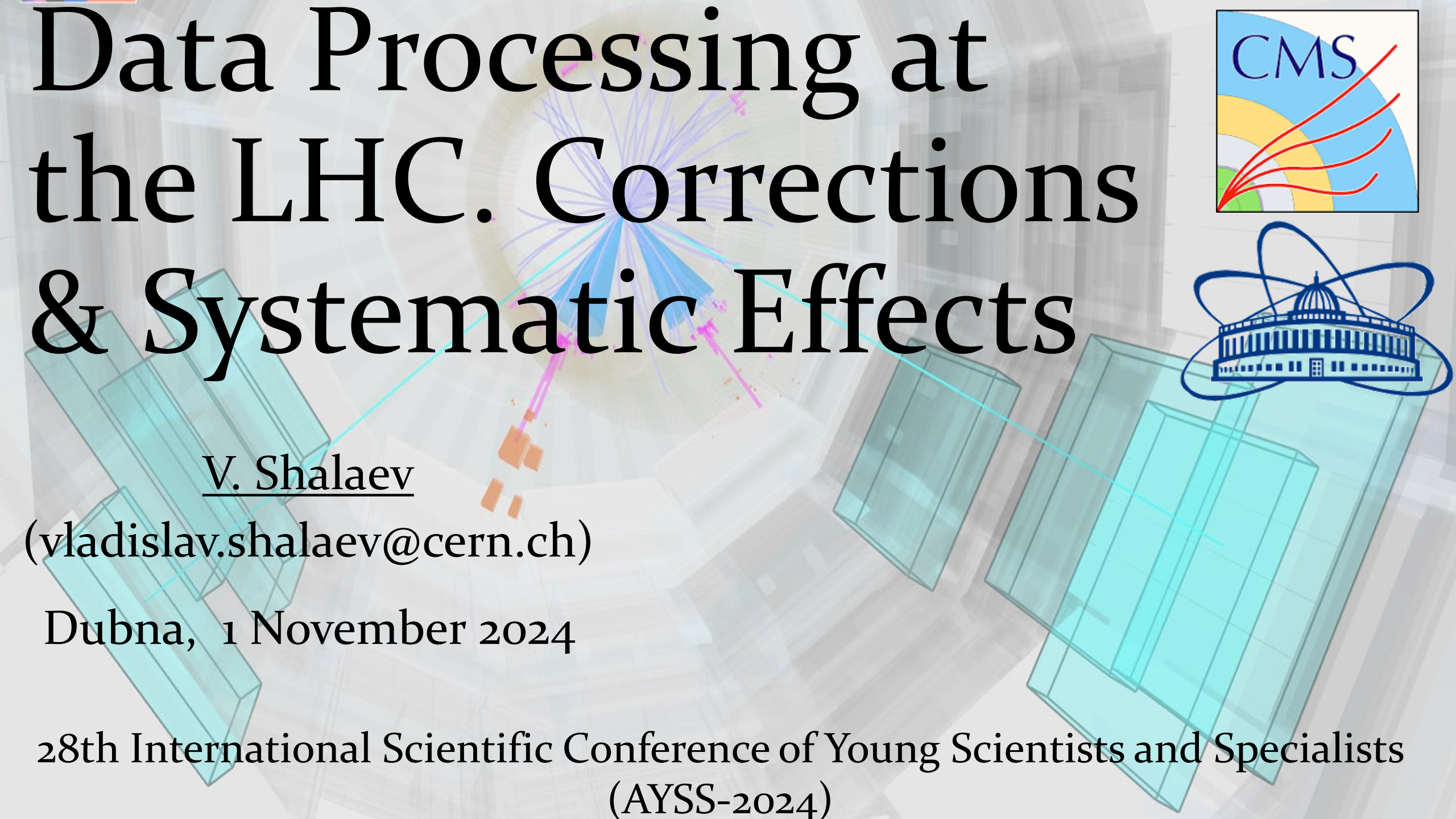
Data Processing at the LHC. Corrections & Systematic Effects



V. Shalaev

(vladislav.shalaev@cern.ch)

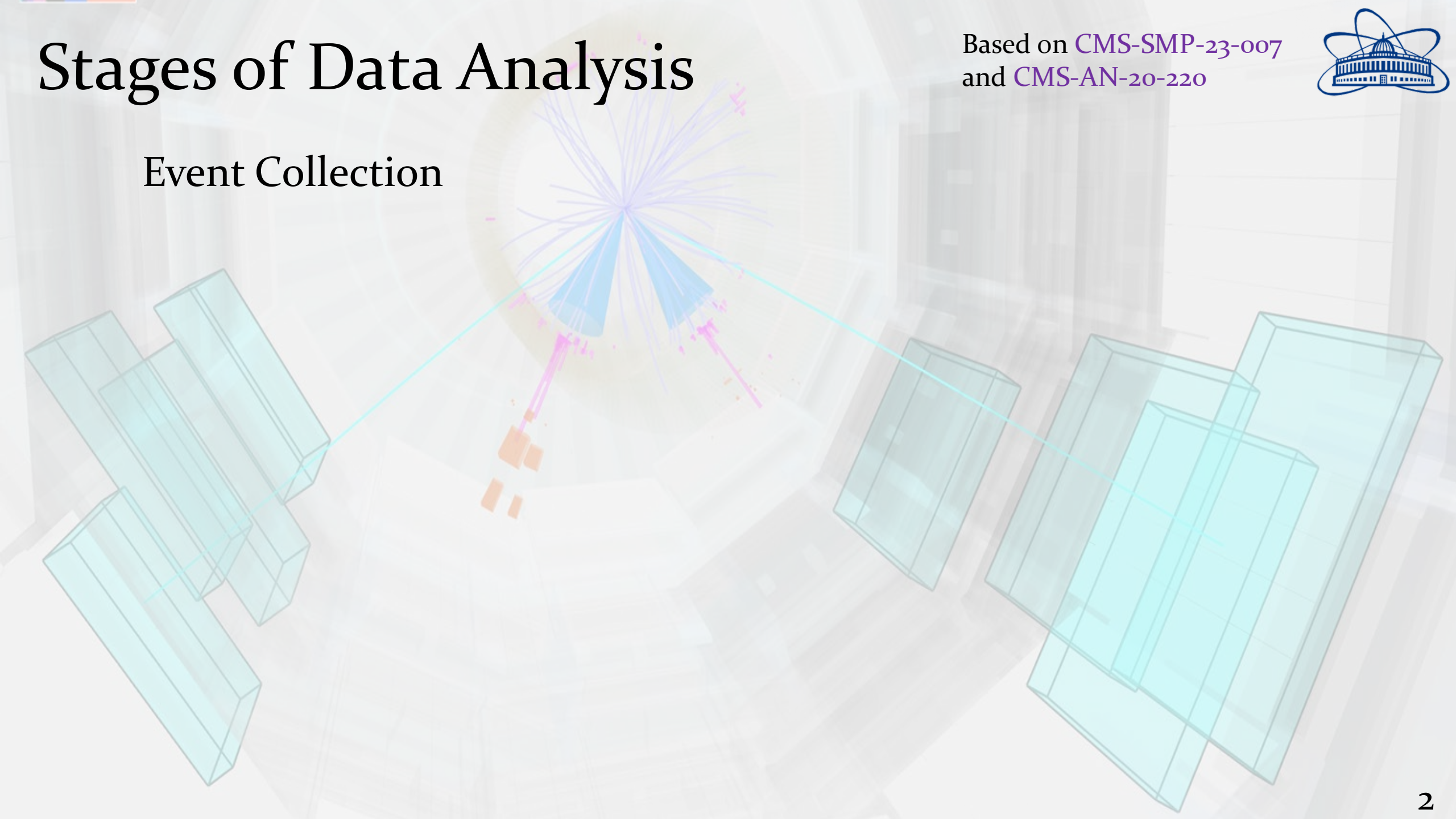
Dubna, 1 November 2024

28th International Scientific Conference of Young Scientists and Specialists (AYSS-2024)

Stages of Data Analysis

Based on CMS-SMP-23-007
and CMS-AN-20-220

Event Collection



Stages of Data Analysis

Based on CMS-SMP-23-007
and CMS-AN-20-220

Event Collection

⬇

Data Certification

Simulation

Stages of Data Analysis

Based on CMS-SMP-23-007
and CMS-AN-20-220

Event Collection

⬇

Data Certification

Simulation

⬇

Event selection

Stages of Data Analysis

Based on CMS-SMP-23-007
and CMS-AN-20-220

Event Collection

⬇

Data Certification

Simulation

⬇

Event selection

⬇

Data and modeling corrections

Stages of Data Analysis

Based on CMS-SMP-23-007 and CMS-AN-20-220

Event Collection

⬇

Data Certification

Simulation

⬇

Event selection

⬇

Data and modeling corrections

⬇

Measurement

Stages of Data Analysis

Based on CMS-SMP-23-007 and CMS-AN-20-220

Event Collection

⬇

Data Certification Simulation

⬇

Event selection

⬇

Data and modeling corrections

⬇

Measurement

⬇

Uncertainty estimation

Stages of Data Analysis

Based on CMS-SMP-23-007 and CMS-AN-20-220

Event Collection

⬇

Data Certification

Simulation

⬇

Event selection

⬇

Data and modeling corrections

⬇

Measurement

⬇

Uncertainty estimation

Stages of Data Analysis

Based on CMS-SMP-23-007 and CMS-AN-20-220

Event Collection

⬇

Data Certification

Simulation

⬇

Event selection

⬇

Data and modeling corrections

⬇

Measurement

⬇

Uncertainty estimation

- Pileup
- Misalignment
- Efficiency

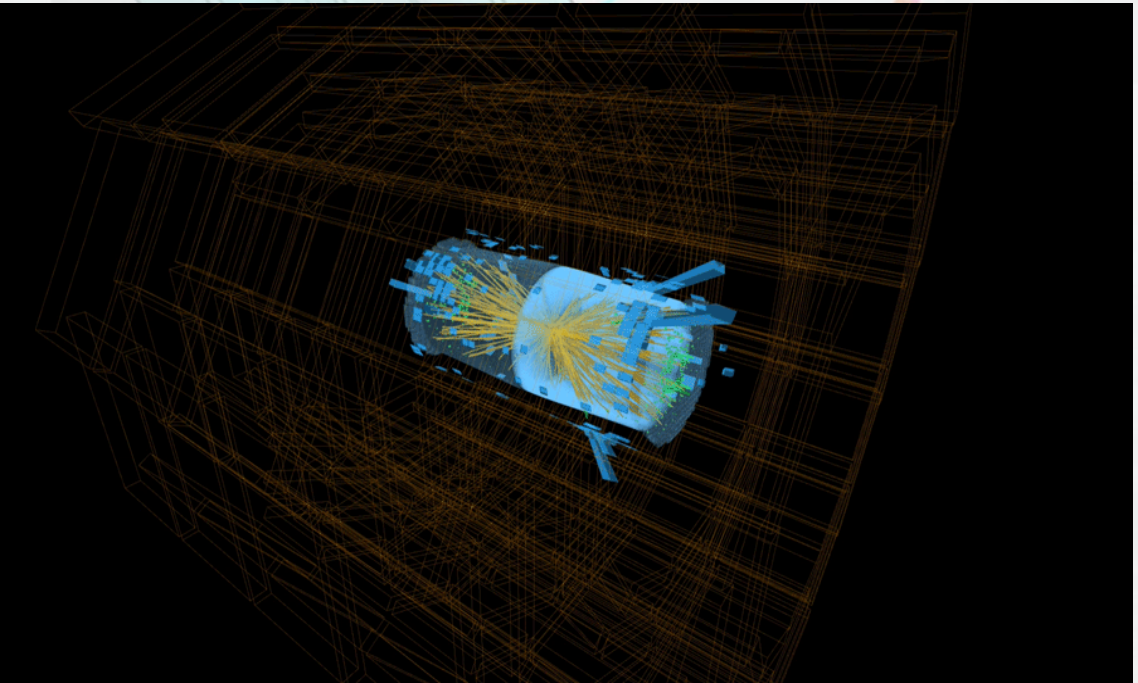
- Prefiring
- Integral Luminosity
- PDF and α_S
- Simulation Cross section
- QCD Scale Factors

Pileup

Origin: $L_{inst} \sim 10^{34}\ cm^{-2}s^{-1}$

- Bunch crossing (BX) every **25 ns**
- ~**60-70** pp-collisions per BX
- ~ **60** charged particles per pp-collision

~**1500-2000** charged particles per BX or $\mathbf{10^{11}}$ per second

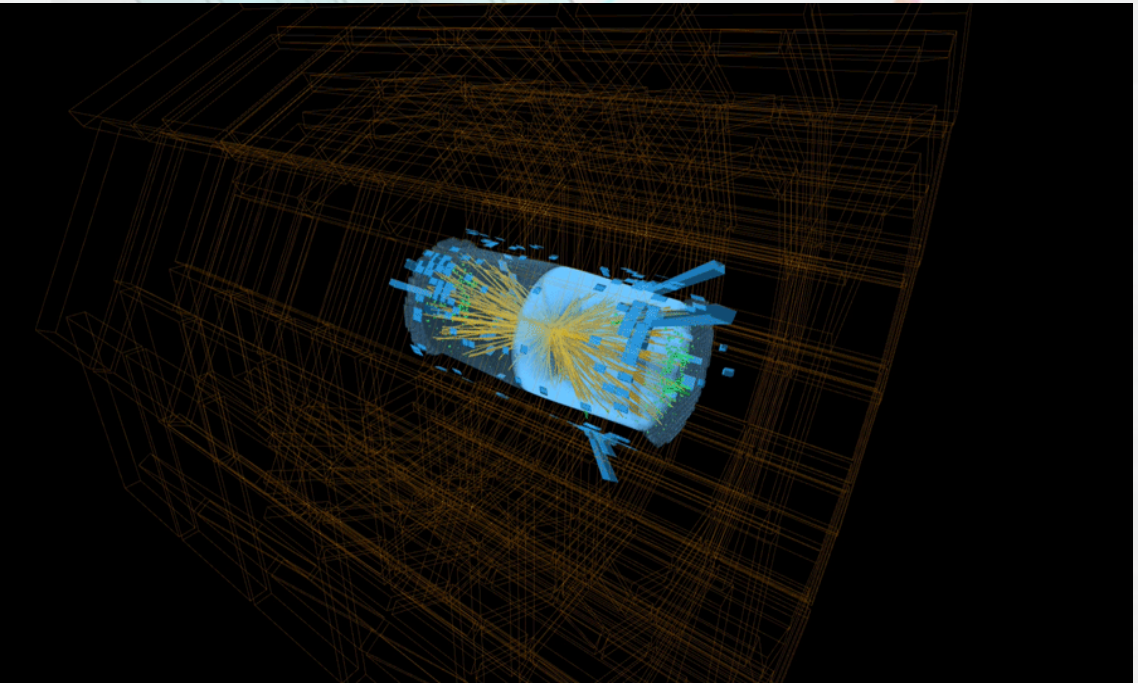


Pileup

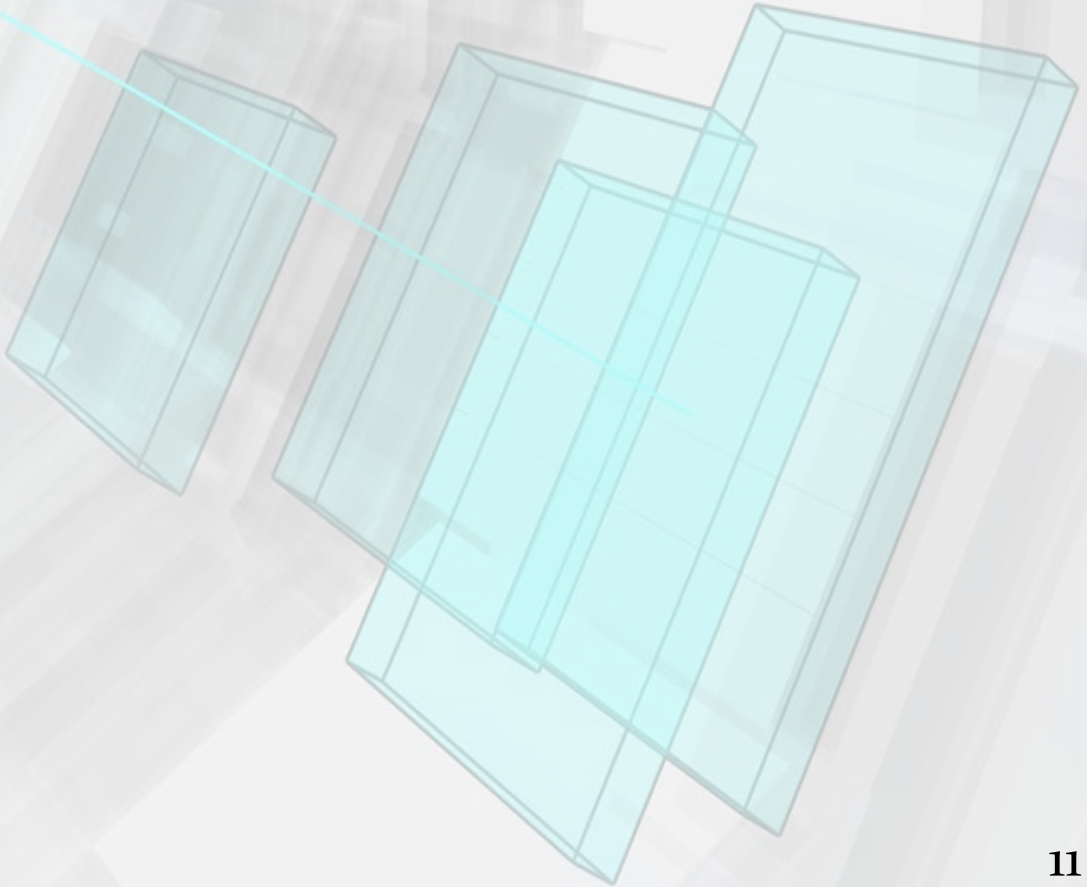
Origin: $L_{inst} \sim 10^{34}\ cm^{-2}s^{-1}$

- Bunch crossing (BX) every **25 ns**
- ~**60-70** pp-collisions per BX
- ~ **60** charged particles per pp-collision

~**1500-2000** charged particles per BX or $\mathbf{10^{11}}$ per second



Correction: Average pileup is reweighted in MC in according with Data

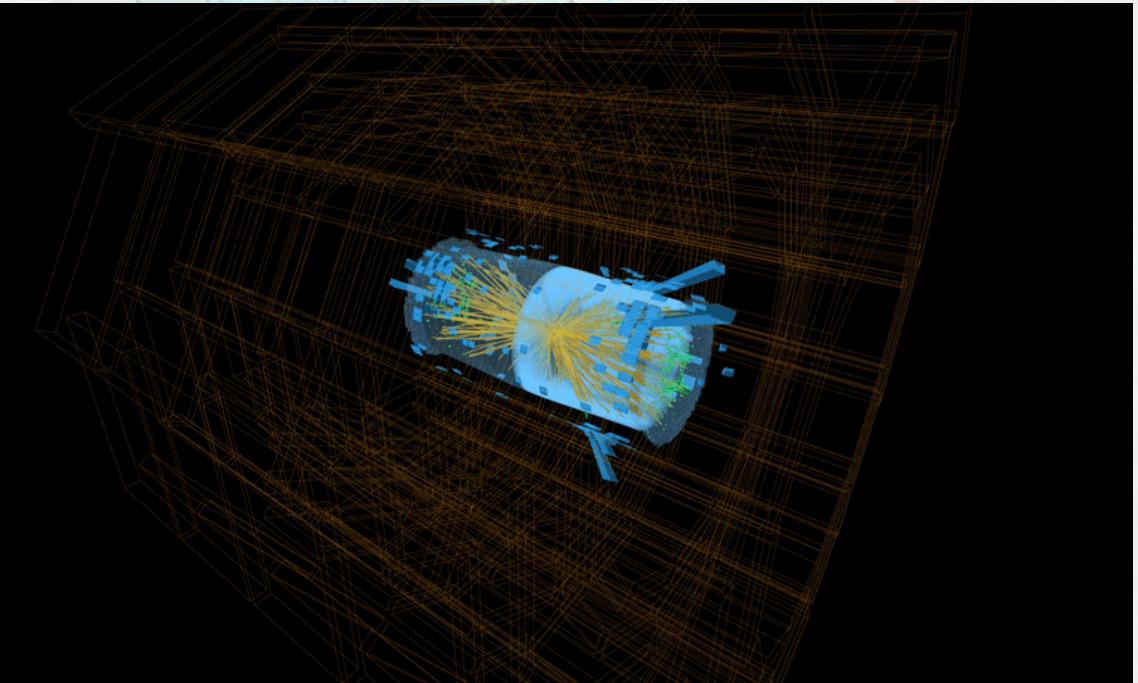


Pileup

Origin: $L_{inst} \sim 10^{34}\ cm^{-2}s^{-1}$

- Bunch crossing (BX) every **25 ns**
- ~**60-70** pp-collisions per BX
- ~ **60** charged particles per pp-collision

~**1500-2000** charged particles per BX or $\mathbf{10^{11}}$ per second



Correction: Average pileup is reweighted in MC in according with Data

Uncertainty: Variation of Minimum biased events XS

Pileup

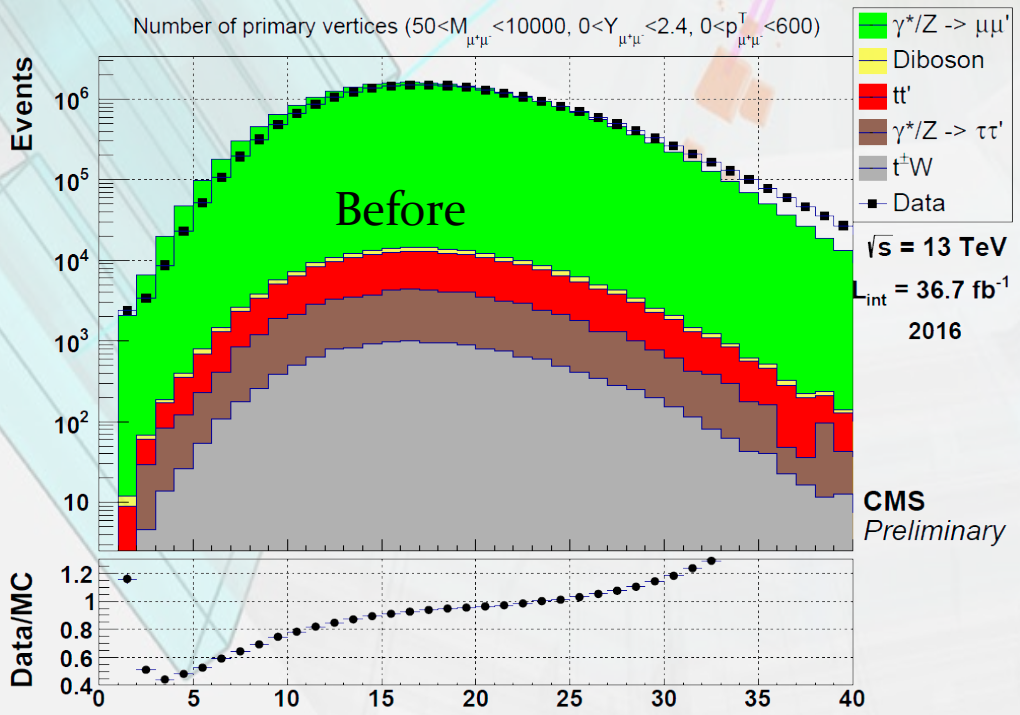
Origin: $L_{inst} \sim 10^{34}\ cm^{-2}s^{-1}$

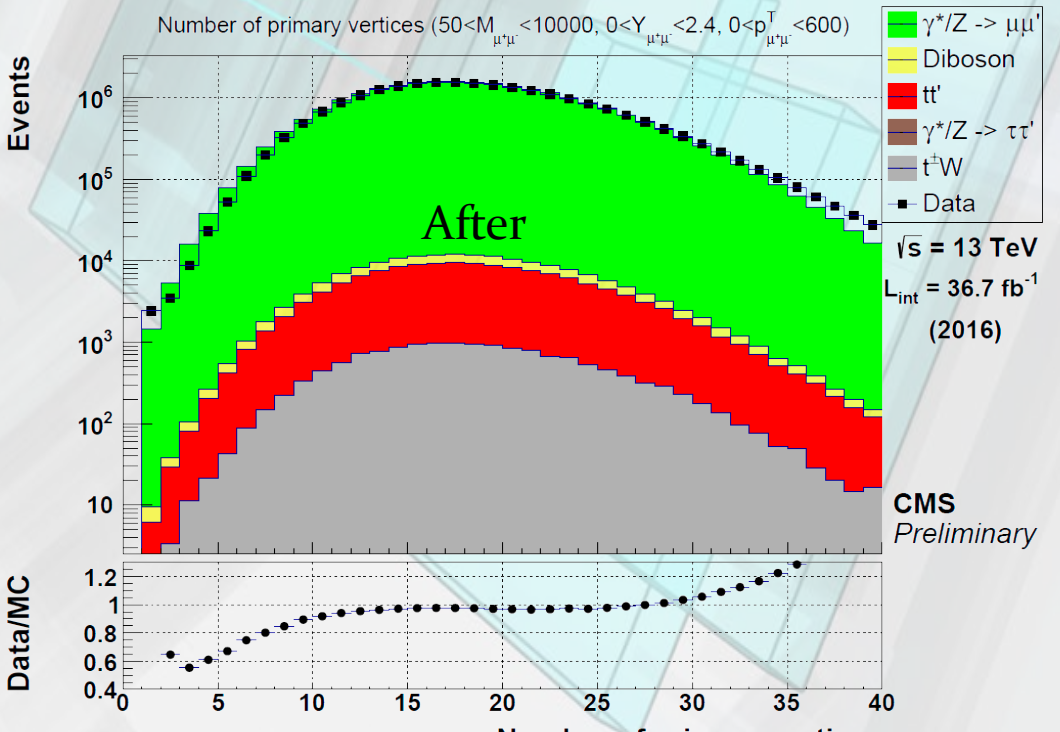
- Bunch crossing (BX) every **25 ns**
- ~**60-70** pp-collisions per BX
- ~ **60** charged particles per pp-collision

~**1500-2000** charged particles per BX or $\mathbf{10^{11}}$ per second

Correction: Average pileup is reweighted in MC in according with Data

Uncertainty: Variation of Minimum biased events XS



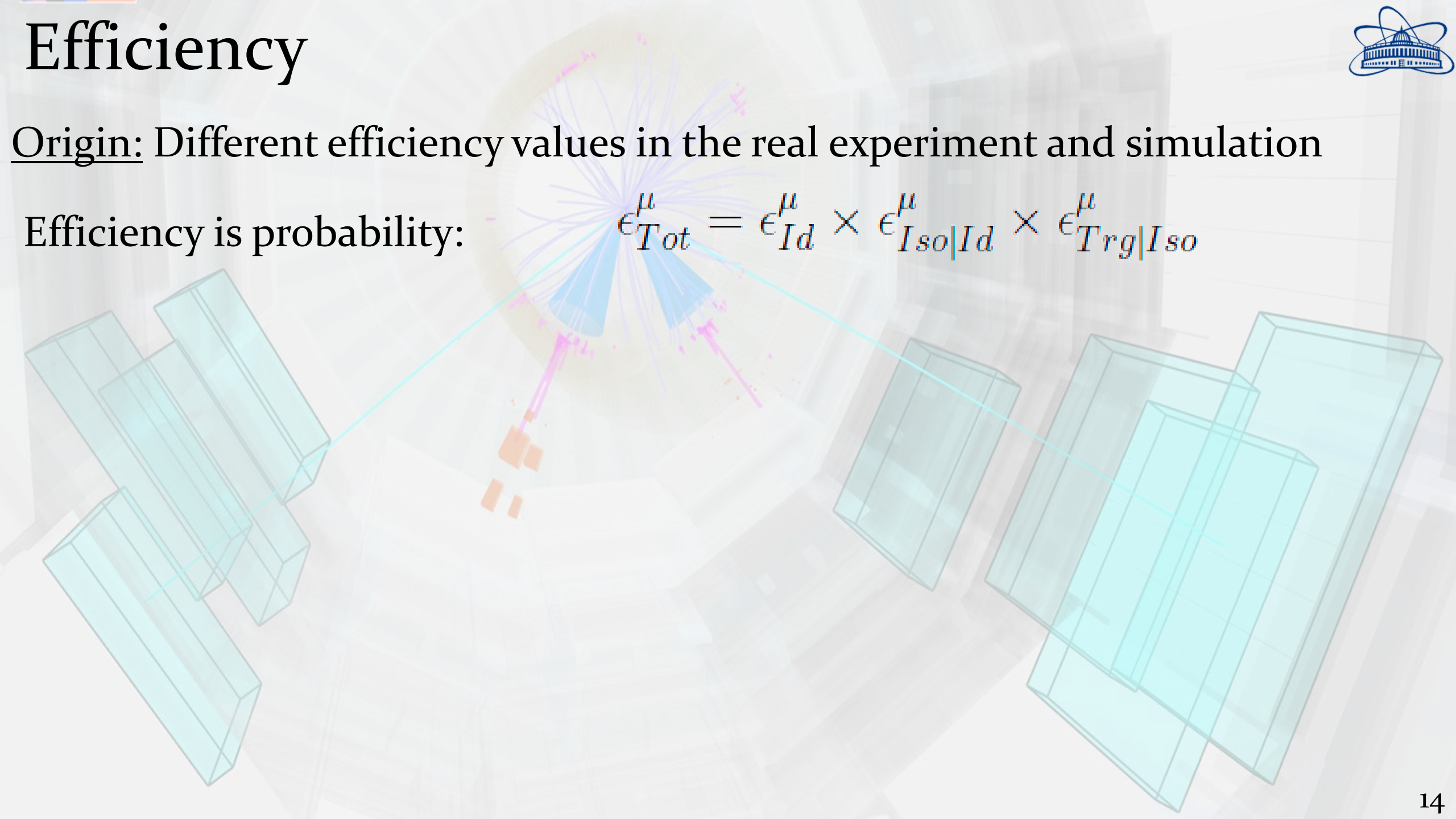


Efficiency

Origin: Different efficiency values in the real experiment and simulation

Efficiency is probability:

$$\epsilon^{\mu}_{Tot} = \epsilon^{\mu}_{Id} \times \epsilon^{\mu}_{Iso|Id} \times \epsilon^{\mu}_{Trg|Iso}$$



Efficiency

Origin: Different efficiency values in the real experiment and simulation

Efficiency is probability: $\epsilon^{\mu}_{Tot} = \epsilon^{\mu}_{Id} \times \epsilon^{\mu}_{Iso|Id} \times \epsilon^{\mu}_{Trg|Iso}$

Correction: Tag & Probe , event reweighting $w^{\mu^+\mu^-}_{Id,Iso,Trg} = \frac{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(Data)}{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(MC)}$

Efficiency

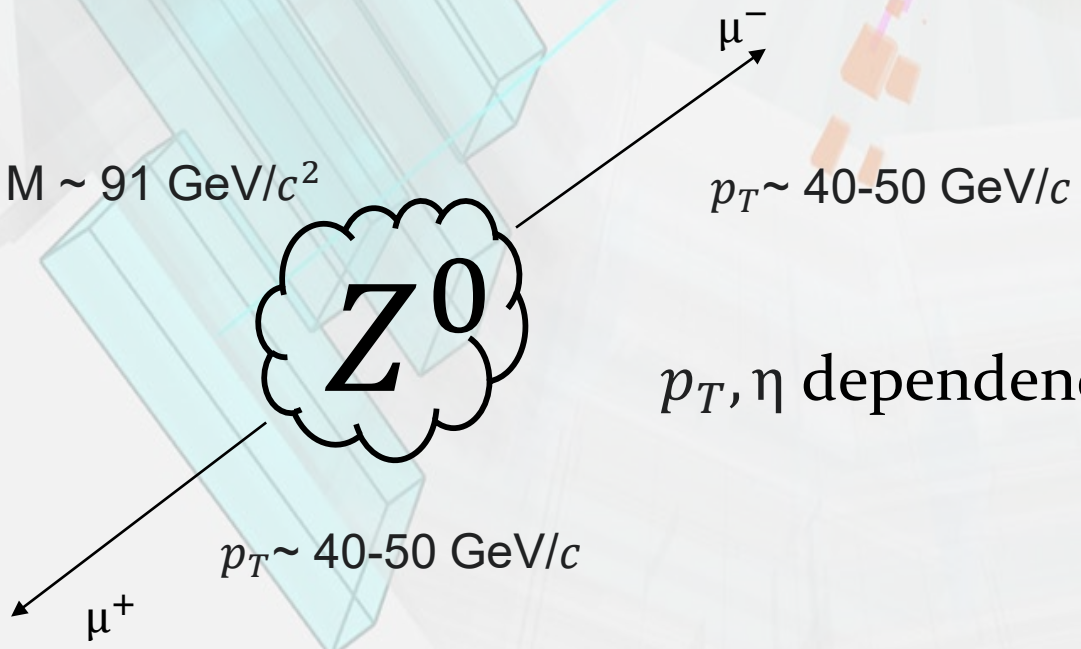
Origin: Different efficiency values in the real experiment and simulation

Efficiency is probability:

$$\epsilon^{\mu}_{Tot} = \epsilon^{\mu}_{Id} \times \epsilon^{\mu}_{Iso|Id} \times \epsilon^{\mu}_{Trg|Iso}$$

Correction: Tag & Probe , event reweighting

$$w^{\mu^+\mu^-}_{Id,Iso,Trg} = \frac{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(Data)}{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(MC)}$$



p_T, η dependence!

Efficiency

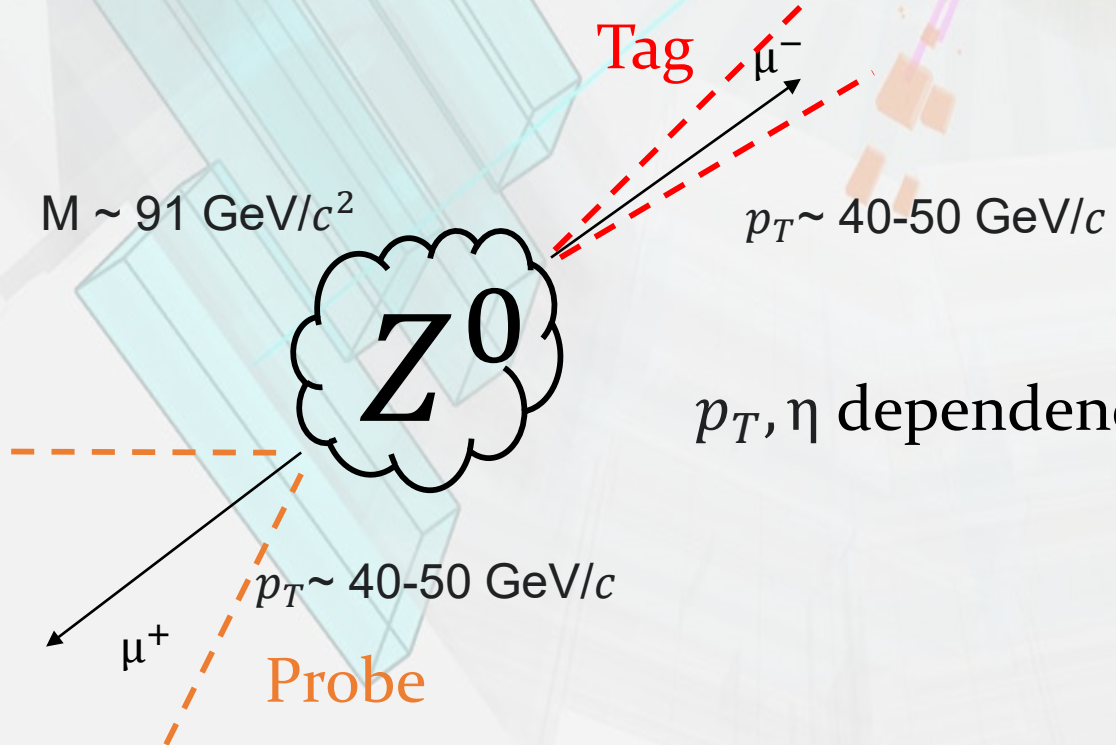
Origin: Different efficiency values in the real experiment and simulation

Efficiency is probability:

$$\epsilon^{\mu}_{Tot} = \epsilon^{\mu}_{Id} \times \epsilon^{\mu}_{Iso|Id} \times \epsilon^{\mu}_{Trg|Iso}$$

Correction: Tag & Probe , event reweighting

$$w^{\mu^+\mu^-}_{Id,Iso,Trg} = \frac{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(Data)}{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(MC)}$$



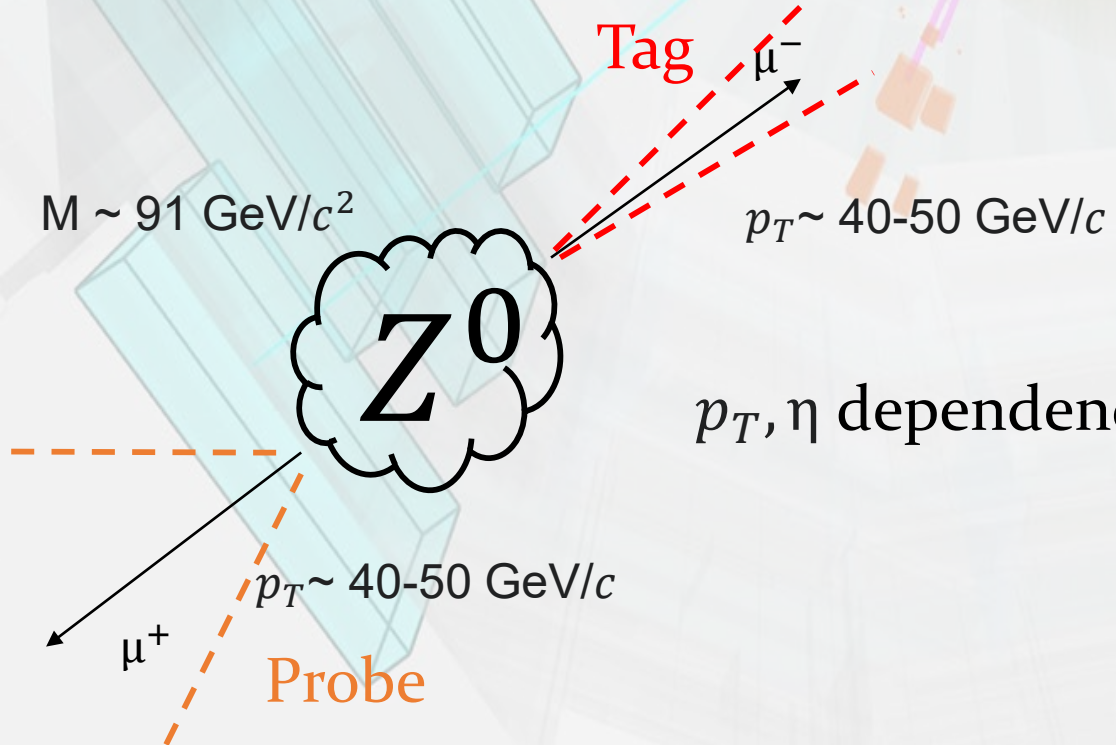
p_T, η dependence!

Efficiency

<u>Origin:</u> Different efficiency values in the real experiment and simulation

Efficiency is probability: $\epsilon^{\mu}_{Tot} = \epsilon^{\mu}_{Id} \times \epsilon^{\mu}_{Iso|Id} \times \epsilon^{\mu}_{Trg|Iso}$

<u>Correction:</u> Tag & Probe , event reweighting $w^{\mu^+\mu^-}_{Id,Iso,Trg} = \frac{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(Data)}{\epsilon^{\mu^+\mu^-}_{Id,Iso,Trg}(MC)}$



p_T, η dependence!

<u>Uncertainty:</u> Variation each of weights. Sum in quadrature

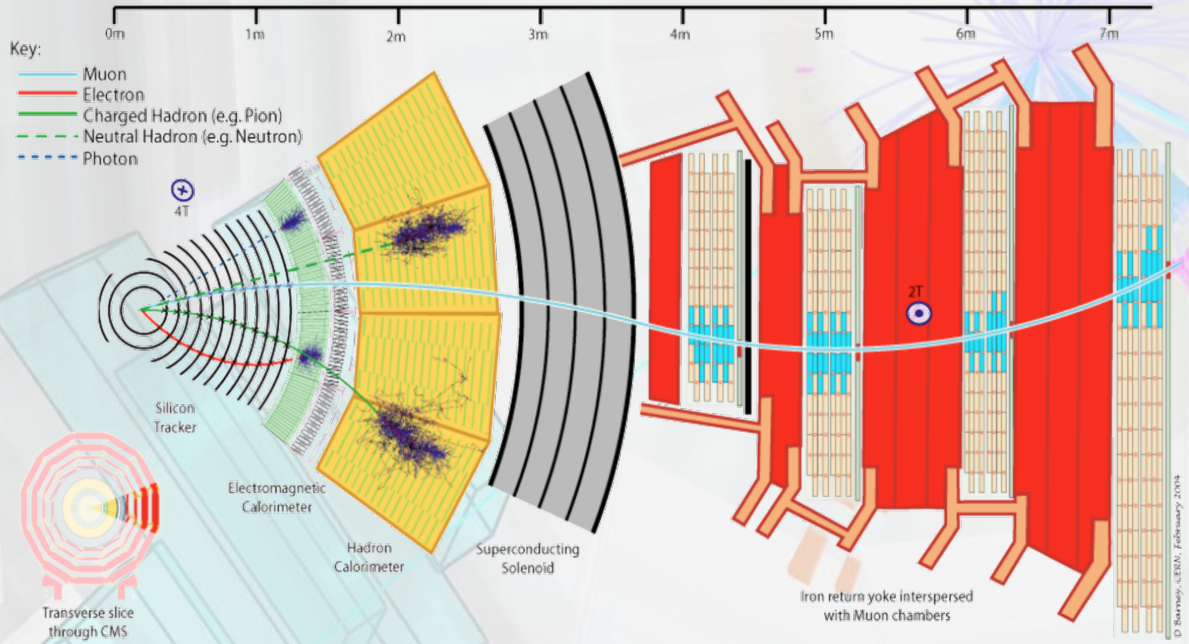
$\Delta = \sqrt{\sum_k^N \delta_k^2}$

- A_i^k - Variated value of A_i coefficients results from variation of uncertainty source
- A_i^0 - central value of A_i coefficient

Most significant type of uncertainty!

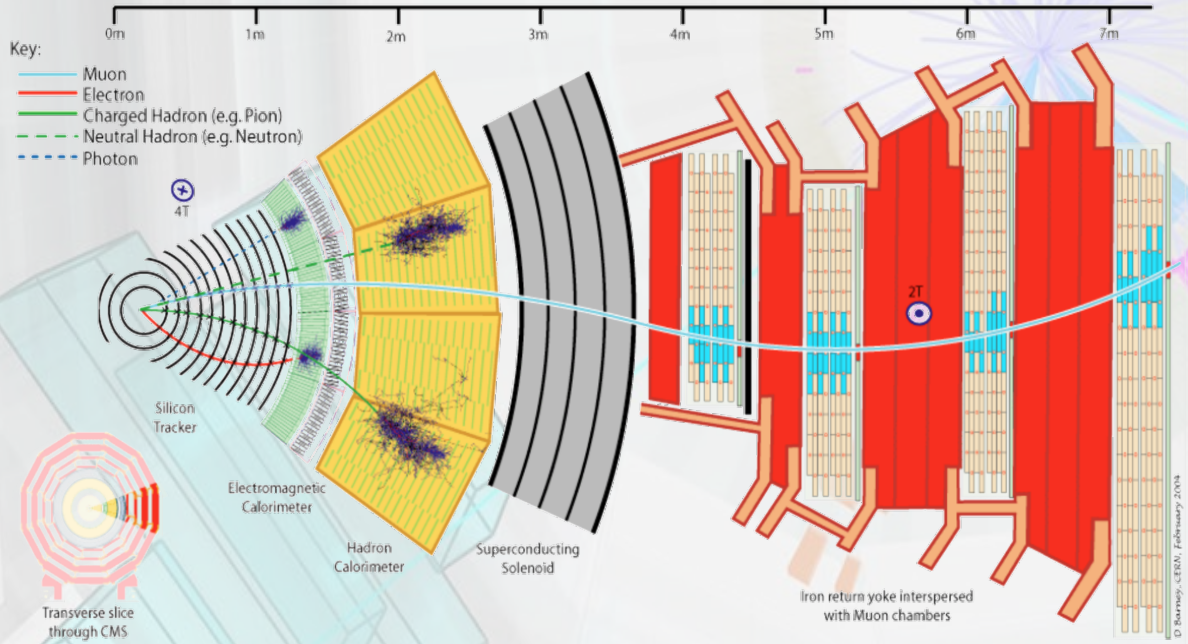
Misalignment

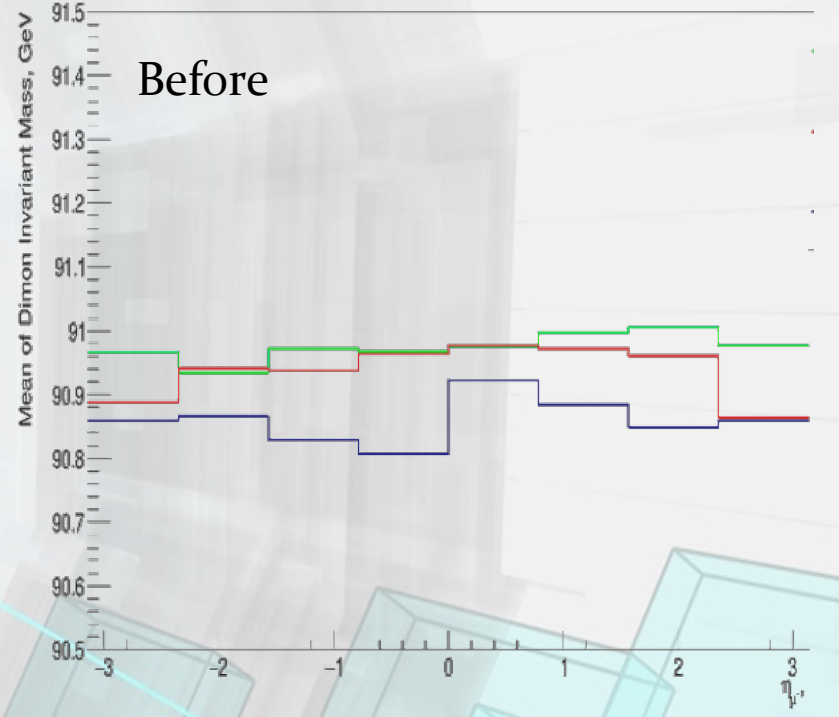
Origin: Non-accuracy of detector model

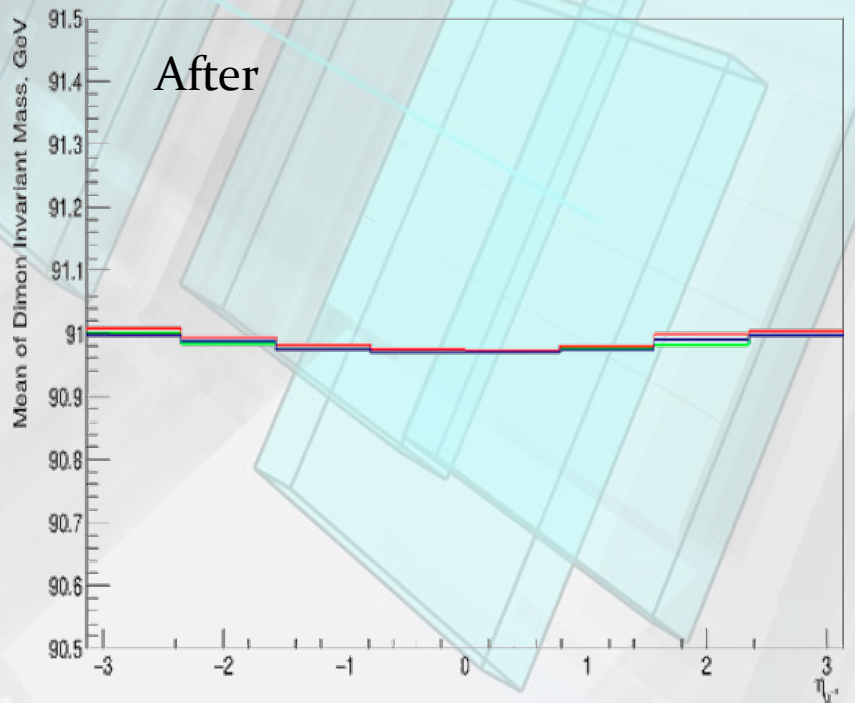


Misalignment

Origin: Non-accuracy of detector model



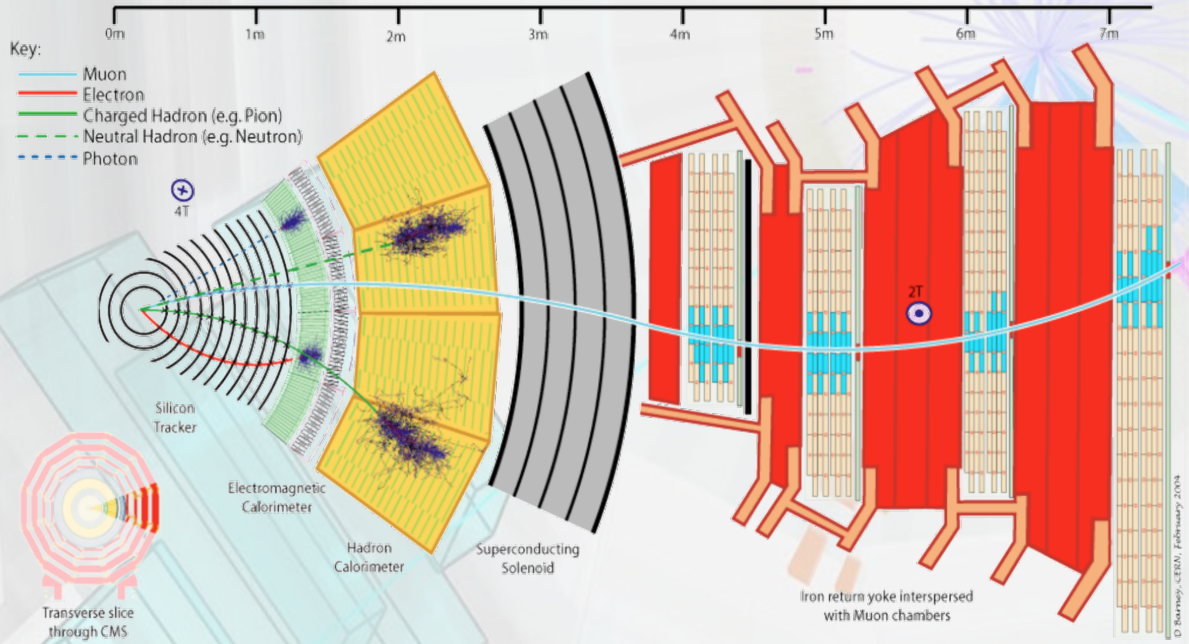




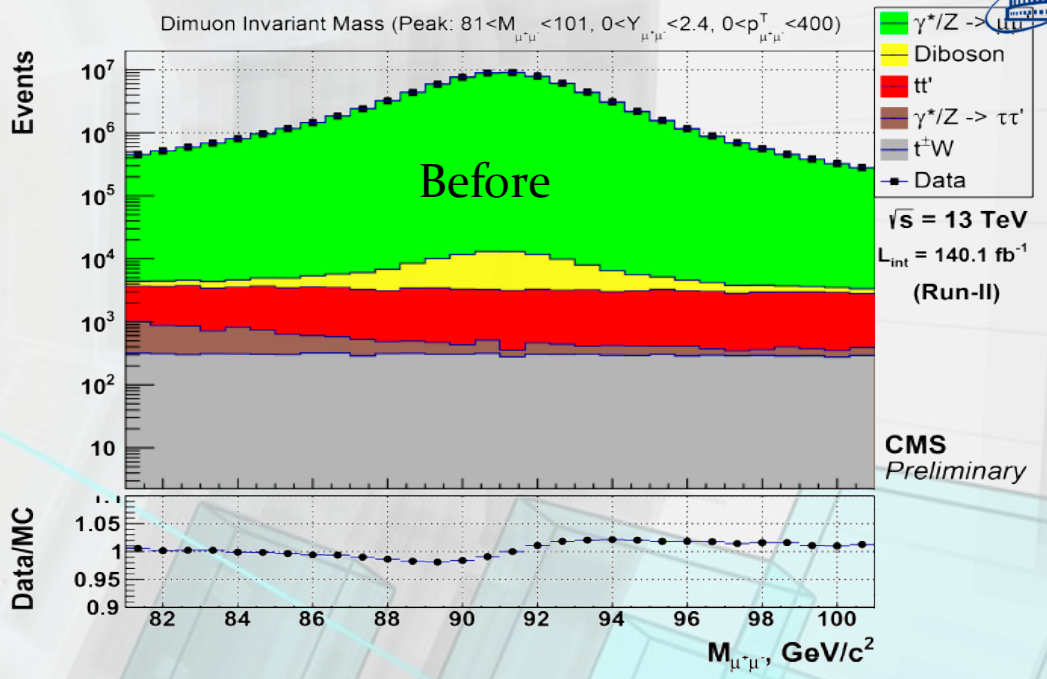
Correction: Tag & Probe with Rochester correction (Eur. Phys. J. C 2012. V. 72, P.)

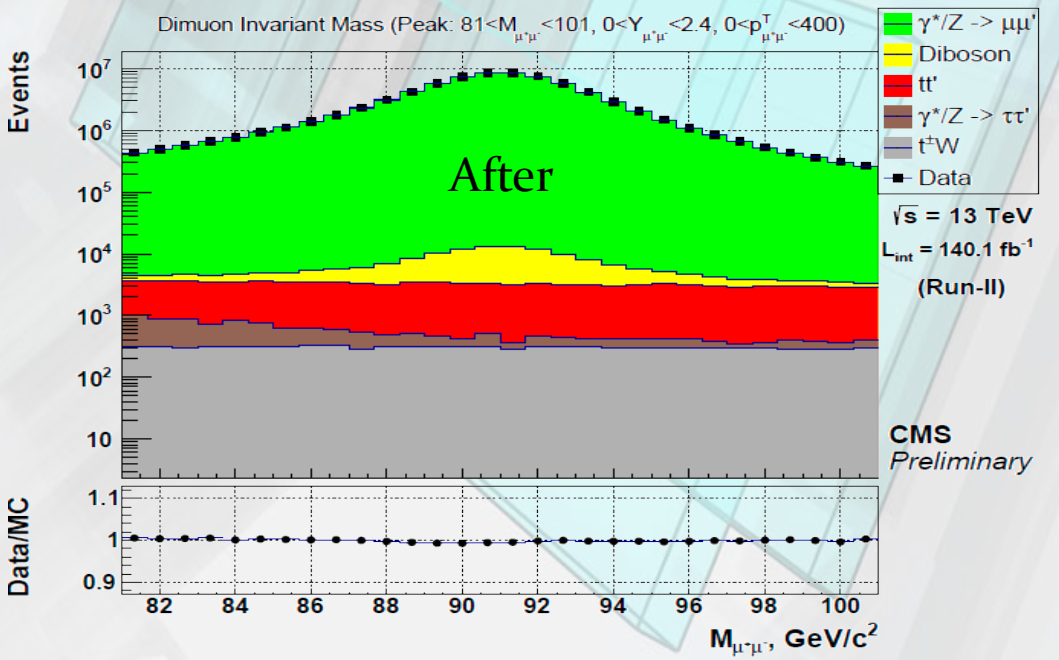
Misalignment

Origin: Non-accuracy of detector model



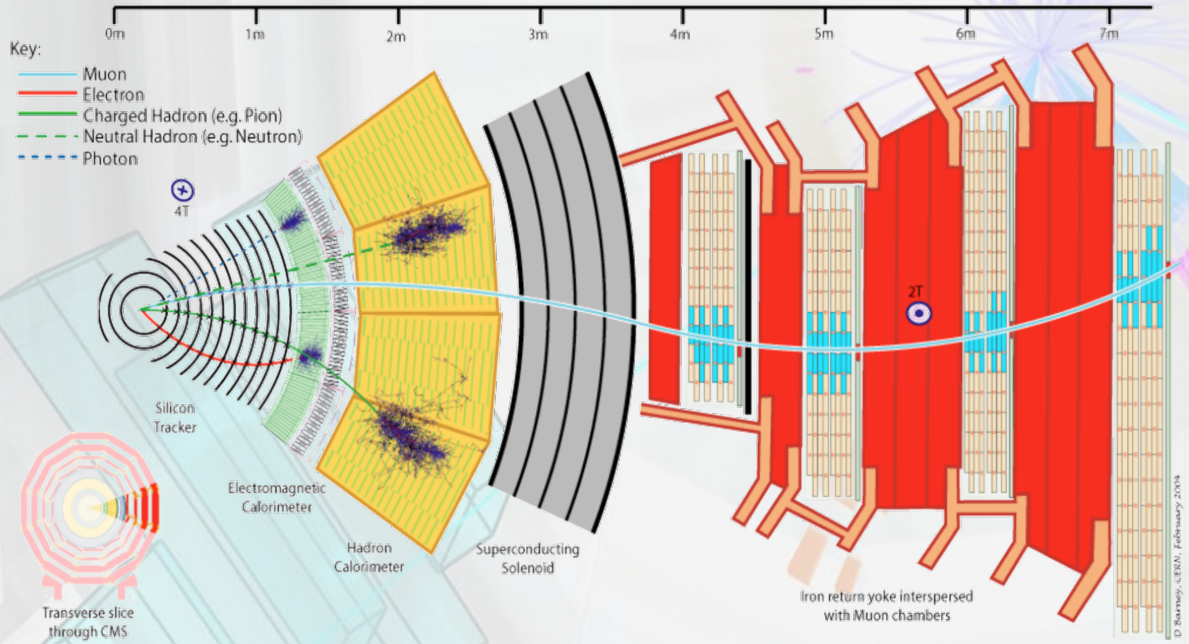
Correction: Tag & Probe with Rochester correction (Eur. Phys. J. C 2012. V. 72, P.)

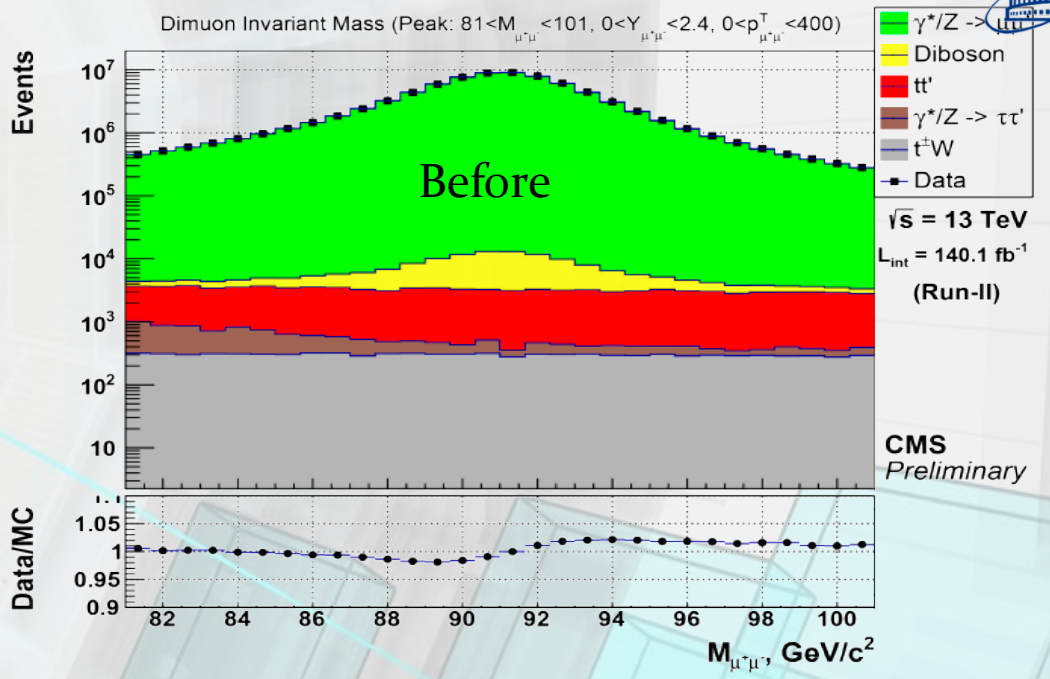


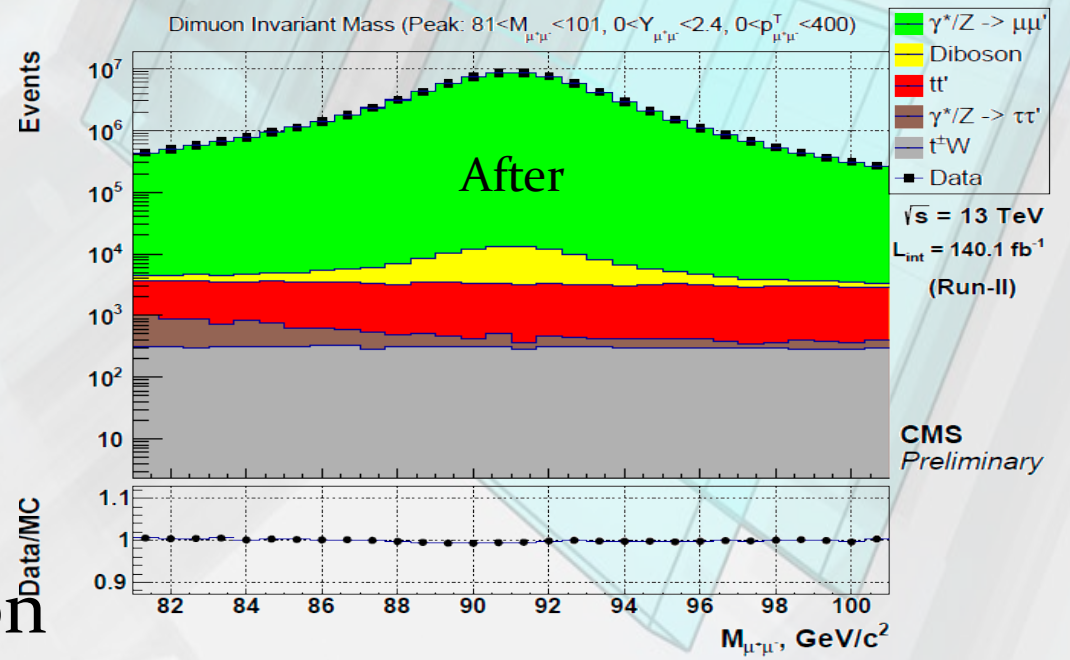


Misalignment

Origin: Non-accuracy of detector model







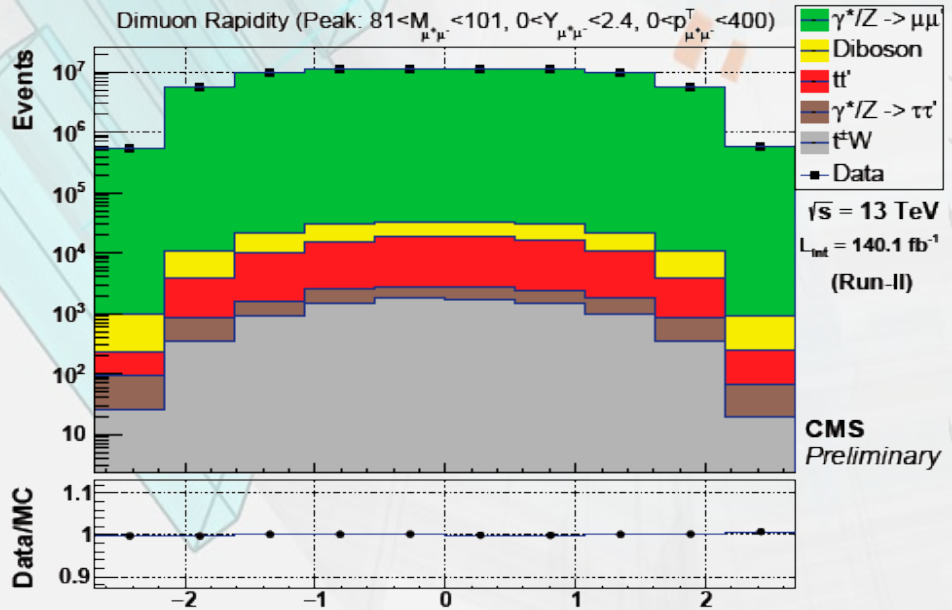
Correction: Tag & Probe with Rochester correction (Eur. Phys. J. C 2012. V. 72, P.)

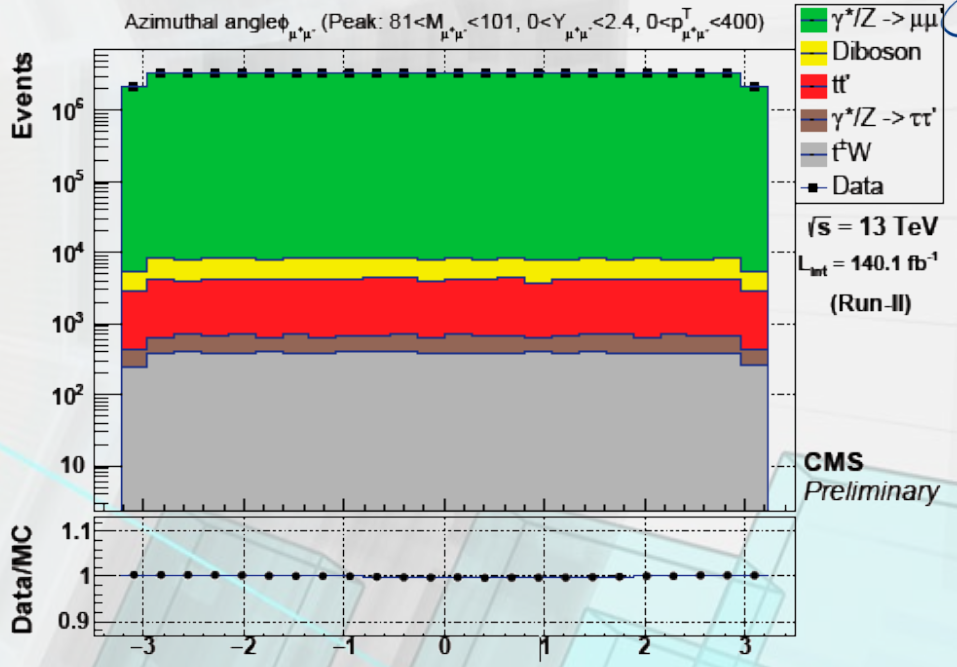
Uncertainty: 100 “toys” and RMS calculation

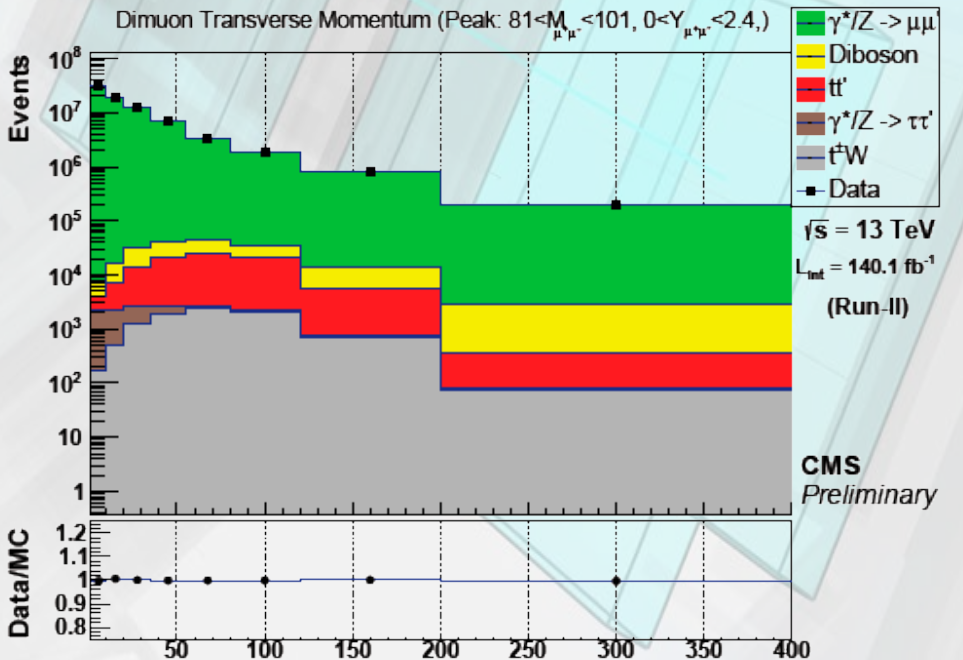
Conclusions

- Effects of difference between data and simulation are constantly being studied and could be successfully corrected.

- In analysis SMP-23-007 this steps are successfully done. Good agreement between data and MC is observed



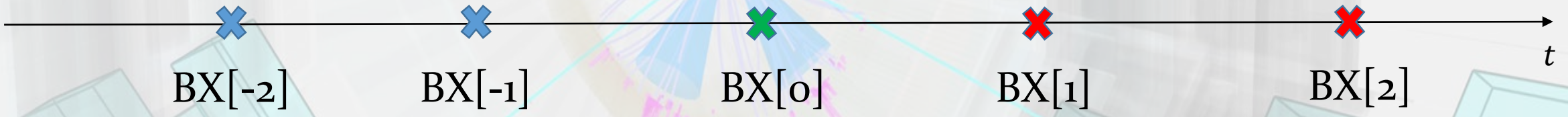




Prefiring

Origin: L_{inst} ~10^{34} $cm^{-2}s^{-1}$ and L1 Trigger time desynchronization

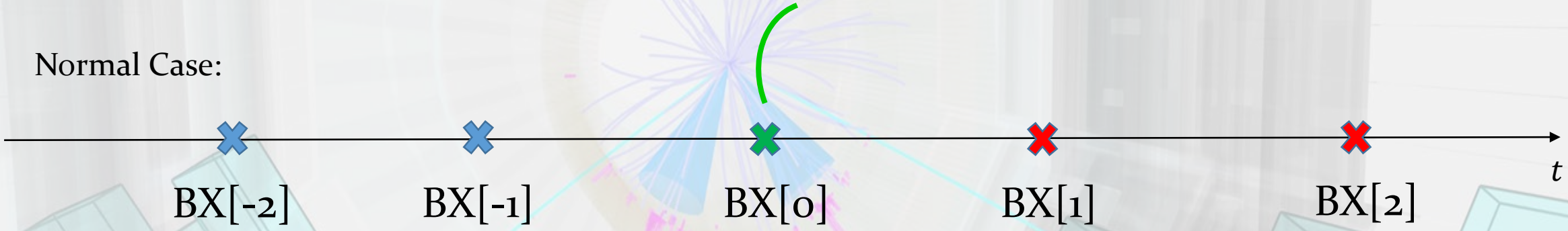
Normal Case:



Prefiring

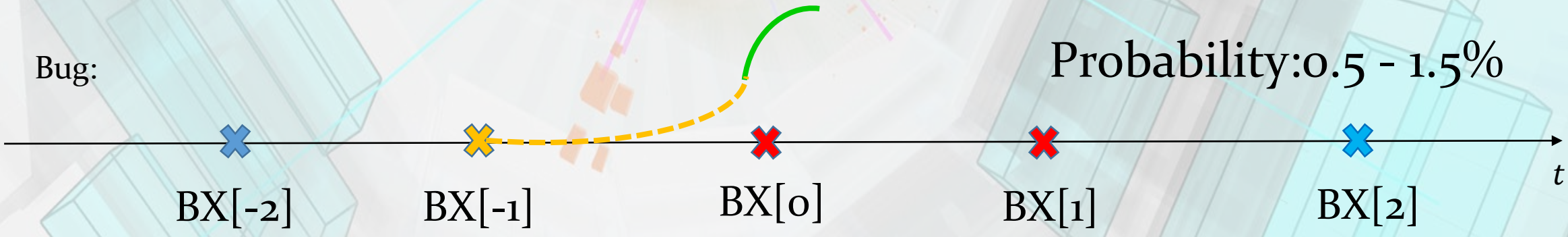
Origin: L_{inst} ~10^{34} $cm^{-2}s^{-1}$ and L1 Trigger time desynchronization

Normal Case:



Bug:

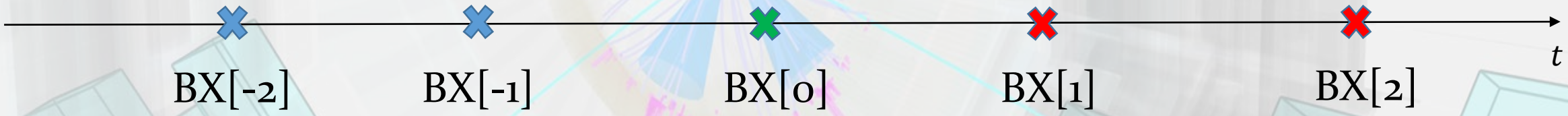
Probability:0.5 - 1.5%



Prefiring

Origin: L_{inst} ~$10^{34}\ cm^{-2}s^{-1}$ and L1 Trigger time desynchronization

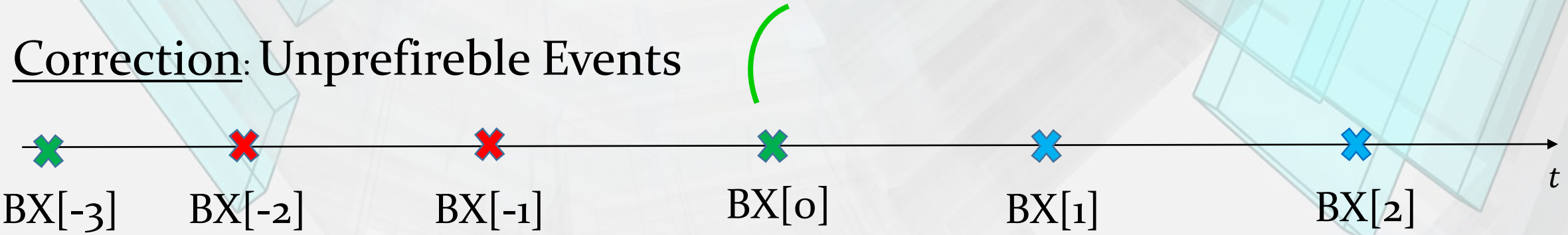
Normal Case:



Bug:

Probability:0.5 - 1.5%

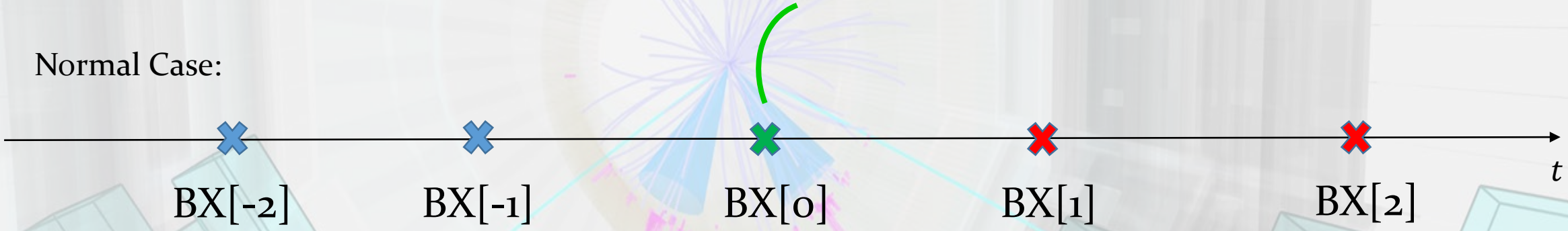
Correction: Unprefireble Events



Prefiring

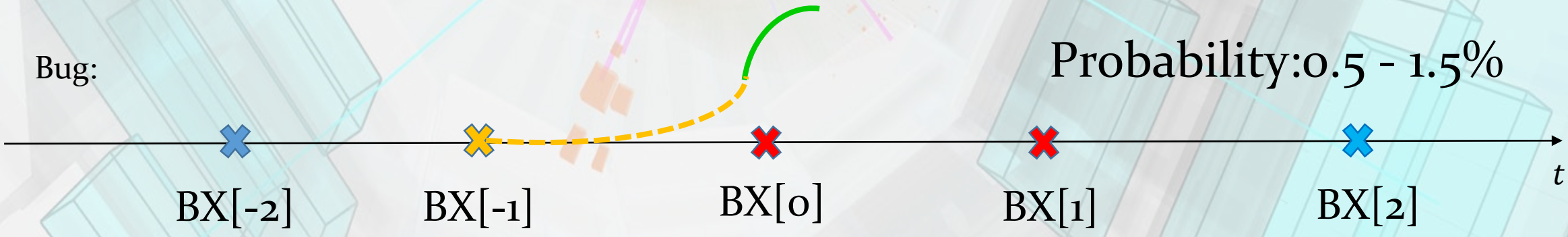
Origin: L_{inst} ~10^{34} $cm^{-2}s^{-1}$ and L1 Trigger time desynchronization

Normal Case:



Bug:

Probability:0.5 - 1.5%



Correction: Unprefireble Events

Uncertainty: variation within unc.