

Quantum measurement retrodiction and entropic uncertainty relations

Jiaxi Kuang^{*}, Kensei Torii[†], and Francesco Buscemi[‡]

Department of Mathematical Informatics, Nagoya University, Japan

Abstract

We study quantum measurement retrodiction using the principle of minimum change. For quantum-to-classical measurement channels, we show that all standard quantum divergences select 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*, for which we also derive a general upper bound that depends only on the prior state and holds for all measurements. This structure leads to two retrodictive entropic uncertainty relations, expressed directly in terms of the prior state and the POVMs, but valid independently of the retrodictive framework and fully compatible with the conventional operational interpretation of entropic uncertainty relations. Finally, we benchmark these relations numerically and find that they provide consistently tighter bounds than existing entropic uncertainty relations over 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:

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

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 [Pól54, Cox61, DF74, Jef65, Zel88, Jef98, Jay03].

Once the prior has been updated to incorporate the newly acquired information, namely the observed outcome $\bar{x}$, it can be used to *predict* the statistics of other observations. Consider, for instance, another measurement with outcomes in the set $\mathcal{Y} = \{y\}$ and likelihood function $\psi(y|s)$. The updated belief $\hat{\gamma}(s|\bar{x})$ allows us to compute the expected probability of each outcome y as

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

This quantity $\Pr\{y|\bar{x}\}$ represents our rational expectation for the probability of obtaining outcome y in the yet-unobserved measurement, given that the first observation has already yielded $\bar{x}$. It encapsulates the predictive power of Bayesian reasoning: from one piece of evidence, we infer a consistent net of expectations for future data.

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

^{*}kuang.jiaxi.g5@s.mail.nagoya-u.ac.jp

[†]toriikensei@nagoya-u.jp

[‡]buscemi@nagoya-u.jp

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 $M_x \geq 0$ such that $\sum_{x \in \mathcal{X}} M_x = \mathbf{1}$. Upon obtaining a specific outcome $\bar{x} \in \mathcal{X}$, the state of subsystem B is updated, according to Born’s rule, to

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

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

In the classical case, the quantity $\hat{\gamma}(s|\bar{x})$ admits another interpretation: it can be viewed not merely as a prediction about future events, but as a *retrodiction*, 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” [Lap86].

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 Ref. [BJP21, JOB24]. 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 Refs. [Wat55, BPJ00]); 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 [Joh07]); or they are derived from a set of *axioms* or desiderata whose appeal and physical justification can vary [PB23].

Only very recently, a fully quantum generalization of Bayes’ rule has been derived from the *minimum change principle* [BBS25]. 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 [Zel88]. 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 [Jay57, Jay03, CG06], although see Refs. [Sei86] and [Uff96].

In this paper we focus on the task of retrodiction based on the outcome of a quantum measurement. By restricting the analysis to quantum-to-classical (measurement) channels, we show that the quantum minimum change principle of Ref. [BBS25] 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. 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 Ref. [BCC⁺10], showing that our approach yields tighter bounds in a wide range of scenarios.

The paper is organized as follows. In Sec. 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 Sec. 3 we examine the resulting retrodicted states and develop the symmetric retrodictive joint probability that underpins our later constructions. In Sec. 4 we introduce the mutual retrodictability $R(\mathbf{M}; \mathbf{N})_\gamma$ and establish its general upper bound in Theorem 3. In Sec. 5 we use this framework to derive the retrodictive entropic uncertainty relations (13) and (19). In Sec. 6 we benchmark our bounds numerically against the entropic uncertainty relation of Berta *et al.* [BCC⁺10]. Finally, in Sec. 7 we summarize our results and outline directions for future work.

2 Minimum change principle for quantum measurements

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 [Jef65]. 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:

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

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

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

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 positive operator-valued measure (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 $p(x) = \text{Tr}\{M_x \gamma_Q\}$. It is convenient to summarize the resulting outcome distribution by introducing a quantum-to-classical channel, namely a completely positive trace-preserving map

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

Here the kets $|x\rangle$ 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.

Ref. [BBS25] developed a framework 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 Ref. [BBS25] 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 [BBS25]. For measurement channels, however, we may introduce the bipartite state

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

It is immediate to verify that $\text{Tr}_Q\{\gamma_{XQ}\} = \sum_x p(x)|x\rangle\langle x|_X$ and $\text{Tr}_X\{\gamma_{XQ}\} = \gamma_Q$, so the marginals recover the expected output distribution and the input state, respectively. Aside from a partial transposition that plays no role here, this coincides with the definition used in [BBS25], based on the Choi isomorphism and the standard purification [Wil17].

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

$$\min_{\sigma_{XQ}} \mathbb{D}(\sigma_{XQ} \| \gamma_{XQ}) , \quad (6)$$

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 $\sigma_X := \text{Tr}_Q\{\sigma_{XQ}\} = \sum_{x \in \mathcal{X}} \tau(x)|x\rangle\langle x|_X$, 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

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

Theorem 1. Assume $\text{Tr}\{\gamma_Q M_x\} > 0$ for all $x \in \mathcal{X}$. For the following quantum divergences (below, A and B denote density matrices, and we assume $B > 0$):

1. Umegaki relative entropy [Ume61]: $\mathbb{D}(A\|B) \equiv D(A\|B) := \text{Tr}\{A(\log_2 A - \log_2 B)\}$;
2. Petz–Rényi relative entropy [Pet85, Pet86a]: $D_\alpha(A\|B) := \frac{1}{\alpha-1} \log_2 \text{Tr}\{A^\alpha B^{1-\alpha}\}$, for $\alpha \in (0, 1) \cup (1, 2]$;
3. sandwiched (panini) Rényi relative entropy [MLDS⁺13, WWY14]: $\tilde{D}_\alpha(A\|B) := \frac{1}{\alpha-1} \log_2 \text{Tr}\{(B^{\frac{1-\alpha}{2\alpha}} A B^{\frac{1-\alpha}{2\alpha}})^\alpha\}$, for $\alpha \in [1/2, 1) \cup (1, \infty)$;
4. geometric Rényi relative entropy [PR98, Mat15]: $\hat{D}_\alpha(A\|B) := \frac{1}{\alpha-1} \log_2 \text{Tr}\{B (B^{-1/2} A B^{-1/2})^\alpha\}$, for $\alpha \in (0, 1) \cup (1, 2]$;
5. Belavkin–Staszewski relative entropy [BS82]: $D_{BS}(A\|B) := \text{Tr}\{A \log_2 (A^{1/2} B^{-1} A^{1/2})\}$;

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

$$\hat{\gamma}_{XQ} := \arg \min_{\sigma_{XQ}} \mathbb{D}(\sigma_{XQ} \| \gamma_{XQ}) = \sum_{x \in \mathcal{X}} \tau(x)|x\rangle\langle x|_X \otimes \frac{\sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}}{\text{Tr}\{M_x \gamma_Q\}} . \quad (7)$$

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

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

so that any state achieving the right-hand side 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:

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

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

$$\sigma_{Q|x} = \frac{\sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}}{\text{Tr } \gamma_Q M_x} \quad (8)$$

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

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

where

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

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

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

for all $x \in \mathcal{X}$, for all $\alpha \in [1/2, 1) \cup (1, \infty)$, thus implying Eq. (8). Similar arguments hold for the other cases. Therefore, we conclude

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

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

Besides relative entropies, the same conclusion holds also when the statistical divergence between forward and reverse processes is measured using Uhlmann fidelity (this follows directly from Ref. [BBS25]) and trace-distance:

Theorem 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 (7).

Proof. Due to the common cq structure, we have

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

where the equality in the second line holds if and only if $\sigma_{Q|x} = \frac{\sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}}{\text{Tr}\{\gamma_Q M_x\}} =: \hat{\gamma}_{Q|x}$, for all x . □

Theorems 1 and 2 reinforce the argument first proposed in [BBS25], which motivates the adoption of the conditional states

$$\hat{\gamma}_{Q|x} := \frac{1}{\text{Tr}\{\gamma_Q M_x\}} \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}, \quad (9)$$

as *the* quantum analogue of the classical Bayesian inverse in Eq. (1). This identification had been previously anticipated in the literature, e.g. [Fuc03, Lei06, Lei07, LS13, Jac19, PB23, BSŠ23, NKWB24], where it was argued largely on the basis of formal analogy with Bayes' rule rather than

as the outcome of a variational principle. We therefore refer to $\hat{\gamma}_{Q|\bar{x}}$ as the *quantum Bayesian inverse* associated with the measurement outcome $\bar{x}$. The results obtained in this semiclassical framework include the classical minimum change principle as a special case. In what follows, we study how the Bayesian inverse $\hat{\gamma}_{Q|\bar{x}}$, conditioned on a measurement outcome, can be used to retrodict information about the system state prior to the measurement.

Before concluding this section, we note that the states defined in (9) coincide with the outputs of the *Petz transpose map* [Pet86b, Pet88] 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 Ref. [NWKB25].

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 belief about the state and the new information about the observed outcome $\bar{x}$ to compute the corresponding retrodicted state and, through it, infer the probabilities associated with the outcomes of *another measurement*. The latter in this context should be understood as a hypothetical probe used to test the informational content of the retrodicted state. It need not be actually performed, nor must it correspond to a measurement carried out after the first one (although it can be *simulated* as such, see Section 5 for further details). Rather, it represents any measurement whose statistics we can compute *as if* the retrodicted state described the system's preparation prior to the first measurement. In this way, it serves as a diagnostic tool: by evaluating the probabilities associated with this secondary measurement, we assess how well the retrodicted state encapsulates the information that the original outcome $\bar{x}$ conveys about the system's earlier condition. Thus, the “other measurement” provides an operational interpretation of $\hat{\gamma}_{Q|\bar{x}}$, as a state that reproduces, through Born's rule, all rational expectations for future or counterfactual observations consistent with the retrodictive update.

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

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

Accordingly, we define the *joint* retrodictive probability of assigning outcome y to N from the information obtained through outcome x of M as

$$\Pr\{y \leftarrow x\} := \text{Tr}\{N_y \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}\} . \quad (10)$$

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

$$\sum_{y \in \mathcal{Y}} \Pr\{y \leftarrow x\} = \text{Tr}\{M_x \gamma_Q\} , \quad \sum_{x \in \mathcal{X}} \Pr\{y \leftarrow x\} = \text{Tr}\{N_y \gamma_Q\} . \quad (11)$$

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 [BKM⁺23]. 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 $\text{Tr}\{\mathcal{J}_x^{\mathbf{M}}(\rho)\} = \text{Tr}\{M_x \rho\}$ 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

$$\Pr\{x \rightarrow y\} = \text{Tr}\{N_y \mathcal{J}_x^{\mathbf{M}}(\gamma_Q)\},$$

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

4 Mutual retrodictability

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

Definition 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. 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

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

Remark 1. 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.

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

$$\Pr\{y \leftarrow x\} = \text{Tr}\{N_y \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}\},$$

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,

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

Moreover, for all operators X and Y ,

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

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

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

Let $\mathcal{N}$ and $\mathcal{M}^T$ denote the qc channels, as defined in Eq. (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:

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

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

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

Combining inequalities yields the claimed bound

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

□

Remark 2. 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.

5 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 Eq. (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 non-negative, we immediately obtain the lower bound

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

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

Theorem 4. 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:

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

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 $H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geq H(XY)$ we have:

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

□

A second retrodictive entropic uncertainty relation follows from the observation that, while in Section 3 we noted that the predictive joint probability distribution $\Pr\{x \rightarrow y\}$ 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 $\sqrt{M_x} \sqrt{\gamma_Q} \rightarrow \sqrt{\gamma_Q} \sqrt{M_x}$. Explicitly, we define

$$\mathcal{J}_x^{\mathbf{M}}(\gamma_Q) := U_{\gamma,x} \sqrt{M_x} \gamma_Q \sqrt{M_x} U_{\gamma,x}^\dagger = \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q}, \quad (14)$$

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:

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

which, by construction, when applied to γ_Q , satisfies

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

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

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

and

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

Using the above, we can establish the following alternative entropic uncertainty relation:

Theorem 5. 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 . Assume that at least one among γ_Q , $\mathbf{M}$, or $\mathbf{N}$ is rank-one. Then,

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geq H(\gamma_Q) + \max\{D(\gamma_Q \| \tilde{\gamma}_Q), D(\gamma_Q \| \tilde{\eta}_Q)\}, \quad (19)$$

where the operators $\tilde{\gamma}_Q$ and $\tilde{\eta}_Q$ are defined by

$$\begin{aligned} \tilde{\gamma}_Q &:= [(\tilde{\mathcal{J}}^{\mathbf{M} \rightarrow \mathbf{N}})^\dagger \circ \tilde{\mathcal{J}}^{\mathbf{M} \rightarrow \mathbf{N}}](\gamma_Q) \\ &= \sum_{x,y} \sqrt{M_x} U_{\gamma,x}^\dagger N_y \sqrt{\gamma_Q} M_x \sqrt{\gamma_Q} N_y U_{\gamma,x} \sqrt{M_x}, \end{aligned} \quad (20)$$

and

$$\begin{aligned}\tilde{\eta}_Q &:= \left[(\tilde{\mathcal{J}}^{\mathbf{N} \rightarrow \mathbf{M}})^\dagger \circ \tilde{\mathcal{J}}^{\mathbf{N} \rightarrow \mathbf{M}} \right] (\gamma_Q) \\ &= \sum_{x,y} \sqrt{N_y} V_{\gamma,y}^\dagger M_x \sqrt{\gamma_Q} N_y \sqrt{\gamma_Q} M_x V_{\gamma,y} \sqrt{N_y},\end{aligned}\tag{21}$$

respectively. Note that, in general, $\text{Tr } \tilde{\gamma}_Q, \text{Tr } \tilde{\eta}_Q \leq 1$.

Proof. Since the same exact reasoning can be repeated substituting $\tilde{\mathcal{J}}^{\mathbf{M} \rightarrow \mathbf{N}}$ with $\tilde{\mathcal{J}}^{\mathbf{N} \rightarrow \mathbf{M}}$, we prove the lower bound (19) only for $\tilde{\gamma}_Q$.

The proof follows the idea in [BDW16]. For a state γ_Q and $\mathcal{E} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}')$ a positive, trace-preserving and sub-unital map, we have:

$$\begin{aligned}H(\mathcal{E}(\gamma_Q)) - H(\gamma_Q) &= \text{Tr}\{\gamma_Q \log \gamma_Q\} - \text{Tr}\{\mathcal{E}(\gamma_Q) \log \mathcal{E}(\gamma_Q)\} \\ &= \text{Tr}\{\gamma_Q \log \gamma_Q\} - \text{Tr}\{\gamma_Q \mathcal{E}^\dagger(\log \mathcal{E}(\gamma_Q))\} \\ &\geq \text{Tr}\{\gamma_Q \log \gamma_Q\} - \text{Tr}\{\gamma_Q \log(\mathcal{E}^\dagger \circ \mathcal{E})(\gamma_Q)\} \\ &= D(\gamma_Q \| (\mathcal{E}^\dagger \circ \mathcal{E})(\gamma_Q)),\end{aligned}\tag{22}$$

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 [Cho74]. The sub-unitality is required for the non-negativity of $D(\gamma_Q \| (\mathcal{E}^\dagger \circ \mathcal{E})(\gamma_Q))$.

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

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

we thus have:

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

where

$$I_{GO}(\gamma_Q; \tilde{\mathcal{J}}^{\mathbf{M} \rightarrow \mathbf{N}}) = H(\gamma_Q) - \sum_{x,y} \text{Pr}(x,y) H\left(\frac{\tilde{\mathcal{J}}_{x,y}^{\mathbf{M} \rightarrow \mathbf{N}}(\gamma_Q)}{\text{Pr}(x,y)}\right)$$

is the Groenewold–Ozawa information gain [Gro71, Oza86].

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

Assume now at least one among γ_Q , $\mathbf{M}$, or $\mathbf{N}$ is rank-one, in this case, $I_{GO}(\gamma_Q; \tilde{\mathcal{J}}^{\mathbf{M} \rightarrow \mathbf{N}}) = H(\gamma_Q)$ [BHH08]. We then have:

$$\begin{aligned}H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma &\geq H(XY)_\sigma \\ &\geq H(\gamma_Q) + D(\gamma_Q \| \tilde{\gamma}_Q).\end{aligned}$$

□

6 Numerical Benchmark Codes and Results

In this section we present the benchmark code together with the corresponding numerical results. The code repository is available at <https://doi.org/10.5281/zenodo.17637006>.

We compare our bounds (13) and (19) with the well known entropic uncertainty relation of Berta *et al.* [BCC⁺10]:

$$H(\mathbf{M})_\gamma + H(\mathbf{N})_\gamma \geq -\max_{x,y} \left\| \sqrt{N_y} \sqrt{M_x} \right\|_\infty^2 + H(\gamma_Q),\tag{23}$$

where the measurements $\{M_x\}_x$ and $\{N_y\}_y$ are rank-one projective valued measurements (PVMs). Our bounds (13) and (19) apply to general POVMs; however, for a fair comparison with (23), we restrict to the rank-one PVM setting in this subsection.

For general POVMs, Eq. (23) does not hold, although it does apply whenever at least one of the POVMs has rank-one elements [CYGG11]. Motivated by this fact, we also perform numerical tests on more general rank-one POVMs.

The three bounds compared in the numerical experiments are:

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

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.

Negative gaps counts	$d = 2, n = 2$	$d = 3, n = 3$	$d = 4, n = 4$
EUR1 < EUR3	76325	44045	39426
EUR2 < EUR3	394	0	0
EUR2 < EUR1	0	0	0

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. A negative gap for “EUR1 < EUR3” means that the bound (13) is numerically weaker than the Berta *et al.* bound (23).

Negative gaps counts	$d = 2, n = 3$	$d = 2, n = 4$	$d = 3, n = 4$
EUR1 < EUR3	46123	44905	41855
EUR2 < EUR3	0	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 $n > d$.

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 (19) is consistently stronger than (23) for all cases with $n \geq 3$, and also always stronger than (13).

We emphasize, however, that in general bound (19) is not strictly comparable with the Berta *et al.* bound (23). Although Table 1 shows no negative gaps for $d \geq 3$ 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 behaviour. Random generation rarely produces such pairs in higher dimensions, which explains their 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 behaviour 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.

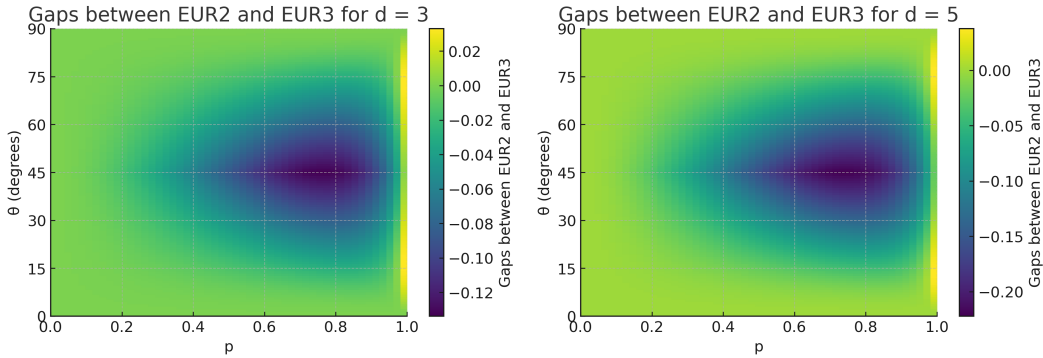


Figure 1: Negative gaps between EUR2 (19) and EUR3 (23) for mutually unbiased PVMs in dimensions $d = 3$ and $d = 5$. The shaded regions correspond to parameter ranges (p, θ) where EUR2 becomes weaker than EUR3. These counterexamples occur for states that interpolate between the two bases and illustrate that EUR2 is not uniformly stronger than the Berta *et al.* bound.

7 Conclusions

Starting from the quantum minimum change principle, we have shown that, for all quantum to classical channels, the information projection (6) onto fixed outcome statistics selects a single retrodictive update independent of the divergence used. This yields a canonical and divergence-independent quantum Bayesian inverse (7) 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 $R(\mathbf{M}; \mathbf{N})_\gamma$. Theorem 3 shows that this quantity is always finite and bounded above by $H(\gamma_Q)$, which implies that the mixedness of the prior is the sole quantity limiting the strength of backward inferences between measurements.

We then used the retrodictive joint distribution to derive two entropic uncertainty relations, Eqs. (13) and (19), 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.

We benchmarked our retrodictive uncertainty relations against the Berta *et al.* bound (23), 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. This contrast highlights a conceptual difference: our retrodictive bounds depend on the joint interplay between the state and the POVMs, whereas the Berta *et al.* relation does not.

More broadly, our results show that the minimum change principle provides a robust and quantitatively strong framework for backward inference in quantum mechanics. They also suggest several directions for further work, including extensions to memory-assisted or sequential measurement scenarios and retrodictive formulations of operational tasks such as quantum cryptography and quantum hypothesis testing.

Acknowledgements

The 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 Grant No. 23K03230. J. K. and K.T. acknowledge support from JST SPRING Grant No. JPMJSP2125.

References

- [BBS25] Ge Bai, Francesco Buscemi, and Valerio Scarani. Quantum Bayes’ rule and Petz transpose map from the minimum change principle. *Physical Review Letters*, 135:090203, 2025. [doi:10.1103/5n4p-bxhm](https://doi.org/10.1103/5n4p-bxhm).

- [BCC⁺10] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M Renes, and Renato Renner. The uncertainty principle in the presence of quantum memory. *Nature Physics*, 6(9):659–662, 2010. doi:10.1038/nphys1734.
- [BDW16] Francesco Buscemi, Siddhartha Das, and Mark M. Wilde. Approximate reversibility in the context of entropy gain, information gain, and complete positivity. *Physical Review A*, 93:062314, 2016. doi:10.1103/PhysRevA.93.062314.
- [BHH08] Francesco Buscemi, Masahito Hayashi, and Michał Horodecki. Global information balance in quantum measurements. *Physical Review Letters*, 100(21):210504, 2008. doi:10.1103/PhysRevLett.100.210504.
- [BJP21] Stephen M. Barnett, John Jeffers, and David T. Pegg. Quantum retrodiction: Foundations and controversies. *Symmetry*, 13(4), 2021. doi:10.3390/sym13040586.
- [BKM⁺23] Francesco Buscemi, Kodai Kobayashi, Shintaro Minagawa, Paolo Perinotti, and Alessandro Tosini. Unifying different notions of quantum incompatibility into a strict hierarchy of resource theories of communication. *Quantum*, 7:1035, 2023. doi:10.22331/q-2023-06-07-1035.
- [BPJ00] Stephen M. Barnett, David T. Pegg, and John Jeffers. Bayes’ theorem and quantum retrodiction. *Journal of Modern Optics*, 47(11):1779–1789, 2000. doi:10.1080/09500340008232431.
- [BS82] Viacheslav P Belavkin and P Staszewski. C^* -algebraic generalization of relative entropy and entropy. In *Annales de l’IHP Physique théorique*, volume 37, pages 51–58, 1982.
- [BSŠ23] Francesco Buscemi, Joseph Schindler, and Dominik Šafránek. Observational entropy, coarse-grained states, and the petz recovery map: information-theoretic properties and bounds. *New Journal of Physics*, 25(5):053002, 2023. doi:10.1088/1367-2630/accd11.
- [CG06] Ariel Caticha and Adom Giffin. Updating probabilities. *AIP Conference Proceedings*, 872(1):31–42, 2006. doi:10.1063/1.2423258.
- [Cho74] Man-Duen Choi. A Schwarz inequality for positive linear maps on C^* -algebras. *Illinois Journal of Mathematics*, 18(4):565–574, 1974.
- [Cox61] R.T. Cox. *Algebra of Probable Inference*. Johns Hopkins University Press, 1961. URL: <https://books.google.it/books?id=dcNpAUU6ACgC>.
- [CYGG11] Patrick J Coles, Li Yu, Vlad Gheorghiu, and Robert B Griffiths. Information-theoretic treatment of tripartite systems and quantum channels. *Physical Review A*, 83(6):062338, 2011. doi:10.1103/PhysRevA.83.062338.
- [DF74] B. De Finetti. *Theory of Probability*. Wiley series in probability and mathematical statistics. Wiley, 1974. URL: <https://books.google.it/books?id=Xtj1mgEACAAJ>.
- [Fuc03] Christopher A. Fuchs. Quantum mechanics as quantum information, mostly. *Journal of Modern Optics*, 50(6-7):987–1023, 2003. doi:10.1080/09500340308234548.
- [Gro71] Hilbrand J Groenewold. A problem of information gain by quantal measurements. *International Journal of Theoretical Physics*, 4(5):327–338, 1971. doi:10.1007/BF00815357.
- [Jac19] Bart Jacobs. The mathematics of changing one’s mind, via Jeffrey’s or via Pearl’s update rule. *Journal of Artificial Intelligence Research*, 65:783–806, 2019. doi:10.1613/jair.1.11349.
- [Jay57] Edwin T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620, 1957.
- [Jay03] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003. doi:10.1017/CB09780511790423.
- [Jef65] R.C. Jeffrey. *The logic of decision*. McGraw-Hill, 1965.

- [Jef98] H. Jeffreys. *The Theory of Probability*. Oxford Classic Texts in the Physical Sciences. OUP Oxford, 1998. URL: <https://books.google.co.jp/books?id=vh9Act9rtzQC>.
- [JOB24] John Jeffers, Daniel K. L. Oi, and Thomas Brougham. Quantum retrodiction. *Philosophical Transactions of the Royal Society A*, 382:20230338, 2024. doi:10.1098/rsta.2023.0338.
- [Joh07] Lars M. Johansen. Quantum theory of successive projective measurements. *Physical Review A*, 76:012119, 2007. doi:10.1103/PhysRevA.76.012119.
- [Lap86] Pierre Simon Laplace. Memoir on the probability of the causes of events. *Statistical Science*, 1(3):364–378, 1986. URL: <http://www.jstor.org/stable/2245476>.
- [Lei06] M. S. Leifer. Quantum dynamics as an analog of conditional probability. *Physical Review A*, 74:042310, 2006. doi:10.1103/PhysRevA.74.042310.
- [Lei07] Matthew S Leifer. Conditional density operators and the subjectivity of quantum operations. In *AIP Conference Proceedings*, volume 889, pages 172–186. American Institute of Physics, 2007.
- [LS13] M. S. Leifer and Robert W. Spekkens. Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference. *Physical Review A*, 88:052130, 2013. doi:10.1103/PhysRevA.88.052130.
- [Mat15] Keiji Matsumoto. A new quantum version of f-divergence. In *Nagoya Winter Workshop: Reality and Measurement in Algebraic Quantum Theory*, pages 229–273. Springer, 2015.
- [MLDS⁺13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013. doi:10.1063/1.4838856.
- [NKWB24] Teruaki Nagasawa, Kohtaro Kato, Eyuri Wakakuwa, and Francesco Buscemi. Generic increase of observational entropy in isolated systems. *Physical Review Research*, 6:043327, 2024. doi:10.1103/PhysRevResearch.6.043327.
- [NWKB25] Teruaki Nagasawa, Eyuri Wakakuwa, Kohtaro Kato, and Francesco Buscemi. Macroscopicity and observational deficit in states, operations, and correlations. *Reports on Progress in Physics*, 88:117601, 2025. doi:10.1088/1361-6633/ae140e.
- [Oza86] Masanao Ozawa. On information gain by quantum measurements of continuous observables. *Journal of mathematical physics*, 27(3):759–763, 1986. doi:10.1063/1.527179.
- [PB23] Arthur J. Parzygnat and Francesco Buscemi. Axioms for retrodiction: achieving time-reversal symmetry with a prior. *Quantum*, 7:1013, 2023. doi:10.22331/q-2023-05-23-1013.
- [Pet85] Dénes Petz. Quasi-entropies for states of a von Neumann algebra. *Publications of the Research Institute for Mathematical Sciences*, 21(4):787–800, 1985. doi:10.2977/prims/1195178929.
- [Pet86a] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports on mathematical physics*, 23(1):57–65, 1986. doi:10.1016/0034-4877(86)90067-4.
- [Pet86b] Dénes Petz. Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Communications in mathematical physics*, 105(1):123–131, 1986. doi:10.1007/BF01212345.
- [Pet88] Dénes Petz. Sufficiency of channels over von Neumann algebras. *The Quarterly Journal of Mathematics*, 39(1):97–108, 1988. doi:10.1093/qmath/39.1.97.
- [Pó154] G. Pólya. *Induction and Analogy in Mathematics*. Induction and Analogy in Mathematics. Princeton University Press, 1954. URL: https://books.google.it/books?id=0mgS_LDFaW8C.

- [PR98] Dénes Petz and Mary Beth Ruskai. Contraction of generalized relative entropy under stochastic mappings on matrices. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 01(01):83–89, 1998. doi:[10.1142/S0219025798000077](https://doi.org/10.1142/S0219025798000077).
- [Sei86] Teddy Seidenfeld. Entropy and uncertainty. *Philosophy of Science*, 53(4):467–491, 1986. doi:[10.1086/289336](https://doi.org/10.1086/289336).
- [Uff96] Jos Uffink. The constraint rule of the maximum entropy principle. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 27(1):47–79, 1996. doi:[10.1016/1355-2198\(95\)00022-4](https://doi.org/10.1016/1355-2198(95)00022-4).
- [Ume61] Hisaharu Umegaki. On information in operator algebras. *Proceedings of the Japan Academy*, 37(8):459–461, 1961. doi:[10.3792/pja/1195523632](https://doi.org/10.3792/pja/1195523632).
- [Wat55] Satoshi Watanabe. Symmetry of Physical Laws. Part III. Prediction and Retrodiction. *Reviews of Modern Physics*, 27:179–186, 1955. doi:[10.1103/RevModPhys.27.179](https://doi.org/10.1103/RevModPhys.27.179).
- [Wil17] Mark M Wilde. *Quantum Information Theory, 2nd edition*. Cambridge University Press, 2017. doi:[10.1017/9781316809976](https://doi.org/10.1017/9781316809976).
- [WWY14] Mark M Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, 2014. doi:[10.1007/s00220-014-2122-x](https://doi.org/10.1007/s00220-014-2122-x).
- [Zel88] Arnold Zellner. Optimal information processing and Bayes’s theorem. *The American Statistician*, 42(4):278–280, 1988.

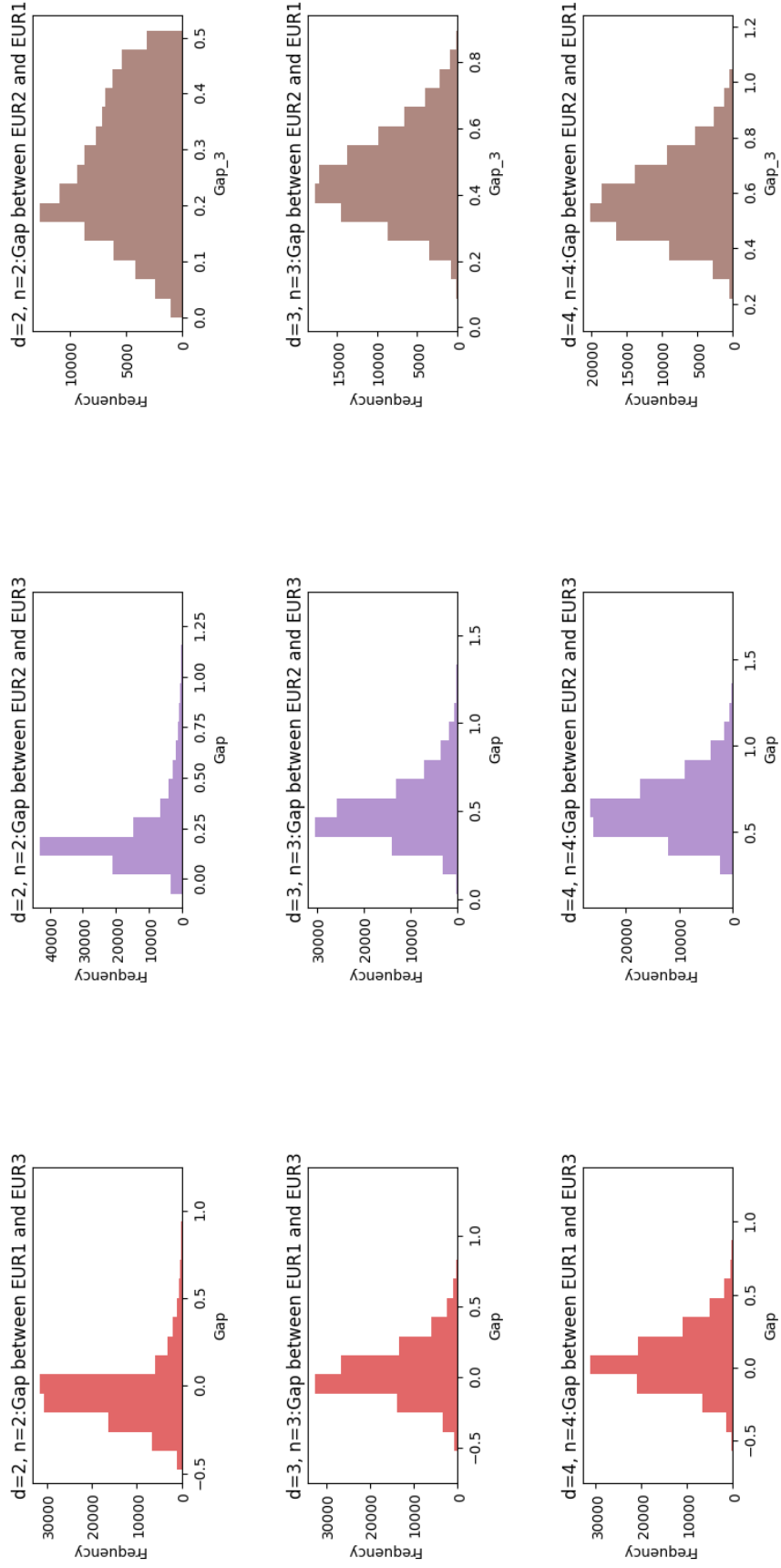


Figure 2: Numerical comparison of the bounds EUR1 (13), EUR2 (19), and EUR3 (23) for random rank-one PVMs. Each point corresponds to one of 100,000 random trials. The plot illustrates that EUR2 is consistently tighter than the Berta *et al.* bound for $d \geq 3$, while EUR1 shows no uniform ordering relative to EUR3.

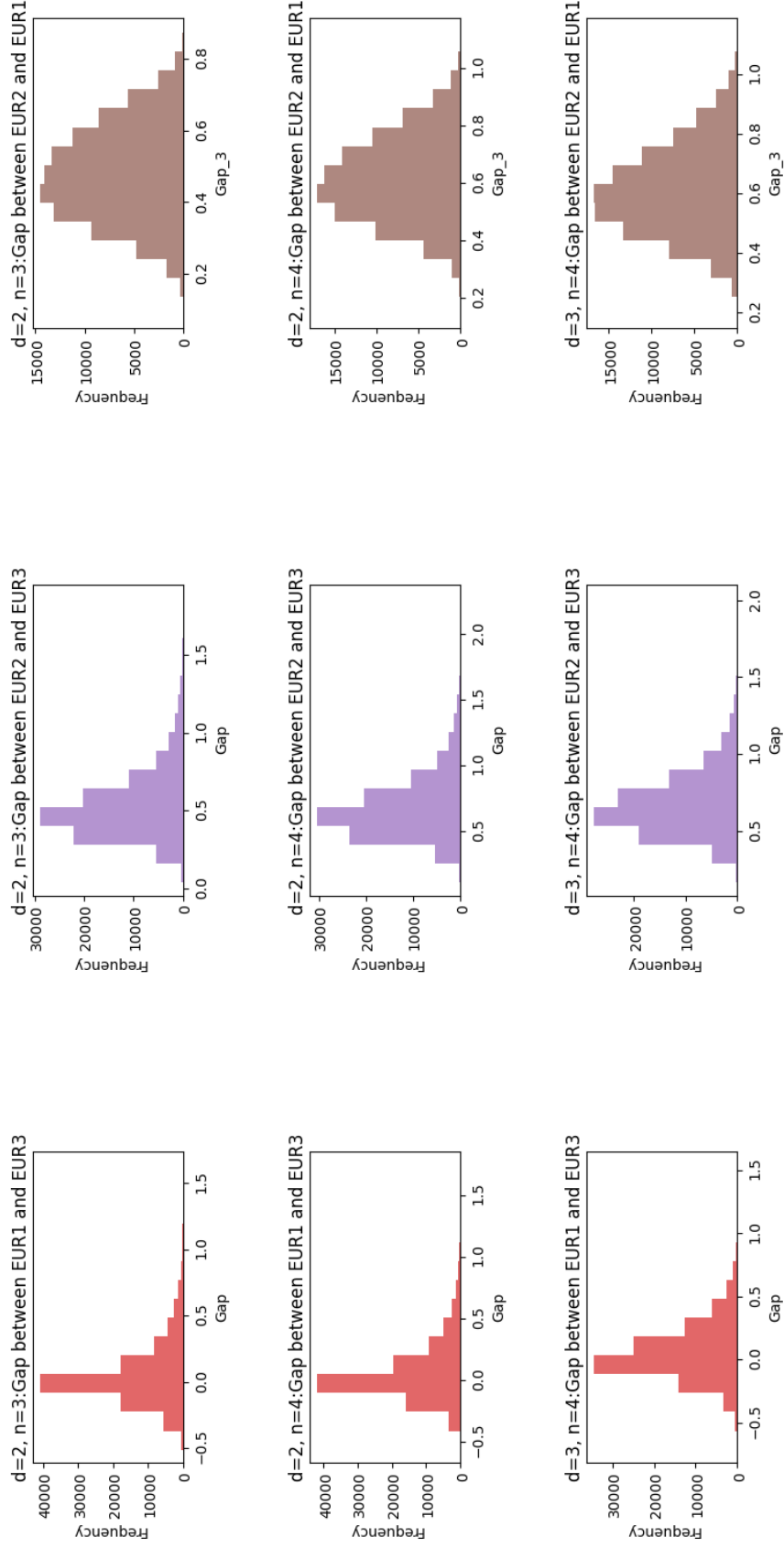


Figure 3: Numerical comparison of EUR1, EUR2, and EUR3 for random rank-one POVMs in various dimensions. The data show the same trend as in the projective case: EUR2 dominates EUR3 in nearly all instances, while EUR1 exhibits mixed behaviour. The increased variety of POVMs broadens the range of achievable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