Identifying the Higgs boson production with Quantum Classifiers

Vasilis Belis (ETH Zürich, CERN), Samuel González-Castillo (University of Oviedo), Christina Reissel (ETH Zürich), Sofia Vallecorsa (CERN), Elías F. Combarro (University of Oviedo, CERN)

Monday, 26th of April 2021

A brief description of the problem

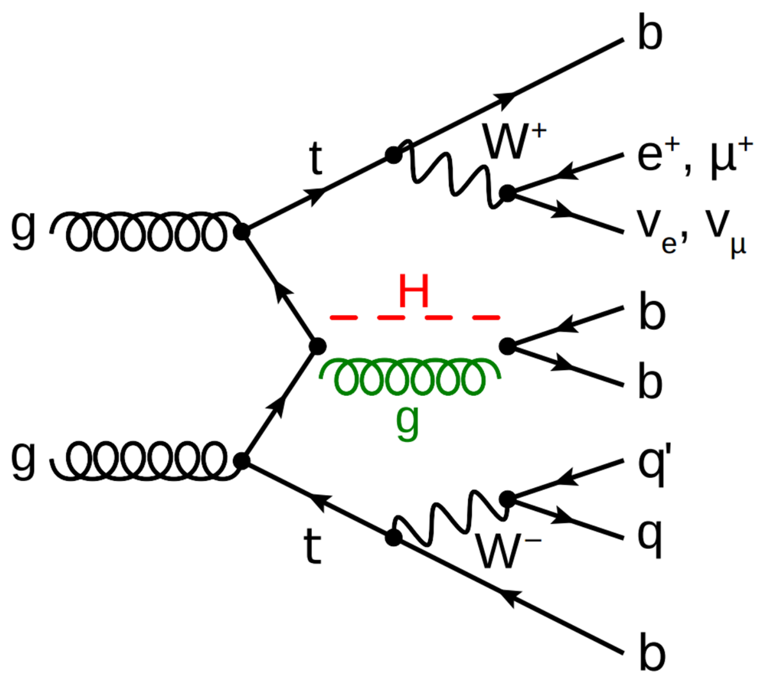
Identifying the Higgs boson production.

Signal vs background

The final state is the same!

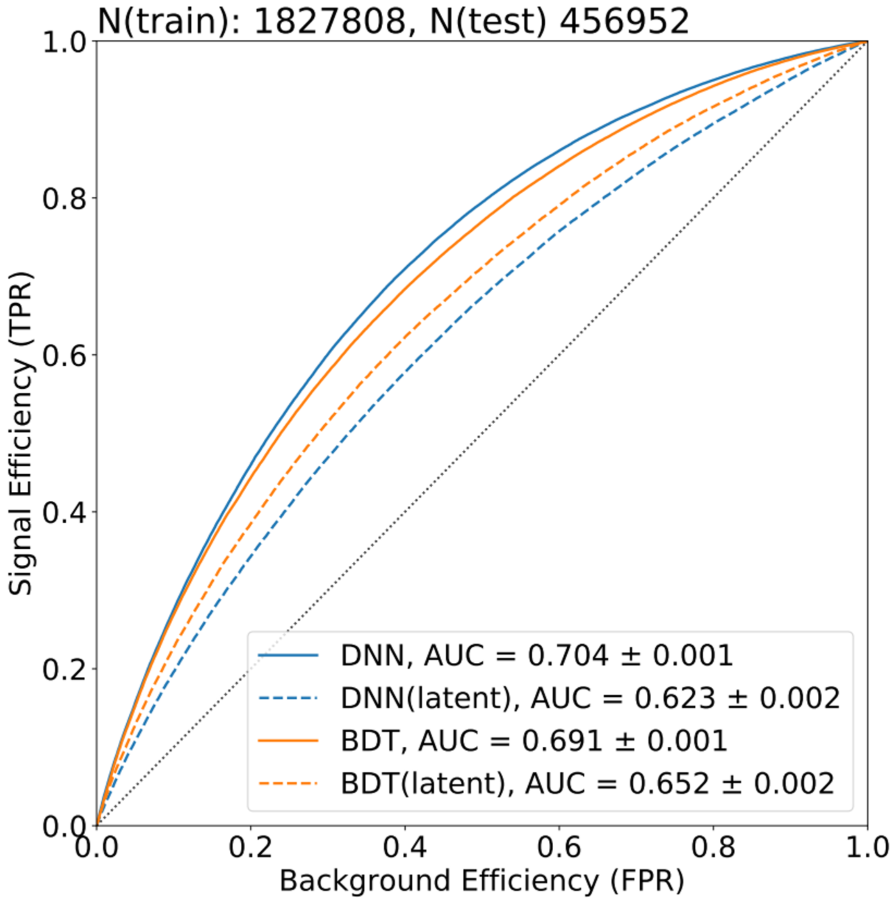
We use 67 physical variables.

7 Jets + MET + Leptons



Performance of classical methods

Using large training datasets

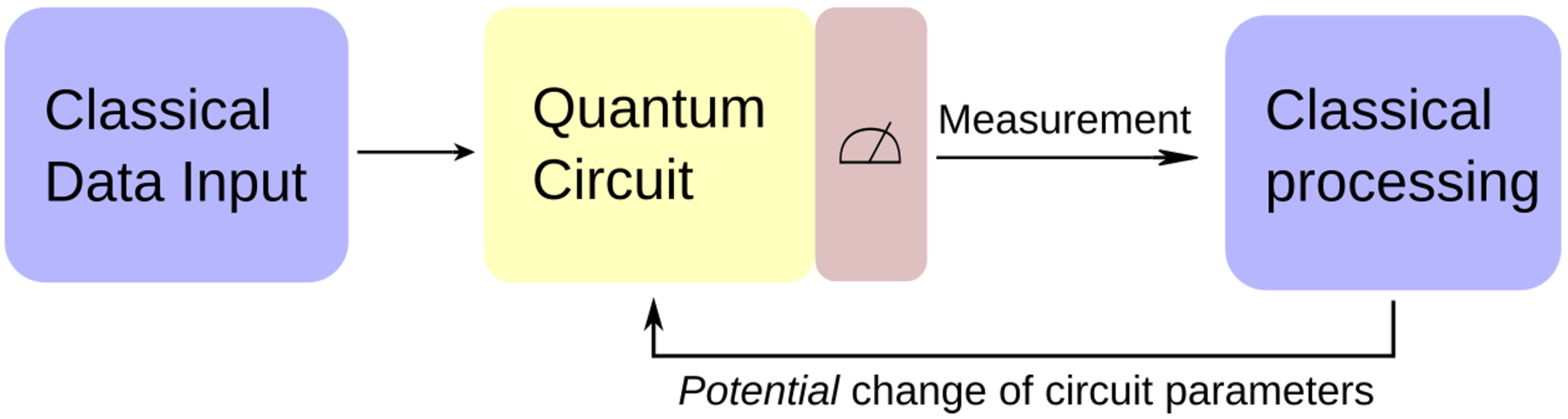


- Deep neural networks (DNN)
- Boosted decision trees (BDT)

- Trained on the original 67 features,
- and on a reduced feature space.

Hybrid quantum-classical ML models

Suitable for noisy intermediate-scale quantum devices.



We have worked with two models:

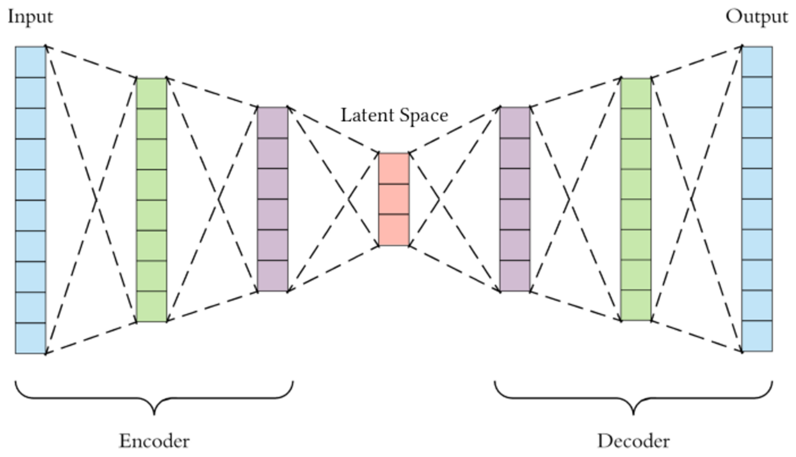
- Kernel methods → Quantum Support Vector Machines (QSVMs).
- Neural networks → Variational Quantum Circuits (VQCs).

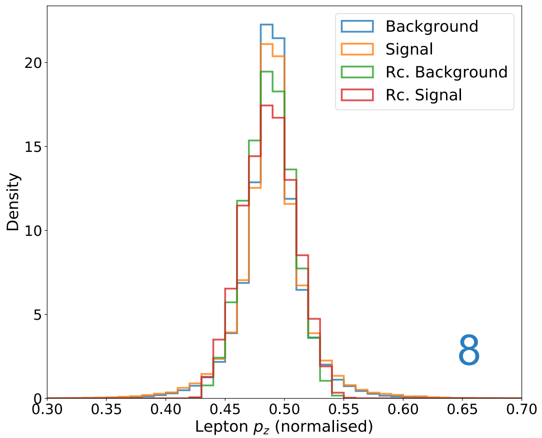
Data preparation

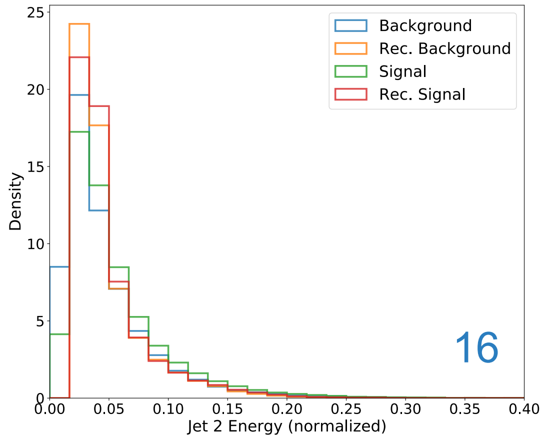
Two different approaches

1. Autoencoders

- Two autoencoders: one with 16 latent space features and one with 8.







2. AUC

- We picked the 8, 16 original variables that had the highest discriminative power according to their AUC score.

Quantum support vector machines

The SVM optimisation problem:

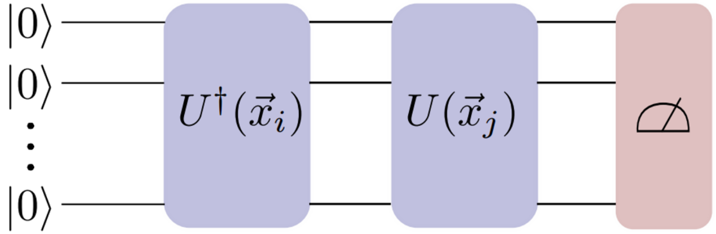
$$\text{maximize } L(c_1 \ldots c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\vec{x}_i \cdot \vec{x}_j) y_j c_j,$$

$$\text{subject to } \sum_{i=1}^{n} c_i y_i = 0, \text{ and } 0 \leq c_i \leq \frac{1}{2n\lambda} \equiv C \text{ for all } i.$$

Kernel substitution trick:

$$(\vec{x}_i \cdot \vec{x}_j) \rightarrow k(\vec{x}_i, \vec{x}_j) \equiv \phi(\vec{x}_i) \cdot \phi(\vec{x}_j)$$

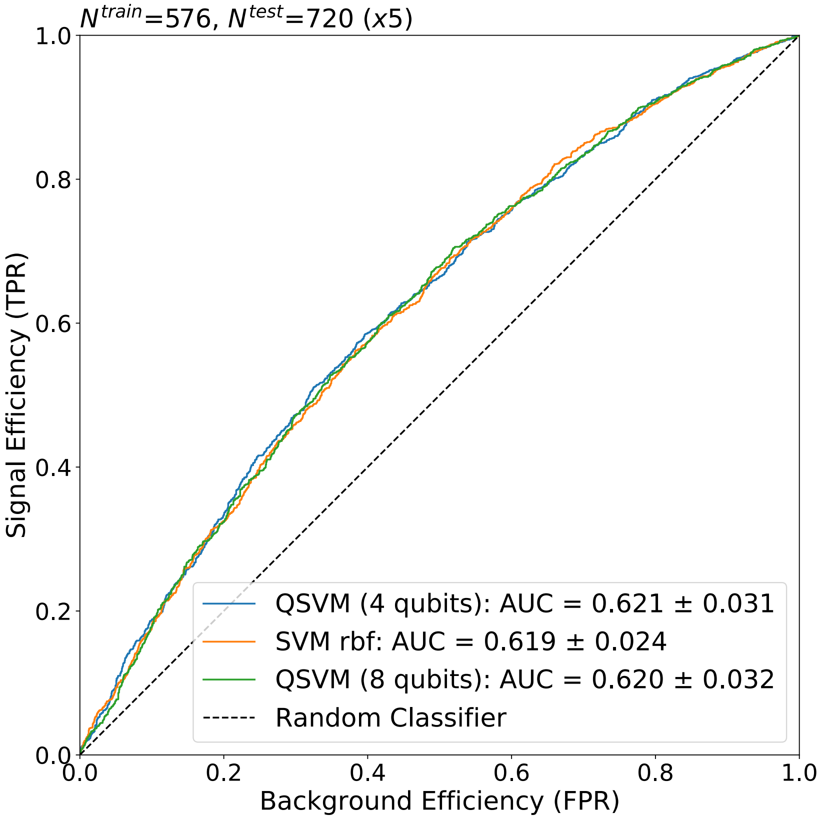
We can use a quantum kernel that would be computed on a quantum device



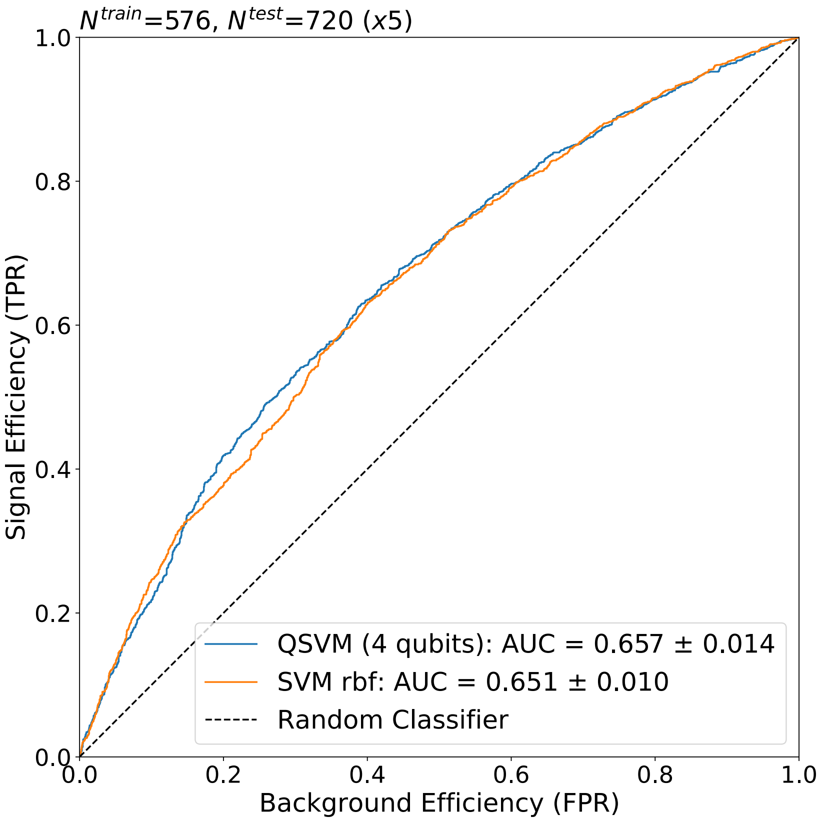
$$\Rightarrow K_{ij} = |\langle 0|U^\dagger(\vec{x}_i) U(\vec{x}_j)|0\rangle|^2$$

while the optimisation process would remain classical.

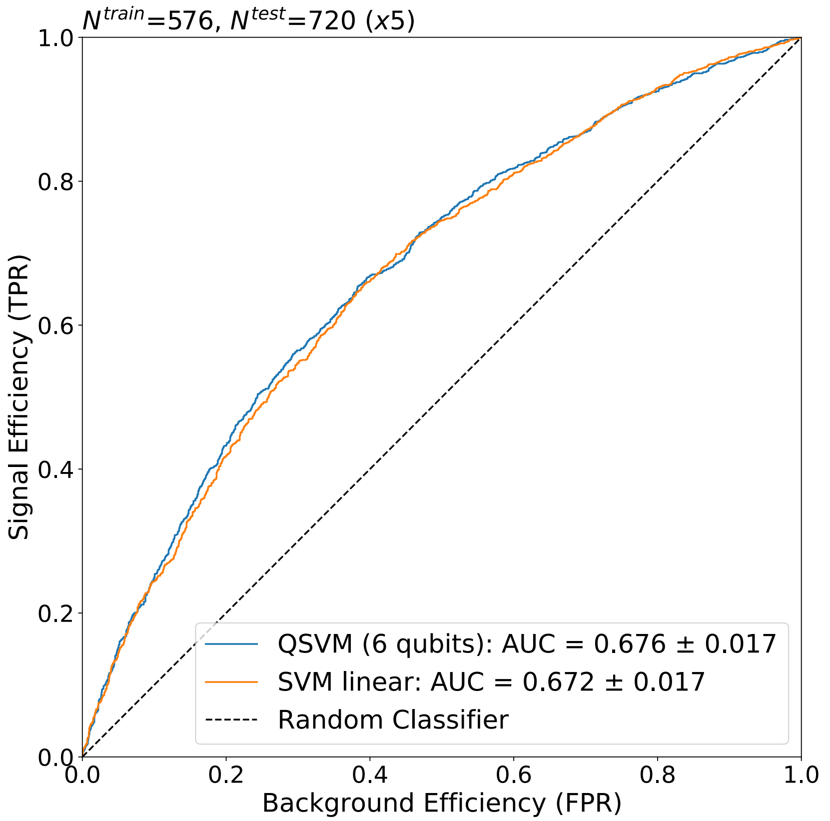
QSVM Results



AE (16 features)

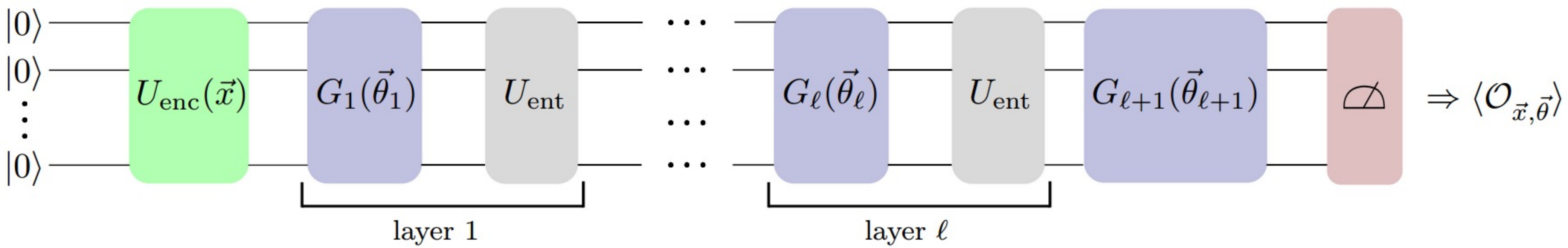


16 original features (AUC)



64 original features (AUC)

Variational quantum circuits



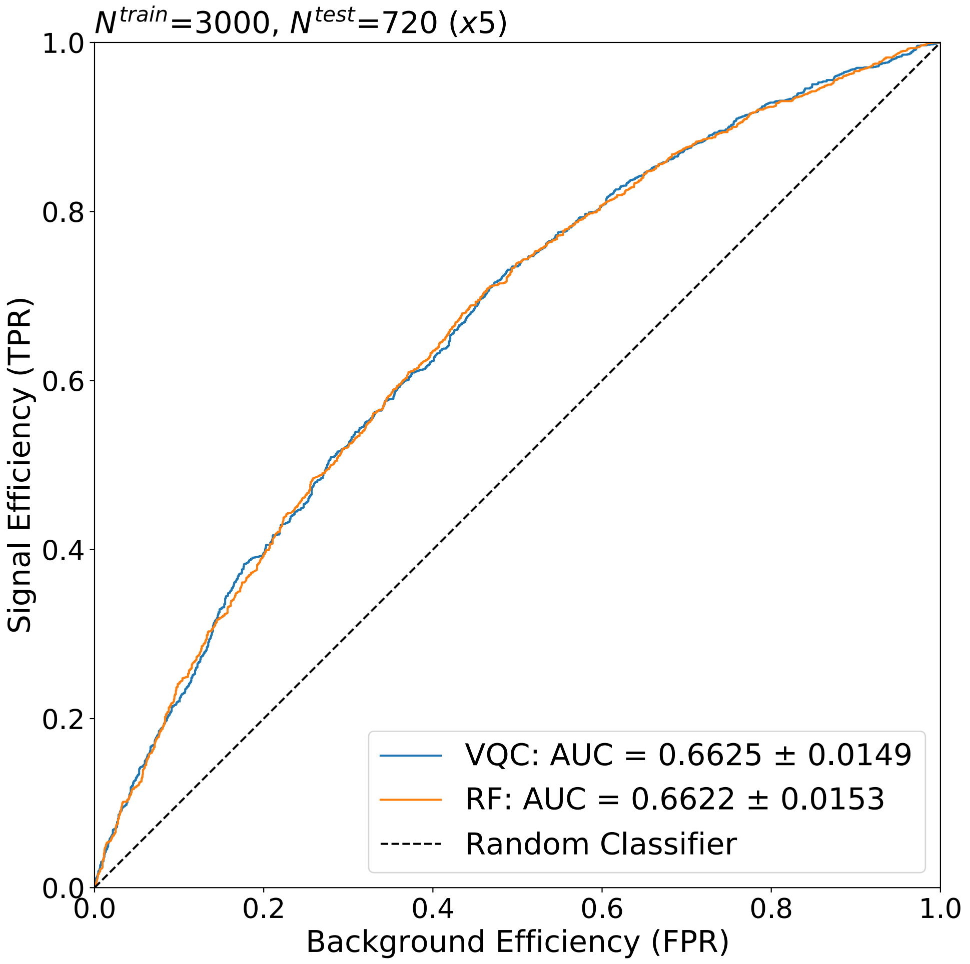
We measure the first qubit and minimise the binary cross entropy loss.

Data reuploading

Several repetitions of the VQC scheme before the measurement.

We can use more variables.

VQC Results



Things to be explored:

- Reducing the size of the training set.
- Different architectures.
- Different number of input features.
- Further adjustments.

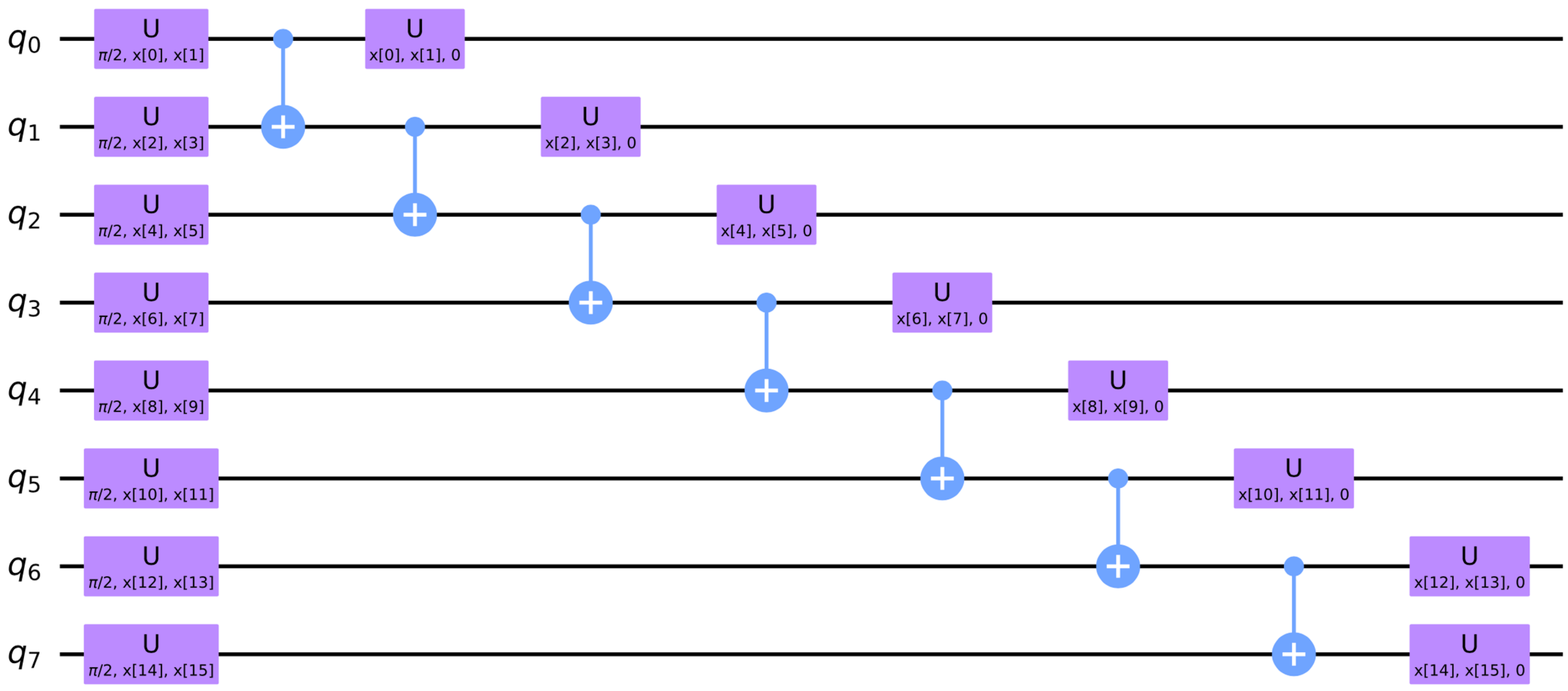
Future studies and outlook

1. Systematic study of data embedding circuits (feature maps).
   - Optimisation for their discrimination power in the quantum Hilbert space.

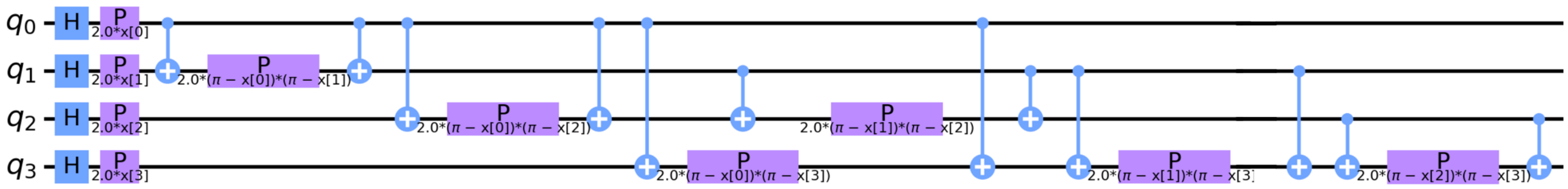
2. Investigation of other input feature reduction methods.
   - Aim for less information loss (classification power) in the reduced feature space.

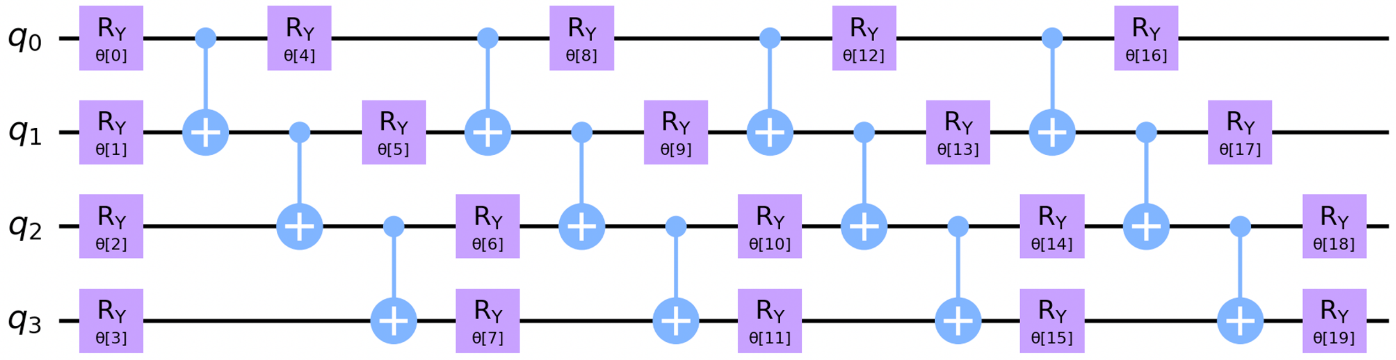
3. Implementation of developed algorithms on NISQ devices.
   - Design algorithms with a limited number of qubits, with a limited number of operations and that are robust against hardware noise.

QSVM feature map



VQC circuits





Full results

Feature selection + Model	AUC
AUC + QSVM	0.66 ± 0.01
PyTorch AE + QSVM	0.62 ± 0.03
AUC + SVM rbf	0.65 ± 0.01
PyTorch AE + SVM rbf	0.62 ± 0.02
KMeans + SVM rbf	0.61 ± 0.02

(a) 16 input variables

Feature selection + Model	AUC
AUC + QSVM	0.68 ± 0.02
AUC + Linear SVM	0.67 ± 0.02
Logistic Regression	0.68 ± 0.02

(b) 64 (QSVM, LSVM) and 67 (LR) input variables

Feature selection + Model	AUC
AUC + VQC	0.66 ± 0.01
AUC + Random Forest	0.66 ± 0.02
KMeans + Log. Regr.	0.64 ± 0.01
TensorFlow AE + AdaBoost	0.63 ± 0.03