

Performance Guarantees for Quantum Neural Estimation of Entropies

Sreejith Sreekumar, Ziv Goldfeld and Mark M Wilde

L2S, CentraleSupélec, CNRS, University of Paris-Saclay and Cornell University

Full version of paper available at [arXiv:2511.19289](https://arxiv.org/abs/2511.19289) [quant-ph] [1]

Introduction: Entropy, a concept rooted in thermodynamics and statistical mechanics, plays a fundamental role in quantifying the amount of uncertainty or disorder in physical systems. Von Neumann and Shannon introduced these concepts to quantum mechanics and classical information theory, now referred to as the von Neumann entropy [2] and Shannon entropy [3], respectively. Rényi generalized this concept axiomatically, showing that any uncertainty measure satisfying a set of natural axioms can be characterized in terms of a single parameter family of entropy measures now known as Rényi entropies [4]. Since their seminal works, related measures of information and discrepancy, such as mutual information, relative entropy [5, 6], and their Rényi counterparts [7–10], have found operational interpretations in classical and quantum hypothesis testing [7, 11–15], channel coding, data compression, and other related tasks.

Given the significance of entropic measures in characterizing the fundamental limits of quantum information systems, computing them is an important problem. However, often the underlying state is unknown, which necessitates estimating these quantities based on measurements of the quantum system. To this end, several estimators have been proposed in both the classical and quantum settings [16–29]. Recently, quantum neural estimation (QNE), employing a hybrid classical-quantum architecture composed of neural nets and quantum circuits, has been proposed for estimating quantum entropies and measured relative entropies [30] (see also [31, 32]). QNE utilizes the variational form of the entropic measure given as the supremum over Hermitian operators of a trace functional with respect to the underlying states [33–35]. For instance, the measured relative entropy $D_M(\rho||\sigma)$ between quantum states ρ and σ on a Hilbert space of dimension d admits the variational form

$$D_M(\rho||\sigma) = \sup_H \text{Tr}[H\rho] - \text{Tr}[e^H\sigma] + 1, \quad (1)$$

where the supremum is over the class of Hermitian operators H . When $\rho, \sigma > 0$, the supremum in (1) is achieved for some $H^*(\rho, \sigma) = \sum_{i=1}^d \lambda_i^* P_i^*$, where $\lambda_i^* := \log \text{Tr}[P_i^* \rho] - \log \text{Tr}[P_i^* \sigma]$ and $\{P_i^*\}_{i=1}^d$ denotes a set of orthogonal rank-one projectors from the spectral decomposition of $H^*(\rho, \sigma)$. By parametrizing the eigenvalues and eigenvectors of H by the neural net class $\mathcal{F}$ and parametrized quantum circuit (PQC) $\mathcal{U}(\Theta)$, indexed by a set Θ , respectively, the QNE leverages their expressivity to construct scalable estimators for $D_M(\rho||\sigma)$ based on n copies of ρ and σ each as follows [30]:

$$\hat{D}_{M,\Theta,\mathcal{F}}^n := \sup_{\theta \in \Theta} \hat{D}_{M,\theta,\mathcal{F}}^n, \quad \text{with} \quad \hat{D}_{M,\theta,\mathcal{F}}^n := \sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{\ell=1}^n f(i_\ell(\theta)) - \frac{1}{n} \sum_{\ell=1}^n e^{f(j_\ell(\theta))} + 1. \quad (2)$$

Here, $\{i_\ell(\theta)\}_{\ell=1}^n$ and $\{j_\ell(\theta)\}_{\ell=1}^n$ denote independent and identically distributed samples according to pmfs p_θ^ρ and q_θ^σ , respectively, with $p_\theta^\rho(i) := \text{Tr}[U(\theta)|i\rangle\langle i|U^\dagger(\theta)\rho]$ and $q_\theta^\sigma(i) := \text{Tr}[U(\theta)|i\rangle\langle i|U^\dagger(\theta)\sigma]$. QNEs for measured Rényi relative entropy of order $\alpha \neq 0, 1, \infty$ can be obtained similarly using the relevant variational forms. In [30], it was shown via numerical simulations that such QNEs are consistent.

Main Contributions: We provide the first theoretical guarantees for QNEs of measured relative entropy and measured Rényi relative entropy in the form of sharp bounds on the expected absolute error. The bounds depend on the number of quantum state copies the algorithm uses, parameters defining the class of underlying density operators, and the classical-quantum circuit parameters. Furthermore, we also derive concentration inequalities for the absolute error about the ground truth, which leads to an upper bound on the copy complexity of QNE, i.e., the minimum number of copies of the quantum states required for estimating measured relative entropies.

To summarize our contributions, let $\mathcal{S}_d(b)$ be the set of density operator pairs (ρ, σ) for a Hilbert space of dimension d , such that their Thompson metric [36] is bounded from above by $\log b$. Assum-

Rényi order and state class	Minimax risk upper bound (achieved by QNE)	Minimax risk lower bound (valid for any estimator)
$\alpha = 1$ and $\mathcal{S} \subseteq \mathcal{S}_d(b)$	$O_b\left(\delta + \sqrt{\frac{d\Theta(\delta, \mathcal{S})}{n}}\right)$	$\Omega_b\left(\frac{1}{\sqrt{n}} + \frac{d}{n \log d}\right)$
$0 < \alpha \neq 1$ and $\mathcal{S} \subseteq \mathcal{S}_d(b)$	$O_{\alpha,b}\left(\delta + \sqrt{\frac{d\Theta(\delta, \mathcal{S})}{n}}\right)$	$\Omega_{\alpha,b}\left(\frac{1}{\sqrt{n}} + \frac{d}{n \log d}\right)$

Table 1: Minimax risk bounds for estimation of measured relative entropies over the class $\mathcal{S} \subseteq \mathcal{S}_d(b)$. Here, b is a uniform upper bound on the Thompson metric between density operator pairs in $\mathcal{S}_d(b)$, δ is an approximation error parameter for PQC parametrized by $\Theta(\delta, \mathcal{S})$, and n is the number of copies of quantum states on a Hilbert space of dimension d used by the QNE. The upper bound holds for any $\mathcal{S} \subseteq \mathcal{S}_d(b)$ while the lower bound holds for any $\mathcal{S} \subseteq \mathcal{S}_d(b)$ containing all state pairs in any fixed eigen-basis, when $db \lesssim n \log d$ and $b \geq (\log d)^2 \vee 2$.

ing that the neural net and quantum circuit are sufficiently expressive and have bounded statistical complexity, we show the following:

- (i) We derive an upper bound on the expected absolute error achieved by QNE for estimating the measured relative entropy and the measured Rényi relative entropy over the class $\mathcal{S}_d(b)$. Further, we establish an exponential deviation inequality that shows that the absolute error has sub-Gaussian tails. To obtain concrete bounds, we specialize the aforementioned results to shallow neural nets with sufficiently large width and deep neural nets with a sufficient number of layers. An upper bound on the absolute error achieved by a QNE with a shallow net over the class $\mathcal{S}_d(b)$ is illustrated in Table 1, along with a lower bound on the minimax risk (valid for any estimator).
- (ii) Let $d = m^N$, with N denoting the number of qudits, and let ρ, σ be density operators defined on N qudits. While the QNE has a worst-case copy complexity that scales exponentially in N , we show that for a subclass (which depends on the quantum circuit) of permutation invariant states across N qudits within $\mathcal{S}_d(b)$, the algorithm has a much better polynomial dependence. This is possible due to Schur–Weyl duality from representation theory, which implies that the number of degrees of freedom for such states scales polynomially in N .

Background and Related Work: Measured relative entropy was introduced in [37, 38] and since then has found operational significance in terms of characterizing the Stein’s and strong converse exponents in centralized, adversarial, and distributed hypothesis testing (see e.g., [39–44]). A pertinent approach for estimating such measures is to use quantum tomography to learn the entire density matrix, from which the desired entropic quantity of interest can be computed. The copy complexity of this approach is linear in the dimension of the state (or, equivalently, exponential in the number of qudits) [45–48]. A possibly more efficient approach is to estimate the entropies directly, without estimating the underlying states. In this regard, several estimators have been proposed and their computational complexities investigated under different input models [16–28]. In particular, [18] provided algorithms for estimating the von Neumann entropy and Rényi entropies, given access to independent copies of the input quantum state, which achieves the copy complexity (sub-linear in dimension) for integer orders greater than one. Entropy estimation under the quantum query model, i.e., where one has access to an oracle that prepares the input quantum state, was explored in [19, 21]. These works also established sub-linear (in dimension) query complexity bounds for estimating the von Neumann and Rényi entropies.

In contrast to the estimators mentioned above, QNE proposed in [30] is based on variational methods, which have become a prominent research area in quantum computing and machine learning [49–52]. Specifically, QNEs utilize a hybrid quantum-classical algorithm that extends the approach from [53–59] to the quantum framework. The appeal of such estimators stems from their scalability to high-dimensional problems and large datasets, as well as their compatibility with modern gradient-based optimization techniques. This makes them particularly relevant in the quantum context, where the

dimension of the underlying Hilbert space scales exponentially in the number of qudits. Although QNE was shown to perform empirically well in [30], theoretical performance guarantees are still lacking, which this work aims to advance upon. We note that variational methods have also been used in [60] for estimating fidelity, entropies and Fisher information, in [31] for estimating quantum mutual information and in [32] for estimating entanglement entropy, which indicates their versatility.

Key Technical Contributions: Our objective is to quantify the performance of QNE in terms of d , n , and the parameters defining $\mathcal{F}$, $\mathcal{U}$, and $\mathcal{S}$. To that end, we split the error of QNE into two components: *approximation* and *statistical*. In practice, there is also an *optimization* error component, which arises when the suprema in (2) are only approximately computed, e.g., by gradient-based methods, whose quantification we leave as part of future work. The approximation error depends on the ability of the classical neural net and PQC to approximate, respectively, the eigenvalues and eigenvectors of the optimizer $H^*(\rho, \sigma)$. The statistical error, on the other hand, is the random component of the overall error that depends on the complexity of the classical-quantum model used by the QNE.

To state the performance guarantees for QNE, we first introduce the required assumptions on the neural net class and PQC. For a positive integer $m \leq d$, let $e_m: \{1, \dots, d\} \mapsto \mathbb{R}^m$ denote an embedding, and let $\mathcal{X} := \{e_m(i)\}_{i=1}^d$ be the set of embedded points. Let $\mathcal{F}_{b,\mathcal{X}}$ denote the class of continuous functions $f: \mathbb{R}^m \mapsto \mathbb{R}$ such that $\|f\|_{\infty,\mathcal{X}} \leq \log b$, which we will simply denote by $\mathcal{F}_b$. We assume that the neural class $\mathcal{F}$ is a closed subset of $\mathcal{F}_b$ and satisfies

$$\log N(2\epsilon \log b, \mathcal{F}, \|\cdot\|_{\infty,\mathcal{X}}) \leq \bar{\kappa}(\mathcal{F}, \epsilon), \quad \forall \epsilon \in (0, 1],$$

where $N(\epsilon, \mathcal{F}, \|\cdot\|_{\infty,\mathcal{X}})$ denotes the ϵ -covering number of $\mathcal{F}$ in $L^\infty(\mathcal{X})$ norm [61] and $\bar{\kappa}(\mathcal{F}, \epsilon) \geq 0$ is such that $\kappa(\mathcal{F}) := \int_0^1 \sqrt{\bar{\kappa}(\mathcal{F}, \epsilon)} d\epsilon < \infty$. Define the following approximation error component:

$$\varepsilon(\mathcal{F}, \mathcal{S}) := \min_{f \in \mathcal{F}} \sup_{(\rho, \sigma) \in \mathcal{S}} \max_{\{\lambda_i^*\}_{i=1}^d} \max_{i \in [1:d]} |f(e_m(i)) - \lambda_i^*|.$$

Further, we denote by $\mathcal{U}(\delta, \mathcal{S})$ a PQC composed of a finite set of unitary operators, parametrized by $\theta \in \Theta(\delta, \mathcal{S})$ which can approximate a unitary $U^*(\rho, \sigma)$ formed from eigenvectors of $H^*(\rho, \sigma)$ for any $(\rho, \sigma) \in \mathcal{S}$ up to δ precision, i.e., there exists $\theta \in \Theta(\delta, \mathcal{S})$ such that $\|U^*(\rho, \sigma) - U(\theta)\| \leq \delta$, where $\|\cdot\|$ denotes the operator norm. Then, we have the following result.

Theorem 1 Consider a QNE for estimating $D_M(\rho\|\sigma)$ over $\mathcal{S} \subseteq \mathcal{S}_d(b)$ with a classical neural net $\mathcal{F}$ and a PQC $\mathcal{U}(\delta, \mathcal{S})$ parametrized by $\Theta(\delta, \mathcal{S})$. Setting $\xi_{\mathcal{F}, \mathcal{S}, \delta, b} := \varepsilon(\mathcal{F}, \mathcal{S})(b+1) + 2\delta(b + \log b)$, we have

$$\begin{aligned} \sup_{(\rho, \sigma) \in \mathcal{S}} \mathbb{E} \left[\left| D_M(\rho\|\sigma) - \hat{D}_{M, \Theta(\delta, \mathcal{S}), \mathcal{F}}^n \right| \right] &\leq \xi_{\mathcal{F}, \mathcal{S}, \delta, b} + 48(b+1)(\log b) \kappa(\mathcal{F}) n^{-\frac{1}{2}}, \\ \sup_{(\rho, \sigma) \in \mathcal{S}} \mathbb{P} \left(\left| D_M(\rho\|\sigma) - \hat{D}_{M, \Theta(\delta, \mathcal{S}), \mathcal{F}}^n \right| \geq \xi_{\mathcal{F}, \mathcal{S}, \delta, b} + 128(b+1)(\log b) (n^{-\frac{1}{2}} \kappa(\mathcal{F}) + z) \right) &\leq 2e^{-nz^2}, \quad z \geq 0. \end{aligned}$$

By specializing the above result to appropriate shallow nets, we obtain as a corollary that with probability at least $p \in (0, 1)$, the QNE estimates $D_M(\rho\|\sigma)$ between $(\rho, \sigma) \in \mathcal{S}$ within an error of ϵ given $O_{b,p}(d|\Theta(\mathcal{U})|\epsilon^{-2})$ copies each of ρ and σ , provided $U^*(\rho, \sigma)$ can be approximated by the PQC $\mathcal{U}$ within an accuracy of δ . This copy complexity can be improved when ρ and σ have additional symmetries such as permutation invariance by pre-processing the states via an appropriate quantum channel.

Future Outlook: Several open questions remain in the context of quantum neural estimation of entropies. One such issue pertains to the quantification of the optimization error and an analysis of optimization dynamics such as the investigation of barren plateaus and local optima, which was neglected here. This would call for an in-depth study of the hybrid optimization landscape induced by the neural net and quantum circuit, and an analysis of specific parameterizations or architectures that lead to relatively fast convergence of the algorithm. Also relevant is the question of estimating the quantum relative entropy (see [29]) and its Rényi generalizations using QNE.

References

- [1] S. Sreekumar, Z. Goldfeld, and M. M. Wilde, “Performance Guarantees for Quantum Neural Estimation of Entropies,” *Quantum*, vol. 10, p. 2113, May 2026.
- [2] J. von Neumann, “Thermodynamik quantenmechanischer gesamtheiten,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1927, pp. 273–291, 1927. [Online]. Available: <http://eudml.org/doc/59231>
- [3] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [4] A. Rényi, “On measures of entropy and information,” *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, vol. 4.1, pp. 547–561, Jan. 1961. [Online]. Available: <https://digicoll.lib.berkeley.edu/record/112906>
- [5] S. Kullback and R. A. Leibler, “On information and sufficiency,” *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951.
- [6] H. Umegaki, “Conditional expectations in an operator algebra IV (entropy and information),” *Kodai Mathematical Seminar Reports*, vol. 14, no. 2, pp. 59–85, 1962.
- [7] D. Petz, “Quasi-entropies for states of a von Neumann algebra,” *Publications of the Research Institute for Mathematical Sciences*, vol. 21, no. 4, pp. 787–800, Aug. 1985.
- [8] ———, “Quasi-entropies for finite quantum systems,” *Reports on Mathematical Physics*, vol. 23, no. 1, pp. 57–65, Feb. 1986.
- [9] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, “On quantum Rényi entropies: A new generalization and some properties,” *Journal of Mathematical Physics*, vol. 54, no. 12, p. 122203, 2013.
- [10] M. M. Wilde, A. Winter, and D. Yang, “Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy,” *Communications in Mathematical Physics*, vol. 331, pp. 593–622, 2014.
- [11] T. Ogawa and H. Nagaoka, “Strong converse and Stein’s lemma in quantum hypothesis testing,” *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2428–2433, 2000.
- [12] T. Ogawa and M. Hayashi, “On error exponents in quantum hypothesis testing,” *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1368–1372, 2004.
- [13] M. Nussbaum and A. Szkoła, “The Chernoff lower bound for symmetric quantum hypothesis testing,” *The Annals of Statistics*, vol. 37, no. 2, pp. 1040–1057, 2009.
- [14] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, “Asymptotic error rates in quantum hypothesis testing,” *Communications in Mathematical Physics*, vol. 279, no. 1, pp. 251–283, 2008.
- [15] M. Berta, F. Brandao, and C. Hirche, “On composite quantum hypothesis testing,” *Communications in Mathematical Physics*, vol. 385, pp. 55–77, Jul. 2021.
- [16] T. Li and X. Wu, “Quantum query complexity of entropy estimation,” *IEEE Transactions on Information Theory*, vol. 65, no. 5, pp. 2899–2921, Mar. 2017.
- [17] Y. Subasi, L. Cincio, and P. J. Coles, “Entanglement spectroscopy with a depth-two quantum circuit,” *Journal of Physics A: Mathematical and Theoretical*, vol. 52, no. 4, p. 44001, Mar. 2019.

- [18] J. Acharya, I. Issa, N. V. Shende, and A. B. Wagner, “Estimating quantum entropy,” *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 2, pp. 454–468, 2020.
- [19] S. Subramanian and M.-H. Hsieh, “Quantum algorithm for estimating α -Rényi entropies of quantum states,” *Physical Review A*, vol. 104, p. 022428, Aug 2021.
- [20] A. Gilyén and T. Li, “Distributional property testing in a quantum world,” in *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), T. Vidick, Ed., vol. 151. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2020, pp. 25:1–25:19.
- [21] T. Gur, M.-H. Hsieh, and S. Subramanian, “Sublinear quantum algorithms for estimating von Neumann entropy,” *arXiv:2111.11139*, Nov. 2021.
- [22] Y. Wang, B. Zhao, and X. Wang, “Quantum algorithms for estimating quantum entropies,” *Physical Review Applied*, vol. 19, no. 4, p. 044041, Apr. 2023.
- [23] Q. Wang, J. Guan, J. Liu, Z. Zhang, and M. Ying, “New quantum algorithms for computing quantum entropies and distances,” *IEEE Transactions on Information Theory*, vol. 70, no. 8, p. 5653–5680, Aug. 2024.
- [24] A. Gilyén and A. Poremba, “Improved quantum algorithms for fidelity estimation,” 2022. [Online]. Available: <https://arxiv.org/abs/2203.15993>
- [25] Y. Wang, L. Zhang, Z. Yu, and X. Wang, “Quantum phase processing and its applications in estimating phase and entropies,” *Physical Review A*, vol. 108, no. 6, p. 062413, Dec. 2023.
- [26] Q. Wang, Z. Zhang, K. Chen, J. Guan, W. Fang, J. Liu, and M. Ying, “Quantum algorithm for fidelity estimation,” *IEEE Transactions on Information Theory*, vol. 69, no. 1, pp. 273–282, Mar. 2023.
- [27] X. Wang, S. Zhang, and T. Li, “A quantum algorithm framework for discrete probability distributions with applications to Rényi entropy estimation,” *IEEE Transactions on Information Theory*, vol. 70, no. 5, pp. 3399–3426, 2024.
- [28] Q. Wang and Z. Zhang, “Time-efficient quantum entropy estimator via sampler,” *IEEE Transactions on Information Theory*, vol. 71, no. 12, pp. 9569–9599, 2025.
- [29] Y. Lu and K. Fang, “Estimating quantum relative entropies on quantum computers,” *Communications Physics*, 2026.
- [30] Z. Goldfeld, D. Patel, S. Sree Kumar, and M. M. Wilde, “Quantum neural estimation of entropies,” *Physical Review A*, vol. 109, p. 032431, Mar 2024.
- [31] M. Shin, J. Lee, and K. Jeong, “Estimating quantum mutual information through a quantum neural network,” *Quantum Information Processing*, vol. 23, no. 2, p. 57, 2024.
- [32] S. Lee, H. Kwon, and J. S. Lee, “Estimating entanglement entropy via variational quantum circuits with classical neural networks,” *Physical Review E*, vol. 109, p. 044117, Apr. 2024.
- [33] D. Petz, “A variational expression for the relative entropy,” *Communications in Mathematical Physics*, vol. 114, no. 2, pp. 345–349, 1988.
- [34] —, *Quantum Information Theory and Quantum Statistics*. Springer Science & Business Media, 2007.
- [35] M. Berta, O. Fawzi, and M. Tomamichel, “On variational expressions for quantum relative entropies,” *Letters in Mathematical Physics*, vol. 107, no. 12, pp. 2239–2265, Dec. 2015.

- [36] A. C. Thompson, “On certain contraction mappings in a partially ordered vector space,” *Proceedings of the American Mathematical Society*, vol. 14, pp. 438–443, 1963.
- [37] M. J. Donald, “On the relative entropy,” *Communications in Mathematical Physics*, vol. 105, no. 1, pp. 13–34, 1986.
- [38] M. Piani, “Relative entropy of entanglement and restricted measurements,” *Physical Review Letters*, vol. 103, no. 16, p. 160504, Oct. 2009.
- [39] F. Hiai and D. Petz, “The proper formula for relative entropy and its asymptotics in quantum probability,” *Communications in Mathematical Physics*, vol. 143, no. 1, pp. 99–114, 1991.
- [40] M. Hayashi, “Optimal sequence of quantum measurements in the sense of Stein’s lemma in quantum hypothesis testing,” *Journal of Physics A: Mathematical and General*, vol. 35, no. 50, p. 10759, Dec. 2002.
- [41] M. Mosonyi and T. Ogawa, “Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies,” *Communications in Mathematical Physics*, vol. 334, pp. 1617–1648, 2015.
- [42] F. Brandão, A. W. Harrow, J. R. Lee, and Y. Peres, “Adversarial hypothesis testing and a quantum Stein’s lemma for restricted measurements,” *IEEE Transactions on Information Theory*, vol. 66, pp. 5037–5054, 2020.
- [43] T. Rippchen, S. Sreekumar, and M. Berta, “Locally-measured Rényi divergences,” *IEEE Transactions on Information Theory*, vol. 71, no. 8, pp. 6105–6133, 2025.
- [44] S. Sreekumar, C. Hirche, H.-C. Cheng, and M. Berta, “Distributed quantum hypothesis testing under zero-rate communication constraints,” *Annales Henri Poincaré*, 10 2025.
- [45] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, “Sample-optimal tomography of quantum states,” *IEEE Transactions on Information Theory*, vol. 63, no. 9, pp. 5628–5641, 2017.
- [46] J. Wright, “How to learn a quantum state,” Ph.D. dissertation, Carnegie Mellon University, 2016. [Online]. Available: <https://people.eecs.berkeley.edu/~jswright/papers/thesis.pdf>
- [47] R. O’Donnell and J. Wright, “Efficient quantum tomography,” in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, ser. STOC ’16, . New York, NY, USA: Association for Computing Machinery, 2016, p. 899–912.
- [48] H. Yuen, “An improved sample complexity lower bound for (fidelity) quantum state tomography,” *Quantum*, vol. 7, p. 890, Jan. 2023.
- [49] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” *Nature*, vol. 549, no. 7671, pp. 195–202, Sep. 2017.
- [50] M. Cerezo, A. Poremba, L. Cincio, and P. J. Coles, “Variational Quantum Fidelity Estimation,” *Quantum*, vol. 4, p. 248, Mar. 2020.
- [51] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, “Variational quantum algorithms,” *Nature Reviews Physics*, vol. 3, no. 9, pp. 625–644, Aug. 2021.
- [52] J. Jäger and R. V. Krems, “Universal expressiveness of variational quantum classifiers and quantum kernels for support vector machines,” *Nature Communications*, vol. 14, no. 1, p. 576, Feb. 2023.

- [53] X. Nguyen, M. J. Wainwright, and M. I. Jordan, “Estimating divergence functionals and the likelihood ratio by convex risk minimization,” *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5847–5861, 2010.
- [54] S. Arora, R. Ge, Y. Liang, T. Ma, and Y. Zhang, “Generalization and equilibrium in generative adversarial nets (GANs),” in *Proceedings of the 34th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 70. PMLR, Aug. 2017, pp. 224–232. [Online]. Available: <https://proceedings.mlr.press/v70/arora17a.html>
- [55] P. Zhang, Q. Liu, D. Zhou, T. Xu, and X. He, “On the discrimination-generalization tradeoff in GANs,” in *6th International Conference on Learning Representations*, Apr. 2018. [Online]. Available: https://openreview.net/forum?id=Hk9Xc_lR-
- [56] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, “Mutual information neural estimation,” in *Proceedings of the 35th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 80. PMLR, Jul. 2018, pp. 531–540. [Online]. Available: <https://proceedings.mlr.press/v80/belghazi18a.html>
- [57] S. Sreekumar, Z. Zhang, and Z. Goldfeld, “Non-asymptotic performance guarantees for neural estimation of f -divergences,” in *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, vol. 130, Apr. 2021, pp. 3322–3330. [Online]. Available: <https://proceedings.mlr.press/v130/sreekumar21a.html>
- [58] S. Sreekumar and Z. Goldfeld, “Neural estimation of statistical divergences,” *Journal of Machine Learning Research*, vol. 23, no. 126, pp. 1–75, 2022. [Online]. Available: <https://www.jmlr.org/papers/v23/21-1212.html>
- [59] J. Birrell, P. Dupuis, M. A. Katsoulakis, L. Rey-Bellet, and J. Wang, “Variational representations and neural network estimation of Rényi divergences,” *SIAM Journal on Mathematics of Data Science*, vol. 3, no. 4, pp. 1093–1116, 2021.
- [60] K. C. Tan and T. Volkoff, “Variational quantum algorithms to estimate rank, quantum entropies, fidelity, and Fisher information via purity minimization,” *Physical Review Research*, vol. 3, p. 033251, Sep 2021.
- [61] A. N. Kolmogorov and V. M. Tikhomirov, “ ε -entropy and ε -capacity of sets in functional spaces,” *Uspekhi Matematicheskikh Nauk*, vol. 14, no. 2, pp. 3–86, 1959. [Online]. Available: <http://mi.mathnet.ru/eng/rm7289>