

Functional information geometry of Euclidean quantum fields

Stefan Floerchinger^{*}

Theoretisch-Physikalisches Institut, Max-Wien Platz 1, 07743 Jena, Germany



(Received 16 July 2024; accepted 2 December 2024; published 26 December 2024)

Information geometry provides differential geometric concepts like a Riemannian metric, connections and covariant derivatives on spaces of probability distributions. We discuss here how these concepts extend in a functional sense to quantum field theories in the Euclidean domain which can also be seen as statistical field theories. The geometry has a dual affine structure corresponding to source fields or expectation value fields seen as coordinates. In the latter version the coordinates label the macrostates of the classical field theory. A key concept is a new generating functional, which is a functional generalization of the Kullback-Leibler divergence. From its functional derivatives one can obtain connected as well as one-particle irreducible correlation functions. It also encodes directly the geometric structure, i.e. the functional Fisher information metric and the two dual connections, and it determines asymptotic probabilities for field configurations through Sanov's theorem. Based on the two dual connections one can construct covariant functional derivatives which allow one to calculate connected and one-particle irreducible correlation functions in general functional coordinate systems.

DOI: [10.1103/PhysRevD.110.125027](https://doi.org/10.1103/PhysRevD.110.125027)

I. INTRODUCTION

Information-theoretic aspects of quantum field theory are getting more important recently, partly triggered by the advancement of quantum information theory. Traditionally this has been discussed in the context of black holes and the information paradox [1–4], but also in the context of nonequilibrium dynamics and thermalization [5,6] or for expanding quantum fields [7–10]. Information-theoretic methods can be used to classify states, for example, in terms of entropy or entanglement entropy, to compare states in terms of relative entropies, etc.

It is interesting in that context that there is actually a sophisticated geometric structure on the spaces of probability distributions. This has been developed mathematically in the field of information geometry, see Refs. [11,12] for recent reviews. More specifically, the Fisher information metric provides a natural Riemannian metric to induce a distance on the space of probabilities and Chentsov's theorem ensures that it is unique in being invariant under sufficient statistics. Beyond this, a pair of dual connections is available in the form of the Amari-Chentsov structure [11,12]. They differ from the Levi-Civita connection and are not metric compatible or, in other words, they have

nonmetricity. Even more, the entire geometric structure is very nicely encoded in terms of the Kullback-Leibler divergence.

On the one side this raises the question whether a similar geometric structure exists on the space of quantum field theoretic states. Such a structure could be highly interesting to better understand nonequilibrium dynamics, for example. One generalization of the Fisher metric is the Kubo-Mori-Bogoliubov metric, which is similarly induced by the quantum relative entropy. However, there are other possible metrics on the space of quantum states, and a uniqueness result as strong as the Chentsov theorem in the classical case is not available, for reviews see [13,14].

Relative entropy and the Fisher information metric have been discussed as an interesting possibility to distinguish between different quantum field theories [15]. Previously information-theoretic metrics have been considered in a finite-dimensional space of coupling constants, for example, in Refs. [16–19]. In this context, infinite information-theoretic distances have been discussed in Refs. [20–22] with a view on applications in the context of the Swampland program, see Refs. [23,24] for reviews. In Ref. [25], Erdmenger *et al.* explored the use of the Fisher information metric in different model systems. In Ref. [26] they studied relative entropy and the connections of neural networks with the renormalization group, see also Ref. [27].

An information-theoretic view on the renormalization group has actually been being developed for many years, see Refs. [28–39]. Recently also connections between the renormalization group and optimal transport have been explored [40,41].

^{*}Contact author: stefan.floerchinger@uni-jena.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

The quantum Fisher information has also been employed in the context of entanglement detection, see e.g. Refs. [42–44] and references therein. It was also discussed in the context of holography, see e.g. Refs. [45,46]. In Refs. [47,48] an emergent metric in large- N scalar theories has been discussed. Similar concepts have also been used to quantify complexity [49–51]. Let us mention here that somewhat different information-theoretic concepts (based on the Wasserstein metric) have also been applied in particle physics [52,53].

Geometric concepts also find applications in thermodynamics, see [54,55] for reviews or to characterize phase transitions [56]. Quantum states have a geometric characterization as well, see, for example, Ref. [14] for an introduction. Relative entropy has recently been used to formulate entropic uncertainty relations for quantum field theories [57].

Previous applications of information geometry to Euclidean quantum field theory have concentrated on the space of theories that can be parametrized by a finite (countably infinite) set of coupling constants. In contrast we develop in the present work a different application. We consider a parametrization of actions in terms of local *fields* and take those as coordinates in a functional sense. We work with Euclidean quantum field theories that directly have a probability interpretation or that can be seen as classical statistical field theories with one more spatial dimension. In one representation the parameter fields are sources that couple to fundamental fields or composite operators and could in the latter case also be seen as space-dependent masses or coupling constants. In another representation, related to the former through a Legendre transform, the information geometric coordinates are expectation value fields. The beauty of geometric concepts is that they apply also in other coordinate systems, however.

In other words, in the present work we discuss the information geometry of statistical field theories which can also be seen as quantum field theories in the Euclidean domain. We explore specifically how source fields or expectation value fields, which parametrize the macrostates of these statistical theories, can be seen as alternative coordinates in the space of functional probability distributions and what kind of geometric structures are associated with them.

In terms of source or expectation value coordinates, the Riemannian Fisher information metric will be seen to correspond to connected two-point correlation functions. The two dual connections are corresponding to different kinds of three-point correlation functions (connected and one-particle irreducible).

We will find very close relations between functionals like the Schwinger functional and the quantum effective action that are used in the quantum field theoretic context to generate correlation functions, effective propagators and vertices, or to investigate thermodynamic properties and information geometric concepts. We will introduce a

functional generalization of the Kullback-Leibler divergence and show that it can itself be used to generate connected correlation functions as well as one-particle irreducible correlation functions, effective vertices, propagators and so on.

This leads to a new, information-theoretic perspective on important concepts of nonperturbative statistical and quantum field theory. In addition also a new geometric view is thereby obtained, that allows one, for example, to go to arbitrary field coordinates but to calculate nevertheless connected or one-particle irreducible correlation functions through the use of appropriate connections and covariant derivatives. Specifically in this regard the newly introduced divergence functional has substantial advantages over the traditionally used Schwinger functional and quantum effective action. When evaluated in arbitrary coordinates it automatically contains the geometric information about the Riemannian metric as well as a set of two dual connections. This information can be used to formulate covariant functional derivatives and to calculate the correlation functions, again from the divergence functional. In this sense the divergence functional is a generating functional for the functional information geometry but also for connected as well as one-particle irreducible correlation functions.

The fact that we concentrate on Euclidean theories with probabilistic interpretation has the advantage that concepts can be taken over from the mathematical literature on classical information geometry rather directly, with the generalization being from functions to functionals.¹

Let us remark that the discussion we present here is from a mathematical point of view heuristic. For example, issues with different topologies or the precise definition of the functional integral will not be discussed. We leave in particular the issue of ultraviolet regularization and renormalization to future work.

This paper is organized as follows. In Sec. II we briefly recall the conceptual setup of Euclidean quantum field theory before discussing a natural affine structure in terms of sources in Sec. III. In Sec. IV we introduce the Fisher information metric and discuss its significance in the field theoretic context. In Sec. V we introduce expectation value fields as alternative coordinates using Legendre transforms and the quantum effective action. The corresponding affine geometry is discussed in Sec. VI. In Sec. VII we discuss a new generating functional for quantum field theory, the functional generalization of the Kullback-Leibler divergence in terms of source coordinates, expectation value coordinates and combinations thereof. Also the information-theoretic significance of this object will be discussed. Subsequently, in Sec. VIII we present functional integral representations of the divergence functional including

¹Note that some of the mathematical literature is formulated based on powerful measure theoretic concepts that also apply formally to functionals [12].

steepest descent or one-loop approximations to it. In Sec. IX we discuss different connections in more detail, before explaining how they can be obtained from the divergence functional in general coordinates in Sec. X. We also introduce there covariant functional derivatives based on these divergences. Finally, we draw some conclusions in Sec. XII.

II. EUCLIDEAN QUANTUM FIELD THEORY

We study here a bosonic quantum field theory in Euclidean spacetime, or statistical field theory. The fluctuating field $\chi_l(x)$ is seen here as a very large set of random variables. In a discretized version of the theory there would be 1 degree of freedom or random variable per field component l and lattice site x_j , but in the field theoretic limits of infinite volume and vanishing lattice constant the number tends to infinity.

We assume here that there is a well-defined functional integral measure $D\chi$, which could, for example, be obtained through the limiting process from the lattice formulation of the theory. With respect to this measure the probability density for the field χ is taken to be

$$p[\chi, J] = \exp \left(-I[\chi] + \int_x \sum_n \{J_n(x) \phi_n[\chi](x)\} - W[J] \right). \quad (1)$$

We have introduced here a set of parameter fields $J_n(x)$ as generalized sources, which are supposed to couple in some way to the microscopic field χ through operators $\phi_n[\chi](x)$. Note that this is a very general setting, because the $\phi_n[\chi](x)$ could be linear or nonlinear in χ . In other words, the sources $J_n(x)$ can couple to fundamental fields but also composite operators of various types. Some sources $J_n(x)$ can also couple to the theory like coupling constants λ or changes in coupling constants $\lambda - \lambda_0$. Of course, coupling constants are usually constant in space, but that just means to restrict $J_n(x)$ accordingly. The functional $I[\chi]$ parametrizes the functional probability density at the point $J_n(x) = 0$, up to an overall constant specified through $W[0]$.

In general the Schwinger functional $W[J]$ is determined such that the probability distribution is properly normalized, as the logarithm of the partition function

$$Z[J] = e^{W[J]} = \int D\chi \exp \left(-I[\chi] + \int_x \sum_n \{J_n(x) \phi_n[\chi](x)\} \right). \quad (2)$$

From the partition function $Z[J]$, or the Schwinger functional $W[J]$ (known as cumulant generating function in statistics), one easily derives expectation values, correlation functions or other interesting observables. One can easily show, using Hölder's inequality, that $W[J]$ is convex,

$$W[(1-t)J' + tJ''] \leq (1-t)W[J'] + tW[J''], \quad (3)$$

with $0 \leq t \leq 1$.

The positions x are supposed to be in a space of dimension d and we use the abbreviation

$$\int_x = \int d^d x \sqrt{g(x)}, \quad (4)$$

with $g(x) = \det(g_{\mu\nu}(x))$ the determinant of the Euclidean spacetime metric. In the following, it will also be convenient to use an abstract index which combines position and field component, $\alpha = (x, n)$, and to use the convention that a combination of an upper and lower abstract index corresponds to an integral over x and a sum over n ,

$$J^\alpha \phi_\alpha = \int d^d x \sqrt{g(x)} \sum_n J_n(x) \phi_n(x), \quad (5)$$

with $J^\alpha = J_n(x)$ and $\phi_\alpha = \phi_n(x)$. In other words, whenever upper and lower Greek indices appear together this implies an integration over position and a summation over discrete field indices. On the other side, a formal Kronecker δ stands for a functional unit matrix,

$$\delta^\alpha_\beta = \frac{1}{\sqrt{g(x)}} \delta^{(d)}(x-y) \delta_{nm}. \quad (6)$$

When convenient one can also use another basis and expand fields in Fourier modes, for example. The abstract index notation needs a moment to get used to, but it is very convenient for a calculus involving many functional differentiations, higher-order correlation functions, etc. It is usually clear from the context how expressions involving abstract indices can be made concrete. Let us stress that we are still dealing with fields and functionals, despite the simplified notation.

A class of probability distributions as in Eq. (1) with a set of sources J^α seen as parameters is known as an exponential family in information geometry [11,12]. We stress that usually the set of coordinates J^α is finite, while we are working here in a functional setup with the notation $J^\alpha = J_n(x)$. Our exponential family has infinitely many parameters.

Let us note here that one may alternatively also consider the operators $\phi_n(x)$ as the field theoretic random variables and perform a change of variables

$$D\chi \exp(-I[\chi]) = D\phi \exp(-S[\phi]). \quad (7)$$

The functional $S[\phi]$ is known in this context as the Euclidean action. The partition function (2) is then represented as

$$Z[J] = e^{W[J]} \\ = \int D\phi \exp \left(-S[\phi] + \int_x \sum_n \{J_n(x) \phi_n(x)\} \right). \quad (8)$$

This notation seems to imply a restriction in the sense that $J_n(x)$ now couples linearly to the fields. Moreover, one sometimes sees sources as auxiliary constructs used to calculate field correlation functions. However, as the discussion above shows, one can take another point of view and parametrize through $J_n(x)$ a general class of deformations of the theory, which is allowed to depend on space position. The only real restriction is that $J_n(x)$ enters linearly any term in the exponential function in Eq. (2) that depends on the field χ .

III. AFFINE GEOMETRY FOR SOURCES

Let us note here immediately that there is an affine structure on this exponential family in the sense that an affine transformation of sources,

$$J^\alpha \rightarrow J'^\alpha = M^\alpha_\beta J^\beta + c^\alpha, \quad (9)$$

with invertible M^α_β , leads again to a probability distribution of the form (1), i.e. in the exponential family. Note that this is not the case for nonlinear transformation of sources.

In a related way, affine transformations of the form (9) preserve the convexity of the Schwinger functional $W[J]$, while more general nonlinear transformations do not.

Moreover, the affine structure allows connecting probability distributions associated with the sources J'^α and J''^α through a distribution with source $J^\alpha(t)$ given by the flat, so-called e -geodesic,²

$$J^\alpha(t) = (1-t)J'^\alpha + tJ''^\alpha, \quad (10)$$

where $0 \leq t \leq 1$. Geometrically this means that the manifold of probability distributions in the exponential family has a flat connection in terms of the coordinates J^α . As trajectories, e -geodesics are characterized by the differential equation

$$\frac{d^2}{dt^2} J^\alpha(t) + (\Gamma_E)_\beta^\alpha{}_\gamma [J] \frac{d}{dt} J^\beta(t) \frac{d}{dt} J^\gamma(t) = 0, \quad (11)$$

where the connection vanishes here, in terms of source coordinates,

$$(\Gamma_E)_\beta^\alpha{}_\gamma [J] = 0. \quad (12)$$

²A geodesic is here a line connecting two points, determined by a connection, and not necessarily the shortest path in any sense.

In another coordinate system, related to J^α in a nonlinear way, this would not be the case, however.

IV. THE FISHER INFORMATION METRIC

Spaces of probability distributions have a natural Riemannian metric, the Fisher information metric. For distributions parametrized by coordinates J^α it is given by

$$G_{\alpha\beta}[J] = \int D\chi p[\chi, J] \frac{\delta}{\delta J^\alpha} \ln p[\chi, J] \frac{\delta}{\delta J^\beta} \ln p[\chi, J] \\ = - \int D\chi p[\chi, J] \frac{\delta^2}{\delta J^\alpha \delta J^\beta} \ln p[\chi, J]. \quad (13)$$

In the last equation we have used the product rule and that $\int D\chi p[\chi, J] = 1$.

The information-theoretic significance of the Fisher metric is that the infinitesimal Fisher-Rao distance of two nearby probability distributions

$$ds^2 = G_{\alpha\beta}[J] dJ^\alpha dJ^\beta \quad (14)$$

gives a measure for how well distributions at J and J' can be distinguished [58]. As usual, in Riemannian geometry, the length of a path is defined through line integrals of ds .

For the exponential family, one finds easily

$$G_{\alpha\beta}[J] = \frac{\delta^2}{\delta J^\alpha \delta J^\beta} W[J] \\ = \langle \phi_\alpha[\chi] \phi_\beta[\chi] \rangle - \langle \phi_\alpha[\chi] \rangle \langle \phi_\beta[\chi] \rangle. \quad (15)$$

The Fisher metric agrees with the connected correlation function of fields. With the exception of Gaussian theories, this correlation function depends still on the source J^α and the metric is therefore not constant. Let us note that the metric (15) can be seen as an extension of Zamolodchikov's metric defined for conformal field theories [16], see [15,20] for further discussion.

We are now in the interesting situation that the space of probability distributions as parametrized by J^α has a nontrivial Riemannian metric $G_{\alpha\beta}[J]$, but at the same time has an affine structure with e -geodesics corresponding to vanishing connection, $(\Gamma_E)_\beta^\alpha{}_\gamma [J] = 0$. It is clear that this connection is, for non-Gaussian theories, not the Levi-Civita connection corresponding to the metric $G_{\alpha\beta}[J]$, and it remains to characterize it further. By a Gaussian theory we mean in this context a theory where the Schwinger functional $W[J]$ is quadratic in J such that $G_{\alpha\beta}$ is independent of J . In that special case the Levi-Civita connection symbols would vanish as well.

Let us note here that for statistical field theories in $d \geq 2$ dimensions two-point correlation functions $G(x-y) = \langle \phi(x) \phi(y) \rangle - \langle \phi(x) \rangle \langle \phi(y) \rangle$ are usually divergent for $x = y$ (ultraviolet divergence) and should be considered as Green's functions which also depend on boundary

conditions [59]. This is a point where the present setup deviates from applications of information geometry to finite systems.

V. EXPECTATION VALUES AS COORDINATES

Instead of the sources J^α one can parametrize the probability distributions for χ or the fields $\phi_\alpha[\chi]$ in terms of the expectation values

$$\Phi_\alpha = \langle \phi_\alpha[\chi] \rangle = \frac{\delta}{\delta J^\alpha} W[J] = \int D\chi p[\chi, J] \phi_\alpha[\chi]. \quad (16)$$

With the exception of Gaussian theories this is a nonlinear change of coordinates from the sources J^α .

One way to describe this change of variables is in terms of Legendre transforms. The quantum effective action or one-particle irreducible effective action is defined as

$$\Gamma[\Phi] = \sup_J (J^\alpha \Phi_\alpha - W[J]), \quad (17)$$

and as a Legendre transform it is a convex functional of expectation values Φ_α . The definition also implies that $\Gamma^{(2)}$ and $W^{(2)}$ are inverse,

$$\frac{\delta^2 \Gamma[\Phi]}{\delta \Phi_\alpha \delta \Phi_\beta} \frac{\delta^2 W[J]}{\delta J^\beta \delta J^\gamma} = \frac{\delta^2 \Gamma[\Phi]}{\delta \Phi_\alpha \delta \Phi_\beta} G_{\beta\gamma}[J] = \delta^\alpha_\gamma. \quad (18)$$

In the context of the theory of large deviations this is known as the Cramér function [60]. The quantum effective action satisfies the field equation

$$\frac{\delta}{\delta \Phi_\alpha} \Gamma[\Phi] = J^\alpha. \quad (19)$$

With these relations one can write the probability density for χ as

$$p[\chi, \Phi] = \exp \left(-I[\chi] + \frac{\delta \Gamma[\Phi]}{\delta \Phi_\alpha} (\phi_\alpha[\chi] - \Phi_\alpha) + \Gamma[\Phi] \right). \quad (20)$$

The normalization condition for the distribution (20), $\int D\chi p[\chi, \Phi] = 1$, gives the well-known background identity for $\Gamma[\Phi]$.

Let us mention here also a relation to a differential entropy for the functional probability distribution in the form (1) or (20). Defining a differential entropy as

$$H = - \int D\chi p[\chi, J] \ln p[\chi, J], \quad (21)$$

we find

$$H = \langle I[\chi] \rangle - J^\alpha \Phi_\alpha + W[J] = \langle I[\chi] \rangle - \Gamma[\Phi]. \quad (22)$$

Note that as a differential entropy H does not have a definite sign.

Somewhat surprisingly one can actually argue under certain conditions that $p[\chi, \Phi]$ must be linear in Φ_α , see e.g. Ref. [12]. To investigate this in more detail we first define Φ_α^{eq} as the expectation value at vanishing source, i.e. the solution of Eq. (19) at $J^\alpha = 0$. We can rewrite (16) in the form

$$\int D\chi (p[\chi, \Phi] - p[\chi, \Phi^{\text{eq}}]) (\phi_\alpha[\chi] - \Phi_\alpha^{\text{eq}}) = \Phi_\alpha - \Phi_\alpha^{\text{eq}}. \quad (23)$$

We now assume that one can find a dual basis $\psi^\beta[\chi]$ to the functionals $(\phi_\alpha[\chi] - \Phi_\alpha^{\text{eq}})$ such that

$$\int D\chi (\phi_\alpha[\chi] - \Phi_\alpha^{\text{eq}}) \psi^\beta[\chi] = \delta^\beta_\alpha, \quad (24)$$

and that allows one to expand

$$p[\chi, \Phi] - p[\chi, \Phi^{\text{eq}}] = c_\beta[\Phi] \psi^\beta[\chi]. \quad (25)$$

Using this in Eq. (23) we find $c_\beta[\Phi] = \Phi_\beta - \Phi_\beta^{\text{eq}}$ or

$$p[\chi, \Phi] = p[\chi, \Phi^{\text{eq}}] + (\Phi_\alpha - \Phi_\alpha^{\text{eq}}) \psi^\alpha[\chi], \quad (26)$$

which shows indeed that $p[\chi, \Phi]$ is linear in Φ_α . Let us now compare this to Eq. (20). One can infer immediately that

$$p[\chi, \Phi^{\text{eq}}] = \exp(-I[\chi] - W[0]). \quad (27)$$

Expanding to linear order in $\Phi_\alpha - \Phi_\alpha^{\text{eq}}$ yields

$$p[\chi, \Phi] = p[\chi, \Phi^{\text{eq}}] + \frac{\delta^2 \Gamma[\Phi^{\text{eq}}]}{\delta \Phi_\alpha \delta \Phi_\beta} (\Phi_\alpha - \Phi_\alpha^{\text{eq}}) (\phi_\beta[\chi] - \Phi_\beta^{\text{eq}}) p[\chi, \Phi^{\text{eq}}]. \quad (28)$$

We see that consistency requires

$$\psi^\alpha[\chi] = \frac{\delta^2 \Gamma[\Phi^{\text{eq}}]}{\delta \Phi_\alpha \delta \Phi_\beta} (\phi_\beta[\chi] - \Phi_\beta^{\text{eq}}) p[\chi, \Phi^{\text{eq}}]. \quad (29)$$

Under the mentioned assumptions, (28) is not only the linear order of an expansion, but actually an exact expression. Higher orders of a formal expansion of Eq. (20) in $\Phi_\alpha - \Phi_\alpha^{\text{eq}}$ need to vanish. We stress, however, that original assumption, namely that a dual basis $\psi^\alpha[\chi]$ with the required properties exists, may well fail. In that case Eq. (20) remains correct, but does not have to be linear in Φ_α .

The class of probability distributions in Eq. (28), written as a linear combination of expectation values, is called a mixture family in the context of information geometry [11,12].

VI. AFFINE GEOMETRY FOR EXPECTATION VALUES

The Fisher metric in terms of expectation value coordinates is easily determined to be

$$G^{\alpha\beta}[\Phi] = - \int D\chi p[\chi, \Phi] \frac{\delta^2}{\delta\Phi_\alpha \delta\Phi_\beta} \ln p[\chi, \Phi] = \frac{\delta^2 \Gamma[\Phi]}{\delta\Phi_\alpha \delta\Phi_\beta}. \quad (30)$$

As a matrix, this is actually the inverse of $G_{\alpha\beta}[J]$, see Eq. (18). In a statistical field theory, the inverse propagator (30) contains usually derivative operators when evaluated in position space.

The Fisher-Rao distance between close by distributions can thus be written in three equivalent ways,

$$ds^2 = G_{\alpha\beta}[J] \delta J^\alpha \delta J^\beta = G^{\alpha\beta}[\Phi] \delta\Phi_\alpha \delta\Phi_\beta = \delta J^\alpha \delta\Phi_\beta. \quad (31)$$

As discussed before, the expectation values Φ_α are related to the sources J^α by a change of variables that is in general nonlinear. In this sense it is surprising that there is an affine structure here, as well. Indeed, transformations of the form

$$\Phi_\alpha \rightarrow \Phi'_\alpha = N_\alpha{}^\beta \Phi_\beta + d_\alpha, \quad (32)$$

with invertible $N_\alpha{}^\beta$, map elements of the mixture family to elements of the mixture family. One may consider so-called m -geodesics connecting two probability distributions with expectation values Φ'_α and Φ''_α of the form

$$\Phi_\alpha(t) = (1-t)\Phi'_\alpha + t\Phi''_\alpha, \quad (33)$$

and the corresponding probability distributions for $0 \leq t \leq 1$ are simply linear superpositions of the distributions at the two end points of the geodesic.

Note that the m -geodesic in (33) is different from the e -geodesic described in (10), even when the end points J'^α and Φ'_α as well as J''^α and Φ''_α correspond to the same probability distributions (see also below).

The m -geodesics can also be characterized in terms of differential equations,

$$\frac{d^2}{dt^2} \Phi_\alpha(t) + (\Gamma_M)^\beta{}_\alpha{}^\gamma[\Phi] \frac{d}{dt} \Phi_\beta(t) \frac{d}{dt} \Phi_\gamma(t) = 0. \quad (34)$$

In terms of expectation values as coordinates the m -connection symbols vanish,

$$(\Gamma_M)^\beta{}_\alpha{}^\gamma[\Phi] = 0. \quad (35)$$

At this point it is instructive to work out the connection symbols of the m -connection in terms of source coordinates,

$$(\Gamma_M)^\beta{}_\alpha{}^\gamma[J] = \frac{\delta\Phi_\alpha}{\delta J^\rho} \frac{\delta\Phi_\gamma}{\delta J^\nu} \frac{\delta J^\beta}{\delta\Phi_\mu} (\Gamma_M)^\rho{}_\mu{}^\nu[\Phi] + \frac{\delta J^\beta}{\delta\Phi_\mu} \frac{\delta^2\Phi_\mu}{\delta J^\alpha \delta J^\gamma}. \quad (36)$$

This uses the general transformation law for a connection under changes of coordinates. As we just argued, in expectation value coordinates the m -connection symbol vanishes, and we thus find in source coordinates

$$(\Gamma_M)^\beta{}_\alpha{}^\gamma[J] = \frac{\delta J^\beta}{\delta\Phi_\mu} \frac{\delta^2\Phi_\mu}{\delta J^\alpha \delta J^\gamma} = G^{\beta\mu} \frac{\delta^3 W[J]}{\delta J^\alpha \delta J^\gamma \delta J^\mu}. \quad (37)$$

Up to a Fisher metric, the m -connection symbol in source coordinates is the connected three-point function. Accordingly, m -geodesics are in source coordinates not straight lines when this three-point function is nonvanishing.

Similarly, one can find the e -connection symbol in expectation value coordinates,

$$\begin{aligned} (\Gamma_E)^\alpha{}_\beta{}^\gamma[\Phi] &= \frac{\delta J^\alpha}{\delta\Phi_\rho} \frac{\delta J^\gamma}{\delta\Phi_\nu} \frac{\delta\Phi_\beta}{\delta J^\mu} (\Gamma_E)^\rho{}_\mu{}^\nu[J] + \frac{\delta\Phi_\beta}{\delta J^\mu} \frac{\delta^2 J^\mu}{\delta\Phi_\alpha \delta\Phi_\gamma} \\ &= G_{\beta\mu} \frac{\delta^3 \Gamma[\Phi]}{\delta\Phi_\alpha \delta\Phi_\beta \delta\Phi_\mu}. \end{aligned} \quad (38)$$

This is essentially the one-particle irreducible three-point function.

VII. DIVERGENCE FUNCTIONAL

A. Divergence functional in source coordinates

Another highly interesting quantity used in information geometry is the Kullback-Leibler divergence or relative information entropy between two probability distributions with coordinates J^α and J'^α ,

$$D[J||J'] = \int D\chi p[\chi, J] \ln(p[\chi, J]/p[\chi, J']). \quad (39)$$

Note the antisymmetric nature of the definition. The Kullback-Leibler divergence is non-negative,

$$D[J||J'] \geq 0, \quad (40)$$

and it vanishes when the distributions $p[\chi, J]$ and $p[\chi, J']$ agree, i.e. for $J^\alpha = J'^\alpha$. Moreover, for $J'^\alpha = J^\alpha + \delta J^\alpha$ one has

$$D[J||J'] = \frac{1}{2} G_{\alpha\beta}[J] \delta J^\alpha \delta J^\beta + \dots, \quad (41)$$

where the ellipses are for terms of cubic and higher order in δJ^α . In other words, for nearby distributions the Kullback-Leibler divergence equals the Fisher-Rao distance squared, up to a factor.

The relative entropy finds an interesting application in terms of Sanov's theorem [58,60]. Consider n random realizations of field configurations χ_k (where $k = 1, \dots, n$) taken from the distribution (1). This gives rise to an "empirical" distribution

$$p_n[\chi] = \frac{1}{n} \sum_{k=1}^n \delta[\chi - \chi_k]. \quad (42)$$

The probability w for this empirical distribution $p_n[\chi]$ to lie in some region Ω of the space of all possible distributions $p[\chi]$ is asymptotically (for large n) constrained by Sanov's theorem. For simplicity let us assume that Ω is closed (with respect to weak topology). Then w is asymptotically constrained by the element $q[\chi]$ of Ω that is closest to $p[\chi]$ in the sense of relative entropy. In other words, one has

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \ln(w) \right) = - \inf_{q \in \Omega} \int D\chi q[\chi] \ln(q[\chi]/p[\chi]). \quad (43)$$

Assuming now in addition that Ω consists of a set of distributions $p[\chi, J]$ that is in the exponential class (1) with J in a set A , or expectation value Φ in the corresponding set B , we can write w in terms of the divergence functional,

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \ln(w) \right) = - \inf_{J \in A} D[J||J'] = - \inf_{\Phi \in B} D[\Phi||\Phi']. \quad (44)$$

Here we assumed that the true distribution is $p[\phi, J'] = p[\phi, \Phi']$. By these arguments it becomes clear that the divergence functional plays a natural role in quantifying the probability for large deviations [60].

For the exponential family one finds easily the relative entropy

$$\begin{aligned} D[J||J'] &= (J^\alpha - J'^\alpha) \Phi_\alpha - W[J] + W[J'] \\ &= (J^\alpha - J'^\alpha) \frac{\delta W[J]}{\delta J^\alpha} - W[J] + W[J']. \end{aligned} \quad (45)$$

The expectation value Φ_α is here with respect to the distribution $p[\chi, J]$, as made explicit in the second line.

The expression in Eq. (45) is up to an interchange of arguments (a duality transform), the Bregman divergence associated with the convex function $W[J]$,

$$D[J||J'] = D_W[J'||J]. \quad (46)$$

In this context it is also clear that the Schwinger functional $W[J]$ is uniquely fixed by $D[J||J']$, up to an additive constant and a linear term in J . In other words, to fully reconstruct $W[J]$ from $D[J||J']$ from one needs additional information such as $W[0]$ and

$$\left. \frac{\delta W[J]}{\delta J^\alpha} \right|_{J=0} = \Phi_\alpha^{\text{eq}}. \quad (47)$$

The largest part of the functional information of $W[J]$ is contained in $D[J||J']$. Concretely, one finds for $J^\alpha = 0$

$$D[0||J'] + J'^\alpha \Phi_\alpha^{\text{eq}} + W[0] = W[J']. \quad (48)$$

First derivatives of the divergence functional with respect to the source coordinates are

$$\begin{aligned} \frac{\delta}{\delta J^\alpha} D[J||J'] &= (J^\lambda - J'^\lambda) \frac{\delta^2}{\delta J^\alpha \delta J^\lambda} W[J] \\ &= (J^\lambda - J'^\lambda) G_{\alpha\lambda}[J], \\ \frac{\delta}{\delta J'^\alpha} D[J||J'] &= - \frac{\delta}{\delta J^\alpha} W[J] + \frac{\delta}{\delta J'^\alpha} W[J'] \\ &= -\Phi_\alpha[J] + \Phi'_\alpha[J'], \end{aligned} \quad (49)$$

and they vanish for $J = J'$, as expected. Second derivatives are

$$\begin{aligned} \frac{\delta^2}{\delta J^\alpha \delta J^\beta} D[J||J'] &= (J^\lambda - J'^\lambda) \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\lambda} W[J] + G_{\alpha\beta}[J], \\ \frac{\delta^2}{\delta J^\alpha \delta J'^\beta} D[J||J'] &= -G_{\alpha\beta}[J], \\ \frac{\delta^2}{\delta J'^\alpha \delta J'^\beta} D[J||J'] &= G_{\alpha\beta}[J']. \end{aligned} \quad (50)$$

Let us also give the third derivatives,

$$\begin{aligned} \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\gamma} D[J||J'] &= (J^\lambda - J'^\lambda) \frac{\delta^4}{\delta J^\alpha \delta J^\beta \delta J^\gamma \delta J^\lambda} W[J] \\ &\quad + 2 \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\gamma} W[J], \\ \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J'^\gamma} D[J||J'] &= - \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J'^\gamma} W[J], \\ \frac{\delta^3}{\delta J^\alpha \delta J'^\beta \delta J'^\gamma} D[J||J'] &= 0, \\ \frac{\delta^3}{\delta J'^\alpha \delta J'^\beta \delta J'^\gamma} D[J||J'] &= \frac{\delta^3}{\delta J'^\alpha \delta J'^\beta \delta J'^\gamma} W[J']. \end{aligned} \quad (51)$$

It is particularly interesting to observe that the derivatives with respect to the second argument J' yield directly the connected correlation functions generated by the Schwinger functional $W[J']$. Similarly, one derivative with respect to J' and several derivatives with respect to J also lead to connected correlation functions. Schematically one has for $n \geq 2$

$$D^{(0,n)}[J||J'] = W^{(n)}[J'], \quad (52)$$

and

$$D^{(n-1,1)}[J||J'] = -W^{(n)}[J]. \quad (53)$$

This shows again how most of the information of $W[J]$ is contained in $D[J||J']$.

B. Divergence functional in expectation value coordinates

The divergence functional can also be expressed in terms of expectation value coordinates, and it is convenient to work then with the quantum effective action defined in (17). One obtains from (45) immediately

$$D[\Phi||\Phi'] = \Gamma[\Phi] - \Gamma[\Phi'] - \frac{\delta\Gamma[\Phi']}{\delta\Phi'_\lambda}(\Phi_\lambda - \Phi'_\lambda). \quad (54)$$

This can be seen as the Bregman divergence associated with $\Gamma[\Phi]$,

$$D[\Phi||\Phi'] = D_\Gamma[\Phi||\Phi']. \quad (55)$$

On the other side one can reconstruct $\Gamma[\Phi]$ up to an additive constant from $D[\Phi||\Phi']$ if additionally also Φ^{eq} satisfying (47) or (19) for $J^\alpha = 0$ is known. Specifically, the difference of the quantum effective action at Φ to the one at Φ^{eq} is the divergence functional,

$$\Gamma[\Phi] - \Gamma[\Phi^{\text{eq}}] = D[\Phi||\Phi^{\text{eq}}]. \quad (56)$$

On similar ground, let us note that from the functional $D[J||J']$ alone one cannot immediately infer the functional $D[\Phi||\Phi']$ (and vice versa) because the information about the relation between the source J^α and the expectation value Φ_α is not contained in $D[J||J']$ (or in $D[\Phi||\Phi']$). It is sufficient, however, to know in addition the expectation value Φ_α^{eq} . From (49) follows a relation for the expectation value as functional of the source,

$$\Phi_\alpha[J] = \Phi_\alpha^{\text{eq}} - \frac{\delta}{\delta J^\alpha} D[J||J']|_{J'=0}, \quad (57)$$

which allows one to implement the change of variables. Similarly, this can be done in the other direction.

Taking functional derivatives of (54) with respect to the two arguments yields

$$\begin{aligned} \frac{\delta}{\delta\Phi_\alpha} D[\Phi||\Phi'] &= \frac{\delta\Gamma[\Phi]}{\delta\Phi_\alpha} - \frac{\delta\Gamma[\Phi']}{\delta\Phi'_\alpha} = J^\alpha - J'^\alpha, \\ \frac{\delta}{\delta\Phi'_\lambda} D[\Phi||\Phi'] &= -\frac{\delta^2\Gamma[\Phi']}{\delta\Phi'_\alpha\delta\Phi'_\lambda}(\Phi_\lambda - \Phi'_\lambda) \\ &= -G^{\alpha\lambda}[\Phi'](\Phi_\lambda - \Phi'_\lambda), \end{aligned} \quad (58)$$

and they vanish for $\Phi = \Phi'$ as expected. Evaluating the first line at $\Phi' = \Phi^{\text{eq}}$ leads to a relation for the source as a functional of the expectation value,

$$J^\alpha[\Phi] = \frac{\delta}{\delta\Phi_\alpha} D[\Phi||\Phi^{\text{eq}}]. \quad (59)$$

Second derivatives are

$$\begin{aligned} \frac{\delta^2}{\delta\Phi_\alpha\delta\Phi_\beta} D[\Phi||\Phi] &= G^{\alpha\beta}[\Phi], \\ \frac{\delta^2}{\delta\Phi_\alpha\delta\Phi'_\beta} D[\Phi||\Phi'] &= -G^{\alpha\beta}[\Phi'], \\ \frac{\delta^2}{\delta\Phi'_\alpha\delta\Phi'_\beta} D[\Phi||\Phi'] &= -\frac{\delta^3}{\delta\Phi'_\alpha\delta\Phi'_\beta\delta\Phi'_\lambda} \Gamma[\Phi'](\Phi_\lambda - \Phi'_\lambda) \\ &\quad + G^{\alpha\beta}[\Phi'], \end{aligned} \quad (60)$$

and third derivatives follow as

$$\begin{aligned} \frac{\delta^3}{\delta\Phi_\alpha\delta\Phi_\beta\delta\Phi_\gamma} D[\Phi||\Phi] &= \frac{\delta^3}{\delta\Phi_\alpha\delta\Phi_\beta\delta\Phi_\gamma} \Gamma[\Phi], \\ \frac{\delta^3}{\delta\Phi_\alpha\delta\Phi_\beta\delta\Phi'_\gamma} D[\Phi||\Phi'] &= 0, \\ \frac{\delta^3}{\delta\Phi'_\alpha\delta\Phi'_\beta\delta\Phi'_\gamma} D[\Phi||\Phi'] &= -\frac{\delta^3}{\delta\Phi'_\alpha\delta\Phi'_\beta\delta\Phi'_\gamma} \Gamma[\Phi'], \\ \frac{\delta^3}{\delta\Phi'_\alpha\delta\Phi'_\beta\delta\Phi'_\gamma} D[\Phi||\Phi'] &= 2\frac{\delta^3}{\delta\Phi'_\alpha\delta\Phi'_\beta\delta\Phi'_\gamma} \Gamma[\Phi'] \\ &\quad - \frac{\delta^4}{\delta\Phi'_\alpha\delta\Phi'_\beta\delta\Phi'_\gamma\delta\Phi'_\lambda} \Gamma[\Phi'](\Phi_\lambda - \Phi'_\lambda). \end{aligned} \quad (61)$$

Here it is particularly interesting that the derivatives with respect to the first argument Φ_α yield the one-particle irreducible correlation functions generated by $\Gamma[\Phi]$. Something similar happens for one derivative with respect to Φ and several with respect to Φ' . Schematically, for $n \geq 2$,

$$D^{(n,0)}[\Phi||\Phi'] = \Gamma^{(n)}[\Phi], \quad (62)$$

and

$$D^{(1,n-1)}[\Phi||\Phi'] = -\Gamma^{(n)}[\Phi']. \quad (63)$$

To summarize, the divergence functional $D[\Phi||\phi']$, when supplemented with one expectation value configuration ϕ_α^{eq} corresponding to vanishing source J^α , contains the same information as the effective action $\Gamma[\Phi]$.

C. Divergence functional in mixed representation

There is also a very elegant mixed representation, where the expectation value coordinate is used for the first argument, and the source coordinate for the second argument of the divergence functional,

$$D[\Phi||J'] = \Gamma[\Phi] + W[J'] - J'^\alpha \Phi_\alpha. \quad (64)$$

Here it is manifest that functional derivatives with respect to the first argument generate one-particle irreducible correlation functions, and those with respect to the second argument generate connected correlation functions. Also, by setting one of the arguments to zero, one obtains the quantum effective action or Schwinger functional up to an additive constant, respectively.

First derivatives are

$$\begin{aligned} \frac{\delta}{\delta \Phi_\alpha} D[\Phi||J'] &= \frac{\delta \Gamma[\Phi]}{\delta \Phi_\alpha} - J'^\alpha = J^\alpha - J'^\alpha, \\ \frac{\delta}{\delta J'^\alpha} D[\Phi||J'] &= \frac{\delta W[J']}{\delta J'^\alpha} - \Phi_\alpha = \Phi'_\alpha - \Phi_\alpha, \end{aligned} \quad (65)$$

and they vanish for $J^\alpha = J'^\alpha$ and $\Phi_\alpha = \Phi'_\alpha$. In particular, when the first line is evaluated at $J' = 0$ one obtains

$$J^\alpha[\Phi] = \frac{\delta}{\delta \Phi_\alpha} D[\Phi||0] = \frac{\delta \Gamma[\Phi]}{\delta \Phi_\alpha}, \quad (66)$$

which is the field equation for the expectation value. Similarly, the second line in (65) gives $\Phi'_\alpha[J']$ when evaluated at some given configuration Φ_α . In this sense the mixed functional $D[\Phi||J']$ contains more information than either $D[J||J']$ or $D[\Phi||\Phi']$, because also the relation between sources and expectation values is directly contained. Note that $D[\Phi||J']$ is not in the form of a Bregman divergence.

Second derivatives are

$$\begin{aligned} \frac{\delta^2}{\delta \Phi_\alpha \delta \Phi_\beta} D[\Phi||J'] &= G^{\alpha\beta}[\Phi], \\ \frac{\delta^2}{\delta \Phi_\alpha \delta J'^\beta} D[\Phi||J'] &= -\delta^\alpha_\beta, \\ \frac{\delta^2}{\delta J'^\alpha \delta J'^\beta} D[\Phi||J'] &= G_{\alpha\beta}[J']. \end{aligned} \quad (67)$$

Third derivatives are simply

$$\begin{aligned} \frac{\delta^3}{\delta \Phi_\alpha \delta \Phi_\beta \delta \Phi_\gamma} D[\Phi||J'] &= \frac{\delta^3}{\delta \Phi_\alpha \delta \Phi_\beta \delta \Phi_\gamma} \Gamma[\Phi], \\ \frac{\delta^3}{\delta \Phi_\alpha \delta \Phi_\beta \delta J'^\gamma} D[\Phi||J'] &= 0, \\ \frac{\delta^3}{\delta \Phi_\alpha \delta J'^\beta \delta J'^\gamma} D[\Phi||J'] &= 0, \\ \frac{\delta^3}{\delta J'^\alpha \delta J'^\beta \delta J'^\gamma} D[\Phi||J'] &= \frac{\delta^3}{\delta J'^\alpha \delta J'^\beta \delta J'^\gamma} W[J']. \end{aligned} \quad (68)$$

More general, one finds here for $n \geq 2$,

$$\begin{aligned} D^{(n,0)}[\Phi||J'] &= \Gamma^{(n)}[\Phi], \\ D^{(0,n)}[\Phi||J'] &= W^{(n)}[J']. \end{aligned} \quad (69)$$

Let us summarize that the divergence functional $D[\Phi||J']$ in the mixed representation of expectation value for the first argument and source for the second argument contains information equivalent to the Schwinger functional $W[J]$ and effective action $\Gamma[\Phi]$, respectively. In fact, it elegantly combines the two functionals.

D. Divergence functional in opposite mixed representation

Finally, let us discuss also a mixed representation with the opposite choice of coordinates

$$\begin{aligned} D[J||\Phi'] &= J^\lambda \frac{\delta W[J]}{\delta J^\lambda} - W[J] + \Phi'_\lambda \frac{\delta \Gamma[\Phi']}{\delta \Phi'_\lambda} \\ &\quad - \Gamma[\Phi'] - \frac{\delta W[J]}{\delta J^\lambda} \frac{\delta \Gamma[\Phi']}{\delta \Phi'_\lambda}. \end{aligned} \quad (70)$$

First derivatives are here

$$\begin{aligned} \frac{\delta}{\delta J^\alpha} D[J||\Phi'] &= (J^\lambda - J'^\lambda) G_{\alpha\lambda}[J], \\ \frac{\delta}{\delta \Phi'_\alpha} D[J||\Phi'] &= -G^{\alpha\lambda}[\Phi'] (\Phi_\lambda - \Phi'_\lambda). \end{aligned} \quad (71)$$

On first sight $D[J||\Phi']$ in this opposite mixed representation seems to contain even less information than $D[J||J']$ or $D[\Phi||\Phi']$, because even with the additional knowledge of the expectation value Φ_α^{eq} corresponding to vanishing source $J^\alpha = 0$, it is not possible to infer from (71) the general relation between expectation values and sources. Only if in addition also the two-point function or Fisher metric $G_{\alpha\lambda}[0]$ at $J = 0$ or its inverse $G^{\lambda\alpha}[\Phi^{\text{eq}}]$ is known, can one evaluate the first line of (71) (at $J = 0$ to yield

$$J'^\lambda[\Phi'] = -G^{\lambda\alpha}[\Phi^{\text{eq}}] \frac{\delta}{\delta J^\alpha} D[J||\Phi']|_{J=0}. \quad (72)$$

Similarly, evaluating the second line at $\Phi'_\alpha = \Phi_\alpha^{\text{eq}}$ gives

$$\Phi_\lambda[J] = \Phi_\lambda^{\text{eq}} - G_{\lambda\alpha}[0] \frac{\delta}{\delta \Phi'_\alpha} D[J||\Phi']|_{\Phi'=\Phi^{\text{eq}}}. \quad (73)$$

However, the necessary information about the Fisher metric is contained in the second functional derivatives of (70), which follow in general as

$$\begin{aligned}
\frac{\delta^2}{\delta J^\alpha \delta J^\beta} D[J||\Phi'] &= (J^\lambda - J'^\lambda) \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\lambda} W[J] + G_{\alpha\beta}[J], \\
\frac{\delta^2}{\delta J^\alpha \delta \Phi'_\beta} D[J||\Phi'] &= -G_{\alpha\lambda}[J] G^{\lambda\beta}[\Phi'], \\
\frac{\delta^2}{\delta \Phi'_\alpha \delta \Phi'_\beta} D[J||\Phi'] &= -\frac{\delta^3}{\delta \Phi'_\alpha \delta \Phi'_\beta \delta \Phi'_\lambda} \Gamma[\Phi'] (\Phi_\lambda - \Phi'_\lambda) \\
&\quad + G^{\alpha\beta}[\Phi'].
\end{aligned} \tag{74}$$

Evaluating the first and third line at $J = 0$ and $\Phi' = \Phi^{\text{eq}}$ yields $G_{\alpha\beta}[0]$ and $G^{\alpha\beta}[\Phi^{\text{eq}}]$, respectively.

Finally, third derivatives are in this representation given by

$$\begin{aligned}
\frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\gamma} D[J||\Phi'] &= 2 \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\gamma} W[J], \\
&\quad - (J^\lambda - J'^\lambda) \frac{\delta^4}{\delta J^\alpha \delta J^\beta \delta J^\gamma \delta J^\lambda} W[J] \\
\frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta \Phi'_\gamma} D[J||\Phi'] &= \frac{-\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\lambda} W[J] \frac{\delta^2}{\delta \Phi'_\gamma \delta \Phi'_\lambda} \Gamma[\Phi'], \\
\frac{\delta^3}{\delta J^\alpha \delta \Phi'_\beta \delta \Phi'_\gamma} D[J||\Phi'] &= \frac{-\delta^2}{\delta J^\alpha \delta J^\lambda} W[J] \frac{\delta^3}{\delta \Phi'_\beta \delta \Phi'_\gamma \delta \Phi'_\lambda} \Gamma[\Phi'], \\
\frac{\delta^3}{\delta \Phi'_\alpha \delta \Phi'_\beta \delta \Phi'_\gamma} D[J||\Phi'] &= 2 \frac{\delta^3}{\delta \Phi'_\alpha \delta \Phi'_\beta \delta \Phi'_\gamma} \Gamma[\Phi'] \\
&\quad - \frac{\delta^4}{\delta \Phi'_\alpha \delta \Phi'_\beta \delta \Phi'_\gamma \delta \Phi'_\lambda} \Gamma[\Phi'] (\Phi_\lambda - \Phi'_\lambda).
\end{aligned} \tag{75}$$

This representation is less elegant and obviously leads to more involved expressions for derivatives.

The question arises to what extent one can recover the Schwinger functional $W[J]$ or quantum effective action $\Gamma[\Phi]$ from the opposite mixed representation functional (70). As we have argued, one can, with the additional information of Φ^{eq} construct the map between expectation values and sources, and the information in $D[J||\Phi']$ is then equivalent to the divergence functional in other representations.

VIII. FUNCTIONAL INTEGRAL REPRESENTATION FOR DIVERGENCE FUNCTIONAL

For some purposes it is useful to have a direct functional integral representation of functionals. From Eq. (45) one finds

$$e^{-D[J||\Phi']} = \frac{e^{W[J] - J^\alpha \Phi_\alpha}}{e^{W[J] - J'^\alpha \Phi_\alpha}}. \tag{76}$$

Here one can use the functional integral representation for the Schwinger functional (2), leading to an intuitive expression for the divergence functional

$$e^{-D[J||\Phi']} = \frac{\int D\chi \exp(-I[\chi] + J^\alpha(\phi_\alpha[\chi] - \Phi_\alpha))}{\int D\tilde{\chi} \exp(-I[\tilde{\chi}] + J'^\alpha(\phi_\alpha[\tilde{\chi}] - \Phi_\alpha))}. \tag{77}$$

Note that the expectation value Φ_α appearing in the numerator as well as in the denominator is with respect to the distribution at source J^α .

At this point it is interesting to undo the coordinate change in (7) which yields

$$e^{-D[J||\Phi']} = \frac{\int D\phi \exp(-S[\Phi] + J^\alpha(\phi_\alpha - \Phi_\alpha))}{\int D\tilde{\phi} \exp(-S[\tilde{\phi}] + J'^\alpha(\tilde{\phi}_\alpha - \Phi_\alpha))}. \tag{78}$$

Assume now that the two functional integrals can be (approximately) evaluated in steepest descent method (one-loop approximation). This yields

$$\begin{aligned}
D[J||\Phi'] &\approx S[\phi[J]] - S[\phi[J']] \\
&\quad - J^\alpha(\phi_\alpha[J] - \Phi_\alpha) + J'^\alpha(\phi_\alpha[J'] - \Phi_\alpha) \\
&\quad + \frac{1}{2} \text{Tr}\{\ln S^{(2)}[\phi[J]] - \ln S^{(2)}[\phi[J']]\},
\end{aligned} \tag{79}$$

where $\phi_\alpha = \phi_\alpha[J]$ is a solution to the classical field equation

$$\frac{\delta}{\delta \phi_\alpha} S[\phi] = J^\alpha. \tag{80}$$

The functional integral representation (77) can be adapted to other coordinate systems by deriving from (61) the relations

$$\begin{aligned}
J^\alpha &= \frac{\delta}{\delta \Phi_\alpha} D[\Phi||\Phi^{\text{eq}}], \\
J'^\alpha &= -\frac{\delta}{\delta \Phi_\alpha} D[\Phi||\Phi']|_{\Phi=\Phi^{\text{eq}}},
\end{aligned} \tag{81}$$

where Φ^{eq} is the expectation value configuration corresponding to vanishing source $J = 0$. Using this, one finds in terms of expectation value coordinates

$$\begin{aligned}
e^{-D[\Phi||\Phi']} &= \frac{\int D\chi \exp(-I[\chi] + \frac{\delta}{\delta \Phi_\alpha} D[\Phi||\Phi^{\text{eq}}](\phi_\alpha[\chi] - \Phi_\alpha))}{\int D\tilde{\chi} \exp(-I[\tilde{\chi}] - \frac{\delta}{\delta \Phi_\alpha} D[\Phi||\Phi']|_{\Phi=\Phi^{\text{eq}}}(\phi_\alpha[\tilde{\chi}] - \Phi_\alpha))}.
\end{aligned} \tag{82}$$

Note that this is an implicit relation, because the divergence functional $D[\Phi||\Phi']$ appears on the right-hand side.

For the steepest descent approximation in (79) one finds to leading order

$$\begin{aligned} \frac{\delta}{\delta\Phi_\alpha} D[J||J'] &= \left(\frac{\delta}{\delta\phi_\beta} S[\varphi[J]] - J^\beta \right) \frac{\delta\varphi_\beta[J]}{\delta\Phi_\alpha} \\ &\quad - \frac{\delta J^\beta}{\delta\Phi_\alpha} (\varphi_\beta[J] - \Phi_\beta) + J^\alpha - J'^\alpha \\ &= - \frac{\delta J^\beta}{\delta\Phi_\alpha} (\varphi_\beta[J] - \Phi_\beta) + J^\alpha - J'^\alpha. \end{aligned} \quad (83)$$

Equation (81) are therefore solved when $\varphi_\beta[J] = \Phi_\beta$. This leads to the leading order steepest descent approximation in expectation value coordinates

$$D[\Phi||\Phi'] \approx S[\Phi] - S[\Phi'] - \frac{\delta}{\delta\Phi_\alpha} S[\Phi'] (\Phi_\alpha - \Phi'_\alpha). \quad (84)$$

This is consistent with (54) and the standard steepest descent or one-loop approximation for the quantum effective action,

$$\Gamma[\Phi] \approx S[\Phi] + \frac{1}{2} \text{Tr}\{\ln S^{(2)}[\Phi]\}. \quad (85)$$

One can also develop the field theoretic perturbation theory to higher order and use it for an approximation to the divergence functional $D[\Phi||\Phi']$. In fact, many methods used in the context of quantum field theory could be transferred to information geometry in this manner.

Let us give for completeness also a functional integral relation in the mixed representation (64),

$$e^{-D[\Phi||J']} = \frac{\int D\chi \exp(-I[\chi] + \frac{\delta}{\delta\Phi_\alpha} D[\Phi||0] (\phi_\alpha[\chi] - \Phi_\alpha))}{\int D\tilde{\chi} \exp(-I[\tilde{\chi}] + J'^\alpha (\phi_\alpha[\tilde{\chi}] - \Phi_\alpha))}, \quad (86)$$

where we used

$$J^\alpha = \frac{\delta}{\delta\Phi_\alpha} D[\Phi||J']|_{J'=0} = \frac{\delta}{\delta\Phi_\alpha} D[\Phi||0]. \quad (87)$$

As seen before the mixed representation is particularly elegant in the sense that it does not need additional information about Φ^{eq} .

IX. CONNECTIONS

Let us now discuss the different connections in more detail. In the information-theoretic context a continuous family of connections, parametrized by a parameter α has been discussed [11]. For the field theoretic context mainly the two boundary cases $\alpha = -1$ and $\alpha = 1$ are important. They correspond to the affine spaces of source and expectation value coordinates. We emphasize that we are concerned here with a infinite-dimensional space, so the connections we discuss are to be understood in a functional way.

We start by working in source coordinates. Associated with the Fisher metric in Eq. (15) is a unique connection that is both torsion-free and metric compatible (i.e. free of nonmetricity). This is the Levi-Civita connection

$$\begin{aligned} (\Gamma_{\text{LC}})_\alpha{}^\beta{}_\gamma[J] &= \frac{1}{2} G^{\beta\lambda}[J] \\ &\quad \times \left(\frac{\delta}{\delta J^\alpha} G_{\gamma\lambda}[J] + \frac{\delta}{\delta J^\gamma} G_{\alpha\lambda}[J] - \frac{\delta}{\delta J^\lambda} G_{\alpha\gamma}[J] \right). \end{aligned} \quad (88)$$

For the exponential family one finds easily

$$(\Gamma_{\text{LC}})_\alpha{}^\beta{}_\gamma[J] = \frac{1}{2} G^{\beta\lambda}[J] \frac{\delta^3}{\delta J^\alpha \delta J^\beta \delta J^\lambda} W[J], \quad (89)$$

where $G^{\beta\lambda}[J]$ is the inverse Fisher metric.

Similarly, the Levi-Civita connection in expectation value coordinates is obtained as

$$\begin{aligned} (\Gamma_{\text{LC}})_\alpha{}^\beta{}_\gamma[\Phi] &= \frac{1}{2} G_{\beta\lambda}[\Phi] \left(\frac{\delta}{\delta\Phi_\alpha} G^{\gamma\lambda}[\Phi] \right. \\ &\quad \left. + \frac{\delta}{\delta\Phi_\gamma} G^{\alpha\lambda}[\Phi] - \frac{\delta}{\delta\Phi_\lambda} G^{\alpha\gamma}[\Phi] \right) \\ &= \frac{1}{2} G_{\beta\lambda}[\Phi] \frac{\delta^3}{\delta\Phi_\alpha \delta\Phi_\beta \delta\Phi_\gamma} \Gamma[\Phi], \end{aligned} \quad (90)$$

where $G^{\alpha\beta}[\Phi]$ is the Fisher metric in terms of expectation value coordinates and $G_{\alpha\beta}[\Phi]$ its inverse.

Starting from the Levi-Civita connection, one can write any other connection as

$$\Gamma_\alpha{}^\beta{}_\gamma = (\Gamma_{\text{LC}})_\alpha{}^\beta{}_\gamma + N_\alpha{}^\beta{}_\gamma, \quad (91)$$

where $N_\alpha{}^\beta{}_\gamma$ is known as the distortion tensor. (The difference of two connections transforms as a tensor under coordinate changes or diffeomorphisms.) The torsion tensor can be expressed through the distortion tensor as

$$T^\alpha{}_{\beta\gamma} = N_\beta{}^\alpha{}_\gamma - N_\gamma{}^\alpha{}_\beta, \quad (92)$$

and the nonmetricity tensor is expressed through the distortion tensor and the metric as

$$B_{\alpha\beta\gamma} = \frac{1}{2} G_{\beta\lambda} N_\alpha{}^\lambda{}_\gamma + \frac{1}{2} G_{\gamma\lambda} N_\alpha{}^\lambda{}_\beta. \quad (93)$$

After these general considerations, let us now address the e - and m -connection. In source coordinates, the e -connection symbols vanish, see Eq. (12). Together with Eq. (89) this implies the distortion tensor in source coordinates

$$(N_E)_{\alpha}^{\beta}{}_{\gamma}[J] = -\frac{1}{2}G^{\beta\lambda}[J]\frac{\delta^3}{\delta J^{\alpha}\delta J^{\beta}\delta J^{\lambda}}W[J]. \quad (94)$$

Torsion vanishes, and the nonmetricity tensor is

$$(B_E)_{\alpha\beta\gamma}[J] = -\frac{1}{2}\frac{\delta^3}{\delta J^{\alpha}\delta J^{\beta}\delta J^{\gamma}}W[J]. \quad (95)$$

Up to a factor, this is simply the connected three-point function.

The m -connection symbols in source coordinates are given by Eq. (37). Accordingly, the distortion tensor is here

$$(N_M)_{\alpha}^{\beta}{}_{\gamma}[J] = \frac{1}{2}G^{\beta\lambda}[J]\frac{\delta^3}{\delta J^{\alpha}\delta J^{\beta}\delta J^{\lambda}}W[J]. \quad (96)$$

The nonmetricity has simply the opposite sign compared to the one of the e -connection,

$$(B_M)_{\alpha\beta\gamma}[J] = -(B_E)_{\alpha\beta\gamma}[J] = \frac{1}{2}\frac{\delta^3}{\delta J^{\alpha}\delta J^{\beta}\delta J^{\gamma}}W[J]. \quad (97)$$

In the context of information geometry, the fully symmetric tensor $T_{\alpha\beta\gamma} = 2(B_E)_{\alpha\beta\gamma} = -2(B_M)_{\alpha\beta\gamma}$ is known as the Amari-Chentsov tensor [11,12].

By similar arguments, or by a change of coordinates, one also obtains the distortion tensors in expectation value coordinates,

$$(N_M)^{\alpha}{}_{\beta}{}^{\gamma}[\Phi] = -(N_E)^{\alpha}{}_{\beta}{}^{\gamma}[\Phi] = \frac{1}{2}G_{\beta\lambda}[\Phi]\frac{\delta^3}{\delta\Phi_{\alpha}\delta\Phi_{\beta}\delta\Phi_{\lambda}}\Gamma[\Phi]. \quad (98)$$

The nonmetricity tensors are here, up to factors, given by the one-particle irreducible three-point function,

$$(B_M)^{\alpha\beta\gamma}[\Phi] = -(B_E)^{\alpha\beta\gamma}[\Phi] = \frac{1}{2}\frac{\delta^3}{\delta\Phi_{\alpha}\delta\Phi_{\beta}\delta\Phi_{\gamma}}\Gamma[\Phi]. \quad (99)$$

We see that the e - and m -connection are dual in the sense that they have opposite nonmetricity and they are both free of torsion. This duality implies that vector fields $V^{\mu}[J]$ and $W^{\nu}[J]$ obey

$$\begin{aligned} \frac{\delta}{\delta J^{\alpha}}(G_{\mu\nu}[J]V^{\mu}[J]W^{\nu}[J]) &= G_{\mu\nu}[J](\nabla_{\alpha}^{(E)}V^{\mu}[J])W^{\nu}[J] \\ &+ G_{\mu\nu}[J]V^{\mu}[J](\nabla_{\alpha}^{(M)}W^{\nu}[J]), \end{aligned} \quad (100)$$

with the covariant functional derivatives associated with the e -connection,

$$\begin{aligned} \nabla_{\alpha}^{(E)}V^{\mu}[J] &= \frac{\delta}{\delta J^{\alpha}}V^{\mu}[J] + (\Gamma_E)_{\alpha}{}^{\mu}{}_{\beta}[J]V^{\beta}[J] \\ &= \nabla_{\alpha}^{(LC)}V^{\mu}[J] + (N_E)_{\alpha}{}^{\mu}{}_{\beta}[J]V^{\beta}[J], \end{aligned} \quad (101)$$

the covariant derivative associated with the m -connection,

$$\begin{aligned} \nabla_{\alpha}^{(M)}W^{\nu}[J] &= \frac{\delta}{\delta J^{\alpha}}W^{\nu}[J] + (\Gamma_M)_{\alpha}{}^{\nu}{}_{\beta}[J]W^{\beta}[J] \\ &= \nabla_{\alpha}^{(LC)}W^{\nu}[J] + (N_M)_{\alpha}{}^{\nu}{}_{\beta}[J]W^{\beta}[J], \end{aligned} \quad (102)$$

and the covariant derivative of the Levi-Civita connection

$$\nabla_{\alpha}^{(LC)}V^{\mu}[J] = \frac{\delta}{\delta J^{\alpha}}V^{\mu}[J] + (\Gamma_{LC})_{\alpha}{}^{\mu}{}_{\beta}[J]V^{\beta}[J]. \quad (103)$$

Equation (100) follows from the well-known relation for the Levi-Civita connection

$$\begin{aligned} \frac{\delta}{\delta J^{\alpha}}(G_{\mu\nu}[J]V^{\mu}[J]W^{\nu}[J]) &= G_{\mu\nu}[J](\nabla_{\alpha}^{(LC)}V^{\mu}[J])W^{\nu}[J] \\ &+ G_{\mu\nu}[J]V^{\mu}[J](\nabla_{\alpha}^{(LC)}W^{\nu}[J]), \end{aligned} \quad (104)$$

and the symmetry properties of the distortion tensors N_E and N_M .

Let us note here that the relation between connected and one-particle irreducible correlation functions is well explored in quantum and statistical field theory. However, a geometric understanding of these relations, as it is offered by information geometry, is new.

X. METRIC AND CONNECTION FROM DIVERGENCE AND GENERAL COORDINATES

Interestingly it is also possible to obtain the connection symbols directly from the divergence functional [11]. From Eq. (50) one can read off that the metric in source coordinates is given by

$$G_{\alpha\beta}[J] = -\frac{\delta^2}{\delta J^{\alpha}\delta J^{\beta}}D[J||J']|_{J=J'}. \quad (105)$$

This is an interesting relation because one can do a change of coordinates, e.g. $J^{\alpha} \rightarrow K^{\alpha} = K^{\alpha}[J]$, and both sides transform automatically in the right way,

$$\begin{aligned} G_{\mu\nu}[K] &= \frac{\delta J^{\alpha}}{\delta K^{\mu}}\frac{\delta J^{\beta}}{\delta K^{\nu}}G_{\alpha\beta}[J[K]] \\ &= -\frac{\delta J^{\alpha}}{\delta K^{\mu}}\frac{\delta J^{\beta}}{\delta K^{\nu}}\frac{\delta}{\delta J^{\alpha}}\frac{\delta}{\delta J^{\beta}}D[J||J'] \\ &= -\frac{\delta}{\delta K^{\mu}}\frac{\delta}{\delta K^{\nu}}D[K||K']|_{K=K'}. \end{aligned} \quad (106)$$

In other words, the divergence functional $D[K||K']$ is in this sense as a generating functional for the metric in arbitrary coordinates. One example for a coordinate change $J \rightarrow K[J]$ is the transition from sources to expectation value fields, but there are of course many more possibilities.

In a related way one can obtain the m -connection symbols

$$(\Gamma_M)_{\alpha\beta\gamma}[J] = G_{\beta\delta}[J](\Gamma_M)_\alpha{}^\delta{}_\gamma[J] \\ = -\frac{\delta^2}{\delta J^\alpha \delta J^\gamma} \frac{\delta}{\delta J^\beta} D[J||J']|_{J=J'}. \quad (107)$$

This has indeed the right transformation law for a connection and can immediately be evaluated in any other coordinate system.

Finally, the symbols of the dual e -connection vanish in source coordinates, but can be written as

$$(\Gamma_E)_{\alpha\beta\gamma}[J] = -\frac{\delta}{\delta J^\alpha} \frac{\delta^2}{\delta J^\beta \delta J^\gamma} D[J||J']|_{J=J'}. \quad (108)$$

This formula easily generalizes to other coordinates, for example, one can obtain (38) from (61). We observe how elegantly information geometry is encoded in a divergence function.

Once the functional m - and e -connection are available in a given coordinate system one can use them to perform covariant functional derivatives, very much as covariant derivatives are used in Riemannian geometry or in the context of gauge theories. It is crucial to realize that after a general (nonlinear) change of field coordinates $J \rightarrow K[J]$:

- (i) Functional derivatives with respect to the sources J^α generalize to covariant functional derivatives based on the e -connection Γ_E . One may call them e -covariant derivatives. With such e -covariant derivatives with respect to the second argument K' of the divergence functional $D[K||K']$ one can generate connected correlation functions in the new field basis.
- (ii) Functional derivatives with respect to expectation values Φ_α generalize to covariant functional derivatives based on the m -connection Γ_M . One may call them m -covariant derivatives. With such m -covariant derivatives with respect to the first argument K of the divergence functional $D[K||K']$ one can generate one-particle irreducible correlation functions in the new field basis.

When working with such a calculus of covariant functional derivatives one must be careful in taking over intuition from standard Riemannian geometry in the sense that both covariant derivatives are not metric compatible.

Let us add another remark here. While it is straightforward to express the divergence functional in other coordinates, e.g.

$$D[K||K'] = D[J[K]||J'[K']], \quad (109)$$

it is in general not possible to write the resulting functional $D[K||K']$ again in the form of a Bregman divergence with reversed arguments as in (45). This works only when the transformation from J^α to K^α is an affine map.

Associated with some connection $\Gamma_\alpha{}^\beta{}_\gamma[J]$ is a Riemann curvature tensor

$$R^\alpha{}_{\beta\gamma\delta}[J] = \frac{\delta}{\delta J^\gamma} \Gamma_\delta{}^\alpha{}_\beta[J] - \frac{\delta}{\delta J^\delta} \Gamma_\gamma{}^\alpha{}_\beta[J] \\ + \Gamma_\gamma{}^\alpha{}_\lambda[J] \Gamma_\delta{}^\lambda{}_\beta[J] - \Gamma_\delta{}^\alpha{}_\lambda[J] \Gamma_\gamma{}^\lambda{}_\beta[J]. \quad (110)$$

However, the curvature tensor associated with the e - and m -connections both vanish [12]. This is immediately clear because $R^\alpha{}_{\beta\gamma\delta}[J]$ is a tensor, and both the e - and the m -connections have coordinate systems where their connection symbols vanish. In contrast, the curvature tensor associated with the Levi-Civita connection has no reason to vanish, and it is given by a combination of three-point functions and the Fisher metric.

XI. EXAMPLE: $O(N)$ MODEL

To exemplify our construction we consider now a Euclidean scalar field theory with $O(N)$ symmetry. We start from the quantum effective action in the second order of a derivative expansion,

$$\Gamma[\phi] = \int d^d x \left\{ U(\rho) + \frac{1}{2} Z(\rho) \delta^{\mu\nu} \partial_\mu \phi_n \partial_\nu \phi_n \right. \\ \left. + \frac{1}{4} Y(\rho) \delta^{\mu\nu} \partial_\mu \rho \partial_\nu \rho \right\}, \quad (111)$$

where $\rho(x) = (1/2) \sum_n \phi_n(x)^2$ is a combination of fields that is invariant under the $O(N)$ transformations,

$$\phi_n(x) \rightarrow \sum_{m=1}^N O_{nm} \phi_m(x), \quad (112)$$

where $O^T O = \mathbb{1}$. The divergence functional in the representation (54) is, up to boundary terms,

$$D[\phi||\phi'] = \Gamma[\phi] - \Gamma[\phi'] \\ - \int d^d x \left\{ \frac{\delta \Gamma[\phi']}{\delta \phi'_n(x)} (\phi_n(x) - \phi'_n(x)) \right\}. \quad (113)$$

Note that in any case $D[\phi||\phi']$ is invariant under $O(N)$ transformation with a single transformation matrix, $\phi_n(x) \rightarrow \sum_m O_{nm} \phi_m(x)$ and $\phi'_n(x) \rightarrow \sum_m O_{nm} \phi'_m(x)$, in the sense that

$$D[\phi||\phi'] = D[O\phi||O\phi'], \quad (114)$$

but not necessarily under a transformation acting on one argument only. In general one has $D[\phi||\phi'] \neq D[O\phi||\phi']$.

An exception is a configuration $\phi_n^{\text{eq}}(x)$ that solves the field equation

$$\left. \frac{\delta \Gamma[\phi']}{\delta \phi'_n(x)} \right|_{\phi'=\phi^{\text{eq}}} = 0. \quad (115)$$

If the second argument of the divergence functional is set to such a field configuration, the second line in (113) vanishes, and one has

$$D[\phi \|\phi^{\text{eq}}] = \Gamma[\phi] - \Gamma[\phi^{\text{eq}}], \quad (116)$$

which is invariant under separate $O(N)$ transformations of the first or second argument,

$$D[\phi \|\phi^{\text{eq}}] = D[O\phi \|\phi^{\text{eq}}] = D[\phi \|\phi^{\text{eq}}]. \quad (117)$$

Note that this holds for any configuration $\phi_n(x)$ as first argument. This property can be used to find solutions to the

field equation (115) if only $D[\phi \|\phi']$, but not $\Gamma[\phi]$, is available.

Combining Eqs. (111) and (113) we find for homogeneous fields ϕ_n and ϕ'_n

$$D[\phi \|\phi'] = \int d^d x \{ U(\rho) - U(\rho') + (2\rho' - (\phi', \phi)) U'(\rho') \}, \quad (118)$$

where $\rho = (1/2) \sum_n \phi_n^2$, $\rho' = (1/2) \sum_n \phi_n'^2$ and where $(\phi', \phi) = \sum_n \phi_n' \phi_n$ is the $O(N)$ invariant scalar product. Because $2\rho' - (\phi', \phi) \geq 0$, positivity $D[\phi \|\phi']$ for $\rho = \rho'$, but arbitrary relative field orientation, requires $U'(\rho) \geq 0$. Expanding to quadratic order in $\Delta\rho = \rho - \rho'$ for $(\phi', \phi) = 2\sqrt{\rho'\rho}$ yields also the condition $U'(\rho) + 2\rho U''(\rho) \geq 0$.

At second order in derivatives, the expression in (118) gets supplemented by

$$\begin{aligned} \Delta D[\phi \|\phi'] = \int d^d x \Big\{ & \frac{1}{2} Z(\rho) \delta^{\mu\nu} \partial_\mu \phi_n \partial_\nu \phi_n + \frac{1}{2} Z(\rho') \delta^{\mu\nu} \partial_\mu \phi'_n \partial_\nu (\phi'_n - 2\phi_n) + \frac{1}{4} Y(\rho) \delta^{\mu\nu} \partial_\mu \rho \partial_\nu \rho \\ & + \frac{1}{4} Y(\rho') \delta^{\mu\nu} \partial_\mu \rho' \partial_\nu (3\rho' - 2(\phi', \phi)) + (2\rho' - (\phi', \phi)) \left[\frac{1}{2} Z'(\rho') \delta^{\mu\nu} \partial_\mu \phi'_n \partial_\nu \phi'_n + \frac{1}{4} Y'(\rho') \delta^{\mu\nu} \partial_\mu \rho' \partial_\nu \rho' \right] \Big\}. \end{aligned} \quad (119)$$

One may check that $D[\phi \|\phi']$ vanishes indeed for $\phi_n(x) = \phi'_n(x)$. To quadratic order in $\phi_n(x) - \phi'_n(x)$, positivity of $D[\phi \|\phi']$ requires that the inverse propagator has non-negative eigenvalues [cf. Eq. (60)]. When one expands around homogeneous field configurations this necessities, besides the above conditions, also $Z(\rho) \geq 0$ (for field perturbations orthogonal to the homogeneous background) and $Z(\rho) + \rho Y(\rho) \geq 0$ (for field perturbations parallel to the background configuration).

One can also take Eqs. (118) and (119) as starting point for coordinate changes, for example, to a parametrization $\phi = (\sigma, \pi)$ with $\sigma = \sqrt{2\rho_0 - \pi^2}$, where $\pi_j(x)$ are field coordinates on the manifold $O(N)/O(N-1)$ where $\rho(x) = \rho_0$. These are typical coordinates used for the formulation of nonlinear σ models. The formalism developed in Sec. X allows one to determine connected or one-particle irreducible correlation functions in field coordinates that differ from the original parametrization in a nonlinear way.

XII. CONCLUSIONS

To conclude, we have discussed here the conceptual setup of Euclidean quantum field theory in the functional integral representation from the point of view of information geometry. It is nice to see how naturally the concepts of information geometry apply in an infinite-dimensional

setting where the coordinates are fields. When source fields, that are usually introduced to calculate correlation functions, are seen as coordinates, a natural and rich geometric picture arises. Dual to this is a description with expectation value fields as functional coordinates. In this dual picture the coordinates parametrize macroscopic fields or, in other words, the macrostates of the statistical field theory.

It is clear to any physicist familiar with general relativity how powerful the concepts of differential geometry can be. It is therefore great to have a similar formalism now also available for Euclidean quantum fields. For example, one can easily go to general coordinates without losing the significance of convexity of generating functionals. One can work with connections and corresponding covariant derivatives very similar as familiar from spacetime geometry.

A particularly interesting feature of the type of information geometry explored here is that the metric as well as the two dual connections arise from the functional derivatives of a divergence functional. The latter corresponds to the relative information entropy between two probability distributions and it has many highly interesting properties. One of them is that it can serve as a generating functional for correlation functions, very much as the Schwinger functional or the quantum effective action. Another is that it is non-negative and of course it has an information-theoretic significance as exemplified by Sanov's theorem.

Another aspect we did not study here is the renormalization group. In fact, the Schwinger functional and quantum effective action are subject to renormalization. This is discussed in detail in particular in the context of the functional renormalization group [61–64]. In a companion publication we have presented a renormalization group flow equation for the divergence functional [65].

For field theories it is in principle conceivable that the effect of quantum or statistical fluctuations modifies the properties of the divergence functional when renormalization is needed in the ultraviolet or infrared. More investigations are needed to clarify this further. So far it seems that the divergence functional is, as a relative entropy, actually better defined than the Schwinger functional or quantum effective action. At least partly this seems to be related to the property that $D[\Phi||\Phi']$ vanishes for $\Phi = \Phi'$.

Finally, information geometry can also be developed for quantum states described by density matrices or reduced

density matrices. Central concepts like the relative entropy and Fisher information metric are defined in that context, as well. In the context of relativistic quantum field theory it is particularly interesting that the relative entropy is well defined also for spatial subregions [66]. Relative entropy can be used to formulate thermodynamics [5] and relativistic fluid dynamics [6] on the basis of quantum field theory. We can well imagine that a quantum extension of information geometry allows one eventually to understand quantum field theory dynamics in much more detail.

ACKNOWLEDGMENTS

The author would like to thank Holger Gies and Markus Schröfl for useful discussions. This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under 273811115—SFB 1225 ISOQUANT.

-
- [1] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Quantum source of entropy for black holes, *Phys. Rev. D* **34**, 373 (1986).
 - [2] M. Srednicki, Entropy and area, *Phys. Rev. Lett.* **71**, 666 (1993).
 - [3] C. Callan and F. Wilczek, On geometric entropy, *Phys. Lett. B* **333**, 55 (1994).
 - [4] E. Witten, APS medal for exceptional achievement in research: Invited article on entanglement properties of quantum field theory, *Rev. Mod. Phys.* **90**, 045003 (2018).
 - [5] S. Floerchinger and T. Haas, Thermodynamics from relative entropy, *Phys. Rev. E* **102**, 052117 (2020).
 - [6] N. Dowling, S. Floerchinger, and T. Haas, Second law of thermodynamics for relativistic fluids formulated with relative entropy, *Phys. Rev. D* **102**, 105002 (2020).
 - [7] J. Berges, S. Floerchinger, and R. Venugopalan, Dynamics of entanglement in expanding quantum fields, *J. High Energy Phys.* **04** (2018) 145.
 - [8] J. Berges, S. Floerchinger, and R. Venugopalan, Thermal excitation spectrum from entanglement in an expanding quantum string, *Phys. Lett. B* **778**, 442 (2018).
 - [9] V. Giansos and N. Tetrads, Entanglement entropy in a four-dimensional cosmological background, *Phys. Lett. B* **833**, 137331 (2022).
 - [10] K. Boutivas, G. Pastras, and N. Tetrads, Entanglement and expansion, *J. High Energy Phys.* **05** (2023) 199.
 - [11] S.-i. Amari, *Information Geometry and Its Applications*, 1st ed. (Springer Publishing Company, Incorporated, New York, 2016).
 - [12] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer, *Information Geometry*, Vol. 64 of *Ergebnisse Der Mathematik Und Ihrer Grenzgebiete* (Springer International Publishing, Cham, 2017), 10.1007/978-3-319-56478-4.
 - [13] D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer, New York, 2008), 10.1007/978-3-540-74636-2.
 - [14] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 2nd ed. (Cambridge University Press, Cambridge, England, 2017), 10.1017/9781139207010.
 - [15] V. Balasubramanian, J. J. Heckman, and A. Maloney, Relative entropy and proximity of quantum field theories, *J. High Energy Phys.* **05** (2015) 104.
 - [16] A. B. Zamolodchikov, Irreversibility of the flux of the renormalization group in a 2D field theory, *JETP Lett.* **43**, 730 (1986) [*Pis'ma Zh. Eksp. Teor. Fiz.* **43**, 565 (1986)].
 - [17] B. P. Dolan, D. A. Johnston, and R. Kenna, The Information geometry of the one-dimensional Potts model, *J. Phys. A* **35**, 9025 (2002).
 - [18] W. Janke, D. A. Johnston, and R. Kenna, The information geometry of the spherical model, *Phys. Rev. E* **67**, 046106 (2003).
 - [19] W. Janke, D. A. Johnston, and R. Kenna, Information geometry and phase transitions, *Physica (Amsterdam)* **336A**, 181 (2004).
 - [20] J. Stout, Infinite distance limits and information theory, *arXiv:2106.11313*.
 - [21] I. Basile, Emergent strings at infinite distance with broken supersymmetry, *Astronomy* **2**, 206 (2023).
 - [22] J. Stout, Infinite distances and factorization, *arXiv:2208.08444*.
 - [23] T. D. Brennan, F. Carta, and C. Vafa, The string landscape, the swampland, and the missing corner, *Proc. Sci. TASI2017* (2017) 015 [*arXiv:1711.00864*].
 - [24] E. Palti, The swampland: Introduction and review, *Fortschr. Phys.* **67**, 1900037 (2019).

- [25] J. Erdmenger, K. Grosvenor, and R. Jefferson, Information geometry in quantum field theory: Lessons from simple examples, *SciPost Phys.* **8**, 073 (2020).
- [26] J. Erdmenger, K. Grosvenor, and R. Jefferson, Towards quantifying information flows: Relative entropy in deep neural networks and the renormalization group, *SciPost Phys.* **12**, 041 (2022).
- [27] K. Grosvenor and R. Jefferson, The edge of chaos: Quantum field theory and deep neural networks, *SciPost Phys.* **12**, 081 (2022).
- [28] L. Diósi, G. Forgács, B. Lukács, and H. L. Frisch, Metrization of thermodynamic-state space and the renormalization group, *Phys. Rev. A* **29**, 3343 (1984).
- [29] J. Gaite and D. O'Connor, Field theory entropy, the H theorem, and the renormalization group, *Phys. Rev. D* **54**, 5163 (1996).
- [30] B. P. Dolan, Renormalisation group flow and geodesics in the $O(N)$ model for large N , *Nucl. Phys.* **B528**, 553 (1998).
- [31] D. C. Brody and A. Ritz, On the symmetry of real-space renormalisation, *Nucl. Phys.* **B522**, 588 (1998).
- [32] B. P. Dolan and A. Lewis, Renormalization group flow and parallel transport with non-metric compatible connections, *Phys. Lett. B* **460**, 302 (1999).
- [33] J. Preskill, Quantum information and physics: Some future directions, *J. Mod. Opt.* **47**, 127 (2000).
- [34] H. Casini and M. Huerta, A c -theorem for entanglement entropy, *J. Phys. A* **40**, 7031 (2007).
- [35] S. Apenko, Information theory and renormalization group flows, *Physica (Amsterdam)* **391A**, 62 (2012).
- [36] C. Bény and T. J. Osborne, Information-geometric approach to the renormalization group, *Phys. Rev. A* **92**, 022330 (2015).
- [37] C. Bény and T. J. Osborne, The renormalization group via statistical inference, *New J. Phys.* **17**, 083005 (2015).
- [38] R. Fowler and J. J. Heckman, Misanthropic entropy and renormalization as a communication channel, *Int. J. Mod. Phys. A* **37**, 2250109 (2022).
- [39] A. Koenigstein, M. J. Steil, N. Wink, E. Grossi, and J. Braun, Numerical fluid dynamics for FRG flow equations: Zero-dimensional QFTs as numerical test cases. II. Entropy production and irreversibility of RG flows, *Phys. Rev. D* **106**, 065013 (2022).
- [40] J. Cotler and S. Rezchikov, Renormalization group flow as optimal transport, *Phys. Rev. D* **108**, 025003 (2023).
- [41] D. S. Berman and M. S. Klinger, The inverse of exact renormalization group flows as statistical inference, *Entropy* **26**, 389 (2024).
- [42] H. Strobel, W. Muessel, D. Linnemann, T. Zibold, D. B. Hume, L. Pezzè *et al.*, Fisher information and entanglement of non-Gaussian spin states, *Science* **345**, 424 (2014).
- [43] P. Hauke, M. Heyl, L. Tagliacozzo, and P. Zoller, Measuring multipartite entanglement through dynamic susceptibilities, *Nat. Phys.* **12**, 778 (2016).
- [44] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, *Rev. Mod. Phys.* **90**, 035005 (2018).
- [45] S. Banerjee, J. Erdmenger, and D. Sarkar, Connecting Fisher information to bulk entanglement in holography, *J. High Energy Phys.* **08** (2018) 001.
- [46] M. Blau, K. S. Narain, and G. Thompson, Instantons, the information metric, and the AdS/CFT correspondence, [arXiv:hep-th/0108122](https://arxiv.org/abs/hep-th/0108122).
- [47] S. Aoki, K. Kikuchi, and T. Onogi, Geometries from field theories, *Prog. Theor. Exp. Phys.* **2015**, 101B01 (2015).
- [48] S. Aoki, J. Balog, T. Onogi, and P. Weisz, Flow equation for the large N scalar model and induced geometries, *Prog. Theor. Exp. Phys.* **2016**, 083B04 (2016).
- [49] S. Chapman, M. P. Heller, H. Marrochio, and F. Pastawski, Toward a definition of complexity for quantum field theory states, *Phys. Rev. Lett.* **120**, 121602 (2018).
- [50] J. Erdmenger, M. Flory, M. Gerbershagen, M. P. Heller, and A.-L. Weigel, Exact gravity duals for simple quantum circuits, *SciPost Phys.* **13**, 061 (2022).
- [51] J. Erdmenger, M. Gerbershagen, M. P. Heller, and A.-L. Weigel, From complexity geometry to holographic space-time, *Phys. Rev. D* **108**, 106020 (2023).
- [52] P. T. Komiske, E. M. Metodiev, and J. Thaler, Metric space of collider events, *Phys. Rev. Lett.* **123**, 041801 (2019).
- [53] P. T. Komiske, E. M. Metodiev, and J. Thaler, The hidden geometry of particle collisions, *J. High Energy Phys.* **07** (2020) 006.
- [54] G. Ruppeiner, Riemannian geometry in thermodynamic fluctuation theory, *Rev. Mod. Phys.* **67**, 605 (1995).
- [55] D. C. Brody and D. W. Hook, Information geometry in vapour-liquid equilibrium, *J. Phys. A* **42**, 023001 (2009).
- [56] A. Carollo, D. Valenti, and B. Spagnolo, Geometry of quantum phase transitions, *Phys. Rep.* **838**, 1 (2020).
- [57] S. Floerchinger, T. Haas, and M. Schröfl, Relative entropic uncertainty relation for scalar quantum fields, *SciPost Phys.* **12**, 089 (2022).
- [58] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, Hoboken, NJ, 2006).
- [59] M. Schröfl and S. Floerchinger, Relative entropy and mutual information in Gaussian statistical field theory, [arXiv:2307.15548](https://arxiv.org/abs/2307.15548).
- [60] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Vol. 38 of Stochastic Modelling and Applied Probability (Springer, Berlin, Heidelberg, 2010), [10.1007/978-3-642-03311-7](https://doi.org/10.1007/978-3-642-03311-7).
- [61] J. Berges, N. Tetradis, and C. Wetterich, Non-perturbative renormalization flow in quantum field theory and statistical physics, *Phys. Rep.* **363**, 223 (2002).
- [62] J. M. Pawłowski, Aspects of the functional renormalisation group, *Ann. Phys. (Amsterdam)* **322**, 2831 (2007).
- [63] H. Gies, Introduction to the functional RG and applications to gauge theories, in *Renormalization Group and Effective Field Theory Approaches to Many-Body Systems*, edited by A. Schwenk and J. Polonyi (Springer, Berlin, Heidelberg, 2012), Vol. 852, pp. 287–348, [10.1007/978-3-642-27320-9_6](https://doi.org/10.1007/978-3-642-27320-9_6).
- [64] B. Delamotte, An introduction to the nonperturbative renormalization group, in *Renormalization Group and Effective Field Theory Approaches to Many-Body Systems*, edited by A. Schwenk and J. Polonyi (Springer, Berlin, Heidelberg, 2012), Vol. 852, pp. 49–132, [10.1007/978-3-642-27320-9_2](https://doi.org/10.1007/978-3-642-27320-9_2).
- [65] S. Floerchinger, Exact flow equation for the divergence functional, *Phys. Lett. B* **846**, 138244 (2023).
- [66] H. Araki, Relative entropy for states of von Neumann algebras. II, *Publ. RIMS* **13**, 173 (1977).