Physicality of evolution and statistical contractivity are equivalent notions of maps

Matteo Scandi,[1, *] Paolo Abiuso,[2, †] Dario De Santis,[3, ‡] and Jacopo Surace[4, §]

[1]*Institute for Cross-Disciplinary Physics and Complex Systems (IFISC) UIB-CSIC,*
Campus Universitat Illes Balears,E-07122 Palma de Mallorca, Spain
[2]*Institute for Quantum Optics and Quantum Information - IQOQI Vienna,*
Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria
[3]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[4]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada*
(Dated: March 29, 2024)

Statistical quantifiers are generically required to contract under physical maps, following the intuition that information should be lost under noisy transformations. This principle is so important in statistics that it even allows to derive uniqueness results based on it: the Chentsov-Petz theorem identifies the Fisher information metrics as the only family, on the space of probability distributions (or density matrices), that contracts under physical maps. This construction could suggest that statistical quantifiers are a derived concept, while the only fundamental object are the physical evolutions. The aim of this work is to disproof this belief. To this end, we prove a statement dual to the Chentsov-Petz theorem, showing that among all possible linear maps, the only ones that contract the Fisher information are exactly the physical ones. This result proves that, contrary to the common opinion, there is no fundamental hierarchy between physical maps and canonical statistical quantifiers, as either of them can be defined in terms of the other.

INTRODUCTION

The result that gave birth to information theory, Shannon's noiseless coding theorem, answers the question of what is the maximum rate at which a message can be transmitted over an ideal transmission line [1, 2]. This can be phrased in terms of the relative entropy, defined for two discrete probability vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ as:

$$D(\boldsymbol{p}||\boldsymbol{q}) := \sum_i p_i \log \frac{p_i}{q_i} , \qquad (1)$$

where p_i/q_i are the components of the two vectors. Then, the minimum compression is connected to the statistical difference, as measured by $D(\boldsymbol{p}||\boldsymbol{q})$, between the code probability distribution and the uniform one. Thanks to the operational interpretation of the relative entropy as the asymptotic difference between the two log-likelihood [3], this result simply means that in order for a probabilistic code to work, it needs to generate correlations significantly different from random bits.

Over time, information theory evolved from its role in telecommunication to become one of the cornerstones of the modern understanding of physics. Indeed, apart from its decisive role in the treatment of statistical mechanics [4–6], its extension to the quantum regime gave rise to an entire new field of physics [7]. Today, the concept of information plays a foundational role very similar to the one of energy for classical mechanics. For example, time-reversal symmetry follows from the requirement that the global evolution should conserve information, constraining the dynamics, thanks to Wigner's theorem [2], to unitaries or antiunitaries transformations. Moreover, much in the same way in which one reconciles the existence of dissipative dynamics with the principle of conservation of energy, we explain the existence of evolutions for which information appears to be lost by postulating that the missing information is simply transferred to some environmental degrees of freedom which we do not have access to. Following this argument, the most generic quantum mechanical evolution Φ takes the form:

$$\Phi(\rho_S) = \mathrm{Tr}_E \left[U_{SE} \left(\rho_S \otimes \omega_E \right) U_{SE}^\dagger \right] , \qquad (2)$$

where ρ_S is the state of the system, ω_E is a generic environmental state and U_{SE} is the global unitary/antiunitary evolution. Moreover, requiring the compatibility with composition of different systems, further restricts the set of global transformations U_{SE} to unitaries only [8, 9]. The class of maps defined by Eq. (2), with U_{SE} restricted to be unitary, is characterised by Stinespring theorem [10] as the set of Completely Positive Trace-Preserving maps (CPTP). When the same reasoning is applied to classical probabilities, one obtains the set of stochastic maps, since these can always be decomposed as:

$$[T(\boldsymbol{p})]_i = \sum_{j,k,l} U_{ij}^{kl} p_k \omega_l , \qquad (3)$$

where $\boldsymbol{p}$ represents the classical probability distribution of the system, $\boldsymbol{\omega}$ the one of the environment, while U is a global isometry.

Despite their different origins, statistical quantifiers and dynamical maps are deeply connected. For one thing, the relative entropy in Eq. (1) (as well as its quantum counterpart $H_L(\rho||\sigma) := \mathrm{Tr}\left[\rho(\log \rho - \log \sigma)\right]$) contracts under physical evolutions:

$$D(\boldsymbol{p}||\boldsymbol{q}) \geq D(T(\boldsymbol{p})||T(\boldsymbol{q})) , \qquad (4)$$

for any stochastic transformation T. Similarly, any sensible generalisation of the relative entropy needs to satisfy a condition akin to Eq. (4), corresponding to the intuition that information cannot increase during noisy transformations. This connection becomes even more compelling when one looks at the local behaviour of statistical quantifiers: a renown theorem by Chentsov establishes that there exists a unique metric on the space of classical probability distributions that is contractive under any stochastic (i.e., physical) map [11, 12]. This takes the name of Fisher information metric, or Fisher-Rao metric, and it is the main object of interest of this work. In this way, dynamical considerations constrain the local behaviour of any possible statistical quantifier. The deep connection between dynamics and Fisher information [13, 14] has been mostly overlooked in the literature, as more operative uses were explored [15–17].

Still, there is another path to the Fisher metric that we want to highlight here. In an effort to gather the many possible statistical quantifiers present in the literature, Csiszár introduced the concept of generalised information divergences [18], a continuous family generalising many of the main properties of the relative entropy [19]. Interestingly though, despite the wide range of possible definitions, it was noticed that for close-by probabilities (in the sense that $\boldsymbol{q} = \boldsymbol{p} + \mathrm{d}\boldsymbol{q}$ for some $||\mathrm{d}\boldsymbol{q}|| \ll 1$) all the Csiszár-divergences reduce to the Fisher information. Indeed, if we take for example the relative entropy in Eq. (1) it is a matter of Taylor-expanding the logarithm to see that:

$$D(\boldsymbol{p}||\boldsymbol{p} + \mathrm{d}\boldsymbol{q}) \simeq \sum_i \frac{\mathrm{d}q_i^2}{2p_i} = \frac{1}{2} \sum_{i,j} \mathrm{d}q_i \, \eta_{i,j}\big|_{\boldsymbol{p}} \, \mathrm{d}q_j \,, \quad (5)$$

where $\eta_{i,j}\big|_{\boldsymbol{p}} := \frac{\delta_{i,j}}{p_i}$ is exactly the Fisher metric, and we do not consider higher order corrections in $\mathrm{d}\boldsymbol{q}$. This shows that the Fisher information metric could also be defined in purely statistical terms, that is, independently of any dynamical considerations.

Indeed, even if one might think that there is a hierarchy between physical maps and Fisher information (and, therefore, statistical quantifiers), in which the latter derives from the first, we show in this work that the situation is actually more intertwined: we prove that physical evolutions can be defined solely in terms of the Fisher information metric. Before doing this, though, we need to introduce the concept of contrast functions and Fisher information metrics for quantum states.

QUANTUM FISHER INFORMATION METRICS

In this section we give a brief overview of the quantum Fisher information metrics, presenting in particular their statistical derivation (Thm. 1) and the dynamical one (Thm. 2). For the sake of brevity, we will only outline the main definitions, and we refer the reader to the review in [14] for further details.

Let us then introduce the concept of contrast functions $H(\rho||\sigma)$ between two quantum states ρ and σ. This notion was first introduced for classical statistics by Csiszár [19], and later extended by Petz to the quantum regime [20], following an effort to give an axiomatic foundation to the many different divergences present in the literature. In particular, the principles proposed are the following: *1) positivity and faithfulness*: $H(\rho||\sigma) \geq 0$, with equality *iff* $\rho \equiv \sigma$; *2) joint convexity*: mixing states should not increase their information content, therefore $H(\lambda\rho_1 + (1 - \lambda)\rho_2||\lambda\sigma_1 + (1 - \lambda)\sigma_2) \leq \lambda H(\rho_1||\sigma_1) + (1 - \lambda)H(\rho_2||\sigma_2)$, for $0 \leq \lambda \leq 1$; *3) monotonicity*: an analogous condition for compatibility with physical evolutions $H(\Phi(\rho)||\Phi(\sigma)) \leq H(\rho||\sigma)$; *4) differentiability*: C^∞-regularity in the arguments of H. These conditions are still not enough to single out a unique family of statistical quantifiers, and for this reason it is customary to also propose an ansatz for $H(\rho||\sigma)$, which encompasses most of the examples defined in the literature. In particular, the one proposed in [20] reads as follows:

$$H_g(\rho||\sigma) := \mathrm{Tr}\left[g\big(\mathbb{L}_\sigma \mathbb{R}_\rho^{-1}\big)[\rho]\right] , \quad (6)$$

where $g : \mathbb{R}^+ \to \mathbb{R}$ is a matrix convex function [21], and $\mathbb{L}_\rho$ and $\mathbb{R}_\rho$ are the left and right multiplication operators, acting as:

$$\mathbb{L}_\rho[A] = \rho\,A\,, \qquad \mathbb{R}_\rho[A] = A\,\rho\,. \quad (7)$$

It should be noticed that despite being a very large family, containing the relative entropy (when $g(x) = -\log x$), and functionals related to the Rényi divergences, the definition in Eq. (6) is still not enough to completely exhaust all the possibilities present in the literature [14]. Still, it is enough to define the entire family of Fisher information metrics. Indeed, we have the following:

Theorem 1 (Lesniewski, Ruskai [22]). *The contrast functions $H_g(\pi + \varepsilon A||\pi + \varepsilon B)$ can be locally approximated up to corrections of order $\mathcal{O}(\varepsilon^3)$ as:*

$$H_g(\pi + \varepsilon A||\pi + \varepsilon B) \simeq \frac{\varepsilon^2}{2} \mathrm{Tr}\left[(A - B)\,\mathbb{J}_f^{-1}\big|_\pi [(A - B)]\right] , \quad (8)$$

where A and B are traceless, Hermitian perturbations, and the superoperator $\mathbb{J}_f\big|_\pi$, called the quantum Fisher operator, is defined as:

$$\mathbb{J}_f\big|_\pi := \mathbb{R}_\pi\, f\big(\mathbb{L}_\pi \mathbb{R}_\pi^{-1}\big)\,. \quad (9)$$

Moreover, f is an operator monotone function connected to g by the equation:

$$f(x) = \frac{(x - 1)^2}{g(x) + x\,g(x^{-1})}\,. \quad (10)$$

This represents the first possible derivation of the Fisher information metrics for quantum systems. Still, in order to claim that the operators in Eq. (9) define all and only the Fisher information metrics we need to resort to their dynamical characterisation, i.e., their contractivity under arbitrary CPTP maps. This is provided by the independent result of Petz:

Theorem 2 (Petz [23]). *All the metrics on quantum states that contract under arbitrary CPTP maps are of the form:*

$$K_{f,\pi}(A, B) := \mathrm{Tr}\left[A\, \mathbb{J}_f^{-1}\big|_\pi[B]\right], \qquad (11)$$

where $f : \mathbb{R}^+ \to \mathbb{R}^+$ is an operator monotone function from the family in Eq. (10).

This result makes the definition of Fisher information metrics in the quantum regime unambiguous. Indeed, even if the single metric $\eta_{i,j}\big|_p$ is now replaced by a continuous family of possible operators $\mathbb{J}_f^{-1}\big|_\pi$, each one describing the local behaviour of a different contrast function (Thm. 1), it would not be correct to postulate some extra axioms to isolate a unique quantity. Indeed, Thm. 2 tells us that narrowing down the family in Eq. (9) would also destroy the connection with the dynamical characterisation of the Fisher information metrics. Several metrics in this family have important applications in information theory, such as the Bures metric (corresponding to $\mathbb{J}_{\frac{1+x}{2}}\big|_\pi[A] = \frac{1}{2}\{\pi, A\}$) in quantum metrology [24], the Kubo-Mori-Bogoliubov metric ($\mathbb{J}_{\frac{x-1}{\ln x}}\big|_\pi[A] = \int_0^1 \mathrm{d}s\, \pi^s A \pi^{(1-s)}$) as the local expansion of the quantum relative entropy, or the Wigner-Yanase metric ($\mathbb{J}_{\frac{(1+\sqrt{x})^2}{4}}\big|_\pi[A] = \frac{1}{4}\{\sqrt{\pi}, \{\sqrt{\pi}, A\}\}$) in hypothesis testing [25]. For a more exhaustive classification, see [14]. Despite the exotic form of these superoperators, it should be noticed that for commuting observables the uniqueness of the classical Fisher information is reattained. That is, if $[A, \pi] = 0$, then: $\mathbb{J}_f^{-1}\big|_\pi[A] = \pi^{-1} A$, irrespectively of the defining function f.

CONTRACTIVITY IMPLIES PHYSICALITY

In this section we prove our main result, that is how the property of statistical contractivity isolates physical evolutions among all possible linear maps. Our main mathematical theorem (Thm. 3) characterizes positive maps via their contraction of the Fisher information metric. This can be then extended to complete positivity (Cor. 3.1) and contrast functions contractivity (Cor. 3.2).

Define $\mathcal{M}_d(\mathbb{C})$ to be the space of $d \times d$ complex matrices, $\mathcal{P}_d \subset \mathcal{M}_d(\mathbb{C})$ to be the subset of positive semidefinite matrices, and $\mathcal{S}_d \subset \mathcal{P}_d$ the space of states, i.e., the matrices in $\mathcal{P}_d$ with trace one. Moreover, denote the interior of a set X as X°. Then, it holds that:
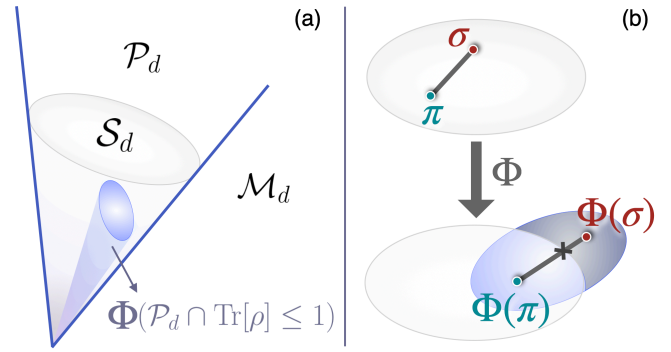

FIG. 1. We consider a generic linear map Φ, and its action on the cone of positive operators $\mathcal{P}_d$ (including normalized states $\mathcal{S}_d$ with trace 1). (a) Depiction of our main result (Thm. 3): the contractivity of the Fisher information metric (12) implies that Φ is positive ($\Phi(\mathcal{P}_d) \subseteq \mathcal{P}_d$) and does not increase the trace. (b) Sketch of the positivity proof, valid also when restricting to $\mathcal{S}_d$ only (Cor. 3.1): thanks to condition 1 one can always find a point π in $\mathcal{P}_d^\circ$ that gets mapped into $\mathcal{P}_d^\circ$. Now assume there is a matrix σ in the $\mathcal{P}_d^\circ$ that gets mapped outside of $\mathcal{P}_d$. Then, the segment connecting π to σ will be mapped into a curve intersecting the boundary of $\mathcal{P}_d$. Since the Fisher information is bounded on the interior of the space of states, and diverges on its boundary, there must be a point for which the Fisher metric increases.

Theorem 3. *Consider a Hermitian preserving, linear map $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$, and any choice f of Fisher information metrics (11). If Φ satisfies the following two properties:*

1. *Φ maps at least one point from $\mathcal{P}_d^\circ$ into $\mathcal{P}_d^\circ$;*

2. *for any matrix $\rho \in \mathcal{P}_d^\circ$ such that $\Phi(\rho) \in \mathcal{P}_d^\circ$, and any vector $\delta\rho$ in the $\mathcal{P}_d$-tangent space, the Fisher information decreases:*

$$K_{f,\rho}(\delta\rho, \delta\rho) \geq K_{f,\Phi(\rho)}(\Phi(\delta\rho), \Phi(\delta\rho)); \qquad (12)$$

then the image of Φ is completely contained in $\mathcal{P}_d$, meaning that Φ is a positive map (P). Moreover, Φ is trace non-increasing.

We refer the reader to Appendix 1 for the precise proof of the theorem. The main idea is explained in Fig. 1, and it leverages the fact that the Fisher information metric is bounded in the interior of $\mathcal{P}_d$, while it diverges on its boundary. Then, one can prove the claim by contradiction: if there is even a single point in $\mathcal{P}_d^\circ$ that gets mapped outside of $\mathcal{P}_d$, thanks to the linearity of Φ, one can define a segment in $\mathcal{P}_d^\circ$ whose image will intersect the boundary of $\mathcal{P}_d$ (see Fig. 1). This allows to find a point for which the Fisher metric expands, contradicting condition 2 and proving the claim.

There are a couple of remarks to be made: first, it should be noticed that since the Fisher information metrics are only defined on $\mathcal{P}_d^\circ$, condition 1 is minimal to introduce the concept of contractivity (encoded by

Eq. (12)). Moreover, the restriction on the set of ρ considered at the beginning of condition 2 is also minimal in order for the two sides of Eq. (12) to be regular, i.e., we only focus on matrices ρ for which both the Fisher metrics in ρ and in $\Phi(\rho)$ are well-defined. Finally, let us notice that the use of $\mathcal{P}_d$ as a domain was chosen to show the mathematical result in full generality, as well as to accommodate potentially trace non-preserving maps. In this sense, the theorem again isolates physical maps, which can only *decrease* the probability-normalization of states [26]. In the corollaries that follows we restrict, for clarity, to standard trace-preserving (TP) maps on the physical set of states $\mathcal{S}_d$.

Corollary 3.1. *Consider a Hermitian preserving, trace-preserving linear map $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$. If Φ satisfies the following two properties:*

 1. *it maps at least one point from $\mathcal{S}_d^\circ$ into $\mathcal{S}_d^\circ$;*

 2. *for all $\rho \in \mathcal{S}_d^\circ$ such that $\Phi(\rho) \in \mathcal{S}_d^\circ$, and for all $\delta\rho$ in the $\mathcal{S}_d$-tangent space, it contracts the Fisher metric as in Eq. (12);*

Then Φ is a positive map (PTP). Moreover, if $\Phi \otimes \mathbb{I}_d$, acting on $(\mathcal{S}_d \otimes \mathcal{S}_d)^\circ$, satisfies the same conditions, then Φ is a completely positive map (CPTP).

The proof of the corollary is contained at the end of Appendix 1, and it is equivalent to the one for Thm. 3. It should be noticed that Thm. 3 and Cor. 3.1 completely characterize physical evolutions: the first gives a condition for positivity, which is sufficient for classical systems, while the latter characterise CPTP maps, the relevant notion of maps for quantum systems.

We conclude by observing that the same results can be phrased in terms of *any of the contrast functions* introduced above. In fact, putting together Thm. 1 with Cor. 3.1, we have:

Corollary 3.2. *Consider a Hermitian preserving, trace preserving, linear map $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ mapping at least one point from $\mathcal{S}_d^\circ$ into $\mathcal{S}_d^\circ$. Consider any contrast function $H_g(\rho||\sigma)$. If for every two states ρ and σ in $\mathcal{S}_d^o$ it holds that:*

$$H_g(\rho||\sigma) \geq H_g(\Phi(\rho)||\Phi(\sigma)), \qquad (13)$$

then Φ is a positive map (PTP). Moreover, if the same holds for $\Phi \otimes \mathbb{I}_d$ then the map is completely positive (CPTP).

We refer to Appendix 2 for the precise proof, but this result simply follows from the fact that all contrast functions locally expand to Fisher metrics, hence it is sufficient to choose ρ and σ to be close-by matrices and then apply Thm. 3 to them. Cor. 3.2 shows that it is enough to consider any divergence to actually characterise the class of positive (or completely positive) maps. In fact, while our results are more general, a specific instance of Cor. 3.2 can be expressed as follows:

If a linear map acting on the set of states (including, in the quantum case, ancillas with trivial dynamics) cannot increase their relative entropy, then such map represents a physical evolution.

CONCLUSION

In the axiomatic definition of quantum statistical divergences, the monotonicity condition stands out as the only request which is not purely statistical in nature: indeed, whereas the other properties can be expressed solely in terms of states, in order to introduce monotonicity one needs to explicitly introduce the concept of noisy transformations. This fact might suggest that there is some hierarchy between physical evolutions and statistical quantifiers, since one needs the first to define the latter. Corroborating this belief, the Chentsov-Petz theorem (Thm. 2) constrains the local behaviour of statistical distances, singling out the Fisher information through a purely dynamical request.

Still, historically most of the contrast functions were introduced in a purely statistical setting, therefore their strong connection with dynamics is *a priori* not obvious and, indeed, quite remarkable. This manuscript, then, aims at clarifying this fortunate conjunction: Thm. 3 and its Cor. 3.1 show that physical evolutions can be singled out as the unique family of linear maps that contract the Fisher information. This result not only justifies the claim that physicality of evolution and statistical contractivity are equivalent notions of maps, but also hints at the source of this connection, as both dynamical maps and statistical distances arise from the underlying notion of conservation of information.

Moreover, Cor. 3.2 also shows how one could in principle define physical evolutions merely in terms of any contrast function, strengthening the message that there is no intrinsic hierarchy of fundamentality between statistical notions and dynamical ones. Still, it should be kept in mind that Thm. 2 is peculiar to the Fisher information, further corroborating the natural role this quantity plays both in quantum and classical physics.

In fact, it was noticed that many dynamical aspects of evolutions, as Markovianity or detailed balance, can be naturally phrased in terms of the behaviour of maps with respect to the Fisher metrics [13, 14]. The results presented here suggest that this is far from being a coincidence, showing that there is a fundamental connection which is worth exploring further in future works. Moreover, in the context of the reconstructions of quantum mechanics [27, 28], the fact that physical evolutions can be singled out in terms of a statistical quantity as the Fisher information might be of particular relevance for the derivation of a sensible notion of dynamics.

ACKNOWLEDGEMENTS

———————

* matteo@ifisc.uib-csic.es
† paolo.abiuso@oeaw.ac.at
‡ dario.desantis@sns.it
§ jsurace@perimeterinstitute.ca

[1] C. E. Shannon, "A mathematical theory of communication," The Bell System Technical Journal **27**, 379–423 (1948).

[2] A. S. Holevo, *Quantum Systems, Channels, Information. A Mathematical Introduction.* (De Gruyter, Berlin, Boston, 2013).

[3] S. Kullback, *Information Theory and Statistics*, Dover Books on Mathematics (Dover Publications, Mineola, NY, 1997).

[4] E. T. Jaynes, "Information theory and statistical mechanics," Phys. Rev. **106**, 620–630 (1957).

[5] H. S. Leff and A. F. Rex, *Maxwell's Demon: Entropy, Information, Computing*, Princeton Series in Physics (Princeton University Press, 2014).

[6] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, "Thermodynamics of information," Nature Physics **11**, 131–139 (2015).

[7] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information* (American Association of Physics Teachers, 2002).

[8] K. Życzkowski and I. Bengtsson, "On duality between quantum maps and quantum states," Open systems & information dynamics **11**, 3–42 (2004).

[9] G. Chiribella, E. Aurell, and K. Życzkowski, "Symmetries of quantum evolutions," Physical Review Research **3**, 033028 (2021).

[10] W. F. Stinespring, "Positive functions on C* -algebras," Proceedings of the American Mathematical Society **6**, 211 (1955).

[11] N. N. Chentsov, *Statistical decision rules and optimal inference* (American Mathematical Soc., 2000).

[12] L. L. Campbell, "An extended Čencov characterization of the information metric," Proceedings of the American Mathematical Society **98**, 135–141 (1986).

[13] Paolo Abiuso, Matteo Scandi, Dario De Santis, and Jacopo Surace, "Characterizing (non-)Markovianity through Fisher information," SciPost Phys. **15**, 014 (2023).

[14] M. Scandi, P. Abiuso, J. Surace, and D. De Santis, "Quantum Fisher information and its dynamical nature," arXiv preprint arXiv:2304.14984 (2023).

[15] H. Cramer, *Mathematical methods of statistics* (Princeton University Press, 1946).

[16] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," Kotz S., Johnson N.L. (eds) Breakthroughs in Statistics. Springer Series in Statistics (Perspectives in Statistics) (1992), 10.1007/978-1-4612-0919-5.

[17] A. Ly, M. Marsman, J. Verhagen, R. P. P. P. Grasman, and E.-J. Wagenmakers, "A Tutorial on Fisher information," Journal of Mathematical Psychology **80**, 40–55 (2017).

[18] Also known as f-divergences.

[19] I. Csiszár and P. C. Shields, *Information theory and statistics: a tutorial*, Vol. 1 (Now Publishers, Inc., 2004) pp. 417–528.

[20] D. Petz, "Quasi-entropies for finite quantum systems," Reports on Mathematical Physics **23**, 57–65 (1986).

[21] F. Hiai and D. Petz, *Introduction to matrix analysis and applications*, Universitext (Springer, 2014).

[22] A. Lesniewski and M. B. Ruskai, "Monotone Riemannian metrics and relative entropy on non-commutative probability spaces," Journal of Mathematical Physics **40**, 5702–5724 (1999).

[23] D. Petz, "Monotone metrics on matrix spaces," Linear Algebra and its Applications **244**, 81–96 (1996).

[24] Matteo G. A. Paris, "Quantum estimation for quantum technology," International Journal of Quantum Information **07**, 125–137 (2009).

[25] Koenraad MR Audenaert, John Calsamiglia, Ramón Munoz-Tapia, Emilio Bagan, Ll Masanes, Antonio Acin, and Frank Verstraete, "Discriminating states: The quantum Chernoff bound," Physical Review Letters **98**, 160501 (2007).

[26] Maps that decrease the trace correspond to evolutions with post-selection, i.e., with a non-null probability of discarding the system.

[27] L. Hardy, "Reconstructing quantum theory," in *Quantum theory: informational foundations and foils* (Springer, 2015) pp. 223–248.

[28] M. P. Mueller and L. Masanes, "Information-theoretic postulates for quantum theory," in *Quantum theory: informational foundations and foils* (Springer, 2015) pp. 139–170.

Methods

1. Proof of Thm. 3 and Cor. 3.1

In this appendix we present the proof of Thm. 3, which we repeat here for convenience:

Theorem. Consider a Hermitian preserving, linear map $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$. If Φ satisfies the following two properties:

    1. Φ maps at least one point from $\mathcal{P}_d^\circ$ into $\mathcal{P}_d^\circ$;

    2. for any matrix ρ in $\mathcal{P}_d^\circ \cap \big(\Phi^{-1}\big(\mathcal{P}_d^\circ \cap \Phi(\mathcal{P}_d)\big)\big)$, and any tangent vector $\delta\rho$, one has:

$$K_{f,\rho}(\delta\rho, \delta\rho) \geq K_{f,\Phi(\rho)}(\Phi(\delta\rho), \Phi(\delta\rho)) \,; \tag{14}$$

then the image of Φ is completely contained in $\mathcal{P}_d$, meaning that Φ is a positive map (P). Moreover, Φ is trace non-increasing.

Proof. We prove Theorem 3 by contradiction: suppose there exists a map Φ that is not positive, but that satisfies Eq. (14). Thanks to condition 1 there exists at least one point π in $\mathcal{P}_d^\circ$ such that its evolution is also in the interior of $\mathcal{P}_d$. Moreover, from the assumption that Φ is not P there is also at least one matrix σ such that $\Phi(\sigma) \notin \mathcal{P}_d$. Without loss of generality one can choose σ to be in the interior of $\mathcal{P}_d$: if this is not the case we take a ball around σ small enough so that its image still lays outside of the state space. Then, by inspecting the intersection between the preimage of this ball and $\mathcal{P}_d^\circ$ one can find a point satisfying the assumption. Since both ρ and σ are in $\mathcal{P}_d^\circ$, the line $\rho_\lambda := (1 - \lambda)\pi + \lambda\sigma$ is completely contained in $\mathcal{P}_d^\circ$ for $\lambda \in [0, 1]$. This implies that the following superior is finite:

$$\sup_{\lambda, \mathrm{Tr}[\delta\rho^2]=1} K_{f,\rho_\lambda}(\delta\rho, \delta\rho) = \sup_{\lambda, \mathrm{Tr}[\delta\rho^2]=1} \mathrm{Tr}\left[\delta\rho \, \mathbb{J}_f^{-1}\big|_{\rho_\lambda}[\delta\rho] \right] < \infty \,, \tag{15}$$

since the Fisher information is a bounded operator when restricted to a closed set completely inside of $\mathcal{P}_d^\circ$ (see [14]). By varying λ, $\Phi(\rho_\lambda)$ interpolates linearly between the positive definite matrix $\Phi(\pi)$ and one with at least one negative eigenvalue, namely $\Phi(\sigma)$. Then, there exists a λ^* such that $\Phi(\rho_{\lambda^*})$ is a matrix with at least one zero-eigenvalue. We are now ready to prove the claim. Set the state ρ_η so that the smallest eigenvalue ϵ_η of $\Phi(\rho_\eta)$ is of order η, where $\eta \ll 1$. Call the associated eigenvector $|\psi_\eta\rangle$. Then, choose a perturbation $\delta\rho_\eta$ such that, for any η, $\Phi(\delta\rho_\eta)$ has a positive and finite contribution along the eigenvectors corresponding to η-eigenvalues (the positivity condition ensures that in the limit $\eta \to 0$ the perturbed state is still in the interior of $\mathcal{P}_d^\circ$ for any finite η, allowing for the definition of the Fisher information). One can always find such a perturbation. Suppose, in fact, the contrary. Then, this would mean that for any $\delta\rho$, it holds that:

$$\lim_{\eta \to 0} \langle\psi_\eta| \Phi(\delta\rho) |\psi_\eta\rangle = 0 \quad \forall \delta\rho. \tag{16}$$

But any $\delta\rho$ can always be rewritten as $\delta\rho = \sigma_1 - \sigma_2$, where $\sigma_{1/2} \in \mathcal{P}_d$. Then, Eq. (16) would imply that for any two matrices in $\mathcal{P}_d$ it holds that:

$$\langle\psi_0| \Phi(\sigma_1) |\psi_0\rangle = \langle\psi_0| \Phi(\sigma_2) |\psi_0\rangle \quad \forall \sigma_1, \sigma_2 \in \mathcal{P}_d \,. \tag{17}$$

Set $\sigma_2 := \rho_{\lambda^*}$, which, by definition, satisfies $\langle\psi_0| \Phi(\rho_{\lambda^*}) |\psi_0\rangle = 0$. This implies that for any matrix in $\sigma \in \mathcal{P}_d$, the image of σ is not full-rank, since $\langle\psi_0| \Phi(\sigma) |\psi_0\rangle = 0$, which in turns contradicts the hypothesis that at least one point in $\mathcal{P}_d^\circ$ is mapped inside $\mathcal{P}_d^\circ$. Thus, there exists at least one $\delta\rho_\eta$ for which Eq. (16) does not hold. Then, choosing such a vector, one can see that the evolved Fisher information is unbounded as $\eta \to 0$ by explicit inspection of its coordinates expression:

$$K_{f,\Phi(\rho_\eta)}(\Phi(\delta\rho_\eta), \Phi(\delta\rho_\eta)) = \sum_{i \neq j} \frac{|\Phi(\delta\rho_\eta)_{ij}|^2}{\Phi(\rho_\eta)_{jj} f(\Phi(\rho_\eta)_{ii}/\Phi(\rho_\eta)_{jj})} + \sum_i \frac{|\Phi(\delta\rho_\eta)_{ii}|^2}{\Phi(\rho_\eta)_{ii}} \tag{18}$$

where we used the formula for the Fisher information in the orthonormal basis of $\Phi(\rho_\eta)$ provided in [14]. All the terms in the two sums are positive. Moreover, the second sum diverges at least as $\mathcal{O}\left(\eta^{-1}\right)$ as $\eta \to 0$, due to the contribution associated to the eigenvalue of $|\psi_\eta\rangle$. Hence, we can always find a η small enough such that:

$$K_{f,\Phi(\rho_\eta)}(\Phi(\delta\rho_\eta), \Phi(\delta\rho_\eta)) > \sup_{\lambda, \mathrm{Tr}[\delta\rho^2]=1} K_{f,\rho_\lambda}(\delta\rho, \delta\rho) \geq K_{f,\rho_\eta}(\delta\rho_\eta, \delta\rho_\eta), \tag{19}$$

contradicting the assumption that Φ contracts the Fisher metric for any two points in $\mathcal{P}_d^\circ$. This proves the Positivity of Φ. Moreover, choosing $\delta\rho \equiv \rho$ in Eq. (14), one obtains:

$$\mathrm{Tr}\,[\rho] = K_{f,\rho}(\delta\rho, \delta\rho) \geq K_{f,\Phi(\rho)}(\Phi(\delta\rho), \Phi(\delta\rho)) = \mathrm{Tr}\,[\Phi(\rho)], \tag{20}$$

thanks to the property that $\mathbb{J}_f^{-1}\big|_\rho[\rho] = \mathbb{I}$ [14]. This proves that Φ is trace non-increasing. $\qquad\square$

Finally, regarding Cor. 3.1, it should be noticed that the only difference is that the tangent space of $\mathcal{S}_d$ is given by traceless Hermitian operators. Thus, whereas tangent vectors to $\mathcal{P}_d$ can be written as $\delta\rho = \sigma_1 - \sigma_2$ for $\sigma_{1/2} \in \mathcal{P}_d$, in the case of $\mathcal{S}_d$ the same decomposition still holds, but with $\sigma_{1/2} \in \mathcal{S}_d$. Then, one can prove the existence of the perturbation $\delta\rho_\eta$ in the very same way as we did for $\mathcal{P}_d$.

2. Proof of Corollary 3.2

We prove Corollary 3.2, which we repeat here for convenience:

Corollary. *Consider a Hermitian preserving, linear map $\Phi : \mathcal{M}_d(\mathbb{C}) \to \mathcal{M}_d(\mathbb{C})$ satisfying condition 1 in Thm. 3. Consider any contrast function $H_g(\rho||\sigma)$. If for every two matrices ρ and σ on $\mathcal{P}_d^o$ it holds that:*

$$H_g(\rho||\sigma) \geq H_g(\Phi(\rho)||\Phi(\sigma)), \tag{21}$$

then Φ is a positive map. Moreover, if the same holds for $\Phi \otimes \mathbb{I}_d$ then the map is completely positive.

Proof. The proof of the corollary directly follows from the one of Thm. 3. Indeed, by contradiction, assume there exists a matrix $\sigma \in \mathcal{P}_d^\circ$ that gets mapped outside of $\mathcal{P}_d$, and consider $\pi \in \mathcal{P}_d^\circ$ such that $\Phi(\pi) \in \mathcal{P}_d^\circ$. Analogously to the proof above, define $\rho_\lambda := (1-\lambda)\,\pi + \lambda\,\sigma$. For ε small enough we can apply Thm. 1 to obtain:

$$\sup_{\lambda, \mathrm{Tr}[\delta\rho^2]=1} H_g(\rho_\lambda||\rho_\lambda + \varepsilon\,\delta\rho) = \sup_{\lambda, \mathrm{Tr}[\delta\rho^2]=1} \frac{\varepsilon^2}{2}\,K_{f,\rho_\lambda}(\delta\rho, \delta\rho) + \mathcal{O}\left(\varepsilon^3\right) < \infty. \tag{22}$$

Then, following the steps presented for Thm. 3, we define a ρ_η such that $\Phi(\rho_\eta)$ is of order η, where $\eta \ll 1$, and a $\delta\rho_\eta$ having a positive finite contribution along the eigenvectors corresponding to η-eigenvalues. The positivity condition ensures that in the limit $\eta \to 0$ the perturbed state is still in the interior of $\mathcal{P}_d$ for any finite η, and ε small enough. Then, $H_g(\Phi(\rho_\eta)||\Phi(\rho_\eta + \varepsilon\,\delta\rho_\eta))$ scales as ε^2/η. Hence, one can find η small enough such that:

$$H_g(\Phi(\rho_\eta)||\Phi(\rho_\eta + \varepsilon\,\delta\rho_\eta)) > \sup_{\lambda, \mathrm{Tr}[\delta\rho^2]=1} H_g(\rho_\lambda||\rho_\eta + \varepsilon\,\delta\rho) \geq H_g(\rho_\eta||\rho_\eta + \varepsilon\,\delta\rho_\eta). \tag{23}$$

This gives the desired contradiction, concluding the proof. $\qquad\square$