Information-geometric approach to the renormalization group

Cédric Bény[1] and Tobias J. Osborne[1,2]

[1]*Institut für Theoretische Physik, Leibniz Universität Hannover, Appelstraße 2, 30167 Hannover, Germany*
[2]*Riemann Center for Geometry and Physics, Leibniz Universität Hannover, Appelstraße 2, 30167 Hannover, Germany*

We propose a general formulation of the renormalization group (RG) as a family of quantum channels which connect the microscopic physical world to the observable world at some scale. By endowing the set of quantum states with an operationally motivated information geometry, we induce the space of Hamiltonians with a corresponding metric geometry. The resulting structure allows one to quantify information loss along RG flows in terms of the distinguishability of thermal states. In particular, we introduce a family of functions, expressible in terms of two-point correlation functions, which are nonincreasing along the flow. Among those, we study the speed of the flow and its generalization to infinite lattices.

The renormalization group (RG), one of the most profound ideas in science, allows us to understand long-range physics without requiring us to completely describe the fundamental constituents of the universe. This is because small-scale microscopic behavior can be absorbed by a handful of parameters of some effective theory.

The RG is typically taken to act on the space of theories or, more concretely, the space $\mathcal{K}$ of *Hamiltonians* and is thus a kind of superoperation which produces from a given initial Hamiltonian an effective Hamiltonian on some larger scale. By repeatedly composing the RG operation one generates a flow or trajectory in $\mathcal{K}$. This flow typically possesses attractive fixed points or surfaces corresponding to the physics at large length scales.

The fact that vastly different theories can converge to the same submanifold characterized by a handful of relevant parameters—the phenomenon of "universality"—invites an information-theoretic explanation [1]. Indeed, because the renormalization procedure *ignores* progressively larger-scale features of the theory, its irreversibility should be directly related to a loss of information about the microscopic physics.

Previous work in this direction has focused on the attempt to derive a function on the manifold $\mathcal{K}$ which always decreases along the RG flow, except at a fixed point, which would prove the irreversibility of the renormalization procedure. For $(1 + 1)$-dimensional quantum fields, this is achieved by Zamolodchikov's c theorem [2], which establishes the existence of such a scalar function, expressible in terms of the correlation functions of the theory. Similar results have been obtained for higher-dimensional theories [3,4]. More direct information-theoretic approaches were also explored in the context of classical field theory [5,6] or using ground-state entanglement [7].

Here we propose a different approach based on an information-theoretic formulation of the renormalization group. We argue that the RG is naturally expressed in terms of a Markov process on the manifold of statistical quantum states. Because thermal states essentially specify their Hamiltonian, we can pull back any information-theoretic structure on the manifold of statistical quantum states to the manifold of Hamiltonians. In this context, the information being lost along the RG flow is that which measures *distinguishability* between different theories. This gives rise to an information metric on $\mathcal{K}$ which allows for the derivation of quantum field theory (QFT) friendly c-like quantities.

Suppose that the fundamental constituents of the system under study are quantum particles or degrees of freedom interacting at some extremely short length scale ε. We write the (possibly mixed) quantum state of this system as ρ_ε. When we perform an experiment at a larger length scale $\ell \gg \varepsilon$ we are making a very imperfect measurement: We are ignoring many unobservable degrees of freedom. The most general procedure to ignore degrees of freedom consistent with the laws of quantum mechanics is described by a *quantum channel* $\mathcal{E}$, i.e., a completely positive and trace-preserving map. We thus write the best description of the quantum state available to our experiments as $\rho_\ell = \mathcal{E}_\ell(\rho_\varepsilon)$, where $\mathcal{E}_\ell$ denotes the quantum channel which ignores (or discards) all the degrees of freedom we cannot observe. Typically in the context of Wilsonian renormalization [8] and the functional RG the map $\mathcal{E}_\ell$ is the partial trace channel over all momenta $|k| \gg 1/\ell$. For our purposes, we suppose that knowing the effective state on any intermediate scale is enough to obtain the effective state on a larger scale. We call $\ell \mapsto \mathcal{E}_\ell$ the *renormalization group* or *process*.

In order to connect this flow on states to a flow on Hamiltonians, we assume that the fundamental state ρ_ε is in thermal equilibrium at a temperature τ and hence has the form $\rho_\varepsilon = e^{-H_\varepsilon/\tau}/Z$ for some Hamiltonian H_ε. Furthermore, if the temperature is defined by a reservoir with which our system is in contact, it is natural that the effective states $\rho_\ell = \mathcal{E}_\ell(\rho_\varepsilon)$ are thermal states at the same temperature τ. The effective Hamiltonian H_ℓ at scale ℓ can then be defined so that $\rho_\ell = e^{H_\ell/\tau}/Z$ or $H_\ell = -\tau \ln(Z\rho_\ell)$. The partition function is independent of ℓ, thanks to the fact that the maps $\mathcal{E}_\ell$ are trace preserving. Since the temperature is assumed to be constant, we can work with the operator $K_\ell = -H_\ell/\tau$ instead of H_ℓ and, abusing language, call those *Hamiltonians*. The basic arena of our study is therefore the manifold $\mathcal{K}$ of Hermitian operators K with the same $Z = \text{Tr } e^K$.

On a finite-dimensional Hilbert space, the process $\mathcal{E}_\ell$ is generated by a Lindblad superoperator $\mathcal{L}_\ell$ so that $\frac{d}{d\ell}\rho_\ell = \mathcal{L}_\ell(\rho_\ell)$. In order to further simplify the presentation, we assume a constant generator $\mathcal{L}$. This may require the parameter of
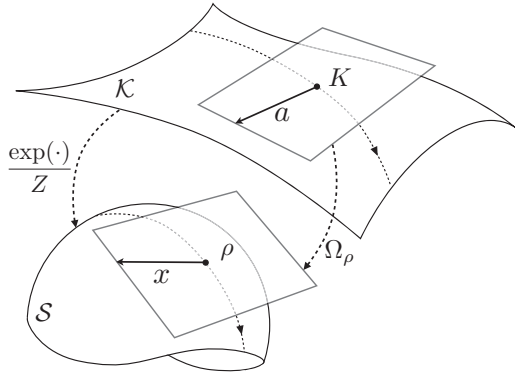
FIG. 1. Artist's impression of the manifold of states $\mathcal{S}$ and the manifold of Hamiltonians $\mathcal{K}$. The manifolds are related by the exponential map and their tangent spaces by its derivative Ω_ρ. Here $\rho = e^K/Z$ and $x = \Omega_\rho(a)$.

the flow to be different from the length scale ℓ, albeit still a monotone function of ℓ. In order to differentiate between them, we therefore denote that parameter t and write $\frac{d}{dt}\rho_t = \mathcal{L}(\rho_t)$ so that the family of channels $\mathcal{E}_t = e^{t\mathcal{L}}$ is the semigroup generated by $\mathcal{L}$.

We define the effective Hamiltonian at the scale indexed by t to be K_t such that $\rho_t = e^{K_t}/Z$. The evolution of the Hamiltonian $K_t = \ln(Z\rho_t)$ under the renormalization group is given by

$$\frac{d}{dt}K_t = \mathcal{B}(K_t),$$

with tangent field $\mathcal{B}(K) = \Omega_{e^K}^{-1}\mathcal{L}(e^K) = \Omega_\rho^{-1}\mathcal{L}(\rho)$, which we expressed in terms of the useful linear superoperator [9],

$$\Omega_\rho^{-1}(x) = \frac{d}{dt}\ln(\rho + tx)\Big|_0 = \int_0^\infty du \frac{1}{\rho + u\mathbf{1}}x\frac{1}{\rho + u\mathbf{1}}. \quad (1)$$

This superoperator can be understood as a noncommutative generalization of the operation of dividing by ρ. In geometric terms, it pushes forward a tangent vector at $\rho = e^K/Z$ to the corresponding tangent vector at K (see Fig. 1). We also need its inverse,

$$\Omega_\rho(a) = \frac{d}{dt}e^{\ln \rho + ta}\Big|_{t=0} = \int_0^1 \rho^s a\rho^{(1-s)}ds.$$

On the manifold of states, since the diffeomorphism $e^{t\mathcal{L}}$ is linear, its push forward on tangent vectors is just itself. Hence, on the manifold of Hamiltonians, it maps the tangent vector a at K to $a_t = \Omega_{\rho_t}^{-1}e^{t\mathcal{L}}\Omega_\rho(a)$ at the point K_t. The infinitesimal form is

$$\begin{aligned}\frac{d}{dt}a_t &= \frac{d}{d\epsilon}\mathcal{B}(K_t + \epsilon a_t)\Big|_{\epsilon=0} \\ &= \left[\Omega_{\rho_t}^{-1}\mathcal{L}\Omega_{\rho_t} - \Omega_{\rho_t}^{-1}\dot{\Omega}_{\rho_t}\right](a_t) \\ &=: \nabla_{\rho_t}a_t.\end{aligned} \quad (2)$$

Although our setup is finite dimensional, it would be natural to identify $\mathcal{B}(K)$ with the β *function* of quantum field theory in the following way. For a translation-invariant QFT one would usually assume that K has the form $K = \int dx \sum_i g_i \Phi_i(x)$ where the family $\{\Phi_i(x)\}_{i=0}^\infty$ spans all local field operators at x

and g_i are the corresponding coupling constants. The generator $\mathcal{B}(K)$ then should have the form $\mathcal{B}(K) = \int dx \sum_i \beta_i(K)\Phi_i(x)$ where the component $\beta_i(K) \in \mathbb{R}$ is the β function for the coupling constant g_i.

I. INFORMATION GEOMETRY

Recall that vectors tangent to the Hamiltonian K are observables with zero expectation value with respect to the state e^K/Z. For any two such observables a and b, we define the bilinear form

$$\langle a,b\rangle_K := -\frac{\partial^2}{\partial t\,\partial s}F(K + ta + sb)\Big|_0 = \text{Tr}\,[a\Omega_\rho(b)], \quad (3)$$

where $F(K) = -\ln \text{Tr}\,e^K$ is the free energy for the Hamiltonian $H = -\tau K$.

This defines a Riemannian metric on the manifold of observables. The corresponding metric on the manifold of states is the Kubo-Mori metric: Given two tangent (traceless) vectors x and y at ρ,

$$\langle x,y\rangle_\rho = \frac{\partial^2}{\partial t\,\partial s}S(\rho + tx + sy\|\rho)\Big|_0 = \text{Tr}\,[x\Omega_\rho^{-1}(y)], \quad (4)$$

where $S(\rho'\|\rho) = \text{Tr}\,[\rho'(\ln \rho' - \ln \rho)]$ is the *relative entropy* [10], an information-theoretic quantity which measures how *distinguishable* the two states ρ and ρ' are [11]. This gives an operational meaning to our metric. We can think of the norm $\sqrt{\langle a,a\rangle_K}$ of a small tangent vector a as measuring how distinguishable the perturbed state e^{K+a}/Z is from e^K/Z.

There are two compatible interpretations of the relation between these bilinear forms, connected to the dual role played by observables in equilibrium statistical mechanics. On one hand, observables are perturbations of Hamiltonians and are hence tangent vectors to the manifold of Hamiltonians, connected to the manifold of states via the exponential diffeomorphism. On the other hand, observables can be seen as cotangent vectors on the manifold of states since they are meant to be combined with states to yield an expectation value.

The main property of these dual metrics is that the norm of a tangent vector cannot increase under the action of a quantum channel [12]. For states this means that for any channel $\mathcal{N}$ and tangent vector x,

$$\langle \mathcal{N}(x),\mathcal{N}(x)\rangle_{\mathcal{N}(\rho)} \leqslant \langle x,x\rangle_\rho. \quad (5)$$

In classical probability theory, there exists only one metric with this property: the Fisher-Rao metric, which has been extensively used in statistical physics (e.g., see Refs. [13–15] in the context of renormalization). The metric defined in Eq. (4) is one of a family of noncommutative generalizations [16] and reduces to the Fisher-Rao metric for commuting operators and states.

It follows from this contraction property that, along the flow generated by $\mathcal{B}$ on the manifold of Hamiltonians, for any tangent vector a,

$$\frac{d}{dt}\langle a_t,a_t\rangle_{K_t} \leqslant 0, \quad (6)$$

which, more explicitly, is equivalent to the statement that for all $\rho = e^K/Z$ and for all tangent vectors a,

$$2\langle a, \nabla_\rho(a)\rangle_K + \text{Tr}\,[a\dot{\Omega}_\rho(a)] \leqslant 0, \qquad (7)$$

where the $\dot{\Omega}_\rho$ denotes the derivative of Ω_ρ with respect to t.

II. NONINCREASING SCALAR

Motivated by Zamolodchikov's c theorem [2], we want to build a scalar field $K \mapsto f(K)$ on the manifold $\mathcal{K}$ of Hamiltonians which has the property that it never increases under the renormalization flow generated by $\mathcal{B}$, i.e., such that

$$\frac{d}{dt} f(K_t) \leqslant 0, \qquad (8)$$

and, if possible, such that it stops decreasing only at a fixed point.

In the case where an exact fixed point of the renormalization semigroup is known, i.e., a state σ such that $\mathcal{L}(\sigma) = 0$, then any contractive measure of distinguishability with respect to σ has that property. An example would be the relative entropy: $f(K) = S(e^K/Z \| \sigma)$. Indeed, $S(\rho_1 \| \rho_2)$ satisfies $S[\mathcal{E}(\rho_1) \| \mathcal{E}(\rho_2)] \leqslant S(\rho_1 \| \rho_2)$ for any quantum channel $\mathcal{E}$. It has a straightforward interpretation: It measures the distinguishability with respect to the fixed point. Moreover, if σ is the maximally mixed state, i.e., the infinite temperature fixed point, then this function is simply a constant minus the entropy. However, this presupposes that the fixed point towards which the system flows is known. In addition, it is a very difficult quantity to compute. We propose instead a quantity which can be computed locally directly in terms of correlation functions.

In view of Eq. (6), the norm of any observable, which is expressed in terms of a Green's function, does decrease along the flow. That quantity is not purely a function of the Hamiltonian K. However, it can be made so if the observable is replaced by a tangent *field* $K \mapsto A(K)$ which is invariant under the flow, i.e., such that, for all t, $A(K_t)$ is equal to the image under the flow of the tangent vector $A(K_0)$. Using Eq. (2), the infinitesimal version of this condition is

$$\frac{d}{dt} A(K_t) = \nabla_{\rho_t}[A(K_t)], \qquad (9)$$

where $\rho_t = e^{K_t}/Z$. This is equivalent to the statement that the Lie derivative of A with respect to $\mathcal{B}$ is zero. An example is $A = \mathcal{B}$ itself, or fields of the form $A(K) = \Omega_{e^K}^{-1}\mathcal{L}'(e^K)$, where $\mathcal{L}'$ is any linear superoperator which commutes with $\mathcal{L}$. Indeed, if a field A satisfies Eq. (9), then Eq. (7) implies that the scalar,

$$f(K) = \langle A(K), A(K)\rangle_K, \qquad (10)$$

where $\rho = e^K/Z$ satisfies Eq. (8). For instance, if we use $A := \mathcal{B}$, then $f(K)$ is the squared *speed* of the flow, $f(K) = \langle \mathcal{B}(K), \mathcal{B}(K)\rangle_K$ measured in the distinguishability metric, which is zero if and only if $\mathcal{B}(K) = 0$, i.e., at a fixed point. In the classical setting, this corresponds to the "velocity function" of Ref. [14].

III. INFINITE LATTICES

We have so far worked within the framework of finite-dimensional Hilbert spaces and for a generic semigroup $\mathcal{E}_t$. However, to study renormalization more specifically, we may need to consider infinite lattices and translation-invariant states

for which quantities, such as those defined in Eq. (10), are usually infinite.

The natural course of action is to consider a *density* corresponding to $f(K)$. But, as we will see, this presents severe issues if the renormalization flow includes an explicit rescaling of the lattice.

Therefore, we first consider the case where the renormalization flow commutes with translations, which means that it consists only of decimation with no active rescaling of length. Explicitly, if T_x is the superoperator implementing a translation of the state by the lattice vector x, we require that $\mathcal{L}T_x = T_x\mathcal{L}$.

In addition, we require that the renormalization flow stays inside a submanifold $\mathcal{M}$ of local translation-invariant Hamiltonians, which implies that, for any $K \in \mathcal{M}$, $\mathcal{B}(K) = \sum_x b_x(K)$ where the operators $b_x(K) = T_x[b(K)]$ act only on a finite number of sites and T_x is a linear superoperator implementing a translation by the lattice vector x.

We then define

$$f(K) = \sum_x \langle b_0(K), b_x(K)\rangle_K = \langle b(K), \mathcal{B}(K)\rangle_K. \qquad (11)$$

This quantity corresponds to the speed *density* of the flow. We show that under the above assumptions, it is indeed nonincreasing along the flow, i.e., $\frac{d}{dt} f(K_t) \leqslant 0$. First we observe that since $[T_x, \nabla_\rho] = 0$ and $T_x(\rho) = \rho$, then

$$\langle \nabla_\rho b_y, b_x\rangle = \langle \nabla_\rho T_y b, b_x\rangle = \langle \nabla_\rho b, T_{-y}b_x\rangle = \langle \nabla_\rho b, b_{x-y}\rangle. \qquad (12)$$

Let $\mathcal{B}^L := \sum_{|x|<L} b_x(K)$, and V_L be the number of translations satisfying $|x| < L$. Then Eq. (7) implies

$$0 \geqslant 2\langle \nabla_\rho \mathcal{B}^L, \mathcal{B}^L\rangle + \text{Tr}\,[\mathcal{B}^L\dot{\Omega}_\rho(\mathcal{B}^L)]$$

$$= \sum_{|y|\leqslant L,|x|\leqslant L} 2\langle \nabla_\rho b, b_{x-y}\rangle + \text{Tr}\,[b\dot{\Omega}_\rho(b_{x-y})]$$

$$= \sum_x f_L(x)\{2\langle \nabla_\rho b, b_x\rangle + \text{Tr}\,[b\dot{\Omega}_\rho(b_x)]\},$$

where $f_L(x)$ counts the number of identical elements in the double sum. It is such that

$$\lim_{L\to\infty} \frac{1}{V_L} \sum_x f_L(x)\{2\langle \nabla_\rho b, b_x\rangle + \text{Tr}\,[b\dot{\Omega}_\rho(b_x)]\}$$

$$= 2\langle \nabla_\rho b, \mathcal{B}\rangle + \text{Tr}\,[b\dot{\Omega}_\rho(\mathcal{B})],$$

which then must also be negative. We conclude by showing that $\frac{d}{dt} f(K_t)$ is equal to that last quantity. The derivative $\frac{d}{dt} f(K_t)|_{t=0}$ is made of three terms, one of which is equal to $\text{Tr}\,[b\dot{\Omega}_\rho(\mathcal{B})]$. The second term is just $\langle b, \frac{d}{dt}\mathcal{B}|_0\rangle = \langle b, \nabla\mathcal{B}\rangle = \langle \nabla b, \mathcal{B}\rangle$ by virtue of Eq. (12). The last term is

$$\left\langle \frac{d}{dt} b(K_t)\Big|_0, \mathcal{B}\right\rangle = \sum_x \left\langle \frac{d}{dt} b(K_t)\Big|_0, T_x b\right\rangle$$

$$= \left\langle \frac{d}{dt}\mathcal{B}(K_t)\Big|_0, b\right\rangle$$

$$= \langle \nabla_\rho(\mathcal{B}), b\rangle = \langle \nabla_\rho(b), \mathcal{B}\rangle,$$

which concludes the proof.

IV. SCALING

The quantity defined in Eq. (11) may fail to be nonincreasing if the flow does not commute with translations. A simple counterexample is given by the one-dimensional Ising model with the flow generated by the superoperator $\mathcal{L}(\rho) = \sum_i [\text{Tr}_i(\rho) - \rho]$ where the sum is over all lattice sites and Tr_i traces out the ith site and then, for all $j \geqslant i$, replaces the jth sites by the states of the $(j + 1)$th site, i.e., it shifts the spins on the right of the ith in order to fill the gap. A good property of this renormalization group is that it does not increase the range of the interactions. Hence the nearest-neighbor Ising Hamiltonian stays within the manifold of commuting translation-invariant Hamiltonians acting only on nearest neighbors, which is three dimensional. The Hamiltonians can be written as $K(J,h,c) = J \sum_i S_i S_{i+1} + h \sum_i S_i + c\mathbf{1}$, where S_i is the diagonal matrix acting on site i with eigenvalues $+1$ and -1. In finite dimensions, the expectation value of $\mathcal{B}(K)$ would always be zero provided that the semigroup is trace preserving. However, here one needs to choose the parameter c so that the free-energy density is zero, i.e., $c = -\ln Z$.

On the stable submanifold defined by $c = -\ln(2\cosh J)$ and $h = 0$, we have $\mathcal{B}(K) = \sum_i b_i(K)$ with $b_i[K(J)] = -e^{-J}\sinh(J)S_i S_{i+1} - \ln[2\cosh(J)]\mathbf{1}$ [17]. We then obtain $f(K) = e^{-2J}[\tanh(J)]^2$. One can check $J > 0$ always decreases along the flow, but $f(K)$ attains a maximum for a finite J, which indeed contradicts the proposition that $f(K)$ be always decreasing along the flow.

Generally, the problem with an active scaling is that the *causal past* Σ' of a region Σ (with respect to the "dynamics" $\mathcal{E}_\ell$) would scale as the volume of Σ rather than its area, and therefore the difference between $\frac{1}{\Sigma}(\rho_\ell)_\Sigma$ and $\frac{1}{\Sigma'}(\rho_\ell)_{\Sigma'}$ does not necessarily vanish for large Σ.

This is a problem because we are generally interested in renormalization groups which converge to interesting fixed points. If no scaling is involved together with the removal of local information, then, strictly speaking, the only possible fixed points contain no correlations.

Let us sketch a possible way of addressing this problem. For simplicity, instead of the differential state-independent quantity $f(K)$, we consider the relative entropy density,

$$D(\rho\|\sigma) = \lim_\Sigma \frac{1}{|\Sigma|} S(\rho_\Sigma\|\sigma_\Sigma),$$

where the limit is over the net of subsets Σ of lattice sites and ρ_Σ denotes the reduced state on Σ for a full state ρ.

The idea is that, instead of implementing the scaling explicitly in the dynamics $\mathcal{E}_\ell$, which is problematic as it requires an unbounded generator as in the previous example, we can more simply implement it in the way that we compare states at different values of ℓ. Because we are dealing with a relative entropy *density*, it has a unit: that of inverse distance. This unit should scale with ℓ appropriately. Assuming that $\mathcal{L}_\ell$ is chosen so that $\mathcal{E}_\ell$ erases information about a *scale* ℓ so that if ϵ is the lattice spacing then $\mathcal{E}_\epsilon(\rho) = \rho$, we define

$$D_\ell(\rho_\ell\|\sigma_\ell) := \ell^d D(\rho_\ell\|\sigma_\ell),$$

where d is the number of spatial dimensions. This quantity has no unit. This tool allows us to say that two states ρ and σ converge "to the same fixed point" if $\lim_{s\to\infty} D_\ell(\rho_\ell\|\sigma_\ell) =$

0. The fixed point itself is represented only as the resulting equivalence class of states.

This quantity does not generally decrease with ℓ. Indeed, we have

$$\frac{d}{d\ell}D_\ell(\rho_\ell\|\sigma_\ell) = \frac{d}{\ell}D_\ell(\rho_\ell\|\sigma_\ell) + \ell^d \frac{d}{d\ell}D(\rho_\ell\|\sigma_\ell). \quad (13)$$

In general we only expect that the last term is negative (or zero), but we do not know that it is negative *enough* to compensate for the positivity of the first term $\frac{d}{\ell}D_\ell(\rho_\ell\|\sigma_\ell)$. However, a simple argument suggests that this term can be canceled by adding the right amount of depolarization to the coarse-graining procedure. Let $\mathcal{K}_\ell(\rho) = \frac{d}{\ell}\sum_i[\mathcal{D}_i(\rho) - \rho]$, where $\mathcal{D}_i$ is the application of the depolarization map $\mathcal{D}(\rho) = \text{Tr}(\rho)\mathbf{1}/\text{Tr}\,\mathbf{1}$ to the ith site and d is the dimension of space. Let ρ_ℓ be the state evolved under the dynamics generated by $\mathcal{L}'_\ell = \mathcal{L}_\ell + \mathcal{K}_\ell$. Furthermore, we write $\rho_{\ell,\epsilon} = \rho_\ell + \epsilon\mathcal{K}_\ell(\rho_\ell)$. From the joint convexity of the relative entropy we obtain, to first order in ϵ,

$$S[(\rho_{\ell,\epsilon})_\Sigma\|(\sigma_{\ell,\epsilon})_\Sigma] \leqslant \left(1 - \epsilon|\Sigma|\frac{d}{\ell}\right)S[(\rho_\ell)_\Sigma\|(\sigma_\ell)_\Sigma]$$

$$+ \epsilon\frac{d}{\ell}\sum_i S[(\rho_\ell)_{\Sigma\setminus i}\|(\sigma_\ell)_{\Sigma\setminus i}]$$

$$= S_\Sigma - \epsilon\frac{d}{\ell}\sum_i(S_\Sigma - S_{\Sigma\setminus i}),$$

where we abbreviated $S_\Sigma = S[(\rho_\ell)_\Sigma\|(\sigma_\ell)_\Sigma]$ and $\Sigma\setminus i$ is the region Σ without the site i. Since also the derivate of the relative entropy is negative with respect to the generator $\mathcal{L}_\ell$, we have altogether,

$$\frac{d}{d\ell}\frac{1}{|\Sigma|}S_\Sigma \leqslant -\frac{d}{\ell}\frac{1}{|\Sigma|}\sum_i(S_\Sigma - S_{\Sigma\setminus i}).$$

Since, for large Σ we expect $S_\Sigma \simeq |\Sigma|D(\rho_\ell\|\sigma_\ell)$, then in the thermodynamic limit, we ought to have

$$\frac{d}{d\ell}D(\rho_\ell\|\sigma_\ell) \leqslant -\frac{d}{s}D(\rho_\ell\|\sigma_\ell),$$

which, using Eq. (13), would imply that $\frac{d}{d\ell}D_\ell(\rho_\ell\|\sigma_\ell) \leqslant 0$.

V. EXAMPLE

Here we introduce an example of a Lindblad generator which can be defined on any lattice and generates a renormalization flow which commutes with translations (and hence does no active rescaling). Given this property, the quantity defined in Eq. (11) is nonincreasing. The Lindblad generator is

$$\mathcal{L}(\rho) = \sum_{ij}[U_{ij}\rho U_{ij}^\dagger - \rho], \quad (14)$$

where the sum is over all neighboring pairs of spins on the lattice and U_{ij} denotes the unitary map swapping sites i and j.

The interpretation of this generator is that, after a short "time" ϵ, we have a probability ϵ of confusing any given neighboring pair of sites, hence losing short-scale information. In fact, as we will see, a finite time t amounts to erasing information up to a length scale $\ell \propto \sqrt{t}$. Hence the corresponding
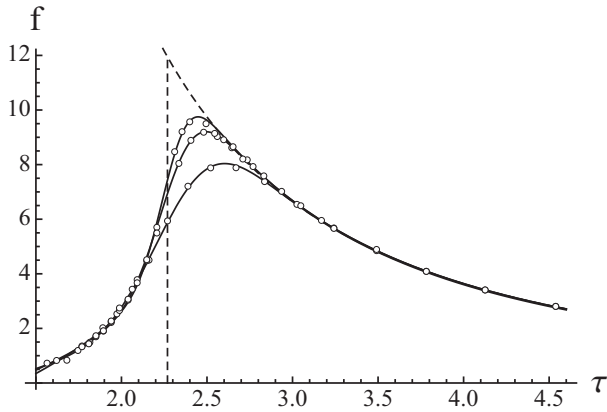
FIG. 2. The quantity $f(K)$ defined in Eq. (11) (per lattice unit) versus the temperature in units of J/k_B for the two-dimensional Ising model on periodic square lattices computed using a Monte Carlo simulation. The vertical dashed line represents the exact transition temperature τ_c. The different curves correspond to lattice sizes 10^2, 20^2, and 40^2. The fitting function is of the form $a/\tau^{2.1}$ (dashed curve) for $\tau > \tau_c$ and $c - d(\tau_c - \tau)^e$ for $\tau < \tau_c$, convoluted with a Gaussian to account for the finite-size effect.

generator for the flow in terms of the length scale ℓ is $\mathcal{L}_\ell = \frac{\ell}{\varepsilon^2}\mathcal{L}$, where ε is the lattice spacing.

The fixed points of this generator, i.e., states σ such that $\mathcal{L}(\sigma) = 0$, are precisely the states invariant under permutations of any two sites. According to the quantum de Finetti theorem, this implies that any reduced state on a finite number of sites has the form $\int d\mu(\rho)\, \rho^{\otimes N}$ where μ is a probability measure on mixed states. Hence, due to the absence of active rescaling, the exact fixed points all correspond to *mean-field* theories. More interesting critical states must be identified through the asymptotic behavior of the flow under a passive rescaling of the lattice spacing, i.e., in the continuum limit.

Let us study the effect of the semigroup on a one-dimensional lattice. It is best interpreted in the Heisenberg picture where it maps observables of the effective coarse-grained theory to actual physical observables of the microscopic theory. For instance, the action of the generator on an observable a_i at site i is $\mathcal{L}^\dagger(a_i) = a_{i+1} - 2a_i + a_{i-1}$.

We recognize a finite-difference approximation of a diffusion equation. Therefore, for long time t, we expect that any *local* operator of the form $a(f) = \sum_i f_i a_i$, where

$f_i > 0$ and $\sum_i f_i = 1$, will be mapped to $e^{t\mathcal{L}^\dagger}[a(f)] \simeq \frac{1}{\sqrt{4\pi t}} \sum_k e^{-[(k-i)^2]/4t} a_k$. In particular, we see that the physical resolution of our observables filtered by $e^{t\mathcal{L}} = \exp\left(\int_\varepsilon^\ell \mathcal{L}_u du\right)$ is indeed $\ell = \varepsilon\sqrt{t}$.

Figure 2 shows the nonincreasing quantity defined in Eq. (11) for the two-dimensional Ising model on a square lattice computed using a Monte Carlo simulation. We see that it peaks at the transition temperature and does not appear to diverge.

To conclude with this example, let us also show that different fixed points of $\mathcal{L}_\ell$ are still distinguishable in terms of the rescaled relative entropy density D_ℓ, even when the depolarization generator $\mathcal{K}_\ell$ introduced above is added to $\mathcal{L}_\ell$. For instance, consider mean-field states of the form $\rho^{\otimes N}$ and $\sigma^{\otimes N}$, where N is the number of lattice sites and ρ and σ are single-site states. First, observe that $\mathcal{L}_\ell$ and $\mathcal{K}_\ell$ commute. Since these states are invariant under $\mathcal{L}_\ell$, the action of the channel,

$$\mathcal{E}_\ell = \exp\left[\int_\varepsilon^\ell (\mathcal{L}_u + \mathcal{K}_u)du\right]$$
$$= \exp\left(\int_\varepsilon^\ell \mathcal{K}_u du\right)\exp\left(\int_\varepsilon^\ell \mathcal{L}_u du\right),$$

on these states is simply given by the depolarizing maps which acts independently on each site. Because the relative entropy is additive on product states, we obtain

$$D_\ell[\mathcal{E}_\ell(\rho^{\otimes\infty})\|\mathcal{E}_\ell(\sigma^{\otimes\infty})] = \frac{\ell^d}{\varepsilon^d}S(\rho_\ell\|\sigma_\ell),$$

where $\rho_\ell = \frac{\varepsilon^d}{\ell^d}\rho + (1 - \frac{\varepsilon^d}{\ell^d})\frac{\mathbf{1}}{\mathrm{Tr}\,\mathbf{1}}$.

Using the joint concavity of the relative entropy for the upper bound and the relation $\frac{1}{2}(\|\rho - \sigma\|_1)^2 \leqslant S(\rho\|\sigma)$ (see Ref. [18]) for the lower bound, we obtain

$$\tfrac{1}{2}(\|\rho - \sigma\|_1)^2 \leqslant D_\ell[\mathcal{E}_\ell(\rho^{\otimes\infty})\|\mathcal{E}_\ell(\sigma^{\otimes\infty})] \leqslant S(\rho\|\sigma).$$

In particular, the lower bound shows that the distance between different mean-field states is asymptotically finite, which shows that this renormalization procedure with passive rescaling can, at the very least, distinguish between different mean-field phases.

[1] J. Preskill, J. Mod. Opt. **47**, 127 (2000).

[2] A. Zamolodchikov, Pis'ma Zh. Eksp. Teor. Fiz. **43**, 565 (1986) [JETP Lett. **43**, 730 (1986)].

[3] J. L. Cardy, Phys. Lett. B **215**, 749 (1988).

[4] Z. Komargodski and A. Schwimmer, J. High Energy Phys. 02 (2011) 019.

[5] J. C. Gaite and D. O'Connor, Phys. Rev. D **54**, 5163 (1996).

[6] S. M. Apenko, Physica A **391**, 62 (2012).

[7] H. Casini and M. Huerta, J. Phys. A: Math. Theor. **40**, 7031 (2007).

[8] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).

[9] M. B. Ruskai, Int. J. Quantum Inform. **03**, 579 (2005).

[10] This duality between the free energy and the relative entropy can be traced to the fact that they are Legendre transforms of each other [19].

[11] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99 (1991).

[12] D. Petz, J. Math. Phys. **35**, 780 (1994).

[13] L. Diósi, G. Forgács, B. Lukács, and H. L. Frisch, Phys. Rev. A **29**, 3343 (1984).

[14] D. C. Brody and A. Ritz, Nucl. Phys. B **522**, 588 (1998).

[15] B. P. Dolan, Proc. R. Soc. London, Ser. A **454**, 2655 (1998).

[16] D. Petz, Linear Algebra Appl. **244**, 81 (1996).

[17] Incidentally, this is the same flow as the one derived in Ref. [14] with a different approach.

[18] M. Ohya and D. Petz, *Quantum Entropy and Its Use* (Springer-Verlag, Berlin, 2004).

[19] F. Hiai and D. Petz, Linear Algebra Appl. **181**, 153 (1993).