

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Faculty of Physics
AI in Physics Research Group

PP

Group A

Vyron Arvanitis
Emanuele Poggi
Paul Hofmann
Moritz Heidtmann
Angelos Aslanidis

Supervisor

Nikolai Krug

July 12, 2025

Contents

1	Introduction	2
2	The Belle experiment	2
2.1	Problem Description	4
3	Theory	5
3.1	Deep Sets	5
3.2	Graph Neural Networks	5
4	Methodology	6
5	Results	6
6	Conclusion	6

1 Introduction

In this lab course, we work with simulated data from the Belle II experiment to develop a fast and efficient simulation pipeline. The goal is to train and compare machine learning models that can predict early in the simulation chain whether an event is likely to survive later selection stages. This allows us to discard unimportant background events early, saving computational resources typically spent on full detector simulation and reconstruction (processes which are far more computationally expensive). The strategy is illustrated in Fig. 1, where a neural network acts as a pre-filter to speed up the overall simulation process.

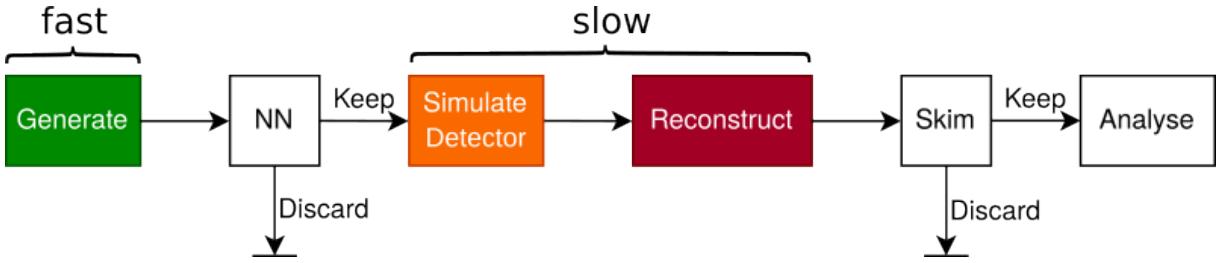


Figure 1: Smart simulation strategy: a neural network filters generated events before expensive detector simulation. Figure adapted from the official LMU AI Lab repository [3].

2 The Belle experiment

The Belle experiment is a so-called *B-Factory*, a type of particle accelerator specifically designed to produce large numbers of *B mesons*. These facilities operate with a center-of-mass energy tuned to the $\Upsilon(4S)$ resonance, a state that decays almost exclusively into pairs of *B* and \bar{B} mesons. This selective production mechanism enables detailed studies of *B* meson decays, allowing physicists to investigate phenomena such as *CP violation*, rare decays, and physics beyond the Standard Model. The Belle II experiment, located at the *SuperKEKB collider in Tsukuba, Japan*, is the successor to the original Belle experiment and significantly improves on its precision and data collection capabilities.

Unlike proton-proton colliders (such as the LHC), Belle II uses electron-positron (e^+e^-) collisions. Electrons and positrons are elementary particles with no internal structure, which offers several advantages. First, their interactions are much cleaner, and the background is easier to interpret—there is a notable absence of complex hadronic activity such as jets. Moreover, the absence of internal substructure allows us to precisely know the initial momenta of the colliding particles, enabling accurate reconstruction of the event kinematics. This precision is crucial for tuning the aforementioned center-of-mass energy to specific resonances, such as the $\Upsilon(4S)$.

B mesons are unstable particles that decay via the weak interaction, and their decays can be classified into three categories: hadronic, semileptonic, and rare decays. In *hadronic decays*, the *B* meson decays into only hadrons, such as combinations of kaons (K), pions (π), and other mesons. *Semileptonic decays* involve both hadrons and leptons, typically resulting in a final state containing a charged lepton (e^\pm or μ^\pm), a neutrino, and hadrons. *Rare decays* may involve photons, multiple leptons, or other suppressed processes that are especially sensitive to

effects beyond the Standard Model. These decay modes provide a rich ground for studying CP violation, flavor physics, and potential new physics.

The Belle II detector is structured as a layered system of subdetectors (like an onion) built around the interaction point. These subdetectors are designed to detect and identify the various particles produced in the collisions. Closest to the interaction region is the Vertex Detector (VXD), composed of the PiXel Detector (PXD) and Silicon Vertex Detector (SVD), which provide precise tracking information for reconstructing decay vertices. Surrounding the VXD is the Central Drift Chamber (CDC), responsible for tracking charged particles and measuring their momenta and energy loss. Particle identification is handled by the Time Of Propagation (TOP) detector in the barrel region and the Aerogel Ring Imaging Cherenkov (ARICH) detector in the forward region. Photons and electrons are measured in the Electromagnetic Calorimeter (ECL) via the production of showers, while muons and long-lived neutral kaons (K_L^0) are identified in the outermost KLong and Muon (KLM) detector. The entire setup is enclosed in a 1.5 T magnetic field provided by a superconducting solenoid, which not only aligns the beams but also enables accurate momentum measurements. An overview of the collider (figure 2) and the particles that can be identified in each subdetector can be seen in table 1

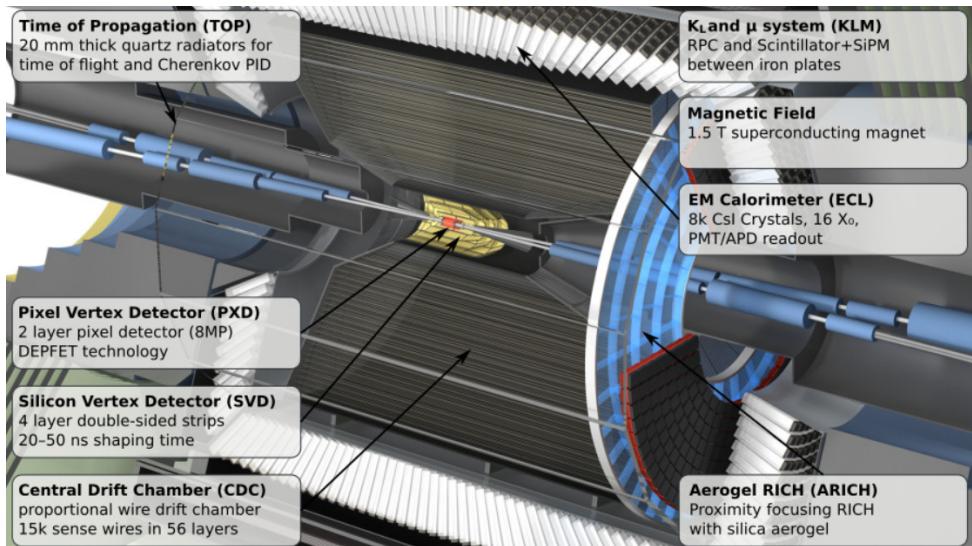


Figure 2: Closeup of the Belle II detector indicating all the different subdetectors [1].

Subdetector	Main Purpose
PXD + SVD (VXD)	Track reconstruction near the IP, measures short-lived particles
CDC	Tracking of charged particles
TOP / ARICH	Particle identification (e.g. distinguish π , K , p) via Cherenkov radiation
ECL	Detection of electrons and photons (electromagnetic showers)
KLM	μ identification and detection of K_L^0

Table 1: Belle II subdetectors and the particles they primarily detect.

2.1 Problem Description

The Belle II experiment collects an enormous amount of data, with its center-of-mass energy tuned to the $\Upsilon(4S)$ resonance *corresponding to a center-of-mass energy of $\sqrt{s} = 10.58 \text{ GeV}$.* While these "resonant" events are the primary source for B physics, Belle II also records other types of data.

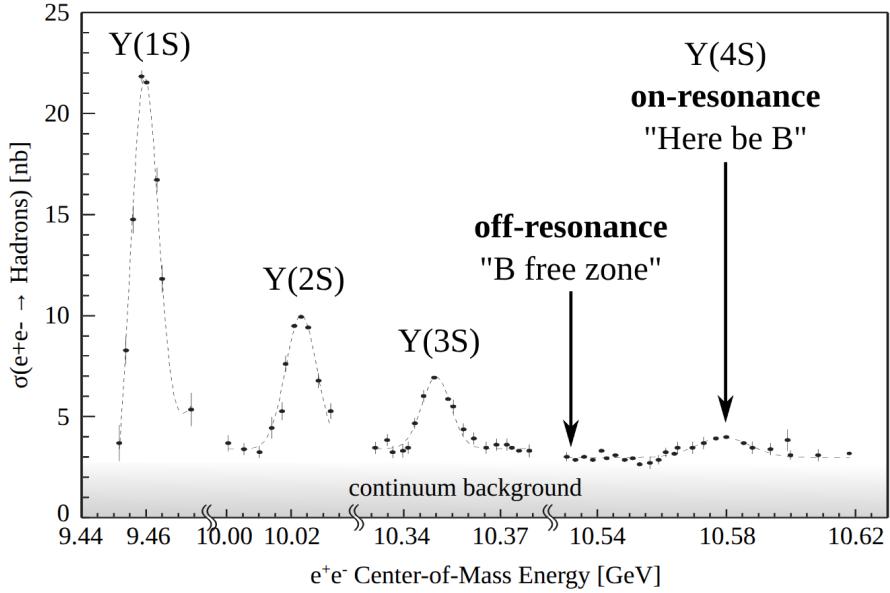


Figure 3: Cross section of e^+e^- with respect to the center-of-mass energy [GeV] [2].

To reduce the data volume and focus on interesting events, Belle II employs a two-level filtering system: the Level-1 hardware trigger (TRG) and the High Level Trigger (HLT), which together reduce the event rate before data is stored. Even after this, the number of events retained is still too large for direct analysis, so further selection — known as skimming — is performed to extract smaller datasets.

In parallel, Monte Carlo (MC) simulations are used to model the detector response and to train and validate analysis strategies. These simulated events must pass through all stages of detector simulation, digitization, reconstruction, and selection, mirroring the real data processing pipeline. However, this process is computationally expensive, especially when simulating large background samples, most of which are eventually discarded by skims.

To address this issue, the goal of this lab course is to explore a technique known as *Smart Background Simulation*. Instead of fully simulating every event, we aim to train a neural network that can predict early on whether a given event will survive a later skim. This skim selects events in which at least one B^0 meson can be reconstructed through hadronic decays. After building several neural networks we evaluate their performance on an independent test set and calculate the speedup offered by each network.

3 Theory

3.1 Deep Sets

In many physical systems, such as collections of particles or events in high-energy physics, the data can be naturally represented as unordered sets. Traditional neural networks, however, are not inherently permutation-invariant. This means that they assume an order in their input. To address this, Deep Sets provide a principled way to process such set-structured data while maintaining invariance to the order of elements. Figure 4 visualizes the structure of a Deep Set model.

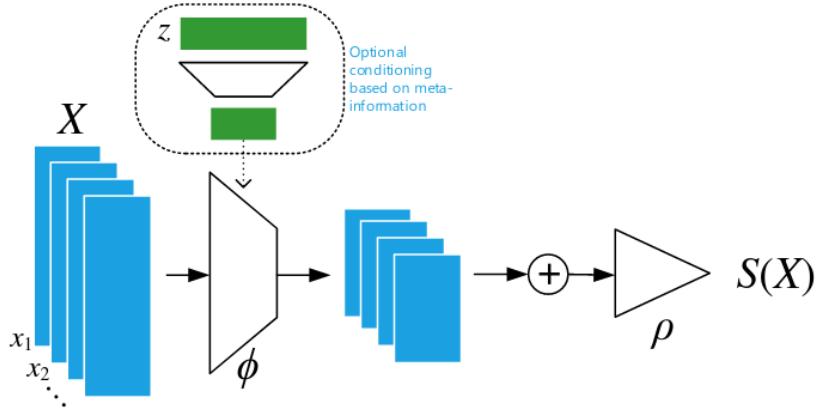


Figure 4: Schematic illustration of a Deep Set model.

A Deep Set model operates by applying a shared function ϕ independently to each element in the set. This produces a new set of embeddings, which are then aggregated using a permutation-invariant operation such as sum, mean, or max. The aggregated result is passed through a second function ρ , typically implemented as a multilayer perceptron (MLP), to produce the final output:

$$f(X) = \rho \left(\sum_{x \in X} \phi(x) \right) \quad (1)$$

This structure ensures that the model respects the symmetry of sets. The theoretical foundation of Deep Sets is discussed in Zaheer et al., 2017, where it is shown that any function on a set that is invariant to permutations can be decomposed in this form.

In our project, we use this architecture to model collections of particles, where the order of particles is physically irrelevant. The Deep Set model thus provides a natural and efficient baseline for learning from such data.

3.2 Graph Neural Networks

In contrast to Deep Set models, which treat the input as an unordered collection, Graph Neural Networks (GNNs) explicitly leverage relational information between data points. This is partic-

ularly important in our context, where particles are connected through a decay tree structure that naturally forms a graph. Each node in the graph represents a particle, and edges represent physical relationships (e.g., parent-child connections in the decay chain).

In terms of Deep Sets, graph convolutions can be seen as permutation-equivariant operations. As long as the aggregation function (e.g., sum, mean, or max) is symmetric, the update remains invariant to the ordering of neighbors.

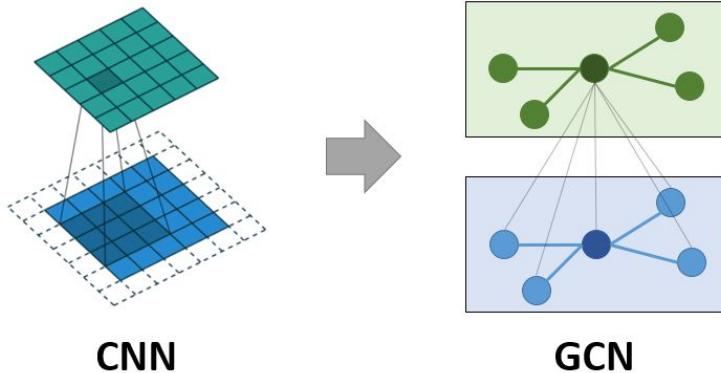


Figure 5: Analogy between CNNs and GCNs. In CNNs, pixels are updated based on spatial neighborhoods; in GCNs, nodes are updated based on graph neighborhoods.

Figure 5 illustrates the analogy between Convolutional Neural Networks (CNNs) and Graph Convolutional Networks (GCNs). In CNNs, pixels are updated based on their spatial neighborhoods, while in GCNs, nodes are updated based on their graph neighborhoods. This allows GNNs to capture complex relationships and dependencies in the data, making them particularly suitable for tasks involving structured data like particle decays.

4 Methodology

5 Results

6 Conclusion

References

- [1] Belle II Collaboration. Belle ii detector overview. https://software.belle2.org/development/sphinx/_images/detector_labeled.jpg, 2023. Accessed: 2025-07-12.
- [2] Belle II Collaboration. Belle ii detector overview. https://software.belle2.org/development/sphinx/_images/Uresonances.svg, 2023. Accessed: 2025-07-12.
- [3] Nikolai Krug. Preparation materials for particle physics lab group a. <https://gitlab.physik.uni-muenchen.de/ai-lab/ss25/pp-lab-group-a/-/blob/main/preparation.md>, 2025. Accessed: July 8, 2025.