Testing and Learning Quantum Juntas Nearly Optimally*

Thomas Chen† Shivam Nadimpalli‡ Henry Yuen§

Abstract

We consider the problem of testing and learning *quantum k-juntas*: n-qubit unitary matrices which act non-trivially on just k of the n qubits and as the identity on the rest. As our main algorithmic results, we give

1. A $\widetilde{O}(\sqrt{k})$-query quantum algorithm that can distinguish quantum k-juntas from unitary matrices that are "far" from every quantum k-junta; and
2. A $O(4^k)$-query algorithm to *learn* quantum k-juntas.

We complement our upper bounds for testing and learning quantum k-juntas with near-matching lower bounds of $\Omega(\sqrt{k})$ and $\Omega(4^k/k)$, respectively. Our techniques are Fourier-analytic and make use of a notion of *influence* of qubits on unitaries.

1 Introduction

Certifying and characterizing the dynamical behavior of quantum systems is a fundamental task in physics which is often achieved via *quantum process tomography* (QPT) [CN97]. However, QPT is extremely resource-intensive. For example, all known methods for learning a classical description of an arbitrary n-qubit unitary operator, given black-box query access to it, require $\Omega(4^n)$ queries to the unitary [GJ14]. On the other hand, this complexity can be significantly reduced if, instead of learning the entire description of the unknown unitary, we want to *test* whether the unitary satisfies a specific property. This naturally leads us to consider the well-studied *property testing* framework in theoretical computer science [Gol10, BY22].

The setup of property testing (in the context of unitary dynamics, as it pertains to this paper) is as follows: Given oracle access[1] to a unitary operator U and its inverse $U^\dagger$, our goal is to determine whether U has a certain property or is "far"[2] from every unitary operator satisfying that property using a small number of calls to the oracles to U and $U^\dagger$. We also allow for the algorithm to output an incorrect answer with some small probability. Several natural properties of unitary dynamics have been studied in this model, such as commutativity, diagonality, membership in the Pauli basis, etc. We refer the interested reader to Section 5.1 of the survey by Montanaro and de Wolf on quantum property testing [MdW16] for more information.

The property we are interested in testing here is that of being a *k-junta*: We say that an n-qubit unitary U is a *k-junta* if it acts "non-trivially" on only k of the n-qubits (see Definition 2.2 for a formal definition). Like Montanaro and Osborne [MO10], we will refer to a unitary k-junta as a *quantum k-junta*, to distinguish it from a k-junta Boolean function (or simply, Boolean k-junta). As a special case, the notion of quantum k-juntas captures the well-studied problem of testing if a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is a k-junta (cf. Question 1.3).

PROBLEM 1.1. (TESTING QUANTUM k-JUNTAS) *Given oracle access to a unitary U and its inverse $U^\dagger$ acting on n qubits and $\varepsilon > 0$, decide with probability at least $9/10$ if U is a k-junta or if $\mathrm{dist}(U,V) \geq \varepsilon$ for all quantum k-juntas V acting on n qubits.*

Our first main result is an algorithm for testing if a unitary U is k-junta using $\widetilde{O}(\sqrt{k})$ queries to U and $U^\dagger$, where $\widetilde{O}(\cdot)$ hides polylogarithmic factors of k. Crucially, the query complexity of the tester is independent of n, the total number of qubits in U. We complement this with a near-matching lower bound of $\Omega(\sqrt{k})$ for the junta testing problem, implying that our algorithm is optimal up to a polylogarithmic factor in k.

*The full version of the paper can be accessed at `https://arxiv.org/abs/2207.05898`

†Columbia University.

‡Columbia University. Supported by NSF grants IIS-1838154, CCF-2106429, CCF-2211238, CCF-1763970, and CCF-2107187.

§Columbia University. Supported by AFOSR award FA9550-21-1-0040, NSF CAREER award CCF-2144219, and the Sloan Foundation.

[1]More formally, an oracle for a unitary U takes in as input a quantum state $|\psi\rangle$ and outputs $U|\psi\rangle$.

[2]See Definition 2.3 for a formal definition of "dist," the distance metric.

	Classical Testing	Quantum Testing	Classical Learning	Quantum Learning
$f : \{0,1\}^n \to \{0,1\}$	$O(k \log k)$ [Bla09]	$\widetilde{O}(\sqrt{k})$ [ABRdW16]	$n^{0.6k}\text{poly}(n)$ [Val15]	$O(2^k)$ [AS07]
	$\Omega(k \log k)$ [Sağ18]	$\Omega(\sqrt{k})$ [BKT17]	$\Omega(2^k + \log n)$ (Folklore)	$\Omega(2^k)$ [AS07]
Unitary $U \in \mathcal{M}_{2^n \times 2^n}$	—	$\widetilde{O}(\sqrt{k})$ (Theorem 3.2)	—	$O(4^k)$ (Theorem 5.1)
		$\Omega(\sqrt{k})$ (Theorem 4.1)		$\Omega(\frac{4^k}{k})$ (Theorem 5.2)

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

THEOREM 1.1. (INFORMAL VERSION OF THEOREMS 3.2 AND 4.1) *Quantum k-juntas can be tested with $\widetilde{O}(\sqrt{k})$ queries. Furthermore, testing quantum k-juntas requires $\Omega(\sqrt{k})$ queries.*

As a remark, our upper bound uses amplitude amplification on a subroutine that queries U. Because amplitude amplification will apply our subroutine and its inverse, we need query access to $U^\dagger$ as well in the formulation of the quantum junta testing problem.

Another natural problem we consider is that of *learning* quantum k-juntas. In particular, the learning problem asks to output an approximation to a quantum k-junta given oracle access to the latter. Unlike the testing problem, the learning problem does not require access to $U^\dagger$.

PROBLEM 1.2. (LEARNING QUANTUM k-JUNTAS) *Given oracle access to a quantum k-junta U acting on n qubits and an error parameter ε, output a unitary $\widehat{U}$ such that* $\text{dist}(U, \widehat{U}) \leq \varepsilon$.

Our second main result is an algorithm to *learn* quantum k-juntas with significantly lower sample complexity than the naive QPT approach; in particular there once again is *no* dependence on the total number of qubits, n.

THEOREM 1.2. (INFORMAL VERSION OF THEOREMS 5.1 AND 5.2) *Given oracle access to a quantum k-junta U acting on n-qubits and $\varepsilon > 0$, there exists an algorithm that makes $O(4^k/\varepsilon^2)$ queries to U and outputs with probability $9/10$ a unitary $\widehat{U}$ such that* $\text{dist}(U, \widehat{U}) \leq \varepsilon$. *Furthermore, $\Omega(4^k/k)$ queries are necessary to learn quantum juntas.*

Our upper bounds for testing and for learning are proved via Fourier-analytic techniques and crucially make use of the notion of *influence of qubits on a unitary*, first introduced by Montanaro and Osborne [MO10] in the context of Hermitian unitary matrices. Our lower bound for testing quantum k-juntas appeals to the lower bound for testing Boolean k-juntas obtained by Bun, Kothari, and Thaler [BKT17], as well as a new structural result for quantum k-juntas. Our lower bound for learning quantum k-juntas arises from the communication complexity of the INPUT GUESSING game [Nay99].

Organization. We briefly recall related work on testing both Boolean and quantum juntas in Section 1.1, and then give a high-level technical overview of our results in Section 1.2. We prove our $\widetilde{O}(\sqrt{k})$ upper bound for testing quantum k-juntas in Section 3, and prove our $\Omega(\sqrt{k})$ lower bound for the same in Section 4. Finally, we present our upper and lower bound on learning quantum k-juntas in Section 5.

1.1 Related Work

We summarize related work as well as our contributions in Table 1.

Classical Testing of Boolean Juntas. We first note that Question 1.1 captures as a special case its Boolean analog, which we state below as Question 1.3. Recall that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is a k-junta if $f(x) = g(x_{i_1}, \ldots, x_{i_k})$ for some $g : \{0,1\}^k \to \{0,1\}$. We also say that for $f, g : \{0,1\}^n \to \{0,1\}$,

$$\text{dist}(f, g) := \mathbf{Pr}[f(\boldsymbol{x}) \neq g(\boldsymbol{x})]$$

for $\boldsymbol{x} \sim \{0,1\}^n$ drawn uniformly at random. (In other words, the distance metric we use for Boolean functions is simply the normalized Hamming distance.)

PROBLEM 1.3. (TESTING BOOLEAN k-JUNTAS) *Given classical or quantum query access to a function* $f : \{0,1\}^n \to \{0,1\}$ *via a unitary* $\mathcal{O}_f$*, decide with constant probability if* f *is a* k*-junta or if* $\mathrm{dist}(f,g) \geq \varepsilon$ *for every* k*-junta* $g : \{0,1\}^n \to \{0,1\}$.

This question has been extensively studied over recent decades, with the first result explicitly related to testing juntas obtained by Parnas, Ron, and Samorodnitsky [PRS02] who gave a *classical* algorithm for testing 1-juntas with $O(1)$ queries. Soon afterwards, Fischer et al. [FKR$^+$04] introduced classical algorithms for testing k-juntas with $\widetilde{O}(k^2)$ queries. The query complexity of classically testing juntas was later improved by Blais [Bla09] who gave a nearly optimal tester which makes $\widetilde{O}(k)$ queries. Blais's tester is asymptotically optimal up to a logarithmic factor, given the $\Omega(k)$ lower bound for classically testing k-juntas by [CG04].

Quantum Testing and Learning of Boolean Juntas. There has also been a long line of work on testing Boolean juntas via *quantum* algorithms, i.e. algorithms with query access to a unitary $\mathcal{O}_f$ representing a function $f : \{0,1\}^n \to \{0,1\}$, allowing the algorithm to query superpositions of inputs. Atıcı and Servedio [AS07] gave an elegant quantum algorithm to test k-juntas using $O(k)$ queries.[3] More recently, Ambainis et al. [ABRdW16] came up with a quantum algorithm to test juntas that makes only $\widetilde{O}(\sqrt{k})$ queries. This was shown to be essentially optimal by Bun, Kothari, and Thaler [BKT17] who proved an $\widetilde{\Omega}(\sqrt{k})$ lower bound for via a reduction from the image size testing problem. Finally, Atıcı and Servedio [AS07] also gave a $O(2^k)$-sample quantum algorithm for *learning* Boolean k-juntas in the PAC model.

Quantum Testing of Quantum Juntas. Returning to Question 1.1, Wang [Wan11] gave a tester for testing whether a unitary operator U is a k-junta or is ε-far from a k-junta that makes $O(k)$ queries, and their algorithm turns out to be a direct generalization of the tester of Atıcı and Servedio [AS07].[4] Finally, Montanaro and Osborne [MO10] had previously studied a different tester for the property of being a "dictatorship," i.e. a 1-junta, but did not prove correctness.

1.2 Our Techniques

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

1.2.1 Testing Quantum Juntas

Our $\widetilde{O}(\sqrt{k})$-query tester for quantum k-juntas can be viewed as direct analog of the $\widetilde{O}(\sqrt{k})$-query tester for Boolean k-juntas obtained by Ambainis, et al. [ABRdW16]. Our tester relies crucially on the notion of *influence of qubits on a unitary*, which was first introduced by Montanaro and Osborne [MO10] for Hermitian unitaries. Informally, the influence of a qubit on a unitary U captures how non-trivially U acts on that qubit; see Section 2.3 for a formal definition as well as useful properties of this notion of influence. Our main technical contributions here are a formulation of the influence of a qubit on an arbitrary unitary and a subroutine INFLUENCE-ESTIMATOR (cf. Section 3.1) to estimate this influence using the *Choi-Jamiołkowski (CJ) isomorphism* between unitary operators on n qubits and pure states in $\mathbb{C}^{2^n} \times \mathbb{C}^{2^n}$. With this in hand, we closely mirror the approach of Ambainis et al. [ABRdW16] in Section 3.2. We essentially used their algorithm as a black-box, but our analysis differs in certain parameters; for completeness, we present the entire analysis with these modifications.

The $\Omega(\sqrt{k})$-query lower bound for testing quantum k-juntas relies on the $\Omega(\sqrt{k})$-query lower bound for testing Boolean k-juntas obtained by Bun, Kothari, and Thaler [BKT17]. We do so via the natural encoding of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ as a unitary U_f given by

$$U_f := \mathrm{diag}((-1)^{f(x)}). \tag{1.1}$$

It is immediate from Equation (1.1) that encoding a Boolean k-junta in this way yields a quantum k-junta. Our main structural result Proposition 4.1, shows that if a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is far from any Boolean k-junta, then U_f is also far from any quantum k-junta. We start by first showing that U_f is far from U_g for every Boolean k-junta $g : \{0,1\}^n \to \{0,1\}$, and then handling arbitrary quantum k-juntas via Lemma 4.1.

[3] At the time [AS07] was written, the best classical upper bound on testing juntas was $\widetilde{O}(k^2)$.

[4] The result originally obtained by Wang had a worse bound of $O(k \log k)$, but this can be improved to $O(k)$ by following the analysis of [AS07] (cf. Section 5.1.6 of [MdW16]). We also note that their query complexity's dependence on ε can be improved via a straightforward application of amplitude amplifiaction.

1.2.2 Learning Quantum Juntas Our learning algorithm, Figure 6, can be viewed as analogous to the algorithm obtained by Atıcı and Servedio [AS07] for Boolean k-juntas. We again make use of the CJ isomorphism between unitary operators on n qubits and pure states in $\mathbb{C}^{2^n} \times \mathbb{C}^{2^n}$, allowing us to techniques used to learn *quantum states* to learn the unitary U. We start by first determining the high-influence qubits of the quantum k-junta U via "Pauli sampling," which is analogous to the Fourier Sampling subroutine used by [AS07]. We then take the CJ isomorphism of U and appropriately trace out the qubits with negligible influence, producing a reduced CJ state on the high influence qubits. Finally, we use a *pure state tomography* procedure to learn the reduced CJ state using $O(4^k/\varepsilon^2)$ samples.

Our lower bound for learning quantum juntas is proved via a reduction to a communication complexity lower bound, namely a quantum lower bound proved by Nayak [Nay99] on the communication required for one party to guess the input of another party. The key idea is that an $\varepsilon-$covering of k-junta unitaries has size at least $\Omega((1/\varepsilon)^{4^k})$. Thus, identifying a certain party's uniformly-selected k-junta unitary in this cover requires at least $\Omega(\log(1/\varepsilon)^{4^k})$ communication rounds in a protocol where that party behaves as a membership-oracle for their unitary.

1.3 Future Work A natural next direction is to consider the testability/learnability of *quantum channels* acting non-trivially on k-qubits. Recall that a quantum channel is a completely positive, trace-preserving linear map; see [Wat18] for a comprehensive introduction to the subject. As noted in [MdW16], there has not been much work on testing properties of quantum channels.

We also remark that (to our knowledge) there has been no work on *tolerant* property testing—for both Boolean functions as well as unitary matrices—via quantum algorithms.[5] The best known classical upper bound for tolerant testing of Boolean k-juntas is $2^{\widetilde{O}(\sqrt{k})}$ due to Iyer, Tal, and Whitmeyer [ITW21]. We also note that a $\Omega(2^{\sqrt{k}})$ lower-bound against classical non-adaptive algorithms for tolerant junta testing was obtained by Pallavoor et al. [PRW19].

Finally, it is unknown whether quantum algorithms offer any advantage in terms of query complexity for the problem of testing Boolean k-juntas in the *distribution-free* setting.[6] In particular, Belovs [Bel19] gave a $O(k)$ quantum tester for Boolean k-juntas in the distribution-free model, matching the query complexity of the best classical algorithms for testing Boolean k-juntas in the distribution-free model due to Bshouty [Bsh19] and Zhang [Zha19].

2 Preliminaries

In this section, we introduce notation and recall useful background. We assume familiarity with elementary quantum computing and quantum information theory, and refer the interested reader to [NC10, Wil17] for more background. For $n \geq 1$, we will write $N = 2^n$. Given $T \subseteq [n]$, we will write $\overline{T} := [n] \setminus T$. We will write I_n to denote the $n \times n$ identity matrix; when n is clear from context, we may write I instead.

2.1 Unitary Operators We will write $\mathcal{M}_{N,N}$ to denote the set of linear operators from $\mathbb{C}^N$ to $\mathbb{C}^N$ and denote by $\mathcal{U}_N$ the set of N-dimensional unitary operators, i.e.

$$\mathcal{U}_N := \left\{U \in \mathcal{M}_{N,N} : UU^\dagger = U^\dagger U = I\right\}.$$

DEFINITION 2.1. *Given a unitary $U \in \mathcal{U}_N$ and $S \subseteq [n]$, we define the operator* $\mathrm{Tr}_S(U)$ *obtained by* tracing out S *to be*

$$\mathrm{Tr}_S(U) = \sum_{k \in \{0,1\}^S} (I_{\overline{S}} \otimes \langle k|) U (I_{\overline{S}} \otimes |k\rangle).$$

In the above definition, we write $|k\rangle$ for $k \in \{0,1\}^S$ to be the $|S|$ qubit state in the computational basis corresponding to the bit-string k. Note that Definition 2.1 aligns with the fact that the trace of a unitary matrix U is given by

$$\mathrm{Tr}(U) = \sum_{k \in \{0,1\}^n} \langle k| U |k\rangle .$$

[5] Recall that in the tolerant model, the tester is asked to distinguish instances that are ε_1-close to the property from instances that are ε_2-far from the property.

[6] In the *distribution-free* model, the distance between two functions is measured with respect to a fixed but unknown distribution.

DEFINITION 2.2. (k-JUNTA) *We say that a unitary $U \in \mathcal{U}_N$ is a* quantum k-junta *if there exists $S \subseteq [n]$ with $|S| = k$ such that*

$$U = V_S \otimes I_{\overline{S}}$$

for some $V_S \in \mathcal{U}_{2^k}$.

In contrast, *classical* k-juntas are Boolean functions $f : \{0,1\}^n \to \{0,1\}$ that depend on only k of their n input variables. More formally, a function $f : \{0,1\}^n \to \{0,1\}$ is a k-junta if there exists $g : \{0,1\}^k \to \{0,1\}$ such that $f(x_1, \dots, x_n) = g(x_{i_1}, \dots, x_{i_k})$ for some fixed $i_1, \dots, i_n \in [n]$ and for all $x \in \{0,1\}^n$.

We will view $\mathcal{M}_{N,N}$ as an inner-product space equipped with the Hilbert–Schmidt inner product

$$\langle A, B \rangle := \mathrm{Tr}\big(A^\dagger B\big).$$

Recall that the Hilbert–Schmidt inner product induces the Hilbert–Schmidt (or Frobenius) norm, which is given by

$$\|A\|^2 := \mathrm{Tr}\big(A^\dagger A\big) = \sum_{i,j=0}^{N-1} |A[i,j]|^2.$$

We will use the following metric to compare the distance between unitary matrices. Note that this metric is *not* the natural metric induced by the Hilbert–Schmidt norm.

DEFINITION 2.3. (DISTANCE BETWEEN UNITARIES) *Given $A, B \in \mathcal{M}_{N,N}$, we define*

$$\mathrm{dist}(A, B) := \min_{\theta \in [0, 2\pi)} \frac{1}{\sqrt{2N}} \|e^{i\theta} A - B\|.$$

*We say that A is ε-*far *from B if* $\mathrm{dist}(A, B) \geq \varepsilon$. *More generally, for any $\mathcal{P} \subseteq \mathcal{M}_{N,N}$ and $A \in \mathcal{M}_{N,N}$, we write*

$$\mathrm{dist}(A, \mathcal{P}) := \min_{B \in \mathcal{P}} \mathrm{dist}(A, B)$$

*and similarly say that A is ε-*far *from $\mathcal{P}$ if* $\mathrm{dist}(A, \mathcal{P}) \geq \varepsilon$.

It can easily be checked that $\mathrm{dist}(A, B) \geq 0$, with equality holding if and only if $A = e^{i\theta} B$ for some $\theta \in [0, 2\pi)$, as well as other standard properties of a metric. Finally, note that $\mathrm{dist}(V_1 \otimes U, V_2 \otimes U) = \mathrm{dist}(V_1, V_2)$ for unitaries U, V_1, V_2.

2.2 The Pauli Decomposition In this section, we introduce a useful orthonormal basis for $\mathcal{M}_{N,N}$ (viewed as a $\mathbb{C}$-vector space) which will be central to what follows. Recall that the set of Pauli operators given by

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

forms an orthonormal basis for $\mathcal{M}_{2,2}$ with respect to the Hilbert–Schmidt inner product. For $x \in \{0,1,2,3\}^n \cong \mathbb{Z}_4^n$, we define $\sigma_x := \sigma_{x_1} \otimes \cdots \otimes \sigma_{x_n}$ and write $\mathrm{supp}(x) := \{i \in [n] : x_i \neq 0\}$. It is then easy to check that the collection

$$\left\{ \frac{1}{\sqrt{N}} \sigma_x \right\}_{x \in \mathbb{Z}_4^n}$$

forms an orthonormal basis for $\mathcal{M}_{N,N}$ with respect to the Hilbert–Schmidt inner product. We will frequently refer to this basis as the *Pauli basis* for $\mathcal{M}_{N,N}$. It follows that we can write any $A \in \mathcal{M}_{N,N}$ as

$$A = \sum_{x \in \mathbb{Z}_4^n} \widehat{A}(x) \sigma_x \qquad \text{where} \qquad \widehat{A}(x) := \frac{1}{N} \langle A, \sigma_x \rangle.$$

We will sometimes refer to $\widehat{A}(x)$ as the *Pauli coefficient of A on x* and will refer to the collection $\{\widehat{A}(x)\}_x$ as the *Pauli spectrum of A*. It is easy to verify that Parseval's and Plancharel's formulas hold in this setting:

$$\frac{1}{N} \|A\|^2 = \sum_{x \in \mathbb{Z}_4^n} |\widehat{A}(x)|^2 \qquad \text{and} \qquad \frac{1}{N} \langle A, B \rangle = \sum_{x \in \mathbb{Z}_4^n} \widehat{A}(x)^* \cdot \widehat{B}(x).$$

In particular, for $U \in \mathcal{U}_N$, we have $\sum_{x \in \mathbb{Z}_4^n} |\widehat{U}(x)|^2 = 1$.

2.3 Influence of Qubits on Unitaries [MO10] introduced a notion of *influence of qubits on unitaries*, in the spirit of the well-studied classical notion of influence of variables on Boolean functions $f : \{0,1\}^n \to \{0,1\}$ (cf. Chapter 2 of [O'D14]). This notion of influence will be central to the testing algorithm presented in Section 3. Although [MO10]'s notion of influence was developed only for Hermitian unitaries (i.e. "Quantum Boolean Functions"), we first present their formulation as it gives good intuition for what influence captures, after which we introduce a more general definition of influence that applies to arbitrary unitaries as well as to more than one qubit.

DEFINITION 2.4. (DERIVATIVE OPERATOR) *The i^{th}* derivative operator D_i *is a superoperator on $\mathcal{M}_{N,N}$ defined through its action on the Pauli basis element σ_x, $x \in \mathbb{Z}_4^n$:*

$$\mathrm{D}_i\sigma_x = \begin{cases} \sigma_x & x_i \neq 0 \\ 0 & x_i = 0 \end{cases}.$$

It follows immediately that for $A \in \mathcal{M}_{N,N}$, $A = \sum_{x\in\mathbb{Z}_4^n} \widehat{A}(x)\sigma_x$, we have

$$\mathrm{D}_i A = \sum_{x:x_i\neq 0} \widehat{A}(x)\sigma_x \tag{2.2}$$

Informally, D_i isolates the part of the Pauli spectrum that acts non-trivially on the i^{th} qubit (i.e. the x such that $\sigma_{x_i} \neq I$). We can now introduce the notion of influence of qubits on unitaries proposed by [MO10].

DEFINITION 2.5. (INFLUENCE OF SINGLE QUBIT) *Given a unitary $U \in \mathcal{U}_N$, the* influence of the i^{th} qubit on U, *written* $\mathbf{Inf}_i[U]$, *is*

$$\mathbf{Inf}_i[U] := \|\mathrm{D}_i U\|^2.$$

At a high level, the influence of the i^{th} qubit on a unitary U captures how non-trivially the unitary U acts on the i^{th} qubit of a quantum state. Note that it is immediate from Equation (2.2) that

$$\mathbf{Inf}_i[U] = \sum_{x:x_i\neq 0} |\widehat{U}(x)|^2.$$

This suggests a natural way to extend the Definition 2.5 to more than one qubit.

DEFINITION 2.6. (INFLUENCE OF MULTIPLE QUBITS) *Given a unitary $U \in \mathcal{U}_N$ and $S \subseteq [n]$, the* influence of S on U, *written* $\mathbf{Inf}_S[U]$, *is*

$$\mathbf{Inf}_S[U] = \sum_{x:\mathrm{supp}(x)\cap S\neq\emptyset} |\widehat{U}(x)|^2. \tag{2.3}$$

The above definition is analogous to the "Fourier formula" for the the influence of a set of variables on a Boolean function (cf. Section 2.4 of [ABRdW16]). Furthermore, as stated earlier, note that these definitions apply to arbitrary unitaries (i.e. we do not require them to be Hermitian). We present an alternative characterization of $\mathbf{Inf}_S[U]$ (which we will not require, but may be of independent interest) in Appendix A. We have the following lemma.

LEMMA 2.1. *For $S, T \subseteq [n]$ and a unitary $U \in \mathcal{U}_N$, we have*

1. *Monotonicity: If $S \subseteq T$, $\mathbf{Inf}_S[U] \leq \mathbf{Inf}_T[U]$; and*
2. *Subadditivity: $\mathbf{Inf}_{S\cup T}[U] \leq \mathbf{Inf}_S[U] + \mathbf{Inf}_T[U]$.*

Note that monotonicity is immediate from the analytic interpretation of influence (cf. Equation (2.3)), and subadditivity follows from the fact that

$$\{(S \cup T) \cap \mathrm{supp}(x)\} = \{S \cap \mathrm{supp}(x)\} \cup \{T \cap \mathrm{supp}(x)\}.$$

As mentioned before, Wang [Wan11] implicitly used this notion of influence to test quantum k-juntas. In particular, Wang proved the following.

LEMMA 2.2. ([WAN11]) *Given a unitary $U \in \mathcal{U}_N$, if U is ε-far from every quantum k-junta V, then for all $T \subseteq [n]$ with $|T| \leq k$, we have that*

$$\mathbf{Inf}_{\overline{T}}[U] \geq \frac{\varepsilon^2}{4}.$$

2.4 Query Complexity of the Composition of Quantum Algorithms Our upper bound for quantum junta testing will rely on the following lemma about the query complexity of the composition of quantum algorithms.

LEMMA 2.3. ([ABRDW16], COROLLARY 2.12) *With $D \subset \{0,1\}^n$, let $F : D \to \{0,1\}$ and G_j be partial Boolean functions $\forall j \in [n]$. Let $Q(F)$ denote the bounded-error quantum query complexity of F. Let T equal to the objective value of a feasible solution (X_j) to the adversarial bound in (2.3) of [ABRdW16]. We let an input variable j be irrelevant for input $z \in D$ if and only if $X_j[z,z] = 0$. Then, we have*

$$Q(F \circ (G_1, ...G_n)) = O\left(T \max_{j \in [n]} Q(G_j)\right).$$

with the function composition done as in Definition 2.10 of [ABRdW16].

Note that the notion of function composition in [ABRdW16] is more subtle than direct function composition, as the latter would incur a logarithmic overhead in query complexity.

Our junta testing algorithms involve composing quantum algorithms, each of which have small error probabilities. The above lemma allows us to compose them without incurring logarithmic overheads; the details of our approach is identical to that of [ABRdW16] to which we defer the technical details.

As a rough overview, [ABRdW16] defines two key notions: First is the notion of "robust conjunctions," which are essentially error-resistant conjunctions, and second is the notion of "irrelevant variables for an algorithm," i.e. input variables which do not affect the algorithm's output. In the context of function composition, we can ensure that F depends only on G_j that are relevant to the tester, and in our analysis we can ignore cases where irrelevant G_j behave unpredictably.

2.5 The Choi-Jamiolkowski Isomorphism In our algorithms, we will encode a unitary as a quantum state using the *Choi-Jamiołkowski isomorphism* [Cho75, Jam72], which is a mapping between $N \times N$ unitary operators and pure states in $\mathbb{C}^N \otimes \mathbb{C}^N$. Concretely, this mapping associates to every unitary $U \in \mathcal{U}_N$ the *Choi-Jamiołkowski state* (which we abbreviate as *CJ state*):

$$|v(U)\rangle := (U \otimes I)\left(\frac{1}{\sqrt{N}} \sum_{0 \leq i < N} |i\rangle\,|i\rangle\right) = \frac{1}{\sqrt{N}} \sum_{0 \leq i,j < N} U[i,j]\ |i\rangle\,|j\rangle\,.$$

The CJ state $|v(U)\rangle$ can be prepared by first creating the maximally entangled state of dimension N, and then querying U on half of the maximally entangled state. Since $N = 2^n$, this is equivalent to preparing n EPR pairs (which altogether forms $2n$ qubits) and applying the unitary U to the n qubits coming from the first half of each of the EPR pairs. As such, each qubit of the unitary U corresponds to two qubits of the state $|v(U)\rangle$. We will refer to qubits in $\{1, \ldots, n\}$ as the ones acted on by the unitary U, and qubits in $\{n+1, \ldots,\ 2n\}$ as the ones acted on by I. We introduce the following notation for convenience.

NOTATION 2.1. *For each qubit $\ell \in [n]$ acted on by the unitary U, there is a pair of corresponding qubits $(\ell, \widetilde{\ell}) \in [n] \times \{n+1, \ldots, 2n\}$ in the state $|v(U)\rangle$. In particular, $\widetilde{\ell}$ and ℓ are related as they formed an EPR pair at the synthesis of the CJ state.*

3 Testing Quantum k-Juntas with $\widetilde{O}(\sqrt{k})$ Queries

As suggested by Lemma 2.1, the notion of influence for unitaries behaves analogously to the "usual" notion of influence for Boolean functions, which was crucial to the $\widetilde{O}(\sqrt{k})$-query k-junta tester for Boolean functions obtained by Ambainis et al. [ABRdW16]. This motivates an analog of the algorithm obtained by Ambainis et al. for quantum juntas, and this is indeed how we obtain a $\widetilde{O}(\sqrt{k})$-tester for quantum k-juntas. In Section 3.1, we present an unbiased estimator for the influence of qubits on a unitary, which we then combine with Ambainis et al.'s tester in Section 3.2 to obtain our quantum k-junta tester.

Input: Oracle access to $U \in \mathcal{U}_N$, $S \subset [n]$

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

RAW-INFLUENCE-ESTIMATOR(U, S):

1. Prepare the Choi-Jamiolkowski state $|v(U)\rangle$ given by
$$|v(U)\rangle = \frac{1}{\sqrt{N}} \sum_{0 \le i,j < N} U[i,j] \, |i\rangle \, |j\rangle \, .$$
This is prepared by querying U once on the maximally entangled state.
2. Measure the $2|S|$ qubits in the registers $S \cup \{\widetilde{\ell} : \ell \in S\}$ in the Bell basis and let $|\varphi\rangle$ denote the post-measurement state.
   (a) Test if $|\varphi\rangle$ is equal to $|\mathrm{EPR}\rangle^{\otimes |S|}$, return 0.
   (b) Otherwise, return 1.

Figure 1: Influence Estimator for Quantum Unitaries

3.1 An Influence Tester for Unitaries We start by describing a subroutine RAW-INFLUENCE-ESTIMATOR(cf. Figure 1) that allows us to estimate the influence of a set of variables $S \subseteq [n]$ on a unitary U.

LEMMA 3.1. *Let X denote the output of* RAW-INFLUENCE-ESTIMATOR(U, S) *for $U \in \mathcal{U}_N$ and $S \subseteq [n]$ as described in Figure 1. Then*
$$\mathbf{E}[X] = \mathbf{Inf}_S[U].$$

Proof. Recall that U can be written in the Pauli basis as $U = \sum_{x \in \mathbb{Z}_4^n} \widehat{U}(x) \sigma_x$. Thus, $|v(U)\rangle$ can be written as
$$\begin{aligned} |v(U)\rangle &= \sum_{x \in \mathbb{Z}_4^n} \widehat{U}(x) \, |v(\sigma_x)\rangle \\ &= \sum_{x : \mathrm{supp}(x) \cap S = \emptyset} \widehat{U}(x) \, |v(\sigma_x)\rangle + \sum_{x : \mathrm{supp}(x) \cap S \neq \emptyset} \widehat{U}(x) \, |v(\sigma_x)\rangle \\ &= \sum_{x : \mathrm{supp}(x) \cap S = \emptyset} \widehat{U}(x) \, |v(\sigma_{x_{\overline{S}}})\rangle \, |v(I^{\otimes |S|})\rangle + \sum_{x : \mathrm{supp}(x) \cap S \neq \emptyset} \widehat{U}(x) \, |v(\sigma_{x_{\overline{S}}})\rangle \, |v(\sigma_{x_S})\rangle \, . \end{aligned}$$

Where $x_S \in \mathbb{Z}_4^S$ is notation for the restriction of x onto the qubits in S. Similarly, σ_{x_S} is the Pauli basis vector given by the tensor product of $|S|$ Pauli matrices according to x_S. Thus, for any $x \in \mathbb{Z}_4^n$ such that $\mathrm{supp}(x) \cap S \neq \emptyset$, the state $|v(\sigma_{x_S})\rangle$ is orthogonal to the state $|v(I^{\otimes |S|})\rangle = |\mathrm{EPR}\rangle^{\otimes |S|}$. Because $\{|\widehat{U}(x)|^2\}_{x \in \mathbb{Z}_4^n}$ forms a probability distribution, when Figure 1 measures the qubits in $S \cup \{\widetilde{\ell} : \ell \in S\}$, it will return 1 with the following probability.
$$\begin{aligned} \mathbf{E}[X] &= \mathbf{Pr}[X = 1] \\ &= \sum_{x : \mathrm{supp}(x) \cap S \neq \emptyset} |\widehat{U}(x)|^2 \\ &= \mathbf{Inf}_S[U] \end{aligned}$$
This completes the proof. ∎

Note that we can boost the probability that RAW-INFLUENCE-ESTIMATOR outputs 1 via amplitude amplification (see, for example, Section 2.2 of [MO10]). In particular, we can amplify the probability of RAW-INFLUENCE-ESTIMATOR outputting 1 from δ to an arbitrary constant (say 0.9) via $O(1/\sqrt{\delta})$ calls to the oracles for the unitary U. Thus, we have the following lemma.

Input: Oracle access to $U, U^\dagger \in \mathcal{U}_N$, $S \subseteq [n]$, $\delta \in (0,1]$

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

INFLUENCE-ESTIMATOR(U, S, δ):

1. Use amplitude amplification with $O(1/\sqrt{\delta})$ calls to RAW-INFLUENCE-ESTIMATOR(U, S).
2. Return the same value as RAW-INFLUENCE-ESTIMATOR(U, S).

Figure 2: Influence Estimator via Amplitude Amplification

LEMMA 3.2. *Let $U \in \mathcal{U}_N$ and $S \subseteq [n]$. If $\mathbf{Inf}_S[U] \geq \delta$, then* INFLUENCE-ESTIMATOR(U, S, δ) *as described in Figure 2 outputs* 1 *with probability at least 9/10, and if $\mathbf{Inf}_S[U] = 0$, then* INFLUENCE-ESTIMATOR(U, S, δ) *always outputs* 0*. Furthermore, the number of queries made to U is $O(1/\sqrt{\delta})$.*

3.2 Reducing to Gapped Group Testing Using our influence estimator INFLUENCE-ESTIMATOR, we can now reduce the problem of testing quantum juntas to that of Gapped Group Testing (GGT), which we define below. Our approach closely follows that of Ambainis et al. [ABRdW16], who reduce the problem of testing k-juntas to GGT. We remark that certain parameters in our adaptation of Ambainis et al.'s algorithm will be worse by a square-root factor, resulting in an overall query complexity of $\widetilde{O}(\sqrt{k}/\varepsilon)$ for testing quantum k-juntas as opposed to $\widetilde{O}(\sqrt{k/\varepsilon})$ as obtained by Ambainis et al. for testing classical juntas.

We first define the exact version of Group Testing.

DEFINITION 3.1. (EGGT) *Let k and d be positive integers, $\mathcal{X}$ consist of all subsets of $[n]$ with size k, and $\mathcal{Y}$ consist of all subsets of $[n]$ of size $k + d$. In the Exact Gapped Group Testing (EGGT) problem, we are given oracle access to the function* Intersects_A, $A \in \mathcal{X} \cup \mathcal{Y}$ *and must decide whether $A \in \mathcal{X}$ or if $A \in \mathcal{Y}$*

The exact GGT will be referenced in the analysis. However, the actual algorithm we will use in our algorithm solves a more general version of EGGT.

DEFINITION 3.2. (GGT) *Let k and d be positive integers. Define two families of functions*

$$\widetilde{\mathcal{X}} = \{f : \{0,1\}^n \to \{0,1\} \mid \exists A \in \mathcal{X}\, \forall S \subset [n] : S \cap A = \emptyset \implies f(S) = 0\}$$

$$\widetilde{\mathcal{Y}} = \{f : \{0,1\}^n \to \{0,1\} \mid \exists B \in \mathcal{Y}\, \forall S \subset [n] : S \cap B \neq \emptyset \implies f(S) = 1\}$$

In an instance of $\text{GGT}(k, d)$*, given oracle access to some function $f \in \widetilde{\mathcal{X}} \cup \widetilde{\mathcal{Y}}$, decide whether $f \in \widetilde{\mathcal{X}}$ or $f \in \widetilde{\mathcal{Y}}$.*

Note that if the function f is in $\widetilde{\mathcal{X}}$, then sets S such that $S \cap A \neq \emptyset$ do not restrict f. They are "irrelevant." Similarly, if f is in $\widetilde{\mathcal{Y}}$, sets S such that $S \cap B = \emptyset$ are "irrelevant" (cf. Section 2.4). More precisely, the sets that are deemed irrelevant follow from the adversary bound and is explained in more detail in Observation 3.9 of [ABRdW16]. Also, note that if we replace implication symbols in Definition 3.2 with equivalence symbols, we recover the EGGT problem. Thus, EGGT is a special case of GGT.

To get some intuition for Definition 3.2, consider the following scenario: Given n soldiers, some of which are sick, you would like to determine whether there are at most k sick soldiers, or if there are at least $k + d$ sick soldiers. You are allowed to test this by pooling blood samples from subsets of the n soldiers, where the pooled test returns positive if the group contains at least one sick soldier.

More precisely, for an unknown $A \subseteq [n]$, we would like to decide if $|A| \leq k$ or $|A| \geq k + d$ given access to the following oracle

$$\text{Intersects}_A(S) := \begin{cases} 1 & A \cap S \neq \emptyset \\ 0 & \text{otherwise} \end{cases}.$$

We briefly explain the connection to junta testing: Given a unitary U and a fixed threshold $\delta > 0$, let $S_\delta \subseteq [n]$ be the set of qubits whose influence is at least δ. Note then that INFLUENCE-ESTIMATOR(U, T, δ) will will return

Input: Oracle access to $U, U^\dagger$, parameter k

Output: "Yes" or "No"

UNITARY-JUNTA-TESTER(U, k)

1. Run TESTER-I(U, k, l) for $l \in \{0, \ldots, \lfloor \log(200k) \rfloor\}$.
2. Run TESTER-II(U, k).
3. Output "Yes" if all $\lfloor \log(200k) \rfloor + 2$ testers above accept, and output "No" otherwise.

Figure 3: Quantum k-Junta Tester

Input: Oracle access to $U, U^\dagger$, parameter k, parameter l

Output: "Yes" or "No"

TESTER-I(U, k, l):

1. Let $d_l = 2^l$ and $\delta_l = \frac{\varepsilon^2}{2^{l+5} \log(400k)}$.
2. Run QUANTUM-GGT with parameters k and $d = d_l$, and query access to the following oracle:

   Given $S \subseteq [n]$, output INFLUENCE-ESTIMATOR(U, S, δ_l)
3. Output "Yes" if GGT accepts, and "No" otherwise.

Figure 4: Tester of the First Kind

1 with high probability if at least one variable in S is in S_δ. In this sense, we have that

$$\text{INFLUENCE-ESTIMATOR}(U, T, \delta) \qquad \approx \qquad \text{Intersects}_{S_\delta}(T)$$

By examining various settings of δ, we can use GGT to infer the "distribution" of influence of a unitary U among its qubits. We will make use of the following quantum algorithm obtained by Ambainis et al. for GGT.

THEOREM 3.1. (THEOREM 3.6 OF [ABRDW16]) *There exists a quantum algorithm* QUANTUM-GGT *that solves* GGT(k, d) *using* $O(\sqrt{1 + k/d})$ *queries.*

Our algorithm for quantum junta testing and analysis thereof closely follow the structure of Ambainis et al.'s algorithm for junta testing and its analysis; we include complete details below for completeness but refer the interested reader to Section 4 of [ABRdW16] for the original algorithm.

Note that because our INFLUENCE-ESTIMATOR serves as a subroutine to the GGT algorithm, there is a need for a careful analysis of the properties of their composition. This is addressed in Section 2.4 at a high level and addressed in more detail in [ABRdW16].

THEOREM 3.2. *Given* $U \in \mathcal{U}_N$*, with high probability* 9/10*, the algorithm* UNITARY-JUNTA-TESTER(U) *outputs "Yes" if* U *is a* k*-junta, and outputs "No" if* U *is* ε*-far from every quantum* k*-junta. Furthermore,* UNITARY-JUNTA-TESTER(U) *makes* $O\left(\frac{\sqrt{k \log k}}{\varepsilon} \log k\right)$ *calls to the unitary* U *and has two-sided error.*

Proof. The setup and analysis of the algorithm (Lemmas 3.3 and 3.4, 3.5) is almost the same as in [ABRdW16], with a few constants changed.

Without loss of generality, we assume that the first K qubits are the most influential ones and are ordered in decreasing amount of influence.

$$\mathbf{Inf}_1[U] \geq \mathbf{Inf}_2[U] \geq \ldots \geq \mathbf{Inf}_K[U] > 0 = \mathbf{Inf}_{K+1}[U] = \ldots = \mathbf{Inf}_n[U].$$

Input: Oracle access to $U, U^\dagger$, parameter k

Output: "Yes" or "No"

TESTER-II(U, k):

1. Estimate acceptance probability of following subroutine up to additive error 0.05:
   - Generate $\boldsymbol{S} \subset [n]$ by adding $i \in [n]$ to $\boldsymbol{S}$ with probability $1/k$ independently.
   - Run INFLUENCE-ESTIMATOR$(U, \boldsymbol{S}, \delta)$ where $\delta := \frac{\varepsilon^2}{16k}$.
2. Output "Yes" if estimated acceptance probability is at most 0.8, and "No" otherwise.

Figure 5: Tester of the Second Kind

Of course, the tester does not know this order. The primary challenge is in showing that if U is ε-far from every quantum k-junta, then at least one of the two subroutines TESTER-I and TESTER-II will output "No" with significant probability. The $\lfloor \log(200k) \rfloor + 2$ tests in the main Figure 3 are tailored for this purpose; in particular, we have the two following cases when U is ε-far from every quantum k-junta:

1. **Case 1:** $\sum_{j=k+1}^{200k} \mathbf{Inf}_j[U] \geq \varepsilon^2/8$. This case is further split into $\lfloor \log 200k \rfloor + 1$ subcases:

$$\left| \left\{ j \in [n] : \mathbf{Inf}_j[U] \geq \frac{\varepsilon^2}{2^{l+5} \log(400k)} \right\} \right| \geq k + 2^l$$

   for $l \in \{0, ..., \lfloor \log(200k) \rfloor\}$. We say that a unitary U is a *non-junta of the first kind* if this is the case for some $l \in \{0, ..., \lfloor \log(200k) \rfloor\}$.

2. **Case 2:** $\sum_{j=k+1}^{200k} \mathbf{Inf}_j[U] \leq \varepsilon^2/8$. We say U is a *non-junta of the second kind* if this is the case.

Lemma 3.3 says that any unitary U that is ε-far from every quantum k-junta satisfies at least one of the two cases above. The correctness and query complexity of Figure 3 now follows from Lemmas 3.3, 3.4 and 3.5. ∎

Finally, we prove the auxiliary lemmas used in the proof of the above theorem. Lemmas 3.3, 3.4 and 3.5 are analogous to Lemmas 4.3 to 4.5 of [ABRdW16].

LEMMA 3.3. *Every U that is ε-far from being a quantum k-junta satisfies one of the two cases above.*

Proof. It suffices to show that if U is a non-junta of the first kind, then at least one of the $\lfloor \log 200k \rfloor + 1$ sub-cases holds. By definition, we have

$$\sum_{j=k+1}^{200k} \mathbf{Inf}_j[U] \geq \varepsilon^2/8.$$

Define

$$\varepsilon' = \frac{\varepsilon^2}{32 \log(400k)}$$

and consider the partition of $[0, 1]$ given by

$$A_\infty = \left[0, \frac{\varepsilon'}{2^{\lfloor \log 200k \rfloor}}\right), \quad A_0 = [\varepsilon', 1], \quad A_l = \left[\frac{\varepsilon'}{2^l}, \frac{\varepsilon'}{2^{l-1}}\right)$$

where $l \in \{\lfloor \log 200k \rfloor, \ldots, 1\}$. Define $B_l := \{j \in \{k+1, ..., 200k\} : \mathbf{Inf}_j[U] \in A_l\}$, and note that each $j \in [n]$ is included in exactly one of the B_l. Writing

$$W_l = \sum_{j \in B_l} \mathbf{Inf}_j[U] \qquad \text{we have} \qquad \sum_l W_l \geq \varepsilon^2/8$$

as U is a non-junta of the first kind. We also have that

$$W_\infty < 200k\left(\frac{\varepsilon^2}{32 \cdot 2^{\lfloor \log 200k \rfloor}}\right) < \frac{\varepsilon^2}{16}.$$

Since the maximum of the W_l's is at least their average, there exists $l^* \in \{0, 1, ...\lfloor \log 200k \rfloor\}$ such that

$$W_{l^*} \geq \frac{\varepsilon^2}{16 \log 400k},$$

which in turn implies

$$|B_l| \geq \frac{\frac{\varepsilon^2}{16 \log 400k}}{\frac{\varepsilon^2 \cdot 2^{1-l}}{32 \log 400k}} = 2^l.$$

Every variable $j \in B_l$ has influence at least $\frac{\varepsilon'}{2^l} =: \delta_l$. Furthermore, since the influence of variables are ordered in decreasing order, each variable $j \in [k]$ also has at least δ_l influence. Thus, there are at least $k + 2^l$ indices j such that $\mathbf{Inf}_j[U] \geq \delta_l$, and U satisfies the first case for this particular l. ∎

LEMMA 3.4. *If U is a k-junta, then all calls to* TESTER-I *will accept with high probability. If U is a non-junta of the first kind, then one of the calls to* TESTER-I *will reject with high probability. Finally, the overall query complexity of all $\lfloor \log 200k \rfloor + 1$ testers of the first kind is*

$$O\left(\frac{\sqrt{k \log k}}{\varepsilon} \log k\right).$$

Proof. The composition in TESTER-I is done as described in Definition 2.10 of [ABRdW16] which allows for a tight query-complexity. Towards this definition, F and (G_j) are defined as follows: The partial function F is the EGGT function from Definition 3.1. F takes in a function h and outputs 0 if $h = \text{Intersects}_A, |A| = k$ and 1 if $h = \text{Intersects}_A, |A| = k + d$. In other cases, F is undefined. For each $S \subset [n]$, the partial function G_S is our INFLUENCE-ESTIMATOR on set S. G_S is partial in that it equals 1 if $\mathbf{Inf}_S[U] \geq \delta$, equals 0 if $\mathbf{Inf}_S[U] = 0$, but is undefined for anything in between. Thus, TESTER-I is equivalent to the following composition:

$$U \to (G_\emptyset(U), G_{\{1\}}(U), G_{\{2\}}(U), ...G_{[n]}(U)) \tag{3.4}$$

The irrelevant variables to the function F correspond to the sets S that do not impact its output; that is, whatever INFLUENCE-ESTIMATOR outputs on these sets do not matter to F. Because we use the same GGT algorithm, derived from the same solution to the adversary bound as [ABRdW16], we have the same irrelevant variables.

1. If the input A is in $\mathcal{X}(|A| = k)$, a set $S \subset [n]$ is irrelevant if $S \cap A \neq \emptyset$. That is, if U is a k-Junta, TESTER-I only looks at sets such that $S \cap A = \emptyset$.
2. If the input A is in $\mathcal{Y}(|A| = k + d)$, a set $S \subset [n]$ is irrelevant if $|S \cap A| \neq 1$. In particular, if U is $\varepsilon-$far from a k-Junta, TESTER-I ignores sets such that $S \cap A = \emptyset$

Suppose U is a non-junta of the first kind, satisfying case l, in the sense of Lemma 3.3. By definition, there is an $A \subset [n], |A| = k + 2^l$ such that for all $j \in A$, $\mathbf{Inf}_j[U] \geq \delta_l$. By the monotonicity of influence, $\mathbf{Inf}_S[U] \geq \delta$ for all S that intersect A. Finally, because the sets that are disjoint from A are irrelevant in the non-junta case, TESTER-I's oracle behaves like an $\text{Intersect}_\text{A}$ oracle that depends on at least $k+2^l$ indices. Thus, this instantiation of Figure 4's GGT will reject with high probability.

Finally, if U is a k-junta, then there is a set $A \subset [n], |A| \leq k$ such that if $S \cap A = \emptyset$, then $\mathbf{Inf}_S[U] = 0$. Because all sets $S \cap A \neq \emptyset$ are irrelevant in the k-junta case, TESTER-I's oracle behaves like an $\text{Intersect}_\text{A}$ oracle that depends on k indices. Thus, all the TESTER-I's will accept with high probability as there are at most k influential variables.

Thus, the tester of the first kind, a group tester instantiated with $d = 2^l$ and δ_l, will be able to distinguish between this case from case where U is a k-junta, where the set of variables of influence at least δ_l is size at most k.

Finally, for a particular value of l, the query complexity of the influence tester is $O({\delta_l}^{-1/2})$ while the query complexity of the corresponding group tester instance is $O(\sqrt{k/d_l})$. It then follows by Lemma 2.3 that the complexity of any tester of the first kind is

$$O\left(\sqrt{\frac{k}{2^l}} \cdot \sqrt{\frac{2^l \log 400k}{\varepsilon^2}}\right) = O\left(\frac{\sqrt{k \log k}}{\varepsilon}\right)$$

giving an overall query complexity of

$$O\left(\frac{\sqrt{k \log k}}{\varepsilon} \log k\right)$$

for all $\lfloor \log(200k) \rfloor + 1$ testers of the first kind. ∎

LEMMA 3.5. *Figure 5 accepts if U is a k-junta and rejects if U is a non-junta of the second kind, and its query complexity is $O(\sqrt{k/\varepsilon})$*

Proof. We show that the procedure described in Item 1 of Figure 5 has acceptance probability *at most* 0.75 if U is a k-junta, and has acceptance probability *at least* 0.85 if U is a non-junta of the second kind.

Suppose U is a k-junta. Then the probability that the set S does not intersect the set J of relevant variables is

$$\left(1 - \frac{1}{k}\right)^{|J|} \geq \left(1 - \frac{1}{k}\right)^{k} \geq \frac{1}{4}.$$

Therefore, with probability at least 0.25, we have $S \cap J = \emptyset$ in which case $\mathbf{Inf}_S[U] = 0$. It follows then that the acceptance probability above is at most 0.75.

Now suppose U is a non-junta of the second kind. For $j \in [n]$, define

$$\underline{\mathbf{Inf}}_j[U] := \begin{cases} 0 & j \leq 200k \\ \sum_{x:\mathrm{supp}(x) \cap \{200k+1 \ldots j\} = \{j\}} |\widehat{U}(x)|^2 & \text{otherwise} \end{cases}.$$

For $S \subset [n]$, define $\underline{\mathbf{Inf}}_S[U] := \sum_{j \in S} \underline{\mathbf{Inf}}_j[U]$. It is easy to see that

$$\underline{\mathbf{Inf}}_S[U] \leq \mathbf{Inf}_S[U],$$

and that for $S, T \subseteq [n]$ with $S \cap T = \emptyset$, we have

$$\underline{\mathbf{Inf}}_{S \cup T}[U] = \underline{\mathbf{Inf}}_S[U] + \underline{\mathbf{Inf}}_T[U].$$

Now, because U is ε-far from every quantum k-junta, by Lemma 2.2, we have that

$$\mathbf{Inf}_{\{k+1 \ldots K\}}[U] \geq \varepsilon^2/4, \tag{3.5}$$

and since U is a non-junta of the second kind,

$$\sum_{j=k+1}^{200k} \mathbf{Inf}_j[U] \leq \frac{\varepsilon^2}{8}. \tag{3.6}$$

Combining Equations (3.5) and (3.6) we get that

$$\begin{aligned} \underline{\mathbf{Inf}}_{[n]}[U] &= \mathbf{Inf}_{\{200k+1 \ldots K\}}[U] \\ &\geq \mathbf{Inf}_{\{k+1 \ldots K\}}[U] - \sum_{j=k+1}^{200k} \mathbf{Inf}_j[U] \\ &\geq \frac{\varepsilon^2}{8}. \end{aligned} \tag{3.7}$$

Consider now the random variable $\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U]$ where $\boldsymbol{S}$ is drawn as described in Figure 5. We have

$$\mu := \mathop{\mathbb{E}}_{\boldsymbol{S}}[\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U]] = \frac{1}{k} \cdot \underline{\mathbf{Inf}}_{[n]}[U] \geq \frac{\varepsilon^2}{8k}.$$

We also have

$$\begin{aligned}
\sigma^2 := \mathbf{Var}\left[\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U]\right] &\leq \frac{1}{k}\sum_j \underline{\mathbf{Inf}}_j[U]^2 \\
&\leq \frac{1}{k}\left(\max_j \underline{\mathbf{Inf}}_j[U]\right) \cdot \underline{\mathbf{Inf}}_{[n]}[U] \\
&\leq \frac{1}{k} \cdot \left(\frac{\varepsilon^2}{4 \cdot 200k}\right) \underline{\mathbf{Inf}}_{[n]}[U] \\
&\leq \frac{\mu^2}{100}.
\end{aligned}$$

It then follows by Chebyshev's inequality that

$$\mathbf{Pr}\left[\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U] < \frac{\varepsilon^2}{16k}\right] \leq \mathbf{Pr}\left[|\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U] - \mu| > \frac{\mu}{2}\right] \tag{3.8}$$

$$\leq \mathbf{Pr}\left[|\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U] - \mu| > 5\sigma\right] \tag{3.9}$$

$$\leq \frac{1}{25}. \tag{3.10}$$

In other words, the probability $\underline{\mathbf{Inf}}_{\boldsymbol{S}}[U] > \varepsilon^2/16k$ is at least 0.96. So the acceptance probability of the subroutine described in Item 1 of Figure 5 on $\boldsymbol{S}$ is at least $0.9 \times 0.96 > 0.85$ if U is a non-junta of the second kind.

Finally, the subroutine of the tester of the second kind only makes queries to Influence-Estimator on $\delta = \frac{\varepsilon^2}{16k}$, which requires complexity $O(\sqrt{k/\varepsilon^2})$, and the outer estimation overhead is a constant. ∎

4 An $\Omega(\sqrt{k})$ Lower Bound for Testing Quantum k-Juntas

In this section, we obtain an $\Omega(\sqrt{k})$ lower bound for testing quantum k-juntas, which shows that the algorithm obtained in Section 3 is essentially optimal (up to polylogarithmic factors in k). Our lower bound follows via a natural reduction from testing classical k-juntas to testing quantum k-juntas, combined with the $\Omega(\sqrt{k})$ lower bound for testing classical k-juntas obtained by Bun, Kothari, and Thaler [BKT17]. The key technical insight here is in Lemma 4.1, which shows that every quantum k-junta is (in a certain sense) "close" to a quantum Boolean function (i.e. a Hermitian quantum unitary).

In what follows, we say that an algorithm is a (k, ε)-*classical* (respectively *quantum*) *junta tester* if, given query access to a Boolean function $f : \{0,1\}^n \to \{0,1\}$ (respectively unitary $U \in \mathcal{U}_N$), with probability at least 9/10 it outputs

- "Yes" if f (respectively U) is a k-junta; and
- "No" if f (respectively U) is ε-far from every k-junta.

Theorem 4.1. *Every T-query $(k, \sqrt{\varepsilon/2})$-quantum junta tester is also a T-query (k, ε)-classical junta tester.*

Note that Theorem 4.1 together with the $\Omega(\sqrt{k})$ lower bound for quantum testing of k-juntas by Bun, Kothari, and Thaler [BKT17] implies the desired lower bound. Before proving Theorem 4.1, we first introduce some notation. Given a Boolean function $f : \{0,1\}^n \to \{0,1\}$, we will write

$$U_f := \mathrm{diag}\Big((-1)^{f(x)}\Big) \tag{4.11}$$

as a diagonal matrix whose diagonal entries are the 2^n values of the function f. Note that U_f is unitary as its singular values are ± 1. Also, given a matrix A, we will use $A[i,j]$ to mean the entry of A at row i and column j.

The transformation we use to reduce Boolean functions to quantum Boolean functions (towards the goal of proving Theorem 4.1) is the natural one given by Equation (4.11). First, if a function $f : \{0,1\}^n \to \{0,1\}$ is a k-junta, then U_f is also a quantum k-junta. To see this, suppose without loss of generality that the last k bits of f are the relevant ones[7], i.e. we have

$$f(x) = \widetilde{f}(x_{n-k+1}, \ldots, x_n)$$

for some $\widetilde{f} : \{0,1\}^k \to \{0,1\}$. It then follows that

$$U_f = I^{\otimes(n-k)} \otimes U_{\widetilde{f}}.$$

The following lemma shows that an analogous statement holds when f is far from being a k-junta, from which Theorem 4.1 is immediate.

PROPOSITION 4.1. *If $f : \{0,1\}^n \to \{0,1\}$ is ε-far from every k-junta, then U_f is $\sqrt{\varepsilon/2}$-far from every quantum k-junta.*

Proof. We will first show if $g : \{0,1\}^n \to \{0,1\}$ is a k-junta, then

$$\operatorname{dist}(U_f, U_g) \geq \sqrt{2\varepsilon}. \tag{4.12}$$

As f is ε-far from g, we have that

$$\mathbf{Pr}\left[f \neq g\right] \geq \varepsilon.$$

Consider U_g, the unitary whose diagonal entries are the values of g, as we did with f above. The distance between U_g and U_f is at least

$$\begin{aligned}
\operatorname{dist}(U_f, U_g)^2 &= \frac{1}{2N} \cdot \min_{\theta} \|e^{i\theta} U_f - U_g\|^2 \\
&= \min\left(\frac{1}{2N}\|U_f - U_g\|^2, \frac{1}{2N}\| - U_f - U_g\|^2\right) \qquad (4.13) \\
&= \min\left(\frac{2}{N}\sum_{x\in\{0,1\}^n}\left(\frac{f(x)-g(x)}{2}\right)^2, \frac{2}{N}\sum_{x\in\{0,1\}^n}\left(\frac{f(x)+g(x)}{2}\right)^2\right) \\
&= 2\min\left(\mathbf{Pr}[f \neq g], \mathbf{Pr}[f = g]\right) \\
&\geq 2\varepsilon. \qquad (4.14)
\end{aligned}$$

Equation 4.13 holds because U_f and U_g are both diagonal with real entries, so the only possible phases that would minimize the Frobenius norm of their difference are $\theta = 0$ or π. Equation 4.14 holds because k-juntas are closed under negation; in more detail, if g is a k-junta, then $1 - g$ is also a k-junta and so

$$\mathbf{Pr}[f = g] = \mathbf{Pr}[f \neq 1 - g] \geq \varepsilon.$$

We thus have that $\operatorname{dist}(U_f, U_g) \geq \sqrt{2\varepsilon}$. In order to prove the lemma, it suffices to show that for any quantum k-junta V, there exists a Boolean k-junta $g : \{0,1\}^n \to \{0,1\}$ such that $\operatorname{dist}(V, U_g) \leq \operatorname{dist}(V, U_f)$. This is proved in Lemma 4.1.

To see why this suffices, note that if this were the case, then by the triangle inequality,

$$\operatorname{dist}(V, U_f) + \operatorname{dist}(V, U_g) \geq \operatorname{dist}(U_f, U_g).$$

However, as $\operatorname{dist}(U_f, U_g) \geq \sqrt{2\varepsilon}$ by Equation (4.14), and as $\operatorname{dist}(V, U_g) \leq \operatorname{dist}(V, U_f)$ by Lemma 4.1, we have that

$$2 \cdot \operatorname{dist}(V, U_f) \geq \sqrt{2\varepsilon}$$

and so the result follows. ∎

[7] We will use this indexing convention for the remaining sections as well.

LEMMA 4.1. *Suppose $f : \{0,1\}^n \to \{0,1\}$ is ε-far from every k-junta. Then, for every quantum k-junta $V \in \mathcal{U}_N$, there exists some Boolean function $g : \{0,1\}^n \to \{0,1\}$ that is a k-junta for which*

$$\operatorname{dist}(V, U_g) \leq \operatorname{dist}(V, U_f).$$

Proof. We can assume without loss of generality that V is a quantum k-junta on the last k qubits. We define $g : \{0,1\}^n \to \{0,1\}$ as follows: Writing $V = I^{\otimes(n-k)} \otimes \widetilde{V}$, let

$$\widetilde{g} = \arg \min_{h \in \{0,1\}^{2^k}} \operatorname{dist}\Big(\widetilde{V}, \operatorname{diag}((-1)^{h(x)})\Big) = \arg \min_{h \in \{0,1\}^{2^k}} \left(\min_{\theta} \|e^{i\theta}\widetilde{V} - \operatorname{diag}((-1)^{h(x)})\| \right) \tag{4.15}$$

and set $g := \widetilde{g}$ where we interpret $\widetilde{g} : \{0,1\}^n \to \{0,1\}$ as a k-junta. We claim that $\operatorname{dist}(U_f, V) \geq \operatorname{dist}(U_g, V)$. First, note that because U_f and U_g are both diagonal matrices, the off-diagonal contributions to

$$\|U_f - V\|^2 = \sum_{0 \leq i,j < N} |U_f[i,j] - V[i,j]|^2$$

is the same as that to $\|U_g - V\|^2$. Moreover, if we multiply the off-diagonal terms of V by a phase $e^{i\theta}$, their contribution to the sum will still be the same as U_f and U_g are zero on their off-diagonal entries. It therefore suffices to compare the diagonal terms of these two quantities. With this in mind, we define the following quantity: For $A, B \in \mathbb{C}^{2^n \times 2^n}$, let $\underline{\operatorname{dist}}(A,B)$ be the sum of diagonal contributions to the Frobenius norm of $A - B$, i.e. we have

$$\underline{\operatorname{dist}}(A,B)^2 := \frac{1}{2N} \min_{\theta} \sum_{0 \leq i < N} |e^{i\theta}A[i,i] - B[i,i]|^2.$$

We then have that

$$\begin{aligned}
\underline{\operatorname{dist}}(U_f, V)^2 &= \frac{1}{2N} \min_{\theta} \sum_{j=0}^{2^{n-k}-1} \sum_{l=0}^{2^k-1} \left| e^{i\theta} V[j \cdot 2^k + l, j \cdot 2^k + l] - U_f[j \cdot 2^k + l, j \cdot 2^k + l] \right|^2 \qquad (4.16)\\
&\geq \frac{1}{2N} \sum_{j=0}^{2^{n-k}-1} \min_{\theta_j} \sum_{l=0}^{2^k-1} \left| e^{i\theta_j} V[j \cdot 2^k + l, j \cdot 2^k + l] - U_f[j \cdot 2^k + l, j \cdot 2^k + l] \right|^2 \\
&\geq \frac{1}{2^{n-k}} \sum_{j=0}^{2^{n-k}-1} \underline{\operatorname{dist}}(U_{\widetilde{g}}, \widetilde{V})^2 \qquad (4.17)\\
&= \underline{\operatorname{dist}}(U_{\widetilde{g}}, \widetilde{V})^2 \\
&= \underline{\operatorname{dist}}(U_g, V)^2.
\end{aligned}$$

In particular, Equation (4.16) rewrites the sum by considering 2^{n-k} blocks of $2^k \times 2^k$ matrices on the diagonal of $e^{i\theta}V - U_f$, and Equation (4.17) follows from the choice of g as a minimizer in Equation (4.15). Because the off-diagonal contributions to the expressions for distance are the same for U_f and U_g, $\operatorname{dist}(U_f, V) \geq \operatorname{dist}(U_g, V)$, completing the proof. ∎

5 Learning Quantum k-Juntas

We present algorithm to learn quantum k-juntas in Section 5.1, and our lower bound for learning quantum k-juntas in Section 5.2.

5.1 Learning Upper Bound In this section, we present our algorithm for learning quantum k-juntas. Our algorithm can be viewed as analogous to the quantum algorithm of Atıcı and Servedio [AS07] for learning classical k-juntas; as such, we start be briefly recalling their high-level approach.

Given query access to a function $f : \{0,1\}^n \to \{0,1\}$, the algorithm of [AS07] first determines the set of all relevant variables of non-negligible influence via "Fourier sampling" from f.[8] It then learns the truth table of the

[8] Recall that Fourier sampling from $f : \{0,1\}^n \to \{0,1\}$ refers to drawing $S \subseteq [n]$ (identified with its 0/1 indicator vector with probability $|\widehat{f}(S)|^2$.

Input: Oracle access to quantum k-junta U, error parameter $\varepsilon > 0$

Output: Classical description of U (as a $2^n \times 2^n$ matrix)

QUANTUM-JUNTA-LEARNER(U, ε):

1. Let $S := \text{PAULI-SAMPLE}\Big(U, \frac{\varepsilon^2}{4k}, k\Big)$.
2. Set $t := O(\frac{4^k}{\varepsilon^2})$. Call QUANTUM-STATE-PREPARATION(U, S) $10t$ times to obtain at least t copies of $|\psi_S\rangle$.
3. Let $|\widehat{\psi}\rangle := \text{TOMOGRAPHY}\Big(|\psi_S\rangle \langle\psi_S|^{\otimes t}, \frac{\varepsilon^2}{4}\Big)$.
4. Return the unitary encoded by $|\widehat{\psi}\rangle$ tensored with $I^{\otimes(n-k)}$

Figure 6: Quantum k-Junta Learner

function f restricted to the at most k relevant variables by querying f on each of the 2^k possible input strings on the relevant variables. Given membership query access to a unitary U, our algorithm proceeds analogously by first learning a set S of relevant qubits with non-negligible influence via "Pauli sampling", a subroutine analogous to Fourier Sampling.[9] Then, we learn an approximation to the part of U that acts only on the subset S, the qubits with nonnegligible influence. We do this by reducing the problem to learning a *quantum state*, a task known as *quantum state tomography*.

The connection between learning the unknown unitary U and learning quantum states comes via the *Choi-Jamiołkowski isomorphism* (described in Section 2.5). In our learning algorithm we will use the following procedure to perform pure state tomography on (copies of) the CJ state $|v(U)\rangle$ in order to learn a description of U:

PROPOSITION 5.1. (PURE STATE TOMOGRAPHY) *There exists a procedure* TOMOGRAPHY [10] *that, given $O(d/\varepsilon)$ samples of an unknown d-dimensional pure state $|\psi\rangle$, outputs with high probability a classical description of a pure state $|\widehat{\psi}\rangle \in \mathbb{C}^d$ such that*

$$\Big|\Big\langle \psi \Big| \widehat{\psi} \Big\rangle\Big|^2 \geq 1 - \varepsilon \ .$$

Our quantum junta learning algorithm is presented in Figure 6 and its properties are established in Theorem 5.1.

THEOREM 5.1. *Given oracle access to a quantum k-junta $U \in \mathcal{U}_N$ and $\varepsilon > 0$,* QUANTUM-JUNTA-LEARNER(U, ε) *(cf. Figure 6) outputs, with probability $9/10$, a unitary $\widehat{U}$ such that $\text{dist}(U, \widehat{U}) \leq \varepsilon$. Furthermore,* QUANTUM-JUNTA-LEARNER *makes $O\Big(\frac{k}{\varepsilon} + \frac{4^k}{\varepsilon^2}\Big)$ queries to U.*

Proof. We will analyze the closeness guarantee and the query complexity separately, starting with the former.

Consider the state $|\psi_S\rangle$ obtained by running QUANTUM-STATE-PREPARATION in Step 2 of Figure 6. $|\psi_S\rangle$ is a pure state with $2k$ qubits; as such, it encodes k-qubit unitary matrix V acting on the qubits in the relevant set $R \subset [n], |R| = k$. We have that

$$\text{dist}\Big(U, V \otimes I^{\otimes(n-k)}\Big) \leq \frac{\varepsilon}{2} \tag{5.18}$$

by Lemma 5.1.

[9] This subroutine is also implicit in [Wan11].

[10] This procedure was first devised by Derka, Bužek, and Ekert [DBE98] (and whose sample complexity was determined by Bruß and Machiavello [BM99]).

Input: Oracle access to unitary U, $S \subseteq [n]$

Output: The quantum state $|\psi_S\rangle$ or "error"

Quantum-State-Preparation(U, S):

1. Prepare the state $|v(U)\rangle = \sum_{x \in \mathbb{Z}_4^n} \widehat{U}(x) |v(\sigma_x)\rangle$.
2. Measure qubits in $\bar{S} \subseteq [n]$ and $\{\tilde{l} : l \in \bar{S}\}$ in the Pauli basis $\{|v(\sigma_x)\rangle\}_{x \in \mathbb{Z}_4^{n-|S|}}$.
   (a) If the measurement result is $|v(I^{\otimes(n-|S|)})\rangle$, let $|\psi_S\rangle$ be the the unmeasured state on $2|S|$ qubits tensored with $(k - |S|)$ EPR pairs. Return $|\psi_S\rangle$.
   (b) Otherwise, return "error".

Figure 7: The Quantum State Preparation Subroutine (cf. Step 2 of Figure 6)

Input: Oracle access to quantum k-junta U on n qubits, threshold $\gamma > 0$

Output: $S \subseteq [n]$

Pauli-Sample(U, γ, k):

1. Initialize $S = \emptyset$.
2. Repeat the following $O\left(\frac{\log k}{\gamma}\right)$ times:
   (a) Prepare the $|v(U)\rangle$ and measure all qubits in the Pauli basis, $\{|v(\sigma_x)\rangle\}_{x \in \mathbb{Z}_4^n}$.
   (b) Given the measurement outcome $|\sigma_x\rangle$, set $S \leftarrow S \cup \mathrm{supp}(x)$
3. Return S.

Figure 8: The Pauli Sampling Subroutine (cf. Step 1 of Figure 6)

Let $\widehat{U}$ be the output of Algorithm 6, and let $\widehat{U} := W \otimes I^{\otimes(n-k)}$, for a k-qubit unitary W on qubits in the relevant set R. To show that $\widehat{U}$ is close to U, we will now show that with probability at least 99/100,

$$\operatorname{dist}(V, W) \leq \frac{\varepsilon}{2}. \tag{5.19}$$

It would then follow from the triangle inequality and Equations (5.18) and (5.19) that

$$\operatorname{dist}(\widehat{U}, U) \leq \varepsilon.$$

To show that V and W are close, consider the output of Tomography in Step 3 of Figure 6. By Proposition 5.1, we have that with $O\left(4^k/\varepsilon^2\right)$ copies of $|\psi_S\rangle$,

$$\left| \left\langle \psi_S \middle| \widehat{\psi} \right\rangle \right|^2 \geq 1 - \varepsilon^2/4$$

Note that $|\widehat{\psi}\rangle$ encodes W and that $|\psi_S\rangle$ encodes V. Writing $K := 2^k$, we have that

$$|\widehat{\psi}\rangle = \sum_{0 \leq i,j < K} \frac{W[i,j]}{\sqrt{K}} |i\rangle |j\rangle \qquad \text{and} \qquad |\psi_S\rangle = \sum_{0 \leq i,j < K} \frac{V[i,j]}{\sqrt{K}} |i\rangle |j\rangle .$$

We then have that

$$\begin{aligned}
\operatorname{dist}(V,W)^2 &= \min_\theta \frac{1}{2K} \|e^{i\theta} V - W\|^2 \\
&= \frac{1}{2} \min_\theta \sum_{0 \leq i,j < K} \left| \frac{e^{i\theta} V[i,j]}{\sqrt{K}} - \frac{W[i,j]}{\sqrt{K}} \right|^2 \\
&= \frac{1}{2} \min_\theta \sum_{0 \leq i,j < K} \left(\left| \frac{e^{i\theta} V[i,j]}{\sqrt{K}} \right|^2 + \left| \frac{W[i,j]}{\sqrt{K}} \right|^2 - 2 \cdot \mathfrak{Re}\left(\frac{e^{i\theta}}{K} V[i,j] W[i,j]^* \right) \right) \\
&= \frac{1}{2} \min_\theta \left(2 - 2 \cdot \mathfrak{Re}\left(\sum_{0 \leq i,j < K} \frac{e^{i\theta}}{K} V[i,j] W[i,j]^* \right) \right) \\
&= \frac{1}{2} \left(2 - 2 \left| \sum_{0 \leq i,j < K} \frac{1}{K} V[i,j] W[i,j]^* \right| \right) \\
&= 1 - |\langle \psi_S | \widehat{\psi} \rangle| \\
&\leq \varepsilon^2/4,
\end{aligned}$$

Finally, we turn to the query complexity of Figure 6. By Lemma 5.2, the query complexity of Step 1 of Figure 6 (Pauli-Sample) is

$$\frac{\log k}{\frac{\varepsilon^2}{4k}} = O\left(\frac{k \log k}{\varepsilon^2} \right).$$

The number of copies required for the tomography subroutine is

$$t := O\left(\frac{4^k}{\varepsilon^2} \right).$$

As Quantum-State-Preparation has a small probability of error ($O(\varepsilon^2)$), we can show by Markov's inequality that with $10t$ calls to Quantum-State-Preparation, we will obtain at least t copies of $|\psi_S\rangle$ with high probability. In more detail, let Y be the random variable indicating the number of failed executions of Quantum-State-Preparation. Then,

$$\mathbf{E}[Y] \leq 10t \cdot \frac{\varepsilon^2}{4} \qquad \text{and so} \qquad \mathbf{Pr}[Y > 9t] \leq \frac{5\varepsilon^2}{18} \ll 0.01.$$

Because each call to Quantum-State-Preparation makes one call to U, the total query complexity of Figure 6 is $O\left(\frac{k \log k}{\varepsilon^2} + \frac{4^k}{\varepsilon^2} \right)$, completing the proof. ∎

We now prove the following lemma that we used in the proof of Theorem 5.1.

LEMMA 5.1. *Let V denote the unitary whose Choi-Jamiolkowski isomorphism is given by $|\psi_S\rangle$, as obtained from the call to* QUANTUM-STATE-PREPARATION *in Step 2 of Figure 6. Then*

$$\operatorname{dist}\Big(U, V \otimes I^{\otimes(n-k)}\Big) \leq \frac{\varepsilon}{2}.$$

Proof. Since U is a k-junta, let $R \subset [n]$ be the set of k relevant variables. Let $S \subset R$ be the set of qubits with nonnegligible influence outputted by PAULI-SAMPLE in Step 1 of Figure 6.

Let $U = U_R \otimes I_{\bar{R}}$, where U_R is a k-qubit unitary acting only on the relevant qubits in R. It is sufficient to show that $\operatorname{dist}(U_R, V) \leq \frac{\varepsilon}{2}$. First, note that

$$|v(U)\rangle = |v(U_R)\rangle\,|v(I^{\otimes(n-k)})\rangle$$

Thus, when we measure qubits in $\bar{R}$ and $\{\tilde{\ell} : \ell \in \bar{R}\}$, we always obtain $|v(I^{\otimes(n-k)})\rangle$, as U acts trivially on qubits outside of R.

Now we will consider what happens when we measure qubits in $R-S$. We will use the following decomposition of $|v(U_R)\rangle$.

$$|v(U_R)\rangle = \sum_{x \in Z_4^k} \widehat{U_R}(x)\,|v(\sigma_x)\rangle \tag{5.20}$$

$$= \sum_{x:\operatorname{supp}(x)\cap\bar{S}=\emptyset} \widehat{U}(x)\,|v(\sigma_x)\rangle + \sum_{x:\operatorname{supp}(x)\cap(R-S)\neq\emptyset} \widehat{U}(x)\,|v(\sigma_x)\rangle \tag{5.21}$$

By Lemma 5.2, S will contain all the qubits with influence larger than $\frac{\varepsilon^2}{4k}$ with high probability. Further, each qubit in S has nonzero influence. This implies that with high probabilty,

$$\sum_{x:\operatorname{supp}(x)\cap(R-S)\neq\emptyset} |\widehat{U}(x)|^2 = \mathbf{Inf}_{\bar{S}}[U] \leq \sum_{i\in\bar{S}} \mathbf{Inf}_i[U] \leq k \cdot \frac{\varepsilon^2}{4k}$$

By the decomposition in Equation (5.21), measuring qubits in $(R-S) \cup \{\tilde{\ell} : \ell \in R-S\}$ yields the state $|v(I^{\otimes(|R|-|S|)})\rangle$ with probability at least $1 - \frac{\varepsilon^2}{4}$. Conditioned on this event, the $2k-$qubit post measurement state is as follows:

$$|\psi_S\rangle = \frac{1}{\sqrt{1 - \mathbf{Inf}_{R-S}[U]}} \sum_{x:\operatorname{supp}(x)\cap\bar{S}=\emptyset} \widehat{U}(x)\,|v(\sigma_x)\rangle \otimes |EPR\rangle^{\otimes(k-|S|)}$$

Let $\alpha := \frac{1}{\sqrt{1-\mathbf{Inf}_{R-S}[U]}}$. Note that $1 \leq \alpha \leq \frac{1}{\sqrt{1-\frac{\varepsilon^2}{4}}}$. Then,

$$\begin{aligned}
2\,\operatorname{dist}^2(V, U_R) &= ||\,|\psi_S\rangle - |v(U_R)\rangle\,||^2 \\
&= ||\sum_{x:\operatorname{supp}(x)\cap\bar{S}=\emptyset} \widehat{U}(x)\,|v(\sigma_x)\rangle + \sum_{x:\operatorname{supp}(x)\cap(R-S)\neq\emptyset} \widehat{U}(x)\,|v(\sigma_x)\rangle \\
&\quad - \alpha \sum_{x:\operatorname{supp}(x)\cap\bar{S}=\emptyset} \widehat{U}(x)\,|v(\sigma_x)\rangle\,||^2 \\
&= (\alpha-1)^2 \sum_{x:\operatorname{supp}(x)\cap\bar{S}=\emptyset} |\widehat{U}(x)|^2 + \sum_{x:\operatorname{supp}(x)\cap(R-S)\neq\emptyset} |\widehat{U}(x)|^2 \\
&\leq \Big(\frac{1}{\sqrt{1-\frac{\varepsilon^2}{4}}} - 1\Big)^2 + \mathbf{Inf}_{R-S}[U] \\
&\leq 2\frac{\varepsilon^2}{4}
\end{aligned}$$

This shows that $\mathrm{dist}(V \otimes I^{\otimes(n-k)}, U) = \mathrm{dist}(V, U_R) \leq \varepsilon/2$. ∎

The following lemma is analogous to Lemma IV.4 in [AS07].

LEMMA 5.2. *Let $U \in \mathcal{U}_N$ be a unitary acting non-trivially on qubits in $R \subset [n]$. Then* PAULI-SAMPLE$(U, \varepsilon, |R|)$ *makes $t = O\left(\frac{\log |R|}{\varepsilon}\right)$ membership queries to U and outputs with high probability a list $S \subset [n]$ that satisfies the following properties:*

1. *S contains all qubits $i \in [n]$ such that $\mathbf{Inf}_i[U] \geq \varepsilon$; and*
2. *All qubits i in S have nonzero influence, i.e. $\mathbf{Inf}_i[U] > 0$.*

Proof. If $\mathbf{Inf}_i[U] \geq \varepsilon$, then the probability i does not occur in S is at most $(1-\varepsilon)^t \leq \frac{1}{100|R|}$. By the union bound, S will contain every i such that $\mathbf{Inf}_i[U] \geq \varepsilon$ with probability at least 99/100. The second item follows from the fact that if $i \in [n]$ is Pauli-sampled, there must exist $x \in \mathbb{Z}_4^n, i \in \mathrm{supp}(x)$ such that $\widehat{U}(x) \neq 0$. ∎

5.2 Learning Lower Bound Finally, we present a nearly-matching lower bound for the query complexity of learning quantum juntas. Although it is commonly stated that process tomography requires $\Omega(4^n)$ queries, we have not been able to identify in the literature a formal lower bound proof. Thus, we provide the following proof for completeness.

THEOREM 5.2. *Any algorithm for learning quantum k-juntas with error ε requires $\Omega(4^k \log(1/\varepsilon)/k)$ queries.*

Proof. We prove this lower bound via a communication complexity argument. In particular, we reduce the INPUT GUESSING game to learning quantum juntas. The INPUT GUESSING game with domain size K is a two-party communication task where one party (named Alice) receives an uniformly random input x from $\{1, 2, \ldots, K\}$ and the other party (named Bob) has to output a guess for x after engaging in two-way communicating with Alice. We consider the model of *quantum communication*, where Alice and Bob can exchange qubits with each other. A classic result of Nayak [Nay99] implies the following lower bound on the communication complexity of the INPUT GUESSING game.

THEOREM 5.3. (LOWER BOUND FOR INPUT GUESSING GAME [NAY99]) *Any quantum communication protocol that solves the* INPUT GUESSING *game with domain size K and success probability p requires exchanging $\log K - \log \frac{1}{p}$ qubits between the parties.*

Let $\mathcal{A}$ denote an algorithm that learns quantum k-juntas, assuming it is told which k of the n qubits are relevant. Note that the problem of learning quantum k-juntas without this additional information is at least as hard. Without loss of generality, assume the first k qubits are relevant.

Suppose $\mathcal{A}$ makes q queries, achieves error ε and achieves constant success probability. Then we construct a quantum communication protocol for the INPUT GUESSING game with domain size $K = \Omega\left((1/\varepsilon)^{4^k}\right)$, communication complexity $O(kq)$, and constant success probability. By Theorem 5.3, this implies that

$$kq \geq \Omega(\log K) = \Omega(4^k \log(1/\varepsilon))$$

which implies the desired lower bound.

Let K denote the size of a maximal ε-packing of the space of k-qubit unitary matrices, with respect to the distance measure $\mathrm{dist}(\cdot, \cdot)$. In other words, this is the maximal number of disjoint ε-balls in the space of k-qubit unitaries. By standard volume arguments (see [Sza97]), since $\mathrm{dist}(\cdot, \cdot)$ is a unitarily invariant distance measure, K is at least $\Omega\left((1/\varepsilon)^{4^k}\right)$. Let $\{U_1, \ldots, U_K\}$ denote an enumeration of the maximal ε-packing.

Suppose Alice gets a random input $x \in \{1, 2, \ldots, K\}$. Bob will simulate the algorithm $\mathcal{A}$. Whenever $\mathcal{A}$ has to make a query to the oracle, Bob sends the first k qubits of his query register to Alice, then Alice applies the k-qubit unitary U_x to the register, and then sends the register back. Indeed, if $\mathcal{A}$'s query register is on n qubits total, then the effective n-qubit unitary applied in this simulated execution of $\mathcal{A}$ is $U = U_x \otimes I$ where U_x acts on the first k qubits and I acts on the remaining $n - k$ qubits. Bob continues in this fashion until the

algorithm $\mathcal{A}$ terminates and outputs (with constant probability) a classical description of a unitary V such that $\text{dist}(U_x \otimes I, V) \leq \varepsilon$ where, by the correctness of the algorithm $\mathcal{A}$, V is a quantum k-junta $V' \otimes I$ that acts trivially on all qubits except the first k. Thus we have that $\text{dist}(U_x, V') \leq \varepsilon$, and by definition of an ε-packing, U_x is the unique member of the packing that has distance ε to V'. Thus Bob can uniquely identify Alice's input x with constant probability. The total communication complexity of this protocol is $2kq$. ∎

Acknowledgements

We would like to thank Rocco A. Servedio and Xi Chen for helpful discussions. We would also like to thank Zongbo Bao for pointing out an error in Figure 1 in an earlier version of this paper, as well as Vishnu Iyer and Michael Whitmeyer for helpful comments.

References

[ABRdW16] Andris Ambainis, Aleksandrs Belovs, Oded Regev, and Ronald de Wolf. Efficient quantum algorithms for (gapped) group testing and junta testing. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 903–922. SIAM, 2016.

[AS07] Alp Atıcı and Rocco A. Servedio. Quantum algorithms for learning and testing juntas. *Quantum Inf. Process.*, 6(5):323–348, 2007.

[Bel19] Aleksandrs Belovs. Quantum algorithm for distribution-free junta testing. In René van Bevern and Gregory Kucherov, editors, *Computer Science - Theory and Applications - 14th International Computer Science Symposium in Russia, CSR 2019, Novosibirsk, Russia, July 1-5, 2019, Proceedings*, volume 11532 of *Lecture Notes in Computer Science*, pages 50–59. Springer, 2019.

[BKT17] Mark Bun, Robin Kothari, and Justin Thaler. The polynomial method strikes back: Tight quantum query bounds via dual polynomials. *CoRR*, abs/1710.09079, 2017.

[Bla09] Eric Blais. Testing juntas nearly optimally. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 151–158. ACM, 2009.

[BM99] Dagmar Bruß and Chiara Macchiavello. Optimal state estimation for d-dimensional quantum systems. *Physics Letters A*, 253(5-6):249–251, 1999.

[Bsh19] Nader H. Bshouty. Almost optimal distribution-free junta testing. In Amir Shpilka, editor, *34th Computational Complexity Conference, CCC 2019, July 18-20, 2019, New Brunswick, NJ, USA*, volume 137 of *LIPIcs*, pages 2:1–2:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[BY22] Arnab Bhattacharyya and Yuichi Yoshida. *Property Testing - Problems and Techniques*. Springer, 2022.

[CG04] Hana Chockler and Dan Gutfreund. A lower bound for testing juntas. *Inf. Process. Lett.*, 90(6):301–305, 2004.

[Cho75] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[CN97] 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.

[DBE98] Radoslav Derka, Vladimir Bužek, and Artur K Ekert. Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement. *Physical Review Letters*, 80(8):1571, 1998.

[FKR+04] Eldar Fischer, Guy Kindler, Dana Ron, Shmuel Safra, and Alex Samorodnitsky. Testing juntas. *J. Comput. Syst. Sci.*, 68(4):753–787, 2004.

[GJ14] Gus Gutoski and Nathaniel Johnston. Process tomography for unitary quantum channels. *Journal of Mathematical Physics*, 55(3):032201, 2014.

[Gol10] Oded Goldreich. A brief introduction to property testing. In Oded Goldreich, editor, *Property Testing - Current Research and Surveys*, volume 6390 of *Lecture Notes in Computer Science*, pages 1–5. Springer, 2010.

[ITW21] Vishnu Iyer, Avishay Tal, and Michael Whitmeyer. Junta distance approximation with sub-exponential queries. In Valentine Kabanets, editor, *36th Computational Complexity Conference, CCC 2021, July 20-23, 2021, Toronto, Ontario, Canada (Virtual Conference)*, volume 200 of *LIPIcs*, pages 24:1–24:38. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[Jam72] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[MdW16] Ashley Montanaro and Ronald de Wolf. A survey of quantum property testing. *Theory Comput.*, 7:1–81, 2016.

[MO10] Ashley Montanaro and Tobias Osborne. Quantum boolean functions. *Chic. J. Theor. Comput. Sci.*, 2010, 2010.

[Nay99] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. In *40th Annual Symposium on Foundations of Computer Science*, pages 369–376. IEEE, 1999.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.
[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.
[PRS02] Michal Parnas, Dana Ron, and Alex Samorodnitsky. Testing basic boolean formulae. *SIAM J. Discret. Math.*, 16(1):20–46, 2002.
[PRW19] Ramesh Krishnan S. Pallavoor, Sofya Raskhodnikova, and Erik Waingarten. Approximating the distance to monotonicity of boolean functions. volume abs/1911.06924, 2019.
[Sağ18] Mert Sağlam. Near log-convexity of measured heat in (discrete) time and consequences. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 967–978, 2018.
[Sza97] Stanislaw J Szarek. Metric entropy of homogeneous spaces. *arXiv preprint math/9701213*, 1997.
[Val15] Gregory Valiant. Finding correlations in subquadratic time, with applications to learning parities and the closest pair problem. *J. ACM*, 62(2):13:1–13:45, May 2015.
[Wan11] Guoming Wang. Property testing of unitary operators. *Physical review. A, Atomic, molecular, and optical physics*, 84(5), November 2011.
[Wat18] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.
[Wil17] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2017.
[Zha19] Xiaojin Zhang. Near-optimal algorithm for distribution-free junta testing. *ArXiv*, `abs/1911.10833`, 2019.

A An Alternative Characterization of Influence

Below is an alternative characterization of the influence of a set of variables on a unitary. Our learning and testing algorithms do not make use of this characterization, but it may be of independent interest.

LEMMA A.1. *(Equivalent characterization of influence) Given $U \in \mathcal{U}_N$ and $j \in [n]$, we have*

$$\mathbf{Inf}_j[U] = 1 - \frac{1}{2^{n+1}}\mathrm{Tr}\big((\mathrm{Tr}_j U^\dagger)(\mathrm{Tr}_j U)\big).$$

More generally, for $S \subseteq [n]$, we have

$$\mathbf{Inf}_S[U] = 1 - \frac{1}{2^{n+|S|}}\mathrm{Tr}\big((\mathrm{Tr}_S U^\dagger)(\mathrm{Tr}_S U)\big).$$

Proof. Note that for $S \subseteq [n]$ and $x \in \mathbb{Z}_4^n$, we have that

$$\mathrm{Tr}_S(\sigma_x) = 0 \qquad \text{if and only if} \qquad S \cap \mathrm{supp}(x) \neq \emptyset.$$

This is immediate from the fact that the only Pauli matrix with non-zero trace is $\sigma_0 = I$. Writing U in the Pauli basis, $\mathrm{Tr}_S(U)$ has the following form

$$\mathrm{Tr}_S(U) = \mathrm{Tr}_S\left(\sum_{x\in\mathbb{Z}_4^n} \widehat{U}(x)\sigma_x\right) = 2^{|S|}\sum_{x:\mathrm{supp}(x)\cap S=\emptyset} \widehat{U}(x)\sigma_x. \tag{A.1}$$

Using this characterization, it follows that

$$\begin{aligned}\frac{1}{2^{n-|S|}}\mathrm{Tr}\big(\mathrm{Tr}_S(U)^\dagger \mathrm{Tr}_S(U)\big) &= \frac{1}{2^{n-|S|}}\langle \mathrm{Tr}_S(U), \mathrm{Tr}_S(U)\rangle \\ &= 2^{2|S|}\sum_{x:\mathrm{supp}(x)\cap S=\emptyset} |\widehat{U}(x)|^2 \\ &= 2^{2|S|}(1 - \mathbf{Inf}_S[U])\end{aligned}$$

where the second equality follows from Parseval's formula and equation A.1, while the final equality is because $\|U\|^2 = N$ for all unitaries $U \in \mathcal{U}_N$. The lemma follows by rearranging the final expression above. ∎