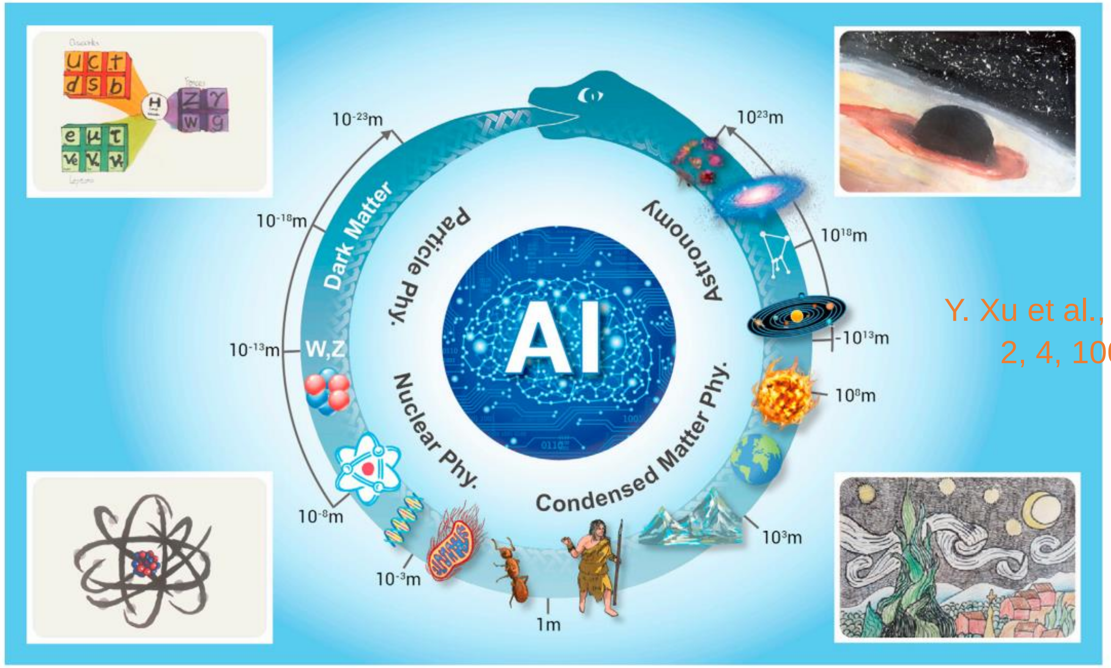
Machine Learning & Quantum Computing Applications at IHEP

DESY Seminar, February 13, 2023

Hideki Okawa

Institute of High Energy Physics, Chinese Academy of Sciences

AI & High Energy Physics



Y. Xu et al., The Innovation, 2, 4, 100179 (2021)

- Applications of AI has exploded in the past decade & beginning to provide a new paradigm in high energy physics as well as other basic sciences.
- Quantum computing may provide another 'leap' & is attracting global attention.
- Today, I will give an overview of recent AI & quantum computing applications at IHEP.

Institute of High Energy Physics (IHEP)

- Belongs to the Chinese Academy of Sciences (~114 institutes)
- **The largest research laboratory on basic science in China.**
- A world-class research center for high energy physics, astrophysics and multi-disciplinary research based on synchrotron radiation and spallation neutron source

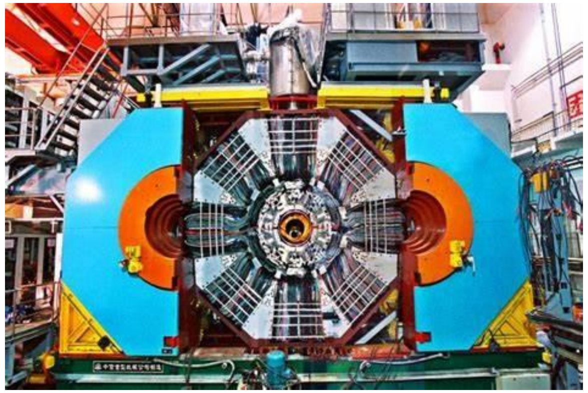
Employee ~ 1500
Student ~ 700
Postdoc ~ 170
Visitor : 300
Budget ~ 3B CNY (€0.4B) in 2021
Academician of CAS, CAE: 7
Large Science Facilities: 10
National & CAS Key Lab: 8

Research Fields

& Computing Technologies, Nuclear Technology and Applications

Particle Physics

- Beijing Spectrometer (BES) III
- Jiangmen Underground Neutrino Observatory (JUNO)
- International experiments (ATLAS, CMS, LHCb, PANDA, DarkSide, Belle-II, etc.)
- Theoretical studies



Accelerator

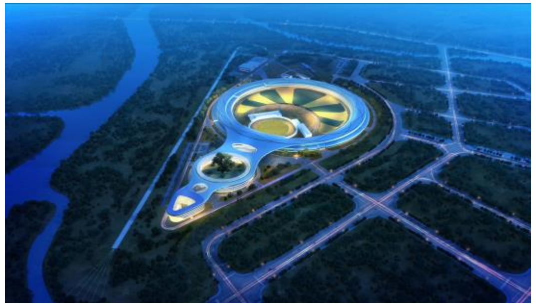
- Beijing Electron Positron Collider (BEPC) II
- China Spallation Neutron Source (CSNS)
- Accelerator-driven Sub-critical System (ADS)

Particle Astrophysics

- Hard X-ray Modulation Telescope (HXMT)
- Large High Altitude Air Shower Observatory (LHAASO)
- etc.

Multi-disciplinary

- High Energy Photon Source (HEPS)
- Beijing Synchrotron Radiation Facility (BSRF)
- Medical studies on nanomaterials, R&D of tumor drugs



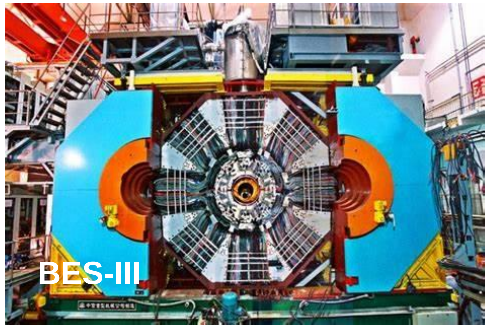
HEPS: Brightest photon source in the world (expected to operate from 2024)

IHEP Facilities

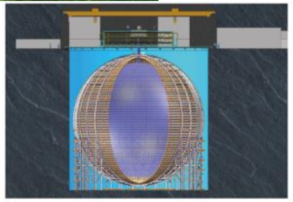
Huairou Campus
(Beijing)

Ali CMB Polarization
Telescope
(Tibet)

Main Campus
(Beijing)



Jinan Campus (Shandong)
Focuses on technology transfer

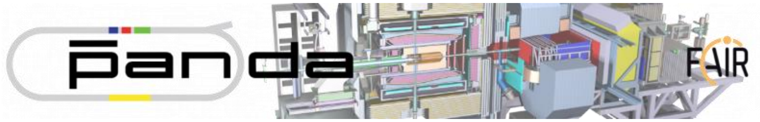


Dongguan Campus
(Guangdong)

LHAASO
(Sichuan)

International Cooperation

- Actively being involved in international collaborations & projects: ATLAS, CMS, LHCb, AMS, PANDA, Belle-II, COMET, EXO, Darkside, ILC, etc.
- Hosting international projects: BES-III, JUNO, LHAASO, CEPC, HERD, eXTP, etc.
- Developing long-term cooperative relationship with Europe, US and Asia.
- Actively recruiting foreign employees at all levels (faculty, engineers, postdocs)
  - Career workshop at CERN on December 2, 2022: https://indico.ihep.ac.cn/event/18414/



Principles of Mutually Beneficial Cooperation in Science – a Helmholtz-CAS Workshop organized by the WIKOOP-INFRA Project

Sep 26, 2022, 7:45 AM → Sep 27, 2022, 12:00 PM Europe/Berlin

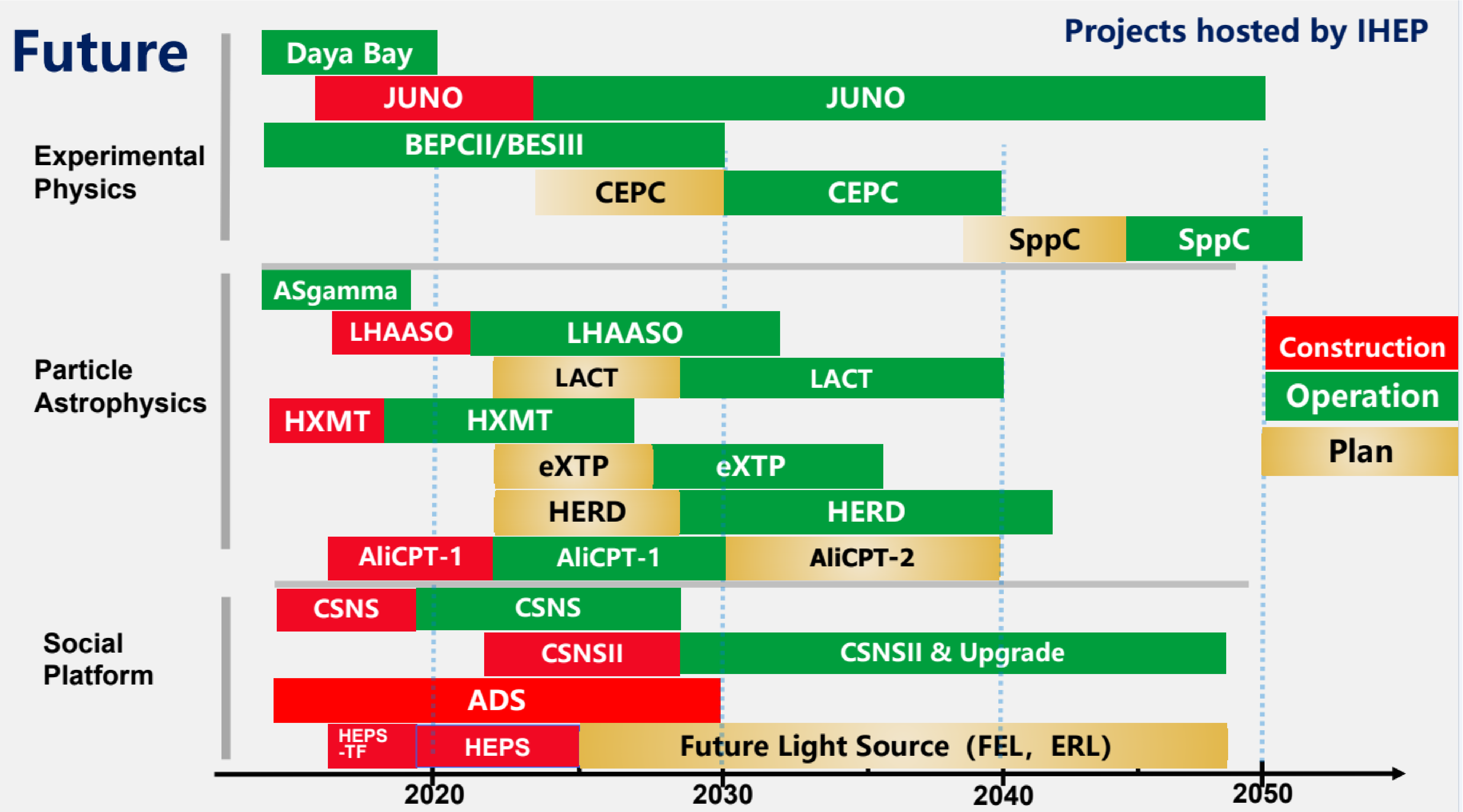
Online (Zoom)

Description The workshop serves as a platform for mutual information about the latest science-related developments at Helmholtz and CAS, with a focus on the field of Matter, and for discussing viable approaches for reaping the benefits of peaceful bilateral research cooperation.

Registration HELMHOLTZ-CAS WORKSHOP Registration — Apply for participation

Contact marcus.conle@desy.de
martin.sandhop@desy.de

International Employees
Faculty: 8
Postdocs: 22 (total of 86 since 2006)



From Liangjian Wen

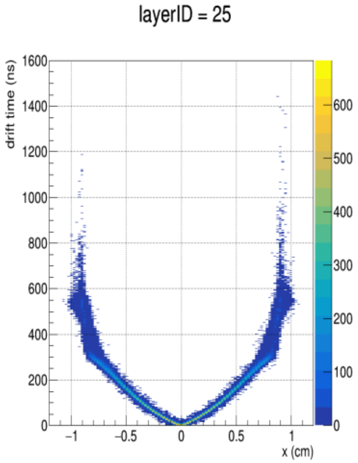
IHEP Computing Center

- Computing: 100k CPU cores, 300 GPU cards for more than 10 experiments
- Storage: 85 PB disk storage, 50 PB tape storage
- Network: LHCONE member, WAN Bandwidth: 40Gbps (LHCONE 20Gbps)
- At Beijing:
  - ATLAS, CMS Tier-2
  - LHCb will build Tier 1
- **Building platforms for AI & quantum computing**

ML & Quantum C. Projects at IHEP

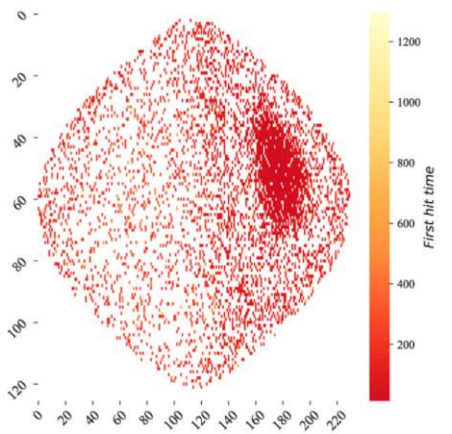
- Machine Learning Working Group (includes quantum computing) is formed at IHEP last Fall to promote discussions across experiments & departments: BES-III, LHC, JUNO, CEPC, LHAASO, HERD,…
- Aiming to establish **platforms** & develop innovative applications

Detector simulation



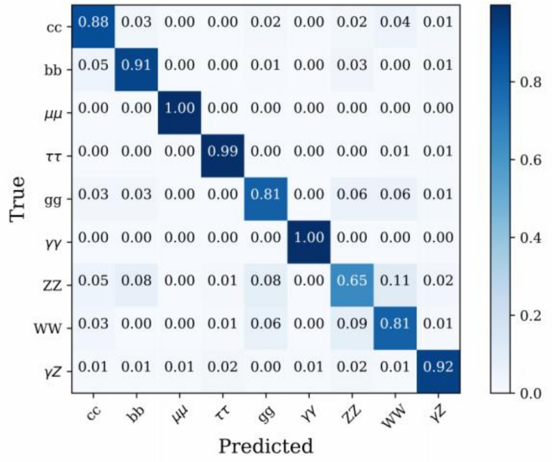
Using normalization flows for calorimeter & drift time simulation (BES-III)

Pattern recognition, reconstruction



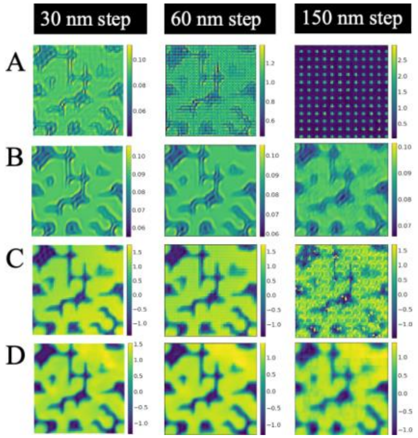
Vertex reconstruction using CNN, ResNet, GNN in JUNO

Object/event classification



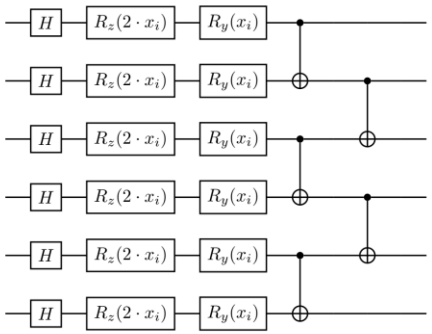
Classification of Higgs decay mode at CEPC

Inference, optimization



Speedy recovery of images (HEPS)

Quantum computing



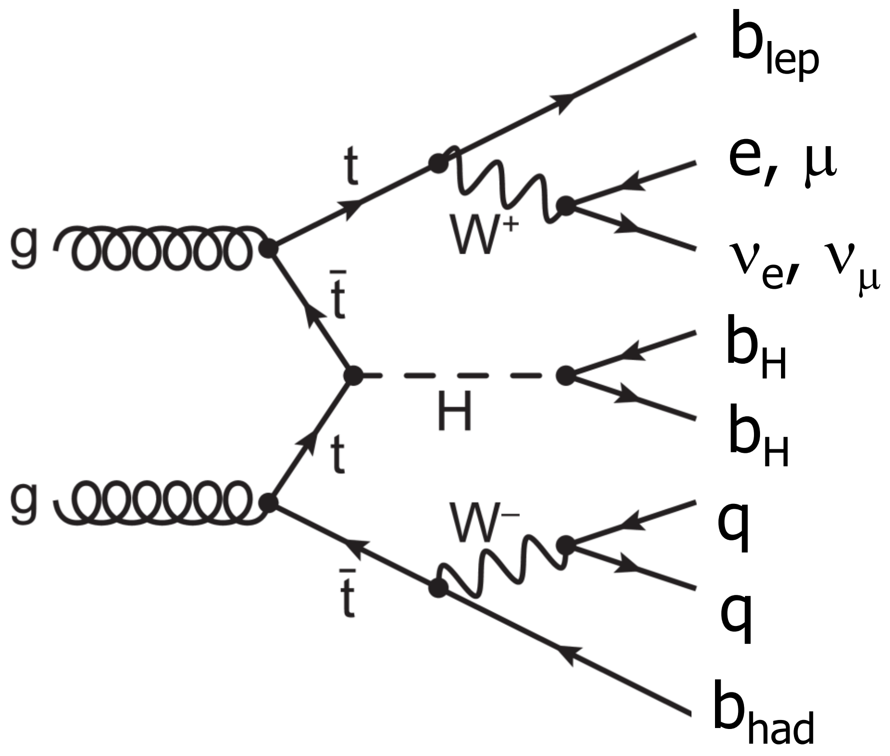
Quantum circuit of quantum feature map (CEPC)

Highlights of Machine Learning Studies from LHC

1. Symmetry Preserving Attention Networks
2. Pileup Mitigation using Attension-based Cloud Networks

Michael Fenton, Alexander Shmakov, **Hideki Okawa (IHEP)**, Yuji Li, Ko-Yang Hsiao, Shih-Chieh Hsu, Daniel Whiteson, Pierre Baldi

1. Symmetry Preserving Attention Networks



- Collaboration w/ UC Irvine & Washington
- Jet-parton assignment (i.e. top reconstruction) is a crucial component in $t\bar{t}$ & $t\bar{t}H$ analyses.
- Standard algorithms compare all possible permutations of jets per event & systematics
  - Combinatoric diverges with jet multiplicity.

Unsorted list of jets

$j_1, j_2, j_3, j_4, j_5, j_6, j_7, j_8$

⟷

Target partons

$b_{lep}, b_H, b_H, \varnothing, b_{had}, \varnothing, q, q$

Attention for Top Reconstruction

- Attention mechanisms are superceding RNNs & LSTMs in neuro linguistic programming.
  - Permutation invariant & can handle variable-length lists
- **Tensor Attention: generalization of attention to encode symmetries** ($t \leftrightarrow \bar{t}$, $b \leftrightarrow \bar{b}$ in H, $q \leftrightarrow \bar{q}'$ in W)

e.g. Two-body decay symmetries (W → $q\bar{q}'$, H → $b\bar{b}'$)

$$S^{ijk} = \frac{1}{2}\left(\theta^{ijk} + \theta^{jik}\right)$$

$$O^{ijk} = X^i_n X^j_m X^k_l S^{nml}.$$

$i \leftrightarrow j$ symmetry

- Allow us to test every possible permutation in a single pass.

Symmetry Preserving Attention Networks (SPA-Net)

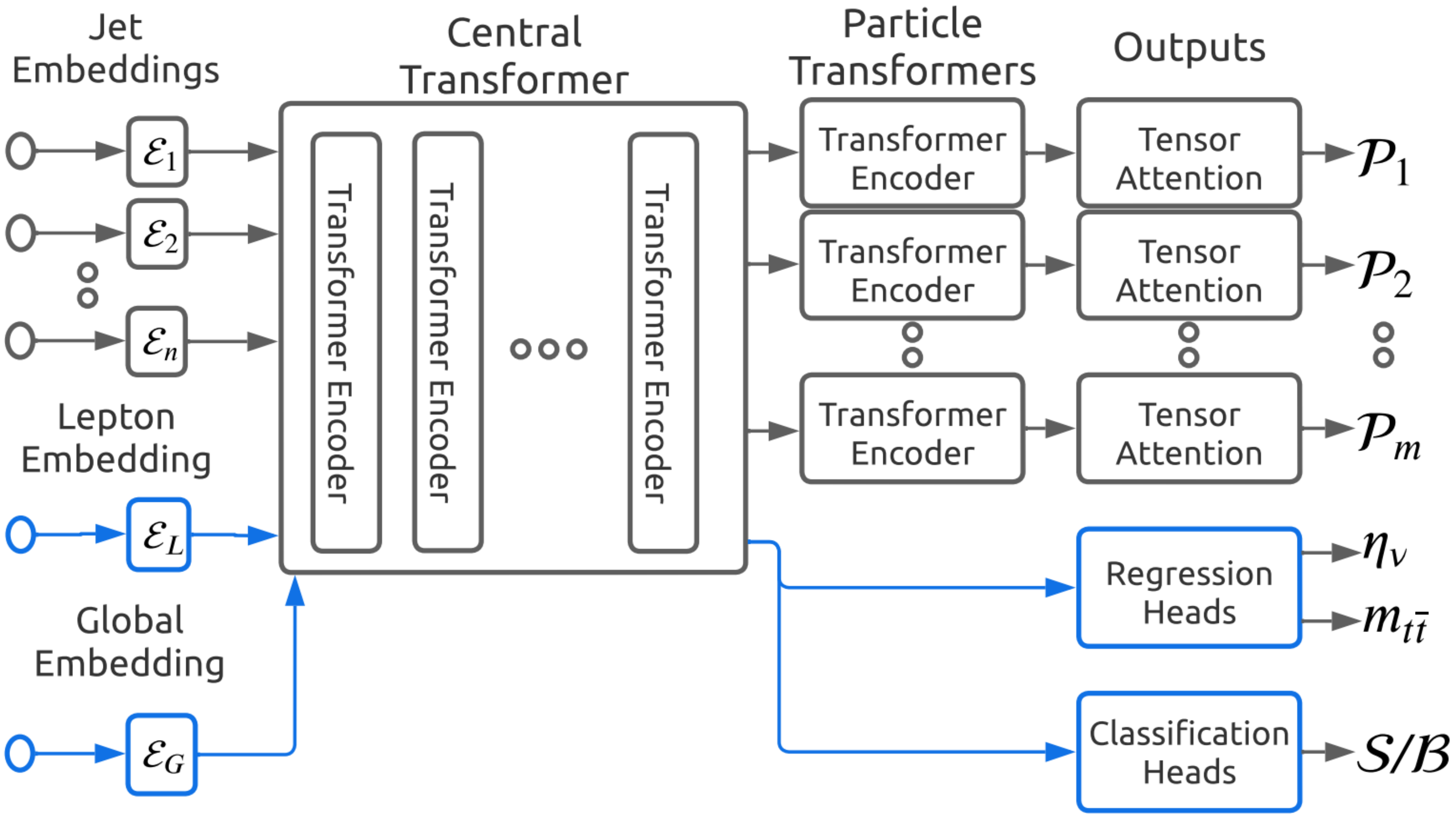
New Version!

Event-level context-aware encoding

Particle-level encoding

Symmetric jet matching

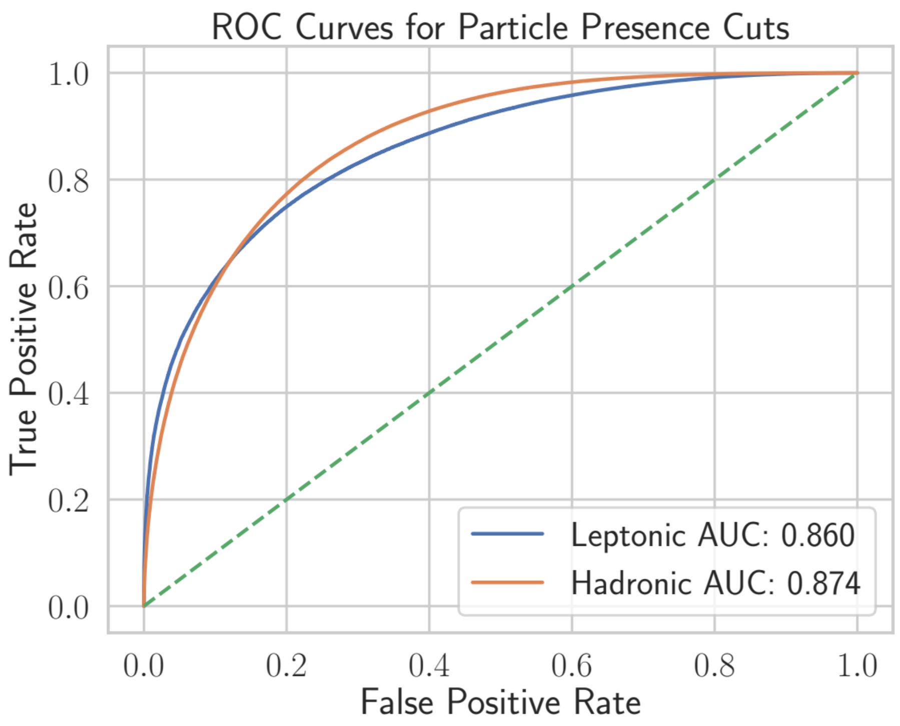
Unordered list of object four-momenta + additional info (e.g. btag) & event-level variables (e.g. MET)



Full jet/doublet/triplet assignment distributions for every particle targe (e.g. t, H)

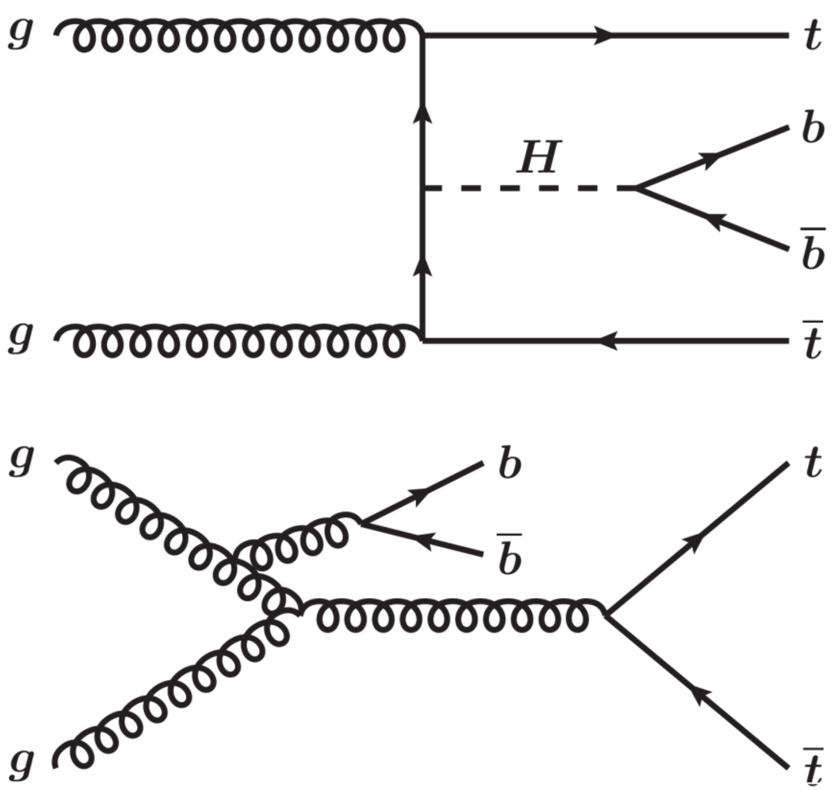
Reconstruction in semi-leptonic $t\bar{t}$ & $t\bar{t}H$

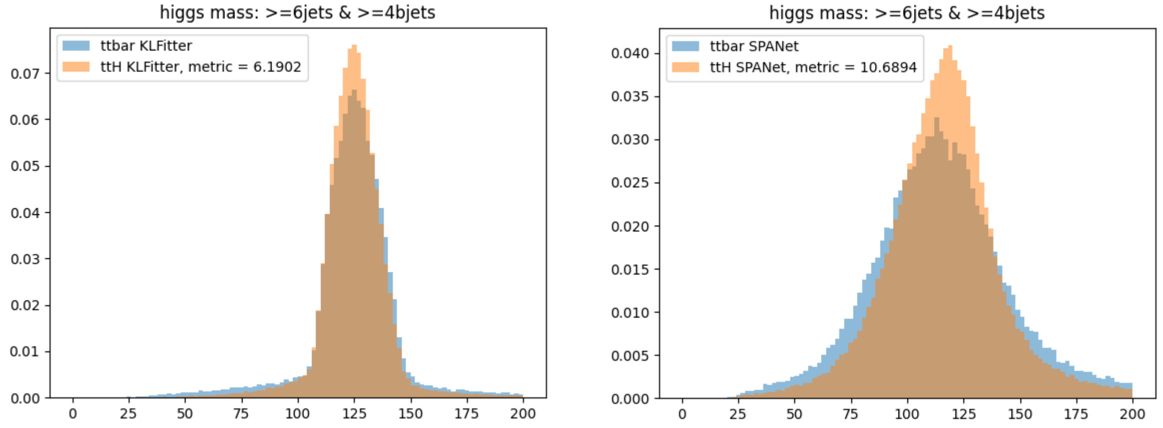
- New version of SPA-Net can handle different types of physics objects (jets, leptons, etc.)
- Can add event-level variables (MET, MET ϕ, etc.)
- $t\bar{t}$
  - SPA-Net: 75.6% full event reconstruction (85.5-59.8% vs NJets)
  - Permutation DNN: 64.9% (80.3-48.8%)
  - KLFitter: 52.1% (77.2-23.7%)
- $t\bar{t}H(\rightarrow b\bar{b})$
  - SPA-Net: 54.2% in 6j, 42.6% in 7j
  - Permutation DNN: 48.8% in 6j, 36.4% in 7j
  - KLFitter: 31.4% in 6j, 17.7% in 7j



ROC from detection probability (explained later)
X: false ID in unreconstructable events
Y: Eff. in reconstructable events

$t\bar{t}$H($\rightarrow b\bar{b}$) semi-leptonic



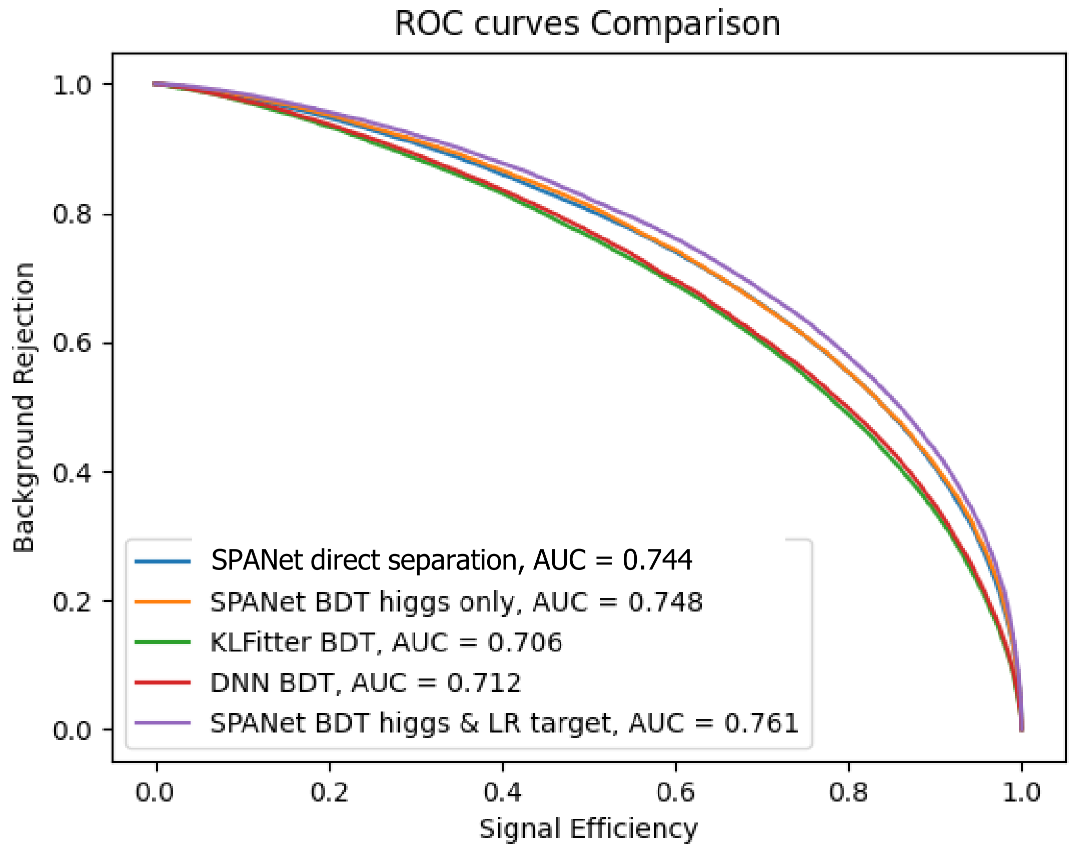


- $t\bar{t}b\bar{b}$ background is rather large with very similar kinematics to $t\bar{t}$H.
- t & H kinematics are main inputs to the BDT.
- However, the fraction of reconstructable events is only 35% in $t\bar{t}$H semi-lep. events.

- **"Goodness" of the jet-parton assignment is also important to remove unreconstructable events.** → i.e. likelihood for KLFitter, a score for permutation DNN

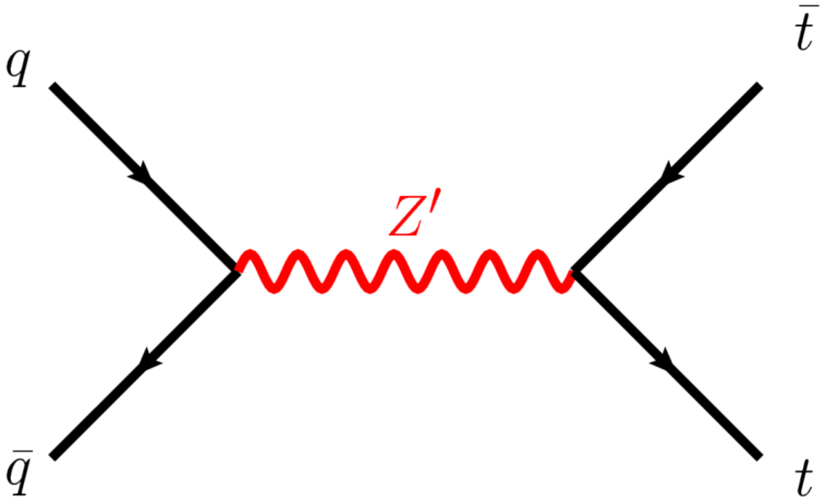
New Features: Signal/BG Discrimination

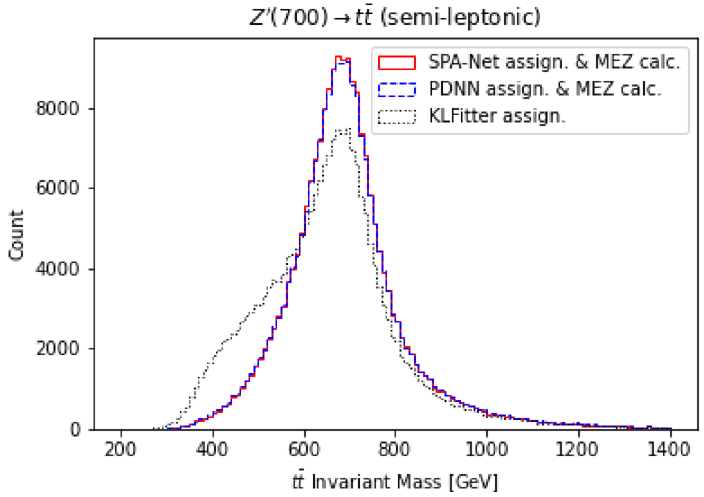
- **SPA-Net for jet-assignment & with all 9 probabilities included in the BDT gives the best results.**
- A new feature in SPA-Net can provide ttH/ttbb discrimination directly. → ~same performance (0.744 vs 0.748) as the full BDT w/ kinematics & Higgs probabilities.
  - **This indicates that SPA-Net is learning the kinematics as well!**

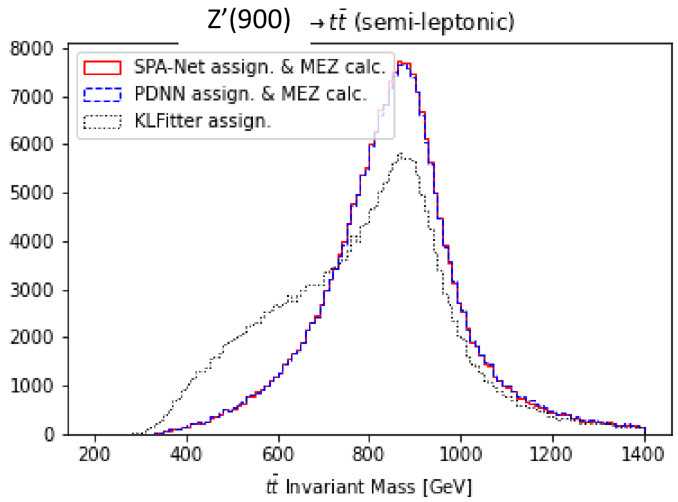


Transformer architecture provides us with meaningful embeddings for every jet, particle, and event: a big benefit over permutation-based models

Z'→$t\bar{t}$ Searches

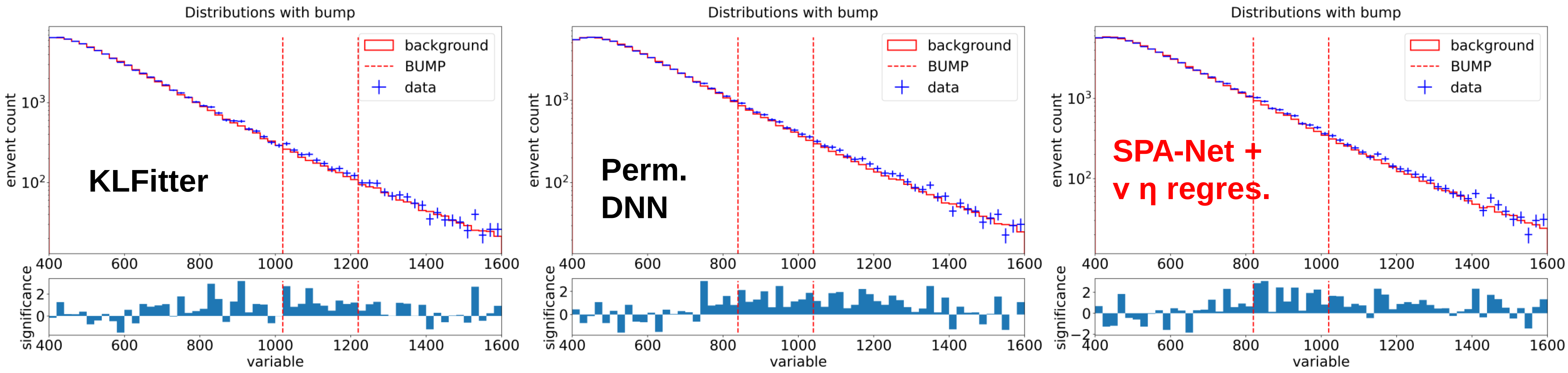






- Successful reconstruction of top quarks is crucial for the $t\bar{t}$ resonance searches (e.g. Z', RS Graviton, Heavy Higgs A/H→$t\bar{t}$).
- More jets are radiated for higher mass resonance. Significantly improved mass reconstruction from KLFitter.

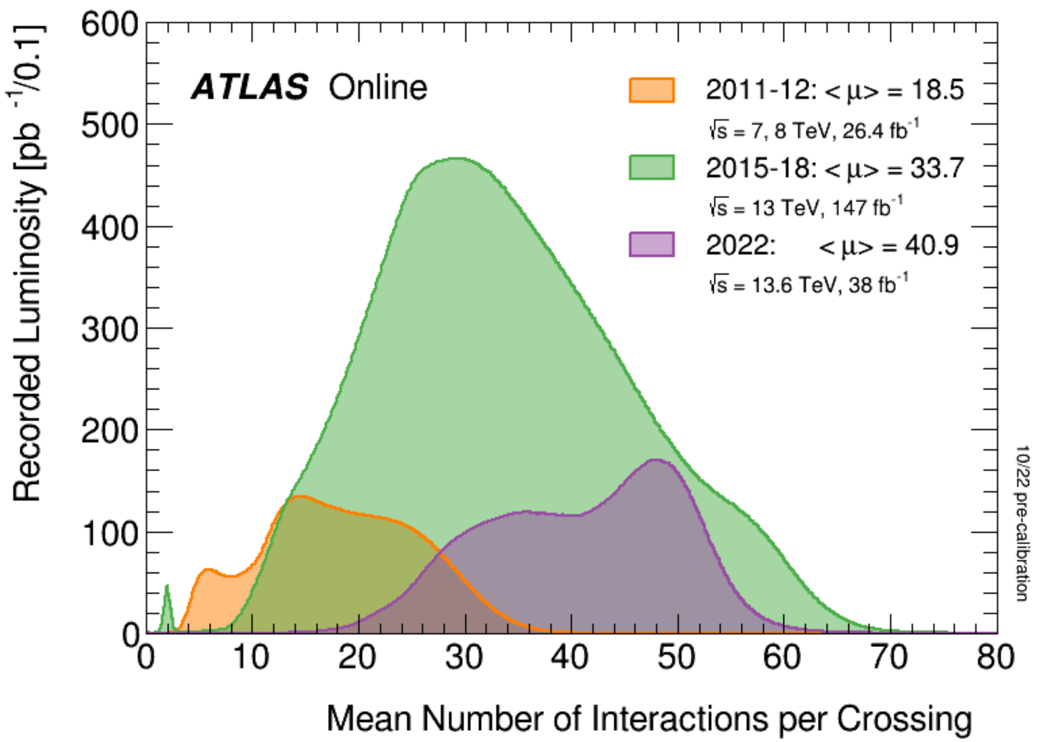
WIP: Impact on Z'→$t\bar{t}$ Searches

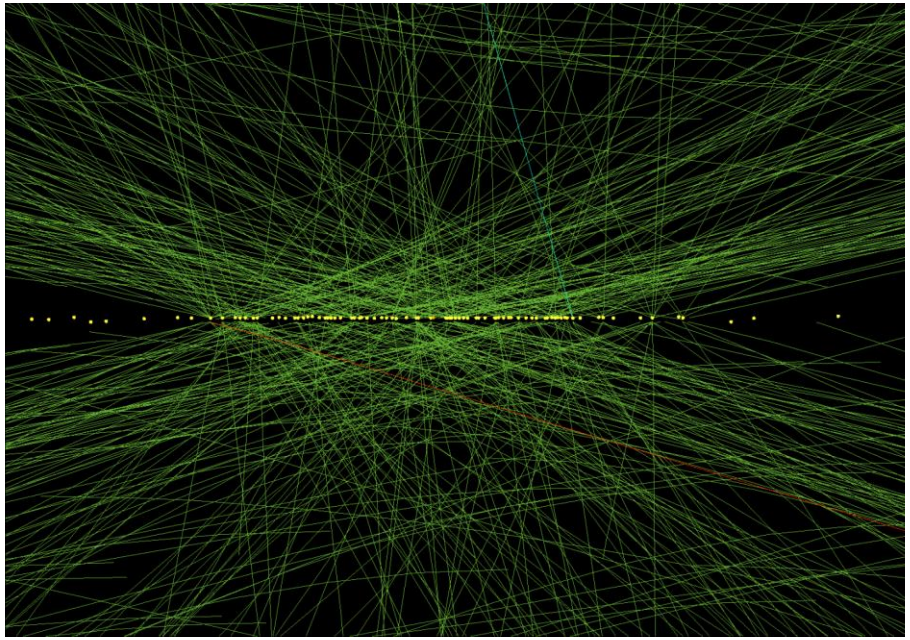


- Visible improvement in signal significance: e.g. global significance = **2.96σ (KLFitter)**, 3.19σ (p.DNN), **3.87σ (SPA-Net), 4.41σ (SPA-Net+ν η regression)** for m(Z')=900 GeV [integ. lumi.=3 fb^{-1} of pseudo-data assumed here]
- Removing unreconstructable events with SPA-Net probabilities should further enhance the sensitivity. (studies under way)

Thanks to Louis Vaslin for feedback on pyBumpHunter

2. Pileup Mitigation at the LHC

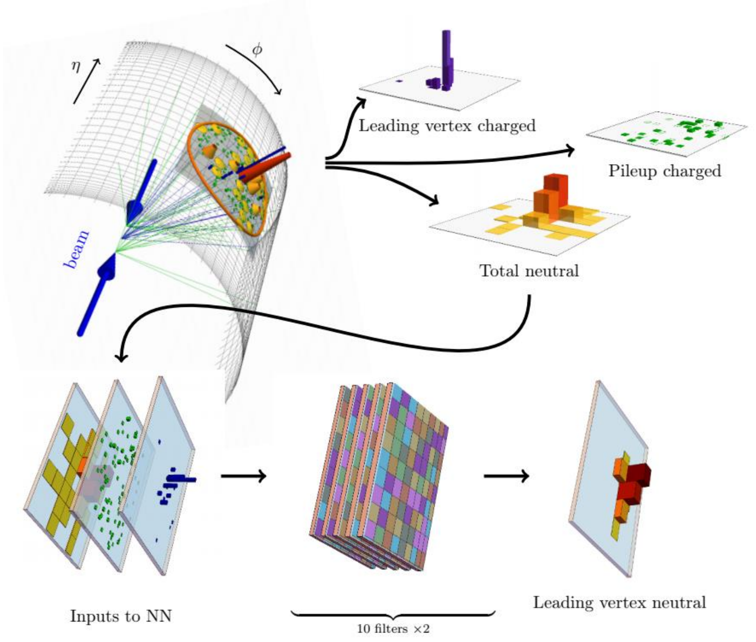




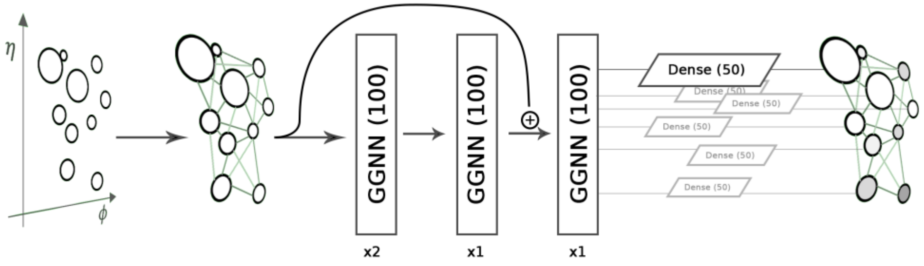
- Pileup (additional pp collisions) has been continuously increasing; Run-3 $\langle\mu\rangle$ will be ~50, will reach $\langle\mu\rangle$=140-200 at the HL-LHC. Mitigation of pileup effects is crucial at the LHC.

Existing Methods

- **PileUp Per Particle Identification (PUPPI):** Non-ML method in CMS
- **PileUp Mitigation with Machine Learning (PUMML):** CNN with squared images. Overlooks the actual detector geometry.
- **PUPPI with Machine Learning (PUPPIML):** Full supervised GNN with truth labels from Delphes. Not feasible in full Geant-4 simulation.
- **Semi-supervised PUPPI:** Semi-supervised GNN. Training done on charged but not neutrals. Requires charged→neutral & central→forward extrapolations. T. Li et al., arXiv:2203.15823



P. Komiske et al., JHEP12(2017)051

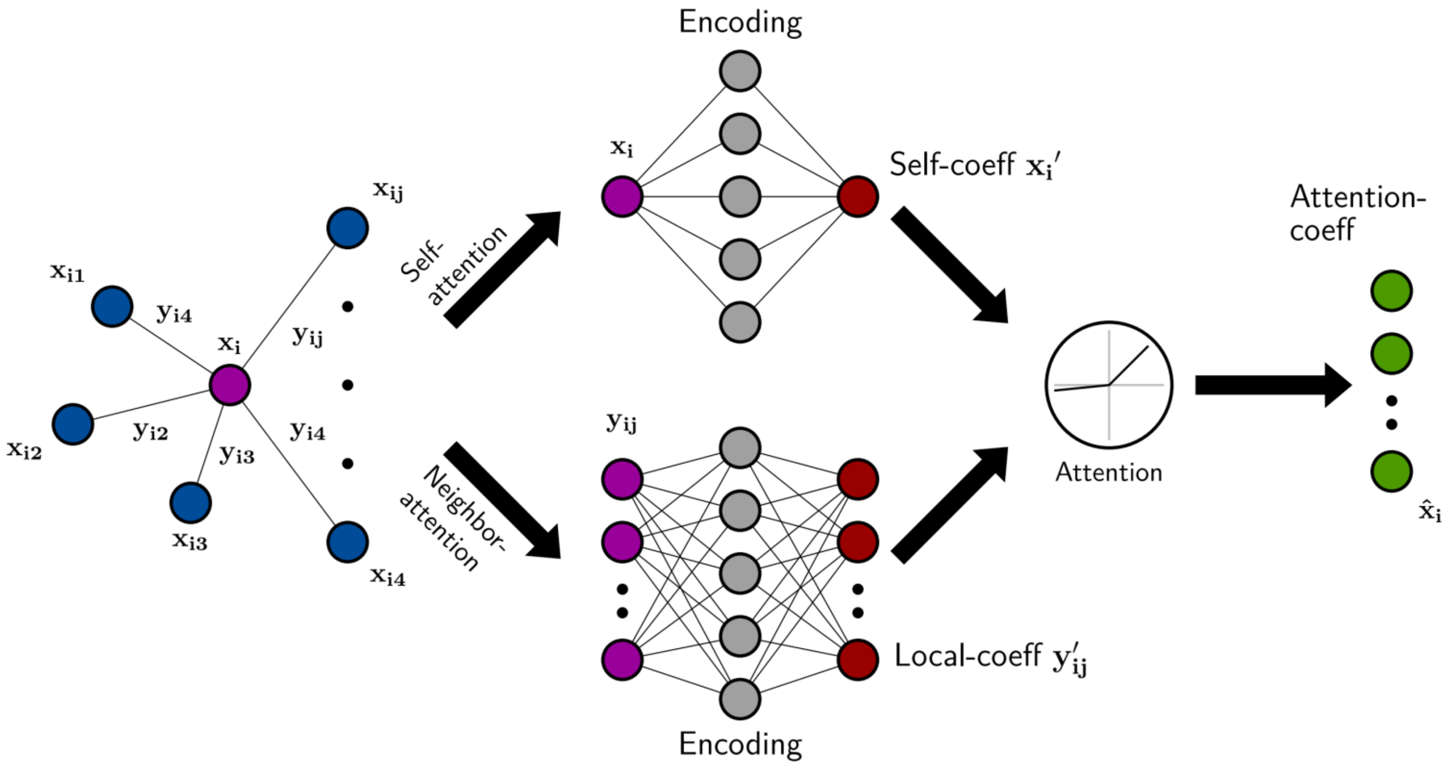


J. Martinez et al., arXiv:1810.07988

Lukas Gouskos , **Fabio Iemmi (IHEP)**, Sascha Liechti,
Benedikt Maier, Vinicius Mikuni , Huilin Qu, arXiv:2211.02029

Attention-Based Cloud Network

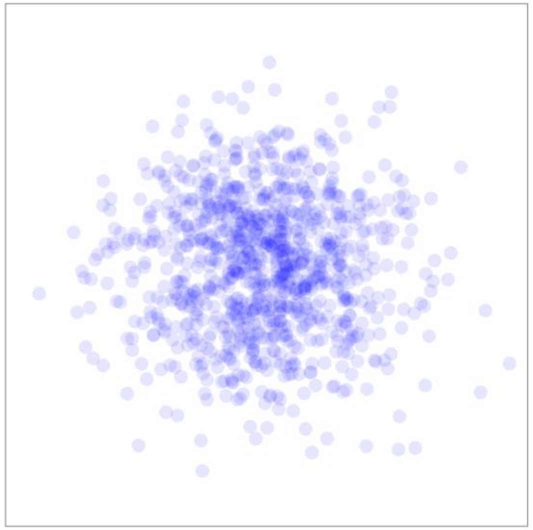
- Collaboration w/ CERN, U. Zurich, NERSC Berkeley.
- Attention-based cloud network (ABCNet) is a GNN enhanced with attention mechanisms

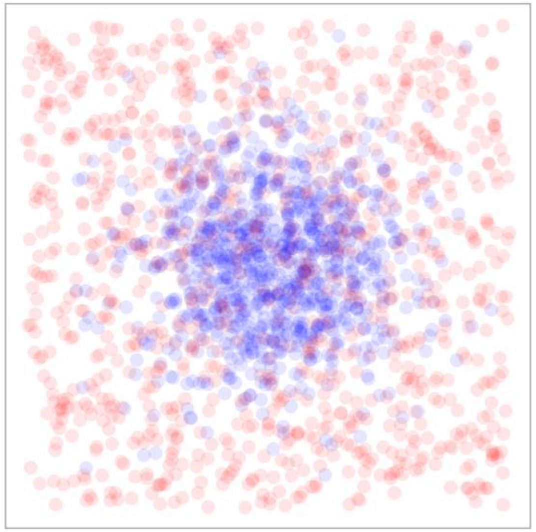


- Treat particle collision data as a set of permutation-invariant objects (p_T, η, φ, E, charge, particle type, dxy & dz impact parameters, PUPPI weight)
- Attention mechanisms filter out particles not relevant for the learning process

- Implemented inside custom graph attention pooling layers

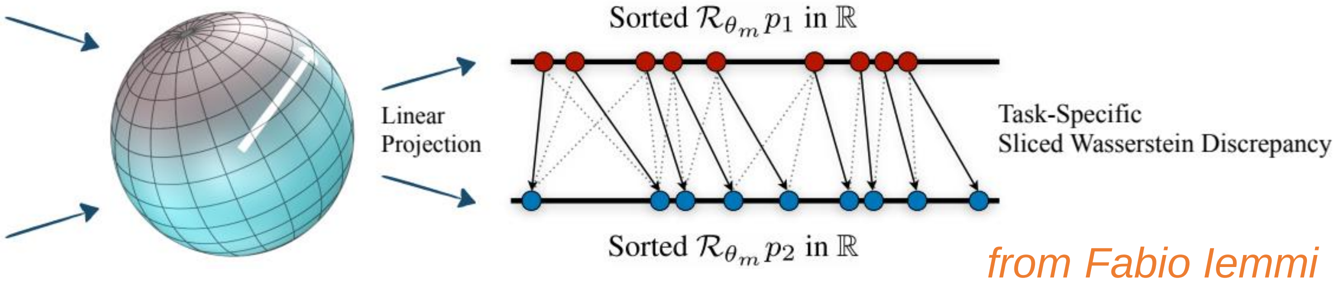
Novel Approaches in ABCNet





- Do not assign per-particle labels → rather assign a global label to samples
- Simulate identical events w/ & w/o PU.
- Learn the differences between the PU & non-PU samples

- Custom loss inspired by optimal transport.
- Match non-PU & PU particles weighted by ABCNet weights
- Used sliced Wasserstein distance → Take n-D feature space & project it to 1D

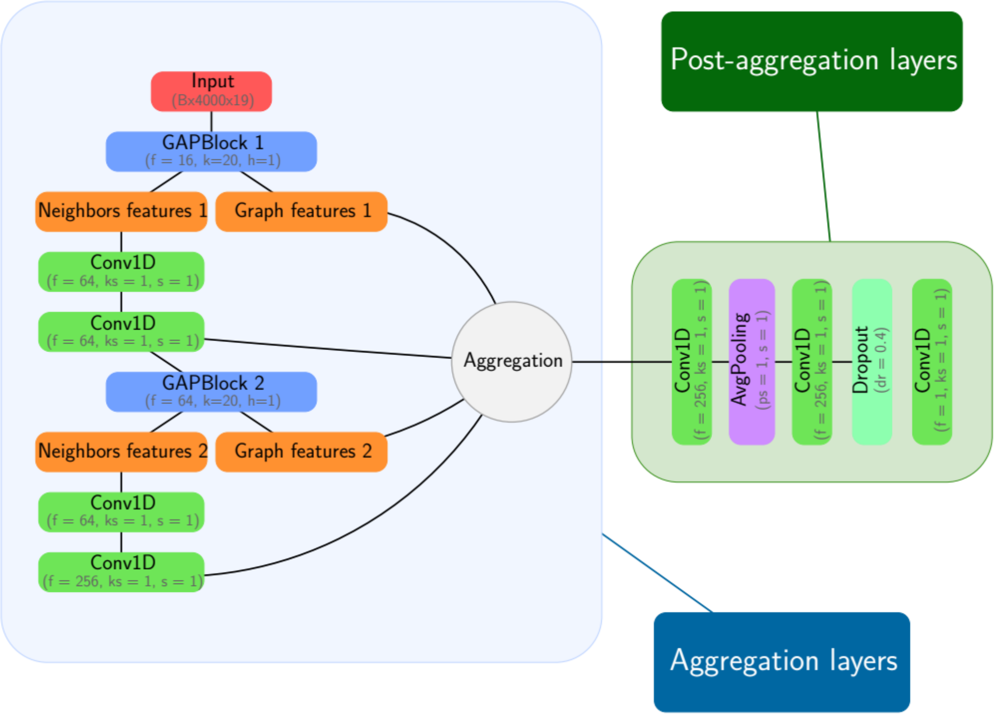


from Fabio Iemmi

$$\mathcal{OT} = \mathrm{SWD}(\vec{x}_p \cdot \vec{\omega}, \vec{x}_{np}) + \mathrm{MSE}\left(\mathrm{MET}(\vec{x}_p \cdot \vec{\omega}), \mathrm{MET}(\vec{x}_{np})\right)$$

Lukas Gouskos , **Fabio Iemmi (IHEP)**, Sascha Liechti, Benedikt Maier, Vinicius Mikuni , Huilin Qu, arXiv:2211.02029

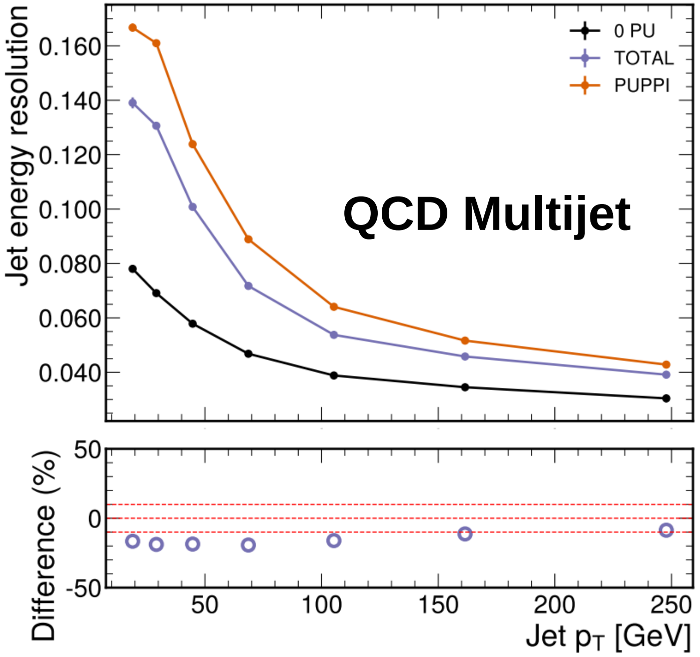
Attention-Based Cloud Network

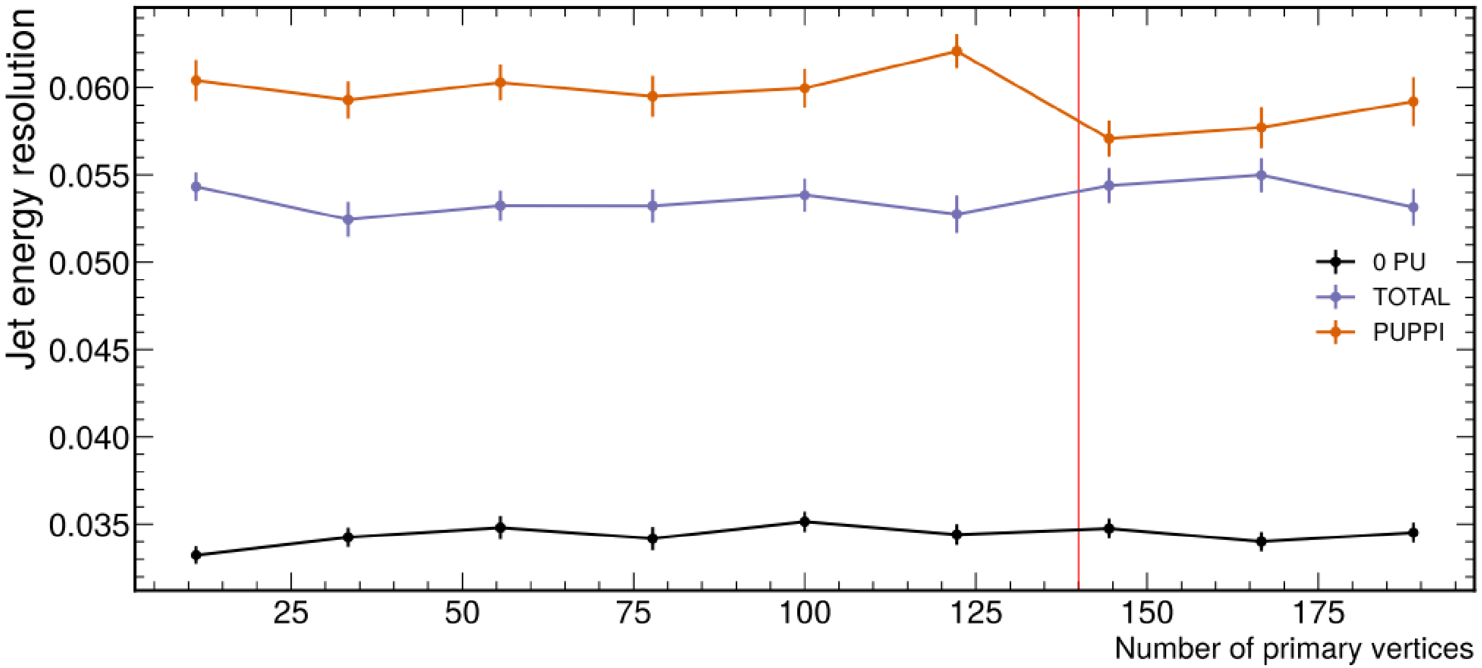


- Input features: p_T, η, φ, E, charge, particle type, dxy & dz impact parameters, PUPPI weight
- Loss: Sliced Wasserstein distance (previous slide)
- Cost function: squared distance
- Sliced features: p_T, η, φ, E
- Output: per particle weight

- Trained on 300k events, QCD multijet, ttbar dileptonic and invisible VBF Higgs
- Consider 9000 particles per event. Gather 20 k-nearest neighbors.

ABCNet Performance





- Improvement up to ~30-40% in JER in QCD & ttbar events, up to 25% in η resolution. ~10% improvement in τ_3/τ_2 resolution. MET improved by 9% from PUPPI.
- Performance stable against PU.

Status of Quantum Computing Studies

Current Activities at IHEP

1. Discussions with Institute of Physics (IOP), CAS to use their quantum computer/cloud platform (Quafu)
2. Establishing a platform for quantum simulation
3. Quantum Machine Learning studies
   a. Event classification at CEPC
   b. Particle Identification at BES III
   c. Quantum GAN for ATLAS calorimeter simulation (ongoing)

Cloud Platforms World Wide

	Superconducting			Trapped ion	Photonic	Annealing	Semiconducting Superconducting	Cloud Service Provider	
Vendor	IBM	rigetti	Google	IONQ	XANADU	D:WAVE	QuTech	amazon	
Cloud Name	IBM Quantum	Quantum Cloud Service(QCS)	Google Cloud	IonQ Quantum Cloud	Xanadu Quantum Cloud	Leap	Quantum Inspire	Braket	Azure Quantum
Latest Processor	Eagle	Aspen	Sycamore	Forte	Borealis	Advantage	Spin-2 Starmon-5	D-Wave IonQ、OQC Rigetti Xanadu	IonQ Quantinuum
Latest Qubit Count	127	80	54	32	216	5000+	2;5	QPU family	QPU family
Cloud Device + Pricing	27q	Aspen-11（40q）、M-1（80q）	No public access	IonQ device (11q)	Borealis	2000Q、Advantage	Non-commercial use	**task+shot**	11;20*
	$ 1.6 per runtime second	task+shot $ 0.3 per task $ 0.00035 per shot		①task+shot $ 0.3 per task $ 0.01 per shot ②QGS=N·C $ 0.00003 1q gate shot $ 0.0003 2q gate shot	①task+shot $ 0.3 per task $ 0.0002 per shot ② $ 100 per 1 million shots	①task+shot $ 0.3 per task $ 0.00019 per shot ② time×rate			**Different Formulas** Quantinuum： $HQC=5+C(N_{1q}+10N_{2q}+5N_m)/5000$

Material from Heng Fan (IOP, CAS)

Cloud Platforms in China

Institute or Company	Platform Name	# of Qubits	Fee
Center for Excellence in Quantum Information & Quantum Physics, Chinese Academy of Sciences (Jinan site)	Quantum Computing Cloud Platform	12	Free
Zhejiang University	Taiyuan	10	
Beijing Academy of Quantum Information Sciences (BAQIS)	**Quafu**	**45, 18, 10**	
Origin Quantum	OriginQ Cloud	6 (Wuyuan), (20 [not free])	
Baidu	Quantum Leaf	Internal	
Huawei	HiQ	Simulator only	
Wuhan Institute of Quantum Technology	Cold Atom Cloud (under preparation)	100+Atom (according to report)	

Material from Heng Fan (IOP, CAS)

1. Quafu Platform

User

User

User

User

User

http://quafu.baqis.ac.cn

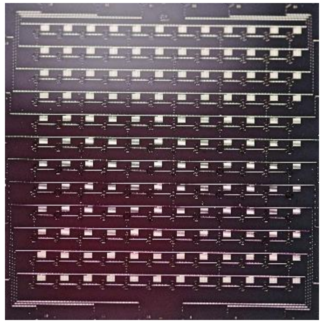
Quafu 量子计算云平台

北京量子信息科学研究院 Beijing Academy of Quantum Information Sciences

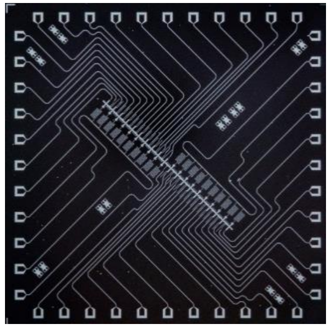
中國科學院物理研究所 Institute of Physics, Chinese Academy of Sciences | 北京凝聚态物理国家研究中心 BEIJING NATIONAL LABORATORY FOR CONDENSED MATTER PHYSICS

Pursue Quantum Computing, Prompt results

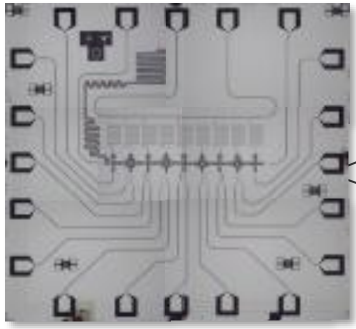
Adopts globally common quantum programming language; convenient & fast



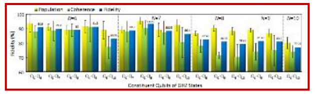
45 qubit chip



18 qubit chip



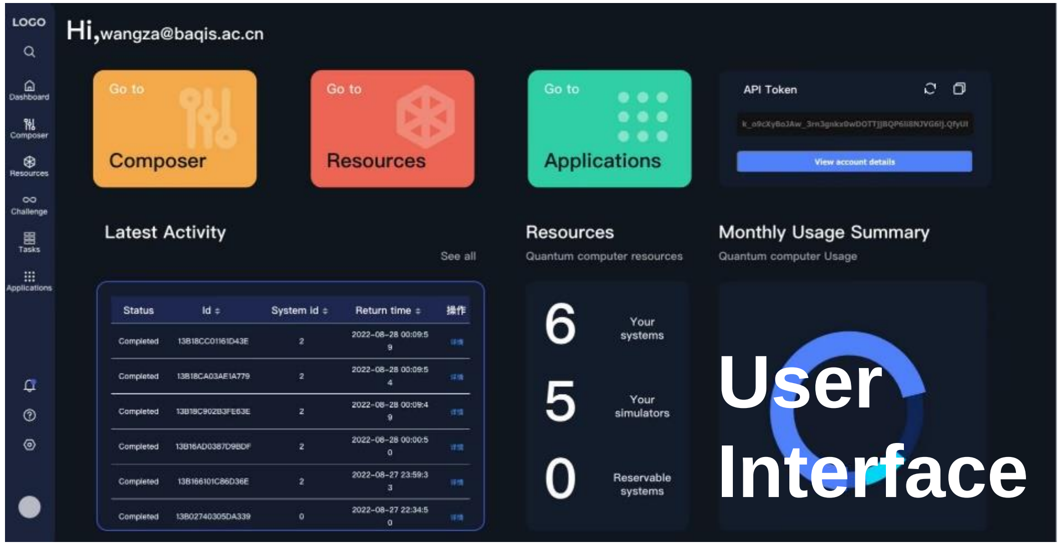
10 qubit chip



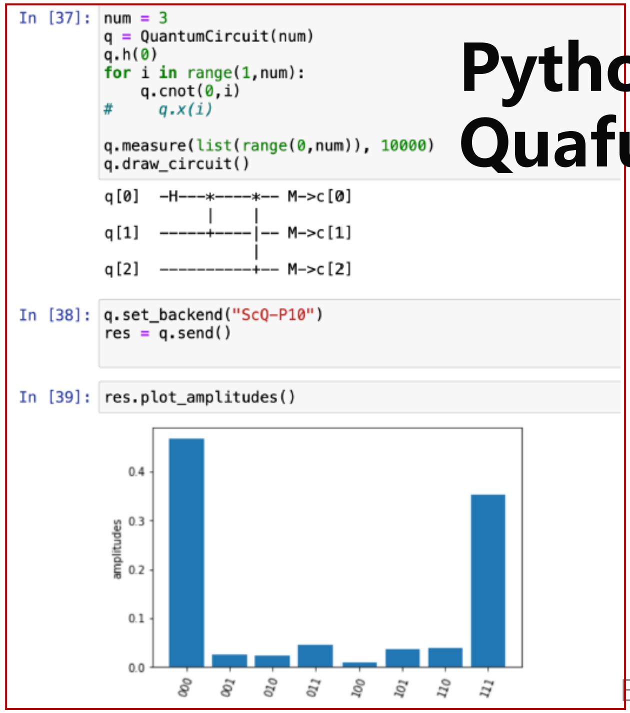
Already published results from cloud, C.-T. Chen et al., Science China-PMA (2022)

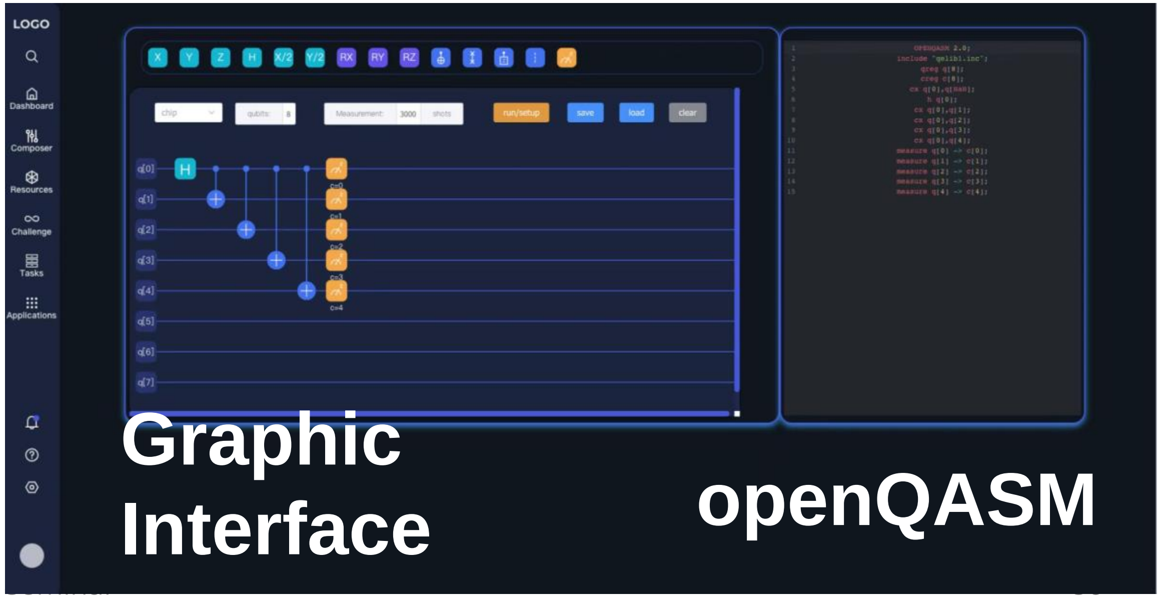
Material from Heng Fan (IOP, CAS)

Quafu Platform



Cloud platform publicly accessible & documentation fully available in English

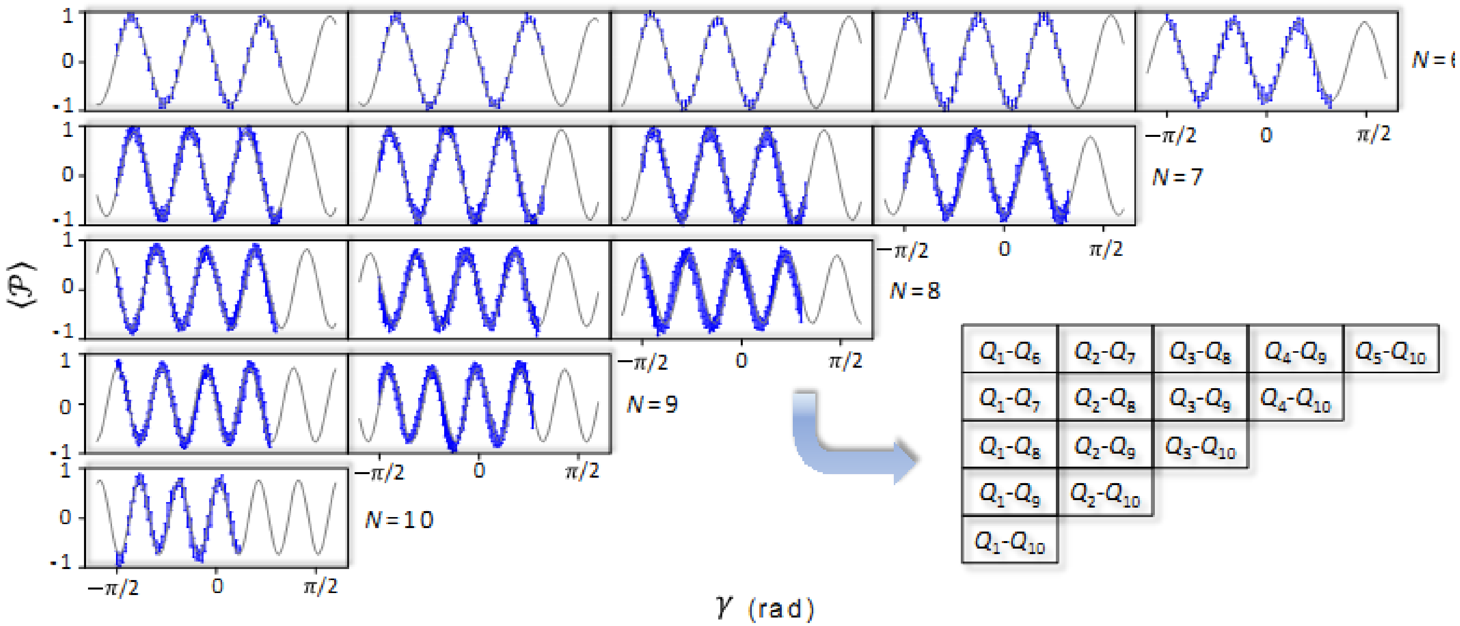


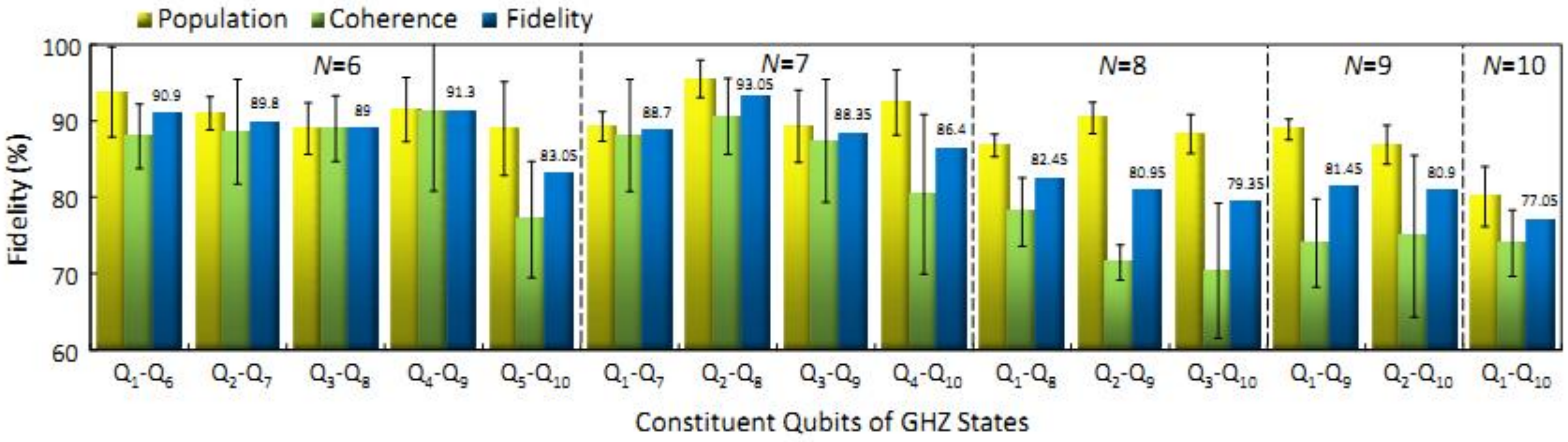


Material from Heng Fan (IOP, CAS)

Cloud Quantum Computation

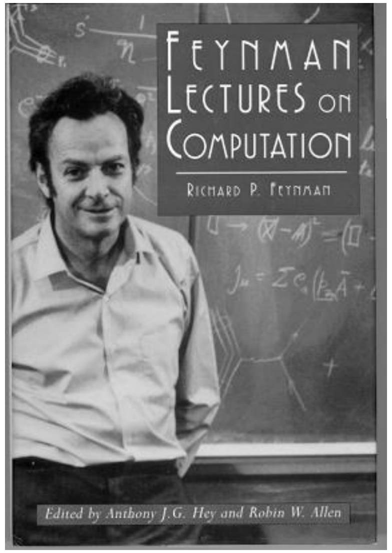
C.-T. Chen et al., Science China-PMA, 65, 110362 (2022)





- Results using the cloud platform
- Users can manipulate qubits graphically or with a QASM programming language.
- Greenberger-Horne-Zeilinger (GHZ) states of up to all 10 qubits in 1D array were generated to demonstrate the capabilities of the platform!

Quantum Simulation



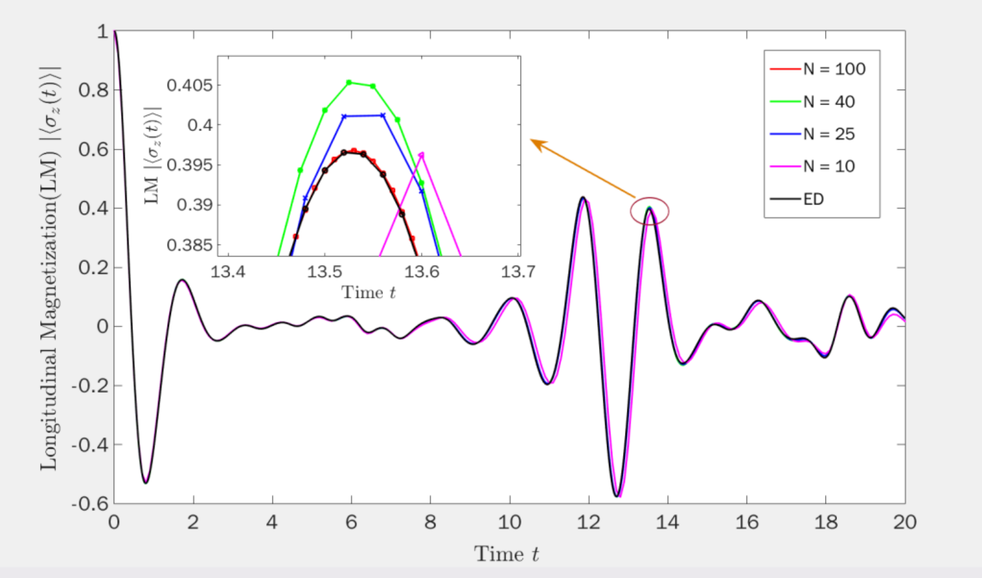
$$i\hbar \frac{d}{dt}|\phi\rangle = H|\varphi\rangle$$

$$|\varphi(t)\rangle = \exp\{-i\hbar Ht\}|\varphi(0)\rangle.$$

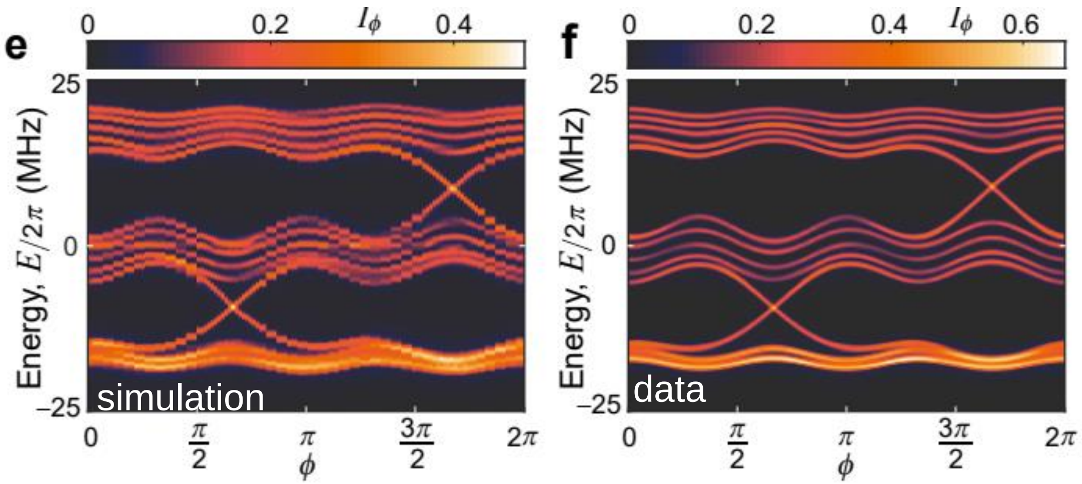
$$U = \exp\{-i\hbar Ht\}$$

$$U(\Delta t) \approx \prod_l \exp\{-i\hbar H_l \Delta t\}.$$

Trotter-Suzuki approximation

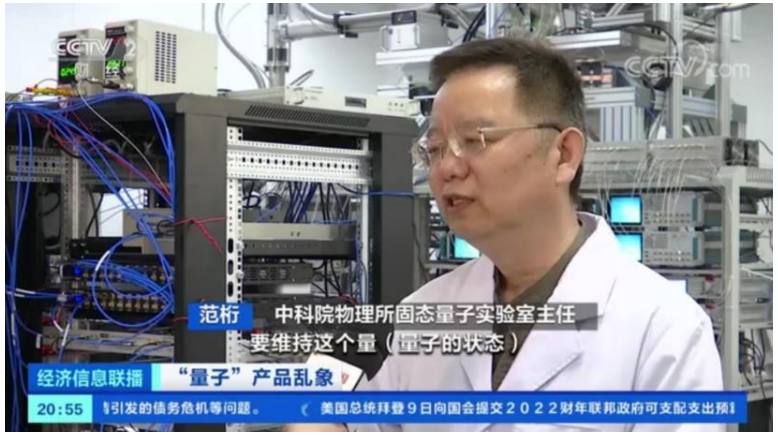


- Simulate the time evolution of Schrödinger equation (Feynman's motivation)
- Some examples: 1. simulation of dynamical phase transition, 2. simulation of band structure of quantum Hall system.



Z.-C. Xiang et al., arXiv:2207.11797

Quafu Platform in Beijing

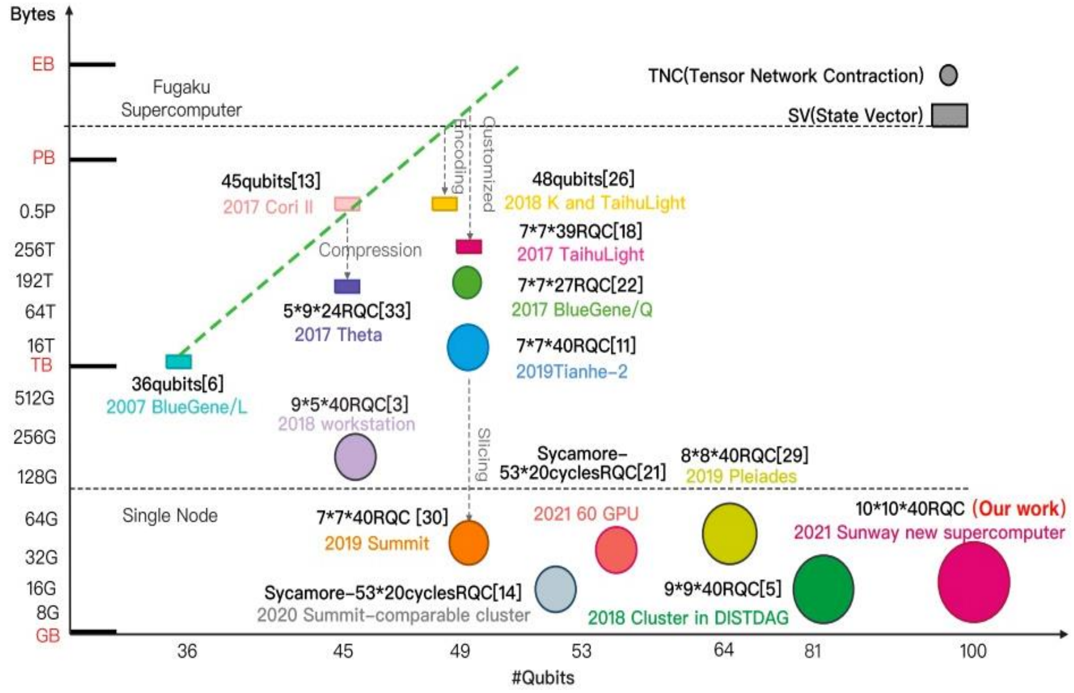


- Platform in Zhongguancun + Huairouyuan District.
- Widely advertised in media (CCTV, etc.)

Yujiang Bi, Wei Sun, Lu Wang (IHEP C.C.)

2. Quantum Simulator Platform

- Platform for quantum simulator is also highly crucial to **validate results from the actual quantum computers & develop quantum algorithms**.
  - Pros: can test perfect qubits; can simulate very deep quantum circuits
  - Cons: can usually simulate up to ~50~60 qubits even with HPC; modeling of noise & error not trivial.
- IHEP Computing Center is developing its platform using HPC CPUs/GPUs
  - Currently only accessible within IHEP; plan to make it public in the future
  - Also planning to collaborate with Quafu to access the real hardware
  - Targeting ≥38 qubit simulator.



Scientific Objectives

Material from Yujiang Bi

Quantum Computing Infrastructures

- User-friendly interactive developing platform
- Distributed heterogeneous computing platform
- Quantum simulating algorithms and implementation

Computing and Simulation Applications

- Possible quantum applications in HEP and QCD
- Common frameworks for general research
- Combining quantum computing in current workflow

Collaboration with industries & institutes

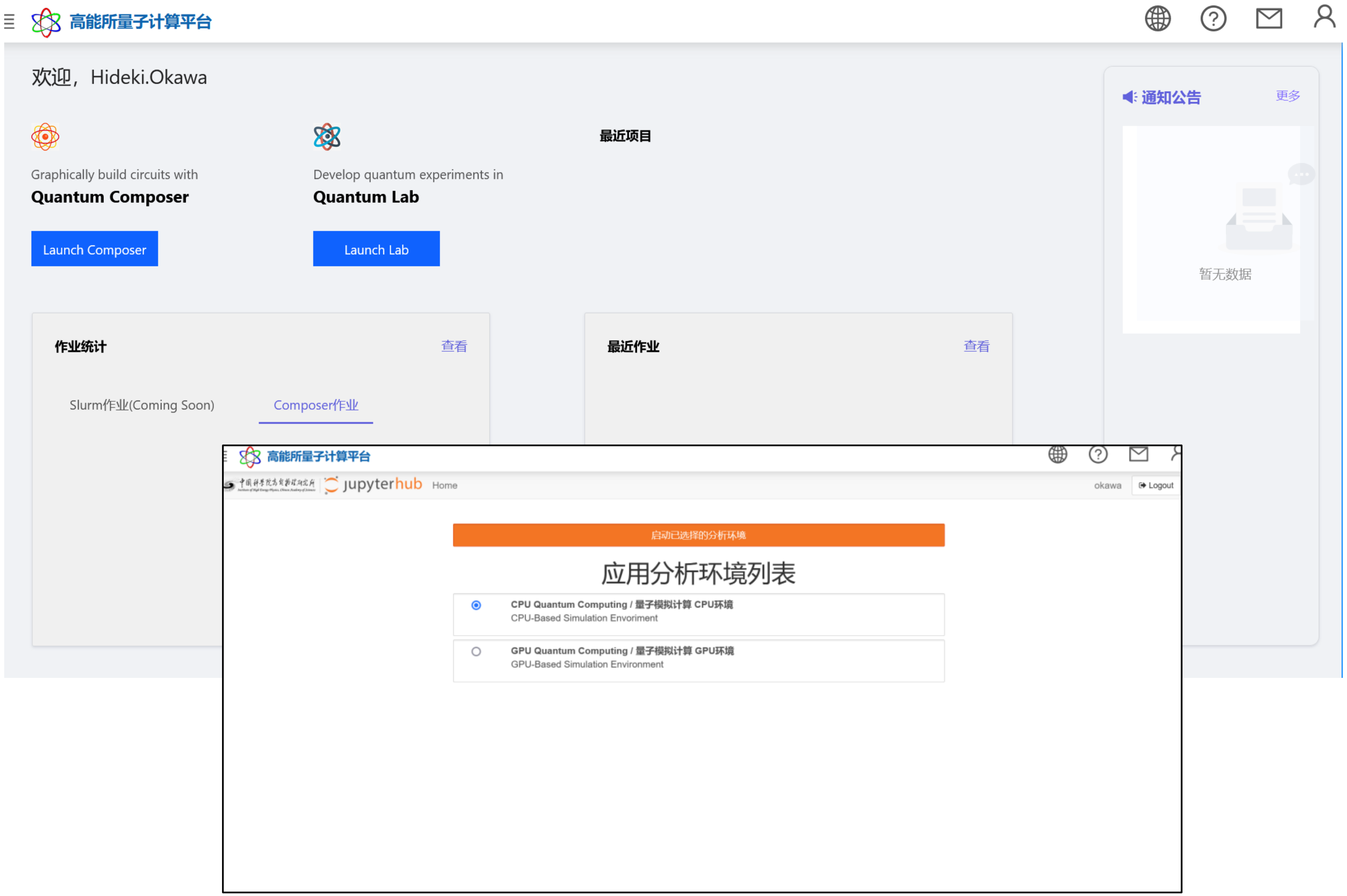
- Collaborating with manufacturers for compatibility
- Collaborative research on various subjects
- Active quantum development community

Education and Training

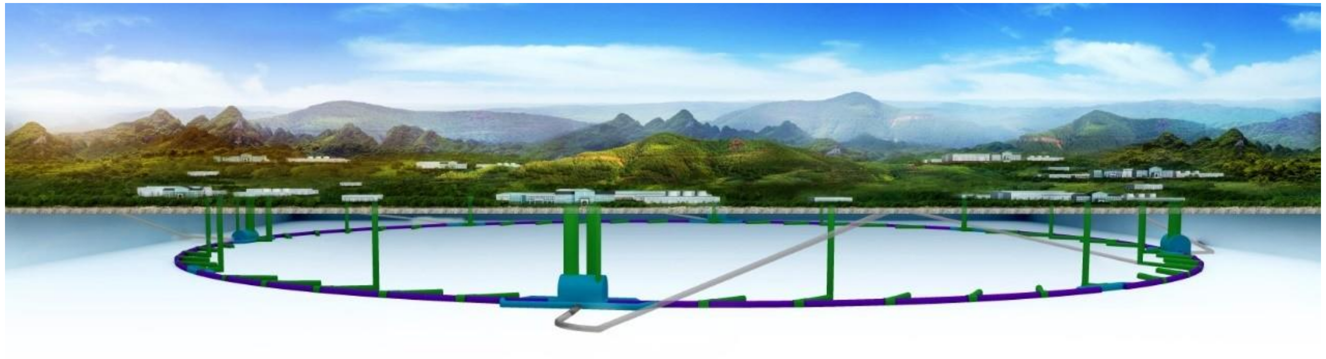
- Introducing quantum computing to the public
- Research introduction and progress presentation
- Providing conditions for learning and discussing

Heterogeneous Interactive Platform

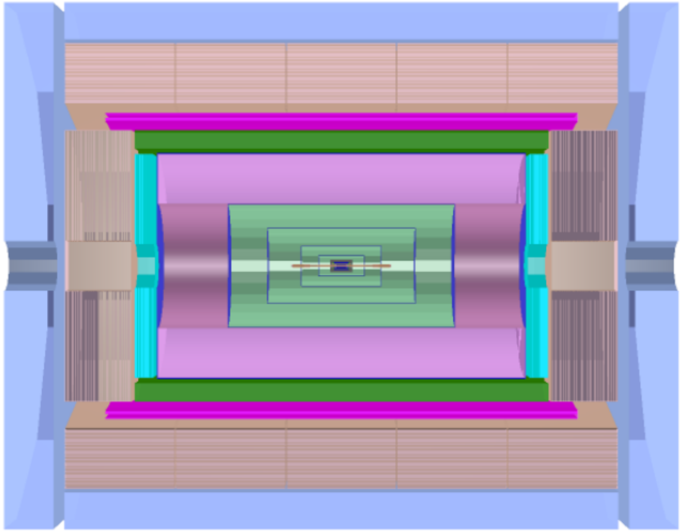
- Platform can access all services: storage, computing, scheduling, dashboard …
- Interactive interfaces
  - **Drag-and-drop** GUI like IBM quantum composer
  - **Jupyter** environment for interactive development
- Heterogeneous HPC clusters for heavy tasks
  - NVIDIA/AMD **GPU** clusters
  - Intel/AMD/**ARM** clusters
- Distributed computing platform
  - Unifying different centers from via global Slurm scheduler



2. CEPC & Quantum Computing



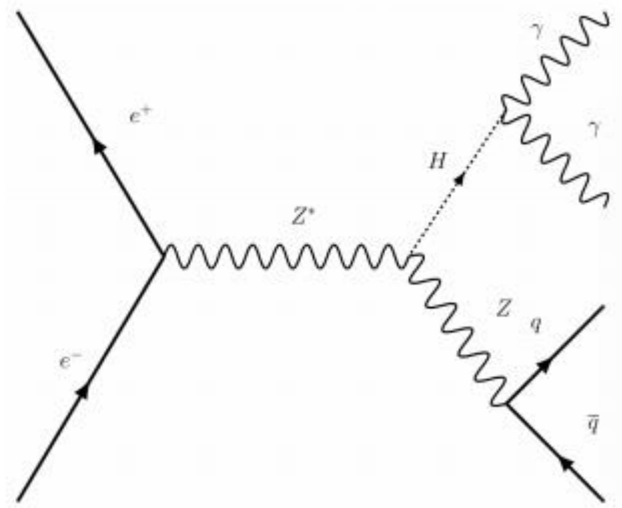
CEPC Operation mode		ZH	Z	W+W-	ttbar
		~ 240	~ 91.2	~ 160	~ 360
Run time [years]		7	2	1	-
CDR (30MW)	L / IP [$\times 10^{34}$ cm^{-2}s^{-1}]	3	32	10	-
	[ab^{-1}, 2 IPs]	5.6	16	2.6	-
	Event yields [2 IPs]	**1×10^6**	**7×10^{11}**	**2×10^7**	-
Run time [years]		10	2	1	5
Latest (50MW)	**L / IP [$\times 10^{34}$ cm^{-2}s^{-1}]**	8.3	192	27	0.83
	[ab^{-1}, 2 IPs]	20	96	7	1
	Event yields [2 IPs]	**4×10^6**	**4×10^{12}**	**5×10^7**	**5×10^5**

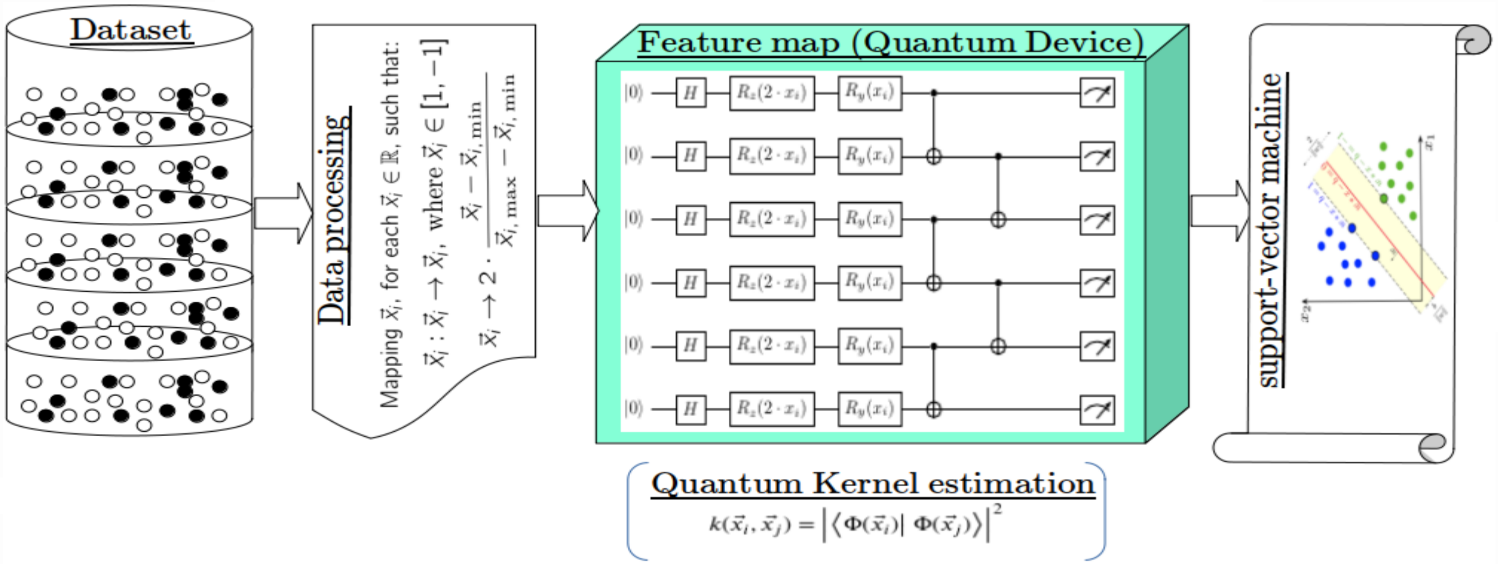


- **100km circumference e^+e^- collider; Z/Higgs/Top factory under consideration in China.**
- Cutting-edge detector and computing technologies will be essential to maximize the potential of the experiment.
  - Z pole measurements will face pileup due to the high luminosity. Innovations in tracking would help.

Abdualazem Fadola, Qiyu Sha, Yaquan Fang, Zhan Lia (IHEP),
Sitian Qian, Yuyang Xiao, Yu Zhang, Chen Zhou, arXiv:2209.12788

Event Classification in $e^+e^- \rightarrow ZH \rightarrow qq\gamma\gamma$

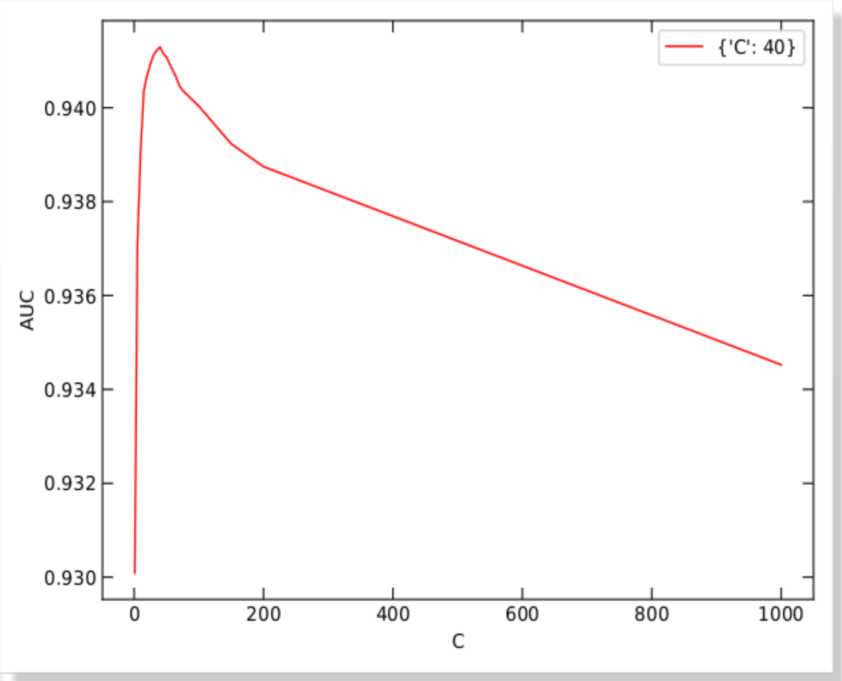


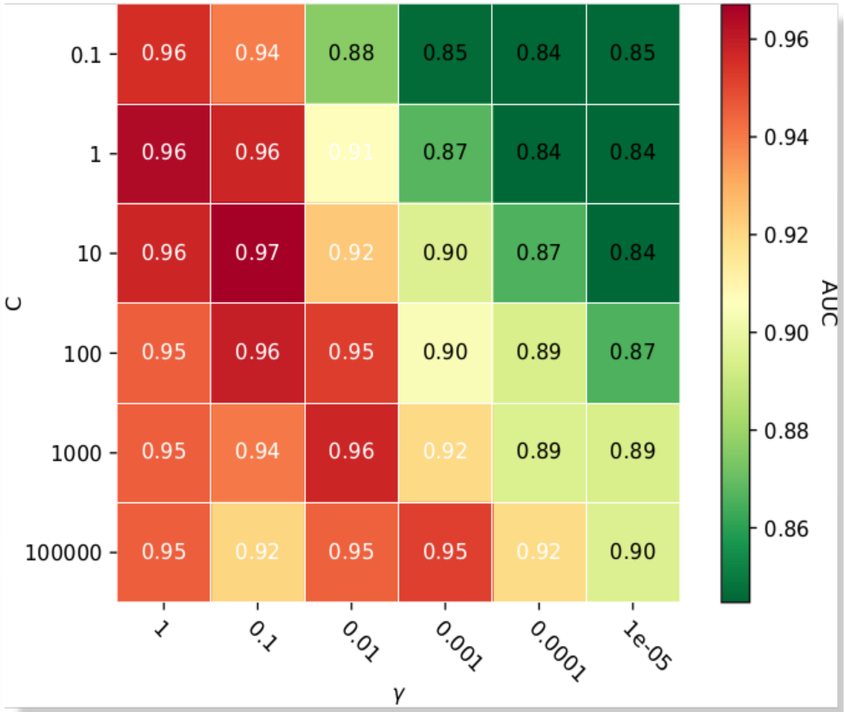


- 5.6 ab-1 at √s=240 GeV at CEPC → >1M Higgs will be produced
- Quantum SVM is considered as a benchmark to compared with the classical SVM.

- Kinematic features $\vec{x}$, event class y∈[-1,1] (-1: BG, 1: signal)
- QSVM maps $\vec{x}$ to a quantum state of N qubits (here N=6).

Hyper-parameter Tuning



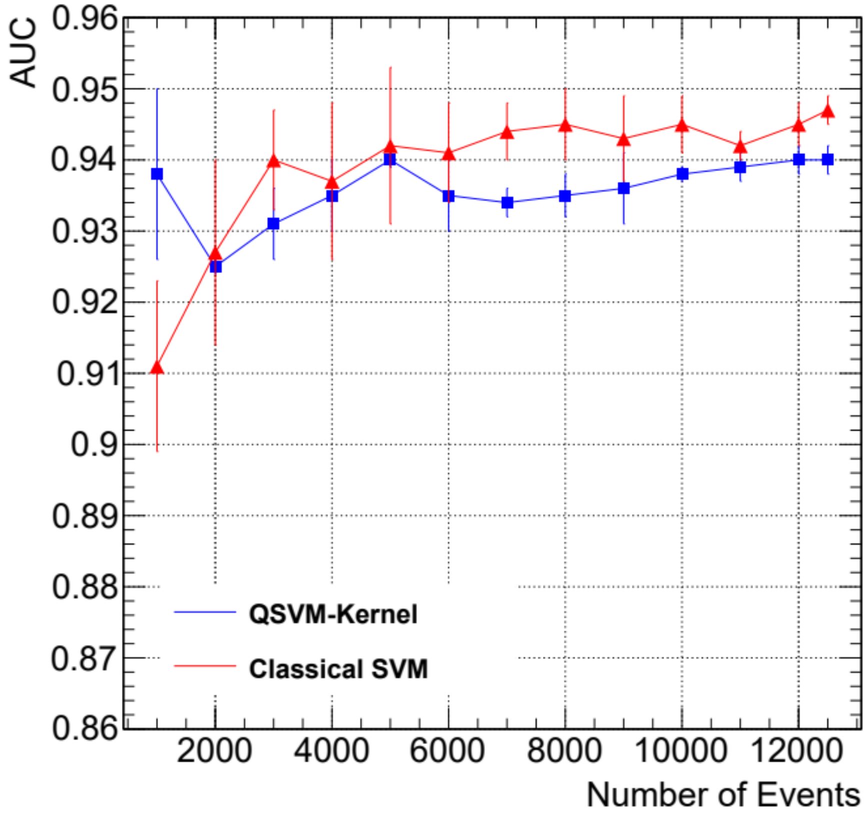


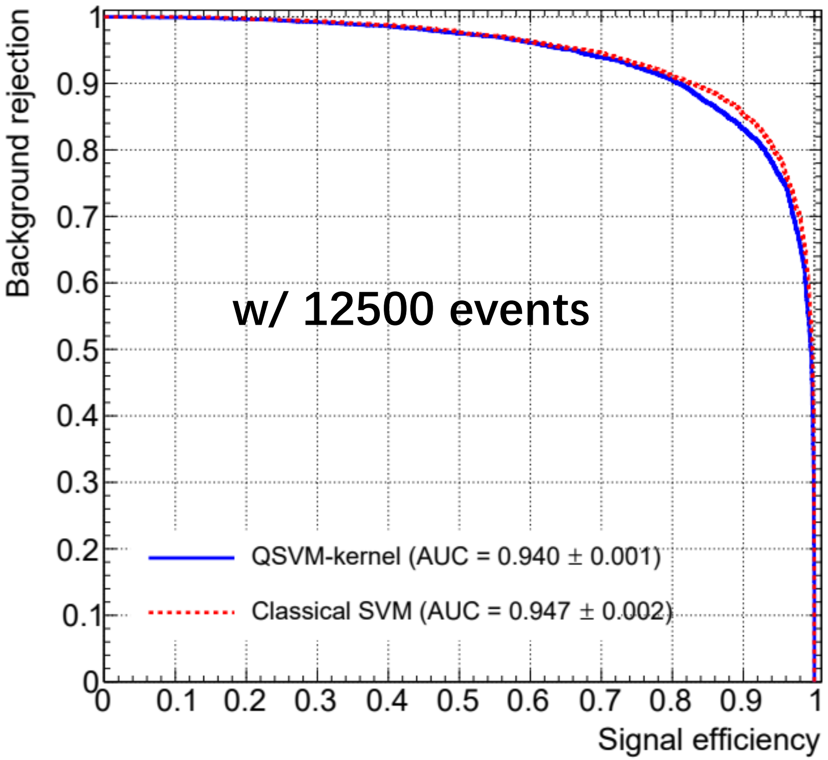
C \ γ	1	0.1	0.01	0.001	0.0001	1e-05
0.1	0.96	0.94	0.88	0.85	0.84	0.85
1	0.96	0.96	0.91	0.87	0.84	0.84
10	0.96	0.97	0.92	0.90	0.87	0.84
100	0.95	0.96	0.95	0.90	0.89	0.87
1000	0.95	0.94	0.96	0.92	0.89	0.89
100000	0.95	0.92	0.95	0.95	0.92	0.90

$$\left(\sum_{i=1}^{t} \gamma_i y_i k(\vec{x}_i, \vec{x}) + C\right) = 0$$

- Hyper-parameters γ and C are optimized to give the best separation

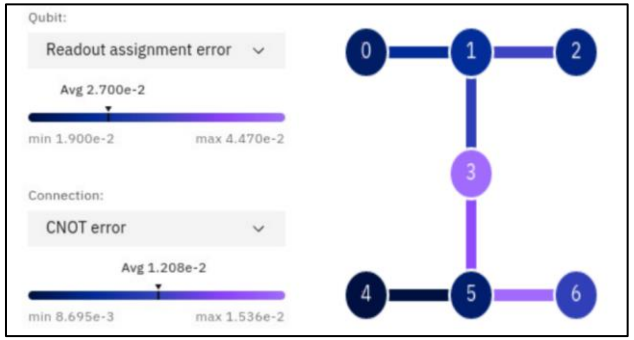
Quantum Simulator

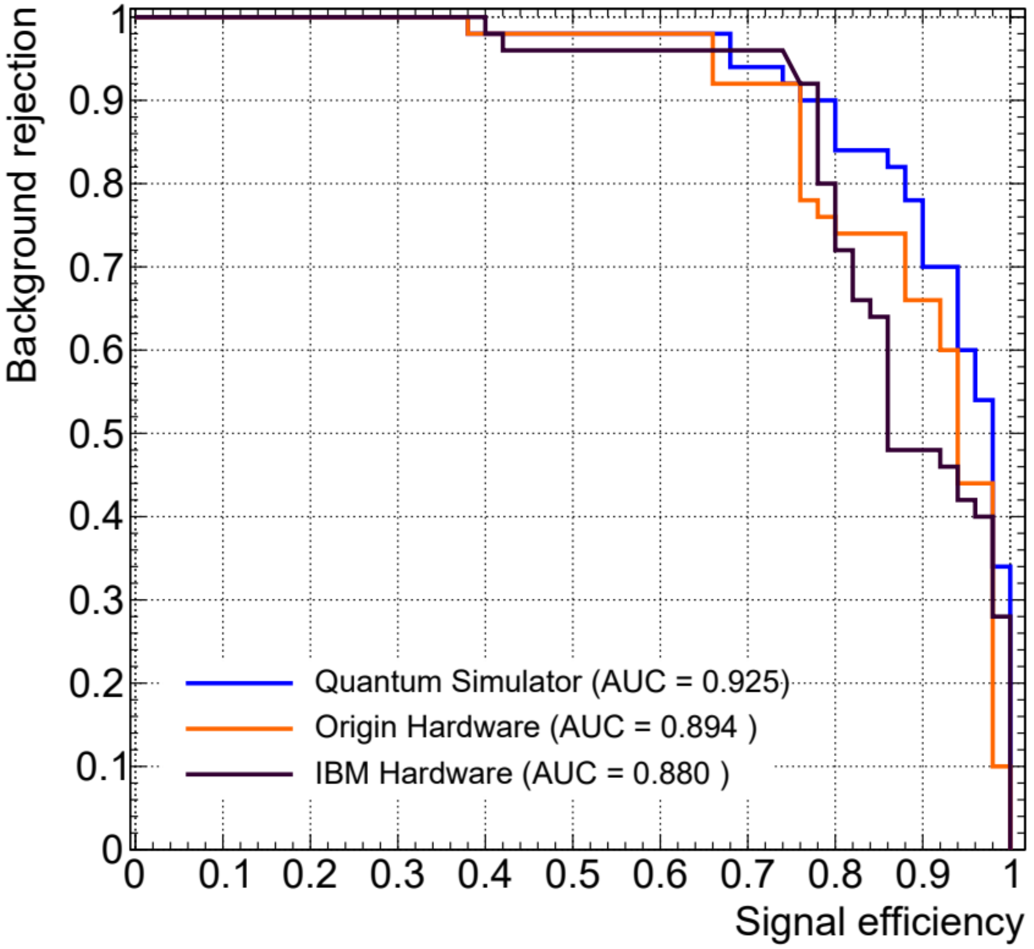




- 6-qubits on StatevectorSimulator from IBM Quantum & 6-qubits on Tensor Network Simulator from Origin QPanda (Z.-Y. Chen, G.-P. Guo, arXiv:1901.08340)
- Hardware noise not considered. Both simulators gave identical results.

Results from Quantum C. Hardware





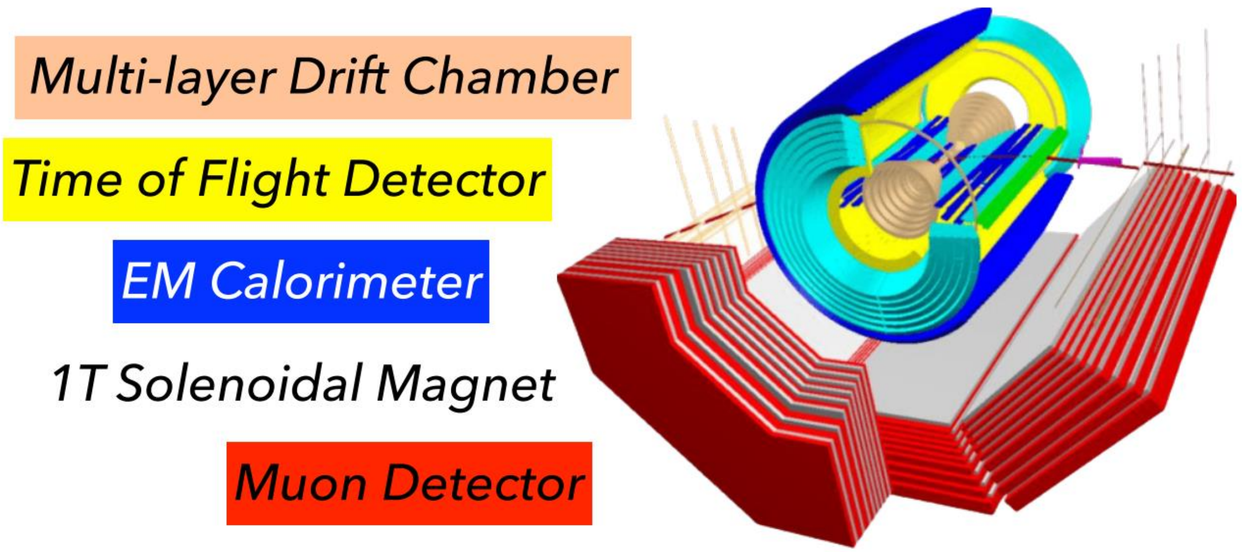
- 7-qubit IBM Nairobi machine (used only 6 qubits [0-5])
- 6-qubit (free usage) Origin Wuyuan machine
- Only using 100 events for training & testing respectively, due to long execution time of hardware.
- Performance approaching the noiseless conditions
- Difference b/w IBM & Origin from hardware noise & statistical fluctuations during the execution time.

Teng Li[1], Zhipeng Yao[1], **Jiaheng Zou[2], Tao Lin[2], Weidong Li[2]**, Xingtao Huang[1]
[1]Shandong U., [2] IHEP

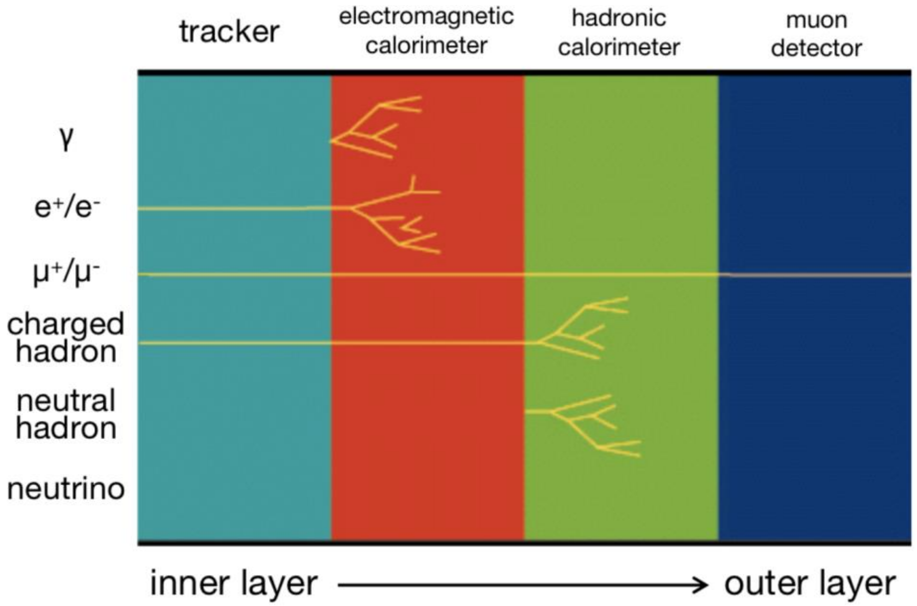
3. Quantum ML in PID for BES III

Beijing Spectrometer III (BES III)
e^+e^- τ-charm factory at 2.0 GeV $\leq \sqrt{s} \leq$ 4.9 GeV
Instantaneous luminosity 10^{33} $cm^{-2}s^{-1}$ at $\sqrt{s}$ = 3.77 GeV

- Machine learning is actively used for particle identification (PID).
- Muon / charged pion separation is a difficult PID task



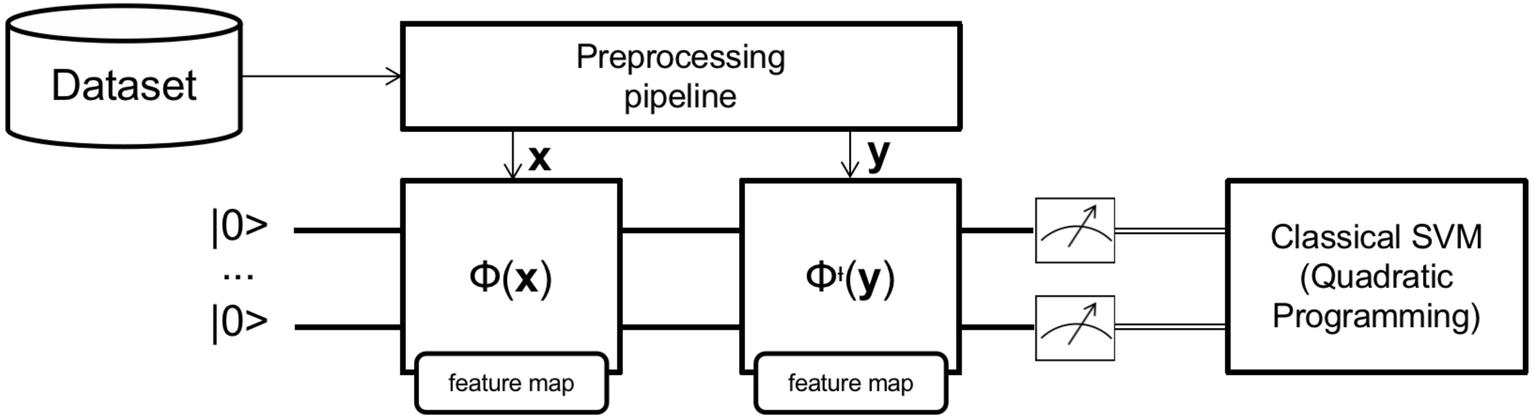
From Marco Scodeggio



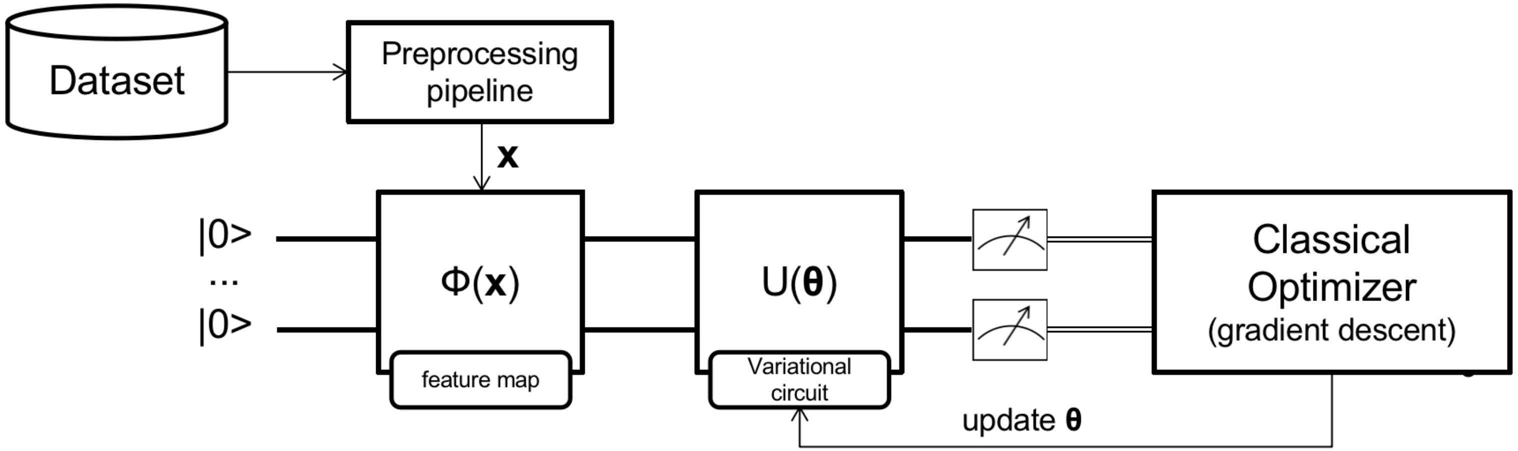
Teng Li[1], Zhipeng Yao[1], **Jiaheng Zou[2], Tao Lin[2], Weidong Li[2]**, Xingtao Huang[1]
[1]Shandong U., [2] IHEP

3. Quantum ML in PID for BES III

- **Method I: Quantum Support Vector Machine**



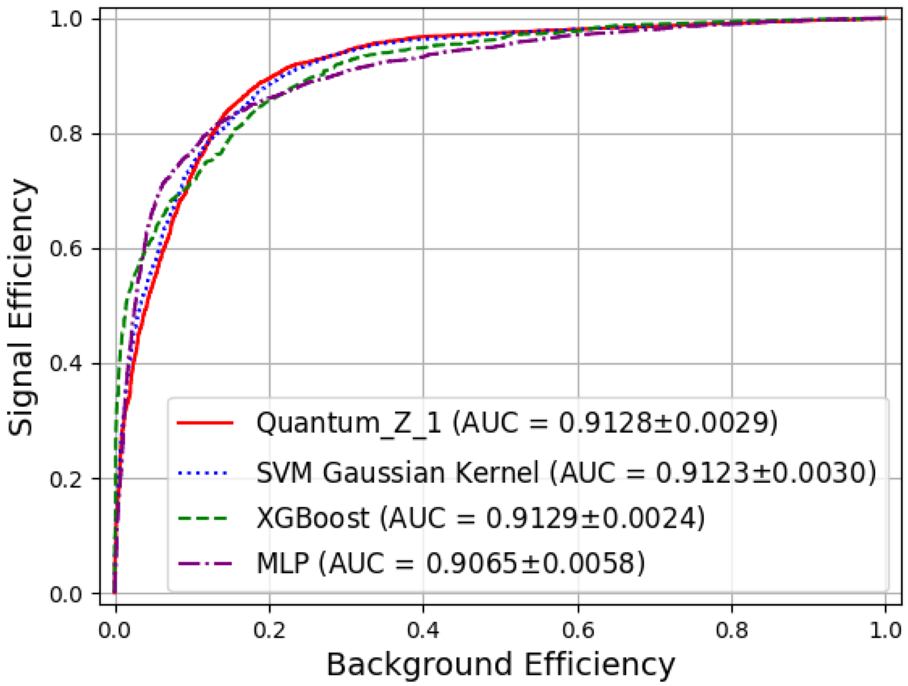
- **Method II: Variational Quantum Classifier**

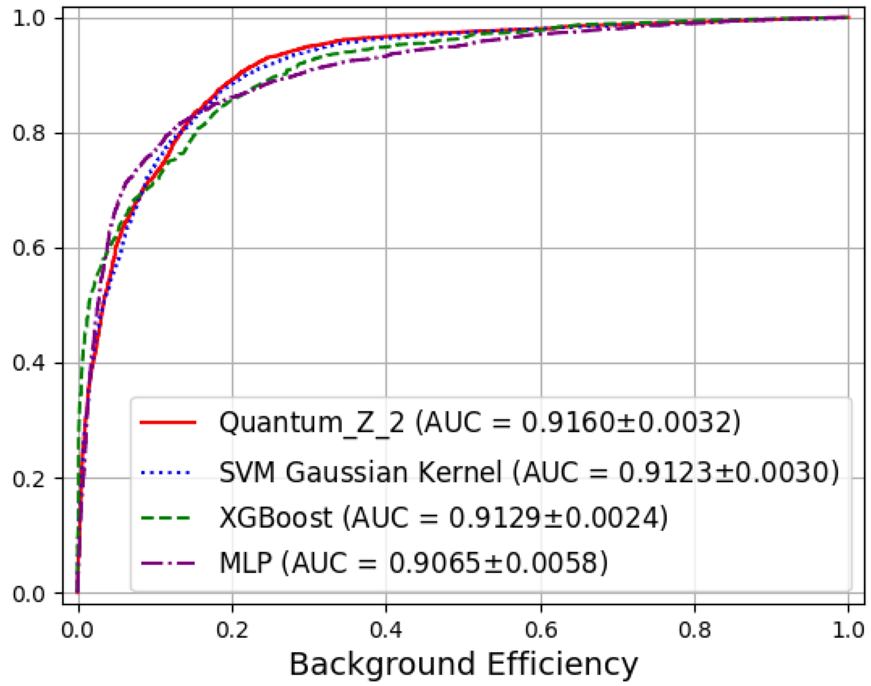


Teng Li[1], Zhipeng Yao[1], Jiaheng Zou[2], Tao Lin[2], Weidong Li[2], Xingtao Huang[1]
[1]Shandong U., [2] IHEP

BES III: Quantum SVM on Simulator

- **The discrimination power of simulated QSM and traditional ML models are compared**
  - Various types of encoding circuits simulated. Complicated circuits tend to lead to ovefitting
  - Similar discrimination power can be achieved using quantum simulator

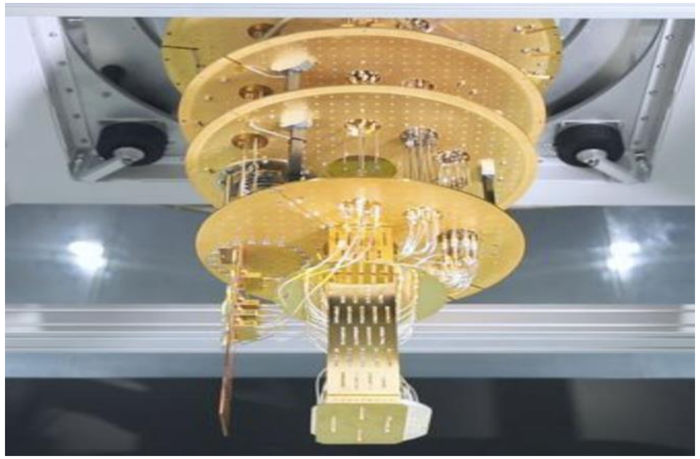


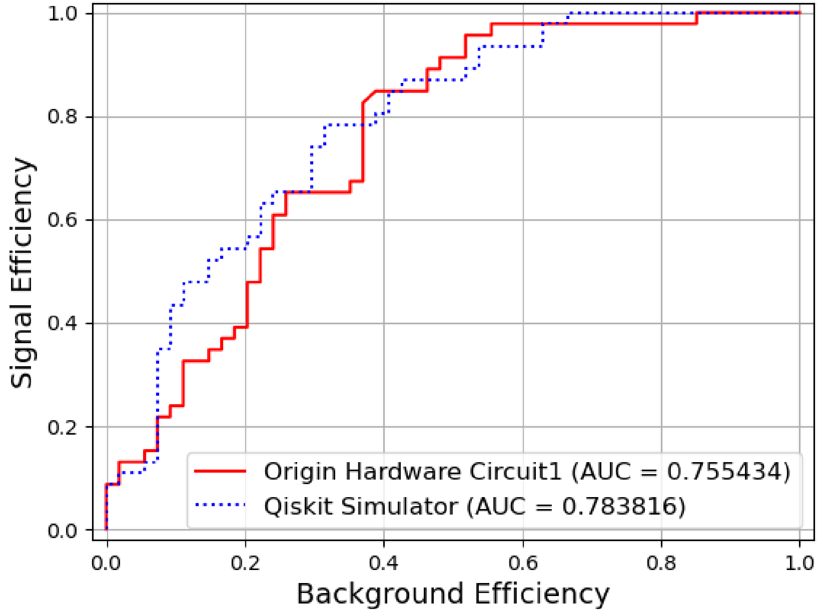


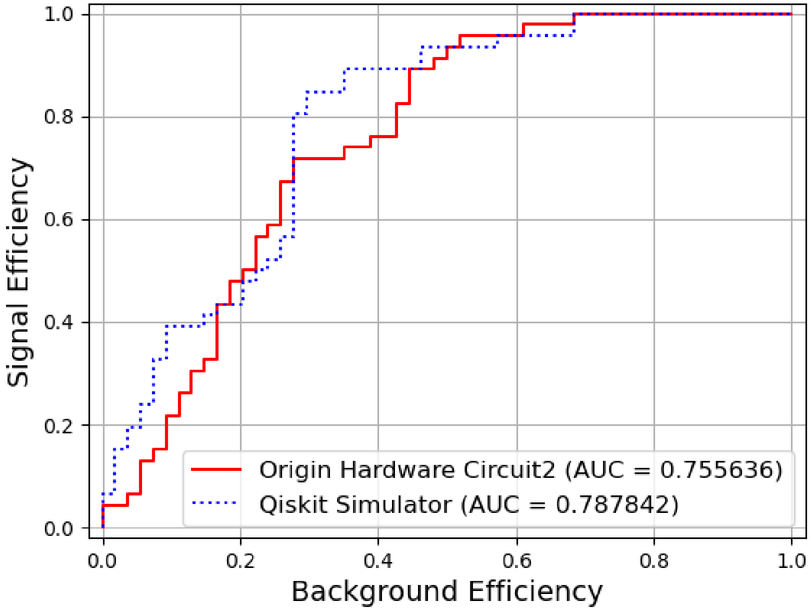
Teng Li[1], Zhipeng Yao[1], Jiaheng Zou[2], Tao Lin[2], Weidong Li[2], Xingtao Huang[1]
[1]Shandong U., [2] IHEP

BES III: Quantum SVM on Hardware

- Using Origin Quantum - Wuyuan system
  - Based on super-conducting technology
  - 6 qubits accessible publicly & freely (fees for higher number of qubits), controlled by QPanda API
- 100 training tracks, 100 testing tracks; averaged from three runs



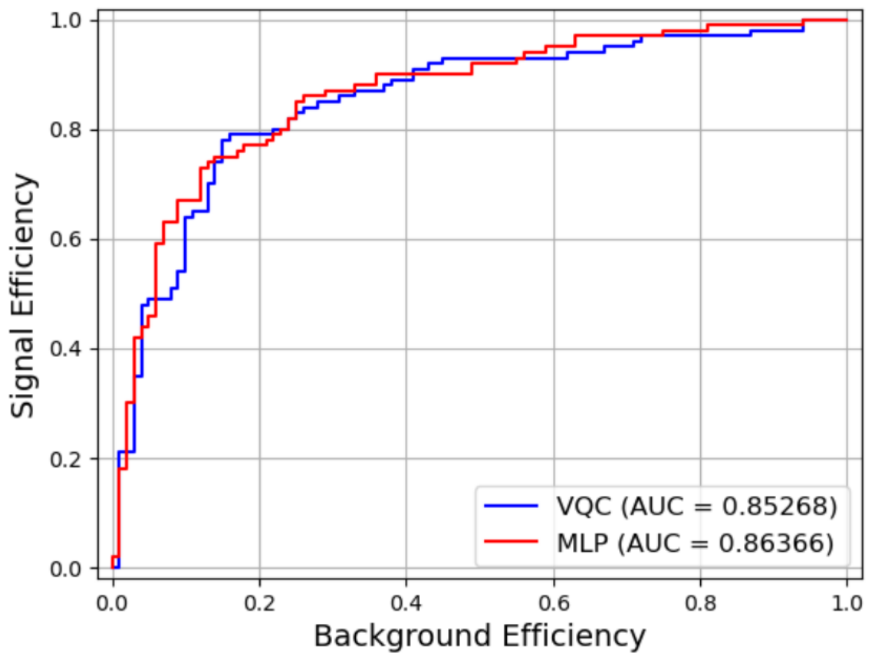




Teng Li[1], Zhipeng Yao[1], Jiaheng Zou[2], Tao Lin[2], Weidong Li[2], Xingtao Huang[1]
[1]Shandong U., [2] IHEP

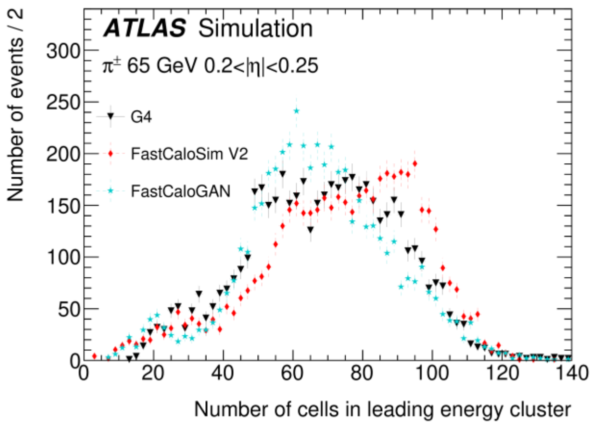
BES III: VQC on Simulator

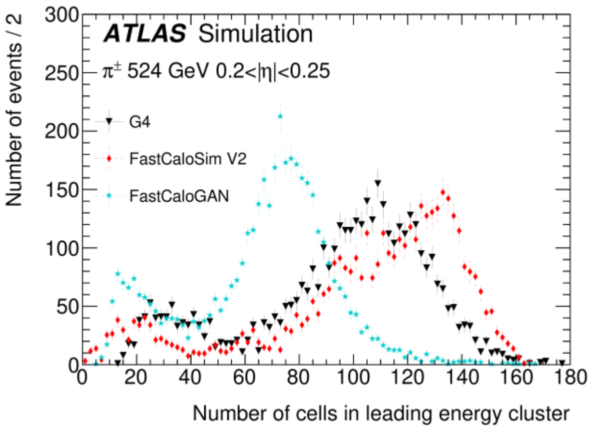
- The optimal VQC is compared with the classical MLP neural network
  - VQC config: EfficientSU2 with X_3 and L_BFGS_B optimizer
  - MLP config: 400x200x100x50x15 (relu, adam)
- On small samples, VQC performs similarly to the classical MLP neural network



4. Quantum GAN for ATLAS calo sim

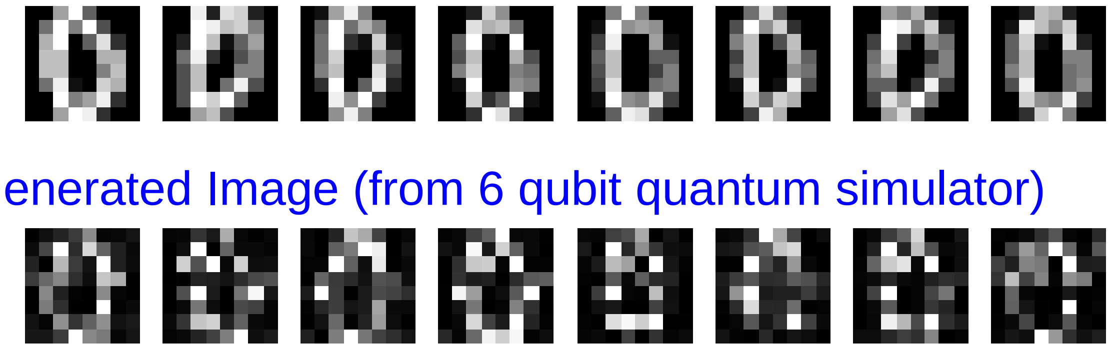
- HL-LHC will require enormous computing resource. **Already in Run 2, ~40% of the CPU is consumed by MC simulation.** (Complicated accordion geometry in EM calo)
  - e.g. 80% of the total simulation time taken by the shower development in the calo for ttbar events.
- Classical GAN is partially used in fast sim (Atlfast3), but is time consuming → Quantum GAN may be able to reduce the training time & improve the accuracy.
- Currently testing two approaches: (1) hybrid of quantum generator & classical discriminator (→for NISQ era), (2) full quantum version of generator & discriminator





Real Image

Generated Image (from 6 qubit quantum simulator)



Plans & Areas of Interest

- Error correction for lattice QCD studies (under investigation)
- Quantum computing for tracking (ATLAS, CEPC)
- Anomaly detection
- …

International Exchanges

- We are strongly interested in international exchanges at all levels (e.g. joint research, joint funding, Helmholtz-China postdoctoral fellowship, etc).
- Also actively recruiting international staffs, engineers & postdocs to work at IHEP.
  - Career workshop at CERN on December 2, 2022: https://indico.ihep.ac.cn/event/18414/
- Please contact me (hideki.okawa@cern.ch) if interested in any aspect!

Summary

- IHEP has established a working group for machine learning and quantum computing.
- Presented a few highlights from our recent studies in the high energy physics domain.
- Still at an early stage, but hope to expand the activities in the coming months & years.
- We are strongly interested in exchanges & collaborations with DESY!

backup – SPA-Net

SPA-Net Dataset & Selection

- Generated MadGraph 5 interfaced with Pythia8 for showering & hadronization
- Detector response with Delphes v3.4.2
- Top mass = 173 GeV

Object selection:

- Object overlap removal done
- Electron, muon p_T>25 GeV, $|\eta|$<2.5.
- Jet p_T>25 GeV, $|\eta|$<2.5 (dR matching considerd for truth jet-parton assignment)

Preselection: =1-lepton, $\geq$4 jets & $\geq$2 b-jets (for both $t\bar{t}$ & $t\bar{t}H$)

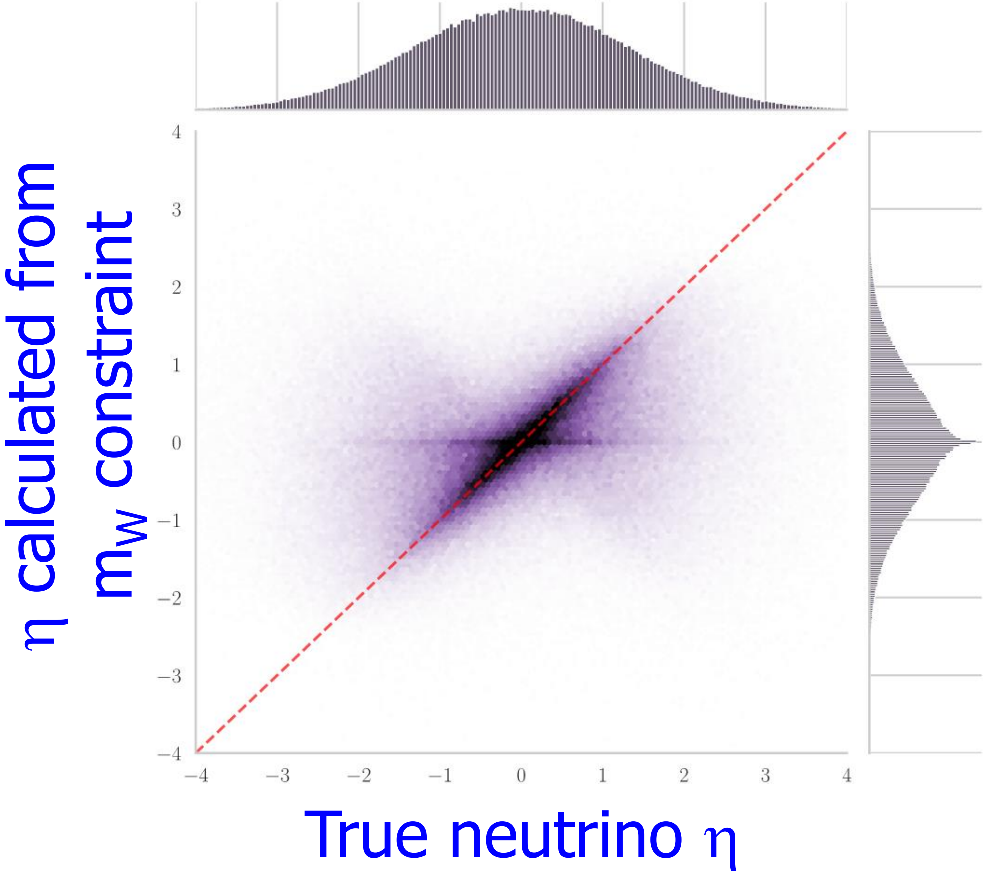
SPA-Net vs Baseline: Existing Methods

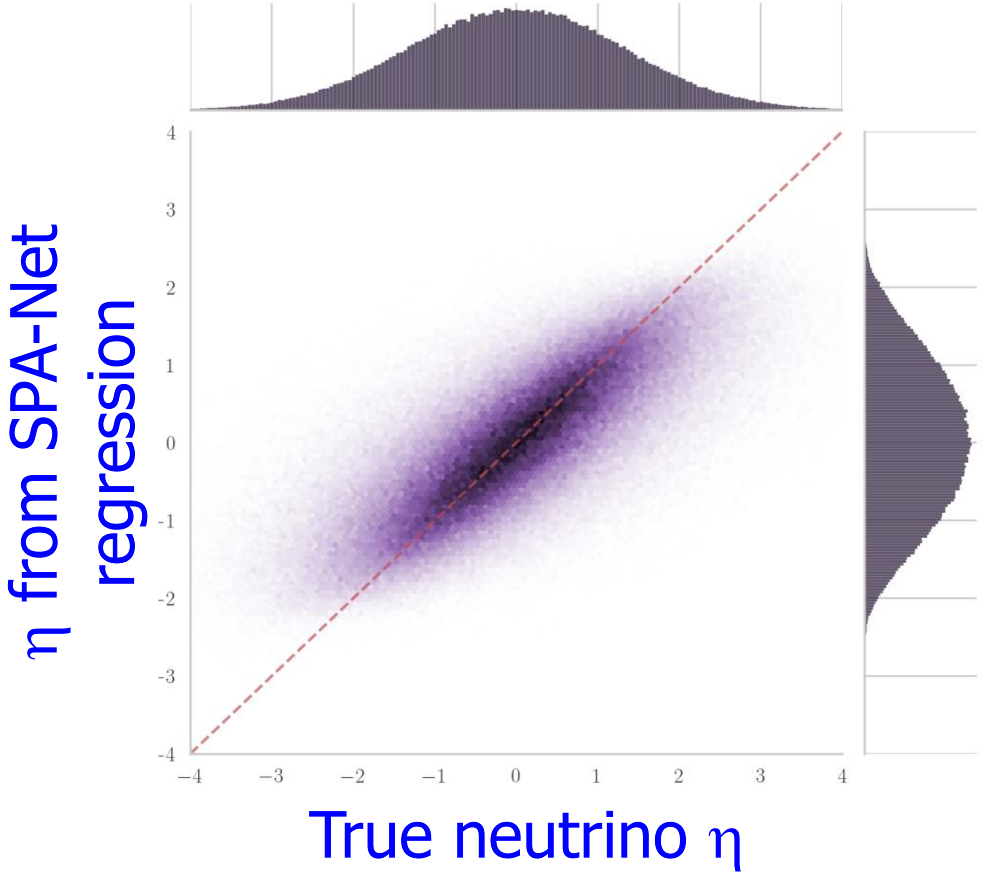
1. (χ^2 minimization: The simplest approach, considered in the previous all-hadronic studies; not considered in this talk)

2. KLFitter: likelihood-based kinematic fitting, assuming or not assuming a specific top mass (in this talk, the former);
   J. Erdmann et al., NIM A 748 (2014) 18

3. Permutation DNN: DNN considering all possible permutations;
   J. Erdmann et al., JINST 14 (2019) P11015

New Features: Regression of Kinematics



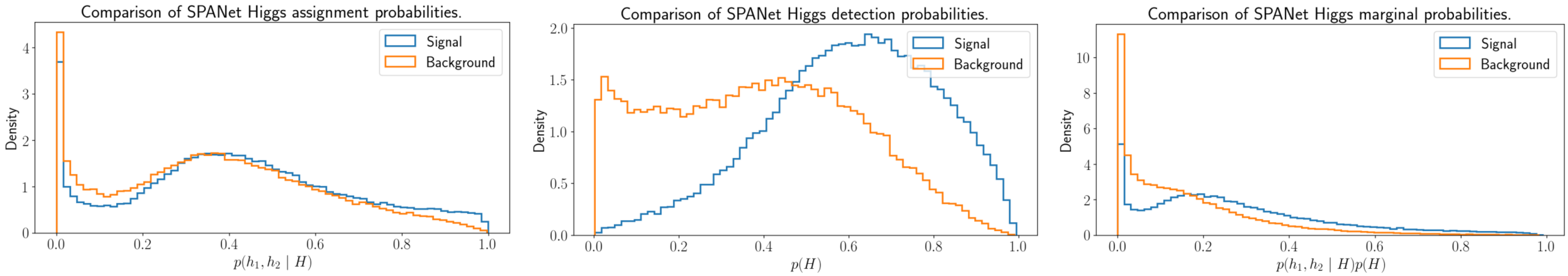


- SPA-Net can now train to reconstruct $t\bar{t}$ & simultaneously train to regress a variable (e.g. neutrino η or p_z, $t\bar{t}$ invariant mass)
- Neutrino η is more diagonal than the traditional method w/ improved RMS (1.39→0.9).

Updates: SPA-Net Probabilities

- Defined for each particle to reconstruct (t, H). **9 probabilities in total for ttH**. **SPA-Net can provide detailed information on the goodness of jet-parton assignment.**

- **Detection probability:** Is t or H reconstructable?
- **Assignment probability:** Given t or H is present, are the predicted jets correct?
- **(Pseudo-)marginal probability:** the product of the two above



SPA-Net Package (new version!)

https://github.com/Alexanders101/SPANet

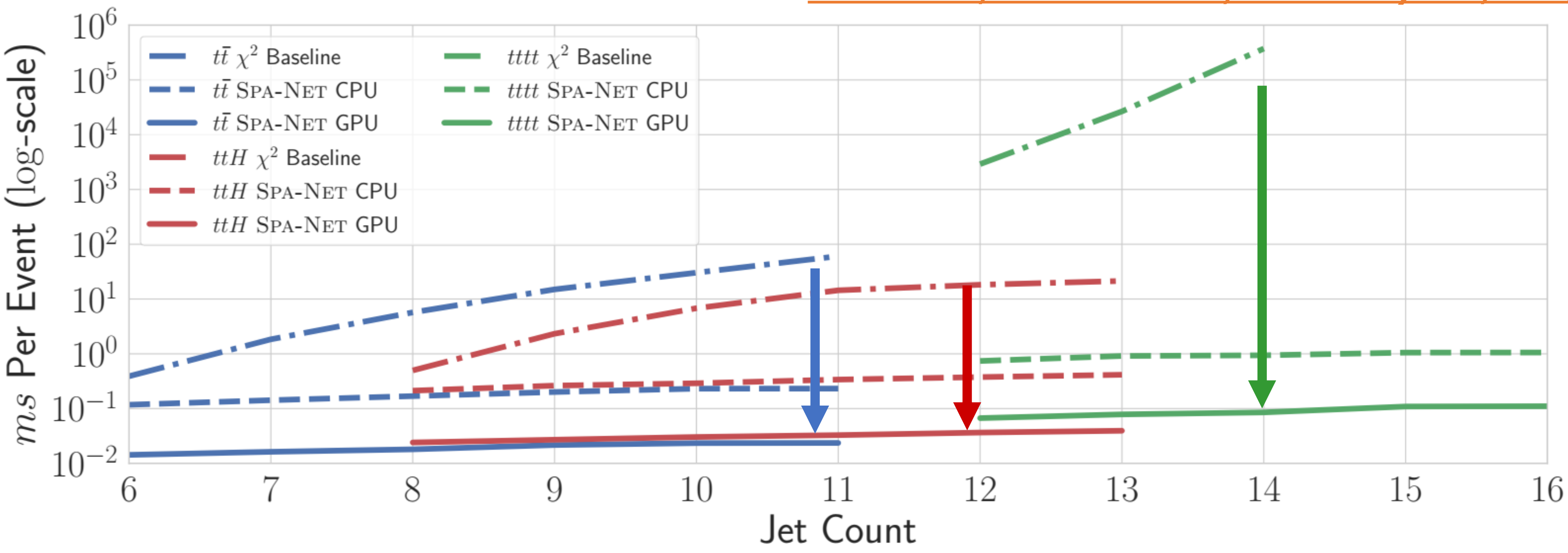
SPA-Net is not limited to top physics! Already considered for a SUSY RPV analysis & HH→4b in ATLAS as well.

New features in v2:

1. New configuration file format with more options on inputs and event topology.
2. Allow for several different inputs, including global inputs (e.g. MET, MET ϕ) for additional context.
3. New Regression and Classification output heads for performing per-event or per-particle predictions.
4. Gated transformers and linear layers for more robust networks. Less hyperparameter optimization.

SPA-Net CPU Time (prev. studies from allhad)

A. Shmakov, M. Fenton et al., SciPost Phys. 12, 178 (2022)



- **A few orders of magnitude improvement w/ SPA-Net compared to χ^2. A further acceleration w/ GPU.**
- Similar improvement expected for semi-leptonic final states & with much more visible improvement compared to KLFitter.

backup – Quantum Computing studies in BES-III

Original talk by Teng Li at HKUST IAS-HEP2022: https://indico.cern.ch/event/1096427/contributions/4671399/attachments/2372927/4052871/mu-pi-QSVM-ias.pdf

Scan of Various Encoding Circuits

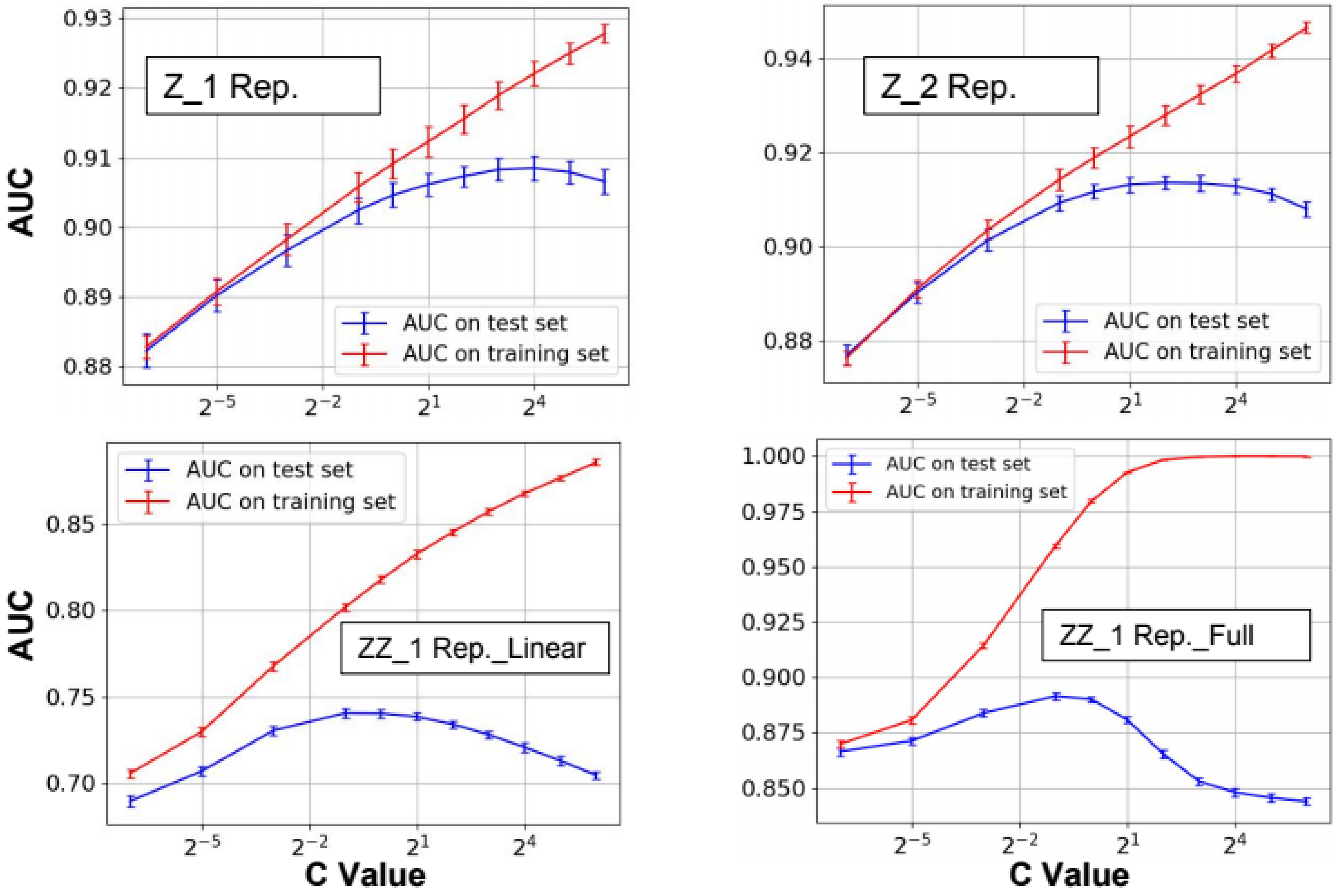
- Various types of encoding circuits are simulated using qiskit
  - A few simple circuits show comparable performance
  - Complicated circuits are prone to overfitting

Classical Model	AUC
SVM	0.91234±0.0030
BDT	0.91292±0.0024

Circuit	Rep.	Entanglement	Test set AUC	Training set AUC
X	2	none	0.90834±0.0030	0.91658±0.0021
	3		0.91238±0.0036	0.93055±0.0027
Z	1	none	0.90834±0.0030	0.91658±0.0020
	2		0.91238±0.0036	0.93055±0.0027
	3		0.89240±0.0036	0.90949±0.0009
XX	2	linear	0.73146±0.0037	0.84744±0.0017
		full	0.86332±0.0052	0.99887±0.0001
	3	linear	0.73198±0.0048	0.93766±0.0009
		full	0.72999±0.0047	0.99970±0.0002
ZZ	1	linear	0.73146±0.0037	0.84744±0.0025
		full	0.86332±0.0052	0.99887±0.0002
	2	linear	0.73198±0.0048	0.93767±0.0009
		full	0.72999±0.0047	0.99970±0.0002
	3	linear	0.67960±0.0040	0.88481±0.0004
		full	0.62707±0.0035	0.99964±0.0001
X, XX	2	linear	0.87201±0.0043	0.97902±0.0012
		full	0.88359±0.0021	0.99974±0.0001
	3	linear	0.84779±0.0010	0.99364±0.0002
		full	0.80892±0.0020	1.00000±0.0000
X, ZZ	1	linear	0.73145±0.0037	0.84745±0.0034
		full	0.86332±0.0052	0.99886± 0.0011
	2	linear	0.84441±0.0029	0.99136±0.0013
		full	0.82086±0.0029	1.00000±0.0000
	3	linear	0.84892±0.0029	0.99551±0.0002
		full	0.70668±0.0031	1.00000±0.0000
Z, ZZ	1	linear	0.87423±0.0024	0.97972±0.0004
		full	0.88378±0.0039	0.99974±0.0001
	2	linear	0.84675±0.0022	0.99253±0.0010
		full	0.80875±0.0032	1.00000±0.0000
	3	linear	0.83464±0.0030	0.99512±0.0003
		full	0.69984±0.0026	1.00000±0.0000

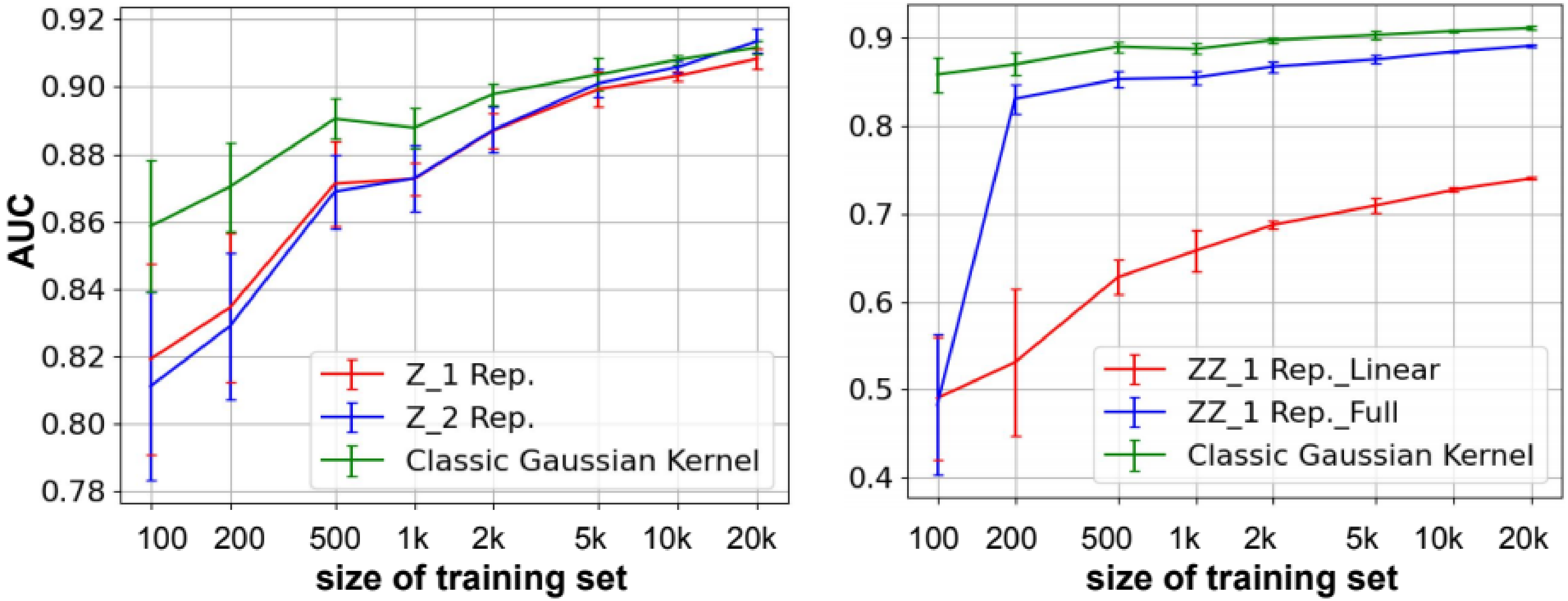
Influence of the Regularization Parameter

- The influence of the SVM regularization parameter can be carefully tuned to handle the overfitting/underfitting trade-off



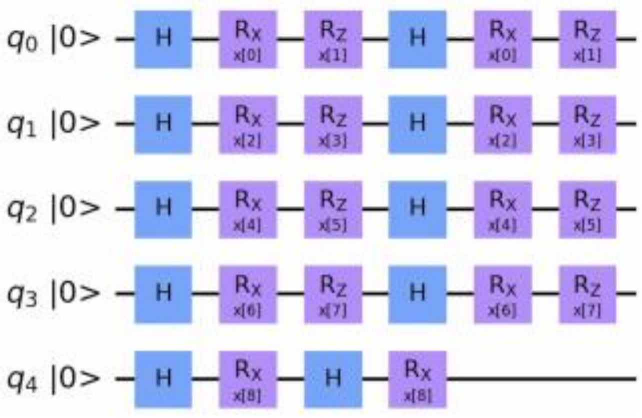
Influence of the training size

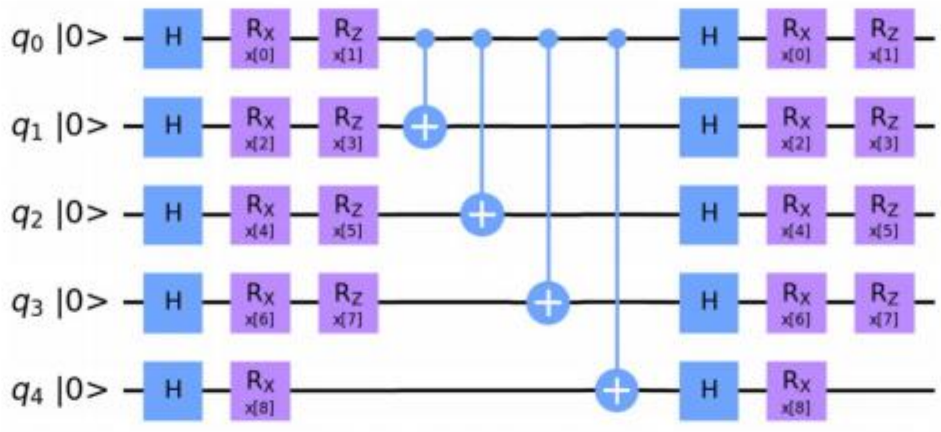
- Different size of the training set are tested
  - The quantum SVM usually shows unstable performance when the training size is small
  - Some circuits start to overtake Gaussian kernel with larger training sets



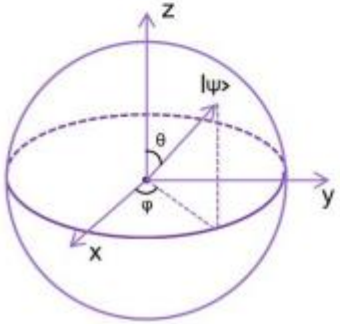
Compressed Feature Map on Quantum Hardware

- Two feature maps are re-designed to meet the limited number of qubits on the Wuyuan system
  - Two features are encoded into each qubit, based on the R_X and R_Z rotations





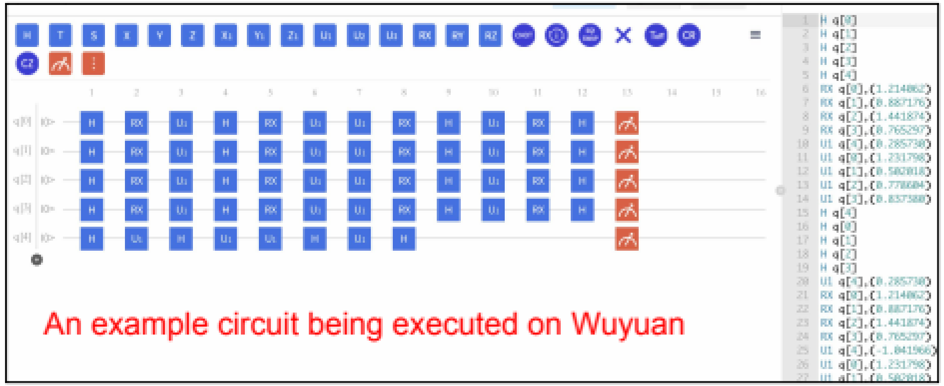
  - The feature map structure is carefully tuned for the best simulation results
    - AUC (1): 0.90373±0.0024
    - AUC (2): 0.91029±0.0023



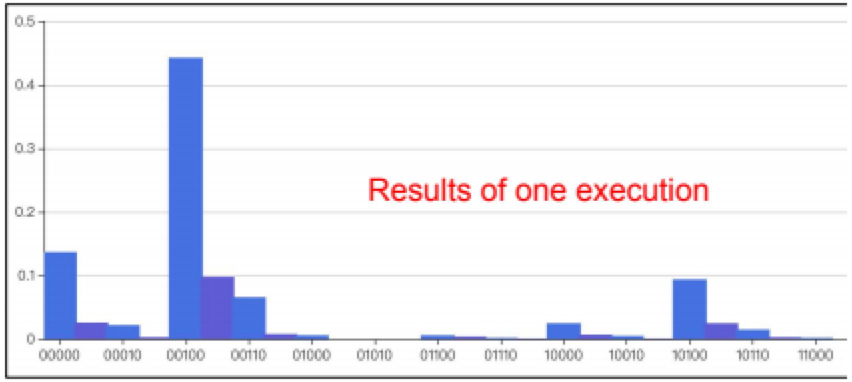
15

Job Execution on the Wuyuan System

- The quantum circuits are generated based on the dataset, then uploaded to the Wuyuan system via QPanda
  - Quantum circuits are automatically optimized against the Wuyuan system



An example circuit being executed on Wuyuan



Results of one execution

  - Results are transferred back to the classical computer for downstream computations

Search of Optimal Encoding Circuits

- A wide range of encoding circuits is simulated as well
  - Relatively simpler (X_3, Z_2) circuits provide better discrimination power. This is simillar with results from qSVM
  - Overfitting is less obvious comparing to qSVM

Circuit	Repetitions[1]	Entanglement[2]	Training set AUC[3]	Test set AUC[3]
X	2	none	0.85733±0.0158	0.83719±0.0334
	3		0.86581±0.0242	0.84354±0.0309
XX	2	linear	0.70366±0.2394	0.62051±0.0511
		full	0.74656±0.1962	0.61870±0.0456
	3	linear	0.69882±0.0331	0.56579±0.0850
		full	0.73998±0.0367	0.53483±0.0422
Z	1	none	0.84914±0.0416	0.81800±0.0607
	2		0.87338±0.0134	0.84729±0.0282
	3		0.75637±0.0349	0.70577±0.0965
ZZ	1	linear	0.69493±0.0274	0.59169±0.0770
		full	0.76208±0.2652	0.61317±0.0334
	2	linear	0.71653±0.0280	0.61720±0.0519
		full	0.72435±0.0267	0.55957±0.0497
	3	linear	0.71762±0.0274	0.59083±0.0377
		full	0.75620±0.0242	0.50290±0.0308
X,XX	2	linear	0.68095±0.0195	0.48709±0.0310
		full	0.74749±0.0258	0.52300±0.0423
	3	linear	0.68118±0.0165	0.51419±0.0288
		full	0.75874±0.0262	0.52024±0.0295
Z,ZZ	1	linear	0.65018±0.0100	0.49202±0.0327
		full	0.75312±0.0161	0.50179±0.0274
	2	linear	0.67567±0.0243	0.50894±0.0372
		full	0.75500±0.0259	0.50336±0.0324
	3	linear	0.65160±0.0179	0.50692±0.0387
		full	0.75223±0.0195	0.50762±0.0359
X,ZZ	1	linear	0.66164±0.0242	0.49825±0.0429
		full	0.74262±0.0202	0.50841±0.0288
	2	linear	0.66667±0.0248	0.50124±0.0381
		full	0.75893±0.0163	0.49098±0.0483
	3	linear	0.68708±0.0228	0.50050±0.0531
		full	0.75371±0.0240	0.48624±0.5777

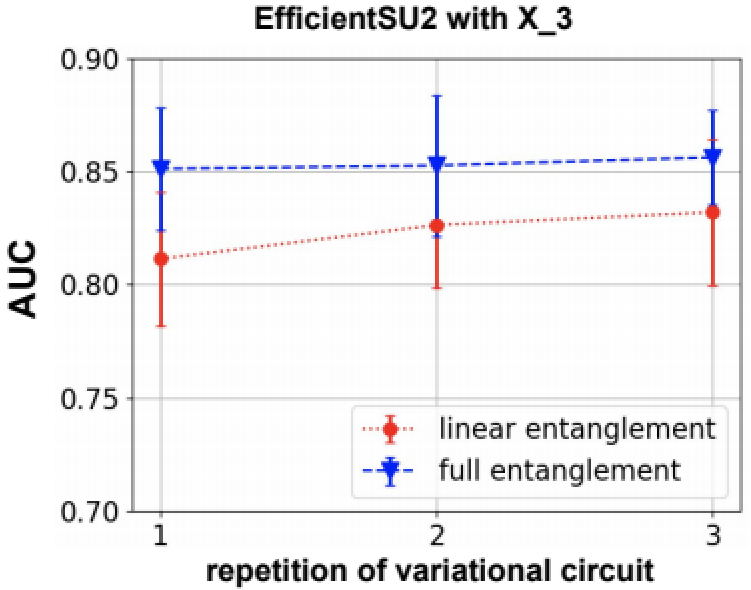
Search of Optimal Variational Circuits

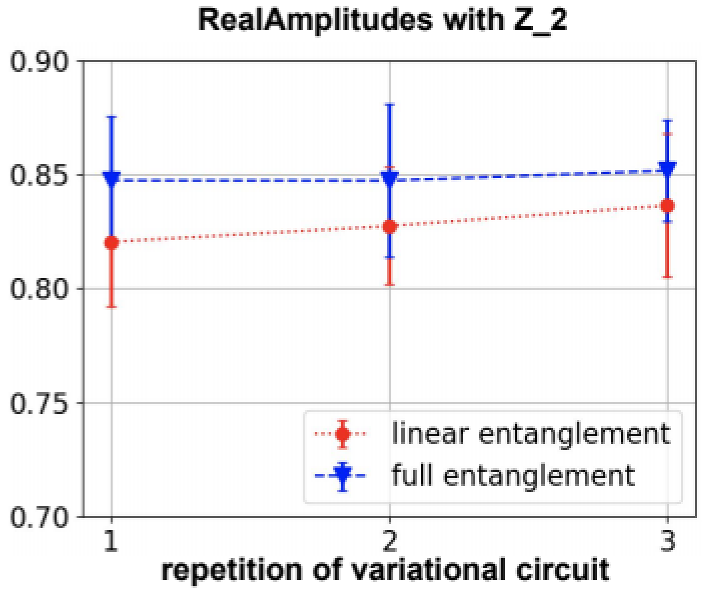
- A set of pre-studied N-Local ansatz are scanned based on the optimal encoding circuits
  - EfficientSU2: ansatz with single qubit spanned by SU(2) and CX
  - PauliTwoDesign: ansatz with single qubit Pauli rotations and pairwise CZ entanglements
  - RealAmplitudes: ansatz with single qubit Y rotations and pairwise CX entanglements
  - TwoLocal: ansatz with flexible rotation layers and entanglement layers
  - ExcitationPreserving: heuristic excitation-preserving wave function ansatz
- For X_3 and Z_2, the best variational circuits are EfficientSU2 and RealAmplitudes, respectively

variational circuit	Test set AUC with X_3	Test set AUC with Z_2
EfficientSU2	0.84983 ± 0.0292	0.84514 ± 0.0365
ExcitationPreserving	0.42660 ± 0.0327	0.52333 ± 0.0280
PauliTwoDesign	0.74614 ± 0.0277	0.76688 ± 0.0425
RealAmplitudes	0.84656 ± 0.0303	0.84711 ± 0.0277
TwoLocal	0.84102 ± 0.0330	0.84707 ± 0.0272

Search of Optimal Variational Cirtcuits

- Performance of different ansatz structures
  - Different entanglement methods and ansatz depth are simulated to study the impact on the performance
  - In general, the full entanglement method, and deeper ansatz (with more trainable parameters) always gives better discrimination power, but also consumes much more computing resource

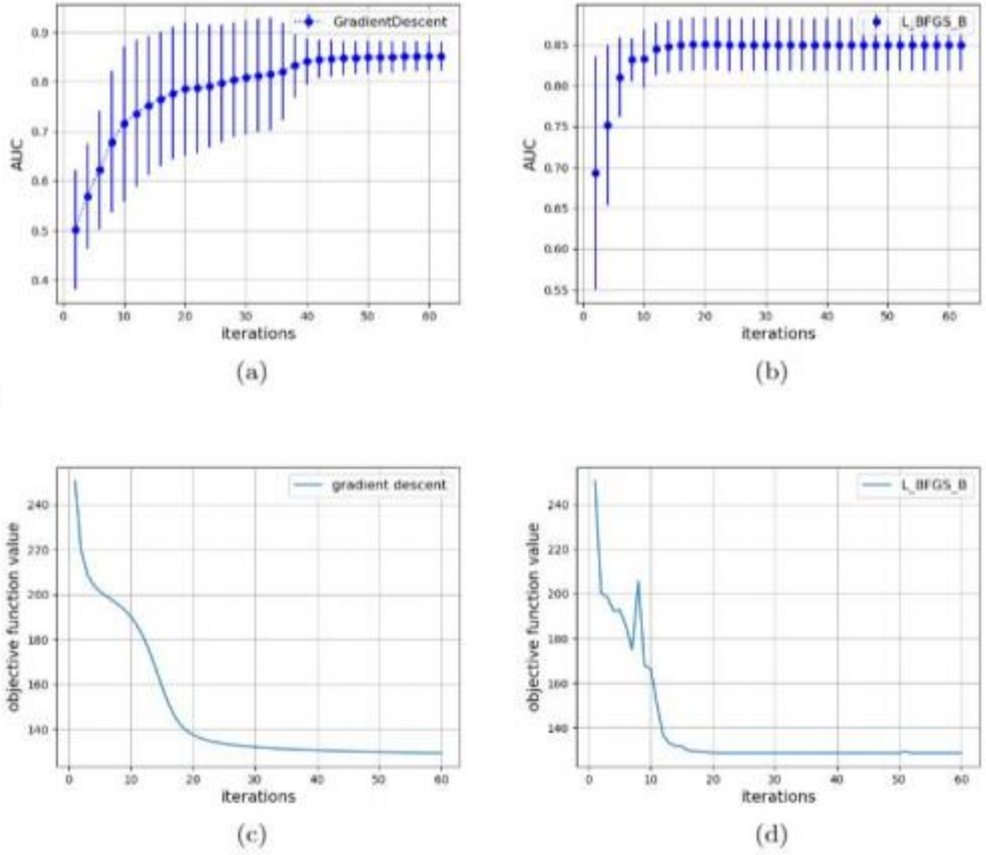




Optimization of VQC

❖ Common gradient descent method and Quasi-Newton method are compared

- Since L_BFGS_B (Quasi-Newton method) invokes the second derivative of the loss function, the convergence can be achieved much faster



(a)(c) Variation of AUC with the number of iterations
(b)(d) Variation in the objective function during the iteration.