IOP Publishing Quantum Sci. Technol. 11 (2026) 025049 https://doi.org/10.1088/2058-9565/ae5fcd

Quantum Science and Technology

Check for updates

OPEN ACCESS

RECEIVED
6 January 2026

REVISED
5 April 2026

ACCEPTED FOR PUBLICATION
15 April 2026

PUBLISHED
11 May 2026

PAPER

Quantum measurement retrodiction and entropic uncertainty relations*

Jiaxi Kuang[**], **Kensei Torii** and **Francesco Buscemi**[**]

Department of Mathematical Informatics, Nagoya University, Furo-cho Chikusa-ku, Nagoya 464-8601, Japan
[**] Authors to whom any correspondence should be addressed.

E-mail: kuang.jiaxi.g5@s.mail.nagoya-u.ac.jp, buscemi@nagoya-u.jp and toriikensei@nagoya-u.jp

Keywords: quantum information theory, quantum measurements, uncertainty relation, quantum Bayesian retrodiction

Abstract

We study quantum measurement retrodiction via the principle of minimum change. For general quantum measurements that are described by quantum-to-classical channels, we show that a broad family of standard quantum divergences selects the same retrodictive update, yielding a unique and divergence-independent quantum Bayesian inverse for any POVM and prior state. Using this update, we construct a symmetric joint distribution for pairs of POVMs and introduce the *mutual retrodictability*, which quantifies how well the two POVMs can retrodict each other's outcome distributions under this update. We also derive a general upper bound on this quantity, which depends only on the prior and holds for all measurements. This framework leads to two *retrodictive entropic uncertainty relations*, expressed directly in terms of the prior and the POVMs, yet valid independently of any retrodictive interpretation and fully compatible with conventional operational formulations. One of these relations links entropic uncertainty, approximate recoverability, and the thermodynamics of measurement through the Groenewold–Ozawa information gain. Finally, numerical benchmarks show that the resulting bounds are consistently tighter than existing entropic uncertainty relations across broad classes of measurements and states.

1. Introduction

Let us begin by considering a finite-state classical system, whose states are labeled by the elements of a finite set $\mathcal{S} = \{s\}$. Suppose that the actual, or 'true,' state of the system is unknown to us. Instead, our uncertainty about it is captured by a prior distribution $\gamma(s)$. The system then undergoes a measurement whose possible outcomes form another finite set $\mathcal{X} = \{x\}$, and whose statistics are governed by a known likelihood function $\varphi(x|s)$, the probability of observing outcome x given that the system is in state s.

When a particular outcome $\bar{x} \in \mathcal{X}$ is observed, Bayes' rule prescribes how our prior beliefs should be updated in light of this new information:

$$\hat{\gamma}(s|\bar{x}) := \frac{\gamma(s)\ \varphi(\bar{x}|s)}{\sum_s \gamma(s)\ \varphi(\bar{x}|s)}\ . \tag{1}$$

This simple expression lies at the heart of all probabilistic inference and numerous rigorous derivations have been proposed to justify its universal applicability and remarkable empirical success across the sciences [1–7].

Once the prior has been updated to incorporate the newly acquired information, namely the observed outcome $\bar{x}$, it can be used to compute the statistics associated with other observations. Consider, for instance, another measurement with outcomes in the set $\mathcal{Y} = \{y\}$ and likelihood function

* We dedicate this work to the memory of Michael J. W. Hall, whose comments, suggestions, and encouragement helped shape this work.

$\psi(y|s)$. The updated belief $\hat{\gamma}(s|\bar{x})$ allows us to compute the expected probability of each outcome y as

$$\Pr\{y|\bar{x}\} = \sum_{s\in\mathcal{S}} \hat{\gamma}(s|\bar{x})\,\psi(y|s)\,. \tag{2}$$

This quantity $\Pr\{y|\bar{x}\}$ represents the probability assigned to outcome y, conditioned on the information provided by the observation $\bar{x}$. At this stage, the role of the variable y is *inferential* rather than *temporal*: it labels the outcomes of another observation whose statistics are evaluated conditional on $\bar{x}$, without yet assigning that observation a specific position in time.

The above argument, a direct consequence of Bayes' rule and the total law of probability, admits a natural generalization to quantum theory. Consider a bipartite quantum system described by a state ρ_{AB}, and suppose we perform a local measurement on subsystem A described by a positive operator-valued measure (POVM) $\mathbf{M} := \{M_x : x \in \mathcal{X}\}$, i.e. a family of positive semidefinite operators $M_x \geqslant 0$ such that $\sum_{x\in\mathcal{X}} M_x = \mathbb{1}$. Upon obtaining a specific outcome $\bar{x} \in \mathcal{X}$, the state of subsystem B is updated, according to Born's rule, to

$$\rho_{B|\bar{x}} := \frac{\mathrm{Tr}_A\{(M_{\bar{x}} \otimes \mathbb{1}_B)\;\rho_{AB}\}}{\mathrm{Tr}\{(M_{\bar{x}} \otimes \mathbb{1}_B)\;\rho_{AB}\}}.$$

This updated state encodes our new expectations about any subsequent measurement that might be performed on B.

In the classical case, the quantity $\hat{\gamma}(s|\bar{x})$ also admits a *retrodictive interpretation*, namely as an inference about the system's state *prior to the measurement*. Retrodictive reasoning underlies the logic of diagnostic sciences, from medicine to forensics. Though the explicit term 'retrodiction' emerged only in the past century, its conceptual roots trace back to Laplace, who described it as 'the probability of causes given events' [8].

In quantum mechanics, however, this *backward-in-time interpretation* encounters deep conceptual subtleties: measurement outcomes do not, in general, reveal preexisting properties of the system. Consequently, the retrodictive use of Bayes' rule in quantum theory has long been a fertile ground for discussion and insight; for a recent overview of these developments, see [9, 10]. A possible source of disagreement in the literature may lie in the variety of approaches that have been adopted. To the best of our knowledge, the quantum versions of classical Bayesian retrodiction proposed so far typically fall into one of the following categories: they are either *semiclassical*, where the underlying process is quantum but the retrodiction concerns only the classical statistics of the outcomes (such as in [11, 12]); or they are defined *by analogy*, in the sense that the resulting expressions reduce to classical retrodiction when operators commute, yet may yield non-positive quasi-probability distributions in the general case (such as in [13]); or they are derived from a set of *axioms* or desiderata whose appeal and physical justification can vary [14].

Only very recently, a fully quantum generalization of Bayes' rule has been derived from the *minimum change principle* [15]. This principle postulates that the update of one's belief should proceed through a reverse (i.e. retrodictive) process that remains consistent with the newly acquired information while, at the same time, minimizing the statistical divergence from the corresponding forward (i.e. predictive) process [5]. In essence, it embodies a conservative epistemic stance, ensuring that the belief revision faithfully reflects the new data without introducing unintended or spurious biases. In this respect, it seems to be related (at least 'morally') with Jaynes' Maximum Entropy Principle [7, 16, 17], although see [18, 19].

In this paper, we focus on the task of retrodiction based on the outcome of a quantum measurement. We show that, by restricting the analysis to quantum-to-classical (measurement) channels, the quantum minimum change principle of [15] naturally extends to a broad family of statistical divergences. All such divergences yield the same optimal retrodicted quantum state, revealing the universality of the retrodictive update rule within this semiclassical setting. This universality shows that the resulting update rule is not an artifact of a particular divergence measure, but rather a robust feature of the underlying variational structure of quantum inference. Building upon this result, we employ the retrodicted state to formulate a new class of *retrodictive entropic uncertainty relations*, which capture the information-theoretic trade-offs inherent in backward-in-time quantum reasoning. In addition, we construct two new entropic uncertainty relations. In particular, theorem 3.7 shows that one of our entropic uncertainty relations admits an extension that is valid for arbitrary states and POVMs and that can be written directly in terms of the Groenewold–Ozawa information gain, thereby providing an explicit bridge to thermodynamic analyses of quantum measurement processes [20–23]. Importantly, although derived through a retrodictive construction, these relations are valid independently of that narrative and admit the conventional operational interpretation of entropic uncertainty relations. Finally, we benchmark these relations against the well-known entropic uncertainty bounds of [24], showing that our approach yields tighter bounds in a wide range of scenarios.

The paper is organized as follows. In section 2.2 we recall the minimum change principle and specialize it to quantum measurement channels, showing that this restriction allows us to extend the principle from fidelity to a broad class of statistical divergences and that all such divergences select the same minimum–change retrodictive update. In section 2.3 we examine the resulting retrodicted states and develop the symmetric retrodictive joint probability that underpins our later constructions. In section 3.1 we introduce the mutual retrodictability $R(\mathbf{M};\mathbf{N})_\gamma$ and establish its general upper bound in theorem 3.2. In section 3.2 we use this framework to derive the retrodictive entropic uncertainty relations (13) and (22). In section 3.3 we benchmark our bounds numerically against the entropic uncertainty relation of Berta *et al* [24]. Finally, in section 4 we summarize our results and outline directions for future work.

2. Framework of retrodiction from quantum measurements

2.1. Preliminaries and problem setup

For completeness, we briefly recall how the minimum change principle provides a general foundation for belief updating in a classical information-theoretic setting. We view $\varphi(x|s)$ as a classical channel from S to X, and $\gamma(s)$ as the input distribution. The joint law

$$P_{\text{fwd}}(s,x) := \varphi(x|s)\,\gamma(s)\,,$$

therefore represents the complete operational process describing the information flow from system states to observable outcomes. Here $\gamma(s)$ encodes the prior belief about S, while the likelihood $\varphi(x|s)$ specifies the channel correlation between S and X.

When new information about X is acquired, it typically takes the form of a constraint on the marginal of X. We denote this target distribution by $\tau(x)$. It may represent either a fully specified measurement result, as in the textbook Bayes update where $\tau(x)=\delta(x,\bar{x})$, or a noisy or coarse-grained form of evidence [4]. The updated belief should incorporate this constraint while deviating from the prior operational process as little as possible.

The requirement of minimum information change can be cast as an information projection problem:

$$\min_R \mathbb{D}(R\|P_{\text{fwd}})\,, \tag{3}$$

where $\mathbb{D}$ is a statistical divergence. The variable R ranges over all joint distributions consistent with the acquired information, i.e. of the form $R(s,x)=\tau(x)\hat{\varphi}(s|x)$ for some normalized conditional distribution $\hat{\varphi}(s|x)$. Thus the optimization in (3) is over the posterior channel $\hat{\varphi}(s|x)$ only. In the language of information geometry, this corresponds to the projection of P_{fwd} onto the affine set $\{R_{SX}:R_X=\tau\}$, being R_X the marginal of R_{SX}.

A key feature of the minimum change principle is that it applies to the full joint input–output process, rather than only to its marginals. In accordance with Bayesian practice, we assume that the prior process assigns nonzero probability to every trajectory, i.e. $P_{\text{fwd}}(s,x)>0$ for all $s\in\mathcal{S}$ and $x\in\mathcal{X}$. Under this regularity assumption, and for a broad class of divergences (including all strictly convex f-divergences), the minimizer is unique and given by

$$P_{\text{rev}} := \arg\min_R \mathbb{D}(R\|P_{\text{fwd}}) = \frac{P_{\text{fwd}}(s,x)}{\sum_{s'}P_{\text{fwd}}(s',x)}\tau(x)\,, \tag{4}$$

thus recovering the classical Bayesian update.

In this paper we focus on a generalization of the minimum change principle to the special case of quantum measurements. We consider a finite-dimensional quantum system Q with Hilbert space $\mathcal{H}_Q$, subjected to a measurement represented by a POVM $\mathbf{M}=\{M_x:x\in\mathcal{X}\}$, where $\mathcal{X}$ is a finite set of outcomes and M_x are positive semidefinite operators satisfying the completeness relation $\sum_{x\in\mathcal{X}}M_x=\mathbb{1}_Q$. We say that the POVM $\mathbf{M}$ is *rank-one* whenever all its non-zero elements are rank-one operators. If the state of the system is represented by a density operator γ_Q, the probability of obtaining the outcome x is given by $p(x) = \mathrm{Tr}\{M_x\gamma_Q\}$. It is convenient to summarize the resulting outcome distribution by introducing a quantum-to-classical channel, namely a completely positive trace-preserving map

$$\mathcal{M}(\gamma_Q) := \sum_{x\in\mathcal{X}} \mathrm{Tr}\{M_x\gamma_Q\}\,|x\rangle\langle x|_X. \tag{5}$$

Here the kets $\{|x\rangle\}_x$ form a fixed orthonormal basis of a space $\mathcal{H}_X \cong \mathbb{C}^{|\mathcal{X}|}$. This is a purely formal device: we will never consider superpositions of such vectors. It provides a compact way to treat the classical random variable X representing the measurement outcome within the quantum formalism.

Finally, we introduce three basic information-theoretic quantities. For a classical random variable X with probability distribution $\Pr\{x\}$, the Shannon entropy is defined as

$$H(X) := -\sum_x \Pr\{x\}\log_2 \Pr\{x\}.$$

Given two classical random variables X and Y with joint distribution $\Pr\{x,y\}$, the classical mutual information is defined as

$$I(X;Y) := H(X) + H(Y) - H(XY),$$

where $H(Y)$ and $H(XY)$ denote the Shannon entropies of the marginal distribution $\Pr\{y\}$ and the joint distribution $\Pr\{x,y\}$, respectively. For the quantum state γ_Q, the von Neumann entropy is defined as

$$H(\gamma_Q) := -\mathrm{Tr}\{\gamma_Q \log_2 \gamma_Q\}.$$

2.2. Minimum change principle for quantum measurements

In [15], a framework was developed for studying the minimum change principle for general quantum channels. In this work, we restrict attention to quantum-to-classical (qc) channels of the form (5). Their semiclassical structure makes them considerably easier to analyze, and allows us to extend the minimum change principle to a broad family of statistical divergences, unlike [15] which focused on fidelity.

Considering only qc channels also provides a natural analogue of the joint input–output distributions used in the classical setting. In fully noncommutative scenarios, the absence of a straightforward quantum analogue of an input-output joint state complicates the formulation [15]. For measurement channels, however, we may introduce the bipartite state

$$\gamma_{XQ} := \sum_{x\in\mathcal{X}} |x\rangle\langle x|_X \otimes \sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}. \tag{6}$$

Up to a partial transposition, which is irrelevant here because γ_{XQ} is separable, this forward-process bipartite state coincides with the definition in [15]. That construction, in turn, is based on the Choi isomorphism and the standard purification, both of which are commonly regarded as *canonical* constructions [25–29]. Although the derivation is straightforward, for the sake of completeness, we report the explicit calculation leading to (6) in appendix A. It is immediate to verify that $\mathrm{Tr}_Q\{\gamma_{XQ}\} = \sum_x p(x)|x\rangle\langle x|_X$ and $\mathrm{Tr}_X\{\gamma_{XQ}\} = \gamma_Q$, so the marginals recover the expected output distribution and the input state, respectively.

The minimum change principle for quantum measurement channels can now be formulated as an information projection

$$\min_{\sigma_{XQ}} \mathbb{D}(\sigma_{XQ}\|\gamma_{XQ}), \tag{7}$$

between the forward bipartite state γ_{XQ}, which encodes the prior process for the measurement of Q, and a backward bipartite state σ_{XQ} satisfying the following conditions:

1. its marginal state satisfies $\sigma_X := \mathrm{Tr}_Q\{\sigma_{XQ}\} = \sum_{x\in\mathcal{X}} \tau(x)|x\rangle\langle x|_X$, representing the newly acquired information about the outcome distribution—this condition is imposed as a constraint in the minimization;
2. the backward channel, the analogue of $\hat{\varphi}(s|x)$ in (3), is now a *classical-to-quantum* (cq) channel that associates to each $x\in\mathcal{X}$ a quantum state $\sigma_{Q|x}$ on $\mathcal{H}_Q$.

Consequently, the bipartite state σ_{XQ} takes the form

$$\sigma_{XQ} = \sum_{x\in\mathcal{X}} \tau(x)\,|x\rangle\langle x|_X \otimes \sigma_{Q|x}.$$

With the above notations and assumptions, we have the following:

Theorem 2.1. *Assume* $\mathrm{Tr}\{\gamma_Q M_x\} > 0$ *for all* $x \in \mathcal{X}$*, and let an arbitrary probability distribution* $\tau(x)$ *represent the newly acquired information. For the following quantum divergences (below, A and B denote density matrices, and we assume* $B > 0$*):*

1. *Umegaki relative entropy [30]:* $\mathbb{D}(A\|B) \equiv D(A\|B) := \mathrm{Tr}\{A(\log_2 A - \log_2 B)\}$*;*
2. *Petz–Rényi relative entropy [31, 32]:* $D_\alpha(A\|B) := \frac{1}{\alpha-1}\log_2 \mathrm{Tr}\{A^\alpha B^{1-\alpha}\}$*, for* $\alpha \in (0,1)\cup(1,2]$*;*
3. *sandwiched Rényi relative entropy [33, 34]:* $\tilde{D}_\alpha(A\|B) := \frac{1}{\alpha-1}\log_2 \mathrm{Tr}\left\{\left(B^{\frac{1-\alpha}{2\alpha}} A B^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right\}$*, for* $\alpha \in [1/2,1)\cup(1,\infty)$*;*
4. *geometric Rényi relative entropy [35, 36]:* $\hat{D}_\alpha(A\|B) := \frac{1}{\alpha-1}\log_2 \mathrm{Tr}\{B\left(B^{-1/2}AB^{-1/2}\right)^\alpha\}$*, for* $\alpha \in (0,1)\cup(1,2]$*;*
5. *Belavkin–Staszewski relative entropy [37]:* $D_{BS}(A\|B) := \mathrm{Tr}\{A\log_2\left(A^{1/2}B^{-1}A^{1/2}\right)\}$*;*

we have

$$\min_{\sigma_{XQ}} \mathbb{D}(\sigma_{XQ}\|\gamma_{XQ}) = \mathbb{D}(\sigma_X\|\gamma_X) \equiv \mathbb{D}(\tau\|p),$$

where $\tau \equiv \tau(x)$ *and* $p \equiv p(x) = \mathrm{Tr}\{\gamma_Q M_x\}$*, and the minimum is uniquely achieved by*

$$\hat{\gamma}_{XQ} := \arg\min_{\sigma_{XQ}} \mathbb{D}(\sigma_{XQ}\|\gamma_{XQ}) = \sum_{x\in\mathcal{X}} \tau(x)\,|x\rangle\langle x|_X \otimes \frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{M_x\gamma_Q\}}. \tag{8}$$

Besides relative entropies, the same conclusion holds also when the statistical divergence between forward and reverse processes is measured by Uhlmann fidelity (this follows directly from [15]) and trace-distance:

Theorem 2.2. *With the same conventions as above, the minimum*

$$\min_{\sigma_{XQ}} \frac{1}{2}\|\sigma_{XQ} - \gamma_{XQ}\|_1 = \frac{1}{2}\sum_{x\in\mathcal{X}} |\tau(x) - p(x)|,$$

is uniquely achieved by $\hat{\gamma}_{XQ}$ *in (8).*

The proofs of theorems 2.1 and 2.2 can be found in appendix B.

Theorems 2.1 and 2.2 reinforce the argument first proposed in [15], which motivates the adoption of the conditional states

$$\hat{\gamma}_{Q|\bar{x}} := \frac{\sqrt{\gamma_Q}M_{\bar{x}}\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_{\bar{x}}\}}, \tag{9}$$

as *the* quantum analogue of the classical Bayesian inverse in equation (1). This identification had been anticipated in earlier work, where expressions equivalent to $\hat{\gamma}_{Q|\bar{x}}$ were motivated primarily by formal analogy with classical Bayes' rule rather than derived from a variational principle; see, e.g. [14, 38–46]. In light of theorems 2.1 and 2.2, we adopt $\hat{\gamma}_{Q|\bar{x}}$ as the canonical update singled out by the minimum-change principle, and we refer to it as the *quantum Bayesian inverse* associated with outcome $\bar{x}$. This update has already proven useful beyond the present setting: [47] showed that it leads naturally to a quantum analogue of classical retrodictive smoothing, and [48] further developed a generalized smoothing framework that treats smoothing as intrinsically retrodictive and is built explicitly on $\hat{\gamma}_{Q|\bar{x}}$. Note also that, within this semiclassical measurement framework, the construction reduces to the classical minimum-change principle in the commuting limit. In what follows, we use the Bayesian inverse $\hat{\gamma}_{Q|\bar{x}}$ to quantify how a measurement outcome can be employed to retrodict information about the system's pre-measurement state.

Before concluding this section, we note that the states defined in (9) coincide with the outputs of the *Petz transpose map* [49, 50] associated with the measurement channel $\mathcal{M}$ in (5) and the prior state

γ_Q, when this map is applied to the classical output states $|x\rangle\langle x|$. For a detailed discussion of this construction, its mathematical and statistical properties, and its physical interpretation, we refer the reader to [51].

2.3. Quantum Bayesian retrodiction from a measurement's outcomes

We start with a few words of warning. The use of retrodictive arguments in quantum theory is conceptually delicate. We are aware that interpreting quantum measurements as providing information about prior states, in particular the interpretation of the retrodicted states $\hat{\gamma}_{Q|x}$, is problematic both operationally and philosophically. Nevertheless, we take a pragmatic view: we adopt the principle of minimum change as a working hypothesis and examine where it leads. Our goal is not to justify retrodictive reasoning but to study its mathematical structure and to identify the regimes where it becomes consistent.

We repeat the setup. A quantum system Q has a state described by a density operator γ_Q, which, in accordance to the Bayesian narrative, we interpret as the prior state. The system undergoes a measurement represented by a POVM $\mathbf{M} = \{M_x : x \in \mathcal{X}\}$, yielding an outcome $\bar{x}$. We now use the prior state together with the new information provided by the observed outcome $\bar{x}$ to compute the corresponding retrodicted state and, from it, the probabilities associated with the outcomes of *another measurement*. Here it is important to distinguish inferential order from physical time order. Inferentially, the logic is fixed: one starts from the prior state, conditions on the observation of $\bar{x}$, forms the retrodicted state, and then evaluates the statistics of another POVM on that retrodicted state. Physically, however, this second measurement need not be actually performed, nor is it assigned, in the general discussion, a fixed temporal position relative to the observation of $\bar{x}$, although in section 3.2 we show that the same joint statistics can be realized through a suitable sequential construction. In this sense, the role of the 'other measurement' is to test how much information the outcome $\bar{x}$ conveys about the system, under the retrodictive update, rather than to represent a second measurement event placed at a definite later time.

If the other measurement is represented by a POVM $\mathbf{N} := \{N_y : y \in \mathcal{Y}\}$, the analogue of equation (2) takes the form

$$\begin{aligned}\Pr\{y|\bar{x}\} &= \mathrm{Tr}\left\{N_y \hat{\gamma}_{Q|\bar{x}}\right\} \\ &= \frac{1}{\mathrm{Tr}\{M_{\bar{x}}\gamma_Q\}} \mathrm{Tr}\left\{N_y \sqrt{\gamma_Q} M_{\bar{x}} \sqrt{\gamma_Q}\right\}.\end{aligned}$$

Representing this as a diagram:

$$\gamma_Q \longrightarrow \text{observe } \bar{x} \longrightarrow \hat{\gamma}_{Q|\bar{x}} \longrightarrow \text{compute } \Pr\{y|\bar{x}\} = \mathrm{Tr}\left\{N_y\, \hat{\gamma}_{Q|\bar{x}}\right\}.$$

Accordingly, we define the *joint* retrodictive probability of outcomes y of $\mathbf{N}$ and x of $\mathbf{M}$ as

$$\Pr\{y \leftarrow x\} := \mathrm{Tr}\left\{N_y \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}\right\}. \tag{10}$$

It is immediate to verify that marginalizing the joint probability $\Pr\{y \leftarrow x\}$ recovers the standard Born rule for both POVMs:

$$\begin{aligned}\sum_{y\in\mathcal{Y}} \Pr\{y \leftarrow x\} &= \mathrm{Tr}\{M_x \gamma_Q\}, \\ \sum_{x\in\mathcal{X}} \Pr\{y \leftarrow x\} &= \mathrm{Tr}\{N_y \gamma_Q\}.\end{aligned} \tag{11}$$

The retrodictive joint probability $\Pr\{y \leftarrow x\}$ thus defines a valid joint distribution for any state γ_Q and any pair of POVMs $\mathbf{M}$ and $\mathbf{N}$, even when these are not jointly measurable [52]. This is possible because, in general, $\Pr\{y \leftarrow x\}$ is not a linear functional of the state γ_Q, although it remains linear in the POVMs.

We also observe that, as a consequence of the cyclic property of the trace, the retrodictive joint probability $\Pr\{y \leftarrow x\}$ is *symmetric* with respect to the two POVMs, so that, in fact

$$\Pr\{y \leftarrow x\} = \Pr\{x \leftarrow y\}.$$

In other words, performing the measurement $\mathbf{M}$ first and using its outcome to retrodict the statistics of $\mathbf{N}$ yields the same joint distribution as performing $\mathbf{N}$ first and using its outcome to retrodict those of $\mathbf{M}$. This symmetry contrasts sharply with the forward-in-time, or predictive, case, where the joint statistics depend on the specific order of the measurements. In the predictive setting, the measurement of a POVM $\mathbf{M} = \{M_x : x \in \mathcal{X}\}$ must be described by a *quantum instrument* $\{\mathcal{J}_x^{\mathbf{M}} : x \in \mathcal{X}\}$, i.e. a collection of completely positive trace-nonincreasing maps satisfying $\mathrm{Tr}\{\mathcal{J}_x^{\mathbf{M}}(\rho)\} = \mathrm{Tr}\{M_x\rho\}$ for all density operators ρ. The joint probability for two sequential measurements associated with the POVMs $\mathbf{M}$ and $\mathbf{N}$, performed in this order, is then given by

$$\Pr\{x \to y\} = \mathrm{Tr}\left\{N_y \mathcal{J}_x^{\mathbf{M}}(\gamma_Q)\right\},$$

which in general depends on the particular instrument chosen for the measurement performed first.

3. Results

3.1. Mutual retrodictability

With a joint probability at hand, it is natural to define a *mutual retrodictability* as follows:

Definition 3.1. Given a state γ_Q and two finite POVMs $\mathbf{M} = \{M_x : x \in \mathcal{X}\}$ and $\mathbf{N} = \{N_y : y \in \mathcal{Y}\}$ on Q, the *mutual retrodictability* of $\mathbf{M}$ and $\mathbf{N}$ with respect to γ_Q is defined as

$$R(\mathbf{M};\mathbf{N})_\gamma := I(X;Y),$$

where the right-hand side is the mutual information computed for the joint probability distribution $\Pr\{y \leftarrow x\}$ given in (10).

The term 'mutual retrodictability' highlights that the relationship between two measurements, $\mathbf{M}$ and $\mathbf{N}$, is not merely statistical but inferential. Through the principle of minimum change, the outcome of one measurement provides the most conservative retrodictive estimate of what the other measurement's outcome statistics must have been, had it probed the same underlying preparation. Accordingly, the mutual information of the retrodictive joint distribution $\Pr\{y \leftarrow x\}$ quantifies how strongly these backward inferences are correlated, measuring the degree to which the two POVMs can infer each other's outcome distributions under the minimum-change retrodictive update.

Theorem 3.2. *For any state γ_Q and any pair of POVMs $\mathbf{M} = \{M_x : x \in \mathcal{X}\}$ and $\mathbf{N} = \{N_y : y \in \mathcal{Y}\}$ on Q, the mutual retrodictability satisfies*

$$0 \leqslant R(\mathbf{M};\mathbf{N})_\gamma \leqslant H(\gamma_Q). \tag{12}$$

The proof is reported in appendix C.

Remark 3.3. The theorem shows that the mutual retrodictability of any two measurements is upper-bounded by a term that depends only on the 'purity' of the prior state. In particular, if the prior state γ_Q is pure, then $H(\gamma_Q) = 0$ and the mutual retrodictability vanishes identically for all pairs of POVMs. In this case, the system is already in a state of complete certainty, leaving no room for non-trivial retrodictive inference. A pure prior state thus acts as an *inferential firewall*: it blocks any attempt to draw further retrodictive conclusions, just as, in the classical setting, no amount of new information can alter a delta-distributed prior.

Remark 3.4. The bounds in (12) are tight. It is straightforward to construct scenarios where $R(\mathbf{M};\mathbf{N})_\gamma = 0$ or $R(\mathbf{M};\mathbf{N})_\gamma = H(\gamma_Q)$. The former occurs, as noted above, when, e.g. γ_Q is a pure state; in this case, the upper bound also vanishes, so the lower and upper bounds coincide. A slightly less trivial example arises when, for instance, $\mathbf{M}$ is a projective measurement on the eigenstates of γ_Q, while $\mathbf{N}$ is mutually unbiased. Conversely, a situation where $R(\mathbf{M};\mathbf{N})_\gamma = H(\gamma_Q) > 0$ may occur when $\mathbf{M} = \mathbf{N}$, and both are projective measurements on the eigenstates of γ_Q.

These observations suggest the possibility that mutual retrodictability could serve as a *state-dependent indicator of compatibility* between two POVMs. When $R(\mathbf{M};\mathbf{N})_\gamma$ is large, the measurements exhibit significant inferential overlap with respect to γ_Q; when it is small, they appear to probe largely independent aspects of the state. We leave this point open for future investigation.

These examples also suggest that large mutual retrodictability corresponds to regimes in which the overall triplet, consisting of the state γ_Q together with the two POVMs $\mathbf{M}$ and $\mathbf{N}$, behaves in a quasi-classical manner. In such cases, the measurement statistics admit an interpretation in terms of classical uncertainty about an underlying ontic configuration, rather than genuine quantum indeterminacy. Put differently, when $R(\mathbf{M};\mathbf{N})_\gamma$ approaches its upper bound, the retrodictive reconstruction of measurement outcomes becomes nearly indistinguishable from standard Bayesian inference over classical variables. The notion of 'retrodiction' thus regains its uncontroversial classical meaning: it no longer reflects an inference about noncommuting observables, but merely the update of classical ignorance concerning a common underlying cause consistent with both measurements.

3.2. Retrodictive entropic uncertainty relations

Before proceeding, we emphasize that the results in this section do not rely on any interpretive use of retrodiction. The quantity $\Pr\{y \leftarrow x\}$ introduced above can be regarded simply as a symmetric joint probability constructed from two POVMs and a state, without invoking any notion of temporal order or inference about past events. The uncertainty relations that follow are purely mathematical consequences of this symmetric construction. They hold regardless of the interpretation given to $\Pr\{y \leftarrow x\}$.

As shown in equation (11), the marginals of the symmetric joint probability $\Pr\{y \leftarrow x\}$ coincide with the standard outcome probabilities of both POVMs. This allows us to identify $H(X)$ and $H(Y)$, computed from $\Pr\{y \leftarrow x\}$, with $H(\mathbf{M})_\gamma$ and $H(\mathbf{N})_\gamma$, respectively. Since the mutual retrodictability is nonnegative, we immediately obtain the lower bound

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant H(XY),$$

where $H(XY)$ is the entropy computed from $\Pr\{y \leftarrow x\}$. This yields our first retrodictive entropic uncertainty relation:

Theorem 3.5. *For any state γ_Q and any pair of POVMs $\mathbf{M} = \{M_x : x \in \mathcal{X}\}$ and $\mathbf{N} = \{N_y : y \in \mathcal{Y}\}$ on Q, the following bound holds:*

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant -\max_{x,y} \log_2 \|\sqrt{N_y}\sqrt{\gamma_Q}\sqrt{M_x}\|_2^2. \tag{13}$$

Proof. The proof is immediate: the joint probability (10) is equivalent to $\|\sqrt{N_y}\sqrt{\gamma_Q}\sqrt{M_x}\|_2^2$, then from the inequality $H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant H(XY)$ we have:

$$\begin{aligned} H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant H(XY) &= -\sum_{x,y} \Pr\{y \leftarrow x\} \log_2 \|\sqrt{N_y}\sqrt{\gamma_Q}\sqrt{M_x}\|_2^2 \\ &\geqslant -\max_{x,y} \log_2 \|\sqrt{N_y}\sqrt{\gamma_Q}\sqrt{M_x}\|_2^2. \end{aligned}$$

□

A second retrodictive entropic uncertainty relation follows from the observation that, while in section 2.3 we noted that the predictive joint probability distribution $\Pr\{x \to y\}$ generally depends on the choice of instrument for the first measurement and is therefore not symmetric, there exists a *particular* instrument that reproduces the retrodictive joint statistics $\Pr\{y \leftarrow x\}$. This instrument can be constructed by starting from the square-root instrument and applying a unitary operator, arising from the polar decomposition, that exchanges the two square roots, namely $\sqrt{M_x}\sqrt{\gamma_Q} \to \sqrt{\gamma_Q}\sqrt{M_x}$. Explicitly, we define

$$\begin{aligned} \mathcal{J}_x^{\mathbf{M}}(\gamma_Q) &:= U_{\gamma,x}\sqrt{M_x}\gamma_Q\sqrt{M_x}U_{\gamma,x}^\dagger \\ &= \sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}, \end{aligned} \tag{14}$$

for all $x \in \mathcal{X}$. The unitary operators $U_{\gamma,x}$ depend on the state γ_Q. Consequently, this construction works for a fixed state, but the map $\gamma_Q \mapsto \sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}$ is not linear in γ_Q, and no single instrument can reproduce it for all quantum states.

Concatenating this instrument with the square-root instrument for $\mathbf{N}$ and adding two classical registers for the measurement outcomes, we obtain a quantum channel with hybrid quantum–classical output:

$$\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}(\cdot) := \sum_{y,x} \sqrt{N_y}U_{\gamma,x}\sqrt{M_x}(\cdot)\sqrt{M_x}U_{\gamma,x}^\dagger\sqrt{N_y} \otimes |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y, \tag{15}$$

which, by construction, when applied to γ_Q, satisfies

$$\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}(\gamma_Q) = \sum_{y,x} \sqrt{N_y}\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}\sqrt{N_y} \otimes |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y. \tag{16}$$

Since the joint probability distribution we aim to reproduce is symmetric under the exchange $\mathbf{M} \leftrightarrow \mathbf{N}$, we can repeat the same reasoning with the roles of the two POVMs reversed, taking $\mathbf{N}$ as the first measurement instead of $\mathbf{M}$. This yields

$$\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}(\cdot) := \sum_{y,x} \sqrt{M_x} V_{\gamma,y} \sqrt{N_y} (\cdot) \sqrt{N_y} V^\dagger_{\gamma,y} \sqrt{M_x} \otimes |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y, \tag{17}$$

and

$$\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}(\gamma_Q) = \sum_{y,x} \sqrt{M_x} \sqrt{\gamma_Q} N_y \sqrt{\gamma_Q} \sqrt{M_x} \otimes |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y. \tag{18}$$

The following argument will utilize the definition of the Groenewold–Ozawa information gain that is recalled here:

Definition 3.6 [Groenewold–Ozawa information gain [53, 54]]. Let γ_Q be a state on Q and

$$\mathcal{I}(\cdot) = \sum_z \mathcal{I}_z(\cdot) \otimes |z\rangle\langle z|_Z$$

an instrument with classical outcomes $z \in \mathcal{Z}$. The *Groenewold–Ozawa information gain* of the instrument $\mathcal{I}$ on the input state γ_Q is

$$I_{GO}(\gamma_Q; \mathcal{I}) := H(\gamma_Q) - \sum_{z:\Pr(z)>0} \Pr(z)\, H\left(\frac{\mathcal{I}_z(\gamma_Q)}{\Pr(z)}\right), \tag{19}$$

where $\Pr(z)$ is the probability of getting the outcome z and $H(\cdot)$ denotes the von Neumann entropy.

We can now state the following alternative retrodictive entropic uncertainty relation:

Theorem 3.7. *Let γ_Q be a state on Q and let $\mathbf{M} = \{M_x : x \in \mathcal{X}\}$ and $\mathbf{N} = \{N_y : y \in \mathcal{Y}\}$ be POVMs on Q. Let $\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}$ and $\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}$ be the instruments defined in equations (15) and (17) above, and define*

$$\tilde{\gamma}_Q := \left[\left(\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right)^\dagger \circ \tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right](\gamma_Q) = \sum_{x,y} \sqrt{M_x}\, U^\dagger_{\gamma,x} N_y \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q} N_y U_{\gamma,x} \sqrt{M_x}, \tag{20}$$

and

$$\tilde{\eta}_Q := \left[\left(\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}\right)^\dagger \circ \tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}\right](\gamma_Q) = \sum_{x,y} \sqrt{N_y}\, V^\dagger_{\gamma,y} M_x \sqrt{\gamma_Q} N_y \sqrt{\gamma_Q} M_x V_{\gamma,y} \sqrt{N_y}. \tag{21}$$

Note that, in general, $\mathrm{Tr}\{\tilde{\gamma}_Q\} \leqslant 1$ and $\mathrm{Tr}\{\tilde{\eta}_Q\} \leqslant 1$. Then the following retrodictive entropic uncertainty relation holds:

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant \max\left\{D\left(\gamma_Q \| \tilde{\gamma}_Q\right) + I_{GO}\left(\gamma_Q; \tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right),\ D\left(\gamma_Q \| \tilde{\eta}_Q\right) + I_{GO}\left(\gamma_Q; \tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}\right)\right\}, \tag{22}$$

where $D(\cdot\|\cdot)$ denotes the Umegaki quantum relative entropy.

We report the proof in appendix D.

Remark 3.8. Two additional observations are worth noting. First, since the channels $\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}(\cdot)$ and $\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}(\cdot)$ in equations (15) and (17) are both *subunital*, i.e. $\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}(\mathbb{1}_Q) \leqslant \mathbb{1}_{QXY}$ and $\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}(\mathbb{1}_Q) \leqslant \mathbb{1}_{QXY}$, both operators $\tilde{\gamma}_Q$ and $\tilde{\eta}_Q$ are subnormalized, i.e. $\mathrm{Tr}\{\tilde{\gamma}_Q\} \leqslant 1$ and $\mathrm{Tr}\{\tilde{\eta}_Q\} \leqslant 1$. Then, by Klein's inequality, this immediately implies that

$$D(\gamma_Q \| \tilde{\gamma}_Q) \geqslant 0 \qquad \text{and} \qquad D(\gamma_Q \| \tilde{\eta}_Q) \geqslant 0,$$

so the relative entropy terms appearing in equation (22) are always non-negative.

Second, the instruments $\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}$ and $\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}$ defined in equations (15) and (17) are *efficient* [55], in the sense that each classical outcome (x, y) is associated with a single Kraus operator. For efficient instruments, the Groenewold–Ozawa information gain is known to be non-negative and, moreover, to coincide with Winter's *intrinsic information* of the corresponding effective POVMs [20, 56, 57]. In the present setting, these POVMs have elements

$$J_{xy} := \sqrt{M_x} U^\dagger_{\gamma,x} N_y U_{\gamma,x} \sqrt{M_x}\,, \qquad \text{and} \qquad J'_{yx} := \sqrt{N_y} V^\dagger_{\gamma,y} M_x V_{\gamma,y} \sqrt{N_y},$$

respectively.

Table 1. Counts of random trials (out of 100 000) in which one uncertainty bound is strictly weaker than another for rank-one PVMs in dimension d with $n = d$ outcomes. 'Number of counts' for, e.g. 'EUR1 < EUR3', means the number of cases in which the bound (13) is numerically smaller (i.e. weaker) than the Berta *et al* bound (24).

Number of counts	$d = 2,\ n = 2$	$d = 3,\ n = 3$	$d = 4,\ n = 4$
EUR1 < EUR3	76 325	44 045	39 426
EUR2 < EUR3	394	0	0
EUR2 < EUR1	0	0	0

Table 2. Counts of negative gaps (out of 100 000 random trials) for rank-one POVMs with varying dimensions d and outcomes $n > d$.

Number of counts	$d = 2, n = 3$	$d = 2, n = 4$	$d = 3, n = 4$
EUR1 < EUR3	46 123	44 905	41 855
EUR2 < EUR3	0	0	0
EUR2 < EUR1	0	0	0

Remark 3.9. The Groenewold–Ozawa information gain terms in equation (22) appear to be related to a measure of classicality, postulated in [58] and derived in [59], with which they coincide in the case of a single projective measurement; see, for example, equation (23) in [59]. In the present setting, however, the Groenewold–Ozawa information gain naturally arises also for general POVMs, generalizing its operational use beyond the projective setting. A thorough examination of this connection is left for future work.

The bound in theorem 3.7 admits a particularly transparent form when at least one element of the measurement-state triple is rank-one, as summarized in the following corollary.

Corollary 3.10. *Assume now at least one among γ_Q, **M**, or **N** is rank-one. In this case [20]*

$$I_{GO}\left(\gamma_Q;\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right) = I_{GO}\left(\gamma_Q;\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}\right) = H(\gamma_Q)\ ,$$

and hence,

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant H(\gamma_Q) + \max\left\{D(\gamma_Q\|\tilde{\gamma}_Q)\ ,\ D(\gamma_Q\|\tilde{\eta}_Q)\right\}. \tag{23}$$

3.3. Numerical benchmark codes and results

In this section we present the benchmark code together with the corresponding numerical results.

We compare our bounds (13) and (22) with the well-known entropic uncertainty relation of Berta *et al* [24]:

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant -\max_{x,y}\log_2\left\|\sqrt{N_y}\sqrt{M_x}\right\|_\infty^2 + H(\gamma_Q), \tag{24}$$

where the measurements $\{M_x\}_x$ and $\{N_y\}_y$ are rank-one projective valued measurements (PVMs). Our bounds (13) and (22) apply to general POVMs; however, for a fair comparison with (24), we restrict to the rank-one PVM setting in this subsection.

For general POVMs, equation (24) does not hold, although it does apply whenever at least one of the POVMs has rank-one elements [60]. Motivated by this fact, we also perform numerical tests on more general rank-one POVMs.

The three bounds compared in the numerical experiments are:

$$\begin{aligned}
\mathrm{EUR1} &:= -\max_{x,y}\log_2\left\|\sqrt{N_y}\sqrt{\gamma_Q}\sqrt{M_x}\right\|_2^2,\\
\mathrm{EUR2} &:= H(\gamma_Q) + \max\left\{D(\gamma_Q\|\tilde{\gamma}_Q), D(\gamma_Q\|\tilde{\eta}_Q)\right\},\\
\mathrm{EUR3} &:= H(\gamma_Q) - \max_{x,y}\log_2\left\|\sqrt{N_y}\sqrt{M_x}\right\|_\infty^2.
\end{aligned}$$

We perform 100 000 random trials for each configuration. The results are summarized in table 1 and figure 2 for rank-one PVMs, and in table 2 and figure 3 for general rank-one POVMs. Here d denotes the Hilbert space dimension and n the number of measurement outcomes.

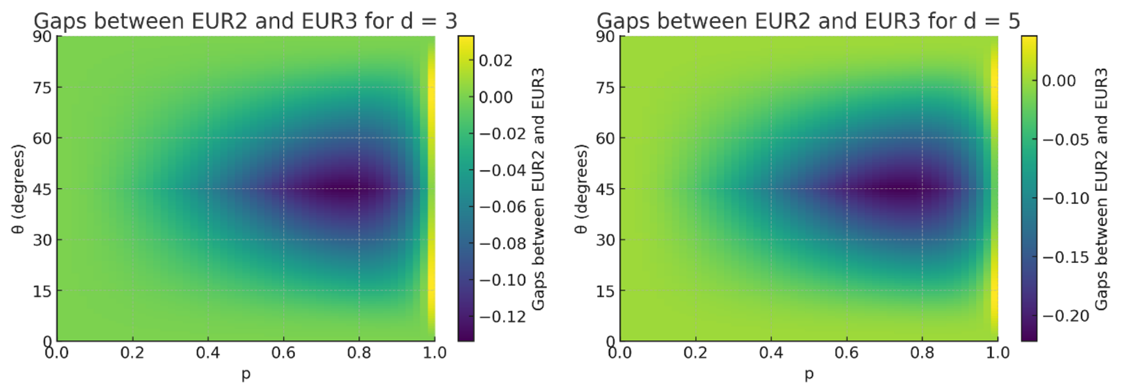

Figure 1. Negative gaps between EUR2 (22) and EUR3 (24) for mutually unbiased PVMs in dimensions $d = 3$ and $d = 5$. The shaded regions correspond to parameter ranges (p, θ) where EUR2 becomes weaker than EUR3. These counterexamples occur for states that interpolate between the two bases and illustrate that EUR2 is not uniformly stronger than the Berta *et al* bound.

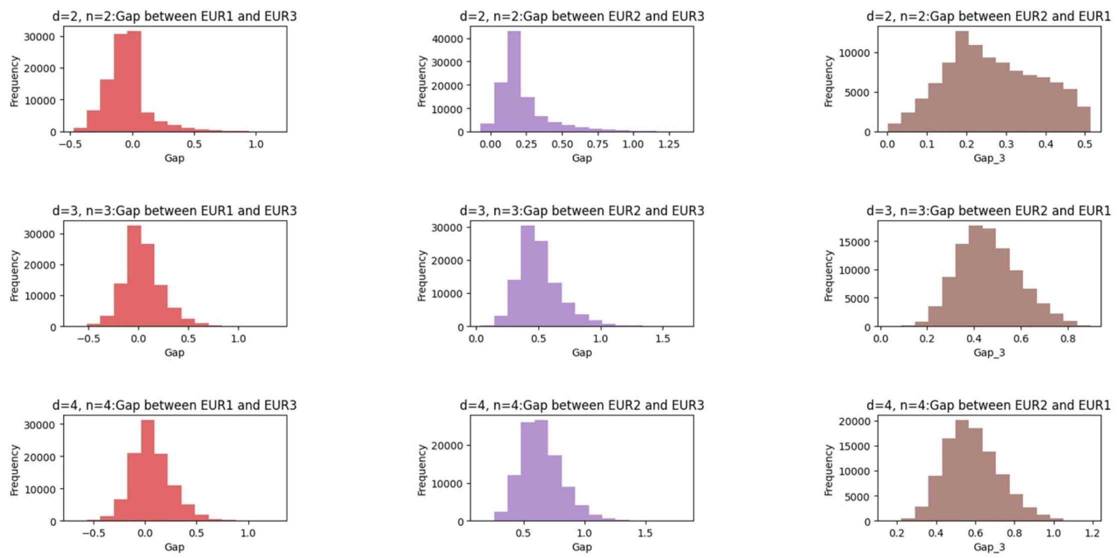

Figure 2. Numerical comparison of the bounds EUR1 (13), EUR2 (22), and EUR3 (24) for random rank-one PVMs. Each point corresponds to one of 100 000 random trials. The plot illustrates that EUR2 is consistently tighter than the Berta *et al* bound for $d \geqslant 3$, while EUR1 shows no uniform ordering relative to EUR3.

In table 1 we generate random rank-one PVMs. Since any rank-one PVM in dimension d must have exactly d outcomes, we only consider the case $n = d$. The numerical results show that the bound (13) is not uniformly comparable to the bound of Berta *et al* while bound (22) is consistently stronger than (24) for all cases with $n \geqslant 3$, and also always stronger than (13).

We emphasize, however, that in general bound (22) is not strictly comparable with the Berta *et al* bound (24). Although table 1 shows no negative gaps for $d \geqslant 3$ in the random tests, explicitly fine-tuned counterexamples can be constructed. Consider mutually unbiased PVMs $\mathbf{M}$ and $\mathbf{N}$ and a state of the form

$$\gamma_Q = (1-p)\,\frac{\mathbb{1}}{d} + p|\psi\rangle\langle\psi|,$$

where

$$|\psi\rangle \propto \cos\theta\,|m_x\rangle + \sin\theta\,|n_x\rangle\,,$$

with $|m_x\rangle$ and $|n_y\rangle$ denoting mutually unbiased measurement bases. For parameters near $p \approx 0.75$ and $\theta \approx 45^\circ$, negative gaps reproducibly appear, as shown in figure 1 for $d = 3$ and $d = 5$.

These figures indicate that such negative gaps persist over a nontrivial region of (p, θ) that depends on the dimension. Exact mutual unbiasedness is not required; near-MUB measurements exhibit the same behavior. Random generation rarely produces such pairs in higher dimensions, which explains their

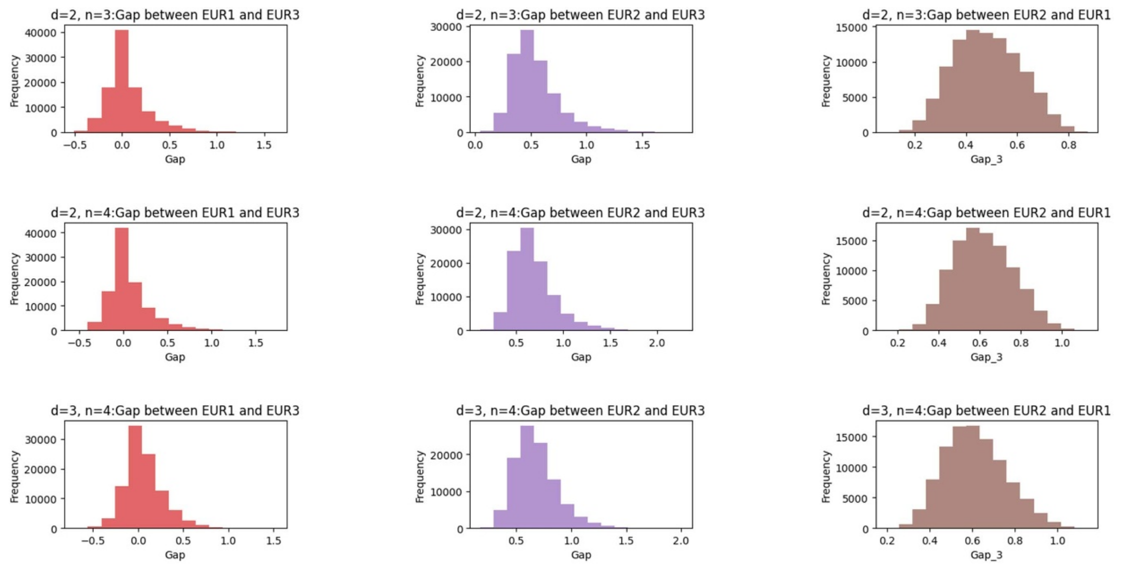

Figure 3. Numerical comparison of EUR1, EUR2, and EUR3 for random rank-one POVMs in various dimensions. The data show the same trend as in the projective case: EUR2 dominates EUR3 in nearly all instances, while EUR1 exhibits mixed behavior. The increased variety of POVMs broadens the range of achievable overlaps and highlights the structural advantage of the retrodictive bound.

absence from the general PVM tests. These counterexamples highlight a key conceptual difference: our retrodictive entropic uncertainty relation depends on the joint interplay between the state and the two measurements, whereas the Berta *et al* relation does not incorporate this dependence.

As we extend to more general rank-one POVMs, similar behavior appears (see table 2 and figure 3). The upper limit of EUR3 is now attained for uniformly split mutually unbiased POVMs of the form $\frac{1}{m}|m_x\rangle\langle m_x|$ and $\frac{1}{m}|n_y\rangle\langle n_y|$ with m being some constant.

4. Conclusion

In this paper, we start from the quantum minimum change principle and shown that, for all quantum-to-classical channels, the information projection (7) onto fixed outcome statistics selects a single retrodictive update independent of the divergence used. This yields a canonical and divergence-independent quantum Bayesian inverse (8) for any POVM and prior state, extending and unifying earlier constructions. Using this update, we introduced a symmetric retrodictive joint distribution (10) for pairs of POVMs and defined the mutual retrodictability $R(\mathbf{M};\mathbf{N})_\gamma$. Then, theorem 3.2 shows that this quantity is always finite and bounded above by $H(\gamma_Q)$.

We then used the retrodictive joint distribution to derive two entropic uncertainty relations, equations (13) and (22), which express uncertainty directly in terms of the prior state and the retrodictive action of the POVMs. The second bound links the entropic trade-off to relative entropies between the prior and its retrodictively updated versions, giving the relation a clear information-theoretic interpretation. Furthermore, the second bound is an always-valid form expressed via the Groenewold–Ozawa information gain, making explicit the connection between our entropic uncertainty relations and information-theoretic approaches to the thermodynamics of quantum measurements [23].

Finally, we benchmarked our retrodictive uncertainty relations against the Berta *et al* bound (24), both for rank-one PVMs and for rank-one POVMs in low dimensions. The numerical evidence shows that our bounds are consistently tighter across broad classes of measurements and states, while also identifying structured counterexamples such as near mutually unbiased settings.

Acknowledgment

J K and K T acknowledge support from JST SPRING Grant No. JPMJSP2125. All authors acknowledge support from MEXT Quantum Leap Flagship Program (MEXT QLEAP) Grant No. JPMXS0120319794;

from MEXT-JSPS Grant-in-Aid for Transformative Research Areas (A) 'Extreme Universe,' No. 21H05183; and from JSPS KAKENHI Grants Nos. 23K03230 and 26K00621.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.17636821 [63].

Appendix A. Derivation of the bipartite state γ_{XQ} in equation (6)

Following the definition of 'forward' output–input bipartite state $\mathcal{Q}_{\mathrm{fwd}}$ in [15], consider the canonical purification of γ_Q:

$$|\sqrt{\gamma_Q}\rangle\rangle := \sum_{ij} \langle i|\sqrt{\gamma_Q}|j\rangle \; |i\rangle_{Q_1}|j\rangle_{Q_2},$$

where $Q_1 \cong Q_2 \cong Q$ and $\{|i\rangle\}$ is the chosen orthonormal basis of $\mathcal{H}_Q$. For the quantum–classical measurement channel defined in (5), we obtain

$$\begin{aligned}
\mathcal{Q}^{\mathcal{M}}_{\mathrm{fwd}} &= (\mathcal{M}\otimes\mathcal{I})\left(|\sqrt{\gamma_Q}\rangle\rangle\langle\langle\sqrt{\gamma_Q}|\right)\\
&= \left(\mathbb{1}_X\otimes\sqrt{\gamma_Q^T}\right) C_{\mathcal{M}} \left(\mathbb{1}_X\otimes\sqrt{\gamma_Q^T}\right)\\
&= \left(\mathbb{1}_X\otimes\sqrt{\gamma_Q^T}\right)\left(\sum_{i,j}\mathcal{M}(|i\rangle\langle j|)\otimes|i\rangle\langle j|\right)\left(\mathbb{1}_X\otimes\sqrt{\gamma_Q^T}\right)\\
&= \left(\mathbb{1}_X\otimes\sqrt{\gamma_Q^T}\right)\left(\sum_{i,j}\sum_{x\in\mathcal{X}}\langle j|M_x|i\rangle\,|x\rangle\langle x|\otimes|i\rangle\langle j|\right)\left(\mathbb{1}_X\otimes\sqrt{\gamma_Q^T}\right)\\
&= \sum_{x\in\mathcal{X}}|x\rangle\langle x|\otimes\sqrt{\gamma_Q^T}\left(\sum_{i,j}\langle j|M_x|i\rangle\,|i\rangle\langle j|\right)\sqrt{\gamma_Q^T}\\
&= \sum_{x\in\mathcal{X}}|x\rangle\langle x|\otimes\sqrt{\gamma_Q^T}M_x^T\sqrt{\gamma_Q^T},
\end{aligned}$$

where $C_{\mathcal{M}} = \sum_{i,j}\mathcal{M}(|i\rangle\langle j|)\otimes|i\rangle\langle j|$ is the Choi operator of the measurement channel. Finally, since in this case the operator $\mathcal{Q}^{\mathcal{M}}_{\mathrm{fwd}}$ is separable, we suppress the partial transposition on the second space for the sake of simplicity and readability, which leads to our definition of γ_{XQ} in equation (6).

Appendix B. Proof of theorems 2.1 and 2.2

Proof of theorem 2.1. The claim follows from three properties satisfied by all divergences listed above: the direct-sum property, monotonicity under partial trace, and faithfulness [61]. In particular, monotonicity implies that

$$\min_{\sigma_{XQ}}\mathbb{D}(\sigma_{XQ}\|\gamma_{XQ}) \geqslant \mathbb{D}(\tau\|p),$$

so that any state that saturates the inequality is necessarily a global minimizer.

The argument proceeds in the same way for each divergence in the list. For illustration, we discuss two representative cases. First, consider the Umegaki relative entropy, for which we have

$$\begin{aligned}
D(\sigma_{XQ}\|\gamma_{XQ}) &= D\left(\sum_{x\in\mathcal{X}}\tau(x)\,|x\rangle\langle x|_X\otimes\sigma_{Q|x}\middle\|\sum_{x\in\mathcal{X}}p(x)\,|x\rangle\langle x|_X\otimes\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right)\\
&= D(\tau\|p) + \sum_{x\in\mathcal{X}}p(x)\,D\left(\sigma_{Q|x}\middle\|\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right),
\end{aligned}$$

where the second equality is verified by the direct-sum property. The minimum of $D(\sigma_{XQ}\|\gamma_{XQ})$ is achieved when the second term vanishes, and from the faithfulness this means

$$\sigma_{Q|x} = \frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}} \tag{B1}$$

for all $x\in\mathcal{X}$. Next, for the sandwiched Rényi relative entropy, we obtain

$$\begin{aligned}\tilde{D}_\alpha(\sigma_{XQ}\|\gamma_{XQ}) &= \tilde{D}_\alpha\left(\sum_{x\in\mathcal{X}}\tau(x)|x\rangle\langle x|_X\otimes\sigma_{Q|x}\middle\|\sum_{x\in\mathcal{X}}p(x)|x\rangle\langle x|_X\otimes\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right)\\ &= \frac{1}{\alpha-1}\log_2\sum_{x\in\mathcal{X}}\tau(x)^\alpha p(x)^{1-\alpha}\,\mathrm{Tr}\left\{\left(\left(\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right)^{\frac{1-\alpha}{2\alpha}}\sigma_{Q|x}\left(\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right)^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right\}\\ &= \frac{1}{\alpha-1}\log_2\sum_{x\in\mathcal{X}}\tau(x)^\alpha p(x)^{1-\alpha}\;2^E,\end{aligned}$$

where

$$E := (\alpha-1)\tilde{D}_\alpha\left(\sigma_{Q|x}\middle\|\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right).$$

The second equality again follows from the direct-sum property. The minimum of $\tilde{D}_\alpha(\sigma_{XQ}\|\gamma_{XQ})$ is achieved when

$$\tilde{D}_\alpha\left(\sigma_{Q|x}\middle\|\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right) = 0$$

for all $x\in\mathcal{X}$ and for all $\alpha\in[1/2,1)\cup(1,\infty)$, thus implying equation (B1). Similar arguments hold for the other cases. Therefore, we conclude

$$\min_{\sigma_{XQ}}\mathbb{D}(\sigma_{XQ}\|\gamma_{XQ}) = \mathbb{D}(\sigma_X\|\gamma_X),$$

and the minimum is achieved by $\sigma_{XQ} = \hat{\gamma}_{XQ}$.

□

Proof of theorem 2.2. Due to the common cq structure, we have

$$\begin{aligned}\|\sigma_{XQ}-\gamma_{XQ}\|_1 &= \sum_{x\in\mathcal{X}}\left\|\tau(x)\sigma_{Q|x}-p(x)\frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}}\right\|_1\\ &\geqslant \sum_{x\in\mathcal{X}}|\tau(x)-p(x)|,\end{aligned}$$

where the equality in the second line holds if and only if

$$\sigma_{Q|x} = \frac{\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}}{\mathrm{Tr}\{\gamma_Q M_x\}} =: \hat{\gamma}_{Q|x},$$

for all x.

□

Appendix C. Proof of theorem 3.2

Proof of theorem 3.2. The mutual retrodictability is defined as the mutual information of the joint distribution

$$\Pr\{y\leftarrow x\} = \mathrm{Tr}\left\{N_y\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}\right\},$$

so non-negativity follows immediately. It remains to prove the upper bound.

Choose an orthonormal basis that diagonalizes γ_Q, writing $\gamma_Q = \sum_i g_i |\psi_i\rangle\langle\psi_i|$. Using this basis, define the canonical purification

$$|\Psi_\gamma\rangle := \sum_i \sqrt{g_i}\,|\psi_i\rangle \otimes |\psi_i\rangle\,.$$

By construction,

$$\mathrm{Tr}_1\{|\Psi_\gamma\rangle\langle\Psi_\gamma|\} = \mathrm{Tr}_2\{|\Psi_\gamma\rangle\langle\Psi_\gamma|\} = \gamma_Q.$$

Moreover, for all operators X and Y,

$$\mathrm{Tr}\{X \otimes Y\,|\Psi_\gamma\rangle\langle\Psi_\gamma|\} = \mathrm{Tr}\{X\sqrt{\gamma_Q}\,Y^T\sqrt{\gamma_Q}\}\,,$$

where the transpose is taken in the eigenbasis of γ_Q. Hence,

$$\Pr\{y \leftarrow x\} = \mathrm{Tr}\left\{N_y \otimes M_x^T\,|\Psi_\gamma\rangle\langle\Psi_\gamma|\right\}\ .$$

Let $\mathcal{N}$ and $\mathcal{M}^T$ denote the qc channels, as defined in equation (5), associated with $\mathbf{N}$ and $\mathbf{M}^T := \{M_x^T : x \in \mathcal{X}\}$, respectively. The above implies

$$R(\mathbf{M};\mathbf{N})_\gamma = I(X;Y)_{(\mathcal{N}\otimes\mathcal{M}^T)(|\Psi_\gamma\rangle\langle\Psi_\gamma|)}\,.$$

We now apply the data-processing inequality for the quantum mutual information:

$$\begin{aligned} R(\mathbf{M};\mathbf{N})_\gamma &= I(Y;X)_{(\mathcal{N}\otimes\mathcal{M}^T)(|\Psi_\gamma\rangle\langle\Psi_\gamma|)} \\ &\leqslant I(Y;Q)_{(\mathcal{N}\otimes\mathcal{I})(|\Psi_\gamma\rangle\langle\Psi_\gamma|)}\,. \end{aligned}$$

Since $(\mathcal{N}\otimes\mathcal{I})(|\Psi_\gamma\rangle\langle\Psi_\gamma|)$ is a cq state with marginal γ_Q, the quantum mutual information satisfies

$$I(Y;Q)_{(\mathcal{N}\otimes\mathcal{I})(|\Psi_\gamma\rangle\langle\Psi_\gamma|)} \leqslant H(\gamma_Q)\,.$$

Combining inequalities yields the claimed bound

$$R(\mathbf{M};\mathbf{N})_\gamma \leqslant H(\gamma_Q)\,.$$

□

Appendix D. Proof of theorem 3.7

Proof of theorem 3.7. Since the same exact reasoning can be repeated substituting $\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}$ with $\tilde{\mathcal{J}}^{\mathbf{N}\to\mathbf{M}}$, we prove the lower bound (22) only for $\tilde{\gamma}_Q$.

The proof follows the idea in [21]. For a state γ_Q and $\mathcal{E} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ a positive, trace-preserving and sub-unital map, we have:

$$\begin{aligned} H(\mathcal{E}(\gamma_Q)) - H(\gamma_Q) &= \mathrm{Tr}\{\gamma_Q \log_2 \gamma_Q\} - \mathrm{Tr}\{\mathcal{E}(\gamma_Q)\log_2 \mathcal{E}(\gamma_Q)\} \\ &= \mathrm{Tr}\{\gamma_Q \log_2 \gamma_Q\} - \mathrm{Tr}\left\{\gamma_Q \mathcal{E}^\dagger(\log_2 \mathcal{E}(\gamma_Q))\right\} \\ &\geqslant \mathrm{Tr}\{\gamma_Q \log_2 \gamma_Q\} - \mathrm{Tr}\left\{\gamma_Q \log\left(\mathcal{E}^\dagger\circ\mathcal{E}\right)(\gamma_Q)\right\} \\ &= D\left(\gamma_Q \| \left(\mathcal{E}^\dagger\circ\mathcal{E}\right)(\gamma_Q)\right), \end{aligned} \tag{D1}$$

where the second equality is from the definition of adjoint and the inequality follows from operator concavity of the logarithm and the operator Jensen inequality for positive unital maps [62]. The sub-unitality is required for the non-negativity of $D(\gamma_Q\|(\mathcal{E}^\dagger\circ\mathcal{E})(\gamma_Q))$.

Noting that the sequential measurement channel $\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}$ is CPTP, sub-unital, and

$$\begin{aligned} \tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}(\gamma_Q) &= \sum_{y,x} \Pr\{y\leftarrow x\}\,\frac{\sqrt{N_y}\sqrt{\gamma_Q}M_x\sqrt{\gamma_Q}\sqrt{N_y}}{\Pr\{y\leftarrow x\}} \otimes |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y \\ &=: \omega_{QXY}, \end{aligned}$$

we thus have

$$\begin{aligned}
H\left(\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}(\gamma_Q)\right) - H(\gamma_Q) &= H(QXY)_\omega - H(\gamma_Q) \\
&= H(XY)_\omega + \sum_{x,y} \Pr\{y \leftarrow x\} H\left(\frac{\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}_{x,y}(\gamma_Q)}{\Pr\{y \leftarrow x\}}\right) - H(\gamma_Q) \\
&= H(XY)_\omega - I_{GO}\left(\gamma_Q; \tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right) \\
&\geqslant D\left(\gamma_Q \middle\| \left[\left(\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right)^\dagger \circ \tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}\right](\gamma_Q)\right) \\
&= D(\gamma_Q \| \tilde{\gamma}_Q),
\end{aligned}$$

where

$$\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}}_{x,y}(\gamma_Q) := \sqrt{N_y}\sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}\sqrt{N_y}\,.$$

Note that $(\tilde{\mathcal{J}}^{\mathbf{M}\to\mathbf{N}})^\dagger$ is trace-non-increasing, so $\mathrm{Tr}\{\tilde{\gamma}_Q\} \leqslant 1$. Since for Umegaki relative entropy, we have $B \leqslant B' \Rightarrow D(A||B) \geqslant D(A||B')$, therefore the use of $\tilde{\gamma}_Q$ actually makes the bound tighter.

Combine the above equation with the well-known bound:

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geqslant H(XY)_\omega\,,$$

and consider the measurement order swapping, we recover the bound (22). □

ORCID iDs

Jiaxi Kuang 0009-0007-3328-9292
Kensei Torii 0009-0008-4072-945X
Francesco Buscemi 0000-0001-9741-0628

References

[1] Pólya G 1954 *Induction and Analogy in Mathematics* (*Induction and Analogy in Mathematics*) (Princeton University Press)
[2] Cox R 1961 *Algebra of Probable Inference* (Johns Hopkins University Press)
[3] De Finetti B 1974 *Theory of Probability* (*Wiley Series in Probability and Mathematical Statistics*) (Wiley)
[4] Jeffrey R 1965 *The Logic of Decision* (McGraw-Hill)
[5] Zellner A 1988 *Am. Stat.* **42** 278
[6] Jeffreys H 1998 *The Theory of Probability* (*Oxford Classic Texts in the Physical Sciences*) (OUP)
[7] Jaynes E T 2003 *Probability Theory: The Logic of Science* (Cambridge University Press)
[8] Laplace P S 1986 *Stat. Sci.* **1** 364
[9] Barnett S M, Jeffers J and Pegg D T 2021 *Symmetry* **13** 586
[10] Jeffers J, Oi D K L and Brougham T 2024 *Phil. Trans. R. Soc. A* **382** 20230338
[11] Watanabe S 1955 *Rev. Mod. Phys.* **27** 179
[12] Barnett S M, Pegg D T and Jeffers J 2000 *J. Mod. Opt.* **47** 1779
[13] Johansen L M 2007 *Phys. Rev. A* **76** 012119
[14] Parzygnat A J and Buscemi F 2023 *Quantum* **7** 1013
[15] Bai G, Buscemi F and Scarani V 2025 *Phys. Rev. Lett.* **135** 090203
[16] Jaynes E T 1957 *Phys. Rev.* **106** 620
[17] Caticha A and Giffin A 2006 *AIP Conf. Proc.* **872** 31
[18] Seidenfeld T 1986 *Phil. Sci.* **53** 467–91
[19] Uffink J 1996 *Stud. Hist. Phil. Sci. B* **27** 47
[20] Buscemi F, Hayashi M and Horodecki M 2008 *Phys. Rev. Lett.* **100** 210504
[21] Buscemi F, Das S and Wilde M M 2016 *Phys. Rev. A* **93** 062314
[22] Danageozian A, Wilde M M and Buscemi F 2022 *PRX Quantum* **3** 020318
[23] Minagawa S, Mohammady M H, Sakai K, Kato K and Buscemi F 2025 *npj Quantum Inf.* **11** 18
[24] Berta M, Christandl M, Colbeck R, Renes J M and Renner R 2010 *Nat. Phys.* **6** 659
[25] Jiang M, Luo S and Fu S 2013 *Phys. Rev. A* **87** 022310
[26] Paulsen V I and Shultz F 2013 *J. Math. Phys.* **54** 072201
[27] Wilde M M 2017 *Quantum Information Theory* 2nd edn (Cambridge University Press)
[28] Kye S-H 2022 *J. Math. Phys.* **63** 092202
[29] Frembs M and Cavalcanti E G 2024 *J. Phys. A: Math. Theor.* **57** 265301
[30] Umegaki H 1961 *Proc. Japan Acad.* **37** 459
[31] Petz D 1985 *Publ. Res. Inst. Math. Sci.* **21** 787
[32] Petz D 1986a *Rep. Math. Phys.* **23** 57
[33] Müller-Lennert M, Dupuis F, Szehr O, Fehr S and Tomamichel M 2013 *J. Math. Phys.* **54** 122203

[34] Wilde M M, Winter A and Yang D 2014 *Commun. Math. Phys.* **331** 593
[35] Petz D and Ruskai M B 1998 Infinite dimensional analysis *Quantum Probab. Relat. Top.* **01** 83
[36] Matsumoto K 2015 *Nagoya Winter Workshop: Reality and Measurement in Algebraic Quantum Theory* (Springer) pp 229–73
[37] Belavkin V P and Staszewski P 1982 *Ann. l'IHP Phys. Théor.* **37** 51–58
[38] Hughston L P, Jozsa R and Wootters W K 1993 *Phys. Lett. A* **183** 14
[39] Hall M J W 1997 *Phys. Rev. A* **55** 100
[40] Fuchs C A 2003 *J. Mod. Opt.* **50** 987
[41] Leifer M S 2006 *Phys. Rev. A* **74** 042310
[42] Leifer M S 2007 *AIP Conf. Proc.* **889** 172–86
[43] Leifer M S and Spekkens R W 2013 *Phys. Rev. A* **88** 052130
[44] Jacobs B 2019 *J. Artif. Intell. Res.* **65** 783
[45] Buscemi F, Schindler J and Šafránek D 2023 *New J. Phys.* **25** 053002
[46] Nagasawa T, Kato K, Wakakuwa E and Buscemi F 2024 *Phys. Rev. Res.* **6** 043327
[47] Liu M, Scarani V, Auffèves A and Laverick K T 2025 *Phys. Rev. A* **112** L030203
[48] Liu M, Bai G and Scarani V 2025 arXiv:2510.08447
[49] Petz D 1986 *Commun. Math. Phys.* **105** 123
[50] Petz D 1988 *Q. J. Math.* **39** 97
[51] Nagasawa T, Wakakuwa E, Kato K and Buscemi F 2025 *Rep. Prog. Phys.* **88** 117601
[52] Buscemi F, Kobayashi K, Minagawa S, Perinotti P and Tosini A 2023 *Quantum* **7** 1035
[53] Groenewold H J 1971 *Int. J. Theor. Phys.* **4** 327
[54] Ozawa M 1986 *J. Math. Phys.* **27** 759
[55] Jacobs K 2009 *Phys. Rev. A* **80** 012322
[56] Winter A 2004 *Commun. Math. Phys.* **244** 157
[57] Wilde M M, Hayden P, Buscemi F and Hsieh M-H 2012 *J. Phys. A: Math. Theor.* **45** 453001
[58] Korzekwa K, Lostaglio M, Jennings D and Rudolph T 2014 *Phys. Rev. A* **89** 042122
[59] Hall M J W 2023 *Phys. Rev. A* **107** 062215
[60] Coles P J, Yu L, Gheorghiu V and Griffiths R B 2011 *Phys. Rev. A* **83** 062338
[61] Khatri S and Wilde M M 2024 arXiv:2011.04672 [quant-ph]
[62] Choi M-D 1974 *Ill. J. Math.* **18** 565
[63] Kuang J 2025 JiaxiKuang/Uncertanty-bounds-compare: Gaps_benchmark_with_heatmapof_MUB (v1.2) *Zenodo* (https://doi.org/10.5281/zenodo.17636821)