# Quantum Spectrum Testing

Ryan O'Donnell* John Wright†

## 1 Introduction

In this work, we study the problem of *quantum spectrum testing*. Here one is given $n$ copies of a mixed state $\rho \in \mathbb{C}^{d \times d}$ and the goal is to determine whether $\rho$'s spectrum satisfies some property $\mathcal{P}$ or is far from satisfying $\mathcal{P}$. Formally:

**Definition 1.1.** A property $\mathcal{P}$ is testable with $f(d, \epsilon)$ copies if for every $d \geq 1, \epsilon > 0$ there is an algorithm $\mathcal{T}$ which, when given $f(d, \epsilon)$ copies of a mixed state $\rho \in \mathbb{C}^{d \times d}$ with spectrum $\eta = (\eta_1, \ldots, \eta_d)$, behaves as follows: (i) If $\eta \in \mathcal{P}$, then $\mathbf{Pr}[\mathcal{T} \text{ accepts}] \geq 2/3$. (ii) If $\eta$ is $\epsilon$-far from $\mathcal{P}$ in total variation distance, then $\mathbf{Pr}[\mathcal{T} \text{ rejects}] \geq 2/3$.

This model is the natural quantum analogue of the classical problem of testing symmetric properties of probability distributions. It can also be shown [Mon14][1] equivalent to the model of testing unitarily invariant properties of mixed states, as proposed by Montanaro and de Wolf [MdW13]. A large focus of this work is addressing questions raised by their survey.

There are two previous results directly relevant to results in our work. The first is an algorithm for learning the spectrum of an unknown mixed state. This algorithm is naturally suggested by the early work of Alicki, Rudnicki, and Sadowski [ARS88] and was explicitly proposed by Keyl and Werner [KW01]. Regarding its performance guarantee, Hayashi and Matsumoto [HM02] gave explicit error bounds and a short proof, but their work contained some small calculational errors, subsequently corrected by Christandl and Mitchison [CM06]. From the last of these it is easy to deduce the following:

**Theorem 1.2.** *There is an algorithm which, given $O(d^2/\epsilon^2 \cdot \ln(d/\epsilon))$ copies of a mixed state $\rho$ with spectrum $\eta$, outputs with high probability an estimate of $\eta$ that is $\epsilon$-close in total variation distance.*

The second result comes from the work of Childs et al. [CHW07]. It can be thought of as a quantum analogue of the (Birthday Paradox-based) fact that $\Theta(\sqrt{r})$ samples are necessary and sufficient to distinguish a distribution which is uniform on half of $[2r]$ from one which is uniform on all of $[2r]$:

**Theorem 1.3.** $\Theta(r)$ *copies of a state $\rho$ are necessary and sufficient to distinguish between the cases when $\rho$'s spectrum is uniform on either $r$ or $2r$ values. (The bound also holds for $r$ vs. $cr$ for integers $c > 2$.)*

Theorem 1.2 gives an upper bound of $O(d^2/\epsilon^2 \cdot \ln(d/\epsilon))$ copies for testing any property in the model of Definition 1.1, whereas Theorem 1.3 gives a lower bound of $\Omega(d)$ for various properties of spectra. Thus, when considering specific properties, we hope for *subquadratic* algorithms, though we usually cannot hope for sublinear algorithms. This is in contrast with property testing of classical probability distributions, in which sublinear algorithms are the main goal, with the Birthday Paradox typically precluding sub-$O(\sqrt{d})$-sample algorithms.

We have four main results. The first concerns the property that Montanaro and de Wolf refer to as **Mixedness**:

*Dept. of Computer Science, Carnegie Mellon University. Supported by NSF grants CCF-0747250 and CCF-1116594. Some work performed while at Boğaziçi University, supported by Marie Curie IIF project #626373.

†Dept. of Computer Science, Carnegie Mellon University. Supported in part by a Simons Graduate Fellowship.

[1] Via an inequality relating trace distance of mixed states to the total variation distance of their sorted spectra.

**Theorem 1.4.** $\Theta(d/\epsilon^2)$ *copies are necessary and sufficient to test whether* $\rho \in \mathbb{C}^{d\times d}$ *is the maximally mixed state.*

(In terms of Definition 1.1, a state is maximally mixed if its spectrum is $\eta = (1/d, \ldots, 1/d)$.) This is the quantum analogue of the result of Paninski [Pan08] which states that $\Theta(\sqrt{d}/\epsilon^2)$ samples are necessary and sufficient to test whether an unknown probability distribution is uniform (see also [GR11, BFR+13]).

Our second result gives new bounds for testing whether a state has low rank.

**Theorem 1.5.** $\Theta(r^2/\epsilon)$ *copies are necessary and sufficient to test whether* $\rho \in \mathbb{C}^{d\times d}$ *has rank* $r$ *with perfect completeness. With imperfect completeness, a lower bound of* $\Omega(r/\epsilon)$ *holds.*

We note that the copy complexity is independent of the ambient dimension $d$. Knowing that a state is low rank can often make solving a given problem much simpler. For example, quantum state tomography can be made more efficient when the state is known to be low-rank [FGLE12].

Next, we extend Theorem 1.3 to $r$ vs. $r'$ for *any* $r+1 \leq r' \leq 2r$. A qualitative difference is seen when $r' = r+1$; namely, nearly quadratically many copies are necessary.

**Theorem 1.6.** *Let* $1 \leq \Delta \leq r$. *Then* $O(r^2/\Delta)$ *copies are sufficient to distinguish between the cases when* $\rho$*'s spectrum is uniform on either* $r$ *or* $r+\Delta$ *eigenvalues; further, a nearly matching lower bound of* $\Omega((r^2/\Delta)^{1-\epsilon})$ *copies holds for any* $\epsilon > 0$.

As above, we note that these bounds are independent of the ambient dimension $d$. The $1-\epsilon$ in the exponent is an artefact of our proof technique; it is an interesting open problem as to whether a tight $\Omega(r^2/\ell)$ lower bound can be obtained.

Our final result shows that the analysis of the algorithm from Theorem 1.2 is tight up to logarithmic factors.

**Theorem 1.7.** *If* $\rho \in \mathbb{C}^{d\times d}$ *is the maximally mixed state, the algorithm from Theorem 1.2* fails *to give an* $\epsilon$*-accurate estimate (with high probability) unless* $\Omega(d^2/\epsilon^2)$ *copies are used.*

To our knowledge, no such lower bound was known previously. We remark that it is an interesting open question as to whether some *other* algorithm can estimate an unknown state's spectrum with from a subquadratic number of copies.

## 2 Techniques

Following [ARS88, Har05, CM06, CHW07], we use techniques from representation theory of the symmetric group $\mathfrak{S}_n$. A basic tool is *Schur-Weyl duality*, which decomposes the space $(\mathbb{C}^d)^{\otimes n}$ as

$$(\mathbb{C}^d)^{\otimes n} \overset{\mathfrak{S}_n \times \mathcal{U}_d}{\cong} \bigoplus_{\lambda \vdash n} \mathcal{P}_\lambda \otimes \mathcal{Q}_\lambda^d, \tag{1}$$

where the subspace $\mathcal{P}_\lambda$ corresponds to the symmetric group, the subspace $\mathcal{Q}_\lambda^d$ corresponds to the unitary group, and $\lambda$ is a partition of $n$ (meaning a tuple $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ satisfying $\lambda_1 \geq \ldots \geq \lambda_\ell \geq 0$ and $\lambda_1 + \ldots + \lambda_\ell = n$). In our testing problem, the tester is provided $\rho^{\otimes n}$, which is invariant under any permutation of the $n$ coordinates, and whether it accepts or rejects should be invariant under any unitary transformation of $\rho$. This means that if we measure $\rho^{\otimes n}$ in the *Schur basis* given by Equation (1), we can throw away the information from the permutation and unitary registers without losing any relevant information. What is left is only the "irrep" label $\lambda$.

The end result is this: there is a sampling algorithm—referred to in [CHW07] as *weak Schur sampling*—which, on input a mixed state $\rho^{\otimes n}$, outputs a random partition $\lambda$ whose distribution depends on the spectrum of $\rho$. We will denote this distribution by $\mathrm{SW}_n(\rho)$. Furthermore, an argument which is essentially from [CHW07] (though see [MdW13, Lemma 19] for a full statement) shows that for any property $\mathcal{P}$, there is an *optimal* tester in the model of Definition 1.1 whose operation is as follows: 1. Sample $\boldsymbol{\lambda} \sim \mathrm{SW}_n(\rho)$. 2. Accept or reject based only on $\boldsymbol{\lambda}$. We may therefore proceed without loss of generality by analyzing only algorithms of this form. This necessitates understanding the distribution $\mathrm{SW}_n(\rho)$.

In case $\rho$ is the maximally mixed state, the distribution $\mathrm{SW}_n(\rho)$ is somewhat well-studied [Bia01, Mél10]. It is known as the *Schur-Weyl* distribution, and we denote it by $\mathrm{SW}_{n,d}$. (In the limit as $d \to \infty$, it approaches the well-known *Plancherel* distribution.) The exact distribution on partitions given by $\mathrm{SW}_{n,d}$ is quite complicated and difficult to work with, and so various works have instead sought to describe large-scale features of a "typical" $\boldsymbol{\lambda} \sim \mathrm{SW}_{n,d}$. For example, Biane [Bia01] has shown that, up to small fluctuations, the "shape" of a random $\boldsymbol{\lambda} \sim \mathrm{SW}_{n,d}$ tends toward a certain limiting shape $\Omega$ which depends only on the ratio $\frac{\sqrt{n}}{d}$. Furthermore, Meliot [Mél10] has characterized these small fluctuations as being distributed like a certain Gaussian process. The second of these results borrows heavily from a proof of the analogous result for the Plancherel measure by Kerov [IO02], and we will give an overview his techniques below.

Kerov's approach involves studying a certain space of symmetric polynomial functions on Young diagrams. For example, if one is interested in showing that a random $\boldsymbol{\lambda} \sim \mathrm{SW}_{n,d}$ tends to have some coordinates which are much larger than the rest, then it would be natural to study "moments" of the form $\sum \boldsymbol{\lambda}_i^k$. However, the approach of Kerov would suggest studying the following "moments" instead:

$$p_k(\boldsymbol{\lambda}) \coloneqq \sum_{i=1}^{\infty} [(\boldsymbol{\lambda}_i - i + \tfrac{1}{2})^k - (-i + \tfrac{1}{2})^k], \text{ for } k \geq 1.$$

The polynomial family $(p_k)$ inhabits (in fact, it generates) the so-called *algebra of polynomial functions on the set of Young diagrams* $\Lambda^*$ (also known as Kerov's *algebra of observables*). There are other important polynomial families within $\Lambda^*$—in addition to the $p_k$ polynomials, our work uses the $\widetilde{p}_k$, $p_\mu^\#$, $s_\mu^*$, and $c_k$ polynomials— and each of these families sheds light on a different aspect of the input partition $\lambda$. For example, though the $p_\mu^\#(\lambda)$ polynomials don't give any obvious information regarding the "shape" of $\lambda$, they are unique in that we can easily compute the expectation $\mathbf{E}_{\boldsymbol{\lambda} \sim \mathrm{SW}_n(\rho)}[p_\mu^\#(\boldsymbol{\lambda})]$ for any mixed state $\rho$. There are various methods for passing from one polynomial family to another, and it is often the case that a problem most easily stated in terms of one polynomial family is most easily solved in terms of another.

Our lower bounds generally have the following outline: 1. Reduce the problem to showing that a certain expression within the algebra of observables is small with high probability. 2. Use various polynomial-estimation techniques developed by Kerov and others for proving concentration of said expression. Thus, e.g., proving the lower bound in Theorem 1.6 roughly reduces to showing that the expression $\sum_{k=1}^{\infty} (-1/r)^{k+1} c_k(\boldsymbol{\lambda})$ is small (with high probability) for $\boldsymbol{\lambda} \sim \mathrm{SW}_{n,d}$. As another example, proving the lower bound in Theorem 1.4 reduces to showing that the expression $\sum_\mu s_\mu^*(\boldsymbol{\lambda}) s_\mu(\epsilon, -\epsilon, \ldots, \epsilon, -\epsilon)/d^{\uparrow \mu}$ is small (with high probability) for $\boldsymbol{\lambda} \sim \mathrm{SW}_{n,d}$. Our upper bounds generally involve analyzing algorithms which accept or reject based on simple statistics of the sampled $\boldsymbol{\lambda} \sim \mathrm{SW}_{n,d}$. For example, the rank tester of Theorem 1.5 accepts iff the sampled $\boldsymbol{\lambda}$ has at most $r$ nonzero parts, and the uniformity tester of Theorem 1.4 accepts iff $c_1(\boldsymbol{\lambda})$ is sufficiently small. As in the lower bounds, analyzing these algorithms uses techniques from the algebra of observables, and we sometimes also require certain combinatorial interpretations of the weak Schur sampling algorithm; e.g., its relationship with the Robinson–Schensted–Knuth "bumping" algorithm.

## References


[ARS88] Robert Alicki, Slawomir Rudnicki, and Slawomir Sadowski. Symmetry properties of product states for the system of N n-level atoms. *Journal of mathematical physics*, 29(5):1158–1162, 1988. 1, 2

[BFR+13] Tuğkan Batu, Lance Fortnow, Ronitt Rubinfeld, Warren D Smith, and Patrick White. Testing closeness of discrete distributions. *Journal of the ACM*, 60(1):4, 2013. 1

[Bia01] Philippe Biane. Approximate factorization and concentration for characters of symmetric groups. *International Mathematics Research Notices*, 2001(4):179–192, 2001. 2

[CHW07] Andrew M Childs, Aram W Harrow, and Paweł Wocjan. Weak Fourier-Schur sampling, the hidden subgroup problem, and the quantum collision problem. In *24th Annual Symposium on Theoretical Aspects of Computer Science*, pages 598–609, 2007. 1, 2, 2

[CM06] Matthias Christandl and Graeme Mitchison. The spectra of quantum states and the Kronecker coefficients of the symmetric group. *Communications in mathematical physics*, 261(3):789–797, 2006. 1, 2

[FGLE12] Steven T Flammia, David Gross, Yi-Kai Liu, and Jens Eisert. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 14(9):095022, 2012. 1

[GR11] Oded Goldreich and Dana Ron. On testing expansion in bounded-degree graphs. In *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 68–75. Springer, 2011. 1

[Har05] Aram W Harrow. *Applications of coherent classical communication and the Schur transform to quantum information theory*. PhD thesis, Massachusetts Institute of Technology, 2005. 2

[HM02] Masahito Hayashi and Keiji Matsumoto. Quantum universal variable-length source coding. *Physical Review A*, 66(2):022311, 2002. 1

[IO02] Vladimir Ivanov and Grigori Olshanski. Kerov's central limit theorem for the Plancherel measure on Young diagrams. In *Symmetric functions 2001: surveys of developments and perspectives*, pages 93–151. Springer, 2002. 2

[KW01] Michael Keyl and Reinhard F Werner. Estimating the spectrum of a density operator. *Physical Review A*, 64(5):052311, 2001. 1

[MdW13] Ashley Montanaro and Ronald de Wolf. A survey of quantum property testing. Technical report, arXiv:1310.2035, 2013. 1, 2

[Mél10] Pierre-Loïc Méliot. Kerov's central limit theorem for Schur-Weyl measures of parameter 1/2. Technical report, arXiv:1009.4034, 2010. 2

[Mon14] Ashley Montanaro. Personal communication, 2014. 1

[Pan08] Liam Paninski. A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Transactions on Information Theory*, 54(10):4750–4755, 2008. 1