A Quantum Outlier Theorem

Samuel Epstein
samepst@jptheorygroup.org

March 8, 2023

Abstract

In recent results, it has been proven that all sampling methods produce outliers. In this paper, we extend this result to quantum information theory. Projectors of large rank must contain pure quantum states in their images that are outlying states. Otherwise, the projectors are exotic, in that they have high mutual information with the halting sequence. Thus quantum coding schemes that use projections, such as Schumacher compression, must communicate using outlier quantum states.

1 Introduction

In algorithmic information theory, the notion of an outlier is modelled using the randomness deficiency. The model is defined by a probability p, over natural numbers, and the data point x is a natural number. The deficiency of randomness is formally defined as

$$\mathbf{d}(x|p) = \log \mathbf{t}_p(x),$$

where $\mathbf{t}_p$ is a universal lower computable p-test. It is a score of how atypical a datapoint is with respect to a model. In [G01́], the quantum notion of randomness deficiency was introduced. This quantum deficiency measures the algorithmic atypicality of a pure or mixed quantum state ρ with respect to a second mixed quantum state σ. It is defined by

$$\mathbf{d}(\rho|\sigma) = \log \mathrm{Tr}\rho\mathbf{t}_\sigma,$$

where $\mathbf{t}_\sigma$ is a universal lower computable Hermitian matrix such that $\mathrm{Tr}\sigma\mathbf{t}_\sigma \leq 1$. The density matrix σ is assumed to be computable. Mixed states are used to model random mixtures $\{p_i\}$ of pure states $\{|\psi_i\rangle\}$, so a quantum deficiency is a score of how atypical a quantum state is with respect to a mixture. In [Eps19], quantum deficiency was extended to uncomputable quantum states.

What are the interesting properties of quantum deficiency? In [Eps19], conservation of quantum deficiency was proven over partial trace and unitary operations. With some work conservation over quantum operations can be proven. One recent result in the classical randomness deficiency case is that sampling methods produce outliers [Eps21]. There are several proofs to this result, with one of them derived from the fact that large sets of natural numbers with low randomness deficiencies are exotic, in that they have high mutual information with the halting sequence.

In this paper, we prove a quantum version of this result. Projections of large rank must contain pure quantum states in their images that are outlying states. Otherwise, the projections are exotic, in that they have high mutual information with the halting sequence. Thus quantum coding schemes that use projections, such as Schumacher compression, must communicate using outlier quantum states.

2 Conventions

The length of a string $x \in \{0,1\}^*$ is $\|x\|$. For positive real function f, $<^+ f$, $>^+ f$, and $=^+ f$ is used to represent $< f + O(1)$, $> f + O(1)$, and $= f \pm O(1)$. For the nonnegative real function f, the terms $<^{\log} f$, $>^{\log} f$, and $=^{\log} f$ represent the terms $<f+O(\log(f+1))$, $>f-O(\log(f+1))$, and $=f\pm O(\log(f+1))$, respectively.

Let $\mathbf{K}(x)$ be the prefix free Kolmogorov complexity. We use $\mathbf{I}(x;\mathcal{H}) = \mathbf{K}(x) - \mathbf{K}(x|\mathcal{H})$ to be the amount of information that the halting sequence $\mathcal{H} \in \{0,1\}^\infty$ has about $x \in \{0,1\}^*$.

We use $\mathcal{H}_n$ to denote a Hilbert space with n dimensions, spanned by bases $|\beta_1\rangle, \ldots, |\beta_n\rangle$. A qubit is a unit vector in the Hilbert space $\mathcal{H}_2$, spanned by vectors $|0\rangle$, $|1\rangle$. To model n qubits, we use a unit vector in $\mathcal{H}_{2^n}$, spanned by basis vectors $|x\rangle$, where x is a string of size n.

A pure quantum state $|\psi\rangle$ of length n is a unit vector in $\mathcal{H}_{2^n}$. Its corresponding element in the dual space is denoted by $\langle\phi|$. The conjugate transpose of a a matrix A is A^*. Tr is used to denote the trace of a matrix. Projection matrices are Hermitian matrices with eigenvalues in $\{0,1\}$.

A complex matrix A is elementary if its entries are complex numbers with rational coefficients and can be encoded as $\langle A\rangle$, and has a Kolmogorov complexity $\mathbf{K}(A)$ and algorithmic probability $\mathbf{m}(A)$.

For positive semidefinite matrices, $\sigma \leq \rho$ iff $\rho - \sigma$ is positive semidefinite. We say program $q \in \{0,1\}^*$ lower computes positive semidefinite matrix σ if, given as input to universal Turing machine U, the machine U reads $\leq \|q\|$ bits and outputs, with or without halting, a sequence of elementary semi-density matrices $\{\sigma_i\}$ such that $\sigma_i \leq \sigma_{i+1}$ and $\lim_{i\to\infty}\sigma_i = \sigma$. A matrix T is lower computable if there is a program that lower computes it. Its complexity is $\mathbf{K}(T) = \min\{\mathbf{K}(q) : q \text{ lower computes } T\}$. Given a density matrix σ, a σ-test is a lower computable semi-definite matrx T such that $\mathrm{Tr}T\sigma = 1$. If σ is computable, there exists a universal σ test $\mathbf{t}_\sigma$, that is lower computable relative to the number of qubits n, $\mathrm{Tr}\sigma\mathbf{t}_\sigma \leq 1$, and for every lower computable σ test T, $O(1)\mathbf{t}_\sigma > 2^{-\mathbf{K}(T|n)}T$. This universal test can be computed in the standard way, analagously to the classical case (see [Ǵ21]).

Definition 1 (Quantum Deficiency) *For mixed states σ and ρ, $\mathbf{d}(\rho|\sigma) = \log \mathbf{t}_\sigma\rho$.*

A probability is elementary if it has finite support and all its values are rational. The deficiency of randomness of a string x with respect to an elementary probability mesaure Q is $\mathbf{d}(x|Q) = \lfloor -\log Q(x)\rfloor - \mathbf{K}(x|\langle Q\rangle)$. The stochasticity of a string is $\mathbf{Ks}(x) = \min_Q\{\mathbf{K}(Q) + 3\log\mathbf{d}(x|Q)\}$.

Lemma 1 ([Eps21, Lev16]) $\mathbf{Ks}(x) <^{\log} \mathbf{I}(x;\mathcal{H})$.

3 Results

Theorem 1 *Relativized to an n qubit mixed state σ, for elementary 2^m rank projector P, $3m - 2n <^{\log} \max_{|\phi\rangle\in\mathrm{Image}(P)} \mathbf{d}(|\phi\rangle\,|\sigma) + \mathbf{I}(\langle P\rangle;\mathcal{H})$.*

Proof. We relativize the universal Turing machine to $\langle\sigma\rangle$ and $(3m-2n)$. Thus it is effectively relativized to m, n, and σ. Let Q and d realize $\mathbf{Ks}(P)$, where $d = \mathbf{d}(P|Q)$. Without loss of generality we can assume that the support of Q is elementary projections of rank 2^m. There are $d2^{n-m}$ rounds. For each round we select an σ-test that is of dimension 1. We now describe the selection process.

Select a random test T to be $2^{m-3}|\psi\rangle\langle\psi|$, where $|\psi\rangle$ is an n qubit state chosen uniformly from the unit sphere, with distribution Λ.

$$\mathbf{E}[\mathrm{Tr}T\sigma] = 2^{m-3}\int \mathrm{Tr}\,\langle\psi|\,\sigma\,|\psi\rangle\, d\Lambda = 2^{m-3}\mathrm{Tr}\sigma\int \mathrm{Tr}\,\langle\psi|\psi\rangle\, d\Lambda = 2^{m-n-3}\mathrm{Tr}\sigma = 2^{m-n-3}.$$

Thus the probability that T is a σ-test is $\geq 1-2^{m-n-3}$. Let I_m be an n-qubit identity matrix with only the first 2^m diagonal elements being non-zero. Let $K_m = I - I_m$. Let $p = 2^{m-n}$ and $\hat{T} = T/2^{m-3}$. For any projection B of rank 2^m,

$$\begin{aligned}
&\Pr(\mathrm{Tr}B\hat{T} \leq .5p)\\
=&\Pr(\mathrm{Tr}I_m\hat{T} \leq .5p)\\
=&\Pr(\mathrm{Tr}K_m\hat{T} \geq 1-.5p)\\
\mathbf{E}[\mathrm{Tr}K_m\hat{T}] &= 1-p\\
\Pr(\mathrm{Tr}K_m\hat{T} \geq 1-.5p) &\leq (1-p)/(1-.5p)\\
\Pr(\mathrm{Tr}B\hat{T} \geq .5p) &= 1-\Pr(\mathrm{Tr}K_m\hat{T} \geq 1-.5p)\\
&\geq 1-(1-p)/(1-.5p)\\
&= .5p/(1-.5p) \geq .25p\\
\Pr(\mathrm{Tr}BT \geq 2^{2m-n-4}) &\geq .25p.
\end{aligned}$$

Let $k = 2^{2m-n-4}$. Let Ω be the space of all matrices of the form $2^{m-3}\left|\phi\right\rangle\left\langle\phi\right|$. Let R be the uniform distribution over Ω. Let $[A,B]$ be 1 if $\mathrm{Tr}AB > k$, and 0 otherwise. By the above equations, for all $A \in \mathrm{Support}(Q)$, $\int_\Omega [A,B]dR(B) \geq .25p$. So $\sum_A \int_\Omega [A,B]Q(A)dR(B) \geq .25p$. For Hermitian matrix A, $\{A\}$ is 1 if $\mathrm{Tr}A\sigma \leq 1$, and 0 otherwise. So $\int_\Omega \{A\}dR(A) \geq (1-p2^{-3})$. Let $f = \max_T \{T\}\sum Q(A)[T,A]$. So

$$\begin{aligned}
.25p &\leq \sum_A \int_\Omega [A,B]Q(A)dR(B)\\
&= \sum_A \int_\Omega \{B\}Q[A,B](A)dR(B) + \sum_A \int_\Omega (1-\{B\})[A,B]Q(A)dR(B)\\
&\leq \sum_A \int_\Omega \{B\}[A,B]Q(A)dR(B) + \int_\Omega (1-\{B\})dR(B)\\
&\leq \sum_A \int_\Omega \{B\}[A,B]Q(A)dR(B) + p2^{-3}\\
p/8 &\leq \sum_A \int_\Omega \{B\}[A,B]Q(A)dR(B) = \int_\Omega \left(\{B\}\sum_A [A,B]Q(A)\right) dR(B) \leq \int_\Omega f dR(B)\\
p/8 &\leq f.
\end{aligned}$$

Thus for each round i, the lower bounds on f proves there exists a one dimensional matrix $T_i = 2^{m-3}\left|\psi\right\rangle\left\langle\psi\right|$ such that $\mathrm{Tr}T_i\sigma \leq 1$ and $\sum_R \{Q(R) : \mathrm{Tr}T_iR \geq k\} \geq p/8 = 2^{m-n-3}$. Such a T_i is selected, and the the Q probability is conditioned on those projections B for which $[T_i,B]=0$. Assuming that there are $d2^{n-m+3}$ rounds, the Q measure of projections B such there does not exist a T_i with $[T_i,B]=1$ is

$$\leq (1-p/8)^{d2^{n-m+3}} \leq e^{-d}.$$

Thus there exists a T_i such that $[T_i,P]=1$, otherwise one can create a Q test t that assigns e^d to all projections B where there does not exist T_i with $[T_i,B]=1$, and 0 otherwise. Then $t(P)=e^d$ so

$$1.44d < \log t(P) <^+ \mathbf{d}(P|Q) =^+ d.$$

This is a contradiction, because without loss of generality, one can assume d is large. Let $T_i = 2^{m-3}\left|\psi\right\rangle\left\langle\psi\right|$ with $[T_i, P] = 1$. Let $\left|\phi\right\rangle = P\left|\psi\right\rangle/\sqrt{\left\langle\psi\right|P\left|\psi\right\rangle}$. So $\left\langle\phi\right|T_i\left|\phi\right\rangle \geq 2^{2m-n-4}$ and $\left|\phi\right\rangle$ is in the image of P. Thus by Lemma 1,

$$
\begin{aligned}
k &\leq \log\left\langle\phi\right|T_i\left|\phi\right\rangle \\
2m-n &<^+ \log\left\langle\phi\right|T_i\left|\phi\right\rangle \\
2m-n &<^+ \log\max_{\left|\phi\right\rangle\in\mathrm{Image}(P)}\left\langle\phi\right|T_i\left|\phi\right\rangle \\
2m-n &<^+ \max_{\left|\phi\right\rangle\in\mathrm{Image}(P)}\mathbf{d}(P|\sigma)+\mathbf{K}(T_i) \\
2m-n &<^+ \max_{\left|\phi\right\rangle\in\mathrm{Image}(P)}\mathbf{d}(P|\sigma)+(n-m)+\log d+\mathbf{K}(d)+\mathbf{K}(Q) \\
3m-2n &<^{\log} \max_{\left|\phi\right\rangle\in\mathrm{Image}(P)}\mathbf{d}(P|\sigma)+\mathbf{I}(P;\mathcal{H}).
\end{aligned}
$$

Note that due to the fact that the left hand side of the equation is $(3m-2n)$ and it has log precision, this enables one to condition the universal Turing machine to $(3m-2n)$. □

3.1 Computable Projections

Theorem 1 is in terms of elementary described projecctions and can be generalized to arbitrarily computable projections. For a matrix M, let $\|M\| = \max_{i,j}|M_{i,j}|$ be the max norm. A program $p \in \{0,1\}^*$ computes a projection P of rank ℓ if it outputs a series of rank ℓ projections $\{P_i\}_{i=1}^{\infty}$ such that $\|P - P_i\| \leq 2^{-i}$. For computable projection operator P, $\mathbf{I}(P;\mathcal{H}) = \min\{\mathbf{K}(p) - \mathbf{K}(p|\mathcal{H}) :$ p is a program that computes $P\}$.

Lemma 2 ([Eps22]) *For partial computable f, $\mathbf{I}(f(a);\mathcal{H}) <^+ \mathbf{I}(a;\mathcal{H}) + \mathbf{K}(f)$.*

Corollary 1 *Relativized to an n qubit mixed state σ, for computable 2^m rank projector P, $3m - 2n <^{\log} \max_{\left|\phi\right\rangle\in\mathrm{Image}(P)}\mathbf{d}(\left|\phi\right\rangle|\sigma) + \mathbf{I}(\langle P\rangle;\mathcal{H})$.*

Proof. Let p be a program that computes P. There is a simply defined algorithm A, that when given p and σ, outputs P_n such that $\max_{\left|\psi\right\rangle\in\mathrm{Image}(P)}\mathbf{d}(\left|\psi\right\rangle|\sigma) =^+ \max_{\left|\psi\right\rangle\in\mathrm{Image}(P_n)}\mathbf{d}(\left|\psi\right\rangle|\sigma)$. Thus by Lemma 2, one gets that $\mathbf{I}(P_n;\mathcal{H}) <^+ \mathbf{I}(P;\mathcal{H})$. The corollary follows from Theorem 1. □

References

[Eps19] S. Epstein. Algorithmic no-cloning theorem. *IEEE Transactions on Information Theory*, 65(9):5925–5930, 2019.

[Eps21] Samuel Epstein. All sampling methods produce outliers. *IEEE Transactions on Information Theory*, 67(11):7568–7578, 2021.

[Eps22] S. Epstein. The outlier theorem revisited. *CoRR*, abs/2203.08733, 2022.

[G0́1] P. Gács. Quantum Algorithmic Entropy. *Journal of Physics A Mathematical General*, 34(35), 2001.

[G2́1] Peter Gács. Lecture notes on descriptional complexity and randomness. *CoRR*, abs/2105.04704, 2021.

[Lev16] L. A. Levin. Occam bound on lowest complexity of elements. *Annals of Pure and Applied Logic*, 167(10):897–900, 2016. And also: S. Epstein and L.A. Levin, Sets have simple members, arXiv preprint arXiv:1107.1458, 2011.