Device-independent tests of quantum channels

Michele Dall'Arno,[1, *] Sarah Brandsen,[1, †] and Francesco Buscemi[2, ‡]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore*
[2] *Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya, 464-8601, Japan*
(Dated: March 17, 2017)

We develop a device-independent framework for testing quantum channels. That is, we falsify a hypothesis about a quantum channel based only on an observed set of input-output correlations. Formally, the problem consists of characterizing the set of input-output correlations compatible with any arbitrary given quantum channel. For binary (i.e., two input symbols, two output symbols) correlations, we show that extremal correlations are always achieved by orthogonal encodings and measurements, irrespective of whether or not the channel preserves commutativity. We further provide a full, closed-form characterization of the sets of binary correlations in the case of: i) any dihedrally-covariant qubit channel (such as any Pauli and amplitude-damping channels), and ii) any universally-covariant commutativity-preserving channel in an arbitrary dimension (such as any erasure, depolarizing, universal cloning, and universal transposition channels).

I. INTRODUCTION

Any physical experiment is based upon the observation of correlations among events at various points in space and time, along with some assumptions about the underlying physics. Naturally, in order to be operational any such assumption must have been tested as a hypothesis in a previous experiment. Ultimately, to break an otherwise circular argument, experiments involving no further assumptions are required – that is, device-independent tests.

Formally, a hypothesis consists of a circuit [9], which is usually assumed to have a global causal structure (following special relativity), and its components, which are usually assumed to be governed by classical or quantum theories and thus representable by channels.

Denoting a hypothesis (circuit) by $\mathcal{X}$, the set of correlations compatible with $\mathcal{X}$ is denoted by $S(\mathcal{X})$. Then, hypothesis $\mathcal{X}$ is falsified, along with any other hypothesis $\mathcal{Y}$ such that $S(\mathcal{Y}) \subseteq S(\mathcal{X})$, as soon as the observed correlation does not belong to $S(\mathcal{X})$ (This inclusion relation induces an ordering among channels which is reminiscent of that introduced by Shannon [1] among classical channels). Therefore, from the theoretical viewpoint, the problem of falsifying a hypothesis $\mathcal{X}$ can be recast [2] as that of characterising the set $S(\mathcal{X})$ of compatible correlations.

Since (discrete, memoryless) classical channels are *by definition* input-output correlations (conditional probabilities), the characterisation of $S(\mathcal{X})$ is trivial in classical theory as it is a polytope easily related to the correlation defining the channel. On the contrary, the problem is far from trivial in quantum theory: due to the existence of superpositions of states and effects, the set $S(\mathcal{X})$ can be strictly convex.

In this work we address the problem of device-independent tests of quantum channels, in particular the characterization of the set $S_m^n(\mathcal{X})$ of m-inputs/n-outputs correlations $p_{j|i}$ obtainable through an arbitrary given channel $\mathcal{X}$, upon the input of an arbitrary preparation $\{\rho_i\}_{i=0}^{m-1}$ and the measurement of an arbitrary POVM $\{\pi_j\}_{j=0}^{n-1}$, that is

$$p_{j|i} := \mathrm{Tr}[\mathcal{X}(\rho_i)\pi_j] \quad = \quad i = \boxed{\rho_i}\!-\!\boxed{\mathcal{X}}\!-\!\boxed{\pi_j} = j . \quad (1)$$

The analogous problems of device-independent tests of quantum states and measurements have been recently addressed in Ref. [15] and Ref. [16], respectively.

An alternative formulation for the problem considered here can be given in terms of a "game" involving two parties: an experimenter, claiming to be able to prepare quantum states, feed them through some quantum channel $\mathcal{X}$, and then perform measurements on the output, and a skeptical theoretician, willing to trust observed correlations only. If the experimenter produces *some* correlations lying outsides of $S_m^n(\mathcal{X})$, then the theoretician must conclude that the actual channel $\mathcal{X}'$ is not worse than $\mathcal{X}$ at producing correlations, but this is not sufficient to support the experimenter's claim. Indeed, in order to convince the theoretician, the experimenter must produce *the entire set* $S_m^n(\mathcal{X})$: in fact, it is sufficient to produce a set of correlations whose convex hull *contains* $S_m^n(\mathcal{X})$. Then, the theoretician must conclude that whatever channel the experimenter actually has is at least as good as $\mathcal{X}$ at producing correlations, and the experimenter's claim is accepted.

It is hence clear that the problem of device-independent tests of quantum channels induces a pre-ordering relation among quantum channels: $\mathcal{X} \succeq \mathcal{Y}$ if and only if $S_m^n(\mathcal{X}) \supseteq S_m^n(\mathcal{Y})$. (The order also depends upon m and n, but for compactness we drop the indexes whenever they are clear from the context). In order to characterize such preorder, for any given channel $\mathcal{X}$, we need to i) provide the experimenter with all the states and measurements generating the extremal correlations of $S_m^n(\mathcal{X})$, and ii) provide the theoretician with a full

closed-form characterization of the set $S_m^n(\mathcal{X})$ of compatible correlations.

As a preliminary result, we find that the sets $S_m^n(\mathcal{X})$ coincide for any d-dimensional unitary and dephasing channels, for any d, m, and n (this is an immediate consequence of a remarkable result by Frenkel and Weiner [17].) Upon considering only the binary case $m = n = 2$, our first result is to show that any correlation on the boundary of $S_2^2(\mathcal{X})$ is achieved by a pair of commuting pure states – irrespective of whether $\mathcal{X}$ is a commutativity-preserving channel. Then, we derive the *complete closed-form* characterization of $S_2^2(\mathcal{X})$ for: i) any given dihedrally-covariant qubit channel, including any Pauli and amplitude-damping channels; and ii) any given universally-covariant commutativity-preserving channel, including any erasure, depolarizing, universal $1 \rightarrow 2$ cloning [18], and universal transposition [19] channels.

Upon specifying $\mathcal{X}$ as the d-dimensional identity channel $\mathcal{I}_d$, one recovers device-independent dimension tests analogous to those discussed in Refs. [20–23], in which case the aforementioned ordering induced by the inclusion $S_m^n(\mathcal{I}_{d_0}) \subseteq S_m^n(\mathcal{I}_{d_1}) \Leftrightarrow d_0 \leq d_1$ is of course total. However, the completeness of our characterization of $S_2^2(\mathcal{X})$ implies that our framework detects *all* correlations incompatible with the given hypothesis, unlike Refs. [20–24] where the set of correlations is tested only along an arbitrarily chosen direction.

Let us provide a preview of some consequences of our results:

- Any Pauli channel $\mathcal{P}^{\vec{\lambda}} : \rho \rightarrow \lambda_0 \rho + \sum_{k=1}^{3} \lambda_k \sigma_k \rho \sigma_k^\dagger$ is compatible with p if and only if
$$\frac{|p_{1|1} - p_{1|2}|}{1 - |p_{1|1} - p_{2|2}|} \leq \max_{k \in [1,3]} |2(\lambda_0 + \lambda_k) - 1|;$$

- any amplitude-damping channel $\mathcal{A}^\lambda : \rho \rightarrow A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$ with $A_0 = |0\rangle\langle 0| + \sqrt{\lambda} |1\rangle\langle 1|$ and $A_1 = \sqrt{1-\lambda} |0\rangle\langle 1|$ is compatible with p if and only if
$$\left(\sqrt{p_{1|2} p_{2|1}} - \sqrt{p_{1|1} p_{2|2}} \right)^2 \leq \lambda;$$

- any d-dimensional erasure channel $\mathcal{E}_d : \rho \rightarrow \lambda \rho \oplus (1 - \lambda) \text{Tr}[\rho] \phi$ for some pure state ϕ is compatible with p if and only if
$$|p_{1|1} - p_{1|2}| \leq \lambda;$$

- any d-dimensional depolarizing channel $\mathcal{D}_d^\lambda : \rho \rightarrow \lambda \rho + (1 - \lambda) \text{Tr}[\rho] \mathbb{1}/d$ is compatible with p if and only if
$$\begin{cases} |p_{1|1} - p_{1|2}| \leq \lambda, \\ \dfrac{|p_{1|1} - p_{1|2}|}{1 - |p_{1|1} - p_{2|2}|} \leq \dfrac{d\lambda}{2 - 2\lambda + d\lambda}; \end{cases}$$

- the d-dimensional universal optimal $1 \rightarrow 2$ cloning [18] channel $\mathcal{C}_d$ is compatible with p if and only if
$$|p_{1|1} - p_{1|2}| \leq \frac{d}{d+1};$$

- any d-dimensional universal optimal transposition [19] channel $\mathcal{T}_d$ is compatible with p if and only if
$$\begin{cases} |p_{1|1} - p_{1|2}| \leq \dfrac{1}{d+1}, \\ \dfrac{|p_{1|1} - p_{1|2}|}{1 - |p_{1|1} - p_{2|2}|} \leq \dfrac{1}{3}. \end{cases}$$

This paper is structured as follows. We will introduce our framework and discuss the case of unitary and trace class channels in Section II. For the binary case, introduced in Section III, we will solve the problem for any qubit dihedrally-covariant channel in Section IV, and for any arbitrary-dimensional universally-covariant commutativity-preserving channel in Section V. In Section VI we will provide a natural geometrical interpretation of our results, and in Section VII we will summarize our results and present further outlooks.

II. GENERAL RESULTS

We will make use of standard definitions and results in quantum information theory [25]. Since $S_m^n(\mathcal{X})$ is convex for any n and m, the *hyperplane separation theorem* [26, 27] states that $p \notin S_m^n(\mathcal{X})$ if and only if there exists an $m \times n$ real matrix w such that
$$p^T \cdot w - W(\mathcal{X}, w) > 0, \tag{2}$$
where $p^T \cdot w := \sum_{i,j} p_{j|i} w_{i,j}$, and
$$W(\mathcal{X}, w) := \max_{q \in S_m^n(\mathcal{X})} w^T \cdot q, \tag{3}$$

We call w a *channel witness* and $W(\mathcal{X}, w)$ its threshold value for channel $\mathcal{X}$.

Although Eq. (2) generally allows one to detect *some* conditional probability distributions p not belonging to $S_m^n(\mathcal{X})$ for any arbitrarily fixed witness w, here our aim is to detect *any* such p. Direct application of Eq. (2) is impractical, as one would need to consider *all* of the infinitely many witnesses w. Notice however that Eq. (2) can be rewritten through negation by stating that $p \in S_m^n(\mathcal{X})$ if and only if for any $m \times n$ witness w one has
$$p^T \cdot w - W(\mathcal{X}, w) \leq 0,$$

We then have our first preliminary result.

Lemma 1. *A channel $\mathcal{X} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{K})$ is compatible with conditional probability distribution p if and only if*
$$\max_w \left[p^T \cdot w - W(\mathcal{X}, w) \right] \leq 0. \tag{4}$$

Let us start by considering an arbitrary d-dimensional unitary channel $\mathcal{U}_d : \rho \to U\rho U^\dagger$, for some unitary $U \in \mathcal{L}(\mathcal{H})$ with $\dim \mathcal{H} = d$. If $d \geq m$, the maximization in Eq. (3) is trivial, since the input labels $i \in [1, m]$ can all be encoded on orthogonal states, so that *any* $m \times n$ conditional probability distribution q can in fact be obtained. However, if $d < m$, the evaluation of the witness threshold $W(\mathcal{U}_d, w)$ for any witness w is far from obvious. The solution immediately follows from a recent, remarkable result by Frenkel and Weiner [17]. It turns out that $W(\mathcal{U}_d, w)$ is attained on extremal conditional probability distributions q compatible with the exchange of a classical d-level system, namely, those q where $q_{j|i} = 0$ or 1 for any i and j, and such that $q_{j|i} \neq 0$ for at most d different values of j. Frenkel and Weiner's result hence guarantees that the threshold $W(\mathcal{U}_d, w)$ can be provided in closed form since, for any m and n, the number of such extremal classical conditional probabilities is finite, i.e., the set $S_m^n(\mathcal{U}_d)$ is a *polytope*. Any probability p lying outside $S_m^n(\mathcal{U}_d)$ can thus be detected by testing the violation of Eq. (4) for a *finite number* of witnesses w, corresponding to the faces of the polytope. Moreover, the set $S_m^n(\mathcal{U}_d)$ of distributions compatible with any d-dimensional unitary channel $\mathcal{U}_d$ coincides with the set $S_m^n(\mathcal{F}_d^\lambda)$ of distributions compatible with any d-dimensional dephasing channel $\mathcal{F}_d^\lambda : \rho \to \lambda\rho + (1-\lambda)\sum_k \langle k|\rho|k\rangle |k\rangle\langle k|$.

At the opposite end of the unitary channels, there sit *trace-class channels* $\mathcal{T} : \rho \to \sigma$ for some arbitrary but fixed state σ. In this case, no information about i (the input label) can be communicated. Of course, the set $S_m^n(\mathcal{T})$ of correlations achievable through any trace-class channel $\mathcal{T}$ does not depend on the particular choice of σ: a trace-class channel simply means that no communication is available. For any trace-class channel $\mathcal{T}$ and any witness w, it immediately follows that the threshold $W(\mathcal{T}, w)$ is achieved by conditional probabilities q such that $q_{j|i} = 1$ for a single value of j, and therefore is given by $W(\mathcal{T}, w) = \max_j \sum_i w_{i,j}$. As a consequence, the set $S_m^n(\mathcal{T})$ is a polytope with n vertices, and any probability p lying outside $S_m^n(\mathcal{T})$ can be detected by testing the violation of Eq. (4) for a *finite* number of witnesses w.

III. BINARY CONDITIONAL PROBABILITY DISTRIBUTION

In the remainder of this work we will consider the case where p is a binary input-output conditional probability distributions (i.e. $m = n = 2$).

First, we show that it suffices to consider diagonal or anti-diagonal witnesses with positive entries summing up to one. Indeed, for any witness w, the witness $w' := \alpha(w + \beta)$, where $\alpha > 0$ and β is such that $\beta_{i,j}$ is independent of j, leaves Eq. (4) invariant for any conditional probability distribution p and channel $\mathcal{X}$, since $w' \cdot p = \alpha(p^T \cdot w + \sum_i \beta_{i,1})$.

By taking $\beta_{i,j} = -\min_k w_{i,k}$ for any i and j, the witness w' is diagonal, anti-diagonal, or has a single non-null column. We first consider the latter case. Clearly, the maximum in Eq. (3) is attained when p is a vertex of the polytope $S_2^2(\mathcal{T})$ of probabilities compatible with any trace-type channel $\mathcal{T}$, and therefore Eq. (4) is always verified. Then we consider the case of diagonal and anti-diagonal witnesses. By taking $\alpha^{-1} = \sum_i |w_{i,1} - w_{i,2}|$ one recovers the normalization condition $\sum_{i,j} w_{i,j} = 1$, thus proving the statement.

Therefore, upon denoting with $w^\pm(\omega)$ the diagonal and anti-diagonal witnesses given by

$$w^+(\omega) := \begin{pmatrix} \frac{1+\omega}{2} & 0 \\ 0 & \frac{1-\omega}{2} \end{pmatrix}, \quad w^-(\omega) := \begin{pmatrix} 0 & \frac{1+\omega}{2} \\ \frac{1-\omega}{2} & 0 \end{pmatrix},$$

where $\omega \in [-1, 1]$, one has the following preliminary result.

Lemma 2. *The maximum in Eq. 4 is attained for a diagonal or anti-diagonal witness, namely*

$$\max_w (p^T \cdot w - W(\mathcal{X}, w))$$
$$= \max_{\omega \in [-1,1]} (p^T \cdot w^\pm(\omega) - W(\mathcal{X}, w^\pm(\omega))).$$

Any extremal distribution q in Eq. (3) can be represented by states ρ_0 and ρ_1 and a POVM $\{\pi_0, \pi_1\}$ such that $q_{j|i} = \text{Tr}[\mathcal{X}(\rho_i)\pi_j]$. Since $w^\pm(\omega)$ is diagonal or anti-diagonal, Eq. (3) represents the maximum probability of success in the discrimination of states $\{\rho_0, \rho_1\}$ with prior probabilities given by the non-null entries of w, in the presence of noise $\mathcal{X}$, namely

$$W(\mathcal{X}, w^\pm(\omega))$$
$$= \frac{1}{2} \max_{\substack{\rho_0, \rho_1 \\ \{\pi_0, \pi_1\}}} [(1 + \omega)\text{Tr}[\mathcal{X}(\rho_0)\pi_0] + (1 - \omega)\text{Tr}[\mathcal{X}(\rho_1)\pi_1]].$$

It is a well-known fact [28] that the solution of the optimization problem over POVMs is given as a function of the Helstrom matrix defined as

$$H_\omega(\rho_0, \rho_1) := \frac{1+\omega}{2}\rho_0 - \frac{1-\omega}{2}\rho_1,$$

as follows

$$W(\mathcal{X}, w^\pm(\omega)) = \frac{1}{2} \max_{\rho_0, \rho_1} [1 + \|\mathcal{X}(H_\omega(\rho_0, \rho_1))\|_1], \quad (5)$$

where $\|\cdot\|_1$ denotes the operator 1-norm.

It is easy to see that without loss of generality one can take ρ_0 and ρ_1 such that $[\rho_0, \rho_1] = 0$. Indeed, let $\{|k\rangle\}$ be a basis of eigenvectors of the Helstrom matrix $H_\omega(\rho_0, \rho_1)$. The complete dephasing channel $\mathcal{F}_d^0$ on the basis $\{|k\rangle\}$ is such that

$$H_\omega(\rho_0, \rho_1) = \mathcal{F}_d^0(H_\omega(\rho_0, \rho_1)) = H_\omega(\sigma_0, \sigma_1),$$

where $\sigma_i := \mathcal{F}_d^0(\rho_i)$ and therefore $[\sigma_0, \sigma_1] = 0$. By applying channel $\mathcal{X}$ we have the following identity

$$\mathcal{X}(H_\omega(\rho_0, \rho_1)) = \mathcal{X}(H_\omega(\sigma_0, \sigma_1))$$

Therefore, the encoding $\{\sigma_i\}$ performs as well as the encoding $\{\rho_i\}$, and thus without loss of generality we can take the supremum in Eq. (5) over commuting encodings only.

Moreover, one can see that without loss of generality one can take σ_i to be orthogonal pure states. Indeed, let $\sigma_i = \sum_k \mu_{k|i} |k\rangle\langle k|$ be a spectral decomposition of σ_i. Due to the convexity of the trace norm we have

$$\|\mathcal{X}(H_\omega(\sigma_0, \sigma_1))\|_1$$
$$= \left\| \sum_{k,l} \mu_{k|0}\mu_{l|1} \mathcal{X}(H_\omega(|k\rangle\langle k|, |l\rangle\langle l|)) \right\|_1$$
$$\leq \sum_{k,l} \mu_{k|0}\mu_{l|1} \|\mathcal{X}(H_\omega(|k\rangle\langle k|, |l\rangle\langle l|))\|_1$$
$$\leq \max_{k,l} \|\mathcal{X}(H_\omega(|k\rangle\langle k|, |l\rangle\langle l|))\|_1 .$$

Then we have the following preliminary result.

Lemma 3. *The maximum in Eq. (3) is given by an orthonormal pure encoding, namely*

$$W\left(\mathcal{X}, w^\pm(\omega)\right) := \max_{\substack{|\phi_0\rangle, |\phi_1\rangle \\ \langle\phi_1|\phi_0\rangle=0}} \frac{1}{2}\left[1 + \|\mathcal{X}(H_\omega(\phi_0, \phi_1))\|_1\right],$$

and by an orthogonal POVM such that π_0 is the projector on the positive part of $H_\omega(\phi_0, \phi_1)$ and $\pi_1 = \mathbb{1} - \pi_0$.

Here, for any pure state $|\phi\rangle$ we denote with $\phi := |\phi\rangle\langle\phi|$ the corresponding projector.

IV. DIHEDRALLY COVARIANT QUBIT CHANNEL

Let us start with the case where $\mathcal{X} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is a qubit channel, i.e. $\dim\mathcal{H} = \dim\mathcal{K} = 2$. Since Pauli matrices span the space of qubit Hermitian operators, any qubit state ρ can be parametrized in terms of Pauli matrices, i.e.

$$\rho = \frac{1}{2}(\mathbb{1} + \vec{\sigma}^T \cdot \vec{x}), \qquad |\vec{x}|_2 \leq 1, \tag{6}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$ and $\vec{x}$ are the vectors of Pauli matrices and their real coefficients, respectively. Analogously, any qubit channel $\mathcal{X}$ can be parametrized in terms of Pauli matrices, i.e.

$$\mathcal{X}(\rho) = \frac{1}{2}\left(\mathbb{1} + \vec{\sigma}^T \cdot (A\vec{x} + \vec{b})\right),$$

where $A_{i,j} = \frac{1}{2}\text{Tr}[\sigma_i \mathcal{X}(\sigma_j)]$ and $b_i = \frac{1}{2}\text{Tr}[\sigma_i \mathcal{X}(\mathbb{1})]$.

With such a parametrization $\mathcal{X}(H_\omega(\phi_0, \phi_1))$ assumes a very simple form given by

$$\mathcal{X}(H_\omega(\phi_0, \phi_1)) = \frac{1}{2}\left[\omega\mathbb{1} + \left(A\vec{x} + \omega\vec{b}\right)^T \cdot \vec{\sigma}\right],$$

whose eigenvalues are $\frac{1}{2}\left(\omega \pm \left|A\vec{x} + \omega\vec{b}\right|_2\right)$. Thus, the witness threshold $W(\mathcal{X}, w^\pm(\omega))$ in Eq. (3) can be readily computed by means of Lemma 3 as

$$W\left(\mathcal{X}, w^\pm(\omega)\right) = \frac{1}{2}\left[1 + \max\left(|\omega|, \max_{\substack{\vec{x} \\ |\vec{x}|_2 \leq 1}} \left|A\vec{x} + \omega\vec{b}\right|_2\right)\right].$$

Notice that this expression is the maximum between two strategies. The first one is given by the trivial POVM and thus corresponds to trivial guessing. The second one can be further simplified by means of the following substitutions. Let $A = VDU$ be a polar decomposition of matrix A with U and V unitaries and D diagonal and positive-semidefinite with eigenvalues $\vec{d}$ (accordingly $\vec{c} := -V^\dagger \vec{b}$). By unitary invariance of the 2-norm one has

$$\max_{\substack{\vec{x} \\ |\vec{x}|_2 \leq 1}} \left|A\vec{x} + \omega\vec{b}\right|_2 = \max_{\substack{\vec{x} \\ |\vec{x}|_2 \leq 1}} |D\vec{x} - \omega\vec{c}|_2 .$$

By defining $\vec{y} := D\vec{x}$ one has

$$\max_{\substack{\vec{x} \\ |\vec{x}|_2 \leq 1}} |D\vec{x} - \omega\vec{c}|_2 = \max_{\substack{\vec{y}, \vec{z} \\ |D^{-1}\vec{y} + (\mathbb{1} - D^{-1}D)\vec{z}|_2 \leq 1}} |\vec{y} - \omega\vec{c}|_2,$$

where $(\cdot)^{-1}$ denotes the Moore-Penrose pseudoinverse. By explicit computation it follows that $[D^{-1}]^T(\mathbb{1} - D^{-1}D) = 0$, and therefore vectors $D^{-1}\vec{y}$ and $(\mathbb{1} - D^{-1}D)\vec{z}$ are orthogonal. Then for any optimal $(\vec{y}, \vec{z})$ one has that $(\vec{y}, 0)$ is also optimal, since $|D^{-1}\vec{y} + (\mathbb{1} - D^{-1}D)\vec{z}|_2 \geq |D^{-1}\vec{y}|_2$. Therefore we have

$$W\left(\mathcal{X}, w^\pm(\omega)\right) = \frac{1}{2}[1 + \max(\omega, \Delta(\omega))], \tag{7}$$

where

$$\Delta(\omega) := \max_{\substack{\vec{y} \\ |D^{-1}\vec{y}|_2 \leq 1}} |\vec{y} - \omega\vec{c}|_2 . \tag{8}$$

The maximum in Eq. (8) is a quadratically constrained quadratic optimization problem, which is known to be NP-hard in general. However, $\Delta(\omega)$ has a simple geometrical interpretation: it is the maximum Euclidean distance of vector $\omega\vec{c}$ and ellipsoid $|D^{-1}\vec{y}|_2 \leq 1$. This interpretation suggests symmetries under which the optimization problem becomes feasible. In particular, we take vector $\vec{c}$ to be parallel to one of the axis of the ellipsoid $|D^{-1}\vec{y}|_2 \leq 1$, namely $c_1 = c_2 = 0$ (up to irrelevant permutations of the computational basis).

This configuration corresponds to a D_2-covariant channel $\mathcal{X}$, where D_2 is the dihedral group of the symmetries of a line segment, consisting of two reflections and a π-rotation. This configuration is depicted in Fig. 1. In particular, a qubit channel $\mathcal{X}$ is D_2-covariant if and only if there exist unitary representations $U_k \in \mathbb{R}^{3\times 3}$ and $V_k \in \mathbb{R}^{3\times 3}$ of D_2 such that

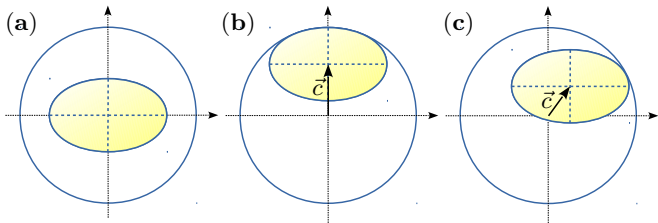$$AU_k\vec{x} + \vec{b} = V_k(A\vec{x} + \vec{b}). \tag{9}$$

Figure 1. Bloch-sphere representation of: [(a), (b)] dihedrally covariant channels $\mathcal{X}$ mapping the sphere into an ellipsoid (a) centered in the Bloch sphere (e.g. any Pauli channel $\mathcal{P}^{\vec{\lambda}}$), or (b) translated by a vector $\vec{c}$ which is parallel to one of the axis of the ellipsoid (e.g. any amplitude damping channel $\mathcal{A}^{\lambda}$); (c) non-dihedrally covariant channel $\mathcal{X}$, as the ellipsoid is translated by a vector $\vec{c}$ which is not parallel to any of the axis of the ellipsoid.

Up to unitaries, the most general unitary representation of D_2 in $\mathbb{R}^{3\times 3}$ is given by

$$W_1 = \sigma_z \oplus 1, \quad W_2 = -\sigma_z \oplus 1, \quad W_3 = -\mathbb{1} \oplus 1,$$

where W_1 and W_2 are reflections and W_3 is a π-rotation. We take $U_k := U^{\dagger} W_k U$ and $V_k := V W_k V^{\dagger}$. Then by explicit computation we have

$$A U_k \vec{x} + \vec{b} = V_k A \vec{x} + \vec{b},$$

where we used the fact that $[D, W_k] = 0$ for any k. Therefore, D_2 covariance expressed by Eq. (9) is equivalent to the requirement $W_k \vec{c} = \vec{c}$, namely $c_1 = c_2 = 0$.

Under the assumption of D_2-covariance, we take without loss of generality $d_2 \geq d_1$ and $c_3 \geq 0$. If also $c_3 = 0$, we further take without loss of generality $d_3 \geq d_2$. Then, as formally proved in the Appendix, the maximum Euclidean distance $\Delta(\omega)$ in Eq. (8) can be explicitly computed, leading to the following result.

Lemma 4. *The witness threshold $W(\mathcal{X}, w^{\pm}(\omega))$ of any qubit D_2-covariant channel $\mathcal{X}$ is given by Eq. (7) where*

$$\Delta(\omega) = \begin{cases} d_2\sqrt{1 + \dfrac{c_3^2\omega^2}{d_2^2 - d_3^2}}, & \text{if } |\omega| < \dfrac{d_2^2 - d_3^2}{d_3 c_3}, \\ d_3 + c_3|\omega|, & \text{otherwise.} \end{cases}$$

The optimal encoding is given by Eq. (6) with $\vec{x} = D^{-1}\vec{y}$ and

$$\vec{y} = \begin{cases} \left(0, \pm d_2\sqrt{1 - \dfrac{c_3^2 d_3^2 \omega^2}{(d_3^2 - d_2^2)^2}}, \dfrac{c_3 d_3^2 \omega}{d_3^2 - d_2^2}\right)^T & \text{if } |\omega| \leq \dfrac{d_2^2 - d_3^2}{d_3 c_3} \\ (0, 0, \pm d_3)^T & \text{otherwise.} \end{cases}$$

Using Lemma 4 and Lemma 1, Eq. (4) becomes the maximum over ω of the minimum of two functions. The maximum is attained either in the maxima 0, $\pm\omega_1$, or ± 1 of the two functions over the domain $[-1, 1]$, where

$$\omega_1 := \frac{(d_2^2 - d_3^2)(p_{1|1} - p_{2|2})}{c_3\sqrt{c_3^2 d_2^2 - (d_2^2 - d_3^2)(p_{1|1} - p_{2|2})^2}},$$

(the limit should be considered if $c_3 = 0$), or in their intersection $\pm\omega_2$ given by

$$\omega_2 := \begin{cases} \sqrt{\dfrac{d_2^2(d_2^2 - d_3^2)}{d_2^2 - d_3^2 - d_2^2 c_3^2}}, & \text{if } (d_2^2 - d_3^2) > d_2^2 c_3, \\ \dfrac{d_3}{1 - c_3}, & \text{otherwise.} \end{cases}$$

We can then state our first main result, formally proved in the Appendix, namely a complete and closed-form characterization of the set $S_2^2(\mathcal{X})$ of conditional probability distributions compatible with any qubit D_2-covariant channel $\mathcal{X}$.

Theorem 1. *Any given binary conditional probability distribution p is compatible with any given qubit D_2-covariant channel $\mathcal{X}$ if and only if*

$$\max_{\omega \in \Omega}(p^T \cdot w^{\pm}(\omega) - W(\mathcal{X}, w^{\pm}(\omega))) \leq 0, \qquad (10)$$

where $\Omega := \{0, \pm\omega_1, \pm\omega_2, \pm 1\} \cap [-1, 1]$.

As applications of Theorem 1, let us explicitly characterize the sets of binary conditional probability distributions compatible with two relevant examples of qubit D_2-covariant channels: the Pauli and amplitude-damping channels.

Any Pauli channel can be written as $\mathcal{P}^{\vec{\lambda}} : \rho \to \lambda_0 \rho + \sum_{k=1}^3 \lambda_k \sigma_k \rho \sigma_k^{\dagger}$, where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. One has that $c_3 = 0$ and $d_3 = \max_{k \in [1,3]} |2(\lambda_0 + \lambda_k) - 1| \geq d_2$, thus $\omega_1 = \infty$ and $\omega_2 = d_3$ and the maximum in Eq. (10) is attained for $\omega = \pm\omega_2$. Thus, upon applying Theorem 1, one has the following result.

Corollary 1. *Any given binary conditional probability distribution p is compatible with the Pauli channel $\mathcal{P}^{\vec{\lambda}}$ if and only if*

$$\frac{|p_{1|1} - p_{1|2}|}{1 - |p_{1|1} - p_{2|2}|} \leq \max_{k \in [1,3]} |2(\lambda_0 + \lambda_k) - 1|.$$

Any amplitude-damping channel can be written as $\mathcal{A}^{\lambda}(\rho) = \sum_{k=0}^1 A_k \rho A_k^{\dagger}$, where $A_0 = |0\rangle\langle 0| + \sqrt{\lambda}|1\rangle\langle 1|$ and $A_1 = \sqrt{1-\lambda}|0\rangle\langle 1|$. As shown in the Appendix, one has that $c_3 = 1 - \lambda$ and $d_3 = \lambda$, $d_2 = d_1 = \sqrt{\lambda}$, and thus the maximum in Eq. (10) is attained for $\omega = \pm\omega_1$ or $\omega = \pm 1$. Thus, upon applying Theorem 1, one has the following result, formally proved in the Appendix.

Corollary 2. *Any given binary conditional probability distribution p is compatible with the amplitude-damping channel $\mathcal{A}^{\lambda}$ if and only if*

$$\left(\sqrt{p_{1|2}p_{2|1}} - \sqrt{p_{1|1}p_{2|2}}\right)^2 \leq \lambda.$$

V. UNIVERSALLY-COVARIANT COMMUTATIVITY-PRESERVING CHANNELS

Let us now move to the arbitrary dimensional case. We trade generality regarding the dimension for generality regarding the symmetry of the channel, and assume

universal covariance. A channel $\mathcal{X} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ is universally covariant if and only if there exist unitary representations $U_g \in \mathcal{L}(\mathcal{H})$ and $V_g \in \mathcal{L}(\mathcal{K})$ of the special unitary group $SU(d)$ with $d := \dim \mathcal{H}$, such that for every state $\rho \in \mathcal{L}(\mathcal{H})$ one has

$$\mathcal{X}(U_g \rho U_g^\dagger) = V_g \mathcal{X}(\rho) V_g^\dagger. \tag{11}$$

From universal covariance it immediately follows that *any* orthonormal pure encoding attains the witness threshold $W(\mathcal{X}, w^\pm(\omega))$ in Eq. (5). Indeed, for any orthonormal pure states $\{\phi_i\}$ let U be the unitary such that $\phi_i = U |i\rangle\langle i| U^\dagger$. Then one has

$$\begin{aligned}
&\|\mathcal{X}(H_\omega(\phi_0, \phi_1))\|_1 \\
&= \|\mathcal{X}(H_\omega(U|0\rangle\langle 0| U^\dagger, U|1\rangle\langle 1| U^\dagger))\|_1 \\
&= \|V\mathcal{X}(H_\omega(|0\rangle\langle 0|, |1\rangle\langle 1|))V^\dagger\|_1 \\
&= \|\mathcal{X}(H_\omega(|0\rangle\langle 0|, |1\rangle\langle 1|))\|_1,
\end{aligned}$$

where the second equality follows from Eq. (11), and the third from the invariance of trace distance under unitary transformations. Then we have the following result.

Lemma 5. *The witness threshold $W(\mathcal{X}, w^\pm(\omega))$ of any universally covariant channel $\mathcal{X}$ is given by*

$$W(\mathcal{X}, w^\pm(\omega)) = \frac{1}{2}[1 + \|\mathcal{X}(H_\omega(|0\rangle\langle 0|, |1\rangle\langle 1|))\|_1]. \tag{12}$$

The optimal encoding is given by any pair of orthonormal pure states.

Equation (12) has a simple dependence on w in the case when channel $\mathcal{X}$ is commutativity preserving, i.e. $[\mathcal{X}(\rho_0), \mathcal{X}(\rho_1)] = 0$ whenever $[\rho_0, \rho_1] = 0$. Notice that it suffices to check commutativity preservation for pure states, indeed a channel $\mathcal{X}$ is commutativity preserving if and only if $[\mathcal{X}(\phi_0), \mathcal{X}(\phi_1)] = 0$ whenever $\langle \phi_1 | \phi_0 \rangle = 0$. Necessity is trivial, and sufficiency follows by assuming $[\rho_0, \rho_1] = 0$, and considering a simultaneous spectral decompositions of $\rho_0 = \sum_k \mu_k \phi_k$ and $\rho_1 := \sum_j \nu_j \phi_j$. Then one has

$$\begin{aligned}
[\mathcal{X}(\rho_0), \mathcal{X}(\rho_1)] &= \sum_{k,l} \mu_k \nu_l [\mathcal{X}(\phi_k), \mathcal{X}(\phi_l)] \\
&= 0,
\end{aligned}$$

where the last inequality follows from the fact that $\langle \phi_l | \phi_k \rangle = \delta_{k,l}$. For a universally covariant channel $\mathcal{X}$, it immediately follows from Eq. (11) that it suffices to check commutativity preservation for an arbitrary pair of orthogonal pure states.

In this case $\mathcal{X}(|0\rangle\langle 0|)$ and $\mathcal{X}(|1\rangle\langle 1|)$ admit a common basis of eigenvectors $\{|k\rangle\}$, and thus a spectral decomposition of the Helstrom matrix $\mathcal{X}(H_\omega(|0\rangle\langle 0|, |1\rangle\langle 1|))$ is given by

$$\mathcal{X}(H_\omega(|0\rangle\langle 0|, |1\rangle\langle 1|)) = \sum_k (\alpha_k \omega + \beta_k)|k\rangle\langle k|,$$

where α_k and β_k are the half-sum and half-difference of the k-th eigenvectors of $\mathcal{X}(|0\rangle\langle 0|)$ and $\mathcal{X}(|1\rangle\langle 1|)$, respectively. Therefore Eq. (12) becomes

$$W(\mathcal{X}, w^\pm(\omega)) = \frac{1}{2}\left(1 + \sum_k |\alpha_k \omega + \beta_k|\right).$$

Then, the optimization problem in Eq. (4) becomes piece-wise linear, thus the maximum is attained on the intersections of the piece-wise components given by $\gamma_k := \beta_k / \alpha_k$ when such values belongs to the domain $[-1, 1]$, or on its extrema. We can then provide our second main result, namely a complete closed-form characterization of the set $S_2^2(\mathcal{X})$ of conditional probability distributions compatible with any arbitrary-dimensional universally-covariant commutativity-preserving channel $\mathcal{X}$.

Theorem 2. *Any given binary conditional probability distribution p is compatible with any given arbitrary-dimensional universally-covariant commutativity-preserving channel $\mathcal{X}$ if and only if*

$$\begin{cases} |p_{1|1} - p_{1|2}| \leq \sum_k |\beta_k|, \\ |p_{1|1} - p_{1|2}| \leq \|\mathcal{X}(H_{\gamma_k}(|0\rangle\langle 0|, |1\rangle\langle 1|))\|_1 - \gamma_k |p_{1|1} - p_{2|2}|, \end{cases}$$

for any k such that $\gamma_k \in [-1, 1]$.

As applications of Theorem 2, let us explicitly compute the binary conditional probability distributions compatible with any erasure, depolarizing, universal optimal $1 \to 2$ cloning, and universal optimal transposition channels. As discussed before, commutativity preservation can be immediately verified for all of these channels by checking that $[\mathcal{X}(|0\rangle\langle 0|), \mathcal{X}(|1\rangle\langle 1|)] = 0$.

Any erasure channel can be written as $\mathcal{E}_d^\lambda : \rho \to \lambda\rho \oplus (1-\lambda)\phi$, where ϕ is some pure state. One can compute that $\vec{\alpha} = \left(\frac{\lambda}{2}, \frac{\lambda}{2}, 0 \times d - 2, 1 - \lambda\right)$ and $\vec{\beta} = \left(\frac{\lambda}{2}, -\frac{\lambda}{2}, 0 \times d - 1\right)$, thus upon applying Theorem 2 one has the following Corollary.

Corollary 3. *Any given binary conditional probability distribution p is compatible with the erasure channel $\mathcal{E}_d^\lambda$ if and only if*

$$|p_{1|1} - p_{1|2}| \leq \lambda.$$

Any depolarizing channel can be written as $\mathcal{D}_d^\lambda : \rho \to \lambda\rho + (1-\lambda)\frac{\mathbb{1}}{d}$. One can compute that $\vec{\alpha} = \left(\frac{\lambda}{2} + \frac{1-\lambda}{d} \times 2, \frac{1-\lambda}{d} \times d - 2\right)$ and $\vec{\beta} = \left(-\frac{\lambda}{2}, \frac{\lambda}{2}, 0 \times d - 2\right)$, thus upon applying Theorem 2 one has the following Corollary.

Corollary 4. *Any given binary conditional probability distribution p is compatible with the depolarizing channel $\mathcal{D}_d^\lambda$ if and only if*

$$\begin{cases} |p_{1|1} - p_{1|2}| \leq \lambda, \\ \dfrac{|p_{1|1} - p_{1|2}|}{1 - |p_{1|1} - p_{2|2}|} \leq \dfrac{d\lambda}{2 - 2\lambda + d\lambda}. \end{cases}$$

The universal optimal $1 \to 2$ cloning channel can be written as $\mathcal{C}_d^\lambda : \rho \to \frac{2}{d+1} P_S(\rho \otimes \mathbb{1}) P_S$. By explicit computation one has

$$\mathcal{C}_d(|i\rangle\langle i|) = \frac{1}{2(d+1)} \sum_k (|k,i\rangle + |i,k\rangle)(\langle k,i| + \langle i,k|),$$

and therefore $[\mathcal{C}_d(|0\rangle\langle 0|), \mathcal{C}_d(|1\rangle\langle 1|)] = 0$, thus the universal optimal $1 \to 2$ cloning $\mathcal{C}_d$ is a commutativity preserving channel. One can compute that $\vec{\alpha} = \left(\frac{1}{d+1} \times 3, \frac{1}{2(d+1)} \times 2(d-2) \right)$ and $\vec{\beta} = \left(-\frac{1}{d+1}, \frac{1}{d+1}, 0, -\frac{1}{2(d+1)} \times d - 2, \frac{1}{2(d+1)} \times d - 2 \right)$, thus upon applying Theorem 2 one has the following Corollary.

Corollary 5. *Any given binary conditional probability distribution p is compatible with the universal optimal $1 \to 2$ cloning channel $\mathcal{C}_d$ if and only if*

$$|p_{1|1} - p_{1|2}| \leq \frac{d}{d+1}.$$

The universal transposition channel can be written as $\mathcal{T}_d : \rho \to \frac{1}{d+1}\left(\rho^T + \mathbb{1}\right)$. One can compute that $\vec{\alpha} = \left(\frac{3}{2(d+1)} \times 2, \frac{1}{d+1} \times d - 2 \right)$ and $\vec{\beta} = \left(\frac{1}{2(d+1)}, -\frac{1}{2(d+1)}, 0 \times d - 2 \right)$, thus upon applying Theorem 2 one has the following Corollary.

Corollary 6. *Any given binary conditional probability distribution p is compatible with the universal transposition channel $\mathcal{T}_d$ if and only if*

$$\begin{cases} |p_{1|1} - p_{1|2}| \leq \frac{1}{d+1}, \\ \dfrac{|p_{1|1} - p_{1|2}|}{1 - |p_{1|1} - p_{2|2}|} \leq \dfrac{1}{3}. \end{cases}$$

The results of Corollaries 1, 2, 3, 4, 5, and 6 are summarized in Table I.

VI. CARTESIAN REPRESENTATION

In this Section we provide a geometrical interpretation of our results. Binary conditional probability distributions are represented by 2×2 real matrices, so they can be regarded as vectors in $\mathbb{R}^4$. However, due to the normalization constraint $\sum_j p_{j|i} = 1$ for any i, they all lie in a bidimensional affine subspace. A natural Cartesian parametrization of such a subspace is given by

$$p_{j|i} = p(x,y) = \frac{1}{2}\left[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + x \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} + y \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right],$$
(13)

and binary conditional probability distributions form the square $|x \pm y| \leq 1$, whose 4 vertices are the right-stochastic matrices with all entries equal to 0 or 1.

As it is clear from Eq. (13):

$\mathcal{X}$	$p \in S_2^2(\mathcal{X})$
$\mathcal{P}^{\vec{\lambda}}$	$|p_{1|1} - p_{1|2}| \leq \max_{k \in [1,3]} |2(\lambda_0 + \lambda_k) - 1|$
$\mathcal{A}^\lambda$	$\left(\sqrt{p_{1|2} p_{2|1}} - \sqrt{p_{1|1} p_{2|2}} \right)^2 \leq \lambda$
$\mathcal{E}_d^\lambda$	$|p_{1|1} - p_{1|2}| \leq \lambda$
$\mathcal{D}_d^\lambda$	$\begin{cases} |p_{1|1} - p_{1|2}| \leq \lambda \\ \dfrac{|p_{1|1}-p_{1|2}|}{1-|p_{1|1}-p_{2|2}|} \leq \dfrac{d\lambda}{2-2\lambda+d\lambda} \end{cases}$
$\mathcal{C}_d$	$|p_{1|1} - p_{1|2}| \leq \frac{d}{d+1}$
$\mathcal{T}_d$	$\begin{cases} |p_{1|1} - p_{1|2}| \leq \frac{1}{d+1} \\ \dfrac{|p_{1|1}-p_{1|2}|}{1-|p_{1|1}-p_{2|2}|} \leq \dfrac{1}{3} \end{cases}$

Table I. Complete closed-form characterization of the set $S_2^2(\mathcal{X})$ of binary conditional probability distributions compatible with channel $\mathcal{X}$, for $\mathcal{X}$ given by the Pauli channel $\mathcal{P}^{\vec{\lambda}}$, the amplitude damping channel $\mathcal{A}^\lambda$, the erasure channel $\mathcal{E}_d^\lambda$, the depolarizing channel $\mathcal{D}_d^\lambda$, the universal $1 \to 2$ cloning channel $\mathcal{C}_d$, and the universal transposer $\mathcal{T}_d$, as given by Corollaries 1, 2, 3, 4, 5, and 6, respectively.

- a permutation of the states $\{\rho_0, \rho_1\}$ corresponds to the transformation $(x,y) \to (x, -y)$;

- a permutation of the effects $\{\pi_0, \pi_1\}$ corresponds to the transformation $(x,y) \to (-x, -y)$;

- a permutation of the states $\{\rho_0, \rho_1\}$ and effects $\{\pi_0, \pi_1\}$ corresponds to the transformation $(x,y) \to (-x, y)$.

Therefore, for any channel $\mathcal{X}$, the set $S_2^2(\mathcal{X})$ of binary conditional probability distributions compatible with $\mathcal{X}$ is symmetric for reflections around the x or y axes (i.e., it is D_2-covariant).

As a consequence of our previous results, the sets $S_2^2(\mathcal{U}_d)$ and $S_2^2(\mathcal{F}_d^\lambda)$ of conditional probability distributions compatible with any unitary and dephasing channels $\mathcal{U}_d$ and $\mathcal{F}_d^\lambda$ coincide with the square $|x \pm y| \leq 1$, for any d and any λ. The set $S_2^2(\mathcal{T})$ of conditional probability distributions compatible with any trace-class channel $\mathcal{T}$ coincide with the segment $x \in [-1, 1]$, $y = 0$.

With the parametrization in Eq. (13), the sets of binary conditional probability distributions compatible with any Pauli, amplitude-damping, erasure, depolarizing, universal $1 \to 2$ cloning, and universal transposition channels as given by Corollaries 1, 2, 3, 4, 5, and 6 respectively, are given in Table II and depicted in Fig. 2.

VII. CONCLUSIONS AND OUTLOOK

In this work, we developed a device-independent framework for testing quantum channels. The problem was framed as a game involving an experimenter, claiming to be able to produce some quantum channel, and a theoretician, willing to trust observed correlations only.

$\mathcal{X}$	$p(x,y) \in S_2^2(\mathcal{X})$
$\mathcal{P}^{\vec{\lambda}}$	$\|y\| \leq \max\limits_{k \in [1,3]} \|2(\lambda_0 + \lambda_k) - 1\|$
$\mathcal{A}^\lambda$	$\frac{1}{4}\left(\sqrt{1 - 2y - x^2 + y^2} - \sqrt{1 + 2y - x^2 + y^2}\right)^2 \leq \lambda$
$\mathcal{E}_d^\lambda$	$\|y\| \leq \lambda$
$\mathcal{D}_d^\lambda$	$\begin{cases} \|y\| \leq \lambda \\ \frac{\|y\|}{1-\|x\|} \leq \frac{d\lambda}{2-2\lambda+d\lambda} \end{cases}$
$\mathcal{C}_d$	$\|y\| \leq \frac{d}{d+1}$
$\mathcal{T}_d$	$\begin{cases} \|y\| \leq \frac{1}{d+1} \\ \frac{\|y\|}{1-\|x\|} \leq \frac{1}{3} \end{cases}$

Table II. Cartesian parametrization of the set $S_2^2(\mathcal{X})$ of binary conditional probability distributions compatible with channel $\mathcal{X}$, for $\mathcal{X}$ given by the Pauli channel $\mathcal{P}^{\vec{\lambda}}$, the amplitude damping channel $\mathcal{A}^\lambda$, the erasure channel $\mathcal{E}_d^\lambda$, the depolarizing channel $\mathcal{D}_d^\lambda$, the universal $1 \to 2$ cloning channel $\mathcal{C}_d$, and the universal transposer $\mathcal{T}_d$.

The optimal strategy consists of i) all the input states and measurements generating the extremal correlations that the experimenter needs to produce, and ii) a full closed-form characterization of the correlations compatible with the claim, that the theoretician needs to compare with the observed correlations. For binary correlations, we explicitly derived the optimal strategy for the cases where the claimed channel is a dihedrally-covariant qubit channel, such as any Pauli and amplitude-damping channels, or an arbitrary-dimensional universally-covariant commutativity-preserving channel, such as any erasure, depolarizing, universal cloning, and universal transposition channels.

Natural generalisation of our results include relaxing the restriction of binary correlations, that is $m = n = 2$, and extending the characterization of $S_m^n(\mathcal{X})$ to other classes of channels. An interesting generalisation would consist of letting the POVM $\{\pi_y\}$ depend upon an input not known during the preparation of $\{\rho_x\}$, as is the case in quantum random access codes. Moreover, the setup in Eq. (1) could be modified to allow for entanglement alongside $\mathcal{X}$, or many parallel or sequential uses of channel $\mathcal{X}$.

We conclude by remarking that our results are particularly suitable for experimental implementation. For any channel $\mathcal{X}$ an experimenter claims to be able to produce, our framework only requires them to prepare orthogonal pure input states and perform orthogonal measurements in order to fully characterize $S_2^2(\mathcal{X})$ and thus device-independently test $\mathcal{X}$.

DATA ACCESSIBILITY

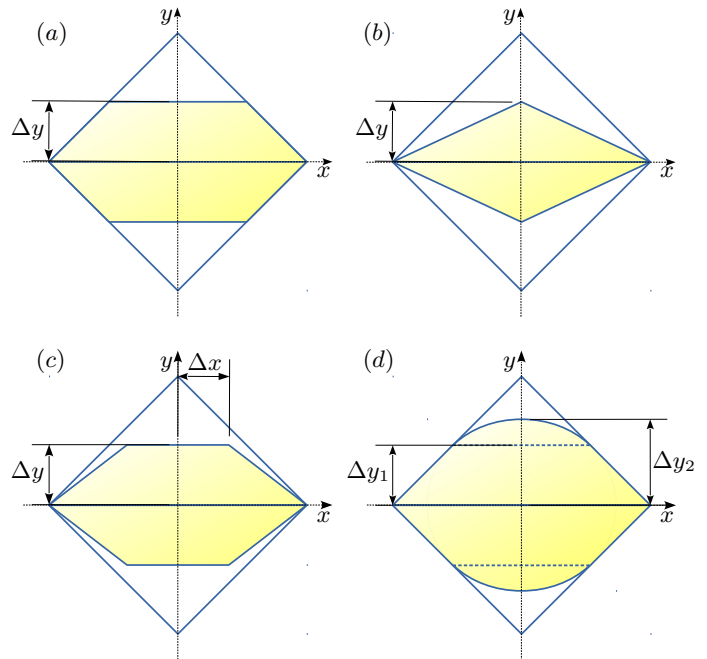This work does not have any experimental data.

Figure 2. Cartesian representation of the space of binary conditional probability distributions p. The outer white square denotes the polytope of all binary conditional probability distributions. The inner yellow region denotes the sets $S_2^2(\mathcal{X})$ of conditional probability distributions compatible with: **(a)** the erasure channel $\mathcal{X} = \mathcal{E}_d^\lambda$ (for $\Delta y = \lambda$) and the universal optimal $1 \to 2$ cloning channel $\mathcal{X} = \mathcal{C}_d$ (for $\Delta y = \frac{d}{d+1}$); **(b)** the Pauli channel $\mathcal{X} = \mathcal{P}^{\vec{\lambda}}$ (for $\Delta y = \max_{k \in [1,3]} \|2(\lambda_0 + \lambda_k) - 1\|$); **(c)** the depolarizing channel $\mathcal{X} = \mathcal{D}_d^\lambda$ (for $\Delta x = \frac{d-2}{d}(1 - \lambda)$ and $\Delta y = \lambda$) and the universal optimal transposition channel $\mathcal{T}_d$ (for $\Delta x = \frac{d-2}{d+1}$ and $\Delta y = \frac{1}{d+1}$); **(d)** the amplitude-damping channel $\mathcal{A}^\lambda$ (for $\Delta y_1 = \lambda$ and $\Delta y_2 = \sqrt{\lambda}$).

COMPETING INTERESTS

The authors declare no competing interests.

AUTHOR'S CONTRIBUTIONS

All authors equally contributed to the original ideas, analytical derivations, and final writing of this manuscript, and gave final approval for publication.

Appendix A: Proofs

In this Section we prove those results reported in the previous Sections for which the proof, being lengthy and not particularly insightful, had only been outlined. The numbering of statements follows that of the previous Sections.

Lemma 4. *The witness threshold $W(\mathcal{X}, w^\pm(\omega))$ of any qubit D_2-covariant channel $\mathcal{X}$ is given by Eq. (7) where*

$$\Delta(\omega) = \begin{cases} d_2\sqrt{1 + \dfrac{c_3^2\omega^2}{d_2^2 - d_3^2}}, & \text{if } |\omega| < \dfrac{d_2^2 - d_3^2}{d_3 c_3}, \\ d_3 + c_3|\omega|, & \text{otherwise}. \end{cases}$$

Proof. Under the assumption of D_2-covariance, take without loss of generality $\vec{c} = (0, 0, c_3)^T$. Then without loss of generality we take $d_2 \geq d_1$ and $c_3 \geq 0$. If $c_3 = 0$ without loss of generality we also take $d_3 \geq d_2$.

First notice that $\vec{y}^*$, which attains the maximum in Eq. (7), lies in the yz plane. Indeed, any ellipse obtained as the intersection of the ellipsoid $|D^{-1}\vec{y}|_2 \leq 1$ and a plane containing the z axis is, up to a z rotation, a subset of the ellipse obtained as the intersection of the ellipsoid $|D^{-1}\vec{y}|_2 \leq 1$ and the yz plane.

The generic vector on the boundary of the yz ellipse can be parametrized as

$$\vec{y} = \left(0, \pm d_2\sqrt{1 - \frac{z^2}{d_3^2}}, z\right)^T,$$

with $z \in [-d_3, d_3]$, and thus the maximum Euclidean distance in Eq. (8) is given by

$$\Delta(\omega) = \max_{z \in [-d_3, d_3]} \sqrt{d_2^2\left(1 - \frac{z^2}{d_3^2}\right) + (z - \omega c_3)^2}. \quad \text{(A1)}$$

By explicit computation one has

$$\frac{d\Delta(\omega)}{dz}$$
$$= \left[d_2^2\left(1 - \frac{z^2}{d_3^2}\right) + (z - \omega c_3)^2\right]^{-\frac{1}{2}} \left[\left(1 - \frac{d_2^2}{d_3^2}\right)z - c_3\omega\right],$$

which is zero for $z^* = \frac{c_3 d_3^2 \omega}{d_3^2 - d_2^2}$, and

$$\frac{d^2\Delta(\omega)}{dz^2}\Big|_{z=z^*} = \left[d_3^2\sqrt{d_2^2\left(1 + \frac{c_3^2\omega^2}{d_2^2 - d_3^2}\right)}\right]^{-1}(d_3^2 - d_2^2),$$

namely z^* attains the maximum in Eq. (A1) whenever $d_2 \geq d_3$. Therefore the maximum is attained by $z = z^*$ iff $-d_3 < z^* \leq d_3$, namely when $|\omega| < \frac{d_2^2 - d_3^2}{d_3 c_3}$, and by $z = \pm d_3$ otherwise. By replacing z^* and $\pm d_3$ in Eq. (A1) the statement follows. □

Theorem 1. *Any given binary conditional probability distribution p is compatible with any given qubit D_2-covariant channel $\mathcal{X}$ if and only if*

$$\max_{\omega \in \Omega}(p \cdot w^\pm(\omega) - W(\mathcal{X}, w^\pm(\omega))) \leq 0,$$

where $\Omega := \{0, \pm\omega_1, \pm\omega_2, \pm 1\} \cap [-1, 1]$.

Proof. The function $f^\pm(\omega) := p^T \cdot w^\pm(\omega) - W(\mathcal{X}, \omega)$ is the minimum of continuous functions $g^\pm(\omega) := p^T \cdot w^\pm(\omega) - \frac{1}{2}(1 + |\omega|)$ and $h^\pm(\omega) := p^T \cdot w^\pm(\omega) - \frac{1}{2}(1 + \Delta(\omega))$. Therefore, $\max_{\omega \in [-1,1]} f^\pm(x)$ is attained by those values of ω maximizing $g^\pm(\omega)$ or $h^\pm(\omega)$, or in the intersections of $g^\pm(\omega)$ and $h^\pm(\omega)$.

The function $g^\pm(\omega)$ is piece-wise linear and attains its maximum on $[-1, 1]$ in 0. The function $h^\pm(\omega)$ is quasi-concave continuous with a continuous derivative. Indeed

$$2\frac{dh^\pm(\omega)}{d\omega}$$
$$= \begin{cases} \pm(p_{1|1} - p_{2|2}) - \dfrac{d_2 c_3^2 \omega}{\sqrt{(d_2^2 - d_3^2)(d_2^2 - d_3^2 + c_3^2\omega^2)}}, & \text{if } |\omega| < \dfrac{d_2^2 - d_3^2}{d_3 c_3}, \\ \pm(p_{1|1} - p_{2|2}) - \text{sgn}(\omega)c_3, & \text{otherwise}, \end{cases}$$

is continuous and

$$2\frac{d^2h^\pm(\omega)}{d\omega^2} = \begin{cases} -\dfrac{d_2 c_3^2 (d_2^2 - d_3^2)^2}{\left[(d_2^2 - d_3^2)(d_2^2 - d_3^2 + c_3^2\omega^2)\right]^{3/2}}, & \text{if } |\omega| < \dfrac{d_2^2 - d_3^2}{d_3 c_3}, \\ 0, & \text{if } |\omega| > \dfrac{d_2^2 - d_3^2}{d_3 c_3}. \end{cases}$$

is non positive. Therefore $h^\pm(\omega)$ attains its maximum on $[-1, 1]$ in 0, ± 1, or in the zero $\pm\omega_1$ of its first derivative.

Due to the piece-wise linearity of $g^\pm(\omega)$ and the quasi-concavity of $h^\pm(\omega)$, since $g^\pm(0) \geq h^\pm(0)$ and $g^\pm(\pm 1) \leq h^\pm(\pm 1)$ one has that $g^\pm(\omega)$ and $h^\pm(\omega)$ intersect in exactly two points $\pm\omega_2 \in [-1, 1]$, thus the statement follows. □

Corollary 2. *Any given binary conditional probability distribution p is compatible with the amplitude-damping channel $\mathcal{A}^\lambda$ if and only if*

$$\left(\sqrt{p_{1|2}p_{2|1}} - \sqrt{p_{1|1}p_{2|2}}\right)^2 \leq \lambda.$$

Proof. One has $c_3 = 1 - \lambda$, $d_2 = \sqrt{\lambda}$, and $d_3 = \lambda$, thus

$$\omega_1 = \sqrt{\frac{\lambda}{(1-\lambda)((1-\lambda) - (p_{1|1} - p_{2|2})^2)}}(p_{1|1} - p_{2|2}),$$

and $\omega_2 = 1$. By explicit computation, the conditions $\omega_1 \in \mathbb{R}$ and $|\omega_1| \leq 1$ are equivalent to $(p_{1|1} - p_{2|2})^2 < 1 - \lambda$ and $(p_{1|1} - p_{2|2})^2 \leq (1 - \lambda)^2$, respectively, thus

$\omega_1 \in [-1, 1]$ is equivalent to $(p_{1|1} - p_{2|2})^2 \leq (1 - \lambda)^2$ for any $\lambda > 0$.

becomes

$$|p_{1|1} - p_{1|2}| - \sqrt{\lambda \left[1 - \frac{(p_{1|1} - p_{2|2})^2}{1 - \lambda}\right]} \leq 0,$$

By explicit computation, the maximum in Eq. (10) is attained at $\omega = \pm\omega_1$ and $\omega = \pm 1$ whenever $|p_{1|1} - p_{2|2}| \leq 1 - \lambda$ and $|p_{1|1} - p_{2|2}| > 1 - \lambda$, respectively. Thus Eq. (10)

whenever $|p_{1|1} - p_{2|2}| \leq 1 - \lambda$, which, by solving in λ, becomes $\lambda_- \leq \lambda \leq \lambda_+$ whenever $\lambda \leq 1 - |p_{1|1} - p_{2|2}|$, where $\lambda_{\pm} = (\sqrt{p_{1|1}p_{2|2}} \pm \sqrt{p_{1|2}p_{2|1}})^2$. By explicit computation $1 - |p_{1|1} - p_{2|2}| \leq \lambda_+$, so the statement follows. $\qquad\square$

[1] Shannon, C. E., *A Note on a Partial Ordering for Communication Channels* , Information and Control **1**, 390 (1958).

[2] As a comparison we notice that, while our approach is top-down, i.e. it aims at characterizing the set of correlations compatible with a given hypothesis, in self-testing [3–8] the approach is bottom-up, i.e. it aims at characterizing the set of hypotheses compatible with a given correlation.

[3] Mayers, D., Yao, A., *Self testing quantum apparatus*, Quantum Information & Computation 4, 273, (2003).

[4] Magniez, F., Mayers, D., Mosca, M., and Ollivier H., *Self-testing of quantum circuits*, in Proceedings of 33rd ICALP, Lecture Notes in Computer Science (Springer, 2006).

[5] Bardyn, C.-E., Liew, T. C. H., Massar, S., McKague, M., and Scarani, V. *Device-independent state estimation based on Bells inequalities*, Phys. Rev. A 80, 062327 (2009).

[6] McKague, M., Yang, T. H., and Scarani, V., *Robust self-testing of the singlet*, J. Phys. A: Math. Theor. 45, 45, 455304 (2012).

[7] Šupić I., Augusiak R., Salavrakos, A., and Acín, A., *Self-testing protocols based on the chained Bell inequalities*, arXiv:1511.09220.

[8] Wang Y., Wu X., and Scarani V., *All the self-testings of the singlet for two binary measurements*, New J. Phys. 18, 025021 (2016).

[9] The role of the circuit within any hypothesis is to describe the space-time structure of the experiment, usually assumed to obey special relativity. Thus, while circuits corresponding to space-like correlations are constrained by the no-signaling principle, those corresponding to time-like correlations are only constrained by the strictly weaker no-signaling-from-the-future principle [10, 11]. As a consequence, the hypotheses falsifiable in a time-like test are inherently more specific than those falsifiable in a space-like test: for instance, while a Bell-test [12–14] can rule out classical theory altogether, a classical model always exists supporting any given time-like correlation.

[10] D'Ariano, G. M., *Operational axioms for C*-algebra representation of transformations*, work presented at the conference proceedings of the *Quantum Theory: Reconsideration of Foundations, 4* held on 11-16 June 2007 at the International Centre for Mathematical Modeling in Physics, Engineering and Cognitive Sciences, Vaxjo University, Sweden.

[11] Ozawa, M., private communication.

[12] J. S. Bell, *On the Einstein-Podolsky-Rosen Paradox*, Physics **1**, 195 (1964).

[13] J. F. Clauser, M. A. Horne, A. Shimony; R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23**, 880 (1969).

[14] B. S. Cirel'son, *Quantum Generalizations of Bell's Inequality*, Lett. Math. Phys. **4**, 93 (1980).

[15] M. Dall'Arno, *Device-independent tests of quantum states*, arXiv:1702.00575.

[16] M. Dall'Arno, S. Brandsen, F. Buscemi, and V. Vedral, arXiv:1609.07846.

[17] P.E. Frenkel, & M. Weiner, *Classical Information Storage in an n-Level Quantum System*, Commun. Math. Phys. **340**, 563 (2015).

[18] Werner, R. F., *Optimal cloning of pure states*, Phys. Rev. A 58, 1827 (1998).

[19] Buscemi, F., D'Ariano, G. M., Perinotti, P., and Sacchi, M. F., *Optimal realization of the transposition maps*, Phys. Lett. A **314**, 374 (2003).

[20] Gallego, R., Brunner, N., Hadley, C., and Acín, A., *Device-Independent Tests of Classical and Quantum Dimensions*, Phys. Rev. Lett. **105**, 230501 (2010).

[21] Hendrych, M., Gallego, R., Mičuda, M., Brunner, N., Acín, A., and Torres, J. P., *Experimental estimation of the dimension of classical and quantum systems*, Nature Phys. **8**, 588-591 (2012).

[22] Ahrens, H., Badziąg, P., Cabello, A. and Bourennane, M. *Experimental Device-independent Tests of Classical and Quantum Dimensions*, Nature Physics **8**, 592 (2012).

[23] Dall'Arno, M., Passaro, E., Gallego, R. and Acín, A., *Robustness of device independent dimension witnesses*, Phys. Rev. A **86**, 042312 (2012).

[24] Chaves, R., Bohr Brask, J., and Brunner, N., *Device-Independent Tests of Entropy*, Phys. Rev. Lett. **115**, 110501 (2015).

[25] Wilde, M. M., *From classical to quantum Shannon theory*, arXiv:1106.1445.

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

[27] Buscemi, F., *Comparison of quantum statistical models: equivalent conditions for sufficiency*, Comm. Math. Phys. **310**(3), 625 (2012).

[28] Helstrom, C. W., *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).