




# Benchmarking regularisation methods for quantum process tomography on NISQ devices

I. J. David<sup>1,a</sup>, I. Sinayskiy<sup>1,2,b</sup> , and F. Petruccione<sup>1,2,3,c</sup>

<sup>1</sup> School of Chemistry and Physics, University of KwaZulu-Natal, Durban, South Africa

<sup>2</sup> National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, South Africa

<sup>3</sup> School of Data Science and Computational Thinking and Department of Physics, Stellenbosch University, Stellenbosch 7604, South Africa

Received 18 April 2023 / Accepted 2 December 2023 / Published online 9 January 2024

© The Author(s) 2024

**Abstract** Quantum process tomography (QPT) is a crucial tool for characterizing and validating quantum devices and quantum algorithms. However, the problem of finite sampling leads to an estimated process matrix which is non-positive semi-definite (non-PSD), which can yield a reconstructed quantum channel that is non-physical. To address this problem, various methods have been proposed to correct the issue of finite sampling in the estimation of the process matrix. In this work, we perform a comparison of regularisation methods that will be used to tackle the problem of finite sampling in QPT. For this comparison we simulate some common single qubit quantum channels. We use two metrics, the minimum eigenvalue of the Choi matrix and the fidelity, to compare the effectiveness of these methods. Our results show that the spectral transformations perform the best overall in dealing with finite sampling present in reconstructing the quantum channel in the NISQ era.

## 1 Introduction

Quantum process tomography (QPT) is an essential tool for characterizing and validating quantum devices and quantum algorithms [1]. It is the process of determining the characteristics of a quantum channel by performing measurements on various input states and output states. The process matrix obtained from QPT contains complete information about the quantum channel, such as the probability of obtaining a specific output state given a particular input state [1, 2].

However, a significant challenge in QPT arises due to the limited number of measurements that can be performed. This results in the problem of finite sampling [3], which can lead to significant errors in the estimated process matrix and leads to a process matrix that is not positive semi-definite (PSD). This yields a reconstructed quantum channel that is non-physical. To address this problem, various methods have been proposed to correct the issue of finite sampling in the estimation of the process matrix [3–5].

There are many situations that arise in science where matrices obtained from experiment are not PSD, one such example is the issue of a non-PSD kernel matrix in quantum machine learning [6]. The authors of [6] proposed various methods to correct the problem of the kernel matrix being non-PSD. These methods make use of semi-definite programming [6] as well as spectral transformations [6–8] to correct the kernel matrix. We propose the use of these methods, as well as the methods used in [3, 4], to correct the non-PSD process matrix obtained after performing the QPT. We shall use the term regularisation methods, as a blanket term that refers to all the methods used to correct the non-PSD process matrix.

In recent years, Noisy Intermediate-Scale Quantum (NISQ) devices have emerged as a promising platform for quantum computing. These devices are characterized by their relatively small size, limited coherence times, and high levels of noise. While the issue of noise is a significant challenge for NISQ devices, it is important to note that finite sampling will also be a challenge for both NISQ and fault-tolerant quantum computers. Researchers have

<sup>a</sup> e-mail: [ian.david@nithecs.ac.za](mailto:ian.david@nithecs.ac.za)

<sup>b</sup> e-mail: [sinayskiy@ukzn.ac.za](mailto:sinayskiy@ukzn.ac.za) (corresponding author)

<sup>c</sup> e-mail: [petruccione@sun.ac.za](mailto:petruccione@sun.ac.za)

developed methods for mitigating noise in NISQ devices, however there have been fewer methods developed for dealing with finite sampling. Making the study of the effectiveness of these regularisation methods with dealing with finite sampling in QPT essential for the NISQ era.

In this work we take inspiration from the article [9], which compares different regularisation methods for correcting the non-PSD kernel matrix obtained from implementing quantum kernels on a quantum computer. We aim to perform a comparison of regularisation methods that will be used to tackle the problem of finite sampling in QPT.

Finite sampling is a problem that is prevalent in both the NISQ and fault tolerant setting as we are only ever able to make a finite number of measurements on our quantum device. However, addressing this problem in the NISQ era is of greater importance as repeating circuit evaluations to combat finite sampling in the NISQ era comes with long execution times and high financial costs. This is the reason why the number of times we can measure our system is often fixed to a maximum number depending on the device [10]. It is for this reason that we require methods to address the problem of finite sampling that are efficient and easily generalise to many qubit systems. Here we have selected a few well known regularisation methods to help combat the problem of finite sampling in QPT. These methods rely on optimisation [3, 4, 6] or spectral transformations [11–14] to regularise the process matrices obtained after QPT.

In this study, we perform comparisons of regularization methods for process matrices obtained in quantum process tomography by simulating the circuits with the `qasm_simulator` from the python package `Qiskit` [15]. This allows us to remove device noise from our experiments and focus solely on the effectiveness of the regularization methods at dealing with finite sampling.

Recent literature explores modern methods to improve the efficiency and applicability of QPT, including Matrix Product State representations [16], machine learning techniques [17–19], Bayesian methods [20] and CPTP projection methods [21]. Matrix Product States offer a more compact representation, machine learning leverages pattern recognition, and Bayesian methods enhance robustness. Despite these advancements, the standard QPT method remains essential for benchmarking and validation. This is why we make use of the the standard QPT [2] so that we can compare only regularisation methods and we do not allow for any optimization or complex transformations in the reconstruction step.

To perform the comparison we shall simulate three quantum channels, the amplitude damping (AD) channel, depolarising (DEP) channel and a Pauli (PAU) channel. We construct quantum circuits for these channels and perform a QPT using `qiskit`. For each channel we apply a regularisation method and obtain a 'corrected' process matrix called the regularised process matrix. To compare the effectiveness of these methods we use two metrics, the minimum eigenvalue of the Choi matrix [22] and the fidelity [23]. We use this regularised process matrix to calculate the Choi matrix and check that its eigenvalues are positive ensuring our channel is completely positive and physical. We shall also compute the fidelity between the analytically obtained process matrix and the regularised process matrix to check the quality of our matrix after regularisation. We then make some recommendations on which is the best regularisation method to use based on these two metrics.

This article is outline as follows: in Sect. 2 we introduce some background information about the QPT. In Sect. 3 we shall briefly discuss the regularisation methods we will use in this work. Section 4 will introduce and define the channels we will simulate. In Sect. 5 we present in detail the metrics we will use to compare the regularisation methods and present the procedure we use to perform this comparison. Section 6 will present the results and the discussion of the methods that performed the best. Lastly, in Sect. 7 we make some concluding remarks.

## 2 Quantum process tomography (QPT)

The reduced system dynamics of open quantum systems are usually described by a dynamical map  $\Lambda_t$  where  $t \geq 0$  and  $\Lambda_0 = \mathbb{1}$  i.e. a family of single parameter completely positive and trace preserving (CPTP) maps. If  $\rho(0)$  is the initial state of the system then  $\rho(t) = \Lambda_t \rho(0)$  represents the density operator at some time  $t$  [24]. A dynamical map is also referred to as a quantum channel, these shall be used interchangeably throughout this work.

It is known that a quantum channel  $\Lambda_t$  has a Kraus representation [25]:

$$\Lambda_t \rho = \sum_{\alpha} \hat{K}_{\alpha} \rho \hat{K}_{\alpha}^{\dagger} \quad (1)$$

where  $\hat{K}_{\alpha}$  are the Kraus operators that satisfy  $\sum_{\alpha} \hat{K}_{\alpha}^{\dagger} \hat{K}_{\alpha} = \mathbb{1}$ . In this work we will consider the case of a single qubit channel then the Kraus operators are  $2 \times 2$  matrices. If we choose a complete basis for the Kraus operators of a single qubit channel as  $\{\sigma_0 = \mathbb{1}, \sigma_1, \sigma_2, \sigma_3\}$ , where  $\sigma_i$  are the usual Pauli matrices. Then we can expand the Kraus operators in terms of this basis to get the process matrix representation of the quantum channel for a single

qubit:

$$\Lambda_t \rho = \sum_{m,n=0}^3 \chi_{mn} \sigma_m \rho \sigma_n. \quad (2)$$

Here  $\chi_{mn}$  is a positive and Hermitian  $4 \times 4$  matrix called the process matrix and shall be determined using a quantum process tomography [1, 2]. Now if we know the process Matrix then we have a complete description of the channel  $\Lambda_t$ .

To determine the elements of the process matrix we need to choose a complete set of input states, we choose the states

$$\mathcal{D} = \left\{ \begin{aligned} |\psi_1\rangle &= |0\rangle, |\psi_2\rangle = |1\rangle, \\ |\psi_3\rangle &= |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ |\psi_4\rangle &= |+_y\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \end{aligned} \right\}. \quad (3)$$

The states from the set  $\mathcal{D}$  form a complete set as the projectors constructed from each ket vector in this set can be used to construct the density operator of any physical single qubit state. Now we send these input states through the channel  $\Lambda_t$ . We can prepare the initial state of the qubit as each of the input states. Using quantum state tomography we reconstruct the state after each input state is passed through the channel [3]. Then the formulas from [1, 2] are used to construct the  $\chi$  matrix which allows us to reconstruct the channel  $\Lambda_t$ .

To perform a QPT on a quantum computer we construct a quantum circuit that implements the channel  $\Lambda_t$ , this is usually done using the Stinespring dilation theorem [26], which requires ancillary qubits to simulate the evolution of system and its environment. The next step is to prepare the system qubit in one of the input states above thereafter we apply the quantum circuit. Lastly, we perform a state tomography on the system qubit for each input state and get the corresponding counts. Using the counts obtained from the tomographic circuits we can construct the initial process matrix [2], denoted by  $\chi_{in}$ . The matrix  $\chi_{in}$  obtained, will not be positive and Hermitian. This is due to the fact that we can only make a finite number of measurements on the system qubit, this is known as the problem of finite sampling. Finite sampling presents an issue in both the NISQ and fault tolerant setting, since even fault tolerant quantum computers are also limited to a finite number of measurements. There are many methods that attempt to fix the problem of finite sampling in quantum tomography by using optimisation techniques [3, 4, 6] and spectral transformations [11–14] to find a process matrix that is positive semi-definite and Hermitian. We use the term regularisation methods as a blanket term that encompasses both optimisation techniques and spectral transformations. The regularisation methods will yield a process matrix, denoted  $\chi_c$ , which is positive semi-definite and Hermitian.

Finite sampling also presents a greater issue in many qubit systems, as the number of measurements and circuit evaluations grow with the number of qubits. Therefore we require regularisation methods which can easily and efficiently generalise to multi-qubit systems, for example the spectral transformations [11–14].

In the following sections we will outline the various regularisation methods that we will compare in this work.

### 3 Regularisation methods For QPT

In this work we make use of six regularisation methods that aim to solve the problem of finite sampling in QPT. In Sect. 5. we shall discuss how we will bench mark and compare each of these regularisation methods. In this section, we will briefly discuss how these methods work and how to implement them. We shall first discuss the regularisation methods that rely on solving convex optimisation problems, these are: Least squares (LS) [4], Maximum Likelihood Estimation (MLE) [3] and Semi-Definite Programming (SDP) [6]. Thereafter, we shall discuss the methods that rely on transforming the spectrum of our initial process matrix, we refer to these as spectral transformations and these are: Thresholding (THR), Tikhonov Regularisation (TIK) and Flipping (FLIP) [11–14]. All of these methods can be easily generalised to many qubit systems. The methods that require optimisation are formulated in general in their respective references [3, 4, 6], and for the spectral transformations, since they only rely on changing the eigenvalues of the process matrix these methods can easily generalise to many qubit systems.

### 3.1 Regularisation by optimisation

#### 3.1.1 Least squares (LS)

To construct the LS optimisation [4] problem we first need to parameterise  $\chi_c$  in terms of parameters that we can optimise. We also define  $\chi_c$  in such a way that it is Hermitian and positive semidefinite for all parameter values. We now parameterise  $\chi_c$  as,

$$\chi_c = \chi_c(x_1, x_2, \dots, x_{16}) = T^\dagger T \quad (4)$$

where  $T = T(x_1, x_2, \dots, x_{16})$  is a  $4 \times 4$  triangular matrix that is a function of 16 real variables  $x_1, x_2, \dots, x_{16}$  and is shown below in matrix form:

$$T = \begin{pmatrix} x_1 & 0 & 0 & 0 \\ x_5 + ix_6 & x_2 & 0 & 0 \\ x_{11} + ix_{12} & x_7 + ix_8 & x_3 & 0 \\ x_{15} + ix_{16} & x_{13} + ix_{14} & x_9 + ix_{10} & x_4 \end{pmatrix}. \quad (5)$$

It is evident from this parameterisation that  $\chi_c$  is positive semidefinite and Hermitian. To find  $\chi_c$  with LS optimisation we need to define an objective function that will be minimised with respect to constraints. We define the objective function by the squared difference between the theoretical and experimental probability distributions for each of the counts obtained from the process tomography. The following projective measurement operators are defined from the set  $\mathcal{D}$  above,

$$\{M_1 = |\psi_1\rangle\langle\psi_1|, M_2 = |\psi_2\rangle\langle\psi_2|, \\ M_3 = |\psi_3\rangle\langle\psi_3|, M_4 = |\psi_4\rangle\langle\psi_4|\}. \quad (6)$$

Next, we consider the input state  $\rho_i = |\psi_i\rangle\langle\psi_i|$  where  $|\psi_i\rangle \in \mathcal{D}$ . Then the theoretical probability of being in the state  $|\psi_j\rangle$  after the application of the channel  $\Lambda_t$  to the initial state  $|\psi_i\rangle$ , denoted  $p_{ij}^{\text{the}}$ , is,

$$p_{ij}^{\text{the}} = \text{Tr}[M_j \Lambda_t(\rho_i)] \quad (7)$$

$$= \text{Tr} \left[ M_j \left( \sum_{m,n=1}^4 (\chi_c)_{mn} \sigma_m \rho_i \sigma_n \right) \right] \quad (8)$$

$$= \sum_{m,n=1}^4 (\chi_c)_{mn} \text{Tr}[M_j \sigma_m \rho_i \sigma_n]. \quad (9)$$

The experimentally obtained probability  $p_{ij}^{\text{exp}}$  of being the initial state  $\rho_i$  and measuring in the state  $|\psi_j\rangle$  is,

$$p_{ij}^{\text{exp}} = \frac{n_{ij}}{N}, \quad (10)$$

where  $n_{ij}$  is the counts obtained from the circuit with input state  $|\psi_i\rangle$  and output state  $|\psi_j\rangle$  and  $N$  being the total number of counts. Now we can define the objective function as,

$$\mathcal{F}(x_1, \dots, x_{16}) = \sum_{i,j=1}^4 (p_{ij}^{\text{exp}} - p_{ij}^{\text{the}})^2 \quad (11)$$

$$= \sum_{i,j=1}^4 \left( \frac{n_{ij}}{N} - \sum_{m,n=1}^4 (\chi_c)_{mn} \text{Tr}[M_j \sigma_m \rho_i \sigma_n] \right)^2. \quad (12)$$

We should also take into consideration that the channel  $\Lambda_t$  should be trace preserving, however this constraint may be too strict when solving the optimisation problem. We weaken this constraint, as they do in [4] and require only that the channel  $\Lambda_t$  be trace non-increasing. This will be one of the constraints that we define for our optimisation

problem. We can now state the optimisation problem that we want to solve that will yield a positive semidefinite and Hermitian matrix  $\chi_c$  that is the closest to  $\chi_{id}$ ,

$$\min_{\{x_1, \dots, x_{16}\}} \mathcal{F}(x_1, \dots, x_{16}) \tag{13}$$

$$\text{such that, } \left( \mathbb{1} - \sum_{m, n=1}^4 (\chi_c)_{mn} \sigma_n \sigma_m \right) \geq 0, \tag{14}$$

$$\chi_c(x_1, \dots, x_{16}) \geq 0. \tag{15}$$

The solution to this problem will yield the optimal values for  $x_1, \dots, x_{16}$  so that  $\chi_c$  is positive semi-definite and Hermitian.

### 3.1.2 Maximum likelihood estimation (MLE)

Maximum likelihood estimation (MLE) is a method of estimating the parameters of an assumed probability distribution, given some observed data [3, 4]. To construct the objective function for MLE one starts with the probability distribution for obtaining the measurement results from the counts obtained in the QPT i.e.  $\prod_{i,j} p_{ij}^{n_{ij}}$ , and takes the negative logarithm which yields,

$$\mathcal{L} = -\log \left( \prod_{i,j} p_{ij}^{n_{ij}} \right) = -\sum_{i,j} n_{ij} \log(p_{ij}), \tag{16}$$

where  $n_{ij}$  is the counts obtained after inputting in the state  $\rho_i$  and measuring the state  $\rho_j$  and  $p_{ij}$  is the probability of measuring the state  $\rho_j$  after inputting  $\rho_i$ . We can now write the objective function in terms of  $\chi_c$ , set up the constraints used in the LS method to yield,

$$\mathcal{L}(x_1, \dots, x_{16}) = -\sum_{i,j} n_{ij} \log(\text{tr}[M_j \Lambda_t(\rho_i)]), \tag{17}$$

$$= -\sum_{i,j} n_{ij} \log \left( \sum_{m,n} \chi_c \text{tr}[M_j \sigma_m \rho_i \sigma_n] \right). \tag{18}$$

We can now define the MLE optimisation problem as,

$$\min_{\{x_1, \dots, x_{16}\}} \mathcal{L}(x_1, \dots, x_{16}) \tag{19}$$

$$\text{such that, } \left( \mathbb{1} - \sum_{m, n=1}^4 (\chi_c)_{mn} \sigma_n \sigma_m \right) \geq 0, \tag{20}$$

$$\chi_c(x_1, \dots, x_{16}) \geq 0. \tag{21}$$

Here, as well we also used the weakened constraint that the channel should be trace non-increasing, just as we did in the case of the least squares method.

### 3.1.3 Semi-definite programming (SDP)

For the Semi-Definite Program (SDP), we parameterise  $\chi_c$  as,

$$\chi_c = \chi_c(y_1, y_2, \dots, y_{16}) \tag{22}$$

where each  $y_i \in \mathbb{C}$  is an element of the  $4 \times 4$  process matrix. The objective function that will be used is

$$\mathcal{S} = \|\chi_c - \chi_{in}\|_F \tag{23}$$

where  $\|A\|_F = \sqrt{\langle A, A \rangle_F}$  denotes the Frobenius norm and  $\langle A, B \rangle_F = \text{Tr}(A^\dagger B)$  is the Frobenius inner product for any matrix  $A$  and  $B$ . This objective function describes the similarity between the initial process matrix,  $\chi_{in}$ , and the parameterised process matrix,  $\chi_c$ .

The SDP can then be formulated as

$$\min_{\{y_1, \dots, y_{16}\}} \mathcal{S}(x_1, \dots, x_{16}) \quad (24)$$

$$\text{such that, } \text{tr}[\chi_c(y_1, \dots, y_{16})] = 1. \quad (25)$$

### 3.2 Regularisation By spectral transformation

The three spectral transformations used in this work first require an eigendecomposition of  $\chi_{in}$ ,

$$\chi_{in} = U D_{in} U^\dagger, \quad (26)$$

where the columns  $\mathbf{u}_1, \dots, \mathbf{u}_4$  of  $U$  are the normalised eigenvectors of  $\chi_{in}$  and

$$D_{in} = \text{diag}(\lambda_1, \dots, \lambda_4) \quad (27)$$

is a diagonal matrix with the corresponding eigenvalues  $\lambda_1, \dots, \lambda_4$  of  $\chi_{in}$ .

Thresholding (THR) requires that the negative eigenvalues are set to zero. Tikhonov Regularisation (TIK) requires that the smallest negative eigenvalue of  $\chi_{in}$  is subtracted from all the eigenvalues of  $\chi_{in}$ . Flipping (FLIP) requires that the negative eigenvalues of  $\chi_{in}$  are multiplied by  $-1$ . When applied, these methods have no effect on  $\chi_{in}$  if  $\chi_{in}$  is already positive semi-definite.

After the spectrum of  $\chi_{in}$  has been transformed by one of the methods (THR, TIK or FLIP),  $\chi_c$  is constructed by the eigendecomposition

$$\chi_{in} = U D' U^\dagger, \quad (28)$$

where  $D'$  is a diagonal matrix with the transformed eigenvalues  $\lambda'_1, \dots, \lambda'_4$ .

Each of these spectral transformations have physical motivations as well. For THR it has been shown that the best approximation of a PSD matrix to a non-PSD matrix is achieved by neglecting all the negative eigenvalues (setting them to zero). In this method, the negative eigenvalues are treated as ‘noise’. This has a motivation because a physical and realizable quantum channel has a positive semi definite matrix and any deviation from this is due to noise. For the FLIP method the motivation is also that a physical channel has a positive semi definite process matrix, and therefore flipping the signs of the negative eigenvalues ensures that the process matrix is PSD while keeping the numerical value of the eigenvalue of the original process matrix. Lastly, for TIK, the motivation for the use of this method is that by subtracting by the smallest eigenvalue we remove the noise obtained during the measurement process.

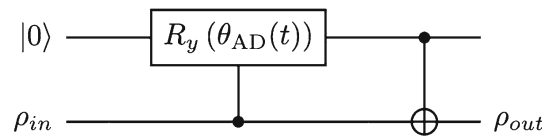
## 4 Single qubit quantum channels to be simulated

In this work we simulate three single qubit channels: Amplitude Damping, Depolarising and a Pauli Channel. This section will briefly define these channels as well as present the quantum circuits we used to simulate these channels.

### 4.1 Amplitude damping (AD) channel

The amplitude damping channel models physical processes such as spontaneous emission and a spin system at high temperature approaching equilibrium with its environment. Of interest to us is the amplitude damping channel for a single qubit. The amplitude-damping channel for a single qubit models energy relaxation from an excited state to the ground state. We can define the AD channel for a single qubit in Kraus form,

$$\Lambda_t^{(\text{AD})}(\rho) = E_0(t)\rho E_0^\dagger(t) + E_1(t)\rho E_1^\dagger(t), \quad (29)$$



**Fig. 1** Circuit implementing the amplitude damping channel for a single qubit, where the probability is given by  $p_{AD}(t)$  and the angle  $\theta_{AD}(t)$  is determined by the formula,  $\theta_{AD}(t) = 2 \arcsin(\sqrt{p_{AD}(t)})$

where,

$$E_0(t) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1 - p_{AD}(t)} \end{pmatrix}, \tag{30}$$

$$E_1(t) = \begin{pmatrix} 0 & \sqrt{p_{AD}(t)} \\ 0 & 0 \end{pmatrix} \tag{31}$$

and,

$$p_{AD}(t) = \frac{2.6(1 - e^{-t})}{3}. \tag{32}$$

To simulate the AD channel on a quantum computer we must construct a quantum circuit for this channel we can do this by using the Stinespring dilation [26]. The quantum circuit is shown in Fig. 1, the angle  $\theta_{AD}(t)$  is determined by from  $p_{AD}(t)$  i.e.

$$\theta_{AD}(t) = 2 \arcsin(\sqrt{p_{AD}(t)}). \tag{33}$$

The qubit in the state  $|\psi\rangle$  is the system qubit and  $|0\rangle$  is the state of the environment.

### 4.2 Depolarising (DEP) channel

The depolarising channel is a model for quantum noise in quantum systems [1]. We can define the DEP channel for a single qubit as,

$$\Lambda_t^{(DEP)}(\rho) = \left(1 - \frac{3}{4}p_{DEP}(t)\right) \rho + \frac{p_{DEP}(t)}{4} \sum_{i=1}^3 \sigma_i \rho \sigma_i, \tag{34}$$

where  $\sigma_i$  for  $i = 1, 2, 3$  are the Pauli matrices and,

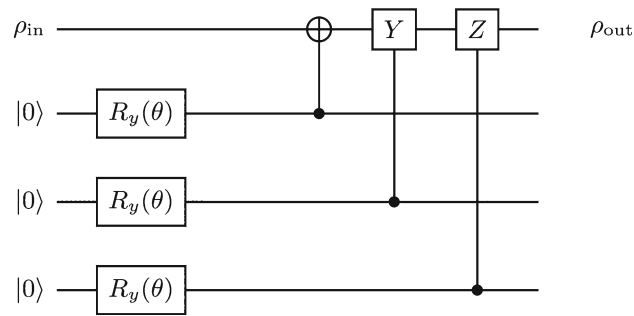
$$p_{DEP}(t) = \frac{1 - e^{-t}}{2}. \tag{35}$$

To construct a quantum circuit for the DEP channel, we use the quantum circuit from [27] which can be seen in Fig. 2.

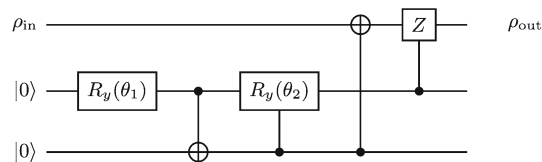
### 4.3 Pauli (PAU) channel

The single qubit Pauli channel applies Pauli matrices to the state with some probability. We can define the Pauli channel as,

$$\Lambda_t^{PAU}(\rho) = p_{PAU}(t)\rho + \left(\frac{1 - p_{PAU}(t)}{2}\right) (\sigma_1 \rho \sigma_1 + \sigma_2 \rho \sigma_2), \tag{36}$$



**Fig. 2** Circuit implementing the depolarizing channel for a single system qubit, for the probability  $p_{\text{DEP}}(t)$ . The angle  $\theta$  is determined by the formula  $\theta(t) = \frac{1}{2} \arccos(1 - 2p_{\text{DEP}}(t))$



**Fig. 3** Quantum circuit implementing the Pauli channel  $\Lambda_t^{(\text{PAU})}$  for probability  $p_{\text{PAU}}(t)$

where,

$$p_{\text{PAU}}(t) = \frac{1 + e^{-t}}{2}. \quad (37)$$

We construct the circuit for the Pauli channel using the approach outlined in [27]. To summarise this approach, we keep in mind that we want to leave the input state  $\rho_{\text{in}}$  unchanged with probability  $p_{\text{PAU}}(t)$  and apply the noise operators  $\sigma_1$  and  $\sigma_2$  to the input state with probability  $(1 - p_{\text{PAU}}(t))/2$ . Refer to Fig. 3. for the quantum circuit that implements the total channel. The parameters  $\theta_1$  and  $\theta_2$  in Fig. 3. are given as,

$$\theta_1(t) = 2 \arccos(\sqrt{p_{\text{PAU}}(t)}), \theta_2 = \frac{\pi}{2}. \quad (38)$$

## 5 Methods for benchmarking the regularisation methods

Here we describe the methods that we will use to compare the regularization methods for the process matrices obtained in QPT. To evaluate the performance of the different regularization methods, we will use two metrics namely fidelity of the process matrix to the analytically obtained process matrix and the minimum eigenvalue of the Choi matrix, which will be obtained from the regularised process matrix. We first describe the experimental pipeline that we used to simulate the quantum channels and compare the regularisation methods. Then we shall discuss the calculation of the two metrics used to benchmark the channels.

### 5.1 Metrics to benchmark regularisation methods

#### 5.1.1 Minimum Eigenvalue of the choi matrix

The problem of finite sampling leads us to reconstruct process matrices that are not positive semi-definite which yield non-physical channels. To check if a channel is completely positive and physical we must check that the Choi matrix [22], denoted  $W(t)$ , is positive semi-definite for all times  $t \geq 0$ . In the case of our simulations we must check that the Choi matrix  $W_c(t)$  obtained from the regularised process matrix  $\chi_c(t)$  is positive semi-definite, to do so we check the minimum eigenvalue of the  $W_c(t)$ . We evaluate the regularisation method by first checking if the method fixes any negative minimum eigenvalues. Also, we want to see how close to the analytical minimum eigenvalue the minimum eigenvalue of  $W_c(t)$  is.



To calculate Choi matrix  $W(t)$  from the process matrix  $\chi(t)$  we make use of the transfer matrix  $F(t)$ , which is a concrete matrix representation of the channel  $\Lambda_t$  [28]. The elements of the transfer matrix are,

$$F_{\alpha,\beta}(t) = \text{Tr}[G_\alpha^\dagger \Lambda_t G_\beta] \tag{39}$$

where  $\{G_\alpha\}$  are a set of orthonormal operators with respect to the Hilbert-Schmidt inner product [28]. We choose the set  $\{G_\alpha\}$  to be the standard matrix basis of  $\mathcal{M}_2(\mathbb{C})$  i.e.  $\{G_1 = |0\rangle\langle 0|, G_2 = |0\rangle\langle 1|, G_3 = |1\rangle\langle 0|, G_4 = |1\rangle\langle 1|\}$ , where  $\{|0\rangle, |1\rangle\}$  are the standard computational basis vectors of the single qubit.

For a given transfer matrix  $F(t)$  we can obtain the Choi matrix  $W(t)$ , for a single qubit this can be written as:

$$W(t) = \frac{1}{2} \sum_{\alpha,\beta=1}^4 F_{\alpha,\beta}(t)(G_\beta \otimes G_\alpha) \tag{40}$$

This is derived by applying  $\Lambda_t$  to a single qubit of the maximally entangled state  $|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ , hence  $W(t) = (\mathbb{1} \otimes \Lambda_t)|\beta_{00}\rangle\langle\beta_{00}|$ .

### 5.1.2 Fidelity of the process matrix

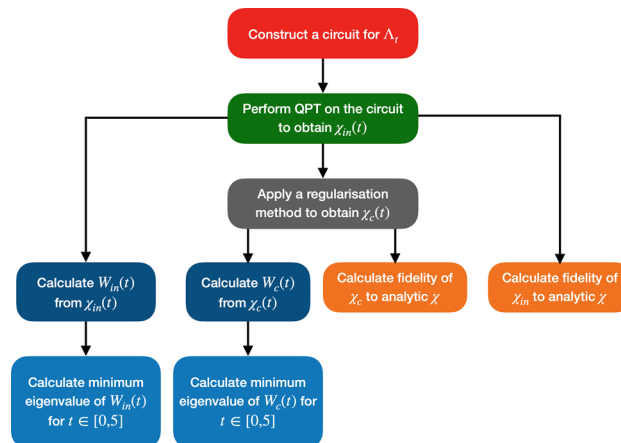
We compute the process fidelity of the process matrix for each time  $t$  after optimisation, using the the following formula [23, 29]:

$$F_p(\chi, \chi_{\text{id}}) = \frac{\text{Tr}[(\sqrt{\chi}\chi_{\text{id}}\sqrt{\chi})^{1/2}]^2}{\text{Tr}[\chi]\text{Tr}[\chi_{\text{id}}]}. \tag{41}$$

This is done to measure the quality of the process matrices obtained after regularisation and to see how close it is to the ideal process matrix  $\chi_{\text{id}}$ . We note that  $F_p \in [0, 1]$ , when  $F_p = 1$  this tells us that the process matrix is the same as the ideal i.e.  $\chi = \chi_{\text{id}}$  and when  $F_p = 0$  the process matrix is far from the ideal process matrix  $\chi_{\text{id}}$ . We acknowledge that the best way to compare the process matrix before and after regularisation would be to use the diamond norm and compare the channels realised by the process matrices. This is because the diamond norm is a completely bounded trace norm and takes into account the channel acting on a subsystem. However, computing the diamond norm requires solving an optimisation problem [30] which adds extra computation that is out of the scope of this work. We instead choose to compute the fidelities of the process matrices before and after regularisation and use this as a simple measure for how well the regularisation methods performed.

## 5.2 Experimental pipeline

The pipeline we shall use to perform our comparison is summarised in Fig. 4. We make use of Python and Qiskit [15] to construct the quantum circuits in Sect. 4 above for each time  $t \in [0, 5]$ , in seconds, with a time step of



**Fig. 4** A flow chart summarising the experimental procedure used to simulate the channel, perform the QPT and regularisation, and compute the metrics we used to compare each of the regularisation methods

0.1 s for the three channels. We shall then perform the measurements that are necessary for a QPT on the circuits. Then, we make use of the `qasm_simulator` in Qiskit which simulates the effect of finite sampling and run the circuits. Once we obtain the results we make use of the formulas from [ ] to construct the process matrix  $\chi_{in}(t)$ . We then apply each of the regularisation methods outlined in Sect. 3 to obtain the process matrix  $\chi_c(t)$ . We then compute  $W_c(t)$  from  $\chi_c(t)$ . We also compute  $W_{in}(t)$ , which is the Choi matrix obtained from  $\chi_{in}(t)$  and will sometimes be referred to as the ‘raw’ Choi matrices, using equations (41) and (42). We then plot the minimum eigenvalue of  $W_{in}(t)$  and  $W_c(t)$  for each regularisation method for each of the three channels. Then we calculate the fidelities of the process matrices  $\chi_c(t)$  and  $\chi_{in}(t)$  with respect to the analytic process matrix and plot the results for each regularisation method and for all three channels.

## 6 Results and discussion

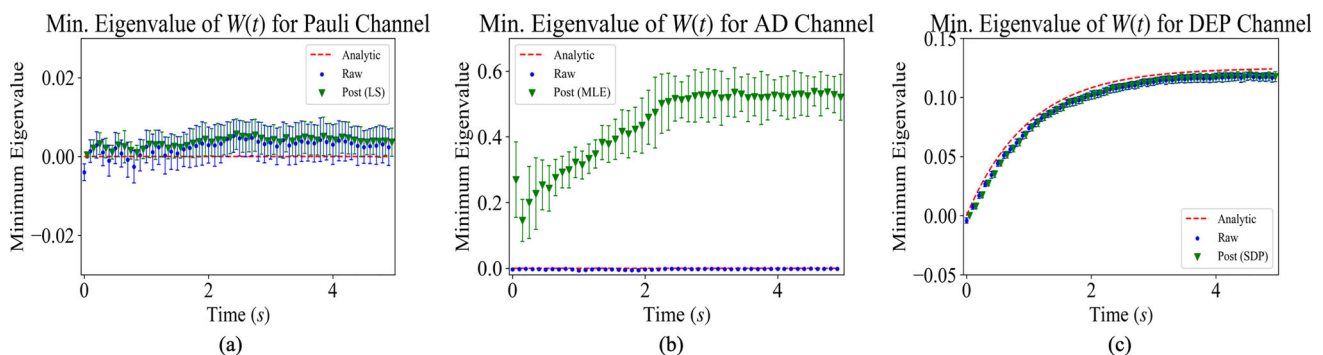
There are two ways in which we will compare the quality of the regularisation methods. First, we will look at the minimum eigenvalue of the Choi Matrices produced from the regularised Process Matrices. Then, we will look at the fidelity of the Choi Matrices to the analytic Choi Matrices compared to the fidelity of the ‘raw’ Choi Matrices to the analytic Choi Matrices. We also only present a portion of the results here, to see the results in detail please refer to supplementary information.

### 6.1 Results

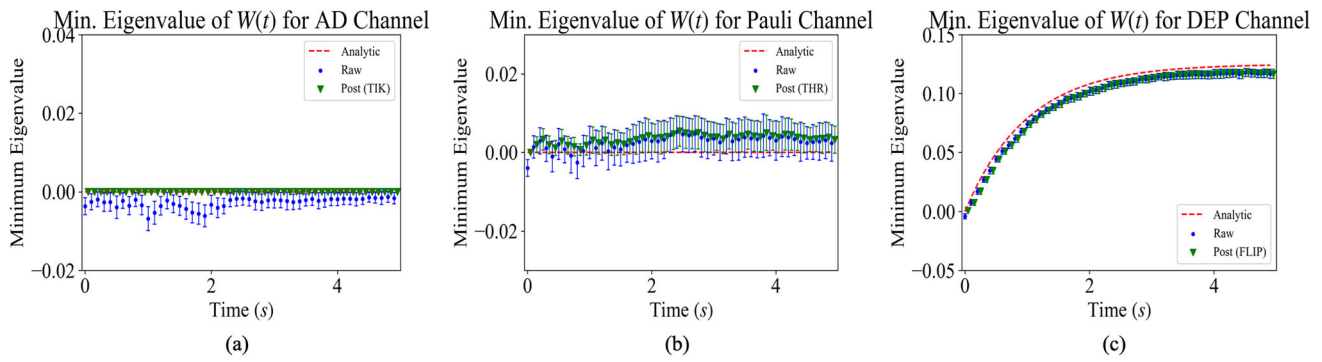
For all channels, the minimum eigenvalues of the Choi Matrices are expected to be greater than or equal to zero in order to satisfy the properties of a physical channel. For the Amplitude Damping Channel as well as the Pauli Channel, the minimum eigenvalues are expected to be zero for all time ( $t \in [0, 5]$ ). However, for the Depolarising Channel, the minimum eigenvalues are expected to follow a curve between 0 and 0.1. This is shown by the analytic curve in Figs. 5 and 6. The mean minimum eigenvalues obtained from the ‘raw’ Choi Matrix for the Amplitude Damping Channel are mostly small negative values, ranging from approximately  $-0.01$  to 0. This is the same for all the regularisation methods except MLE so we only present here the MLE method and TIK method shown in Figs. 5b and 6a, we see that MLE produced much larger deviations than other methods, this is discussed in detail later.

Similarly, for the Pauli Channel in Figs. 5a and 6b, there are a few small negative minimum eigenvalues between approximately  $t = 0$  to  $t = 2$ . Thereafter, the mean minimum eigenvalues are mostly small and positive but not 0. For the Depolarising Channel, the mean minimum eigenvalue for  $t = 0$  is small and negative while the rest of the mean minimum eigenvalues are positive and follow the curve of the analytic minimum eigenvalues.

After the application of the regularisation strategies, we observe that the strategies ensure that the mean minimum eigenvalues are positive for all time ( $t \in [0, 5]$ ). The regularised post-processed Choi Matrices now satisfy the properties of a physical channel for all time. We observe that MLE gives rise to Choi Matrices with



**Fig. 5** Shows the minimum eigenvalues for each of the channels before and after applying the regularisation methods that require optimisation LS, MLE and SDP. **a** A plot of the minimum eigenvalue for the PAU channel after using the LS method. This method works well and the minimum eigenvalue is positive and close to zero. **b** Shows the minimum eigenvalue for the AD channel, we see here that the MLE method performs very poorly as the eigenvalue is positive but deviates very far from zero. **c** Shows the minimum eigenvalue for the DEP channel after SDP, since the minimum eigenvalue was already positive we see that the SDP does not change much but this is good as the regularisation method should not introduce error into our reconstruction



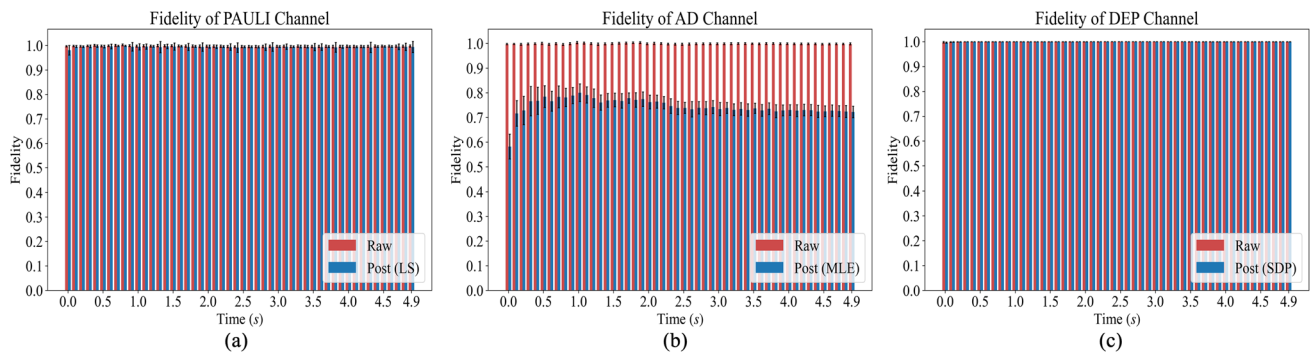
**Fig. 6** Shows the minimum eigenvalues for each of the channels before and after applying the spectral transformation regularisation methods TIK, THR and FLIP. **a** A plot of the minimum eigenvalue for the AD channel after using the TIK method. This method works well and the minimum eigenvalue is zero after regularisation. **b** Shows the minimum eigenvalue for the PAU channel, we see here that the THR method performs very well since the minimum eigenvalue was already positive we see that the SDP does not change much but this is good as the regularisation method should not introduce error into our reconstruction. **c** Shows the minimum eigenvalue for the DEP channel after FLIP, and also leaves the eigenvalues unchanged since they are positive

minimum eigenvalues that are far from their expected values. In addition, there are high standard deviations from the mean minimum eigenvalues.

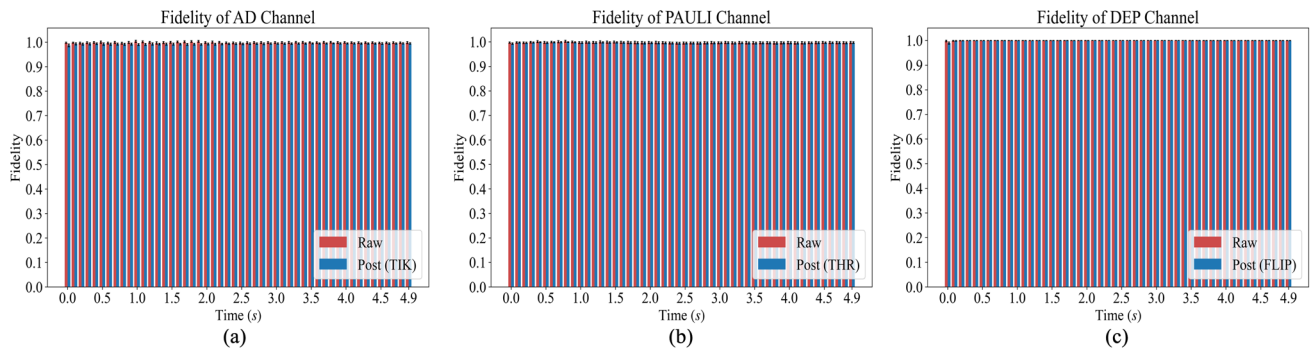
For the Amplitude Damping Channel, TIK and THR produce Choi Matrices with minimum eigenvalues equal to for all time, as expected. This is because these regularisation strategies effectively set the negative eigenvalues to 0. LS, SDP and FLIP produce Choi Matrices with mean minimum eigenvalue that are very close to 0 but not equal to 0.

For the Depolarising Channel in Figs. 5 and 6c, every regularisation strategy except MLE produces a Choi Matrix with mean minimum eigenvalues that follow the analytic curve as well as the curve produced from the ‘raw’ Choi Matrix. For this channel, we observe that the expected minimum eigenvalues are large enough that any deviations that arise from finite sampling do not lead to negative eigenvalues.

For the Pauli Channel, it is observed that while the initial mean minimum eigenvalue is now positive, the rest of the minimum eigenvalues are small positive numbers but not equal to 0. This is because the SDP and spectral transformations, as seen in Fig. 6b where we have plot the minimum eigenvalues after using the THR method, will leave the matrices unchanged if all the eigenvalues are already positive, that is, if the matrices are positive semi-definite. The ‘raw’ matrices for all time  $t \geq 0$  have positive eigenvalues and so these matrices are left the same.



**Fig. 7** Shows the fidelities for each of the process matrices before and after applying the regularisation methods that require optimisation LS, MLE and SDP. **a** A plot of the fidelity for the PAU channel after using the LS method. This method works well and the fidelity is close to one. **b** Shows the fidelity for the AD channel, we see here that the MLE method performs very poorly as the fidelity deviates alot after MLE. **c** Shows the fidelity for the DEP channel after SDP, and the fidelity has the value of 1 for all time  $t \in [0, 5]$



**Fig. 8** Shows the fidelities for each of the process matrices before and after applying the regularisation methods that require optimisation LS, MLE and SDP. **a** A plot of the fidelity for the AD channel after using the TIK method. This method works well and the fidelity is close to one. **b** Shows the fidelity for the AD channel, we see here that the THR method performs very well as the fidelity is close to 1. **c** Shows the fidelity for the DEP channel after FLIP, and the fidelity has the value of 1 for all time  $t \in [0, 5]$

## 6.2 Fidelity

We observe that, for all channels, the fidelity of the ‘raw’ Choi Matrices to the analytic Choi Matrices is approximately 1 for all time ( $t \in [0, 5]$ ). This reveals that the fidelity does not reflect that, in some cases, physical channels are not being simulated.

After the application of the regularisation strategies, we observe that the only strategy that consistently and significantly decreases the fidelity to the analytic process Matrices is MLE as seen for the amplitude damping channel in Fig. 7b.

For the Amplitude Damping Channel, LS also decreases the fidelity but not to any value below 0.9. SDP and the spectral transformations lead to only slight reductions in the fidelity with very small standard deviations.

For the Depolarising Channel, the spectral transformations and LS lead to very slight reductions in the fidelity for time  $t = 0$  and thereafter lead to fidelities of approximately 1. The SDP achieves fidelities equal to 1, that match the fidelities of the ‘raw’ process Matrices to the analytic process Matrices.

This is similar to the Pauli Channel, for which TIK and FLIP lead to very slight reductions in the fidelity for time  $t = 0$  and thereafter lead to fidelities of approximately 1. Thereafter, SDP and THR leads to fidelities equal to 1 which can be seen in Fig. 8.

Any reductions in the fidelities at time  $t = 0$  can be explained by the transformations that are applied to the minimum eigenvalue of the Choi Matrices at that time.

## 6.3 Run times

The regularisation methods were implemented using two python packages `scipy`, `numpy` and `cvxpy`. We used `scipy` and `numpy` for the MLE, LS methods and spectral transformations and we used `cvxpy` to implement the SDP method. All of the regularisation method code was run on a personal computer with an AMD Ryzen 9 7950X 16 core CPU with 128 GB of RAM and an Nvidia 4090 GPU.

While the MLE and LS required run times of approximately 45 mins per channel: that is, approximately 0.54 s per Chi Matrix. The SDP required a run time of 3 mins per channel: that is, approximately 0.036 s per Chi Matrix. Each of the spectral transformations required a run time of 7 s per channel, that is 0.0014 s per Chi Matrix.

## 6.4 Discussion

Although commonly used for solving the problem of finite sampling in quantum state tomographies, we have found that MLE does not well when applied to process tomographies. MLE may be good for state reconstruction since the assumption that the counts are sampled from the same distribution as the theoretically obtained probabilities holds true. However, this assumption does not hold true for channel reconstruction.

LS and SDP solve the problem of the negative eigenvalues by producing Choi Matrices with only positive eigenvalues. However, for the Amplitude Damping Channel and the Pauli Channel, when the minimum eigenvalues are expected to be 0, LS and SDP produce Choi matrices with small positive eigenvalues close to the eigenvalues produced by the ‘raw’ data. It should be noted that SDP achieves fidelities greater than or the same as LS, even though the SDP requires a much shorter run time.

Overall, for the methods that require optimisation, the SDP may be the best choice for channel reconstruction. This is interesting as the SDP merely finds the closest possible positive semi-definite process matrix to the process matrix obtained from the ‘raw’ data, while the LS takes into account the distance between the experimentally obtained probability distribution for the counts and the theoretical distribution. LS squares has access to more information about the channel we are reconstructing, where as SDP only considers the properties of the process matrix and has no extra information about the channel itself.

Similarly, the spectral transformations also only ensure that the process matrices are positive semi-definite and account for no other information about the channel itself. These methods do not alter the process matrices unless the eigenvalues are negative. While the TIK and THR both change the minimum eigenvalues to 0, FLIP flips the sign of the minimum eigenvalue. It is for this reason that TIK and THR are able to produce the correct minimum eigenvalues of 0 while FLIP results in the minimum eigenvalue being small and positive. Overall, these transformations do not often reduce the fidelities and any reductions in the fidelity are very small. These methods are also favourable for channel reconstruction because they require very short run times.

For channel reconstruction, we recommend using TIK and THR as they yielded the best fidelities, and solved the problem of negative eigenvalues in the shortest run times.

## 7 Conclusion

In this work we compared various regularisation methods that deal with the problem of finite sampling in QPT. We compared methods that require solving an optimisation problem i.e. MLE, LS and SDP, as well as spectral transformations. Using the the metrics on the minimum eigenvalue of the Choi matrix and the fidelity we have come to the conclusion that the spectral transformations perform the best overall, when dealing with finite sampling present in reconstructing the quantum channel. The spectral transformations not only yield the best results but they also have the shortest run times making them the most effective for regularising a process matrix obtained from QPT. However, we should acknowledge the fact that we have only compared these methods on three single qubit channels and further testing on more complex channels as well as channels of higher dimension to gain more information on the effectiveness of the regularisation methods.

Future work could look into comparing the regularisation methods and how they deal with both finite sampling and device noise. Another problem one could tackle would be finding better measures for the quality of the process matrices obtained after QPT, as we have seen that high fidelities do not necessarily mean that the reconstructed channel is physical.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1140/epjs/s11734-023-01067-1>.

**Acknowledgements** We would like to thank Ms. S. M. Pillay for her helpful discussions and assistance in proofreading the manuscript.

## Author contributions

IJD and IS came up with the idea for the project. IJD implemented the code and wrote up the manuscript. All Authors discussed the results and reviewed the manuscript.

**Funding** Open access funding provided by University of KwaZulu-Natal. This work is based upon research supported by the National Research Foundation of the Republic of South Africa. Support from the NICIS (National Integrated Cyber Infrastructure System) e-research grant QICSA is kindly acknowledged.

**Data availability** The data for this research is available from the corresponding author upon reasonable request.

## Declarations

**Conflict of interest** Francesco Petruccione the Chair of the Scientific Board and Co-Founder of QUNOVA computing. The authors declare no other competing interests.

**Code availability** The codes used to generate all the data for this research are available from the corresponding author upon reasonable request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were

made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information*, vol. 10, Anniversary. (Cambridge University Press, Cham, 2010)
2. I.L. Chuang, M.A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.* **44**(11–12), 2455–2467 (1997)
3. D.F. James, P.G. Kwiat, W.J. Munro, A.G. White, On the measurement of qubits, in *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*. (World Scientific, Cham, 2005), pp.509–538
4. X.-L. Huang, J. Gao, Z.-Q. Jiao, Z.-Q. Yan, Z.-Y. Zhang, D.-Y. Chen, X. Zhang, L. Ji, X.-M. Jin, Reconstruction of quantum channel via convex optimization. *Sci. Bull.* **65**(4), 286–292 (2020)
5. T. O. Maciel, R. O. Vianna, Optimal estimation of quantum processes using incomplete information: variational quantum process tomography. arXiv preprint [arXiv:1007.2395](https://arxiv.org/abs/1007.2395) (2010)
6. T. Hubregtsen, D. Wierichs, E. Gil-Fuster, P.-J.H. Derks, P.K. Faehrmann, J.J. Meyer, Training quantum embedding kernels on near-term quantum computers. *Phys. Rev. A* **106**(4), 042431 (2022)
7. A.N. Tikhonov, On the solution of ill-posed problems and the method of regularization, in *Doklady Akademii Nauk*, vol. 151, 3rd edn. (Russian Academy of Sciences, 1963), pp.501–504
8. T. Graepel, R. Herbrich, P. Bollmann-Sdorra, K. Obermayer, Classification on pairwise proximity data, in *Advances in Neural Information Processing Systems*, vol. 11, ed. by M. Kearns, S. Solla, D. Cohn (MIT Press, 1998)
9. S.M. Pillay, I. Sinayskiy, E. Jembere, F. Petruccione, Implementing quantum-kernel-based classifiers in the nisq era, in *Artificial Intelligence Research: Second Southern African Conference, SACAIR, Durban, South Africa, December 6–10, 2021, Proceedings*. (Springer, Cham, 2021), pp.257–273
10. D. Kreplin, M. Roth, Reduction of finite sampling noise in quantum neural networks. arXiv preprint [arXiv:2306.01639](https://arxiv.org/abs/2306.01639) (2023)
11. G. Wu, E.Y. Chang, Z. Zhang, An analysis of transformation on non-positive semidefinite similarity matrix for kernel machines, in *Proceedings of the 22nd International Conference on Machine Learning*. (Citeseer, Cham, 2005)
12. V. Roth, J. Laub, M. Kawanabe, J.M. Buhmann, Optimal cluster preserving embedding of nonmetric proximity data. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(12), 1540–1551 (2003)
13. T. Graepel, R. Herbrich, P. Bollmann-Sdorra, K. Obermayer, Classification on pairwise proximity data. *Adv. Neural Inf. Process. Syst.* **11**, 5 (1998)
14. E. Pekalska, P. Paclik, R.P. Duin, A generalized kernel approach to dissimilarity-based classification. *J. Mach. Learn. Res.* **2**, 175–211 (2001)
15. H. Abraham, AduOffei, R. Agarwal et al., Qiskit: An open-source framework for quantum computing. (2019)
16. Z. Yang, Z.-Y. Fan, L.-Z. Mu, H. Fan, Approximate quantum state reconstruction without a quantum channel. *Phys. Rev. A* **98**(6), 062315 (2018)
17. S. Ahmed, F. Quijandría, A.F. Kockum, Gradient-descent quantum process tomography by learning kraus operators. *Phys. Rev. Lett.* **130**(15), 150402 (2023)
18. S. Xue, Y. Liu, Y. Wang, P. Zhu, C. Guo, J. Wu, Variational quantum process tomography of unitaries. *Phys. Rev. A* **105**(3), 032427 (2022)
19. E. Canonici, S. Martina, R. Mengoni, D. Ottaviani, F. Caruso, Machine-learning based noise characterization and correction on neutral atoms nisq devices. arXiv preprint [arXiv:2306.15628](https://arxiv.org/abs/2306.15628) (2023)
20. K. Schultz, Exponential families for Bayesian quantum process tomography. *Phys. Rev. A* **100**(6), 062316 (2019)
21. G.C. Knee, E. Bolduc, J. Leach, E.M. Gauger, Quantum process tomography via completely positive and trace-preserving projection. *Phys. Rev. A* **98**(6), 062336 (2018)
22. M.-D. Choi, Completely positive linear maps on complex matrices. *Linear Algebra Appl.* **10**(3), 285–290 (1975)
23. R. Jozsa, Fidelity for mixed quantum states. *J. Mod. Opt.* **41**(12), 2315–2323 (1994)
24. H. Breuer, F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2002)
25. K. Kraus, General state changes in quantum theory. *Ann. Phys.* **64**(2), 311–335 (1971)
26. W.F. Stinespring, Positive functions on  $C^*$ -algebras. *Proc. Am. Math. Soc.* **6**(2), 211–216 (1955)
27. G. García-Pérez, M.A. Rossi, S. Maniscalco, Ibm q experience as a versatile experimental testbed for simulating open quantum systems. *Npj Quantum Inf.* **6**(1), 1–10 (2020)
28. M. M. Wolf, Quantum channels & operations: Guided tour. Lecture notes available at <http://www-m5.ma.tum.de/fo/wiki/pubM>, vol. 5 (2012)
29. S.A. Uriri, F. Wudarski, I. Sinayskiy, F. Petruccione, M.S. Tame, Experimental investigation of Markovian and non-Markovian channel addition. *Phys. Rev. A* **101**(5), 052107 (2020)
30. J. Watrous, Simpler semidefinite programs for completely bounded norms. *Chic. J. Theor. Comput. Sci.* **8**, 1–19 (2013)