Re-discovery of $Z_c(3900)$ at BESIII Based on Quantum Support Vector Machine

Siyang Wu[1], Jiaheng Zou[2], Zhaoxia Meng[3], Teng Li[1], Hideki Okawa[2], Tao Lin[2], Zhiqing Liu[1], Weidong Li[2], Xingtao Huang[1]

1.Shandong University 2. Institute of High Energy Physics 3. University of Jinan

2026.08.17, Shenzhen, QCML 2026

Quantum Machine Learning for HEP

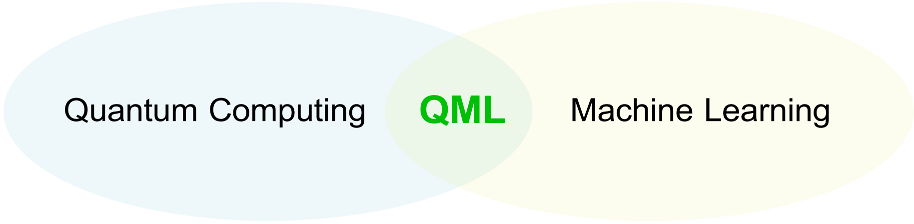
- **Machine Learning:**

Learn from data

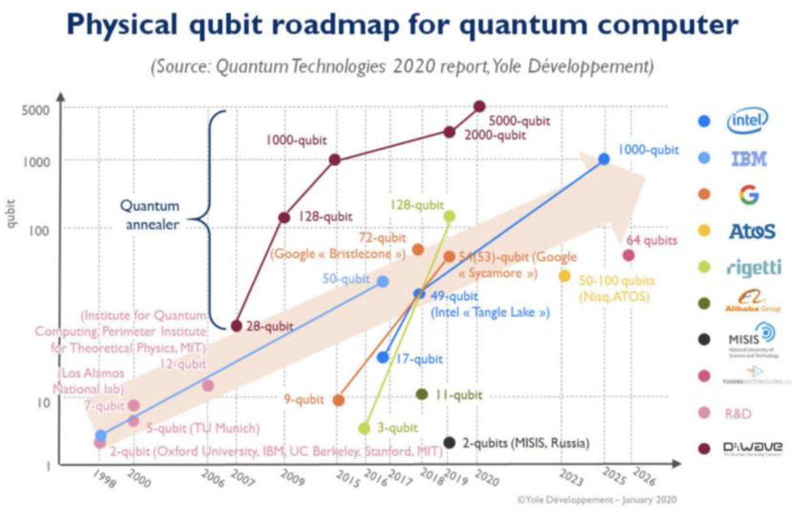
- **Quantum Computing:**

$$n \text{ qubits} \rightarrow 2^n \text{ states} \sim 2^n \text{ bits}$$

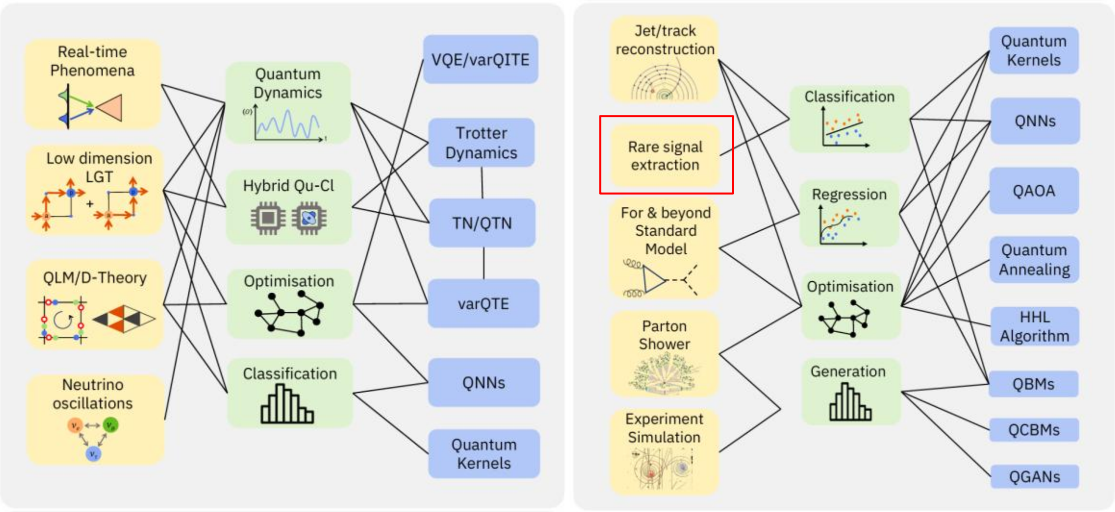
- **Quantum Machine Learning, QML:**



CERN Quantum Technology Initiative[1]



QC4HEP Working Group[2]



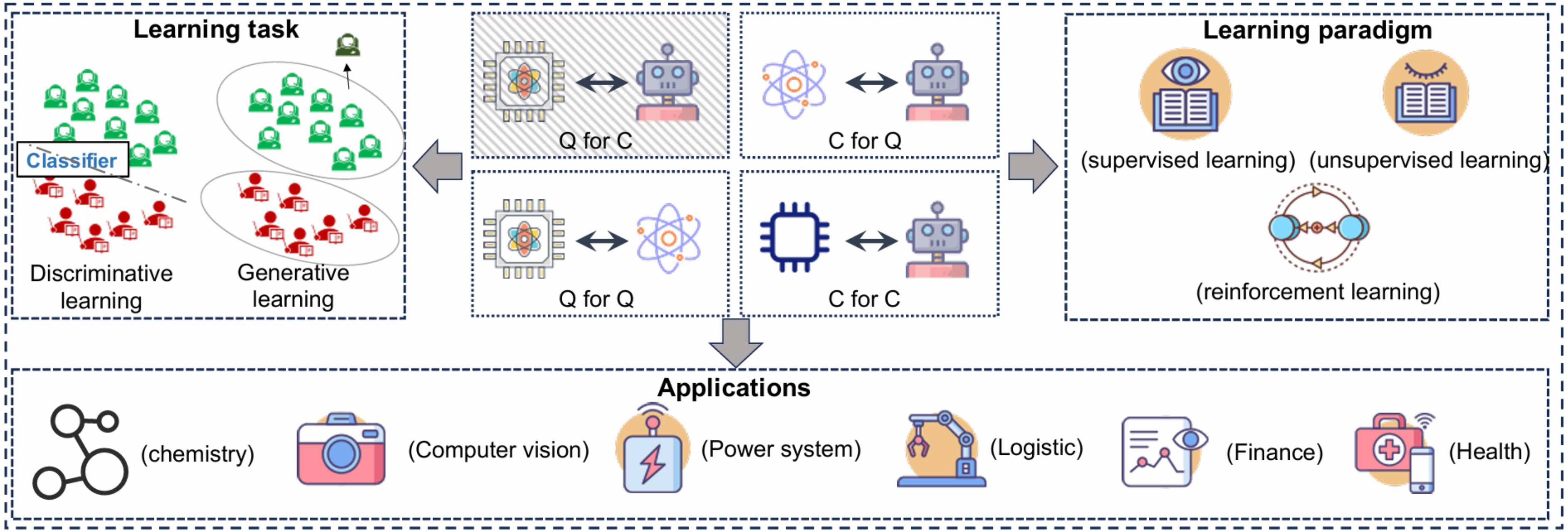
[1] *https://quantum.cern/quantum-computing-and-algorithms*

[2] *DOI: 10.1103/PRXQuantum.5.037001*

Quantum Machine Learning at NISQ era

Current characteristics of quantum machine learning:

- Limited circuit depth and qubit number.
- Hybrid quantum-classical architecture.
- Quantum processor: runs a parameterized quantum circuit (ansatz).
- Classical optimizer: adjusts the parameters to achieve the optimization goal.



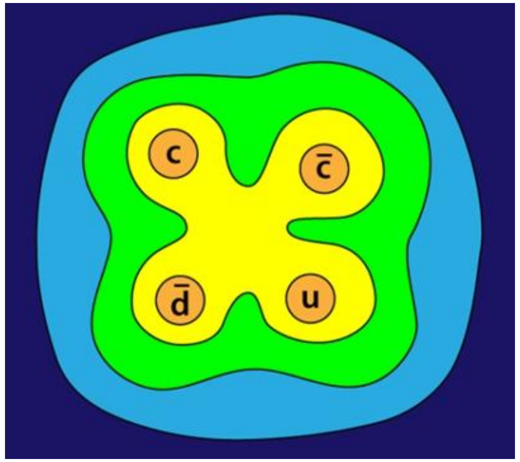
Quantum Machine Learning Categorized by Different Dimensions: learner and system, learning tasks, learning paradigms. QML has diverse applications. [3]

[3] *DOI: 10.48550/arXiv.2502.01146*

$Z_c(3900)$ Analysis as a Baseline

The discovery of $Z_c(3900)$

- Initially discovered by BESIII at 2013 [4].
- The first confirmed tetraquark state.
- Top of the eleven international physics breakthroughs of 2013 (Physics, APS).
- Clear analysis conditions, reliable references.

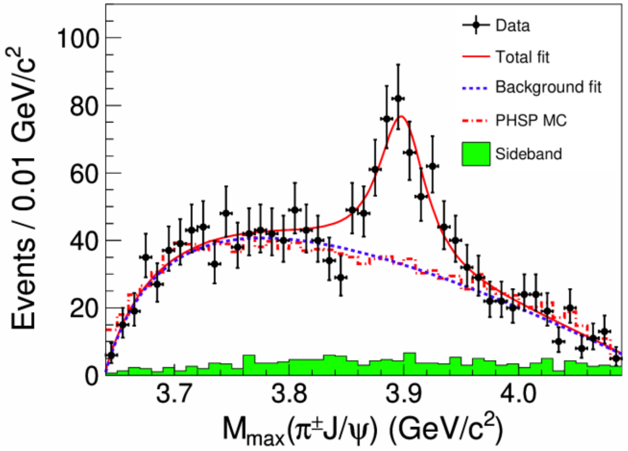


The probably structure of $Z_c(3900)$

Goal of this work

Apply QSVM method to $Z_c(3900)$ analysis workflow.

- Using the same dataset.
- Building the analysis chains with QSVM.
- Experimental validation of QML feasibility.



$Z_c(3900)$ signal observed by BESIII

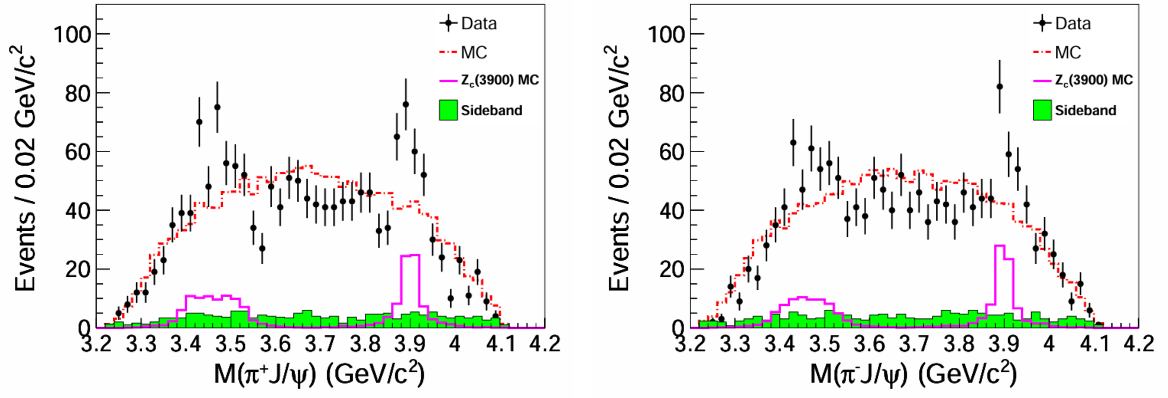
[4] *DOI: 10.1103/PhysRevLett.110.252001*

$Z_c(3900)$ Discovery at BESIII

Decay process:

$$e^+e^- \to Y(4260) \to Z_c(3900)\ \pi^{\pm} \to J/\psi\ \pi^{\mp}\pi^{\pm} \to l^+l^-\pi^+\pi^-,\ (l = e \text{ or } \mu)$$

The final state consists of 4 charged tracks, which leptons and pions are separated clearly by the momentum as shown later.



Event selection criteria:

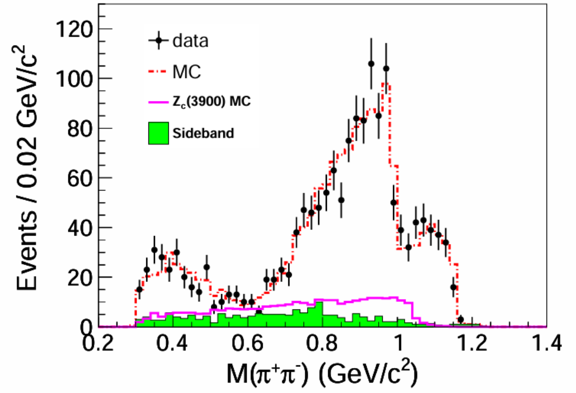
A. 4 tracks with netcharge = 0.

B. Four-constraint (4C) kinematic.

Traditional cut-based method for J/ψ selection:

- Signal: $3.08\ \text{GeV}/c^2 < M_{l^+l^-} < 3.12\ \text{GeV}/c^2$
- Background: $3.0\ \text{GeV}/c^2 < M_{l^+l^-} < 3.06\ \text{GeV}/c^2$ and

  $3.14\ \text{GeV}/c^2 < M_{l^+l^-} < 3.20\ \text{GeV}/c^2$



Quantum Support Vector Machine

Support Vector Machine (SVM) contains:

SVM Classification (SVC) SVM Regression (SVR)

For a given dataset

$$\mathcal{D}=\{(\boldsymbol{x}^{(i)},y^{(i)})\}_{i=1}^{n},\ \boldsymbol{x}^{(i)}\in\mathbb{R}^{d},\ y^{(i)}\in\{-1,1\}$$

If the dataset is linearly separable, there exists a hyperplane

$$\boldsymbol{\omega}\cdot\boldsymbol{x}+b=0$$

such that

$$y^{(i)}(\boldsymbol{\omega}\cdot\boldsymbol{x}^{(i)}+b)\geqslant 0$$

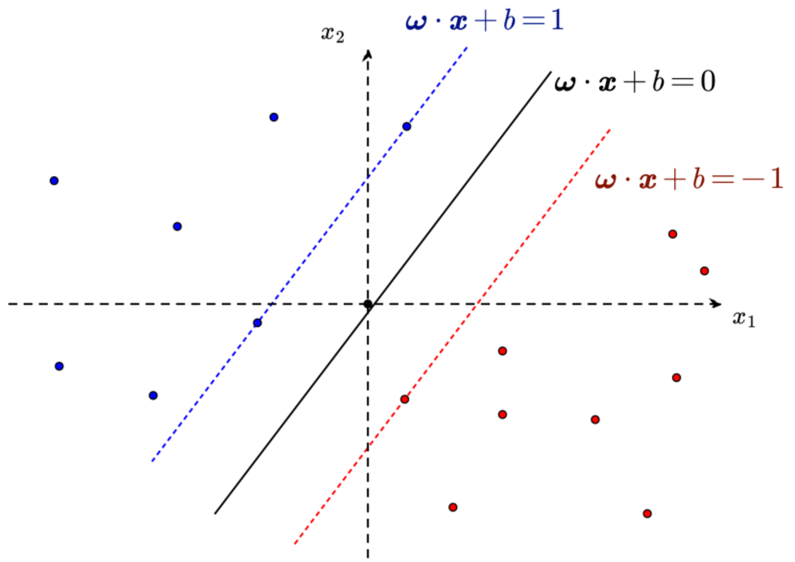
In general, the goal of an SVM classification task is to minimize

$$\frac{|\boldsymbol{\omega}|^{2}}{2}+C\sum_{i=1}^{n}\xi_i$$

$$\xi_i=\max(0,1-y^{(i)}(\boldsymbol{\omega}\cdot\boldsymbol{x}^{(i)}+b))$$

Hyperparameter C controls the trade-off:

- maximize the margin $2/|\boldsymbol{\omega}|$.
- minimize falling within the margin or are misclassified.



Dual form:

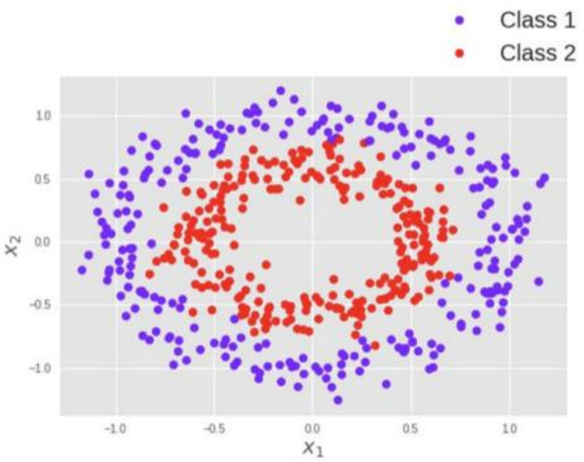
$$\mathcal{L}(\boldsymbol{a})=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a^{(i)}a^{(j)}y^{(i)}y^{(j)}\boldsymbol{x}^{(i)}\cdot\boldsymbol{x}^{(j)}-\sum_{i=1}^{n}a^{(i)}$$

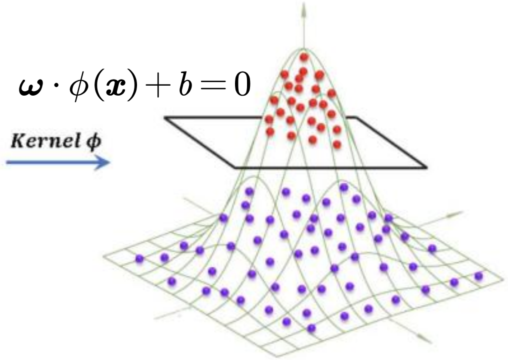
subject to

$$\sum_{i=1}^{n}a^{(i)}y^{(i)}=0,\ 0\leqslant a^{(i)}\leqslant C$$

Quantum Support Vector Machine

For non-linear task, a feature map $\phi: \boldsymbol{x} \rightarrow \phi(\boldsymbol{x}) \in \mathbb{R}^D$ need to be introduced, that the hyperplane can be





The dual form loss function then is

$$\mathcal{L}(\boldsymbol{a}) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a^{(i)} a^{(j)} y^{(i)} y^{(j)} \phi^{(i)}(\boldsymbol{x}) \cdot \phi^{(j)}(\boldsymbol{x}) - \sum_{i=1}^{n} a^{(i)}$$

To circumvent the computational cost of explicitly calculating the feature, define

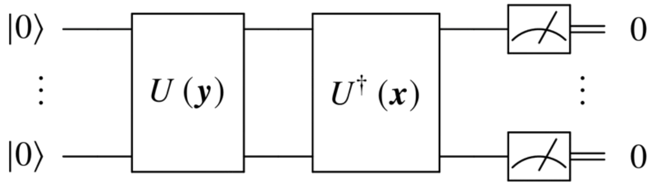
$$K_{ij} = \phi^{(i)}(\boldsymbol{x}) \cdot \phi^{(j)}(\boldsymbol{x})$$

For QSVM method, a feature vector $\boldsymbol{x}$ will be encoded into a quantum state $|\psi(\boldsymbol{x})\rangle$ by the circuit $U(\boldsymbol{x})$.

Kernel matrix K_Q is the inner product of two quantum state:

$$K_Q: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$$
$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto K_Q(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{Tr}(\rho(\boldsymbol{x})\rho(\boldsymbol{y})) = |\langle \psi(\boldsymbol{x}) | \psi(\boldsymbol{y}) \rangle|^2$$

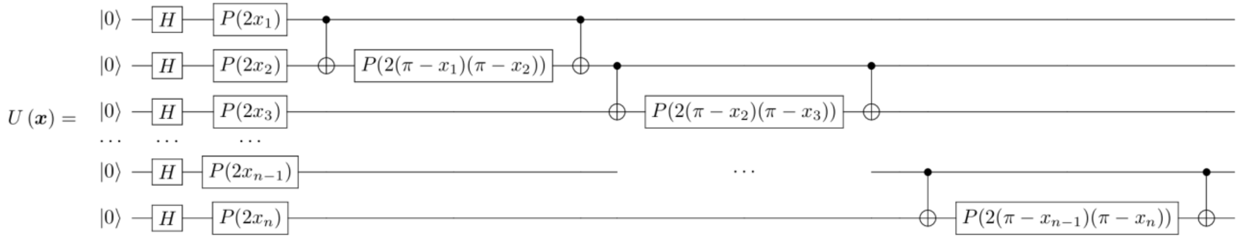
To get the inner product form K_Q, the Loschmidt echo test need to be performed:



Measuring the state, the probability that all the qubits in $|0\rangle$ is

$$|\langle \psi | U^\dagger(\boldsymbol{x}) U(\boldsymbol{y}) | \psi \rangle|^2 = |\langle \psi(\boldsymbol{x}) | \psi(\boldsymbol{y}) \rangle|^2 = K_Q(\boldsymbol{x}, \boldsymbol{y})$$

In this task, the feature encoding circuit is ZZFeatureMap:



Workflow

1. Event pre-selection.

- Dataset: 525 pb^{-1} collected at 4.26 GeV.
- J/ψ mass cut.

2. Training dataset selection.

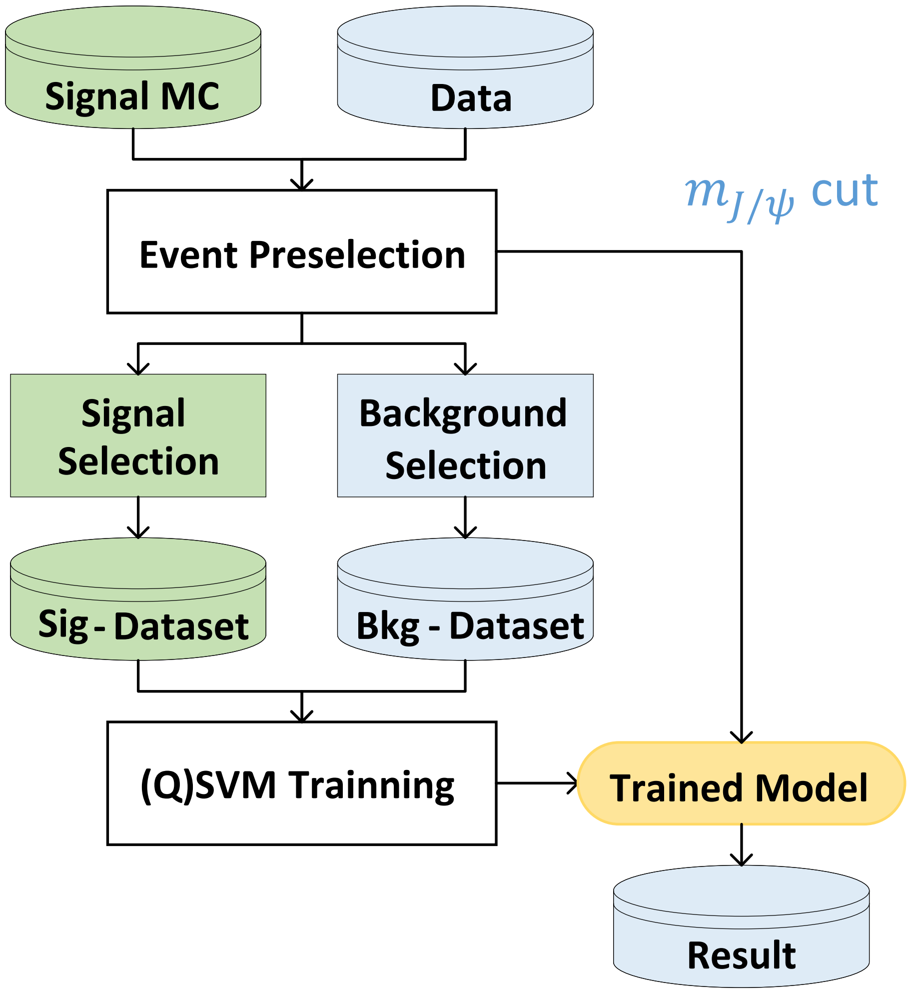
- Signal: Monte Carlo (MC) data in the signal region.
- Background: BESIII events in the sideband background region.

3. (Q)SVM model training and optimization.

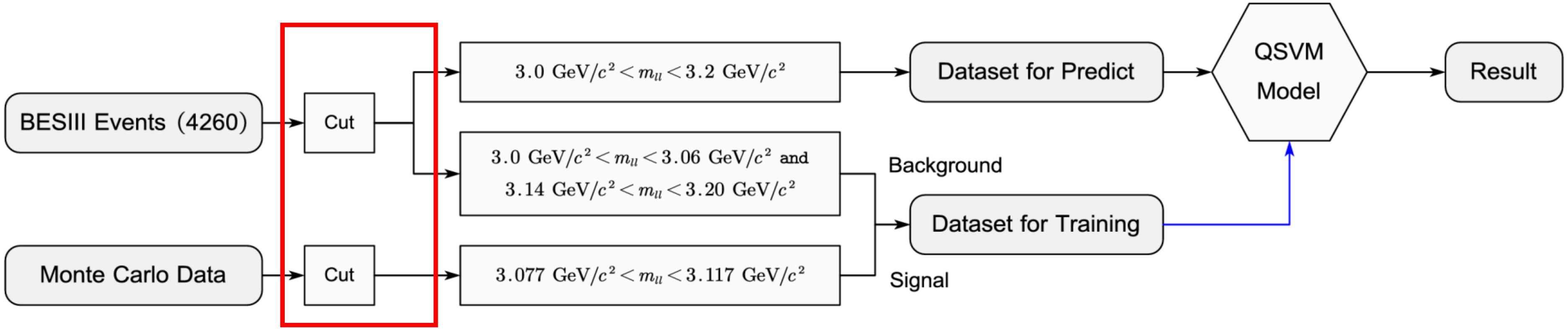
- Feature importance ranking.
- QSVM model optimization.
- SVM model optimization.

4. Model selection for J/ψ events.

5. Fit the $Z_c(3900)$ signal and get the result.



Event Pre-selection



A. 4 tracks with netcharge = 0.

Sort charged particles by their momentum, and denoted l_1, l_2, π_1, π_2 separately.

- Momentum:

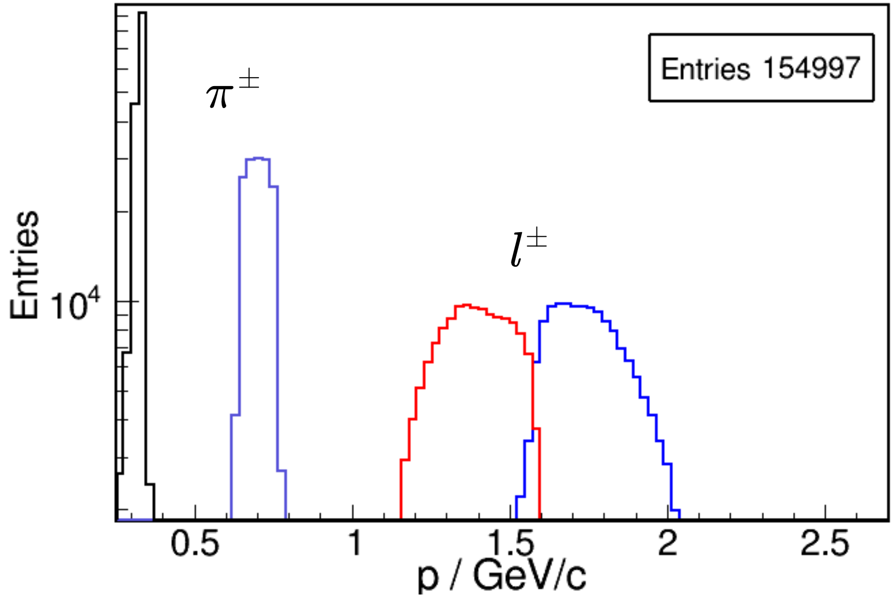
$$\text{For } l_1 \text{ and } l_2: P > 1.0\ \text{GeV},$$
$$\text{For } \pi_1 \text{ and } \pi_2: P < 1.0\ \text{GeV}.$$

- Energy deposition in EMC

$$ee \text{ mode}: E_{l_1} > 1.1\ \text{GeV},\ E_{l_2} > 0.8\ \text{GeV},$$
$$\mu\mu \text{ mode}: E_{l_1} < 0.35\ \text{GeV},\ E_{l_2} < 0.35\ \text{GeV}.$$

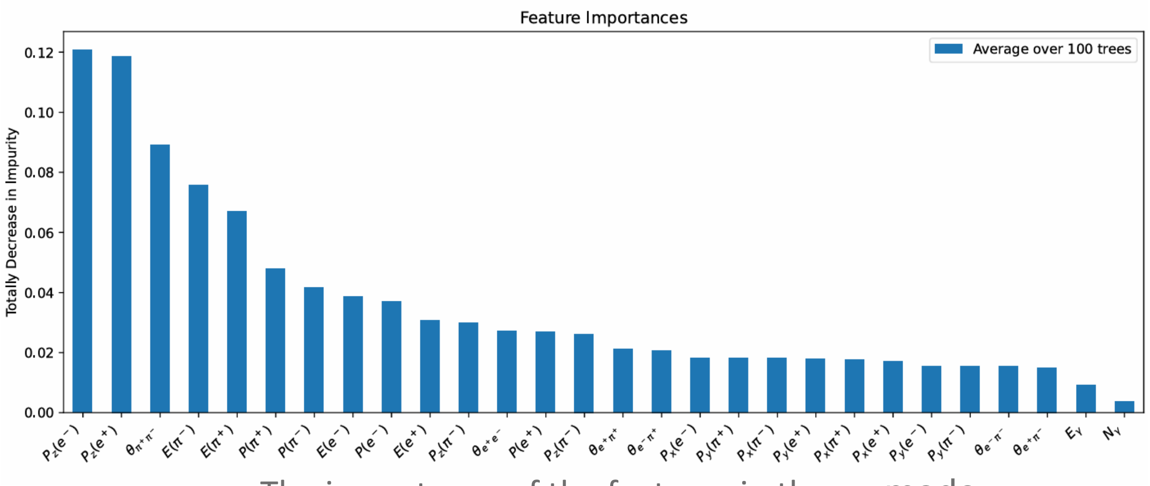
- Angles between the tracks

$$ee \text{ mode: } \cos(\theta_{e^+\pi^-}),\ \cos(\theta_{e^-\pi^+}),\ \cos(\theta_{\pi^-\pi^+}) < 0.98,$$
$$\mu\mu \text{ mode: } \cos(\theta_{\pi^-\pi^+}) < 0.98.$$

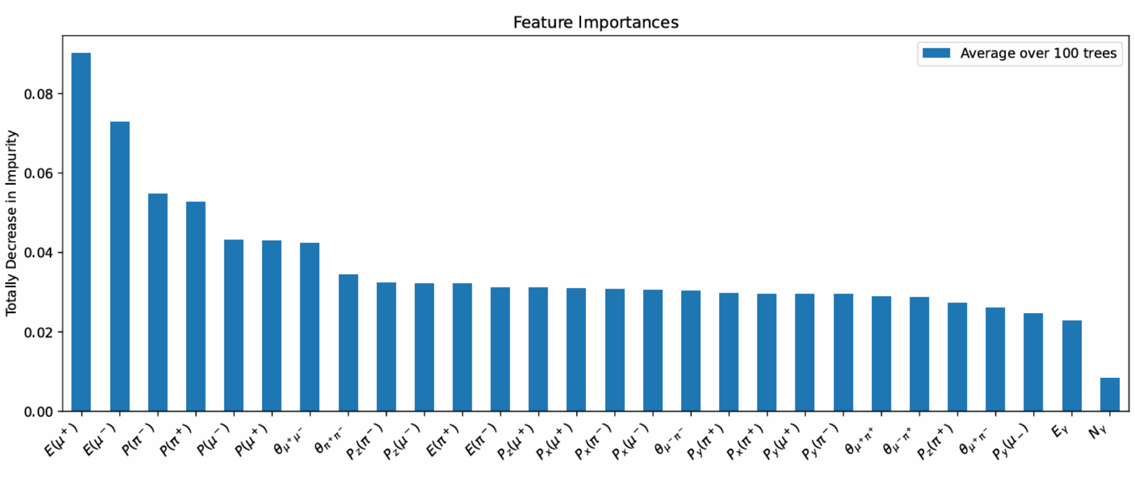


B. Apply 4C kinematic fitting.

Feature Importance Ranking



The importance of the features in the ee mode.



The importance of the features in the $\mu\mu$ mode.

Totally 28 feature candidates:

- Momentum:

$$|P(l_1)|,\ P_x(l_1),\ P_y(l_1),\ P_z(l_1).$$
$$|P(l_2)|,\ P_x(l_2),\ P_y(l_2),\ P_z(l_2).$$
$$|P(\pi_1)|,\ P_x(\pi_1),\ P_y(\pi_1),\ P_z(\pi_1).$$
$$|P(\pi_2)|,\ P_x(\pi_2),\ P_y(\pi_2),\ P_z(\pi_2).$$

- Energy:

$$E(l_1),\ E(l_2),\ E(\pi_1),\ E(\pi_2).$$

- Angle:

$$\theta_{l_1 l_2},\ \theta_{l_1 \pi_1},\ \theta_{l_1 \pi_2},\ \theta_{l_2 \pi_1},\ \theta_{l_2 \pi_2},\ \theta_{\pi_1 \pi_2}.$$

- Photons:

$$N_\gamma,\ E_{\gamma\,\max}.$$

Reduce the input dimensionality.

QSVM Model Optimization

Adjustable Variables:

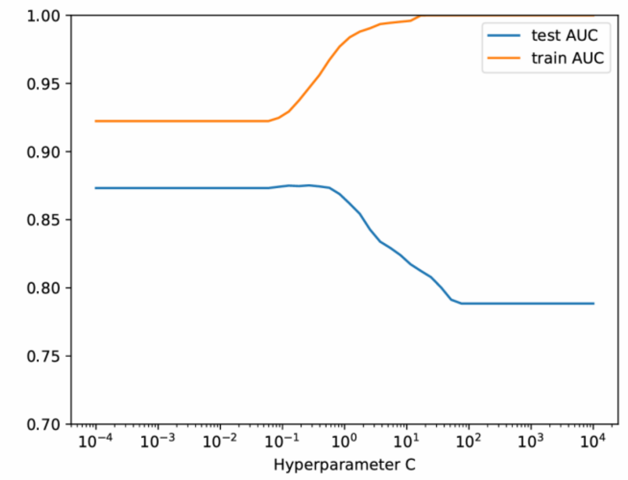
- Feature Candidates.
- Hyperparameter C.
- Feature Scaling Method.

Optimal Selection:

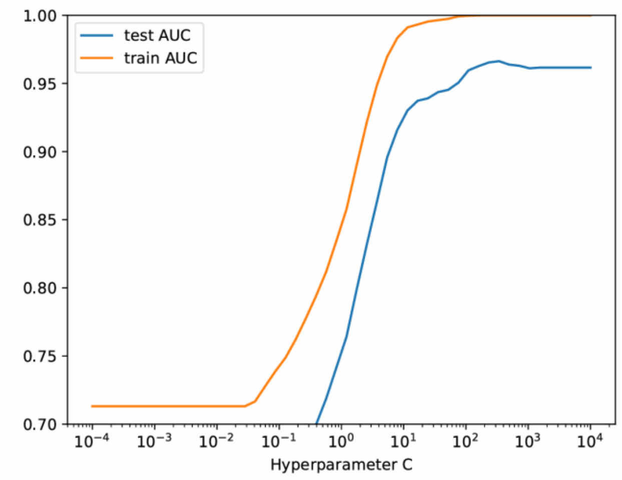
- ee mode: 6 features, scaled to $(0, \pi)$.
- $\mu\mu$ mode: 7 features, without scaler.

Conclusions:

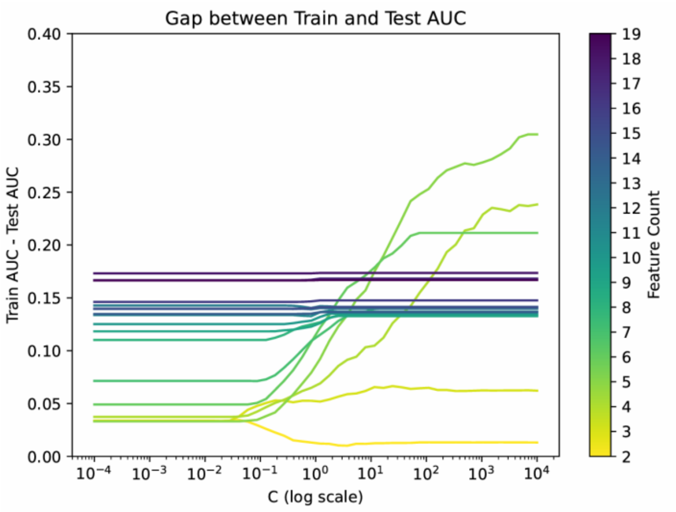
- More features cause stronger overfitting.
- ee mode: lower C performs better.
- $\mu\mu$ mode: higher C performs better.



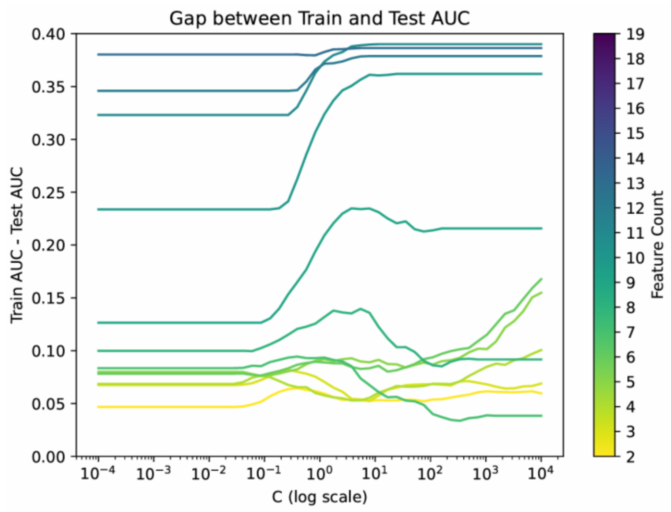
AUC curves for train and test with C (ee mode).



AUC curves for train and test with C ($\mu\mu$ mode).



AUC gap between training and testing in the ee mode.



AUC gap between training and testing in the $\mu\mu$ mode.

SVM Model Optimization

Same Parameters as QSVM:

- Selected features
- Feature Scaling Method

Differences from QSVM:

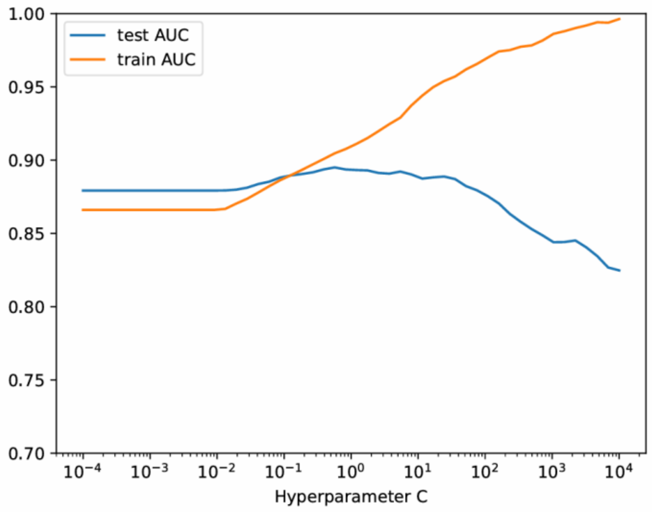
- Kernel function choice is different.
- C tuned to the optimal value for SVM.

Purpose:

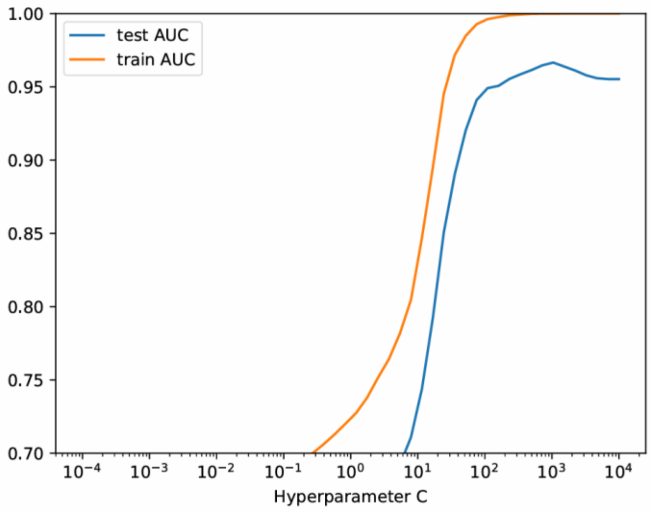
- Compare the performance of QSVM and SVM under the same dataset conditions.

Conclusions:

- SVM performs reliably on the test set, shows limited overfitting.



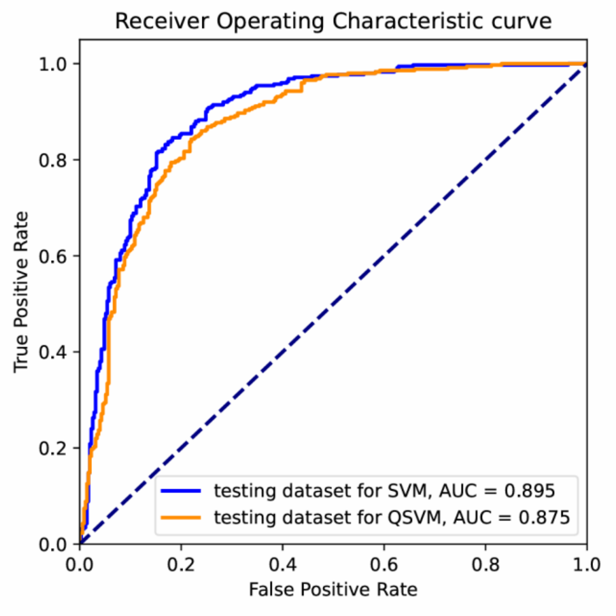
AUC curves for train and test with C (ee mode).



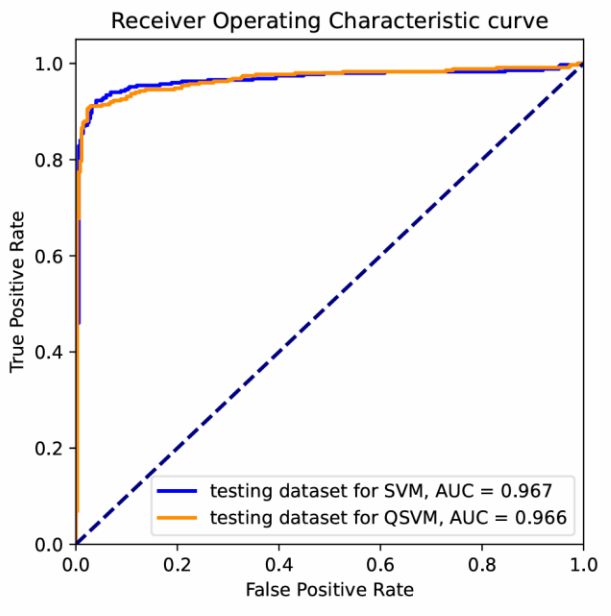
AUC curves for train and test with C ($\mu\mu$ mode).

Model Performance Comparison

- Both QSVM and SVM are trained under their respective optimal configurations.
- ROC curves are used as a comprehensive performance indicator.
- QSVM and SVM models exhibit consistent performance.



QSVM vs. SVM test ROC (ee mode).



QSVM vs. SVM test ROC ($\mu\mu$ mode).

Optimal parameters for QSVM training in ee and $\mu\mu$ modes.

mode	feature candidates	hyperparameter C
ee	$P_z(e^-), P_z(e^+), \theta_{\pi^+\pi^-}, E(\pi^-), E(\pi^+), P(\pi^+)$	2.683×10^{-1}
$\mu\mu$	$E(\mu^+), E(\mu^-), P(\pi^-), P(\pi^+), P(\mu^-), P(\mu^+), \theta_{\mu^+\mu^-}$	3.393×10^2

Preliminary Result

Combine signal data from ee and $\mu\mu$ modes,
fit $M_{\max}(\pi^{\pm}J/\psi)$:

Fitting Function:

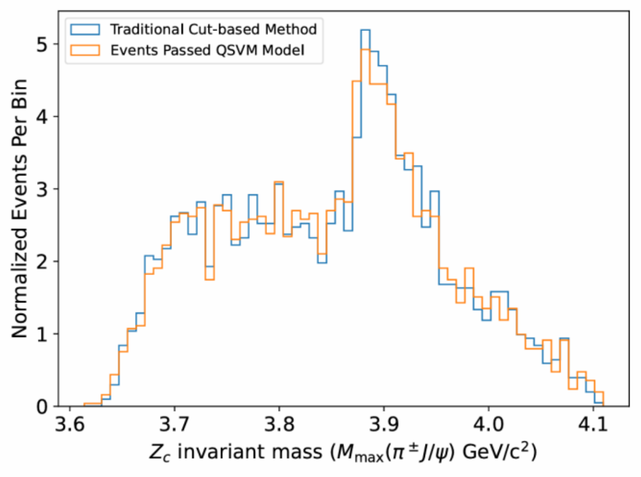
- Signal: Convolution of Breit-Wigner and Gaussian.
- Background: Third-order Chebychev polynomial.

Results:

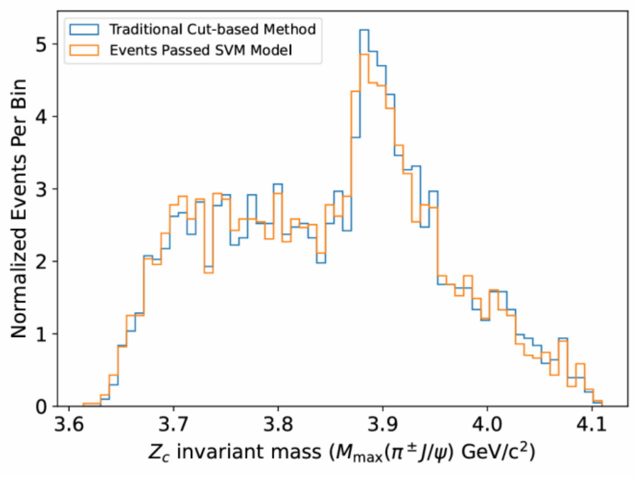
- Highly consistent with traditional analysis and PDG data.

Comparison of $Z_c(3900)$ parameters with different methods.

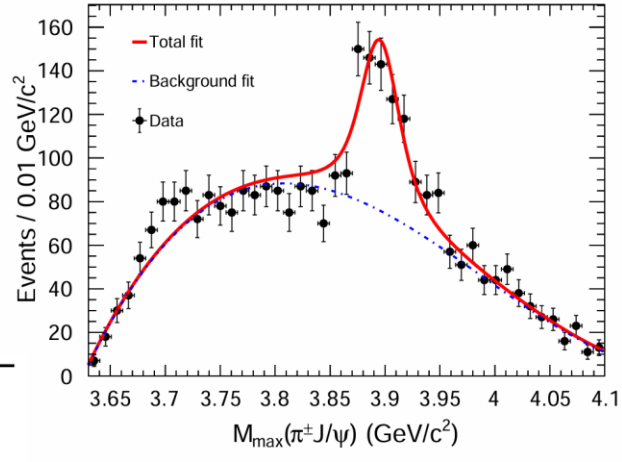
Method / (BOSS Version)	Cut-based / (6.6.3)	QSVM / (7.0.3)	SVM / (7.0.3)	PDG data
Mass (MeV/c^2)	3899.0 ± 3.6	3895.3 ± 2.6	3895.8 ± 2.6	3887.1 ± 2.6
Width (MeV)	46 ± 10	30.23 ± 7.66	29.71 ± 7.53	28.4 ± 2.6



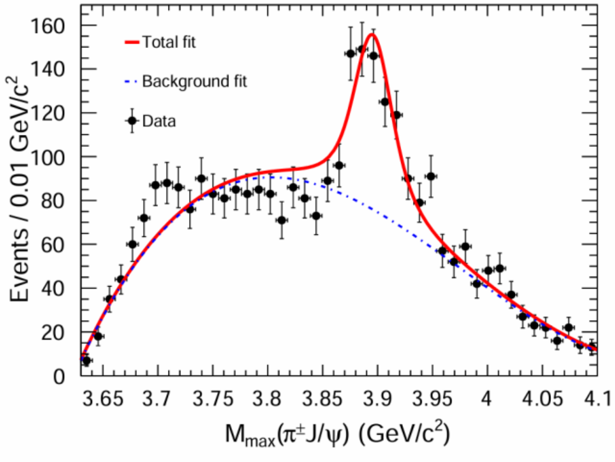
QSVM-predicted $Z_c(3900)$ distribution compared with traditional method.



SVM-predicted $Z_c(3900)$ distribution compared with traditional method.



QSVM fitting result.



SVM fitting result.

Summary and Outlook

Summary:

- The QSVM-based $Z_c(3900)$ analysis chain has been established for the first time at the BESIII experiment.
- The QSVM and SVM models have been trained and optimized to extract J/ψ signal events for the rediscovery of $Z_c(3900)$.
- The key parameters of $Z_c(3900)$, mass and width obtained from QSVM and SVM, are consistent with both traditional methods and the values reported by the PDG.
- This work further demonstrates the feasibility of quantum machine learning methods for high-energy physics data analysis.

Next to do:

- Subsequently, this analysis will be executed and validated on real quantum hardware to investigate the effects of realistic noise on the results.

山东大学
SHANDONG UNIVERSITY

Thank You!