Higgs analysis with quantum classifiers

vCHEP 2021

Vasilis Belis (ETH Zurich,CERN)

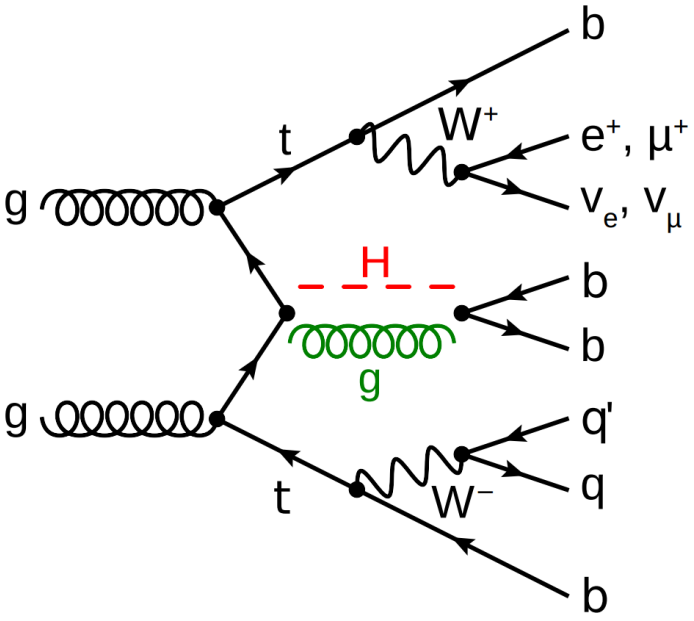
Samuel G. Castillo (U. Oviedo), Christina Reissel (ETH Zurich), Sofia Vallecorsa (CERN), Elias F. Combarro (U. Oviedo,CERN)

Günther Dissertori (ETH Zurich), Florentin Reiter (ETH Zurich)

Physical analysis and the classification task

Identify the Higgs boson production

Signal vs background



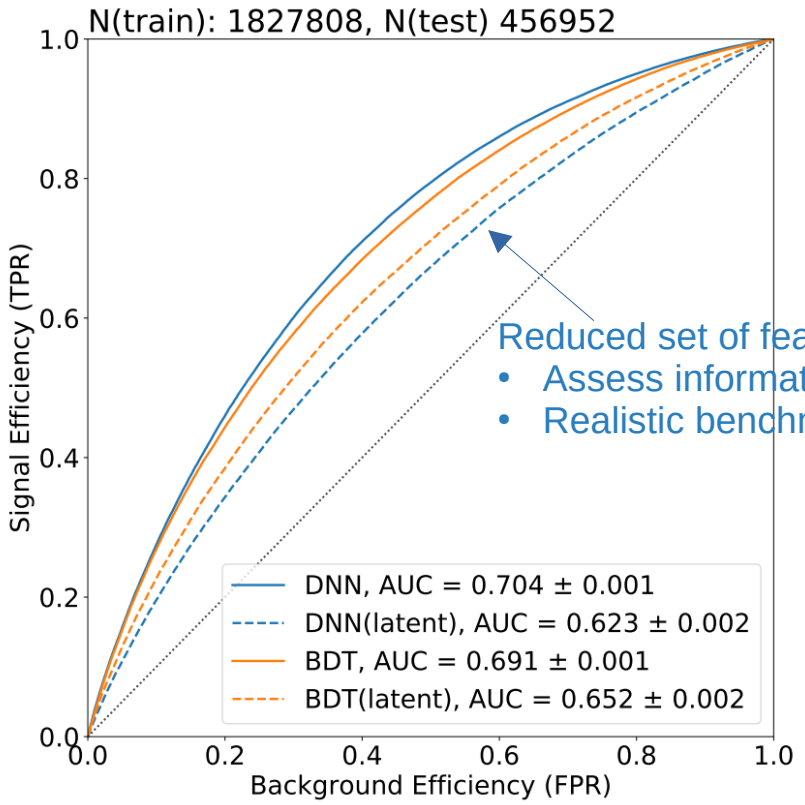
Features per event: (p_T, E, b-tag, etc.)

$$67 = 8 \times 7\,(\text{jets}) + 7\,(1\,\text{lepton}) + 4\,(\text{MET})$$

Typical approaches in HEP :

→ Deep Neural Networks (DNN)

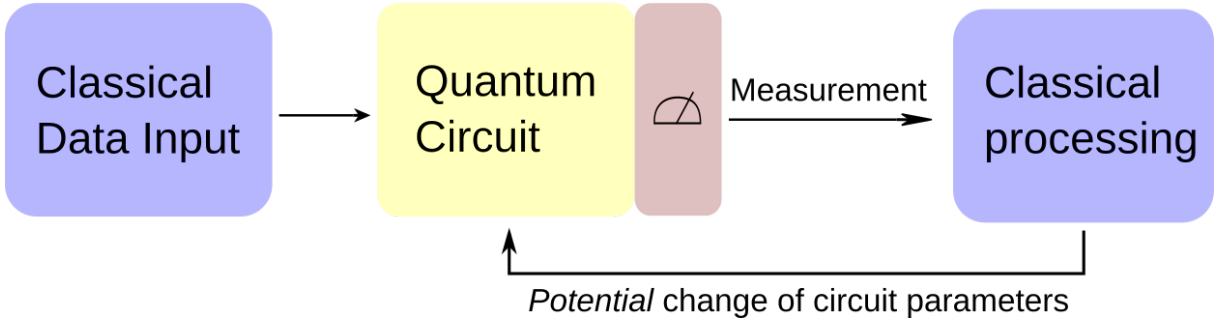
→ Boosted Decision Trees (BDT)



Reduced set of features (16)
- Assess information loss
- Realistic benchmark for QML

Hybrid Quantum-Classical machine learning models

Implementing quantum algorithms on *Noisy Intermediate Scale Quantum* (NISQ) devices :



→ Limited number of qubits

→ Compact circuits (circuit depth and qubit connectivity)

Quantum Machine Learning models for classification :

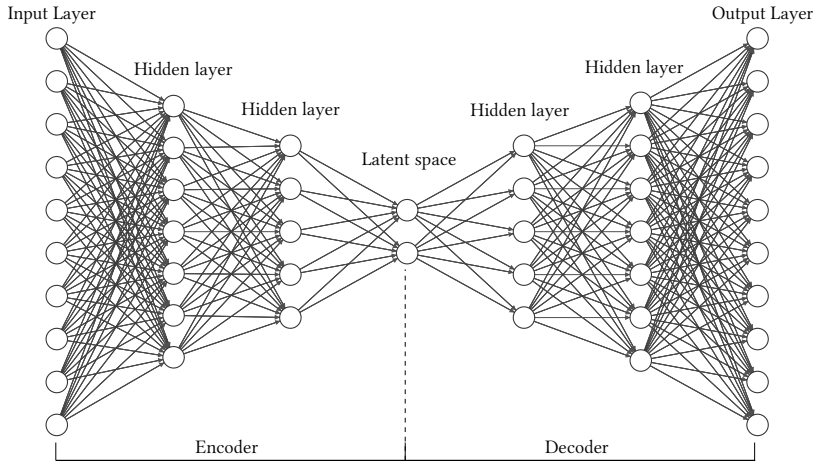
- Kernel methods → Quantum Support Vector Machines (QSVM)
- Quantum Neural Networks → Variational Quantum Circuits (VQC)

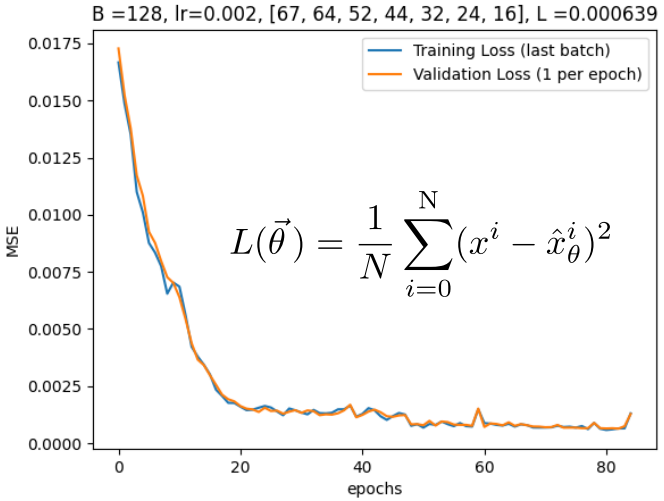
Feature reduction

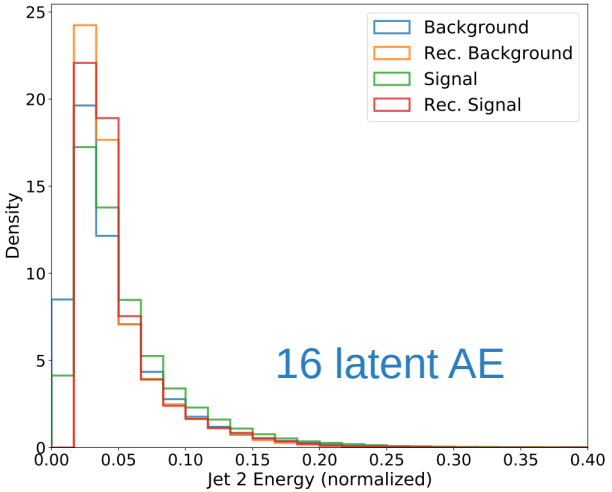
Two different approaches

1. Auto-Encoders (AE)

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







2. Feature selection

- We picked the 8, 16 original variables that had the highest discriminative power according to their AUC score (Area Under Receiver Operating Characteristic curve).

Quantum Support Vector Machines

SVM quadratic optimization problem :

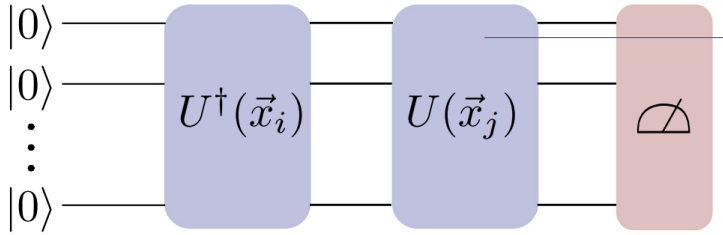
$$\text{maximize } L(c_1 \dots c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\vec{x}_i \cdot \vec{x}_j) y_j c_j,$$

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

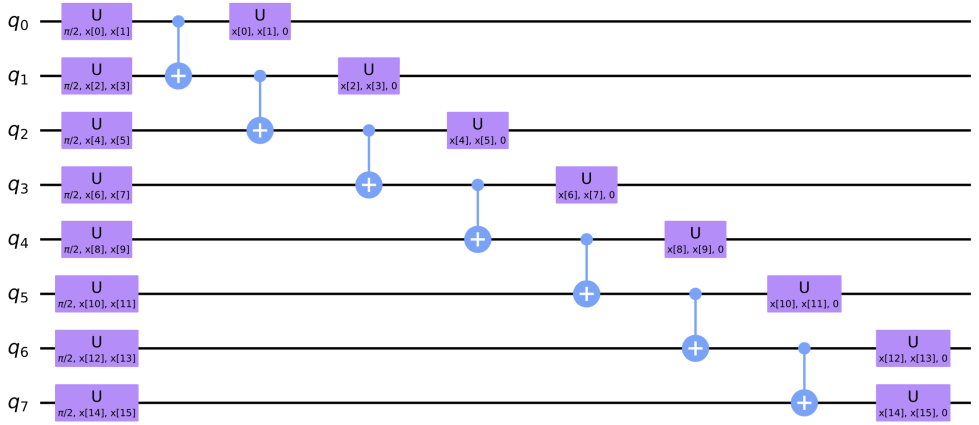
→ Kernel substitution trick :

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

Substitute Kernel with a *quantum* one !



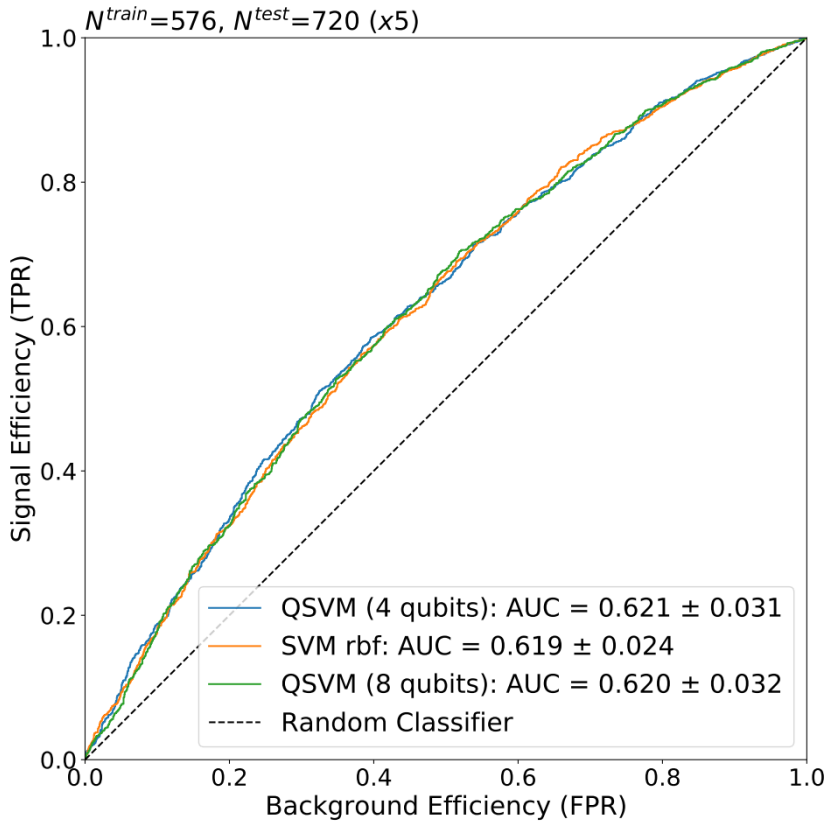
$$\Rightarrow K_{ij} = |\langle 0|U^\dagger(\vec{x}_i)U(\vec{x}_j)|0\rangle|^2$$



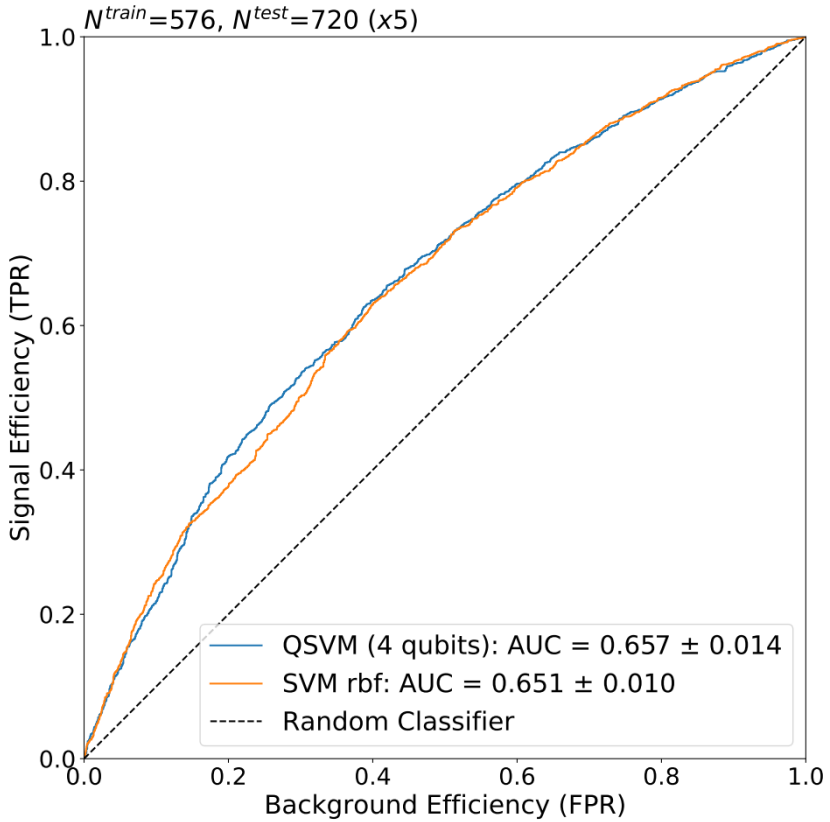
→ Sample kernel matrix with a quantum device

→ Maximize objective function of SVM on a classical computer

QSVM Results

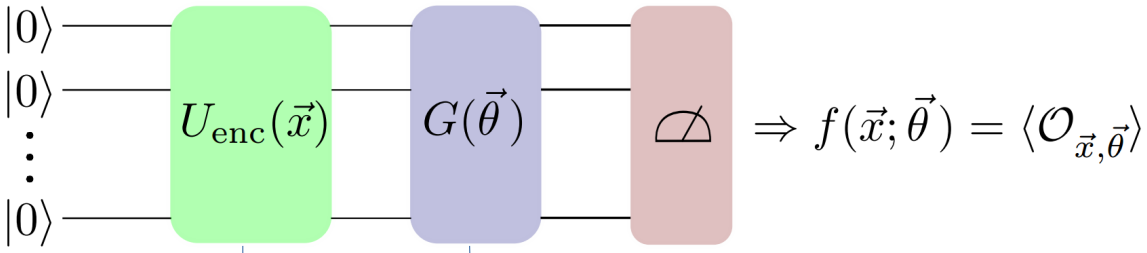


AE (16 features)

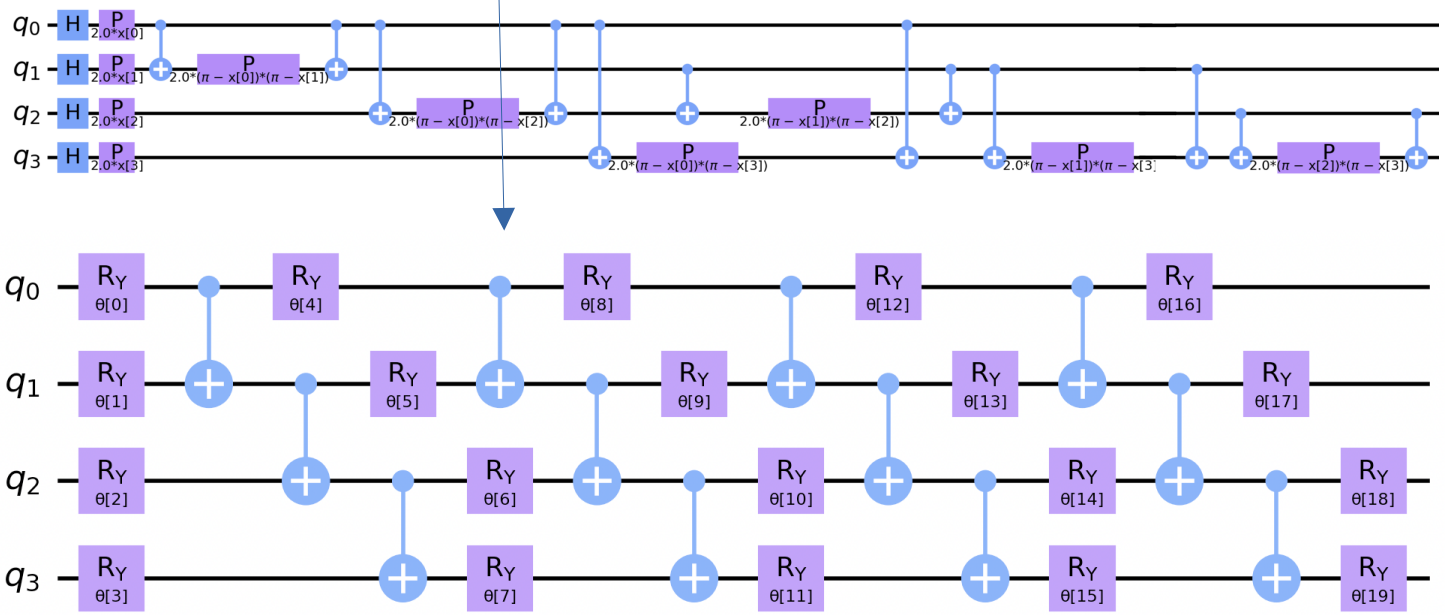


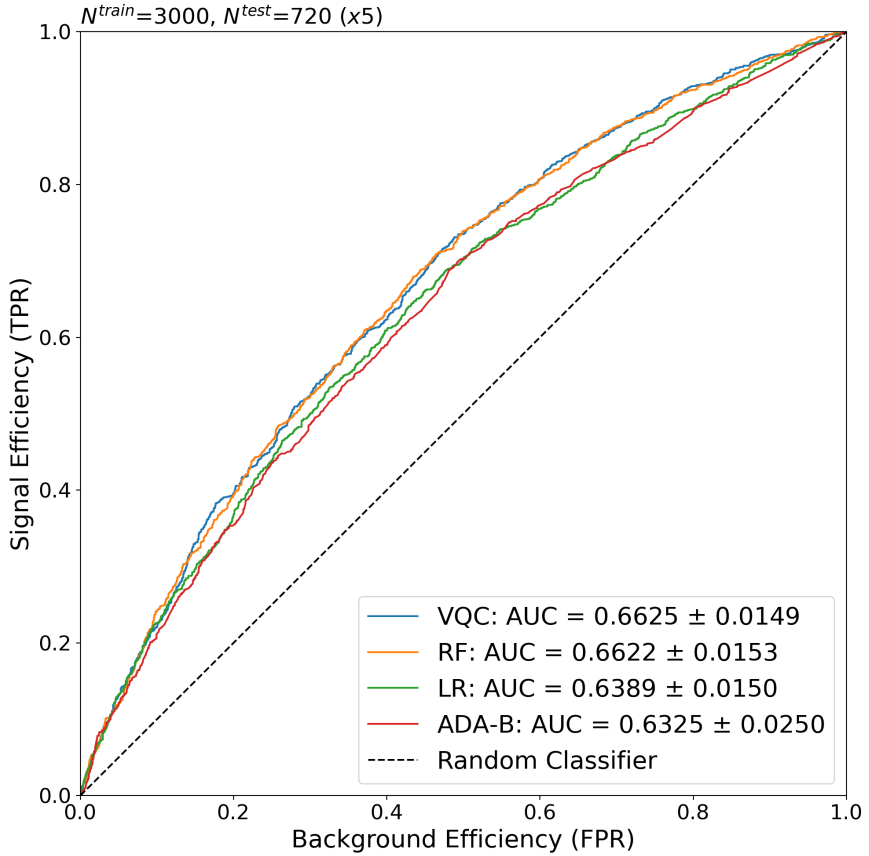
16 original features (AUC)

Variational Quantum Circuits



We measure the first qubit (z-basis) and minimise the binary cross entropy loss.





Future studies and outlook

1. Implementation of developed algorithms on NISQ devices

   → Design algorithms with limited number of qubits, limited number of operation and robust against hardware noise

2. Investigation of other input feature reduction methods

   → Aim for less information loss (classification power) in the reduced space

3. Systematic study of data embedding circuits (feature maps)

   → Optimization for their discrimination power in the quantum Hilbert space

Thank you!

Questions ?

(arXiv:2104.07692)

Back up

Feature reduction and classical models 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

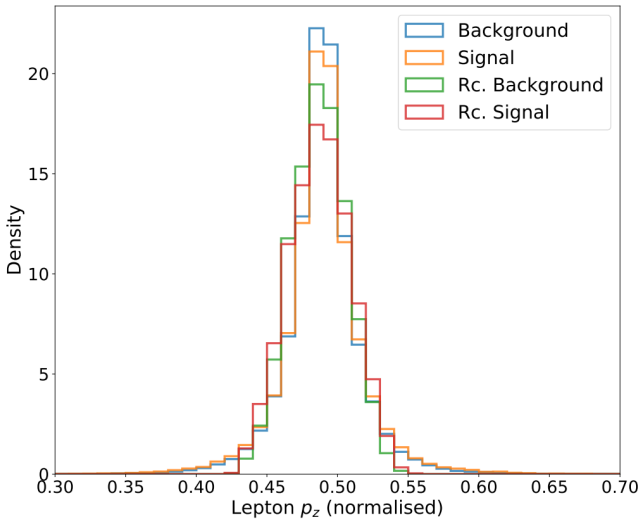
(b) 64 (QSVM, LSVM) and 67 (LR) input variables

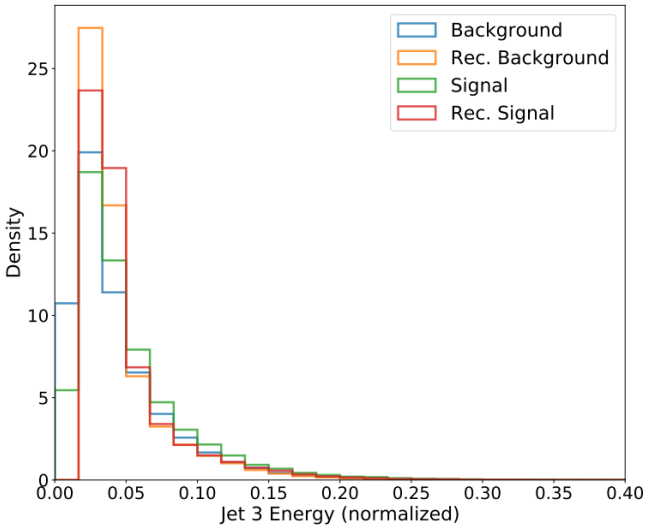
Feature selection + Model	AUC
AUC + VQC	0.66 ± 0.01
AUC + Random Forest	0.66 ± 0.02
KMeans + Log. Regr.	0.64 ± 0.01
TensorFlow AE + AdaBoost	0.63 ± 0.03

Auto-Encoder

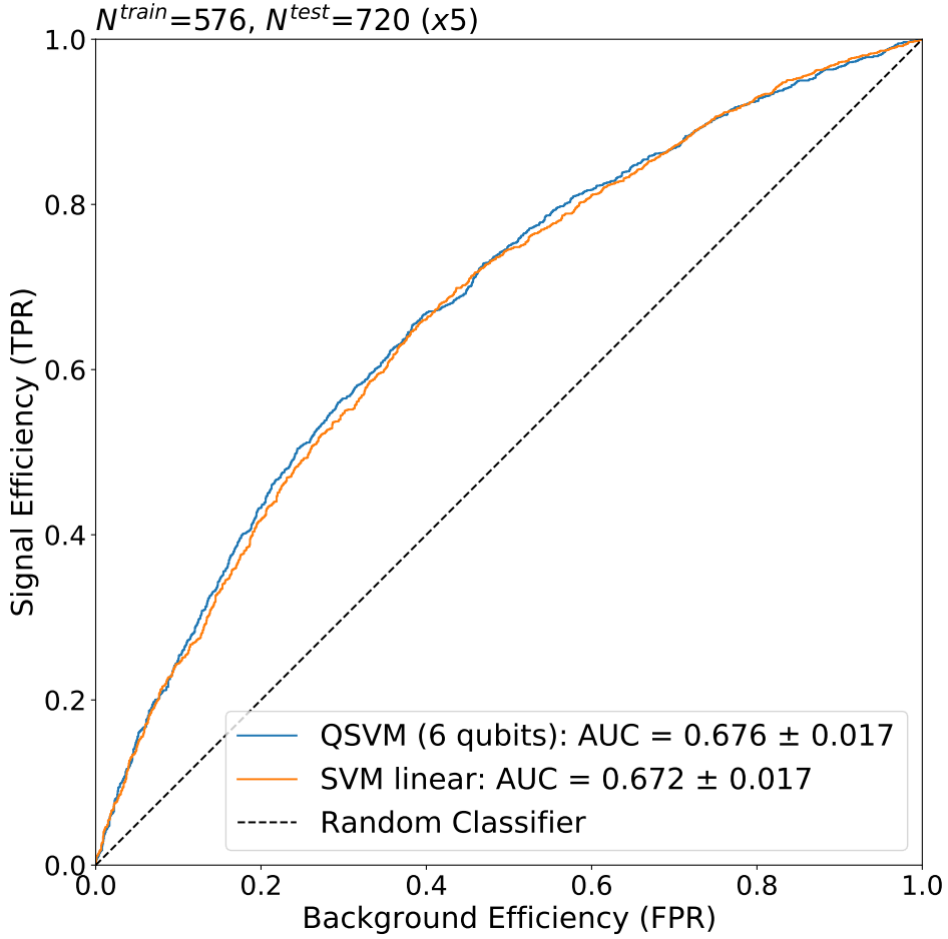
AE for feature reduction : 67 → 16 latent space features :

- Input physical obs. normalized to [0,1]
- Latent space dim. = 16 (Sigmoid activation in latent and output space)
- ELU activation functions



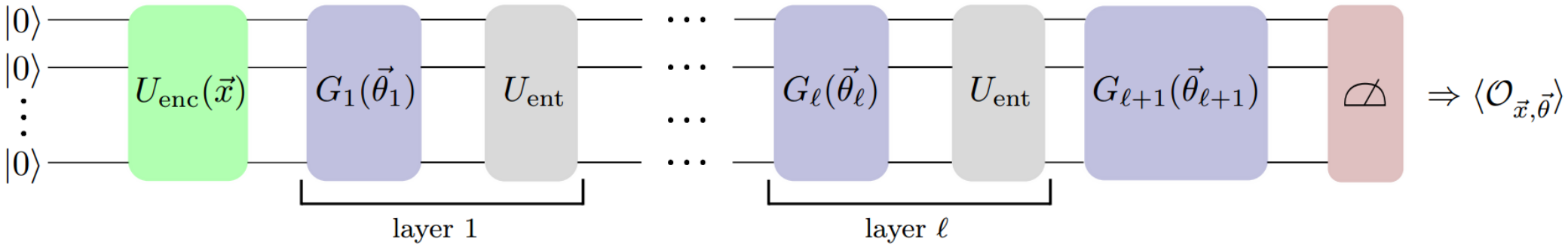


QSVM with more features



64 original features (AUC)

Variational quantum circuits



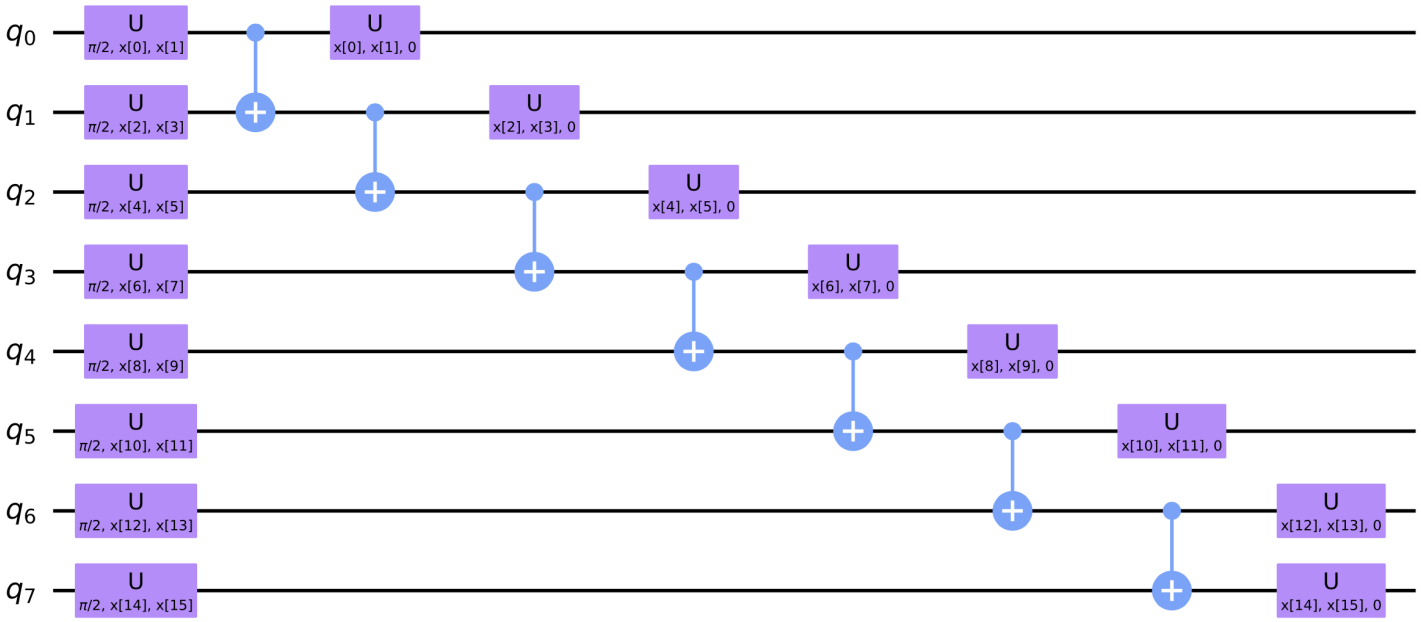
We measure the first qubit and minimise the binary cross entropy loss.

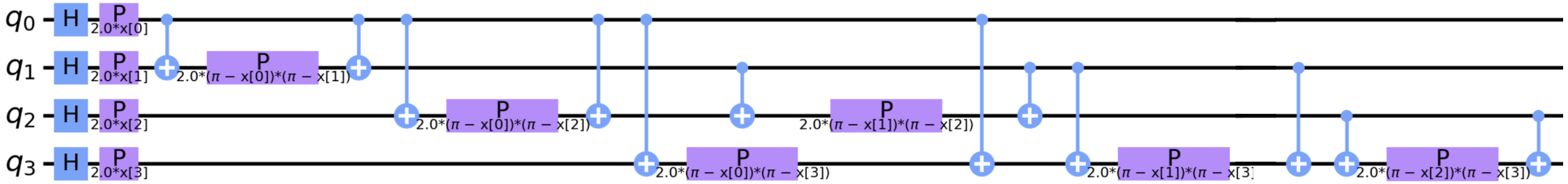
Data reuploading

Several repetitions of the VQC scheme before the measurement.

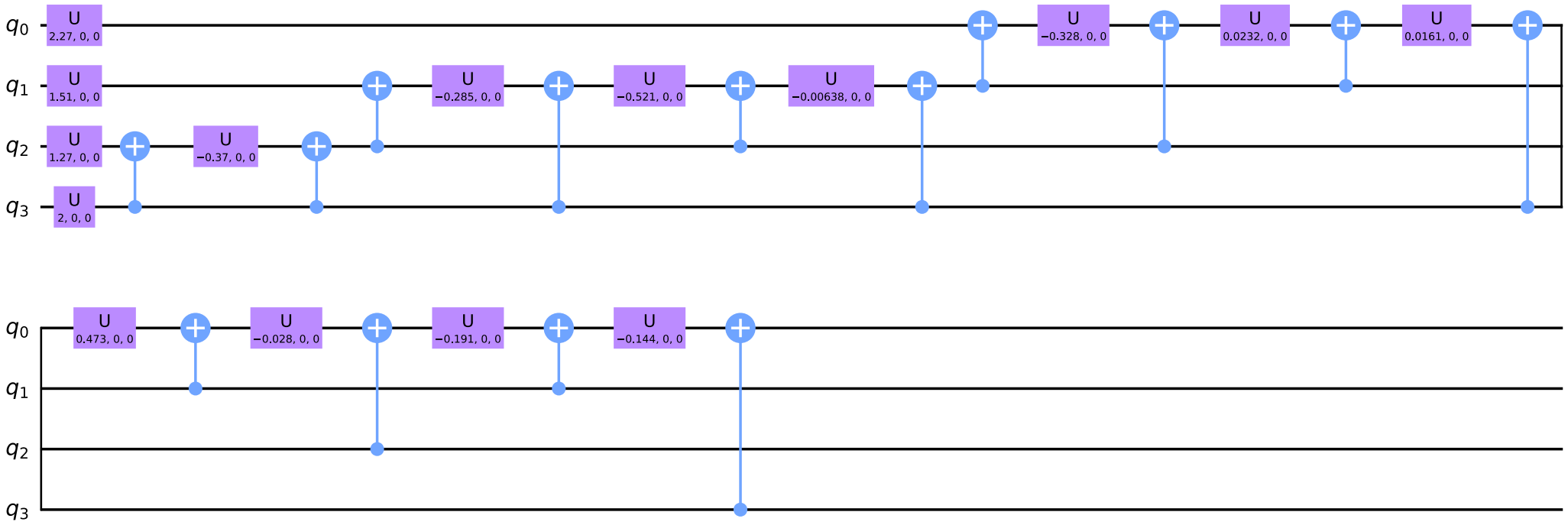
We can use more variables.

Feature maps circuits





Feature maps circuits (Cont.)



Amplitude encoding circuit