Short proofs of the Quantum Substate Theorem

Rahul Jain and Ashwin Nayak

Abstract—**The Quantum Substate Theorem due to Jain, Radhakrishnan, and Sen (2002) gives us a powerful operational interpretation of relative entropy, in fact, of the observational divergence of two quantum states, a quantity that is related to their relative entropy. Informally, the theorem states that if the observational divergence between two quantum states ρ, σ is small, then there is a quantum state ρ' close to ρ in trace distance, such that ρ' when scaled down by a small factor becomes a substate of σ. We present new proofs of this theorem. The resulting statement is optimal up to a constant factor in its dependence on observational divergence. In addition, the proofs are both conceptually simpler and significantly shorter than the earlier proof.**

Index Terms—**Quantum information theory, observational divergence, relative entropy, substate theorem, smooth relative min-entropy.**

I. The Quantum Substate Theorem

CONSIDER quantum states $\rho, \sigma \in \mathrm{D}(\mathcal{H})$, where $\mathcal{H}$ is a finite dimensional Hilbert space $\mathcal{H}$, and $\mathrm{D}(\mathcal{H})$ denotes the set of all quantum states with support in $\mathcal{H}$, i.e., the set of unit trace positive semi-definite operators on $\mathcal{H}$. We say that ρ is a c-substate of σ if $\rho \preceq 2^c \sigma$, where $\preceq$ represents the Löwner partial order on operators on $\mathcal{H}$. We may equivalently express this condition in terms of measurement outcomes ("POVM elements") as follows. Let

$$\mathrm{P}(\mathcal{H}) \quad \overset{\mathsf{def}}{=} \quad \{ M \in \mathrm{L}(\mathcal{H}) \; : \; O \preceq M \preceq \mathrm{I} \} \quad ,$$

denote the set of POVM elements on $\mathcal{H}$, where $\mathrm{L}(\mathcal{H})$ is the space of linear operators and I is the identity operator on $\mathcal{H}$. The state ρ is a c-substate of σ iff for every measurement outcome $M \in \mathrm{P}(\mathcal{H})$, the probability $\mathrm{Tr}(M\sigma)$ of observing M when σ is measured according to the POVM $\{M, \mathrm{I} - M\}$ is at least $\mathrm{Tr}(M\rho)/2^c$, a $1/2^c$ fraction of the probability of observing M when ρ is measured. Morally, the state σ may be decomposed as $\sigma = \alpha\rho + (1-\alpha)\tilde{\sigma}$, for some $\tilde{\sigma} \in \mathrm{D}(\mathcal{H})$, with $\alpha \geq 1/2^c$. This in turn may be used to construct the state ρ from σ through quantum analogues of rejection sampling. For example, we may apply the quantum measurement specified by the Kraus operators $\left\{ \sqrt{\alpha}\, \rho^{1/2}\sigma^{-1/2}, \sqrt{1-\alpha}\, \tilde{\sigma}^{1/2}\sigma^{-1/2} \right\}$, or go through a purification of σ [1], [2].

Given arbitrary quantum states $\rho, \sigma \in \mathrm{D}(\mathcal{H})$ we are interested in how well σ masquerades as ρ in the above sense. In other words, we are interested in the least c such that ρ is a c-substate of σ. We call this quantity the *relative min-entropy* $\mathsf{S}_\infty(\rho\|\sigma)$ of the two states. A generalization of this notion to bipartite states has been studied by Renner [3, Chapter 3], and the notion itself has been studied by Datta [4] as "max-relative entropy". For typical applications, such as privacy trade-offs in communication protocols [1], [2], it suffices to construct an approximation ρ' to ρ, with respect to a metric on quantum states. This leads us to the notion of the *smooth relative min-entropy* $\mathsf{S}^\varepsilon_\infty(\rho\|\sigma)$ of the two states, a quantity implicitly studied by Jain, Radhakrishnan, and Sen [1], [2] and later explicitly by Renner [3, Chapter 3] and Datta [4]. The metric initially used for the smoothness parameter ε was the trace distance. The fidelity of quantum states gives us a more natural metric in typical applications, and we adopt this measure of closeness in the article.

Let $\varepsilon \in (0,1)$ and $\rho, \sigma \in \mathrm{D}(\mathcal{H})$ be such that $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. We may express the ε-smooth relative min-entropy $\mathsf{S}^\varepsilon_\infty(\rho\|\sigma)$ as the base 2 logarithm of the value of following optimization problem with variables $\rho' \in \mathrm{D}(\mathcal{H})$ and $\kappa \in \mathbb{R}$:

$$\begin{array}{rcl}
\text{minimize:} & \kappa & \\
\text{subject to:} & & \\
\rho' & \preceq & \kappa\,\sigma \\
\mathrm{Tr}\,\rho' & = & 1 \\
\mathsf{F}(\rho,\rho') & \geq & 1-\varepsilon \\
\rho' \in \mathrm{L}(\mathcal{H}), & \rho' \succeq 0 & \\
\kappa \in \mathbb{R}, & \kappa \geq 0 &
\end{array} \qquad \text{(P1)}$$

Here $\mathsf{F}(\rho',\rho) \overset{\mathsf{def}}{=} \left\| \sqrt{\rho'}\sqrt{\rho} \right\|^2_{\mathrm{tr}}$, denotes the fidelity between the two quantum states, and $\|M\|_{\mathrm{tr}} \overset{\mathsf{def}}{=} \mathrm{Tr}\sqrt{M^\dagger M}$ denotes the trace norm of the linear operator $M \in \mathrm{L}(\mathcal{H})$. The existence of a pair ρ', κ that are feasible for the problem (P1) means that there is a quantum state ρ' with fidelity $\mathsf{F}(\rho',\rho) \geq 1-\varepsilon$ that is also a $(\log_2 \kappa)$-substate of σ. The substate constraint implies that $\kappa \geq 1$.

The program (P1) is feasible, as $\rho' \overset{\mathsf{def}}{=} \rho$ and $\kappa \overset{\mathsf{def}}{=} 1/\lambda$, where λ is the smallest non-zero eigenvalue of σ, satisfy all the constraints. Therefore we may restrict the optimization to $\kappa \in [0, 1/\lambda]$ and the compact set of quantum states with fidelity at least $1-\varepsilon$ with ρ. The ε-smooth relative min-entropy between the two states is thus always achieved.

If ρ is a c-substate of σ, i.e., their relative min-entropy is at most c, their relative entropy $\mathsf{S}(\rho\|\sigma) \overset{\mathsf{def}}{=} \mathrm{Tr}\,\rho(\log_2 \rho - \log_2 \sigma)$

Centre for Quantum Technologies and Department of Computer Science, National University of Singapore, Block S15, 3 Science Drive 2, Singapore 11754. Email: rahul@comp.nus.edu.sg. Work done in part while visiting the Institute for Quantum Computing, University of Waterloo.

Department of Combinatorics and Optimization, and Institute for Quantum Computing, University of Waterloo, 200 University Ave. W., Waterloo, ON, N2L 3G1, Canada. Email: ashwin.nayak@uwaterloo.ca. Work done in part at Center for Quantum Technologies, National University of Singapore, and at Perimeter Institute for Theoretical Physics. Research supported in part by NSERC Canada, CIFAR, an ERA (Ontario), QuantumWorks, MITACS, and ARO (USA). Research at Perimeter Institute is supported in part by the Government of Canada through Industry Canada and by the Province of Ontario through MRI.

is also at most c. Jain *et al.* [1], [2] gave a weak converse to this relation via the Quantum Substate Theorem, which gives a bound on the ε-smooth relative min-entropy in terms of the more familiar notion of relative entropy. This theorem may also be viewed as a handy operational interpretation of the rather abstract notion of relative entropy.

The substate theorem (classical or quantum) lies at the heart of a growing number of applications [2, Section 1]. These include privacy trade-offs in communication protocols for computing relations [5], message compression leading to direct sum theorems in classical and quantum communication complexity [5], impossibility results for bit-string commitment [6], the communication complexity of remote state preparation [7], and direct product theorems for classical communication complexity [8], [9]. To highlight one of these examples, the Quantum Substate Theorem enables (non-oblivious) compression of an ensemble of mixed quantum states to within a constant factor of the Holevo information of the ensemble, given access to shared entanglement and classical communication, when we are allowed a small loss of fidelity in the compression process. In contrast, the compression of arbitrary ensembles of mixed quantum states to the Holevo limit remains an open problem in quantum information theory.

Jain *et al.* formulated their bound in terms of a new information theoretic quantity, *observational divergence* $\mathsf{D}(\rho\|\sigma)$, rather than relative entropy.

Definition 1 (Observational divergence)**.** *Let $\rho, \sigma \in \mathrm{D}(\mathcal{H})$. Their* observational divergence *is defined as*

$$\mathsf{D}(\rho\|\sigma) \stackrel{\mathsf{def}}{=} \sup\left\{ (\mathrm{Tr}\, M\rho) \log_2 \frac{\mathrm{Tr}\, M\rho}{\mathrm{Tr}\, M\sigma} \;:\; M \in \mathrm{P}(\mathcal{H}),\ \mathrm{Tr}\, M\sigma \neq 0 \right\} .$$

The supremum in the definition above is achieved if and only if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. As is evident, this quantity is a scaled measure of the maximum factor by which $\mathrm{Tr}(M\rho)$ may exceed $\mathrm{Tr}(M\sigma)$ for any measurement outcome M of interest. Observational divergence is related to relative entropy. In particular, $\mathsf{D}(\rho\|\sigma) \leq \mathsf{S}(\rho\|\sigma) + 1$. However, it could be smaller than relative entropy by a factor proportional to the dimension [2, Proposition 4] (see also [10]).

We present alternative proofs of the Quantum Substate Theorem, also strengthening it in the process.

Theorem 1. *Let $\mathcal{H}$ be a Hilbert space, and let $\rho, \sigma \in \mathrm{D}(\mathcal{H})$ be quantum states such that $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. For any $\varepsilon \in (0,1)$, there is a quantum state ρ' with fidelity $\mathsf{F}(\rho', \rho) \geq 1-\varepsilon$ such that $\rho' \preceq \kappa\sigma$, where*

$$\kappa = \frac{1}{1-\varepsilon}\, 2^{\mathsf{D}(\rho\|\sigma)/\varepsilon} .$$

Equivalently, for any $\varepsilon \in (0,1)$,

$$\mathsf{S}^{\varepsilon}_{\infty}(\rho\|\sigma) \leq \frac{\mathsf{D}(\rho\|\sigma)}{\varepsilon} + \log_2 \frac{1}{1-\varepsilon} .$$

The proofs that we present are both shorter and conceptually simpler than the original proof. The proof due to Jain *et al.* consists of a number of technical steps, several of which are bundled into a "divergence lifting" theorem that reduces the problem to one in which ρ is a pure state (a rank one quantum state). Finally, the pure state case is translated into a problem in two dimensions which is solved by a direct calculation. Divergence lifting involves going from a construction of a suitable state ρ' for a fixed POVM element to one that is independent of the POVM element, by appealing to a minimax theorem from Game Theory. We show that this minimax theorem can be applied directly to establish the Quantum Substate Theorem. The resulting statement is stronger in its dependence on observational divergence. The original bound read as $\mathsf{S}^{\varepsilon}_{\infty}(\rho\|\sigma) \leq d'/\varepsilon - \log_2(1-\varepsilon)$, where $d' \stackrel{\mathsf{def}}{=} d + 4\sqrt{d+2} + 2\log_2(d+2) + 6$ with $d \stackrel{\mathsf{def}}{=} \mathsf{D}(\rho\|\sigma)$. The formulation in terms of fidelity also allows us to show that the dependence on observational divergence in Theorem 1 is optimal up to a constant factor.

Theorem 2. *Let $\mathcal{H}$ be a Hilbert space, and let $\rho, \sigma \in \mathrm{D}(\mathcal{H})$ be quantum states such that $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. Suppose $k \in \mathbb{R}$ is such that for any $\varepsilon \in (0,1)$, there is a quantum state ρ' with fidelity $\mathsf{F}(\rho', \rho) \geq 1-\varepsilon$ such that $\rho' \preceq \kappa\sigma$, where*

$$\kappa = \frac{1}{1-\varepsilon}\, 2^{k/\varepsilon} ,$$

or equivalently,

$$\mathsf{S}^{\varepsilon}_{\infty}(\rho\|\sigma) \leq \frac{k}{\varepsilon} + \log_2 \frac{1}{1-\varepsilon} .$$

Then $\mathsf{D}(\rho\|\sigma) \leq 4k + 3$.

We thus settle two questions posed by Jain *et al.* [2, Section 5].

For the first proof (Section II), we start by converting the convex minimization problem (P1) into a min-max problem through a simple duality argument. The minimax theorem now applies and reduces the problem of construction of a suitable state ρ' to one that works for a fixed POVM element. The latter task turns out to be similar to proving the Classical Substate Theorem. This proof is thus shorter and conceptually simpler than the original one, and also leads to a tighter dependence on observational divergence. We present a second proof based on semi-definite programming (SDP) duality (Section III). We believe that both approaches have their own merits. The first approach is more intuitive in that once the problem is formulated as a min-max program, the subsequent steps emerge naturally. The second approach has the appeal of relying on the more standard SDP duality. These routes to the theorem may prove useful in its burgeoning list of applications, as also in the study of smooth relative min-entropy.

II. A proof based on min-max duality

In this section, we present an alternative proof of the Quantum Substate Theorem. It hinges on a powerful minimax theorem from game theory, which is a consequence of the Kakutani fixed point theorem in real analysis [11, Propositions 20.3 and 22.2].

Theorem 3. *Let A_1, A_2 be non-empty, convex and compact subsets of $\mathbb{R}^n$ for some positive integer n. Let $u : A_1 \times A_2 \to \mathbb{R}$ be a continuous function such that*

- *$\forall a_2 \in A_2$, the set $\{a_1 \in A_1 \; : \; (\forall a_1' \in A_1) \; u(a_1, a_2) \geq u(a_1', a_2)\}$ is convex, i.e., for every $a_2 \in A_2$, the set of points $a_1 \in A_1$ such that $u(a_1, a_2)$ is maximum is a convex set; and*
- *$\forall a_1 \in A_1$, the set $\{a_2 \in A_2 \; : \; (\forall a_2' \in A_2) \; u(a_1, a_2) \leq u(a_1, a_2')\}$ is convex, i.e., for every $a_1 \in A_1$, the set of points $a_2 \in A_2$, such that $u(a_1, a_2)$ is minimum is a convex set.*

Then, there is an $(a_1^*, a_2^*) \in A_1 \times A_2$ such that

$$\begin{aligned} \max_{a_1 \in A_1} \min_{a_2 \in A_2} u(a_1, a_2) &= u(a_1^*, a_2^*) \\ &= \min_{a_2 \in A_2} \max_{a_1 \in A_1} u(a_1, a_2) \; . \end{aligned}$$

We start with the following lemma which bounds the distance between a quantum state and its normalized projection onto a subspace in which it has "large" support. It is a variant of the "gentle measurement lemma" due to Winter [12].

Lemma 4. *Let $\rho \in \mathrm{D}(\mathcal{H})$ be a quantum state in the Hilbert space $\mathcal{H}$. Let Π be an orthogonal projection onto a subspace of $\mathcal{H}$ such that $\mathrm{Tr}\, \Pi \rho = \delta < 1$. Let $\rho'' = (\mathrm{I} - \Pi)\rho(\mathrm{I} - \Pi)$ be the projection of ρ onto the orthogonal subspace, and let $\rho' = \frac{\rho''}{\mathrm{Tr}\, \rho''}$ be this state normalized. Then $\mathsf{F}(\rho, \rho') \geq 1 - \delta$.*

Proof: Let $\mathcal{K}$ be a Hilbert space with $\dim(\mathcal{K}) = \dim(\mathcal{H})$. Let $|v\rangle \in \mathcal{K} \otimes \mathcal{H}$ be a purification of ρ [13]. Let $|v''\rangle = (\mathrm{I} \otimes (\mathrm{I} - \Pi))|v\rangle$. Let $|v'\rangle = |v''\rangle / \left\| v'' \right\|$. Observe that $\mathrm{Tr}_{\mathcal{K}} |v''\rangle\langle v''| = \rho''$, so

$$\begin{aligned} \left\| v'' \right\|^2 &= \mathrm{Tr}\, |v''\rangle\langle v''| \\ &= \mathrm{Tr}\, \rho'' \quad = \quad \mathrm{Tr}\, \rho - \mathrm{Tr}\, \Pi \rho \\ &= 1 - \delta \; , \end{aligned}$$

and $\mathrm{Tr}_{\mathcal{K}} |v'\rangle\langle v'| = \rho'$. Now,

$$\begin{aligned} \mathsf{F}(\rho, \rho') &\geq \mathsf{F}(|v\rangle\langle v|, |v'\rangle\langle v'|) \\ &= |\langle v | v' \rangle|^2 \quad = \quad \left\| v'' \right\|^2 \\ &= 1 - \delta \; , \end{aligned}$$

where the first inequality follows from the monotonicity of fidelity under completely positive trace preserving (CPTP) operations [13]. ■

The next lemma is an important step in the proof, and along with the minimax theorem (Theorem 3) yields the Quantum Substate Theorem. It mimics the proof of the Classical Substate Theorem with respect to a particular operator $M \succeq 0$, which may be viewed as an unnormalized POVM element. Namely, we decompose M into its diagonal basis, and imagine measuring with respect to this basis. If the observational divergence of ρ with respect to σ is small, then for most of the basis elements, the probability of the outcome for ρ is not too large relative to the probability for σ. Projecting ρ onto the space spanned by these basis elements gives us a state ρ', close to ρ, for which $\mathrm{Tr}\, M\rho'$ is correspondingly bounded, relative to $\mathrm{Tr}\, M\sigma$.

Lemma 5. *Suppose $\rho, \sigma \in \mathrm{D}(\mathcal{H})$ are quantum states in the Hilbert space $\mathcal{H}$ such that $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. Let $d = \mathsf{D}(\rho \| \sigma)$, $\varepsilon \in (0, 1)$, and $M \succeq 0$ be an operator on $\mathcal{H}$. There exists a quantum state $\rho' \in \mathrm{D}(\mathcal{H})$ such that $\mathsf{F}(\rho', \rho) \geq 1 - \varepsilon$ and*

$$(1 - \varepsilon) \cdot \mathrm{Tr}\, M\rho' \quad \leq \quad 2^{d/\varepsilon} \cdot \mathrm{Tr}\, M\sigma \; .$$

Proof: Consider M in its diagonal form $M = \sum_{i=1}^{\dim(\mathcal{H})} p_i |v_i\rangle\langle v_i|$, where the (p_i) are the eigenvalues of M corresponding to the orthonormal eigenvectors $(|v_i\rangle)$. Let

$$\begin{aligned} B &\overset{\mathsf{def}}{=} \\ &\left\{ i \; : \; \langle v_i | \rho | v_i \rangle > 2^{d/\varepsilon} \cdot \langle v_i | \sigma | v_i \rangle, \; 1 \leq i \leq \dim(\mathcal{H}) \right\} \; . \end{aligned}$$

Let $\Pi = \sum_{i \in B} |v_i\rangle\langle v_i|$ be the projector onto the space spanned by vectors specified by B. Then $\mathrm{Tr}\, \Pi\rho > 2^{d/\varepsilon} \cdot \mathrm{Tr}\, \Pi\sigma$ and hence,

$$d \quad \geq \quad (\mathrm{Tr}\, \Pi\rho) \log_2 \frac{\mathrm{Tr}\, \Pi\rho}{\mathrm{Tr}\, \Pi\sigma} \quad > \quad (\mathrm{Tr}\, \Pi\rho) \cdot \frac{d}{\varepsilon} \; .$$

This implies that $\mathrm{Tr}\, \Pi\rho < \varepsilon$. Let $\rho'' = (\mathrm{I} - \Pi)\rho(\mathrm{I} - \Pi)$ and $\rho' = \frac{\rho''}{\mathrm{Tr} \rho''} \prec \frac{\rho''}{1 - \varepsilon}$. From Lemma 4 we have $\mathsf{F}(\rho, \rho') \geq 1 - \varepsilon$. Finally, by the definition of Π,

$$\begin{aligned} (1 - \varepsilon) \cdot \mathrm{Tr}\, M\rho' &\leq \mathrm{Tr}\, M\rho'' \\ &= \sum_{i \notin B} p_i \langle v_i | \rho | v_i \rangle \\ &\leq 2^{d/\varepsilon} \sum_{i \notin B} p_i \langle v_i | \sigma | v_i \rangle \\ &\leq 2^{d/\varepsilon} \cdot \mathrm{Tr}\, M\sigma \; . \end{aligned}$$

■

We now prove the main result, Theorem 1. For this, it suffices to produce a state close to ρ that when scaled suitably is a substate of σ. The condition $\rho' \preceq \kappa\sigma$ is equivalent to $\mathrm{Tr}\, M\rho' \leq \kappa$ for all $M \succeq 0$ with $\mathrm{Tr}\, M\sigma \leq 1$. We use this dual view of the substate condition to convert the minimization problem (P1) into a min-max optimization problem. We then use the minimax theorem, Theorem 3, to drastically simplify the search for a suitable state ρ'. As a consequence, it suffices to produce a state ρ' close to ρ such that $\mathrm{Tr}\, M\rho' \leq \kappa$ for an arbitrary but fixed $M \succeq 0$ with $\mathrm{Tr}\, M\sigma \leq 1$.

Proof of Theorem 1: We first massage the program (P1) into a form to which Theorem 3 applies. If a pair ρ', κ are feasible for (P1), then $\mathrm{supp}(\rho') \subseteq \mathrm{supp}(\sigma)$. By taking $\mathcal{H} = \mathrm{supp}(\sigma)$ if necessary, we may therefore assume that $\sigma \succ 0$, i.e., σ has full support. It is straightforward to check that for any given $\rho' \in \mathrm{D}(\mathcal{H})$,

$$\min_{\kappa \; : \; \rho' \preceq \kappa\sigma} \kappa \quad = \quad \max_{M \succeq 0 \; : \; \mathrm{Tr}\, M\sigma \leq 1} \mathrm{Tr}\, M\rho' \; .$$

Hence we may rewrite $\mathsf{S}_\infty^\varepsilon(\rho \| \sigma)$ as the base 2 logarithm of

$$\min_{\substack{\rho' \succeq 0 \; : \; \mathrm{Tr}\, \rho' = 1, \\ \mathsf{F}(\rho', \rho) \geq 1 - \varepsilon}} \quad \max_{M \succeq 0 \; : \; \mathrm{Tr}\, M\sigma \leq 1} \quad \mathrm{Tr}\, M\rho' \; .$$

Viewing ρ' and M as elements of the real vector space of Hermitian operators in $\mathrm{L}(\mathcal{H})$, noting that fidelity is concave in each of its arguments [13] and that the trace function is bilinear, we may apply Theorem 3 to the resulting optimization problem. We get

$$2^{\mathsf{S}^{\varepsilon}_{\infty}(\rho\|\sigma)} = \max_{M\succeq 0 \,:\, \operatorname{Tr} M\sigma \leq 1} \quad \min_{\substack{\rho'\succeq 0 \,:\, \operatorname{Tr}\rho' = 1, \\ \mathsf{F}(\rho',\rho)\geq 1-\varepsilon}} \quad \operatorname{Tr} M\rho' \ .$$

By Lemma 5, for every $M \succeq 0$ with $\operatorname{Tr} M\sigma \leq 1$, there is a quantum state ρ', with $\mathsf{F}(\rho',\rho) \geq 1-\varepsilon$, such that $(1-\varepsilon)\operatorname{Tr} M\rho' \leq 2^{d/\varepsilon}$, where $d = \mathsf{D}(\rho\|\sigma)$. The desired result now follows. ■

Combining Theorem 1 and the Uhlmann theorem [13] immediately gives us the following statement. The Quantum Substate Theorem is often used in this form in its applications.

Corollary 6. *Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces with $\dim(\mathcal{K}) \geq \dim(\mathcal{H})$, and let $\rho,\sigma \in \mathrm{D}(\mathcal{H})$ be quantum states such that $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$. Let $d = \mathsf{D}(\rho\|\sigma)$, $\varepsilon \in (0,1)$, and $|v\rangle \in \mathcal{K}\otimes\mathcal{H}$ be a purification of ρ. Then there is a pure state $|v'\rangle \in \mathcal{K}\otimes\mathcal{H}$ with $\mathsf{F}(|v\rangle\langle v|, |v'\rangle\langle v'|) \geq 1-\varepsilon$, and a pure state $|w'\rangle \in \mathcal{K}\otimes\mathcal{H}$ such that $|w\rangle \in \mathbb{C}^2\otimes\mathcal{K}\otimes\mathcal{H}$ defined as*

$$|w\rangle \quad = \quad \sqrt{\alpha}\;|0\rangle|v'\rangle + \sqrt{1-\alpha}\;|1\rangle|w'\rangle \ ,$$

with $\alpha = (1-\varepsilon)2^{-d/\varepsilon}$, is a purification of σ.

Proof: Let ρ' be a state given by Theorem 1 such that fidelity $\mathsf{F}(\rho',\rho) \geq 1-\varepsilon$ and $\alpha\rho' \preceq \sigma$. Then we can decompose σ as

$$\sigma \quad = \quad \alpha\rho' + (1-\alpha)\theta \ ,$$

where $\theta \in \mathsf{D}(\mathcal{H})$ is some quantum state. By the Uhlmann theorem [13] there is a purification $|v'\rangle \in \mathcal{K}\otimes\mathcal{H}$ of ρ' such that $\mathsf{F}(|v\rangle\langle v|, |v'\rangle\langle v'|) = \mathsf{F}(\rho,\rho') \geq 1-\varepsilon$. Let $|w'\rangle \in \mathcal{K}\otimes\mathcal{H}$ be any purification of θ. Then we may verify that $|w\rangle$ as defined in the statement of the corollary is a purification of σ. ■

The dependence of the bound on the ε-smooth relative min-entropy in Theorem 1 in terms of observational divergence is optimal up to a constant factor, as stated in Theorem 2. We start its proof with the following lemma.

Lemma 7. *Let $\delta,\delta' \in [0,1]$ and $\beta \in [0,1/4]$ such that*

$$\left(\sqrt{\delta\,\delta'} + \sqrt{(1-\delta)(1-\delta')}\right)^2 \quad \geq \quad 1-\beta\delta \ .$$

Then $\delta' \geq \left(1-\sqrt{\beta}\right)^2\delta$.

Proof: Let $u = \left(\sqrt{\delta},\sqrt{1-\delta}\right)^{\mathrm{T}}$ and $u' = \left(\sqrt{\delta'},\sqrt{1-\delta'}\right)^{\mathrm{T}}$ be vectors in $\mathbb{R}^2$. Let $\phi,\phi' \in [0,\pi/2]$ be the angles u,u' make with $(1,0)^{\mathrm{T}}$, respectively. By hypothesis, $\langle u|u'\rangle^2 \geq 1-\beta\delta$. Let $\theta \in [0,\pi/2]$ be the angle between u,u', so that $\cos^2\theta \geq 1-\beta\delta$.

We wish to bound $\delta' = \cos^2\phi'$ from below given that $|\phi'-\phi| = \theta$. Observe that $\cos(\phi+\theta) \geq \sqrt{\delta(1-\beta\delta)} - \sqrt{(1-\delta)\beta\delta} \geq 0$, so that $\phi+\theta \leq \pi/2$. Therefore, δ' takes its minimum value when $\phi' = \phi+\theta$.

We may now bound δ' as follows.

$$\begin{aligned}
\delta' \quad = \quad \cos^2\phi' \quad & \geq \quad \cos^2(\phi+\theta) \\
& \geq \quad \left(\sqrt{\delta(1-\beta\delta)} - \sqrt{(1-\delta)\beta\delta}\right)^2 \\
& = \quad (1+\beta)\delta - 2\beta\delta^2 \\
& \qquad -2\sqrt{\beta}\,\delta\left(1-(1+\beta)\delta+\beta\delta^2\right)^{1/2} \\
& \geq \quad (1+\beta)\delta - 2\beta\delta^2 - 2\sqrt{\beta}\,\delta(1-\delta)^{1/2} \\
& \geq \quad (1+\beta)\delta - 2\beta\delta^2 - 2\sqrt{\beta}\,\delta(1-\delta/2) \\
& = \quad \left(1-\sqrt{\beta}\right)^2\delta + (\sqrt{\beta}-2\beta)\delta^2 \\
& \geq \quad \left(1-\sqrt{\beta}\right)^2\delta \ ,
\end{aligned}$$

since $\beta \leq 1/4$. ■

We are now ready to prove the optimality of Theorem 1.

Proof of Theorem 2: It suffices to prove that for any POVM element $M \in \mathrm{P}(\mathcal{H})$ with $\operatorname{Tr}(M\rho) \neq 0$,

$$\operatorname{Tr}(M\rho)\ \log\frac{\operatorname{Tr}(M\rho)}{\operatorname{Tr}(M\sigma)}$$

is bounded by $4k+3$ from above.

Fix such a POVM element M, let $\delta = \operatorname{Tr}(M\rho)$, and $\varepsilon = \beta\delta$ for some $\beta \in (0,1)$ to be specified later. By hypothesis, there is a quantum state $\rho' \in \mathrm{D}(\mathcal{H})$ with $\mathsf{F}(\rho',\rho) \geq 1-\varepsilon$ and $\rho' \preceq \kappa\,\sigma$, where

$$\kappa \quad = \quad \frac{2^{k/\varepsilon}}{1-\varepsilon} \ .$$

Let $\delta' = \operatorname{Tr}(M\rho')$. By the monotonicity of fidelity under CPTP operations [13], we have

$$\begin{aligned}
\left(\sqrt{\delta\,\delta'} + \sqrt{(1-\delta)(1-\delta')}\right)^2 \quad & \geq \quad \mathsf{F}(\rho',\rho) \\
& \geq \quad 1-\varepsilon \quad = \quad 1-\beta\delta \ .
\end{aligned}$$

By Lemma 7, we have $\operatorname{Tr}(M\rho') = \delta' \geq (1-\sqrt{\beta})^2\delta$ if $\beta \leq 1/4$.

We set $\beta = 1/4$, so that $\operatorname{Tr}(M\rho') \geq \delta/4$. Furthermore,

$$\begin{aligned}
\operatorname{Tr}(M\sigma) \quad & \geq \quad \frac{(1-\varepsilon)}{2^{k/\varepsilon}}\operatorname{Tr}(M\rho') \\
& \geq \quad \frac{(1-\delta/4)}{2^{4k/\delta}}(\delta/4) \quad \geq \quad \frac{\delta}{2^{3+4k/\delta}} \ ,
\end{aligned}$$

as $\delta \leq 1$. So

$$\operatorname{Tr}(M\rho)\ \log\frac{\operatorname{Tr}(M\rho)}{\operatorname{Tr}(M\sigma)} \quad = \quad \delta\ \log\frac{\delta}{\operatorname{Tr}(M\sigma)} \quad \leq \quad 4k+3 \ .$$

■

III. A PROOF BASED ON SDP DUALITY

In this section we present a second alternative proof of the Quantum Substate Theorem, Theorem 1. The proof is based on a formulation of smooth relative min-entropy as a semi-definite program.

The optimization problem (P1) in Section I is seen to be an SDP once we express the fidelity constraint as a semi-definite inequality. This is based on a formulation due to

Watrous [14] of the fidelity of two quantum states as an SDP. For completeness, we include a proof of its correctness.

Lemma 8 (Watrous)**.** *Suppose $\rho, \rho' \in \mathrm{D}(\mathcal{H})$ are quantum states in the Hilbert space $\mathcal{H}$. The fidelity $\mathsf{F}(\rho, \rho')$ of the two states equals the square of the optimum of the following SDP over the variable $X \in \mathrm{L}(\mathcal{H})$.*

$$\begin{aligned} &\textit{maximize:} \quad \frac{1}{2}\left(\operatorname{Tr} X + \operatorname{Tr} X^\dagger\right) \\ &\textit{subject to:} \\ &\qquad \begin{pmatrix} \rho' & X \\ X^\dagger & \rho \end{pmatrix} \succeq 0 \qquad \text{(P2)} \\ &\qquad X \in \mathrm{L}(\mathcal{H}) \end{aligned}$$

Proof: By Theorem IX.5.9 in the text [15], the matrix inequality in the program (P2) holds iff there is an operator $Y \in \mathrm{L}(\mathcal{H})$ such that $\|Y\| \leq 1$ and $X = \sqrt{\rho'}\, Y \sqrt{\rho}$. Since $\mathsf{F}(\rho', \rho) = \left\|\sqrt{\rho'}\sqrt{\rho}\right\|_{\mathrm{tr}}^2$ and we may characterize trace norm as $\|M\|_{\mathrm{tr}} = \max\left\{|\operatorname{Tr}(ZM)| \,:\, Z \in \mathrm{L}(\mathcal{H}), \|Z\| \leq 1\right\}$ for any $M \in \mathrm{L}(\mathcal{H})$, the lemma follows. ■

The problem (P1) may now be formulated as the following SDP with variables $\kappa \in \mathbb{R}, \rho' \in \mathrm{L}(\mathcal{H}), X \in \mathrm{L}(\mathcal{H})$ in the primal problem, and variables $Z_1, Z_2 \in \mathrm{L}(\mathcal{H})$ and $z_3, z_4 \in \mathbb{R}$ in the dual, where Z_1, Z_2 are Hermitian.

P3 Primal problem

$$\begin{aligned} &\text{minimize:} \quad \kappa \\ &\text{subject to:} \\ &\qquad \rho' \preceq \kappa\,\sigma \\ &\qquad \operatorname{Tr}\rho' = 1 \\ &\qquad \begin{pmatrix} \rho' & X \\ X^\dagger & \rho \end{pmatrix} \succeq 0 \\ &\qquad \operatorname{Tr} X + \operatorname{Tr} X^\dagger \geq 2\sqrt{1-\varepsilon} \\ &\qquad \rho' \in \mathrm{L}(\mathcal{H}), \quad \rho' \succeq 0 \\ &\qquad \kappa \in \mathbb{R}, \quad \kappa \geq 0 \\ &\qquad X \in \mathrm{L}(\mathcal{H}) \end{aligned}$$

P3 Dual problem

$$\begin{aligned} &\text{maximize:} \quad z_4 + 2z_3\sqrt{1-\varepsilon} + \operatorname{Tr}(Z_2\rho) \\ &\text{subject to:} \\ &\qquad \operatorname{Tr}(Z_1\sigma) \leq 1 \\ &\qquad \begin{pmatrix} z_4\mathrm{I} - Z_1 & z_3\mathrm{I} \\ z_3\mathrm{I} & Z_2 \end{pmatrix} \preceq 0 \\ &\qquad Z_1 \in \mathrm{L}(\mathcal{H}), \qquad Z_1 \succeq 0 \\ &\qquad z_3, z_4 \in \mathbb{R}, \qquad z_3 \geq 0 \\ &\qquad Z_2 \in \mathrm{L}(\mathcal{H}), \qquad Z_2 \text{ Hermitian} \end{aligned}$$

The equivalence of the problems (P1) and (P3) follows from Lemma 8 and paves the way for the second proof.

Proof of Theorem 1: We may verify that strong duality holds since the P3 primal program is feasible, and the dual is strictly feasible [14], [16]. Therefore, it suffices to bound the dual objective function for any set of dual feasible variables (Z_1, Z_2, z_3, z_4).

By Lemma 5, there is a quantum state ρ', with $\mathsf{F}(\rho', \rho) \geq 1 - \varepsilon$, such that $(1-\varepsilon)\operatorname{Tr}(Z_1\rho') \leq 2^{d/\varepsilon}\operatorname{Tr}(Z_1\sigma) \leq 2^{d/\varepsilon}$, where $d = \mathsf{D}(\rho\|\sigma)$.

Since $\mathsf{F}(\rho', \rho) \geq 1 - \varepsilon$, by Lemma 8, there is an operator $X \in \mathrm{L}(\mathcal{H})$ such that

$$\begin{pmatrix} \rho' & X \\ X^\dagger & \rho \end{pmatrix} \succeq 0 \ ,$$

and $\operatorname{Tr} X + \operatorname{Tr} X^\dagger \geq 2\sqrt{1-\varepsilon}$. Therefore,

$$\operatorname{Tr}\begin{pmatrix} \rho' & X \\ X^\dagger & \rho \end{pmatrix}\begin{pmatrix} z_4\mathrm{I} - Z_1 & z_3\mathrm{I} \\ z_3\mathrm{I} & Z_2 \end{pmatrix} \leq 0 \ ,$$

In other words,

$$z_4 - \operatorname{Tr}(Z_1\rho') + z_3(\operatorname{Tr} X + \operatorname{Tr} X^\dagger) + \operatorname{Tr}(Z_2\rho) \leq 0 \ ,$$

which implies that the dual objective function is bounded as

$$z_4 + 2z_3\sqrt{1-\varepsilon} + \operatorname{Tr}(Z_2\rho) \leq \operatorname{Tr}(Z_1\rho') \leq \frac{2^{d/\varepsilon}}{1-\varepsilon} \ .$$

This completes the proof. ■

IV. Conclusion

We presented two alternative proofs of the Quantum Substate Theorem due to Jain, Radhakrishnan, and Sen [1], [2]. In addition to giving bounds on the smooth relative min-entropy of two quantum states, this gives us a powerful operational interpretation of relative entropy and observational divergence. In the process, we resolve two questions left open by Jain *et al.*.

The crucial insight here is that the we may express smooth relative min-entropy as a convex or semi-definite program and appeal to duality theory. In this respect, we join a growing number of applications of convex and semi-definite programming to quantum information processing. This approach can be extended to the more general notion of smooth relative min-entropy studied by Renner [3] to get similar bounds on this quantity. This view of the quantity may shed light on its numerous applications.

Acknowledgement

For helpful discussions and pointers to the literature, we thank Matthias Christandl, Rajat Mittal, Marco Piani, Renato Renner, Pranab Sen, and John Watrous.

Our first attempt at this work involved an awkward formulation of smooth relative min-entropy defined with trace distance as a semi-definite program. We set aside this formulation in favour of the min-max program, which we believe also leads to a more intuitive proof. We thank John Watrous for sharing a simple SDP formulation (with a bound on trace distance) and the resulting proof. Our current presentation involving fidelity supersedes this.

Finally, we thank the referees for their extensive feedback. It lead not only to what we hope is a vastly improved presentation of the work, but also to the discovery of a tight connection between smooth relative min-entropy and observational divergence.

References

[1] R. Jain, J. Radhakrishnan, and P. Sen, “Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states,” in *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002, pp. 429–438.

[2] ——, “A new information-theoretic property about quantum states with an application to privacy in quantum communication,” *Journal of the ACM*, vol. 56, no. 6, Sep. 2009, article no. 33.

[3] R. Renner, “Security of quantum key distribution,” Ph.D. dissertation, Eidgenössische Technische Hochschule (ETH) Zürich, 2005.

[4] N. Datta, “Min- and max-relative entropies and a new entanglement monotone,” *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2816–2826, May 2009.

[5] R. Jain, J. Radhakrishnan, and P. Sen, “Prior entanglement, message compression and privacy in quantum communication,” in *Proceedings of the 20th Annual IEEE Conference on Computational Complexity*. IEEE Computer Society Press, Los Alamitos, CA, USA, 2005, pp. 285–296.

[6] R. Jain, “New binding-concealing trade-offs for quantum string commitment,” *Journal of Cryptology*, vol. 24, no. 4, pp. 579–592, 2008.

[7] ——, “Communication complexity of remote state preparation with entanglement,” *Quantum Information and Computation*, vol. 6, no. 4–5, pp. 461–464, Jul. 2006.

[8] R. Jain, H. Klauck, and A. Nayak, “Direct product theorems for communication complexity via subdistribution bounds,” in *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*. New York, NY: ACM Press, May 17–20, 2008, pp. 599–608.

[9] R. Jain, “New strong direct product results in communication complexity,” Electronic Colloquium on Computational Complexity, `http://www.eccc.uni-trier.de/`, Tech. Rep. TR11-024, Feb. 25, 2011.

[10] R. Jain, A. Nayak, and Y. Su, “A separation between divergence and Holevo information for ensembles,” *Mathematical Structures in Computer Science*, vol. 20, no. 5, pp. 977–993, Oct.27, 2010, special issue devoted to selected articles from *Theory and Applications of Models of Computation (TAMC 2008, 2009)*.

[11] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

[12] A. Winter, “Coding theorem and strong converse for quantum channels,” *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2481–2485, Nov. 1999.

[13] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, UK: Cambridge University Press, 2000.

[14] J. Watrous, “Lecture notes, CS 766/QIC 820 Theory of Quantum Information, University of Waterloo,” Fall 2011, see lecture 7, Semidefinite programming, and Lecture 8, Semidefinite Programs for Fidelity and Optimal Measurements. [Online]. Available: http://www.cs.uwaterloo.ca/˜watrous/CS766/

[15] R. Bhatia, *Matrix Analysis*, ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 1997, vol. 169.

[16] L. Lovász, “Semidefinite programs and combinatorial optimization,” in *Recent Advances in Algorithms and Combinatorics*, ser. CMS Books Math./Ouvrages Math. SMC, B. A. Reed and C. Linhares-Sales, Eds. New York: Springer, 2003, vol. 11, ch. 6, pp. 137–194.