**Discover** Applied Sciences

Research

# Emerging generalization advantage of quantum-inspired machine learning in the diagnosis of hepatocellular carcinoma

**Domenico Pomarico[1,2] · Alfonso Monaco[1,2] · Nicola Amoroso[2,3] · Loredana Bellantuono[2,4] · Antonio Lacalamita[1,2] · Marianna La Rocca[1,2] · Tommaso Maggipinto[1,2] · Ester Pantaleo[1,2] · Sabina Tangaro[2,5] · Sebastiano Stramaglia[1,2] · Roberto Bellotti[1,2]**

Received: 6 June 2024 / Accepted: 21 February 2025
Published online: 06 March 2025


**Abstract**
Research into quantum advantage is increasingly taking on an interdisciplinary character. In particular, quantum machine learning shows promising generalization capabilities, which we have exploited in the classification of hepatocellular carcinoma tissue based on microarray gene expressions. By using previously characterized genetic communities, we minimize the computational complexity associated with the number of qubits, enabling the execution of quantum-inspired algorithms on classical machines. We consider two categories of such algorithms: parameterized quantum circuits (PQC) and tensor networks. The variational optimization of PQCs achieves better accuracy than classical counterparts on the independent test set, reaching an advantage equal to 11% in accuracy, while tensor networks offer equivalent performance with fewer parameters.

**Article highlights**

- We applied QML methods in the classification of hepatocellular carcinoma tissue based on microarray data.
- Quantum circuits shows higher performance on independent sets compared to classical methods.
- Tensor network methods can match the performance of artificial neural networks (ANNs) while using fewer parameters.

**Keywords** Quantum machine learning · Genes community · Generalization

Sebastiano Stramaglia and Roberto Bellotti have contributed equally to this work.

✉ Alfonso Monaco, alfonso.monaco@ba.infn.it; Domenico Pomarico, domenico.pomarico@ba.infn.it; Nicola Amoroso, nicola.amoroso@uniba.it; Loredana Bellantuono, loredana.bellantuono@ba.infn.it; Antonio Lacalamita, antonio.lacalamita@ba.infn.it; Marianna La Rocca, marianna.larocca@uniba.it; Tommaso Maggipinto, tommaso.maggipinto@ba.infn.it; Ester Pantaleo, ester.pantaleo@uniba.it; Sabina Tangaro, sonia.tangaro@ba.infn.it; Sebastiano Stramaglia, sebastiano.stramaglia@ba.infn.it; Roberto Bellotti, roberto.bellotti@ba.infn.it | [1]Dipartimento Interateneo di Fisica, Università degli Studi di Bari, 70125 Bari, Italy. [2]Istituto Nazionale di Fisica Nucleare, Sezione di Bari, 70125 Bari, Italy. [3]Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari, 70125 Bari, Italy. [4]Dipartimento di Biomedicina Traslazionale e Neuroscienze (DiBraiN), Università degli Studi di Bari, 70124 Bari, Italy. [5]Dipartimento Di Scienze Del Suolo, Della Pianta e Degli Alimenti, Università degli Studi di Bari, 70125 Bari, Italy.

## 1 Introduction

The constantly growing availability of data and the increasing complexity in analyzing them are driving the development of new technologies that leverage quantum computing to achieve a more efficient integration of data with respect to classical approaches. Quantum applications support these efforts in two ways: by exploiting parallel quantum state superposition [1] and by estimating unknown parameters with unparalleled efficiency [2, 3].

The literature presents a plethora of quantum machine learning (QML) applications addressing various issues, including optimal embeddings [4], multipartite entanglement [5], Fourier decompositions via feature re-uploading [6, 7], data symmetries [8–10], barren plateaus [11–13], capacity [14], and expressivity [15], among others. This list being far from exhaustive.

The mapping of classical data into quantum states leads to the definition of quantum kernels, which form the basis for quantum support vector machines (SVM) [16–19]. Once the kernel is defined, classical SVM optimization methods can be applied using training data, though they may entail high computational complexity due to the high dimensions of the quantum state space. In this work, we will exploit variational methods [19–22], which offer advantages in terms of computational time resources, even if they may involve a lower precision.

A typical limitation of QML applications is the noise present in noisy intermediate-scale quantum (NISQ) devices [23]. This characteristic challenge of open quantum systems is addressed by tensor networks, either as stand-alone quantum-inspired numerical tools [24–29] or by utilizing classical resources whose output is converted into a parametric quantum circuit (PQC) [30–32]. These hybrid approaches leverage tensor networks computational capabilities to minimize the number of gates composing a quantum circuit in order to bound noise sources affecting results precision. Another strategy consists in the implementation of error correction, which can incorporate causal relationships in the learning framework, since the associated tensor renormalization procedure effectively iterates feature selection from causally related groups, leading to the extraction of classification labels [33–35].

Machine learning (ML) methods are proving useful in quantum many-body systems [36, 37] and are currently applied to a variety of problems [38]. Classifiers based on quantum models exhibit an interesting behavior concerning an emerging training size advantage [39–41]. Similarly, entanglement-based discrimination between quantum channels is characterized by an exponential learning advantage in the required number of samples for both discrete and continuous variable platforms [42–44]. Quantum resources also exponentially improve unsupervised clustering methods [45–47], which is particularly advantageous in the biomedical field, given the specific need for explainability [48–51].

In the biological field, it is beneficial to illustrate how quantum resources scale concerning the complexity of the problem, using small molecules as an example [52]. Efficient results have been achieved with small superconducting circuits consisting of six qubits, whereas a significantly higher number is necessary for tasks such as protein folding [53, 54]. Tensor network methods are also utilized to model both covalent and ionic bonds between a large number of orbitals [55, 56]. This natural progression extends to assisted drug discovery in pharmacology, leveraging the interplay between quantum chemical models and genomic big data machine learning tools [57]. The complexity inherent in biological data aligns well with the high number of qubits offered by platforms like D-Wave, processed through quantum annealing [39, 41, 58].

This paper adopts a methodological perspective focusing on the comparison of the classification performance between quantum and conventional algorithms applied to genomics and oncology [39, 41, 59–68]. The considered application scenario refers to gene expressions from control and hepatocellular carcinoma (HCC) tissues recently analysed in [51] using complex networks and ML to identify pivotal communities linked to the HCC phenotype. This characterisation of gene datasets represents statistical preprocessing required for the development of supervised, quantum-inspired ML. Our study primarily focuses on the performance comparison between classical and QML applications, while discussion of the biological implications is beyond the scope of this article. In [39, 41] quantum annealing is applied for the classification of gene expression and multiomics tumor data, unveiling promising performances already comparable with conventional ML algorithms. A similar molecular subtype classification for lung cancer in [63] is targeted by a hybrid combination of classical feature selection and a quantum Boltzmann machine. In the framework of breast cancer features selection, including clinical data, is presented in [62] as well, in order to minimize circuit complexity in a quantum-inspired framework. Genes subset selection is targeted in [64] according to a hybrid scheme, highlighting again the importance of dimensional reduction to prepare input data of a quantum algorithm.

The input gene expression data and the associated communities are briefly presented in Sect. 2. In addition, the chosen PQC encoding for three features per qubit [7] is presented as implemented in PennyLane [69], while the tensor network methods [70] are based on TorchMPS [71–73] and equivalent tensor trains on TensorFlow [74–78].

Conventional ML algorithms such as SVM, random forest (RF) and artificial neural networks (ANN) are used as benchmarks. In Sect. 3 we present and discuss our findigs, consisting in good performances of the considered variational PQC on the independent set, also known as generalization [79–83], with a significant benefit from features re-uploading, implemented by a repeated application of gates used for features encoding.

## 2 Methods

Two microarray public datasets, GSE102079 and GSE54236, were considered (https://www.ncbi.nlm.nih.gov/geo/). The GSE102079 dataset comprises gene expression data from liver tissue samples of 152 patients who underwent liver resection. These samples were analyzed using the GPL570 Affymetrix Human Genome U133 Plus 2.0 array. This dataset includes 152 tumor tissues and 91 adjacent liver tissues from patients with HCC and 14 adjacent liver tissues from patients with colorectal cancer metastases who had not undergone chemotherapy. The GSE54236 dataset contains gene expression data from 156 samples, 78 HCC tumor tissues and 78 adjacent non-tumor tissues. These data were obtained using the GPL6480 Agilent-014850 Whole Human Genome Microarray. We used this dataset for independent testing purposes.

Microarray sequencing technologies simultaneously measure the expression level of thousands of genes. Both datasets contain over 10, 000 gene expressions, so a straightforward mapping of these data is unfeasible, because a too high number of qubits is required, or alternatively in a quantum-inspired setting a huge memory storage should be available. For this reason, we processed the data as in [51] in two steps: (i) a data normalization step based on multiarray analysis (robust multiarray analysis, RMA) with background correction of the original data, log2 transformation and quantile normalization; (ii) a community detection phase in which we first created the gene co-expression matrix and then applied the Leiden algorithm to find stable gene communities with highly correlated gene expression profiles, as pictorially represented in the scheme shown in the left part of Fig. 1.

Based on the results reported in [51], we selected the 20 best performing communities out of 46 where the RF algorithm discriminated HCC from control patients with an accuracy greater than 90%. We used these communities to feed two families of quantum algorithms, namely variational quantum circuits and tensor networks, and compared their performance with 3 conventional ML methods, namely RF, SVM and ANN. In Table 1 we list the 20 communities with corresponding cardinalities.

To verify the better performance on smaller training datasets of quantum algorithms than conventional ML methods [41], we trained the selected algorithms with an increasing proportion of the training set and tested their performance on the independent dataset. Specifically, the random cuts are repeated 20 times for each fraction from 20 to 90% with step 10%, for each classifier and for each community listed in Table 1. Figure 1 shows a pictorial representation of the implemented procedure.

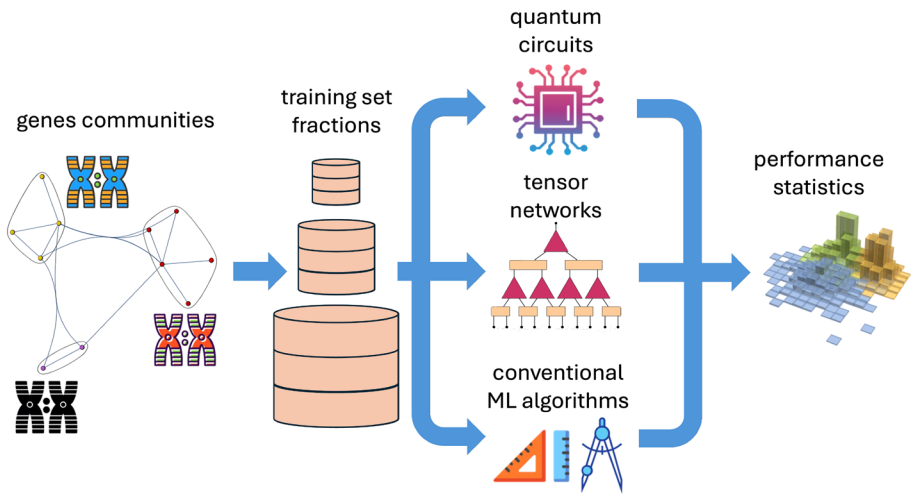


**Fig. 1** A flowchart representing the adopted workflow, based on genes communities identified in [51]. For each features subset we sampled random cuts with an increasing fraction of the training set fed to the considered three families of classification algorithms, quantum circuits and tensor networks compared with conventional ML algorithms. Collected performances are evaluated in order to establish statistically significant different behaviors

**Table 1** Summary of the considered genes communities with related number of qubits according to the chosen encoding in PQCs

| Community | # Genes | # Qubits | Community | # Genes | # Qubits |
|---|---|---|---|---|---|
| C8 | 28 | 10 | C29 | 48 | 16 |
| C12 | 47 | 16 | C30 | 32 | 11 |
| C14 | 31 | 11 | C31 | 25 | 9 |
| C15 | 25 | 9 | C32 | 35 | 12 |
| C16 | 34 | 12 | C33 | 35 | 12 |
| C17 | 26 | 9 | C35 | 31 | 11 |
| C23 | 31 | 11 | C40 | 29 | 10 |
| C24 | 23 | 8 | C41 | 48 | 16 |
| C27 | 36 | 12 | C42 | 33 | 11 |
| C28 | 35 | 12 | C43 | 32 | 11 |

### 2.1 PQC variational methods

Quantum models for QML in our classification problems yield a score dependent on both input features $\boldsymbol{x}$ and variational parameters $\theta$ characterizing the considered PQC

$$f_{\theta}(\boldsymbol{x}) = \langle 0|U^{\dagger}(\boldsymbol{x},\theta)\mathcal{M}U(\boldsymbol{x},\theta)|0\rangle, \tag{1}$$

where $U(\boldsymbol{x},\theta)$ is a sequence of encoding and parametrized gates, $|0\rangle$ is the initial state and $\mathcal{M}$ is an observable of the qubits system [4, 6, 19, 20, 22].

Usual mappings of one feature $x$ per qubit are implemented according to $R_X(x)$ or $R_Y(x)$ [19]. A minimum number of qubits $n$ required for each gene community is ensured in our case by the adopted mapping of three features per qubit [7], thus involving a system consisting in $n = \lceil \frac{N}{3} \rceil$ qubits

$$\begin{aligned}\mathcal{G}(x_1,x_2,x_3) =& R_Z(x_3)R_Y(x_2)R_Z(x_1)\\ =& e^{-\frac{i}{2}x_3 Z}e^{-\frac{i}{2}x_2 Y}e^{-\frac{i}{2}x_1 Z},\end{aligned} \tag{2}$$

composed according to the PQC schemes in Fig. 2 with variational parameters ruling the second kind of single qubit gates

$$\begin{aligned}\mathcal{W}(\theta_1,\theta_2,\theta_3) =& R_Z(\theta_3)R_Y(\theta_2)R_Z(\theta_1)\\ =& e^{-\frac{i}{2}\theta_3 Z}e^{-\frac{i}{2}\theta_2 Y}e^{-\frac{i}{2}\theta_1 Z},\end{aligned} \tag{3}$$

where $X$, $Y$ and $Z$ are Pauli matrices. An equivalent expression for these $SU(2)$ transformations is discussed in Appendix A.

The chosen PQC schemes presented in Fig. 2 are characterized by a circuit structure with

(a) separable states with no entanglement between qubits at the PQC output and measured through $\mathcal{M} = \frac{1}{n}\sum_{i=1}^{n} Z_i$;
(b) the previous circuit structure followed by an isometry layer, composed by controlled $\mathcal{CW}$, with measurements $\mathcal{M} = \frac{1}{n}\sum_{i=1}^{\lceil n/2 \rceil} Z_{2i-1}$;
(c) *CNOT* layers are introduced after the application of parametrized gates, with measurement $\mathcal{M} = \frac{1}{n}\sum_{i=1}^{n} Z_i$.

In practice we measure an average magnetization $\mathcal{M}$ at the PQC output, taking into account an additive contribution from each circuit line. The optimal parameters for the training stage correspond to 200 epochs with batch size equal to 2/3 of the available training set and learning rate $10^{-1}$, the loss function consists in the hinge loss. The tuning of each PQC is performed by means of Adam optimizer. Noise models for each PQC are not included in order to avoid performance results affected by environment effects.

### 2.2 Tensor network methods

The memory overhead in a classical computer for systems composed by a high number of qubits can be managed by specific dimensionality reduction techniques, named tensor networks and based on iterated truncations over less entangled degrees of freedom [35]. We consider three tensor network algorithms: matrix product states (MPS), tensor train (TT) and compressed tensor train (cTT). Tensor networks for one dimensional quantum systems are called MPS and are implemented

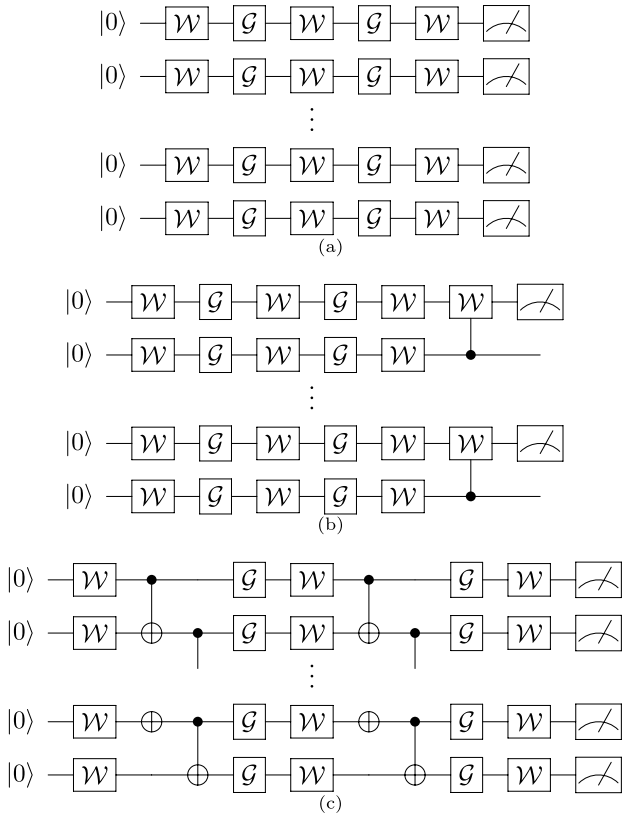


**Fig. 2** Representation of the three chosen PQCs: **a** features re-uploading with separable states, **b** the previous scheme followed by an isometry layer, **c** parametrized gates followed by a *CNOT* layer

in our case by means of TorchMPS [71–73]. The lower requirement for memory storage compared to methods that do not perform dimensionality reduction allows us to map each feature into a single qubit

$$\Phi(\boldsymbol{x})^{\boldsymbol{s}} = \bigotimes_{i=1}^{N} \phi(x_i)^{s_i} = \bigotimes_{i=1}^{N} \begin{pmatrix} x_i \\ 1 - x_i \end{pmatrix}, \tag{4}$$

where $s_i = 0, 1$, $\boldsymbol{s} = (s_1, \ldots, s_N)$ and the normalization with respect to the $L^1$ norm ensures numerical stability [70, 73]. This input state is a separable quantum spin state, according to the pictorial notation in Fig. 3. The definition of a score is based on the action of a tensor $W$ [24]

$$f^m(\boldsymbol{x}) = \sum_{s_1,\ldots,s_N} W^m_{s_1,\ldots,s_N} \phi(x_1)^{s_1} \otimes \cdots \otimes \phi(x_N)^{s_N}, \tag{5}$$

where the index $m$ runs over available classes in the classification problem. The tensor $W$ is converted into MPS format, as shown in Fig. 3**c**, such that the contraction of input indices is performed in parallel. The assigned label is $\widetilde{m} = \text{argmax}_m f^m(\boldsymbol{x})$ and for our binary classification model we define the score $f(\boldsymbol{x}) = f_1(\boldsymbol{x}) - f_0(\boldsymbol{x})$.

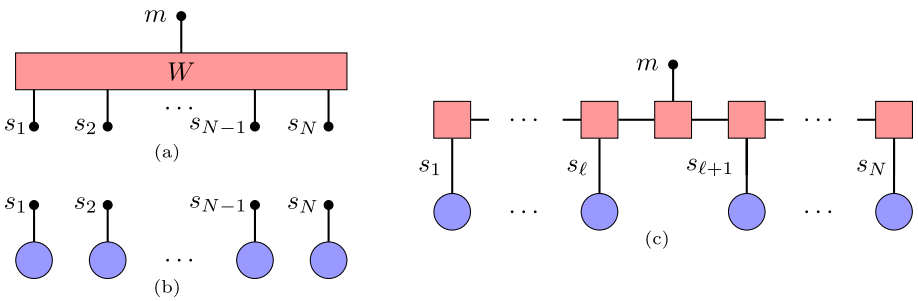


**Fig. 3** Schematic representation of the tensor $W$ trained for classification in panel (**a**) and of features mapped into a separable state $\Phi(\boldsymbol{x})$ of qubits in panel (**b**). Indices $(s_1, \ldots, s_N)$ contraction is optimized by an MPS approximation of the tensor $W$ shown in panel (**c**). The output index $m$ separates two regions $(s_1, \ldots, s_\ell)$ and $(s_{\ell+1}, \ldots, s_N)$ in TorchMPS [73]

The same matrices decomposition is possible without any encoding into quantum states, as implemented in TensorFlow [75–78] according to the TT format. Densely connected layers in ANN are compactly represented by means of these rank reduction techniques. It is possible to directly train a layer in TT format or alternatively we convert a dense layer after training in a cTT. The optimal bond dimension, namely the number of singular values in each tensor block after truncation, used for MPS is equal to 20, while it is 16 for TT and cTT. The number of epochs is 200 with batch size equal to 1/2 of the available traning set, learning rate is equal to $10^{-3}$ for MPS and cTT, while for TT it is $10^{-2}$. The loss function consists in the cross entropy, where MPS is endowed with a regularization penalty in order to avoid overfitting.

### 2.3 Conventional machine learning methods setup

We compared the performances of the chosen QML algorithms to three conventional ML methods, namely SVM [84], RF [85] and ANN [86].

SVM is a machine learning technique that utilizes mathematical functions known as kernels to transform data into a new hyperspace, making it easier to identify complex patterns within the data. When aiming to distinguish data from two clusters, SVM identifies the functional equation that best separates these clusters. With additional variables, the separation boundary evolves from a line to a plane, and with even more variables, it becomes a hyperplane. This hyperplane is derived from a subset of points from the different classes, known as support vectors. We considered the simplest example, which consists of a linear kernel enhanced by a regularization penalty to increase the performance in classifying independent groups.

RF is an ensemble learning method that constructs multiple classification trees using bootstrapped samples of the training dataset. During tree construction, at each node, a random subset of features is chosen, leading to trees that are only weakly correlated with one another. Generally, RF classifiers are straightforward to tune, highly resistant to overfitting, and especially effective when the number of features exceeds the number of observations. In our configuration we considered 100 trees and the Gini impurity criterion for splitting the nodes with no restrictions on the minimum number of samples in each leaf. The number of randomly selected features at each split is equal to $\sqrt{N}$, where we recall that $N$ is the total number of features.

ANNs were conceived as a model that mimics associative memory processes by storing a set of patterns. A layered network consisting of nodes connected by a set of weights allows the above patterns to be encoded during the training phase, based on the minimization of a loss function according to backpropagation. The output of each layer is fed with an activation function that provides the input for the next layer. We choose an ANN with a single hidden layer (with $N$ neurons) whose connections with the inputs are weighted by ReLU followed by softmax activations for the output nodes and cross entropy as a loss function.

In order to ensure a maximal adaptability for unseen data we do not impose any constraint ruling the training phase for each conventional ML algorithm.

## 3 Results and discussion

A common challenge in ML is balancing the trade-off between overfitting and generalizability. This issue is also relevant in QML [79–83], which additionally benefits from the potential training size advantage of quantum algorithms [39–41]. This advantage presents a promising perspective for biomedical applications. In our study, we will address both aspects by comparing performance during the training phase through cross-validation. We will also describe how classification accuracy varies on an independent set as the size of the training set increases.

For each random slice of 90% of the training set, we evaluate the performance considering the remaining 10% (see Table 2) in cross-validation. The conventional ML methods ANN, RF and SVM show high classification performance as measured by the reported accuracies. These performances are comparable with those obtained through TT and cTT tensor networks.

MPS tensor networks demonstrates excellent performance in almost all cases, excluding the C17 community. On average, the variational PQC (re-upload, isometry, *CNOT* layer) exhibits lower performances. However, on certain communities, such as C29 and C41, quantum learning methods achieve high classification accuracy comparable to conventional methods, confirming their rich information content as discussed in [51]. Similarly, the PQCs with re-upload and isometry

**Table 2** Training set (GSE102079 dataset) classification accuracy for each gene community in cross validation considering 90% of the training set and averaged over 20 random cuts, including the standard deviation

| | Parametrized Quantum Circuits | | | Tensor Networks | | | Conventional Methods | | |
|---|---|---|---|---|---|---|---|---|---|
| | re-upload | isometry | CNOT layer | MPS | TT | cTT | ANN | RF | SVM |
| C8 | 0.85±0.08 | 0.71±0.11 | 0.71±0.13 | 0.78±0.26 | 0.90±0.01 | 0.90±0.02 | 0.91±0.02 | 0.92±0.01 | 0.86±0.01 |
| C12 | 0.73±0.13 | 0.76±0.13 | 0.71±0.13 | 0.83±0.11 | 0.87±0.02 | 0.87±0.01 | 0.88±0.01 | 0.90±0.01 | 0.83±0.01 |
| C14 | 0.85±0.11 | 0.91±0.07 | 0.72±0.16 | 0.91±0.06 | 0.90±0.02 | 0.91±0.01 | 0.92±0.01 | 0.92±0.02 | 0.89±0.02 |
| C15 | 0.77±0.10 | 0.75±0.11 | 0.64±0.14 | 0.86±0.16 | 0.90±0.01 | 0.91±0.01 | 0.93±0.01 | 0.92±0.01 | 0.91±0.01 |
| C16 | 0.80±0.12 | 0.78±0.11 | 0.71±0.14 | 0.85±0.08 | 0.86±0.01 | 0.86±0.02 | 0.88±0.01 | 0.92±0.01 | 0.79±0.02 |
| C17 | 0.80±0.08 | 0.71±0.13 | 0.72±0.13 | 0.68±0.20 | 0.89±0.02 | 0.89±0.02 | 0.91±0.01 | 0.90±0.01 | 0.87±0.02 |
| C23 | 0.86±0.06 | 0.84±0.08 | 0.85±0.09 | 0.87±0.05 | 0.90±0.02 | 0.90±0.01 | 0.91±0.01 | 0.90±0.01 | 0.92±0.01 |
| C24 | 0.85±0.12 | 0.83±0.10 | 0.75±0.15 | 0.86±0.14 | 0.88±0.02 | 0.89±0.02 | 0.90±0.01 | 0.93±0.01 | 0.86±0.02 |
| C27 | 0.91±0.07 | 0.89±0.08 | 0.87±0.09 | 0.86±0.07 | 0.88±0.02 | 0.88±0.01 | 0.90±0.01 | 0.91±0.01 | 0.86±0.01 |
| C28 | 0.95±0.05 | 0.92±0.06 | 0.91±0.07 | 0.95±0.05 | 0.94±0.01 | 0.95±0.01 | 0.95±0.01 | 0.97±0.01 | 0.96±0.01 |
| C29 | 0.98±0.04 | 0.98±0.04 | 0.85±0.12 | 0.92±0.06 | 0.94±0.01 | 0.94±0.01 | 0.94±0.01 | 0.97±0.00 | 0.90±0.01 |
| C30 | 0.87±0.10 | 0.90±0.08 | 0.83±0.11 | 0.91±0.08 | 0.95±0.01 | 0.96±0.01 | 0.95±0.01 | 0.92±0.01 | 0.95±0.01 |
| C31 | 0.89±0.09 | 0.85±0.09 | 0.88±0.09 | 0.89±0.08 | 0.88±0.01 | 0.89±0.02 | 0.90±0.01 | 0.93±0.00 | 0.88±0.01 |
| C32 | 0.90±0.06 | 0.87±0.09 | 0.74±0.14 | 0.93±0.06 | 0.94±0.01 | 0.95±0.01 | 0.96±0.01 | 0.93±0.01 | 0.94±0.01 |
| C33 | 0.89±0.09 | 0.86±0.09 | 0.87±0.11 | 0.89±0.06 | 0.94±0.01 | 0.94±0.01 | 0.95±0.01 | 0.95±0.01 | 0.92±0.02 |
| C35 | 0.90±0.09 | 0.89±0.09 | 0.89±0.09 | 0.91±0.07 | 0.88±0.02 | 0.89±0.02 | 0.89±0.01 | 0.91±0.01 | 0.86±0.01 |
| C40 | 0.84±0.10 | 0.86±0.10 | 0.79±0.10 | 0.93±0.06 | 0.90±0.01 | 0.91±0.02 | 0.92±0.01 | 0.90±0.01 | 0.91±0.01 |
| C41 | 0.98±0.03 | 0.99±0.03 | 0.99±0.03 | 0.95±0.07 | 0.95±0.01 | 0.96±0.01 | 0.96±0.01 | 0.96±0.01 | 0.95±0.01 |
| C42 | 0.86±0.08 | 0.85±0.08 | 0.80±0.07 | 0.90±0.06 | 0.91±0.02 | 0.91±0.01 | 0.93±0.01 | 0.92±0.01 | 0.89±0.02 |
| C43 | 0.79±0.09 | 0.81±0.10 | 0.77±0.11 | 0.86±0.11 | 0.89±0.01 | 0.89±0.02 | 0.90±0.02 | 0.89±0.01 | 0.88±0.02 |

Graded colors from blue to brown are referred to increasing mean values

show robust performance, with notable differences only in C8, C14, and C17. In contrast, the PQC *CNOT* layer generally exhibits lower performances.

We tested the generalisation properties in terms of classification performances on the independent set presented in Sect. 2 and evaluated the performances using gradual reduction of the training set size. In Table 3 we summarised the accuracy performances corresponding to the largest training set (90% of the input dataset).

Among classical algorithms, both ANN and RF consistently outperform SVM in numerous communities, with statistically significant differences. ANN only surpasses RF in two gene communities, specifically C28 and C43. Tensor network methods exhibit performances consistent with the aforementioned algorithms, with MPS generally yielding equal or lower performances, except for C35 and C42. The TT and cTT methods demonstrate the ability to achieve comparable performance to ANN with fewer parameters, partially mitigating barren plateaus. Particularly interesting behavior is observed in PQCs, which generally achieve accuracy equal to or higher than classical benchmarks, albeit with fluctuations towards lower values in C14 and C29. This highlights the privileged role of the feature re-uploading schemes, identified by testing MPS, endowed with a single feature upload, with a modified encoding of three features per qubit (see Appendix A), observing no significant improvement.

We investigated the bilayer structure of PQCs, shown in Fig. 2, as an enhancement over the monolayer configuration. However, further increasing the number of layers results in performance saturation. Regarding biologically significant communities such as C29 and C41, as characterized in [51], it is notable that variational PQCs perform comparably to classical methods, affirming their rich information content. When comparing classifiers, it's essential to consider additional evaluation metrics like the Area Under the Receiver Operating Characteristic (ROC) curve (AUC) and the $F_1$ score. A comprehensive overview about the implemented metrics can be found in Appendix B. Table 3 shows that PQCs outperforms classical ML methods in particular for 8 communities namely C8, C17, C23, C24, C31, C32, C33 and C42. This behavior is statistically characterized by defining the generalization gap as the difference between training set

**Table 3** Independent set (GSE54236 dataset) classification accuracy for each gene community considering 90% of the training set and averaged over 20 random cuts, including the standard deviation

| | Parametrized Quantum Circuits | | | Tensor Networks | | | Conventional Methods | | |
|---|---|---|---|---|---|---|---|---|---|
| | re-upload | isometry | CNOT layer | MPS | TT | cTT | ANN | RF | SVM |
| C8 | 0.58±0.04 | 0.60±0.03 | 0.50±0.04 | 0.42±0.05 | 0.47±0.03 | 0.48±0.03 | 0.47±0.02 | 0.55±0.03 | 0.43±0.03 |
| C12 | 0.68±0.01 | 0.63±0.02 | 0.65±0.03 | 0.58±0.03 | 0.58±0.03 | 0.57±0.03 | 0.57±0.03 | 0.67±0.02 | 0.53±0.03 |
| C14 | 0.50±0.03 | 0.39±0.02 | 0.54±0.04 | 0.40±0.03 | 0.52±0.06 | 0.52±0.06 | 0.53±0.06 | 0.51±0.02 | 0.51±0.05 |
| C15 | 0.56±0.05 | 0.52±0.04 | 0.56±0.04 | 0.51±0.03 | 0.55±0.03 | 0.54±0.03 | 0.53±0.03 | 0.55±0.02 | 0.54±0.03 |
| C16 | 0.66±0.01 | 0.65±0.02 | 0.63±0.06 | 0.68±0.01 | 0.67±0.03 | 0.67±0.03 | 0.67±0.03 | 0.65±0.03 | 0.68±0.03 |
| C17 | 0.65±0.03 | 0.68±0.02 | 0.63±0.02 | 0.50±0.02 | 0.55±0.02 | 0.56±0.04 | 0.57±0.02 | 0.62±0.02 | 0.49±0.02 |
| C23 | 0.76±0.02 | 0.73±0.02 | 0.74±0.01 | 0.61±0.03 | 0.68±0.01 | 0.68±0.02 | 0.70±0.01 | 0.69±0.02 | 0.64±0.02 |
| C24 | 0.71±0.05 | 0.68±0.04 | 0.72±0.07 | 0.56±0.03 | 0.63±0.03 | 0.64±0.03 | 0.63±0.03 | 0.59±0.02 | 0.63±0.03 |
| C27 | 0.62±0.02 | 0.61±0.02 | 0.60±0.05 | 0.62±0.03 | 0.61±0.04 | 0.61±0.03 | 0.62±0.02 | 0.60±0.02 | 0.56±0.04 |
| C28 | 0.81±0.01 | 0.79±0.01 | 0.81±0.01 | 0.73±0.02 | 0.78±0.02 | 0.78±0.02 | 0.79±0.02 | 0.73±0.02 | 0.74±0.01 |
| C29 | 0.77±0.01 | 0.73±0.01 | 0.73±0.04 | 0.77±0.02 | 0.78±0.02 | 0.78±0.02 | 0.78±0.01 | 0.79±0.01 | 0.73±0.02 |
| C30 | 0.70±0.01 | 0.66±0.02 | 0.64±0.07 | 0.63±0.04 | 0.63±0.03 | 0.64±0.02 | 0.65±0.02 | 0.67±0.02 | 0.61±0.03 |
| C31 | 0.77±0.01 | 0.76±0.01 | 0.74±0.02 | 0.67±0.02 | 0.63±0.05 | 0.68±0.03 | 0.68±0.03 | 0.68±0.02 | 0.63±0.02 |
| C32 | 0.66±0.01 | 0.68±0.02 | 0.62±0.04 | 0.60±0.04 | 0.58±0.05 | 0.60±0.05 | 0.59±0.04 | 0.64±0.02 | 0.56±0.03 |
| C33 | 0.76±0.02 | 0.71±0.01 | 0.78±0.01 | 0.59±0.03 | 0.61±0.03 | 0.63±0.03 | 0.65±0.02 | 0.67±0.02 | 0.58±0.02 |
| C35 | 0.81±0.02 | 0.79±0.02 | 0.76±0.04 | 0.73±0.02 | 0.67±0.04 | 0.69±0.03 | 0.70±0.03 | 0.79±0.01 | 0.61±0.04 |
| C40 | 0.67±0.01 | 0.66±0.01 | 0.63±0.04 | 0.61±0.03 | 0.60±0.04 | 0.62±0.03 | 0.64±0.02 | 0.69±0.02 | 0.61±0.04 |
| C41 | 0.81±0.01 | 0.80±0.01 | 0.82±0.01 | 0.78±0.03 | 0.83±0.01 | 0.83±0.02 | 0.84±0.01 | 0.83±0.02 | 0.81±0.02 |
| C42 | 0.67±0.02 | 0.69±0.01 | 0.68±0.01 | 0.65±0.02 | 0.59±0.02 | 0.60±0.02 | 0.61±0.02 | 0.63±0.02 | 0.53±0.04 |
| C43 | 0.63±0.02 | 0.62±0.02 | 0.59±0.03 | 0.53±0.02 | 0.62±0.03 | 0.62±0.02 | 0.62±0.02 | 0.58±0.01 | 0.56±0.04 |

Graded colors from blue to brown are referred to increasing mean values

accuracy and independent set accuracy for each dataset. This quantity is described in Fig. 4, where we consider mean values listed in Tables 2 and 3. To quantify this difference we comprehensively collect values per panel in Fig. 4 and evaluate Kruskal-Wallis p-values for each pair: PQC obtain a p-value lower than 1% for both comparisons, namely with tensor network methods and conventional ML algorithms, while the p-value associated with the comparison of these two classes is higher than 1%. In Fig. 5 we report the classification accuracy obtained for these communities as function of the training set fraction.

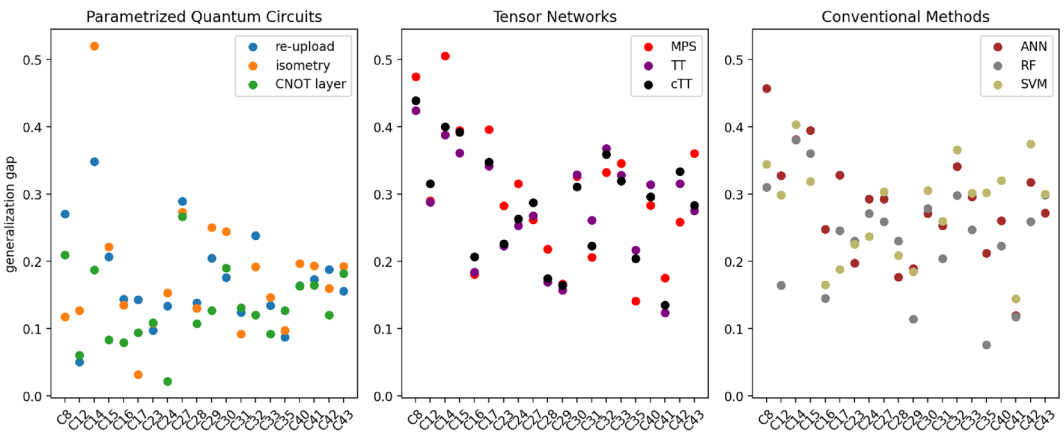


**Fig. 4** Generalization gap, defined as the difference between training set accuracy and independent set accuracy, is reported for each pair of elements in Tables 2 and 3, grouped in the three methods categories per panel

The better generalisation properties [79–83] of PQCs methods are evident, in most cases for small training sets. To highlight this behavior we carried out the Kruskal-Wallis test between the accuracy distributions obtained through the PQCs methods, for the 8 considered communities, with those relating to the best conventional ML method. Figure 6 reports the test results in terms of p-values for the methods re-upload, isometry and *CNOT* layer, as a function of training set fraction.

For the communities C31, C33 and C23, the statistical significance of 1% is already reached for small training sets. The centre and right panels of Fig. 6 refer to the isometric and *CNOT* layer PQCs that introduce entanglement between qubits. We can highlight that a reduced number of control gates characterising isometric PQCs ensures, on average, lower p-values for the eight communities considered, while for the *CNOT*-layer the statistical significance is lost in C8, C17 and C32. This behaviour can be explained by interpreting the last isometry layer as an effective selection of the feature subspace, which is endowed with a higher information content. Instead, the *CNOT* layer allows the exploration of a larger part of the quantum state space without focusing on the meaningful genes. PQC methods tend to achieve almost the same classification performances whether using small or large training set portions.

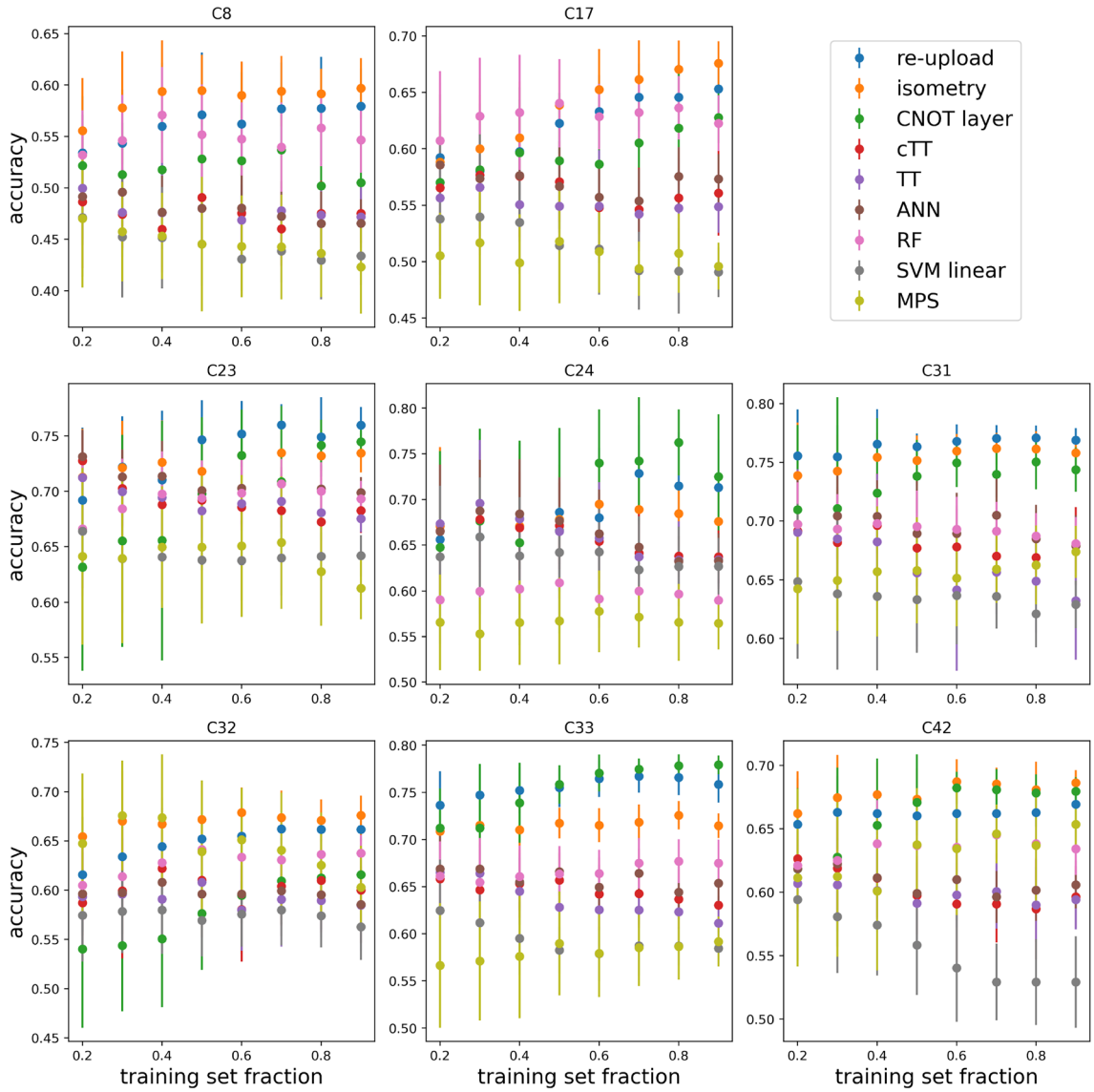


**Fig. 5** Independent set classification accuracy for the eight gene communities showing the best performances of PQCs with respect to classical algorithms, where error bars are referred to the standard deviation

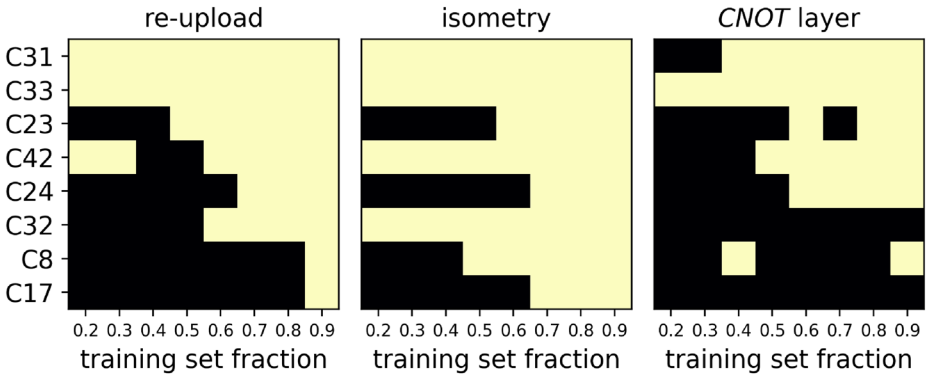


**Fig. 6** Maximum Kruskal-Wallis p-values selected between those yielded by comparisons with both RF and ANN, for re-upload, isometry and *CNOT* layer. Bright areas correspond to a p-value lower than 1%, while in black ones this statistical significance is lost. Communities are ordered with respect to p-values obtained by re-upload PQCs

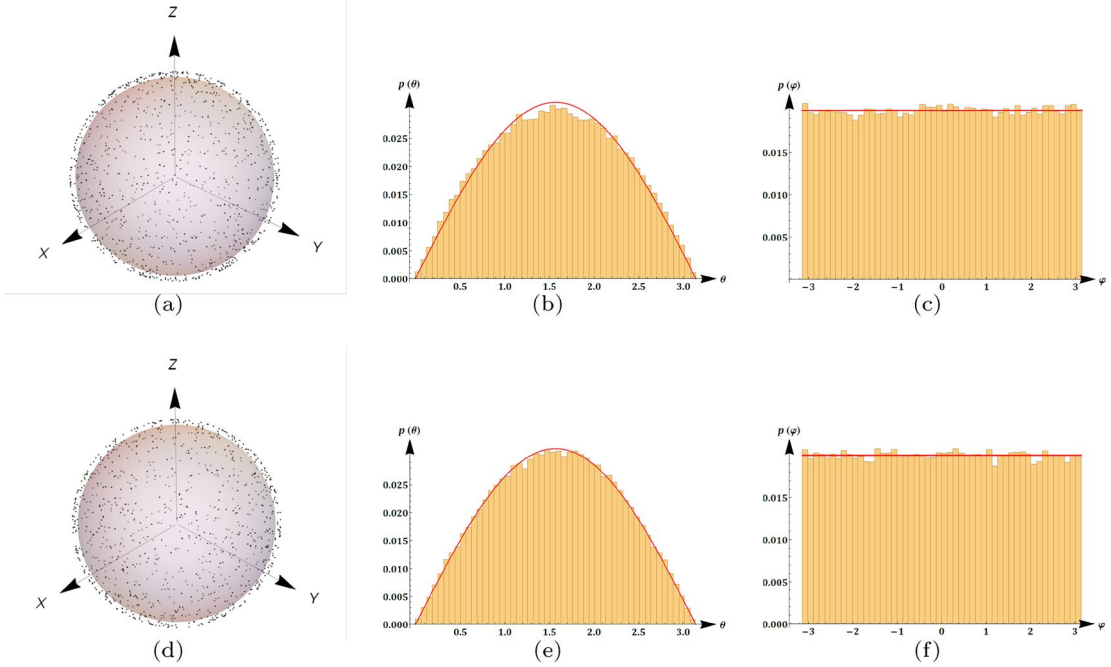


**Fig. 7** Expressivity for one (**a**–**c**) and two layers (**d**–**f**) for the PQC presented in Fig. 2**a**. The action of operators in Eq. (A13)–(A14), with random input features and parameters, over $|0\rangle$ is sampled for 1000 times in panels (**a**) and (**d**), while 100000 times in remaining panels

## 4 Conclusion

Our study aimed to compare classical and quantum-inspired ML algorithms for diagnosing hepatocellular carcinoma based on gene expression in liver tissues. Quantum circuits exhibit higher performance on independent sets, also when trained on small samples. Tensor network methods can achieve performances comparable to ANNs but with fewer parameters.

Future developments will involve utilizing quantum hardware. We also intend to apply synergistic approaches on individual communities in order to identify higher order correlations. Quantum algorithms might be particularly advantageous when applied on communities with synergies as many-body correlations could be exploited.

**Acknowledgements** DP acknowledges the support by PNRR MUR project CN00000041-"National Center for Gene Therapy and Drugs based on RNA Technology". We are grateful to Giovanni Gramegna and Giuseppe Magnifico for useful discussions and comments. Computational resources were provided by ReCaS Bari [87]. Some icons in Fig. 1 by https://freeicons.io/profile/433683, https://freeicons.io/profile/8939, https://freeicons.io/profile/3031 on https://freeicons.io. Authors were supported by the Italian funding within the "Budget MIUR - Dipartimenti di Eccellenza 2023 - 2027" (Law 232, 11 December 2016) - Quantum Sensing and Modelling for One-Health (QuaSiModO), CUP:H97G23000100001. Authors want to thank the project "Genoma mEdiciNa pERsonalizzatA -GENERA", local project code T3-AN-04 - CUP H93C22000500001, financed under the Health Development and Cohesion Plan 2014-2020, Trajectory 3 "Regenerative, predictive and personalized medicine" - Action line 3.1 "Creation of a precision medicine program for the mapping of the human genome on a national scale", referred to in the Notice of

the Ministry of Health published in the Official Journal no. 46 of 24 February 2021. Authors want to thank the Funder: Project funded under the National Recovery and Resilience Pan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender No. 3138 of 16 December 2021 of Italian Ministry of University and Research funded by the European Union - NextGenerationEU (award number/project code: CN00000013), and Concession Decree No. 1031 of 17 February 2022 adopted by the Italian Ministry of University and Research (CUP: D93C22000430001), Project title: "National Centre for HPC, Big Data and Quantum Computing".

**Author contributions** Conceptualization: D.P., A.M.; methodology: D.P., A.M., A.L., ; software: D.P. ; formal analysis: D.P.; writing—original draft preparation: D.P., A.M., ; writing—review and editing: D.P. , A.M., N.A., L.B., A.L., M.LR., T.M., E.P., S.T., S.S, R.B.; visualization: DP; supervision: A.M., S.S., R.B.

**Data availability** All codes and data used to perform the analysis are available upon request.

## Declarations

**Competing interests** The authors declare no competing interests.



## Appendix A: Features encoding and parametrized gates in variational PQCs

The encoding of three features per qubit presented in the main text can be resumed as a unique $SU(2)$ transformation by exploiting the composition rule [7]

$$e^{iT_1(\hat{n}\cdot\boldsymbol{\sigma})}e^{iT_2(\hat{m}\cdot\boldsymbol{\sigma})} = e^{iT_3(\hat{k}\cdot\boldsymbol{\sigma})}, \tag{A1}$$

where $T_1, T_2, T_3$ are transformation parameters, $\boldsymbol{\sigma} = (X, Y, Z)$ and we obtain for the spherical law of cosines

$$\cos T_3 = \cos T_1 \cos T_2 - \hat{n}\cdot\hat{m}\sin T_1 \sin T_2, \tag{A2}$$

furthermore yielding

$$\hat{k} = \frac{1}{\sin T_3}\left(\hat{n}\sin T_1\cos T_2 + \hat{m}\cos T_1 \sin T_2 - \hat{n}\times\hat{m}\sin T_1 \sin T_2\right). \tag{A3}$$

Features encoding in Eq. (2) is resumed as a unique a transformation by first composing

$$R_Y(x_2)R_Z(x_1) = e^{-\frac{i}{2}x_2 Y}e^{-\frac{i}{2}x_1 Z} = e^{iT_3(\hat{k}\cdot\boldsymbol{\sigma})}, \tag{A4}$$

deduced by setting $T_1 = -\frac{x_2}{2}, T_2 = -\frac{x_1}{2}, \hat{n} = \hat{y}$ and $\hat{m} = \hat{z}$

$$\cos T_3 = \cos\frac{x_2}{2}\cos\frac{x_1}{2}, \tag{A5}$$

$$\hat{k} = \frac{-1}{\sin T_3}\left(\hat{y}\sin\frac{x_2}{2}\cos\frac{x_1}{2} + \hat{z}\cos\frac{x_2}{2}\sin\frac{x_1}{2} + \hat{x}\sin\frac{x_2}{2}\sin\frac{x_1}{2}\right). \tag{A6}$$

The next step evaluates the last composition

$$\mathcal{G}(x_1, x_2, x_3) = R_Z(x_3)e^{iT_3(\hat{k}\cdot\boldsymbol{\sigma})} = e^{-\frac{i}{2}x_3 Z}e^{iT_3(\hat{k}\cdot\boldsymbol{\sigma})} = e^{iT_3'(\hat{k}'\cdot\boldsymbol{\sigma})}, \tag{A7}$$

implying $T_1 = -\frac{x_3}{2}, T_2 = T_3, \hat{n} = \hat{z}$ and $\hat{m} = \hat{k}$ according to the notation in Eq. (A1), thus imposing

$$\cos T_3' = \cos\frac{x_2}{2}\cos\frac{x_1+x_3}{2}, \tag{A8}$$

$$\begin{aligned}\hat{k}' = \frac{-1}{\sin T_3'}\Big(&\hat{z}\cos\frac{x_2}{2}\sin\frac{x_1+x_3}{2} + \hat{y}\sin\frac{x_2}{2}\cos\frac{x_3-x_1}{2}\\ &+\hat{x}\sin\frac{x_2}{2}\sin\frac{x_3-x_1}{2}\Big).\end{aligned} \tag{A9}$$

It is possible to explicitly consider the combination with Eq. (3) proposed in [7]

$$\mathcal{W}(\boldsymbol{\theta})\mathcal{G}(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}T(\boldsymbol{x},\boldsymbol{\theta})(\hat{k}(\boldsymbol{x},\boldsymbol{\theta})\cdot\boldsymbol{\sigma})}, \tag{A10}$$

adopting the compact notation $\boldsymbol{x} = (x_1, x_2, x_3)$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. The aforementioned $SU(2)$ transformations composition allows us to obtain

$$\begin{aligned}\cos T(\boldsymbol{x},\boldsymbol{\theta}) = &\cos\frac{x_2}{2}\cos\frac{\theta_2}{2}\cos\frac{x_1+x_3+\theta_1+\theta_3}{2}\\ &-\sin\frac{x_2}{2}\sin\frac{\theta_2}{2}\cos\frac{x_1-x_3-\theta_1+\theta_3}{2},\end{aligned} \tag{A11}$$

$$\hat{k}(\boldsymbol{x},\boldsymbol{\theta}) = \frac{-1}{\sin T(\boldsymbol{x},\boldsymbol{\theta})}\begin{pmatrix}\cos\frac{x_2}{2}\sin\frac{\theta_2}{2}\sin\frac{x_1+x_3-\theta_1+\theta_3}{2} - \sin\frac{x_2}{2}\cos\frac{\theta_2}{2}\sin\frac{x_1-x_3+\theta_1+\theta_3}{2}\\ \sin\frac{x_2}{2}\cos\frac{\theta_2}{2}\cos\frac{x_1-x_3+\theta_1+\theta_3}{2} + \cos\frac{x_2}{2}\sin\frac{\theta_2}{2}\cos\frac{x_1+x_3-\theta_1+\theta_3}{2}\\ \sin\frac{x_2}{2}\sin\frac{\theta_2}{2}\sin\frac{x_1-x_3-\theta_1+\theta_3}{2} + \cos\frac{x_2}{2}\cos\frac{\theta_2}{2}\sin\frac{x_1+x_3+\theta_1+\theta_3}{2}\end{pmatrix}, \tag{A12}$$

where rows are ordered according to $\hat{x}$, $\hat{y}$ and $\hat{z}$. This composition of transformations ensures the implementation of a single layer neural network according to the quantum formalism [7].

Our layered structures are inspired by Fourier decompositions [6] and they represent two examples of gate sequences $U(\boldsymbol{x},\boldsymbol{\theta})$ introduced in Eq. (1) required for the definition of quantum models

$$\mathcal{L}_1 = \mathcal{W}(\theta_1,\theta_2,\theta_3)\mathcal{G}(x_1,x_2,x_3)\mathcal{W}(\theta_4,\theta_5,\theta_6), \tag{A13}$$

$$\mathcal{L}_2 = \mathcal{W}(\theta_1,\theta_2,\theta_3)\mathcal{G}(x_1,x_2,x_3)\mathcal{W}(\theta_4,\theta_5,\theta_6)\mathcal{G}(x_1,x_2,x_3)\mathcal{W}(\theta_7,\theta_8,\theta_9), \tag{A14}$$

whose expressivity properties are shown in Fig. 7, panels (**a**–**c**) and (**d**–**f**) respectively. The uniform distribution over the Bloch sphere for a single qubit in panels (**a**) and (**d**) is referred to parameters in Eq. (A13)–(A14) sampled form a uniform distribution over $[0, 4\pi]$, while features $(x_1, x_2, x_3)$ are sampled from a standard normal distribution, since we adopt a standard rescaling for genes expression. Panels (**b**, **c**) and (**e**, **f**) are useful in order to verify the uniform distribution over the Bloch sphere, according to the Haar measure: considering the solid angle $p(\Omega)\mathrm{d}\Omega = \frac{1}{4\pi}\mathrm{d}\Omega$, we convert it into the joint density for polar and azimuthal angles $(\theta, \varphi)$ through the Jacobian $p(\Omega)\mathrm{d}\Omega = p(\theta,\varphi)\mathrm{d}\theta\mathrm{d}\varphi = \frac{1}{4\pi}\sin\theta\mathrm{d}\theta\mathrm{d}\varphi$, represented as red curves $p(\theta) = \frac{1}{2}\sin\theta$ and $p(\varphi) = \frac{1}{2\pi}$, which ensures the absence of any bias in our PQC.

## Appendix B: Evaluation metrics

The classification performances are evaluated in terms of

$$\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{B15}$$

where *TP* and *TN* stand for true positives (number of tissues with HCC correctly classified) and true negatives (number of control tissues correctly classified), while *FP* (number of control tissues identified as positive) and *FN* (number of tissues with HCC identified as control) are false positives and false negatives, respectively.

Since a classifier provides a score $f(\boldsymbol{x})$ for each element, it is useful to evaluate a metric able to take into account the variation of applied threshold $c$, namely AUC. Its definition is based on sensitivity (also known as recall) and specificity

$$\text{sensitivity} = \frac{TP}{TP + FN} = P(f(\boldsymbol{x}) > c | y = 1), \tag{B16}$$

$$\text{specificity} = \frac{TN}{TN + FP} = P(f(\boldsymbol{x}) \leq c | y = 0), \tag{B17}$$

representing two conditional probabilities, such that we define

$$\text{AUC} = \int_0^1 \text{sensitivity}(c)\, \text{d}(1 - \text{specificity}(c)), \tag{B18}$$

where $\text{d}(1 - \text{specificity}(c)) = \frac{\text{d}(1-\text{specificity}(c))}{\text{d}c}\text{d}c$ denoted the differential.

In order to quantify the ability of a classifier in selecting positive cases in a diagnostic binary decision problem, we introduce

$$\text{precision} = \frac{TP}{TP + FP} = P(y = 1 | f(\boldsymbol{x}) > c), \tag{B19}$$

such that we define the $F_1$ score as its harmonic mean with sensitivity

$$F_1 = \frac{2}{\text{precision}^{-1} + \text{sensitivity}^{-1}} = \frac{2\,TP}{2\,TP + FP + FN}. \tag{B20}$$

**Table 4** Independent set classification AUC for each gene community considering 90% of the training set and averaged over 20 random cuts

| | Parametrized Quantum Circuits | | | Tensor Networks | | | Conventional Methods | | |
|---|---|---|---|---|---|---|---|---|---|
| | re-upload | isometry | CNOT layer | MPS | TT | cTT | ANN | RF | SVM |
| C8 | 0.61±0.03 | 0.67±0.05 | 0.51±0.06 | 0.50±0.06 | 0.51±0.05 | 0.53±0.04 | 0.52±0.03 | 0.63±0.04 | 0.50±0.04 |
| C12 | 0.76±0.01 | 0.70±0.02 | 0.73±0.03 | 0.64±0.03 | 0.62±0.03 | 0.61±0.03 | 0.62±0.03 | 0.74±0.02 | 0.57±0.04 |
| C14 | 0.50±0.04 | 0.50±0.02 | 0.50±0.09 | 0.50±0.04 | 0.56±0.08 | 0.57±0.06 | 0.57±0.08 | 0.49±0.03 | 0.55±0.07 |
| C15 | 0.61±0.03 | 0.52±0.04 | 0.61±0.04 | 0.51±0.03 | 0.60±0.04 | 0.59±0.04 | 0.58±0.03 | 0.57±0.03 | 0.60±0.04 |
| C16 | 0.74±0.01 | 0.73±0.02 | 0.72±0.05 | 0.74±0.01 | 0.73±0.03 | 0.74±0.02 | 0.75±0.03 | 0.73±0.02 | 0.75±0.03 |
| C17 | 0.73±0.01 | 0.74±0.02 | 0.68±0.05 | 0.56±0.06 | 0.58±0.03 | 0.59±0.05 | 0.61±0.03 | 0.68±0.03 | 0.52±0.03 |
| C23 | 0.81±0.01 | 0.80±0.01 | 0.81±0.01 | 0.67±0.03 | 0.75±0.02 | 0.76±0.02 | 0.78±0.01 | 0.80±0.01 | 0.70±0.02 |
| C24 | 0.78±0.02 | 0.76±0.02 | 0.78±0.08 | 0.69±0.02 | 0.72±0.05 | 0.75±0.05 | 0.76±0.03 | 0.75±0.02 | 0.74±0.04 |
| C27 | 0.70±0.01 | 0.65±0.02 | 0.70±0.03 | 0.66±0.04 | 0.68±0.04 | 0.69±0.04 | 0.71±0.03 | 0.70±0.03 | 0.59±0.04 |
| C28 | 0.88±0.00 | 0.86±0.00 | 0.88±0.00 | 0.83±0.01 | 0.84±0.01 | 0.85±0.01 | 0.86±0.01 | 0.84±0.01 | 0.82±0.01 |
| C29 | 0.84±0.01 | 0.82±0.01 | 0.81±0.02 | 0.82±0.01 | 0.84±0.01 | 0.84±0.01 | 0.84±0.01 | 0.87±0.00 | 0.82±0.01 |
| C30 | 0.80±0.01 | 0.74±0.01 | 0.74±0.02 | 0.74±0.04 | 0.67±0.04 | 0.70±0.03 | 0.73±0.02 | 0.82±0.01 | 0.66±0.03 |
| C31 | 0.84±0.00 | 0.82±0.01 | 0.82±0.01 | 0.76±0.01 | 0.73±0.04 | 0.77±0.02 | 0.77±0.02 | 0.78±0.01 | 0.70±0.03 |
| C32 | 0.71±0.01 | 0.73±0.01 | 0.68±0.05 | 0.70±0.05 | 0.64±0.04 | 0.66±0.06 | 0.65±0.04 | 0.71±0.02 | 0.59±0.05 |
| C33 | 0.83±0.01 | 0.78±0.01 | 0.83±0.00 | 0.64±0.02 | 0.66±0.03 | 0.70±0.03 | 0.71±0.02 | 0.76±0.01 | 0.64±0.02 |
| C35 | 0.88±0.00 | 0.84±0.01 | 0.85±0.02 | 0.78±0.02 | 0.75±0.03 | 0.76±0.03 | 0.77±0.02 | 0.87±0.01 | 0.68±0.05 |
| C40 | 0.73±0.01 | 0.72±0.01 | 0.70±0.02 | 0.68±0.02 | 0.64±0.05 | 0.68±0.03 | 0.70±0.02 | 0.76±0.01 | 0.68±0.03 |
| C41 | 0.90±0.00 | 0.89±0.00 | 0.89±0.00 | 0.86±0.02 | 0.88±0.01 | 0.88±0.01 | 0.89±0.00 | 0.90±0.01 | 0.87±0.00 |
| C42 | 0.72±0.01 | 0.75±0.01 | 0.74±0.01 | 0.70±0.03 | 0.69±0.03 | 0.68±0.03 | 0.70±0.03 | 0.74±0.01 | 0.55±0.05 |
| C43 | 0.67±0.02 | 0.65±0.02 | 0.63±0.02 | 0.56±0.01 | 0.70±0.02 | 0.71±0.02 | 0.72±0.02 | 0.73±0.02 | 0.67±0.04 |

**Table 5** Independent set classification $F_1$ score for each gene community considering 90% of the training set and averaged over 20 random cuts

| | Parametrized Quantum Circuits | | | Tensor Networks | | | Conventional Methods | | |
|---|---|---|---|---|---|---|---|---|---|
| | re-upload | isometry | CNOT layer | MPS | TT | cTT | ANN | RF | SVM |
| C8 | 0.58±0.03 | 0.48±0.06 | 0.43±0.09 | 0.38±0.22 | 0.58±0.03 | 0.59±0.03 | 0.59±0.03 | 0.65±0.02 | 0.56±0.03 |
| C12 | 0.67±0.02 | 0.63±0.03 | 0.61±0.04 | 0.63±0.02 | 0.64±0.02 | 0.62±0.03 | 0.63±0.02 | 0.70±0.02 | 0.61±0.03 |
| C14 | 0.62±0.03 | 0.52±0.02 | 0.53±0.08 | 0.48±0.02 | 0.59±0.05 | 0.60±0.04 | 0.61±0.05 | 0.64±0.02 | 0.59±0.03 |
| C15 | 0.63±0.03 | 0.59±0.01 | 0.58±0.06 | 0.54±0.18 | 0.61±0.03 | 0.60±0.03 | 0.59±0.03 | 0.64±0.02 | 0.59±0.02 |
| C16 | 0.68±0.01 | 0.66±0.02 | 0.65±0.02 | 0.69±0.02 | 0.69±0.03 | 0.68±0.03 | 0.69±0.02 | 0.66±0.02 | 0.70±0.02 |
| C17 | 0.56±0.07 | 0.61±0.04 | 0.53±0.06 | 0.41±0.24 | 0.60±0.02 | 0.62±0.03 | 0.63±0.02 | 0.70±0.01 | 0.57±0.02 |
| C23 | 0.77±0.02 | 0.75±0.01 | 0.74±0.01 | 0.66±0.02 | 0.70±0.01 | 0.71±0.02 | 0.72±0.01 | 0.74±0.01 | 0.67±0.02 |
| C24 | 0.73±0.03 | 0.70±0.02 | 0.71±0.06 | 0.67±0.01 | 0.70±0.02 | 0.70±0.02 | 0.71±0.02 | 0.70±0.01 | 0.69±0.02 |
| C27 | 0.70±0.01 | 0.68±0.01 | 0.69±0.01 | 0.64±0.03 | 0.66±0.03 | 0.66±0.02 | 0.68±0.02 | 0.68±0.02 | 0.61±0.03 |
| C28 | 0.81±0.01 | 0.80±0.01 | 0.79±0.01 | 0.76±0.02 | 0.79±0.02 | 0.79±0.02 | 0.80±0.01 | 0.76±0.01 | 0.77±0.01 |
| C29 | 0.76±0.01 | 0.72±0.01 | 0.75±0.02 | 0.75±0.03 | 0.78±0.02 | 0.78±0.02 | 0.78±0.01 | 0.79±0.01 | 0.76±0.02 |
| C30 | 0.66±0.02 | 0.62±0.04 | 0.71±0.03 | 0.69±0.02 | 0.68±0.02 | 0.69±0.02 | 0.71±0.02 | 0.74±0.01 | 0.66±0.02 |
| C31 | 0.77±0.01 | 0.74±0.02 | 0.75±0.02 | 0.71±0.01 | 0.68±0.03 | 0.72±0.03 | 0.72±0.02 | 0.74±0.02 | 0.70±0.01 |
| C32 | 0.64±0.02 | 0.66±0.02 | 0.55±0.04 | 0.66±0.04 | 0.63±0.03 | 0.64±0.04 | 0.63±0.03 | 0.71±0.01 | 0.62±0.03 |
| C33 | 0.76±0.01 | 0.71±0.02 | 0.76±0.01 | 0.63±0.02 | 0.65±0.03 | 0.66±0.03 | 0.68±0.02 | 0.72±0.02 | 0.64±0.01 |
| C35 | 0.81±0.02 | 0.78±0.02 | 0.79±0.02 | 0.71±0.03 | 0.70±0.03 | 0.71±0.03 | 0.72±0.03 | 0.79±0.02 | 0.66±0.03 |
| C40 | 0.68±0.01 | 0.67±0.02 | 0.68±0.02 | 0.66±0.02 | 0.63±0.03 | 0.65±0.03 | 0.68±0.02 | 0.73±0.01 | 0.64±0.03 |
| C41 | 0.79±0.01 | 0.77±0.02 | 0.81±0.02 | 0.79±0.02 | 0.83±0.01 | 0.82±0.02 | 0.83±0.01 | 0.84±0.02 | 0.82±0.01 |
| C42 | 0.68±0.03 | 0.69±0.01 | 0.66±0.01 | 0.71±0.01 | 0.66±0.02 | 0.66±0.02 | 0.67±0.02 | 0.71±0.01 | 0.59±0.03 |
| C43 | 0.70±0.01 | 0.68±0.01 | 0.63±0.02 | 0.59±0.02 | 0.67±0.02 | 0.67±0.02 | 0.68±0.01 | 0.69±0.01 | 0.63±0.03 |

The useful property of harmonic mean in this framework is referred to the inequalities min(precision, sensitivity) $\leq F_1 \leq 2$ min(precision, sensitivity), so $F_1$ is dominated by the minimum of its arguments, both expressing the classifier ability in selecting HCC tissues based on a different conditional event.

Collected performances corresponding to the maximal considered fraction of the training set in terms of AUC and $F_1$ score are resumed in Tables 4 and 5, respectively. The gained information regarding a proper selection of HCC tissues allows us to restrict our focus on just C23, C31 and C33 among the eight communities discussed in Sect. 3. In Figs. 8 and 9 we can verify that variational PQCs advantage could be partially confirmed in the considered eight communities by looking at AUC, but $F_1$ score underlines a comparable selection of positive cases with classical counterparts in most genes communities, a behavior which is explained given the imbalanced training set.

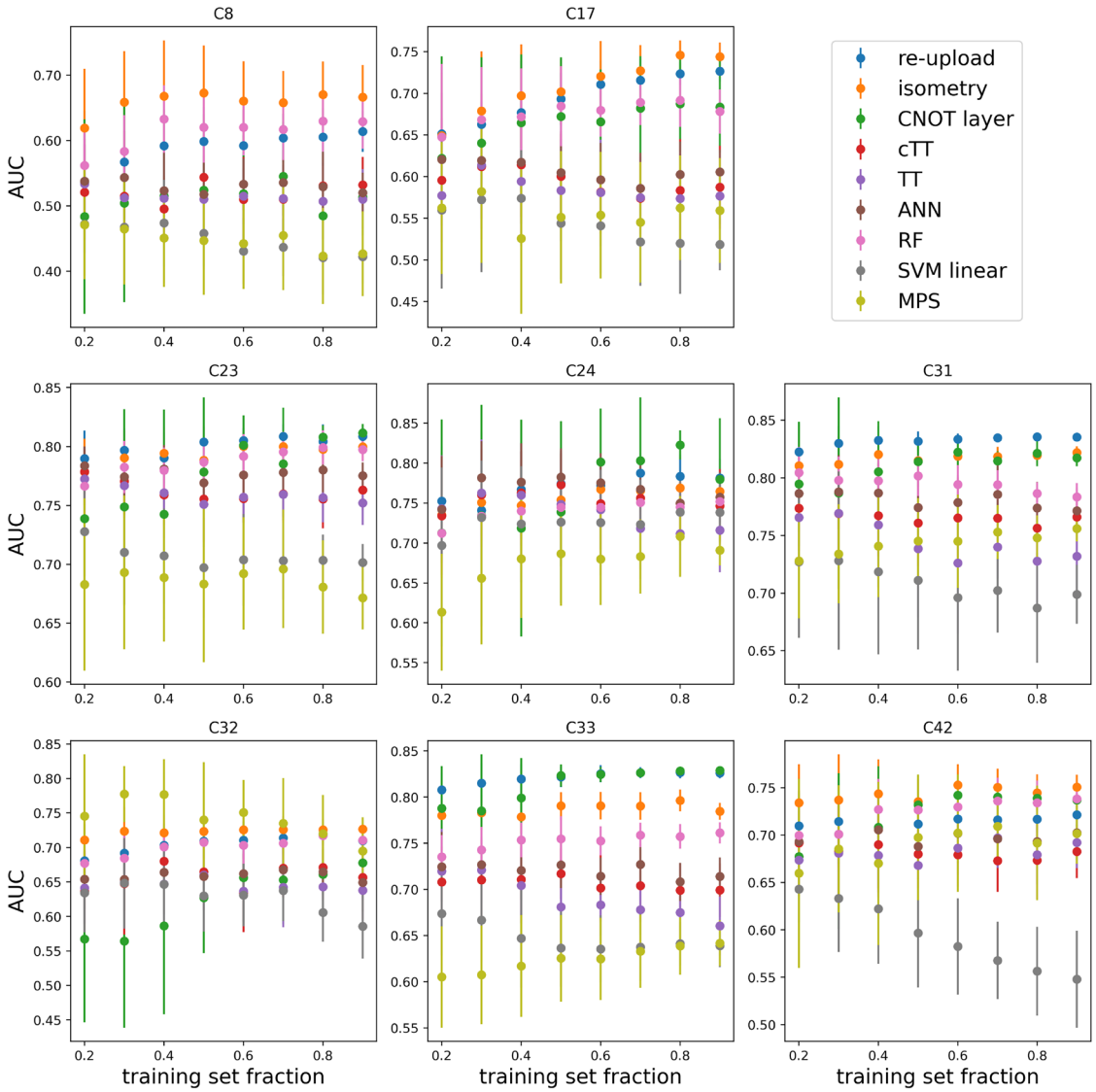


**Fig. 8** Independent set classification AUC for the eight gene communities showing the best performances of PQCs with respect to classical algorithms

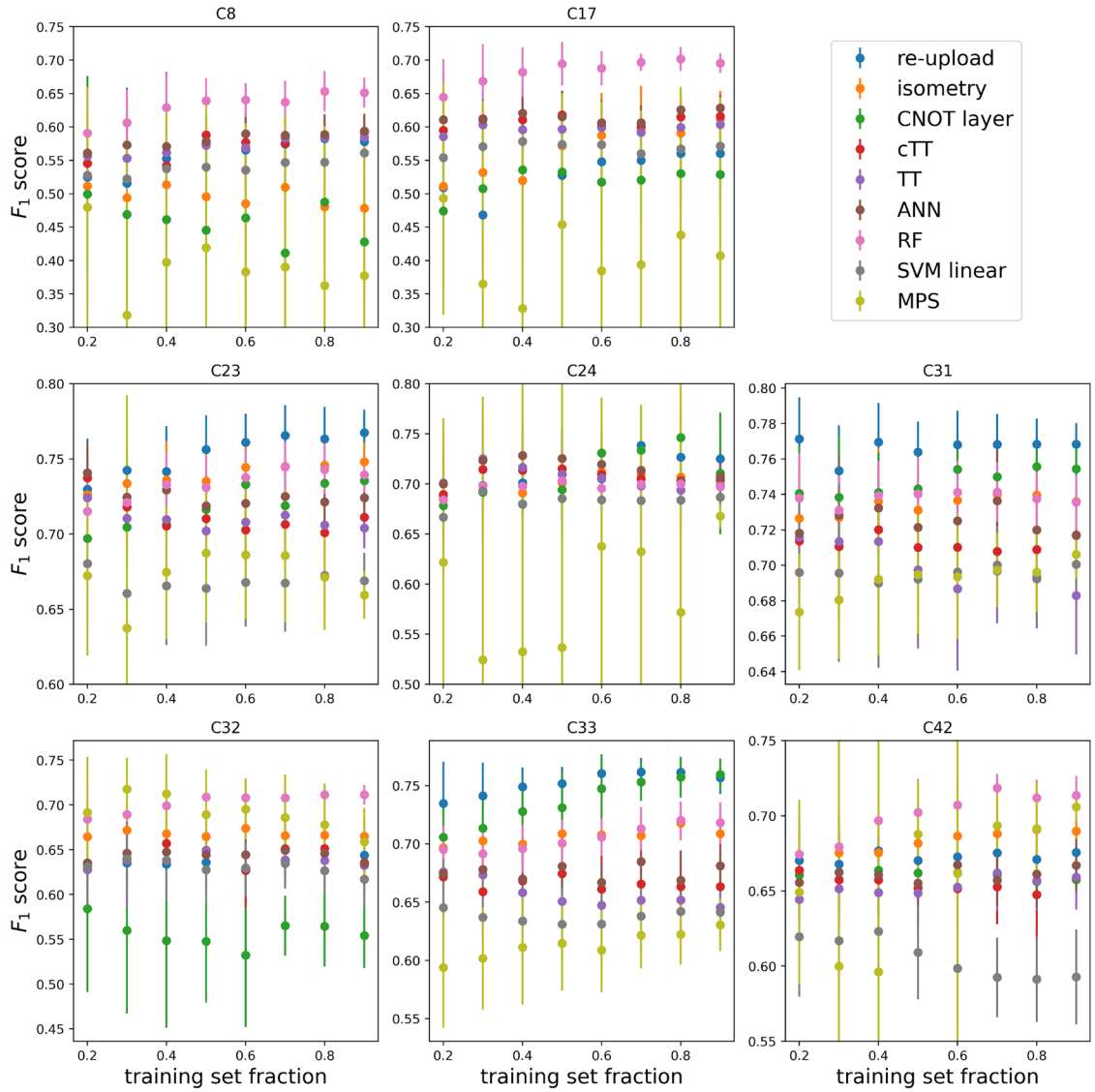


**Fig. 9** Independent set classification $F_1$ score for the eight gene communities showing the best performances of PQCs with respect to classical algorithms

## References

1. Arute F, et al. Quantum supremacy using a programmable superconducting processor. Nature. 2019;574:505–10. https://doi.org/10.1038/s41586-019-1666-5.
2. Giovannetti V, Lloyd S, Maccone L. Advances in quantum metrology. Nat Photon. 2011;5:222–9. https://doi.org/10.1038/nphoton.2011.35.
3. Ortolano G, Napoli C, Harney C, Pirandola S, Leonetti G, Boucher P, Losero E, Genovese M, Ruo-Berchera I. Quantum-enhanced pattern recognition. Phys Rev Appl. 2023;20:024072. https://doi.org/10.1103/PhysRevApplied.20.024072.
4. Lloyd S, Schuld M, Ijaz A, Izaac J, Killoran N. Quantum embeddings for machine learning. 2020. arXiv: https://arxiv.org/abs/2001.03622.
5. Tacchino F, Macchiavello C, Gerace D, Bajoni D. An artificial neuron implemented on an actual quantum processor. NPJ Quantum Inf. 2019. https://doi.org/10.1038/s41534-019-0140-4.
6. Schuld M, Sweke R, Meyer JJ. Effect of data encoding on the expressive power of variational quantum-machine-learning models. Phys Rev A. 2021;103:032430. https://doi.org/10.1103/PhysRevA.103.032430.
7. Pérez-Salinas A, Cervera-Lierta A, Gil-Fuster E, Latorre JI. Data re-uploading for a universal quantum classifier. Quantum. 2020;4:226. https://doi.org/10.22331/q-2020-02-06-226.

8. Larocca M, Sauvage F, Sbahi FM, Verdon G, Coles PJ, Cerezo M. Group-invariant quantum machine learning. PRX Quantum. 2022;3:030341. https://doi.org/10.1103/PRXQuantum.3.030341.
9. Meyer JJ, Mularski M, Gil-Fuster E, Mele AA, Arzani F, Wilms A, Eisert J. Exploiting symmetry in variational quantum machine learning. PRX Quantum. 2023;4:010328. https://doi.org/10.1103/PRXQuantum.4.010328.
10. Larocca M, Ju N, García-Martín D, Coles PJ, Cerezo M. Theory of over parametrization in quantum neural networks. Nat Comput Sci. 2023;3:542–51. https://doi.org/10.1038/s43588-023-00467-6.
11. McClean JR, Boixo S, Smelyanskiy VN, Babbush R, Neven H. Barren plateaus in quantum neural network training landscapes. Nat Commun. 2018;9:4812. https://doi.org/10.1038/s41467-018-07090-4.
12. Cerezo M, Sone A, Volkoff T, Cincio L, Coles PJ. Cost function dependent barren plateaus in shallow parametrized quantum circuits. Nat Commun. 2021;12:1791. https://doi.org/10.1038/s41467-021-21728-w.
13. Larocca M, Czarnik P, Sharma K, Muraleedharan G, Coles PJ, Cerezo M. Diagnosing Barren plateaus with tools from quantum optimal control. Quantum. 2022;6:824. https://doi.org/10.22331/q-2022-09-29-824.
14. Benatti F, Gramegna G, Mancini S. Pattern capacity of a single quantum perceptron. J Phys A Math Theoret. 2022;55(15):155301. https://doi.org/10.1088/1751-8121/ac58d1.
15. Sim S, Johnson PD, Aspuru-Guzik A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. Adv Quantum Technol. 2019;2:1900070. https://doi.org/10.1002/qute.201900070.
16. Rebentrost P, Mohseni M, Lloyd S. Quantum support vector machine for big data classification. Phys Rev Lett. 2014;113:130503. https://doi.org/10.1103/PhysRevLett.113.130503.
17. Schuld M, Killoran N. Quantum machine learning in feature Hilbert spaces. Phys Rev Lett. 2019;122:040504. https://doi.org/10.1103/PhysRevLett.122.040504.
18. Liu Y, Arunachalam S, Temme K. A rigorous and robust quantum speed-up in supervised machine learning. Nat Phys. 2021;17:1013–7. https://doi.org/10.1038/s41567-021-01287-z.
19. Schuld M. Supervised quantum machine learning models are kernel method. 2021. arXiv: https://arxiv.org/abs/2101.11020.
20. Mitarai K, Negoro M, Kitagawa M, Fujii K. Quantum circuit learning. Phys Rev A. 2018;98:032309. https://doi.org/10.1103/PhysRevA.98.032309.
21. Beer K, Bondarenko D, Farrelly T, Osborne TJ, Salzmann R, Scheiermann D, Wolf R. Training deep quantum neural networks. Nat Commun. 2020. https://doi.org/10.1038/s41467-020-14454-2.
22. Schuld M, Bocharov A, Svore KM, Wiebe N. Circuit-centric quantum classifiers. Phys Rev A. 2020;101:032308. https://doi.org/10.1103/PhysRevA.101.032308.
23. Pomarico D, Cosmai L, Facchi P, Lupo C, Pascazio S, Pepe FV. Dynamical quantum phase transitions of the Schwinger model: real-time dynamics on IBM quantum. Entropy. 2023;25(4):608. https://doi.org/10.3390/e25040608.
24. Huggins W, Patil P, Mitchell B, Whaley KB, Stoudenmire EM. Towards quantum machine learning with tensor networks. Quantum Sci Technol. 2019;4(2):024001. https://doi.org/10.1088/2058-9565/aaea94.
25. Felser T, Trenti M, Sestini L, Gianelle A, Zuliani D, Lucchesi D, Montangero S. Quantum-inspired machine learning on high-energy physics data. NPJ Quantum Inf. 2021. https://doi.org/10.1038/s41534-021-00443-w.
26. Dborin J, Barratt F, Wimalaweera V, Wright L, Green AG. Matrix product state pre-training for quantum machine learning. Quantum Sci Technol. 2022;7(3):035014. https://doi.org/10.1088/2058-9565/ac7073.
27. Ballarin M, Mangini S, Montangero S, Macchiavello C, Mengoni R. Entanglement entropy production in quantum neural networks. Quantum. 2023;7:1023. https://doi.org/10.22331/q-2023-05-31-1023.
28. Collura M, Dell'Anna L, Felser T, Montangero S. On the descriptive power of neural-networks as constrained tensor networks with exponentially large bond dimension. SciPost Phys Core. 2021;4:1. https://doi.org/10.21468/SciPostPhysCore.4.1.001.
29. Glasser I, Pancotti N, Cirac JI. From probabilistic graphical models to generalized tensor networks for supervised learning. IEEE Access. 2020;8:68169–82. https://doi.org/10.1109/ACCESS.2020.2986279.
30. Ran S-J. Encoding of matrix product states into quantum circuits of one- and two-qubit gates. Phys Rev A. 2020;101:032310. https://doi.org/10.1103/PhysRevA.101.032310.
31. Rudolph MS, Chen J, Miller J, Acharya A, Perdomo-Ortiz A. Decomposition of matrix product states into shallow quantum circuits. Quantum Sci Technol. 2023;9(1):015012. https://doi.org/10.1088/2058-9565/ad04e6.
32. Rudolph MS, Miller J, Motlagh D, Chen J, Acharya A, Perdomo-Ortiz A. Synergistic pretraining of parametrized quantum circuits via tensor networks. Nat Commun. 2023;14:8367. https://doi.org/10.1038/s41467-023-43908-6.
33. Vidal G. Entanglement renormalization. Phys Rev Lett. 2007;99:220405. https://doi.org/10.1103/PhysRevLett.99.220405.
34. Vidal G. Class of quantum many-body states that can be efficiently simulated. Phys Rev Lett. 2008;101:110501. https://doi.org/10.1103/PhysRevLett.101.110501.
35. Pomarico D. Multiscale entanglement renormalization ansatz: causality and error correction. Dynamics. 2023;3:622–35. https://doi.org/10.3390/dynamics3030033.
36. Carleo G, Troyer M. Solving the quantum many-body problem with artificial neural networks. Science. 2017;355:602–6. https://doi.org/10.1126/science.aag2302.
37. Tibaldi S, Magnifico G, Vodola D, Ercolessi E. Unsupervised and supervised learning of interacting topological phases from single-particle correlation functions. SciPost Phys. 2023;14:005. https://doi.org/10.21468/SciPostPhys.14.1.005.
38. Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L, Zdeborová L. Machine learning and the physical sciences. Rev Mod Phys. 2019;91:045002. https://doi.org/10.1103/RevModPhys.91.045002.
39. Li RY, Di Felice R, Rohs R, Lidar D. Quantum annealing versus classical machine learning applied to a simplified computational biology problem. NPJ Quantum Inf. 2018. https://doi.org/10.1038/s41534-018-0060-8.
40. Mott A, Job J, Vlimant J-R, Lidar D, Spiropulu M. Solving a Higgs optimization problem with quantum annealing for machine learning. Nature. 2017;550:375–9. https://doi.org/10.1038/nature24047.

41. Li RY, Gujja S, Bajaj SR, Gamel OE, Cilfone N, Gulcher JR, Lidar D, Chittenden TW. Quantum processor-inspired machine learning in the biomedical sciences. Patterns. 2021;2(6):100246. https://doi.org/10.1016/j.patter.2021.100246.
42. Kardashin AS, Vlasova AV, Pervishko AA, Yudin D, Biamonte JD. Quantum-machine-learning channel discrimination. Phys Rev A. 2022;106:032409. https://doi.org/10.1103/PhysRevA.106.032409.
43. Chen S, Zhou S, Seif A, Jiang L. Quantum advantages for Pauli channel estimation. Phys Rev A. 2022;105:032435. https://doi.org/10.1103/PhysRevA.105.032435.
44. Oh C, Chen S, Wong Y, Zhou S, Huang H-Y, Nielsen JAH, Liu Z-H, Neergaard-Nielsen JS, Andersen UL, Jiang L, Preskill J. Entanglement-enabled advantage for learning a bosonic random displacement channel. 2024. arXiv: https://arxiv.org/abs/2402.18809.
45. Lloyd S, Mohseni M, Rebentrost P. Quantum principal component analysis. Nature Phys. 2014;10:631–3. https://doi.org/10.1038/nphys3029.
46. Huang H-Y, Broughton M, Cotler J, Chen S, Li J, Mohseni M, Neven H, Babbush R, Kueng R, Preskill J, McClean JR. Quantum advantage in learning from experiments. Science. 2023;376:1182–6. https://doi.org/10.1126/science.abn7293.
47. Lazarev ID, Narozniak M, Byrnes T, Pyrkov AN. Hybrid quantum-classical unsupervised data clustering based on the Self-Organizing Feature Map. 2023. arXiv: https://arxiv.org/abs/2009.09246.
48. Lombardi A, Diacono D, Amoroso N, Monaco A, Tavares JMRS, Bellotti R, Tangaro S. Explainable deep learning for personalized age prediction with brain morphology. Front Neurosci. 2021. https://doi.org/10.3389/fnins.2021.674055.
49. Massafra R, et al. Analyzing breast cancer invasive disease event classification through explainable artificial intelligence. Front Med. 2023. https://doi.org/10.3389/fmed.2023.1116354.
50. Amoroso N, Quarto S, La Rocca M, Tangaro S, Monaco A, Bellotti R. An eXplainability artificial intelligence approach to brain connectivity in Alzheimer's disease. Front Aging Neurosci. 2023. https://doi.org/10.3389/fnagi.2023.1238065.
51. Lacalamita A, Serino G, Pantaleo E, Monaco A, Amoroso N, Bellantuono L, Piccinno E, Scalavino V, Dituri F, Tangaro S, Bellotti R, Giannelli G. Artificial intelligence and complex network approaches reveal potential gene biomarkers for hepatocellular carcinoma. Int J Mol Sci. 2023;24(20):15286. https://doi.org/10.3390/ijms242015286.
52. Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow JM, Gambetta JM. Hardware-efficient variational quantum Eigensolver for small molecules and quantum magnets. Nature. 2017;549:242–6. https://doi.org/10.1038/nature23879.
53. Perdomo-Ortiz A, Dickson N, Drew-Brook M, Rose G, Aspuru-Guzik A. Finding low-energy conformations of lattice protein models by quantum annealing. Sci Rep. 2012. https://doi.org/10.1038/srep00571.
54. Robert A, Barkoutsos PK, Woerner S, Tavernelli I. Resource-efficient quantum algorithm for protein folding. NPJ Quantum Inf. 2021. https://doi.org/10.1038/s41534-021-00368-4.
55. Nakatani N, Chan GK-L. Efficient tree tensor network states (TTNS) for quantum chemistry: generalizations of the density matrix renormalization group algorithm. J Chem Phys. 2013;138(13):134113. https://doi.org/10.1063/1.4798639.
56. Murg V, Verstraete F, Schneider R, Nagy PR, Legeza O. Tree tensor network state with variable tensor order: an efficient multireference method for strongly correlated systems. J Chem Theory Comput. 2015;11(3):1027–36. https://doi.org/10.1021/ct501187j.
57. Batra K, Zorn KM, Foil DH, Minerali E, Gawriljuk VO, Lane TR, Ekins S. Quantum machine learning algorithms for drug discovery applications. J Chem Inf Model. 2021;61(6):2641–7. https://doi.org/10.1021/acs.jcim.1c00166.
58. Willsch D, Willsch M, De Raedt H, Michielsen K. Support vector machines on the D-wave quantum Annealer. Comput Phys Commun. 2020;248:107006. https://doi.org/10.1016/j.cpc.2019.107006.
59. Maheshwari D, Garcia-Zapirain B, Sierra-Sosa D. Quantum machine learning applications in the biomedical domain: a systematic review. IEEE Access. 2022;19:80463–84. https://doi.org/10.1109/ACCESS.2022.3195044.
60. Rahimi M, Farkhondeh A. Oncological applications of quantum machine learning. Tech Canc Res Treat. 2023. https://doi.org/10.1177/15330338231215214.
61. Pyrkov A, Aliper A, Bezrukov D, Podolskiy D, Ren F, Zhavoronkov A. Complexity of life sciences in quantum and AI era. WIREs Comput Mol Sci. 2024;14(1):1701. https://doi.org/10.1002/wcms.1701.
62. Pomarico D, Fanizzi A, Amoroso N, Bellotti R, Biafora A, Bove S, Didonna V, La Forgia D, Pastena MI, Tamborra P, Zito A, Lorusso V, Massafra R. A proposal of quantum-inspired machine learning for medical purposes: an application case. Mathematics. 2021. https://doi.org/10.3390/math9040410.
63. Jain S, Ziauddin J, Leonchyk P, Yenkanchi S, Geraci J. Quantum and classical machine learning for the classification of non-small-cell lung cancer patients. SN Appl Sci. 2020. https://doi.org/10.1007/s42452-020-2847-4.
64. Dabba A, Tari A, Meftali S. Hybridization of moth flame optimization algorithm and quantum computing for gene selection in microarray data. J Ambient Intell Human Comput. 2021;12:2731–50. https://doi.org/10.1007/s12652-020-02434-9.
65. Sergioli G, Militello C, Rundo L, Minafra L, Torrisi F, Russo G, Chow KL, Giuntini R. A quantum-inspired classifier for clonogenic assay evaluations. Sci Rep. 2021. https://doi.org/10.1038/s41598-021-82085-8.
66. Cavinato S, Felser T, Fusella M, Paiusco M, Montangero S. Optimizing radiotherapy plans for cancer treatment with tensor networks. Phys Med Biol. 2021;66(12):125015. https://doi.org/10.1088/1361-6560/ac01f2.
67. Esposito F, Gillis N, Del Buono N. Orthogonal joint sparse NMF for microarray data analysis. J Math Biol. 2019;79:223–47. https://doi.org/10.1007/s00285-019-01355-2.
68. Selicato L, Esposito F, Gargano G, Vegliante MC, Opinto G, Zaccaria GM, Ciavarella S, Guarini A, Del Buono N. A new ensemble method for detecting anomalies in gene expression matrices. Mathematics. 2021;9(8):882. https://doi.org/10.3390/math9080882.
69. Bergholm V, et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. 2022. arXiv: https://arxiv.org/abs/1811.04968.
70. Efthymiou S, Hidary J, Leichenauer S. TensorNetwork for Machine Learning. 2019. arXiv: https://arxiv.org/abs/1906.06329.
71. Selvan R, Dam EB. tensor networks for medical image classification. In: Proceedings of the third conference on medical imaging with deep learning. 2020. https://proceedings.mlr.press/v121/selvan20a.html.
72. Selvan R, Dam EB, Petersen J. Segmenting two-dimensional structures with strided tensor networks. In: International Conference on Information Processing in Medical Imaging. 2021. https://doi.org/10.1007/978-3-030-78191-0_31.
73. Miller J. TorchMPS. GitHub. 2019. https://github.com/jemisjoky/torchmps.

74. Tucker LR. Some mathematical notes on three-mode factor analysis. Psychometrika. 1966;31:279–311. https://doi.org/10.1007/BF02289464.
75. Oseledets IV. Tensor-train decomposition. SIAM J Sci Comput. 2011;33(5):2295–317. https://doi.org/10.1137/090752286.
76. Novikov A, Podoprikhin D, Osokin A, Vetrov D. Tensorizing neural networks. MIT Press, Cambridge, MA, USA. In: Proceedings of the 28th International Conference on Neural Information Processing Systems, Vol 1. 2015. https://dl.acm.org/doi/10.5555/2969239.2969289.
77. Novikov A, Izmailov P, Khrulkov V, Figurnov M, Oseledets I. Tensor Train Decomposition on TensorFlow (T3F). 2020. http://jmlr.org/papers/v21/18-008.html.
78. Novikov A, Izmailov P, Khrulkov V, Figurnov M, Oseledets I. Tensor Train Decomposition on TensorFlow (T3F). GitHub. https://github.com/Bihaqo/t3f.
79. Caro MC, Huang H, Cerezo M, Sharma K, Sornborger A, Cincio L, Coles PJ. Generalization in quantum machine learning from few training data. Nat Commun. 2022;13:4919. https://doi.org/10.1038/s41467-022-32550-3.
80. Gibbs J, Holmes Z, Caro MC, Ezzell N, Huang H-Y, Cincio L, Sornborger AT, Coles PJ. Dynamical simulation via quantum machine learning with provable generalization. Phys Rev Res. 2024;6:013241. https://doi.org/10.1103/PhysRevResearch.6.013241.
81. Peters E, Schuld M. Generalization despite overfitting in quantum machine learning models. Quantum. 2023;7:1210. https://doi.org/10.22331/q-2023-12-20-1210.
82. Bowles J, Wright VJ, Farkas M, Killoran N, Schuld M. Contextuality and inductive bias in quantum machine learning. 2023. arXiv: https://arxiv.org/abs/2302.01365.
83. Gil-Fuster E, Eisert J, Bravo-Prieto C. Understanding quantum machine learning also requires rethinking generalization. Nat Commun. 2024. https://doi.org/10.1038/s41467-024-45882-z.
84. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273–97.
85. Breiman L. Random forests. Mach Learn. 2001;45:5–32.
86. Jain AK, Mao J, Mohiuddin KM. Artificial neural networks: a tutorial. Computer. 1996;29(3):31–44.
87. ReCaS Bari: Data Center. https://www.recas-bari.it/index.php/en/