Off-shell Higgs into 4 leptons & electron tracking in ATLAS

Arnaud Maury

PhD supervised by **David Rousseau**

PHENIICS Fest 2022

Laboratoire des 2 Infinis Irène Joliot-Curie
Orsay

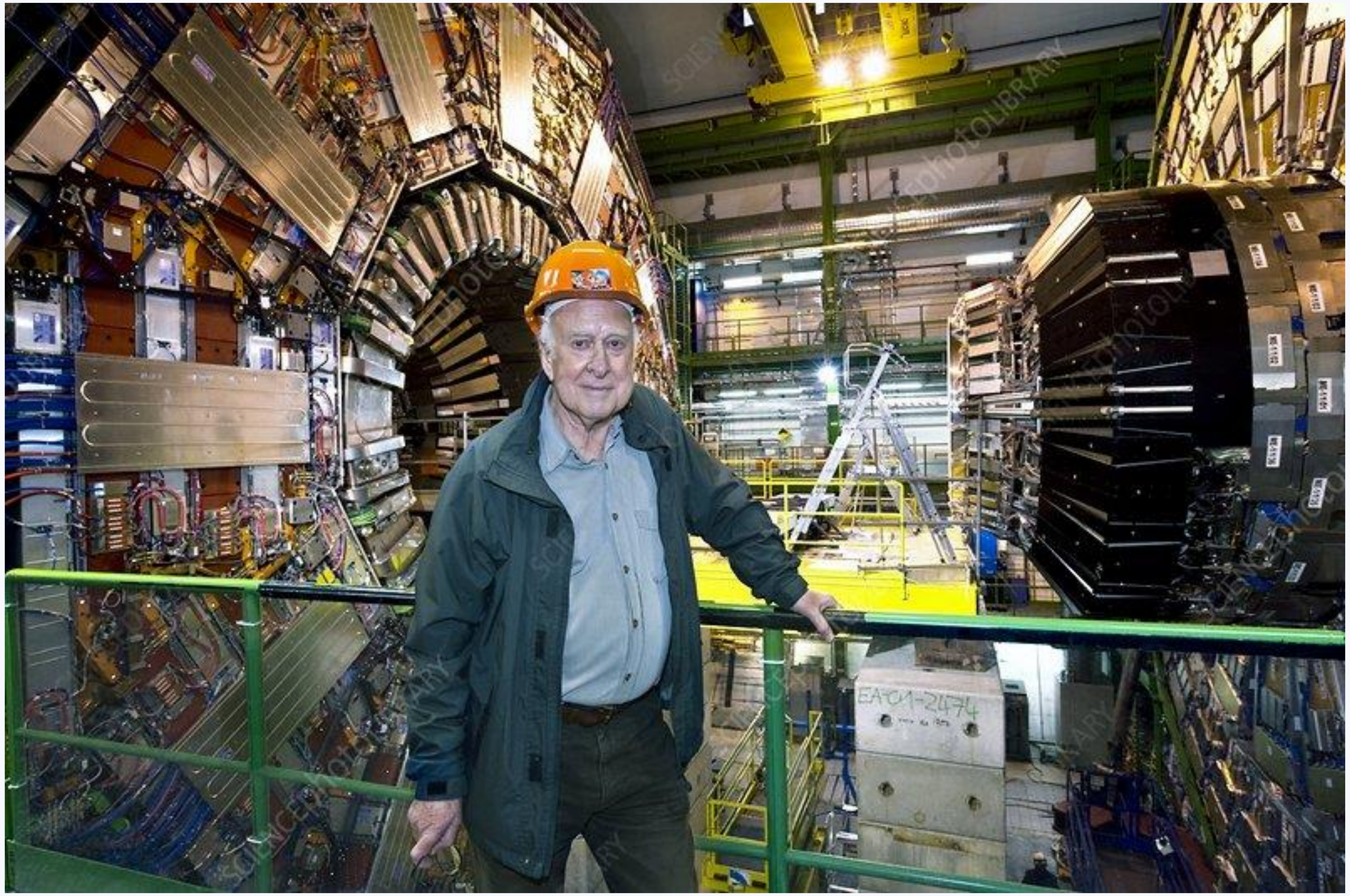
Off-shell Higgs into 4 leptons

- EFT Monte-Carlo generation of off-shell Higgs:
  - Higgs: a bit of history
  - H4l: the golden channel
  - Off-shell Higgs
  - What is SMEFT?
  - Basics of Monte-Carlo generation for HEP
  - Validation plots

- Simulation-based inference (with ML)

Higgs: a bit of history

Discovered at CERN in 2012!



H4l decay mode

- H -> ZZ* -> $\ell^+\ell^-\ \ell^+\ell^-$ (⚠ : ℓ = e or μ)

- Called the "*golden channel*"
  - yields a very clean signal
  - is one of the Higgs decay modes with the best resolution

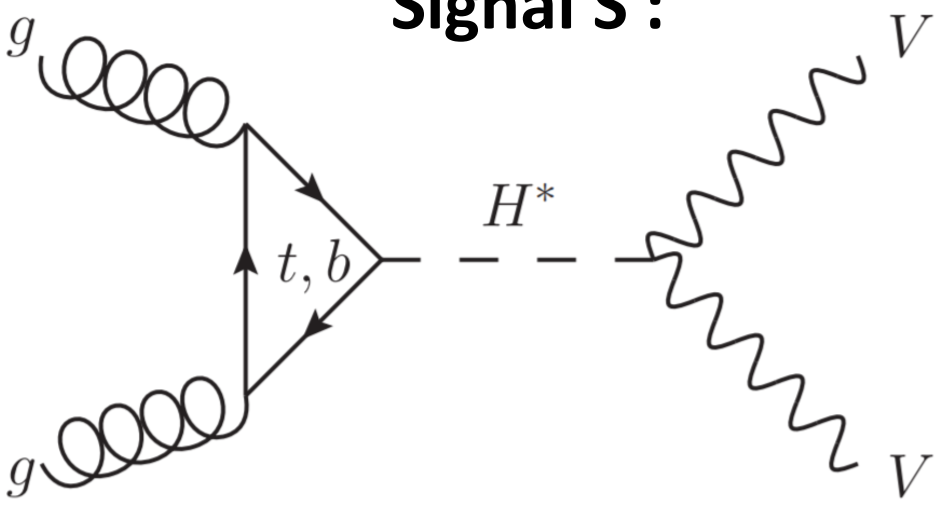
- Downside: low number of events
  - BR(H -> ZZ) = 2.6% (in comparison, BR(H -> bb) = 58%, BR(H -> W^+W^-) = 21%)
  - BR(Z -> l^+l^-) = 6.7%

  (at m_H = 125 GeV)

IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

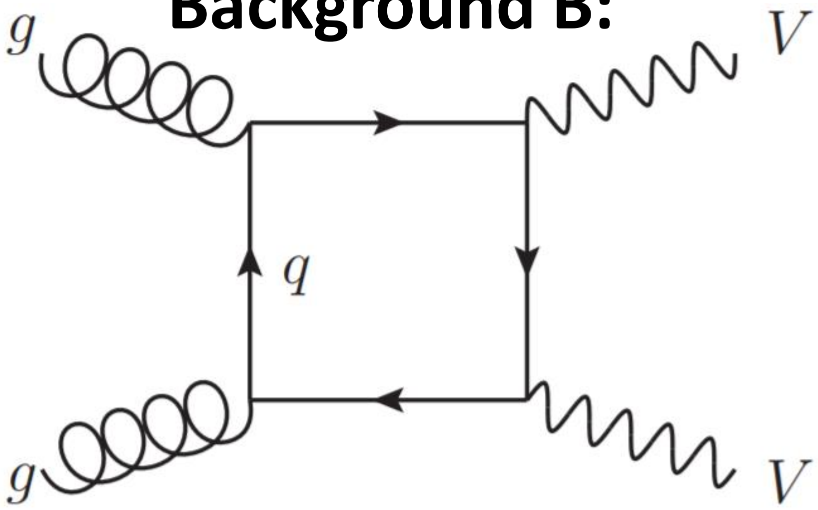
Interference and background(s)

Signal S :



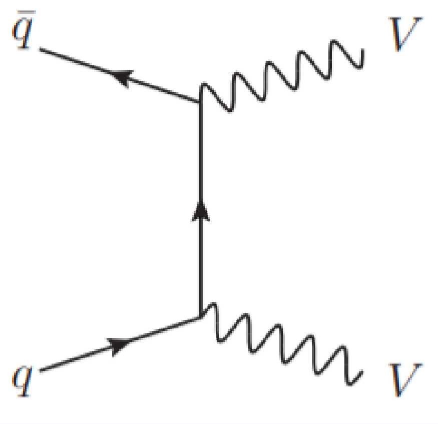
+

Background B:



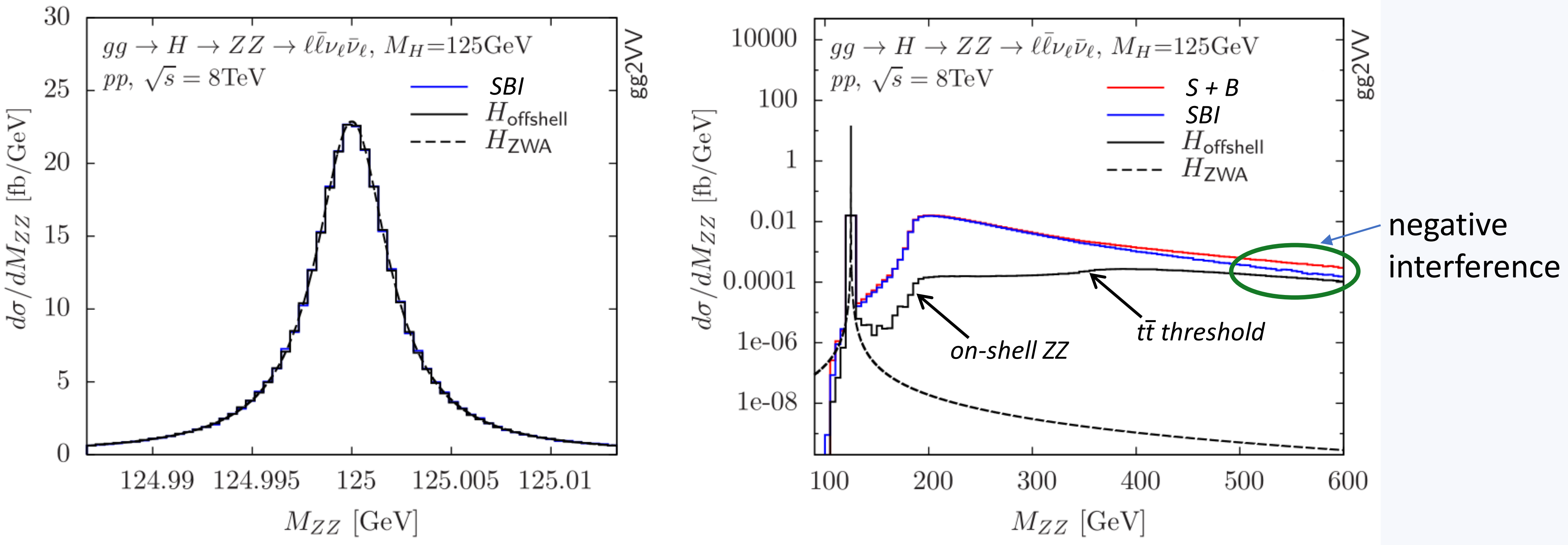
Total signal SBI (negative interference)

+ Background qqZZ:



Off-shell Higgs

- *High-Mass (HM) region:* m_{4l} > 220 GeV
- Off-shell Higgs production is not negligible (kinematics: on-shell ZZ at 180 GeV, $t\bar{t}$ threshold at ~350 GeV).
- But still, m_H = 125 GeV while Γ_H = 4.1 MeV…
- Zero-Width Approximation works well around 125 GeV, but not in the High-Mass region



(a) Near the Higgs resonance (b) In the high-mass region

Inadequacy of zero-width approximation for a light Higgs boson signal, N. Kauer, G. Passarino, arXiv:1206.4803 **[hep-ph]**

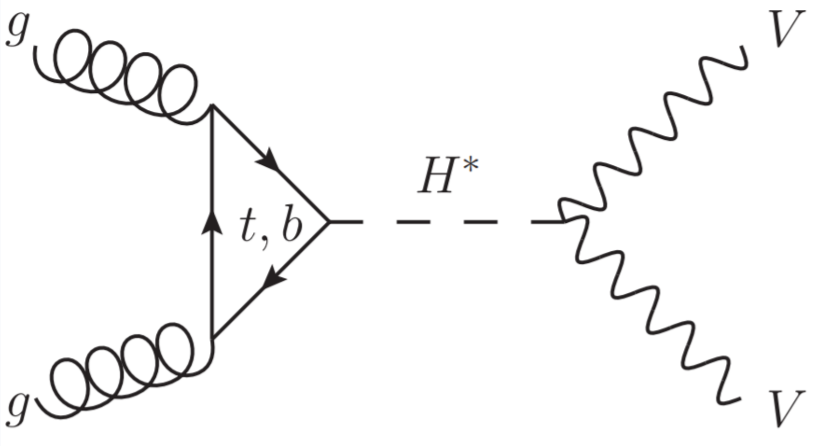
SMEFT: an Effective Field Theory for the SM

- $$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{1}{\Lambda}\sum_i c_i^{(5)} O_i^{(5)} + \frac{1}{\Lambda^2}\sum_i c_i^{(6)} O_i^{(6)} + \dots$$
- C_i : Wilson coefficients O_i : Wilson operators
- We treat the SM as an effective theory.
- We only consider dimension-6 operators.
- **Considered operators :**
  - O_pg, O_pt, O_tp

Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

Monte-Carlo generation 101

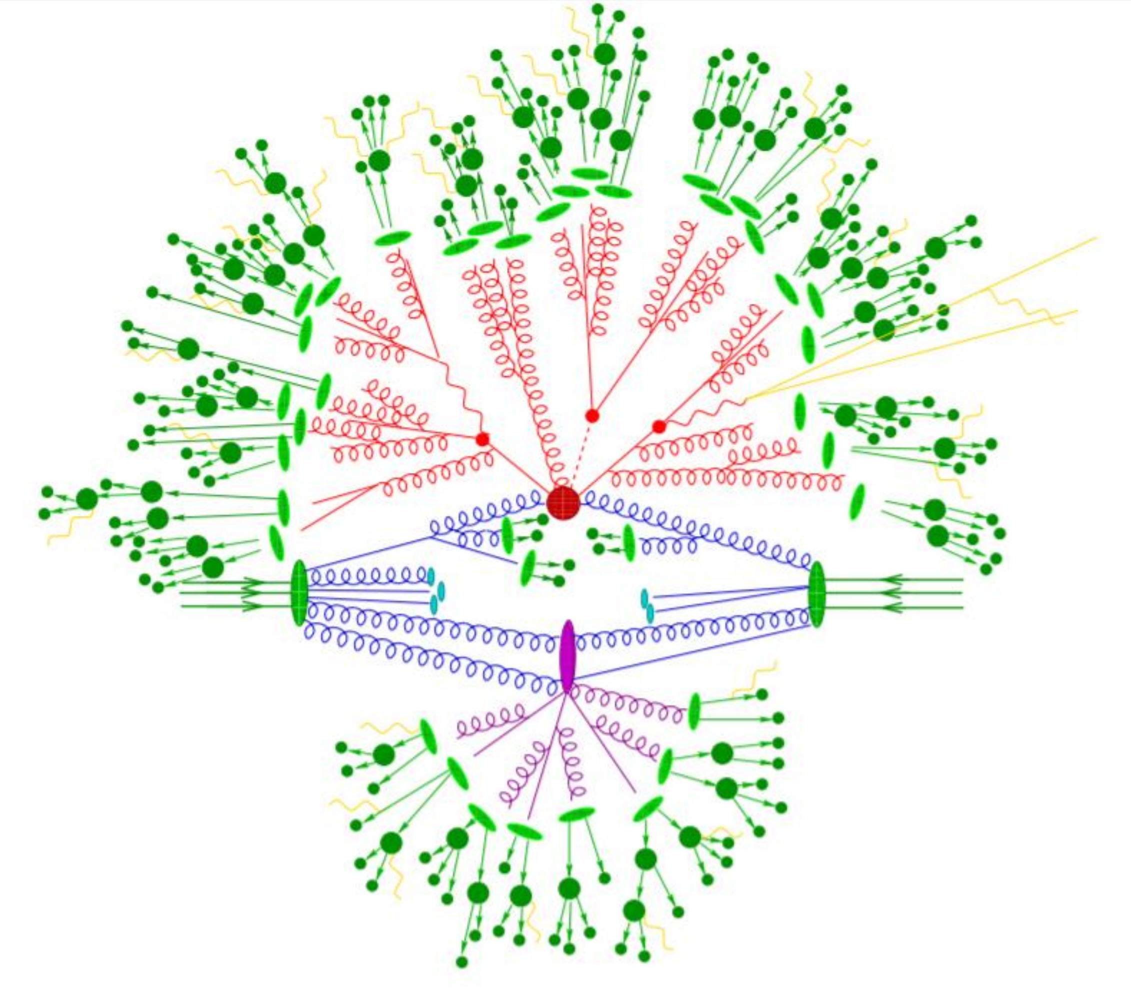
- Parton level process (hard scattering) : MadGraph



- Parton shower (Pythia)

Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

Monte-Carlo generation 101



Introduction to parton-shower event generators, Stefan Höche, *arXiv:1411.4085* **[hep-ph]**

Validation plots

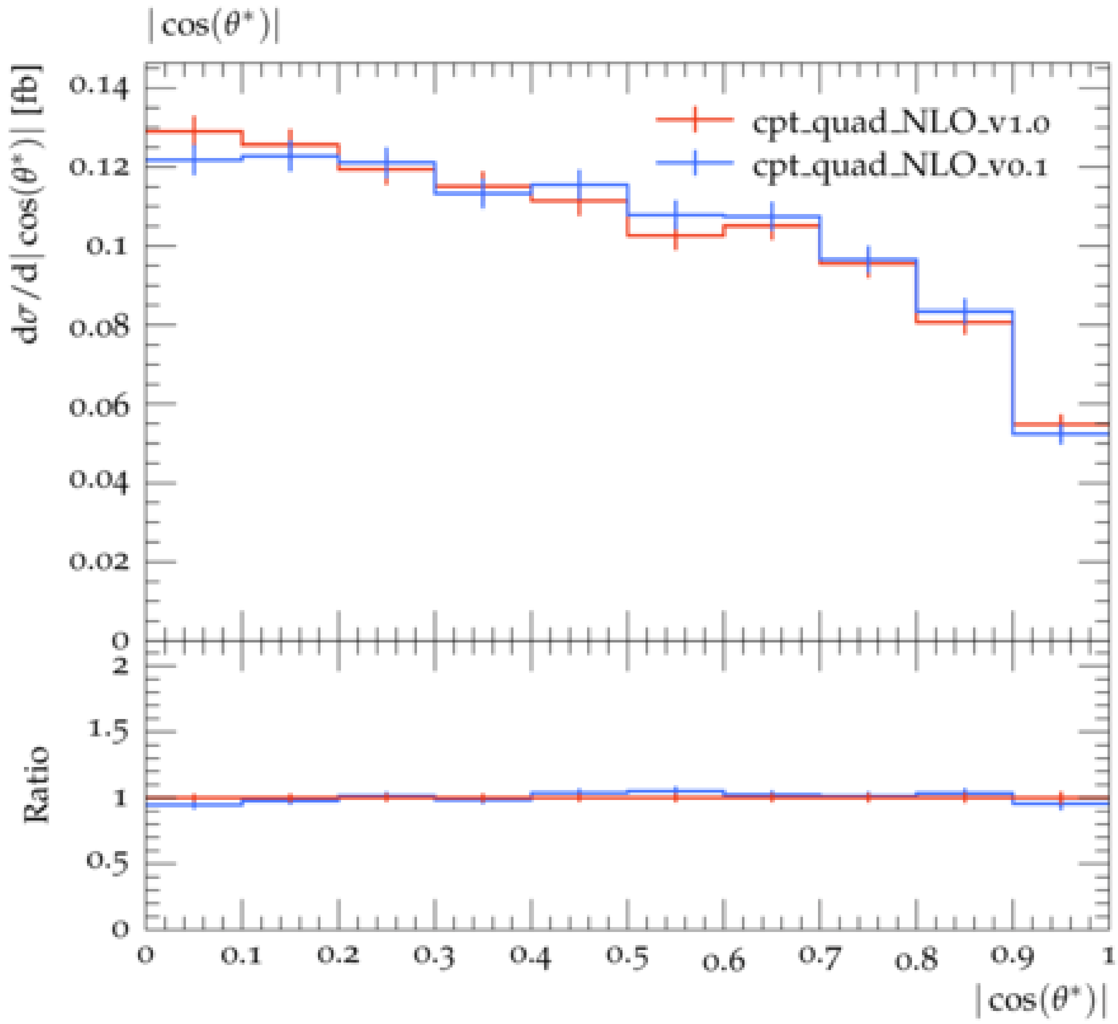
- The main goal: compare two versions of the « SMEFTatNLO » model (v0.1 vs v1.0).
- Two versions of MadGraph were finally used: MGv2.7.3 and MGv2.9.3
- Plots for these observables: **$|\cos\theta^*|$**, $|y_{4l}|$, $\cos(\theta_1)$, $\cos(\theta_2)$, m_{4l}, $p_T(4l)$, Φ
- Relevant EFT operators: Opt, Otp, OpG

Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

SMEFTatNLO v1.0 vs v0.1

Same observable

With MadGraph v2.9.3 (release 21.6.71) *Observable: |cos θ*|* *Operator: **cpt*** *Quadratic contribution*

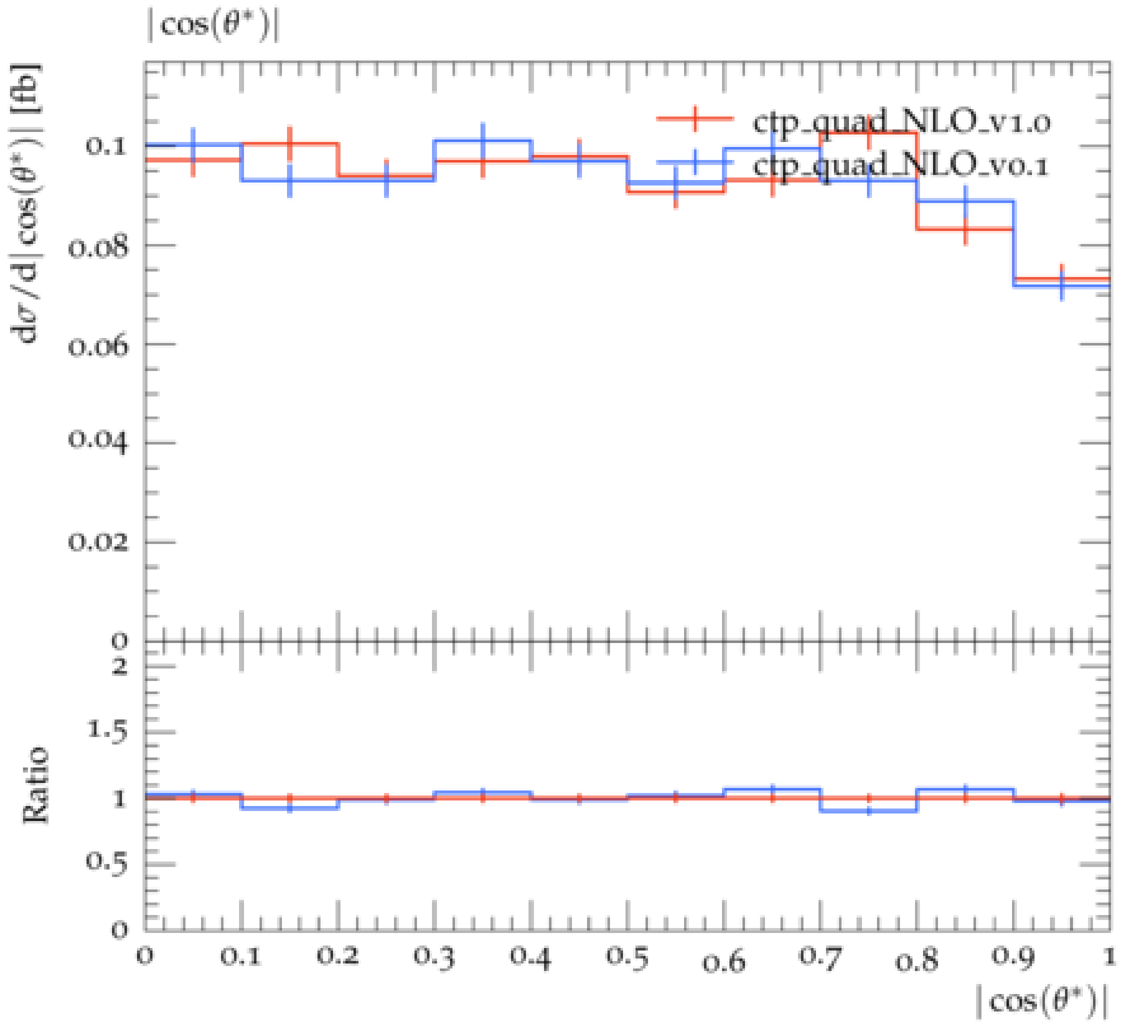


IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

SMEFTatNLO v1.0 vs v0.1

Same observable

With MadGraph v2.9.3 (release 21.6.71) *Observable:* ***|cos θ*|*** *Operator:* ***ctp*** *Quadratic contribution*



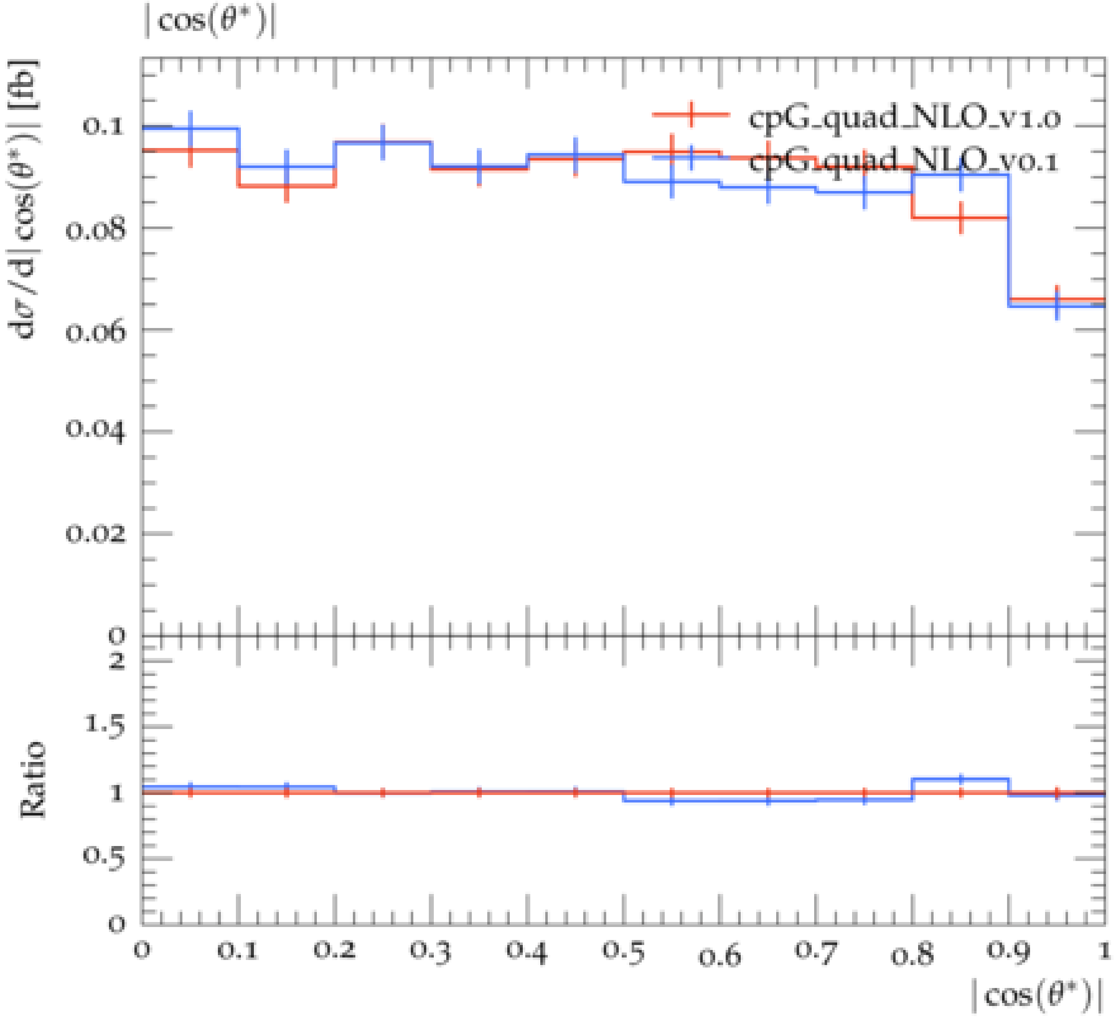
IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

SMEFTatNLO v1.0 vs v0.1

Same observable

With MadGraph v2.9.3 (release 21.6.71)

Observable: ***|cos θ*|*** *Operator:* ***cpG*** *Quadratic contribution*



Laboratoire de Physique des 2 Infinis Irène Joliot-Curie

Simulation-based inference

What we are trying to achieve:

Theory parameters:

- Signal strength μ
- Couplings of the Higgs with other particles
- EFTs: Wilson coefficients
- …

Physical observables:

- Fireworks

IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

Simulation-based inference

What we are trying to achieve:

Theory parameters:

- Signal strength μ
- Couplings of the Higgs with other particles
- EFTs: Wilson coefficients
- …

Physical observables:

- ~~Fireworks~~
- Mass, energy, trajectories of particles (interaction with the detector)

Simulation-based inference

What we are trying to achieve:

Theory parameters:

- Signal strength μ
- Couplings of the Higgs with other particles
- EFTs: Wilson coefficients
- …

Simulation →

Physical observables:

- ~~Fireworks~~
- Mass, energy, trajectories of particles (interaction with the detector)

IJCLab Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

Simulation-based inference

What we are trying to achieve:

Theory parameters:

- Signal strength μ
- Couplings of the Higgs with other particles
- EFTs: Wilson coefficients
- …

Simulation →

← **Likelihood**

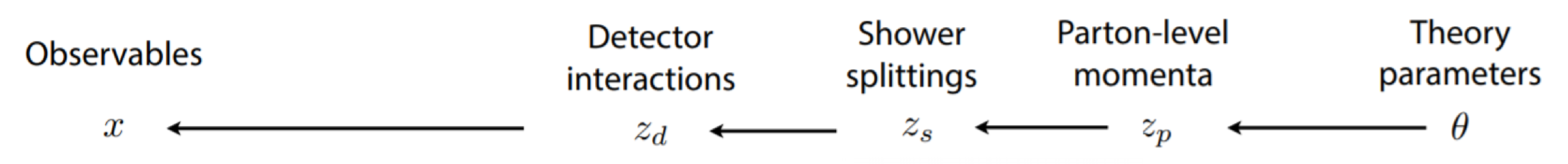
Physical observables:

- ~~Fireworks~~
- Mass, energy, trajectories of particles (interaction with the detector)

Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

Simulation-based inference

- Intractable problem (like in thermodynamics, epidemiology, meteorology, etc)
- Here, the likelihood function is **not tractable**, because of the high dimension of the problem (dimensionality curse), non-analytical processes (e.g. parton shower), incomplete computation (only upto NLO diagrams), experimental effects…



- Main idea: train DNNs to estimate several probability ratios

Mixture Model (1/2)

- *"Mixture model"* idea: train 3 different DNNs to estimate 3 probability ratios:
  - ➢ SBI vs S: $\frac{p(x|SBI)}{p(x|S)}$
  - ➢ B vs S: $\frac{p(x|B)}{p(x|S)}$
  - ➢ qqZZ vs S: $\frac{p(x|qqZZ)}{p(x|S)}$.

➢ Reminder: There are 4 processes studied here: S, B, SBI and qqZZ.

IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

Mixture Model (2/2)

- *“Mixture model”* idea:
  re-express the likelihood ratio into those 3 probability ratios.

3 ratios: $\frac{p(x|SBI)}{p(x|S)}$, $\frac{p(x|B)}{p(x|S)}$ *and* $\frac{p(x|qqZZ)}{p(x|S)}$.

$$\frac{p(\vec{x}|\theta)}{p(\vec{x}|1)} = \frac{\sigma_{\text{obs}}(1)}{\sigma_{\text{obs}}(\theta)} \frac{(\theta^2-\theta)\,\sigma_{\text{S}} + \theta\,\sigma_{\text{SBI}}\frac{p(x|SBI)}{p(x|S)} + (1-\theta)\,\sigma_B\frac{p(x|B)}{p(x|S)} + \sigma_{\text{qqZZ}}\frac{p(x|qqZZ)}{p(x|S)}}{\sigma_{\text{SBI}}\frac{p(x|SBI)}{p(x|S)} + \sigma_{\text{qqZZ}}\frac{p(x|qqZZ)}{p(x|S)}}$$

Simulation-based inference applied to off-shell H4l to find 1 parameter (signal strength θ): work by Jay Sandesara

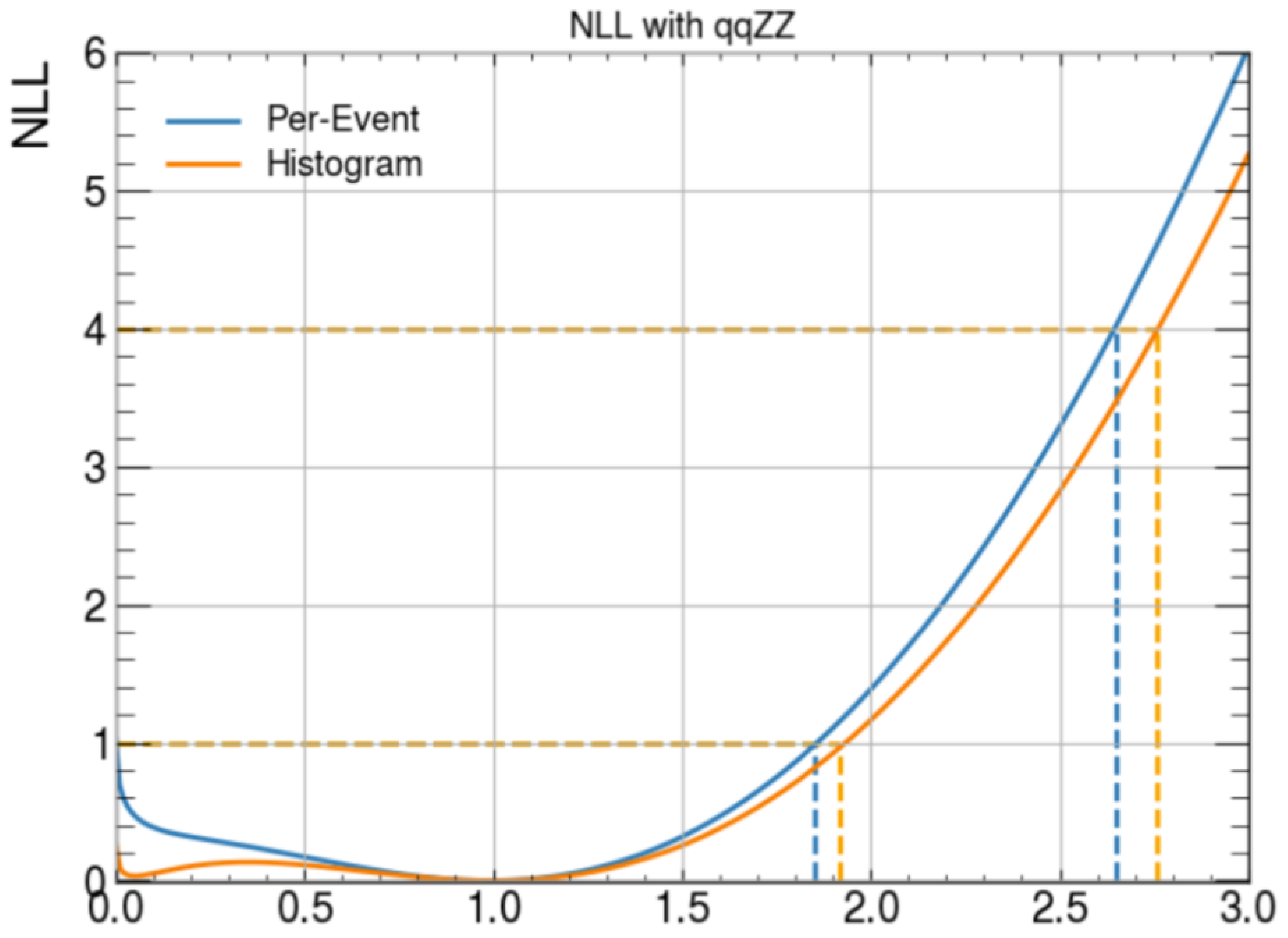
3 ratios: $\frac{p(x|SBI)}{p(x|S)}$, $\frac{p(x|B)}{p(x|S)}$ *and* $\frac{p(x|qqZZ)}{p(x|S)}$.

$$\frac{p(\vec{x}|\theta)}{p(\vec{x}|1)} = \frac{\sigma_{\text{obs}}(1)}{\sigma_{\text{obs}}(\theta)} \frac{(\theta^2 - \theta)\,\sigma_S + \theta\,\sigma_{\text{SBI}} \frac{p(x|SBI)}{p(x|S)} + (1-\theta)\,\sigma_B \frac{p(x|B)}{p(x|S)} + \sigma_{\text{qqZZ}} \frac{p(x|qqZZ)}{p(x|S)}}{\sigma_{\text{SBI}} \frac{p(x|SBI)}{p(x|S)} + \sigma_{\text{qqZZ}} \frac{p(x|qqZZ)}{p(x|S)}}$$

$\frac{p(x|\theta)}{p(x|1)}$ is our main goal: the likelihood ratio for the parameter strength.

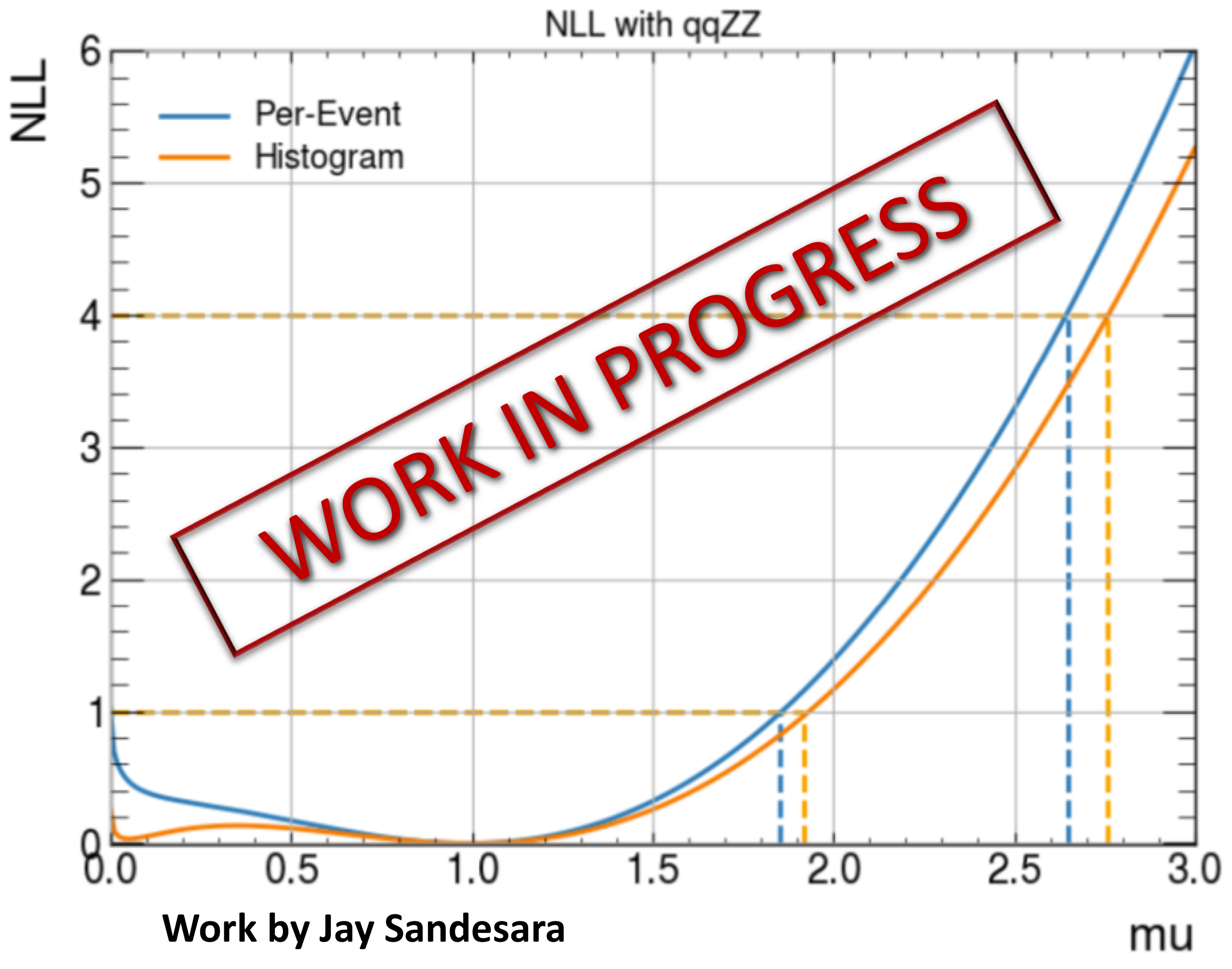
This is what we can and will use to analyse experimental data (one day).

Comparison between "conventional" histogram method and this NN/statistics method:



Work by Jay Sandesara

Comparison between "conventional" histogram method and this NN/statistics method:



Work by Jay Sandesara

- Need to take into account systematics
- NN training: the problem of negative weights
- This method still is not very stable: either too much bias or too much variance…

IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

Electron tracking in ATLAS (ITk)

- ITk, a new all-silicon tracker for ATLAS:
  - Tracking: what is a track?
  - How to find tracks
  - The problem of electrons (spoiler alert, it's because of bremsstrahlung)

- The ACTS software
  - Adaptation to ITK and integration to Athena
  - Integration of a new track fitting algorithm (GSF)
  - Pull plots, a useful diagnostic tool

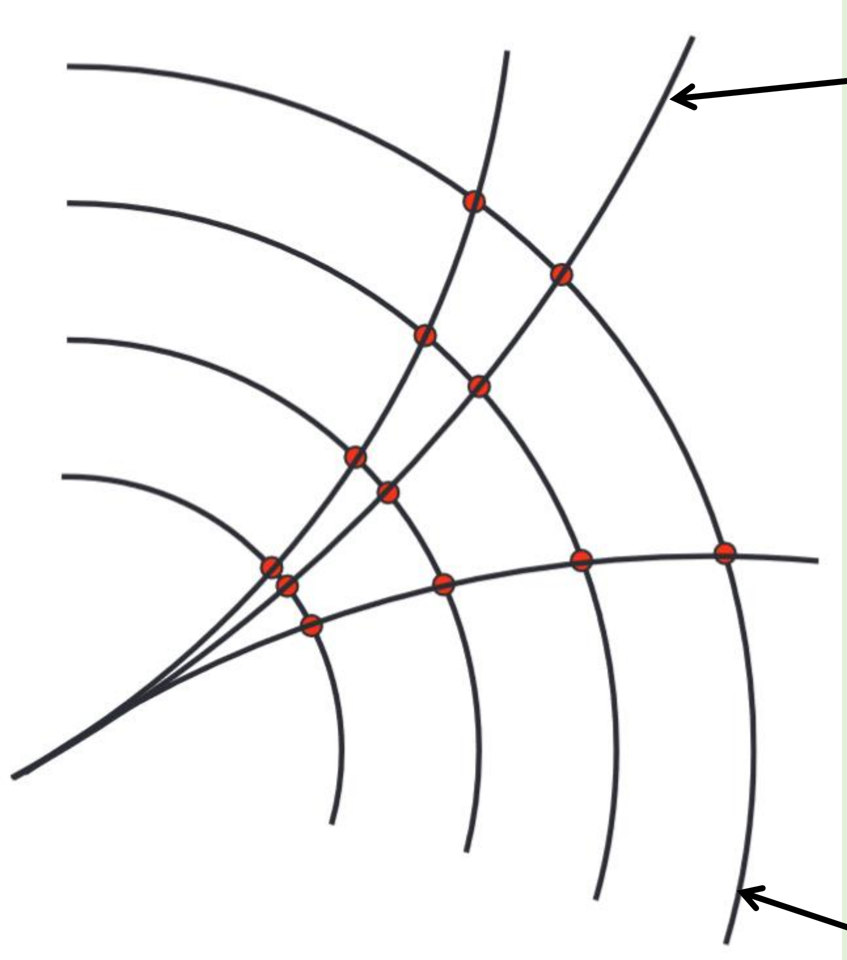
IJCLab Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

The ACTS software

- **A Common Tracking Software**
- Originates from the indigenous tracking software developed at ATLAS.
- Was then restarted from scratch as an experiment-independent software
- The project now:
  - integrate to ACTS a new track fitting algorithm (GSF);
  - integrate it back to ATLAS (for ITk);
  - **Validation:** Physics must be accurate, execution should be efficient.

IJCLab
Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

What is a track?



Trajectory of the particle: a track

Detector layer

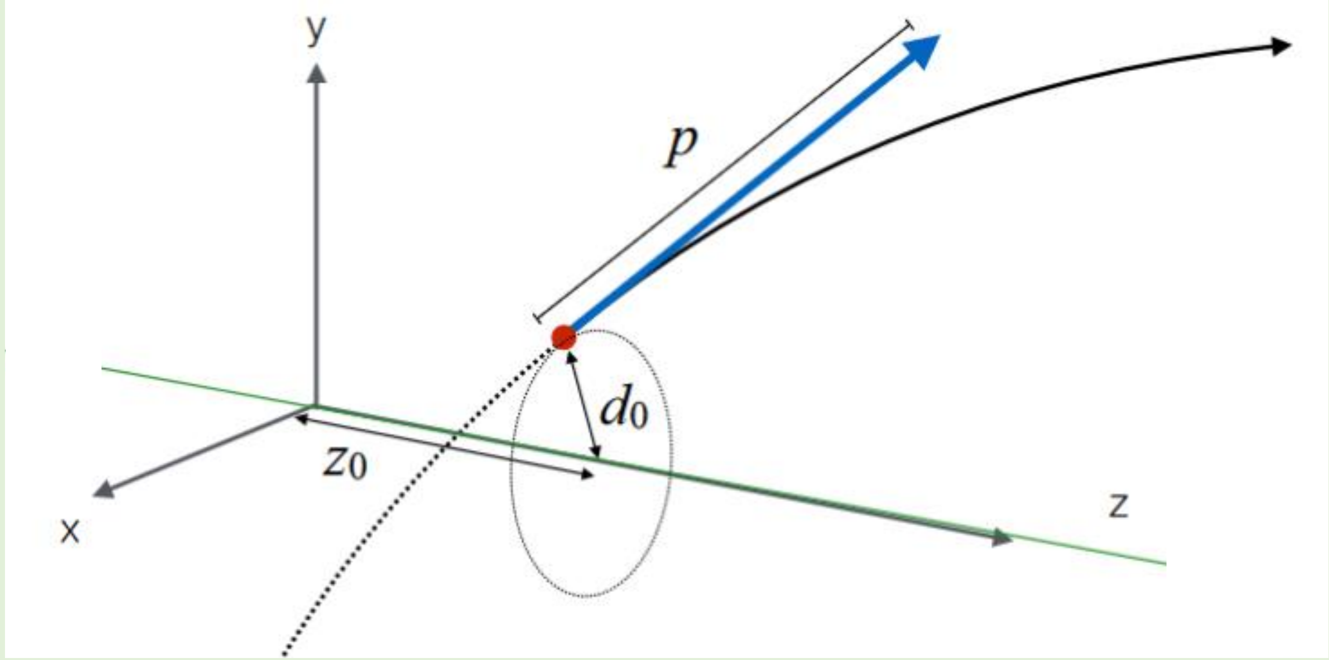
We can only measure points, at the intersection of the track and the detectors !

IJCLab Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis

Figure from Andreas Salzburger (Track and Vertex Reconstruction - HCPSS Aug 11-22, 2014)

What is a track?

Each track is (uniquely) defined by 5 parameters:



- d0
- z0
- θ : angle with the beam axis
- ϕ : angle in the normal plan of the beam axis
- q/p : electric charge over momentum

Figure from Andreas Salzburger (Track and Vertex Reconstruction - HCPSS Aug 11-22, 2014)

How to find tracks

From those points, we must find a track. But this is not always trivial…

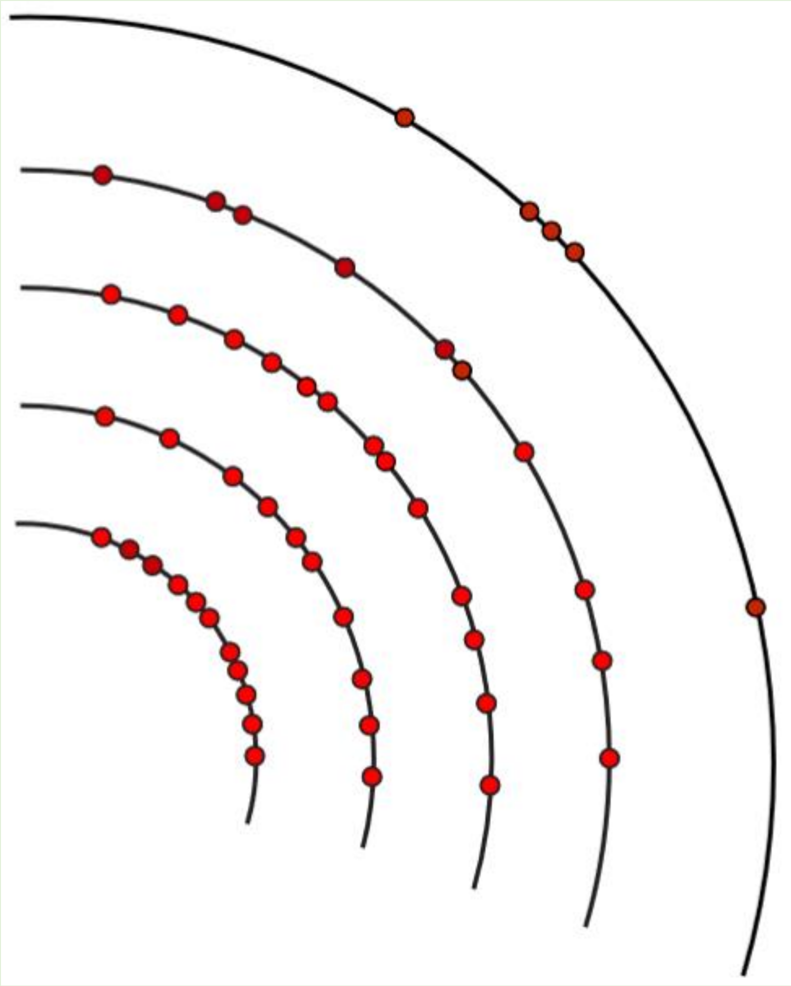


Figure from Andreas Salzburger (Track and Vertex Reconstruction - HCPSS Aug 11-22, 2014)

Basic idea:

- Find a seed (a crude first guess)
- Refine this "idea" by propagation (B field map, magnetic force, Runge-Kutta…)
- Do a chi-square minimization
- *Current default method: Combinatorial Kalman Filter (**CKF**)*

The problem of (low energy) electrons

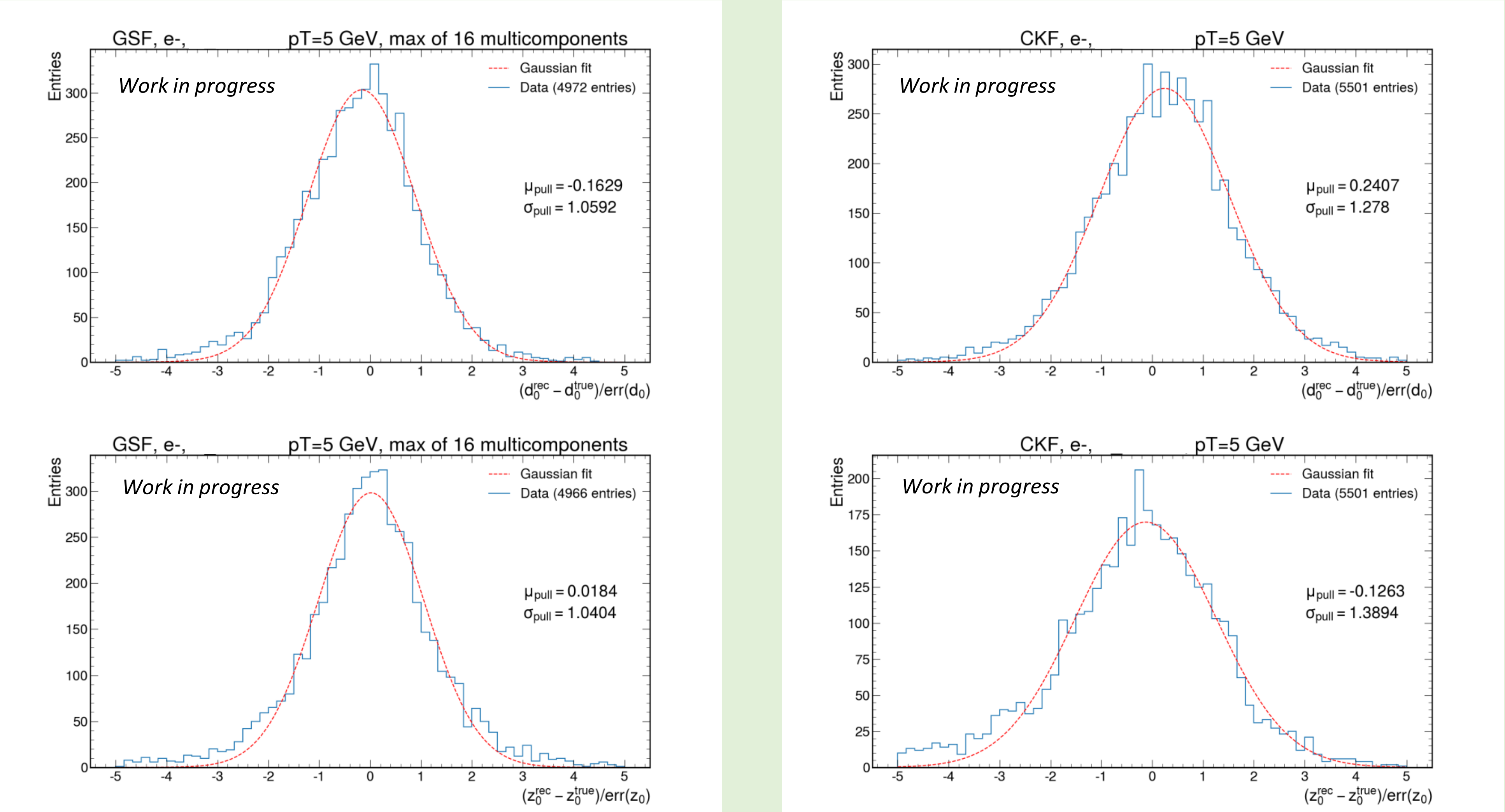
- At “low” energies (in the order of 5 to 10 GeV), electrons radiate and lose a lot of energy by bremsstrahlung.
- Reminder: the curvature is inversely proportional to the momentum.
- Consequence: the trajectory gets more curved than expected.
- Solution: use a fitting algorithm that can take this phenomenon into account and better fit electron tracks: **Gaussian Sum Filtering (GSF)**

IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis

Pull plots

- Pull defined as: (true_value – fitted_value) / error_value
- For each track, for each parameter (one of 6), we have a pull value.
- Plot the histogram. Must be ideally:
  - a gaussian
  - centred
  - a standard deviation of 1.

Electrons: GSF vs CKF, p_T=5 GeV



Better performance from the GSF (left)

Conclusion

- There are interesting new studies to be done in the **off-shell Higgs**’ study!
- Monte-Carlo generation is challenging especially for EFT samples.
- But MC is our main source of data when developing new stat/ML techniques or tracking algorithms.
- With HL-LHC coming up, aggressive R&D is a must. Computational progress will not be able to keep up.

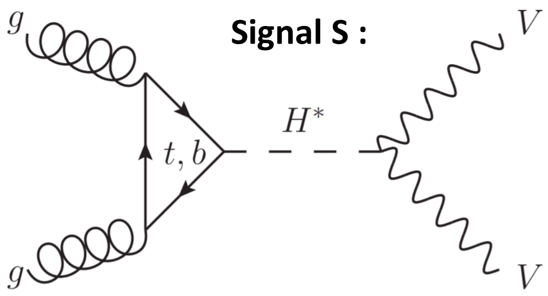
- Run 3 and the ITk upgrade are coming up soon.
- Work to integrate a new track fitting algorithm (GSF) to ACTS.
- Work to make ACTS work for ITk.
- In progress: validate ACTS GSF vs the existing ATLAS GSF in Athena, the software versioning framework of ATLAS.

Thank you for your attention

Backup

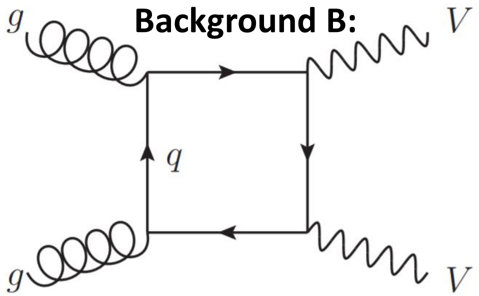
Interference and background(s)

Signal S :



+

Background B:



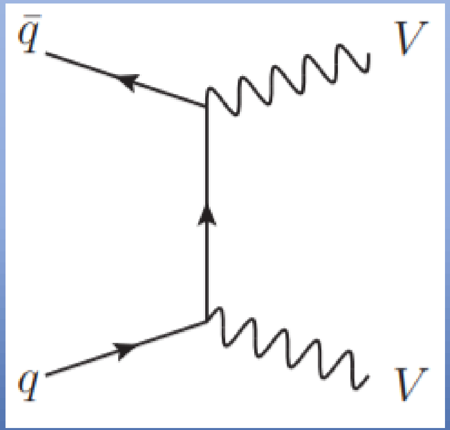
Total signal SBI (negative interference)

$$|\mathcal{M}_{\rm SBI}|^2 = |\mathcal{M}_S + \mathcal{M}_B|^2$$
$$= |\mathcal{M}_S|^2 + 2\cos(\theta)\cdot|\mathcal{M}_S\mathcal{M}_B| + |\mathcal{M}_B|^2$$

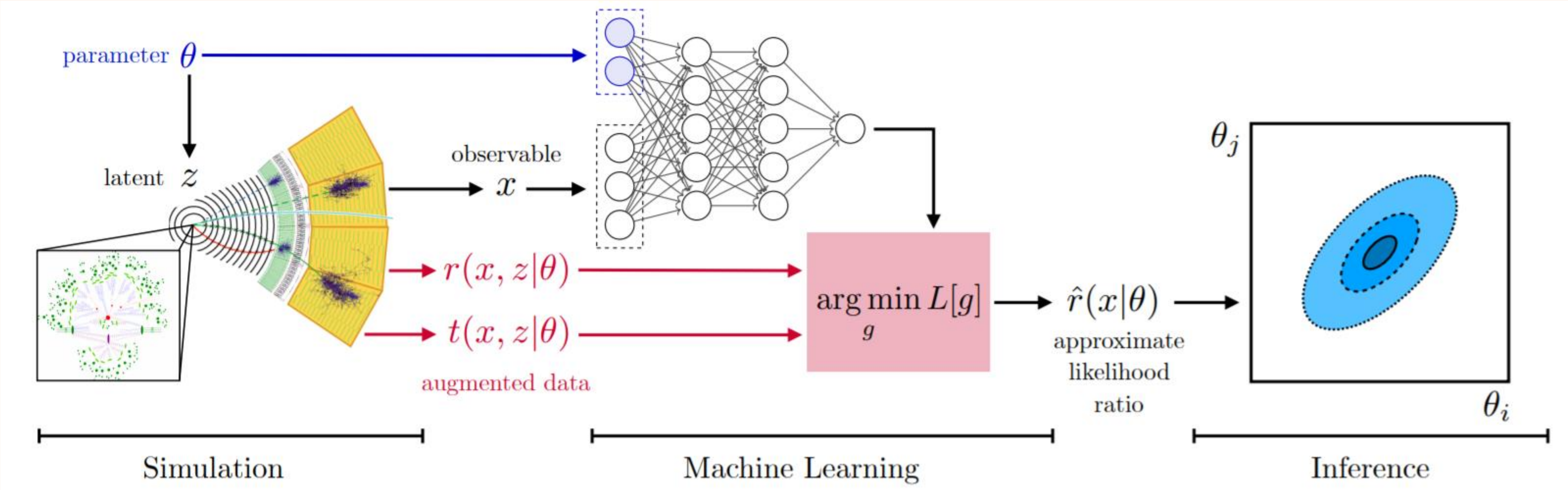
$$SBI = S + I + B$$

Note that $\boldsymbol{I}$ can be **negative** (here, it is the case).

+ Background B2:



Simulation-based inference



Constraining Effective Field Theories with Machine Learning,
J. Brehmer, K. Cranmer, G. Louppe and J. Pavez
arXiv:1805.00013 **[hep-ph]**

IJCLab Irène Joliot-Curie
Laboratoire de Physique des 2 Infinis