Distance is employed as a cost function in recent variations of the trace distance, showing that bounding the trace distance between two states, i.e., \( D(\rho, \sigma) = \frac{1}{2} \text{Tr}[\rho - \sigma] = \frac{1}{2} \|\rho - \sigma\|_1 \),

(1)

where \( \|M\|_1 = \text{Tr}\sqrt{M^\dagger M} \) is called the 1-norm or trace norm. Computing the trace distance may, in general, involve diagonalizing the matrix \( \Delta = \rho - \sigma \), and hence can be computationally difficult.

An alternative metric is the Hilbert-Schmidt distance:

\[ D_{\text{HS}}(\rho, \sigma) = \text{Tr}[\rho - \sigma]^2 = \|\rho - \sigma\|_2^2, \]

(2)

where \( \|M\|_2 = \sqrt{\text{Tr}(M^\dagger M)} \) is called the 2-norm or Frobenius norm. While the Hilbert-Schmidt distance does not share the operational relevance of the trace distance [6], it has the benefit that one can compute it without doing matrix diagonalization, and furthermore it is known to be efficiently computable on a quantum computer [7, 8]. Because of the latter, the Hilbert-Schmidt distance is employed as a cost function in recent variational hybrid quantum-classical algorithms [9, 10].

In many applications, one is interested in upper-bounding the trace distance between two states, i.e., showing that \( D(\rho, \sigma) \leq \epsilon \) for some small number \( \epsilon \). Hence, in this work, we explore upper bounds on \( D(\rho, \sigma) \), and in particular, upper bounds formulated in terms of \( D_{\text{HS}}(\rho, \sigma) \), since the latter may be computable.

For example, the norm equivalence [11] of the trace norm and Frobenius norm can be written as

\[ \|M\|_2 \leq \|M\|_1 \leq \sqrt{\text{rank}(M)} \|M\|_2. \]

(3)

Applying this to \( M = \Delta \) gives

\[ (1/4)D_{\text{HS}}(\rho, \sigma) \leq D(\rho, \sigma)^2 \leq (d/4) \cdot D_{\text{HS}}(\rho, \sigma), \]

(4)

where \( d \) is the Hilbert space dimension, and we used the inequality \( \text{rank}(\Delta) \leq d \), which is the best one can do without any prior knowledge about \( \Delta \). As \( d \) grows exponentially in the number of quantum subsystems, \( d \) can be very large, making the upper bound in (4) very weak. Hence it is natural to ask whether the upper bound in (4) can be tightened for certain states \( \rho \) and \( \sigma \).

The goal of this article is to explore possible tightenings of this upper bound when \( \rho \) and/or \( \sigma \) are low-rank states or low-entropy states. Such tightening would be expected since the extreme case of pure states, \( \rho = \ket{\psi}\bra{\psi} \) and \( \sigma = \ket{\phi}\bra{\phi} \), gives \( D(\rho, \sigma)^2 = 1 - |\bra{\psi}\phi\rangle|^2 \) and \( D_{\text{HS}}(\rho, \sigma) = 2 - 2|\bra{\psi}\phi\rangle|^2 \). Hence for pure states we have

\[ D(\rho, \sigma)^2 = (1/2)D_{\text{HS}}(\rho, \sigma), \]

(5)

and no dimension-dependent factor appears here.

In what follows, we first discuss tightenings of (4) for low-rank states, including our main result of an upper bound that is stronger than what one could obtain via norm equivalence. Then we explore some bounds for the more general scenario of low-entropy states.

I. INTRODUCTION

The trace distance is employed in the definition of security in quantum cryptography [1, 2], is related to the error probability in quantum state discrimination [3], and generally has operational relevance to many quantum information protocols [4, 5]. It is defined by

\[ D(\rho, \sigma) = \frac{1}{2} \text{Tr}[\rho - \sigma] = \frac{1}{2} \|\rho - \sigma\|_1, \]

(1)

where \( \|M\|_1 = \text{Tr}\sqrt{M^\dagger M} \) is called the 1-norm or trace norm. Computing the trace distance may, in general, involve diagonalizing the matrix \( \Delta = \rho - \sigma \), and hence can be computationally difficult.

An alternative metric is the Hilbert-Schmidt distance:

\[ D_{\text{HS}}(\rho, \sigma) = \text{Tr}[\rho - \sigma]^2 = \|\rho - \sigma\|_2^2, \]

(2)

where \( \|M\|_2 = \sqrt{\text{Tr}(M^\dagger M)} \) is called the 2-norm or Frobenius norm. While the Hilbert-Schmidt distance does not share the operational relevance of the trace distance [6], it has the benefit that one can compute it without doing matrix diagonalization, and furthermore it is known to be efficiently computable on a quantum computer [7, 8]. Because of the latter, the Hilbert-Schmidt distance is employed as a cost function in recent variational hybrid quantum-classical algorithms [9, 10].

In many applications, one is interested in upper-bounding the trace distance between two states, i.e., showing that \( D(\rho, \sigma) \leq \epsilon \) for some small number \( \epsilon \). Hence, in this work, we explore upper bounds on \( D(\rho, \sigma) \), and in particular, upper bounds formulated in terms of \( D_{\text{HS}}(\rho, \sigma) \), since the latter may be computable.

For example, the norm equivalence [11] of the trace norm and Frobenius norm can be written as

\[ \|M\|_2 \leq \|M\|_1 \leq \sqrt{\text{rank}(M)} \|M\|_2. \]

(3)

Applying this to \( M = \Delta \) gives

\[ (1/4)D_{\text{HS}}(\rho, \sigma) \leq D(\rho, \sigma)^2 \leq (d/4) \cdot D_{\text{HS}}(\rho, \sigma), \]

(4)

where \( d \) is the Hilbert space dimension, and we used the inequality \( \text{rank}(\Delta) \leq d \), which is the best one can do without any prior knowledge about \( \Delta \). As \( d \) grows exponentially in the number of quantum subsystems, \( d \) can be very large, making the upper bound in (4) very weak. Hence it is natural to ask whether the upper bound in (4) can be tightened for certain states \( \rho \) and \( \sigma \).

The goal of this article is to explore possible tightenings of this upper bound when \( \rho \) and/or \( \sigma \) are low-rank states or low-entropy states. Such tightening would be expected since the extreme case of pure states, \( \rho = \ket{\psi}\bra{\psi} \) and \( \sigma = \ket{\phi}\bra{\phi} \), gives \( D(\rho, \sigma)^2 = 1 - |\bra{\psi}\phi\rangle|^2 \) and \( D_{\text{HS}}(\rho, \sigma) = 2 - 2|\bra{\psi}\phi\rangle|^2 \). Hence for pure states we have

\[ D(\rho, \sigma)^2 = (1/2)D_{\text{HS}}(\rho, \sigma), \]

(5)

and no dimension-dependent factor appears here.

In what follows, we first discuss tightenings of (4) for low-rank states, including our main result of an upper bound that is stronger than what one could obtain via norm equivalence. Then we explore some bounds for the more general scenario of low-entropy states.

II. RANK-BASED BOUNDS

Consider the following useful lemma.

Lemma 1. Let \( \Delta = \rho - \sigma = \Delta_+ - \Delta_- \). Here \( \Delta_+ \) and \( \Delta_- \) respectively correspond to the positive and negative part of \( \Delta \), with \( \Delta_+ \geq 0 \), \( \Delta_- \geq 0 \), and \( \Delta_+ \Delta_- = 0 \). Then we have:

\[ \text{rank}(\Delta_+) \leq \text{rank}(\rho), \]

(6)

\[ \text{rank}(\Delta_-) \leq \text{rank}(\sigma). \]

(7)

Proof. Let \( \{r_j\}, \{s_j\} \), and \( \{\delta_j\} \) respectively denote the eigenvalues of \( \rho \), \( \sigma \), and \( \Delta \), where the eigenvalues in each set are ordered in decreasing order. Weyl’s inequality [12] applied to \( \rho = \Delta + \sigma \) gives

\[ r_j \geq \delta_j + s_d, \quad \forall j. \]

(8)

Because \( \sigma \geq 0 \), we have \( s_d \geq 0 \), and hence

\[ r_j \geq \delta_j, \quad \forall j. \]

(9)

Since the eigenvalues of \( \rho \) are bigger than the eigenvalues of \( \Delta \), then when \( r_j = 0 \) we must have \( \delta_j \leq 0 \). This implies that \( \text{rank}(\rho) \geq \text{rank}(\Delta_+) \).
Similarly, we can define \( \Delta = \sigma - \rho = \Delta_+ - \Delta_- \), where \( \Delta_- \) corresponds to the positive part of \( \Delta \). Writing \( \sigma = \overline{\Delta} + \rho \), and applying Weyl’s inequality gives

\[
 s_j \geq \overline{\sigma}_j, \quad \forall j, 
\]

where \( \{\overline{\sigma}_j\} \) are the ordered eigenvalues of \( \overline{\Delta} \). Since the eigenvalues of \( \sigma \) are bigger than those of \( \overline{\Delta} \), then when \( s_j = 0 \) we must have \( \overline{\sigma}_j \leq 0 \), implying \( \operatorname{rank}(\sigma) \geq \operatorname{rank}(\Delta_-) \).

Combining this lemma with the norm equivalence in (3) gives the following result.

**Proposition 1.** For any two quantum states \( \rho \) and \( \sigma \),

\[
 D(\rho, \sigma)^2 \leq \left( \frac{\operatorname{rank}(\rho) + \operatorname{rank}(\sigma)}{4} \right) D_{\text{HS}}(\rho, \sigma). \tag{11} \]

**Proof.** From Lemma 1, we have

\[
 \operatorname{rank}(\Delta) = \operatorname{rank}(\Delta_+) + \operatorname{rank}(\Delta_-) \tag{12} 
\]

\[
 \leq \operatorname{rank}(\rho) + \operatorname{rank}(\sigma). \tag{13} 
\]

The upper bound in (3) then gives:

\[
 4D(\rho, \sigma)^2 \leq \operatorname{rank}(\Delta)D_{\text{HS}}(\rho, \sigma) \tag{14} 
\]

\[
 \leq (\operatorname{rank}(\rho) + \operatorname{rank}(\sigma))D_{\text{HS}}(\rho, \sigma). \tag{15} 
\]

(11) gets rid of the dimension-dependent factor in (4) and replaces it with a rank-based factor. This can potentially improve the bound, and indeed one can see from (5) that (11) is tight when both \( \rho \) and \( \sigma \) are pure.

On the other hand, there are many other cases where (11) is loose. For example, as we will soon see from our main result below, the following inequality holds when, say, \( \rho \) is pure but \( \sigma \) is an arbitrary state (i.e., when only one of the two states is pure):

\[
 D(\rho, \sigma)^2 \leq D_{\text{HS}}(\rho, \sigma). \tag{16} 
\]

In this special case, \( \operatorname{rank}(\sigma) \) can grow in proportion to the dimension \( d \) and hence the bound in (11) can be extremely loose.

The looseness of (11) motivates looking for a tightening of the bound that is not based on the norm equivalence in (3). Indeed, that is what we do in the following theorem, which is our main result.

**Theorem 1.** For any two quantum states \( \rho \) and \( \sigma \),

\[
 D(\rho, \sigma)^2 \leq R \cdot D_{\text{HS}}(\rho, \sigma), \tag{17} 
\]

where we refer to \( R \) as the reduced rank (defined analogously to the reduced mass in physics):

\[
 R = \frac{\operatorname{rank}(\rho)\operatorname{rank}(\sigma)}{\operatorname{rank}(\rho) + \operatorname{rank}(\sigma)}. \tag{18} 
\]

**Proof.** Let \( \tau_+ = \Delta_+/\operatorname{Tr}(\Delta_+) \) and \( \tau_- = \Delta_-/\operatorname{Tr}(\Delta_-) \). Note that both \( \tau_+ \) and \( \tau_- \) are density matrices (positive semidefinite with trace one). Since the purity of a density matrix is lower bounded by the inverse of its rank, we have

\[
 \operatorname{Tr}(\tau_+^2) \geq 1/\operatorname{rank}(\Delta_+) \geq 1/\operatorname{rank}(\rho), \tag{19} 
\]

\[
 \operatorname{Tr}(\tau_-^2) \geq 1/\operatorname{rank}(\Delta_-) \geq 1/\operatorname{rank}(\sigma), \tag{20} 
\]

where we have employed Lemma 1. Using the definitions of \( \tau_+ \) and \( \tau_- \), we have

\[
 \operatorname{Tr}(\Delta_+^2) \geq \operatorname{Tr}(\Delta_+)^2/\operatorname{rank}(\rho), \tag{21} 
\]

\[
 \operatorname{Tr}(\Delta_-^2) \geq \operatorname{Tr}(\Delta_-)^2/\operatorname{rank}(\sigma). \tag{22} 
\]

Summing these two inequalities gives:

\[
 \operatorname{Tr}(\Delta_+^2) + \operatorname{Tr}(\Delta_-^2) \geq \frac{\operatorname{Tr}(\Delta_+)^2}{\operatorname{rank}(\rho)} + \frac{\operatorname{Tr}(\Delta_-)^2}{\operatorname{rank}(\sigma)}. \tag{23} 
\]

The left-hand-side of this inequality is \( D_{\text{HS}}(\rho, \sigma) \), while \( \operatorname{Tr}(\Delta_+) = \operatorname{Tr}(\Delta_-) = D(\rho, \sigma) \). Hence we have

\[
 D_{\text{HS}}(\rho, \sigma) \geq D(\rho, \sigma)^2 \left( \frac{1}{\operatorname{rank}(\rho)} + \frac{1}{\operatorname{rank}(\sigma)} \right), \tag{24} 
\]

which is equivalent to (17).

Let us now show that (17) is stronger than (11). First note that

\[
 2\operatorname{rank}(\rho)\operatorname{rank}(\sigma) \leq \operatorname{rank}(\rho)^2 + \operatorname{rank}(\sigma)^2, \tag{25} 
\]

which implies that

\[
 4\operatorname{rank}(\rho)\operatorname{rank}(\sigma) \leq (\operatorname{rank}(\rho) + \operatorname{rank}(\sigma))^2. \tag{26} 
\]

Dividing through by \( 4(\operatorname{rank}(\rho) + \operatorname{rank}(\sigma)) \) gives

\[
 R \leq (\operatorname{rank}(\rho) + \operatorname{rank}(\sigma))/4, \tag{27} 
\]

which shows that (17) implies (11).

Next let us consider some examples. Consider the example noted above in (16), where \( \rho \) is pure. In this case, the reduced rank is

\[
 R = \frac{\operatorname{rank}(\sigma)}{1 + \operatorname{rank}(\sigma)} \leq 1, \tag{28} 
\]

and hence (17) implies (16). This is a dramatic improvement compared to the bound in (11).

Consider another example where \( \rho = \Pi/r \) is proportional to a rank \( r \) projector \( \Pi \) and \( \sigma = \mathbb{1}/d \) is maximally mixed. In this case,

\[
 D(\rho, \sigma) = \frac{d - r}{d}, \quad R = \frac{dr}{d + r}, \quad D_{\text{HS}}(\rho, \sigma) = \frac{d - r}{dr}. 
\]

This gives

\[
 \frac{D(\rho, \sigma)^2}{D_{\text{HS}}(\rho, \sigma)} = R \left( \frac{d^2 - r^2}{d^2} \right), \tag{29} 
\]
where the right-hand-side is approximately $R$ when $r \ll d$, and hence (17) becomes tight in this limit.

We remark that the reduced rank has the property

$$R \leq \min \{ \text{rank}(\rho), \text{rank}(\sigma) \},$$

(30)

and hence provides a stronger bound than simply taking the minimum of the two ranks.

### III. ENTROPY-BASED BOUNDS

The rank of a matrix is an abruptly changing function of its eigenvalues. So it makes sense to consider a smoother function such as entropy. While there are many such entropy functions, we consider the linear entropy

$$S_L(\rho) = 1 - \text{Tr}(\rho^2)$$

(31)

for two reasons: it is the natural entropy to associate with the Hilbert-Schmidt distance, and it is efficient to compute. This is in the spirit that we would like to upper-bound the trace distance with a quantity that is easily computable.

Consider the following bound.

**Proposition 2.** For any two quantum states $\rho$ and $\sigma$,

$$D(\rho, \sigma)^2 \leq \frac{1}{2}[D_{\text{HS}}(\rho, \sigma) + S_L(\rho) + S_L(\sigma)].$$

(32)

**Proof.** Let $\{\delta^+_j\}$ and $\{\delta^-_j\}$ respectively be the eigenvalues of $\Delta_+$ and $\Delta_-$. Since $D(\rho, \sigma) = \text{Tr}(\Delta_+) = \text{Tr}(\Delta_-)$, we can write

$$D(\rho, \sigma)^2 = \frac{1}{2}[\text{Tr}(\Delta_+)^2 + \text{Tr}(\Delta_-)^2]$$

(33)

$$= \frac{1}{2} \left[ \sum_{jk} \delta^+_j \delta^+_k + \sum_{jk} \delta^-_j \delta^-_k \right]$$

(34)

$$= \frac{1}{2} \left[ D_{\text{HS}}(\rho, \sigma) + \sum_{k \neq j} \delta^+_j \delta^+_k + \sum_{k \neq j} \delta^-_j \delta^-_k \right],$$

(35)

where we used the fact that $D_{\text{HS}}(\rho, \sigma) = \sum_j [\delta^+_j]^2 + [\delta^-_j]^2$. Next we use Weyl’s inequality (see proof of Lemma 1), which implies that:

$$r_j \geq \delta^+_j, \quad \forall j$$

(36)

$$s_j \geq \delta^-_j, \quad \forall j,$$

(37)

where we assume the eigenvalues are ordered in decreasing order. Since these eigenvalues are non-negative, we have

$$D(\rho, \sigma)^2 \leq \frac{1}{2} \left[ D_{\text{HS}}(\rho, \sigma) + \sum_{k \neq j} r_j r_k + \sum_{k \neq j} s_j s_k \right].$$

(38)

Finally, note that

$$S_L(\rho) = (\text{Tr}(\rho))^2 - \text{Tr}(\rho^2) = \sum_{k \neq j} r_j r_k,$$

and similarly $S_L(\sigma) = \sum_{k \neq j} s_j s_k$. Plugging these relations into (38) proves the result. \qed

Equation (32) is tight when both $\rho$ and $\sigma$ are pure, since $S_L(\rho) = S_L(\sigma) = 0$ in this case. On the other hand, when only $\rho$ is pure, the bound in (32) becomes $(1/2)[D_{\text{HS}}(\rho, \sigma) + S_L(\sigma)]$, which does not reduce to (16). In this case, (32) can be either stronger or weaker than (16).

The following is an alternative bound.

**Proposition 3.** For any two quantum states $\rho$ and $\sigma$,

$$D(\rho, \sigma)^2 \leq D_{\text{HS}}(\rho, \sigma) + \min \{ S_L(\rho), S_L(\sigma) \}.$$  \hspace{1cm} (40)

**Proof.** We will prove that

$$D(\rho, \sigma)^2 \leq D_{\text{HS}}(\rho, \sigma) + S_L(\rho).$$

(41)

Then by symmetry the roles of $\rho$ and $\sigma$ can be interchanged to give:

$$D(\rho, \sigma)^2 \leq D_{\text{HS}}(\rho, \sigma) + S_L(\sigma).$$

(42)

Taking the minimum of the two bounds in (41) and (42) then gives (40).

To prove (41), we follow a similar approach as the proof of Proposition 2. Namely, we write

$$D(\rho, \sigma)^2 = \text{Tr}(\Delta_+)^2$$

(43)

$$= \sum_{jk} \delta^+_j \delta^+_k$$

(44)

$$= \sum_{j} (\delta^+_j)^2 + \sum_{k \neq j} \delta^+_j \delta^+_k$$

(45)

$$\leq D_{\text{HS}}(\rho, \sigma) + \sum_{k \neq j} \delta^+_j \delta^+_k.$$  \hspace{1cm} (46)

Then from (36) we have

$$D(\rho, \sigma)^2 \leq D_{\text{HS}}(\rho, \sigma) + \sum_{k \neq j} r_j r_k,$$

(47)

and inserting (39) we obtain (41). \qed

Equation (40) does not have the factor of $(1/2)$ in front of $D_{\text{HS}}(\rho, \sigma)$ like (32) does. On the other hand, it does reduce to (16) when $\rho$ is pure.

There are cases in which (32) is stronger than (40), and there are cases where the reverse is true. Hence, in general, one can take the minimum of the two bounds from (32) and (40) for our strongest entropy-based bound.

We do remark on the following issue. Both (32) and (40) are additive bounds, where the entropy is added to the Hilbert-Schmidt distance. This is in contrast to our
rank-based bounds, which were multiplicative. Typically multiplicative bounds are more desirable than additive bounds, since one would like the bound to go to zero when the Hilbert-Schmidt distance goes to zero. For this reason, we propose that future work is needed to search for multiplicative entropy-based bounds, where the entropy term multiplies the Hilbert-Schmidt distance. The existence of such bounds remains an interesting open question [13].

IV. CONCLUSIONS

Low-rank and low-entropy quantum states show up naturally in, for example, condensed matter physics, where bipartite cuts of condensed matter ground states lead to weakly entangled subsystems [14]. Low-rank states also appear in data science, where the covariance matrix has only a small number of principal components due to redundant features [15, 16]. These applications motivate the work in this article, where we gave improved upper bounds on the trace distance for low-rank or low-entropy states. We focused on upper bounds involving the Hilbert-Schmidt distance, since this quantity can be efficiently computed on a quantum computer. Hence, our bounds may benefit the field of quantum algorithms, by making it easier for quantum algorithms to provide a tight bound on the operationally relevant trace distance.

Our main result, Theorem 1, gave a bound involving the reduced rank (a quantity analogous to the reduced mass in physics). This bound was stronger than what one could obtain directly from the equivalence of the Frobenius and trace norms. We also gave bounds involving the linear entropy in Propositions 2 and 3, where the linear entropy appeared as an additive term in the bound. An open question is whether multiplicative entropy-based bounds exist, and we believe this an important direction for future work.

V. ACKNOWLEDGEMENTS

We thank Rolando Somma and Marco Tomamichel for helpful conversations. All authors acknowledge support from the LDRD program at Los Alamos National Laboratory (LANL). PJC additionally acknowledges support from the LANL ASC Beyond Moore’s Law project. LC was also supported by the DOE through the J. Robert Oppenheimer fellowship.

[1] M. Ben-Or, M. Horodecki, D. W. Leung, D. Mayers, and J. Oppenheim, “The universal composable security of quantum key distribution,” in Theory of Cryptography, edited by J. Kilian (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005) pp. 386–406.
[2] R. Renner and R. König, “Universally composable privacy amplification against quantum adversaries,” in Theory of Cryptography, edited by J. Kilian (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005) pp. 407–425.
[3] Carl W. Helstrom, “Quantum detection and estimation theory,” J. Stat. Phys. 1, 231–252 (1969).
[4] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press (2010).
[5] M. M. Wilde, Quantum Information Theory, 2nd ed. (Cambridge University Press, 2017).
[6] M. Ozawa, “Entanglement measures and the Hilbert–Schmidt distance,” Phys. Lett. A 268, 158–160 (2000).
[7] J. C. García-Escartin and P. Chamorro-Posada, “Swap test and Hong-Ou-Mandel effect are equivalent,” Phys. Rev. A 87, 052330 (2013).
[8] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, “Learning the quantum algorithm for state overlap,” New J. Phys. 20, 113022 (2018).
[9] R. LaRose, A. Tikku, É. O’Neel-Judy, L. Cincio, and P. J. Coles, “Variational Quantum State Diagonalization,” arXiv:1810.10506 (2018).
[10] A. Arrasmith, L. Cincio, A. T. Sornborger, W. H. Zurek, and P. J. Coles, “Variational Consistent Histories: A Hybrid Algorithm for Quantum Foundations,” arXiv:1812.10759 (2018).
[11] R. A. Horn and C. R. Johnson, Matrix analysis (Cambridge University Press, 1990).
[12] H. Weyl, “Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohiraustrahlung),” Mathematische Annalen 71, 441–479 (1912).
[13] For example, the following bound would be natural conjecture: $D(\rho, \sigma)^2 \leq \frac{D_{\text{HS}}(\rho, \sigma)}{\text{Tr}(\rho^2)}$. However, we numerically found this bound to be violated.
[14] H. Li and F. D. M. Haldane, “Entanglement Spectrum as a Generalization of Entanglement Entropy: Identification of Topological Order in Non-Abelian Fractional Quantum Hall Effect States,” Phys. Rev. Lett. 101, 010504 (2008).
[15] C. M. Bishop, Pattern recognition and machine learning (Springer, 2006).
[16] K. P. Murphy, Machine learning: a probabilistic perspective (MIT press, 2012).