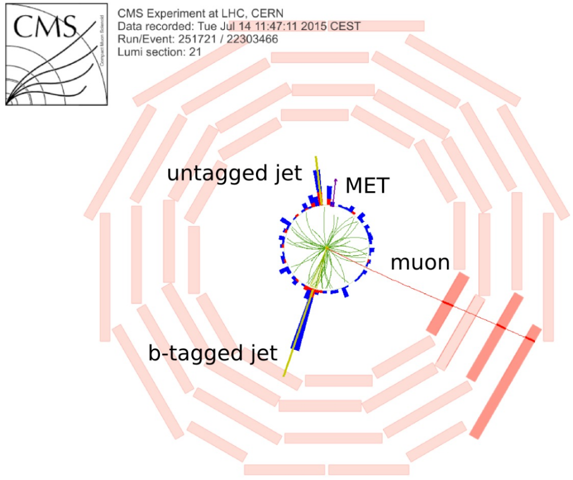
ETH zürich

Quantum machine learning for LHC data:
Classification & Anomaly Detection

Vasilis Belis

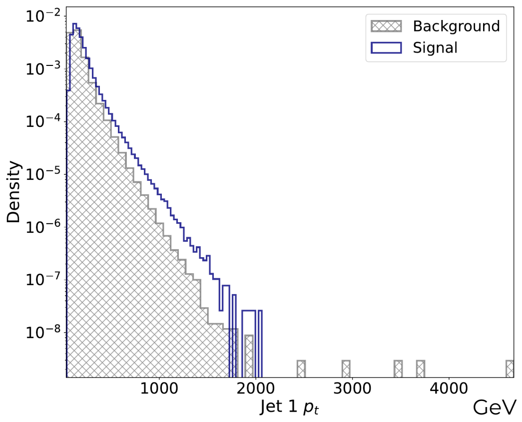
IPA-ML workshop
March 22nd 2023
ETH Zurich

Typical workflow: Model-dependent searches



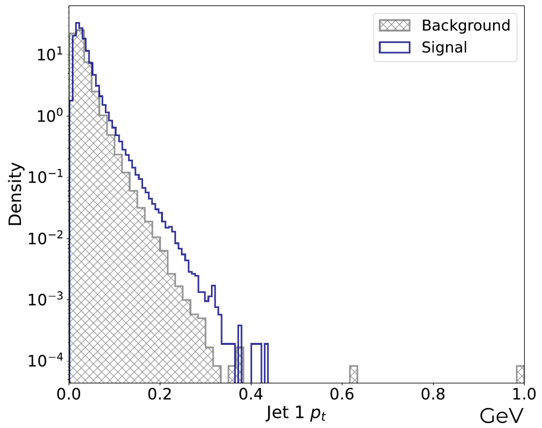
MC Simulation

compute prob. distributions →



• • •

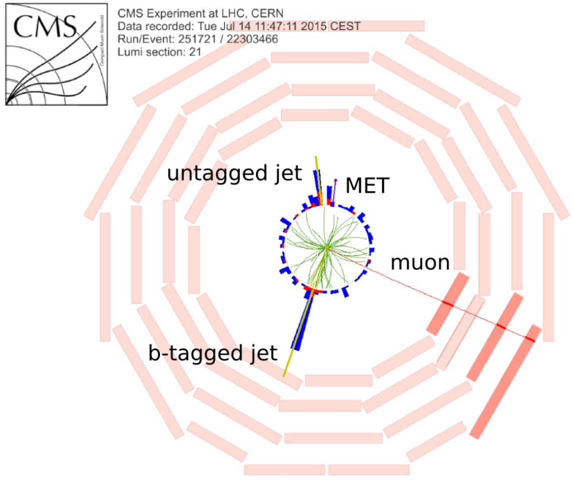
apply cuts and normalise →



• • •

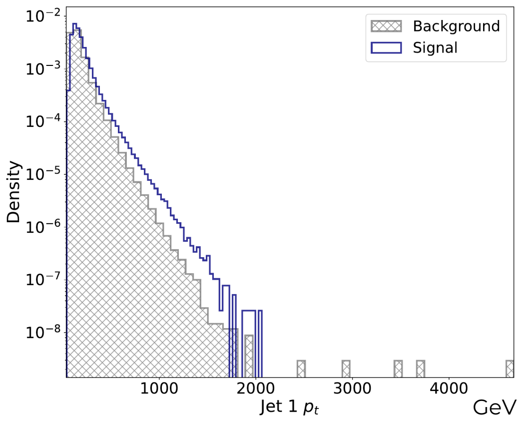
Input for classifiers models

Typical workflow: Model-dependent searches

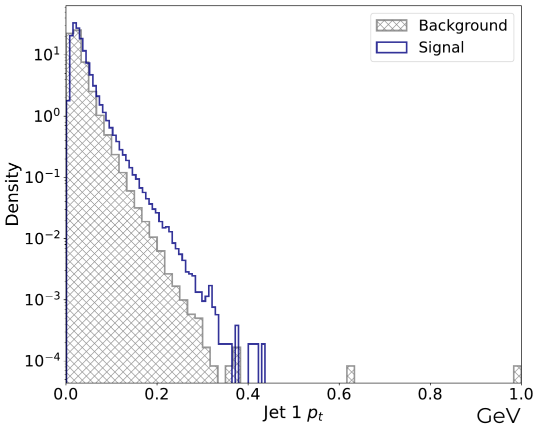


MC Simulation

compute prob. distributions



apply cuts and normalise

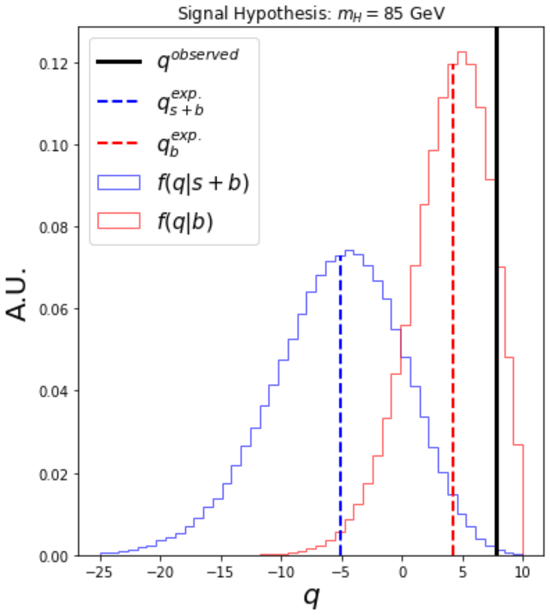


Input for classifiers models

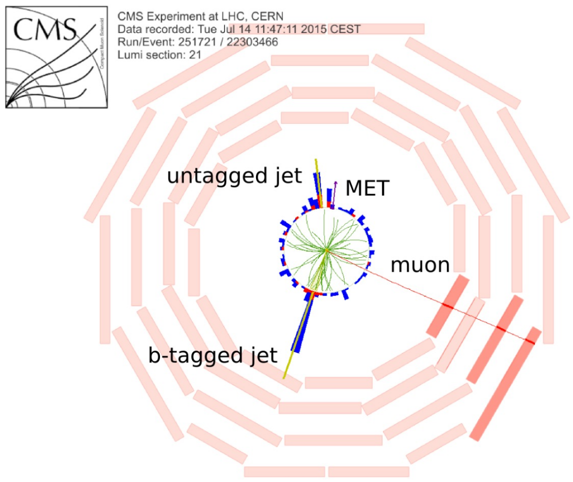
Classifier models:

cut-based or Machine Learning (BDT, NN, graph-net, etc.)

produce the *test statistic*

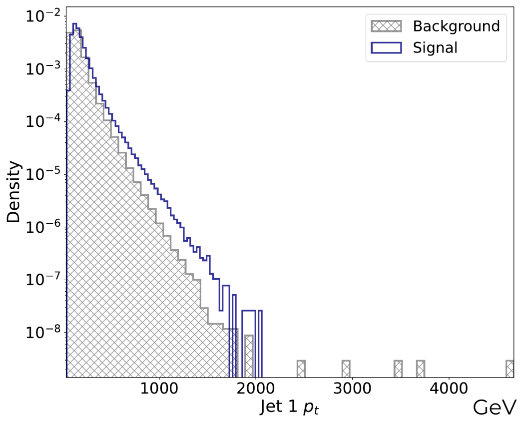


Typical workflow: Model-dependent searches

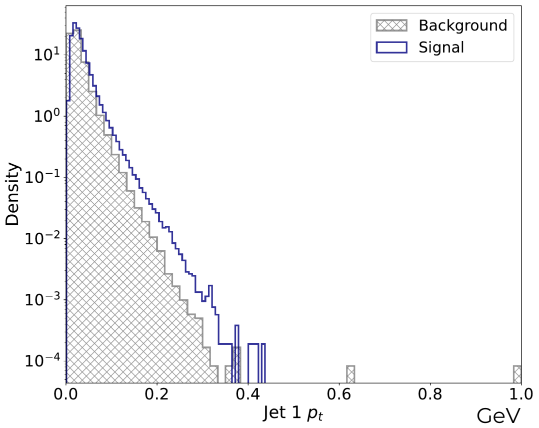


MC Simulation

compute prob. distributions



apply cuts and normalise

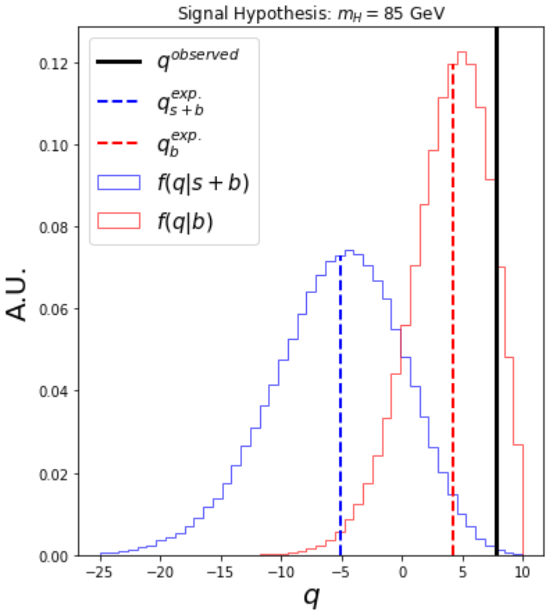


Input for classifiers models

Classifier models:

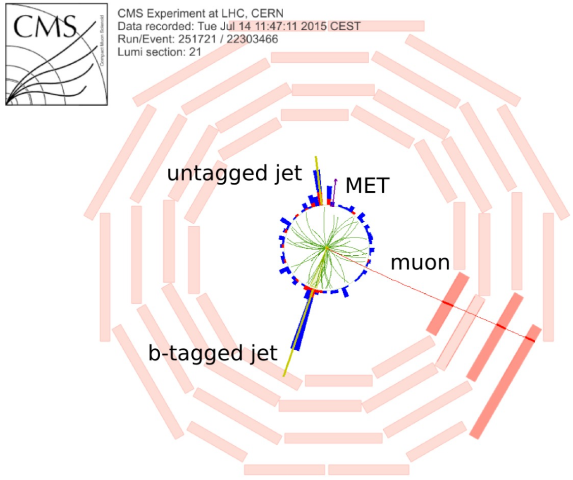
cut-based or Machine Learning (BDT, NN, graph-net, etc.)

produce the *test statistic*



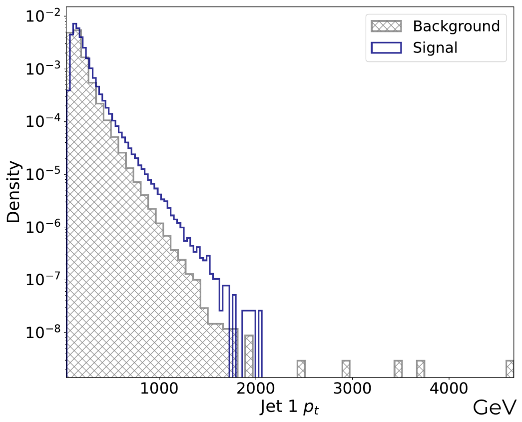
What if you do not know the signal?

Workflow: Model-dependent searches

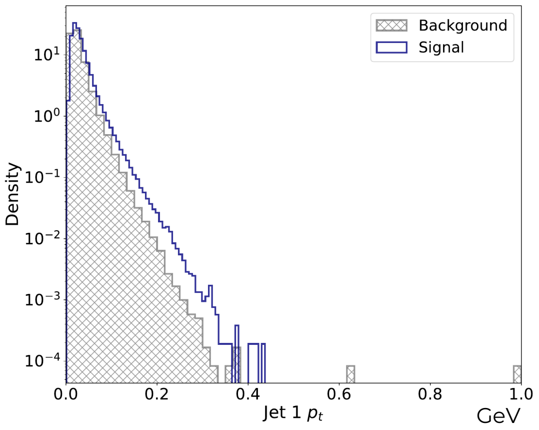


MC Simulation

compute prob. distributions



apply cuts and normalise

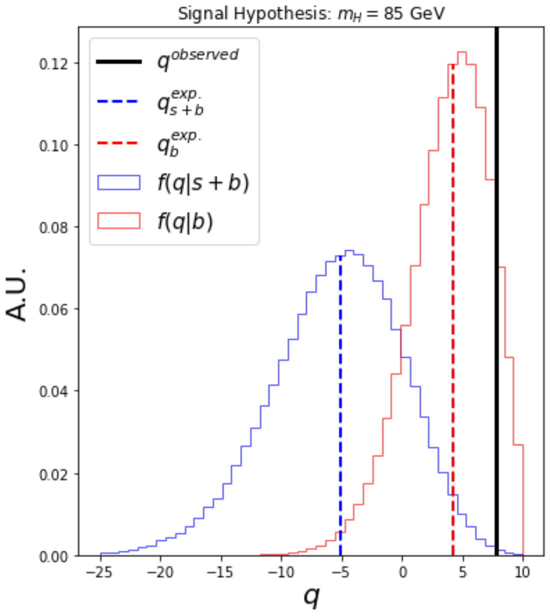


Input for classifiers models

Classifier models:

cut-based or Machine Learning (BDT, NN, graph-net, etc.)

produce the *test statistic*

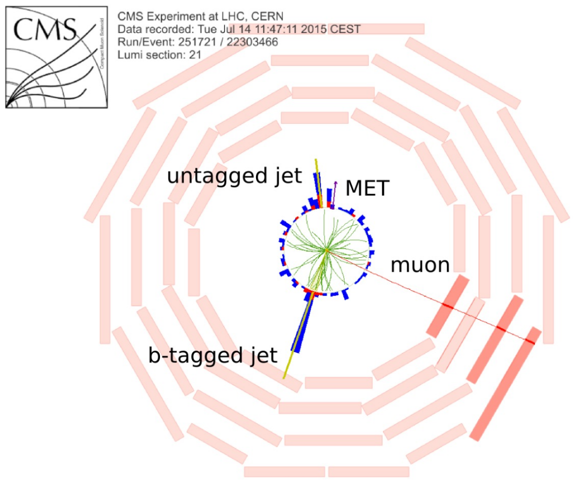


What if you do not know the signal?

Bias is **not necessarily bad**. It can be **great!**

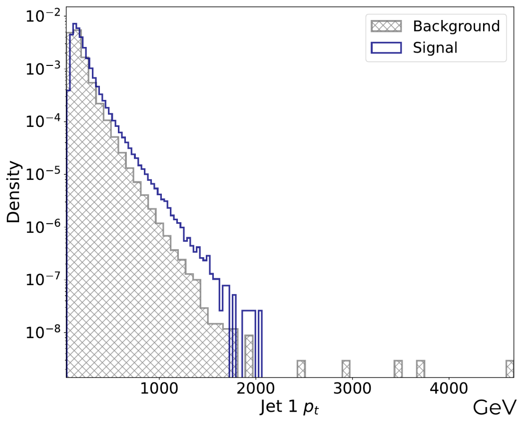
Higher bias ⟶ lower variance of the statistical model [Cramér–Rao bound]

Typical workflow: Model-dependent searches

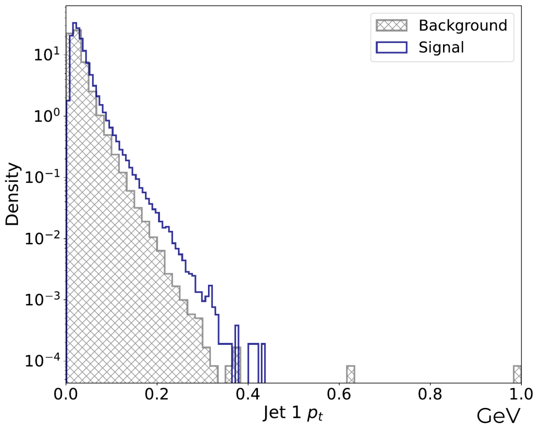


MC Simulation

compute prob. distributions



apply cuts and normalise

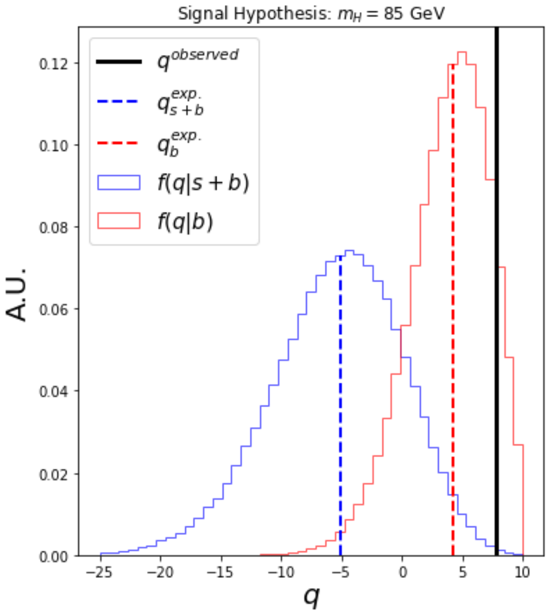


Input for classifiers models

Classifier models:

cut-based or Machine Learning (BDT, NN, graph-net, etc.)

produce the *test statistic*



What if you do not know the signal?

Bias is **not necessarily bad**. It can be **great!**

Higher bias → lower variance of the statistical model [Cramér–Rao bound]

Look at nature with minimal bias.

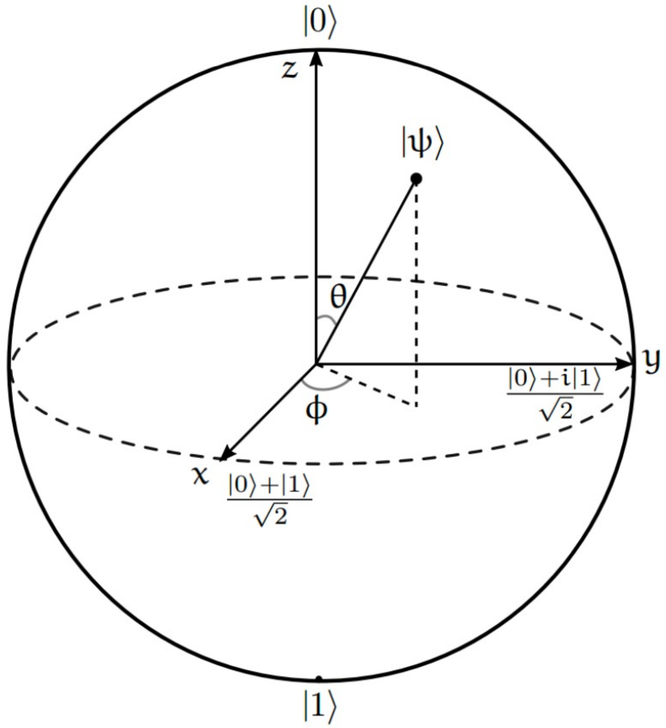
One possible solution:
Anomaly detection (ML/DL)

Quantum machine learning

Basics of quantum information processing

The qubit:

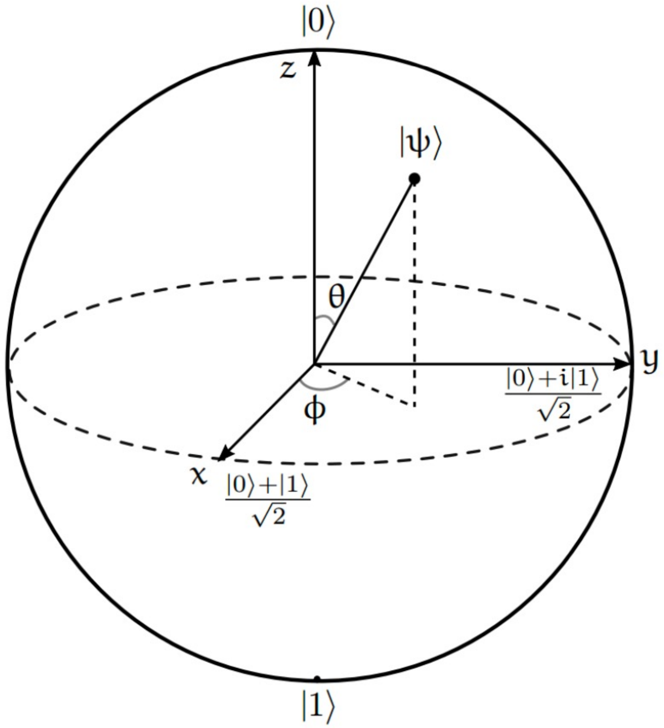
$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \;\equiv \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$



Basics of quantum information processing

The qubit:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \;\equiv \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$



Generic qubit operations (quantum gates)
$U = e^{-i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \in \mathrm{SU}(2)$:

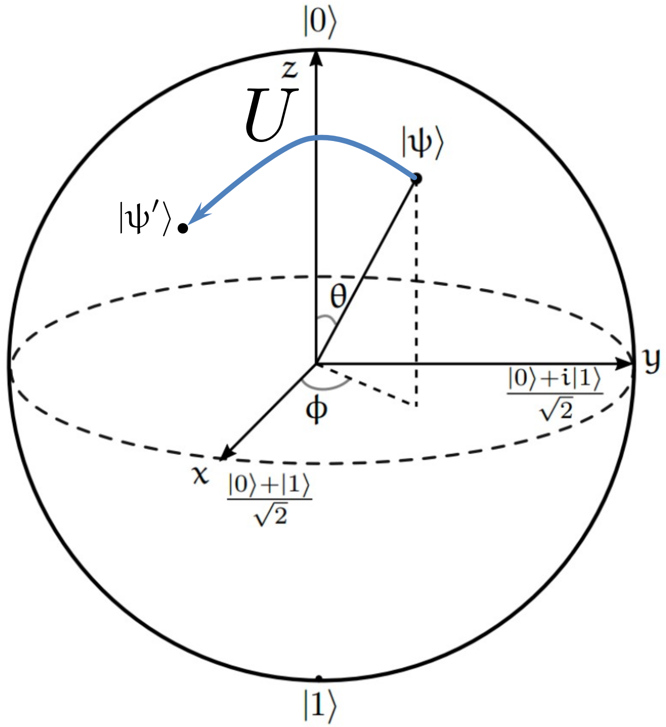
$$U(\theta,\phi,\lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda}\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}$$

Construct all possible gates from $U(\theta,\phi,\lambda)$

Basics of quantum information processing

The qubit:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \;\equiv \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$



Generic qubit operations (quantum gates) $U = e^{-i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \in \mathrm{SU}(2)$:

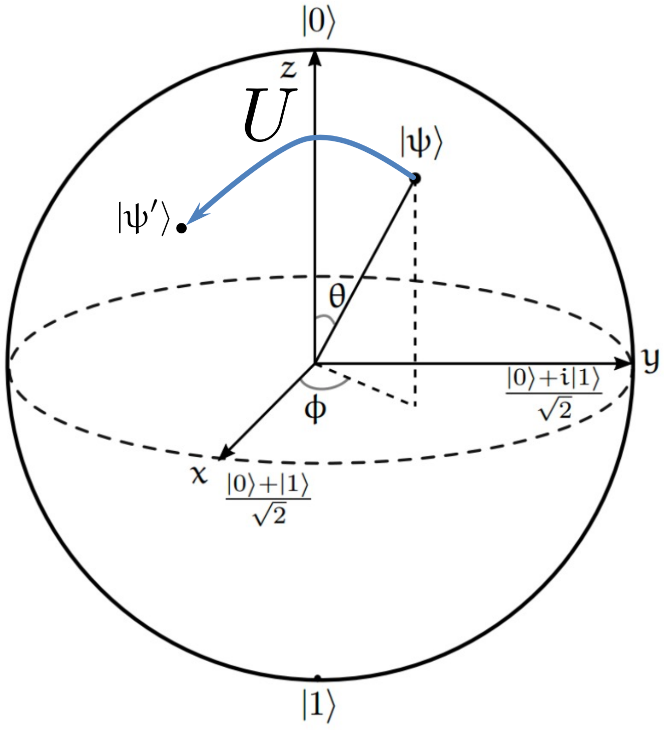
$$U(\theta,\phi,\lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda}\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}$$

Construct all possible gates from $U(\theta,\phi,\lambda)$

Basics of quantum information processing

The qubit:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \equiv \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$

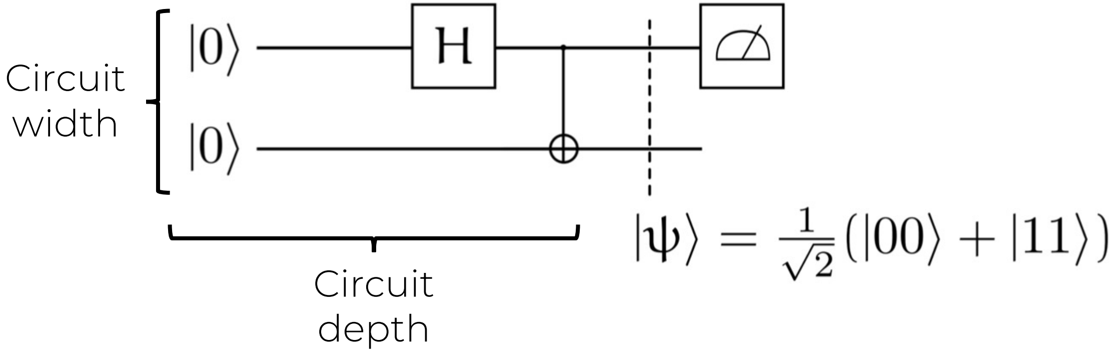


Generic qubit operations (quantum gates) $U = e^{-i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \in \mathrm{SU}(2)$:

$$U(\theta, \phi, \lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda}\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}$$

Construct all possible gates from $U(\theta, \phi, \lambda)$

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \equiv U\left(\frac{\pi}{2}, 0, \pi\right)$$



Quantum gates and universality

Single qubit gates

Decomposition of unitary up to global phase

$$U(\theta, \phi, \lambda) \sim R_z(\lambda) R_y(\theta) R_z(\phi)$$

True for any other non-parallel axes.

Quantum gates and universality

Single qubit gates

Decomposition of unitary up to global phase

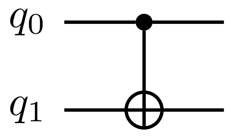
$$U(\theta, \phi, \lambda) \sim R_z(\lambda) R_y(\theta) R_z(\phi)$$

True for any other non-parallel axes.

Multi-qubit gates

Controlled-X gate (CNOT)

$$CX_{q_0,q_1} = |0\rangle\langle 0| \otimes \mathbb{I}_{2\times 2} + |1\rangle\langle 1| \otimes X$$
$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} =$$



Quantum gates and universality

Single qubit gates

Decomposition of unitary up to global phase

$$U(\theta, \phi, \lambda) \sim R_z(\lambda) R_y(\theta) R_z(\phi)$$

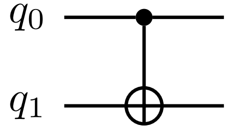
True for any other non-parallel axes.

Multi-qubit gates

Controlled-X gate (CNOT)

$$CX_{q_0,q_1} = |0\rangle\langle 0| \otimes \mathbb{I}_{2\times 2} + |1\rangle\langle 1| \otimes X$$

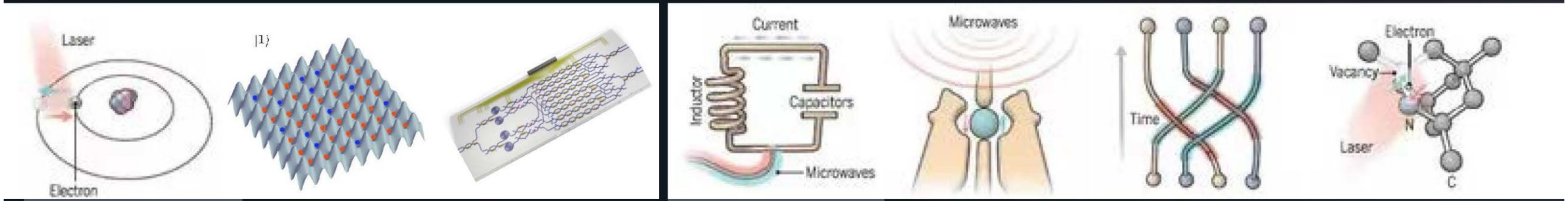
$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} =$$



Hardware implementation: decompose quantum circuit to the physical (implemented) set of gates.

The above “building blocks” can construct any n-qubit quantum circuit by operating on **at most** two qubits at a time [DiV95].

Quantum computer technologies



	Natural Qubits			Synthetic Qubits			
	Trapped Ions	**Neutral Atoms**	**Photonics**	**Superconducting Loops**	**Silicon Quantum Dots**	**Topological Qubits**	**Diamond Vacancies**
	Electrically charged atoms, or ions, are held in place with electric fields. Qubits are stored in electronic states. Ions are pushed with laser beams to allow the qubits to interact.	Neutral atoms, like ions, store qubits within electronic states. Laser activates the electrons to create interaction between qubits.	Photonic qubits (light particles) are sent through a maze of optical channels on a chip to interact. At the end of the maze, the distribution of photons is measured as an output.	A resistance-free current oscillates back and forth around a circuit loop. An injected microwave signal excites the current into super-position states.	These "artificial atoms" are made by adding an electron to a small piece of pure silicon. Microwaves control the electron's quantum state.	Quasiparticles can be seen in the behavior of electrons channeled through semi-conductor structures. Their braided paths can encode quantum information.	A nitrogen atom and a vacancy add an electron to a diamond lattice. Its quantum spin state, along with those of nearby carbon nuclei, can be controlled with light.
Qubit Coherence Time (sec)	>1000	1	--	0.00005	0.03	N/A	10
Fidelity	99.9%	97%	--	99.4%	~99%	N/A	99.2%
Qubits Connected	High	Very high; low individual control	--	High	Very Low	N/A	Low
Company Support	IONQ, AQT, Honeywell, Oxford Ionics	Atom Computing, ColdQuanta, QuEra	Psiquantum, Xanadu	Google, IBM, QCI, Rigetti	HRL, Intel, SQC	Microsoft	Quantum Diamond Technologies
Pros	Very stable. Highest achieved gate fidelities.	Many qubits, 2D and maybe 3D.	Linear optical gates, integrated on-chip.	Can lay out physical circuits on chip.	Borrows from existing semiconductor industry.	Greatly reduce errors.	Can operate at room temperature.
Cons	Slow operation. Many lasers are needed.	Hard to program and control individual qubits; prone to noise.	Each program requires its own chip with unique optical channels. No memory.	Must be cooled to near absolute zero. High variability in fabrication. Lots of noise.	Only a few connected. Must be cooled to near absolute zero. High variability in fabrication.	Existence not yet confirmed.	Difficult to create high numbers of qubits, limiting compute capacity.

Source: Science, Dec. 2016

Motivation

Why quantum machine learning? Why for HEP?

Practical and exploratory answer

Investigate a new set of ML techniques to assess for advantages. Why not?

Motivation

Why quantum machine learning? Why for HEP?

Practical and exploratory answer

Investigate a new set of ML techniques to assess for advantages. Why not?

Fundamental motivation

Exploitation of information and correlations (quantum remnants) inherent in HEP data?
Some natural applications:
Quantum simulation of parton shower [Phys. Rev. D 106, 056002], Simulating Lattice Field Theory [arXiv:2302.00467]

Motivation

Why quantum machine learning? Why for HEP?

Practical and exploratory answer

Investigate a new set of ML techniques to assess for advantages. Why not?

Fundamental motivation

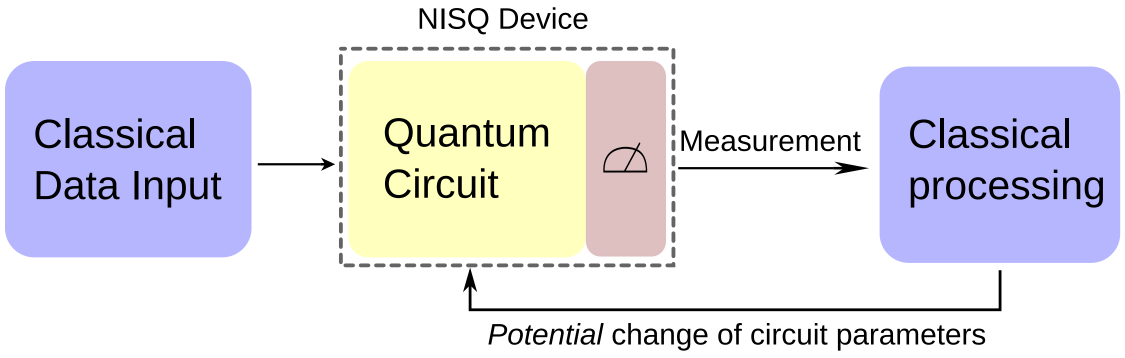
Exploitation of information and correlations (quantum remnants) inherent in HEP data?
Some natural applications:
Quantum simulation of parton shower [Phys. Rev. D 106, 056002], Simulating Lattice Field Theory [arXiv:2302.00467]

Theoretical results

Generalisation with few data, computational speed-ups, uncover correlations unrecognisable to classical methods, etc.

[M. Caro et al., Nature Communications 13, 4919 (2022)]
[A. Abbas et al., Nature Computational Science 1, 403 (2021)]
[Y. Liu et al., Nature Physics 17, 1013 (2021)]
[H. Huang et al., Nature Communications 12, 2631 (2021)]
[H . Huang et al.,, Science 376, 1182 (2022)]
[N. Pirnay et al., arXiv: 2212.08678 (2022)]
Among others...

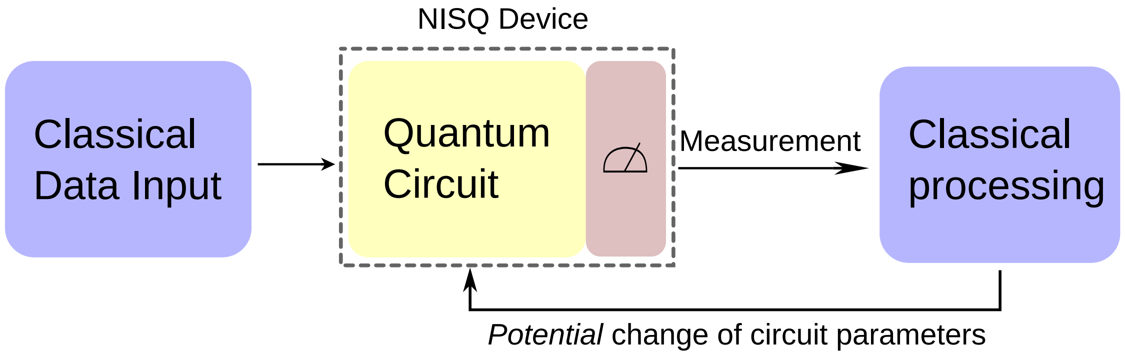
Quantum machine learning algorithms



Noisy intermediate scale quantum devices

- Circuit width: limited number of qubits.
- Circuit depth: limited number of operations per qubit (small decoherence times).
- Hardware noise.

Quantum machine learning algorithms

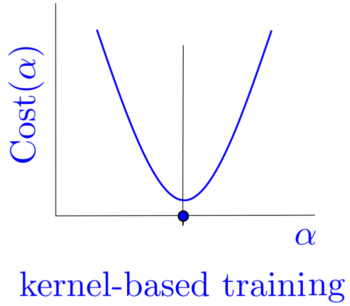


Noisy intermediate scale quantum devices

- Circuit width: limited number of qubits.
- Circuit depth: limited number of operations per qubit (small decoherence times).
- Hardware noise.

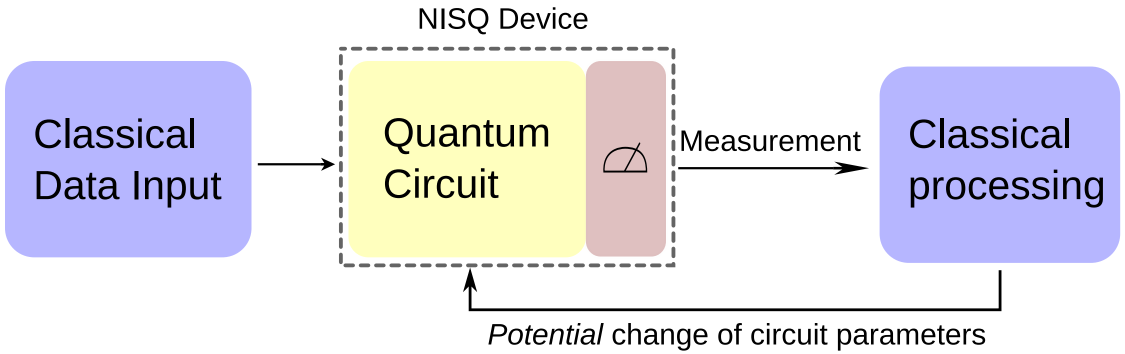
Kernel methods

Quantum Support Vector Machines



- Heavily dependent on the choice of the kernel.
- Theoretically provable quantum advantages.
- $\mathcal{O}(N_{\text{train}}^2)$

Quantum machine learning algorithms

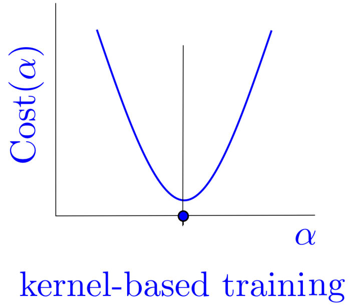


Noisy intermediate scale quantum devices

- Circuit width: limited number of qubits.
- Circuit depth: limited number of operations per qubit (small decoherence times).
- Hardware noise.

Kernel methods

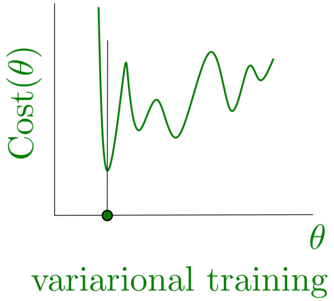
Quantum Support Vector Machines



- Heavily dependent on the choice of the kernel.
- Theoretically provable quantum advantages.
- $\mathcal{O}(N_{\text{train}}^2)$

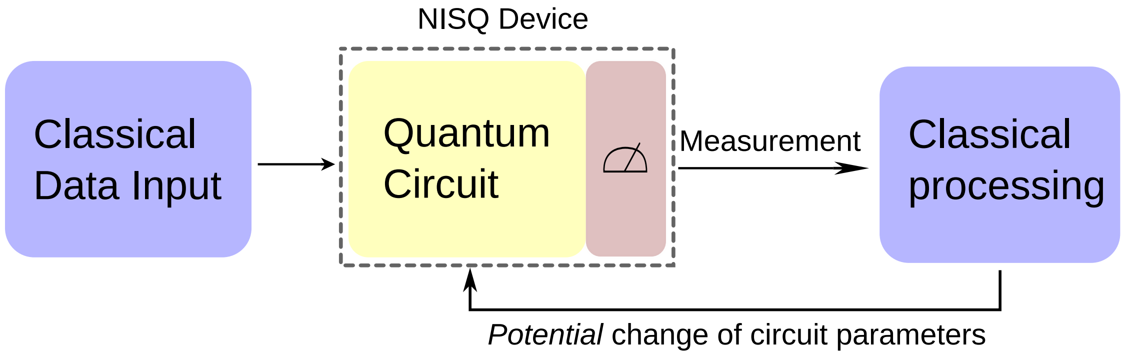
Variational/Parametrised Quantum Circuits

Quantum Neural Networks



- Model flexibility.
- Trainability issues (vanishing gradients)
- $\mathcal{O}(N_{\text{train}})$

Quantum machine learning algorithms

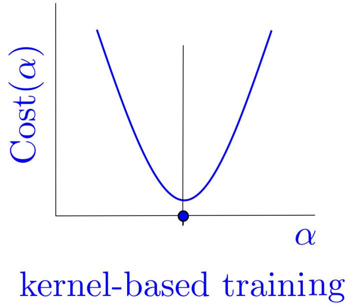


Noisy intermediate scale quantum devices

- Circuit width: limited number of qubits.
- Circuit depth: limited number of operations per qubit (small decoherence times).
- Hardware noise.

Kernel methods

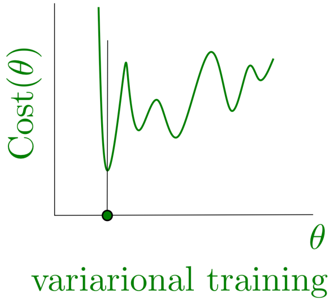
Quantum Support Vector Machines



- Heavily dependent on the choice of the kernel.
- Theoretically provable quantum advantages.
- $\mathcal{O}(N_{\text{train}}^2)$

Variational/Parametrised Quantum Circuits

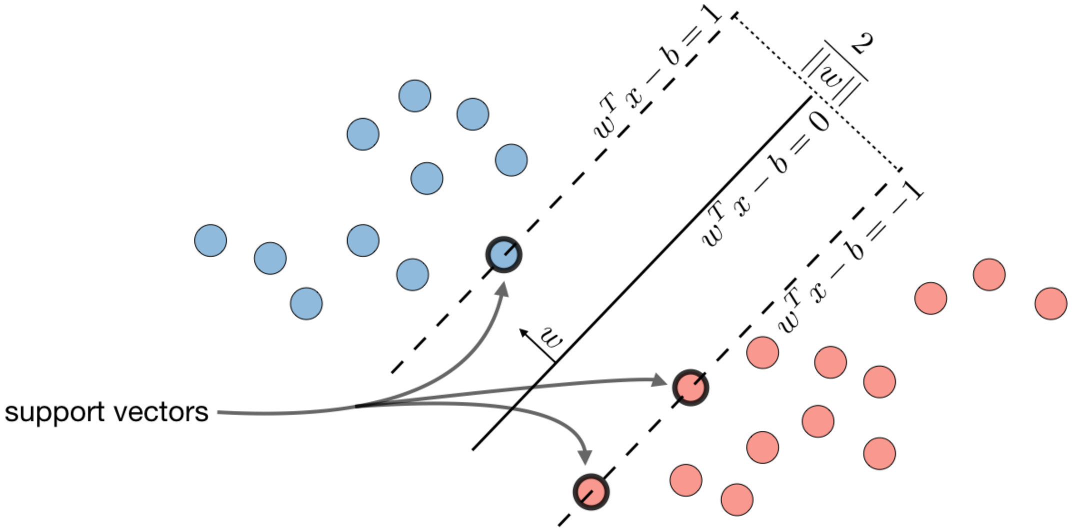
Quantum Neural Networks



- Model flexibility.
- Trainability issues (vanishing gradients)
- $\mathcal{O}(N_{\text{train}})$

Current hardware limitations: feature reduction presently needed for realistic datasets.

Quantum Support Vector Machines



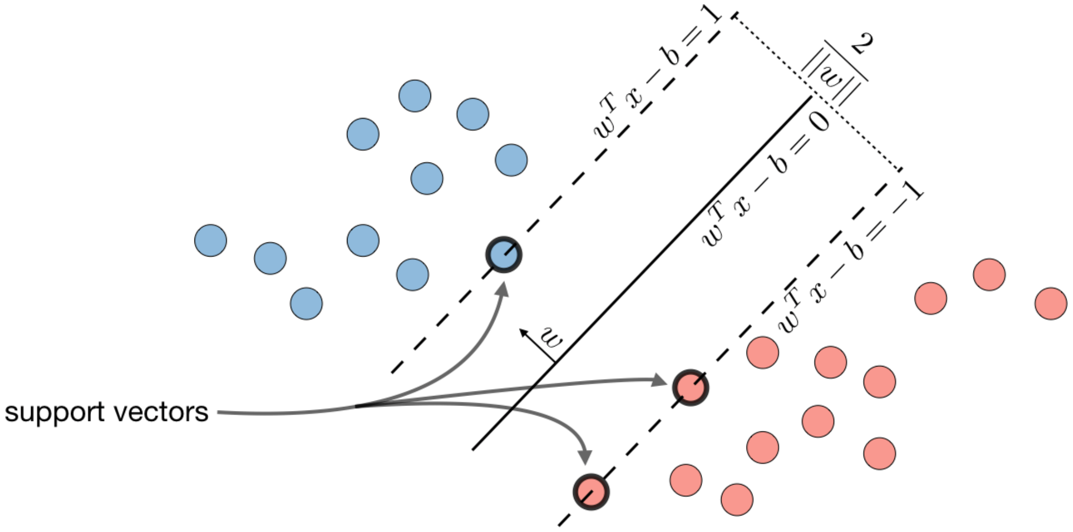
SVM objective function is equivalent to (dual Lagrangian):

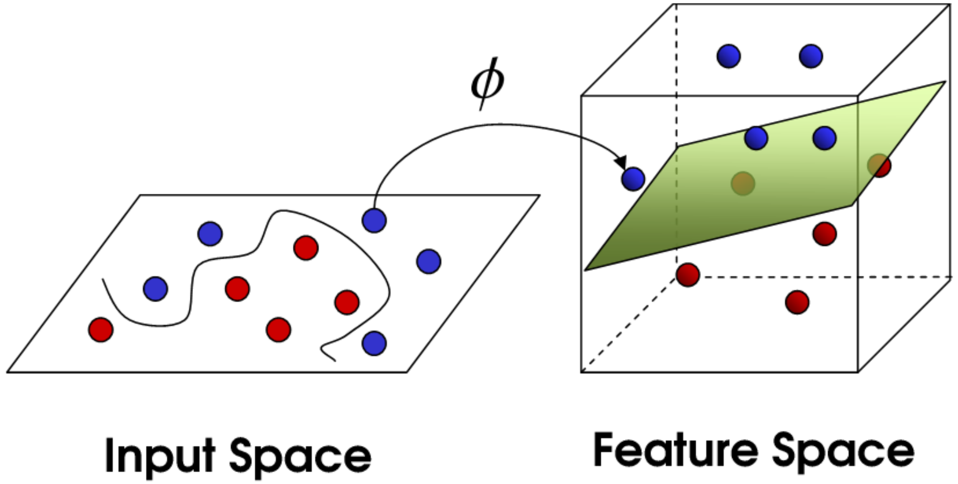
maximize $$L(c_1, \ldots, c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\vec{x}_i \cdot \vec{x}_j) y_j c_j$$

subject to $$\sum_{i=1}^{n} c_i y_i = 0, \text{ and } 0 \leq c_i \leq C, \ \forall i$$

Kernel trick: $$(\vec{x}_i \cdot \vec{x}_j) \mapsto k(\vec{x}_i \cdot \vec{x}_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j)$$

Quantum Support Vector Machines





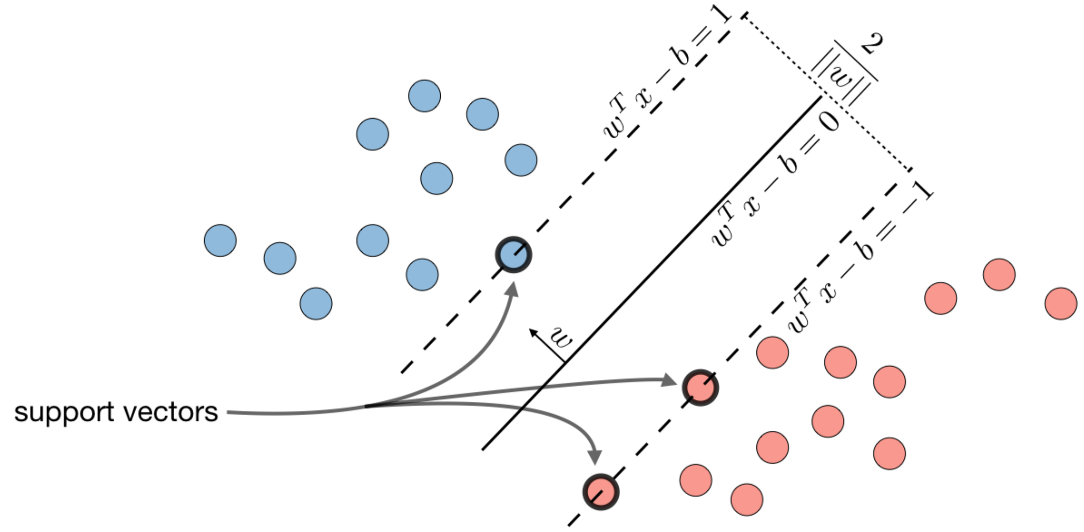
SVM objective function is equivalent to (dual Lagrangian):

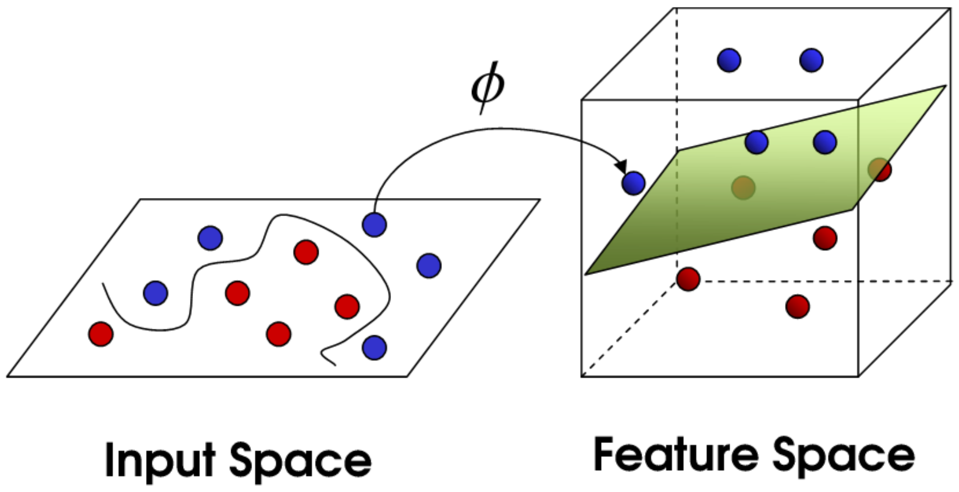
maximize $L(c_1, \ldots, c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\vec{x}_i \cdot \vec{x}_j) y_j c_j$

subject to $\sum_{i=1}^{n} c_i y_i = 0, \text{ and } 0 \leq c_i \leq C, \ \forall i$

Kernel trick: $(\vec{x}_i \cdot \vec{x}_j) \mapsto k(\vec{x}_i \cdot \vec{x}_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j)$

Quantum Support Vector Machines





SVM objective function is equivalent to (dual Lagrangian):

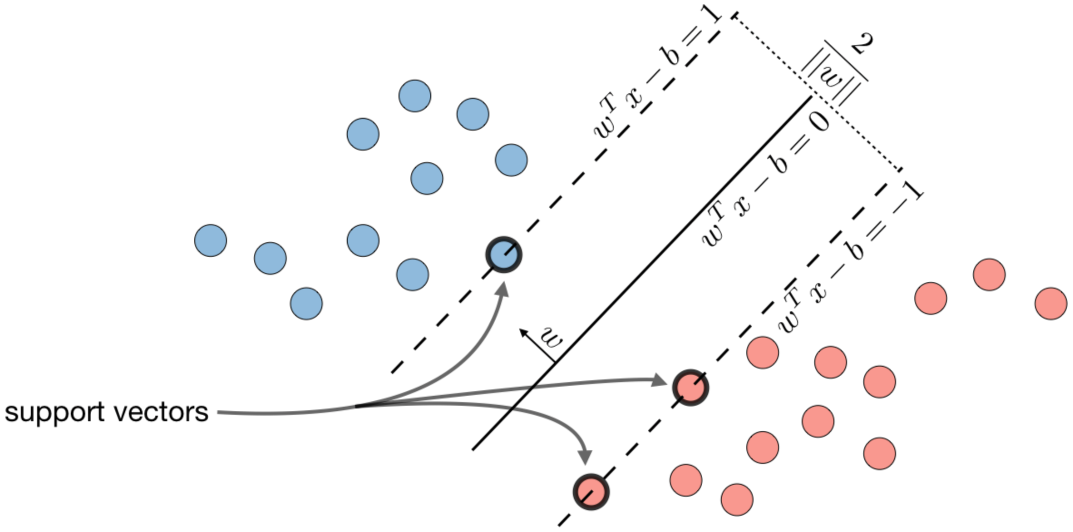
maximize $$L(c_1, \ldots, c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\vec{x}_i \cdot \vec{x}_j) y_j c_j$$

subject to $$\sum_{i=1}^{n} c_i y_i = 0, \text{ and } 0 \leq c_i \leq C, \ \forall i$$

Kernel trick: $$(\vec{x}_i \cdot \vec{x}_j) \mapsto k(\vec{x}_i \cdot \vec{x}_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j)$$

Make the kernel **quantum**

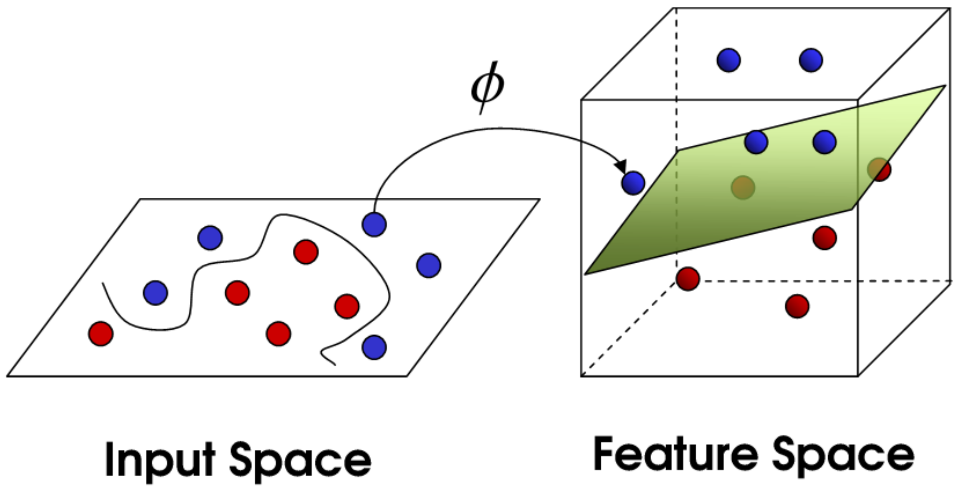
Quantum Support Vector Machines



SVM objective function is equivalent to (dual Lagrangian):

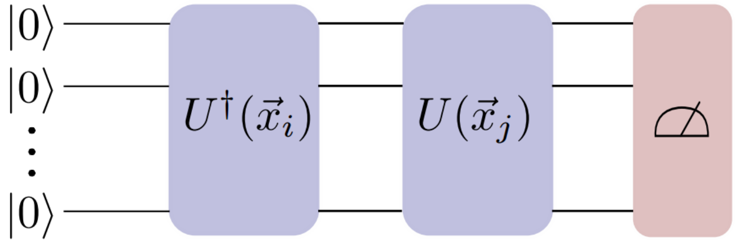
maximize $$L(c_1, \dots, c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i (\vec{x}_i \cdot \vec{x}_j) y_j c_j$$

subject to $$\sum_{i=1}^{n} c_i y_i = 0, \text{ and } 0 \leq c_i \leq C, \ \forall i$$



Kernel trick: $$(\vec{x}_i \cdot \vec{x}_j) \mapsto k(\vec{x}_i \cdot \vec{x}_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j)$$

Make the kernel **quantum**

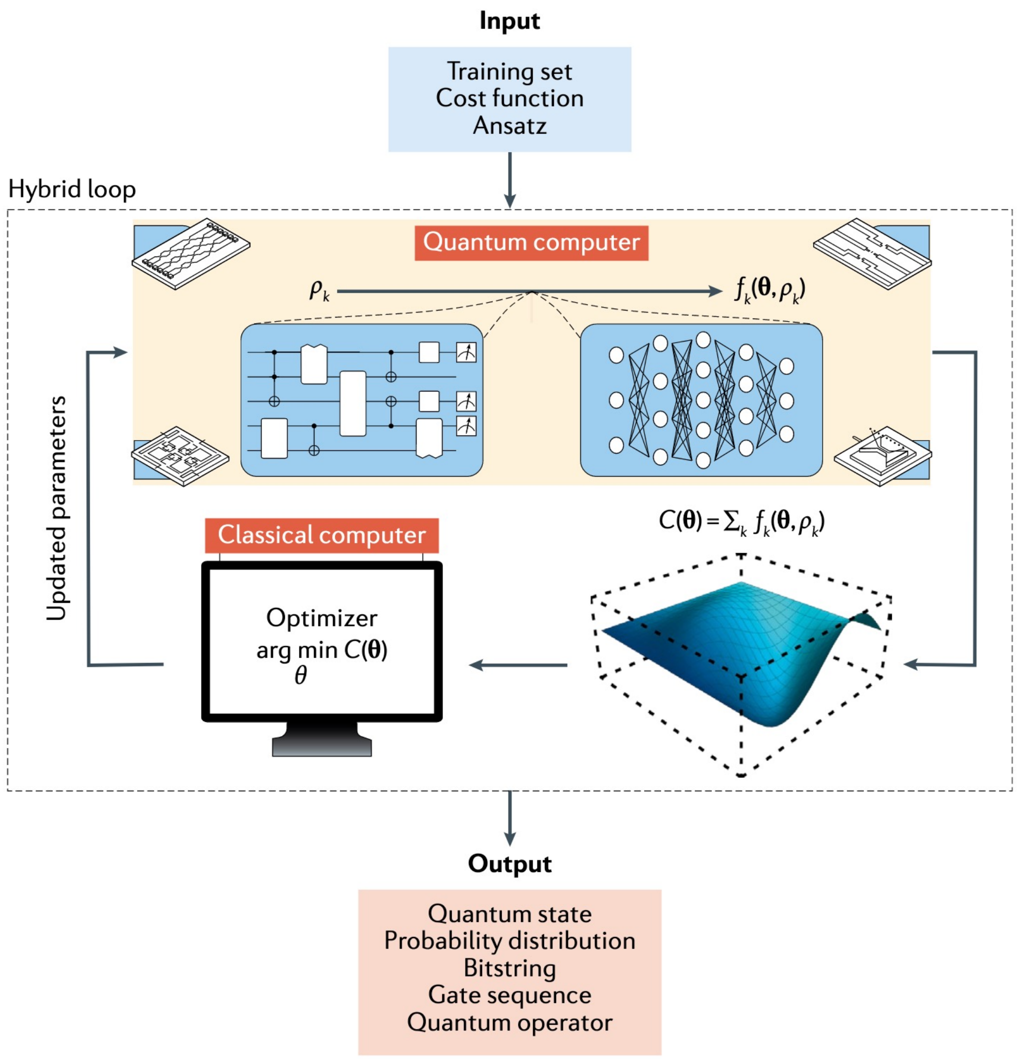


$$\Rightarrow K_{ij} = |\langle 0 | U^\dagger(\vec{x}_i) U(\vec{x}_j) | 0 \rangle|^2$$

***Can be generalised to unsupervised learning**

Quantum Neural Networks

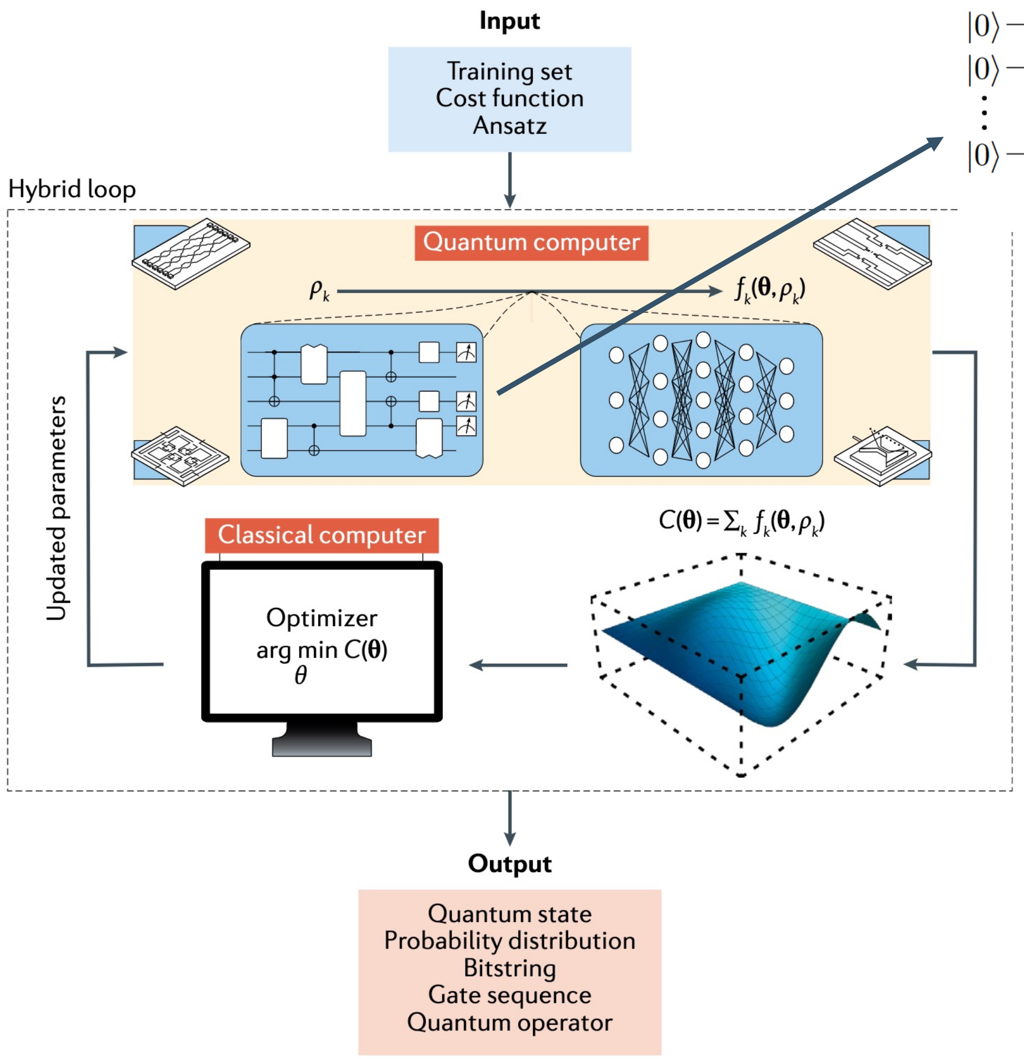
Variational quantum algorithm workflow



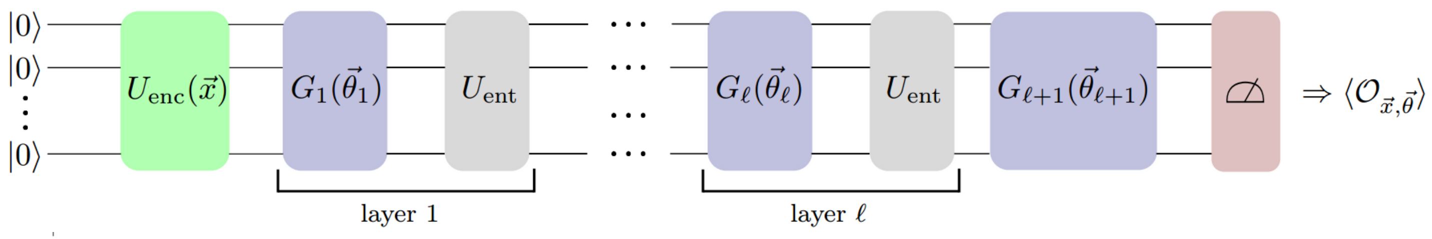
[M. Cerezo, et al. Nat. Rev. Phys. 3, 625–644 (2021)]

Quantum Neural Networks

Variational quantum algorithm workflow



[M. Cerezo, et al. Nat. Rev. Phys. 3, 625–644 (2021)]



1. Choose loss function.
   Task dependent: classification, reconstruction, generative modeling, etc.

2. Embed classical data to quantum states: $U_{\text{enc}}(x)$

3. Process quantum state with parametrized quantum gates.

4. Update trainable parameters

$$\Theta_{t+1} \leftarrow \Theta_t - \eta \nabla_\Theta \mathcal{L}[\langle \mathcal{O}(x;\Theta) \rangle]$$

Results

Finding new-physics in dijet events with QML

Identifying new-physics with quantum models

Anomaly detection with quantum machine learning

Background: QCD dijet events. $n^{\text{features}} = 300$ per jet → Too many for current hardware.

BSM anomalies: Graviton **&** New Scalar Boson → Multi-jet final state

$G \to W^- W^+$ $A \to HZ \to ZZZ$

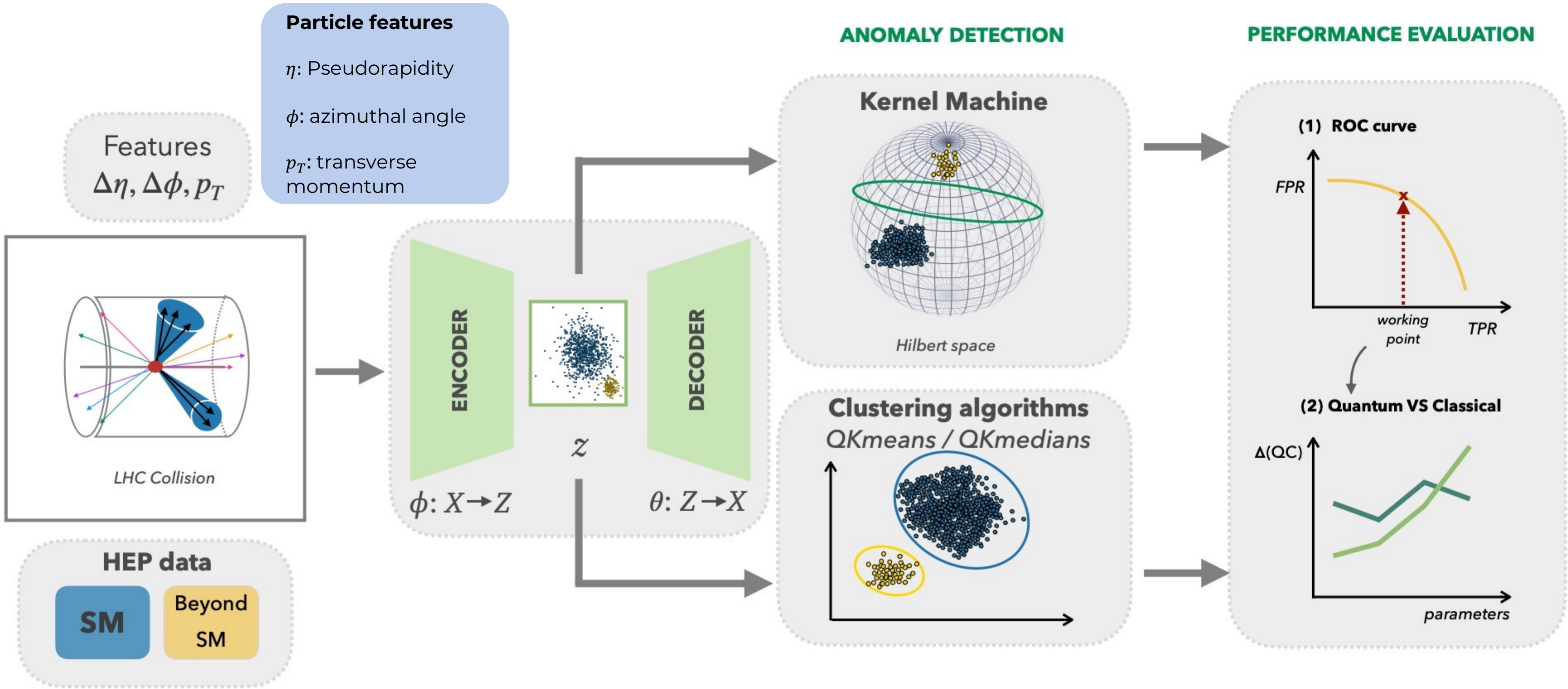
Identifying new-physics with quantum models

Anomaly detection with quantum machine learning

Background: QCD dijet events. $n^{\text{features}} = 300$ per jet ⟶ Too many for current hardware.

BSM anomalies: Graviton & New Scalar Boson ⟶ Multi-jet final state

$G \to W^- W^+$ $\quad A \to HZ \to ZZZ$



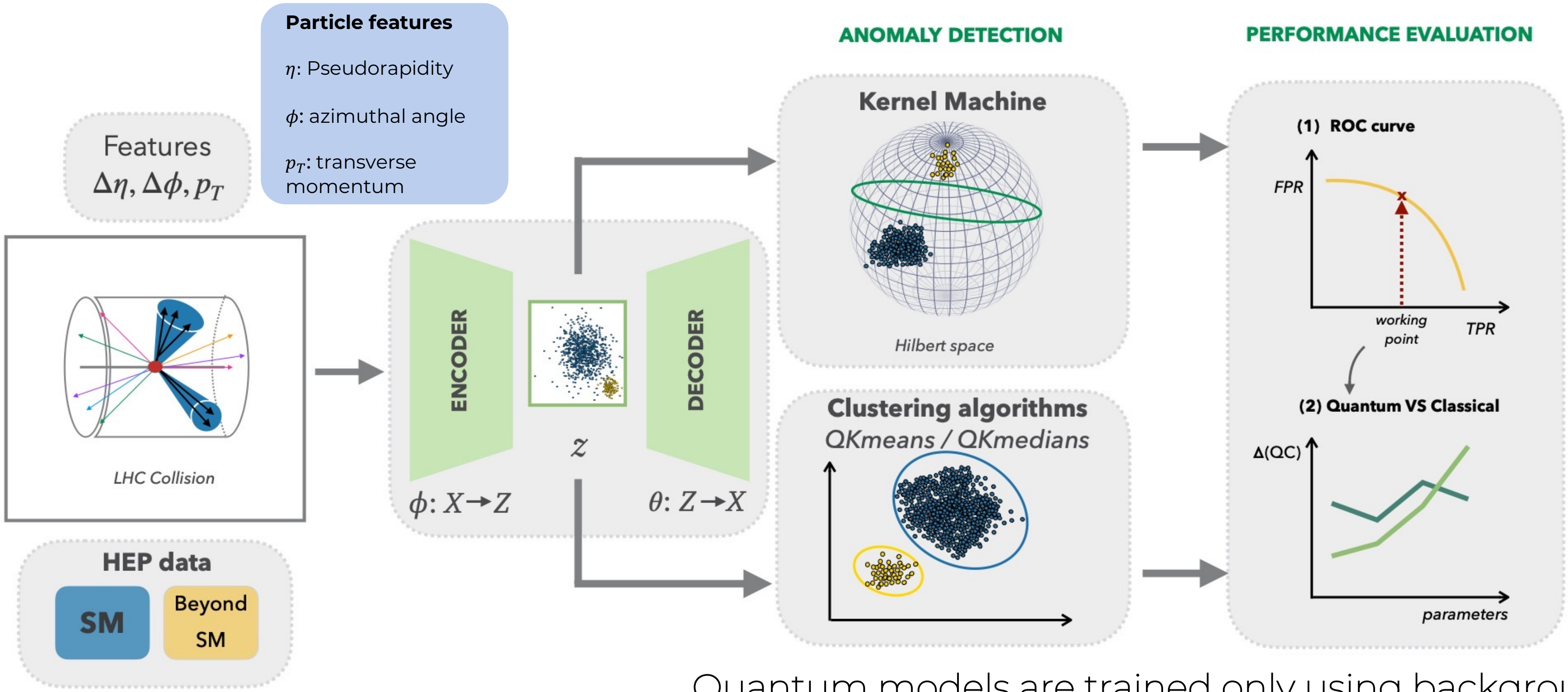
Identifying new-physics with quantum models

Anomaly detection with quantum machine learning

Background: QCD dijet events. $n^{\text{features}} = 300$ per jet ⟶ Too many for current hardware.

BSM anomalies: Graviton & New Scalar Boson ⟶ Multi-jet final state

$G \to W^- W^+$ $\quad$ $A \to HZ \to ZZZ$



Quantum models are trained only using background data and learn to extract a metric of "typicality".

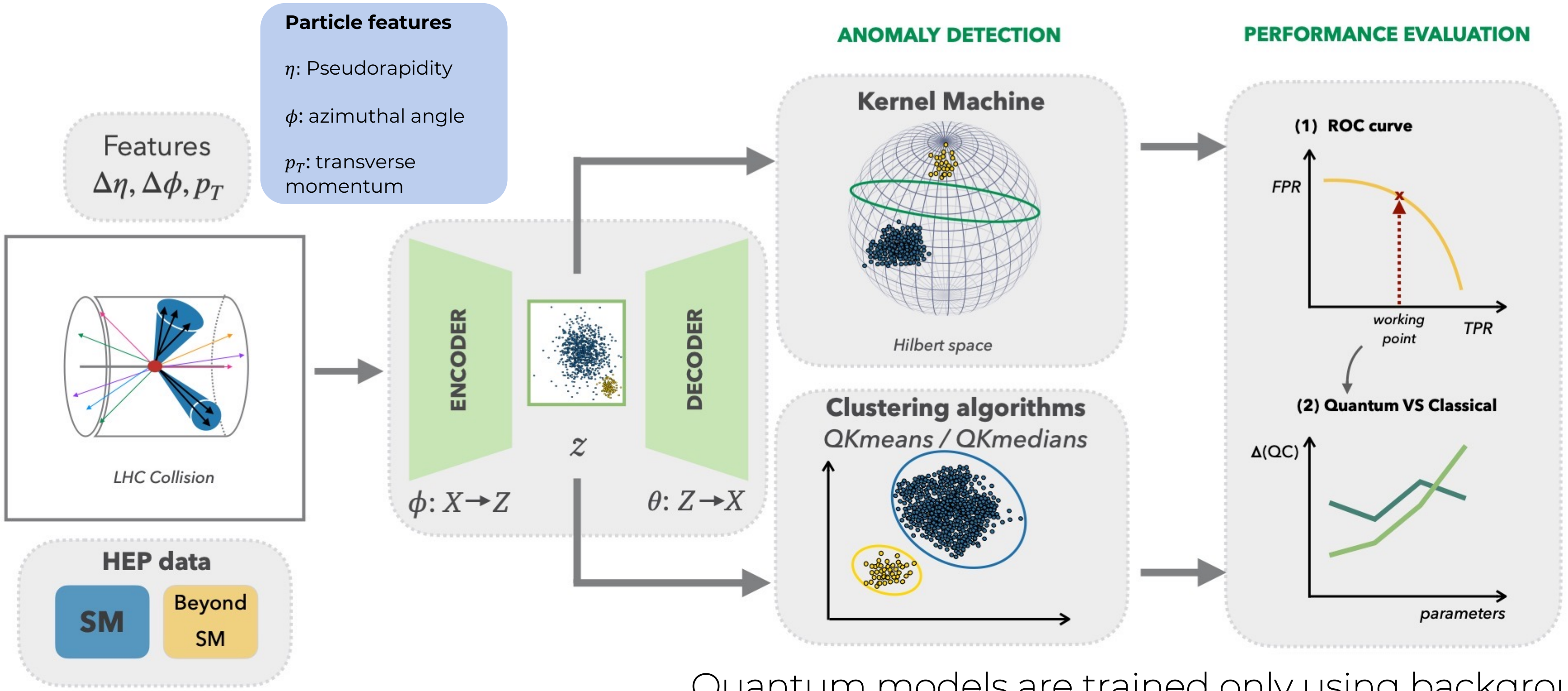
Identifying new-physics with quantum models

Anomaly detection with quantum machine learning

Background: QCD dijet events. $n^{\text{features}} = 300$ per jet ⟶ Too many for current hardware.

BSM anomalies: Graviton & New Scalar Boson ⟶ Multi-jet final state

$G \to W^- W^+$ $\quad A \to HZ \to ZZZ$



Suitable metric for anomaly detection

Background rejection @ working point

$$\varepsilon_{\mathrm{b}}^{-1}(\varepsilon_s; \mathcal{M})$$

Compare models

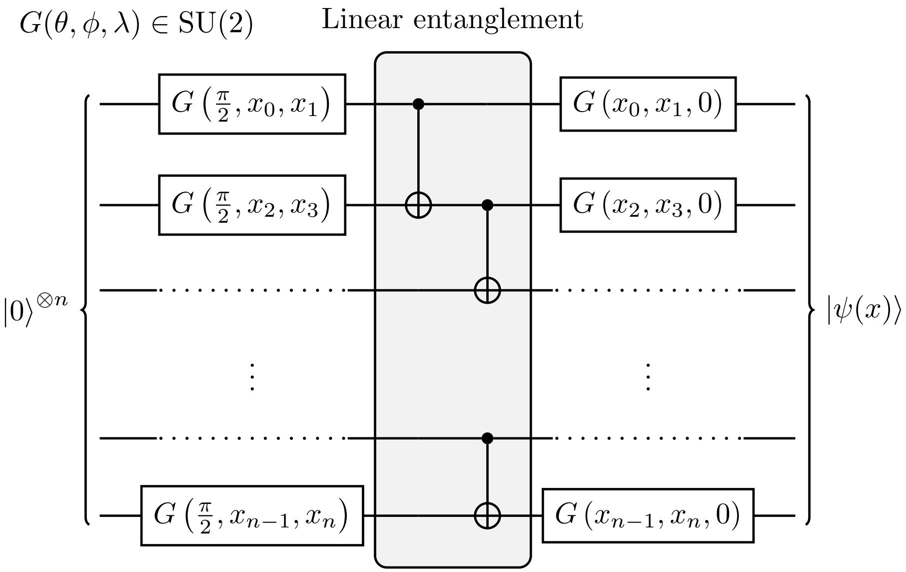
$$\Delta_{\mathrm{QC}}(\varepsilon_s) = \frac{\varepsilon_{\mathrm{b}}^{-1}(\varepsilon_s; Q)}{\varepsilon_{\mathrm{b}}^{-1}(\varepsilon_s; C)}$$

Quantum models are trained only using background data and learn to extract a metric of "typicality".

Kernel-based quantum anomaly detection

Unsupervised quantum kernel machine $\mathrm{K}_{ij} = |\langle 0|U^\dagger(\vec{x}_i)U(\vec{x}_j)|0\rangle|^2$

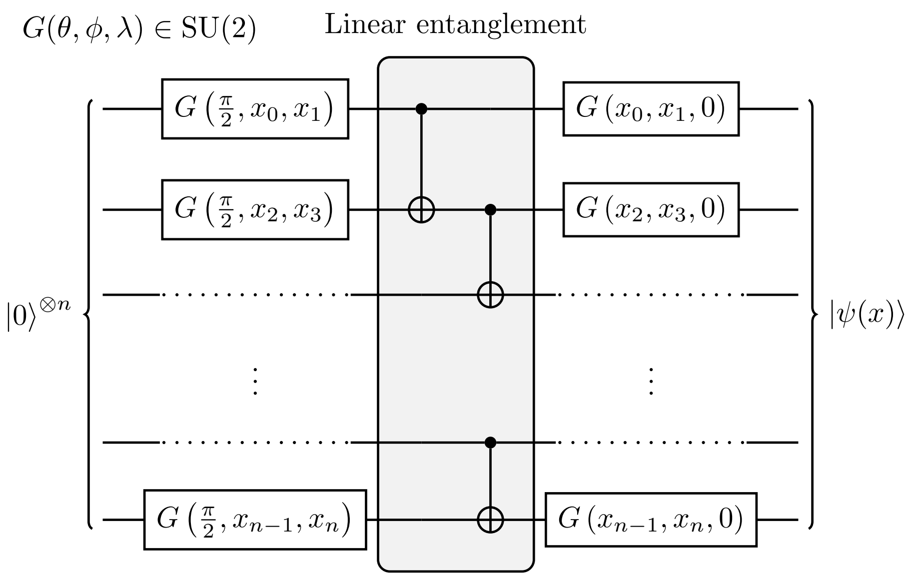
Designed data encoding circuit



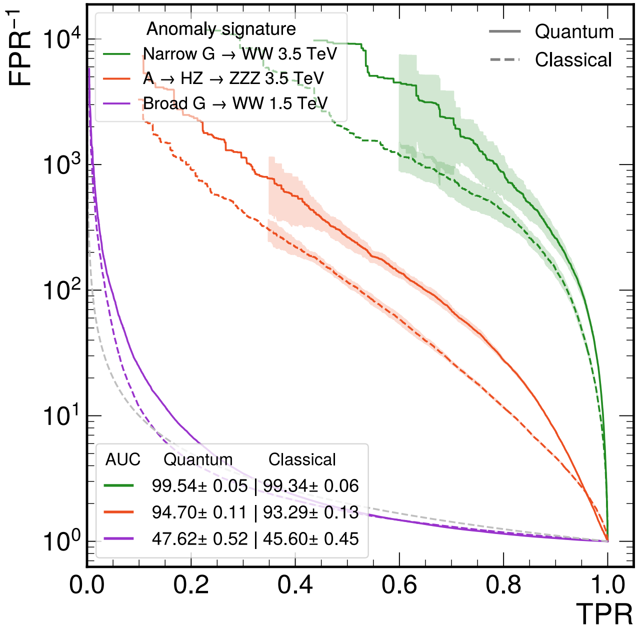
Kernel-based quantum anomaly detection

Unsupervised quantum kernel machine $\mathrm{K}_{ij} = |\langle 0|U^\dagger(\vec{x}_i)U(\vec{x}_j)|0\rangle|^2$

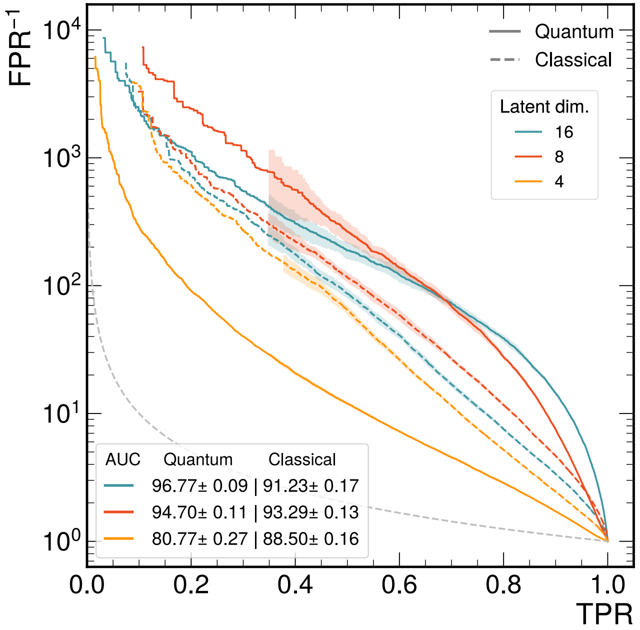
Designed data encoding circuit



Different BSM scenarios



Different qubit number

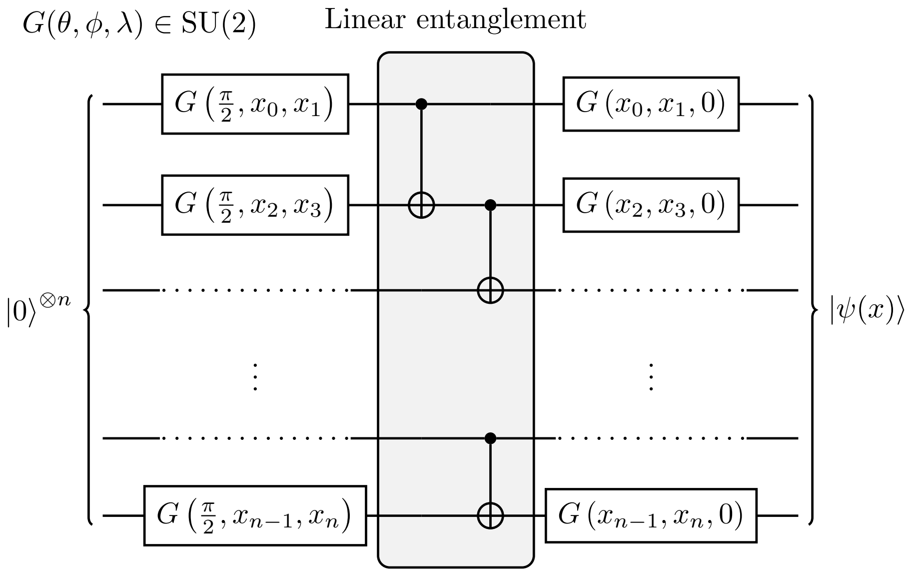


[K.A. Wozniak*, VB*, E. Puljak*, et al., arXiv: 2301.10780]

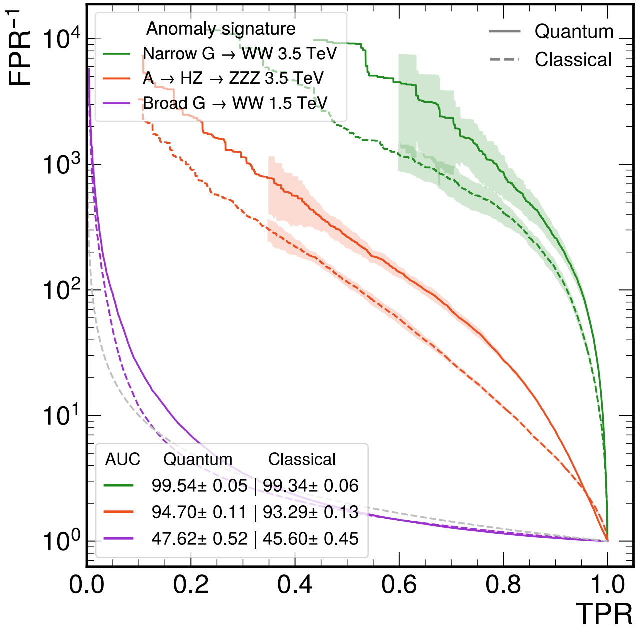
Kernel-based quantum anomaly detection

Unsupervised quantum kernel machine $\mathrm{K}_{ij} = |\langle 0|U^\dagger(\vec{x}_i)U(\vec{x}_j)|0\rangle|^2$

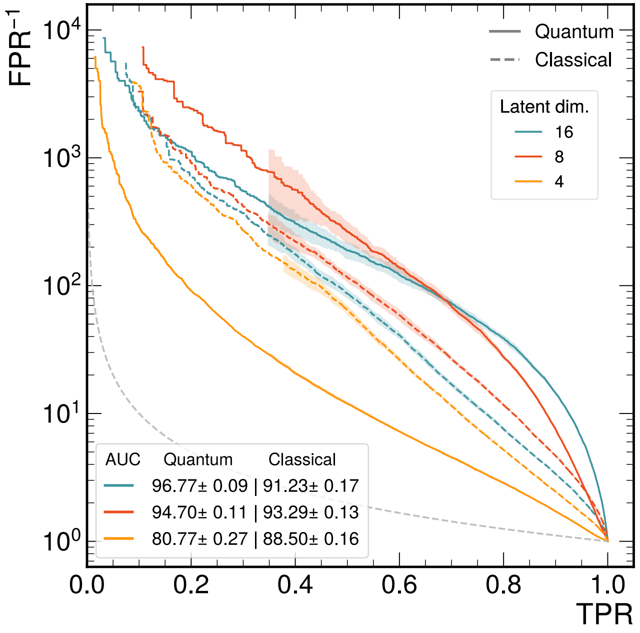
Designed data encoding circuit



Different BSM scenarios



Different qubit number



[K.A. Wozniak*, VB*, E. Puljak*, et al., arXiv: 2301.10780]

Instance of significant and consistent quantum performance advantage!

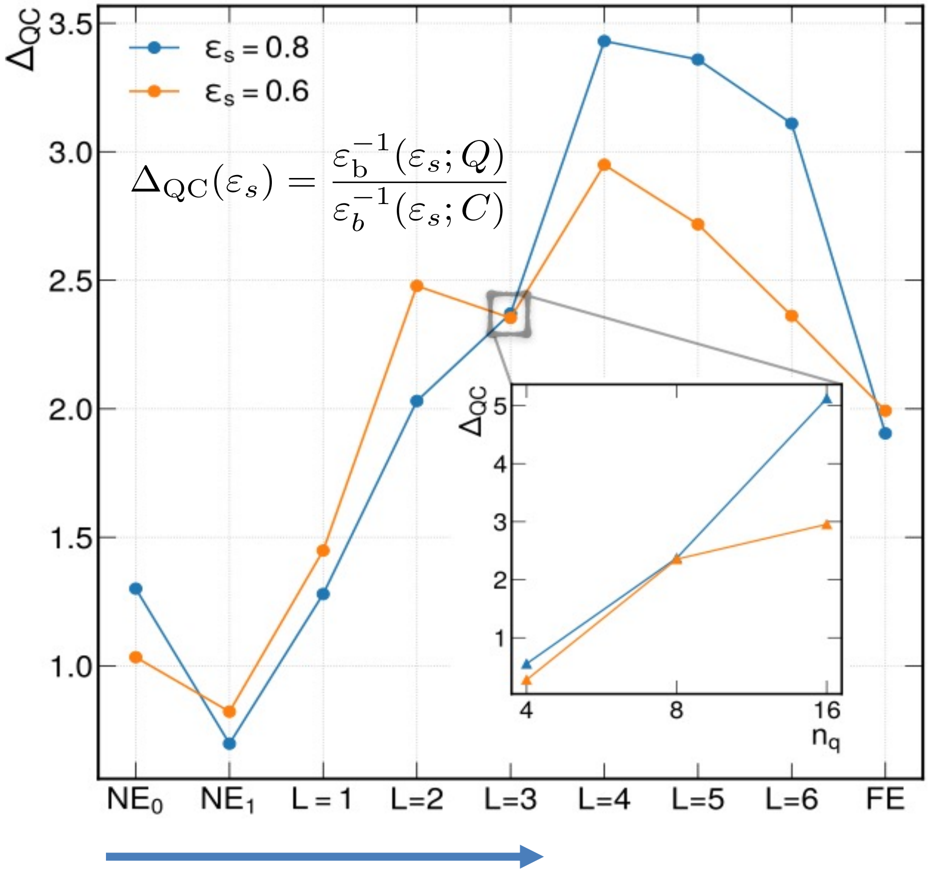
Very exciting and first of its kind result (HEP + Anomaly detection)!

Quantum circuit properties vs. performance

[K.A. Wozniak*, VB*, E. Puljak*, et al., arXiv: arXiv:2301.10780]

Performance vs. circuit architectures
Analysing circuit depth (expressibility) and amount entanglement

Importance of intrinsically quantum properties of the feature map.



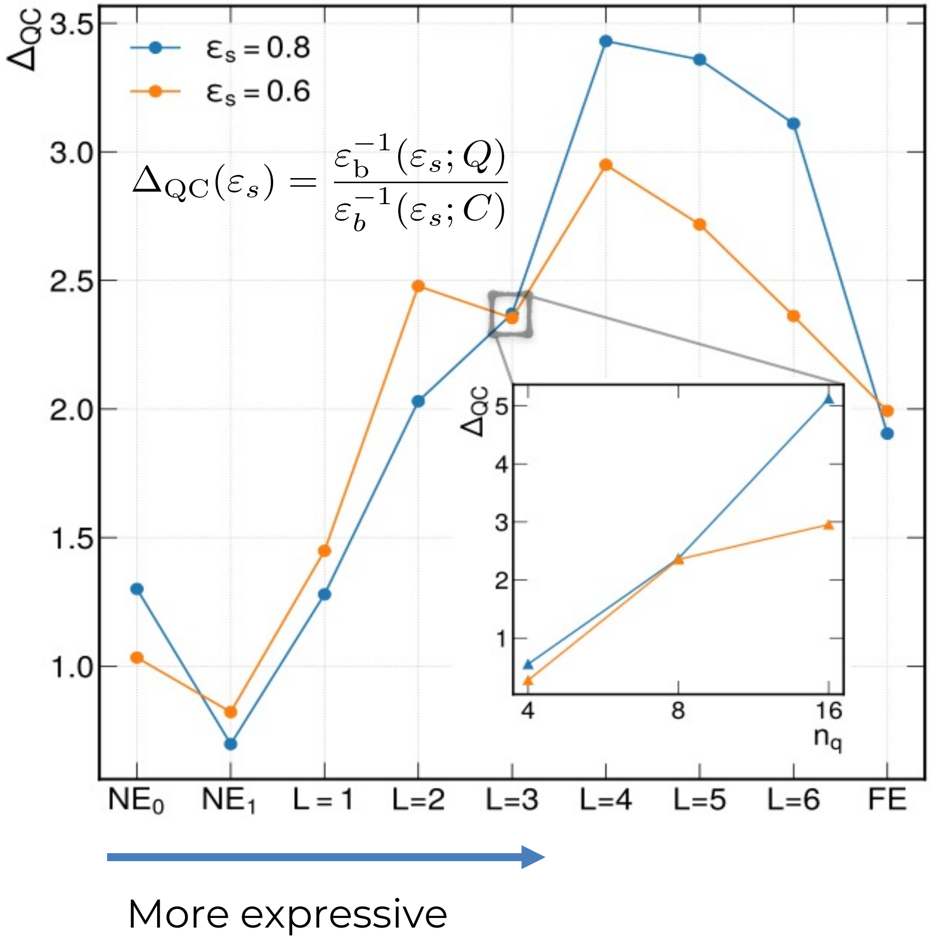
More expressive

[K.A. Wozniak*, VB*, E. Puljak*, et al., arXiv: arXiv:2301.10780]

Performance vs. circuit architectures
Analysing circuit depth (expressibility) and amount entanglement

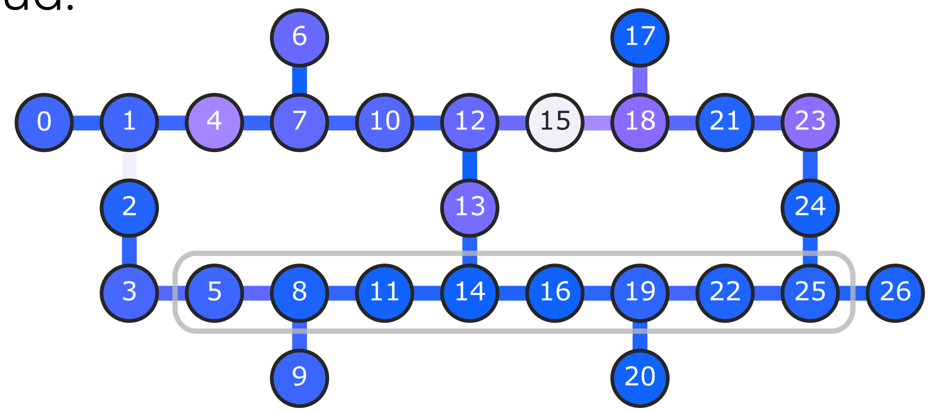
Importance of intrinsically quantum properties of the feature map.

Up to **five times** the performance of the classical model for 16 qubits!



Quantum hardware runs

Submit jobs to a real machine (ibm_toronto) using IBMQ cloud. (CERN quantum-hub)



Superconducting qubits connectivity topology

Map algorithm to hardware qubits.

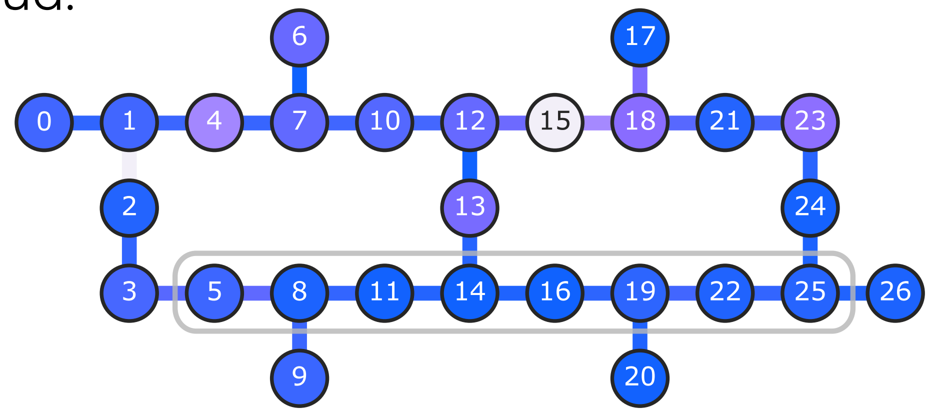
Minimal instance 100 + 100 (train + test) datapoints.

Quantum hardware runs

Submit jobs to a real machine (ibm_toronto) using IBMQ cloud.
(CERN quantum-hub)

Map algorithm to hardware qubits.

Minimal instance 100 + 100 (train + test) datapoints.



Superconducting qubits connectivity topology

Kernel Machine Run	AUC	$\langle \mathrm{tr}\rho^2 \rangle$
Hardware $L = 1$	0.844	0.271(6)
Ideal $L = 1$	0.999	1
Hardware $L = 3$	0.997	0.15(2)
Ideal $L = 3$	1.0	1
Classical	0.998	-

Purity of fully mixed state: $1/2^{n_q} \approx 0.39 \times 10^{-2}$
(decoherence = loss of "quantumness")

$$\langle \mathrm{tr}\rho^2 \rangle = \langle K(x_i, x_i) \rangle$$
$$\rho(x_i) = U(x_i)|0\rangle\langle 0|U^\dagger(x_i)$$

Proposed data encoding circuit realistic and suitable for current devices

Conclusions & Outlook

QML for (un)supervised learning: Fundamentally different data representation and processing.

Promising results identifying a **significant and consistent advantage** in anomaly detection!

Conclusions & Outlook

QML for (un)supervised learning: Fundamentally different data representation and processing.

Promising results identifying a **significant and consistent advantage** in anomaly detection!

Identify HEP problems that quantum models could have a natural inductive bias.

Parallel rapid development of quantum hardware.

Conclusions & Outlook

QML for (un)supervised learning: Fundamentally different data representation and processing.

Promising results identifying a **significant and consistent advantage** in anomaly detection!

Identify HEP problems that quantum models could have a natural inductive bias.

Parallel rapid development of quantum hardware.

For more details checkout:

- K.A. Wozniak*, VB*, E. Puljak*, et al., **Quantum anomaly detection in the latent space of proton collision events at the LHC,** arXiv:2301.10780
- J. Shuhmacher, L. Bogia, VB, et al. **Unravelling physics beyond the standard model with classical and quantum anomaly detection,** arXiv: 2301.10787
- VB, S. González-Castillo, et al., **Higgs analysis with quantum classifiers,** *EPJ Web Conf.*, 251 (2021) 03070

Conclusions & Outlook

QML for (un)supervised learning: Fundamentally different data representation and processing.

Promising results identifying a **significant and consistent advantage** in anomaly detection!

Identify HEP problems that quantum models could have a natural inductive bias.

Parallel rapid development of quantum hardware.

For more details checkout:

- K.A. Wozniak*, VB*, E. Puljak*, et al., **Quantum anomaly detection in the latent space of proton collision events at the LHC,** arXiv:2301.10780
- J. Shuhmacher, L. Bogia, VB, et al. **Unravelling physics beyond the standard model with classical and quantum anomaly detection,** arXiv: 2301.10787
- VB, S. González-Castillo, et al., **Higgs analysis with quantum classifiers,** *EPJ Web Conf.*, 251 (2021) 03070

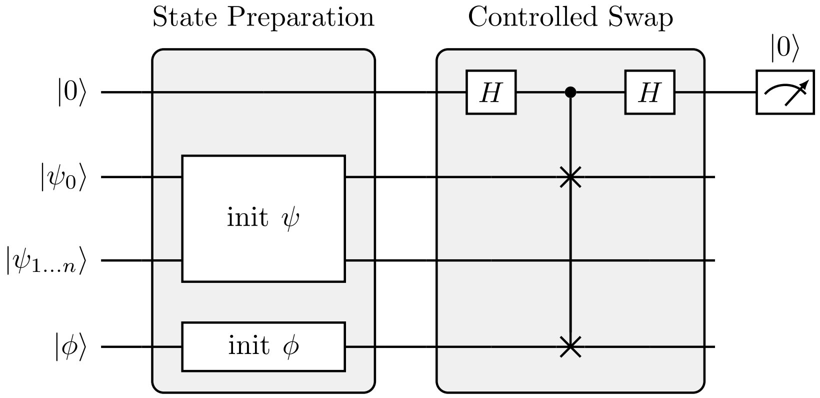
Questions?

Backup slides

Quantum clustering for anomaly detection

Construct clusters in the Hilbert space

Quantum distance calculation from clusters:

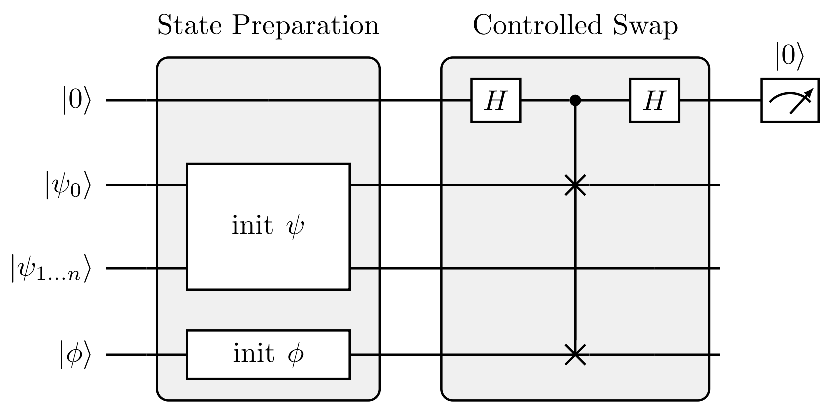


Minimise the distance with **quantum** (QK-means) or hybrid/**classical** (QK-medians) optimisation algorithms

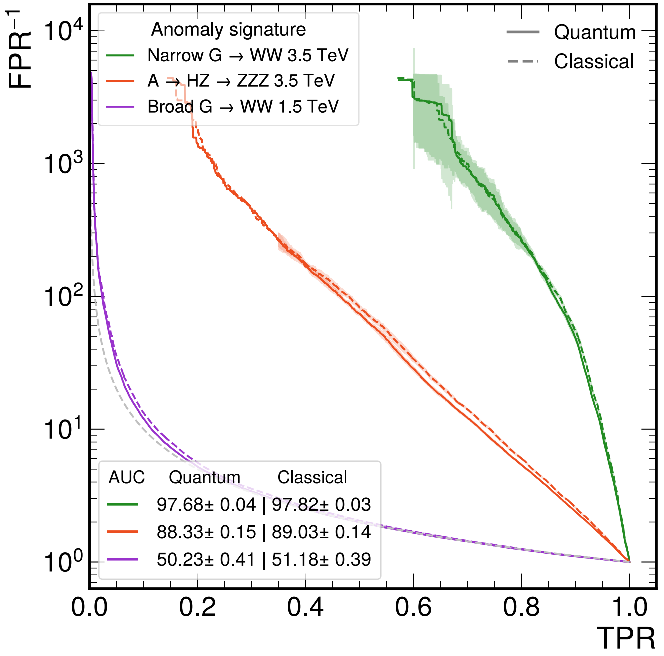
Quantum clustering for anomaly detection

Construct clusters in the Hilbert space

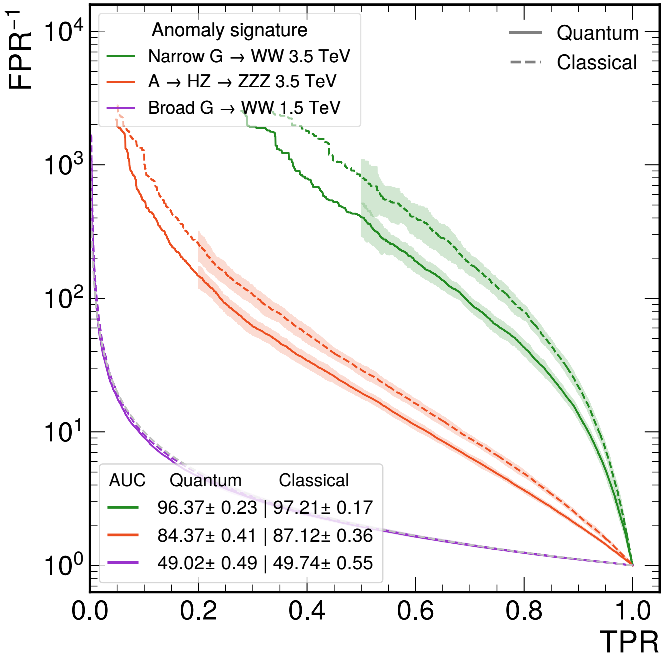
Quantum distance calculation from clusters



Quantum K-medians



Quantum K-means

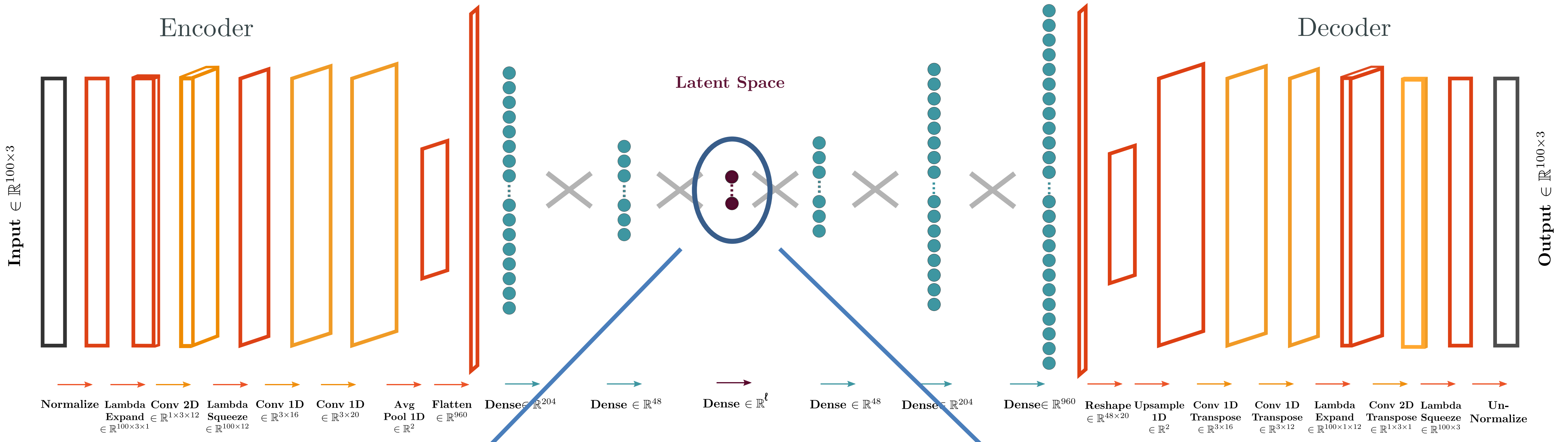


[K.A. Wozniak*, VB*, E. Puljak*, et al., arXiv: 2301.10780]

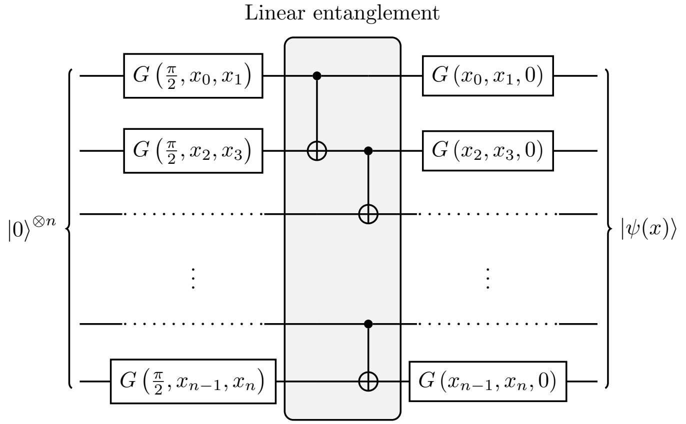
Minimise the distance with **quantum** (QK-means) or hybrid/**classical** (QK-medians) optimisation algorithms

Quantum and classical anomaly detection has similar performance.

Convolutional autoencoder architecture



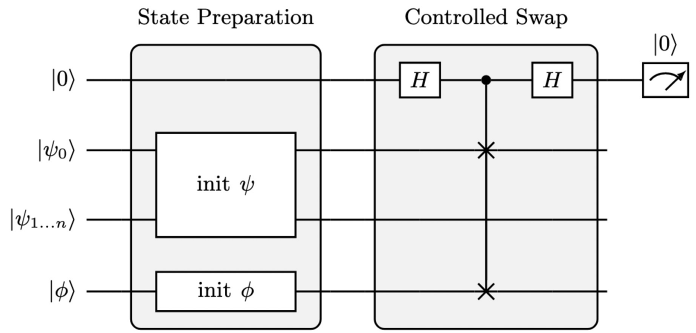
Unsupervised kernel machine



Training size:

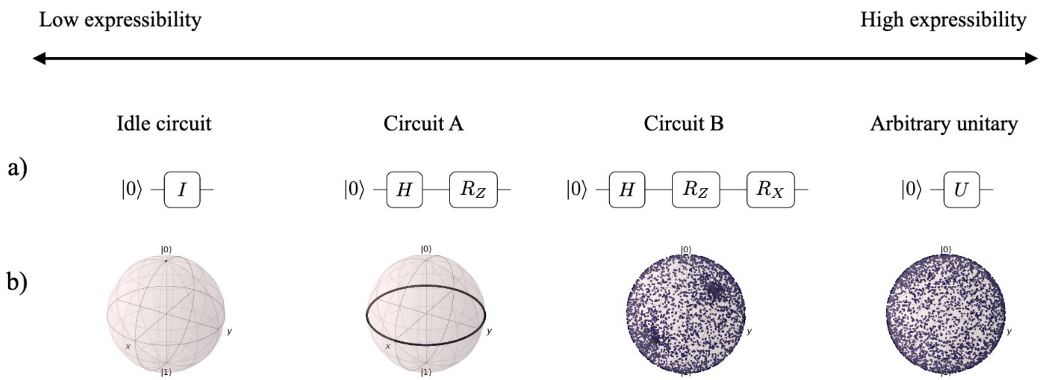
$\mathcal{O}(10^2) - \mathcal{O}(10^3)$

Quantum clustering algorithms

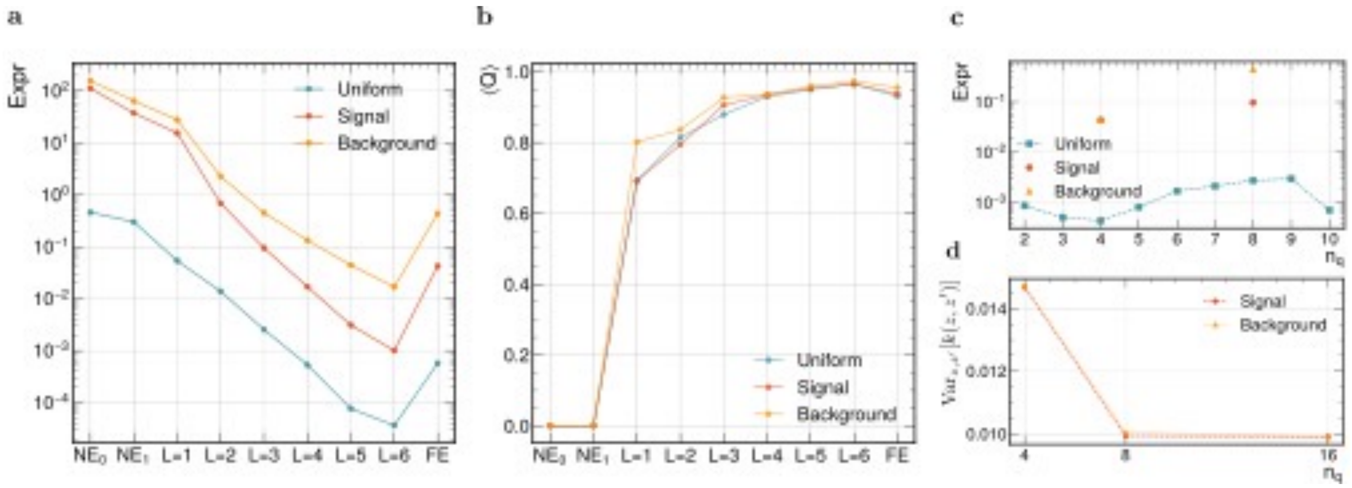


Expressibility and entanglement capability

Expressibility [S. Sim, et al., Adv. Quantum Technol. 2 (2019) 1900070]



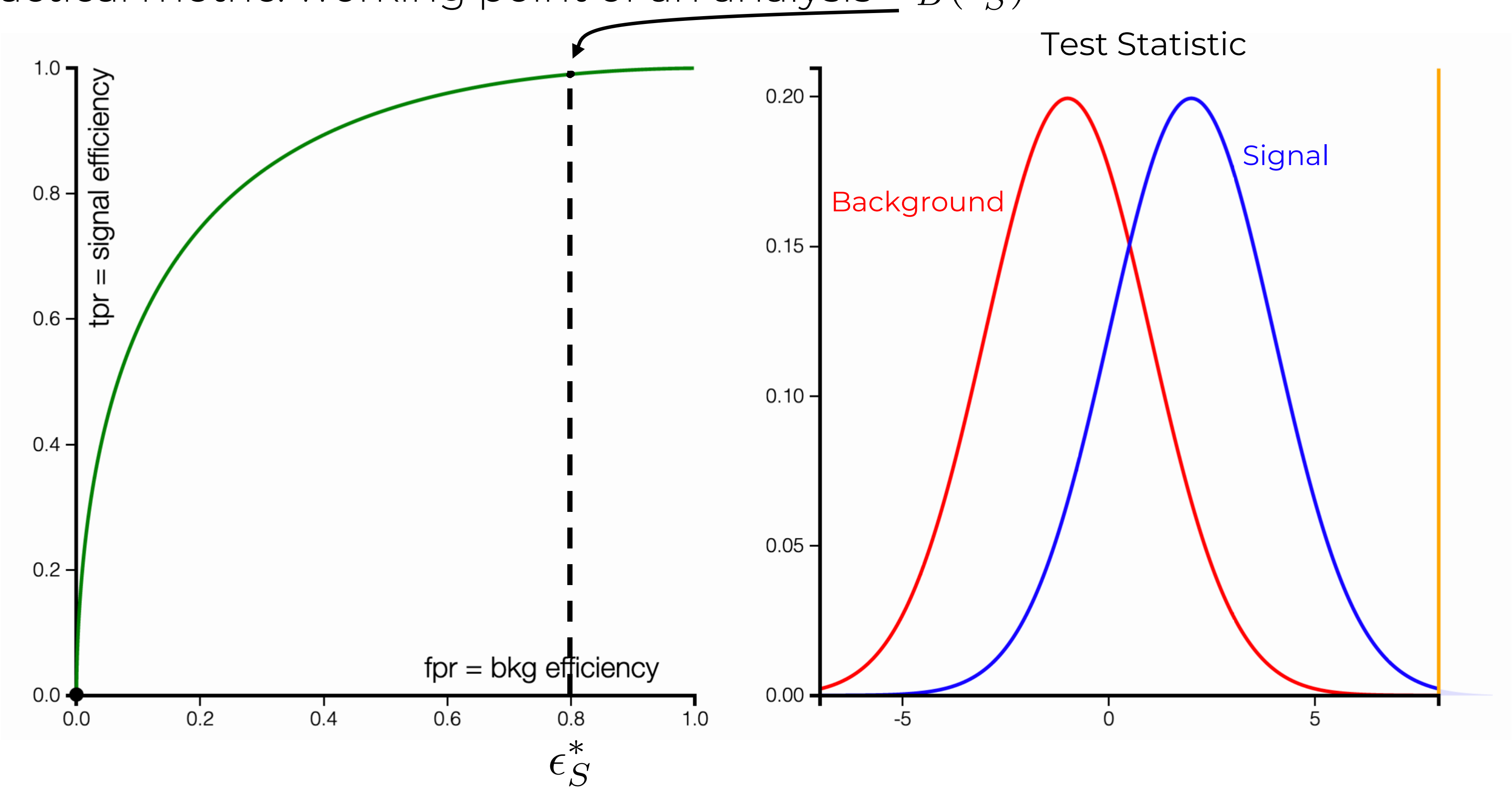
Expressibility & Entanglement capability of our data encoding circuit



[K.A. Wozniak*, **VB***, E. Puljak*, et al., arXiv:2301.10780]

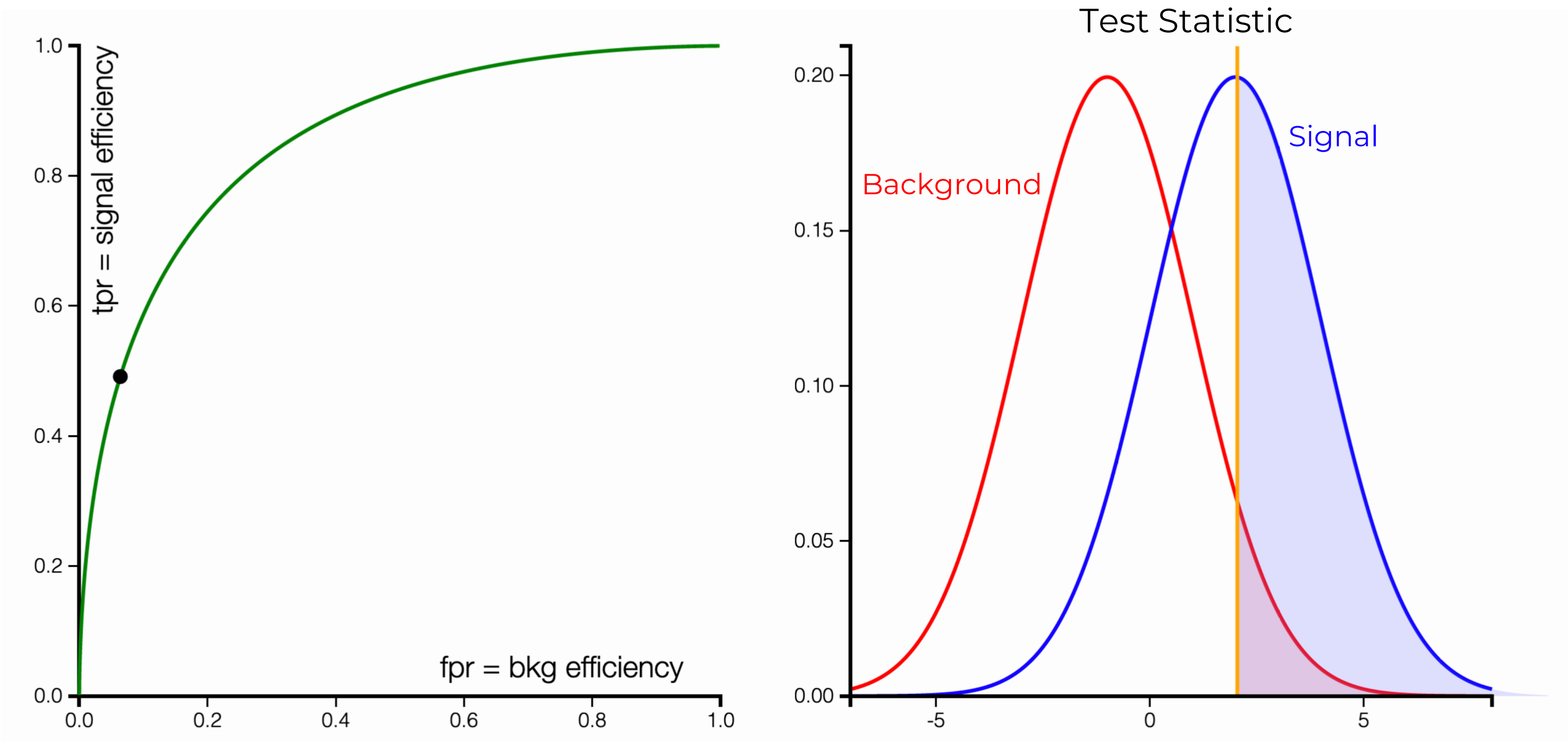
Performance Metric

- The normalised data samples are split into training, validation, and testing data sets.
- Classification power metric: Receiver Operating Characteristic (ROC) curve.
- More compact metric: Area Under Curve (AUC) of the ROC curve.
- More practical metric: working point of an analysis $\epsilon_B(\epsilon_S^*)$



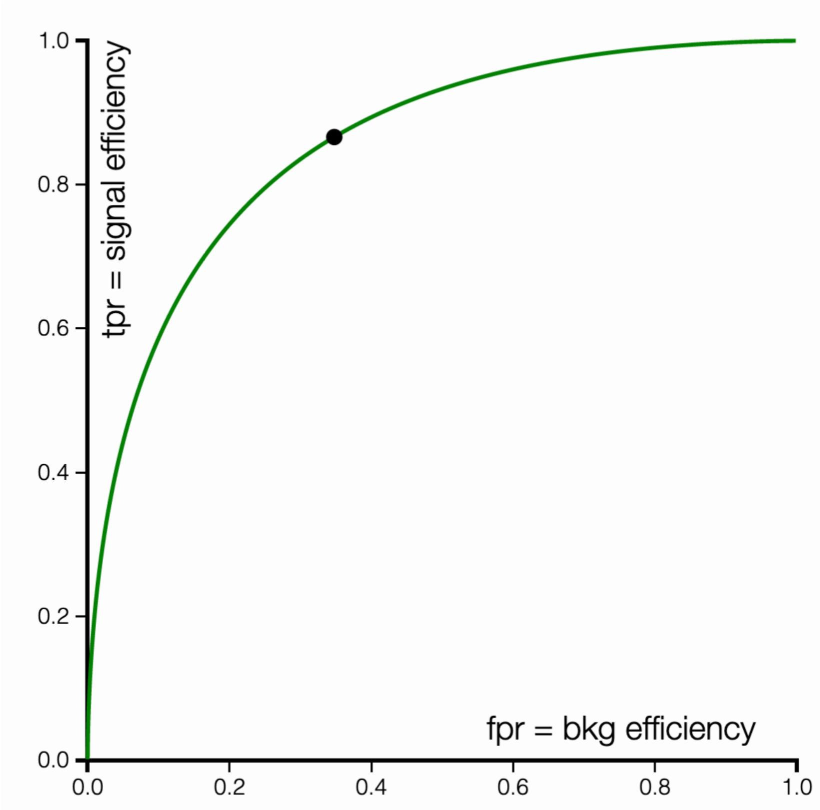
Performance Metric

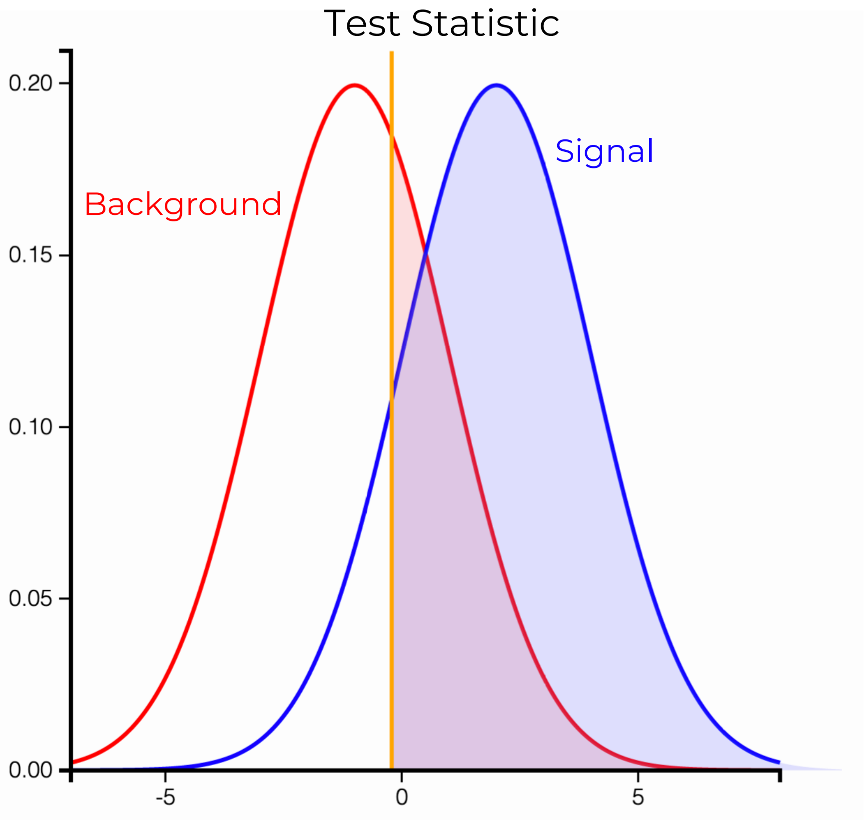
- The normalised data samples are split into training, validation, and testing data sets.
- Classification power metric: Receiver Operating Characteristic (ROC) curve.
- More compact metric: Area Under Curve (AUC) of the ROC curve.
- More practical metric: working point of an analysis $\epsilon_B(\epsilon_S^*)$



Performance Metric

- The normalised data samples are split into training, validation, and testing data sets.
- Classification power metric: Receiver Operating Characteristic (ROC) curve.
- More compact metric: Area Under Curve (AUC) of the ROC curve.
- More practical metric: working point of an analysis $\epsilon_B(\epsilon_S^*)$





Performance Metric

- The normalised data samples are split into training, validation, and testing data sets.
- Classification power metric: Receiver Operating Characteristic (ROC) curve.
- More compact metric: Area Under Curve (AUC) of the ROC curve.
- More practical metric: working point of an analysis $\epsilon_B(\epsilon_S^*)$