IEEE Transactions on Quantum Engineering

Received 27 October 2023; revised 31 March 2024; accepted 14 April 2024; date of publication 25 April 2024; date of current version 7 May 2024.

Digital Object Identifier 10.1109/TQE.2024.3393416

Variational Estimation of Optimal Signal States for Quantum Channels

LEONARDO OLEYNIK, JUNAID UR REHMAN, HAYDER AL-HRAISHAWI (Senior Member, IEEE), AND SYMEON CHATZINOTAS (Fellow, IEEE)

Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, L-1855 Esch-sur-Alzette, Luxembourg

Corresponding author: Leonardo Oleynik (e-mail: leonardo.oleynik@uni.lu).

ABSTRACT This article 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. Furthermore, 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–quantum computing, classical communication, 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], [2], [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 fibers 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 [9], [11] and operate under the shot noise level (or standard quantum limit) [12], [13], [14], having the potential application in communication scenarios where the received signal is extremely weak, such as satellite-to-ground, underwater, 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 [15]. 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 [16], 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 analyze specific properties [17], [18]. While this approach is computationally more feasible, it only provides bounds for the capacity of particular channels [19], [20]. 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], [15]. This task is more straightforward than the information-theoretic approach, as it revolves around a *quantum hypothesis testing scenario* [21, Sec. 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 [22, Ch. 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 [23]. As shown in [24] and [25], 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 minimize 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 [23] 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 article can be summarized as follows.

1) We develop a framework to identify the optimal signal states of binary quantum communication systems by estimating and maximizing 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.
2) 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.
3) We also demonstrate the convergence behavior of our developed framework in terms of maximizing the trace distance and Holevo information, achieving Holevo's capacity for Pauli channels.

The rest of this article is organized as follows. In Section II, we provide preliminaries and set some notation. In Section III, the problem of finding the optimal encoding for maximizing the classical communication rate of noisy channels is presented. We present the signal processing framework in Section IV, applying it to the amplitude damping and Pauli channels in Section V. Finally, Section VI concludes this article and provides possible future directions.

II. PRELIMINARIES

We employ the following notations throughout this article. 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 $\mathcal{D}(\mathcal{H}) = \{\rho \in \mathcal{B}(\mathcal{H})\,;\ \rho \geq 0 \text{ and } \mathrm{Tr}\,\rho = 1\}$ the set of *quantum states*.[1] Specifically, a pure state ψ is a state for which we can associate a normalized 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 $\sum_j \Pi_j = \mathbb{I}$, where $\mathbb{I}$ denotes the identity.

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

$$\rho = \frac{\mathbb{I} + \mathbf{r}\cdot\boldsymbol{\sigma}}{2} \tag{1}$$

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

[1]The mathematical description of a physical system.

vector is given by

$$\mathbf{r} = \mathrm{Tr}\,\rho\boldsymbol{\sigma}. \tag{2}$$

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 $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ 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

$$\mathcal{N}(O_A) = \sum_{l=0}^{d-1} V_l O_A V_l^\dagger \tag{3}$$

where $O_A \in \mathcal{B}(\mathcal{H}_A), V_l \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_B)$, for all $l \in \{0, \ldots, d-1\}$, $\sum_{l=0}^{d-1} V_l^\dagger V_l = \mathbb{I}$, 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:

$$\mathcal{N}_P(\rho) = \sum_{\mu} p_\mu \sigma_\mu \rho \sigma_\mu \tag{4}$$

in which $\{\sqrt{p_\mu}\sigma_\mu\}_{\mu=0}^3$ are the Kraus operators. Here, $\sigma_0 = \mathbb{I}$, σ_k, $k \in \{x, y, z\}$ are Pauli matrices, and p_μ 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 generalization 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.,

$$\mathcal{N}_{PF}(\rho) = (1-p)\rho + p\sigma_z\rho\sigma_z. \tag{5}$$

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

Another useful parameterization for the Pauli channel is choosing p_μ to be the eigenvalues of the following $d^2 \times d^2$ exponential correlation matrix [26], [27]:

$$\Phi(\gamma) = \frac{1}{d^2}\left[\gamma^{|i-j|}\right]_{0\leq i,j\leq d^2-1} \tag{6}$$

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. In addition, as γ increases, the channel parameters become increasingly ordered, resulting in channels with reduced noise levels [26].

The Pauli channel, or dephasing channel, characterizes 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 [28]. 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 modeling scenarios involving open systems where environmental interactions lead to energy dissipation. A typical application of this channel is in modeling the spontaneous decay of an excited quantum state. In such phenomena, a physical system in an excited state $|1\rangle$ decays to its ground state $|0\rangle$ with some probability κ, i.e., after this process, the system is more likely to be measured in the ground state than before [22]. 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 κ, act on a qubit decreasing the probabilities of being in an excited state—as explained in [22].

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 $O \in \mathcal{B}(\mathcal{H})$ is then expressed as

$$\mathcal{U}_g(O) = U_g O U_g^\dagger \tag{7}$$

where U_g satisfies $U_g U_g^\dagger = U_g^\dagger U_g = \mathbb{I}$. 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 2-D Hilbert space, one can represent its transformations as rotations in $\mathbb{R}^3$. We define a rotation of angle θ along the axis $\mathbf{e}$ by $R_{\mathbf{e}}(\theta)$ and express the rotated qubit by $R_{\mathbf{e}}(\theta)\rho R_{\mathbf{e}}^\dagger(\theta)$. A rotation of a system comprising more than one qubit will be denoted by $\mathcal{U}_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ indicates that more than one parameter θ is needed.

The trace norm of an operator $O \in \mathcal{B}(\mathcal{H})$ is defined as $\|O\|_1 := \mathrm{Tr}\sqrt{\mathrm{O}^\dagger\mathrm{O}}$. This norm naturally induces a distance measure $\|O - O'\|_1$, called the trace distance. The *normalized* trace distance between two states ρ_0 and ρ_1 is denoted by

$$D(\rho_0, \rho_1) := \frac{1}{2}\|\rho_0 - \rho_1\|_1. \tag{8}$$

For a pair of qubits, ρ_0 and ρ_1, the above equation equals the Euclidean distance between their corresponding Bloch vectors, i.e.,

$$D(\rho_0, \rho_1) = \|\mathbf{r}_0 - \mathbf{r}_1\|_2 \tag{9}$$

where $\mathbf{r}_0 = (x_0, y_0, z_0)$ and $\mathbf{r}_1 = (x_1, y_1, z_1)$ are the Bloch vectors of the states ρ_0 and ρ_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 [21] 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 ψ_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 $\mathcal{N}$, the received states are noisy and, therefore, described by a set of density operators $\{\rho_0, \rho_1\}$, in which $\rho_x = \mathcal{N}(\psi_x)$.

B. QUANTUM BINARY DETECTION THEORY

According to quantum detection theory [29], [30], the detection system used by Bob for choosing among the possible states $\{\rho_0, \rho_1\}$ is characterized by a POVM $\{\Pi_0, \Pi_1\}$. The probability that the detection device outputs the symbol x, provided that the received quantum state is ρ_y, is given by

$$p(x|y) = \mathrm{Tr}\left(\Pi_x \rho_y\right), \qquad x, y \in \{0, 1\}. \tag{10}$$

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

$$p_c = \frac{1}{2}[p(0|0) + p(1|1)] = \frac{1}{2}[\mathrm{Tr}\Pi_0\rho_0 + \mathrm{Tr}\Pi_1\rho_1]. \tag{11}$$

The optimization of the detection system reduces to finding the POVM elements Π_0 and Π_1 that maximize (11). The optimal probability of correct detection is [22, Sec. 9.3]

$$P_c = \max_{\Pi_0, \Pi_1} p_c \tag{12}$$

$$= \frac{1}{2}[1 + D(\rho_0, \rho_1)]. \tag{13}$$

The latter expression shows the operational meaning of the trace distance. The probability of correct decision is one when ρ_0 and ρ_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 $\mathcal{N}$, maximizes $D(\rho_0, \rho_1)$ and therefore (13). Formally, we are seeking for

$$\mathcal{Y}^* = \arg\max_{\psi_0, \psi_1} D(\rho_0, \rho_1) \tag{14}$$

in which $\rho_x = \mathcal{N}(\psi_x)$. As we are constrained to an equiprobable distribution, these two states generally do *not* attain the Holevo capacity. Nevertheless, as outlined in [31], a collection of two states is sufficient to attain Holevo's capacity for the Pauli channel—one of the channels we will analyze.

IV. MAXIMIZATION 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:

$$D(\rho_0, \rho_1) = \sqrt{\sum_k \alpha_k^2 \left[(\mathbf{r}_0 - \mathbf{r}_1) \cdot \mathbf{e}_k\right]^2} \tag{15}$$

where $\alpha_k = \frac{1}{2}\mathrm{Tr}\mathcal{N}(\sigma_k)\sigma_k$ and $\alpha_0 = 1$ are the eigenvalues of the channel $\mathcal{N}$, and $\mathbf{r}_0 - \mathbf{r}_1$ is the Pauli vector of $\psi_0 - \psi_1$ found by (2). Furthermore, $\alpha_k \in \mathbb{R}$ since the Pauli matrix σ_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 α_k^2. The orthogonality $\mathbf{r}_0 = -\mathbf{r}_1$ of any pair of maximally distinguishable states follows from maximizing the second factor in (15), i.e., the maximal of (15) only occurs for orthogonal states. Moreover, these states are associated with a direction in $\mathbb{R}^3$, for which the corresponding Pauli vectors are aligned. Specifically, these vectors lie along the direction $\mathbf{e}_{k^*}$, where k^* is such that $\alpha_{k^*}^2 = \max_k\{\alpha_k^2\}$. This condition defines three cases based on the relations between coefficients α_k.

1) If all the coefficients are equal, any orthogonal pair of states is maximally distinguishable. That is, if $\alpha_x = \alpha_y = \alpha_z$, then $\alpha_{k^*}^2 = \{\alpha_x^2, \alpha_y^2, \alpha_z^2\}$ and $\mathbf{e}_{k^*}$ can be any linear combination of $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$.
2) If two coefficients are equal and greater than the third one, then any pair of orthogonal states, with vectors lying on a plane defined by the greater coefficients, is maximally distinguishable. That is, if $\alpha_k = \alpha_{k'} > \alpha_{k''}$, then $\alpha_{k^*}^2 = \{\alpha_k^2, \alpha_{k'}^2\}$ and $\mathbf{e}_{k^*}$ can be any linear combination of $\mathbf{e}_k$ and $\mathbf{e}_{k'}$, where $k, k', k'' \in \{x, y, z\}$;
3) If all the coefficients are different to each other, then the pair of orthogonal states, with Pauli vectors aligned to $\mathbf{e}_{k^*}$, is maximally distinguishable. That is, if $\alpha_x \neq \alpha_y \neq \alpha_z$, then $\mathbf{e}_{k^*}$ is uniquely defined by $\alpha_{k^*}^2$.

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

B. VTDE ALGORITHM

Chen et al. [23] have introduced a variational quantum algorithm for estimating the trace norm of a Hermitian operator H. Their method employs a *classical* optimization of a parameterized 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 $\mathcal{H}_S$ and $\mathcal{H}_A$ denote the Hilbert spaces associated with the ancillary qubit and the system of interest, respectively. Specializing H as $(\rho_0 - \rho_1)/2$, where $\rho_0, \rho_1 \in \mathcal{D}(\mathcal{H}_A)$, they showed

$$D(\rho_0, \rho_1) = \max_{\boldsymbol{\theta}}[L_{\boldsymbol{\theta}}(\rho_0, \rho_1)] \tag{16}$$

IEEE Transactions on Quantum Engineering

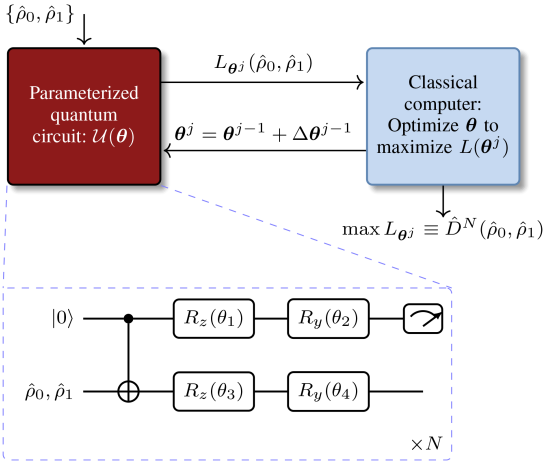

FIGURE 1. VTDE quantum algorithm [23]. A parameterized quantum circuit estimates the loss function L_{θ^j}, which is maximized on a classical computer. Red (blue) indicates quantum (classical) operations, and dashed boxes represent iteration counts for each part of the algorithm.

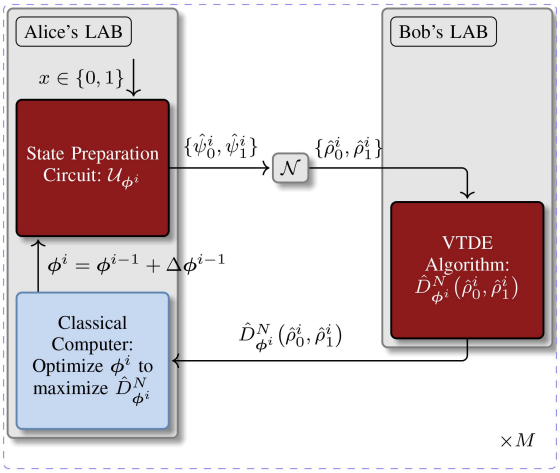

FIGURE 2. Variational quantum algorithm designed to optimize signal states. In Alice's lab, the binary alphabet {0, 1} is encoded into estimated quantum states $\{\hat{\psi}_0^i, \hat{\psi}_1^i\}$ using a parameterized circuit $\mathcal{U}_{\phi^i}$ before transmitting them through the quantum channel $\mathcal{N}$. In Bob's lab, the VTDE algorithm estimates the received signal pair's distance $\hat{D}^N_{\phi^i}(\hat{\rho}_0^i, \hat{\rho}_1^i)$. This assessment is then maximized on a classical computer in Alice's lab, updating the state preparation circuit with ϕ^i. Red (blue) indicates quantum (classical) operations, and the dashed box represents iteration counts.

in which $L_{\boldsymbol{\theta}}(\rho_0, \rho_1) := p(0|\bar{\rho}_0) - p(0|\bar{\rho}_1)$. In the latter expression

$$\bar{\rho}_x = \mathrm{Tr}_A \mathcal{U}_{\boldsymbol{\theta}}(\rho_x \otimes \Pi_0) \in \mathcal{D}(\mathcal{H}_S) \tag{17}$$

denotes the final state of the ancillary system and

$$p(0|\bar{\rho}_x) = \mathrm{Tr}\Pi_0 \bar{\rho}_x \tag{18}$$

the probability distribution of obtaining the value zero by measuring $\bar{\rho}_x$. This enables them to estimate the trace distance by optimizing over parameterized unitary maps $\mathcal{U}_{\boldsymbol{\theta}}$. Moreover, by adopting a hardware-efficient ansatz [32], [33], they could reduce the set of unitaries to a combination of parameterized single-qubit rotations, along with CNOT gates on adjacent qubits as entanglement gates. For our case, in which ρ_0 and ρ_1 are qubits, the global unitary $U_{\boldsymbol{\theta}}$ corresponding to $\mathcal{U}_{\boldsymbol{\theta}}$ is depicted in Fig. 1 —$U_{\boldsymbol{\theta}}$ is defined in terms of single-qubit rotations $R_y(\theta)$ and $R_z(\theta)$, each controlled by a classical parameter θ. At the jth iteration, $\theta^j = \theta^{j-1} + \Delta\theta^{j-1}$, where θ^0 is generated randomly and $\Delta\theta^{j-1}$ is a small increment obtained using the constrained optimization by linear approximation (COBYLA) minimization algorithm.

In practice, they compute an approximation of $D(\rho_0, \rho_1)$ whose deviation from the exact value decreases as the number of experimental 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 experimental 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.,

$$\lim_{N\to\infty} \hat{D}^N(\rho_0, \rho_1) = D(\rho_0, \rho_1). \tag{19}$$

C. VARIATIONAL OPTIMAL SIGNAL STATE ESTIMATION ALGORITHM

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

$$\hat{\mathcal{Y}}^* = \arg\max_{\boldsymbol{\phi}} \hat{D}^N_{\boldsymbol{\phi}}(\rho_0, \rho_1) \tag{20}$$

in which $\rho_x = \mathcal{N}[\mathcal{U}_{\boldsymbol{\phi}_x}(|0\rangle\langle 0|)]$ and $\hat{\mathcal{Y}}^* \cong \mathcal{Y}^*$. The parameter $\boldsymbol{\phi} = (\boldsymbol{\phi}_0, \boldsymbol{\phi}_1)$ parameterized the global unitary map $\mathcal{U}_{\boldsymbol{\phi}} = \mathcal{U}_{\boldsymbol{\phi}_0} \otimes \mathcal{U}_{\boldsymbol{\phi}_1}$. This allowed us to determine the optimal pair of signal states employing a parameterized state preparation circuit, the VTDE algorithm [23], and a classical maximization routine, as depicted in Fig. 2.

Our algorithm accuracy is associated with the number of experimental 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 $\boldsymbol{\phi}$ of our classical optimizer (cf. Fig. 2). Similar to θ, $\phi^i = \phi^{i-1} + \Delta\phi^{i-1}$ in the ith iteration, where $\Delta\phi^{i-1}$ is a small increment found using the COBYLA algorithm mentioned before.

The optimization problem defined in (20) is convex in the asymptotic limit of infinitely many rounds, i.e., the objective function $\hat{D}^N_{\boldsymbol{\phi}}(\rho_0, \rho_1)$ is convex for any arbitrary ansatz $\boldsymbol{\phi} = (\boldsymbol{\phi}_0, \boldsymbol{\phi}_1)$ as $N \to \infty$. This conclusion stems from the strong convexity of the exact trace distance and (19). Essentially, as $D(\rho_0, \rho_1)$ is strongly convex concerning any pair of states, and $\hat{D}^\infty(\rho_0, \rho_1) = D(\rho_0, \rho_1)$ according to (19), it

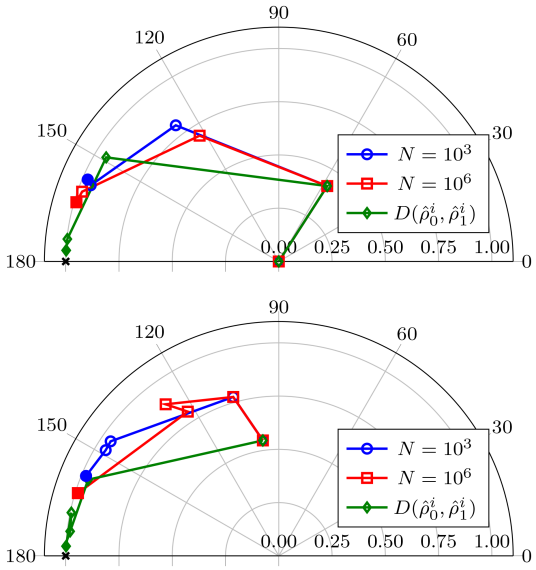

FIGURE 3. In this polar representation, we visualize the convergence of our estimations toward the maximally distinguishable pair for the amplitude damping channel ($\kappa = 0.9$) using two distinct ansatz (top and bottom). The radial and angular axes are represented as $\hat{R}^i$ and $\hat{\beta}^i$, respectively (cf. Section IV-D). The blue (red) line illustrates the 1st, 5th, 10th, 15th, and 20th estimations based on 10^3 (10^6) experimental runs, capturing the evolution of the estimations toward 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.

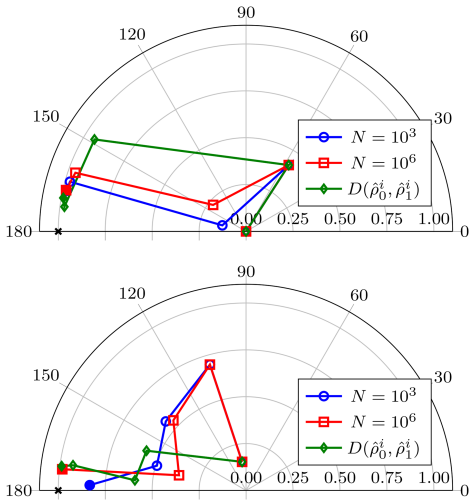

FIGURE 4. In this polar representation, we visualize the convergence of our estimations towards the maximally distinguishable pair for the Pauli channel ($\gamma = 0.9$) using two distinct ansatz (top and bottom). The radial and angular axes are represented as $\hat{X}^i$ and $\hat{\beta}^i$, respectively (cf. Section IV-D). The blue (red) line illustrates the 1st, 5th, 10th, 15th, and 20th estimations based on 10^3 (10^6) experimental runs, capturing the evolution of the estimations toward 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.

follows that $\hat{D}^{\infty}(\rho_0, \rho_1)$ also maintains strong convexity. In addition, for every ansatz $\boldsymbol{\phi}$, we can correlate a pair of states $\rho_{0,1} = \mathcal{N}[\mathcal{U}_{\boldsymbol{\phi}_{0,1}}(|0\rangle\langle 0|)]$, implying that $\hat{D}^{\infty}_{\boldsymbol{\phi}}(\rho_0, \rho_1)$ remains strongly convex concerning any arbitrary ansatz.

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–Kraus 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 ith estimation of the optimal ensemble using a given estimation $\hat{D}^N$. More specifically, a point in Fig. 3 is defined in terms of the estimated Pauli vectors $\hat{\mathbf{r}}_0^i$ and $\hat{\mathbf{r}}_1^i$ in the following way: its coordinates $\hat{\beta}^i$ and $\hat{R}^i$ are defined by the relative angle between $\hat{\mathbf{r}}_0^i$ and $\hat{\mathbf{r}}_1^i$ and their relative distance on the xy plane, i.e.,

$$\hat{\beta}^i := \arccos(\hat{\mathbf{r}}_0^i \cdot \hat{\mathbf{r}}_1^i) \tag{21}$$

and

$$\hat{R}^i := \frac{1}{2}\sqrt{(\hat{x}_1^i - \hat{x}_0^i)^2 + (\hat{y}_1^i - \hat{y}_0^i)^2}. \tag{22}$$

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

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 lies on the equatorial plane of the Bloch sphere. As discussed in Section IV-A, the orthogonal pair of states that maximizes (15) have their Pauli vector components in $\mathbf{e}_{k^*}$. For the amplitude damping channel, these coefficients are

$$\alpha_{x,y} = \sqrt{1-\kappa} \qquad \alpha_z = 1 - \kappa \tag{23}$$

implying in $\alpha_{x,y} \geq \alpha_z$ since $\kappa \in [0, 1]$.

In this relation, the equality holds for $\kappa = 0$ (the ideal channel) and $\kappa = 1$ (the completely depolarizing channel), in which any and none pair of states maximizes (15), respectively. For $\kappa \in (0, 1)$, we have $\alpha_x = \alpha_y > \alpha_z$, which implies by the rationale exposed in Section IV-A that any pair of antilinear vectors expressed as a linear combination of $\mathbf{e}_x$ and $\mathbf{e}_y$ corresponds to a maximally distinguishable pair of states. Therefore, we say for short that any pair of states lying on the equatorial plane (defined by $\mathbf{e}_x$, $\mathbf{e}_y$) maximizes (15).

For the Pauli channels parameterized according to (6), the maximally distinguishable pair of states is $\{|+\rangle\langle +|, |-\rangle\langle -|\}$.

IEEE Transactions on Quantum Engineering

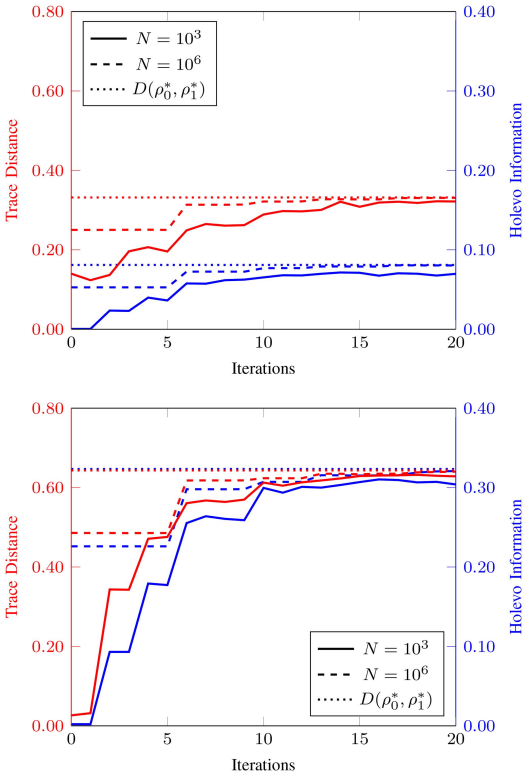

FIGURE 5. Optimal signaling for the Pauli channel. Red (blue) lines are the trace distance (Holevo information) estimations as a function of the output states of the *i*th iteration of our algorithm (cf. Fig. 2) for the Pauli channel with $\gamma = 0.3$ (top) and $\gamma = 0.6$ (bottom). In dotted, we have the corresponding exact values. Data shown here are the average estimation of ten executions.

As before, this comes from an inspection of the coefficients α_k. For a general Pauli channel with probability distribution p_μ, we have [34]

$$\alpha_k = p_0 + p_k - \sum_{k' \neq k} p_{k'}. \tag{24}$$

Choosing p_μ as the ordered eigenvalues of (6) implies in $\alpha_x^2 \geq \alpha_y^2 \geq \alpha_z^2$, where the equality holds $\gamma = 0$ and $\gamma = 1$, the completely depolarizing and ideal channels, respectively. For $\gamma \in (0, 1)$, $\alpha_x^2 \geq \alpha_y^2 \geq \alpha_z^2$ and $\alpha_{k^*}^2 = \alpha_1^2$. Therefore, the pair of states in the direction $\mathbf{e}_x$ will maximize (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. Section 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 Section IV-B. By comparing trace distance

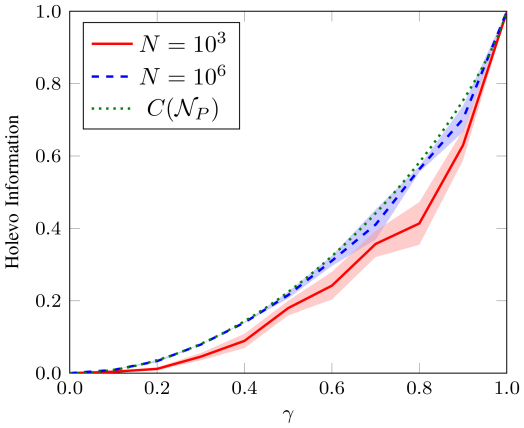

FIGURE 6. Comparing signal processing and information theory approaches for the Pauli channel. The green line illustrates the classical capacity of the Pauli channel [34] as a function of the correlation matrix parameter γ [cf. (6)]. The red and blue lines represent the maximum Holevo information for the 20th estimation of the maximally distinguishable pair with respect to γ, with solid and dashed lines denoting results based on 10^3 and 10^6 experimental runs, respectively.

estimations based on 10^3 and 10^6 experimental 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 utilizing 10^6 experimental 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. 5 and 6 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 maximizes (13) and approximates (14) but also corresponds to the optimal ensemble needed to attain classical capacity for Pauli channels, as emphasized in Fig. 6. Although valid for all Pauli channels, this conclusion cannot be straightforwardly generalized to other channels since an ensemble consisting solely of two equiprobable orthogonal states can be insufficient to achieve Holevo's capacity [31].

VI. CONCLUSION

By employing the VTDE algorithm presented in [23], 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

numerically. More specifically, we showed that our method approximates the analytical predictions (which we have also derived) as the number of experimental runs N and classical iterations M increases for specific cases of the above-mentioned channels. 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.

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 under the constraints of NISQ devices. This 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 nonequiprobable binary communication systems. This would broaden the range of potential signaling encodings, possibly achieving higher values of Holevo information.

REFERENCES

[1] N. Gisin and R. Thew, "Quantum communication," *Nature Photon.*, vol. 1, no. 3, pp. 165–171, Mar. 2007, doi: 10.1038/nphoton.2007.22.

[2] H. J. Kimble, "The quantum internet," *Nature*, vol. 453, no. 7198, pp. 1023–1030, Jun. 2008, doi: 10.1038/nature07127.

[3] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, no. 6412, Sep. 2018, Art. no. eaam9288, doi: 10.1126/science.aam9288.

[4] C. Simon, "Towards a global quantum network," *Nature Photon.*, vol. 11, no. 11, pp. 678–680, Nov. 2017, doi: 10.1038/s41566-017-0032-0.

[5] N. Cerf, G. Leuchs, and E. Polzik, *Quantum Information With Continuous Variables of Atoms and Light*. London, U.K.: Imperial College Press, Feb. 2007, doi: 10.1142/p489.

[6] M. S. Winnel, J. J. Guanzon, N. Hosseinidehaj, and T. C. Ralph, "Achieving the ultimate end-to-end rates of lossy quantum communication networks," *npj Quantum Inf.*, vol. 8, no. 1, Nov. 2022, Art. no. 129, doi: 10.1038/s41534-022-00641-0.

[7] H. Al-Hraishawi, J. U. Rehman, M. Razavi, and S. Chatzinotas, "Characterizing and utilizing the interplay between quantum technologies and non-terrestrial networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1937–1957, 2024, doi: 10.1109/OJCOMS.2024.3380508.

[8] D. F. Walls, "Squeezed states of light," *Nature*, vol. 306, no. 5939, pp. 141–146, Nov. 1983, doi: 10.1038/306141a0.

[9] G. Cariolaro and G. Pierobon, "Performance of quantum data transmission systems in the presence of thermal noise," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 623–630, Feb. 2010, doi: 10.1109/TCOMM.2010.02.080013.

[10] G. Cariolaro and G. Pierobon, "Theory of quantum pulse position modulation and related numerical problems," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1213–1222, Apr. 2010, doi: 10.1109/TCOMM.2010.04.090103.

[11] R. L. Cook, P. J. Martin, and J. M. Geremia, "Optical coherent state discrimination using a closed-loop quantum measurement," *Nature*, vol. 446, no. 7137, pp. 774–777, Apr. 2007, doi: 10.1038/nature05655.

[12] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum metrology," *Phys. Rev. Lett.*, vol. 96, no. 1, Jan. 2006, Art. no. 010401, doi: 10.1103/PhysRevLett.96.010401.

[13] F. Benatti and D. Braun, "Sub-shot-noise sensitivities without entanglement," *Phys. Rev. A*, vol. 87, no. 1, Jan. 2013, Art. no. 012340, doi: 10.1103/PhysRevA.87.012340.

[14] M. L. Shcherbatenko, M. S. Elezov, G. N. Goltsman, and D. V. Sych, "Sub-shot-noise-limited fiber-optic quantum receiver," *Phys. Rev. A*, vol. 101, no. 3, Mar. 2020, Art. no. 032306, doi: 10.1103/PhysRevA.101.032306.

[15] A. Vázquez-Castro and B. Samandarov, "Quantum advantage of binary discrete modulations for space channels," *IEEE Wireless Commun. Lett.*, vol. 12, no. 5, pp. 903–906, May 2023, doi: 10.1109/LWC.2023.3249282.

[16] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, "Quantum-process tomography: Resource analysis of different strategies," *Phys. Rev. A*, vol. 77, no. 3, Mar. 2008, Art. no. 032322, doi: 10.1103/PhysRevA.77.032322.

[17] P. Horodecki and A. Ekert, "Method for direct detection of quantum entanglement," *Phys. Rev. Lett.*, vol. 89, no. 12, Aug. 2002, Art. no. 127902, doi: 10.1103/PhysRevLett.89.127902.

[18] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, "Direct estimations of linear and nonlinear functionals of a quantum state," *Phys. Rev. Lett.*, vol. 88, no. 21, May 2002, Art. no. 217901, doi: 10.1103/PhysRevLett.88.217901.

[19] C. Macchiavello and M. F. Sacchi, "Detecting lower bounds to quantum channel capacities," *Phys. Rev. Lett.*, vol. 116, no. 14, Apr. 2016, Art. no. 140501, doi: 10.1103/PhysRevLett.116.140501.

[20] J. U. Rehman, Y. Jeong, and H. Shin, "Directly estimating the Holevo capacity of discrete Weyl channels," *Phys. Rev. A*, vol. 99, no. 4, Apr. 2019, Art. no. 042312, doi: 10.1103/PhysRevA.99.042312.

[21] J. Bae and L.-C. Kwek, "Quantum state discrimination and its applications," *J. Phys. A: Math. Theor.*, vol. 48, no. 8, Jan. 2015, Art. no. 083001, doi: 10.1088/1751-8113/48/8/083001.

[22] M. M. Wilde, *Quantum Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2013, doi: 10.1017/CBO9781139525343.

[23] R. Chen, Z. Song, X. Zhao, and X. Wang, "Variational quantum algorithms for trace distance and fidelity estimation," *Quantum Sci. Technol.*, vol. 7, no. 1, Dec. 2022, Art. no. 015019, doi: 10.1088/2058-9565/ac38ba.

[24] R. LaRose, A. Tikku, É. O'Neel-Judy, L. Cincio, and P. J. Coles, "Variational quantum state diagonalization," *npj Quantum Inf.*, vol. 5, no. 1, Jun. 2019, Art. no. 57, doi: 10.1038/s41534-019-0167-6.

[25] M. Cerezo et al., "Variational quantum algorithms," *Nature Rev. Phys.*, vol. 3, no. 9, pp. 625–644, Sep. 2021, doi: 10.1038/s42254-021-00348-9.

[26] H. Shin and M. Z. Win, "MIMO diversity in the presence of double scattering," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2976–2996, Jul. 2008, doi: 10.1109/TIT.2008.924672.

[27] J. U. Rehman and H. Shin, "Simultaneous communication and parameter estimation of Pauli channels," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2022, pp. 648–653, doi: 10.1109/ICC45855.2022.9838702.

[28] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge, U.K.: Cambridge Univ. Press, 2010, doi: 10.1017/CBO9780511976667.

[29] C. W. Helstrom, "Quantum detection and estimation theory," *J. Statist. Phys.*, vol. 1, no. 2, pp. 231–252, Jun. 1969, doi: 10.1007/BF01007479.

[30] H. Yuen, R. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 125–134, Mar. 1975, doi: 10.1109/TIT.1975.1055351.

[31] D. W. Berry, "Qubit channels that achieve capacity with two states," *Phys. Rev. A*, vol. 71, no. 3, Mar. 2005, Art. no. 032334, doi: 10.1103/PhysRevA.71.032334.

[32] A. Kandala et al., "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, vol. 549, no. 7671, pp. 242–246, Sep. 2017, doi: 10.1038/nature23879.

[33] B. Commeau, M. Cerezo, Z. Holmes, L. Cincio, P. Coles, and A. Sornborger, "Variational Hamiltonian diagonalization for dynamical quantum simulation," 2020, *arXiv:2009.02559*, doi: 10.48550/arXiv.2009.02559.

[34] K. Siudzińska, "Classical capacity of generalized Pauli channels," *J. Phys. A: Math. Theor.*, vol. 53, no. 44, Oct. 2020, Art. no. 445301, doi: 10.1088/1751-8121/abb276.