

CERN Open Data

for teaching and outreach

Partikeldagarna 2026, Sept 18
Lunds Universitet

Christian Ohm, KTH

Christian Bierlich & Alice Ohlson, Lund



LUNDS
UNIVERSITET



CERN OpenData

“The CERN Open Data Policy reflects values that have been enshrined in the CERN Convention for more than sixty years that were reaffirmed in the European Strategy for Particle Physics (2020), and aims to empower the LHC experiments to adopt a consistent approach towards the openness and preservation of experimental data. Making data available [...] allows the maximum realisation of their scientific potential and the fulfillment of the collective moral and fiduciary responsibility to member states and the broader global scientific community.”

[CERN OpenData policy \(2020\)](#)

OpenData released by ATLAS, CMS, ALICE, LHCb for

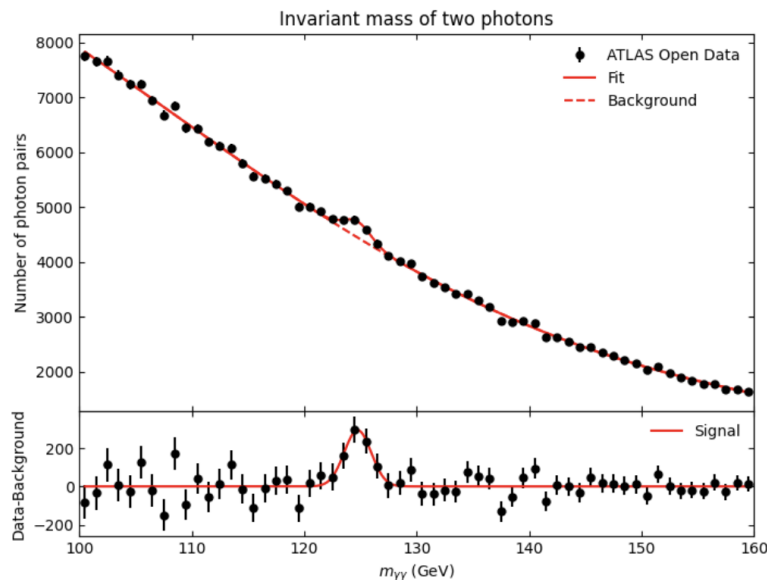
Research, Education, and Outreach

as well as software tools and tutorials for analysing the data

[OpenData portal](#)

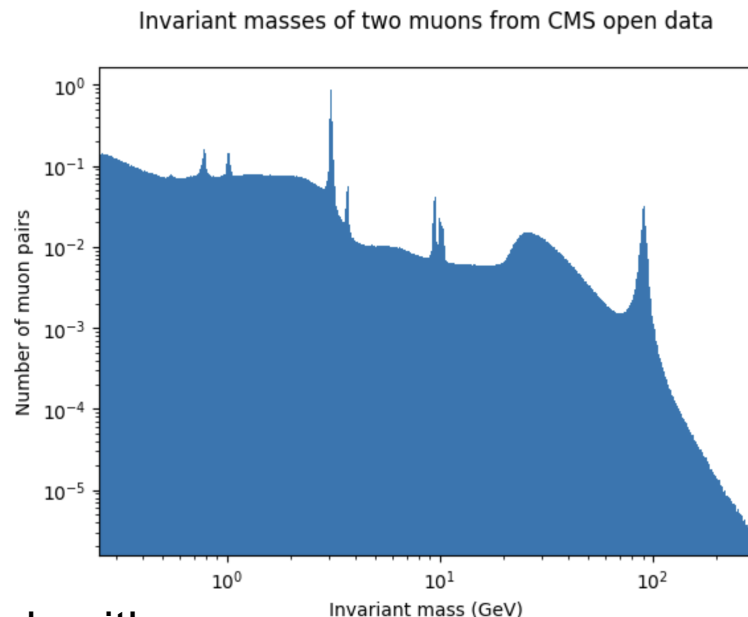
Short exercise for bachelors students: ATLAS $H \rightarrow \gamma\gamma$

- Intended audience: introductory course for bachelors students, short in-class exercise (2-3 hours)
- Based on [ATLAS tutorial](#), but reworked as a student exercise tuned to course context
- No fancy software needed, only Python and Jupyter notebooks, data is provided in a csv text file
- Guide students through analysis steps (but not fill-in-the-blank): Read in photon data, evaluate invariant mass, apply cuts, plot diphoton spectrum, fit



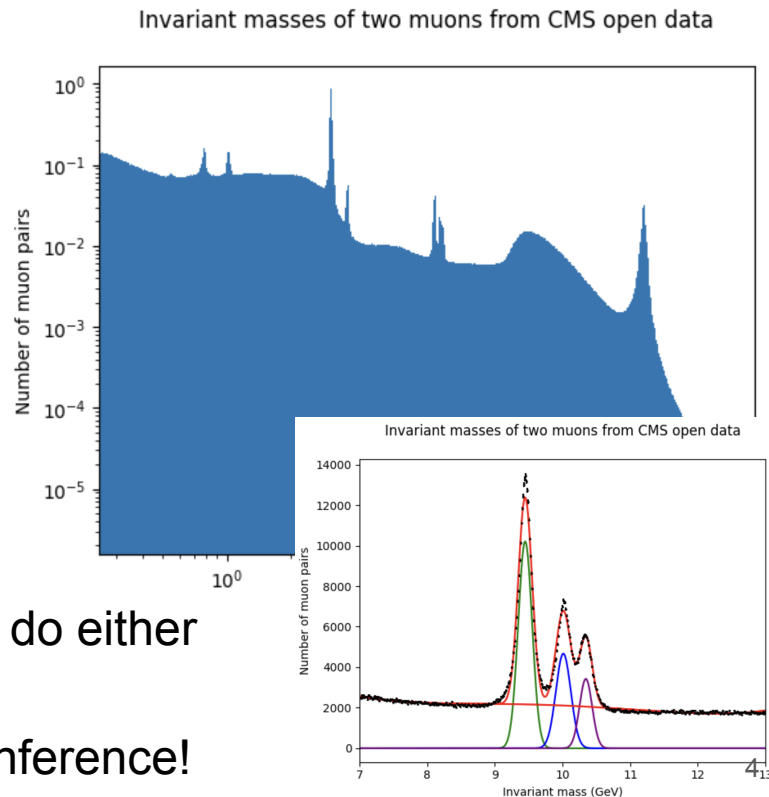
Short exercise for bachelors students: CMS dimuons

- Based on [CMS tutorial](#), but rewritten as a student exercise (Python) and data filtered into a csv file
- Students read in data, calculate invariant mass, plot dimuon spectrum, identify peaks
- Suggestions given for “more exploration”:
 - Calculate S/B ratio, then modify cuts and see how purity changes
 - Zoom in on Y mass and resolve 1s, 2s, 3s states
- Idea for the future (thanks to LKB): students do either CMS or ATLAS exercise (or LHCb, ALICE), present to each other via posters → mini conference!



Short exercise for bachelors students: CMS dimuons

- Based on [CMS tutorial](#), but rewritten as a student exercise (Python) and data filtered into a csv file
- Students read in data, calculate invariant mass, plot dimuon spectrum, identify peaks
- Suggestions given for “more exploration”:
 - Calculate S/B ratio, then modify cuts and see how purity changes
 - Zoom in on Y mass and resolve 1s, 2s, 3s states
- Idea for the future (thanks to LKB): students do either CMS or ATLAS exercise (or LHCb, ALICE), present to each other via posters → mini conference!



Project: MSc course on computational physics & ML

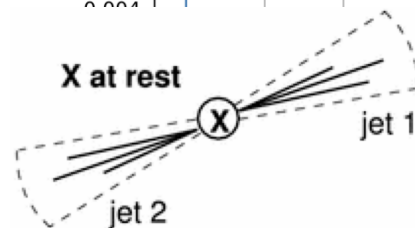
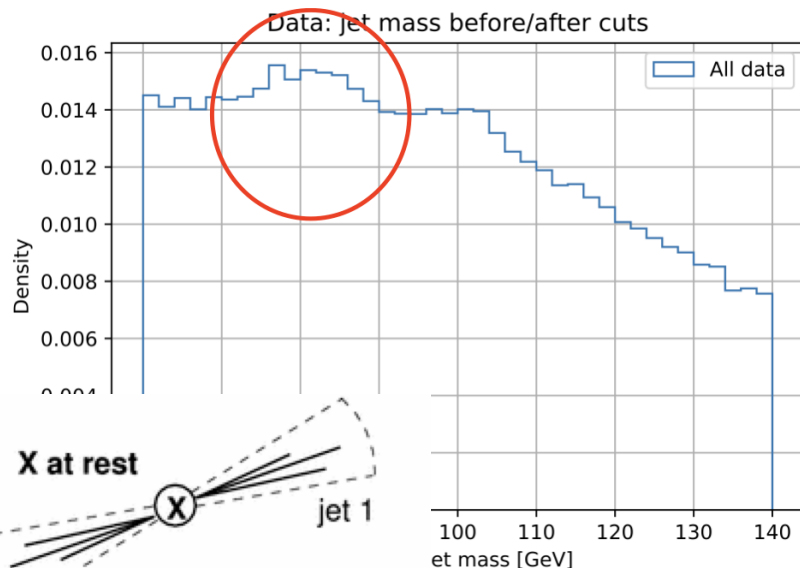
A Neural-Network tagger for LHC

Separate $W/Z \rightarrow$ jets from background

ATLAS open data, flattened format, own Monte Carlo

Build NN; train, test, validate on MC; apply to data – ***can you beat the simple cuts?***

Students have already done a toy NN separation task and implemented anti-kT from scratch.



<https://gitlab.com/comp-phys-ai> (apply for access and email me: christian.bierlich@fysik.lu.se)

Students like working a real NN problem

Real physics data & MC makes **project authentic**

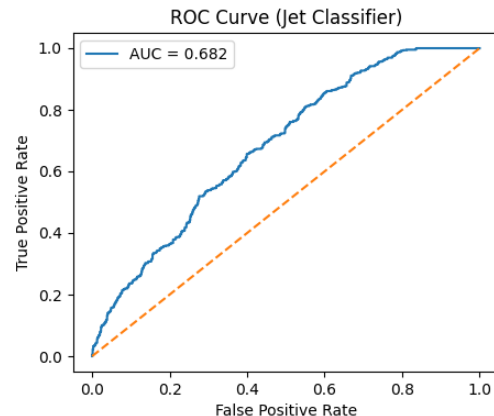
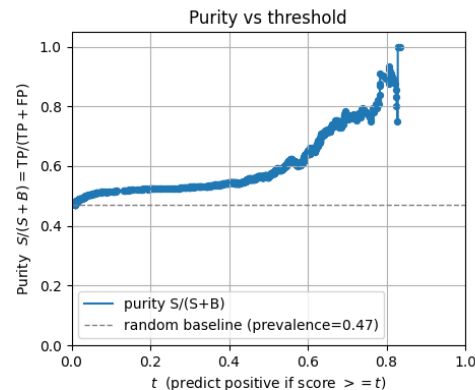
By choice: It is not easy to obtain good separation

The project was appropriately scoped in terms of workload and difficulty.	Number of responses
I strongly agree	12 (57,1%)
I somewhat agree	8 (38,1%)
I somewhat disagree	0 (0,0%)
I strongly disagree	1 (4,8%)
Total	21 (100,0%)

I really liked this project, well structured and interesting.

No notes! Good project! Frustrating that the model kind of sucked...

<https://gitlab.com/comp-phys-ai> (apply for access and email me: christian.bierlich@fysik.lu.se)



ATLAS open-data labs at KTH - measure the Z mass

- Part of [SH1015 Applied Modern Physics](#), for 3rd year bachelor students
- Over 100 students per year
- Designed to teach
 - statistical data analysis
 - mathematical modeling & fitting
- Includes lots of physics:
 - Different Z shapes for ee and $\mu\mu$ - bremsstrahlung, resolution, etc
 - Width-lifetime connection
 - Background processes, small but still important to model
- Constructed in 2018, by now around 1000 students have done it (time to update..)
- [Notebooks](#) in docker or Google Colab

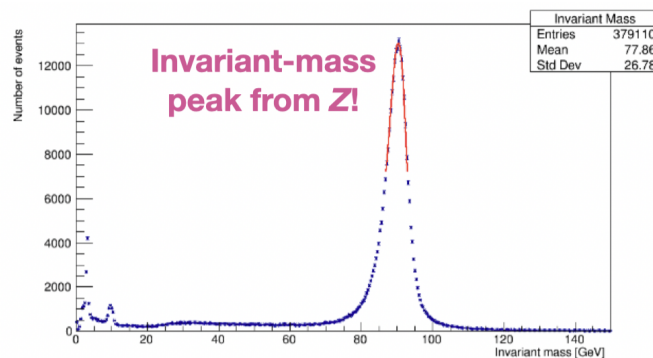
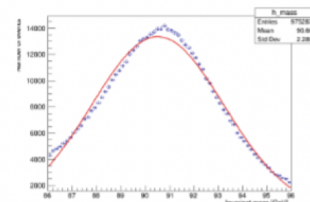


Figure 3: Invariant mass of the Z^0 boson, approximated as the function $\mathcal{A}$, on the range [87, 93] [GeV].

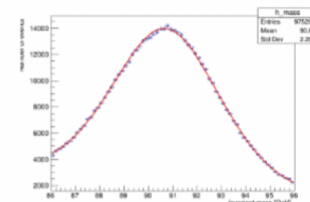
Z^0 boson mass

SH1015

HT 2018



(a) Gaussian,
 $\chi^2/DF = 24.4$



(b) Breit-Wigner + background,
 $\chi^2/DF = 1.2$

Figure 5: Fitting different distributions to the results. DF signifies the number of degrees of freedom

ATLAS open-data labs at KTH - explore and discover

- Part of [SH2103 Subatomic Physics](#), for 1st year master students (30-40 per year)
- Since 2019, uses same sw stack ([repo](#))
- Uses Z as an example, then ***let's the students themselves decide what particle they want to try to find*** in the data.
 - Open task - piques the interest
 - Freedom leads to curiosity/creativity
- Students are asked to come up with a proposal for a particle, what final state, thoughts about backgrounds, etc - then check-in with teacher (and discuss/tune if needed) before starting

ATLAS open-data labs at KTH - explore and discover

- Part of [SH2103 Subatomic Physics](#), for 1st year master students (30-40 per year)
- Since 2019, uses same sw stack ([repo](#))
- Uses Z as an example, then ***let's the students themselves decide what particle they want to try to find*** in the data.
 - Open task - piques the interest
 - Freedom leads to curiosity/creativity
- Students are asked to come up with a proposal for a particle, what final state, thoughts about backgrounds, etc - then check-in with teacher (and discuss/tune if needed) before starting

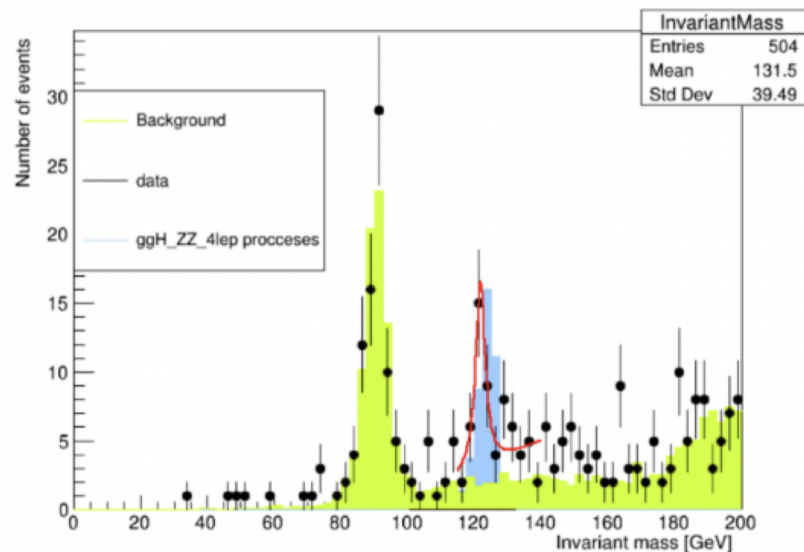
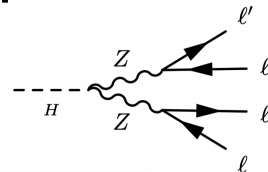


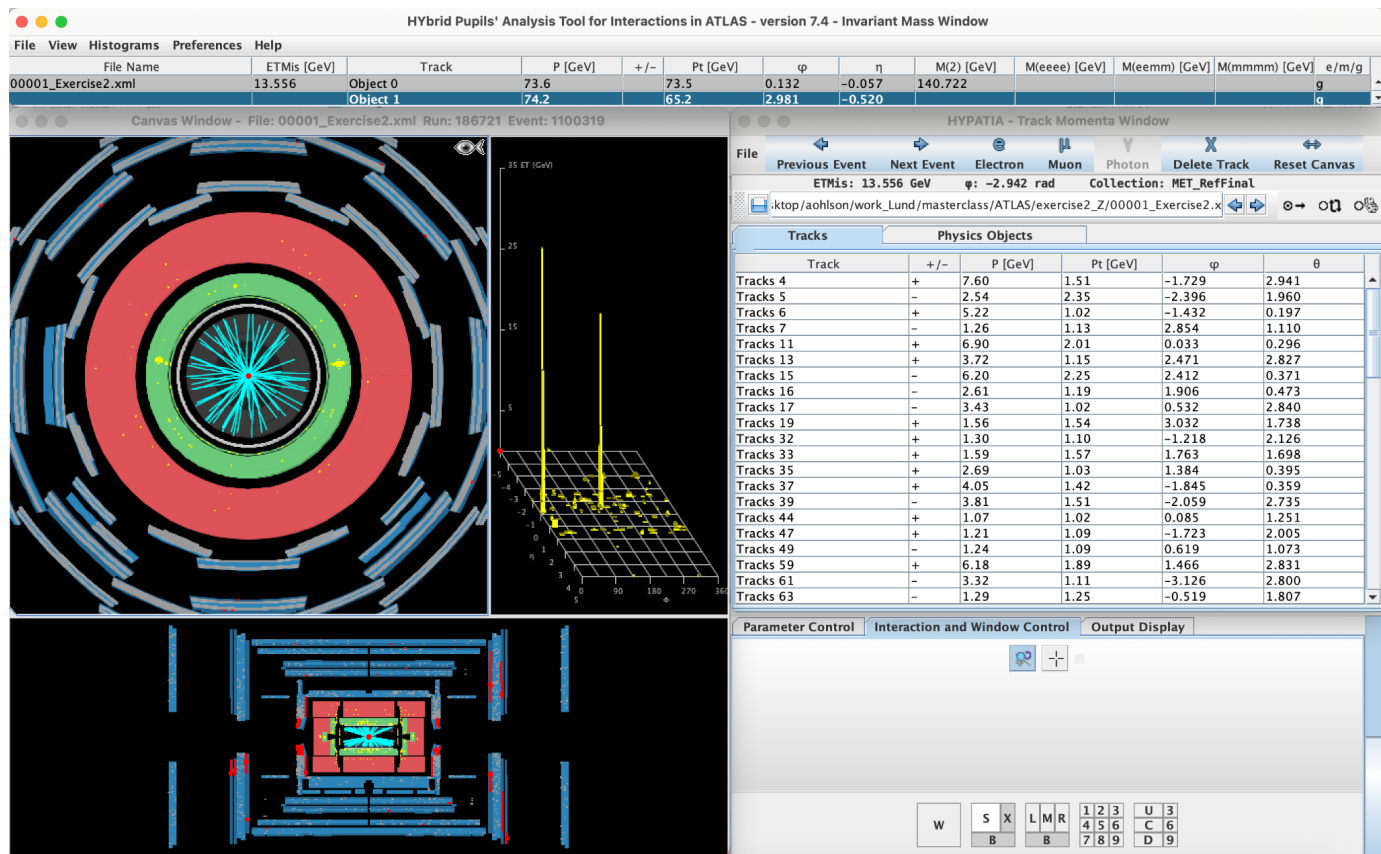
Figure 2.2.7: The simulated invariant mass histogram for the production of the Higgs boson and the simulated invariant mass histogram for background processes resulting in 4 lepton on top of the experimentally attained invariant mass histogram for the production of 4 leptons.

IPPOG masterclasses - “Bli partikelfysiker för en dag!”

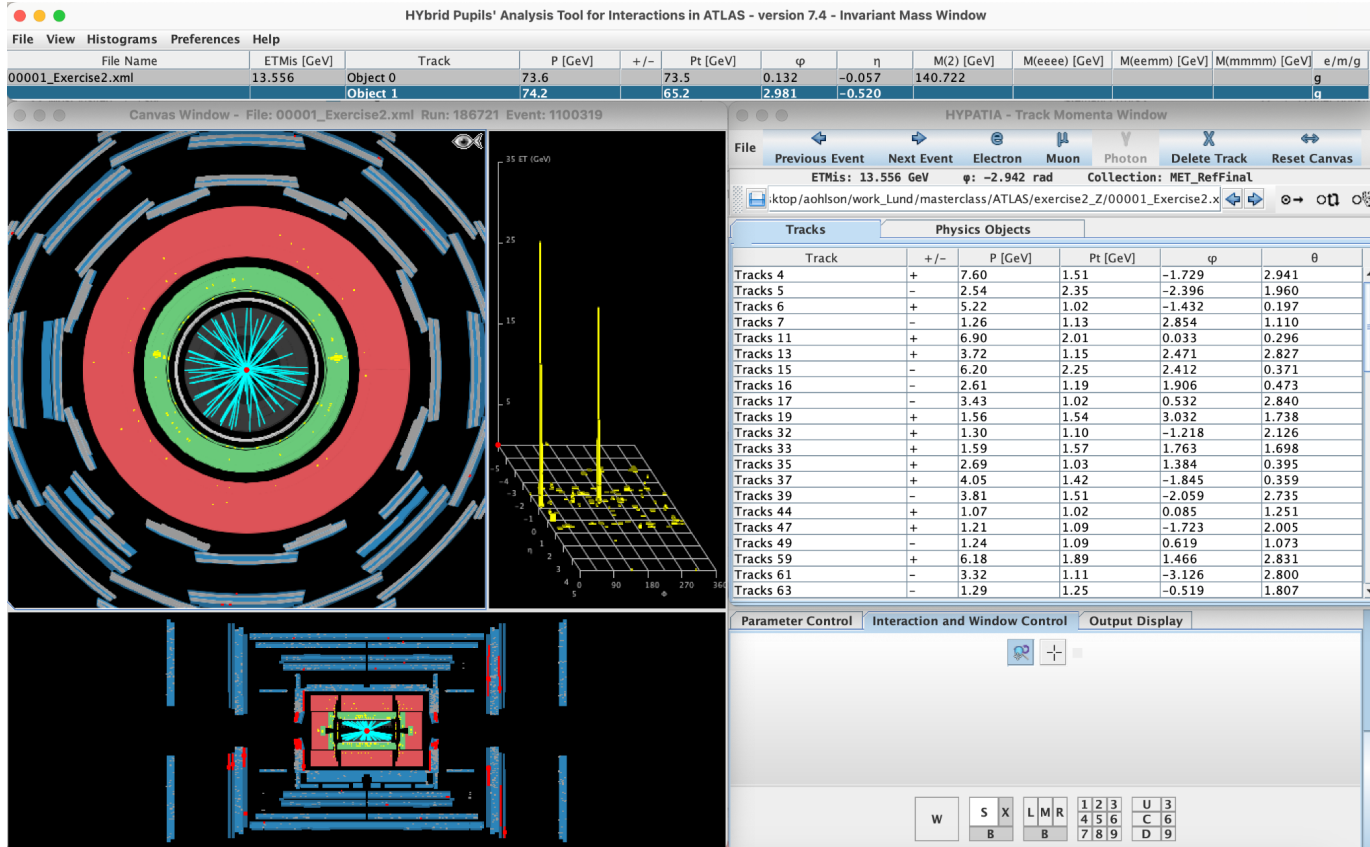
- Organized by International Particle Physics Outreach Group ([IPPOG](#)) and [CERN](#)
- Hosted in Lund and in Stockholm, students come to the university for one day
 - Morning: lectures (intro to the Standard Model and experiments, etc)
 - Afternoon: computer exercise to look for the Higgs in ATLAS data
 - Followed by a video conference with CERN and other student groups
- Target audience: 3rd year high school (gymnasium) students
- The aim:
 - Students learn about modern ideas in particle physics from active researchers
 - Apply what they have learned in school
 - Hands-on experimental experience: analyse real data
 - Collaborate with students from Sweden and abroad
 - Discuss work and careers in particle physics, and science more generally

If you're working on ATLAS, you can also volunteer to be a moderator for the zoom conference

$H \rightarrow \gamma\gamma, H \rightarrow ZZ \rightarrow 4 \text{ leptons (e or } \mu)$



$H \rightarrow \gamma\gamma, H \rightarrow ZZ \rightarrow 4 \text{ leptons (e or } \mu)$



Tue, Mar 03

VC 1: ATLAS Z

16:00 CET

Alexander

Shalini

Graz



Lund



Trieste



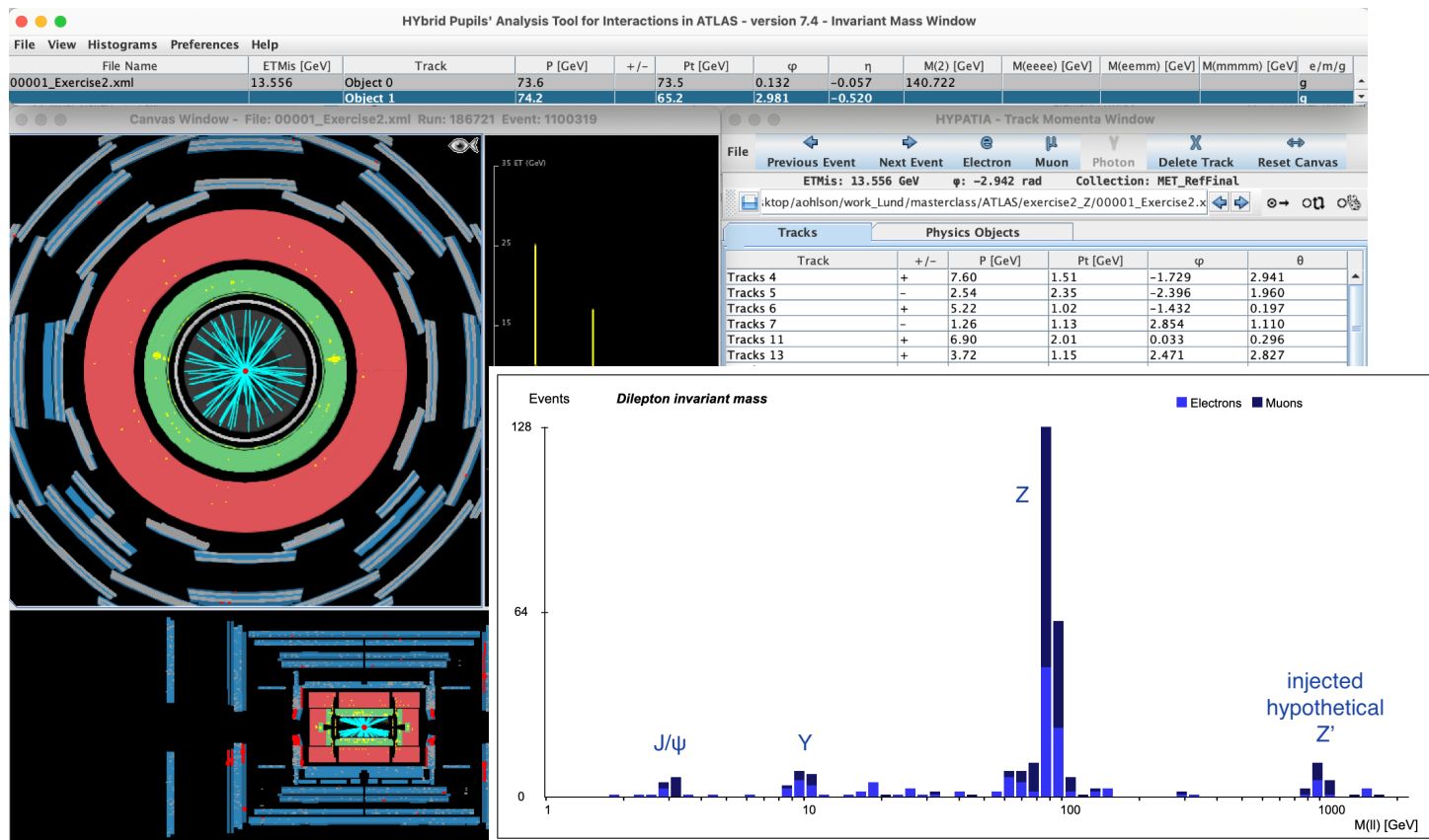
Heidelberg



Prague, Charles U.



$H \rightarrow \gamma\gamma, H \rightarrow ZZ \rightarrow 4 \text{ leptons (e or } \mu)$



Tue, Mar 03

VC 1: ATLAS Z

16:00 CET

Alexander

Shalini

Graz



Lund



Trieste



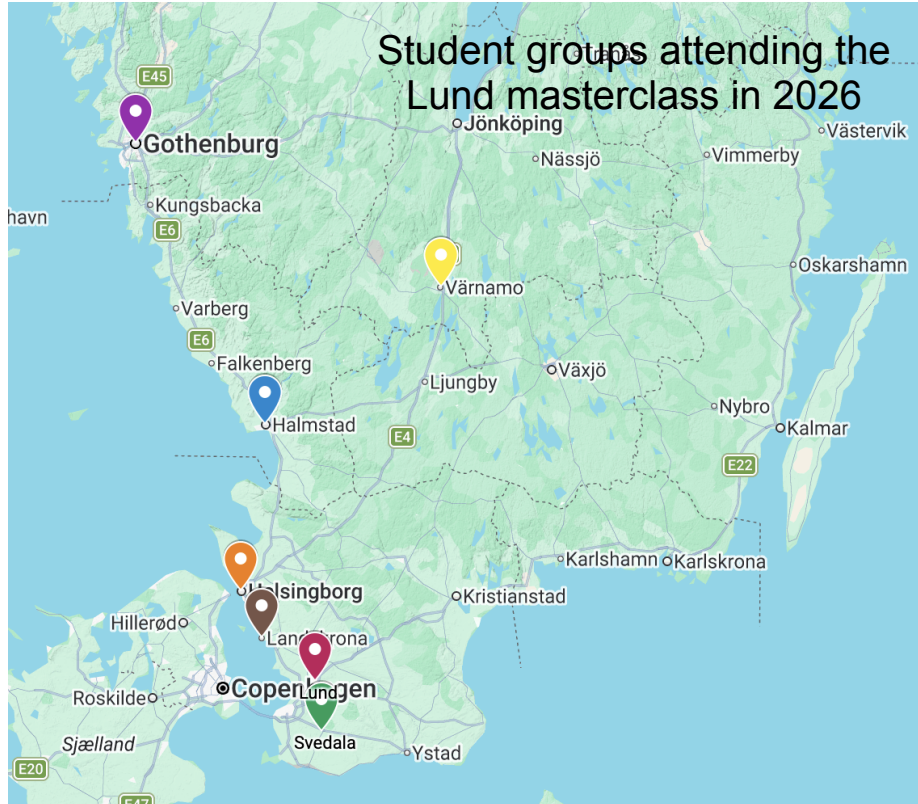
Heidelberg



Prague, Charles U.



A great way to reach and inspire tomorrow's scientists!



“Tack för att vi får möjlighet att återigen anmäla elever till mästartklassen i partikelfysik. Förra året fick tre elever vara med på mästartklassen och idag studerar två av dem fysik och den tredje studerar matematik på LTH.”

Volunteers always needed!

And if you are interested in organizing the masterclasses elsewhere, please talk to us!

Reminder about IPPOG

- International collaboration with 45 members: 34 countries, 10 experiments, 1 International lab
- What it does
 - Collects resources that can be used for a range of outreach activities
 - Shares good examples and inspiring success stories of outreach initiatives
 - aims to ***make it easier for more of us to do more and better outreach***
- Collaboration meets twice per year, last one in June in Ljubljana ([indico](#))
- Plenty of opportunities to drive ideas and connect with others interested in outreach, see ippog.org (sign up for newsletter and email list)
 - What would you like IPPOG to do? Ideas and initiatives are very welcome
 - Hot topic: critical to be able to communicate motivation for future accelerators to all kinds of audiences!
- Financed through membership fee (thanks UU, LU pays next year)

Closing remarks

- Communicating what we do to various stakeholders is important — also *why* and *how* we do research
- CERN Open Data provides powerful resources
 - Extensive library of datasets from various experiments
 - Lots of tools, packages and technical instructions greatly reducing the thresholds for setting up exercises
- The high school master classes uses open data from CERN, and give students a feeling for how it is to do research in particle physics — and work in an international collaboration!
- IPPOG exists to help everyone do more research — use this, and propose activities, initiatives and improvements if you see opportunities for improvements!