Relative entropy and a substate theorem about quantum states

Rahul Jain Jaikumar Radhakrishnan Pranab Sen*

School of Technology and Computer Science
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400005, India
email: {rahulj, jaikumar, pranab}@tcs.tifr.res.in

*Work done while the author was at LRI, UMR 8623, Université de Paris–Sud, 91405 Orsay, France, supported by the EU 5th framework program QAIP IST-1999-11234, and by CNRS/STIC 01N80/0502 and 01N80/0607 grants.

Relative entropy

Classical setting: Let P, Q be probability distributions over the finite sample space $[n]$. Their relative entropy (aka information divergence or Kulback-Leibler divergence) is defined as

$$S(P\|Q) \triangleq \sum_{i\in[n]} P(i)(\log P(i) - \log Q(i)).$$

Quantum setting: Let ρ, σ be density matrices in the same finite dimensional Hilbert space $\mathcal{H}$. Their relative entropy $S(\rho\|\sigma)$ is defined as

$$S(\rho\|\sigma) \triangleq \mathrm{Tr}\ \rho(\log\rho - \log\sigma).$$

A substate theorem about relative entropy

Theorem (informally): If $S(\rho\|\sigma) \leq k$, then $\frac{\rho}{2^{O(k)}}$ is 'approximately' a 'substate' of σ i.e. $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

For Hermitian operators A, B, $A \leq B$ is a shorthand for "B - A is positive semidefinite".

Note that by the *operator-monotonicity* of log (i.e. $A \leq B \Rightarrow \log A \leq \log B$), if $\frac{\rho}{2^k} \leq \sigma$, $S(\rho\|\sigma) \leq k$. The substate theorem can be thought of as a converse.

The substate theorem thus gives a new intuitive way of understanding what relative entropy really *means*.

Substate theorem: formal statements

Classical setting: Suppose P, Q are probability distributions on $[n]$. If $S(P\|Q) \leq k$, then for all $r > 1$, there exists a probability distribution P' on [n] such that $\|P - P'\|_1 \leq \frac{2}{r}$, and $\frac{r-1}{r2^{rk'}}P' \leq Q$, where $k' \triangleq k + 1$.

Quantum setting: Suppose ρ, σ are quantum states in the same finite dimensional Hilbert space $\mathcal{H}$. If $S(\rho\|\sigma) \leq k$, then for all $r > 1$, there exists a quantum state ρ' in $\mathcal{H}$ such that $\|\rho - \rho'\|_t \leq \frac{2}{\sqrt{r}}$, and $\frac{r-1}{r2^{rk'}}\rho' \leq \sigma$, where $k' \triangleq 8k + 14$.

Proof in classical setting

$$S(P\|Q) \triangleq \sum_{i\in[n]} P(i)\log\frac{P(i)}{Q(i)} \leq k.$$

Fix any $r > 1$. Let

$$\mathsf{Good} \triangleq \{i : P(i)/2^{r(k+1)} \leq Q(i)\}.$$

$P(i \notin \mathsf{Good}) < 1/r$, using a standard classical information-theoretic inequality. Let $P'(i) \triangleq P(i \mid i \in \mathsf{Good})$. Then,

$$\frac{r-1}{r2^{r(k+1)}}P' \leq Q \quad \text{and} \quad \|P - P'\|_1 \leq \frac{2}{r}.$$

This completes the proof of the substate theorem in the classical case.

Proof in quantum setting: I

The proof method of the classical setting fails to work, as ρ and σ need not be simultaneously diagonalisable.

We use an indirect approach. First, define the *observational divergence* of ρ and σ as follows.

$$D(\rho \| \sigma) \stackrel{\Delta}{=} \sup_F \mathrm{Tr}\ (F\rho) \log \frac{\mathrm{Tr}\ (F\rho)}{\mathrm{Tr}\ (F\sigma)},$$

where the supremum is over all POVM elements F s.t. Tr $(F\sigma) \neq 0$.

$D(\rho\|\sigma) < S(\rho\|\sigma) + 1$, using Lindblad-Uhlmann monotonicity.
Remark: Note that in general, $D(\rho\|\sigma)$ can be much smaller than $S(\rho\|\sigma)$.

Proof in quantum setting: II

We then show the following two results.

Substate theorem (first state pure): Suppose ρ is a pure state and $k \triangleq D(\rho\|\sigma)$. Then for all $r > 1$, there exists a pure state ρ' such that $\|\rho - \rho'\|_t \leq \frac{2}{\sqrt{r}}$ and $\left(\frac{r-1}{r2^{rk}}\right)\rho' \leq \sigma$.

Observational Divergence lifting theorem: Suppose $\mathcal{K}$ is another finite dimensional Hilbert space, and $\dim(\mathcal{H}) \leq \dim(\mathcal{K})$. Let $|\psi\rangle$ be a purification of ρ in $\mathcal{H} \otimes \mathcal{K}$. Then there exists a quantum state ω in $\mathcal{H} \otimes \mathcal{K}$ such that $\mathrm{Tr}_{\mathcal{K}}\, \omega = \sigma$ and $D((|\psi\rangle\langle\psi|)\|\,\omega) < 8D(\rho\|\sigma) + 6$.
Remark: Note that we do not yet know how to 'lift' relative entropy.

Combining the above three statements and tracing out $\mathcal{K}$, we finally prove the substate theorem in the quantum case.

Two consequences of the substate theorem

Substate theorem: If $S(\rho\|\sigma) \leq k$, $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

If $S(\rho\|\sigma) \leq k$, $\|\rho - \sigma\|_t \leq 2 - 2^{-O(k)}$.

If $S(\rho\|\sigma) \leq k$, $F(\rho, \sigma) \geq 2^{-O(k)}$, where $F(\rho, \sigma)$ is the Jozsa fidelity of ρ and σ.

The index function problem and privacy

Substate theorem: If $S(\rho\|\sigma) \leq k$, $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

The problem: Alice has a bit string $x \in \{0, 1\}^n$ and Bob has an index $i \in [n]$. They must exchange messages according to a quantum protocol $\mathcal{P}$ so that in the end Bob knows x_i.

The privacy model: A player turns malicious (i.e. strays away from $\mathcal{P}$) and wants to gain information about the input of the other player, without him realising that some cheating is going on.

The question: Suppose Bob 'leaks' at most b bits of information about i to a malicious Alice. How much information about x does Alice 'leak' to a malicious Bob?

Privacy tradeoff: main idea

Substate theorem: If $S(\rho\|\sigma) \leq k$, $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

Consider two-round protocols $\mathcal{P}$ only. Let ρ_i be Bob's message when his index is i. Let $\rho \triangleq \frac{1}{n}\sum_{i\in[n]} \rho_i$ be the average message. Let $k_i \triangleq S(\rho_i\|\rho)$. Since the information about i in Bob's message is at most k, $\frac{1}{n}\sum_{i\in[n]} k_i \leq k$. By Markov's inequality, $k_i \leq 10k$ for at least 90% of the i's (call them *secretive*).

A malicious Bob can cheat as follows. He sends ρ irrespective of his input i. At the end, he tries to infer x_j for all secretive j's, using the qubits in his possession. Since ρ_j and ρ are 'close' if j is secretive, this is possible with reasonable probability. Since 90% of the j's are secretive, Alice must 'leak' a large amount of information about x.

What does 'close' mean?

Substate theorem: If $S(\rho \| \sigma) \leq k$, $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

Since we will allow k to be super constant, $\left\| \rho_j - \rho \right\|_t \approx 2$ even for secretive j's. So trace distance is not the correct notion of 'closeness'. The substate theorem tells us, for every secretive j, that ρ contains ρ_j with weight factor $2^{-O(k)}$. Thus, there is the following 'split' of ρ.

$$\rho = \alpha_j \rho'_j + (1 - \alpha_j) \rho''_j,$$

where $\alpha_j = 2^{-O(k)}$, ρ'_j, ρ''_j are quantum states, and $\left\| \rho_j - \rho'_j \right\|_t$ is small.

Bob tries to guess x_j as in $\mathcal{P}$ if the 'split' gives ρ'_j; else, he tosses a fair coin. This gives a success probability of $2^{-1} + 2^{-O(k)}$.

The final privacy tradeoff result

Substate theorem: If $S(\rho\|\sigma) \leq k$, $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

Formalising the above ideas, gives us the following result.
Theorem: For the index function problem, if Bob 'leaks' at most b bits of information about i to a malicious Alice, then Alice 'leaks' at least $n/2^{O(b)}$ bits of information about x to a malicious Bob.
Remark: This tradeoff is optimal!

Corollary: For the index function problem, if Bob sends at most b qubits, Alice must send at least $n/2^{O(b)}$ qubits.

This theorem generalises earlier work (in both classical and quantum settings) by Miltersen, Nisan, Safra and Wigderson, and by Nayak.

Importance of ‘splitting’

Substate theorem: If $S(\rho\|\sigma) \leq k$, $\frac{\rho'}{2^{O(k)}} \leq \sigma$, where ρ' is close to ρ.

The privacy tradeoff argument for index function crucially depends on the fact that, for a secretive j, one can split the average message ρ into a ‘good’ state ρ'_j and a ‘bad’ state ρ''_j, the ‘good’ state appearing with reasonable weight factor $\alpha_j = 2^{-O(k)}$. Lindblad-Uhlmann monotonicity does not give this ‘split’; we require the substate theorem for this purpose.

The ‘splitting’ approach via the substate theorem also plays a crucial role in proving lower bounds for bounded error quantum protocols for the pointer chasing problem, full pointer version.

Conclusions and Open Problems

A substate theorem about quantum states.
Optimal privacy tradeoff for the index function problem.
Tight lower bound for bounded error quantum communication protocols for pointer chasing.

Improving the parameters in the substate theorem (e.g. the dependence on r)?
Privacy tradeoffs for other problems?
More applications of the substate theorem?