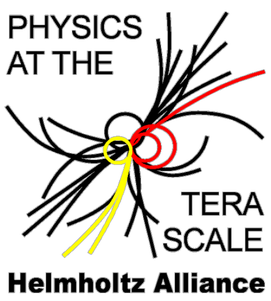
2nd Pan-european Statistics School
28-30 March 2022

INTRODUCTION

Olaf Behnke (DESY)



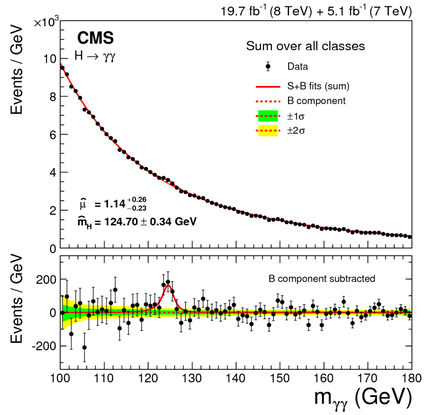
DESY orga team: Carsten Burgard, Alexander Glazov, Dirk Kruecker, Isabell Melzer-Pellmann

Setting the scene

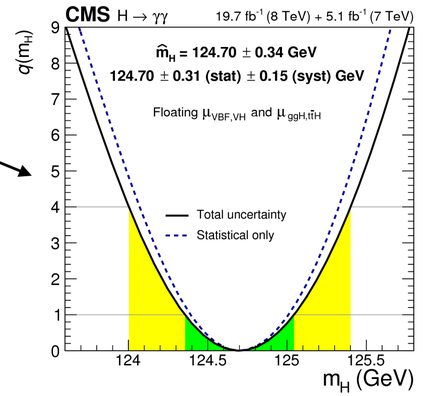
Typical HEP analysis: aiming at measuring a signal s over background b

Analysis flow

- Event Selection
- MVA enhancement of s over b
- Fitting to discriminator spectrum
- Signal inference
- Systematic uncertainties
- Differential signal cross sections
- Data combination



EPJ C74 (2014) 3076



School topics

1. Modeling of data

Relevant in each analysis step

2. Data combination

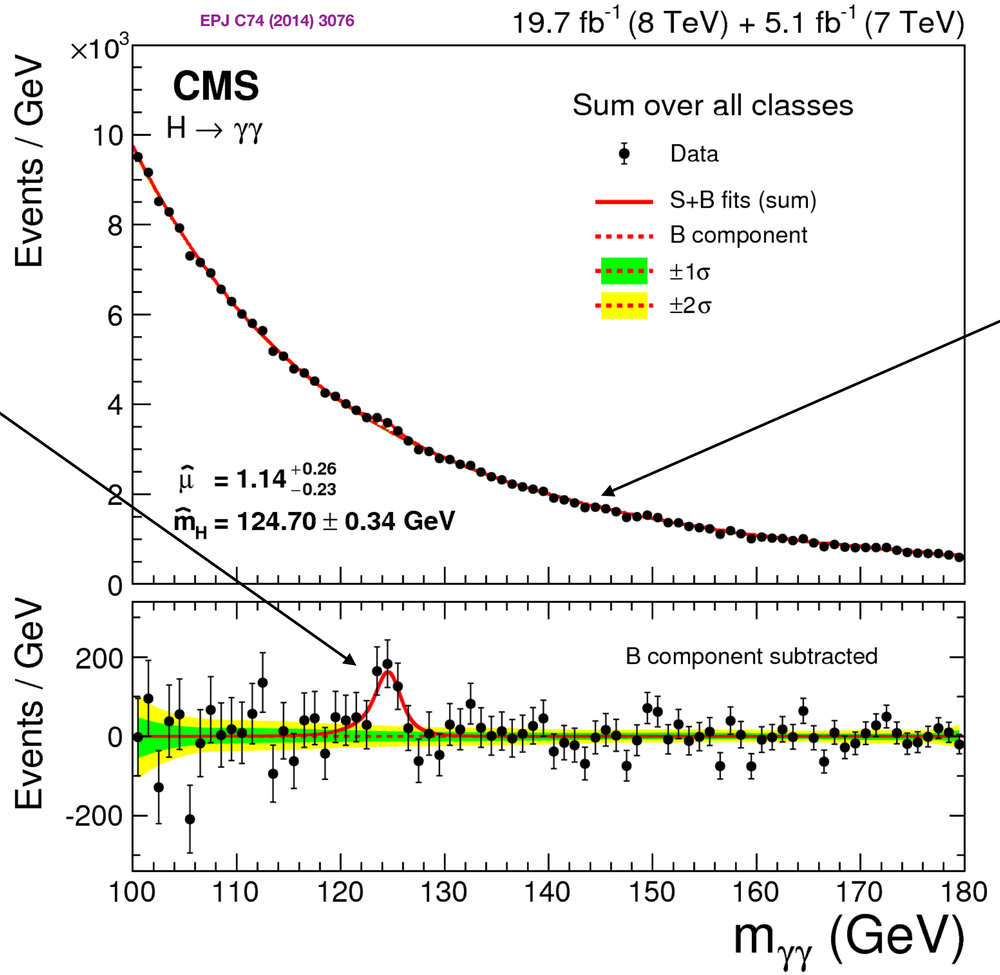
Form the legacy of experiments

Modeling of data

- Modeling of statistical fluctuations: Poisson statistics ✓
- Shape modeling:

Signal peak

Usually well understood



Background distribution

Our main topic here!

School lectures mainly by statisticians, not neccessarily focused on HEP examples

1. Modeling of data

Lectures today 28 March 2-6pm CEST

- W. Rolke: Goodness-of-Fit tests for validating H0 hypothesis ***(e.g. Standard model only background hypothesis in searches for new physics)***

- L. Wasserman: Statistical modelling ***(e.g. parametric vs. non parametric shape modeling)***

Lectures wednesday 30 March 2-6pm CEST

- L. Wasserman: optimal transport ***(e.g. for template morphing of background shapes vs. systematics nuisance parameters)***

- M. Kuusela: gaussian processes ***(interpolation/smoothing e.g. for multi-dimensional background template modeling vs. nuisance parameters)***

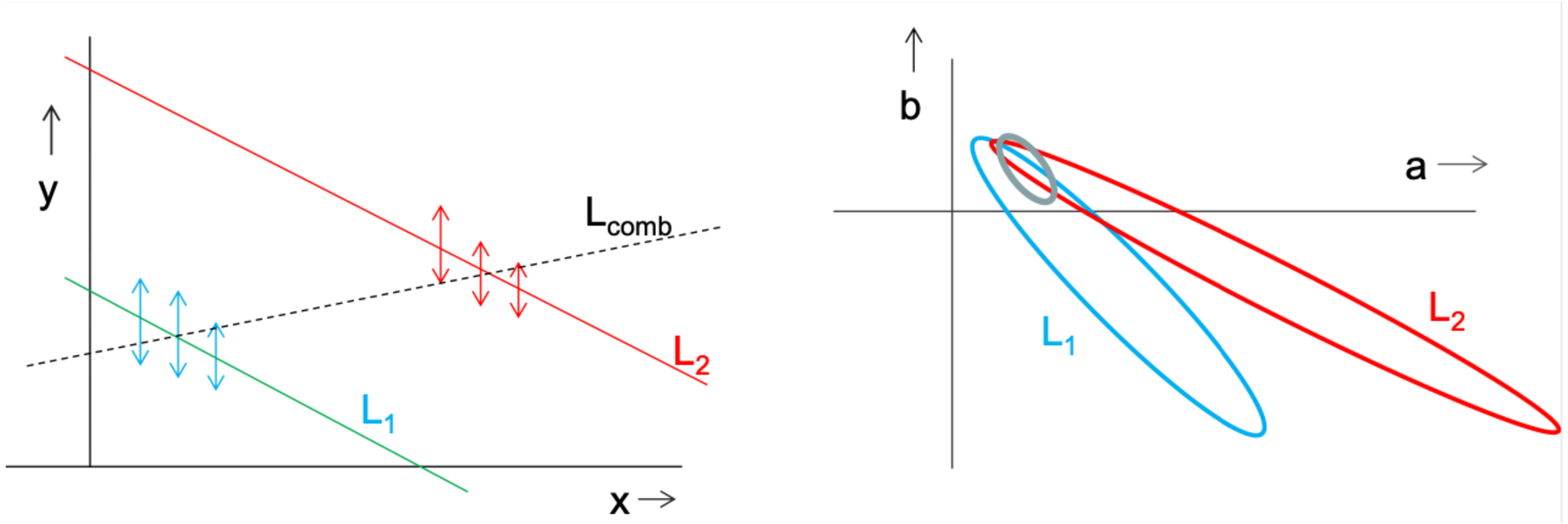
- R. Balasubramanian: EFT Lagrangian Morphing ***(signal template morphing vs. Strength of new physics contact interactions)***

2. Data combination

* 2-D fit of straight line y = a + bx

  a = parameter of interest, b = nuisance param

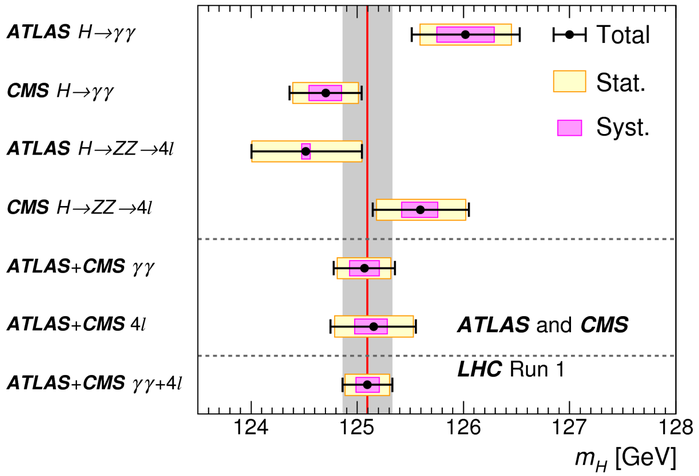
* Track hits in 2 subdetectors, each of 3 planes



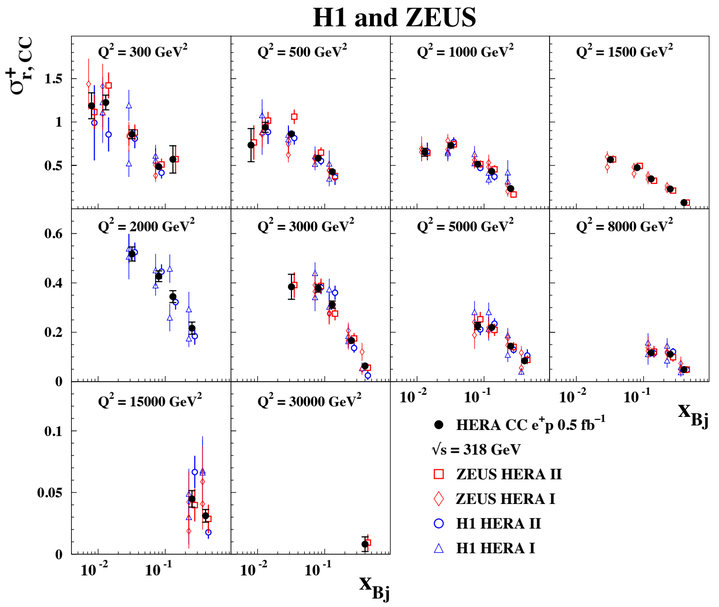
👉 Combining data with complementary info can vastly reduce uncertainties

2. Data combination

Phys. Rev. Lett. 114 (2015) 191803



EPJ C75 (2015) 580



Key issue for any combination: quantify systematic uncertainty correlations of the input measurements

Lectures tomorrow 29 March 2-6pm CEST

- G. Cowan: Data combination introduction **(from basics to errors on errors)**
- J. Kieseler: Data combination in practice
- Carsten Burgard & Lukas Heinrich: what to publish **(to preserve measurement information for the future)**

Zoom session handling

- Per lecture 15+15 mins time for questions+discussion
- Please ask questions also during the lectures!
- For questions and comments:
  - During the lecture: please write in the chat box, we will interrupt the lecturer for answering the question
  - At the end of the lecture: write in the chat box and/or raise your hand ✋

For questions/comments/suggestion on the school please write to olaf.behnke@desy.de

Enjoy the school!