Quantum Sampling and Entropic Uncertainty

Walter O. Krawec
Computer Science & Engineering Department
University of Connecticut
Storrs, CT 06268
Email: walter.krawec@uconn.edu

Abstract—In this paper, we show an interesting connection between a quantum sampling technique and quantum uncertainty. Namely, we use the quantum sampling technique, introduced by Bouman and Fehr, to derive a simple proof of a version of the Maassen and Uffink uncertainty relation. We also use quantum sampling to prove a finite, non-i.i.d., entropic uncertainty relation based on smooth min entropy.

I. INTRODUCTION

In this paper, we revisit a famous entropic uncertainty relation proven by Maassen and Uffink in [1] (which followed a conjecture by Kraus in [2] and was also an improvement of an entropic uncertainty relation first proposed by Deutsch [3]). Given a quantum system $\rho$ and two projective measurements (PMs) $\{M_x\}$ and $\{N_y\}$ (where $M_x = |\mu_x\rangle \langle \mu_x|$ and $N_y = |\nu_y\rangle \langle \nu_y|$ for some orthonormal bases $\{|\mu_x\rangle\}$ and $\{|\nu_y\rangle\}$), then one cannot necessarily be certain of the outcome of both measurements. More specifically, the relation states:

$$H(M) + H(N) \geq - \log_2 c,$$  \hspace{1cm} (1)

where $c$ is a function of the two measurements, namely:

$$c = \max_{x,y} \langle \mu_x | \nu_y \rangle^2.$$  \hspace{1cm} (2)

This relation, and numerous others like it, are not only interesting in and of themselves, but also have numerous applications throughout quantum information science and quantum cryptography. For a general survey of entropic uncertainty relations, the reader is referred to [4], [5], [6].

In this paper, we derive a novel, and in our opinion simpler, proof of Equation 1 for projective measurements over two-dimensional systems using a quantum sampling technique first introduced in [7]. This sampling technique has been applied to the study of quantum cryptography, however to our knowledge, it has not seen application to more broad areas of quantum information. We show how this technique can be used to prove Equation 1 in this case. Along the way, we also derive a new entropic uncertain bound for smooth min entropy with a direct connection to sampling strategies and for states which are not necessarily i.i.d.

There are several contributions in this work. First, we show a rather interesting connection between quantum sampling and quantum uncertainty and use this to derive a much simpler proof of a particular case of the above bound. We discover a rather fascinating connection between quantum sampling and information theory and derive a novel entropic uncertainty bound (involving smooth min entropy and applicable to non-i.i.d. states) directly related to sampling strategies. The techniques we use in this paper may find application to other areas of quantum information science and may eventually lead to better bounds for quantum cryptography in the finite key setting.

Notation and Definitions: Let $\mathcal{A}$ be a finite alphabet of size $d$. Then if $q \in \mathcal{A}^n$ and $\tau = \{\tau_1, \ldots, \tau_n\} \subset \{1, \ldots, n\}$, we write $q_\tau$ to mean the sub-string of $q$ indexed by $\tau$, namely $q_\tau = (q_{\tau_1}, \ldots, q_{\tau_n})$. We write $q_{-\tau}$ to mean the sub-string of $q$ indexed by the complement of $\tau$.

If $\mathcal{A} = \{0, 1\}$, the Hamming weight of the string $q$ is defined to be the number of non-zero elements in $q$. For arbitrary $\mathcal{A}$ and for any $a \in \mathcal{A}$, we define the relative $a$-Hamming weight, which we denote by $w_a(q)$, to be the number of letters in $q$ not equal to $a$ and that quantity divided by the length of $q$. Namely: $w_a(q) = |\{i \mid q_i \neq a\}|/|q|$, where $|q|$ denotes the length of the string $q$.

A density operator acting on Hilbert space $\mathcal{H}$ is a Hermitian positive semi-definite operator of unit trace. Given $|\psi\rangle \in \mathcal{H}$ we write $|\psi\rangle$ to mean $|\psi\rangle \langle \psi|$. We define a Projective Measurement or PM over a $d$-dimensional Hilbert space $\mathcal{H}$ to be a set of projectors $N = \{\phi_1, \ldots, \phi_d\}$, where $\{|\phi_i\rangle\}_{i=1}^d$ form an orthonormal basis of $\mathcal{H}$. It is not difficult to see that we may treat a measurement outcome of $|\phi_{j_1}\rangle \otimes \cdots \otimes |\phi_{j_n}\rangle$ as the classical string $j = j_1 \cdots j_n$. We often write $\mathcal{H}_A$ to mean a $d$-dimensional Hilbert space.

We denote $H(X)$ to be the Shannon entropy of random variable $X$. If $\rho$ is a density operator acting on Hilbert space $\mathcal{H}$ and if $N$ is a PM over $\mathcal{H}$, we write $H(N)_\rho$ to mean the Shannon entropy of the random variable induced by measuring $\rho$ using PM $N$. Similarly, if $|\psi\rangle$ is a pure state in $\mathcal{H}$ we write $H(N)_\psi$ to mean the entropy of the result of measuring $|\psi\rangle$ using PM $N$. For technical reasons later, we define an extended binary entropy function, denoted $H(x)$ which is defined to be $H(x, 1-x)$ if $x \in [0, 1/2]$; otherwise, if $x < 0$, $H(x) = 0$ and if $x > 1/2$, then $H(x) = 1$.

Given a density operator $\rho_{A_E}$, the conditional quantum min entropy [8], denoted $H_\infty(A|E)_\rho$, is defined to be $H_\infty(A|E)_\rho = \sup_{\sigma_E} \max \{\lambda \in \mathbb{R} \mid 2^{-\lambda} I_A \otimes \sigma_E - \rho_{A_E} \geq 0\}$. If the $E$ system is trivial then we may write $H_\infty(A)_\rho$ in which case we have $H_\infty(A)_\rho = -\log \lambda_{\max}(\rho)$, where $\lambda_{\max}(\rho)$ is the maximal eigenvalue of $\rho$ (note that all logarithms in this paper are base 2 unless otherwise stated). If $N$ is a PM on $\mathcal{H}$ and $\rho$ is a density operator on $\mathcal{H} \otimes_{\sigma}$, then we use $H_\infty(N)_\rho$ to mean the min entropy of the resulting state following the measurement...
of each of the $n$ sub-spaces $\rho$ acts on using PM $N$. If $p(j)$ is the probability of observing outcome $j = j_1 \cdots j_n$ (i.e., after measuring, one observes the quantum state $|\phi_{j_1}\rangle \otimes \cdots \otimes |\phi_{j_n}\rangle$) it is not difficult to see that: $H_\infty(N)_\rho = -\log \max_j p(j)$.

Given a density operator $\rho_{ABC}$ acting on $H_A \otimes H_B \otimes H_C$, where the $C$ portion is classical (namely, we may write $\rho_{ABC} = \sum_{c} \sigma_{c}^{AB} \otimes |c\rangle$, where $\{|c\rangle\}$ is an orthonormal basis of $H_C$ and each $\sigma_{c}^{AB}$ is an arbitrary, though not normalized density operator acting on $H_A \otimes H_B$) then the conditional min entropy $H_\infty(A|BC)_\rho$ is:

$$H_\infty(A|BC)_\rho \geq \inf_c H_\infty(A|B)_{\sigma(c)},$$

(this also applies if the $B$ portion is trivial). The above can be proven from Lemma 3.1.8 in [8] and the definition of conditional min entropy.

Finally, the $\epsilon$-smooth min entropy, denoted $H^\epsilon_\infty(\rho)$ is defined to be: $H^\epsilon_\infty(\rho) = \sup_{\sigma \in \Gamma_\epsilon(\rho)} H_\infty(\sigma)$, where $\Gamma_\epsilon(\rho)$ is the set of all density operators $\epsilon$ close to $\rho$ as measured by the trace distance; i.e., $\Gamma_\epsilon(\rho) = \{\sigma \mid \|\sigma - \rho\| \leq \epsilon\}$, and $||A||$ is the trace distance of $A$. We define $H^\epsilon_\infty(N)_\rho$, similarly to $H_\infty(N)_\rho$, described earlier whenever $N$ is a PM. The conditional smooth min entropy, $H^\epsilon_\infty(A|B)$ is defined similarly.

An important result, which we will use later, was proven in [7] (based on a Lemma in [8]) and allows one to compute the min entropy of a superposition of states:

**Lemma 1.** (From [7]): Let $H$ be a $d$-dimensional Hilbert space with orthonormal basis $\{|i\rangle\}_{i=1}^d$ and let $H_E$ be an arbitrary finite dimensional Hilbert space. Then, for any pure state $|\psi\rangle = \sum_{i,j} \alpha_i |i\rangle \otimes |\phi_i\rangle_E \in H \otimes H_E$, if we define:

$$\rho = \sum_{i,j} |\alpha_i|^2 |i\rangle \langle i| \otimes |\phi_i\rangle_E,$$

it holds that for any PM $N$ on $H$:

$$H_\infty(N|E)_\psi \geq H_\infty(N|E)_\rho - \log_2 |J|.$$  

The above lemma will allow us to bound the min entropy of a superposition of states, by computing, instead, the min entropy in a suitable mixed state.

**Quantum Sampling:** Since our proof relies on the quantum sampling technique introduced in [7], we now review this subject here. All information in this subsection is derived from [7] (we make only a few changes in notation and some generality) and is meant only as a review of this material for completeness.

Let $A$ be a finite alphabet of size $d$, and let $a \in A$, and $k \in \mathbb{N}$. We assume $d$, $a$, and $k$ are arbitrary, but fixed. A sampling strategy is a pair $\Sigma = (P^k_F, F^k_a)$ where $P^k_F$ is a distribution over all subsets of $\{1, \cdots , n\}$ of size $k$ and $F^k_a$ is a function which, given a subset of a sample $q \in A^n$ (i.e., given $q_r$), will output a guess of the value $w_a(q_r)$. That is, given a randomly chosen sample $q_r$ (where $\tau$ was drawn according to $P^k_F$, $F^k_a$ will estimate the value of $w_a$ in the remaining portion of $q$. When it is clear, we will often forgo writing the superscript, and simply write $F_a$.

Define $B^k_{\tau,a}(\Sigma)$ to be the set of all words in $A^n$ such that the estimate provided by $F_a$ is $\delta$ close to the actual value given a fixed subset $\tau \subset \{1, \cdots , n\}$ of size $k$. That is, let:

$$B^k_{\tau,a}(\Sigma) = \{q \in A^n \mid |F_a(q_r) - w_a(q_r)| \leq \delta\}.$$  

Informally, if we have a fixed subset $\tau$ with $|\tau| = k$, then the set $B^k_{\tau,a}(\Sigma)$ defines the set of all “good” strings; i.e., strings for which the sampling strategy $\Sigma$ provides an accurate estimate of $w_a$, up to an error of $\delta$ assuming $\tau$ was the chosen subset.

From this, the error probability of $\Sigma$ is defined to be:

$$\epsilon^c_\delta = \max_{q \in A^n} Pr(q \notin B^k_{\tau,a}(\Sigma)),$$

where the probability is over all subsets $\tau$ chosen according to $P^k_F$ (i.e., we treat $B^k_{\tau,a}$ as a random variable induced by choosing subsets $\tau$ according to $P^k_F$). From this definition, it is clear that for any word $q \in A^n$, the estimated value of $w_a$, given by the sampling strategy $\Sigma$, is $\delta$ close to the real value in the remainder of the string (i.e., in the portion of the string that was not used in the test set $\tau$), except with probability $\epsilon^c_\delta$. Note the superscript “$c$” is used to show this is the error probability of a classical sampling strategy.

One important sampling strategy we will make use of is the following: Let $P^k_{\text{uni}}$ be the uniform distribution over all subsets $\tau \subset \{1, \cdots , N\}$ with $|\tau| = k$; i.e., $Pr(P^k_{\text{uni}}) = 1/\binom{N}{k}$.

Then, given a string $q \in A^N$, the function $F$ is defined simply to be: $F_a(q_r) = w_a(q_r)$. That is, the sampling strategy is to choose a random subset, uniformly at random, evaluate $w_a$ on that subset, and output, as an estimate of the value $w_a(q_r)$, the value $w_a(q_r)$. The following Lemma was proven in [7] (see Appendix B in the extended, online version, of that reference):

**Lemma 2.** (From [7]): Let $\delta > 0$ be given and $\Sigma$ be as described above in the text. If $|\tau| = k \leq N/2$ then for any $d$ and $a$, it holds that:

$$\epsilon^c_\delta \leq 2 \exp\left(-\frac{\delta^2 k N}{N + 2}\right).$$

These notions can be extended to the quantum domain [7]. Consider an orthonormal basis $\{|a\rangle \mid a \in A\}$ and let $H_A$ be the $d$-dimensional Hilbert space spanned by this basis. Let $U$ be a unitary operator acting on $H_A$. Then, we may define an orthonormal basis:

$$B = \{U^{\otimes n} |b_1 \cdots b_n\rangle \mid U |b_1\rangle \otimes \cdots \otimes U |b_n\rangle \mid b_i \in A\},$$

of the Hilbert space $H_A^{\otimes n}$. Then, given a state $|\psi\rangle \in H_A^{\otimes n} \otimes H_E$, it is said to have relative $a$-Hamming weight $\beta$ in $A$ with respect to basis $B$, if we can write $|\psi\rangle = U^{\otimes n} |b_1 \cdots b_n\rangle \otimes |\phi\rangle_E$ with $w_a(b) = \beta$. Note that we are allowed an additional, arbitrary, system in some Hilbert space $H_E$ (this may be the trivial space if it is not needed). Also, notice that this definition is dependent on the choice of basis.

By abusing notation slightly, we may also define span$(B^k_{\tau,a})$ to be:

$$\text{span} \left( \{U^{\otimes n} |q\rangle \mid q \in A^n \text{ and } |w_a(q_r) - w_a(q_{r-})| \leq \delta \} \right)$$
Note that if \( |\psi\rangle \in \text{span}(B_{\text{ref}}^m) \otimes \mathcal{H}_E \) then, if sampling is done by measuring in the \( B \) basis on subset \( \tau \), it is guaranteed that the state collapses to a superposition of states which are \( \rho \)-close to the observed \( \alpha \)-Hamming weight (with respect to basis \( B \)). Also note we will drop the \( \delta \) superscript when the context is clear.

Using the above results, the main result from [7] is as follows:

**Theorem 1.** (From [7], though reworded for our application in this paper and our specific sampling strategy): Let \( k \leq n/2 \) be given and consider sampling strategy \( \Sigma \) as described above. Then, for every pure state \( |\psi\rangle \in \mathcal{H}_A^{2^{n+m}} \otimes \mathcal{H}_E \), there exists a collection of “ideal states” \( \{ |\phi_i\rangle \} \) where the index is over all subsets \( \tau \) of size \( k \) and each \( |\phi_i\rangle \in \text{span}(B_{\text{ref}}^m) \otimes \mathcal{H}_E \) such that: 
\[
\left| \frac{1}{T} \sum_{\tau} |\tau\rangle \otimes |\phi_i\rangle - \frac{1}{T} \sum_{\tau} |\tau\rangle \otimes |\phi_i\rangle \right| \leq \sqrt{\epsilon^2},
\]
where \( T = \binom{n}{k} \) and the sum is over all subsets of size \( k \).

The above result states that, on average over the choice of subset \( \tau \), the real system \( |\psi\rangle \) is \( \epsilon \)-close to an ideal state, where the ideal state is defined to be one where the sampling strategy always works (i.e., where, after sampling, regardless of the subset choice, the state collapses to one which is a superposition of states \( \rho \)-close to the estimate). The error probability \( \epsilon \) can be computed from the classical error probability.

**II. MAIN RESULT**

We are now in a position to show how quantum sampling can be used to derive a simple proof of the quantum uncertainty relation shown in Equation 1. We first prove the following intermediate result which may be of independent interest which bounds the smooth min entropy of a potentially non-i.i.d. state, in terms of the successful operation of the quantum sampling strategy.

**Theorem 2.** Let \( \epsilon \geq \epsilon > 0 \), \( a \in \{0,1\} \), \( 0 < \beta < 1/2 \), and \( \rho \) a density operator acting on Hilbert space \( \mathcal{H}_A^{2^{n+m}} \otimes \mathcal{H}_E \) with \( m \leq n \). Also, let \( M = \{|\mu_0\rangle,|\mu_1\rangle\} \) and \( N = \{|v_0\rangle,|v_1\rangle\} \), be two projective measurements. If a subset \( t \) of size \( m \) of \( \rho \) is measured using \( M \) resulting in outcome \( q \) we denote by \( \rho(t,q) \) to be the post-measurement state (this is well defined given \( \rho \)). Let \( \epsilon' = 2e + 2\epsilon \beta \) and \( \epsilon'' = \epsilon' - 2\beta \), then it holds that:
\[
Pr \left[ H_\infty(N)_{\rho(t,q)} + nH(w_a(q) + \delta) \geq -n \log c \right] \geq 1 - \epsilon''
\]
where the probability is over all choice of subsets and resulting measurement outcomes. Above, \( c \) is defined in Equation 2 and:
\[
\delta = \sqrt{\frac{(n+m+2) \ln(2/\epsilon')}{m(n+m)}}.
\]  

**Proof.** We first consider the case when \( \rho \) is pure; that is, \( \rho = |\psi\rangle \langle \psi| \) for some \( |\psi\rangle \in \mathcal{H}_A^{2^{n+m}} \). Then, applying Theorem 1 to \( \rho \), using the sampling strategy described in the previous section for a sample subset size of \( m \), it follows that there exists an “ideal” state \( \sigma \) of the form: 
\[
\sigma = \frac{1}{T} \sum_{\tau} |\tau\rangle \otimes |\phi_i\rangle,
\]
where \( T \) is the number of possible subsets (i.e., \( T = \binom{n+m}{m} \)); the summation is over all possible subsets \( t \) of \( \{1, \ldots, n + m\} \)
which are of size \( m \) (we expand the underlying Hilbert space to include this auxiliary subspace \( \mathcal{H}_E \) spanned by orthonormal basis \( \{|t\} \mid t \in \{1, \ldots, n + m\}, |t| = m \}; and, finally, each \( |\phi_i\rangle \in \text{span}(B_{\text{ref}}^m) \)). This ideal state satisfies the following:
\[
||\sigma - \frac{1}{T} \sum_{\tau} |\tau\rangle \otimes |\phi_i\rangle|| \leq \sqrt{\epsilon^2}.
\]
Given \( \delta \) as in Equation 7, and also given Lemma 2, it holds that \( \sqrt{\epsilon^2} = \epsilon \).

Consider the following experiment: First, run the sampling strategy, choosing a random subset \( t \) (which is chosen by measuring the auxiliary \( \mathcal{H}_E \) subspace) and performing a measurement in the \( M \) basis resulting in outcome \( q \) (note that \( q \) depends on the subset chosen and the intrinsic randomness of the measurement itself). Let \( \rho(t,q) \) be the post-measurement state if this experiment is performed on the true state \( \rho = |\psi\rangle \). Likewise, let \( \sigma(t,q) \) be the post measurement state if this experiment is performed on the ideal state \( \sigma \). Both post-measurement states are well defined given both \( t \) and \( q \) (though, of course, the post-measurement state may be a superposition, they are, however, exactly defined pure states, conditioning on the outcome of \( t \) and \( q \)).

We first show:
\[
H_\infty(N)_{\sigma(t,q)} \geq -n \log c - nH(w_a(q) + \delta). \tag{8}
\]
That is, with certainty, for any subset \( t \) and observed value \( q \), Equation 8 holds in the ideal case.

Let \( t \) be the chosen subset, thus the measurement in basis \( M \) is performed on the pure state \( |\phi_i\rangle \). Since \( |\phi_i\rangle \in \text{span}(B_{\text{ref}}^m) \), it follows that the post measurement state, after observing value \( q \), collapses to a superposition of the form:
\[
|\phi_i\rangle = \sum_{i \in J} \alpha_i |\mu_{i_1}, \ldots, \mu_{i_m}\rangle,
\]
where \( J \subset I = \{i \in \{0,1\}^n \mid |w_a(i) - w_a(q)| \leq \delta \} \) and normalization requires \( \sum_{i \in J} |\alpha_i|^2 = 1 \). Of course \( \sigma(t,q) = |\phi_i\rangle \).

Now, consider the mixed state:
\[
\chi = \sum_{i \in J} |\alpha_i|^2 |\mu_{i_1}, \ldots, \mu_{i_m}\rangle.
\]

By applying Lemma 1, we have:
\[
H_\infty(N)_{\chi} = H_\infty(N)_{\chi} \geq H_\infty(N)_{\chi} - \log |J|. \tag{10}
\]
We now compute \( H_\infty(N)_{\chi} \). Let \( \chi_N \) be the result of measuring \( \chi \) using PM \( N \). It is not difficult to see that this state is simply:
\[
\chi_N = \sum_{i \in J} |\alpha_i|^2 \left( \sum_{j \in \{0,1\}^n} p(j|i) |\nu_{j_1}, \ldots, \nu_{j_m}\rangle \right) = \sum_{j \in \{0,1\}^n} p(j) |\nu_{j_1}, \ldots, \nu_{j_m}\rangle,
\]
where we define \( p(j|i) = p(j_1 \cdots j_n |i_1 \cdots i_n) \) to be the probability of observing \( |\nu_{j_1}, \ldots, \nu_{j_m}\rangle \) if given an input state of \( |\mu_{i_1}, \ldots, \mu_{i_m}\rangle \). We define \( p(j) = \sum_{i \in \{0,1\}^n} |\alpha_i|^2 p(j|i) \). It is straightforward to compute \( p(j|i) \):
\[
p(j|i) = p(j_1 \cdots j_n |i_1 \cdots i_n) = \prod_{l=1}^{n} \langle \nu_{j_l}|\mu_{i_l}\rangle^2 \tag{11}
\]
Since $\chi_N$ is a classical system, we have:

$$H_\infty(N)_\chi = -\log \max_j p(j) = -\log \max_j \left( \sum_{i \in J} |\alpha_i|^2 p(j|i) \right).$$

Let $p^* = \max_{j \in J} p(j|i)$ (where the maximum is over all $i \in J$ and $j \in \{0,1\}^n$). Then it is clear that:

$$\max_j p(j) = \max_j \left[ \sum_{i \in J} |\alpha_i|^2 p(j|i) \right] \leq p^*,
$$

(recall that $\sum_i |\alpha_i|^2 = 1$) and thus:

$$H_\infty(N)_\chi = -\log \max_j p(j) \geq -\log p^*.$$

Finally, we compute a bound on $p^*$ as:

$$p^* = \max_{j \in \{0,1\}^n} \prod_{l=1}^n \left| \langle \nu_j_l | \mu_{ji} \rangle \right|^2 \leq e^n,$$

where $c = \max_{x,y} |\langle \nu_x | \mu_{yx} \rangle|^2$. Thus:

$$H_\infty(N)_\chi \geq -\log p^* \geq -n \log c. \tag{12}$$

It is clear that $J \subset \{i \in \{0,1\}^n \mid w_a(i) \leq w_a(q) + \delta \}$ and so using the well-known bound on the volume of a Hamming ball we have $|J| \leq 2^n H(w_a(q) + \delta)$ (note we are using our “extended” version $H$ here to avoid the issue when $w_a(q) + \delta > 1/2$; indeed, if that is the case then $H(\cdot) = 1$ and so the bound holds trivially), we may combine this with Equations 10 and 12 to derive:

$$H_\infty(N)_{\sigma(t,q)} \geq -n \log c - nH(w_a(q) + \delta).$$

Of course, the above analysis only considered the ideal state from which we are guaranteed that the sampling strategy was successful. We now consider the “real” state $\rho = |\psi\rangle$.

Consider the real state $\frac{1}{T} \sum_t |t\rangle \otimes |\psi\rangle$. The process of choosing a subset $t$, measuring, and observing $q$ (resulting in post-measurement state $\rho(t,q)$) may be described, entirely, by the mixed state: $\rho_{TQR} = \frac{1}{T} \sum_t |t\rangle \langle t| \sum_q \rho(q|t) |q\rangle \otimes \rho(t,q)$, where $p(q|t)$ is the probability of observing outcome $q$ given subset $t$ was sampled; here we use “$R$” to denote the “remainder” - that is the portion of the state not yet measured. Likewise, the ideal state, after performing this experiment, may be written as the mixed state: $\sigma_{TQR} = \frac{1}{T} \sum_t |t\rangle \langle t| \sum_q \bar{\rho}(q|t) |q\rangle \otimes \sigma(t,q)$. Since quantum operations cannot increase trace distance, we have $||\rho_{TQR} - \sigma_{TQR}|| \leq \epsilon$. By basic properties of trace distance:

$$\epsilon \geq \frac{1}{T} \sum_t \sum_q ||p(q|t) \rho(t,q) - \bar{p}(q|t) \sigma(t,q)||. \tag{13}$$

Of course, it holds that $\frac{1}{T} \sum_t \sum_q |p(q|t) - \bar{p}(q|t)| \leq \epsilon$ (this follows by tracing out the unmeasured portion “$R$” of $\rho_{TQR}$ and $\sigma_{TQR}$ and again realizing that quantum operations, such as partial trace, do not increase trace distance). Let $\bar{p}(q|t) = p(q|t) + \epsilon_q.t$ where $\epsilon_q.t \in [-\epsilon, \epsilon]$. Then, the above inequality of course implies $\frac{1}{T} \sum_t \sum_q |\epsilon_q.t| \leq \epsilon$.

Returning to Equation 13 we then find:

$$\epsilon \geq \frac{1}{T} \sum_t \sum_q ||p(q|t)(\rho(t,q) - \sigma(t,q)) - \epsilon_q.t \sigma(t,q)||
\geq \frac{1}{T} \sum_t \sum_q |p(q \land t)| \Delta_{q,t} - \epsilon,
\tag{14}$$

where we define $\Delta_{q,t} = \frac{1}{2} ||\rho(t,q) - \sigma(t,q)|| \leq 1$. Note that, above, we made use of the reverse triangle inequality and the fact that $||\sigma(t,q)|| = tr\sigma(t,q) = 1$ since $\sigma(t,q)$ is a positive operator of unit trace. We also used the fact that $p(q \land t) = p(q|t)p(t) = p(q|t) \cdot \frac{1}{T}$ (here, $p(q \land t)$ is the probability of sampling subset $t$ and observing $q$). Of course, the above implies: $\sum_q p(q \land t)\Delta_{q,t} \leq \epsilon$.

Now, let us consider $\Delta_{q,t}$ as a random variable over the choice of all subsets $t$ and measurement outcomes on that subset $q$. The expected value is easily seen to be $\mathbb{E}(\Delta_{q,t}) = \mu \leq \epsilon$. We also compute the variance $V^2$:

$$V^2 = \sum_q \sum_{q,t} p(q \land t)\Delta_{q,t}^2 - \mu^2 \leq \sum_q \sum_{q,t} p(q \land t)\Delta_{q,t} - \mu^2 = \mu(1 - \mu) \leq \mu \leq \epsilon,$$

where, above, we used the fact that $\Delta_{q,t} \leq 1$ and so $\Delta_{q,t}^2 \leq \Delta_{q,t}$.

Now, by Chebyshev’s inequality, we have:

$$Pr \left( |\Delta_{q,t} - \mu| \geq \epsilon^2 \right) \leq \frac{V^2}{\epsilon^2} \leq 1 - \epsilon^2 \leq 1 - 2\beta,$$

(the last inequality follows since $\beta < 1/2$; note that this probability is over all subsets $t$ and measurement outcomes $q$). Thus, except with probability at most $1 - 2\beta$, after choosing $t$ and observing $q$, it holds that $|\Delta_{q,t} - \mu| \leq \epsilon^2$ which, of course, implies:

$$\frac{1}{2} ||\rho(t,q) - \sigma(t,q)|| = \Delta_{q,t} \leq \mu + \epsilon^2 \leq \epsilon + \epsilon^2.$$

Since, in this case we have $\sigma(t,q) \in \Gamma_{2\epsilon + 2\epsilon^2}(\rho(t,q))$, it holds:

$$H_{\infty}^{2\epsilon + 2\epsilon^2}(N)_{\rho(t,q)} \geq H_{\infty}(N)_{\sigma(t,q)} \geq -n \log c - H(w_a(q) + \delta),$$

completing the proof when the case $\rho$ is pure.

Now consider the case when $\rho$ is not pure. In this case, let $|\psi\rangle_{HC}$ be a purification of $\rho$, where the $H$ portion is the original $H_{(2^n+2n)}$ space and the $C$ portion lives in an extra Hilbert space ($\mathcal{H}_C$) needed to purify $\rho$. As before, using quantum sampling, there exists an ideal state $\sigma$ where, now, each of the $|\phi^i\rangle \in \text{span}(B_{2^{n+1}}) \otimes \mathcal{H}_C$.

Let us consider running the same experiment as before on this ideal state (where, now, the experiment consists only of measuring the $H$ portion, not the $C$ portion). Let $t$ be the chosen subset and $q$ the observed value. Then, in the ideal case, the state collapses to a pure state of the form:

$$|\phi^i\rangle_{HC} = \sum_{i \in J} \alpha_i |\mu_{i_1}, \cdots, \mu_{i_n}\rangle \otimes |C_i\rangle,$$

where $J$ is defined as before and the states $|C_i\rangle$ are arbitrary (not necessarily normalized, nor orthogonal) states in $\mathcal{H}_C$. Let
\[ \chi_{HC} = \sum_{i \in J} |\alpha_i|^2 [\mu_1, \ldots, \mu_n] \otimes [C_i]. \]
From Lemma 1, we have:
\[ H_\infty(N|C)_{\psi} \geq H_\infty(N|C)_X - \log |J|. \]
We add an additional system \( I \) spanned by orthonormal basis \( \{I_i\}_{i \in J} \) and define the following state:
\[ \chi_{HCI} = \sum_{i \in J} |\alpha_i|^2 [\mu_1, \ldots, \mu_n] \otimes [C_i] \otimes [I_i] \]
Measuring this state using PM \( N \) yields:
\[ \chi_{NCI} = \sum_{i \in J} |\alpha_i|^2 [I_i] \otimes [C_i] \otimes \sum_{j \in \{0,1\}^n} p(j|i) |\nu_{j_1}, \ldots, \nu_{j_m}|, \]
where \( p(j|i) \) is defined as before in Equation 11 (also, note that we permuted the ordering of the sub-spaces above only for clarity). Define the sub-normalized states \( \chi_{NC,i} \) as:
\[ \chi_{NC,i} = |\alpha_i|^2 [C_i] \otimes \sum_{j \in \{0,1\}^n} p(j|i) |\nu_{j_1}, \ldots, \nu_{j_m}|, \]
from which we may write
\[ \chi_{NCI} = \sum_{i \in J} [I_i] \otimes \chi_{NC,i}. \]
From Equation 3, we have:
\[ H_\infty(N|CI)_X \geq \inf_{i \in J} H_\infty(N|C)_{\chi_{NC,i}}, \]
where \( H_\infty(N|C)_{\chi_{NC,i}} \) is taken to mean the min entropy in the sub-normalized state \( \chi_{NC,i} \). Note that this is a product state with the \( C \) part independent of \( N \); thus the min entropy is simply the min entropy in the \( N \) portion: \( H_\infty(N|C)_{\chi_{NC,i}} = H_\infty(N|C)_{X,i} \). Thus:
\[ H_\infty(N|CI)_X \geq \inf_{i \in J} H_\infty(N|C)_{\chi_{NC,i}}, \]
\[ = \inf_i (\log \max_j |\alpha_i|^2 p(j|i)) \]
\[ \geq -\log p^* \geq -n \log c. \]
Finally, from the strong subadditivity of min entropy [8]:
\[ H_\infty(N|\psi) \geq H(N|C)_{\psi} \geq H(N|C)_X - \log |J| \]
\[ \geq H(N|CI)_X - \log |J| \geq -n \log c - \log |J| \]
\[ \geq -n \log c - nH(w_{a}(q) + \delta), \]
The above analysis only utilized the ideal state from which sampling is guaranteed to succeed. However, the analysis of the real state follows identically as earlier (when we considered an initial pure state), thus completing the proof. \[ \square \]

The above result gives us an interesting entropic uncertainty bound in terms of smooth entropy and also in terms of the success of a quantum sampling strategy. Beyond its independent interest, it also, as a simple corollary, gives us the usual Maassen and Uffink entropic relation in the case of i.i.d. states.

**Corollary II.1.** Let \( M \) and \( N \) be two PMs and \( \rho \) a qubit density operator. Then, except with arbitrarily small probability, it holds that:
\[ H(M)_\rho + H(N)_\rho \geq -\log c. \]

**Proof.** Let \( \rho \) be a density operator on \( \mathcal{H}_2 \) and consider the state \( \rho^{2n} \). Let \( a = \max_x t_r(\{\mu_x \} \cdot \rho) \); in particular, if measuring \( \rho \) using \( M \) the probability of observing \( |\mu_a| \) is no less than \( 1/2 \). Note that this “\( a \)” need not be known to users making the measurement, however it clearly exists. Since \( \rho^{2n} \) is i.i.d., for any subset \( t \) of size \( n \) and any measurement outcome \( q \) on that subset, the post-measurement state is simply \( \rho^{2n-2}. \)

Fix \( \varepsilon > 0 \) and \( 0 < \beta < 1/2 \). Then, for any \( n \) and \( \epsilon \leq \varepsilon \), Theorem 2 implies that, except with probability at most \( \varepsilon^{1-2\beta} \), the following inequality holds:
\[ \frac{1}{n} H_\infty^{2\varepsilon+2\beta}(N)_\rho^{2n} + H(w_{a}(q) + \delta) \geq -\log c, \]
where \( q \) is the observed value after measuring using \( M \) and:
\[ \delta = \sqrt{\frac{\gamma + 1}{n^2}}. \]
(We used \( m = n \) when applying the theorem.) By the asymptotic equipartition property [9], we have
\[ \lim_{n \to \infty} \lim_{\rho \to \infty} \frac{1}{n} H_\infty^{2\varepsilon+2\beta}(N)_\rho^{2n} = H(N|\rho). \]
By the law of large numbers, we have
\[ \lim_{n \to \infty} w_{a}(q) = p_{1-a}. \]
Note that by definition of \( a \), we have \( p_{1-a} \leq 1/2 \) thus allowing us to replace \( H(\cdot) \) with \( H(p_{1-a} ; p_a) = H(M|\rho). \)
Finally, \( \delta \to 0 \) as \( n \to \infty \). Given fixed \( \varepsilon \) the above holds; of course \( \varepsilon \) may be made arbitrarily small, thus yielding the result. \[ \square \]

**III. CLOSING REMARKS**

In this paper we showed an interesting connection between quantum sampling and quantum uncertainty. We used the quantum sampling technique introduced in [7] to prove a finite non-i.i.d. case of an entropic uncertainty relation based on smooth min entropy. From this we were able to derive an alternative, and simple, proof for the Maassen and Uffink bound first proven in [1]. This connection of quantum sampling and entropic uncertainty may hold wider application to other areas of quantum information science. Note that, though we only proved the qubit case of the Maassen and Uffink entropic uncertainty relation, we strongly suspect this technique can be used to prove the higher dimensional case also. It would also be interesting to see if quantum sampling can yield a simple proof for the conditional version of the uncertainty relation, namely
\[ H(M|B) + H(N|E) \geq -\log c \]
We are currently investigating this, also, as future work.

**Acknowledgments:** WK would like to acknowledge support from National Science Foundation grant number 1812070.

**References**

[1] H. Maassen and J. B. M. Uffink, “Generalized entropic uncertainty relations,” Phys. Rev. Lett., vol. 60, pp. 1103–1106, Mar 1988.
[2] K. Kraus, “Complementary observables and uncertainty relations,” Phys. Rev. D, vol. 35, pp. 3070–3075, May 1987.
[3] D. Deutsch, “Uncertainty in quantum measurements,” Phys. Rev. Lett., vol. 50, pp. 631–633, Feb 1983.
[4] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, “Entropic uncertainty relations and their applications,” Rev. Mod. Phys., vol. 89, p. 015002, Feb 2017.
[5] I. Bialynicki-Birula and L. Rudnicki, “Entropic uncertainty relations in quantum physics,” in Statistical Complexity. Springer, 2011, pp. 1–34.
[6] S. Wehner and A. Winter, “Entropic uncertainty relations on a survey,” New Journal of Physics, vol. 12, no. 2, p. 025009, 2010.
[7] N. J. Bouman and S. Fehr, “Sampling in a quantum population, and applications,” in Annual Cryptology Conference. Springer, 2010, pp. 724–741. [Online]. Available: arXiv preprint arXiv:0907.4246
[8] R. Renner, “Security of quantum key distribution,” Ph.D. dissertation, Citeseer, 2005.
[9] M. Tomamichel, R. Colbeck, and R. Renner, “A fully quantum asymptotic equipartition property,” IEEE Transactions on Information Theory, vol. 55, no. 12, pp. 5840–5847, 2009.
[10] M. Tomamichel and R. Renner, “Uncertainty relation for smooth entropies,” Physical review letters, vol. 106, no. 11, p. 110506, 2011.