New Journal of Physics

The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG Institute of Physics

TOPICAL REVIEW • OPEN ACCESS

From classical to quantum information geometry: a guide for physicists

To cite this article: J Lambert and E S Sørensen 2023 New J. Phys. 25 081201

View the article online for updates and enhancements.

You may also like

- Ground-state information geometry and quantum criticality in an inhomogeneous spin model Yu-Quan Ma and
- Information geometry in vapour-liquid equilibrium Dorje C Brody and Daniel W Hook
- Information geometry in the population dynamics of bacteria Zihan Zhang, Shaohua Guan and Hualin Shi

New Journal of Physics

The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED 28 February 2023

REVISED 9 July 2023

ACCEPTED FOR PUBLICATION 27 July 2023

PUBLISHED 8 August 2023

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



TOPICAL REVIEW

From classical to quantum information geometry: a guide for physicists

J Lambert* 💿 and E S Sørensen 💿

Department of Physics & Astronomy, McMaster University, 1280 Main St. W., Hamilton, ON L8S 4M1, Canada * Author to whom any correspondence should be addressed.

E-mail: lambej3@mcmaster.ca

Keywords: quantum Fisher information, quantum information geometry, multipartite entanglement, quantum metrology, fidelity susceptibility, quantum variance

Abstract

Recently, there has been considerable interest in the application of information geometry to quantum many body physics. This interest has been driven by three separate lines of research, which can all be understood as different facets of quantum information geometry. First, the study of topological phases of matter characterized by Chern number is rooted in the symplectic structure of the quantum state space, known in the physics literature as Berry curvature. Second, in the study of quantum phase transitions, the fidelity susceptibility has gained prominence as a universal probe of quantum criticality, even for systems that lack an obviously discernible order parameter. Finally, the study of quantum Fisher information in many body systems has seen a surge of interest due to its role as a witness of genuine multipartite entanglement and owing to its utility as a quantifier of quantum resources, in particular those useful in quantum sensing. Rather than a thorough review, our aim is to connect key results within a common conceptual framework that may serve as an introductory guide to the extensive breadth of applications, and deep mathematical roots, of quantum information geometry, with an intended audience of researchers in quantum many body and condensed matter physics.

Contents

1. Overview and motivations	2
2. Introduction to information geometry	5
2.1. Classical information geometry	6
2.1.1. Shannon information	6
2.1.2. Fisher information	7
2.1.3. From information to geometry	8
2.2. Quantum information geometry	11
2.2.1. von Neumann entropy	11
2.2.2. QFI	12
2.2.3. Quantum geometric tensor	12
2.2.4. Topological invariants and volumes	15
3. Information geometry in practice	16
3.1. Relationship to linear response	16
3.2. The quantum variance	17
3.3. QFI and entanglement	19
3.4. Fidelity susceptibility	23
3.5. Quantum sensing	25
3.6. Chern number and the Euler characteristic	25
3.7. State space curvature	26
4. Summary and outlook	27
Data availability statement	27
Appendix A. Properties of Shannon information	27
Appendix B. Example calculations	27
B.1. Single qubit	27
References	29

1. Overview and motivations

The importance of geometry to the study of quantum many body systems has emerged in several major contexts. First, in the groundbreaking work of Thouless et al [1], who noticed the relationship between the geometric phase [2, 3] of a single electron wave function, and the topology of the band structure in two-dimensional insulators. Second, in the work of Zanardi and Paunković [4], who applied the notion of state space geometry to the study of quantum phase transitions, and termed the resulting geometric quantity fidelity susceptibility [5]. The fidelity susceptibility and geometric phase are two halves of a deeper whole. The former defines the Riemannian structure of the state space [6] while the latter defines the symplectic structure. Taken together, the study of these two structures is the subject of quantum information geometry [7]. The word information here arises due to the relationship between the geometry of the state space, and the degree to which the parameters that define the location of a state in state space may be estimated in a fixed number of measurements. In classical statistics this is the concept of Fisher information [8], which was generalized to the quantum context for pure states by Wootters [9], and then to the case of mixed states by Braunstein and Caves [10]. As we shall see, the relationship between geometry and information has important implications for a number of fields, including quantum many body physics, quantum sensing and quantum computing.

Many of the topics touched upon in the present guide have been the subject of a dedicated review. For example, for a review of fidelity susceptibility one can consult [5] or more recently [11], and for a review of the applications of the geometric phase to condensed matter systems one might consult [12]. A recent and very thorough review of quantum Fisher information (QFI) in the context of metrology can be found in [13]. For a mathematically rigorous review of the classical Fisher information in the context of statistical physics, one can consult [14]. In its bibliographical notes, this review also contains a very complete history of the field of information geometry from a mathematical perspective. Nonetheless, over the last decade a number of important connections have emerged between state space geometry, quantum sensing [15–18], multipartite entanglement [19, 20], topology [21–23], and response functions [24, 25], which demand a contemporary perspective, easily accessible to a quantum many-body audience. The present guide intends to serve two purposes. First, as a pedagogical introduction, using a unified notation and conceptual framework, to information geometry in quantum many body physics. Second, as a guide where key results in the field from the last two decades are gathered in one place, offering new researchers an opportunity to orient themselves in the field, and providing experienced researchers with an entry point for future inquiries. While

we avoid rigorous proofs, we strive to summarize the key foundational ideas in the mathematics literature, providing references to the original research literature.

Much of the utility of quantum information geometry arises from the distinction between classical fluctuations and quantum fluctuations that will play an important role in the application of quantum information geometry to condensed matter systems. To get a sense of what is meant by this distinction, we introduce the following example, which we refer to as the *classical case*. Imagine a collection of N rotors obeying the laws of classical mechanics, with orientations, \vec{n} , and angular momenta, \hat{L} , in three dimensions given by, $x \equiv \{\vec{n}_1, \ldots, \vec{n}_N, \hat{L}_1, \ldots, \hat{L}_N\} \in \Omega$, where Ω is the classical phase space. The time dependence of the state can be found by solving the classical equations of motion for a given set of initial conditions, yielding a particular trajectory $\chi(t)$. Then, any macroscopic property of the system can be described by the function, $\Lambda\{\chi(t)\}$. In practice, it is intractable in most cases to specify the initial conditions for any macroscopic collection of particles. Even theoretically, solving the equations of motion is only possible for systems with sufficient symmetry. Thus, in both theory and practice, we consider averages of this function over either fixed intervals of time, or ensembles of possible states, which are equivalent as long as the conditions for ergodicity are met [26]. Classical systems are thus modelled by a probability distribution p(x), called a state. Observable properties are characterized by their averages over this state,

$$\langle \Lambda \rangle_p = \int p(x) \Lambda(x) \mathrm{d}x$$
 (1)

and their variances and covariances,

$$\operatorname{Var}_{p}\{\Lambda\} = \langle \Lambda^{2} \rangle_{p} - \langle \Lambda \rangle_{p}^{2}$$

$$(2a)$$

$$\operatorname{Cov}_{p}\{\Lambda_{1},\Lambda_{2}\} = \langle \Lambda_{1},\Lambda_{2}\rangle_{p} - \langle \Lambda_{1}\rangle_{p}\langle \Lambda_{2}\rangle_{p}.$$

$$(2b)$$

In the classical case, fluctuations arise out of our ignorance of the precise conditions of the system at any given instant. We call these fluctuations *classical fluctuations*, or sometimes, *incoherent fluctuations*. The distribution for a system of rotors in equilibrium with a bath of inverse temperature β is captured by the Gibbs distribution,

$$p(x) = \frac{e^{-\beta E(x)}}{\mathcal{Z}} \tag{3}$$

where,

$$\mathcal{Z} = \int e^{-\beta E(x)} \,\mathrm{d}x. \tag{4}$$

Fluctuations that arise due to equilibration with a bath (whether or particles, energy, volume, etc...) are termed thermal fluctuations. All thermal fluctuations are classical, though not all classical fluctuations are thermal.

Classically, a pure state is one that is specified without any degree of uncertainty (so far as our physical theory allows). For any particular solutions of the equations of motion $\chi(t)$, we have the pure state,

$$p(x;\chi(t)) = \delta(x - \chi(t)).$$
(5)

An important albeit obvious property of classical pure states is that they are pure on all subsystems. Consider a subsystem V and its complement \overline{V} , such that the phase space is decomposed into $\Omega = \Omega_V \oplus \Omega_{\overline{V}}$, with points in phase space factorizing as $x = x_V \oplus x_{\overline{V}}$ and trajectories as $\chi(t) = \chi_V(t) \oplus \chi_{\overline{V}}(t)$. Then the state of the subsystem will be given by,

$$p_V(x_V;\chi_V(t)) = \int_{\Omega_{\bar{V}}} p(x;\chi(t)) \mathrm{d}x_{\bar{V}} = \delta(x_V - \chi_V(t)).$$
(6)

As long as our above factorizations are possible, the above state will also be pure. This is what we mean when we say that a classical pure state is pure on its subsystems.

The revelation of quantum mechanics is that, at the time-energy scales of Planck's constant, the properties of even a single particle undergoing the simplest imaginable dynamics are subject to persistent, irrevocable fluctuations. Where classically we defined a point in phase space explicitly in terms of the microscopic degrees of freedom that we might, in principle, know and measure, the quantum state space is defined by a unit norm vector in Hilbert space called the wavefunction, $|\Psi\rangle$ in a Hilbert space \mathcal{H} [27]. The

3

wavefunction induces a probability distribution on an observable represented a Hermitian operator $\hat{\Lambda}$ via the amplitude of projection onto the eigenbasis of $\hat{\Lambda}$, $p(\lambda) = |\langle \Psi | \lambda \rangle|^2$.

Once we grant that the state of a quantum system is given by a vector in a Hilbert space rather than by a complete set of observable values, all the reasoning of the classical case can be applied. Namely, it is practically very difficult to determine precisely what state a quantum many-body system is in at any given moment, and consequently we consider mixtures of quantum systems represented by the density matrix,

$$\hat{\rho} = \sum_{j} p_{j} |j\rangle \langle j| \tag{7}$$

where p_j is a classical probability distribution over the pure states labelled by j. Such mixtures introduce incoherent fluctuations, that is, fluctuations not caused by the wave-like nature of quantum phenomena. Just as in the classical case, the mixture in equation (7) can be reduced to a pure state, $|\psi\rangle$ by taking $p_j = \delta_{j,\psi}$. In this case, fluctuations arise entirely from quantum mechanical effects, as we shall see. The extent to which a quantum state represented by the density matrix $\hat{\rho}$ is pure is captured by the purity,

$$\mathcal{P} = \text{Tr}\{\hat{\rho}^2\}.$$
(8)

A purity of 1 is only achieved if $p_j = \delta_{j,\psi}$, corresponding to the density matrix associated to the pure state $|\psi\rangle$. An important example of a quantum mixed state is that of a system governed by a Hamiltonian \hat{H} in thermal equilibrium with a bath of inverse temperature β , which we refer to as a Gibbs state,

$$\hat{\rho}_{\text{Gibbs}} = \frac{\exp\{-\beta \hat{H}\}}{\mathcal{Z}},\tag{9}$$

where,

$$\mathcal{Z} = \text{Tr}\{\exp{-\beta\hat{H}}\}\tag{10}$$

is the partition function.

This formal transition, from points in classical phase space to vectors in Hilbert space, facilitates the introduction of two new physical principles. The first is the principle of superposition. Whereas the pure states on the right-hand side of equation (7) are combined incoherently, quantum states can also combine coherently in superposition,

$$|\psi\rangle = \sum_{j} c_{j} |j\rangle \tag{11}$$

where $c_j = \langle j | \psi \rangle$ are complex numbers such that $\sum_j |c_j|^2 = 1$. Unlike the mixture of states in equation (7), the superposition of states in equation (11) preserves the wave-like nature of quantum effects, allowing for the resulting fluctuations to become correlated in ways not possible when considering incoherent superpositions. One way to understand the difference is that the mixture decreases the purity, whereas super position does not. It is the principle of superposition that gives rise to entanglement. The second is the principle of uncertainty. That is the fact that for pairs of observables, say $\hat{\Lambda}_1$ and $\hat{\Lambda}_2$, have a minimum amount of simultaneous fluctuation,

$$\operatorname{Var}_{\hat{\rho}}\{\hat{\Lambda}_1\}\operatorname{Var}_{\hat{\rho}}\{\hat{\Lambda}_2\} \geqslant \frac{1}{4} \left| \left\langle \left[\hat{\Lambda}_1, \hat{\Lambda}_2\right] \right\rangle_{\hat{\rho}} \right|^2.$$
(12)

Hence, if $\hat{\Lambda}_1$ and $\hat{\Lambda}_2$ do not commute, $[\hat{\Lambda}_1, \hat{\Lambda}_2] \neq = 0$, this establishes limits on $\operatorname{Var}_{\hat{\rho}}\{\hat{\Lambda}_1\}$ and $\operatorname{Var}_{\hat{\rho}}\{\hat{\Lambda}_2\}$. The fluctuations that arise from these two principles are collectively termed quantum fluctuations.

An important implication of the principle of superposition is the emergence of entanglement. In direct analogy with the discussion surrounding equation (6), we may consider the state of a many-body pure state on one of its subsystems, *V*. In direct analogy with equation (6) we have,

$$\hat{\rho}_V^{\psi} = \operatorname{Tr}_{\bar{V}}\{\hat{\rho}^{\psi}\}.$$
(13)

Unlike in the case of classical pure states, the state $\hat{\rho}_V^{\psi}$ will only have unit purity if the state $|\psi\rangle$ admits the decomposition $|\psi\rangle = |\phi_V\rangle |\phi_{\bar{V}}\rangle$. If $\hat{\rho}^{\psi}$ is not pure, we say that the subsystems V and \bar{V} are entangled in the state $|\psi\rangle$.

Entanglement comes down to the producibility of a state, that is, the extent to which it can be factored into subsystems. Since the study of entanglement is a major application of information geometry, we take a



Figure 1. (a) Statistical structure of classical mechanics, with all observable subordinated to a single classical probability distribution. (b) Statistical structure of quantum mechanics, showing the subordination of different sets of observables to different distributions generated by the state $\hat{\rho}$.

moment to formalize this idea. Let $|\Psi\rangle$ be an *N* body state. Now consider a factorization of this state into *M* subsystems,

$$|\Psi\rangle = \bigotimes_{l=1}^{M} |\psi_l\rangle \tag{14}$$

where each subsystem has N_l particles with $\sum_l N_l = N$. Now let *m* be an integer such that $m \ge N_l$ for all ℓ . We then say that $|\Psi\rangle$ is *m*-producible. If a state is *m*-producible but no (m-1)-producible, we say that $|\Psi\rangle$ is *m*-partite entangled.

Thus, we have two kinds of fluctuations that coexist in many-body systems. The classical (thermal) fluctuations caused by limitations in our precise knowledge of the state of the system at a given instant, and the quantum fluctuations that seem to be intrinsic to the definition of the states themselves. These quantum fluctuations can give rise to correlations beyond classical physics, of which entanglement is a particular example. The Relationship between the two is depicted schematically in figures 1(a) and (b). In classical statistical mechanics, a single underlying probability distribution governs the outcome of all possible observables by assigning a probability to the classical pure states. In the classical case discussed above this means assigning a probability to each set of orientations and angular momenta. In the quantum case, however, each pure state gives rise to a probability distribution for each non-commuting set of observables. We will return to this fact in section 2.2.2 when we generalize classical Fisher information to the quantum case. Using information theory, we can understand these fluctuations, both classical and quantum, as the notion of distance that induces the aforementioned geometrical structure, and in turn, reveal the presence of quantum correlations.

The guide is laid out as follows. In the next section, we introduce the concept of quantum information geometry, by beginning with the classical case and then progressing to the quantum case. In section three, we go over key results in several different fields. Finally, we offer some discussion of the overall state of the field and suggest possible new lines of research. In addition, for readers interested in a more thorough discussion of many aspects of quantum information geometry, we note here some textbook length works that we recommend and cite throughout [30–32].

2. Introduction to information geometry

The study of information geometry, in both the quantum and classical contexts, has a rich mathematical history that predates quantum mechanics itself, with the work of Fubini [33] and Study [34], who first studied the Riemannian structure of complex projective spaces. Soon after this, the notion of Fisher information [8] emerged, which quantifies the precision with which a parameter can be estimated based on measurements sampled from a distribution which depends on the parameter (e.g. we might try to estimate the mean of a Gaussian distribution based on the sample average). The idea of Fisher information explicitly developed into a geometrical idea with the work of Rao [35], and hence, the classical notion of information

geometry was born. Some time later, mathematical physicists made the final connection, generalizing the classical Fisher information to the quantum context and, in the process, rediscovering the Fubini-Study metric [6, 9]. For a thorough review of the history of information geometry, one can refer to the bibliographic notes in [14].

Before introducing the notion of Fisher information, which will lead ultimately to our geometrical picture of the space of probability distributions, we review the more family concept of Shannon information.

2.1. Classical information geometry

2.1.1. Shannon information

In developing our notions of information, it helps to consider a particular problem. We imagine a meteorological laboratory with a number of field stations in various different locations. Each hour, the stations transmit the current weather conditions. The transmissions occur in binary, with each possible forecast in the set $\{x_i\}_{i=1}^{D}$ having encodings with lengths $\{L_i\}_{i=1}^{D}$. For example, x_1 , which might correspond to clear skies, could be encoded as 001 (with $L_1 = 3$), while x_2 , which might correspond to overcast skies, could be encoded as 0100 (with $L_2 = 4$), and so on. When determining an encoding of forecasts into binary, it is important that we ensure that no forecast's encoding is contained in the prefix of another forecast's encoding, lest we confuse our receiver (for example we do not want to use 1 as an encoding for clear skies if 10 is an encoding for cloudy skies). Codes with this property are called prefix codes, and they are always possible provided that the lengths of the encodings, $\{L_1, \ldots, L_D\}$, satisfy the inequality [36],

$$\sum_{i=1}^{D} 2^{-L_i} \leqslant 1. \tag{15}$$

This inequality can be generalized to alphabets of any size by replacing 2 with the length of the alphabet. In the interest of efficiency, we should aim to assign the shortest encoding to the most likely forecast, so that on average we transmit as few bits as possible. If the probabilities for each forecast are given by the set $\{p_i\}_{i=1}^{D}$, then we want to minimize the expected length of a transmission,

$$\langle L \rangle = \sum_{i=1}^{D} p_i L_i \tag{16}$$

subject to the constraint given by equation (15). In doing so, we find that the optimal encoding length of a forecast occurring with probability p_i is,

$$\ell_i = -\log_2(p_i). \tag{17}$$

In general we might imagine that our probabilities, $p_i = p(x_i | \vec{\lambda})$ depend on some other parameters, $\vec{\lambda}$ (such as temperature). In this case, equation (17) is called the *log-likelihood*. It can be read as the logarithm of the likelihood that $\vec{\lambda}$ has a particular value given the observed result x_i . We will encounter it again shortly. From equation (16) we can see that the optimal average message length is bounded from below by,

$$\langle L \rangle \geqslant \sum_{i=1}^{D} p_i \ell_i. \tag{18}$$

This lower bound is our first measure of information,

$$S = -\sum_{i} p_i \log_2(p_i) \tag{19}$$

and, as it was introduced by Shannon [37] a little less than a year before Kraft introduced the inequality in equation (15) [36], we call it the *Shannon Information*. To gain an understanding of its properties, let us first consider the trivial case where there is a single possible forecast with probability one hundred percent, i.e. $p_i = \delta_{1i}$ where δ_{1i} is the Kronecker delta function. Then the shortest possible encoding has average length zero. After all, what would be the point in even transmitting? Assuming now a uniform probability distribution on s_i , we see that the average length of the encoding will grow monotonically with the number of forecasts, *N*. This property is called monotonicity, and it is a natural property to demand from a measure of information.

The Shannon Information can tell us much more, however, than the average length of an optimal encoding. Considering the case of two possible forecasts, it is clear that the minimal encoding will be one bit (say dot for x_1 and dash x_2), regardless of the relative values of p_1 and p_2 . If, however, we know that p_1 is very

nearly 1 and p_2 is very close to zero, we can see that the Shannon information tells us something about how much the receiver learns from reading the message. If, at station *P*, the forecast is sunny nine out of every ten days, we learn relatively little when we receive a forecast of sunshine. The power of Shannon information is that it captures an intuitive aspect of information: the less probable an event is thought to be, the more information is gained when that event is observed.

For the mathematically inclined reader, one may wonder to what degree Shannon information is unique. In 1961, Rényi [38] found an entire family of information measures by imposing a number of conditions (see appendix A) that are considered sensible for a measure of information. These so-called Rényi entropies contain the Shannon information as a special case. The Shannon information is singled out by its property of recursion (its appearance in the prefix coding problem above).

Physicists will recognize that the Shannon information is a generalization of the concept of entropy. There is no particular reason why we should have to measure information in bits, and we can just as well take the logarithm in equation (19) to have base *e*, (in which case we say we are measuring information in *nats*). Given a thermodynamic ensemble of *W* possible microstates, all taken to be equally likely with $p_i = \frac{1}{W}$, we see that the Shannon information is simply the Boltzmann entropy,

$$S = k \ln(W). \tag{20}$$

For this reason we tend to use the words entropy and information interchangeably. Since changing the base of the logarithm simply amount to a redefinition of the units of information, we will do so frequently in the following text as a matter of convenience.

2.1.2. Fisher information

At our meteorological laboratory, we are interested in producing models that determine the distribution of forecasts from our field stations. To this end, we make a statistical model $P(\vec{\lambda}) = \{p_i(\lambda)\}$ where $\vec{\lambda} \in \mathbb{R}^d$ is a parameterization of possible distributions. As we receive data from the field stations, we want to determine as accurately as possible the values of $\vec{\lambda}$ that fit the data. For simplicity, let us consider the case of only one parameter. Let λ represent a parameter such as a temperature, and let $\tilde{\lambda}$ be an estimator for this parameter (such as the sample mean of some number of temperature measurements). The function $p(x|\lambda)$ gives the likelihood of the temperature *x* given the value of the mean temperature in the model λ . We assume our estimator is unbiased, which means that $\langle \tilde{\lambda} \rangle = \lambda$ (this condition can be relaxed without substantially changing the discussion),

$$0 = \int dx (\tilde{\lambda}(x) - \lambda) p(x|\lambda).$$
(21)

Differentiating both sides with respect to λ , we can write,

$$0 = \int \mathrm{d}x \, (\tilde{\lambda}(x) - \lambda) \partial_{\lambda} p(x|\lambda) - \int \mathrm{d}x \, p(x|\lambda). \tag{22}$$

Notice that we assume that the estimator itself does not depend on λ . This makes sense since the estimator is the processing that we do to our data to estimate λ and if it depended on λ it would be ill-defined. Performing the integral in the second term and rearranging we arrive at,

$$\int \mathrm{d}x \, (\tilde{\lambda}(x) - \lambda) p(x|\lambda) \left(\frac{1}{p(x|\lambda)} \partial_{\lambda} p(x|\lambda)\right) = 1.$$
(23)

Now we notice that the left-hand side of the above equation is just the covariance of the estimator with the logarithmic derivative of the probability. Squaring both sides of the above equation and applying the Cauchy-Schwarz inequality, yields

$$1 \leq \operatorname{Var}_{P}\{\lambda\}\operatorname{Var}_{P}\{\partial_{\lambda}\ln(p(x|\lambda))\}.$$
(24)

Notice here that we are assuming that the expected value of the derivative of the log-likelihood is zero. This is straightforward to show,

$$\langle \partial_{\lambda} \ln(p(x|\lambda)) \rangle = \int p(x|\lambda) \left(\partial_{\lambda} \ln(p(x|\lambda)) \right) dx = \partial_{\lambda} \int p(x|\lambda) dx = 0$$
(25)

The variance of $\tilde{\lambda}$ is thus bounded from below as,

$$\operatorname{Var}_{P}(\tilde{\lambda}) \geq \frac{1}{\operatorname{Cov}_{P}\left\{\tilde{\lambda}, \partial_{\lambda} \ln(p(x|\lambda))\right\}}.$$
(26)





Reintroducing the multidimensional nature of the problem and the definition of the Fisher information, we arrive at the Cramér-Rao bound,

$$\operatorname{Cov}_{P}\{\tilde{\lambda}_{\mu}, \tilde{\lambda}_{\nu}\} \geqslant \frac{1}{4F_{\mu\nu}}$$
(27)

where,

$$F_{\mu\nu} = \frac{1}{4} \int_{\Omega} \mathrm{d}x \frac{1}{p(x|\vec{\lambda})} \frac{\partial p(x|\vec{\lambda})}{\partial \lambda_a} \frac{\partial p(x|\vec{\lambda})}{\partial \lambda_b}$$
(28)

is the definition of the *Fisher information* [39]. We've added the factor of $\frac{1}{4}$ for reasons that will become apparent in a moment. The inequality in equation (27) is known as the Cramér-Rao bound [35]. It quantifies how much information can be learned about the parameters $\vec{\lambda}$ by sampling from $p(x|\vec{\lambda})$. The appearance of the derivative of the log-likelihood implies a connection between the Shannon information, which is the expectation value of the log-likelihood, and the Fisher information which is the covariance of the partial derivatives of the log-likelihood, and we will explore this connection in more detail momentarily.

We can begin to get a sense of the meaning of the Fisher information by considering the case of a weather station with only two forecasts. Let λ be the probability of Sun and $\lambda - 1$ be the probability of rain. We can see that when $\lambda = 0.5$, the Shannon information—the information we learn by reading the forecast—is maximal, whereas the Fisher information—our ability to estimate λ —is minimal. As λ approaches 1 or 0, the Shannon information decreases while the Fisher information grows. This behavior is shown in figure 2 where the Shannon information is given by,

$$S = -\lambda \log_2(\lambda) - (1 - \lambda) \log_2(1 - \lambda), \tag{29}$$

and the Fisher information for each measurement is,

$$F_{\lambda\lambda} = \frac{1}{4} \frac{1}{\lambda(1-\lambda)}.$$
(30)

2.1.3. From information to geometry

To better understand the connection between Shannon information, Fisher information and geometry, we consider the above scenario, that of estimating the value of the parameter λ , from a slightly different perspective. Considering again the model $p(x|\lambda)$, we can ask how distinguishable $P = p(x|\lambda)$ is from

 $Q = p(x|\lambda')$. The degree of distinguishability can be quantified in terms of the informational difference between *P* and *Q* by introducing the *relative entropy* [40],

$$S(P|Q) = \sum_{j=1}^{D} p_j \log_2\left(\frac{p_j}{q_j}\right).$$
(31)

If P = Q then the relative entropy is zero, and the distributions are indistinguishable. On the other hand, we can imagine that *P* corresponds to a place that is always sunny, while *Q* corresponds to a place that is always rainy. Letting x_1 be sunshine and x_2 be rain. Then $p_i = \delta_{1i}$ and $q_i = \delta_{2i}$, and we can see that the relative entropy is infinite.

Let us imagine now that *P* and *Q* correspond to the information arriving from different field stations. What are the chances that we might mistake one for the other? The relative entropy can be used to bound the probability, *E*, that we would obtain a sequence of reports corresponding to the expected sequences for the distribution, *P*, when sampling \mathcal{N} times from the distribution *Q*,

$$E \leqslant (\mathcal{N}+1)^D 2^{-\mathcal{N}S(P|Q)}. \tag{32}$$

This result is called Sanov's theorem [41]. In our above example, there is no chance, even on a single transmission, of mistaking the forecast for a town that always rains for a town that is always sunny.

Still considering a situation where the only two possible forecasts are Sun or rain, we consider a place P which has, on any given day, a fifty percent chance of either, and compare it to a place Q where the probability of Sun is λ . If we take the limit $\lambda \to 1$, then the relative entropy $S(P|Q) \to \infty$. The asymmetry of the relative entropy is apparent here, as S(Q|P) = 1. We interpret this to mean that there is zero chance that the typical sequence of N reports from P would be recovered by Q if Q is always sunny. By contrast, there is a chance, which decays with 2^N , that the typical sequence of Q will be produced by P, albeit one that is vanishingly small.

This notion of distinguishability is at the heart of information geometry. Intuitively, the more easily distinguishable two distributions, the further they ought to be. This geometry is interesting because it arises naturally in the course of scientific investigation. The example of weather modelling is just one example, and will see several more in the quantum context. But first, we must explain exactly how distinguishability can be used as a metric.

The relative entropy is, asymmetric, and thus unsuitable as a metric in and of itself. However, we can do slightly better if we consider the case where *P* is varied only infinitesimally. Concretely we define the distributions, $P = \{p_i\}$ and $P + dP = \{p_i + dp_i\}$, and expand the relative entropy in powers of dp_i ,

$$S(P|P+dP) = \sum_{i=1}^{D} p_i \ln\left(\frac{p_i}{p_i+dp_i}\right) \approx \frac{1}{2} \sum_{i=1}^{D} \frac{dp_i dp_i}{p_i}.$$
(33)

Here we have changed to measuring our relative entropy in nats for convenience. This infinitesimal expansion is symmetric, with the asymmetry of the relative entropy entering only at higher order. If we imagine that this infinitesimal expansion is a distance, then we can infer the line element,

$$ds^{2} = \sum_{i,j=1}^{D} F_{ij} dp_{i} dp_{j} = \frac{1}{4} \sum_{i,j=1}^{D} \frac{dp_{i} dp_{i}}{p_{i}}$$
(34)

with the Fisher-Rao metric,

$$F_{ij} = \frac{1}{4} \frac{\delta_{ij}}{p_i}.$$
(35)

We can normalize the metric in any way we like, but the choice of $\frac{1}{4}$ is particularly convenient. To see why, let us make a change of variables $u_i = \sqrt{p_i}$. Now we find, that the line element is,

$$\mathrm{d}s^2 = \sum_{i,j=1}^D \mathrm{d}u_i \,\mathrm{d}u_i \tag{36}$$

subject to the constraint,

$$1 = \sum_{i=1}^{D} u_i^2$$
 (37)

coming from the normalization of the probability distribution. We recognize immediately the equation for an *N*-sphere of unit radius. The constraint that $0 \le u_i \le 1$ confines us to the first octant of this sphere. Each point on the octant corresponds to a probability distribution, and the geodesic distance between two distributions in terms of the probabilities is simply,

$$D_{\rm B}(P,Q) = \cos^{-1}\left(\sum_{i=1}^{D} \sqrt{p_i q_i}\right) \tag{38}$$

where we recognize the argument of the inverse cosine as the standard dot product in terms the u_j variables. We call this dot product the *classical fidelity* and the corresponding distance the *Bhattacharyya distance* [42]. From equation (38) we can see that pure states are always separated by $\frac{\pi}{2}$. Geometrically, the fact that all classical mixtures of pure points are unique means that the space of classical probability distributions is a simplex [43]. We make these points because they will no longer be true when we pass from classical to quantum mechanics.

The Fisher-Rao metric is unique in the sense that it is the only Riemannian metric on the space of probability distributions that is also monotone. The meaning of this is as follows. Let *T* be a stochastic map, then a monotone function is one for which, given any distributions *P* and *Q*, $D_B(TP, TQ) \leq D_B(P, Q)$. This property corresponds to the introduction of noise over the channels connecting our laboratory to the field stations. As the channels become noisier, all the incoming transmissions should become less distinguishable. If distinguishability is our notion of distance, then under such stochastic mappings, $\lim_{n\to\infty} D_B(T^nP, T^n, Q) \to 0$.

To complete the connection between the Fisher-Rao metric and the Fisher information derived in the previous section, let us re-express the Fisher-Rao metric in terms of the parameters $\vec{\lambda}$. By the chain rule, the metric in equation (35) induces a metric on the space of parameterizations as,

$$F_{\mu\nu} = \frac{1}{4} \sum_{i,j=1}^{D} \frac{\partial p_i}{\partial \lambda_{\mu}} \frac{\partial p_j}{\partial \lambda_{\nu}} F_{ij}.$$
(39)

Using our definition of the Fisher-Rao metric we arrive at the *classical Fisher information metric* (CFIM), which we give in both discrete and continuous form,

$$F_{\mu\nu} = \frac{1}{4} \sum_{i=1}^{D} \frac{1}{p_i(\vec{\lambda})} \frac{\partial p_i(\vec{\lambda})}{\partial \lambda_{\mu}} \frac{\partial p_i(\vec{\lambda})}{\partial \lambda_{\nu}},\tag{40a}$$

$$F_{\mu\nu} = \frac{1}{4} \int_{\Omega} \mathrm{d}x \frac{1}{p(x|\vec{\lambda})} \frac{\partial p(x|\vec{\lambda})}{\partial \lambda_{\mu}} \frac{\partial p(x|\vec{\lambda})}{\partial \lambda_{\nu}}.$$
(40b)

here Ω is sample space if x is continuous [32]. Up to the normalization, this metric is the Fisher information that appears in the Cramér-Rao bound from equation (27). From now on we will write $\frac{\partial}{\partial \lambda} \equiv \partial_{\mu}$

There are two other ways of expressing the metric above that we will find useful later. The first takes advantage of the log-likelihood introduced in equation (17). Notice that equations (40) are, in fact, covariances in the log-likelihood's partial derivatives (recalling equation (25)),

$$F_{\mu\nu} = \operatorname{Cov}_P \left\{ \partial_\mu \ell, \partial_\nu \ell \right\}. \tag{41}$$

The covariance can be taken as an inner product, and the derivatives of the log-likelihoods as tangent vectors, allowing us to recover a familiar form of the metric from differential geometry. What is the base space? The Shannon information is the mean of the log-likelihoods, and so we can imagine that the log-likelihood of a probability p_i is the information conveyed by that particular event. The derivatives of the log-likelihood are then tangent vectors on this space of information.

For certain special choices of coordinate system, it becomes possible for us to represent the metric in terms of a potential function. While these coordinate systems may not always be physically relevant, they can always be defined over a given space of probability distributions. For example, we might parametrize the space of probability distributions by creating linear mixtures of some subset of probability distributions (for example, the pure points). This is the case of the *mixture family* of probability distributions,

$$p(x|\vec{\lambda}) = \sum_{\mu=1}^{N} \lambda_{\mu} p_{\mu}(x), \qquad (42)$$

where $p_{\mu}(x)$ are some probability distributions (possible discrete or continuous, possibly mixed or pure) and the parameters λ_{μ} are chosen such that normalization is respected. The two-forecast distribution discussed above can be considered as a mixture family by taking $p_1 = \{p(Sun) = 1, p(Rain) = 0\}$ and $p_2 = \{p(Sun) = 0, p(Rain) = 1\}$. If one considers the mixture family in equation (42) then the metric can be given by the Hessian of the Shannon information [32],

$$F_{\mu\nu} = -\frac{1}{4}\partial_{\mu}\partial_{\nu}S. \tag{43}$$

This relationship is not true in general, in particular it fails whenever the second derivative of the probability distribution with respect to the parameters is non-zero. This equation can be checked for the two forecast example by comparing equations (29) and (30).

A similar statement can be made when one instead considers the exponential family of distributions,

$$p(x|\vec{\lambda}) = q(x) \exp\left\{\sum_{\mu=1}^{N} \lambda_{\mu} H_{\mu}(x) - \psi(\vec{\lambda})\right\},\tag{44}$$

which are especially relevant in thermodynamics. Here, *x* is some state in the phase space Ω , q(x) gives the distribution at $\vec{\lambda} = 0$ and $\psi(\vec{\lambda})$ is the normalization which plays the role of the thermodynamic potential or free energy,

$$\psi(\vec{\lambda}) = \ln\left(\int_{\Omega} \exp\left\{\sum_{\mu=1}^{N} \lambda_{\mu} H_{\mu}\right\} q(x) \mathrm{d}x\right). \tag{45}$$

In this case we can treat $\psi(\vec{\lambda})$ as the potential for the metric via [14],

$$F_{\mu\nu} = -\frac{1}{4} \partial_{\mu} \partial_{\nu} \psi(\vec{\lambda}) \tag{46}$$

For a detailed discussion on the metric potential one can see [32]. For a discussion of the role of that this relationship plays in classical thermodynamics, one can refer to [14].

2.2. Quantum information geometry

We now generalize the concepts developed above in a classical setting to the quantum case.

2.2.1. von Neumann entropy

We first consider the quantum analogue of the Shannon information, which will allow us to ultimately quantify the entanglement between subsystems. Starting from equation (19), we can apply a heuristic mapping from the classical to the quantum,

$$p(x) \to \hat{\rho}, \qquad \Lambda(x) \to \hat{\Lambda}, \qquad \int (\cdot) \, \mathrm{d}x \to \mathrm{Tr}\{\cdot\}.$$
 (47)

We immediately arrive at the von Neumann entropy,

$$S = -\text{Tr}\{\hat{\rho}\ln(\hat{\rho})\}\tag{48}$$

where we measure the information in nats. In doing so, we introduce our convention that the quantum analogue of a classical information measure is written in script font. We can see that the von Neumann entropy will be zero for pure quantum states, just as the Shannon entropy was zero for pure classical states.

At first glance, the von Neumann entropy does not seem particularly useful for detecting entanglement, since it will give zero for all pure states. In fact, this is exactly its utility. If the von Neumann entropy of the subsystem state $\hat{\rho}_V^{\Psi}$ is non-zero, then we know that the subsystem V is entangled with its complement, \bar{V} as per our earlier discussion. For this reason, the von Neumann entropy of a subsystem S_V is called the *entanglement entropy*. Useful in this interpretation is the fact that $S_V = S_{\bar{V}}$, that is, V is entangled with \bar{V} exactly as much as \bar{V} is entangled with V. This is true even if V and \bar{V} are not the same size, as long as they form a disjoint and complete partition of the system.

To see how the entanglement entropy works, we can consider the simple example of the entanglement between the two spins for the state,

$$|\mathrm{GHZ}_N\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle^{\otimes N} + |\downarrow\rangle^{\otimes N}\right) \tag{49}$$

with N = 2. In this case, the von Neumann entropy between the subsystems will be 1 bit.

The entanglement entropy has seen extensive application in quantum information and many-body physics. A significant triumph is the scaling laws of entanglement entropy with subsystem size, which characterize the central charge of conformal field theories in models where such theories describe the low energy degrees of freedom [44–46]. For many-body systems with local interactions, the entanglement entropy tends to scale with the area of the subsystem under consideration [47]. This fact is crucial to the efficiency of the density matrix renormalization group family of algorithms [48].

Taking equation (48) as the expectation value of the operator $-\ln(\hat{\rho})$, we define the latter as the entanglement Hamiltonian,

$$\hat{H}_{\rm ent} = -\ln(\hat{\rho}). \tag{50}$$

Phases exhibiting symmetry protected topological order have been shown to exhibit non-trivial degeneracies in the spectrum of the entanglement Hamiltonian (the entanglement spectrum) [49, 50]. Another significant aspect of the entanglement entropy is that, for phases with long range topological order in two dimensions and greater, the area law scaling of the entanglement entropy is modified by a negative subleading correction, the *topological entanglement entropy* [51].

These facts have made the entanglement entropy ubiquitous as a theoretical probe of quantum fluctuations and correlations. Experimentally, however, the application of equation (48) is quite limited. In particular, computing the entanglement entropy requires complete knowledge of the density matrix, something which is only possible in highly controlled experimental settings.

2.2.2. QFI

When generalizing the concept of Fisher information to the quantum case, we must address the fact that quantum mechanics allows us to choose between sets of measurements that are decoherent with one another. For example, we might orient our Stern–Gerlach devices along any spin axis we like, and in doing so will be sampling from distinct probability distributions corresponding to the different choices of measurement basis. To formalize this, we introduce the concept of the positive operator valued measure (POVM). See for instance [52, 53]. Let $\{|\xi\rangle\langle\xi|\}_{\xi=1}^{M}$ be a set of positive semi definite operators (hereafter projectors) acting on the Hilbert space satisfying the resolution of the identity,

$$\mathbb{I} = \sum_{\xi} |\xi\rangle\langle\xi| \tag{51}$$

In general we do not require that the projectors be orthogonal as they would be if they represented the spectrum of a Hermitian operator. Equation (51) represents the most general possible kind of measurement that can be made on a quantum state. For each set of observables that commute with each other, there exists a corresponding POVM which we might label with an additional superscript, α . Referring back to figure 1(b), each of the probability distribution branching off of the density matrix correspond formally to,

$$p(\xi^{\alpha}) = \operatorname{Tr}\{|\xi^{\alpha}\rangle\langle\xi^{\alpha}|\hat{\rho}\}\tag{52}$$

When considering the quantum generalization of the Fisher information, we see that we would in fact have a QFI for each possible choice of POVM. What is conventionally called the QFI then corresponds to maximizing the classical Fisher information over all possible choices of POVM [10, 54]. We make this point only briefly because it is relevant in the case of quantum sensing, which we return to later. For now, we move straight to the quantum generalization of the geometrical structure introduce in section 2.1.3.

2.2.3. Quantum geometric tensor

In contrast to classical states, quantum states can vary in two distinct ways. To see this we consider a general mixed state depending on the parameter λ ,

$$\hat{\rho}(\lambda) = \sum_{j} p_{j}(\lambda) |j(\lambda)\rangle \langle j(\lambda)|.$$
(53)

A change in the parameter λ might cause a change in the eigenvalues and/or a change in the eigenvectors. The former of these changes induces a distance identical to the kind that we considered in the classical case. Those parameterizations that change the eigenvectors give rise to quantum information geometry. We can study this case in detail by first considering a pure state, $|\Psi(\vec{\lambda})\rangle$. Now we can expand this state to leading order for a small change in the component of the parameter λ_{μ} ,

$$|\Psi(\lambda_{\mu} + d\lambda_{\mu})\rangle = |\Psi(\lambda_{\mu})\rangle + d\lambda_{\mu}|\partial_{\mu}\Psi(\lambda)\rangle, \tag{54}$$

where $\partial_{\mu} \equiv \partial_{\lambda_{\mu}}$. Intuitively, the distance between states is the size of the change in the state. We need to take care however to ensure that our notion of distance is independent of the global phase, since physical states are defined by rays in the Hilbert space. We thus define the quantum geometric tensor (QGT),

$$Q_{\mu\nu} = \langle \partial_{\mu} \Psi | (\mathbb{I} - |\Psi\rangle \langle \Psi |) | \partial_{\mu} \Psi \rangle$$
(55)

where we have subtracted off the projection of the change in the state $|\partial_{\mu}\Psi\rangle$ onto the subspace of the unchanged state defined by the projector $|\Psi\rangle\langle\Psi|$.

The object in equation (55) admits a decomposition into a real symmetric component and an imaginary antisymmetric component,

$$Q_{\mu\nu} = \mathcal{F}_{\mu\nu} + i\Omega_{\mu\nu}.$$
(56)

The real part, $\mathcal{F}_{\mu\nu}$ gives the Riemannian metric on the quantum state space [33, 34], while the imaginary antisymmetric part, $\Omega_{\mu\nu}$ gives the Berry curvature [3]. From this point on, we focus in particular on the real symmetric component, which we term the QFI metric (QFIM). The reason for this decomposition is that the Hilbert space in which we have formulated our geometry is complex. Consequently, the metric tensor is required to be Hermitian, but not purely real. The imaginary component will not contribute to the line element due to the fact that it is antisymmetric. The state space thus has two parallel geometric structures. On the one hand, the real Riemannian structure, $\mathcal{F}_{\mu\nu}$, and on the other hand the symplectic structure determined by $\Omega_{\mu\nu}$

The distance on the space of pure states is a direct analogue of the Bhattacharyya distance in equation (38) [9],

$$D_{\rm QB}(\Psi_1, \Psi_2) = \cos^{-1}(|\langle \Psi_1 | \Psi_2 \rangle|)$$
(57)

where the argument of the inverse cosine is now the quantum analogue of the fidelity. In [55], the quantum fidelity is defined at $|\langle \Psi_1 | \Psi_1 \rangle|^2$ in place of the classical fidelity. The distance between pure states is sometimes called the quantum angle. Where classically, pure states are always separated by the maximal distance of $\frac{\pi}{2}$, now it is only orthogonal pure states that are separated by $\frac{\pi}{2}$. Instead of the positive octant of a hypersphere, the space of quantum pure states can be visualized as a Bloch hypersphere, with anti-podal points identified as orthogonal states.

All physical pure states must be normalized, and so in the pure state limit only unitary parameterizations are relevant. Without loss of generality, we can always express a unitary parameterization via the Schrödinger equation,

$$\partial_{\mu}|\Psi\rangle = -i\hat{\Lambda}_{\mu}|\Psi\rangle \tag{58}$$

where $\hat{\Lambda}_{\mu}$ is a Hermitian operator corresponding to a change in the parameter λ_{μ} . By substituting equation (58) into equation (55) we see that the QFIM for unitary parameterizations is given by,

$$\mathcal{F}_{\mu\nu}^{\text{unitary}} = 4\text{Re}[\text{Cov}_{\Psi}\{\hat{\Lambda}_{\mu}, \hat{\Lambda}_{\nu}\}]$$
(59)

For mixed states, unitary parameterizations are slightly more subtle. As a first attempt, we may try to take equations (40) and directly apply the mapping in equation (47). The problem with this is that, in general, $[\hat{\rho}, \partial_{\mu}\hat{\rho}] \neq 0$. Resolving this ambiguity is, to a degree, a matter of convention. We proceed along the lines of Braunstein and Caves [10], who resolve this ambiguity by introducing the superoperator,

$$\mathcal{L}_{\hat{\rho}}^{-1}\{\hat{O}\} = \frac{1}{2}(\hat{\rho}\hat{O} + \hat{O}\hat{\rho}).$$
(60)

In the eigenspace of the density matrix, the inverse of this operator is given by,

$$\mathcal{L}_{\hat{\rho}}(\hat{O}) = \sum_{j,k} \frac{2}{p_j + p_k} O_{jk} |j\rangle \langle k|, \qquad (61)$$

with the sum running only over terms with $p_j + p_k > 0$. Notice that the above form is the inverse of equation (60) in the sense that $\mathcal{L}_{\hat{\rho}} \{ \mathcal{L}_{\hat{\rho}}^{-1} \{ \hat{O} \} \} = \hat{O}$. Applying the inverse to the derivative of the density matrix,

$$\mathcal{L}_{\hat{\rho},\mu} = \mathcal{L}_{\rho} \left(\frac{\partial \hat{\rho}}{\partial \lambda_{\mu}} \right) \tag{62}$$

Table 1. A partial reproduction of the table in [25] listing several different forms of the QFIM along with the associated monotone function.

QFIM	Monotone function
Symmetric logarithmic derivative Bogoliubov-Kubo-Mori	$\frac{\frac{1+x}{2}}{\frac{x-1}{\log(x)}}$
Right logarithmic derivative	x
Left logarithmic derivative	1
Skew information ($\alpha = 1$)	$\frac{(\sqrt{x}+1)^2}{4}$

gives us a logarithmic derivative wherein the ambiguity in the ordering of the operators is resolved symmetrically, hence we call equation (62) the *symmetric logarithmic derivative* (SLD). Now we can write the QFIM for mixed states as,

$$\mathcal{F}_{\mu\nu} = 4\text{Re}\left[\text{Cov}_{\hat{\rho}}\left\{\mathcal{L}_{\hat{\rho},\mu}, \mathcal{L}_{\hat{\rho},\nu}\right\}\right].$$
(63)

The above equation can be taken as the quantum analogue of equation (41). It includes the special case where only the eigenvalues of the density matrix are modified, and in that instance reduces to the classical Fisher information matrix (CFIM). Parameterizations that modify the eigenvectors of the density matrix are captured by the von Neumann equation,

$$\partial_{\mu}\hat{\rho} = i \left| \hat{\rho}, \hat{\Lambda}_{\mu} \right|. \tag{64}$$

which described unitary evolutions generated by the Hermitian operator $\hat{\Lambda}_{\mu}$

As we mentioned, the choice of resolving the ambiguity in ordering through the SLD is just one choice. Morozova and Chentsov [56] demonstrated that for every operator monotone function, we can define a metric on the space of mixed states. Skipping the rigorous details, we present here the relationship between metrics and operator monotone functions. First we introduce the superoperators [25],

$$R_{\hat{\rho}}(\hat{O}) = \hat{O}\hat{\rho} \tag{65a}$$

$$L_{\hat{\rho}}(\hat{O}) = \hat{\rho}\hat{O} \tag{65b}$$

$$(\mathcal{L}_{\hat{\rho}}^{f})^{-1} = f(L_{\hat{\rho}}R_{\hat{\rho}}^{-1})R_{\hat{\rho}}.$$
(65c)

Here f is the operator monotone function. For the case of the SLD, we can find that,

$$f^{\rm SLD}(x) = \frac{1+x}{2}.$$
 (66)

There are many other possible choices (we list several in table 1). These types of operator monotones were studied extensively in the work of Dénes Petz, see [57–59].

The QFIM is distinguished by the fact that it finds the maximal distance between states, and is, in this sense, unique [32].

The SLD QFIM can be used to define the Bures distance between mixed states in analogy with equation (57) [60],

$$D_{\text{Bures}}(\hat{\rho}_1, \hat{\rho}_2) = \cos^{-1}\left(\sqrt{F(\hat{\rho}_1, \hat{\rho}_2)}\right),$$
(67)

where

$$F(\hat{\rho}_1, \hat{\rho}_2) = \operatorname{Tr}(\sqrt{\sqrt{\hat{\rho}_1}\hat{\rho}_2\sqrt{\hat{\rho}_1}})^2$$
(68)

is the mixed state generalization of the quantum fidelity, often termed the Uhlmann fidelity [61]. The above expression can be interpreted as the transition amplitude between mixed states, in exactly the same way as the quantum fidelity can be interpreted as the transition amplitude between pure states. Taking the SLD QFIM for granted, the Bures distance is the broadest generalization of Riemannian distance on the quantum state space.

Focusing on the symmetric part of equation (55) it is useful to explicitly decompose the QFIM into a component that arises due to changes in the eigenvectors of the density matrix (the unitary component), and

a component that arises due to changes in the eigenvalues (the non-unitary component) (see for example the supplementary material of [24]),

$$\mathcal{F}_{\mu\nu} = \mathcal{F}_{\mu\nu}^{\text{non-unitary}} + \mathcal{F}_{\mu\nu}^{\text{unitary}},\tag{69}$$

where,

$$\mathcal{F}_{\mu\nu}^{\text{non-unitary}} = \frac{1}{4} \sum_{j} \frac{1}{p_j} \partial_{\mu} p_j \partial_{\nu} p_j, \tag{70}$$

and

$$\mathcal{F}_{\mu\nu}^{\text{unitary}} = 2\sum_{j,k} \frac{(p_j - p_k)^2}{p_j + p_k} [\Lambda_{\mu}]_{jk} [\Lambda_{\nu}]_{kj}^*.$$
(71)

Equation (70) is the direct analogue of equations (40). Equation (71) has no direct classical analogue. It emerges specifically due to the non-commutativity of quantum observables. Often times, the expression in equation (71) is called the QFI, due to this fact. To maintain our symmetry with the notation used in the discussion of von Neumann entropy, we prefer to call the entire metric, containing both unitary and non-unitary components the QFIM, and instead refer to the unitary and non-unitary components as needed. Because the Berry curvature is antisymmetric, there is no such ambiguity.

2.2.4. Topological invariants and volumes

The QFIM allows for the construction of two distinct but related topological invariants. We discuss these invariants specifically in the case of a two-dimensional submanifold of states, \mathcal{M} , which might be given by the parameters of a Hamiltonian, or by the first Brillouin zone.

The first emerges from the Berry curvature, which, when integrated over a two-dimensional manifold, gives the Chern number of the manifold,

$$C = \frac{1}{2\pi} \int_{\mathcal{M}} \Omega_{\mu\nu} d\lambda_1 d\lambda_2.$$
(72)

The second is the Euler characteristic, which is given by Gauss-Bonnet theorem [62],

$$\chi = \frac{1}{2\pi} \left(\int_{\mathcal{M}} K \, \mathrm{d}A + \oint_{\partial \mathcal{M}} k_g \, \mathrm{d}\ell \right). \tag{73}$$

In the above expression, K is the *Gaussian curvature* and k_g the *geodesic curvature* of the boundary. In order to express these explicitly in terms of the QFIM, we model our notation on [63]. Let us write the metric in first fundamental form, by expressing the line element as

$$\mathrm{d}s^2 = E\mathrm{d}\lambda_1^2 + 2F\mathrm{d}\lambda_1\mathrm{d}\lambda_2 + G\mathrm{d}\lambda_2. \tag{74}$$

We denote the determinant of the metric by omitting the Greek indices,

$$\mathcal{F} = \det \mathcal{F}_{\mu\nu} = EG - F^2. \tag{75}$$

We also recall the Christoffel symbols,

$$\Gamma^{k}_{\ ij} = \frac{1}{2} \mathcal{F}^{km} \left(\partial_{j} \mathcal{F}_{im} + \partial_{i} \mathcal{F}_{jm} - \partial_{m} \mathcal{F}_{ij} \right).$$
(76)

Now we can express the Gaussian curvature and the geodesic curvature,

$$K = \frac{1}{\sqrt{\mathcal{F}}} \left(\partial_2 \left(\frac{\sqrt{\mathcal{F}} \Gamma^2_{11}}{E} \right) - \partial_1 \left(\frac{\sqrt{\mathcal{F}} \Gamma^2_{12}}{E} \right) \right)$$
(77*a*)

$$k_g = \sqrt{\mathcal{F}} G^{-\frac{3}{2}} \Gamma^1{}_{22}, \tag{77b}$$

along with the line and area elements,

$$\mathrm{d}A = \sqrt{\mathcal{F}}\mathrm{d}\lambda_1\mathrm{d}\lambda_2 \tag{78a}$$

$$d\ell = \sqrt{G} d\lambda_2 \tag{78b}$$

The metric can of also be used to evaluate the volume of the parameter manifold in *n* dimensions via,

$$\operatorname{Vol}_{g} = \int \sqrt{\operatorname{det}(g)} d\lambda_{1} \wedge d\lambda_{2} \wedge \ldots \wedge d\lambda_{n}$$
(79)

In sections 3.6 and 3.7 we offer further discussion of the analysis of these topological invariants in several systems.

3. Information geometry in practice

So far, we have laid out the geometrical structure of the quantum state space. In this section, we try to understand what this structure means from a physical perspective, with a particular focus on the applications of this structure to many-body systems. There are two main topics to be addressed here. The first is understanding what, in physical terms, the quantum information geometry is actually describing. A hint is already given by the fact that the metric is the real part of the covariance. The second is understanding several fundamental choices of parameter manifold, and the contexts in which these emerge.

For the reader's convenience, we identify three main categories of manifolds that appear in the literature. The first is the Brillouin zone in a band system parametrized by the momenta (k_x, k_y) with $k_\alpha \in [0, 2\pi)$. The non-trivial topology of the band structure, which leads to the classification of topological phases of matter, emerges from the symplectic component of the QGT: the Berry curvature. Several results in this direction have found interested geometrical properties associated with this geometry [21, 64–71].

In the second case are manifolds of ground state parametrized by fields and couplings in a parent Hamiltonian (see, for example, references [4, 5, 72–86]). This kind of geometry emerges especially in the context of the Fidelity susceptibility, and we treat in some details the structure of the QGT in section 3.4. Finally, in the context first of metrology and increasingly in many body physics, are manifolds generated by the application of some unitary operator $\hat{U} = exp\{-i\lambda\hat{\Lambda}\}$ to the state $|\psi\rangle$ with $\hat{\Lambda}$ a bounded Hermitian operator with spectrum.

In the case of metrology, $\hat{\Lambda}$ might play the role of the coupling between the probe state and the system being studied, with λ the parameter to be estimated. In the context of many body spin systems, $\hat{\Lambda}$ is usually a sum over operators defined on each degree of freedom in the system, such as the collective spin operator,

$$\hat{\Lambda} = \sum_{j} S_{j}^{\alpha} \tag{80}$$

along some particular axis $\alpha \in \{x, y, z\}$ (see for example [19, 20, 24, 87–95]). In this case the components of the quantum metric correspond to the integrated dynamical response of the system. Operators of the form given in equation (80) are called local, because the summand depends on a bounded number of sites even in the thermodynamic limit. Along this line of reasoning, [96] found super-extensive scaling of the QFI in the topologically non-trivial Kitaev wire by choosing a non-local operators,

$$\hat{\Lambda}^{(\alpha)} = \sum_{j} (-i)^{\delta_{\alpha,x}} \left(\hat{a}_{j}^{\dagger} e^{i\pi \sum_{i=1}^{j-1} \hat{n}_{i}} + (-1)^{\delta_{\alpha,y}} e^{-i\pi \sum_{i=1}^{j-1} \hat{n}} \hat{a}_{j} \right)$$
(81)

where \hat{a}_j is the fermionic annihilation operator and \hat{n} is the number operator. We discuss the relationship between these three branches of research more in the following sections.

3.1. Relationship to linear response

Considering the fact that the QFIM quantifies the degree of distinguishability along the path generated by the hermitian operator $\hat{\Lambda}_{\mu}$, it is not surprising that the unitary component of QFI has a convenient representation in terms of the dissipative part of the linear response [24],

$$\mathcal{F}_{\mu\nu}^{\text{unitary}} = \frac{4}{\pi} \int_0^\infty d\omega \tanh\left(\frac{\omega}{2T}\right) \chi_{\mu\nu}^{\prime\prime}(\omega). \tag{82}$$

where,

$$\chi(\omega) = \int_0^\infty \mathrm{d}t e^{-i\omega t} \mathrm{Tr}\left\{\hat{\rho}\left[\hat{\Lambda}_\mu(t), \hat{\Lambda}_\nu\right]\right\}.$$
(83)

Analogous formulae can be shown to hold for all QFIM's (that is, for any choice of monotone function f). The general relationship is given by,

$$\mathcal{F}_{\mu\nu}^{\text{unitary}}\{f\} = \frac{2}{\pi} \int_0^\infty \mathrm{d}\omega \, \frac{1 - e^{-\beta\omega}}{f(e^{-\beta\omega})} \chi_{\mu\nu}^{\prime\prime}.$$
(84)

The connection between quantum geometry and response functions has been exploited in a number of proposed [22, 23, 97] and realized [70, 98] experiments.

This connection is particularly significant in light of the relationship between quantum geometry and entanglement, which we discuss further in section 3.3. In particular, the low temperature properties of the entanglement are strongly controlled by the static structure factor, as demonstrated in [88, 99].

3.2. The quantum variance

In the introduction, we promised that quantum information geometry would offer us a means by which to separate the classical fluctuations from the quantum fluctuations in a given system. The quantum covariance [100, 101] makes this concept explicit by decomposing the covariance between two operators as,

$$\operatorname{Cov}_{\hat{\rho}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\} = \operatorname{Cov}_{\hat{\rho}}^{\mathcal{Q}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\} + \operatorname{Cov}_{\hat{\rho}}^{\mathcal{C}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\}$$
(85)

where $\operatorname{Cov}_{\hat{\rho}}^{\mathcal{Q}}\{\cdot,\cdot\}$ is the quantum component of the covariance and $\operatorname{Cov}_{\hat{\rho}}^{\mathcal{C}}\{\cdot,\cdot\}$ is the classical component of the covariance.

To define the classical covariance, we begin by considering the mixed state defined in equation (9),

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{\mathcal{Z}} \tag{86}$$

where,

$$\mathcal{Z} = \text{Tr}\{e^{-\beta \hat{H}}\}\tag{87}$$

which can be used to describe any mixed state for a suitable choice of \hat{H} . Now add to the Hamiltonian the operators $\hat{\vec{\Lambda}}$ which are coupled via the parameters $\hat{\vec{\lambda}}$

$$\hat{\rho}(\vec{\lambda}) = \frac{e^{-\beta(\hat{H} + \vec{\lambda} \cdot \vec{\Lambda})}}{\mathcal{Z}}$$
(88*a*)

$$\mathcal{Z}(\vec{\lambda}) = \operatorname{Tr}\{e^{-\beta(\hat{H}+\vec{\lambda}\cdot\vec{\Lambda})}\}.$$
(88b)

Using the free energy,

$$F = -\frac{1}{\beta} \ln(\mathcal{Z}(h)).$$
(89)

we can now define the classical covariance as,

$$\operatorname{Cov}_{\hat{\rho}}^{\mathcal{C}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\} = \frac{1}{\beta} \frac{\partial^2 F}{\partial \lambda_{\mu} \partial \lambda_{\nu}}$$
(90)

where the second derivative of the free energy can be recognized as the static thermal susceptibility. The quantum covariance is then defined directly by rearranging equation (85),

$$\operatorname{Cov}_{\hat{\rho}}^{\mathcal{Q}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\} = \operatorname{Cov}_{\hat{\rho}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\} - \operatorname{Cov}_{\hat{\rho}}^{\mathcal{C}}\{\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}\}$$
(91)

To understand the meaning of this definition, let us first consider the case where $[\hat{H}, \vec{\Lambda}] = 0$. From the perspective of information geometry, this kind of perturbation will change the eigenvalues of the density matrix, but not the eigenvectors. In this case, the classical covariance will be precisely the covariance, and the quantum covariance will be zero.

In the case where the commutator of the Hamiltonian with the perturbation is non-zero. The evaluation of the second derivative in equation (90) must now be performed by taking a Taylor series expansion of the exponential about the point where $\vec{\lambda} = \vec{0}$. In general the thermal covariance will be strictly less than the total covariance (except at infinite temperature), and the difference is made up for by the quantum fluctuations

quantified by the quantum covariance. Notice that the quantum variance is just the diagonal of the quantum covariance matrix,

$$\operatorname{Var}_{\hat{o}}^{\mathcal{Q}}\{\hat{\Lambda}_{\mu}\} \equiv \operatorname{Cov}_{\hat{o}}^{\mathcal{Q}}\{\hat{\Lambda}_{\mu}, \hat{\Lambda}_{\mu}\}$$

$$\tag{92}$$

The physical meaning of the quantum covariance at finite temperature can be gleaned by considering the case of a free particle in equilibrium with a thermal bath. The quantum variance associated with the position operator is then [100],

$$\operatorname{Var}_{\hat{\rho}}^{\mathcal{Q}}(\hat{x};\hat{\rho}) = \frac{1}{24\pi} \lambda_{dB}^2$$
(93)

where $\lambda_{dB} = \sqrt{\frac{2\pi\hbar^2}{mk_BT}}$ is the *thermal de Broglie wavelength*. It decreases monotonically to zero as the temperature increases, and thus indicates the falling off of quantum fluctuations at finite temperature.

The quantum covariance appears to be behaving exactly as we would expect the unitary component of the QGT to behave. To better understand the relationship between the two, we recall a special case of the QGT, the Wigner-Yanase-Dyson (WYD) skew information, [102, 103],

$$\mathcal{I}_{\alpha,\hat{\rho}}(\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}) = -\frac{1}{2} \operatorname{tr}\left(\left[\hat{\Lambda}_{\mu},\hat{\rho}^{\alpha}\right]\left[\hat{\Lambda}_{\nu},\rho^{\hat{1}-\alpha}\right]\right).$$
(94)

This QGT corresponds to the following choice of monotone function [25],

$$f_{\alpha}(x) = \alpha (1 - \alpha) \frac{(x - 1)^2}{(x^{\alpha} - 1)(x^{1 - \alpha} - 1)}.$$
(95)

It is shown in [100] that the quantum covariance may be written as,

$$\operatorname{Cov}_{\hat{\rho}}^{\mathcal{Q}}(\hat{\Lambda}_{\mu},\hat{\Lambda}_{\nu}) = \int_{0}^{1} \mathrm{d}\alpha \, \mathcal{I}_{\alpha}(\hat{\Lambda};\hat{\rho}).$$
(96)

So we see that the quantum variance is the average over a particular family of QGT's. The diagonal entries of the SLD QFIM are bounded from above and below by the diagonal entries of the quantum covariance (recall that the Berry curvature does not have diagonal entries),

$$4\operatorname{Var}_{\hat{\rho}}^{\mathcal{Q}}(\hat{\Lambda}_{\mu}) \leqslant \mathcal{F}_{\mu\mu} \leqslant 12\operatorname{Var}_{\hat{\rho}}^{\mathcal{Q}}(\hat{\Lambda}_{\mu}) \tag{97}$$

Using the representation of the WYD information in terms of response functions, equation (96) can be written as [104],

$$\operatorname{Var}_{\mathcal{Q}}(\hat{\Lambda};\hat{\rho}) = \hbar \int_{0}^{\infty} \frac{\mathrm{d}\omega}{\pi} L\left(\frac{\beta\hbar\omega}{2}\right) \chi_{\hat{\Lambda}}^{\prime\prime}(\omega) \tag{98}$$

where,

$$L(x) = \coth(x) - \frac{1}{x}$$
(99)

is the Langevin function. We thus see that the quantum variance corresponds directly to the unitary component of the QFIM described in equation (71), making explicit the fact that this component corresponds to quantum fluctuations by subtracting off the thermal fluctuations that would give rise to the non-unitary component of the QFIM in equation (70).

To understand the difference between the quantum variance and SLD QFIM we consider a single qubit in thermal equilibrium with a bath of inverse temperature β . We denote explicitly the gap of the qubit with Δ

$$\rho = \frac{1}{\mathcal{Z}} e^{-\beta \Delta \sigma^z}.$$
(100)

It is easy to rewrite the density matrix explicitly as,

$$\rho = \frac{1}{2} (\mathbb{I} - \tanh(|\Delta|\beta)\hat{\sigma}^z).$$
(101)

We consider parameterizations of the Hilbert space generated unitarily by the magnetization operator in direction \hat{n} ,

$$\Lambda(\theta,\phi) = \hat{n} \cdot \vec{\sigma} = \cos(\theta)\sin(\phi)\hat{\sigma}^x + \sin(\theta)\sin(\phi)\hat{\sigma}^y + \cos(\phi)\hat{\sigma}^z.$$
(102)



Below we calculate the variance, its quantum and thermal components, and the QFI as a function of β , θ , and ϕ . It is helpful for this purpose to compute the partition function with a source term,

$$\mathcal{Z}(h) = \operatorname{tr}\{\exp(-\beta(\Delta\hat{\sigma}^z - h\hat{\Lambda}(\theta, \phi)))\}.$$

The variance, QV (from equation (91)), and QFI (from equation (71)), for $\phi = \frac{\pi}{2}$ and $\phi = \frac{\pi}{4}$ are shown in figure 3 with details of the calculations given in appendix B.1. We can see that regardless of the orientation, the variance will always saturate at its maximal value in the limit of infinite temperature ($\beta \rightarrow 0$). If we imagine the Bloch sphere with its interior as the prototypical representation of the space of pure and mixed states, then the infinite temperature limit corresponds to the center of this sphere. From an information theory perspective, it is the equivalent of a black hole, as all states taken to this limit become indistinguishable. This manifests in the QFI going to zero. We see that the quantum variance also tends to zero. While both the QV and QFIM must go to zero in this limit, they do so at slightly different rates. Their difference is not constant but peaks at some value of β that scales with the energy gap of the system.

3.3. QFI and entanglement

Among the most significant aspects of the geometrical structure of quantum state space is its explicit relationship to multipartite which was, as far as we are aware, first made explicit in [105], and later elaborated on in [19, 20]. The diagonal components of the QFIM may be related explicitly to bounds on the amount of multipartite entanglement. These bounds are significant for three reasons. First, they apply even at finite temperature. Second, they are directly accessible through the linear response functions discussion in section 3.1. Third, the detection of entanglement depends in particular on the direction in state space that we choose to examine, thus allowing us to consider a richer structure of entanglement than would be accessible if we only considered the von Neumann entropy. Below we derive the bounds explicitly following the discussion in [20].

Consider an N body system and a Hermitian operator composed of a sum of one-body operators,

$$\hat{\Lambda} = \sum_{j=1}^{N} \hat{A}_j \tag{103}$$

each with maximum eigenvalue a_{max} , minimum eigenvalue a_{min} . Let $\Delta_a = a_{max} - a_{min}$ be the spectral width. The first bound we derive relies only on the component of the QFIM associated with the operator in equation (103), which we hereafter refer to simply as the QFI, \mathcal{F} , omitting the indices for simplicity. We also note the following

- (a) The set of *k*-producible states is convex. That is to say, the producibility is non-increasing under mixture.
- (b) The Fisher information is convex under mixture, $\mathcal{F}\{p\hat{\rho}_1 + (1-p)\hat{\rho}_2\} \leq p\mathcal{F}\{\hat{\rho}_1\} + (1-p)\mathcal{F}\{\hat{\rho}_2\}$ for $p \in [0,1]$
- (c) For a *N* body state $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ we have $\operatorname{Var}_{|\Psi\rangle}(\hat{\Lambda}) = \operatorname{Var}_{|\psi_A\rangle}(\hat{\Lambda}^{(A)}) + \operatorname{Var}_{|\psi_B\rangle}(\hat{\Lambda}^{(B)})$
- (d) For a *N* body pure state $|\Psi\rangle$, $4\text{Var}_{|\Psi\rangle}(\hat{\Lambda}) \leq \Delta_a N^2$ with the maximum achieved by the GHZ state,

$$|\text{GHZ}_N\rangle = \frac{1}{\sqrt{2}} \left(|a_{\min}\rangle^{\otimes N} + |a_{\max}\rangle^{\otimes N} \right).$$
 (104)

By (a) and (b) the maximal QFI will occur for pure states, so we limit our discussion to that case. Now we consider the *k* producible state,

$$|\Psi\rangle = \bigotimes_{\ell=1}^{M} |\psi_{\ell}\rangle \tag{105}$$

where each factor $|\psi_{\ell}\rangle$ contains N_{ℓ} particles and the largest N_l is equal to k. For a given value of k, the largest possible QFI is achieved by a state with $s = \lfloor \frac{N}{k} \rfloor$ factors in a $|\text{GHZ}_k\rangle$ state with the remaining particles in a $|\text{GHZ}_r\rangle$ state where r = N - sk. The QFI for a k-producible state is thus bounded from above,

$$\mathcal{F} \leqslant \Delta_a (sk^2 + r^2) \tag{106}$$

with the bound for separable states given by [105],

$$\mathcal{F}(\hat{\rho}_{\text{separable}}) \leqslant N.$$
 (107)

In [19, 20], a second inequality is derived, which is aimed specifically at systems whose degrees of freedom can be represented in terms of the spin. Let's consider the case where the local operators in equation (103) are given by,

$$A_j = \mathbf{n}_j \cdot \mathbf{S}_j \tag{108}$$

where $\mathbf{n}_j = (\cos(\theta_j)\sin(\phi_j), \sin(\theta_j)\sin(\phi_j), \cos(\phi_j))$ where $\theta \in [0, 2\pi)$ and $\phi \in [0, \pi)$. Let us consider the case where the orientation on each site is the same by dropping the subscript *j* on the unit vector. We denote the corresponding generator, $\hat{\Lambda}_{\mathbf{n}}$ and the corresponding QFI $\mathcal{F}_{\mathbf{n}}$. We can now consider the direction averaged QFI,

$$\bar{\mathcal{F}} = \int P(\mathbf{n}) \mathcal{F}_{\mathbf{n}} \sin(\phi) \mathrm{d}\theta \mathrm{d}\phi$$
(109)

where $P(\mathbf{n})$ is a function such that $1 = \int P(\mathbf{n}) \sin(\phi) d\theta d\phi$. If we choose the case of $P(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{a})$ we find the QFI for the particular orientation \mathbf{a} of the generator and the inequality in equation (106) would apply. Instead, we consider the uniform case where $P(\mathbf{n}) = \frac{1}{4\pi}$, for which

$$\bar{\mathcal{F}} = \frac{\mathcal{F}_{\mathbf{x}} + \mathcal{F}_{\mathbf{y}} + \mathcal{F}_{\mathbf{z}}}{3}.$$
(110)

In [20] the above expression is given the following bound for k-producible states,

$$\bar{\mathcal{F}}(\hat{\rho}_{k-\text{producible}}) = \frac{1}{3} \left(s \left(k^2 + 2k - \delta_{k,1} \right) + r^2 + 2r - \delta_{r,1} \right)$$
(111)

where $s = \lfloor \frac{N}{k} \rfloor$ and r = N - sk are defined as earlier. For fully separable states the above bound reduces to,

$$\bar{\mathcal{F}}(\hat{\rho}_{\text{separable}}) = \frac{2}{3}N.$$
(112)

To illustrate the applications of equations (106) and (112) we consider the case of 2 partite states. In the following discussion, we consider the QFI density, which we indicate with $f = \frac{F}{N}$ or $\overline{f} = \frac{\overline{F}}{N}$.





The extent to which the above inequalities detect entanglement depend sensitively on the choice of the operator $\hat{\Lambda}$ in equation (103), or equivalently, on the direction in state space in which we choose to perturb the state. To see this, consider the following two states,

$$|\psi(\alpha)\rangle = \cos(\alpha)|\uparrow_A\downarrow_B\rangle - \sin(\alpha)|\downarrow_A\uparrow_B\rangle, \qquad (113a)$$

$$|\phi(\alpha)\rangle = \cos(\alpha)|\uparrow_A\downarrow_B\rangle + \sin(\alpha)|\downarrow_A\downarrow_B\rangle \tag{113b}$$

on the interval $\alpha \in [0, \frac{\pi}{2}]$. At $\alpha = \frac{\pi}{4}$ the first state is the singlet, which we know is maximally entangled. This state is maximally entangled, containing one bit of entanglement entropy, as can be seen from the solid red curve in figure 4. The second state is separable for all possible values of λ and therefor has an entanglement entropy of zero. For both states, we can compute the QFI associated with the parameter λ and find that they are identical. The QFI associated with this parameter tells us nothing about the entanglement content of the state. It's easy to see that any single parameter family of states that interpolates smoothly between two orthogonal states in the form given about will always yield the QFI,

$$\mathcal{F}_{\alpha\alpha} = \sec^2(\alpha)\csc^2(\alpha). \tag{114}$$

To distinguish between entangled and separable spin states, we must consider a path generated by an operator that is local in the spin degrees of freedom. Even for two spins, there is considerable freedom in which operators we might choose. Two natural choices are the ferromagnetic and anti-ferromagnetic combinations of each of the three spin operators,

$$\hat{\Lambda}^{\pm}_{\mu} = \hat{S}^{\mu}_A \pm \hat{S}^{\mu}_B \tag{115}$$

where $\mu \in \{x, y, z\}$.

Each of these generators defines a unitary operator, which induces an infinitesimal displacement in the state along a path parameterized by λ ,

$$|\alpha, \mathbf{d}\lambda_{\mu}\rangle = e^{i\mathbf{d}\lambda_{\mu}\Lambda_{\mu}^{\pm}}|\alpha\rangle. \tag{116}$$

The distance between $|\lambda, 0\rangle$ and $|\lambda, dh\rangle$ is then given by (following equation (59)),

$$\mathcal{F}^{\pm}_{\mu\mu}\{\psi\} = 4\text{Cov}_{\psi}(\hat{\Lambda}^{\pm}_{\mu}, \hat{\Lambda}^{\pm}_{\mu}).$$
(117)

The QFI densities generated by the ferromagnetic and anti-ferromagnetic operators are give in figure 5.





While the separable state, $|\phi\rangle$ never exceeds the bound given in equation (107), we can see the for the entangled state $|\psi\rangle$ the bound is exceeded by three of the different paths chosen. First, for the path given in terms of the *z* component for the anti-ferromagnetic spin combination, bipartite entanglement is detected for a window the parameter space $\lambda \in \left[\frac{\pi}{8}, \frac{3\pi}{8}\right]$. The bound is also always violated for the parameterizations generated by the anti-ferromagnetic *x* and *y* operators for the entire span of λ except at the separable points.

The three ferromagnetic operators defined in equation (115) (or, equivalently, the three anti-ferromagnetic spin operators), are related to each other by global rotations. We can thus consider the QFIM corresponding to the three-dimensional submanifold whose parameterization is generated by these operators and take its trace. The results for the ferromagnetic and anti-ferromagnetic operators are given in figure 6. The bound corresponding to equation (112) is again given in gray. Here we see that only the anti-ferromagnetic spin combinations detect entanglement.

Notice that at the point, $\alpha = \frac{\pi}{4}$ all three components of the QFIM for the ferromagnetic combination are zero. The singlet is an eigenstate of the magnetization operator, having total magnetization zero in any of the three spin directions. This special case implies a general principle. If a state $|\psi\rangle$ is an eigenstate of the generator of a parameterization, then its QFI along that parameterization will be zero. This is easy to see if one considers the fact that the QFI for a pure state is just the covariance of the generator for that state.

More generally, if the Hamiltonian of a system commutes with the generator, then the ground state of that Hamiltonian will tend to exhibit zero QFI. An exception is in the case where the ground state is degenerate. For example, the ferromagnetic Heisenberg chain has a three-fold degenerate ground state manifold, which can be used to construct entangled states for which the QFI associated with the ferromagnetic generator does detect entanglement.



Figure 6. Trace of the QFIM matrix for the three symmetric (blue) and antisymmetric (red) generators defined in equation (115). The gray region indicates the bound on entanglement given in equation (112) for detecting 2-partite entanglement. Notice that only the symmetric choice detects the bipartite entanglement in the state $|\psi\rangle$.

3.4. Fidelity susceptibility

The geometrical structure of the Hilbert space introduced in section 2 applies to any parameterization of the state space. One case of particular physical interest is when that parameterization is generated by considering the ground state of the Hamiltonian,

$$\hat{H}(\vec{\Lambda}) = H_0 + \vec{\lambda} \cdot \vec{\Lambda} \tag{118}$$

where the parameters $\overline{\lambda}$ can be used to drive phase transitions in the ground state manifold. In this case, the QFIM is called the *fidelity susceptibility* [4, 5, 75] (up to a conventional factor of four, which we omit). It is also possible to consider the fidelity susceptibility with respect to parameters that are not driving the phase transition, such as a twist in the boundary conditions [72, 106, 107], in which case it would be more appropriate to associate it with the QFIM.

To construct the QGT explicitly, let $|\Psi_0(\lambda)\rangle$ be the ground state of equation (118) (we consider one component for simplicity). The infinitesimally neighboring state $|\Psi(\lambda + d\lambda)\rangle$ can be found by applying perturbation theory,

$$|\Psi_{0}(\lambda_{\mu} + d\lambda_{\mu})\rangle = |\Psi_{0}(\lambda_{\mu})\rangle + d\lambda_{\mu} \sum_{n \neq 0} \frac{\langle \Psi_{n} | \hat{\Lambda}_{\mu} | \Psi_{0} \rangle | \Psi_{n} \rangle}{E_{n} - E_{0}}$$
(119)

where, for simplicity, we consider one parameter at a time. Then, applying equation (55) and taking the real part, we recognize the QGT as,

$$\tilde{\mathcal{Q}}_{\mu\nu} = \sum_{n\neq 0} \frac{\langle \Psi_0 | \hat{\Lambda}_\mu | \Psi_n \rangle \langle \Psi_n | \hat{\Lambda}_\nu | \Psi_0 \rangle}{(E_n - E_0)^2}.$$
(120)

Provided that we avoid level crossings, the change in the parameters λ can be represented by the unitary operator [78],

$$\hat{U} = \sum_{n} |\Psi(\vec{\lambda} + d\vec{\lambda})\rangle \langle \Psi(\vec{\lambda})|.$$
(121)

We may identify the associated Hermitian operator via,

$$\tilde{\hat{\Lambda}}_{\mu} = i(\partial_{\mu}\hat{U})\hat{U}$$
(122)

and rewrite the QGT in the covariance form,

$$\tilde{\mathcal{Q}}_{\mu\nu} = \operatorname{Cov}_{\Psi_0} \left\{ \hat{\tilde{\Lambda}}_{\mu}, \hat{\tilde{\Lambda}}_{\nu} \right\}.$$
(123)

We denote the above QGT with a tilde in order to make a distinction between two different but related approaches to quantum metrology. Both approaches involve first defining a Hermitian operator $\hat{\Lambda}$. In the first approach, the associated unitary $exp\{-i\lambda\hat{\Lambda}\}$ is applied directly to our initial state. The resulting QGT is denoted $Q_{\mu\nu}$, where

$$\mathcal{Q}_{\mu\nu} = \langle \Psi_0 | \hat{\Lambda}_{\mu} \hat{\Lambda}_{\nu} | \Psi_0 \rangle - \langle \Psi_0 | \hat{\Lambda}_{\mu} | \Psi_0 \rangle \langle \Psi_0 | \hat{\Lambda}_{\nu} | \Psi_0 \rangle$$

= $\text{Cov}_{\Psi_0} \left\{ \hat{\Lambda}_{\mu}, \hat{\Lambda}_{\nu} \right\}.$ (124)

In the second approach, the initial state is taken to be the ground state of a Hamiltonian to which $\tilde{\Lambda}$ is applied as a perturbation. This is the QGT denoted $\tilde{Q}_{\mu\nu}$ in equation (123). The two are in fact related by a series of inequalities [78] which we reproduce here for completeness.

Let $\epsilon_{nm} = E_n - E_m$ and let ϵ_{10} be the smallest spectral gap. Now we may rewrite equation (120) as

$$\begin{split} \tilde{\mathcal{Q}}_{\mu\nu} &= \sum_{n\neq 0} \frac{\langle \Psi_0 | \hat{\Lambda}_\mu | \Psi_n \rangle \langle \Psi_n | \hat{\Lambda}_\nu | \Psi_0 \rangle}{\epsilon_{n0}^2} \\ &= \frac{1}{\epsilon_{n'0}^2} \sum_{n\neq 0} \frac{\epsilon_{10}^2}{\epsilon_{n0}^2} \langle \Psi_0 | \hat{\Lambda}_\mu | \Psi_n \rangle \langle \Psi_n | \hat{\Lambda}_\nu | \Psi_0 \rangle \\ &\leqslant \frac{1}{\epsilon_{10}^2} \sum_{n\neq 0} \langle \Psi_0 | \hat{\Lambda}_\mu | \Psi_n \rangle \langle \Psi_n | \hat{\Lambda}_\nu | \Psi_0 \rangle \\ &= \frac{1}{\epsilon_{10}^2} \left(\sum_n \langle \Psi_0 | \hat{\Lambda}_\mu | \Psi_n \rangle \langle \Psi_n | \hat{\Lambda}_\nu | \Psi_0 \rangle - \langle \Psi_0 | \hat{\Lambda}_\mu | \Psi_0 \rangle \langle \Psi_0 | \hat{\Lambda}_\nu | \Psi \rangle_0 \right) \\ &= \frac{1}{\epsilon_{10}^2} \mathcal{Q}_{\mu\nu}. \end{split}$$
(125)

Thus we find that, $\tilde{\mathcal{Q}}_{\mu\nu} \leq \frac{1}{\epsilon_{10}^2} \mathcal{Q}_{\mu\nu}$.

At a level crossing critical point, we expect the QGT to diverge. The nature of this divergence was determined in [74, 85, 108]. In order to analyze the scaling of the QGT it is useful to consider the QGT density,

$$\tilde{q}_{\mu\nu} = L^{-d} \tilde{\mathcal{Q}}_{\mu\nu} \tag{126}$$

where *d* is the spatial dimension of the system. Following [108], we consider a scale transform $x \to \alpha x$ and $\tau \to \alpha^{\zeta} \tau$. Here ζ is the dynamical critical exponent. Let Δ_{μ} be the scaling dimension of the operator $\hat{\Lambda}_{\mu}$. If λ_c is the critical point, the correlation length will diverge as $\xi \sim |\lambda - \lambda_c|^{-\nu}$. In the region approaching the critical point the QGT density will diverge as,

$$\tilde{q}_{\mu\nu} \sim |\lambda - \lambda_{c}|^{\nu(\Delta_{\mu} + \Delta_{\nu} - 2\zeta - d)}.$$
(127)

This also leads to a finite size scaling relation at the critical point itself where,

$$\tilde{q}_{\mu\nu} \sim L^{-(\Delta_{\mu} + \Delta_{\nu} - 2\zeta - d)}.$$
(128)

If the operator $\hat{\Lambda}_{\mu}$ drives the transition, as is the case when considering the fidelity susceptibility, then $\Delta_{\mu} = d + \zeta - 1/\nu$ and the diagonal terms in equation (128) become [74, 85]:

$$\tilde{q} \sim L^{2/\nu - d}.\tag{129}$$

These scaling relations can be contrasted with the scaling of the QGT for a unitary perturbation generated directly by the operator $\hat{\Lambda}$. In this case the dynamical critical exponent does not enter in, and we instead have,

$$q_{\mu\nu} \sim |\lambda - \lambda_c|^{\nu(\Delta_\mu + \Delta_\nu - d)} \tag{130}$$

and at the critical point itself the finite size scaling,

$$a_{\mu\nu} \sim L^{-(\Delta_{\mu} + \Delta_{\nu} - d)}.\tag{131}$$

As above, if we assume $\Delta_{\mu} = \Delta_{\nu} = 2 + \zeta - 1/\nu$, we obtain for the diagonal term:

$$q \sim L^{2/\nu - d - 2\zeta}.\tag{132}$$

For comparison, it is useful to note that the closely related energy susceptibility, $\chi^e = -\partial^2 e_0 / \partial \lambda^2$, with e_0 the ground-state energy per site, scales as

$$\chi^e \sim L^{2/\nu - d - \zeta}.\tag{133}$$

This result follows immediately from the scaling of the critical free energy density, $f_s \sim |\lambda - \lambda_c|^{\nu(d+\zeta)}$, with χ^e the second derivative of f_s . Notice that the dynamical critical exponent appears in equations (127) and (128) but not in equations (130) and (131). This is due to the fact that the two metrics consider different parameterizations. In particular, the metric $\tilde{Q}_{\mu\nu}$ determines the distance between the ground state of a Hamiltonian at some value of the Hamiltonian parameter λ_{μ} and some nearby value $\lambda_{\mu} + d\lambda_{\mu}$, where λ_{μ} couples to the operator $\hat{\Lambda}_{\mu}$. The metric $Q_{\mu\nu}$, however, determines the distance between a state $|\psi\rangle$ and the infinitesimally neighboring state $exp\{-id\lambda_{\mu}\hat{\Lambda}_{\mu}\}|\psi\rangle$. As we describe in equation (122), the perturbation to the Hamiltonian may be recast as a unitary transformation on the ground state provided that the change in parameters occurs adiabatically. The contribution of the dynamical critical exponent in the Fidelity susceptibility reflects the necessity of the transformation being adiabatic if we are to keep ourselves in the ground state. This difference in scaling turns out to have important implications for quantum sensing [109].

The above scaling discussion coupled with the relationship between the QFIM and genuine multipartite entanglement implies that as long as $\hat{\Lambda}$ is relevant in the sense of the renormalization group, the associated entanglement will diverge. In particular, if one had knowledge of both the dynamical critical exponent and the fidelity susceptibility, one could determine the scaling of a lower bound on the QFI associated with the operator $\hat{\Lambda}$ and hence the multipartite entanglement. Reciprocally, if one knew both the scaling of the QFI and the dynamical critical exponent, one could compute the scaling of an upper bound on the associated fidelity susceptibility.

3.5. Quantum sensing

Quantum sensing is the utilization of quantum resources, such as entanglement, to enhance the precision and sensitivity of measurements [15, 16]. The example of the weather stations describing in section 3 can be mapped directly onto a general quantum framework, whereby a probe is made to interact with a system for some amount of time, during which the probe acquires a parameter dependent phase [110]. The probe is then measured and the phase estimated. The maximum precision with which the parameter can be estimated is given by the Cramér-Rao bound,

$$\operatorname{Cov}_{P}\{\tilde{\lambda}_{\mu}, \tilde{\lambda}_{\nu}\} \geqslant \frac{1}{4F_{\mu\nu}}.$$
(134)

which we derived in section 2.1.2. For an extensive review of quantum sensing, see [18] as well as [17].

In [111], the authors introduce a bound on the metrological usefulness of a state $\hat{\rho}$ for a parameterization generated by the operator $\hat{\Lambda}$. In particular, if $g_{\lambda\lambda}^{\text{separable}}$ is a component of the metric generated by $\hat{\Lambda}$ optimized over all separable states, then the metrological gain, \mathcal{G} can be quantified as,

$$\mathcal{G}\{\hat{\Lambda}\} := \frac{g_{\lambda\lambda}(\hat{\rho})}{g_{\lambda\lambda}^{\text{separable}}}.$$
(135)

For any state with $\mathcal{G} \ge 1$ there is some metrologically useful entanglement.

The bounds in equations (107) and (112) imply that the sensitivity of a probe to quantum measurements is enhanced by the presence of entanglement. This fact can be used to exploit quantum critical regions for increased measurement sensitivity, as discussed in [112]. As mentioned in section 3.4, the difference in the scaling relations for parameterizations controlled adiabatically by a parent Hamiltonian, and those generated directly by a bounded unitary operator lead to different scaling in measurement sensitivity at the critical point [109]. Topological spins chains [96] and atomic ensembles have also been proposed as useful platforms for quantum sensing due to their geometrical properties [113].

Recent work has also examined the applications of the QFIM to the detection of periodically oscillating magnetic fields, using spin chains as intermediates [114]. Recently, the dynamical phase corresponding to a time crystal has also been shown to have a robust quantum advantage in measurement sensitivity.

3.6. Chern number and the Euler characteristic

Topology is playing an increasingly central role in the study of condensed matter and many body systems [115–117]. Quantum information geometry facilitates a link between the geometry of state space manifolds, response function, and the topological invariants that characterize those systems. A recent achievement was the classification of N level fermionic and bosonic systems according to a generalization of the Bloch-sphere [21].

Another studies have found that the Euler Characteristic in the first Brillouin zone is also indicative of topologically non-trivial phases [66, 118]. In this line of research, one considers the topology and geometry of the single particle states in a system that is well described by band theory. Recently such systems have been studied in the context of quantum sensing, see for example [64]. An especially important result in this context is given by [119] which bounds the volume of a two-dimensional parameter manifold (see equation (79)) with the Chern number (see equation (72)),

$$\operatorname{Vol}_{g} \geqslant \pi |C|. \tag{136}$$

The quantum geometric tensor associated with non-trivial band structures has also been measured experimentally using the relationship between the quantum geometric tensor and the response function discussed in section 3.1 (see [70]).

3.7. State space curvature

The curvature associated with different many-body phases has also been examined in several contexts [77, 89, 120]. While the physical implications of the state space curvature are not yet fully understood, it has been shown to clearly indicate quantum criticality, including in systems that lack a local order parameter [87]. In [77] it was shown that the state space curvature—with inverse temperature as one parameter, and the driving parameter of a quantum phase transition as the other—detects different scaling regimes in the vicinity of the critical point.

In [89] some progress was made in interpreting the state space curvature through its apparent relationship to multipartite entanglement. By considering spherical slices of state space generated by perturbing the Hamiltonian with the staggered magnetization operator,

$$\hat{\Lambda}_{\mathbf{n}(\theta,\phi)} = \sum_{r} (-1)^{r} \mathbf{n}(\theta,\phi) \cdot \mathbf{S}.$$
(137)

where $\hat{n} = (\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi))$, coupled to a source Hamiltonian via the field *h*,

$$H = H_0 + h\tilde{\Lambda}_{\mathbf{n}(\theta,\phi)} \tag{138}$$

The size of these manifolds can then be computed using equation (79),

$$Vol_{\psi} = \int_{\Omega} \sqrt{\det(\mathcal{F}_{\mu\nu}\{\psi(\theta,\phi)\})} d\theta d\phi$$
(139)

where $\mathcal{F}_{\mu\nu}$ is the metric for the sphere with tangent operators,

$$\hat{\Lambda}_{\theta} = \sum_{r} (-1)^{r} \partial_{\theta} \mathbf{n} \cdot \mathbf{S}_{r}$$
(140)

$$\hat{\Lambda}_{\phi} = \sum_{r} (-1)^{r} \partial_{\phi} \mathbf{n} \cdot \mathbf{S}_{r}$$
(141)

and the curvature is estimated by comparing the scaling of this volume as a function of the coupling field to the scaling of the volume (surface area) of a sphere in flat space. For the 2-manifold parameterized by the orientation of **n**, we expect the volume to scale as,

$$V_{\text{flat}} = 4\pi \,(ah)^2 \tag{142}$$

where *a* is some constant that depends on the spin moment. The ratio of the volume centered on the ground state, ψ of the source Hamiltonian in equation (138) to the volume flat space is then given by,

$$\frac{V_{\psi}}{V_{\text{flat}}} \approx 1 - bRh^2 \tag{143}$$

with *b* a constant and *R* the scalar curvature.

In the XY chain, the Euler characteristic was examined for a manifold parameterized by the spin anisotropy and by a global rotation. The topology of the resulting manifold was shown to change depending on whether or not the model was in a paramagnetic or ordered phase, with the former phase resulting in an irrational Euler characteristic due to the emergence of a cusp like singular [65].

Another application of quantum curvature is in the context of variational quantum Monte Carlo, where the QFIM is used to find the energetic minimum using stochastic reconfiguration [121]. In this case, the local curvature of the state space can be used to infer the robustness of the variational calculation, as discussed in [122].

4. Summary and outlook

In this paper, we have presented a pedagogical introduction to information geometry. Though the formal foundations of information geometry were laid over a century ago, it is in the last two decades that we have seen a rapid growth in research interest in the many fields to which this geometry is applicable. In particular, there are now an abundance of studies examining Fidelity susceptibility and Fisher information in various models from condensed matter systems to cold atom setups. We expect that this interest will only grow. Beyond equilibrium physics, concepts from quantum information geometry are finding application in the study of dynamical phases [86] and quantum chaos [123].

There is still much to be learned about the geometrical structure of quantum state space and in particular about interpretations of geometric invariants such as the curvature as discussed in [120]. If one recalls equation (43), we see that in classical information geometry, the metric is sourced by the Shannon information. In quantum information geometry, the analogous quantity is the Kähler potential (see, for example, [32]). Investigation of this quantity in many-body systems is only just beginning, and there are no doubt many exciting results to be found in this direction.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Properties of Shannon information

In this appendix we discuss some properties of information measures. We use $P = \{p_1, ..., p_N\}$ to denote a discrete distribution. The Shannon information defined in equation (19) has the following properties [32],

- Positivity, $S(P) \ge 0$
- *Continuity*, *S*(*P*) will vary continuous with the distribution *P*.
- *Expansibility*, $S(p_1, ..., p_N) = S(p_1, ..., p_N, 0)$.
- *Concavity*, $S(\alpha P_1 + (1 \alpha)P_2) \ge \alpha S(P_1) + (1 \alpha)S(P_2)$. The mixing of distributions always increases the information.
- *Subadditivity* For two random variables, not necessarily independent, where p_{12} is the joint probability distribution,

$$S(P_{12}) \leqslant S(P_1) + S(P_2) \tag{A1}$$

with equality only if the two distributions are independent.

• *Recursion*. We can coarse grain a distribution $P = p_1, \ldots, p_N$, via,

$$q_j = \sum_{i=k_{j-1}+1}^{k_j} p_i$$
 (A2)

where j = 1, ..., r and $k_0 = 0$. The discrete distribution has been partition according to the sum $N = \sum_{i=1}^{r} k_i$. Then the Shannon information is given by,

$$S(P) = S(Q) + q_1 S\left(\frac{p_1}{q_1}, \dots, \frac{p_{k_1}}{q_1}\right) + \dots + q_r S\left(\frac{p_{k_N - k_r + 1}}{q_r}, \dots, \frac{p_N}{q_r}\right)$$
(A3)

It is the property of recursion that makes the Shannon information unique amongst all possible choices of information measure. It appears in many other interesting contexts. In particular in the field of signal transmission where it was originally developed.

Appendix B. Example calculations

B.1. Single qubit

We may compute the trace explicitly by diagonalizing the Hamiltonian with source term. We denote the coefficients of the magnetization operator as m_{α} (no hat) and temporarily suppress the explicit dependence on θ and ϕ

$$\hat{H} \doteq \begin{pmatrix} \Delta - hm_z & h(m_x - im_y) \\ h(m_x + im_y) & -(\Delta - hm_z) \end{pmatrix}$$

We can immediately read off the eigenvalues for this matrix,

$$\lambda_{\pm} = \pm \sqrt{(\Delta - hm_z)^2 + h^2(m_x - im_y)(m_x + im_y)}$$
$$= \pm \sqrt{\Delta^2 - 2\Delta m_z h + h^2}$$

and thus find,

$$\mathcal{Z}(h) = 2\cosh(\beta\sqrt{\Delta^2 - 2\Delta m_z h + h^2}) \tag{B1}$$

The variance is given by,

$$Var(\hat{M}) = tr(\rho \hat{M}^2) - tr(\rho \hat{M})^2$$

= 1 - tanh²(\beta|\Delta|) cos²(\phi) (B2)

The thermal variance can be evaluated by first computing the first and second derivatives of the partition function with respect to the driving field,

$$\partial_{h=0} \mathcal{Z} = \partial_{h=0} 2 \cosh(\beta \sqrt{\Delta^2 - 2\Delta m_z h + h^2})$$

= $2\beta \left(\frac{\sinh(\beta \sqrt{\Delta^2 - 2\Delta m_z h + h^2})}{\sqrt{\Delta^2 - 2\Delta m_z h + h^2}} (h - \Delta m_z) \right)_{h=0}$
= $-2\beta \operatorname{sgn}(\Delta) m_z \sinh(\beta |\Delta|)$ (B3a)

$$\partial_{h=0}^{2} \mathcal{Z} = \left(2\beta^{2} \frac{\cosh(\beta\sqrt{\Delta^{2} - 2\Delta m_{z}h + h^{2}})}{\Delta^{2} - 2\Delta m_{z}h + h^{2}} (h - \Delta m_{z})^{2} - 2\beta \frac{\sinh(\beta\sqrt{\Delta^{2} - 2\Delta m_{z}h + h^{2}})}{(\Delta^{2} - 2\Delta m_{z}h + h^{2})^{\frac{3}{2}}} (h - \Delta m_{z})^{2} + 2\beta \frac{\sinh(\beta\sqrt{\Delta^{2} - 2\Delta m_{z}h + h^{2}})}{\sqrt{\Delta^{2} - 2\Delta m_{z}h + h^{2}}} \right)_{h=0} = 2\beta^{2} \cosh(\beta|\Delta|)m_{z}^{2} - 2\beta \frac{\sinh(\beta|\Delta|)}{|\Delta|}m_{z}^{2} + 2\beta \frac{\sinh(\beta|\Delta|)}{|\Delta|}$$
$$= 2\beta^{2} \cosh(\beta|\Delta|)m_{z}^{2} + 2\beta(1 - m_{z}^{2})\frac{\sinh(\beta|\Delta|)}{|\Delta|}.$$
(B3b)

No we substitute these relations into the definition of the thermal variance which we restate below,

$$\begin{aligned} \operatorname{Var}_{\mathcal{T}}(\hat{M};\hat{\rho}) &= \frac{1}{\beta^2} \left(\frac{1}{\mathcal{Z}(h=0)} \left(\frac{\partial^2 \mathcal{Z}}{\partial h^2} \right)_{h=0} - \frac{1}{\mathcal{Z}^2(h=0)} \left(\frac{\partial \mathcal{Z}}{\partial h} \right)_{h=0}^2 \right) \\ &= \frac{1}{\beta^2} \left(\beta^2 m_z^2 + \beta (1-m_z^2) \frac{\tanh(\beta|\Delta|)}{|\Delta|} - \beta^2 m_z^2 \tanh^2(\beta|\Delta|) \right) \\ &= m_z^2 + (1-m_z^2) \frac{\tanh(\beta|\Delta|)}{\beta|\Delta|} - m_z^2 \tanh^2(\beta|\Delta|) \\ &= m_z^2 (1-\tanh^2(\beta|\Delta|)) + (1-m_z^2) \frac{\tanh(\beta|\Delta|)}{\beta|\Delta|}. \end{aligned}$$
(B4)

The QV can be computed by substituting equations (B2) and (B4) into equation (91),

$$\operatorname{Var}_{\mathcal{Q}}(\hat{M};\hat{\rho}) = 1 - \operatorname{tanh}^{2}(\beta|\Delta|)m_{z}^{2} - m_{z}^{2}(1 - \operatorname{tanh}^{2}(\beta|\Delta|)) - (1 - m_{z}^{2})\frac{\operatorname{tanh}(\beta|\Delta|)}{\beta|\Delta|}$$
$$= (1 - m_{z}^{2})\left(1 - \frac{\operatorname{tanh}(\beta|\Delta|)}{\beta|\Delta|}\right). \tag{B5}$$

Using equation (71), we can calculate an explicit form of the QFI for the single qubit as well. Since diagonal entries do not contribute to the QFI, we consider only off diagonal terms, and denote by $\lambda_{\pm} = \pm \Delta$ the two eigenvalues and eigenstates of the Hamiltonian. We denote the corresponding eigenstates of σ^z by $|\pm\rangle$. The QFI given by,

(B6)

$$\begin{split} \mathcal{F}(\hat{M};\hat{\rho}) &= 2\sum_{\lambda,\lambda'} \frac{(p_{\lambda} - p_{\lambda'})^2}{p_{\lambda} + p_{\lambda'}} |\langle \lambda | \hat{M} | \lambda' \rangle|^2 \\ &= 4\frac{(p_+ - p_-)^2}{p_+ + p_-} |\langle + | \hat{M} | - \rangle|^2 \\ &= \frac{4}{\mathcal{Z}} \frac{(e^{-\beta |\Delta|} - e^{\beta |\Delta|})^2}{e^{-\beta |\Delta|} + e^{\beta |\Delta|}} (m_x^2 + m_y^2) \\ &= \frac{4}{\mathcal{Z}} \frac{4\sinh^2(\beta |\Delta|)}{2\cosh(\beta |\Delta|)} (1 - m_z^2) \\ &= 4\tanh^2(\beta |\Delta|)(1 - m_z^2) \\ &= 4\tanh^2(\beta |\Delta|)\sin^2(\phi). \end{split}$$

ORCID iDs

J Lambert (a) https://orcid.org/0000-0002-0888-5582 E S Sørensen (a) https://orcid.org/0000-0002-5956-1190

References

- Thouless D J, Kohmoto M, Nightingale M P and den Nijs M 1982 Quantized Hall conductance in a two-dimensional periodic potential *Phys. Rev. Lett.* 49 405–8
- [2] Pancharatnam S 1956 Generalized theory of interference and its applications: Part II. Partially coherent pencils Proc. Indian Acad. Sci. A 44 398–417
- [3] Berry M V 1984 Quantal phase factors accompanying adiabatic changes Proc. R. Soc. A 392 45–57
- [4] Zanardi P and Paunković N 2006 Ground state overlap and quantum phase transitions Phys. Rev. E 74 031123
- [5] Gu S-J 2010 Fidelity approach to quantum phase transitions Int. J. Mod. Phys. B 24 4371–458
- [6] Provost J P and Vallee G 1980 Riemannian structure on manifolds of quantum states Commun. Math. Phys. 76 289-301
- [7] Brody D C and Hughston L P 2001 Geometric quantum mechanics J. Geom. Phys. 38 19–53
- [8] Fisher R A 1925 Theory of statistical estimation Math. Proc. Camb. Phil. Soc. 22 700-25
- [9] Wootters W K 1981 Statistical distance and Hilbert space Phys. Rev. D 23 357
- [10] Braunstein S L and Caves C M 1994 Statistical distance and the geometry of quantum states Phys. Rev. Lett. 72 3439
- [11] Carollo A, Valenti D and Spagnolo B 2020 Geometry of quantum phase transitions Phys. Rep. 838 1-72
- [12] Böhm A, Mostafazadeh A, Koizumi H, Niu Q and Zwanziger J 2003 The Geometric Phase in Quantum Systems: Foundations, Mathematical Concepts and Applications in Molecular and Condensed Matter Physics (Springer)
- [13] Liu J, Yuan H, Lu X-M and Wang X 2019 Quantum Fisher information matrix and multiparameter estimation *J. Phys. A: Math. Theor.* 53 023001
- [14] Brody D C and Hook D W 2008 Information geometry in vapour-liquid equilibrium J. Phys. A: Math. Theor. 42 023001
- [15] Giovannetti V, Lloyd S and Maccone L 2006 Quantum metrology Phys. Rev. Lett. 96 010401
- [16] Giovannetti V, Lloyd S and Maccone L 2011 Advances in quantum metrology Nat. Photon. 5 222-9
- [17] Tóth G and Apellaniz I 2014 Quantum metrology from a quantum information science perspective J. Phys. A: Math. Theor. 47 424006
- [18] Degen C L, Reinhard F and Cappellaro P 2017 Quantum sensing Rev. Mod. Phys. 89 035002
- [19] Tóth G 2012 Multipartite entanglement and high-precision metrology Phys. Rev. A 85 022322
- [20] Hyllus P, Laskowski W, Krischek R, Schwemmer C, Wieczorek W, Weinfurter H, Pezzé L and Smerzi A 2012 Fisher information and multiparticle entanglement Phys. Rev. A 85 022321
- [21] Kemp C J D, Cooper N R and Nur Ünal F 2021 Nested spheres description of the N-level Chern number and the generalized Bloch hypersphere (arXiv:2110.06934)
- [22] Klees R L, Rastelli G, Cuevas J C and Belzig W 2020 Microwave spectroscopy reveals the quantum geometric tensor of topological Josephson matter Phys. Rev. Lett. 124 197002
- [23] Klees R L, Cuevas J C, Belzig W and Rastelli G 2021 Ground-state quantum geometry in superconductor-quantum dot chains Phys. Rev. B 103 014516
- [24] Hauke P, Heyl M, Tagliacozzo L and Zoller P 2016 Measuring multipartite entanglement through dynamic susceptibilities Nat. Phys. 12 778–82
- [25] Shitara T and Ueda M 2016 Determining the continuous family of quantum Fisher information from linear-response theory Phys. Rev. A 94 062316
- [26] Birkhoff G D 1931 Proof of the ergodic theorem Proc. Natl Acad. Sci. 17 656-60
- [27] In fact the particular formalism is not so important. We can define quantum systems in classical state space through the Wigner quasi probability distribution function, [28, 29], but, the Kolmogorov axioms must be relaxed to allow for negative probabilities. We can also define quantum states in real spaces as opposed to complex spaces, by imposing additional constraints on the allowed observables [32].
- [28] Groenewold H J 1946 On the principles of elementary quantum mechanics On the Principles of Elementary Quantum Mechanics (Springer) pp 1–56
- [29] Moyal J E 1949 Quantum mechanics as a statistical theory Math. Proc. Camb. Phil. Soc. 45 99–124
- [30] Amari S-ichi and Nagaoka H 2000 Methods of Information Geometry (American Mathematical Society and Oxford University Press)
- [31] Amari S-ichi 2016 Information Geometry and Its Applications (Springer)
- [32] Bengtsson I and Życzkowski K 2017 Geometry of Quantum States: An Introduction to Quantum Entanglement (Cambridge University Press)

- [33] Fubini G 1904 Sulle Metriche Definite da una Forma Hermitiana: Nota (Office graf. C. Ferrari)
- [34] Study E 1905 Kürzeste wege im komplexen gebiet Math. Ann. 60 321-78
- [35] Rao C R 1945 Information and the accuracy attainable in the estimation of statistical parameters Reson. J. Sci. Educ. 20 78–90
- [36] Kraft L G 1949 A device for quantizing, grouping, and coding amplitude-modulated pulses PhD Thesis Massachusetts Institute of Technology
- [37] Shannon C E 1948 A mathematical theory of communication Bell Syst. Tech. J. 27 379-423
- [38] Rényi A 1961 On measures of entropy and information *Proc. 4th Berkeley Symp. on Mathematical Statistics and Probability* (*Berkeley, California, USA*) vol 1 pp 547–61
- [39] Fisher R A 1922 On the mathematical foundations of theoretical statistics Phil. Trans. R. Soc. A 222 309-68
- [40] Kullback S and Leibler R A 1951 On information and sufficiency Ann. Math. Stat. 22 79-86
- [41] Sanov I N 1958 On the Probability of Large Deviations of Random Variables (United States Air Force, Office of Scientific Research)
- [42] Bhattacharyya A 1946 On a measure of divergence between two multinomial populations Sankhya: Indian J. Stat. 7 401-6
- [43] A simplex is a convex set wherein which each mixture of pure points is unique.
- [44] Holzhey C, Larsen F and Wilczek F 1994 Geometric and renormalized entropy in conformal field theory Nucl. Phys. B 424 443–67
- [45] Calabrese P and Cardy J 2004 Entanglement entropy and quantum field theory J. Stat. Mech. P06002
- [46] Calabrese P and Cardy J 2009 Entanglement entropy and conformal field theory J. Phys. A: Math. Theor. 42 504005
- [47] Eisert J, Cramer M and Plenio M B 2010 Colloquium: Area laws for the entanglement entropy Rev. Mod. Phys. 82 277
- [48] White S R 1992 Density matrix formulation for quantum renormalization groups Phys. Rev. Lett. 69 2863-6
- [49] Li H and Haldane. F D M 2008 Entanglement spectrum as a generalization of entanglement entropy: identification of topological order in non-Abelian fractional quantum Hall effect states *Phys. Rev. Lett.* 101 010504
- [50] Pollmann F, Turner A M, Berg E and Oshikawa M 2010 Entanglement spectrum of a topological phase in one dimension Phys. Rev. B 81 064439
- [51] Kitaev A and Preskill J 2006 Topological entanglement entropy Phys. Rev. Lett. 96 110404
- [52] Holevo A S 2003 *Statistical Structure of Quantum Theory* vol 67 (Springer Science & Business Media)
- [53] Holevo A S 2011 Probabilistic and Statistical Aspects of Quantum Theory vol 1 (Springer Science & Business Media)
- [54] Braunstein S L, Caves C M and Milburn G J 1996 Generalized uncertainty relations: theory, examples and Lorentz invariance Ann. Phys., NY 247 135–73
- [55] Jozsa R 1994 Fidelity for mixed quantum states J. Mod. Opt. 41 2315-23
- [56] Morozova E A and Chentsov N N 1989 Markov invariant geometry on state manifolds Itogi Nauki Tekh. 36 69-102
- [57] Petz D 1996 Monotone metrics on matrix spaces *Linear Algebr. Appl.* 244 81–96
- [58] Petz D and Sudár C 1996 Geometries of quantum states J. Math. Phys. 37 2662-73
- [59] Petz D and Hasegawa H 1996 On the Riemannian metric of α -entropies of density matrices Lett. Math. Phys. 38 221–5
- [60] Bures D 1969 An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w*-algebras Trans. Am. Math. Soc. 135 199–212
- [61] Uhlmann A 1976 The "transition probability" in the state space of a *-algebra Rep. Math. Phys. 9 273-9
- [62] Do Carmo M P 2016 Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition (Courier Dover Publications)
- [63] Kolodrubetz M, Gritsev V and Polkovnikov A 2013 Classifying and measuring geometry of a quantum ground state manifold Phys. Rev. B 88 064304
- [64] Sarkar S, Mukhopadhyay C, Alase A and Bayat A 2022 Free-fermionic topological quantum sensors (arXiv:2201.07102)
- [65] Kolodrubetz M, Sels D, Mehta P and Polkovnikov A 2017 Geometry and non-adiabatic response in quantum and classical systems Phys. Rep. 697 1–87
- [66] Ma Y-Q, Gu S-J, Chen S, Fan H and Liu W-M 2013 The Euler number of Bloch states manifold and the quantum phases in gapped fermionic systems *Europhys. Lett.* 103 10008
- [67] Ozawa T and Goldman N 2018 Extracting the quantum metric tensor through periodic driving Phys. Rev. B 97 201117
- [68] Mera B and Ozawa T 2021 K\u00e4hler geometry and Chern insulators: relations between topology and the quantum metric Phys. Rev. B 104 045104
- [69] McClarty P 2021 Topological magnons: a review (arXiv:2106.01430)
- [70] Gianfrate A et al 2020 Measurement of the quantum geometric tensor and of the anomalous Hall drift Nature 578 381-5
- [71] Tan X et al 2019 Experimental measurement of the quantum metric tensor and related topological phase transition with a superconducting qubit Phys. Rev. Lett. 122 210401
- [72] Thesberg M and Sørensen E S 2011 General quantum fidelity susceptibilities for the *j*₁-*j*₂ chain *Phys. Rev.* B 84 224435
- [73] Tzeng Y-C and Yang M-F 2008 Scaling properties of fidelity in the spin-1 anisotropic model Phys. Rev. A 77 012311
- [74] Fabricio Albuquerque A, Alet F, Sire C and Capponi S 2010 Quantum critical scaling of fidelity susceptibility Phys. Rev. B 81 064418
- [75] Gu S-J, Kwok H-M, Ning W-Q and Lin H-Q 2008 Fidelity susceptibility, scaling and universality in quantum critical phenomena Phys. Rev. B 77 245109
- [76] Gu S-J and Lin H-Q 2009 Scaling dimension of fidelity susceptibility in quantum phase transitions Europhys. Lett. 87 10003
- [77] Zanardi P, Venuti L C and Giorda P 2007 Bures metric over thermal state manifolds and quantum criticality Phys. Rev. A 76 062318
- [78] Zanardi P, Giorda P and Cozzini M 2007 Information-theoretic differential geometry of quantum phase transitions Phys. Rev. Lett. 99 100603
- [79] You W-L, Li Y-W and Gu S-J 2007 Fidelity, dynamic structure factor and susceptibility in critical phenomena Phys. Rev. E 76 022101
- [80] Yang S, Gu S-J, Sun C-P and Lin H-Q 2008 Fidelity susceptibility and long-range correlation in the Kitaev honeycomb model Phys. Rev. A 78 012304
- [81] Garnerone S, Abasto D, Haas S and Zanardi P 2009 Fidelity in topological quantum phases of matter Phys. Rev. A 79 032302
- [82] Wang L, Liu Y-H, Imriška J, Ma P N and Troyer M 2015 Fidelity susceptibility made simple: a unified quantum Monte Carlo approach Phys. Rev. X 5 031007
- [83] Cozzini M, Giorda P and Zanardi P 2007 Quantum phase transitions and quantum fidelity in free fermion graphs Phys. Rev. B 75 014439
- [84] Langari A, Pollmann F and Siahatgar M 2013 Ground-state fidelity of the spin-1 Heisenberg chain with single ion anisotropy: quantum renormalization group and exact diagonalization approaches J. Phys.: Condens. Matter 25 406002

- [85] Schwandt D, Alet F and Capponi S 2009 Quantum Monte Carlo simulations of fidelity at magnetic quantum phase transitions Phys. Rev. Lett. 103 170501
- [86] Mera B, Vlachou C, Paunković N, Vieira V R and Viyuela O 2018 Dynamical phase transitions at finite temperature from fidelity and interferometric Loschmidt echo induced metrics *Phys. Rev. B* 97 094110
- [87] Lambert J and Sørensen E S 2020 Revealing divergent length scales using quantum Fisher information in the Kitaev honeycomb model Phys. Rev. B 102 224401
- [88] Lambert J and Sørensen E S 2019 Estimates of the quantum Fisher information in the s = 1 antiferromagnetic Heisenberg spin chain with uniaxial anisotropy *Phys. Rev.* B 99 045117
- [89] Lambert J and Sørensen E S 2022 State space geometry of the spin-1 antiferromagnetic Heisenberg chain (arXiv:2209.05005)
- [90] Yin S, Song J, Zhang Y and Liu S 2019 Quantum Fisher information in quantum critical systems with topological characterization Phys. Rev. B 100 184417
- [91] Laurell P, Scheie A, Mukherjee C J, Koza M M, Enderle M, Tylczynski Z, Okamoto S, Coldea R, Alan Tennant D and Alvarez G 2021 Quantifying and controlling entanglement in the quantum magnet Cs₂ CoCl₄ Phys. Rev. Lett. 127 037201
- [92] Scheie A, Laurell P, Samarakoon A M, Lake B, Nagler S E, Granroth G E, Okamoto S, Alvarez G and Tennant D A 2021 Witnessing entanglement in quantum magnets using neutron scattering Phys. Rev. B 103 224434
- [93] Li S-S, Yi H-G and Chen R-H 2013 Spin squeezing and quantum Fisher information in the Lipkin-Meshkov-Glick model Int. J. Theor. Phys. 52 1175–81
- [94] Ma J and Wang X 2009 Fisher information and spin squeezing in the Lipkin-Meshkov-Glick model Phys. Rev. A 80 012318
- [95] Liu W-F, Ma J and Wang X 2013 Quantum Fisher information and spin squeezing in the ground state of the XY model J. Phys. A: Math. Theor. 46 045302
- [96] Pezze L, Gabbrielli M, Lepori L and Smerzi A 2017 Multipartite entanglement in topological quantum phases Phys. Rev. Lett. 119 250401
- [97] Bleu O, Solnyshkov D D and Malpuech G 2018 Measuring the quantum geometric tensor in two-dimensional photonic and exciton-polariton systems Phys. Rev. B 97 195422
- [98] Yu M et al 2020 Experimental measurement of the quantum geometric tensor using coupled qubits in diamond Natl Sci. Rev. 7 254–60
- [99] Menon V, Sherman N E, Dupont M, Scheie A O, Alan Tennant D and Moore J E 2023 Multipartite entanglement in the one-dimensional spin-1 2 Heisenberg antiferromagnet Phys. Rev. B 107 054422
- [100] Frérot I and Roscilde T 2016 Quantum variance: a measure of quantum coherence and quantum correlations for many-body systems Phys. Rev. B 94 075121
- [101] Malpetti D and Roscilde T 2016 Quantum correlations, separability and quantum coherence length in equilibrium many-body systems Phys. Rev. Lett. 117 130401
- [102] Wigner E P and Yanase M M 1997 Information contents of distributions Part I: Particles and Fields. Part II: Foundations of Quantum Mechanics (Springer) pp 452–60
- [103] Wigner E P and Yanase M M 1964 On the positive semidefinite nature of a certain matrix expression Can. J. Math. 16 397–406
- [104] Frérot I and Roscilde T 2019 Reconstructing the quantum critical fan of strongly correlated systems using quantum correlations Nat. Commun. 10 1–7
- [105] Pezzé L and Smerzi A 2009 Entanglement, nonlinear dynamics and the Heisenberg limit Phys. Rev. Lett. 102 100401
- [106] Laflorencie N, Capponi S and Sørensen. E S 2001 Finite size scaling of the spin stiffness of the antiferromagnetic s = 1/2 XXZ chain *Eur. Phys. J.* B 24 77–84
- [107] Thesberg M and Sørensen E S 2014 A quantum fidelity study of the anisotropic next-nearest-neighbour triangular lattice Heisenberg model J. Phys.: Condens. Matter 26 425602
- [108] Venuti L C and Zanardi P 2007 Quantum critical scaling of the geometric tensors Phys. Rev. Lett. 99 095701
- [109] Rams M M, Sierant P, Dutta O, Horodecki P and Zakrzewski J 2018 At the limits of criticality-based quantum metrology: apparent super-Heisenberg scaling revisited Phys. Rev. X 8 021022
- [110] Pezze L and Smerzi A 2014 Quantum theory of phase estimation (arXiv:1411.5164)
- [111] Tóth G, Vértesi T, Horodecki P and Horodecki R 2020 Activating hidden metrological usefulness Phys. Rev. Lett. 125 020402
- [112] Frérot I and Roscilde T 2018 Quantum critical metrology Phys. Rev. Lett. 121 020402
- [113] Pezze L, Smerzi A, Oberthaler M K, Schmied R and Treutlein P 2018 Quantum metrology with nonclassical states of atomic ensembles *Rev. Mod. Phys.* 90 035005
- [114] Mishra U and Bayat A 2021 Driving enhanced quantum sensing in partially accessible many-body systems Phys. Rev. Lett. 127 080504
- [115] Zahid Hasan M and Kane C L 2010 Colloquium: Topological insulators Rev. Mod. Phys. 82 3045
- [116] Sato M and Ando Y 2017 Topological superconductors: a review Rep. Prog. Phys. 80 076501
- [117] Rachel S 2018 Interacting topological insulators: a review Rep. Prog. Phys. 81 116501
- [118] Zhang A 2022 Revealing Chern number from quantum metric Chin. Phys. B 31 040201
- [119] Ozawa T and Mera B 2021 Relations between topology and the quantum metric for Chern insulators (arXiv:2103.11582)
- [120] Erdmenger J, Grosvenor K and Jefferson R 2020 Information geometry in quantum field theory: lessons from simple examples SciPost Phys. 8 073
- [121] Sorella S 2001 Generalized Lanczos algorithm for variational quantum Monte Carlo Phys. Rev. B 64 024512
- [122] Park C-Y and Kastoryano M J 2020 Geometry of learning neural quantum states Phys. Rev. Res. 2 023232
- [123] Wang X-Q, Ma J, Zhang X-H and Wang X-G 2011 Chaos and quantum Fisher information in the quantum kicked top Chin. Phys. B 20 050510