On the Algorithmic Content of Quantum Measurements

Samuel Epstein*

May 16, 2021

Abstract

We show that given a quantum measurement, for an overwhelming majority of pure states, no meaningful information is produced. This is independent of the number of outcomes of the quantum measurement. Due to conservation inequalities, such random noise cannot be processed into coherent data.

1 Introduction

Quantum information theory studies the limits of communicating through quantum channels. In [Hol73], the Holevo bound was proven, providing an upper bound on the amount of classical information shared between two parties that can prepare and measure mixed states. The Holevo bound states that only n bits of classical information can be accessed from n qubits. The main result of this paper shows the limitations of the algorithmic content of measurements of pure quantum states. Given a measurement apparatus E, there is only a tiny fraction of quantum pure states on which E's application produces coherent information. This is independent of the number of measurement outcomes of E.

In this paper we introduce a novel definition of algorithmic information between probabilities. This definition uses the information of two strings x and y, $\mathbf{i}(x : y)$, defined in Section 3.

Definition. *The amount of algorithmic information between two probabilities p and q over Σ^* is* $\mathbf{i}(p : q) = \log \sum_{x,y \in \Sigma^*} 2^{\mathbf{i}(x \,:\, y)} p(x) q(y)$.

If a probability p has low self information $\mathbf{i}(p : p)$ then it has no meaningful information. Generally speaking, such probabilities either have high measure for simple strings and/or very low measure over complex strings. As shown in the appendix, the information between two probabilities is conserved over randomized transformations. Therefore there are no means to increase the self-information of a probability.

In quantum mechanics, given a quantum state $|\psi\rangle$, a measurement, or POVM, E produces a probability measure $E|\psi\rangle$ over strings. This probability represents the classical information produced from the measurement. The exact structure of POVMs is described in Section 5. The main theorem of this paper states that given a measurement E, for an overwhelming majority of quantum states $|\psi\rangle$, the probability produced will have no meaningful information, i.e. $\mathbf{i}(E|\psi\rangle : E|\psi\rangle)$ is negligible.

Theorem. *Let Λ be the uniform distribution on the unit sphere of an n qubit space. Relativized to POVM E, $\int 2^{\mathbf{i}(E|\psi\rangle : E|\psi\rangle)} d\Lambda = O(1)$.*

---

*JP Theory Group, samepst@jptheorygroup.org

2 Related Work

For information about the history and foundation of algorithmic information theory, we refer readers to the textbooks [DH10] and [LV08]. There are several definitions that model the algorithmic content of a quantum state. In [BvL01], the complexity of a quantum state is equal to the size of the smallest quantum Turing machine that can approximate the state to a given fidelity. In [MB05], the algorithmic complexity of a quantum state is equal to the minimal length of an encoding of the preparation of the state through quantum gates. In [Gac01], the algorithmic entropy of a quantum state is measured by the negative logarithmic of the state multiplied by a universal lower computable semi-density matrix. In [Vit00], the entropy of a pure quantum state is equal to the classical complexity of an elementary approximating state plus the negative logarithm of their fidelity. A quantum version of Brudno's theorem was proven in [BKM$^+$06]. Randomness for infinite quantum spin chains, called quantum Martin Löf random sequences, was introduced in [NS19]. An infinite version of algorithmic entropy can be found at [BOD14].

3 Conventions

We use Σ^* to denote finite strings. The length of a string $x \in \Sigma^*$ is $\|x\|$. For positive real function f, $<^+ f$, $>^+ f$, and $=^+ f$ is used to represent $< f + O(1)$, $> f + O(1)$, and $= f \pm O(1)$. In addition $\overset{*}{<}f$, $\overset{*}{>}f$ denote $< f/O(1)$, $> f/O(1)$. The terms $\overset{*}{=}f$ denotes $\overset{*}{<}f$ and $\overset{*}{>}f$. The encoding of $x \in \Sigma^*$ is $\langle x \rangle = 1^{\|x\|}0x$.

For strings $x, y \in \Sigma^*$, the output of algorithm T on input x and auxiliary input y is denoted $T_y(x)$. An algorithm T is prefix free if for strings $x, y, s \in \Sigma^*, \neq \emptyset$, if $T_y(x)$ halts then $T_y(xs)$ does not halt. There exists a universal prefix free algorithm U, where for all prefix-free algorithms T, there exists a $t \in \Sigma^*$, where for all $x, y \in \Sigma^*$, $U_y(tx) = T_y(x)$. This U is used to define Kolmogorov complexity, with $\mathbf{K}(x/y) = \min\{\|p\| : U_y(x) = p\}$. The universal probability of $x \in \Sigma^*$, conditional to $y \in \Sigma^*$, is $\mathbf{m}(x/y) = \sum\{2^{-\|p\|} : U_y(p) = x\}$. The algorithmic information between two strings is $\mathbf{i}(x : y) = \mathbf{K}(x) + \mathbf{K}(y) - \mathbf{K}(x, y)$. The expression "relativized to an elementary object" seen in theorems is equivalent to saying there is an encoding of the elementary object on an auxiliary tape of the universal algorithm U.

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 $\mathcal{G} = \mathcal{H}_2$, 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 $\langle\psi|$. The conjugate transpose of a matrix A is A^*. The tensor product of two matrices A and B is $A \otimes B$. Tr is used to denote the trace of a matrix, and for Hilbert space $\mathcal{H}_X \otimes \mathcal{H}_Y$, the partial trace with respect to $\mathcal{H}_Y$ is Tr_Y.

For positive semi-definite matrices A and B, we say $B \preceq A$, if $A - B$ is positive semi-definite. For functions f whose range are Hermitian matrices, we use $\overset{*}{<}f$ and $\overset{*}{>}f$ to denote $\preceq f/O(1)$ and $\succeq f/O(1)$. We use $\overset{*}{=}f$ to denote $\overset{*}{<}f$ and $\overset{*}{>}f$.

Density matrices are used to represent mixed states, and are self-adjoint, positive definite matrices with trace equal to 1. Semi-density matrices are density matrices except they may have a trace in [0,1].

Pure and mixed quantum states are elementary if their values are complex numbers with rational coefficients, and thus they can be represented with finite strings. Thus elementary quantum states $|\phi\rangle$ and ρ can be encoded as strings, $\langle|\phi\rangle\rangle$ and $\langle\rho\rangle$, and assigned Kolmogorov complexities $\mathbf{K}(|\phi\rangle)$,

$\mathbf{K}(\rho)$ and algorithmic probabilities $\mathbf{m}(|\phi\rangle)$ and $\mathbf{m}(\rho)$. They are equal to the complexity (and algorithmic probability) of the strings that encodes the states.

More generally, 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 $\mathbf{m}(A)$.

In [Gac01], a universal lower computable semi-density matrix, $\boldsymbol{\mu}$ was introduced. It can be defined (up to a multiplicative constant) by $\boldsymbol{\mu} = \sum_{\text{elementary } |\phi\rangle} \mathbf{m}(|\phi\rangle /n) |\phi\rangle \langle\phi|$, where the summation is over all n qubit elementary pure quantum states.

We say a semi-density matrix ρ is lower computable if there a program $p \in \Sigma^*$ such that when given to the universal Turing machine U, outputs, with or without halting, a finite or infinite sequence of elementary matrices ρ_i such that $\rho_i \preceq \rho_{i+1}$ and $\lim_{i \to \infty} \rho_i = \rho$. If U reads $\leq \|p\|$ bits on the input tape, then we say p lower computes ρ. From [Gac01] Theorem 2, if q lower computes ρ, then $\mathbf{m}(q/n)\rho \overset{*}{<} \boldsymbol{\mu}$.

We say a semi-density matrix $\rho \otimes \sigma$ is upper computable if there a program $p \in \Sigma^*$ such that when given to the universal Turing machine U, outputs, with or without halting, a finite or infinite sequence of elementary matrices $\rho_i \otimes \sigma_i$ such that $\rho_{i+1} \otimes \sigma_{i+1} \preceq \rho_i \otimes \sigma_i$ and $\lim_{i \to \infty} \rho_i \otimes \sigma_i = \rho \otimes \sigma$. If U reads $\leq \|p\|$ bits on the input tape, then we say p upper computes $\rho \otimes \sigma$. The upper probability of an upper computable mixed state $A \otimes B$ is defined by $\overline{\mathbf{m}}(A \otimes B/x) = \sum \{\mathbf{m}(q/x) : q \text{ upper computes } A \otimes B\}$.

4 Information

Let $\mathcal{C}_{C \otimes D}$ be the set of all upper computable matrices (tests) of the form $A \otimes B$, where $\text{Tr}(A \otimes B)(C \otimes D) \leq 1$. Let $\mathfrak{C}_{C \otimes D} = \sum_{A \otimes B \in \mathcal{C}_{C \otimes D}} \overline{\mathbf{m}}(A \otimes B/n)(A \otimes B)$ be an aggregation of upper computable $C \otimes D$ tests of the form $A \otimes B$, weighted by their upper probability. The information between semi-density matrices A and B is $\mathbf{I}(A : B) = \log \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}}(A \otimes B)$.

Proposition 1 $\mathbf{I}(2^{-n} : 2^{-n}) = O(1)$.

Proof. $1 \geq \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}}(\boldsymbol{\mu} \otimes \boldsymbol{\mu}) \overset{*}{>} \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}}(2^{-n}I \otimes 2^{-n}I) \overset{*}{>} 2^{\mathbf{I}(2^{-n}I : 2^{-n}I)}$.

$\square$

Lemma 1 *Let Λ be the uniform distribution on the unit sphere of an n qubit space.*
$\int 2^{\mathbf{I}(|\psi\rangle : |\psi\rangle)} d\Lambda = O(1)$.

Proof. Using [Gac01] Section 5 and [BvL01] Section 6.3, we have that $\int |\psi\rangle\langle\psi| \otimes |\psi\rangle \langle\psi| d\Lambda = \int |\psi\psi\rangle\langle\psi\psi| d\Lambda = \binom{2^n + 1}{2}^{-1} P$, where P is the projection onto the space of pure states $|\psi\psi\rangle$. Using Proposition 1,

$$\int 2^{\mathbf{I}(|\psi\rangle : |\psi\rangle)} d\Lambda = \int \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}} |\psi\rangle\langle\psi| \otimes |\psi\rangle \langle\psi| d\Lambda = \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}} \int |\psi\rangle\langle\psi| \otimes |\psi\rangle \langle\psi| d\Lambda$$

$$= \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}} \binom{2^n + 1}{2}^{-1} P \overset{*}{<} \text{Tr} \mathfrak{C}_{\boldsymbol{\mu} \otimes \boldsymbol{\mu}} 2^{-2n} I \overset{*}{=} 2^{\mathbf{I}(2^{-n}I : 2^{-n}I)} = O(1).$$

$\square$

5 Measurements

A POVM E is a finite set of positive definite matrices $\{E_k\}$ such that $\sum_k E_k = I$. For a given density matrix σ, a POVM E induces a probability measure over strings, where $E\sigma(k) = \mathrm{Tr} E_k \sigma$. This can be seen as the probability of seeing measurement k given quantum state σ and measurement E. An elementary POVM has each E_k being elementary. We introduce a novel definition to algorithmic information theory, the amount of algorithmic mutual information between two probabilities.

Definition 1 (Information, Probabilities)
For probabilities p and q over Σ^*, $\mathbf{i}(p:q) = \log \sum_{x,y \in \Sigma^*} 2^{\mathbf{i}(x\,:\,y)} p(x) q(y)$.

Theorem 2 in the appendix proves conservation of information of probabilities transformed by random channels. A channel $f : \Sigma^* \times \Sigma^* \to \mathbb{R}$, such that $f(\cdot|x)$ is a probability for all $x \in \Sigma^*$, transforms a probability p by $fp(x) = \sum_y f(x|y)p(y)$. Conservation occurs over f, $\mathbf{i}(fp:q) <^+ \mathbf{i}(p:q)$.

Lemma 2 *Relativized to POVM E, $\mathbf{i}(E\,|\psi\rangle\,{:}\,E\,|\psi\rangle) <^+ \mathbf{I}(|\psi\rangle\,{:}\,|\psi\rangle)$.*

Proof. Since $z(k) = \mathrm{Tr}\boldsymbol{\mu} E_k$ is lower semi-computable and $\sum_k z(k) < 1$, $\mathbf{m}(k/n) \overset{*}{>} \mathrm{Tr}\boldsymbol{\mu} E_k$, and so $1 > 2^{\mathbf{K}(k/n)-O(1)}\mathrm{Tr}\boldsymbol{\mu} E_k$. So $\nu_{i,j} = 2^{\mathbf{K}(i/n)+\mathbf{K}(j/n)-O(1)}(E_i \otimes E_j) \in \mathcal{C}_{\boldsymbol{\mu}\otimes\boldsymbol{\mu}}$, with $\overline{\mathbf{m}}(\nu_{i,j}/n) \overset{*}{>} \mathbf{m}(i,j/n)$.

$$
\begin{aligned}
\mathbf{I}(|\psi\rangle\,{:}\,|\psi\rangle) &= \log \sum_{A\otimes B \in \mathcal{C}_{\boldsymbol{\mu}\otimes\boldsymbol{\mu}}} \overline{\mathbf{m}}(A \otimes B/n)(A \otimes B)(|\psi\rangle\langle\psi|\,|\psi\rangle\langle\psi|) \\
&>^+ \log \mathrm{Tr} \sum_{ij} \nu_{i,j}\overline{\mathbf{m}}(\nu_{i,j}/n)(|\psi\rangle\langle\psi|\,|\psi\rangle\langle\psi|) \\
&>^+ \log \sum 2^{\mathbf{K}(i/n)+\mathbf{K}(j/n)}\mathbf{m}(i,j/n)E\,|\psi\rangle\,(i)E\,|\psi\rangle\,((j) \\
&>^+ \mathbf{i}(E\,|\psi\rangle\,{:}\,E\,|\psi\rangle).
\end{aligned}
$$

$\square$

Note that the number of qubits n is simple relative to E, thus the complexity terms in the proof are relativized to n. An implication of Lemma 2 is that most pure quantum states have almost no self classical information. That is for an overwhelming majority of pure quantum states, the probabilities induced by a measurement will have negligible self information, as shown in Theorem 1.

Theorem 1 *Let Λ be the uniform distribution on the unit sphere of an n qubit space. Relativized to POVM E, $\int 2^{\mathbf{i}(E|\psi\rangle:E|\psi\rangle)}d\Lambda = O(1)$.*

Proof. By Lemma 2, $2^{\mathbf{I}(|\psi\rangle:|\psi\rangle)} \overset{*}{>} 2^{\mathbf{i}(E|\psi\rangle:E|\psi\rangle)}$. From Lemma 1, $\int 2^{\mathbf{I}(|\psi\rangle:|\psi\rangle)}d\Lambda = O(1)$. The integral $\int 2^{\mathbf{i}(E|\psi\rangle:E|\psi\rangle)}d\Lambda$ is well defined because $2^{\mathbf{i}(E|\psi\rangle:E|\psi\rangle)} = \mathrm{Tr}\sum_{i,j}\nu_{i,j}2^{-\mathbf{K}(i,j)}(|\psi\rangle\langle\psi| \otimes |\psi\rangle\langle\psi|)$, which can be integrated over Λ. $\square$

Theorem 1 says that given a measurement apparatus, the overwhelming majority of pure states, when measured, will produce classical probabilities with no self-information, i.e. random noise. Theorem 2 shows that there is no randomized way to process the probabilities to produce more self-information, i.e. process the random noise.

A Conservation of Information Between Probabilities

In this section, we show that the information between probabilities cannot be increased through randomized transformations. We recall that information $\mathbf{i}$ was introduced in Section 5, where for probabilities p and q, $\mathbf{i}(p:q) = \sum_{x,y} 2^{\mathbf{i}(x:y)} p(x) q(y)$. The information between strings $x, y \in \Sigma^*$ is $\mathbf{i}(x:y) = \mathbf{K}(x) + \mathbf{K}(y) - \mathbf{K}(x,y)$. A probability p is transformed by channel $f : \Sigma^* \times \Sigma^* \to \mathbb{R}_{\geq 0}$, by $fp(x) = \sum_y f(x|y) p(y)$. For channel f, $f(\cdot|y)$ is a conditional probability given $y \in \Sigma^*$.

Lemma 3 ([Lev84]) *For $x, y, z \in \Sigma^*$, $\mathbf{i}(x:y) <^+ \mathbf{i}((x,z):y)$.*

Lemma 4 *Let ψ_d be a semi-measure, lower computable relative to $d \in \Sigma^*$. For $a, b \in \Sigma^*$*
$\sum_{c \in \Sigma^*} 2^{\mathbf{i}((a,c):b)} \psi_a(c) \overset{*}{<} 2^{\mathbf{i}(a:b)}$.

Proof. This requires a slight modification of the proof of Proposition 2 in [Lev84]. We need to show $\mathbf{m}(a,b)/(\mathbf{m}(a)\mathbf{m}(b)) \overset{*}{>} \sum_c (\mathbf{m}(a,b,c)/(\mathbf{m}(b)\mathbf{m}(a,c)))\psi_a(c)$, or $\sum_c (\mathbf{m}(a,b,c)/\mathbf{m}(a,c))\mathbf{m}(c|a) \overset{*}{<}$ $\mathbf{m}(a,b)/\mathbf{m}(a)$, since $\mathbf{m}(c|a) \overset{*}{>} \psi_a(c)$. Rewrite it $\sum_c \mathbf{m}(c|a)\mathbf{m}(a,b,c)/\mathbf{m}(a,c) \overset{*}{<} \mathbf{m}(a,b)/\mathbf{m}(a)$ or $\sum_c \mathbf{m}(c|a)\mathbf{m}(a)\mathbf{m}(a,b,c)/\mathbf{m}(a,c) \overset{*}{<} \mathbf{m}(a,b)$. The latter is true since $\mathbf{m}(c|a)\mathbf{m}(a) \overset{*}{<} \mathbf{m}(a,c)$ and $\sum_c \mathbf{m}(a,b,c) \overset{*}{<} \mathbf{m}(a,b)$. $\square$

Theorem 2 *For probabilities p and q, relativized to channel f, $\mathbf{i}(fp:q) <^+ \mathbf{i}(p:q)$.*

Proof. Using Lemma 3,

$$\mathbf{i}(fp:q) = \log \sum_{x,y} 2^{\mathbf{i}(x:y)} \sum_z f(x|z) p(z) q(y) <^+ \log \sum_{y,z} q(y) p(z) \sum_x 2^{\mathbf{i}((z,x):y)} f(x|z).$$

Using Lemma 4,

$$\mathbf{i}(fp:q) <^+ \log \sum_{z,y} q(y) p(z) 2^{\mathbf{i}(z:y)} =^+ \mathbf{i}(p:q).$$

References

[BKM⁺06] F. Benatti, T. Krüger, M. Müller, R. Siegmund-Schultze, and A. Szkola. Entropy and Quantum Kolmogorov Complexity: A Quantum Brudno's Theorem. *Communications in Mathematical Physics*, 265(2), 2006.

[BOD14] F. Benatti, S. K. Oskouei, and A. S. Deh Abad. Gacs Quantum Algorithmic Entropy in Infinite Dimensional Hilbert Spaces. *Journal of Mathematical Physics*, 55(8), 2014.

[BvL01] A. Berthiaume, W. van Dam, and S. Laplante. Quantum Kolmogorov Complexity. *Journal of Computer and System Sciences*, 63(2), 2001.

[DH10] R. G. Downey and D.R. Hirschfeldt. *Algorithmic Randomness and Complexity*. Theory and Applications of Computability. Springer New York, 2010.

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

[GÍ3] P. Gács. Lecture notes on descriptional complexity and randomness, 2013.

[Gac01] P. Gacs. Quantum algorithmic entropy. *Journal of Physics A Mathematical General*, 34(35), 2001.

[Hol73] A. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9, 1973.

[Lev84] L. A. Levin. Randomness conservation inequalities; information and independence in mathematical theories. *Information and Control*, 61(1):15–37, 1984.

[LV08] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer Publishing Company, Incorporated, 3 edition, 2008.

[MB05] C. Mora and H. Briegel. Algorithmic Complexity and Entanglement of Quantum States. *Phys. Rev. Lett.*, 95, 2005.

[NS19] A. Nies and V. Scholz. Martin-löf random quantum states. *Journal of Mathematical Physics*, 60(9), 2019.

[Sch95] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51, 1995.

[Utt12] S. Uttam. Chapter 4 - Introduction to Quantum Information Processing. In I. Djordjevic, editor, *Quantum Information Processing and Quantum Error Correction*. 2012.

[Vit00] P. Vitanyi. Three Approaches to the Quantitative Definition of Information in an Individual Pure Quantum State. In *Proceedings 15th Annual IEEE Conference on Computational Complexity*, 2000.