UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Scuola di Scienze
Dipartimento di Fisica «Giuseppe Occhialini»
Corso di Laurea Magistrale in Fisica

QUANTUM TECHNOLOGY INITIATIVE

STUDY OF FEYNMAN DIAGRAMS WITH QUANTUM GRAPH NEURAL NETWORK

Tesi di Laurea Magistrale
Lorenzo Ballerio
Matr. 838676

Relatore Interno: *Prof. Andrea Giachero*
Relatore Esterno: *Dott. Michele Grossi*
Correlatore: *Dott. Roberto Moretti*
Correlatore: *Prof. Pietro Govoni*
Correlatore: *Dott. Danilo Labranca*

QUANTUM TECHNOLOGY INITIATIVE

Anno Accademico 2022/2023 – Sessione di Novembre

Summary

Idea of the Project

Idea of the Model

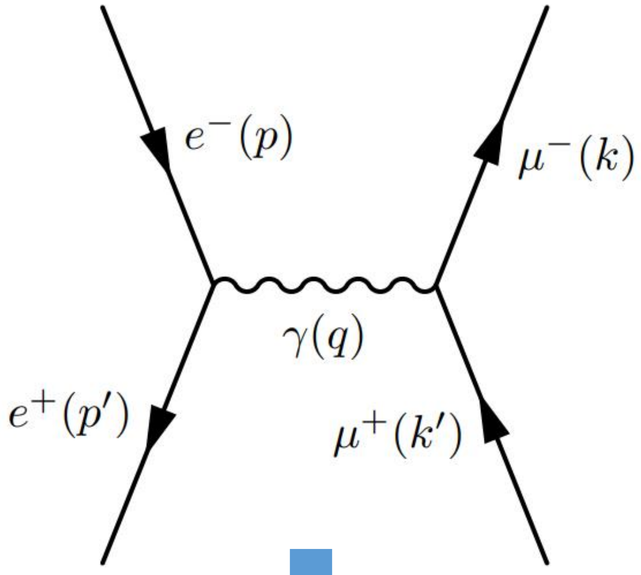
Our goal:

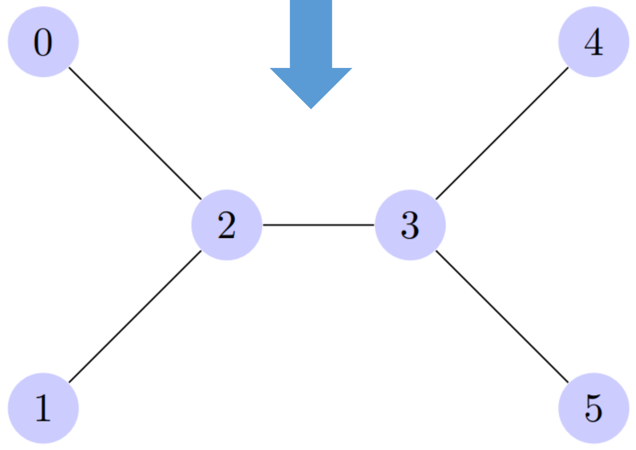
- Build **a quantum supervised regression model** that can compute the scattering amplitude squared for a Feynman diagram.

Strategy:

- Encode the diagram as a labeled **graph**

- **Topology** encoded in the adjacency matrix of the graph
- **Particles (m,Q,S)** encoded in the edges
- **Time flow** (initial state, interaction vertex, final state) encoded in the vertices





Scattering Amplitudes in Quantum Field Theory (QFT)

Scattering Theory I

Scattering Process: fundamental process where initial particles interact with one another by colliding and changing their direction or momentum**.**

Three different kinds of Scattering Processes:

- **Elastic Scattering**
- **Inelastic Scattering**
- **Deep Inelastic Scattering**

A scattering process is studied by measuring its **cross section**, related to the **likelihood of the process to occur**:

$$d\sigma = \frac{1}{(2E_1E_2)|\vec{v}_1 - \vec{v}_2|}|\langle f|M|i\rangle|^2 \prod_{final\ particles\ j} \frac{1}{2E_{p_j}} \frac{d^3\overrightarrow{p_j}}{(2\pi)^3}(2\pi)^4\delta^4\left(\sum_{init} p_i - \sum_{fin} p_k\right)$$

The cross section gives information about **fundamental interactions**.

Scattering Theory II

We are interested in $|\langle \boldsymbol{f}|\boldsymbol{M}|\boldsymbol{i}\rangle|^2 \equiv |\boldsymbol{M}|^2$, the **scattering amplitude squared** of the process:

- **Probability that from the initial state $|\boldsymbol{i}\rangle$ we can get to the final state $|\boldsymbol{f}\rangle$.**

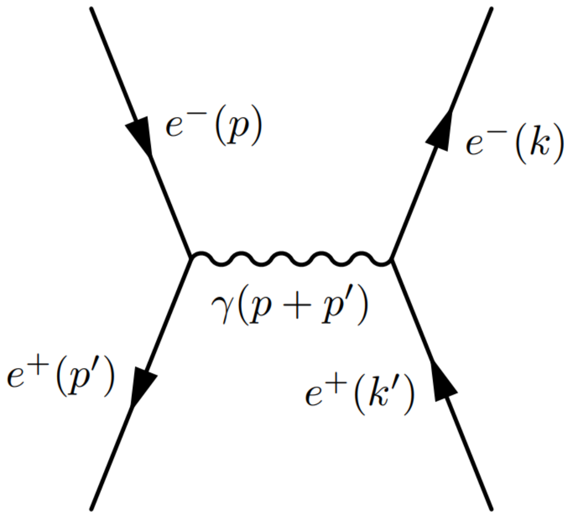
- It describes the **dynamics** of the scattering

- Computed by using **perturbation theory**

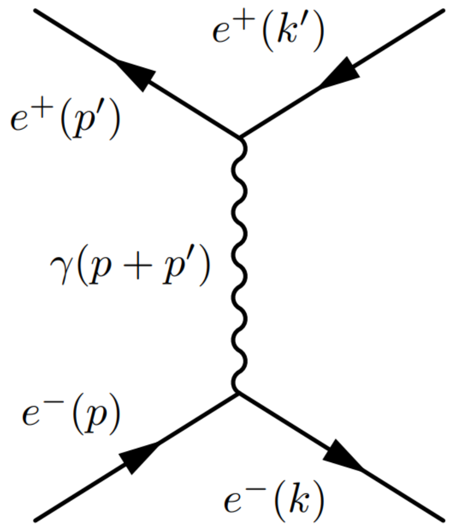
  $M = \sum_{n=1}^{\infty} M^{(n)}$ where $M^{(n)}$ is the n-th order perturbation term

- **Feynman diagrams** formalism:

  Associating a mathematical expression to each graph component: **Feynman Rules**



(a)



(b)

Scattering Theory III

A generic scattering process at fixed perturbation order may include multiple Feynman diagrams:

$$|M_{tot}|^2 = |M_a + M_b|^2 = |M_a|^2 + |M_b|^2 + M_a M_b^* + M_a^* M_b$$

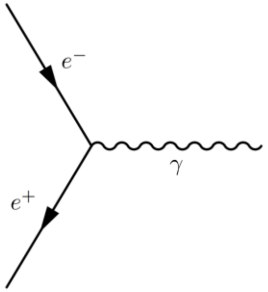
Appereance of an interference term, quantum effect

- **Quantum Electrodynamics (QED):** $\mathcal{L}_{\text{QED}} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \bar{\psi}(i\partial_\mu \gamma^\mu - m)\psi + e\bar{\psi} A_\mu \gamma^\mu \psi$

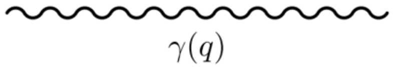
Interaction term $\mathcal{L}_{int}$

Feynman rules for QED in momentum space:

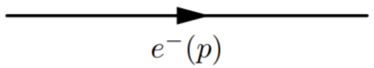
- **Interaction vertex**: $ie\gamma_\mu$



- **Photon propagator** (photonic internal line): $-\frac{ig_{\mu\nu}}{q^2}$



- **Fermionic propagator** (fermionic internal line): $\frac{i}{p_\mu \gamma^\mu - m}$



- **External line**:

e^- in	p →●	$u_s(p)$
e^- in	p ●→	$\bar{u}_s(p)$
e^+ in	p ←●	$\bar{v}_s(p)$
e^+ out	p ●←	$v_s(p)$
γ in	k ∿∿●	$\varepsilon_s(k)$
γ out	k ●∿∿	$\varepsilon_s(k)$

Quantum Computing and Quantum Graph Neural Networks

Qubits

Qubit:

Two-level quantum systems:

$$|\boldsymbol{\psi}\rangle = \boldsymbol{\alpha}|\mathbf{0}\rangle + \boldsymbol{\beta}|\mathbf{1}\rangle$$

with $|\boldsymbol{\alpha}|^2 + |\boldsymbol{\beta}|^2 = \mathbf{1}, \quad \boldsymbol{\alpha},\ \boldsymbol{\beta} \in \mathbb{C}$

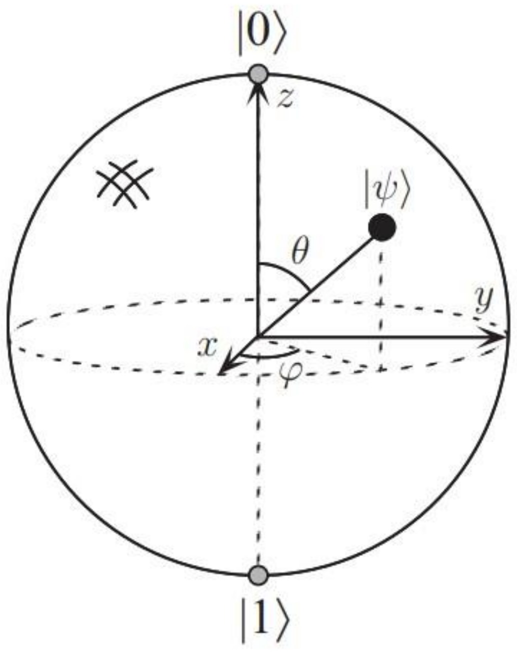
Quantum Computing (QC) is the paradigm that exploits quantum mechanical properties of matter in order to do calculations.

- Qubits are the **fundamental unit of quantum computation**
- $\{|0\rangle, |1\rangle\}$ is called computational basis and corresponds to the logical bits $0,1$

- Qubits live in $\mathcal{H} \sim \mathbb{C}^2$ with 3 real d.o.f
- Can be defined as a point on a spherical surface with unit norm:

$$|\psi\rangle = e^{i\gamma}\left[\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle\right]$$

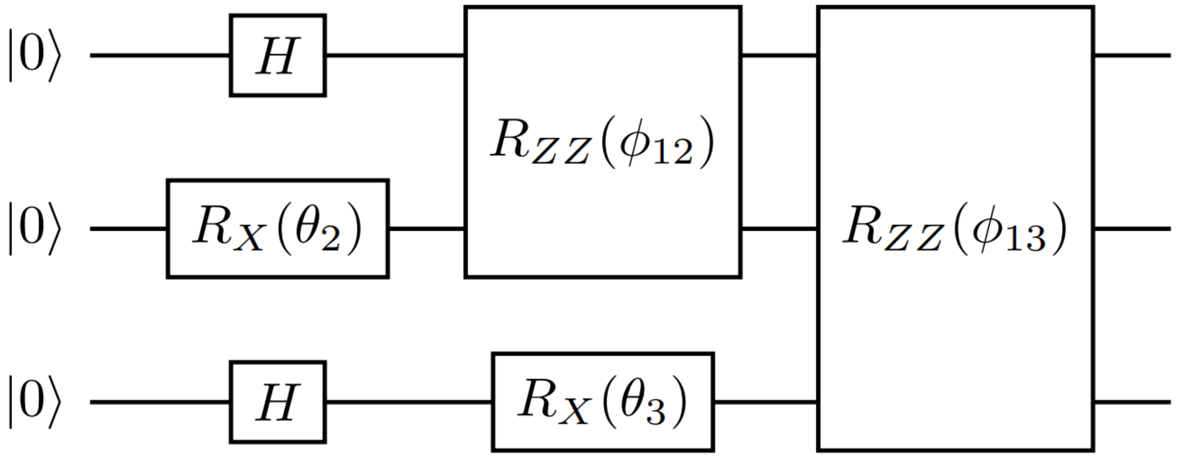
with $\theta, \varphi, \gamma \in \mathbb{R}$



Bloch Sphere

Quantum Circuits

- Information processed via unitary operators called **quantum gates**, collected in quantum circuits.



- Quantum Computing is by definition **reversible** and **universal**.
  - can be defined a finite set of gates that can approximate any unitary operation (any function).

- Examples of quantum gates:
  - $H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ Hadamard gate:
    $$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle$$
    $$H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle$$
  - Single qubit Pauli rotation along $\vec{n}$:
    $$R_{\vec{n}}(\theta) = e^{\frac{i\theta}{2}\vec{\sigma}\cdot\vec{n}}$$
  - $R_2(\gamma, \rho) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & e^{-i\rho} \\ e^{i\gamma} & e^{i(\rho+\gamma)} \end{bmatrix}$
  - $R_1(\tau) = e^{i\tau}R_Z(\tau) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\tau} \end{bmatrix}$
  - $R_{ZZ}(\phi) = \exp\left(-i\frac{\phi}{2}Z \otimes Z\right)$ Ising ZZ coupling gate

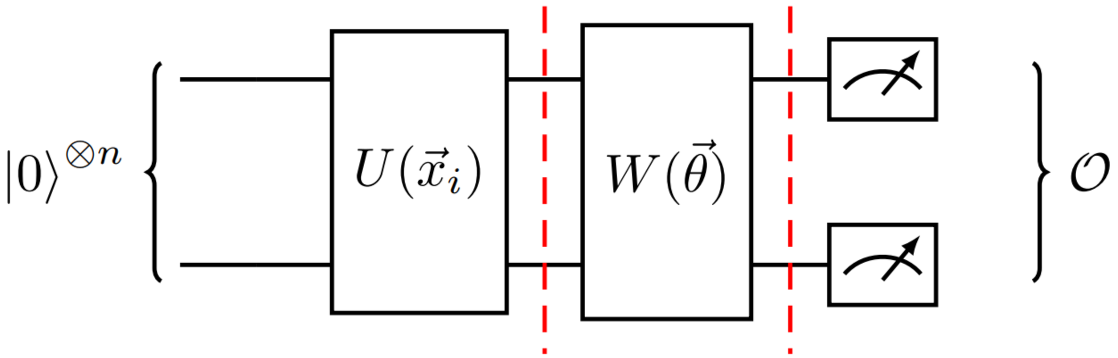
2-qubits operators, creates **entanglement** between qubits.

Quantum Neural Network (QNN)

Quantum Machine Learning (QML) combines Machine Learning and Quantum Computing with the aim of exploiting the principles of quantum mechanics to develop new learning models.

Quantum Neural Networks are n-qubits circuits, usually divided in four sections:

- **State initialization**
- **Feature Map** $\boldsymbol{U(\vec{x_i})}$, encoding of the input data $\vec{x_i}$
- **Ansatz** $\boldsymbol{W(\vec{\theta})}$, **Parametrized Quantum Circuit** that depends on trainable parameters $\vec{\theta}$
- Output extraction, measurement of an **observable** $\boldsymbol{O}$



Scheme for a generic QNN

QGNN for Feynman diagrams

Datasets

Data structure: $[\,(\boldsymbol{G}, \boldsymbol{\theta}, \boldsymbol{p}), \hat{\boldsymbol{y}}]$

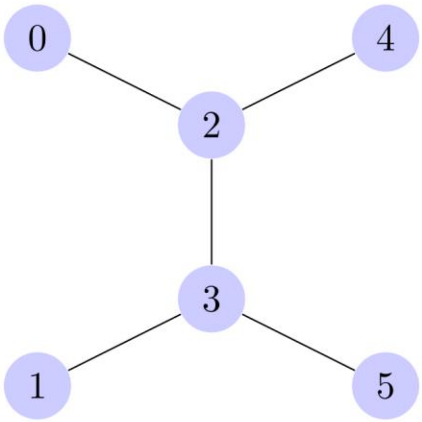
- G labeled graph $\boldsymbol{G = (V, E, X_V, X_E)}$
  - **V** set of vertices
  - **E** set of edges
  - $\boldsymbol{X_V}$ vertex feature set (1-hot encoding)
  - $\boldsymbol{X_E}$ edge features set (m, Q, S)
- $\boldsymbol{\theta}$ scattering angle
- **p** momentum of the initial particles
- $\hat{\boldsymbol{y}}$ theoretical value of $\overline{|M|}^2$ computed analitycally

Our usecase **tree level QED**:

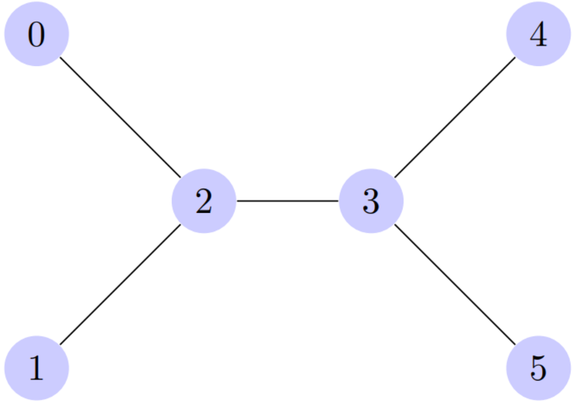
- Bhabha scattering (s-channel and t-channel)
- $e^+e^- \to \mu^+\mu^-$

- Any graph has the same number of vertices $|\boldsymbol{V(G)}| = \boldsymbol{6}$, the QGNN has a constant number of qubits.
- Very similar edge sets E_s and E_t

When $\vec{x_i}$ is a graph-structured data this model is called **Quantum Graph Neural Network (QGNN)**.



t-channel

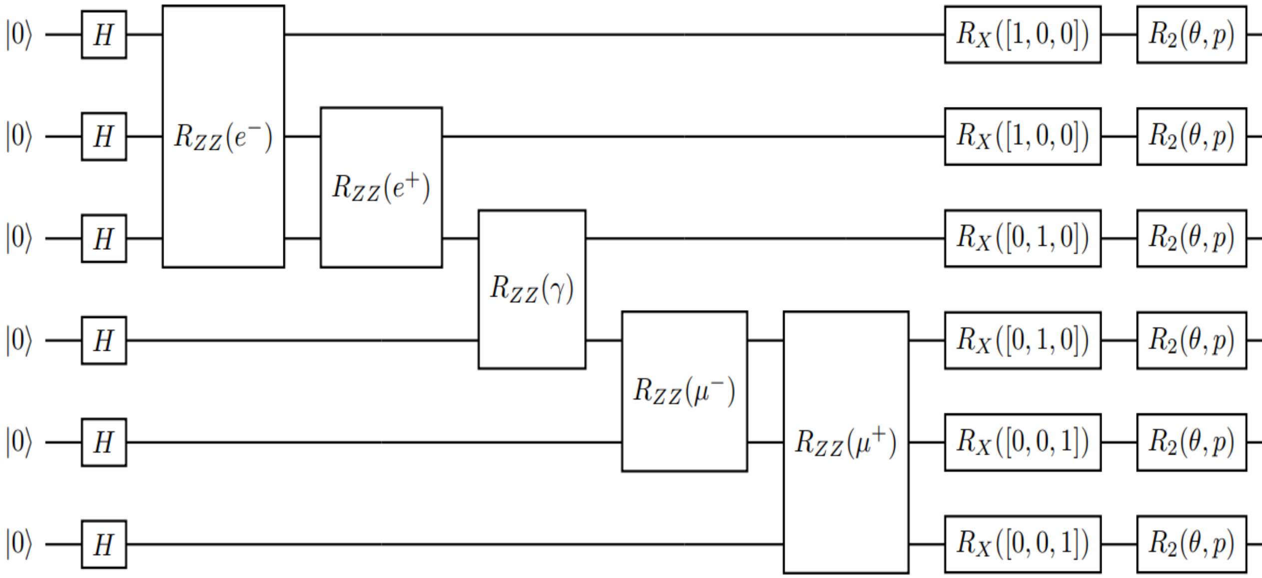


s-channel

Feature Map

Best model to hande this kind of data is a **QGNN**.
Model Architecture:

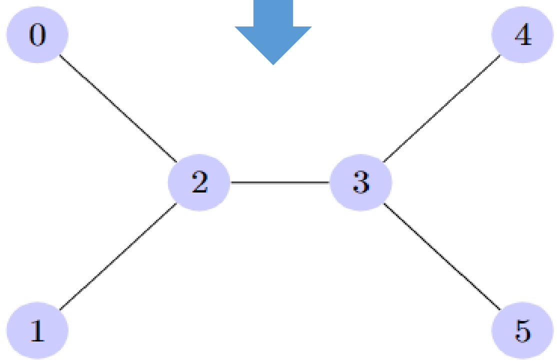
- Data encoding → **Feature Map**



$$U(G,\theta,p,\alpha)=\prod_{v\in V(g)} R_2^{(v)}(\theta,p)\prod_{j\in V(G)} R_X^k\left(x_V^k\cdot\begin{bmatrix}\frac{\pi}{2}\\ \pi\\ \frac{3\pi}{2}\end{bmatrix}\right)\prod_{(i,j)\in E(G)} R_{ZZ}^{(i,j)}(x_E^{(i,j)}\cdot\alpha)\prod_{a\in V(G)} H^{(a)}$$

The qubits correspond to the **vertices** of the graph.

Connectivity of the **circuit equal to the topology** of the graph.

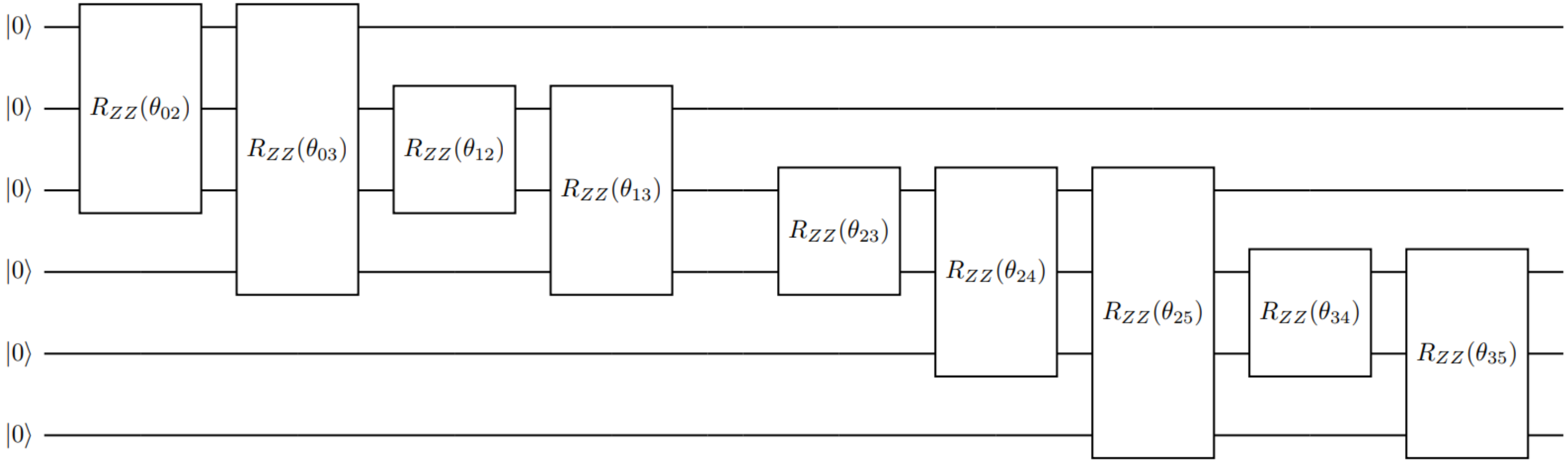


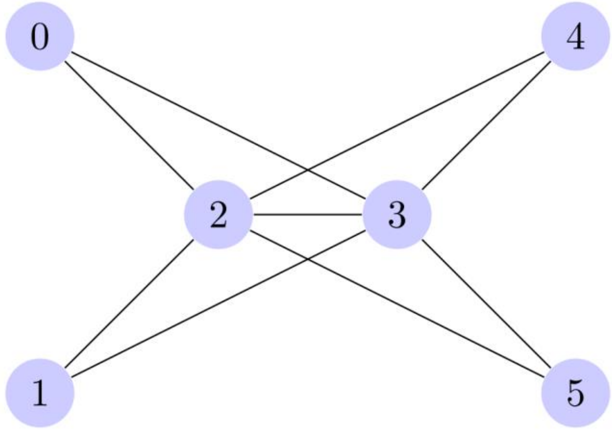
Ansatz of the Model

It includes different graph **topologies**

- Trainable part of the circuit → **Ansatz**
- Similar architecture of the Feature Map, but it has more connectivity among qubits.

$$W(\omega) = \prod_{l=1}^{L}\left(\prod_{v\in V(\bar{G})} R_2^{(v)}(\omega_{|E(\bar{G})|+|V(\bar{G})|+v,l}) \prod_{k\in V(\bar{G})} R_X^{(k)}(\omega_{|E(\bar{G})|+k,l}) \prod_{(i,j)\in E(\bar{G})} R_{ZZ}^{(i,j)}(\omega_{ij,k})\right)$$





Graph that includes all the possible QED processes at tree level

Output of the QGNN

- **Output of the circuit:**

$$f(G,\theta_i,p_i) \equiv \langle O \rangle = \langle \psi | O | \psi \rangle = \left\langle 0 \middle| U^\dagger(G,\theta_i,p_i,\alpha) W^\dagger(\omega) O(a,b) W(\omega) U(G,\theta_i,p_i,\alpha) \middle| 0 \right\rangle \equiv \overline{|M|}^2(G,\theta_i,p_i)$$

Where the **observable $\boldsymbol{O}$** is defined as:

$$O = \tilde{O}^{(1)} \otimes \mathbb{I}^{\otimes 5}, \tilde{O} = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix} \Rightarrow \overline{|M|}^2(G,\theta_i,p_i) = |\alpha(G,\theta_i,p_i)|^2 a^2 + |\beta(G,\theta_i,p_i)|^2 b^2$$

Unpolarized scattering amplitude:

$$\overline{|M|}^2 = \frac{1}{N_{init}} \sum_{init} \sum_{fin} |M|^2$$

Where $a, b \in \Re^+$, $a, b \neq 1$ are trainable parameters too.

Training process:

- 0.8-0.2 split for training and test set
- 5-fold cross validation with the training set
- Optimize the trainable tensor ω with respect to the cumulative squared loss function:

$$SSE = \sum_{i}^{batch-size} (f(G,\theta_i,p_i) - \hat{y}_i)^2$$

Performance of the Quantum Graph Neural Network

Performance for Bhabha s-channel I

Bhabha s-channel:

$$e^+e^- \to e^+e^-$$

- high energy limit: electrons are **massless**

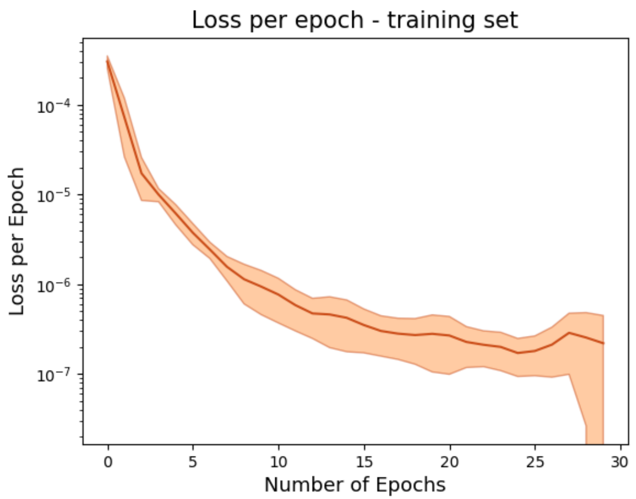
$$\overline{|M|}^2 = e^4[1 + \cos^2(\theta)]$$

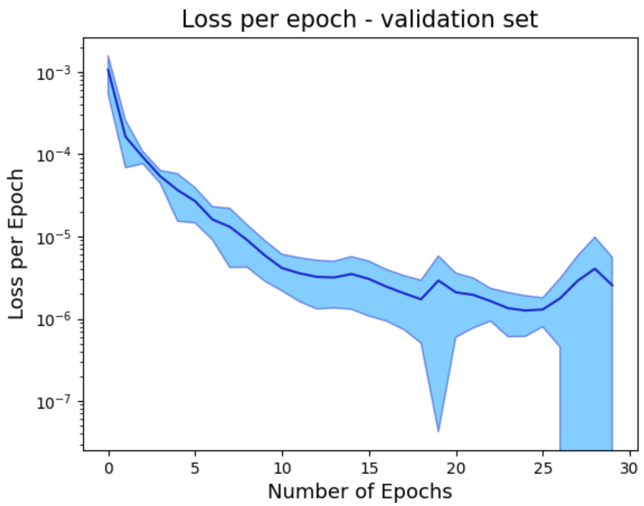
- $\theta \in [0.5, \pi]$

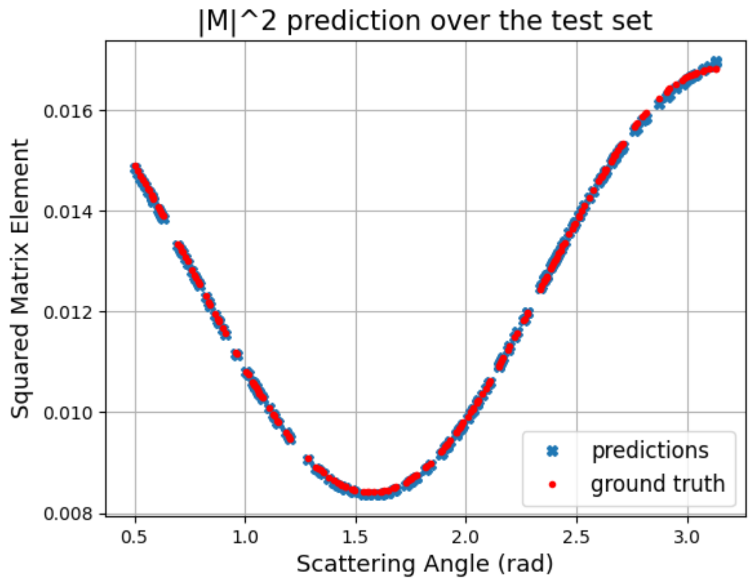
- 3-layers (repetition of the ansatz)

Training MSE	Validation MSE
$(1.104 \pm 1.140) \times 10^{-8}$	$(1.597 \pm 1.907) \times 10^{-8}$

MSE after test evaluation: 1.838×10^{-9}







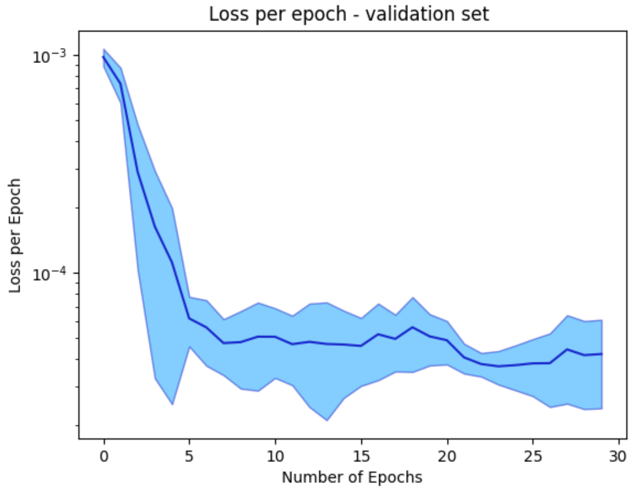
Performance for Bhabha s-channel II

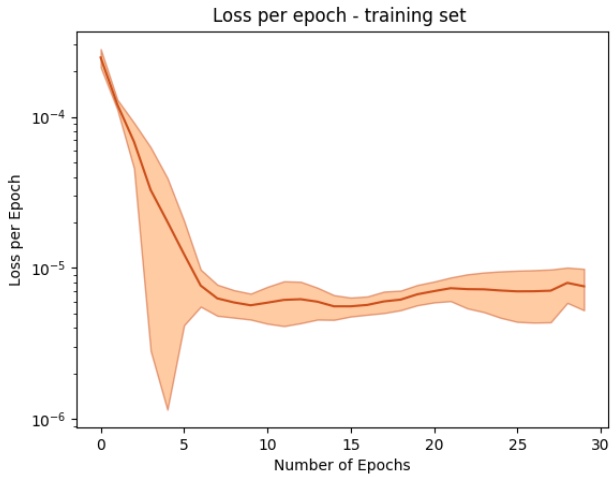
Bhabha s-channel low-energy limit:

$$\overline{|\boldsymbol{M}|^2} = e^4\left[1 + \frac{p^4\cos^2(\theta)}{(p^2+m^2)^2} + \frac{m^2}{2(p^2+m^2)}\right]$$

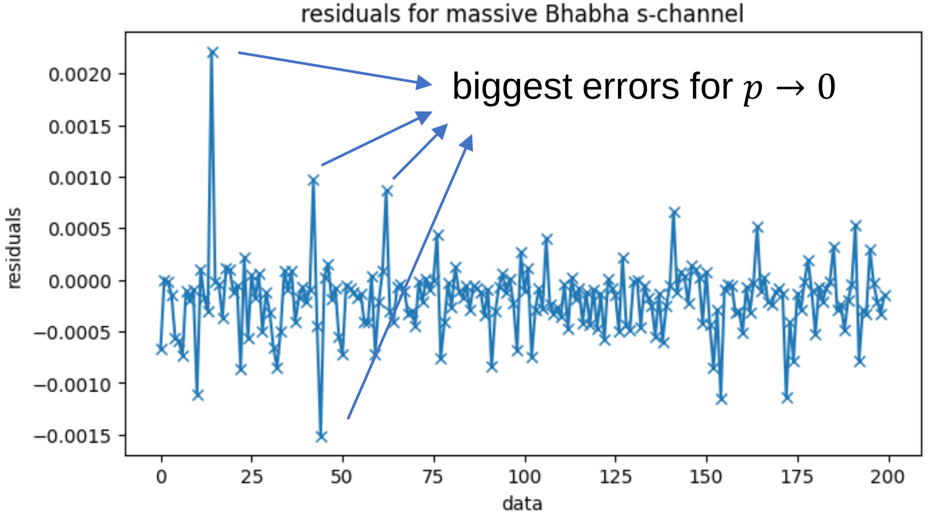
- $p \in [0.0, 5.0]\text{MeV}, \;\; \theta \in [0.5, \pi]$

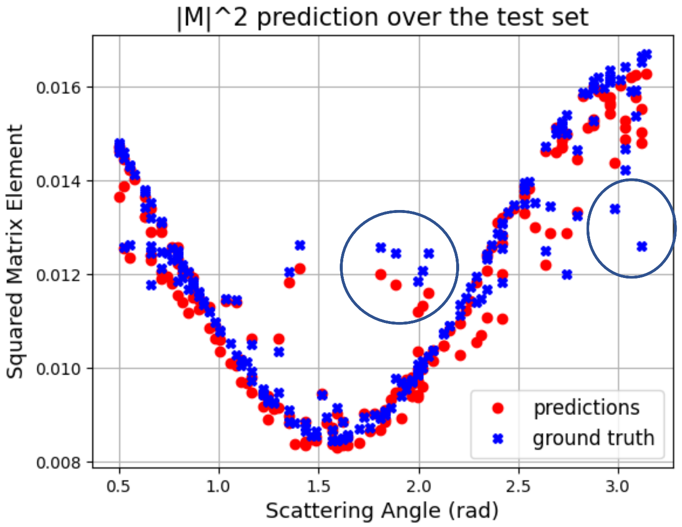
Training MSE	Validation MSE
$(3.781 \pm 1.154) \times 10^{-7}$	$(2.646 \pm 1.154) \times 10^{-7}$

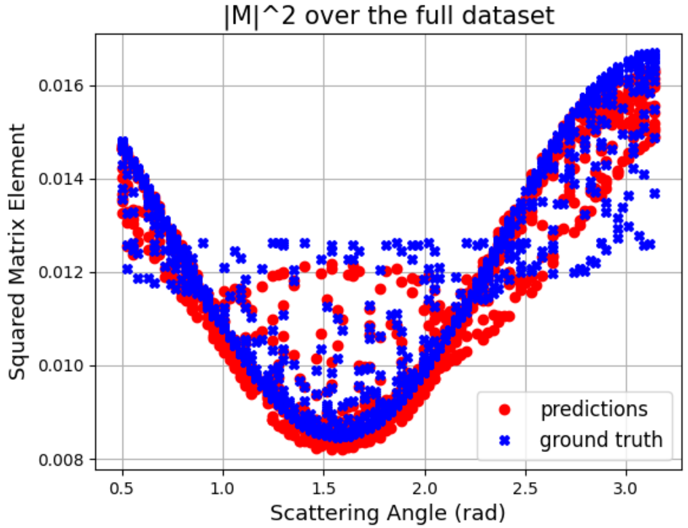




- **MSE** after test evaluation: 1.686×10^{-7}







QGNN for multiple Feynman diagrams

Bhabha s-channel and Bhabha t-channel in high energy limit are put together in a single QGNN:

- Data element: $[G_k, \theta_i, p_i]$, where $G_k = G_s \vee G_t$
- Study the **discriminative power** of the circuit for two **different graphs**.

The two graphs are very similar, distiguish them is a complex task.

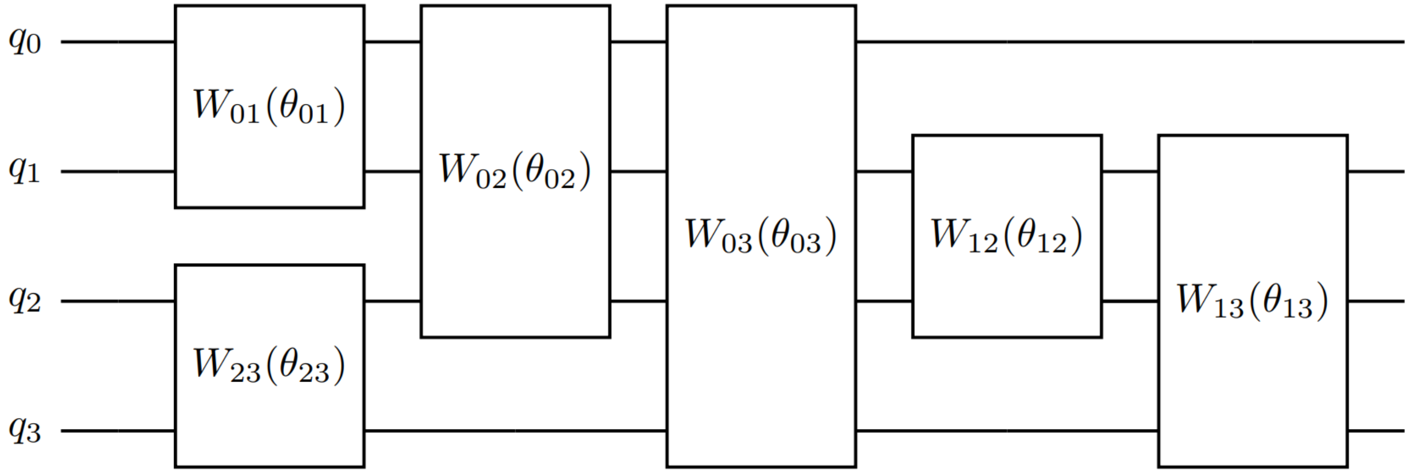
In order to augment the **expressibility** of the circuit the ansatz is changed:

- Fully connected Ansatz (full connection between all qubits)

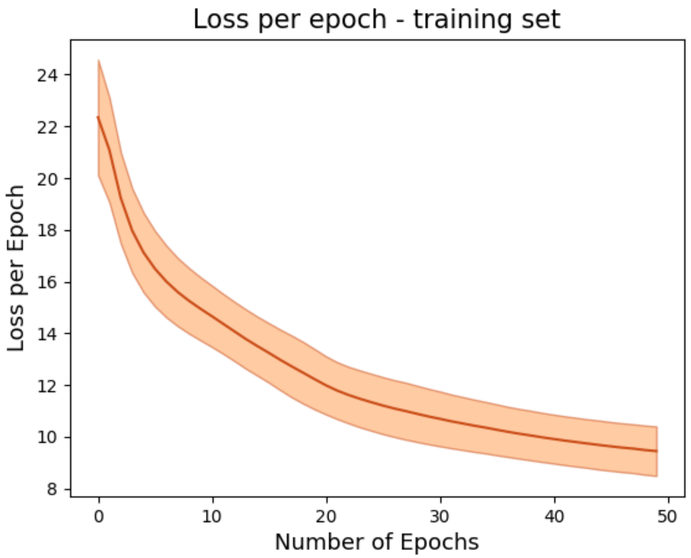
all-to-all **Graph Transformer attention** and **message passing** mechanisms.

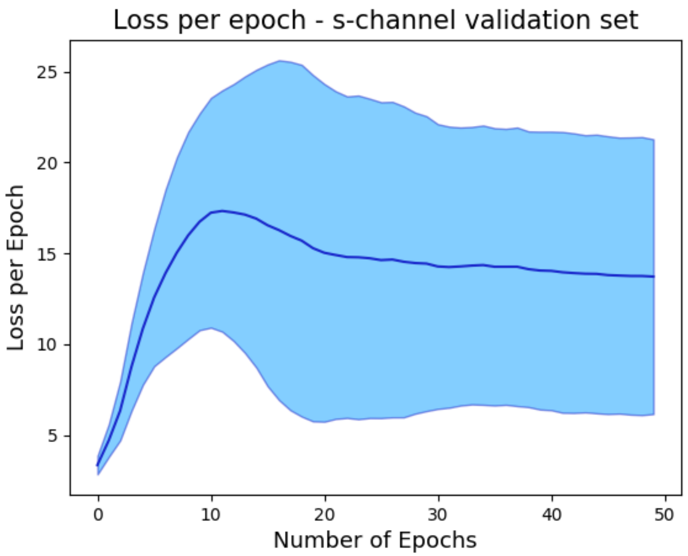
S-channel	T-channel
[(0,2), (1,2), (2,3), (3,4), (3,5)]	[(0,2), (1,3), (2,3), (2,4), (3,5)]

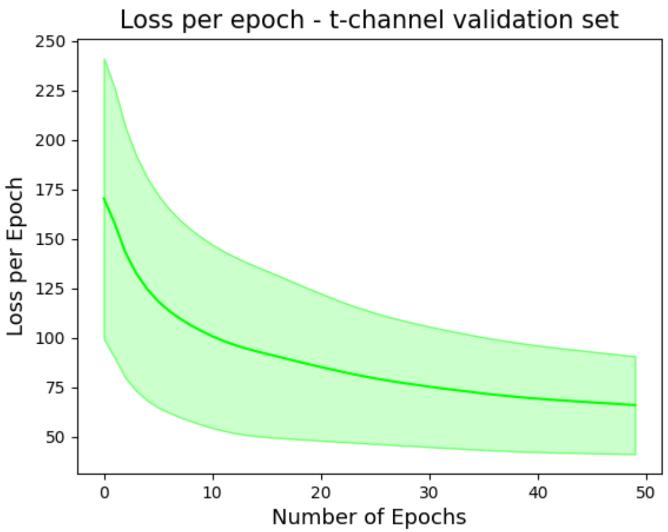
Edge lists of the two channels



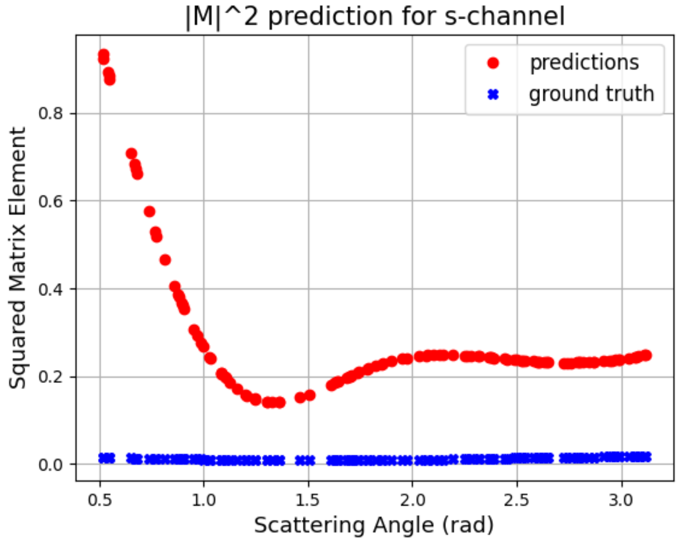
Performance for multiple Feynman diagrams II

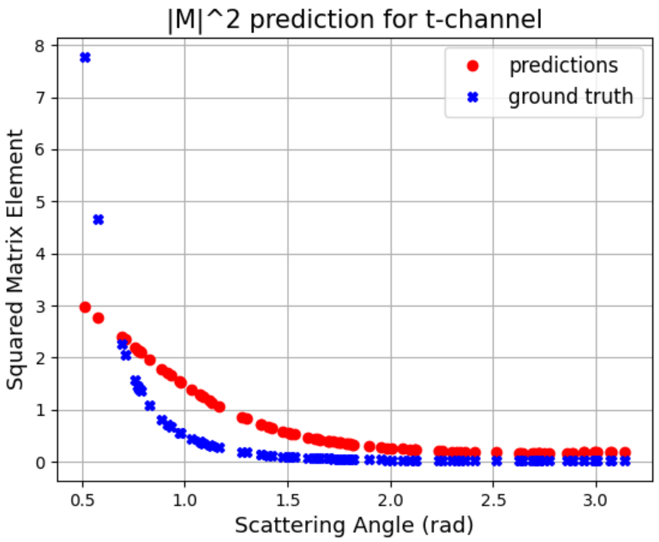






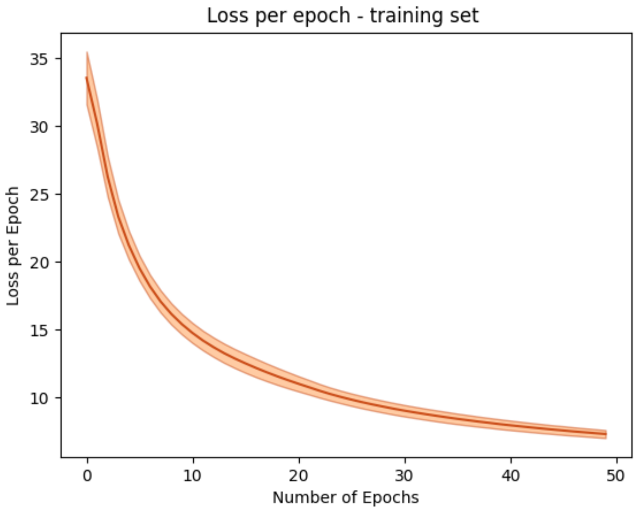
MSE training	MSE s-validation	MSE t-validation
0.472 ± 0.047	0.121 ± 0.067	0.758 ± 0.284

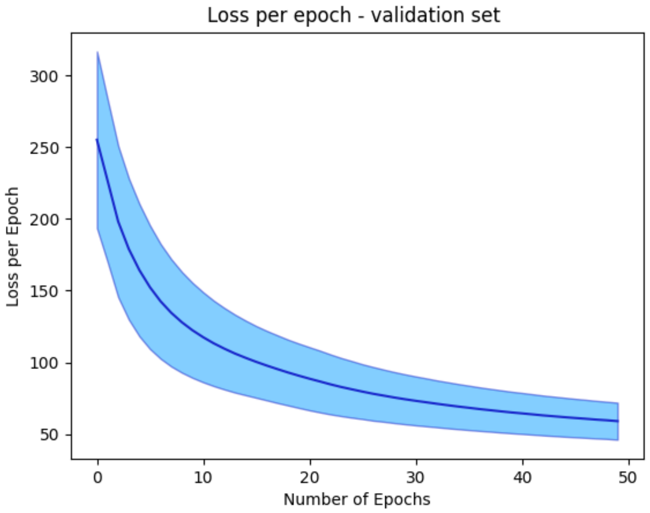


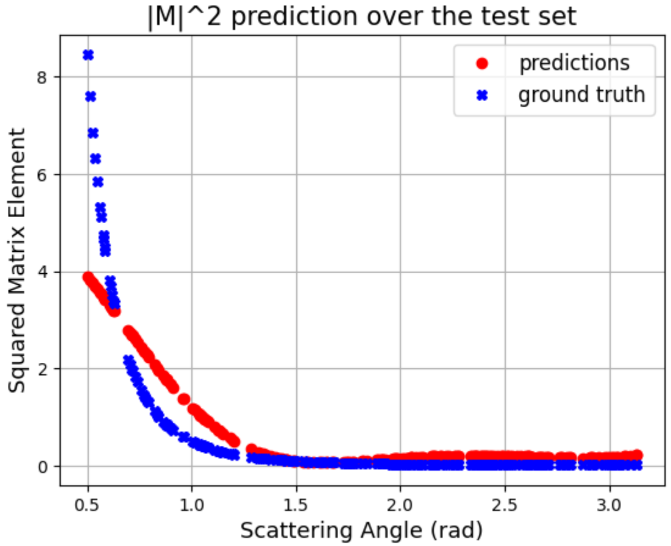


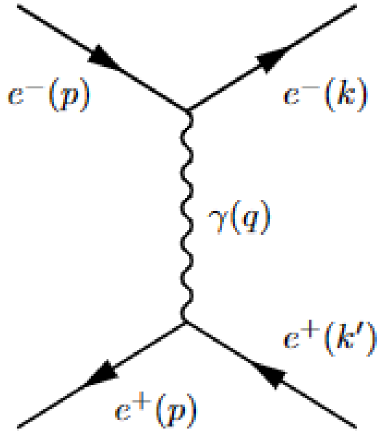
- **MSE s-test** set: 0.104
- **MSE t-test** set: 0.527

Performance for Bhabha t-channel









Bhabha t-channel:

$$e^+e^- \to e^+e^-$$

- high energy limit: electrons are massless

$$\overline{|\boldsymbol{M}|^2} = \frac{2e^4}{(1-\cos(\theta))^2}\left[4 + (1+\cos(\theta))^2\right]$$

Divergence for $\theta = 0$ (cutoff for $\overline{\boldsymbol{\theta}} = \mathbf{0.5}\, \boldsymbol{rad}$)

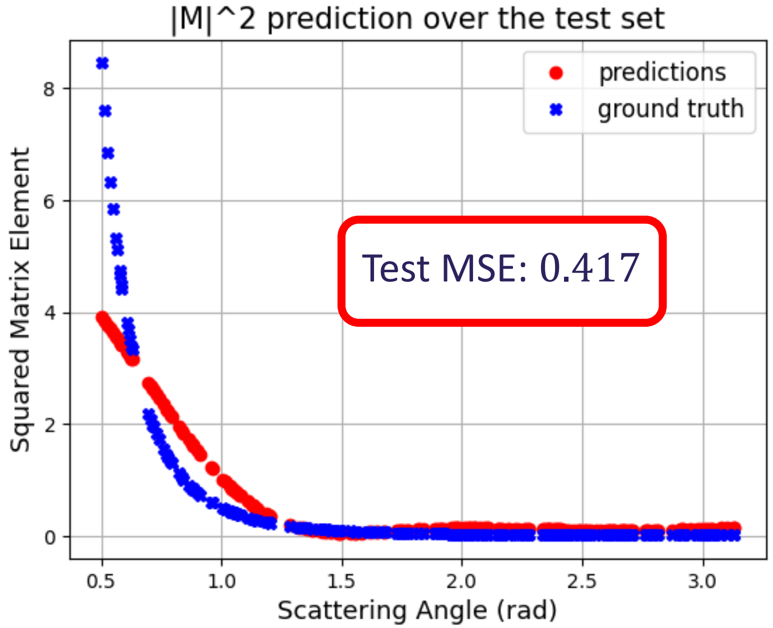
- Architecture: 3-layers QGNN

Training MSE	Validation MSE
0.365 ± 0.015	0.369 ± 0.080

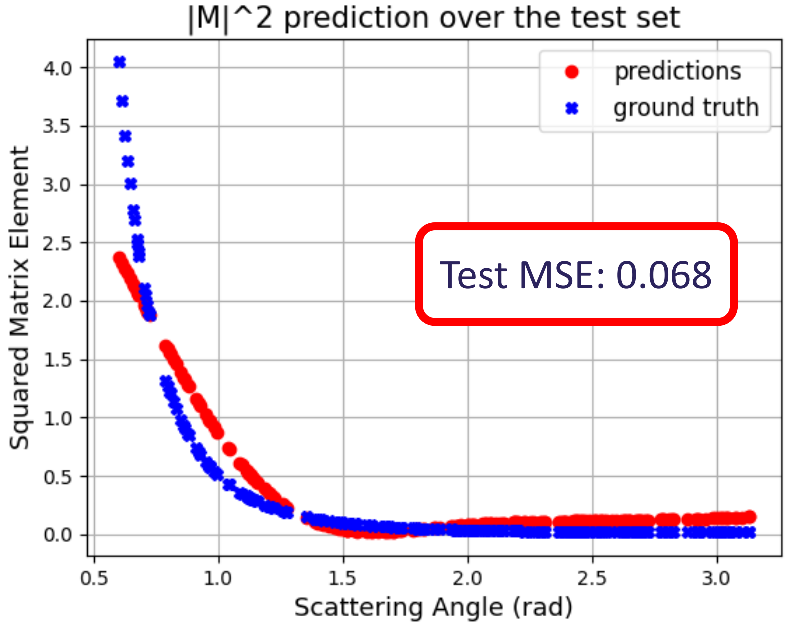
MSE after test evaluation: 0.461

Improvement of the model II

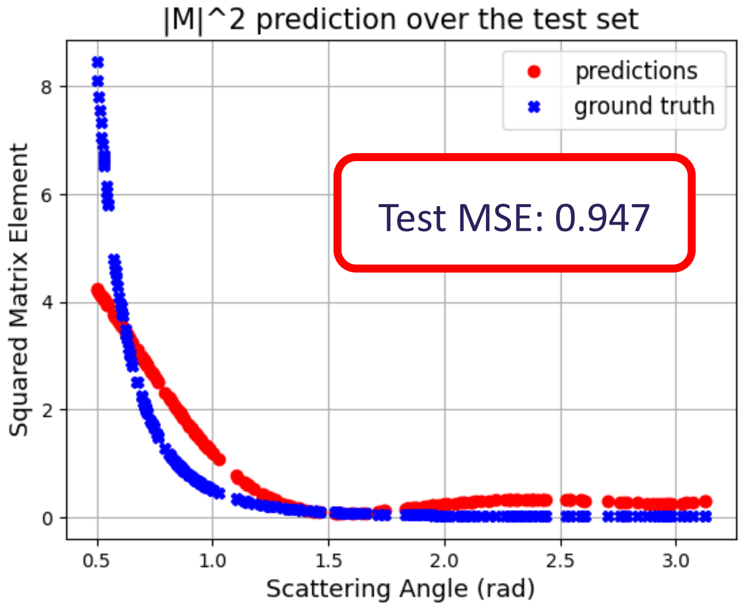
- Performance for a **5-layers QGNN**:



- Performance for a 3-layer QGNN with **different cutoff angle** ($\bar{\theta} = 0.5 \rightarrow \bar{\theta}' = 0.6$):



- Performance for a 3-layer QGNN with denser datapoints in the asymptotical region:



The performance does not improve with these different methods, we can conclude that the QGNN has saturated is expressive power; the asymptotical region is basically predicted in the same way.

Conclusions and Future Directions

Conclusions and Future Directions

Quantum Graph Neural Network:

- **Strong performance** of the QGNN for Bhabha s-channel, both in **high** (MSE $\sim 1.8 \times 10^{-9}$) and **low energy limit** (MSE $\sim 1.7 \times 10^{-7}$):
  - ➢ Slight decrease of performance in the low energy case, **difficulties for $\boldsymbol{p \to 0}$**.
- Fails in **distinguishing different topologies**:
  - ➢ Harnessing of a Quantum Graph Transformer with all-to-all qubit connections.
- Fails in the case of $\overline{|M|}^2$ with **vertical asymptote** (MSE ~ 0.461)
  - ➢ it seems not to significantly benefit from methods that balance the divergence (e.g. change of the cutoff angle, denser sampling, deeper QGNN).

Future directions:

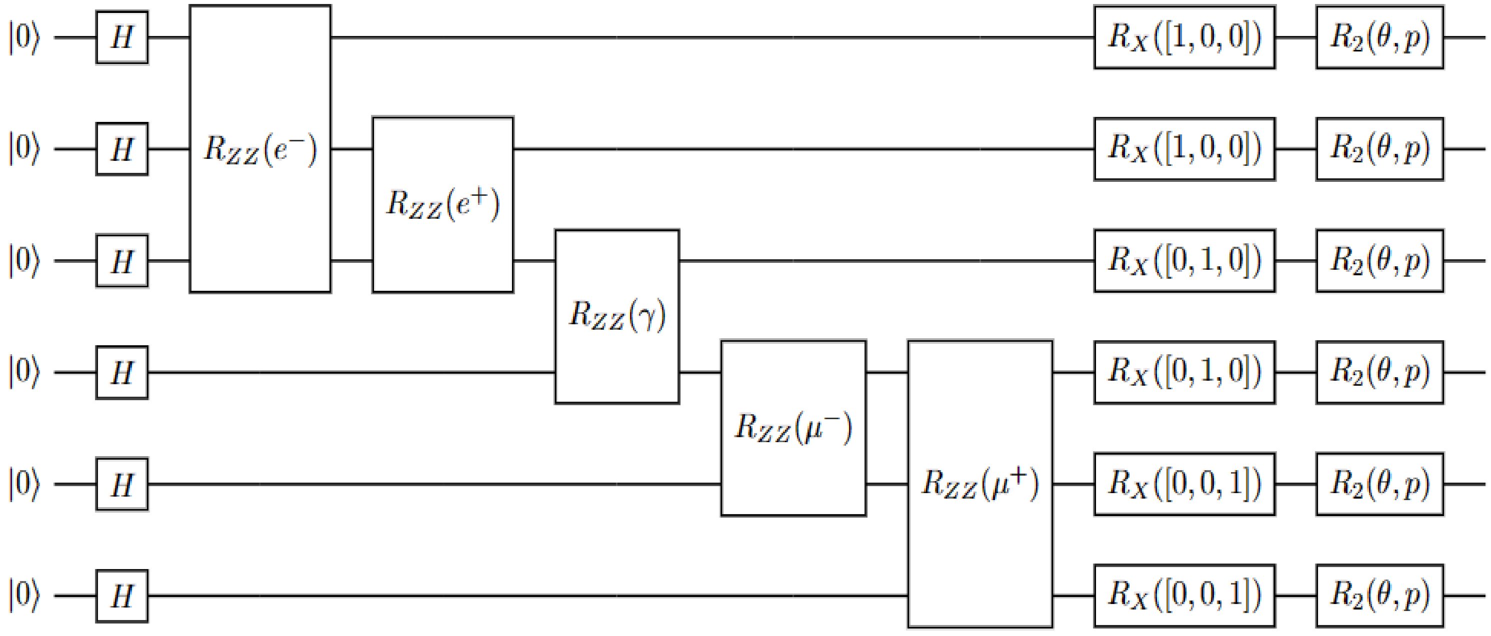
Exploit quantum computing properties in order to **compute for free interference term** between different Feynman diagrams (**can not be done in ML**):

- ➢ There is a **phase ambiguity** due to the way the QGNN computes $\overline{|M|}^2$, the model is not suitable for this application.

Backups

Feature Map II

Feature Map circuit $U(G, \theta, p, \alpha)$:



- $H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ Hadamard gate
- $R_{ZZ}(\phi) = \exp\left(-i\frac{\phi}{2} Z \otimes Z\right)$ Ising ZZ coupling gate
- $R_X(\theta) = \exp\left(-i\frac{\theta}{2} X\right)$ single qubit rotation along X-axis
- $R_2(\gamma, \rho) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & e^{-i\rho} \\ e^{i\gamma} & e^{i(\rho+\gamma)} \end{bmatrix}$ for massive fermions
- $R_1(\tau) = e^{i\tau} R_Z(\tau) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\tau} \end{bmatrix}$ for massles fermions

Improvement of the model I

How to deal with the vertical asymptote of $\overline{|M|_t}^2$?

- Try to enlarge the circuit: 5-layers QGNN

Training MSE	Validation MSE
0.322 ± 0.015	0.321 ± 0.075

MSE after test evaluation: 0.417

- Take more points on the ansatz (similar approach as adaptive quadrature)

MSE during training	MSE during validation
0.811 ± 0.038	0.826 ± 0.155

MSE after test evaluation: 0.947

- Put a regulator in the denominator:

$$\frac{1}{(1-\cos(\theta))^2} \to \frac{1}{(1-\cos(\theta))^2+\varepsilon^2} \quad \text{with } \varepsilon \to 0$$

Even with relatively small ε the shape of $\overline{|M|}^2$ changes drastically

- Change the cutoff angle $\bar{\theta} \to \bar{\theta}'$

Training MSE	Validation MSE
$(0.558 \pm 0.020) \times 10^{-1}$	$(0.564 \pm 0.130) \times 10^{-1}$

MSE after test evaluation: 0.068

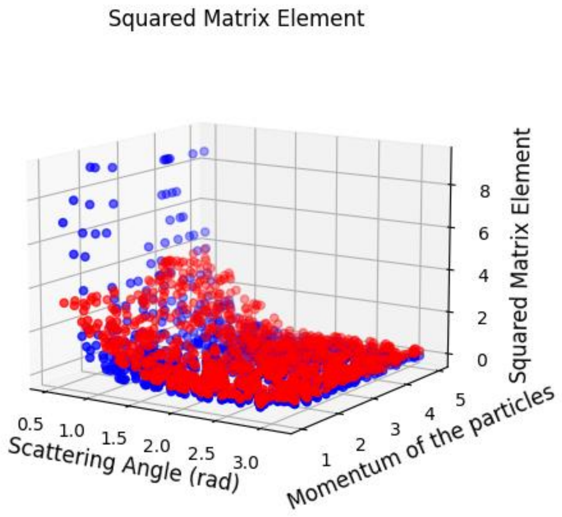
Performance for single Feynman diagram

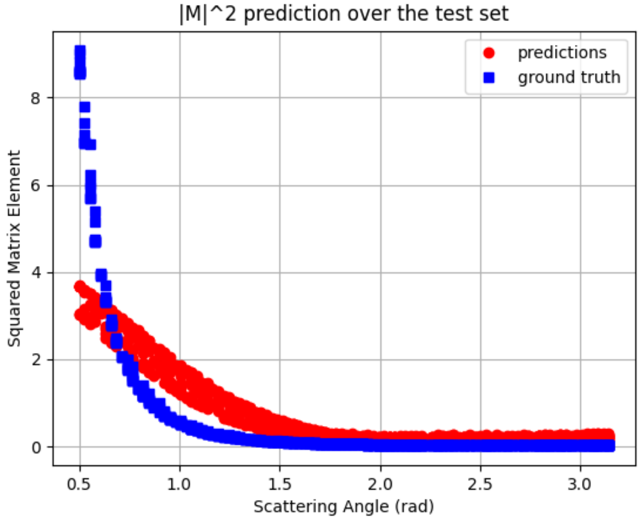
- Architecture: 3-layers QGNN:

MSE during training	MSE during validation
0.783 ± 0.037	0.771 ± 0.140

- Architecture: 5-layers QGNN:

MSE during training	MSE during validation
0.681 ± 0.036	0.654 ± 0.113





Bhabha t-channel:

$$e^+e^- \rightarrow e^+e^-$$

- low energy limit: electrons are massive

$$\overline{|M|^2} = \frac{2e^4}{p^4(1-\cos(\theta))^2}\left[(p^2(1+\cos(\theta))+2m^2)^2\right.$$

MSE after test evaluation:

- 3-layers= 1.167
- 5-layers = 1.016

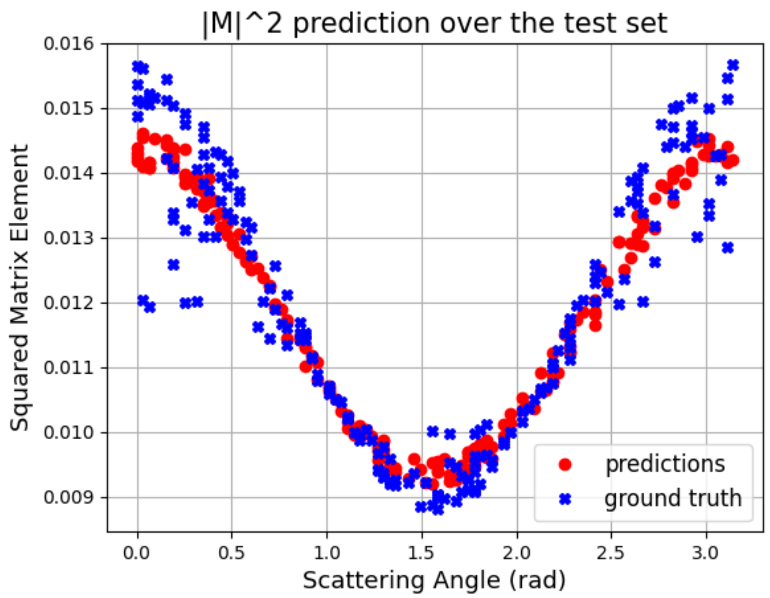
μ-pair Production

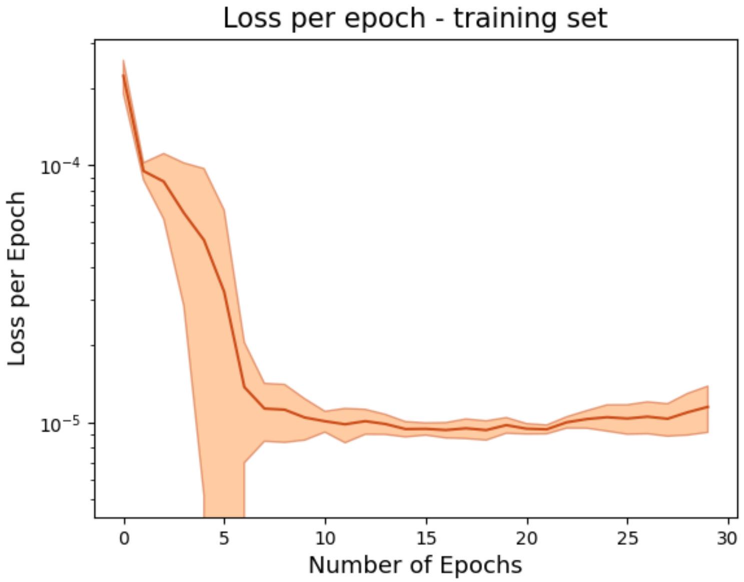
Bhabha s-channel low-energy limit:

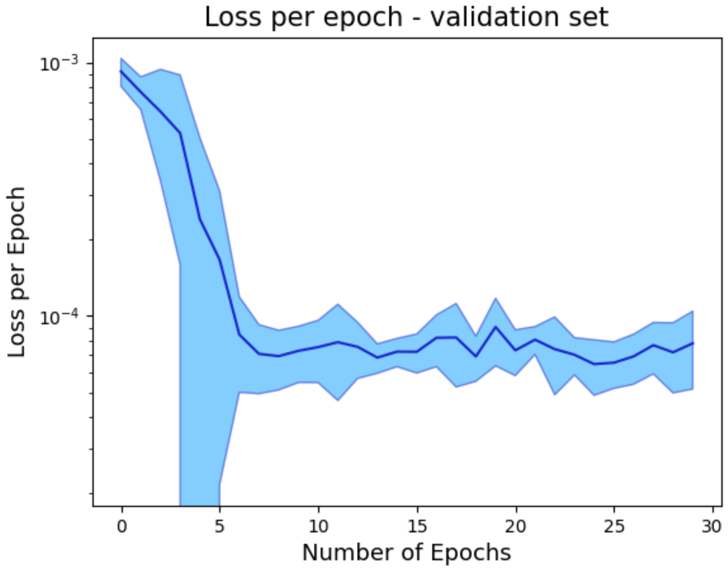
$$\overline{|\boldsymbol{M}|^2} = e^4 \left[1 + p^2 \cos^2(\theta) \left(\frac{1}{p^2 + m_e^2} - \frac{M_\mu^2}{(p^2 + m_e^2)^2} \right) + \frac{m_e^2 + M_\mu^2}{4(p^2 + m_e^2)} \right]$$

This process has an energy threshold: $E_{tot} \geq 2M_\mu \Rightarrow p \geq \sqrt{M_\mu^2 - m_e^2}$

- $p \in [120\,,\ 250]$ MeV, $\theta \in [0, \pi]$







Training MSE	Validation MSE
$(5.745 \pm 1.173) \times 10^{-7}$	$(4.871 \pm 1.654) \times 10^{-7}$

- MSE after test evaluation: 4.331×10^{-7}

The Impossibility for an Interference Protocol

Phase Ambiguity

Exploit Quantum Machine Learning properties to predict the interference term between two or more Feynman diagrams.

Interference: $2Re(M_1^* M_2) \equiv M_1^* M_2 + M_1 M_2^* \qquad M_1, M_2 \in \mathbb{C}$

Take 2 pre-trained QGNNs for two Feynman diagrams $[G_1, G_2, \theta, p]$:

$$\overline{|M_1|}^2(G_1,\theta,p) \equiv \langle O_1 \rangle = |\alpha(G_1,\theta,p)|^2 a^2 + |\beta(G_1,\theta,p)|^2 b^2$$

$$\overline{|M_2|}^2(G_2,\theta,p) \equiv \langle O_2 \rangle = |\gamma(G_2,\theta,p)|^2 c^2 + |\delta(G_2,\theta,p)|^2 d^2$$

$$\begin{cases} Re(M_1) = \sqrt{|\alpha|^2 a^2 + |\beta|^2 b^2}\cos(\varphi) \equiv \overline{|M_1|}\cos(\varphi) \\ Im(M_1) = \sqrt{|\alpha|^2 a^2 + |\beta|^2 b^2}\sin(\varphi) \equiv \overline{|M_1|}\sin(\varphi) \end{cases}$$

$$\begin{cases} Re(M_2) = \sqrt{|\gamma|^2 c^2 + |\delta|^2 d^2}\cos(\rho) \equiv \overline{|M_2|}\cos(\rho) \\ Im(M_2) = \sqrt{|\gamma|^2 c^2 + |\delta|^2 d^2}\sin(\rho) \equiv \overline{|M_2|}\sin(\rho) \end{cases}$$

$$2Re(M_1 M_2^*) = 2\big(Re(M_1)Re(M_2) + Im(M_1)Im(M_2)\big) = \overline{|M_1||M_2|}\cos(\varphi - \rho)$$

It's necessary to tune for each couple of Feynman diagrams $[G_1, G_2, \theta, p]$ the angle $\tau = \varphi - \rho$.