High-expressibility Quantum Neural Networks using only classical resources

Marco Maronese[1], Francesco Ferrari[1], Matteo Vandelli[1], and Daniele Dragoni[1,2]

[1]Quantum Computing Solutions, Leonardo S.p.A., Via R. Pieragostini 80, Genova, Italy
[2]Hyper-Computing Continuum Unit, Leonardo S.p.A., Via R. Pieragostini 80, Genova, Italy

Quantum neural networks (QNNs), as currently formulated, are near-term quantum machine learning architectures that leverage parameterized quantum circuits with the aim of improving upon the performance of their classical counterparts. In this work, we show that some desired properties attributed to these models can be efficiently reproduced without necessarily resorting to quantum hardware. We indeed study the expressibility of parametrized quantum circuit commonly used in QNN applications and contrast it to those of two classes of states that can be efficiently simulated classically: matrix-product states (MPS), and Clifford-enhanced MPS (CMPS), obtained by applying a set of Clifford gates to MPS. In addition to expressibility, we assess the level of primary quantum resources, entanglement and non-stabilizerness (a.k.a. "magic"), in random ensembles of such quantum states, tracking their convergence to the Haar distribution. While MPS require a large number of parameters to reproduce an arbitrary quantum state, we find that CMPS approach the Haar distribution more rapidly, in terms of both entanglement and magic. Our results indicate that high expressibility in QNNs is attainable with purely classical resources.

Quantum computation leverages the capability of qubits to encode exponentially more information than the same number of classical bits [1]. This property motivated the development of quantum machine learning (QML) models [2, 3, 4], which exploit the exponentially large Hilbert space of qubits as effective latent space [5, 6, 7]. Specifically, Quantum Neural Networks (QNNs) are learning models that mimic the functionality of classical neural networks within a quantum computational framework [8, 9, 10, 11, 12]. QNNs are typically composed of three stages: embedding of the input data in the Hilbert space of the qubits; processing of the quantum state by a parametrized quantum circuit (PQC), namely a unitary operator $U(\boldsymbol{\theta})$; finally, information extraction by measurements of certain observables of choice [13]. QNN models have been proposed as candidates to enhance expressivity compared to classical neural networks, e.g. for classification or generative tasks [14, 15].

In analogy to the universal approximation theorem of classical statistical learning theory [16, 17, 18], a desirable property of QNN architectures is the ability to approximate a wide range of functions [19, 20]. This property is closely related to the *expressibility* of the underlying PQC, which is the capability to uniformly represent any pure quantum state in the Hilbert space [21, 22]. The latter task requires a number of parameters that grows exponentially with the number of qubits [1, 23]. However, in practical applications, PQCs should be designed to allow efficient implementation on actual quantum hardware, using a manageable number of one- and two-qubit operations. This raises the question of which classes of PQCs are best suited to effectively span the Hilbert space.

Previous works have explored the connection between expressibility of representative PQC architectures and their degree of quantumness, as measured for instance by entanglement [21, 24]. Recently, a second measure of non-classicality of quantum states has emerged: non-stabilizerness, also known as *magic*, which quantifies the amount of non-Clifford gates involved in a PQC [25, 1, 26]. Both of these resources, entanglement and magic, are deeply linked to the ability to efficiently emulate quantum states on a classical computer. Specifically, states with low entanglement can be efficiently simulated using tensor-

network (TN) methods, e.g. by matrix-product states (MPS) [27], regardless of their level of magic (see Fig. 1). Conversely, zero-magic states, commonly known as stabilizer states, can be simulated in polynomial time by leveraging the Gottesman-Knill theorem independently of their entanglement [25, 28]. Given that randomly sampled states typically exhibit large values of magic and entanglement [29, 30], it is an open question whether expressibility is ultimately rooted in both these non-classical resources. To this end, we characterize the quantum resource content and the expressibility of three different classes of quantum states. On the one hand, we consider states obtained by a prototypical quantum circuit for QML applications that is classically hard to emulate but runs on quantum hardware with polynomial resources. We refer to this architecture as fully-quantum neural network (fQNN). On the other hand, we consider quantum states that can be efficiently simulated classically: MPS and Clifford-enhanced matrix-product states (CMPS). The latter are obtained by applying a Clifford circuit to an MPS and exhibit large entanglement along with substantial magic. Expectation values over these states can be efficiently computed on a classical computer [31]. We generate random samples from the different classes of states and analyze the properties of these ensembles compared to uniformly distributed states as drawn from the Haar measure [32]. Our findings demonstrate that, using only polynomial classical resources, CMPS are able to compete with quantum native architectures like fQNN in terms of expressibility and quantum resources.

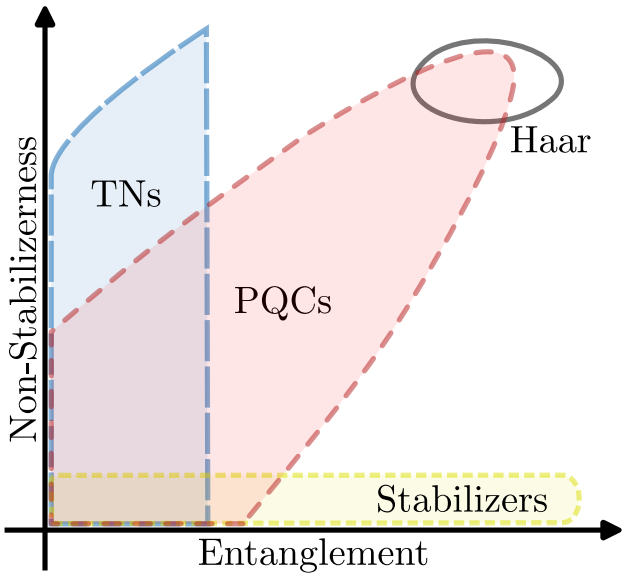

Figure 1: Qualitative representation of the phase space defined by non-stabilizerness (magic) and entanglement, illustrating the regions typically explored by different classes of wave function Ansätze. Tensor networks (TNs), shown in blue, can access areas with relatively low entanglement, but potentially high magic [33]. Parameterized quantum circuits (PQCs), depicted in red, generally have a layered structure that allows them to simultaneously increase both magic and entanglement, when the circuit is made deeper. They tend to reach the quantum resources of the Haar distribution (region bounded in gray) more efficiently than TNs. Stabilizer states, shown in yellow, have zero magic.

1 Theory

1.1 Expressibility of quantum neural networks

QNNs can be defined as parametrized learning models in the form:

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = \mathrm{Tr}\left[OU(\boldsymbol{\theta})\overline{\rho}(\boldsymbol{x})U^{\dagger}(\boldsymbol{\theta})\right], \qquad (1)$$

where $\boldsymbol{x}$ indicates the input data and $\boldsymbol{\theta}$ are the model parameters. Such a model is to be interpreted as a composition of operations:

1. *Embedding layer* (often referred to as feature map): the input $\boldsymbol{x}$ is encoded into the quantum state $|\varphi(\boldsymbol{x})\rangle$ in a n-qubit Hilbert space $\mathcal{H}_n$ with dimension $d = 2^n$ [5]. Its density matrix is $\overline{\rho}(\boldsymbol{x}) = |\varphi(\boldsymbol{x})\rangle\langle\varphi(\boldsymbol{x})|$;

2. *Parametrized unitary*: a unitary operator $U(\boldsymbol{\theta})$ acting on the states in $\mathcal{H}_n$, where $\boldsymbol{\theta}$ is a vector of $\mathcal{P}$ real parameters. We denote by $\Omega \in R^{\mathcal{P}}$ the domain of the parameters.

3. *Activation function*: a measurement-induced activation function $\rho \mapsto \mathrm{Tr}\left[O\rho\right]$ that extracts information from the Hilbert space by measuring an observable O on the state with density-matrix ρ.

The learning performance of the QNN model naturally depends on all three stages. However, for a fixed embedding layer, the capability of the QNN to effectively explore the quantum Hilbert space depends on the form of the parametrized unitary and is tied to the notion of *expressibility*, which quantifies how well $U(\boldsymbol{\theta})$ can generate a distribution of states which are uniformly distributed across the Hilbert space.

To rigorously define this concept, we utilize the machinery of the Haar measure [32], which identifies the unique left- and right-invariant probability measure on the group of unitary operators $\mathrm{U}(d)$. Applying unitary operators drawn from the Haar measure to a certain reference state[1], we obtain the uniform distribution of quantum states across the Hilbert space, henceforth denoted as μ_{H}. Similarly, we introduce the distributions of states μ obtained by the parametrized unitary operator of a chosen QNN model, namely the distribution of states $|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})\,|0\rangle^{\otimes n}$ when varying the parameters $\boldsymbol{\theta}$. The ability of the distribution μ to approximate μ_{H} quantifies the expressibility of $U(\boldsymbol{\theta})$. We aim at quantifying the distance between μ and μ_{H} by examining the discrepancy between the moments of the two distributions.

To this aim, we introduce the quantity

$$A_t = \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu}\left[|\psi\rangle\langle\psi|^{\otimes t}\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu_{\mathrm{H}}}\left[|\psi\rangle\langle\psi|^{\otimes t}\right], \quad (2)$$

where $\mathbb{E}[\cdot]$ denotes the expectation value over a distribution of quantum states. A_t quantifies the difference between the t-th moment of μ and μ_{H}. The distribution μ is said to form a t-state design if the statistical moments up to order t match those of the uniform distribution μ_{H}, i.e. if $A_t = \mathbf{0}$ [21]. In practical terms, parametrized states $|\psi(\boldsymbol{\theta})\rangle$ are generated by sampling the parameters $\boldsymbol{\theta}$ from a distribution η_θ defined over the parameter space Ω. The expression in Eq. (2) for the matrix A_t effectively translates to

$$A_t = \mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\eta_\theta}\left[|\psi(\boldsymbol{\theta})\rangle\langle\psi(\boldsymbol{\theta})|^{\otimes t}\right] - \mathop{\mathbb{E}}_{|\psi\rangle\sim\mu_{\mathrm{H}}}\left[|\psi\rangle\langle\psi|^{\otimes t}\right]. \quad (3)$$

where the expectation value $\mathbb{E}_{\boldsymbol{\theta}\sim\eta_\theta}[\cdot]$ is over the distributions of parameters η_θ defining the quantum state $|\psi(\boldsymbol{\theta})\rangle$. To represent the difference in Eq. (3) by a single scalar value, we evaluate the Hilbert-Schmidt norm of A_t

$$\|A_t\|_{HS}^2 := \mathrm{Tr}\left[A_t^\dagger A_t\right] = \mathcal{F}_\mu^{(t)} - \mathcal{F}_{\mathrm{H}}^{(t)}, \quad (4)$$

where $\mathcal{F}_\mu^{(t)}$ is the so-called t-frame potential of the μ distribution and is defined as

$$\mathcal{F}_\mu^{(t)} := \mathop{\mathbb{E}}_{\substack{\boldsymbol{\theta}\sim\eta_\theta \\ \boldsymbol{\phi}\sim\eta_\phi}}\left[|\langle\psi(\boldsymbol{\theta})|\psi(\boldsymbol{\phi})\rangle|^{2t}\right]. \quad (5)$$

[1]The choice of the reference state is arbitrary and does not lead to a loss of generality for the present discussion.

The derivation of Eq. (4) from Eq. (3) is shown in Ref. [32]. The frame potentials for the Haar distributions are analytically known $\mathcal{F}_{\mathrm{H}}^{(t)} = \frac{t!(d-1)!}{(d-1+t)!}$. Frame potentials can be interpreted as the moments of the distribution of fidelities $P_\mu(F = |\langle\psi(\boldsymbol{\theta})|\psi(\boldsymbol{\phi})\rangle|^2)$ between pairs of states extracted from the same distribution μ. Indeed, $\mathcal{F}_\mu^{(t)} = \mathbb{E}_{F\sim P_\mu}[F^t]$. For what concerns the Haar distribution, the analytical expression of the distribution of fidelities is known to be $P_{\mathrm{H}}(F) = (d-1)(1-F)^{d-2}$ [32].

We note that, since $\|A\|_{HS}^2$ is positive semidefinite, the condition $\mathcal{F}_\mu^{(t)} \geq \mathcal{F}_{\mathrm{H}}^{(t)}$ strictly holds for any distribution of states μ. For this reason, $\mathcal{F}_{\mathrm{H}}^{(t)}$ is a lower bound (called the *Welch bound*). If the equality $\mathcal{F}_\mu^{(t)} = \mathcal{F}_{\mathrm{H}}^{(t)}$ holds for each $t^* \leq t$, then μ is a t-state design. The equality holds for all t if and only if μ is the Haar distribution. In this framework, we can estimate expressibility by computing the first t frame potentials, and thus determining if the distribution μ is a t-state design. In circuit-based quantum computation models, the question has been raised whether it is possible to approximately saturate the Welch bound using PQCs with a polynomial gate scaling $O(\mathrm{poly}(n))$. To approximate a t-design, the depth of a PQC must grow with both the number of qubits n and the order t as $O(\mathrm{poly}(n,t))$ [34]. For many practical applications, logarithmic or polynomial scaling in n is sufficient for low-order designs ($t \leq 3$).

A more holistic approach to quantifying the mismatch between μ and the Haar distribution involves directly estimating the distance between the corresponding fidelity distributions $P_\mu(F)$ and $P_{\mathrm{H}}(F)$. One suitable metric for this purpose is the Kullback–Leibler divergence, which provides a quantitative measure of the difference between these distributions

$$D_{\mathrm{KL}}(P_\mu \parallel P_{\mathrm{H}}) = \sum_F P_\mu(F)\log\left(\frac{P_\mu(F)}{P_{\mathrm{H}}(F)}\right). \quad (6)$$

The smaller the value of D_{KL}, the closer the distribution of fidelities between states sampled by the distribution μ is to that of the Haar-random states, indicating higher expressibility [21].

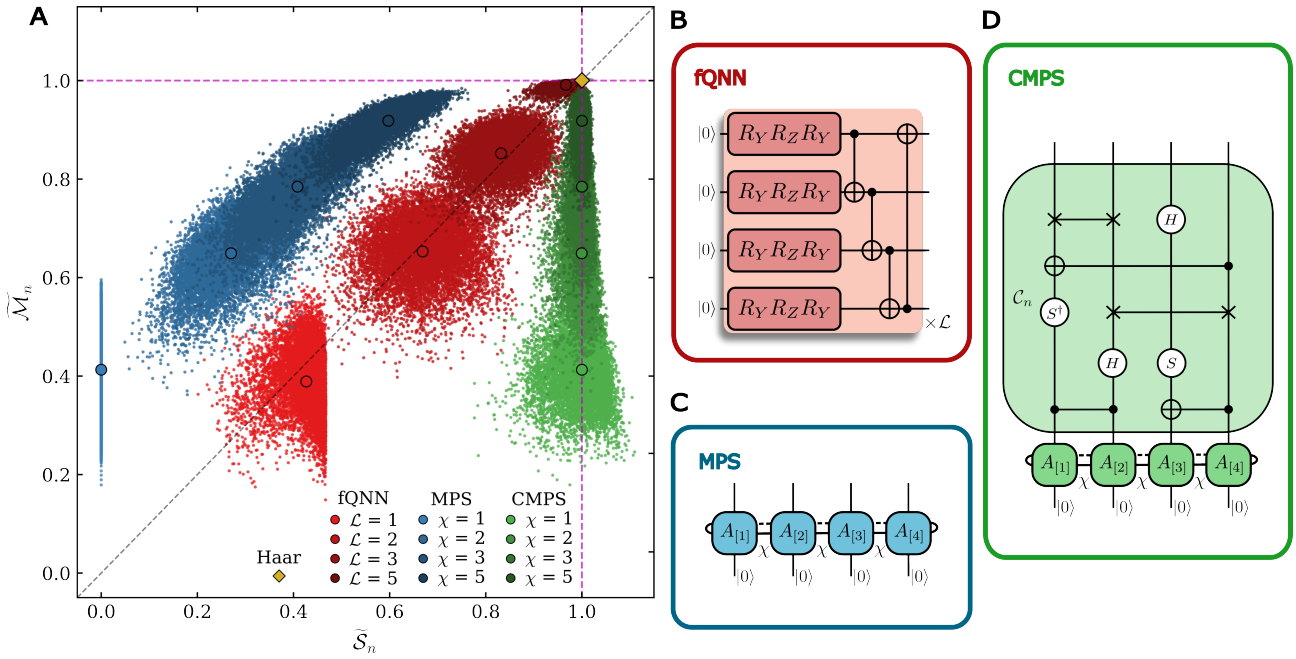

Figure 2: **A** Quantum states samples in the entanglement-magic phase space $\widetilde{\mathcal{S}} - \widetilde{\mathcal{M}}$. The results are obtained for $n = 10$ qubits. States sampled from the Haar measure form an extremely localized distribution [29] centered at $(1, 1)$ (marked by the gold diamond). The other colors correspond to the various architectures shown in **B**,**C**,**D**. Red colors indicate fQNN states (**B**) with different number of layers, blue colors indicate MPS states (**C**) with different bond dimension, while green colors denote CMPS (**D**) results. Empty circles correspond to the average values of $\widetilde{\mathcal{S}}$ and $\widetilde{\mathcal{M}}$ for the various sets of samples.

1.2 Quantum resources: Entanglement and magic

Alongside expressibility, we seek to characterize the non-classicality inherent in the distribution of parameterized quantum states μ. To this end, we quantitatively evaluate two fundamental quantum resources: entanglement and magic.

To quantify the degree of entanglement in a quantum state, we use the (bipartite) entanglement entropy. This is defined by first partitioning the state into two subsystems. To do this, we label the qubits with sequential integer indices $l = 1, \ldots, n$. We then consider the various bipartitions $\{A_l, B_l\}_{l=1,\ldots,n-1}$ that cut the system between the l-th and $(l+1)$-th qubit, defined by $A_l = [1{:}l]$ and $B_l = [l+1{:}n]$. The entanglement entropy is taken to be the maximum entropy over all such bipartitions [24]

$$\mathcal{S}_n = \max_{\{A_l, B_l\}} \mathrm{Tr}_{A_l}[\rho_{A_l} \log_2 \rho_{A_l}], \tag{7}$$

where $\rho_{A_l} = \mathrm{Tr}_{B_l}[\rho(\boldsymbol{\theta})]$ with $\rho(\boldsymbol{\theta}) = |\psi(\boldsymbol{\theta})\rangle\langle\psi(\boldsymbol{\theta})|$.

Turning to magic, we utilize the stabilizer Rényi entropy as a measure of non-stabilizerness of a given quantum state [35]. To introduce this quantity, we briefly review the theory of stabilizers [36, 25]. We begin by defining the set of Pauli strings for a system of n qubits as

$$P_n = \{\zeta\, \sigma_1 \otimes \sigma_2 \otimes \cdots \otimes \sigma_n\}. \tag{8}$$

with global phase $\zeta \in \{\pm 1, \pm i\}$ and Pauli matrices $\sigma_l \in \{\mathbb{1}, X, Y, Z\}$. A quantum state $|\psi\rangle$ is called a *stabilizer state* if there exist an Abelian subgroup $\Sigma \subset P_n$ containing d Pauli strings such that $s\,|\psi\rangle = |\psi\rangle\,,\ \forall s \in \Sigma$. Stabilizer states are transformed into other stabilizer states under the application of a special class of unitary operators, forming the so-called Clifford group, which is the normalizer group of P_n. Clifford operators are those that can be constructed by a combination of only Hadamard, S and CNOT gates. In this respect, any stabilizer state can be obtained by applying a certain Clifford operator to the reference state $|0\rangle^{\otimes n}$. Following the Gottesman–Knill theorem, any quantum circuit composed solely of stabilizer state preparations, Clifford gates and Pauli measurements can be simulated in polynomial time on a classical computer [25, 28]. For this reason, quantifying the amount of non-stabilizerness (i.e., magic) of a quantum state is considered a way to assess its degree of non-classicality. In-

deed, stabilizer states and Clifford gates are not sufficient for universal quantum computation [1].

Among the different metrics for magic [26, 37, 38, 39, 40, 41, 7], we select the stabilizer 2-Rényi entropy [35, 42]

$$\mathcal{M}_n = -\log_2 \sum_{P\in P_n} \Xi_P^2 - n \tag{9}$$

where $\Xi_P = d^{-1}|\langle\psi(\boldsymbol{\theta})|P|\psi(\boldsymbol{\theta})\rangle|^2$. This quantity is zero if and only if the quantum state is a stabilizer and is additive with respect to the tensor product of states. Most importantly, it is invariant under the application of Clifford operators to the state $|\psi(\boldsymbol{\theta})\rangle$.

To simplify the following discussion, we introduce normalized entanglement and magic entropy with respect to the values obtained for the Haar distribution, i.e.

$$\widetilde{\mathcal{M}}_n = \frac{\mathcal{M}_n}{\mathop{\mathbb{E}}_{|\psi\rangle\sim\mu_{\mathrm{H}}}[\mathcal{M}_n]}, \quad \widetilde{\mathcal{S}}_n = \frac{\mathcal{S}_n}{\mathop{\mathbb{E}}_{|\psi\rangle\sim\mu_{\mathrm{H}}}[\mathcal{S}_n]} \tag{10}$$

As the number of qubits increases, the entanglement and non-stabilizerness of a Haar-distributed state ensemble quickly become extremely localized around the asymptotic values $\mathcal{M}_n = n-2$ and $\mathcal{S}_n = n/2$, respectively [29, 30, 43]. The intensive quantities $\widetilde{\mathcal{M}}_n$ and $\widetilde{\mathcal{S}}_n$ indicate how close magic and entanglement of a given circuit are to the asymptotic values of the Haar ensemble.

1.3 Classically-hard *vs* classically simulatable QNN models

In this work, we consider three different types of states, which can be used as QNN architectures.

fQNN: a widely used PQC for QML. The name stands for fully-quantum neural network. The network is characterized by a number $\mathcal{L}$ of layers. Each layer is defined by generic $SU(2)$ rotations applied independently to each qubit and decomposed using the Euler angles decomposition $R_Y(\alpha)R_Z(\beta)R_Y(\gamma)$. These operations are followed by CNOT gates arranged in a ring topology, as sketched in Fig. 2**B**. These states cannot be represented efficiently on a classical computer, but can be executed on quantum hardware with polynomial resources. An fQNN with $\mathcal{L}$ layers is specified by a total of $\mathcal{P} = 3n\mathcal{L}$ angles as parameters. We generate random fQNN by indepently sampling each of these angles from the uniform distribution on the interval $[0, 2\pi)$.

MPS: matrix-product states with periodic boundary conditions defined as

$$|\mathrm{MPS}\rangle = \sum_{\{\mathbf{s}\}} \mathrm{Tr}\left[A^{s_1}_{[1]} A^{s_2}_{[2]} \dots A^{s_n}_{[n]}\right] |s_1 \dots s_n\rangle\,, \tag{11}$$

where, for fixed $s_i \in \{0,1\}$ (physical dimension), $A^{s_i}_{[i]}$ are $\chi\times\chi$ complex-valued matrices (as schematically illustrated in Fig. 2**C**). The hyperparameter χ is conventionally called bond dimension. In principle, MPS can span the full Hilbert space $\mathcal{H}$ if equipped with an exponentially large bond dimension, i.e. if $\chi \sim 2^{n/2}$ [44]. However, for practical application, polynomial bond dimensions $\chi = \mathrm{poly}(n)$ are employed. In this case, MPS can be efficiently represented and manipulated on a classical computer, allowing for tractable computation of expectation values of local observables, by means of tensor contraction techniques [27, 45]. In general, MPS provide an efficient representation for quantum states with limited (e.g., area-law) entanglement [46]. In the special case $\chi = 1$, $|\mathrm{MPS}\rangle$ reduces to a simple product state, meaning it has no entanglement. The MPS ansätze of Eq. (11) requires a total of $\mathcal{P} = 4n\chi^2$ real parameters, coming from its $2n$ $\chi\times\chi$ complex-valued matrices. We generate random MPS by independently sampling real and imaginary parts of each entry in the $A^{s_i}_{[i]}$ matrices from a normal distribution $\mathcal{N}(0,1)$ [47].

CMPS: states obtained by applying a Clifford operator to an MPS state (as exemplified in Fig. 2**D**) [48, 49, 50, 51]. Explicitly, we can write

$$|\mathrm{CMPS}\rangle = U_C|\mathrm{MPS}\rangle. \tag{12}$$

Although the direct representation of these states has an exponential footprint on classical memory, we can efficiently compute expectation values of observables [31] required in QNN applications as

$$\begin{aligned}\langle\mathrm{CMPS}|O|\mathrm{CMPS}\rangle &= \\ &= \langle\mathrm{MPS}|\left[U_C^\dagger\, O\, U_C\right]|\mathrm{MPS}\rangle.\end{aligned} \tag{13}$$

Indeed, the observable O is typically a linear combination of a polynomial number of Pauli strings. Therefore, the action of $\mathcal{U}_C$ on O can be obtained by simple swap rules for Pauli operators under the application of Clifford circuits, which can be efficiently performed classically. The resulting expectation value of the transformed operator $U_C^\dagger\, O\, U_C$ on the MPS part of the state can

be computed by standard tensor contraction operations. Random sampling of CMPS states requires specifying $\mathcal{P} = 4n\chi^2$ continuous parameters, which fully characterize the MPS part and are sampled as previously described. Additionally, the Clifford unitary U_C needs to be chosen. Since the Clifford group for n qubits is finite (and thus countable), each Clifford unitary can be uniquely defined by its action on the generators of the Pauli group, or equivalently, by a binary symplectic matrix, up to a global phase [52]. In this framework, random Clifford circuits can be sampled by adopting the method of Ref. [53]. On the contrary, we note that sticking to a single fixed Clifford unitary $\overline{U}_C$ is not suitable for our analysis, as it would yield the same values of $\mathcal{F}^{(t)}$ and D_{KL} as in the MPS case, as can be easily verified by substitution in Eqs. (5) and (6).

2 Computational details

Our calculations are based on the Pennylane Python software library for quantum computing [54], interfaced with JAX [55]. Random Clifford unitaries for the CMPS ensembles are generated using the `Tableau.random` subroutine of the Stim library [56], which implements the method of Ref. [53], and the Cirq library [57]. For the calculations presented in this work, we used a single node of our proprietary *davinci-1* cluster equipped with 2 AMD EPYC Rome 7402 @ 2.80 GHz CPUs (24 cores each), 512 GB of RAM and 4 NVIDIA A100 GPUs with 40 GB of dedicated memory.

3 Results

3.1 Magic *vs* Entanglement

We begin by assessing the quantum resource content of the three different kinds of states ensembles $\mu = \{\mathrm{fQNN}, \mathrm{MPS}, \mathrm{CMPS}\}$, to understand where these are located in the entanglement-magic phase space. The results are shown in Fig. 2**A**, for a system of $n = 10$ qubits. Specifically, we present a scatter plot illustrating the distribution of 10,000 sample instances (for each circuit type) in the normalized entanglement and magic plane $\widetilde{\mathcal{S}} - \widetilde{\mathcal{M}}$, for different number of layers $\mathcal{L}$ (for fQNN) and different bond dimensions χ (for MPS, CMPS). We note that CMPS samples are obtained by taking the MPS states from the MPS samples and applying a random Clifford operator to them. Therefore, as expected, the magic content of the CMPS samples is equal to that of the MPS samples, due to the invariance of $\mathcal{M}_n$ under the application of a Clifford operator. As previously shown [29], Haar distributed states tend to form an extremely localized distribution in the entanglement-magic plane, which in our plot is centered around the saturation values $\widetilde{\mathcal{M}}_{\mathrm{H}} = 1$ and $\widetilde{\mathcal{S}}_{\mathrm{H}} = 1$. The distribution of resources for all the μ state ensembles considered here localize more and more towards those of the Haar distribution, as their number of parameters $\mathcal{P}$ increases. Magic and entanglement of the different classes of states converge towards the values of the Haar distribution in different ways. For the fQNN case, we note that the distribution with $\mathcal{L} = 1$ exhibits a sharp bound in entanglement because the circuit only has a single layer of CNOT gates. For $\mathcal{L} \geq 2$, instead, the convergence is roughly balanced in both resources, and the samples align along the diagonal of the plot. For what concerns MPS, $\chi = 1$ corresponds to the case of simple product states, which have exactly zero entanglement. MPS samples approach the Haar limit from the upper-left direction of the $\widetilde{\mathcal{S}} - \widetilde{\mathcal{M}}$ plane, while CMPS states lie in the right-most side of the plane.

To better illustrate the convergence of quantum resources distributions, in Fig. 3**A** we provide a detailed view of how their average values approach the asymptotic values of the Haar samples, for different number of qubits n. The average magic $\widetilde{\mathcal{M}}$ is shown as golden triangles and the entanglement $\widetilde{\mathcal{S}}$ as violet circles for each state type μ. The fQNN resources are plotted in terms of the normalized number of layers $\mathcal{L}/n$, while the MPS and CMPS results are shown as a function of the normalized bond dimension χ/n. With this choice, the results are independent on n and the data points of different system sizes fall on the same curve.

We observe that fQNN states converge in both resources at the same rate with increasing layers, as the points for $\widetilde{\mathcal{M}}_{\mathrm{fQNN}}$ and $\widetilde{\mathcal{S}}_{\mathrm{fQNN}}$ overlap when plotted against $\mathcal{L}/n$ (see Fig. 3**A**, left-most panel). On the other hand, the MPS states saturate rather quickly in terms of magic but their entanglement converges slowly as the number of parameters increases, compared to the other states.

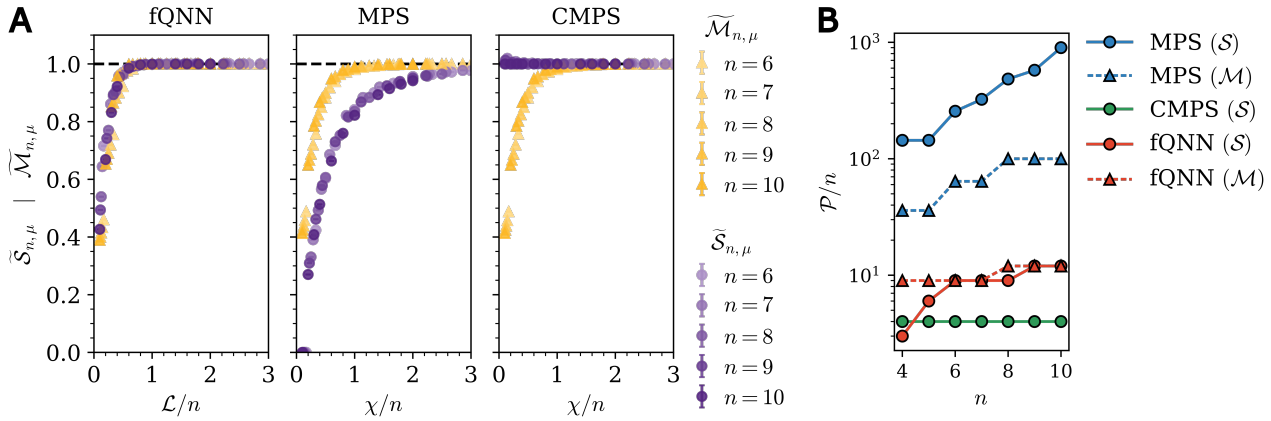

Figure 3: **A** Average values of $\widetilde{\mathcal{M}}$ (golden triangles) and $\widetilde{\mathcal{S}}$ (violet circles) for fQNN, MPS and CMPS ensembles, as a function of the number of layers $\mathcal{L}$ or the bond dimension χ. Different shades of the colors indicate different numbers of qubits n. **B** Number of parameters $\mathcal{P}$ that are needed such that magic and entanglement of the various ensembles reach 90% of the asymptotic values of the Haar distribution, namely $\widetilde{\mathcal{M}}_{n,\mu} = 0.9$ and $\widetilde{\mathcal{S}}_{n,\mu} = 0.9$. Triangles and circles represent magic and entanglement, respectively. Different colors indicate different classes of states. Here $\widetilde{\mathcal{M}}_{n,\mathrm{CMPS}}$ is omitted since $\widetilde{\mathcal{M}}_{n,\mathrm{CMPS}} = \widetilde{\mathcal{M}}_{n,\mathrm{MPS}}$ by construction.

Finally, we find that the CMPS displays a peculiar behavior. Indeed, the mean entanglement reaches an entanglement content compatible with the Haar asymptotic value already at $\chi = 1$. From this perspective, increasing χ only improves the content of magic $\widetilde{\mathcal{M}}_{\mathrm{CMPS}}$. Since $\widetilde{\mathcal{S}}_{\mathrm{CMPS}}$ is always compatible with $\widetilde{\mathcal{S}}_{\mathrm{H}} = 1$ and $\widetilde{\mathcal{M}}_{\mathrm{CMPS}}$ saturates quickly, we observe that the CMPS contains a high degree of quantumness despite being classically simulatable for the purpose of QNNs.

With an eye on QNN applications, in Fig. 3**B**, we discuss the normalized number of parameters $\mathcal{P}/n$ necessary to reach the magic and entanglement values of the Haar distribution. As a criterion, we plot at which value of $\mathcal{P}/n$ the various classes of states reach 90% of the resources of the Haar-distribution, in analogy with the definition of *entangling layers* used in Ref. [24]. Although limited to a small range of qubits n, we can estimate the scaling of $\mathcal{P}/n$ for the various resources. Concerning $\widetilde{\mathcal{M}}$, we observe that the scaling behaviors as a function of n appear to be compatible with a polynomial scaling for all μ. This observation also applies to MPS and CMPS. On the other hand, for what concerns entanglement, we show that the value of $\mathcal{P}/n$ needed to reach $\widetilde{\mathcal{S}} = 0.9$ for MPS is compatible with an exponential scaling as a function of n, in accordance with the known theoretical bound [58]. This means that an exponential $\mathcal{P}$ is needed to effectively reproduce the non-classicality of the Haar distribution as the feature space $\mathcal{H}_n$ increases in size.

3.2 Frame potentials and Welch bounds

As discussed previously, a way to estimate the deviations of a distribution of states from the Haar distribution is in terms of the frame potentials $\mathcal{F}_\mu^{(t)}$. We collect samples of fidelities $F = |\langle\psi(\boldsymbol{\theta})|\psi(\boldsymbol{\phi})\rangle|^2$ by generating 10^7 pairs of states and we estimate the frame potentials as the moments of the distribution $P(F)$, according to Eq. (5). For a given distribution μ, all moments $\mathcal{F}_\mu^{(t)}$ are estimated using the same set of fidelity sample pools. Since the Welch bound $\mathcal{F}_{\mathrm{H}}^{(t)}$ is a lower bound of the frame potentials, we define rescaled potentials $\widetilde{\mathcal{F}}_\mu^{(t)} = \mathcal{F}_\mu^{(t)}/\mathcal{F}_{\mathrm{H}}^{(t)}$, which are computed for the different classes of states as a function of the order t. The lower-bound condition is then expressed in terms of the normalized potentials as $\widetilde{\mathcal{F}}_\mu^{(t)} \geq 1$, with $\widetilde{\mathcal{F}}_{\mathrm{H}}^{(t)} \equiv 1$. In Fig. 4**A** we show the numerical estimation of the frame potentials for $n = 10$ qubits as a function of the order t. In Fig. 4**B**, we provide a zoom of the results in the region around $\widetilde{\mathcal{F}}^{(t)} \approx 1$ for $t = 4, 5, 6$ to resolve small differences from the Welch bound. All frame potentials are shown for selected values of $\mathcal{L}$ (for fQNN) and χ (for MPS and CMPS), which are chosen such that the total

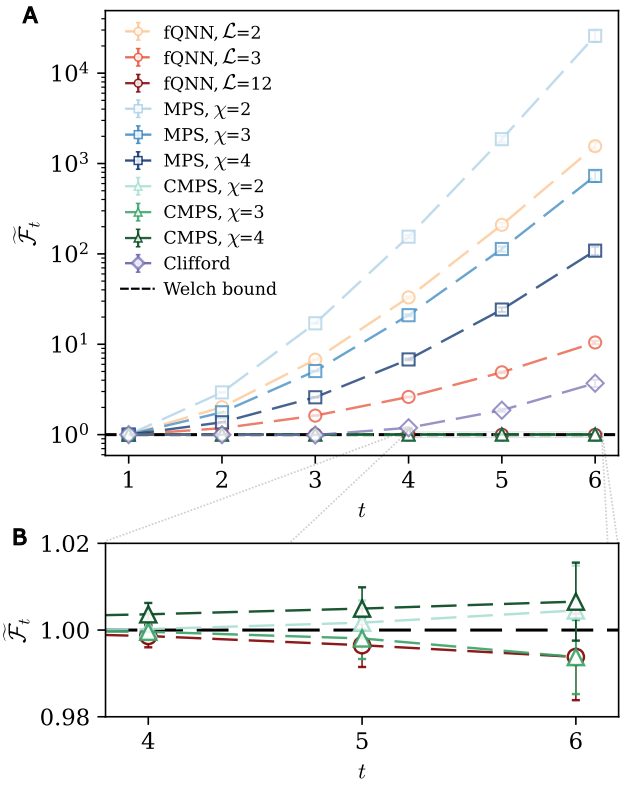

Figure 4: Rescaled frame potentials $\widetilde{\mathcal{F}}_\mu^{(t)}$ in function of the t, numerically estimated with 10^7 samples of fidelities F. In panel **A** we show results for $n = 10$ qubits and different classes of states. The Welch bound is marked by the black horizontal line. In panel **B** we show a zoom in a region close to the Welch bound for $t = 4, 5, 6$. The results for CMPS are compatible with the Haar moments distribution even for a bond dimension of $\chi = 2$. For a fair comparison, we show, for instance, the case of fQNN with $\mathcal{L} = 12$ layers because it has the same number of parameters, $\mathcal{P} = 360$, as a CMPS with bond dimension $\chi = 3$.

number of parameters $\mathcal{P}$ of the different architectures is comparable.

We observe that fQNN with $\mathcal{L} = 2, 3$ and MPS at low χ diverge super-exponentially from the Welch bound as a function of t. These state ensembles fail to reproduce moments of the Haar distribution beyond $t = 1$. However, fQNN with $\mathcal{L} = 12$ (or $\mathcal{P} = 360$ parameters) outperforms the MPS both with $\chi = 3$ (same number of parameters) and $\chi = 4$ ($\mathcal{P} = 640$).

In Fig. 4**B**, we highlight how the frame potentials $\widetilde{\mathcal{F}}^{(t)}$ of a fQNN with $\mathcal{L} = 12$ are numerically compatible with the Welch bound within the error bars for all values of t considered here. Similarly, also the frame potentials of CMPS agree with the Welch bound already for $\chi = 2$ within the same range of t. This suggests that CMPS require a smaller number of parameters $\mathcal{P}$ than fQNN to converge to the Welch bound. For completeness, we also numerically demonstrate that stabilizer states generated by Clifford unitaries form a 3-design, in agreement with known analytical results [59, 60].

3.3 Quantifying expressibility by Kullback-Leibler divergence

Expressibility evaluated as the Kullback-Leibler divergence between the fidelity ditributions $P_\mu(F)$ and $P_{\mathrm{H}}(F)$ [Eq. (6)] provides a simple metric, but needs to be evaluated numerically in the case of an unknown distribution $P_\mu(F)$. To this end, we sample 10^5 pairs of states and compute the corresponding fidelities. Creating a histogram from these samples to approximate the distribution $P_\mu(F)$ requires defining a binning strategy. For distributions with sharp peaks or rapid variations, selecting the bin width and number of bins is a particularly delicate task, as it strongly influences the accuracy and usefulness of the histogram as an approximation of the underlying distribution. As the number of qubits increases, the underlying distribution $P_\mu(F)$ quickly becomes extremely peaked at $F \sim 0$, similar to the fidelities of Haar-distributed states $P_{\mathrm{H}}(F) \sim O(d(1-F)^d)$. Therefore, instead of creating a histogram, we employ the Gaussian Kernel Density Estimation (KDE) [61] method to approximate $P_\mu(F)$ from the samples of fidelities. KDE provides a smoother and more accurate representation of the distribution, especially for datasets with sharp peaks or rapid variations. We denote distributions numerically estimated by KDE with a hat, e.g. $\hat{P}_\mu(F)$. The latter, together with the analytically known $P_{\mathrm{H}}(F)$, are used to compute the Kullback-Leibler divergence of Eq. (6) for the different classes of states. The results are shown in Fig. 5, for $n = 6, 8, 10$ qubits.

Knowing the analytical formula for $P_{\mathrm{H}}(F)$, we are able to assess the error introduced by the finite sampling. This is achieved by computing the Kullback-Leibler divergence between the exact probability $P_{\mathrm{H}}(F)$ and the numerically estimated $\hat{P}_{\mathrm{H}}(F)$, obtained by sampling 10^5 pairs of Haar-distributed states. We obtain a discrepancy $D_{\mathrm{KL}}(\hat{P}_{\mathrm{H}} \,\|\, P_{\mathrm{H}}) \approx 3 \times 10^{-4}$ in the cases of $n = 6, 8, 10$ qubits, indicated by the shaded area in Fig. 5 (which accounts for error bars). This means that $D_{\mathrm{KL}}(\hat{P}_\mu \,\|\, P_{\mathrm{H}})$ computed from 10^5

samples is approximately bounded from below by this region, regardless of the ensemble of states considered. We observe how the values of D_{KL} of the different classes of states approach the lower bound as the number of parameters is increased. The results for fQNN and CMPS architectures converge significantly faster than MPS.

Considering expressibility as a relative metric that allows us to evaluate which architectures better approximate the Haar distribution, we choose a threshold value of $D_{\mathrm{KL}} = 10^{-3}$ to assess the characteristic number of parameters needed to reach this level of approximation as the state size increases, and thus the number of qubits. As shown in Fig. 5, although fQNN is much faster than MPS in reaching the chosen bound, CMPS already surpasses the threshold for $\chi = 1$. We observe that in the range $n = 4$ to $n = 10$, a bond dimension $\chi = 1$ is always sufficient for CMPS, unlike the cases of fQNN and MPS where the number of parameters (normalized by the number of qubits) increases within this range of instances.

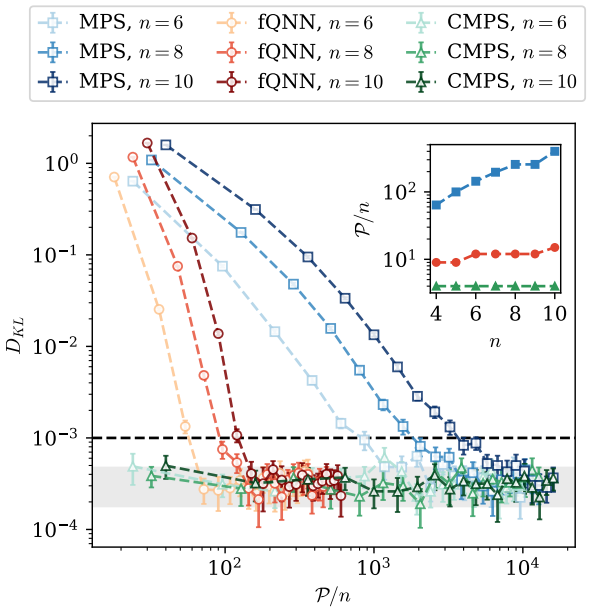

Figure 5: Kullback-Leibler divergence D_{KL} between numerically estimated fidelity distributions $\hat{P}_\mu(F)$, for different classes of states, and the exact Haar fidelity distribution $P_{\mathrm{H}}(F)$. Error bars are computed using jackknife resampling. Lower D_{KL} divergence signifies higher expressibility of the architecture in approximating the Haar distribution on pure quantum states. The gray shaded region represents the KL divergence between the true Haar fidelity distribution $P_{\mathrm{H}}(F)$ and the one obtained by sampling 10^5 Haar-distributed fidelities for $n = 6, 8, 10$ qubits, accounting for the errors in this estimate. In the inset, the number of parameters, normalized by the number of qubits, required by different architectures to reach a threshold of $D_{\mathrm{KL}} = 10^{-3}$ (black dashed line in the main plot) is shown.

4 Discussion

In terms of quantum resources, the fQNN quantum circuit discussed in this work can be seen as a sequence of alternating layers that inject magic, via single-qubit rotations, and generate entanglement, through the ring of CNOT gates. In this respect, an isolated layer of the fQNN is equivalent to a CMPS with bond dimension 1, where the Clifford operator is the entangling map of CNOTs. The repeated application of these layers underscores the difficulty of classically simulating such architectures. In this paper we show that the high expressibility achieved by quantum circuits that alternate phases of magic and entanglement injection can still be obtained with CMPS architectures that are classically simulatable with polynomial computational cost and memory requirements. The key insight is that it suffices to partition the model so that the two resources, magic and entanglement, are injected sequentially, as is done in CMPS. In particular, the process involves first using an MPS to inject magic, which grows rapidly with increasing bond dimension, and then applying a Clifford circuit to generate large entanglement (leaving the magic invariant). This approach potentially enables the construction of QNN with high expressibility, while retaining efficient classical simulability.

5 Conclusions

In this work, we assessed the amount of magic and entanglement of three architectures for QNN applications: fQNN, MPS and CMPS. The first requires polynomial resources on a quantum computer, but is exponentially hard to emulate on classical hardware. In contrast, MPS and CMPS can be efficiently simulated on a classical computer. Our calculations show that CMPS display resource characteristics comparable to fQNN, using a similar number of parameters. In the context of QML, these findings challenge the assump-

tion that states with both high entanglement and high magic are not efficiently accessible using only classical resources.

To further investigate the connection between quantum resources and expressive power for QML, we assess how these architectures perform in terms of expressibility, namely the ability of a parameterized quantum state to approximate arbitrary random states. We find that both architectures achieve comparable levels of expressibility, according to common metrics based on the distribution of fidelities. This suggests that while fQNNs are promising for future quantum machine learning, CMPS can offer a practical, classically efficient alternative that requires no quantum hardware.

In addition to that, CMPS may also be useful in an integrated quantum-classical hybrid workflow, as they can be compiled in terms of quantum gates [62, 63, 64]. This might enable efficient training using only a classical computer [51], alleviating the burden of extensive measurements on a quantum hardware. Instead, the inference stage may be performed on quantum hardware, potentially leading to significant energy savings and speed improvements. Future work involves the applications of these quantum-inspired methods to practical use cases, as they may help circumventing known issues of quantum on-chip training of QNN models.

Acknowledgement

We acknowledge the whole Quantum Computing Solutions group at Leonardo S.p.A., in particular Francesco Turro, for insightful discussions. We are grateful to Damiano Paniccia for helping us with the JAX implementation.

References

[1] Michael A. Nielsen and Isaac L. Chuang. "Quantum computation and quantum information: 10th anniversary edition". Cambridge University Press. (2010).

[2] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. "Quantum machine learning". Nature **549**, 195–202 (2017).

[3] Maria Schuld and Francesco Petruccione. "Machine learning with quantum computers". Springer. (2021).

[4] M. Cerezo, Guillaume Verdon, Hsin-Yuan Huang, Lukasz Cincio, and Patrick J. Coles. "Challenges and opportunities in quantum machine learning". Nature Computational Science **2**, 567–576 (2022).

[5] Maria Schuld and Nathan Killoran. "Quantum machine learning in feature hilbert spaces". Phys. Rev. Lett. **122**, 040504 (2019).

[6] Seth Lloyd, Maria Schuld, Aroosa Ijaz, Josh Izaac, and Nathan Killoran. "Quantum embeddings for machine learning" (2020). arXiv:2001.03622.

[7] Arash Ahmadi and Eliska Greplova. "Quantifying non-stabilizerness via information scrambling". SciPost Phys. **16**, 043 (2024).

[8] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. "Parameterized quantum circuits as machine learning models". Quantum science and technology **4**, 043001 (2019).

[9] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. "The quest for a quantum neural network". Quantum Information Processing **13**, 2567–2586 (2014).

[10] Carlo Ciliberto, Mark Herbster, Alessandro Davide Ialongo, Massimiliano Pontil, Andrea Rocchetto, Simone Severini, and Leonard Wossnig. "Quantum machine learning: a classical perspective". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **474**, 20170551 (2018).

[11] Marco Maronese and Enrico Prati. "A continuous rosenblatt quantum perceptron". International Journal of Quantum Information **19**, 2140002 (2021).

[12] Marco Maronese, Claudio Destri, and Enrico Prati. "Quantum activation functions for quantum neural networks". Quantum Information Processing **21**, 128 (2022).

[13] Maria Schuld, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. "Circuit-centric quantum classifiers". Phys. Rev. A **101**, 032308 (2020).

[14] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, and Dacheng Tao. "Expressive power of parametrized quantum circuits". Phys. Rev. Res. **2**, 033125 (2020).

[15] Amira Abbas, David Sutter, Christa Zoufal, Aurelien Lucchi, Alessio Figalli, and Stefan Woerner. "The power of quantum neural networks". Nature Computational Science **1**, 403–409 (2021).

[16] Vladimir Vapnik. "The nature of statistical learning theory". Springer science & business media. (1999).

[17] V.N. Vapnik. "An overview of statistical learning theory". IEEE Transactions on Neural Networks **10**, 988–999 (1999).

[18] Rodrigo Fernandes de Mello and Moacir Antonelli Ponti. "Machine learning: A practical approach on the statistical learning theory". Springer Publishing Company, Incorporated. (2018). 1st edition.

[19] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. "Effect of data encoding on the expressive power of variational quantum-machine-learning models". Phys. Rev. A **103**, 032430 (2021).

[20] Lukas Gonon and Antoine Jacquier. "Universal approximation theorem and error bounds for quantum neural networks and quantum reservoirs". IEEE Transactions on Neural Networks and Learning SystemsPages 1–14 (2025).

[21] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms". Advanced Quantum Technologies **2**, 1900070 (2019).

[22] Thomas Hubregtsen, Josef Pichlmeier, Patrick Stecher, and Koen Bertels. "Evaluation of parameterized quantum circuits: on the relation between classification accuracy, expressibility, and entangling capability". Quantum Machine Intelligence **3**, 9 (2021).

[23] Aram W. Harrow and Richard A. Low. "Efficient quantum tensor product expanders and k-designs". In Irit Dinur, Klaus Jansen, Joseph Naor, and José Rolim, editors, Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. Pages 548–561. Berlin, Heidelberg (2009). Springer Berlin Heidelberg. url: https://doi.org/10.1007/978-3-642-03685-9_41.

[24] Marco Ballarin, Stefano Mangini, Simone Montangero, Chiara Macchiavello, and Riccardo Mengoni. "Entanglement entropy production in Quantum Neural Networks". Quantum **7**, 1023 (2023).

[25] Daniel Gottesman. "The heisenberg representation of quantum computers" (1998). arXiv:quant-ph/9807006.

[26] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. "The resource theory of stabilizer quantum computation". New Journal of Physics **16**, 013009 (2014).

[27] Ulrich Schollwöck. "The density-matrix renormalization group in the age of matrix product states". Annals of Physics **326**, 96–192 (2011).

[28] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". Phys. Rev. A **70**, 052328 (2004).

[29] Dominik Szombathy, Angelo Valli, Cătălin Paşcu Moca, Lóránt Farkas, and Gergely Zaránd. "Independent stabilizer rényi entropy and entanglement fluctuations in random unitary circuits" (2025). arXiv:2501.11489.

[30] Daniele Iannotti, Gianluca Esposito, Lorenzo Campos Venuti, and Alioscia Hamma. "Entanglement and stabilizer entropies of random bipartite pure quantum states" (2025). arXiv:2501.19261.

[31] Guglielmo Lami, Tobias Haug, and Jacopo De Nardis. "Quantum state designs with clifford-enhanced matrix product states". PRX Quantum **6**, 010345 (2025).

[32] Antonio Anna Mele. "Introduction to Haar measure tools in quantum information: A beginner's tutorial". Quantum **8**, 1340 (2024).

[33] Liyuan Chen, Roy J. Garcia, Kaifeng Bu, and Arthur Jaffe. "Magic of random matrix product states". Phys. Rev. B **109**, 174207 (2024).

[34] Aram W. Harrow and Richard A. Low. "Random quantum circuits are approximate 2-designs". Communications in Mathematical Physics **291**, 257–302 (2009).

[35] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. "Stabilizer rényi entropy". Phys. Rev. Lett. **128**, 050402 (2022).

[36] Daniel Gottesman. "Stabilizer codes and quantum error correction". California Institute of Technology. (1997). url: https:

//resolver.caltech.edu/CaltechETD:etd-07162004-113028.

[37] Michael Beverland, Earl Campbell, Mark Howard, and Vadym Kliuchnikov. "Lower bounds on the non-clifford resources for quantum computations". Quantum Science and Technology **5**, 035009 (2020).

[38] Oliver Hahn, Alessandro Ferraro, Lina Hultquist, Giulia Ferrini, and Laura García-Álvarez. "Quantifying qubit magic resource with gottesman-kitaev-preskill encoding". Phys. Rev. Lett. **128**, 210502 (2022).

[39] Tobias Haug and Lorenzo Piroli. "Stabilizer entropies and nonstabilizerness monotones". Quantum **7**, 1092 (2023).

[40] Tobias Haug and M.S. Kim. "Scalable measures of magic resource for quantum computers". PRX Quantum **4**, 010301 (2023).

[41] Xhek Turkeshi, Marco Schirò, and Piotr Sierant. "Measuring nonstabilizerness via multifractal flatness". Phys. Rev. A **108**, 042408 (2023).

[42] Lorenzo Leone and Lennart Bittel. "Stabilizer entropies are monotones for magic-state resource theory". Phys. Rev. A **110**, L040403 (2024).

[43] Xhek Turkeshi, Anatoly Dymarsky, and Piotr Sierant. "Pauli spectrum and nonstabilizerness of typical quantum many-body states". Phys. Rev. B **111**, 054301 (2025).

[44] Jens Eisert. "Entanglement and tensor network states". In Eva Pavarini, Erik Koch, and Ulrich Schollwöck, editors, Emergent Phenomena in Correlated Matter: Modeling and Simulation. Volume 3. Verlag des Forschungszentrum Jülich (2013). url: https://www.cond-mat.de/events/correl13/manuscripts/.

[45] Mario Collura, Guglielmo Lami, Nishan Ranabhat, and Alessandro Santini. "Tensor network techniques for quantum computation". SISSA Medialab s.r.l. (2024).

[46] J. Eisert, M. Cramer, and M. B. Plenio. "Colloquium: Area laws for the entanglement entropy". Rev. Mod. Phys. **82**, 277–306 (2010).

[47] Cécilia Lancien and David Pérez-García. "Correlation length in random mps and peps". Annales Henri Poincaré **23**, 141–222 (2022).

[48] Antonio Francesco Mello, Alessandro Santini, Guglielmo Lami, Jacopo De Nardis, and Mario Collura. "Clifford dressed time-dependent variational principle". Phys. Rev. Lett. **134**, 150403 (2025).

[49] Sergi Masot-Llima and Artur Garcia-Saez. "Stabilizer tensor networks: Universal quantum simulator on a basis of stabilizer states". Phys. Rev. Lett. **133**, 230601 (2024).

[50] Antonio Francesco Mello, Alessandro Santini, and Mario Collura. "Hybrid stabilizer matrix product operator". Phys. Rev. Lett. **133**, 150604 (2024).

[51] Xiangjian Qian, Jiale Huang, and Mingpu Qin. "Augmenting density matrix renormalization group with clifford circuits". Phys. Rev. Lett. **133**, 190402 (2024).

[52] Erik Hostens, Jeroen Dehaene, and Bart De Moor. "Stabilizer states and clifford operations for systems of arbitrary dimensions and modular arithmetic". Phys. Rev. A **71**, 042315 (2005).

[53] Sergey Bravyi and Dmitri Maslov. "Hadamard-free circuits expose the structure of the clifford group". IEEE Transactions on Information Theory **67**, 4546–4563 (2021).

[54] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M. Sohaib Alam, Guillermo Alonso-Linaje, B. AkashNarayanan, Ali Asadi, et al. "Pennylane: Automatic differentiation of hybrid quantum-classical computations" (2022). arXiv:1811.04968.

[55] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. "JAX: composable transformations of Python+NumPy programs". (2018). url: http://github.com/jax-ml/jax.

[56] Craig Gidney. "Stim: a fast stabilizer circuit simulator". Quantum **5**, 497 (2021).

[57] Cirq Developers. "Cirq". Zenodo. (2025).

[58] F. Verstraete and J. I. Cirac. "Matrix product states represent ground states faithfully". Phys. Rev. B **73**, 094423 (2006).

[59] Zak Webb. "The clifford group forms a unitary 3-design". Quantum Info. Comput. **16**, 1379–1400 (2016).

[60] Huangjun Zhu. "Multiqubit clifford groups are unitary 3-designs". Phys. Rev. A **96**, 062336 (2017).

[61] David W. Scott. "Multivariate density estimation: Theory, practice, and visualization". Wiley Series in Probability and Statistics. John Wiley & Sons, Inc. New York, Chichester (2015).

[62] Shi-Ju Ran. "Encoding of matrix product states into quantum circuits of one- and two-qubit gates". Phys. Rev. A **101**, 032310 (2020).

[63] Daniel Malz, Georgios Styliaris, Zhi-Yuan Wei, and J. Ignacio Cirac. "Preparation of matrix product states with log-depth quantum circuits". Phys. Rev. Lett. **132**, 040404 (2024).

[64] Kevin C. Smith, Abid Khan, Bryan K. Clark, S.M. Girvin, and Tzu-Chieh Wei. "Constant-depth preparation of matrix product states with adaptive quantum circuits". PRX Quantum **5**, 030344 (2024).