BDT-(un)folding of $\frac{d\sigma(H \rightarrow \gamma\gamma)}{dp_T^{\gamma\gamma}}$ measurements

J. W. Andrejkovic

Joint Annual Meeting of SPS and ÖPG, 21 - 25 August 2017 in Genève

Motivation

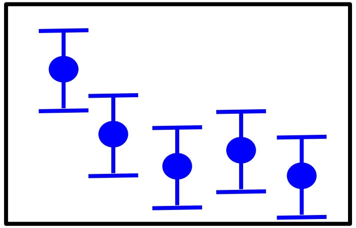
detector folding:

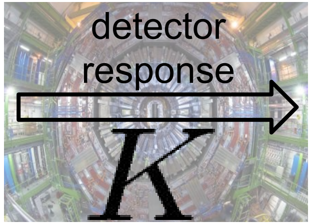
$$K\vec{x}_{\mathrm{true}} = \vec{y}_{\mathrm{data}}$$

nature

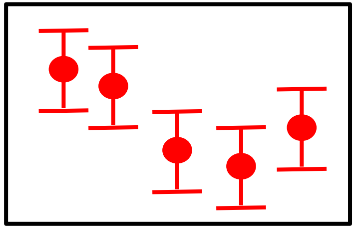
subject to stat. fluctuation

particle level
“un-smeared space”





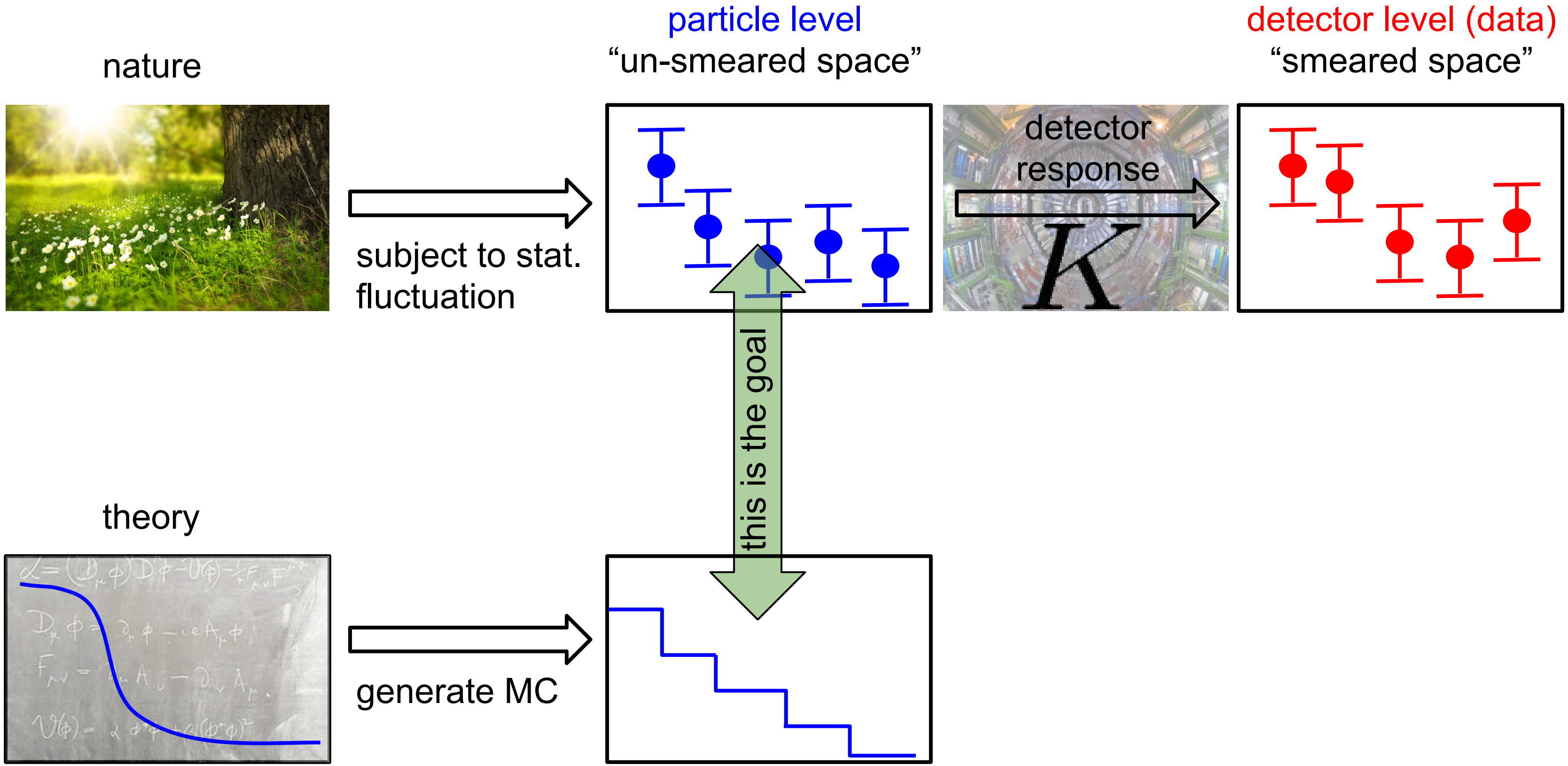
detector level (data)
“smeared space”



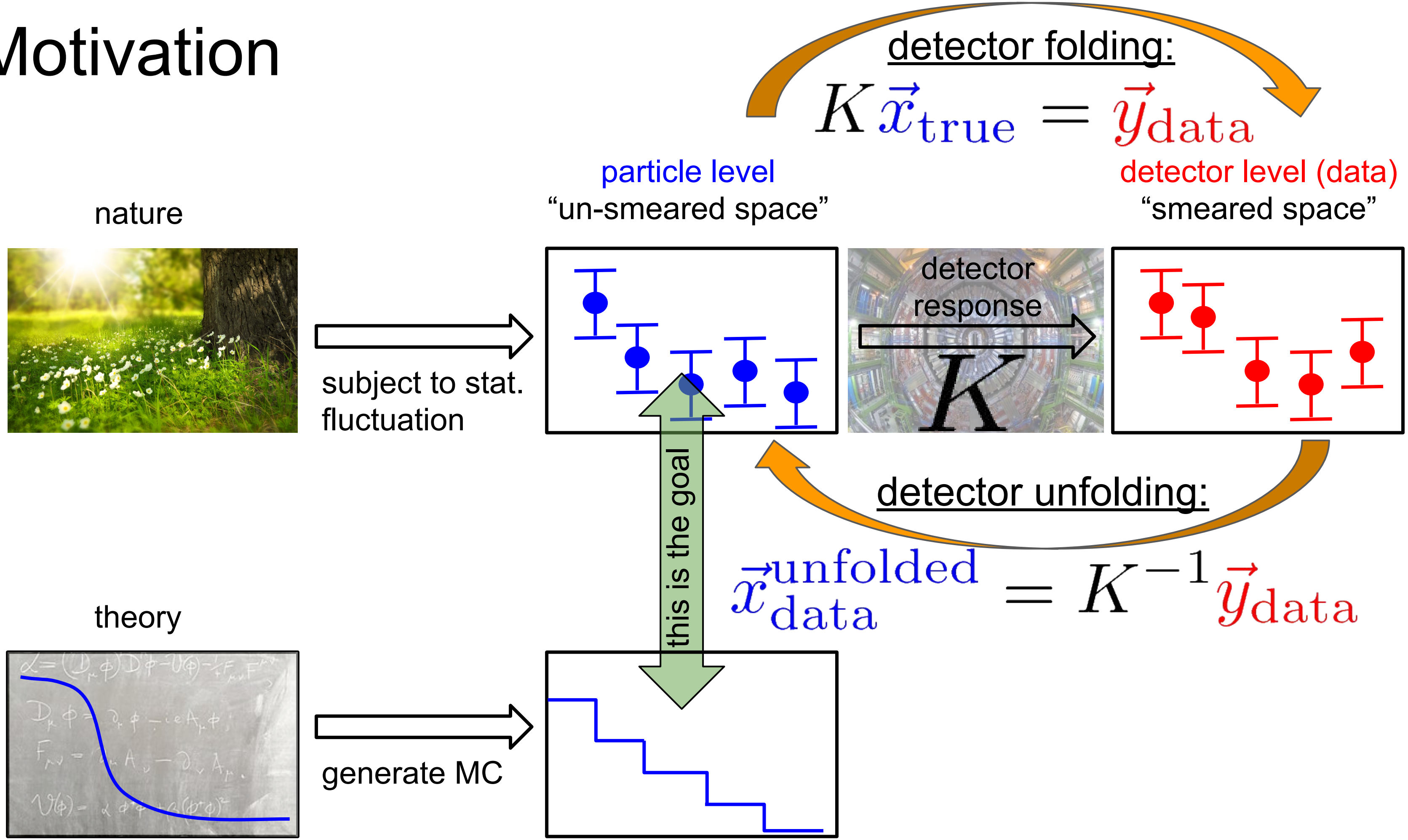
Motivation

detector folding:

$$K\vec{x}_{\mathrm{true}} = \vec{y}_{\mathrm{data}}$$



Motivation



detector folding:
$K\vec{x}_{\mathrm{true}} = \vec{y}_{\mathrm{data}}$
nature
particle level
“un-smeared space”
detector level (data)
“smeared space”
subject to stat. fluctuation
detector response
K
this is the goal
detector unfolding:
$\vec{x}^{\,\mathrm{unfolded}}_{\mathrm{data}} = K^{-1}\vec{y}_{\mathrm{data}}$
theory
generate MC

Motivation

Issues with unfolding:

- Neighbouring bins are correlated
- Unfolded spectra can be biased towards MC predictions
- Unfolded spectra can exhibit oscillations if not regularized

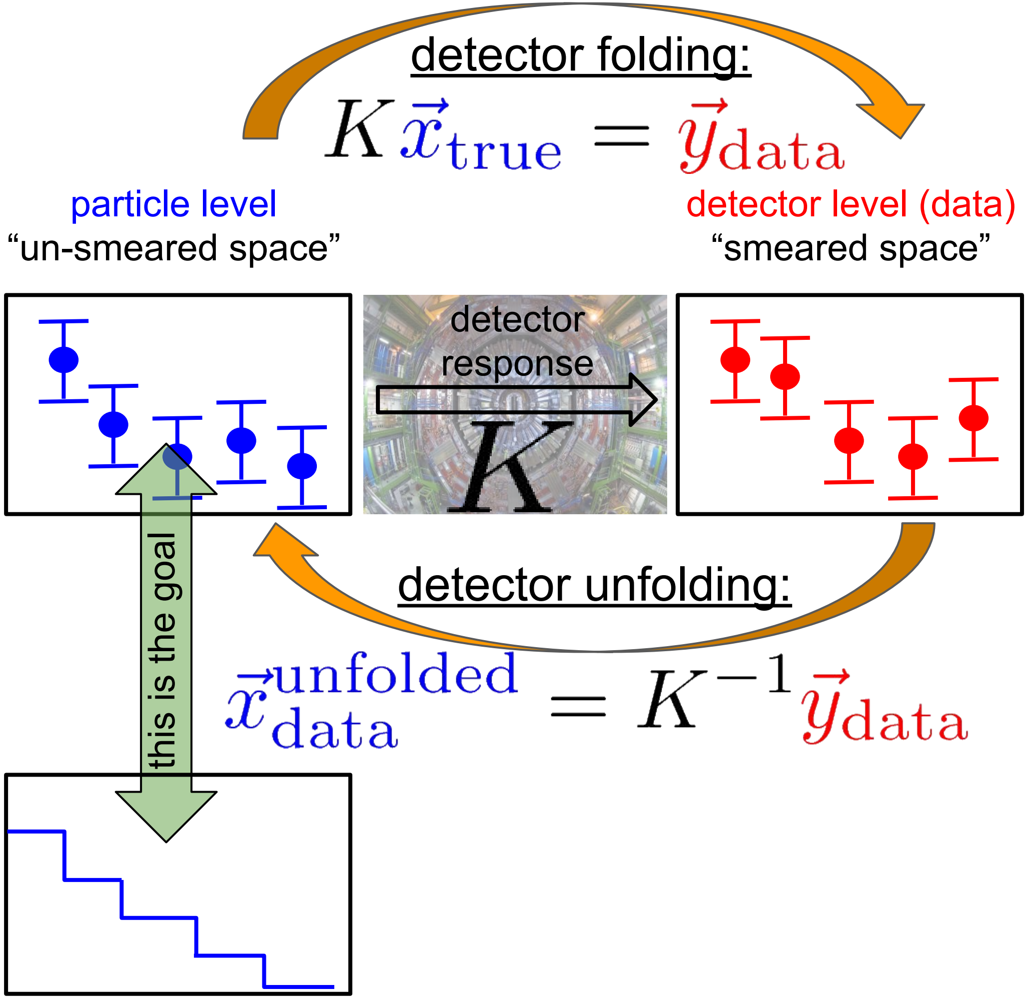
nature

subject to stat.

theory

generate MC

TODAY → Proof-of-principle that the folding approach is feasible



Motivation

detector folding:

$$K\vec{x}_{\text{true}} = \vec{y}_{\text{data}}$$

Issues with unfolding:

- Neighbouring bins are correlated
- Unfolded spectra can be biased towards MC predictions
- Unfolded spectra can exhibit oscillations if not regularized

nature

subject to stat. fluctuation

theory

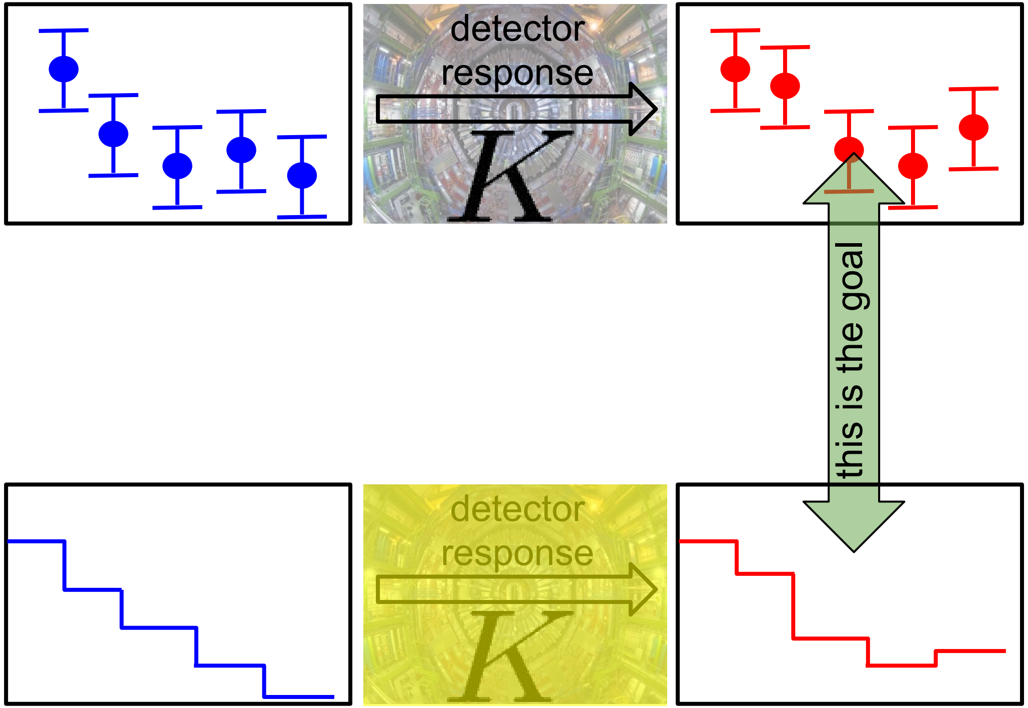
generate MC

TODAY

→ Proof-of-principle that the folding approach is feasible

particle level
“un-smeared space”

detector level (data)
“smeared space”

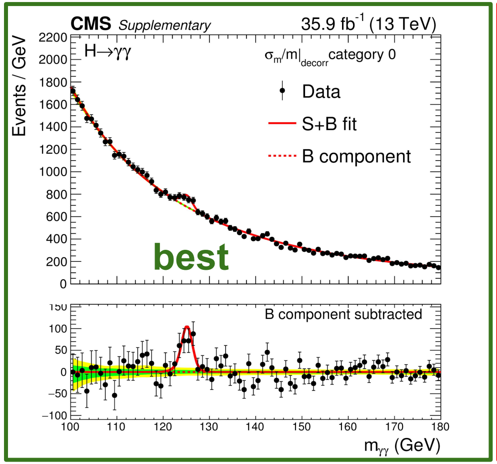


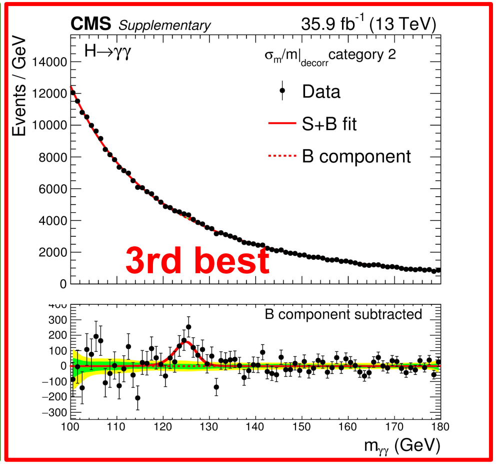
$\frac{d\sigma(\mathrm{H} \to \gamma\gamma)}{dp_{\mathrm{T}}^{\gamma\gamma}}$ measurement (part of HIG-17-015)

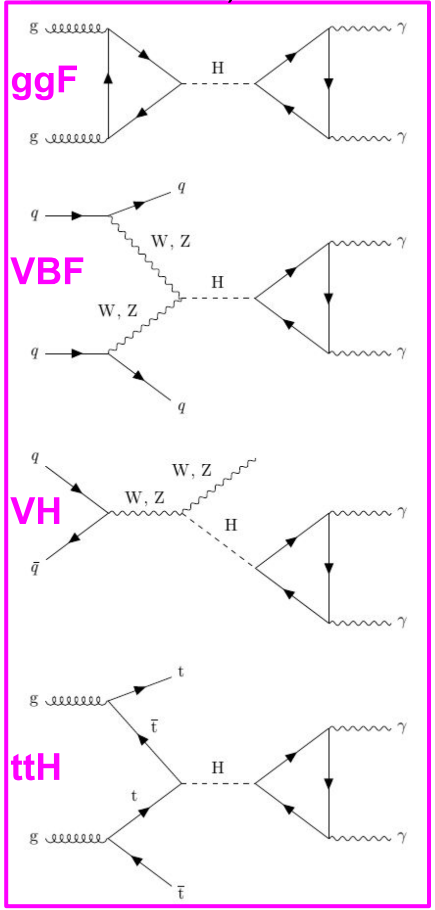
Higgs signal is a small bump over large background

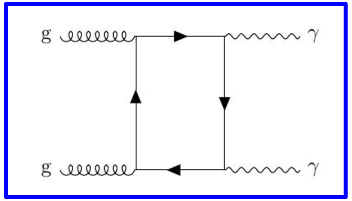
Gain in sensitivity when putting events in categories of mass resolution $\frac{\sigma_M}{M}$:

- **best**
- **2nd best**
- **3rd best**







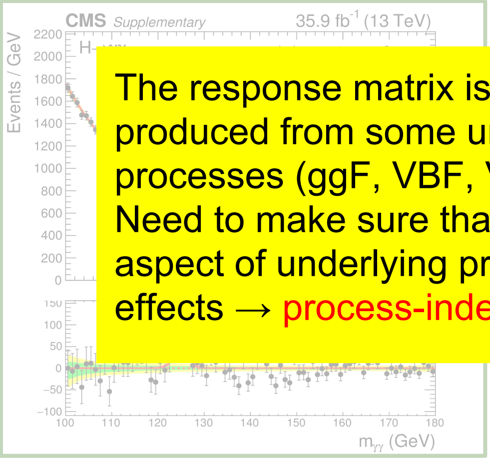


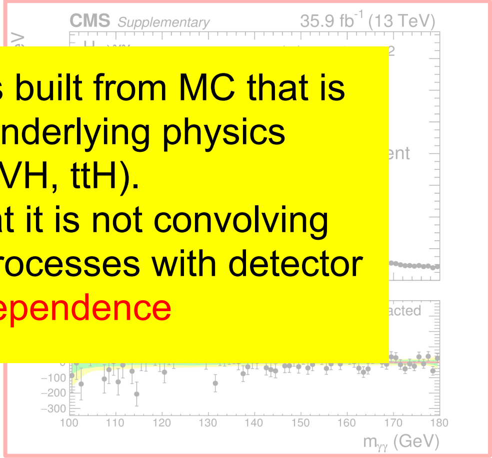
$\frac{d\sigma(\mathrm{H}\to\gamma\gamma)}{dp_{\mathrm{T}}^{\gamma\gamma}}$ measurement (part of HIG-17-015)

Higgs signal is a small bump over large background

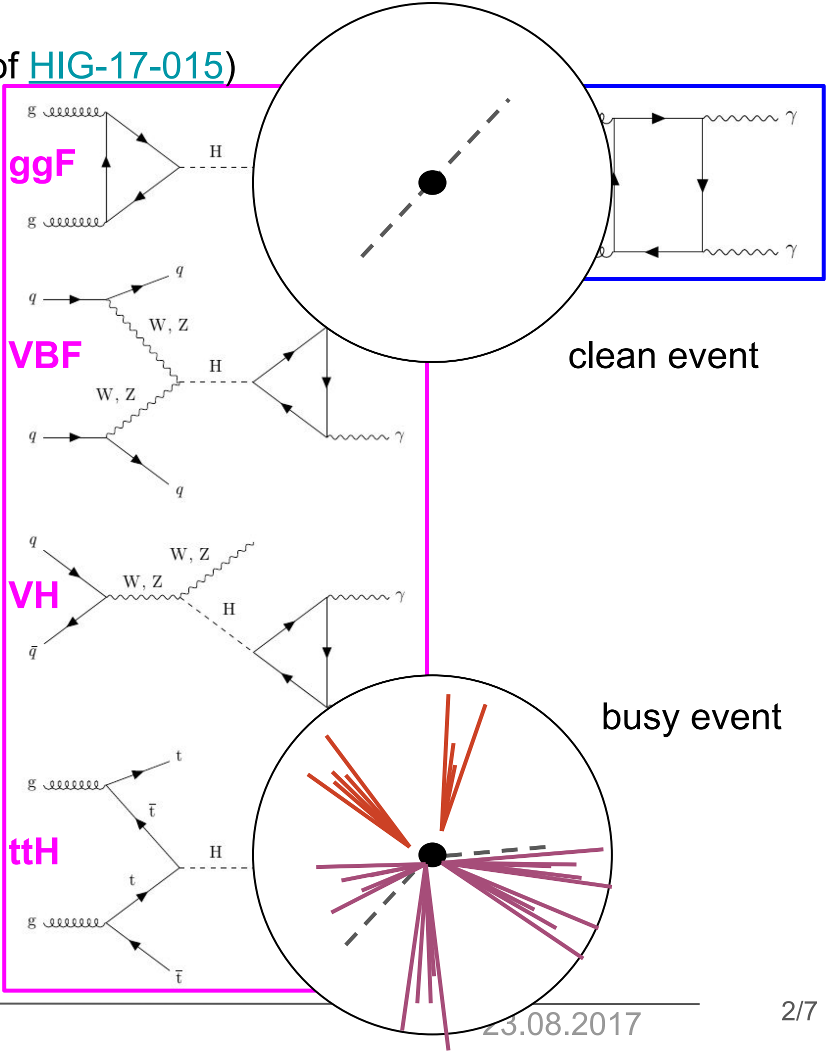
Gain in sensitivity when putting events in categories of mass resolution $\frac{\sigma_M}{M}$:

- best
- 2nd best
- 3rd best





The response matrix is built from MC that is produced from some underlying physics processes (ggF, VBF, VH, ttH).
Need to make sure that it is not convolving aspect of underlying processes with detector effects → process-independence



Developing the machinery

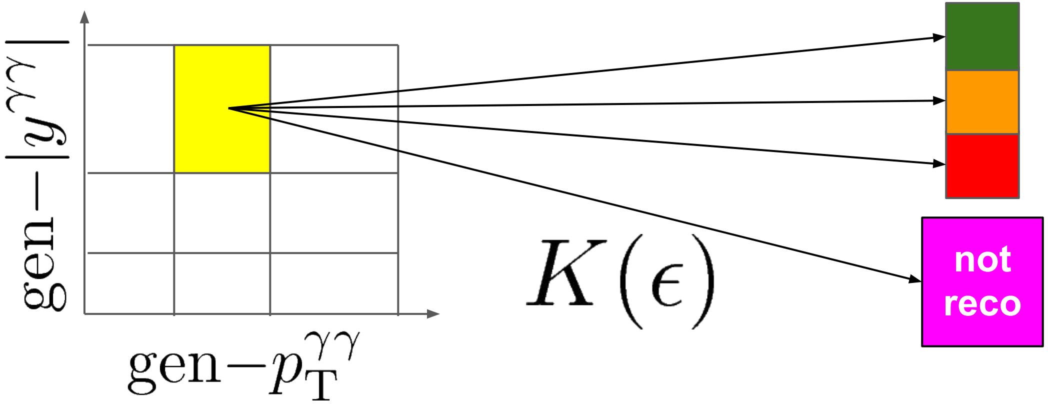
classes

- **best**
- **2nd best**
- **3rd best**
- **not-reconstructed**

Diphoton event characterized by particle-level variables

ϕ phase space $\begin{cases} \text{gen}-p_{\mathrm{T}}^{\gamma\gamma} \\ \text{gen}-|y^{\gamma\gamma}| \end{cases}$

detector folding



ϵ^0 probability to be mapped to best $\frac{\sigma_M}{M}$ category

ϵ^1 probability to be mapped to 2nd best $\frac{\sigma_M}{M}$ category

ϵ^2 probability to be mapped to 3rd best $\frac{\sigma_M}{M}$ category

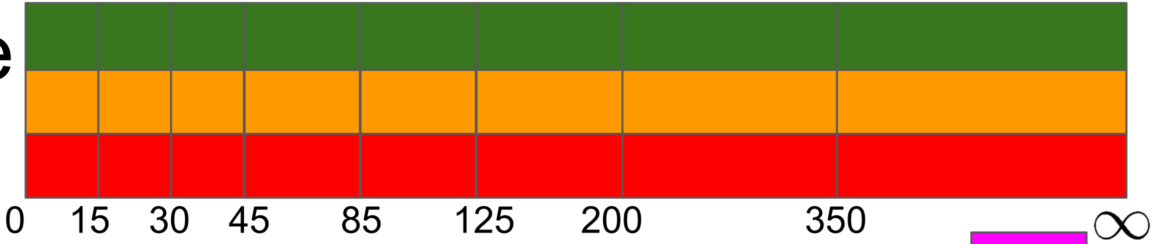
ϵ^3 probability to be not reconstructed

How to choose the partitioning in a smart way?
Gets more and more difficult in higher dimensions

Idea: Use BDT to find optimal partitioning:
Feed gen-variables as inputs, classes as targets and extract the probabilities (efficiencies)

$$\boxed{\epsilon \approx \mathrm{BDT}^{(\mathrm{SM})}}$$

Adding observable

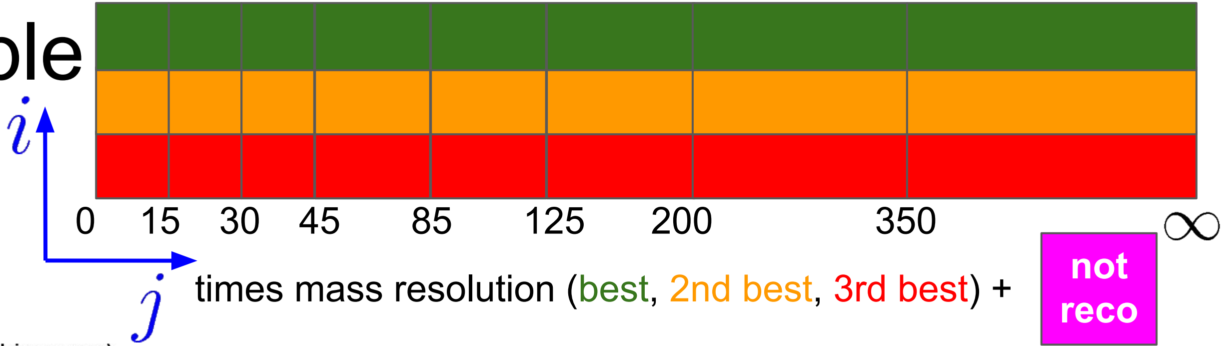


Diphoton transverse momentum $\text{reco}-p_{\mathrm{T}}^{\gamma\gamma}$

times mass resolution (best, 2nd best, 3rd best) + not reco

→ 25 classes

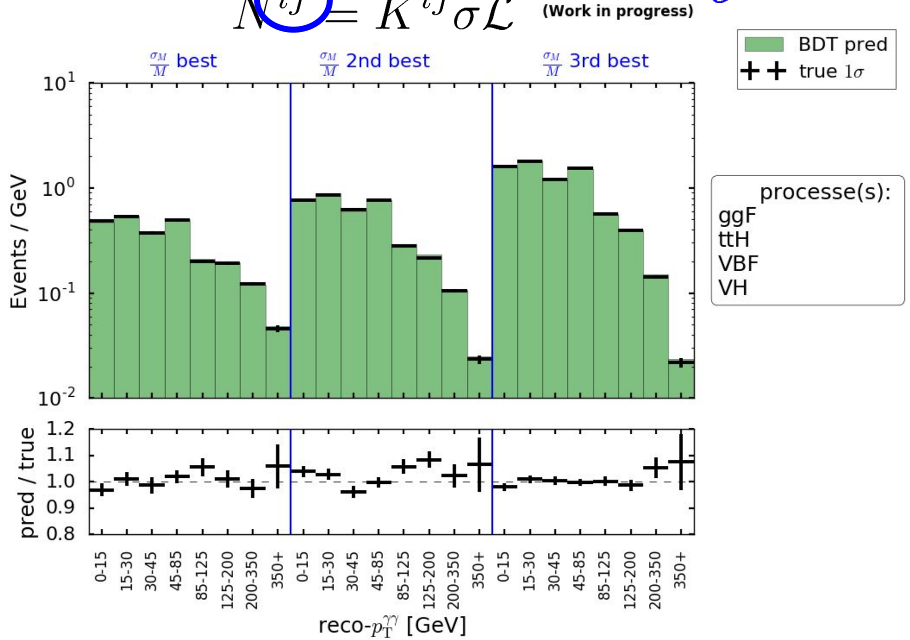
Adding observable



Diphoton transverse momentum $\text{reco}-p_T^{\gamma\gamma}$

→ 25 classes

$$N^{ij} = K^{ij}\sigma\mathcal{L}$$



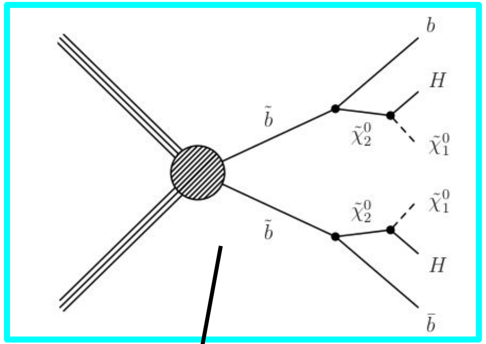
From histograms one can build the response matrices by normalization:

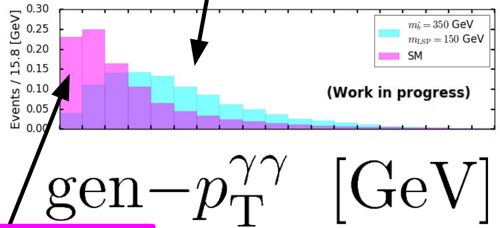
- True gives the **"SM response"**, i.e. the true response matrix using MC from all main production mechanisms
- BDT gives the **"BDT response"**

With the BDT we parametrize the response matrix and it is easy to re-compute the response matrix with an arbitrary model

Test BDT on unknown topology

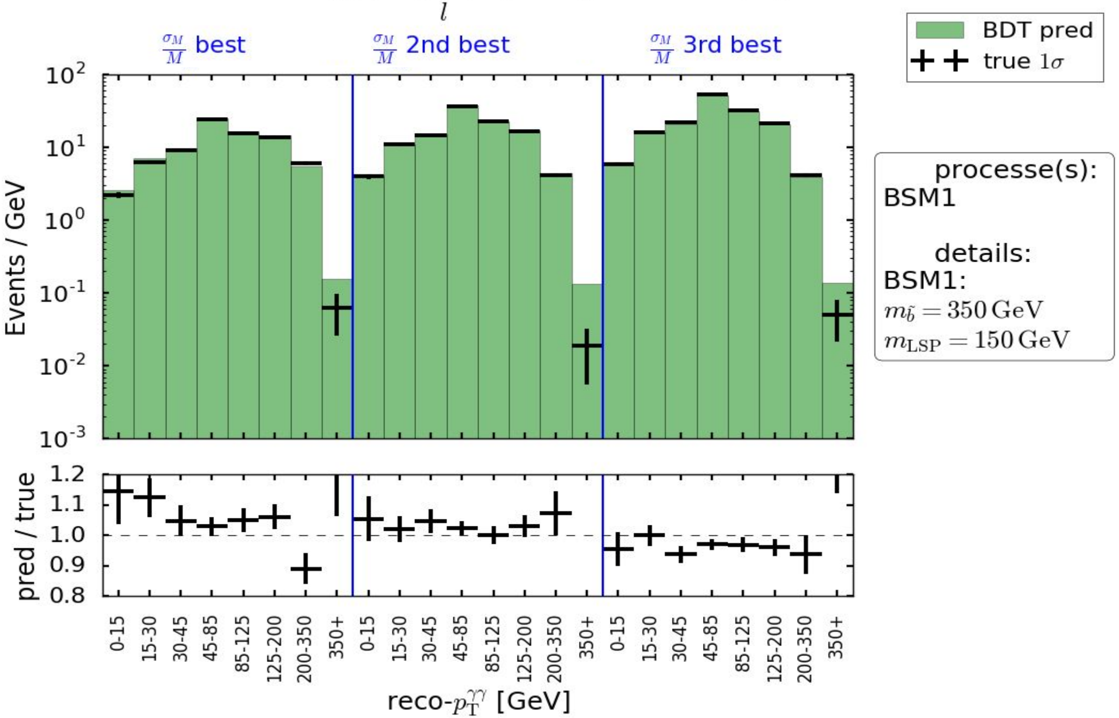
preliminary SUSY study





SM distribution

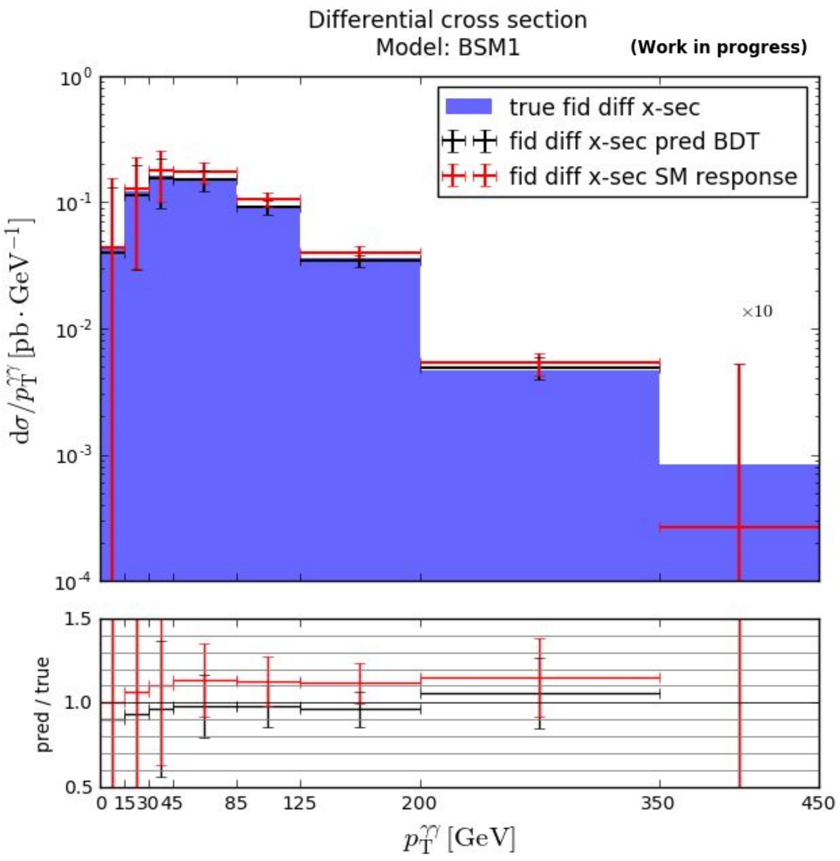
$$N^{ij}_{\mathrm{reco}} = \sum_l K^{ij}_l \mathcal{L}\, \sigma_l$$



BDT trained with SM MC can predict correctly the BSM spectrum

→ BDT selection efficiencies are modeled process-independently

Bias correction within unfolding procedure



Black: response matrix is computed from evaluating the BDT on the BSM MC sample (no re-training)

Red: the true response matrix using SM MC samples is used (introduce a bias). (See slide 4)

→ BDT approach allows for the correction of this bias

Summary

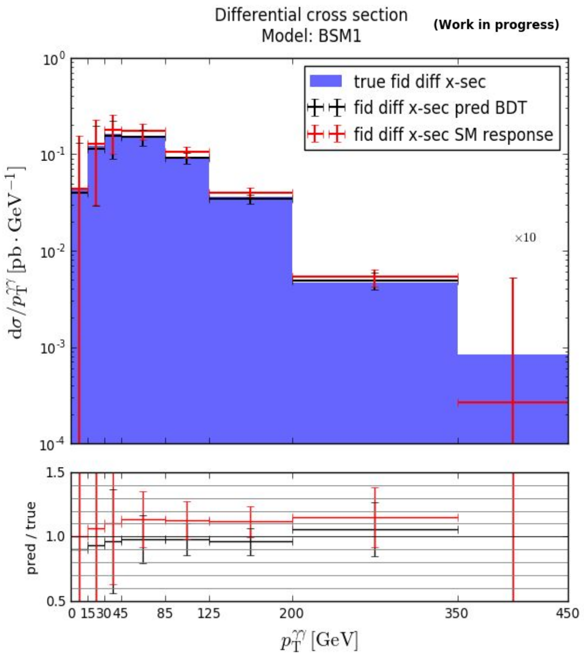
- Shown how to parametrize the detector response in a process-independent way
  - Selection efficiencies trained using SM MC samples
  - Works also on topologies never seen to the BDT (BSM)
- Bias correction within the unfolding

- Reduces model-dependence of unfolded results in the context of differential cross section measurements.
- The whole description is very general and can be applied elsewhere (e.g. other decay channels)
- Allows a comparison of theory and CMS data in the “smeared space”
- Could be used in a “fast folding” kind of way to get a lot of reco-MC from gen-MC by mapping to detector level histograms

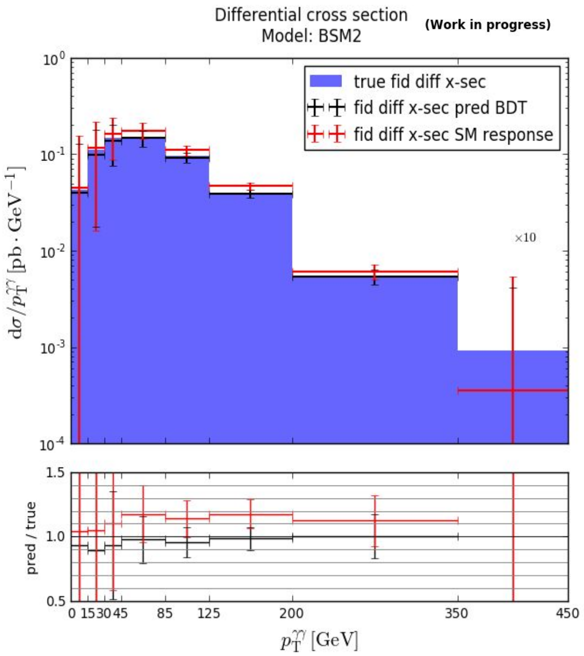
back-up

Bias correction within unfolding procedure

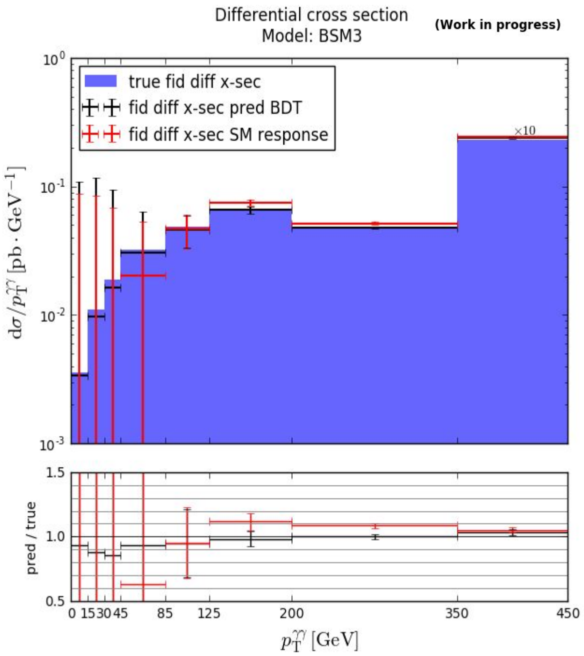
BSM1



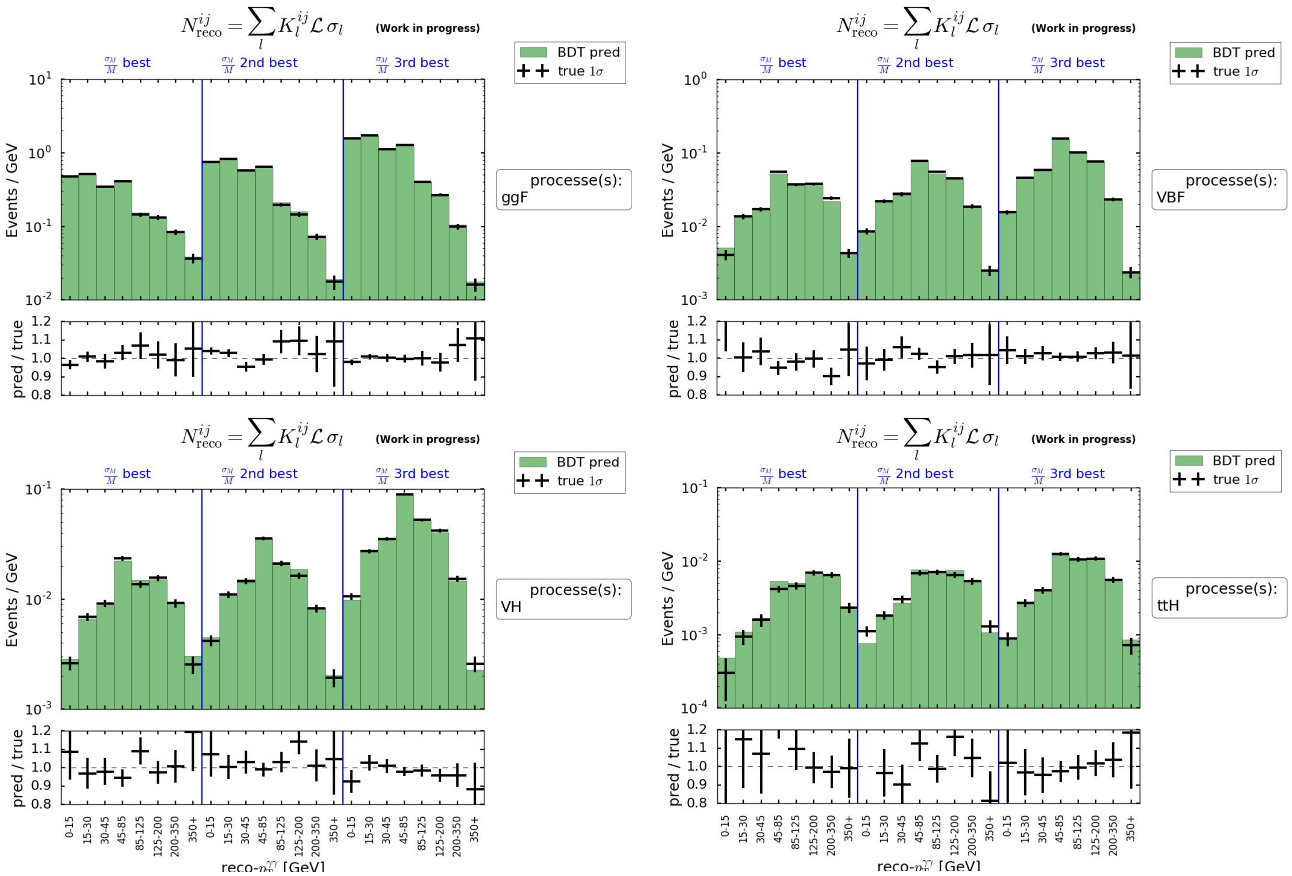
BSM2



BSM3



Examine each production process



Good agreement indicates process-independence of the BDT selection efficiencies

BSM closure test

SM training set

$$\epsilon^{ij}(\phi) \approx \mathrm{BDT}^{ij}(\phi)^{(\mathrm{SM})}$$

SM test set

$$N^{ij}_{\mathrm{reco,pred}} = \sum_l \underbrace{\left(\int_l \mathrm{d}\phi \, \rho^{(\mathrm{SM})}(\phi) \, \overbrace{\epsilon^{ij}(\phi)}^{\approx \mathrm{BDT}^{ij}(\phi)^{(\mathrm{SM})}} \right)}_{K_l^{ij}} \mathcal{L} \, \sigma_l$$

$\rho^{(\mathrm{ggF})}(\phi)$

$\rho^{(\mathrm{ttH})}(\phi)$

$\rho^{(\mathrm{VBF})}(\phi)$

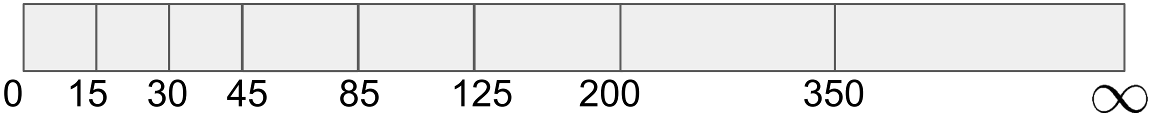
$\rho^{(\mathrm{VH})}(\phi)$

BSM set

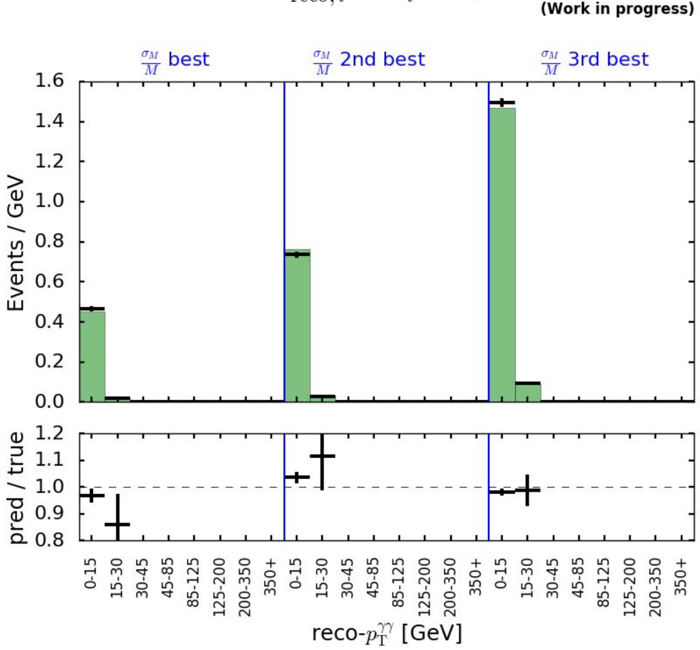
$$N^{ij}_{\mathrm{reco,pred}} = \sum_l \underbrace{\left(\int_l \mathrm{d}\phi \, \rho^{(\mathrm{BSM})}(\phi) \, \overbrace{\epsilon^{ij}(\phi)}^{\approx \mathrm{BDT}^{ij}(\phi)^{(\mathrm{SM})}} \right)}_{K_l^{ij}} \mathcal{L} \, \sigma_l$$

Go differential

$$\mathrm{gen}\!-\!p_{\mathrm{T}}^{\gamma\gamma}$$



$$N^{ij}_{\mathrm{reco},\,l} = K^{ij}_{l}\,\mathcal{L}\,\sigma_l$$



BDT pred
true 1σ

$$K^{ij}_{l} = \int_{l} \mathrm{d}\phi\, \rho(\phi)^{(\mathrm{SM})}\, \epsilon^{ij}(\phi)$$

$l=$ gen-$p_{\mathrm{T}}^{\gamma\gamma}$: 0-15 GeV

processe(s):
ggF
ttH
VBF
VH