



*entropy*



Article

---

# Superdense Coding Using Higher Dimensional Embedding

---






Elio Thadhani, Sharjeel Ahmad, Hussain Ali Razvi, Facundo Martin Lopez and Eric Chitambar



<https://doi.org/10.3390/e28040387>

## Article

# Superdense Coding Using Higher Dimensional Embedding

Elio Thadhani <sup>1</sup>, Sharjeel Ahmad <sup>2</sup>, Hussain Ali Razvi <sup>3</sup>, Facundo Martin Lopez <sup>4</sup> and Eric Chitambar <sup>5,\*</sup><sup>1</sup> Department of Physics, Harvey Mudd College, Claremont, CA 91711, USA; ethadhani@hmc.edu<sup>2</sup> Department of Physics, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA; ahmad18@illinois.edu<sup>3</sup> Department of Computer Science, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA; hrazvi2@illinois.edu<sup>4</sup> Department of Physics and Astronomy, Carleton College, Northfield, MN 55057, USA; lopezf@carleton.edu<sup>5</sup> Department of Electrical and Computer Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

\* Correspondence: echitamb@illinois.edu

## Abstract

Quantum dense coding is a foundational protocol in quantum communication, allowing two classical bits to be transmitted by sending a single qubit when a maximally entangled pair is shared. In this work, we consider Embedded Dense Coding (EDC)—a generalization of deterministic dense coding that embeds one subsystem into a higher-dimensional Hilbert space. To assess the operational advantage of EDC compared to standard dense coding, we consider the probability of transmission error when fixing the rate of entanglement consumed per classical message sent. We first demonstrate that EDC enables a smaller one-shot transmission error compared to standard dense coding when using quantum channels with nonzero rates of dephasing and loss. We then demonstrate that even with noiseless communication channels, EDC leads to smaller overall errors when the sender and receiver have noisy local processors. This advantage is shown through concrete implementations of EDC on IBM's Heron processor.

**Keywords:** quantum communication; superdense coding; quantum entanglement; quantum channel; qubit; qudit

## 1. Introduction

Entanglement is a central resource in quantum communication, enabling protocols such as quantum teleportation and quantum key distribution [1,2]. Shared entanglement, through a process called quantum dense coding, allows classical information to be transmitted more efficiently than is possible classically [3]. In its usual form, quantum dense coding allows Alice to transmit  $2 \log_2 d$  bits of classical data to Bob using a  $d$ -dimensional quantum channel combined with  $\log_2 d$  entangled bits (ebits) shared between them. Several generalizations of dense coding have been explored, including protocols with asymmetric dimensions, probabilistic protocols, and alternative resource trade-offs [3–8].

From a conceptual point of view, we can describe the resource trade-off of dense coding as [1,9]

$$\log d[\text{qq}] + \log d[\text{q} \rightarrow \text{q}] \geq 2 \log d[\text{c} \rightarrow \text{c}]. \quad (1)$$

Here,  $\log d[\text{qq}]$  expresses a total of  $\log d$  ebits, each ebit having the canonical two-qubit form  $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ , while  $\log d[\text{q} \rightarrow \text{q}]$  denotes a noiseless quantum channel that can transmit  $\log d$  qubits, and  $2 \log d[\text{c} \rightarrow \text{c}]$  expresses a noiseless classical channel that can



Academic Editor: Shenghao Yang

Received: 26 January 2026

Revised: 26 February 2026

Accepted: 12 March 2026

Published: 1 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

transmit  $2 \log d$  bits. The inequality means that having the communication resources on the LHS is sufficient for simulating the communication resources on the RHS. We refer to the dense coding protocol described in Equation (1) as Standard Dense Coding (SDC). The relevant feature of SDC for our purposes is that the entangled system on the sender's side has the same dimension as the input system to the quantum channel.

However, more general dense coding protocols, such as the one proposed in Ref. [4], can achieve the resource trade-off

$$\log d[\text{qq}] + \log D[\text{q} \rightarrow \text{q}] \geq \log dD[\text{c} \rightarrow \text{c}]. \quad (2)$$

In this case, the size of the entanglement shared between the sender and receiver has dimension  $d$  (more precisely, its Schmidt rank is  $d$ ), while the noiseless quantum channel transmits  $D$ -dimensional quantum states.

This paper considers dense coding scenarios similar to the one captured in Equation (2). Specifically, we are interested in the situation when  $d < D$  and exploring what advantages can be attained when the shared entanglement is smaller than the dimension of the connecting quantum channel. From a practical perspective, this is a very natural scenario to consider since storing high-dimensional entanglement over large distances can be very challenging, and thus in the foreseeable future all quantum communication networks will likely satisfy  $d \ll D$ .

In order to perform dense coding when  $d < D$ , the  $d$ -dimensional entanglement must be embedded into a larger  $D$ -dimensional system for transmission over the channel. We refer to protocols of this form as Embedded Dense Coding (EDC), and we describe such a protocol in this paper. Unlike previous works that studied higher-dimensional dense coding protocols [4,5], our interest lies in comparing the performance of EDC with SDC.

To make a fair comparison between different dense coding protocols that use the same  $D$ -dimensional quantum channel, we fix the entanglement-message rate of a given protocol, which is its *rate of entanglement consumed per classical message sent*. Here, we choose to quantify entanglement in terms of its Schmidt rank, which is a bona fide entanglement measure that is suitable for providing a coarse-grained description of entanglement in the one-shot setting [10,11]. Different dense coding protocols with the same entanglement-message rate are on the same level in terms of resource consumption, and thus it is fair to compare them in terms of other communication metrics such as transmission error. SDC and EDC protocols using a  $D$ -dimensional quantum channel have the same entanglement-message rate:

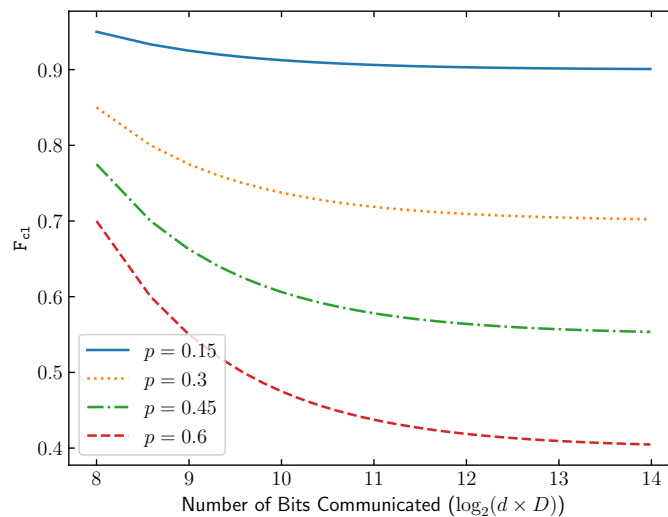
$$\frac{1}{D} = \frac{D}{D^2} = \frac{d}{dD}. \quad (3)$$

When the quantum channel is noiseless, both these protocols transmit classical messages at this rate without any transmission error. However, if we deviate from the ideal scenario and consider a noisy quantum channel, it is not clear whether SDC and EDC have the same performance in terms of decoding error. The fundamental question studied in this paper is whether EDC offers any advantage over SDC in noisy environments by reducing the transmission error at a fixed entanglement-message rate. We find that this is indeed possible when the quantum communication channel or the local encoding/decoding operations are noisy. Hence, in all realistic scenarios, EDC offers a communication advantage over SDC.

In this paper, we present an Embedded Dense Coding protocol which utilizes isometric embedding (Section 4) and verify that it enables the zero-error transmission of messages under noiseless conditions. We then analytically analyze the performance of EDC when the sender Alice communicates to the receiver Bob over a type of noisy quantum channel known as the dephasure channel (Section 5). The dephasure channel encompasses errors

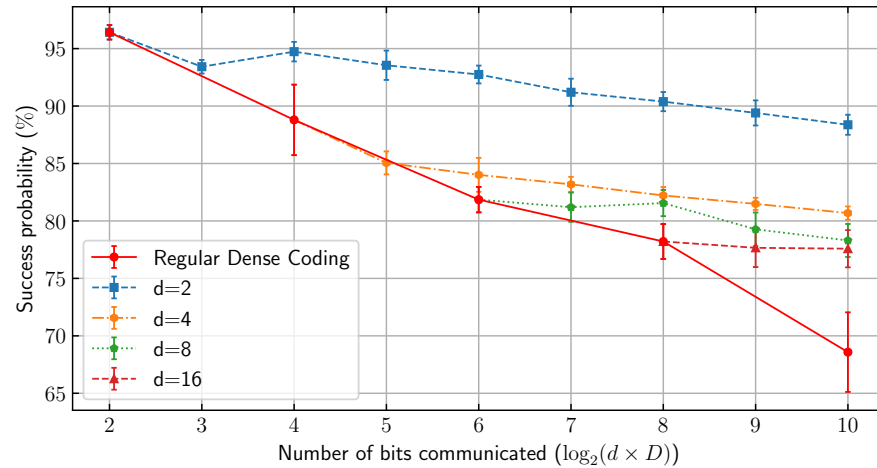
due to both dephasing and loss, and it provides a realistic model for studying quantum communication effects [12].

Our main findings are as follows. We first observe that the tolerance to dephasing error increases as the embedding dimension  $D$  increases. Moreover, at a fixed entanglement-message rate  $\frac{1}{D}$ , EDC always obtains a strictly larger transmission success probability in the presence of dephasing compared to SDC (see Proposition 1). This makes EDC a versatile protocol for one-shot entanglement-assisted classical communication, as it allows for sacrificing the number of bits transmitted in exchange for a higher one-shot success probability when the transmission is subject to dephasing noise (see Figure 1). We then experimentally analyze the performance of EDC when Alice and Bob have noisy local encoding/decoding and their shared entanglement is stored in noisy quantum memory (Section 6). This is done by implementing EDC on IBM’s Heron Processor. By providing an explicit construction, we show that EDC can be realized using a fewer number of encoding/decoding gates, which ultimately leads to a smaller transmission error when compared to conventional dense coding (see Figure 2).



**Figure 1.** Comparison of the Classical Correlation Fidelity  $F_{c1}$  and the number of bits sent for a given EDC protocol with fixed channel dimension  $D = 128$ . The dimension of entanglement,  $d$ , is varied to communicate  $\log_2(d \times D)$  classical bits. We show various probabilities of a dephasing  $p$  and assume an ideal resource as seen in Proposition 1. Here, we can see the trade-off that, using an EDC protocol, one can sacrifice the number of bits transmitted in exchange for a higher one-shot success probability if transmission is subject to dephasing noise.

Throughout this work, quantum systems will be denoted by  $A, B, C$ , etc., and classical systems will be denoted by  $X, Y, Z$ , etc. The Hilbert space associated with a bipartite system  $AB$  will be defined as  $\mathcal{H}_{AB} := \mathcal{H}_B \otimes \mathcal{H}_A$  [13]. For a quantum system  $A$  with associated Hilbert space  $\mathcal{H}_A$ , we denote its dimension by  $|A| := \dim \mathcal{H}_A$ . Systems whose representation differs only by primes have equal dimension, i.e.,  $|A| = |A'| = |A''|$ . Systems whose representation differs by a subscript, i.e.,  $A$  and  $A_e$ , do not necessarily have the same dimension and are used to differentiate the same system after some projection or embedding process. The space of operators acting on  $\mathcal{H}_A$  is defined as  $\mathcal{B}(\mathcal{H}_A)$ . A pure state  $|\psi\rangle_A \in \mathcal{H}_A$  has associated density operator  $\psi_A = |\psi\rangle\langle\psi|_A \in \mathcal{B}(\mathcal{H}_A)$ . The identity operator acting on  $\mathcal{H}_A$  is denoted  $\mathbb{I}_A$ , and the identity map on  $\mathcal{B}(\mathcal{H}_A)$  is denoted  $\text{id}_A$ . Mathematically, every quantum channel  $\mathcal{N}_{A \rightarrow B} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$  is described by a completely positive trace-preserving (CPTP) map.



**Figure 2.** Success probabilities for EDC schemes with varying dimensions implemented on IMB’s Heron processor through IBM<sub>Torino</sub>. Due to hardware limitations, logical qudits are encoded as physical qubits. We plot the success probability of a given EDC scheme that can communicate a set number of classical bits utilizing an entanglement resource  $\rho \in \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_A) : |A| = |B| = d$ . The embedding dimension  $D$  is varied to compare success probability based on the number of bits communicated and the starting resource. Each plotted branch represents a different starting resource dimension  $d$ , while the conventional dense coding branch utilizes  $d = D$ . For example, 6 bits can be communicated utilizing  $\{d = 8, D = 8\}$ ,  $\{d = 4, D = 16\}$ , and  $\{d = 2, D = 32\}$ . Each data point was collected over six independent trials, each consisting of 10,000 shots.

## 2. Background

The generalized Pauli operators provide a useful set of orthogonal operators that act on qudit ( $d$ -dimensional) quantum state space. These operators play a significant role in the creation of generalized communication protocols.

**Definition 1** (Pauli Operators [14]). Let the Generalized Pauli  $X$  and  $Z$  operators acting on a  $d$ -dimensional system  $\mathcal{H}_B$  be defined as  $X_d^m$  and  $Z_d^n$  where

$$X_d^m = \sum_{j=0}^{d-1} |j \oplus m\rangle\langle j| \quad \text{for } m = 0, \dots, d-1 \tag{4}$$

and

$$Z_d^n = \sum_{k=0}^{d-1} \omega^{kn} |k\rangle\langle k| : \omega = e^{i\frac{2\pi}{d}} \quad \text{for } m = 0, \dots, d-1. \tag{5}$$

### 2.1. Entanglement

Quantum entanglement has become a fundamental aspect of quantum information theory. References [15–18] help define different aspects of quantum entanglement, including discussions of maximally entangled pure states and entangled mixed states. In general, a bipartite state  $\rho_{AB}$  is considered to be unentangled if and only if it can be expressed as a convex combination of product states [19]. Throughout this text, we will focus primarily on utilizing maximally entangled pure states (MES) as communication resources. For  $d$ -dimensional systems  $C$  and  $C'$ , a state  $|\Phi\rangle_{CC'}$   $\in \mathcal{H}_C \otimes \mathcal{H}_{C'}$  is called maximally entangled if it has the form

$$|\Phi\rangle_{CC'} = U_C \otimes \mathbb{I} \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle_C \otimes |k\rangle_{C'} \tag{6}$$

where  $U_C$  is an arbitrary unitary operator. The dimension  $d$  is also called the Schmidt rank of state  $|\Phi\rangle$ . A canonical maximally entangled state is given by the choice  $U_C = \mathbb{I}$ , which we denote  $|\Phi^+\rangle_{CC'}$ .

As discussed by [18], for some bipartite state  $\rho_{AB}$ , we can define the Fully Entangled Fraction. The Fully Entangled Fraction (FEF) is notably not a valid entanglement measure but rather a lower bound on the amount of entanglement present for a given resource [20]. However, the FEF is still a meaningful quantity in the study of quantum communication protocols and is shown to provide useful information when evaluating entanglement applications [21].

**Definition 2** (FEF [18]). *The Fully Entangled Fraction of some state  $\rho \in \mathcal{B}(\mathcal{H}_C \otimes \mathcal{H}_{C'})$  is given by*

$$F_{|C|}(\rho) = \max_{\Phi} \langle \Phi | \rho | \Phi \rangle \tag{7}$$

where the maximum is over all  $|C|$ -dimension maximally entangled states  $|\Phi\rangle_{CC'} \in \mathcal{H}_C \otimes \mathcal{H}_{C'}$ .

### 2.2. Dense Coding

Standard dense coding (SDC), originally proposed by [3], enables Alice to communicate  $\log_2(d^2)$  classical bits by sending Bob one-half of a  $d$ -dimensional maximally entangled state over a noiseless quantum channel. In the ideal case, Alice and Bob have the maximally entangled state as a shared resource in advance. However, the notion of dense coding can be defined more generally for an arbitrary bipartite state  $\rho_{AB}$  and quantum channel  $\mathcal{N}_{A_e \rightarrow B_e}$ . For completion, we state the definition here.

**Definition 3.** *An  $N$  message dense coding protocol from Alice to Bob over channel  $\mathcal{N}_{A_e \rightarrow B_e}$  is defined in terms of a tuple  $\mathfrak{D} = (\rho_{AB}, \{\mathcal{E}^i\}_{i=0}^{N-1}, \{\Pi^j\}_{j=0}^{N-1})$ , where  $\rho_{AB}$  is a quantum state shared between Alice and Bob,  $\{\mathcal{E}^i\}_{i=0}^{N-1}$  is a family of encoding channels with  $\mathcal{E}^i : A \rightarrow A_e$ , and  $\{\Pi^j\}_{j=0}^{N-1}$  is a POVM acting on systems  $BB_e$ . The protocol goes as follows:*

1. *Alice and Bob each hold their respective subsystems of the bipartite entangled state  $\rho_{AB}$ , and Alice has a classical message  $i \in [N] := \{0, \dots, N - 1\}$  that she wishes to communicate to Bob.*
2. *Alice encodes the message  $i$  into her part of the entangled state  $\rho_{AB}$  using the encoding operator  $\mathcal{E}^i : A \rightarrow A_e$ .*
3. *Alice sends Bob subsystem  $A_e$  over channel  $\mathcal{N}_{A_e \rightarrow B_e}$ .*
4. *Bob measures the joint system  $BB_e$  using the POVM  $\{\Pi^j\}_{j=0}^{N-1}$  and guesses that Alice sent message  $j$  whenever he obtains measurement outcome  $j$ .*

Overall, this generates a classical channel  $W : [N] \rightarrow [N]$  with transition probabilities given by

$$W(j|i) = \text{Tr} \left[ \Pi^j \cdot \mathcal{N} \circ \mathcal{E}^i \otimes \text{id}_B(\rho_{AB}) \right]. \tag{8}$$

When the dimension of the channel input,  $D = |A_e|$ , is larger than the dimension of the system storing Alice’s entanglement,  $d = |A|$ , the encoders  $\mathcal{E}^i$  are embedding maps. In this work, we will be exclusively interested in isometric embeddings. These are easily constructed as follows. Let  $\{|\theta_i\rangle\}_{i=0}^{d-1}$  and  $\{|\phi_i\rangle\}_{i=0}^{D-1}$  both be an orthonormal basis for systems  $A$  and  $A_e$ , respectively, with  $d \leq D$ . Then, the isometric embedding operator  $V_{A \rightarrow A_e} : \mathcal{H}_A \mapsto \mathcal{H}_{A_e}$  can be constructed as

$$V_{A \rightarrow A_e} = \sum_{i=0}^{d-1} |\phi_i\rangle \langle \theta_i|. \tag{9}$$

For convenience, we denote the action of the isometry on density matrices as the CPTP map  $\mathcal{V}(\rho_{AB}) = V(\rho_{AB})V^\dagger$ .

### 3. Performance Criteria

In this work, we evaluate the performance of a dense coding protocol in terms of its average success probability in transmitting classical messages from Alice to Bob, assuming a uniform prior distribution. We refer to this as the Classical Correlation Fidelity,  $F_{c1}$ .

**Definition 4** (Classical Correlation Fidelity [13]). *For a classical channel  $W : [N] \rightarrow [N]$  with transition probabilities  $W(j|i)$ , its Classical Correlation Fidelity  $F(W) \equiv F_{c1}$  is defined as*

$$F_{c1} = \frac{1}{N} \sum_{i=0}^{N-1} W(i|i). \tag{10}$$

For a dense coding protocol  $\mathcal{D}$  defined in Definition 3, its Classical Correlation Fidelity is given by

$$F_{c1} = \frac{1}{N} \sum_{i=0}^{N-1} \text{Tr} \left[ \Pi^i \cdot \mathcal{N} \circ \mathcal{E}^i \otimes \text{id}_B(\rho_{AB}) \right]. \tag{11}$$

We stress that the choice of performance metric chosen here is best suited for the one-shot setting. This refers to the fact that we are considering dense coding protocols that encode/decode over a single use of both the resource state  $\rho_{AB}$  and the channel  $\mathcal{N}_{A_e \rightarrow B_e}$ . In contrast, one could consider applying block-encoding across multiple uses of the resource state or the quantum channel, which is a more familiar information-theoretic problem. When performing asymptotic block encoding, the more relevant figure of merit for classical communication is the Holevo information [22] or the quantum mutual information [23] for entanglement-assisted channel coding.

For example, prior work on Hiroshima has identified dense coding protocols that maximize the Holevo information for a given resource state  $\rho_{AB}$  [8]. The construction presented in that work involves a group-covariant unitary encoding. Rather remarkably, this encoding/decoding strategy is optimal for every state  $\rho_{AB}$  with fixed dimensions  $|A|$  and  $|B|$ . However, the construction will not necessarily be optimal when considering the Classical Correlation Fidelity  $F_{c1}$  as a figure of merit, and, in general, optimizing  $F_{c1}$  requires solving a bilinear optimization problem (see Ref. [13] for details). Nevertheless, the EDC protocol we propose in the next section has the same symmetric form as Hiroshima’s, and it turns out to indeed optimize  $F_{c1}$  for the special case of sending half a maximally entangled state over the dephasure channel.

### 4. Embedded Dense Coding

We now describe an Embedded Dense Coding (EDC) protocol from Alice to Bob, which is defined for a shared resource state  $\rho_{AB}$  and any channel  $\mathcal{N}_{A_e \rightarrow B_e}$  with  $|A| = |B| = d$  and  $|A_e| = |B_e| = D > d$ . In this protocol, Alice attempts to send Bob one of  $d \cdot D$  possible messages  $\mu \in \{(k, j) : k \in [d], j \in [D]\}$ . The mapping of system  $A$  into  $A_e$  is done via an isometric embedding.

Alice first performs the generalized Pauli  $Z_A^k$  operation on her subsystem, encoding the  $k$ -dit element of his message. As a result, the joint bipartite system is now

$$\rho_k = (\mathbb{I} \otimes Z_A^k) \rho (\mathbb{I} \otimes Z_A^{k\dagger}). \tag{12}$$

She then embeds her subsystem into a higher-dimensional space  $\mathcal{H}_{A_e}$  using an isometric embedding  $\mathcal{V}$ , as described in Equation (9). We write  $\varrho := \mathcal{V}(\rho)$  to denote the embedded

form of any density matrix  $\rho$ . However, note that it is equivalent for Alice to encode  $k$  after the embedding process using the operator

$$\hat{Z}_{A_e}^k = V_{A \rightarrow A_e} Z_A^k V_{A \rightarrow A_e}^\dagger. \tag{13}$$

That is,  $\rho_k = \hat{Z}^k \rho \hat{Z}^{k\dagger}$ . We will utilize this alternative convention for notational simplicity.

Alice next encodes the index  $j$  of her message by performing a generalized Pauli  $X_{A_e}^j$  operation on her subsystem. The resulting joint state is now

$$\rho^\mu := (\mathbb{I} \otimes X_{A_e}^j \hat{Z}_{A_e}^k) \rho (\mathbb{I} \otimes \hat{Z}_{A_e}^{k\dagger} X_{A_e}^{j\dagger}). \tag{14}$$

In total, the encoder for message  $\mu = (k, j)$  is given by the CPTP map  $\mathcal{E}^\mu : A \rightarrow A_e$  given by

$$\mathcal{E}^\mu(\cdot) = X_{A_e}^j \hat{Z}_{A_e}^k (\cdot) (X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger. \tag{15}$$

Note that this encoder is an isometry and the set of encoding operators  $X_{B_e}^j \hat{Z}_{B_e}^k$  are mutually orthogonal. This is consistent with the encoding strategy that maximizes the Holevo information [8].

The orthogonality of the  $X_{A_e}^j \hat{Z}_{A_e}^k$  naturally suggests the following strategy for decoding. First, let us denote the encoded states as

$$\rho^\mu = \text{id} \otimes \mathcal{E}^\mu \circ \mathcal{V}(\rho). \tag{16}$$

Notice that when choosing  $\rho = |\Psi\rangle\langle\Psi|$  to be maximally entangled with Schmidt rank  $d$ , the encoded states  $\rho^\mu$  form an orthonormal set. We can therefore define a decoding POVM for Bob by

$$\Pi_{BB_e}^\mu = \text{id} \otimes \mathcal{E}^\mu \circ \mathcal{V}(|\Psi\rangle\langle\Psi|), \tag{17}$$

plus an additional effect  $\mathbb{I} - \sum_\mu \Pi_{BB_e}^\mu$ , which projects onto the space orthogonal to all the encoded states. Observe that this construction leads to an error-free transmission of all  $d \cdot D$  messages when all resources and communication are noiseless since

$$\text{tr}[\Pi_{BB_e}^\mu \rho^\mu] = \text{tr}[\Pi_{BB_e}^\mu \text{id} \otimes \mathcal{E}^\mu \circ \mathcal{V}(|\Psi\rangle\langle\Psi|)] = \text{tr}[|\Psi\rangle\langle\Psi|] = 1, \quad \forall \mu. \tag{18}$$

While the EDC protocol here is clearly optimal when using a noiseless quantum channel  $\text{id}_{A_e \rightarrow B_e}$ , it is also shown in Appendix A to be optimal for certain types of noisy channels and resources.

**Remark 1.** *There is freedom in the choice of maximally entangled state  $|\Psi\rangle$  when defining the decoder in an EDC protocol. For communication over noisy channels, this choice will affect the Classical Correlation Fidelity  $F_{c1}$ . Hence, in general, one will need to optimize over the choice of  $|\Psi\rangle$  to maximize  $F_{c1}$ .*

For a general bipartite state  $\rho_{AB}$  and a noiseless quantum channel  $\text{id}_{A_e \rightarrow B_e}$ , the Classical Correlation Fidelity of EDC can be expressed in terms of the singlet fraction.

**Lemma 1.** *EDC over a  $D$ -dimensional noiseless quantum channel using bipartite state  $\rho_{AB}$  with  $d = |A| = |B|$  satisfies  $F_{c1} = F_d(\rho)$ .*

**Proof.** Let  $|\Psi\rangle$  be the  $d$ -dimensional maximally entangled state such that  $F_d(\rho) = \langle\Psi|\rho|\Psi\rangle$ . We use this state to define the decoding POVM in EDC, as in Equation (17). Then for every message  $\mu$ , we have

$$\begin{aligned}
 p(\mu|\mu) &= \text{Tr}[\Pi^\mu \rho^\mu] \\
 &= \text{Tr}[\mathcal{E}^\mu \circ \mathcal{V}(|\Psi\rangle\langle\Psi|) \mathcal{E}^\mu \circ \mathcal{V}(\rho)] \\
 &= \langle \Psi | \rho | \Psi \rangle = F_d(\rho),
 \end{aligned}
 \tag{19}$$

where we have used the fact that the EDC encoder and decoder are isometries, and they therefore preserve inner products. Hence, by the definition of the Classical Correlation Fidelity  $F_{c1}$ , we have  $F_{c1} = F_d(\rho)$ .  $\square$

### 5. Noisy Transmission: Dephasure Quantum Channel

Having established that the EDC protocol enables perfect transmission of  $d \cdot D$  classical messages in the case of noiseless communication resources, we now turn to the question whether EDC is robust to noise. We are specifically interested in comparing the performance of EDC to standard dense coding (SDC). Suppose Alice can send Bob quantum information over a  $D$ -dimensional quantum channel. Recall that SDC involves using an entangled system that has the same dimension as the quantum channel connecting Alice to Bob. We are interested in understanding whether or not there is any advantage in using entanglement and channels with mismatched dimensions, thereby leveraging an EDC protocol. This is a highly practical question since storing high-dimensional entanglement for use in distributed communication protocols is experimentally demanding. We demonstrate that indeed an advantage in using EDC can be found when communication is subject to dephasing and erasure.

More precisely, we consider that Alice and Bob are connected by a dephasure channel [12], which is a two-parameter CPTP map transforming  $A_e \rightarrow B_e$  and having the form

$$\mathcal{N}_{p,q}(\rho) = (1 - q)((1 - p)\rho + \frac{p}{D} \sum_{\ell=0}^{D-1} Z^\ell \rho Z^{\ell\dagger}) + q \text{Tr}(\rho) |e\rangle\langle e|.
 \tag{20}$$

with  $D = |A_e|$ . Here,  $|e\rangle$  is a vector orthogonal to the support of  $\rho$ , serving as an erasure flag that the receiver can detect. The parameter  $q \in [0, 1]$  represents the probability of erasure, while  $p \in [0, 1]$  represents the probability of completely dephasing the state.

Standard dense coding over a  $D$ -dimensional dephasure channel is performed on a  $D \times D$  state  $\rho_{AB}^{(D)}$  and attempts to transmit  $D^2$  messages. EDC over the same channel uses a  $d \times d$  state  $\rho_{AB}^{(d)}$  with  $d < D$  and sends a correspondingly smaller number of  $d \cdot D$  messages. Note that in both these scenarios, SDC and EDC have the same entanglement-message rate of  $\frac{1}{D}$ . They are thus comparable in terms of resource consumption. Nevertheless, as one of our main findings, we observe that they differ in their average transmission error.

**Proposition 1.** *For a  $D$ -dimensional dephasure channel  $\mathcal{N}_{p,q}$ , SDC and EDC protocols using arbitrary entangled resources  $\rho^{(D)}$  and  $\rho^{(d)}$ , respectively, have their Classical Correlation Fidelities upper bounded as*

$$F_{c1}^{(SDC)} \leq (1 - q)((1 - p)F_D(\rho^{(D)}) + \frac{p}{D}) + \frac{q}{D^2}
 \tag{21}$$

$$F_{c1}^{(EDC)} \leq (1 - q)((1 - p)F_d(\rho^{(d)}) + \frac{p}{d}) + \frac{q}{dD}.
 \tag{22}$$

Moreover, these upper bounds can be made tight by using a maximally entangled resource state. In this case, EDC has a greater  $F_{c1}$  whenever  $p \neq 0$ .

**Proof of Proposition 1.** We prove the upper bound of Equation (22); the argument for Equation (21) follows analogously. As in the proof for Lemma 1, let  $|\Psi\rangle$  be such that

$F_d(\rho) = \langle \Psi | \rho | \Psi \rangle$ . We use this state to build the decoder of an EDC protocol with the decoding POVM given by Equation (17).

Starting with the state  $\rho^{(d)}$ , let  $\rho^\mu$  denote the state after Alice applies encoder  $\mathcal{E}^\mu \circ \mathcal{V}$  for message  $\mu$ . System  $A_e$  is then subject to the dephasing channel  $\mathcal{N}_{p,q}$ , yielding the transformed state  $\hat{\rho}^\mu$ . Thus, the final state held by Bob prior to the decoding measurement is

$$\begin{aligned} \hat{\rho}^\mu &= (\text{id}_B \otimes \mathcal{N}_{p,q})(\rho^\mu) \\ &= (1-q) \left( (1-p)(\text{id} \otimes \mathcal{E}^\mu)(\rho) \right. \\ &\quad \left. + \frac{p}{D} \sum_{\ell=0}^{D-1} (\mathbb{I} \otimes Z_{A_e}^\ell)(\text{id} \otimes \mathcal{E}^\mu)(\rho)(\mathbb{I} \otimes Z_{A_e}^{\ell\dagger}) \right) + q\rho_B \otimes |e\rangle\langle e|, \end{aligned} \tag{23}$$

where  $\rho_B = \text{tr}_A[\rho^{(d)}]$ .

**Lemma 2.** *The encoding operator  $\mathcal{E}^\mu(\cdot) = X_{A_e}^j \hat{Z}_{A_e}^k(\cdot)(X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger$  defined in Equation (15) commutes under conjugation with any generalized canonical Pauli  $Z_{A_e}$  operator as defined in Definition 1.*

**Proof.** Let  $\mathcal{E}^\mu(\cdot) = X_{A_e}^j \hat{Z}_{A_e}^k(\cdot)(X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger$  and  $Z_{A_e}^\ell$  be some generalized pauli operator such that  $Z_{A_e}^\ell \in \mathcal{B}(A_e)$ . Then,

$$Z_{A_e}^\ell \mathcal{E}^\mu(\cdot) Z_{A_e}^{\ell\dagger} = Z_{A_e}^\ell X_{A_e}^j \hat{Z}_{A_e}^k(\cdot)(X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger Z_{A_e}^{\ell\dagger}. \tag{24}$$

Note, by the Weyl Commutation relation,  $Z_{A_e}^\ell X_{A_e}^j = \omega Z_{A_e}^\ell X_{A_e}^j$  for some  $\omega \in \mathbb{C}$  where  $|\omega| = 1$ . Additionally note that  $[Z_{A_e}^\ell, \hat{Z}_{A_e}^k] = 0$  since both  $Z_{A_e}^\ell$  and  $\hat{Z}_{A_e}^k$  are diagonal. It therefore follows that

$$\begin{aligned} Z_{A_e}^\ell \mathcal{E}^\mu(\cdot) Z_{A_e}^{\ell\dagger} &= Z_{A_e}^\ell X_{A_e}^j \hat{Z}_{A_e}^k(\cdot)(X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger Z_{A_e}^{\ell\dagger} \\ &= X_{A_e}^j \hat{Z}_{A_e}^k Z_{A_e}^\ell(\cdot)(X_{A_e}^j \hat{Z}_{A_e}^k Z_{A_e}^\ell)^\dagger = \mathcal{E}^\mu(Z_{A_e}^\ell(\cdot) Z_{A_e}^{\ell\dagger}). \end{aligned} \tag{25}$$

□

By Lemma 2, we know that the encoding operator  $\mathcal{E}^\mu$  commutes with the Pauli Z twirling channel; thus

$$\begin{aligned} \hat{\rho}^\mu &= (\text{id}_A \otimes \mathcal{N}_{p,q})(\rho^\mu) \\ &= (1-q) \left( (1-p)(\text{id} \otimes \mathcal{E}^\mu)\rho + \frac{p}{D} (\text{id} \otimes \mathcal{E}^\mu) \left( \sum_{\ell=0}^{D-1} (\mathbb{I} \otimes Z_{A_e}^\ell)\rho(\mathbb{I} \otimes Z_{A_e}^{\ell\dagger}) \right) \right) \\ &\quad + q\rho_B \otimes |e\rangle\langle e|. \end{aligned} \tag{26}$$

The general effect of Pauli Z twirling on an arbitrary bipartite state  $\rho_{BA_e}$  is given by

$$\rho_{BA_e} \mapsto \frac{1}{|A_e|} \sum_{\ell=0}^{|A_e|-1} (\mathbb{I} \otimes Z^\ell)\rho_{BA_e}(\mathbb{I} \otimes Z^{\ell\dagger}) = \sum_{\kappa=0}^{|A_e|-1} \rho_B^\kappa \otimes |\kappa\rangle\langle\kappa|_{A_e} \tag{27}$$

where  $\rho_B^\kappa := \text{Tr}_{A_e}[(\mathbb{I}_B \otimes |\kappa\rangle\langle\kappa|)\rho_{BA_e}]$ . Substituting this into Equation (26) yields

$$\hat{\rho}^\mu = (1-q)(\text{id} \otimes \mathcal{E}^\mu) \left( (1-p)\rho + \frac{p}{D} \left( \sum_{\kappa=0}^{D-1} \rho_B^\kappa \otimes |\kappa\rangle\langle\kappa|_{A_e} \right) \right) + q\rho_B \otimes |e\rangle\langle e|. \tag{28}$$

We can now calculate the Classical Correlation Fidelity, which we recall is defined as

$$F_{c1} = \frac{1}{dD} \sum_{\mu=0}^{dD-1} p(\mu|\mu). \tag{29}$$

By the symmetry of states and assuming equal priors, we have that  $F_{c1} = p(\mu|\mu)$  for all  $\mu$ . Furthermore, note that our encoding operators preserve the inner product on the pre-image of our embedding, and, in the case of an erasure flag,  $|e\rangle\langle e|$ , we will assume the message is randomly guessed with success probability  $1/dD$ . Thus,

$$\begin{aligned} F_{c1}^{(EDC)} &= p(\mu|\mu) \\ &= \text{Tr} \left[ \hat{\rho}^\mu \Pi_{BB_e}^\mu \right] \\ &= (1-q)(1-p)F_d(\rho) + (1-q)p \text{Tr} \left[ \left( \sum_{\kappa=0}^{D-1} \rho_B^\kappa \otimes |\kappa\rangle\langle\kappa|_{A_e} \right) \mathcal{V}(|\Psi\rangle\langle\Psi|) \right] + \frac{q}{dD} \end{aligned} \tag{30}$$

$$\leq (1-q)(1-p)F_d(\rho) + \frac{1}{d}(1-q)p + \frac{q}{dD}, \tag{31}$$

where the upper bound follows from the fact that

$$\sum_{\kappa=0}^{D-1} \text{Tr} \left[ \rho_B^\kappa \otimes |\kappa\rangle\langle\kappa|_{A_e} \mathcal{V}(|\Psi\rangle\langle\Psi|) \right] \leq \sum_{\kappa=0}^{D-1} \frac{1}{d} \langle \kappa | \rho_{A_e} | \kappa \rangle = \frac{1}{d} \tag{32}$$

since the largest overlap of a  $d$ -dimensional maximally entangled state with a product state is  $\frac{1}{d}$ .

We now show that the upper bound in Equation (31) is tight for when using the maximally entangled state  $|\Psi\rangle$  as the resource state. In this case, observe that

$$\rho_B^\kappa = \text{Tr}_{A_e} [(\mathbb{I}_B \otimes |\kappa\rangle\langle\kappa|) \mathcal{V}(|\Psi\rangle\langle\Psi|)] = \frac{1}{d} \mathcal{V}^\dagger(|\kappa\rangle\langle\kappa|)^*. \tag{33}$$

Substituting this into Equation (30) gives

$$\begin{aligned} F_{c1}^{(EDC)} &= (1-q)(1-p) + (1-q) \frac{p}{d} \text{Tr} \left[ \left( \sum_{\kappa=0}^{D-1} \mathcal{V}^\dagger(|\kappa\rangle\langle\kappa|)^* \otimes \mathcal{V}^\dagger(|\kappa\rangle\langle\kappa|) \right) |\Psi\rangle\langle\Psi| \right] + \frac{q}{dD} \\ &= (1-q) \left[ (1-p) + \frac{p}{d} \right] + \frac{q}{dD}. \end{aligned} \tag{34}$$

Note that SDC is the special case of EDC where  $d = D$ . A depiction of communication fidelity to message size trade-off can be seen in Figure 1.  $\square$

*Noisy Singlet Resource*

To show a more general result, we will examine the Classical Correlation Fidelity of EDC and SDC in the case where Alice and Bob’s shared resource is a noisy singlet of the form

$$\rho^{(d)} = \theta |\Psi\rangle\langle\Psi| + (1-\theta) \frac{\mathbb{I} \otimes \mathbb{I}}{d^2} \text{ such that } 0 < \theta \leq 1. \tag{35}$$

Observe that

$$\rho_B^\kappa = \text{Tr}_{A_e} [(\mathbb{I}_B \otimes |\kappa\rangle\langle\kappa|) \mathcal{V}(|\Psi\rangle\langle\Psi|)] = \frac{1}{d} \mathcal{V}^\dagger(|\kappa\rangle\langle\kappa|)^*. \tag{36}$$

Substituting this into Equation (30) gives

$$\begin{aligned}
 F_{\text{cl}}^{(\text{EDC})} &= (1-q)(1-p) + (1-q)\frac{p}{d} \text{Tr} \left[ \left( \sum_{\kappa=0}^{D-1} \mathcal{V}^\dagger(|\kappa\rangle\langle\kappa|)^* \otimes \mathcal{V}^\dagger(|\kappa\rangle\langle\kappa|)|\Psi\rangle\langle\Psi| \right) \right] + \frac{q}{dD} \\
 &= (1-q)\theta[(1-p) + p\frac{1}{d}] + (1-q)(1-\theta)\frac{1}{d^2} + \frac{q}{dD}.
 \end{aligned} \tag{37}$$

We can therefore say, the Classical Correlation Fidelity when using a one parameter noisy singlet state is

$$F_{\text{cl}}^{(\text{EDC})} = (1-q)((1-p) + \theta\frac{p}{d} + (1-\theta)\frac{p}{d^2}) + \frac{q}{dD}. \tag{38}$$

Furthermore, since SDC is a special case of EDC where  $d = D$  we can generalize to say

$$F_{\text{cl}}^{(\text{SDC})} = (1-q)((1-p) + \theta\frac{p}{D} + (1-\theta)\frac{p}{D^2}) + \frac{q}{dD}. \tag{39}$$

This more general result demonstrates that the advantage identified in Proposition 1 are not exclusive to the ideal resource case but generalized to additional non-ideal resource states. This extension of our result strengthens our claim that EDC provides a general communication advantage to SDC. Note that, as shown in Appendix A, this represents the optimal fidelity achievable by the noisy singlet resource since the utilized POVM is optimized.

## 6. Noisy Local Operations: Hardware Implementation

We now evaluate EDC in the case where Alice and Bob are connected by a noiseless quantum channel but have limited local capabilities and are susceptible to local errors. To accomplish this analysis, we implement a modified version of EDC on current-day quantum hardware. We find that EDC serves as a more efficient alternative to SDC in terms of transmission success probability when implemented using non-ideal noisy local systems.

### 6.1. Decoding Gate Volume and Noise Accumulation

We implemented an adapted Embedded Dense Coding (EDC) scheme on IBM's Heron processor to explore its advantages in noisy settings. Throughout this section, we describe the protocol in terms of logical qudits. On current superconducting hardware, these logical qudits are realized using collections of physical qubits; the details of this encoding are discussed in Section 6.2. This section evaluates the relationship between the resulting gate volume and the accumulation of noise during execution. Experimentally, it is hard to isolate the impact from a specific type of quantum noise, such as dephasing. Results in this section will therefore include the accumulation of all types of noise affecting modern quantum hardware.

Practically, the primary source of noise for protocols on current hardware is the use of two-qubit gates, typically with a probability of error on the order of  $10^{-3}$  per control Z gate (basis two-qubit gate) on IBM Eagle and Heron processors [24]. In dense coding protocols, these are used for the preparation of the initial entangled state and Alice's decoding of the received message. Therefore, the quantity of these two-qubit operations scales linearly with the dimension of the entanglement dimension  $d$ , contributing directly to the accumulated physical error.

Embedded Dense Coding reduces this overhead by using an embedded resource state, which lowers the entanglement dimension  $d$  required for transmitting the same effective message dimension. Because the total accumulated error grows approximately

linearly with the number of noisy two-qudit gates, a reduction in decoding depth yields a proportional reduction in the internal hardware noise.

Thus, EDC's reduced decoding Gate Volume translates to a lower expected error contribution from native device noise without requiring any change to the physical hardware. This expectation is supported by the results of our implementation on IBM's Heron processor through IBM<sub>Torino</sub> in Figure 2.

### 6.2. Implementation and Results

Due to modern hardware constraints, we utilize multiple physical qubits to encode logical qudits rather than employing native higher-dimensional physical qudits. This means that, to simulate a  $D$ -dimensional qudit, we use  $\log_2(D)$  qubit channels, which have a total of  $D$  possible configurations. For example, a ququart is encoded using two qubit channels as such:

$$|0\rangle \mapsto |00\rangle \quad |1\rangle \mapsto |01\rangle \quad |2\rangle \mapsto |10\rangle \quad |3\rangle \mapsto |11\rangle \quad (40)$$

Further adaptations followed from qubit encoding, such as how Pauli gates and higher-dimensional embeddings are applied, which can be found in Appendix B.

Using this implementation, we conducted a comparison of success probabilities when transmitting the same number of classical bits of information via EDC protocols with various entanglement dimensions  $d$ . We performed six independent trials of 10,000 shots per trial, which yielded an average variance between trials of approximately 1.2%, confirming the stability of the simulation. The results of these tests can be seen in Figure 2, where each trajectory reflects a distinct embedding strategy and illustrates the associated success probability as the number of encoded messages increases. Figure 2 highlights the inherent trade-off between success rate and entanglement dimension introduced by the embedding process. Specifically, the observed divergence in success probabilities provides insight into the performance constraints encountered when attempting to encode messages in extended Hilbert space architectures.

Although these results are indicative of EDC advantages, due to the difficulty of implementing true qudit quantum codes, we can not make concrete assertions about EDC's performance if implemented via true qudit quantum codes. These results have been included due to their suggestive nature but require additional exploration in the future as quantum hardware advances.

## 7. Conclusions

Embedded dense coding is a generalization of deterministic dense coding, which embeds one subsystem into a higher-dimensional Hilbert space, increasing the classical capacity of an entanglement resource. In this work, we compare the performance of EDC with SDC and find that EDC offers fidelity advantages and additional flexibility when compared to SDC.

To assess performance in realistic conditions, we first evaluated EDC under a dephasing channel acting on the transmitted system while assuming ideal local operations. To enable a meaningful comparison with SDC, we fixed the classical communication rate per shared entangled resource and compared the resulting one-shot Classical Correlation Fidelity. Under this criterion, we show that when the entanglement dimension is strictly smaller than the channel dimension, EDC achieves a lower transmission error than SDC in the presence of dephasing and loss. This result establishes a clear noise-rate trade-off: decreasing the entanglement-to-channel dimension ratio reduces the number of distinguishable messages but improves robustness to dephasing.

We additionally implemented an adapted EDC scheme on IBM's Heron processors to examine its behavior under contemporary hardware noise. In this setting, the communication channel itself is effectively noiseless, and imperfections arise primarily from gate errors and circuit depth. Consequently, the experimental study does not directly simulate the dephasing model analyzed theoretically. Instead, it evaluates how the structural differences between EDC and SDC influence accumulated hardware error. We observe that embeddings requiring reduced decoding gate volume yield higher empirical success probabilities compared to conventional dense coding circuits of comparable dimension. Although constrained by the lack of native qudit support, these results demonstrate that the architectural features of EDC can translate into measurable performance gains on current qubit-based platforms.

Embedded Dense Coding provides a dimension-expanding reformulation of dense coding that preserves optimal noiseless performance while introducing a tunable trade-off between communication rate and noise robustness. Analytically, it reduces transmission error relative to SDC under dephasing noise at a fixed rate, and experimentally, it benefits from reduced decoding complexity under hardware noise. Together, these results clarify the operational role of entanglement dimension in dense coding protocols and highlight embedding as a meaningful design parameter for near-term quantum communication settings.

**Author Contributions:** Conceptualization, E.T., S.A., H.A.R., F.M.L. and E.C.; methodology, E.T.; software, H.A.R. and F.M.L.; validation, E.T., S.A., H.A.R. and F.M.L.; formal analysis, E.T., S.A. and H.A.R.; investigation, E.T., S.A., H.A.R. and F.M.L.; resources, E.C.; data curation, E.T. and F.M.L.; writing—original draft preparation, E.T., S.A., H.A.R. and F.M.L.; writing—review and editing, E.T., S.A., H.A.R., F.M.L. and E.C.; visualization, E.T. and F.M.L.; supervision, E.C.; project administration, S.A. and E.C.; funding acquisition, E.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded in part by the National Science Foundation Research Experience for Undergraduates (NSF-REU) under Grant No. PHY-2244433. This work was supported in part by the National Science Foundation through the HQAN Quantum Leap Challenge Institute (QLCI) under the Open Quantum Initiative Undergraduate Fellowship program. Additional support was received from the University of Illinois at Urbana-Champaign.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

**Acknowledgments:** The authors thank Felix Leditzky for his guidance, discussions, and valuable feedback throughout the development of this work.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

EDC	Embedded Dense Coding
SDC	Standard Dense Coding
FEF	Fully Entangled Fraction
CCF	Classical Correlation Fidelity
MES	Maximally Entangled State

### Appendix A. Optimality Constraints

We will show that the physically motivated POVM utilized in EDC is optimal for any dephrasing channel  $\mathcal{N}_{p,q}$  if Alice and Bob share an entangled resource of the form of the canonical one-parameter noisy singlet:

$$\rho^{(d)} = \theta|\Psi\rangle\langle\Psi| + (1 - \theta)\frac{\mathbb{I} \otimes \mathbb{I}}{d^2} \text{ such that } 0 < \theta \leq 1. \tag{A1}$$

**Proof.** As shown by [25], a necessary and sufficient condition for some POVM  $\{\Pi^{\mu_i}\}_{i=0}^{dD-1}$  to be optimal in the state discrimination problem  $\{e^{\mu_i}\}_{i=0}^{dD-1}$  is for

$$\sum_i p_i \hat{e}^{\mu_i} \Pi^{\mu_i} - p_j \hat{e}^{\mu_j} \geq 0, \quad \forall j. \tag{A2}$$

This requirement is equivalent to the condition

$$\langle\psi|\frac{1}{dD}\sum_i \hat{e}^{\mu_i} \Pi^{\mu_i}|\psi\rangle \geq \frac{1}{dD}\langle\psi|\hat{e}^{\mu_j}|\psi\rangle \quad \forall|\psi\rangle, \forall j. \tag{A3}$$

Note that, without loss of generality, since any POVM that detects erasure will perform equally in the case of erasure, we can assume  $q = 0$ . Starting with the LHS, we can use our result from Equation (28) to say,

$$\begin{aligned} \text{LHS} &= \langle\psi|\frac{1}{dD}\sum_i \hat{e}^{\mu_i} \Pi^{\mu_i}|\psi\rangle \\ &= \frac{1}{dD}\langle\psi|\sum_{i=0}^{dD-1} (\text{id} \otimes \mathcal{E}^{\mu_i})\left((1-p)\mathcal{V}(\rho^{(d)}) + p\left(\sum_{\kappa=0}^{D-1} e_B^\kappa \otimes |\kappa_{A_e}\rangle\langle\kappa_{A_e}|\right)\right)(\text{id} \otimes \mathcal{E}^{\mu_i})\mathcal{V}(\Psi)|\psi\rangle \\ &= \frac{1}{dD}\langle\psi|\sum_{i=0}^{dD-1} (\text{id} \otimes \mathcal{E}^{\mu_i})\left[\left((1-p)\mathcal{V}(\rho^{(d)})\right.\right. \\ &\quad \left.\left.+ p\left(\sum_{\kappa=0}^{D-1} e_B^\kappa \otimes |\kappa_{A_e}\rangle\langle\kappa_{A_e}|\right)\right)(\mathbb{I} \otimes (X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger (X_{A_e}^j \hat{Z}_{A_e}^k))\mathcal{V}(\Psi)\right]|\psi\rangle \end{aligned} \tag{A4}$$

Note that  $\mathbb{I} \otimes (X_{A_e}^j \hat{Z}_{A_e}^k)^\dagger (X_{A_e}^j \hat{Z}_{A_e}^k) = VV^\dagger$  is the projection operator onto the image  $\text{Im}(V)$ , and since  $\mathcal{V}(\Psi)$  is in the image of  $V$ , we can say,

$$\begin{aligned} \text{LHS} &= \frac{1}{dD}\langle\psi|\sum_{i=0}^{dD-1} (\text{id} \otimes \mathcal{E}^{\mu_i})\left[\left((1-p)\mathcal{V}(\rho^{(d)}) + p\left(\sum_{\kappa=0}^{D-1} e_B^\kappa \otimes |\kappa_{A_e}\rangle\langle\kappa_{A_e}|\right)\right)\mathcal{V}(\Psi)\right]|\psi\rangle \\ &= \frac{1}{dD}\langle\psi|\sum_{i=0}^{dD-1} (\text{id} \otimes \mathcal{E}^{\mu_i})\left[\left((1-p)\mathcal{V}(\rho^{(d)})\mathcal{V}(\Psi) + p\left(\sum_{\kappa=0}^{D-1} e_B^\kappa \otimes |\kappa_{A_e}\rangle\langle\kappa_{A_e}|\right)\mathcal{V}(\Psi)\right)\right]|\psi\rangle \\ &= \frac{1}{dD}\langle\psi|\sum_{i=0}^{dD-1} (\text{id} \otimes \mathcal{E}^{\mu_i})\left[\left((1-p)\mathcal{V}(\theta\Psi + \frac{(1-\theta)}{d^2}\Psi) + p\mathcal{V}\left(\frac{\theta}{d}\Psi + \frac{(1-\theta)}{d^2}\Psi\right)\right)\right]|\psi\rangle \\ &= \frac{1}{dD}\langle\psi|\sum_{i=0}^{dD-1} (\text{id} \otimes \mathcal{E}^{\mu_i})\left[\mathcal{V}\left(\theta\left(1-p + \frac{p}{d}\right)\Psi + \frac{(1-\theta)}{d^2}\Psi\right)\right]|\psi\rangle \end{aligned} \tag{A5}$$

Note, that  $\{(\text{id} \otimes \mathcal{E}^{\mu_i})\mathcal{V}(\Psi)\}_{i=1}^{dD-1}$  forms a complete basis in our Hilbert space. Thus,

$$\text{LHS} = \frac{1}{dD}\left(\theta - \theta p + p\frac{\theta}{d} + \frac{(1-\theta)}{d^2}\right). \tag{A6}$$

We can now evaluate the RHS,

$$\begin{aligned} \text{RHS} &= \frac{1}{dD} \langle \psi | \hat{Q}^{\mu_j} | \psi \rangle \\ &= \frac{1}{dD} \langle \psi | (\text{id} \otimes \mathcal{E}^{\mu_j}) \left( (1-p)\mathcal{V}(\rho^{(d)}) + p \left( \sum_{\kappa=0}^{D-1} \varrho_B^\kappa \otimes |\kappa_{A_e}\rangle\langle\kappa_{A_e}| \right) \right) | \psi \rangle. \end{aligned} \tag{A7}$$

We will now find the maximum value of the RHS. Note that our encoding operator, as an isometry, should not affect the maximum overlap; thus, we can say,

$$\max_{|\psi\rangle} \text{RHS} = \frac{1}{dD} \max_{|\psi\rangle} \langle \psi | \left( (1-p)\mathcal{V}(\rho^{(d)}) + p \left( \sum_{\kappa=0}^{D-1} \varrho_B^\kappa \otimes |\kappa_{A_e}\rangle\langle\kappa_{A_e}| \right) \right) | \psi \rangle. \tag{A8}$$

This is maximized when  $|\psi\rangle\langle\psi| \in \text{Im } \mathcal{V}(\cdot)$ , thus we can utilize our pre-embedding dimension since the isometry preserves the inner product,

$$\begin{aligned} \max_{|\psi\rangle} \text{RHS} &= \frac{1}{dD} \max_{|\psi\rangle} \langle \psi | \left( (1-p)\rho + p \left( \sum_{\kappa=0}^{d-1} \rho_B^\kappa \otimes |\kappa_A\rangle\langle\kappa_A| \right) \right) | \psi \rangle \\ &= \frac{1}{dD} \max_{|\psi\rangle} \left[ (1-p) \left( \theta \langle \psi | \Psi | \psi \rangle + \frac{(1-\theta)}{d^2} \right) \right. \\ &\quad \left. + p \left( \theta \langle \psi | \sum_{\kappa=0}^{d-1} \Psi_B^\kappa \otimes |\kappa_A\rangle\langle\kappa_A| | \psi \rangle + \frac{(1-\theta)}{d^2} \right) \right] \\ &= \frac{1}{dD} \max_{|\psi\rangle} \left[ (1-p) \left( \theta \langle \psi | \Psi | \psi \rangle \right) + p \left( \theta \langle \psi | \sum_{\kappa=0}^{d-1} \Psi_B^\kappa \otimes |\kappa_A\rangle\langle\kappa_A| | \psi \rangle \right) + \frac{(1-\theta)}{d^2} \right]. \end{aligned} \tag{A9}$$

We can see that this is maximized when  $|\psi\rangle = |\Psi\rangle$ , and, thus,

$$\max_{|\psi\rangle} \text{RHS} = \frac{1}{dD} \left( (1-p)\theta + p \frac{\theta}{d} + \frac{(1-\theta)}{d^2} \right). \tag{A10}$$

From Equations (A6) and (A10), we can therefore see that LHS = max RHS meaning

$$\sum_i p_i \hat{Q}^{\mu_i} \Pi^{\mu_i} - p_j \hat{Q}^{\mu_j} \geq 0, \quad \forall j. \tag{A11}$$

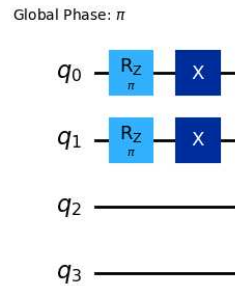
is true and our POVM is optimized for a noisy singlet regardless of dephasure.  $\square$

### Appendix B. Implementation on Quantum Hardware

Due to the limitation that qubit encoding must be used, we performed adjustments to the rest of the protocol to ensure it still behaved as expected, the first of these being the way embedding is handled. In the typical EDC Protocol, higher-dimensional embedding is only applied after the Z gates are applied and involves the introduction of another, higher-dimensional quantum channel. In our implementation, Alice’s qudit begins by possessing the necessary quantum channels to simulate a  $D$ -dimensional qudit; they just remain unused until the X gates are applied, conserving the notion that embedding only happens after entanglement and the Z operation. In other words, qubit channels corresponding to the extra  $D - d$  dimensions added after embedding remain untouched until the X gate is applied.

Another adaptation was applied to the way generalized Pauli Z and Pauli X gates are applied. Custom unitary Qiskit gates using the matrix definitions shown in Equations (4) and (5) involve the use of many native gates and, in turn, make the protocol very noisy. Instead, the  $d$  different Z gates we define are simply the  $d$  possible combinations

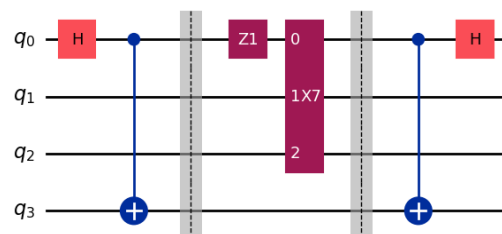
of applying qubit Z gates to each qubit of our encoding. For example, for  $d = 4$ ,  $Z_4^0$  is defined as not applying any Z gate to any qubit of the encoding, and  $Z_4^3$  is defined as applying a Z gate to both qubits of the encoding. For the  $X_D$  gates, the same applies only with  $D$  possibilities instead of  $d$ , usually involving more qubit channels due to the encoding (unless  $d = D$ ). An example of a  $Z_4^3$  and  $X_4^3$  gate is shown in Figure A1.



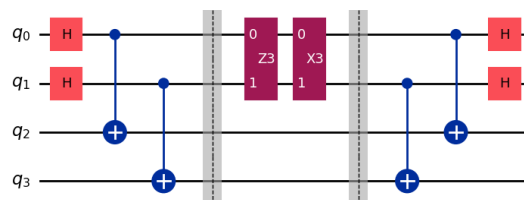
**Figure A1.** A  $Z_4^3$  (left) and  $X_4^3$  (right) gate used in a  $d = D = 4$  EDC protocol. Both gates are expressed in terms of the native gates on the IBM Heron processor. As is seen, when  $d = D = 4$ , a maximum of only four native gates is used for the encoding of the message.

This way of expressing gates is only possible when both  $d$  and  $D$  are powers of two. Otherwise, custom unitary gates must be used to ensure that no qudit reaches a forbidden state (such as  $|11\rangle$  in the case of  $D = 3$  qutrits). As these gates present much worse performance, we restrict ourselves to only cases where  $d$  and  $D$  are powers of two.

Restricting  $D$  and  $d$  allows for Bob’s decoding of the message to be very simple. Reversing the order of the gates used for entanglement is enough to decode the message such that none of Bob’s particles are found in a superposition. Bob can therefore measure both particles and always get the same outcome (assuming no noise). A few examples of EDC protocols with different values of  $d, D$  can be seen in Figures A2 and A3.



**Figure A2.** Example of Embedded Dense Coding implementation on IBM Heron processor through  $IMB_{Torino}$  using starting dimension  $d = 2$  and embedding dimension  $D = 8$ , resulting in a total of four classical bits of information being transmitted.



**Figure A3.** Example of conventional Dense Coding implementation on IBM Heron processor through  $IMB_{Torino}$  using starting dimension  $d = 4$ , resulting in a total of four classical bits of information being transmitted. Note, this represents the special case of EDC where  $d = D = 4$ .

**References**

1. Bennett, C.H.; Brassard, G.; Crépeau, C.; Jozsa, R.; Peres, A.; Wootters, W.K. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **1993**, *70*, 1895–1899. [CrossRef]

2. Ekert, A.K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **1991**, *67*, 661–663. [[CrossRef](#)] [[PubMed](#)]
3. Bennett, C.H.; Wiesner, S.J. Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **1992**, *69*, 2881. [[CrossRef](#)] [[PubMed](#)]
4. Barenco, A.; Ekert, A.K. Dense coding based on quantum entanglement. *J. Mod. Opt.* **1995**, *42*, 1253–1259. [[CrossRef](#)]
5. Mozes, S.; Oppenheim, J.; Reznik, B. Deterministic dense coding with partially entangled states. *Phys. Rev. A* **2005**, *71*, 012311. [[CrossRef](#)]
6. Pati, A.K.; Parashar, P.; Agrawal, P. Probabilistic superdense coding. *Phys. Rev. A* **2005**, *72*, 012329. [[CrossRef](#)]
7. Scholin, O.; Lynn, T.W. Exchange-symmetrized qudit Bell bases and Bell-state distinguishability. *Phys. Rev. Res.* **2025**, *7*, 033124. [[CrossRef](#)]
8. Hiroshima, T. Optimal dense coding with mixed state entanglement. *J. Phys. A Math. Gen.* **2001**, *34*, 6907. [[CrossRef](#)]
9. Devetak, I.; Harrow, A.W.; Winter, A.J. A Resource Framework for Quantum Shannon Theory. *IEEE Trans. Inf. Theory* **2008**, *54*, 4587–4618. [[CrossRef](#)]
10. Buscemi, F.; Datta, N. Entanglement Cost in Practical Scenarios. *Phys. Rev. Lett.* **2011**, *106*, 130503. [[CrossRef](#)]
11. Yue, Q.; Chitambar, E. The zero-error entanglement cost is highly non-additive. *J. Math. Phys.* **2019**, *60*, 112204. [[CrossRef](#)]
12. Leditzky, F.; Leung, D.; Smith, G. Dephasure Channel and Superadditivity of Coherent Information. *Phys. Rev. Lett.* **2018**, *121*, 160501. [[CrossRef](#)]
13. Chitambar, E.; Leditzky, F. On the Duality of Teleportation and Dense Coding. *IEEE Trans. Inf. Theory* **2024**, *70*, 3529–3537. [[CrossRef](#)]
14. Pudda, F.; Chizzini, M.; Crippa, L. Generalised Quantum Gates for Qudits and their Application in Quantum Fourier Transform. *arXiv* **2024**, arXiv:2410.05122. [[CrossRef](#)]
15. Wilde, M.M. From classical to quantum Shannon theory. *arXiv* **2011**, arXiv:1106.1445. .. [[CrossRef](#)]
16. Wei, T.C.; Nemoto, K.; Goldbart, P.M.; Kwiat, P.G.; Munro, W.J.; Verstraete, F. Maximal entanglement versus entropy for mixed quantum states. *Phys. Rev. A* **2003**, *67*, 022110. [[CrossRef](#)]
17. Cavalcanti, D.; Brandão, F.G.S.L.; Terra Cunha, M.O. Are all maximally entangled states pure? *Phys. Rev. A* **2005**, *72*, 040303. [[CrossRef](#)]
18. Badziag, P.; Horodecki, M.; Horodecki, P.; Horodecki, R. Local environment can enhance fidelity of quantum teleportation. *Phys. Rev. A* **2000**, *62*, 012311. [[CrossRef](#)]
19. Horodecki, R.; Horodecki, M. Information-theoretic aspects of inseparability of mixed states. *Phys. Rev. A* **1996**, *54*, 1838–1843. [[CrossRef](#)] [[PubMed](#)]
20. Bennett, C.H.; DiVincenzo, D.P.; Smolin, J.A.; Wootters, W.K. Mixed-state entanglement and quantum error correction. *Phys. Rev. A* **1996**, *54*, 3824–3851. [[CrossRef](#)]
21. Grondalski, J.; Etlinger, D.; James, D. The fully entangled fraction as an inclusive measure of entanglement applications. *Phys. Lett. A* **2002**, *300*, 573–580. [[CrossRef](#)]
22. Holevo, A. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory* **1998**, *44*, 269–273. [[CrossRef](#)]
23. Bennett, C.; Shor, P.; Smolin, J.; Thapliyal, A. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Trans. Inf. Theory* **2002**, *48*, 2637–2655. [[CrossRef](#)]
24. AbuGhanem, M. IBM quantum computers: Evolution, performance, and future directions. *J. Supercomput.* **2025**, *81*, 687. [[CrossRef](#)]
25. Barnett, S.M.; Croke, S. Quantum state discrimination. *Adv. Opt. Photonics* **2009**, *1*, 238–278. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.