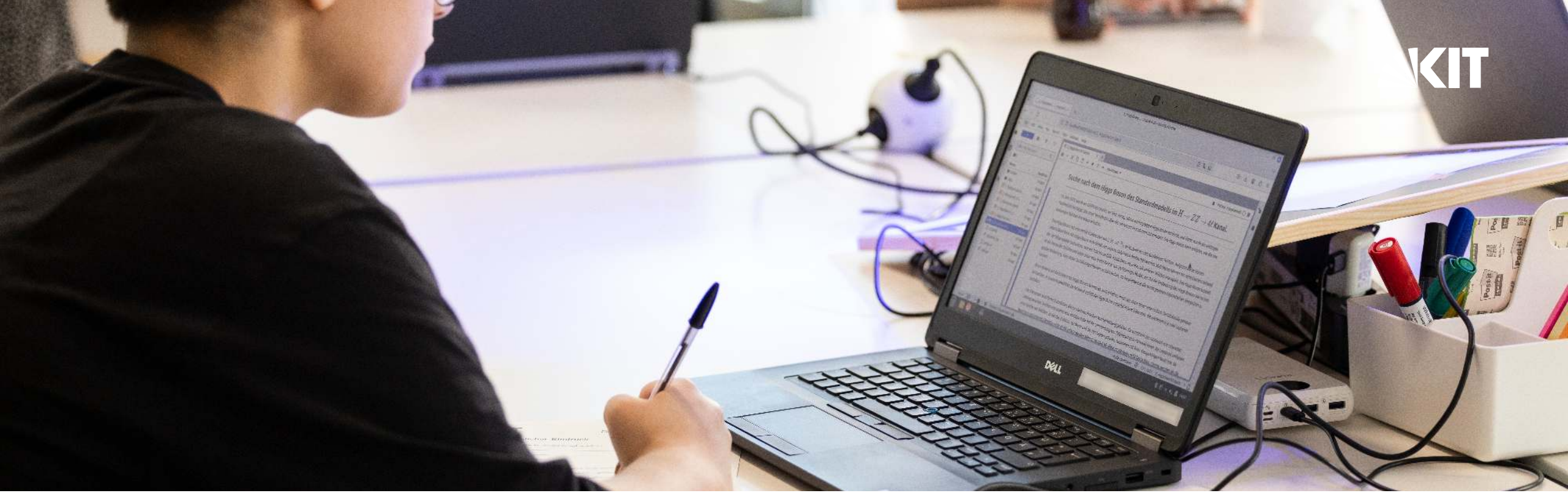


CMS Masterclass 2: Ein Erfahrungsbericht

Das Higgs-Boson entdecken mit Python Notebooks und CMS OpenData

Artur Monsch, Günter Quast, Cedric Verstege, **Christian Winter** | 4. April 2025

Institute of Experimental Particle Physics

Science Camp

Teilchen- & Astroteilchenphysik

Eckdaten

SCHÜLERAKADEMIE KARLSRUHE

- 20 Schüler:innen
- Jugendliche ab 16 (Klasse 9 – 11)
- 6 Tage
- Förderung durch Stiftungen

ETP Institute of Experimental Particle Physics

Inhalt

- Vorträge von Professoren
- Exkursionen
- Masterclass
- Gruppenprojekte
- Abschlusspräsentation

Mehr Infos: DPG 2024

Gruppenprojekt

CMS Masterclass 2: Higgs Suche

Basiswissen erforderlich

- python™ Kenntnisse
- Standardmodell der Teilchenphysik
- CMS Experiment
- Typische Größen in der Teilchenphysik

Voraussetzung

- CMS Masterclass
- Python Kurs

Link zum Material

Suche nach dem Higgs Boson des Standardmodells i

Im Jahr 2012 wurde am CERN das bereits vor über vierzig Jahren vorhergesagte Higgs-Boso
uns bekannten Teilchen ihre Masse erhalten.

Das **Higgs-Boson** hat eine mittlere Lebenszeit von $1.56 \cdot 10^{-22}$ s, also ein sehr kurzlebiges T
der Teilchen nachweisen, in die es sich umwandelt (umgangsprachlich wird auch oft von Zerf
der Teilchen. Es gibt also eine breite Palette von Umwandlungsmöglichkeiten. Bei der Entde
überprüfen zu können.

In dieser Aufgabe schauen wir uns jetzt den Kanal an, in dem sich das Higgs-Boson zunächst

Die **Z-Bosonen** sind schwere Elementarteilchen, die zur elektroschwachen Wechselwirkung
Kernreaktionen (z.B. beim radioaktiven β-Zerfall). Die **Leptonen** umfassen eine Familie von l
erfasst werden können, werden wir diese im Weiteren nicht betrachten. Ebenso werden wir
Rahmen sprengen. Daher beschränken wir uns auf die Elektronen und Muonen. Die Muonen
vollständig durchqueren können.

Die Arbeitsmethode für dieses Notebook wird darin bestehen, dass wir es gemeinsam im Dia
Anschluss daran werdet Ihr auch Abschnitte eigenständig oder in kleinen Gruppen bearbeite

Unser Hauptaugenmerk wird auf den simulierten Datensätzen liegen, und am Ende werden
entsprechenden Kollisionen zeilenweise auflisten. Genauere Details dazu werden im nächste

> Warum verwenden wir eigentlich simulierte Daten und nicht gleich die gemessenen?

Das Ziel wird darin bestehen, die Empfindlichkeit für Signalereignisse zu erhöhen, indem wir
Untergrundereignissen erreichen - im besten Fall nur von diesen -, während die Signalereigni

> Ist eine solche Vorgehensweise überhaupt möglich?

Am Ende werden wir unsere Filtermethoden auf die tatsächlich durchgeführte Messung anw
Überschusses werden wir auf einfache Weise quantifizieren.

In der folgenden Zelle importieren wir Bibliotheken und Hilfsfunktionen, die wir benötigen

[1]:
```
import pandas as pd  # Für die Manipulation von Daten in Tabellenform
from IPython.display import display  # Für die visualle Darstellung von Da

import numpy as np  # Für die effiziente Berechnung von unterschiedlichen

from pprint import pprint  # ggf. für eine schönere print-Ausgabe, pprint
from copy import deepcopy # Für das Kopieren von Objekten

from utils import (
    # Hilfsfunktionen für die durchgeführte Rekonstruktion
    get_lepton_names_by_flavour,  # Gibt uns die Namen der Leptonen in ein
    # Funktionssammlung an Filtern, die wir auf die Ereignisse anwenden kö
    Filter,  # Sammlung von Filtern, auf die wir nach
    # Hilfsfunktionen für die Visualisierung
    plot_masses,   # Visualisierung der invarianten Massen
```

CMS Masterclass

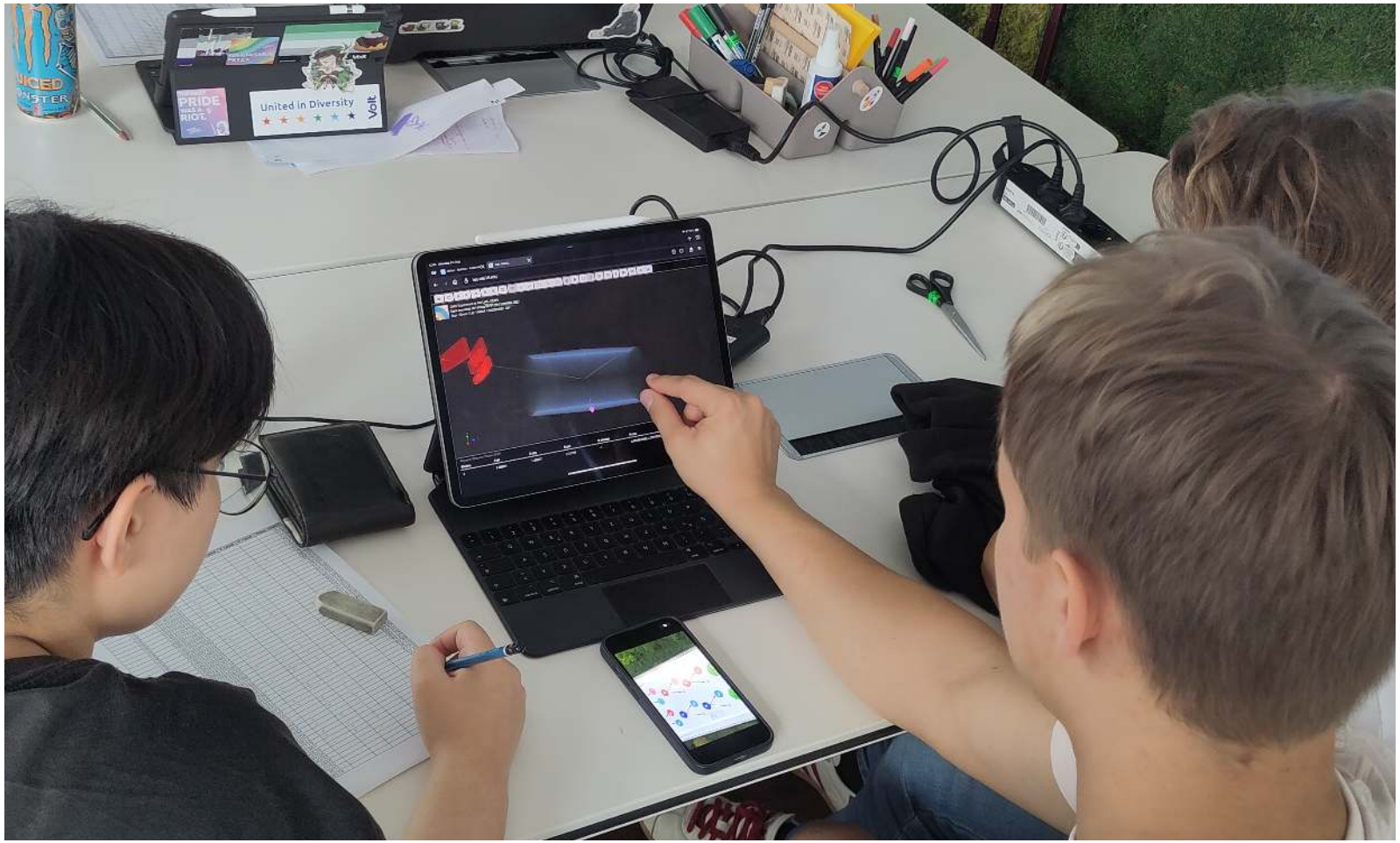
Ablauf

1. Vortrag: Standardmodell & CMS

CMS Masterclass

Ablauf

1. Vortrag: Standardmodell & CMS
2. Eventdisplays kategorisieren



CMS Masterclass

Ablauf

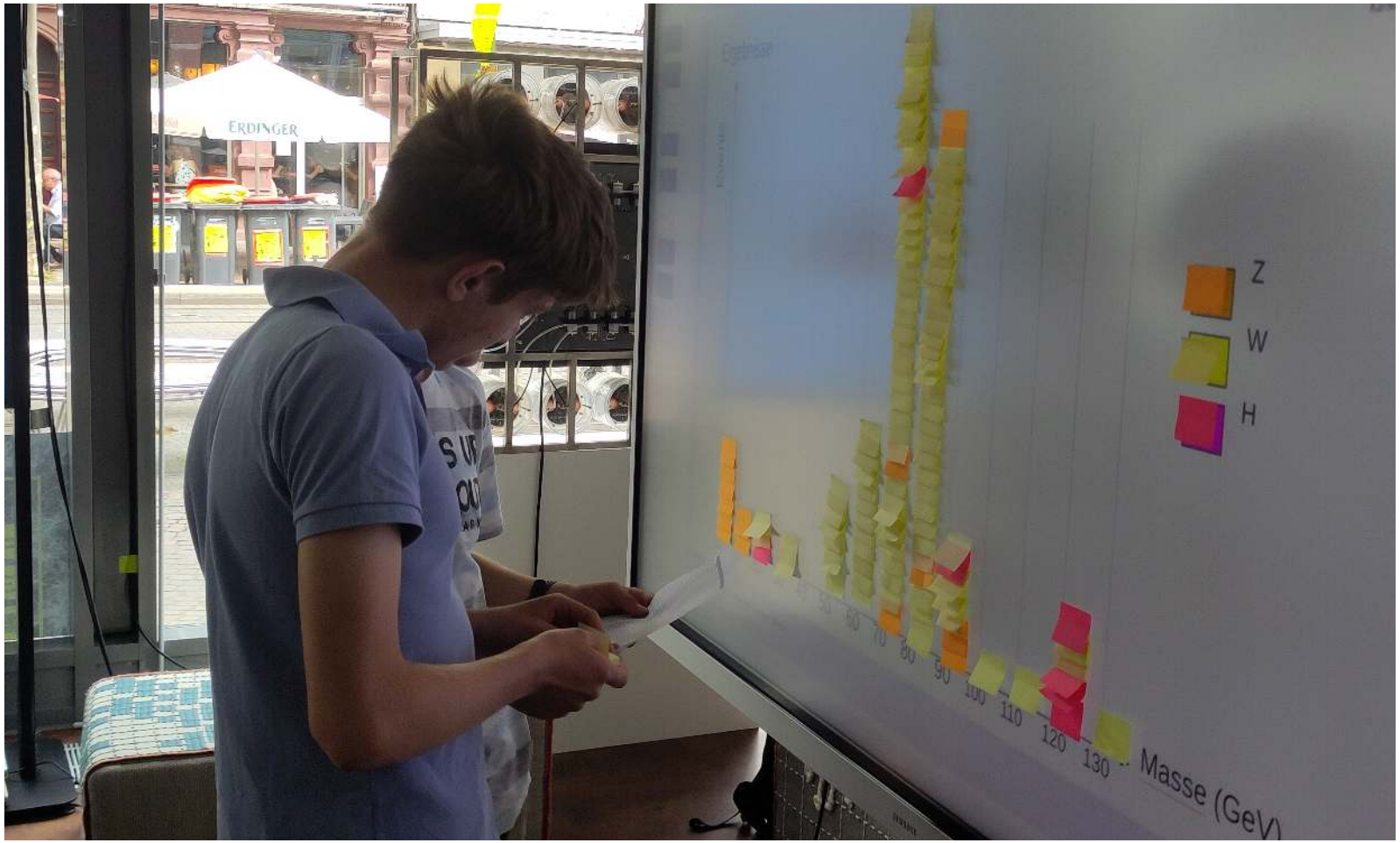
1. Vortrag: Standardmodell & CMS
2. Eventdisplays kategorisieren
3. Histogramm entwickeln

Dauer: $\sim$ 4h

Übermittelte Konzepte

- Klassifikation von Events
- Statistische Idee
- Zählexperiment
- Typische physikalische Größen

Link zum Material



Pythonkurs

1. Notebook: Python Grundlagen

- Kommentare
- Variablen und Datentypen
- Operatoren
- Bedingungen, Schleifen
- Funktionen
- Module
- Zusatzwissen für die Schnellen

Ziele

- Python Grundlagen bereits gesehen
- Python Code verstehen
- Fortgeschrittene: Selbständig Python programmieren

2. Notebook: Statistik in Python

- Zufallsexperiment (Würfelspiel)
- Histogramme
- Normalverteilung
- Matplotlib

Ziele

- Gelerntes anwenden
- Mit Histogrammen arbeiten
- Statistik-Grundlagen

Link zum Kurs

CMS Masterclass 2

Ablauf

- Schrittweise gemeinsam durch die Textpassagen
- In Gruppen durch die Codeblöcke

Inhalt

- Dataframes (Schülergerecht präparierte CMS OpenData)
  - Signal Simulation
  - Untergrund Simulation
  - Messung

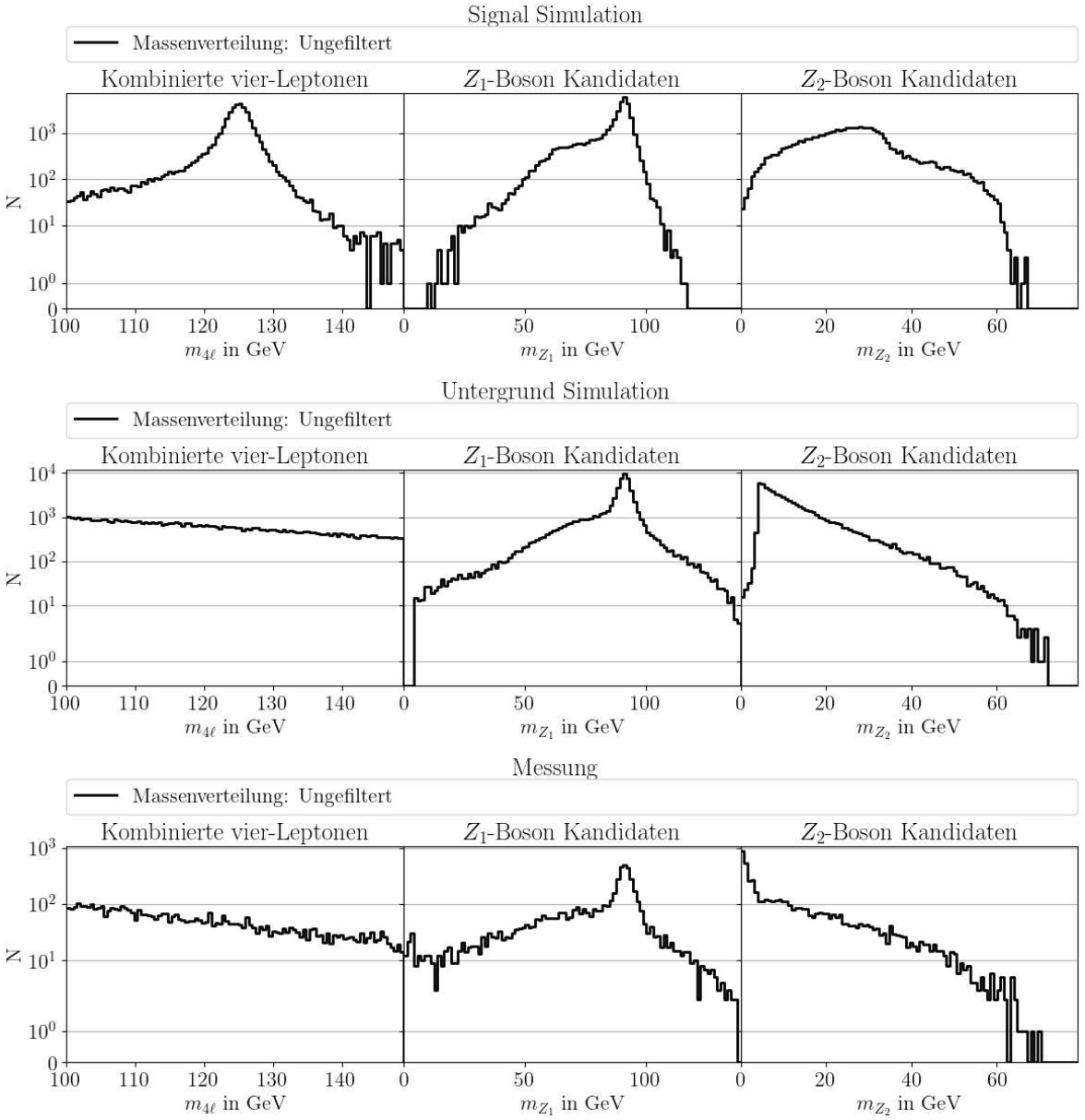
CMS Masterclass 2

Ablauf

- Schrittweise gemeinsam durch die Textpassagen
- In Gruppen durch die Codeblöcke

Inhalt

- Dataframes (Schülergerecht präparierte CMS OpenData)
  - Signal Simulation
  - Untergrund Simulation
  - Messung
- Invariante Masse rekonstruieren (15min)
  - 2 Z-Kandidaten aus 4 Leptonen (Kombinatorik)
  - 1 Higgs-Kandidat



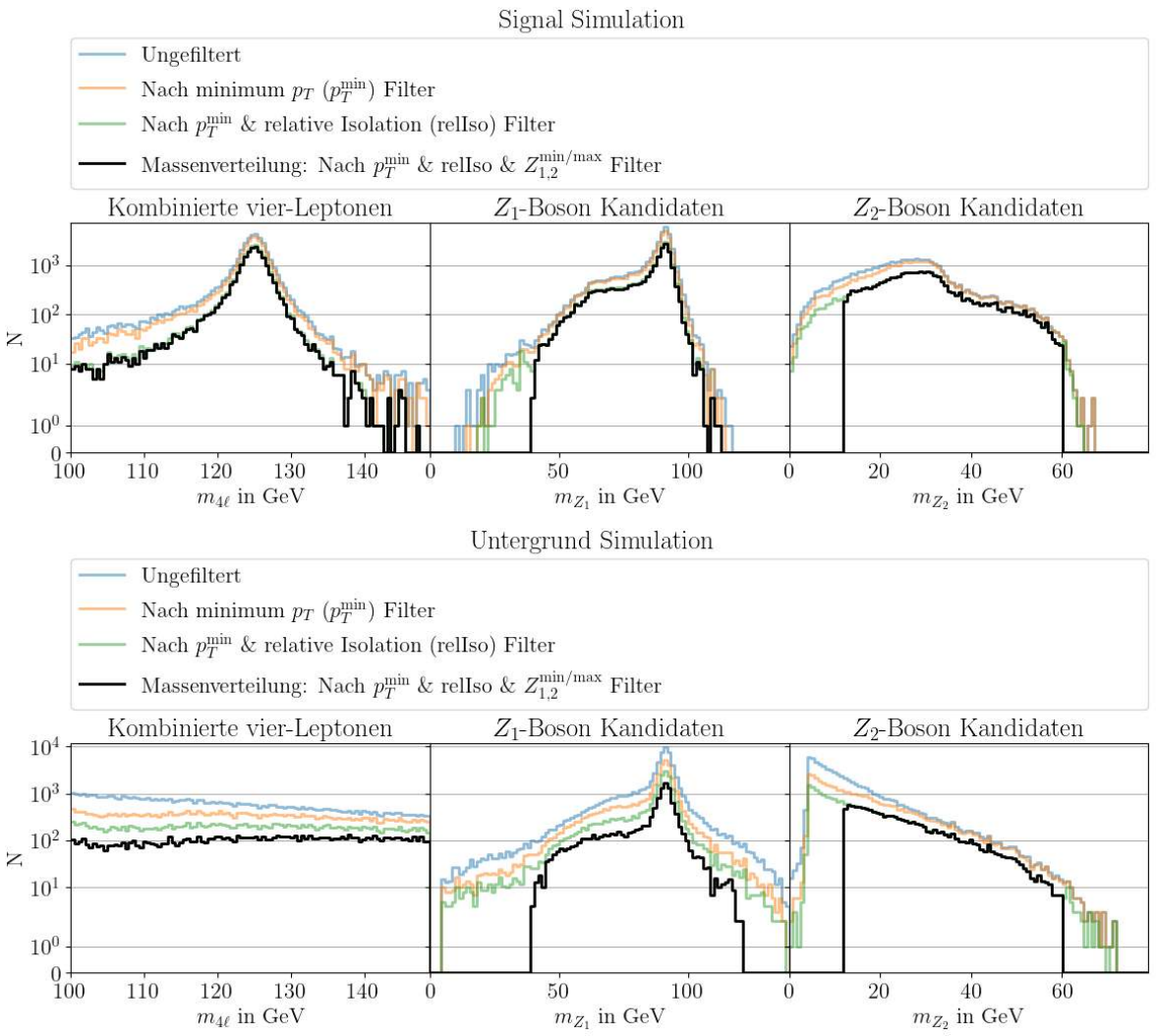
CMS Masterclass 2

Ablauf

- Schrittweise gemeinsam durch die Textpassagen
- In Gruppen durch die Codeblöcke

Inhalt

- Dataframes (Schülergerecht präparierte CMS OpenData)
  - Signal Simulation
  - Untergrund Simulation
  - Messung
- Invariante Masse rekonstruieren (15min)
  - 2 Z-Kandidaten aus 4 Leptonen (Kombinatorik)
  - 1 Higgs-Kandidat
- Schrittweise Selektion auf Simulation anwenden



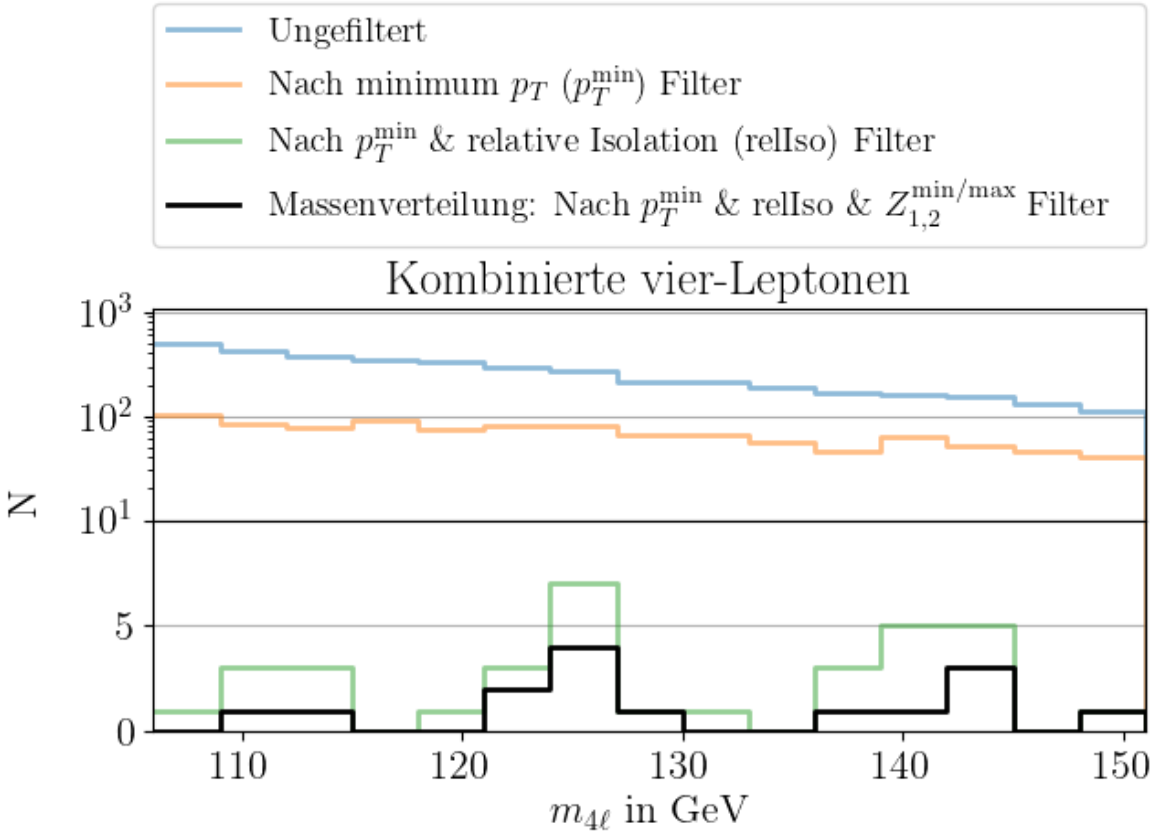
CMS Masterclass 2

Ablauf

- Schrittweise gemeinsam durch die Textpassagen
- In Gruppen durch die Codeblöcke

Inhalt

- Dataframes (Schülergerecht präparierte CMS OpenData)
  - Signal Simulation
  - Untergrund Simulation
  - Messung
- Invariante Masse rekonstruieren (15min)
  - 2 Z-Kandidaten aus 4 Leptonen (Kombinatorik)
  - 1 Higgs-Kandidat
- Schrittweise Selektion auf Simulation anwenden
- Anwenden auf Daten



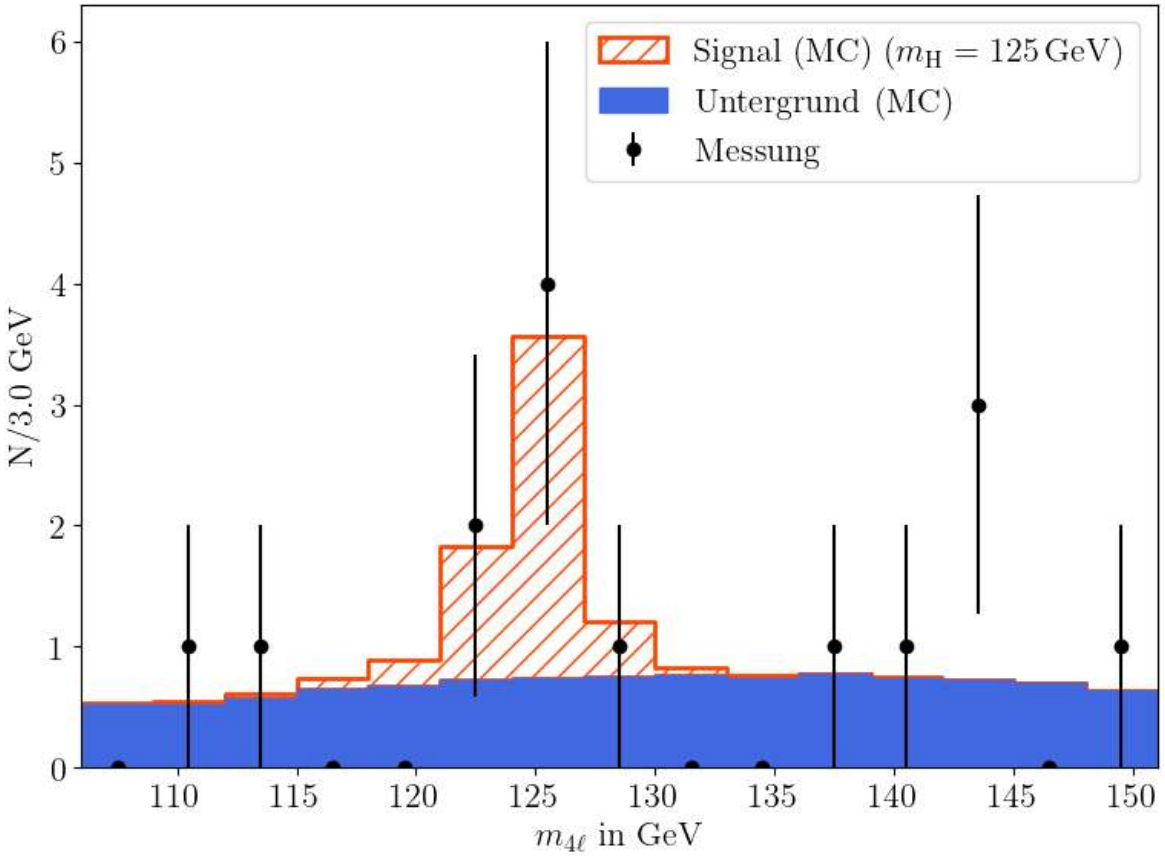
CMS Masterclass 2

Ablauf

- Schrittweise gemeinsam durch die Textpassagen
- In Gruppen durch die Codeblöcke

Inhalt

- Dataframes (Schülergerecht präparierte CMS OpenData)
  - Signal Simulation
  - Untergrund Simulation
  - Messung
- Invariante Masse rekonstruieren (15min)
  - 2 Z-Kandidaten aus 4 Leptonen (Kombinatorik)
  - 1 Higgs-Kandidat
- Schrittweise Selektion auf Simulation anwenden
- Anwenden auf Daten
- Vergleich Messung und Simulation
- Signifikanz berechnen ($\sim 3\sigma$)
- Optimieren



CMS Masterclass 2

Ablauf

- Schrittweise gemeinsam durch die Textpassagen
- In Gruppen durch die Codeblöcke

Inhalt

- Dataframes (Schülergerecht präparierte CMS OpenData)
  - Signal Simulation
  - Untergrund Simulation
  - Messung
- Invariante Masse rekonstruieren (15min)
  - 2 Z-Kandidaten aus 4 Leptonen (Kombinatorik)
  - 1 Higgs-Kandidat
- Schrittweise Selektion auf Simulation anwenden
- Anwenden auf Daten
- Vergleich Messung und Simulation
- Signifikanz berechnen ($\sim 3\sigma$)
- Optimieren

Übermittelte Konzepte

- Cut Based Analysis (Filtern, Signal anreichern)
- Blinde Analyse (Simulation $\rightarrow$ echte Messdaten)
- Signal von Untergrund trennen
- Signifikanz einer Messung

Benötigt

- Laptops
- Jupyter (remote oder lokal)
- Keine Anforderungen an Speicher (< 20 MB)
- Zwei Betreuer auf ca. 8 Teilnehmer:innen
- Beamer

Dauer: $\sim$ 9h

Mehr Infos: DPG 2023

Link zum Kurs

Erfahrungen

Science Camp 2023

- Funktioniert, aber zu wenig Zeit
- Sehr heterogene Gruppe: Mühevoll wenn noch keine Python Kenntnisse

Konsequenzen

- Mehr Zeit und Betreuer eingeplant
- Formulierungen vereinfacht
- Zusatzinfos in ausklappbaren Boxen
- Gemeinsame Besprechung der Textblöcke

Science Camp 2024

- Sehr gut funktioniert und sehr beliebt
- Teilnehmer:innen sehr selbstsicher
- Wenig Zeit für Python Einführung

Fazit

- Sehr positives Feedback von den Teilnehmer:innen (Rating > **4.5/5**)
- Sehr gute Vermittlung der Arbeitsweise in der Teilchenphysik
- Sehr gute Vermittlung von Higgs-Physik
- Heterogene Python Kenntnisse sind eine Herausforderung
- *„Ich werde jetzt definitiv Physik studieren!“*
- **Öffentlich verfügbar und bereit von anderen Gruppen genutzt zu werden**
- **Einfach von anderen Experimenten adaptierbar (Aufbereitung der Daten erforderlich)**