Measuring Quantum Entropy

Jayadev Acharya*, Ibrahim Issa†, Nirmal V. Shende*, and Aaron B. Wagner*
*School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853
Email: {acharya, nvs25, wagner}@cornell.edu
†American University of Beirut, Beirut, Lebanon. Email: ii19@aub.edu.lb

***Abstract*—The entropy of a quantum system is a measure of its randomness and is useful in quantifying entanglement. We study the problem of measuring the von Neumann and Rényi entropies of an unknown mixed quantum state given access to independent copies of the state. For Rényi entropy of integral order exceeding one, we determine the order-optimal copy complexity and show that it is strictly lower than the number of copies required to learn the underlying state. The main technical innovation is a concentration result for certain polynomials that arise in the Kerov algebra of Young diagrams, which is proven using the cycle structure of compositions of certain types of permutations. For von Neumann entropy and Rényi entropy of non-integral orders, we provide upper and lower bounds on the sample complexity of the Empirical Young Diagram (EYD) algorithm, which is the analogue of the empirical plug-in estimator in classical estimation.**

I. Introduction and Results

We consider how to estimate the mixedness or noisiness of a quantum state given independent copies of the state. Mixed quantum states can arise in practice in various ways: classical stochasticity can be intentionally introduced when the state is originally prepared; pure states can become mixed by a quantum measurement; and the states of the subsystems of bipartite states can be mixed even when the overall bipartite state is pure, which forms the basis for purification.

In the third case, the level of mixedness of the subsystems indicates the level of entanglement in the pure, bipartite system. The possibility of entanglement of two separated systems is arguably the most curious, and the most powerful, way in which quantum systems differ from classical ones. Indeed, entanglement has been fruitfully exploited as a resource in a number of quantum information processing protocols [1]–[5]. The subsystems of a pure bipartite state are pure if and only if the bipartite state itself is unentangled, and likewise they are maximally mixed if and only if the bipartite state is maximally entangled. Thus the mixedness of the subsystems' states can be used as a measure of entanglement of the bipartite system.

Mixedness can be measured in multiple ways. We shall use the von Neumann and (the family of) Rényi entropies, which correspond to the classical Shannon and (the family of) Rényi entropies of the eigenvalues of the density operator of the state, respectively. A density matrix (or operator) ρ is a complex positive semidefinite matrix with unit trace; thus its eigenvalues are nonnegative and sum to one. The von Neumann entropy of a density matrix ρ is

$$S(\rho) \stackrel{\text{def}}{=} -\mathrm{tr}(\rho \log \rho).$$

For $\alpha > 0, \alpha \neq 1$, the Rényi entropy of order α of ρ is

$$S_\alpha(\rho) \stackrel{\text{def}}{=} \frac{1}{1-\alpha} \log \mathrm{tr}(\rho^\alpha).$$

Quantum entropy can be justified operationally as a measure of compressibility [6]–[8], and as noted earlier, entanglement [9].

In principle, both the von Neumann and Rényi entropies for a quantum state ρ can be computed if the state is known. We consider how to estimate these quantities for an unknown state given independent copies of the state, to which arbitrary quantum measurements followed by arbitrary classical computation can be applied. This problem arises when characterizing a completely unknown system and when one seeks to experimentally verify that a system is behaving as desired. Since generating independent copies of a state can be quite costly in the quantum setting [10], [11], it is desirable to minimize the number of independent copies of the state that are required to estimate the von Neumann and Rényi entropies to a desired precision and confidence. We thus adopt this *copy complexity* as our figure-of-merit.

Using standard results in quantum state estimation, we reduce our problem to one that is fully classical. We first describe this classical problem, which is potentially of interest in its own right.[1]

A. Quantum-Free Formulation

Let p be a distribution over $[d] \stackrel{\text{def}}{=} \{1, \ldots, d\}$. A property $f(p)$ is a mapping of distributions to real numbers. A property f is said to be symmetric (or label-invariant) if it is a function of only the multiset of probability values, and not the ordering.

Classical symmetric property estimation. We are given independent samples $X^n \stackrel{\text{def}}{=} X_1, \ldots, X_n$ from an unknown distribution p, and the goal is to estimate a symmetric property $f(p)$ up to an additive ε error, with probability at least 2/3.

Quantum state property estimation. The problem of estimating von Neumann and Rényi entropies of a quantum state with eigenvalues $\boldsymbol{\eta}$ can be shown to be equivalent to estimating a symmetric property $f(\boldsymbol{\eta})$. However, instead of being given independent samples $X_1, \ldots, X_n$ from the distribution $\boldsymbol{\eta}$ as in the classical case, we are given access to a function $\boldsymbol{\lambda}(X^n) = \boldsymbol{\lambda}_1 \geq \boldsymbol{\lambda}_2 \geq \ldots$ of X^n. Here $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots$ are integers satisfying the following property.

- For any $k \geq 1$, $\sum_{i=1}^{k} \boldsymbol{\lambda}_i$ is equal to the largest possible sum of the lengths of k disjoint non-decreasing subsequences of X^n.

[1]The full version of this paper is available online [12].

Equivalently, we may view the observations as the output of the Robinson–Schensted–Knuth (RSK) algorithm applied to the sequence X^n, instead of being X^n itself. The reader is referred to [13] for more details on the procedure. The copy complexity of estimating quantum entropy turns out to be equivalent to the problem of estimating classical entropy when given access to $\boldsymbol{\lambda}(X^n)$. A simple data processing implication of the form $\boldsymbol{\eta} \to X^n \to \boldsymbol{\lambda}(X^n)$ shows that the complexity of estimating quantum state property is at least as hard as estimating the same property in the classical setting.

B. Our Results

We consider the following framework.

> $\Pi(f, d, \varepsilon)$: Given a property f, and access to independent copies of a d-dimensional mixed state ρ (e.g. output of some quantum experiment), how many copies are needed to estimate $f(\rho)$ to within $\pm\varepsilon$?[2]

The copy complexity, denoted by $C(f, d, \varepsilon)$, is the minimum number of copies required for an algorithm that solves $\Pi(f, d, \varepsilon)$.

We study the copy complexity of estimating the entropy of a mixed state of dimension d. We will use the standard asymptotic notations, and are interested in characterizing the dependence of $C(S, d, \varepsilon)$, and $C(S_\alpha, d, \varepsilon)$, as a function of d and ε. We assume the parameter α to be a constant, and focus on only the growth rate as a function of d and ε.

We will now discuss our results, which are summarized in Table I and Table II. For comparison purposes, it is useful to recall the copy complexity of quantum tomography, in which the goal is to learn the entire density matrix ρ. This problem has been studied in various works using various distance measures; and up to poly-logarithmic factors, for the standard distance measures, the copy complexity depends quadratically on the dimension d. Namely, it is $\tilde{O}(d^2)$. Similar to the sample complexity of estimating Rényi entropies of classical distributions from samples, our bounds are also dependent on whether α is less than one, and whether it is an integer.

1) Rényi Entropy, Integral $\alpha > 1$: We obtain our most optimistic and conclusive results in this case.

Theorem 1. *For $\alpha \in \mathbb{N}\backslash\{1\}$,*

$$C(S_\alpha, d, \varepsilon) = \Theta\left(\max\left\{\frac{d^{1-1/\alpha}}{\varepsilon^2}, \frac{d^{2-2/\alpha}}{\varepsilon^{2/\alpha}}\right\}\right),$$

where the hidden constants depend only on α.

We note that the lower bounds here hold for *all estimators*, not just for the estimators used in the upper bound. Furthermore, these bounds are sub-quadratic in d, namely we can estimate the Rényi entropy of integral orders even before we have enough copies to perform full tomography. The upper bounds are established by analyzing certain polynomials from representation theory that are related to the central characters of the symmetric group. The main contribution is to analyze the variance of these estimators, for which we draw upon various results from Kerov's algebra. For the lower bound, we design the spectrums of two mixed states such that their Rényi entropy differ by at least ε, but such that they require a large copy complexity to distinguish between them. For this we use various properties of Schur polynomials and other properties of integer partitions [14], [15].

Remark 1. The first term in the complexity expression in Theorem 1 dominates when $\varepsilon < 1/\sqrt{d}$, and is identical to the sample complexity of estimating Rényi entropy in the classical setting.

For estimating $\mathcal{S}_\alpha(\rho)$ for $\alpha \leq 1$ and non-integral $\alpha > 1$, we analyze the Empirical Young Diagram (EYD) algorithm [16], [17] . The EYD algorithm is similar to using a plug-in estimate of the empirical distribution to estimate properties in classical distribution property estimation.

2) Rényi Entropy, $\alpha < 1$: We show that $C(S_\alpha, d, \varepsilon) = O(d^{2/\alpha}/\varepsilon^{2/\alpha})$. Since $\alpha < 1$, this growth is more than quadratic, namely the EYD algorithm requires more copies than is required for tomography. We complement this by showing that in fact the EYD algorithm requires $\Omega(d^{1+1/\alpha}/\varepsilon^{1/\alpha})$ copies, showing that the super-quadratic dependence on d is inherent to the EYD algorithm.

Theorem 2. *The empirical estimator of $\mathcal{S}_\alpha(\rho)$ outputs a $\pm\varepsilon$ estimate with $O\Big((d/\varepsilon)^{2/\alpha}\Big)$ copies. Moreover, the EYD algorithm requires at least $\Omega(d^{1+1/\alpha}/\varepsilon^{1/\alpha})$ copies to estimate $\mathcal{S}_\alpha(\rho)$ to $\pm\varepsilon$.*

In comparison, in the classical setting the tight exponent of d in the sample complexity for $\alpha < 1$ is $1/\alpha$.

3) von Neumann entropy ($\alpha = 1$): Again using the EYD algorithm, we show that $C(S, d, \varepsilon) = O(d^2/\varepsilon^2)$. We formulate an optimization problem whose solutions are an upper bound on the bias of the empirical estimate, and we bound the variance by proving that the estimator has a small bounded difference constant.

Theorem 3. *For von Neumann entropy ($\alpha = 1$), using the empirical entropy estimate:*

$$C(S, d, \varepsilon) = O\left(\frac{d^2}{\varepsilon^2} + \frac{\log^2(1/\varepsilon)}{\varepsilon^2}\right).$$

Moreover, there is a constant ε_0, such that for $\varepsilon < \varepsilon_0$, the empirical estimate of entropy requires at least $\Omega(d^2/\varepsilon)$ samples to estimate von Neumann entropy.

This complexity is still similar to that of full quantum tomography.

4) Rényi Entropy, Non integral $\alpha > 1$: Again using the EYD algorithm, in Theorem 4, we show that $C(S_\alpha, d, \varepsilon) = O(d^2/\varepsilon^2)$. We also provide a lower bound of $\Omega(d^2/\varepsilon)$ for the EYD estimator:

Theorem 4. *For $\alpha > 1$, the empirical estimator of $\mathcal{S}_\alpha(\rho)$ outputs a $\pm\varepsilon$ estimate with $O\Big(\frac{d^2}{\varepsilon^2}\Big)$ copies of ρ with probability*

[2]We seek success with probability at least $2/3$, which can be boosted to $1-\delta$ by repeating the algorithm $O(\log(1/\delta))$ times and taking the median.

Upper Bound	Lower Bound
$O\left(\max\left\{\frac{d^{2-\frac{2}{\alpha}}}{\varepsilon^{\frac{2}{\alpha}}}, \frac{d^{1-\frac{1}{\alpha}}}{\varepsilon^2}\right\}\right)$	$\Omega\left(\max\left\{\frac{d^{2-\frac{2}{\alpha}}}{\varepsilon^{\frac{2}{\alpha}}}, \frac{d^{1-\frac{1}{\alpha}}}{\varepsilon^2}\right\}\right)$

TABLE I
COPY COMPLEXITY OF $\mathcal{S}_\alpha(\rho)$ FOR INTEGRAL $\alpha > 1$.

α	Upper Bound	Lower Bound
$\alpha > 1$	$O(d^2/\varepsilon^2)$	$\Omega(d^2/\varepsilon)$
$\alpha < 1$	$O(d^{2/\alpha}/\varepsilon^{2/\alpha})$	$\Omega(d^{1+1/\alpha}/\varepsilon^{1/\alpha})$
$\alpha = 1$	$O(d^2/\varepsilon^2)$	$\Omega(d^2/\varepsilon)$

TABLE II
COPY COMPLEXITY OF EMPIRICAL ESTIMATORS.

at least 2/3. Also, there is a constant ε_0, such that for $\varepsilon < \varepsilon_0$, the empirical estimate of entropy requires $\Omega(d^2/\varepsilon)$ samples to estimate any Rényi entropy of order greater than one.

In addition to these results, we improve the error probability of the lower bounds on the convergence of EYD algorithm to the true spectrum. In particular, for the uniform distribution [18] shows that unless the number of copies is at least $\Omega(d^2/\varepsilon^2)$, the EYD has a total variation distance of at least ε with probability at least 0.01. We show that in fact unless the number of copies is at least $\Omega(d^2/\varepsilon^2)$ the trace distance is at least ε with probability at least $1 - \exp(c \cdot d)$ for some constant c.

C. Related Work

Our work is related to symmetric distribution property estimation in the classical setting, property estimation of classical distributions using quantum queries, and the property estimation of quantum states (as in the set-up of this paper). We briefly mention some closely related works. The survey by Montanaro and de Wolf [19] and the thesis by Wright [20] are recommended for a fuller account of the literature.

The copy complexity of quantum tomography (where the goal is to learn the entire density matrix ρ) is quadratic in d, and the complexity for tomography in various distance measures have been studied in [21]–[23].

Testing whether ρ has a particular unitarily invariant property of interest was studied in [18] for a number of properties. They show that for testing whether ρ is maximally mixed, namely whether all elements of $\boldsymbol{\eta}$ are $1/d$, requires $\Theta(d/\varepsilon^2)$ copies. They also studied the problem of testing the rank of ρ, and also provide bounds on the performance of the EYD algorithm for estimating the spectrum. Recently, [24] obtained tight bounds on the copy complexity of testing whether an unknown density matrix is equal to a known density matrix. The optimal measurement schemes for some of these problems can be quite involved. Testing properties under simpler *local measurements* was studied recently in [25].

In a personal communication, Bavarian, Mehraban, and Wright [26] claim an algorithm with copy complexity $O(d^2/\varepsilon)$ for the von Neumann entropy estimation, which is an ε factor improvement over our bound.

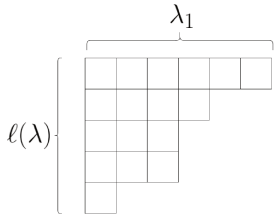

Fig. 1. English Young diagram for the partition $\boldsymbol{\lambda} = (6, 4, 3, 3, 1)$.

Testing and estimating distribution properties using quantum queries has been considered by various authors. Problems of testing properties such as uniformity, identity, closeness under the regular quantum query model, and conditional quantum query models have been studied in [27]–[29]. Recently Li and Wu [30] studied the quantum query complexity of estimating entropy of discrete distributions.

II. PRELIMINARIES

A. Unitarily Invariant Properties

A property f is a mapping from the set of density operators to real numbers. Let $U(d)$ be the set of all $d \times d$ unitary matrices.

Definition 1. A property $f(\rho)$ is called *unitarily invariant* if $f(U\rho U^\dagger) = f(\rho)$ for all $U \in U(d)$.

Let $\boldsymbol{\eta} = \{\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_d\}$ be the multiset the eigenvalues (also called the *spectrum*) of ρ. Two density matrices ρ, and σ have the same spectrum if and only if there is a unitary matrix U such that $\sigma = U\rho U^\dagger$. Therefore, unitarily invariant properties are functions of only the spectrum of the density matrix.

B. Schur Polynomials and Power-Sum Polynomials

A *partition* $\boldsymbol{\lambda}$ of n is a collection of non-negative integers $\boldsymbol{\lambda}_1 \geq \boldsymbol{\lambda}_2 \geq \ldots$ that sum to n. We write $\boldsymbol{\lambda} \vdash n$ and we write Λ_n for the set of all partitions of n. We denote the number of positive integers in $\boldsymbol{\lambda}$ by $\ell(\boldsymbol{\lambda})$, which we call its *length*. A partition $\boldsymbol{\lambda}$ can be depicted with an English *Young diagram*, which consists of a row of $\boldsymbol{\lambda}_1$ boxes above a row of $\boldsymbol{\lambda}_2$ boxes, etc., as shown in Fig. 1. The partition associated with a Young diagram is called its *shape*. A *Young tableau* over alphabet $[d]$ is a Young diagram in which each box has been filled with an element of $[d]$. A Young tableau is *semistandard* if it is strictly increasing top-to-bottom down each column and nondecreasing left-to-right across each row. Given $\boldsymbol{\lambda} \vdash n$ and d, the *Schur polynomial* is the polynomial in the variables $x_1, x_2, \ldots, x_d$ defined by

$$s_{\boldsymbol{\lambda}}(x) = \sum_T \prod_{i=1}^{d} x_i^{\#(T,i)}, \tag{1}$$

where the sum is over the set of semistandard Young Tableaus over alphabet $[d]$ corresponding to the partition $\boldsymbol{\lambda}$ and $\#(T, i)$ is the number of times i appears in T. Schur polynomials turn

out to be symmetric in $x_1, \ldots, x_d$. We shall also consider polynomials obtained from power sums. Given $\alpha \in \mathbb{R}_{\geq 0}$ and a distribution $\boldsymbol{\eta}$ on $[d]$, define $M_\alpha(\boldsymbol{\eta}) \stackrel{\text{def}}{=} \sum_{i=1}^d \boldsymbol{\eta}_i^\alpha$. Given $\boldsymbol{\lambda} \vdash r$, we define the *power sum* polynomial by $M_{\boldsymbol{\lambda}}(\boldsymbol{\eta}) \stackrel{\text{def}}{=} \prod_{i=1}^{\ell(\boldsymbol{\lambda})} M_{\lambda_i}(\boldsymbol{\eta})$. We remark that when $\boldsymbol{\eta}$ is the distribution of the eigenvalues of ρ, obtaining additive estimates of $\mathcal{S}_\alpha(\rho)$ is equivalent to obtaining multiplicative estimates for $M_\alpha(\boldsymbol{\eta})$.

Schur polynomials and power-sum polynomials are related through a change of basis. There exists a function $\chi_\cdot(\cdot) : \Lambda_n^2 \mapsto \mathbb{R}$ such that [31, Theorem 7.17.3]

$$M_{\boldsymbol{\mu}}(\cdot) = \sum_{\boldsymbol{\lambda}} \chi_{\boldsymbol{\lambda}}(\boldsymbol{\mu}) s_{\boldsymbol{\lambda}}(\cdot). \tag{2}$$

The quantity $\chi_{\boldsymbol{\lambda}}(\boldsymbol{\mu})$ is difficult to compute in general [32], although we shall only be interested in particular $\boldsymbol{\mu}$, as follows. Let $\dim(\boldsymbol{\lambda})$ denote the number of standard Young tableaux over alphabet $[n]$ with shape $\boldsymbol{\lambda}$. For $\boldsymbol{\lambda} \vdash n$ and $\boldsymbol{\mu} \vdash r$ define

$$p^{\#}_{\boldsymbol{\mu}}(\boldsymbol{\lambda}) \stackrel{\text{def}}{=} \begin{cases} n^{\underline{r}} \cdot \frac{\chi_{\boldsymbol{\lambda}}(\boldsymbol{\mu} \cup 1^{n-r})}{\dim(\boldsymbol{\lambda})} & \text{if } n \geq r, \\ 0 & \text{otherwise.} \end{cases}$$

where $n^{\underline{r}}$ is the *falling power*, i.e., $n^{\underline{r}} = n \cdot (n-1) \cdot (n-2) \cdot (n-(r-1))$ and $\boldsymbol{\mu} \cup 1^{n-r}$ denotes the partition of $[n]$ consisting of $\boldsymbol{\mu}$ followed by $n-r$ ones. These polynomials are very useful since they will give unbiased estimates of $M_\alpha(\boldsymbol{\eta})$, and can be used to estimate $\mathcal{S}_\alpha(\rho)$.

1) Weak Schur Sampling (WSS): Weak Schur Sampling is a measurement that takes n independent copies of a mixed state ρ (denoted $\rho^{\otimes n}$), and outputs a $\boldsymbol{\lambda} \vdash n$ (see [19, Section 4.2.2], [20, Chapter 3]). The output distribution over partitions is called the Schur-Weyl distribution, denoted $SW_{\boldsymbol{\eta}}$, and the probability of $\boldsymbol{\lambda} \vdash n$ is given by

$$SW_{\boldsymbol{\eta}}(\boldsymbol{\lambda}) = \dim(\boldsymbol{\lambda}) \cdot s_{\boldsymbol{\lambda}}(\boldsymbol{\eta}). \tag{3}$$

We are interested in Weak Schur Sampling due to the following powerful result [33]–[36] (See [19, Section 4.2.2] for details).

Lemma 1. *Weak Schur sampling is optimal for estimating unitarily invariant properties.*

The $p^{\#}_{\boldsymbol{\mu}}(\boldsymbol{\lambda})$ polynomial defined in the last section is useful to us due to the following lemma, which states that the (normalized) polynomial $p^{\#}_{(r)}(\boldsymbol{\lambda})$ is an unbiased estimator of the rth moment of $\boldsymbol{\eta}$. The lemma follows from the definitions and results already mentioned, and is implicit in [37], and explicit in [20, Proposition 3.8.3].

Lemma 2. *Fix a distribution, $\boldsymbol{\eta}$, and a natural number, r. If $\boldsymbol{\lambda}$ is randomly generated according to the distribution in (3) then* $\mathbb{E}\left[p^{\#}_{(r)}(\boldsymbol{\lambda})\right] = n^{\underline{r}} M_r(\boldsymbol{\eta})$.

The Empirical Young Diagram (EYD) algorithm is a quantum analogue of the classical empirical/plug-in estimator, which works as follows. Consider the weak Schur sampling procedure explained in Section II-B1, which outputs $\boldsymbol{\lambda} \vdash n$. The EYD algorithm computes the empirical distribution, which assigns probability $\boldsymbol{\lambda}_i/n$ to the symbol i, and outputs the property f of a mixed state with eigenvalues equal to $\boldsymbol{\lambda}_i/n$.

III. Our Techniques

In this section, we provide a high-level overview of the techniques used to prove our results.

A. Rényi Entropy for Integral $\alpha > 1$

Estimating Rényi entropy is equivalent to obtaining estimates of the power sum $M_\alpha(\boldsymbol{\eta}) \stackrel{\text{def}}{=} \sum \boldsymbol{\eta}_i^\alpha$. In the classical setting, it turns out that for integral $\alpha > 1$, there are simple unbiased estimators of $M_\alpha(\boldsymbol{\eta})$. In the quantum setting, for integral α, (appropriately scaled) $p^{\#}$ polynomials over Young tableaus obtained from Kerov's algebras described earlier are unbiased estimators (see Lemma 2). Our Rényi entropy estimator is described in Algorithm 1.

Algorithm 1 Estimating Rényi entropy for integral α's.

1: **Input:** n independent copies of the state ρ, and $\alpha \in \mathbb{N}$.
2: Run weak Schur sampling to obtain $\boldsymbol{\lambda} \vdash n$.
3: Let (α) be the partition of α with one part.
4: Compute $p^{\#}_{(\alpha)}(\lambda) = n^{\underline{\alpha}} \cdot \frac{\chi^{\lambda}_{(\alpha) \cup 1^{n-\alpha}}}{\dim(\lambda)}$.
5: **Output:** $\frac{1}{1-\alpha} \log\left(\frac{p^{\#}_{(\alpha)}(\lambda)}{n^{\underline{\alpha}}}\right)$.

The estimator is known from existing results to be unbiased. The challenge lies in bounding its variance $\mathrm{Var}\left(p^{\#}_{(\alpha)}(\lambda)\right) = \mathbb{E}\left[p^{\#}_{(\alpha)}(\lambda)^2\right] - \mathbb{E}\left[p^{\#}_{(\alpha)}(\lambda)\right]^2$. It will suffice to show that the variance expression above satisfies $\mathrm{Var}\left(p^{\#}_{(\alpha)}(\boldsymbol{\lambda})\right) \leq \left(\varepsilon \mathbb{E}\left[p^{\#}_{(\alpha)}(\boldsymbol{\lambda})\right]\right)^2$, which will hold when $C_\alpha \cdot n^\alpha \left(1 + n^{\alpha-1} M_{2\alpha-1}(\boldsymbol{\eta})\right) \leq \left(\varepsilon M_\alpha(\boldsymbol{\eta})\right)^2$. Using inequalities between power sums, it can be shown that when n is more than the complexity of Theorem 1, the variance is indeed small, proving the upper bound.

For the lower bound for integral α, the first term $\frac{d^{1-1/\alpha}}{\varepsilon^2}$ follows from the classical lower bounds, and the fact that estimation is easier in the classical setting than in the quantum setting. To prove a lower bound equal to the second term, we invoke the classical Le Cam's method. In particular, for the following two spectrums:

$$\boldsymbol{\eta} = \left(\frac{1 + (\varepsilon d)^{1/\alpha}}{d}, \frac{1 - \frac{(\varepsilon d)^{1/\alpha}}{d-1}}{d}, \ldots, \frac{1 - \frac{(\varepsilon d)^{1/\alpha}}{d-1}}{d}\right), \text{ and}$$

$$\boldsymbol{\nu} = \left(\frac{1}{d}, \ldots, \frac{1}{d}\right)$$

we show that $\mathcal{S}_\alpha(\boldsymbol{\eta}) - \mathcal{S}_\alpha(\boldsymbol{\nu}) = \Theta(\varepsilon)$, and $d_{TV}(SW_{\boldsymbol{\eta}}, SW_{\boldsymbol{\nu}}) < 0.1$, unless $n = \Omega(d^{2-2/\alpha}/\varepsilon^{2/\alpha})$. This proves that unless n is large enough, there is no classifier that can test between the spectrums $\boldsymbol{\eta}$ and $\boldsymbol{\nu}$ with probability greater than $2/3$, implying our lower bound.

Our upper bounds for von Neumann entropy and for non-integral α use the EYD algorithm. Our upper bounds require various bias and concentration results on the Young-tableaux. Fortunately, in the recent works of O'Donnell and Wright, a number of such bounds were proved. We build upon their

results, and prove some additional results to show the copy complexity bounds for the EYD algorithm.

To prove the lower bounds for the EYD algorithm, we design eigenvalues such that unless the number of copies is large enough, the EYD algorithm cannot concentrate around the true entropy. For $\alpha \geq 1$, we use the uniform distribution (maximally mixed state), and for $\alpha < 1$, we design a distribution that is uniform, except for one large eigenvalue.

One of our contributions pertains to the convergence of the empirical Young diagram to the true distribution. A lower bound of d^2/ε^2 was shown by [18]. However, their results only holds with a constant probability (with probability 0.01 to be precise). We show the following sharp concentration. There are constants $\varepsilon_0 > 0$, c_1, and c_2 such that when $\varepsilon \leq \varepsilon_0$, and $n < c_1 d^2/\varepsilon^2$, and ρ is maximally mixed,

$$\Pr\left(\sum_{i=1}^{d}\left|\frac{\boldsymbol{\lambda}_i}{n} - \frac{1}{d}\right| > \varepsilon\right) > 1 - \exp(-c_2 \cdot d).$$

Note that the right-hand side does not depend on n. We show that unless the number of samples is more than d^2/ε^2, the empirical Young diagram's lower bound holds with probability $1 - \exp(-cd)$ for some constant c. This exponential concentration result is of independent interest.

ACKNOWLEDGMENT

The authors thank John Wright for helpful comments on an earlier draft of this paper. This research was supported by the US National Science Foundation under grant 1815893.

REFERENCES

[1] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," *Phys. Rev. Lett.*, vol. 69, pp. 2881–2884, Nov 1992.

[2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar 1993.

[3] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Trans. Info. Theory*, vol. 48, no. 10, pp. 2637–2655, Oct 2002.

[4] I. Devetak, A. W. Harrow, and A. Winter, "A family of quantum protocols," *Phys. Rev. Lett.*, vol. 93, p. 230504, Dec 2004.

[5] M. H. Hsieh and M. M. Wilde, "Entanglement-assisted communication of classical and quantum information," *IEEE Trans. Info. Theory*, vol. 56, no. 9, pp. 4682–4704, Sept 2010.

[6] B. Schumacher, "Quantum coding," *Physical Review A*, vol. 51, no. 4, pp. 2738–2747, 1995.

[7] R. Jozsa and B. Schumacher, "A new proof of the quantum noiseless coding theorem," *Journal of Modern Optics*, vol. 41, no. 12, pp. 2343–2349, 1994.

[8] H.-K. Lo, "Quantum coding theorem for mixed states," *Optics Communications*, vol. 119, no. 5-6, pp. 552–556, 1995.

[9] J. Cardy, "Measuring quantum entanglement," Max Born Lecture, 2012.

[10] X.-S. Ma, T. Herbst, T. Scheidl, D. Wang, S. Kropatschek, W. Naylor, B. Wittmann, A. Mech, J. Kofler, E. Anisimova, V. Makarov, T. Jennewein, R. Ursin, and A. Zeilinger, "Quantum teleportation over 143 kilometres using active feed-forward," *Nature*, vol. 489, no. 7415, pp. 269–273, 2012.

[11] H. Haeffner, W. Haensel, C. F. Roos, J. Benhelm, D. C. al kar, M. Chwalla, T. Koerber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, "Scalable multi-particle entanglement of trapped ions," *Nature*, vol. 438, pp. 643–646, 2005.

[12] J. Acharya, I. Issa, N. V. Shende, and A. B. Wagner, "Measuring Quantum Entropy," Nov 2017. [Online]. Available: http://arxiv.org/abs/1711.00814

[13] R. O'Donnell and J. Wright, "Guest column: A primer on the statistics of longest increasing subsequences and quantum states (shortened version)," *SIGACT News*, vol. 48, no. 3, pp. 37–59, September 2017.

[14] I. G. Macdonald, *Symmetric functions and Hall polynomials*. Oxford university press, 1998.

[15] G. H. Hardy and S. Ramanujan, "Asymptotic formulaæ in combinatory analysis," *Proceedings of the London Mathematical Society*, vol. 2, no. 1, pp. 75–115, 1918.

[16] R. Alicki, S. Rudnicki, and S. Sadowski, "Symmetry properties of product states for the system of n n-level atoms," *Journal of mathematical physics*, vol. 29, no. 5, pp. 1158–1162, 1988.

[17] M. Hayashi and K. Matsumoto, "Quantum universal variable-length source coding," *Physical Review A*, vol. 66, no. 2, p. 022311, 2002.

[18] R. O'Donnell and J. Wright, "Quantum spectrum testing," in *STOC*. ACM, 2015, pp. 529–538.

[19] A. Montanaro and R. de Wolf, "A survey of quantum property testing," *Theory of Computing, Graduate Surveys*, vol. 7, pp. 1–81, 2016.

[20] J. Wright, "How to learn a quantum state," Ph.D. dissertation, Carnegie Mellon University, 2016.

[21] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, "Sample-optimal tomography of quantum states," *IEEE Trans. Info. Theory*, vol. 63, no. 9, pp. 5628–5641, Sep. 2017.

[22] R. O'Donnell and J. Wright, "Efficient quantum tomography," in *ACM STOC*, 2016, pp. 899–912.

[23] ——, "Efficient quantum tomography II," in *ACM STOC*, 2017, pp. 962–974.

[24] C. Bădescu, R. O'Donnell, and J. Wright, "Quantum state certification," *arXiv preprint arXiv:1708.06002*, 2017.

[25] S. Pallister, N. Linden, and A. Montanaro, "Optimal verification of entangled states with local measurements," *Phys. Rev. Lett.*, vol. 120, p. 170502, Apr 2018.

[26] M. Bavarian, S. Mehraban, and J. Wright, Personal Communication, 2016.

[27] S. Bravyi, A. W. Harrow, and A. Hassidim, "Quantum algorithms for testing properties of distributions," *IEEE Trans. Info. Theory*, vol. 57, no. 6, pp. 3971–3981, 2011.

[28] S. Chakraborty, E. Fischer, A. Matsliah, and R. De Wolf, "New results on quantum property testing," *arXiv preprint arXiv:1005.0523*, 2010.

[29] I. S. Sardharwalla, S. Strelchuk, and R. Jozsa, "Quantum conditional query complexity," *Quantum Information and Computation*, vol. 17, no. 7&8, pp. 541–567, 2017.

[30] T. Li and X. Wu, "Quantum query complexity of entropy estimation," *arXiv preprint quant-ph/1710.06025*, 2017.

[31] R. P. Stanley, *Enumerative Combinatorics: Volume 2*. New York, NY, USA: Cambridge University Press, 1999.

[32] C. T. Hepler, "On the complexity of computing characters of finite groups," 1994.

[33] M. Keyl and R. F. Werner, "Estimating the spectrum of a density operator," *Physical Review A*, vol. 64, no. 5, p. 052311, 2001.

[34] A. Childs, A. Harrow, and P. Wocjan, "Weak Fourier-Schur sampling, the hidden subgroup problem, and the quantum collision problem," *STACS 2007*, pp. 598–609, 2007.

[35] A. W. Harrow, "Applications of coherent classical communication and the schur transform to quantum information theory," *arXiv preprint quant-ph/0512255*, 2005.

[36] M. Christandl, "The structure of bipartite quantum states-insights from group theory and cryptography," *arXiv preprint quant-ph/0604183*, 2006.

[37] V. Ivanov and G. Olshanski, "Kerov's central limit theorem for the Plancherel measure on Young diagrams," *Symmetric functions 2001: surveys of developments and perspectives*, vol. 74, pp. 93–151, 2002.