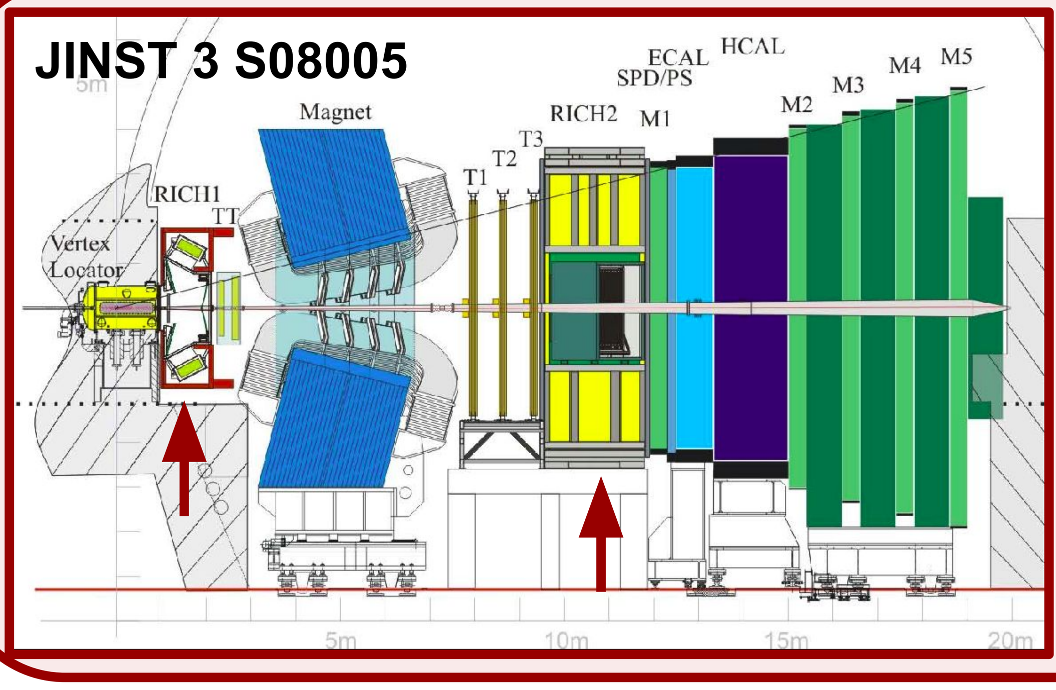
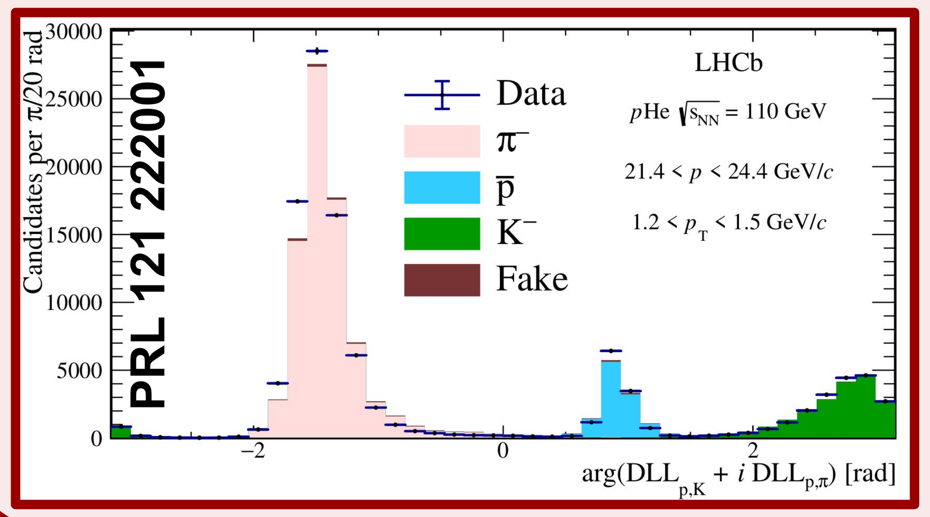
A Neural-Network-defined Gaussian Mixture Model for particle identification applied to the LHCb fixed-target programme[1]

Saverio Mariani, INFN Sezione di Firenze, on behalf of the LHCb collaboration
01/12/2021, 2021 ACAT workshop
saverio.mariani@cern.ch



- **LHCb[2,3] :** spectrometer instrumenting η ϵ [2, 5]
- **Some detectors,** as the **RICH system[4]**, are devoted to **Particle identification (PID)**
- PID classifiers are built as the **log-likelihood difference** between two particle hypotheses (*e.g.* $\mathbf{DLL_{p,\pi}}$ for the p-π separation)

- By injecting noble gases in the LHC beam-pipe, LHCb is performing from 2015 a **unique fixed-target programme** [5, 6] (*p* or Pb beams onto He, Ar, Ne)



- The PID performance affects the measurement of cross-sections, such as $\sigma(p\text{He} \to \bar{p}X, \sqrt{s_{NN}} = 110\,\text{GeV})$[7]:
  - **Limited *p*He data PID calibration statistics**
  - *pp* PID calibration cannot be used because of the **phase-space differences** (higher energy and detector occupancy, lower PVz spread)

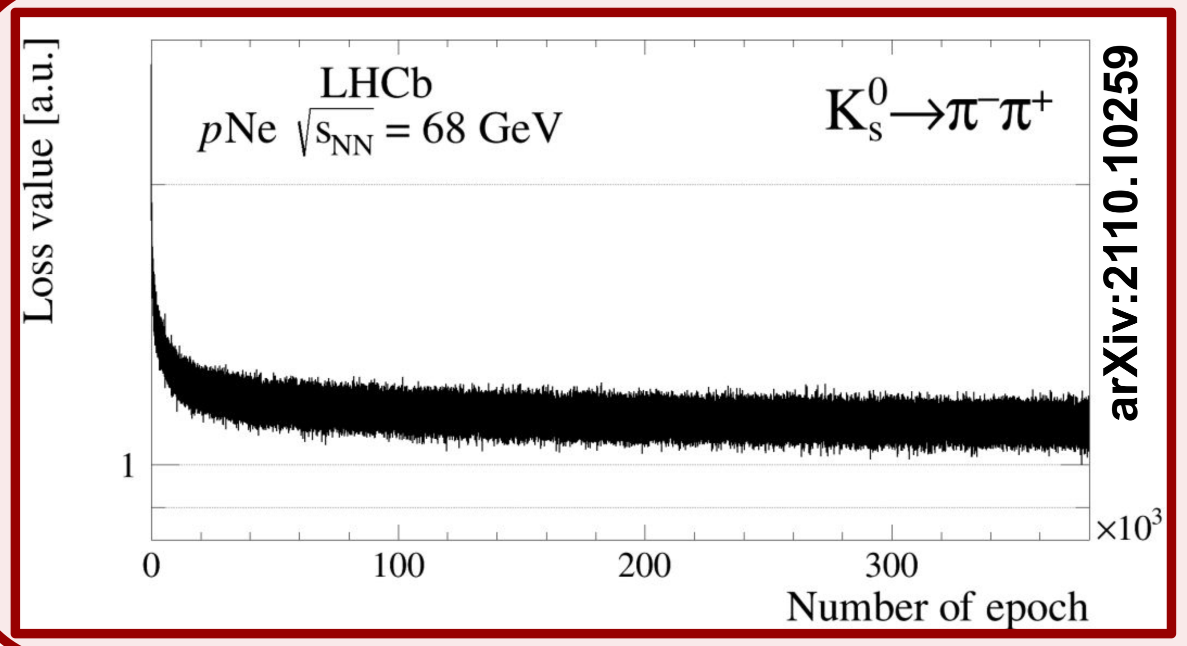
Proposed approach: Model, with **machine-learning** techniques, how the PID classifiers depend on a set of relevant features and **predict their *pdf*** on different channels
Use-case: Train on $K^0_s \to \pi^-\pi^+, \Lambda^0 \to \bar{p}\pi^+$ and $\phi \to K^-K^+$ decays reconstructed and selected in the ***p*Ne** data and apply to smaller-size ***p*He/*p*Ar** samples of different energy

$$\underline{x}_p \sim \sum_{j=1}^{N_{g,p}} \alpha_{j,p}(\underline{\theta}) \frac{\exp(-\frac{1}{2}(\underline{x}_p - \underline{\mu}_{j,p}(\underline{\theta}))^T \Sigma^{-1}_{j,p}(\underline{\theta})\,(\underline{x}_p - \underline{\mu}_{j,p}(\underline{\theta})))}{2\pi\sqrt{\det(\Sigma_{j,p}(\underline{\theta}))}}$$

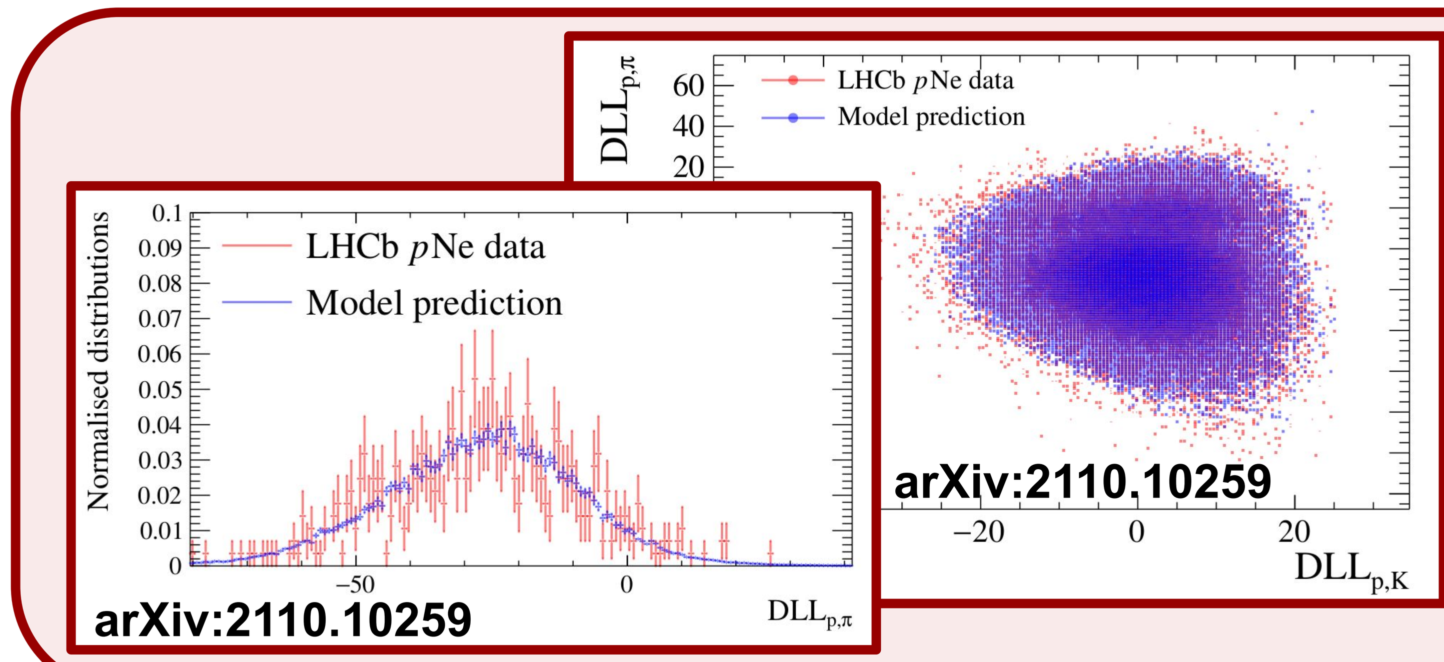
- Bidimensional target $\underline{x}_p$ ($\text{DLL}_{p,\pi}$, $\text{DLL}_{p,K}$) is described as a **Gaussian Mixture Model** (GMM). All parameters of the Gaussian distributions depend on the relevant features $\underline{\theta}$. The $\underline{x}_p(\underline{\theta})$ relation is obtained through **a set of Neural Networks (NNs)** and learned

$$\mathcal{L} = -\sum_{i=1}^{n_p} w_i \log\left[\sum_{j=1}^{N_{g,p}} \alpha_{j,p}(\underline{\theta}_i)\mathcal{G}(\underline{x}_i, \mu_{j,p}(\underline{\theta}_i), \sigma_{j,p}(\underline{\theta}_i))\right]$$

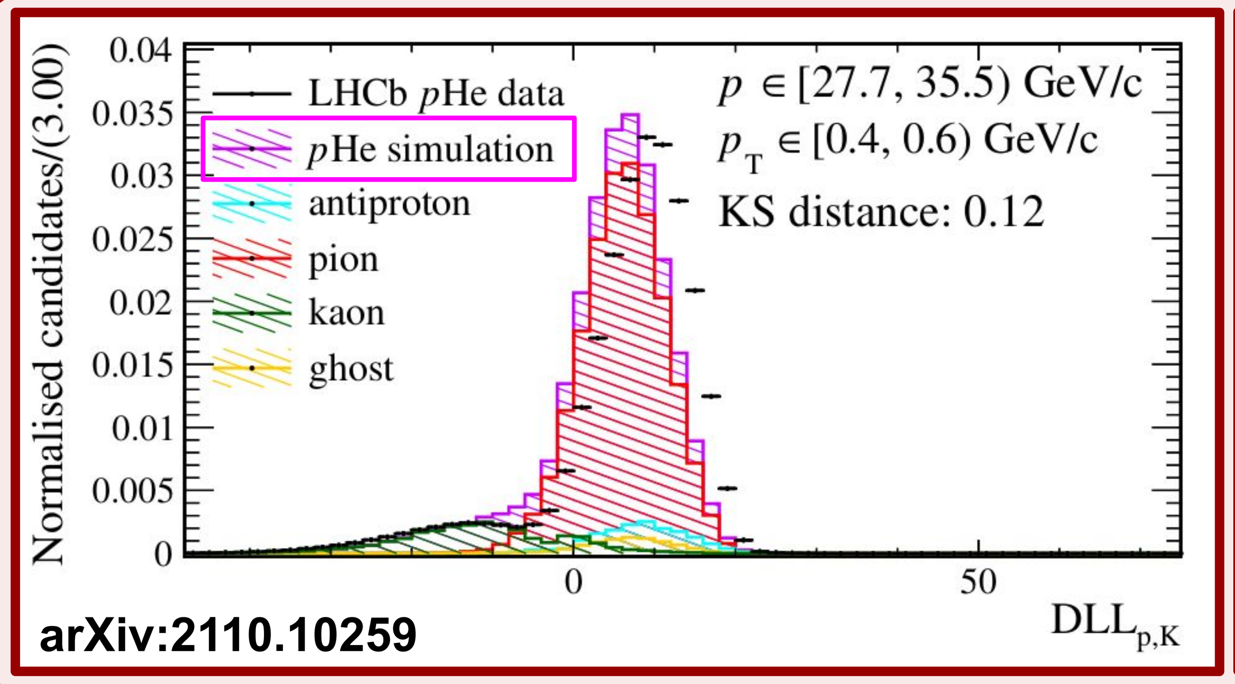
- **The loss function is the negative log-likelihood** for the n_p training events
- Weights w_i can be introduced[8] to statistically subtract calibration background candidates (*e.g.* with *sPlot* [9])

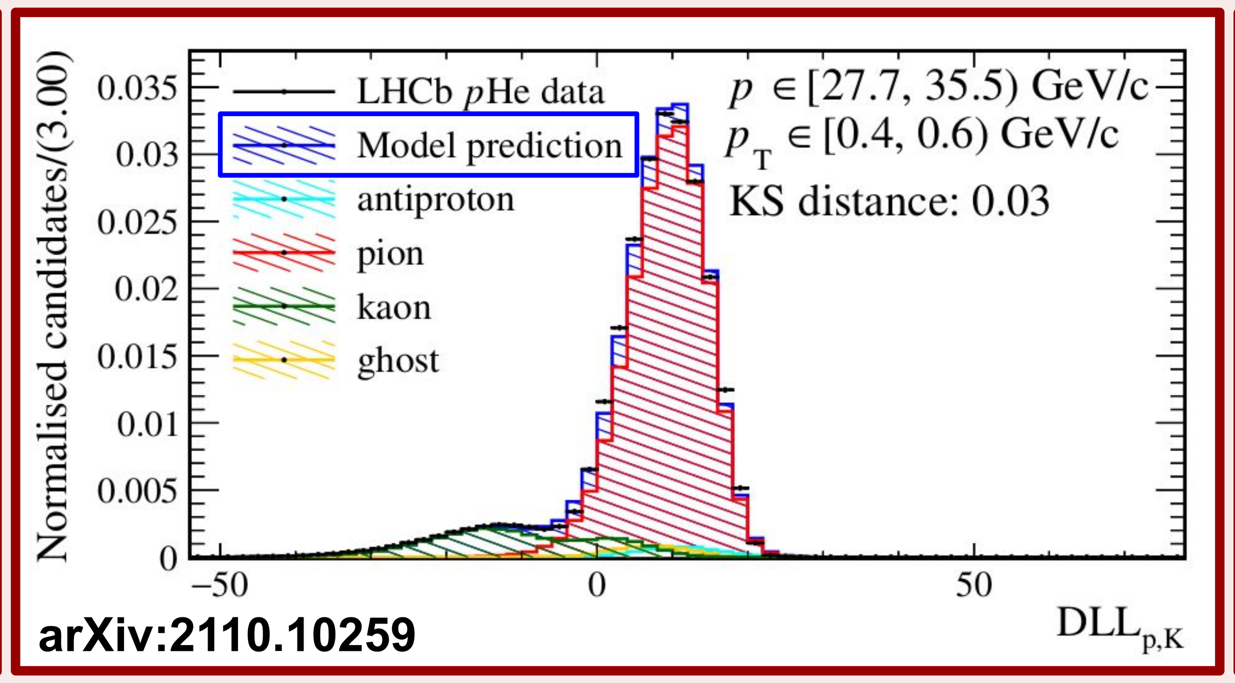


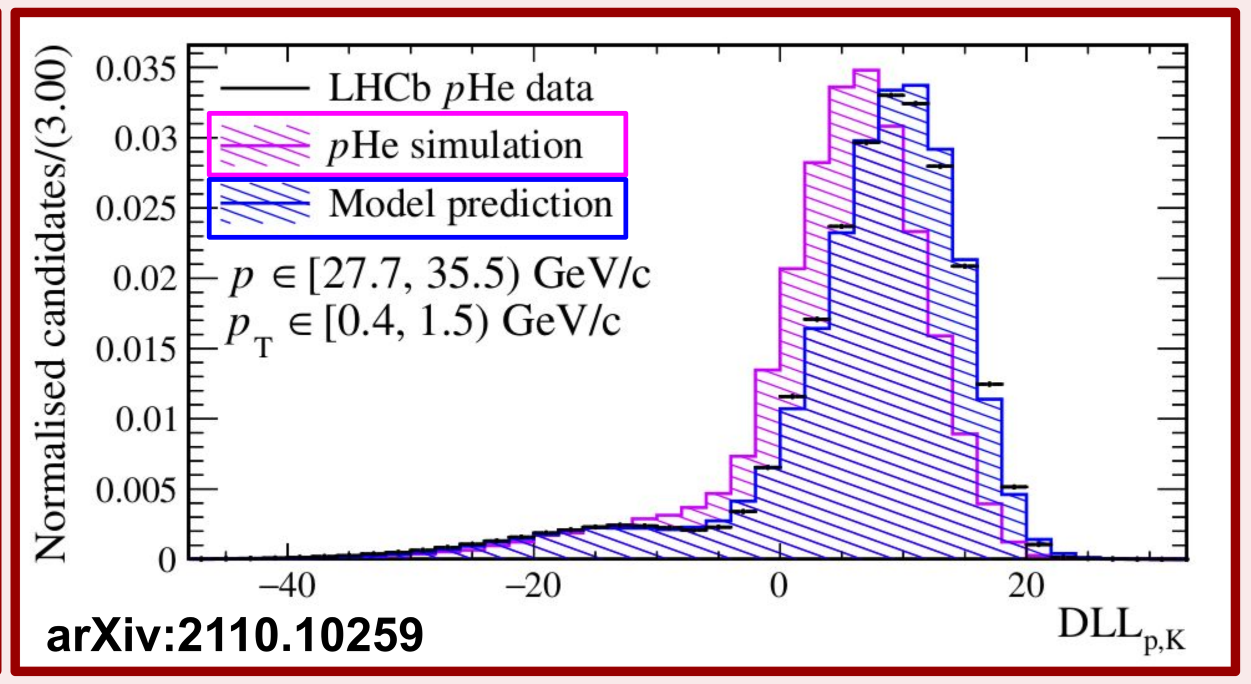
- **Considered features:** particle kinematics (*p*, p_T), occupancy in the RICH, number of tracks in the event, collision geometry and reconstruction quality
- **Training** performed with **mini-batches gradient descent** with a user-defined number of Gaussians and NN structure



- Learned PID dependence on $\underline{\theta}$ **is validated** comparing **training data** with the **model prediction** in intervals of all possible feature pairs
- **Smooth templates** are generated even in poorly-populated regions







- **Fit to *p*He data** with the composition of templates **predicted by the model** and drawn in the LHCb **detailed simulation** compared
- Data description quality **improved by the model**
- **Several use-cases** proposed/ongoing for analysis of fixed-target samples and description of systematics effects in *pp* data

[1] Mariani, S. *et. al.,* (2021) arXiV:2110.10259
[2] LHCb, (2008) JINST **3** S08005
[3] LHCb, (2015) Int. J. Mod. Phys. **A30** 1530022
[4] Adinolfi, M. *et. al.,* (2013) Eur. Phys. J **C73** 2431
[5] LHCb, (2019) LHCb-TDR-020
[6] Bursche, A. *et. al.,* (2018) LHCb-PUB-2018-015
[7] LHCb, (2018) PRL **121** 222001
[8] Borisyak, M. and Kazeev N., (2019) JINST **14** 08020
[9] Pivk, M. and Le Diberder F. R., (2005), NIM **A555** 356