Proceedings of Machine Learning Research vol 195:1–31, 2023 36th Annual Conference on Learning Theory

On Testing and Learning Quantum Junta Channels

Zongbo Bao BAOZB0407@GMAIL.COM
State Key Laboratory for Novel Software Technology, Nanjing University,China

Penghui Yao PHYAO1985@GMAIL.COM
State Key Laboratory for Novel Software Technology, Nanjing University, China
Hefei National Laboratory, Hefei 230088, China

Editors: Gergely Neu and Lorenzo Rosasco

Abstract

We consider the problems of testing and learning quantum k-junta channels, which are n-qubit to n-qubit quantum channels acting non-trivially on at most k out of n qubits and leaving the rest of qubits unchanged. We show the following.

1. An $\widetilde{O}(k)$-query algorithm to distinguish whether the given channel is k-junta channel or is *far* from any k-junta channels, and a lower bound $\Omega(\sqrt{k})$ on the number of queries;
2. An $\widetilde{O}(4^k)$-query algorithm to learn a k-junta channel, and a lower bound $\Omega(4^k/k)$ on the number of queries.

This gives the first junta channel testing and learning results, and partially answers an open problem raised by Chen et al. (2023). In order to settle these problems, we develop a Fourier analysis framework over the space of superoperators and prove several fundamental properties, which extends the Fourier analysis over the space of operators introduced in Montanaro and Osborne (2010).

Keywords: Quantum Channels, Junta Channels, Fourier Analysis, Influence

1. Introduction

It is crucial in quantum computing to understand the behavior of a quantum process, which is also modeled as a quantum channel, in a black-box manner. The most general method for doing this is quantum process tomography (QPT). But it requires a large amount of computational resources, which is exponential in the number of qubits it acts on, as noted by Chuang and Nielsen (1997) and Gutoski and Johnston (2014).

A quantum channel is referred to as a k-junta channel if it acts non-trivially on up to k out of n qubits, leaving the rest qubits unchanged. Characterizing a k-junta channel is easier if k is small, hence it is interesting to find efficient algorithms to test whether a quantum channel is a k-junta channel and learn k-junta channels. The problems of testing and learning k-junta boolean functions is also an important problem in theoretical computer science, having a rich history of research, see Goldreich (2017) and Bhattacharyya and Yoshida (2022). More recently, testing and learning k-junta unitaries has been explored by Montanaro and Osborne (2010); Wang (2011); Chen et al. (2023)

In this paper, we are concerned about the testing and learning k-junta channels. The setting is as follows. Given oracle access to a quantum channel Φ, the algorithm is supposed to output an answer about the channel Φ, where *access* means the algorithm queries the oracle with any n-qubit quantum state ρ and obtains $\Phi(\rho)$ as an output. For both problems, it requires a distance function $\mathrm{dist}(\cdot,\cdot)$ to

© 2023 Z. Bao & P. Yao.

formulate *far* and *close* rigorously. With the assistance of the oracle, we are supposed to determine whether Φ is a k-junta channel or *far* from any k-junta channels in the testing problem or output a description of $\widetilde{\Phi}$, which is close to Φ with respect to the distance. In this work we choose the distance function induced by the inner product over superoperators, which will be formally defined in Section 3.

The first main result is an algorithm testing whether a given black-box channel is a k-junta channel or far from any k-junta channel with $\widetilde{O}(k)$ queries where $\widetilde{O}(\cdot)$ hides ε and logarithmic factor of k.

Theorem 1 (Testing Quantum k-Junta Channels) *There exists an algorithm such that, given oracle access to an n-qubit to n-qubit quantum channel Φ, it makes $O(k \log k/\varepsilon^2)$ queries and determines whether Φ is a k-junta channel or dist$(\Phi, \Psi) \geq \varepsilon$ for any k-junta channel Ψ with probability at least $9/10$. Furthermore, Any quantum algorithm achieving this task requires $\Omega(\sqrt{k})$ queries.*

Our second main result is a learning algorithm, which is given a black-box k-junta channel and outputs a description of a channel close to the channel with $O(4^k/\varepsilon^2)$ queries. We also show that the algorithm is almost optimal. Hence we can learn a k-junta channel efficiently, especially without dependence on the total number of qubits.

Theorem 2 (Learning Quantum k-Junta Channels) *There exists an algorithm, given oracle access to an n-qubit to n-qubit k-junta channel Φ, it makes $O(4^k/\varepsilon^2)$ queries and outputs a description of channel Ψ satisfying dist$(\Phi, \Psi) \leq \varepsilon$ with probability at least $9/10$. Furthermore, any quantum algorithm achieving this task requires $\Omega\left(4^k/k\right)$ queries.*

Both algorithms are proved via Fourier-analytic techniques over superoperators defined in Section 3. In particular, we turn both problems to estimating the influence of a superoperator, which is a generalization of the influence of boolean functions O'Donnell (2014) and the influence of operators Montanaro and Osborne (2010). We prove a series of fundamental properties of the Fourier analysis and the influence of superoperators extending similar results on operators. The lower bound on testing k-junta channels combines the result of testing boolean k-juntas obtained by Bun et al. (2020) and a structural result for k-junta channels. The lower bound on learning k-junta channels is obtained by a reduction from learning k-junta unitaries.

Besides, we exhibit a simple Influence-Estimator to estimate the influence of channels in Appendix F. Compared with the estimator in Chen et al. (2023), where it requires entanglement and 2-qubit operations, our Influence-Estimator requires only single-qubit operations and is as efficient as theirs. Therefore, it might be easily implemented in the lab.

Contributions

1. We develop Fourier analysis over superoperators and prove several basic properties around influence, which are an extension of Fourier analysis over operators Montanaro and Osborne (2010) and may be of independent interest;

2. We present the first k-junta channel testing algorithm and a lower bound for this problem, partially answering an open problem raised by Chen et al. (2023). In addition, we show an almost optimal algorithm for k-junta channel learning problems;

3. We construct a new and simple Influence-Estimator, which may be easy to implement in the lab since it includes only single-qubit operations.

Organization In Section 1.1, we present a brief overview of related works. Section 1.2 provides a high-level overview of the proof techniques. After establishing some preliminaries regarding quantum channels in Section 2, we demonstrate our Fourier-analytic techniques in Section 3, including properties of our distance function. In Section 4 and Section 5, we prove our k-junta channel testing and learning results respectively. Finally, we conclude in Section 6.

1.1. Related Work

A boolean function $f : \{0,1\}^n \to \{0,1\}$ is a k-junta if its value only depends on at most k coordinates of the inputs. Testing and learning boolean juntas has been extensively studied for decades. The first result explicitly related to testing juntas is obtained by Parnas et al. (2002), where an $O(1)$-queries algorithm is given to test 1-juntas. Then Fischer et al. (2004) turned their eyes onto k-junta testing problem and gave an $\widetilde{O}(k^2)$-queries algorithm. This upper bound was improved by Blais (2009) to a nearly optimal algorithm which requires only $O(k \log k)$ queries, provided an $\Omega(k)$ lower bound proved by Chockler and Gutfreund (2004). More recently, Sağlam (2018) gave an $\Omega(k \log k)$ lower bound, which closed the gap. Junta testing has also been investigated in the setting where only non-adaptive queries are allowed Servedio et al. (2015), Chen et al. (2018) and Liu et al. (2019). Learning k-junta boolean function has spawned a large body of work. The learning algorithm obtained by Mossel et al. (2003) was a breakthrough and followed by a series of work Lipton et al. (2005), Arpe and Reischuk (2007), Arpe and Mossel (2008) and Arvind et al. (2009) discussing k-junta learning problem under different circumstances, such as learning symmetric juntas, learning with noise, agnostically learning and considering parameterized learnability. Meanwhile, Bshouty and Costa (2018) tried to understand this problem in the membership query model, and Levi and Waingarten (2019), Blais et al. (2019) and De et al. (2019) turned their eyes to tolerant learning k-juntas. More recently, people paid more and more attention to learning k-junta distribution, see Aliakbarpour et al. (2016) and Chen et al. (2021) for more details.

It is expected that a speedup can be obtained when we use a quantum computer to test or learn boolean juntas. In Atici and Servedio (2007), Atici and Servedio gave the first quantum algorithm, which tests k-junta boolean functions with $O(k)$ queries. More recently, Ambainis et al. (2016) constructed a quantum algorithm which needs only $\widetilde{O}(\sqrt{k})$ queries. This was shown to be optimal up to a polylogarithmic factor Bun et al. (2020) In addition, Atici and Servedio (2007) also gave an $O(2^k)$-sample quantum algorithm for learning k-junta boolean function in the PAC model.

In quantum computing, it is natural to consider the situation where behind the oracle is a quantum operation instead of a boolean function. The quantum junta unitary testing problem is to decide if a unitary U with oracle access is a k-junta or ε-far from any k-junta unitary.

Wang gave an algorithm testing k-junta unitaries with $O(k)$ queries in Wang (2011). Montanaro and Osborne gave a different tester for dictatorship, i.e., 1-junta in Montanaro and Osborne (2010). Recently, Chen, Nadimpalli and Yuen have settled both the quantum testing and learning of quantum juntas problem providing nearly tight upper and lower bounds in Chen et al. (2023). See Table 1.1 for more details.

The algorithms of testing and learning boolean juntas heavily rely on the Fourier analysis of boolean functions, which is nowadays a rich theory and has wide applications in many branches of theoretical computer science. Readers may refer to O'Donnell (2014) or de Wolf (2008) for more details. Fourier analysis on quantum operations has received increasing attention in the past couple of years. Montanaro and Osborne (2010) initiated the study of Fourier analysis on the space of operators and established several interesting properties. Influence is a key notion in Fourier analysis, which describes how much the function value is affected by some subset of inputs and has many applications in theoretical computer science. The analogous notion in the space of operators has also played a crucial role in designing testing and learning algorithms of k-junta unitaries in Wang (2011), Chen et al. (2023).

We summarize related works in Table 1.1.

Table 1: Our contributions and prior work on testing and learning boolean and quantum k-juntas.

	Classical Testing	Quantum Testing	Quantum Learning
$f : \{0,1\}^n \to \{0,1\}$	$O(k \log k)$ Blais (2009)	$\widetilde{O}(\sqrt{k})$ Ambainis et al. (2016)	$O(2^k)$ Atici and Servedio (2007)
	$\Omega(k \log k)$ Sağlam (2018)	$\Omega(\sqrt{k})$ Bun et al. (2020)	$\Omega(2^k)$ Atici and Servedio (2007)
Unitary $U \in \mathcal{M}_{2^n \times 2^n}$	—	$\widetilde{O}(\sqrt{k})$ Chen et al. (2023)	$O(4^k)$ Chen et al. (2023)
	—	$\Omega(\sqrt{k})$ Chen et al. (2023)	$\Omega(4^k/k)$ Chen et al. (2023)
Channel Φ, n to n qubits	—	$\widetilde{O}(k)$ **this work**	$O(4^k)$ **this work**
	—	$\Omega(\sqrt{k})$ **this work**	$\Omega(4^k/k)$ **this work**

1.2. Techniques

In this section, we give a high-level technical overview of our main results.

1.2.1. Testing Junta Channels

Our junta testing algorithm is inspired by the algorithm for k-junta boolean function testing by Atici and Servedio (2007). The algorithm deeply relies on the notion of the influence of superoperators, which captures how much a subset of input qubits affect the output of a channel; see Section 3.1 for more details. The influence of a superoperator is defined through the formal Fourier analysis framework over superoperators. We prove in Section 3.1 that it has many properties similar to the influence of boolean functions and unitaries.

To prove the lower bound, we reduce k-junta channel testing to k-junta boolean function testing, which has a lower bound $\Omega(\sqrt{k})$ by Bun et al. (2020). To make the reduction work, we prove that a tester for k-junta channels is also a tester of k-junta boolean function, if we view a boolean function as a quantum channel. Moreover, we also show that our algorithm naturally induces a tester for k-junta unitaries.

1.2.2. Learning Junta Channels

The learning algorithm is inspired by the algorithms in Atici and Servedio (2007); Chen et al. (2023). We apply PAULI-SAMPLE to the Choi-state of the channel to find the high-influence registers. Then we apply the efficient quantum state tomography algorithm by O'Donnell and Wright (2017) to learn the reduced density operator on the qubits with high influence. The lower bound of learning k-junta channels is obtained by reducing learning k-junta unitaries to learning k-junta channels.

1.2.3. INFLUENCE ESTIMATOR

We propose a new Influence-Estimator to estimate influence for channels, which only needs single-qubit operations. To achieve this, we utilize the ideas from the CSS code and quantum money. The estimator uses Hadamard operators to exchange bit-flip effects and phase-flip effects imposed by the channel and finally decides whether the channel changes the target subset of input qubits too much; see Section F.

2. Preliminary

We assume that readers are familiar with elementary quantum computing and information theory. Readers may refer to Chapters 1 and 2 of Nielsen and Chuang (2000) and Chapters 1 and 2 of Watrous (2018) for more detailed backgrounds. For natural number $n \geq 1$, $[n]$ represents $\{1, 2, \ldots, n\}$. I_n represents an $n \times n$ identity matrix. The subscript may be omitted whenever it is clear from the context. We say a Hermitian matrix is a positive semidefinite matrix (PSD) if all the eigenvalues are nonnegative.

Throughout the paper, we assume that the whole quantum system has n qubits. Let $N = 2^n$ be the dimension of the system. Denote $\Sigma = [N]$ $X = \mathbb{C}^{\Sigma}$. $L(X)$ represents the set of all the linear maps from X to X itself. Therefore all n-qubit quantum states are a subset of $L(X)$. We note that $L(X)$ is isomorphic to $\mathbb{C}^{N \times N}$, the set of $N \times N$ matrices. For any $A \in L(X)$, let $\text{vec}(A) = (A \otimes I) \sum_{i=1}^{N} |i, i\rangle$ be the "stretching" column vector of A. For $x \in \mathbb{Z}_4^n$ and $T \subseteq [n]$, let $x_T \in \mathbb{Z}_4^T$ be the substring of x obtained by restricting x to all the coordinates in T . We write $0^T \in \mathbb{Z}_4^T$ to denote all zero string on coordinates in T. The superscript may be dropped whenever it is clear from the context. We use A^* to stand for the conjugate transpose of A.

Recall the definition of Pauli operators given by

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = X, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = Y, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = Z.$$

It forms an orthogonal basis for $L(\mathbb{C}^2)$ (over $\mathbb{C}$) with respect to the Hilbert-Schmidt inner product. For any $x \in \mathbb{Z}_4^n$, let $\sigma_x = \otimes_{i=1}^n \sigma_{x_i}$. It is easy to check $\{\sigma_x\}_{x \in \mathbb{Z}_4^n}$ is an orthogonal basis for $L(\mathbb{C}^N) = L(X)$.

For $x \in \mathbb{Z}_4^n$, $|v(\sigma_x)\rangle$ represents the quantum state corresponding to column vector $\text{vec}(\sigma_x)$. It is easy to check $\{|v(\sigma_x)\rangle\}_{x \in \mathbb{Z}_4^n}$ is an orthogonal basis in $\mathbb{C}^{2^{2n}}$.

2.1. Superoperators and quantum channels

A superoperator on $L(X)$ is a linear map from $L(X)$ to itself. $T(X)$ represents the set of all superoperators on $L(X)$. A quantum channel $\Phi : L(X) \to L(X)$ is completely positive and a trace preserving superoperator. In this work, we concern ourselves with the channels mapping n qubits to n qubits. We use $C(X)$ to denote the set of all quantum channels from $L(X)$ to itself. For any unitary $U \in L(X)$, Φ_U represents the channel which acts U on the state, i.e., $\Phi_U(\rho) = U\rho U^*$. For any boolean function $g : \{0, 1\}^n \to \{0, 1\}$, we define $\Phi_g = \Phi_{U_g}$, where U_g is the unitary defined to be $U_g |x\rangle = (-1)^{g(x)} |x\rangle$ for $x \in \{0,1\}^n$. Next, we introduce the Kraus representation and the Choi representation of superoperators. The properties and relations around two representations are postponed to Appendix A.

Definition 3 (Kraus representations, Choi Representations, Choi states) *Given superoperator $\Phi \in T(X)$, its Kraus representation is*

$$\Phi(\rho) = \sum_{s\in\Sigma} A_s \rho B_s^*$$

where $A_s, B_s \in L(X)$. Its Choi representation is

$$J(\Phi) = \sum_{a,b\in\Sigma} \Phi(|a\rangle\langle b|) \otimes |a\rangle\langle b| = (\Phi \otimes I)\left(\sum_{a,b\in\Sigma} |a\rangle\langle b| \otimes |a\rangle\langle b|\right) \in L(X \otimes X),$$

where J is a linear map from $T(X)$ to $L(X \otimes X)$.

For a quantum channel Φ, the Choi state $v(\Phi)$ is defined to be

$$v(\Phi) = \frac{J(\Phi)}{\operatorname{Tr} J(\Phi)}.$$

The Choi state of unitaries is defined similarly. Note that for a unitary U, its Choi state is a pure state, denoted by $|v(U)\rangle$.

By Fact 25, $v(\Phi)$ is a density operator if Φ is a quantum channel.

At the end of this section we introduce k-junta channels.

Definition 4 (k-Junta Channels) *Given $\Phi \in C(X)$ and a subset $T \subseteq [n]$, we say Φ is a T-junta channel if $\Phi = \Phi_T \otimes I_{T^c}$. Φ is a k-junta channel if Φ is a T-junta channel for some $T \subseteq [n]$ of size k.*

3. Fourier Analysis over superoperators

We are ready to introduce the Fourier analysis over superoperators. For any superoperators $\Phi, \Psi \in T(X)$, define the inner product $\langle \Phi, \Psi\rangle = \langle J(\Phi), J(\Psi)\rangle = \operatorname{Tr} J(\Phi)^* J(\Psi)$. It is easy to verify that $\langle\cdot,\cdot\rangle$ is an inner product and $(T(X), \langle\cdot,\cdot\rangle)$ forms a finite-dimensional Hilbert space. The norm of Φ is defined to be $\|\Phi\| = \sqrt{\langle\Phi,\Phi\rangle} = \|J(\Phi)\|_2$, where $\|\cdot\|_2$ is the Frobenius norm. The distance between Φ and Ψ is defined to be

$$D(\Phi,\Psi) = \frac{1}{N\sqrt{2}}\|\Phi - \Psi\| = \frac{1}{N\sqrt{2}}\|J(\Phi) - J(\Psi)\|_2 \tag{1}$$

The normalizer $N\sqrt{2}$ simply keeps the distance between two quantum channels in $[0,1]$.

Provided the definitions above, we are going to introduce an orthogonal basis.

Definition 5 (Orthogonal Basis for Superoperators) *For any $x, y \in \mathbb{Z}_4^n$, let*

$$\Phi_{x,y}(\rho) = \sigma_x \rho \sigma_y. \tag{2}$$

Proposition 6 *$\{\Phi_{x,y}\}_{x,y\in\mathbb{Z}_4^n}$ forms an orthogonal basis in $(T(X),\langle\cdot,\cdot\rangle)$. Besides, $\|\Phi_{x,y}\| = N$ for all $x, y \in \mathbb{Z}_4^n$.*

The proof is deferred to Appendix B. We are ready to define the Fourier expansions of superoperators now.

Definition 7 (Fourier Expansion of Superoperators) *For superoperator $\Phi \in T(X)$, the Fourier expansion of Φ is defined to be*

$$\Phi = \sum_{x,y\in\mathbb{Z}_4^n} \widehat{\Phi}(x,y)\Phi_{x,y}$$

where $\Phi_{x,y}$ is defined by Eq. (2). $\widehat{\Phi}(x,y)$'s are the Fourier coefficients of Φ and $\widehat{\Phi}(x,y) = \frac{1}{N^2}\langle\Phi_{x,y},\Phi\rangle$. Moreover, we define $\widehat{\Phi}$ to be the $N^2\times N^2$ matrix with entries $\left(\widehat{\Phi}(x,y)\right)_{x,y\in\mathbb{Z}_4^n}$.

Lemma 8 *There exists unitary U such that $\widehat{\Phi} = \frac{1}{N}U^*J(\Phi)U$. Therefore, $\widehat{\Phi}$ is PSD if and only if $J(\Phi)$ is PSD. In particular, if $J(\Phi)$ is PSD, then $\widehat{\Phi}(x,x)\in\mathbb{R}$ for all $x\in\mathbb{Z}_4^n$. For a quantum channel Φ, we have $0\le\widehat{\Phi}(x,x)\le 1$ for all $x\in\mathbb{Z}_4^n$ and $\sum_{x\in\mathbb{Z}_4^n}\widehat{\Phi}(x,x) = \operatorname{Tr}\widehat{\Phi} = 1$.*

The proof is deferred to Appendix C.

3.1. Influence

Given superoperator $\Phi\in T(X)$ and a subset $S\subseteq[n]$, the influence of S on Φ measures how much the qubits in S affect Φ. It is an extension of the influence on operators introduced by Montanaro and Osborne (2010), which, in turn, is inspired by the analogous notion for boolean functions. We will establish several properties of the influence on quantum channels, which enable us to design both testing algorithms and learning algorithms for k-junta channels.

Definition 9 (Influence of superoperators) *Given superoperator $\Phi\in T(X)$, $S\subseteq[n]$, the influence of Φ on S is defined as*

$$\mathrm{Inf}_S[\Phi] = \sum_{x\in\mathbb{Z}_4^n;\,x_S\neq\mathbf{0}} \widehat{\Phi}(x,x).$$

We use $\mathrm{Inf}_i[\Phi]$ to represent $\mathrm{Inf}_{\{i\}}[\Phi]$ for convenience.

Notice that the influence of a superoperator can be negative, which is different from operators in Montanaro and Osborne (2010) or boolean functions. However, we only concern ourselves about completely positive superoperators, whose influence is always nonnegative by Lemma 8.

The following proposition follows from Lemma 8 directly.

Proposition 10 *Given quantum channel $\Phi\in C(X)$, $S\subseteq[n]$, it holds that $0\le\mathrm{Inf}_S[\Phi]\le 1$.*

The following are some basic properties of influence, which can be easily derived from the definition and Lemma 8.

Proposition 11 *Given quantum channel $\Phi\in C(X)$ and $S,T\subseteq[n]$, we have*

1. $S \subseteq T \Rightarrow \mathrm{Inf}_S[\Phi] \leq \mathrm{Inf}_T[\Phi]$;
2. $\mathrm{Inf}_S[\Phi] + \mathrm{Inf}_T[\Phi] \geq \mathrm{Inf}_{S\cup T}[\Phi]$;
3. $\mathrm{Inf}_\varnothing[\Phi] = 0$, $\mathrm{Inf}_{[n]}[\Phi] = 1$.

The following key theorem states that the closeness between a quantum channel and juntas is captured by the influence.

Theorem 12 (Influence and Distance from k-Junta Channels) *Let $\Phi \in C(X)$ be a quantum channel. If there exists a subset $T \subseteq [n]$ satisfying that $\mathrm{Inf}_{T^c}[\Phi] \leq \varepsilon$ for $0 \leq \varepsilon < 1$, then there exists a T-junta channel Φ'' such that $\mathrm{D}(\Phi, \Phi'') \leq \sqrt{\varepsilon} + \varepsilon/\sqrt{2}$.*

To obtain this theorem, we construct a T-junta channel Φ'' explicitly from Φ by two steps. Firstly we construct a T-junta "sub-channel" Φ' and then complement it into a T-junta channel Φ''. The proof of Theorem 12 is deferred to Appendix D.

Corollary 13 *Given quantum channel Φ, if Φ is ε-far from any k-junta channels, then $\mathrm{Inf}_{T^c}[\Phi] \geq \varepsilon^2/4$ for all $T \subseteq [n]$ with $|T| \leq k$.*

3.2. Characterizations of Distance Function

In this section, we will compare the distance given in Eq. (1) with other metrics measuring the distances between two quantum channels. All the proofs in this section can be found in Appendix E.

Chen et al. (2023) introduced a distance $\mathrm{dist}(\cdot,\cdot)$ between unitaries, with which the authors gave optimal testing and learning algorithms for k-junta unitaries. The distance $\mathrm{dist}(\cdot,\cdot)$ is defined as follows.

$$\mathrm{dist}(U, V) = \frac{1}{\sqrt{2N}} \min_{\theta\in[0,2\pi)} \|U - e^{i\theta}V\|_2 \tag{3}$$

The following lemma asserts that the distance $D(\cdot,\cdot)$ in Eq. (1) and $\mathrm{dist}(\cdot,\cdot)$ in Eq. (3) are equivalent when considering unitary operations. Recall that Φ_U is defined in section 2.1.

Lemma 14 (Related to distance between Unitaries) *For unitary matrices U and V, it holds that*

$$\mathit{dist}(U, V) \leq D(\Phi_U, \Phi_V) \leq \sqrt{2}\,\mathit{dist}(U, V).$$

The following proposition proves that $D(\cdot,\cdot)$ captures the average operator distance between two channels. We expect that our distance function could be used in other channel property testing problems.

Proposition 15 (Related to average-case operator distance) *For quantum channels Φ and Ψ, it holds that*

$$\int_\psi \|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 d\psi = \frac{2N}{N+1} D(\Phi,\Psi)^2 + \frac{1}{N(N+1)}\|\Phi(I) - \Psi(I)\|_2^2,$$

where the integral is taken over the Haar measure on all the unit vectors ψ.

Especially for unital channels Φ and Ψ, i.e., $\Phi(I) = \Psi(I) = I$, we have

$$\int_\psi \|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 d\psi = \frac{2N}{N+1} D(\Phi,\Psi)^2$$

Similar properties have been established for the distance $\mathrm{dist}(\cdot,\cdot)$ between two unitaries in Proposition 21 of Montanaro and de Wolf (2016). We refer interested readers to the discussion about the reason for the chosen distances in Section 5.1.1 of Montanaro and de Wolf (2016)

Finally, we prove that $D(\cdot,\cdot)$ can be very far from the worst-case operator norm. Here we consider the 1 to 2 diamond norm.

Definition 16 (1 to 2 Diamond Norm) *Given $\Phi \in T(X)$, its 1 to 2 diamond norm is defined to be*

$$\|\Phi\|_{\diamond,1\to 2} = \|\Phi \otimes \mathbb{1}_X\|_{1\to 2} = \max_{\rho:\|\rho\|_1=1} \{\|(\Phi \otimes \mathbb{1}_X)(\rho)\|_2\}$$

Proposition 17 (Related to worst-case operator distance) *For quantum channels Φ and Ψ, it holds that*

$$\sqrt{2}D(\Phi,\Psi) \leq \|\Phi - \Psi\|_{\diamond,1\to 2} \leq N \cdot \sqrt{2}D(\Phi,\Psi)$$

Both equalities above can be achieved.

4. Testing k-Junta Quantum Channels

In this section, we show an $\widetilde{O}(k)$-query k-junta channel testing algorithm and an $\Omega(\sqrt{k})$ lower bound. First, we prove an upper bound on the sample complexity by presenting a k-junta channel tester, where the analysis of the algorithm relies on the Fourier analysis of superoperators. The lower bound is obtained by reducing k-junta boolean function testing to k-junta channel testing. Finally, we show that the k-junta channels testing problem is the natural extension of k-junta unitary testing problem under our distance function of channels, which gives an alternative proof of the lower bound.

4.1. $\widetilde{O}(k)$ Upper Bound and $\Omega(\sqrt{k})$ Lower Bound

We firstly show our k-junta channel tester. Our tester is inspired by Atici and Servedio (2007) with minor changes.

Algorithm 1: PAULI-SAMPLE(Φ,γ)

Input : Oracle access to quantum channel $\Phi \in C(X)$, γ
Output : $S \subseteq [n]$

1: Initialize $S = \varnothing$;
2: Repeat the following for $O(1/\gamma)$ times;
- Prepare n EPR states and apply Φ to the half of them to obtain $v(\Phi)$;
- Measure all qubits in the Pauli basis, $\{|v(\sigma_x)\rangle\}_{x\in\mathbb{Z}_4^n}$;
- Given the measurement outcome x, set $S \leftarrow S \cup \mathrm{supp}(x)$;

3: Return S.

Theorem 18 (Property of Algorithm 2, Restatement of Theorem 1) *Given quantum channel $\Phi \in C(X)$, with probability at least $9/10$, the algorithm* JUNTA-CHANNEL-TESTER(Φ,k,ε) *outputs "Yes" if Φ is a k-junta, and outputs "No" if Φ is ε-far from any k-junta channel. The algorithm makes $O\big(k\log k/\varepsilon^2\big)$ queries to the channel Φ.*

Algorithm 2: JUNTA-CHANNEL-TESTER(Φ, k, ε)

Input : Oracle access to quantum channel $\Phi \in C(X)$, k, ε
Output : "Yes" or "No"
1: Let $S = $ PAULI-SAMPLE$(\Phi, \varepsilon^2/8k \log k)$;
2: Output "Yes" if $|S| \leq k$, or else output "No".

An algorithm is a (k, ε)-channel junta tester if it can distinguish whether the given channel is k-junta or is ε-far from any k-junta channels. (k, ε)-classical junta testers and (k, ε)-unitary junta testers are defined similarly.

Lemma 19 *A $(k, \sqrt{\varepsilon/2})$-channel junta tester is a (k, ε)-classical junta tester.*

Combining Lemma 19 with the $\Omega(\sqrt{k})$ lower bound on testing k-junta boolean function proved by Bun et al. (2020), we obtain an $\Omega(\sqrt{k})$ lower bound on testing k-junta channels. Our key technical lemma is as follows. Recall that Φ_g is defined in Section 2.1 for boolean function g.

Lemma 20 *For a k-junta channel Φ, there exists a k-junta boolean function g' satisfying that $D(\Phi, \Phi_{g'}) = \min_g D(\Phi, \Phi_g)$, where the minimization is over all boolean functions $g : \{0,1\}^n \to \{0,1\}$.*

With the assistance of this result around the distance structure of k-junta channels, we obtain the desired reduction in Lemma 19. See Appendix G for the detailed proofs.

4.2. Reduction from k-Junta Unitary Testing

To show our distance function induced by Fourier analysis over superoperators is a natural extension of the distance function on unitaries discussed in Montanaro and de Wolf (2016), we provide an extra reduction from k-junta unitary testing. It gives an alternative proof of our testing lower bound. All the proofs can be found in Appendix G.

Lemma 21 (Reduction from Testing k-Junta Unitaries to Testing k-Junta Channels) *A (k, ε)-channel junta tester is naturally a $(k, \varepsilon/2)$-unitary junta tester.*

The key technical result is as follows:

Lemma 22 *For every k-junta channel Φ', there exists a k-junta unitary V, such that $D(\Phi', \Phi_V) = \min_V D(\Phi', \Phi_V)$, where the minimization is over all unitaries V.*

5. Learning k-Junta Quantum Channels

In this section, we prove a nearly tight bound on k-junta learning problem. Our algorithm is inspired by the learning algorithms in Atici and Servedio (2007) and Chen et al. (2023). We describe the algorithm JUNTA-CHANNEL-LEARNER as follows.

Algorithm 3: JUNTA-CHANNEL-LEARNER(Φ, k, ε)

Input : Oracle access to k-junta channel $\Phi \in C(X)$, ε

Output : A classical description of Φ in the form of its Choi representation, a $4^n \times 4^n$ matrix

1: Let $S =$ PAULI-SAMPLE$(\Phi, \varepsilon^2/8k \log k)$;

2: Set $t = O(4^k/\varepsilon^2)$. Call QUANTUM-STATE-PREPARATION(Φ, S) for $10t$ times to obtain at least t copies of quantum state ψ_S;

3: Return CHANNEL-TOMOGRAPHY$(\psi_S^{\otimes t}, \varepsilon) \otimes v(I^{S^c})$ as the result.

Algorithm 4: QUANTUM-STATE-PREPARATION$(\Phi, S \subseteq [n])$

Input : Oracle access to k-junta channel $\Phi \in C(X)$, γ

Output : A $2|S|$-qubit quantum state, or "error"

1: Prepare the state $v(\Phi)$;

2: Measure $2|S^c|$ qubits in S^c onto the Pauli basis $\{|\sigma_x\rangle\}_{x \in \mathbb{Z}_4^{|S^c|}}$;

3: If the measurement result is $\mathbf{0}^{S^c}$, return the untouched $2|S|$ qubits. Otherwise, return "error".

Theorem 23 (Property of Algorithm 3, Restatement of Theorem 1) *Given oracle access to k-junta channel Φ, with probability at least $9/10$,* JUNTA-CHANNEL-LEARNER(Φ, k, ε) *outputs the description of quantum channel Ψ such that $D(\Phi, \Psi) \leq \varepsilon$. Furthermore, this algorithm makes $O(4^k/\varepsilon^2)$ queries.*

As for the k-junta channel learning lower bound, recall Lemma 14 shows that our distance function over channels is equivalent to the distance between unitaries used in Chen et al. (2023), up to a constant factor, it is very natural to reduce learning k-junta unitaries to learning k-junta channels, and therefore the following lower bound follows.

Theorem 24 (Lower Bound on Learning k-Junta Channels) *Any algorithm learning k-junta channels within precision ε under $D(\cdot, \cdot)$ requires $\Omega(4^k \log(1/\varepsilon)/k)$ queries.*

6. Conclusion

We exhibit two algorithms, one for testing k-junta channels and one for learning k-junta channels and lower bounds respectively. The k-junta channel learning algorithm is nearly optimal. Our algorithms generalize the work Atici and Servedio (2007); Chen et al. (2023) about testing and learning k-junta unitaries and k-junta boolean function. To design the algorithms and prove the lower bounds, we introduce the Fourier analysis over the space of superoperators, which extends the Fourier analysis over operators in Montanaro and Osborne (2010). As Montanaro and de Wolf (2016) mentioned, there was not much work on testing the properties of quantum channels. We expect more applications in designing algorithms for testing and learning quantum channels through the Fourier analysis presented in this paper.

Algorithm 5: CHANNEL-TOMOGRAPHY($\psi^{\otimes O(4^k/\varepsilon^2)}$, ε)

Input : Independent copies of ψ and ε, enough for TOMOGRAPHY

Output : A classical description of ψ

1: Run TOMOGRAPHY($\psi^{\otimes O(4^k/\varepsilon^2)}, 0.04\varepsilon$) to obtain a description of state ψ;

2: Find out, by only local calculation, the Choi state closest to ψ and return the description.

Acknowledgments

We would thank Jingquan Luo for pointing out an error in the previous version of this work. We thank the anonymous reviewers for their careful reading and helpful comments. We also would like to express our gratitude for the insightful discussions with Eric Blais and Nengkun Yu. The first author would also like to extend special thanks to Minglong Qing, Mingnan Zhao and Haochen Xu for their invaluable support in problem-solving and the writing of this paper. This work was supported by National Natural Science Foundation of China (Grant No. 61972191) and Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0302900).

References

Maryam Aliakbarpour, Eric Blais, and Ronitt Rubinfeld. Learning and testing junta distributions. In Vitaly Feldman, Alexander Rakhlin, and Ohad Shamir, editors, *Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016*, volume 49 of *JMLR Workshop and Conference Proceedings*, pages 19–46. JMLR.org, 2016.

Andris Ambainis, Aleksandrs Belovs, Oded Regev, and Ronald de Wolf. Efficient quantum algorithms for (gapped) group testing and junta testing. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 903–922. ACM, New York, 2016.

Jan Arpe and Elchanan Mossel. Agnostically learning juntas from random walks. *CoRR*, abs/0806.4210, 2008.

Jan Arpe and Rüdiger Reischuk. Learning juntas in the presence of noise. *Theoret. Comput. Sci.*, 384(1):2–21, 2007.

V. Arvind, Johannes Köbler, and Wolfgang Lindner. Parameterized learnability of juntas. *Theoret. Comput. Sci.*, 410(47-49):4928–4936, 2009.

Alp Atici and Rocco A. Servedio. Quantum algorithms for learning and testing juntas. *Quantum Inf. Process.*, 6(5):323–348, 2007.

Arnab Bhattacharyya and Yuichi Yoshida. *Property testing—problems and techniques*. Springer, Singapore, 2022.

Eric Blais. Testing juntas nearly optimally. In *STOC'09—Proceedings of the 2009 ACM International Symposium on Theory of Computing*, pages 151–157. ACM, New York, 2009.

Eric Blais, Clément L. Canonne, Talya Eden, Amit Levi, and Dana Ron. Tolerant junta testing and the connection to submodular optimization and function isomorphism. *ACM Trans. Comput. Theory*, 11(4):Art. 24, 33, 2019.

Nader H. Bshouty and Areej Costa. Exact learning of juntas from membership queries. *Theoret. Comput. Sci.*, 742:82–97, 2018.

Mark Bun, Robin Kothari, and Justin Thaler. The polynomial method strikes back: tight quantum query bounds via dual polynomials. *Theory Comput.*, 16:Paper No. 10, 71, 2020.

Thomas Chen, Shivam Nadimpalli, and Henry Yuen. Testing and learning quantum juntas nearly optimally. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1163–1185. SIAM, 2023.

Xi Chen, Rocco A. Servedio, Li-Yang Tan, Erik Waingarten, and Jinyu Xie. Settling the query complexity of non-adaptive junta testing. *J. ACM*, 65(6):Art. 40, 18, 2018.

Xi Chen, Rajesh Jayaram, Amit Levi, and Erik Waingarten. Learning and testing junta distributions with sub cube conditioning. In Mikhail Belkin and Samory Kpotufe, editors, *Conference on Learning Theory, COLT 2021, 15-19 August 2021, Boulder, Colorado, USA*, volume 134 of *Proceedings of Machine Learning Research*, pages 1060–1113. PMLR, 2021.

Hana Chockler and Dan Gutfreund. A lower bound for testing juntas. *Inform. Process. Lett.*, 90(6): 301–305, 2004.

Isaac L Chuang and Michael A Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

Anindya De, Elchanan Mossel, and Joe Neeman. Junta correlation is testable. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science*, pages 1549–1563. IEEE Comput. Soc. Press, Los Alamitos, CA, 2019.

Ronald de Wolf. A brief introduction to fourier analysis on the boolean cube. *Theory Comput.*, 1: 1–20, 2008.

Eldar Fischer, Guy Kindler, Dana Ron, Shmuel Safra, and Alex Samorodnitsky. Testing juntas. *J. Comput. System Sci.*, 68(4):753–787, 2004.

Oded Goldreich. *Introduction to property testing*. Cambridge University Press, Cambridge, 2017.

Gus Gutoski and Nathaniel Johnston. Process tomography for unitary quantum channels. *J. Math. Phys.*, 55(3):032201, 19, 2014.

Amit Levi and Erik Waingarten. Lower bounds for tolerant junta and unateness testing via rejection sampling of graphs. In *10th Innovations in Theoretical Computer Science*, volume 124 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 52, 20. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2019.

Richard J. Lipton, Evangelos Markakis, Aranyak Mehta, and Nisheeth K. Vishnoi. On the fourier spectrum of symmetric boolean functions with applications to learning symmetric juntas. In *20th Annual IEEE Conference on Computational Complexity (CCC 2005), 11-15 June 2005, San Jose, CA, USA*, pages 112–119. IEEE Computer Society, 2005.

Zhengyang Liu, Xi Chen, Rocco A. Servedio, Ying Sheng, and Jinyu Xie. Distribution-free Junta testing. *ACM Trans. Algorithms*, 15(1):Art. 1, 23, 2019.

Ashley Montanaro and Ronald de Wolf. *A Survey of Quantum Property Testing*. Number 7 in Graduate Surveys. Theory of Computing Library, 2016.

Ashley Montanaro and Tobias J. Osborne. Quantum Boolean functions. *Chic. J. Theoret. Comput. Sci.*, pages Article 1, 41, 2010.

Elchanan Mossel, Ryan O'Donnell, and Rocco P. Servedio. Learning juntas. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, pages 206–212. ACM, New York, 2003.

Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.

Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.

Ryan O'Donnell and John Wright. Efficient quantum tomography II. In *STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 962–974. ACM, New York, 2017.

Michal Parnas, Dana Ron, and Alex Samorodnitsky. Testing basic Boolean formulae. *SIAM J. Discrete Math.*, 16(1):20–46, 2002.

Mert Sağlam. Near log-convexity of measured heat in (discrete) time and consequences. In *59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018*, pages 967–978. IEEE Computer Soc., Los Alamitos, CA, 2018.

Rocco A. Servedio, Li-Yang Tan, and John Wright. Adaptivity helps for testing juntas. In *30th Conference on Computational Complexity*, volume 33 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 264–279. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2015.

Guoming Wang. Property testing of unitary operators. *Physical Review A, Atomic, molecular, and optical physics*, 84(5):052328, November 2011.

John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

Appendix A. Properties on Kraus and Choi Representations

In this section, we list some basic properties of Kraus and Choi representations, whose proofs can be found in Section 2.2 in Watrous (2018).

Fact 25 *Given superoperators $\Phi \in T(X), \Phi' \in T(X')$, it holds that*

1. *Φ is completely positive if and only if it has a Kraus representation $\Phi(\rho) = \sum_{s\in\Sigma} A_s \rho A_s^*$. It is trace preserving if and only if its Kraus representation $\Phi(\rho) = \sum_{s\in\Sigma} A_s \rho B_s^*$. satisfies that $\sum_{s\in\Sigma} B_s^* A_s = I$;*

2. *Φ is completely positive if and only if $J(\Phi)$ is PSD. It is trace preserving if and only if $\mathrm{Tr}_{X_1} J(\Phi) = I_{X_2}$, where J is viewed as a map from $L(X)$ to $L(X_1) \otimes L(X_2)$ with $X_1 = X_2 = X$;*

3. *If* $\Phi(\rho) = \sum_{s\in\Sigma} A_s\rho B_s^*$, *we have*

$$J(\Phi) = \sum_{s\in\Sigma} vec(A_s)vec(B_s)^*;$$

4. $J(\Phi \otimes \Phi') = J(\Phi) \otimes J(\Phi')$.

Appendix B. Fourier Basis of Superoperators Is Well-defined

Here we list some basic properties of the inner product and the norm introduced in Section 3, which are easy to verify by the definitions.

Fact 26 (Properties of Inner Product and Norm)

1. *Given* $\Phi(\rho) = A\rho B^*$ *and* $\Psi(\rho) = C\rho D^*$, *we have* $\langle \Phi, \Psi\rangle = \langle A, C\rangle \cdot \langle D, B\rangle$;
2. *For* $\Phi(\rho) = A\rho B^*$, *we have* $\|\Phi\| = \sqrt{\langle \Phi, \Phi\rangle} = \sqrt{\langle A, A\rangle \cdot \langle B, B\rangle} = \|A\|_2 \cdot \|B\|_2$;
3. *Suppose* $\Phi = \Phi_1 \otimes \Phi_2$ *and* $\Psi = \Psi_1 \otimes \Psi_2$. *We have* $\langle \Phi, \Psi\rangle = \langle \Phi_1, \Psi_1\rangle \cdot \langle \Phi_2, \Psi_2\rangle$.

We are going to prove Proposition 6 now.

Proposition 6 $\{\Phi_{x,y}\}_{x,y\in\mathbb{Z}_4^n}$ *forms an orthogonal basis in* $(T(X), \langle\cdot,\cdot\rangle)$. *Besides,* $\|\Phi_{x,y}\| = N$ *for all* $x, y \in \mathbb{Z}_4^n$.

Proof

Norm. $\forall x, y \in \mathbb{Z}_4^n$, $\|\Phi_{x,y}\| = \|\sigma_x\|_2\|\sigma_y\|_2 = N$ using Fact 26.
Orthogonality. $\forall x, x', y, y' \in \mathbb{Z}_4^n$,$x \neq x'$ or $y \neq y'$, we have

$$\begin{aligned}\left\langle \Phi_{x,y}, \Phi_{x',y'}\right\rangle &= \prod_{i\in[n]} \left\langle \Phi_{x_i,y_i}, \Phi_{x_i',y_i'}\right\rangle \\ &= \prod_{i\in[n]} \left\langle \sigma_{x_i}, \sigma_{x_i'}\right\rangle \cdot \left\langle \sigma_{y_i'}, \sigma_{y_i}\right\rangle \\ &= 0\end{aligned}$$

All equalities follow from Fact 26 directly. Note that for non-zero vectors, orthogonality implies linear independence.

Basis, spanning the whole space. The dimension of $T(X)$ is $N^4 = 2^{4n}$ and we have $4^{2n} = 2^{4n}$ linearly independent vectors in $\{\Phi_{x,y}\}_{x,y\in\mathbb{Z}_4^n}$.

■

Appendix C. Properties of Fourier Expansions of Superoperators

Lemma 8 *There exists unitary U such that $\widehat{\Phi} = \frac{1}{N}U^*J(\Phi)U$. Therefore, $\widehat{\Phi}$ is PSD if and only if $J(\Phi)$ is PSD. In particular, if $J(\Phi)$ is PSD, then $\widehat{\Phi}(x,x) \in \mathbb{R}$ for all $x \in \mathbb{Z}_4^n$. For a quantum channel Φ, we have $0 \le \widehat{\Phi}(x,x) \le 1$ for all $x \in \mathbb{Z}_4^n$ and $\sum_{x\in\mathbb{Z}_4^n} \widehat{\Phi}(x,x) = \operatorname{Tr}\widehat{\Phi} = 1$.*

Proof By the definition of $\widehat{\Phi}$, we have

$$\begin{aligned}\widehat{\Phi}(x,y) &= \frac{1}{N^2}\langle \Phi_{x,y}, \Phi\rangle \\ &= \frac{1}{N^2}\operatorname{Tr}(\mathrm{vec}(\sigma_x)\mathrm{vec}(\sigma_y)^*)^* J(\Phi) \\ &= \frac{1}{N^2}\mathrm{vec}(\sigma_x)^* J(\Phi)\mathrm{vec}(\sigma_y),\end{aligned}$$

where the second equality is by the definition of the inner product and the fact that $J(\Phi_{x,y}) = \mathrm{vec}(\sigma_x)\mathrm{vec}(\sigma_y)^*$. Therefore

$$\widehat{\Phi} = \frac{1}{N}U^*J(\Phi)U$$

where $U = [vec(\sigma_x)/\sqrt{N}]_{x\in\mathbb{Z}_4^n}$ is a unitary.

■

The next corollary follows from the properties of Kraus and Choi representations in Fact 25. We note that $\Phi(\rho) = \sum_{x,y\in\mathbb{Z}_4^n} \widehat{\Phi}(x,y)\sigma_x\rho\sigma_y$ is a Kraus representation of Φ. Therefore $\Phi \in T(X)$ is trace preserving if and only if $\sum_{x,y\in\mathbb{Z}_4^n} \widehat{\Phi}(x,y)\sigma_y\sigma_x = I$.

Corollary 27 *Let $\Phi \in T(X)$ be a superoperator. The following statements are equivalent.*

1. *$\Phi \in T(X)$ is completely positive.*
2. *$\widehat{\Phi}$ is PSD.*

The following statements are equivalent as well.

1. *$\Phi \in T(X)$ is trace preserving.*
2. *$\sum_{x,y\in\mathbb{Z}_4^n} \widehat{\Phi}(x,y)\sigma_y\sigma_x = I$.*

Corollary 28 (Relations between Fourier Expansion and Norm and Distance) *Let $\Phi, \Psi \in T(X)$ be superoperators and $\widehat{\Phi}$, $\widehat{\Psi}$ be the corresponding Fourier expansions. Then*

1. *$\|\Phi\| = N\left\|\widehat{\Phi}\right\|_2 = N\sqrt{\sum_{x,y\in\mathbb{Z}_4^n}\left|\widehat{\Phi}(x,y)\right|^2}$;*
2. *$\mathrm{D}(\Phi,\Psi) = \frac{1}{\sqrt{2}}\left\|\widehat{\Phi} - \widehat{\Psi}\right\|_2 = \frac{1}{\sqrt{2}}\sqrt{\sum_{x,y\in\mathbb{Z}_4^n}\left|\widehat{\Phi}(x,y) - \widehat{\Psi}(x,y)\right|^2}$.*

Appendix D. Proof of Theorem 12

Theorem 12 (Influence and Distance from k-Junta Channels) *Let $\Phi \in C(X)$ be a quantum channel. If there exists a subset $T \subseteq [n]$ satisfying that $\mathrm{Inf}_{T^c}[\Phi] \le \varepsilon$ for $0 \le \varepsilon < 1$, then there exists a T-junta channel Φ'' such that $\mathrm{D}(\Phi, \Phi'') \le \sqrt{\varepsilon} + \varepsilon/\sqrt{2}$.*

Proof We need two steps to construct Φ'' explicitly. Firstly we construct a k-junta *sub-channel* Φ', which is completely positive and trace non-increasing, and then turn it to a channel Φ''.

Construction of *sub-channel* Φ' Let

$$\Phi'(\rho) = \sum_{x,y \in \mathbb{Z}_4^n; x_{T^c} = y_{T^c} = \mathbf{0}} \widehat{\Phi}(x,y) \sigma_x \rho \sigma_y$$

Notice that Φ is a quantum channel. By Fact 26 and Corollary 27, it is easy to see Φ' is a T-junta sub-channel. Notice that $\widehat{\Phi}'$ is a principle submatrix of PSD matrix $\widehat{\Phi}$, which implies $\widehat{\Phi}'$ is PSD. Then again by Corollary 27, $J(\widehat{\Phi}')$ is also PSD. Now we bound the distance between Φ and Φ' from above.

By Corollary 28, we have

$$2 \cdot \mathrm{D}(\Phi, \Phi')^2 = \sum_{x,y \in \mathbb{Z}_4^n; x_{T^c} \neq \mathbf{0} \text{ or } y_{T^c} \neq \mathbf{0}} \left| \widehat{\Phi}(x,y) \right|^2$$

For any $x, y \in \mathbb{Z}_4^n$ we have $\left| \widehat{\Phi}(x,y) \right|^2 \le \widehat{\Phi}(x,x)\widehat{\Phi}(y,y)$ since $\widehat{\Phi}$ is a PSD matrix. This implies

$$\begin{aligned} \sum_{x,y \in \mathbb{Z}_4^n; x_{T^c} \neq \mathbf{0} \text{ or } y_{T^c} \neq \mathbf{0}} \left| \widehat{\Phi}(x,y) \right|^2 &\le \sum_{x,y \in \mathbb{Z}_4^n : x_{T^c} \neq \mathbf{0} \text{ or } y_{T^c} \neq \mathbf{0}} \widehat{\Phi}(x,x)\widehat{\Phi}(y,y) \\ &\le \left(\sum_{x,y \in \mathbb{Z}_4^n : x_{T^c} \neq \mathbf{0}} + \sum_{x,y \in \mathbb{Z}_4^n : y_{T^c} \neq \mathbf{0}} \right) \widehat{\Phi}(x,x)\widehat{\Phi}(y,y). \qquad (4) \end{aligned}$$

Notice $\sum_{x \in \mathbb{Z}_4^n} \widehat{\Phi}(x,x) = 1$ by Lemma 8. We have

$$\text{RHS of Eq. (4)} = 2 \sum_{x \in \mathbb{Z}_4^n; x_{T^c} \neq \mathbf{0}} \widehat{\Phi}(x,x)$$

To summarize, we have

$$2 \cdot \mathrm{D}(\Phi, \Phi')^2 \le 2 \sum_{x \in \mathbb{Z}_4^n; x_{T^c} \neq \mathbf{0}} \widehat{\Phi}(x,x) = 2 \cdot \mathrm{Inf}_{T^c}[\Phi] \le 2\varepsilon$$

We claim that $\sum_{x,y \in \mathbb{Z}_4^n; x_{T^c} = y_{T^c} = \mathbf{0}} \widehat{\Phi}(x,y) \sigma_y \sigma_x \le I$.

Let

$$A = \sum_{x,y\in\mathbb{Z}_4^n, x_{T^c}=y_{T^c}=\mathbf{0}} \widehat{\Phi}(x,y)\sigma_y\sigma_x, \tag{5}$$
$$B = \sum_{x,y\in\mathbb{Z}_4^n, x_{T^c}\neq\mathbf{0}, y_{T^c}\neq\mathbf{0}} \widehat{\Phi}(x,y)\sigma_y\sigma_x,$$
$$C_1 = \sum_{x,y\in\mathbb{Z}_4^n, x_{T^c}=\mathbf{0}, y_{T^c}\neq\mathbf{0}} \widehat{\Phi}(x,y)\sigma_y\sigma_x,$$
$$C_2 = \sum_{x,y\in\mathbb{Z}_4^n, x_{T^c}\neq\mathbf{0}, y_{T^c}=\mathbf{0}} \widehat{\Phi}(x,y)\sigma_y\sigma_x.$$

We note that $A = \sum_{x',y'\in\mathbb{Z}_4^T} \widehat{\Phi}(x'\circ 0^{T^c}, y'\circ 0^{T^c})\sigma_{y'}\sigma_{x'}\otimes I^{T^c} =: A'\otimes I^{T^c}$, where $x'\circ 0^{T^c}$ is the concatenation of x' and 0^{T^c}. Same for $y'\circ 0^{T^c}$. To see $A\leq I$, it is enough to show $A'\leq I$, which is equivalent to $\operatorname{Tr} A' \left|\phi\right\rangle\!\left\langle\phi\right| \leq 1$ for any quantum state $|\phi\rangle_T$.

Let I_{T^c} be a $2^{|T^c|}\times 2^{|T^c|}$ identity matrix. Notice that $\operatorname{Tr} A(\left|\phi\right\rangle\!\left\langle\phi\right| \otimes \frac{I_{T^c}}{2^{|T^c|}}) = \operatorname{Tr}(A'\cdot \left|\phi\right\rangle\!\left\langle\phi\right|)$. It suffices to prove that $\operatorname{Tr} A(\left|\phi\right\rangle\!\left\langle\phi\right| \otimes \frac{I_{T^c}}{2^{|T^c|}}) \leq 1$ for arbitrary quantum state $|\phi\rangle$. To this end,

$$\begin{aligned}
1 &= \operatorname{Tr}\Phi(\left|\phi\right\rangle\!\left\langle\phi\right| \otimes \frac{I_{T^c}}{2^{|T^c|}}) \\
&= \frac{1}{2^{|T^c|}}(\operatorname{Tr} A(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) + \operatorname{Tr} B(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) + \operatorname{Tr} C_1(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) + \operatorname{Tr} C_2(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c})) \\
&\geq \frac{1}{2^{|T^c|}}(\operatorname{Tr} A(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) + \operatorname{Tr} C_1(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) + \operatorname{Tr} C_2(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c})) \\
&= \frac{1}{2^{|T^c|}}(\operatorname{Tr} A(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}))
\end{aligned}$$

where the first inequality is because B is a principle sub-matrix of $\widehat{\Phi}$, which is also PSD by Corollary 27; the last equality is because $\operatorname{Tr} C_1(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) + \operatorname{Tr} C_2(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) = 0$. To see this, we will prove that $\operatorname{Tr} C_1(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) = 0$.

$$\begin{aligned}
\operatorname{Tr} C_1(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) &= \sum_{x,y\in\mathbb{Z}_4^n, x_{T^c}=\mathbf{0}, y_{T^c}\neq\mathbf{0}} \widehat{\Phi}(x,y)\operatorname{Tr}\sigma_y\sigma_x(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) \\
&= \sum_{x,y\in\mathbb{Z}_4^n, x_{T^c}=\mathbf{0}, y_{T^c}\neq\mathbf{0}} \widehat{\Phi}(x,y)\left\langle\phi\right|\sigma_{y_T}\sigma_{x_T}\left|\phi\right\rangle\cdot\langle\sigma_{y_{T^c}},\sigma_{x_{T^c}}\rangle \\
&= 0
\end{aligned}$$

$\operatorname{Tr} C_2(\left|\phi\right\rangle\!\left\langle\phi\right|\otimes I_{T^c}) = 0$ follows from the same argument. Therefore $A\leq I$.

Construction of *Channel* Φ'' We set

$$\Phi''(\rho) = \Phi'(\rho) + \sqrt{I-A}\rho\sqrt{I-A},$$

where A is given in Eq. (5). By Corollary 27, we have $J(\Phi'') = J(\Phi') + \operatorname{vec}(\sqrt{I-A})\operatorname{vec}(\sqrt{I-A})^*$, which is PSD. Notice that $A = A'\otimes I_{T^c}$. Thus Φ'' is also a T-junta completely positive map. To prove Φ'' is a channel, it suffices to prove that Φ'' is trace-preserving. By the Kraus representation of Φ''

$$\Phi''(\rho) = \sum_{x,y\in\mathbb{Z}_4^n; x_{T^c}=y_{T^c}=\mathbf{0}} \widehat{\Phi}(x,y)\sigma_x\rho\sigma_y + \sqrt{I-A}\rho\sqrt{I-A},$$

we have

$$\sum_{x,y\in\mathbb{Z}_4^n;x_{T^c}=y_{T^c}=\mathbf{0}} \widehat{\Phi}(x,y)\sigma_y\sigma_x + \sqrt{I-A}\sqrt{I-A} = A + \sqrt{I-A}\sqrt{I-A} = I,$$

which implies Φ'' is trace preserving according to Fact 25.

Next we bound the distance between Φ' and Φ'' from above. Note that $J(\Phi'') = J(\Phi') + \mathrm{vec}\left(\sqrt{I-A}\right)\mathrm{vec}\left(\sqrt{I-A}\right)^*$, we have

$$\begin{aligned} D(\Phi'',\Phi') &= \frac{1}{N\sqrt{2}}\left\|J(\Phi'')-J(\Phi')\right\|_2 = \frac{1}{N\sqrt{2}}\left\|\mathrm{vec}\left(\sqrt{I-A}\right)\mathrm{vec}\left(\sqrt{I-A}\right)^*\right\|_2 \\ &= \frac{1}{N\sqrt{2}}\left\|\sqrt{I-A}\right\|_2^2 = \frac{1}{N\sqrt{2}}\operatorname{Tr}(I-A) \end{aligned}$$

From the definition of A, we have $\frac{1}{N}\operatorname{Tr} A = \sum_{x\in\mathbb{Z}_4^n,x_{T^c}=\mathbf{0}} \widehat{\Phi}(x,x) = 1 - \mathrm{Inf}_{T^c}[\Phi]$, which implies

$$\frac{1}{N\sqrt{2}}\operatorname{Tr}(I-A) = \frac{1}{\sqrt{2}}\mathrm{Inf}_{T^c}[\Phi].$$

Therefore

$$D(\Phi'',\Phi') \leq \frac{1}{\sqrt{2}}\mathrm{Inf}_{T^c}[\Phi] \leq \frac{\varepsilon}{\sqrt{2}}.$$

In conclusion, Φ'' is a T-junta channel and $D(\Phi,\Phi'') \leq \sqrt{\varepsilon} + \varepsilon/\sqrt{2}$, which completes the proof.

■

Appendix E. Characterization of Distance Function

E.1. Proof of Lemma 14

Lemma 14 (Related to distance between Unitaries) *For unitary matrices U and V, it holds that*

$$\mathit{dist}(U,V) \leq D(\Phi_U,\Phi_V) \leq \sqrt{2}\,\mathit{dist}(U,V).$$

Proof It's easy to see

$$\mathrm{dist}(U,V) = \sqrt{1-\frac{1}{N}|\langle U,V\rangle|}$$

and

$$D(\Phi_U,\Phi_V) = \sqrt{1-\frac{1}{N^2}|\langle U,V\rangle|^2}$$

Let $\alpha = \frac{1}{N}|\langle U,V\rangle| \in [0,1]$. Lemma 14 follows from the inequality $\sqrt{1-\alpha} \leq \sqrt{1-\alpha^2} \leq \sqrt{2}\sqrt{1-\alpha}$.

■

E.2. Comparison with other operator norms

Proposition 15 (Related to average-case operator distance) *For quantum channels Φ and Ψ, it holds that*

$$\int_\psi \|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 d\psi = \frac{2N}{N+1} D(\Phi,\Psi)^2 + \frac{1}{N(N+1)}\|\Phi(I) - \Psi(I)\|_2^2,$$

where the integral is taken over the Haar measure on all the unit vectors ψ.

Especially for unital channels Φ and Ψ, i.e., $\Phi(I) = \Psi(I) = I$, we have

$$\int_\psi \|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 d\psi = \frac{2N}{N+1} D(\Phi,\Psi)^2$$

Proof Let $J = J(\Phi) - J(\Psi) = J(\Phi - \Psi) = \sum_{i,j\in[N]} J_{i,j} \otimes |i\rangle\langle j|$, where $J_{i,j} = (\Phi - \Psi)(|i\rangle\langle j|)$.

$$\begin{aligned}\|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 &= \|\mathrm{Tr}_{X_2}(J \cdot (I \otimes |\psi\rangle\langle\psi|))\|_2^2 \\ &= \left\|\sum_{i,j\in[N]} J_{i,j} \langle j|\psi\rangle \langle\psi|i\rangle\right\|_2^2 \\ &= \sum_{i,j,i',j'\in[N]} \langle J_{i,j}, J_{i',j'}\rangle \langle i|\psi\rangle \langle\psi|j\rangle \langle j'|\psi\rangle \langle\psi|i'\rangle\end{aligned}$$

Note that $\langle i|\psi\rangle \langle\psi|j\rangle \langle j'|\psi\rangle \langle\psi|i'\rangle = \mathrm{Tr}\,(|j\rangle\langle i| \otimes |i'\rangle\langle j'|) \cdot (|\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|)$, we have

$$\begin{aligned}&\int_\psi \|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 d\psi \\ =& \sum_{i,j,i',j'\in[N]} \langle J_{i,j}, J_{i',j'}\rangle \mathrm{Tr}\left(|j\rangle\langle i| \otimes |i'\rangle\langle j'|\right) \cdot \int_\psi |\psi\rangle\langle\psi|^{\otimes 2} d\psi \\ =& \sum_{i,j,i',j'\in[N]} \langle J_{i,j}, J_{i',j'}\rangle \mathrm{Tr}\left(|j\rangle\langle i| \otimes |i'\rangle\langle j'|\right) \cdot \frac{I+F}{N(N+1)} \\ =& \frac{1}{N(N+1)}\left(\sum_{i,j\in[N]} \langle J_{i,j}, J_{i,j}\rangle + \sum_{i,j\in[N]} \langle J_{i,i}, J_{j,j}\rangle\right)\end{aligned}$$

In the second equality we use the fact that $\int_\psi |\psi\rangle\langle\psi|^{\otimes 2} d\psi = (I + F)/N(N + 1)$, where F is the swap operator which interchanges two n-qubit quantum systems; see Lemma 7.24 of Watrous (2018). The third equality follows from $\mathrm{Tr}((A \otimes B)F) = \mathrm{Tr}\, AB$.

By the definition of J, we have

$$\|J(\Phi) - J(\Psi)\|_2^2 = \|J\|_2^2 = \langle J, J\rangle = \sum_{i,j\in[N]} \langle J_{i,j}, J_{i,j}\rangle$$

$$\|\Phi(I) - \Psi(I)\|_2^2 = \left\|\sum_{i\in[N]} J_{i,i}\right\|_2^2 = \left\langle \sum_{i\in[N]} J_{i,i}, \sum_{j\in[N]} J_{j,j}\right\rangle = \sum_{i,j\in[N]} \langle J_{i,i}, J_{j,j}\rangle$$

and therefore

$$\begin{aligned}
&\int_{\psi} \|\Phi(|\psi\rangle\langle\psi|) - \Psi(|\psi\rangle\langle\psi|)\|_2^2 d\psi \\
=&\frac{1}{N(N+1)}\left(\sum_{i,j\in[N]} \langle J_{i,j}, J_{i,j}\rangle + \sum_{i,j\in[N]} \langle J_{i,i}, J_{j,j}\rangle\right) \\
=&\frac{1}{N(N+1)}\left(\|J(\Phi) - J(\Psi)\|_2^2 + \|\Phi(I) - \Psi(I)\|_2^2\right) \\
=&\frac{2N}{N+1} D(\Phi,\Psi)^2 + \frac{1}{N(N+1)}\|\Phi(I) - \Psi(I)\|_2^2
\end{aligned}$$

■

Proposition 17 (Related to worst-case operator distance) *For quantum channels Φ and Ψ, it holds that*

$$\sqrt{2}D(\Phi,\Psi) \le \|\Phi - \Psi\|_{\diamond,1\to 2} \le N \cdot \sqrt{2}D(\Phi,\Psi)$$

Both equalities above can be achieved.

Proof We will show

$$\frac{1}{N}\|J(\Phi) - J(\Psi)\|_2 \le \|(\Phi - \Psi)\otimes \mathbb{1}_X\|_{1\to 2} \le \|J(\Phi) - J(\Psi)\|_2.$$

Let $|\Psi_0\rangle = \frac{1}{\sqrt{N}}\sum_{i\in X}|ii\rangle$, we have

$$\frac{1}{N}\|J(\Phi) - J(\Psi)\|_2 = \|((\Phi-\Psi)\otimes I_X)(|\Psi_0\rangle\langle\Psi_0|)\|_2 \le \|(\Phi - \Psi)\otimes I_X\|_{1\to 2}$$

and the first inequality follows immediately. To prove the next inequality, By the following fact, the $1 \to 2$ diamond norm can be achieved by a rank-1 Hermitian matrix.

Fact 29 (Theorem 3.51 in Watrous (2018)) *There exists an unit vector $u \in X \otimes X$, which satisfies that $\|\Phi - \Psi\|_{\diamond,1\to 2} = \|((\Phi-\Psi)\otimes I_X)(uu^*)\|_2$.*

Let u be the unit vector in Fact 29 and A be a matrix satisfying that $u = \mathrm{vec}(A) = \sqrt{N}(A \otimes I)|\Psi_0\rangle = \sqrt{N}(I \otimes A^T)|\Psi_0\rangle$. We have

$$\begin{aligned}
\|\Phi - \Psi\|_{\diamond,1\to 2} &= \|((\Phi-\Psi)\otimes I)(\mathrm{vec}(A)\mathrm{vec}(A)^*)\|_2 \\
&= N \cdot \left\|((\Phi-\Psi)\otimes I)((I\otimes A^T)|\Psi_0\rangle\langle\Psi_0|(I\otimes A^T)^*)\right\|_2 \\
&= N \cdot \left\|(I\otimes A^T)((\Phi-\Psi)\otimes I)(|\Psi_0\rangle\langle\Psi_0|)(I\otimes A^T)^*\right\|_2 \\
&= \left\|(I\otimes A^T)(J(\Phi) - J(\Psi))(I\otimes A^T)^*\right\|_2
\end{aligned}$$

Applying the norm inequality $\|ABC\|_2 \le \|A\|_\infty \|B\|_2 \|C\|_\infty$ and $\left\|I\otimes A^T\right\|_\infty = \left\|A^T\right\|_\infty = \|A\|_\infty \le \|A\|_2$, we have

$$\begin{aligned}
\left\|(I\otimes A^T)(J(\Phi) - J(\Psi))(I\otimes A^T)^*\right\|_2 &\le \left\|I\otimes A^T\right\|_\infty \cdot \|J(\Phi) - J(\Psi)\|_2 \cdot \left\|(I\otimes A^T)^*\right\|_\infty \\
&\le \|A\|_2^2 \cdot \|J(\Phi) - J(\Psi)\|_2 \\
&= \|J(\Phi) - J(\Psi)\|_2
\end{aligned}$$

The last equality is because $\|A\|_2 = \|vec(A)\|_2 = \|u\|_2 = 1$.

To see the tightness of the first inequality, let $\Phi = \Phi_U$ where $U = X \otimes I^{\otimes n-1}$ and Φ_U is defined in section 2.1. Let Ψ be an identity channel, it's easy to check $\|J(\Phi) - J(\Psi)\|_2/N = \|\Phi - \Psi\|_{\diamond,1\to 2} = \sqrt{2}$. For second inequality, let Φ be

$$\begin{aligned}
&\Phi(|1\rangle\langle 1|) = |2\rangle\langle 2|\,, \Phi(|2\rangle\langle 2|) = |1\rangle\langle 1| \\
&\Phi(|i\rangle\langle i|) = |i\rangle\langle i|\,, \forall i \neq 1, 2 \\
&\Phi(|i\rangle\langle j|) = 0^{N\times N}, \forall i \neq j
\end{aligned}$$

and Ψ be

$$\begin{aligned}
&\Psi(|i\rangle\langle i|) = |i\rangle\langle i|\,, \forall i \\
&\Psi(|i\rangle\langle j|) = 0^{N\times N}, \forall i \neq j
\end{aligned}$$

To verify Φ and Ψ are quantum channels, note that $J(\Phi)$ and $J(\Psi)$ are both PSD and they are trace preserving obviously. Meanwhile, $\|J(\Phi) - J(\Psi)\|_2 = \sqrt{2} = \|(\Phi - \Psi)(|1\rangle\langle 1|)\|_2$.

■

Appendix F. A Simple Influence Estimator

In this section, we will describe a new influence estimator. The estimator only includes single-qubit operations though it fulfills the same function efficiently as the raw influence estimator in Chen et al. (2023), which needs two-qubit operations and maximally entanglement states.

Algorithm 6: INFLUENCE-ESTIMATOR(Φ, S)

Input : Oracle access to quantum channel $\Phi \in C(X)$, $S \subseteq [n]$

Output : $Y \in \{0, 1\}$

1: Uniformly randomly choose $i \in \{0,1\}^S$, $j \in \{0,1\}^{S^c}$. Prepare state $|i\rangle_S |j\rangle_{S^c}$;

2: Query Φ to obtain $\Phi(|i\rangle\langle i|_S \otimes |j\rangle\langle j|_{S^c})$;

3: Measure qubits in S over computational basis, set $Y_1 = 0$ if the result is i, otherwise set $Y_1 = 1$;

4: Uniformly randomly choose $i \in \{0,1\}^S$, $j \in \{0,1\}^{S^c}$. Prepare state $|i\rangle_S |j\rangle_{S^c}$. Apply Hadamard gates to obtain $H^{\otimes S} |i\rangle_S \otimes |j\rangle_{S^c}$;

5: Query Φ to obtain $\Phi(H^{\otimes S} |i\rangle\langle i|_S H^{\otimes S} \otimes |j\rangle\langle j|_{S^c})$, and then apply Hadamard gates to obtain

$$H^{\otimes S} \otimes I_{S^c} \Phi(H^{\otimes S} |i\rangle\langle i|_S H^{\otimes S} \otimes |j\rangle\langle j|_{S^c}) H^{\otimes S} \otimes I_{S^c}$$

6: Measure qubits in S over computational basis. Set $Y_2 = 0$ if the result is i, otherwise set $Y_2 = 1$;

7: Return $Y = Y_1 \vee Y_2$.

Theorem 30 *Given quantum channel Φ and $S \subseteq [n]$, let Y be the output of Algorithm 6. For arbitrary $\delta > 0$, we have:*

1. $\mathrm{Inf}_S[\Phi] = 0 \Rightarrow Y = 0$ *with probability 1;*

2. $\mathrm{Inf}_S[\Phi] \geq \delta \Rightarrow \mathbb{E}[Y] \geq \delta/2$.

Proof In the first case, when $\mathrm{Inf}_S[\Phi] = 0$, we know $\Phi = \widetilde{\Phi}_{S^c} \otimes I_S$, therefore $Y_1 = Y_2 = 0$ with probability 1 and thus $Y = 0$ with probability 1. We focus on the second case.

$$\begin{aligned}
\Pr[Y_1 = 0] &= \frac{1}{2^n} \sum_{i \in \{0,1\}^S} \sum_{j \in \{0,1\}^{n-S}} \Pr[Y_1 = 0 \mid i, j] \\
&= \frac{1}{2^n} \sum_i \sum_j \mathrm{Tr}\left((\langle i|_S \otimes I_{S^c}) \cdot \Phi(|i,j\rangle\langle i,j|) \cdot (|i\rangle_S \otimes I_{S^c}) \right) \\
&= \frac{1}{2^n} \sum_i \sum_j \sum_{x,y \in \mathbb{Z}_4^n} \mathrm{Tr}\, \langle i| \otimes I \cdot \widehat{\Phi}(x,y) \sigma_x \,|i,j\rangle\langle i,j|\, \sigma_y \cdot |i\rangle \otimes I \\
&= \frac{1}{2^n} \sum_i \sum_j \sum_{x,y \in \mathbb{Z}_4^n} \widehat{\Phi}(x,y) \,\langle i|\, \sigma_{x_S} \,|i\rangle \,\langle i|\, \sigma_{y_S} \,|i\rangle \cdot \langle j|\, \sigma_{y_{S^c}} \sigma_{x_{S^c}} \,|j\rangle \\
&= \frac{1}{2^n} \sum_{x,y \in \mathbb{Z}_4^n} \widehat{\Phi}(x,y) \left(\sum_i \langle i|\, \sigma_{x_S} \,|i\rangle \,\langle i|\, \sigma_{y_S} \,|i\rangle \right) \cdot \left(\sum_j \langle j|\, \sigma_{y_{S^c}} \sigma_{x_{S^c}} \,|j\rangle \right)
\end{aligned}$$

For the summation in the first bracket, We note that if $x_S \notin \{0,3\}^S$, say, x_S contains 1 or 2, $\langle i|\, \sigma_{x_S} \,|i\rangle = 0$ for all i. Same for y_S. Thus if $x_S \neq y_S \in \{0,3\}^S$, then $\sum_i \langle i|\, \sigma_{x_S} \,|i\rangle \,\langle i|\, \sigma_{y_S} \,|i\rangle = 0$. For the summation in the second bracket, we have $\sum_j \langle j|\, \sigma_{y_{S^c}} \sigma_{x_{S^c}} \,|j\rangle = \langle \sigma_{y_{S^c}}, \sigma_{x_{S^c}} \rangle$, which is zero if $y_{S^c} \neq x_{S^c}$ and 2^{n-S} if $y_{S^c} = x_{S^c}$. Hence,

$$\begin{aligned}
\Pr[Y_1 = 0] &= \frac{1}{2^n} \sum_{x,y \in \mathbb{Z}_4^n} \widehat{\Phi}(x,y) \left(\sum_i \langle i|\, \sigma_{x_S} \,|i\rangle \,\langle i|\, \sigma_{y_S} \,|i\rangle \right) \cdot \left(\sum_j \langle j|\, \sigma_{y_{S^c}} \sigma_{x_{S^c}} \,|j\rangle \right) \\
&= \sum_{x \in \mathbb{Z}_4^n; x_S \in \{0,3\}^S} \widehat{\Phi}(x,x)
\end{aligned}$$

To bound $\Pr[Y_2 = 0]$, the argument is almost the same except that we actually query $\Phi' = \sum_{x,y} \widehat{\Phi}'(x,y) \Phi_{x,y}$ instead of $\Phi = \sum_{x,y} \widehat{\Phi}(x,y) \Phi_{x,y}$, where $\widehat{\Phi}'(x', y') = \Phi(x, y)$ for all x, y, x', y', where x', y' are obtained from x, y, respectively, by flipping 1 to 3 and 3 to 1. Therefore

$$\Pr[Y_2 = 0] = \sum_{x \in \mathbb{Z}_4^n; x_S \in \{0,1\}^S} \widehat{\Phi}(x,x)$$

Recall that

$$1 - \mathrm{Inf}_S[\Phi] = \sum_{x \in \mathbb{Z}_4^n; x_S = \mathbf{0}} \widehat{\Phi}(x,x)$$

By the union bound, we have that

$$\mathbb{E}[Y_1 + Y_2] = \Pr[Y_1 = 1] + \Pr[Y_2 = 1] \geq \mathrm{Inf}_S[\Phi]$$

and

$$\mathbb{E}[Y] = \mathbb{E}\left[\frac{1}{2}(Y_1 + Y_2) + \frac{1}{2}(Y_1 - Y_2)^2\right] \geq \frac{1}{2}\mathrm{Inf}_S[\Phi]$$

The conclusion follows.

■

Appendix G. Testing Quantum k-Junta Channels

G.1. $\widetilde{O}(k)$ Upper Bound on Testing k-Junta Quantum Channels, Proof of Theorem 18

Theorem 18 (Property of Algorithm 2, Restatement of Theorem 1) *Given quantum channel $\Phi \in C(X)$, with probability at least $9/10$, the algorithm* JUNTA-CHANNEL-TESTER(Φ, k, ε) *outputs "Yes" if Φ is a k-junta, and outputs "No" if Φ is ε-far from any k-junta channel. The algorithm makes $O\big(k \log k/\varepsilon^2\big)$ queries to the channel Φ.*

Proof Let R be the subset of $[n]$, over which Φ acts non-trivially. Recall that S is the output of the call to Algorithm 2, PAULI-SAMPLE in line 1. It is easy to see $S \subseteq R$. We will show with probability at least 0.9, $\mathrm{Inf}_{R-S}[\Phi] = \mathrm{Inf}_{S^c}[\Phi] \leq \varepsilon^2/8$.

For all $i \in [n]$ and $\mathrm{Inf}_i[\Phi] \geq \varepsilon^2/8k$, the probability that i never occurs in $\mathrm{supp}(x)$ is at most $\big(1 - \varepsilon^2/8k\big)^{O(8k\log k/\varepsilon^2)} \leq 0.1/k$. By a union bound, with probability 0.9 for all $i \in [n]$, $\mathrm{Inf}_i[\Phi] \geq \varepsilon^2/8k$ we have $i \in S$ and therefore $\mathrm{Inf}_{S^c}[\Phi] \leq \sum_{i\in S^c} \mathrm{Inf}_i[\Phi] \leq \varepsilon^2/8$ with probability at least 0.9.

If Φ is a k-junta, $|S| \leq |R| \leq k$ and therefore the tester will always outputs "Yes". In other case, if Φ is ε-far from any k-junta channel, according to Corollary 13, any subset $T \subseteq [n]$ with $|T| \leq k$, $\mathrm{Inf}_{T^c}[\Phi] \geq \varepsilon^2/4$ must hold. This fact induces that thet tester will output "No" with probability at least 0.9 since $|S| > k$ if $\mathrm{Inf}_{S^c}[\Phi] \leq \varepsilon^2/8$ occurs.

Besides, JUNTA-CHANNEL-TESTER(Φ, k, ε) makes $O(k \log k/\varepsilon^2)$ queries to Φ. Theorem 18 follows.

■

G.2. $\Omega(\sqrt{k})$ Lower Bound on Testing Quantum k-Junta Channels, Proof of Lemma 20 and Lemma 19

Before proving Lemma 20, we need the following technical lemma.

Lemma 31 *Let n, m be natural numbers and $r_{a,b} \in \mathbb{R}, r_{a,a} \geq 0$ for $a, b \in [n]$. The maximum value*

$$\max_x \sum_{a,b\in[n]} r_{a,b} x_a x_b$$
$$s.t.\ x_a \in [-m, m], \forall a \in [n]$$

can be achieved if we restrict x satisfying that $x_a \in \{m, -m\}$, for all $a \in [n]$.

Proof For arbitrary $a \in [n]$,

$$\frac{\partial^2 \sum_{a,b\in[n]} r_{a,b} x_a x_b}{(\partial x_a)^2} = 2r_{a,a} \geq 0.$$

Thus the objective function is convex in x_a for all $a \in [n]$. The conclusion follows. ■

We will prove our key technical lemma, Lemma 20.

Lemma 20 *For a k-junta channel Φ, there exists a k-junta boolean function g' satisfying that $D(\Phi, \Phi_{g'}) = \min_g D(\Phi, \Phi_g)$, where the minimization is over all boolean functions $g : \{0,1\}^n \to \{0,1\}$.*

Proof For any k-junta channel Φ, let boolean function g minimize $D(\Phi, \Phi_g)$. We will show g could be a k-junta boolean function.

$$\begin{aligned}
g &= \arg\min_g D(\Phi, \Phi_g) \\
&= \arg\min_g \left\| \sum_{a,b\in\{0,1\}^n} (\Phi(|a\rangle\langle b|) - \Phi_g(|a\rangle\langle b|)) \otimes |a\rangle\langle b| \right\|_2^2 \\
&= \arg\min_g \sum_{a,b\in\{0,1\}^n} \|\Phi(|a\rangle\langle b|) - \Phi_g(|a\rangle\langle b|)\|_2^2 \\
&= \arg\min_g \sum_{a,b\in\{0,1\}^n} \left(\langle a| \Phi(|a\rangle\langle b|) |b\rangle - (-1)^{g(a)+g(b)} \right)^2
\end{aligned}$$

Since Φ is a k-junta channel, there exists $T \subseteq [n]$,$|T| = k$, such that $\Phi(\rho) = \widetilde{\Phi}(\rho_T) \otimes \rho_{T^c}$. We have

$$\begin{aligned}
g &= \arg\min_g \sum_{a,b\in\{0,1\}^n} \left(\langle a| \Phi(|a\rangle\langle b|) |b\rangle - (-1)^{g(a)+g(b)} \right)^2 \\
&= \arg\min_g \sum_{\substack{a',b'\in\{0,1\}^T \\ a'',b''\in\{0,1\}^{T^c}}} \left(\langle a', a''| \widetilde{\Phi}(|a'\rangle\langle b'|) \otimes |a''\rangle\langle b''| \, |b', b''\rangle - (-1)^{g(a',a'')+g(b',b'')} \right)^2 \\
&= \arg\min_g \sum_{\substack{a',b'\in\{0,1\}^T \\ a'',b''\in\{0,1\}^{T^c}}} \left(\langle a'| \widetilde{\Phi}(|a'\rangle\langle b'|) |b'\rangle - (-1)^{g(a',a'')+g(b',b'')} \right)^2 \\
&= \arg\max_g \sum_{\substack{a',b'\in\{0,1\}^T \\ a'',b''\in\{0,1\}^{T^c}}} \left(\langle a'| \widetilde{\Phi}(|a'\rangle\langle b'|) |b'\rangle + \overline{\langle a'| \widetilde{\Phi}(|a'\rangle\langle b'|) |b'\rangle} \right) \cdot (-1)^{g(a',a'')} \cdot (-1)^{g(b',b'')} \\
&= \arg\max_g \sum_{a',b'\in\{0,1\}^T} \left(\langle a'| \widetilde{\Phi}(|a'\rangle\langle b'|) |b'\rangle + \overline{\langle a'| \widetilde{\Phi}(|a'\rangle\langle b'|) |b'\rangle} \right) \cdot g'(a') \cdot g'(b')
\end{aligned}$$

where $g'(a') = \sum_{a'' \in \{0,1\}^{T^c}} (-1)^{g(a',a'')}, a' \in \{0,1\}^T$. Let $r_{a',b'} = 2\mathrm{Re}\Big(\langle a'| \widetilde{\Phi}(|a'\rangle\langle b'|) |b'\rangle\Big)$, we know $r_{a',a'} \geq 0$ since $\widetilde{\Phi}(|a'\rangle\langle b'|)$ is PSD.

Combining with Lemma 31, we know that there exists g' which achieves the maximum satisfying that $g'(a') = 2^{n-k}$ or $g'(a') = -2^{n-k}$, for all $a' \in \{0,1\}^T$. Thus, we can take k-junta boolean function g to obtain the minimum of $D(\Phi, \Phi_g)$. ■

Before we prove Lemma 19, we need the following lemma.

Lemma 32 *Given boolean function f, if f is ε-far from any k-junta boolean function, then for any k-junta boolean function g, we have $D(\Phi_f, \Phi_g) \geq \sqrt{2\varepsilon}$.*

Proof For k-junta boolean function g, if $1/2 \geq D(f,g) \geq \varepsilon$, we claim that $D(\Phi_f, \Phi_g) \geq \sqrt{2D(f,g)} \geq \sqrt{2\varepsilon}$. Recall that $D(f,g) = \Pr_x[f(x) \neq g(x)]$.

$$\begin{aligned}
D(\Phi_f, \Phi_g) &= \frac{1}{\sqrt{2}N} \|\mathrm{vec}(U_f)\mathrm{vec}(U_f)^* - \mathrm{vec}(U_g)\mathrm{vec}(U_g)^*\|_2 \\
&= \frac{1}{\sqrt{2}N} \sqrt{\sum_{a,b \in \{0,1\}^n} \left[(-1)^{f(a)+f(b)} - (-1)^{g(a)+g(b)}\right]^2} \\
&= \frac{\sqrt{2}}{N} \sqrt{\sum_{a,b \in \{0,1\}^n} (\mathbb{1}[f(a) \neq g(a)] \cdot \mathbb{1}[f(b) = g(b)] + \mathbb{1}[f(a) = g(a)] \cdot \mathbb{1}[f(b) \neq g(b)])} \\
&= \frac{\sqrt{2}}{N} \sqrt{2(1 - D(f,g))D(f,g) \cdot N^2} \\
&\geq \sqrt{2\varepsilon}
\end{aligned}$$

■

Lemma 19 *A $(k, \sqrt{\varepsilon/2})$-channel junta tester is a (k, ε)-classical junta tester.*

Proof We will analyze the output of a $(k, \sqrt{\varepsilon/2})$-channel junta tester given oracle to Φ_f for some boolean function $f : \{0,1\}^n \to \{0,1\}$.

If f is a k-junta, it is easy to see Φ_f is also a k-junta. Thus the channel junta tester outputs "Yes" with probability at least $9/10$.

If f is ε-far from any k-junta boolean function, we are going to show Φ_f is $\sqrt{\varepsilon/2}$-far from any k-junta channel. We give an illustration of our proof as Figure 1. Our goal is show for any k-junta channel Φ', $D(\Phi_f, \Phi') = d_3$ is large. We firstly show, from Lemma 32, that for any k-junta boolean function g, $D(\Phi_f, \Phi_g) = d_1$ is also large (as Step 1 in figure 1). Next we show for any k-junta channel Φ', there exists k-junta boolean function g such that $D(\Phi', \Phi_g) \leq D(\Phi_f, \Phi')$, i.e., $d_2 \leq d_3$. Finally, we conclude that $d_3 \geq (d_2 + d_3)/2 \geq d_1$.

For any k-junta channel Φ', let $g = \arg\min_{g \text{ is a boolean function}} D(\Phi', \Phi_g)$. Then g is a k-junta by Lemma 20. We have

$$D(\Phi', \Phi_g) \leq D(\Phi', \Phi_f)$$

and since g is a k-junta, according to Lemma 32,

$$D(\Phi_f, \Phi_g) \geq \sqrt{2\varepsilon}$$

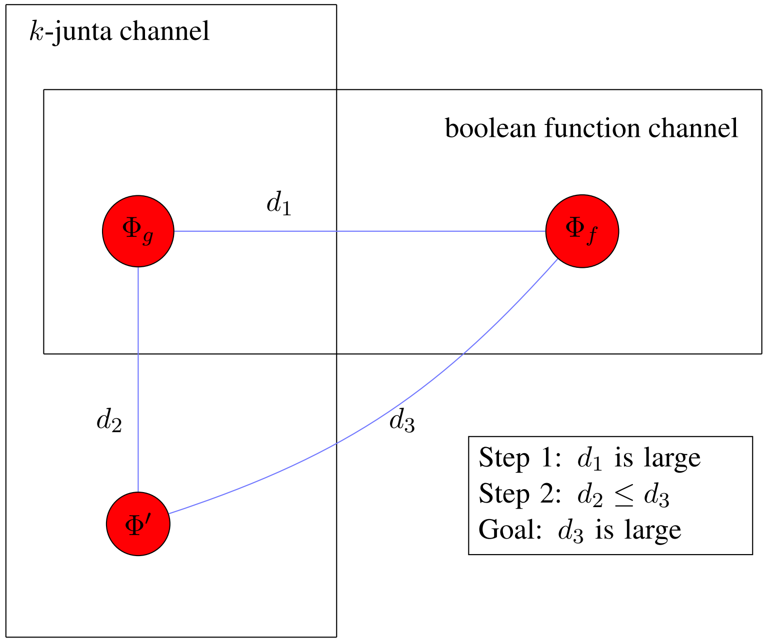

Figure 1: Illustration of proof of Lemma 19

To conclude, we have

$$D(\Phi', \Phi_f) \geq \frac{1}{2}\big(D(\Phi', \Phi_f) + D(\Phi', \Phi_g)\big) \geq \frac{1}{2} D(\Phi_f, \Phi_g) \geq \sqrt{\frac{\varepsilon}{2}}$$

for any k-junta channel Φ'.

■

G.3. Reduction from k-Junta Unitary Testing, Proof of Lemma 22 and Lemma 21

The proof of Lemma 22 follows the same line as Lemma 20. We first show a lemma similar to Lemma 31, which will be used later.

Lemma 33 *Let n be a natural number. For $a, b \in [n]$, let $A_{a,b} \in \mathbb{C}^{n\times n}$ be an $n \times n$ matrix. Set $A = \sum_{a,b\in[n]} A_{a,b} \otimes |a\rangle\langle b| \in \mathbb{C}^{n^2\times n^2}$ to be the Choi representation of a quantum channel Φ. In other words, A is PSD and there exists $(B_s)_s$, $B_s \in \mathbb{C}^{n\times n}$, s.t., $A = \sum_s vec(B_s)vec(B_s)^*$ and $\sum_s B_s^* B_s = I$. The maximum value*

$$\max_{V\in\mathbb{C}^{n\times n}} \sum_{a,b\in[n]} \langle a| V^* A_{a,b} V |b\rangle$$
$$s.t.\ V^*V \leq I$$

can be achieved if we restrict V satisfying that $V^*V = I$.

Proof Note that

$$\begin{aligned}\sum_{a,b\in[n]} \langle a| V^* A_{a,b} V |b\rangle &= \mathrm{vec}(V)^* A \mathrm{vec}(V) \\ &= \sum_s \mathrm{vec}(V)^* \mathrm{vec}(B_s) \mathrm{vec}(B_s)^* \mathrm{vec}(V) \\ &= \sum_s |\langle V, B_s\rangle|^2,\end{aligned}$$

where V takes over all matrices in

$$\begin{aligned}&\{V \mid V \in \mathbb{C}^{n\times n}, V^*V \leq I\} \\ &\qquad = \{W\Sigma W' \mid W, W' \text{ are unitaries}, \Sigma \text{ is a diagonal real matrix}, -I \leq \Sigma \leq I\}\end{aligned}$$

by the SVD decomposition. Suppose $\Sigma = \mathsf{Diag}(x_1, \ldots, x_n)$. It is not hard to see

$$\sum_s |\langle V, B_s\rangle|^2 = \sum_s \left|\left\langle W\Sigma W' \mid W, B_s\right\rangle\right|^2$$

is a quadratic form in $x_1, \ldots, x_n$. The coefficients of x_a is $B'^2_{s,aa} \geq 0$, where $B'_s = W^* B_s W'^*$. By Lemma 22, the maximum can be achieved if $x_a = \pm 1$. We conclude the result. ■

Lemma 22 *For every k-junta channel Φ', there exists a k-junta unitary V, such that $D(\Phi', \Phi_V) = \min_V D(\Phi', \Phi_V)$, where the minimization is over all unitaries V.*

Proof For any k-junta channel Φ', let unitary V minimize $D(\Phi', \Phi_V)$. We will show V could be a k-junta unitary.

$$\begin{aligned}V &= \arg\min_V D(\Phi', \Phi_V) \\ &= \arg\min_V \left\| \sum_{a,b\in\{0,1\}^n} \big(\Phi'(|a\rangle\langle b|) - \Phi_V(|a\rangle\langle b|)\big) \otimes |a\rangle\langle b| \right\|_2^2 \\ &= \arg\min_V \sum_{a,b\in\{0,1\}^n} \left\|\Phi'(|a\rangle\langle b|) - \Phi_V(|a\rangle\langle b|)\right\|_2^2 \\ &= \arg\min_V \sum_{a,b\in\{0,1\}^n} \left\|\Phi'(|a\rangle\langle b|) - V|a\rangle\langle b| V^*\right\|_2^2\end{aligned}$$

Since Φ' is a k-junta, there exists $T \subseteq [n]$,$|T| = k$, s.t., $\Phi' = \widetilde{\Phi}'_T \otimes I_{T^c}$. Let $V = \sum_{x\in\mathbb{Z}_4^{T^c}} V_x \otimes \sigma_x$. Besides, because $\sum_{a,b} \|V |a\rangle\langle b| V^*\|_2^2 = 4^n$, we have:

$$\begin{aligned}
V &= \arg\min_V \sum_{a,b\in\{0,1\}^n} \left\|\Phi'(|a\rangle\langle b|) - V |a\rangle\langle b| V^*\right\|_2^2 \\
&= \arg\max_V \sum_{a,b\in\{0,1\}^n} \text{Re}\left\{\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle\right\} \\
&= \arg\max_V \sum_{a,b\in\{0,1\}^n} \langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle \\
&= \arg\max_V \sum_{a',b'\in\{0,1\}^T} \sum_{a'',b''\in\{0,1\}^{T^c}} \sum_{x,y\in\mathbb{Z}_4^{T^c}} \left\langle b'\right| V_x^*\widetilde{\Phi}'(\left|b'\right\rangle\!\left\langle a'\right|)V_y \left|a'\right\rangle \cdot \left\langle b''\right| \sigma_x \left|b''\right\rangle \cdot \left\langle a''\right| \sigma_y \left|a''\right\rangle \\
&= \arg\max_V \sum_{a',b'\in\{0,1\}^T} \sum_{x,y\in\mathbb{Z}_4^{T^c}} \left\langle b'\right| V_x^*\widetilde{\Phi}'(\left|b'\right\rangle\!\left\langle a'\right|)V_y \left|a'\right\rangle \cdot \text{Tr}\,\sigma_x \cdot \text{Tr}\,\sigma_y \\
&= \arg\max_V \sum_{a',b'\in\{0,1\}^T} \left\langle b'\right| V_{0^{T^c}}^*\widetilde{\Phi}'(\left|b'\right\rangle\!\left\langle a'\right|)V_{0^{T^c}} \left|a'\right\rangle
\end{aligned}$$

where the second equality follows from

$$\begin{aligned}
\left\|\Phi'(|a\rangle\langle b|) - V |a\rangle\langle b| V^*\right\|_2^2 &= \left\|\Phi'(|a\rangle\langle b|)\right\|_2^2 + \|V |a\rangle\langle b| V^*\|_2^2 - 2\cdot\text{Re}\left\{\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle\right\} \\
&= \left\|\Phi'(|a\rangle\langle b|)\right\|_2^2 + \||a\rangle\langle b|\|_2^2 - 2\cdot\text{Re}\left\{\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle\right\}
\end{aligned}$$

and the first two terms have nothing to do with V. The third equality is because, if $a = b$, $\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle \in \mathbb{R}$ and $\text{Re}\left\{\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle\right\} = \langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle$. If $a \neq b$,

$$\begin{aligned}
\sum_{a,b\in\{0,1\}^n, a\neq b} \langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle &= \sum_{a,b\in\{0,1\}^n, a<b} \langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle + \langle a| V^*\Phi'(|a\rangle\langle b|)V |b\rangle \\
&= \sum_{a,b\in\{0,1\}^n, a<b} \langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle + \overline{\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle} \\
&= \sum_{a,b\in\{0,1\}^n, a<b} 2\cdot\text{Re}\big\{\langle b| V^*\Phi'(|b\rangle\langle a|)V |a\rangle\big\} \in \mathbb{R}
\end{aligned}$$

Notice that $\text{Tr}_{T^c} V^*V = \sum_{x,y\in\mathbb{Z}_4^{T^c}} V_x^*V_y\langle\sigma_x, \sigma_y\rangle = 2^{n-k}\sum_{x\in\mathbb{Z}_4^{T^c}} V_x^*V_x = 2^{n-k}I_T$. By Lemma 33, the maximum can be achieved when $V_{0^{T^c}}^*V_{0^{T^c}} = I_T$, which implies $V_x = 0$ for $x \neq 0^{T^c}$. Thus, we can take k-junta unitary V to obtain the minimum of $D(\Phi', \Phi_V)$. ■

The proof of Lemma 21 is similar to Lemma 19.

Lemma 21 (Reduction from Testing k-Junta Unitaries to Testing k-Junta Channels) *A (k, ε)-channel junta tester is naturally a $(k, \varepsilon/2)$-unitary junta tester.*

Proof Let U be a unitary matrix. It suffices to show that if U is ε-far from any k-junta unitary, then Φ_U is $\varepsilon/2$-far from any k-junta channel. We firstly show that for any k-junta unitary V,

$D(\Phi_U, \Phi_V)$ is large. Next, we show for any k-junta channel Φ', there exists k-junta unitary V such that $D(\Phi', \Phi_V) \leq D(\Phi', \Phi_U)$.

For any k-junta channel Φ', let $V = \arg\min_{\text{unitary } V} D(\Phi', \Phi_V)$. By Lemma 22, V is k-junta. For any unitary U, if U is ε-far from any k-junta unitary, then therefore $D(\Phi_U, \Phi_V) \geq \varepsilon$. Thus

$$D(\Phi_U, \Phi') \geq \frac{1}{2}(D(\Phi_U, \Phi') + D(\Phi', \Phi_V)) \geq \frac{1}{2} D(\Phi_U, \Phi_V) \geq \varepsilon/2.$$

The last inequality follows directly from Lemma 14. We complete the proof. ■

Appendix H. $O(4^k/\varepsilon^2)$ Upper Bound on Learning Quantum k-Junta Channels, Proof of Theorem 23

Before describing the learning algorithm JUNTA-CHANNEL-LEARNER, we introduce a tomography algorithm from O'Donnell and Wright (2017).

Fact 34 (Corollary 1.4 of O'Donnell and Wright (2017)) *There exists an algorithm* TOMOGRAPHY, *which is given $O(4^k/\varepsilon^2)$ copies of an unknown $2k$ qubit state ρ and outputs the description an estimated state $\widetilde{\rho}$ satisfying that $\|\rho - \widetilde{\rho}\|_2 \leq \varepsilon$, with probability at least 0.99.*

Now we are ready to prove Theorem 23.

Theorem 23 (Property of Algorithm 3, Restatement of Theorem 1) *Given oracle access to k-junta channel Φ, with probability at least $9/10$,* JUNTA-CHANNEL-LEARNER(Φ, k, ε) *outputs the description of quantum channel Ψ such that $D(\Phi, \Psi) \leq \varepsilon$. Furthermore, this algorithm makes $O(4^k/\varepsilon^2)$ queries.*

Proof Let R be the subset of $[n]$, over which Φ acts non-trivially. Recall that S is the output of the call to Algorithm 1, PAULI-SAMPLE in line 1 of JUNTA-CHANNEL-LEARNER(Φ, k, ε). It is easy to see $S \subseteq R$. With the similar analysis as the proof of Theorem 18, with probability at least 0.99, $\mathrm{Inf}_{R-S}[\Phi] \leq \varepsilon^2/8$ holds.

Let $\Phi = \Phi_R \otimes I^{R^c}$ and $v(\Phi) = v(\Phi_R) \otimes v(I^{R^c})$. Consider the quantum state ψ returned by QUANTUM-STATE-PREPARATION. We conclude that the probability that it does not output "error" is

$$\begin{aligned}
\operatorname{Tr} v(\Phi) \cdot \left(I^S \otimes \left|v(I^{S^c})\right\rangle\!\left\langle v(I^{S^c})\right|\right) &= \sum_{x,y \in \mathbb{Z}_4^n} \widehat{\Phi}(x,y) \operatorname{Tr} v(\Phi_{x_S, y_S}) \cdot \langle v(\sigma_{\mathbf{0}})|\, v(\Phi_{x_{S^c}, y_{S^c}})\, |v(\sigma_{\mathbf{0}})\rangle \\
&= \sum_{x \in \mathbb{Z}_4^n, x_{S^c} = \mathbf{0}} \widehat{\Phi}(x,x) \\
&= \mathrm{Inf}_S[\Phi] \geq 1 - \varepsilon^2/8.
\end{aligned}$$

The step 2 of JUNTA-CHANNEL-LEARNER collects t copies of ψ in $10t$ calls to the preparation subroutine with probability at least 0.99 for large enough k, since the expectation of successful collections is at least $(1 - \varepsilon^2/8) \cdot 10t \geq 8t$. We will show $\psi \otimes v(I^{R-S})$ is close to $v(\Phi_R)$.

It is easy to calculate that

$$\begin{aligned}
\psi &= \frac{1}{\mathrm{Inf}_S\Phi} I^S \otimes \langle v(\sigma_{\mathbf{0}})| \cdot v(\Phi) \cdot I^S \otimes |v(\sigma_{\mathbf{0}})\rangle \\
&= \frac{1}{\mathrm{Inf}_S\Phi} \sum_{x,y\in\mathbb{Z}_4^n} \widehat{\Phi}(x,y) v(\Phi_{x_S,y_S}) \cdot \langle v(\sigma_{\mathbf{0}})| \cdot v(\Phi_{x_{S^c},y_{S^c}}) \cdot |v(\sigma_{\mathbf{0}})\rangle \\
&= \frac{1}{\mathrm{Inf}_S\Phi} \sum_{x,y\in\mathbb{Z}_4^n, x_{S^c}=y_{S^c}=\mathbf{0}} \widehat{\Phi}(x,y) v(\Phi_{x_S,y_S})
\end{aligned}$$

Let $\psi' = \mathrm{Inf}_S[\Phi]\cdot\psi = \sum_{x,y\in\mathbb{Z}_4^n, x_{S^c}=y_{S^c}=\mathbf{0}} \widehat{\Phi}(x,y) v(\Phi_{x_S,y_S})$. We have

$$\begin{aligned}
\left\|\psi \otimes v(I^{S^c}) - v(\Phi)\right\|_2 &\le \left\|\psi \otimes v(I^{S^c}) - \psi' \otimes v(I^{S^c})\right\|_2 + \left\|\psi' \otimes v(I^{S^c}) - v(\Phi)\right\|_2 \\
&= (1 - \mathrm{Inf}_S[\Phi])\left\|\psi \otimes v(I^{S^c})\right\|_2 + \left\|\psi' \otimes v(I^{S^c}) - v(\Phi)\right\|_2 \\
&\le \frac{\varepsilon^2}{8} + \left\|\psi' \otimes v(I^{S^c}) - v(\Phi)\right\|_2
\end{aligned}$$

and

$$\begin{aligned}
\left\|\psi' \otimes v(I^{S^c}) - v(\Phi)\right\|_2^2 &= \left\| \sum_{x,y\in\mathbb{Z}_4^n, x_{S^c}\neq\mathbf{0} \text{ or } y_{S^c}\neq\mathbf{0}} \widehat{\Phi}(x,y) v(\Phi_{x,y}) \right\|_2^2 \\
&= \sum_{x,y\in\mathbb{Z}_4^n, x_{S^c}\neq\mathbf{0} \text{ or } y_{S^c}\neq\mathbf{0}} \left|\widehat{\Phi}(x,y)\right|^2 \\
&\le \sum_{x,y\in\mathbb{Z}_4^n, x_{S^c}\neq\mathbf{0} \text{ or } y_{S^c}\neq\mathbf{0}} \widehat{\Phi}(x,x)\widehat{\Phi}(y,y) \\
&\le 2 \sum_{x,y\in\mathbb{Z}_4^n, x_{S^c}\neq\mathbf{0}} \widehat{\Phi}(x,x)\widehat{\Phi}(y,y) \\
&= 2\mathrm{Inf}_{S^c}[\Phi] \le \frac{\varepsilon^2}{4}
\end{aligned}$$

Therefore

$$\frac{1}{\sqrt{2}}\left\|\psi \otimes v(I^{S^c}) - v(\Phi)\right\|_2 \le \frac{1}{\sqrt{2}} \cdot \left(\frac{\varepsilon^2}{8} + \frac{\varepsilon}{2}\right) \le 0.45\varepsilon$$

By the step 1 of TOMOGRAPHY, we get a description of quantum state ϕ with probability 0.99 s.t. $\|\phi - \psi\|_2 \le 0.04\varepsilon$ and $\left\|\phi \otimes v(I^{S^c}) - v(\Phi)\right\|_2/\sqrt{2} \le 0.49\varepsilon$. After we find the closest Choi state ϕ' to ϕ in the step 2 of TOMOGRAPHY, we are sure that $\left\|\phi' \otimes v(I^{S^c}) - v(\Phi)\right\|_2 \le \varepsilon$ and the returned channel is close to Φ with distance at most ε with probability at least $9/10$.

To see the query complexity, the call to PAULI-SAMPLE costs only $O\big(k\log k/\varepsilon^2\big)$ queries to Φ and the preparation and tomography need $O\big(4^k/\varepsilon^2\big)$ queries. The total queries are $O\big(4^k/\varepsilon^2\big)$.

■