이학박사학위논문

# Adaptive Compressive Tomography with No *a priori* Information

## 사전 정보 없이 수행되는 적응 압축 토모그래피

2019년 8월

서울대학교 대학원
물리·천문학부
안대건

# Adaptive Compressive Tomography with No *a priori* Information

**Daekun Ahn**

Supervised by

Professor **Hyunseok Jeong**

A Dissertation

Submitted to the Faculty of

Seoul National University

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

August 2019

Department of Physics and Astronomy

The Graduate School

Seoul National University

# Abstract

# Adaptive Compressive Tomography with No *a priori* Information

Daekun Ahn
Department of Physics and Astronomy
The Graduate School
Seoul National University

In quantum information processing, quantum-state tomography is widely used for verification and preparation of quantum states by performing quantum measurement whose number of configurations grow polynomially in the dimension of the system. In this thesis, inspired by compressed sensing theory, we propose an adaptive compressive tomography scheme which recovers an arbitrary rank-deficient unknown true quantum state with a substantially reduced number of measurement configurations. This scheme requires no *a priori* information about the quantum state apart from the dimension of the system, which is needed to define our compressive measurements.

We introduce two major components of the adaptive scheme. One is an informational completeness certification that adopts semi-definite programming and the other is an adaptive strategy that takes entropy minimization pro-

cedure. We carry out both noiseless simulations and orbital angular momentum single photon experiments for 4-qubit systems and compare the compression efficiency of informationally complete data with random Pauli-projective measurements. From the comparison, we find that our adaptive scheme always outperforms the random measurement scheme. We also confirm by comparing between simulation and experimental results that the adaptive scheme is robust against noise arising from real experiments.

In addition to the aforementioned number of qubits, we further consider higher numbers of qubits for noiseless simulations. We numerically confirm that for all tested low rank quantum states, both entangled and product adaptive bases measurement schemes have higher compression efficiency than random bases measurement schemes, and also exhibit compressive behaviors comparable to recently proposed compressive element-probing measurements. Moreover, we supply numerical conjectures of asymptotic compression scaling behaviors that lose dependancy on dimension in the high dimension limit of the system. Finally, as a natural development, we establish a faster hybrid compressive tomography scheme that first performs random bases measurement, and later adaptive bases measurement.

**Keywords :** Compressive tomography, *a priori* information, informational completeness certification, adaptive strategy

**Student Number :** 2012-20368

# Contents

# List of Figures

# Chapter 1

# Introduction

Quantum-state tomography pertains to reconstruction of an unknown true quantum state from data acquired after performing a set of measurement outcomes. This task is important in quantum information processing for preparation and verification of quantum states. Given an unknown quantum state $\rho_t$ in the Hilbert space of dimension $d$, one typically performs a probability-operator measurement (POM) that fully characterizes $d^2 - 1$ free parameters of $\rho_t$ and by using the data acquired from the measurement, one gets a unique estimator $\widehat{\rho}$ that converges to the true state $\rho_t$ when the measurement is noiseless. However, the uniqueness of the estimator is guaranteed only when the size of the set of independent measurement is not smaller than $d^2$, which grows polynomially with $d$. Thus, in a real experiment for large $d$, it is inevitable to encounter technical limitations on the preparation of the measurement setup for a full reconstruction of the state. On the other hand, if one adhere to feasible measurement setup, estimators corresponding to an informationally incomplete dataset are not unique.

To overcome this resource-intensive practical problem, the theory of compressed sensing (CS) [7, 8, 9] adopted to quantum-state tomography, which provides a specific compressed sensing map from a quantum state to data, and

a state estimation procedure with respect to the data, both of which are based on the assumed *a priori* information that $\text{rank}\{\rho_t\} \leq r \ll d$, where, $r \ll d$ refers to the assumption that the true state $\rho_t$ is close to a pure state, which is often taken for granted when dealing with quantum information experiments targeting pure states. However, this assumed upper bound threshold of rank requires verification since the CS procedure is theoretically valid only when the *a priori* information is true.

Random Pauli measurement is a popular complete measurement choice for CS tomography [17], which is the set of randomly picked product Pauli observables. It was shown that $M = O(rd(\log d)^2)$ random Pauli observables yield perfect state recovery [17]. [14] generalized the measurement into the set of randomly picked product of projectors on Pauli bases, that is random Pauli-projective measurement and numerically confirmed that under the quantum positivity constraint, CS procedure shows compressive behavior on the number of measurement configurations needed for perfect state recovery.

Meanwhile, without adopting CS estimation procedure, *Flammia et al.* and *Goyeneche et al.* came up with measurement for rank-1 (pure) states [21, 22]. Upon generalizing their results, *Baldwin et al.* mathematically constructed two specific compressive measurements for arbitrary rank-$r$ states $\rho_t$ with the number of measurement configurations respectively $M = 2rd - r^2$ and $M = (4r+1)d$ [20]. These two measurements yield the unique rank-$r$ estimator with respect to the whole state space.

However, both measurement schemes require the knowledge of the rank

of $\rho_t$, which may differ by the experimental situation or perpective of the experimentalist. Thus, it requires justification of such *a priori* information of $\rho_t$ in advance.

In this thesis, we develop a new *adaptive compressive tomography* (ACT) scheme that explicitly recovers the unknown true state with no *a priori* information about the state.

In chapter 2, we briefly review the basic concepts of quantum-state tomography and introduce the notion of "informational completeness". Then we present more details about compressed sensing theory with the introduced feasible compressive measurement schemes.

In chapter 3, we explain the two main components that constitute ACT: self-consistent informational completeness certification (ICC) and adaptive strategy to specify the next optimal measurement *via* an entropy minimization procedure. For the verification of the scheme, we perform a single-photon experiment for the 4-qubit case ($d = 16$) with the orbital angular momentum (OAM) platform, which encodes information on the OAM degrees of freedom of single photons, and compare the experimental results with the simulation results. We further carry out simulations for different numbers of qubits, and compare the schemes with previous known compressive measurements. In particular, we introduce a "random-adaptive" hybrid scheme that combines the high speed of random measurement and the high compressive efficiency to get a unique estimator.

Finally in chapter 4, we conclude the thesis with remarks on the results.

This thesis is based on the following publications:

1. D. Ahn, Y.S. Teo, H. Jeong, F. Bouchard, F. Hufnagel, E. Karimi, D. Koutný, J. Řeháček, Z. Hradil, G. Leuchs, and L.L. Sánchez-Soto, "Adaptive Compressive Tomography with No *a priori* Information," Phys. Rev. Lett. **122**, 100404 (2019).

2. D. Ahn, Y.S. Teo, H. Jeong, D. Koutný, J. Řeháček, Z. Hradil, G. Leuchs, and L.L. Sánchez-Soto, "Adaptive compressive tomography: a numerical study," arXiv:1905.01488.

# Chapter 2

# Background

## 2.1 Concept of quantum-state tomography

Quantum-state tomography is a task of measurement and estimation of the unknown quantum state of a physical system. This requires quantum measurements and state estimation technique that reconstructs the state from the data obtained through the measurements. An unknown quantum state $\rho_t$ within Hilbert space $\mathscr{H}_d$ of dimension $d$ may be expressed as

$$\rho_t = \sum_{j,k} |j\rangle \, \rho_{t\,jk} \, \langle k| \,, \tag{2.1}$$

, where $(\rho_{t\,jk})$ is a density matrix given basis representation for a computational basis $\{|i\rangle\}$ of the space $\mathscr{H}_d$. Since quantum mechanics constrains $\rho_t$ to be a positive operator of unit trace, it is unambiguously specified by $d^2 - 1$ free parameters.

To reconstruct the state $\rho_t$, one performs a quantum measurement, represented by a probability-operator measurement (POM) $\Pi = \{\Pi_j\}$, whose

elements satisfy

$$\Pi_j \geq 0 \quad \text{and} \quad \sum_{j=1}^{M} \Pi_j = 1, \quad 0 \leq j \leq M \tag{2.2}$$

and correspond to measurement probabilities $\{p_j\}$. Here, $M$ is the total number of measurement configurations. The probabilities are related to $\Pi_j$ and $\rho_t$ *via* Born's Rule as follows:

$$p_j = \mathrm{tr}\{\rho_t \Pi_j\}. \tag{2.3}$$

The relation between the Born probabilities and unknown quantum state is formally represented by the map

$$\mathscr{M}[\rho_t] = \boldsymbol{p_t}. \tag{2.4}$$

A POM is said to be "informationally complete (IC)" if and only if for an arbitrary $\rho_t$, there is no quantum state other than $\rho_t$ that can be mapped into the same Born probabilities by $\mathscr{M}$. This condition can be satisfied when the number of linearly independent POM elements are not less than the number of free parameters of quantum states $d^2 - 1$. Owing to the sum-to-unity constraint for POM elements that is $\sum_{j=1}^{M} \Pi_j = 1$, at least one POM element is linearly dependent to the others, and this requires $M \geq d^2$ for the POM to be IC.

In practice. one performs a POM for a finite number of copies ($N$) of

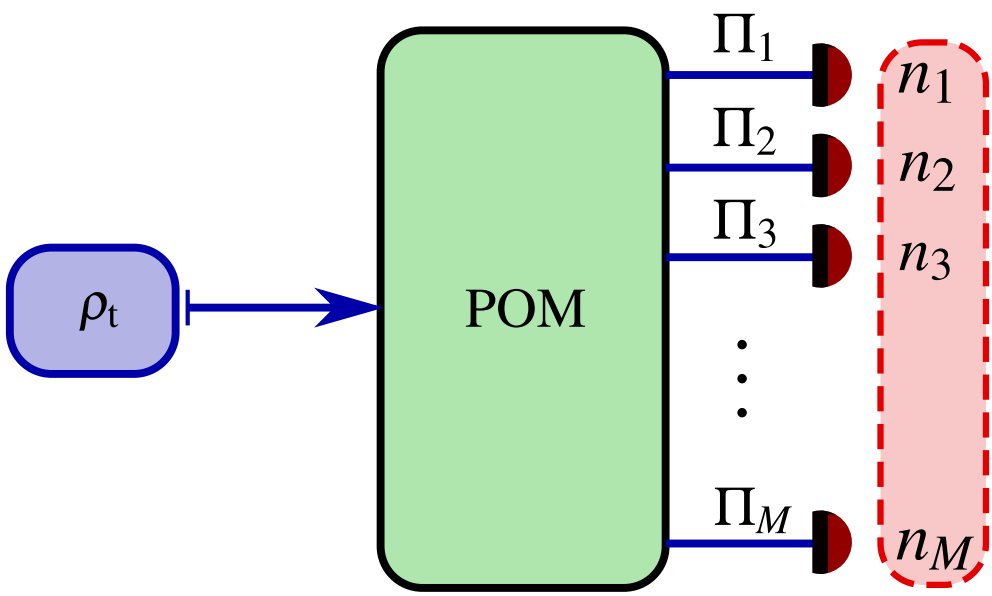


Figure 1: Schematic diagram of standard quantum state tomography. $\rho_t$ represents the unknown quantum state, and $\Pi_j$ and $n_j$ represent respectively the $j$th POM element and coressponding measurement frequency for $0 \leq j \leq M$

unknown true state $\rho_t$ and obtains the measurement frequencies $n_j$ for the $j$th measurement outcome with $\sum_{j=1}^{M} n_j = N$. Figure 1 shows a schematic diagram of the experimental procedure. A real dataset $\mathbb{D}$ consists of relative frequencies $f_j = n_j/N$, which deviates from the true Born probabilities described in Eq. 2.3 due to statistical noise for finite $N$, and approaches to the true probabilities only in the limit of large $N$.

The main aim of quantum-state tomography is to esimate the unknown quantum state $\rho_t$ from the dataset $\mathbb{D}$. $\mathbb{D}$ is said to be IC if the corresponding POM is IC. For noiseless IC data, linear inversion technique which takes the inverse of the map $\mathscr{M}$ described in 2.4 leads to perfect recovery of $\rho_t$. However, for real noisy IC data of relative measurement frequencies, this technique almost always leads to non-positive estimators, which are inconsistent with any quantum states [1]. Therefore, more sophisticated statistical methods are required to obtain an admissible quantum-state estimator.

The maximum likelihood estimation has been widely used for state estimation to cope with the quantum positivity constraint [2, 3, 4, 5, 6]. This method inspects the likelihood function $\mathscr{L}$, which is defined as the probability of getting the dataset $\mathbb{D}$ given the state $\rho$. For the frequency of the $j$th outcome $n_j$, the likelihood function is written as

$$\mathscr{L}\{\rho;\mathbb{D}\} = \prod_{j=1}^{M} {p_j}^{n_j} \tag{2.5}$$

where, $p_j$s are Born probabilities of $\rho$. If the whole dataset "$\mathbb{D}$" is IC, then

maximizing $\mathscr{L}$ over the entire space of positive unit-trace operators gives a unique maximum-likelihood (ML) state estimator. It is known that the ML estimator asymptotically converges to the true state in the large-$N$ limit [5]. The corresponding ML probabilities (collected into $\boldsymbol{p}_{\mathrm{ML}}$) are now the physical representatives of the actual relative frequencies for all measured outcomes.

## 2.2 Compressive tomography

For complex quantum systems in high-dimensional states, an IC measurement with $M \geq d^2$ gives a unique estimator which recovers $\rho_t$ for noiseless data but quickly turns into a resource-intensive practical problem, whereas a non-IC measurement with $M \ll d^2$ yields a non-singleton data convex set $\mathscr{C}$, closed under convex sum ot the estimator elements, in which all the elements of the set possess identical Born probabilities. This is summarized in Fig. 2. To overcome this problem, since a low rank state is often taken as the main target to be experimentally recovered, $\rho_t$ is often assumed to have low rank.

In this section, to cope with compressive tomography, which pertains to the reconstruction of a low rank state $\rho_t$, we review the general compressive estimation theory and introduce compressive measurements. In the first subsection, we summarize the mathematical framework of "compressed sensing (CS)", and in the second subsection, several compressive measurements are discussed with the estimation technique adopted from CS. In the last subsection, we introduce two types of POMs, each of which gives perfect recovery

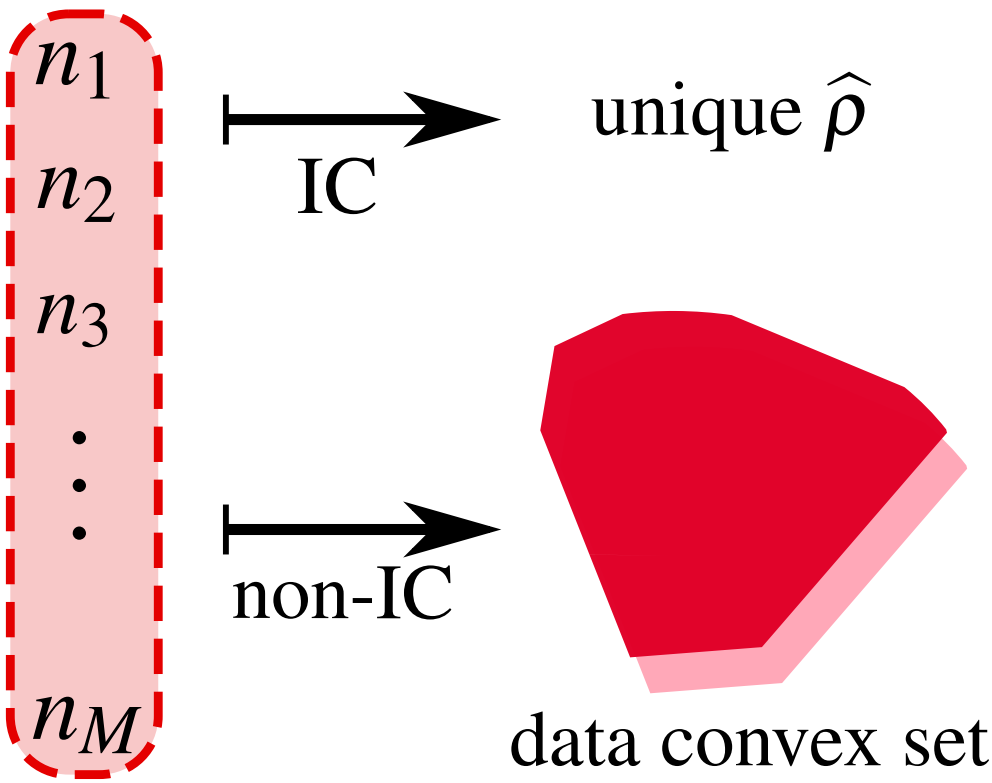


Figure 2: Schematic diagram of IC and non-IC measurements for quantum state tomography. Measurement scheme to obtain data is the same as in Fig. 1

for low-rank states, and compare them in terms of informational completeness with respect to the low-rank assumption.

### 2.2.1 Compressed sensing theory

The technique of compressed sensing (CS) developed by Candès, Donoho, and Tao [7, 8, 9] allows one to perform recovery of an unknown sparse signal with very few linear measurement configurations. The theory, which assumes a *known degree of sparsity* and defines a small set of specialized CS measurement, has been widely adopted in signal processing [10]. When the theory is applied to quantum state tomography [14, 17, 18, 19, 20], the relevent estimators form the set of $d \times d$ Hermitian matrices under the low-rank assumption for the unknown quantum state $\rho_t$ ($\mathrm{rank}\{\rho_t\} \leq r$) for a sufficiently small integer $r \ll d$. Based on the assumed prior information, the CS measurement map $\mathscr{M}_{CS}$ from a Hermitian matrix $A$ to noiseless data $\mathbb{D}$ is designed to satisfy the restricted isometry property (RIP) [11, 12] with $M = O(rd)$ as follows:

$$\begin{aligned} &\text{If } \mathscr{M}_{CS} : \rho \mapsto \mathbb{D} \text{ satisfies RIP, there exists some } 0 \leq \delta_r < 1 \text{ such that} \\ &\qquad (1-\delta_r) \parallel A \parallel_{HS}^2 \leq \parallel \mathscr{M}_{CS}\{A\} \parallel_2^2 \leq (1+\delta_r) \parallel A \parallel_{HS}^2 \\ &\qquad\qquad \text{for any Hermitian matrix } A \text{ such that } \mathrm{rank}\{A\} \leq r\,, \end{aligned} \tag{2.6}$$

where, $\| \cdot \|_{HS} = \sum_{j,k} | [\cdot]_{jk} |^2$ and $\| \cdot \|_2 = \sum_j | [\cdot]_j |^2$ are Hilbert–Schmidt and $l_2$ norms, respectively. The smallest constant $\delta_r$ for RIP is called the

"restricted isometry constant", which represents the degree of isometry for the rank-$r$ class Hermitian matrices corresponding to a given fixed noiseless dataset $\mathbb{D}$ with one of the Hermitian matrices being the unknown state $\rho_t$. For small $\delta_r$, all the corresponding rank-$r$ class Hermitian matrices giving the same dataset $\mathbb{D}$ have approximately the same $l_2$ norm.

This geometrical structure of the CS map with respect to $\mathbb{D}$ offers the possibility of getting a unique estimator among the Hermitian matrices by using a relevant choice of estimation method. Consistent with the prior information about rank, one seeks the unique unit-trace positive Hermitian estimator $\widehat{\rho}$ from a noiseless $\mathbb{D} = \mathscr{M}_{CS}[\widehat{\rho}]$ such that $\widehat{\rho}$ has the lowest rank. This estimator is given by

$$\widehat{\rho} : \arg\min \operatorname{rank}\{\rho\} \quad \text{subject to } \mathscr{M}_{CS}\{\rho\} = \mathbb{D},\ \operatorname{tr}\{\rho\} = 1. \tag{2.7}$$

However, since the matrix-rank minimization problem is known to be NP-hard [13], it was heuristically replaced by the trace-norm minimization problem [9, 15]. It was further shown that the trace-norm is the largest convex function that gives the tightest lower bound for the rank function [16] and that the trace-norm ($\|\cdot\|_{\text{tr}} = \operatorname{tr}\sqrt{(\cdot)(\cdot)^\dagger}$) minimization algorithm, which is

described as

$$\widehat{\rho} : \arg\min \ \| \rho \|_{\mathrm{tr}} \quad \text{subject to } \mathscr{M}_{CS}\{\rho\} = \mathbb{D},\ \mathrm{tr}\{\rho\} = 1 \tag{2.8}$$

, is equivalent to Eq. 2.7 when $\delta_{5r} < 1$ [16], where both Eqs. 2.7 and 2.8 lead to perfect recovery of $\rho_t$ [16].

The most recent study of the CS theory [14] focuses more on the state space and utilizes convex optimization in the presence of positivity constraint $\rho \geq 0$ for an arbitrary convex function $f(\rho)$ and CS measurement maps, which is depicted as

$$\widehat{\rho} : \arg\min f(\rho) \quad \text{subject to } \mathscr{M}_{CS}\{\rho\} = \mathbb{D},\ \mathrm{tr}\{\rho\} = 1,\ \text{and } \rho \geq 0 \tag{2.9}$$

and was rigorously proven to uniquely recover $\rho_t$.

For a more intuitive understanding on how the above convex optimization routine works, a pictorial description is presented in Fig. 3 that explains the CS procedure of trace-norm minimization on the recovery of $\rho_t$ subject to the positivity constraint $\rho \geq 0$. In the figure, the combined constraints $\mathscr{M}_{CS}\{\rho\} = \mathbb{D}$ and $\rho \geq 0$ are represented by the red convex set whose alignment relative to the set of the class of rank-$r$ states $S_r = \{\rho \in \mathscr{H}_d \mid \mathrm{rank}\{\rho\} \leq r\}$ is characterized by the property of the CS map (RIP) with $\rho_t$ in it. The blue

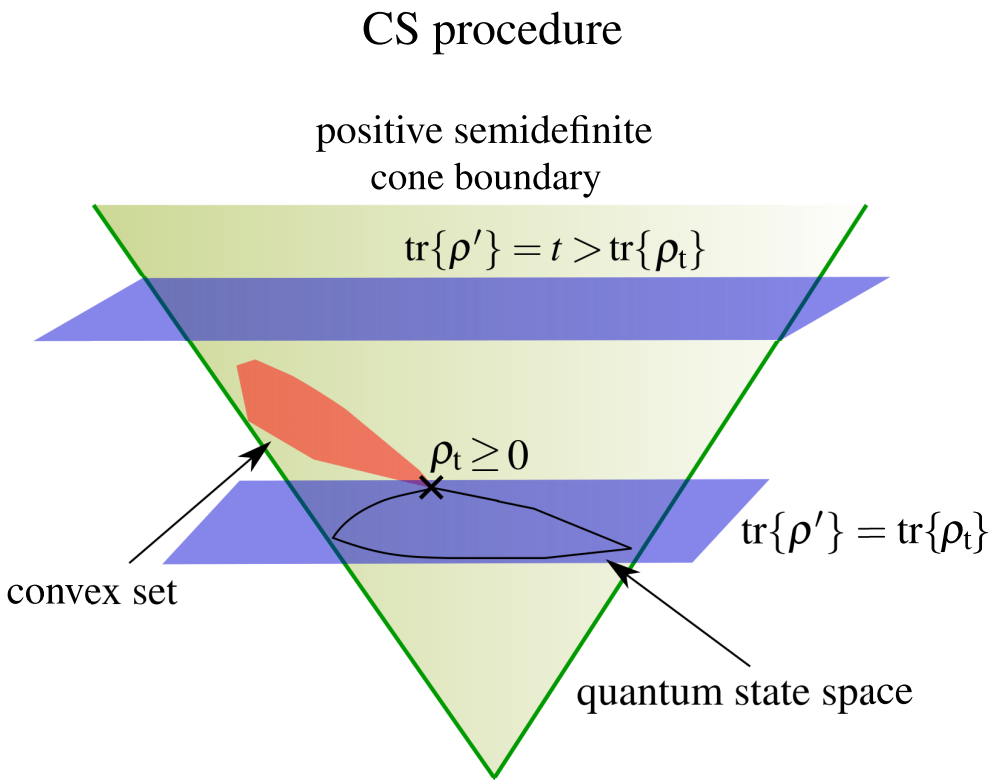


Figure 3: Pictorial review of CS. Noiseless data measured with a CS POM recovers $\rho_t$ through a convex optimization (typically trace-norm minimization) with positivity constraint

planes denote surfaces of equal trace (which is the same as trace-norm within the set of positive Hermitian matrices), each of which taking a different trace value. The geometrical relation between the red convex set and equal trace surfaces within the positive Hermitian space shows how the trace norm minimization procedure subject to the positivity constraint guides the estimator close to $\rho_t$ and ultimately recovers $\rho_t$ for which the convex set coincides with the unit trace surface.

The CS theory, based on the prior information $\text{rank}\{\rho_t\} \leq r$, introduces an appropriate convex optimization method to recover $\rho_t$ using a CS measurement map $\mathscr{M}_{CS}$ specialized for the rank-$r$ subspace. In the next subsection, we discuss specific compressive measurements that can recover the unknown state in CS.

### 2.2.2 Random Pauli measurements

Pauli (expectation value) measurements are popular for being one of the most feasible measurements in actual experimental situation for qubit systems, which make use of Pauli observables $\{\sigma_x, \sigma_y, \sigma_z\}$. The set of IC Pauli measurement for $n$-qubit systems with $d = 2^n$ is composed of all possible configurations of $n$ product observables,

$$E = \{\bigotimes_{j=1}^{n} \sigma_{k_j} \mid k_j = 0,1,2,3\}, \tag{2.10}$$

where, for each qubit, $\sigma_0, \sigma_1, \sigma_2$, and $\sigma_3$ are respectively $1, \sigma_x, \sigma_y$, and $\sigma_z$ for each $j$th qubit. Here, the identity observable 1 refers to the absense of a measurement on the qubit system which relies on the assumption that all measurements are local in that such a measurement on any other qubits does not change the qubit. The resulting set of Pauli observables $E$ is linearly independent with $M = 4^n = d^2$, which confirms that $E$ is IC.

The random Pauli measurement consists of observables randomly selected out of the IC Pauli measurement. This measurement for unknown quantum state $\rho_t$ gives *via* Born's rule (noiseless) data $\mathbb{D}$ of expectation value of the corresponding observable. The trace-norm minimization from CS estimation procedure for this measurement [14] was shown to result in a unique $\rho_t$ using $M = O(rd(\log d)^2)$ measurement configurations [14], where the algorithm is precisely given by

$$\widehat{\rho} : \arg\min \ \| \rho \|_{\mathrm{tr}} \quad \text{subject to } \mathrm{tr}\{\rho E_j\} = \mathrm{tr}\{\rho_t E_j\},\ \mathrm{tr}\{\rho\} = 1 \qquad (2.11)$$

An alternative version to the random Pauli measurement as yet another feasible compressive measurement scheme is the random Pauli-projective measurement, proposed by [14]. A Pauli-projective measurement is a set of projectors on the eigenbasis the of Pauli observables. For $n$-qubit systems ($d = 2^n$), it is composed of all possible products of eigenbasis projectors of Pauli

observable,

$$E' = \{\bigotimes_{j=1}^{n} |l\rangle_{k_j} \langle l| \mid k_j = 1,2,3 \text{ and } l = 0,1\}, \tag{2.12}$$

where, for the $j$th qubit, set of indices $\{k_j = 1,2,3\}$ denotes for three possible choices of Pauli observables $\{\sigma_x, \sigma_y, \sigma_z\}$ and set of indices $\{l = 0,1\}$ describes the two possible Pauli projectors that correspond to the eigenbasis of the chosen observable $\{|\downarrow\rangle_{k_j}, |\uparrow\rangle_{k_j}\}$, which lead to a total number of the measurement configurations $M = 3^n \cdot 2^n = 6^n$ that is larger than $d^2 = 4^n$. Since each eigenbasis is complete, the measurement set is POM. Whilst it might not always be straightforward to count the maximum number of linearly independent elements for the POM, *via* gram-matrix method, it is confirmed that the entire set of $6^n$ outcomes is IC, which implies that the measurement includes a set of $d^2 - 1 = 4^n - 1$ linearly independent elements.

Random Pauli-projective measurement is constructed by randomly picking Pauli observables. For the random selection of the observables out of $3^n$ configurations, the selected configurations of Pauli observables are taken to be mutually distinct without replacement so that none of two configurations coincides. Numerical simulation of CS estimation procedure under positivity constraint (Eq. 2.9) with respect to the measurement revealed compressive behavior that $\rho_t$ is recovered with $M = 20$ for 3 qubit system [14].

In the context of IC, the uniqueness of a CS estimator converging to true unknown state $\rho_t$ under the positivity constraint is IC with respect to the set

of rank-$r$ states $S_r$ within quantum state space. So IC, in this case, is really "rank-r strictly IC" [14, 20], which guarantees not only the uniqueness of the estimator within $S_r$ but also that there is no other quantum state estimators outside $S_r$. In fact, all CS measurements for $S_r$ are known to be rank-$r$ strictly IC but the converse is not so trivial. In the next subsection, we introduce two additional compressive measurements that are rank-$r$ strictly IC.

### 2.2.3 Strictly-complete measurements

An arbitrary rank-$r$ state $\rho_r$ is characterized by using the set of non-zero eigenvalues $\{p_1, p_2, \cdots, p_r\}$ and corresponding set of eigenvectors $\{|\Phi_1\rangle, |\Phi_2\rangle, \cdots, |\Phi_r\rangle\}$ as,

$$\rho_r = p_1|\Phi_1\rangle\langle\Phi_1| + p_2|\Phi_2\rangle\langle\Phi_2| + \cdots + p_r|\Phi_r\rangle\langle\Phi_r|, \tag{2.13}$$

Considering unit trace constraint imposed on each pure state $|\Phi_j\rangle\langle\Phi_j|$ and $\rho_r$, $\{|\Phi_j\rangle\langle\Phi_j|\}$ and $\{p_1, p_2, \cdots, p_r\}$ have respectively $2d-2$ (up to global phase) and $r-1$ free parameters. This yields, with further constraint of $\{|\Phi_j\rangle\}$ being mutually orthogonal, $2rd - r^2 - 1$ free parameters for $\rho_r$, and thus requires that the number of rank-$r$ IC POM elements to be no less than $2rd - r^2$. But as long as the set of rank-$r$ states is not a vector space, this bound presents only a necessary condition.

In this section, we introduce rank-$r$ strictly IC POMs constructed by [20]. These POMs are element-probing (EP-POM), which directly address

the $2rd - r^2 - 1$ independent elements of the density matrix representation of an unknown $\rho_r$ *via* Born's rule.

Firstly, given an orthonormal computational basis $\{|1\rangle, |2\rangle, \cdots, |d\rangle\}$, *Baldwin–Flammia* EP-POM (BF) $E_{BF}$ for an arbitrary rank $r$ state $\rho_r = \sum_{j,k} |j\rangle \rho_{r\,jk} \langle k|$ is describes as,

$$
\begin{aligned}
E_j &= a_j |j\rangle\langle j|, \ j = 1, \cdots, r \\
E_{j,k} &= b_j (1 + |j\rangle\langle k| + |k\rangle\langle j|), \ k = j+1, \cdots, d \\
\tilde{E}_{j,k} &= b_j (1 - i|j\rangle\langle k| + i|k\rangle\langle j|), \ k = j+1, \cdots, d \\
E_{2dr-r^2+1} &= 1 - \sum_{j=1}^{r} [E_j + \sum_{k=j+1}^{d} (E_{j,k} + \tilde{E}_{j,k})] ,
\end{aligned}
\tag{2.14}
$$

with $a_j$ and $b_j$ chosen such that $E_{2rd-r^2+1} \geq 0$. The Born probability from $E_j$ recovers the $j$th diagonal element of $\rho_r$ ($\rho_{r\,jj}$), while Born probabilities from $E_{j,k}$ and $\tilde{E}_{j,k}$ recover off-diagonal elements of $\rho_r$ ($\rho_{r\,jk}$), each of which corresponds to respectively $\mathrm{Re}(\rho_{r\,jk})$ and $\mathrm{Im}(\rho_{r\,jk})$. The entry $\rho_{r\,kj}$ is clearly just the conjugate of $\rho_{r\,jk}$. With all the POM elements of $E_{BF}$, all $\rho_{r\,jk}$s for $\{(j,k) \mid j = 1, \cdots, r \text{ or } k = 1, \cdots, r\}$ are recovered. The remaining $(d-r)^2$ $\rho_{r\,jk}$s for $\{(j,k) \mid j = r+1, \cdots, d \text{ and } k = r+1, \cdots, d\}$ can be derived from the previously estimated elements by using the rank additivity property of Schur complement for block-diagonal components of the density matrix . Overall, the total number of POM elements is $M_{BF} = 2rd - r^2 + 1$, which exceeds the necessary bound $2rd - r^2$ by 1.

Secondly, there is also the *Baldwin–Goyeneche* EP-POM (BG) $E_{BG}$ for an arbitrary rank-$r$ state $\rho_r$ given dimension $d = 2^n$, which is constructed with $4r+1$ specifically chosen orthonormal bases (or equivalently $M_{BG} = (4r+1)d$). These are defined as,

$$
\begin{aligned}
&B_0 = \{|1\rangle, \cdots, |d\rangle\}, \\
&{B_1}^{(l)} = \{\frac{1}{\sqrt{2}}(|j\rangle \pm |j+l\rangle) \mid j = 1, \cdots, m, 2m+1, \cdots, 3m, \cdots, d-2m+1, \cdots, d-m\}, \\
&{B_2}^{(l)} = \{\frac{1}{\sqrt{2}}(|j\rangle \pm i|j+l\rangle) \mid j = 1, \cdots, m, 2m+1, \cdots, 3m, \cdots, d-2m+1, \cdots, d-m\}, \\
&{B_3}^{(l)} = \{\frac{1}{\sqrt{2}}(|j\rangle \pm |j+l\rangle) \mid j = m+1, \cdots, 2m, 3m+1, \cdots, 4m, \cdots, d-m+1, \cdots, d\}, \\
&{B_4}^{(l)} = \{\frac{1}{\sqrt{2}}(|j\rangle \pm i|j+l\rangle) \mid j = m+1, \cdots, 2m, 3m+1, \cdots, 4m, \cdots, d-m+1, \cdots, d\}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad l = 1, \cdots, r,
\end{aligned}
\tag{2.15}
$$

with $m = 2^p$ and $p$ is the largest integer such that $l/2^p$ is an integar. The Born probabilities of projectors on $B_0$ recover the diagonal elements of $\rho_r$, whereas those on ${B_1}^{(l)}, {B_2}^{(l)}, {B_3}^{(l)}$ and ${B_4}^{(l)}$ recover the $l$th and $(d-l)$th off-diagonal elements of $\rho_r$ ($\rho_{r_{jj+l}}$ with $j = 1, \cdots, d$), since the probabilities of $\frac{1}{\sqrt{2}}(|j\rangle \pm |j+l\rangle)$ and $\frac{1}{\sqrt{2}}(|j\rangle \pm i|j+l\rangle)$ respectively address $\mathrm{Re}(\rho_{r_{jj+l}})$ and $\mathrm{Im}(\rho_{r_{jj+1}})$. Thus by measuring all the POM elements of $E_{BG}$, all $\rho_{r_{jj+l}}$s for $j = 1, \cdots, d$ and $l = 0, \cdots, r$ are recovered. The remaining $\rho_{r_{jj+l}}$s for $l = r+1, \cdots, d$ can be derived from the previously estimated elements in a very similar way to BF after rearranging the unmeasured elements through

unitary operations. The number of bases, $4r+1$, is valid regardless of the dimension of the state space $d$.

Upon comparing $M_{BF}$ and $M_{BG}$ with respect to the range of $r$, $M_{BF}$ and $M_{BG}$ exceed $d^2$ respectively at $r = d$ and $r = \frac{d}{4}$, which implies that $E_{BF}$ has a wider rank range where compressive effect exists. However, since some elements of $E_{BF}$ are not projectors, and hence difficult to be realized, $E_{BG}$ , whose elements are projectors on bases, remains the only feasible rank-$r$ strictly IC POM out of the two.

# Chapter 3

# Adaptive compressive tomography with no *a priori* information

In chapter 2, we have summarized the basic contents of quantum state tomography and studied the theory of compressed sensing by introducing several known compressive measurements. We reiterate that compressed sensing assumes the prior knowledge about the rank of the true state, which must be verified before the real experiment for compressed sensing tomography is performed. If the upper bound of the rank is too low, the chosen compressive measurement might be informationally incomplete and one would then require more measurement outcomes to get a unique estimator. On the other hand, if the upper bound is too high, it yields an informationally overcomplete measurement that overuses measurement resources.

In this chapter, we introuduce a new adaptive tomographic scheme for compressive state reconstruction that is partially inspired by the current compressed-sensing paradigm, but uses no prior information about true state. We explain how the new scheme is favorable by studying both real experimental and numerical simulation results and show that our scheme solves the existing problem in the compressed sensing theory.

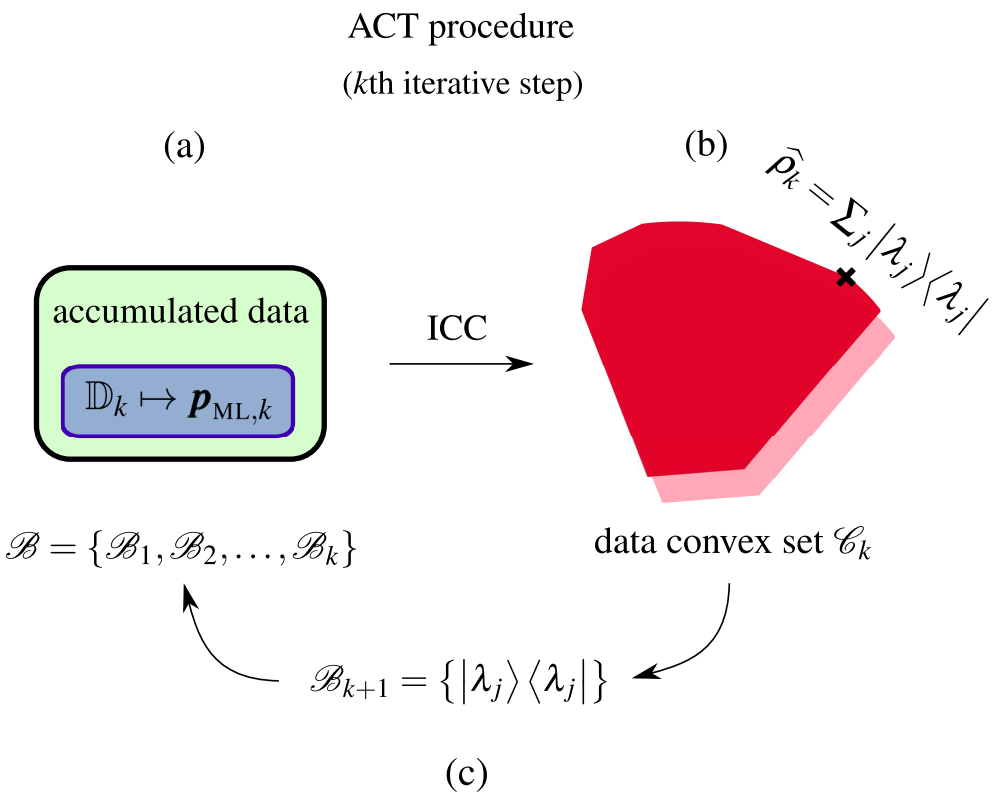


Figure 4: Schematic diagram of ACT at the $k$th iterative step. Whether the size of the data convex set $\mathscr{C}_k$ is non-zero or not is checked by ICC and next good measurement basis is chosen by relevant adaptive strategy

## 3.1 Adaptive compressive tomography

ACT consists of iterative steps of adaptive orthonormal basis measurements. To search for the final unique estimator, at every iterative step, the maximum-likelihood (ML) estimation method is adopted. The ACT scheme involves two main procedures; informational completeness certification (ICC) and adaptive strategy, which are sketched in Fig. 4. Given an unknown rank-$r$ state $\rho_r$ of

Hilbert-space dimension $d$, as the first iterative step, the scheme starts with a measurement of the computational basis $\mathscr{B}_1 = \{|1\rangle, |2\rangle, \cdots, |d\rangle\}$ and one acquires first data in terms of relative frequencies $\mathbb{D}_1 = \{\nu_{j'1} \mid \sum_{j'=0}^{d-1} \nu_{j'1} = 1\}$ and their corresponding Born probabilities $\boldsymbol{p}_{\mathrm{ML},1}$. This yields an ML data convex set $\mathscr{C}_1 = \{\rho' \mid \mathscr{M}_1[\rho'] = \boldsymbol{p}_{\mathrm{ML},1}\}$ within state space, where $\mathscr{M}_1$ is a measurement map for $\mathscr{B}_1$. When $\mathbb{D}_1$ is assumed to be noiseless, the ML probabilities coincide with the true Born probabilities as $\boldsymbol{p}_{\mathrm{ML},1} = \boldsymbol{p}_{t,1} = (\langle 1|\rho_r|1\rangle, \langle 2|\rho_r|2\rangle, \cdots, \langle d|\rho_r|d\rangle)$, with the set $\mathscr{C}_1$ containing not only the true state $\rho_r$ but also other infinitely many ML estimators consistent with the measured data. The size of the first data convex set $\zeta_1$ is, of course, confirmed to be non-zero by ICC and an appropriate next (the second iterative step) measurement basis $\mathscr{B}_2$ is chosen *via* the adaptive strategy. As the scheme progresses, at the $k$th iterative step, one collects the $k$th set of relative frequencies $\mathbb{F}_k = \{\nu_{j'k} \mid \sum_{j'=0}^{d-1} \nu_{j'k} = 1\}$ from measuring the basis $\mathscr{B}_k$, where $\mathscr{B}_k$ is adaptively chosen based on the $(k-1)$th accumulated dataset $\mathbb{D}_{k-1}$. Then, after renormalizing the frequencies, one acquires the $k$th accumulated data $\mathbb{D}_k = (\frac{k-1}{k})\mathbb{D}_{k-1} \cup (\frac{1}{k})\mathbb{F}_k$ and corresponding Born probabilities $\boldsymbol{p}_{\mathrm{ML},k}$. This yields, along with the measurement bases $\{\mathscr{B}_1, \mathscr{B}_2, \cdots, \mathscr{B}_k\}$, the ML data convex set $\mathscr{C}_k = \{\rho' \mid \mathscr{M}_k[\rho'] = \boldsymbol{p}_{\mathrm{ML},k}\}$, where $\mathscr{M}_k$ is a measurement map for $\{\mathscr{B}_1, \mathscr{B}_2, \cdots, \mathscr{B}_k\}$. When the accumulated dataset $\mathbb{D}_k$ is noiseless, $\boldsymbol{p}_{\mathrm{ML},k} = \frac{1}{k}(\boldsymbol{p}_{t,1}, \boldsymbol{p}_{t,2}, \cdots, \boldsymbol{p}_{t,k})$ holds along with $\boldsymbol{p}_{t,k'}$ being true Born probabilities with respect to $\mathscr{B}_{k'}$ $(1 \leq k' \leq k)$. The iterative procedure terminates when ML data convex set shrinks to a point with the size of the convex set confirmed to be zero by ICC. We

shall denote the number of measurement bases at this termination stage as $k_{\mathrm{IC}}$ and from the singleton $\mathscr{C}_{k_{\mathrm{IC}}}$, we get the unique estimator that converges to the true rank-$r$ state $\rho_r$ for noiseless measurements.

In the following subsections, we will explain how we specified two main components (ICC and adaptive strategy) that completely define ACT.

### 3.1.1 Informational completeness certification (ICC)

Given the ML data convex set $\mathscr{C}_k$ at the $k$th iterative step, ICC plays the role of certifying whether the size of the convex set is zero or not. To carry out the procedure, we introduce a linear function $f_Z(\rho') = \mathrm{tr}\{\rho' Z\}$, where $Z$ is a full-rank quantum state such that $Z \neq 1/d$, and define size function as $s_{\mathrm{CVX},k} = (f_{max,k} - f_{min,k})/(f_{max,1} - f_{min,1})$. Here, we denote $f_{max,k}$ and $f_{min,k}$ as the maximum and minimum values of the function $f$ in $\mathscr{C}_k$ respectively. Since the set $\mathscr{C}_k$ is convex (is closed under convex sum of the set elements), maximization and minimization of $f$ over $\rho'$ in $\mathscr{C}_k$ yield unique $\rho_{max,k}$ and $\rho_{min,k}$ that respectively give $f_{max,k}$ and $f_{min,k}$, which can be visually understood in Fig. 5. For noiseless measurements, with increasing number of iteration $k$, since more linear data constraints are added, it is clear that $\mathscr{C}_1 \supseteq \mathscr{C}_2 \supseteq, \cdots, \supseteq \mathscr{C}_{k_{\mathrm{IC}}}$. This implies a monotonically decreasing $f_{max,k}$ and increasing $f_{min,k}$, and so that the size function $s_{\mathrm{CVX},k}$ such that $s_{\mathrm{CVX},k+1} \leq s_{\mathrm{CVX},k}$.

As long as linear function is taken into account, ICC corresponds to semi-definite programming (SDP), which is known to be solvable within polynomial time [23]. This is markedly different from checking RIP of a CS POM,

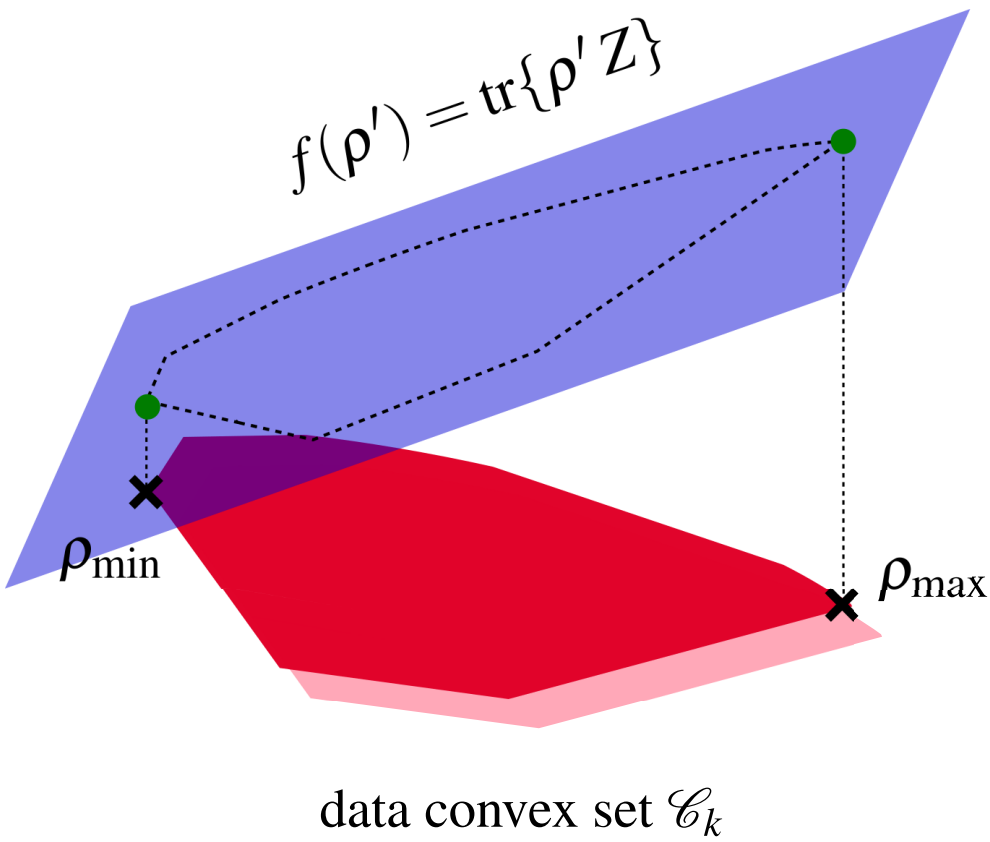


Figure 5: The ICC procedure. The linear function $f(\rho') = \mathrm{tr}\{\rho' Z\}$ (represented by a hyperplane) has global optima at the edges of $\mathscr{C}_k$. The corresponding auxiliary extrema states $\rho_{\max}$ and $\rho_{\min}$ determine whether $\mathscr{C}_k$ is a singleton or not through $s_{\mathrm{CVX},k}$.

which is an NP-hard problem [24]. The accumulated data and/or measurement are considered to be IC when, with a small finite threshold $\varepsilon$, $s_{\mathrm{CVX},k} \leq \varepsilon$, and ACT scheme terminates. The whole ICC algorithm is summarized below.

---

**ICC in the $k$th step**

1. Maximize and minimize $f_Z(\rho') = \mathrm{tr}\{\rho' Z\}$ for a fixed, randomly-chosen full-rank state $Z \neq 1/d$ to obtain $f_{\max,k}$ and $f_{\min,k}$ subject to
   - $\rho' \geq 0$ , $\mathrm{tr}\{\rho'\} = 1$ ,
   - $\mathscr{M}_k[\rho'] = \boldsymbol{p}_{\mathrm{ML},k}$ for $\boldsymbol{p}_{\mathrm{ML},k}$ obtained from the accumulated $\mathbb{D}_k$ thus far .
2. Compute $0 \leq s_{\mathrm{CVX},k} \leq 1$ and check if it is smaller than some threshold $\varepsilon$.
3. If $s_{\mathrm{CVX},k} < \varepsilon$, terminate ACT. Continue otherwise.

---

### 3.1.2 Adaptive strategy

Given the ML data convex set $\mathscr{C}_k$ at the $k$th iterative step, once the set is confirmed to have a non-zero size through ICC, the adaptive process commences and the next $(k+1)$th measurement basis shall be appropriately chosen. Here, we introduce an adaptive strategy that picks the $k$th *a posteriori* estimator $\widehat{\rho}_k$ from $\mathscr{C}_k$ and takes its eigenbasis as the next measurement basis. This is motivated by the fact that if the chosen measurement basis is the eigenbasis of $\rho_r$, we acquire probabilities of all the non-zero (diagonal) entries of the density matrix of $\rho_r$ in its eigenbasis representation. To be more specific, if we take eigenbasis of $\rho_r$, thereby specifying all its $r$ positive eigenvalues, as one of the measurement bases, there remains only $r^2-r$ free parameters of the unit-trace positive $\widehat{\rho}$ that gives the same Born probabilities as $\rho_r$, apart from the eigenvalues. Thus, including the eigenbasis, one only needs $k_0 = \lceil (r^2-r)/(d-1) \rceil + 1$ linearly independent bases to unambiguously characterize $\widehat{\rho} = \rho_r$ for noiseless data. However, this assumes the unrealistic situation of picking precisely the eigenbasis of $\rho_r$ without any assistance.

Instead, inspired by the CS procedure, one may seek the lowest-rank estimator in $\mathscr{C}_k$ that is expected to approach $\rho_r$ in the last iterative step. To establish a feasible scheme for the procedure, we introduce a protocol of minimizing a concave function in $\mathscr{C}_k$ , which is designed to possess the rank-minimizing property, and take the optimal state as $\widehat{\rho}_k$. For the choice of concave functions, we shall consider two exemplifying functions in quantum in-

formation theory: the von Neumann entropy $S(\rho') = -\mathrm{tr}\{\rho' \log \rho'\}$ and the linear entropy $S_{\mathrm{L}}(\rho') = 1 - \mathrm{tr}\{\rho'^2\}$, both of which are similar in value for nearly-pure states and have the global minimum as zero when evaluated with pure states.

As long as the set $\mathscr{C}_k$ is in quantum state space, this is similar to conventional CS procedure that minimizes convex function subject to data and positivity constraints. In fact, entropy minimization is numerically found in [25] to be more compressive on the matrix recovery than trace-norm minimization in CS.

Our adaptive strategy over the ML data convex set gives an effective compression on the IC data, which will be demonstrated by both simulations and experiments in the next section. This compressive effect arises from the action of the positivity constraint that strongly restricts the area of the rank-deficient boundary of $\mathscr{C}_k$, on which all *a posteriori* estimators lie. As the iteration proceeds, the *a posteriori* estimator progressively approaches the true state $\rho_r$, which is depicted in Fig. 6.

In practical many-body experiments, tensor-product bases, rather than general entangled ones, are much more easily measurable. We can accordingly adjust the adaptive strategy of our ACT scheme to cater for product-bases measurements (PACT) by choosing some product state $\widehat{\rho}_{prod,k}$ that is closest to the actual optimal $\widehat{\rho}_k$ from ACT. One alternative would be to pick a new tensor product seed state whose Hilbert–Schmidt distance from $\widehat{\rho}_k$ is the smallest, and minimize either the von Neumann or linear entropy by taking

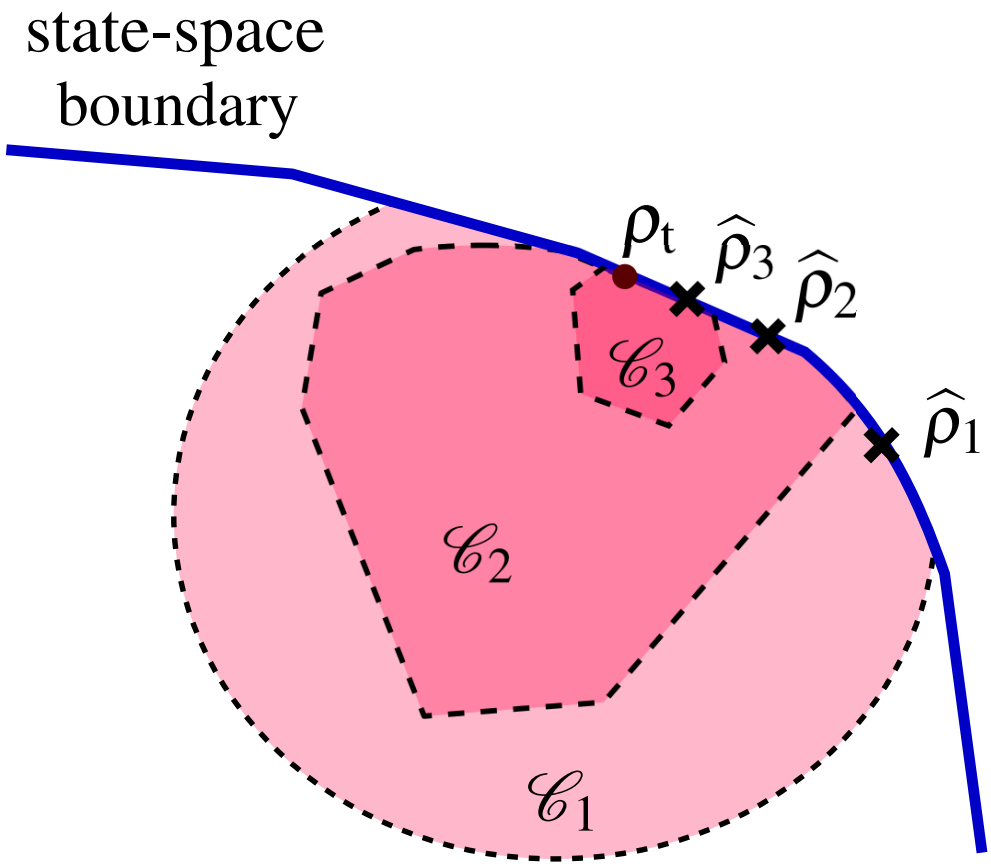


Figure 6: The progress of the adaptive strategy in ACT. In each step, (linear-)entropy minimization is performed under both positivity and data (ML) constraints to give the lowest-(linear-)entropy $\widehat{\rho}_k$. The data convex set $\mathscr{C}_k$ makes a common boundary with the state space. This boundary shrinks as $k$ increases, so that the $\widehat{\rho}_k$s converge to $\rho_r$.

the seed state as a starting point of the procedure.

We may state a very similar routine for PACT below:

---

**(P)ACT procedure**

Beginning with $k = 1$ and a random computational basis $\mathscr{B}_1$:

1. Measure $\mathscr{B}_k$ and collect the relative frequency data $\sum_{j'=0}^{d-1} \nu_{j'k} = 1$.
2. From $\{\nu_{0k'}, \ldots, \nu_{d-1\,k'}\}_{k'=1}^{k}$, obtain $kd$ physical ML probabilities.
3. Perform ICC with the ML probabilities and compute $s_{\mathrm{CVX},k}$:
   - **If** $s_{\mathrm{CVX},k} < \varepsilon$, terminate ACT and take $\rho_{\max} \approx \rho_{\min}$ as the estimator and report $s_{\mathrm{CVX},k}$.
   - **Else** Proceed.
4. Choose a lowest-(linear-)entropy $\widehat{\rho}_k \in \mathscr{C}_k$ in $\mathscr{C}_k$
5. Define $\mathscr{B}_{k+1}$ to be the eigenbasis of $\widehat{\rho}_k$ for ACT, or a local basis close to it for PACT *via* some prechosen distance minimization technique.
6. Set $k = k+1$ and repeat.

---

## 3.2 Simulation and experiment

### 3.2.1 Simulation and experimental setup

For ICC, the size function is calculated with SDP by using the CVXOPT package for MATLAB [26, 27]. ICC was applied to both PACT and RP for both the simulations as well as experiments. The threshold $\varepsilon$ for ICC was set to 0.001. For the adaptive strategy, von Neumann entropy minimization was performed for (P)ACT using a recent superfast algorithm that exploits the "accelerated gradient method" to greatly accelerate the maximization of the likelihood function [6]. To apply this algorithm to the minimization of von Neumann entropy over ML data convex set, we utilized numerical concepts introduced in [4].

To perform actual experiments, single-photon qubits with the orbital angular momentum (OAM) are measured. We considered the OAM degree of freedom of Laguerre-Gauss (LG) modes. For the ease of control, radial mode indices are fixed to $p = 0$ with azimuthal mode $\ell$ varied freely. Information about OAM states is specifically encoded in the azimuthal mode $\ell$ in the form of a phase factor of the helical wave front of the mode given by $e^{i\ell\phi}$, where $\phi$ is the azimuthal coordinate. Figure 7 describes the schematics of the experiemental setup, where two spatial light modulators (SLM-A and SLM-B) and single-mode fiber (SMF) are used. OAM states are automatically generated by using a holographic technique called "intensity masking" [28] with the corresponding intensity pattern being modulated in SLM-A. A pair of

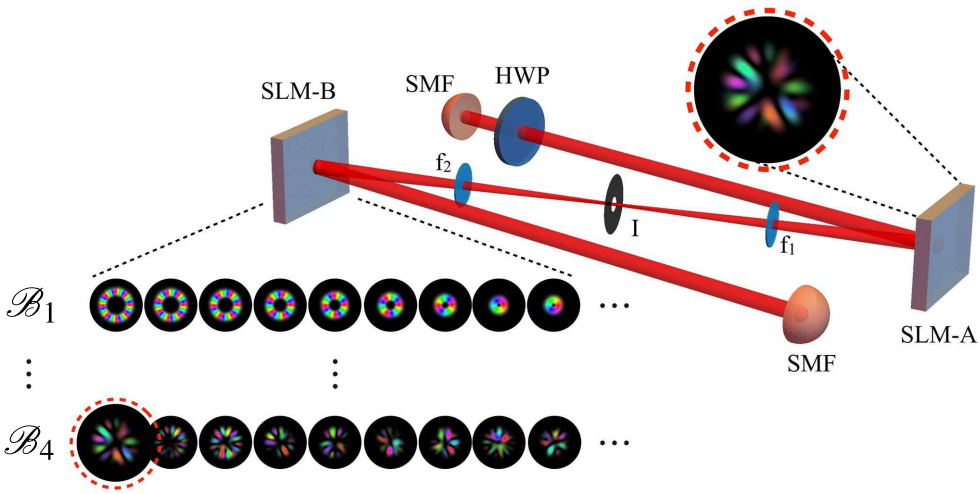


Figure 7: Schematic of the OAM-based experimental setup. A 16-dimensional OAM state is generated at SLM-A using a holographic technique that allows the tailoring of the intensity and phase profile of the incoming beam. The modulated first-order of diffraction is filtered out using an iris (I) and a pair of lenses ($f_1$ and $f_2$). A similar holographic technique is used at the second SLM-B to measure the state in a given basis. The first measurement basis, $\mathscr{B}_1$, is given by the OAM computational basis. In the case of the rank 1 state shown on SLM-A, the corresponding eigenbasis is achieved after the fourth iteration.

lenses and iris between SLM-A and SLM-B filter out the first order of diffraction where all information is encoded, and SLM-B and SMF together play the role in carrying out projection on a given basis *via* an optical technique called "intensity flattening" [29]. Single photons in the beginning are generated by pumping a 3 mm $\beta$-barium borate type I nonlinear crystal with a quasi-continuous wave laser at a wavelength of 355 nm.

### 3.2.2 Simulation and experimental result

We performed both numerical noiseless simulations and experiments for 4-qubit ($d = 16$) case of ranks 1 to 3. For each rank, 5 true states are randomly selected and probed from the group of states uniform in Hilbert–Schmidt distance. To check the efficiency of our (P)ACT scheme, we make comparisons with the random-Pauli projective measurement (RP).

Firstly, the size function $s_{\mathrm{CVX},k}$ and fidelity between $\rho_r$ and $\widehat{\rho}_k$ against the number of measured bases $k$ is illustrated in Fig. 8. Filled markers refer to rank-1 states while unfilled markers refer to rank-3 states. It is clearly confirmed that in terms of the converging speed of both $s_{\mathrm{CVX},k}$ and fidelity, ACT is most efficient, giving the highest compression efficiency (lowest $k_{\mathrm{IC}}$).

Secondly, directly comparing the compressive efficiency, the number of measured IC bases $k_{\mathrm{IC}}$ and fidelity between $\rho_r$ and $\widehat{\rho}_{k_{\mathrm{IC}}}$ (IC fidelity) from rank-1 to rank-3 are compared in Fig. 9, plotted with the known BG benchmark scaling. We can also confirm that ACT is the most favorable scheme in that it has the lowest $k_{\mathrm{IC}}$. That ACT has a higher IC fidelity than any other scheme

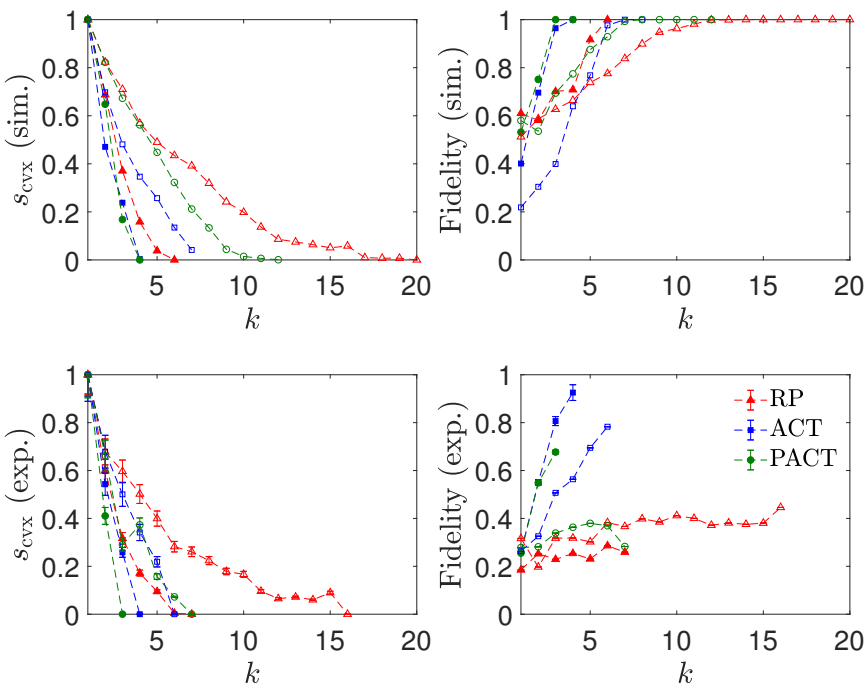


Figure 8: Plots of noiseless simulation and experimental results for $s_{\mathrm{CVX}}$ and fidelity between $\rho_r$ and $\widehat{\rho}_k$.

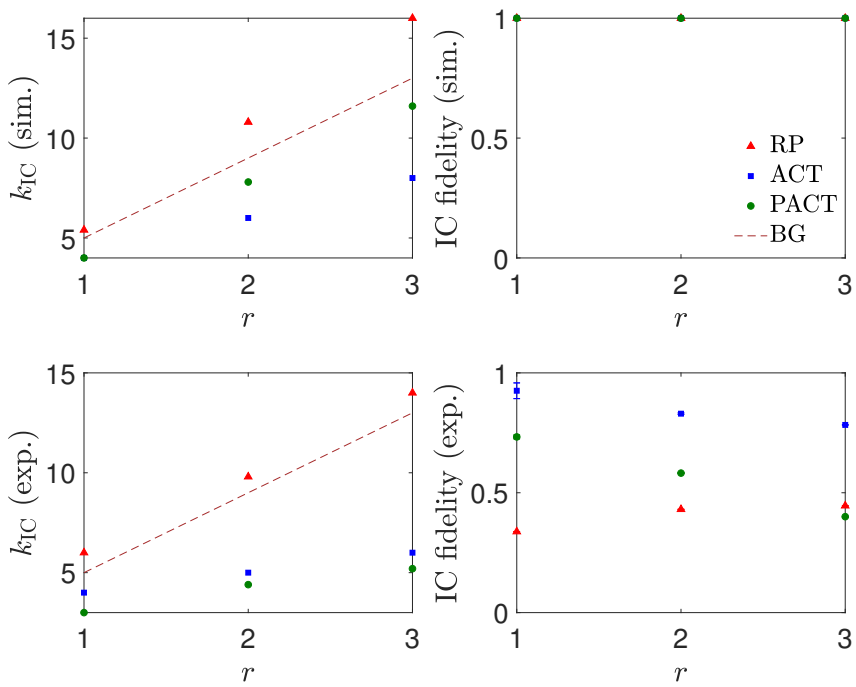


Figure 9: Plots of noiseless simulation and experimental result for $k_{\mathrm{IC}}$ and fidelity between $\rho_r$ and $\widehat{\rho}_{k_{\mathrm{IC}}}$.

from the experimental results also implies that ACT is most robust against noise for the OAM setup.

Moreover, from Figs. 8 and 9, fidelity behaviors of PACT and RP for the experimental results are poor, and so do their corresponding IC fidelities. On the other hand, the experimental ACT shows a gradually increasing fidelity behavior over $k$ that finally reaches a high IC fidelity. This difference is due to the technical bias of the OAM set up appearing for finite $N$ when measured bases are far from the eigenbasis of $\rho_r$; while random eigenbasis of product

Pauli observable for RP remains far from the eigenbasis of $\rho_r$, for ACT, the eigenbasis of $\widehat{\rho}_k$ gets closer to that of $\rho_r$ as $k$ increases. Therefore, less technical bias appears for ACT than for RP, which explains ACT's outperformance of the rest.

So far, the discussion was done for 5 random states of 4-qubits. In the next section, we further generate noiseless simulation data for different number of qubits and compare PACT with random schemes and previously known compressive measurements discussed in Sec. 2.2.3.

## 3.3 Further simulation and result

### 3.3.1 Simulation speculation

To investigate more about the compression efficiency and scaling with respect to dimension, we numerically compared PACT with RP and two kinds of random bases measurements : random (local-) Haar-uniform bases measurement (Random (local-) Haar) and random full rank states' eigenbases measurement (Random). We carried out noiseless numerical simulations on states with $1 \leq r \leq 6$ respectively for 3 qubits ($d = 8$), 4 qubits ($d = 16$), 5 qubits ($d = 32$), and 6 qubits ($d = 64$). For each number of qubits and rank, simulation results are averaged over 100 random target states. As in previous simulations, presented in Sec. 3.2.2, tested quantum states are randomly selected from the set of uniformly distributed states in Hilbert–Schmidt measure [30]. Further technical details for the generation of random Haar-uniform

bases and random states are considered in appendix A and B.

### 3.3.2 $s_{\mathrm{CVX}}$ and $k_{\mathrm{IC}}$ comparison

First of all, we compare (P)ACT with the other schemes in terms of $s_{\mathrm{CVX}}$ and $k_{\mathrm{IC}}$. For (P)ACT, we present here the scheme that minimizes von Neumann entropy during adaptive procedure as a representative.

Firstly, $s_{\mathrm{CVX}}$ is compared for 4, 5, and 6 qubits in Fig. 10. Top, middle, and bottom rows refer to rank-1, rank-2 and rank-3 states respectively. For all schemes and numbers of qubits, we confirmed that $s_{\mathrm{CVX}}$ is monotonically decreasing. Comparing between different schemes, in terms of the converging speed, (P)ACT outperforms RP. A faster convergence for $s_{\mathrm{CVX}}$ is equivalent to a lower $k_{\mathrm{IC}}$, both of which imply higher compression efficiency.

Secondly, compression efficiencies are more directly compared for several different schemes *via* $k_{\mathrm{IC}}$ in Fig. 11. It more clearly shows that ACT has the strongest compression efficiency, while RP has the weakest. More specifically, in terms of compression, ACT is more efficient than random Haar, whereas PACT gives almost the same $k_{\mathrm{IC}}$s, and so the same compression efficiencies as those of random local-Haar. We infer that this similarity comes from the restriction of measurements to product bases, which effectively incurs very similar intrinsic randomness on product measurement choices for both PACT and random local-Haar.

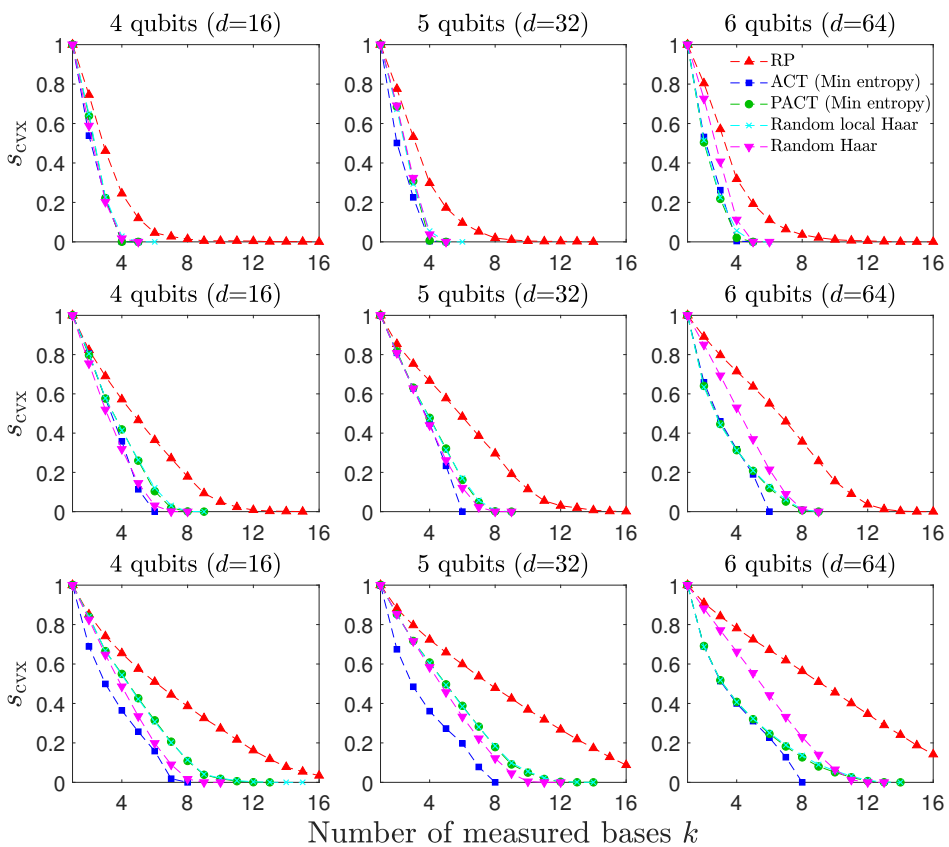


Figure 10: Plots of $s_{\mathrm{CVX}}$ for 4, 5, and 6 qubits. Top, middle, and bottom rows refer to rank-1, rank-2 and rank-3 states respectively. For each number of qubits and rank, data markers are averaged over 100 random states from uniform distribution with respect to Hilbert–Schmidt measure.

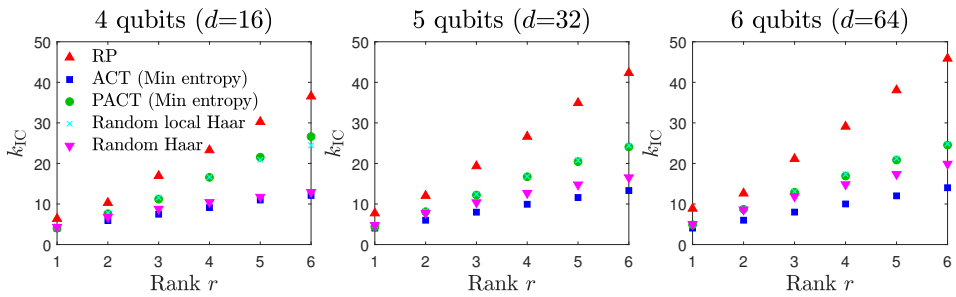


Figure 11: Plots of noiseless simulation of $k_{\mathrm{IC}}$ for 4, 5, 6 qubits. For each number of qubits and rank, data markers are averaged over 100 random states from uniform distribution with respect to Hilbert–Schmidt measure.

### 3.3.3 $k_{\mathrm{IC}}$ behavior

Since we confirmed that (P)ACT is more efficient than RP, we focused on the $k_{\mathrm{IC}}$ behavior and compared the adaptive schemes with the mentioned two random bases measurement schemes and compressive measurements designed by *Baldwin et al.* (BF and BG), where all the comparisons are summarized in Fig. 12. Here we present two different ACT schemes, which respectively adopt the minimization of von Neumann entropy $S$ and linear entropy $S_{\mathrm{L}}$ in the adaptive procedure. From the relation between linear entropy $S_{\mathrm{L}}$ and purity $P$ described as $S_{\mathrm{L}}(\rho') = 1 - P(\rho')$, where $P(\rho') = \mathrm{tr}\{\rho'^2\}$, the latter ACT scheme which minimizes $S_{\mathrm{L}}$ also maximizes $P$.

Firstly, comparing two ACT schemes with BF on the left of Fig. 12, we found out that two ACT schemes show almost the same $k_{\mathrm{IC}}$ behavior, where both behaviors are in good agreement with that of the shifted BF $\frac{M_{BF}}{d} + 2 = (2r+2) - \frac{r^2-1}{d}$, indicating that the number of projectors chosen by our ACT

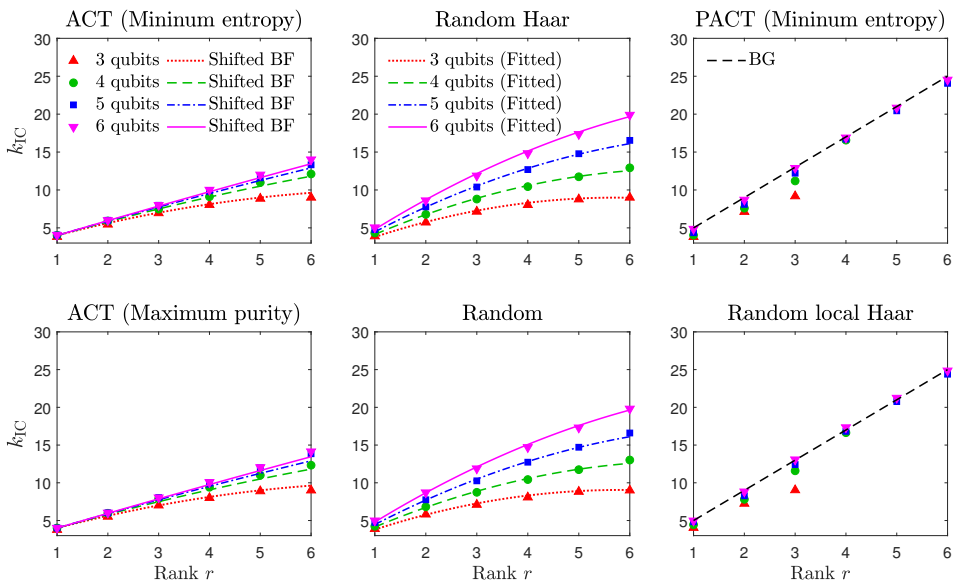


Figure 12: $k_{\mathrm{IC}}$ behavior comparison of various schemes for 3, 4, 5, and 6 qubits. For each number of qubits and rank, data markers are averaged over 100 random states from uniform distribution with respect to Hilbert–Schmidt measure.

exceeds the number of IC POM elements of BF by only 2 bases.

Secondly, comparing PACT and random local Haar with BG on the right of Fig. 12, we founded out that the $k_{\mathrm{IC}}$ behaviors of both these product schemes asymptotically approach that of BG $\frac{M_{BG}}{d} = 4r+1$.

Thirdly, the performances of two random bases measurement schemes in the middle of Fig. 12 (random Haar and random) show similar $k_{\mathrm{IC}}$ behaviors, which are well fitted *via* linear regression with expression $k_{\mathrm{IC}} = f(r)\log d + g(r)$, where $f(r)$ and $g(r)$ are respectively linear and quadratic polynomial in $r$.

Finally, from the comparison of (P)ACT with BF and BG, we propose numerical conjectures for the asymptotic $k_{\mathrm{IC}}$ behaviors of ACT and PACT in the large $d$ limit $d \gg r$ to be

$$\begin{aligned} k_{\mathrm{IC}} &\to 2r+2 \text{ for ACT} \\ k_{\mathrm{IC}} &\to 4r+1 \text{ for PACT.} \end{aligned} \tag{3.1}$$

These aymptotic relations are more clearly confirmed in Fig. 13. The $k_{\mathrm{IC}}$ plot for 7 qubits ($d = 128$) not only validates the numerical conjecture for (P)ACT with a higher confidence, but also shows the discrepancy between PACT and random local Haar in the large $d$ limit, which affirms the optimality of the adaptive scheme adjusted to product-basis measurements.

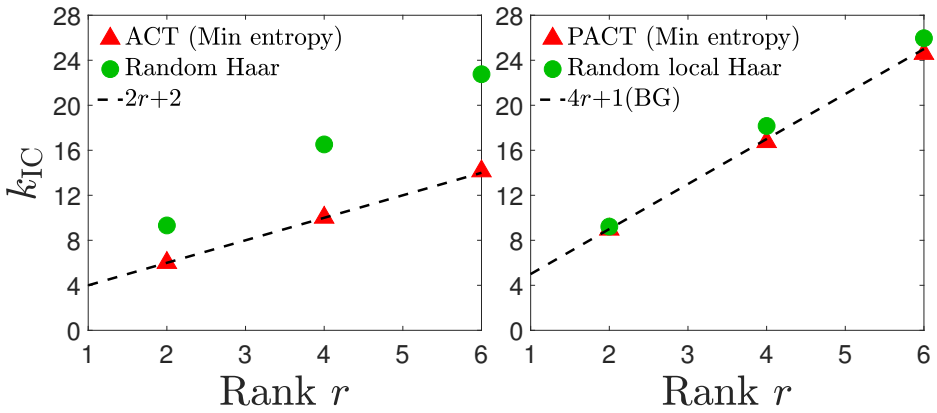


Figure 13: $k_{\mathrm{IC}}$ behavior plot for 7 qubits. For each number of qubits and rank, data markers are averaged over 50 random states from uniform distribution with respect to Hilbert–Schmidt measure.

### 3.3.4 $\widehat{\rho}_k$ evolution of ACT

In the previous subsection, we observed that ACT shows the highest compression efficiency with the lowest IC. Figure 14 describes the evolution of $\widehat{\rho}_k$ from the simulation results as ACT progresses with increasing $k$. It shows that the fidelity between $\widehat{\rho}_k$ and $\rho_r$, rank$\{\widehat{\rho}_k\}$, and $S(\widehat{\rho}_k)$ monotonically increase and ultimately reach respectively the values of $\rho_r$. In particular, sanity checks from Fig. 14 verify the obvious fact that $\widehat{\rho}_k \to \rho_r$ at $k = k_{\mathrm{IC}} - 1$. Reminding ourselves that once eigenbasis of $\rho_r$ is included in the measurement bases, $k_0 = \lceil (r^2 - r)/(d-1) \rceil + 1$ linearly independent measurement bases is sufficient to a unique $\widehat{\rho} = \rho_r$ for noiseless data, it is the *adaptive strategy* that gives the lowest $k_{\mathrm{IC}}$ by providing the fastest convergence of $\widehat{\rho}_k$ to $\rho_r$.

In fact, the adaptive strategy of entropy minimization under the positivity

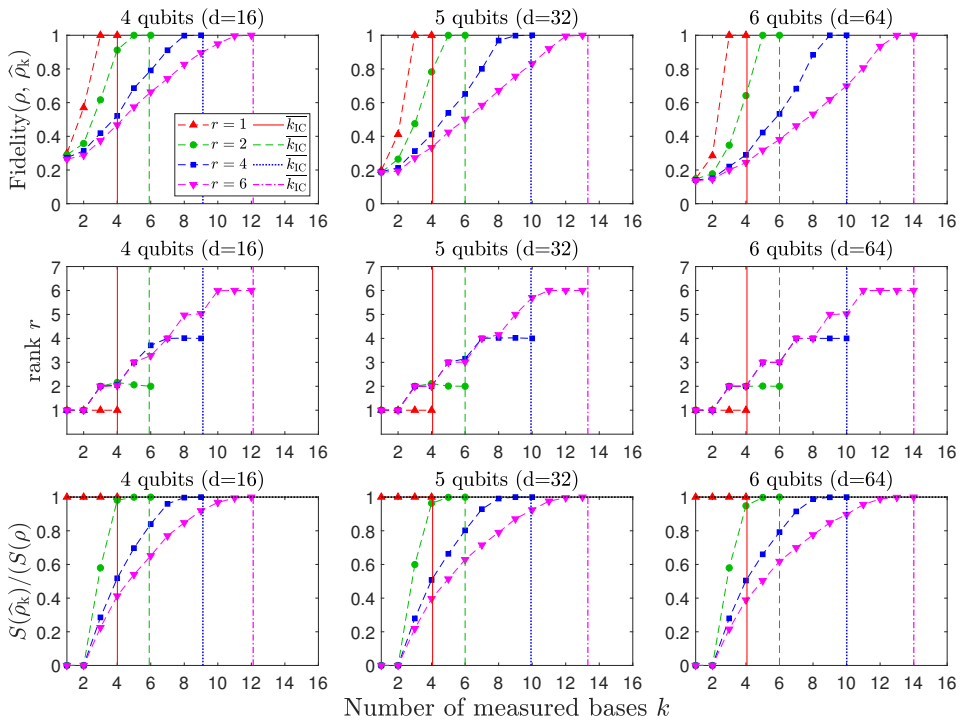


Figure 14: Noiseless evolution of $\widehat{\rho}_k$ with respect to the number of the iterative step $k$. For each number of qubits and rank, data markers are averaged over 100 random states from uniform distribution with respect to Hilbert–Schmidt measure.

constraint, which gives the highest converging speed of $\widehat{\rho}_k$ among all tested schemes, is a variation of the conventional CS procedure where its convex optimization routine is replaced with a non-convex one. This implies that the CS procedure, even without knowing the exact prior information and whether the measurement map satisfy corresponding RIP condition, may be utilized to perform compressive tomography assisted by the ICC procedure that directly verifies whether the data are IC.

### 3.3.5 Hybrid compressive tomography

In spite of the substantial reduction in number of IC measurement $M \ll d^2$, to perform ACT in practice, both the ICC and adaptive procedures are time-consuming at each iteration step of ACT. On the other hand, a fully random measurement scheme picks bases randomly at each step, which takes negligible time compared to that of ACT. To overcome this inefficiency of ACT in time, we establish a hybrid compressive tomography (HCT) scheme that combines the good qualities of both random and adaptive measurements. HCT starts with random bases measurement scheme that proceeds with high speed. After gaining a certain amount of information about $\rho_r$, the scheme becomes adaptive and makes use of the acquired information to adaptively choose the next measurement basis. We may set the transition point, for instance, using some prechosen threshold value $s_{\text{th}}$ of $s_{\text{CVX},k}$, where a lower threshold value leads to a greater gain of information but less effort on the adaptation. The explicit HCT procedure is detailed below.

**HCT procedure**

Beginning with $k = 1$, a random computational basis $\mathscr{B}_1$ and some positive threshold value $s_{\text{th}}$:

1. Measure $\mathscr{B}_k$ and collect the relative frequency data $\sum_{j'=0}^{d-1} \nu_{j'k} = 1$.

2. From $\{\nu_{0k'}, \ldots, \nu_{d-1\,k'}\}_{k'=1}^{k}$, obtain $kd$ physical ML probabilities.

3. Perform ICC with the ML probabilities and compute $s_{\text{CVX},k}$:

   - **If** $s_{\text{CVX},k} < \varepsilon$, terminate ACT and take $\rho_{\max} \approx \rho_{\min}$ as the estimator and report $s_{\text{CVX},k}$.
   - **Else** Proceed.

4. If $s_{\text{CVX},k} > s_{\text{th}}$:

   - Assign a random basis to $\mathscr{B}_{k+1}$, perhaps from appendix A or B.

   Else:

   - Assign an optimal basis obtained using the adaptive strategy in (P)ACT to $\mathscr{B}_{k+1}$.

5. Set $k = k + 1$ and repeat.

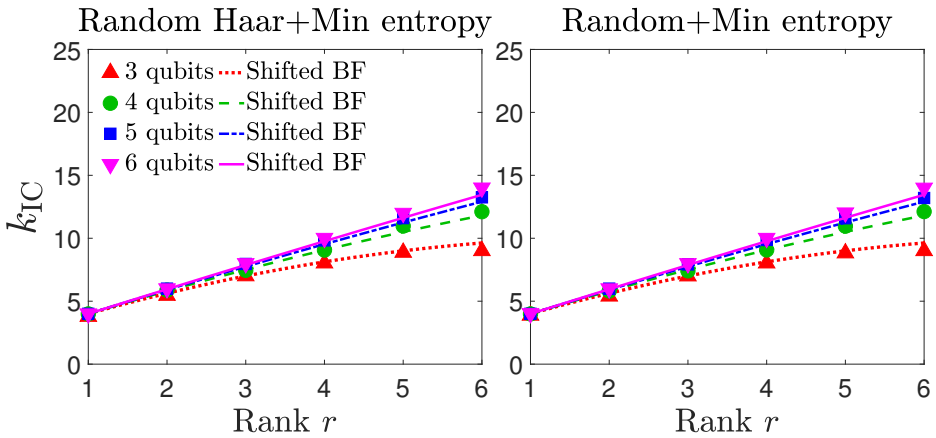


Figure 15: $k_{\mathrm{IC}}$ behavior of two different type of HCT for 3, 4, 5, and 6 qubits. For each number of qubits and rank, data markers are averaged over 100 random states from uniform distribution with respect to Hilbert–Schmidt measure.

Figure 15 shows $k_{\mathrm{IC}}$ behaviors of two types of HCT schemes that start with two types of random bases measurements mentioned previously. From the figure, the $k_{\mathrm{IC}}$ behaviors of the HCT schemes turn out to be almost identical to that of ACT. From this observation, we may argue that the amount of information gain from ACT and HCT until the transition point of HCT are the same. This further implies that without any given *a priori* information, effective adapation of the measurement is valid only after gaining a sufficient amount of information, so that adaptive methods are then worth the effort.

# Chapter 4

# Conclusion

We have constructed a novel and feasible *adaptive compressive tomography* scheme to reconstruct arbitrary low-rank quantum states. This scheme consists of an iteration of measurements of either entangled or product bases that are adaptively chosen. Unlike previous compressive schemes, our scheme requires no *a priori* information as if only makes use of *a posteriori* information of accumulated measured data for deciding the next measurement basis, and the informational completeness of the measured data is self-consistently certified with semi-definite programming.

In the real experiment and simulation, entropy minimization was taken as the de facto adaptive strategy with respect to the maximum likelihood estimation. This is shown to give a reliable compressive effect on the number of measured informationally complete bases for both noiseless simulations and single-photon orbital angular momentum experiments on 4 qubit system, which confirms adaptive tomography's superior compressive efficiency to the popular random Pauli-projective measurement. This confirms that the introduced measurement scheme is valid in the presence of noise in the data.

More noiseless simulations of adaptive compressive tomography were performed for various multi-qubit systems to show that our adaptive schemes

possess higher efficiency in terms of the compression of informational complete data compared to other known random compressive measurements: namely the random Pauli-projective measurement, random Haar uniform bases measurement and random full rank states' eigenbases measurement. Particularly, the averaged number of informationally complete bases $k_{\mathrm{IC}}$ for the entangled-bases adaptive scheme shows good agreement with the number of measurement configurations of the Baldwin-Flammia measurement shifted by $2d$ for all tested ranks $r$ and dimensions $d$, which leads to an averaged $k_{\mathrm{IC}} \to 2r+2$ in the large $d$ limit $d \gg r$. The averaged $k_{\mathrm{IC}}$ for the product-bases adaptive scheme is always lower than and asymptotically approaches that of the Baldwin-Goyeneche measurement with $k_{\mathrm{IC}} \to 4r+1$ when $d \gg r$.

Finally, we suggested a hybrid measurement scheme that starts with picking random measurement bases, later turns into an adaptive scheme at some transition point. This gives almost the same averaged $k_{\mathrm{IC}}$ as a fully adaptive scheme. Thus, the hybrid scheme may be used to speed up the compressive tomography process.

# Appendix A

# Construction of random Haar-uniform bases

It is well-known that the QR decomposition generates unitary operators distributed according to the Haar measure [31], so that the following routine applies:

---

**Constructing a random Haar-uniform basis**

Starting from a reference basis $\mathscr{B}_{\text{ref}} = \{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$:

1. Generate a random $d \times d$ matrix $A$ with entries i.i.d. standard Gaussian distribution.
2. Compute $Q$ and $R$ from the QR decomposition $A = QR$.
3. Define $R_{\text{diag}} = \text{diag}\{R\}$.
4. Define $L = R_{\text{diag}} \oslash |R_{\text{diag}}|$ ($\oslash$ refers to the Hadamard division).
5. Define $U_{\text{Haar}} = QL$.

Thereafter construct the new basis $\mathscr{B}_{\text{Haar}} = U_{\text{Haar}}\mathscr{B}_{\text{ref}}$.

---

Steps 3–5 enforce a QR decomposition procedure that produces an effective

"$R$" matrix that has positive diagonal entries, which is the correct decomposition procedure we need to generate $U_{\text{Haar}}$.

# Appendix B

# Construction of random states

The well-known pseudocode that generates a random rank-$r$ state $\rho_r$ from uniform distribution with respect to Hilbert–Schmidt measure [30] is as follows:

---

**Constructing a random rank-$r$ state**

Starting from a reference basis $\mathscr{B}_{\text{ref}} = \{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$:

1. Generate a random $d \times r$ matrix $A$ with entries i.i.d. standard Gaussian distribution.
2. Obtain $\rho_r = \dfrac{A^\dagger A}{\text{tr}\{A^\dagger A\}}$.

---

This pseudocode is used for both generation of random rank-$r$ target states as well as realization of random full rank states' eigenbases measurement. For the realization of the measurement, random full rank state $\rho_d$ is first generated evaluating $r = d$ in the code, and eigenbasis of $\rho_d$ is taken.

# Bibliography


[1] C. Schwemmer, L. Knips, D. Richart, H. Weinfurter, T. Moroder, M. Kleinmann, and O. Gühne, "Systematic Errors in Current Quantum State Tomography Tools," Phys. Rev. Lett. **114**, 080403 (2015).

[2] Z. Hradil, "Quantum-state estimation," Phys. Rev. A **55**, R1561 (1997).

[3] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky, "Diluted maximum-likelihood algorithm for quantum tomography," Phys. Rev. A **75**, 042108 (2007).

[4] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil, "Quantum-State Reconstruction by Maximizing Likelihood and Entropy," Phys. Rev. Lett. **107**, 020404 (2011).

[5] Y. S. Teo, *Introduction to Quantum-State Estimation* (World Scientific Publishing Co., Singapore, 2015).

[6] J. Shang, Z. Zhang, and H. K. Ng, "Superfast maximum-likelihood reconstruction for quantum tomography," Phys. Rev. A **95**, 062336 (2017).

[7] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**, 1289 (2006)

[8] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," IEEE Trans. Inf. Theory **52**, 5406 (2006).

[9] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," Found. Comput. Math. **9**, 717 (2009).

[10] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," IEEE Access **6**, 4875 (2019).

[11] E. J. Candès and T. Tao, "Decoding by linear programming," IEEE Trans. Inf. Th. **51**, 4203 (2005).

[12] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Communications on Pure and Applied Mathematics **LIX**, 1207 (2006).

[13] L. Vandenberghe and S. Boyd, "Semidefinite programming," SIAM Review **38(1)**, 49-95, March (1996).

[14] A. Kalev, R. L. Kosut, and I. H. Deutsch, "Quantum tomography protocols with positivity are compressed sensing protocols," npj Quantum Inf. **1**, 15018 (2015).

[15] M. Fazel, "Matrix rank minimization with applications," Ph.D. thesis, Stanford University (2002).

[16] B. Recht, M. Fazel, P. Parrilo, "Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization," SIAM Review **52(3)**, 471-501 (2010).

[17] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert, "Quantum State Tomography via Compressed Sensing," Phys. Rev. Lett. **105**, 150401 (2010).

[18] A. Steffens, C. A. Riofrío, W. McCutcheon, I. Roth, B. A. Bell, A. McMillan, M. S. Tame, J. G. Rarity, and J. Eisert, "Experimentally exploring compressed sensing quantum tomography," Quantum Sci. Technol. **2**, 025005 (2017).

[19] C. A. Riofrío, D. Gross, S. T. Flammia, T. Monz, D. Nigg, R. Blatt, and J. Eisert, "Experimental quantum compressed sensing for a seven-qubit system," Nat. Commun. **8**, 15305 (2017).

[20] Charles H. Baldwin, Ivan H. Deutsch, and Amir Kalev, "Strictly-complete measurements for bounded-rank quantum-state tomography," Phys. Rev. A **93**, 052105 (2016).

[21] S. T. Flammia, A. Silberfarb, and C. M. Caves, "Minimal Informationally Complete Measurements for Pure States," Found. Phys. **35**, 1985 (2005).

[22] D. Goyeneche, G. Cañas, S. Etcheverry, E.S. Gómez, G.B. Xavier, G. Lima, and A. Delgado, "Five Measurement Bases Determine Pure Quantum States on Any Dimension," Phys. Rev. Lett. **115**, 090401 (2015).

[23] L. Vandenberghe and S. Boyd, "Semidefinite programming," SIAM Review **38**, 49 (1996).

[24] A. S. Bandeira, E. Dobriban, D. G. Mixon, and W. F. Sawin, "Certifying the restricted isometry property is hard," IEEE Trans. Inf. Theory **59**, 3448 (2013).

[25] D. N. Tran, S. Huang, S. P. Chin, and T. D. Tran, "Low-rank matrices recovery via entropy function," IEEE (ICASSP 2016), 4064 (2016).

[26] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," `http://cvxr.com/cvx` (2014).

[27] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag Limited, 2008) pp. 95–110.

[28] E. Bolduc, N. Bent, E. Santamato, E. Karimi, and R. W. Boyd, "Exact solution to simultaneous intensity and phase encryption with a single phase-only hologram," Opt. Lett. **38**, 3546 (2013).

[29] F. Bouchard, N. H. Valencia, F. Brandt, R. Fickler, M. Huber, and M. Malik, "Measuring azimuthal and radial modes of photons," Opt. Express **26**, 31925 (2018).

[30] K. Życzkowski and H.-J. Sommers, "Hilbert–Schmidt volume of the set of mixed quantum states," J. Phys. A: Math. Gen. **36**, 10115 (2003).

[31] F. Mezzadri, "How to generate random matrices from the classical compact groups," Not. AMS **54**, 592 (2007).

# 국문초록

양자 정보 과정에서 양자 상태 토모그래피는 양자 상태를 준비하고 검증하는데에 폭넓게 활용된다. 이 때 양자 측정이 수행되는데 필요한 측정 구성의 수는 양자 계의 차원의 다항함수에 비례하여 증가한다. 본 학위 논문에서는 압축센싱 이론에 영감을 받아 현저하게 줄어든 측정 구성의 수를 이용하여 임의의 계수 결핍된 양자 상태를 복원하는 적응 압축 토모그래피 양식을 제안한다. 제안한 양식은 현존하는 압축 측정과 압축센싱 이론에 필수적인 양자 계의 차원을 제외한 양자 상태에 관한 사전 정보를 필요로 하지 않는다.

우리는 이 적응 양식을 구성하는 두 가지 주된 요소를 제시한다. 하나는 준완전 프로그래밍을 적용한 정보적 완비성 인증, 나머지는 엔트로피 최소화 과정의 적응 전략이다. 우리는 4 큐비트 계에 대해 무잡음 시뮬레이션과 단일 광자 각운동량 실험을 수행하여 무작위 파울리 정사영 측정 양식과 정보적으로 완비된 데이터의 압축 효율성을 비교하였다. 이 비교로부터, 우리는 적응 압축 토모그래피 양식이 무작위 파울리 정사영 측정 양식을 능가한다는 사실을 발견하였다. 또한, 우리는 시뮬레인션과 실제 실험 결과를 비교하여 제안한 적응 양식이 실제 실험에 존재하는 잡음에 대해 견고함을 확인하였다.

언급한 큐비트 수 이외에도 우리는 무잡음 시뮬레이션에 여러 개의 더 큰 큐비트 수도 고려하였다. 그리하여 우리는 수치적으로 작은 계수의 양자 상태들에 대해 결맞은 기저와 곱 기저 적응 측정 양식 모두 무작위 기저 측

정 양식보다 높은 압축 효율성을 보이고 최근에 설계된 요소 탐지 압축 측정 방식과 비슷한 압축 양상을 보임을 확인하였다. 더욱이, 우리는 앙자 계의 차원이 큰 극한에서, 차원 의존성을 잃어버리는 점근 압축 규모 양상에 관한 수치적인 추측값 또한 제시하였다. 마지막으로 제안한 적응 양식을 변형하여, 측정 양식의 계산 속도가 빨라지도록 무작위 기저 측정에서 시작하여 적응 기저 측정으로 전환하는 혼종 토모그래피 양식을 설립하였다.

**주요어 :** 압축 토모그래피, 사전 정보, 정보적 완비성 인증, 적응 전략

**학번 :** 2012-20368