Quantifying the relative incompatibility of quantum observables

Georgios Styliaris and Paolo Zanardi
Department of Physics and Astronomy, and Center for Quantum Information Science and Technology, University of Southern California, Los Angeles, California 90089-0484
(Dated: January 21, 2019)

Motivated by quantum resource theories, we introduce a preorder over orthonormal bases in a $d$-dimensional Hilbert space which intuitively corresponds to the degree of incompatibility between observables that are diagonal in those bases. The preorder is directly related to multivariate majorization and gives rise to families of monotones, i.e., scalar quantifiers that preserve the ordering.

We relate the preorder with measures of quantum coherence, the strength of quantum fluctuations and entropic uncertainty relations.

Introduction.— One of the cornerstones of quantum theory is the concept of incompatibility between quantum observables [1]. A pair of quantum observables is deemed incompatible if the corresponding self-adjoint operators fail to commute. Operationally, incompatibility implies that it is impossible to simultaneously predict with certainty two sequences of measurement outcomes, where each sequence corresponds to orthogonal measurements of one of the observables, over identically prepared states. Finite-dimensional observables that share the same eigenbasis are fully compatible, while any pair of observables associated with bases that are mutually unbiased are maximally incompatible: certain knowledge for the outcome of one assures complete randomness for the possible outcomes of the other.

Incompatibility of quantum observables is famously captured through uncertainty relations, that may involve variances [2][4], entropic quantities [5][10] or other information-theoretic quantities [11][15]. Alternative approaches yielding a quantitative description of incompatibility exist in the context of state discrimination and quantum steering [16][24], as well as in more general operational theories [25][26].

In this work, we set a framework for the quantification of the relative incompatibility of quantum observables, relevant in the context of orthogonal measurements, by following the paradigm of quantum resource theories [27]. There, a quantum property (e.g., entanglement) is fully described by the interconversion relations among states under a class of quantum processes that, suitably chosen, cannot enhance the desired property. The induced relations among quantum states can be mathematically described by a preorder [28], which is equivalent to (a usually infinite) family of scalar functions that jointly quantify the resourcefulness of states. Imitating this approach, we will equip classes of commuting observables with an operationally motivated preorder, which captures the notion of incompatibility. The preorder will naturally give rise to families of scalar functions, that jointly quantify incompatibility. Our approach utilizes tools from the theory of multivariate majorization and provides a quantitative, as well as conceptual, connection between incompatibility, quantum coherence and uncertainty relations.

Preliminaries.— Consider an observable $A$ over a finite dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^d$ with spectral decomposition $A = \sum_{i=1}^{d} a_i P_i$ (we denote $P_i := |i\rangle\langle i|$. The role of the eigenvalues $a_i$ is to label the possible outcomes and, as long as they are distinct, this role is unimportant from the point of view of the measurement process, since the probability distribution $p_B(\rho)$ with components $[p_B(\rho)]_i := \text{Tr} (P_i \rho)$ (representing a measurement of $A$ in state $\rho$) only depends on the set of projectors $B = \{P_i\}_i$ [29]. We will henceforth use the term basis (always meaning orthonormal) to refer to a set of rank-1 orthogonal projectors $\{P_i\}_{i=1}^d$, with $\sum_i P_i = I$ [30]. We associate with every basis $B$ the real abelian algebra of observables $A_B$ generated by (real linear combinations of) $\{P_i\}_i$. The set of bases over the Hilbert space is denoted by $\mathcal{M}(\mathcal{H})$.

Any basis $B$ gives rise to a corresponding dephasing or measurement quantum map (completely positive and trace preserving superoperator)

$$D_B(X) := \sum_i P_i X P_i. \quad (1)$$

The latter can be thought of as a non-selective orthogonal measurement of any non-degenerate observable belonging in $A_B$, while a composition $D_{B_n} \cdots D_{B_1}$ represents the quantum operation associated with $n$ such successive measurements. Moreover, the basis $B$ corresponding to a dephasing map $D_B$ is unique, i.e., the mapping $B \mapsto D_B$ is injective [31], and similarly for $B \mapsto A_B$ [32]. This establishes a correspondence between bases, classes of commuting observables and measurement superoperators.

Preorder and monotones.— The idea of deriving families of scalar functions that quantify some feature (for instance, the degree of uniformity of a probability distribution) by invoking a preorder has its roots in the mathematical theory of majorization [33]. Such a paradigm has been extensively employed in quantum information in the context of resource theories for comprehending and also quantifying features of quantum systems, such as entanglement [34], coherence [35] and out-of-equilibrium thermodynamics [36].

In this approach, one distinguishes a class of quantum operations, deemed as “easy”, motivated by some
practical consideration. For example, in the case of entanglement, the easy operations are local quantum operations between two parties together with classical communication (LOCC), which reflect the fact that establishing quantum channels between spatially separated parties poses a difficulty. This set of maps induces a preorder "\( \geq \)" in the set of quantum states, defined by the allowed transitions under easy operations, namely \( \rho \geq \sigma \) if and only if there exists an easy operation \( E \) such that \( \sigma = E(\rho) \). The binary relation induced is indeed a preorder since, by definition, the identity quantum channel is always an easy operation and also the composition of easy operations is again an easy operation. Moreover, \( \rho \geq \sigma \) should intuitively correspond to a statement like "\( \rho \) is more entangled than \( \sigma \)." This quantification is rigorously captured by the notion of monotones, i.e., scalar functions \( f \) over states, non-increasing under allowed state transitions \( \rho \geq \sigma \implies f(\rho) \geq f(\sigma) \) [27]. In general, no single monotone function can fully capture the induced preorder (in the sense \( f(\rho) \geq f(\sigma) \iff \rho \geq \sigma \)) as that would imply the ordering is necessarily total. In words, there is no unique scalar measure for the property captured by a preorder, as for instance, there is no unique measure of entanglement. Nevertheless, families of monotones \( \{f_\alpha\}_\alpha \) can form a complete set, namely satisfy \( f_\alpha(\rho) \geq f_\alpha(\sigma) \forall \alpha \iff \rho \geq \sigma \) [27].

A preorder over orthonormal bases.— We would like to give meaning to statements such as "the basis \( B_0 \) is more compatible with \( B_1 \) than with \( B_2 \)." Intuitively, such a statement should imply that for states \( \rho_0 = \sum_i p_i \rho_i^{(0)} \in \mathcal{A}_{B_0} \), the probability distribution \( p \) is "more similar" to \( p_{B_1}(\rho_0) \) than \( p_{B_2}(\rho_0) \) (\( p_E(\rho) \) denotes the probability distribution associated with an \( \mathcal{A}_E \) measurement in state \( \rho \)). For instance, consider the case \( B_0 = B_1 \) (maximal compatibility) and \( B_2 \) a basis mutually unbiased to \( B_0 \) (maximal incompatibility). Indeed, in this extremal case, \( p_{B_1}(\rho_0) = p \) while \( p_{B_2}(\rho_0) = \frac{1}{d}(1, \ldots, 1) \) is drastically altered to the maximally mixed one.

We now give the general definition. We say that \( B_0 \) is be more compatible with \( B_1 \) than \( B_2 \) if, for any initial state diagonal in \( B_0 \), the output of a non-selective \( B_2 \) measurement can be emulated by a non-selective \( B_1 \) measurement, followed possibly by an additional sequence of non-selective measurements and a unitary rotation. More precisely, we have the following.

**Definition 1.** We denote \( B_1 \succ_{B_0} B_2 \) if and only if there exist a unitary superoperator \( U \) and a (possibly trivial) finite sequence of measurements \( \{D_{B_{\alpha}}\}_\alpha \) such that

\[
D_{B_2}D_{B_0} = U \prod_\alpha D_{B_{\alpha}} D_{B_1} D_{B_0} .
\]  

(2)

The definition has the following immediate consequences.

**Proposition 1.** (i) The binary relation "\( \succ_{B_0} \)" over \( \mathcal{M}(\mathcal{H}) \) is a preorder; (ii) \( B_0 \succ_{B_0} B \) for all bases \( B \); (iii) \( B \succ_{B_0} B_{MU} \) for all bases \( B \), where \( B_{MU} \) is any mutually unbiased basis of \( B_0 \).

All proofs can be found in the appendix.

The preorder "\( \succ_{B_0} \)" is not, in general, a partial order [38]. For any pair of bases mutually unbiased to \( B_0 \) it simultaneously holds \( B_{MU_1} \succ_{B_0} B_{MU_2} \) and \( B_{MU_1} \succ_{B_0} B_{MU_2} \) without necessarily \( B_{MU_1} = B_{MU_2} \). The case \( d = 2 \) is simple, as one can invoke the usual Bloch ball representation of quantum states \( \rho = \frac{1}{2}(I + \mathbf{v} \cdot \sigma) \), where different bases are in one to one correspondence with lines passing at the center. In this representation, the action of \( D_{B_1} \) on a state \( \rho \) coincides with projecting \( \mathbf{v} \) onto the \( B_1 \) line while the action of \( U \) is translated into an \( SO(3) \) rotation. Clearly, Eq. (2) can be satisfied (in fact, by means of a single \( D_{B_{1}} \)) if and only if \( \theta_1 \leq \theta_2 \); here \( \theta_i \) is the (acute) angle between the lines corresponding to \( B_0 \) and \( B_i \). In particular, for \( d = 2 \) "\( \succ_{B_0} \)" is a total preorder but not for \( d > 2 \), as it will become clear momentarily.

"\( \succ_{B_0} \)" and multivariate majorization.— Given any state \( \rho_0 = \sum_i p_i \rho_i^{(0)} \) diagonal in the \( B_0 \) basis, the probability distribution \( p_{B_1}(\rho_0) \) of possible measurement outcomes in some other basis \( B_1 \) is more "uniform". This is precisely captured by the majorization state- ment \( p_{B_{0}}(\rho_0) \succ p_{B_{1}}(\rho_0) \) that is true for any basis \( B_1 = \{P_i^{(1)}\}_i \). More explicitly,

\[
[p_{B_1}(\rho_0)]_i = \sum_j X_{ij}(B_1, B_0) p_j ,
\]

(3)

for \( X_{ij}(B_1, B_0) := \text{Tr} \left( P_i^{(1)} P_j^{(0)} \right) \), hence \( p_{B_1}(\rho_0) \) is obtained from \( p \) by the action of a bistochastic matrix, from which majorization follows [39]. Notice that the ordering of the projectors is arbitrary, therefore the X matrix is non-unique up to permutations.

As we will show now, \( B_1 \succ_{B_0} B_2 \) implies that the bistochastic matrix \( X(B_2, B_0) \) factorizes in the sense that there exists a matrix \( M \), also bistochastic, such that \( X(B_2, B_0) = MX(B_1, B_0) \). We remind the reader that a bistochastic matrix \( A_{ij} \) is unistochastic [40] if there exists a unitary matrix \( U_{ij} \) such that \( A_{ij} = ||U_{ij}||^2 \).

**Proposition 2.** Let \( X_{ij}(B_\beta, B_\alpha) := \text{Tr} \left( P_i^{(\beta)} P_j^{(\alpha)} \right) \) denote the bistochastic matrix resulting from an ordering of the bases \( B_\alpha = \{P_i^{(\alpha)}\}_i \) and \( B_\beta = \{P_i^{(\beta)}\}_i \). Then, \( B_1 \succ_{B_0} B_2 \) if and only if there exists a sequence of unistochastic matrices \( \{A^{(\alpha)}\}_\alpha \) such that

\[
X(B_2, B_0) = \prod_\alpha A^{(\alpha)} X(B_1, B_0) .
\]

(4)

A similar preorder between matrices has been studied in the literature in the context of multivariate majorization, called matrix majorization [39]. We write \( A \succ C \) for matrices \( A \) and \( C \) if there exists a bistochastic \( B \) such
that \( C = BA \). Clearly, Eq. \([4]\) implies matrix majorization \( X(B_1, B_0) \succ X(B_2, B_0) \). The following Lemma will help establish a converse statement.

**Lemma 1.** Every bistochastic matrix can be approximated arbitrarily well by a product of unistochastic matrices.

As a result, if one broadens \([1]\) to include the cases where \( ||DB_2DB_0 - UD \prod_{i \neq j} DB_i DB_j || \) can be made arbitrarily small, the resulting preorder over \( X(B_1, B_0) \) and \( X(B_2, B_0) \) coincides with matrix majorization.

**Proposition 2** has the following direct implication.

**Corollary 1.** Let \( B_1 \succ B_0 \). Then, for any state of the form \( \rho_0 = \sum_i p_i \rho_i \), there exists a bistochastic matrix \( M \), independent of \( p \), such that

\[
P_{B_2}(\rho_0) = MP_{B_1}(\rho_0).
\]

In this case, the preorder \( \rho_0 \) is equivalent to \( \rho_0 \). Therefore, for Schur-concave functions, which for instance include Rényi entropies for \( \alpha \geq 0 \) (\( \alpha = 1 \) corresponds to the usual Shannon entropy), satisfy \( f(p_{B_1}(\rho_0)) \leq f(p_{B_2}(\rho_0)) \). The converse statement in terms of majorization has been stated. A state \( \rho_0 \) is not enough to assure \( B_1 \succ B_2 \). A specific counterexample was constructed in the context of multivariate majorization by Horn in \([11]\).

**Measures of relative (in)compatibility.**— A preorder gives rise to a distinguished class of scalar functions, i.e., monotones. We adopt the following definition.

**Definition 2.** A function \( f_{B_2} : \mathcal{M}(\mathcal{H}) \to \mathbb{R}^+ \) is measure of relative compatibility (incompatibility) if it is convex (concave) with respect to the preorder \( a \succ B_0 \) and vanishes for all pairs of bases mutually unbiased to \( B_0 \) (vanishes for \( B_0 \), i.e., \( B_1 \succ B_2 \) \( \implies \) \( f_{B_2}(B_1) \geq f_{B_2}(B_2) \) \( (B_1 \succ B_2 \implies f_{B_2}(B_1) \leq f_{B_2}(B_2)) \) and \( f_{B_2}(B_{MU}) = 0 \) \( (f_{B_2}(B_0) = 0) \). Moreover, if \( f(B_0, B_1) = f_{B_2}(B_1) = f_{B_1}(B_0) \), we call it a symmetric measure of relative compatibility (incompatibility).

The following Proposition gives a construction for measures of relative compatibility arising from convex functions. Clearly, the analogous claims hold for the incompatibility case in terms of concave functions.

**Proposition 3.** Let \( \phi : \mathbb{R}^d \to \mathbb{R} \) be a continuous convex function with \( \phi([0,1,\ldots,1]) = 0 \). Then,

\[
f_{B_0}^{(2)}(B_1) \coloneqq \sum_{i=1}^d \phi(X_{i}(B_1, B_0))
\]

is a measure of relative compatibility; here, \( X_{i} \) stand for the row vectors of the matrix \( X_{i,j} \).

In fact, the family \( \{f_{B_0}^{(2)}(B_1)\}_\phi \) for all continuous convex \( \phi \) turns out to be a complete family of monotones for matrix majorization \([22]\), hence also for the preorder \( a \succ B_0 \) if one includes approximate relations (see earlier discussion).

One can restrict the family of monotones \( f_{B_0}^{(2)}(B_1) \) by demanding \( \phi(\rho) = \sum_i \phi_i\langle v_i | \) for convex \( \phi_i : \mathbb{R} \to \mathbb{R}^+ \). By a choice \( \phi_i = \psi \), one obtains

\[
g_\psi(B_0, B_1) := f_{B_0}^{(2)}(B_1) = f_{B_1}^{(2)}(B_0)
\]

which are symmetric measures of the relative compatibility.

**Incompatibility and coherence.**— Quantum coherence refers to the property of quantum systems to exist in a linear superposition of different physical states. It is a notion defined with respect to some preferred, physically relevant basis, which we will denote as \( B_0 \), unless otherwise stated. A state \( \rho \) is said to be coherent if there exist non-vanishing off-diagonal elements when \( \rho \) is expressed as a matrix in \( B_0 \). Recently, coherence was formulated as a resource theory \([19]\) and this framework gave rise to families of functions that are monotonic with respect to the preorder induced by state conversion under free operations. One of the central measures in the theory is relative entropy of coherence \( c_{B_0}^{(rel)}(\rho) := S(\rho \| DB_0 \rho) \) that admits several operational interpretations in terms of conversion rates \([41, 42, 45]\). Later, we will also invoke the 2-coherence \( c_{B_0}^{(2)}(\rho) := \sum_{i \neq j} |\rho_{ij}|^2 \).

The amount of coherence of \( \rho_0 \) in some basis \( B_1 \) is closely related with the mixedness of the probability distribution \( p_{B_1}(\rho_0) \), and therefore should also be connected with the degree of incompatibility between \( B_1 \) and \( B_0 \). For instance, the Schur-Horn theorem \([17]\) implies that \( p_{B_1}(\rho_0) \) is the least mixed when \( \rho_0 \) is diagonal (i.e., coherent) in \( B_1 \), namely \( p_{B_1}(\rho_0) = p_{B_1}(\rho_0) \) for all \( B_1 \). We now establish a first connection between coherence and the preorder \( a \succ B_0 \).

**Proposition 4.** (i) Let \( B_1 \succ B_0 \) and \( \rho_0 = \sum_i p_i \rho_{0,i} \). Then, \( c_{B_1}^{(rel)}(\rho_0) \leq c_{B_2}^{(rel)}(\rho_0) \). (ii) Similarly, \( c_{B_1}^{(2)}(\rho_0) \leq c_{B_2}^{(2)}(\rho_0) \). (iii) The average relative entropy of coherence \( \langle c_{B_1}^{(rel)}[DB_0(\rho_0)\rho_0)] \rangle \) over Haar distributed pure states \( |\psi\rangle \) is the measure of relative incompatibility between \( B_0 \) and \( B_1 \). (iv) Similarly, for the average 2-coherence \( \langle c_{B_1}^{(2)}[DB_0(\rho_0)\rho_0)] \rangle \) over Haar distributed pure states \( |\psi\rangle \) is the measure of relative incompatibility between \( B_0 \) and \( B_1 \).

In addition to the interpretation of Proposition 4 in the framework of coherence, one can also infer from (i) above that a \( DB_1 \) measurement disturbs less \( \rho_0 \) compared to a \( DB_2 \) measurement, if \( B_1 \succ B_0 \), as it is precisely captured by statistical meaning of the relative entropy \([15]\).

**Incompatibility and uncertainty.**— Incompatibility of quantum observables has been, since the early stages of
quantum theory, associated with the existence of quantum fluctuations. Let us consider the quantity

$$Q_{B_0}(B_1) := \sup_{A \in A_{B_0}, \|A\|_1 = 1} \max_{i=1, \ldots, d} \text{Var}_i(A),$$

(8)

where $$\text{Var}_i(A) := \left| \text{Tr} \left( P_i^{(0)} A^2 \right) - \left( \text{Tr} \left( P_i^{(0)} A \right) \right) ^2 \right|,$$

that captures the strength of the uncertainty of a pure state diagonal in $$B_0$$ over a $$B_1$$ measurement. In the appendix we derive the upper bound

$$Q_{B_0}(B_1) \leq 1 - \lambda_{\min}(X(B_1, B_0)X^T(B_1, B_0)) := q_{B_0}(B_1)$$

(9)

($$\lambda_{\min}(X)$$ stands for the minimum eigenvalue of $$X$$). The bound satisfies $$q_{B_0}(B_1) = 0$$ if and only if $$B_1 = B_0$$, hence it vanishes if and only if $$Q_{B_0}(B_1)$$ vanishes. In addition, $$q_{B_0}(B_1)$$ admits an alternative interpretation as a measure of the ability of the unitary connecting the two bases to generate coherence out of incoherent states [19, 40].

An alternative way to quantify quantum fluctuations over different bases involves entropic uncertainty relations [10]. There, one tries to impose bounds over entropic quantities, such as $$S_{ab}(p_{B_1}(\rho_0)) + S_{ab}(p_{B_2}(\rho_0)) \leq r_{B_0}(B_2, B_1) = -\log(\max_{i,j} X_{ij}(B_2, B_1))$$ for any $$\alpha, \beta \geq 1/2$$ with $$1/\alpha + 1/\beta = 2$$.

The bound has recently been improved by Coles et al. [8] for the Shannon entropy case, as

$$S(p_{B_1}(\rho_0)) + S(p_{B_2}(\rho_0)) + S(\rho_0) \leq r^{(MU)}(B_2, B_1).$$

In all cases, the preorder $$B_1 \succ B_0, B_2$$ implies that the corresponding bounds looses.

**Proposition 5.** (i) If $$B_1 \succ B_0, B_2$$, then $$q_{B_0}(B_1) \leq q_{B_0}(B_2)$$ and also $$q_{B_0}(B_1) = q_{B_0}(B_0)$$, i.e., it is a symmetric measure of relative incompatibility. (ii) If $$B_2 \succ B_1, B_3$$ then $$r^{(MU)}(B_2, B_1) \leq r^{(MU)}(B_3, B_1)$$ and $$r^{(MU)}(B_1, B_2) = r^{(MU)}(B_2, B_1)$$, i.e., it is a symmetric measure of relative incompatibility.

**Mutual unbiasedness.**— In [51], the authors considered a geometrically motivated measure of “mutual unbiasedness” between pairs of orthonormal bases. Their measure is a distance and arises naturally by treating each basis as a subspace ($$B = \text{Span}\{P_i\}_{i=1}^d$$) and then embedding the corresponding Grassmannian into a vector space equipped with a Euclidean distance. Their resulting measure is $$f_{\text{DEBH}}(B_0, B_1) = d - \text{Tr} \left( X(B_1, B_0)X^T(B_1, B_0) \right)$$ (see also [49] for an interpretation of the same function in terms of coherence).

Incompatibility between bases is closely related to mutually unbiased bases (e.g., as in Proposition 1), hence it is no surprise that the measure arises as a prototypical case out of the (concave equivalent) family [6]. In our notation, $$f_{\text{DEBH}} = g^{\psi_2}(B_0, B_1)$$ for $$\psi_2(x) = \frac{1}{2} - x^2$$.

**Other measures.**— Utilizing Proposition 2 one can construct additional measures of incompatibility. For instance, the function $$F(B_0, B_1) := |\text{det}(X(B_1, B_0))|$$ is a symmetric measure of relative compatibility, that behaves similar to a fidelity function between bases. In fact, $$\text{arccos}(F(B_1, B_0))$$ is distance over $$\mathcal{M}(\mathcal{H})$$, in precise correspondence with the usual Fubini-Study one, as demonstrated by the geometric construction in [32]. By the usual geometrical meaning of the determinant, $$F(B_1, B_0)$$ admits a direct interpretation as the fraction of the Euclidean volume that the set $$\{p_{B_1}(\rho_0)\}_{\rho_0}$$ occupies in the $$(d-1)$$-dimensional simplex (embedded in $$\mathbb{R}^d$$) for all states $$\rho_0$$ diagonal in $$B_0$$. As a result, “$$\succ B_0$$” implies that the more incompatible a basis is with respect to $$B_0$$, the more “shrunk” the set of accessible distributions becomes.

**Conclusion.**— Quantum resource theories seem to suggest that an appropriate quantification of quantum properties, even conceptually simple ones such as the “uniformity” of a state [52], cannot be achieved by means of a single scalar quantifier. Instead, only an infinite set of functions is able to capture such properties in their wholeness, as they naturally result out of preorders. In this work, we defined an operationally motivated preorder on the set of orthonormal bases that captures the notion of relative incompatibility between the bases, or equivalently, commuting families of quantum observables. Our approach uncovers a quantitative, as well as conceptual, connection between incompatibility, uncertainty relations and quantum coherence unified under the prism of multivariate majorization.

**Acknowledgements.**— G.S. is thankful to N.A. Rodríguez-Briones for helpful discussions. P.Z. acknowledges partial support from the NSF award PHY-1819189.

---

[1] D. Bohm, *Quantum theory* (Courier Corporation, 2012).
[2] H. P. Robertson, Phys. Rev. 34, 163 (1929).
[3] E. Schrödinger, Sitzungsberichte der Preußischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse 14, 296.
[4] L. Maccone and A. K. Pati, Phys. Rev. Lett. 113, 260401 (2014).
[5] D. Deutsch, Phys. Rev. Lett. 50, 631 (1983).
[6] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. 60, 1103 (1988).
[7] M. Christandl and A. Winter, in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on* (IEEE, 2005) pp. 879–883.
[8] P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak, Phys. Rev. Lett. 108, 210405 (2012).
[9] K. Korzekwa, M. Lostaglio, D. Jennings, and T. Rudolph, Phys. Rev. A 89, 042122 (2014).
Appendix: Proofs

Proposition 1. (i) The binary relation "\(\succ^{B_0}\)" over \(\mathcal{M}(\mathcal{H})\) is a preorder; (ii) \(B_0 \succ^{B_0} B\) for all bases \(B\); (iii) \(B \succ^{B_0} B_{\text{MU}}\) for all bases \(B\), where \(B_{\text{MU}}\) is any mutually unbiased basis of \(B_0\).

Proof. (i) We have to show that the binary relation is reflexive and transitive. Reflexivity follows by choosing \(U = I\), which is a valid identity operator for any basis element in \(\mathcal{M}(\mathcal{H})\). Transitivity follows by considering three bases \(B_0, B_1, B_2\) and two operators \(\sigma, \tau\) belonging to different bases. The transitivity condition can be written as:

\[
B_0 \succ^{B_0} B_1 \quad \text{and} \quad B_1 \succ^{B_1} B_2 \implies B_0 \succ^{B_0} B_2.
\]

To prove this, we need to show that, given any two operators \(\sigma, \tau\) that satisfy the preorder conditions for bases \(B_0, B_1, B_2\), there exists a transformation \(U\) such that:

\[
U \sigma U^\dagger \succ^{B_2} \tau U^\dagger.
\]

This can be achieved by finding a unitary operator \(U\) that rotates \(\sigma\) in such a way that it becomes comparable to \(\tau\) in the basis \(B_2\). This process is possible due to the freedom of reordering and phase factors, which is a consequence of the preorder's transitive property.
Let \( (iii) \) together with

\[
\mathcal{D}_{B_3} \mathcal{D}_{B_0} = \mathcal{U} \left[ \prod_{\alpha} \mathcal{D}_B \right] \mathcal{D}_{B_1} \mathcal{D}_{B_0},
\]

for appropriate unitaries and dephasing channels. Substituting and using Eq. (10), it follows that

\[
\mathcal{D}_{B_3} \mathcal{D}_{B_0} = \mathcal{U} \left[ \prod_{\alpha} \mathcal{D}_{U(\alpha)} \right] \mathcal{D}_{U(\alpha)} \mathcal{D}_{B_1} \mathcal{D}_{B_0}
\]

hence \( B_1 \succ B_0 \) \( B_3 \).

(ii) Using the fact that dephasing channels are projectors, hence \( \mathcal{D}_B^2 = \mathcal{D}_B \), and by choosing a single \( \mathcal{D}_{B_1'} = \mathcal{D}_B \) together with \( \mathcal{U} = \mathcal{I} \).

(iii) Let \( B_0 = \{ P_i \}_{i=1}^d \) and \( B_{MU} = \{ \tilde{P}_i \}_{i=1}^d \), such that \( \text{Tr} (P_i \tilde{P}_j) = \frac{1}{d} \). Then,

\[
\mathcal{D}_{BMU} \mathcal{D}_{B_0} (X) = \sum_{i,j} \text{Tr} (P_i X) \text{Tr} (P_i \tilde{P}_j) \tilde{P}_j = \text{Tr}(X) \frac{I}{d}.
\]

Therefore, the equation

\[
\mathcal{D}_{BMU} \mathcal{D}_{B_0} = \mathcal{D}_{B_1'} \mathcal{D}_B \mathcal{D}_{B_0}
\]

holds for a single measurement \( \mathcal{D}_{B_1'} \) chosen over a mutually unbiased basis of \( B \).

**Proposition 2.** Let \( X_{ij}(B_\beta, B_\alpha) := \text{Tr} (P_i^{(\beta)} P_j^{(\alpha)}) \) denote the bistochastic matrix resulting from an ordering of the bases \( B_\alpha = \{ P_i^{(\alpha)} \}_{i=1}^d \) and \( B_\beta = \{ P_i^{(\beta)} \}_{i=1}^d \). Then, \( B_1 \succ B_0 \) \( B_2 \) if and only if there exists a sequence of unistochastic matrices \( \{ A^{(\alpha)} \}_\alpha \) such that

\[
X(B_2, B_0) = \left[ \prod_\alpha A^{(\alpha)} \right] X(B_1, B_0).
\]

**Proof.** Eq. (3) holds if and only if the action of the LHS and the RHS on any \( P_i^{(0)} \) coincide. This is because \( \mathcal{D}_{B_0} \) is a projector and hence the action is non-trivial only over the image \( \text{Im}(\mathcal{D}_{B_0}) = \text{Span} \{ P_i^{(0)} \}_i \). We have,

\[
\text{LHS: } \mathcal{D}_{B_2} \mathcal{D}_{B_0} P_i^{(0)} = \sum_j X_{ji}(B_2, B_0) P_j^{(2)}
\]

\[
\text{RHS: } \mathcal{U} \left[ \prod_\alpha \mathcal{D}_{B_\alpha} \right] \mathcal{D}_{B_1} \mathcal{D}_{B_0} P_i^{(0)} = \sum_{j,a} \left[ \prod_{a=1}^{\alpha_{\text{max}}-1} X_{j_{\alpha_{\text{max}}+1} a_{\alpha_{\text{max}}} (B'_{\alpha_{\text{max}}+1}, B'_{\alpha})} \right] X_{ji}(B_1, B_0) \mathcal{U} (P_j^{(\alpha_{\text{max}})})
\]

and hence a choice for \( \mathcal{U} \) is such that \( B_2 = \mathcal{U}(B'_{\alpha_{\text{max}}}) \).

Now, notice that all \( X \) matrices are unistochastic, since they can be written as \( X_{ij}(B_2, B_1) = |\langle j^{(1)} | U(B_2, B_1) | i^{(1)} \rangle|^2 \), where \( U(B_2, B_1) \) is a unitary matrix connection the two bases, i.e., such that \( U(B_2, B_1) P_i^{(1)} U^\dagger(B_2, B_1) P_i^{(2)} \forall i \). This proves the necessity of the form [4] for \( A^{(\alpha)} = X(B'_{\alpha_{\text{max}}+1}, B'_{\alpha}) \).

The sufficiency follows by the fact that for unistochastic matrix \( A^{(\alpha)} \) there exists a \( U^{(\alpha)} \) such that \( A^{(\alpha)}_{ij} = |\langle j^{(0)} | U^{(\alpha)} | i^{(0)} \rangle|^2 \). The series of dephasing bases in Eq. (3) is then obtained recursively by \( B'_1 = U^{(1)}(B_0) \) and \( B'_{\alpha+1} = U^{(\alpha+1)}(B'_\alpha) \), for \( 1 \leq \alpha \leq \alpha_{\text{max}} - 1 \).
Lemma 1. Every bistochastic matrix can be approximated arbitrarily well by a product of unistochastic matrices. 

Proof. Assume $M$ is a bistochastic matrix such that $M_{ij} > 0$ for all $i,j$. Then, $M$ can be expanded into a finite product of T-transforms, which are unistochastic matrices. This is because T-transforms act non-trivially only on a 2-dimensional subspace and all bistochastic matrices in $d = 2$ are unistochastic.

The set of bistochastic matrices forms a convex polytope and hence in any $\epsilon$-neighbourhood (as defined, e.g., by the $l_1$ norm) of a $M$ that fails the above condition, there exists some $M'$ that fulfills it. ■

Proposition 3. Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be a continuous convex function with $\phi(\frac{1}{d}(1, \ldots, 1)) = 0$. Then, 

$$f^\phi_{B_0}(B_1) := \sum_{i=1}^d \phi(X^R_i(B_1, B_0))$$

is a measure of relative compatibility; here, $X^R_i$ stand for the row vectors of the matrix $X_{ij}$.

Proof. Let $B_1 \succ B_0 B_2$. Then, there exists a bistochastic matrix $M$ (in fact, a product of unistochastic ones) such that $X(B_2, B_0) = MX(B_1, B_0)$. For any continuous convex function $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$\sum_{i=1}^{d} \phi(X^R_i(B_2, B_0)) = \sum_{i} \phi \left( \sum_{k} M_{ik}X^R_k(B_1, B_0) \right) \leq \sum_{i,k} M_{ik} \phi \left( X^R_k(B_1, B_0) \right) = \sum_{i} \phi \left( X^R_i(B_1, B_0) \right).$$

The bases $B_0$ and $B_1$ are mutually unbiased if and only if $X_{ij}(B_1, B_0) = \frac{1}{\sqrt{d}}$, from which the normalization condition follows. The additional condition for $\phi$ implies $f^\phi_{B_0}(B_1) = \sum_{i,j} \phi_0(X_{ij}(B_1, B_0))$ which assures symmetry since $X(B_0, B_1) = X^T(B_1, B_0)$. ■

Proposition 4. (i) Let $B_1 \succ B_0 B_2$ and $\rho_0 = \sum_i p_i \rho_i^{(0)}$. Then, $c_{B_1}^{(rel)}(\rho_0) \leq c_{B_2}^{(rel)}(\rho_0)$. (ii) Similarly, $c_{B_1}^{(2)}(\rho_0) \leq c_{B_2}^{(2)}(\rho_0)$. (iii) The average relative entropy of coherence $\langle c_{B_1}^{(rel)}[D_{B_0}(|\psi\rangle \langle \psi|)]|\psi\rangle$ over Haar distributed pure states $|\psi\rangle$ is the measure of relative incompatibility between $B_0$ and $B_1$. (iv) Similarly, for the average 2-coherence $\langle c_{B_1}^{(2)}[D_{B_0}(|\psi\rangle \langle \psi|)]|\psi\rangle$.

Proof. (i) We have

$$S(\rho_0 || D_{B_2}\rho_0) = -S(\rho_0) - Tr(\rho_0 \log[D_{B_1}(\rho_0)])$$

$$= -S(\rho_0) - Tr(D_{B_1}(\rho_0) \log[D_{B_1}(\rho_0)])$$

$$= S(D_{B_1}(\rho_0)) - S(\rho_0).$$

Since von Neumann entropy is a Schur-concave function, the assumption $B_1 \succ B_0 B_2$ implies $S(D_{B_1}(\rho_0)) \leq S(D_{B_2}(\rho_0))$ from which the claim follows.

(ii) In the following, we use the operator 2-norm $||X||_2 := Tr(X^2)$. We have,

$$c_{B_2}^{(2)}(\rho_0) = ||I - D_{B_2}\rho ||_2^2 = ||\rho_0||_2^2 - ||D_{B_2}\rho ||_2^2 = ||\rho_0||_2^2 - \left\| \sqrt{\prod_{\alpha} D_{B_\alpha}} D_{B_1}\rho \right\|_2^2$$

$$\geq ||\rho_0||_2^2 - ||D_{B_1}\rho ||_2^2 = c_{B_1}^{(2)}(\rho_0).$$

The inequality follows since the 2-norm is submultiplicative over unital CPTP maps.

(iii) We first notice that, by use of Eq. [10],

$$\langle c_{B_1}^{(rel)}[D_{B_0}(|\psi\rangle \langle \psi|)]|\psi\rangle = \langle c_{B_1}^{(rel)}[\sqrt{V}D_{B_1}|\psi\rangle \langle \psi|\sqrt{V}^T]\rangle|\psi\rangle = \langle c_{B_1}^{(rel)}[\sqrt{V}D_{B_1}|\psi\rangle \langle \psi|\sqrt{V}^T]\rangle|\psi\rangle,$$

where $V(B_1) = B_0$. The above mean was calculated in [53],

$$\langle c_{B_1}^{(rel)}[D_{B_0}(|\psi\rangle \langle \psi|)]|\psi\rangle = \frac{1}{d} \sum_i Q(X^R_i(B_1, B_0)),$$
where

\[ Q(p) := -\sum_{k=1}^{d} \left( \prod_{i\neq k} \frac{p_k}{p_i} \right) p_k \log p_k \]

(11)
is the subentropy of the probability vector \( p \). In [55], it was shown that the subentropy is a concave function. Hence, the claim follows by Proposition 3.

(iv) Like in (iii), we write

\[ \langle c^{(2)}_{B_1} [D_{B_0} (|\psi\rangle \langle \psi|)]_\psi \rangle = \langle c^{(2)}_{B_1} [VD_{B_1} (|\psi\rangle \langle \psi|)]_\psi \rangle. \]

The mean was calculated in the context of coherence in [56].

\[ \langle c^{(2)}_{B_1} [D_{B_0} (|\psi\rangle \langle \psi|)]_\psi \rangle = \frac{1}{d(d+1)} \sum_{ij} \left( \frac{1}{d} - X^2_{ij} (B_1, B_0) \right) \]
hence it is of the form (7) with \( \psi(x) = \frac{1}{d(d+1)} \left( \frac{1}{4} - x^2 \right) \), which is concave.

**Proposition.** \( Q_{B_0}(B_1) \leq q_{B_0}(B_1) \)

**Proof.** One has for \( A = \sum_k a_k P^{(1)}_k \),

\[ \text{Var}_i(A) = \text{Tr} \left( P^{(0)}_i A^2 - \left( \sum_k a_k P^{(1)}_k \right) \right)^2 \\
= \sum_k a_k^2 \text{Tr} \left( P^{(0)}_i P^{(1)}_k \right) - \sum_{k,l} a_k a_l \text{Tr} \left( P^{(0)}_i P^{(1)}_k \right) \text{Tr} \left( P^{(0)}_l P^{(1)}_k \right) \]

hence

\[ Q_{B_0}(B_1) \leq \sup_{A \in A_{B_1}, \|A\|_2 = 1} \sum_i \text{Var}_i(A) \\
= \sup_{A \in A_{B_1}, \|A\|_2 = 1} \left( 1 - \|X^T(B_1, B_0) a\|^2 \right) \\
\leq 1 - \lambda_{\min} \left( X(B_1, B_0) X^T(B_1, B_0) \right) \]

which is the desired bound.

**Proposition 5.** (i) If \( B_1 \succ B_0 \), then \( q_{B_0}(B_1) \leq q_{B_0}(B_2) \) and also \( q_{B_0}(B_1) = q_{B_0}(B_0) \), i.e., it is a symmetric measure of relative incompatibility. (ii) If \( B_2 \succ B_1 \) then \( r^{(MU)}(B_2, B_1) \leq r^{(MU)}(B_3, B_1) \) and \( r^{(MU)}(B_1, B_2) = r^{(MU)}(B_2, B_1) \), i.e., it is a symmetric measure of relative incompatibility.

**Proof.** (i) If \( B_1 \succ B_0 \), then by Proposition 2 there exists a bistochastic matrix \( M \) such that \( X(B_2, B_0) = MX(B_1, B_0) \). We need to show that this implies \( \lambda_{\min} \left( X(B_1, B_0) X^T(B_1, B_0) \right) = s_d^2 \left( X(B_1, B_0) \right) \) (the minimum singular value) satisfies \( s_d \left( X(B_1, B_0) \right) \geq s_d \left( X(B_2, B_0) \right) \). Indeed, this is guaranteed by the Gel’fand-Naimark inequality which states that (for the singular values sorted in decreasing order) \( \prod_{j=1}^k s_{ij}(A) \leq \prod_{j=1}^k s_j(A) \prod_{j=1}^k s_{ij}(B) \) for all \( 1 \leq i_1 \leq \ldots \leq i_k \leq n \) and \( k = 1, \ldots, n \) (in our case we set \( k = 1 \) and \( i_1 = n \)). Notice that \( s_1(M) = 1 \) since \( M \) is bistochastic.

(ii) Proposition 2 implies that \( \max_{i,j} X_{ij}(B_3, B_1) \leq \max_{i,j} X_{ij}(B_2, B_1) \), hence the result follows from the monotonicity of the log function. Symmetry follows from \( X(B_2, B_1) = X^T(B_2, B_1) \).

**Proposition.** \( F(B_0, B_1) := |\det \left( X(B_1, B_0) \right) | \) is a symmetric measure of relative compatibility.

**Proof.** From Proposition 2, \( B_1 \succ B_0 \) implies that \( X(B_2, B_0) = MX(B_1, B_0) \) for some bistochastic \( M \). Monotonicity follows since \( |\det(M)| \leq 1 \) and symmetry from the fact that \( X(B_2, B_1) = X^T(B_1, B_2) \).