On the Probabilistic Quantum Error Correction

Ryszard Kukulski, Łukasz Pawela, and Zbigniew Puchała

***Abstract*— Probabilistic quantum error correction is an error-correcting procedure which uses postselection to determine if the encoded information was successfully restored. In this work, we analyze the probabilistic version of the error-correcting procedure for general noise. We generalize the Knill-Laflamme conditions for probabilistically correctable errors. We show that for some noise channels, the initial information has to be encoded into a mixed state to maximize the probability of successful error correction. Finally, the probabilistic error-correcting procedure offers an advantage over the deterministic procedure. Reducing the probability of successful error correction allows for correcting errors generated by a broader class of noise channels. Significantly, if the errors are caused by a unitary interaction with an auxiliary qubit system, we can probabilistically restore a qubit state by using only one additional physical qubit.**

***Index Terms*— Quantum error correction, postselection, channel-adapted error correction.**

I. Introduction

QUANTUM error correction (QEC) is an encoding-decoding procedure that protects quantum information from errors arising due to quantum noise. Similarly as in classical computations, this procedure is essential to develop fully operational quantum computers [1]. The theory of QEC, initialized by the work of Shor [2], covers a wide range of coding techniques: Calderbank-Shor-Steane codes [3], [4], [5], stabilizer codes [6], topological codes [7], subsystem codes [8], entanglement-assisted quantum error-correcting codes [9], [10], quantum low-density parity-check (LDPC) codes [11], quantum maximum distance separable codes [12] and many more (for a review see [13]).

In this work, we study a particular QEC procedure called *probabilistic quantum error correction* (pQEC) [14], [15], [16]. To outline how the pQEC procedure works, let us present an example of classical probabilistic error correction. Consider the scenario when the encoded data is harmed by a single bit error; that is with the probability $p \in [0,1]$ an arbitrary bit will be flipped. To secure a single bit of information, we use two physical bits. If we expect that $p \leq \frac{2}{3}$, then we can encode $0 \to 00$ and $1 \to 11$. If we receive information 00 at the decoding stage, we are certain the encoded message was 0 (and 1 for 11). We dismiss the cases 01 and 10 as they do not give conclusive answers. Otherwise, if $p > \frac{2}{3}$ it would be beneficial to use encoding $0 \to 00$ and $1 \to 01$ with the accepting states 10 and 11, respectively. It is worth mentioning that to secure one bit of information perfectly, it is necessary to use three physical bits, for example, $0 \to 000, 1 \to 111$.

Let us return to the quantum case. At the heart of the pQEC procedure lies the probabilistic decoding operation [17], [18]. This operation uses a classical postselection to determine if the encoded information was successfully restored. The clear drawback is that the procedure may fail with some probability. In such a case, we should reject the output state and ask for retransmission [19]. In the context of QEC, probabilistic decoding operations have found application in stabilizer codes [20] especially for iterative probabilistic decoding in LDPC codes [11], [21], [22], error decoding [23], [24] or environment-assisted error correction [25]. Moreover, it was noted that they have the potential to increase the spectrum of correctable errors [15] and are useful when the number of qubits is limited [14]. It is also worth mentioning that they were used with success in other fields of quantum information theory, *e.g.* probabilistic cloning [26], learning of unknown quantum operations [27] or measurement discrimination [28].

Even though the pQEC procedure has been studied in the literature for a while, there is a lack of a formal description of its application for a general noise model. In this work, we fill this gap. Inspired by the celebrated Knill-Laflamme conditions [29], we provide conditions (Theorem 1) to check, when probabilistic error correction is possible. We discover that optimal error-correcting codes are not always generated with the usage of isometric encoding operations. We give an explicit example of noise channels family (Section V), such that to maximize the probability of successful error correction, we need to encode the quantum information into a mixed state. Moreover, we discuss the advantage of the pQEC procedure over the deterministic one with a formal statement in Theorem 7. We show, in Theorem 17, how to correct noise channels with bounded Choi rank. Also, we observe the advantage of the pQEC procedure for random noise channels, which is presented in Theorem 21. Finally, if the errors are caused by a unitary interaction with an auxiliary qubit system, we show that it is possible to restore a qubit logical state by using only two physical qubits. We present a procedure how to achieve this in Algorithm 1.

The rest of the paper is organized as follows. In Section II we introduce the notation and define the pQEC protocol. In Section III we present equivalent conditions for probabilistically correctable noise channels. Then, we investigate a realization of the pQEC procedure in Section IV. In Section V we present a family of noise channels for which, it is necessary to use mixed state encoding to maximize the probability of

Manuscript received 18 July 2022; revised 6 January 2023; accepted 1 March 2023. Date of publication 8 March 2023; date of current version 16 June 2023. This work was supported in part by the project “Near-Term Quantum Computers: Challenges, Optimal Implementations and Applications,” which is carried out within the Team-Net Programme of the Foundation for Polish Science, under Grant POIR.04.04.00-00-17C1/18-00; and in part by the European Union under the European Regional Development Fund. *(Corresponding author: Ryszard Kukulski.)*

The authors are with the Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, 44-100 Gliwice, Poland (e-mail: rkukulski@iitis.pl).

Communicated by S. Mancini, Associate Editor for Quantum.

Digital Object Identifier 10.1109/TIT.2023.3254054

successful error correction. Then, we study the advantage of the pQEC procedure in Section VI and Section VII. In Section VIII we define a generalization of the pQEC protocol. Finally, we place all proofs in Appendix.

II. Preliminaries

A. Mathematical Framework

In this section, we will introduce the notation and recall necessary basic facts of quantum information theory. We will denote complex Euclidean spaces by symbols $\mathcal{X}, \mathcal{Y}, \ldots$. In a given space $\mathcal{X}$ we distinguish the computational basis $\{|i\rangle \in \mathcal{X} : i \in \{0, \ldots, \dim(\mathcal{X}) - 1\}\}$, that is a set of vectors, whose components are all zero, except one that equals 1, which span $\mathcal{X}$. The set of linear operators $M : \mathcal{X} \to \mathcal{Y}$ will be written as $\mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{M}(\mathcal{X}) := \mathcal{M}(\mathcal{X}, \mathcal{X})$. The identity operators will be denoted by $\mathbb{1}_\mathcal{X} \in \mathcal{M}(\mathcal{X})$. For any operator $M \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ we will consider its vectorization $|M\rangle \in \mathcal{Y} \otimes \mathcal{X}$, which is defined as

$$|M\rangle := \left(\mathbb{1}_\mathcal{Y} \otimes M^\top\right) \sum_{i=0}^{\dim(\mathcal{Y})-1} |i\rangle \otimes |i\rangle. \tag{1}$$

For any $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, $B \in \mathcal{M}(\mathcal{Y}, \mathcal{Z})$ and $C \in \mathcal{M}(\mathcal{Z}, \mathcal{T})$ it holds $|CBA\rangle = \left(C \otimes A^\top\right) |B\rangle$ [30].

In the space $\mathcal{M}(\mathcal{X})$, we distinguish the set of positive semi-definite operators $\mathcal{P}(\mathcal{X})$, the space of Hermitian operators $\mathcal{H}(\mathcal{X})$ and the set of unitary operators $\mathcal{U}(\mathcal{X})$. We use the convention that for non-invertible operator M, by M^{-1}, we denote its Moore-Penrose pseudo-inverse [30]. The set of quantum states, that is, the set of positive semi-definite operators with unit trace will be denoted by $\mathcal{D}(\mathcal{X})$. We say that a quantum state ρ is a pure state if $\operatorname{rank}(\rho) = 1$, otherwise, if $\operatorname{rank}(\rho) > 1$, we say that ρ is a mixed state. The maximally mixed state will be denoted by $\rho^*_\mathcal{X} := \frac{1}{\dim(\mathcal{X})} \mathbb{1}_\mathcal{X}$.

We also consider transformations between linear operators. We denote by $\mathcal{I}_\mathcal{X} : \mathcal{M}(\mathcal{X}) \to \mathcal{M}(\mathcal{X})$ the identity map. Let us define the set of quantum subchannels $s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ [31]. A quantum subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ is a linear map $\Phi : \mathcal{M}(\mathcal{X}) \to \mathcal{M}(\mathcal{Y})$, which is completely positive [30, Theorem 2.22], *i.e.*

$$(\Phi \otimes \mathcal{I}_\mathcal{X})(Q) \in \mathcal{P}(\mathcal{Y} \otimes \mathcal{X}) \quad \text{for any } Q \in \mathcal{P}(\mathcal{X} \otimes \mathcal{X}) \tag{2}$$

and trace non-increasing

$$\operatorname{tr}(\Phi(\rho)) \leq 1 \quad \text{for any } \rho \in \mathcal{D}(\mathcal{X}). \tag{3}$$

In particular, the subchannel Φ which is trace preserving, *i.e.*

$$\operatorname{tr}(\Phi(\rho)) = 1 \quad \text{for any } \rho \in \mathcal{D}(\mathcal{X}) \tag{4}$$

will be called a quantum channel. We denote by $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ the set of quantum channels $\Phi : \mathcal{M}(\mathcal{X}) \to \mathcal{M}(\mathcal{Y})$. We will also use the following notation, $s\mathcal{C}(\mathcal{X}) := s\mathcal{C}(\mathcal{X}, \mathcal{X})$ and $\mathcal{C}(\mathcal{X}) := \mathcal{C}(\mathcal{X}, \mathcal{X})$.

In this work, we will consider the following representations of subchannels:

- Kraus representation: Each subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ can be defined by a collection of Kraus operators $(K_i)_{i=1}^r \subset \mathcal{M}(\mathcal{X}, \mathcal{Y})$, such that $\Phi(X) = \sum_{i=1}^r K_i X K_i^\dagger$ for $X \in \mathcal{M}(\mathcal{X})$ and $r \in \mathbb{N}$. The operators K_i satisfy the condition $\sum_{i=1}^r K_i^\dagger K_i \leq \mathbb{1}_\mathcal{X}$. We say that the subchannel Φ is given in a canonical Kraus representation $(K_i)_{i=1}^r$, if it holds that $\operatorname{tr}(K_j^\dagger K_i) \propto \delta_{ij}$ and $K_i \neq 0$ for each $i \leq r$. To represent the subchannel Φ by its Kraus representation $(K_i)_{i=1}^r$, we introduce the notation $\mathcal{K} : \mathcal{M}(\mathcal{X}, \mathcal{Y})^{\times r} \to s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ given by $\Phi = \mathcal{K}\left((K_i)_{i=1}^r\right)$. Finally, the Kraus representation is not unique. It holds that $\mathcal{K}\left((K_i)_{i=1}^r\right) = \mathcal{K}\left((K_i')_{i=1}^r\right)$ if and only if there exists $U \in \mathcal{U}(\mathbb{C}^r)$ such that $K_i' = \sum_j U_{ij} K_j$ for any i.
- Choi-Jamiołkowski representation: Each subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ can be uniquely described by its Choi-Jamiołkowski operator $J(\Phi) \in \mathcal{M}(\mathcal{Y} \otimes \mathcal{X})$, which is defied as $J(\Phi) := (\Phi \otimes \mathcal{I}_\mathcal{X})(|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|)$. The rank of $J(\Phi)$ is called the Choi rank and it determines the minimal number r of Kraus operators K_i needed to describe Φ in the Kraus form $\Phi = \mathcal{K}\left((K_i)_{i=1}^r\right)$. Therefore, if the Kraus representation $(K_i)_{i=1}^r$ is canonical, then $r = \operatorname{rank}(J(\Phi))$.
- Stinespring representation: By the Stinespring Dilatation Theorem any subchannel $\Phi \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ can be defined as $\Phi(X) = \operatorname{tr}_2\left(AXA^\dagger\right)$ for $X \in \mathcal{M}(\mathcal{X})$, where $A \in \mathcal{M}(\mathcal{X}, \mathcal{Y} \otimes \mathbb{C}^r)$ and tr_2 is the partial trace over the second subsystem $\mathbb{C}^r$. The minimal dimension r of the auxiliary system is equal to the Choi rank. In particular, for $\Phi \in \mathcal{C}(\mathcal{X})$, the Stinespring representation of Φ can be written in the form $\Phi(X) = \operatorname{tr}_2\left(U(X \otimes |\psi\rangle\langle\psi|)U^\dagger\right)$, where $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^r)$ and $U \in \mathcal{U}(\mathcal{X} \otimes \mathbb{C}^r)$.

B. Problem Formulation

In this work, we consider the following procedure of probabilistic quantum error correction. We are given a noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and a Euclidean space $\mathcal{X}$. The goal of pQEC is to choose an appropriate encoding operation $\mathcal{S} \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and decoding operation $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$, such that for any state $\rho \in \mathcal{D}(\mathcal{X})$ we have $\mathcal{RES}(\rho) \propto \rho$. In this protocol, the pair $(\mathcal{S}, \mathcal{R})$ represents the error-correcting scheme and the quantity $\operatorname{tr}(\mathcal{RES}(\rho))$ represents the probability of successful error correction. This protocol may fail with the probability $1 - \operatorname{tr}(\mathcal{RES}(\rho))$. In such a case, the output state is rejected. To exclude a trivial, null strategy, we add the constrain that a valid error-correcting scheme must satisfy $\operatorname{tr}(\mathcal{RES}(\rho)) > 0$ for any $\rho \in \mathcal{D}(\mathcal{X})$.

In this set-up, the probability of successful error correction does not depend on the input state ρ (see Lemma 24 in Appendix A). We use this fact to standardize the definition of pQEC. From now, we say that $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ is probabilistically correctable for $\mathcal{X}$, if there exists an error-correcting scheme $(\mathcal{S}, \mathcal{R})$ such that

$$0 \neq \mathcal{RES} \propto \mathcal{I}_\mathcal{X}. \tag{5}$$

We say that $\mathcal{E}$ is correctable perfectly if $\mathcal{RES} = \mathcal{I}_\mathcal{X}$. In this work, we will denote by $p_\mathcal{X}(\mathcal{E})$ the maximal probability of successful error correction for a given $\mathcal{E}$ and $\mathcal{X}$, that is

$$p_\mathcal{X}(\mathcal{E}) := \max\left\{p : \mathcal{RES} = p\mathcal{I}_\mathcal{X}, \; (\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})\right\}. \tag{6}$$

III. Probabilistic Quantum Error Correction

To inspect the pQEC procedure, first, we should state conditions which determine when a given noise channel is probabilistically correctable. For deterministic QEC, such conditions have been known for a long time as the Knill-Laflamme conditions [29]. Let $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y})$ be a given noise channel. Then, according to the Knill-Laflamme Theorem, $\mathcal{E}$ is perfectly correctable for $\mathcal{X}$ if and only if

$$S^\dagger E_j^\dagger E_i S \propto \mathbb{1}_\mathcal{X} \quad (7)$$

for all i, j and some isometry operator $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$. In the following theorem we generalize the above, to cover probabilistically correctable noise channels.

Theorem 1 (Equivalent Conditions for pQEC): Let $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y})$. The following conditions are equivalent:

(A) There exist error-correcting scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ and $p > 0$ such that

$$\mathcal{R}\mathcal{E}\mathcal{S} = p\mathcal{I}_\mathcal{X}. \quad (8)$$

(B) There exist $\mathcal{S} = \mathcal{K}((S_k)_k) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $R \in \mathcal{P}(\mathcal{Y})$, such that $R \leq \mathbb{1}_\mathcal{Y}$, for which it holds

$$\mathcal{K}\left(\left(\sqrt{R} E_i S_k\right)_{i,k}\right) = \mathcal{K}((A_i)_i), \quad (9)$$

where $A_i \neq 0$ and $A_j^\dagger A_i \propto \delta_{ij} \mathbb{1}_\mathcal{X}$.

(C) There exist $\mathcal{S} = \mathcal{K}((S_k)_k) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$, $R \in \mathcal{P}(\mathcal{Y})$, such that $R \leq \mathbb{1}_\mathcal{Y}$ and a matrix $M = [M_{jl,ik}]_{jl,ik} \neq 0$, for which it holds

$$\forall_{i,j,k,l} \quad S_l^\dagger E_j^\dagger R E_i S_k = M_{jl,ik} \mathbb{1}_\mathcal{X}. \quad (10)$$

(D) There exist $S_* \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ such that

$$\forall_i \quad R_* E_i S_* \propto \mathbb{1}_\mathcal{X} \quad (11)$$

and there exists i_0, for which it holds $R_* E_{i_0} S_* \neq 0$.

Moreover, if point (A) holds for $\mathcal{S} = \mathcal{K}((S_k)_k)$ and $\mathcal{R} = \mathcal{K}((R_l)_l)$, then $R \in \mathcal{P}(\mathcal{Y})$ from points (B) and (C) can be chosen to satisfy $R = \sum_l R_l^\dagger R_l$. It also holds that $R_l E_i S_k \propto \mathbb{1}_\mathcal{X}$ for any i, k, l.

The proof of Theorem 1 is presented in Appendix B. Let us discuss the meaning of the conditions stated in Theorem 1. The condition (B) presents a general form of probabilistically correctable noise channels $\mathcal{E}$. Such channels, after applying post-processing $\sqrt{R}$ behave as mixed isometry operations. They hide parts of an initial quantum system on orthogonal subspaces. The condition (C) may be used to obtain the form of a recovery subchannel $\mathcal{R}$ in the following way (see Appendix B):

1) Let $M = U^\dagger D U$ be the spectral decomposition of M.
2) Define $A_{ii'} = \sum_{a,b} \overline{U_{ii',ab}} \sqrt{R} E_a S_b$.
3) For each $A_{ii'} \neq 0$ define $\alpha_{ii'}$: $A_{ii'}^\dagger A_{ii'} = \alpha_{ii'} \mathbb{1}_\mathcal{X}$.
4) The recovery subchannel is given as
$$\mathcal{R} = \mathcal{K}\left(\left(\alpha_{ii'}^{-1/2} A_{ii'}^\dagger \sqrt{R}\right)_{i,i'}\right).$$

Finally, the condition (D) gives us a simple method to check if $\mathcal{E} = \mathcal{K}((E_i)_{i=1}^r)$ is probabilistically correctable for $\mathcal{X}$.

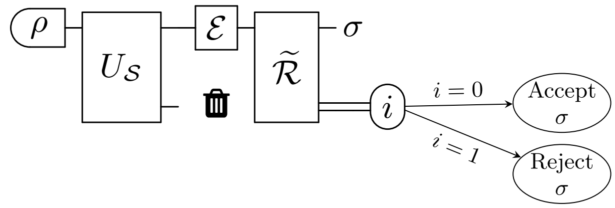

Fig. 1. Schematic realization of the pQEC procedure for the noise channel $\mathcal{E}$.

Let us compare the point (D) with the Knill-Laflamme conditions. The latter, is a constraint satisfaction problem with r^2 quadratic constrains $S^\dagger E_j^\dagger E_i S \propto \mathbb{1}_\mathcal{X}$ for the variable $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, which satisfies $S \neq 0$. The parameters $E_j^\dagger E_i$ constitute a $^\dagger-$closed algebra $\mathcal{A}$, such that $\mathbb{1}_\mathcal{Y} \in \mathcal{A}$. In comparison, the conditions in the point (D) represent a constraint satisfaction problem with r bilinear constrains $R E_i S \propto \mathbb{1}_\mathcal{X}$ for the variables $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$. Additionally, it must hold $R E_{i_0} S \neq 0$ for some $i_0 \in \{1, \ldots, r\}$. In this problem, the parameters E_i are arbitrary operators from $\mathcal{M}(\mathcal{Y})$, which satisfy $\operatorname{span}\left(\operatorname{im}(E_i^\dagger) : i = 1, \ldots, r\right) = \mathcal{Y}$ (although a stronger condition holds $\sum_i E_i^\dagger E_i = \mathbb{1}_\mathcal{Y}$, we will see in Section VI, it is more convenient to use the weaker version).

IV. Realization of pQEC Procedure

In this section, we will investigate the form of error-correcting scheme $(\mathcal{S}, \mathcal{R})$ which provides the maximal probability of successful error correction. For perfectly correctable noise channels, the encoding $\mathcal{S}$ can be realized by an isometry channel. This observation meaningfully reduces the complexity of finding error-correcting schemes – it is enough to consider a vector representation of pure states. Inspired by that, we ask if a similar behavior occurs in the probabilistic quantum error correction. The following proposition gives us some insight in the form of encoding and decoding.

Proposition 2: For a given channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, let us fix an error-correcting scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = p\mathcal{I}_\mathcal{X}$, for some $p > 0$. Then, the following holds:

(A) There exist $\widetilde{\mathcal{S}} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\widetilde{\mathcal{R}} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}} = p\mathcal{I}_\mathcal{X}$.

(B) If $\mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, then there exists $\widetilde{\mathcal{S}} = \mathcal{K}\left((\widetilde{S})\right) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ such that $\mathcal{R}\mathcal{E}\widetilde{\mathcal{S}} = \mathcal{I}_\mathcal{X}$.

(C) If $p = 1$, then there exist $\widetilde{\mathcal{S}} = \mathcal{K}\left((\widetilde{S})\right) \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\widetilde{\mathcal{R}} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}} = \mathcal{I}_\mathcal{X}$.

The proof of Proposition 2 is presented in Appendix C. We may use Proposition 2 (A) to state a realization of the pQEC procedure (see Figure 1). For a given noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ let $(\mathcal{S}, \mathcal{R}) \in \mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ be an error-correcting scheme for which $\mathcal{R}\mathcal{E}\mathcal{S} = p\mathcal{I}_\mathcal{X}$, where $p > 0$. The encoding channel $\mathcal{S}$ can be realized using the Stinespring representation given in the form $\mathcal{S}(X) = \operatorname{tr}_2\left(U_\mathcal{S} X U_\mathcal{S}^\dagger\right)$. The state is then sent through $\mathcal{E}$. The decoding subchannel $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ can be realized by implementing the channel $\widetilde{\mathcal{R}} \in \mathcal{C}(\mathcal{Y}, \mathcal{X} \otimes \mathbb{C}^2)$ given in the form $\widetilde{\mathcal{R}}(Y) = \mathcal{R}(Y) \otimes |0\rangle\langle 0| + \Psi(Y) \otimes |1\rangle\langle 1|$, where $\Psi \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $(\mathcal{R} + \Psi) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$.

In summary, the output of the whole procedure consists of a quantum state $\sigma \in \mathcal{D}(\mathcal{X})$ and a classical label $i \in \{0,1\}$. If the label $i=0$ is obtained, we know that $\sigma \propto \mathcal{RES}(\rho) = p\rho$, and hence, the output state can be accepted. Otherwise, if $i=1$, the output state $\sigma \propto \Psi\mathcal{ES}(\rho)$ should be rejected, as in general it may differ from ρ.

In Proposition 2 (C), we observed that using non-isometric channels $\mathcal{S}$ or formal subchannels $\mathcal{R}$ for perfectly correctable noise channels provides no advantage. Moreover, according to Theorem 1 (D), to predict if a noise channel is probabilistically correctable, we may consider only single Kraus encoding operations. However, among all conditions presented in Proposition 2 there is no condition, which in general allows us to restrict our attention to an isometry channel realization of $\mathcal{S}$. Indeed, there is a class of noise channels $\mathcal{E}$ for which, in order to maximize the probability p of successful error correction, we need to consider a general channel realization of $\mathcal{S}$. Paraphrasing, to obtain the best performance, we have to encode the initial state $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ into the mixed state $\mathcal{S}(|\psi\rangle\langle\psi|)$. In Section V we will present a family of noise channels for which it is necessary to use mixed state encoding.

V. Need for Mixed State Encoding

In this section, we provide an example of a parametrized family of noise channels $\{\mathcal{E}_R\}_R$ for which the mixed state encoding improves the probability of successful error correction. In our example we assume that $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$. For each $R \in \mathcal{P}(\mathbb{C}^4)$ satisfying $R \le \mathbb{1}_{\mathbb{C}^4}$ let us define a noise channel $\mathcal{E}_R \in \mathcal{C}(\mathbb{C}^4)$ given by the equation

$$\begin{aligned}\mathcal{E}_R(Y) = |0\rangle\langle 0| \otimes \operatorname{tr}_1\left(\sqrt{R}Y\sqrt{R}\right)\\ + |1\rangle\langle 1| \otimes \operatorname{tr}\left([\mathbb{1}_{\mathbb{C}^4} - R]Y\right)\rho_2^*.\end{aligned} \tag{12}$$

For a given R we define the optimal probability p_0 of successful error correction as

$$\begin{aligned}p_0(R) \coloneqq \max\{p:\ \mathcal{RE}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2},\\ (\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathbb{C}^2,\mathbb{C}^4) \times s\mathcal{C}(\mathbb{C}^4,\mathbb{C}^2)\}.\end{aligned} \tag{13}$$

We also define the optimal probability p_1 of successful error correction restricted to the pure state encoding:

$$\begin{aligned}p_1(R) \coloneqq \max\{p:\ \mathcal{RE}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2}, \mathcal{R} \in s\mathcal{C}(\mathbb{C}^4,\mathbb{C}^2),\\ \mathcal{S} = \mathcal{K}((S)) \in s\mathcal{C}(\mathbb{C}^2,\mathbb{C}^4),\ \ S \in \mathcal{M}(\mathbb{C}^2,\mathbb{C}^4)\}.\end{aligned} \tag{14}$$

Our claim, which we will prove, is that there exists a family of operators R for which $p_0(R) > p_1(R)$.

We start with the following lemma, where we show the optimal error-correcting scheme $(\mathcal{S},\mathcal{R})$ and a simplified version of the maximization problem $p_0(R)$.

Lemma 3: Let $R \in \mathcal{P}(\mathbb{C}^4)$ and $R \le \mathbb{1}_{\mathbb{C}^4}$. Define Π_R as a projector on the support of R. For $\mathcal{E}_R$ defined in (12) we have the following simplified form of the maximization problem $p_0(R)$:

$$\begin{aligned}p_0(R) = \max\Big\{\operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^2),\\ \operatorname{tr}_1\left(R^{-1}(P \otimes \mathbb{1}_{\mathbb{C}^2})\right) \le \mathbb{1}_{\mathbb{C}^2},\\ \forall_{X \in \mathcal{M}(\mathbb{C}^2)}\ \Pi_R(P \otimes X)\Pi_R = P \otimes X\Big\}.\end{aligned} \tag{15}$$

An optimal scheme $(\mathcal{S},\mathcal{R})$ which achieves the probability $p_0(R)$, that is $\mathcal{RE}_R\mathcal{S} = p_0(R)\mathcal{I}_{\mathbb{C}^2}$, can be taken as

$$\begin{aligned}\mathcal{S}(X) = \sqrt{R}^{-1}(P_0 \otimes X)\sqrt{R}^{-1},\\ \mathcal{R}(Y) = \operatorname{tr}_1\left(Y\left(|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2}\right)\right),\end{aligned} \tag{16}$$

where P_0 is an argument maximizing $p_0(R)$ in (15). Moreover, if there exists another optimal scheme $(\widetilde{\mathcal{S}},\widetilde{\mathcal{R}})$, that is $\widetilde{\mathcal{R}}\mathcal{E}_R\widetilde{\mathcal{S}} = p_0(R)\mathcal{I}_{\mathbb{C}^2}$, then $\operatorname{rank}(J(\mathcal{S})) \le \operatorname{rank}(J(\widetilde{\mathcal{S}}))$.

The proof of Lemma 3 is presented in Appendix D. Let us separately consider two cases: $\operatorname{rank}(R) < 4$ and $\operatorname{rank}(R) = 4$. The first one will be discussed briefly as it will not support our claim.

Corollary 4: Let us take $R \in \mathcal{P}(\mathbb{C}^4)$ such that $R \le \mathbb{1}_{\mathbb{C}^4}$ and $\operatorname{rank}(R) < 4$. Define Π_R as a projector on the support of R. For the noise channel defined in (12) we have $p_0(R) = p_1(R)$. Moreover, it holds

- If $\operatorname{rank}(R) \le 1$, then $p_0(R) = 0$.
- If $\operatorname{rank}(R) = 2, \Pi_R \ne |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}, |\psi\rangle \in \mathbb{C}^2$, then $p_0(R) = 0$.
- If $\operatorname{rank}(R) = 2, \Pi_R = |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}, |\psi\rangle \in \mathbb{C}^2$, then $p_0(R) = \|\operatorname{tr}_1\left(R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.
- If $\operatorname{rank}(R) = 3, \Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\alpha\rangle\langle\alpha|$ and $|\alpha\rangle \in \mathbb{C}^4$ is entangled, then $p_0(R) = 0$.
- If $\operatorname{rank}(R) = 3, \Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\psi^\perp\rangle\langle\psi^\perp| \otimes |\phi\rangle\langle\phi|$, where $|\psi^\perp\rangle, |\phi\rangle \in \mathbb{C}^2, |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$, then $p_0(R) = \|\operatorname{tr}_1\left(R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

The proof of Corollary 4 is presented in Appendix E.

In the case when the operator R is invertible, the situation is more interesting. Let us focus on $p_0(R)$ obtained in (15). As $\Pi_R = \mathbb{1}_{\mathbb{C}^4}$, the equation $\Pi_R(P \otimes X)\Pi_R = P \otimes X$ is always satisfied. For a given P, we can take $\rho \in \mathcal{D}(\mathbb{C}^2)$ such that $P = \operatorname{tr}(P)\rho$. The inequality $\operatorname{tr}(P)\operatorname{tr}_1\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right) \le \mathbb{1}_{\mathbb{C}^2}$ is then equivalent to $\operatorname{tr}(P) \le \|\operatorname{tr}_1\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$. Hence, we get

$$p_0(R) = \max\left\{\|\operatorname{tr}_1\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} :\ \rho \in \mathcal{D}(\mathbb{C}^2)\right\}. \tag{17}$$

To calculate $p_1(R)$ it will be sufficient to add the constraint $\mathcal{S} = \mathcal{K}((S))$. According to Lemma 3 the optimal $\mathcal{S}$ is of the form $\mathcal{S}(X) = \sqrt{R}^{-1}(P \otimes X)\sqrt{R}^{-1}$. As R is invertible, $\mathcal{S} = \mathcal{K}((S))$ if and only if $P = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathbb{C}^2$. Then, we have

$$\begin{aligned}p_1(R) = \max\big\{\|\operatorname{tr}_1\left(R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} :\\ |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)\big\}.\end{aligned} \tag{18}$$

Proposition 5: Let us define a unitary matrix $U \in \mathcal{U}(\mathbb{C}^4)$ which columns form the magic basis [32]

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ 0 & i & -1 & 0 \\ 1 & 0 & 0 & -i \end{bmatrix}. \tag{19}$$

Let us also define a diagonal operator $D(\lambda) \coloneqq \sum_{i=1}^4 \lambda_i |i\rangle\langle i|$, which is parameterized by a 4−dimensional real vector $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, for which it holds $0 < \lambda_i \le 1$. For

$R = UD(\lambda)U^\dagger$ and the noise channel $\mathcal{E}_R$ defined in (12) we have

$$p_0(R) = \frac{4}{\mathrm{tr}(R^{-1})},$$
$$p_1(R) = \frac{4}{\mathrm{tr}(R^{-1}) + c}, \tag{20}$$

where

$$c = \min\left\{\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|, \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right|\right\}. \tag{21}$$

The proof of Proposition 5 is presented in Appendix F. We can clearly see that in the case $\mathrm{rank}(R) = 4$, there are operators R, for which the mixed state encoding improves the probability of successful error correction over the pure state encoding, $p_0(R) > p_1(R)$. In general, the maximization problem in (17) intuitively supports the inequality $p_0(R) > p_1(R)$. The function $\rho \mapsto \| \mathrm{tr}_1 \left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right) \|_\infty$ is convex, so it is possible, that the minimal value of it will be achieved for some mixed state ρ. We observed such behavior in Proposition 5 for R given in the spectral decomposition $R = UD(\lambda)U^\dagger$. The introduced family of noise channels is parameterized by a $4-$dimensional vector $\lambda = (\lambda_1, \ldots, \lambda_4)$, such that $\lambda_i \in (0, 1]$. For almost all such λ we have $p_0(R) > p_1(R)$. The only exception is the $3-$dimensional subset defined by the relation

$$\frac{1}{\lambda_1} + \frac{1}{\lambda_4} = \frac{1}{\lambda_2} + \frac{1}{\lambda_3} \vee \left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| = \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|, \tag{22}$$

which describes the situation, when the pure state encoding match the mixed state encoding, $p_0(R) = p_1(R)$. In an extremal case, *e.g.* for $\lambda = (\frac{1}{2N}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, $N \in \mathbb{N}$, we get $p_1(R) = \frac{1}{N+1}$ and $p_0(R) = \frac{2}{N+3}$. Especially, when $N \to \infty$ the mixed state encoding provides the advantage, $p_0(R)/p_1(R) \to 2$.

The family of parameters R introduced in Proposition 5 is not the only one for which the minimum value of $\| \mathrm{tr}_1 \left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right) \|_\infty$ is achieved for a mixed state ρ. Let $R^{-1} \propto (\mathcal{I}_{\mathbb{C}^2} \otimes \Phi)(|\mathbb{1}_{\mathbb{C}^2}\rangle\langle\mathbb{1}_{\mathbb{C}^2}|)$ for some $\Phi \in \mathcal{C}(\mathbb{C}^2)$. Then, $\| \mathrm{tr}_1 \left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right) \|_\infty \propto \|\Phi(\rho^\top)\|_\infty$. Therefore, the value of $p_0(R)$ is one-to-one related with the maximum value of the output min-entropy of the channel Φ (see for instance [33]). Especially, we can see, if the image of the Bloch ball under Φ is a three dimensional ellipsis and contains the maximally mixed state ρ_2^* in its interior, then the mixed state encoding provides benefits.

Finally, the noise channel $\mathcal{E}_R$ defined for R from Proposition 5 is perfectly correctable for $\mathcal{X} = \mathbb{C}^2$ if and only if $R = \mathbb{1}_{\mathbb{C}^4}$. Interestingly, this suggests that perfectly correctable noise channels may constitute only a small subset of probabilistically correctable noise channels. This behavior will be the object of our investigation in the next section.

VI. Advantage of pQEC Procedure

The goal of this section is to show that the pQEC procedure corrects a wider class of noise channels than the QEC procedure based on the Knill-Laflamme conditions (7). For any Euclidean spaces $\mathcal{X}, \mathcal{Y}$ let us define two families of noise channels: these which are probabilistically correctable for $\mathcal{X}$, denoted as $\xi(\mathcal{X}, \mathcal{Y})$, and these which are correctable perfectly for $\mathcal{X}$, denoted as $\xi_1(\mathcal{X}, \mathcal{Y})$:

$$\begin{aligned} &\xi(\mathcal{X}, \mathcal{Y}) \coloneqq \\ &\{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) : \exists_{(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})} \ 0 \neq \mathcal{RES} \propto \mathcal{I}_\mathcal{X}\}, \\ &\xi_1(\mathcal{X}, \mathcal{Y}) \coloneqq \\ &\{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) : \exists_{(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})} \ \mathcal{RES} = \mathcal{I}_\mathcal{X}\}. \end{aligned} \tag{23}$$

We begin our analysis with some observations.

Proposition 6: For any $\mathcal{X}$, $\mathcal{Y}$ we have the following properties:

(A) $\xi_1(\mathcal{X}, \mathcal{Y}) \subset \xi(\mathcal{X}, \mathcal{Y})$,
(B) If $\dim(\mathcal{X}) > \dim(\mathcal{Y})$, then $\xi(\mathcal{X}, \mathcal{Y}) = \emptyset$,
(C) If $\dim(\mathcal{X}) \le \dim(\mathcal{Y})$, then $\xi_1(\mathcal{X}, \mathcal{Y}) \neq \emptyset$,
(D) If $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, then $\xi_1(\mathcal{X}, \mathcal{Y}) = \xi(\mathcal{X}, \mathcal{Y})$.

The proof of Proposition 6 is presented in Appendix G. We see that if $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, then there is no need to consider the pQEC procedure. The situation changes if we encode the initial information into a larger space, $\dim(\mathcal{Y}) > \dim(\mathcal{X})$. In the following theorem, we will show that $\xi_1(\mathcal{X}, \mathcal{Y}) \subsetneq \xi(\mathcal{X}, \mathcal{Y})$ for $\dim(\mathcal{Y}) > \dim(\mathcal{X})$.

Theorem 7: Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces for which $\dim(\mathcal{X}) < \dim(\mathcal{Y})$. Then, the set $\xi_1(\mathcal{X}, \mathcal{Y})$ is a nowhere dense subset of $\xi(\mathcal{X}, \mathcal{Y})$.

The proof of Theorem 7 is presented in Appendix H.

A. Choi Rank of Correctable Noise Channels

The intensity of a noise channel $\mathcal{E}$ can be connected with its Choi rank $r = \mathrm{rank}(J(\mathcal{E}))$. Given $\mathcal{E}$ in the Stinespring form, the Choi rank describes the dimension of an environment system which unitarily interacts with the encoded information. If the interaction is the weakest ($r = 1$) we deal with unitary noise channels, which are always perfectly correctable. The strongest interaction ($r = \dim(\mathcal{Y})^2$) is a property of noise channels that are difficult to correct. For example, the maximally depolarizing channel $\mathcal{E}(Y) = \mathrm{tr}(Y)\rho_\mathcal{Y}^*$, which can not be corrected, has the maximal Choi rank. In the following theorem, we investigate the maximum Choi rank of probabilistically correctable noise channels $\xi(\mathcal{X}, \mathcal{Y})$ and compare it with the maximum Choi rank for $\xi_1(\mathcal{X}, \mathcal{Y})$.

Theorem 8: Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that $\dim(\mathcal{Y}) \ge \dim(\mathcal{X})$. The following relations hold:

$$\begin{aligned} &\max\{\mathrm{rank}(J(\mathcal{E})) : \mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})\} \\ &= \dim(\mathcal{Y})^2 - \dim(\mathcal{Y})\dim(\mathcal{X}) + \left\lfloor \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})} \right\rfloor, \\ &\max\{\mathrm{rank}(J(\mathcal{E})) : \mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})\} \\ &= \dim(\mathcal{Y})^2 - \dim(\mathcal{X})^2 + 1. \end{aligned} \tag{24}$$

The proof of Theorem 8 is presented in Appendix I. In Proposition 6 we showed that if $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, then the the pQEC procedure gives us no advantage. Indeed, the only reversible noise channels, in this case, are unitary noise channels, that is channels with the Choi rank equal to one. We can ask, what is the maximum value of $r \in \mathbb{N}$, such that all noise channels which Choi rank is less or equal r, can be corrected perfectly or probabilistically, respectively. Formally speaking, for any $\mathcal{X}$ and $\mathcal{Y}$ we define the following quantities:

$$\begin{aligned} &r_1(\mathcal{X},\mathcal{Y}) := \\ &\max\left\{r\in\mathbb{N}:\forall_{\mathcal{E}\in\mathcal{C}(\mathcal{Y})}\operatorname{rank}(J(\mathcal{E}))\le r \implies \mathcal{E}\in\xi_1(\mathcal{X},\mathcal{Y})\right\},\\ &r(\mathcal{X},\mathcal{Y}) := \\ &\max\left\{r\in\mathbb{N}:\forall_{\mathcal{E}\in\mathcal{C}(\mathcal{Y})}\operatorname{rank}(J(\mathcal{E}))\le r \implies \mathcal{E}\in\xi(\mathcal{X},\mathcal{Y})\right\}. \end{aligned} \tag{25}$$

The quantity $r_1(\mathcal{X},\mathcal{Y})$ for a general noise model was studied in [34] and [35]. The authors of [34] calculated a lower bound for $r_1(\mathcal{X},\mathcal{Y})$ by using a technique of noise diagonalization along with Tverberg's theorem. They obtained the following result

$$r_1(\mathcal{X},\mathcal{Y}) \ge \max\left\{r\in\mathbb{N}:\dim(\mathcal{X})\le \frac{\left\lceil\frac{\dim(\mathcal{Y})}{r^2}\right\rceil + r^2}{r^2+1}\right\}. \tag{26}$$

It implies that $r_1(\mathcal{X},\mathcal{Y}) \ge \left\lfloor \sqrt[4]{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}}\right\rfloor$. On the other hand, by using the quantum packing bound [35] we may gain some insight of the upper bound for $r_1(\mathcal{X},\mathcal{Y})$. If we assume that we are allowed to use only non-degenerated codes, then for perfectly correctable $\mathcal{E}$ we have a bound of the form $\operatorname{rank}(J(\mathcal{E})) \le \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}$. In the next part of this section, we will improve the upper bound of $r_1(\mathcal{X},\mathcal{Y})$ without putting any additional assumptions. We also will estimate the behavior of $r(\mathcal{X},\mathcal{Y})$. In the particular case $\mathcal{X}=\mathbb{C}^2$ and $\mathcal{Y}=\mathbb{C}^4$, we will also show that $r_1(\mathcal{X},\mathcal{Y}) < r(\mathcal{X},\mathcal{Y})$.

Let us start with the following, simple but important properties, required to study $r(\mathcal{X},\mathcal{Y})$.

Lemma 9: Let $\mathcal{X},\mathcal{Y}$ be Euclidean spaces. Define $Q\in\mathcal{M}(\mathcal{Y})$ such that $0<Q\le \mathbb{1}_\mathcal{Y}$. Take $\mathcal{E}\in\mathcal{C}(\mathcal{Y})$ and $\mathcal{F}\in s\mathcal{C}(\mathcal{Y})$ given by $\mathcal{F}(Y)=\mathcal{E}(QYQ)$. Then, $\mathcal{E}\in\xi(\mathcal{X},\mathcal{Y})$ if and only if there exists a scheme $(\mathcal{S},\mathcal{R})\in s\mathcal{C}(\mathcal{X},\mathcal{Y})\times s\mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $0\ne\mathcal{RFS}\propto\mathcal{I}_\mathcal{X}$.

Directly from Lemma 9 we receive the monotonicity of $r(\mathcal{X},\mathcal{Y})$ w.r.t. the dimension of $\mathcal{Y}$. Let $\mathcal{Y},\mathcal{Y}'$ be such Euclidean spaces that $\dim(\mathcal{Y})\le\dim(\mathcal{Y}')$. Take $\mathcal{E}=\mathcal{K}((E_i)_i)\in\mathcal{C}(\mathcal{Y}')$. There exist two projectors $\Pi_1,\Pi_2\in\mathcal{P}(\mathcal{Y}')$, such that $\operatorname{rank}(\Pi_1)=\operatorname{rank}(\Pi_2)=\dim(\mathcal{Y})$ and for $\mathcal{F}=\mathcal{K}((\Pi_2E_i\Pi_1)_i)$ we have $\operatorname{rank}(\operatorname{tr}_1(J(\mathcal{F})))=\dim(\mathcal{Y})$. Hence, if there exists a scheme $(\mathcal{S},\mathcal{R})$ such that $0\ne\mathcal{RFS}\propto\mathcal{I}_\mathcal{X}$, then $\mathcal{E}\in\xi(\mathcal{X},\mathcal{Y}')$. Finally, we have

$$r(\mathcal{X},\mathcal{Y})\le r(\mathcal{X},\mathcal{Y}'). \tag{27}$$

B. From Bi-Linear to Linear Problem

In general, the difficulty of finding error-correcting schemes $(\mathcal{S},\mathcal{R})$ comes from bi-linearity of the problem (11). Calculating the maximal probability of successful error correction $p_\mathcal{X}(\mathcal{E})$ defined in (6) is even harder task. However, if we fix an encoding operation $\mathcal{S}$ (or decoding $\mathcal{R}$), it is possible to calculate $p_\mathcal{X}(\mathcal{ES})$ (or $p_\mathcal{X}(\mathcal{RE})$) using SDP programming. In this section, we extend the definition of $p_\mathcal{X}$ provided in (6) so that by $p_\mathcal{X}(\mathcal{F})$, where $\mathcal{F}\in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ we mean $p_\mathcal{X}(\mathcal{F})=\max\{p:\ \mathcal{RFS}=p\mathcal{I}_\mathcal{X},\ (\mathcal{S},\mathcal{R})\in s\mathcal{C}(\mathcal{X},\mathcal{Y})\times s\mathcal{C}(\mathcal{X})\}$. In the same way, for $\mathcal{F}\in s\mathcal{C}(\mathcal{X},\mathcal{Y})$ we will use the notation $p_\mathcal{X}(\mathcal{F})=\max\Big\{p:\ \mathcal{RFS}=p\mathcal{I}_\mathcal{X},\ (\mathcal{S},\mathcal{R})\in s\mathcal{C}(\mathcal{X})\times s\mathcal{C}(\mathcal{Y},\mathcal{X})\Big\}$.

Lemma 10: Let $\mathcal{F}=\mathcal{K}\left((F_i)_{i=0}^{r-1}\right)\in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. Then, it holds

$$\begin{aligned} &p_\mathcal{X}(\mathcal{F})\\ &=\max\{\operatorname{tr}(P):\ P\in\mathcal{P}(\mathbb{C}^r),\operatorname{tr}_1(R_F(P\otimes\mathbb{1}_\mathcal{X}))\le\mathbb{1}_\mathcal{X},\\ &\quad(\Pi_F\otimes\mathbb{1}_\mathcal{X})(P\otimes|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|)(\Pi_F\otimes\mathbb{1}_\mathcal{X})=P\otimes|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|\}, \end{aligned} \tag{28}$$

where $R_F=(FF^\dagger)^{-1},\Pi_F=FF^{-1}$ for $F=\sum_{i=0}^{r-1}|i\rangle\otimes F_i\in\mathcal{M}(\mathcal{Y},\mathbb{C}^r\otimes\mathcal{X})$. Moreover, if $p_\mathcal{X}(\mathcal{F})>0$, then

$$\|R_F\|_\infty^{-1}\le p_\mathcal{X}(\mathcal{F})\le\|R_F^{-1}\|_\infty. \tag{29}$$

The proof of Lemma 10 is presented in Appendix J.

Corollary 11: Let $\mathcal{F}=\mathcal{K}\left((F_i)_{i=0}^{r-1}\right)\in s\mathcal{C}(\mathcal{X},\mathcal{Y})$. Then, it holds

$$\begin{aligned} &p_\mathcal{X}(\mathcal{F})=\max\Big\{\operatorname{tr}(P):\ P\in\mathcal{P}(\mathbb{C}^r),P\otimes\mathbb{1}_\mathcal{X}\le\widetilde{F}\widetilde{F}^\dagger,\\ &(\Pi_{\widetilde{F}}\otimes\mathbb{1}_\mathcal{X})(P\otimes|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|)(\Pi_{\widetilde{F}}\otimes\mathbb{1}_\mathcal{X})=P\otimes|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|\}, \end{aligned} \tag{30}$$

where $\Pi_{\widetilde{F}}=\widetilde{F}\widetilde{F}^{-1}$ for $\widetilde{F}=\sum_{i=0}^{r-1}|i\rangle\otimes F_i^\dagger\in\mathcal{M}(\mathcal{Y},\mathbb{C}^r\otimes\mathcal{X})$.

The proof of Corollary 11 is presented in Appendix K. One can note that it is possible to use a sequence of optimization procedures presented in Lemma 10 and Corollary 11 to increase the probability of successful error correction. In more details, if we have fixed decoding operation $\mathcal{R}_0$ we run the procedure presented in Lemma 10 for $\mathcal{R}_0\mathcal{E}$ to calculate an encoding operation $\mathcal{S}_0$. Then, for $\mathcal{ES}_0$ we run the procedure presented in Corollary 11 to calculate $\mathcal{R}_1$, and so on until the obtained sequence of probability values will converge.

Let us now consider a particular class of noise channels $\mathcal{E}\in\mathcal{C}(\mathcal{Y})$ satisfying $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y}))=\dim(\mathcal{X})$. For each such channel $\mathcal{E}=\mathcal{K}((E_i)_i)$, we may consider an associated channel $\mathcal{F}=\mathcal{K}\left((V^\dagger E_i)_i\right)\in\mathcal{C}(\mathcal{Y},\mathcal{X})$, where $V\in\mathcal{M}(\mathcal{X},\mathcal{Y})$ is an isometry operator with the image on the support of $\mathcal{E}(\mathbb{1}_\mathcal{Y})$. It turns out that for $\mathcal{F}\in\mathcal{C}(\mathcal{Y},\mathcal{X})$ we can easily rewrite the bi-linear problem as a linear one and exploit Lemma 10.

Corollary 12: Let $\mathcal{F}=\mathcal{K}\left((F_i)_{i=0}^{r-1}\right)\in\mathcal{C}(\mathcal{Y},\mathcal{X})$. Define $\Pi_F=FF^{-1}$, where $F=\sum_{i=0}^{r-1}|i\rangle\otimes F_i\in\mathcal{M}(\mathcal{Y},\mathbb{C}^r\otimes\mathcal{X})$. Then, it holds $p_\mathcal{X}(\mathcal{F})\in\{0,1\}$. Moreover, $\mathcal{F}$ is perfectly correctable for $\mathcal{X}$ if and only if there exists $0\ne|\psi\rangle\in\mathbb{C}^r$ such that $(\Pi_F\otimes\mathbb{1}_\mathcal{X})(|\psi\rangle\otimes|\mathbb{1}_\mathcal{X}\rangle)=|\psi\rangle\otimes|\mathbb{1}_\mathcal{X}\rangle$.

The proof of Corollary 12 is presented in Appendix L.

Proposition 13: Let $\mathcal{X}$ and $\mathcal{Y}$ be some complex Euclidean spaces and $\dim(\mathcal{X})\le\dim(\mathcal{Y})$.

(A) If $\mathcal{E}\in\mathcal{C}(\mathcal{Y})$ is a noise channel such that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y}))=\dim(\mathcal{X})$ and $\operatorname{rank}(J(\mathcal{E}))<\frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$, then $\mathcal{E}\in\xi_1(\mathcal{X},\mathcal{Y})$.

(B) There exists a noise channel $\mathcal{E}\in\mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_\mathcal{Y}))=\dim(\mathcal{X})$ and $\operatorname{rank}(J(\mathcal{E}))\ge\frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$, for which we have $\mathcal{E}\notin\xi(\mathcal{X},\mathcal{Y})$.

The proof of Proposition 13 is presented in Appendix M.

C. Schur Noise Channels

In this subsection, we restrict our attention to a particular family of noise channels whose Kraus operators are diagonal in the computational basis. In the literature, these channels are referred to as Schur channels [30, Theorem 4.19]. We use them to study an upper bound for $r(\mathcal{X},\mathcal{Y})$ and $r_1(\mathcal{X},\mathcal{Y})$.

Lemma 14: Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces such that $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. Then, there exists a Schur channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(J(\mathcal{E})) = \left\lceil \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1} \right\rceil$ and $\mathcal{E} \notin \xi(\mathcal{X},\mathcal{Y})$. Moreover, there exists a Schur channel $\mathcal{F} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(J(\mathcal{F})) = \left\lceil \sqrt{\left\lceil \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1} \right\rceil} \right\rceil$ and $\mathcal{F} \notin \xi_1(\mathcal{X},\mathcal{Y})$. Especially, that implies

$$r(\mathcal{X},\mathcal{Y}) < \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1},$$
$$r_1(\mathcal{X},\mathcal{Y}) < \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}}. \tag{31}$$

The proof of Lemma 14 is presented in Appendix N. The bounds obtained in Lemma 14 are asymptotically tight for Schur noise channels with $\dim(\mathcal{Y}) \to \infty$. To prove the tightness of the bound for perfectly correctable noise channels, we may use the construction provided in [34]. Hence, if we take a Schur channel $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y})$, such that $\operatorname{rank}(J(\mathcal{E})) \approx \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}}$, we obtain $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$. In the following proposition we will prove the tightness for probabilistically correctable Schur noise channels.

Proposition 15: Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces and $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$. For any Schur channels $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, such that $\operatorname{rank}(J(\mathcal{E})) < \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}$, it holds $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$.

The proof of Proposition 15 is presented in Appendix O. In the case of Schur channels we have a clear separation between probabilistically and perfectly correctable noise channels. It is worth mentioning that the proof of Proposition 15 is constructive, that is we provide a method to calculate $S \in \mathcal{M}(\mathcal{X},\mathcal{Y})$ and $R \in \mathcal{M}(\mathcal{Y},\mathcal{X})$, such that they satisfy the condition Theorem 1 (D). It turns out that R can be taken as a binary matrix, such that in each column there is at most one non-zero element. If we fix an appropriate decoding matrix R we can use Lemma 10 and SDP programming to calculate the lower bound for $p_{\mathcal{X}}(\mathcal{E})$.

Corollary 16: Let $\mathcal{Y}$ be an Euclidean space such that $\dim(\mathcal{Y}) \geq 2$ and let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ be a Schur channel. Then, $\mathcal{E} \in \xi(\mathbb{C}^2,\mathcal{Y})$ if and only if $\dim(\mathcal{Y}) > \operatorname{rank}(J(\mathcal{E}))$. Moreover, if $\mathcal{E} \in \xi(\mathbb{C}^2,\mathcal{Y})$ then $p_{\mathbb{C}^2}(\mathcal{E}) \geq \frac{1}{\operatorname{rank}(J(\mathcal{E}))^2}$.

The proof of Corollary 16 is presented in Appendix P.

D. Correctable Noise Channels With Bounded Choi Rank

In this subsection we will study the behavior of $r(\mathcal{X},\mathcal{Y})$ and $r_1(\mathcal{X},\mathcal{Y})$. Lower and upper bounds for both quantities will be summarized in the following theorem.

Theorem 17: Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. Then, we have

$$\left\lfloor \sqrt[4]{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}} \right\rfloor \leq r_1(\mathcal{X},\mathcal{Y}) \leq \left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil - 1$$
$$\leq r(\mathcal{X},\mathcal{Y}) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}. \tag{32}$$

The proof of Theorem 17 is presented in Appendix Q. Unfortunately, according to this theorem, there is no clear separation of $r(\mathcal{X},\mathcal{Y})$ and $r_1(\mathcal{X},\mathcal{Y})$ for arbitrary $\mathcal{X}$ and $\mathcal{Y}$.

Let us briefly discuss computability of pQEC codes for noise channels $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y})$ satisfying

$$\operatorname{rank}(J(\mathcal{E}))^2(\dim(\mathcal{X})-1) < \dim(\mathcal{Y}). \tag{33}$$

According to the proof of Theorem 17 we should first diagonalize the noise channel $\mathcal{E}$. To do it, let us consider $A = \Big\{ s \in \mathbb{N} : \exists_{\Pi_s \in \mathcal{P}(\mathcal{Y})}\, \Pi_s = \Pi_s^2, \operatorname{rank}(\Pi_s) = s, \operatorname{rank}(\mathcal{E}^\dagger(\Pi_s)) = \dim(\mathcal{Y}) \Big\}$. We need to find $s_0 = \min(A)$ and a corresponding projector $\Pi_{s_0} \in \mathcal{P}(\mathcal{Y})$, such that $\operatorname{rank}(\Pi_{s_0}) = s_0$ and $\operatorname{rank}(\mathcal{E}^\dagger(\Pi_{s_0})) = \dim(\mathcal{Y})$. It might be a challenge to find directly such a projector but we can do it indirectly by sampling random projectors.

Lemma 18: Let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$. Assume that there exists a projector Π_0 such that $\operatorname{rank}(\Pi_0) = s$ and $\operatorname{rank}(\mathcal{E}^\dagger(\Pi_0)) = \dim(\mathcal{Y})$. Let us consider a random projector Π sampled according to the Haar measure, such that $\operatorname{rank}(\Pi) = s$. Then, almost surely it holds $\operatorname{rank}(\mathcal{E}^\dagger(\Pi)) = \dim(\mathcal{Y})$.

We can use Lemma 18 to find an appropriate projector Π_{s_0} and use it to diagonalize $\mathcal{E}$ according to the proof of Theorem 17. As a consequence, we will get a Schur subchannel. In the second step of our construction, for the given Schur subchannel we use directly the proof of Proposition 15 which is constructive.

For now, we will calculate explicitly $r(\mathcal{X},\mathcal{Y})$ and $r_1(\mathcal{X},\mathcal{Y})$ for $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^3, \mathbb{C}^4$.

Proposition 19: For all $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ satisfying $\operatorname{rank}(J(\mathcal{E})) \leq 2$ we have $\mathcal{E} \in \xi(\mathbb{C}^2,\mathbb{C}^4)$.

The proof of Proposition 19 is presented in Appendix R. By using Theorem 17 and Proposition 19 we get the following advantage of the pQEC protocol for $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$.

Corollary 20: For $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$ we have

$$r_1(\mathcal{X},\mathcal{Y}) < r(\mathcal{X},\mathcal{Y}). \tag{34}$$

In particular, it holds

$$\begin{array}{ll} r_1(\mathbb{C}^2,\mathbb{C}^3) = 1 & r(\mathbb{C}^2,\mathbb{C}^3) = 1 \\ r_1(\mathbb{C}^2,\mathbb{C}^4) = 1 & r(\mathbb{C}^2,\mathbb{C}^4) = 2 \end{array} \tag{35}$$

E. Random Noise Channels

In this subsection, we will show the advantage of the pQEC procedure for randomly generated noise channels. We will follow the procedure of sampling quantum channels considered in [36], [37], and [38].

Let $r \in \mathbb{N}$ and let $(G_i)_{i=1}^r \subset \mathcal{M}(\mathcal{Y})$ be a tuple of random and independent Ginibre matrices (matrices with independent and identically distributed entries drawn from standard complex normal distribution). Define $Q = \sum_{i=1}^r G_i^\dagger G_i$. We define a random channel $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$ given as

$$\mathcal{E}_r = \mathcal{K}\left((G_i Q^{-1/2}))_{i=1}^r \right). \tag{36}$$

This sampling procedure induces the measure $\mathcal{P}$ on $\mathcal{C}(\mathcal{Y})$ whose support is defined on $\{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) : \operatorname{rank}(J(\mathcal{E})) \le r\}$.

Theorem 21: Let $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$ be a random quantum channel defined according to (36). Then, the following two implications hold

$$r < \frac{\dim(\mathcal{X})\dim(\mathcal{Y})}{\dim(\mathcal{X})^2 - 1} \implies \mathcal{P}\left(\mathcal{E}_r \in \xi(\mathcal{X},\mathcal{Y})\right) = 1,$$
$$\mathcal{P}\left(\mathcal{E}_r \in \xi_1(\mathcal{X},\mathcal{Y})\right) = 1 \implies r < \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X}) - 1}}. \quad (37)$$

The proof of Theorem 21 is presented in Appendix S.

Corollary 22: Let $\mathcal{E}_r = \mathcal{K}\left((E_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y})$ be a random quantum channel defined according to (36) and assume that $r \le \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}$. Define a sequence $V_1, V_2, \dots$ of random isometry matrices sampled according to the Haar measure, such that $V_n \in \mathcal{M}(\mathcal{X},\mathcal{Y})$. Let $R_{F_n} = (F_n F_n^\dagger)^{-1}$ for $F_n = \sum_{i=0}^{r-1} |i\rangle \otimes V_n^\dagger E_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Then, almost surely it holds

$$\begin{aligned} &p_{\mathcal{X}}(\mathcal{E}_r) \\ &\ge \sup_{n\in\mathbb{N}} \max\left\{\operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \operatorname{tr}_1\left(R_{F_n}(P \otimes \mathbb{1}_{\mathcal{X}})\right) \le \mathbb{1}_{\mathcal{X}}\right\} \\ &\ge \max\left\{\lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V^\dagger) E E^\dagger (\mathbb{1}_{\mathbb{C}^r} \otimes V)) :\right. \\ &\qquad \left. V \in \mathcal{M}(\mathcal{X},\mathcal{Y}), V^\dagger V = \mathbb{1}_{\mathcal{X}}\right\}, \end{aligned} \quad (38)$$

where $\lambda_{\min}$ is the smallest eigenvalue and $E = \sum_{i=0}^{r-1} |i\rangle \otimes E_i$.

The proof of Corollary 22 is presented in Appendix T. Below, we provide a numerical analysis of a lower bound for $p_{\mathcal{X}}(\mathcal{E}_r)$, where we calculate $\max_{n=1,\dots,N} \max\{\operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \operatorname{tr}_1\left(R_{F_n}(P \otimes \mathbb{1}_{\mathcal{X}})\right) \le \mathbb{1}_{\mathcal{X}}\}$. For a given tuple (y, r, x), where $x = \dim(\mathcal{X})$ and $y = \dim(\mathcal{Y})$ we will sample M random channels $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$. For each random channel we will also sample N random isometry matrices $V \in \mathcal{M}(\mathcal{X},\mathcal{Y})$. As a result we will obtain a list of probability values $L = (p_i)_{i=1}^M$ calculated with a precision $\epsilon = 10^{-5}$. For each L we will provide: minimal element $\min(L)$, maximal element $\max(L)$, mean $\operatorname{mean}(L)$ and quartiles $Q_1(L), Q_2(L), Q_3(L)$ with a precision $\epsilon = 10^{-3}$.

- $(y, r, x) = (4, 2, 2)$ with $N = 10$ and $M = 1000$:

$\min(L)$	$\max(L)$	$\operatorname{mean}(L)$	$Q_1(L)$	$Q_2(L)$	$Q_3(L)$
0.197	0.819	0.505	0.436	0.498	0.570

- $(y, r, x) = (8, 2, 2)$ with $N = 20$ and $M = 500$:

$\min(L)$	$\max(L)$	$\operatorname{mean}(L)$	$Q_1(L)$	$Q_2(L)$	$Q_3(L)$
0.486	0.790	0.599	0.564	0.597	0.629

- $(y, r, x) = (8, 3, 2)$ with $N = 20$ and $M = 500$:

$\min(L)$	$\max(L)$	$\operatorname{mean}(L)$	$Q_1(L)$	$Q_2(L)$	$Q_3(L)$
0.293	0.541	0.399	0.371	0.396	0.427

- $(y, r, x) = (10, 5, 2)$ with $N = 30$ and $M = 100$:

$\min(L)$	$\max(L)$	$\operatorname{mean}(L)$	$Q_1(L)$	$Q_2(L)$	$Q_3(L)$
0.179	0.284	0.213	0.198	0.209	0.224

- $(y, r, x) = (10, 3, 3)$ with $N = 50$ and $M = 100$:

$\min(L)$	$\max(L)$	$\operatorname{mean}(L)$	$Q_1(L)$	$Q_2(L)$	$Q_3(L)$
0.178	0.329	0.226	0.207	0.220	0.241

VII. Example of pQEC Qubit Code

Consider the following scenario. You have a task to transfer a given qubit state $\rho \in \mathcal{D}(\mathbb{C}^2)$ through a quantum communication line represented by a noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ of the form $\mathcal{E}(Y) = \operatorname{tr}_2\left(U(Y \otimes |\psi\rangle\langle\psi|)U^\dagger\right)$, where $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$ and $U \in \mathcal{U}(\mathcal{Y} \otimes \mathbb{C}^2)$. At this point a natural question arises. What is the minimal size of the communication line, $\dim(\mathcal{Y})$, which is large enough to recover the state ρ with the pQEC procedure?

To answer this question, observe that the channel $\mathcal{E}$ satisfies $\operatorname{rank}(J(\mathcal{E})) \le 2$. In Proposition 19 we noticed that such channels are probabilistically correctable for a given input space $\mathbb{C}^2$, if $\dim(\mathcal{Y}) = 4$ (in fact, from monotonicity for $\dim(\mathcal{Y}) \ge 4$). Therefore, to correctly transfer a qubit state through $\mathcal{E}$, we may define an error-correcting scheme with only two physical qubits.

It is worth mentioning that some channels $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ which satisfy $\operatorname{rank}(J(\mathcal{E})) = 2$ are perfectly correctable for a space $\mathbb{C}^2$. If $\mathcal{E}$ is not an extreme point in the set of all channels $\mathcal{C}(\mathbb{C}^4)$, then $\mathcal{E}$ is mixed-unitary channel of the form $\mathcal{E}(Y) = pUYU^\dagger + (1-p)VYV^\dagger$ for some $U, V \in \mathcal{U}(\mathbb{C}^4)$ and $p \in (0,1)$ [39]. In that case, it was shown in [40] that $\mathcal{E} \in \xi_1(\mathbb{C}^2, \mathbb{C}^4)$. Nevertheless, if we consider a random channel $\mathcal{E}_2 \in \mathcal{C}(\mathbb{C}^4)$ defined according to (36) we will see that almost surely it is an extremal channel (see Lemma 25 in Appendix U). What is more, by Theorem 21 we even know that $\mathcal{P}\left(\mathcal{E}_2 \in \xi_1(\mathbb{C}^2, \mathbb{C}^4)\right) < 1$. A particular example of a Schur channel which is not perfectly correctable for $\mathbb{C}^2$ was constructed in the proof of Lemma 14. In fact, it follows from this construction that almost all Schur channels are not perfectly correctable.

We provide the following pQEC procedure based on Proposition 19 to probabilistically correct any $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ which satisfy $\operatorname{rank}(J(\mathcal{E})) = 2$.

VIII. Generalization of pQEC Procedure

Let us denote by Υ an arbitrary family of noise channels, that is $\Upsilon \subset \mathcal{C}(\mathcal{Y})$. In this section, we ask if there exists error-correcting scheme $(\mathcal{S}, \mathcal{R})$, such that all noise channels $\mathcal{E} \in \Upsilon$ we have $\mathcal{R}\mathcal{E}\mathcal{S} = p_{\mathcal{E}}\mathcal{I}_{\mathcal{X}}$, for some $p_{\mathcal{E}} \ge 0$. Note, that $p_{\mathcal{E}}$ may differ for different noise channels $\mathcal{E}$, hence, we shall introduce a quantity to "globally" control the effectiveness of $(\mathcal{S}, \mathcal{R})$. We propose the following approach.

Let μ be some probability measure defined on the set Υ. We assume that noise channels $\mathcal{E} \in \Upsilon$ are probed according to μ. The scheme $(\mathcal{S}, \mathcal{R})$ will be a valid error-correcting scheme for Υ and μ if in average, the probability of successful error correction is non zero, that is

$$\int_{\Upsilon} p_{\mathcal{E}} \mu(d\mathcal{E}) > 0. \quad (39)$$

Without loss of the generality we may assume that Υ is convex. Additionally, we assume that the support of μ is equal

Algorithm 1 Probabilistic QEC Qubit Code

Input: $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ such that $\mathrm{rank}(J(\mathcal{E})) \leq 2$.
Output: The pQEC procedure with success probability $p > 0$.

Let $\mathcal{E} = \mathcal{K}\left((E_0, E_1)\right)$.
Define $S_* \in \mathcal{M}(\mathbb{C}^2, \mathbb{C}^4)$ and $R_* \in \mathcal{M}(\mathbb{C}^4, \mathbb{C}^2)$, such that $R_* E_0 S_* \propto \mathbb{1}_{\mathbb{C}^2}$, $R_* E_1 S_* \propto \mathbb{1}_{\mathbb{C}^2}$ and $R_* E_0 S_* \neq 0 \vee R_* E_1 S_* \neq 0$ according to Appendix R.
Define

$$\begin{aligned} Q &= S_*^\dagger S_*, \\ S &= S_* Q^{-1/2}, \\ R &= \frac{Q^{1/2} R_*}{\|Q^{1/2} R_*\|_\infty}. \end{aligned}$$

Calculate $p \in (0, 1]$, such that $R\left(\mathcal{E}\left(SXS^\dagger\right)\right)R^\dagger = pX$ for any $X \in \mathcal{M}(\mathbb{C}^2)$.
Define $U_S \in \mathcal{U}(\mathbb{C}^4)$ which satisfies $U_S(\mathbb{1}_{\mathbb{C}^2} \otimes |0\rangle) = S$.
Let $R = \sigma_1 |z_1\rangle\langle t_1| + \sigma_2 |z_2\rangle\langle t_2|$ be the singular value decomposition of R. Define $U_R \in \mathcal{U}(\mathbb{C}^4)$ which satisfies

$$\begin{aligned} U_R|t_1\rangle &= |0, 0\rangle, \\ U_R|t_2\rangle &= |1, 0\rangle. \end{aligned}$$

Define $R' = RU_R^\dagger(\mathbb{1}_{\mathbb{C}^2} \otimes |0\rangle)$.
Define $V_R \in \mathcal{U}(\mathbb{C}^4)$ which satisfies $(\mathbb{1}_{\mathbb{C}^2} \otimes \langle 0|)V_R(\mathbb{1}_{\mathbb{C}^2} \otimes |0\rangle) = R'$.
Run the QEC procedure presented in Figure 2 for $|\psi\rangle, U_S, U_R, V_R$.
Let σ_{exp} be the output state of the procedure presented in Figure 2. Use the post-processing of the measurements' output (i, j) according to the following table:

Labels	$(i, j) = (0, 0)$	$(i, j) \neq (0, 0)$
Probability	p	$1-p$
Status	QEC succeeded	QEC failed
Action	Accept σ_{exp}	Reject σ_{exp}
Result	$\sigma_{\mathrm{exp}} = \|\psi\rangle\langle\psi\|$	σ_{exp} ? $\|\psi\rangle\langle\psi\|$

to Υ. Usually, we can take μ as the flat measure, representing the maximal uncertainty in the process of probing random noise channels $\mathcal{E}$ from Υ. Let us define the average noise channel of Υ with respect to μ

$$\bar{\mathcal{E}} = \int_\Upsilon \mathcal{E}\mu(d\mathcal{E}). \tag{40}$$

We will show that we can correct *all* noise channels from the family Υ, whenever $\bar{\mathcal{E}}$ is probabilistically correctable for $\mathcal{X}$. We put this statement as the following proposition.

Proposition 23: Let $\Upsilon \subset \mathcal{C}(\mathcal{Y})$ be a nonempty and convex family of noise channels. Define μ to be a probability measure defined on Υ and assume that the support of μ is equal to Υ. Let $\bar{\mathcal{E}} = \int_\Upsilon \mathcal{E}\mu(d\mathcal{E}) \in \mathcal{C}(\mathcal{Y})$ and fix $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X}, \mathcal{Y}) \times s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. The following conditions are equivalent:

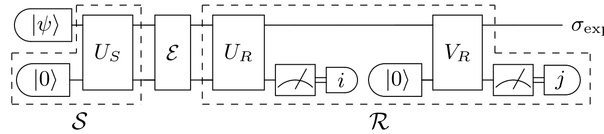

Fig. 2. The circuit representing the pQEC procedure. We have access to two physical qubits. The first qubit is in the state $|\psi\rangle$. This state will be encoded. The second state we set equal to $|0\rangle$. We implement the two-qubit encoding unitary operator U_S. Then, the encoded state, $U_S(|\psi\rangle \otimes |0\rangle)$, is affected by the noise channel $\mathcal{E}$. After that, we start the decoding procedure. We implement the two-qubit unitary rotation U_R. We measure the second qubit in the standard basis and obtain a classical label $i \in \{0, 1\}$. We prepare a third qubit in the state $|0\rangle$ and implement a two-qubit unitary rotation V_R. We measure the third qubit in the standard basis and obtain a classical label $j \in \{0, 1\}$. If $(i, j) = (0, 0)$ we accept the output state, otherwise, we reject it and request resend.

(A) For each $\mathcal{E} \in \Upsilon$ there exists $p_\mathcal{E} \geq 0$ such that $\mathcal{RES} = p_\mathcal{E}\mathcal{I}_\mathcal{X}$ and $\int_\Upsilon p_\mathcal{E}\mu(d\mathcal{E}) > 0$.
(B) It holds that $0 \neq \mathcal{R}\bar{\mathcal{E}}\mathcal{S} \propto \mathcal{I}_\mathcal{X}$.

The proof of Proposition 23 is presented in Appendix V.

IX. Conclusion

In this work, we analyzed the pQEC procedure for a general noise model. We established the necessary and sufficient conditions to check if a given noise channel is probabilistically correctable. Moreover, we showed that mixed state encoding should be taken into account when maximizing the probability of successful error correction. Finally, we pointed the advantage of the probabilistic error-correcting procedure over the deterministic one. We saw a clear separation especially for a correction of Schur noise channels and random noise channels. We obtained the maximum value of Choi rank of probabilistically correctable noise channels and provided a method how to probabilistically correct noise channels with bounded Choi rank.

There are many directions for further study that still remain to be explored. It would be interesting to strengthen Theorem 17 and show the separation between $r(\mathcal{X}, \mathcal{Y})$ and $r_1(\mathcal{X}, \mathcal{Y})$ by improving the proposed proof technique in Appendix Q. We obtained such separation for $\mathcal{X} = \mathbb{C}^2$ and $\mathcal{Y} = \mathbb{C}^4$ in Corollary 20. Another promising direction is to propose tools for the numerical analysis of the pQEC protocols, based on Theorem 1. Such tools would help us estimate the value of $r(\mathcal{X}, \mathcal{Y})$ and gain an insight into probabilistically correctable noises that require mixed state encoding. Last but not least, we would like to calculate the worst-case probability of successful error correction for a given noise intensity $r \leq r(\mathcal{X}, \mathcal{Y})$. For example, as we showed in Proposition 19, the errors caused by a unitary interaction with an auxiliary qubit system ($r = 2$), can be corrected by using only two physical qubits ($\dim(\mathcal{Y}) = 4$). We can ask, how many times in average the procedure presented in Algorithm 1 needs to be repeated.

Appendix A
Constant Probability of Successful Error Correction

Lemma 24: Let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, $\mathcal{S} \in s\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{R} \in s\mathcal{C}(\mathcal{Y}, \mathcal{X})$. If for any pure state $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ it holds $\mathcal{RES}(|\psi\rangle\langle\psi|) \propto |\psi\rangle\langle\psi|$, then there exists $p \in [0, 1]$ such that $\mathcal{RES} = p\mathcal{I}_\mathcal{X}$.

Proof: Let $\mathcal{L} = \mathcal{RES}$ and for any unitary operator $U \in \mathcal{U}(\mathcal{X})$ and $i = 0, \ldots, \dim(\mathcal{X}) - 1$ define $p_{U,i} \in [0,1]$ by $\mathcal{L}(U|i\rangle\langle i|U^\dagger) = p_{U,i}U|i\rangle\langle i|U^\dagger$. We have $\mathcal{L}(\mathbb{1}_\mathcal{X}) = U\left(\sum_i p_{U,i}|i\rangle\langle i|\right)U^\dagger$ for any U and hence, there exists $p \in [0,1]$ such that $\mathcal{L}(\mathbb{1}_\mathcal{X}) = p\mathbb{1}_\mathcal{X}$. That means, $p_{U,i} = p$ for any U and i, so $\mathcal{L}(|\psi\rangle\langle\psi|) = p|\psi\rangle\langle\psi|$ for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$. We obtain the thesis by noting that $\operatorname{span}_\mathbb{C}(|\psi\rangle\langle\psi|) = \mathcal{M}(\mathcal{X})$. □

APPENDIX B
PROOF OF THEOREM 1

Theorem 1: Let $\mathcal{E} = \mathcal{K}((E_i)_i) \in \mathcal{C}(\mathcal{Y})$. The following conditions are equivalent:

(A) There exist error-correcting scheme $(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})$ and $p > 0$ such that
$$\mathcal{RES} = p\mathcal{I}_\mathcal{X}. \tag{41}$$

(B) There exist $\mathcal{S} = \mathcal{K}((S_k)_k) \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$ and $R \in \mathcal{P}(\mathcal{Y})$, such that $R \le \mathbb{1}_\mathcal{Y}$, for which it holds
$$\mathcal{K}\left(\left(\sqrt{R}E_iS_k\right)_{i,k}\right) = \mathcal{K}((A_i)_i), \tag{42}$$
where $A_i \neq 0$ and $A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_\mathcal{X}$.

(C) There exist $\mathcal{S} = \mathcal{K}((S_k)_k) \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$, $R \in \mathcal{P}(\mathcal{Y})$, such that $R \le \mathbb{1}_\mathcal{Y}$ and a matrix $M = [M_{jl,ik}]_{jl,ik} \neq 0$, for which it holds
$$\forall_{i,j,k,l} \quad S_l^\dagger E_j^\dagger R E_i S_k = M_{jl,ik}\mathbb{1}_\mathcal{X}. \tag{43}$$

(D) There exist $S_* \in \mathcal{M}(\mathcal{X},\mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y},\mathcal{X})$ such that
$$\forall_i \quad R_*E_iS_* \propto \mathbb{1}_\mathcal{X} \tag{44}$$
and there exists i_0, for which it holds $R_*E_{i_0}S_* \neq 0$.

Moreover, if point (A) holds for $\mathcal{S} = \mathcal{K}((S_k)_k)$ and $\mathcal{R} = \mathcal{K}((R_l)_l)$, then $R \in \mathcal{P}(\mathcal{Y})$ from points (B) and (C) can be chosen to satisfy $R = \sum_l R_l^\dagger R_l$. It also holds that $R_lE_iS_k \propto \mathbb{1}_\mathcal{X}$ for any i,k,l.

Proof: In order to show that $(A) \iff (B) \iff (C)$, in all implications presented below, we will use the same encoding $\mathcal{S} = \mathcal{K}((S_k)_k) \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$. Hence, to simplify the proof, we introduce the notation of $\mathcal{F} \coloneqq \mathcal{ES}$ given in the form $\mathcal{F} = \mathcal{K}((F_i)_i)$.

$(B) \implies (A)$

Let us define $\alpha_i > 0$ to satisfy $A_i^\dagger A_i = \alpha_i\mathbb{1}_\mathcal{X}$ and a map $\mathcal{R}: \mathcal{M}(\mathcal{Y}) \to \mathcal{M}(\mathcal{X})$ given by
$$\mathcal{R} = \mathcal{K}\left(\left(\alpha_i^{-1/2}A_i^\dagger\sqrt{R}\right)_i\right). \tag{45}$$
We will check that $\mathcal{R}$ is a subchannel. First, from the definition of $\mathcal{R}$, it follows that $\mathcal{R}$ is completely positive. Second, from the assumption (B), operators $\alpha_i^{-1}A_iA_i^\dagger \in \mathcal{P}(\mathcal{Y})$ are orthogonal projectors and hence
$$\sum_i \alpha_i^{-1}\sqrt{R}A_iA_i^\dagger\sqrt{R} = \sqrt{R}\left(\sum_i \alpha_i^{-1}A_iA_i^\dagger\right)\sqrt{R} \le R \le \mathbb{1}_\mathcal{Y}. \tag{46}$$
It means that $\mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. Finally, it holds
$$\mathcal{RF} = \mathcal{K}\left(\left(\alpha_j^{-1/2}A_j^\dagger\sqrt{R}F_i\right)_{i,j}\right) = \mathcal{K}\left(\left(\alpha_j^{-1/2}A_j^\dagger A_i\right)_{i,j}\right) = \mathcal{K}\left(\left(\alpha_i^{1/2}\mathbb{1}_\mathcal{X}\right)_i\right) = p\mathcal{I}_\mathcal{X}, \tag{47}$$
where we introduced $p \coloneqq \sum_i \alpha_i > 0$.

$(A) \implies (B)$

Let $\mathcal{R} = \mathcal{K}((R_k)_k)$ and take $R = \sum_k R_k^\dagger R_k \in \mathcal{P}(\mathcal{Y})$. From $\mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ it follows $R \le \mathbb{1}_\mathcal{Y}$. Define Π_R to be the projector on the support of R. One can show that $R_k\Pi_R = R_k$ for each k. We define $\widetilde{\mathcal{R}} = \mathcal{K}\left(\left(\widetilde{R_k}\right)_k\right)$, where $\widetilde{R_k} \coloneqq R_k\sqrt{R}^{-1}$. From the definition of $\widetilde{\mathcal{R}}$ we have $\sum_k \widetilde{R_k}^\dagger\widetilde{R_k} = \sqrt{R}^{-1}R\sqrt{R}^{-1} = \Pi_R$. Using the assumption (A) we get $p\mathcal{I}_\mathcal{X} = \mathcal{RF} = \widetilde{\mathcal{R}} \circ \mathcal{K}\left((\sqrt{R}F_i)_i\right)$. As we have $p > 0$, it follows that $\mathcal{K}\left((\sqrt{R}F_i)_i\right) \neq 0$. Hence, there exists a canonical decomposition $\mathcal{K}\left((\sqrt{R}F_i)_i\right) = \mathcal{K}((A_i)_i)$, where $A_i \neq 0$ and $\operatorname{tr}(A_j^\dagger A_i) = 0$ for $i \neq j$. From the relationship between Kraus representations, it follows that A_i satisfy $\Pi_RA_i = A_i$. Then, by the assumption (A) we get
$$p|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}| = (\mathcal{RF}\otimes\mathcal{I}_\mathcal{X})(|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|) = \sum_i(\widetilde{\mathcal{R}}\otimes\mathcal{I}_\mathcal{X})(|A_i\rangle\langle A_i|). \tag{48}$$
Therefore, from the extremality of the point $|\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|$ in $\mathcal{P}(\mathcal{X}\otimes\mathcal{X})$ we have $\widetilde{R_k}A_i \propto \mathbb{1}_\mathcal{X}$ for any i,k. On the one hand we get
$$\sum_k A_j^\dagger\widetilde{R_k}^\dagger\widetilde{R_k}A_i \propto \mathbb{1}_\mathcal{X} \tag{49}$$
and on the other hand
$$\sum_k A_j^\dagger\widetilde{R_k}^\dagger\widetilde{R_k}A_i = A_j^\dagger\Pi_RA_i = A_j^\dagger A_i. \tag{50}$$
The above conditions provide that $A_j^\dagger A_i = c_{ji}\mathbb{1}_\mathcal{X}$, for some $c_{ji} \in \mathbb{C}$. Then, for $i \neq j$ we have $0 = \operatorname{tr}(A_j^\dagger A_i) = c_{ji}\dim(\mathcal{X})$ and finally $A_j^\dagger A_i = 0$.

$(B) \implies (C)$

Let us define $\alpha_k > 0$ to satisfy $A_k^\dagger A_k = \alpha_k\mathbb{1}_\mathcal{X}$. From the relationship between Kraus decompositions $\mathcal{K}\left((\sqrt{R}F_i)_i\right)$ and $\mathcal{K}((A_i)_i)$, there exists an isometry operator U, such that
$$\sqrt{R}F_i = \sum_k U_{ik}A_k. \tag{51}$$
Therefore, it holds
$$F_j^\dagger RF_i = \sum_{k,k'} U_{ik}\overline{U_{jk'}}A_{k'}^\dagger A_k = \sum_k U_{ik}\overline{U_{jk}}\alpha_k\mathbb{1}_\mathcal{X}. \tag{52}$$
Let us define a matrix $M = [M_{j,i}]_{j,i}$ where $M_{j,i} = \sum_k U_{ik}\overline{U_{jk}}\alpha_k$. Note, that
$$\operatorname{tr}(M) = \sum_{i,k}|U_{ik}|^2\alpha_k = \sum_k \alpha_k > 0. \tag{53}$$

$(C) \implies (B)$

Let $\mathcal{F} = \mathcal{K}((F_i)_{i=1}^r)$ for some $r \in \mathbb{N}$. Define an operator

$F=\sum_{i=1}^{r}\langle i|\otimes F_i \in \mathcal{M}(\mathbb{C}^r\otimes\mathcal{X},\mathcal{Y})$. From the assumption (C) it follows

$$F^\dagger RF = M\otimes \mathbb{1}_{\mathcal{X}}. \tag{54}$$

That implies $M\geq 0$. Take the spectral decomposition $M=U^\dagger DU$ and define

$$A_i=\sum_k \overline{U_{ik}}\sqrt{R}F_k. \tag{55}$$

Observe that $\mathcal{K}\left((\sqrt{R}F_i)_i\right)=\mathcal{K}\left((A_i)_i\right)$. We obtain

$$\begin{aligned}A_j^\dagger A_i =&\sum_{k,k'}\overline{U_{ik}}U_{jk'}F_{k'}^\dagger RF_k=\sum_{k,k'}\overline{U_{ik}}U_{jk'}M_{k'k}\mathbb{1}_{\mathcal{X}}\\ =&D_{ji}\mathbb{1}_{\mathcal{X}}.\end{aligned} \tag{56}$$

This is equivalent to $A_j^\dagger A_i\propto\delta_{ij}\mathbb{1}_{\mathcal{X}}$. Finally, $A_i\neq 0$ if and only if $D_{ii}>0$ and by the fact $M\neq 0$ we conclude the set $\{A_i : A_i\neq 0\}$ is non-empty.

$(A)\implies(D)$

Let $\mathcal{S}=\mathcal{K}((S_k)_k)$ and $\mathcal{R}=\mathcal{K}((R_l)_l)$. We have $(\mathcal{RES}\otimes\mathcal{I}_{\mathcal{X}})\left(|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|\right)=\sum_{l,i,k}|R_lE_iS_k\rangle\langle R_lE_iS_k|=p|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|$. Therefore, from the extremality of the point $|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|$ in $\mathcal{P}(\mathcal{X}\otimes\mathcal{X})$ we obtain $R_lE_iS_k\propto\mathbb{1}_{\mathcal{X}}$. There exist l_0,i_0,k_0 such that $R_{l_0}E_{i_0}S_{k_0}\neq 0$. We can take $S_*=S_{k_0}$ and $R_*=R_{l_0}$.

$(D)\implies(A)$

There exist $q_0,q_1>0$ for which $\mathcal{S}\coloneqq q_0\mathcal{K}\left((S_*)\right)\in s\mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\mathcal{R}\coloneqq q_1\mathcal{K}\left((R_*)\right)\in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. One may note that $0\neq\mathcal{RES}\propto\mathcal{I}_{\mathcal{X}}$.

$(*)$

Assume that $\mathcal{S}=\mathcal{K}\left((S_k)_k\right)$, $\mathcal{R}=\mathcal{K}\left((R_l)_l\right)$ and it holds (A). From the proof of implications $(A)\implies(B)$ and $(B)\implies(C)$ it follows that R can be chosen as $R=\sum_l R_l^\dagger R_l$. The relation $R_lE_iS_k\propto\mathbb{1}_{\mathcal{X}}$ was proven in $(A)\implies(D)$. □

Appendix C
Proof of Proposition 2

Proposition 2: For a given channel $\mathcal{E}\in\mathcal{C}(\mathcal{Y})$, let us fix an error-correcting scheme $(\mathcal{S},\mathcal{R})\in s\mathcal{C}(\mathcal{X},\mathcal{Y})\times s\mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\mathcal{RES}=p\mathcal{I}_{\mathcal{X}}$, for some $p>0$. Then, the following holds:

(A) There exist $\widetilde{\mathcal{S}}\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\widetilde{\mathcal{R}}\in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}}=p\mathcal{I}_{\mathcal{X}}$.

(B) If $\mathcal{R}\in\mathcal{C}(\mathcal{Y},\mathcal{X})$, then there exists $\widetilde{\mathcal{S}}=\mathcal{K}\left((\widetilde{S})\right)\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ such that $\mathcal{RE}\widetilde{\mathcal{S}}=\mathcal{I}_{\mathcal{X}}$.

(C) If $p=1$, then there exist $\widetilde{\mathcal{S}}=\mathcal{K}\left((\widetilde{S})\right)\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\widetilde{\mathcal{R}}\in\mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}}=\mathcal{I}_{\mathcal{X}}$.

Proof: (A)

Let $\mathcal{S}=\mathcal{K}\left((S_k)_k\right)$ and $S=\sum_k S_k^\dagger S_k\leq\mathbb{1}_{\mathcal{X}}$. Using Theorem 1 one can show that there exists k_0 for which $\mathrm{rank}(S_{k_0})=\dim(\mathcal{X})$. Hence, S is invertible. Define $\widetilde{\mathcal{S}}\in\mathcal{C}(\mathcal{X},\mathcal{Y}),\widetilde{\mathcal{R}}\in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ given by the equations

$$\begin{aligned}\widetilde{\mathcal{S}}(X)&=\mathcal{S}\left(S^{-1/2}XS^{-1/2}\right),\\ \widetilde{\mathcal{R}}(Y)&=S^{1/2}\mathcal{R}(Y)S^{1/2}.\end{aligned} \tag{57}$$

We obtain $\widetilde{\mathcal{R}}\mathcal{E}\widetilde{\mathcal{S}}(X)=S^{1/2}(\mathcal{RES})\left(S^{-1/2}XS^{-1/2}\right)S^{1/2}=pX$.

(B)

Let $\mathcal{S}=\mathcal{K}\left((S_k)_k\right)$ and define $\mathcal{S}_k(X)=S_kXS_k^\dagger$. From Theorem 1 there exists k_0 such that $\mathcal{RES}_{k_0}=p_{k_0}\mathcal{I}_{\mathcal{X}}$, for some $p_{k_0}>0$. For any $|\psi\rangle\langle\psi|\in\mathcal{D}(\mathcal{X})$ it holds then

$$\begin{aligned}p_{k_0}&=\mathrm{tr}\left(\mathcal{RES}_{k_0}(|\psi\rangle\langle\psi|)\right)=\mathrm{tr}\left(\mathcal{S}_{k_0}(|\psi\rangle\langle\psi|)\right)\\ &=\langle\psi|S_{k_0}^\dagger S_{k_0}|\psi\rangle.\end{aligned} \tag{58}$$

Hence, we get $S_{k_0}^\dagger S_{k_0}=p_{k_0}\mathbb{1}_{\mathcal{X}}$. Define $\widetilde{\mathcal{S}}=\frac{1}{p_{k_0}}\mathcal{S}_{k_0}\in\mathcal{C}(\mathcal{X},\mathcal{Y})$ and note that $\mathcal{RE}\widetilde{\mathcal{S}}=\mathcal{I}_{\mathcal{X}}$.

(C)

Let $\mathcal{S}=\mathcal{K}\left((S_k)_k\right)$ and $\mathcal{R}=\mathcal{K}\left((R_k)_k\right)$. For any $|\psi\rangle\langle\psi|\in\mathcal{D}(\mathcal{X})$ we have

$$1=\mathrm{tr}(\mathcal{RES}(|\psi\rangle\langle\psi|))\leq\mathrm{tr}(\mathcal{S}(|\psi\rangle\langle\psi|))\leq 1. \tag{59}$$

Therefore, for any $|\psi\rangle\langle\psi|\in\mathcal{D}(\mathcal{X})$ we get $\langle\psi|\left(\sum_k S_k^\dagger S_k\right)|\psi\rangle=1$, which implies $\mathcal{S}\in\mathcal{C}(\mathcal{X},\mathcal{Y})$. Let $R=\sum_k R_k^\dagger R_k\leq\mathbb{1}_{\mathcal{Y}}$. Then, it holds $\mathrm{tr}\left((\mathbb{1}_{\mathcal{Y}}-R)\mathcal{ES}(X)\ \right)=0$. Define $\widetilde{\mathcal{R}}\in\mathcal{C}(\mathcal{Y},\mathcal{X})$ by the equation

$$\widetilde{\mathcal{R}}(Y)\ =\mathcal{R}(Y)+\mathrm{tr}\left((\mathbb{1}_{\mathcal{Y}}-R)Y\right)\rho_{\mathcal{X}}^*. \tag{60}$$

Observe that $\widetilde{\mathcal{R}}\mathcal{ES}=\mathcal{I}_{\mathcal{X}}$. The rest of the proof follows from (B). □

Appendix D
Proof of Lemma 3

Lemma 3: Let $R\in\mathcal{P}(\mathbb{C}^4)$ and $R\leq\mathbb{1}_{\mathbb{C}^4}$. Define Π_R as a projector on the support of R. For $\mathcal{E}_R$ defined as

$$\mathcal{E}_R(Y)=|0\rangle\langle 0|\otimes\mathrm{tr}_1\left(\sqrt{R}Y\sqrt{R}\right)+|1\rangle\langle 1|\otimes\mathrm{tr}\left([\mathbb{1}_{\mathbb{C}^4}-R]Y\right)\rho_2^* \tag{61}$$

we have the following simplified form of the maximization problem (13):

$$\begin{aligned}p_0(R)=\max\big\{&\mathrm{tr}(P): P\in\mathcal{P}(\mathbb{C}^2),\\ &\mathrm{tr}_1\left(R^{-1}(P\otimes\mathbb{1}_{\mathbb{C}^2})\right)\leq\mathbb{1}_{\mathbb{C}^2},\\ &\forall_{X\in\mathcal{M}(\mathbb{C}^2)}\ \Pi_R(P\otimes X)\Pi_R=P\otimes X\big\}.\end{aligned} \tag{62}$$

An optimal scheme $(\mathcal{S},\mathcal{R})$ which achieves the probability $p_0(R)$, that is $\mathcal{RE}_R\mathcal{S}=p_0(R)\mathcal{I}_{\mathbb{C}^2}$, can be taken as

$$\begin{aligned}\mathcal{S}(X)&=\sqrt{R}^{-1}(P_0\otimes X)\sqrt{R}^{-1},\\ \mathcal{R}(Y)&=\mathrm{tr}_1\left(Y\left(|0\rangle\langle 0|\otimes\mathbb{1}_{\mathbb{C}^2}\right)\right),\end{aligned} \tag{63}$$

where P_0 is an argument maximizing $p_0(R)$ in (62). Moreover, if there exists another optimal scheme $(\widetilde{\mathcal{S}},\widetilde{\mathcal{R}})$, that is $\widetilde{\mathcal{R}}\mathcal{E}_R\widetilde{\mathcal{S}}=p_0(R)\mathcal{I}_{\mathbb{C}^2}$, then $\mathrm{rank}(J(\mathcal{S}))\leq\mathrm{rank}(J(\widetilde{\mathcal{S}}))$.

Proof: Let us investigate the form of an optimal scheme $(\mathcal{S},\mathcal{R})$ that maximizes the probability p of successful error correction, $\mathcal{RE}_R\mathcal{S}=p\mathcal{I}_{\mathbb{C}^2}$. First, one can note that $\mathcal{R}$ must be of the form $\mathcal{R}(A\otimes B)=\mathrm{tr}\left(A|0\rangle\langle 0|\right)\widetilde{\mathcal{R}}(B)$, where $\widetilde{\mathcal{R}}=\mathcal{K}\left((\widetilde{R_k})_k\right)\in s\mathcal{C}(\mathbb{C}^2)$. Let us introduce a map $\mathcal{F}=\mathcal{K}\left((F_i)_i\right)\in s\mathcal{C}(\mathbb{C}^2)$ given by $\mathcal{F}(X)=\mathrm{tr}_1\left(\sqrt{R}\mathcal{S}(X)\sqrt{R}\right)$.

We obtain $p\mathcal{I}_{\mathbb{C}^2} = \mathcal{R}\mathcal{E}_R\mathcal{S} = \widetilde{\mathcal{R}}\mathcal{F}$. From Theorem 1 we have $\widetilde{R}_k F_i \propto \mathbb{1}_{\mathbb{C}^2}$ and there are k_0, i_0 such that $\widetilde{R}_{k_0} F_{i_0} \neq 0$. Hence, for each k we have $\widetilde{R}_k \propto F_{i_0}^{-1}$. That implies the map $\widetilde{\mathcal{R}}$ can be written as $\widetilde{\mathcal{R}}(X) = \widetilde{R} X \widetilde{R}^\dagger$. Now, consider another scheme $(\mathcal{S}', \mathcal{R}')$, where $\mathcal{R}'(A \otimes B) = \operatorname{tr}(A|0\rangle\langle 0|) B$ and $\mathcal{S}'(X) = \mathcal{S}\left(\widetilde{R} X \widetilde{R}^\dagger\right) \in s\mathcal{C}(\mathbb{C}^2, \mathbb{C}^4)$. We get

$$\begin{aligned} \mathcal{R}'\mathcal{E}_R\mathcal{S}'(X) &= \operatorname{tr}_1\left(\sqrt{R}\mathcal{S}\left(\widetilde{R} X \widetilde{R}^\dagger\right)\sqrt{R}\right) = \mathcal{F}\left(\widetilde{R} X \widetilde{R}^\dagger\right) \\ &= \left(\widetilde{R}\right)^{-1}\left(\widetilde{\mathcal{R}}\mathcal{F}\left(\widetilde{R} X \widetilde{R}^\dagger\right)\right)\left(\widetilde{R}^\dagger\right)^{-1} = pX. \end{aligned} \tag{64}$$

Therefore, the scheme $(\mathcal{S}', \mathcal{R}')$ is also optimal and $\operatorname{rank}(J(\mathcal{S}')) \leq \operatorname{rank}(J(\mathcal{S}))$.

To sum up, from now, we will consider the optimal scheme $(\mathcal{S}, \mathcal{R})$, where $\mathcal{R}(Y) = \operatorname{tr}_1(Y(|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2}))$. The equation $\mathcal{R}\mathcal{E}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2}$ can be rewritten as

$$\operatorname{tr}_1\left(\sqrt{R}\mathcal{S}(X)\sqrt{R}\right) = pX, \tag{65}$$

for any $X \in \mathcal{M}(\mathbb{C}^2)$. According to Theorem 1 we have $\sqrt{R}\mathcal{S}(X)\sqrt{R} = \sum_i A_i X A_i^\dagger$, where $A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_{\mathbb{C}^2}$. Using Theorem 1 to the equation $\operatorname{tr}_1\left(\sum_i A_i X A_i^\dagger\right) = pX$ we obtain that $A_i = |v_i\rangle \otimes \mathbb{1}_{\mathbb{C}^2}$ for some orthogonal vectors $|v_i\rangle \in \mathbb{C}^2$. Let $P = \sum_i |v_i\rangle\langle v_i|$. We get $\sqrt{R}\mathcal{S}(X)\sqrt{R} = P \otimes X$. Without loss of the generality we may consider $\mathcal{S}$ such that $\Pi_R \mathcal{S}(X) \Pi_R = \mathcal{S}(X)$ (one can note that $\operatorname{rank}(J(\mathcal{S}))$ will not increase). Hence, the equation $\sqrt{R}\mathcal{S}(X)\sqrt{R} = P \otimes X$ implies $\mathcal{S}(X) = \sqrt{R}^{-1}(P \otimes X)\sqrt{R}^{-1}$. The condition $\sqrt{R}\mathcal{S}(X)\sqrt{R} = P \otimes X$ becomes now equivalent to $\Pi_R(P \otimes X)\Pi_R = P \otimes X$. By the Choi isomorphism the condition $\mathcal{S} \in s\mathcal{C}(\mathbb{C}^2, \mathbb{C}^4)$ is then equivalent to $\operatorname{tr}_1\left(R^{-1}(P \otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}$. Therefore, basing on (65) we can express the probability $p_0(R)$ as:

$$\begin{aligned} p_0(R) = \max\{&p : \mathcal{R}\mathcal{E}_R\mathcal{S} = p\mathcal{I}_{\mathbb{C}^2}, \\ &(\mathcal{S}, \mathcal{R}) \in s\mathcal{C}(\mathbb{C}^2, \mathbb{C}^4) \times s\mathcal{C}(\mathbb{C}^4, \mathbb{C}^2)\} \\ = \max\Big\{&p : \forall_{X \in \mathcal{M}(\mathbb{C}^2)} \operatorname{tr}_1\left(\sqrt{R}\mathcal{S}(X)\sqrt{R}\right) = pX, \\ &\mathcal{S} \in s\mathcal{C}(\mathbb{C}^2, \mathbb{C}^4)\Big\} \\ = \max\Big\{&\operatorname{tr}(P) : P \in \mathcal{P}(\mathbb{C}^2), \\ &\operatorname{tr}_1\left(R^{-1}(P \otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}, \\ &\forall_{X \in \mathcal{M}(\mathbb{C}^2)} \Pi_R(P \otimes X)\Pi_R = P \otimes X\Big\}. \end{aligned} \tag{66}$$

□

Appendix E
Proof of Corollary 4

Corollary 4: Let us take $R \in \mathcal{P}(\mathbb{C}^4)$ such that $R \leq \mathbb{1}_{\mathbb{C}^4}$ and $\operatorname{rank}(R) < 4$. Define Π_R as a projector on the support of R. For the noise channel defined as

$$\mathcal{E}_R(Y) = |0\rangle\langle 0| \otimes \operatorname{tr}_1\left(\sqrt{R}Y\sqrt{R}\right) + |1\rangle\langle 1| \otimes \operatorname{tr}\left([\mathbb{1}_{\mathbb{C}^4} - R]Y\right)\rho_2^* \tag{67}$$

we have $p_0(R) = p_1(R)$, where $p_0(R)$ and $p_1(R)$ are defined in (13) and (14), respectively. Moreover, it holds

- If $\operatorname{rank}(R) \leq 1$, then $p_0(R) = 0$.
- If $\operatorname{rank}(R) = 2, \Pi_R \neq |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}, |\psi\rangle \in \mathbb{C}^2$, then $p_0(R) = 0$.
- If $\operatorname{rank}(R) = 2, \Pi_R = |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}, |\psi\rangle \in \mathbb{C}^2$, then $p_0(R) = \|\operatorname{tr}_1\left(R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.
- If $\operatorname{rank}(R) = 3, \Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\alpha\rangle\langle\alpha|$ and $|\alpha\rangle \in \mathbb{C}^4$ is entangled, then $p_0(R) = 0$.
- If $\operatorname{rank}(R) = 3, \Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\psi^\perp\rangle\langle\psi^\perp| \otimes |\phi\rangle\langle\phi|$, where $|\psi^\perp\rangle, |\phi\rangle \in \mathbb{C}^2, |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$, then $p_0(R) = \|\operatorname{tr}_1\left(R^{-1}(|\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

Proof: The proof is based on Lemma 3. Let us investigate the value of $p_0(R)$. We will consider three cases depending on $\operatorname{rank}(R)$.

In the first case, we assume that $\operatorname{rank}(R) \in \{0, 1\}$. Then, for P satisfying $\Pi_R(P \otimes X)\Pi_R = P \otimes X$ we have

$$\begin{aligned} 2\operatorname{rank}(P) &= \operatorname{rank}(P \otimes \mathbb{1}_{\mathbb{C}^2}) = \operatorname{rank}(\Pi_R(P \otimes \mathbb{1}_{\mathbb{C}^2})\Pi_R) \\ &\leq \operatorname{rank}(\Pi_R) \leq 1. \end{aligned} \tag{68}$$

Hence, we obtain $\operatorname{rank}(P) \leq \frac{1}{2}$ which implies $P = 0$. In this case $p_0(R) = 0$.

In the second case, we assume that $\operatorname{rank}(R) = 2$. Using the same argumentation for P as in the first case, we get $\operatorname{rank}(P) \leq 1$. We can write $P = |x\rangle\langle x|$ for $|x\rangle \in \mathbb{C}^2$. Note that, if $P \neq 0$, then from the equality $\Pi_R|x, y\rangle = |x, y\rangle$ for $|y\rangle \in \mathbb{C}^2$ we get $\Pi_R = |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}$, for $|\psi\rangle = \frac{1}{\|x\|}|x\rangle$. Therefore, if for all $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^2)$ it holds $\Pi_R \neq |\psi\rangle\langle\psi| \otimes \mathbb{1}_{\mathbb{C}^2}$, we have $p_0(R) = 0$. Otherwise, if $\Pi_R = |\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2}$ for $|\psi_0\rangle\langle\psi_0| \in \mathcal{D}(\mathbb{C}^2)$, we take $P = p|\psi_0\rangle\langle\psi_0|$ for $p \geq 0$. From the assumption $p\operatorname{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}$ we get $p_0(R) = \|\operatorname{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

In the third case, we assume that $\operatorname{rank}(R) = 3$. Again, P can be written in the form $P = |x\rangle\langle x|$ for $|x\rangle \in \mathbb{C}^2$. Let $\Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\xi\rangle\langle\xi|$, where $|\xi\rangle\langle\xi| \in \mathcal{D}(\mathbb{C}^4)$. If $P \neq 0$, then from the equality $\Pi_R|x, y\rangle = |x, y\rangle$ for $|y\rangle \in \mathbb{C}^2$ we get $\langle\xi|x, y\rangle = 0$, for $|y\rangle \in \mathbb{C}^2$, and hence, $|\xi\rangle \propto |x^\perp\rangle \otimes |y\rangle$. Therefore, if $|\xi\rangle$ is entangled, we have $p_0(R) = 0$. Otherwise, if $\Pi_R = \mathbb{1}_{\mathbb{C}^4} - |\psi_0^\perp\rangle\langle\psi_0^\perp| \otimes |\phi_0\rangle\langle\phi_0|$ for $|\psi_0^\perp\rangle\langle\psi_0^\perp|, |\phi_0\rangle\langle\phi_0| \in \mathcal{D}(\mathbb{C}^2)$, we take $P = p|\psi_0\rangle\langle\psi_0|$ for $p \geq 0$. The assumption $p\operatorname{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right) \leq \mathbb{1}_{\mathbb{C}^2}$ implies $p_0(R) = \|\operatorname{tr}_1\left(R^{-1}(|\psi_0\rangle\langle\psi_0| \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1}$.

To prove that $p_0(R) = p_1(R)$ observe that in each case the argument maximizing $p_0(R)$ has the form $P_0 = p|\psi_0\rangle\langle\psi_0|$. According to Lemma 3 the optimal scheme $(\mathcal{S}, \mathcal{R})$ can be taken as $\mathcal{S}(X) = \sqrt{R}^{-1}(P_0 \otimes X)\sqrt{R}^{-1}$ and $\mathcal{R}(Y) = \operatorname{tr}_1(Y(|0\rangle\langle 0| \otimes \mathbb{1}_{\mathbb{C}^2}))$. As the pair $(\mathcal{S}, \mathcal{R})$ belongs to the optimization domain of (14) we achieve the desired equality. □

Appendix F
Proof of Proposition 5

Proposition 5: Let us define a unitary matrix $U \in \mathcal{U}(\mathbb{C}^4)$ which columns form the magic basis

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ 0 & i & -1 & 0 \\ 1 & 0 & 0 & -i \end{bmatrix}. \tag{69}$$

Let us also define a diagonal operator $D(\lambda) := \sum_{i=1}^4 \lambda_i |i\rangle\langle i|$, which is parameterized by a $4-$dimensional real vector $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, for which it holds $0 < \lambda_i \leq 1$. For $R = UD(\lambda)U^\dagger$ and the noise channel $\mathcal{E}_R$ defined as

$$\mathcal{E}_R(Y) = |0\rangle\langle 0| \otimes \mathrm{tr}_1 \left(\sqrt{R} Y \sqrt{R}\right) + |1\rangle\langle 1| \otimes \mathrm{tr}\left([\mathbb{1}_{\mathbb{C}^4} - R]Y\right) \rho_2^* \tag{70}$$

we have

$$p_0(R) = \frac{4}{\mathrm{tr}(R^{-1})}, \quad p_1(R) = \frac{4}{\mathrm{tr}(R^{-1}) + c}, \tag{71}$$

where

$$c = \min\left\{ \left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|, \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right| \right\}. \tag{72}$$

Proof: First, we calculate $p_0(R)$. Let $|x\rangle = (x_0, x_1)^\top$. Then, we have

$$(\langle x| \otimes \mathbb{1}_{\mathbb{C}^2}) R^{-1} (|x\rangle \otimes \mathbb{1}_{\mathbb{C}^2}) = \frac{1}{2} \begin{bmatrix} M_{0,0} & M_{0,1} \\ \overline{M_{0,1}} & M_{1,1} \end{bmatrix}, \tag{73}$$

where

$$M_{0,0} = \frac{|x_0|^2}{\lambda_1} + \frac{|x_1|^2}{\lambda_2} + \frac{|x_1|^2}{\lambda_3} + \frac{|x_0|^2}{\lambda_4}, \quad M_{1,1} = \frac{|x_1|^2}{\lambda_1} + \frac{|x_0|^2}{\lambda_2} + \frac{|x_0|^2}{\lambda_3} + \frac{|x_1|^2}{\lambda_4}, \quad M_{0,1} = \frac{x_1 \bar{x}_0}{\lambda_1} + \frac{x_0 \bar{x}_1}{\lambda_2} - \frac{x_0 \bar{x}_1}{\lambda_3} - \frac{x_1 \bar{x}_0}{\lambda_4}. \tag{74}$$

We obtain $\mathrm{tr}\left((\langle x| \otimes \mathbb{1}_{\mathbb{C}^2}) R^{-1} (|x\rangle \otimes \mathbb{1}_{\mathbb{C}^2})\right) = \frac{1}{2}\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right)\|x\|_2^2 = \frac{1}{2}\mathrm{tr}(R^{-1})\|x\|_2^2$. Hence, for any $\rho \in \mathcal{D}(\mathbb{C}^2)$ we have $\mathrm{tr}\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right) = \frac{1}{2}\mathrm{tr}(R^{-1})$. Finally, we obtain the following upper bound

$$\|\mathrm{tr}_1\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} \leq 2\left(\mathrm{tr}\left(R^{-1}(\rho \otimes \mathbb{1}_{\mathbb{C}^2})\right)\right)^{-1} = 4\left(\mathrm{tr}(R^{-1})\right)^{-1}. \tag{75}$$

That means, $p_0(R) \leq 4\left(\mathrm{tr}(R^{-1})\right)^{-1}$. To saturate this bound, we take the maximally mixed state $\rho = \rho_2^*$ and by using (73) we calculate

$$\|\mathrm{tr}_1\left(R^{-1}(\rho_2^* \otimes \mathbb{1}_{\mathbb{C}^2})\right)\|_\infty^{-1} = 2\|\mathrm{tr}_1\left(R^{-1}\right)\|_\infty^{-1} = 2\left\|\frac{1}{2}\mathrm{tr}(R^{-1})\mathbb{1}_{\mathbb{C}^2}\right\|_\infty^{-1} = 4\left(\mathrm{tr}(R^{-1})\right)^{-1}. \tag{76}$$

Therefore, we showed that $p_0(R) = 4\left(\mathrm{tr}(R^{-1})\right)^{-1}$.

In the case of $p_1(R)$, to calculate the largest eigenvalue of $\mathrm{tr}_1\left(R^{-1}(|x\rangle\langle x| \otimes \mathbb{1}_{\mathbb{C}^2})\right)$ we use (73) for $|x\rangle = (|x_0|, |x_1|\alpha)^\top$, such that $|x_0|^2 + |x_1|^2 = 1$ and $|\alpha| = 1$. One may calculate that the largest eigenvalue minimized over α is given by $\frac{1}{4}\left(\mathrm{tr}(R^{-1}) + \left[\left(\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_4}\right) - \left(\frac{1}{\lambda_2} + \frac{1}{\lambda_3}\right)\right)^2 (|x_0|^2 - |x_1|^2)^2 + 4\left(\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right)^2 |x_0|^2|x_1|^2\right]^{-1}\right)$. It turns out, there are only two situations when this expression is minimized:

- For $|x_0| = 0$ and $|x_1| = 1$ (or equivalently $|x_0| = 1$ and $|x_1| = 0$), we obtain

$$\frac{1}{4}\left(\mathrm{tr}(R^{-1}) + \left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|\right). \tag{77}$$

- For $|x_0| = |x_1| = \frac{1}{\sqrt{2}}$, we obtain

$$\frac{1}{4}\left(\mathrm{tr}(R^{-1}) + \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right|\right). \tag{78}$$

Hence, the optimal value $p_1(R)$ equals $p_1(R) = \frac{4}{\mathrm{tr}(R^{-1}) + \min\left\{\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2} - \frac{1}{\lambda_3} + \frac{1}{\lambda_4}\right|, \left|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_4}\right| - \left|\frac{1}{\lambda_2} - \frac{1}{\lambda_3}\right|\right|\right\}}$. □

Appendix G
Proof of Proposition 6

Proposition 6: For any $\mathcal{X}, \mathcal{Y}$ and $\xi(\mathcal{X}, \mathcal{Y}), \xi_1(\mathcal{X}, \mathcal{Y})$ defined in (23) we have the following properties:

(A) $\xi_1(\mathcal{X}, \mathcal{Y}) \subset \xi(\mathcal{X}, \mathcal{Y})$,
(B) If $\dim(\mathcal{X}) > \dim(\mathcal{Y})$, then $\xi(\mathcal{X}, \mathcal{Y}) = \emptyset$,
(C) If $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$, then $\xi_1(\mathcal{X}, \mathcal{Y}) \neq \emptyset$,
(D) If $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, then $\xi_1(\mathcal{X}, \mathcal{Y}) = \xi(\mathcal{X}, \mathcal{Y})$.

Proof: (D)
Let us take $\mathcal{E} = \mathcal{K}((E_i)_i) \in \xi(\mathcal{X}, \mathcal{Y})$. From Theorem 1 (D) there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$ such that $R_* E_i S_* \propto \mathbb{1}_\mathcal{X}$, and there exists i_0 for which it holds $R_* E_{i_0} S_* \neq 0$. It implies that R_* and S_* are invertible, so for all i we have $E_i \propto R_*^{-1} S_*^{-1}$. Hence, $\mathrm{rank}(J(\mathcal{E})) = 1$, so we can write $\mathcal{E}(X) = EXE^\dagger$, for $E \in \mathcal{U}(\mathcal{X})$. By taking $\mathcal{R} = \mathcal{I}_\mathcal{X}$ and $\mathcal{S} = \mathcal{E}^\dagger$ we get $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$.

□

Appendix H
Proof of Theorem 7

Theorem 7: Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces for which $\dim(\mathcal{X}) < \dim(\mathcal{Y})$ and let $\xi(\mathcal{X}, \mathcal{Y}), \xi_1(\mathcal{X}, \mathcal{Y})$ be defined as in (23). Then, the set $\xi_1(\mathcal{X}, \mathcal{Y})$ is a nowhere dense subset of $\xi(\mathcal{X}, \mathcal{Y})$.

Proof: First, we will prove that $\xi_1(\mathcal{X}, \mathcal{Y})$ is a closed set. Define a sequence $(\mathcal{E}_n)_{n\in\mathbb{N}} \subset \xi_1(\mathcal{X}, \mathcal{Y})$ that converges to $\mathcal{E} = \lim_{n\to\infty} \mathcal{E}_n \in \mathcal{C}(\mathcal{Y})$. From Proposition 2 there exist two sequences $(\mathcal{S}_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $(\mathcal{R}_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\mathcal{Y}, \mathcal{X})$ such that $\mathcal{R}_n \mathcal{E}_n \mathcal{S}_n = \mathcal{I}_\mathcal{X}$ for $n \in \mathbb{N}$. Both sets $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ and $\mathcal{C}(\mathcal{Y}, \mathcal{X})$ are compact, so there exists a subsequence $(n_k)_{k\in\mathbb{N}}$, such that $(\mathcal{S}_{n_k})_{k\in\mathbb{N}}$, $(\mathcal{R}_{n_k})_{k\in\mathbb{N}}$ converge to some $\mathcal{S} \in \mathcal{C}(\mathcal{X}, \mathcal{Y}), \mathcal{R} \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$, respectively. Hence, we obtain $\mathcal{RES} = \lim_{k\to\infty} \mathcal{R}_{n_k} \mathcal{E}_{n_k} \mathcal{S}_{n_k} = \mathcal{I}_\mathcal{X}$. That ends this part of the proof.

To show that $\xi_1(\mathcal{X}, \mathcal{Y})$ is nowhere dense in $\xi(\mathcal{X}, \mathcal{Y})$, it is enough to prove $\mathrm{int}\left(\xi_1(\mathcal{X}, \mathcal{Y})\right) = \emptyset$. Therefore, for any $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$ we will construct a sequence of channels $(\mathcal{E}_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\mathcal{Y})$ that converges to $\mathcal{E}$ and for which $\mathcal{E}_n \in \xi(\mathcal{X}, \mathcal{Y})$, and $\mathcal{E}_n \notin \xi_1(\mathcal{X}, \mathcal{Y})$, for $n \in \mathbb{N}$.

Fix $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$. From Proposition 2 there exist $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\mathcal{R} \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = \mathcal{I}_\mathcal{X}$. From Theorem 1 we have

$$\mathcal{E}\mathcal{S} = \mathcal{K}((A_i)_i): \quad A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij} \mathbb{1}_\mathcal{X}. \tag{79}$$

As $\dim(\mathcal{X}) < \dim(\mathcal{Y})$, there exists $|y\rangle\langle y| \in \mathcal{D}(\mathcal{Y})$ such that $\langle y|A_1 = 0$. Let us define a sequence of channels $\mathcal{E}_n \in \mathcal{C}(\mathcal{Y})$ given by

$$\mathcal{E}_n(Y) = \frac{n}{n+1}\mathcal{E}(Y) + \frac{\operatorname{tr}(Y)}{n+1}|y\rangle\langle y|. \tag{80}$$

One can note that $\lim_{n\to\infty} \mathcal{E}_n = \mathcal{E}$. We take $\mathcal{S}_n = \mathcal{S}$ and $\mathcal{R}_n = \mathcal{K}\left((A_1^\dagger)\right)$ for $n \in \mathbb{N}$ and obtain

$$\mathcal{R}_n\mathcal{E}_n\mathcal{S}_n(X) = \frac{n}{n+1}A_1^\dagger \mathcal{E}\mathcal{S}(X)A_1 = \frac{n}{n+1}\|A_1\|_\infty^4 X. \tag{81}$$

As $A_1 \neq 0$, it follows that $\mathcal{E}_n \in \xi(\mathcal{X},\mathcal{Y})$. Now, for each $n \in \mathbb{N}$, let $\widetilde{\mathcal{S}}_n \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\widetilde{\mathcal{R}}_n \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ be arbitrary maps satisfying $0 \neq \widetilde{\mathcal{R}}_n\mathcal{E}_n\widetilde{\mathcal{S}}_n \propto \mathcal{I}_\mathcal{X}$. It holds that $\widetilde{\mathcal{R}}_n(|y\rangle\langle y|) = 0$. Finally, for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ we have $\operatorname{tr}\left(\widetilde{\mathcal{R}}_n\mathcal{E}_n\widetilde{\mathcal{S}}_n(|\psi\rangle\langle\psi|)\right) = \frac{n}{n+1}\operatorname{tr}\left(\widetilde{\mathcal{R}}_n\mathcal{E}\widetilde{\mathcal{S}}_n(|\psi\rangle\langle\psi|)\right) \leq \frac{n}{n+1}$. Hence, we obtain $\mathcal{E}_n \notin \xi_1(\mathcal{X},\mathcal{Y})$. □

Appendix I
Proof of Theorem 8

Theorem 8: Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. The following relations hold:

$$\begin{aligned} &\max\{\operatorname{rank}(J(\mathcal{E})) : \mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})\} \\ &= \dim(\mathcal{Y})^2 - \dim(\mathcal{Y})\dim(\mathcal{X}) + \left\lfloor \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})} \right\rfloor, \\ &\max\{\operatorname{rank}(J(\mathcal{E})) : \mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})\} \\ &= \dim(\mathcal{Y})^2 - \dim(\mathcal{X})^2 + 1. \end{aligned} \tag{82}$$

Proof: Let us define $d = \dim(\mathcal{X})$, $s = \dim(\mathcal{Y})$ and $k = \left\lfloor \frac{s}{d} \right\rfloor$.

Take $\mathcal{E} = \mathcal{K}((E_i)_{i=1}^r) \in \xi_1(\mathcal{X},\mathcal{Y})$, where $r = \operatorname{rank}(J(\mathcal{E}))$. From Proposition 2 there exist $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\mathcal{R} \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = \mathcal{I}_\mathcal{X}$. According to Theorem 1 it holds

$$\mathcal{K}((E_iS)_{i=1}^r) = \mathcal{K}\left((A_i)_{i=1}^{r'}\right): \quad A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_\mathcal{X}. \tag{83}$$

If $r' < r$, then let us define $A_i = 0$ for $i = r'+1,\ldots,r$. As $\mathcal{K}\left((A_i)_{i=1}^{r'}\right) = \mathcal{K}((A_i)_{i=1}^r)$, there exists a Kraus decomposition $\mathcal{E} = \mathcal{K}((E_i')_{i=1}^r)$ such that $A_i = E_i'S$ for each $i \leq r$. For $A_i \neq 0$ images of A_i are orthogonal and $\operatorname{rank}(A_i) = d$. Hence, $r'd \leq s$ which is equivalent to $r' \leq k$. For $i > r'$ it holds that $\left(\mathbb{1}_\mathcal{Y} \otimes S^\top\right)|E_i'\rangle = 0$. Note that the Kraus operators E_i' are linearly independent and it holds

$$\begin{aligned} \dim(\ker(\mathbb{1}_\mathcal{Y} \otimes S^\top)) &= s^2 - \operatorname{rank}(\mathbb{1}_\mathcal{Y} \otimes S^\top) \\ &= s^2 - \operatorname{rank}(\mathbb{1}_\mathcal{Y})\operatorname{rank}(S) = s^2 - sd. \end{aligned} \tag{84}$$

Therefore, we get $r - r' = \dim(\operatorname{span}(E_i', i > r')) \leq \dim(\ker(\mathbb{1}_\mathcal{Y} \otimes S^\top)) = s^2 - sd$ and eventually $r \leq s^2 - sd + k$. To saturate this bound, let us define $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ given by

$$\mathcal{E}(Y) = \sum_{i=0}^{k-1} E_i Y E_i^\dagger + \operatorname{tr}((\mathbb{1}_\mathcal{Y} - \Pi)Y)\,\rho_\mathcal{Y}^*, \tag{85}$$

where

$$\begin{aligned} E_i &= \frac{1}{\sqrt{k}}\sum_{j=0}^{d-1} |j + id\rangle\langle j| \in \mathcal{M}(\mathcal{Y}), \qquad \text{for } i = 0,\ldots,k-1, \\ \Pi &= \sum_{j=0}^{d-1} |j\rangle\langle j| \in \mathcal{P}(\mathcal{Y}). \end{aligned} \tag{86}$$

Note that $\Pi = \sum_{i=0}^{k-1} E_i^\dagger E_i$ and $(\mathbb{1}_\mathcal{Y} \otimes \Pi)|E_i\rangle = |E_i\rangle$. Therefore, we obtain

$$\begin{aligned} \operatorname{rank}(J(\mathcal{E})) &= \operatorname{rank}\left(\sum_{i=0}^{k-1} |E_i\rangle\langle E_i| + \rho_\mathcal{Y}^* \otimes (\mathbb{1}_\mathcal{Y} - \Pi)\right) \\ &= \operatorname{rank}\left(\sum_{i=0}^{k-1} |E_i\rangle\langle E_i|\right) + \operatorname{rank}\left(\rho_\mathcal{Y}^* \otimes (\mathbb{1}_\mathcal{Y} - \Pi)\right) \\ &= s^2 - sd + k. \end{aligned} \tag{87}$$

Finally, let us define $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$, where $S = \sum_{j=0}^{d-1} |j\rangle_\mathcal{Y}\langle j|_\mathcal{X}$, and $\mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ given by $\mathcal{R}(Y) = kS^\dagger\left(\sum_{i=0}^{k-1} E_i^\dagger Y E_i\right)S$. We can observe that $\mathcal{R}\mathcal{E}\mathcal{S} = \mathcal{I}_\mathcal{X}$, so $\mathcal{E} \in \xi_1(\mathcal{X},\mathcal{Y})$.

Now, take $\mathcal{E} = \mathcal{K}((E_i)_{i=1}^r) \in \xi(\mathcal{X},\mathcal{Y})$, where $r = \operatorname{rank}(J(\mathcal{E}))$. According to Theorem 1 (D) there exist $S_* \in \mathcal{M}(\mathcal{X},\mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y},\mathcal{X})$ such that $R_*E_iS_* \propto \mathbb{1}_\mathcal{X}$, and there exists i_0 for which it holds $R_*E_{i_0}S_* \neq 0$. We may assume that $\|R_*\|_\infty \leq 1$ and $\|S_*\|_\infty \leq 1$. Hence, according to Theorem 1 (B) we get

$$\mathcal{K}\left(\left(\sqrt{R_*^\dagger R_*}E_iS_*\right)_{i=1}^r\right) = \mathcal{K}\left((A_i)_{i=1}^{r'}\right) \tag{88}$$

for $A_i \neq 0, A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_\mathcal{X}$. If $r' < r$, then let us define $A_i = 0$ for $i = r'+1,\ldots,r$. As $\mathcal{K}\left((A_i)_{i=1}^{r'}\right) = \mathcal{K}((A_i)_{i=1}^r)$, there exists a Kraus decomposition $\mathcal{E} = \mathcal{K}((E_i')_{i=1}^r)$ such that $A_i = \sqrt{R_*^\dagger R_*}E_i'S_*$ for each $i \leq r$. Let Π be the projector on the support of $R_*^\dagger R_*$. Observe that $\operatorname{rank}(\Pi) = d$. Then, for each $i \leq r$ we have $\Pi A_i = A_i$ and for $i \leq r'$ we have $\operatorname{rank}(A_i) = d$. The relation $A_j^\dagger A_i \propto \delta_{ij}\mathbb{1}_\mathcal{X}$ implies that there exists exactly one $A_i \neq 0$, hence $r' = 1$. For $i > 1$ we have $\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right)|E_i'\rangle = 0$. Note that the Kraus operators E_i' are linearly independent and it holds

$$\begin{aligned} \dim\left(\ker\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right)\right) &= s^2 - \operatorname{rank}\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right) \\ &= s^2 - d^2. \end{aligned} \tag{89}$$

Therefore, we obtain $r - 1 = \dim(\operatorname{span}(E_i', i > 1)) \leq \dim\left(\ker\left(\sqrt{R_*^\dagger R_*} \otimes S_*^\top\right)\right) = s^2 - d^2$ and eventually

$r \leq s^2 - d^2 + 1$. To saturate this bound, we define $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ given by

$$\mathcal{E}(Y) = \frac{\Pi Y \Pi + \mathrm{tr}\,(\Pi Y)\,(\mathbb{1}_{\mathcal{Y}} - \Pi)}{s-d+1} + \mathrm{tr}\,((\mathbb{1}_{\mathcal{Y}} - \Pi)Y)\,\rho^*_{\mathcal{Y}}, \tag{90}$$

where $\Pi = \sum_{j=0}^{d-1} |j\rangle\langle j| \in \mathcal{P}(\mathcal{Y})$. Note, that

$$\begin{aligned}&\mathrm{rank}(J(\mathcal{E}))\\ =&\mathrm{rank}\left(\frac{|\Pi\rangle\langle\Pi| + (\mathbb{1}_{\mathcal{Y}} - \Pi)\otimes\Pi}{s-d+1} + \rho^*_{\mathcal{Y}} \otimes (\mathbb{1}_{\mathcal{Y}} - \Pi)\right)\\ =&\mathrm{rank}(|\Pi\rangle\langle\Pi|) + \mathrm{rank}((\mathbb{1}_{\mathcal{Y}} - \Pi)\otimes\Pi)\\ +&\mathrm{rank}(\rho^*_{\mathcal{Y}} \otimes (\mathbb{1}_{\mathcal{Y}} - \Pi)) = s^2 - d^2 + 1.\end{aligned} \tag{91}$$

Define $\mathcal{S} = \mathcal{K}((S)) \in \mathcal{C}(\mathcal{X},\mathcal{Y})$, where $S = \sum_{j=0}^{d-1} |j\rangle_{\mathcal{Y}}\langle j|_{\mathcal{X}}$ and $\mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$ given by $\mathcal{R}(Y) = S^\dagger Y S$. We can observe that $\mathcal{RES} = \frac{\mathcal{I}_{\mathcal{X}}}{s-d+1}$, so $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$. □

Appendix J
Proof of Lemma 10

Lemma 10: Let $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. Then, it holds

$$\begin{aligned}&p_{\mathcal{X}}(\mathcal{F})\\ =&\max\{\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1\,(R_F(P\otimes\mathbb{1}_{\mathcal{X}})) \leq \mathbb{1}_{\mathcal{X}},\\ &(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_F\otimes\mathbb{1}_{\mathcal{X}}) = P\otimes|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|\},\end{aligned} \tag{92}$$

where $R_F = (FF^\dagger)^{-1}, \Pi_F = FF^{-1}$ for $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Moreover, if $p_{\mathcal{X}}(\mathcal{F}) > 0$, then

$$\|R_F\|_\infty^{-1} \leq p_{\mathcal{X}}(\mathcal{F}) \leq \|R_F^{-1}\|_\infty. \tag{93}$$

Proof: This proof is an extension of the proof of Lemma 3 presented in Appendix D. Let us fix $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{Y},\mathcal{X})$. Assume that for some $\widetilde{\mathcal{S}} \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$ and $\widetilde{\mathcal{R}} \in s\mathcal{C}(\mathcal{X})$ it holds $\widetilde{\mathcal{R}}\mathcal{F}\widetilde{\mathcal{S}} = p\mathcal{I}_{\mathcal{X}} \neq 0$. Then, we know that $\widetilde{\mathcal{R}} = \mathcal{K}\left((\widetilde{R})\right)$ and $\mathcal{F}\widetilde{\mathcal{S}}\widetilde{\mathcal{R}} = p\mathcal{I}_{\mathcal{X}}$. Therefore, $p_{\mathcal{X}}(\mathcal{F}) = \max\{p : \mathcal{FS} = p\mathcal{I}_{\mathcal{X}}, \mathcal{S} \in s\mathcal{C}(\mathcal{X},\mathcal{Y})\}$. Define

$$\begin{aligned}F &= \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X}),\\ \Pi_F &= FF^{-1},\\ R_F &= (FF^\dagger)^{-1}.\end{aligned} \tag{94}$$

We can write $\mathcal{F}(Y) = \mathrm{tr}_1(FYF^\dagger)$. If for some $\mathcal{S} \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$ it holds $\mathrm{tr}_1(F\mathcal{S}(X)F^\dagger) = pX$, then according to Theorem 1 we have $F\mathcal{S}(X)F^\dagger = P \otimes X$, where $P \in \mathcal{P}(\mathbb{C}^r)$. $\mathcal{S}$ which satisfies this equation is equal to $\mathcal{S}(X) = F^{-1}(P\otimes X)(F^{-1})^\dagger$. Moreover it holds $\Pi_F(P\otimes X)\Pi_F = P \otimes X$ for all X and $p = \mathrm{tr}(P)$. Eventually, we obtain

$$\begin{aligned}&p_{\mathcal{X}}(\mathcal{F})\\ =&\max\{p : \mathcal{FS} = p\mathcal{I}_{\mathcal{X}}, \mathcal{S} \in s\mathcal{C}(\mathcal{X},\mathcal{Y})\}\\ =&\max\{\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r),\\ &\mathcal{S}(\cdot) = F^{-1}(P\otimes \cdot)(F^{-1})^\dagger \in s\mathcal{C}(\mathcal{X},\mathcal{Y}),\\ &\forall_X\ \Pi_F(P\otimes X)\Pi_F = P\otimes X\}\\ =&\max\{\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1\,(R_F(P\otimes\mathbb{1}_{\mathcal{X}})) \leq \mathbb{1}_{\mathcal{X}},\\ &(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_F\otimes\mathbb{1}_{\mathcal{X}}) = P\otimes|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|\}.\end{aligned} \tag{95}$$

Assume now that $p_{\mathcal{X}}(\mathcal{F}) > 0$, that is, we can find $0 \neq P \in \mathcal{P}(\mathbb{C}^r)$ satisfying $(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_F\otimes\mathbb{1}_{\mathcal{X}}) = P\otimes|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|$ and $\mathrm{tr}_1\,(R_F(P\otimes\mathbb{1}_{\mathcal{X}})) \leq \mathbb{1}_{\mathcal{X}}$. Let $P = p\rho$ for $\rho \in \mathcal{D}(\mathbb{C}^r)$ and define $\widetilde{P} = \|\mathrm{tr}_1\,(R_F(\rho\otimes\mathbb{1}_{\mathcal{X}}))\|_\infty^{-1}\,\rho$. Observe that $\widetilde{P}$ also belongs to the optimization domain of $p_{\mathcal{X}}(\mathcal{F})$. Hence, we get

$$p_{\mathcal{X}}(\mathcal{F}) \geq \|\mathrm{tr}_1\,(R_F(\rho\otimes\mathbb{1}_{\mathcal{X}}))\|_\infty^{-1} \geq \|R_F\|_\infty^{-1}. \tag{96}$$

On the other hand, it holds $\|R_F^{-1}\|_\infty^{-1}\Pi_F \leq R_F$. Hence, for any P which belongs to the optimization domain of $p_{\mathcal{X}}(\mathcal{F})$ it holds

$$\|R_F^{-1}\|_\infty^{-1}\,\mathrm{tr}(P)\mathbb{1}_{\mathcal{X}} \leq \mathrm{tr}_1\,(R_F(P\otimes\mathbb{1}_{\mathcal{X}})) \leq \mathbb{1}_{\mathcal{X}}. \tag{97}$$

It implies that $p_{\mathcal{X}}(\mathcal{F}) \leq \|R_F^{-1}\|_\infty$. □

Appendix K
Proof of Corollary 11

Corollary 11: Let $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$. Then, it holds

$$\begin{aligned}p_{\mathcal{X}}(\mathcal{F}) = \max\Big\{&\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r), P\otimes\mathbb{1}_{\mathcal{X}} \leq \widetilde{F}\widetilde{F}^\dagger,\\ &(\Pi_{\widetilde{F}} \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_{\widetilde{F}}\otimes\mathbb{1}_{\mathcal{X}}) = P\otimes|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|\Big\},\end{aligned} \tag{98}$$

where $\Pi_{\widetilde{F}} = \widetilde{F}\widetilde{F}^{-1}$ for $\widetilde{F} = \sum_{i=0}^{r-1} |i\rangle \otimes F_i^\dagger \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$.

Proof: This proof is based on the proof of Lemma 10 presented in Appendix J. Let us define $\widetilde{F} = \sum_{i=0}^{r-1} |i\rangle \otimes F_i^\dagger \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$ and $\Pi_{\widetilde{F}} = \widetilde{F}\widetilde{F}^{-1}$ for a given $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in s\mathcal{C}(\mathcal{X},\mathcal{Y})$. One may note that $\mathcal{RF} = p\mathcal{I}_{\mathcal{X}}$ if and only if $\mathcal{F}^\dagger\mathcal{R}^\dagger = p\mathcal{I}_{\mathcal{X}}$. Therefore, we obtain

$$\begin{aligned}&p_{\mathcal{X}}(\mathcal{F}) = \max\{p : \mathcal{F}^\dagger\mathcal{R}^\dagger = p\mathcal{I}_{\mathcal{X}}, \mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X})\}\\ &=\max\Big\{\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r), \mathcal{R}^\dagger(\cdot) = \widetilde{F}^{-1}(P\otimes\cdot)(\widetilde{F}^{-1})^\dagger,\\ &\quad\mathcal{R} \in s\mathcal{C}(\mathcal{Y},\mathcal{X}),\ \forall_X\ \Pi_{\widetilde{F}}(P\otimes X)\Pi_{\widetilde{F}} = P\otimes X\Big\}\\ &=\max\Big\{\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r), \widetilde{F}^{-1}(P\otimes\mathbb{1}_{\mathcal{X}})(\widetilde{F}^{-1})^\dagger \leq \mathbb{1}_{\mathcal{Y}},\\ &\quad\forall_X\ \Pi_{\widetilde{F}}(P\otimes X)\Pi_{\widetilde{F}} = P\otimes X\Big\}\\ &=\max\Big\{\mathrm{tr}(P):\ P \in \mathcal{P}(\mathbb{C}^r), P\otimes\mathbb{1}_{\mathcal{X}} \leq \widetilde{F}\widetilde{F}^\dagger,\\ &\quad(\Pi_{\widetilde{F}} \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_{\widetilde{F}}\otimes\mathbb{1}_{\mathcal{X}}) = P\otimes|\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|\Big\}.\end{aligned} \tag{99}$$

□

Appendix L
Proof of Corollary 12

Corollary 12: Let $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y},\mathcal{X})$. Define $\Pi_F = FF^{-1}$, where $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Then, it holds $p_{\mathcal{X}}(\mathcal{F}) \in \{0,1\}$. Moreover, $\mathcal{F}$ is perfectly correctable for $\mathcal{X}$ if and only if there exists $0 \neq |\psi\rangle \in \mathbb{C}^r$ such that $(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(|\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle) = |\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle$.

Proof: Let us assume that for a given $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y},\mathcal{X})$ there exists error-correcting scheme $(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{X})$ such that $\mathcal{RFS} = p\mathcal{I}_{\mathcal{X}} \neq 0$.

From the proof of Lemma 10, without loss of the generality we may take $\mathcal{R} = \mathcal{I}_{\mathcal{X}}$. Hence, from Proposition 2 we have $p_{\mathcal{X}}(\mathcal{F}) = 1$. Now, from Lemma 10 we know that $p_{\mathcal{X}}(\mathcal{F}) > 0$ if and only if there exists $0 \neq P \in \mathcal{P}(\mathbb{C}^r)$ such that $(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|)(\Pi_F \otimes \mathbb{1}_{\mathcal{X}}) = P \otimes |\mathbb{1}_{\mathcal{X}}\rangle\langle\mathbb{1}_{\mathcal{X}}|$. This condition is equivalent to $(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(|\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle) = |\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle$ for some $0 \neq |\psi\rangle \in \mathbb{C}^r$. □

Appendix M
Proof of Proposition 13

Proposition 13: Let $\mathcal{X}$ and $\mathcal{Y}$ be some complex Euclidean spaces and $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$.

(A) If $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ is a noise channel such that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_{\mathcal{Y}})) = \dim(\mathcal{X})$ and $\operatorname{rank}(J(\mathcal{E})) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$, then $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$.

(B) There exists a noise channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_{\mathcal{Y}})) = \dim(\mathcal{X})$ and $\operatorname{rank}(J(\mathcal{E})) \geq \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$, for which we have $\mathcal{E} \notin \xi(\mathcal{X}, \mathcal{Y})$.

Proof: (A)

Let us take $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and denote $r = \operatorname{rank}(J(\mathcal{E}))$. Assume that $\operatorname{rank}(\mathcal{E}(\mathbb{1}_{\mathcal{Y}})) = \dim(\mathcal{X})$ and $r < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$. Consider an associated to $\mathcal{E}$ channel $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$. Define $\Pi_F = FF^{-1}$, where $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Observe that $\dim(\ker((\mathbb{1}_{\mathbb{C}^r \otimes \mathcal{X}} - \Pi_F) \otimes \mathbb{1}_{\mathcal{X}})) = \dim(\mathcal{Y})\dim(\mathcal{X})$ and $\dim(\operatorname{span}(|\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle : |\psi\rangle \in \mathbb{C}^r)) = r$. Therefore, as $\dim(\mathcal{Y})\dim(\mathcal{X}) + r > r\dim(\mathcal{X})^2$ there exists $0 \neq |\psi\rangle \in \mathbb{C}^r$, such that $(\Pi_F \otimes \mathbb{1}_{\mathcal{X}})(|\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle) = |\psi\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle$. It follows from Corollary 12 that $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$.

(B)

Let us take $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y}, \mathcal{X})$ defined as in the part (A) of the proof. We have that $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if there exists $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ such that $F_i S = c_i \mathbb{1}_{\mathcal{X}}$ and $c_{i_0} \neq 0$ for some i_0. Let $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$ and $|c\rangle = \sum_{i=0}^{r-1} c_i |i\rangle$. Hence, $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if it holds $FS = |c\rangle \otimes \mathbb{1}_{\mathcal{X}} \neq 0$. This is equivalent to

$$\mathcal{E} \notin \xi(\mathcal{X}, \mathcal{Y}) \iff ((F \otimes \mathbb{1}_{\mathcal{X}})|S\rangle = |c\rangle \otimes |\mathbb{1}_{\mathcal{X}}\rangle \implies |S\rangle = 0). \tag{100}$$

Therefore, in this proof, we will construct appropriate operator F. Formally, the operator F should be an isometry operator, but by Lemma 9, it is enough to define F such that $\operatorname{rank}(F) = \dim(\mathcal{Y})$.

Let $d = \dim(\mathcal{X}), s = \dim(\mathcal{Y})$ and fix $r \in \mathbb{N}$, such that $r \geq \frac{sd}{d^2-1}$. We start with the case $s = kd$ for $k \in \mathbb{N}$. Consider the decomposition $F = \sum_{i=0}^{r-1} |i\rangle \otimes F_i$, where $F_i \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$. For $i = 0, \ldots, k-1$ we define

$$F_i = \langle i| \otimes \mathbb{1}_{\mathcal{X}}. \tag{101}$$

Let $\left\{\mathbb{1}_{\mathcal{X}}, (M_j)_{j=0}^{d^2-2}\right\} \subset \mathcal{M}(\mathcal{X})$ be a basis of $\mathcal{M}(\mathcal{X})$. For each $i = k, \ldots, r-1$ we define

$$F_i = \sum_{j=0}^{d^2-2} \delta(j + (i-k)(d^2-1) < k) \langle j + (i-k)(d^2-1)| \otimes M_j. \tag{102}$$

Observe, that $\operatorname{rank}(F) = s$. Let us take S which satisfies $F_i S \propto \mathbb{1}_{\mathcal{X}}$ for each i. Basing on the equations with indices $i = 0, \ldots, k-1$ we get $S = |c\rangle \otimes \mathbb{1}_{\mathcal{X}}$ for some $|c\rangle = \sum_{j=0}^{k-1} c_j |j\rangle$. Note, that if for any $i = k, \ldots, r-1$ it holds $F_i S \propto \mathbb{1}_{\mathcal{X}}$, then $c_j = 0$ for each $j = (i-k)(d^2-1), \ldots, d^2-2+(i-k)(d^2-1)$. From the assumption $r \geq \frac{sd}{d^2-1}$ we have $(r-k)(d^2-1) \geq k$, hence, all entries c_j are zeroed. It implies $S = 0$.

The case $s = kd + l$ for $l = 1, \ldots, d-1$ is more technically engaging than the previous case but it is based on the same idea. It will be only briefly discussed. For $i = 0, \ldots, k-1$ we can define F_i similarly as in the previous case, that is $F_i \sim \langle i| \otimes \mathbb{1}_{\mathcal{X}}$. The operator F_k has a special form, $F_k \sim (\langle k| \otimes \sum_{j=0}^{l-1} |j\rangle\langle j|) + N$, where the image of N is contained in $\operatorname{span}(|j\rangle : j \geq l)$. Here, the operator S which satisfy $F_i S \propto \mathbb{1}_{\mathcal{X}}$ has the form $S \sim |c\rangle \otimes \mathbb{1}_{\mathcal{X}}$ for some $|c\rangle = \sum_{j=0}^{k} c_j |j\rangle$. We can choose N such that $d(d-l)$ entries c_j will be zeroed if $F_k S \propto \mathbb{1}_{\mathcal{X}}$. Finally, operators F_i for $i = k+1, \ldots, r-1$ has the analogous form as (102) – each nullify (d^2-1) entries. In total, the number of entries c_j which can be zeroed is not less than $k+1$. Indeed, it holds

$$d(d-l) + (r-k-1)(d^2-1) \geq k+1. \tag{103}$$

Therefore, $S = 0$, which ends the proof. □

Appendix N
Proof of Lemma 14

Lemma 14: Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces such that $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. Then, there exists a Schur channel $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(J(\mathcal{E})) = \left\lceil \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1} \right\rceil$ and $\mathcal{E} \notin \xi(\mathcal{X}, \mathcal{Y})$. Moreover, there exists a Schur channel $\mathcal{F} \in \mathcal{C}(\mathcal{Y})$ such that $\operatorname{rank}(J(\mathcal{F})) = \left\lceil \sqrt{\left\lceil \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1} \right\rceil} \right\rceil$ and $\mathcal{F} \notin \xi_1(\mathcal{X}, \mathcal{Y})$. This implies

$$r(\mathcal{X}, \mathcal{Y}) < \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1},$$
$$r_1(\mathcal{X}, \mathcal{Y}) < \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}}. \tag{104}$$

Proof: Let $d = \dim(\mathcal{X})$, $s = \dim(\mathcal{Y})$ and $s = k(d-1) - w$, where $k = \left\lceil \frac{s}{d-1} \right\rceil$ and $w \in \{0, \ldots, d-2\}$. First, we will show that $r(\mathcal{X}, \mathcal{Y}) < k$. Define a Schur channel $\mathcal{E} = \mathcal{K}\left((E_i)_{i=0}^{k-1}\right) \in \mathcal{C}(\mathcal{Y})$ given by

$$E_i = \sum_{j=0}^{d-2} |j + (d-1)i\rangle\langle j + (d-1)i|, \quad i = 0, \ldots, k-2,$$
$$E_{k-1} = \sum_{j=0}^{d-2-w} |j + (d-1)(k-1)\rangle\langle j + (d-1)(k-1)|. \tag{105}$$

Observe that $\operatorname{rank}(J(\mathcal{E})) = k$. From Theorem 1 (D) we know that $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$ if and only if there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$, such that $R_* E_i S_* \propto \mathbb{1}_{\mathcal{X}}$ for all i and there exists i_0 for which it holds $R_* E_{i_0} S_* \neq 0$. As $\operatorname{rank}(E_i) \leq d-1$, if we have $R_* E_i S_* \propto \mathbb{1}_{\mathcal{X}}$, then $R_* E_i S_* = 0$ for all i. That implies $\mathcal{E} \notin \xi(\mathcal{X}, \mathcal{Y})$.

Now, let us define $l = \left\lceil \sqrt{k} \right\rceil$. We will prove that $r_1(\mathcal{X}, \mathcal{Y}) < l$. Due to the relation $\operatorname{span}_{\mathbb{C}}(|\psi\rangle\langle\psi| : |\psi\rangle\langle\psi| \in \mathcal{D}(\mathbb{C}^l)) = \mathcal{M}(\mathbb{C}^l)$, we may define unit vectors $|\psi_a\rangle$, for $a = 0, \ldots, l^2 - 1$, such that $\operatorname{span}_{\mathbb{C}}(\{|\psi_a\rangle\langle\psi_a|\}_a) = \mathcal{M}(\mathbb{C}^l)$. Let us define $F_i \in \mathcal{M}(\mathcal{Y})$ for $i = 0, \ldots, l-1$ given by

$$F_i = \sum_{a=0}^{k-1} \langle\psi_a|i\rangle E_a, \tag{106}$$

for E_a defined in (105). Observe that F_i are linearly independent. We have that

$$\begin{aligned}\sum_{i=0}^{l-1} F_i^\dagger F_i &= \sum_{i=0}^{l-1} \sum_{a,b=0}^{k-1} \langle i|\psi_b\rangle\langle\psi_a|i\rangle E_b^\dagger E_a \\ &= \sum_{i=0}^{l-1} \sum_{a=0}^{k-1} \langle i|\psi_a\rangle\langle\psi_a|i\rangle E_a = \sum_{a=0}^{k-1} E_a = \mathbb{1}_{\mathcal{Y}}.\end{aligned} \tag{107}$$

Now, we introduce a Schur channel $\mathcal{F} = \mathcal{K}\left((F_i)_{i=0}^{l-1}\right) \in \mathcal{C}(\mathcal{Y})$. Assume indirectly that $\mathcal{F} \in \xi_1(\mathcal{X}, \mathcal{Y})$. Then, according to Proposition 2 and Theorem 1 there exists $S \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$, which satisfies $S^\dagger S = \mathbb{1}_{\mathcal{X}}$ and $M \in \mathcal{M}(\mathbb{C}^l)$, such that $S^\dagger F_j^\dagger F_i S = M_{ji}\mathbb{1}_{\mathcal{X}}$. Therefore, we get

$$\begin{aligned}M \otimes \mathbb{1}_{\mathcal{X}} &= \sum_{j,i} |j\rangle\langle i| \otimes S^\dagger F_j^\dagger F_i S \\ =&(\mathbb{1} \otimes S^\dagger) \sum_{j,i} \left(|j\rangle\langle i| \otimes \sum_{a=0}^{k-1} \langle j|\psi_a\rangle\langle\psi_a|i\rangle E_a \right) (\mathbb{1} \otimes S) \\ =& \sum_{a=0}^{k-1} |\psi_a\rangle\langle\psi_a| \otimes S^\dagger E_a S.\end{aligned} \tag{108}$$

For each $a = 0, \ldots, k-1$ we can use Gram-Schmidt orthogonalization to define X_a, such that $\operatorname{tr}(X_a|\psi_a\rangle\langle\psi_a|) \neq 0$ and $\operatorname{tr}(X_a|\psi_b\rangle\langle\psi_b|) = 0$ whenever $a \neq b$. Hence, we obtain $\operatorname{tr}(X_a M)\mathbb{1}_{\mathcal{X}} = \operatorname{tr}(X_a|\psi_a\rangle\langle\psi_a|) S^\dagger E_a S$. As $\operatorname{rank}(E_a) \leq d-1$ we get $S^\dagger E_a S = 0$ for all a. It implies that $0 = \sum_{a=0}^{k-1} S^\dagger E_a S = S^\dagger S = \mathbb{1}_{\mathcal{X}}$, which gives the contradiction. That means $\mathcal{F} \notin \xi_1(\mathcal{X}, \mathcal{Y})$. It is enough to observe that $r_1(\mathcal{X}, \mathcal{Y}) < \operatorname{rank}(J(\mathcal{F})) = l$. $\square$

Appendix O
Proof of Proposition 15

Proposition 15: Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces and $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$. For any Schur channels $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$, such that $\operatorname{rank}(J(\mathcal{E})) < \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}$, it holds $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$.

Proof: Let $\Delta \in \mathcal{C}(\mathcal{Y})$ be the maximally dephasing channel, that is $\Delta(Y) = \sum_i |i\rangle\langle i|Y|i\rangle\langle i|$. Let us fix r such that $r < \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}$. We will show that if $\mathcal{E} = \mathcal{K}((E_i)) \in \mathcal{C}(\mathcal{Y})$, such that $E_i = \Delta(E_i)$ for each i and $\operatorname{rank}(J(\mathcal{E})) \leq r$, then $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$. Observe that the thesis is true in two particular situations:

- For $\dim(\mathcal{X}) = 1$ and $\dim(\mathcal{Y}) \geq 1$.
- For $r = 1$ and $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$.

Let us take $\mathcal{E} = \mathcal{K}((E_i)) \in \mathcal{C}(\mathcal{Y})$, such that $\operatorname{rank}(J(\mathcal{E})) \leq r$ and $E_i = \Delta(E_i)$ for each i. We may assume that $\operatorname{rank}(J(\mathcal{E})) = r$. Therefore, there exists a projector $\Pi \in \mathcal{P}(\mathcal{Y})$, such that $\operatorname{rank}(\Pi) = r$ and $\Delta(\Pi) = \Pi$, and for which the operators $\Pi E_i \Pi$ are linearly independent. Let us consider the map $\mathcal{F} = \mathcal{K}\left((\Pi^\perp E_i \Pi^\perp)_{i=1}^r\right)$. Define $\mathcal{X}' = \mathbb{C}^{\dim(\mathcal{X})-1}$. By the recurrence and Theorem 1 for $\mathcal{F}$ there exist $S'_* \in \mathcal{M}(\mathcal{X}', \mathcal{Y})$ and $R'_* \in \mathcal{M}(\mathcal{Y}, \mathcal{X}')$, such that $R'_* \Pi^\perp E_i \Pi^\perp S'_* = c_i \mathbb{1}_{\mathcal{X}'}$ and $c_{i_0} \neq 0$ for some i_0. Let $|s\rangle \in \mathcal{C}(\mathcal{Y})$ be the flat superposition. As $\Pi E_i \Pi$ are diagonal and linearly independent, there exists the vector $|r\rangle$ such that $\langle r|\Pi E_i \Pi|s\rangle = c_i$. We may define an encoding operator S_* by adding a column $\Pi|s\rangle$ to the operator $\Pi^\perp S'_*$. In the same manner, we may construct R_* by adding a row $\langle r|\Pi$ to the operator $R'_* \Pi^\perp$. It is easy to check that S_*, R_* satisfy Theorem 1 (D), so $\mathcal{E} \in \xi(\mathcal{X}, \mathcal{Y})$. $\square$

Appendix P
Proof of Corollary 16

Corollary 16: Let $\mathcal{Y}$ be an Euclidean space such that $\dim(\mathcal{Y}) \geq 2$ and let $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ be a Schur channel. Then, $\mathcal{E} \in \xi(\mathbb{C}^2, \mathcal{Y})$ if and only if $\dim(\mathcal{Y}) > \operatorname{rank}(J(\mathcal{E}))$. Moreover, if $\mathcal{E} \in \xi(\mathbb{C}^2, \mathcal{Y})$ then $p_{\mathbb{C}^2}(\mathcal{E}) \geq \frac{1}{\operatorname{rank}(J(\mathcal{E}))^2}$.

Proof: Let $r = \operatorname{rank}(J(\mathcal{E}))$. If $\dim(\mathcal{Y}) > r$, then from Proposition 15 if follows $\mathcal{E} \in \xi(\mathbb{C}^2, \mathcal{Y})$. Assume now that $\dim(\mathcal{Y}) = r$. Let $\mathcal{E} = \mathcal{K}((E_i)_{i=1}^r)$ and define $M = [M_{i,j}]_{i,j=1,\ldots,r}$ such that $\operatorname{rank}(M) = r$. Fix $S \in \mathcal{M}(\mathbb{C}^2, \mathcal{Y})$ and $R \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^2)$ and observe that the following conditions are equivalent:

- For all i it holds $RE_iS = c_i\mathbb{1}_{\mathbb{C}^2}$ and $c_{i_0} \neq 0$ for some i_0.
- For all i it holds $R\sum_j M_{i,j}E_jS = d_i\mathbb{1}_{\mathbb{C}^2}$ and $d_{i_0} \neq 0$ for some i_0.

Since $\mathcal{E}$ is a Schur channel we can take M such that for all i it holds $\sum_j M_{i,j}E_j = |i\rangle\langle i|$. It implies that $\mathcal{E} \notin \xi(\mathbb{C}^2, \mathcal{Y})$.

Now, we will prove that $p_{\mathbb{C}^2}(\mathcal{E}) \geq \frac{1}{\operatorname{rank}(J(\mathcal{E}))^2}$ for $\dim(\mathcal{Y}) > r$. It is enough to show this inequality for $\mathcal{Y} = \mathbb{C}^{r+1}$. Let us fix a Schur channel $\mathcal{E} = \mathcal{K}\left((E_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathbb{C}^{r+1})$. For $i \in \{0, \ldots, r\}$ define

$$|x_i\rangle \coloneqq \sum_{j=0}^{r-1} (E_j)_{i,i}|j\rangle. \tag{109}$$

Observe that $\langle x_i|x_i\rangle = 1$ for all i. First we will show that there exist $i_0 \in \{0, \ldots, r\}$ and a vector $|v_{i_0}\rangle = [(v_{i_0})_i]_{i\neq i_0}$ such that $|x_{i_0}\rangle = \sum_{i\neq i_0}(v_{i_0})_i|x_i\rangle$ and $\langle v_{i_0}|v_{i_0}\rangle \leq r$. Naturally, this statement is true for $r = 1$. By induction, we assume that this statement is true for $r-1$ and we will show that it implies its validity for r. In the first case, assume that there is $i_0 \in \{0, \ldots, r\}$ such that vectors $|x_i\rangle$ for $i \neq i_0$ are linearly dependent. That means, we have r vectors $|x_i\rangle, i \neq i_0$ which belong to a some subspace $\mathbb{C}^{r-1}$. We may use the induction step to prove the correctness of our statement. In the second case, we assume that for all i_0 the vectors $|x_i\rangle$ for $i \neq i_0$ are linearly independent. That means, for each $i_0 \in \{0, \ldots, r\}$ the vector $|x_{i_0}\rangle$ can be uniquely expressed as a linear combination of $|x_i\rangle$ for $i \neq i_0$ with coefficients $[(v_{i_0})_i]_{i\neq i_0}$. Let us define $Q = \sum_{i=0}^r |x_i\rangle\langle x_i|$ and $Q_i = Q - |x_i\rangle\langle x_i| > 0$. One can show that $\langle v_i|v_i\rangle = \langle x_i|Q_i^{-1}|x_i\rangle$ for all i. We obtain

$$\langle v_i|v_i\rangle = \operatorname{tr}(Q_i^{-1}|x_i\rangle\langle x_i|) = \operatorname{tr}(Q_i^{-1}(Q - Q_i))$$

$$
\begin{aligned}
&= \mathrm{tr}(\sqrt{Q}Q_i^{-1}\sqrt{Q}) - r \\
&= \mathrm{tr}\left(\left(\sqrt{Q}^{-1}(Q - |x_i\rangle\langle x_i|)\sqrt{Q}^{-1}\right)^{-1}\right) - r \\
&= \mathrm{tr}\left(\left(\mathbb{1}_{\mathbb{C}^r} - \sqrt{Q}^{-1}|x_i\rangle\langle x_i|\sqrt{Q}^{-1}\right)^{-1}\right) - r \\
&= \frac{\langle x_i|Q^{-1}|x_i\rangle}{1 - \langle x_i|Q^{-1}|x_i\rangle}.
\end{aligned} \tag{110}
$$

On the other hand, we have $\sum_{i=0}^{r}\langle x_i|Q^{-1}|x_i\rangle = \mathrm{tr}(Q^{-1}Q) = r$. There exists $i_0 \in \{0,\ldots,r\}$ such that $\langle x_i|Q^{-1}|x_i\rangle \leq \frac{r}{r+1}$ and hence, $\langle v_{i_0}|v_{i_0}\rangle \leq r$ which ends the proof of our statement.

Now, without loss of generality we assume that there is a vector $|v\rangle = \sum_{i=1}^{r} v_i|i\rangle$ such that $|x_0\rangle = \sum_{i=1}^{r} v_i|x_i\rangle$ and $\langle v|v\rangle \leq r$. Define $c = \max(1, \|v\|_2)$ and let us take

$$
\begin{aligned}
R &= \frac{1}{\sqrt{rc}}|0\rangle\langle 0| + \frac{1}{\sqrt{r}}|1\rangle\left(\sum_{i=1}^{r}\langle i|\right) \in \mathcal{M}(\mathbb{C}^{r+1}, \mathbb{C}^2), \\
S &= |0\rangle\langle 0| + \frac{1}{c}|v\rangle\langle 1| \in \mathcal{M}(\mathbb{C}^2, \mathbb{C}^{r+1}).
\end{aligned} \tag{111}
$$

Observe that $\|S\|_\infty, \|R\|_\infty \leq 1$ and for any $j \in \{0,\ldots,r-1\}$ we have $RE_jS = \frac{(E_j)_{0,0}}{\sqrt{rc}}\mathbb{1}_{\mathbb{C}^2}$. Eventually, we obtain

$$
p_{\mathbb{C}^2}(\mathcal{E}) \geq \sum_{j=0}^{r-1}\frac{|(E_j)_{0,0}|^2}{rc^2} = \frac{1}{rc^2} \geq \frac{1}{r^2}. \tag{112}
$$

□

Appendix Q
Proof of Theorem 17

Theorem 17: Let $\mathcal{X}$ and $\mathcal{Y}$ be some Euclidean spaces such that $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. Then, we have

$$
\begin{aligned}
\left\lfloor \sqrt[4]{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}} \right\rfloor &\leq r_1(\mathcal{X},\mathcal{Y}) \leq \left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil - 1 \\
&\leq r(\mathcal{X},\mathcal{Y}) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2 - 1}.
\end{aligned} \tag{113}
$$

Proof: The inequality $\left\lfloor \sqrt[4]{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}} \right\rfloor \leq r_1(\mathcal{X},\mathcal{Y})$ follows directly from [34]. The inequalities $r_1(\mathcal{X},\mathcal{Y}) \leq \left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil - 1$ and $r(\mathcal{X},\mathcal{Y}) < \frac{\dim(\mathcal{Y})\dim(\mathcal{X})}{\dim(\mathcal{X})^2-1}$ follow from Lemma 14 and Proposition 13, respectively.

Now, we will show that $\left\lceil \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})-1}} \right\rceil - 1 \leq r(\mathcal{X},\mathcal{Y})$. Take arbitrary $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ such that $\mathrm{rank}(J(\mathcal{E}))^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y})$. We will show $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$. Let us denote $r = \mathrm{rank}(J(\mathcal{E}))$. Consider a Kraus representation $\mathcal{E} = \mathcal{K}\left((E_j)_{j=1}^r\right)$ and define the following set

$$
\begin{aligned}
A = \{&s \in \mathbb{N} : \exists_{\Pi_s \in \mathcal{P}(\mathcal{Y})}\ \Pi_s = \Pi_s^2, \mathrm{rank}(\Pi_s) = s, \\
&\mathrm{rank}(\mathcal{E}^\dagger(\Pi_s)) = \dim(\mathcal{Y})\}.
\end{aligned} \tag{114}
$$

Observe that $\dim(\mathcal{Y}) \in A$ and if some $s \in A$, then $sr \geq \dim(\mathcal{Y})$. Define $s_0 = \min(A)$ and consider a corresponding projector $\Pi_{s_0} \in \mathcal{P}(\mathcal{Y})$, such that $\mathrm{rank}(\Pi_{s_0}) = s_0$ and $\mathrm{rank}(\mathcal{E}^\dagger(\Pi_{s_0})) = \dim(\mathcal{Y})$. Let us take a orthonormal collection of vectors $|v_i\rangle$, where $i = 1,\ldots,s_0$ for which we have $\Pi_{s_0} = \sum_{i=1}^{s_0}|v_i\rangle\langle v_i|$. From the assumption $s_0 = \min(A)$, for any i we get $\mathrm{rank}(\mathcal{E}^\dagger(\Pi_{s_0} - |v_i\rangle\langle v_i|)) < \dim(\mathcal{Y})$. Therefore, we may define vectors $\mathcal{Y} \ni |w_i\rangle \neq 0$ such that $\mathcal{E}^\dagger(\Pi_{s_0} - |v_i\rangle\langle v_i|)|w_i\rangle = 0$. Observe that for each i, there exists E_j for which $\langle v_i|E_j|w_i\rangle \neq 0$. Let us define $F_j = [\langle v_a|E_j|w_b\rangle]_{a,b=1,\ldots,s_0}$ for $j = 1,\ldots,r$. Note, that F_j are diagonal operators and it holds $\sum_j F_j^\dagger F_j > 0$. From $r^2(\dim(\mathcal{X}) - 1) < \dim(\mathcal{Y})$ and $s_0 r \geq \dim(\mathcal{Y})$ we have

$$
r(\dim(\mathcal{X}) - 1) < \frac{\dim(\mathcal{Y})}{r} \leq s_0. \tag{115}
$$

Utilizing Proposition 15, Lemma 9 and Theorem 1 there exist $S_* \in \mathcal{M}(\mathcal{X}, \mathbb{C}^{s_0})$ and $R_* \in \mathcal{M}(\mathbb{C}^{s_0}, \mathcal{X})$, such that $R_*F_jS_* \propto \mathbb{1}_{\mathcal{X}}$ and there exists j_0, for which it holds $R_*F_{j_0}S_* \neq 0$. That implies $\mathcal{E} \in \xi(\mathcal{X},\mathcal{Y})$. □

Appendix R
Proof of Proposition 19

Proposition 19: For all $\mathcal{E} \in \mathcal{C}(\mathbb{C}^4)$ satisfying $\mathrm{rank}(J(\mathcal{E})) \leq 2$ we have $\mathcal{E} \in \xi(\mathbb{C}^2, \mathbb{C}^4)$.

Proof: Let us fix $\mathcal{E} = \mathcal{K}((E_0, E_1)) \in \mathcal{C}(\mathbb{C}^4)$. From the equality $E_0^\dagger E_0 + E_1^\dagger E_1 = \mathbb{1}_{\mathbb{C}^4}$ we may write the singular decomposition of E_0, E_1 in the form: $E_0 = U_0D_0V$ and $E_1 = U_1D_1V$, where $U_0, U_1, V \in \mathcal{U}(\mathbb{C}^4)$ and $D_0, D_1 \in \mathcal{P}(\mathbb{C}^4)$ are diagonal operators satisfying $D_0^2 + D_1^2 = \mathbb{1}_{\mathbb{C}^4}$. In order to show that $\mathcal{E} \in \xi(\mathbb{C}^2, \mathbb{C}^4)$ we will use Theorem 1 (D). We will prove that there exist $S_* \in \mathcal{M}(\mathbb{C}^2, \mathbb{C}^4)$ and $R_* \in \mathcal{M}(\mathbb{C}^4, \mathbb{C}^2)$, such that $R_*E_0S_* = c_0\mathbb{1}_{\mathbb{C}^2}$, $R_*E_1S_* = c_1\mathbb{1}_{\mathbb{C}^2}$ for some $c_0, c_1 \in \mathbb{C}$ satisfying $(c_0, c_1) \neq (0, 0)$. Let us introduce the following notation

$$
\begin{aligned}
|x_i\rangle &= (D_0)_{ii}U_0|i\rangle, \quad i = 0,\ldots,3, \\
|y_i\rangle &= (D_1)_{ii}U_1|i\rangle, \quad i = 0,\ldots,3.
\end{aligned} \tag{116}
$$

Note that vectors $|x_i\rangle$ are orthogonal (the same holds for $|y_i\rangle$) and for each $i = 0,\ldots,3$ we have $|x_i\rangle \neq 0$ or $|y_i\rangle \neq 0$. We may write S_* and R_* in the following form

$$
\begin{aligned}
S_* &= V^\dagger(|S_0\rangle\langle 0| + |S_1\rangle\langle 1|), \\
R_* &= |0\rangle\langle R_0| + |1\rangle\langle R_1|,
\end{aligned} \tag{117}
$$

for some vectors $|S_0\rangle, |S_1\rangle, |R_0\rangle, |R_1\rangle \in \mathbb{C}^4$. The rest of the proof will be divided into three cases.

In the first case, we assume there exists $i_3 \in \{0,\ldots,3\}$ such that vectors $|x_{i_3}\rangle, |y_{i_3}\rangle$ are linearly independent. Define indices $i_0, i_1, i_2 \in \{0,\ldots,3\}$ as the remaining labels, such that $\{i_0,\ldots,i_3\}$ covers the whole set $\{0,\ldots,3\}$. Let $(a_0, a_1, a_2)^\top \in \mathbb{C}^3$ be a normalized vector orthogonal to vectors $(\langle y_{i_3}|x_{i_0}\rangle, \langle y_{i_3}|x_{i_1}\rangle, \langle y_{i_3}|x_{i_2}\rangle)^\dagger$ and $(\langle x_{i_3}|y_{i_0}\rangle, \langle x_{i_3}|y_{i_1}\rangle, \langle x_{i_3}|y_{i_2}\rangle)^\dagger$. Take $|S_1\rangle = |i_3\rangle$ and $|S_0\rangle = a_0|i_0\rangle + a_1|i_1\rangle + a_2|i_2\rangle$. Define $|x\rangle = a_0|x_{i_0}\rangle + a_1|x_{i_1}\rangle + a_2|x_{i_2}\rangle$ and $|y\rangle = a_0|y_{i_0}\rangle + a_1|y_{i_1}\rangle + a_2|y_{i_2}\rangle$. We obtain

$$
\begin{aligned}
E_0S_* &= |x\rangle\langle 0| + |x_{i_3}\rangle\langle 1|, \\
E_1S_* &= |y\rangle\langle 0| + |y_{i_3}\rangle\langle 1|.
\end{aligned} \tag{118}
$$

It is not hard to observe that $|x\rangle \neq 0$ or $|y\rangle \neq 0$. If $|x\rangle \neq 0$, take $|R_0\rangle = |x\rangle$, else take $|R_0\rangle = |y\rangle$. As the vectors

$|x_{i_3}\rangle, |y_{i_3}\rangle$ are linearly independent we may define

$$(b_0, b_1)^\top := \begin{bmatrix} \langle x_{i_3}|x_{i_3}\rangle & \langle y_{i_3}|x_{i_3}\rangle \\ \langle x_{i_3}|y_{i_3}\rangle & \langle y_{i_3}|y_{i_3}\rangle \end{bmatrix}^{-1} (\langle R_0|x\rangle, \langle R_0|y\rangle)^\top \quad (119)$$

Take $|R_1\rangle = \bar{b}_0|x_{i_3}\rangle + \bar{b}_1|y_{i_3}\rangle$. Finally, we may check that it holds

$$R_*E_0S_* = (|0\rangle\langle R_0| + |1\rangle\langle R_1|)(|x\rangle\langle 0| + |x_{i_3}\rangle\langle 1|) = \langle R_0|x\rangle \mathbb{1}_{\mathbb{C}^2},$$
$$R_*E_1S_* = (|0\rangle\langle R_0| + |1\rangle\langle R_1|)(|y\rangle\langle 0| + |y_{i_3}\rangle\langle 1|) = \langle R_0|y\rangle \mathbb{1}_{\mathbb{C}^2}. \quad (120)$$

In the second case, we assume that there exists a pair of vectors $|y_{i_0}\rangle, |y_{i_1}\rangle$ for $i_0 \neq i_1$, such that $|y_{i_0}\rangle = |y_{i_1}\rangle = 0$. Then, the vectors $|x_{i_0}\rangle, |x_{i_1}\rangle$ are orthonormal. We simply define $|S_0\rangle = |i_0\rangle, |S_1\rangle = |i_1\rangle, |R_0\rangle = |x_{i_0}\rangle$ and $|R_1\rangle = |x_{i_1}\rangle$. One can calculate that $R_*E_0S_* = \mathbb{1}_{\mathbb{C}^2}$ and $R_*E_1S_* = 0$.

In the third case, for all $i \in \{0, \ldots, 3\}$ vectors $|x_i\rangle, |y_i\rangle$ are not linearly independent and there is at most one zero vector $|y_{i_3}\rangle$ for some $i_3 \in \{0, \ldots, 3\}$. Define indices $i_0, i_1, i_2 \in \{0, \ldots, 3\}$ as the remaining labels, such that $\{i_0, \ldots, i_3\}$ covers the whole set $\{0, \ldots, 3\}$. Define the matrix

$$M = \begin{bmatrix} \langle y_{i_0}|x_{i_0}\rangle & \langle y_{i_1}|x_{i_1}\rangle & \langle y_{i_2}|x_{i_2}\rangle \\ \langle y_{i_0}|y_{i_0}\rangle & \langle y_{i_1}|y_{i_1}\rangle & \langle y_{i_2}|y_{i_2}\rangle \end{bmatrix}. \quad (121)$$

In the first sub-case we assume that $\mathrm{rank}(M) = 1$. Define $b = \frac{\langle y_{i_1}|y_{i_1}\rangle}{\langle y_{i_0}|y_{i_0}\rangle}$. We can take $|S_0\rangle = |i_0\rangle$, $|S_1\rangle = |i_1\rangle$, $|R_0\rangle = |y_{i_0}\rangle$ and $|R_1\rangle = \frac{1}{b}|y_{i_1}\rangle$. One can calculate that $R_*E_0S_* = \langle y_{i_0}|x_{i_0}\rangle \mathbb{1}_{\mathbb{C}^2}$ and $R_*E_1S_* = \langle y_{i_0}|y_{i_0}\rangle \mathbb{1}_{\mathbb{C}^2}$.

In the second sub-case we assume that $\mathrm{rank}(M) = 2$. Define indices $j_1, j_2 \in \{0, 1, 2\}$, such that

$$\mathrm{rank}\left(\begin{bmatrix} M_{0,j_1} & M_{0,j_2} \\ M_{1,j_1} & M_{1,j_2} \end{bmatrix}\right) = 2. \quad (122)$$

Define $j_0 \in \{0, 1, 2\}$ as the remaining label, such that $\{j_0, j_1, j_2\}$ covers the whole set $\{0, 1, 2\}$. Take $|S_0\rangle = |i_{j_0}\rangle$, $|R_0\rangle = |y_{i_{j_0}}\rangle$ and define

$$(b_1, b_2)^\top := \begin{bmatrix} \langle y_{i_{j_1}}|x_{i_{j_1}}\rangle & \langle y_{i_{j_2}}|x_{i_{j_2}}\rangle \\ \langle y_{i_{j_1}}|y_{i_{j_1}}\rangle & \langle y_{i_{j_2}}|y_{i_{j_2}}\rangle \end{bmatrix}^{-1} \begin{pmatrix} \langle y_{i_{j_0}}|x_{i_{j_0}}\rangle \\ \langle y_{i_{j_0}}|y_{i_{j_0}}\rangle \end{pmatrix}. \quad (123)$$

We may take $|S_1\rangle = |i_{j_1}\rangle + |i_{j_2}\rangle$ and $|R_1\rangle = \bar{b}_1|y_{i_{j_1}}\rangle + \bar{b}_2|y_{i_{j_2}}\rangle$. Direct calculations reveal that $R_*E_0S_* = \langle y_{i_{j_0}}|x_{i_{j_0}}\rangle \mathbb{1}_{\mathbb{C}^2}$ and $R_*E_1S_* = \langle y_{i_{j_0}}|y_{i_{j_0}}\rangle \mathbb{1}_{\mathbb{C}^2}$. □

Appendix S
Proof of Theorem 21

Theorem 21: Let $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$ be a random quantum channel defined according to (36). Then, the following two implications hold

$$r < \frac{\dim(\mathcal{X})\dim(\mathcal{Y})}{\dim(\mathcal{X})^2 - 1} \implies \mathcal{P}\left(\mathcal{E}_r \in \xi(\mathcal{X}, \mathcal{Y})\right) = 1,$$
$$\mathcal{P}\left(\mathcal{E}_r \in \xi_1(\mathcal{X}, \mathcal{Y})\right) = 1 \implies r < \sqrt{\frac{\dim(\mathcal{Y})}{\dim(\mathcal{X}) - 1}}. \quad (124)$$

Proof: For $r \in \mathbb{N}$ satisfying $r < \frac{\dim(\mathcal{X})\dim(\mathcal{Y})}{\dim(\mathcal{X})^2-1}$, let $(G_i)_{i=1}^r \subset \mathcal{M}(\mathcal{Y})$ be a tuple of random and independent Ginibre matrices and $Q = \sum_{i=1}^r G_i^\dagger G_i$. Define the projector $\Pi = \sum_{i=0}^{\dim(\mathcal{X})-1} |i\rangle\langle i|$ and consider the set

$$A = \Bigg\{ (G_i)_{i=1}^r : \quad \mathrm{rank}(Q) = \dim(\mathcal{Y}),$$
$$\mathrm{rank}\left(\sum_{i=1}^r G_i^\dagger \Pi G_i\right) = \min\{r\dim(\mathcal{X}), \dim(\mathcal{Y})\} \Bigg\}. \quad (125)$$

One can observe that $\mathcal{P}((G_i)_{i=1}^r \in A) = 1$. Let $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$ be a random channel defined according to (36) for $(G_i)_{i=1}^r \in A$, that is $\mathcal{E}_r(Y) = \sum_{i=1}^r \left(G_i Q^{-1/2}\right) Y \left(G_i Q^{-1/2}\right)^\dagger$. Define $S = Q^{1/2}\tilde{S}$ for $\tilde{S} \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$ and $R = \tilde{R}\Pi$ for $\tilde{R} \in \mathcal{M}(\mathcal{Y}, \mathcal{X})$. We obtain $RG_iQ^{-1/2}S = \tilde{R}\Pi G_i \tilde{S}$. Utilizing Lemma 9, Proposition 13 and Theorem 1 (D) for $\tilde{\mathcal{E}} = \mathcal{K}\left((\Pi G_i)_{i=1}^r\right) \in s\mathcal{C}(\mathcal{Y})$, there exist $\tilde{S}, \tilde{R}$, such that $\tilde{R}\Pi G_i \tilde{S} \propto \mathbb{1}_\mathcal{X}$ and $\tilde{R}\Pi G_{i_0}\tilde{S} \neq 0$ for some i_0. Finally, $\mathcal{E}_r \in \xi(\mathcal{X}, \mathcal{Y})$.

Now, for a given $r \in \mathbb{N}$ let us define $B = \{\mathcal{E}_r : \mathcal{E}_r \in \xi_1(\mathcal{X}, \mathcal{Y})\}$. From the assumption $\mathcal{P}(B) = 1$, we obtain that B is a dense subset of $\{\mathcal{E} \in \mathcal{C}(\mathcal{Y}) : \mathrm{rank}(J(\mathcal{E})) \leq r\}$. Imitating the proof of Theorem 7, we get that if $\mathcal{E} \in \mathcal{C}(\mathcal{Y})$ and $\mathrm{rank}(J(\mathcal{E})) \leq r$, then $\mathcal{E} \in \xi_1(\mathcal{X}, \mathcal{Y})$. That implies $r \leq r_1(\mathcal{X}, \mathcal{Y})$. By using Lemma 14 we obtain the desired inequality. □

Appendix T
Proof of Corollary 22

Corollary 22: Let $\mathcal{E}_r = \mathcal{K}\left((E_i)_{i=0}^{r-1}\right) \in \mathcal{C}(\mathcal{Y})$ be a random quantum channel defined according to (36) and assume that $r \leq \frac{\dim(\mathcal{Y})}{\dim(\mathcal{X})}$. Define a sequence $V_1, V_2, \ldots$ of random isometry matrices sampled according to the Haar measure, such that $V_n \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$. Let $R_{F_n} = (F_nF_n^\dagger)^{-1}$ for $F_n = \sum_{i=0}^{r-1} |i\rangle \otimes V_n^\dagger E_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Then, almost surely it holds

$$\begin{aligned} &p_\mathcal{X}(\mathcal{E}_r) \\ &\geq \sup_{n\in\mathbb{N}} \max\left\{\mathrm{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1\left(R_{F_n}(P \otimes \mathbb{1}_\mathcal{X})\right) \leq \mathbb{1}_\mathcal{X}\right\} \\ &\geq \max\big\{\lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V^\dagger)EE^\dagger(\mathbb{1}_{\mathbb{C}^r} \otimes V)) : \\ &\qquad V \in \mathcal{M}(\mathcal{X}, \mathcal{Y}), V^\dagger V = \mathbb{1}_\mathcal{X}\big\}, \end{aligned} \quad (126)$$

where $\lambda_{\min}$ is the smallest eigenvalue and $E = \sum_{i=0}^{r-1} |i\rangle \otimes E_i$.

Proof: For $r \in \mathbb{N}$ satisfying $r\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$, let $(G_i)_{i=0}^{r-1} \subset \mathcal{M}(\mathcal{Y})$ be a tuple of random and independent Ginibre matrices and $Q = \sum_{i=0}^{r-1} G_i^\dagger G_i$. Define a sequence $V_1, V_2, \ldots$ of random isometry matrices sampled according to the Haar measure, such that $V_n \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$. Consider the following sets

$$\begin{aligned} A_0 &= \{\left((G_i)_{i=0}^{r-1}, (V_n)_{n\in\mathbb{N}}\right) : \mathrm{rank}(Q) = \dim(\mathcal{Y})\}, \\ A_m &= \{\left((G_i)_{i=0}^{r-1}, (V_n)_{n\in\mathbb{N}}\right) : \\ &\qquad \mathrm{rank}\left(\sum_{i=0}^{r-1} G_i^\dagger V_m V_m^\dagger G_i\right) = r\dim(\mathcal{X})\}. \end{aligned} \quad (127)$$

One can observe that $\mathcal{P}(\bigcap_{n\geq 0} A_n) = 1$. For $n \in \mathbb{N}$ let $R_{F_n} = (F_nF_n^\dagger)^{-1}, \Pi_{F_n} = F_nF_n^{-1}$, where $F_n = \sum_{i=0}^{r-1} |i\rangle \otimes V_n^\dagger E_i \in \mathcal{M}(\mathcal{Y}, \mathbb{C}^r \otimes \mathcal{X})$. Utilizing Lemma 10 we obtain

$$p_\mathcal{X}(\mathcal{E}_r)$$

$$\geq \sup_{n\in\mathbb{N}} \max \{\mathrm{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1 (R_{F_n}(P \otimes \mathbb{1}_\mathcal{X})) \leq \mathbb{1}_\mathcal{X}, (\Pi_{F_n} \otimes \mathbb{1}_\mathcal{X})(P \otimes |\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|)(\Pi_{F_n} \otimes \mathbb{1}_\mathcal{X}) = P \otimes |\mathbb{1}_\mathcal{X}\rangle\langle\mathbb{1}_\mathcal{X}|\} . \tag{128}$$

It holds that $F_n^\dagger F_n = Q^{-1/2} \sum_{i=0}^{r-1} G_i^\dagger V_n V_n^\dagger G_i Q^{-1/2}$. Hence, almost surely $\mathrm{rank}(F_n) = r \dim(\mathcal{X})$ which implies that $\Pi_{F_n} = \mathbb{1}_{\mathbb{C}^r \otimes \mathcal{X}}$. Therefore, we get

$$p_\mathcal{X}(\mathcal{E}_r) \geq \sup_{n\in\mathbb{N}} \max \{\mathrm{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1 (R_{F_n}(P \otimes \mathbb{1}_\mathcal{X})) \leq \mathbb{1}_\mathcal{X}\} . \tag{129}$$

To prove the second inequality note that $\max \{\mathrm{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1 (R_{F_n}(P \otimes \mathbb{1}_\mathcal{X})) \leq \mathbb{1}_\mathcal{X}\} > 0$ for each $n \in \mathbb{N}$. By Lemma 10 and the fact that $R_{F_n} > 0$ for each $n \in \mathbb{N}$ we get

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \max \{\mathrm{tr}(P) : P \in \mathcal{P}(\mathbb{C}^r), \mathrm{tr}_1 (R_{F_n}(P \otimes \mathbb{1}_\mathcal{X})) \leq \mathbb{1}_\mathcal{X}\} \\
&\geq \sup_{n\in\mathbb{N}} \|R_{F_n}\|_\infty^{-1} = \sup_{n\in\mathbb{N}} \|((\mathbb{1}_{\mathbb{C}^r} \otimes V_n^\dagger)EE^\dagger(\mathbb{1}_{\mathbb{C}^r} \otimes V_n))^{-1}\|_\infty^{-1}. \\
&= \sup_{n\in\mathbb{N}} \lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V_n^\dagger)EE^\dagger(\mathbb{1}_{\mathbb{C}^r} \otimes V_n)) \\
&= \max \{\lambda_{\min}((\mathbb{1}_{\mathbb{C}^r} \otimes V^\dagger)EE^\dagger(\mathbb{1}_{\mathbb{C}^r} \otimes V)) : V \in \mathcal{M}(\mathcal{X},\mathcal{Y}), V^\dagger V = \mathbb{1}_\mathcal{X}\} ,
\end{aligned} \tag{130}$$

where in the last equality we used the fact that the subset $\{V_n : n \in \mathbb{N}\}$ is almost surely dense in the set of isometry matrices and the fact that $\lambda_{\min}$ is a continuous function. □

APPENDIX U
EXTREMALITY OF RANDOM CHANNELS

Lemma 25: Let $r \in \mathbb{N}$ and let $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$ be a random quantum channel defined according to (36). Then, almost surely it holds

- $\mathrm{rank}(J(\mathcal{E}_r)) = \min(r, \dim(\mathcal{Y})^2)$,
- $\mathrm{rank}(J(\mathcal{E}_r^\dagger\mathcal{E}_r)) = \min(r^2, \dim(\mathcal{Y})^2)$.

In particular, if $r \leq \dim(\mathcal{Y})$, then $\mathcal{E}_r$ is almost surely an extremal channel.

Proof: The channel $\mathcal{E}_r \in \mathcal{C}(\mathcal{Y})$ is given as $\mathcal{E}_r = \mathcal{K}\left((G_i Q^{-1/2}))_{i=1}^r\right)$, where $Q = \sum_{i=1}^r G_i^\dagger G_i$ and $(G_i)_{i=1}^r \subset \mathcal{M}(\mathcal{Y})$ is a tuple of random and independent Ginibre matrices. As $\mathrm{rank}(Q) = \dim(\mathcal{Y})$ almost surely, then

$$\begin{aligned}
\mathrm{rank}(J(\mathcal{E}_r)) &= \mathrm{rank}\left(\sum_{i=1}^r |G_i Q^{-1/2}\rangle\langle i|\right) \\
&=\mathrm{rank}\left(\sum_{i=1}^r |G_i\rangle\langle i|\right) = \min(r, \dim(\mathcal{Y})^2).
\end{aligned} \tag{131}$$

Similarly, we get

$$\begin{aligned}
\mathrm{rank}(J(\mathcal{E}_r^\dagger\mathcal{E}_r)) &= \mathrm{rank}\left(\sum_{i,j=1}^r |Q^{-1/2}G_j^\dagger G_i Q^{-1/2}\rangle\langle j,i|\right) \\
&=\mathrm{rank}\left(\sum_{i,j=1}^r |G_j^\dagger G_i\rangle\langle j,i|\right) \\
&=\mathrm{rank}\left(\sum_{a=0}^{\dim(\mathcal{Y})-1} \sum_{i,j=1}^r (G_j^\dagger|a\rangle \otimes \overline{G_i^\dagger|a\rangle})\langle j,i|\right) \\
&=\mathrm{rank}\left(\sum_{a=0}^{\dim(\mathcal{Y})-1} K_a \otimes \overline{K_a}\right),
\end{aligned} \tag{132}$$

where $(K_a)_{a=0}^{\dim(\mathcal{Y})-1} \subset \mathcal{M}(\mathbb{C}^r,\mathcal{Y})$ is a tuple of random and independent Ginibre matrices. As $\mathrm{rank}(K_a) = \min(r,\dim(\mathcal{Y}))$ it follows that $\mathrm{rank}(J(\mathcal{E}_r^\dagger\mathcal{E}_r)) = \min(r^2,\dim(\mathcal{Y})^2)$. □

APPENDIX V
PROOF OF PROPOSITION 23

Proposition 23: Let $\Upsilon \subset \mathcal{C}(\mathcal{Y})$ be a nonempty and convex family of noise channels. Define μ to be a probability measure defined on Υ and assume that the support of μ is equal to Υ. Let $\bar{\mathcal{E}} = \int_\Upsilon \mathcal{E}\mu(d\mathcal{E}) \in \mathcal{C}(\mathcal{Y})$ and fix $(\mathcal{S},\mathcal{R}) \in s\mathcal{C}(\mathcal{X},\mathcal{Y}) \times s\mathcal{C}(\mathcal{Y},\mathcal{X})$. The following conditions are equivalent:

(A) For each $\mathcal{E} \in \Upsilon$ there exists $p_\mathcal{E} \geq 0$ such that $\mathcal{R}\mathcal{E}\mathcal{S} = p_\mathcal{E}\mathcal{I}_\mathcal{X}$ and $\int_\Upsilon p_\mathcal{E}\mu(d\mathcal{E}) > 0$.
(B) It holds that $0 \neq \mathcal{R}\bar{\mathcal{E}}\mathcal{S} \propto \mathcal{I}_\mathcal{X}$.

Proof: $(B) \implies (A)$
Let us assume that $\mathcal{R}\bar{\mathcal{E}}\mathcal{S} = p\mathcal{I}_\mathcal{X}$ for $p > 0$. There exists a k dimensional affine subspace $\mathcal{L}$ such that $\Upsilon \subset \mathcal{L}$ and $\mathrm{int}(\Upsilon) \neq \emptyset$. Take an arbitrary $\mathcal{E}_0 \in \Upsilon$. There exist $\mathcal{E}_1,\dots,\mathcal{E}_k \in \Upsilon$ such that convex hull of points $\mathcal{E}_0,\dots,\mathcal{E}_k$ is a k-dimensional simplex Δ_k. For any state $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ it holds

$$\begin{aligned}
p|\psi\rangle\langle\psi| = \mathcal{R}\bar{\mathcal{E}}\mathcal{S}(|\psi\rangle\langle\psi|) &= \int_\Upsilon \mathcal{R}\mathcal{E}\mathcal{S}(|\psi\rangle\langle\psi|)\mu(d\mathcal{E}) \\
&\geq \int_{\Delta_k} \mathcal{R}\mathcal{E}\mathcal{S}(|\psi\rangle\langle\psi|)\mu(d\mathcal{E}).
\end{aligned} \tag{133}$$

Inside Δ_k each $\mathcal{E}$ can be uniquely represented as $\sum_{i=0}^k q_i(\mathcal{E})\mathcal{E}_i$, where $(q_i(\mathcal{E}))_{i=0}^k$ is a probability vector which depends on $\mathcal{E}$. Hence,

$$\begin{aligned}
p|\psi\rangle\langle\psi| &\geq \sum_{i=0}^k \int_{\Delta_k} q_i(\mathcal{E})\mathcal{R}\mathcal{E}_i\mathcal{S}(|\psi\rangle\langle\psi|)\mu(d\mathcal{E}) \\
&\geq \left(\int_{\Delta_k} q_0(\mathcal{E})\mu(d\mathcal{E})\right)\mathcal{R}\mathcal{E}_0\mathcal{S}(|\psi\rangle\langle\psi|).
\end{aligned} \tag{134}$$

There exists ϵ small ball B_ϵ around $\mathcal{E}_0$, such that for each channel $\mathcal{E} \in B_\epsilon \cap \Delta_k$ it holds $q_0(\mathcal{E}) \geq \frac{1}{2}$. Hence, $\int_{\Delta_k} q_0(\mathcal{E})\mu(d\mathcal{E}) \geq \frac{1}{2}\mu(B_\epsilon \cap \Delta_k) > 0$, where in the last inequality we used the fact that the support of μ is equal to Υ. Therefore, it holds that for any $|\psi\rangle\langle\psi| \in \mathcal{D}(\mathcal{X})$ we have $\mathcal{R}\mathcal{E}_0\mathcal{S}(|\psi\rangle\langle\psi|) \propto |\psi\rangle\langle\psi|$ and from Lemma 24 there exists $p_{\mathcal{E}_0} \geq 0$ such that $\mathcal{R}\mathcal{E}_0\mathcal{S} = p_{\mathcal{E}_0}\mathcal{I}_\mathcal{X}$. The instant relation $\int_\Upsilon p_\mathcal{E}\mu(d\mathcal{E}) = p > 0$ ends the proof. □

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their insightful comments and suggestions that allowed them to improve this article considerably. They are also grateful to Karol Życzkowski for numerous discussions on the probability of successful error correction.

REFERENCES

[1] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018.

[2] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Phys. Rev. A, Gen. Phys.*, vol. 52, no. 4, pp. 2493–2496, Oct. 1995.

[3] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 2, p. 1098, 1996.

[4] A. M. Steane, "Error correcting codes in quantum theory," *Phys. Rev. Lett.*, vol. 77, no. 5, p. 793, 1996.

[5] A. Steane, "Multiple-particle interference and quantum error correction," *Proc. R. Soc. Lond. A, Math., Phys. Eng. Sci.*, vol. 452, pp. 2551–2577, Nov. 1996.

[6] D. Gottesman, "Stabilizer codes and quantum error correction," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, USA, 1997.

[7] H. Bombin and M. A. Martin-Delgado, "Topological quantum error correction with optimal encoding rate," *Phys. Rev. A, Gen. Phys.*, vol. 73, no. 6, Jun. 2006, Art. no. 062303.

[8] D. Kribs, R. Laflamme, and D. Poulin, "Unified and generalized approach to quantum error correction," *Phys. Rev. Lett.*, vol. 94, no. 18, May 2005, Art. no. 180501.

[9] T. A. Brun, I. Devetak, and M.-H. Hsieh, "Correcting quantum errors with entanglement," *Science*, vol. 314, no. 5798, pp. 436–439, Oct. 2006.

[10] T. A. Brun, I. Devetak, and M.-H. Hsieh, "Catalytic quantum error correction," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3073–3089, Mar. 2014.

[11] D. MacKay, G. Mitchison, and P. McFadden, "Sparse-graph codes for quantum error correction," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2315–2330, Oct. 2004.

[12] F. Huber and M. Grassl, "Quantum codes of maximal distance and highly entangled subspaces," *Quantum*, vol. 4, p. 284, Jun. 2020.

[13] D. A. Lidar and T. A. Brun, *Quantum Error Correction*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[14] M. Koashi and M. Ueda, "Reversing measurement and probabilistic quantum error correction," *Phys. Rev. Lett.*, vol. 82, no. 12, pp. 2598–2601, Mar. 1999.

[15] J. Fern and J. Terilla, "Probabilistic quantum error correction," 2002, *arXiv:quant-ph/0209058*.

[16] P. Barberis-Blostein, D. G. Norris, L. A. Orozco, and H. J. Carmichael, "From quantum feedback to probabilistic error correction: Manipulation of quantum beats in cavity QED," *New J. Phys.*, vol. 12, no. 2, Feb. 2010, Art. no. 023002.

[17] X. Xiao and Y.-L. Li, "Protecting qutrit-qutrit entanglement by weak measurement and reversal," *Eur. Phys. J. D*, vol. 67, no. 10, pp. 1–7, Oct. 2013.

[18] S.-C. Wang, Z.-W. Yu, W.-J. Zou, and X.-B. Wang, "Protecting quantum states from decoherence of finite temperature using weak measurement," *Phys. Rev. A, Gen. Phys.*, vol. 89, no. 2, Feb. 2014, Art. no. 022318.

[19] A. Ashikhmin, "Fidelity of a quantum ARQ protocol," in *Proc. IEEE Inf. Theory Workshop*, Mar. 2006, pp. 42–46.

[20] A. J. Scott, "Probabilities of failure for quantum error correction," *Quantum Inf. Process.*, vol. 4, no. 5, pp. 399–431, Nov. 2005.

[21] T. Camara, H. Ollivier, and J.-P. Tillich, "A class of quantum LDPC codes: Construction and performances under iterative decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 811–815.

[22] K. Kasai, M. Hagiwara, H. Imai, and K. Sakaniwa, "Quantum error correction beyond the bounded distance decoding limit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1223–1230, Feb. 2012.

[23] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, "Quantum error detection. I. Statement of the problem," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 778–788, May 2000.

[24] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, "Quantum error detection. II. Bounds," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 789–800, May 2000.

[25] K. Wang, X. Zhao, and T. Yu, "Environment-assisted quantum state restoration via weak measurements," *Phys. Rev. A, Gen. Phys.*, vol. 89, no. 4, Apr. 2014, Art. no. 042320.

[26] L.-M. Duan and G.-C. Guo, "Probabilistic cloning and identification of linearly independent quantum states," *Phys. Rev. Lett.*, vol. 80, p. 4999, Jun. 1998.

[27] M. Sedlák, A. Bisio, and M. Ziman, "Optimal probabilistic storage and retrieval of unitary channels," *Phys. Rev. Lett.*, vol. 122, no. 17, May 2019, Art. no. 170502.

[28] Z. Puchala, L. Pawela, A. Krawiec, R. Kukulski, and M. Oszmaniec, "Multiple-shot and unambiguous discrimination of von Neumann measurements," *Quantum*, vol. 5, p. 425, Apr. 2021.

[29] E. Knill and R. Laflamme, "Theory of quantum error-correcting codes," *Phys. Rev. A, Gen. Phys.*, vol. 55, no. 2, p. 900, 1997.

[30] J. Watrous, *The Theory of Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2018.

[31] K. E. Hellwig and K. Kraus, "Pure operations and measurements," *Commun. Math. Phys.*, vol. 11, no. 3, pp. 214–220, Sep. 1969.

[32] S. Hill and W. K. Wootters, "Entanglement of a pair of quantum bits," *Phys. Rev. Lett.*, vol. 78, no. 26, pp. 5022–5025, Jun. 1997.

[33] M. Muller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *J. Math. Phys.*, vol. 54, no. 12, Dec. 2013, Art. no. 122203.

[34] E. Knill, R. Laflamme, and L. Viola, "Theory of quantum error correction for general noise," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2525–2528, Mar. 2000.

[35] G. Chiribella, M. Dall'Arno, G. M. D'Ariano, C. Macchiavello, and P. Perinotti, "Quantum error correction with degenerate codes for correlated noise," *Phys. Rev. A, Gen. Phys.*, vol. 83, no. 5, May 2011, Art. no. 052305.

[36] W. Bruzda, V. Cappellini, H.-J. Sommers, and K. Zyczkowski, "Random quantum operations," *Phys. Lett. A*, vol. 373, no. 3, pp. 320–324, Jan. 2009.

[37] I. Nechita, Z. Puchala, L. Pawela, and K. Zyczkowski, "Almost all quantum channels are equidistant," *J. Math. Phys.*, vol. 59, no. 5, May 2018, Art. no. 052201.

[38] R. Kukulski, I. Nechita, L. Pawela, Z. Puchala, and K. Zyczkowski, "Generating random quantum channels," *J. Math. Phys.*, vol. 62, no. 6, Jun. 2021, Art. no. 062201.

[39] M. Girard et al., "On the mixed-unitary rank of quantum channels," *Commun. Math. Phys.*, vol. 394, pp. 919–951, Sep. 2022.

[40] M.-D. Choi, D. W. Kribs, and K. Zyczkowski, "Quantum error correcting codes from the compression formalism," *Rep. Math. Phys.*, vol. 58, no. 1, pp. 77–91, Aug. 2006.

Ryszard Kukulski was born in Poland, in 1995. He received the M.Sc. degree in mathematics from the Faculty of Science and Technology, University of Silesia, Katowice, Poland, in 2019. He is currently pursuing the Ph.D. degree with the Doctoral School, Silesian University of Technology, Gliwice, Poland. Since 2019, he has been with the Quantum Information Group, Institute of Theoretical and Applied Informatics, Polish Academy of Sciences. His research interests include mathematical aspects of quantum information theory, with a main focus on quantum error correction and the theory of random quantum operations.

Łukasz Pawela was born in Poland, in 1987. He received the M.Sc. degree in physics from the Faculty of Applied Mathematics, Silesian University of Technology, Gliwice, Poland, in 2006, the Ph.D. degree in computer science from the Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Gliwice, in 2017, and the Habilitation degree in computer science from the Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, in 2020. He was a Post-Doctoral Researcher with the National Quantum Information Centre, Gdańsk, Poland. Since 2020, he has been an Associate Professor with the Institute of Theoretical and Applied Informatics, Polish Academy of Sciences. His research interests include numerical investigations in quantum information theory, with a main focus on developing novel and fast algorithms for simulating quantum systems.

Zbigniew Puchała received the M.Sc. degree in mathematics from the University of Wrocław, Poland, in 2003, the Ph.D. degree in mathematics from the Institute of Mathematics, University of Wrocław, in 2007, and the Habilitation degree in computer science from the Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice, Poland, in 2014. He is currently an Associate Professor with the Institute of Theoretical and Applied Informatics, Polish Academy of Sciences. His research interests include the intersection of mathematics and quantum information theory, with a main focus on the foundations of the theory of quantum information, for example, the study of quantum state and channels discrimination, entropic uncertainty relations, and geometrical methods in quantum information and computations.