*Review*

# Quantum Tomography: From Markovianity to Non-Markovianity

Tian Luan [1,2,3,*], Zetong Li [2], Congcong Zheng [2], Xueheng Kuang [1], Xutao Yu [2] and Zaichen Zhang [2]

[1] Yangtze Delta Region Industrial Innovation Center of Quantum Science and Technology, Suzhou 215000, China; kuangxueheng@tgqs.net
[2] Quantum Information Center, Southeast University, Nanjing 210096, China; 230218079@seu.edu.cn (Z.L.); zhengcongcong@seu.edu.cn (C.Z.); yuxutao@seu.edu.cn (X.Y.); zczhang@seu.edu.cn (Z.Z.)
[3] China Academy of Electronics and Information Technology, Beijing 100041, China
[*] Correspondence: luantian0417@126.com or luantian@tgqs.net

**Abstract:** The engineering of quantum computers requires the reliable characterization of qubits, quantum operations, and even the entire hardware. Quantum tomography is an indispensable framework in quantum characterization, verification, and validation (QCVV), which has been widely accepted by researchers. According to the tomographic target, quantum tomography can be categorized into quantum state tomography (QST), quantum process tomography (QPT), gate set tomography (GST), process tensor tomography (PTT), and instrument set tomography (IST). Standard quantum tomography toolkits generally consist of basic linear inverse methods and statistical maximum likelihood estimation (MLE)-based methods. Furthermore, the performance of standard methods, including effectiveness and efficiency, has been further developed by exploiting Bayesian estimation, neural networks, matrix completion techniques, etc. In this review, we introduce the fundamental quantum tomography techniques, including QST, QPT, GST, PTT, and IST. We first introduce the details of basic linear inverse methods. Then, the framework of MLE methods with constraints is summarized. Finally, we briefly introduce recent further research in developing the performance of tomography, utilizing some symmetry properties of the target. This review provides a primary getting-start in developing quantum tomography, which promotes quantum computer development.

**Keywords:** quantum characterization, verification, and validation; quantum tomography; linear inverse; maximum likelihood estimation

## 1. Introduction

Quantum tomography, as a significant framework in quantum characterization, verification, and validation (QCVV), is indispensable for manufacturing and developing quantum computers [1,2]. The term "tomography" originates from the Greek words "tomos", meaning "description" or "writing". In the context of quantum tomography as depicted in Figure 1, the term is used to describe a process that involves reconstructing a quantum state by obtaining information about its various slices or projections. The results of tomography are generally considered as the foundation of quantum device calibration [3–7], quantum error mitigation (QEM) [8], and quantum error correction (QEC) [9–11], providing detailed information about qubits, quantum gates, as well as the quantum noise. In this framework, an experimenter prepares a set of experiments consisting of quantum states, circuits, and measurements. Then, a set of data is collected by executing the prepared experiments. Results of interest are derived by performing estimation algorithms.

Based on the common skeleton but different targets, the main techniques are quantum state tomography (QST) [6,12–22], quantum process tomography (QPT) [23–31], and gate set tomography (GST) [5,32–35], process tensor tomography (PTT) [36,37], and instrument set tomography (IST) [38]. The effectiveness of these methods in characterizing the target components within specified assumptions has been proven. Furthermore, some methods have been used in developing real quantum hardware [4,39–41].

Tomography tools typically operate under the assumption that components other than the target are knowable. This implies that the experimenter possesses complete knowledge of non-target components or can acquire their information implicitly through learning algorithms. Relaxing knowable assumptions indicates the progress of comprehensiveness and practicality at a cost of complexity in terms of the number of experiments required and the post-processing.

The tomography first emerged by proposing QST [23–31] with the assumption that quantum gates and measurements are known to the experimenter. Therefore, one can construct a complete basis of the Hilbert–Schmidt space corresponding to the target quantum state. The black-box quantum state is mathematically reconstructed utilizing measurement probabilities and the knowledge of quantum gates and measurements. Analogously, the QPT [23–26] follows this approach, where the target is a black box quantum process, and the knowable assumptions are made to non-target quantum gates, quantum states, and measurements. Utilizing non-target quantum gates to compose symmetry state preparation and measurement (SPAM) circuits and combining them with the known quantum states and measurements, complete bases are constructed before and after the black-box quantum process, respectively. Consequently, the QPT tries to recover a completely positive (CP) trace-preserving (TP) map between the input and output quantum states to represent the target process.

Developing from the QPT, GST was proposed based on the fact that all components are noisy. While applying tomography to quantum gates, QPT becomes inaccurate in the presence of SPAM errors. GST addressed this issue by independently modeling quantum gates, initial states, and measurements as tomography targets. Therefore, the GST explicitly separates the SPAM error into intrinsic and extrinsic errors. Intrinsic errors are inherent in the state preparation and measurement process, while extrinsic errors are introduced by performing quantum gates in SPAM circuits, facilitating QEC and QEM. Furthermore, GST exhibits gauge freedom up to a similarity transformation by a gauge matrix. This implies that an initialization error cannot be easily distinguished from a faulty measurement. As a result, GST provides a systematic method to simultaneously characterize quantum gates, initial states, and measurements. However, the scaling with the system size is polynomially worse than QPT due to the requirement to characterize multiple gates at once.

QPT and GST do not account for the dynamic influence of the environment on quantum processes. Nevertheless, compelling evidence, as highlighted in recent research [37], suggests the existence of multi-time system–environment (SE) correlations. Neglecting these SE correlations may introduce systematic errors in the tomography process. For instance, using QPT or GST to perform tomography on a black-box quantum process with SE correlations in the state preparation may cause the theoretical violation of CP constraints [37]. This can disrupt tomography under the Markovian model and degrade the effectiveness of QEC codes.

PTT [36,37] initially aims to operationally characterize SE correlations. Based on the quantum stochastic process [42], a multi-time-step non-Markovian system can be modeled by a principal quantum system [43], an environment with the same dimensions as the system, as SE unitaries act on both the system and environment dimensions between adjacent time steps, and instruments that the experimenter applies to the system at time steps. In NISQ settings, those instruments, quantum gates, and measurements are the only components accessible to the experimenter. The PTT exploits the process tensor to represent all inaccessible parts consisting of the initial SE state and SE unitaries. By combining informationally complete instruments that span the space of CP and trace non-increasing (TNI) quantum operations, the process tensor is recovered. However, the PTT requires an exponential number of experiments with respect to the Markovian order, which limits large-scale implementations.

Growing out of PTT, the IST [38] was proposed to perform GST in non-Markovian situations. It tackles the inconsistency in characterizing SE correlations and instruments. For example, two inconsistent process tensors may be generated by two sets of imperfect

instruments. Both SE correlations and instruments are targets. The IST successfully estimates the instrument and the SE correlations that are consistent with the measurement probabilities. However, it still suffers from the exponential number of experiments and the high complexity of post-processing.

In this review, we provide a fundamental introduction to quantum tomography techniques, including QST, QPT, GST, PTT, and IST. We briefly introduce the basics and direction of the development of these techniques for readers to facilitate systematic comprehension. This review contains many abbreviations. To facilitate prompt reference for readers, we have compiled the key abbreviations in Table 1, aiming to enhance accessibility and comprehension.

This review is structured as follows. In Section 2, we briefly introduce mathematical representations for quantum tomography. Then, the basic frameworks for Markovian and non-Markovian quantum tomography are introduced in Sections 3 and 4, respectively. In Section 5, we introduce the maximum likelihood estimation method as an indispensable statistical tool in the standard toolkit. Finally, further developments to enhance the performance of tomography are discussed in Section 6.

**Table 1.** The major abbreviations of this review.

| Terminology Type | Abbreviation | Explanation |
|---|---|---|
| Tomography | (L)QST | (Linear inverse) quantum state tomography |
| | (L)QPT | (Linear inverse) quantum process tomography |
| | (L)GST | (Linear inverse) gate set tomography |
| | PPT | Process tensor tomography |
| | IST | Instrument set tomography |
| | SPAM | State preparation and measurement |
| Mathematical | CP | Completely positive |
| | TP | Trace-preserving |
| | PTM | Pauli transfer matrix |
| | CJI | Choi–Jamiolkowski isomorphism |
| | MLE | Maximum likelihood estimation |
| Noise-related | SE | System–environment |
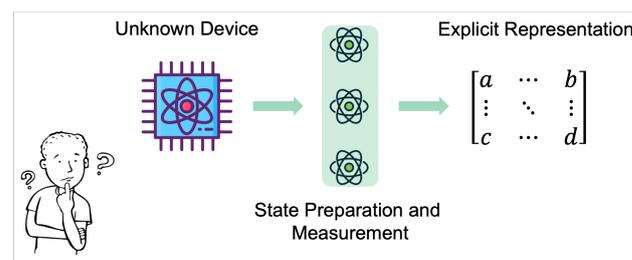


**Figure 1.** Schematic view of quantum tomography. For an unknown quantum device, which can be a quantum state preparation circuit, quantum process, etc., one can obtain the explicit representation with some SPAM circuits.

## 2. Preliminaries

Prior to delving into quantum tomography techniques, it is essential to establish mathematical representations for physical entities. This section introduces the mathematical expressions for quantum states, gates, and measurements. Subsequently, a calligraphic notation will be employed to depict the physical entities associated with quantum operations, such as quantum gates ($\mathcal{G}$) and measurements ($\mathcal{M}$). The quantum state is written in the Roman alphabet, such as $\rho$. We primarily utilize the Pauli transfer matrix (PTM) formalism. Then, the Choi–Jamiolkowski isomorphism (CJI) representation is additionally introduced [37,44].

In the PTM representation, a quantum state $\rho$ on a $d$-dimensional Hilbert space $\mathcal{H}_d$ is represented by a superoperator $|\rho\rangle\rangle$ as a vector in the $d^2$-dimensional Hilbert–Schmidt space. We specify the basis of Hilbert–Schmidt space to be Pauli basis without loss of generality. Hence, the elements of $|\rho\rangle\rangle$ represent the coefficients of the corresponding normalized Pauli matrix [5],

$$|\rho\rangle\rangle = \sum_k \text{Tr}(P_k\rho)|k\rangle\rangle, \tag{1}$$

where $P_k \in \{I/\sqrt{2}, X/\sqrt{2}, Y/\sqrt{2}, Z/\sqrt{2}\}^{\otimes n}$ is the $k$-th normalized Pauli matrix. Furthermore, a quantum measurement, which is described as an operator $M$, can be represented by a set of row vectors [5], i.e.,

$$\langle\langle M| = \sum_k \text{Tr}(P_k M)\langle\langle k|, \tag{2}$$

where $\langle\langle k| = |k\rangle\rangle^\dagger$. Let $\langle\langle A|B\rangle\rangle$ denote $\text{Tr}(AB)$, the probability of performing measurement $M$ on $\rho$, be represented as [5]

$$p = \langle\langle M|\rho\rangle\rangle. \tag{3}$$

A quantum process $\mathcal{A}$ can be conveniently represented by a PTM defined as

$$(A)_{ij} = \text{Tr}\{P_i \mathcal{A}(P_j)\}. \tag{4}$$

This definition describes the mapping between input and output quantum states. Notably, the output state $\rho_{out}$ is represented by the left product of input state $\rho_{in}$ that

$$|\rho_{out}\rangle\rangle = A|\rho_{in}\rangle\rangle. \tag{5}$$

Hence, the experimental probability result of the experiment, which involves a quantum state $\rho$, quantum process $\mathcal{A}$, and a measurement $\mathcal{M}$, can be represented as

$$p = \text{Tr}\{\mathcal{M} \circ \mathcal{A}(\rho_i)\} = \text{Tr}\{M\mathcal{A}(\rho_i)\} = \langle\langle M|A|\rho\rangle\rangle. \tag{6}$$

The PTM constructs a mapping from the input state to the output state. It is easy to check and constrain the trace characteristics by examining the first row of PTM. For example, the trace-preserving (TP) constraints indicate the first row of PTM to be $[1, 0, 0, \ldots, 0]$ [5]. However, it is difficult to intuitively represent the CP property.

Therefore, we also introduce the CJI representations. For a quantum process $\mathcal{G}$, the CJI matrix is defined as [37,44]

$$\xi_\mathcal{G} = \sum_{ij} \mathcal{G}(|i\rangle\langle j|) \otimes |i\rangle\langle j|, \tag{7}$$

where the output state can be determined by

$$\rho_{out} = \text{Tr}_{in}[(I \otimes \rho_{in}^T)\xi_\mathcal{G}]. \tag{8}$$

Since the CJI matrix represents a quantum state without normalization, it is also referred to as the Choi state. The CP constraints for $\mathcal{G}$ are equivalent to the positive semi-definiteness of the $\xi_\mathcal{G}$.

Additionally, PTM representation $A$ can be easily transformed into Choi–Jamiolkowski isomorphism (CJI) [37,44] representation $\xi_\mathcal{A}$, and vice versa, that

$$\xi_{\mathcal{A}} = \frac{1}{d^2} \sum_{ij} A_{ij} P_j \otimes P_i^T, \tag{9}$$

$$A_{ij} = \mathrm{Tr}\left\{ \left( P_j \otimes P_i^T \right) \xi_{\mathcal{A}} \right\}. \tag{10}$$

### 3. Basic Markovian Tomography Framework

Considering a Markovian quantum experiment conducted by an experimenter with specified components consisting of quantum state $\rho$, circuit $\mathcal{G}$, and measurement $\mathcal{M}$, the probability of the experiment, whose samples are the only accessible outputs to the experimenter, is given by [5]

$$p = \mathrm{Tr}[\mathcal{M} \circ \mathcal{G}(\rho)], \tag{11}$$

where the PTM representation is

$$p = \langle\!\langle M | G | \rho \rangle\!\rangle. \tag{12}$$

Current quantum devices generally provide limited initial quantum states and quantum measurements. Therefore, the specified quantum state and measurement are referred to as the state preparation and measurement (SPAM) circuits that act on the initial states and native measurements. Then, the probability can be rewritten in the following form [5]

$$p = \mathrm{Tr}[\mathcal{M} \circ \mathcal{E} \circ \mathcal{G} \circ \mathcal{F}(\rho)] \tag{13}$$
$$= \langle\!\langle M | E\, G\, F | \rho \rangle\!\rangle, \tag{14}$$

where $\mathcal{E}$ and $\mathcal{F}$ represent the state preparation circuit and measurement circuit, respectively. In the following, the quantum state and measurement are referred to as the SPAM circuits acting on the initial state and native measurement, respectively. Moreover, the specified quantum states and measurements with index in experiments are implemented by specifying the SPAM circuits without loss of generality [5]

$$\rho_i : |\rho_i\rangle\!\rangle = F_i |\rho\rangle\!\rangle, \tag{15}$$
$$\mathcal{M}_i : \langle\!\langle M_i | = \langle\!\langle M | E_i. \tag{16}$$

The basic Markovian quantum tomography framework can be summarized in the following steps:

(1) *Experiment preparation*. Prepare a set of experiments, where each experiment consists of the quantum state, circuit, and measurement.
(2) *Data collection*. Execute the prepared experiments and record the measurement samples.
(3) *Tomography reconstruction*. Reconstruct the tomographic target by performing a post-process algorithm based on the collected data.

In general, quantum tomography techniques assume that components except for tomographic targets are perfectly implemented as the knowledge to the experimenter, or implicitly learnable by learning algorithms. Experimenters usually prepare experiments that form a tomographically complete basis for the target to obtain sufficient information.

In this section, based on the linear inversion, we will introduce the basic methods of QST, QPT, and GST, i.e., LQST, LQPT, and LGST, respectively.

### 3.1. Basic QST

The task of QST is to find the explicit representation of an unknown quantum state $\sigma \in \mathcal{H}_d$ [6,13–18]. The construction of a tomographically complete basis for the state is constructed by combining quantum circuits and measurements, which can be solely accomplished by measurements with identity circuits.

Let the set of measurement $\{\mathcal{M}_i\}_{i=0}^{d^2}$ be tomographically complete so that $\{\langle\!\langle M_i|\}_{i=0}^{d^2}$ are linear independent. The probability of sampling a desired output value of $\mathcal{M}_i$ is

$$p_i = \langle\!\langle M_i|\sigma\rangle\!\rangle, \tag{17}$$

which indicates the decomposition of the unknown quantum state $\sigma$. Since the measurements are known to the experimenter, the dual set of $\{\langle\!\langle M_i|\}_{i=0}^{d^2-1}$ can be determined as $\{|D_j\rangle\!\rangle\}_{j=0}^{d^2-1}$ by performing the linear inversion, where $\langle\!\langle M_i|D_j\rangle\!\rangle = \delta_{ij}$.

Then, the unknown quantum state can be reconstructed by

$$|\hat{\sigma}\rangle\!\rangle = \sum_{i=0}^{d^2-1} \hat{p}_i|D_i\rangle\!\rangle, \tag{18}$$

where $\hat{p}_i = n_i/n_t$ is the estimated probability, $n_s$ is the number of the desired output value of $\mathcal{M}_i$ in samples, and $n_t$ is the total sampling times.

The LQST requires at least $d^2$ experiments to collect complete information on the unknown quantum state, and the linear inverse process on a $d^2 \times d^2$ matrix. The reconstructed state $\hat{\sigma}$ is represented by a $d^2$-dimensional real vector.

*3.2. Basic QPT*

The task of QPT is to find the explicit representation of an unknown quantum process $\mathcal{G} \in \mathcal{B}(\mathcal{H}_d)$, where $\mathcal{B}(\bullet)$ represents the space of bounded linear operator on the operand space [23–25,27–31]. The process $\mathcal{G}$ can be treated as a map from the space of the input state to the space of the output state. Therefore, to construct a tomographically complete basis for the unknown quantum process, QPT requires a set of quantum states $\{\rho_i\}_{i=0}^{d^2-1}$ and a set of quantum measurements $\{\mathcal{M}_j\}_{i=0}^{d^2-1}$ that simultaneously span $\mathcal{H}_d$, i.e., both $\{|\rho_i\rangle\!\rangle\}_{i=0}^{d^2-1}$ and $\{\langle\!\langle M_j|\}_{j=0}^{d^2-1}$ are linear independent.

Given $\rho_i$ and $\mathcal{M}_j$, the probability is

$$p_{ij} = \langle\!\langle M_j|G|\rho_i\rangle\!\rangle, \tag{19}$$

which indicates the coefficient of the map from $\rho_i$ to $\mathcal{M}_j$. Based on the known knowledge of quantum states and measurements, the dual sets of $\{|\rho_i\rangle\!\rangle\}_{i=0}^{d^2-1}$ and $\{\langle\!\langle M_i|\}_{j=0}^{d^2-1}$ can be determined by $\{\langle\!\langle \beta_k|\}_{k=0}^{d^2-1}$ and $\{|D_l\rangle\!\rangle\}_{l=0}^{d^2-1}$, respectively, by performing the linear inversion. Then, the quantum process can be reconstructed by

$$\hat{G} = \sum_{i,j=0}^{d^2-1} \hat{p}_{ij}|D_j\rangle\!\rangle\langle\!\langle \beta_i|, \tag{20}$$

where $\hat{p}_{ij}$ is the estimated probability.

It can be observed that the construction of a tomographically complete basis requires $d^4$ experiments and processes the linear inversion on the $d^2 \times d^2$ matrix twice. The reconstructed quantum process $\hat{G}$ is saved as a $d^2 \times d^2$-dimensional PTM.

*3.3. Basic GST*

QST and QPT typically assume the perfect implementation of non-target components, serving as knowledge for the experimenter. However, nothing is absolutely perfect. While using QPT to characterize a set of quantum gates that are available to a quantum device, errors introduced by SPAM circuits impact the self-consistency of tomography, which motivate the development of GST [5,32–35]. Compared to QST and QPT, the goal of GST is to self-consistently and completely characterize a set of $n_k$ quantum gates $\{\mathcal{G}_k\}$, an initial

quantum state $\rho$, and a native quantum measurement $\mathcal{M}$. Based on the probability defined in (11), a quantum gate set is defined as [5]

$$\mathbb{G} := \{\rho, \mathcal{M}, \mathcal{G}_0, \cdots, \mathcal{G}_K\}, \tag{21}$$

where $\mathcal{M}$ is a two-outcome measurement and $\mathcal{G}_0$ is the *null gate*, which does nothing for no time. Based on the given gate set $\mathbb{G}$, SPAM gates are defined as

$$\mathbb{F} = \{\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_{d^2}\}, \tag{22}$$

where $\mathcal{F}_i$ is composed of gates in the set $\{\mathcal{G}_k\}$, that is,

$$\mathcal{F}_i = \mathcal{G}_{i_{L_i}} \circ \cdots \circ \mathcal{G}_{i_2} \circ \mathcal{G}_{i_1}, \tag{23}$$

where $L_i$ is the length of the $i$-th SPAM gate. Then, we can also obtain $d^2$ tomographically complete input quantum states $\{\mathcal{F}_j(\rho)\}$ and $d^2$ measurement operators $\{\mathcal{M} \circ \mathcal{F}_i\}$. Therefore, we can obtain $n_k \times d^4$ probabilities,

$$p_{ikj} = \langle\!\langle M | F_i G_k F_j | \rho \rangle\!\rangle \tag{24}$$

$$= \sum_{rs} \langle\!\langle M | F_i | r \rangle\!\rangle \langle\!\langle r | G_k | s \rangle\!\rangle \langle\!\langle s | F_j | \rho \rangle\!\rangle. \tag{25}$$

Likewise, we can rewrite the above equality in the matrix form,

$$R_k = AG_kB, \tag{26}$$

where

$$A = \sum_i |i\rangle\!\rangle \langle\!\langle M | F_i, \quad B = \sum_j F_j | \rho \rangle\!\rangle \langle\!\langle j |, \tag{27}$$

$$(R_k)_{ij} = p_{ikj}. \tag{28}$$

Experimentally measuring the values $p_{ikj}$, we can obtain matrices $\{S_k\}$, whose elements satisfy

$$\mathbb{E}[S_k] = R_k, \quad \forall\, k. \tag{29}$$

Moreover, we define

$$g = S_0, \quad \mathbb{E}[g] = AB. \tag{30}$$

Furthermore, we define $|\tilde{\rho}\rangle\!\rangle$ and $\langle\!\langle \tilde{M}|$, which can be experimentally estimated with and satisfy

$$\mathbb{E}[|\tilde{\rho}\rangle\!\rangle] = \sum_i |i\rangle\!\rangle \langle\!\langle M | F_i | \rho \rangle\!\rangle, \tag{31}$$

$$\mathbb{E}[\langle\!\langle \tilde{M}|] = \sum_j \langle\!\langle M | F_j | \rho \rangle\!\rangle \langle\!\langle j |, \tag{32}$$

where $\langle\!\langle M | F_i | \rho \rangle\!\rangle$ is measurable.

Then, the unknown gate set can be recovered by performing convex optimization technology. Based on high-quality quantum devices, it is known a priori that the measured gates will differ from an ideal (target) set of gates by some very small error. Therefore, we have a target set $\mathbb{G}' = \{|\rho'\rangle\!\rangle, \langle\!\langle E'|, G_0', \cdots, G_K'\}$ and define

$$\tilde{G}_k := g^{-1}S_k \approx B^{-1}R_kB. \tag{33}$$

Then, we can find the estimation of matrix $B$ via the following convex optimization problem,

$$\hat{B} = \arg\min_{\tilde{B}} \sum_{k=1}^{K+1} \text{Tr}\left[\left(\tilde{G}_k - \tilde{B}^{-1}G_k'\tilde{B}\right)^2\right], \tag{34}$$

where $G_{K+1}' = |\rho'\rangle\rangle\langle\langle M'|$ and $\tilde{G}_k = g^{-1}|\tilde{\rho}\rangle\rangle\langle\langle\tilde{M}|$. Lastly, we can recover the unknown gate set with

$$|\hat{\rho}\rangle\rangle := \hat{B}|\tilde{\rho}\rangle\rangle, \tag{35}$$

$$\langle\langle\hat{M}| := \langle\langle\tilde{M}|\hat{B}^{-1}, \tag{36}$$

$$\hat{G}_k := \hat{B}\tilde{G}_k\hat{B}^{-1}. \tag{37}$$

## 4. Basic Non-Markovian Tomography Framework

No system is isolated. Based on the quantum stochastic process theory [42,45], an open quantum system subjected to $k$-time-step interventions by quantum operations (instruments), such as quantum processes and measurements, can be modeled as shown in Figure 2. It utilizes a principal quantum system and an environment to represent the entire SE dimensions. At time step $t$, the experimenter applies an instrument $\mathcal{A}_{x_t}^{(t)}$ from the $t$-available set $\mathbb{J}^{(t)} := \left\{\mathcal{A}_{x_t}^{(t)}\right\}$ to the system. The instrument transforms the system state and outputs a value $x_t$ as the observation of the experimenter. Without loss of generality, the output $x_t$ is treated as the index of the instrument applied at time step $t$. Notably, there exists an SE unitary between adjacent time steps to represent the non-Markovian SE evolution. Therefore, the probability of obtaining a sequence of output value $\boldsymbol{x} := [x_0, x_1, ..., x_{k-1}]$ is represented by

$$p_{\boldsymbol{x}} = \text{Tr}\left\{\mathcal{A}_{x_{k-1}}^{(k-1)}\mathcal{U}_{k-2:k-1}\dots\mathcal{A}_{x_1}^{(1)}\mathcal{U}_{0:1}\mathcal{A}_{x_0}^{(0)}\left(\rho_{SE}^{(0)}\right)\right\} \tag{38}$$

$$= \langle\langle 0_{SE}|A_{x_{k-1}}^{(k-1)}\prod_{t=0}^{k-2} U_{t:t+1}A_{x_t}^{(t)}|\rho_{SE}^{(0)}\rangle\rangle, \tag{39}$$

where we simplify the $\mathcal{A}_{x_t}^{(t)}\otimes\mathcal{I}$ and $A_{x_t}^{(t)}\otimes I$ into $\mathcal{A}_{x_t}^{(t)}$ and $A_{x_t}^{(t)}$, respectively, without confusing them.
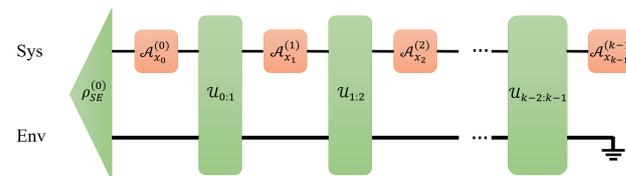


**Figure 2.** Model of operational open quantum process. Components marked by green blocks are inaccessible to the experimenter, while instruments marked by red blocks are the only accessible components.

Non-negligible non-Markovian SE correlations of open quantum systems theoretically impact the tomography techniques designed under Markovian settings. A simple example to demonstrate the impact is the theoretical violation of CP constraints given in Example 1 in [33,37] (c.f. Example A1 in Appendix A). The requirement of characterization of non-Markovian quantum correlations motivates non-Markovian quantum tomography techniques, such as process tensor tomography (PTT) and instrument set tomography (IST). Non-Markovian quantum tomography also exploits the three-step skeleton as shown in Section 3, where each experiment in the first step of *Experiment Preparation* consists of instruments.

The target in the non-Markovian tomography should be further exclaimed. Based on the criterion of accessibility, the components are classified into accessible and inaccessible

parts. Specifically, the only accessible components to the experimenter are the instruments acting on the principal quantum system, while the initial SE state and SE unitaries remain inaccessible. Quantum stochastic process theory [42,45] utilizes the process tensor to represent all inaccessible parts as a map from the product of $k$ spaces of CP and trace non-increasing (TNI) bounded operators on quantum states to the space of output quantum states. In this review, we principally adopt the definitions of process tensor $\mathcal{T}_{0:k}$ that

$$p_x = \mathcal{T}_{0:k}(\mathbf{A}_x) = \mathcal{T}_{0:k}\left(\mathcal{A}_{x_0}^{(0)}, \ldots, \mathcal{A}_{x_{k-1}}^{(k-1)}\right), \tag{40}$$

with physical properties:

(P1) *Linearity.* $\mathcal{T}_{0:k}(a\mathbf{A} + b\mathbf{B}) = \mathcal{T}_{0:k}(a\mathbf{A}) + \mathcal{T}_{0:k}(b\mathbf{B})$, for any $a.b \in \mathbb{R}$.

(P2) *Complete positivity.* $\mathcal{T}_{0:k} \otimes \mathcal{I}_{anc}(\mathbf{A}_{S,anc}) \geq 0$, where $\mathcal{I}_{anc}$ is the identity process on the ancilla, for any instruments $\mathbf{A}_{S,anc}$ act on the system and the ancilla.

(P3) *Containment.* $\mathcal{T}_{i:j}$ is contained in $\mathcal{T}_{l:k}$, where $l \leq i \leq j \leq k$.

Recently, the most acceptable mathematical representation of process tensor is the CJI matrix

$$Y_{\mathcal{T}_{0:k}} = \text{Tr}_E[\xi_{\mathcal{U}_{k-2:k-1}} \star \cdots \star \xi_{\mathcal{U}_{0:1}} \star \rho_{SE}^{(0)}], \tag{41}$$

where $\star$ represents the link product defined in [37], $\xi_{\mathcal{U}}$ is the Choi state of SE unitary $\mathcal{U}$. This formulation indicates that the SE correlations of an open quantum system with $k$ time step instrument interventions can be represented by a Choi state consisting of $2k + 1$ subsystems labeled by $\{i_{k-1}, o_{k-2} \ldots, i_1, o_0\}$, respectively. In this way, the output state can be determined by

$$p_x = \text{Tr}\left[\Xi_x^T Y_{\mathcal{T}_{0:k}}\right], \tag{42}$$

where $\Xi_x = \xi_{x_{k-1}}^{(k-1)} \otimes \cdots \otimes \xi_{x_0}^{(0)}$, and $\xi_{x_t}^{(t)}$ represent the Choi state of $\mathcal{A}_{x_t}^{(t)}$. It is easy to verify the linearity and complete positivity of $Y_{\mathcal{T}_{0:k}}$. Furthermore, the process tensor of a sub-time span can be derived as

$$\begin{aligned}
&Y_{\mathcal{T}_{i:j}}(\Xi_{x_l,\ldots,x_{i-1}}) \\
&= \text{Tr}_l\left[\left(\Xi_{x_{k-1},\ldots,x_j} \otimes I_l \otimes \Xi_{x_l,\ldots,x_{i-1}}\right)^T Y_{\mathcal{T}_{l:k}}\right],
\end{aligned} \tag{43}$$

for any $\Xi_{x_{k-1},\ldots,x_j}$ because of the causality, where $l = \overline{i_{j-1}o_{j-2} \ldots i_{i+1}o_i}$, and partial trace with label $\bar{\bullet}$ means partial trace other dimensions except for $\bullet$.

In this section, based on the linear inversion, we will introduce the basic methods of PTT and IST, i.e., LPTT and LIST, respectively.

### *4.1. Basic PTT*

Linear inverse process tensor tomography (LPTT) provides a common and quick method for PTT [36,42]. The process tensor can be treated as a map from the space of $k$-time-step instruments to the probability. The tomography of process tensor requires the informationally complete (IC) available sets at all time steps, which means that the instruments in each available set span the space of the bounded linear operator on the quantum state. Furthermore, we assume that instruments are perfectly implemented as the knowledge to the experimenter.

The experimenter prepares a set of experiments labeled by $x$ that consist of all combinations of instruments. By performing the experiments, the experimenter collects estimated probabilities $\hat{p}_x$.

Let $\mathbb{D}^{(t)} := \{\chi_{x_t}^{(t)}\}$ be the dual set of $t$-available set $\mathbb{J}^{(t)} := \{\xi_{x_t}^{(t)}\}$ that $\text{Tr}(\chi_i^{(t)T}\xi_j^{(t)}) = \delta_{ij}$. The process tensor can be reconstructed by

$$\hat{Y}_{\mathcal{T}_{0:k}} = \sum_{\boldsymbol{x}} \hat{p}_{\boldsymbol{x}} \left( \chi_{x_{k-1}}^{(k-1)} \otimes \cdots \otimes \chi_{x_0}^{(0)} \right). \tag{44}$$

Given a sequence of output $\boldsymbol{y}$, it can be verified that

$$\text{Tr}\left[ \Xi_{\boldsymbol{y}}^T \hat{Y}_{\mathcal{T}_{0:k}} \right] \tag{45}$$

$$= \sum_{\boldsymbol{x}} p_{\boldsymbol{x}} \prod_t \text{Tr}[\chi_{x_t}^{(t)T}\xi_{y_t}^{(t)}] \tag{46}$$

$$= \sum_{\boldsymbol{x}} p_{\boldsymbol{x}} \prod_t \delta_{x_t y_t} \tag{47}$$

$$= p_{\boldsymbol{y}}. \tag{48}$$

Since the IC available set at each time step consists of $d^2 \times d^2$ instruments, LPTT requires $d^{4k}$ experiments to characterize the entire process tensor. However, the linear inverse procedures are independently applied on each time step to determine the dual set of the available set. This indicates $k$ independent linear inverse procedures on the $d^4 \times d^4$ complex matrix. Nevertheless, it still requires the reconstruction of the $d^{4k}$ matrix of the process tensor.

### 4.2. Basic IST

The emergence of instrument set tomography (IST) results from non-avoidable imperfect implementations of instruments while performing PTT [38]. The differences between the knowledge and the practical implementations of instruments may lead to the systematic inconsistency of tomographic results. A simple example to show the inconsistency is that PTT generates two different process tensors utilizing two sets of imperfectly implemented IC instruments, respectively. Moreover, it is not always reasonable to require the quantum device to be characterized as this can formulate an IC available set at each time step. Recent superconductive quantum devices can hardly provide quantum gates that span the space of unitary when a time step is specified to be a time slot for applying a quantum gate [33,38]. Tomographic results should be responsible for all available quantum instruments that the quantum device can provide, which results in the requirement of the self-consistency of the IST.

Equation (38) indicates that, given the initial SE state, SE unitaries, and instruments, the probability can be determined. By exploiting the process tensor to represent all inaccessible components, the probability can be determined by specifying the instruments and the process tensor. Therefore, the full and reduced instrument set can be defined as

$$\mathfrak{I}_{\text{full}} := \left\{ \mathbb{J}, \mathbb{U}, \rho_{SE}^{(0)} \right\}, \tag{49}$$

$$\mathfrak{I}_{\text{reduced}} := \{ \mathbb{J}, \mathcal{T} \}, \tag{50}$$

respectively, where $\mathbb{J} := \left\{ \mathcal{J}^{(t)} \right\}_{t=0}^{k-1}$ represents available sets, $\mathbb{U} := \{ \mathcal{U}_{t:t+1} \}_{t=0}^{k-2}$ is the set of SE unitaries, and $\rho_{SE}^{(0)}$ represents the initial SE state.

LIST consists of two sub-procedures. First, LIST fixes the linear relationship of available sets by measurement probabilities and knowledge of instruments. For time step $t$, the probability described in (39) can be reformed as

$$p_{\boldsymbol{x}^+\boldsymbol{x}^-}(x_t) = \text{Tr}\left[ B_{\boldsymbol{x}^+\boldsymbol{x}^-} A_{x_t}^{(t)} \right], \tag{51}$$

where $x^+$ and $x^-$ represent the output values before and after time step $t$ in a $k$-time step experiment, respectively. This indicates the decomposition of $A^{(t)}_{x_t}$ on the inaccessible non-orthogonal basis

$$\mathbb{B}^{(t)} = \{B_{x^+x^-}\}. \tag{52}$$

By adjusting $x^+$ and $x^-$, the LIST tries to make the basis $\mathbb{B}^{(t)}$ complete. Utilizing a unique integer $\alpha$ to denote $x^+x^-$ as the index of the basis matrix, the LIST connects the measured probabilities and the basis as

$$\boldsymbol{p}^{(t)}_{x_t} = \begin{bmatrix} (\boldsymbol{b}^{(t)}_0)^\dagger \\ (\boldsymbol{b}^{(t)}_1)^\dagger \\ \vdots \\ (\boldsymbol{b}^{(t)}_{d^4-1})^\dagger \end{bmatrix} \boldsymbol{a}^{(t)}_{x_t} = B^{(t)} \boldsymbol{a}^{(t)}_{x_t}, \tag{53}$$

where $\boldsymbol{a}^{(t)}_{x_t}$ and $\boldsymbol{b}^{(t)}_\alpha$ represent the vectorization of the $A^{(t)}_{x_t}$ and $B^{(t)}_\alpha$, respectively. If $B^{(t)}$ is invertible, we can obtain the instruments

$$\Xi^{(t)} = \left(B^{(t)}\right)^{-1} \Gamma^{(t)}, \tag{54}$$

where $\Xi^{(t)} = [\boldsymbol{a}^{(t)}_0, \boldsymbol{a}^{(t)}_1, \ldots, \boldsymbol{a}^{(t)}_{m_t-1}]$ and $\Gamma^{(t)} = [\boldsymbol{p}^{(t)}_0, \boldsymbol{p}^{(t)}_1, \ldots, \boldsymbol{p}^{(t)}_{m_t-1}]$.

Assuming that the quantum instruments are well implemented and approximately close to the ideal instruments, the inaccessible basis matrix can be optimized by

$$B^{(t)} = \arg\min_X \sum_t \left\| X\Gamma^{(t)} - \Xi^{(t)}_{\text{knowledge}} \right\|_{F'} \tag{55}$$

where $\Xi^{(t)}_{\text{knowledge}}$ represents the knowledge of instruments to the experimenter. Then, the PTMs of instruments are recovered by the devectorization of determined $\boldsymbol{a}^{(t)}_{x_t}$ in $\Xi^{(t)}$.

Then, the LIST performs the process tensor reconstruction to recover the process tensor. Utilized instruments are the maximum linear independent set of instruments in the corresponding available sets. Details of process tensor reconstruction can be referred to in Section 4.1.

Note that the requirement of basis matrix optimization implies gauge freedom, which is also present in the GST. The optimal basis matrix is non-unique and results in the non-uniqueness of tomographic results. This implies that correlations introduced before the time step cannot be distinguished by correlations introduced after.

## 5. Maximum Likelihood Estimation-Based Methods

Maximum likelihood estimation (MLE)-based methods [5,32,37,38,46,47] grew out of the drawbacks of basic linear inverse methods. It should be highlighted that the precision of linear inverse methods highly relies on the accuracy of estimated probability generated by measurement sampling, which is of the order $O(1/n_s)$, where $n_s$ is the number of samples. The probability estimation of finite measurement sampling introduces the sampling error, leading to a biased estimated probability. Subsequently, the result of the tomography target may not be physical, such as the CP violation of the estimated quantum state in QST. Experimenters may be interested in utilizing more experiment data to enhance the performance of the tomography. However, linear inverse methods lack the capability of extracting additional information from overcomplete data.

Addressing these issues, statistical methods are proposed based on the MLE, which enables the exploitation of overcomplete data to reduce the error of estimation. These methods compose the measurement data into a likelihood function $\mathcal{L}(\hat{T})$ with respect to the estimated target $\hat{T}$ which requires maximization. The parameterization of the

tomographic target also varies from the method. Typically, quantum states, processes, and measurements can be directly parameterized by $d^2$-dimensional real column vectors, $d^2 \times d^2$-dimensional real matrices, and $d^2$-dimensional real row vectors, respectively, as their PTM representations.

The parameterization of SE correlations in non-Markovian situations is more complicated. While representing SE correlations by the process tensor, it can be modeled by the Cholesky decomposition of the Choi state which naturally satisfies the CP constraints. The PTM representation of the process tensor is also available as described in Section 4. Furthermore, the SE correlations can also be represented by the initial SE state and SE unitaries, where the SE state can be represented by a $d^4$-dimensional real column vector and each SE unitary can be modeled by a $(d^4 - 1)$-dimensional real vector $\boldsymbol{\alpha}$ whose $i$-th element represents the rotation angle on the $(i + 1)$-th Pauli matrix, i.e., the estimated SE unitary is

$$\hat{V}(\boldsymbol{\alpha}) := \mathbf{PTM}(\exp(\iota \sum_i \alpha_i P_{i+1})), \tag{56}$$

where $\iota^2 = -1$.

Although several kinds of likelihood functions are proposed, we apply the most common likelihood as described in Appendix B, where maximizing $\mathcal{L}(\hat{T})$ is equivalent to minimizing $l(\hat{T})$ with respect to the desired estimator $\hat{p}$ as described in their basic framework. Remarkably, constraints are introduced in MLE methods to guarantee that the result is physical. The details of constraints for tomography components are listed below:

**State:** A quantum state $\hat{\rho}$ is constrained to be completely positive, meaning its density matrix must be positive semi-definite with a unit trace,

$$\hat{\rho} \succeq 0, \tag{57}$$

$$\mathrm{Tr}\, \hat{\rho} = 1. \tag{58}$$

This guarantees that the probabilities of the system are positive with summation 1. Note that TP constraints can be efficiently represented by

$$|\hat{\rho}\rangle\rangle = 1. \tag{59}$$

**Process:** A quantum process is constrained to be CPTP. The CP constraints require the Choi state of $\hat{\mathcal{G}}$ to be completely positive,

$$\hat{\xi}_{\mathcal{G}} \succeq 0, \tag{60}$$

which indicates that the process remains the CP property of the quantum state. Meanwhile, entries of the first row of its PTM are limited as

$$\hat{G}_{0j} = \delta_{0j} \tag{61}$$

to guarantee the TP property, which indicates that the summation of probabilities of the quantum state is 1.

**Measurement:** A quantum measurement $\hat{\mathcal{M}}$ and its complementary are constrained by a complete positive that

$$\hat{M} = \sum_i \langle\langle \hat{M}|_i P_i \succeq 0, \tag{62}$$

$$I - \hat{M} \succeq 0, \tag{63}$$

representing the positive outcome of probabilities.

**Process tensor:** The process tensor $\hat{\mathcal{T}}$ should be constrained by CP and causality. The CP constraints require the Choi state of the process tensor to be positive semi-definite,

$$\hat{Y}_{\mathcal{T}} \succeq 0. \tag{64}$$

Causality constraints result from the fact that future events do not influence past statistics, as a generalization of trace preservation, which in the Choi state representation is

$$\text{Tr}_{o_{k-1}}\left[\hat{Y}_{\mathcal{T}_{0:k}}\right] = I_{i_{k-1}} \otimes \hat{Y}_{\mathcal{T}_{0:k-1}}. \tag{65}$$

By superoperators of PTM representation, the causality constraints can be conveniently represented by

$$\langle\langle \hat{Y}_{\mathcal{T}_{0:k}} | P_{\text{ban}} \rangle\rangle = 0, \tag{66}$$

$$P_{\text{ban}} = I^{\otimes 2t+1} \otimes \left(\tilde{Q}_{2t+2} \otimes Q_{2t+3} \otimes \cdots \otimes Q_{2k-1}\right), \forall t, \tag{67}$$

$$\tilde{Q} \in \left\{P_1, \ldots, P_{d^2-1}\right\}, \tag{68}$$

$$Q \in \left\{P_0, P_1, \ldots, P_{d^2-1}\right\}. \tag{69}$$

Then, an optimization algorithm is performed to optimize the optimization problem

$$\min_{\hat{T}} \; l(\hat{T}) \tag{70}$$

$$s.t. \;\; \text{Constraints of } \hat{T}. \tag{71}$$

The workflow of MLE methods is demonstrated in Figure 3.
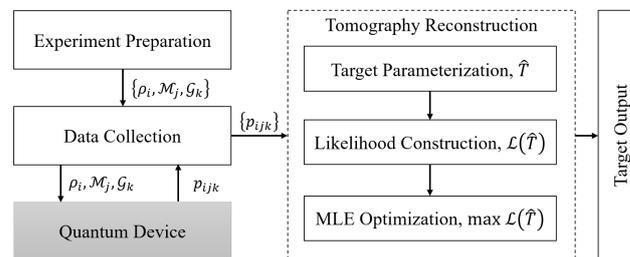


**Figure 3.** Workflow of MLE-based quantum tomography. Components are specified in the Markovian settings for facilitating understanding, although the workflow is general for both Markovian and non-Markovian situations.

MLE methods fill the blank of constrained statistical methods within the standard quantum tomography toolkits and obtained wild acceptance from researchers. They provide physical implementable results by introducing constraints. The capability of utilizing overcomplete data promotes the accuracy of the tomography by extracting additional information compared with basic linear inverse methods. However, MLE methods are quite computationally resource-consuming. In the classic memory aspect, all MLE methods require saving exponentially large data with respect to the system dimension $d$ to perform the optimization algorithm. In the time aspect, MLE methods require more experiments than the linear inverse methods which is exponentially large to $d$. Moreover, solving constrained optimization problems with an exponentially large number of parameters with respect to $d$ is extremely time-consuming. Therefore, the recent applications of MLE methods are limited to small-scale quantum systems.

## 6. Further Developments

The advancement of quantum tomography performance necessitates further development built upon standard methods, including linear inverse and maximum likelihood estimation (MLE) techniques. First, the effectiveness of frequency-based standard methods

requires further consideration. The estimated probabilities are biased with finite sampling, which results in the possibility of the violation of physical implementability as it may occur in linear inverse methods. It seems that this problem has been solved by MLE methods. However, it can be verified that the MLE outputs the same result as linear inverse methods when using the same set of measurement data without constraints. The introduction of constraints forces the output to satisfy the physical requirement, which means that the result lies on the boundary of feasible solutions, as shown in Figure 4. This indicates that no reasonable error bar can be applied in this estimation.
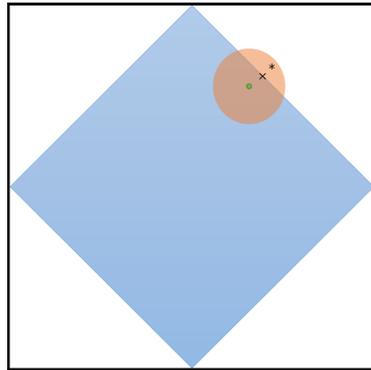


**Figure 4.** Utilizing the black box to represent the whole space of the tomographic target, the space of physical implementable results is represented by the blue region. The green dot is the ideal tomographic target to be characterized. Due to the probabilistic sampling, the sampled target may lie in the red region. Assuming the sampled target to be ∗ out of physical implementable space, the MLE result lies on the boundary of the implementable region, which can be represented by ×.

Bayesian tomographic methods are proposed for this systematic problem [14,26,48,49]. These methods propose meaningful formations for the estimation error of the target. The tomographic targets are assumed to obey a probability distribution. After optimization, the final target is set as the expectation of the distribution, which naturally has a reasonable error bar corresponding to the variance. Nevertheless, Bayesian methods are quite time-consuming due to the construction of distribution. Efficient Bayesian estimation for quantum tomography needs further research.

Then, the efficiency indicated by the accuracy gained from the unit number of experiments and the unit time of the post-process algorithm should be further improved. As previously introduced, the linear inverse provides relatively quick methods for tomography, but the accuracy is bounded by the reciprocal of the number of samples without the capability of utilizing additional data to improve the accuracy. Moreover, sampling errors may cause the physical violations of results. The MLE methods have the capability to utilize additional data with physical constraints. However, performing optimization algorithms to solve constrained optimization problems with an exponentially large number of parameters is time-consuming, which limits these methods to small-scale systems. Recently, researchers have proposed several methods to enhance the efficiency of quantum tomography.

Following the trend of the development of machine learning, a quantum tomography scheme utilizing the neural network has been proposed [50–52]. Typically, training and testing datasets consist of measurement results and prior tomographic results. Then, the neural network explicitly learns the map from measurement data to the tomographic target, and implicitly learns the knowledge of non-target components. To a certain extent, the absence of knowledge of non-target components mitigates the impact of imperfect implementations. Remarkably, when the training has been finished, neural-network-based methods have obtained higher accuracy with the same number of probability data inputs compared with the linear inverse method. It should be noted that neural-network-based methods may not generate physically implementable results due to the absence of constraints. However,

physical constraints can be introduced into the cost function while performing neural network training, which results in time-consuming neural network training.

Another approach to enhance efficiency involves leveraging the prior symmetrical structure of the tomographic target. In other words, incorporating certain assumptions about the tomographic target can prove beneficial. One well-known assumption is employed in compressed tomography [13,53,54], which posits that the tomographic target exhibits a low-rank characteristic. This methodology has evolved from matrix completion techniques, where the complete target matrix is reconstructed from incomplete elements using random sampling and a nuclear norm minimization algorithm. For instance, consider an unknown quantum state $\rho$ with $\text{rank}(\rho) \leq r$. By employing this approach, an estimate $\hat{\rho}$ can be derived with only $O(rd \log^2 d)$ measurement settings. Additionally, other prior symmetry structural assumptions, such as matrix product states [55], permutationally invariant states [56], and similar concepts, also significantly contribute to enhancing efficiency.

**Author Contributions:** Conceptualization, T.L. and Z.L.; writing—original draft preparation, T.L., Z.L. and C.Z.; writing—review and editing, T.L., Z.L., C.Z., X.K. and X.Y.; supervision, T.L. and Z.Z. All authors have read and agreed to the published version of the manuscript.

**Data Availability Statement:** There are no data associated with this work.

**Conflicts of Interest:** Tian Luan and Xueheng Kuang were employed by the research institute, Yangtze Delta Region Industrial Innovation Center of Quantum Science and Technology, which is not a company. All of authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Appendix A. Examples

**Example A1 (CP violation).** *Let the QPT target with SE correlations be*

$$G_{SE} = \prod_{P=X,Y,Z} \{\cos(wt)I_S \otimes I_E - \sin(wt)P_S \otimes P_E\}, \tag{A1}$$

*which is a unitary. Let the initial SE state be*

$$\rho_{SE}^{(0)} = \frac{1}{4}[I_S \otimes I_E + \vec{a}\,\vec{\sigma}_S \otimes I_E + gY_S \otimes Z_E], \tag{A2}$$

*where $\vec{a}\,\vec{\sigma}_S = [a_x, a_y, a_z][X_S, Y_S, Z_S]^{\dagger} = a_x X_S + a_y Y_S + a_z$. Since the experimenter can only intervene in the system state, we prepare states by setting $\vec{a}$ to be $[\pm 1, 0, 0]^T$, $[0, 1, 0]^T$, and $[0, 0, 1]^T$. Hence, the Choi state of tomographic result $\hat{G}$ is*

$$\xi_{\mathcal{G}} = \frac{1}{2}\begin{bmatrix} 1+c^2 & 0 & -gcs & 2c^2 \\ 0 & 1-c^2 & 0 & -gcs \\ -gcs & 0 & 1-c^2 & 0 \\ 2c^2 & -gcs & 0 & 1+c^2 \end{bmatrix}, \tag{A3}$$

*where $c = \cos(2wt)$ and $s = \sin(2wt)$. There exist eigenvalues $\frac{1}{2}[1 - c^2 \pm 2cs]$ that may be negative and cause the theoretical CP violation.*

**Appendix B. Likelihood Function**

In this review, we utilize the likelihood function as described below. Let the probability be measured by repeating the experiment $n_s$ times and recording $n_i$, which denotes how many times the desired output $i$ occurs. Therefore, we use the general likelihood function

$$\mathcal{L}(\hat{T}) = \prod_i (\hat{p}_i)^{n_i} (1 - \hat{p}_i)^{n_s - n_i}, \tag{A4}$$

where $\hat{p}_i$ is the probability estimator modeled by the parameters of tomography target $\hat{T}$.

By exploiting the central limit theorem, each term of the likelihood can be rewritten as a normal distribution,

$$\mathcal{L}(\hat{T}) = \prod_i \exp\left[ -\frac{(\tilde{p}_i - \hat{p}_i)^2}{\sigma_i^2} \right], \tag{A5}$$

where $\tilde{p}_i = n_i / n_s$ represents the measured probability, $\sigma_i^2 = \tilde{p}_i(1 - \tilde{p}_i)/n_s$ is the sampling variance in the measurement $m_i$. Exploiting the monotonic logarithm function, maximizing $\mathcal{L}$ is equivalent to minimizing the weighted mean square error

$$l(\hat{T}) = -\log(\mathcal{L}(\hat{T})) = \sum_i \frac{(\tilde{p}_i - \hat{p}_x)^2}{\sigma_i^2}. \tag{A6}$$

## References

1. Kliesch, M.; Roth, I. Theory of Quantum System Certification. *PRX Quantum* **2021**, *2*, 010201. [CrossRef]
2. Eisert, J.; Hangleiter, D.; Walk, N.; Roth, I.; Markham, D.; Parekh, R.; Chabaud, U.; Kashefi, E. Quantum Certification and Benchmarking. *Nat. Rev. Phys.* **2020**, *2*, 382–390. [CrossRef]
3. D'Ariano, G.M.; Paris, M.G.A.; Sacchi, M.F. Quantum Tomography. *Adv. Imaging Electron Phys.* **2003**, *128*, 205–308. [CrossRef]
4. Häffner, H.; Hänsel, W.; Roos, C.F.; Benhelm, J.; Chek-al-kar, D.; Chwalla, M.; Körber, T.; Rapol, U.D.; Riebe, M.; Schmidt, P.O.; et al. Scalable Multiparticle Entanglement of Trapped Ions. *Nature* **2005**, *438*, 643–646. [CrossRef]
5. Greenbaum, D. Introduction to Quantum Gate Set Tomography. *arXiv* **2015**, arXiv:1509.02921.
6. O'Donnell, R.; Wright, J. Efficient Quantum Tomography. In Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, New York, NY, USA, 2016; STOC'16; pp. 899–912. [CrossRef]
7. Granade, C.; Ferrie, C.; Flammia, S.T. Practical Adaptive Quantum Tomography. *New J. Phys.* **2017**, *19*, 113017. [CrossRef]
8. Cai, Z.; Babbush, R.; Benjamin, S.C.; Endo, S.; Huggins, W.J.; Li, Y.; McClean, J.R.; O'Brien, T.E. Quantum Error Mitigation. *Rev. Mod. Phys.* **2023**, *95*, 045005. [CrossRef]
9. Lidar, D.A.; Brun, T.A. *Quantum Error Correction*; Cambridge University Press: Cambridge, UK, 2013.
10. Roffe, J. Quantum Error Correction: An Introductory Guide. *Contemp. Phys.* **2019**, *60*, 226–245. [CrossRef]
11. Cory, D.G.; Price, M.D.; Maas, W.; Knill, E.; Laflamme, R.; Zurek, W.H.; Havel, T.F.; Somaroo, S.S. Experimental Quantum Error Correction. *Phys. Rev. Lett.* **1998**, *81*, 2152–2155. [CrossRef]
12. Vogel, K.; Risken, H. Determination of Quasiprobability Distributions in Terms of Probability Distributions for the Rotated Quadrature Phase. *Phys. Rev. A* **1989**, *40*, 2847–2849. [CrossRef] [PubMed]
13. Gross, D.; Liu, Y.K.; Flammia, S.T.; Becker, S.; Eisert, J. Quantum State Tomography via Compressed Sensing. *Phys. Rev. Lett.* **2010**, *105*, 150401. [CrossRef]
14. Blume-Kohout, R. Optimal, Reliable Estimation of Quantum States. *New J. Phys.* **2010**, *12*, 043034. [CrossRef]
15. Christandl, M.; Renner, R. Reliable Quantum State Tomography. *Phys. Rev. Lett.* **2012**, *109*, 120403. [CrossRef] [PubMed]
16. Blume-Kohout, R. Hedged Maximum Likelihood Quantum State Estimation. *Phys. Rev. Lett.* **2010**, *105*, 200504. [CrossRef] [PubMed]
17. Gale, W.; Guth, E.; Trammell, G.T. Determination of the Quantum State by Measurements. *Phys. Rev.* **1968**, *165*, 1434–1436. [CrossRef]
18. Banaszek, K.; Cramer, M.; Gross, D. Focus on Quantum Tomography. *New J. Phys.* **2013**, *15*, 125020. [CrossRef]
19. Haah, J.; Harrow, A.W.; Ji, Z.; Wu, X.; Yu, N. Sample-Optimal Tomography of Quantum States. *IEEE Trans. Inf. Theory* **2017**, *63*, 5628–5641. [CrossRef]
20. James, D.F.V.; Kwiat, P.G.; Munro, W.J.; White, A.G. Measurement of Qubits. *Phys. Rev. A* **2001**, *64*, 052312. [CrossRef]
21. Smolin, J.A.; Gambetta, J.M.; Smith, G. Efficient Method for Computing the Maximum-Likelihood Quantum State from Measurements with Additive Gaussian Noise. *Phys. Rev. Lett.* **2012**, *108*, 070502. [CrossRef]
22. Webb, A.E.; Webster, S.C.; Collingbourne, S.; Bretaud, D.; Lawrence, A.M.; Weidt, S.; Mintert, F.; Hensinger, W.K. Resilient Entangling Gates for Trapped Ions. *Phys. Rev. Lett.* **2018**, *121*, 180501. [CrossRef]

23. Mohseni, M.; Rezakhani, A.T.; Lidar, D.A. Quantum-Process Tomography: Resource Analysis of Different Strategies. *Phys. Rev. A* **2008**, *77*, 032322. [CrossRef]
24. Merkel, S.T.; Gambetta, J.M.; Smolin, J.A.; Poletto, S.; Córcoles, A.D.; Johnson, B.R.; Ryan, C.A.; Steffen, M. Self-Consistent Quantum Process Tomography. *Phys. Rev. A* **2013**, *87*, 062119. [CrossRef]
25. Riebe, M.; Kim, K.; Schindler, P.; Monz, T.; Schmidt, P.O.; Körber, T.K.; Hänsel, W.; Häffner, H.; Roos, C.F.; Blatt, R. Process Tomography of Ion Trap Quantum Gates. *Phys. Rev. Lett.* **2006**, *97*, 220407. [CrossRef] [PubMed]
26. Schultz, K. Exponential Families for Bayesian Quantum Process Tomography. *Phys. Rev. A* **2019**, *100*, 062316. [CrossRef]
27. Chuang, I.L.; Nielsen, M.A. Prescription for Experimental Determination of the Dynamics of a Quantum Black Box. *J. Mod. Opt.* **1997**, *44*, 2455–2467. [CrossRef]
28. Childs, A.M.; Chuang, I.L.; Leung, D.W. Realization of Quantum Process Tomography in NMR. *Phys. Rev. A* **2001**, *64*, 012314. [CrossRef]
29. Bialczak, R.C.; Ansmann, M.; Hofheinz, M.; Lucero, E.; Neeley, M.; O'Connell, A.D.; Sank, D.; Wang, H.; Wenner, J.; Steffen, M.; et al. Quantum Process Tomography of a Universal Entangling Gate Implemented with Josephson Phase Qubits. *Nat. Phys.* **2010**, *6*, 409–413. [CrossRef]
30. Bendersky, A.; Pastawski, F.; Paz, J.P. Selective and Efficient Estimation of Parameters for Quantum Process Tomography. *Phys. Rev. Lett.* **2008**, *100*, 190403. [CrossRef]
31. Altepeter, J.B.; Branning, D.; Jeffrey, E.; Wei, T.C.; Kwiat, P.G.; Thew, R.T.; O'Brien, J.L.; Nielsen, M.A.; White, A.G. Ancilla-Assisted Quantum Process Tomography. *Phys. Rev. Lett.* **2003**, *90*, 193601. [CrossRef]
32. Nielsen, E.; Gamble, J.K.; Rudinger, K.; Scholten, T.; Young, K.; Blume-Kohout, R. Gate Set Tomography. *Quantum* **2021**, *5*, 557. [CrossRef]
33. Rudinger, K.; Ribeill, G.J.; Govia, L.C.; Ware, M.; Nielsen, E.; Young, K.; Ohki, T.A.; Blume-Kohout, R.; Proctor, T. Characterizing Midcircuit Measurements on a Superconducting Qubit Using Gate Set Tomography. *Phys. Rev. Appl.* **2022**, *17*, 014014. [CrossRef]
34. Gu, Y.; Mishra, R.; Englert, B.G.; Ng, H.K. Randomized Linear Gate-Set Tomography. *PRX Quantum* **2021**, *2*, 030328. [CrossRef]
35. Blume-Kohout, R.; Gamble, J.K.; Nielsen, E.; Rudinger, K.; Mizrahi, J.; Fortier, K.; Maunz, P. Demonstration of Qubit Operations below a Rigorous Fault Tolerance Threshold with Gate Set Tomography. *Nat. Commun.* **2017**, *8*, 14485. [CrossRef] [PubMed]
36. Pollock, F.A.; Rodríguez-Rosario, C.; Frauenheim, T.; Paternostro, M.; Modi, K. Non-Markovian Quantum Processes: Complete Framework and Efficient Characterization. *Phys. Rev. A* **2018**, *97*, 012127. [CrossRef]
37. White, G.; Pollock, F.; Hollenberg, L.; Modi, K.; Hill, C. Non-Markovian Quantum Process Tomography. *PRX Quantum* **2022**, *3*, 020344. [CrossRef]
38. Li, Z.T.; Zheng, C.C.; Meng, F.X.; Zeng, H.; Luan, T.; Zhang, Z.C.; Yu, X.T. Non-Markovian Quantum Gate Set Tomography. *arXiv* **2023**, arXiv:2307.14696. https://doi.org/10.48550/arXiv.2307.14696.
39. Carolan, J.; Meinecke, J.D.A.; Shadbolt, P.J.; Russell, N.J.; Ismail, N.; Wörhoff, K.; Rudolph, T.; Thompson, M.G.; O'Brien, J.L.; Matthews, J.C.F.; et al. On the Experimental Verification of Quantum Complexity in Linear Optics. *Nat. Photonics* **2014**, *8*, 621–626. [CrossRef]
40. Lvovsky, A.I.; Raymer, M.G. Continuous-Variable Optical Quantum-State Tomography. *Rev. Mod. Phys.* **2009**, *81*, 299–332. [CrossRef]
41. Bellini, M.; Coelho, A.S.; Filippov, S.N.; Man'ko, V.I.; Zavatta, A. Towards Higher Precision and Operational Use of Optical Homodyne Tomograms. *Phys. Rev. A* **2012**, *85*, 052129. [CrossRef]
42. Milz, S.; Modi, K. Quantum Stochastic Processes and Quantum Non-Markovian Phenomena. *PRX Quantum* **2021**, *2*, 030201. [CrossRef]
43. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information: 10th Anniversary Edition*; Cambridge University Press: Cambridge, UK, 2010.
44. Jamiołkowski, A. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.* **1972**, *3*, 275–278. [CrossRef]
45. Davies, E.B. Quantum Stochastic Processes. *Commun. Math. Phys.* **1969**, *15*, 277–304. [CrossRef]
46. Hradil, Z. Quantum-State Estimation. *Phys. Rev. A* **1997**, *55*, R1561–R1564. [CrossRef]
47. Shang, J.; Zhang, Z.; Ng, H.K. Superfast Maximum-Likelihood Reconstruction for Quantum Tomography. *Phys. Rev. A* **2017**, *95*, 062336. [CrossRef]
48. Huszár, F.; Houlsby, N.M.T. Adaptive Bayesian Quantum Tomography. *Phys. Rev. A* **2012**, *85*, 052120. [CrossRef]
49. Mahler, D.H.; Rozema, L.A.; Darabi, A.; Ferrie, C.; Blume-Kohout, R.; Steinberg, A.M. Adaptive Quantum State Tomography Improves Accuracy Quadratically. *Phys. Rev. Lett.* **2013**, *111*, 183601. [CrossRef]
50. Neugebauer, M.; Fischer, L.; Jäger, A.; Czischek, S.; Jochim, S.; Weidemüller, M.; Gärttner, M. Neural-Network Quantum State Tomography in a Two-Qubit Experiment. *Phys. Rev. A* **2020**, *102*, 042604. [CrossRef]
51. Palmieri, A.M.; Kovlakov, E.; Bianchi, F.; Yudin, D.; Straupe, S.; Biamonte, J.D.; Kulik, S. Experimental Neural Network Enhanced Quantum Tomography. *NPJ Quantum Inf.* **2020**, *6*, 20. [CrossRef]
52. Torlai, G.; Mazzola, G.; Carrasquilla, J.; Troyer, M.; Melko, R.; Carleo, G. Neural-Network Quantum State Tomography. *Nat. Phys.* **2018**, *14*, 447–450. [CrossRef]
53. Kliesch, M.; Kueng, R.; Eisert, J.; Gross, D. Guaranteed Recovery of Quantum Processes from Few Measurements. *Quantum* **2019**, *3*, 171. [CrossRef]

54. Roth, I.; Kueng, R.; Kimmel, S.; Liu, Y.K.; Gross, D.; Eisert, J.; Kliesch, M. Recovering Quantum Gates from Few Average Gate Fidelities. *Phys. Rev. Lett.* **2018**, *121*, 170502. [CrossRef] [PubMed]
55. Cramer, M.; Plenio, M.B.; Flammia, S.T.; Somma, R.; Gross, D.; Bartlett, S.D.; Landon-Cardinal, O.; Poulin, D.; Liu, Y.K. Efficient Quantum State Tomography. *Nat. Commun.* **2010**, *1*, 149. [CrossRef] [PubMed]
56. Tóth, G.; Wieczorek, W.; Gross, D.; Krischek, R.; Schwemmer, C.; Weinfurter, H. Permutationally Invariant Quantum Tomography. *Phys. Rev. Lett.* **2010**, *105*, 250403. [CrossRef] [PubMed]