Sample Efficient Identity Testing and Independence Testing of Quantum States

Nengkun Yu
Centre for Quantum Software and Information, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia
nengkunyu@gmail.com

Abstract

In this paper, we study the quantum identity testing problem, i.e., testing whether two given quantum states are identical, and quantum independence testing problem, i.e., testing whether a given multipartite quantum state is in tensor product form. For the quantum identity testing problem of $\mathcal{D}(\mathbb{C}^d)$ system, we provide a deterministic measurement scheme that uses $\mathcal{O}(\frac{d^2}{\epsilon^2})$ copies via independent measurements with d being the dimension of the state and ϵ being the additive error. For the independence testing problem $\mathcal{D}(\mathbb{C}_1^d \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_m})$ system, we show that the sample complexity is $\tilde{\Theta}(\frac{\Pi_{i=1}^m d_i}{\epsilon^2})$ via collective measurements, and $\mathcal{O}(\frac{\Pi_{i=1}^m d_i^2}{\epsilon^2})$ via independent measurements. If randomized choice of independent measurements are allowed, the sample complexity is $\Theta(\frac{d^{3/2}}{\epsilon^2})$ for the quantum identity testing problem, and $\tilde{\Theta}(\frac{\Pi_{i=1}^m d_i^{3/2}}{\epsilon^2})$ for the quantum independence testing problem.

2012 ACM Subject Classification Theory of computation → Quantum computation theory; Mathematics of computing → Probability and statistics

Keywords and phrases Quantum property testing

Digital Object Identifier 10.4230/LIPIcs.ITCS.2021.11

Related Version A full version of the paper is available at `https://arxiv.org/abs/1904.03218`.

Funding *Nengkun Yu*: N. Y. is supported by ARC Discovery Early Career Researcher Award DE180100156 and ARC Discovery Program DP210102449.

Acknowledgements We thank Youming Qiao for his helpful comments on the previous version of this manuscript. We thank Tongyang Li for pointing out relevant reference [38]. We thank Ryan O'Donnell and John Wright for telling us the Sanov's theorem and its relation to [44]. We are grateful for the reviewer to point out that using random independent measurement can provide a tight bound for quantum state certification.

1 Introduction

1.1 Classical Property Testing

The ability to test whether an unknown object satisfies a hypothetical model based on observed data plays a particularly important role in science [50]. Initially proposed by Rubinfeld and Sudan [60, 61] to test algebraic properties of polynomials, the concept of property testing has been extended to many objects of computer science: graphs, Boolean functions, and so on [41, 40]. Property testing and distribution testing are intricately connected. At the beginning of this century, Batu *et al*. introduced the problem of testing properties associated with discrete probability distributions [14, 15]. In other words, how many samples from a collection of probability distributions are needed to determine whether those distributions satisfy a particular property with high confidence? Over the past two decades, this area has become an extremely well-studied and successful branch of property testing due in part to

the ongoing data science revolution. Never have computationally-efficient algorithms, a.k.a, testers, that can identify and/or classify properties using as few samples as possible been in higher demand.

A direct approach to distribution property testing is to reconstruct the given distributions from sufficiently many samples. It is well known that, after taking $\Theta(\epsilon^{-2} \cdot d)$ samples from a d-dimensional probability distribution p, the empirical distribution is, with high probability, ϵ-close to p in total variance distance [29, pages 10 and 31]. Therefore, finding algorithms that use $o(d)$ samples for testing problems is highly desirable. Surprisingly, algorithms using less number of samples than $\Theta(\epsilon^{-2} \cdot d)$ exist for many important properties.

The equality, or identity, of distributions is a central problem in this branch of study, and one that is frequently revisited with different approaches due to its importance. In an important work [40], Goldreich and Ron found that the ℓ_2 norm can be estimated from $O(\epsilon^{-2} \cdot \sqrt{d})$ samples. This led to an algorithm for uniformity testing, *i.e.*, to determine whether a probability distribution is a uniform using $O(\epsilon^{-4} \cdot \sqrt{d})$ samples. Paninski [58] and Valiant and Valiant [63] showed that the complexity of uniformity is $\Theta(\epsilon^{-2} \cdot \sqrt{d})$. If one distribution is an arbitrary known distribution, Batu *et al.* [14, 13] presented an ℓ_2-identity tester and used it to build an ℓ_1 estimator using $O(\epsilon^{-2} \cdot \sqrt{d} \log d)$ samples; later in [65], Valiant and Valiant showed the sample complexity of this problem is $\Theta(\epsilon^{-2} \cdot \sqrt{d})$. If both distributions are unknown, Batu *et al.* provided a tester in [14] using $O(\epsilon^{-8/3} \cdot d^{2/3} \log d)$ samples; In 2014, Chan *et al.*, in [23], showed the complexity of the identity testing is $\Theta(\max(\epsilon^{-2} \cdot \sqrt{d}, \epsilon^{-4/3} \cdot d^{2/3}))$.

The idea of identity testing has been extensively explored in studying other property testing problems. Independence testing and conditional independence testing are among the most important ones. In [13], Batu *et al* presented an independence tester for bipartite independence testing over $[d_1] \times [d_2]$ with a sample complexity of $\tilde{O}(d_1^{2/3} d_2^{1/3}) \cdot \mathrm{Poly}(\epsilon^{-1})$, for $d_1 \geq d_2$. Levi, Ron and Rubinfeld in [51] showed a lower bound $\Omega(\sqrt{d_1 d_2})$ for all $d_1 \geq d_2$ and $\Omega(d_1^{2/3} d_2^{1/3})$ for $d_1 = \Omega(d_2 \log d_2)$. Acharya *et al.* [6] introduced a tester for multipartite independence testing over $\times_{j=1}^{m} [d_j]$ with sample complexity $O(\epsilon^{-2} \cdot \sqrt{\Pi_{j=1}^{m} d_j} + \epsilon^{-2} \cdot \sum_{j=1}^{m} d_j)$. In their important work [31], Diakonikolas and Kane demonstrated a unified approach to resolve the sample complexity of a wide variety of testing problems based on their alternative proof for identity testing. In particular, they showed that the sample complexity of independence tesing is $\Theta(\max_k \{\epsilon^{-2} \sqrt{\Pi_{j=1}^{m} d_j}, \epsilon^{-4/3} \cdot d_k^{2/3} \Pi_{j=1}^{m} d_j^{1/3}\})$. Canonne *et al.* [22] initiated the study of the conditional independence within property testing framework. Notably, for the very important $[2] \times [2] \times [n]$, they showed that the sample complexity for this problem is $\Theta(\max\{\epsilon^{-2} \cdot \sqrt{n}, \min\{\epsilon^{-1} \cdot n^{7/8}, \epsilon^{-8/7} \cdot n^{6/7}\}\})$.

Besides the mentioned works, a very incomplete list of works of distributional property testing includes [12, 15, 68, 5, 64, 51, 46, 26, 32, 48, 66, 70, 30, 67, 28, 39, 33, 7, 21, 27, 34, 24], and two excellent surveys include more [59, 20].

1.2 Quantum Property Testing

Quantum property testing has been extensively studied. At this stage of development of quantum computation, testing the properties of new devices as they are built is a basic problem as illustrated in Montanaro and de Wolf's comprehensive survey [53]. A standard quantum device outputs some known d-dimensional (mixed) state $\sigma \in \mathcal{D}(\mathbb{C}^d)$ but inevitably, the results are noisy such that the actual output state $\rho \in \mathcal{D}(\mathbb{C}^d)$ is not equal σ, maybe not even close to. Similar to property testing with classical distributions, properties of ρ need to be verified by accessing the device, say, m times, to derive $\rho^{\otimes m}$.

Starting from the very basic problem of quantum state tomography, a fundamental problem is to decide how many copies of an unknown mixed quantum state $\rho \in \mathcal{D}(\mathbb{C}^d)$ is necessary and sufficient to output a good approximation of ρ in trace distance, with high probability. This problem has been studied extensively since the birth of quantum information theory. The main-stream approach is through independent measurement, i.e., measurement on each copy of the state. A sequence of work [42, 36, 69, 49] is dedicated to showing that $O(\epsilon^{-2} \cdot d^3)$ copies are sufficient in an ℓ_1 distance of no more than ϵ. Haah *et al.* [44] showed that this is tight for independent measurement. For joint measurement, Haah *et al.* [44] proved that $O(\epsilon^{-1} \cdot d^2 \log(\epsilon^{-1} \cdot d))$ copies are sufficient to obtain an infidelity of no more than ϵ, which can be regarded as a quantum generalization of Sanov's theorem [62]. By combining the lower bound of [44] and upper bound of [56, 57], the sample complexity of state tomography with joint measurement is $\widetilde{\Theta}(\epsilon^{-2} \cdot d^2)$ in a ℓ_1 distance error of less than ϵ with high probability.

A more direct approach to quantum property testing is to estimate ρ by sampling from $\rho^{\otimes m}$, which also means one could check any property of interest. However, like classical property testing, this idea is not optimal for a general property. One problem that has received much attention is quantum identity testing. Suppose we are given query access to two states $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$, and we want to test whether they are equal or have a large ℓ_1 distance. For practical purposes, the results from cases where σ is a known pure state have been extensively studied, in the independent measurement setting [37, 25, 10]. [55] solved the problem, in the joint measurement setting, where σ is a maximally mixed state case by showing that $\Theta(\epsilon^{-2} \cdot d)$ copies are necessary and sufficient. Importantly, the sample complexity of the general problem was proven to be $\Theta(\epsilon^{-2} \cdot d)$ in [18] by providing an efficient ℓ_2 distance estimator between two unknown quantum states.

In [4], Aaronson initialized the study of the learnability of quantum state, whose goal is to output good estimations of a set of measurements simultaneously. In [1], Aaronson provided an efficient procedure of the quantum shadow tomography. A connection between quantum learning and differential privacy was established in [3]. In [2], the online learning of quantum states was studied.

Entanglement is a ubiquitous phenomenon in quantum information theory. A multipartite pure state $|\psi\rangle \in (\mathbb{C}^d)^{\otimes m}$ is not entangled if it can be written as $|\psi\rangle = \otimes_{j=1}^m |\psi_j\rangle$ for some $|\psi_j\rangle \in \mathbb{C}^d$. Pure entanglement testing was first discussed by Mintert *et al* [52]. Harrow and Montanaro [45] subsequently proved that $O(\epsilon^{-2})$ copies are sufficient and used that to study the quantum complexity theory. In [17], it was proved that $\Omega(d^2/\epsilon^2)$ copies are necessary to test separability of quantum states in $\mathbb{C}^d \otimes \mathbb{C}^d$ for not small ϵ.

Acharya *et al.* [8] estimated the von Neumann entropy of general quantum states. Gross *et al.* [43] showed that "stabilizerness" can be tested efficiently. One research direction is to study the potential speed-up of distributional property testing using quantum algorithms where the distribution is given in the form of a quantum oracle [16, 38].

1.3 Measurement Schemes

A significant difference between quantum property testing and classical property testing is the way the objects are sampled. In classical property testing, each sample is output with a classical index according to the probability distribution and given a fixed number of samples, the output string obeys the product probability distribution. However, with quantum property testing, the sampling methods have much richer structures. This difference together with others prevents the potential to design algorithms for quantum property testing from ingenious ideas and techniques of distribution testing.

Measurement	Complexity	Dimension	Implementation
Joint	Low	d^m	Hard even in the future
Independent	Medium	d	Available in the future

Among the many available sampling methods for quantum property testing (given a fixed number of copies, says k, of the states $\rho \in \mathcal{D}(\mathbb{C}^d)$), the two listed in Table 1 are of particular interests, i.e., joint measurement, and independent measurement. Joint measurement, the most general, allows arbitrary measurements of $\mathbb{C}^{d^m}$. Independent measurement only allows non-adaptive measurements on each copy of ρ, which results in, n measurements of $\mathbb{C}^d$.

Joint measurement has the potential to provide the optimal number of samples, but there are two caveats. "Optimal" joint measurement algorithms usually require an *exponential* number of copies of the quantum state to produce optimal results. They are also based on the assumption of noiseless, universal quantum computation on the *exponential* number of copies of the quantum state. For instance, the optimal tomography algorithms of k-qubit quantum state in [44, 56, 57] require a joint measurement on $\Theta(\epsilon^{-2} \cdot k2^{2k})$ qubits. Even in the future when quantum computers become a reality, implementing optimal joint measurement would be extremely hard given these conditions. General independent measurements are not feasible with currently-available technology. To implement a two-outcome measurement on the k-qubit system, one needs to implement a $k+1$ qubit unitary. Implementing a general $k+1$ qubit unitary requires a circuit consisting of at least $\Omega(4^k)$ elementary gates, which could also be hard.

In this paper, we study the quantum identity testing problem, i.e., testing whether two given quantum states are identical, and the quantum independence testing problem, i.e., testing whether a given multipartite quantum state is in tensor product form. For the quantum identity testing problem of $\mathcal{D}(\mathbb{C}^d)$ system, we provide a measurement scheme that uses $\mathcal{O}(\frac{d^2}{\epsilon^2})$ copies via independent measurements with d being the dimension of the state and ϵ being the additive error. For the independence testing problem $\mathcal{D}(\mathbb{C}_1^d \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_m})$ system, we show that the sample complexity is $\tilde{\Theta}(\frac{\Pi_{i=1}^m d_i}{\epsilon^2})$ via collective measurements, and $\mathcal{O}(\frac{\Pi_{i=1}^m d_i^2}{\epsilon^2})$ via independent measurements. Further, we initialize the study of the property testing problems of classical-quantum states, motivated by the potential applications of classical-quantum states. Our main tool is a measurement that "preserves" the ℓ_2 distance, which invokes an immediate connection between quantum and classical property testing.

1.4 Our contributions

Identify whether two quantum states are equal or not is called *quantum identity testing problem.*

▶ **Problem 1.** *Given two unknown quantum mixed states $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$, they satisfy either $\rho = \sigma$ or $||\rho - \sigma||_1 > \epsilon$ for a given $\epsilon > 0$. How many copies of ρ and σ are needed to distinguish these two cases, with high probability?*

This problem under joint measurement setting is solved in [18]. In this paper, we study this problem using independent measurement. To reach this goal, we observe the following lemma. It maintains interesting relations between the ℓ_2 distance of quantum states and the ℓ_2 distance of the generated corresponding probability distributions. Given that ℓ_2 distance plays a central role in classical property testing [31], our approach invokes an immediate connection between quantum and classical property testing. Previous research into quantum property testing has always been in isolation of classical property testing, whereas this scheme

opens up the potential to design quantum property tester from ingenious ideas and techniques of distribution testing. Further, this is a fixed measurement scheme that does not depend on the property to be tested, which makes our algorithms a perfect fit for implementation with current experiments.

▶ **Lemma 2.** *For d being power of 2, there is a measurement*

$$\mathcal{M} = (M_1, M_2, \ldots, M_{d(d+1)}) : \mathcal{D}(\mathbb{C}^d) \mapsto \Delta(d(d+1))$$

whose outcome lies in $\Delta(d(d+1))$, the $d(d+1)$-dimensional probability simplex, such that, for any quantum states $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$

$$||p - q||_2 = \frac{||\rho - \sigma||_2}{d+1}, \qquad ||p||_2, ||q||_2 \leq \frac{\sqrt{2}}{d+1}, \tag{1}$$

where $p = (p_1, p_2, \ldots, p_{d(d+1)})$ and $q = (q_1, q_2, \ldots, q_{d(d+1)})$ with $p_i = \mathrm{Tr}(\rho M_i)$ and $q_i = \mathrm{Tr}(\sigma M_i)$.

We employ mutually unbiased bases (MUB) to construct such measurement. MUB in Hilbert space $\mathbb{C}^d$ are two orthonormal bases $\{|e_1\rangle, \ldots, |e_d\rangle\}$ and $\{|f_1\rangle, \ldots, |f_d\rangle\}$ such that the square of the magnitude of the inner product between any basis states $|e_j\rangle$ and $|f_k\rangle$ equals the inverse of the dimension d. These bases are unbiased in the following sense: if a system is prepared in a state belonging to one of the bases, then all outcomes of the measurement with respect to the other basis are predicted to occur with equal probability.

For $d = 2^n$, there are $2^n + 1$ mutually unbiased bases in $\mathbb{C}^d$. Therefore, the density matrices of these MUBs form a linear basis of $\mathcal{D}(\mathbb{C}^d)$ in this case. Each measurement operator M_i is proportional to a density matrix of a MUB element. Therefore, after the measurement, there is no more information left because applying measurement in other MUB basis would output uniform distribution.

The upper bound of ℓ_2 norms of the output probability distribution is essential in designing an efficient quantum tester by lifting classical property tester because a small ℓ_2 norms ensures that the tester could use a smaller number of samples for the distributional identity testing problem as illustrated in [31], and distributional independence testing problem studied in [22].

Using Lemma 2 and the result of classical property testing, a tester using independent measurement for Problem 2 can be obtained as follows.

▶ **Theorem 3.** *For $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$, $O(\epsilon^{-2} \cdot d^2)$ copies are sufficient to distinguish via deterministic independent measurements, with at least a $\frac{2}{3}$ probability of success, the cases where $\rho = \sigma$ from the cases where $||\rho - \sigma||_1 > \epsilon$.*

This is better than directly using the SWAP test which uses $O(\frac{d^2}{\epsilon^4})$ copies, although the SWAP test is already a joint measurement.

Entanglement is a central feature in quantum information science. Certification of entanglement has received great amount of effort. This motivates us to study the following *quantum independence testing problem.*

▶ **Problem 4.** *Given an unknown quantum mixed states $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n})$, they satisfy either $\rho = \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_n$ for some $\sigma_i \in \mathcal{D}(\mathbb{C}^{d_i})$ or for all $\sigma_i \in \mathcal{D}(\mathbb{C}^{d_i})$, $||\rho - \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_n||_1 > \epsilon$ for a given $\epsilon > 0$. How many copies of ρ are needed to distinguish these two cases, with high probability?*

The above ℓ_1 identity testers for independent measurement together with the ℓ_1 identity tester of [18] for joint measurement enable us to derive the following result.

▶ **Theorem 5.** *The sample complexity of quantum independence testing problem for n-qubit quantum state is $\Theta(\epsilon^{-2} \cdot 2^n)$.*

For general n-partite system $\mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n})$ with $d_1 \geq d_2 \geq \cdots \geq d_n$, the sample complexity of quantum independence testing problem is $\tilde{\Theta}(\epsilon^{-2} \cdot \Pi_{i=1}^n d_i)$ with joint measurements; and $O(\epsilon^{-2} \cdot \Pi_{i=1}^n d_i^2)$ with deterministic independent measurements, where $\tilde{\Theta}$ hides a factor between $(\log^3 d_1 \cdot \log\log d_1)^{-1}$ and 1.

In n-qubit system, the lower bound of the quantum independence for joint measurement comes from a reduction from determining whether a given state is a maximally mixed state.

In a general system, the lower bound is derived using an additional technique called dimension splitting which regards a d_1-dimensional system as $\log d_1$ qubits system.

1.5 Other Results

It is widely believed that the fully-fledged quantum computer will be controlled through a classical system. Therefore, the data generated by quantum computers would be modeled by classical-quantum states, e.g., classical collections of quantum states. The importance of classical-quantum states also comes from its central role in studying quantum communication complexity [47, 9]. In classical property testing, Levi, Ron, and Rubinfeld initialized the study of property testing of collections of distributions in their pioneering work [51]. This motivates us to study the property testing problems of classical-quantum states.

In the query model, there are m states $\rho_1, \rho_2, \cdots, \rho_n$. We can choose $1 \leq i \leq n$ to obtain a copy of ρ_i. A motivation of studying this model is the quantum state preparation. Suppose there are different ways of generating a quantum state. We want to know whether these methods all work well. This problem can be formulated as *the independence testing of collections of quantum states.*

▶ **Problem 6.** *Given unknown quantum mixed states $\rho_1, \rho_2, \cdots, \rho_m \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n})$ and a given distribution $p = (p_1, p_2, \cdots, p_m)$, they satisfy either there exist $\sigma_{k,i} \in \mathcal{D}(\mathbb{C}^{d_k})$ for $1 \leq k \leq n$ such that for all $1 \leq i \leq m$ $\rho_i = \otimes_{k=1}^n \sigma_{k,i}$, or for any $\sigma_{k,i} \in \mathcal{D}(\mathbb{C}^{d_k})$, $\sum_{i=1}^m p_i ||\rho_i - \otimes_{k=1}^n \sigma_{k,i}||_1 > \epsilon$, for a given $\epsilon > 0$. How many queries are needed to distinguish these two cases, with high probability?*

Combing the framework in [31] and our independece testers, we obtain

▶ **Theorem 7.** *The sample complexity of the independence testing of collections of quantum states is $\tilde{\Theta}(\epsilon^{-2} \cdot d)$ with joint measurement; $O(\epsilon^{-2} \cdot d^2)$ with determinstic independent measurement.*

Like their classical counterparts, the complexity does not depend on the number of states n. Similarly, this idea can be used for the independence testing of collections of quantum states.

In further work, we explore the problem of testing conditional independence with classical-quantum-quantum states. This question naturally arises in studying distributed quantum computing. One typical example is environment assisted entanglement distribution. Suppose ρ_{ABC} is a tripartite state. We want to reach the goal of sharing a bipartite state σ_{AB}. C should perform a measurement on its system, now the state becomes classical-quantum-quantum.

▶ **Problem 8.** *Given an unknown classical-quantum-quantum mixed states*

$$\rho_{ABC} = \sum_{i=1}^{m} p_m \rho_{AB,i} \otimes |i\rangle\langle i| \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}) \otimes \Delta(C)$$

with $\Delta(C)$ being the probabilistic simplex of C and $|C| = m$ and $|i\rangle$ being a basis of C, we want to distinguish whether ρ_{ABC} is conditional independence, that is $\rho = \sum_{i=1}^{m} p_m \rho_{A,i} \otimes \rho_{B,i} \otimes |i\rangle\langle i|$, or for any conditional independent classical-quantum-quantum state σ_{ABC} $||\rho - \sigma||_1 > \epsilon$, for a given $\epsilon > 0$. How many queries are needed to distinguish these two cases, with high probability?

This problem is a generalization of the independence testing of collections of quantum states in the sense that the prior coefficient of the ℓ_1 distance is not given explicitly but may be approximated through sampling. One motivation for studying this problem is a simplified version of the conditional independence of general tripartite quantum states, which a fundamental concept in theoretical physics and quantum information theory.

More specifically, we modify the ℓ_2 estimator developed in [18] for joint measurement and develop a finer ℓ_2 estimator for independent measurement. Then we plug that estimator into the classical conditional independence testing framework developed in [22].

▶ **Theorem 9.** *For classical-quantum-quantum state $\rho_{ABC} \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}) \otimes \Delta(C)$, the sample complexity of testing whether A and B are conditionally independent given C is*

- $O(\max\{\frac{\sqrt{n}d_1 d_2}{\epsilon^2}, \min\{\frac{d_1^{\frac{4}{7}} d_2^{\frac{4}{7}} n^{\frac{6}{7}}}{\epsilon^{\frac{8}{7}}}, \frac{\sqrt{d_1 d_2} n^{\frac{7}{8}}}{\epsilon}\}\})$ *with joint measurement; and*
- $O(\max\{\frac{\sqrt{n}d_1^2 d_2^2}{\epsilon^2}, \min\{\frac{d_1^{\frac{6}{7}} d_2^{\frac{6}{7}} n^{\frac{6}{7}}}{\epsilon^{\frac{8}{7}}}, \frac{d_1^{\frac{3}{4}} d_2^{\frac{3}{4}} n^{\frac{7}{8}}}{\epsilon}\}\})$ *with independent measurement.*

1.6 Organization of this paper

Section 2 recalls the basic definitions of distance with discrete distributions and quantum states and presents some formal tools from earlier work that are used here. In Section 3, we state technical lemmata about the independence and conditional independence of quantum states. Section 4 demonstrates Lemma 2. Section 5 contains the results of identity testing and Theorems 3. In Section 6, we discuss the advantage of using random choice of independent measurements. Detail proofs of Lemmata, Theorem 7 and Theorem 9 can be found in the full version [71].

2 Preliminaries

This section begins with some standard notations and definitions used throughout the paper.

2.1 Basic facts for probability distributions

For $m \in \mathbb{N}$, $[m]$ denotes the set $\{1, \cdots, m\}$, and log denotes the binary logarithm. A probability distribution over discrete domain Ω is a function $p : \Omega \mapsto [0, 1]$ such that $\sum_{\omega \in \Omega} p(\omega) = 1$. $|\Omega|$ is the cardinality of set Ω. $\Delta(\Omega)$ denotes the set of probability distributions over Ω, i.e., the probability simplex of Ω. The marginal distributions $p_1 \in \Delta(A)$ and $p_2 \in \Delta(B)$ of a bipartite distribution $p_{1,2} \in \Delta(A \times B)$ can be defined as $p_1(a) = \sum_{b \in B} p_{1,2}(a, b)$, $p_1(b) = \sum_{a \in A} p_{1,2}(a, b)$. The product distribution $q_1 \otimes q_2$ of distributions $q_1 \in \Delta(A)$ and $q_2 \in \Delta(B)$ can be defined as $[q_1 \otimes q_2](a, b) = q_1(a)q_2(b)$, for every $(a, b) \in A \times B$.

The ℓ_1 distance between two distributions $p, q \in \Delta(\Omega)$ is $||p - q||_1 = \sum_{\omega \in \Omega} |p(\omega) - q(\omega)|$, and their ℓ_2 distance is $||p - q||_2 = \sqrt{\sum_{\omega \in \Omega} (p(\omega) - q(\omega))^2}$.

2.2 Basic quantum mechanics

An isolated physical system is associated with a Hilbert space, which is called the *state space*. A *pure state* of a quantum system is a normalized vector in its state space, and a *mixed state* is represented by a density operator on the state space. Here, a density operator ρ on d-dimensional Hilbert space $\mathbb{C}^d$ is a semi-definite positive linear operator such that $\mathrm{Tr}(\rho) = 1$. We let

$$\mathcal{D}(\mathbb{C}^d) = \{\rho : \rho \text{ is } d-\text{dimensional density operator of } \mathbb{C}^d\}$$

denote the set of quantum states.

The reduced quantum state of a bipartite mixed state $\rho_{1,2} \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$ on the second system is the density operators $\rho_2 := \mathrm{Tr}_1 \rho_{1,2} = \sum_i \langle i|A|i\rangle$, where $\{|i\rangle\}$ is the orthonormal basis of $\mathbb{C}^{d_1}$. The partial trace of $\rho_1 := \mathrm{Tr}_2 \rho_{1,2}$ can be similarly defined, note that partial trace functions are also independent of the selected orthonormal basis. This definition can be directly generalized into multipartite quantum states.

2.3 Quantum measurement

A positive-operator valued measure (POVM) is a measure whose values are non-negative self-adjoint operators in a Hilbert space $\mathbb{C}^d$, which is described by a collection of matrices $\{M_i\}$ with $M_i \geq 0$ and $\sum_i M_i = I_d$. If the state of a quantum system was ρ immediately before measurement $\{M_i\}$ was performed on it, then the probability of that result i recurring is $p(i) = \mathrm{Tr}(M_i\rho)$.

2.4 ℓ_1 distance

ℓ_1 distance is used to characterize the difference between quantum states. The ℓ_1 distance between ρ and σ is defined as $||\rho - \sigma||_1 \equiv \mathrm{Tr}|\rho - \sigma|$ where $|A| \equiv \sqrt{A^\dagger A}$ is the positive square root of $A^\dagger A$.

Given a general operator A, the ℓ_1 norm is defined as $||A||_1 = \mathrm{Tr}|A|$. And Lemma 10 always applies:

▶ **Lemma 10** ([54]). *The ℓ_1 distance is decreasing under partial trace. That is*

$$||\rho_1 - \sigma_1||_1, ||\rho_2 - \sigma_2||_1 \leq ||\rho_{1,2} - \sigma_{1,2}||_1.$$

Their ℓ_2 distance is defined as $||\rho - \sigma||_2 = \sqrt{\mathrm{Tr}(\rho - \sigma)^2}$. For $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$, we have the following relation between ℓ_1 and ℓ_2 distances, $||\rho - \sigma||_2 \leq ||\rho - \sigma||_1 \leq \sqrt{d}||\rho - \sigma||_2$. Given a subset $\mathcal{P} \subsetneq \mathcal{D}(\mathbb{C}^d)$, the ℓ_1 distance between ρ and $\mathcal{P}$ is defined as $||\rho - \mathcal{P}||_1 = \inf_{\sigma \in \mathcal{P}} ||\rho - \sigma||_1$. If $||\rho - \mathcal{P}||_1 > \epsilon$, we say that ρ is ϵ-far from $\mathcal{P}$; otherwise, it is ϵ-close.

2.5 Mutually unbiased bases

In quantum information theory, mutually unbiased bases (MUB) in d-dimensional Hilbert space are two orthonormal bases $\{|e_1\rangle, \ldots, |e_d\rangle\}$ and $\{|f_1\rangle, \ldots, |f_d\rangle\}$ such that the square of the magnitude of the inner product between any basis states $|e_j\rangle$ and $|f_k\rangle$ equals the inverse of the dimension d:

$$|\langle e_j|f_k\rangle|^2 = \frac{1}{d}, \quad \forall j, k \in \{1, \ldots, d\}.$$

These bases are unbiased in the following sense: if a system is prepared in a state belonging to one of the bases, then all outcomes of the measurement with respect to the other basis will occur with equal probability. It is known that, for $d = p^n$ with prime p, there exists $d + 1$ MUBs [35].

2.6 Quantum property testing

Let $\mathcal{D}(\mathbb{C}^d)$ denote the set of mixed states in Hilbert space $\mathbb{C}^d$, and let a known $\mathcal{T} \subset \mathcal{D}(\mathbb{C}^d)$ be the working domain of the quantum states. In a standard of property testing scenario, a testing algorithm for a property $\mathcal{P} \subset \mathcal{T}$ would be an algorithm that, when granted access to independent samples from an unknown quantum state $\rho \in \mathcal{T}$ as well as an ℓ_1 distance parameter of $0 < \epsilon \leq 1$, outputs either "Yes" or "No", with the following guarantees:

- If $\rho \in \mathcal{P}$, then it outputs "Yes" with a probability of at least $\frac{2}{3}$.
- If ρ is ϵ-far from $\mathcal{P}$, then it outputs "No" with a probability of at least $\frac{2}{3}$.

Our interest is in designing computational efficient algorithms with the smallest sample complexity (i.e., the smallest number of samples drawn of ρ.).

Confidence of $\frac{2}{3}$ is not essential here, it could be replaced by any constant greater than $\frac{1}{2}$. This would only change the sample complexity by a multiplicative constant. According to the Chernoff bound, the probability of success becomes $1 - 2^{-\Omega(k)}$, after repeating the algorithm k times.

2.7 Tools from earlier work

The following results were established in earlier work, and are used within this paper.

▶ **Theorem 11** ([11]). *The Pauli group $\mathcal{P}_k = \{I, X, Y, Z\}^{\otimes n}$ of order 4^n can be divided into $2^n + 1$ Abelian subgroups with an order of 2^n, say, $G_0, \ldots, G_{2^n}$ such that $G_i \bigcap G_j = \{I_2^{\otimes n}\}$ for $i \neq j$. Each subgroup can be simultaneously diagonalizable by a corresponding basis. All these $2^n + 1$ bases form $2^n + 1$ MUBs.*

▶ **Theorem 12** ([55, 18]). *$100\frac{d}{\epsilon^2}$ copies are sufficient and $0.15\frac{d}{\epsilon^2}$ copies are necessary to test whether $\rho \in \mathcal{D}(\mathbb{C}^d)$ is the maximally mixed state $\frac{I_d}{d}$ or $||\rho - \frac{I_d}{d}||_1 > \epsilon$ with at least a $2/3$ probability of success. Generally, $O(\frac{d}{\epsilon^2})$ copies of ρ and σ are sufficient to test whether $\rho = \sigma$ or $||\rho - \sigma||_1 > \epsilon$*

Algorithm 1 A Mixness Test.

Input: $100\frac{d}{\epsilon^2}$ copies of $\rho \in \mathcal{D}(\mathbb{C}^d)$
Output: "Yes" with a probability of at least $\frac{2}{3}$ if $\rho = \frac{I_d}{d}$; and "No" with a probability of at least $\frac{2}{3}$ if $||\rho - \frac{I_d}{d}||_1 > \epsilon$.

Algorithm 2 A Identity Test with Joint Measurement.

Input: $O(\frac{d}{\epsilon^2})$ copies of $\rho \in \mathcal{D}(\mathbb{C}^d)$ and $O(\frac{d}{\epsilon^2})$ copies of $\sigma \in \mathcal{D}(\mathbb{C}^d)$
Output: "Yes" with a probability of at least $\frac{2}{3}$ if $\rho = \sigma$; and "No" with a probability of at least $\frac{2}{3}$ if $||\rho - \sigma||_1 > \epsilon$.

▶ **Theorem 13** ([23]). *For n-dimensional probability distributions of p and q, $O(\frac{b}{\epsilon^2})$ samples are sufficient to distinguish, with at least a $\frac{2}{3}$ probability, the cases where $p = q$ from the cases where $||p - q||_2 > \epsilon$, where $b \geq ||p||_2, ||q||_2$.*

Algorithm 3 An ℓ_2 norm Identity Test.

Input: $O(\frac{b}{\epsilon^2})$ copies of p and $O(\frac{b}{\epsilon^2})$ copies of q
Output: "Yes" with probability at least $\frac{2}{3}$ if $p = q$, "No" with probability at least $\frac{2}{3}$ if $||p - q||_2 > \epsilon$.

3 Quantum Independence and Technical Lemmata

3.1 Bipartite independence and approximate independence

We say that $\rho_{1,2} \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^{d_2})$ is independent if $\rho_{1,2} = \sigma_1 \otimes \sigma_2$ for some $\sigma_i \in \mathcal{D}(\mathbb{C}^{d_i})$. One can directly verify that, if $\rho_{1,2}$ is independent, then $\rho = \rho_1 \otimes \rho_2$ with ρ_1 and ρ_2 being the reduced density matrices of $\rho_{1,2}$.

We say that ρ is ϵ-independent with respect to the ℓ_1 distance if there is an independent state σ such that $||\rho - \sigma||_1 \leq \epsilon$. We say that ρ is ϵ-far from being independent with respect to the ℓ_1 distance if $||\rho - \sigma||_1 > \epsilon$ for any independent state σ.

▶ **Proposition 14.** *Let ρ and σ be bipartite states of $\mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^{d_2})$. If $||\rho - \sigma||_1 \leq \epsilon/3$ and σ is independent, then $||\rho - \rho_1 \otimes \rho_2||_1 \leq \epsilon$.*

▶ **Lemma 15.** $||\rho_1 \otimes \rho_2 - \sigma_1 \otimes \sigma_2||_1 \leq ||\rho_1 - \sigma_1||_1 + ||\rho_2 - \sigma_2||_1$.

3.2 Multipartite independence and approximate independence

We say that $\rho \in \mathcal{D}(\mathbb{C}^d \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n})$ is n-partite independent if $\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$, and that ρ is ϵ-independent with respect to the ℓ_1 distance if there is a state σ that is m-partite independent and $||\rho - \sigma||_1 \leq \epsilon$. We say that ρ is ϵ-far from being independent with respect to the ℓ_1 distance if $||\rho - \sigma||_1 > \epsilon$ for any m-partite independent state σ .

▶ **Proposition 16.** *Let ρ and σ be n-partite states, if $||\rho - \sigma||_1 \leq \epsilon$, and σ is m-partite independent, then $||\rho - \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n||_1 \leq (n+1)\epsilon$.*

▶ **Lemma 17.** $||\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n - \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_n||_1 \leq \sum_{i=1}^n ||\rho_i - \sigma_i||_1$.

Proposition 18 establishes a connection between bipartite independence and multipartite independence. Specifically, it shows that if an n-partite state is close to bipartite independence in any 1 versus $n-1$ cut, it is close to being n partite independent.

▶ **Proposition 18.** *Let ρ be an n-partite states. If for any $1 \leq i \leq n$, there exists a state $\sigma_i^{(i)}$ of party i, and a state $\psi_{[n]\setminus\{i\}}$ of parties $[n] \setminus \{i\}$ such that $||\rho - \sigma_i^{(i)} \otimes \psi_{[n]\setminus\{i\}}||_1 \leq \epsilon$, then $||\rho - \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n||_1 \leq 5n\epsilon$.*

4 Connections between Quantum Property Testing and Distribution Testing

Mutually unbiased bases (MUBs) are used to map the quantum states of $\mathcal{D}(\mathbb{C}^d)$ into $d(d+1)$ dimensional probability distributions. Without loss of generality, assume $d = 2^n$, and we let the Pauli group $\mathcal{P}_k = \{I, X, Y, Z\}^{\otimes n}$ be to the order of 4^n. According to Theorem 11, any state $\rho \in \mathcal{D}(\mathbb{C}^d)$ can be written as

$$\rho = \sum_{P \in \mathcal{P}_n} \eta_p P = \frac{I_d}{d} + \sum_{a=0}^{d} \sum_{\substack{P \in G_a, \\ P \neq I_d}} \eta_p P = \frac{I_d}{d} + \sum_{i,j} \mu_{i,j} |\beta_{i,j}\rangle\langle\beta_{i,j}|,$$

where G_a are the Abelian subgroups with an order of $2^n = d$ such that $\cup G_a = \mathcal{P}_n$ and $G_a \bigcap G_b = \{I_2^{\otimes n}\}$ for $a \neq b$. The equation is due to the simultaneous spectrum decomposition of G_a through the MUBs bases. That is, for $0 \leq i \neq s \leq d, 1 \leq j, t \leq d$,

$$|\langle \beta_{i,j}, \beta_{s,t} \rangle| = \frac{1}{\sqrt{d}}.$$

In addition, it is verifiable that $\sum_{j=1}^{d} \mu_{i,j} = 0$ for all i by the traceless property of $P \neq I_d$. Therefore, we can obtain the following constraint on $\mu_{i,j}$ using $\sum_{j=1}^{d} \mu_{i,j} = 0$ for all i,

$$\operatorname{Tr} \rho^2 = \operatorname{Tr} \frac{I_d}{d^2} + \sum_{i,j,s,t} \mu_{i,j}\mu_{s,t} |\langle \beta_{i,j}, \beta_{s,t} \rangle|^2 = \frac{1}{d} + \sum_{i,j} \mu_{i,j}^2 \leq 1.$$

$\mathcal{M} = \{M_{ij} = \frac{|\beta_{i,j}\rangle\langle\beta_{i,j}|}{d+1} : 0 \leq i \leq d,\ 1 \leq j \leq d\}$ can be used to map the d-dimensional quantum state ρ into a $d(d+1)$ dimensional probabilistic distribution. The corresponding probability distribution $p = (p(0,1), \ldots, p(d,d))$ satisfies

$$p(i,j) = \frac{\operatorname{Tr}(\rho |\beta_{i,j}\rangle\langle\beta_{i,j}|)}{d+1} = \frac{\mu_{i,j} + \frac{1}{d}}{d+1},$$

note that other terms are orthogonal or cancel out due to the property of MUBs and the equations $\sum_{j=1}^{d} \mu_{i,j} = 0$ for all i.

Then the ℓ_2 norm of p can be bounded with

$$\frac{\sqrt{\sum_{i,j} (\mu_{i,j} + \frac{1}{d})^2}}{d+1} = \frac{\sqrt{\sum_{i,j} \mu_{i,j}^2 + \frac{d(d+1)}{d^2} + \frac{2\sum_{i,j} \mu_{i,j}}{d}}}{d+1} = \frac{\sqrt{\sum_{i,j} \mu_{i,j}^2 + \frac{d+1}{d}}}{d+1} \leq \frac{\sqrt{2}}{d+1}.$$

More importantly, this map preserves the ℓ_2 distance, in the sense that the ℓ_2 distance between the image probability distributions is exactly the same as the ℓ_2 distance between the pre-image quantum states with a scaling of $\frac{1}{d+1}$.

For any two states $\rho = \frac{I_d}{d} + \sum_{i,j} \mu_{i,j} |\beta_{i,j}\rangle\langle\beta_{i,j}|$ and $\sigma = \frac{I_d}{d} + \sum_{i,j} \nu_{i,j} |\beta_{i,j}\rangle\langle\beta_{i,j}|$, we have that

$$||\rho - \sigma||_2 = ||\sum_{i,j} (\mu_{i,j} - \nu_{i,j}) |\beta_{i,j}\rangle\langle\beta_{i,j}|||_2 = \sqrt{\sum_{i,j} (\mu_{i,j} - \nu_{i,j})^2},$$

where the other terms are orthogonal or cancel out due to the property of MUBs and the equation $\sum_{j=1}^{d} \mu_{i,j} = 0$ for all i.

Using the measurement $\mathcal{M}$, the corresponding probability distributions can be obtained: $p = (p(0,1), \ldots, p(d,d))$ and $q = (q(0,1), \ldots, q(d,d))$ with

$$p(i,j) = \frac{\operatorname{Tr}(\rho |\beta_{i,j}\rangle\langle\beta_{i,j}|)}{d+1} = \frac{\mu_{i,j} + \frac{1}{d}}{d+1}, \quad q(i,j) = \frac{\operatorname{Tr}(\sigma |\beta_{i,j}\rangle\langle\beta_{i,j}|)}{d+1} = \frac{\nu_{i,j} + \frac{1}{d}}{d+1}.$$

The following equality proves Lemma 2. $||p - q||_2 = \frac{\sqrt{\sum_{i,j} (\mu_{i,j} - \nu_{i,j})^2}}{d+1} = \frac{||\rho - \sigma||_2}{d+1}$.

5 Quantum State Certification

The connections developed in Section 4, together with the ℓ_2-identity tester of probability distributions provided in [23], also make efficient identity testing of quantum states possible.

Proof of Theorem 3. First map the state into probability distributions, say p and q, through independent measurement with Theorem 2, and follow by executing Algorithm 4.

Algorithm 4 A Identity Test with Independent Measurement.

Input: $O(\frac{d^2}{\epsilon^2})$ copies of $\rho \in \mathcal{D}(\mathbb{C}^d)$ and $O(\frac{d}{\epsilon^2})$ copies of $\sigma \in \mathcal{D}(\mathbb{C}^d)$
Output: "Yes" with a probability of at least $\frac{2}{3}$ if $\rho = \sigma$; and "No" with a probability of at least $\frac{2}{3}$ if $||\rho - \sigma||_1 > \epsilon$.

1 Run Algorithm 3 to distinguish between $p = q$ and $||p - q||_2 \geq \frac{\epsilon}{\sqrt{d}(d+1)}$;

/* p and q are the probability distributions obtained by measuring ρ and σ through the independent measurement with Theorem 2, respectively. */

According to $||p - q||_2 = \frac{||\rho - \sigma||_2}{d+1}$, we only need to distinguish cases where $p = q$ from cases where $||p - q||_2 \geq \frac{||\rho - \sigma||_1}{\sqrt{d}(d+1)} \geq \frac{\epsilon}{\sqrt{d}(d+1)}$. Choosing $b = \frac{\sqrt{2}}{d+1} \geq ||p||_2, ||q||_2$ and invoking Theorem 13, we have

$$O(\frac{b}{(\frac{\epsilon}{\sqrt{d}(d+1)})^2}) = O(\frac{d^2}{\epsilon^2})$$

which is a sufficient number of copies. ◀

According to [44], the sample complexity for tomography is $\rho \in \mathcal{D}(\mathbb{C}^d)$ is $\Theta(\frac{d^3}{\epsilon^2})$, which makes Algorithm 4 a better choice for identity testing after tomography.

As mentioned in the introduction, Algorithm 4 should be significantly easier to implement because it does not demand noiseless, universal quantum computation with an *exponential* number of qubits.

6 Independence Testing

The goal of independence testing is to determine whether a fixed multipartite state ρ is independent, i.e., in tensor product form, or far from being independent. Hence, in this section, we outline a series of testing algorithms and almost matching lower bounds in joint measurement setting, and independent measurement setting.

We start with an algorithm for the bipartite case of Theorem 5.

Algorithm 5 A Bipartite Independence Testing with Joint Measurement.

Input: $n = O(\frac{d_1 d_2}{\epsilon^2})$ copies of $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$
Output: "Yes" with a probability of at least $\frac{2}{3}$ if ρ is independent; and "No" with a probability of at least $\frac{2}{3}$ if $||\rho - \sigma||_1 > \epsilon$ for any independent σ.

1 Use $\frac{n}{3}$ copies of ρ to generate ρ_1;
`/* Trace out system 2 */`
2 Use $\frac{n}{3}$ copies of ρ to generate ρ_2;
`/* Trace out system 1 */`
3 Run Algorithm 2 on $\frac{n}{3}$ copies of ρ and $\frac{n}{3}$ copies of $\rho_1 \otimes \rho_2$ with the parameter $\epsilon/3$;

Proof. The correctness of this algorithm accords with Theorem 1 by note that

- If ρ is independent, then $\rho = \rho_1 \otimes \rho_2$, and this algorithm will output "Yes" with high probability.
- If $||\rho - \sigma||_1 > \epsilon$ for any independent σ, then $||\rho - \rho_1 \otimes \rho_2||_1 > \epsilon/3$ by Proposition 14, and this algorithm will output "No" with high probability.

We can derive an independent measurement tester by replacing the identity tester in Algorithm 2 with Algorithm 4. From a similar analysis to the above, $O(\frac{d_1^2 d_2^2}{\epsilon^2})$ is a sufficient number of copies. ◀

The obvious generalization of the bipartite independence testing to m-partite would work using bipartite independence in any $n-1$ parties versus 1 party. Our goal is to test independence in this scenario with an accuracy of $O(\frac{\epsilon}{n})$ and at least a $1 - \frac{1}{n^2}$ probability of success. The correctness of the algorithm follows from Proposition 18, and the generalization incurs an $O(n^3 \log n)$ factor. For constant n, $O(n^3 \log n)$ is still constant. Thus, the complexity of the different algorithm variants would be $O(\frac{\Pi_{i=1}^n d_i}{\epsilon^2})$ with joint measurement, and $O(\frac{\Pi_{i=1}^n d_i^2}{\epsilon^2})$ with independent measurement. With a super-constant n, algorithms could be built that achieve the same complexity using Diakonikolas and Kane's [31] recursion idea coupled with our previous bipartite independence tester.

We only prove the lower bound part of Theorem 5 for bipartite systems here. The general version can be proved similarly. In cases where d_1 and d_2 are both very large, the bound is derived from the mixness test of Theorem 12 in [55], where the constant 2000 comes from the upper and lower bound of the constant in that theorem. To deal with "unbalanced" cases where only d_1 or d_2 is small–here, let us says d_2–we split the d_1 system into many systems of dimension d_2, which transforms the original unbalance of a bipartite problem into a problem of "balanced" multipartite independence testing. Then, we use Proposition 18.

Proof. First, note that it suffices to consider cases where $d_1 d_2$ are sufficiently large. To show the lower bound for a general d_1 and d_2, assume there is an algorithm, Algorithm A, that uses $f(d_1, d_2, \epsilon)$ copies to decide whether a given $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$ is independent or ϵ-far from being independent with at least a 2/3 probability of successful. By using Algorithm A as an oracle, the following algorithm can distinguish cases where $\rho = \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}$ from cases where $||\rho - \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}||_1 > \epsilon$ for any $t > 1$.

To see this algorithm to succeed at detecting whether ρ is maximally mixed with high probability, note that: If $\rho = \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}$, in Line 1, the algorithm will output $\rho_1 = \frac{I_{d_1}}{d_1}$ with a probability of at least $\frac{20}{27}$; in Line 5, the algorithm will output $\rho_2 = \frac{I_{d_2}}{d_2}$ with a probability of at least $\frac{27}{28}$; in Line 9, ρ will be independent with a probability of at least $\frac{28}{30}$. Overall, Algorithm 6 will output "Yes" with a probability of at least $\frac{2}{3}$.

If $||\rho - \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}||_1 > \epsilon$, then one of the following three statements will be true: ρ_1 is ϵ/t-far from $\frac{I_{d_1}}{d_1}$; or ρ_2 is $\frac{(t-1)\epsilon}{4t}$-far from $\frac{I_{d_2}}{d_2}$; or ρ is $\frac{(t-1)\epsilon}{4t}$-far from being independent. Otherwise, assume that there exists an σ_1 and an σ_2, such that $||\rho - \sigma_1 \otimes \sigma_1||_1 < \frac{(t-1)\epsilon}{4t}$, $||\rho_1 - \frac{I_{d_1}}{d_1}||_1 < \frac{\epsilon}{t}$ and $||\rho_2 - \frac{I_{d_2}}{d_2}||_1 < \frac{(t-1)\epsilon}{4t}$. According to Proposition 14, we have $||\rho - \rho_1 \otimes \rho_2||_1 < \frac{3(t-1)\epsilon}{4t}$. Then by the triangle inequality and Lemma 15, we have

$$||\rho - \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}||_1 \leq ||\rho - \rho_1 \otimes \rho_1||_1 + ||\frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2} - \rho_1 \otimes \rho_1||_1 < \epsilon.$$

Contradiction! Therefore, in this case, the algorithm outputs "No" with a probability of at least $\min\{\frac{20}{27}, \frac{27}{28}, \frac{28}{30}\} > \frac{2}{3}$.

Algorithm 6 A Bipartite Identity test A for a maximally mixed state.

Input: $n = 100f(d_1, d_2, \frac{(t-1)\epsilon}{4t}) + 300t^2\frac{d_1}{\epsilon^2} + \Theta(\frac{d_2}{t^2(t-1)^2\epsilon^2})$ copies of $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$

Output: "Yes" with a probability of at least $\frac{2}{3}$ if $\rho = \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}$; and "No" with a probability of at least $\frac{2}{3}$ if $||\rho - \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}||_1 > \epsilon$.

```
Repeat Algorithm 1, with 100t^2 d_1/ε^2 copies of ρ, three times to test whether ρ_1 = I_{d_1}/d_1
  or ||ρ_1 − I_{d_1}/d_1||_1 > ε/t with at least a 20/27 probability of success;
if "No" then
    Return "No";
else
    Employ Algorithm 1 with Θ(t^2 d_2/((t−1)^2 ε^2)) copies of ρ to test whether ρ_2 = I_{d_2}/d_2 or
      ||ρ_1 − I_{d_1}/d_1||_1 > (t−1)ε/4t with at least a 27/28 probability of success;
    if "No" then
        Return "No";
    else
        Run Algorithm A 100 times to test whether ρ is independent or is (t−1)ε/4t-far
          from being independent with at least a 28/30 probability of success;
        if "Yes" then
            Return "Yes";
        else
            Return "No";
```

This algorithm uses $n = 100f(d_1, d_2, \frac{(t-1)\epsilon}{4t}) + 300t^2\frac{d_1}{\epsilon^2} + \Theta(\frac{t^2 d_2}{(t-1)^2\epsilon^2})$ copies of ρ. Invoking Theorem 12, we know that $0.15\frac{d_1 d_2}{\epsilon^2}$ copies are necessary to test, with at least a $2/3$ probability of success, whether ρ is the maximally mixed state or whether it is ϵ-far.

We must have

$$100f(d_1, d_2, \frac{(t-1)\epsilon}{4t}) + 300t^2\frac{d_1}{\epsilon^2} + \Theta(t^2\frac{d_2}{(t-1)^2\epsilon^2}) \geq 0.15\frac{d_1 d_2}{\epsilon^2}.$$

If d_1 and d_2 are both sufficiently large, we can choose a constant t such that $300t^2\frac{d_1}{\epsilon^2} + \Theta(t^2\frac{d_2}{(t-1)^2\epsilon^2}) = o(\frac{d_1 d_2}{\epsilon^2})$, which implies

$$f(d_1, d_2, \epsilon) \geq \Omega(\frac{16t^2 d_1 d_2}{(t-1)^2\epsilon^2}) = \Omega(\frac{d_1 d_2}{\epsilon^2}).$$

If d_1 is sufficiently large and d_2 is not sufficiently large but $d_2 > 2000$, we can choose $t = \sqrt{\frac{2000.5}{2000}}$, then

$$f(d_1, d_2, c\epsilon) \geq 0.15\frac{d_1 d_2}{\epsilon^2} - 300t^2\frac{d_1}{\epsilon^2} + \Omega(t^2\frac{d_2}{(t-1)^2\epsilon^2}) = \Omega(\frac{d_1}{\epsilon^2}) = \Omega(\frac{d_1 d_2}{\epsilon^2}),$$

with the constant $c = \frac{t-1}{4t}$. Thus, for $d_2 > 2000$,

$$f(d_1, d_2, \epsilon) \geq \Omega(\frac{d_1 d_2}{\epsilon^2}).$$

The above technique does not work with a small d_2, because the number of copies required to test a d_1 system $300t^2\frac{d_1}{\epsilon^2}$ and the number of copies required to test a total system of $0.15\frac{d_1 d_2}{\epsilon^2}$ are of the same order.

To deal with this unbalanced case, we develped a dimension splitting technique that transforms bipartite independence into k-partite independence. First observe that the sample complexity for independence testing in $\mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$ is no less than the sample complexity for an independence test of $\mathcal{D}(\mathbb{C}^{d} \otimes \mathbb{C}^{2})$ for $d = 2^{[\log d_1]} \leq d_1$. Therefore, without loss of generality, assume that $d_2 = 2$ and $d_1 = 2^k$ instead of $d_2 \leq 2000$, and that d_1 is sufficiently large.

We still assume that there is an Algorithm A that uses $f(2^k, 2, \epsilon)$ copies to decide, with at least a $2/3$ probability of success, whether a given $\rho \in \mathcal{D}(\mathbb{C}^{2^k \times 2^k} \otimes \mathbb{C}^{2\times 2})$ is independent or ϵ-far from independent in the 2^k and 2 bipartitions. Any such ρ can be regarded as a $k+1$ qubit state, and the qubit systems will be labeled as $S = \{1, 2, \ldots, k, k+1\}$. ρ_i denotes the reduced density matrix of the i-th qubit of ρ. Algorithm A is a bipartite independence tester for a $k+1$ qubit system in the bipartition of k qubits and 1 qubit. In the following, Algorithm A is applied as a black box to the bipartition i and $S \setminus \{i\}$ for any i to test the identity of ρ and $\frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}$.

Algorithm 7 A Bipartite Identity Test B for a maximally mixed state.

Input: $n = \Theta[(k+1)\log k f(2^k, 2, \frac{\epsilon}{6(k+1)})] + \Theta[(k+1)\log k\frac{(k+1)^2}{\epsilon^2}]$ copies of ρ.

Output: "Yes" with a probability of at least $\frac{2}{3}$ if $\rho = \otimes_{i=1}^{k+1}\otimes\frac{I_2}{2}$; and "No" with a probability of at least $\frac{2}{3}$ if $||\rho - \otimes_{i=1}^{k+1}\otimes\frac{I_2}{2}||_1 > \epsilon$.

```
for i ← 1 to k + 1 do
    Repeat Algorithm 1, with Θ((k+1)^2/ε^2) copies of ρ each time, Θ(log k) times to test
     whether ρ_i = I_2/2 or ||ρ_i − I_2/2||_1 > ε/(6(k+1)) at least a 1 − 1/k^2 probability of success;
    if No then
        Return "No";
    else
        Run Algorithm A Θ(log k) times, with f(2^k, 2, ε/(6(k+1))) copies each time, to
         test whether ρ is independent or ε/(6(k+1))-far from being independent in the
         bipartition {i} and S \ {i} with at least a 1 − 1/k^2 probability of success;
        if No then
            Return "No";
Return "Yes";
```

To see this algorithm succeed in detecting whether ρ is maximally mixed with high probability, we note that

If $\rho = \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}$, then $\rho_i = \frac{I_2}{2}$ when ρ is regarded as a $k+1$ qubit state. It is independent in any bipartition $\{i\}$ and $S \setminus \{i\}$. For each i, the passing probability of the test $\rho_i = \frac{I_2}{2}$ is at least $1 - \frac{1}{k^2}$. For each i, the passing probability of the independence test in the bipartition $\{i\}$ and $S \setminus \{i\}$ is at least $1 - \frac{1}{k^2}$. In total, Algorithm 7 will accept with a probability of at least $(1 - \frac{1}{k^2})^{O(k)} = 1 - o(1) > \frac{2}{3}$.

If $||\rho - \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}||_1 > \epsilon$, at least one of the following two statements is true:

- $||\rho_i - \frac{I_2}{2}||_1 > \frac{\epsilon}{6(k+1)}$ for some $1 \leq i \leq k+1$; and/or
- ρ is $\frac{\epsilon}{6(k+1)}$-far from independent in the bipartition $\{i\}$ and $S \setminus \{i\}$ for some $1 \leq i \leq k+1$.

Otherwise, $||\rho_i - \frac{I_2}{2}||_1 \leq \frac{\epsilon}{6(k+1)}$ and ρ is $\frac{\epsilon}{6(k+1)}$ close to being independent in the bipartition $\{i\}$ and $S \setminus \{i\}$ for all $1 \leq i \leq k+1$.

According to Proposition 18, we have

$$||\rho - \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_{k+1}||_1 \leq 5(k+1)\frac{\epsilon}{6(k+1)} = \frac{5\epsilon}{6}.$$

By Lemma 17, we have

$$\begin{aligned}
& ||\rho - \frac{I_{d_1}}{d_1} \otimes \frac{I_{d_2}}{d_2}||_1 \\
= \; & ||\rho - \frac{I_2}{2} \otimes \frac{I_2}{2} \otimes \cdots \otimes \frac{I_2}{2}||_1 \\
\leq \; & ||\rho - \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_{k+1}||_1 + ||\frac{I_2}{2} \otimes \frac{I_2}{2} \otimes \cdots \otimes \frac{I_2}{2} - \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_{k+1}||_1 \\
\leq \; & \frac{5\epsilon}{6} + \sum_{i=1}^{k+1} ||\rho_i - \frac{I_2}{2}||_1 \\
\leq \; & \epsilon.
\end{aligned}$$

Contradiction! Therefore, the algorithm outputs "No" with a probability of at least $1 - \frac{1}{k^2} > \frac{2}{3}$ in this case.

Invoking Theorem 12, we know that $\Theta(\frac{d_1 d_2}{\epsilon^2}) = \Theta(\frac{2^{k+1}}{\epsilon^2})$ copies are necessary to test whether ρ is a maximally mixed state or ϵ-far with at least a $2/3$ probability of success. Algorithm 7 uses $\Theta[(k+1)\log k f(2^k, 2, \frac{\epsilon}{6(k+1)})] + \Theta[(k+1)\log k \frac{(k+1)^2}{\epsilon^2}]$ copies of ρ. We must have

$$\begin{aligned}
& \Theta[(k+1)\log k f(2^k, 2, \frac{\epsilon}{6(k+1)})] + \Theta[(k+1)\log k \frac{(k+1)^2}{\epsilon^2}] \geq \Theta(\frac{2^{k+1}}{\epsilon^2}) \\
\Rightarrow \; & f(2^k, 2, \frac{\epsilon}{6(k+1)}) \geq \Theta(\frac{2^k}{k \log k \epsilon^2}) \\
\Rightarrow \; & f(2^k, 2, \epsilon) \geq \Theta(\frac{2^k}{k^3 \log k \epsilon^2}) \\
\Rightarrow \; & f(d_1, d_2, \epsilon) = \Omega(\frac{d_1}{\log^3 d_1 \log\log d_1 \epsilon^2}) = \Omega(\frac{d_1 d_2}{\log^3 d_2 \log\log d_1 \epsilon^2})
\end{aligned}$$

That is, if d_2 is a small constant, $\Omega(\frac{d_1 d_2}{\epsilon^2 \log^3 d_1 \log\log d_1})$ copies are necessary to test the independence of $\rho \in \mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2})$. ◀

7 Discussion

If we can use random measurements, a fewer number of copies are needed for quantum identity testing and independence testing.

In [19], it is proved that using non-adaptive independent measurements, to test whether a quantum state $\rho \in \mathcal{D}(\mathbb{C}^d)$ is equal to or ϵ-far in trace distance from the maximally mixed state, $\Omega(\frac{d^{3/2}}{\epsilon^2})$, and this complexity can be achieved via Haar-random orthogonal POVMs. The measurement can be implemented by randomly choosing a unitary U applied on ρ and measuring $U\rho U^\dagger$ in computational basis many times. The last step is to test whether the resulting d-dimensional probability dsitribution p_U is equal to or $\frac{\epsilon}{\sqrt{d}}$-far from uniform distribution u. The correctness of this algorithm comes from the concentration of measure and

$$\mathbb{E}_U ||p_U - u||_2^2 = \frac{||\rho - \frac{I}{d}||_2^2}{d+1}.$$

This method can be used to the general quantum identity testing problem: Randomly choosing a unitary U applied on ρ and σ respectively, then measuring $U\rho U^\dagger$ and $U\sigma U^\dagger$ in computational basis many times. The last step is to test whether the resulting d-dimensional probability dsitributions p_U and q_U are equal or $\frac{\epsilon}{\sqrt{d}}$-far. One can verify

$$\mathbb{E}_U ||p_U - q_U||_2^2 = \frac{||\rho - \sigma||_2^2}{d+1},$$

and

$$\mathbb{E}_U ||p_U||_2^2, ||q_U||_2^2 \leq O(\frac{1}{d+1}).$$

Using concentration of measure, we know that $||p_U - q_U||_2^2 \geq \frac{||\rho-\sigma||_2^2}{2d+1}$ and $||p_U||_2^2, ||q_U||_2^2 \leq O(\frac{1}{d+1})$ are valid with high probability. The rest is to run Algorithm 3 from [23] with $O(\frac{d^{3/2}}{\epsilon^2})$ samples. We can conclude that

▶ **Theorem 19.** *The sample complexity of quantum identity testing is* $\Theta(\frac{d^{3/2}}{\epsilon^2})$ *for non-adaptive independent measurements.*

This continuous randomness can be discretized by randomly choosing the MUB basis presented in Section 4.

Plug in our method of quantum independence testing, we know that

▶ **Theorem 20.** *The sample complexity of quantum independence testing of* $\mathcal{D}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_m})$ *is* $\tilde{\Theta}(\frac{\Pi_{i=1}^m d_i^{3/2}}{\epsilon^2})$ *for non-adaptive independent measurements.*

References

1. S. Aaronson. Shadow tomography of quantum states. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, pages 325–338, 2018.
2. S. Aaronson, X. Chen, E. Hazan, S. Kale, and A. Nayak. Online learning of quantum states. In *Advances in Neural Information Processing Systems 31*, pages 8962–8972, 2018.
3. S. Aaronson and G. Rothblum. Gentle measurement of quantum states and differential privacy. In *Proceedings of the 51th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, 2019.
4. Scott Aaronson. The learnability of quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2088):3089–3114, September 2007.
5. J. Acharya, H. Das, A. Jafarpour, A. Orlitsky, and S. Pan. Competitive closeness testing. In *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19, pages 47–68, 2011.
6. J. Acharya, C. Daskalakis, and G. Kamath. Optimal testing for properties of distributions. In *Advances in Neural Information Processing Systems 28*, pages 3591–3599, 2015.

7 J. Acharya, I. Diakonikolas, J. Li, and L. Schmidt. Sample-optimal density estimation in nearly-linear time. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, pages 1278–1289, 2017.

8 J. Acharya, I. Issa, N. Shende, and A. B. Wagner. Measuring quantum entropy. *1711.00814*, 2017.

9 Anurag Anshu, Dave Touchette, Penghui Yao, and Nengkun Yu. Exponential separation of quantum communication and classical information. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 277–288, New York, NY, USA, 2017. ACM. doi:10.1145/3055399.3055401.

10 Leandro Aolita, Christian Gogolin, Martin Kliesch, and Jens Eisert. Reliable quantum certification of photonic state preparations. *Nature Communications*, 6, 2015.

11 Bandyopadhyay, Boykin, Roychowdhury, and Vatan. A new proof for the existence of mutually unbiased bases. *Algorithmica*, 34(4):512–528, 2002.

12 T. Batu, S. Dasgupta, R. Kumar, and R. Rubinfeld. The complexity of approximating entropy. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 678–687, 2002.

13 T. Batu, L. Fortnow, E. Fischer, R. Kumar, R. Rubinfeld, and P. White. Testing random variables for independence and identity. In *Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science*, FOCS '01, pages 442–451, 2001.

14 T. Batu, L. Fortnow, R. Rubinfeld, W. D. Smith, and P. White. Testing that distributions are close. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, FOCS'00, pages 259–269, 2000.

15 T. Batu, R. Kumar, and R. Rubinfeld. Sublinear algorithms for testing monotone and unimodal distributions. In *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing*, STOC '04, pages 381–390, 2004.

16 S. Bravyi, A. W. Harrow, and A. Hassidim. Quantum algorithms for testing properties of distributions. *IEEE Transactions on Information Theory*, 57(6):3971–3981, 2011.

17 C. Bădescu and R. O'Donnell. Lower bounds for testing complete positivity and quantum separability. In *14th Latin American Theoretical Informatics Symposium*, 2020.

18 C. Bădescu, R. O'Donnell, and J. Wright. Quantum state certification. In *Proceedings of the Forty-Nineth Annual ACM on Symposium on Theory of Computing*, STOC '19, 2019.

19 Sébastien Bubeck, Sitan Chen, and Jerry Li. Entanglement is necessary for optimal quantum property testing. In *FOCS 2020*, April 2020.

20 C. L. Canonne. A survey on distribution testing: Your data is big. but is it blue? *Electronic Colloquium on Computational Complexity (ECCC)*, 22:63, 2015.

21 C. L. Canonne, I. Diakonikolas, T. Gouleakis, and R. Rubinfeld. Testing shape restrictions of discrete distributions. *Theory of Computing Systems*, 62(1):4–62, 2018.

22 C. L. Canonne, I. Diakonikolas, D. M. Kane, and A. Stewart. Testing conditional independence of discrete distributions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, pages 735–748, 2018.

23 S. Chan, I. Diakonikolas, G. Valiant, and P. Valiant. Optimal algorithms for testing closeness of discrete distributions. In *Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 1193–1203, 2014.

24 Y. Cheng, I. Diakonikolas, and R. Ge. High-dimensional robust mean estimation in nearly-linear time. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2019.

25 Marcus P. da Silva, Olivier Landon-Cardinal, and David Poulin. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.*, 107:210404, 2011.

26 C. Daskalakis, I. Diakonikolas, R. A. Servedio, G. Valiant, and P. Valiant. Testing k-modal distributions: Optimal algorithms via reductions. In *Proceedings of the Twenty-fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '13, pages 1833–1852, 2013.

27 C. Daskalakis, N. Dikkala, and G. Kamath. Testing ising models. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '18, pages 1989–2007, 2018.

28 C. Daskalakis and Q. Pan. Square hellinger subadditivity for bayesian networks and its applications to identity testing. In *Proceedings of the 2017 Conference on Learning Theory*, volume 65, pages 697–703, 2017.

29 L. Devroye and G. Lugosi. *Combinatorial Methods in Density Estimation*. Springer, 2001.

30 I. Diakonikolas, G. Kamath, D. M. Kane, J. Li, A. Moitra, and A. Stewart. Robust estimators in high dimensions without the computational intractability. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 655–664, 2016.

31 I. Diakonikolas and D. Kane. A new approach for testing properties of discrete distributions. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 685–694, 2016.

32 I. Diakonikolas, D. M. Kane, and V. Nikishkin. Optimal algorithms and lower bounds for testing closeness of structured distributions. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1183–1202, 2015.

33 I. Diakonikolas, D. M. Kane, and V. Nikishkin. Near-Optimal Closeness Testing of Discrete Histogram Distributions. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80, pages 8:1–8:15, 2017.

34 I. Diakonikolas, D. M. Kane, and A. Stewart. Sharp bounds for generalized uniformity testing. In *Advances in Neural Information Processing Systems 31*, pages 6201–6210, 2018.

35 T. Durt, B. Englert, I. Bengtsson, and Zyczkowski K. On mutually unbiased bases. *International Journal of Quantum Information*, pages 535–640, 2010.

36 S. T. Flammia, D. Gross, Y. Liu, and J. Eisert. Quantum tomography via compressed sensing: Error bounds, sample complexity, and efficient estimators. *New J. Phys.*, 14:095022, 2012.

37 Steven T. Flammia and Yi-Kai Liu. Direct fidelity estimation from few pauli measurements. *Phys. Rev. Lett.*, 106:230501, 2011.

38 András Gilyén and Tongyang Li. Distributional Property Testing in a Quantum World. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151, pages 25:1–25:19, 2020.

39 O. Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017.

40 O. Goldreich and D. Ron. *On Testing Expansion in Bounded-Degree Graphs*, volume 6650 of *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation, Lecture Notes in Computer Science*. Springer, 2000.

41 Oded Goldreich, Shari Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, July 1998.

42 D. Gross, Y. Liu, S. T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105(150401), 2010.

43 D. Gross, S. Nezami, and M. Walter. Schur-weyl duality for the clifford group with applications: Property testing, a robust hudson theorem, and de finetti representations. *arXiv*, 2017. `arXiv:1712.08628`.

44 J. Haah, A. W. Harrow, Z. Ji, X. Wu, , and N. Yu. Sample-optimal tomography of quantum states. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '16, pages 913–925, 2016.

45 Aram W. Harrow and Ashley Montanaro. Testing product states, quantum merlin-arthur games and tensor optimization. *J. ACM*, 60(1):3:1–3:43, 2013.

46 P. Indyk, R. Levi, and R. Rubinfeld. Approximating and testing k-histogram distributions in sub-linear time. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, PODS '12, pages 15–22, 2012.

47 R. Jain, J. Radhakrishnan, and Sen P. A lower bound for the bounded round quantum communication complexity of set disjointness. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 220–229, October 2003. `doi:10.1109/SFCS.2003.1238196`.

48 J. Jiao, K. Venkat, Y. Han, and T. Weissman. Minimax estimation of functionals of discrete distributions. *IEEE Transactions on Information Theory*, 61(5):2835–2885, 2015.

49 R. Kueng, H. Rauhut, and U. Terstiege. Low rank matrix recovery from rank one measurements. *Applied and Computational Harmonic Analysis*, 42:88–116, 2017.

50 E. L. Lehmann and Joseph P. Romano. *Testing statistical hypotheses*. Springer Texts in Statistics. Springer, New York, 2005.

51 R. Levi, D. Ron, and R. Rubinfeld. Testing properties of collections of distributions. In *proceedings of the Second Symposium on Innovations in Computer Science*, ICS '11, pages 179–194, 2011.

52 Florian Mintert, Marek Kuś, and Andreas Buchleitner. Concurrence of mixed multipartite quantum states. *Phys. Rev. Lett.*, 95:260502, 2005.

53 A. Montanaro and R. de Wolf. A survey of quantum property testing. *Theory of Computing Graduate Surveys*, 7, 2016.

54 M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 10th edition, 2011.

55 R. O'Donnell and J. Wright. Quantum spectrum testing. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '15, pages 529–538, 2015.

56 R. O'Donnell and J. Wright. Efficient quantum tomography. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '16, pages 899–912, 2016.

57 R. O'Donnell and J. Wright. Efficient quantum tomography ii. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '17, pages 962–974, 2017.

58 L. Paninski. A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Trans. Inf. Theor.*, 54(10):4750–4755, 2008.

59 R. Rubinfeld. Taming big probability distributions. *XRDS*, 19(1):24–28, 2012.

60 R. Rubinfeld and M. Sudan. Self-testing polynomial functions efficiently and over rational domains. In *Proceedings of the 3rd Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '92, pages 23–32, 1992.

61 R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

62 I. N. Sanov. On the probability of large deviations of random variables. *Mat. Sbornik*, 42:11–44, 1957.

63 G. Valiant and P. Valiant. Estimating the unseen: An n/log(n)-sample estimator for entropy and support size, shown optimal via new clts. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, STOC '11, pages 685–694, 2011.

64 G. Valiant and P. Valiant. The power of linear estimators. In *Proceedings of the 2011 IEEE 52Nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pages 403–412, 2011.

65 G. Valiant and P. Valiant. An automatic inequality prover and instance optimal identity testing. In *Proceedings of the 2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, FOCS '14, pages 51–60, 2014.

66 G. Valiant and P. Valiant. Instance optimal learning of discrete distributions. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 142–155, 2016.

67 G. Valiant and P. Valiant. Estimating the unseen: Improved estimators for entropy and other properties. *Journal of the ACM*, 64(6), 2017.

68 P. Valiant. Testing symmetric properties of distributions. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 383–392, 2008.

69 V. Voroninski. Quantum tomography from few full-rank observables, 2013. `arXiv:1309.7669`.

70 Y. Wu and P. Yang. Minimax rates of entropy estimation on large alphabets via best polynomial approximation. *IEEE Transactions on Information Theory*, 62(6):3702–3720, 2016.

71 Nengkun Yu. Quantum closeness testing: A streaming algorithm and applications, 2020. `arXiv:1904.03218`.