Incompatibility of unbiased qubit observables and Pauli channels

Teiko Heinosaari

Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, FI-20014 Turku, Finland

Daniel Reitzner, Tomáš Rybár, and Mário Ziman

RCQI, Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 845 11 Bratislava, Slovakia

A quantum observable and a channel are considered compatible if they form parts of the same measurement device, otherwise they are incompatible. Constrains on compatibility between observables and channels can be quantified via relations highlighting the necessary tradeoffs between noise and disturbance within quantum measurements. In this paper we shall discuss the general properties of these compatibility relations and then fully characterize the compatibility conditions for an unbiased qubit observable and a Pauli channel. The implications of the characterization are demonstrated on some concrete examples.

I. INTRODUCTION

The paradigm of quantum incompatibility stands behind many quantum phenomena and quantum information no-go theorems [1]. One of the most paradigmatic manifestations of incompatibility is the *no-information-without-disturbance theorem*. It states that a unitary channel, i.e., a channel that does not cause an irreversible disturbance, is compatible only with trivial observables [2–4]. A trivial observable corresponds to a coin tossing measurement and hence does not give any information on the input state. One therefore concludes that if a measurement gives some information, it must cause disturbance.

The tradeoff between information and disturbance is relevant at least in two different scenarios. First, suppose an unwanted disturbance is identified in a communication channel and it may have been caused by the actions of an eavesdropper. In this case, it is relevant to know what sort of information the eavesdropper may have obtained. This means that we want to know all possible measurements that may have caused the disturbance. Second, we may plan to perform a measurement of an observable A. The measurement causes a necessary state perturbation, but the form of this perturbation can be partly controlled by choosing the way in which we measure A. In this case, it is relevant to know all channels that are compatible with A.

There are several studies in the literature where the information-disturbance relation is investigated by first quantifying information and disturbance and then deriving an inequality for those measures. In this work we follow a structural approach [5,6] that does not commit to any specific quantifications of information and disturbance. The main idea is to determine if a channel and an observable can be parts of the same measurement process or not. After presenting the general characterization of compatible pairs of observables and channels (Sec. II), we concentrate on the cases in which the implemented channel is a Pauli channel (Sec. III). In particular, we derive a complete criterion when a noisy version of a binary qubit observable is compatible with a given Pauli channel.

Finally, we demonstrate the consequences of these results on concrete examples (Sec. IV).

II. INCOMPATIBILITY OF CHANNELS AND OBSERVABLES

A. Two equivalent definitions of incompatibility

When considering (in)compatibility of channels and observables, we can start either from the concept of a *measurement model* or from an *instrument* [7–9]. The first one explains the physical meaning of compatibility, while the latter is more convenient from the mathematical point of view. In the following we recall these two equivalent ways to define (in)compatibility.

A measurement model (see Fig. 1) is a formalized description of a measurement process. A measured system, associated with a Hilbert space $\mathcal{H}$, is coupled to an ancillary system associated with $\mathcal{H}_a$. The composite system undergoes a unitary evolution $\varrho \otimes \xi \mapsto U(\varrho \otimes \xi)U^*$, which is called measurement coupling. After that the ancillary system is measured with a pointer observable F. A measurement model is hence specified by the ancillary Hilbert space $\mathcal{H}_a$, initial state of the ancilla ξ, unitary operator U on $\mathcal{H} \otimes \mathcal{H}_a$, and pointer observable F. From these it is straightforward to determine both the average state disturbance Λ and the effective observable A of the measured system initialized in the state ϱ. In particular, for the probability of getting an outcome x we need to have

$$\mathrm{tr}[\varrho \mathsf{A}(x)] = \mathrm{tr}[U(\varrho \otimes \xi)U^*(\mathbb{1} \otimes \mathsf{F}(x))]. \tag{1}$$

This equation is required to be valid for all input states ϱ, so it in fact determines the observable A. Mathematically, A is a positive operator valued measure (POVM). The average state disturbance on the system is given as

$$\Lambda(\varrho) = \mathrm{tr}_{\mathcal{H}_a}[U(\varrho \otimes \xi)U^*]. \tag{2}$$

Mathematically, Λ is a channel, i.e., a trace-preserving, completely positive map.
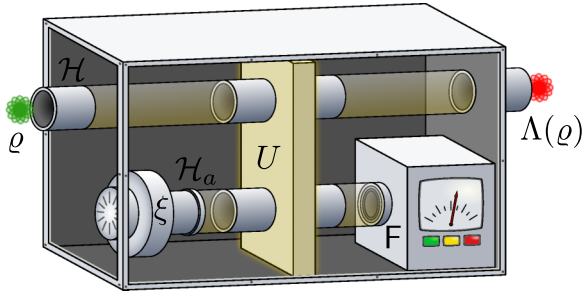
FIG. 1. A measurement model describes a measurement of a system via coupling it to an ancillary system. It gives a description of both the state change $\Lambda(\varrho)$ and the effective measurement A on a given input state ϱ. This model exemplifies the compatibility of channel Λ and observable A.

Physically speaking, the observable A is related to the information that we can extract from the system's initial state ϱ, while the channel Λ describes the average perturbation of the state caused by the measurement process, i.e., $\Lambda(\varrho)$ is the output state of the system when no postselection took place.

In what follows we are interested in the inverse problem: given an observable A and a channel Λ, is there a measurement model [specified by a quadruple $(\mathcal{H}_a, \xi, U, \mathsf{F})$] such that A is given by (1) and Λ is given by (2)? If that is the case, we say that A and Λ are *compatible*; otherwise they are *incompatible*.

The compatibility relation can be concisely expressed in terms of an instrument. An instrument is a map $x \mapsto \Phi_x$ that assigns a trace-non-increasing, completely positive map to each measurement outcome, and such that the sum $\sum_x \Phi_x$ is a trace-preserving map. The instrument related to a measurement model specified by a quadruple $(\mathcal{H}_a, \xi, U, \mathsf{F})$ is

$$\Phi_x(\varrho) = \mathrm{tr}_{\mathcal{H}_a}[U(\varrho \otimes \xi)U^*(\mathbb{1} \otimes \mathsf{F}(x))]. \quad (3)$$

Conversely, for every instrument it is possible to find some measurement model in this way, and one can even choose ξ to be a pure state and F to be a projection valued measure [10]. The compatibility of a channel Λ and an observable A is hence equivalent to the existence of an instrument $x \mapsto \Phi_x$ such that

$$\sum_x \Phi_x(\varrho) = \Lambda(\varrho) \quad \text{and} \quad \mathrm{tr}[\Phi_x(\varrho)] = \mathrm{tr}[\varrho \mathsf{A}(x)] \quad (4)$$

for all outcomes x and input states ϱ.

It is reasonable to expect that simultaneous unitary transformations of both the observable and the channel should not change the compatibility relation of the two. This is formalized in the proposition we present after first fixing some notation. For a unitary operator V, we denote by $\widetilde{V}$ the corresponding unitary channel, i.e.,

$$\widetilde{V}(\varrho) = V\varrho V^*. \quad (5)$$

The functional composition of two channels is denoted by $\circ$. Hence, for two unitary operators V, W and a channel Λ, the composition $\widetilde{W} \circ \Lambda \circ \widetilde{V}$ denotes the channel

$$\varrho \mapsto W\Lambda(V\varrho V^*)W^*. \quad (6)$$

Further, if A is an observable, then we denote by $V^*\mathsf{A}V$ the observable consisting of operators $V^*\mathsf{A}(x)V$.
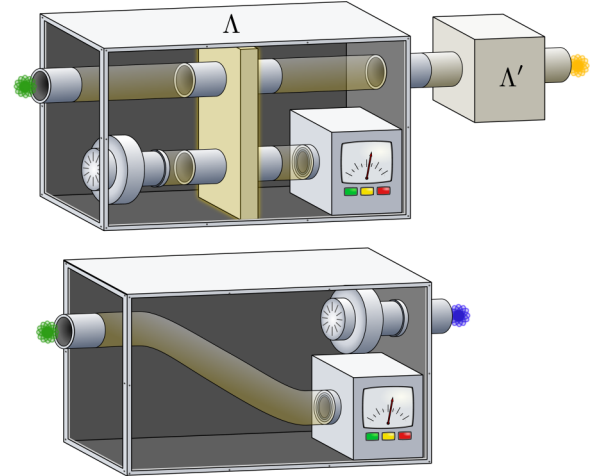

FIG. 2. The set $\mathfrak{C}_\mathsf{A}$ is a left ideal of the set of all channels, using some channel Λ' after a channel Λ compatible with A produces a new channel compatible with A (upper figure). The set $\mathfrak{C}_\mathsf{A}$ always contains completely depolarizing channels Γ (lower figure), for which $\Gamma(\varrho) = \varrho_0$ for some fixed state ϱ_0.

Proposition 1. For any unitary operators V and W, the following holds: A channel Λ is compatible with an observable A if and only if the channel $\widetilde{W} \circ \Lambda \circ \widetilde{V}$ is compatible with the observable $V^*\mathsf{A}V$.

Proof. Suppose that $x \mapsto \Phi_x$ is an instrument such that Λ and A satisfy (4). In that case, the instrument $x \mapsto \Phi'_x$, defined as $\Phi'_x(\varrho) = W\Phi_x(V\varrho V^*)W^*$, demonstrates the compatibility of $\widetilde{W} \circ \Lambda \circ \widetilde{V}$ and $V^*\mathsf{A}V$. We can run the same argument for the inverse operators V^* and W^*; hence the converse holds also. ∎

B. Channels compatible with given observable

Every observable A has a collection of compatible channels, denoted by $\mathfrak{C}_\mathsf{A}$. The set $\mathfrak{C}_\mathsf{A}$ specifies what kinds of perturbations are possible when A is measured. There are many ways (by means of measurement models, or instruments) to measure A, and for this reason $\mathfrak{C}_\mathsf{A}$ contains many channels. We will limit $\mathfrak{C}_\mathsf{A}$ to channels that have the same input and output spaces, although generally one could allow arbitrary output spaces [5]. For each observable A, the set $\mathfrak{C}_\mathsf{A}$

  (i) is *convex*, i.e., if Λ_1 and Λ_2 are compatible with A, then also all their mixtures $t\Lambda_1 + (1-t)\Lambda_2, 0 < t < 1$, are compatible with A;

  (ii) is a *left ideal* of the set of all channels (see Fig. 2), i.e., if Λ is compatible with A and Λ' is any other channel, then their concatenation $\Lambda' \circ \Lambda$ is compatible with A as well;

  (iii) contains all *completely depolarizing channels* (see Fig. 2) $\varrho \mapsto \varrho_0$, where ϱ_0 is an arbitrary fixed state;

  (iv) contains the *Lüders channel* $\mathcal{L}_\mathsf{A}$ of A, which is defined as $\mathcal{L}_\mathsf{A}(\varrho) = \sum_x \sqrt{\mathsf{A}(x)}\varrho\sqrt{\mathsf{A}(x)}$.

The following result, proved in [5], characterizes the set $\mathfrak{C}_\mathsf{A}$ completely.

Theorem 1. There is a channel $\Lambda_\mathsf{A}: \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ and a Hilbert space $\mathcal{K}$ such that every channel Λ compatible with A is of the form $\Lambda = \Lambda' \circ \Lambda_\mathsf{A}$, where $\Lambda': \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H})$ is some channel.

The "mother channel" mentioned in Theorem 1 is given as

$$\Lambda_{\mathsf{A}}(\varrho) = \sum_x \hat{\mathsf{A}}(x) T \varrho T^* \hat{\mathsf{A}}(x), \tag{7}$$

where $(\mathcal{K}, \hat{\mathsf{A}}, T)$ is some minimal Naimark dilation of A, i.e., $T : \mathcal{H} \to \mathcal{K}$ is an isometry, $\hat{\mathsf{A}}$ is a projection-valued measure (PVM) on $\mathcal{K}$, and $\mathsf{A}(x) = T^* \hat{\mathsf{A}}(x) T$ for all $x \in \Omega_{\mathsf{A}}$. In a finite-dimensional case we can write a concrete form of the minimal Naimark dilation as follows. We fix a spectral decomposition for each operator $\mathsf{A}(x)$,

$$\mathsf{A}(x) = \sum_{k=1}^{r_x} |\phi_{x,k}\rangle\langle\phi_{x,k}|, \tag{8}$$

where r_x is the rank of $\mathsf{A}(x)$. We then choose $\mathcal{K} = \mathbb{C}^{r_1} \oplus \cdots \oplus \mathbb{C}^{r_n}$ and we fix orthonormal bases $\{e_{x,k}\}_{k=1}^{r_x}$ for each $\mathbb{C}^{r_x}$. We define a linear map $T : \mathcal{H} \to \mathcal{K}$ as $T\psi = \sum_{x,k} \langle \phi_{x,k} | \psi \rangle e_{x,k}$. Its adjoint $T^* : \mathcal{K} \to \mathcal{H}$ is given as $T^* e_{x,k} = \phi_{x,k}$. The sharp observable $\hat{\mathsf{A}}$ that dilates A is

$$\hat{\mathsf{A}}(x) = \sum_{k=1}^{r_x} |e_{x,k}\rangle\langle e_{x,k}|, \tag{9}$$

and hence we obtain

$$\Lambda_{\mathsf{A}}(\varrho) = \sum_x \sum_{k,\ell}^{r_x} \langle \phi_{x,k} | \varrho \phi_{x,\ell} \rangle |e_{x,k}\rangle\langle e_{x,\ell}|. \tag{10}$$

By Theorem 1 any $\Lambda \in \mathfrak{C}_{\mathsf{A}}$ can thus be written as

$$\Lambda(\varrho) = \sum_x \sum_{k,\ell}^{r_x} \langle \phi_{x,k} | \varrho \phi_{x,\ell} \rangle \Lambda'(|e_{x,k}\rangle\langle e_{x,\ell}|) \tag{11}$$

for some channel $\Lambda' : \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H})$.

C. Observables compatible with a given channel

Let us look at the converse to the previous consideration; we fix a channel Λ and denote by $\mathfrak{D}_\Lambda$ the set of all observables compatible with Λ. Also the set $\mathfrak{D}_\Lambda$ has some elementary properties, namely,

(i) $\mathfrak{D}_\Lambda$ is *convex*, i.e., if A_1 and A_2 are compatible with Λ, then also all their mixtures $t\mathsf{A}_1 + (1-t)\mathsf{A}_2, 0 < t < 1$ are compatible with Λ,

(ii) $\mathfrak{D}_\Lambda$ is *closed under postprocessing*, i.e., if $\mathsf{A} \in \mathfrak{D}_\Lambda$, then also $\mu \circ \mathsf{A} \in \mathfrak{D}_\Lambda$ for all postprocessings μ. A postprocessing μ is given by the formula

$$(\mu \circ \mathsf{A})(x) = \sum_y \mu_{xy} \mathsf{A}(y), \tag{12}$$

where μ_{xy} is a stochastic matrix,

(iii) $\mathfrak{D}_\Lambda$ contains all *trivial observables* T, $\mathsf{T}(x) = p(x)\mathbb{1}$ for some probability distribution p.

The structure of $\mathfrak{D}_\Lambda$ can be inferred from the results presented in [11]. However, we find it useful to give a self-contained derivation of the characterization of $\mathfrak{D}_\Lambda$. To formulate it, we recall that any channel $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ has a Stinespring dilation $(\mathcal{K}, V)$, where $V : \mathcal{H} \to \mathcal{H} \otimes \mathcal{K}$ is an isometry and $\Lambda(\varrho) = \text{tr}_{\mathcal{K}}[V\varrho V^*]$. The dilation also gives another channel $\bar{\Lambda} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ by formula $\bar{\Lambda}(\varrho) = \text{tr}_{\mathcal{H}}[V\varrho V^*]$. This channel is called a *conjugate channel* (or
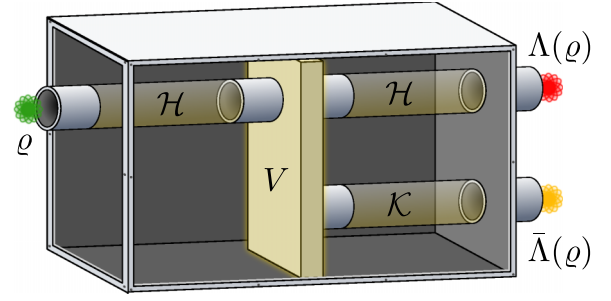

FIG. 3. Every channel Λ has a Stinespring dilation expanding the Hilbert space $\mathcal{H}$ to $\mathcal{H} \otimes \mathcal{K}$ by using an isometry V. In this way we can define the conjugate channel $\bar{\Lambda}$ of Λ.

complementary channel) of Λ (see Fig. 3). We further say that a conjugate channel of Λ is minimal if it is related to a minimal Stinespring dilation of Λ.

Theorem 2. Let $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be a channel. Observable A is compatible with channel Λ if and only if the observable can be written in the form $\mathsf{A}(x) = \bar{\Lambda}^*(\mathsf{A}'(x))$, where $\bar{\Lambda}^* : \mathcal{L}(\mathcal{K}) \mapsto \mathcal{L}(\mathcal{H})$ is a fixed (minimal) conjugate channel of Λ in the Heisenberg picture and A' is some observable on $\mathcal{K}$.

Proof. Assume that A is compatible with Λ. By the definition, this means that the conditions (1) and (2) hold for some $\mathcal{H}_a, U, \xi$, and F, where ξ is a pure state, i.e., $\xi = |\phi\rangle\langle\phi|$. We then define an operator $W : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}_a$ as

$$W\psi = U(\psi \otimes \phi). \tag{13}$$

This operator satisfies

$$W^*W = \mathbb{1}_{\mathcal{H}} \quad \text{and} \quad W\varrho W^* = U(\varrho \otimes \xi)U^* \tag{14}$$

for all states ϱ on $\mathcal{H}$. We conclude that $(\mathcal{H}_a, W)$ is a Stinespring dilation of Λ and

$$\mathsf{A}(x) = W^*(\mathbb{1}_{\mathcal{H}} \otimes \mathsf{F}(x))W. \tag{15}$$

Let us then fix the minimal Stinespring dilation $(\mathcal{K}, V)$ for the channel Λ. From the minimality it follows that $W = (\mathbb{1}_{\mathcal{H}} \otimes V')V$, where $V' : \mathcal{K} \to \mathcal{H}_a$ is an isometry. Therefore,

$$\mathsf{A}(x) = V^*(\mathbb{1}_{\mathcal{H}} \otimes V'^* \mathsf{F}(x)V')V. \tag{16}$$

Denoting $\mathsf{A}'(x) = V'^* \mathsf{F}(x)V$ we obtain the claimed form.

Conversely, assume that A' is some observable on $\mathcal{K}$ and $\mathsf{A}(x) = \bar{\Lambda}^*(\mathsf{A}'(x))$ for some (minimal) Stinespring dilation $(\mathcal{K}, V)$ of Λ. Then we define an instrument

$$\Phi_x(\varrho) = \text{tr}_{\mathcal{K}}[V\varrho V^*(\mathbb{1}_{\mathcal{H}} \otimes \mathsf{A}'(x))]. \tag{17}$$

This instrument fulfills the conditions in (4), hence showing that A and Λ are compatible. ∎

Let us note that the condition on the minimality of the dilation in Theorem 2 is not necessary. However, we found it convenient to use a concrete form of the minimal Stinespring dilation $(\mathcal{K}, V)$ of Λ. In particular, fix an orthonormal basis $\{e_k\}_{k=1}^N$ for $\mathcal{K}$ and for each k we define an operator $M_k \in \mathcal{L}(\mathcal{H})$ via

$$\langle \psi | M_k \varphi \rangle = \langle \psi \otimes e_k | V\varphi \rangle, \quad \psi, \varphi \in \mathcal{H}. \tag{18}$$

Then the operators M_k form a minimal set of Kraus operators for the channel Λ and we obtain

$$\mathsf{A}(x) = \sum_{k,l} \langle e_k \mid \mathsf{A}'(x) e_l \rangle M_k^* M_l. \tag{19}$$

Here again A' is some observable on $\mathcal{K}$.

III. INCOMPATIBILITY OF UNBIASED QUBIT OBSERVABLES AND PAULI CHANNELS

A. Unbiased qubit observables and Pauli channels

An observable A, acting on a Hilbert space $\mathcal{H}$, is called *unbiased* if it maps the maximally mixed state $\frac{1}{d}\mathbb{1}$ to the uniform probability distribution of outcomes x, i.e., $\mathrm{tr}[\mathsf{A}(x)] = \frac{d}{n}$, where n is the number of outcomes and d is the dimension of $\mathcal{H}$. In the case of qubit observables the unbiasedness condition implies that effects have the form $\mathsf{A}(x) = \frac{1}{n}[\mathbb{1} + \vec{a}(x) \cdot \vec{\sigma}]$, where $\vec{a}(x) \in \mathbb{R}^3$ and $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli operators. In what follows we will use the notation $\sigma_0 = \mathbb{1}$.

In particular, for binary (i.e., two outcome) qubit observables the unbiasedness condition means that observables are of the form

$$\mathsf{A}_{s,\vec{n}}(\pm) = \tfrac{1}{2}(\mathbb{1} \pm s\vec{n} \cdot \vec{\sigma}) \tag{20}$$

for $s \in [0,1]$ and $\vec{n} \in \mathbb{R}^3$, $\|\vec{n}\| = 1$. We will also use the notation X_s, Y_s, and Z_s for observables $\mathsf{A}_{s,\vec{n}}$ with $\vec{n} = (1,0,0)$, $\vec{n} = (0,1,0)$, and $\vec{n} = (0,0,1)$, respectively.

We notice that an observable $\mathsf{A}_{t,\vec{n}}$ is a postprocessing of another observable $\mathsf{A}_{s,\vec{n}}$ if and only if $t \leqslant s$. Namely, if $s \neq 0$, then for any $t \in [0,1]$ the operator $\mathsf{A}_{t,\vec{n}}(+)$ can be written as a linear combination of $\mathsf{A}_{s,\vec{n}}(+)$ and $\mathsf{A}_{s,\vec{n}}(-)$ in a unique way:

$$\mathsf{A}_{t,\vec{n}}(+) = \frac{s+t}{2s}\mathsf{A}_{s,\vec{n}}(+) + \frac{s-t}{2s}\mathsf{A}_{s,\vec{n}}(-). \tag{21}$$

This is a valid postprocessing if and only if

$$0 \leqslant (s \pm t)/2s \leqslant 1, \tag{22}$$

which is equivalent to $t \leqslant s$. We can therefore interpret the parameter s as the degree of noise inherent in $\mathsf{A}_{s,\vec{n}}$.

Let us note that the set of effects $\mathsf{A}_{s,\vec{n}}(\pm)$ (thus also the set of unbiased binary qubit observables) is convex. Indeed, the effects $\mathsf{A}_{s,\vec{n}}(\pm)$ are positive operators of unit trace; hence, they formally correspond to density operators and as such, can be visualized as points inside the Bloch ball. An unbiased binary qubit observable thus corresponds to a pair of points inside the Bloch ball, and the points are symmetric with respect to the origin.

Using the analogy with observables, we say that a channel Λ is unbiased if it keeps the maximally mixed state invariant, i.e., $\Lambda(\frac{1}{d}\mathbb{1}) = \frac{1}{d}\mathbb{1}$. This property is obviously equivalent with unitality; thus, the notion of unbiased channels is just a synonym for *unital channels*. In the case of qubits it is further known that the set of unital channels coincides with the set of *random unitary channels* [12,13]. A prominent class of random unitary qubit channels are the so-called Pauli channels, and in the following we shall concentrate on that class.
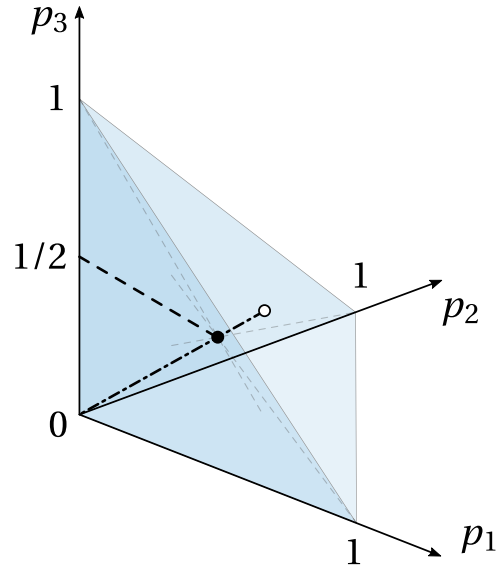

FIG. 4. Pauli channels can be parametrized by the points inside the tetrahedron. The identity channel corresponds to the origin. The three Pauli unitaries correspond to the remaining three vertices of the tetrahedron. The edges correspond to Pauli channels having two of the probabilities zero. Special classes of Pauli channels are: (1) the class of partially depolarizing channels (dot-dashed line), including the totally depolarizing channel (solid point) and the quantum NOT (hollow point); (2) the class of measure-and-prepare channels (dashed line), here shown for measurements in the z direction.

A *Pauli channel* $\Psi_{\vec{p}}$ is a qubit channel of the form

$$\Psi_{\vec{p}}(\varrho) = \sum_{j=0}^{3} p_j \, \sigma_j \varrho \sigma_j, \tag{23}$$

where $\vec{p} \in \mathbb{R}^4$ is a probability vector, i.e., $0 \leqslant p_j \leqslant 1$ and $\sum_{j=0}^{3} p_j = 1$. Due to the normalization of $\vec{p}$, a Pauli channel $\Psi_{\vec{p}}$ is determined already by three of the components, e.g., p_1, p_2, p_3. We can therefore visualize the set of Pauli channels as a tetrahedron in $\mathbb{R}^3$; see Fig. 4. We denote by Γ the completely depolarizing channel on the maximally mixed state $\frac{1}{2}\mathbb{1}$, and it corresponds to the probability vector $\vec{p} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

B. Compatibility condition

Let A be a qubit observable compatible with a Pauli channel $\Psi_{\vec{p}}$. By concatenating $\Psi_{\vec{p}}$ with a Pauli unitary channel we generate three new Pauli channels that are compatible with A. Using Proposition 1 with $W = \sigma_i$ and $U = \mathbb{1}$ we conclude the following:

Proposition 2. Let A be a qubit observable compatible with a Pauli channel $\Psi_{\vec{p}}$, $\vec{p} = (p_0, p_1, p_2, p_3)$. Then A is also compatible with Pauli channels with the following probability vectors:

(i) (p_1, p_0, p_3, p_2),

(ii) (p_2, p_3, p_0, p_1),

(iii) (p_3, p_2, p_1, p_0).

In conclusion, for a fixed qubit observable A, the probability vectors $\vec{p} = (p_0, p_1, p_2, p_3)$ that correspond to Pauli channels $\Psi_{\vec{p}}$ compatible with the observable form a convex region
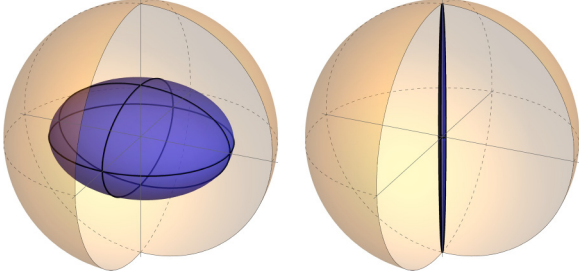
FIG. 5. Compatibility region for Pauli channel $\Psi_{\vec{p}}$ and observable $\mathsf{A}_{s,\vec{n}}$ is in general described as an ellipsoid for allowed Bloch vectors of the observable (left figure). In specific cases it collapses to a line (right figure), as e.g., in the case of phase-damping channels.

FIG. 6. Compatibility region for Pauli channel $\Psi_{\vec{p}}$ and observable X_s with $s = 0.8$ for allowed $\vec{p}$ vectors of the Pauli channel. Here, unlike in Fig. 4, the region of vectors $\vec{p}$ is depicted as a simplex. The darkened areas are parts laying on the faces of the simplex, i.e., when (at least) one of the p_j's equals zero.

inside of the tetrahedron in Fig. 4, and this region has the permutational symmetry described in Proposition 2.

Let then Λ be a qubit channel compatible with an unbiased qubit observable $\mathsf{A}_{s,\vec{n}}$. As we have seen earlier, any observable $\mathsf{A}_{t,\vec{n}}$ with $t \leqslant s$ is a postprocessing of $\mathsf{A}_{s,\vec{n}}$. It follows that also $\mathsf{A}_{t,\vec{n}}$ is compatible with Λ. For any unit vector $\vec{n}$, it thus makes sense to seek for the largest s such that $\mathsf{A}_{s,\vec{n}}$ and Λ are compatible.

The next result gives a sufficient and necessary condition for an unbiased binary qubit observable $\mathsf{A}_{s,\vec{n}}$ and Pauli channel $\Psi_{\vec{p}}$ to be (in)compatible. The compatibility properties of a Pauli channel $\Psi_{\vec{p}}$ are determined by the vector $\vec{p}$, and to formulate the condition we denote

$$
\begin{aligned}
p_{\pm}[1] &:= 2(\sqrt{p_0 p_1} \pm \sqrt{p_2 p_3}), \\
p_{\pm}[2] &:= 2(\sqrt{p_0 p_2} \pm \sqrt{p_1 p_3}), \\
p_{\pm}[3] &:= 2(\sqrt{p_0 p_3} \pm \sqrt{p_1 p_2}).
\end{aligned}
\tag{24}
$$

We observe that $|p_-[j]| \leqslant p_+[j]$ for $j = 1,2,3$. Furthermore all $p_j[+]$ are invariant under the permutations described in Proposition 2.

Theorem 3. An unbiased binary qubit observable $\mathsf{A}_{s,\vec{n}}$ and a Pauli channel $\Psi_{\vec{p}}$ are compatible if and only if

$$
\frac{s^2 n_1^2}{p_+[1]^2} + \frac{s^2 n_2^2}{p_+[2]^2} + \frac{s^2 n_3^2}{p_+[3]^2} \leqslant 1 .
\tag{25}
$$

This inequality is understood in a way that if $p_+[j] = 0$, then necessarily the whole term vanishes.

Before we present the proof of this statement, let us discuss the content of inequality (25). First, suppose that a Pauli channel $\Psi_{\vec{p}}$ is fixed and that $p_+[j] \neq 0$ for every $j = 1,2,3$. For a vector $s\vec{n}$, the inequality (25) is a solid ellipsoid (see Fig. 5). Second, for at least one of $p_+[j]$ to be zero, we need to have at least two of the components of $\vec{p}$ zero, which always makes at least two of $p_+[j]$ zero. In such cases the inequality (25) does not represent a solid ellipsoid but only a line in one of the canonical directions. In the most extreme case we have $p_+[j] = 0$ for all $j = 1,2,3$, which occurs when only one of the p_j's is nonzero. Then the ellipsoid collapses to a point $s\vec{n} = \vec{0}$. This equation is satisfied only when $s = 0$, and this is consistent with our earlier discussion that only a trivial observable is compatible with a unitary channel.
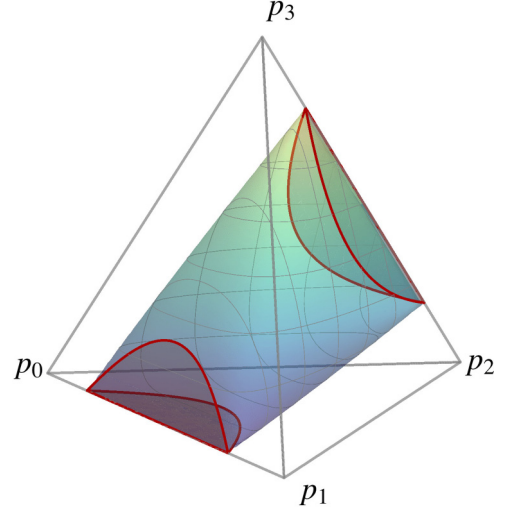
While for constant Pauli channel the set of compatible unbiased qubit observables is rather simply visualized as an ellipsoid, the region of Pauli channels compatible with some given unbiased qubit observable that we get from (25) is less illustrative. Clearly, for trivial observable ($s = 0$) the compatibility region is the whole set of Pauli channels (up to unitary rotations forming a tetrahedron). For sharp unbiased observables ($s = 1$) the compatibility region contains only one element—the completely depolarizing channel ($p_0 = p_1 = p_2 = p_3 = 1/4$). For an intermediate value of $s = 0.8$ the compatibility region within the Pauli channels is illustrated in Fig. 6.

C. Proof of Theorem 3

Let $\Psi_{\vec{p}}$ be a Pauli channel. In the following we will assume that $p_j \neq 0$ for every $j = 0, \ldots, 3$. The required modifications to the proof in the other cases shall be obvious.

The minimal set of Kraus operators for $\Psi_{\vec{p}}$ is given as $M_k = \sqrt{p_k}\sigma_k$, $k = 0,1,2,3$. Using Theorem 2 and the formula (19) derived from it, we obtain observables compatible with $\Psi_{\vec{p}}$ by inserting various choices for A', which is an observable acting on $\mathbb{C}^4$.

We fix a unit vector $\vec{n}$ and we seek for allowed s such that $\mathsf{A}_{s,\vec{n}}$ and $\Psi_{\vec{p}}$ are compatible. It is useful to define a vector $\vec{n}' \in \mathbb{R}^3$ as

$$
n'_j = \frac{n_j}{p_+[j]} \left(\frac{n_1^2}{p_+[1]^2} + \frac{n_2^2}{p_+[2]^2} + \frac{n_3^2}{p_+[3]^2} \right)^{-1/2},
\tag{26}
$$

and then an operator $\mathsf{A}'(+)$ as

$$
\mathsf{A}'(+) = \frac{1}{2}
\begin{pmatrix}
1 & n'_1 & n'_2 & n'_3 \\
n'_1 & 1 & -in'_3 & in'_2 \\
n'_2 & in'_3 & 1 & -in'_1 \\
n'_3 & -in'_2 & in'_1 & 1
\end{pmatrix},
\tag{27}
$$

and set $A'(-) = \mathbb{1} - A'(+)$. Since $\vec{n}'$ is a unit vector, one can verify that $A'(\pm)$ are projections and hence form a binary observable. Applying Eq. (19) we then get

$$\sum_{k,l} \langle e_k | A'(x) e_l \rangle \sqrt{p_k p_l} \sigma_k \sigma_l$$

$$= \frac{1}{2} \left[\mathbb{1} + \left(\frac{n_1^2}{p_+[1]^2} + \frac{n_2^2}{p_+[2]^2} + \frac{n_3^2}{p_+[3]^2} \right)^{-\frac{1}{2}} (\vec{n} \cdot \vec{\sigma}) \right].$$
(28)

This means that $A_{s,\vec{n}}$ and $\Psi_{\vec{p}}$ are compatible for

$$s = s_{\max} \equiv \left(\frac{n_1^2}{p_+[1]^2} + \frac{n_2^2}{p_+[2]^2} + \frac{n_3^2}{p_+[3]^2} \right)^{-\frac{1}{2}},$$
(29)

and hence also for any value $s \leqslant s_{\max}$. We have thus seen that the inequality (25) is a sufficient condition for the compatibility of $A_{s,\vec{n}}$ and $\Psi_{\vec{p}}$.

In order to prove that the inequality (25) is also necessary for the compatibility of $A_{s,\vec{n}}$ and $\Psi_{\vec{p}}$, we will formulate the problem in terms of a semidefinite program (SDP). Any feasible instance of primal SDP problem will give us a lower bound on the largest possible s, while any feasible instance of the dual SDP problem will give us an upper bound on the largest possible s. If both bounds coincide, we have found the optimal solution with the largest possible s. And this will be indeed the case.

1. Primal SDP problem

For given vectors $\vec{n}$ and $\vec{p}$, we want to find the largest s such that $A_{s,\vec{n}}$ and $\Psi_{\vec{p}}$ are compatible. Since $A_{s,\vec{n}}(-) = \mathbb{1} - A_{s,\vec{n}}(+)$, we can formulate the question in terms of the effect $A_{s,\vec{n}}(+)$ only. Further, we will now understand $\vec{n}$ and s as parameters of the SDP problem and will omit the subscripts from now on. We will thus denote $A \equiv A_{s,\vec{n}}(+)$, and we observe that $s = \mathrm{tr}[A(\vec{n} \cdot \vec{\sigma})]$.

In summary, we are trying to find the maximum of $\mathrm{tr}[A(\vec{n} \cdot \vec{\sigma})]$ over all effects A such that the corresponding binary observable $A_{s,\vec{n}}$ is compatible with $\Psi_{\vec{p}}$. Using Theorem 2, this is equivalent to

$$\max_{\mathbb{1} \geqslant A' \geqslant 0} \mathrm{tr}[\overline{\Psi}_{\vec{p}}^*(A')(\vec{n} \cdot \vec{\sigma})],$$
(30)

where A' is an effect on the Hilbert space defined by the minimal conjugate channel $\overline{\Psi}_{\vec{p}}^*$ of $\Psi_{\vec{p}}$. We take the conjugate channel that is related to the Kraus operators $M_k = \sqrt{p_k} \sigma_k$ of $\Psi_{\vec{p}}$ and denote $\Sigma_i = \overline{\Psi}_{\vec{p}}(\sigma_i)$. Then

$$\Sigma_i = \sum_{j,k,n=0}^{3} \sqrt{p_k p_n} \mathrm{tr}[\sigma_k \sigma_i \sigma_n] |e_k\rangle \langle e_n|,$$
(31)

and we obtain

$$\Sigma_1 = 2 \begin{pmatrix} 0 & \sqrt{p_0 p_1} & 0 & 0 \\ \sqrt{p_0 p_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & -i\sqrt{p_2 p_3} \\ 0 & 0 & i\sqrt{p_2 p_3} & 0 \end{pmatrix},$$
(32)

$$\Sigma_2 = 2 \begin{pmatrix} 0 & 0 & \sqrt{p_0 p_2} & 0 \\ 0 & 0 & 0 & i\sqrt{p_1 p_3} \\ \sqrt{p_0 p_2} & 0 & 0 & 0 \\ 0 & -i\sqrt{p_1 p_3} & 0 & 0 \end{pmatrix},$$
(33)

$$\Sigma_3 = 2 \begin{pmatrix} 0 & 0 & 0 & \sqrt{p_0 p_3} \\ 0 & 0 & -i\sqrt{p_1 p_2} & 0 \\ 0 & i\sqrt{p_1 p_2} & 0 & 0 \\ \sqrt{p_0 p_3} & 0 & 0 & 0 \end{pmatrix}.$$
(34)

We can finally write the primal SDP problem as

$$s_{\mathrm{P}} := \max_{A'} \mathrm{tr}[A'(\vec{n} \cdot \vec{\Sigma})]$$
(35)

$$\text{subject to} \quad -A' \leqslant 0,$$
$$A' - \mathbb{1} \leqslant 0,$$
$$\mathrm{tr}[A'(\vec{n}_1 \cdot \vec{\Sigma})] = 0,$$
$$\mathrm{tr}[A'(\vec{n}_2 \cdot \vec{\Sigma})] = 0.$$

The vectors $\vec{n}_{1,2}$ are orthogonal to $\vec{n}$ and linearly independent. The last two constraints ensure that the resulting effect A is indeed in the direction defined by $\vec{n}$.

The solution given in (27) is a feasible solution for this primal SDP problem and it thus means that $s_{\mathrm{P}} \geqslant s_{\max}$, where $s_{\max}$ is given in (29).

2. Dual SDP problem

The previous convex optimization problem can be transformed into a dual problem by standard methods [14] (see Appendix):

$$s_{\mathrm{D}} := \min_{\lambda, \vec{m}} \mathrm{tr}[\lambda]$$
(36)

$$\text{subject to} \quad \lambda \geqslant 0,$$
$$\lambda \geqslant \vec{m} \cdot \vec{\Sigma},$$
$$\vec{m} \cdot \vec{n} = 1.$$

It is always true that $s_{\mathrm{D}} \geqslant s_{\mathrm{P}}$. Since we have already shown that $s_{\mathrm{P}} \geqslant s_{\max}$, the remaining thing is to show that $s_{\max} \geqslant s_{\mathrm{D}}$. We do this by providing a corresponding feasible solution for the dual SDP problem.

We denote by Q the diagonal matrix with entries $Q_{ij} = \delta_{ij} p_i[+]$. With this notation the vector $\vec{n}'$ in (26) can be concisely written as $\vec{n}' = Q^{-1} \vec{n} / \|Q^{-1} \vec{n}\|$. We then set

$$\vec{m} = \frac{Q^{-2} \vec{n}}{\|Q^{-1} \vec{n}\|^2}, \qquad \lambda = A'(\vec{m} \cdot \vec{\Sigma}) A',$$
(37)

where $A' = A'(+)$ is given in (27). This choice of $\vec{m}$ fulfills the third dual condition and, moreover, we have

$$\mathrm{tr}[\lambda] = \mathrm{tr}[A'(\vec{m} \cdot \vec{\Sigma}) A'] = \mathrm{tr}[A'(\vec{m} \cdot \vec{\Sigma})]$$
$$= \mathrm{tr}[A'(\vec{n} \cdot \vec{\Sigma})] = s_{\max},$$
(38)

where we first used the definition of λ, then in the next equality we used that $(A')^2 = A'$, and in the third equality we used the fact that the trace for components of $\vec{m}$ orthogonal to $\vec{n}$ is zero.

It remains to show that $\lambda \geqslant 0$ and $\lambda \geqslant \vec{m} \cdot \vec{\Sigma}$. For this we will rewrite λ in the basis of eigenvectors of A'. The operator A' is a two-dimensional projection, and hence it has doubly

degenerate eigenvalues 1 and 0. The eigenvalue 1 eigenvectors can be chosen to be

$$v_\pm = \frac{1}{2\sqrt{1 \pm n_1'}} \begin{pmatrix} n_2' \pm i n_3' \\ \mp n_2' - i n_3' \\ 1 \pm n_1' \\ \pm i(1 \pm n_1') \end{pmatrix}. \qquad (39)$$

The eigenvalue 0 eigenvectors can be chosen to be

$$u_\pm = \frac{1}{2\sqrt{1 \pm n_1'}} \begin{pmatrix} -n_2' \pm i n_3' \\ \mp n_2' + i n_3' \\ 1 \pm n_1' \\ \mp i(1 \pm n_1') \end{pmatrix}. \qquad (40)$$

Note that these eigenvectors are valid when $n_1' \neq \pm 1$; the excluded points have a different eigenbasis, however the remainder of the solution is analogous to what follows. The eigenbasis (v_+, v_-, u_+, u_-) is orthonormal and the operator $\vec{m} \cdot \vec{\Sigma}$ in this basis has a block-diagonal form,

$$\vec{m} \cdot \vec{\Sigma} = \begin{pmatrix} M & 0 \\ 0 & -M^* \end{pmatrix}, \qquad (41)$$

where

$$M = \frac{s}{2}(\mathbb{1} + \vec{g} \cdot \vec{\sigma}) \qquad (42)$$

is a qubit operator with

$$\vec{g} = \frac{1}{\sqrt{1 - n_1'^2}} \begin{pmatrix} \frac{p_-[2]}{p_+[2]} n_2'^2 - \frac{p_-[3]}{p_+[3]} n_3'^2 \\ \left(\frac{p_-[2]}{p_+[2]} + \frac{p_-[3]}{p_+[3]}\right) n_2' n_3' \\ -n_1' \sqrt{1 - n_1'^2} \frac{p_-[1]}{p_+[1]} \end{pmatrix}. \qquad (43)$$

Hence we have

$$\lambda = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix}, \quad \lambda - \vec{m} \cdot \vec{\Sigma} = \begin{pmatrix} 0 & 0 \\ 0 & M^* \end{pmatrix}. \qquad (44)$$

Therefore we need to check only the positivity of M. The positivity of M is in this case equivalent to the condition $\|\vec{g}\| \leqslant 1$. Using the fact that $(p_-[j]/p_+[j])^2 \leqslant 1$, this is easily checked.

To sum up, we have shown that our choices for primary and dual variables are feasible solutions that lead to the same values. Hence, this choice is optimal and the boundary given by $s_{\max}$ given in (29) is not only sufficient but also necessary.

IV. EXAMPLES

We will now demonstrate the use of the presented compatibility condition by looking at some concrete classes of qubit channels.

A. Partially depolarizing channels

A partially depolarizing channel is an example of a Pauli channel. It is constructed as a mixture of the identity channel and the completely depolarizing channel Γ to the maximally mixed state $\frac{1}{2}\mathbb{1}$. As the mixing weight can vary, we get a one-parameter class of channels

$$\Gamma_p(\varrho) = (1 - 4p)\varrho + 2p\mathbb{1}, \qquad (45)$$

where $p \in [0, 1/4]$. The channel Γ_p is, in fact, a Pauli channel with the probability vector $(1 - 3p, p, p, p)$, see also Fig. 4. Actually, the map Γ_p defined in (45) is a valid channel for any $p \in [0, 1/3]$, although the interpretation as a partially depolarizing channel holds only for $p \in [0, 1/4]$. Further, we can start from any unitary Pauli channel instead of the identity channel; however, in that case the depolarization is with respect to a different basis.

For a channel Γ_p we have

$$p_+[j] = 2[p + \sqrt{p(1 - 3p)}] \qquad (46)$$

for every $j = 1, 2, 3$. From Theorem 3 we conclude that the set of unbiased qubit observables compatible with Γ_p corresponds to the shrunken Bloch ball with the radius $2(p + \sqrt{p(1 - 3p)})$. The identity channel $id = \Gamma_0$ shrinks the compatibility Bloch ball region to the central point, while the completely depolarizing channel $\Gamma = \Gamma_{1/4}$ keeps the Bloch ball invariant. This is consistent with our earlier observations.

An interesting special case is the universal quantum NOT channel, which transforms any qubit input state to as close as possible to its orthogonal complement [15]. This operation cannot be perfect for any input and for general input is described by a Pauli channel falling under the case presented in this section, where $p = 1/3$; in this case $s \leqslant 2/3$.

B. Phase-damping channels and Lüder's channels

A *phase-damping channel* is a map that damps the off-diagonal elements of a density matrix in a specific basis. We fix the basis to be the eigenbasis of σ_3. The action of a phase-damping channel Φ_p is then such that in the σ_3 eigenbasis the density matrices retain their diagonals, but the off-diagonal elements acquire a factor of $2p - 1$. Thus, we have a one-parameter class of Pauli channels Φ_p and the corresponding probability vector is $(p, 0, 0, 1 - p)$. For $p \in [1/2, 1]$ the action of the channel describes pure damping, while for $p \in [0, 1/2]$ the damping is complemented with the inversion of the off-diagonal elements. The extreme case $p = 0$ corresponds to the inversion in the xy plane without any damping.

For a phase-damping channel Φ_p we have $p_+[1] = p_+[2] = 0$ and $p_+[3] = 2\sqrt{p(1 - p)}$. Therefore, using Theorem 3, we conclude that an observable $\mathsf{A}_{s, \vec{n}}$ is compatible with Φ_p if and only if $\vec{n} = (0, 0, 1)$ and $s \leqslant 2\sqrt{p(1 - p)}$.

Specific cases are the identity channel ($p = 1$) and the NOT channel ($p = 0$) for which $s = 0$. On the other end lies the case of $p = 1/2$, which is the completely phase-damping channel that zeros all off-diagonal elements and conserves the diagonal which contains all the information about z direction; this means that all z measurements are compatible with this channel ($s \leqslant 1$)—see Fig. 5 on the right for this example.

An interesting class of channels falling into this category are Lüder's channels of Z_t, given as

$$\mathcal{L}_{\mathsf{Z}_t}(\varrho) = \sqrt{\mathsf{Z}_t(+)}\varrho\sqrt{\mathsf{Z}_t(+)} + \sqrt{\mathsf{Z}_t(-)}\varrho\sqrt{\mathsf{Z}_t(-)}. \qquad (47)$$

This is a phase damping channel Φ_p in the σ_3-eigenbasis with $p = \frac{1}{2}(\sqrt{1 - t^2} + 1)$. One direct consequence is hence that an observable X_s is compatible with the Lüder's channel $\mathcal{L}_{\mathsf{Z}_t}$ if and only if $s = 0$. This result stands in contrast to the compatibility at the level of observables, as X_s and Z_t are compatible if

and only if $s^2 + t^2 \leqslant 1$ [16]. This means that if we want to implement a joint measurement of X_s and Z_t with $s, t \neq 0$, the joint measurement process cannot contain a Lüders channel $\mathcal{L}_{\mathsf{Z}_t}$ or $\mathcal{L}_{\mathsf{X}_s}$. We remark that it has been earlier shown that a joint measurement process cannot contain both $\mathcal{L}_{\mathsf{Z}_t}$ and $\mathcal{L}_{\mathsf{X}_s}$ [17]. Our new result hence strengthens that observation.

C. Measure-and-prepare channels

Let us consider the class of measure-and-prepare channels related to observables Z_t, $0 \leqslant t \leqslant 1$. For each $0 \leqslant t \leqslant 1$, we define a map Θ_t as

$$\Theta_t(\varrho) = \mathrm{tr}[\varrho \mathsf{Z}_t(+)]\mathsf{Z}_1(+) + \mathrm{tr}[\varrho \mathsf{Z}_t(-)]\mathsf{Z}_1(-). \quad (48)$$

This is a measure-and-prepare channel, which can be implemented by first measuring the observable Z_t and then preparing either pure state $\mathsf{Z}_1(+)$ or $\mathsf{Z}_1(-)$, depending on the outcome of the measurement. It is easy to see that Θ_t is unital for any t. Also, this channel is a composition of a partially depolarizing channel and a completely phase-damping channel in z direction. Overall, Θ_t is a Pauli channel defined by the probability vector

$$\vec{p} = \tfrac{1}{4}(1 + t, 1 - t, 1 - t, 1 + t). \quad (49)$$

See also Fig. 4.

When considering the compatibility of Θ_t with an observable $\mathsf{A}_{s,\vec{n}}$, Theorem 3 gives

$$\frac{n_1^2 + n_2^2}{1 - t^2} + n_3^2 \leqslant \frac{1}{s^2}. \quad (50)$$

After some manipulation we get

$$s^2 + t^2 - s^2 t^2 \cos^2 \vartheta \leqslant 1, \quad (51)$$

where ϑ is the angle between the Bloch vector $\vec{n}$ and the z axis. This condition is equivalent to the compatibility between the observables given in [16]. Specifically, for observables X_s and Z_t, the previous condition on compatibility gives

$$s^2 + t^2 \leqslant 1. \quad (52)$$

This is what one would have expected due to the physical nature of Θ_t.

V. SUMMARY

We addressed the question of compatibility of unbiased qubit channels and observables. Although our analysis was made explicitly for Pauli channels, i.e., random mixtures of Pauli unitaries, the conclusions hold for a more general case as well—without lost of generality [12], any qubit unital channel Φ can be expressed as a convex combination of (at most) four orthogonal unitary channels induced by unitary operators $U_j = U\sigma_j V^*$, where U, V are suitable unitary operators. It follows that $\mathrm{tr}[U_j^* U_k] = 0$ for $j \neq k$ and $\Phi(\varrho) = U\Psi_{\vec{p}}[V^*\varrho V]U^*$.

We have derived a compatibility formula [Eq. (19)] for the case of unbiased qubit channels and observables. We have shown (Theorem 3) that for a given unital qubit channel the set of compatible unbiased observables forms an ellipsoid (see Fig. 5) naturally embedded inside the Bloch sphere representing the set of binary unbiased observables, while for the inverse

problem the region of unbiased channels compatible with a given unbiased observable is less illustrative (see Fig. 6).

Let us stress that concerning the observable compatibility by Proposition 1, the rotation induced by U is irrelevant. Therefore, the ellipsoid for Φ is connected to the ellipsoid for the Pauli channels $\Psi_{\vec{p}}$ by the unitary rotation V only. For a given qubit unital channel the sharpest (least noisy) compatible observable (quantified by parameter s) is oriented along the main axis of the ellipsoid of compatible observables with a given channel Φ.

Because of the mentioned unitary freedom, the least noisy compatible channel is not unique. However, in accordance with noise-disturbance qualitative tradeoffs it follows that the more noisy the fixed observable is, the less noisy the compatible channel can be.

APPENDIX: THE DUAL SDP PROBLEM

Let us recall the primal problem

$$s_{\mathrm{P}} := \max_{\mathsf{A}'} \mathrm{tr}[\mathsf{A}'(\vec{n} \cdot \vec{\Sigma})] \quad (A1)$$

$$\text{subject to} \quad -\mathsf{A}' \leqslant 0,$$
$$\mathsf{A}' - \mathbb{1} \leqslant 0,$$
$$\mathrm{tr}[\mathsf{A}'(\vec{n_1} \cdot \vec{\Sigma})] = 0,$$
$$\mathrm{tr}[\mathsf{A}'(\vec{n_2} \cdot \vec{\Sigma})] = 0.$$

The Lagrangian is then

$$L(\mathsf{A}', \lambda_1, \lambda_2, \mu, \nu)$$
$$= \mathrm{tr}[\mathsf{A}'(-(\vec{n} + \mu\vec{n_1} + \nu\vec{n_2}) \cdot \vec{\Sigma} - \lambda_1 + \lambda_2)] - \mathrm{tr}[\lambda_2], \quad (A2)$$

where $\lambda_i \geqslant 0$. The Lagrange dual function $g(\lambda_1, \lambda_2, \mu, \nu)$ is then infimum over A' of the Lagrangian. Since it is linear in A' we get that

$$g(\lambda_1, \lambda_2, \mu, \nu) = \begin{cases} -\mathrm{tr}[\lambda_2] & -(\vec{n} + \mu\vec{n_1} + \nu\vec{n_2}) \cdot \vec{\Sigma} \\ & \qquad\qquad -\lambda_1 + \lambda_2 = 0 \\ -\infty & \text{otherwise.} \end{cases} \quad (A3)$$

Thus the function g is nontrivial only when

$$\lambda_2 - (\vec{n} + \mu\vec{n_1} + \nu\vec{n_2}) \cdot \vec{\Sigma} = \lambda_1 \geqslant 0. \quad (A4)$$

Let $\vec{n} + \mu\vec{n_1} + \nu\vec{n_2} = \vec{m}$. Note that $\vec{m} \cdot \vec{n} = 1$. We can include this condition into the constraints of the dual problem, which

is now stated as

$$s_{\mathrm{D}} := \max_{\lambda_1, \lambda_2, \vec{m}} (-\mathrm{tr}[\lambda_2])$$

$$\text{subject to} \quad \lambda_2 \geqslant 0,$$

$$\lambda_2 - \vec{m} \cdot \vec{\Sigma} = \lambda_1,$$

$$\lambda_1 \geqslant 0$$

$$\vec{m} \cdot \vec{n} = 1, \tag{A5}$$

or in its simplified form in (36).

---

[1] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A **49**, 123001 (2016).

[2] D. Kretschmann, D. Schlingemann, and R. F. Werner, The information-disturbance tradeoff and the continuity of Stinespring's representation, IEEE Trans. Inf. Theory **54**, 1708 (2008).

[3] P. Busch, No information without disturbance: Quantum limitations of measurement, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle*, edited by J. Christian and W. Myrvold (Springer-Verlag, Berlin, 2009).

[4] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Probabilistic theories with purification, Phys. Rev. A **81**, 062348 (2010).

[5] T. Heinosaari and T. Miyadera, Qualitative noise-disturbance relation for quantum measurements, Phys. Rev. A **88**, 042117 (2013).

[6] T. Heinosaari, T. Miyadera, and D. Reitzner, Strongly incompatible quantum devices, Found. Phys. **44**, 34 (2014).

[7] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer-Verlag, Berlin, 1996), 2nd rev. ed.

[8] A. S. Holevo, *Statistical Structure of Quantum Theory* (Springer-Verlag, Berlin, 2001).

[9] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory* (Cambridge University Press, Cambridge, UK, 2012).

[10] M. Ozawa, Quantum measuring processes of continuous observables, J. Math. Phys. **25**, 79 (1984).

[11] T. Heinosaari and T. Miyadera, Incompatibility of quantum channels, J. Phys. A **50**, 135302 (2017).

[12] M. B. Ruskai, S. Szarek, and E. Werner, An analysis of completely positive trace-preserving maps on $\mathcal{M}_2$, Linear Algebra Appl. **347**, 159 (2002).

[13] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, UK, 2006).

[14] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2004).

[15] V. Bužek, M. Hillery, and R. F. Werner, Optimal manipulations with qubits: Universal-NOT gate, Phys. Rev. A **60**, R2626(R) (1999).

[16] P. Busch, Unsharp reality and joint measurements for spin observables, Phys. Rev. D **33**, 2253 (1986).

[17] T. Heinosaari, M. A. Jivulescu, D. Reitzner, and M. Ziman, Approximating incompatible von Neumann measurements simultaneously, Phys. Rev. A **82**, 032328 (2010).