# Hands-on: Systematics framework

Frank Meier

Belle II Summer Workshop
Durham
24 – 28 July 2023


Duke
UNIVERSITY

# Introduction

- Physics analyses have two sources of uncertainties: statistical and systematic
- various sources of systematic uncertainties must be considered
- pyhf hands-on session showed how they can be directly incorporated into likelihood fit
- physics performance group provides recipes, recommendations, values, and size of many systematic uncertainties
- check out conference readiness confluence page for detailed information
- PID (and in the future many more) systematics are provided via the "Systematics Framework"
- documentation can be found here: https://syscorrfw.readthedocs.io/en/latest/index.html

# Idea of Systematics Framework

- PID distributions may differ between experimental data and simulation
- naive PID selection might introduce large systematic uncertainty
- mitigation of data-MC differences by using the following control samples
  - for kaon PID: $D^{*+} \to [D^0 \to K^- \pi^+]\pi^+$
  - for proton PID: $\Lambda^0 \to p\pi^-$
  - for pion PID: $K^0_S \to \pi^+\pi^-$
  - for lepton PID: $e^+e^- \to [\tau^\pm \to 3\pi^\pm \nu_\tau][\tau^\mp \to l^\mp \nu_l \nu_\tau]$
  - for electron PID: $J/\psi \to e^+e^-$, $e^+e^- \to e^+e^-e^+e^-$
  - for muon PID: $J/\psi \to \mu^+\mu^-$, $e^+e^- \to \mu^+\mu^-\gamma$, $e^+e^- \to e^+e^-\mu^+\mu^-$
- isolate signal tracks by fitting invariant mass distribution of control sample and computing sWeights
- perform data-driven corrections to MC distributions
  - reweighting in momentum, polar angle, and potentially other variables
  - replacing with values sampled from calibration samples
  - transforming (not yet implemented)

# Sweights

- disentangle signal from background in discriminating variable
- via signal weights effectively subtract background from data sample
- study signal-only distributions of control variables
- most prominent example: use invariant mass distribution and apply weights to decay time distribution
- original paper: "SPlot: A Statistical tool to unfold data distributions" by M. Pivk and F. R. Le Diberder arXiv:physics/0402083
- generalization of concept: "Custom Orthogonal Weight functions (COWs) for Event Classification" arXiv:2112.04574
- sWeight calculation implemented in RooStats and in python package sweights
- warning: naive covariance calculation invalid for weighted fits
  - see "Parameter uncertainties in weighted unbinned maximum likelihood fits" for details arXiv:1911.01303
  - in RooFit use `AsymptoticError(true)`, in sweights package `cov_correct`
- run sweights.ipynb notebook available on indico or at ~fmeier/B2SW2023/sweights.ipynb on kekcc

# Setting up local installation of systematics framework

- git clone git@gitlab.desy.de:belle2/performance/systematic_corrections_framework.git
- cd systematic_corrections_framework
- bash setup_pypath.sh
- You'll be asked whether you want to download and build the Meerkat library. For the purposes of this hands-on session we don't need it, so please type "n".
- Try out some of the notebooks. Recommendation: run particleid_performance_analysis.ipynb
- even without kekcc account you can read and see output on web page