Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 157 (2024)

On estimating the trace of quantum state powers

Yupan Liu*[1] and Qisheng Wang†[2,1]

[1]Graduate School of Mathematics, Nagoya University
[2]School of Informatics, University of Edinburgh

Abstract

We investigate the computational complexity of estimating the trace of quantum state powers $\mathrm{tr}(\rho^q)$ for an n-qubit mixed quantum state ρ, given its state-preparation circuit of size $\mathrm{poly}(n)$. This quantity is closely related to and often interchangeable with the Tsallis entropy $\mathrm{S}_q(\rho) = \frac{1-\mathrm{tr}(\rho^q)}{q-1}$, where $q = 1$ corresponds to the von Neumann entropy. For any non-integer $q \geq 1 + \Omega(1)$, we provide a quantum estimator for $\mathrm{S}_q(\rho)$ with time complexity $\mathrm{poly}(n)$, *exponentially* improving the prior best results of $\exp(n)$ due to Acharya, Issa, Shende, and Wagner (ISIT 2019), Wang, Guan, Liu, Zhang, and Ying (TIT 2024), Wang, Zhang, and Li (TIT 2024), and Wang and Zhang (ESA 2024). Our speedup is achieved by introducing efficiently computable *uniform approximations* of positive power functions into quantum singular value transformation.

Our quantum algorithm reveals a sharp phase transition between the case of $q = 1$ and constant $q > 1$ in the computational complexity of the Quantum q-Tsallis Entropy Difference Problem ($\mathrm{TsallisQED}_q$), particularly deciding whether the difference $\mathrm{S}_q(\rho_0) - \mathrm{S}_q(\rho_1)$ is at least 0.001 or at most -0.001:

- For any $1+\Omega(1) \leq q \leq 2$, $\mathrm{TsallisQED}_q$ is BQP-complete, which implies that Purity Estimation is also BQP-complete.
- For any $1 \leq q \leq 1+\frac{1}{n-1}$, $\mathrm{TsallisQED}_q$ is QSZK-hard, leading to hardness of approximating the von Neumann entropy because $\mathrm{S}_q(\rho) \leq \mathrm{S}(\rho)$, as long as $\mathsf{BQP} \subsetneq \mathsf{QSZK}$.

The hardness results are derived from reductions based on new inequalities for the quantum q-Jensen-(Shannon-)Tsallis divergence with $1 \leq q \leq 2$, which are of independent interest.

*Email: yupan.liu.e6@math.nagoya-u.ac.jp
†Email: QishengWang1994@gmail.com

ISSN 1433-8092

Contents

1 Introduction

In recent years, the development of quantum devices has posed an intriguing challenge of verifying their intended functionality. Typically, a quantum device is designed to prepare an n-qubit (mixed) state ρ. The problem of (tolerant) quantum state testing aims to design algorithms that can efficiently test whether a quantum state approximately has a certain property, assuming the state either nearly has the property or is somehow "far" from having it. This problem is a quantum (non-commutative) generalization of classical (tolerant) distribution testing (see [Can20]) and classical property testing in general (see [Gol17]). Furthermore, this problem is an instance of the emerging field of quantum property testing (see [MdW16]), which focuses on devising (efficient) quantum testers for properties of quantum objects.

The general upper bound for (tolerant) quantum state testing depends (at least) linearly on the dimension (e.g., [MdW16, Section 4.2]), whereas some properties of quantum states can be tested significantly more efficiently than the general case. A simple and interesting example is the property PURITY, where ρ satisfies the property if and only if it is a pure state. This example is essentially an instance of estimating the trace of quantum state powers, specifically $\mathrm{tr}(\rho^2)$. A natural approach to test PURITY is to apply the SWAP test [BCWdW01] to two copies of ρ, and this algorithm accepts with probability $(1+\mathrm{tr}(\rho^2))/2$, which is equal to 1 if and only if ρ is pure. Further analysis deduces that PURITY can be tolerantly tested with $O(1/\epsilon^2)$ copies of ρ.[1] Meanwhile, Ekert et al. [EAO+02] presented an efficient quantum algorithm for estimating $\mathrm{tr}(\rho^q)$ where $q > 1$ is an integer. These two fundamental works raise two interesting questions:

(i) Is there an efficient quantum algorithm for estimating the trace of quantum state powers $\mathrm{tr}(\rho^q)$ for any non-integer $q > 1$?

(ii) Can estimating the trace of quantum state powers, e.g., $\mathrm{tr}(\rho^2)$, fully capture the computational power of quantum computing, namely BQP-complete?

Notably, the trace of quantum state powers $\mathrm{tr}(\rho^q)$ is closely related to the *power* quantum entropy of order q. Particularly, the quantum q-Tsallis entropy $\mathrm{S}_q(\rho)$, which is a non-additive (but still concave) generalization of the von Neumann entropy $\mathrm{S}(\rho)$, with the von Neumann entropy being the limiting case of the quantum q-Tsallis entropy as q approaches 1:

$$\mathrm{S}_q(\rho) = \frac{1-\mathrm{tr}(\rho^q)}{q-1} \quad \text{and} \quad \lim_{q\to 1} \mathrm{S}_q(\rho) = \mathrm{S}(\rho) = -\mathrm{tr}(\rho\log(\rho)).$$

As a consequence, $\mathrm{S}_q(\rho)$ can naturally provide a lower bound for $\mathrm{S}(\rho)$ when considering $\mathrm{S}_q(\rho)$ with $q = 1+\epsilon$, where ϵ can be a small constant, such as $q = 1.0001$. This observation serves as the first reason motivating Question (i).

The study of power entropy dates back to Havrda and Charvát [HC67]. Since then, it has been rediscovered independently by Daróczy [Dar70], and finally popularized by Tsallis [Tsa88]. Raggio [Rag95] expanded on this study by introducing the quantum Tsallis entropy. Tsallis entropy has been particularly useful in physics for describing systems with non-extensive properties, such as long-range interactions, in statistical mechanics (see [Tsa01]).

A notable example is the Tsallis entropy $\mathrm{H}_q(p)$ with $q = 3/2$, which is useful for modeling systems where both frequent and rare events matter.[2] For instance, in fluid dynamics, the distribution that maximizes $\mathrm{H}_{3/2}$ helps model velocity changes in turbulent flows [Bec02]. This example provides the second reason motivating Question (i), as existing efficient quantum algorithms [BCWdW01, EAO+02] are designed only for integer $q \geq 2$. Estimating $\mathrm{S}_q(\rho)$ for non-integer q between 1 and 2, therefore, appears to be computationally challenging.

In this paper, we focus on estimating the trace of quantum state powers, or equivalently,

[1] The sample (or query) complexity for PURITY differs between one-sided or two-sided error scenarios. Our upper bound applies to the later, while the sample complexity for the former is $O(1/\epsilon)$ [MdW16, Section 4.2].

[2] In contrast, the Tsallis entropy with $q = 2$ (Gini impurity) is very sensitive to rare events.

the Quantum q-Tsallis Entropy Difference Problem (TsallisQED_q) and the Quantum q-Tsallis Entropy Approximation Problem (TsallisQEA_q). These two problems constitute the (white-box) quantum state testing problem with respect to the quantum q-Tsallis entropy. For TsallisQED_q, we consider two polynomial-size quantum circuits (devices), denoted as Q_0 and Q_1, which prepare n-qubit quantum states ρ_0 and ρ_1, respectively, with access to the descriptions of these circuits. Our goal is to decide whether the difference $\mathrm{S}_q(\rho_0) - \mathrm{S}_q(\rho_1)$ is at least 0.001 or at most -0.001.[3] The setting of TsallisQEA_q is similar to TsallisQED_q, except that we only consider a single n-qubit quantum state ρ, and the task is to decide whether the difference $\mathrm{S}_q(\rho) - t(n)$ is at least 0.001 or at most -0.001, where $t(n)$ is a known threshold.

Next, we will state our main results and then provide justifications for their significance.

1.1 Main results

We begin by presenting our first main result, which provides a positive answer to Question (i) for the regime $q \geq 1 + \Omega(1)$:[4]

Theorem 1.1 (Quantum estimator for q-Tsallis entropy)**.** *Given quantum query access to the state-preparation circuit of an n-qubit quantum state ρ, for any $q \geq 1+\Omega(1)$, there is a quantum algorithm for estimating $\mathrm{S}_q(\rho)$ to additive error 0.001 with query complexity $O(1)$. Moreover, if the description of the state-preparation circuit is of size $\mathrm{poly}(n)$, then the time complexity of the quantum algorithm is $\mathrm{poly}(n)$. Consequently, for any $q \geq 1 + \Omega(1)$, TsallisQED_q and TsallisQEA_q are in BQP.*

More specifically, when the desired additive error is set to ϵ, the explicit query complexity of Theorem 1.1 becomes $O(1/\epsilon^{1+\frac{1}{q-1}})$, or expressed as $\mathrm{poly}(1/\epsilon)$ (see Theorem 3.2). Moreover, if the state-preparation circuit of ρ is of size $L(n) = \mathrm{poly}(n)$, Theorem 1.1 provides a quantum algorithm with time complexity $O(L/\epsilon^{1+\frac{1}{q-1}})$, or equivalently, $\mathrm{poly}(n, 1/\epsilon)$. Using the same idea, we can also derive an upper bound $\widetilde{O}(1/\epsilon^{3+\frac{2}{q-1}})$, or expressed as $\mathrm{poly}(1/\epsilon)$, for the sample complexity needed to estimate $\mathrm{S}_q(\rho)$ (see Theorem 3.3). This is achieved by applying the *samplizer* from [WZ24c], which allows a quantum query-to-sample simulation.

There are several quantum algorithms for estimating the q-Tsallis entropy of an n-qubit mixed quantum state ρ for non-integer constant $q > 1$ proposed in [AISW20, WGL+24, WZL24, WZ24c], all of which turn out to have time complexity $\exp(n)$ in the setting that ρ is given by its state-preparation circuit of size $\mathrm{poly}(n)$.

- In [AISW20, Theorem 3] and [WZ24c, Theorem 1.2], for non-integer constant $q > 1$, they proposed quantum algorithms for estimating the q-Rényi entropy of an n-qubit quantum state ρ by using $\mathsf{S} = \mathrm{poly}(1/\epsilon) \cdot \exp(n)$ samples of ρ and $\mathsf{T} = \mathrm{poly}(1/\epsilon) \cdot \exp(n)$ quantum gates.[5] Their result implies an estimator for $\mathrm{S}_q(\rho)$ with the same complexity, because any estimator for q-Rényi entropy implies an estimator for q-Tsallis entropy with the same parameter for $q > 1$ (as noted in [AOST17, Appendix A]). By preparing each sample of ρ using its state-preparation circuit of size $\mathrm{poly}(n)$, one can estimate $\mathrm{S}_q(\rho)$ by using their estimators with overall time complexity $\mathsf{S} \cdot \mathrm{poly}(n) + \mathsf{T} = \mathrm{poly}(1/\epsilon) \cdot \exp(n)$.

[3] It is noteworthy that 0.001 is just an arbitrary constant for the precision parameter, which can be replaced by any inverse polynomial function in general. See Definition 5.1 and Definition 5.2 for formal definitions.

[4] We implicitly assume that q satisfies $1+\Omega(1) \leq q \leq O(1)$. Since $\mathrm{S}_q(\rho) \leq o(1)$ when $q = \omega(1)$, it is reasonable to consider constantly large q.

[5] The explicit sample complexities of the approaches of [AISW20, Theorem 3] and [WZ24c, Theorem 2] are $O(2^{2n}/\epsilon^2)$ and $O(2^{(\frac{4}{q}-2)n}/\epsilon^{1+\frac{4}{q}} \cdot \mathrm{poly}(n, \log(1/\epsilon)))$, respectively, both of which are $\mathrm{poly}(1/\epsilon) \cdot \exp(n)$. The number of quantum states in the approach of [AISW20, Theorem 3] was mentioned in [WZ24c] to be $O((2^{2n}/\epsilon^2)^3 \cdot \mathrm{polylog}(2^n, 1/\epsilon)) = \mathrm{poly}(1/\epsilon) \cdot \exp(n)$ by using the weak Schur sampling in [MdW16, Section 4.2.2] and the quantum Fourier transform over symmetric groups [KS16]. Another possible implementation noted in [Hay24] is to use the Schur transform in [Ngu23], resulting in $O(2^{2n}/\epsilon^2 \cdot 2^{4n} \cdot \mathrm{polylog}(2^n, 1/\epsilon)) = \mathrm{poly}(1/\epsilon) \cdot \exp(n)$.

- In [WGL+24, Theorem III.9], for non-integer constant $q > 1$, they proposed a quantum algorithm for estimating $\mathrm{S}_q(\rho)$ with query complexity $\widetilde{O}(r^{1/\{\frac{q-1}{2}\}}/\epsilon^{1+1/\{\frac{q-1}{2}\}}) = \mathrm{poly}(r, 1/\epsilon)$, where r is (an upper bound on) the rank of ρ and $\{x\} \coloneqq x - \lfloor x \rfloor$ denotes the fractional part of x. In [WZL24, Corollary 5], for non-integer constant $q > 1$, they proposed a quantum algorithm for estimating the q-Rényi entropy of a quantum state with query complexity $\widetilde{O}(r/\epsilon^{1+\frac{1}{q}}) = \mathrm{poly}(r, 1/\epsilon)$, which also implies a quantum algorithm for estimating $\mathrm{S}_q(\rho)$ with query complexity $\mathrm{poly}(r, 1/\epsilon)$ (the reason has been discussed in the last item). For n-qubit quantum state ρ without prior knowledge, by taking $r = 2^n$, their query compelxity is then $\mathrm{poly}(2^n, 1/\epsilon) = \mathrm{poly}(1/\epsilon) \cdot \exp(n)$, which is exponentially larger than our $\mathrm{poly}(n, 1/\epsilon)$.

Our efficient quantum estimator for $\mathrm{S}_q(\rho)$ where $q \geq 1 + \Omega(1)$ (Theorem 1.1), combined with our hardness results for $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$ (Theorem 1.2), indicates a sharp phase transition between the case of $q = 1$ and constant $q > 1$ and answers to Question (i) and (ii). For clarity, we summarize our main results in Table 1.

	$q = 1$	$1 < q \leq 1 + \frac{1}{n-1}$	$1 + \Omega(1) \leq q \leq 2$	$q > 2$
$\textsc{TsallisQED}_q$	QSZK-complete [BASTS10]	QSZK-hard Theorem 1.2(2)	BQP-complete Theorem 1.1 and Theorem 1.2(1)	in BQP Theorem 1.1
$\textsc{TsallisQEA}_q$	NIQSZK-complete [BASTS10, CCKV08]	NIQSZK-hard* Theorem 1.2(2)	BQP-complete Theorem 1.1 and Theorem 1.2(1)	in BQP Theorem 1.1

Table 1: Computational hardness of $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$.

Here, QSZK and NIQSZK are the classes of promise problems possessing quantum statistical zero-knowledge and non-interactive quantum statistical zero-knowledge, respectively, as introduced by [Wat02, Wat09] and [Kob03]. The asterisk in Table 1 indicates that $\textsc{TsallisQEA}_q$ is NIQSZK-hard for a specific $q(n) = 1 + \frac{1}{n-1}$, as detailed in Theorem 1.2(2).

For the case of $q = 1$, $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$ coincide with the Quantum Entropy Difference Problem (QED) and the Quantum Entropy Approximation Problem (QEA) introduced in [BASTS10], respectively. Moreover, QED is complete for the class QSZK [BASTS10], whereas QEA is complete for the class NIQSZK [BASTS10, CCKV08]. These two classes contain BQP and are seemingly much harder than BQP.[6] Meanwhile, the best known upper bound for QSZK is QIP(2) with a quantum linear-space honest prover [LGLW23], and the best known upper bound for NIQSZK is qq-QAM [KLGN19], both of which are contained in $\mathsf{QIP(2)} \subseteq \mathsf{PSPACE}$ [JUW09].

In terms of *quantitative* bounds on quantum query and sample complexities, QSZK-hard or NIQSZK-hard in the white-box setting correspond to rank-dependent complexities in black-box settings. Specifically, we establish lower bounds for both the easy regime $q \geq 1 + \Omega(1)$ and the hard regime $1 < q \leq 1 + \frac{1}{n-1}$, with the upper bounds for the hard regime derived from those for estimating quantum Rényi entropy, as detailed in Table 2.

On the other hand, understanding why the regime $q \geq 1 + \Omega(1)$ is computationally easy can be illustrated by the case of $q = 2$ (Purity Estimation), particularly deciding whether $\mathrm{tr}(\rho^2)$ is at least $2/3$ or at most $1/3$. Let $\{\lambda_k\}_{k \in [2^n]}$ be the eigenvalues of an n-qubit quantum state ρ.

[6] Following the oracle separation between NISZK and PP [BCH+19], it holds that $\mathsf{NIQSZK}^{\mathcal{O}} \not\subseteq \mathsf{PP}^{\mathcal{O}}$ and likewise $\mathsf{QSZK}^{\mathcal{O}} \not\subseteq \mathsf{PP}^{\mathcal{O}}$ for some classical oracle $\mathcal{O}$.

[7] In these bounds, $c > 0$ is a constant that can be made arbitrarily small, and we set $c' = 3c$.

[8] In the regime $1 \leq q \leq 1 + \frac{1}{n-1}$, as the rank r approaches 2^n, a sample complexity upper bound of $O(4^n/\epsilon^2)$ with better dependence on ϵ was given in [AISW20].

[9] As the rank r approaches 2^n, a better query complexity upper bound of $\widetilde{O}(2^n/\epsilon^{1.5})$ was shown in [GL20].

Regime of q	Query Complexity		Sample Complexity	
	Upper Bound	Lower Bound	Upper Bound	Lower Bound
$q \geq 1 + \Omega(1)$	$O(1/\epsilon^{1+\frac{1}{q-1}})$ Theorem 3.2	$\Omega(1/\sqrt{\epsilon})$ Theorem 5.12	$\widetilde{O}(1/\epsilon^{3+\frac{2}{q-1}})$ Theorem 3.3	$\Omega(1/\epsilon)$ Theorem 5.15
$1 < q \leq 1 + \frac{1}{n-1}$	$\widetilde{O}(r/\epsilon^2)$ [WZL24]	$\Omega(r^{0.17-c})$[7] Theorem 5.13	$\widetilde{O}(r^2/\epsilon^5)$[8] [WZ24c]	$\Omega(r^{0.51-c'})$[7] Theorem 5.16
$q = 1$	$\widetilde{O}(r/\epsilon^2)$[9] [WGL+24]	$\widetilde{\Omega}(\sqrt{r})$ [BKT20]	$\widetilde{O}(r^2/\epsilon^5)$[8] [WZ24c]	$\Omega(r/\epsilon)$ [WZ24c]

Table 2: (Rank-dependent) bounds on query and sample complexities for estimating $\mathrm{S}_q(\rho)$.

For any quantum state $\hat{\rho}$ having eigenvalues at most $1/n$, it follows that $\mathrm{tr}(\hat{\rho}^2) = \sum_{k\in[2^n]} \lambda_k^2 \leq n \cdot n^{-2} = 1/n$, hence 0 provides a good estimate of $\mathrm{tr}(\hat{\rho}^2)$ to within additive error $1/3$. This intuition implies that only sufficiently large eigenvalues contribute to estimating the value of $\mathrm{tr}(\rho^2)$. Consequently, the computational complexity of Purity Estimation is supposed to be independent of the rank r.

However, this argument is just the first step towards establishing an efficient quantum estimator for $\mathrm{S}_q(\rho)$.[10] We also need to estimate $\sum_{k\in\mathcal{I}_{\mathrm{large}}} \lambda_k^q$, where $\mathcal{I}_{\mathrm{large}}$ is the index set for sufficiently large eigenvalue λ_k. For the case of integer $q > 1$, the approach of [BCWdW01, EAO+02] equipped with quantum amplitude estimation [BHMT02] provides a solution, whereas the case of non-integer $q \geq 1 + \Omega(1)$ is more challenging and requires more sophisticated techniques. Notably, the task is finally resolved by our first main result (Theorem 1.1).

Lastly, we provide our second main result, namely the computational hardness for TsallisQED$_q$ and TsallisQEA$_q$, as stated in Theorem 1.2. Let ConstRankTsallisQED$_q$ and ConstRankTsallisQEA$_q$ denote restricted variants of TsallisQED$_q$ and TsallisQEA$_q$, respectively, such that the ranks of the states of interest are at most $O(1)$.

Theorem 1.2 (Computational hardness for TsallisQED$_q$ and TsallisQEA$_q$, informal)**.** *The promise problems* TsallisQED$_q$ *and* TsallisQEA$_q$ *capture the computational power of their respective complexity classes in the corresponding regimes of* q*:*[11]

(1) **Easy regimes:** *For any* $q \in [1, 2]$, ConstRankTsallisQED$_q$ *is* BQP*-hard under Karp reduction, and consequently,* ConstRankTsallisQEA$_q$ *is* BQP*-hard under Turing reduction. As a corollary,* TsallisQED$_q$ *and* TsallisQEA$_q$ *are* BQP*-complete for* $1 + \Omega(1) \leq q \leq 2$.

(2) **Hard regimes:** *For any* $q \in \left(1, 1 + \frac{1}{n-1}\right]$, TsallisQED$_q$ *is* QSZK*-hard under Karp reduction, and consequently,* TsallisQEA$_q$ *is* QSZK*-hard under Turing reduction. Furthermore, for* $q = 1 + \frac{1}{n-1}$, TsallisQEA$_q$ *is* NIQSZK*-hard under Karp reduction.*

It is noteworthy that BQP-hardness under Turing reduction is as strong as BQP-hardness under Karp reduction, due to the BQP subroutine theorem [BBBV97].[12] Moreover, Theorem 1.2 implies a direct corollary, offering a positive answer to Question (ii):

Corollary 1.3. Purity Estimation *is* BQP*-hard.*

[10] A similar argument also applies to the classical Tsallis entropy, see [AOST17, Section III.C]. Nevertheless, this type of argument does not extend to von Neumann entropy ($q = 1$), see [QKW24, Section 7].

[11] For detailed definitions of Karp reduction and Turing reduction, please refer to Section 2.2.1.

[12] Once we have an efficient quantum algorithm $\mathcal{A}$ for TsallisQEA$_q$, any problem in BQP can be solved using $\mathcal{A}$ as a subroutine. The BQP subroutine theorem, as stated in [BBBV97, Section 4], implies that $\mathsf{BQP}^{\mathcal{A}} \subseteq \mathsf{BQP}$.

Interestingly, the BQP-hardness for a similar problem, specifically deciding whether $\mathrm{tr}(\rho_0\rho_1)$ is at least $2/3$ or at most $1/3$, turns out to be not difficult to show.[13] However, this result does not imply Corollary 1.3.

1.2 Proof techniques: BQP containment for q constantly larger than 1

The proof of Theorem 1.1 consists of an efficient quantum (query) algorithm for estimating the value of $\mathrm{tr}(\rho^q)$ for $q > 1$, given quantum query access to the state-preparation circuit Q of the mixed quantum state ρ. Our approach to estimating $\mathrm{tr}(\rho^q)$ is via one-bit precision phase estimation [Kit95], also known as the Hadamard test [AJL09], equipped with the quantum singular value transformation (QSVT) [GSLW19]. Our algorithm is sketched in the following four steps (see Section 3 for more details):

1. Find a good polynomial approximation of x^{q-1}.
2. Implement a unitary block-encoding U of ρ^{q-1} using QSVT, with the state-preparation circuit Q.
3. Perform the Hadamard test on U and ρ with outcome $b \in \{0, 1\}$.
4. One can learn the value of $\mathrm{tr}(\rho^q)$ from a good estimate of b via quantum amplitude estimation.

The idea is simple. Similar ideas were ever used to estimate the fidelity [GP22], trace distance [WZ24a, LGLW23], and von Neumann entropy [LGLW23, WZ24c]. However, all of the aforementioned quantum algorithms have query or time complexity polynomials in the rank r of quantum states. Additionally, all these prior works rely on the quantum singular value transformation [GSLW19], which is a technique for designing quantum algorithms by approximating the target functions.[14] The main technical reason is that the functions to be approximated in their key steps are not smooth in the whole range of $[0, 1]$, so they have to use the polynomial approximations of piece-wise smooth functions in [GSLW19, Corollary 23] to avoid the bad part (which is actually the regime of tiny eigenvalues);[15] this results in an estimation error dependent on r because, technically, the error for each bad eigenvalue has to be bounded individually (there are at most r bad eigenvalues), thereby introducing an (at least) linear r-dependence. Specifically, in their approaches, a target function $f(x)$ is specified and the goal is to estimate the value of $\mathrm{tr}(\rho f(\rho))$. For example, $f(x) = -\log(x)$ for estimating the von Neumann entropy. The target function $f(x)$ is usually only approximated well in the range $x \in [\delta, 1]$ for some parameter δ, while leaving the rest range of x unspecified; more precisely, $f(x)$ is approximated by a polynomial $P(x)$ by, e.g., [GSLW19, Corollary 23], such that

$$\max_{x\in[\delta,1]} |P(x) - f(x)| \le \epsilon, \quad \max_{x\in[-1,1]} |P(x)| \le 1, \text{ and } \deg(P) = O\left(\frac{1}{\delta}\log\frac{1}{\epsilon}\right). \tag{1.1}$$

Then, they instead estimate the value of $\mathrm{tr}(\rho P(\rho))$. The intrinsic error turns out to be

$$|\mathrm{tr}(\rho f(\rho)) - \mathrm{tr}(\rho P(\rho))| \le \sum_{\lambda_j<\delta} |\lambda_j f(\lambda_j) - \lambda_j P(\lambda_j)| + \sum_{\lambda_j\ge\delta} |\lambda_j f(\lambda_j) - \lambda_j P(\lambda_j)| \le r \cdot \mathrm{poly}(\delta) + O(\epsilon).$$

Here, $\{\lambda_j\}_{1\le j\le 2^n}$ are the eigenvalues of the state ρ, with each λ_j satisfying $0 \le \lambda_j \le 1$. To make the intrinsic error bounded, δ must be sufficiently small, e.g., $\delta = 1/\mathrm{poly}(r)$.

The above standard method has drawbacks: the intrinsic error is $r \cdot \mathrm{poly}(\delta)$ for the small-eigenvalue part and $O(\epsilon)$ for the large-eigenvalue part. While the ϵ-dependence in the approxi-

[13] For any BQP circuit C_x, the acceptance probability $\||1\rangle\langle 1|_{\mathrm{out}} C_x|\bar{0}\rangle\|_2^2 = \mathrm{tr}\big(|1\rangle\langle 1|_{\mathrm{out}} C_x|\bar{0}\rangle\langle\bar{0}|C_x^\dagger\big) = \mathrm{tr}(\rho_0\rho_1)$, where $\rho_0 \coloneqq |1\rangle\langle 1|_{\mathrm{out}}$ and $\rho_1 \coloneqq \mathrm{tr}_{\overline{\mathrm{out}}}\big(C_x|\bar{0}\rangle\langle\bar{0}|C_x^\dagger\big)$. Similar observations appeared in [Kob03, Theorem 9].

[14] For example, estimating the fidelity and trace distance requires to approximate the sign function; and estimating the von Neumann entropy requires to approximate the logarithmic function.

[15] These eigenvalues correspond to the inputs of the target function.

mation degree is logarithmic (and thus not the dominating term), the δ-dependence is significant. This suggests the need for the following trade-off: Can we reduce the error caused by the small-eigenvalue part, at the cost of a possibly worse error caused by the large-eigenvalue part?

To make this trade-off possible for our purpose, we turn to find polynomials that uniformly approximate the positive power functions. This is inspired by the Stone-Weierstrass theorem, stating that any continuous function (e.g., x^q) on a closed interval (e.g., $[0,1]$) can be uniformly approximated by polynomials. The study of the *best uniform approximation* (by polynomials)[16] of positive power functions was initiated by Bernstein [Ber14, Ber38] almost a century ago in an abstract manner.[17] The best uniform approximation polynomial of x^q was shown with a non-constructive proof in [Tim63, Section 7.1.41], stating that there is a family of polynomials $P_d(x)$ of degree d such that

$$\max_{x\in[0,1]} |P_d(x) - x^q| \to \frac{1}{d^q}, \text{ as } d \to \infty, \tag{1.2}$$

whose approximation range is in sharp contrast to that in Equation (1.1). However, the coefficients of the leading error terms and the explicit construction of these polynomial approximations seem still not fully understood (e.g., [Gan02]). Consequently, it is somewhat challenging to directly use such polynomial approximations (e.g., [Tim63, Section 7.1.41]) in a time-efficient manner.

Inspired by the result of the best uniform approximation of positive power functions in [Tim63], we, instead, aim to find a good enough uniform approximation that is also efficiently computable. This is achieved by employing the construction of asymptotically best uniform approximation via combining Chebyshev truncations and the de La Vallée Poussin partial sum (cf. [Riv90, Chapter 3]). Finally, we obtain a family of efficiently computable uniform approximation polynomials of (scaled) x^q that are suitable for QSVT:

$$\max_{x\in[0,1]} \left| P(x) - \frac{1}{2}x^q \right| \leq \epsilon, \quad \max_{x\in[-1,1]} |P(x)| \leq 1, \text{ and } \deg(P) = O\left(\frac{1}{\epsilon^{1/q}}\right). \tag{1.3}$$

Using these efficiently computable uniform approximation polynomials, we are able to give a quantum algorithm for estimating $\mathrm{tr}(\rho^q)$. First, we approximate the function x^{q-1} in the range $[0,1]$ to error ϵ by a polynomial of degree $O(1/\epsilon^{\frac{1}{q-1}})$. Then, we can apply the algorithm sketched at the very beginning of this subsection. With further analysis, we can estimate the value of $\mathrm{tr}(\rho^q)$ to additive error ϵ with quantum query complexity $O(1/\epsilon^{1+\frac{1}{q-1}})$ (see Theorem 3.2). Using the same idea, we can also estimate $\mathrm{tr}(\rho^q)$ to additive error ϵ by using $\widetilde{O}(1/\epsilon^{3+\frac{2}{q-1}})$ copies of ρ through the samplizer [WZ24c] (see Theorem 3.3).

To conclude this subsection, it can be seen that our quantum algorithm for estimating $\mathrm{tr}(\rho^q)$ is naturally applicable to solving $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$. Particularly for the precision in the regime $1/\operatorname{poly}(n) \leq \epsilon \leq 1$, the efficiently-computability of the uniform approximation polynomials in Equation (1.3) ensures that the description of the quantum circuit of our algorithm can be computed by a classical deterministic Turing machine in $\operatorname{poly}(n)$ time, which is a significant step to show the BQP-completeness of $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$ for $1+\Omega(1) \leq q \leq 2$ and precision $1/\operatorname{poly}(n) \leq \epsilon \leq 1$.

1.3 Proof techniques: Hardness via QJT_q-based reductions

Before we proceed with the proof of Theorem 1.2, we start by defining the (white-box) quantum state testing problem with respect to the trace distance, which was first introduced in [Wat02].

[16] The best uniform approximation polynomial of a continuous function $f(x)$ on $[-1,1]$ is a degree-d polynomial that minimizes $\max_{x\in[-1,1]} |f(x) - P_d(x)|$ over all degree-d polynomials P_d. For a formal definition, see Section 2.3.1.

[17] Actually, the function $|x|^q$ for $x \in [-1,1]$ is commonly considered in the literature. Nevertheless, we are only interested in the non-negative part, i.e., the range $[0,1]$.

Let ρ_0 and ρ_1 be n-qubit quantum states such that their purifications can be prepared by polynomial-size quantum circuits Q_0 and Q_1, respectively. The Quantum State Distinguishability Problem (QSD) is to decide whether the trace distance $\mathrm{T}(\rho_0, \rho_1)$ is at least $1-\epsilon(n)$ or at most $\epsilon(n)$. Furthermore, we need the other two restricted versions of QSD, see Section 2.2 for formal definitions:

- PureQSD: Both ρ_0 and ρ_1 are pure states.
- QSCMM: ρ_1 is fixed to be the n-qubit maximally mixed state.[18]

The proof of Theorem 1.2, particularly the hardness results under Karp reduction, utilizes reductions from the aforementioned variants of QSD to $\textsc{TsallisQED}_q$ or $\textsc{TsallisQEA}_q$ for the respective ranges of q. Next, we will specify two main technical challenges related to the corresponding inequalities necessary for establishing Theorem 1.2:

(1) For $\textsc{ConstRankTsallisQED}_q$ and $\textsc{TsallisQED}_q$, the key ingredient of these reductions is the quantum q-Jensen-(Shannon-)Tsallis divergence (QJT_q, see Definition 2.10), first introduced in [BH09]. We notice that QJT_q can be viewed as a *distance version* of the quantum q-Tsallis entropy difference for $1 \leq q \leq 2$,[19] and consequently, these reductions heavily rely on the inequalities between QJT_q and the trace distance. However, such inequalities are only known for the case of $q = 1$ [FvdG99, Hol73a, BH09], presenting the first technical challenge.

(2) For $\textsc{TsallisQEA}_q$, the reduction essentially relies on the lower and upper bounds on the quantum q-Tsallis entropy of a quantum state ρ in terms of the trace distance between the state and the maximally mixed state, when the trace distance is promised to be a fixed value. These bounds are also only known for the case of $q = 1$ [Vaj70, CCKV08, KLGN19], leading to the second technical challenge.

For clarity, we summarize the correspondence between our reductions for establishing Theorem 1.2 and the new inequalities in Table 3, where the q-logarithm $\ln_q(x) := \frac{1-x^{1-q}}{q-1}$.[20]

Problem	Regime of q	Reduction from	New inequalities
$\textsc{ConstRank}$ $\textsc{TsallisQED}_q$ Theorem 1.2(1)	$1 \leq q \leq 2$	PureQSD is BQP-hard adapted from [RASW23]	$\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1-\mathrm{T}}{2}\right) \leq \mathrm{QJT}_q \leq \mathrm{H}_q\left(\frac{1}{2}\right)\mathrm{T}^q$ Theorem 4.1
$\textsc{TsallisQED}_q$ Theorem 1.2(2)	$1 \leq q \leq 1+\frac{1}{n-1}$	QSD is QSZK-hard [Wat02, Wat09]	$\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1-\mathrm{T}}{2}\right) \leq \mathrm{QJT}_q$ Theorem 4.1
$\textsc{TsallisQEA}_q$ Theorem 1.2(2)	$q = 1+\frac{1}{n-1}$	QSCMM is NIQSZK-hard [Kob03, BASTS10, CCKV08]	$\left(1-\mathrm{T}-\frac{1}{2^n}\right)\ln_q(2^n) \leq \mathrm{S}_q \leq \ln_q(2^n(1-\mathrm{T}))$ Lemma 4.10

Table 3: Reductions for $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$, and the corresponding inequalities.

Once we have established these new inequalities, together with our new bounds for the Tsallis binary entropy $\mathrm{H}_q(x) \leq \mathrm{H}_q\left(\frac{1}{2}\right)\sqrt{4x(1-x)}$ (see Theorem 4.2, where previously only the case of $q = 1$ was known [Lin91, Top01]), we can establish our three hardness results under Karp reduction in Theorem 1.2 through relatively complicated and detailed analyses. The additional

[18] Precisely speaking, the problem called Quantum State Closeness to Maximally Mixed State (QSCMM) is to decide whether $\mathrm{T}(\rho, (I/2)^{\otimes n})$ is at most $1/n$ or at least $1-1/n$, which is the complement of QSD concerning the same states of interest.

[19] For the case of $q = 1$, similar observations are implicitly used to show that QED is QSZK-hard [BASTS10], and recently explicitly emphasized in [Liu23], leading to a simple proof for the QSZK hardness of QED.

[20] As q approaches 1, the q-logarithm becomes the natural logarithm. For further details and references on q-logarithm, please refer to the beginning of Section 2.1.

two hardness results for ConstRankTsallisQEA$_q$ and TsallisQEA$_q$ under Turing reduction in Theorem 1.2 follow straightforwardly from a binary search for promise problems.

In the remainder of this subsection, we provide insights into proving the new inequalities in Table 3. The first technical challenge involves establishing the inequalities between QJT$_q$ and the trace distance. The main barrier is to provide the data-processing inequality $\mathrm{QJT}_q(\Phi(\rho_0),\Phi(\rho_1)) \le \mathrm{QJT}_q(\rho_0,\rho_1)$ for $1 < q \le 2$.[21] This implies that applying any quantum channel Φ on states ρ_0 and ρ_1 does not increase the divergence between them. For $q = 1$, the quantum Jensen-Shannon divergence (QJS), defined in [MLP05], can be decomposed into a sum of quantum relative entropy $\mathrm{D}(\rho_0\|\rho_1)$:

$$\mathrm{QJS}(\rho_0,\rho_1) \coloneqq \mathrm{S}\left(\frac{\rho_0+\rho_1}{2}\right) - \frac{\mathrm{S}(\rho_0)+\mathrm{S}(\rho_1)}{2} = \frac{1}{2}\left(\mathrm{D}\left(\rho_0\middle\|\frac{\rho_0+\rho_1}{2}\right) + \mathrm{D}\left(\rho_1\middle\|\frac{\rho_0+\rho_1}{2}\right)\right). \quad (1.4)$$

Since the data-processing inequality (essentially, the joint convexity) for the quantum relative entropy was established decades ago [Lie73, Uhl77], and given the equality in Equation (1.4), it directly follows that the data-processing inequality also holds for QJS. However, a similar decomposition does not apply to the quantum q-Tsallis entropy when $q \neq 1$. Fortunately, the joint convexity of QJT$_q$ for $1 \le q \le 2$, specifically,

$$\mathrm{QJT}_q\big((1-\lambda)\rho_0+\lambda\rho_0', (1-\lambda)\rho_1+\lambda\rho_1'\big) \le (1-\lambda)\mathrm{QJT}_q\big(\rho_0,\rho_1\big) + \lambda\mathrm{QJT}_q(\rho_0',\rho_1'),$$

was established few years ago [CT14, Vir19], where $0 < \lambda < 1$. As a consequence, once we establish the data-processing inequality for QJT$_q$, we can then generalize the inequalities between QJS and the trace distance to QJT$_q$ for $1 \le q \le 2$, using the same approach applied to QJS.

For the second technical challenge, specifically the bounds for $\mathrm{S}_q(\rho)$ when $\mathrm{T}(\mathrm{S}_q(\rho),(I/2)^{\otimes n}) = \gamma$ is fixed, it suffices to focus on the classical counterpart,[22] as the maximally mixed state commutes with any state ρ. The lower bound can be established by following the approach in [KLGN19] for $q = 1$. On the other hand, the upper bound for $q = 1$ can be derived using Vajda's inequality [Vaj70], but similar results for $q \neq 1$ are unknown. However, by assuming an appropriate condition between q and the fixed distance γ, we can deduce an upper bound analogous to the $q = 1$ case.

1.4 Discussion and open problems

Our first main theorem (Theorem 1.1) provides an efficiently computable lower bound for the von Neumann entropy $\mathrm{S}(\rho)$. This naturally raises the question:

(i) Is there an efficiently computable upper bound for $\mathrm{S}(\rho)$, perhaps based on some relaxed notion of the von Neumann entropy?

The quantum Tsallis entropy $\mathrm{S}_q(\rho)$ in the regime $1 < q < 2$ exhibits distinct behavior compared to both $\mathrm{S}(\rho)$ and $\mathrm{S}_2(\rho) = 1 - \mathrm{tr}(\rho^2)$, leading to another open problem:

(ii) Can we find further applications of estimating $\mathrm{S}_q(\rho)$ in the regime $1 < q < 2$?

Moreover, two open problems arise regarding quantitative bounds and (NI)QSZK containments:

(iii) Can the query and sample bounds in Table 2 be improved, especially for $q \ge 1 + \Omega(1)$?

(iv) Can we establish that TsallisQED$_q$ (or TsallisQEA$_q$) in the regime $1 < q < 1 + \frac{1}{n-1}$, as specified in Theorem 1.2(2), is also contained in QSZK (or NIQSZK)?

[21] We generalize the approach in [BH09] for $q = 1$. Using the data-processing inequality with a measurement channel, we can establish the lower bound via the measured version of QJT$_q$ (see Equation (2.2)) and the classical counterpart inequality for JT$_q$ in [BH09]. For the upper bound, we construct new states $\hat{\rho}_0$ and $\hat{\rho}_1$ with an ancillary qubit, making $\mathrm{QJT}_q(\hat{\rho}_0,\hat{\rho}_1)$ related to the trace distance for $1 < q \le 2$ (and coincide with the trace distance for $q = 1$). Applying the data-processing inequality with the partial trace, we obtain $\mathrm{QJT}_q(\rho_0,\rho_1) \le \mathrm{QJT}_q(\hat{\rho}_0,\hat{\rho}_1)$.

[22] More specifically, let p denote the distribution of the eigenvalues of ρ, and let ν be the uniform distribution over 2^n items. This task is exactly equivalent to proving the bounds for $\mathrm{H}_q(p)$ when $\mathrm{TV}(p,\nu) = \gamma$ is fixed.

Lastly, it is natural to consider generalizations of the von Neumann entropy tighter than $\mathrm{S}_q(\rho)$ for $q > 1$, particularly $\mathrm{S}_q(\rho)$ for $0 < q < 1$ and the quantum Rényi entropy $\mathrm{S}^R_\alpha(\rho) := \frac{\ln \mathrm{tr}(\rho^\alpha)}{1-\alpha}$:

(v) What are the containment and hardness of estimating $\mathrm{S}_q(\rho)$ in the regime $0 < q < 1$?

(vi) Since RÉNYIQEA$_\alpha$ for $1 < \alpha \leq \frac{1}{n-1}$ is *intuitively* QSZK-hard, as per Theorem 1.2(2), can we obtain (rigorous) computational hardness results for estimating $\mathrm{S}^R_\alpha(\rho)$ with $\alpha > 0$?

1.5 Related works

(Quantum) property testing for probability distributions. (Near-)optimal classical estimators are known for Shannon, Rényi, and Tsallis entropies [JVHW15, JVHW17, WY16, AOST17]. Quantum testers for classical probability distributions were initiated in [BHH11]. Quantum algorithms for ℓ_1 distance of probability distributions were investigated in [BHH11, CFMdW10, Mon15, GL20, LWL24]. Quantum estimators for the Shannon and Rényi entropies were proposed in [LW19, GL20, WZL24]. Notably, matching query lower bounds for estimating the Shannon entropy and ℓ_1 distance from the uniform distribution were shown in [BKT20].

Quantum property testing for quantum states. Quantum sample complexities for a series of problems have been studied in the literature. For von Neumann entropy and Rényi entropy estimations, the dimension-dependence was studied in [AISW20], the dependence on the reciprocal of the minimum non-zero eigenvalue of the quantum state was studied in [WZW23], and the rank-dependence and time-efficiency were studied in [WZ24c]. Other problems include tomography [HHJ$^+$17, OW16], spectrum testing [OW21], closeness testing or estimation with respect to fidelity and trace distance [BOW19, GP22, WZ24a, WZ24b, LWWZ25]. Quantum inner product estimation is a basic task and is well-known to be solved by the SWAP test [BCWdW01]. Recently, a distributed quantum algorithm for quantum inner product estimation was given in [ALL22], where they also provided a matching lower bound; this was later generalized to fidelity estimation between pure states with limited quantum communications [AS24]. As a special case of quantum inner product estimation, tight bounds for purity estimation with and without restricted quantum measurements were shown in [CWLY23, GHYZ24, LGDC24].

The quantum query complexities are also extensively studied. For von Neumann entropy estimation, the dimension-dependence was studied in [GL20], the dependence on the reciprocal of the minimum non-zero eigenvalue of the quantum state was studied in [CLW20], the multiplicative error-dependence was studied in [GHS21], and the rank-dependence was studied in [WGL$^+$24]. In [SLLJ24], they presented a rank-dependent estimator for the q-Tsallis entropy with integer q larger than the rank of quantum states. Additionally, a dimension-dependent estimator for the quantities $\{\mathrm{tr}(\rho^k)\}_{k=1}^N$ was given in [WSP$^+$24]. For Rényi entropy estimation, the query complexity was first studied in [SH21], the rank-dependence was studied in [WGL$^+$24], and was later improved in [WZL24]. Other problems include tomography [vACGN23], and the estimations of fidelity and trace distance [WZC$^+$23, WGL$^+$24, GP22, WZ24a, Wan24, LWWZ25].

In [GH20], the Quantum Entropy Difference Problem (with respect to von Neumann entropy) with shallow circuits was shown to have (conditional) hardness. The computational complexity of the space-bounded versions of the Quantum Entropy Difference Problem and Quantum State Distinguishability Problem were studied in [LGLW23].

Quantum algorithms for estimating the Schatten p-norm. Estimating the Schatten p-norm, $\mathrm{tr}(|A|^p)$, of an $O(\log n)$-local Hermitian matrix A on n qubits to within an additive error of $2^{n-p}\epsilon\|A\|^p$, where $\epsilon(n) \leq 1/\mathrm{poly}(n)$ and real $p(n) \leq \mathrm{poly}(n)$, was proven to be DQC1-complete in [CM18]. In [LS20], with a unitary block-encoding of a matrix A, a quantum algorithm was proposed for estimating the Schatten p-norm, $(\mathrm{tr}(|A|^p))^{1/p}$, to relative error ϵ for integer p. This algorithm requires a condition number κ such that $A \geq I/\kappa$, particularly when p is odd.

2 Preliminaries

We assume that the reader is familiar with quantum computation and the theory of quantum information. For an introduction, the textbook [NC10] provides a good starting point.

We adopt the following conventions throughout the paper: (1) we denote $[n] \coloneqq \{1, 2, \ldots, n\}$; (2) we use both notations, $\log(x)$ and $\ln(x)$, to represent the natural logarithm for any $x \in \mathbb{R}^+$; (3) the notation $\widetilde{O}(f)$ is defined as $O(f \operatorname{polylog}(f))$; (4) we utilize the notation $|\bar{0}\rangle$ to represent $|0\rangle^{\otimes a}$ with $a > 1$; and (5) we use $\|A\|$ to denote the operator norm (equivalently, the Schatten ∞-norm) of a matrix A.

Notions on linear maps and quantum channels. We recommend [AS17, Section 2.3] as an introduction on superoperators and quantum channels. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be finite-dimensional Hilbert spaces with $\dim(\mathcal{H}_i) = N_i = 2^{n_i}$ for $i \in \{1, 2\}$. Let $\mathrm{L}(\mathcal{H}_1, \mathcal{H}_2)$ denote linear maps from $\mathcal{H}_1$ to $\mathcal{H}_2$, and specifically, let $\mathrm{L}(\mathcal{H})$ denote linear maps from $\mathcal{H}$ to $\mathcal{H}$. A map $\Phi\colon \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$ is called *self-adjointness-preserving* if $\Phi(X^\dagger) = (\Phi(X))^\dagger$ for any $X \in \mathrm{L}(\mathcal{H}_1)$. We further say that a self-adjointness-preserving map $\Phi\colon \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$ is a *quantum channel* if Φ is a completely positive trace-preserving map. Here, a map Φ is *trace-preserving* if $\operatorname{tr}(\Phi(X)) = \operatorname{tr}(X)$ for any $X \in \mathrm{L}(\mathcal{H}_1)$. Let $\{|v_i\rangle\}_{i \in [N_1]}$ be an orthonormal basis of $\mathcal{H}_1$, and a map Φ is *completely positive* if $\Phi \otimes I_n$ is positive for any $n \in \mathbb{N}$, where I_n is the identity matrix of the dimension n.

Let $\mathrm{D}(\mathcal{H})$ be the set of all density operators, which are semi-positive and trace-one matrices on $\mathcal{H}$. Let the trace norm of a linear map X be $\|X\|_1 \coloneqq \operatorname{tr}\left(\sqrt{X^\dagger X}\right)$. For any quantum channels $\mathcal{E}$ and $\mathcal{F}$ that act on $\mathrm{D}(\mathcal{H})$, the *diamond norm distance* between them is defined as

$$\|\mathcal{E} - \mathcal{F}\|_\diamond \coloneqq \sup_{\rho \in \mathrm{D}(\mathcal{H} \otimes \mathcal{H}')} \|(\mathcal{E} \otimes \mathcal{I}_{\mathcal{H}'})(\rho) - (\mathcal{F} \otimes \mathcal{I}_{\mathcal{H}'})(\rho)\|_1.$$

2.1 Closeness measures for distributions and quantum states

Since both classical and quantum Tsallis entropy are central to this paper, we introduce the *q-logarithm* function $\ln_q\colon \mathbb{R}^+ \to \mathbb{R}$ for any real $q \neq 1$:

$$\forall x \in \mathbb{R}^+, \quad \ln_q(x) \coloneqq \frac{1 - x^{1-q}}{q - 1}.$$

The q-logarithm is a generalization of the natural logarithm, as it is straightforward to verify that $\lim_{q \to 1} \ln_q(x) = \ln(x)$ for any $x \in \mathbb{R}^+$. However, the q-logarithm exhibits different behavior when $q \neq 1$, for instance, $\ln_q(xy) = \ln_q(x) + \ln_q(y) + (1-q)\ln_q(x)\ln_q(y)$. For additional properties of the q-logarithm, see the references [Tsa01, Appendix] and [Yam02].

In the rest of this subsection, we provide useful closeness measures for probability distributions in Section 2.1.1 and for quantum states in Section 2.1.2. For convenience, we use a general convention $D_1 \leq D_2$ to denote an inequality between two distances or divergences, whether classical or quantum. In particular, this notation, as seen in *the titles of technical lemmas* (e.g., Lemma 2.5), indicates that D_1 is bounded above by a function f of D_2, i.e., $D_1 \leq f(D_2)$; or that D_2 is bounded below by a function g of D_1, i.e., $g(D_1) \leq D_2$.

2.1.1 Closeness measures for classical probability distributions

We begin by defining the total variation distance:

Definition 2.1 (Total variation distance)**.** *Let p_0 and p_1 be two probability distributions over $[N]$. The total variation distance between two p_0 and p_1 is defined by*

$$\mathrm{TV}(p_0, p_1) \coloneqq \frac{1}{2}\|p_0 - p_1\|_1 = \frac{1}{2}\sum_{x \in [N]} |p_0(x) - p_1(x)|.$$

Then we define the Tsallis entropy and provide useful properties of the Tsallis entropy.

Definition 2.2 (q-Tsallis entropy and Shannon entropy)**.** *Let p be a probability distribution over $[N]$. The q-Tsallis entropy of p is defined by*

$$\mathrm{H}_q(p) \coloneqq \frac{1-\sum_{x\in[N]} p(x)^q}{q-1} = -\sum_{x\in[N]} p(x)^q \ln_q(p(x)).$$

The Shannon entropy is the limiting case of the q-Tsallis entropy as $q \to 1$:

$$\mathrm{H}_1(p) \coloneqq \lim_{q\to 1} \mathrm{H}_q(p) = \mathrm{H}(p) = -\sum_{x\in[N]} p(x)\ln(p(x)).$$

For $N = 2$, we slightly abuse the notation by writing the q-Tsallis binary entropy and the (Shannon) binary entropy as $\mathrm{H}_q(p_0) = \mathrm{H}_q(1-p_0) = \mathrm{H}_q(p)$ and $\mathrm{H}(p_0) = \mathrm{H}(1-p_0) = \mathrm{H}(p)$, respectively.

It is noteworthy that the properties in Lemma 2.3 were also provided in [Tsa88] without proofs. In addition, by considering the eigenvalues of any quantum state, Lemma 2.3 straightforwardly extends to quantum q-Tsallis entropy (see Definition 2.8).

Lemma 2.3 (Basic properties of Tsallis entropy, partially adapted from [Dar70])**.** *Let p and p' be two probability distributions over $[N]$ with $N \geq 2$, and let ν be the uniform distribution over $[N]$. We have the following properties of the Tsallis entropy $\mathrm{H}_q(p)$ with $q > 0$:*

- **Concavity**: *For any $\lambda \in [0,1]$, $\mathrm{H}_q((1-\lambda)p+\lambda p') \geq (1-\lambda)\mathrm{H}_q(p)+\lambda \mathrm{H}_q(p')$. Equivalently, $F(q;x) \coloneqq \frac{x-x^q}{q-1}$ is concave in $x \in [0,1]$ for any fixed $q > 0$, and $\mathrm{H}_q(p) = \sum_{i\in[N]} F(q;p(i))$.*
- **Extremes**: $0 \leq \mathrm{H}_q(p) \leq \mathrm{H}_q(\nu) = \frac{1-n^{1-q}}{q-1}$. *Specifically, $\mathrm{H}_q(p) = \mathrm{H}_q(\nu)$ occurs when $p = \nu$, and $\mathrm{H}_q(p) = 0$ occurs when $p(i) = \begin{cases} 1, & i = k \\ 0, & i \neq 1 \end{cases}$ for any $k \in [N]$.*
- **Monotonicity**: *For any q and q' satisfying $0 < q \leq q'$, $\mathrm{H}_q(p) \geq \mathrm{H}_{q'}(p)$.*

Proof. For the first item, by inspecting the proof of [Dar70, Theorem 6], we know that $\frac{x-x^q}{1-2^{1-q}} \cdot \frac{1-2^{1-q}}{q-1} = F(q;x)$ is concave in $x \in [0,1]$ for any fixed $q \neq 1$. It is easy to verify that $\mathrm{H}_q(p) = \sum_{i\in[N]} F(q;p(i))$, we have that $\mathrm{H}_q(p)$ is concave.

For the second item, note that $\frac{q-1}{1-2^{1-q}} \geq 0$ for $q \neq 1$ and $\lim_{q\to 1} \frac{q-1}{1-2^{1-q}} = \frac{1}{\ln 2}$. Hence, by [Dar70, Theorem 6], we deduce $0 \leq \mathrm{H}_q(p) \leq \mathrm{H}_q(\nu)$. Moreover, because $F(q;x)$ is non-negative and $F(q;x) = 0$ occurs when $x = 1$, we conclude that $\mathrm{H}_q(p) = 0$ occurs when p satisfies the desired condition.

For the third item, since $\lim_{q\to 1} \mathrm{H}_q(x) = \mathrm{H}(x)$, it is enough to show that $\frac{\partial}{\partial q}\mathrm{H}_q(x) \leq 0$ for any $q \neq 1$ and $x \in [0,1]$. Given that $\mathrm{H}_q(p) = \sum_{i\in[n]} F(q;p(i))$, it remains to prove that $\frac{\partial}{\partial q}F(q;x) \leq 0$, specifically:

$$\frac{\partial}{\partial q}F(q;x) = -\frac{x-x^q}{(q-1)^2} - \frac{x^q\log(x)}{q-1} \leq 0 \Leftrightarrow G(q;x) \coloneqq x^q - x - (q-1)x^q\log(x) \leq 0. \tag{2.1}$$

A direct calculation implies that $\frac{\partial}{\partial q}G(q;x) = -(q-1)x^q\ln^2(x)$ for any $x \in [0,1]$. This inequality shows that for any fixed $x \in [0,1]$, $G(q;x)$ is monotonically increasing for $0 < q \leq 1$ and monotonically decreasing for $q > 1$. Hence, by noticing $\max_{q\geq 0} G(q;x) \leq G(1;x) = 0$ for any $x \in [0,1]$, we establish Equation (2.1) and the monotonicity. □

Next, we define a variant of the Jensen-Shannon divergence based on the q-Tsallis entropy:

Definition 2.4 (q-Jensen-(Shannon-)Tsallis divergence, adapted from [BR82])**.** *Let p_0 and p_1 be two probability distributions over $[N]$. The q-Jensen-(Shannon-)Tsallis divergence between p_0 and p_1 is defined as*

$$\mathrm{JT}_q(p_0,p_1) \coloneqq \begin{cases} \mathrm{H}_q\left(\frac{p_0+p_1}{2}\right) - \frac{1}{2}\left(\mathrm{H}_q(p_0)+\mathrm{H}_q(p_1)\right), & q \neq 1 \\ \mathrm{H}\left(\frac{p_0+p_1}{2}\right) - \frac{1}{2}\left(\mathrm{H}(p_0)+\mathrm{H}(p_1)\right), & q = 1 \end{cases}.$$

Specifically, the Jensen-Shannon divergence $\mathrm{JS}(p_0,p_1) = \mathrm{JT}_1(p_0,p_1)$.

Lastly, we provide a useful bound the divergence JT_q, which generalizes the bound $\mathrm{H}\left(\frac{1}{2}\right) - \mathrm{H}\left(\frac{1}{2} - \frac{\mathrm{TV}(p_0,p_1)}{2}\right) \leq \mathrm{JS}(p_0,p_1)$ in [Top00, Theorem 5] for the case of $q = 1$:

Lemma 2.5 ($\mathrm{TV} \leq \mathrm{JT}_q$, adapted from [BH09, Theorem 9])**.** *Let p_0 and p_1 be two probability distributions over $[N]$. For any $1 \leq q \leq 2$, we have the following inequality:*[23]

$$\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1}{2} - \frac{\mathrm{TV}(p_0,p_1)}{2}\right) \leq \mathrm{JT}_q(p_0,p_1).$$

It is important to note, the joint convexity of JT_q [BR82, Corollary 1] plays a key role in proving Lemma 2.5. And additionally, for $N \geq 3$, the joint convexity of JT_q holds if and only if $q \in [1,2]$, as stated in [BR82, Corollary 2].

2.1.2 Closeness measures for quantum states

We start by defining the trace distance and providing useful properties of this distance:

Definition 2.6 (Trace distance)**.** *The trace distance between two quantum states ρ_0 and ρ_1 is*

$$\mathrm{T}(\rho_0,\rho_1) \coloneqq \frac{1}{2}\operatorname{tr}(|\rho_0-\rho_1|) = \frac{1}{2}\operatorname{tr}\left(\left((\rho_0-\rho_1)^\dagger(\rho_0-\rho_1)\right)^{1/2}\right).$$

Importantly, the trace distance is a measured version of the total variation distance [NC10, Theorem 9.1]. In particular, for any classical f-divergence $\mathrm{D}_f(\cdot,\cdot)$, let ρ_0 and ρ_1 be two N-dimensional quantum states that are mixed in general, we can define the *measured quantum f-divergence* by considering the probability distributions induced by the POVM $\mathcal{M}$:

$$\mathrm{D}_f^{\mathrm{meas}}(\rho_0,\rho_1) = \sup_{\mathrm{POVM}\ \mathcal{M}} \mathrm{D}_f\left(p_0^{(\mathcal{M})}, p_1^{(\mathcal{M})}\right) \text{ where } p_b^{(\mathcal{M})} \coloneqq \left(\operatorname{tr}(\rho_b M_1), \cdots, \operatorname{tr}(\rho_b M_N)\right). \tag{2.2}$$

Moreover, the trace distance is a distance metric (e.g., [Wil13, Lemma 9.1.8]). In addition, as indicated in [Wil13, Equation (9.134)], for pure states $|\psi_0\rangle$ and $|\psi_1\rangle$, we have

$$\mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) = \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}.$$

Additionally, the trace distance characterizes the maximum success probability of discriminating quantum states in quantum hypothesis testing, as explained in [Wil13, Section 9.1.4]:

Lemma 2.7 (Helstrom-Holevo bound [Hel67, Hol73b])**.** *Suppose that a mixed quantum state ρ is given such that either $\rho = \rho_0$ or $\rho = \rho_1$ with equal probability. Then, any POVM distinguishes the two cases with success probability upper bounded by*

$$p_{\mathrm{succ}} \leq \frac{1}{2} + \frac{1}{2}\mathrm{T}(\rho_0,\rho_1).$$

Next, we define the quantum q-Tsallis entropy, generalizing the von Neumann entropy:

[23] It is evident that $\mathrm{H}_q\left(\frac{1-x}{2}\right) = \mathrm{H}_q\left(\frac{1+x}{2}\right)$ for any $x \in [0,1]$. Moreover, the proof of the lower bound in [BH09, Theorem 9] uses the notation $V(p_0,p_1) \coloneqq \sum_{i=1}^n |p_0(i) - p_1(i)| = \|p_0 - p_1\|_1 = 2\mathrm{TV}(p_0,p_1)$ defined in [Top00], where p_0 and p_1 are probability distributions over $[N]$.

Definition 2.8 (Quantum q-Tsallis entropy and von Neumann entropy)**.** *Let ρ be a (mixed) quantum state. The quantum q-Tsallis entropy of ρ is defined by*

$$\mathrm{S}_q(\rho) \coloneqq \frac{1-\mathrm{tr}(\rho^q)}{q-1} = -\mathrm{tr}(\rho^q \ln_q(\rho)).$$

Furthermore, as $q \to 1$, the quantum q-Tsallis entropy coincides with the von Neumann entropy:

$$\mathrm{S}_1(\rho) \coloneqq \lim_{q\to 1} \mathrm{S}_q(\rho) = \mathrm{S}(\rho) = -\mathrm{tr}(\rho \ln(\rho)).$$

Lemma 2.9 (Pseudo-additivity of S_q, adapted from [Rag95, Lemma 3])**.** *For any quantum states ρ_0 and ρ_1, and any $q \geq 1$, we have*:

$$\mathrm{S}_q(\rho_0 \otimes \rho_1) = \mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1) - (q-1)\mathrm{S}_q(\rho_0)\mathrm{S}_q(\rho_1).$$

Specifically, the equality $\mathrm{S}_q(\rho_0 \otimes \rho_1) = \mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)$ holds if and only if (a) *$q = 1$, or* (b) *for $q > 1$, either of the states ρ_0 or ρ_1 is pure.*

Now we define a variant of the quantum Jensen-Shannon divergence [MLP05] based on the quantum q-Tsallis entropy, as stated in Definition 2.10. Notably, the study of quantum analogs of the Jensen-Shannon divergence could date back to the Holevo bound [Hol73a].[24]

Definition 2.10 (Quantum q-Jensen-(Shannon-)Tsallis Divergence, adapted from [BH09])**.** *Let ρ_0 and ρ_1 be two quantum states that are mixed in general. The quantum q-Jensen-(Shannon-)Tsallis divergence between ρ_0 and ρ_1 is defined by*

$$\mathrm{QJT}_q(\rho_0, \rho_1) \coloneqq \begin{cases} \mathrm{S}_q\left(\frac{\rho_0+\rho_1}{2}\right) - \frac{1}{2}\left(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)\right), & q \neq 1 \\ \mathrm{S}\left(\frac{\rho_0+\rho_1}{2}\right) - \frac{1}{2}\left(\mathrm{S}(\rho_0) + \mathrm{S}(\rho_1)\right), & q = 1 \end{cases}.$$

Specifically, for pure states $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$, $\mathrm{QJT}_q(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) = \mathrm{S}_q\left(\frac{|\psi_0\rangle\langle\psi_0|+|\psi_1\rangle\langle\psi_1|}{2}\right)$. Furthermore, the quantum Jensen-Shannon divergence $\mathrm{QJS}(\rho_0, \rho_1) = \mathrm{QJT}_1(\rho_0, \rho_1)$.

It is worth noting that the square root of QJT_q is a distance metric when $0 \leq q \leq 2$ [Sra21] (also see [Vir21] for the $q = 1$ case). Moreover, whereas QJS can be expressed as a symmetrized version of the quantum relative entropy $\mathrm{D}(\rho_0\|\rho_1) \coloneqq \mathrm{tr}(\rho_0(\ln(\rho_0) - \ln(\rho_1)))$ by

$$\mathrm{QJS}(\rho_0, \rho_1) = \frac{1}{2}\left(\mathrm{D}\left(\rho_0 \middle\| \frac{\rho_0+\rho_1}{2}\right) + \mathrm{D}\left(\rho_1 \middle\| \frac{\rho_0+\rho_1}{2}\right)\right), \tag{2.3}$$

a similar equality does not hold for QJT_q with respect to the quantum Tsallis relative entropy $\mathrm{D}_q(\rho_0\|\rho_1) \coloneqq \frac{1-\mathrm{tr}(\rho_0^q \rho_1^{1-q})}{1-q}$ (see, e.g., [FYK04]).[25] In addition to QJS, the work of [FvdG99] studied a measured variant of the Jensen-Shannon divergence $\mathrm{QJS}^{\mathrm{meas}}(\rho_0, \rho_1)$ in terms of Equation (2.2), namely the *quantum Shannon distinguishability*.

Lastly, we provide more useful properties of QJT_q. By combining [FYK07, Theorem 1.5] and [Fur05, Remark V.3], we can immediately derive Lemma 2.11 and Lemma 2.12. In particular, the equality in Lemma 2.12 holds even in a stronger form: $\mathrm{S}_q\left(\sum_{i\in[k]} \mu_i^q \rho_i\right) = \mathrm{H}_q(\mu) + \sum_{i\in[k]} \mu_i \mathrm{S}_q(\rho_i)$ for orthogonal quantum states $\rho_1, \cdots, \rho_k$. Additionally, it is noteworthy that Lemma 2.12 admits a simple proof in [Kim16, Lemma 1].

Lemma 2.11 (Unitary invariance of QJT_q, adapted from [FYK07, Theorem 1.5])**.** *For any quantum states ρ_0 and ρ_1, and any unitary transformation U acting on ρ_0 or ρ_1, it holds that*:

$$\mathrm{QJT}_q(U^\dagger \rho_0 U, U^\dagger \rho_1 U) = \mathrm{QJT}_q(\rho_0, \rho_1).$$

[24] The quantum Jensen-Shannon divergence (QJS) is a special case of the Holevo χ quantity (the right-hand side of the Holevo bound [Hol73a]). Following the notations in [NC10, Theorem 12.1], QJS coincides with the Holevo χ quantity on size-2 ensembles with a uniform distribution.

[25] A symmetrized version of the quantum Tsallis relative entropy will lead to a different quantity, see [JMDA21].

Lemma 2.12 (Joint q-Tsallis entropy theorem, adapted from [FYK07, Theorem 1.5]). *Let k be an integer, and let $\{\rho_i\}_{i\in[k]}$ be a set of (mixed) quantum states. Let k-tuple $\mu := (\mu_1, \cdots, \mu_k)$ be a probability distribution. Then, for any $q \geq 0$, we have the following*:

$$\mathrm{S}_q\left(\sum_{i\in[k]} \mu_i |i\rangle\langle i| \otimes \rho_i\right) = \mathrm{H}_q(\mu) + \sum_{i\in[k]} \mu_i^q \mathrm{S}_q(\rho_i).$$

Following the discussion in [Ras11, Section 3], Fannes' inequality for QJT_q, where $0 \leq q \leq 2$, was established in [FYK07, Theorem 2.4]. Notably, for QJT_q with $q > 1$, a sharper Fannes-type inequality was provided in [Zha07, Theorem 2]:

Lemma 2.13 (Fannes' inequality for QJT_q, adapted from Theorem 2 and Corollary 2 in [Zha07]). *For any quantum states ρ_0 and ρ_1 of dimension N, we have:*

$$\forall q > 1,\ |\mathrm{S}_q(\rho_0) - \mathrm{S}_q(\rho_1)| \leq \mathrm{T}(\rho_0, \rho_1)^q \cdot \ln_q(N-1) + \mathrm{H}_q(\mathrm{T}(\rho_0, \rho_1)).$$

Moreover, for the case of $q = 1$ (von Neumann entropy), we have:

$$|\mathrm{S}(\rho_0) - \mathrm{S}(\rho_1)| \leq \mathrm{T}(\rho_0, \rho_1) \cdot \ln(N-1) + \mathrm{H}(\mathrm{T}(\rho_0, \rho_1)).$$

2.2 Closeness testing of quantum states via state-preparation circuits

We begin by defining the closeness testing of quantum states with respect to the trace distance, denoted as $\mathrm{QSD}[\alpha, \beta]$,[26] along with a variant of this promise problem, as described in Definition 2.14. In particularly, we say that $\mathcal{P} = (\mathcal{P}_{\text{yes}}, \mathcal{P}_{\text{no}})$ is a *promise* problem, if it satisfies the conditions $\mathcal{P}_{\text{yes}} \cap \mathcal{P}_{\text{no}} = \emptyset$ and $\mathcal{P}_{\text{yes}} \cup \mathcal{P}_{\text{no}} \subseteq \{0,1\}^*$.

Definition 2.14 (Quantum State Distinguishability, QSD, adapted from [Wat02, Section 3.3]). *Let Q_0 and Q_1 be quantum circuits acting on m qubits ("input length") and having n specified output qubits ("output length"), where $m(n)$ is a polynomial function of n. Let ρ_i denote the quantum state obtained by running Q_i on state $|0\rangle^{\otimes m}$ and tracing out the non-output qubits. Let $\alpha(n)$ and $\beta(n)$ be efficiently computable functions. Decide whether*:

- Yes: *A pair of quantum circuits (Q_0, Q_1) such that $\mathrm{T}(\rho_0, \rho_1) \geq \alpha(n)$;*
- No: *A pair of quantum circuits (Q_0, Q_1) such that $\mathrm{T}(\rho_0, \rho_1) \leq \beta(n)$.*

Furthermore, we denote the restricted version, where ρ_0 and ρ_1 are pure states, as PureQSD.

In addition to QSD, we can similarly define the closeness testing of a quantum state to the maximally mixed state (with respect to the trace distance), denoted as $\mathrm{QSCMM}[\beta, \alpha]$:

Definition 2.15 (Quantum State Closeness to Maximally Mixed State, QSCMM, adapt from [Kob03, Section 3]). *Let Q be a quantum circuit acting on m qubits and having n specified output qubits, where $m(n)$ is a polynomial function of n. Let ρ denote the quantum state obtained by running Q on state $|0\rangle^{\otimes m}$ and training out the non-output qubits. Let $\alpha(n)$ and $\beta(n)$ be efficiently computable functions. Decide whether*:

- Yes: *A quantum circuit Q such that $\mathrm{T}(\rho, (I/2)^{\otimes n}) \leq \beta(n)$;*
- No: *A quantum circuit Q such that $\mathrm{T}(\rho, (I/2)^{\otimes n}) \geq \alpha(n)$.*

[26] While Definition 2.14 aligns with the classical counterpart of QSD defined in [SV03, Section 2.2], it is slightly less general than the definition in [Wat02, Section 3.3]. Specifically, Definition 2.14 assumes that the input length m and the output length n are *polynomially equivalent*, whereas [Wat02, Section 3.3] allows for cases where the output length (e.g., a single qubit) is *much smaller* than the input length.

2.2.1 Input models and the concept of reductions

In this work, we consider the *purified quantum access input model*, as defined in [Wat02], in both white-box and black-box scenarios:

- **White-box input model**: The input of the problem QSD consists of descriptions of polynomial-size quantum circuits Q_0 and Q_1. Specifically, for $b \in \{0,1\}$, the description of Q_b includes a sequence of polynomially many 1- and 2-qubit gates.
- **Black-box input model**: In this model, instead of providing the descriptions of the quantum circuits Q_0 and Q_1, only query access to Q_b is allowed, denoted as O_b for $b \in \{0,1\}$. For convenience, we also allow query access to $Q_b^\dagger$ and controlled-Q_b, denoted by $O_b^\dagger$ and controlled-O_b, respectively.

Next, the concept of *reductions* between promise problems is used to address computational hardness in the context of the while-box input model, particularly in relation to complexity classes. Following the definitions in [Gol08, Section 2.2.1], we introduce two types of reductions from a promise problem $\mathcal{P} = (\mathcal{P}_{\text{yes}}, \mathcal{P}_{\text{no}})$ to another promise problem $\mathcal{P}' = (\mathcal{P}'_{\text{yes}}, \mathcal{P}'_{\text{no}})$:

- **Karp reduction**. A deterministic polynomial-time computable function f is called a *Karp reduction* from a promise problem $\mathcal{P}$ to another promise problem $\mathcal{P}'$ if, for every x, the following holds: $x \in \mathcal{P}_{\text{yes}}$ if and only if $f(x) \in \mathcal{P}'_{\text{yes}}$, and $x \in \mathcal{P}_{\text{no}}$ if and only if $f(x) \in \mathcal{P}'_{\text{no}}$.
- **Turing reduction**. A promise problem $\mathcal{P}$ is *Turing-reducible* to a promise problem $\mathcal{P}'$ if there exists a deterministic polynomial-time oracle machine $\mathcal{A}$ such that, for every function f that solves $\mathcal{P}'$ it holds that $\mathcal{A}^f$ solves $\mathcal{P}$. Here, $\mathcal{A}^f(x)$ denotes the output of machine $\mathcal{A}$ on input x when given oracle access to f.

It is noteworthy that Karp reduction is a special case of Turing reduction.

2.2.2 Computational hardness of QSD and QSCMM

Note that the polarization lemma for the total variation distance [SV03] and the trace distance [Wat02, Wat09] have the same inequalities. Consequently, using the parameters chosen in [BDRV19, Theorem 3.14], we obtain the QSZK hardness of QSD:

Lemma 2.16 (QSD is QSZK-hard)**.** *Let $\alpha(n)$ and $\beta(n)$ be efficiently computable functions satisfying $\alpha^2(n) - \beta(n) \geq 1/O(\log n)$. For any constant $\tau \in (0, 1/2)$, $\text{QSD}[\alpha, \beta]$ is QSZK-hard under Karp reduction when $\alpha(n) \leq 1 - 2^{-n^\tau}$ and $\beta(n) \geq 2^{-n^\tau}$ for every $n \in \mathbb{N}$.*

Following the construction in [RASW23, Theorem 12] (see also [LGLW23, Lemma 17] and [WZ24a, Theorem IV.1]), we can establish that PureQSD is BQP-hard under Karp reduction:

Lemma 2.17 (PureQSD is BQP-hard)**.** *Let $\alpha(n)$ and $\beta(n)$ be efficiently computable functions such that $\alpha(n) - \beta(n) \geq 1/\operatorname{poly}(n)$. For any polynomial $l(n)$, let $n' := n+1$, $\text{PureQSD}[\alpha(n'), \beta(n')]$ is BQP-hard when $\alpha(n') \leq \sqrt{1 - 2^{-2l(n'-1)}}$ and $\beta(n') \geq 2^{-(l(n'-1)+1)/2}$ for every integer $n' \geq 2$. Specifically, by choosing $l(n'-1) = n'$, it holds that: For every integer $n' \geq 2$,*

$$\text{PureQSD}\left[\sqrt{1 - 2^{-2n'}}, 2^{-(n'+1)/2}\right] \text{ is BQP-hard under Karp reduction.}$$

Proof. Since BQP is closed under complement, it suffices to show that PureQSD is coBQP-hard under Karp reduction. For any promise problem $(\mathcal{P}_{yes}, \mathcal{P}_{no}) \in \text{coBQP}[b(n), a(n)]$ with $a(n) - b(n) \geq 1/\operatorname{poly}(n)$, we assume without loss of generality that the coBQP circuit $\hat{C}_x$ has an output length of n. Leveraging error reduction for coBQP via a sequential repetition, for any polynomial $l(n)$, we can achieve that the acceptance probability $\Pr[C_x \text{ accepts}] \leq 2^{-l(n)}$ for *yes* instances, whereas $\Pr[C_x \text{ accepts}] \geq 1 - 2^{-l(n)}$ for *no* instances.

Next, we construct a new quantum circuit C'_x with an additional single-qubit register F initialized to zero. The circuit C'_x is defined as $C'_x := C_x^\dagger X_{\mathsf{O}}^\dagger \mathrm{CNOT}_{\mathsf{O}\to\mathsf{F}} X_{\mathsf{O}} C_x$, where the single-qubit register O corresponds to the output qubit. It is evident that the output length n' of C'_x satisfies $n' = n+1$. We say that C'_x accepts if the measurement outcomes of all qubits are all zero. Then, we have:

$$\Pr[C'_x \text{ accepts}] = \left\| (|\bar{0}\rangle\langle\bar{0}| \otimes |0\rangle\langle 0|_{\mathsf{F}}) C'_x (|\bar{0}\rangle \otimes |0\rangle_{\mathsf{F}}) \right\|_2^2 = \left| \langle \bar{0}| C_x^\dagger |1\rangle\langle 1|_{\mathsf{O}} C_x |\bar{0}\rangle \right|^2 = \Pr[C_x \text{ accepts}]^2. \quad (2.4)$$

Here, the second equality owes to $\mathrm{CNOT}_{\mathsf{O}\to\mathsf{F}} = |0\rangle\langle 0|_{\mathsf{O}} \otimes I_{\mathsf{F}} + |1\rangle\langle 1|_{\mathsf{O}} \otimes X_{\mathsf{F}}$. By defining two pure states $|\psi_0\rangle := |\bar{0}\rangle \otimes |0\rangle_{\mathsf{F}}$ and $|\psi_1\rangle := C'_x(|\bar{0}\rangle \otimes |0\rangle_{\mathsf{F}})$ corresponding to $Q_0 = I$ and $Q_1 = C'_x$, respectively, we can derive the following:

$$\Pr[C'_x \text{ accepts}] = |\langle \psi_0|\psi_1\rangle|^2 = 1 - \mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|). \quad (2.5)$$

Combining Equation (2.4) and Equation (2.5), we conclude that:

- For *yes* instances, $\Pr[C_x \text{ accepts}] = |\langle \psi_0|\psi_1\rangle| \le 2^{-l(n)}$ implies that
$$\mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \ge \sqrt{1 - 2^{-2l(n)}} \ge \sqrt{1 - 2^{-2l(n'-1)}}.$$
- For *no* instances, $\Pr[C_x \text{ accepts}] = |\langle \psi_0|\psi_1\rangle| \ge 1 - 2^{-l(n)}$ yields that
$$\mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \le \sqrt{1 - (1 - 2^{-l(n)})^2} \le 2^{-(l(n)+1)/2} \le 2^{-(l(n'-1)+1)/2}.$$ □

Lastly, combining the proof strategy outlined in [Kob03, Section 3] and the reduction from QEA to QSCMM in [BASTS10, Section 5.3], the NIQSZK hardness of QSCMM was established in [CCKV08, Section 8.1] with an appropriate parameter trade-off:

Lemma 2.18 (QSCMM is NIQSZK-hard, adapted from [CCKV08, Section 8.1])**.**

For any $n \ge 3$, $\mathrm{QSCMM}[1/n, 1-1/n]$ is NIQSZK-hard under Karp reduction.

2.2.3 Query and sample complexity lower bounds for states and distributions

We begin by stating a query complexity lower bound for QSD. Note that an n-qubit maximally mixed state $(I/2)^{\otimes n}$ is commutative with any n-qubit quantum states ρ. Consider the spectral decomposition $\rho = \sum_{i\in[2^n]} \mu_i |v_i\rangle\langle v_i|$, where $\{|v_i\rangle\}_{i\in[2^n]}$ is an orthonormal basis, we have $\mathrm{T}(\rho, (I/2)^{\otimes n}) = \mathrm{TV}(\mu, U_{2^n})$, where U_{2^n} is a uniform distribution over $[2^n]$. Leveraging a similar argument for ρ_{U}, as in Lemmas 2.19 and 2.21, where the eigenvalues of ρ_{U} form a uniform distribution on the support of ρ, we can obtain:

Lemma 2.19 (Query complexity lower bound for QSD, adapted from [CFMdW10, Theorem 2])**.** *For any $\epsilon \in (0, 1/2]$, there exists an n-qubit quantum state ρ of rank r and the corresponding n-qubit state ρ_{U} such that the quantum query complexity to decide whether $\mathrm{T}(\rho, \rho_{\mathsf{U}})$ is at least ϵ or exactly 0, in the purified quantum query access model, is $\Omega(r^{1/3})$.*

It is noteworthy that the quantum query model used in [CFMdW10] differs from the purified quantum query access model. Nevertheless, this lower bound also applies to our query model, as the discussion after Definition 3 in [GL20].

Next, we introduce a query complexity lower bound of distinguishing probability distributions provided in [Bel19], which will be used to prove the quantum query complexity lower bound for estimating the quantum Tsallis entropy.

Lemma 2.20 (Query complexity for distinguishing probability distributions, [Bel19, Theorem 4])**.** *Suppose that U_p and U_q are two unitary operators such that*

$$U_p|0\rangle = \sum_{j\in[N]} \sqrt{p(j)}|j\rangle|\varphi_j\rangle \text{ and } U_q|0\rangle = \sum_{j\in[N]} \sqrt{q(j)}|j\rangle|\psi_j\rangle.$$

Here, p and q are probability distributions on $[N]$, and $\{|\varphi_j\rangle\}$ and $\{|\psi_j\rangle\}$ are orthonormal bases. Then, any quantum query algorithm that distinguishes U_p and U_q requires query complexity

$$\Omega(1/d_{\mathrm{H}}(p,q)).$$

Here, the Hellinger distance is defined as

$$d_{\mathrm{H}}(p,q) = \sqrt{\frac{1}{2}\sum_{j\in[n]}\left(\sqrt{p_j}-\sqrt{q_j}\right)^2}.$$

It is noteworthy that Lemma 2.20 was ever used as a tool to prove the quantum query complexity lower bounds for the closeness testing of probability distributions [LWL24] and the estimations of trace distance and fidelity [Wan24].

Furthermore, we also need a sample complexity lower bound for QSD, which follows from [OW21, Theorem 4.2] and is specified in Lemma 2.21. Here, *sample complexity* denotes the number of copies of ρ required to accomplish a specific closeness testing task.

Lemma 2.21 (Sample complexity lower bound for QSD, adapted from [OW21, Corollary 4.3])**.** *For any $\epsilon \in (0, 1/2]$, there exists an n-qubit quantum state ρ of rank r and the corresponding n-qubit state $\rho_{\mathtt{U}}$ such that the quantum sample complexity to decide whether $\mathrm{T}(\rho, \rho_{\mathtt{U}})$ is at least ϵ or exactly 0 is $\Omega(r/\epsilon^2)$.*

2.3 Polynomial approximations

We provide several useful results and tools for polynomial approximations in this subsections.

2.3.1 Best uniform polynomial approximations

Let $f(x)$ be a continuous function defined on the interval $[-1, 1]$ that we aim to approximate using a polynomial of degree at most d. We define P_d^* as a *best uniform approximation* on $[-1, 1]$ to f of degree d if, for any degree-d polynomial approximation P_d of f, the following holds:

$$\max_{x\in[-1,1]} |f(x) - P_d^*(x)| \leq \max_{x\in[-1,1]} |f(x) - P_d(x)|.$$

Let $\mathbb{R}_d[x]$ be the set of all polynomials (with real coefficients) of degree at most d. Equivalently, the best uniform approximation P_d^* to f is the polynomial that solves the minimax problem

$$\min_{P_d\in\mathbb{R}_d[x]} \max_{x\in[-1,1]} |f(x) - P_d(x)|.$$

Especially, we need the best uniform polynomial approximation of positive constant powers:

Lemma 2.22 (Best uniform approximation of positive constant powers, adapted from [Tim63, Section 7.1.41])**.** *For any positive (constant) integer r and order $\alpha \in (-1, 1)$, let $P_d^* \in \mathbb{R}[x]$ be the best uniform polynomial approximation for $f(x) = x^{r-1}|x|^{1+\alpha}$ of degree $d = \left\lceil (\beta_{\alpha,r}/\epsilon)^{\frac{1}{r+\alpha}} \right\rceil$, where $\beta_{\alpha,r}$ is a constant depending on $r + \alpha$. Then, for sufficiently small ϵ, it holds that*

$$\max_{x\in[-1,1]} |P_d^*(x) - f(x)| \leq \epsilon.$$

2.3.2 Chebyshev expansion and truncations

We introduce Chebyshev polynomial and an averaged variant of the Chebyshev truncation. We recommend [Riv90, Chapter 3] for a comprehensive review of Chebyshev expansion.

Definition 2.23 (Chebyshev polynomials)**.** *The Chebyshev polynomials (of the first kind) $T_k(x)$ are defined via the following recurrence relation: $T_0(x) := 1$, $T_1(x) := x$, and $T_{k+1}(x) := 2xT_k(x) - T_{k-1}(x)$. For $x \in [-1, 1]$, an equivalent definition is $T_k(\cos\theta) = \cos(k\theta)$.*

To use Chebyshev polynomials (of the first kind) for Chebyshev expansion, we first need to define an inner product between two functions, f and g, as long as the following integral exists:

$$\langle f, g\rangle \coloneqq \frac{2}{\pi}\int_{-1}^{1}\frac{f(x)g(x)}{\sqrt{1-x^2}}\mathrm{d}x. \tag{2.6}$$

The Chebyshev polynomials form an orthonormal basis in the inner product space induced by $\langle\cdot,\cdot\rangle$ defined in Equation (2.6). As a result, any continuous and integrable function $f:[-1,1]\to\mathbb{R}$ whose Chebyshev coefficients satisfy $\lim_{k\to\infty}c_k=0$, where c_k is defined in Equation (2.7), has a Chebyshev expansion given by:

$$f(x)=\frac{1}{2}c_0T_0(x)+\sum_{k=1}^{\infty}c_kT_k(x),\ \text{where } c_k\coloneqq\langle T_k,f\rangle. \tag{2.7}$$

Instead of approximating functions directly by the Chebyshev truncation $\tilde{P}_d=c_0/2+\sum_{k=1}^{d}c_kT_k$, we use an average of Chebyshev truncations, known as the *de La Vallée Poussin partial sum*, we obtain the degree-d *averaged Chebyshev truncation* $\hat{P}_{d'}$, which is a polynomial of degree $d'=2d-1$:

$$\hat{P}_{d'}(x)\coloneqq\frac{1}{d}\sum_{l=d}^{d'}\tilde{P}_l(x)=\frac{\hat{c}_0}{2}+\sum_{k=1}^{d'}\hat{c}_kT_k(x),\ \text{where } \hat{c}_k=\begin{cases}c_k, & 0\le k\le d'\\ \frac{2d-k}{d}c_k, & k>d\end{cases}, \tag{2.8}$$

we can achieve the truncation error 4ϵ for any function that admits Chebyshev expansion.

Lemma 2.24 (Asymptotically best approximation by averaged Chebyshev truncation, adapted from Exercises 3.4.6 and 3.4.7 in [Riv90])**.** *For any function f that has a Chebyshev expansion, consider the degree-d averaged Chebyshev truncation $\hat{P}_{d'}$ defined in Equation (2.8). Let $\varepsilon_d(f)$ be the truncation error corresponds to the degree-d best uniform approximation on $[-1,1]$ to f. If there is a degree-d polynomial $P_d^*\in\mathbb{R}[x]$ such that $\max_{x\in[-1,1]}|f(x)-P_d^*(x)|\le\epsilon$, then*

$$\max_{x\in[-1,1]}\left|f(x)-\hat{P}_{d'}(x)\right|\le 4\varepsilon_d(f)\le 4\max_{x\in[-1,1]}|f(x)-P_d^*(x)|\le 4\epsilon.$$

2.4 Quantum algorithmic toolkit

In this subsection, we provide several quantum algorithmic tools: the quantum singular value transformation, three useful quantum algorithmic subroutines, and the quantum samplizer, which enables a quantum query-to-sample simulation.

2.4.1 Quantum singular value transformation

We begin by introducing the notion of block-encoding:

Definition 2.25 (Block-encoding)**.** *A linear operator A on an $(n+a)$-qubit Hilbert space is said to be an (α,a,ϵ)-block-encoding of an n-qubit linear operator B, if*

$$\|\alpha(\langle 0|^{\otimes a}\otimes I_n)A(|0\rangle^{\otimes a}\otimes I_n)-B\|\le\epsilon,$$

where I_n is the n-qubit identity operator and $\|\cdot\|$ is the operator norm.

Then, we state the quantum singular value transformation:

Lemma 2.26 (Quantum singular value transformation, [GSLW19, Theorem 31])**.** *Suppose that unitary operator U is an (α,a,ϵ)-block-encoding of Hermitian operator A, and $P\in\mathbb{R}[x]$ is a polynomial of degree d with $|P(x)|\le\frac{1}{2}$ for $x\in[-1,1]$. Then, we can implement a quantum circuit $\tilde{U}$ that is a $(1,a+2,4d\sqrt{\epsilon/\alpha}+\delta)$-block-encoding of $P(A/\alpha)$, by using $O(d)$ queries to U and $O((a+1)d)$ one- and two-qubit quantum gates. Moreover, the classical description of $\tilde{U}$ can be computed in deterministic time* $\mathrm{poly}(d,\log(1/\delta))$.

2.4.2 Quantum subroutines

The first subroutine is the quantum amplitude estimation:

Lemma 2.27 (Quantum amplitude estimation, [BHMT02, Theorem 12]). *Suppose that U is a unitary operator such that*

$$U|0\rangle|0\rangle = \sqrt{p}|0\rangle|\phi_0\rangle + \sqrt{1-p}|1\rangle|\phi_1\rangle,$$

where $|\phi_0\rangle$ and $|\phi_1\rangle$ are normalized pure quantum states and $p \in [0,1]$. Then, there is a quantum query algorithm using $O(M)$ queries to U that outputs $\tilde{p}$ such that

$$\Pr\left[|\tilde{p}-p| \leq \frac{2\pi\sqrt{p(1-p)}}{M} + \frac{\pi^2}{M^2}\right] \geq \frac{8}{\pi^2}.$$

Moreover, if U acts on n qubits, then the quantum query algorithm can be implemented by using $O(Mn)$ one- and two-qubit quantum gates.

The second subroutine prepares a purified density matrix, originally stated in [LC19]:

Lemma 2.28 (Block-encoding of density operators, [GSLW19, Lemma 25]). *Suppose that U is an $(n+a)$-qubit unitary operator that prepares a purification of an n-qubit mixed quantum state ρ. Then, we can implement a unitary operator W by using 1 query to each of U and $U^\dagger$ such that W is a $(1, n+a, 0)$-block-encoding of ρ.*

The third subroutine is a specific version of one-bit precision phase estimation [Kit95], often referred to as the Hadamard test [AJL09], as stated in [GP22]:

Lemma 2.29 (Hadamard test for block-encodings, adapted from [GP22, Lemma 9]). *Suppose that unitary operator U is a $(1, a, 0)$-block-encoding of an n-qubit operator A. Then, we can implement a quantum circuit that, on input an n-qubit mixed quantum state ρ, outputs 0 with probability $\frac{1}{2} + \frac{1}{2}\operatorname{Re}[\operatorname{tr}(A\rho)]$ (resp., $\frac{1}{2} + \frac{1}{2}\operatorname{Im}[\operatorname{tr}(A\rho)]$), by using 1 query to controlled-U and $O(1)$ one- and two-qubit quantum gates.*

Moreover, if an $(n+a)$-qubit unitary operator $\mathcal{O}$ prepares a purification of ρ, then, by combining Lemma 2.27, we can estimate $\operatorname{tr}(A\rho)$ to within additive error ϵ by using $O(1/\epsilon)$ queries to each of U and $\mathcal{O}$ and $O((n+a)/\epsilon)$ one- and two-qubit quantum gates.

2.4.3 Quantum samplizer

We introduce the notion of samplizer in [WZ24c], which helps us establish the sample complexity upper bound from the query complexity upper bound.

Definition 2.30 (Samplizer). *A samplizer $\mathsf{Samplize}_*\langle * \rangle$ is a mapping that converts quantum query algorithms (quantum circuit families with query access to quantum unitary oracles) to quantum sample algorithms (quantum channel families with sample access to quantum states) such that: For any $\delta > 0$, quantum query algorithm $\mathcal{A}^U$, and quantum state ρ, there exists a unitary operator U_ρ that is a $(2, a, 0)$-block-encoding of ρ for some $a > 0$, satisfying*

$$\left\|\mathsf{Samplize}_\delta\langle \mathcal{A}^U\rangle[\rho] - \mathcal{A}^{U_\rho}\right\|_\diamond \leq \delta,$$

where $\|\cdot\|_\diamond$ is the diamond norm and $\mathcal{E}[\rho](\cdot)$ is a quantum channel $\mathcal{E}$ with sample access to ρ.

Then, we include an efficient implementation of the samplizer in [WZ24c], which is based on quantum principal component analysis [LMR14, KLL+17] and generalizes [GP22, Corollary 21] and [WZ23, Theorem 1.1].

Lemma 2.31 (Optimal samplizer, [WZ24c, Theorem 4]). *There is a samplizer $\mathsf{Samplize}_*\langle * \rangle$ such that for $\delta > 0$ and quantum query algorithm $\mathcal{A}^U$ with query complexity Q, the implementation of $\mathsf{Samplize}_\delta\langle \mathcal{A}^U\rangle[\rho]$ uses $\widetilde{O}(Q^2/\delta)$ samples of ρ.*

3 Efficient quantum algorithms for estimating q-quantum Tsallis entropy

In this section, we propose efficient quantum algorithms for estimating the quantum Tsallis entropy $\mathrm{S}_q(\rho)$ when $q \geq 1+\Omega(1)$, using either queries to the state-preparation circuit or samples of the state ρ. The key ingredient underlying our algorithms is an *efficient* uniform approximation of positive constant power functions. Specifically, our polynomial approximation (Lemma 3.1) is "full-range", meaning it maintains a uniform error bound across the entire interval $[-1,1]$. This differs from the polynomial approximations commonly used in QSVT, which typically provide separate error bounds for the intervals $[-\delta,\delta]$ and $[-1,-\delta)\cup(\delta,1]$.

Utilizing our "full-range" polynomial approximation, we construct a query-efficient quantum algorithm for estimating $\mathrm{tr}(\rho^q)$, as established in Theorem 3.2. Consequently, our quantum query algorithm (Theorem 3.2) directly leads to BQP containments for the promise problems $\textsc{TsallisQEA}_q$ and $\textsc{TsallisQED}_q$, defined in Section 5. Furthermore, by employing the samplizer in [WZ24c], we develop a sample-efficient quantum algorithm for estimating $\mathrm{tr}(\rho^q)$, as presented in Theorem 3.3.

3.1 Efficient uniform approximations to positive constant power functions

We provide an efficiently computable uniform approximation of positive constant powers:

Lemma 3.1 (Efficient uniform polynomial approximation of positive constant powers)**.** *Let r be a positive (constant) integer and let α be a real number in $(-1,1)$. For any $\epsilon \in (0,1/2)$, there is a degree-d polynomial $P_d \in \mathbb{R}[x]$, where $d = \left\lceil \left(\beta'_{\alpha,r}/\epsilon\right)^{\frac{1}{r+\alpha}} \right\rceil$ and $\beta'_{\alpha,r}$ is a constant depending on $r+\alpha$, that can be deterministically computed in $\widetilde{O}(d)$ time. For sufficiently small ϵ, it holds that*:

$$\max_{x\in[-1,1]} \left| \frac{1}{2} x^{r-1}|x|^{1+\alpha} - P_d(x) \right| \leq \epsilon \quad \textit{and} \quad \max_{x\in[-1,1]} |P_d(x)| \leq 1.$$

Furthermore, P_d has the same parity as the integer $r-1$.

Proof. Let $f(x) \coloneqq \frac{1}{2}x^{r-1}|x|^{1+\alpha}$. For any $\tilde{\epsilon} \in (0,1/8)$, using Lemma 2.22, we obtain the degree-$\tilde{d}$ best polynomial approximation $P^*_{\tilde{d}}(x)$, where $\tilde{d} = \left\lceil \left(\beta_{\alpha,r}/\tilde{\epsilon}\right)^{\frac{1}{r+\alpha}} \right\rceil$ and $\beta_{\alpha,r}$ is a constant depending on $r+\alpha$, such that

$$\max_{x\in[-1,1]} \left| \frac{1}{2} x^{r-1}|x|^{1+\alpha} - P^*_{\tilde{d}}(x) \right| \leq \tilde{\epsilon} \quad \text{and} \quad \max_{x\in[-1,1]} \left| P^*_{\tilde{d}}(x) \right| \leq \frac{1}{2} + \tilde{\epsilon}. \tag{3.1}$$

Next, we consider the degree-$\tilde{d}$ averaged Chebyshev truncation (Equation (2.8)) of $f(x)$. In particular, let $d \coloneqq 2\tilde{d}-1 = \left\lceil \left(\beta'_{\alpha,r}/\epsilon\right)^{\frac{1}{r+\alpha}} \right\rceil$, where $\beta'_{\alpha,r}$ is another constant depending on $r+\alpha$ and ϵ will be specified later. We obtain the following degree-d polynomial:

$$P_d(x) = \frac{\hat{c}_0}{2} + \sum_{k=1}^{d} \hat{c}_k T_k(x), \text{ where } \hat{c}_k \coloneqq \begin{cases} c_k, & 0 \leq k \leq \tilde{d} \\ \frac{2\tilde{d}-k}{\tilde{d}} c_k, & k > \tilde{d} \end{cases} \text{ and } c_k \coloneqq \langle T_k, f \rangle. \tag{3.2}$$

Using the asymptotically best approximation by averaged Chebyshev truncation (Lemma 2.24) and Equation (3.1), we can derive that $P_d(x)$ satisfies the following:

$$\max_{x\in[-1,1]} \left| \frac{1}{2} x^{r-1}|x|^{1+\alpha} - P_d(x) \right| \leq 4\tilde{\epsilon} \coloneqq \epsilon \text{ and } \max_{x\in[-1,1]} |P_d(x)| \leq \frac{1}{2} + 4\tilde{\epsilon} = \frac{1}{2} + \epsilon < 1.$$

It remains to show that $P_d(x)$ can be computed in deterministic time $\widetilde{O}(d)$. A direct calcu-

lation implies that the Chebyshev coefficient $\{c_k\}_{0\le k\le d}$ in Equation (3.2) satisfy the following:

$$c_{2l+1} = c_{2l-1}\cdot\frac{r+\alpha-2l+1}{r+\alpha+2l+1}, \quad c_{2l} = c_{2l-2}\cdot\frac{r+\alpha-2l+2}{r+\alpha+2l},$$
$$c_0 = \frac{2}{\pi}\int_{-1}^{1}\frac{\frac{1}{2}x^{r-1}|x|^{1+\alpha}\cdot T_0(x)}{\sqrt{1-x^2}}\mathrm{d}x = -\frac{-1+(-1)^r}{2\sqrt{\pi}}\cdot\frac{\Gamma\left(\frac{1}{2}(r+\alpha+1)\right)}{\Gamma\left(\frac{1}{2}(r+\alpha+2)\right)},$$
$$c_1 = \frac{2}{\pi}\int_{-1}^{1}\frac{\frac{1}{2}x^{r-1}|x|^{1+\alpha}\cdot T_1(x)}{\sqrt{1-x^2}}\mathrm{d}x = \frac{1+(-1)^r}{2\sqrt{\pi}}\cdot\frac{\Gamma\left(\frac{1}{2}(r+\alpha+2)\right)}{\Gamma\left(\frac{1}{2}(r+\alpha+3)\right)}.$$

Here, the Gamma function $\Gamma(x) \coloneqq \int_0^\infty t^{x-1}e^{-x}\mathrm{d}t$ for any $x > 0$.

Consequently, we can recursively compute the averaged Chebyshev coefficient $\{\hat{c}_k\}_{0\le k\le d}$ in deterministic time $\widetilde{O}(d)$. We complete the proof by noting that the Chebyshev polynomials $\{T_k(x)\}_{0\le k\le d}$ also can be recursively computed in deterministic time $\widetilde{O}(d)$. □

3.2 Quantum q-Tsallis entropy approximation for q constantly larger than 1

3.2.1 Query-efficient quantum algorithm for estimating $\mathrm{tr}(\rho^q)$

We now present efficient quantum query algorithms for estimating the q-Tsallis entropy of a mixed quantum state. For readability, their framework is given in Algorithm 1.

Algorithm 1 A framework for estimating q-Tsallis entropy for $q \ge 1+\Omega(1)$ (query access).

Input: A quantum circuit Q that prepares a purification of an n-qubit mixed quantum state ρ, and a precision parameter $\epsilon \in (0,1)$.

Output: A single bit $b \in \{0,1\}$ such that $\Pr[b=0] \approx \frac{1}{2} + \frac{1}{8}\mathrm{tr}(\rho^q)$.

1: Implement a unitary operator U_ρ that is a block-encoding of ρ by Lemma 2.28, using $O(1)$ queries to Q.
2: Let $P(x)$ be a polynomial that approximates $\frac{1}{4}x^{q-1}$ in the range $[0,1]$, where $P(x)$ is determined according to ϵ, n, and q. More precisely, for constant $q > 1$, $P(x)$ is chosen by Lemma 3.1.
3: Implement a unitary operator $U_{P(\rho)}$ that is a block-encoding of $P(\rho)$ by quantum singular value transformation (Lemma 2.26), using $O(\deg(P))$ queries to U_ρ.
4: Perform the Hadamard test on ρ and $U_{P(\rho)}$ by Lemma 2.29, and return the measurement outcome.

Theorem 3.2 (Trace estimation of quantum state constant powers via queries)**.** *Suppose that Q is a unitary operator that prepares a purification of mixed quantum state ρ. For every $q \ge 1+\Omega(1)$, there is a quantum query algorithm that estimates $\mathrm{tr}(\rho^q)$ to within additive error ϵ by using $O(1/\epsilon^{1+\frac{1}{q-1}})$ queries to Q.*

Proof. Let Q be an $(n+a)$-qubit unitary operator that prepares a purification of the n-qubit mixed quantum state ρ. Then, by Lemma 2.28, we can implement a unitary operator U_ρ that is a $(1, n+a, 0)$-block-encoding of ρ, by using $O(1)$ queries to Q.

Let $\epsilon_p \in (0,1)$ be a parameter to be determined later. By Lemma 3.1 with $r \coloneqq \max\{\lfloor q-1\rfloor, 1\}$, $\alpha \coloneqq q-1-r$, and $\epsilon \coloneqq \epsilon_p$, there exists a polynomial $P \in \mathbb{R}[x]$ of degree $d = O(1/\epsilon_p^{\frac{1}{q-1}})$ such that

$$\max_{x\in[0,1]}\left|P(x) - \frac{1}{2}x^{q-1}\right| \le \epsilon_p, \quad \text{and} \max_{x\in[-1,1]}|P(x)| \le 1.$$

By Lemma 2.26 with $P \coloneqq \frac{1}{2}P$, $\alpha \coloneqq 1$, $a \coloneqq n+a$, $\epsilon \coloneqq 0$ and $d \coloneqq O(1/\epsilon_p^{\frac{1}{q-1}})$, we can implement a quantum circuit $U_{P(\rho)}$ that is a $(1, n+a+2, \delta)$-block-encoding of $\frac{1}{2}P(\rho)$, by using $O(1/\epsilon_p^{\frac{1}{q-1}})$

queries to U_ρ. Moreover, the classical description of $U_{P(\rho)}$ can be computed in deterministic time $\mathrm{poly}(1/\epsilon_p, \log(1/\delta))$.

Suppose that $U_{P(\rho)}$ is a $(1, n+a+2, 0)$-block-encoding of A, i.e., $\|A - \frac{1}{2}P(\rho)\| \leq \delta$. Then, by Lemma 2.29, we can obtain an estimate $\tilde{x}$ of $\mathrm{tr}(A\rho)$ to within additive error ϵ_H by using $O(1/\epsilon_H)$ queries to each of $U_{P(\rho)}$ and Q such that

$$\Pr\left[|\tilde{x} - \mathrm{tr}(A\rho)| \leq \epsilon_H\right] \geq \frac{2}{3}. \tag{3.3}$$

It can be seen that, in the overall quantum circuit to obtain $\tilde{x}$, the number of queries to Q is

$$O\left(\frac{1}{\epsilon_H}\right) \cdot O\left(\frac{1}{\epsilon_p^{\frac{1}{q-1}}}\right) = O\left(\frac{1}{\epsilon_H \epsilon_p^{\frac{1}{q-1}}}\right),$$

and the number of one- and two-qubit quantum gates is

$$O\left(\frac{n+a}{\epsilon_H \epsilon_p^{\frac{1}{q-1}}}\right).$$

Moreover, the classical description of the overall quantum circuit can be computed in deterministic time $\mathrm{poly}(1/\epsilon_p, 1/\epsilon_H, \log(1/\delta))$.

On the other hand, we have

$$\left|\mathrm{tr}(A\rho) - \mathrm{tr}\left(\frac{1}{2}P(\rho)\rho\right)\right| \leq \left\|A - \frac{1}{2}P(\rho)\right\| \leq \delta, \tag{3.4}$$

where we use the inequality $|\mathrm{tr}(AB)| \leq \|A\| \mathrm{tr}(|B|)$ (which is a special case of the matrix Hölder inequality, e.g., [Bau11, Theorem 2]). We also have

$$\left|\mathrm{tr}\left(\frac{1}{2}P(\rho)\rho\right) - \mathrm{tr}\left(\frac{1}{4}\rho^q\right)\right| \leq \frac{1}{2}\epsilon_p. \tag{3.5}$$

To see Equation (3.5), suppose that $\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$ is the spectrum decomposition of ρ with $\lambda_j \geq 0$ for all j and $\sum_j \lambda_j = 1$. Then,

$$\begin{aligned}
\left|\mathrm{tr}\left(\frac{1}{2}P(\rho)\rho\right) - \mathrm{tr}\left(\frac{1}{4}\rho^q\right)\right| &= \left|\sum_j \left(\frac{1}{2}P(\lambda_j)\lambda_j - \frac{1}{4}\lambda_j^q\right)\right| \\
&\leq \sum_j \frac{1}{2}\lambda_j \left|P(\lambda_j) - \frac{1}{2}\lambda_j^{q-1}\right| \\
&\leq \frac{1}{2}\sum_j \lambda_j \epsilon_p = \frac{1}{2}\epsilon_p.
\end{aligned}$$

Finally, by combining Equations (3.3) to (3.5), we obtain

$$\Pr\left[|4\tilde{x} - \mathrm{tr}(\rho^q)| \leq 2\epsilon_p + 4\epsilon_H + 4\delta\right] \geq \frac{2}{3}.$$

To make $4\tilde{x}$ an ϵ-estimate of $\mathrm{tr}(\rho^q)$ with high probability, it is sufficient to take $\epsilon_p = \epsilon_H = \delta = \epsilon/10$, thereby using

$$O\left(\frac{1}{\epsilon^{1+\frac{1}{q-1}}}\right)$$

queries to Q. □

3.2.2 Sample-efficient quantum algorithm for estimating $\mathrm{tr}(\rho^q)$

We also study the sample complexity for the trace estimation of quantum state powers, which is obtained by extending the quantum query algorithm in Theorem 3.2 via the samplizer in

Lemma 2.31. An illustrative framework is given in Algorithm 2.

Algorithm 2 A framework for estimating q-Tsallis entropy for $q > 1 + \Omega(1)$ (sample access).

Input: Independent and identical samples of an n-qubit mixed quantum state ρ, and parameters $q > 1$ and $\delta, \epsilon_p, \delta_p \in (0,1)$.

Output: A single bit $b \in \{0,1\}$ such that $\Pr[b=0] \approx \frac{1}{2} + \frac{1}{2^{q+3}} \operatorname{tr}(\rho^q)$.

1: **function** $\texttt{ApproxPower}(q, \epsilon_p, \delta_p)^U$
   Input: A unitary $(1, a, 0)$-block-encoding U of A, and parameters $q > 1, \epsilon_p, \delta_p \in (0,1)$.
   Output: A unitary operator $\widetilde{U}$.
2: Let $P(x)$ be a polynomial of degree $d = O(1/\epsilon_p^{\frac{1}{q-1}})$ such that $\max_{x\in[0,1]}|P(x) - \frac{1}{2}x^{q-1}| \le \epsilon_p$ and $\max_{x\in[-1,1]}|P(x)| \le 1$ (by Lemma 3.1).
3: Construct a unitary $(1, a+2, \delta_p)$-block-encoding $\widetilde{U}$ of $\frac{1}{2}P(A)$ (by Lemma 2.26).
4: **return** $\widetilde{U}$.
5: **end function**
6: Let b' be the outcome of the Hadamard test (by Lemma 2.29) performing on the quantum state ρ and $\mathsf{Samplize}_\delta\langle \texttt{ApproxPower}(q, \epsilon_p, \delta_p)^U \rangle[\rho]$ (as if it were unitary).
7: **return** b'.

Theorem 3.3 (Trace estimation of quantum state constant powers via samples)**.** *For every $q \ge 1 + \Omega(1)$, there is a quantum sample algorithm that estimates $\operatorname{tr}(\rho^q)$ to within additive error ϵ by using $\widetilde{O}(1/\epsilon^{3+\frac{2}{q-1}})$ samples of ρ.*

Proof. Let unitary operator U be a $(1, a, 0)$-block-encoding of A for some $a > 0$ and let $\epsilon_p, \delta_p \in (0,1)$ be parameters to be determined. By Lemma 3.1 with $r := \max\{\lfloor q-1 \rfloor, 1\}$, $\alpha := q - 1 - r$, and $\epsilon := \epsilon_p$, there is a polynomial $P \in \mathbb{R}[x]$ of degree $d = O(1/\epsilon_p^{\frac{1}{q-1}})$ such that

$$\max_{x\in[0,1]} \left| P(x) - \frac{1}{2}x^{q-1} \right| \le \epsilon_p, \quad \text{and} \quad \max_{x\in[-1,1]} |P(x)| \le 1.$$

By Lemma 2.26 with $P := \frac{1}{2}P$, $\alpha := 1$, $a := n + a$, $\epsilon := 0$, $\delta := \delta_p$ and $d := O(1/\epsilon_p^{\frac{1}{q-1}})$, we can implement a quantum circuit $U_{P(A)}$ that is a $(1, n+a+2, \delta_p)$-block-encoding of $\frac{1}{2}P(A)$, by using $O(1/\epsilon_p^{\frac{1}{q-1}})$ queries to U. Moreover, the classical description of $U_{P(A)}$ can be computed in deterministic time $\operatorname{poly}(1/\epsilon_p, \log(1/\delta_p))$. Let $\texttt{ApproxPower}(q, \epsilon_p, \delta_p)^U$ denote the procedure of implementing $U_{P(A)}$ by using queries to U.

For our purpose, we take $A := \rho/2$. Suppose that $U_{P(\frac{\rho}{2})}$ is a $(1, n+a+2, 0)$-block-encoding of B, then $\|B - \frac{1}{2}P(\frac{\rho}{2})\| \le \delta_p$. Let $b \in \{0,1\}$ be the outcome of the Hadamard test (by Lemma 2.29) on ρ and $U_{P(\frac{\rho}{2})}$, then

$$\Pr[b=0] = \frac{1}{2} + \frac{1}{2}\operatorname{Re}[\operatorname{tr}(B\rho)]. \tag{3.6}$$

Let $\delta \in (0,1)$ be a parameter to be determined, and let $b' \in \{0,1\}$ be the outcome of the Hadamard test (by Lemma 2.29) on ρ and $\mathsf{Samplize}_\delta\langle \texttt{ApproxPower}(q, \epsilon_p, \delta_p)^U \rangle[\rho]$ (as if it were $U_{P(\frac{\rho}{2})}$). Then,

$$\left|\Pr[b=0] - \Pr[b'=0]\right| \le \delta. \tag{3.7}$$

Now we repeat the Hadamard test k times, obtaining outcomes $b'_1, b'_2, \ldots, b'_k \in \{0,1\}$, where k is an integer to be determined. Let $X = \frac{1}{k}\sum_{j=1}^k b'_j$. Then, by Hoeffding's inequality ([Hoe63,

Theorem 2]), we have

$$\Pr\big[\big|X - \mathbb{E}[b']\big| \le \epsilon_H\big] \ge 1 - 2\exp\big(-2k\epsilon_H^2\big). \tag{3.8}$$

On the other hand, similar to the proof of Theorem 3.2, we have

$$\left|\mathrm{Re}[\mathrm{tr}(B\rho)] - \mathrm{tr}\left(\frac{1}{2}P\left(\frac{\rho}{2}\right)\rho\right)\right| \le \left|\mathrm{tr}(B\rho) - \mathrm{tr}\left(\frac{1}{2}P\left(\frac{\rho}{2}\right)\rho\right)\right| \le \left\|B - \frac{1}{2}P\left(\frac{\rho}{2}\right)\right\| \le \delta_p. \tag{3.9}$$

We also have

$$\left|\mathrm{tr}\left(\frac{1}{2}P\left(\frac{\rho}{2}\right)\rho\right) - \mathrm{tr}\left(\frac{1}{2^{q+2}}\rho^q\right)\right| \le \frac{1}{2}\epsilon_p. \tag{3.10}$$

To see Equation (3.5), suppose that $\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$ is the spectrum decomposition of ρ with $\lambda_j \ge 0$ for all j and $\sum_j \lambda_j = 1$. Then,

$$\begin{aligned}
\left|\mathrm{tr}\left(\frac{1}{2}P\left(\frac{\rho}{2}\right)\rho\right) - \mathrm{tr}\left(\frac{1}{2^{q+2}}\rho^q\right)\right| &= \left|\sum_j \left(\frac{1}{2}P\left(\frac{\lambda_j}{2}\right)\lambda_j - \frac{1}{2^{q+2}}\lambda_j^q\right)\right| \\
&\le \sum_j \frac{1}{2}\lambda_j \left|P\left(\frac{\lambda_j}{2}\right) - \frac{1}{2}\left(\frac{\lambda_j}{2}\right)^{q-1}\right| \\
&\le \frac{1}{2}\sum_j \lambda_j \epsilon_p = \frac{1}{2}\epsilon_p.
\end{aligned}$$

Finally, by combining Equations (3.6) to (3.10), we obtain

$$\Pr\big[\big|2^{q+2}(1-2X) - \mathrm{tr}(\rho^q)\big| \le 2^{q+1}(4\delta + 4\epsilon_H + 2\delta_p + \epsilon_p)\big] \ge 1 - 2\exp(-2k\epsilon_H^2).$$

By taking $\delta = \epsilon_H = \delta_p = \epsilon_p := 2^{-q-5}\epsilon$ and $k := \left\lceil \frac{\ln(6)}{2\epsilon_H^2} \right\rceil$, we have

$$\Pr\big[\big|2^{q+2}(1-2X) - \mathrm{tr}(\rho^q)\big| \le \epsilon\big] \ge \frac{2}{3},$$

which means that $2^{q+2}(1-2X)$ is an ϵ-estimate of $\mathrm{tr}(\rho^q)$ with high probability.

To complete the proof, we analyze the sample complexity of our algorithm. The algorithm consists of k repetitions of the Hadamard test, and each Hadamard test uses one sample of ρ and one call to $\mathsf{Samplize}_\delta\langle \mathtt{ApproxPower}(q, \epsilon_p, \delta_p)^U\rangle[\rho]$. Here, $\mathtt{ApproxPower}(q, \epsilon_p, \delta_p)^U$ uses $O(1/\epsilon_p^{\frac{1}{q-1}})$ queries to U, and thus by Lemma 2.31 we can implement $\mathsf{Samplize}_\delta\langle \mathtt{ApproxPower}(q, \epsilon_p, \delta_p)^U\rangle[\rho]$ by using $\widetilde{O}(1/(\delta\epsilon_p^{\frac{2}{q-1}}))$ samples of ρ. Therefore, the total number of samples of ρ is

$$k \cdot \widetilde{O}\left(\frac{1}{\delta\epsilon_p^{\frac{2}{q-1}}}\right) = \widetilde{O}\left(\frac{1}{\epsilon^{3+\frac{2}{q-1}}}\right). \qquad \square$$

4 Properties of quantum Jensen-Tsallis divergence and Tsallis entropy

In this section, we present inequalities between the quantum q-Jensen-Tsallis divergence ($1 \le q \le 2$) and the trace distance. Our results (Theorem 4.1) extend the previous results for the quantum Jensen-Shannon divergence ($q = 1$), as stated in [BH09, Theorem 14].

Theorem 4.1 (QJT_q vs. T)**.** *For any quantum states ρ_0 and ρ_1, and $1 \le q \le 2$, we have*:

$$\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1 - \mathrm{T}(\rho_0, \rho_1)}{2}\right) \le \mathrm{QJT}_q(\rho_0, \rho_1) \le \mathrm{H}_q\left(\frac{1}{2}\right) \cdot \mathrm{T}(\rho_0, \rho_1)^q.$$

To prove Theorem 4.1, we first need to prove the data-processing inequality for QJT_q (Lemma 4.5), which crucially relies on the relatively recent results on the *joint convexity* of

QJT_q [CT14, Vir19]. Consequently, we can establish Theorem 4.1 by proving the inequalities in Section 4.2. In particular, the lower bound on QJT_q in terms of T (Lemma 4.6) holds for $q \in [1, 2]$, and the upper bound on QJT_q in terms of T (Lemma 4.7) for the same range of q.

Next, to utilize Lemma 4.6, we provide bounds of the Tsallis binary entropy in Section 4.3:

Theorem 4.2 (Tsallis binary entropy bounds)**.** *For any $p = (x, 1-x)$, let $\mathrm{H}_q(x)$ denote the Tsallis binary entropy with $1 \le q \le 2$, we have*:

$$\mathrm{H}_q(1/2) \cdot 4x(1-x) \le \mathrm{H}_q(x) \le \mathrm{H}_q(1/2) \cdot (4x(1-x))^{1/2}.$$

It is noteworthy that the best known bounds for the Shannon binary entropy ($q = 1$) are $\mathrm{H}(1/2) \cdot 4x(1-x) \le \mathrm{H}(q) \le \mathrm{H}(1/2) \cdot (4x(1-x))^{\frac{1}{2\mathrm{H}(1/2)}}$, as shown in [Top01, Theorem 1.2]. Our lower bound on the Tsallis binary entropy (Lemma 4.8) matches the case of $q = 1$, whereas our upper bound (Lemma 4.9) only aligns with a weaker bound $\mathrm{H}(q) \le \mathrm{H}(1/2) \cdot (4x(1-x))^{1/2}$ in [Lin91, Theorem 8] and the proof of Lemma 4.9 is more complicated than in the case of $q = 1$.

Lastly, we provide the inequalities between the Tsallis entropy of a distribution p and the total variation distance between p and the uniform distribution ν of the same dimension, as stated in Lemma 4.10. By adding an additional assumption regarding q and $\mathrm{TV}(p, \nu)$, this lemma partially generalizes the previous result for the case of $q = 1$ (cf. [CCKV08, Fact 8.4] and [KLGN19, Lemma 16]) to the case of $q > 1$.

4.1 Data-processing inequality for QJT_q from the joint convexity

With the correspondence between QJS and the quantum relative entropy (Equation (2.3)), the joint convexity of QJS directly follows from the joint convexity of the quantum relative entropy [Lie73, Uhl77] (see also [Rus22] for a simple proof). However, since QJT_q does not correspond to a Tsallis variant of quantum relative entropy (e.g., quasi-entropy [Pet07, Equation (3.23)]) in this sense, the joint convexity of QJT_q can only be established by the recent results of [CT14, Vir19]:

Lemma 4.3 (Joint convexity of QJT_q, adapted from [CT14, Vir19])**.** *Let k be an integer. For any $i \in [k]$, let $\rho_0^{(i)}$ and $\rho_1^{(i)}$ be two quantum states. Let k-tuple $\mu := (\mu_1, \cdots, \mu_k)$ be a probability distribution. Then, for any $q \in [1, 2]$ and $t \in (0, 1)$, the joint convexity of QJT_q holds*:

$$\mathrm{QJT}_q\left(\sum_{i\in[k]} \mu_i \rho_0^{(i)}, \sum_{i\in[k]} \mu_i \rho_1^{(i)}\right) \le \sum_{i\in[k]} \mu_i \mathrm{QJT}_q\left(\rho_0^{(i)}, \rho_1^{(i)}\right).$$

Proof. Following [CT14, Theorem 2.3(2)], we know that the quantum q-Tsallis entropy $\mathrm{S}_q(\rho)$ for $1 \le q \le 2$ is in the Matrix Entropy Class [CT14, Definition 2.2] (or [Vir19, Definition 2]). Therefore, as a corollary of [Vir19, Theorem 1], we can obtain: for any $1 \le q \le 2$ and $0 < \lambda < 1$,

$$\mathrm{QJT}_q\big((1-\lambda)\rho_0 + \lambda\rho_0', (1-\lambda)\rho_1 + \lambda\rho_1'\big) \le (1-\lambda)\mathrm{QJT}_q(\rho_0, \rho_1) + \lambda \mathrm{QJT}_q(\rho_0', \rho_1'). \tag{4.1}$$

Hence, we can complete the proof by applying Equation (4.1) inductively. □

Remark 4.4 (Data-processing inequality for $\mathrm{QJT}_{q,t}$). It is noteworthy that Lemma 4.3 applies to a generalized version of QJT_q, denoted as $\mathrm{QJT}_{q,t}$, such that $\mathrm{QJT}_{q,1/2} = \mathrm{QJT}_q$:

$$\forall t \in (0,1),\ \mathrm{QJT}_{q,t} := \mathrm{S}_q\left((1-t)\rho_0 + t\rho_1\right) - (1-t)\mathrm{S}_q(\rho_0) - t\mathrm{S}_q(\rho_1).$$

Lemma 2.11 also directly extends to $\mathrm{QJT}_{q,t}$, and consequently, Lemma 4.5 holds for $\mathrm{QJT}_{q,t}$ with $1 \le q \le 2$. However, the inequalities between QJT_q and the trace distance provided in this work, particularly Lemma 4.6 and Lemma 4.7, do not extend to $\mathrm{QJT}_{q,t}$ for $1 \le q \le 2$.

Lemma 4.5 (Data-processing inequality for QJT_q)**.** *For any quantum state ρ_0 and ρ_1, any quantum channel Φ, and $1 \le q \le 2$, we have*

$$\mathrm{QJT}_q(\Phi(\rho_0), \Phi(\rho_1)) \le \mathrm{QJT}_q(\rho_0, \rho_1).$$

Interestingly, the inequality in Lemma 4.5 *cannot* hold for $0 \le q < 1$. We can see this by considering pure states $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$, and their average $\hat{\rho}_{\psi,\phi} \coloneqq \frac{1}{2}\big(|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi|\big)$, then $\mathrm{QJT}_q(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = \mathrm{S}_q\big(\hat{\rho}_{\psi,\phi}\big)$. Following Lemma 4.5, we have $\mathrm{S}_q\big(\Phi\big(\hat{\rho}_{\psi,\phi}\big)\big) \le \mathrm{S}_q\big(\hat{\rho}_{\psi,\phi}\big)$ for $q \in [1, 2]$. However, using [FYK04, Corollary 2.6], we have $\mathrm{S}_q\big(\Phi\big(\hat{\rho}_{\psi,\phi}\big)\big) \ge \mathrm{S}_q\big(\hat{\rho}_{\psi,\phi}\big)$ for $q \in [0, 1)$.

Proof of Lemma 4.5. The case of $q = 1$ coincides with the quantum Jensen-Shannon divergence: Using Equation (2.3), $\mathrm{QJS}(\Phi(\rho_0), \Phi(\rho_1)) \le \mathrm{QJS}(\rho_0, \rho_1)$ follows from the data-processing inequality of the quantum relative entropy [Lin75, Uhl77] (see also [Pet07, Theorem 3.9]).

It remains to prove the case for $1 < q \le 2$. We use the standard proof strategy to derive the data-processing inequality from joint convexity, as in [FYK04, Theorem 2.5].

First, we consider the case of the partial trace tr_{B} on the quantum registers A and B, where $\rho_0, \rho_1 \in \mathrm{L}\big(\mathcal{H}_{\mathsf{AB}}\big)$ and $\dim(\mathcal{H}_{\mathsf{B}}) = N_{\mathsf{B}}$. Since $\mathrm{QJT}_q(\rho_0 \otimes \tilde{I}_{\mathsf{B}}, \rho_1 \otimes \tilde{I}_{\mathsf{B}}) = \mathrm{QJT}_q(\rho_0, \rho_1)$ where $\tilde{I}_{\mathsf{B}}$ is the maximally mixed state in B, it suffices to consider a quantum channel on registers A and B that is completely depolarizing on B and identity on A, denoted as $\Phi_{\mathrm{tr}_{\mathsf{B}}}$. Noting that $\Phi_{\mathrm{tr}_{\mathsf{B}}}$ can be expressed as a convex combination of unitary channels (e.g., [Wil13, Exercise 4.4.9] or [Rus22, Equation (9)]), for any quantum state ρ_{AB} on registers A and B, we can obtain:

$$\Phi_{\mathrm{tr}_{\mathsf{B}}}(\rho_{\mathsf{AB}}) \coloneqq \mathrm{tr}_{\mathsf{B}}(\rho_{\mathsf{AB}}) \otimes \mathrm{tr}(\rho_{\mathsf{AB}})\tilde{I}_{\mathsf{B}} = \sum_{l \in [N_{\mathsf{B}}^2]} \frac{1}{N_{\mathsf{B}}^2} (I_{\mathsf{A}} \otimes U_l)\rho_{\mathsf{AB}}(I_{\mathsf{A}} \otimes U_l)^\dagger,$$

where U_l is a unitary operator on B for each $l \in [N_{\mathsf{B}}^2]$.

Leveraging the joint convexity (Lemma 4.3) and the unitary invariance (Lemma 2.11) of QJS_q, we derive the following data-processing inequality concerning the quantum channel $\Phi_{\mathrm{tr}_{\mathsf{B}}}$:

$$\begin{aligned}
\mathrm{QJT}_q\left(\mathrm{tr}_{\mathsf{B}}(\rho_0), \mathrm{tr}_{\mathsf{B}}(\rho_1)\right) &= \mathrm{QJT}_q(\Phi_{\mathrm{tr}_{\mathsf{B}}}(\rho_0), \Phi_{\mathrm{tr}_{\mathsf{B}}}(\rho_1)) \\
&\le \sum_{l \in [N_{\mathsf{B}}^2]} \frac{1}{N_{\mathsf{B}}^2} \mathrm{QJT}_q\left((I_{\mathsf{A}} \otimes U_l)\rho_0(I_{\mathsf{A}} \otimes U_l)^\dagger, (I_{\mathsf{A}} \otimes U_l)\rho_1(I_{\mathsf{A}} \otimes U_l)^\dagger\right) \\
&= \sum_{l \in [N_{\mathsf{B}}^2]} \frac{1}{N_{\mathsf{B}}^2} \mathrm{QJT}_q\left(\rho_0, \rho_1\right) \\
&= \mathrm{QJT}_q(\rho_0, \rho_1).
\end{aligned} \tag{4.2}$$

Next, we move to the general case. Using the Stinespring dilation theorem (e.g., [AS17, Theorem 2.25]), for any quantum channel Φ on the registers (A, B), we have the following representation with some unitary U_Φ on the registers $(\mathsf{A}, \mathsf{B}, \mathsf{E})$ where $\dim(\mathcal{H}_{\mathsf{E}}) \le \dim(\mathcal{H}_{\mathsf{AB}})^2$:

$$\Phi(\rho_{\mathsf{AB}}) = \mathrm{tr}_{\mathsf{E}}\left(U_\Phi(\rho_{\mathsf{AB}} \otimes |\bar{0}\rangle\langle\bar{0}|_{\mathsf{E}})U_\Phi^\dagger\right).$$

Consequently, we can obtain the following for any quantum channel Φ:

$$\begin{aligned}
\mathrm{QJT}_q(\Phi(\rho_0), \Phi(\rho_1)) &\le \mathrm{QJT}_q\left(U_\Phi(\rho_0 \otimes |\bar{0}\rangle\langle\bar{0}|_{\mathsf{E}})U_\Phi^\dagger, U_\Phi(\rho_1 \otimes |\bar{0}\rangle\langle\bar{0}|_{\mathsf{E}})U_\Phi^\dagger\right) \\
&= \mathrm{QJT}_q\left(\rho_0 \otimes |\bar{0}\rangle\langle\bar{0}|_{\mathsf{E}}, \rho_1 \otimes |\bar{0}\rangle\langle\bar{0}|_{\mathsf{E}}\right) \\
&= \mathrm{QJT}_q(\rho_0, \rho_1).
\end{aligned}$$

Here, the first line owes to Equation (4.2), the second line is due to the unitary invariance of QJT_q (Lemma 2.11), and the last line is because $\mathrm{tr}\left((\rho_b \otimes |\phi\rangle\langle\phi|_{\mathsf{E}})^q\right) = \mathrm{tr}(\rho_b^q)$ for any $b \in \{0, 1\}$ and $q \in [1, 2]$. We now complete the proof. □

4.2 Inequalities between the trace distance and QJT_q

We begin by establishing the lower bound on QJT_q in terms of the trace distance, as stated in Lemma 4.6. The measured variant of the q-Jensen-Tsallis divergence (JT_q), denoted by $\mathrm{QJT}_q^{\mathrm{meas}}$, is derived from the definition provided in Equation (2.2).

Lemma 4.6 ($\mathrm{T} \leq \mathrm{QJT}_q$)**.** *For any quantum states ρ_0 and ρ_1, we have the following inequality*:

$$\forall q \in [1,2],\ \mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1}{2} - \frac{\mathrm{T}(\rho_0,\rho_1)}{2}\right) \leq \mathrm{QJT}_q^{\mathrm{meas}}(\rho_0,\rho_1) \leq \mathrm{QJT}_q(\rho_0,\rho_1).$$

Proof. The case of $q = 1$ follows from [BH09, Theorem 14], and can also be derived by combining [FvdG99, Theorem 1] with the Holevo bound (see [Liu23, Lemma 2.4] for details).

Our focus will be on the cases where $1 < q \leq 2$. We first prove the second inequality. Let $\mathcal{M}^*$ be an optimal POVM corresponding to $\mathrm{QJT}_q^{\mathrm{meas}}(\rho_0,\rho_1)$, then this POVM $\mathcal{M}^*$ corresponds to a quantum-to-classical channel $\Phi_{\mathcal{M}^*}(\rho) = \sum_{i=1}^{N} |i\rangle\langle i| \operatorname{tr}(\rho M_i^*)$ (e.g., [AS17, Equatin (2.41)]). Using the data-processing inequality for QJT_q (Lemma 4.5), for $1 < q \leq 2$, we obtain:

$$\mathrm{QJT}_q^{\mathrm{meas}}(\rho_0,\rho_1) = \mathrm{QJT}_q\left(\Phi_{\mathcal{M}^*}(\rho_0), \Phi_{\mathcal{M}^*}(\rho_1)\right) \leq \mathrm{QJT}_q(\rho_0,\rho_1).$$

Next, let us move to the first inequality. Let $p_b^{\mathcal{M}}$ be the induced distribution with respect to the POVM $\mathcal{M}$ of ρ_b for any $b \in \{0,1\}$. Utilizing Lemma 2.5, for $1 < q \leq 2$, we can derive that:

$$\mathrm{QJS}_{q,\mathcal{M}^*}^{\mathrm{meas}}(\rho_0,\rho_1) \geq \mathrm{QJS}_{q,\mathcal{M}}^{\mathrm{meas}}(\rho_0,\rho_1) = \mathrm{JT}_q\big(p_0^{\mathcal{M}}, p_1^{\mathcal{M}}\big) \geq \mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1}{2} - \frac{\mathrm{TV}\left(p_0^{\mathcal{M}}, p_1^{\mathcal{M}}\right)}{2}\right). \quad (4.3)$$

We then consider the function $g(q;x)$ and its first derivative $\frac{\partial}{\partial x} g(q;x)$:

$$g(q;x) \coloneqq \mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1-x}{2}\right) = \frac{2^{-q}}{q-1}\left((1+x)^q + (1-x)^q - 2\right),$$
$$\frac{\partial}{\partial x} g(q;x) = \frac{2^{-q} q}{q-1}\left((1+x)^{q-1} - (1-x)^{q-1}\right).$$

Since it is easy to see that $\frac{\partial}{\partial x} g(q;x) \geq 0$ for $0 \leq x \leq 1$ when $1 < q \leq 2$, we know that $g(q;x)$ is monotonically increasing for $0 \leq x \leq 1$. Noting that Equation (4.3) holds for arbitrary POVM $\mathcal{M}$, and the trace distance is the measured version of the total variation distance (e.g., [NC10, Theorem 9.1]), we thus complete the proof by choosing the POVM that maximizes $\mathrm{T}(\rho_0,\rho_1)$. □

Next, we demonstrate the upper bound on QJT_q in terms of the trace distance:

Lemma 4.7 ($\mathrm{QJT}_q \leq \mathrm{T}$)**.** *For any quantum states ρ_0 and ρ_1, and any $1 \leq q \leq 2$, we have*:

$$\mathrm{QJT}_q(\rho_0,\rho_1) \leq \mathrm{H}_q\left(\frac{1}{2}\right) \cdot \frac{1}{2}\operatorname{tr}(|\rho_0 - \rho_1|^q) \leq \mathrm{H}_q\left(\frac{1}{2}\right) \cdot \mathrm{T}(\rho_0,\rho_1)^q.$$

Proof. We begin with the construction for establishing $\mathrm{QJT}_q \leq \ln 2 \cdot \mathrm{T}$ for $q = 1$ as in [BH09, Theorem 14]. Our analysis differs since we need to address the cases of $1 \leq q \leq 2$. Consider a single qutrit register B with basis vectors $|0\rangle, |1\rangle, |2\rangle$. Define $\hat{\rho}_0$ and $\hat{\rho}_1$ on $\mathcal{H} \otimes \mathcal{B}$ as below, where $\mathcal{B} = \mathbb{C}^3$ is the Hilbert space corresponding to the register B:

$$\hat{\rho}_0 \coloneqq \frac{\rho_0 + \rho_1 - |\rho_0 - \rho_1|}{2} \otimes |2\rangle\langle 2| + \frac{\rho_0 - \rho_1 + |\rho_0 - \rho_1|}{2} \otimes |0\rangle\langle 0| \coloneqq \quad \sigma_2 \otimes |2\rangle\langle 2| + \sigma_0 \otimes |0\rangle\langle 0|,$$
$$\hat{\rho}_1 \coloneqq \frac{\rho_0 + \rho_1 - |\rho_0 - \rho_1|}{2} \otimes |2\rangle\langle 2| + \frac{\rho_1 - \rho_0 + |\rho_0 - \rho_1|}{2} \otimes |1\rangle\langle 1| \coloneqq \quad \sigma_2 \otimes |2\rangle\langle 2| + \sigma_1 \otimes |1\rangle\langle 1|.$$

Intuitively, σ_b represents the case where ρ_b is "larger than" ρ_{1-b} for $b \in \{0,1\}$ (i.e., ρ_0 and ρ_1 are "distinguishable"), while σ_2 represents the case where ρ_0 is "indistinguishable" from ρ_1. This construction generalizes the proof of the classical analogs (e.g., [Vad99, Claim 4.4.2]).

Noting that QJT_q is contractive when applying a partial trace (Lemma 4.5), we obtain:

$$\begin{aligned}\mathrm{QJT}_q(\rho_0,\rho_1) &= \mathrm{QJT}_q(\mathrm{tr}_{\mathsf{B}}(\hat{\rho}_0),\mathrm{tr}_{\mathsf{B}}(\hat{\rho}_1))\\ &\le \mathrm{QJT}_q(\hat{\rho}_0,\hat{\rho}_1)\\ &= \frac{1}{q-1}\left(\mathrm{tr}\left(\left(\frac{\hat{\rho}_0+\hat{\rho}_1}{2}\right)^q\right)-\frac{1}{2}\mathrm{tr}(\hat{\rho}_0^q)-\frac{1}{2}\mathrm{tr}(\hat{\rho}_1^q)\right).\end{aligned} \tag{4.4}$$

Noting that $\sigma_0\otimes|0\rangle\langle 0|$, $\sigma_1\otimes|1\rangle\langle 1|$, and $\sigma_2\otimes|2\rangle\langle 2|$ are orthogonal to each other, we have:

$$\mathrm{tr}\left(\left(\frac{\hat{\rho}_0+\hat{\rho}_1}{2}\right)^q\right)=\mathrm{tr}\left(\left(\sigma_2\otimes|2\rangle\langle 2|+\sum_{b\in\{0,1\}}\frac{\sigma_b}{2}\otimes|b\rangle\langle b|\right)^q\right)=\mathrm{tr}\left(\sigma_2^q+\frac{\sigma_0^q}{2^q}+\frac{\sigma_1^q}{2^q}\right), \tag{4.5}$$

$$\forall b\in\{0,1\},\ \mathrm{tr}(\hat{\rho}_b^q)=\mathrm{tr}\left((\sigma_2\otimes|2\rangle\langle 2|+\sigma_b\otimes|b\rangle\langle b|)^q\right)=\mathrm{tr}\left(\sigma_2^q+\sigma_b^q\right).$$

Plugging Equation (4.5) and the equality $\mathrm{H}_q\left(\frac{1}{2}\right)=\frac{1-2^{1-q}}{q-1}$ into Equation (4.4), we obtain:

$$\begin{aligned}\mathrm{QJT}_q(\rho_0,\rho_1) &\le \mathrm{H}_q\left(\frac{1}{2}\right)\cdot\frac{1}{2}\mathrm{tr}(\sigma_0^q+\sigma_1^q)\\ &\le \mathrm{H}_q\left(\frac{1}{2}\right)\cdot\frac{1}{2}(\mathrm{tr}(\sigma_0)^q+\mathrm{tr}(\sigma_1)^q)\\ &= \mathrm{H}_q\left(\frac{1}{2}\right)\cdot\mathrm{T}(\rho_0,\rho_1)^q.\end{aligned} \tag{4.6}$$

Here, the second line is due to the monotonicity of the Schatten p-norm (e.g., [AS17, Equation (1.31)]), equivalently, $\mathrm{tr}(M^q)\le\mathrm{tr}(M)^q$ for any positive semi-definite matrix M and $q\ge 1$. The last line owes to the fact that

$$\mathrm{tr}(\sigma_b)=(-1)^b\,\mathrm{tr}\left(\frac{\rho_0-\rho_1}{2}\right)+\frac{1}{2}\mathrm{tr}(|\rho_0-\rho_1|)=\frac{1}{2}\mathrm{tr}(|\rho_0-\rho_1|)\text{ for }b\in\{0,1\}.$$

Lastly, since σ_0 and σ_1 are orthogonal to each other, we complete the proof by plugging the equality $\mathrm{tr}(\sigma_0^q+\sigma_1^q)=\mathrm{tr}((\sigma_0+\sigma_1)^q)=\mathrm{tr}(|\rho_0-\rho_1|^q)$ into the first line in Equation (4.6). □

4.3 Bounds for the Tsallis binary entropy

In this subsection, we establish lower and upper bounds (Lemma 4.8 and Lemma 4.9, respectively) for the Tsallis binary entropy, which are useful when applying the lower bound on QJT_q in terms of the trace distance (Lemma 4.6).

We begin by proving an lower bound, which extends the bound $\mathrm{H}(1/2)\cdot 4x(1-x)\le\mathrm{H}(x)$ from the specific case of $q=1$, as stated in [Top01, Theorem 1.2], to a broader range of q:

Lemma 4.8 (Tsallis binary entropy lower bound)**.** *For any $p=(x,1-x)$, let $\mathrm{H}_q(x)$ denote the Tsallis binary entropy with $q\in[0,2]\cup[3,+\infty)$, we have*:

$$\mathrm{H}_q(1/2)\cdot 4x(1-x)\le\mathrm{H}_q(x).$$

Proof. We need only consider the cases where $q\in\mathcal{I}:=[0,1)\cup(1,2]\cup[3,+\infty)$, as the case $q=1$ directly follows from [Top01, Theorem 1.2]. Our proof strategy is inspired by the approach used in that theorem. We start by defining functions $F(q;x)$ and $G(q;x)$ on $0\le x\le 1$ and $q\in\mathcal{I}$:

$$F(q;x):=\frac{\mathrm{H}_q(x)}{x(1-x)}=\frac{1-x^q-(1-x)^q}{(q-1)x(1-x)}\text{ and }G(q;x):=\frac{x^{q-1}-1}{(q-1)(x-1)}.$$

It is evident that $F(q;0)=F(q;1)=\infty$ and $F(q;1/2)=4\mathrm{H}_q(1/2)$. We then assume that $G(q;x)$ is convex on $x\in[0,1]$ for any fixed $q\in\mathcal{I}$:

$$\text{For any }x\in[0,1]\text{ and }q\in(1,2],\ \frac{\partial^2 G(q;x)}{\partial x^2}=\frac{(q-2)x^{q-3}}{x-1}-\frac{2x^{q-2}}{(x-1)^2}+\frac{2(x^{q-1}-1)}{(x-1)^3(q-1)}\ge 0. \tag{4.7}$$

Since $F(q;x) = G(q;x)+G(q;1-x)$, Equation (4.7) implies that $F(q,x)$ is convex on $x \in [0,1]$ for any fixed $q \in \mathcal{I}$. By noticing that $F(q;x) = F(q;1-x)$ for any $x \in [0,1]$, we can obtain that: for any $q \in \mathcal{I}$, $F(q;x)$ is monotonically decreasing on $x \in (0,1/2)$ and monotonically increasing on $x \in (1/2,1)$. Consequently, we establish the lower bound by noticing that:

$$\text{For any } x \in [0,1] \text{ and } q \in \mathcal{I}, F(q;x) \geq F(q;1/2) = 4\mathrm{H}_q(1/2).$$

It remains to prove Equation (4.7). Noting that $(x-1)^3 \leq 0$ for any $0 \leq x \leq 1$, Equation (4.7) holds if and only if the following holds:

$$f(q;x) \coloneqq (q-2)(x-1)^2x^{q-3} - 2(x-1)x^{q-2} + \frac{2(x^{q-1}-1)}{q-1} \leq 0.$$

A direct calculation implies that $\frac{\partial}{\partial x} f(q;x) = (q-3)(q-2)(x-1)^2x^{q-4} \geq 0$ for any $q \in \mathcal{I}$ and $x \in [0,1]$ since $\mathcal{I} \cup (2,3) = \emptyset$. Hence, for any fixed $q \in \mathcal{I}$, $f(q;x)$ is monotonically increasing for any $x \in (0,1)$. Therefore, we complete the proof by concluding that

$$\max_{x \in [0,1]} f(q;x) \leq f(q,1) = 0. \qquad \square$$

Next, we will demonstrate an upper bound for the range of $1 < q \leq 2$ that is weaker than the best known upper bound for the case of $q = 1$ as shown in [Top01, Theorem 1.2]:[27]

Lemma 4.9 (Tsallis binary entropy upper bound)**.** *For any $p = (x, 1-x)$, let $\mathrm{H}_q(x)$ denote the Tsallis binary entropy with $1 \leq q \leq 2$, we have*:

$$\mathrm{H}_q(x) \leq \mathrm{H}_q(1/2) \cdot (4x(1-x))^{1/2}.$$

Proof. The case of $q = 1$ follows directly from [Lin91, Theorem 8], it remains to address the range $1 < q \leq 2$. We will establish the bound separately for $x \in \mathcal{I}_{\text{inner}}$ and $x \in \mathcal{I}_{\text{outer}}$, where $\mathcal{I}_{\text{inner}} \cup \mathcal{I}_{\text{outer}} = [0,1]$. Specifically, these intervals are defined as $\mathcal{I}_{\text{inner}} \coloneqq [0,1/8] \cup [1/8,1]$ and $\mathcal{I}_{\text{outer}} \coloneqq [1/2 - \tau(q), 1/2 + \tau(q)]$, where $\tau(q)$ will be specified latter.

The outer interval case. We start with the case of $x \in \mathcal{I}_{\text{outer}}$. Since $\mathrm{H}_q(x) = \mathrm{H}_q(1-x)$ for any $0 \leq x \leq 1$, it is sufficient to consider the case of $0 \leq x \leq 1/8$. Noting that $q - 1 \geq 0$, it suffices to show that: For any $0 \leq x \leq 1/8$ and $1 < q \leq 2$,

$$(q-1)\left(\mathrm{H}_q\left(\frac{1}{2}\right)\sqrt{4x(1-x)} - \mathrm{H}_q(x)\right) = \left(2 - 2^{2-q}\right)\sqrt{x(1-x)} - (1 - x^q - (1-x)^q) \geq 0. \quad (4.8)$$

Leveraging the Taylor expansion of $1-(1-x)^q$ at $x = 0$, we obtain that:

$$1 - (1-x)^q = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k!} \prod_{r=0}^{k-1}(q-r)x^k \coloneqq \sum_{k=1}^{\infty} \alpha_k x^k \leq qx. \quad (4.9)$$

Here, notice that $1 < q \leq 2$, the last inequality owes to the fact that $\alpha_1 = q > 0$ and $\alpha_k \leq 0$ for all integer $k \geq 2$. Plugging Equation (4.9) into Equation (4.8), it remains to prove that:

$$F_1(q;x) \coloneqq \frac{-x^q + qx}{\sqrt{x(1-x)}} \leq 2 - 2^{2-q}.$$

A direct calculation implies that $F_1(q;1/8) = (q - 2^{3-3q})/\sqrt{7}$ satisfies $2 - 2^{2-q} - F_1(q;1/8) > 0$ for $1 < q \leq 2$.[28] As a consequence, it is enough to show that $F_1(q,x)$ is monotonically non-

[27] Numerical evidence suggests that Lemma 4.9 can be improved to $\mathrm{H}_q(x) \leq \mathrm{H}_q\left(\frac{1}{2}\right) \cdot (4x(1-x))^{\frac{1}{2\mathrm{H}_q(1/2)}}$ for any $0 \leq x \leq 1$ and $1 \leq q \leq 2$, which coincides with the bound $\mathrm{H}(q) \leq \mathrm{H}(1/2)(4x(1-x))^{\frac{1}{2\mathrm{H}_q(1/2)}}$ in [Top01].

[28] It is noteworthy that $2 - 2^{2-q} - F_1(q;1/2) = 2 - 2^{1-q} - q < 0$ for $1 < q \leq 2$, and consequently, the outer interval case is not enough to establish our Tsallis binary entropy upper bound for any $0 \leq x \leq 1$.

decreasing on $x \in [0, 1/8]$ for any fixed $q \in (1, 2]$, specifically:

$$\frac{\partial}{\partial x} F_1(q; x) = \frac{1}{2}(x(1-x))^{-3/2}\left(qx + (1 + 2q(x-1) - 2x)x^q\right) \geq 0. \tag{4.10}$$

Noting that $\frac{1}{2}(x(1-x))^{-3/2} \geq 0$, Equation (4.10) holds if and only if the following holds:

$$F_2(q; x) \coloneqq (2(1-q)x + 2q - 1)\, x^{q-1} \leq q.$$

A direct calculation implies that $F_2(q; 1/8) = 2^{1-3q}(7q - 3)$ satisfies that $q - F_2(q; 1/8) > 0$ for $1 < q \leq 2$. Consequently, it suffices to show that $F_2(q; x)$ is monotonically non-decreasing on $x \in [0, 1/8]$ for any fixed $q \in (1, 2]$, particularly:

$$\frac{\partial}{\partial x} F_2(q; x) = x^{q-2}(q-1)(2q(1-x) - 1) \geq 0. \tag{4.11}$$

Since $x^{q-2}(q-1) > 0$ for any $q \in (1, 2]$ and $x \in [0, 1/8]$, Equation (4.11) holds if and only if $F_3(q; x) \coloneqq 2q(1-x) - 1 \geq 0$. It is evident that $F_3(q; x) \geq 0$ is equivalent to $x \leq 1 - 1/2q < 1/2$ for $1 < q \leq 2$, and thus we complete the proof of the outer interval case.

The inner interval case. Next, we move to the case of $x \in \mathcal{I}_{\text{inner}}$. Let $x = (1+t)/2$, then it suffices to consider the case of $0 \leq t \leq 1$ since $\mathrm{H}_q(x) = \mathrm{H}_q(1-x)$ for any $0 \leq x \leq 1$. Noting that $2^q/(q-1) > 0$ for $1 < q \leq 2$, it suffices to show that: For any $0 \leq t \leq 2\tau(q)$ and $1 < q \leq 2$,

$$\frac{2^q}{q-1}\left(\mathrm{H}_q\Big(\frac{1}{2}\Big)\sqrt{4x(1-x)} - \mathrm{H}_q(x)\right) = (1-t)^q + (1+t)^q - \left(2^q + (2-2^q)\sqrt{1-t^2}\right) \geq 0. \tag{4.12}$$

Utilizing the Taylor expansion of $(1-t)^q + (1+t)^q$ at $t = 0$, we obtain that:

$$(1-t)^q + (1+t)^q = \sum_{k=0}^{\infty} \frac{2}{(2k)!} \prod_{r=0}^{2k-1} (q-r)t^{2k} \coloneqq \sum_{k=0}^{\infty} \beta_k t^{2k} \geq \beta_0 + \beta_1^2 = 2 + q(q-1)t^2. \tag{4.13}$$

Here, the last inequality is because $\beta_k \geq 0$ for all integer $k \geq 0$. Substituting Equation (4.13) into Equation (4.12), it remains to show that:

$$2 + q(q-1)t^2 \geq 2^q + (2-2^q)\sqrt{1-t^2}. \tag{4.14}$$

A direct calculation implies that Equation (4.14) holds for the following range of t:

$$|t| \leq \frac{\sqrt{(2q^2 - 2q + 2 - 2^q)(2^q - 2)}}{q(q-1)} = 2\tau(q) \text{ where } \tau(q) \coloneqq \frac{\sqrt{(q^2 - q + 1 - 2^{q-1})(2^{q-1} - 1)}}{q(q-1)}.$$

It is easy to see that $\lim_{q \to 1^+} \tau(q) = \sqrt{\ln 2(1 - \ln 2)} \approx 0.4612$ and $\tau(2) = 1$. Assume that $\tau(q)$ is monotonically non-decreasing for $q \in (1, 2]$, we obtain that $[1/2 - \sqrt{\ln 2(1-\ln 2)}, 1/2 + \sqrt{\ln 2(1-\ln 2)}] \subseteq \mathcal{I}_{\text{inner}}$, and consequently, $\mathcal{I}_{\text{inner}} \cup \mathcal{I}_{\text{outer}} = [0, 1]$.

It is left to show that $\tau(q)$ is monotonically non-decreasing for $q \in (1, 2]$, specifically:

$$\frac{\mathrm{d}}{\mathrm{d}q}\tau(q) = \frac{2}{q^3(q-1)^3} \underbrace{\left(2^q - q^2 + q - 2\right)}_{g_1(q)} \underbrace{\left(2 - 2^q - 2^q q^2 \ln 2 + q\left(2^{1+q} + 2q\ln 2 - 4\right)\right)}_{g_2(q)} \geq 0. \tag{4.15}$$

Note that $g_1(q) = 0$ corresponds to an intersection between a quadratic function and an exponential function, indicating that $g_1(q)$ has at most three zeros. It is evident that $g_1(1) = g_1(2) = g_1(3) = 0$ and $g(3/2) = 2\sqrt{2} - 11/4 \approx 0.078$, and thus $g_1(q) \geq 0$ for $1 \leq q \leq 2$.

For $g_2(q)$, notice that $g_2(1) = 0$. Assuming that $g_2(q)$ is monotonically non-decreasing on $1 \leq q \leq 2$, we obtain $g_2(q) \geq g_2(1) = 0$ for $1 \leq q \leq 2$. In particular, it remains to prove that:

$$\text{For any } q \in [1, 2],\ g_3(q) \coloneqq \frac{\mathrm{d}}{\mathrm{d}q} g_2(q) = 2^q(-(\ln 2)^2 q^2 + (\ln 2)^2 q + 2) - 4 \geq 0. \tag{4.16}$$

Since $g_3(q) + 4$ is the product of a quadratic function and an exponential function, $g_3(q)$

has at most two zeros. Therefore, we establish Equation (4.16), and thus Equation (4.15), by noticing that $g_3(1) = 0$, $g_3(2) = 4 - 8(\ln 2)^2 > 0$, and $g_3(3) = 12 - 48(\ln 2)^2 < 0$. □

4.4 Useful bounds on Tsallis entropy

In this subsection, we present a useful bound on Tsallis entropy. Lemma 4.10 establishes inequalities between the Tsallis entropy of a distribution p and the total variation distance between p and the uniform distribution of the same dimension.

Lemma 4.10 (Tsallis entropy bounds by closeness to uniform distribution)**.** *Let p be a probability distribution over $[N]$ with $N \geq 2$, and let ν be the uniform distribution over $[N]$. Then, for any $q > 1$ and $0 \leq \mathrm{TV}(p, \nu) \leq 1 - 1/N$, it holds that*:

$$(1 - \mathrm{TV}(p, \nu) - 1/N) \ln_q(N) \leq \mathrm{H}_q(p).$$

Moreover, for any $q > 1$ and N satisfying $1/q \leq \mathrm{TV}(p, \nu) \leq 1 - 1/N$, it holds that:

$$\mathrm{H}_q(p) \leq \ln_q\big(N(1 - \mathrm{TV}(p, \nu))\big).$$

Proof. Let $\gamma \coloneqq \mathrm{TV}(p, \nu)$, and let Δ_N be the set of probability distributions of dimension N. It is evident that $0 \leq \mathrm{TV}(p, \nu) \leq 1 - 1/N$. To establish the lower bound, it suffices to minimize the Tsallis entropy $\mathrm{H}_q(p)$ subject to the constraint $\mathrm{TV}(p, \nu) = \gamma$, which is equivalent to solve the convex optimization problem in Equation (4.17).[29]

$$\begin{aligned} \text{minimize} \quad & \mathrm{H}_q(p') \\ \text{subject to} \quad & p' \in \Delta_n, \\ & \mathrm{TV}(p', \nu) \leq \gamma \end{aligned} \tag{4.17}$$

$$\begin{aligned} p_{\min}(i) &= \begin{cases} \frac{1}{N}, & \text{if } i \in [k_{\min}] \\ \frac{1}{N} + \gamma, & \text{if } i = k_{\min} + 1 \\ \frac{\varepsilon}{N}, & \text{if } i = k_{\min} + 2 \\ 0, & \text{otherwise} \end{cases}, \\ \text{where} \quad & k_{\min} \coloneqq \lfloor N(1-\gamma) \rfloor - 1, \\ & \varepsilon \coloneqq N(1-\gamma) - \lfloor N(1-\gamma) \rfloor. \end{aligned} \tag{4.18}$$

Note that $\mathrm{H}_q(p)$ is concave (Lemma 2.3) for any fixed $q > 1$, and the constraints in Equation (4.17) form a closed convex set. Since the minimum of a concave function is attained at some extreme point (e.g., [Roc70, Corollary 32.3.1]) and the Tsallis entropy is permutation-invariant, we deduce an optimal solution $p_{\min}$ to Equation (4.17), as stated in Equation (4.18).

Next, we can deduce the lower bound of the Tsallis entropy by evaluating $\mathrm{H}_q(p_{\min})$:

$$\begin{aligned} \mathrm{H}_q(p_{\min}) &= \frac{1}{q-1}\left(1 - k_{\min}\Big(\frac{1}{N}\Big)^q - \Big(\frac{1}{N} + \gamma\Big)^q - \Big(\frac{\varepsilon}{N}\Big)^q\right) \\ &\geq \frac{1}{q-1}\left(\frac{(\lfloor N(1-\gamma) \rfloor - 1)}{N} + \frac{\varepsilon}{N} - (\lfloor N(1-\gamma) \rfloor - 1 + \varepsilon^q)\Big(\frac{1}{N}\Big)^q\right) \\ &\geq \frac{1}{q-1}\left(1 - \gamma - \frac{1}{N} - \Big(1 - \gamma - \frac{1}{N}\Big)\Big(\frac{1}{N}\Big)^{q-1}\right) \\ &= \left(1 - \gamma - \frac{1}{N}\right)\ln_q(N). \end{aligned}$$

Here, the second line excludes terms corresponding to $p_{\min}(k_{\min} + 1)$, and the third line follows from the fact that $\varepsilon^q \leq \varepsilon$ for $q \geq 1$ and $0 \leq \varepsilon \leq 1$.

To demonstrate the upper bound, it remains to maximize the Tsallis entropy $\mathrm{H}_q(p)$ subject to the constraint $\mathrm{TV}(p, \nu) = \gamma$, which is equivalent to solve a *non-convex* optimization problem

[29] A similar formulation also appeared in the proof of [KLGN19, Lemma 16].

analogous to Equation (4.17). This task is challenging in general, but we consider only the regime $\mathrm{TV}(p,\nu) \geq 1/q$.[30]. Particularly, we focus on the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & \mathrm{H}_q(p') \\ \text{subject to} \quad & p' \in \Delta_n, \\ & \mathrm{TV}(p',\nu) \geq \gamma \geq 1/q \end{aligned} \tag{4.19}$$

It is not too hard to obtain an optimal solution $p_{\max}$ to Equation (4.19), where ε is defined as in Equation (4.18), as stated in Proposition 4.10.1. The proof is deferred in Appendix A.1.

Proposition 4.10.1. *For the optimization problem presented in Equation (4.19), an optimal solution is the distribution provided in Equation (4.20), where $\varepsilon = N(1-\gamma) - \lfloor N(1-\gamma) \rfloor$:*

$$p_{\max}(i) = \begin{cases} \frac{1}{N} + \frac{\gamma}{k_{\max}}, & \text{if } i \in [k_{\max}] \\ \frac{\varepsilon}{N(N-k_{\max})}, & \text{otherwise} \end{cases}, \text{where } k_{\max} \coloneqq \lfloor N(1-\gamma) \rfloor. \tag{4.20}$$

Consequently, we can derive the upper bound of the Tsallis entropy by evaluating $\mathrm{H}_q(p_{\max})$:

$$\begin{aligned} \mathrm{H}_q(p_{\max}) &= \frac{1}{q-1}\left(1 - k_{\max}\Big(\frac{1}{N} + \frac{\gamma}{k_{\max}}\Big)^q - (N-k_{\max})\Big(\frac{\varepsilon}{N(N-k_{\max})}\Big)^q\right) \\ &= \frac{1}{q-1}\left(1 - \Big(1-\frac{\varepsilon}{N}\Big)^q\Big(\frac{1}{(N(1-\gamma)-\varepsilon}\Big)^{q-1} - \Big(\frac{\varepsilon}{N}\Big)^q\Big(\frac{1}{N\gamma+\varepsilon}\Big)^{q-1}\right) \\ &\leq \frac{1}{q-1}\left(1 - \Big(\frac{1}{(N(1-\gamma)}\Big)^{q-1}\right) \\ &= \ln_q\Big(N(1-\gamma)\Big). \end{aligned}$$

Let $F(q;N,\varepsilon,\gamma) \coloneqq \big(1-\frac{\varepsilon}{N}\big)^q (N(1-\gamma)-\varepsilon)^{1-q} + \big(\frac{\varepsilon}{N}\big)^q (N\gamma+\varepsilon)^{1-q}$, then the third line holds by assuming that $F(q;N,\varepsilon,\gamma)$ is monotonically non-decreasing on $0 \leq \varepsilon \leq 1$ for any fixed γ, q, and N satisfying $q\gamma \geq 1$ and $N \geq q/(q-1)$.

It remains to prove $\frac{\partial}{\partial\varepsilon}F(q;N,\varepsilon,\gamma) \geq 0$ the aforementioned range of x, γ, q, and N. By a direct calculation, we complete the proof by noticing all terms in the following are non-negative:

$$\frac{\partial}{\partial\varepsilon}F(q;N,\varepsilon,\gamma) = \Big(1-\frac{\epsilon}{N}\Big)^q \frac{(N(\gamma q-1)+\epsilon)}{(N-\epsilon)(N(1-\gamma)-\epsilon)^q} + \Big(\frac{\epsilon}{N}\Big)^q \frac{(\gamma N q+\epsilon)}{\epsilon(\gamma N+\epsilon)^q} \geq 0. \qquad \square$$

5 Hardness and lower bounds via QJT_q-based reductions

In this section, we will establish reductions from the closeness testing of quantum states via the trace distance to testing via the quantum q-Tsallis entropy difference. Our proof crucially depends on the properties of the quantum Jensen-Tsallis divergence (QJT_q) demonstrated in Section 4. Using these reductions, we will prove computational hardness results and query complexity lower bounds for several problems related to the quantum q-Tsallis entropy difference under various circumstances.

We begin by defining the Quantum q-Tsallis Entropy Difference and the Quantum q-Tsallis Entropy Approximation, denoted by $\mathrm{TsallisQED}_q[g(n)]$ and $\mathrm{TsallisQEA}_q[t(n),g(n)]$, respectively. These definitions generalize the counterpart definitions in [BASTS10] from the von Neumann entropy (i.e., QJT_q with $q=1$) to the quantum q-Tsallis entropy for $1 \leq q \leq 2$.

Definition 5.1 (Quantum q-Tsallis Entropy Difference, $\mathrm{TsallisQED}_q$)**.** *Let Q_0 and Q_1 be quantum circuits acting on m qubits and having n specified output qubits, where $m(n)$ is a polynomial in n. Let ρ_i be the quantum state obtained by running Q_i on $|0\rangle^{\otimes m}$ and tracing out the non-output qubits. Let $g(n)$ be a positive efficiently computable function. Decide whether*:

[30] For the regime $0 \leq \mathrm{TV}(p,\nu) \leq 1/q$, the optimal solution to Equation (4.19) depends on the choice of q.

- Yes: *A pair of quantum circuits* (Q_0, Q_1) *such that* $\mathrm{S}_q(\rho_0) - \mathrm{S}_q(\rho_1) \geq g(n)$*;*
- No: *A pair of quantum circuits* (Q_0, Q_1) *such that* $\mathrm{S}_q(\rho_1) - \mathrm{S}_q(\rho_0) \geq g(n)$*.*

Definition 5.2 (Quantum q-Tsallis Entropy Approximation, $\textsc{TsallisQEA}_q$)**.** *Let Q be a quantum circuit acting on m qubits and having n specified output qubits, where $m(n)$ is a polynomial in n. Let ρ be the quantum state obtained by running Q on $|0\rangle^{\otimes m}$ and tracing out the non-output qubits. Let $g(n)$ and $t(n)$ be positive efficiently computable functions. Decide whether*:

- Yes: *A quantum circuit Q such that* $\mathrm{S}_q(\rho) \geq t(n) + g(n)$*;*
- No: *A quantum circuit Q such that* $\mathrm{S}_q(\rho) \leq t(n) - g(n)$*.*

Notably, the quantum q-Tsallis entropy of any pure state is zero. Hence, similar to Section 2.2, it is reasonable to define *constant-rank* variants of $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$:

(1) $\textsc{ConstRankTsallisQED}_q$: the ranks of ρ_0 and ρ_1 are at most $O(1)$.

(2) $\textsc{ConstRankTsallisQEA}_q$: the rank of ρ is at most $O(1)$.

Next, we present the main theorem in this section:

Theorem 5.3 (Computational hardness for $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$)**.** *The promise problems $\textsc{TsallisQED}_q$ and $\textsc{TsallisQEA}_q$ capture the computational power of their respective complexity classes in the corresponding regimes of q*:

(1) *For any $q \in [1, 2]$ and $n \geq 3$, it holds that: For $1/\operatorname{poly}(n) \leq g_q(n) \leq 2^q \mathrm{H}_q(1/2)\left(1 - 2^{-\frac{qn}{2}+1}\right)$, $\textsc{ConstRankTsallisQED}_q[g_q(n)]$ is BQP-hard under Karp reduction. Consequently, $\textsc{ConstRankTsallisQEA}_q$ with $g(n) = \Theta(1)$ is BQP-hard under Turing reduction.*

(2) *For any $q \in \left(1, 1 + \frac{1}{n-1}\right]$ and $n \geq 90$, it holds that: For $1/\operatorname{poly}(n) \leq g(n) \leq 1/400$, $\textsc{TsallisQED}_q[g(n)]$ is QSZK-hard under Karp reduction. Consequently, $\textsc{TsallisQEA}_q$ with $g(n) = \Theta(1)$ is QSZK-hard under Turing reduction.*

(3) *For any $n \geq 5$, it holds that: For $1/\operatorname{poly}(n) \leq g(n) \leq 1/150$, $\textsc{TsallisQEA}_{1+\frac{1}{n-1}}$ with $g(n)$ is NIQSZK-hard.*

In particular, Theorem 5.3(1) is derived from the pure-state reduction (Lemma 5.4), and the detailed statements are Theorem 5.7 and Theorem 5.8. Moreover, Theorem 5.3(2) is obtained through a mixed-state reduction (Lemma 5.5), and the detailed statements are Theorem 5.9 and Theorem 5.10. Furthermore, Theorem 5.3(3) follows from a tailor-made mixed state reduction for QSCMM (Lemma 5.6), and the detailed statement is Theorem 5.11.

Lastly, using the reductions in Lemma 5.5, we derive lower bounds on the quantum query and sample complexity for estimating $\mathrm{S}_q(\rho)$ where $1 < q \leq 1 + \frac{1}{n-1}$, as presented in Theorem 5.13 and Theorem 5.16. These theorems build on prior works in quantum query complexity [CFMdW10] and sample complexity [OW21] lower bounds for the trace distance. In addition, we provide quantum query and sample complexity lower bounds for estimating $\mathrm{S}_q(\rho)$ when $q \geq 1 + \Omega(1)$, leveraging the hard instances from [Bel19], as detailed in Theorem 5.12 and Theorem 5.15.

5.1 Pure-state reduction: $\textsc{PureQSD} \leq \textsc{ConstRankTsallisQED}_q$ for $1 \leq q \leq 2$

The reduction in Lemma 5.4 is from the trace distance between two n-qubit pure states ($\textsc{PureQSD}$) to the quantum q-Tsallis entropy difference between two new constant-rank $(n+1)$-qubit states ($\textsc{ConstRankTsallisQED}_q$), for $1 \leq q \leq 2$.

Lemma 5.4 ($\textsc{PureQSD} \leq \textsc{ConstRankTsallisQED}_q$)**.** *Let Q_0 and Q_1 be quantum circuits acting on n qubits and having the same number of output qubits. Let $|\psi_i\rangle$ be the quantum state obtained by running Q_i on $|0\rangle^{\otimes n}$. For any $b \in \{0, 1\}$, there is a new quantum circuit Q'_b acting*

on $n+3$ qubits, using $O(1)$ queries to controlled-Q_0 and controlled-Q_1, as well as $O(1)$ one- and two-qubit gates. The circuit Q'_b prepares a new quantum state ρ'_b, which has constant rank and acts on $n' := n+1$ qubits, such that for any efficiently computable functions $\alpha(n)$ and $\beta(n)$, where $\beta(n) + \sqrt{1-\alpha(n)^2} < 1$, and any $q \in [1,2]$, the following holds:

$$\begin{aligned} \mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \geq \alpha(n) \;&\Rightarrow\; \mathrm{S}_q(\rho'_0) - \mathrm{S}_q(\rho'_1) \geq g_q(n') = g_q(n+1), \\ \mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \leq \beta(n) \;&\Rightarrow\; \mathrm{S}_q(\rho'_1) - \mathrm{S}_q(\rho'_0) \geq g_q(n') = g_q(n+1), \end{aligned}$$

where $g_q(n+1) := 2^{-q} \cdot \mathrm{H}_q(1/2) \cdot \left(1 - \beta(n)^q - \sqrt{1-\alpha(n)^2}\right)$.

Proof. Our proof strategy is inspired by the proof of [Liu23, Corollary 4.3 and Lemma 4.4]. We begin by considering the following constant-rank quantum states ρ'_0 and ρ'_1, which can be prepared by the quantum circuits Q'_0 and Q'_1, respectively:

$$\begin{aligned} \rho'_0 &:= (p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1|) \otimes \frac{1}{2}\left(|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|\right) \\ \rho'_1 &:= \frac{1}{2}|0\rangle\langle 0| \otimes |\psi_0\rangle\langle\psi_0| + \frac{1}{2}|1\rangle\langle 1| \otimes |\psi_1\rangle\langle\psi_1|. \end{aligned}$$

Here, (p_0, p_1) is some two-element probability distribution that will be specified later. Moreover, for any $b \in \{0,1\}$, the quantum circuit Q'_b uses $O(1)$ queries to controlled-Q_0 and controlled-Q_1 as well as $O(1)$ one- and two-qubit gates, as presented in [Liu23, Figure 1 and Figure 2].

Using the pseudo-additivity of S_q (Lemma 2.9), we can obtain that:

$$\begin{aligned} \mathrm{S}_q(\rho'_0) &= \mathrm{H}_q(p_0) + \mathrm{S}_q\left(\frac{|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|}{2}\right) - (q-1)\cdot \mathrm{H}_q(p_0) \cdot \mathrm{S}_q\left(\frac{|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|}{2}\right) \\ &= \mathrm{H}_q(p_0) + (1-(q-1)\mathrm{H}_q(p_0)) \cdot \mathrm{S}_q\left(\frac{|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|}{2}\right). \end{aligned} \tag{5.1}$$

By the joint q-Tsallis entropy theorem (Lemma 2.12), we have:

$$\mathrm{S}_q(\rho'_1) = \mathrm{H}_q(1/2) + 2^{-q}(\mathrm{S}_q(|\psi_0\rangle\langle\psi_0|) + \mathrm{S}_q(|\psi_1\rangle\langle\psi_1|)) = \mathrm{H}_q(1/2). \tag{5.2}$$

Combining Equation (5.1) and Equation (5.2), we conclude that:

$$\begin{aligned} \mathrm{S}_q(\rho'_0) - \mathrm{S}_q(\rho'_1) &= (1-(q-1)\mathrm{H}_q(p_0)) \cdot \mathrm{S}_q\left(\frac{|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|}{2}\right) + \mathrm{H}_q(p_0) - \mathrm{H}_q\left(\frac{1}{2}\right) \\ &= (1-(q-1)\mathrm{H}_q(p_0)) \cdot \mathrm{QJT}_q(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) + \mathrm{H}_q(p_0) - \mathrm{H}_q\left(\frac{1}{2}\right). \end{aligned} \tag{5.3}$$

Next, we choose $p_0 \in (0, 1/2)$ satisfying the following equality:

$$\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q(p_0) = \frac{1-(q-1)\mathrm{H}_q(p_0)}{2}\left(\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1-\alpha}{2}\right) + \mathrm{H}_q\left(\frac{1}{2}\right)\cdot\beta^q\right). \tag{5.4}$$

As a consequence, we can derive that:

- For the case where $\mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \geq \alpha$, plugging the lower bound on QJT_q in terms of the trace distance (Lemma 4.6) into Equation (5.3) and Equation (5.4), we obtain

$$\begin{aligned} \mathrm{S}_q(\rho'_0) - \mathrm{S}_q(\rho'_1) &\geq (1-(q-1)\mathrm{H}_q(p_0)) \cdot \left(\mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1-\alpha}{2}\right)\right) + \mathrm{H}_q(p_0) - \mathrm{H}_q\left(\frac{1}{2}\right) \\ &= \frac{1-(q-1)\mathrm{H}_q(p_0)}{2}\left(\mathrm{H}_q\left(\frac{1}{2}\right)\cdot(1-\beta^q) - \mathrm{H}_q\left(\frac{1-\alpha}{2}\right)\right) := \tilde{g}_q. \end{aligned}$$

- For the case where $\mathrm{T}(|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|) \leq \beta$, plugging the upper bound on QJT_q in terms

of the trace distance (Lemma 4.7) into Equation (5.3) and Equation (5.4), we obtain

$$\begin{aligned}\mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1') &\le (1-(q-1)\mathrm{H}_q(p_0)) \cdot \beta^q \cdot \mathrm{H}_q\Big(\frac{1}{2}\Big) + \mathrm{H}_q(p_0) - \mathrm{H}_q\Big(\frac{1}{2}\Big) \\ &= -\frac{1-(q-1)\mathrm{H}_q(p_0)}{2}\left(\mathrm{H}_q\Big(\frac{1}{2}\Big)\cdot(1-\beta^q) - \mathrm{H}_q\Big(\frac{1-\alpha}{2}\Big)\right) = -\tilde{g}_q.\end{aligned}$$

It is left to show a lower bound on $\tilde{g}(n)$. Using $\mathrm{H}_q(x) \le \mathrm{H}_q(1/2)$ in Lemma 2.3, we have

$$\frac{1-(q-1)\cdot \mathrm{H}_q(p_0)}{2} \ge \frac{1}{2} - \frac{q-1}{2}\cdot \mathrm{H}_q\Big(\frac{1}{2}\Big) = 2^{-q}. \tag{5.5}$$

Plugging the Tsallis binary entropy upper bound (Lemma 4.9) and Equation (5.5) into $\tilde{g}(n)$, we complete the proof by concluding the following:

$$\tilde{g}_q(n) \ge 2^{-q}\cdot \mathrm{H}_q(1/2)\cdot\left(1-\beta(n)^q - \sqrt{1-\alpha^2(n)}\right) \coloneqq g_q(n+1) = g_q(n'). \qquad \square$$

5.2 Mixed-state reductions

In this subsection, we present two reductions for mixed states. The first reduction is from the trace distance between two n-qubit states (QSD), to the quantum q-Tsallis entropy difference between two new $(n+1)$-qubit states ($\textsc{TsallisQED}_q$), for $1 \le q \le 2$, under appropriate assumptions about $\mathrm{S}_q(\rho_0)$ and $\mathrm{S}_q(\rho_1)$, as stated in Lemma 5.5. The second reduction is from the trace distance between an n-qubit quantum state (QSCMM) and the n-qubit maximally mixed state to the quantum q-Tsallis entropy of the state ($\textsc{TsallisQEA}_q$) for $q = 1 + \frac{1}{n-1}$, as state in Lemma 5.6.

5.2.1 QSD $\le$ $\textsc{TsallisQED}_q$ for $1 \le q \le 2$

Lemma 5.5 (QSD $\le$ $\textsc{TsallisQED}_q$)**.** *Let Q_0 and Q_1 be quantum circuits acting on m qubit, defined in Definition 5.1, that prepares the purification of n-qubit mixed states ρ_0 and ρ_1, respectively. For any $b \in \{0,1\}$, there is a new quantum circuits Q_b' acting on $m+3$ qubits, requiring $O(1)$ queries to controlled-Q_0 and controlled-Q_1, as well as $O(1)$ one- and two- qubit gates, that prepares a new n'-qubit mixed state ρ_b', where $n' \coloneqq n+1$, such that: For any ρ_0 and ρ_0 satisfying $\max\{\mathrm{S}_q(\rho_0), \mathrm{S}_q(\rho_1)\} \le \gamma(n)$ with $\mathrm{S}_q(I/2) \le \gamma(n) \le \mathrm{S}_q((I/2)^{\otimes n})$, any $\varepsilon(n) \in (0, 1/2)$, and any $q \in [1,2]$, there is a $g(n) > 0$ with appropriate ranges of γ, ε, and n such that*

$$\begin{aligned}&\mathrm{T}(\rho_0,\rho_1) \ge 1-\varepsilon(n) \quad &\Rightarrow\ \mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1') \ge g_q(n') = g_q(n+1),\\ &\mathrm{T}(\rho_0,\rho_1) \le \varepsilon(n) \quad &\Rightarrow\ \mathrm{S}_q(\rho_1') - \mathrm{S}_q(\rho_0') \ge g_q(n') = g_q(n+1),\end{aligned}$$

where $g_q(n) \coloneqq \frac{1}{2}\mathrm{H}_q\big(\frac{1}{2}\big) - \gamma(n)\big(\frac{1}{2} - \frac{1}{2^q}\big) - \big(\frac{1}{2}+\frac{1}{2^q}\big)\Big(\frac{\varepsilon(n)^q}{2^q}\ln_q\big(2^n\big) + \mathrm{H}_q\big(\frac{1}{2}\big)\sqrt{\varepsilon(n)(2-\varepsilon(n))}\Big)$.

Proof. Our proof strategy is somewhat inspired by [BASTS10, Section 5.4]. We start by considering the following mixed states ρ_0' and ρ_1':

$$\begin{aligned}\rho_0' &\coloneqq (\vartheta|0\rangle\langle 0| + (1-\vartheta)|1\rangle\langle 1|)\otimes \rho_+, \text{ where } 2\mathrm{H}_q(\vartheta) = \mathrm{H}_q\left(\frac{1}{2}\right) \text{ and } \rho_+ \coloneqq \frac{\rho_0+\rho_1}{2},\\ \rho_1' &\coloneqq \frac{1}{2}|0\rangle\langle 0|\otimes\rho_0 + \frac{1}{2}|1\rangle\langle 1|\otimes\rho_1.\end{aligned}$$

These states ρ_0' and ρ_1' can be prepared by the quantum circuits Q_0' and Q_1', respectively. For instance, adapting the constructions in [Liu23, Figure 1 and Figure 2], for any $b \in \{0,1\}$, the quantum circuit Q_b' uses $O(1)$ queries to controlled-Q_0 and controlled-Q_1, as well as $O(1)$ one- and two-qubit gates.

Utilizing the pseudo-additivity of S_q (Lemma 2.9), we have:

$$\mathrm{S}_q(\rho_0') = \mathrm{H}_q(\vartheta) + (1-(q-1)\mathrm{H}_q(\vartheta))\mathrm{S}_q(\rho_+) = \frac{1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) + \left(1 - \frac{q-1}{2}\cdot\mathrm{H}_q\left(\frac{1}{2}\right)\right)\mathrm{S}_q(\rho_+). \tag{5.6}$$

Using the joint q-Tsallis entropy theorem (Lemma 2.12), we obtain:

$$\mathrm{S}_q(\rho_1') = \mathrm{H}_q\left(\frac{1}{2}\right) + \frac{1}{2^q}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)). \tag{5.7}$$

Combining Equation (5.7) and Equation (5.6), we obtain:

$$\mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1') = \left(1 - \frac{q-1}{2} \cdot \mathrm{H}_q\left(\frac{1}{2}\right)\right)\mathrm{S}_q(\rho_+) - \frac{1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) - \frac{1}{2^q}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) \tag{5.8}$$

Next, we can consider the following two cases:

- For the case where $\mathrm{T}(\rho_0, \rho_1) \geq 1 - \varepsilon$, using the lower bound on QJT_q(Lemma 4.6), we have:

$$\mathrm{S}_q(\rho_+) - \frac{1}{2}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) = \mathrm{QJT}_q(\rho_0, \rho_1) \geq \mathrm{H}_q\left(\frac{1}{2}\right) - \mathrm{H}_q\left(\frac{1 - \mathrm{T}(\rho_0, \rho_1)}{2}\right). \tag{5.9}$$

Substituting Equation (5.9) into Equation (5.8), we obtain:

$$\begin{aligned}
&\mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1') \\
&\geq \left(1 - \tfrac{q-1}{2} \cdot \mathrm{H}_q\left(\tfrac{1}{2}\right)\right)\left(\tfrac{1}{2}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) + \mathrm{H}_q\left(\tfrac{1}{2}\right) - \mathrm{H}_q\left(\tfrac{\varepsilon}{2}\right)\right) - \tfrac{1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right) - \tfrac{1}{2^q}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) \\
&\geq \left(\tfrac{1}{2} - \tfrac{1}{2^q} - \tfrac{q-1}{4}\mathrm{H}_q\left(\tfrac{1}{2}\right)\right)(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) + \left(1 - \tfrac{q-1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right)\right)\mathrm{H}_q\left(\tfrac{1}{2}\right)\left(1 - \sqrt{\varepsilon(2-\varepsilon)}\right) - \tfrac{1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right) \\
&\geq \left(\tfrac{1}{2} + \tfrac{1}{2^q}\right)\mathrm{H}_q\left(\tfrac{1}{2}\right)\left(1 - \sqrt{\varepsilon(2-\varepsilon)}\right) - \tfrac{1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right) \\
&= \tfrac{1}{2^q}\mathrm{H}_q\left(\tfrac{1}{2}\right) - \left(\tfrac{1}{2} + \tfrac{1}{2^q}\right)\mathrm{H}_q\left(\tfrac{1}{2}\right)\sqrt{\varepsilon(2-\varepsilon)} \coloneqq \tilde{g}_q^{\mathtt{Y}}(\varepsilon).
\end{aligned}$$

Here, the third line uses the Tsallis binary entropy upper bound (Lemma 4.9) and the fact that $1 - \frac{q-1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) > 0$ for $q \in [1, 2]$. The last line relies on the following facts: (a) $\mathrm{S}_q(\rho) \geq 0$ for any state ρ; (b) $2\left(\frac{1}{2} - \frac{1}{2^q} - \frac{q-1}{4}\mathrm{H}_q\left(\frac{1}{2}\right)\right) = \frac{1}{2} - \frac{1}{2^q} \geq 0$ for $q \in [1, 2]$; and (c) $1 - \frac{q-1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) = \frac{1}{2} + \frac{1}{2^q}$;

- For the case where $\mathrm{T}(\rho_0, \rho_1) \leq \varepsilon$, by Fannes' inequality for QJT_q (Lemma 2.13), we have:

$$\begin{aligned}
\mathrm{S}_q(\rho_+) &\leq \frac{|\mathrm{S}_q(\rho_+) - \mathrm{S}_q(\rho_0)|}{2} + \frac{|\mathrm{S}_q(\rho_+) - \mathrm{S}_q(\rho_1)|}{2} + \frac{\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)}{2} \\
&\leq \mathrm{T}(\rho_+, \rho_b)^q \cdot \ln_q(2^n - 1) + \mathrm{H}_q(\mathrm{T}(\rho_+, \rho_b)) + \frac{\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)}{2} \\
&\leq \left(\frac{\mathrm{T}(\rho_0, \rho_1)}{2}\right)^q \ln_q(2^n) + \mathrm{H}_q\left(\frac{\mathrm{T}(\rho_0, \rho_1)}{2}\right) + \frac{\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)}{2}
\end{aligned} \tag{5.10}$$

Here, the first line is due to the triangle inequality, and the last line is because $\ln_q(x)$ is monotonically increasing on $x > 0$ for any fixed $q > 1$.

Plugging Equation (5.10) into Equation (5.8), we can derive that:

$$\begin{aligned}
&\mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1') \\
&\leq \left(1 - \tfrac{q-1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right)\right)\left(\left(\tfrac{\varepsilon}{2}\right)^q \ln_q(2^n) + \mathrm{H}_q\left(\tfrac{\varepsilon}{2}\right) + \tfrac{1}{2}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1))\right) - \tfrac{1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right) - \tfrac{1}{2^q}(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) \\
&\leq \left(\tfrac{1}{2} - \tfrac{1}{2^q} - \tfrac{q-1}{4}\mathrm{H}_q\left(\tfrac{1}{2}\right)\right)(\mathrm{S}_q(\rho_0) + \mathrm{S}_q(\rho_1)) + \left(1 - \tfrac{q-1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right)\right)\left(\left(\tfrac{\varepsilon}{2}\right)^q \ln_q(2^n) + \mathrm{H}_q\left(\tfrac{1}{2}\right)\sqrt{\varepsilon(2-\varepsilon)}\right) - \tfrac{1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right) \\
&\leq \left(\tfrac{1}{2} - \tfrac{1}{2^q}\right) \cdot \gamma + \left(\tfrac{1}{2} + \tfrac{1}{2^q}\right)\left(\left(\tfrac{\varepsilon}{2}\right)^q \cdot \ln_q(2^n) + \mathrm{H}_q\left(\tfrac{1}{2}\right)\sqrt{\varepsilon(2-\varepsilon)}\right) - \tfrac{1}{2}\mathrm{H}_q\left(\tfrac{1}{2}\right) \coloneqq -\tilde{g}_q^{\mathtt{N}}(\varepsilon, n, \gamma).
\end{aligned}$$

Here, the third line uses the Tsallis binary entropy upper bound (Lemma 4.9) and the fact that $1 - \frac{q-1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) > 0$ for $q \in [1, 2]$. The last line relies on the following facts: (a) $1 - \frac{q-1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) = \frac{1}{2} + \frac{1}{2^q}$; (b) $2\left(\frac{1}{2} - \frac{1}{2^q} - \frac{q-1}{4}\mathrm{H}_q\left(\frac{1}{2}\right)\right) = \frac{1}{2} - \frac{1}{2^q} \geq 0$ for $q \in [1, 2]$; and (c) $\mathrm{S}_q(\rho) \leq \gamma \leq \mathrm{S}_q((I/2)^{\otimes n})$ for any n-qubit state ρ.

It is evident that $\tilde{g}_q^{\mathtt{N}}(\varepsilon, n, \gamma)$ is monotonically decreasing on $\gamma \geq 0$ for any fixed q, ε, and n. Consequently, it remains to show that $\tilde{g}_q^{\mathtt{Y}}(\varepsilon) \geq \tilde{g}_q^{\mathtt{N}}(\varepsilon, n, \mathrm{H}_q(1/2)) \geq \tilde{g}_q^{\mathtt{N}}(\varepsilon, n, \gamma)$ for $\mathrm{H}_q(1/2) =$

$\mathrm{S}_q(I/2) \le \gamma \le \mathrm{S}_q((I/2)^{\otimes n})$. In particular, by noting that $(\varepsilon/2)^q \cdot \ln_q\left(2^n\right) \ge 0$ for $q \ge 1$ and $\varepsilon \ge 0$, we obtain:

$$\begin{aligned}\tilde{g}_q^{\mathtt{Y}}(\varepsilon) - \tilde{g}_q^{\mathtt{N}}\left(\varepsilon, n, \mathrm{H}_q\left(\frac{1}{2}\right)\right) &= \frac{1}{2^q}\mathrm{H}_q\left(\frac{1}{2}\right) + \left(\frac{1}{2}+\frac{1}{2^q}\right)\left(\frac{\varepsilon}{2}\right)^q \ln_q(2^n) - \frac{1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) + \left(\frac{1}{2}-\frac{1}{2^q}\right)\mathrm{H}_q\left(\frac{1}{2}\right) \\ &= \left(\frac{1}{2}+\frac{1}{2^q}\right)\left(\frac{\varepsilon}{2}\right)^q \ln_q(2^n) \\ &\ge 0.\end{aligned}$$

Therefore, we complete the proof by choosing $g_q(n) = \tilde{g}_q^{\mathtt{N}}(\varepsilon(n), n, \gamma(n))$, specifically:

$$g_q(n) \coloneqq \frac{1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) - \gamma(n)\left(\frac{1}{2}-\frac{1}{2^q}\right) - \left(\frac{1}{2}+\frac{1}{2^q}\right)\left(\frac{\varepsilon(n)^q}{2^q}\ln_q(2^n) + \mathrm{H}_q\left(\frac{1}{2}\right)\sqrt{\varepsilon(n)(2-\varepsilon(n))}\right). \quad \square$$

5.2.2 QSCMM $\le$ TsallisQEA$_q$ for $q(n) = 1 + \frac{1}{n-1}$

Lemma 5.6 (QSCMM $\le$ TsallisQEA$_q$)**.** *Let Q be a quantum circuit acting on m qubit, defined in Definition 5.2, that prepares the purification of n-qubit mixed states ρ, respectively. For any ρ, any $n \ge 5$, and any $q(n) = 1 + 1/(n-1)$, let $t(n) \coloneqq \frac{1}{4}\left(3n - n^{1+\frac{1}{n}} - 1\right)$, we have:*

$$\begin{aligned}&\mathrm{T}\left(\rho, (I/2)^{\otimes n}\right) \le 1/n && \Rightarrow\ \mathrm{S}_q(\rho) > t(n) + 1/150, \\ &\mathrm{T}\left(\rho, (I/2)^{\otimes n}\right) \ge 1 - 1/n && \Rightarrow\ \mathrm{S}_q(\rho) < t(n) - 1/150.\end{aligned}$$

Proof. Let $\rho = \sum_{i\in[2^n]} \lambda_i |v_i\rangle\langle v_i|$ be the spectral decomposition of ρ, where $\{v_i\}_{i\in[2^n]}$ is an orthonormal basis and $p \coloneqq (\lambda_1, \cdots, \lambda_{2^n})$ is a probability distribution of dimension 2^n. And let ν be the uniform distribution of dimension 2^n. Noting that ρ and $(I/2)^{\otimes n}$ commute, we have $\mathrm{T}(\rho, (I/2)^{\otimes n}) = \mathrm{TV}(p, \nu)$ and $\mathrm{S}_q(\rho) = \mathrm{H}_q(p)$.

Let $t(n) \coloneqq \frac{1}{4}\left(3n - n^{1+\frac{1}{n}} - 1\right)$. Next, we can consider the following two cases:

- For the case where $\mathrm{T}(\rho, (I/2)^{\otimes n}) \le 1/n$, by the lower bound on $\mathrm{H}_q(p)$ in Lemma 4.10, it follows that

$$\begin{aligned}\mathrm{S}_q(\rho) &\ge \ln_{1+\frac{1}{n-1}}\left(2^n\right) \cdot \left(1 - \mathrm{T}\left(\rho, (I/2)^{\otimes n}\right) - 2^{-n}\right) \\ &\ge (n-1)\left(1 - \frac{1}{2}\cdot\left(\frac{1}{2}\right)^{\frac{1}{n-1}}\right)\left(1 - \frac{1}{n} - 2^{-n}\right) \coloneqq \tau_{\mathtt{Y}}(n).\end{aligned}$$

By a direct calculation, we obtain:

$$\begin{aligned}&\mathrm{S}_q(\rho) - t(n) \ge \tau_{\mathtt{Y}}(n) - t(n) = g_1(n) + g_2(n) + g_3(n) - \frac{7}{4}, \\ &\text{where } g_1(n) \coloneqq 2^{-n} + \frac{1 - 2^{\frac{n}{1-n}}}{n} + 2^{\frac{n^2}{1-n}}(n-1) + \frac{n}{4}\left(1 - 2^{\frac{1}{1-n}}\right), \\ &\qquad g_2(n) \coloneqq 2^{\frac{1}{1-n}} - 2^{-n}n, \quad g_3(n) \coloneqq \frac{n}{4}\left(n^{\frac{1}{n}} - 2^{\frac{1}{1-n}}\right).\end{aligned} \tag{5.11}$$

Through a fairly tedious calculation, we know that $g_1(n)$, $g_2(n)$, and $g_3(n)$ defined in Equation (5.11) satisfy the properties in Fact 5.6.1, and the proof is deferred in Appendix A.2.

Fact 5.6.1. *Let $g_1(n)$, $g_2(n)$, and $g_3(n)$ be functions defined in Equation (5.11). It holds that:*

(1) *For $n \ge 3$, $g_1(n) \ge 0$.*

(2) *For $n \ge 3$, $g_2(n)$ and $g_3(n)$ are monotonically increasing.*

Combining Equation (5.11) and Fact 5.6.1, we obtain that:

$$\forall n \ge 5,\ \mathrm{S}_q(\rho) - t(n) \ge \tau_{\mathtt{Y}}(n) - t(n) \ge g_2(n) + g_3(n) - \frac{7}{4} > \frac{1}{150}. \tag{5.12}$$

- For the case where $\mathrm{T}(\rho,(I/2)^{\otimes n}) \geq 1-1/n$, by noting $\mathrm{T}(\rho,(I/2)^{\otimes n})q \geq \left(1-\frac{1}{n}\right)\left(1+\frac{1}{n-1}\right) = 1$ and using the upper bound on $\mathrm{H}_q(p)$ in Lemma 4.10, it holds that

$$\begin{aligned}\mathrm{S}_q(\rho) &\leq \ln_{1+\frac{1}{n-1}}\left(2^n\left(1-\mathrm{T}\left(\rho,(I/2)^{\otimes n}\right)\right)\right)\\ &\leq \ln_{1+\frac{1}{n}}\left(2^n\left(1-\mathrm{T}\left(\rho,(I/2)^{\otimes n}\right)\right)\right)\\ &\leq n\left(1-\frac{1}{2}\cdot n^{1/n}\right) \coloneqq \tau_{\mathtt{N}}(n).\end{aligned}$$

  Here, the second line is because $\ln_q(x) < \ln_{q'}(x)$ for $q > q' > 0$ and $\frac{1}{n-1} > \frac{1}{n}$.

  Similarly, a direct calculation implies that:

$$t(n) - \mathrm{S}_q(\rho) \geq t(n) - \tau_{\mathtt{N}}(n) = \frac{g_4(n)-1}{4}, \text{ where } g_4(n) \coloneqq n\left(n^{\frac{1}{n}}-1\right). \tag{5.13}$$

  Next, we will prove that $g_4(n)$ is monotonically non-decreasing for $n \geq 2$. We proceed by expressing the first and second derivative of $g_4(n)$:

$$\frac{\mathrm{d}}{\mathrm{d}n}g_4(n) = \frac{n^{\frac{1}{n}}}{n}(n-\log(n)+1)-1, \text{ and } \frac{\mathrm{d}^2}{\mathrm{d}n^2}g_4(n) = \frac{n^{\frac{1}{n}}}{n^3}\left((\log(n)-1)^2-n\right).$$

  Since $\sqrt{n} > \log n$, we know that $\frac{\mathrm{d}^2}{\mathrm{d}n^2}g_4(n)$ has one zero at $n=1$. As $\frac{\mathrm{d}^2}{\mathrm{d}n^2}g_4(n)\big|_{n=e} = -e < 0$, we have that $\frac{\mathrm{d}}{\mathrm{d}n}g_4(n)$ is monotonically decreasing for $n \geq 2$, and thus, $\frac{\mathrm{d}}{\mathrm{d}n}g_4(n) \geq \lim_{n\to\infty}\frac{\mathrm{d}}{\mathrm{d}n}g_4(n) = 0$ for $n \geq 2$. Hence, we conclude that $g_4(n)$ is monotonically non-decreasing for $n \geq 2$. Consequently, combining with Equation (5.13), we obtain:

$$\forall n \geq 3,\ t(n) - \mathrm{S}_q(\rho) \geq t(n) - \tau_{\mathtt{N}}(n) = \frac{g_4(n)-1}{4} > \frac{1}{13}. \tag{5.14}$$

  Lastly, we finish the proof by comparing Equation (5.12) with Equation (5.14). □

5.3 Computational hardness results

In this subsection, we present the computational hardness results for various settings of TsallisQED$_q$ and TsallisQEA$_q$ by using our reductions established in Section 5.1 and Section 5.2.

5.3.1 BQP hardness results

Theorem 5.7 (ConstRankTsallisQED$_q$ is BQP-hard for $1 \leq q \leq 2$)**.** *For any $q \in [1,2]$ and any $n \geq 3$, the following holds*:

$$\forall g_q(n) \in \left[\frac{1}{\mathrm{poly}(n)}, 2^{-q}\mathrm{H}_q\left(\frac{1}{2}\right)\left(1-2^{-\frac{qn}{2}+1}\right)\right],\ \textsc{ConstRankTsallisQED}_q[g_q(n)] \textit{ is } \mathsf{BQP}\textit{-hard.}$$

Proof. Using Lemma 2.17, we have that PureQSD$\left[\sqrt{1-2^{-2\hat{n}}}, 2^{-(\hat{n}+1)/2}\right]$ is BQP-hard for $\hat{n} \geq 2$. Let Q_0 and Q_1 be the corresponding BQP-hard instance such that these circuits are polynomial-size and prepare the pure states $|\psi_0\rangle\langle\psi_0|$ and $|\psi_1\rangle\langle\psi_1|$, respectively. Leveraging the reduction from PureQSD to ConstRankTsallisQED$_q$ (Lemma 5.4), there are two polynomial-size quantum circuits Q_0' and Q_1', which prepares the purifications of constant-rank states ρ_0' and ρ_1', such that: For any $1 \leq q \leq 2$ and any $n = \hat{n}+1 \geq 3$,

$$\begin{aligned}&\mathrm{T}(|\psi_0\rangle\langle\psi_0|,|\psi_1\rangle\langle\psi_1|) \geq \sqrt{1-2^{-2\hat{n}}} &&\Rightarrow\ \mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1') \geq g_q(n) = g_q(\hat{n}+1),\\ &\mathrm{T}(|\psi_0\rangle\langle\psi_0|,|\psi_1\rangle\langle\psi_1|) \leq 2^{-(\hat{n}+1)/2} &&\Rightarrow\ \mathrm{S}_q(\rho_1') - \mathrm{S}_q(\rho_0') \leq g_q(n) = g_q(\hat{n}+1).\end{aligned}$$

Hence, we complete the proof by a direct calculation:

$$g_q(n) \coloneqq 2^{-q}\cdot\mathrm{H}_q(1/2)\cdot\left(1-2^{-\frac{qn}{2}}-\sqrt{1-\left(1-2^{-2(n-1)}\right)}\right) \geq 2^{-q}\cdot\mathrm{H}_q(1/2)\cdot\left(1-2^{-\frac{qn}{2}+1}\right).\ \square$$

Theorem 5.8 ($\textsc{ConstRankTsallisQEA}_q$ is BQP-hard under Turing reduction for $1 \le q \le 2$)**.** *For any $q \in [1,2]$ and any $n \ge 3$, the following holds*:

$$\textsc{ConstRankTsallisQEA}_q \text{ with } g(n) = \Theta(1) \text{ is } \mathsf{BQP}\text{-hard under Turing reduction.}$$

Proof. For any $1 \le q \le 2$ and $n \ge 3$, since $\textsc{ConstRankTsallisQED}_q[\hat{g}_q(n)]$ is BQP-hard under Karp reduction (Theorem 5.7), where $\hat{g}_q(n) \coloneqq 2^{-q}\mathrm{H}_q(1/2)\big(1-2^{-n/2+1}\big)$, it suffices to provide an algorithm for $\textsc{ConstRankTsallisQED}_q[\hat{g}_q(n)]$ by using $\textsc{ConstRankTsallisQEA}_q[t(n), g(n)]$ as subroutines, with appropriately adaptive choices of $t(n)$ and $g(n)$.

Let Q_0 and Q_1 be the corresponding BQP-hard instance such that these circuits are polynomial-size and prepare the constant-rank states ρ_0 and ρ_1, respectively. Let $\texttt{TsallisQEA}_q(Q, t(n), g(n))$ be the subroutine for decide whether $\mathrm{S}_q(\rho) \ge t(n) + g(n)$ or $\mathrm{S}_q(\rho) \le t(n) - g(n)$. Next, we estimate $\mathrm{S}_q(\rho_b)$ to within additive error $\hat{g}_q(n)/2$ for $b \in \{0,1\}$. This procedure, inspired by [Amb14, Appendix A.2 Part 1], is denoted by `BiSearch`, as presented in Algorithm 3.

Algorithm 3 Tsallis entropy estimation $\texttt{BiSearch}(Q, \tau, g)$ via queries to $\texttt{TsallisQEA}_q$.

Input: A quantum circuit Q that prepares the purification of ρ, an upper bound τ on the q-Tsallis entropy $\mathrm{S}_q(\rho)$, and a precision parameter g.
Output: Return t such that $|t - \mathrm{S}_q(\rho)| \le g/2$.

```
Let δ ← g/2, and set the interval [a, b] ← [0, τ].
while b − a > ĝ/2 do
    Query TsallisQEA_q(Q, (a+b)/2, δ/4) to decide whether S_q(ρ) ≥ (a+b)/2 + δ/4 or S_q(ρ) ≤ (a+b)/2 − δ/4.
    if S_q(ρ) ≥ (a+b)/2 + δ/4 then
        [a, b] ← [(a+b)/2 − ĝ/4, b].
    else
        [a, b] ← [a, (a+b)/2 + ĝ/4].
    end if
end while
return (a+b)/2.
```

To solve $\textsc{ConstRankTsallisQED}_q[\hat{g}_q(n)]$, noting that $\max\{\mathrm{rank}(\rho_0), \mathrm{rank}(\rho_1)\} \le r \le O(1)$, we choose $\tau(n) = \mathrm{S}_q((I/2)^{\otimes r})$. Subsequently, let $t_0(n) = \texttt{BiSearch}(Q_0, \tau(n), \hat{g}_q(n))$ and $t_1(n) = \texttt{BiSearch}(Q_1, \tau(n), \hat{g}_q(n))$, we obtain:

$$\begin{aligned}
&\mathrm{S}_q(\rho_0) - \mathrm{S}_q(\rho_1) \ge \hat{g}_q(n) \quad \Rightarrow \quad t_0(n) - t_1(n) \ge \mathrm{S}_q(\rho_0) - \frac{\hat{g}_q(n)}{2} - \left(\mathrm{S}_q(\rho_1) + \frac{\hat{g}_q(n)}{2}\right) \ge 0, \\
&\mathrm{S}_q(\rho_0) - \mathrm{S}_q(\rho_1) \le -\hat{g}_q(n) \quad \Rightarrow \quad t_0(n) - t_1(n) \le \mathrm{S}_q(\rho_0) + \frac{\hat{g}_q(n)}{2} - \left(\mathrm{S}_q(\rho_1) - \frac{\hat{g}_q(n)}{2}\right) \le 0.
\end{aligned} \tag{5.15}$$

Note that $\hat{g}_q(n) = 2^{-q}\mathrm{H}_q(1/2)(1-2^{-n/2+1}) \ge \frac{2-\sqrt{2}}{2^{q+1}}\mathrm{H}_q(1/2)$ for $n \ge 3$ and $\tau(n) \le \mathrm{S}((I/2)^r) \le O(1)$. Since each query to $\texttt{TsallisQEA}_q$ in `BiSearch` decreases the size of the interval $[a,b]$ by almost a half, we can conclude that the number of adaptive queries to $\texttt{TsallisQEA}_q$ in $\texttt{BiSearch}(Q_0, \tau(n), \hat{g}_q(n))$ and $\texttt{BiSearch}(Q_1, \tau(n), \hat{g}_q(n))$ is $O(\log(1/\hat{g}_q(n))) = O(1)$. □

5.3.2 QSZK hardness results

Theorem 5.9 ($\textsc{TsallisQED}_q$ is QSZK-hard for $1 < q \le 1 + \frac{1}{n-1}$)**.** *For any $q \in \left(1, 1+\frac{1}{n-1}\right]$ and any $n \ge 90$, it holds that*

$$\forall g(n) \in [1/\operatorname{poly}(n), 1/400],\ \textsc{TsallisQED}_q[g(n)] \text{ is } \mathsf{QSZK}\text{-hard.}$$

Proof. Following Lemma 2.16, we have that $\mathrm{QSD}\left[1-2^{-\hat{n}^{0.49}}, 2^{-\hat{n}^{0.49}}\right]$ is QSZK-hard for $\hat{n} \ge 1$. Let Q_0 and Q_1 be the corresponding QSZK-hard instance such that these circuits are polynomial-

size and prepare the purification of ρ_0 and ρ_1, respectively. Leveraging the reduction from QSD to $\textsc{TsallisQED}_q$ (Lemma 5.5), there are two polynomial-size quantum circuits Q_0' and Q_1', which prepare the purifications of n-qubit ρ_0' and ρ_1' where $n \coloneqq \hat{n}+1$, respectively, such that:

$$\begin{aligned} &\mathrm{T}(\rho_0,\rho_1) \geq 1-2^{-\hat{n}^{0.49}} &&\Rightarrow\ \mathrm{S}_q(\rho_0')-\mathrm{S}_q(\rho_1') \geq g_q(n) = g_q(\hat{n}+1),\\ &\mathrm{T}(\rho_0,\rho_1) \leq 2^{-\hat{n}^{0.49}} &&\Rightarrow\ \mathrm{S}_q(\rho_1')-\mathrm{S}_q(\rho_0') \leq g_q(n) = g_q(\hat{n}+1). \end{aligned}$$

Since $\sqrt{2^{-\hat{n}^{0.49}}\left(2-2^{-\hat{n}^{0.49}}\right)} \leq 2^{\frac{1-\hat{n}^{0.49}}{2}}$ and $\gamma(n) \leq \mathrm{S}_q\left((I/2)^{\otimes\hat{n}}\right) = \frac{1-2^{\hat{n}(1-q)}}{q-1}$, we have

$$g_q(\hat{n}) \geq \underbrace{\frac{1}{2}\mathrm{H}_q\left(\frac{1}{2}\right) - \frac{1-2^{\hat{n}(1-q)}}{q-1}\left(\frac{1}{2}-\frac{1}{2^q}\right)}_{G_1(q;\hat{n})} - \underbrace{\left(\frac{1}{2}+\frac{1}{2^q}\right)\frac{2^{-\hat{n}^{0.49}q}}{2^q}\ln_q\left(2^{\hat{n}}\right)}_{G_2(q;\hat{n})} - \underbrace{\left(\frac{1}{2}+\frac{1}{2^q}\right)\mathrm{H}_q\left(\frac{1}{2}\right)2^{\frac{1-\hat{n}^{0.49}}{2}}}_{G_3(q;\hat{n})}.$$

It remains to show that $g_q(\hat{n}) \geq G_1(q;\hat{n}) - G_2(q;\hat{n}) - G_3(q;\hat{n}) > 0$ for $1 \leq q \leq 1+\frac{1}{\hat{n}}$ and large enough n. By the Taylor expansion of $G_1(q;\hat{n})$, $G_2(q;\hat{n})$, and $G_3(q;\hat{n})$ at $q=1$, we obtain:

$$\begin{aligned} g_q(\hat{n}) &\geq G_1(q;\hat{n}) - G_2(q;\hat{n}) - G_3(q;\hat{n})\\ &\geq \left(\frac{\log(2)}{2} - \frac{1}{4}(2\hat{n}+1)\log^2(2)(q-1)\right) - \frac{\log(2)}{2}\cdot\hat{n}2^{-\hat{n}^{0.49}} - \log(2)\cdot 2^{\frac{1-\hat{n}^{0.49}}{2}} \coloneqq G(q;\hat{n}) \end{aligned}$$

Noting that $\frac{\partial}{\partial q}G(q;\hat{n}) = -\frac{1}{4}(2\hat{n}+1)\log^2(2) < 0$ for $\hat{n} \geq 1$, we know that $G(q;\hat{n})$ is monotonically decreasing on $q>1$ for any fixed $\hat{n}\geq 1$. As a consequence, as $1 \leq q \leq 1+\frac{1}{\hat{n}}$, it is left to show that $G\left(1+\frac{1}{\hat{n}};\hat{n}\right) > 0$ for large enough $\hat{n}$, specifically:

$$G\left(1+\frac{1}{\hat{n}};\hat{n}\right) = \frac{\log(2)}{4}\left(2 - 2\log(2) - 2^{1-\hat{n}^{0.49}}\hat{n} - 4\cdot 2^{\frac{1-\hat{n}^{0.49}}{2}} - \frac{\log(2)}{\hat{n}}\right) > 0.$$

A direct calculation implies that

$$\frac{\mathrm{d}}{\mathrm{d}\hat{n}}G\left(1+\frac{1}{\hat{n}};\hat{n}\right) = \frac{\log(2)}{200}\left(49\sqrt{2^{1-\hat{n}^{0.49}}}\hat{n}^{1.49}\log(2) + 2^{-\hat{n}^{0.49}}\hat{n}^2\left(49\hat{n}^{0.49}\log(2) - 100\right) + 50\log(2)\right).$$

Since it is evident that $49\hat{n}^{0.49}\log(2) - 100 > 0$, we can deduce that $\frac{\mathrm{d}}{\mathrm{d}\hat{n}}G\left(1+\frac{1}{\hat{n}};\hat{n}\right) > 0$. As $49\hat{n}^{0.49}\log(2) - 100 > 0$ holds when $\hat{n} \geq 10$, we obtain that $G\left(1+\frac{1}{\hat{n}};\hat{n}\right)$ is monotonically increasing for $\hat{n}\geq 10$. Therefore, we complete the proof by noticing $\hat{n} = n-1$ and the following:

$$\text{For any } q \in \left(1, 1+\frac{1}{\hat{n}}\right] \text{ and } \hat{n} \geq 89,\ g_q(\hat{n}) \geq G(q;\hat{n}) \geq G\left(1+\frac{1}{\hat{n}};\hat{n}\right) \geq G\left(1+\frac{1}{89};89\right) > \frac{1}{400}.\ \square$$

Theorem 5.10 ($\textsc{TsallisQEA}_q$ is QSZK-hard under Turing reduction for $1 < q \leq 1+\frac{1}{n-1}$). *For any $q \in \left(1, 1+\frac{1}{n-1}\right]$ and any $n \geq 90$, the following holds:*

$$\textsc{TsallisQEA}_q \text{ with } g(n) = \Theta(1) \text{ is } \mathsf{QSZK}\text{-hard under Turing reduction.}$$

Proof. This proof is very similar to the proof of Theorem 5.8. For any $1 < q \leq 1+\frac{1}{n-1}$ and $n \geq 90$, since $\textsc{TsallisQED}_q[\hat{g}_q(n)]$ is QSZK-hard under Karp reduction (Theorem 5.9), where $\hat{g}_q(n) = 1/400$, it suffices to provide an algorithm for $\textsc{TsallisQED}_q[\hat{g}_q(n)]$ by using $\textsc{TsallisQEA}_q[t(n),g(n)]$ as subroutines, with appropriately adaptive choices of $t(n)$ and $g(n)$.

Let Q_0 and Q_1 be the corresponding QSZK-hard instance such that these circuits are polynomial-size and prepare the states ρ_0 and ρ_1, respectively. Let $\texttt{TsallisQEA}_q(Q,t(n),g(n))$ be the subroutine for decide whether $\mathrm{S}_q(\rho) \geq t(n)+g(n)$ or $\mathrm{S}_q(\rho) \leq t(n)-g(n)$. Next, we estimate $\mathrm{S}_q(\rho_b)$ to within additive error $\hat{g}_q(n)/2$ for $b \in \{0,1\}$ via the procedure `BiSearch`, as specified in Algorithm 3. To solve $\textsc{TsallisQED}_q[\hat{g}_q(n)]$, noting that $\max\{\mathrm{rank}(\rho_0),\mathrm{rank}(\rho_1)\} \leq 2^n$, we choose $\tau(n) = \mathrm{S}_q((I/2)^{\otimes n})$. Subsequently, let $t_0(n) = \texttt{BiSearch}(Q_0,\tau(n),\hat{g}_q(n))$ and $t_1(n) = \texttt{BiSearch}(Q_1,\tau(n),\hat{g}_q(n))$, we obtain the same inequalities in Equation (5.15).

Note that $\hat{g}_q(n) = 1/400$ for $n \geq 90$ and $\tau(n) \leq \mathrm{S}((I/2)^n) < 1/(q-1) \leq O(1)$. Since each query to $\texttt{TsallisQEA}_q$ in `BiSearch` decreases the size of the interval $[a,b]$ by almost a half,

we complete the proof by concluding that the number of adaptive queries to $\texttt{TsallisQEA}_q$ in $\texttt{BiSearch}(Q_0, \tau(n), \hat{g}_q(n))$ and $\texttt{BiSearch}(Q_1, \tau(n), \hat{g}_q(n))$ is $O(\log(1/g_q(n))) = O(1)$. □

5.3.3 NIQSZK hardness result

Theorem 5.11 ($\textsc{TsallisQEA}_q$ is NIQSZK-hard for $q = 1+\frac{1}{n-1}$)**.** *For any $n \geq 5$, it holds that:*

$$\forall g(n) \in [1/\operatorname{poly}(n), 1/150],\ \textsc{TsallisQEA}_{1+\frac{1}{n-1}} \text{ with } g(n) \text{ is NIQSZK-hard.}$$

Proof. Utilizing Lemma 2.18, we know that QSCMM$[1/n, 1-1/n]$ is NIQSZK-hard for $n \geq 3$. Following the reduction from QSCMM to $\textsc{TsallisQEA}_{1+\frac{1}{n-1}}$ for $n \geq 5$ (Lemma 5.6), and the specific choice of $t(n)$ in the reduction, we can conclude that $g(n) \geq 1/150$. □

5.4 Quantum query complexity lower bounds

In this subsection, we present two quantum query complexity lower bounds for estimating the quantum Tsallis entropy $\mathrm{S}_q(\rho)$: When q is constantly larger than 1, the lower bound is *independent* of the rank of ρ (Theorem 5.12). However, when $q > 1$ is inverse-polynomially close to 1 or even closer, the lower bound *depends polynomially* on the rank of ρ (Theorem 5.13).

Theorem 5.12 (Query complexity lower bound for estimating quantum Tsallis entropy with q constantly above 1)**.** *For any $q \geq 1 + \Omega(1)$ and sufficiently small $\epsilon > 0$, the quantum query complexity for estimating the q-Tsallis entropy of a quantum state to within additive error ϵ, in the purified quantum query access model, is $\Omega(1/\sqrt{\epsilon})$.*

Proof. Consider the task of distinguishing two quantun unitary operators U_ϵ and U_0 corresponding to two probability distributions p_ϵ and p_0, where $p_x \coloneqq (1-x, x)$, U_x is a unitary operator satisfying

$$U_x|0\rangle = \sqrt{1-x}|0\rangle|\varphi_0\rangle + \sqrt{x}|1\rangle|\varphi_1\rangle,$$

with $|\varphi_0\rangle$ and $|\varphi_1\rangle$ being any orthogonal unit vectors. By the quantum query complexity of distinguishing probability distributions given in Lemma 2.20, we know that distinguishing U_ϵ and U_0 requires quantum query complexity $\Omega(1/d_{\mathrm{H}}(p_\epsilon, p_0))$, where $d_{\mathrm{H}}(\cdot,\cdot)$ is the Hellinger distance between two probability distributions. Direct calculation shows that if $\epsilon \in (0, 1)$,

$$d_{\mathrm{H}}(p_\epsilon, p_0) = \frac{1}{\sqrt{2}}\sqrt{\left(\sqrt{1-\epsilon}-1\right)^2 + \left(\sqrt{\epsilon}-0\right)^2} \leq \sqrt{\epsilon}.$$

Thus the query complexity of distinguishing U_0 and U_ϵ is $\Omega(1/\sqrt{\epsilon})$.

On the other hand, U_x prepares a purification of $\rho_x \coloneqq (1-x)|0\rangle\langle 0| + x|1\rangle\langle 1|$. Then, for sufficiently small $\epsilon > 0$, we have

$$|\mathrm{S}_q(\rho_\epsilon) - \mathrm{S}_q(\rho_0)| = \frac{1-(1-\epsilon)^q - \epsilon^q}{q-1} = \Omega(\epsilon).$$

Therefore, any quantum query algorithm that can compute the q-Tsallis entropy of a quantum state to within additive error $\Theta(\epsilon)$ can be used to distinguish U_ϵ and U_0, thus requiring query complexity $\Omega(1/\sqrt{\epsilon})$. □

Theorem 5.13 (Query complexity lower bound for estimating quantum Tsallis entropy with $q > 1$ near 1)**.** *For any $q \in \left(1, 1+\frac{1}{n-1}\right]$, there exists a mixed quantum state ρ of sufficiently large rank r such that the quantum query complexity for estimating $\mathrm{S}_q(\rho)$, in the purified quantum query access model, is $\Omega(r^{0.17-c})$ for any constant $c > 0$.*

Remark 5.14 (τ-dependence in the lower bounds). The lower bounds on query and sample complexities in Theorems 5.13 and 5.16 are $\Omega\left(r^{\frac{1-\tau}{3}-c}\right)$ and $\Omega(r^{1-\tau-c'})$, respectively, where $\tau = 0.49$ is chosen to establish the QSZK hardness (Theorem 5.9) and $c' = 3c$. Notably, these

bounds can be further improved by selecting a smaller τ that still satisfies all requirements in the reduction (Lemma 5.5), which is left for future work.

Proof of Theorem 5.13. By Lemma 2.19 with $\epsilon = 1/2$, there exists an $\hat{n}$-qubit state $\hat{\rho}$ of rank $\hat{r} \geq 2$ and the corresponding "uniform" state $\hat{\rho}_{\mathsf{U}}$ of rank r on the same support as $\hat{\rho}$ such that the quantum sample complexity to decide whether $\mathrm{T}\big(\hat{\rho}, \hat{\rho}_{\mathsf{U}}\big)$ is at least $1/2$ or exactly 0 is $\Omega(\hat{r}^{1/3})$. We apply the polarization lemma for the trace distance to the states $\hat{\rho}$ and $\hat{\rho}_{\mathsf{U}}$, particularly using only the direct product lemma [Wat02, Lemma 8]. Let $\rho \coloneqq \hat{\rho}^{\otimes \hat{r}^k}$ and $\rho_{\mathsf{U}} \coloneqq \hat{\rho}_{\mathsf{U}}^{\otimes \hat{r}^k}$ be the resulting states, where k is a parameter to be determined later. Then, for any constant $k > \frac{\tau}{1-\tau}$ with $\tau = 0.49$ and for sufficiently large $\hat{r}$, the following holds:

$$\begin{aligned} \mathrm{T}(\hat{\rho}, \hat{\rho}_{\mathsf{U}}) \geq 1/2 \quad &\Rightarrow \quad \mathrm{T}(\rho, \rho_{\mathsf{U}}) \geq 1 - \exp(-\hat{r}^k/8) \geq 1 - 2^{-r^\tau}, \\ \mathrm{T}(\hat{\rho}, \hat{\rho}_{\mathsf{U}}) = 0 \quad &\Rightarrow \quad \mathrm{T}(\rho, \rho_{\mathsf{U}}) \leq \hat{r}^k \cdot 0 = 0 \leq 2^{-r^\tau}. \end{aligned}$$

Hence, the sample complexity of deciding whether $\mathrm{T}(\rho, \rho_{\mathsf{U}})$ is at least $1 - 2^{-r^\tau}$ or at most 2^{-r^τ} is $\Omega\big(r^{\frac{1}{3(1+k)}}\big)$, where $r \coloneqq \hat{r} \cdot \hat{r}^k = \hat{r}^{1+k}$. For any $q \in \big(1, 1 + \frac{1}{n-1}\big] \subseteq \big(1, 1 + \frac{1}{r-1}\big]$, using the reduction from QSD to $\textsc{TsallisQED}_q$ (Lemma 5.5) with parameters from Theorem 5.9, there are two corresponding states ρ_0' and ρ_1' of rank at most $2r$ such that the quantum query complexity for deciding whether $\mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1')$ is at least $1/400$ or at most $-1/400$ is $\Omega\big(r^{\frac{1}{3(1+k)}}\big) = \Omega\big(r^{\frac{1-\tau}{3} - c}\big) = \Omega(r^{0.17-c})$ for any constant $c > 0$. Thus, estimating $\mathrm{S}_q(\rho_b')$ for $b \in \{0, 1\}$ to within additive error $1/800$ requires at least the same number of quantum queries. □

5.5 Quantum sample complexity lower bounds

In this subsection, we present two quantum sample complexity lower bounds for estimating the quantum Tsallis entropy $\mathrm{S}_q(\rho)$: When q is constantly larger than 1, the lower bound is *independent* of the rank of ρ (Theorem 5.15). However, when $q > 1$ is inverse-polynomially close to 1, the lower bound *depends polynomially* on the rank of ρ (Theorem 5.16).

Theorem 5.15 (Sample complexity lower bound for estimating quantum Tsallis entropy with q constantly above 1)**.** *For any $q \geq 1 + \Omega(1)$ and sufficiently small $\epsilon > 0$, the quantum sample complexity for estimating the quantum q-Tsallis entropy of a quantum state to within additive error ϵ is $\Omega(1/\epsilon)$.*

Proof. Consider the hypothesis testing problem where the given quantum state ρ is promised to be either ρ_0 or ρ_ϵ, each with equal probability. Specifically, the states are defined as

$$\forall x \in [0, 1], \quad \rho_x \coloneqq (1 - x)|0\rangle\langle 0| + x|1\rangle\langle 1|.$$

For sufficiently small $\epsilon > 0$, we know that $|\mathrm{S}_q(\rho_\epsilon) - \mathrm{S}_q(\rho_0)| = \Omega(\epsilon)$, as shown in the proof of Theorem 5.12. Now, assume that there is a quantum estimator for $\mathrm{S}_q(\rho)$ to within additive error $\Theta(\epsilon)$ with sample complexity S. This estimator can then be used to distinguish these two states ρ_0 and ρ_ϵ with success probability $p_{\mathrm{succ}} \geq 2/3$. On the other hand, by Lemma 2.7, we have

$$p_{\mathrm{succ}} \leq \frac{1}{2} + \frac{1}{2}\mathrm{T}\Big(\rho_0^{\otimes \mathsf{S}}, \rho_\epsilon^{\otimes \mathsf{S}}\Big).$$

By applying the Fuchs–van de Graaf inequalities [FvdG99, Theorem 1], we have

$$\mathrm{T}\Big(\rho_0^{\otimes \mathsf{S}}, \rho_\epsilon^{\otimes \mathsf{S}}\Big) \leq \sqrt{1 - \mathrm{F}\Big(\rho_0^{\otimes \mathsf{S}}, \rho_\epsilon^{\otimes \mathsf{S}}\Big)^2},$$

where $\mathrm{F}(\rho, \sigma) = \mathrm{tr}(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}})$ is the fidelity of quantum states. A direct calculation shows that $\mathrm{F}(\rho_0, \rho_\epsilon) = \sqrt{1 - \epsilon}$, which gives that

$$p_{\mathrm{succ}} \leq \frac{1}{2} + \frac{1}{2}\sqrt{1 - (1 - \epsilon)^{\mathsf{S}}}.$$

By combining this with the condition $p_{\mathrm{succ}} \geq 2/3$, we conclude that $\mathsf{S} = \Omega(1/\epsilon)$. □

Theorem 5.16 (Sample complexity lower bound for estimating quantum Tsallis entropy with $q > 1$ near 1)**.** *For any $q \in \left(1, 1+\frac{1}{n-1}\right]$, there exists a mixed quantum state ρ of sufficiently large rank r such that the quantum sample complexity for estimating $\mathrm{S}_q(\rho)$ is $\Omega(r^{0.51-c})$ for any constant $c > 0$.*

Notably, Remark 5.14 on the τ-dependence in the lower bound also applies to Theorem 5.16. Moreover, the proof strategy of Theorem 5.16 is similar to that of Theorem 5.13, as both rely on the direct product lemma for the trace distance [Wat02, Lemma 8].[31]

Proof of Theorem 5.16. By Lemma 2.21 with $\epsilon = 1/2$, there exists an $\hat{n}$-qubit state $\hat{\rho}$ of rank $\hat{r} \geq 2$ and the corresponding "uniform" state $\hat{\rho}_{\mathtt{U}}$ of rank r on the same support as $\hat{\rho}$ such that the quantum sample complexity to decide whether $\mathrm{T}\big(\hat{\rho}, \hat{\rho}_{\mathtt{U}}\big)$ is at least $1/2$ or exactly 0 is $\Omega(\hat{r})$. We apply the direct product lemma [Wat02, Lemma 8] to the states $\hat{\rho}$ and $\hat{\rho}_{\mathtt{U}}$. Let $\rho \coloneqq \hat{\rho}^{\otimes \hat{r}^k}$ and $\rho_{\mathtt{U}} \coloneqq \hat{\rho}_{\mathtt{U}}^{\otimes \hat{r}^k}$ be the resulting states, where k is a parameter to be determined later. Then, for any constant $k > \frac{\tau}{1-\tau}$ with $\tau = 0.49$ and for sufficiently large $\hat{r}$, the following holds:

$$\begin{aligned} &\mathrm{T}(\hat{\rho}, \hat{\rho}_{\mathtt{U}}) \geq 1/2 \quad &\Rightarrow \quad &\mathrm{T}(\rho, \rho_{\mathtt{U}}) \geq 1 - \exp(-\hat{r}^k/8) \geq 1 - 2^{-r^\tau}, \\ &\mathrm{T}(\hat{\rho}, \hat{\rho}_{\mathtt{U}}) = 0 \quad &\Rightarrow \quad &\mathrm{T}(\rho, \rho_{\mathtt{U}}) \leq \hat{r}^k \cdot 0 = 0 \leq 2^{-r^\tau}. \end{aligned}$$

As a consequence, the sample complexity of deciding whether $\mathrm{T}(\rho, \rho_{\mathtt{U}})$ is at least $1 - 2^{-r^\tau}$ or at most 2^{-r^τ} is $\Omega(r^{\frac{1}{1+k}})$, where $r \coloneqq \hat{r} \cdot \hat{r}^k = \hat{r}^{1+k}$. For any $q \in \left(1, 1+\frac{1}{n-1}\right] \subseteq \left(1, 1+\frac{1}{r-1}\right]$, utilizing the reduction from QSD to $\textsc{TsallisQED}_q$ (Lemma 5.5) with parameters from Theorem 5.9, there are two corresponding states ρ_0' and ρ_1' of rank at most $2r$ such that the quantum sample complexity for deciding whether $\mathrm{S}_q(\rho_0') - \mathrm{S}_q(\rho_1')$ is at least $1/400$ or at most $-1/400$ is $\Omega(r^{\frac{1}{1+k}}) = \Omega(r^{1-\tau-c}) = \Omega(r^{0.51-c})$ for any constant $c > 0$. Therefore, estimating $\mathrm{S}_q(\rho_b')$ for $b \in \{0, 1\}$ to within additive error $1/800$ requires at least the same number of copies of ρ. □

Acknowledgments

The authors express their gratitude to the anonymous reviewers for their valuable comments, particularly for suggesting the inclusion of further explanation of the importance of estimating $\mathrm{S}_q(\rho)$ for non-integer $q > 1$. The authors also thank Kean Chen for pointing out a parameter misuse in Theorems 3.2 and 3.3 in an earlier version of this paper.

This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Quantum Leap Flagship Program (MEXT Q-LEAP) under Grant JPMXS0120319794. The work of Yupan Liu was also supported in part by JST, the establishment of University fellowships towards the creation of science technology innovation, under Grant JPMJFS2125. The work of Qisheng Wang was also supported in part by the Engineering and Physical Sciences Research Council under Grant EP/X026167/1.

References

[AISW20] Jayadev Acharya, Ibrahim Issa, Nirmal V. Shende, and Aaron B. Wagner. Estimating quantum entropy. *IEEE Journal on Selected Areas in Information Theory*, 1(2):454–468, 2020. Preliminary version in *ISIT 2019*. arXiv:1711.00814, doi:10.1109/JSAIT.2020.3015235. 2, 3, 9

[AJL09] Dorit Aharonov, Vaughan Jones, and Zeph Landau. A polynomial quantum algorithm for approximating the Jones polynomial. *Algorithmica*, 55(3):395–421,

[31] This inequalities can also be derived using the polarization lemma for the measured quantum triangular discrimination, specifically combining Theorem 3.3 and Lemma 4.11 in [Liu23].

2009. Preliminary version in *STOC 2006*. arXiv:quant-ph/0511096, doi:10.1007/s00453-008-9168-0. 5, 19

[ALL22] Anurag Anshu, Zeph Landau, and Yunchao Liu. Distributed quantum inner product estimation. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 44–51, 2022. arXiv:2111.03273, doi:10.1145/3519935.3519974. 9

[Amb14] Andris Ambainis. On physical problems that are slightly more difficult than QMA. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 32–43. IEEE, 2014. arXiv:1312.4758, doi:10.1109/ccc.2014.12. 39

[AOST17] Jayadev Acharya, Alon Orlitsky, Ananda Theertha Suresh, and Himanshu Tyagi. Estimating Renyi entropy of discrete distributions. *IEEE Transactions on Information Theory*, 63(1):38–56, 2017. arXiv:1408.1000, doi:10.1109/TIT.2016.2620435. 2, 4, 9

[vACGN23] Joran van Apeldoorn, Arjan Cornelissen, András Gilyén, and Giacomo Nannicini. Quantum tomography using state-preparation unitaries. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1265–1318, 2023. arXiv:2207.08800, doi:10.1137/1.9781611977554.ch47. 9

[AS17] Guillaume Aubrun and Stanisław J Szarek. *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*, volume 223 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2017. doi:10.1090/surv/223. 10, 26, 27, 28

[AS24] Srinivasan Arunachalam and Louis Schatzki. Distributed inner product estimation with limited quantum communication. ArXiv e-prints, 2024. arXiv:2410.12684. 9

[BASTS10] Avraham Ben-Aroya, Oded Schwartz, and Amnon Ta-Shma. Quantum expanders: motivation and construction. *Theory of Computing*, 6(3):47–79, 2010. Preliminary version in *CCC 2008*. doi:10.4086/toc.2010.v006a003. 3, 7, 16, 32, 35

[Bau11] Bernhard Baumgartner. An inequality for the trace of matrix products, using absolute values. ArXiv e-prints, 2011. arXiv:1106.6189. 22

[BBBV97] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997. arXiv:quant-ph/9701001, doi:10.1137/S0097539796300933. 4

[BCH+19] Adam Bouland, Lijie Chen, Dhiraj Holden, Justin Thaler, and Prashant Nalini Vasudevan. On the power of statistical zero knowledge. *SIAM Journal on Computing*, 49(4):FOCS17-1–FOCS17-58, 2019. Preliminary version in *FOCS 2017*. arXiv:1609.02888, doi:10.1137/17M1161749. 3

[BCWdW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001. arXiv:quant-ph/0102001, doi:10.1103/PhysRevLett.87.167902. 1, 4, 9

[BDRV19] Itay Berman, Akshay Degwekar, Ron D Rothblum, and Prashant Nalini Vasudevan. Statistical difference beyond the polarizing regime. In *Proceedings of the 17th International Conference on Theory of Cryptography Conference*, pages 311–332. Springer, 2019. ECCC:TR19-038. doi:10.1007/978-3-030-36033-7_12. 15

[Bec02] Christian Beck. Generalized statistical mechanics and fully developed turbulence. *Physica A: Statistical Mechanics and its Applications*, 306:189–198, 2002. arXiv:cond-mat/0110073, doi:10.1016/s0378-4371(02)00497-1. 1

[Bel19] Aleksandrs Belovs. Quantum algorithms for classical probability distributions. In *Proceedings of the 27th Annual European Symposium on Algorithms, ESA 2019*, volume 144 of *LIPIcs*, pages 16:1–16:11. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. arXiv:1904.02192, doi:10.4230/LIPICS.ESA.2019.16. 16, 33

[Ber14] Serge Bernstein. Sur la meilleure approximation de $|x|$ par des polynomes de degrés donnés. *Acta Mathematica*, 37(1):1–57, 1914. doi:10.1007/BF02401828. 6

[Ber38] Serge Bernstein. Sur la meilleure approximation de $|x|^p$ par des polynômes de degrés très élevés. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 2(2):169–190, 1938. URL: https://www.mathnet.ru/eng/im3513. 6

[BH09] Jop Briët and Peter Harremoës. Properties of classical and quantum Jensen-Shannon divergence. *Physical Review A*, 79(5):052311, 2009. arXiv:0806.4472, doi:10.1103/PhysRevA.79.052311. 7, 8, 12, 13, 24, 27

[BHH11] Sergey Bravyi, Aram W. Harrow, and Avinatan Hassidim. Quantum algorithms for testing properties of distributions. *IEEE Transactions on Information Theory*, 57(6):3971–3981, 2011. Preliminary version in *STACS 2010*. arXiv:0907.3920, doi:10.1109/TIT.2011.2134250. 9

[BHMT02] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. In Samuel J. Lomonaco, Jr. and Howard E. Brandt, editors, *Quantum Computation and Information*, volume 305 of *Contemporary Mathematics*, pages 53–74. AMS, 2002. arXiv:quant-ph/0005055, doi:10.1090/conm/305/05215. 4, 19

[BKT20] Mark Bun, Robin Kothari, and Justin Thaler. The polynomial method strikes back: tight quantum query bounds via dual polynomials. *Theory of Computing*, 16(10):1–71, 2020. Preliminary version in *STOC 2018*. arXiv:1710.09079, doi:10.4086/toc.2020.v016a010. 4, 9

[BOW19] Costin Bădescu, Ryan O'Donnell, and John Wright. Quantum state certification. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 503–514, 2019. arXiv:1708.06002, doi:10.1145/3313276.3316344. 9

[BR82] J. Burbea and C. Rao. On the convexity of some divergence measures based on entropy functions. *IEEE Transactions on Information Theory*, 28(3):489–495, 1982. doi:10.1109/tit.1982.1056497. 12

[Can20] Clément L. Canonne. A survey on distribution testing: your data is big. but is it blue? In *Theory of Computing Library*, number 9 in Graduate Surveys, pages 1–100. University of Chicago, 2020. ECCC:TR15-063. doi:10.4086/toc.gs.2020.009. 1

[CCKV08] André Chailloux, Dragos Florin Ciocan, Iordanis Kerenidis, and Salil Vadhan. Interactive and noninteractive zero knowledge are equivalent in the help model. In *Proceedings of the Fifth Theory of Cryptography Conference*, pages 501–534. Springer, 2008. IACR ePrint:2007/467. doi:10.1007/978-3-540-78524-8_28. 3, 7, 16, 25

[CFMdW10] Sourav Chakraborty, Eldar Fischer, Arie Matsliah, and Ronald de Wolf. New results on quantum property testing. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010)*, volume 8

of *LIPIcs*, pages 145–156. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2010. arXiv:1005.0523, doi:10.4230/LIPICS.FSTTCS.2010.145. 9, 16, 33

[CLW20] Anirban N. Chowdhury, Guang Hao Low, and Nathan Wiebe. A variational quantum algorithm for preparing quantum Gibbs states. ArXiv e-prints, 2020. arXiv:2002.00055. 9

[CM18] Chris Cade and Ashley Montanaro. The quantum complexity of computing Schatten p-norms. In *Proceedings of the 13th Conference on the Theory of Quantum Computation, Communication and Cryptography*, volume 111 of *LIPIcs*, pages 4:1–4:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. arXiv:1706.09279, doi:10.4230/LIPICS.TQC.2018.4. 9

[CT14] Richard Y. Chen and Joel A. Tropp. Subadditivity of matrix ϕ-entropy and concentration of random matrices. *Electronic Journal of Probability*, 19(27):1–30, 2014. arXiv:1308.2952, doi:10.1214/ejp.v19-2964. 8, 25

[CWLY23] Kean Chen, Qisheng Wang, Peixun Long, and Mingsheng Ying. Unitarity estimation for quantum channels. *IEEE Transactions on Information Theory*, 69(8):5116–5134, 2023. arXiv:2212.09319, doi:10.1109/TIT.2023.3263645. 9

[Dar70] Zoltán Daróczy. Generalized information functions. *Information and Control*, 16(1):36–51, 1970. doi:10.1016/s0019-9958(70)80040-7. 1, 11

[EAO+02] Artur K. Ekert, Carolina Moura Alves, Daniel K. L. Oi, Michał Horodecki, Paweł Horodecki, and Leong Chuan Kwek. Direct estimations of linear and nonlinear functionals of a quantum state. *Physical Review Letters*, 88(21):217901, 2002. arXiv:quant-ph/0203016, doi:10.1103/physrevlett.88.217901. 1, 4

[FvdG99] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999. arXiv:quant-ph/9712042, doi:10.1109/18.761271. 7, 13, 27, 42

[Fur05] Shigeru Furuichi. On uniqueness theorems for Tsallis entropy and Tsallis relative entropy. *IEEE Transactions on Information Theory*, 51(10):3638–3645, 2005. arXiv:cond-mat/0410270, doi:10.1109/tit.2005.855606. 13

[FYK04] Shigeru Furuichi, Kenjiro Yanagi, and Ken Kuriyama. Fundamental properties of Tsallis relative entropy. *Journal of Mathematical Physics*, 45(12):4868–4877, 2004. arXiv:cond-mat/0406178, doi:10.1063/1.1805729. 13, 26

[FYK07] Shigeru Furuichi, Kenjiro Yanagi, and Ken Kuriyama. A generalized Fannes' inequality. *Journal of Inequalities in Pure and Applied Mathematics*, 8(1):5, 2007. URL: https://www.emis.de/journals/JIPAM/article818.html?sid=818, arXiv:1001.1390. 13, 14

[Gan02] Michael I. Ganzburg. The Bernstein constant and polynomial interpolation at the Chebyshev nodes. *Journal of Approximation Theory*, 119(2):193–213, 2002. doi:10.1006/jath.2002.3729. 6

[GH20] Alexandru Gheorghiu and Matty J. Hoban. Estimating the entropy of shallow circuit outputs is hard. ArXiv e-prints, 2020. arXiv:2002.12814. 9

[GHS21] Tom Gur, Min-Hsiu Hsieh, and Sathyawageeswar Subramanian. Sublinear quantum algorithms for estimating von Neumann entropy. ArXiv e-prints, 2021. arXiv:2111.11139. 9

[GHYZ24] Weiyuan Gong, Jonas Haferkamp, Qi Ye, and Zhihan Zhang. On the sample complexity of purity and inner product estimation. ArXiv e-prints, 2024. arXiv:2410.12712. 9

[GL20] András Gilyén and Tongyang Li. Distributional property testing in a quantum world. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151 of *LIPIcs*, pages 25:1–25:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. arXiv:1902.00814, doi:10.4230/LIPIcs.ITCS.2020.25. 3, 9, 16

[Gol08] Oded Goldreich. *Computational Complexity: A Conceptual Perspective*. Cambridge University Press, 2008. doi:10.1017/CBO9780511804106. 15

[Gol17] Oded Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017. doi:10.1017/9781108135252. 1

[GP22] András Gilyén and Alexander Poremba. Improved quantum algorithms for fidelity estimation. ArXiv e-prints, 2022. arXiv:2203.15993. 5, 9, 19

[GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 193–204, 2019. arXiv:1806.01838, doi:10.1145/3313276.3316366. 5, 18, 19

[Hay24] Masahito Hayashi. Measuring quantum relative entropy with finite-size effect. ArXiv e-prints, 2024. arXiv:2406.17299. 2

[HC67] Jan Havrda and František Charvát. Quantification method of classification processes. concept of structural a-entropy. *Kybernetika*, 3(1):30–35, 1967. URL: https://eudml.org/doc/28681. 1

[Hel67] Carl W. Helstrom. Detection theory and quantum mechanics. *Information and Control*, 10(3):254–291, 1967. doi:10.1016/S0019-9958(67)90302-6. 12

[HHJ+17] Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, 63(9):5628–5641, 2017. Preliminary version in *STOC 2016*. arXiv:1508.01797, doi:10.1109/TIT.2017.2719044. 9

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:10.1080/01621459.1963.10500830. 24

[Hol73a] Alexander S Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973. URL: https://www.mathnet.ru/eng/ppi903. 7, 13

[Hol73b] Alexander S. Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, 1973. doi:10.1016/0047-259X(73)90028-6. 12

[JMDA21] Reza Asgharzadeh Jelodar, Hossein Mehri-Dehnavi, and Hamzeh Agahi. Some properties of Tsallis and Tsallis–Lin quantum relative entropies. *Physica A: Statistical Mechanics and its Applications*, 567:125719, 2021. doi:10.1016/j.physa.2020.125719. 13

[JUW09] Rahul Jain, Sarvagya Upadhyay, and John Watrous. Two-message quantum interactive proofs are in PSPACE. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 534–543. IEEE, 2009. arXiv:0905.1300, doi:10.1109/focs.2009.30. 3

[JVHW15] Jiantao Jiao, Kartik Venkat, Yanjun Han, and Tsachy Weissman. Minimax estimation of functionals of discrete distributions. *IEEE Transactions on Information Theory*, 61(5):2835–2885, 2015. arXiv:1406.6956, doi:10.1109/TIT.2015.2412945. 9

[JVHW17] Jiantao Jiao, Kartik Venkat, Yanjun Han, and Tsachy Weissman. Maximum likelihood estimation of functionals of discrete distributions. *IEEE Transactions on Information Theory*, 63(10):6774–6798, 2017. arXiv:1406.6959, doi:10.1109/TIT.2017.2733537. 9

[Kim16] Jeong San Kim. Tsallis entropy and general polygamy of multiparty quantum entanglement in arbitrary dimensions. *Physical Review A*, 94(6):062338, 2016. arXiv:1612.04480, doi:10.1103/physreva.94.062338. 13

[Kit95] Alexei Yu. Kitaev. Quantum measurements and the Abelian stabilizer problem. ArXiv e-prints, 1995. arXiv:quant-ph/9511026. 5, 19

[KLGN19] Hirotada Kobayashi, François Le Gall, and Harumichi Nishimura. Generalized quantum Arthur–Merlin games. *SIAM Journal on Computing*, 48(3):865–902, 2019. Preliminary version in *CCC 2015*. arXiv:arXiv:1312.4673, doi:10.1137/17m1160173. 3, 7, 8, 25, 31

[KLL+17] Shelby Kimmel, Cedric Yen-Yu Lin, Guang Hao Low, Maris Ozols, and Theodore J. Yoder. Hamiltonian simulation with optimal sample complexity. *npj Quantum Information*, 3(1):1–7, 2017. arXiv:1608.00281, doi:10.1038/s41534-017-0013-7. 19

[Kob03] Hirotada Kobayashi. Non-interactive quantum perfect and statistical zero-knowledge. In *Proceedings of the 14th International Symposium on Algorithms and Computation*, pages 178–188. Springer, 2003. arXiv:quant-ph/0207158, doi:10.1007/978-3-540-24587-2_20. 3, 5, 7, 14, 16

[KS16] Yasuhito Kawano and Hiroshi Sekigawa. Quantum Fourier transform over symmetric groups — improved result. *Journal of Symbolic Computation*, 75:219–243, 2016. doi:10.1016/j.jsc.2015.11.016. 2

[LC19] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019. arXiv:1707.05391, doi:10.22331/q-2019-07-12-163. 19

[LGDC24] Zhenhuan Liu, Weiyuan Gong, Zhenyu Du, and Zhenyu Cai. Exponential separations between quantum learning with and without purification. ArXiv e-prints, 2024. arXiv:2410.17718. 9

[LGLW23] François Le Gall, Yupan Liu, and Qisheng Wang. Space-bounded quantum state testing via space-efficient quantum singular value transformation. ArXiv e-prints, 2023. arXiv:2308.05079. 3, 5, 9, 15

[Lie73] Elliott H. Lieb. Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Advances in Mathematics*, 11(3):267–288, 1973. doi:10.1016/0001-8708(73)90011-x. 8, 25

[Lin75] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40:147–151, 1975. doi:10.1007/bf01609396. 26

[Lin91] Jianhua Lin. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi:10.1109/18.61115. 7, 25, 29

[Liu23] Yupan Liu. Quantum state testing beyond the polarizing regime and quantum triangular discrimination. ArXiv e-prints, 2023. arXiv:2303.01952. 7, 27, 34, 35, 43

[LMR14] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, 2014. arXiv:1307.0401, doi:10.1038/nphys3029. 19

[LS20] Alessandro Luongo and Changpeng Shao. Quantum algorithms for spectral sums. ArXiv e-prints, 2020. arXiv:2011.06475. 9

[LW19] Tongyang Li and Xiaodi Wu. Quantum query complexity of entropy estimation. *IEEE Transactions on Information Theory*, 65(5):2899–2921, 2019. arXiv:1710.06025, doi:10.1109/TIT.2018.2883306. 9

[LWL24] Jingquan Luo, Qisheng Wang, and Lvzhou Li. Succinct quantum testers for closeness and k-wise uniformity of probability distributions. *IEEE Transactions on Information Theory*, 70(7):5092–5103, 2024. arXiv:2304.12916, doi:10.1109/TIT.2024.3393756. 9, 17

[LWWZ25] Nana Liu, Qisheng Wang, Mark M. Wilde, and Zhicheng Zhang. Quantum algorithms for matrix geometric means. *npj Quantum Information*, page to appear, 2025. arXiv:2405.00673. 9

[MLP05] Ana P. Majtey, Pedro W. Lamberti, and Domingo P. Prato. Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states. *Physical Review A*, 72(5):052310, 2005. arXiv:quant-ph/0508138, doi:10.1103/PhysRevA.72.052310. 8, 13

[Mon15] Ashley Montanaro. Quantum speedup of Monte Carlo methods. *Proceedings of the Royal Society A*, 471(2181):20150301, 2015. arXiv:1504.06987, doi:10.1098/rspa.2015.0301. 9

[MdW16] Ashley Montanaro and Ronald de Wolf. A survey of quantum property testing. In *Theory of Computing Library*, number 7 in Graduate Surveys, pages 1–81. University of Chicago, 2016. arXiv:1310.2035, doi:10.4086/toc.gs.2016.007. 1, 2

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010. doi:10.1017/CBO9780511976667. 10, 12, 13, 27

[Ngu23] Quynh T. Nguyen. The mixed Schur transform: efficient quantum circuit and applications. ArXiv e-prints, 2023. arXiv:2310.01613. 2

[OW16] Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing*, pages 899–912, 2016. arXiv:1508.01907, doi:10.1145/2897518.2897544. 9

[OW21] Ryan O'Donnell and John Wright. Quantum spectrum testing. *Communications in Mathematical Physics*, 387(1):1–75, 2021. Preliminary version in *STOC 2015*. arXiv:1501.05028, doi:10.1007/s00220-021-04180-1. 9, 17, 33

[Pet07] Dénes Petz. *Quantum Information Theory and Quantum Statistics*. Springer, 2007. doi:10.1007/978-3-540-74636-2. 25, 26

[QKW24] Yihui Quek, Eneet Kaur, and Mark M. Wilde. Multivariate trace estimation in constant quantum depth. *Quantum*, 8:1220, 2024. arXiv:2206.15405, doi:10.22331/Q-2024-01-10-1220. 4

[Rag95] Guido A. Raggio. Properties of q-entropies. *Journal of Mathematical Physics*, 36(9):4785–4791, 1995. doi:10.1063/1.530920. 1, 13

[Ras11] Alexey E. Rastegin. Some general properties of unified entropies. *Journal of Statistical Physics*, 143:1120–1135, 2011. arXiv:1012.5356, doi:10.1007/s10955-011-0231-x. 14

[RASW23] Soorya Rethinasamy, Rochisha Agarwal, Kunal Sharma, and Mark M. Wilde. Estimating distinguishability measures on quantum computers. *Physical Review A*, 108(1):012409, 2023. arXiv:2108.08406, doi:10.1103/PhysRevA.108.012409. 7, 15

[Riv90] Theodore J. Rivlin. *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*. Courier Dover Publications, 1990. 6, 17, 18

[Roc70] Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970. doi:10.1515/9781400873173. 31

[Rus22] Mary Beth Ruskai. Yet another proof of the joint convexity of relative entropy. *Letters in Mathematical Physics*, 112(4):81, 2022. arXiv:2112.13763, doi:10.1007/s11005-022-01562-x. 25, 26

[SH21] Sathyawageeswar Subramanian and Min-Hsiu Hsieh. Quantum algorithm for estimating α-Renyi entropies of quantum states. *Physical Review A*, 104(2):022428, 2021. doi:10.1103/PhysRevA.104.022428. 9

[SLLJ24] Myeongjin Shin, Junseo Lee, Seungwoo Lee, and Kabgyun Jeong. Rank is all you need: Estimating the trace of powers of density matrices. ArXiv e-prints, 2024. arXiv:2408.00314. 9

[Sra21] Suvrit Sra. Metrics induced by Jensen-Shannon and related divergences on positive definite matrices. *Linear Algebra and its Applications*, 616:125–138, 2021. arXiv:1911.02643, doi:10.1016/j.laa.2020.12.023. 13

[SV03] Amit Sahai and Salil Vadhan. A complete problem for statistical zero knowledge. *Journal of the ACM*, 50(2):196–249, 2003. Preliminary version in *FOCS 1997*. ECCC:TR00-084. doi:10.1145/636865.636868. 14, 15

[Tim63] Aleksandr F. Timan. *Theory of Approximation of Functions of a Real Variable*, volume 34 of *International Series of Monographs on Pure and Applied Mathematics*. Pergamon Press, 1963. doi:10.1016/c2013-0-05307-8. 6, 17

[Top00] Flemming Topsøe. Some inequalities for information divergence and related measures of discrimination. *IEEE Transactions on Information Theory*, 46(4):1602–1609, 2000. doi:10.1109/18.850703. 12

[Top01] Flemming Topsøe. Bounds for entropy and divergence for distributions over a two-element set. *Journal of Inequalities in Pure and Applied Mathematics*, 2(2), 2001. URL: https://eudml.org/doc/122035. 7, 25, 28, 29

[Tsa88] Constantino Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52:479–487, 1988. doi:10.1007/bf01016429. 1, 11

[Tsa01] Constantino Tsallis. *Nonextensive Statistical Mechanics and Its Applications*, chapter I. Nonextensive Statistical Mechanics and Thermodynamics: Historical Background and Present Status, page 3–98. Springer, 2001. doi:10.1007/3-540-40919-x_1. 1, 10

[Uhl77] A. Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics*, 54:21–32, 1977. doi:10.1007/BF01609834. 8, 25, 26

[Vad99] Salil Pravin Vadhan. *A Study of Statistical Zero-Knowledge Proofs*. Phd thesis, Massachusetts Institute of Technology, 1999. URL: https://people.seas.harvard.edu/~salil/research/phdthesis.pdf. 27

[Vaj70] Igor Vajda. Note on discrimination information and variation. *IEEE Transactions on Information Theory*, 16(6):771–773, 1970. doi:10.1109/TIT.1970.1054557. 7, 8

[Vir19] Dániel Virosztek. Jointly convex quantum Jensen divergences. *Linear Algebra and its Applications*, 576:67–78, 2019. arXiv:1712.05324, doi:10.1016/j.laa.2018.03.002. 8, 25

[Vir21] Dániel Virosztek. The metric property of the quantum Jensen-Shannon divergence. *Advances in Mathematics*, 380:107595, 2021. arXiv:1910.10447, doi:10.1016/j.aim.2021.107595. 13

[Wan24] Qisheng Wang. Optimal trace distance and fidelity estimations for pure quantum states. *IEEE Transactions on Information Theory*, 2024. arXiv:2408.16655, doi:10.1109/TIT.2024.3447915. 9, 17

[Wat02] John Watrous. Limits on the power of quantum statistical zero-knowledge. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 459–468. IEEE, 2002. arXiv:quant-ph/0202111, doi:10.1109/SFCS.2002.1181970. 3, 6, 7, 14, 15, 42, 43

[Wat09] John Watrous. Zero-knowledge against quantum attacks. *SIAM Journal on Computing*, 39(1):25–58, 2009. Preliminary version in *STOC 2006*. arXiv:quant-ph/0511020, doi:10.1137/060670997. 3, 7, 15

[WGL+24] Qisheng Wang, Ji Guan, Junyi Liu, Zhicheng Zhang, and Mingsheng Ying. New quantum algorithms for computing quantum entropies and distances. *IEEE Transactions on Information Theory*, 70(8):5653–5680, 2024. arXiv:2203.13522, doi:10.1109/TIT.2024.3399014. 2, 3, 4, 9

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013. doi:10.1017/9781316809976. 12, 26

[WSP+24] Rafael Wagner, Zohar Schwartzman-Nowik, Ismael L Paiva, Amit Te'eni, Antonio Ruiz-Molero, Rui Soares Barbosa, Eliahu Cohen, and Ernesto F Galvão. Quantum circuits for measuring weak values, Kirkwood–Dirac quasiprobability distributions,

and state spectra. *Quantum Science and Technology*, 9(1):015030, 2024. arXiv:2302.00705, doi:10.1088/2058-9565/ad124c. 9

[WY16] Yihong Wu and Pengkun Yang. Minimax rates of entropy estimation on large alphabets via best polynomial approximation. *IEEE Transactions on Information Theory*, 62(6):3702–3720, 2016. arXiv:1407.0381, doi:10.1109/TIT.2016.2548468. 9

[WZ23] Qisheng Wang and Zhicheng Zhang. Quantum lower bounds by sample-to-query lifting. ArXiv e-prints, 2023. arXiv:2308.01794. 19

[WZ24a] Qisheng Wang and Zhicheng Zhang. Fast quantum algorithms for trace distance estimation. *IEEE Transactions on Information Theory*, 70(4):2720–2733, 2024. arXiv:2301.06783, doi:10.1109/TIT.2023.3321121. 5, 9, 15

[WZ24b] Qisheng Wang and Zhicheng Zhang. Sample-optimal quantum estimators for pure-state trace distance and fidelity via samplizer. ArXiv e-prints, 2024. arXiv:2410.21201. 9

[WZ24c] Qisheng Wang and Zhicheng Zhang. Time-efficient quantum entropy estimator via samplizer. In *Proceedings of the 32nd Annual European Symposium on Algorithms*, pages 101:1–101:15, 2024. arXiv:2401.09947, doi:10.4230/LIPIcs.ESA.2024.101. 2, 4, 5, 6, 9, 19, 20

[WZC+23] Qisheng Wang, Zhicheng Zhang, Kean Chen, Ji Guan, Wang Fang, Junyi Liu, and Mingsheng Ying. Quantum algorithm for fidelity estimation. *IEEE Transactions on Information Theory*, 69(1):273–282, 2023. arXiv:2103.09076, doi:10.1109/TIT.2022.3203985. 9

[WZL24] Xinzhao Wang, Shengyu Zhang, and Tongyang Li. A quantum algorithm framework for discrete probability distributions with applications to Rényi entropy estimation. *IEEE Transactions on Information Theory*, 70(5):3399–3426, 2024. arXiv:2212.01571, doi:10.1109/TIT.2024.3382037. 2, 3, 4, 9

[WZW23] Youle Wang, Benchi Zhao, and Xin Wang. Quantum algorithms for estimating quantum entropies. *Physical Review Applied*, 19(4):044041, 2023. arXiv:2203.02386, doi:10.1103/PhysRevApplied.19.044041. 9

[Yam02] Takuya Yamano. Some properties of q-logarithm and q-exponential functions in tsallis statistics. *Physica A: Statistical Mechanics and its Applications*, 305(3-4):486–496, 2002. doi:10.1016/s0378-4371(01)00567-2. 10

[Zha07] Zhengmin Zhang. Uniform estimates on the Tsallis entropies. *Letters in Mathematical Physics*, 80:171–181, 2007. doi:10.1007/s11005-007-0155-1. 14

A Omitted proofs

A.1 Omitted proof in Section 4

Proposition 4.10.1. *For the optimization problem presented in Equation (4.19), an optimal solution is the distribution provided in Equation (4.20), where* $\varepsilon = N(1-\gamma) - \lfloor N(1-\gamma) \rfloor$*:*

$$p_{\max}(i) = \begin{cases} \frac{1}{N} + \frac{\gamma}{k_{\max}}, & \text{if } i \in [k_{\max}] \\ \frac{\varepsilon}{N(N-k_{\max})}, & \text{otherwise} \end{cases}, \text{where } k_{\max} \coloneqq \lfloor N(1-\gamma) \rfloor. \tag{4.20}$$

Proof. We begin by noting that $\mathrm{H}_q(p) = \frac{1}{q-1}\left(1 - \sum_{i\in[N]} p(i)^q\right)$ is concave (Lemma 2.3) for any fixed $q > 1$. Consequently, an optimal solution $p_{\max}$ to the optimization problem specified in Equation (4.19) has a particular form. Specifically, $p_{\max}$ is one of probability distributions $p^{(k)}$ for integer $k \in [\lfloor N(1-\gamma)\rfloor]$ defined in Equation (A.1) with a maximum Tsallis entropy:[32]

$$\mathrm{H}_q(p_{\max}) = \max_{k\in[\lfloor N(1-\gamma)\rfloor]} \mathrm{H}_q\left(p^{(k)}\right), \text{ where } p^{(k)}(i) \coloneqq \begin{cases} \frac{1}{N} + \frac{\gamma}{k}, & \text{if } i \in [k] \\ \frac{1}{N} - \frac{\gamma}{N-k}, & \text{otherwise} \end{cases}. \tag{A.1}$$

Plugging Equation (A.1) into Equation (4.19), it suffices to solve the following optimization problem with $q > 1$:

$$\begin{aligned} \text{minimize} \quad & F_q(N,k,\gamma) \coloneqq \sum_{i\in[N]} p(i)^q = k\cdot\left(\frac{1}{N}+\frac{\gamma}{k}\right)^q + (N-k)\cdot\left(\frac{1}{N}-\frac{\gamma}{N-k}\right)^q \\ \text{subject to} \quad & 1/q \le \gamma \le 1 - 1/N, \\ & 1 \le k \le \lfloor N(1-\gamma)\rfloor, \\ & k, N \in \mathbb{Z}_+ \end{aligned} \tag{A.2}$$

To establish that Equation (4.20) is an optimal solution to Equation (A.2), it remains to show that the objective function $F_q(N,k,\gamma)$ is monotonically non-increasing in k for N, γ, and $q > 1$ satisfying the constraints in Equation (A.2). Equivalently, it needs to be shown that $\frac{\partial}{\partial k} F_q(N,k,\gamma) \le 0$ for $1/q \le \gamma \le 1 - 1/N$ and $1 \le k \le \lfloor N(1-\gamma)\rfloor$, specifically:

$$\frac{\partial}{\partial k} F_q(N,k,\gamma) = \frac{(k-\gamma N(q-1))\left(\frac{\gamma}{k}+\frac{1}{N}\right)^q}{k+\gamma N} + \frac{\left(\frac{1}{N}-\frac{\gamma}{N-k}\right)^q(\gamma N(q-1)+N-k)}{\gamma N-(N-k)} \le 0. \tag{A.3}$$

Since it is evident that $\frac{\gamma}{k}+\frac{1}{N} \ge 0$, $k+\gamma N \ge 0$, and $k \le \lfloor N(1-\gamma)\rfloor \le N(1-\gamma)$, we can deduce Equation (A.3) by combining the following inequalities:

$$\begin{aligned} k - \gamma N(q-1) &\le N(1-\gamma) - \gamma N(q-1) = N(1-q\gamma) \le 0, \\ \frac{1}{N} - \frac{\gamma}{N-k} &\ge \frac{1}{N} - \frac{\gamma}{N-N(1-\gamma)} = 0, \\ \gamma N(q-1) + N - k &\ge N(q-1) + N - N(1-\gamma) = Nq\gamma \ge N > 0, \\ \gamma N - (N-k) &\le N - (N - N(1-\gamma)) = 0. \end{aligned}$$

Here, the first and the third line hold also due to $\gamma \ge 1/q$. This completes the proof. □

A.2 Omitted proof in Section 5

Fact 5.6.1. *Let $g_1(n)$, $g_2(n)$, and $g_3(n)$ be functions defined in Equation (5.11). It holds that:*

(1) *For $n \ge 3$, $g_1(n) \ge 0$.*

(2) *For $n \ge 3$, $g_2(n)$ and $g_3(n)$ are monotonically increasing.*

Proof. We begin by defining $f_1(n) \coloneqq 2^{-n} + \frac{1-2^{\frac{n}{1-n}}}{n}$, $f_2(n) \coloneqq 2^{\frac{n^2}{1-n}}(n-1)$, and $f_3(n) \coloneqq \frac{n}{4}\left(1 - 2^{\frac{1}{1-n}}\right)$ such that $g_1(n) = f_1(n) + f_2(n) + f_3(n)$. We then prove the first item separately:

- For $f_1(n)$, since $2^{\frac{n}{1-n}} = 2^{-\left(1+\frac{1}{n-1}\right)}$, we know that $f_1(n)$ is monotonically decreasing for $n \ge 2$, and thus, $f_1(n) \ge \lim_{n\to\infty} f_1(n) = 0$ for $n \ge 2$.
- For $f_2(n)$, noting that $\frac{\mathrm{d}}{\mathrm{d}n} f_2(n) = \frac{2^{n^2/(1-n)}}{n-1}\left(-\log(2)n^2 + (1+\log(2))n - 1\right)$, we obtain that $f_2(n)$ is monotonically decreasing for $n \ge 3 > \frac{1+2\log(2)+\sqrt{1+4\log(2)^2}}{2\log(2)} \approx 2.9544$, and consequently, $f_2(n) \ge \lim_{n\to\infty} g_2(n) = 0$ for $n \ge 3$.

[32] It is easy to verify that $\frac{1}{N} - \frac{\gamma}{N-k} \ge 0$ holds if and only if $k \le N(1-\gamma)$ holds.

- For $f_3(n)$, it suffices to show that $2^{1/(1-n)} \leq 1$ for $n \geq 3$. Since $2^{1/(1-n)}$ is monotonically increasing for $n \geq 3$, we prove the first item by noting that $2^{1/(1-n)} \leq \lim_{n\to\infty} 2^{1/(1-n)} = 1$.

For $g_2(n)$, noting that $\frac{\mathrm{d}}{\mathrm{d}n} g_3(n) = \frac{2^{\frac{1}{1-n}} \log(2)}{(n-1)^2} + 2^{-n}(n\log(2) - 1)$ and $n\log(2) \geq 1$ for $n \geq 2$, we obtain that $g_3(n)$ is monotonically increasing for $n \geq 2$.

For $g_3(n)$, since $2^{-1/x}$ is monotonically increasing for $x \geq 1$, we have $g_3(n) \geq \frac{1}{4}n(n^{1/n} - 2^{-1/n})$. It remains to show that $\tilde{g}_3(n) \coloneqq \frac{1}{4}n(n^{1/n} - 2^{-1/n})$ are monotonically increasing for $n \geq 3$, namely:

$$\frac{\mathrm{d}}{\mathrm{d}n}\tilde{g}_3(n) = \frac{1}{4n}\underbrace{\left(n^{1/n} - 2^{-1/n}\log(2)\right)}_{f_4(n)} + \underbrace{\frac{1}{4n}\left(n^{1/n}n - n^{1/n}\log(n) - n2^{-1/n}\right)}_{f_5(n)} \geq 0. \tag{A.4}$$

Noting that $\frac{\mathrm{d}}{\mathrm{d}n} f_4(n) = -\frac{1}{n^2}\left(n^{1/n}(\log(n) - 1) + 2^{-1/n}\log^2(2)\right) < 0$ for $n > e$, namely $f_4(n)$ is monotonically decreasing for $n \geq 3$, we obtain that $\frac{f_4(n)}{4n} \geq \frac{1}{4n}\lim_{n\to\infty} f_4(n) = \frac{1}{4n} > 0$. Let $f_6(n) \coloneqq \left(\frac{1}{2n}\right)^{1/n}$. Notice that $\frac{\mathrm{d}}{\mathrm{d}n} f_6(n) = 2^{-1/n}\left(\frac{1}{n}\right)^{\frac{1}{n}+2}\left(-\log\left(\frac{1}{n}\right) - 1 + \log(2)\right) \geq 0$ for $n \geq e/2$, we have that $f_6(n) = \left(\frac{1}{2n}\right)^{1/n} \geq f_6(2) = 1/2$ for $n \geq 2$. Consequently, we can derive that:

$$\left(\frac{1}{n}\right)^{\frac{1}{n}}\frac{\mathrm{d}f_5(n)}{\mathrm{d}n} = \frac{(\log(n)-1)\log(n) - \left(\frac{1}{2n}\right)^{\frac{1}{n}} n\log(2)}{4n^3} \leq \frac{1}{4n^2}\left(\frac{(\log(n)-1)\log(n)}{n} - \frac{\log(2)}{2}\right) < 0.$$

Here, the last inequality follows by assuming $f_7(n) \coloneqq \frac{\log(n)(\log(n)-1)}{n} < \frac{\log(2)}{2}$. A direct calculation implies that $\frac{\mathrm{d}}{\mathrm{d}n} f_7(n) = -\frac{1}{n^2}((\log(n) - 3)\log(n) + 1) = 0$ have two zeros at $n = \exp(\frac{3\pm\sqrt{5}}{2})$. Therefore, we establish Equation (A.4) by noticing

$$f_7(n) \leq \max\left\{f_7\left(\frac{3-\sqrt{5}}{2}\right), f_7\left(\frac{3+\sqrt{5}}{2}\right)\right\} < \frac{\log(2)}{2}. \qquad \square$$

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il