מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

# A data-directed search for $e \leftrightarrow \mu$ asymmetry in events containing $\tau$'s

Shikma Bressler

The 16th International Workshop on Tau Lepton Physics (TAU2021) (Virtual Edition)

# Frame setting


- **Flavor physics** - excellent arena for discovery of NP
  Relates to some of the most fundamental mysteries of the universe

- **Lepton number and lepton flavo**r - accidental symmetries of the SM
  $U(1)_\tau \times U(1)_\mu \times U(1)_e$ → Easily violated by NP

- **Neutrino oscillations** → lepton flavor is not a symmetry of nature
  Can be accounted for with new physics at the seesaw scale ($\Lambda = \mathcal{O}(10^{15}\ GeV)$)
  → no visible imprint in colliders

- **Hints for Lepton non-universality in b-hadron decays** $R_{D^{(*)}}$ and $R_{K^{(*)}}$
  → confirmed or rejected in the coming years
  If real → upper bounds on the scale of NP at $\Lambda = \mathcal{O}(5\ TeV)$ and $\Lambda = \mathcal{O}(35\ TeV)$
  → directly accessible at the (HL)-LHC

# $e \leftrightarrow \mu \leftrightarrow \tau$ asymmetry – a tool for NP discovery

- **Hints for Lepton non-universality in b-hadron decays** $R_{D^{(*)}}$ and $R_{K^{(*)}}$
  → Demonstrate the role that $e \leftrightarrow \mu \leftrightarrow \tau$ asymmetry could play in the discovery of NP

$$R(D^{(*)}) \equiv \frac{\Gamma(B \to D^{(*)}\tau\bar{\nu})}{\Gamma(B \to D^{(*)}\ell\bar{\nu})}, \quad (\ell = e \text{ or } \mu) \qquad R_K = \frac{\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)} \Bigg/ \frac{\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathcal{B}(B^+ \to J/\psi(\to e^+e^-)K^+)}$$

# $e \leftrightarrow \mu \leftrightarrow \tau$ asymmetry – a tool for NP discovery

- **Hints for Lepton non-universality in b-hadron decays** $R_{D^{(*)}}$ and $R_{K^{(*)}}$
  → Demonstrate the role that $e \leftrightarrow \mu \leftrightarrow \tau$ asymmetry cou [illegible] very of NP

$$R(D^{(*)}) \equiv \frac{\Gamma(B \to D^{(*)}\tau\bar{\nu})}{\Gamma(B \to D^{(*)}\ell\bar{\nu})}, \quad (\ell = e \text{ or } \mu)$$

Special feature of the SM

- **Part of a bigger picture**

| interaction | fermions | force carrier | coupling | SM flavor | BSM flavor | Possible models |
|---|---|---|---|---|---|---|
| Electromagnetic | $\ell$ | $A^0$ | $eQ$ | universal | universal | |
| NC weak | all | $Z^0$ | $\frac{e(T_3 - s_W^2 Q)}{s_W c_W}$ | universal | ? | Different $T_3$ i.e. VLL |
| CC weak | $\bar{\ell}\nu$ | $W^{\pm}$ | $g$ | universal | ? | $\Delta m_\nu \neq 0$ |
| Yukawa | $\ell$ | $h$ | $y_\ell$ | diagonal | ? | 2HDM VLL |

→ Any observed $e \leftrightarrow \mu \leftrightarrow \tau$ asymmetry is an indication for NP

→ The search for it is highly motivated *BUT* (too?) many possible searches

# The experimental challenge

- Individual searches might pose specific challenges
  - Solutions, if available, will eventually be found

- Many possible final states
- (Too?) Many potential searches & limited manpower
  → Impossible to search for all of them
  → Impossible to cover all the observable space
- Theoretical predictions could provide some guidance
  - Already (Too?) many predictions
  - Many models haven't been written yet
- NP might be out there but missed detection

- This concern is common to any BSM search → Calls for new approaches

# The Data-Directed Paradigm (DDP)

**arXiv:2107.11573**

- Complementary to the blind-analysis paradigm
  - In which the data is only looked at the final stage of the analysis
  - After most of the time and effort have been invested

- Goal: Identify regions in the data which exhibit significant deviation from a SM prediction
  - Rather than exhaustively study huge number of exclusive selections –
    first find those that matter and invest the effort there

# The Data-Directed Paradigm (DDP)

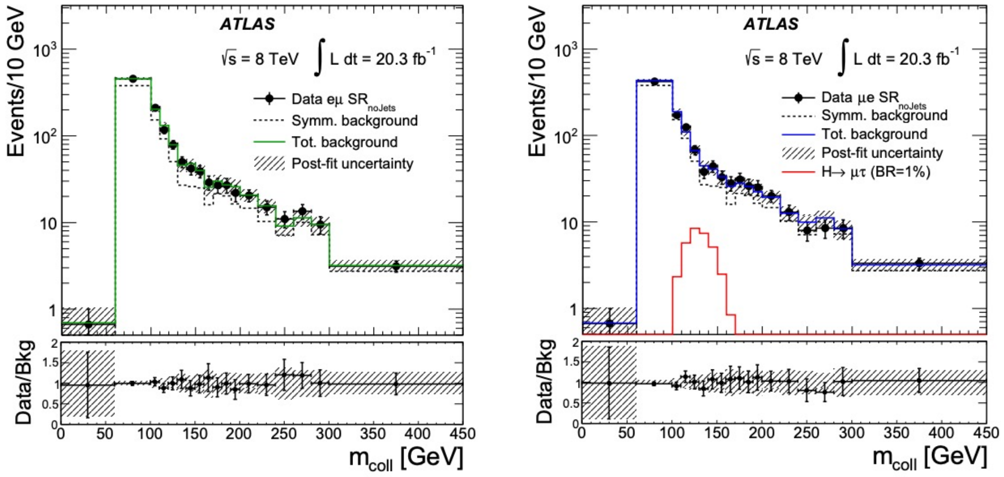
**arXiv:2107.11573**

- Complementary to the blind-analysis paradigm
  - In which the data is only looked at the final stage of the analysis
  - After most of the time and effort have been invested

- Goal: Identify regions in the data which exhibit significant deviation from a SM prediction
  - Rather than exhaustively study huge number of exclusive selections – first find those that matter and invest the effort there
- Two key ingredients
  - A property of the SM based on which deviations can be looked for
    Here: assume flavor symmetry and search for $e \leftrightarrow \mu \leftrightarrow \tau$ asymmetry
  - An efficient tool to search for a deviation

- In arXiv:2107.11573 the DDP is demonstrated based on the bump hunting concept
  - Exploiting the fact that within the SM, in absence of resonances, almost any invariant mass distribution at the LHC data is smoothly falling

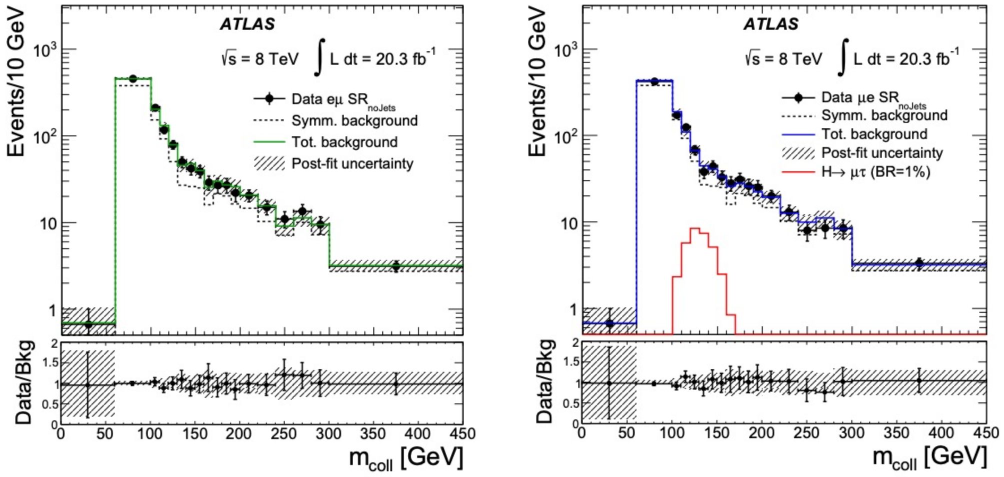
# A DDP search for $e \leftrightarrow \mu_{(\leftrightarrow \tau)}$ asymmetry




Ingredient I: a property of the SM:

- Symmetry to the replacement of prompt $e \leftrightarrow \mu$
  arXiv:1405.4545
  - Theory wise
    - Gauge coupling are universal
    - Yukawa and phase space effects are negligible
      And can be accounted for if needed
  - Detector effects invalidating the symmetry can be accounted for
    - Different trigger and offline efficiencies
    - Different fake and non-prompt rates
    - Different momentum resolution and rate and spectrum of Bremsstrahlung radiation (relevant in part of the observable space)
  - Proved for final state containing exactly one e and od one $\mu$ in the final state
    arXiv:1604.07730 – ATLAS search for Higgs LFV decays

# A DDP search for $e \leftrightarrow \mu_{(\leftrightarrow \tau)}$ asymmetry




Ingredient I: a property of the SM:

- Symmetry to the replacement of prompt $e \leftrightarrow \mu$
  arXiv:1405.4545
  - Theory wise
    - Gauge coupling are universal
    - Yukawa and phase space effects are negligible
      And can be accounted for if needed
  - Detector effects invalidating the symmetry can be accounted for
    - Different trigger and offline efficiencies
    - Different fake and non-prompt rates
    - Different momentum resolution and rate and spectrum of Bremsstrahlung radiation (relevant in part of the observable space)
  - Proved for final state containing exactly one e and od one $\mu$ in the final state
    arXiv:1604.07730 – ATLAS search for Higgs LFV decays

How does it work?

# Ingredient I: $e \leftrightarrow \mu$ symmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Divide the data into two mutually exclusive samples
  - *e-based* dataset and the corresponding μ-based dataset
  - Exactly the same selection criteria replacing $e \leftrightarrow \mu$
- The symmetry assumption – the same number of event in the IP: $N_{IP}^{e-based} = N_{IP}^{\mu-based} \equiv N_0$
- Measured number of events affected by efficiencies and fake rates

$$N^{e-based} = N_0 \times \varepsilon^{e-based} + N_{fake}^{e}$$
$$N^{\mu-based} = N_0 \times \varepsilon^{\mu-based} + N_{fake}^{\mu}$$

# Ingredient I: $e \leftrightarrow \mu$ symmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Divide the data into two mutually exclusive samples
  - *e-based* dataset and the corresponding μ-based dataset
  - Exactly the same selection criteria replacing $e \leftrightarrow \mu$
- The symmetry assumption – the same number of events in the IP: $N_{IP}^{e-based} = N_{IP}^{\mu-based} \equiv N_0$
- Measured number of events affected by efficiencies and fake rates

$$N^{e-based} = N_0 \times \varepsilon^{e-based} + N_{fake}^{e}$$
$$N^{\mu-based} = N_0 \times \varepsilon^{\mu-based} + N_{fake}^{\mu}$$

$$N^{e-based} = (N^{\mu-based} - N_{fake}^{\mu}) \times \frac{\varepsilon^{e-based}}{\varepsilon^{\mu-based}} + N_{fake}^{e}$$

- Efficiencies and fake rate are properties of the lepton
  - Up to some corrections and systematic uncertainties
- $\mu$ dataset can be used to estimate the number of events in the e dataset

# Ingredient I: $e \leftrightarrow \mu$ symmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Divide the data into two mutually exclusive samples
  - *e-based* dataset and the corresponding μ-based dataset
  - Exactly the same selection criteria replacing $e \leftrightarrow \mu$
- The symmetry assumption – the same number of events in the IP: $N_{IP}^{e-based} = N_{IP}^{\mu-based} \equiv N_0$
- Measured number of events affected by efficiencies and fake rates

$$N^{e-based} = N_0 \times \varepsilon^{e-based} + N_{fake}^{e}$$
$$N^{\mu-based} = N_0 \times \varepsilon^{\mu-based} + N_{fake}^{\mu}$$

$$\boxed{N^{e-based} = (N^{\mu-based} - N_{fake}^{\mu}) \times \mathfrak{R}^{\varepsilon} + N_{fake}^{e}}$$

- Efficiencies and fake rate are properties of the lepton
  - Up to some corrections and systematic uncertainties
- $\mu$ dataset can be used to estimate the number of events in the e dataset

# Ingredient I: $e \leftrightarrow \mu$ symmetry

Restoring the $e \leftrightarrow \mu$ symmetry

- Divide the data into two mutually exclusive samples
  - *e-based* dataset and the corresponding μ-based dataset
  - Exactly the same selection criteria replacing $e \leftrightarrow \mu$
- The symmetry assumption – the same number of events in the IP: $N_{IP}^{e-based} = N_{IP}^{\mu-based} \equiv N_0$
- Measured number of events affected by efficiencies and fake rates

$$N^{e-based} = N_0 \times \varepsilon^{e-based} + N_{fake}^{e}$$
$$N^{\mu-based} = N_0 \times \varepsilon^{\mu-based} + N_{fake}^{\mu}$$

$$N^{e-based} = (N^{\mu-based} - N_{fake}^{\mu}) \times \Re^{\varepsilon} + N_{fake}^{e}$$

- Efficiencies and fake rate are properties of the lepton
  - Up to some corrections and systematic uncertainties
- $\mu$ dataset can be used to estimate the number of events in the e dataset
- True for any possible distribution

# Ingredient I: $e \leftrightarrow \mu$ symmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Correction can be made on an event-by-event basis
  - Replacing to total number of events by summation on event weights
    Excluding the fake contribution for simplicity

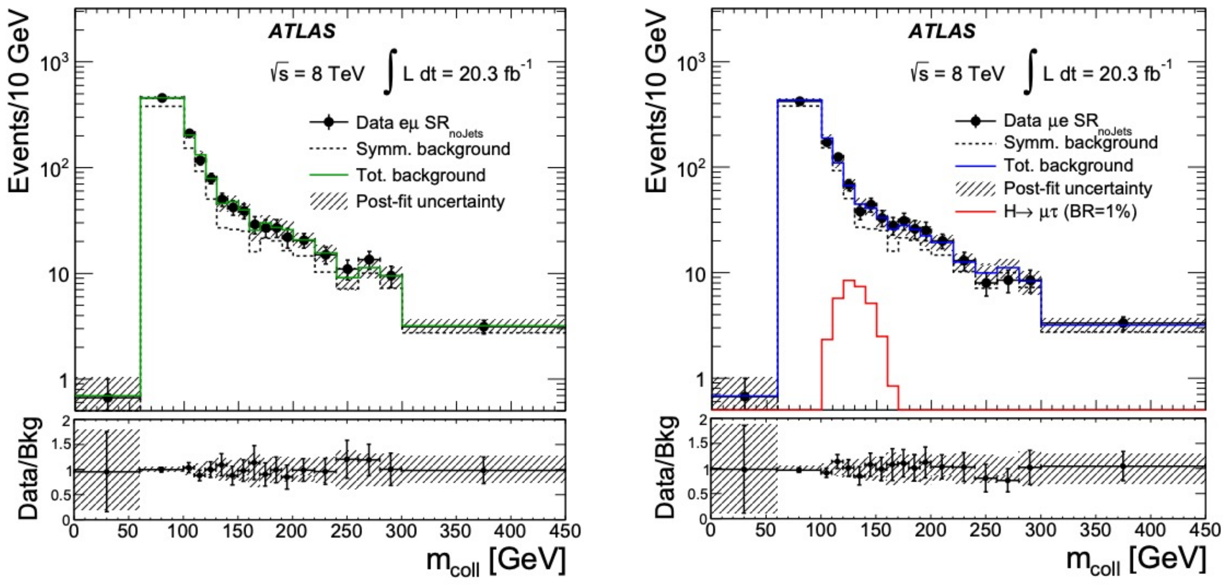
$$N^{e-based} = N^{\mu-based} \times \Re^{\mathcal{E}} \rightarrow \sum_{i}^{N^{e-based}} 1 = \sum_{i}^{N^{\mu-based}} \Re_i^{\mathcal{E}}$$

- True for any possible kinematic distribution
  - Not necessary of the leptons
    e.g. the $p_T$ distribution of the leading jet, number of jets, MET, invariant mass, …

# Ingredient I: $e \leftrightarrow \mu$ symmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

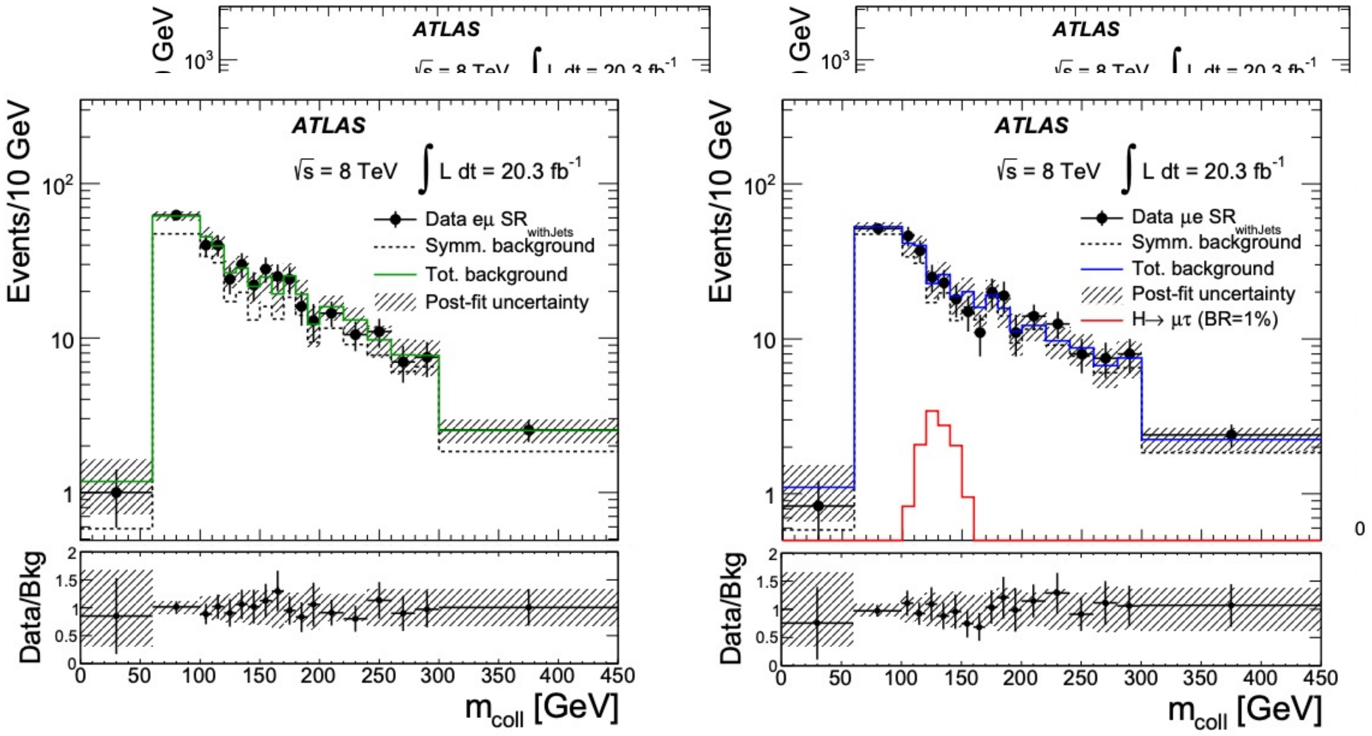
- In arXiv:1604.07730 we showed that for $e\mu \leftrightarrow \mu e$ events
- For every distribution in every final state
- E.g. (I can only present approved plots)
  - Collinear mass in events with 0 jets



# Ingredient I: $e \leftrightarrow \mu$ symmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

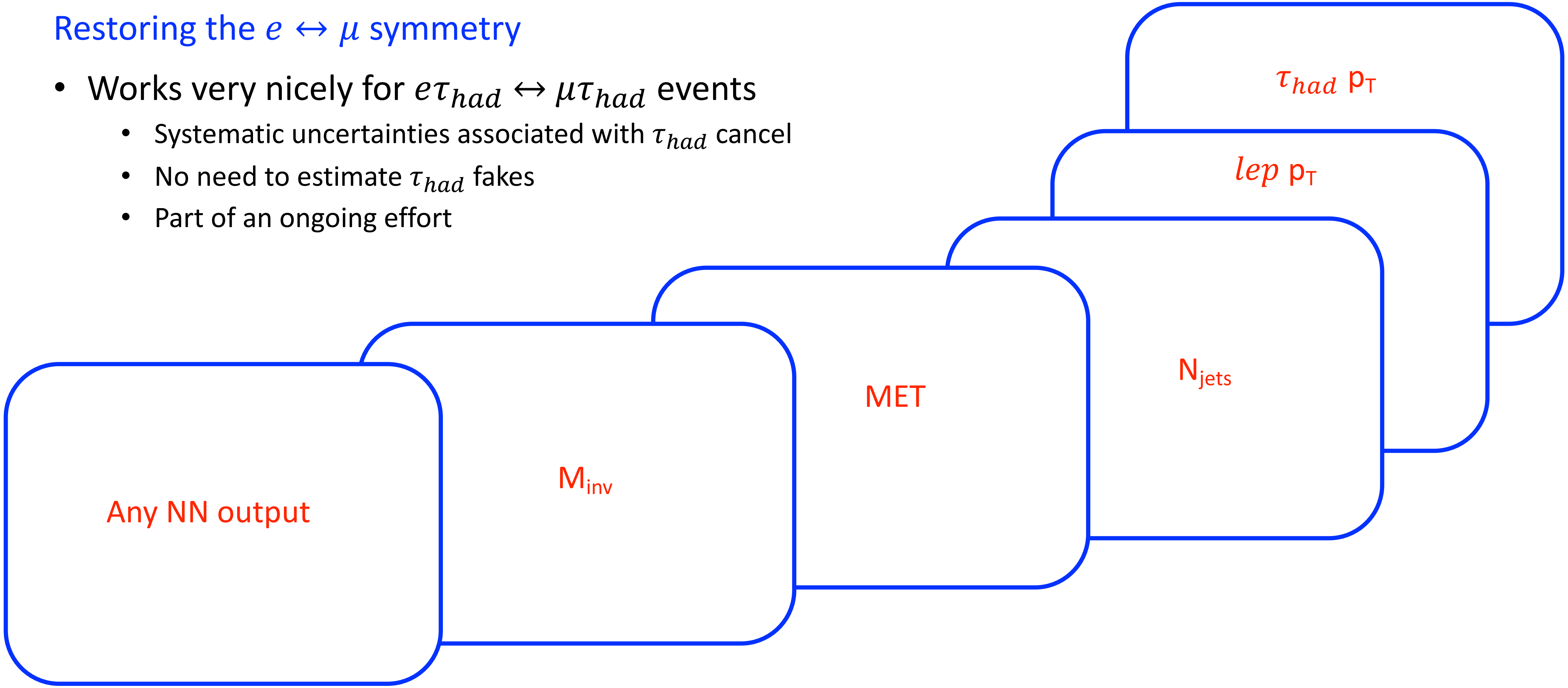
- In arXiv:1604.07730 we showed that for $e\mu \leftrightarrow \mu e$ events
- For every distribution in every final state
- E.g. (I can only present approved plots)
    - Collinear mass in events with 0 jets
    - Colinear mass in events with jets



# Ingredient I: $e\tau_{had} \leftrightarrow \mu\tau_{had}$ symmetry

Restoring the $e \leftrightarrow \mu$ symmetry

- Works very nicely for $e\tau_{had} \leftrightarrow \mu\tau_{had}$ events
  - Systematic uncertainties associated with $\tau_{had}$ cancel
  - No need to estimate $\tau_{had}$ fakes
  - Part of an ongoing effort



# A DDP search for $e \leftrightarrow \mu_{(\leftrightarrow \tau)}$ asymmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Correction can be made on an event-by-event basis
  - Replacing to total number of events by summation of event weights
    Excluding the fake contribution for simplicity

$$N^{e-based} = N^{\mu-based} \times \Re^{\mathcal{E}} \rightarrow \sum_{i}^{N^{e-based}} 1 = \sum_{i}^{N^{\mu-based}} \Re_i^{\mathcal{E}}$$

- True for any possible kinematic distribution
  - Not necessary of the leptons
    e.g. the $p_T$ distribution of the leading jet, number of jets, MET, invariant mass, …

# A DDP search for $e \leftrightarrow \mu_{(/\tau)}$ asymmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Correction can be made on an event-by-event basis
  - Replacing to total number of events by summation of event weights
    Excluding the fake contribution for simplicity

$$N^{e-based} = N^{\mu-based} \times \Re^{\mathcal{E}}$$

$$\sum_{i}^{N^{e-based}} 1 = \sum_{i}^{N^{\mu-based}} \Re_i^{\mathcal{E}}$$

- True for any possible kinematic distribution
  - Not necessary of the leptons
    e.g. the pT distribution of the leading jet, number of jets, MET, invariant mass, …
- We are left with two datasets s.t. any measured asymmetry between them is ***potentially interesting***

# A DDP search for $e \leftrightarrow \mu_{(/\tau)}$ asymmetry

## Restoring the $e \leftrightarrow \mu$ symmetry

- Correction can be made on an event-by-event basis
  - Replacing to total number of events by summation of event weights
    Excluding the fake contribution for simplicity

$$N^{e-based} = N^{\mu-based} \times \Re^{\mathcal{E}} \rightarrow \sum_i^{N^{e-based}} 1 = \sum_i^{N^{\mu-based}} \Re_i^{\mathcal{E}}$$

- True for any possible kinematic distribution
  - Not necessary of the leptons
    e.g. the pT distribution of the leading jet, number of jets, MET, invariant mass, …
- We are left with two datasets s.t. any measured asymmetry between them is ***potentially interesting***
- Need a tool to identify efficiently asymmetries between two data set
  → **Ingredient II**

# Ingredient II:Identifying asymmetries efficiently

A few notes

- The development of a tool to ***enhance*** asymmetries between two datasets is left to a future work
- A tool that identifies asymmetries efficiently allows fast scanning of many exclusive selections
  $\rightarrow$ Cover large region of the observable space
  - E.g. with and without jets, Different cut on jet $p_T$, angular distribution, MET, …
  - In the current implementation, the separation power is in all selected data
- "The tool" – a procedure applied to the two datasets that 'tells' asymmetry from symmetry
- The tool can be used to search for any asymmetries, not only $e \leftrightarrow \mu$
  - Forward backward, hemispheres, CP, …

# Ingredient II:Identifying asymmetries efficiently

Interpretation

- Most asymmetries are due to statistical fluctuations
  - Should be washed out with more data
  - We can start with Run-2 data and test over Run-3 data
- Some asymmetries are due to systematic uncertainties
  - Should appear also in SM MC
- Systematic uncertainties not modeled by the MC are likely to be detector/experiment specific
  - Can be confirmed or rejected with the data of other experiments

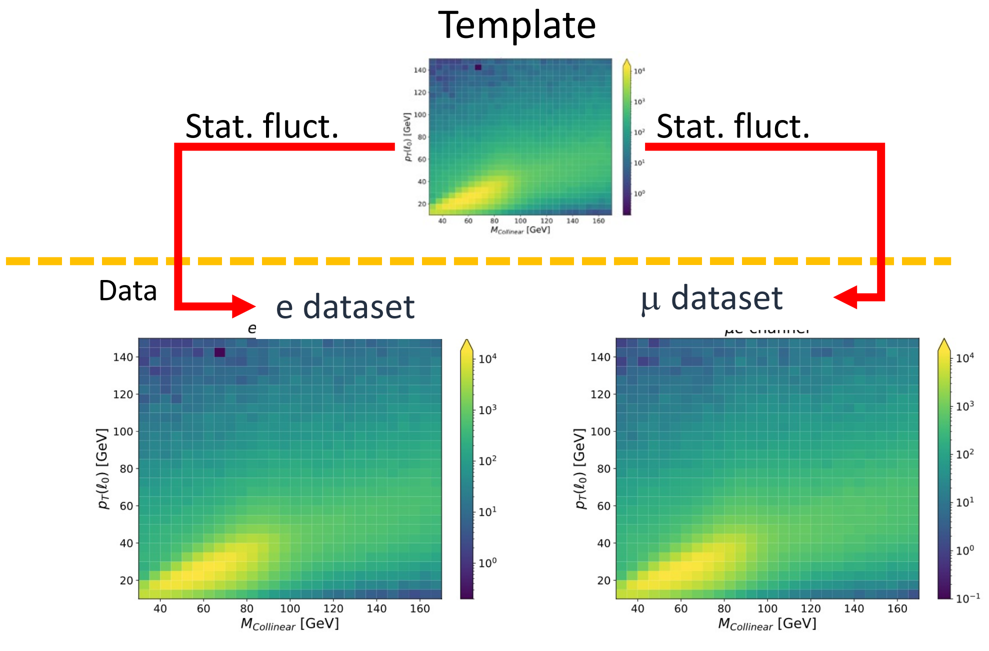
Our goal: Identify regions in the data which exhibit significant deviation from a SM prediction

- Rather than exhaustively study huge number of exclusive selections –
  **first find those that matter and invest the effort there**

# Ingredient II: Identifying asymmetries

## An example

- Project the two datasets into two N-Dim Matrices (here 2D for simplicity)
- Symmetry → the two matrices originate from a single, inaccessible, template matrix
- Various test statistics can be developed $t(A,B) = \frac{1}{M^2}\sum_{i,j}^{M}\frac{A_{ij}-B_{ij}}{\sqrt{B_{ij}}}$
  - In 1-Dim we could have used KS
- Is a given value of $t(A,B)$ interesting or not?



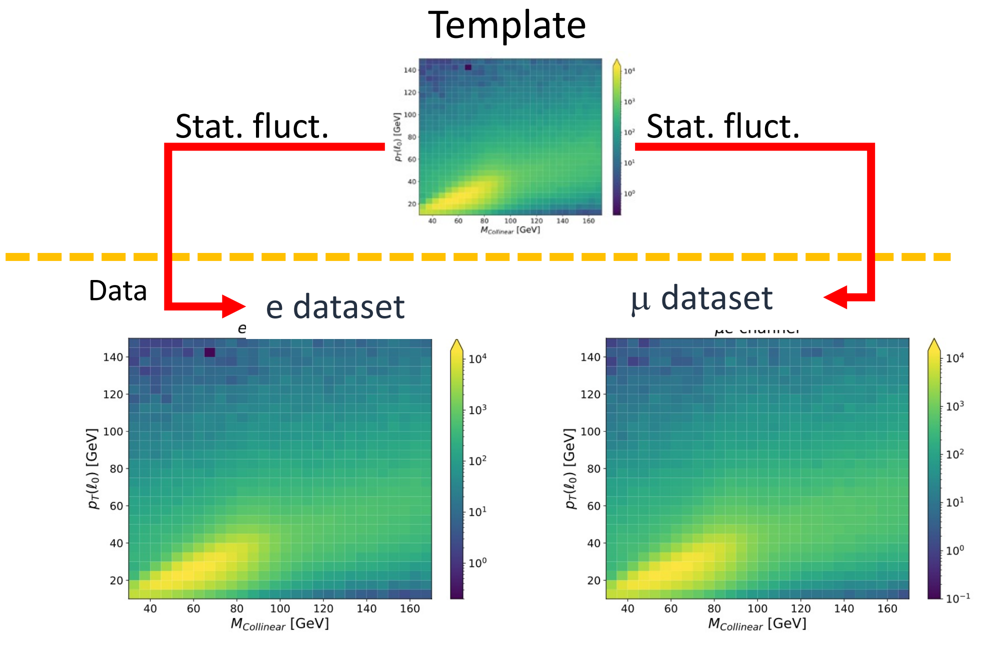

(a) $e\mu$ matrix

(b) $\mu e$ matrix

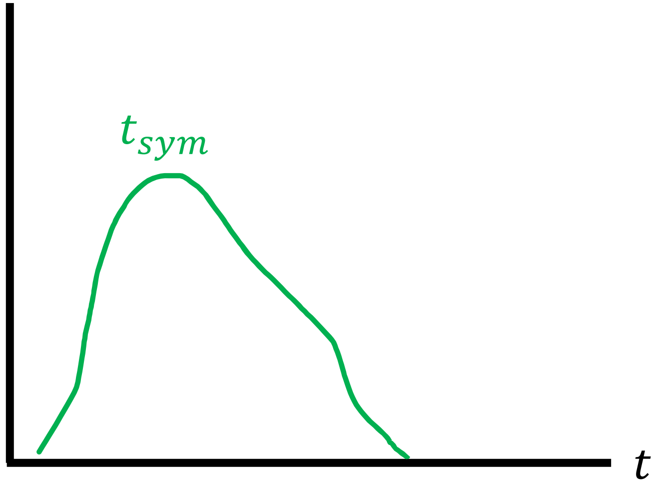
# Ingredient II: Identifying asymmetries

The procedure – for a given matrix T

- Evaluate the “symmetry pdf”
  - Draw two Matrices A and B and calculate $t(A, B) \equiv t_{sym}$
  - Repeat many times and draw the pdf of $t_{sym}$



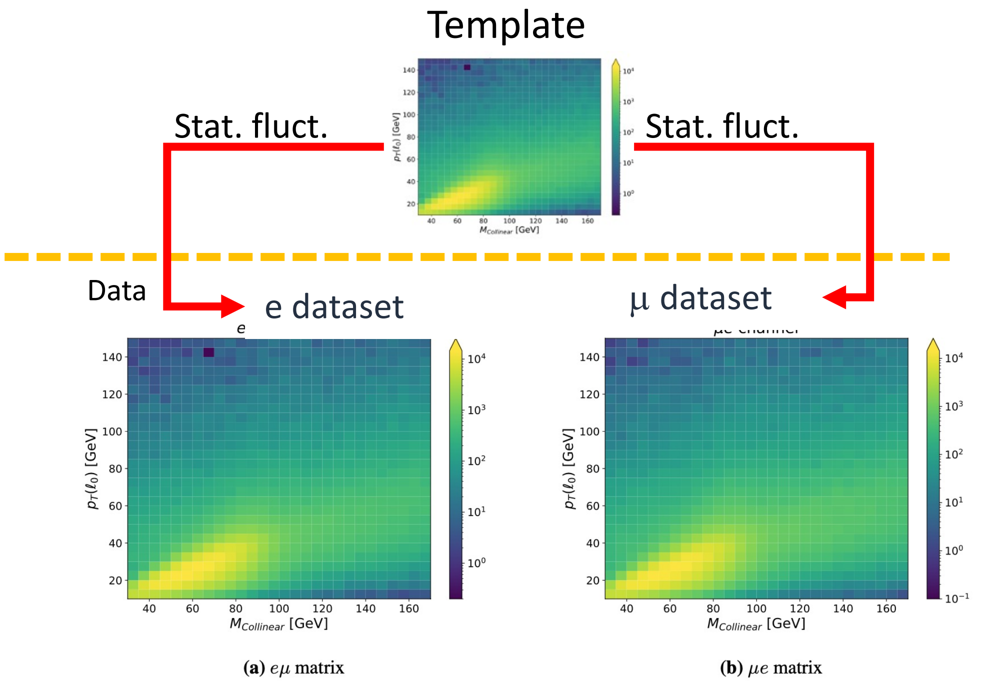

**(a)** $e\mu$ matrix **(b)** $\mu e$ matrix

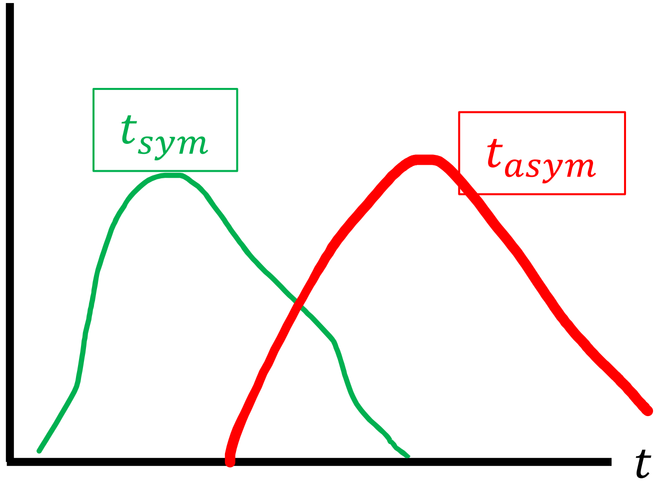


# Ingredient II: Identifying asymmetries

The procedure – for a given matrix T

- Evaluate the "symmetry pdf"
  - Draw two Matrices A and B and calculate $t(A, B) \equiv t_{sym}$
  - Repeat many times and draw the pdf of $t_{sym}$
- Evaluate the "asymmetry pdf"
  - Draw two Matrices A and B add a signal to B
    - Use Profile likelihood ratio test statistics to add a signal at a known significance
  - Calculate $t(A, B + S) \equiv t_{asym}$
  - Repeat many times and draw the pdf of $t_{asym}$

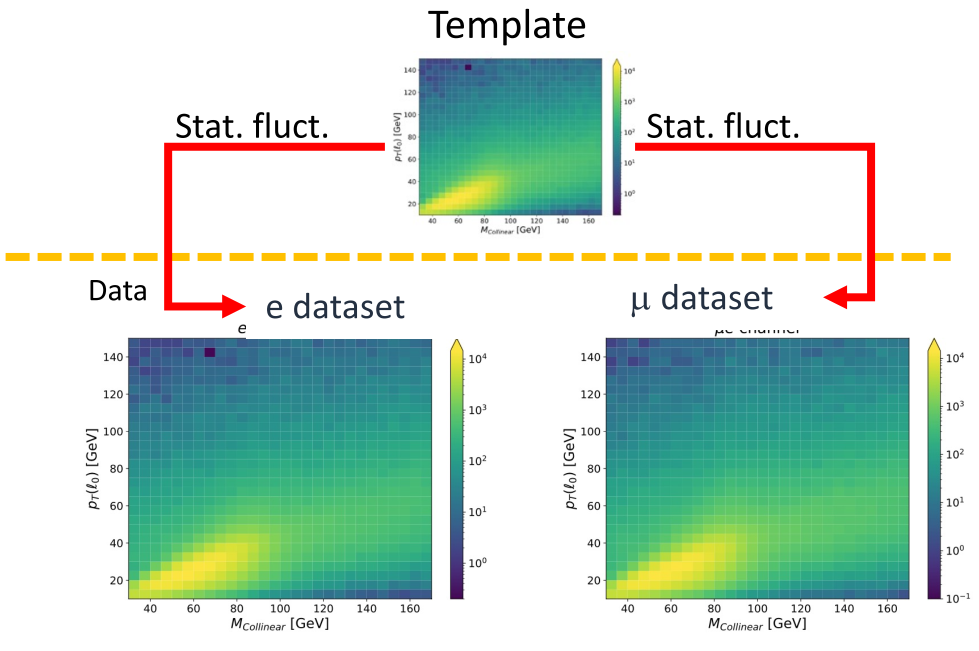


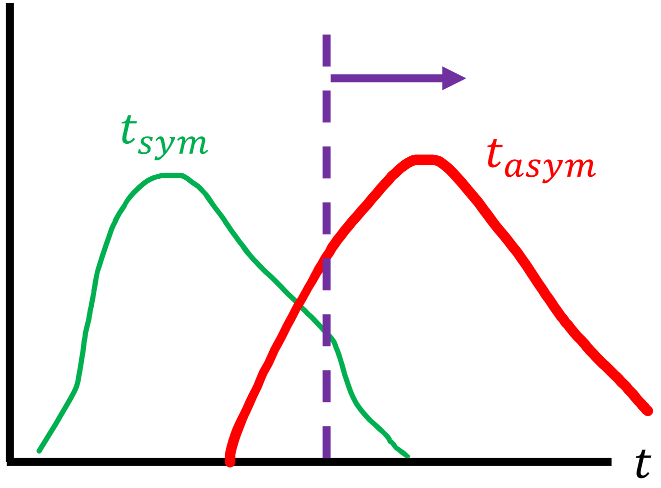
# Ingredient II: Identifying asymmetries

The procedure – for a given matrix T

- Evaluate the “symmetry pdf”
  - Draw two Matrices A and B and calculate $t(A,B) \equiv t_{sym}$
  - Repeat many times and draw the pdf of $t_{sym}$
- Evaluate the “asymmetry pdf”
  - Draw two Matrices A and B add a signal to B
    - Use Profile likelihood ratio test statistics to add a signal at a known significance
  - Calculate $t(A,B+S) \equiv t_{asym}$
  - Repeat many times and draw the pdf of $t_{asym}$
- Evaluate performance using ROC curves



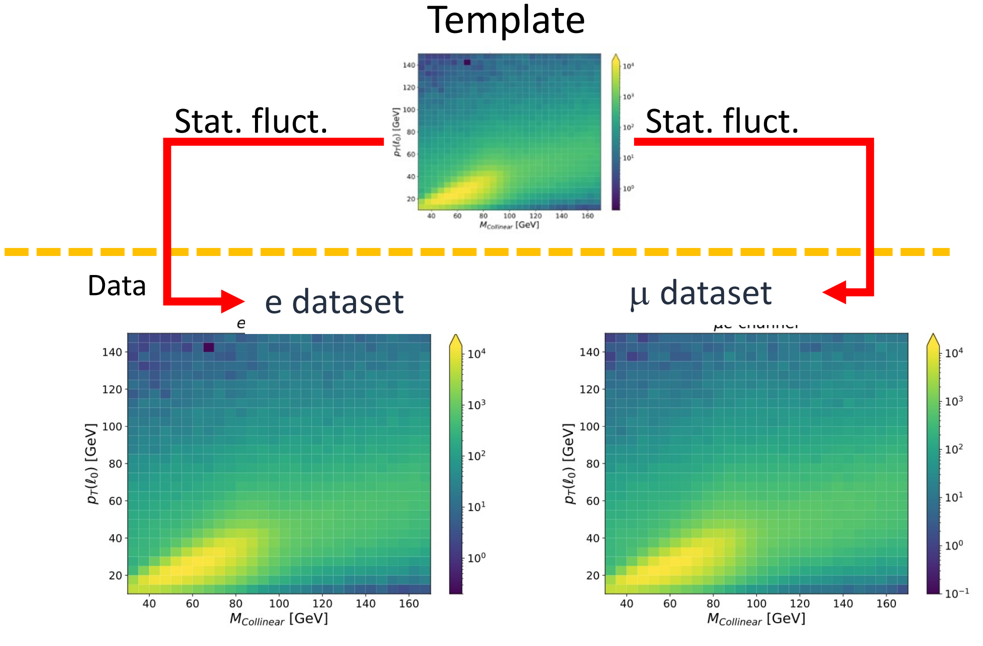

(a) $e\mu$ matrix (b) $\mu e$ matrix



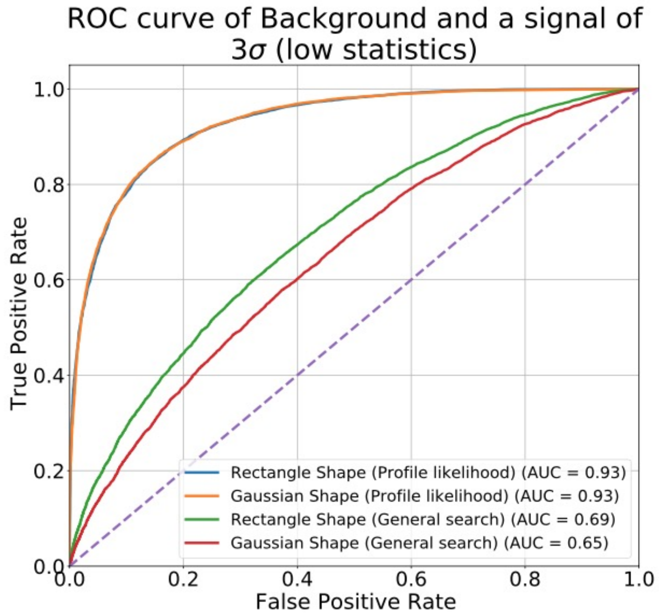
# Ingredient II: Identifying asymmetries

## The procedure – for a given matrix T

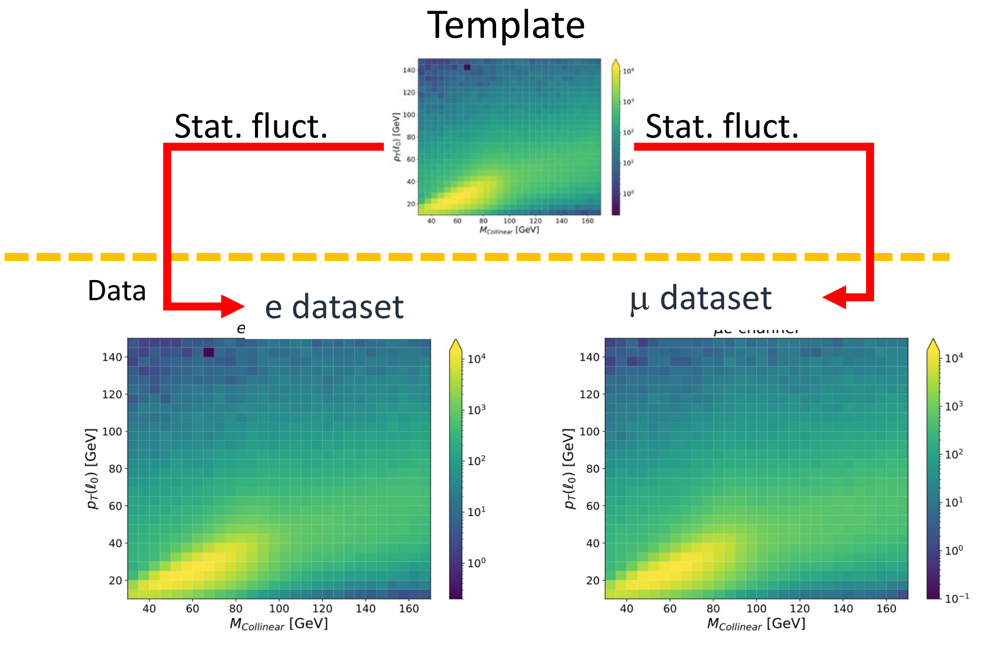
- Evaluate the “symmetry pdf”
  - Draw two Matrices A and B and calculate $t(A,B) \equiv t_{sym}$
  - Repeat many times and draw the pdf of $t_{sym}$
- Evaluate the “asymmetry pdf”
  - Draw two Matrices A and B add a signal to B
    - Use Profile likelihood ratio test statistics to add a signal at a known significance
  - Calculate $t(A,B+S) \equiv t_{asym}$
  - Repeat many times and draw the pdf of $t_{asym}$
- Evaluate performance using ROC curves
  - Compared to Profile Likelihood (PL) ratio – known **background and signal and no uncertainties**
- PL is more sensitive; ~30% better than the test statistics used




(a) $e\mu$ matrix (b) $\mu e$ matrix



# Ingredient II: Identifying asymmetries

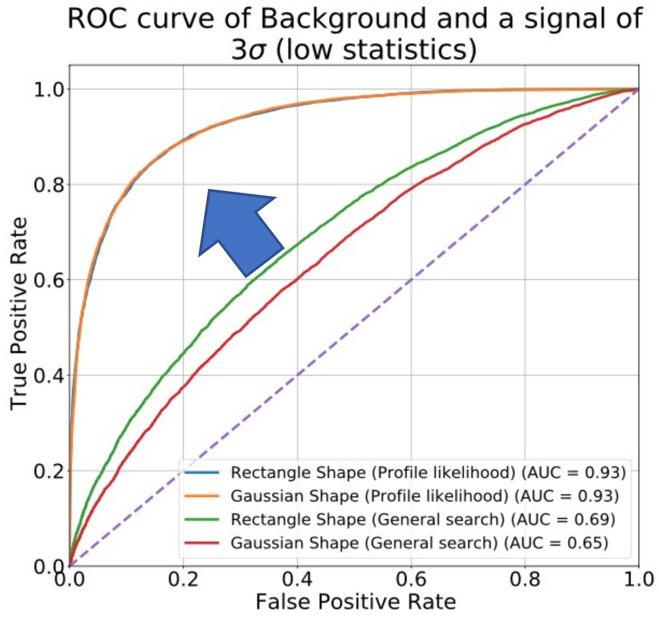



(a) $e\mu$ matrix

(b) $\mu e$ matrix

## The procedure – for a given matrix T

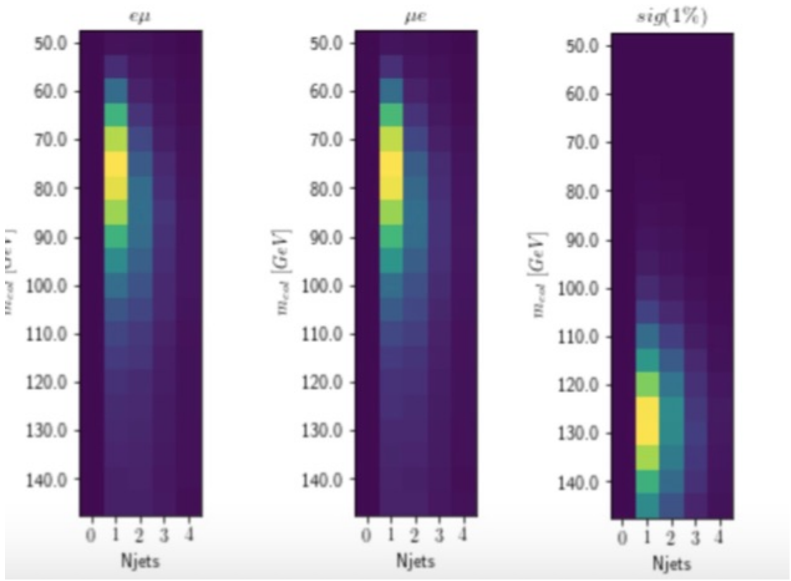
- Evaluate the "symmetry pdf"
  - Draw two Matrices A and B and calculate $t(A,B) \equiv t_{sym}$
  - Repeat many times and draw the pdf of $t_{sym}$
- Evaluate the "asymmetry pdf"
  - Draw two Matrices A and B add a signal to B
    - Use Profile likelihood ratio test statistics to add a signal at a known significance
  - Calculate $t(A,B+S) \equiv t_{asym}$
  - Repeat many times and draw the pdf of $t_{asym}$
- Evaluate performance using ROC curves
  - Compared to Profile Likelihood (PL) ratio – known **background and signal and no uncertainties**
- PL is more sensitive; ~30% better than the test statistics used
- Plenty of room for improvement
  - NN techniques could be used

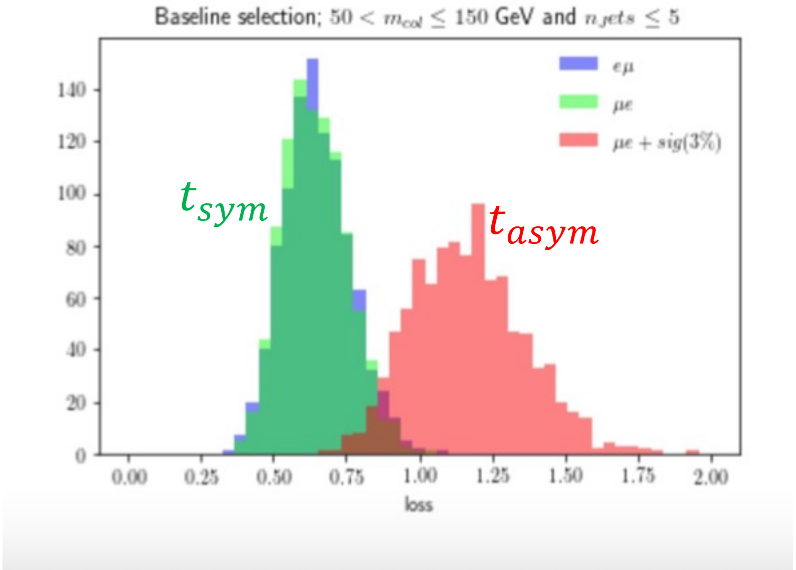


# Ingredient II:
# Identifying asymmetries

Another example

- Use $A$ as an estimator of the template matrix T
- Draw many matrices from $A$ - $C_A$
- Use the $C_A$-matrices to train an autoencoder
- Draw many matrices from $B$ - $D_B$
- Draw the pdf of the loss value of the autoencoder acting on $D_B \rightarrow t_{sym}$
- Add signal to the $D_B$ matrices - $D_{B+s}$
- Draw the pdf of the loss value of the autoencoder acting on $D_{B+s} \rightarrow t_{asym}$



# Ingredient II: Identifying asymmetries

Another example

- Calculate the error matrix, $\epsilon$, entry by entry $\epsilon_{ij} = \frac{A_{ij}-B_{ij}}{\sqrt{A_{ij}+B_{ij}}}$
- Symmetry $\rightarrow \epsilon_{ij}$ is consistent with an error distribution ("noise")
- Independent of the original template
- Train autoencoder / GAN / BIGAN to discriminate noise matrices from matrices containing something on top of the noise
- Analogous problems: taking two picture of the same chair and identify a mosquito in one of them…
  - There should be an elegant solution to this problem

# Take home messages

- In search for NP, we must leave no stone unturned
- The data-directed paradigm should be exploited
  - Complementary to the blind analysis paradigm
- The goal: define signal hypotheses on the basis of collected data
  - Exclusive selections which exhibit significant deviation from a SM property
- Two key ingredients
  - A property of the SM based on which deviations can be searched for
  - An efficient tool to find deviations
- The $e \leftrightarrow \mu \leftrightarrow \tau$ symmetry is a powerful tool in search for BSM physics
- Can and should be exploited in the context of the data directed paradigm
  - Other properties of interest are bumps and many other symmetries

**We have open PhD and postdoc positions to work on these projects**

Contact me at shikma.bressler@cern.ch