Optimization of the LHCb calorimeter

Alexey Boldyrev[1], Denis Derkach, Fedor Ratnikov, Andrey Shevelev
on behalf of the MODE collaboration

[1] *alexey.boldyrev@cern.ch*

Introduction

Calorimeters are a crucial component for most detectors mounted on modern colliders. Their tasks include identifying and measuring the energy of photons and neutral hadrons, recording energetic hadronic jets, and contributing to the identification of electrons, muons, and charged hadrons. To fulfill these many tasks while keeping costs reasonable, the calorimeter construction requires good and thoughtful balancing with other components of the detector.

Much harder operation conditions during LHC's high luminosity Run 5 and beyond (Upgrade II conditions) imply new technological and computational challenges. This requires optimization of technologies, layouts, readouts, reconstruction algorithms to achieve the best overall physics performance for the limited cost.

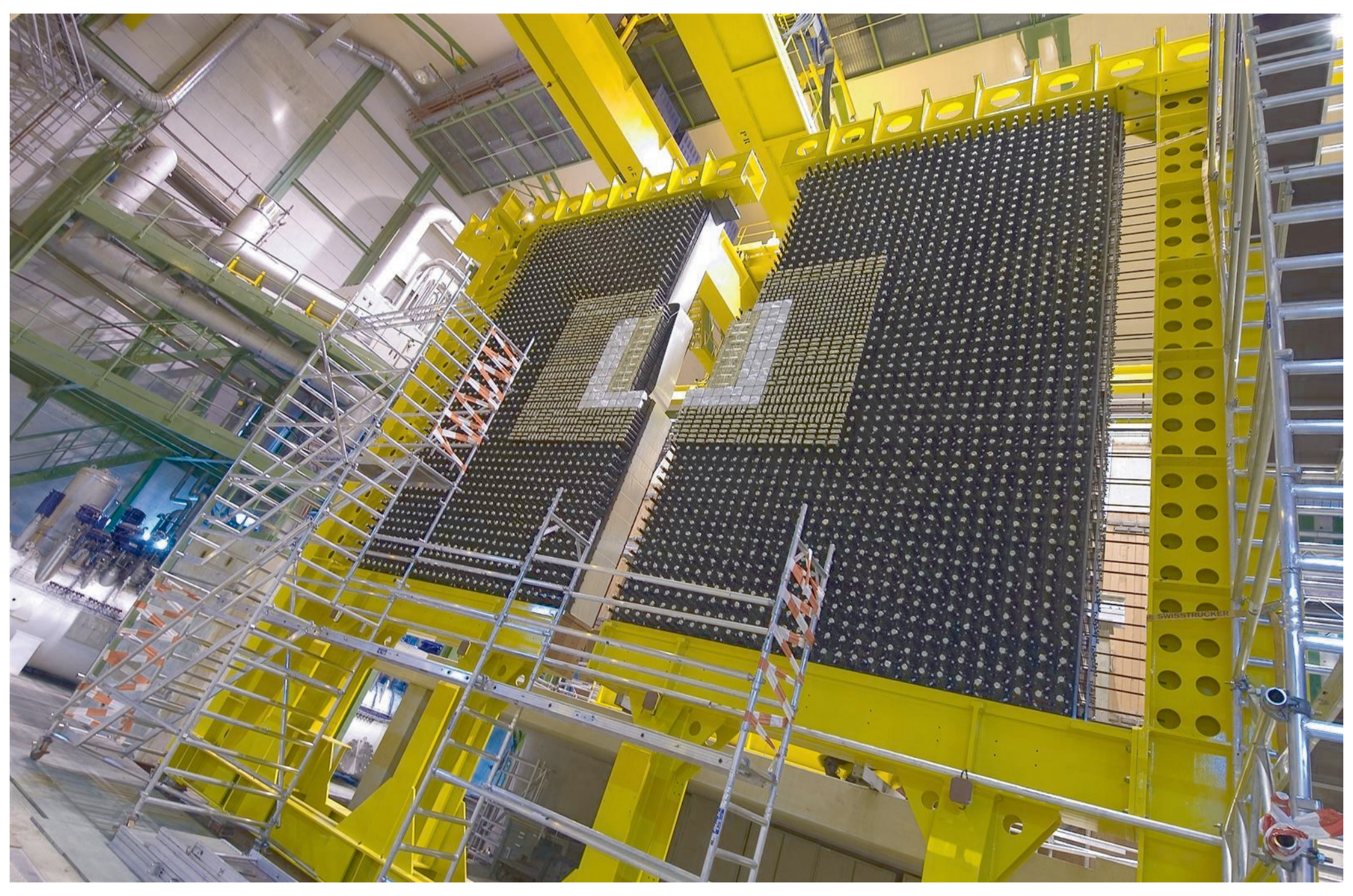
LHCb detector

LHCb is one of four major LHC experiments and provides:

- precise tests for Standard Model verification
- detailed studies of Charm and Beauty physics
- precise measurements of CP violation effects

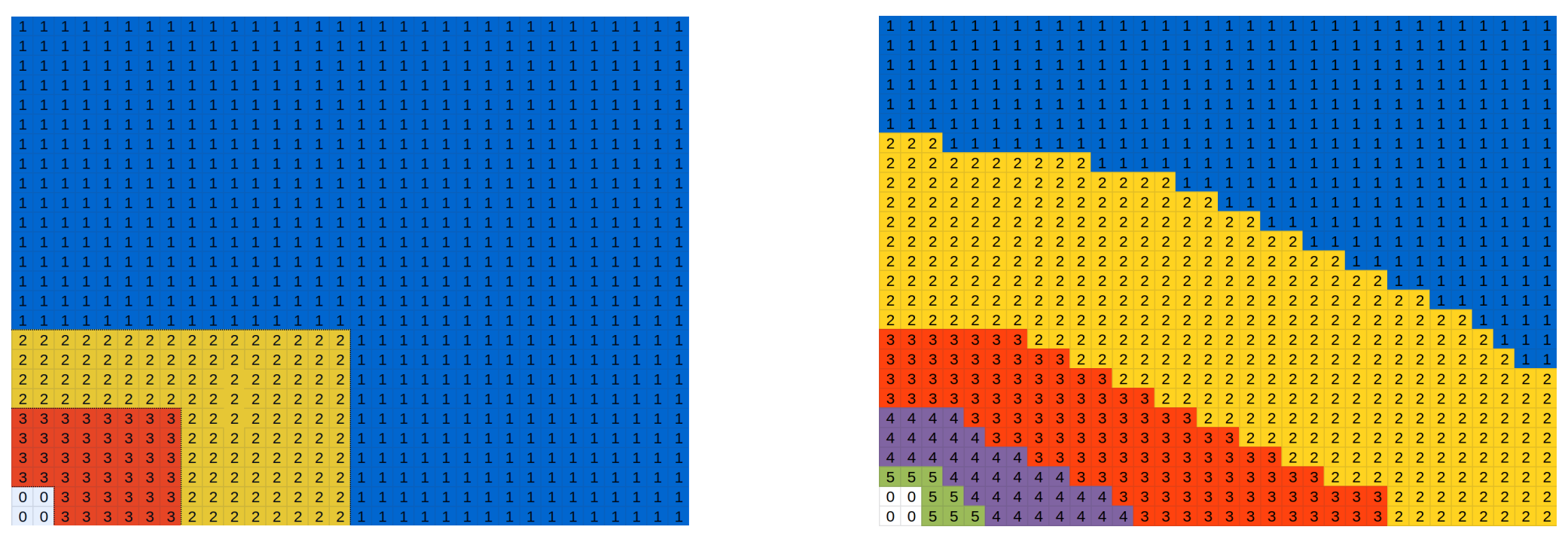
LHCb ECAL

The current ECAL is based on Shashlik-type modules of 3 granularities, and contains 1536/1792/2688 cells in its inner/middle/outer regions, respectively.



LHCb calorimeter wall

Several technological options for the upgraded calorimetry are foreseen. The most severe requirements for radiation tolerance can be met by SpaCal modules consist of longitudinal fibres acting both as scintillator and light-transporting medium.



Baseline configuration (left) and the first level of optimisation of the cell sizes for LHC Upgrade II conditions (right). Upper right quarter of the calorimeter wall is shown.

Figures of merit

The performance of the calorimeter consists of:

- Radiation tolerance to sustain the expected lifetime span;
- Energy and spatial resolution for good photon reconstruction and electron identification;
- High granularity and longitudinal segmentation to facilitate better precision, both spatial and in energy, which in turn improves reconstruction algorithms;
- Timing resolution enough to facilitate pileup suppression in high-occupancy areas as well as better matching of separate signal components.

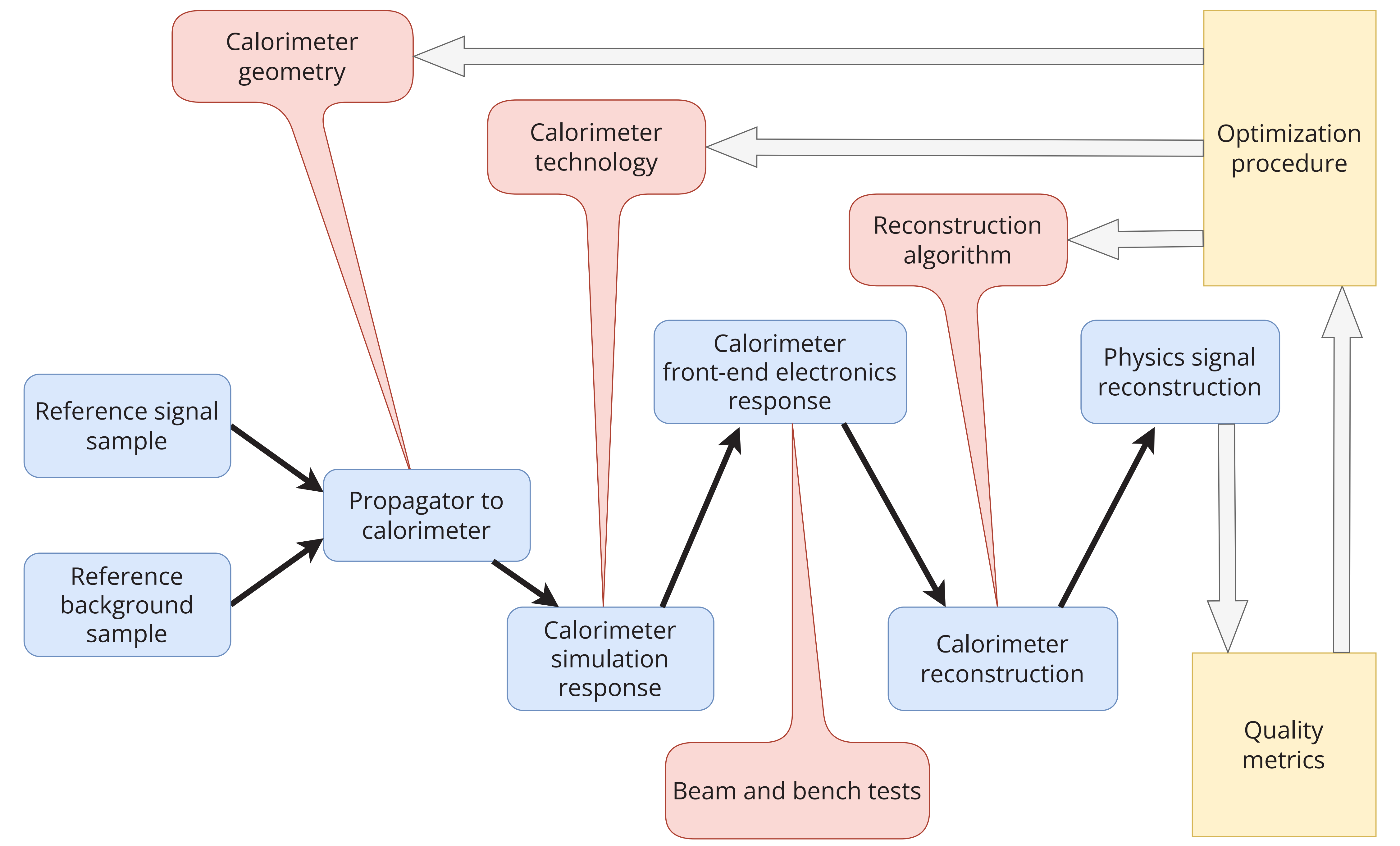
However, the ultimate goal for the optimization process is to achieve the dependency of the physics performance on the cost of the configuration of the detector under study.

Models in the optimization cycle

Model	**Primary implementation**	**Surrogate impl.**
Detector response	GEANT4	GAN
Detector reconstruction	NA	ML regressor
Performance metric aggregator	Physics significance vs. cost	Not needed

Models are required to be differentiable with respect to θ_{model} in vicinity of the current optimization point. For non-differentiable models the differentiable machine learning (ML) based surrogates can be used.

General pipeline of the calorimeter optimization



Calorimeter geometry

Calorimeter technology

Reconstruction algorithm

Optimization procedure

Reference signal sample

Reference background sample

Propagator to calorimeter

Calorimeter front-end electronics response

Physics signal reconstruction

Calorimeter simulation response

Calorimeter reconstruction

Quality metrics

Beam and bench tests

A general pipeline for the calorimeter optimization includes several steps. Blue blocks indicate data processing pipeline steps; pink callouts represent configurations and conditions for pipeline steps; yellow blocks close the optimization loop.

Workflow steps

- Selected event samples, both signal and background, are used to initiate an optimization cycle for comparing the performance in terms of signal recovery and background suppression;
- the calorimeter is usually installed downstream of the detector, so a propagation of events from their origin to the calorimeter is necessary. This step is dependent on the properties of the elements of the detector between origin and calorimeter. Additionally, if the calorimeter detector has a non-homogeneous configuration, the details of the global geometry are to be accounted for in this step for the optimization to be based on physics performance quality metrics;
- The construction technology used for the individual calorimeter modules is a central point for the detector R&D. To evaluate the impact of a choice of construction technology, we need to simulate its effect on observable event characteristics. This is done using response simulation models, typically based on GEANT4. The details of the calorimeter technology used drive such simulation;
- The behaviour of the front-end electronics is another important contribution to the physics quality of the detector. Although the properties driving the behaviour are hard to simulate, good data samples may be obtained from beam or bench tests;
- A reconstruction algorithm is absolutely necessary to evaluate the quality of converting the detector response into the physics objects;
- Physics quality metrics may be calculated using reconstructed objects and it can be used as a target function for the optimization procedure;
- All aspects of the calorimeter may be optimized: the details of the calorimeter technology, the geometrical layout, and the possible reconstruction algorithms.

Evaluation of the physics performance

To evaluate the physics performance of a particular configuration for a possible future calorimeter detector, one needs to run the optimization cycle described above. A good fine-tuning of the individual blocks is important to properly propagate the properties of the configuration under study to the ultimate physics performance. For the regular operation of a stable detector, these blocks are carefully tuned based on the actual detector configuration. In contrast, for the R&D of a new detector, many different possible configurations are studied simultaneously during the optimization stage. Nevertheless, reasonable representations of the simulation and reconstruction steps, which are tuned for each of the configurations studied, are necessary for inferring consistent conclusions about the physics performance of these configurations. This is a time consuming work, if done manually. Fortunately, these studies use well-labelled data sets either from MC simulation or from test beam measurements. Thus surrogate models may be built and trained on labelled data using regular ML approaches. This makes it possible to speed up model building for different pipeline steps. Importantly, such training may be automated and requires minor expert supervision.

Conclusion

- LHCb ECAL optimization is a good use case for the generic problem of comprehensive optimization of the complex physics detector;
- Optimization pipeline is established and used to practically find optimal configurations in reduced phase space of considered parameters to satisfy LHCb experiment needs;
- Black box optimization is straightforward but requires huge resources in case of multidimensional parameter space;
- Differentiable optimization is a promising approach to speed up the optimization cycle. It requires all involved models to be explicitly differentiable otherwise, ML dynamically trained surrogates might be a solution.