text{fin}} = \ln(\xi_{\text{fin}})). \quad (67)$$

The Kullback-Leibler divergence S_{KL} becomes

$$S_{KL}(y_{\text{in}}, y_{\text{fin}}) = e^{(y_{\text{fin}} - y_{\text{in}})} - (y_{\text{fin}} - y_{\text{in}}) - 1. \quad (68)$$

Now, consider the following Lagrangian on $T\mathcal{N}$:

$$\mathcal{Q}_{KL}(y, u) = e^u - u - 1, \quad (69)$$

and let us calculate the Hamilton principal function associated with \mathcal{Q}_{KL} . Since \mathcal{Q}_{KL} depends only on the velocity coordinate u , it is an alternative Lagrangian for the 1-dimensional free-particle on $\mathcal{N} = \mathbb{R}$ (Ref. 20). Consequently, the dynamical trajectories of the system coincide with the geodesics of the Euclidean metric, that is, they are straight lines $\gamma_c(t) = v_{\text{in}}t + y_{\text{in}}$. Setting $t_{\text{in}} = 0$, $t_{\text{fin}} = 1$, the connection

between y_{in}, y_{fin} and v_{in} is easily seen to be $v_{in} = y_{fin} - y_{in}$. Furthermore, \mathcal{Q}_{KL} is a constant of the motion, and thus, it can be brought out from the integral defining the Hamilton principal function

$$I_{KL}(y_{in}, y_{fin}) = \mathcal{Q}_{KL}(y_{in}, v_{in}(y_{in}, y_{fin})) = e^{(y_{fin} - y_{in})} - (y_{fin} - y_{in}) - 1. \tag{70}$$

Confronting Eqs. (68) and (70), we conclude that the Hamilton principal function associated with \mathcal{Q}_{KL} is precisely the Kullback-Leibler divergence S_{KL} as claimed.

Now we will point out an interesting connection between the Lagrangian \mathcal{Q}_{KL} associated with the Kullback-Leibler divergence of Eq. (65) and the Lagrangian \mathcal{Q}_α associated with the potential function of Eq. (61). At this purpose, let us apply the tangent lift of the diffeomorphism given by Eq. (66) to the Lagrangian \mathcal{Q}_α of Eq. (56),

$$\mathcal{Q}_\alpha(y, u) = \frac{u^2}{2} - \frac{\alpha}{3} u^3. \tag{71}$$

Now, let us perform a series expansion of \mathcal{Q}_{KL} around $u = 0$,

$$\mathcal{Q}_{KL}(y, u) = \frac{u^2}{2} + \frac{u^3}{6} + \mathcal{O}(u^4). \tag{72}$$

Confronting Eqs. (71) and (72), we immediately see that, upon taking $\alpha = -\frac{1}{2}$, the Lagrangian \mathcal{Q}_α is precisely the third order approximation of \mathcal{Q}_{KL} .

We can push this line of reasoning a little further, and show that, if the Kullback-Leibler divergence S_{KL} generating the statistical structure of a statistical manifold (\mathcal{M}, g, T) is the Hamilton principal function associated with a Lagrangian \mathcal{Q}_{KL} , then the Lagrangian \mathcal{Q}_α we have proposed here is the third order approximation of \mathcal{Q}_{KL} up to a constant factor, provided we choose $\alpha = -\frac{1}{2}$.

Let us consider a statistical manifold (\mathcal{M}, g, T) and the Kullback-Leibler divergence S_{KL} generating the statistical structure of \mathcal{M} . Let us assume that S_{KL} admits a Lagrangian \mathcal{Q}_{KL} such that S_{KL} is the Hamilton principal function associated with \mathcal{Q}_{KL} . Assuming \mathcal{Q}_{KL} analytic in \mathbf{v} as in Sec. II, an expansion of \mathcal{Q}_{KL} in a power series of the velocity vector \mathbf{v} around $\mathbf{v} = 0$ gives

$$\mathcal{Q}_{KL} = \mathcal{Q}_{KL}|_{\mathbf{v}=0} + \left. \frac{\partial \mathcal{Q}_{KL}}{\partial v^j} \right|_{\mathbf{v}=0} v^j + \left. \frac{\partial^2 \mathcal{Q}_{KL}}{\partial v^j \partial v^k} \right|_{\mathbf{v}=0} \frac{v^j v^k}{2!} + \left. \frac{\partial^3 \mathcal{Q}_{KL}}{\partial v^j \partial v^k \partial v^l} \right|_{\mathbf{v}=0} \frac{v^j v^k v^l}{3!} + \mathcal{O}(\mathbf{v}^4).$$

We will now examine the terms of the expansion up to the third order.

Concerning the first order term, we recall that S_{KL} must have a minimum on the diagonal $\mathbf{q}_{in} = \mathbf{q}_{fin}$ of $\mathcal{M} \times \mathcal{M}$. Therefore, recalling Eqs. (20) and (13) and expressing the derivatives with respect to v^j by means of the derivatives with respect to q_{fin}^j , we have

$$\left. \frac{\partial \mathcal{Q}_{KL}}{\partial v^j} \right|_{\mathbf{v}=0} = - \left. \frac{\partial S_{KL}}{\partial q_{in}^j} \right|_{\mathbf{q}_{in}=\mathbf{q}_{fin}} = 0. \tag{73}$$

This equation implies that \mathcal{Q}_{KL} is at least of second order in \mathbf{v} . Consequently, the analysis of Sec. II for the functional dependence between \mathbf{v} and \mathbf{q}_{fin} stemming from the Euler-Lagrange equations can be analogously repeated in order to give the first order relation

$$\left. \frac{\partial q_{fin}^j}{\partial v^j} \right|_{\mathbf{v}=0} = \delta_k^j. \tag{74}$$

This result, together with Eqs. (13), (14), and (73), allows us to see that the second order term of \mathcal{Q}_{KL} becomes

$$\left. \frac{\partial^2 \mathcal{Q}_{KL}}{\partial v^j \partial v^k} \right|_{\mathbf{v}=0} = - \left. \frac{\partial^2 S_{KL}}{\partial q_{fin}^r \partial q_{in}^j} \right|_{\mathbf{q}_{in}=\mathbf{q}_{fin}} \left. \frac{\partial q_{fin}^r}{\partial v^k} \right|_{\mathbf{v}=0} = g_{jk}. \tag{75}$$

Now that we have the second-order term of the Lagrangian, we can proceed in the analysis of the functional dependence between \mathbf{v} and \mathbf{q}_{fin} and find that

$$\left. \frac{\partial^2 q_{fin}^l}{\partial v^k \partial v^j} \right|_{\mathbf{v}=0} = -g^l{}_{jk}. \tag{76}$$

Consequently, the third order term of the Lagrangian is

$$\begin{aligned} \left. \frac{\partial^3 \mathfrak{L}_{KL}}{\partial v^j \partial v^k \partial v^l} \right|_{\mathbf{v}=0} &= - \left. \frac{\partial^3 S_{KL}}{\partial q_{\text{fin}}^r \partial q_{\text{fin}}^n \partial q_{\text{in}}^j} \right|_{\mathbf{q}_{\text{in}}=\mathbf{q}_{\text{fin}}} \left(\frac{\partial q_{\text{fin}}^r}{\partial v^k} \frac{\partial q_{\text{fin}}^n}{\partial v^l} \right) \Big|_{\mathbf{v}=0} - \left. \frac{\partial^2 S_{KL}}{\partial q_{\text{fin}}^r \partial q_{\text{in}}^j} \right|_{\mathbf{q}_{\text{in}}=\mathbf{q}_{\text{fin}}} \left. \frac{\partial^2 q_{\text{fin}}^r}{\partial v^l \partial v^k} \right|_{\mathbf{v}=0} = \\ &= - \left. \frac{\partial^3 S_{KL}}{\partial q_{\text{in}}^j \partial q_{\text{fin}}^k \partial q_{\text{fin}}^l} \right|_{\mathbf{q}_{\text{in}}=\mathbf{q}_{\text{fin}}} - {}_g\Gamma_{jkl} = -\frac{T_{jkl}}{2}, \end{aligned} \tag{77}$$

where, in the last equality, we have used the fact that

$$\left. \frac{\partial^3 S_{KL}}{\partial q_{\text{in}}^j \partial q_{\text{fin}}^k \partial q_{\text{fin}}^l} \right|_{\mathbf{q}_{\text{in}}=\mathbf{q}_{\text{fin}}} = {}_g\Gamma_{jkl} + \frac{T_{jkl}}{2}, \tag{78}$$

which is formula 4 of Lemma 2.1 in Ref. 11.

Collecting the results, we can write

$$\mathfrak{L}_{KL} = \mathfrak{L}_{KL}|_{\mathbf{v}=0} + g_{jk} \frac{v^j v^k}{2!} - \frac{T_{jkl}}{2} \frac{v^j v^k v^l}{3!} + \mathcal{O}(\mathbf{v}^4), \tag{79}$$

from which it follows that, choosing $\alpha = -\frac{1}{2}$, the Lagrangian \mathfrak{L}_α differs from the third order approximation of \mathfrak{L}_{KL} only by the constant factor $\mathfrak{L}_{KL}|_{\mathbf{v}=0}$ as claimed. It would be very interesting to understand the conditions under which the Kullback-Leibler divergence of a given statistical manifold is the Hamilton principal function of some suitably defined Lagrangian.

This appears to be relevant in the context of quantum information geometry of mixed states (Ref. 6, Chapter 7). Here, the quantum counterpart of the Fisher-Rao metric has been studied extensively by Petz and co-workers (Refs. 13 and 14 and references therein), who found that there is an infinite-number of metrics providing a meaningful generalization of the classical Fisher-Rao metric. Furthermore, unlike the classical case, there is no preferred definition for a skewness tensor T . This means that there is a good amount of freedom in the choice of a statistical structure on the space of quantum states. The usual way in which a statistical structure is defined is to start considering a generalization of some classical divergence function, and then derive a metric g and a tensor T on the space of quantum states. Interestingly, it is possible to use well known examples of quantum relative entropies as quantum divergence functions (Ref. 22). As it is clear, in this quantum setting, the statistical structure of the space of quantum states depends on the explicit form of the quantum divergence one starts with. Now, we have seen that the tensors g and T of a statistical manifold, being it classical or quantum, are completely encoded in the first four terms of the expansion of the Lagrangian associated with a divergence function; therefore, it is reasonable to argue that there must be some other geometrical information hidden in the divergence function that is not fully captured by g and T alone. We believe that the dynamical characterization of the quantum divergences stemming from the Hamilton-Jacobi approach outlined here can be fruitfully exploited to better understand the relations between the quantum divergences, the statistical structure they induce, and the geometrical structure of the space of quantum states. However, a careful analysis of these problems requires the use of a more advanced formulation of the Hamilton-Jacobi theory (Refs. 15 and 23) and will be left for further work.

Finally, let us comment on the other approaches to the definition of a potential (contrast) function for a statistical manifold (\mathcal{M}, g, T) . In Ref. 12, a canonical contrast function is constructed using the arclength functional l associated with the metric g and the so-called inverse exponential map Exp_{∇}^{-1} associated with an affine connection ∇ defined in terms of T and ${}_g\Gamma$. In this case, information about the geometrical structures of (\mathcal{M}, g, T) is taken into account separately. Specifically, the arclength functional l carries information about the metric tensor g , while the inverse exponential map Exp_{∇}^{-1} carries information on the affine connection ∇ and thus on the symmetric covariant tensor T . The exponential map Exp_{∇} provides a correspondence between tangent vectors at a point $m_{\text{in}} \in \mathcal{M}$ and points in \mathcal{M} . Essentially, given a tangent vector v_{in} , the image of the exponential map $Exp(v_{\text{in}})$ is the point $m_{\text{fin}} \in \mathcal{M}$ that is reached from m_{in} moving along the ∇ -geodesic $\gamma_{m_{\text{in}}, v_{\text{in}}}$ with initial velocity v_{in} when $t = 1$, that is, $m_{\text{fin}} = Exp_{\nabla}(v_{\text{in}}) = \gamma_{m_{\text{in}}, v_{\text{in}}}(1)$. The inverse Exp_{∇}^{-1} of this map

gives us a correspondence between a point $m_{\text{fin}} \in \mathcal{M}$ and a tangent vector v_{in} at $m_{\text{in}} \in \mathcal{M}$. Writing $X(m_{\text{in}}, m_{\text{fin}}) = \text{Exp}_{\nabla}^{-1}(m_{\text{fin}})$, the canonical contrast function S constructed in Ref. 12 reads

$$S(m_{\text{in}}, m_{\text{fin}}) := \int_0^1 g(X(\gamma(t), m_{\text{fin}}), \dot{\gamma}(t)) dt. \tag{80}$$

Following a similar line of reasoning, it is possible to construct another kind of potential function for (\mathcal{M}, g, T) using the metric Lagrangian \mathcal{L}_g associated with g and the inverse of the exponential map associated with the affine connection $\nabla \equiv \nabla_{\alpha=1}$. We recall that the affine connection $\nabla \equiv \nabla_{\alpha=1}$ is particularly relevant when dealing with families of exponential probability distributions since for them it turns to be flat (Ref. 6). The equations of motion for the ∇ -geodesics are

$$\ddot{q}^j(t) = -\Gamma_{kl}^j(q(t))\dot{q}^k(t)\dot{q}^l(t) = -\Gamma_{kl}^j(q(t))v^k(t)v^l(t). \tag{81}$$

Hence, a series expansion of $q^j(t)$ around $t = 0$ gives

$$q^j(t) = q_{\text{in}}^j + v_{\text{in}}^j t - \frac{t^2}{2} \Gamma_{kl}^j(q_{\text{in}})v_{\text{in}}^k v_{\text{in}}^l + O(\|tv_{\text{in}}\|^3), \tag{82}$$

where $q^j(0) = q_{\text{in}}$ and $\dot{q}^j(0) = v_{\text{in}}$, and the higher order terms are always a product of some functions of the $q^j(t)$ with the functions $v^j(t) = \dot{q}^j(t)$. Consequently, the exponential map $\text{Exp}_{\nabla}(v_{\text{in}})$ reads

$$q_{\text{fin}}^j = q^j(1) = q_{\text{in}}^j + v_{\text{in}}^j - \frac{1}{2} \Gamma_{kl}^j(q_{\text{in}})v_{\text{in}}^k v_{\text{in}}^l + O(\|v_{\text{in}}\|^3), \tag{83}$$

from which we immediately obtain

$$\left. \frac{\partial q_{\text{fin}}^j}{\partial q_{\text{in}}^k} \right|_{v_{\text{in}}=0} = \left. \frac{\partial q_{\text{fin}}^k}{\partial v_{\text{in}}^k} \right|_{v_{\text{in}}=0} = \delta_k^j, \tag{84}$$

$$\left. \frac{\partial^2 q_{\text{fin}}^j}{\partial q_{\text{in}}^k \partial q_{\text{in}}^l} \right|_{v_{\text{in}}=0} = \left. \frac{\partial^2 q_{\text{fin}}^j}{\partial v_{\text{in}}^k \partial q_{\text{in}}^l} \right|_{v_{\text{in}}=0} = 0 \quad \left. \frac{\partial^2 q_{\text{fin}}^j}{\partial v_{\text{in}}^k \partial v_{\text{in}}^l} \right|_{v_{\text{in}}=0} = -\Gamma_{kl}^j(q_{\text{in}}), \tag{85}$$

and

$$\left. \frac{\partial v_{\text{in}}^j}{\partial q_{\text{fin}}^k} \right|_d = - \left. \frac{\partial v_{\text{in}}^j}{\partial q_{\text{in}}^k} \right|_d = \delta_k^j, \tag{86}$$

$$\left. \frac{\partial^2 v_{\text{in}}^j}{\partial q_{\text{in}}^k \partial q_{\text{in}}^l} \right|_d = - \left. \frac{\partial^2 v_{\text{in}}^j}{\partial q_{\text{in}}^k \partial q_{\text{fin}}^l} \right|_d = \Gamma_{kl}^j(q_{\text{in}}), \tag{87}$$

where $|_d$ denotes the evaluation on the diagonal $q_{\text{in}} = q_{\text{fin}} = q$ of $\mathcal{M} \times \mathcal{M}$. At this point, we define

$$S(q_{\text{in}}, q_{\text{fin}}) := \frac{1}{2} g_{jk}(q_{\text{in}}) v_{\text{in}}^j(q_{\text{in}}, q_{\text{fin}}) v_{\text{in}}^k(q_{\text{in}}, q_{\text{fin}}), \tag{88}$$

where the $v_{\text{in}}^j(q_{\text{in}}, q_{\text{fin}})$ are determined by the inverse of the exponential map. Then, a careful application of the chain rule and relations (86) and (87) to Eq. (88) gives

$$\left. \frac{\partial^2 S}{\partial q_{\text{fin}}^k \partial q_{\text{fin}}^j} \right|_d = \left. \frac{\partial^2 S}{\partial q_{\text{in}}^k \partial q_{\text{in}}^j} \right|_d = - \left. \frac{\partial^2 S}{\partial q_{\text{in}}^k \partial q_{\text{fin}}^j} \right|_d = g_{jk}(q), \tag{89}$$

and

$$\left. \frac{\partial^3 S}{\partial q_{\text{fin}}^l \partial q_{\text{fin}}^k \partial q_{\text{in}}^j} \right|_d = -g \Gamma_{jkl} + T_{jkl}, \tag{90}$$

from which it follows that S is a potential function for (\mathcal{M}, g, T) as claimed.

As it is clear, the non-uniqueness of the potential (contrast) function S implies that the ‘‘inverse problem’’ has many alternative solutions, and all solutions are to be considered permissible. However, the Hamilton-Jacobi approach to potential functions outlined above has the advantage to clearly point out the mathematical regularity conditions needed to consider the problem well-posed from a not strictly local point of view. Let us indulge a little on this subject. On the one hand, the algorithm

constructed in Ref. 12, as well as the one given above, heavily depends on the existence, uniqueness, and differentiability properties of the exponential map Exp_∇ and of its inverse Exp_∇^{-1} . These are strong assumptions that, from a global point of view, must be checked using a case-by-case analysis. Furthermore, all the regularity requirements are relative to the affine connection ∇ , and thus the geometrical information encoded in g and T seems to be uncorrelated, which is in contrast with the fact that the geometrical structure of a statistical manifold (\mathcal{M}, g, T) considers g and T on the same footing. On the other hand, the Hamilton-Jacobi approach to potential functions completely depends on a complete solution to the Hamilton-Jacobi problem for the Lagrangian \mathcal{L} , which is a well-known problem. The mathematical requirement for its well-posedness is the complete integrability (Refs. 15 and 16) of the dynamical vector field associated with the Lagrangian \mathcal{L} . Consequently, once we write down the Lagrangian \mathcal{L} , we are immediately able to pose the problem in a mathematically rigorous way, even if its explicit solution could be very hard to find. Furthermore, the Lagrangian \mathcal{L} contains the information about g and T together, which means that it contains all the information on the geometrical structure of (\mathcal{M}, g, T) as a whole. Having a Lagrangian, it would be possible to use the tools of symmetries and constants of the motion characteristic of Lagrangian mechanics to better understand the symmetry properties of the potential function S_α .

IV. CONCLUSIONS AND OUTLOOKS

We have shown that a solution of the Hamilton-Jacobi problem for a Lagrangian \mathcal{L}_α defined in terms of g and T is a potential function for the statistical manifold (\mathcal{M}, g, T) . This dynamical perspective naturally leads to new questions, and, furthermore, paves the way to an interesting interchange of tools and methods between information geometry and the theory of dynamical systems. For instance, the “unfolding-reduction” attitude towards dynamical systems clearly illustrated in 19 could be a powerful technique in the search of potential functions. Let us briefly comment on this point.

Let us consider a two dimensional sphere embedded into \mathbb{R}^3 through the map $i_{\mathbb{S}^2} : \mathbb{S}^2 \rightarrow \mathbb{R}^3$. A local expression of this map is given by

$$\begin{cases} x^1 = \sin \theta \cos \phi \\ x^2 = \sin \theta \sin \phi \\ x^3 = \cos \theta \end{cases}, \quad (91)$$

where $\theta \in]0, \pi[$ and $\phi \in]0, 2\pi[$. By means of this immersion, it is possible to pull-back covariant tensors on \mathbb{R}^3 to \mathbb{S}^2 .

Let us consider the following statistical manifold: \mathbb{R}^3 equipped with the Euclidean metric $g = \delta_{jk} dx^j \otimes dx^k$ and the skewness tensor $T = dx^1 \otimes dx^1 \otimes dx^1 + dx^2 \otimes dx^2 \otimes dx^2 + dx^3 \otimes dx^3 \otimes dx^3$. According to the prescription outlined in this paper, a canonical potential function is

$$S(x_{\text{in}}, x_{\text{fin}}) = \delta_{jk} (x_{\text{fin}}^j - x_{\text{in}}^j)(x_{\text{fin}}^k - x_{\text{in}}^k) + \frac{\alpha}{6} \left((x_{\text{fin}}^1 - x_{\text{in}}^1)^3 + (x_{\text{fin}}^2 - x_{\text{in}}^2)^3 + (x_{\text{fin}}^3 - x_{\text{in}}^3)^3 \right).$$

By means of the previous immersion one can pull back this potential to \mathbb{S}^2 obtaining the following function:

$$\begin{aligned} S_{\mathbb{S}^2}((\theta_0, \phi_0), (\theta_1, \phi_1)) &= \frac{1}{2} (\sin \theta_0 \sin \theta_1 \cos(\phi_1 - \phi_0) + \cos \theta_0 \cos \theta_1) + \\ &+ \frac{\alpha}{6} \left((\sin \theta_0 \cos \phi_0 - \sin \theta_1 \cos \phi_1)^3 + (\sin \theta_0 \sin \phi_0 - \sin \theta_1 \sin \phi_1)^3 + (\cos \theta_0 - \cos \theta_1)^3 \right). \end{aligned}$$

A direct computation shows that this is a potential function on the submanifold \mathbb{S}^2 and it generates a metric tensor $g_{\mathbb{S}^2}$ and a skewness tensor $T_{\mathbb{S}^2}$ which coincide with the pull-back to \mathbb{S}^2 of the metric and skewness tensors on \mathbb{R}^3 . Indeed

$$g_{\mathbb{S}^2} = d\theta \otimes d\theta + (\sin \theta)^2 d\phi \otimes d\phi, \quad (92)$$

$$\begin{aligned}
T_{\mathbb{S}^2} = & - \left(\cos^3 \phi \cos^3 \theta + \sin^3 \phi \cos^3 \theta - \sin^3 \theta \right) d\theta \otimes d\theta \otimes d\theta - \\
& - \cos^2 \theta \sin \theta \sin \phi \cos \phi (\cos \phi - \sin \phi) d\theta \otimes d\theta \otimes d\phi + \\
& + \sin^2 \theta \cos \theta \sin \phi \cos \phi (\cos \phi + \sin \phi) d\theta \otimes d\phi \otimes d\phi + \\
& + \sin^3 \theta (\sin^3 \phi - \cos^3 \phi) d\phi \otimes d\phi \otimes d\phi.
\end{aligned} \tag{93}$$

This simple example shows that in some cases, it is possible to obtain a tensor which is no more constant, the metric tensor on the sphere, starting from an Euclidean space, and the potential on the Euclidean space induces a potential on the submanifold. However one could also invert this procedure. If one starts from a manifold with a non-constant tensor, it is possible to enlarge this manifold to a larger space equipped with a constant metric tensor: this is the meaning of the word “unfolding” in such a context. These methods can be useful, for instance, in information geometry in relation with the description of curved exponential families, which are submanifolds of the statistical manifold of the exponential distribution. Another possible application is related to the Hamilton-Jacobi approach described in Secs. II and III. Indeed, we could enlarge the initial carrier space to a bigger space on which metric and skewness tensors are generated by a simple Lagrangian, simple in the sense of easily solvable. However, a better understanding of this situation can be achieved only by adopting an intrinsic language, and this is one of the possible developments we are working on.

Furthermore, the dynamical picture described in this contribution seems to suggest that the tangent bundle $T\mathcal{M}$ of the statistical manifold \mathcal{M} plays an active role in the research of a contrast function S for (\mathcal{M}, g, T) . Consequently, it is natural to ask for a more clear interpretation of the tangent vectors to a probability distribution.

The transition from the classical to the quantum setting is still to be fully worked out. There are different aspects that need to be completely understood. For instance, in the quantum setting, the manifold \mathcal{M} is the manifold of states of the system; hence, its points are no more probability distributions as in the classical case, they are probability amplitudes. Indeed, denoting with $\psi(x)$, the wave function associated with a quantum state, it is well known that the square modulus $|\psi|^2 \equiv p(x)$ of $\psi(x)$ is a genuine probability distribution. Accordingly, the wave function can be written as $\psi(x) = \sqrt{p(x)} e^{i\alpha(x)}$, and thus a phase term $e^{i\alpha(x)}$ arises. In Ref. 5, it is shown that this phase term enters into the definition of the Fubini-Study metric g , as well as in the definition of the symplectic form ω . This, in turn, calls for a deeper understanding of the phase term in relation with the geometric structure of the manifold of quantum states.

Another question is related to some results known in information geometry and described, for instance, in Refs. 6 and 24, i.e., it is possible to use well-studied relative-entropies as contrast functions on a statistical manifold (examples are the Shannon relative entropy or the Tsallis q -relative entropy). Since we have interpreted contrast (potential) functions as solutions of a Hamilton-Jacobi problem, it is reasonable to ask whether relative entropies are generating functions of canonical transformations, and what such a transformation would do. This could then lead to a formulation of thermodynamics as a dynamical theory, entropy providing the action functional. Attempts in such direction have already been done. For instance, Souriau in Ref. 25 described thermodynamical evolutions in terms of symplectic scattering processes in a relativistic framework. However a deeper analysis in such a direction is necessary.

It would be interesting to apply the Hamilton-Jacobi procedure to formulate an inverse problem for well-known divergence functions (relative entropies) such as the Kullback-Leibler divergence. Specifically, once a particular divergence function S is fixed, the problem we have in mind is to find a suitable Lagrangian \mathcal{Q} such that the associated Hamilton principal function I is precisely the divergence function S of the problem. A preliminary step in this direction has been made in Sec. III, where it is shown that, if the Kullback-Leibler divergence S_{KL} is the Hamilton principal function associated with a Lagrangian \mathcal{Q}_{KL} , then the Lagrangian \mathcal{Q}_α of Eq. (22) differs from the third order approximation of \mathcal{Q}_{KL} only by the constant factor $\mathcal{Q}_{KL}|_{v=0}$.

Finally, let us note that the Hamilton-Jacobi approach to potential functions introduced here makes use only of the geometric information contained in the metric tensor g and the skewness tensor T , and it is independent on whether or not the manifold is a statistical manifold equipped with the Fisher-Rao metric. This paves the way to a deeper analysis of the space of states of quantum

mechanics where these structures are also present but do not necessarily coincide with the “classical” ones that appear in the context of statistical manifolds.

ACKNOWLEDGMENTS

We thank the anonymous referee for the valuable comments. G. Marmo would like to acknowledge the support provided by the Banco de Santander-UCIIM “Chairs of Excellence” Programme 2016-2017. J.M.P.-P. is partly supported by the Spanish MINECO Grant Nos. MTM2014-54692-P and QUITEMAD+, S2013/ICE-2801.

- ¹ I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).
- ² A. Ashtekar and T. A. Schilling, “Geometrical formulation of quantum mechanics,” in *On Einstein’s Path: Essays in Honor of Engelbert Schucking* (Springer, New York, 1999), pp. 23–65.
- ³ E. Ercolessi, G. Marmo, and G. Morandi, “From the equations of motion to the canonical commutation relations,” *Riv. Nuovo Cimento Soc. Ital. Fis.* **33**, 401–590 (2010).
- ⁴ J. P. Provost and G. Vallee, “Riemannian structure on manifolds of quantum states,” *Commun. Math. Phys.* **76**, 289–301 (1980).
- ⁵ P. Facchi, R. Kulkarni, V. I. Man’ko, G. Marmo, E. C. G. Sudarshan, and F. Ventriglia, “Classical and quantum fisher information in the geometrical formulation of quantum mechanics,” *Phys. Lett. A* **374**(48), 4801–4803 (2010).
- ⁶ S. I. Amari and H. Nagaoka, *Methods of Information Geometry* (American Mathematical Society, Providence, Rhode Island, 2000).
- ⁷ J. Zhang and F. Li, “Symplectic and Kähler structures on statistical manifolds induced from divergence functions,” in *Geometric Science of Information* (Springer-Verlag, Berlin, Heidelberg, 2013), pp. 595–603.
- ⁸ S. I. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen, and C. R. Rao, *Differential Geometry in Statistical Inference* (Institute of Mathematical Statistics, Hayward, California, 1987).
- ⁹ N. Ay and W. Tuschmann, “Dually flat manifolds and global information geometry,” *Open Syst. Inf. Dyn.* **9**(2), 195–200 (2002).
- ¹⁰ N. Ay and W. Tuschmann, “Duality versus dual flatness in quantum information geometry,” *J. Math. Phys.* **44**(4), 1512–1518 (2003).
- ¹¹ T. Matumoto, “Any statistical manifold has a contrast function: On the c^3 -functions taking the minimum at the diagonal of the product manifold,” *Hiroshima Math. J.* **23**(2), 327–332 (1993).
- ¹² N. Ay and S. I. Amari, “A novel approach to canonical divergences within information geometry,” *Entropy* **17**(12), 8111–8129 (2015).
- ¹³ D. Petz and C. Sudár, “Monotone metrics on matrix spaces,” *Linear Algebra Appl.* **244**, 81–96 (1996).
- ¹⁴ D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer, Berlin, Heidelberg, 2007).
- ¹⁵ G. Marmo, G. Morandi, and N. Mukunda, “A geometrical approach to the Hamilton-Jacobi form of dynamics and its generalizations,” *Riv. Nuovo Cimento Soc. Ital. Fis.* **13**, 1–74 (1990).
- ¹⁶ V. I. Arnol’d, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978).
- ¹⁷ R. Abraham and J. E. Marsden, *Foundations of Mechanics*, 2nd ed. (Addison-Wesley Publishing Co., Inc., Redwood City, CA, 1978).
- ¹⁸ C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto Press, Toronto, 1952).
- ¹⁹ J. F. Cariñena, A. Ibort, G. Marmo, and G. Morandi, *Geometry from Dynamics, Classical and Quantum* (Springer, Berlin, 2015).
- ²⁰ G. Morandi, C. Ferrario, G. Lo Vecchio, G. Marmo, and C. Rubano, “The inverse problem in the calculus of variations and the geometry of the tangent bundle,” *Phys. Rep.* **188**, 147–284 (1990).
- ²¹ W. K. Wootters, “Statistical distance and Hilbert space,” *Phys. Rev. D* **23**(2), 357 (1981).
- ²² V. I. Man’ko, G. Marmo, F. Ventriglia, and P. Vitale, “Metric on the space of quantum states from relative entropy. Tomographic reconstruction,” e-print [arXiv:1612.07986](https://arxiv.org/abs/1612.07986) [quant-ph].
- ²³ J. F. Cariñena, X. Gràcia, G. Marmo, E. Martínez, M. C. Muñoz Lecanda, and N. Román-Roy, “Geometric Hamilton-Jacobi theory,” *Int. J. Geom. Methods Modern Phys.* **03**(07), 1417–1458 (2006).
- ²⁴ S. I. Amari, *Information Geometry and its Application* (Springer, Japan, 2016).
- ²⁵ J. M. Souriau, “Thermodynamique et Geometrie,” in *Differential Geometrical Methods in Mathematical Physics II: Proceedings, University of Bonn, July 13–16, 1977* (Springer-Verlag, Berlin, Heidelberg, 1978), pp. 369–397.
- ²⁶ These are metric tensors satisfying the so-called monotonicity property, i.e., the scalar product they induce on tangent vectors does not increase under the action of completely positive trace-preserving (CPTP) maps.
- ²⁷ Note that, at this level, formulas (33) and (34) are valid for every dynamical curve γ .