Variational Estimation of Optimal Signal States for Quantum Channels

Leonardo Oleynik\textsuperscript{1}, Junaid Ur Rehman\textsuperscript{1}, Hayder Al-Hraishawi\textsuperscript{1}, and Symeon Chatzinotas\textsuperscript{1}

\textsuperscript{1}Affiliation not available

January 18, 2024

Abstract

This paper explores the performance of quantum communication systems in the presence of noise and focuses on finding the optimal encoding for maximizing the classical communication rate, approaching the classical capacity in some scenarios. Instead of theoretically bounding the ultimate capacity of the channel, we adopt a signal-processing perspective to estimate the achievable performance of a physically available but otherwise unknown quantum channel. By employing a variational algorithm to estimate the trace distance between quantum states, we numerically determine the optimal encoding protocol for the amplitude damping and Pauli channels. Our simulations demonstrate the convergence and accuracy of the method with a few iterations, confirming that optimal conditions for binary quantum communication systems can be variationally determined with minimal computation. Further, since the channel knowledge is not required at the transmitter or at the receiver, these results can be employed in arbitrary quantum communication systems, including satellite-based communication systems, a particularly relevant platform for the quantum Internet.
Variational Estimation of Optimal Signal States for Quantum Channels

Leonardo Oleynik, Junaid ur Rehman, Hayder Al-Hraishawi, Senior Member, IEEE, Symeon Chatzinotas Fellow, IEEE

Abstract—This paper explores the performance of quantum communication systems in the presence of noise and focuses on finding the optimal encoding for maximizing the classical communication rate, approaching the classical capacity in some scenarios. Instead of theoretically bounding the ultimate capacity of the channel, we adopt a signal-processing perspective to estimate the achievable performance of a physically available but otherwise unknown quantum channel. By employing a variational algorithm to estimate the trace distance between quantum states, we numerically determine the optimal encoding protocol for the amplitude damping and Pauli channels. Our simulations demonstrate the convergence and accuracy of the method with a few iterations, confirming that optimal conditions for binary quantum communication systems can be variationally determined with minimal computation. Further, since the channel knowledge is not required at the transmitter or at the receiver, these results can be employed in arbitrary quantum communication systems, including satellite-based communication systems, a particularly relevant platform for the quantum Internet.

Index Terms—classical communication, classical-quantum computing, quantum channels, variational algorithms,

I. INTRODUCTION

The ultimate form of quantum communication systems is the vision of quantum Internet that offers a number of unique quantum advantages including security, efficiency, and enabling distributed quantum computing and sensing [1]–[3]. A critical challenge in deploying the quantum Internet lies in establishing reliable and long-range communication links between any two points on Earth [4]. This implies envisioning solutions that can overcome path loss in quantum communication systems, the primary noise source through telecom fibres and free space medium [5], [6]. To address this, the most promising approach involves leveraging the low channel losses offered by free-space optical links to envision networks where satellites serve as intermediate nodes, connecting distant locations seamlessly [4], [7]. This approach requires us to focus on satellite-to-ground and deep-space communication systems, schemes particularly relevant for classical communications over quantum channels. By addressing these facets, we can pave the way for next-generation communication networks and the quantum Internet.

The transmission of classical information through a quantum channel offers unique advantages in free-space optical communication systems. The main reason lies in the fundamental nature of the signal states and the devices used to detect them, i.e., while classical optical communication systems are limited to detection schemes operating over the shot noise level and classical states of an electromagnetic field, quantum communication systems can employ fundamentally different states of light [8] and more accurate detection schemes [9], [10]. In particular, one can achieve a lower bit-error rate compared with the classical counterpart [11], [12] and operate under the shot-noise level (or standard quantum limit) [13]–[15], having the potential application in communication scenarios where the received signal is extremely weak, such as satellite-to-ground, under-water and deep-space communication. Regarding the first, a comparison of quantum and classical detection for communication employing different modulations has recently been made for illustrative system parameters for different satellites’ orbits [16]. Given the fundamentally different nature of quantum channels, employing them in a communication scenario offers promising applications.

Evaluating a communication system’s performance can be approached through two distinct perspectives: information theory and signal processing. Information theory primarily seeks to determine the maximum reliable transmission rate, known as the channel’s capacity. However, calculating this capacity often necessitates a comprehensive tomographic reconstruction of the underlying channel [17], which is often impractical in real-world scenarios. Alternatively, one can apply application-specific conditions to narrow the range of possible channels, making it easier to analyse specific properties [18], [19]. While this approach is computationally more feasible, it only provides bounds for the capacity of particular channels [20], [21]. We embrace a signal-processing perspective to circumvent these challenges, focusing on directly determining optimal signal states. This framework’s main objective is the optimal discrimination of quantum signal states [9], [10], [19]. This task is more straightforward than the information-theoretic approach, as it revolves around a quantum hypothesis testing scenario [22] Section 2, where the goal is to minimize the probability of error when determining the transmitted state at the receiver. Consequently, from a signal-processing standpoint, we aim not to theoretically bind the channel’s ultimate capacity but to assess its real-world performance.

We show that the performance of a binary quantum communication system can be evaluated under the constraints of noisy intermediate-scale quantum (NISQ) computers. It is well known that the maximum probability of discriminating two input states correctly in a binary quantum testing scenario is linearly related to the trace distance between the output states [23] Chapter 9. Besides the panoply of mathematical properties, the trace distance between arbitrary states can
be efficiently estimated using a variational trace distance estimation (VTDE) algorithm \cite{24}. As shown in \cite{25, 26}, variational algorithms fit in a hybrid quantum-classical computation framework, a well-suited framework for the NISQ era. Therefore, since hybrid algorithms can efficiently estimate the trace norm, the performance of a binary quantum modulator can be empirically assessed under the constraints of today’s quantum computers.

Given the importance of classical communication over quantum channels for next-generation connectivity, we consider a general communication scenario in the presence of noise whose description is unknown but stationary. We are interested in finding binary signal states that render optimal communication, i.e., that minimise the probability of error in the discrimination of received states and possibly achieve Holevo’s channel capacity. Based on the close relation between optimal detection and quantum state distinguishability, we implemented the VTDE algorithm presented in \cite{24} to numerically find the optimal encoding for the amplitude damping and Pauli channels. In particular, we show the convergence of our method with a few iterations. Addressing this problem is of utmost importance, not only by the significant challenge of designing optimal communication schemes but also by fully unlocking the applications of free-space communication systems, such as the quantum Internet. In short, the key technical contributions of this paper can be summarised as follows:

- We develop a framework to identify the optimal signal states of binary quantum communication systems by estimating and maximising the trace distance at the receiver’s end. In particular, we do not make any assumption on the underlying channel, except that we restrict our exploration to the binary case alone. This assumption limits our results to quantum channels where the binary signal states are optimal.

- We demonstrate the efficacy of our framework by considering the amplitude damping and Pauli channels. We evaluate two different ansatzes as the starting point of optimization and compare the obtained performance in terms of convergence speed and the quality of obtained solutions.

- We also demonstrate the convergence behaviour of our developed framework in terms of maximising the trace distance and Holevo information, achieving Holevo’s capacity for Pauli channels.

The remainder of this paper is organized as follows. In Section \ref{rel}, we provide preliminaries and set some notation. In Section \ref{rel}, the problem of finding the optimal encoding for maximising the classical communication rate of noisy channels is presented. We present the signal-processing framework in Section \ref{rel} applying it to the amplitude damping and Pauli channels in Section \ref{rel}. We conclude our discussion and provide possible future directions in Section \ref{rel}.

\section*{II. Preliminaries}

We employ the following notation throughout the paper. For a finite Hilbert space $\mathcal{H}$ we denote by $\mathcal{B}(\mathcal{H})$ the algebra of bounded linear operators acting on $\mathcal{H}$; by $\mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$ the space of linear operators taking $\mathcal{H}_A$ to $\mathcal{H}_B$; and by $\mathcal{D}(\mathcal{H}) = \{ \rho \in \mathcal{B}(\mathcal{H}) : \rho \geq 0 \text{ and } \text{Tr} \rho = 1 \}$ the set of quantum states\footnote{The mathematical description of a physical system.}. Specifically, a pure state $\psi$ is a state for which we can associate a normalised vector $|\psi\rangle \in \mathcal{H}$, such that $\psi = |\psi\rangle \langle \psi|$. We use subscripts to identify different Hilbert spaces that are unclear from the context, e.g., $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces associated with systems $A$ and $B$, respectively. A positive operator-valued measure (POVM) is a set of operators $\{\Pi_j\}$ such that $\Pi_j \geq 0 \forall j$ and $\sum_j \Pi_j = I$, where $I$ denotes the identity.

Any physical system of two levels $\rho$, known as qubit, can be decomposed in the so-called Bloch representation

$$\rho = \frac{1}{2} \left( \mathbb{1} + \mathbf{r} \cdot \mathbf{\sigma} \right),$$

where $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$, $\mathbf{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Bloch and Pauli vectors respectively. In this representation, a Bloch vector is given by

$$\mathbf{r} = \text{Tr} \rho \mathbf{\sigma}.$$ 

Moreover, any qubit is mapped to a point in a sphere of radius 1, the Bloch sphere, with pure states (mixed) lying on its surface (interior).

A quantum channel $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_B)$ is a linear, completely positive and trace-preserving map, where $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces associated with the systems $A$ and $B$. Moreover, every quantum channel admits a Choi-Kraus decomposition:

$$\mathcal{N}(O_A) = \sum_{l=0}^{d-1} V_l O_A V_l^\dagger,$$ 

where $O_A \in \mathcal{B}(\mathcal{H}_A)$; $V_l \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$, for all $l \in \{0, \ldots, d-1\}$, $\sum_{l=0}^{d-1} V_l^\dagger V_l = \mathbb{1}$ and $d \leq \dim(\mathcal{H}_A) \dim(\mathcal{H}_B)$. In what follows, we comment on the Pauli and amplitude damping channel.

Pauli channels have the following Choi-Kraus decomposition

$$\mathcal{N}_P(\rho) = \sum_{\mu} p_\mu \sigma_\mu \rho \sigma_\mu,$$

in which $\{\sqrt{p_\mu} \sigma_\mu\}_{\mu=0}^3$ are the Kraus operators. Here, $\sigma_0 = \mathbb{1}$, $\sigma_k$, $k \in \{x, y, z\}$ are Pauli matrices, and $p_\mu$ is a probability distribution (i.e., $\sum_\mu p_\mu = 1$ such that $p_\mu \geq 0$). Pauli channels can be seen as a mathematical generalisation of the bit flip and phase flip channels. By specifying $p_\mu = (1-p, 0, 0, p)$ in (4), one finds the phase flip channel, i.e.,

$$\mathcal{N}_{PF}(\rho) = (1-p)\rho + p \sigma_z \rho \sigma_z.$$ 

The expression shows that the state remains unchanged with probability $(1-p)$ and undergoes a sign inversion with probability $p$.

Another useful parametrisation for the Pauli channel is choosing $p_\mu$ to be the eigenvalues of the following $d^2 \times d^2$ exponential correlation matrix \cite{27, 28}:

$$\Phi(\gamma) = \frac{1}{d^2} \begin{bmatrix} \gamma^{i-j} \end{bmatrix}_{0 \leq i, j \leq d^2-1}$$

where $\gamma$ is exponential correlation matrix \cite{27, 28}.
in ascending order. It is important to note that when $\gamma = 0$, [6] gives a completely depolarizing channel characterized by high noise levels. Conversely, $\gamma = 1$ yields an ideal, noiseless channel. Additionally, as $\gamma$ increases, the channel parameters become increasingly ordered, resulting in channels with reduced noise levels [27].

The Pauli channel, or dephasing channel, characterises the type of noise that can occur during the transmission of quantum information. Importantly, it represents a distinctly quantum process wherein information within a quantum state is lost without any energy dissipation [29]. Due to its particularities, the Pauli channel is frequently employed to simulate and understand errors in various quantum applications.

The amplitude damping channel describes the stochastic degradation of quantum information resulting from energy dissipation within quantum systems. This channel is a valuable tool for modelling scenarios involving open systems where environmental interactions lead to energy dissipation. A typical application of this channel is in modelling the spontaneous decay of an excited quantum state. In such phenomena, a physical system in an excited state $|1\rangle$ decay to its ground state $|0\rangle$ with some probability $\kappa$, i.e., after this process, the system is more likely to be measured in the ground state than before [23]. The Kraus operators $A_0 = \sqrt{\kappa}|0\rangle\langle 1|$ and $A_1 = |0\rangle\langle 0| + \sqrt{1 - \kappa}|1\rangle\langle 1|$ of the amplitude damping channel, defined in terms of $\kappa$, act on a qubit decreasing the probabilities of being in a excited state — as explained in [23].

Unitary evolutions are a special kind of quantum channel for which one associates a group $G$ with elements $g$. A unitary evolution of an operator $O \in B(H)$ is then expressed as

$$U_g(O) = U_g O U_g^\dagger,$$  

(7)

where $U_g$ satisfies $U_g U_g^\dagger = U_g^\dagger U_g = \mathbb{1}$. Following the representation [1], the unitary evolution of a qubit can be pictured as a rotation of its associated Bloch vector. Since the group of special unitaries $SU(2)$ is associated with a two-dimensional Hilbert space, one can represent its transformations as rotations in $\mathbb{R}^3$. We define a rotation of angle $\theta$ along the axis $e$ by $R_e(\theta)$ and express the rotated qubit by $R_e(\theta)\rho R_e^\dagger(\theta)$. A rotation of a system comprising more than one qubit will be denoted by $U_\theta$, where $\theta$ indicates that more than one parameter $\theta$ is needed.

The trace norm of an operator $O \in B(H)$ is defined as $||O||_1 := \text{Tr} \sqrt{O^\dagger O}$. This norm naturally induces a distance measure $||O - O'||_1$, called the trace distance. The normalized trace distance between two states $\rho_0$ and $\rho_1$ is denoted by

$$D(\rho_0, \rho_1) := \frac{1}{2}||\rho_0 - \rho_1||_1,$$  

(8)

For a pair of qubits, $\rho_0$ and $\rho_1$, the above equation equals the Euclidean distance between their corresponding Bloch vectors, i.e.,

$$D(\rho_0, \rho_1) = ||r_0 - r_1||_2,$$  

(9)

where $r_0 = (x_0, y_0, z_0)$ and $r_1 = (x_1, y_1, z_1)$ are the Bloch vectors of the states $\rho_0$ and $\rho_1$ respectively. This gives us a straightforward geometric interpretation of the trace distance, which we will use to find our analytical results.

III. Problem Statement: Quantum Communication Systems with Noise

This Section presents the formalism of two-state minimum-error discrimination and its relation to the trace distance. The reader is referred to [22] for a detailed treatment.

A. System Model: Binary Communication over Quantum Channels

In a binary quantum communication system, a classical source emits a symbol $x$ drawn from the alphabet $\{0, 1\}$ with probabilities $q_0$ and $q_1 = 1 - q_0$. Based on the emitted symbol, the transmitter (Alice) prepares and sends a quantum state $\psi_x$ through a quantum channel. On the other side of the communication line, the receiver (Bob) performs a set of measurements to guess the received state and the original symbol. While for the ideal channel, the transmitted states are received at the receiver end without any noise, for general channels $N$, the received states are noisy and therefore described by a set of density operators $\{\rho_0, \rho_1\}$, in which $\rho_x = N(\psi_x)$.

B. Quantum Binary Detection Theory

According to quantum detection theory [30], [31], the detection system used by Bob for choosing among the possible states $\{\rho_0, \rho_1\}$ is characterised by a POVM $\{\Pi_0, \Pi_1\}$. The probability that the detection device outputs the symbol $x$, provided that the received quantum state is $\rho_y$, is given by

$$p(x|y) = \text{Tr} (\Pi_x \rho_y), \quad x, y \in \{0, 1\}.$$  

(10)

In particular, for an equiprobable binary communication system, the probability of correct detection is

$$p_c = \frac{1}{2} [p(0|0) + p(1|1)] = \frac{1}{2} [\text{Tr} \Pi_0 \rho_0 + \text{Tr} \Pi_1 \rho_1]$$  

(11)

The optimisation of the detection system reduces to finding the POVM elements $\Pi_0$ and $\Pi_1$ that maximise (11). The optimal probability of correct detection is [23] [Sec. 9.3]

$$P_c = \max_{\Pi_0, \Pi_1} p_c$$  

(12)

$$= \frac{1}{2} [1 + D(\rho_0, \rho_1)].$$  

(13)

The latter expression shows the operational meaning of the trace distance. The probability of correct decision is one when $\rho_0$ and $\rho_1$ are orthogonal, i.e., maximally distinguishable $D(\rho_0, \rho_1) = 1$.

Our main goal is to devise analytical and numerical means for finding the set of optimal input states $\{\psi_0, \psi_1\}$ that, after being transmitted through a quantum channel $N$, maximises $D(\rho_0, \rho_1)$ and therefore (13). Formally, we are seeking for

$$\mathcal{Y}^* = \arg \max_{\psi_0, \psi_1} D(\rho_0, \rho_1),$$  

(14)

in which $\rho_x = N(\psi_x)$. As we are constrained to an equiprobable distribution, these two states generally do not attain the Holevo capacity. Nevertheless, as outlined in [32], a collection of two states is sufficient to attain Holevo’s capacity for the Pauli channel — one of the channels we will analyse.
IV. Maximisation of State Distinguishability

This section explores the analytical and numerical means for finding the maximally distinguishable pair of states for amplitude damping and Pauli channels.

A. Analytical Solution

The normalized trace distance between the output qubits \(\{\rho_0, \rho_1\}\) has a geometrical interpretation in terms of the input states in the following way

\[
D(\rho_0, \rho_1) = \sqrt{\sum_k \alpha_k^2 \left(\left(|r_0 - r_1| \cdot e_k\right)^2\right)},
\]

where \(\alpha_k = \frac{1}{2} \text{Tr} N(\sigma_k)\sigma_k\) and \(\alpha_0 = 1\) are the eigenvalues of the channel \(N\); and \(r_0 - r_1\) is the Pauli vector of \(\psi_0 - \psi_1\) found by (2). Further, \(\alpha_k \in \mathbb{R}\) since the Pauli matrix \(\sigma_k\) is hermitian. Therefore, all terms in the above expression are positive.

The maximally distinguishable pairs of states are orthogonal and have their Pauli vector components aligned toward the greatest \(\alpha_k^2\). The orthogonality \(r_0 = -r_1\) of any pair of maximally distinguishable states follows from maximising the second factor in (15), i.e., the maximal of (15) only occurs for orthogonal states. Moreover, these states must have their Pauli vector components in the direction \(e_k^*\), where \(k^*\) is such that \(\alpha_{k^*}^2 = \max_k \{\alpha_k^2\}\). This condition defines three situations:

1) All the coefficients are identical, and any orthogonal pair is maximally distinguishable;
2) Two of the \(\alpha_k^2\) are identical, and any orthogonal pair lying in the plane defined by them is maximally distinguishable;
3) All the coefficients are different, and a unique orthogonal pair in the direction \(e_k^*\) is maximally distinguishable.

It is important to emphasise that these states are optimal in the sense of (14) as mentioned earlier, they do not necessarily attain Holevo capacity.

B. Variational Trace Distance Estimation Algorithm

In [24], the authors have introduced a variational quantum algorithm for estimating the trace norm of a hermitian operator \(H\). Their method employs a classical optimisation of a parameterised quantum circuit and requires only single-qubit measurements of an arbitrary ancillary pure state. Fulfilling these criteria, their algorithm constitutes an efficient algorithm for NISQ devices.

Let \(H_S\) and \(H_A\) denote the Hilbert spaces associated with the ancillary qubit and the system of interest, respectively. Specializing \(H\) as \((\rho_0 - \rho_1)/2\), where \(\rho_0, \rho_1 \in \mathcal{D}(H_A)\), they showed

\[
D(\rho_0, \rho_1) = \max_\theta \left[\mathcal{L}_{\theta}(\rho_0, \rho_1)\right],
\]

in which \(\mathcal{L}_{\theta}(\rho_0, \rho_1) := p(0|\tilde{\rho}_0) - p(0|\tilde{\rho}_1)\). In the latter expression,

\[
\tilde{\rho}_x = \text{Tr}_A U_\theta (\rho_x \otimes \Pi_0) \in \mathcal{D}(H_S)
\]

denotes the final state of the ancillary system and

\[
p(0|\tilde{\rho}_x) = \text{Tr} \Pi_0 \tilde{\rho}_x
\]

the probability distribution of obtaining the value zero by measuring \(\tilde{\rho}_x\). This enables them to estimate the trace distance by optimising over parameterised unitary maps \(U_\theta\). Moreover, by adopting a hardware-efficient ansatz [33], [34], they could reduce the set of unitaries to a combination of parameterised single-qubit rotations, along with CNOT gates on adjacent qubits as entanglement gates. For our case, in which \(\rho_0, \rho_1\) are qubits, the global unitary \(U_\theta\) corresponding to \(U_\theta\) is depicted in Fig. 1 — \(U_\theta\) is defined in terms of single qubit rotations \(R_x(\theta)\) and \(R_y(\theta)\), each controlled by a classical parameter \(\theta\).

In practice, they compute an approximation of \(D(\rho_0, \rho_1)\) whose deviation from the exact value decreases as the number of experiment runs \(N\) increases. This follows from the fact that in practice, we can only compute the frequency of successful outcomes corresponding to (18) and not its exact value. However, due to the law of large numbers, we know that this approximation is sufficiently accurate for many experiment runs, and so is the corresponding approximation of \(D(\rho_0, \rho_1)\).

In our analysis, \(N\) defines an empirical estimation \(\hat{D}^N(\rho_0, \rho_1)\) of the exact value of \(D(\rho_0, \rho_1)\) that arbitrarily approximates the second as \(N\) increases, i.e.,

\[
\lim_{N \to \infty} \hat{D}^N(\rho_0, \rho_1) = D(\rho_0, \rho_1).
\]

C. Variational optimal signal states estimation algorithm

Our algorithm employs the above VTDE algorithm as an intermediate step to determine the optimal signal states of a given channel. Parametrisating the input states in terms of unitary maps \(U_{\phi}\), and using the VTDE algorithm to estimate the trace distance between two output states, (14) becomes

\[
\hat{\gamma}^* = \arg \max_{\phi} \hat{D}^N_{\phi}(\rho_0, \rho_1),
\]

in which \(\rho_x = N [U_{\phi_0}(|0\rangle |0\rangle)]\) and \(\hat{\gamma}^* \approx \gamma^*\). The parameter \(\phi = (\phi_0, \phi_1)\) parametrized the global unitary map \(U_{\phi} = U_{\phi_0} \otimes U_{\phi_1}\). This allowed us to determine the optimal pair.

![Variational Trace Distance Estimation VTDE algorithm](image-url)
of signal states employing a parametrised state preparation

circuit, the VTDE algorithm [24], and a classical maximisation

routine, as depicted in Fig. 2.

Our algorithm accuracy is associated with the number of

experiment runs $N$ and the number of $M$ iterations. While the

former regards the number of measurements required to build

up the statistics of each probability distribution in the VTDE

algorithm, as explained before, the latter represents the number

of updates on the parameter $\phi$ of our classical optimiser (cf.

Fig. 2).

The algorithm can also be regarded as the initial calibration

phase within a communication process, where optimal signaling

states (encoding) are selected over a sequence of $M$ iterations.

In this context, it is crucial to assume stationary channels, which signifies channels whose Choi-Krauss decomposition

is constant over time.

D. Representation of optimal states

We chose the polar plane to show how our empirical

estimations predict the optimal ensemble. In this plane, a point

corresponds to the $i$-th estimation of the optimal ensemble

using a given estimation $\{\hat{\psi}_0^i, \hat{\psi}_1^i\}$ in terms of the estimated Pauli vectors $\hat{r}_0^i, \hat{r}_1^i$

in the following way: its coordinates $\beta^i$ and $\hat{R}^i$ are defined by the relative angle between $\hat{r}_0^i$ and $\hat{r}_1^i$ and their relative distance on the $x-y$ plane, i.e.,

$$\beta^i := \arccos \left( \hat{r}_0^i \cdot \hat{r}_1^i \right)$$

(21)

and

$$\hat{R}^i := \frac{1}{2} \sqrt{ (\hat{x}_1^i - \hat{x}_0^i)^2 + (\hat{y}_1^i - \hat{y}_0^i)^2 }.$$  

(22)

Likewise in Fig. 4, but the radial distance being $\hat{X}^i := \frac{1}{2} | \hat{x}_1^i - \hat{x}_0^i |$.

V. APPLICATION: OPTIMAL BINARY QUANTUM

COMMUNICATION SYSTEMS

In this section, we assume a binary quantum communication

system as presented in Section III and find the optimal signal ensembles for the amplitude damping and Pauli channels, described

in Section II using the variational algorithm presented in Section IV. We compare these numerical results with the analytical ones derived from [15]. Finally, we evaluate the performance of our methods.

A. Optimal ensembles

For the amplitude damping channel, the maximally distinguishable pair of states lie on the equatorial plane of the Bloch sphere. As discussed in Subsection IV-A, the orthogonal pair of states that maximises [15] have their Pauli vector components in the $e_1$-$e_2$. For the amplitude damping channel, these coefficients are

$$\alpha_{x,y} = \sqrt{1 - \kappa} \quad \text{and} \quad \alpha_z = 1 - \kappa,$$  

(23)

implying in $\alpha_{x,y} \leq \kappa$ since $\kappa \in [0,1]$. In this relation, the equality holds for $\kappa = 0$ (the ideal channel) and $\kappa = 1$ (the completely depolarising channel), in which any and none pair of states maximises [15] respectively. For $\kappa \in (0,1)$, $\alpha_{x,y} > \alpha_z$ implies that any pair states lying on the equatorial plane will maximise equation (15).

For the Pauli channels parametrized according to (6), the maximaly distinguishable pair of states is $\{|+\rangle |+\rangle, |\rangle |\rangle\rangle$.

Figure 3. In this polar representation, we visualise the convergence of our estimations towards the maximally distinguishable pair for the amplitude damping channel ($\kappa = .9$) using two distinct ansatz (top and bottom). The radial and angular axes are represented as $R^i$ and $\beta^i$, respectively (cf. IV-D). The blue (red) line illustrates the 1st, 5th, 10th, 15th, and 20th estimations based on $10^3$ ($10^6$) experiment runs, capturing the evolution of the estimations towards the maximally distinguishable pair, marked in black. We also display estimations obtained using the exact trace distance expression, displayed in green, to provide a reference.
As before, this comes from an inspection of the coefficients \(\alpha_k\). For a general Pauli channel with probability distribution \(p_\mu\) we have [35]

\[
\alpha_k = p_0 + p_k - \sum_{k' \neq k} p_{k'}.
\]

Choosing \(p_\mu\) as the ordered eigenvalues of \(\rho\) implies in \(\alpha_x^2 \geq \alpha_y^2 \geq \alpha_z^2\), where the equality holds \(\gamma = 0\) and \(\gamma = 1\), the completely depolarising and ideal channels respectively. For \(\gamma \in (0, 1)\), \(\alpha_x^2 \geq \alpha_y^2 \geq \alpha_z^2\) and \(\alpha_x^2 = \alpha_y^2\). Therefore, the pair of states in the direction \(e_x\) will maximise equation (15), i.e., the pair \(\{+\rangle\langle+|, -\rangle\langle-|\}\).

In Figs. 3 and 4, we depict how our empirical estimations predict the maximally distinguishable set of states for the amplitude damping and Pauli channels, respectively. As discussed above, these sets, respectively, are any pair on the equatorial plane of the Bloch sphere and the pair \(\{+\rangle\langle+|, -\rangle\langle-|\}\). We can see that, according to our polar representations (cf. Subsection IV-D), in both cases, all the estimations converge.

**B. Algorithm Accuracy**

In red in Fig. 5 we ratify the role of the parameters \(N\) and \(M\) in our algorithm accuracy in finding the optimal pair of states as discussed in Subsection IV-B. By comparing trace distance estimations based on \(10^3\) and \(10^6\) experiment runs to the exact value \(D(\rho_0^*, \rho_1^*)\), we discern a noteworthy pattern: while all estimations ultimately converge asymptotically to the exact value as the number of iterations increases, those utilising \(10^6\) experiment runs achieve this convergence with significantly fewer iterations than their \(10^3\) counterparts. This demonstrates that \(N\) is closely related to the algorithm performance, as discussed before. Moreover, the figure shows a rapid improvement in all the estimations after the fifth iteration. This observation indicates that the algorithm’s computational complexity remains within the practical limits of the NISQ era. In other words, the algorithm does not require excessive computational resources, making it feasible and efficient for real-world applications.

Furthermore, Figs. 6 and 7 compare the signal processing and information-theoretical viewpoints for identifying the optimal signal states for Pauli channels. The first figure shows Holevo Information estimations (depicted in blue) following the convergence of trace distance estimations (shown in red). This indicates that the maximally distinguishable pair of states determined not only maximises (13) and approximates (14) but also corresponds to the optimal ensemble needed to attain classical capacity for Pauli channels, as emphasised in Fig. 6.

Although valid for all Pauli channels, this conclusion cannot
be straightforwardly generalised to other channels since an ensemble consisting solely of two equiprobable orthogonal states can be insufficient to achieve Holevo’s capacity [22].

VI. CONCLUSION

By employing the VTDE algorithm presented in [24], we developed a numerical method for finding the optimal encoding of a binary quantum communication system. We applied our approach for the amplitude damping and Pauli channels and demonstrated its convergence and accuracy. More specifically, we showed that our method approximates the analytical predictions (which we have also derived) in these cases as the number of experiment runs \( N \) and classical iterations \( M \) increases. This indicates that the maximally distinguishable pair of states determined approximates [14] and corresponds to the optimal ensemble needed to attain classical capacity for Pauli channels. Therefore, our algorithm can efficiently estimate the Holevo’s capacity of an unknown and stationary channel, which has numerous applicability in various quantum communication systems encompassing satellite-based communication platforms. More importantly, since our approach is based on a hybrid classical-quantum architecture, these results show that the performance of quantum communication systems can be efficiently evaluated using NISQ devices. These results ratify the importance of a signal-processing approach in fostering the unique advantages of classical communication over quantum channels. Efforts can be undertaken to extend the proposed scheme to accommodate non-equiprobable binary communication systems. This would broaden the range of potential signalling encodings, possibly achieving higher values of Holevo information.

REFERENCES