

Yandex



LHCb opendata and outreach highlights.

Personal perspective.

2016-10 International Workshop on Nuclear Emulsions, Napoli

Andrey Ustyuzhanin,
Yandex School of Data Analysis, Visiting Researcher INFN
Napoli

Example 1: D0-Masterclass

| Dataset:

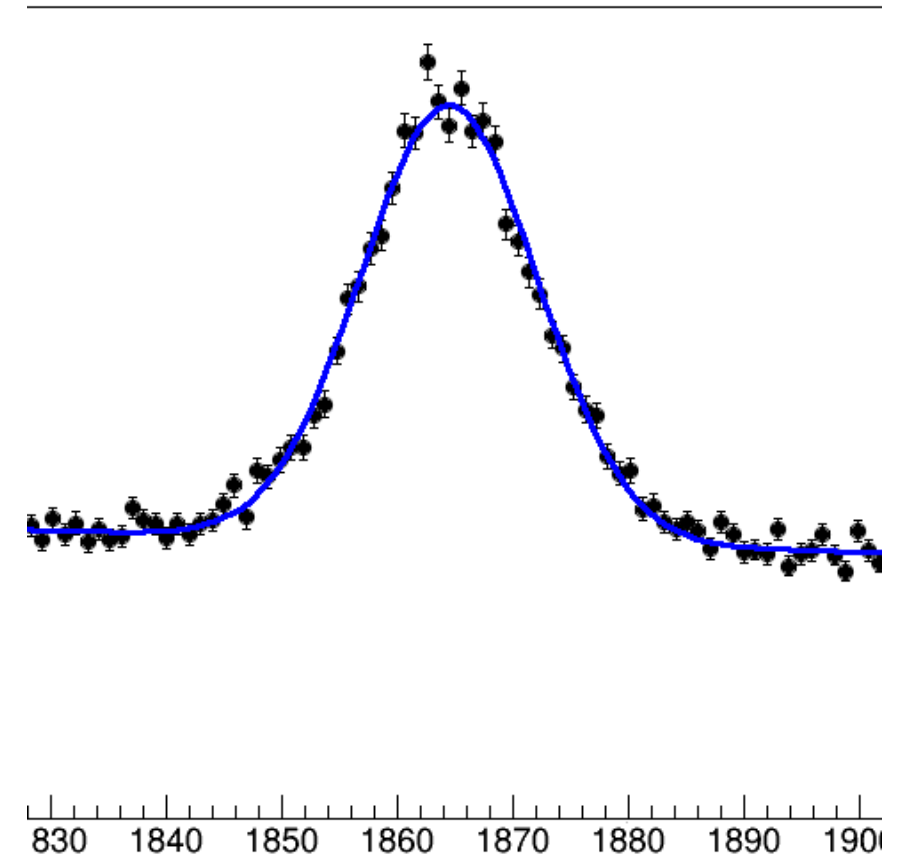
- › $D^0 \rightarrow K^+ \pi^-$
Format: root n-tuple
- › Year: 2010

| Masterclass

- › Goal: measure mass of D^0 , lifetime
- › Target audience: high-school students
- › International event, several times a year

<http://opendata.cern.ch/about/LHCb>

A RooPlot of "D0_MM"



Example 2: CP asymmetry

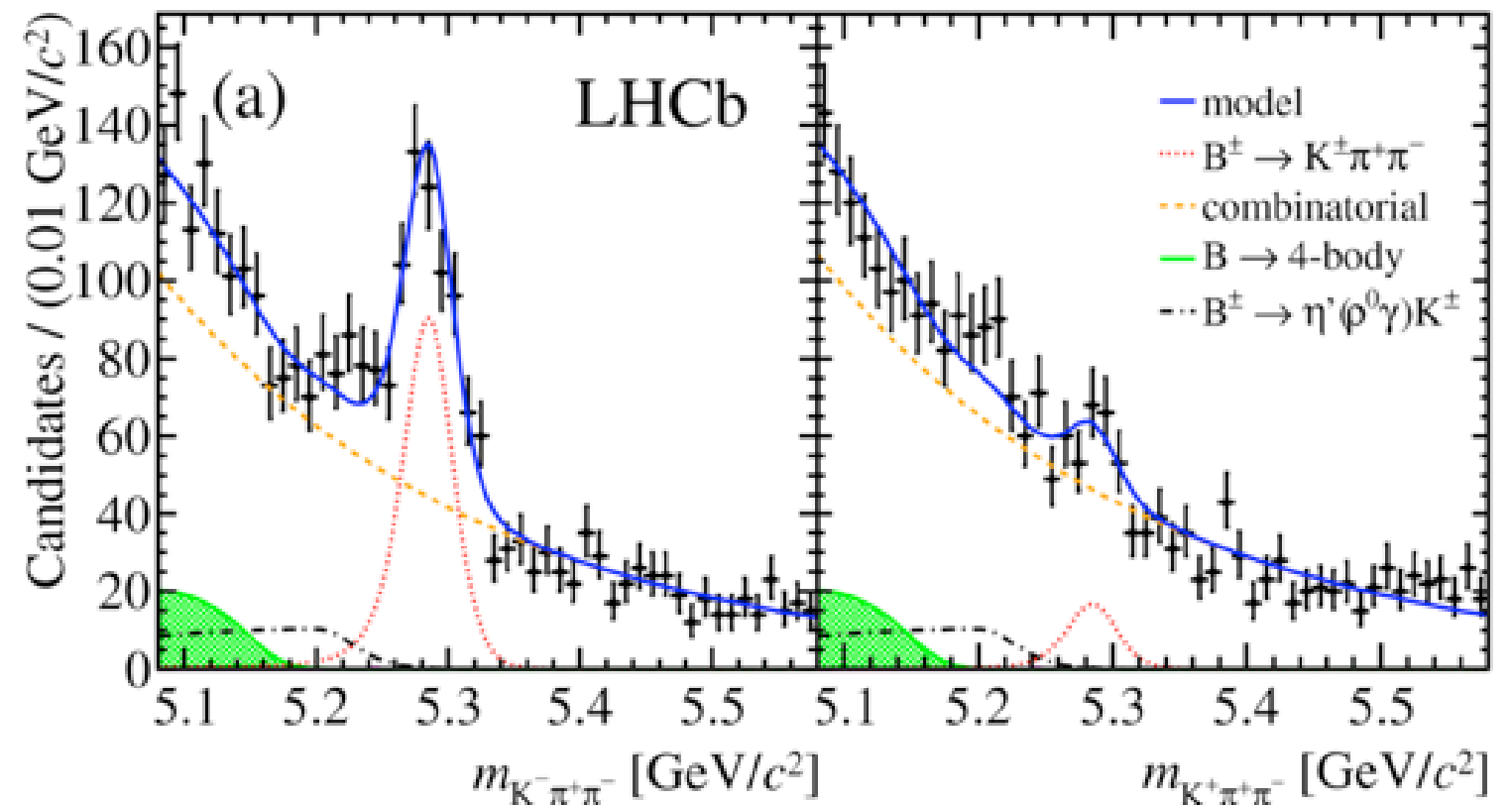


Dataset:

- › Format: root n-tuple
- › Year: 2011

Masterclass

- › Goal: measure CP violation
- › Target audience: Bachelor students, societies (astronomy)
- › Initially created by Manchester Uni



It used to take > 1 week to complete!

Starting Masterclass

The screenshot shows a web browser window with several tabs. The active tab is 'lhcb/opendata...' with a '502 Bad Gateway' error. The browser address bar shows 'https://github.com/lhcb'. The page displays the GitHub interface for the repository 'lhcb / opendata-project'. It includes navigation links for 'Code', 'Issues 1', 'Pull requests 0', 'Projects 0', 'Wiki', 'Pulse', and 'Graphs'. A message states 'No description or website provided.' Below this, repository statistics are shown: 134 commits, 3 branches, 0 releases, 7 contributors, and GPL-2.0 license. Action buttons include 'New pull request', 'Create new file', 'Upload files', 'Find file', and 'Clone or download'. A commit history table is visible at the bottom.

Commit	Message	Time
chrisburr committed on GitHub	Update .gitignore	Latest commit 656c964
Background-Information-Notebooks	Many changes in preparation for beta release (#7)	3 mon
Data	Correctly revert the main Dockerfile	3 mon
Images	Add changes from @ChrisParkes	4 mon
.gitignore	Update .gitignore	3 mon
Dockerfile	jupyter start update to support push (#11)	3 mon
Example-Analysis.ipynb	Move setup code into an external file (#9)	3 mon

Supplementary tutorial

Яндекс Home Example-An... LHCb_Oper_Da... ManchesterTut... https://...minais/1

Яндекс <https://everware.rep.school.yandex.net>

jupyter Example-Analysis (autosaved) Python 2

File Edit View Insert Cell Kernel Widgets Help

+ ↶ ↷ ↵ ↴ ↵ ↴ Code CellToolbar

▼ Analysis of Nobel prize winners

Welcome to the programming example page. This page shows an example analysis of Nobel prize winners. The coding commands and techniques that are demonstrated in this analysis are similar to those that are needed for your particle physics analysis.

IMPORTANT: For every code box with code already in it, like the one below you must click in and press shift+enter to run the code. This is how you also run your own code.

If the In [x]: to the left of a codebox changes to In [*]: that means the code in that box is currently running

If you ever want more space to display output of code you can press the + button in the toolbar to the right of the save button to create another input box.

<https://everware.rep.school.yandex.net/user/anaderi/notebooks/Example-Analysis.ipynb#>

Interacting with notebook

Яндекс Яндекс Home 502 Bad Gateway LHCb_Open_Da... ManchesterTut... https://...minals/1

Яндекс <https://everware.rep.school.yandex.net>

jupyter Control Panel Logout

Files Running Clusters

Select items to perform actions on them. Upload New ↕ ↻

- Background-Information-Notebooks
- Data
- Images
- Example-Analysis.ipynb
- LHCb_Open_Data_Project.ipynb** Running
- _start_jupyter.sh
- B2HHH_MagnetDown.root
- B2HHH_MagnetUp.root
- Dockerfile

https://everware.rep.school.yandex.net/user/anaderi/notebooks/LHCb_Open_Data_Project.ipynb

Technological landscape

Entirely web-based

removes OS dependency,
user IT requirements



Hosted by Yandex

Jupyter Notebooks

› Instructions, hints, user code

Docker

› Environment

GitHub

› Code storage

CERN open data portal

› Data (to appear, for the moment uses cernbox)

Everware

› Bridge between Github/
Docker/Jupyter

Example 3: Web-based Event Display

LHCb:

<https://lbevent.cern.ch/EventDisplay/index.html>

Other experiments:

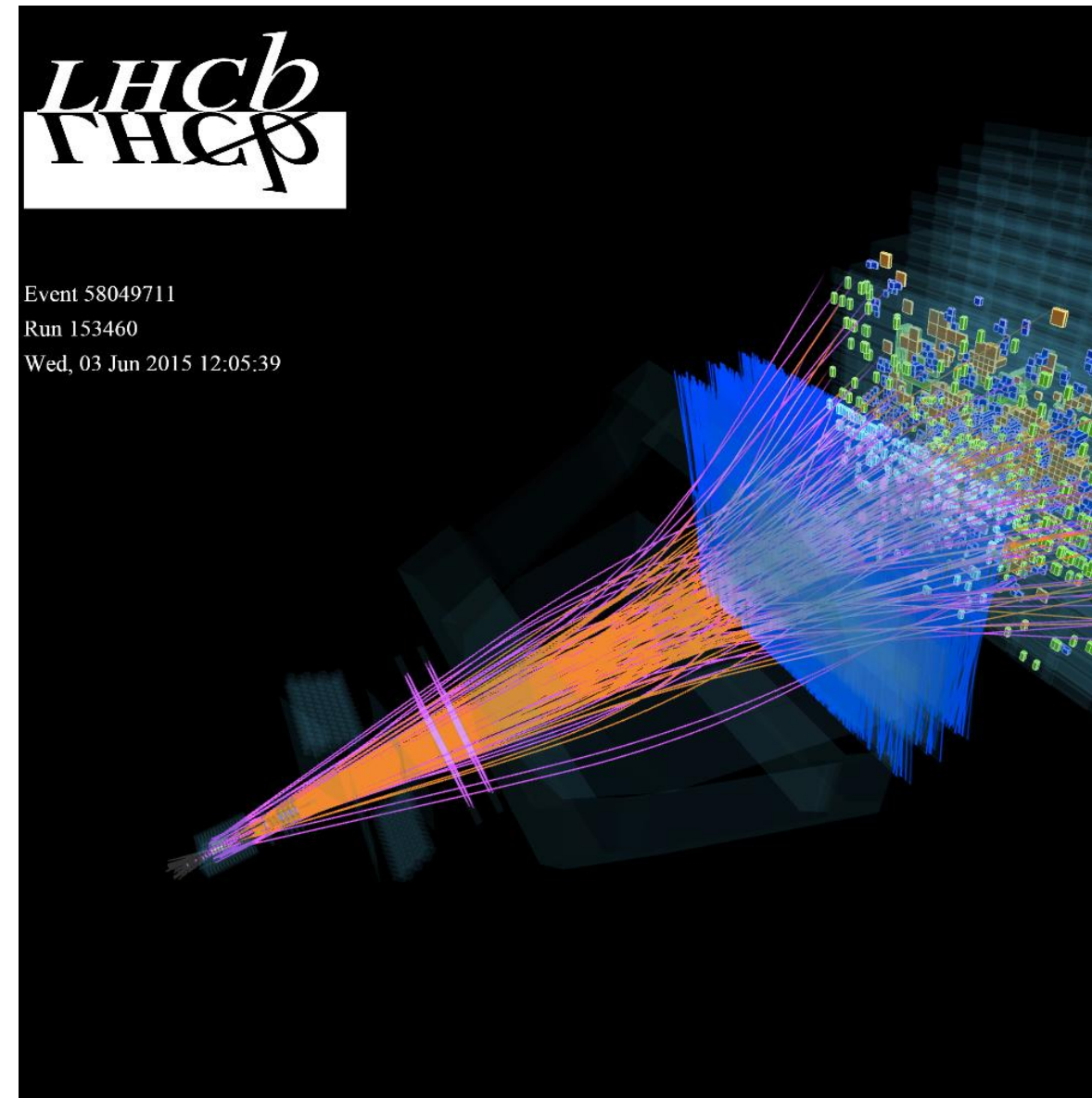
<http://atlas.cern/atlas-live>

<http://mcmstv.web.cern.ch/mcmstv/#home>

<http://cern.ch/ispy-webgl>

Technologies:

› Javascript, WebGL, Unity, ...



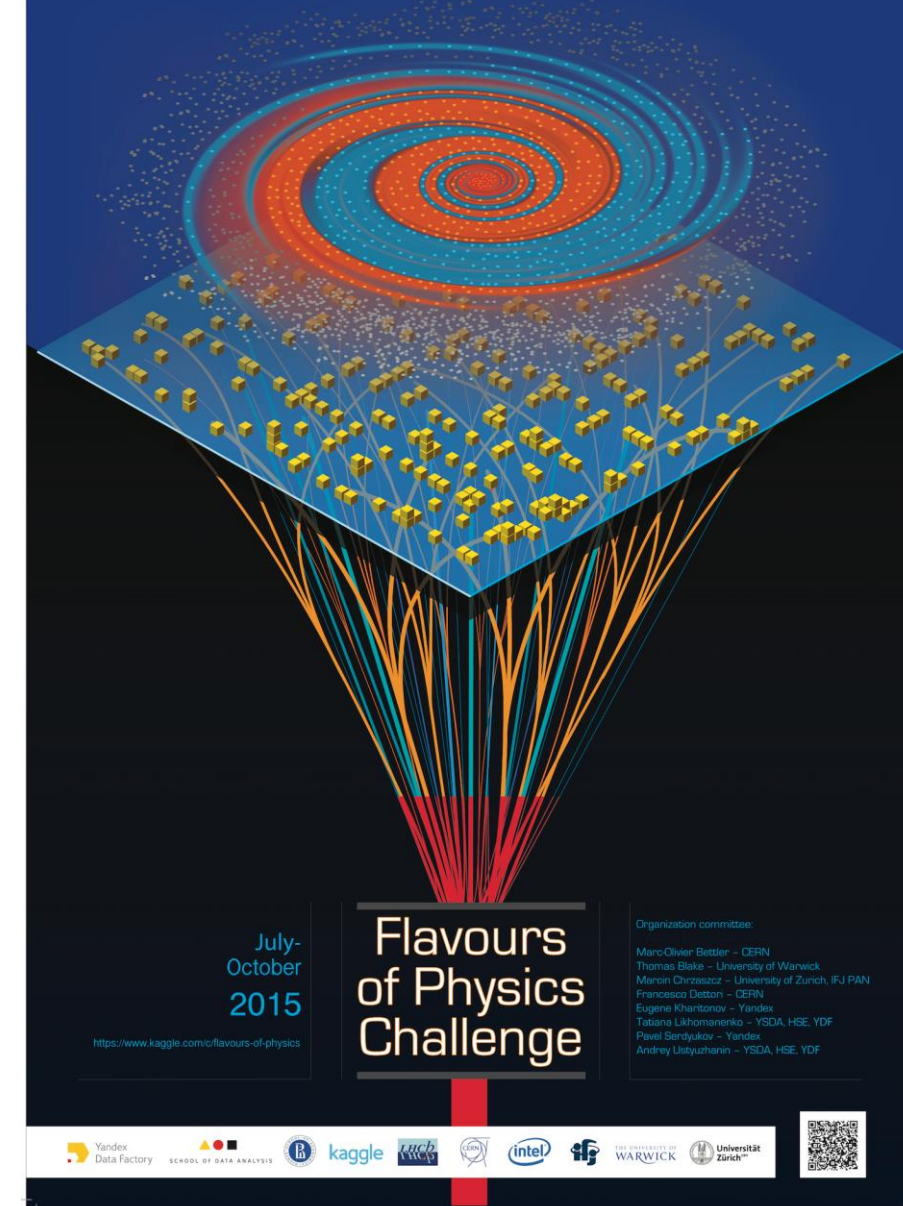
Example 4: $\tau \rightarrow 3\mu$

I Dataset:

- › $D_s \rightarrow \varphi\pi, \tau \rightarrow 3\mu$, MC (Simulated)
- › Format: CSV
- › Year: 2012

I Challenge

- › Goal: design algorithm looking for $\tau \rightarrow 3\mu$,
- › Target audience: data science students, ex physicists,
- › International challenge: 2015,
- › Complex metrics,
- › Interesting solutions, outcomes.



Conclusion

- | Formats of interaction vary in wide range
 - > Event display,
 - > Interact with root files using ROOT
 - > Interact with root files using Python
 - > Challenge(s)
- | Every format takes considerable of efforts
 - > Technological,
 - > Administrative,
 - > Marketing (main audience driver!)
- | Other LHC experiments have plenty of useful experience (CMS)

Thank you for attention!

Andrey Ustyuzhanin, andrey.ustyuzhanin@cern.ch