Almost all quantum channels are equidistant

Ion Nechita,[1,2] Zbigniew Puchała,[3,4] Łukasz Pawela,[3,5,a)]
and Karol Życzkowski[4,6]
[1]*Zentrum Mathematik, M5, Technische Universität München, Boltzmannstrasse 3, 85748 Garching, Germany*
[2]*CNRS, Laboratoire de Physique Théorique, IRSAMC, Université de Toulouse, UPS, F-31062 Toulouse, France*
[3]*Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, ulica Bałtycka 5, 44-100 Gliwice, Poland*
[4]*Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, ulica prof. Stanisława Łojasiewicza 11, 30-348 Kraków, Poland*
[5]*Faculty of Applied Physics and Mathematics, National Quantum Information Center, Gdańsk University of Technology, 80-233 Gdańsk, Poland*
[6]*Center for Theoretical Physics, Polish Academy of Sciences, aleja Lotników 32/46, 02-668 Warszawa, Poland*

(Received 13 December 2017; accepted 13 April 2018; published online 8 May 2018)

In this work, we analyze properties of generic quantum channels in the case of large system size. We use random matrix theory and free probability to show that the distance between two independent random channels converges to a constant value as the dimension of the system grows larger. As a measure of the distance we use the diamond norm. In the case of a flat Hilbert-Schmidt distribution on quantum channels, we obtain that the distance converges to $\frac{1}{2} + \frac{2}{\pi}$, giving also an estimate for the maximum success probability for distinguishing the channels. We also consider the problem of distinguishing two random unitary rotations. *Published by AIP Publishing.*
https://doi.org/10.1063/1.5019322

I. INTRODUCTION

For any linear map $\Phi : M_{d_1}(\mathbb{C}) \to M_{d_2}(\mathbb{C})$, we define its *Choi-Jamiołkowski matrix* as

$$J(\Phi) := \sum_{i,j=1}^{d_1} |i\rangle\langle j| \otimes \Phi(|i\rangle\langle j|) \in M_{d_1}(\mathbb{C}) \otimes M_{d_2}(\mathbb{C}). \tag{1}$$

Here $A \otimes B$ denotes the Kronecker product. This isomorphism was first studied by Choi[11] and Jamiołkowski.[26] Note that some authors prefer to add a normalization factor of d_1^{-1} in front of the expression for $J(\Phi)$. Other authors use the other order for the tensor product factors, a choice resulting in an awkward order for the space in which $J(\Phi)$ lives.

The rank of the matrix $J(\Phi)$ is called the *Choi rank* of Φ; it is the minimum number r such that the map Φ can be written as

$$\Phi(\cdot) = \sum_{i=1}^{r} A_i \cdot B_i^*,$$

for some operators $A_i, B_i \in M_{d_2 \times d_1}(\mathbb{C})$.

The *diamond norm* was introduced in quantum information theory by Kitaev,[25] Sec. 3.3, as a counterpart to the 1-norm in the task of distinguishing quantum channels. First, define the $1 \to 1$ norm of a linear map $\Phi : M_{d_1}(\mathbb{C}) \to M_{d_2}(\mathbb{C})$ as

$$\|\Phi\|_{1\to 1} := \sum_{X \neq 0} \frac{\|\Phi(X)\|_1}{\|X\|_1}.$$

a)lpawela@iitis.pl

Kitaev noticed that the $1 \to 1$ norm is not stable under tensor products (as it can easily be seen by looking at the transposition map) and considered the following "regularization":

$$\|\Phi\|_\diamond := \sup_{n\geq 1} \|\Phi \otimes \mathrm{id}_n\|_{1\to 1}.$$

In operator theory, the diamond norm was known before as the *completely bounded trace norm*; indeed, the $1 \to 1$ norm of an operator is the $\infty \to \infty$ norm of its dual, hence the diamond norm of Φ is equal to the completely bounded (operator) norm of Φ^* (see Ref. 35, Chap. 3).

We shall need two simple properties of the diamond norm. First, note that the supremum in the definition can be replaced by taking the value $n = d_1$ (recall that d_1 is the dimension of the input Hilbert space of the linear map Φ); actually, one could also take n equal to the Choi rank of the map Φ; see Ref. 43, Theorem 3.3, or Ref. 44, Theorem 3.66. Second, using the fact that the extremal points of the unit ball of the 1-norm are unit rank matrices, we always have

$$\|\Phi\|_\diamond = \sup\{\|(\Phi \otimes \mathrm{id}_{d_1})(|x\rangle\langle y|)\|_1 \,:\, x, y \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_1},\ \|x\| = \|y\| = 1\}.$$

Moreover, if the map Φ is Hermiticity-preserving (e.g., Φ is the difference of two quantum channels), one can optimize over $x = y$ in the formula above; see Ref. 44, Theorem 3.53.

Given a map Φ, it is in general difficult to compute its diamond norm. Computationally, there is a semidefinite program for the diamond norm,[45] which has a simple form and which has been implemented in various places (see, e.g., Ref. 29). We will bind the diamond norm in terms of the partial trace of the absolute value of the Choi-Jamiołkowski matrix.

The diamond norm finds applications in the problem of quantum channel discrimination. Suppose we have an experiment in which our goal is to distinguish between two quantum channels Φ and Ψ. Each of the channels may appear with probability $\frac{1}{2}$. Then, celebrated results by Helstrom,[21] Holevo,[23] and Kitaev[25] give an upper bound on the probability of correct discrimination

$$p \leq \frac{1}{2} + \frac{1}{4}\|\Phi - \Psi\|_\diamond. \tag{2}$$

The main goal of this work is to study the asymptotic behavior of the diamond norm of the difference of two independent quantum channels. To achieve this, in Sec. II, we find a new upper bound of on the diamond norm of a general map. In our case, it has a nice form

$$\|\Phi - \Psi\|_\diamond \leq \|\mathrm{Tr}_2|J(\Phi - \Psi)|\|_\infty. \tag{3}$$

Next, in Sec. IV A, we prove that the well-known lower bound on the diamond norm, $\|J(\Phi - \Psi)\|_1 \leq \|\Phi - \Psi\|_\diamond$, converges to a finite value for random independent quantum channels Φ and Ψ in the limit $d_{1,2} \to \infty$. We obtain that for a channel sampled from the flat Hilbert-Schmidt distribution, the value of the lower bound is

$$\lim_{d_{1,2}\to\infty} \frac{1}{d_1}\|J(\Phi - \Psi)\|_1 = \frac{1}{2} + \frac{2}{\pi} \text{ a.s.} \tag{4}$$

Finally, in Sec. IV B, we show that the upper bound (3) also converges to the same value as the lower bound. From these results, we infer that for independent random quantum channels sampled from the Hilbert-Schmidt distribution, we have

$$\lim_{d_{1,2}\to\infty} \|\Phi - \Psi\|_\diamond = \frac{1}{2} + \frac{2}{\pi} \text{ a.s.} \tag{5}$$

In particular, the optimal success probability of distinguishing the two channels (in the asymptotical regime) is

$$p \leq \frac{1}{2} + \frac{1}{4}\left(\frac{1}{2} + \frac{2}{\pi}\right) = \frac{5}{8} + \frac{1}{2\pi} \approx 0.7842. \tag{6}$$

Several generalizations of these types of results are gathered in Theorem 7, the main result of this paper.

In Secs. V and VII, we address, respectively, two similar problems: distinguishing a random quantum channels from the maximally depolarizing channel and distinguishing two random unitary channels.

II. SOME USEFUL BOUNDS FOR THE DIAMOND NORM

We discuss in this section some bounds for the diamond norm. For a matrix X, we denote by $\sqrt{X^*X}$ and $\sqrt{XX^*}$ its right and left absolute values, i.e.,

$$\sqrt{X^*X} = V\Sigma V^* \qquad \text{and} \qquad \sqrt{XX^*} = U\Sigma U^*,$$

when $X = U\Sigma V^*$ is the singular value decomposition (SVD) of X. In the case where X is self-adjoint, we obviously have $\sqrt{X^*X} = \sqrt{XX^*}$.

In the result below, the lower bound is well known, while the upper bound appeared in a weaker and less general form in Ref. 27, Theorem 2.

Proposition 1. For any linear map $\Phi : M_{d_1}(\mathbb{C}) \to M_{d_2}(\mathbb{C})$, *we have*

1.
$$\frac{1}{d_1}\|J(\Phi)\|_1 \le \|\Phi\|_\diamond \le \frac{\|\operatorname{Tr}_2 \sqrt{J(\Phi)^*J(\Phi)}\|_\infty + \|\operatorname{Tr}_2 \sqrt{J(\Phi)J(\Phi)^*}\|_\infty}{2}. \tag{7}$$

2. *The above bounds are equal if and only if the positive semidefinite (PSD) matrices* $\varphi := \operatorname{Tr}_2 \sqrt{J(\Phi)^*J(\Phi)}$ *and* $\psi := \operatorname{Tr}_2 \sqrt{J(\Phi)J(\Phi)^*}$ *are both scalar.*

Proof. We start by proving item 1. Consider the semidefinite programs for the diamond norm given in Ref. 45, Sec. 3.2,

$$\begin{array}{llll} & \underline{\text{Primal problem}} & & \underline{\text{Dual problem}} \\ \text{maximize} & \frac{1}{2}\langle X, J(\Phi)\rangle + \frac{1}{2}\langle X^*, J(\Phi)^*\rangle, & \text{minimize} & \frac{1}{2}\|\operatorname{Tr}_2 Y_0\|_\infty + \frac{1}{2}\|\operatorname{Tr}_2 Y_1\|_\infty, \\ \text{subject to} & \begin{bmatrix} \rho_0 \otimes I_{d_2} & X \\ X^* & \rho_1 \otimes I_{d_2} \end{bmatrix} \ge 0, & \text{subject to} & \begin{bmatrix} Y_0 & -J(\Phi) \\ -J(\Phi)^* & Y_1 \end{bmatrix} \ge 0, \\ & \rho_0, \rho_1 \in M_{d_1}^{1,+}(\mathbb{C}), & & Y_0, Y_1 \in M_{d_1 d_2}^{+}(\mathbb{C}). \\ & X \in M_{d_1 d_2}(\mathbb{C}). & & \end{array}$$

The lower and upper bounds will follow from very simple feasible points for the primal, respectively, the dual problems. Let $J(\Phi) = U\Sigma V^*$ be a SVD of the Choi-Jamiołkowski state of the linear map. For the primal problem, consider the feasible point $\rho_{0,1} = d_1^{-1} I_{d_1}$ and $X = d_1^{-1} UV^*$. The value of the primal problem at this point is

$$\frac{1}{2d_1}\langle UV^*, UV^*|J(\Phi)|\rangle + \frac{1}{2d_1}\langle VU^*, |J(\Phi)|VU^*\rangle = \frac{1}{d_1}\|J(\Phi)\|_1,$$

showing the lower bound.

For the upper bound, set $Y_0 = \sqrt{J(\Phi)J(\Phi)^*} = U\Sigma U^*$ and $Y_1 = \sqrt{J(\Phi)^*J(\Phi)} = V\Sigma V^*$, both being PSD matrices. The condition in the dual problem is satisfied,

$$\begin{bmatrix} Y_0 & -J(\Phi) \\ -J(\Phi)^* & Y_1 \end{bmatrix} = \begin{bmatrix} U\Sigma U^* & -U\Sigma V^* \\ -V\Sigma U^* & V\Sigma V^* \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \cdot \left(\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \otimes \Sigma \right) \cdot \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix}^* \ge 0,$$

and the proof of item 1 is complete.

To show statement in item 2 note that the lower bound in (7) can be rewritten as

$$\frac{1}{d_1}\|J(\Phi)\|_1 = \frac{1}{d_1}\operatorname{Tr}\varphi = \frac{1}{d_1}\operatorname{Tr}\psi,$$

and the two bounds are equal exactly when the spectra of φ and ψ are flat. This is also the necessary and sufficient condition for the saturation of the lower bound; see Refs. 30 and 32. ∎

Corollary 2. If the map Φ *is Hermiticity-preserving [i.e., the matrix* $J(\Phi)$ *is self-adjoint], then the inequality in the statement reads simply as*

$$\frac{1}{d_1}\|J(\Phi)\|_1 \le \|\Phi\|_\diamond \le \|\operatorname{Tr}_2 |J(\Phi)|\|_\infty.$$

Let us now characterize the maps Φ for which the upper bound in (7) is saturated. Since our proof is semidefinite programming (SDP)-based, we use the same technique as in Ref. 30, Theorem 18.

Proposition 3. A map Φ saturates the upper bound in (7) if and only if there exist unit vectors $a, b \in \mathbb{C}^{d_1}$ and a unitary operator $W \in \mathcal{U}_{d_1 d_2}$ with the following properties [we write $J = J(\Phi)$]:

- *The vector a achieves the operator norm for* $\operatorname{Tr}_2 \sqrt{JJ^*}$.
- *The vector b achieves the operator norm for* $\operatorname{Tr}_2 \sqrt{J^*J}$.
- $(aa^* \otimes I_{d_2})W = W(bb^* \otimes I_{d_2})$.
- *$J = WP$ for some positive semidefinite operator P; in other words, W is the angular part in some polar decomposition of J.*

Proof. The reasoning follows closely the proof of Ref. 30, Theorem 18, and we only sketch the main lines. Write the SDP in the standard form (see also Ref. 45, Sec. 3.2, for the notation). Optimal matrices for the primal and the dual program are, respectively,

$$A_{opt} = \begin{bmatrix} \rho_0 & . & . & . \\ . & \rho_1 & . & . \\ . & . & \rho_0 \otimes I_{d_2} & W \\ . & . & W^* & \rho_1 \otimes I_{d_2} \end{bmatrix}, B_{opt} = \frac{1}{2}\begin{bmatrix} \|\operatorname{Tr}_2 \sqrt{JJ^*}\|_\infty & . & . & . \\ . & \|\operatorname{Tr}_2 \sqrt{J^*J}\|_\infty & . & . \\ . & . & \sqrt{JJ^*} & . \\ . & . & . & \sqrt{J^*J} \end{bmatrix},$$

where . denotes an unimportant element. Since strong duality holds for our primal-dual pair (Ref. 45, Sec. 3.2), *complementary slackness* holds and we have

$$\begin{aligned} \left(\|\operatorname{Tr}_2 \sqrt{JJ^*}\|_\infty I - \operatorname{Tr}_2 \sqrt{JJ^*}\right)\rho_0 &= 0,\\ \left(\|\operatorname{Tr}_2 \sqrt{J^*J}\|_\infty I - \operatorname{Tr}_2 \sqrt{J^*J}\right)\rho_1 &= 0,\\ U\Sigma U^* R &= U\Sigma V^*(\rho_1 \otimes I_{d_2}),\\ V\Sigma V^* R^* &= V\Sigma U^*(\rho_0 \otimes I_{d_2}), \end{aligned}$$

where $J = U\Sigma V^*$ is the singular value decomposition of J. Using an approximation argument, we can assume J (and thus Σ) is invertible, and thus $W = UV^*$ is unique. We then set $\rho_0 = aa^*$ and $\rho_1 = bb^*$, and the result follows. ∎

Remark 4. The upper bound in (7) can be seen as a strengthening of the following inequality $\|\Phi\|_\diamond \le \|J(\Phi)\|_1$, which already appeared in the literature (e.g., Ref. 44, Sec. 3.4). Indeed, again in terms of φ and ψ, we have $\|\varphi\|_\infty \le \|\varphi\|_1$ and $\|\psi\|_\infty \le \|\psi\|_1$. The inequality in (7) is much stronger: for example, it is always saturated for tensor product matrices $J = J_1 \otimes J_2$ (W from the result above is also a product), whereas the weaker inequality $\|\Phi\|_\diamond \le \|J(\Phi)\|_1$ is saturated in this case only when J_1 has rank one; see Refs. 30 and 32.

III. DISCRIMINATING RANDOM QUANTUM CHANNELS

A. Probability distributions on the set of quantum channels

There are several ways to endow the convex body of quantum channels with probability distributions. In this section, we discuss several possibilities and the relations between them.

Recall that the Choi-Jamiołkowski isomorphism puts into correspondence a quantum channel $\Phi : M_{d_1}(\mathbb{C}) \to M_{d_2}(\mathbb{C})$ with a bipartite matrix $J(\Phi) \in M_{d_1}(\mathbb{C}) \otimes M_{d_2}(\mathbb{C})$ having the following two properties:

- $J(\Phi)$ is positive semidefinite.
- $\operatorname{Tr}_2 J(\Phi) = I_{d_1}$.

The above two properties correspond, respectively, to the fact that Φ is complete positive and trace preserving. Hence, it is natural to consider probability measures on quantum channels obtained as the image measures of probabilities on the set of bipartite matrices with the above properties. Henceforth we will denote the set of all quantum channels as $\Theta(d_1, d_2)$.

Given some fixed dimensions d_1, d_2 and a parameter $s \geq d_1 d_2$, let $G \in M_{d_1 d_2 \times s}(\mathbb{C})$ be a random matrix having independent identically distributed (i.i.d.) standard complex Gaussian entries; such a matrix is called a *Ginibre random matrix*. Define then as

$$W := GG^* \in M_{d_1}(\mathbb{C}) \otimes M_{d_2}(\mathbb{C}), \tag{8}$$

$$D := \left((\mathrm{Tr}_2 W)^{-1/2} \otimes I_{d_2}\right) W \left((\mathrm{Tr}_2 W)^{-1/2} \otimes I_{d_2}\right) \in M_{d_1}(\mathbb{C}) \otimes M_{d_2}(\mathbb{C}). \tag{9}$$

The random matrices W and D are called, respectively, *Wishart* and *partially normalized Wishart*. The inverse square root in the definition of D uses the Moore-Penrose convention if W is not invertible; note however that this is almost never the case since the Wishart matrices with parameter s larger than its size is invertible with unit probability. It is for this reason we do not consider here smaller integer parameters s. Note that the matrix D satisfies the two conditions discussed above: it is positive semidefinite and its partial trace over the second tensor factor is the identity

$$\begin{aligned} \mathrm{Tr}_2\, D &= \mathrm{Tr}_2 \left[\left((\mathrm{Tr}_2 W)^{-1/2} \otimes I_{d_2}\right) W \left((\mathrm{Tr}_2 W)^{-1/2} \otimes I_{d_2}\right)\right] \\ &= (\mathrm{Tr}_2 W)^{-1/2}\,(\mathrm{Tr}_2\, W)\,(\mathrm{Tr}_2 W)^{-1/2} = I_{d_1}. \end{aligned}$$

Hence, there exists a quantum channel Φ_G such that $J(\Phi_G) = D$ (note that D and thus Φ are functions of the original Ginibre random matrix G).

Definition 1. The image measure of the Gaussian standard measure through the map $G \mapsto \Phi_G$ defined in (8), (9), and the equation $J(\Phi_G) = D$ is called the partially normalized Wishart measure *and is denoted by $\gamma^W_{d_1,d_2,s}$.*

Of particular interest is the case $s = d_1 d_2$; the measure obtained in this case will be called the *Hilbert-Schmidt measure* as it is induced from the Hilbert-Schmidt measure on the space of bipartite quantum states by partial normalization[10] and will be denoted by γ^{HS} (see Ref. 40 for the case of random quantum states).

Let us mention here also other measures in the space of quantum operations discussed in the literature. One can use the Stinespring dilation theorem:[41] for any channel $\Phi : M_{d_1}(\mathbb{C}) \to M_{d_2}(\mathbb{C})$, there exists, for some given $s \leq d_1 d_2$, an isometry $V : \mathbb{C}^{d_1} \to \mathbb{C}^{d_2} \otimes \mathbb{C}^s$ such that

$$\Phi(\cdot) = \mathrm{Tr}_2(V \cdot V^*). \tag{10}$$

Definition 2. For any integer parameter s, let $\gamma^{Haar}_{d_1,d_2,s}$ be the image measure of the Haar distribution on isometries V through the map in (10).

Finally, one can consider the Lebesgue measure on the convex body of quantum channels, $\gamma^L_{d_1,d_2}$ which leads to the Euclidean geometry of this set.[42] In this work, we shall however be concerned only with the measure γ^W coming from normalized Wishart matrices. The relations between all these probability measures on the set of quantum channels shall be investigated in some future work.

B. The (two-parameter) subtracted Marčenko–Pastur distribution

In this section, we introduce and study the basic properties of a two-parameter family of probability measures which will appear later in the paper. This family generalizes the symmetrized Marčenko–Pastur distributions from Ref. 37; see also Refs. 18 and 34 for other occurrences of some special cases. Before we start, recall that the Marčenko–Pastur (of free Poisson) distribution of parameter $x > 0$ has density given by Ref. 33, Proposition 12.11,

$$d\mathcal{MP}_x = \max(1-x,0)\delta_0 + \frac{\sqrt{4x-(u-1-x)^2}}{2\pi u} 1_{[a,b]}(u)\, du,$$

where $a = (\sqrt{x}-1)^2$ and $b = (\sqrt{x}+1)^2$.

Definition 3. Let a, b be two free random variables having Marčenko–Pastur distributions with their respective parameters x and y. The distribution of the random variable $a/x - b/y$ is called the subtracted Marčenko–Pastur distribution *with parameters x, y and is denoted by $\mathcal{SMP}_{x,y}$. In other words,*

$$\mathcal{SMP}_{x,y} = D_{1/x}\mathcal{MP}_x \boxplus D_{-1/y}\mathcal{MP}_y, \tag{11}$$

where $D_c\mathcal{P}$ is a distribution of a random variable $Z' = cZ$ provided Z is distributed according to $\mathcal{P}$.

We have the following result.

Proposition 5. Let W_x (respectively, W_y) be two Wishart matrices of parameters (d, s_x) [respectively, (d, s_y)]. Assuming that $s_x/d \to x$ and $s_y/d \to y$ for some constants $x, y > 0$, then, almost surely as $d \to \infty$, we have

$$\lim_{d\to\infty} \|(xd^2)^{-1}W_x - (yd^2)^{-1}W_y\|_1 = \int |u|\, d\mathcal{SMP}_{x,y}(u) =: \Delta(x,y).$$

Proof. The proof follows from standard arguments in random matrix theory and from the fact that the Schatten 1-norm is the sum of the singular values, which are the absolute values of the eigenvalues in the case of self-adjoint matrices. ■

We gather next some properties of the probability measure $\mathcal{SMP}_{x,y}$. Examples of this distribution are shown in Fig. 1.

Proposition 6. Let $x, y > 0$. Then we have the following:

1. *If $x + y < 1$, then the probability measure $\mathcal{SMP}_{x,y}$ has exactly one atom, located at 0, of mass $1-(x+y)$. If $x+y \geq 1$, then $\mathcal{SMP}_{x,y}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

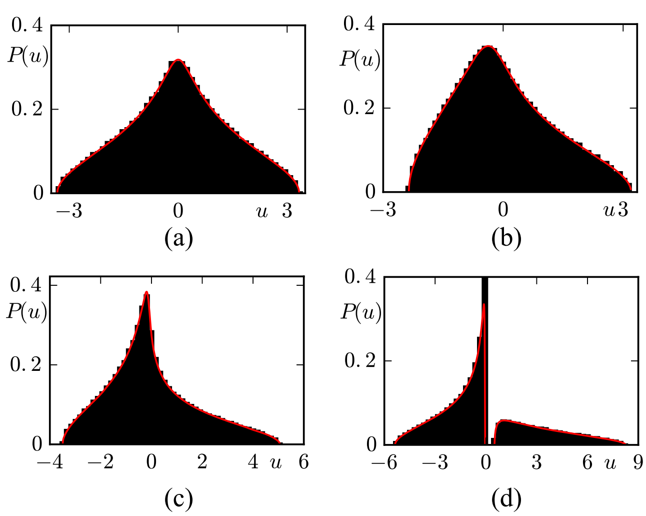

FIG. 1. Subtracted Marčenko–Pastur distribution for (x, y) = (1,1) (a), (1, 2) (b), (0.5, 1) (c), and (0.25, 0.5) (d). The red curve is the plot of (14), while the black histogram corresponds to Monte Carlo simulations. Notice the Dirac mass at zero in the last example.

2. *Define*

$$\begin{aligned} a_{x,y} &= (x-y)(2x+y)(x+2y), \\ b_{x,y} &= 2x^3+2y^3+(x+y)^2+xy(x+y+2), \\ c_{x,y} &= (x-y)(x+y+1-2(x+y)^2), \\ U_{x,y}(u) &= -u^3 a_{x,y}+3u^2 b_{x,y}+3uc_{x,y}+2(x+y-1)^3, \\ T_{x,y}(u) &= (x+y-1-u(x-y))^2+3u(y-x+uxy), \\ Y_{x,y}(u) &= U_{x,y}(u)+\sqrt{\left[U_{x,y}(u)\right]^2-4\left[T_{x,y}(u)\right]^3}. \end{aligned} \tag{12}$$

The support of the absolutely continuous part of $\mathcal{SMP}_{x,y}$ *is the set*

$$\{u : \left[U_{x,y}(u)\right]^2-4\left[T_{x,y}(u)\right]^3 \geq 0\}. \tag{13}$$

3. *On its support, the density of* $\mathcal{SMP}_{x,y}$ *is given by*

$$\frac{d\mathcal{SMP}_{x,y}}{du} = \left|\frac{\left[Y_{x,y}(u)\right]^{\frac{2}{3}}-2^{\frac{2}{3}}T_{x,y}(u)}{2^{\frac{4}{3}}\sqrt{3}\pi u\left[Y_{x,y}(u)\right]^{\frac{1}{3}}}\right|. \tag{14}$$

Proof. The statement regarding the atoms follows from Ref. 5, Theorem 7.4. The formula for the density and Eq. (13) comes from Stieltjes inversion; see, e.g., Ref. 33, Lecture 12. Indeed, since the R-transform of the Marčenko–Pastur distribution $\mathcal{MP}_x$ reads as $R_x(z) = x/(1-z)$, the R-transform of the subtracted measure reads as

$$R(z) = \frac{1}{1-z/x} - \frac{1}{1+z/y}.$$

The Cauchy transform G of $\mathcal{SMP}_{x,y}$ can be obtained from the functional equation

$$R(G) + 1/G = z.$$

This leads to the following third degree polynomial equation for G:

$$zG^3 + [x(1-z)+y(1+z)-1]G^2 + (x-y-xyz)G + xy = 0.$$

Using Cardano's formulas for the solutions of a cubic, we can solve this equation and obtain a solution $G(z)$. The last step is to perform the Stieltjes inversion

$$\mathcal{SMP}_{x,y}(u) = -\frac{1}{\pi}\lim_{\varepsilon\to 0}\Im G(u+\mathrm{i}\varepsilon).$$

■

In the case where $x = y$, some of the formulas from the above result become simpler (see also Ref. 37). When $x = y > 1/2$, the distribution of $\mathcal{SMP}_{x,x}$ is supported between

$$u_\pm = \pm\frac{\sqrt{2+10x-x^2+(x+4)^{\frac{3}{2}}\sqrt{x}}}{\sqrt{2}x}.$$

When $x = y \leq 1/2$, $\mathcal{SMP}_{x,x}$ has an atom in 0 of mass $1-2x$, and its absolutely continuous part is supported on $[u_-, v_-] \cup [v_+, u_+]$, where

$$v_\pm = \pm\frac{\sqrt{2+10x-x^2-(x+4)^{\frac{3}{2}}\sqrt{x}}}{\sqrt{2}x}.$$

Finally, in the case when $x = y = 1$, which corresponds to a flat Hilbert-Schmidt measure on the set of quantum channels, we get that $\Delta(1,1) = \frac{1}{2} + \frac{2}{\pi}$.

IV. THE ASYMPTOTIC DIAMOND NORM OF THE DIFFERENCE OF TWO INDEPENDENT RANDOM QUANTUM CHANNELS

We state here the main result of the paper. For the proof, see Subsections IV A and IV B, each providing one of the bounds needed to conclude.

Theorem 7. *Let* Φ*, respectively,* Ψ*, be two* independent *random quantum channels from* $\Theta(d_1, d_2)$ *having* γ^W *distribution with parameters* (d_1, d_2, s_x)*, respectively,* (d_1, d_2, s_y)*. Then, almost surely as* $d_{1,2} \to \infty$ *in such a way that* $s_x/(d_1 d_2) \to x$*,* $s_y/(d_1 d_2) \to y$ *(for some positive constants* x*,* y*) and* $d_1 \ll d_2^2$*,*

$$\lim_{d_{1,2}\to\infty} \|\Phi - \Psi\|_\diamond = \Delta(x, y) = \int |u| \, d\mathcal{SMP}_{x,y}(u).$$

Proof. The proof follows from Theorems 10 and 14, which give the same asymptotic value. ■

Remark 8. We think that the condition $d_1 \ll d_2^2$ *in the statement is purely technical and could be replaced by a much weaker condition.*

Corollary 9. Combining Theorem 7 with Helstrom's theorem for quantum channels, we get that the optimal probability p of distinguishing two quantum channels is equal to

$$p = \frac{5}{8} + \frac{1}{2\pi}. \tag{15}$$

Additionally, any maximally entangled state may be used to achieve this value.

A. The lower bound

In this section, we compute the asymptotic value of the lower bound in Theorem 7. Given two random quantum channels Φ, Ψ, we are interested in the asymptotic value of the quantity $d_1^{-1} \|J(\Phi - \Psi)\|_1$.

Theorem 10. *Let* Φ*, respectively,* Ψ*, be two* independent *random quantum channels from* $\Theta(d_1, d_2)$ *having* γ^W *distribution with parameters* (d_1, d_2, s_x)*, respectively,* (d_1, d_2, s_y)*. Then, almost surely as* $d_{1,2} \to \infty$ *in such a way that* $s_x/(d_1 d_2) \to x$ *and* $s_y/(d_1 d_2) \to y$ *for some positive constants* x*,* y*,*

$$\lim_{d_{1,2}\to\infty} \frac{1}{d_1} \|J(\Phi - \Psi)\|_1 = \Delta(x, y) = \int |u| \, d\mathcal{SMP}_{x,y}(u).$$

The proof of this result (as well as the proof of Theorem 10) uses in a crucial manner the approximation result for partially normalized Wishart matrices.

Proposition 11. Let $W \in M_{d_1}(\mathbb{C}) \otimes M_{d_2}(\mathbb{C})$ *be a random Wishart matrix of parameters* $(d_1 d_2, s)$ *and consider its "partial normalization" D as in (9). Then, almost surely as* $d_{1,2} \to \infty$ *in such a way that* $s \sim t d_1 d_2$ *for a fixed parameter* $t > 0$*,*

$$\|D - (t d_1 d_2^2)^{-1} W\|_\infty = O(d_2^{-2}).$$

Note that in the statement above, the matrix W is not normalized; we have

$$\frac{1}{d_1 d_2} \sum_{i=1}^{d_1 d_2} \delta_{\lambda_i((d_1 d_2)^{-1} W)} \to \mathcal{MP}_t,$$

the Marčenko–Pastur distribution of parameter t. In other words, $W = GG^*$, where G is a random matrix of size $d_1 d_2 \times s$, having i.i.d. standard complex Gaussian entries.

Let us introduce the random matrices

$$X = (td_1 d_2^2)^{-1} \operatorname{Tr}_2 W \quad \text{and} \quad Y = X^{-1/2} \otimes I_{d_2}.$$

The first observation we make is that the random matrix X is also a (rescaled) Wishart matrix. Indeed, the partial trace operation can be seen, via duality, as a matrix product, so we can write

$$X = \frac{1}{td_1 d_2^2} \tilde{G}\tilde{G}^*,$$

where $\tilde{G}$ is a complex Gaussian matrix of size $d_1 \times d_1 s$; remember that s scales like $td_1 d_2$. Since, in our model, both d_1, d_2 grow to infinity, the behavior of the random matrix X follows from Ref. 15.

Lemma 12. As $d_{1,2} \to \infty$, the random matrix $\sqrt{t}d_2(X - I_{d_1})$ converges in moments toward a standard semicircular distribution. Moreover, almost surely, the limiting eigenvalues converge to the edges of the support of the limiting distribution,

$$\begin{aligned} \sqrt{t}d_2 \lambda_{\min}(X - I_{d_1}) &\to -2, \\ \sqrt{t}d_2 \lambda_{\max}(X - I_{d_1}) &\to 2. \end{aligned}$$

Proof. The proof is a direct application of Ref. 15, Corollary 2.5 and Theorem 2.7; we just need to check the normalization factors. In the setting of Ref. 15, Sec. 2, the Wishart matrices are not normalized, so the convergence result deals with the random matrices (here $d = d_1$ and $s = td_1 d_2^2$),

$$\sqrt{t}d_1 d_2 \left(\frac{\tilde{G}\tilde{G}^*}{td_1^2 d_2^2} - \frac{I_{d_1}}{d_1} \right) = \sqrt{t}d_2(X - I_{d_1}).$$

■

We look now for a similar result for the matrix Y; the result follows by functional calculus.

Lemma 13. Almost surely as $d_{1,2} \to \infty$, the limiting eigenvalues of the random matrix $\sqrt{t}d_2(Y - I_{d_1 d_2})$ converge, respectively, to ± 1,

$$\begin{aligned} \sqrt{t}d_2 \lambda_{\min}(Y - I_{d_1 d_2}) &\to -1, \\ \sqrt{t}d_2 \lambda_{\max}(Y - I_{d_1 d_2}) &\to 1. \end{aligned}$$

Proof. By functional calculus, we have $\lambda_{\max}(Y) = [\lambda_{\min}(X)]^{-1/2}$, so, using the previous lemma, we get

$$\lambda_{\max}(Y) = \left[1 - \frac{2}{\sqrt{t}d_2} + o(d_2^{-1}) \right]^{-1/2} = 1 + \frac{1}{2}\frac{2}{\sqrt{t}d_2} + o(d_2^{-1}),$$

and the conclusion follows. The case of $\lambda_{\min}(Y)$ is similar. ■

We have now all the ingredients to prove Proposition 11.

Proof of Proposition 11. We have

$$\begin{aligned} \|D - (td_1 d_2^2)^{-1} W\|_\infty &= \|(td_1 d_2^2)^{-1} (YWY - W)\|_\infty \\ &= (td_1 d_2^2)^{-1} \|(Y_i - I)W_i Y_i + W_i(Y_i - I)\|_\infty \\ &\le (td_1 d_2^2)^{-1} \|Y_i - I\|_\infty \|W_i\|_\infty (1 + \|Y_i\|_\infty) \\ &= \frac{t^{-3/2}}{d_2^2} \cdot \sqrt{t}d_2 \|Y_i - I\|_\infty \cdot (d_1 d_2)^{-1} \|W_i\|_\infty \cdot (1 + \|Y_i\|_\infty). \end{aligned}$$

Note that, almost surely, the three random matrix norms in the last equation above converge, respectively, to the following finite quantities:

$$\sqrt{t} d_2 \|Y_i - I\|_\infty \to 1,$$
$$(d_1 d_2)^{-1} \|W_i\|_\infty \to (\sqrt{t} + 1)^2,$$
$$1 + \|Y_i\|_\infty \to 1.$$

The first and the third limit above follow from Lemma 13, while the second one is the Bai-Yin theorem, Ref. 3, Theorem 2, or Ref. 2, Theorem 5.11. ■

Let us now prove Theorem 10.

Proof of Theorem 10. The result follows easily by approximating the partially normalized Wishart matrices with scalar normalizations. By the triangle inequality, with $D_x := J(\Phi)$ and $D_y := J(\Psi)$, we have

$$\begin{aligned} &\left| \frac{1}{d_1} \|D_x - D_y\|_1 - \frac{1}{d_1} \|(x d_1 d_2^2)^{-1} W_x - (y d_1 d_2^2)^{-1} W_y\|_1 \right| \\ &\quad \le \frac{1}{d_1} \|D_x - (x d_1 d_2^2)^{-1} W_x\|_1 + \frac{1}{d_1} \|D_y - (y d_1 d_2^2)^{-1} W_y\|_1 \\ &\quad \le d_2 \|D_x - (x d_1 d_2^2)^{-1} W_x\|_\infty + d_2 \|D_y - (y d_1 d_2^2)^{-1} W_y\|_\infty. \end{aligned}$$

The conclusion follows from Propositions 5 and 11. ■

B. The upper bound

The core technical result of this work consists of deriving the asymptotic value of the upper bound in Theorem 7. Given two random quantum channels Φ, Ψ, we are interested in the asymptotic value of the quantity $\|\mathrm{Tr}_2 |J(\Phi - \Psi)|\|_\infty$.

Theorem 14. *Let* Φ, *respectively,* Ψ, *be two* independent *random quantum channels from* $\Theta(d_1, d_2)$ *having* γ^W *distribution with parameters* (d_1, d_2, s_x), *respectively,* (d_1, d_2, s_y). *Then, almost surely as* $d_{1,2} \to \infty$ *in such a way that* $s_x/(d_1 d_2) \to x$, $s_y/(d_1 d_2) \to y$ *(for some positive constants* x, y*) and* $d_1/d_2^2 \to 0$,

$$\lim_{d_{1,2}\to\infty} \| \mathrm{Tr}_2 |J(\Phi - \Psi)| \|_\infty = \Delta(x, y) = \int |u| \, d\mathcal{SMP}_{x,y}(u).$$

The proof of Theorem 14 is presented at the end of this section. It is based on the following lemma which appears in Ref. 17; see also Ref. 7, Eq. (5.10), or Ref. 6, Chap. X.

Lemma 15. For any matrices A, B *of size* d, *the following holds:*

$$\| |A| - |B| \|_\infty \le C \log d \, \|A - B\|_\infty, \tag{16}$$

for a universal constant C *which does not depend on the dimension* d.

For the sake of completeness, we give here a proof, relying on a similar estimate for the Schatten classes proved in Ref. 17.

Proof. Using Ref. 17, Theorem 8, we have, for any $p \in [2, \infty)$

$$\begin{aligned} \| |A| - |B| \|_\infty &\le \| |A| - |B| \|_p \\ &\le 4(1 + cp)\|A - B\|_p \\ &\le 4(1 + cp) d^{1/p} \|A - B\|_\infty, \end{aligned}$$

for some universal constant $c \ge 1$. Choosing $p = \log d$ gives the desired bound, for d large enough. The case of small values of d is obtained by a standard embedding argument. ■

Proof of Theorem 14. Using the triangle inequality and Lemma 15, we first prove an approximation result [as before, we write $D_x := J(\Phi)$ and $D_y := J(\Psi)$]

$$
\begin{aligned}
\big| \, \| \operatorname{Tr}_2 |D_x - D_y| \, \|_\infty - \| \operatorname{Tr}_2 |(xd_1d_2^2)^{-1}W_x - (yd_1d_2^2)^{-1}W_y| \, \|_\infty \big| \\
&\leq \| \operatorname{Tr}_2 |D_x - D_y| - \operatorname{Tr}_2 |(xd_1d_2^2)^{-1}W_x - (yd_1d_2^2)^{-1}W_y| \|_\infty \\
&= \| \operatorname{Tr}_2 \left(|D_x - D_y| - |(xd_1d_2^2)^{-1}W_x - (yd_1d_2^2)^{-1}W_y| \right) \|_\infty \\
&\leq d_2 \| \, |D_x - D_y| - |(xd_1d_2^2)^{-1}W_x - (yd_1d_2^2)^{-1}W_y| \, \|_\infty \\
&\leq C d_2 \log(d_1d_2) \| (D_x - D_y) - ((xd_1d_2^2)^{-1}W_x - (yd_1d_2^2)^{-1}W_y) \|_\infty \\
&\leq C d_2 \log(d_1d_2) \left(\|D_x - (xd_1d_2^2)^{-1}W_x\|_\infty + \|D_y - (yd_1d_2^2)^{-1}W_y\|_\infty \right) \\
&= \frac{\log(d_1d_2)}{d_2} O(1) \to 0,
\end{aligned}
$$

where we have used Proposition 11 and the fact that $d_1 \ll d_2^2 \implies \log(d_1) \ll d_2$. This proves the approximation result, and we focus now on the simpler case of Wishart matrices. Let us define

$$
\begin{aligned}
Z &:= (xd_1d_2)^{-1}W_x - (yd_1d_2)^{-1}W_y, \\
\tilde{Z}_1 &:= \operatorname{tr}_2(|Z|) = \operatorname{Tr}_2 |(xd_1d_2^2)^{-1}W_x - (yd_1d_2^2)^{-1}W_y|.
\end{aligned}
$$

It follows from Ref. 22, Proposition 4.4.9, that the random matrix Z converges almost surely (see Appendix A for the definition of almost sure convergence for a sequence of random matrices) to a non-commutative random variable having distribution $\mathcal{SMP}_{x,y}$; see (11). Moreover, using a standard strong convergence argument,[31] the extremal eigenvalues of Z converge almost surely to the extremal points of the support of the limiting probability measure $\mathcal{SMP}_{x,y}$. Hence, the almost sure convergence extends from the traces of the powers of Z to any continuous bounded function (on the support of $\mathcal{SMP}_{x,y}$), in particular, to the absolute value, i.e., to $|Z|$. From Proposition 23, the asymptotic spectrum of the random matrix $\tilde{Z}_1$ is flat, with all the eigenvalues being equal to

$$
a = \lim_{d_1,d_2 \to \infty} \mathbb{E} \frac{\operatorname{Tr} |(xd_1d_2)^{-1}W_x - (yd_1d_2)^{-1}W_y|}{d_1d_2} = \int |u| d\mathcal{SMP}_{x,y}(u),
$$

which, by Proposition 5, is equal to $\Delta(x, y)$, finishing the proof. ∎

V. DISTANCE TO THE DEPOLARIZING CHANNEL

In this section, we derive the asymptotic distance between a random quantum channel Φ and the *maximally depolarizing channel*

$$
\Psi_{\text{dep}} : M_{d_1}(\mathbb{C}) \to M_{d_2}(\mathbb{C}), \qquad \Psi_{\text{dep}}(X) = \frac{\operatorname{Tr}(X)}{d_2} I_{d_2}.
$$

Let us define the function g: $(1/4, \infty) \to (0, \infty)$,

$$
\begin{aligned}
g(x) := \frac{3}{2} - x + \frac{\sqrt{4x-1}(2x+1)}{2\pi x} \\
- \frac{1}{\pi} \left((x-1) \arctan\left(\frac{3x-1}{(x-1)\sqrt{4x-1}} \right) + \arctan\left(\sqrt{4x-1}\right) + x \arctan\left(\frac{1}{\sqrt{4x-1}} \right) \right).
\end{aligned}
$$

Theorem 16. *Let Φ be a random quantum channel from $\Theta(d_1, d_2)$ having distribution γ^W with parameters (d_1, d_2, s_x). Then, almost surely as $d_1, d_2 \to \infty$ and $s_x \sim xd_1d_2$, we have*

$$
\lim_{d_1,d_2 \to \infty} \|\Phi - \Psi_{dep}\|_\diamond = \int \left| \frac{u}{x} - 1 \right| d\mathcal{MP}_x(u) = \begin{cases} 2(1-x) & \text{if } x \in (0, 1/4], \\ g(x) & \text{if } x \in (1/4, 1), \\ g(x) + x - 1 & \text{if } x \in [1, \infty). \end{cases} \tag{17}
$$

In the case $x = 1$, the limit above reads as $3\sqrt{3}/(2\pi)$.

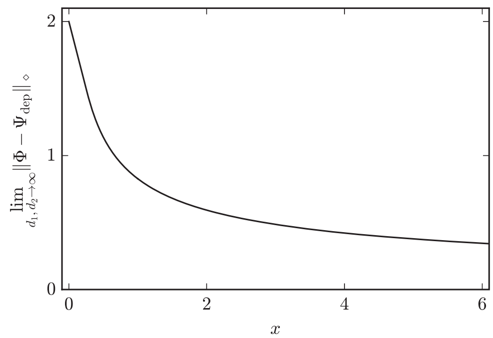

FIG. 2. The asymptotic diamond-norm distance between a random quantum channel and the maximally depolarizing channel as a function of the channel parameter x.

Remark 17. We plot in Fig. 2 the value of the limit in (17) as a function of x. One can show that the limit is a decreasing function of x, converging to 0 *as* $x \to \infty$. *The function behaves as* $8/(3\pi)x^{-1/2}$ *as* $x \to \infty$.

Proof. We analyze separately the lower bound and the upper bound from Proposition 1. First, let us denote by D_x the Choi-Jamiołkowski matrix of the channel Φ, and note that $J(\Psi_{\text{dep}}) = d_2^{-1} I_{d_1 d_2}$. For the lower bound, first show that we can approximate the random matrix D_x by a rescaled Wishart matrix,

$$\begin{aligned}\left| \frac{1}{d_1} \|D_x - d_2^{-1} I_{d_1 d_2}\|_1 - \frac{1}{d_1} \|(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}\|_1 \right| &\le \frac{1}{d_1} \|D_x - (x d_1 d_2^2)^{-1} W_x\|_1 \\ &\le d_2 \|D_x - (x d_1 d_2^2)^{-1} W_x\|_\infty,\end{aligned}$$

which converges almost surely to 0, by Proposition 11. The quantity with which we approximate is then

$$\frac{1}{d_1} \|(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}\|_1 = \frac{1}{d_1 d_2} \sum_{i=1}^{d_1 d_2} |\lambda_i[(x d_1 d_2)^{-1} W_x - I_{d_1 d_2}]|. \tag{18}$$

The quantity above converges almost surely, as $d_1 d_2 \to \infty$, towards

$$\int \left| \frac{u}{x} - 1 \right| d\mathcal{MP}_x(u).$$

Let us now show that the upper bound from Proposition 1 converges to the same quantity. We follow the same steps as in the proof of Theorem 14: we first approximate the matrix D_x by a rescaled Wishart random matrix, and then we argue that the partial trace appearing in the bound has "flat" eigenvalues, allowing us to replace the operator norm by the normalized trace. For the approximation step, we get, using again Proposition 11,

$$\begin{aligned}&\left| \, \| \operatorname{Tr}_2 |D_x - d_2^{-1} I_{d_1 d_2}| \, \|_\infty - \| \operatorname{Tr}_2 |(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}| \, \|_\infty \right| \\ &\quad \le \|\operatorname{Tr}_2 \left(|D_x - d_2^{-1} I_{d_1 d_2}| - |(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}| \right) \|_\infty \\ &\quad \le d_2 \| \, |D_x - d_2^{-1} I_{d_1 d_2}| - |(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}| \, \|_\infty \\ &\quad \le C d_2 \log(d_1 d_2) \|D_x - (x d_1 d_2^2)^{-1} W_x\|_\infty \\ &\quad = \frac{\log(d_1 d_2)}{d_2} O(1) \to 0 \quad \text{almost surely.}\end{aligned}$$

We focus now on the quantity $\| \operatorname{Tr}_2 |(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}| \, \|_\infty$. From Proposition 23, the spectrum of the random matrix $\operatorname{Tr}_2 |(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}|$ is flat, so its operator norm has the same limit as $d_1^{-1} \operatorname{Tr} |(x d_1 d_2^2)^{-1} W_x - d_2^{-1} I_{d_1 d_2}|$, which is the same as (18), finishing the proof. ■

VI. DISTANCE TO THE NEAREST UNITARY CHANNEL

In this section, we consider an asymptotic distance between a random quantum channel $\Phi : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ and a unitary channel. First we note that if a quantum channel Φ is an interior point of the set of channels, then the best distinguishable one Ψ is some unitary channel.[36] Below we show that in the case of $d \to \infty$ almost all quantum channels are perfectly distinguishable from any unitary channel. To see it, we write

$$\begin{aligned}\min_{\Psi_U} \|\Phi - \Psi_U\|_\diamond &\geq \frac{1}{d} \min_{\Psi_U} \|J(\Phi) - J(\Psi_U)\|_1 = \min_U \frac{1}{d} \|J(\Phi) - |U\rangle\langle U|\|_1 \geq \min_{|x\rangle} \frac{1}{d} \|J(\Phi) - d|x\rangle\langle x|\|_1 \\ &\geq \min_{|x\rangle} 2[1 - F(J(\Phi)/d, |x\rangle\langle x|)] = 2 - 2\|J(\Phi)/d\|_\infty. \end{aligned} \tag{19}$$

In the above we have used the inequality between the diamond norm and the trace norm of Choi-Jamiołkowski matrices, see Proposition 1, and next the Fuchs-van de Graaf inequality[19] involving the trace norm and fidelity function $F(\rho, \sigma) = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2$. Next we use the fact that the largest eigenvalue os matrix $J(\Phi)/d$ tends to 0 almost surely.

VII. DISTANCE BETWEEN RANDOM UNITARY CHANNELS

We consider in this section the problem of distinguishing two unitary channels,

$$\Phi(X) = UXU^* \qquad \text{and} \qquad \Psi(X) = VXV^*, \tag{20}$$

where U, V are two $d \times d$ unitary operators. The diamond norm of the difference $\|\Phi - \Psi\|_\diamond$ has already been considered in the literature, and we gather below the results from Ref. 44, Theorem 3.57, and Ref. 28, Theorem 12.

Proposition 18. For any two unitary operators U, V, the diamond norm of the difference of the unitary channels induced by U, V is given by

- $\|\Phi - \Psi\|_\diamond = 2\sqrt{1 - \nu(U^*V)^2}$, *where $\nu(U^*V)$ is the smallest absolute value of an element in the numerical range of the unitary operator U^*V. In other words, $\nu(U^*V)$ is the radius of the largest open disc centered at the origin which does not intersect the convex hull of the eigenvalues of U^*V (*i.e., *the numerical range).*
- $\|\Phi - \Psi\|_\diamond = 2R(U^*V)$, *where $R(U^*V)$ is the radius of the smallest disc (not necessarily centered at the origin) containing all the eigenvalues of U^*V.*
- *Let 2α be the smallest arc containing the spectrum of U^*V. Then,*

$$\|\Phi - \Psi\|_\diamond = \begin{cases} 2\sin\alpha, & \textit{if } \alpha < \pi/2, \\ 2, & \textit{if } \alpha \geq \pi/2. \end{cases}$$

We represent in Fig. 3 the eigenvalues of the operator $W := U^*V$ and the numerical range of W. Recall that the *numerical range* of an operator A is the set

$$\mathcal{N}(A) = \{\langle x, Ax\rangle : x \in \mathbb{C}^d, \|x\| = 1\}.$$

The numerical range is a convex body (Ref. 24, Chap. 1), and in the case where A is a normal operator ($AA^* = A^*A$) it coincides with the convex hull of the spectrum. One remarkable fact about the results in the proposition above is that two unitary operations Φ and Ψ become *perfectly* distinguishable as soon as the convex hull of the eigenvalues of U^*V contains the origin, (Ref. 44, Theorem 3.57) and (Ref. 28, Theorem 12).

We consider next random unitary operators U, V. We analyze Haar-distributed operators and then the case where U and V are sampled from the distribution of two independent unitary Brownian

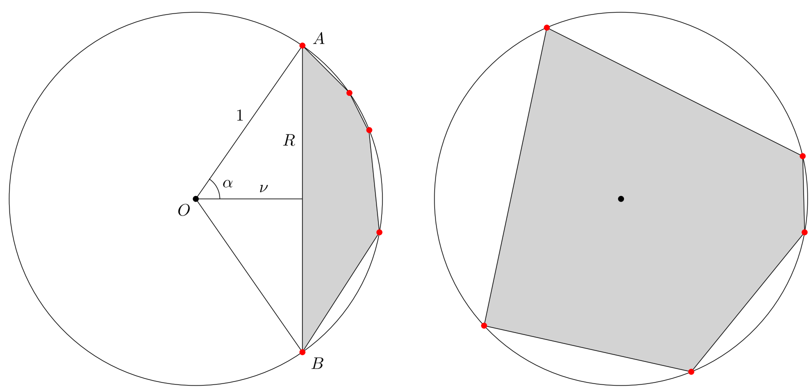

FIG. 3. The eigenvalues of the unitary operator $A = U^*V$ (red dots, here $d = 5$) and its numerical range $\mathcal{N}(A)$ (gray filled area). On the left, the eigenvalues span an arc of length smaller than π, so the quantities ν and R are non-trivial. On the right, the eigenvalues span an arc larger than half a circle, so the origin belongs to the numerical range; here, $\nu = 0$ and $R = 1$.

motions stopped at different times. For independent, Haar-distributed unitary operators, in the limit of large dimension, the corresponding channels become perfectly distinguishable.

Proposition 19. Let $U, V \in \mathcal{U}(d)$ be two independent random variables, at least one of them being Haar-distributed. Then, with overwhelming probability as $d \to \infty$, the quantum channels Φ and Ψ from (20) become perfectly distinguishable: for d large enough,

$$\mathbb{P}\left[\|\Phi - \Psi\|_\diamond = 2\right] \geq 1 - \exp\left(-\frac{\log 2}{2} d^2\right).$$

Remark 20. The statement above includes the case where U is a Haar-distributed random unitary matrix, and V is the identity operator (hence, Ψ is the identity channel).

Proof. From the hypothesis and the left/right invariance of the Haar distribution, it follows that the random matrix $W = U^*V$ is Haar-distributed. The estimate follows from Ref. 4, Sec. 3.1, where the probability of a Haar unitary matrix not having any eigenvalues in a given arc is related to a Toeplitz determinant; see Eq. (3.1) in Ref. 4. ∎

Let us now consider the case where the operators U and V are elements of two independent unitary Brownian motion processes. We shall not give the definition of this process, referring the reader to, e.g., Refs. 8, 9, and 38. We shall only need here the following result of Biane, giving the asymptotic support of a unitary Brownian motion stopped at time t.

Proposition 21 (Ref. 9, Proposition 10). *Let $(U_t)_{t\geq 0}$ be a unitary Brownian motion on $\mathcal{U}(d)$ starting at the identity. Then, asymptotically as $d \to \infty$, the support of the eigenvalue distribution (on the unit circle) of the operator U_t is the full circle if $t \geq 4$ and the arc*

$$\left\{\exp(i\alpha) : |\alpha| \leq \frac{1}{2}\sqrt{t(4-t)} + \arccos(1 - t/2)\right\} \qquad \text{if } 0 \leq t < 4.$$

As a direct application of this result, we obtain the diamond norm of the difference of two unitary quantum channels stemming from independent unitary Brownian motions.

Proposition 22. Let $(U_s)_{s\geq 0}$ and $(V_t)_{t\geq 0}$ be two independent unitary Brownian motions and consider the random unitary quantum channels Φ_s and Ψ_t from (20) obtained from the operators U_s and, respectively, V_t. Then, almost surely,

$$\lim_{d\to\infty} \|\Phi_s - \Psi_t\|_\diamond = \begin{cases} 2\sin\left[\frac{1}{2}\sqrt{(s+t)(4-s-t)} + \arccos(1-(s+t)/2)\right], & \text{if } s+t < \tau, \\ 2, & \text{if } s+t \geq \tau, \end{cases}$$

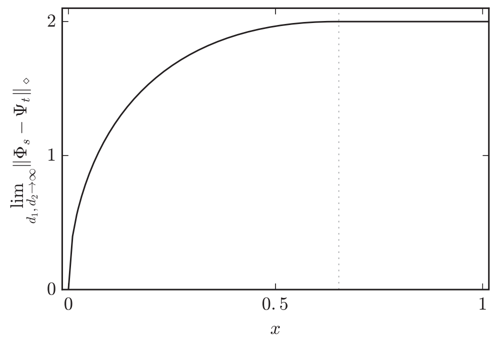

FIG. 4. The diamond norm of a difference of two random unitary channels coming from two independent unitary Brownian motions stopped at times s and t, as a function of $s + t$.

where $\tau \approx 0.6528$ *is the unique solution of the equation*

$$\frac{1}{2}\sqrt{t(4-t)} + \arccos(1 - t/2) = \pi/2$$

on $(0, 4)$.

Proof. The proof is an easy consequence of Biane's result (more precisely, of its "strong" formulation from Ref. 12, Theorem 1.1), once we notice that the random unitary matrix $U_s^* V_t$ has the same distribution as W_{s+t}, where $W_\cdot$ is another unitary Brownian motion. We plot the diamond norm as a function of $s + t$ in Fig. 4. ■

VIII. CONCLUDING REMARKS

In this work, we analyzed properties of generic quantum channels concentrating on the case of large system size. Using tools provided by the theory of random matrices and the free probability calculus, we showed that the diamond norm of the difference between two random channels asymptotically tends to a constant specified in Theorem 7. In the case of channels corresponding to the simplest case $x = y = 1$, the limit value of the diamond norm of the difference is $\Delta(1, 1) = 1/2 + 2/\pi$. Based on these results, in Fig. 5, we provide a sketch of the set of quantum channels. In Fig. 6, we illustrate the convergence of the upper and lower bound to the value $1/2 + 2/\pi$. This statement allows us to quantify the mean distinguishability between two random channels

To arrive at this result, we considered an ensemble of normalized random density matrices, acting on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ and distributed according to the flat (Hilbert-Schmidt) measure. Such matrices can be generated with the help of a complex Ginibre matrix G as $\rho = GG^*/\mathrm{Tr}GG^*$.

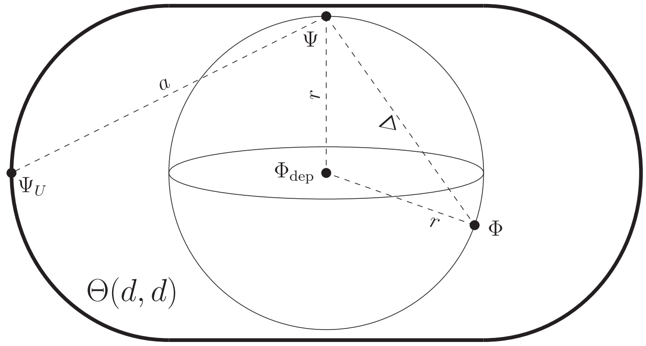

FIG. 5. Sketch of the set $\Theta(d, d)$ of all channels acting on d-dimensional states. A generic channel Φ belongs to a sphere of radius $r = 3\sqrt{3}/2\pi$, centered at the maximally depolarizing channel, Φ_{dep}, in the metric induced by the diamond norm. The distance between generic channels, Φ, Ψ is $\Delta = 1/2 + 2/\pi$, while the distance to the nearest unitary channel reads as $a = 2$.

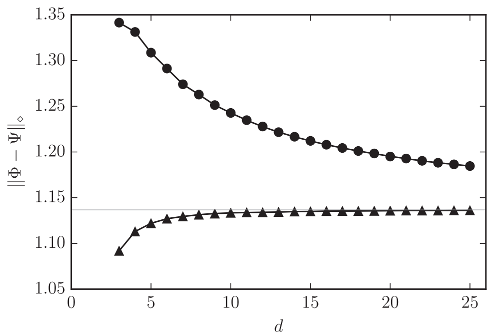

FIG. 6. The convergence of upper (circles) and lower (triangles) bounds on the distance between two random quantum channels sampled from the Hilbert-Schmidt distribution ($d_1 = d_2 = d$). The results were obtained via Monte Carlo simulation with 100 samples for each data point.

In the simplest case of square matrices G of order $d = d_1^2$, the average trace distance of a random state ρ from the maximally mixed state $\rho_* = I/d$ behaves asymptotically as $||\rho - \rho_*||_1 \to 3\sqrt{3}/4\pi$.[37] However, analyzing both reduced matrices $\rho_A = \mathrm{Tr}_B\rho$ and $\rho_B = \mathrm{Tr}_A\rho$ we can show that they become ϵ close to the maximally mixed state in sense of the operator norm so that their smallest and largest eigenvalues do coincide. This is visualized in Fig. 7.

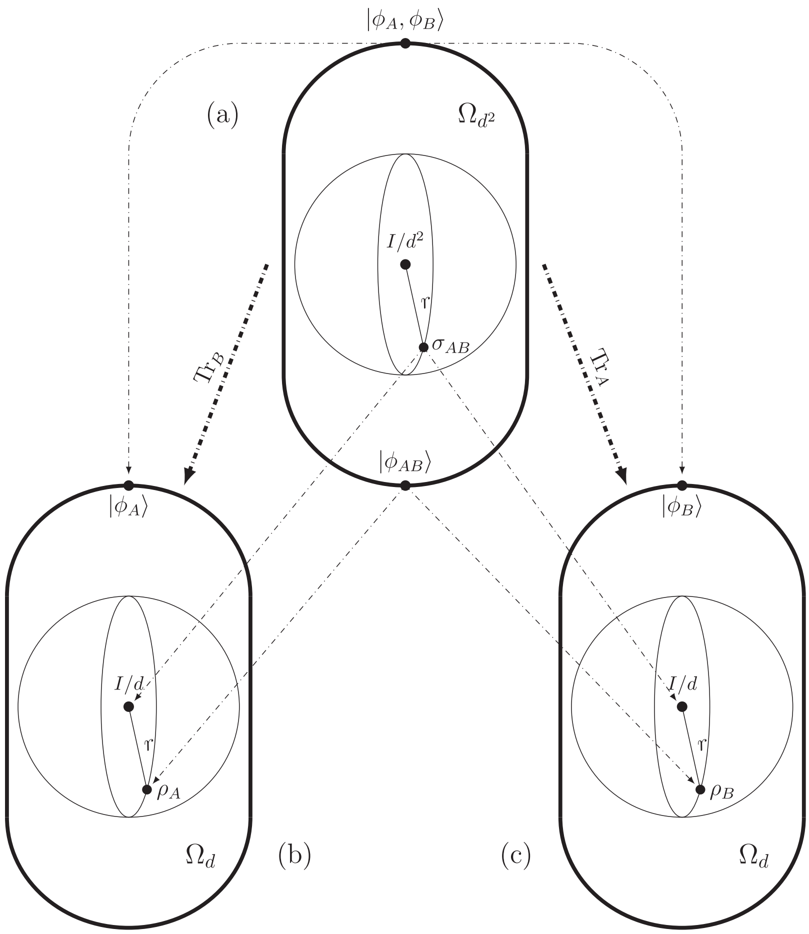

FIG. 7. Set of all bipartite quantum states of dimension d^2, Ω_{d^2}, (a) and its partial traces (b) and (c) containing states of dimension d. A generic bipartite state σ_{AB}, distant $r = 3\sqrt{3}/4\pi$ from the maximally mixed state I/d^2, is mapped into $\sigma_A \approx \sigma_B \approx I/d$, while a typical pure state $|\phi_{AB}\rangle$ is sent into a generic mixed state $\rho_A \equiv \rho_B$ distant r from I/d.

This observation implies that the state ρ can be directly interpreted as a Jamiołkowski state J representing a stochastic map Φ, as its partial trace ρ_A is proportional to identity. Furthermore, as it becomes asymptotically equal to the other partial trace ρ_B, it follows that a generic quantum channel (stochastic map) becomes unital and thus bistochastic.

The partial trace of a random bipartite state is shown to be close to identity provided the support of the limiting measure characterizing the bipartite state is bounded. In particular, this holds for a family of subtract Marčenko–Pastur distributions defined in Eq. (11) as a free additive convolution of two rescaled Marčenko–Pastur distributions with different parameters and determining the density of a difference of two random density matrices. In this way, we could establish the upper bound for the average diamond norm between two channels and show that it asymptotically converges to the lower bound $\Delta(x, y)$ given in Theorem 10. The results obtained can be understood as an application of the measure concentration paradigm[1] to the space of quantum channels.

SUPPLEMENTARY MATERIAL

See supplementary material for Mathematica notebook.

ACKNOWLEDGMENTS

I.N. would like to thank Anna Jenčová and David Reeb for very insightful discussion regarding the diamond norm, which led to several improvements and simplifications of the proof of Proposition 1. I.N.'s research has been supported by the ANR Project Nos. RMTQIT ANR-12-IS01-0001-01 and StoQ ANR-14-CE25-0003-01, as well as by a von Humboldt fellowship. Financial support by the Polish National Science Centre under Project No. 2016/22/E/ST6/00062 (Z.P.), 2015/17/B/ST6/01872 (Ł.P.), and 2011/02/A/ST1/00119 (K.Ż.) is gratefully acknowledged.

APPENDIX A: ON THE PARTIAL TRACES OF UNITARILY INVARIANT RANDOM MATRICES

In this section, we show a general result about unitarily invariant random matrices: under some technical convergence assumptions, the partial trace of a unitarily invariant random matrix is "flat," i.e., it is close in norm to its average.

Recall that the normalized trace functional can be extended to arbitrary permutations as follows: for a matrix $X \in M_d(\mathbb{C})$, write

$$\mathrm{tr}_\pi(X) := \prod_{c \in \pi} \frac{1}{d} \mathrm{Tr}(X^{|c|}).$$

Recall the following definition from Ref. 22, Sec. 4.3.

Definition 4. A sequence of random matrices $X_d \in M_d(\mathbb{C})$ is said to have almost surely limit distribution *μ if,*

$$\forall p \geq 1, \quad a.s. - \lim_{d\to\infty} \mathrm{tr}(X^p) = \int x^p \, d\mu(x).$$

Proposition 23. Consider a sequence of hermitian random matrices $A_d \in M_{d_1(d)}(\mathbb{C}) \otimes M_{d_2(d)}(\mathbb{C})$ and assume the following:

1. *Both functions $d_{1,2}(d)$ grow to infinity, in such a way that $d_1/d_2^2 \to 0$.*
2. *The matrices A_d are unitarily invariant.*
3. *The family (A_d) has almost surely limit distribution μ, for some compactly supported probability measure μ.*

   Then, the normalized partial traces $B_d := d_2^{-1}[\mathrm{id} \otimes \mathrm{Tr}](A_d)$ converge almost surely to multiple of the identity matrix

$$a.s. - \lim_{d\to\infty} \|B_d - aI_{d_1(d)}\| = 0,$$

where a is the average of μ,

$$a := \int x d\mu(x).$$

Proof. In the proof, we shall drop the parameter $d \to \infty$, but the reader should remember that the matrix dimensions $d_{1,2}$ are functions of d and that all the matrices appearing are indexed by d. To conclude, it is enough to show that

$$\mathbb{P} - \lim_{d_1 \to \infty} \lambda_{\max}(B_d) = a$$

since the statement for the smallest eigenvalue follows in a similar manner. Let us denote by

$$b := \frac{1}{d_1} \sum_{i=1}^{d_1} \lambda_i(B) = \mathrm{tr}_{(1)}(B), \tag{A1}$$

$$v := \frac{1}{d_1} \sum_{i=1}^{d_1} (\lambda_i(B) - b)^2 = \frac{1}{d_1} \sum_{i=1}^{d_1} \lambda_i(B)^2 - \left[\frac{1}{d_1} \sum_{i=1}^{d_1} \lambda_i(B) \right]^2 = \mathrm{tr}_{(12)}(B) - [\mathrm{tr}_{(1)}(B)]^2 \tag{A2}$$

the average eigenvalue and, respectively, the variance of the eigenvalues of B; these are real random variables (actually, sequences of random variables indexed by d). By Chebyshev's inequality, we have a bound

$$\lambda_{\max}(B) \le b + \sqrt{v}\sqrt{d_1}. \tag{A3}$$

Note that one could replace the $\sqrt{d_1}$ factor in the inequality above by $\sqrt{d_1 - 1}$ by using Samuelson's inequality,[39,46] but the weaker version is enough for us.

We shall prove now that $b \to a$ almost surely and later that $d_1 v \to 0$ almost surely, which is what we need to conclude. To do so, we shall use the Weingarten formula.[16,47] In the graphical formalism for the Weingarten calculus introduced in Ref. 13, the expectation value of an expression involving a random Haar unitary matrix can be computed as a sum over diagrams indexed by permutation matrices; we refer the reader to Ref. 13 or Ref. 14 for details.

Using the unitary invariance of A, we write $A = U\,\mathrm{diag}(\lambda)U^*$, for a Haar-distributed random unitary matrix $U \in \mathcal{U}(d_1 d_2)$ and some (random) eigenvalue vector λ. Note that traces of powers of A depend only on λ, so we shall write $\mathrm{tr}_\pi(\lambda) := \mathrm{tr}_\pi(A)$. We apply the Weingarten formula to a general moment of B, given by a permutation π,

$$\mathbb{E}_U \mathrm{tr}_\pi(B) = \prod_{i=1}^{\#\pi} \frac{1}{d_1} \mathbb{E}_U \operatorname{Tr} B^{|c_i|},$$

where $c_1, \ldots, c_{\#\pi}$ are the cycles of $\pi \in \mathcal{S}_p$ and $\mathbb{E}_U$ denotes the conditional expectation with respect to the Haar random unitary matrix U. From the graphical representation of the Weingarten formula (Ref. 13, Theorem 4.1), we can compute the conditional expectation over U (note that below, the vector of eigenvalues λ is still random),

$$\mathbb{E}_U \mathrm{tr}_\pi(B) = d_1^{-\#\pi} d_2^{-p} \sum_{\alpha,\beta \in \mathcal{S}_p} d_1^{\#(\pi^{-1}\alpha)} d_2^{\#\alpha} (d_1 d_2)^{\#\beta} \mathrm{tr}_\beta(\lambda)\, \mathrm{Wg}_{d_1 d_2}(\alpha^{-1}\beta). \tag{A4}$$

Above, Wg is the Weingarten function[16] and $\mathrm{tr}_\beta(\lambda)$ is the moment of the diagonal matrix $\mathrm{diag}(\lambda)$ corresponding to the permutation β. The combinatorial factors $d_1^{\#(\pi^{-1}\alpha)}$ and $d_2^{\#\alpha}$ come from the initial wirings of the boxes respective to the vector spaces of dimensions d_1 (initial wiring given by π) and d_2 (initial wiring given by the identity permutation); see Fig. 8. The pre-factors $d_1^{-\#\pi} d_2^{-p}$ contain the normalization from the (partial) traces. Finally, the (random) factors $\mathrm{tr}_\beta(\lambda)$ are the normalized power

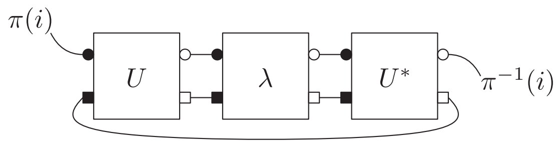

FIG. 8. The ith group in the diagram corresponding to $m_\pi(B)$.

sums of λ,

$$\mathrm{tr}_\beta(\lambda) = \prod_{i=1}^{\#\beta} (d_1 d_2)^{-1} \sum_{j=1}^{d_1 d_2} \lambda_j^{|w_i|},$$

where $w_1, \ldots, w_{\#\beta}$ are the cycles of β. Recall that we have assumed almost sure convergence for the sequence (A_d) [and, thus, for (λ_d)],

$$\forall \pi, \quad \text{a.s.} - \lim_{d\to\infty} \mathrm{tr}_\beta(\lambda) = \prod_{i=1}^{\#\beta} \int x^{|w_i|} d\mu(x) =: m_\pi(\mu). \tag{A5}$$

As a first application of the Weingarten formula (A4), let us find the distribution of the random variable $b = \mathrm{Tr}(B)/d_1$. Obviously,

$$\mathbb{E}_U b = \mathbb{E}_U \mathrm{tr}_{(1)}(B) = d_1^{-1} d_2^{-1} d_1^2 d_2^2 \mathrm{tr}_{(1)}(\lambda) \frac{1}{d_1 d_2} = \mathrm{tr}_{(1)}(\lambda). \tag{A6}$$

Actually, b does not depend on the random unitary matrix U since

$$b = \frac{1}{d_1} \mathrm{Tr}(B) = \frac{1}{d_1 d_2} \mathrm{Tr}(A) = \frac{1}{d_1 d_2} \mathrm{Tr}(U \,\mathrm{diag}(\lambda) U^*) = \frac{1}{d_1 d_2} \sum_{i=1}^{d_1 d_2} \lambda_i = \mathrm{tr}_{(1)}(\lambda).$$

From the hypothesis (A5) [with $\pi = (1)$], we have that, almost surely as $d \to \infty$, the random variable b converges to the scalar $a = m_{(1)}(\mu)$.

Let us now move on to the variance v of the eigenvalues. First, we compute its expectation $\mathbb{E}_U v = \mathrm{tr}_{(12)}(B) - \mathrm{tr}_{(1)(2)}(B)$. We apply now the Weingarten formula (A4) for $\mathbb{E}_U \mathrm{tr}_{(12)}(B)$; the sum has $2!^2 = 4$ terms, which we compute below:

- $\alpha = \beta = (1)(2)$: $T_1 = d_1^2 d_2^2 \mathrm{tr}_{(1)(2)}(\lambda) \frac{1}{d_1^2 d_2^2 - 1}$.
- $\alpha = (1)(2)$, $\beta = (12)$: $T_2 = -\mathrm{tr}_{(12)}(\lambda) \frac{1}{d_1^2 d_2^2 - 1}$.
- $\alpha = (12)$, $\beta = (1)(2)$: $T_3 = -d_1^2 \mathrm{tr}_{(12)}(\lambda) \frac{1}{d_1^2 d_2^2 - 1}$.
- $\alpha = \beta = (12)$: $T_4 = d_1^2 \mathrm{tr}_{(12)}(\lambda) \frac{1}{d_1^2 d_2^2 - 1}$.

Combining the expressions above with (A6), we get

$$\mathbb{E}_U v = \frac{d_1^2 - 1}{d_1^2 d_2^2 - 1} (\mathrm{tr}_{(12)}(\lambda) - \mathrm{tr}_{(1)(2)}(\lambda)).$$

Using the hypothesis (A5), we have thus, as $d_{1,2} \to \infty$,

$$\mathbb{E} v = (1 + o(1)) d_2^{-2} (m_{(12)}(\mu) - m_{(1)(2)}(\mu)).$$

Let us now proceed and estimate the variance of v; more precisely, let us compute $\mathbb{E}(v^2)$. As before, we shall compute the expectation in two steps: first with respect to the random Haar unitary matrix U and then using our assumption (A5), with respect to λ, in the asymptotic limit. To perform the unitary integration, note that the Weingarten sum is indexed by a couple $(\alpha, \beta) \in \mathcal{S}_4^2$, so it contains $4!^2 = 576$ terms; see the supplementary material. In Appendix B, we have computed the variance of v with the usage of symmetry arguments. The result, to the first order, reads as

$$\mathrm{Var}(v) = \mathbb{E}_U(v^2) - (\mathbb{E}_U v)^2 = (1 + o(1))2d_1^{-2}d_2^{-4}[m_{(12)}(\lambda) - m_{(1)(2)}(\lambda)]^2.$$

Taking the expectation over λ and the limit (we are allowed to, by dominated convergence), we get

$$\mathrm{Var}(v) = (1 + o(1))2d_1^{-2}d_2^{-4}[m_{(12)}(\mu) - m_{(1)(2)}(\mu)]^2.$$

We put now all the ingredients together,

$$\mathbb{P}(\sqrt{d_1}\sqrt{v} \geq \varepsilon) = \mathbb{P}(v \geq \varepsilon^2 d_1^{-1}) \leq \frac{\mathrm{Var}(v)}{[\varepsilon^2 d_1^{-1} - v\mathbb{E}]^2} \sim \frac{Cd_1^{-2}d_2^{-4}}{[\varepsilon^2 d_1^{-1} - (1 + o(1))C' d_2^{-2}]^2},$$

where C, C' are non-negative constants depending on the limiting measure μ. Using $d_1 \ll d_2^2$, the dominating term in the denominator above is $\varepsilon^2 d_1^{-1}$, and thus we have,

$$\mathbb{P}(\sqrt{d_1}\sqrt{v} \geq \varepsilon) \lesssim C\varepsilon^{-4}d_2^{-4}.$$

Since the series $\sum d_2^{-4}$ is summable, we obtain the announced almost sure convergence by the Borel-Cantelli lemma, finishing the proof. ∎

APPENDIX B: CALCULATION OF THE VARIANCE Var(v)

In this appendix, we compute the centered second moment of the variable v defined in (A2) necessary to show almost sure convergence $d_1 v \to 0$. We remind here that $A \in M_{d_1}(\mathbb{C}) \otimes M_{d_2}(\mathbb{C})$ and $B = d_2^{-1}[\mathrm{id} \otimes \mathrm{Tr}](A)$. Because we assume that A has unitarily invariant distribution, we can write

$$A = U\mathrm{diag}(\lambda)U^\dagger = \sum_{i=0}^{d_1 d_2 - 1} \lambda_i |U_i\rangle\langle U_i|, \tag{B1}$$

where $|U_i\rangle = U|i\rangle$ is the ith column of matrix U, and

$$B = d_2^{-1}\mathrm{Tr}_2 A = d_2^{-1} \sum_{i=0}^{d_1 d_2 - 1} \lambda_i \mathrm{Tr}_2 |U_i\rangle\langle U_i|. \tag{B2}$$

We denote $\rho_i = \mathrm{Tr}_2|U_i\rangle\langle U_i|$ and consider mixed moments computed in Lemma 25,

$$\mathcal{M}(i,j,k,l) = \mathbb{E}_U \mathrm{Tr}(\rho_i\rho_j)\mathrm{Tr}(\rho_k\rho_l), \tag{B3}$$

where $\mathbb{E}_U$ denotes the conditional expectation with respect to the Haar random unitary matrix U. We also define symmetric mixed moments

$$\mathcal{SM}(i,j,k,l) = \mathbb{E}_U \mathrm{Tr}(\rho_i\rho_j)\mathrm{Tr}(\rho_k\rho_l) - \mathbb{E}_U \mathrm{Tr}(\rho_i\rho_j)\, \mathbb{E}_U \mathrm{Tr}(\rho_k\rho_l). \tag{B4}$$

Proposition 24. Let $v = \frac{1}{d_1}\mathrm{Tr}B^2 - \left(\frac{1}{d_1}\mathrm{Tr}B\right)^2$. *Denoting* $\mathrm{Var}_U(v) = \mathbb{E}_U v^2 - (\mathbb{E}_U v)^2$, *we have*

$$\mathrm{Var}_U(v) = \frac{2(\mu_1^2 - \mu_2)^2}{d_1^2 d_2^4}(1 + o(1)) \tag{B5}$$

as $d_1, d_2 \to \infty$, *in the above* $\mu_k = \frac{1}{d_1 d_2}\sum \lambda_i^k$.

Direct computations with the usage of symmetric moments $\mathcal{SM}$ give us

$$\begin{aligned}
\mathrm{Var}_U(v) = &\frac{1}{d_1^2 d_2^4}\Big[(d_1d_2)^4\mu_1^4\mathcal{SM}(0,1,2,3)\\
&+2(d_1d_2)^3\mu_2\mu_1^2\big(\mathcal{SM}(0,0,1,2)+2\mathcal{SM}(0,1,0,2)-3\mathcal{SM}(0,1,2,3)\big)\\
&+4(d_1d_2)^2\mu_3\mu_1\big(\mathcal{SM}(0,0,0,1)-\mathcal{SM}(0,0,1,2)-2\mathcal{SM}(0,1,0,2)+2\mathcal{SM}(0,1,2,3)\big)\\
&+(d_1d_2)^2\mu_2^2\big(\mathcal{SM}(0,0,1,1)-2\mathcal{SM}(0,0,1,2)+2\mathcal{SM}(0,1,0,1)-4\mathcal{SM}(0,1,0,2)+3\mathcal{SM}(0,1,2,3)\big)\\
&+d_1d_2\mu_4\big(\mathcal{SM}(0,0,0,0)-4\mathcal{SM}(0,0,0,1)-\mathcal{SM}(0,0,1,1)+4\mathcal{SM}(0,0,1,2)-2\mathcal{SM}(0,1,0,1)\\
&\qquad+8\mathcal{SM}(0,1,0,2)-6\mathcal{SM}(0,1,2,3)\big)\Big]\\
=&\frac{2\left(d_1^2-1\right)\left(d_2^2-1\right)}{d_2^2\left(d_1^2d_2^2-1\right)^2\left(d_1^4d_2^4-13d_1^2d_2^2+36\right)}\Big(d_1^4d_2^4\left(\mu_1^2-\mu_2\right)^2\\
&+d_1^2d_2^2\left(11\mu_1^4-22\mu_2\mu_1^2+20\mu_3\mu_1-4\mu_2^2-5\mu_4\right)+5\left(3\mu_2^2-4\mu_1\mu_3+\mu_4\right)\Big).
\end{aligned} \tag{B6}$$

The above formula gives the exact result for $\mathrm{Var}_U(v)$. Considering the limiting behaviour $d_1, d_2 \to \infty$ we get

$$\mathrm{Var}_U(v) = \frac{2(\mu_1^2-\mu_2)^2}{d_1^2d_2^4}(1+o(1)), \tag{B7}$$

which completes the proof of Proposition 24.

The moments computed here are used in Eq. (B6) to obtain the variance $\mathrm{Var}_U(v)$.

Lemma 25. We have the following formulas for mixed moments defined in Eq. (B3). Note that because of symmetry we cover all possible cases,

$$\begin{aligned}
\mathcal{M}(0,0,0,0) &= \frac{d_2d_1^3+2\left(d_2^2+2\right)d_1^2+d_2\left(d_2^2+10\right)d_1+4d_2^2+2}{(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)},\\
\mathcal{M}(0,0,0,1) &= \frac{\left(d_2^2-1\right)(d_1(d_1+d_2)(d_1d_2+4)+2)}{(d_1d_2-1)(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)},\\
\mathcal{M}(0,0,1,1) &= \frac{(d_1d_2(d_1d_2+2)-4)(d_1+d_2)^2+4}{d_1d_2(d_1d_2-1)(d_1d_2+2)(d_1d_2+3)},\\
\mathcal{M}(0,0,1,2) &= \frac{\left(d_2^2-1\right)(d_1(d_1+d_2)(d_1d_2(d_1d_2+4)+2)-2)}{d_1d_2(d_1d_2-1)(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)},\\
\mathcal{M}(0,1,0,1) &= \frac{\left(d_2^2-1\right)\left(d_1\left(6d_2+d_1\left(d_2^2(d_1d_2+5)-2\right)\right)+2\right)}{d_1d_2(d_1d_2-1)(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)},\\
\mathcal{M}(0,1,0,2) &= \frac{\left(d_2^2-1\right)\left(d_1\left(d_1\left(d_2\left(d_1\left(3d_2^2+d_1\left(d_2^2-1\right)d_2-4\right)-3d_2\right)+2\right)-8d_2\right)-2\right)}{d_1d_2(d_1d_2-2)(d_1d_2-1)(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)},\\
\mathcal{M}(0,1,2,3) &= \frac{\left(d_2^2-1\right)\left(d_2^2\left(d_2^2-1\right)d_1^4+2\left(7-6d_2^2\right)d_1^2+22\right)}{d_1^2d_2^2\left(d_1^2d_2^2-7\right)^2-36}.
\end{aligned} \tag{B8}$$

The rest of this section is devoted to the proof of the above lemma. We will omit the subscript U in the expectation because matrices ρ_i depend only on the Haar unitary matrix U. Following the result of Giraud,[20] we find the second moment of the purity

$$\mathcal{M}(0,0,0,0) = \mathbb{E}(\mathrm{Tr}\rho_0^2)^2 = \frac{d_2 d_1^3 + 2\left(d_2^2+2\right)d_1^2 + d_2\left(d_2^2+10\right)d_1 + 4d_2^2 + 2}{(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)}, \tag{B9}$$

where $\rho_0 = \mathrm{Tr}_2|U_0\rangle\langle U_0|$.

Next we consider the moments $\mathcal{M}(0,0,0,1)$ defined in (B3),

$$\mathcal{M}(0,0,0,1) = \mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_0\rho_1) = \frac{1}{d_1d_2-1}\mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_0(d_2 I_{d_1} - \rho_0)). \tag{B10}$$

The above follows from the fact that we have invariance with respect to the permutation of columns of U, and therefore $\mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_0\rho_1) = \mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_0\rho_2)$. Next we note that $\sum_{i=0}^{d_1d_2-1}\rho_i = d_2 I_{d_1}$. Using the above equation, we obtain

$$\begin{aligned}
\mathcal{M}(0,0,0,1) &= \frac{1}{d_1d_2-1}\left(d_2\mathbb{E}\mathrm{Tr}\rho_0^2 - \mathbb{E}(\mathrm{Tr}\rho_0^2)^2\right)\\
&= \frac{1}{d_1d_2-1}\left(d_2\frac{d_1+d_2}{d_1d_2+1} - \mathcal{M}(0,0,0,0)\right)\\
&= \frac{\left(d_2^2-1\right)\left(d_1\left(d_1+d_2\right)\left(d_1d_2+4\right)+2\right)}{(d_1d_2-1)(d_1d_2+1)(d_1d_2+2)(d_1d_2+3)}.
\end{aligned} \tag{B11}$$

In order to get other mixed moments, we need to perform another integration. We start with expectations of the following kind:

$$\mathcal{M}(0,0,1,1) = \mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}\rho_1^2 = \mathbb{E}\mathrm{Tr}(\mathrm{Tr}_2|U_0\rangle\langle U_0|)^2\mathrm{Tr}(\mathrm{Tr}_2|U_1\rangle\langle U_1|)^2. \tag{B12}$$

Note that if we multiply matrix U by a unitary matrix which does not change the first column, we will not change the expectation value. In fact we can integrate over the subgroup of matrices which does not change the first column of U. Now for a moment we fix the matrix U and consider the expectation value

$$\mathrm{Tr}(\mathrm{Tr}_2|U_0\rangle\langle U_0|)^2\mathbb{E}_V[\mathrm{Tr}(\mathrm{Tr}_2 UV|1\rangle\langle 1|V^\dagger U^\dagger)^2], \tag{B13}$$

where matrices V are in the form

$$V = \begin{pmatrix} 1 & 0 & 0 & \dots & 0\\ 0 & v_{1,1} & v_{1,2} & \cdots & v_{1,d_1d_2-1}\\ 0 & v_{2,1} & v_{1,2} & \cdots & v_{2,d_1d_2-1}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & v_{d_1d_2-1,1} & v_{d_1d_2-1,2} & \cdots & v_{d_1d_2-1,d_1d_2-1}\end{pmatrix}. \tag{B14}$$

The $\mathbb{E}_V$ is an expectation with respect to the Haar measure on $U(d_1d_2-1)$ embedded in $U(d_1d_2)$, in the above way. Note that the vector $UV|1\rangle$ represents a random orthogonal vector to the $|U_0\rangle = U|0\rangle$.

First we calculate

$$\begin{aligned}
&\mathbb{E}_V(UV|1\rangle\langle 1|V^\dagger U^\dagger)\otimes(UV|1\rangle\langle 1|V^\dagger U^\dagger)\\
&= \mathbb{E}_V(U|V_1\rangle\langle V_1|U^\dagger)\otimes(U|V_1\rangle\langle V_1|U^\dagger)\\
&= (U\otimes U)\mathbb{E}_V|V_1\rangle\langle V_1|\otimes|V_1\rangle\langle V_1|(U\otimes U)^\dagger.
\end{aligned} \tag{B15}$$

Now, using standard integrals, we obtain

$$\mathbb{E}_V|V_1\rangle\langle V_1|\otimes|V_1\rangle\langle V_1| = \frac{1}{(d_1d_2-1)d_1d_2}\sum_{i_1i_2j_1j_2}\left(\delta_{i_1j_1}\delta_{i_1j_1} + \delta_{i_1j_2}\delta_{i_2j_1}\right)\theta(i_1i_2j_1j_2)|i_1i_2\rangle\langle j_1j_2|, \tag{B16}$$

where $\theta(x) = (1-\delta_{x,0})$ and incorporates the condition that first element of vector $|V_1\rangle$ is zero. Now we obtain, after elementary calculations, using the fact that U is unitary

$$
\begin{aligned}
&(U\otimes U)\mathbb{E}_V|V_1\rangle\langle V_1|\otimes|V_1\rangle\langle V_1|(U\otimes U)^\dagger\\
&=\frac{1}{(d_1d_2-1)d_1d_2}\sum_{i_1i_2j_1j_2}\Big((\delta_{i_1j_1}-u_{i_1,0}u_{j_1,0})(\delta_{i_2j_2}-u_{i_2,0}u_{j_2,0})\\
&\qquad+(\delta_{i_1j_2}-u_{i_1,0}u_{j_2,0})(\delta_{i_2j_1}-u_{i_2,0}u_{j_1,0})\Big)\theta(i_1i_2j_1j_2)|i_1i_2\rangle\langle j_1j_2| \\
&=\frac{1}{(d_1d_2-1)d_1d_2}\Big(I_{d_1d_2d_1d_2}+|U_0\rangle\langle U_0|\otimes|U_0\rangle\langle U_0|-I_{d_1d_2}\otimes|U_0\rangle\langle U_0|-|U_0\rangle\langle U_0|\otimes I_{d_1d_2}\\
&\quad+S_{d_1d_2}(I_{d_1d_2d_1d_2}+|U_0\rangle\langle U_0|\otimes|U_0\rangle\langle U_0|-I_{d_1d_2}\otimes|U_0\rangle\langle U_0|-|U_0\rangle\langle U_0|\otimes I_{d_1d_2})\Big),
\end{aligned}
\tag{B17}
$$

where S_N is a swap operation on two systems of dimensions N each, i.e., $S=\sum_{i_1,i_2=0}^{N-1}|i_1i_2\rangle\langle i_2i_1|$. So we get

$$
\begin{aligned}
&(U\otimes U)\mathbb{E}_V|V_1\rangle\langle V_1|\otimes|V_1\rangle\langle V_1|(U\otimes U)^\dagger\\
&=\frac{1}{(d_1d_2-1)d_1d_2}(I_{d_1d_2d_1d_2}+S_{d_1d_2})\Big(I_{d_1d_2}+|U_0\rangle\langle U_0|\otimes|U_0\rangle\langle U_0|-I_{d_1d_2}\otimes|U_0\rangle\\
&\quad\times\langle U_0|-|U_0\rangle\langle U_0|\otimes I_{d_1d_2}\Big).
\end{aligned}
\tag{B18}
$$

We are going to use several times the following identity often used in quantum information. For two square matrices ρ_1, ρ_2 of size N,

$$\mathrm{Tr}\rho_1\rho_2=\mathrm{Tr}S_N(\rho_1\otimes\rho_2). \tag{B19}$$

This identity allows us to obtain

$$
\begin{aligned}
&\mathbb{E}_V\mathrm{Tr}(\mathrm{Tr}_2U|V_1\rangle\langle V_1|U^\dagger)^2=\mathbb{E}\mathrm{Tr}S_{d_1}(\mathrm{Tr}_2U|V_1\rangle\langle V_1|U^\dagger)\otimes(\mathrm{Tr}_2U|V_1\rangle\langle V_1|U^\dagger)\\
&=\mathbb{E}_V\mathrm{Tr}S_{d_1}\mathrm{Tr}_{2,4}(U|V_1\rangle\langle V_1|U^\dagger\otimes U|V_1\rangle\langle V_1|U^\dagger)\\
&=\mathrm{Tr}S_{d_1}\mathrm{Tr}_{2,4}(U\otimes U)\mathbb{E}_V|V_1\rangle\langle V_1|\otimes|V_1\rangle\langle V_1|(U\otimes U)^\dagger\\
&=\frac{1}{(d_1d_2-1)d_1d_2}\mathrm{Tr}S_{d_1}\mathrm{Tr}_{2,4}(I_{d_1d_2d_1d_2}+S_{d_1d_2})\Big(I_{d_1d_2d_1d_2}+|U_0\rangle\langle U_0|\otimes|U_0\rangle\langle U_0|\\
&\quad-I_{d_1d_2}\otimes|U_0\rangle\langle U_0|-|U_0\rangle\langle U_0|\otimes I_{d_1d_2}\Big).
\end{aligned}
\tag{B20}
$$

After performing partial trace over subsystems 2 and 4, we get

$$
\begin{aligned}
&\mathbb{E}_V\mathrm{Tr}(\mathrm{Tr}_2U|V_1\rangle\langle V_1|U^\dagger)^2\\
&=\frac{1}{(d_1d_2-1)d_1d_2}\Big(d_1d_2^2+\mathrm{Tr}\rho_0^2-d_2\mathrm{Tr}\rho_0-d_2\mathrm{Tr}\rho_0+d_1^2d_2+\mathrm{Tr}(\rho_0')^2-d_1\mathrm{Tr}\rho_0'-d_1\mathrm{Tr}\rho_0'\Big)\\
&=\frac{1}{(d_1d_2-1)d_1d_2}\Big(d_1d_2^2+d_1^2d_2-2d_1-2d_2+2\mathrm{Tr}\rho_0^2\Big).
\end{aligned}
\tag{B21}
$$

In the above formulas, we used $\rho_0'=\mathrm{Tr}_1|U_0\rangle\langle U_0|$ and the fact that two partial traces of a pure bi-partite state have the same purity $\mathrm{Tr}\rho_0^2=\mathrm{Tr}(\rho_0')^2$. Using the above we find the desired expectation

$$
\begin{aligned}
\mathcal{M}(0,0,1,1)&=\mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}\rho_1^2=\mathbb{E}\mathrm{Tr}\rho_0^2\mathbb{E}_V\mathrm{Tr}(\mathrm{Tr}_2U|V_1\rangle\langle V_1|U^\dagger)^2\\
&=\frac{1}{(d_1d_2-1)d_1d_2}\Big((d_1d_2^2+d_1^2d_2-2d_1-2d_2)\mathbb{E}\mathrm{Tr}\rho_0^2+2\mathbb{E}(\mathrm{Tr}\rho_0^2)^2\Big)\\
&=\frac{1}{(d_1d_2-1)d_1d_2}\Big((d_1d_2^2+d_1^2d_2-2d_1-2d_2)\frac{d_1+d_2}{d_1d_2+1}+2\mathcal{M}(0,0,0,0)\Big)\\
&=\frac{(d_1d_2\,(d_1d_2+2)-4)\,(d_1+d_2)^2+4}{d_1d_2\,(d_1d_2-1)\,(d_1d_2+2)\,(d_1d_2+3)}.
\end{aligned}
\tag{B22}
$$

Using the inner integral, we can also calculate the other mixed moments

$$\begin{aligned}
&\mathcal{M}(0,1,0,1) = \mathbb{E}\mathrm{Tr}\rho_0\rho_1\mathrm{Tr}\rho_0\rho_1 = \mathbb{E}\mathrm{Tr}(\rho_0\otimes\rho_0)(\rho_1\otimes\rho_1)\\
&= \mathbb{E}\mathrm{Tr}(\rho_0\otimes\rho_0)\mathrm{Tr}_{2,4}(U\otimes U)\mathbb{E}_V|V_1\rangle\langle V_1|\otimes|V_1\rangle\langle V_1|(U\otimes U)^\dagger\\
&= \frac{1}{(d_1d_2-1)d_1d_2}\mathbb{E}\mathrm{Tr}(\rho_0\otimes\rho_0)\mathrm{Tr}_{2,4}(I_{d_1d_2d_1d_2}+S_{d_1d_2,d_1d_2})\\
&\quad\times\left(I_{d_1d_2d_1d_2}+|U_0\rangle\langle U_0|\otimes|U_0\rangle\langle U_0|-I_{d_1d_2}\otimes|U_0\rangle\langle U_0|-|U_0\rangle\langle U_0|\otimes I_{d_1d_2}\right)\\
&= \frac{1}{(d_1d_2-1)d_1d_2}\mathbb{E}\Big(\mathrm{Tr}(\rho_0\otimes\rho_0)(d_2^2I_{d_1^2}+\rho_0\otimes\rho_0-d_2I_{d_1}\otimes\rho_0-d_2\rho_0\otimes I_{d_2})\\
&+\mathrm{Tr}S_{d_1d_2,d_1d_2}\left(I_{d_1d_2d_1d_2}+|U_0\rangle\langle U_0|\otimes|U_0\rangle\langle U_0|-I_{d_1d_2}\otimes|U_0\rangle\langle U_0|-|U_0\rangle\langle U_0|\otimes I_{d_1d_2}\right)\\
&\quad\times(\rho_0\otimes I_{d_2}\otimes\rho_0\otimes I_{d_2})\Big) \qquad\qquad (B23)\\
&= \frac{1}{(d_1d_2-1)d_1d_2}\mathbb{E}\Big(d_2^2+(\mathrm{Tr}\rho_0^2)^2-2d_2\mathrm{Tr}\rho_0\mathrm{Tr}\rho_0^2\\
&+\mathrm{Tr}(\rho_0^2\otimes I_{d_2})+\mathrm{Tr}(|U_0\rangle\langle U_0|\rho_0\otimes I_{d_2})^2-2\mathrm{Tr}(\rho_0\otimes I_{d_2})(|U_0\rangle\langle U_0|\rho_0\otimes I_{d_2})\Big)\\
&= \frac{1}{(d_1d_2-1)d_1d_2}\mathbb{E}\Big(d_2^2-2d_2\mathrm{Tr}\rho_0^2+(\mathrm{Tr}\rho_0^2)^2+d_2\mathrm{Tr}\rho_0^2+(\mathrm{Tr}\rho_0^2)^2-2\mathrm{Tr}\rho_0^3\Big)\\
&= \frac{1}{(d_1d_2-1)d_1d_2}\left(d_2^2-d_2\frac{d_1+d_2}{d_1d_2+1}+2\mathcal{M}(0,0,0,0)-2\mathbb{E}\mathrm{Tr}\rho_0^3\right)\\
&= \frac{\left(d_2^2-1\right)\left(d_1\left(6d_2+d_1\left(d_2^2\left(d_1d_2+5\right)-2\right)\right)+2\right)}{d_1d_2\left(d_1d_2-1\right)\left(d_1d_2+1\right)\left(d_1d_2+2\right)\left(d_1d_2+3\right)}.
\end{aligned}$$

This is because $\mathbb{E}\mathrm{Tr}\rho_0^3 = \frac{(d_1+d_2)^2+d_1d_2+1}{(d_1d_2+1)(d_1d_2+2)}$; see Ref. 40.

Using the above results, we obtain other moments

$$\begin{aligned}
\mathcal{M}(0,0,1,2) &= \mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_1\rho_2) = \frac{1}{d_1d_2-2}\mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_1(d_2I_{d_1}-\rho_0-\rho_1))\\
&= \frac{1}{d_1d_2-2}\left(d_2\mathbb{E}\mathrm{Tr}\rho_0^2-\mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}(\rho_1\rho_0)-\mathbb{E}\mathrm{Tr}\rho_0^2\mathrm{Tr}\rho_1^2\right)\\
&= \frac{1}{d_1d_2-2}\left(d_2\frac{d_1+d_2}{d_1d_2+1}-\mathcal{M}(0,0,0,1)-\mathcal{M}(0,0,1,1)\right)\\
&= \frac{\left(d_2^2-1\right)\left(d_1\left(d_1+d_2\right)\left(d_1d_2\left(d_1d_2+4\right)+2\right)-2\right)}{d_1d_2\left(d_1d_2-1\right)\left(d_1d_2+1\right)\left(d_1d_2+2\right)\left(d_1d_2+3\right)}.
\end{aligned} \tag{B24}$$

Next we consider the mixed moment of type (0, 1, 0, 2),

$$\begin{aligned}
\mathcal{M}(0,1,0,2) &= \mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\mathrm{Tr}(\rho_0\rho_2) = \frac{1}{d_1d_2-2}\mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\mathrm{Tr}(\rho_0(d_2I_{d_1}-\rho_0-\rho_1))\\
&= \frac{1}{d_1d_2-2}\left(d_2\mathbb{E}\mathrm{Tr}(\rho_0\rho_1)-\mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\mathrm{Tr}(\rho_0^2)-\mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\mathrm{Tr}(\rho_0\rho_1)\right)\\
&= \frac{\left(d_2^2-1\right)\left(d_1\left(d_1\left(d_2\left(d_1\left(3d_2^2+d_1\left(d_2^2-1\right)d_2-4\right)-3d_2\right)+2\right)-8d_2\right)-2\right)}{d_1d_2\left(d_1d_2-2\right)\left(d_1d_2-1\right)\left(d_1d_2+1\right)\left(d_1d_2+2\right)\left(d_1d_2+3\right)}.
\end{aligned} \tag{B25}$$

Consider now the last case of all different indices

$$\begin{aligned}\mathcal{M}(0,1,2,3) &= \mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\mathrm{Tr}(\rho_2\rho_3)\\ &= \frac{1}{d_1d_2-3}\mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\mathrm{Tr}\left(\,\rho_2(d_2I_{d_1}-\rho_0-\rho_1-\rho_2\right)\\ &= \frac{1}{d_1d_2-3}\mathbb{E}\mathrm{Tr}(\rho_0\rho_1)\left(d_2-\mathrm{Tr}\rho_2\rho_0-\mathrm{Tr}\rho_2\rho_1-\mathrm{Tr}\rho_2\rho_2\right)\\ &= \frac{\left(d_2^2-1\right)\left(d_2^2\left(d_2^2-1\right)d_1^4+2\left(7-6d_2^2\right)d_1^2+22\right)}{d_1^2d_2^2\left(d_1^2d_2^2-7\right)^2-36}.\end{aligned} \tag{B26}$$

In this way, we calculated all the moments defined in Eq. (B3). Symmetrizing them according Eq. (B4), they can be used in Eq. (B6) to establish Proposition 24.

[1] Aubrun, G. and Szarek, S., *Alice and Bob Meet Banach* (American Mathematical Society, 2017).
[2] Bai, Z. and Silverstein, J. W., *Spectral Analysis of Large Dimensional Random Matrices*, Springer Series in Statistics (Springer, 2010).
[3] Bai, Z. D. and Yin, Y. Q., "Limit of the smallest eigenvalue of large dimensional covariance matrix," Ann. Probab. **21**, 1275–1294 (1993).
[4] Ben Arous, G. and Bourgade, P., "Extreme gaps between eigenvalues of random matrices," Ann. Probab. **41**, 2648–2681 (2013).
[5] Bercovici, H. and Voiculescu, D., "Regularity questions for free convolution," Oper. Theory **104**, 37–47 (1998).
[6] Bhatia, R., *Matrix Analysis*, Graduate Texts in Mathematics Vol. 169 (Springer-Verlag, New York, 1997).
[7] Bhatia, R., "Matrix factorizations and their perturbations," Linear Algebra Appl. **197-198**, 245–276 (1994).
[8] Biane, P., "Free Brownian motion, free stochastic calculus and random matrices," in *Free Probability Theory*, Volume 12 of Fields Institute Communications (American Mathematical Society, Providence, Waterloo, ON, 1995, 1997), pp. 1–19.
[9] Biane, P., "Segal-Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems," J. Funct. Anal. **144**, 232–286 (1997).
[10] Bruzda, W., Cappellini, V., Sommers, H.-J., and Życzkowski, K., "Random quantum operations," Phys. Lett. A **373**, 320–324 (2009).
[11] Choi, M.-D., "Completely positive linear maps on complex matrices," Linear Algebra Appl. **10**, 285 (1975).
[12] Collins, B., Dahlqvist, A., and Kemp, T., "Strong convergence of unitary Brownian motion," preprint arXiv:1502.06186.
[13] Collins, B. and Nechita, I., "Random quantum channels. I: Graphical calculus and the Bell state phenomenon," Commun. Math. Phys. **297**, 345–370 (2010).
[14] Collins, B. and Nechita, I., "Random matrix techniques in quantum information theory," J. Math. Phys. **57**, 015215 (2016).
[15] Collins, B., Nechita, I., and Ye, D., "The absolute positive partial transpose property for random induced states," Random Matrices: Theory Appl. **01**, 1250002 (2012).
[16] Collins, B. and Śniady, P., "Integration with respect to the Haar measure on unitary, orthogonal and symplectic group," Commun. Math. Phys. **264**, 773–795 (2006).
[17] Davies, E. B., "Lipschitz continuity of operators in the Schatten classes," J. London Math. Soc. **s2-37**, 148–157 (1988).
[18] Deya, A. and Nourdin, I., "Convergence of Wigner integrals to the Tetilla law," Latin Am. J. Probab. Math. Stat. **9**, 101–127 (2012).
[19] Fuchs, C. A. and van De Graaf, J., "Cryptographic distinguishability measures for quantum-mechanical states," IEEE Trans. Inf. Theory **45**(4), 1216–1227 (1999).
[20] Giraud, O., "Distribution of bipartite entanglement for random pure states," J. Phys. A: Math. Theor. **40**, 2793 (2007).
[21] Helstrom, C. W., *Quantum Detection and Estimation Theory* (Academic Press, 1976).
[22] Hiai, F. and Petz, D., *The Semicircle Law, Free Random Variables, and Entropy* (AMS, 2000).
[23] Holevo, A. S., "An analogue of statistical decision theory and noncommutative probability theory," Tr. Mosk. Mat. O-va. **26**, 133–149 (1972).
[24] Horn, R. and Johnson, C., *Topics in Matrix Analysis* (Cambridge University Press, 1991).
[25] Kitaev, A. Y., "Quantum computations: Algorithms and error correction," Russ. Math. Surv. **52**, 1191–1249 (1997).
[26] Jamiołkowski, A., "Linear transformations which preserve trace and positive semidefiniteness of operators," Rep. Math. Phys. **3**, 275–278 (1972).
[27] Jenčová, A. and Plávala, M., "Conditions for optimal input states for discrimination of quantum channels," J. Math. Phys. **57**, 122203 (2016).
[28] Johnston, N., Kribs, D., and Paulsen, V., "Computing stabilized norms for quantum operations via the theory of completely bounded maps," Quantum Inf. Comput. **9**, 0016–0035 (2009).
[29] Johnston, N., QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9. http://www.qetlab.com, 12 January 2016.
[30] Kliesch, M., Kueng, R., Eisert, J., and Gross, D., "Improving compressed sensing with the diamond norm," IEEE Trans. Inf. Theory **62**, 7445–7463 (2016).
[31] Male, C., "The norm of polynomials in large random and deterministic matrices," Probab. Theory Relat. Fields **154**, 477–532 (2012).
[32] Michel, U., Kliesch, M., Kueng, R., and Gross, D., "Note on the saturation of the norm inequalities between diamond and nuclear norm," preprint arXiv:1612.07931.
[33] Nica, A. and Speicher, R., *Lectures on the Combinatorics of Free Probability* (Cambridge University Press, 2006).
[34] Nica, A. and Speicher, R., "Commutators of free random variables," Duke Math. J. **92**, 553–592 (1998).

[35] Paulsen, V., *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2002).
[36] Puchała, Z., Jenčová, A., Sedlák, M., and Ziman, M., "Exploring boundaries of quantum convex structures: Special role of unitary processes," Phys. Rev. A **92**(1), 012304 (2015).
[37] Puchała, Z., Pawela, Ł., and Życzkowski, K., "Distinguishability of generic quantum states," Phys. Rev. A **93**, 062112 (2016).
[38] Rains, E. M., "Combinatorial properties of Brownian motion on the compact classical groups," J. Theor. Probab. **10**, 659–679 (1997).
[39] Samuelson, P. A., "How deviant can you be?," J. Am. Stat. Assoc. **63**, 1522–1525 (1968).
[40] Sommers, H.-J. and Życzkowski, K., "Statistical properties of random density matrices," J. Phys. A: Math. Gen. **37**, 8457 (2004).
[41] Stinespring, W. F., "Positive functions on C*-algebras," Proc. Am. Math. Soc. **6**, 211–216 (1955).
[42] Szarek, S., Werner, E., and Życzkowski, K., "Geometry of sets of quantum maps: A generic positive map acting on a high-dimensional system is not completely positive," J. Math. Phys. **49**, 032113 (2008).
[43] Timoney, R. M., "Computing the norms of elementary operators," Ill. J. Math. **47**, 1207–1226 (2003).
[44] Watrous, J., *Theory of Quantum Information* (Cambridge University Press, in press), book draft available at https://cs.uwaterloo.ca/watrous/TQI/.
[45] Watrous, J., "Simpler semidefinite programs for completely bounded norms," Chicago J. Theor. Comput. Sci. **2013**, 8.
[46] Wolkowicz, H. and Styan, G. P. H., "Bounds for eigenvalues using traces," Linear Algebra Appl. **29**, 471–506 (1980).
[47] Weingarten, D., "Asymptotic behavior of group integrals in the limit of infinite rank," J. Math. Phys. **19**, 999–1001 (1978).