Observational entropy with general quantum priors

Ge Bai[1], Dominik Šafránek[2,3], Joseph Schindler[4], Francesco Buscemi[5], and Valerio Scarani[1,6]

[1]Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543
[2]Center for Theoretical Physics of Complex Systems, Institute for Basic Science (IBS), Daejeon 34126, Korea
[3]Basic Science Program, Korea University of Science and Technology (UST), Daejeon - 34113, Korea
[4]Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[5]Department of Mathematical Informatics, Nagoya University, Furo-cho Chikusa-ku, Nagoya 464-8601, Japan
[6]Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542

arXiv:2308.08763v3 [quant-ph] 11 Nov 2024

Observational entropy captures both the intrinsic uncertainty of a thermodynamic state and the lack of knowledge due to coarse-graining. We demonstrate two interpretations of observational entropy, one as the statistical deficiency resulting from a measurement, the other as the difficulty of inferring the input state from the measurement statistics by quantum Bayesian retrodiction. These interpretations show that the observational entropy implicitly includes a uniform reference prior. Since the uniform prior cannot be used when the system is infinite-dimensional or otherwise energy-constrained, we propose generalizations by replacing the uniform prior with arbitrary quantum states that may not even commute with the state of the system. We propose three candidates for this generalization, discuss their properties, and show that one of them gives a unified expression that relates both interpretations.

1 Introduction

A few pages after defining the entropy that nowadays bears his name, von Neumann warns the reader that the quantity that he just defined is, in fact, unable to capture the phenomenological behavior of thermodynamic entropy [1]. More precisely, while the von Neumann entropy $S(\rho) := -\operatorname{Tr}[\rho \ln \rho]$ is always invariant in a closed system as a consequence of its invariance under unitary evolutions, the thermodynamic entropy of a closed system can instead increase, as it happens for example in the free expansion of an ideal gas. The explanation that von Neumann gives for this apparent paradox is the following: thermodynamic entropy includes not only the intrinsic ignorance associated with the microscopic state ρ of the system, but also the lack of knowledge arising from a macroscopic coarse-graining of it. The latter lack of knowledge becomes worse as the gas expands. This observation leads him to introduce an alternative quantity, that he calls *macroscopic entropy*, for which an H-theorem can be proved [2].

In recent years, von Neumann's macroscopic entropy and a generalization thereof called *observational entropy* (OE) has been the object of renewed interest [3–8], finding a number of applications [9–21]. So far, even when the narrative is based on a quantum state ρ being subject to a measurement M, all the definitions fit in classical stochastic thermodynamics.

In this paper, we explore possible generalizations of OE. We note that the original OE includes an implicit prior belief about the state, which is the uniform distribution. Since in several applications the uniform prior cannot be used, e.g., in infinite-dimensional or continuous variable systems, or does not play well with other physical constraints, e.g., in thermodynamic systems with a nondegenerate Hamiltonian at finite temperature, we allow the observer to have a non-uniform prior. More generally, we consider the possibility that the observer has a reference prior described by an arbitrary density operator, which may not even commute with the state of the system. In this case, classical probability distribu-

Francesco Buscemi: buscemi@nagoya-u.jp
Valerio Scarani: physv@nus.edu.sg

tions may not be sufficient to describe the non-commutativity between the state and the reference, and thus the original definition of OE is not applicable.

2 Classical OE and reference states

In what follows, we restrict our attention to finite-dimensional quantum systems, with Hilbert space $\mathbb{C}^d$, and finite measurements, i.e., positive operator-valued measures (POVMs) $\mathsf{M} = \{\Pi_y\}_y$ labeled by the elements of a finite set $y \in \{1, \dots, m\}$. In this context, the definition of OE is

$$S_{\mathsf{M}}(\rho) \coloneqq -\sum_{y=1}^{m} p_y \ln \frac{p_y}{V_y} \,, \tag{1}$$

where $p_y \coloneqq \mathrm{Tr}[\rho\Pi_y]$ and $V_y \coloneqq \mathrm{Tr}[\Pi_y]$. One of the conceptual advantages of OE is that it is able to "interpolate" between Boltzmann and Gibbs–Shannon entropies. On the one hand, if the measurement is so coarse-grained that one of its elements (say Π_1) is the projector on the support of ρ, then $S_{\mathsf{M}}(\rho) = \ln V_1$ takes the form of a Boltzmann entropy. If, on the other hand, the measurement is projective and rank-one (i.e., $V_y = 1$ for all y), then $S_{\mathsf{M}}(\rho)$ coincides with the Shannon entropy of the probability distribution $\{p_y\}$, which is equal to $S(\rho)$ when $\rho = \sum_y p_y \Pi_y$.

In general, it holds that [4]

$$\Sigma_{\mathsf{M}}(\rho) \coloneqq S_{\mathsf{M}}(\rho) - S(\rho) \geq 0 \,. \tag{2}$$

If $S(\rho)$ represents, in von Neumann's original narrative, the *least* uncertainty that an observer, able to perform *any* measurement in principle allowed by quantum theory, has about the state of the system, then the additional uncertainty $\Sigma_{\mathsf{M}}(\rho)$ included in OE is a consequence of observing the system through the "lens" provided by the given measurement M. Thus, in this sense, OE can be seen as a measure of how inadequate a given measurement M is with respect to the state ρ.

2.1 OE from statistical deficiency

The above discussion suggests one possible generalization of OE, starting from the re-writing of (2) recently noticed by some of us [8]. Consider the *measurement channel* $\mathcal{M}$ associated to the measurement M, defined as

$$\mathcal{M}(\rho) \coloneqq \sum_y \mathrm{Tr}[\Pi_y \rho] |y\rangle\langle y| \,, \tag{3}$$

where $\{|y\rangle\}$ is an arbitrary but fixed orthonormal basis of the system that records the measurement outcome. By further noticing that $V_y = d\,\mathrm{Tr}[\Pi_y u]$ with $u = \mathbb{1}/d$ the maximally mixed state, one obtains

$$\Sigma_{\mathsf{M}}(\rho) = D(\rho\|u) - D(\mathcal{M}(\rho)\|\mathcal{M}(u)) \,, \tag{4}$$

where

$$D(\rho\|\sigma) \coloneqq \mathrm{Tr}[\rho(\ln\rho - \ln\sigma)] \tag{5}$$

is the Umegaki quantum relative entropy between states ρ and σ [22, 23], which generalizes the relative entropy (a.k.a. Kullback-Leibler divergence)

$$D(p\|q) \coloneqq \sum_i p_i \ln \frac{p_i}{q_i} \tag{6}$$

between probability distributions $\{p_i\}$ and $\{q_i\}$ [24].

The expression (4) makes it clear that the quantity $\Sigma_{\mathsf{M}}(\rho)$ exactly equals the loss of distinguishability between the signal ρ and the totally uniform background u that occurs when the measurement M is used instead of the best possible measurement allowed by quantum theory. In statistical jargon, we thus say that $\Sigma_{\mathsf{M}}(\rho)$ measures the *statistical deficiency* of the measurement M in distinguishing ρ against u.

This observation enlightens something implicit in the original definition (1) of OE: the coarse-graining is captured by the "volumes" $V_y = \mathrm{Tr}[\Pi_y]$ only because the maximally mixed state is chosen as the reference background. It is thus natural to try to incorporate more general references in the definition of OE. A direct generalization could be obtained, therefore, by replacing u with another reference state γ in (4), so that

$$\begin{aligned} S_{\mathsf{M},\gamma}(\rho) &\coloneqq S(\rho) + \Sigma_{\mathsf{M},\gamma}(\rho) \\ &\coloneqq S(\rho) + \mathfrak{D}(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \,, \end{aligned} \tag{7}$$

where $\mathfrak{D}(\cdot\|\cdot)$ represents some non-commutative generalization of the Kullback–Leibler divergence, not necessarily Umegaki's one.

2.2 OE from irretrodictability

There exists another evocative re-writing of (2), which in turn suggests that further structures may play a role in the definition of OE. Specifically, here we exhibit a *dynamical* interpretation of OE, based on a measurement process defined as follows.

Let $\rho = \sum_{x=1}^{d} \lambda_x |\psi_x\rangle\langle\psi_x|$ be a diagonal decomposition of the state of the system. We consider a stochastic map associated to a prepare-and-measure protocol: with probability λ_x, the state $|\psi_x\rangle$ is prepared, and it is then measured with the POVM $\mathsf{M} = \{\Pi_y\}_y$, yielding outcome y with a probability given by the Born rule, that is

$$P_F(y|x) \coloneqq \langle\psi_x| \Pi_y |\psi_x\rangle \ , \tag{8}$$

$$P_F(x,y) \coloneqq \lambda_x \langle\psi_x| \Pi_y |\psi_x\rangle \ . \tag{9}$$

The subscript F stands for "forward". This is because, as we will see in what follows, the quantity $\Sigma_{\mathsf{M}}(\rho)$ in (2) emerges also from a comparison between the forward map defined above and a suitably defined "reverse" map.

Traditionally, the definition of the reverse process, given a forward process, relies on a detailed knowledge of the physical dynamics involved [25–27]. In the absence of such knowledge, which is typically the case for a system interacting with a complex environment, one must resort to physical intuition, plausibility arguments, or, failing that, arbitrary assumptions. In order to avoid all this, a systematic recipe has recently been found [28, 29], which allows to define the reverse process only from unavoidable rules of *logical retrodiction*: specifically, Jeffrey's theory of probability kinematics [30] or, equivalently, Pearl's virtual evidence method [31, 32].

The idea is as follows: given a forward conditional probability $P_F(y|x)$, how should information, obtained at later times about the final outcome and encoded in an arbitrary distribution q_y, be propagated back to the initial state in a way that ensures logical consistency? Jeffrey's theory of probability kinematics, which is equivalent to Pearl's virtual evidence method [31, 32], stipulates that the only logically consistent back-propagation rule is what is now known as *Jeffrey's update*: starting from an arbitrarily chosen reference prior on the initial state x, say γ_x, one constructs the Bayesian inverse of $P_F(y|x)$, i.e.,

$$P_R^{\gamma}(x|y) \coloneqq \frac{\gamma_x P_F(y|x)}{\sum_{x'} \gamma_{x'} P_F(y|x')} \ , \tag{10}$$

and uses that as a *stochastic channel* to back-propagate the new information q_y from the final outcome y to the initial state x, so that as the reverse process we obtain

$$P_R^{\gamma}(y,x) = q_y \, P_R^{\gamma}(x|y) \ . \tag{11}$$

Jeffrey's update constitutes a generalization of Bayes' theorem, as the latter is recovered as a special case of the former when $q_y = \delta_{y,y_0}$, i.e., when the information about the final outcome is definite [32].

An important point to emphasize here is that the reference prior γ_x used to construct the retrodictive channel in (10) is merely a formal device needed to establish a mathematical correspondence between forward and backward process: it need not be related in any way with the "true" distribution λ_x. Likewise, the distribution q_y represents *new and completely arbitrary* information, which need not correspond to any input distribution under P_F, that is, there may exist no distribution q'_x such that $q_y = \sum_x P_F(y|x) q'_x$. Moreover, in principle, q_y may also be incompatible with the reference prior γ_x, in the sense that it could happen that, for some y, $q_y > 0$ but $\sum_x \gamma_x P_F(y|x) = 0$. In such a situation, one would conclude that the data *falsify* the inferential model, but for simplicity we will avoid such cases by assuming that all probabilities are strictly greater than zero (though possibly arbitrarily small).

We now go back to our specific forward process (9), i.e., $P_F(y|x) = \langle\psi_x| \Pi_y |\psi_x\rangle$. If we choose as reference the uniform distribution $\gamma = u$, i.e., $\gamma_x = 1/d$ for all x, and as new information information the outcomes' expected probability of occurrence, i.e., $q_y = p_y = \mathrm{Tr}[\rho\Pi_y]$, by direct substitution in (10) and (11), we obtain

$$P_R^{u}(y,x) = p_y \langle\psi_x| \frac{\Pi_y}{\mathrm{Tr}[\Pi_y]} |\psi_x\rangle \ . \tag{12}$$

The above can also be read as a prepare-and-measure process, in which the state $\sigma_y \coloneqq \frac{\Pi_y}{\mathrm{Tr}[\Pi_y]}$ is prepared with probability p_y, and later measured in the basis $\{|\psi_x\rangle\}$. The process in (12) is the process that a retrodictive agent would infer, knowing only the forward process (9) and the

outcome distribution p_y, but completely ignoring the actual distribution λ_x, so that the latter is replaced by the uniform distribution.

Using (9) and (12), it is straightforward to check that

$$\Sigma_{\mathsf{M}}(\rho) = D(P_F \| P_R^u)\,. \tag{13}$$

The above relation suggests an alternative interpretation for the difference $\Sigma_{\mathsf{M}}(\rho)$, as the degree of statistical distinguishability between a predictive process, i.e., $P_F(x,y)$, and a retrodictive process constructed from a uniform reference, i.e., $P_R^u(y,x)$. Thus, the larger $\Sigma_{\mathsf{M}}(\rho)$, the more *irretrodictable* the process becomes [33, 34].

Eq. (13) also offers an alternative way of thinking about generalizations of OE, where the uniform reference is again replaced by an arbitrary state, as was done in Section 2.1, but this time for the purpose of constructing another reverse process. That is, one could also consider generalizations such as

$$\widetilde{S}_{\mathsf{M},\gamma}(\rho) \coloneqq S(\rho) + \mathfrak{D}(Q_F \| Q_R^\gamma)\,, \tag{14}$$

where $\mathfrak{D}$ is again some quantum relative entropy (not necessarily Umegaki's), Q_F is an input-output description of the quantum process consisting in preparing the state ρ and measuring it with the channel $\mathcal{M}$, and Q_R^γ is the description of the corresponding reverse process computed with respect to the reference prior γ. All these ingredients will be rigorously defined in Section 4.

3 A definition of OE for priors γ such that $[\rho,\gamma]=0$

As we have seen, in the case of conventional OE, the statistical deficiency approach and the irretrodictability approach coincide, i.e.

$$\Sigma_{\mathsf{M}}(\rho) = D(\rho\|u) - D(\mathcal{M}(\rho)\|\mathcal{M}(u)) = D(P_F\|P_R^u)\,.$$

The same holds for any prior γ that commutes with ρ. Indeed, assuming $[\rho,\gamma]=0$, let us write $\gamma = \sum_{x=1}^d \gamma_x |\psi_x\rangle\langle\psi_x|$ using the same vectors that diagonalize ρ. The reverse process of P_F [Eq. (9)] becomes

$$\begin{aligned} P_R^\gamma(x|y) &\coloneqq \frac{\gamma_x \langle\psi_x|\,\Pi_y\,|\psi_x\rangle}{\mathrm{Tr}[\Pi_y\gamma]} \\ P_R^\gamma(x,y) &= p_y P_R^\gamma(x|y)\,, \end{aligned} \tag{15}$$

and it is straightforward to verify that

$$D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) = D(P_F||P_R^\gamma)\,. \tag{16}$$

Therefore, when $[\rho,\gamma]=0$, the expression

$$\begin{aligned} &S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho) \\ &\coloneqq\ S(\rho) + D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) && (17)\\ &=\ S(\rho) + D(P_F||P_R^\gamma) && (18)\\ &=\ -\mathrm{Tr}[\rho\ \ln\gamma] - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))\,. && (19) \end{aligned}$$

is a generalized OE that fits both the statistical deficiency approach and the irretrodictability approach. As we are going to see, it is not obvious to ensure both interpretations when $[\rho,\gamma]\neq 0$.

Note that nothing has been said about the measurement M, which may well not commute with either ρ or γ. As a case study, let us look closely to the *fully classical case* in which, besides having $[\rho,\gamma]=0$, the measurement is a projective measurement in their same eigenbasis $\{|k\rangle\}$:

$$\mathsf{M} = \{\Pi_y\}_y \quad \text{with} \quad \Pi_y = \sum_{k\in K(y)} |k\rangle\langle k| \tag{20}$$

where the index sets $K(y)$ are disjoint and complete as to form a POVM, i.e. $\bigcup_y K(y) = 1,...,d$, and $K(y)\cap K(y') = \emptyset$ for $y\neq y'$. Then, denoting by $\{r_k\}$ and $\{g_k\}$ the eigenvalues of ρ and γ respectively, we have $p_y = \sum_{k\in K(y)} r_k$ and $\mathrm{Tr}[\gamma\Pi_y] \equiv G_y = \sum_{k\in K(y)} g_k$, and Eq. (19) yields

$$S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho) = -\sum_k r_k \log g_k - D(\{p_y\}||\{G_y\})\,. \tag{21}$$

While the second term depends only on the observed statistics $\{p_y\}$ by construction, the first term depends in general on the full information $\{r_k\}$. In fact, $S^{\mathrm{clax}}_{\mathsf{M},\gamma}$ depends only on the $\{p_y\}$ if and only if the g_k are uniform in each $K(y)$ subspace; that is, $g_k = G_y/|K(y)|$ for every $k\in K(y)$ and for every y, or, equivalently, $\gamma = \sum_y G_y\Pi_y/|K(y)|$. In this case, $S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho) = S_{\mathsf{M}}(\rho)$: the weights G_y of the prior do not matter. The interpretation of this observation is clear: the observation gives precisely the weights p_y to be attributed to each Π_y, trumping any prior belief on those weights.

When the dependence on the $\{r_k\}$ is present, it is a mild one: for instance, in the paradigmatic case where γ is thermal, the term $-\sum_k r_k \log g_k$ is (up to an additive constant) the average energy,

often assumed as known in thermodynamics. A purely "observational" character could be recovered with minor modifications of the definition; we leave this aside.

4 Definitions of OE with an arbitrary quantum reference prior

In this section, we introduce the mathematical notations and backgrounds, and propose some candidates for OE with a general reference prior state.

4.1 Input-output description of quantum processes

Let A and B respectively denote the input and output systems of the measurement channel $\mathcal{M}$ in Eq. (3). The general recipe for the retrodiction of a quantum process $\mathcal{M}$ [8, 28, 35] is defined via its Petz recovery map [36, 37] as

$$\widetilde{\mathcal{M}}^{\gamma}(\tau) \coloneqq \gamma^{1/2}\mathcal{M}^{\dagger}(\mathcal{M}(\gamma)^{-1/2}\tau\mathcal{M}(\gamma)^{-1/2})\gamma^{1/2} \quad (22)$$

where $\tau := \sum_y q_y|y\rangle\langle y|$ encodes the distribution $\{q_y\}$, describing the retrodictor's knowledge, cfr. Eq. (11). We will mainly discuss the case where $\tau = \mathcal{M}(\rho)$, namely $q_y = p_y = \mathrm{Tr}[\Pi_y\rho]$. For the measurement channel given in (3), the Petz recovery map can be written as

$$\widetilde{\mathcal{M}}^{\gamma}(\tau) = \sum_y \frac{\langle y|\,\tau\,|y\rangle}{\mathrm{Tr}[\Pi_y\gamma]}\sqrt{\gamma}\Pi_y\sqrt{\gamma}\,. \quad (23)$$

As an ingredient for later constructions, we introduce the Choi operator [38], defined for the process $\mathcal{M}$ from system A to system B as

$$C_{\mathcal{M}} \coloneqq \sum_{i,j} \mathcal{M}(|i\rangle\langle j|) \otimes |i\rangle\langle j|\,, \quad (24)$$

where $|i\rangle$ and $|j\rangle$ belong to an arbitrary but fixed orthonormal basis of the input Hilbert space $\mathcal{H}_A$ of system A. The reverse process has the following Choi operator

$$C_{\widetilde{\mathcal{M}}^{\gamma}} \coloneqq \sum_{k,l} |k\rangle\langle l| \otimes \widetilde{\mathcal{M}}^{\gamma}(|k\rangle\langle l|)\,, \quad (25)$$

where $|k\rangle$ and $|l\rangle$ belong to an arbitrarily fixed orthonormal basis of the Hilbert space $\mathcal{H}_B$. Note that we put system B first and system A second, in order to have the same ordering of systems for both $C_{\mathcal{M}}$ and $C_{\widetilde{\mathcal{M}}^{\gamma}}$.

With such a definition, the Choi operators of the forward and reverse processes are related by the following lemma (proof in Appendix A):

Lemma 1. *For a quantum channel $\mathcal{M}$ and its Petz map $\widetilde{\mathcal{M}}^{\gamma}$, their Choi operators $C_{\mathcal{M}}$ and $C_{\widetilde{\mathcal{M}}^{\gamma}}$ are related as*

$$C^{T}_{\widetilde{\mathcal{M}}^{\gamma}} = \left(\mathcal{M}(\gamma)^{-1/2} \otimes \sqrt{\gamma^{T}}\right) C_{\mathcal{M}} \left(\mathcal{M}(\gamma)^{-1/2} \otimes \sqrt{\gamma^{T}}\right)\,, \quad (26)$$

where the superscript $\bullet^T$ denotes the matrix transposition done with respect to the fixed bases used in Eqs. (24) and (25).

We now want to construct two objects, Q_F and Q_R^{γ}, which, analogously to the joint distributions P_F and P_R^{γ}, are able to capture *both* the input *and* output of the forward and reverse processes. Specifically, the marginals of the operator Q_F should recover the input state ρ and the output $\mathcal{M}(\rho)$ respectively, and analogously for Q_R^{γ}.

One choice is to define

$$Q_F := (\mathbb{1}_B \otimes \sqrt{\rho^{T}})C_{\mathcal{M}}(\mathbb{1}_B \otimes \sqrt{\rho^{T}})\,. \quad (27)$$

Such an operator is indeed able to capture the input and output of the forward process, in the sense that:

$$\mathrm{Tr}_A[Q_F] = \mathcal{M}(\rho), \quad \mathrm{Tr}_B[Q_F] = \rho^{T}\,. \quad (28)$$

We define the representation for the reverse process $\widetilde{\mathcal{M}}^{\gamma}$ similarly as

$$Q_R^{\gamma} := (\sqrt{\tau} \otimes \mathbb{1}_A)\, C^{T}_{\widetilde{\mathcal{M}}^{\gamma}}\, (\sqrt{\tau} \otimes \mathbb{1}_A)\,, \quad (29)$$

where $\tau = \mathcal{M}(\rho)$ is the input of the reverse process. We use the transpose of the Choi operator of the reverse process so that it can be linked to $C_{\mathcal{M}}$ by Lemma 1 in the following way:

$$Q_R^{\gamma} = \left(\sqrt{\tau}\sqrt{\mathcal{M}(\gamma)}^{-1} \otimes \sqrt{\gamma^{T}}\right) C_{\mathcal{M}} \left(\sqrt{\mathcal{M}(\gamma)}^{-1}\sqrt{\tau} \otimes \sqrt{\gamma^{T}}\right)\,.$$

This operator captures the input and output of the reverse process:

$$\mathrm{Tr}_A[Q_R^{\gamma}] = \tau, \quad \mathrm{Tr}_B[Q_R^{\gamma}] = [\widetilde{\mathcal{M}}^{\gamma}(\tau)]^{T}\,. \quad (30)$$

The operators Q_F, Q_R^{γ} just defined are analogous to the state over time proposed by Leifer and Spekkens [39, 40] up to a partial transpose.

Other definitions of input-output operators may satisfy nice properties. An alternative choice is, for example,

$$ {}^tQ_F := \sqrt{C_{\mathcal{M}}}(\mathbb{1}_B \otimes \rho^T)\sqrt{C_{\mathcal{M}}} \tag{31} $$

and

$$ {}^tQ_R^\gamma := \sqrt{C_{\mathcal{M}}}(\mathcal{M}(\gamma)^{-1/2}\tau\mathcal{M}(\gamma)^{-1/2} \otimes \gamma^T)\sqrt{C_{\mathcal{M}}}\,. \tag{32} $$

The superscript ${}^t\bullet$ in (31) and (32) (not to be confused with $\bullet^T$) is used because the operators ${}^tQ_F$ and ${}^tQ_R^\gamma$ are, in a loose sense, a "transposition" of Q_F and Q_F^γ, respectively. If Π_y, ρ, γ do not commute, in general ${}^tQ_F \neq Q_F$ and ${}^tQ_R^\gamma \neq Q_R^\gamma$. For example, $\mathrm{Tr}_B[{}^tQ_F] = (\sum_y \sqrt{\Pi_y}\rho\sqrt{\Pi_y})^T$ which in general differs from ρ^T. Yet, they are similar, in the sense that Q_F and ${}^tQ_F$ (resp. Q_R^γ and ${}^tQ_R^\gamma$) share the same eigenvalues, and are thus unitarily equivalent, as it happens when doing a proper transposition. Therefore, ${}^tQ_F$ and ${}^tQ_R^\gamma$ can be viewed as legitimate representations (up to unitaries) of the forward and reverse processes, and they will be useful in the irretrodictability interpretation of OE.

4.2 Candidates for generalized OE

Eqs. (7) and (14) provide two forms of the observational entropy: Eq. (7), arising from the statistical deficiency approach, is the difference between relative entropies evaluated on the input system and the output system; Eq. (14), arising from the irretrodictability approach, is the relative entropy between the forward and reverse processes. In the remainder of this section, we will propose generalizations of OE that take either or both of these forms.

4.2.1 Candidate #1: difference between input/output Umegaki entropies

A first fully quantum generalisation of OE may just be obtained by replacing the reference state u in Eq. (4) with a general reference state γ:

$$ S^{(1)}_{\mathsf{M},\gamma}(\rho) := S(\rho) + \Sigma^{(1)}_{\mathsf{M},\gamma}(\rho) \quad \text{with } \Sigma^{(1)}_{\mathsf{M},\gamma}(\rho) = D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))\,, \tag{33} $$

that is, $S^{(1)}_{\mathsf{M},\gamma}(\rho) = -\mathrm{Tr}[\rho\ \ln\gamma] - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$, cf. Eq. (19), though this time it may be that $[\rho,\gamma] \neq 0$. This definition has the form of Eq. (7) with $\mathfrak{D}$ taken to be the Umegaki relative entropy (5). Notice that, while $D(\rho\|\gamma)$ is a fully quantum relative entropy, $D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$ is in fact classical, since all the outputs of the channel $\mathcal{M}$ are diagonal in the same basis.

4.2.2 Candidate #2: Umegaki relative entropy between forward/reverse processes

Another option is to define a OE through Eq. (14), thus choosing to prioritize irretrodictability. For this, one needs to choose a relative entropy and representations of the forward and reverse processes. Using the Umegaki relative entropy and the representations defined in (27) and (29), we get

$$ S^{(2)}_{\mathsf{M},\gamma}(\rho) := S(\rho) + \Sigma^{(2)}_{\mathsf{M},\gamma}(\rho) \quad \text{with } \Sigma^{(2)}_{\mathsf{M},\gamma}(\rho) = D(Q_F\|Q_R^\gamma). \tag{34} $$

However, we will show in the following sections that this candidate lacks some of the properties we desire: we introduced it mainly for comparison with other candidates.

4.2.3 Candidate #3: Belavkin–Staszewski relative entropy

Besides Umegaki relative entropy, there are other choices for the quantum relative entropy between the representations of the forward and reverse processes, and between the states ρ and γ. One such choice is the Belavkin–Staszewski relative entropy [41], defined as

$$ D_{\mathrm{BS}}(\rho\|\sigma) := \mathrm{Tr}[\rho\ \ln\rho\sigma^{-1}]. \tag{35} $$

The Belavkin–Staszewski relative entropy coincides with the Umegaki relative entropy and the classical relative entropy when ρ and σ commute, otherwise in general it is never smaller than Umegaki's. For a summary of the main properties of Belavkin–Staszewski relative entropy, and its relations with other quantum relative entropies, we refer the interested reader to Ref. [42].

Inserting D_{BS} into Eq. (7), we obtain

$$ S^{(3)}_{\mathsf{M},\gamma}(\rho) := S(\rho) + \Sigma^{(3)}_{\mathsf{M},\gamma}(\rho) \quad \text{with } \Sigma^{(3)}_{\mathsf{M},\gamma}(\rho) = D_{\mathrm{BS}}(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))\,. \tag{36} $$

Remarkably, it turns out that the above definition recovers the form of Eq. (14). Assuming that Q_R^γ,

and thus ${}^tQ_R^\gamma$, is full-rank, one has

$$D_{\mathrm{BS}}(\rho\|\gamma) - D_{\mathrm{BS}}(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) = D_{\mathrm{BS}}\left({}^tQ_F\middle\|{}^tQ_R^\gamma\right), \tag{37}$$

where $D_{\mathrm{BS}}(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) = \mathcal{D}(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$ since those states commute, and where ${}^tQ_F$ and ${}^tQ_R^\gamma$ were defined in (31) and (32). The proof of the identity (37) is given in Appendix B. Thus, $\Sigma^{(3)}_{\mathsf{M},\gamma}$ indeed admits both the statistical deficiency and the irretrodictability interpretations.

5 Properties

We proceed now to discuss the properties of the three candidates $S^{(j)}_{\mathsf{M},\gamma}$ $(j = 1, 2, 3)$ defined above, with a comparison between them summarized in Table 1. The main properties to consider for any candidate generalized OE are the following:

(i) When the reference prior is the uniform distribution (maximally mixed state), the candidate should recover the original definition (1). This is true for $S^{(1)}_{\mathsf{M},\gamma}$ and $S^{(3)}_{\mathsf{M},\gamma}$, namely when $\gamma = u := \mathbb{1}/d$,

$$S^{(1,3)}_{\mathsf{M},u}(\rho) = S_{\mathsf{M}}(\rho)\,. \tag{38}$$

Instead, in order to recover the conventional OE, $S^{(2)}_{\mathsf{M},\gamma}$ further requires that $[\rho, \Pi_y] = 0$ for all y.

(ii) More generally, when $[\rho, \gamma] = 0$, one has

$$S^{(1,3)}_{\mathsf{M},\gamma}(\rho) = S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho)\,. \tag{39}$$

Instead, the condition $S^{(2)}_{\mathsf{M},\gamma}(\rho) = S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho)$ in general requires $[\rho, \gamma] = [\rho, \Pi_y] = [\gamma, \Pi_y] = 0$ for all y.

(iii) Like the original OE, all of them are lower-bounded by the von Neumann entropy:

$$S^{(j)}_{\mathsf{M},\gamma}(\rho) \geq S(\rho)\,. \tag{40}$$

Thus, the OEs retain the desirable property that one cannot have less uncertainty than the von Neumann entropy.

The proofs of the above properties are in Appendix C.

Other non-essential, yet desirable properties include:

(iv) $S^{(j)}_{\mathsf{M},\gamma}(\rho)$ admits both interpretations, as statistical deficiency, i.e. Eq. (7), and irretrodictability, i.e., Eq. (14). This property is satisfied by $S^{(3)}_{\mathsf{M},\gamma}$, with suitable definitions of the input-output descriptions.

(v) $S^{(j)}_{\mathsf{M},\gamma}(\rho)$ satisfies the Petz recovery criterion: $S^{(j)}_{\mathsf{M},\gamma}(\rho) = S(\rho)$ if and only if $\widetilde{\mathcal{M}}^\gamma(\mathcal{M}(\rho)) = \rho$, where $\widetilde{\mathcal{M}}^\gamma$ is the Petz map of $\mathcal{M}$ with reference γ defined in (22). This property is satisfied by all candidates, as shown in Appendix D.

(vi) $S^{(j)}_{\mathsf{M},\gamma}(\rho)$ is non-decreasing under stochastic post-processing. We say $\mathsf{M}' = \{\Pi'_z\}$ is a post-processing of M if its outcome can be obtained by applying a stochastic map on the outcome of M, namely there exists a stochastic matrix w with $\sum_z w_{zy} = 1$ for all y satisfying

$$\Pi'_z = \sum_i w_{zy}\Pi_y,\ \forall y\,. \tag{41}$$

This property for $S^{(j)}_{\mathsf{M},\gamma}(\rho)$ says that, for any M' that is a post-processing of M, one has

$$S^{(j)}_{\mathsf{M}',\gamma}(\rho) \geq S^{(j)}_{\mathsf{M},\gamma}(\rho)\,. \tag{42}$$

This property is satisfied by $j = 1, 3$, with proofs in Appendix E.

Finally, we notice that while the original OE, Eq. (1), is upper bounded as $S_{\mathsf{M}}(\rho) \leq \ln d$, in general, for a non-uniform reference γ, the same bound does not hold, as expected. However

$$S^{(1)}_{\mathsf{M},\gamma}(\rho) \leq S^{(3)}_{\mathsf{M},\gamma}(\rho) \tag{43}$$

holds because the Belavkin-Staszewski relative entropy bounds the Umegaki one from above [43, 44]. Also

$$S^{(1)}_{\mathsf{M},\gamma}(\rho) \leq S^{(2)}_{\mathsf{M},\gamma}(\rho) \tag{44}$$

holds due to joint convexity of the relative entropy (proof in Appendix F).

6 Examples

6.1 Gibbs prior

In the presence of a Hamiltonian $H = \sum_{n=0}^{d-1} E_n|n\rangle\langle n|$, a very natural choice of non-uniform prior is the Gibbs state

$$\gamma := e^{-\beta H}/\mathrm{Tr}[e^{-\beta H}],\ \beta > 0. \tag{45}$$

Definition	Deficiency interpretation	Irretrodictability interpretation	Equal to $S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho)$ when	Petz recovery criterion	Non-decreasing under stochastic post-processing
$S^{(1)}_{\mathsf{M},\gamma}$ (33)	$D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$	N/A	$[\rho,\gamma]=0$	Yes	Yes
$S^{(2)}_{\mathsf{M},\gamma}$ (34)	N/A	$D(Q_F\|Q_R^\gamma)$	ρ,γ,Π_y commute	Yes	No
$S^{(3)}_{\mathsf{M},\gamma}$ (36)	$D_{\mathrm{BS}}(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$	$D_{\mathrm{BS}}({}^tQ_F\|{}^tQ_R^\gamma)$	$[\rho,\gamma]=0$	Yes	Yes

Table 1: Properties of $S^{(j)}_{\mathsf{M},\gamma}$. The expressions for the statistical deficiency and irretrodictability interpretations do not match if one uses the Umegaki relative entropy. On the other hand, the use of the Belavkin-Staszewski relative entropy gives an expression that unifies both interpretations.

We also consider the measurement in the energy eigenbasis $\mathsf{M} := \{|0\rangle\langle 0|, \dots, |d-1\rangle\langle d-1|\}$, but to move far away from the classical case we assume that the input state is pure and maximally unbiased with the energy eigenbasis:

$$\rho := |\psi\rangle\langle\psi| \text{ with } |\psi\rangle := \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} |n\rangle\,. \tag{46}$$

With these assumptions, the first definition yields

$$S^{(1)}_{\mathsf{M},\gamma}(\rho) = S_{\mathsf{M}}(\rho) = \ln d\,, \tag{47}$$

which is also the case if ρ is a mixture of maximally unbiased states. Like $S^{\mathrm{clax}}_{\mathsf{M},\gamma}$, $S^{(1)}_{\mathsf{M},\gamma}$ reduces to the original S_{M} when the prior is a convex sum of the measurement elements.

The second definition yields

$$S^{(2)}_{\mathsf{M},\gamma}(\rho) = \infty\,, \tag{48}$$

for any pure state, since the support of $Q_R^\gamma = \frac{1}{d}\sum_n |n\rangle\langle n| \otimes |n\rangle\langle n|$ does not contain the support of $Q_F = \frac{1}{d}\mathbb{1} \otimes |\psi\rangle\langle\psi|$. We shall comment on this result after the next example.

Finally, the third definition yields

$$\begin{aligned} S^{(3)}_{\mathsf{M},\gamma}(\rho) &= \ln \mathrm{Tr}[\gamma^{-1}] + \frac{1}{d}\mathrm{Tr}[\ln\gamma] && (49)\\ &= \ln \frac{1-e^{\beta\omega d}}{1-e^{\beta\omega}} - \frac{\beta\omega(d-1)}{2}\,. && (50)\end{aligned}$$

where the first line is general, while the second is the expression for equidistant spectrum $E_n = n\omega$. Thus $S^{(3)}_{\mathsf{M},\gamma}$ is more sensitive than $S^{(1)}_{\mathsf{M},\gamma}$ to quantum situations.

6.2 Three-qubit encoding

The following example is inspired by a simple error-correcting code, the three-qubit encoding of a pure qubit:

$$\alpha|0\rangle + \beta|1\rangle \mapsto \alpha|000\rangle + \beta|111\rangle\,,\ |\alpha|^2 + |\beta|^2 = 1\,. \tag{51}$$

Suppose ρ is the encoded state

$$\rho := (\alpha|000\rangle + \beta|111\rangle)(\alpha^*\langle 000| + \beta^*\langle 111|)\,. \tag{52}$$

We consider the measurement of each qubit in the $\{|+\rangle, |-\rangle\}$ basis, i.e. the POVM elements are projectors on the basis vectors

$$\{|+++\rangle, |++-\rangle, |+-+\rangle, \dots, |---\rangle|\}\,. \tag{53}$$

As for the prior, we suppose that the observer knows the encoding of the error correction code, and expect ρ to be more probably in the subspace spanned by $|000\rangle$ and $|111\rangle$, possibly with a bias towards one of those product states; whence

$$\begin{aligned}\gamma &:= p_0|000\rangle\langle 000| + p_1|111\rangle\langle 111| \\ &\quad + \frac{1-p_0-p_1}{6}(\mathbb{1} - |000\rangle\langle 000| - |111\rangle\langle 111|)\,, \qquad (54)\\ &p_0, p_1 > 0,\ p_0 + p_1 < 1\,.\end{aligned}$$

In this case, the three definitions proposed here yield

$$S^{(1)}_{\mathsf{M},\gamma}(\rho) = |\alpha|^2 \ln\frac{1}{p_0} + |\beta|^2 \ln\frac{1}{p_1} - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))\,, \tag{55}$$

$$S^{(2)}_{\mathsf{M},\gamma}(\rho) = \infty\,, \tag{56}$$

$$S^{(3)}_{\mathsf{M},\gamma}(\rho) = \ln\left(\frac{|\alpha|^2}{p_0} + \frac{|\beta|^2}{p_1}\right) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \tag{57}$$

with

$$\begin{aligned} &D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \\ &= |\alpha+\beta|^2 \ln|\alpha+\beta| + |\alpha-\beta|^2 \ln|\alpha-\beta|\,. \end{aligned} \tag{58}$$

$S^{(1)}_{\mathsf{M},\gamma}$ and $S^{(3)}_{\mathsf{M},\gamma}$ differ in the first term, as long as $p_0 \neq p_1$: for $p_0 = p_1 = p$, both yield $\ln(1/p)$.

In particular, when $p = \frac{1}{8}$, $\gamma = u$ and therefore $S^{(1)}_{\mathsf{M},\gamma}(\rho) = S^{(3)}_{\mathsf{M},\gamma}(\rho) = S_{\mathsf{M}}(\rho)$. We see that $S^{(2)}_{\mathsf{M},\gamma}$ is still infinite, for the same reason of support mismatch as in the previous example. From the examples, we observe that $S^{(2)}_{\mathsf{M},\gamma}$ is often overly sensitive to the non-commutativity between ρ and γ. This suggests that, instead of the natural choice of Q_F and Q^{γ}_R as input-output representations, one could opt for representations whose supports are more aligned, such as ${}^tQ_F$ and ${}^tQ^{\gamma}_R$, which relate to $S^{(3)}_{\mathsf{M},\gamma}$ via Eq. (37).

7 Conclusions

The original definition [Eq. (1)] of observational entropy (OE) was known to be lower-bounded by the von Neumann entropy. Here we have first brought to the fore that the excess term $\Sigma_{\mathsf{M}}(\rho)$ can be interpreted in two ways: as a statistical deficiency (4), quantifying the decrease of state distinguishability induced by the measurement; and as irretrodictability (13), quantifying the hardness of retrodicting the input from the output statistics. While it is intuitive that recovering the input state is harder if the measurement makes states less distinguishable, the exact coincidence of the quantifiers is of interest.

In both interpretations, we observe that the uniform state u plays the role of reference, or prior, knowledge. This may not represent the proper knowledge of the physical situation: for instance, for systems in contact with a thermal bath, it may be more natural to choose the Gibbs prior. Based on this, we have studied generalisations of OE, in which the prior knowledge can be an arbitrary state γ.

When $[\rho, \gamma] = 0$, we find an obvious generalisation of the excess term [Eq. (16)] that retains both interpretations of statistical deficiency and irretrodictability. This is no longer straightforward for a general quantum prior. Technically, one of the main difficulty lies in that the irretrodictability quantifier is a relative entropy between joint input-and-output objects, whose definition in quantum theory is a current topic of research. We have explored three possible definitions of generalized OE (Table 1): two specifically designed to satisfy one of the interpretations but lacking the other; the third retaining both by replacing the usual Umegaki relative entropy with the Belavkin-Staszewski version. Thus we have a novel fully quantum object, that quantifies simultaneously the loss of distinguishability by the measurement and the hardness to retrodict the input knowing the output. Being built from information-theoretical considerations, our new formulation of OE may also hold significance in physical (thermodynamical) contexts, such as its relationship with work extraction [19]. We leave a deeper exploration of the physical implications of OE for future research.

Acknowledgments

We thank Clive Aw, Fumio Hiai, Anna Jenčová and Andreas Winter for discussions.

G.B. and V.S. are supported by the National Research Foundation, Singapore and A*STAR under its CQT Bridging Grant; and by the Ministry of Education, Singapore, under the Tier 2 grant "Bayesian approach to irreversibility" (Grant No. MOE-T2EP50123-0002). D.Š. acknowledges the support from the Institute for Basic Science in Korea (IBS-R024-D1). F.B. acknowledges 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 No. 20K03746 and No. 23K03230. J.S. acknowledges support by MICIIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat de Catalunya.

References

[1] John von Neumann. "Mathematical foundations of quantum mechanics". Princeton university press. (1955).

[2] John von Neumann. "Proof of the ergodic theorem and the H-theorem in quantum mechanics. Translation of: Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik". European Physical Journal H **35**, 201–237 (2010).

[3] Dominik Šafránek, J. M. Deutsch, and Anthony Aguirre. "Quantum coarse-grained entropy and thermodynamics". Phys. Rev. A **99**, 010101 (2019). arXiv:1707.09722.

[4] Dominik Šafránek, J. M. Deutsch, and Anthony Aguirre. "Quantum coarse-grained

entropy and thermalization in closed systems". Phys. Rev. A **99**, 012103 (2019). arXiv:1803.00665.

[5] Dominik Šafránek, Anthony Aguirre, Joseph Schindler, and J. M. Deutsch. "A Brief Introduction to Observational Entropy". Foundations of Physics **51**, 101 (2021). arXiv:2008.04409.

[6] Philipp Strasberg and Andreas Winter. "First and second law of quantum thermodynamics: A consistent derivation based on a microscopic definition of entropy". PRX Quantum **2**, 030202 (2021).

[7] Dominik Šafránek and Juzar Thingna. "Quantifying information extraction using generalized quantum measurements". Physical Review A **108**, 032413 (2023). arXiv:2007.07246.

[8] 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**, 053002 (2023).

[9] Andreu Riera-Campeny, Anna Sanpera, and Philipp Strasberg. "Quantum systems correlated with a finite bath: Nonequilibrium dynamics and thermodynamics". PRX Quantum **2**, 010340 (2021). arXiv:2008.02184.

[10] Dominik Šafránek, Anthony Aguirre, and J. M. Deutsch. "Classical dynamical coarse-grained entropy and comparison with the quantum version". Phys. Rev. E **102**, 032106 (2020). arXiv:1905.03841.

[11] Joshua M. Deutsch, Dominik Šafránek, and Anthony Aguirre. "Probabilistic bound on extreme fluctuations in isolated quantum systems". Phys. Rev. E**101**, 032112 (2020). arXiv:1806.08897.

[12] Dana Faiez, Dominik Šafránek, J. M. Deutsch, and Anthony Aguirre. "Typical and extreme entropies of long-lived isolated quantum systems". Phys. Rev. A**101**, 052101 (2020). arXiv:1908.07083.

[13] Charlie Nation and Diego Porras. "Taking snapshots of a quantum thermalization process: Emergent classicality in quantum jump trajectories". Phys. Rev. E**102**, 042115 (2020). arXiv:2003.08425.

[14] Philipp Strasberg, María García Díaz, and Andreu Riera-Campeny. "Clausius inequality for finite baths reveals universal efficiency improvements". Phys. Rev. E **104**, L022103 (2021). arXiv:2012.03262.

[15] Ryusuke Hamazaki. "Speed Limits for Macroscopic Transitions". PRX Quantum **3**, 020319 (2022). arXiv:2110.09716.

[16] Ranjan Modak and S. Aravinda. "Observational-entropic study of anderson localization". Phys. Rev. A **106**, 062217 (2022).

[17] Sreeram PG, Ranjan Modak, and S. Aravinda. "Witnessing quantum chaos using observational entropy". Phys. Rev. E **107**, 064204 (2023).

[18] Joseph Schindler and Andreas Winter. "Continuity bounds on observational entropy and measured relative entropies". Journal of Mathematical Physics**64** (2023). arXiv:2302.00400.

[19] Dominik Šafránek, Dario Rosa, and Felix C. Binder. "Work extraction from unknown quantum sources". Phys. Rev. Lett. **130**, 210401 (2023).

[20] Dominik Šafránek and Dario Rosa. "Measuring energy by measuring any other observable". Phys. Rev. A **108**, 022208 (2023).

[21] Dominik Šafránek. "Ergotropic interpretation of entanglement entropy" (2023). arXiv:2306.08987.

[22] Hisaharu Umegaki. "On information in operator algebras". Proc. Japan Acad. **37**, 459–461 (1961).

[23] Hisaharu Umegaki. "Conditional expectation in an operator algebra. iv. entropy and information". Kodai Mathematical Journal**14** (1962).

[24] S. Kullback and R. A. Leibler. "On information and sufficiency". Ann. Math. Statist. **22**, 79–86 (1951).

[25] Gavin E. Crooks. "Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems". Journal of Statistical Physics **90**, 1481–1487 (1998).

[26] Massimiliano Esposito, Upendra Harbola, and Shaul Mukamel. "Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems". Rev. Mod. Phys. **81**, 1665–1702 (2009).

[27] Gabriel T. Landi and Mauro Paternostro. "Irreversible entropy production: From clas-

sical to quantum". Rev. Mod. Phys. **93**, 035008 (2021).

[28] Francesco Buscemi and Valerio Scarani. "Fluctuation theorems from bayesian retrodiction". Phys. Rev. E **103**, 052111 (2021).

[29] Clive Cenxin Aw, Francesco Buscemi, and Valerio Scarani. "Fluctuation theorems with retrodiction rather than reverse processes". AVS Quantum Science **3**, 045601 (2021). arXiv:https://doi.org/10.1116/5.0060893.

[30] Richard Carl Jeffrey. "The logic of decision". McGraw-Hill. (1965).

[31] Judea Pearl. "Probabilistic reasoning in intelligent systems: networks of plausible inference". Elsevier. (1988).

[32] Hei Chan and Adnan Darwiche. "On the revision of probabilistic beliefs using uncertain evidence". Artificial Intelligence **163**, 67 – 90 (2005).

[33] Satosi Watanabe. "Symmetry of physical laws. part iii. prediction and retrodiction". Rev. Mod. Phys. **27**, 179–186 (1955).

[34] Satosi Watanabe. "Conditional probabilities in physics". Progr. Theor. Phys. Suppl. **E65**, 135–160 (1965).

[35] Arthur J. Parzygnat and Francesco Buscemi. "Axioms for retrodiction: achieving time-reversal symmetry with a prior". Quantum **7**, 1013 (2023).

[36] Denes Petz. "Sufficient subalgebras and the relative entropy of states of a von neumann algebra". Comm. Math. Phys. **105**, 123–131 (1986).

[37] Denes Petz. "Sufficiency of channels over von Neumann algebras". The Quarterly Journal of Mathematics **39**, 97–108 (1988).

[38] Man-Duen Choi. "Completely positive linear maps on complex matrices". Linear Algebra and its Applications **10**, 285–290 (1975).

[39] M. S. Leifer. "Conditional density operators and the subjectivity of quantum operations". In AIP Conference Proceedings. Volume 889, page 172–186. AIP (2007).

[40] 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** (2013).

[41] Viacheslav P. Belavkin and P. Staszewski. "C^*-algebraic generalization of relative entropy and entropy". In Annales de l'IHP Physique théorique. Volume 37 no. 1, pages 51–58. (1982). url: http://www.numdam.org/item/AIHPA_1982__37_1_51_0/.

[42] Sumeet Khatri and Mark M. Wilde. "Principles of quantum communication theory: A modern approach" (2020). arXiv:2011.04672.

[43] Keiji Matsumoto. "A new quantum version of f-divergence". In Reality and Measurement in Algebraic Quantum Theory. Page 229–273. Springer Singapore (2018).

[44] Fumio Hiai and Milán Mosonyi. "Different quantum f-divergences and the reversibility of quantum operations". Reviews in Mathematical Physics **29**, 1750023 (2017).

[45] Andreas Bluhm, Ángela Capel, Paul Gondolf, and Antonio Pérez-Hernández. "Continuity of quantum entropic quantities via almost convexity". IEEE Transactions on Information Theory **69**, 5869–5901 (2023).

[46] Fumio Hiai, Milán Mosonyi, Dénes Petz, and Cédric Bény. "Quantum f-divergences and error correction". Reviews in Mathematical Physics **23**, 691–747 (2011).

[47] Andreas Bluhm and Ángela Capel. "A strengthened data processing inequality for the belavkin–staszewski relative entropy". Reviews in Mathematical Physics **32**, 2050005 (2019).

A Proof of Lemma 1

Lemma 1 relates the Choi operators of the forward and reverse processes. This can be shown using their definitions (24) and (25).

Proof. Let the Kraus representation of $\mathcal{M}$ be $\mathcal{M}(\rho) = \sum_k K_k \rho K_k^\dagger$. First, observe the following identity

$$\sum_{i,j} A\left|i\right\rangle\left\langle j\right| B \otimes \left|i\right\rangle\left\langle j\right| = \sum_{i,j} \left|i\right\rangle\left\langle j\right| \otimes A^T \left|i\right\rangle\left\langle j\right| B^T \tag{59}$$

for any operators A and B. Using this identity twice, the right-hand side of Eq. (26) equals to

$$\begin{aligned}
&\left(\mathcal{M}(\gamma)^{-1/2} \otimes \sqrt{\gamma^T}\right) C_{\mathcal{M}} \left(\mathcal{M}(\gamma)^{-1/2} \otimes \sqrt{\gamma^T}\right) \\
&= \sum_{i,j} \mathcal{M}(\gamma)^{-1/2} \mathcal{M}(\left|i\right\rangle\left\langle j\right|) \mathcal{M}(\gamma)^{-1/2} \otimes \sqrt{\gamma^T} \left|i\right\rangle\left\langle j\right| \sqrt{\gamma^T} \\
&= \sum_{i,j,k} \mathcal{M}(\gamma)^{-1/2} \left|i\right\rangle\left\langle j\right| \mathcal{M}(\gamma)^{-1/2} \otimes \sqrt{\gamma^T} K_k^T \left|i\right\rangle\left\langle j\right| K_k^* \sqrt{\gamma^T} \\
&= \sum_{i,j,k} \left|i\right\rangle\left\langle j\right| \otimes \sqrt{\gamma^T} K_k^T (\mathcal{M}(\gamma)^{-1/2})^T \left|i\right\rangle\left\langle j\right| (\mathcal{M}(\gamma)^{-1/2})^T K_k^* \sqrt{\gamma^T}
\end{aligned} \tag{60}$$

On the other hand,

$$\begin{aligned}
C_{\widetilde{\mathcal{M}}^\gamma} &= \sum_{i,j} \left|i\right\rangle\left\langle j\right| \otimes \widetilde{\mathcal{M}}^\gamma(\left|i\right\rangle\left\langle j\right|) \\
&= \sum_{i,j,k} \left|i\right\rangle\left\langle j\right| \otimes \sqrt{\gamma} K_k^\dagger \mathcal{M}(\gamma)^{-1/2} \left|i\right\rangle\left\langle j\right| \mathcal{M}(\gamma)^{-1/2} K_k \sqrt{\gamma} \\
&= \sum_{i,j,k} \left|j\right\rangle\left\langle i\right| \otimes \sqrt{\gamma} K_k^\dagger \mathcal{M}(\gamma)^{-1/2} \left|j\right\rangle\left\langle i\right| \mathcal{M}(\gamma)^{-1/2} K_k \sqrt{\gamma}
\end{aligned} \tag{61}$$

Notice that Eq. (61) is the transpose of Eq. (60). This proves Eq. (26). □

B Proof of Eq. (37)

The most important observation for Eq. (37) is that, one of the $\sqrt{C_{\mathcal{M}}}$ in the definitions of ${}^tQ_F$ (31) and ${}^tQ_R$ (32) will cancel each other in the expression of $D_{\mathrm{BS}}({}^tQ_F \| {}^tQ_R^\gamma)$, leaving a tensor product inside the logarithm. That is to say,

$$\begin{aligned}
&\ln {}^tQ_F ({}^tQ_R^\gamma)^{-1} \\
&= \ln \sqrt{C_{\mathcal{M}}}(\mathbb{1} \otimes \rho^T)\sqrt{C_{\mathcal{M}}} \left(\sqrt{C_{\mathcal{M}}}(\mathcal{M}(\gamma)^{-1/2} \tau \mathcal{M}(\gamma)^{-1/2} \otimes \gamma^T)\sqrt{C_{\mathcal{M}}}\right)^{-1} \\
&= \ln \sqrt{C_{\mathcal{M}}}(\mathcal{M}(\gamma)^{1/2} \tau^{-1} \mathcal{M}(\gamma)^{1/2} \otimes \rho^T (\gamma^T)^{-1})\sqrt{C_{\mathcal{M}}}^{-1} \\
&= \sqrt{C_{\mathcal{M}}}(\ln \mathcal{M}(\gamma)^{1/2} \tau^{-1} \mathcal{M}(\gamma)^{1/2} \otimes \rho^T (\gamma^T)^{-1})\sqrt{C_{\mathcal{M}}}^{-1} \\
&= \sqrt{C_{\mathcal{M}}} \left(\ln \mathcal{M}(\gamma)^{1/2} \tau^{-1} \mathcal{M}(\gamma)^{1/2} \otimes \mathbb{1} + \mathbb{1} \otimes \ln \rho^T (\gamma^T)^{-1}\right) \sqrt{C_{\mathcal{M}}}^{-1}
\end{aligned} \tag{62}$$

Notice that ${}^tQ_R^\gamma$ being full-rank implies $C_{\mathcal{M}}$ and γ being full-rank. Putting Eq. (62) into the definition

of D_{BS}, the left-hand side of Eq. (37) is

$$\begin{aligned}
&D_{\mathrm{BS}}\left({}^tQ_F\middle\|{}^tQ_R^\gamma\right)\\
&=\mathrm{Tr}[{}^tQ_F\ \ln{}^tQ_F({}^tQ_R^\gamma)^{-1}]\\
&=\mathrm{Tr}\left[(\mathbb{1}\otimes\rho^T)C_{\mathcal{M}}\left(\ln\mathcal{M}(\gamma)^{1/2}\tau^{-1}\mathcal{M}(\gamma)^{1/2}\otimes\mathbb{1}+\mathbb{1}\otimes\ln\rho^T(\gamma^T)^{-1}\right)\right]\\
&=\mathrm{Tr}[\mathrm{Tr}_B[(\mathbb{1}\otimes\rho^T)C_{\mathcal{M}}]\ \ln\rho^T(\gamma^T)^{-1}]+\mathrm{Tr}[\mathrm{Tr}_A[(\mathbb{1}\otimes\rho^T)C_{\mathcal{M}}]\ \ln\mathcal{M}(\gamma)^{1/2}\tau^{-1}\mathcal{M}(\gamma)^{1/2}]\\
&=\mathrm{Tr}[\rho^T\ \ln\rho^T(\gamma^T)^{-1}]+\mathrm{Tr}[\mathcal{M}(\rho)\ \ln\mathcal{M}(\gamma)^{1/2}\tau^{-1}\mathcal{M}(\gamma)^{1/2}]\\
&=D_{\mathrm{BS}}(\rho\|\gamma)-\mathrm{Tr}[\mathcal{M}(\rho)\ \ln\mathcal{M}(\gamma)^{-1/2}\tau\mathcal{M}(\gamma)^{-1/2}]
\end{aligned}\tag{63}$$

Recall that we choose the input of the reverse process to be $\tau=\mathcal{M}(\rho)$. Noting that $\mathcal{M}(\rho)$ commutes with $\mathcal{M}(\gamma)$, the second term equals to

$$\begin{aligned}
&\mathrm{Tr}[\mathcal{M}(\rho)\ \ln\mathcal{M}(\gamma)^{-1/2}\tau\mathcal{M}(\gamma)^{-1/2}]\\
&=\mathrm{Tr}[\mathcal{M}(\rho)(\ln\mathcal{M}(\rho)-\ln\mathcal{M}(\gamma))]\\
&=D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))\,.
\end{aligned}\tag{64}$$

This proves Eq. (37).

C Proof of properties (i)-(iii)

Since $S^{\mathrm{clax}}_{\mathsf{M},u}(\rho)=S_{\mathsf{M}}(\rho)$ and $[\rho,u]=0$, property (i) is a special case of property (ii), so we prove (ii) directly.

When $[\rho,\gamma]=0$, $D(\rho\|\gamma)$ and $D_{\mathrm{BS}}(\rho\|\gamma)$ are both equal to the relative entropy between the eigenvalues of ρ and γ. By comparing their definitions Eqs. (17), (33) and (36), we obtain $S^{(1)}_{\mathsf{M},\gamma}(\rho)=S^{(3)}_{\mathsf{M},\gamma}(\rho)=S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho)$.

For $S^{(2)}_{\mathsf{M},\gamma}$, we further need to use that $[\rho,\Pi_y]=[\gamma,\Pi_y]=0$. This condition indicates that $C_{\mathcal{M}},(\mathbb{1}\otimes\rho),(\mathbb{1}\otimes\gamma)$ all commute, and therefore $Q_F=(\mathbb{1}\otimes\rho^T)C_{\mathcal{M}}$, $Q_R^\gamma=(\mathcal{M}(\rho)\mathcal{M}(\gamma)^{-1}\otimes\gamma^T)C_{\mathcal{M}}$, and

$$\begin{aligned}
\Sigma^{(2)}_{\mathsf{M},\gamma}(\rho)&=D(Q_F\|Q_R^\gamma)\\
&=\mathrm{Tr}[Q_F(\ln Q_F-\ln Q_R^\gamma)]\\
&=\mathrm{Tr}[Q_F(\ln(\mathbb{1}\otimes\rho^T)+\ln C_{\mathcal{M}})]-\mathrm{Tr}\left[Q_F\left(\ln(\mathcal{M}(\rho)\mathcal{M}(\gamma)^{-1}\otimes\gamma^T)+\ln C_{\mathcal{M}}\right)\right]\\
&=\mathrm{Tr}\left[Q_F\left(\mathbb{1}\otimes(\ln\rho^T-\ln\gamma^T)\right)\right]-\mathrm{Tr}[Q_F(\ln\mathcal{M}(\rho)\mathcal{M}(\gamma)^{-1}\otimes\mathbb{1})]\\
&=D(\rho\|\gamma)-D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))\\
&=S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho)-S(\rho)
\end{aligned}\tag{65}$$

Notice that the above equality holds as long as the support of Q_F is contained in that of Q_R^γ without assuming Q_R^γ to be invertible, since the Umegaki relative entropy is continuous with respect to both arguments [45]. Therefore, $S^{(2)}_{\mathsf{M},\gamma}(\rho)=S^{\mathrm{clax}}_{\mathsf{M},\gamma}(\rho)$ and property (ii) holds for $j=2$.

Property (iii) is equivalent to say that $\Sigma^{(j)}_{\mathsf{M},\gamma}$ is non-negative.

$\Sigma^{(2)}_{\mathsf{M},\gamma}(\rho)=D(Q_F\|Q_R^\gamma)$ is non-negative by the non-negativity of relative entropy between two unit-trace positive operators.

$\Sigma^{(1)}_{\mathsf{M},\gamma}(\rho)=D(\rho\|\gamma)-D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$ is non-negative by the data-processing inequality of Umegaki relative entropy. Last, by Eq. (43), $\Sigma^{(3)}_{\mathsf{M},\gamma}(\rho)\geq\Sigma^{(1)}_{\mathsf{M},\gamma}(\rho)\geq 0$.

D Proof of Petz recovery criteria (v)

We first show that both $S^{(1)}_{\mathsf{M},\gamma}$ and $S^{(3)}_{\mathsf{M},\gamma}$ are equal to $S(\rho)$ if and only if $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$. The $j = 2$ case is addressed later.

D.1 Property (v) of $S^{(1)}_{\mathsf{M},\gamma}$

The property $D(\rho\|\gamma) = D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \Leftrightarrow \widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$ is shown, for the larger family called f-divergences, by Theorem 5.1 in Ref. [46] and Theorem 3.18 in Ref. [44]. Therefore, the property (v) for $S^{(1)}_{\mathsf{M},\gamma}$ is proved since $S^{(1)}_{\mathsf{M},\gamma}(\rho) = S(\rho)$ is equivalent to $D(\rho\|\gamma) = D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$.

D.2 Property (v) of $S^{(3)}_{\mathsf{M},\gamma}$

For $S^{(3)}_{\mathsf{M},\gamma}$, first assume $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$. By the data-processing inequality of the Belavkin-Staszewski relative entropy [44, 47],

$$\begin{aligned} D_{\mathrm{BS}}(\rho\|\gamma) &\geq D_{\mathrm{BS}}(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \\ &\geq D_{\mathrm{BS}}\left(\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho))\middle\|\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\gamma))\right) \\ &= D_{\mathrm{BS}}(\rho\|\gamma)\,, \end{aligned} \tag{66}$$

where $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\gamma)) = \gamma$ by definition. Therefore, $D_{\mathrm{BS}}(\rho\|\gamma) = D_{\mathrm{BS}}(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$ and thus $S^{(3)}_{\mathsf{M},\gamma}(\rho) = S(\rho)$.

On the other hand, if $S^{(3)}_{\mathsf{M},\gamma}(\rho) = S(\rho)$, by Eq. (43) and property (iii), $S^{(3)}_{\mathsf{M},\gamma}(\rho) \geq S^{(1)}_{\mathsf{M},\gamma}(\rho) \geq S(\rho)$, one infers $S^{(1)}_{\mathsf{M},\gamma}(\rho) = S(\rho)$, and therefore $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$ by property (v) of $S^{(1)}_{\mathsf{M},\gamma}$.

D.3 Property (v) of $S^{(2)}_{\mathsf{M},\gamma}$

Next, we prove property (v) for $S^{(2)}_{\mathsf{M},\gamma}$. Before that, we notice that the Petz recovery condition $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$ indicates that $[\rho, \gamma] = 0$.

Lemma 2. *For a measurement channel $\mathcal{M}$, $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$ implies $[\rho, \gamma] = 0$.*

Proof. By property (v) of $S^{(1)}_{\mathsf{M},\gamma}$ and $S^{(3)}_{\mathsf{M},\gamma}$, one has $\Sigma^{(1)}_{\mathsf{M},\gamma}(\rho) = \Sigma^{(3)}_{\mathsf{M},\gamma}(\rho) = 0$. That is,

$$\begin{aligned} D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) &= D_{\mathrm{BS}}(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \\ D(\rho\|\gamma) &= D_{\mathrm{BS}}(\rho\|\gamma) \end{aligned} \tag{67}$$

Because $D(\rho\|\gamma) = D_{\mathrm{BS}}(\rho\|\gamma)$ if and only if $[\rho, \gamma] = 0$ (Ref. [44] Theorem 4.3), one has $[\rho, \gamma] = 0$. □

Now, we prove property (v). We first prove the "only if" part.

Suppose $S^{(2)}_{\mathsf{M},\gamma} = S(\rho)$. This is equivalent to that $\Sigma^{(2)}_{\mathsf{M},\gamma}(\rho) = D(Q_F\|Q_R^{\gamma}) = 0$, which is equivalent to that $Q_F = Q_R^{\gamma}$. Taking the partial trace over system B, we get $\rho^T = \mathrm{Tr}_B[Q_F] = \mathrm{Tr}_B[Q_R^{\gamma}] = \widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho))^T$. Therefore, $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$.

For the "if" part, we will show that $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$ implies $Q_F = Q_R^{\gamma}$, which is equivalent to $S^{(2)}_{\mathsf{M},\gamma} = S(\rho)$.

Suppose $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$. By Lemma 2, $[\rho, \gamma] = 0$, and thus we can diagonalize them in the same basis:

$$\rho = \sum_x \lambda_x |\psi_x\rangle\langle\psi_x|\,, \tag{68}$$

$$\gamma = \sum_x \gamma_x |\psi_x\rangle\langle\psi_x|\,. \tag{69}$$

Next, we construct a new POVM by taking the diagonal elements of the POVM $\mathsf{M} = \{\Pi_y\}$ in the above basis:

$$\mathsf{M}' \coloneqq \{\Pi'_y\}, \quad \Pi'_y \coloneqq \sum_x |\psi_x\rangle\langle\psi_x|\Pi_y|\psi_x\rangle\langle\psi_x| \,. \tag{70}$$

Let $\mathcal{M}'(\rho) \coloneqq \sum_y \mathrm{Tr}[\Pi'_y\rho]|y\rangle\langle y|$ be the measurement channel and $\widetilde{\mathcal{M}'}^{\gamma}$ be its Petz map with reference γ. Notice that $\mathrm{Tr}[\Pi_y\rho] = \mathrm{Tr}[\Pi'_y\rho]$, $\mathrm{Tr}[\Pi_y\gamma] = \mathrm{Tr}[\Pi'_y\gamma]$, and therefore $\mathcal{M}'(\rho) = \mathcal{M}(\rho)$, $\mathcal{M}'(\gamma) = \mathcal{M}(\gamma)$.

Since $[\rho,\gamma] = [\rho,\Pi'_y] = [\gamma,\Pi'_y] = 0$, using property (ii) for $\mathcal{M}'$, one has

$$\begin{aligned} S^{\mathrm{clax}}_{\mathsf{M}',\gamma}(\rho) - S(\rho) &= D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \\ &= D(Q_{F'}\|Q^{\gamma}_{R'}) \,, \end{aligned} \tag{71}$$

where $Q_{F'} \coloneqq (\mathbb{1}\otimes\rho^T)C_{\mathcal{M}'}$ and $Q^{\gamma}_{R'} \coloneqq (\mathcal{M}'(\rho)\mathcal{M}'(\gamma)^{-1}\otimes\gamma^T)C_{\mathcal{M}'}$ are the representations of $\mathcal{M}'$ and $\widetilde{\mathcal{M}'}^{\gamma}$, which are simplified using the commutativity of ρ,γ and Π'_y.

Note that the expression Eq. (71) is the same as $\Sigma^{(1)}_{\mathsf{M},\gamma}(\rho) = D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))$. Since $\widetilde{\mathcal{M}}^{\gamma}(\mathcal{M}(\rho)) = \rho$, by property (v) for $S^{(1)}_{\mathsf{M},\gamma}$, we have $\Sigma^{(1)}_{\mathsf{M},\gamma}(\rho) = 0$. Combining this with Eq. (71) indicates that $D(Q_{F'}\|Q^{\gamma}_{R'}) = 0$ and $Q_{F'} = Q^{\gamma}_{R'}$. Expanding this with respect to the basis $\{|y\rangle\otimes|\psi_x\rangle\}$, one get

$$\langle y|\langle\psi_x| Q_{F'} |y\rangle|\psi_x\rangle = \langle y|\langle\psi_x| Q^{\gamma}_{R'} |y\rangle|\psi_x\rangle \tag{72}$$

$$\varphi(y|x)\lambda_x = \frac{\varphi(y|x)\gamma_x \mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}, \ \forall x,y \tag{73}$$

where $\varphi(y|x) = \langle\psi_x|\Pi'_y|\psi_x\rangle = \langle\psi_x|\Pi_y|\psi_x\rangle$.

Now, fix y, and consider the support of Π'_y, which is spanned by all $|\psi_x\rangle$ such that $\varphi(y|x) \neq 0$. In this subset of x, one could cancel out $\varphi(y|x)$ in both sides and Eq. (73) becomes

$$\lambda_x = \frac{\mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}\gamma_x, \ \text{for } x : |\psi_x\rangle \in \mathrm{Supp}(\Pi'_y) \,, \tag{74}$$

where $\mathrm{Supp}(\Pi'_y)$ denotes the support of Π'_y. Taking the square root of this equation, one gets

$$\sqrt{\lambda_x} = \sqrt{\frac{\mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}}\sqrt{\gamma_x}, \ \text{for } x : |\psi_x\rangle \in \mathrm{Supp}(\Pi'_y) \,. \tag{75}$$

Define $(\Pi'_y)^0 \coloneqq \sum_{x:\varphi(y|x)\neq 0} |\psi_x\rangle\langle\psi_x|$ as the projector onto $\mathrm{Supp}(\Pi'_y)$. The above equation can be rewritten as

$$\sqrt{\rho}(\Pi'_y)^0 = \sqrt{\frac{\mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}}\sqrt{\gamma}(\Pi'_y)^0 \,, \tag{76}$$

since multiplication with $(\Pi'_y)^0$ selects the eigenvectors of $\sqrt{\rho}$ and $\sqrt{\gamma}$ in $\mathrm{Supp}(\Pi'_y)$, which are $\sqrt{\lambda_x}$ and $\sqrt{\gamma_x}$ in Eq. (75).

Since Π_y is positive, its off-diagonal elements are cross terms in $\mathrm{Supp}(\Pi'_y)$:

$$\Pi_y = \Pi'_y + \sum_{\substack{x_1,x_2:x_1\neq x_2,\\ |\psi_{x_1}\rangle,|\psi_{x_2}\rangle\in\mathrm{Supp}(\Pi'_y)}} \pi_{y,x_1,x_2} |\psi_{x_1}\rangle\langle\psi_{x_2}| \tag{77}$$

for some complex numbers π_{y,x_1,x_2}. Therefore, $\Pi_y(\Pi'_y)^0 = (\Pi'_y)^0\Pi_y = \Pi_y$.

By this and Eq. (76),

$$\begin{aligned}\sqrt{\rho}\Pi_y\sqrt{\rho} &= \sqrt{\rho}(\Pi'_y)^0\Pi_y(\Pi'_y)^0\sqrt{\rho}\\ &= \frac{\mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}\sqrt{\gamma}(\Pi'_y)^0\Pi_y(\Pi'_y)^0\sqrt{\gamma}\\ &= \frac{\mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}\sqrt{\gamma}\Pi_y\sqrt{\gamma}\end{aligned} \tag{78}$$

Last, by $\mathrm{Tr}[\Pi_y\rho] = \mathrm{Tr}[\Pi'_y\rho]$, $\mathrm{Tr}[\Pi_y\gamma] = \mathrm{Tr}[\Pi'_y\gamma]$ and Eq. (78), one gets

$$\begin{aligned}Q_F &= (\mathbb{1}\otimes\sqrt{\rho^T})C_{\mathcal{M}}(\mathbb{1}\otimes\sqrt{\rho^T})\\ &= \sum_y |y\rangle\!\langle y| \otimes \sqrt{\rho^T}\Pi_y^T\sqrt{\rho^T}\\ &= \sum_y |y\rangle\!\langle y| \otimes \frac{\mathrm{Tr}[\Pi'_y\rho]}{\mathrm{Tr}[\Pi'_y\gamma]}\sqrt{\gamma^T}\Pi_y^T\sqrt{\gamma^T}\\ &= \sum_y |y\rangle\!\langle y| \otimes \frac{\mathrm{Tr}[\Pi_y\rho]}{\mathrm{Tr}[\Pi_y\gamma]}\sqrt{\gamma^T}\Pi_y^T\sqrt{\gamma^T}\\ &= (\mathcal{M}(\rho)\mathcal{M}(\gamma)^{-1}\otimes\sqrt{\gamma^T})C_{\mathcal{M}}(\mathbb{1}\otimes\sqrt{\gamma^T})\\ &= Q_R^\gamma\end{aligned} \tag{79}$$

Therefore, $Q_F = Q_R^\gamma$, and $S^{(2)}_{\mathsf{M},\gamma} = S(\rho)$.

E Proof of monotonicity under stochastic post-processing (vi)

Let $\mathcal{W}$ be the linear map describing the post-processing w, which satisfies

$$\mathcal{W}(|y\rangle\!\langle y|) = \sum_z w_{zy}|z\rangle\!\langle z| \tag{80}$$

$$\mathcal{W}(|y\rangle\langle y'|) = 0, \quad \text{for } y \neq y'\,. \tag{81}$$

Since w is a stochastic matrix, $\mathcal{W}$ is completely positive and trace-preserving. The measurement channel of M' is then described by

$$\begin{aligned}(\mathcal{P}\circ\mathcal{M})(\rho) &= \sum_y \mathrm{Tr}[\Pi_y\rho]\mathcal{P}(|y\rangle\!\langle y|)\\ &= \sum_z \mathrm{Tr}[\Pi'_z\rho]|z\rangle\!\langle z|\end{aligned} \tag{82}$$

Property (vi) is equivalent to that $S^{(j)}_{\mathsf{M}',\gamma}(\rho) - S^{(j)}_{\mathsf{M},\gamma}(\rho) \geq 0$. For $j = 1, 3$, they have the same form:

$$\begin{aligned}&S^{(j)}_{\mathsf{M}',\gamma}(\rho) - S^{(j)}_{\mathsf{M},\gamma}(\rho)\\ &= D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) - D((\mathcal{P}\circ\mathcal{M})(\rho)\|(\mathcal{P}\circ\mathcal{M})(\gamma))\\ &\geq 0\end{aligned} \tag{83}$$

The inequality is due to the data-processing inequality of relative entropy.

F Proof of Eq. (44)

By definition of Q_F and Q_R^γ in Eqs. (27) and (29),

$$Q_F = \sum_y |y\rangle\langle y| \otimes \sqrt{\rho^T}\Pi_y^T\sqrt{\rho^T} \tag{84}$$

$$Q_R = \sum_y |y\rangle\langle y| \otimes \frac{\mathrm{Tr}[\Pi_y\rho]}{\mathrm{Tr}[\Pi_y\gamma]}\sqrt{\gamma^T}\Pi_y^T\sqrt{\gamma^T} \tag{85}$$

and thus

$$\begin{aligned}
D(Q_F\|Q_R) &= \mathrm{Tr}\left[\sum_y |y\rangle\langle y| \otimes \sqrt{\rho^T}\Pi_y^T\sqrt{\rho^T}\left(\ln\sqrt{\rho^T}\Pi_y^T\sqrt{\rho^T} - \ln\frac{\mathrm{Tr}[\Pi_y\rho]\sqrt{\gamma^T}\Pi_y^T\sqrt{\gamma^T}}{\mathrm{Tr}[\Pi_y\gamma]}\right)\right] \\
&= \sum_y \mathrm{Tr}\left[\sqrt{\rho}\Pi_y\sqrt{\rho}\left(\ln\frac{\sqrt{\rho}\Pi_y\sqrt{\rho}}{\mathrm{Tr}[\Pi_y\rho]} - \ln\frac{\sqrt{\gamma}\Pi_y\sqrt{\gamma}}{\mathrm{Tr}[\Pi_y\rho]} - (\ln\mathrm{Tr}[\Pi_y\rho] - \ln\mathrm{Tr}[\Pi_y\gamma])\mathbb{1}\right)\right] \\
&= \sum_y D\left(\frac{\sqrt{\rho}\Pi_y\sqrt{\rho}}{\mathrm{Tr}[\Pi_y\rho]}\middle\|\frac{\sqrt{\gamma}\Pi_y\sqrt{\gamma}}{\mathrm{Tr}[\Pi_y\rho]}\right) - \sum_y \mathrm{Tr}[\sqrt{\rho}\Pi_y\sqrt{\rho}](\ln\mathrm{Tr}[\Pi_y\rho] - \ln\mathrm{Tr}[\Pi_y\gamma]) \\
&\geq D\left(\sum_y \frac{\sqrt{\rho}\Pi_y\sqrt{\rho}}{\mathrm{Tr}[\Pi_y\rho]}\middle\|\sum_y \frac{\sqrt{\gamma}\Pi_y\sqrt{\gamma}}{\mathrm{Tr}[\Pi_y\rho]}\right) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma)) \\
&= D(\rho\|\gamma) - D(\mathcal{M}(\rho)\|\mathcal{M}(\gamma))
\end{aligned} \tag{86}$$

where the inequality comes from the joint convexity of the Umegaki relative entropy. Similar inequalities hold for any other relative entropies that satisfy the joint convexity, such as the Belavkin-Staszewski one. Adding $S(\rho)$ to both sides of Eq. (86) gives Eq. (44).