# From the analysis of the observed Higgs boson coupling structure to the search for more Higgs bosons

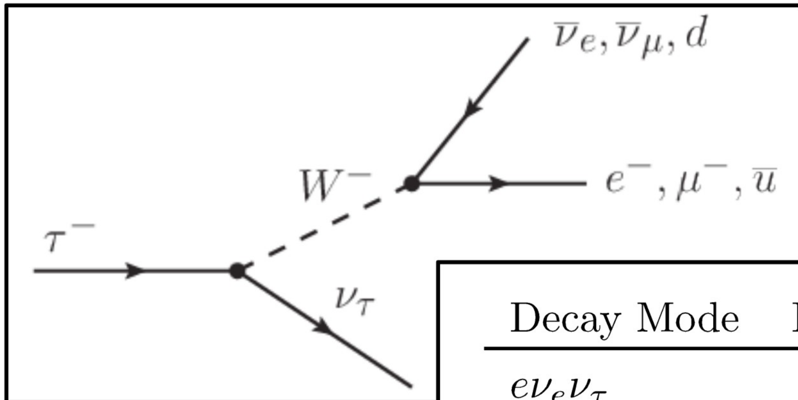
## – Higgs boson analyses in the di-τ final state –

**Roger Wolf**

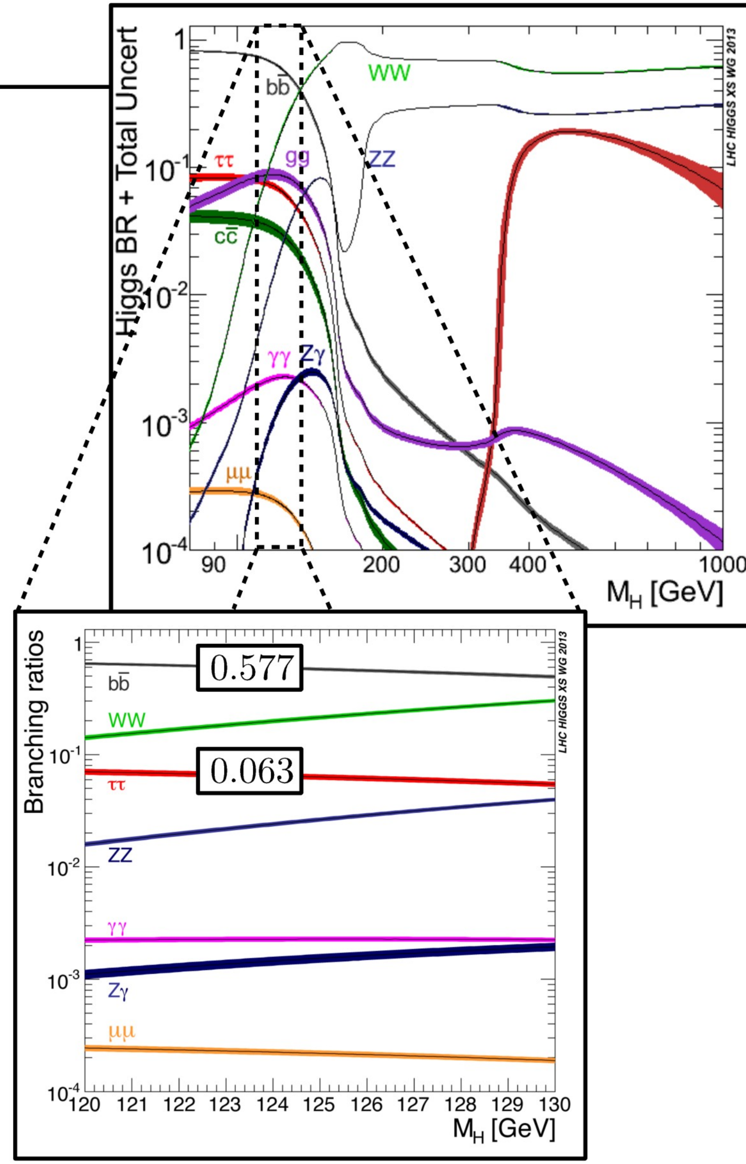
**23. Mai 2019**

# $\tau$-leptons & LHC Higgs physics

- With 1.77 GeV the **heaviest known lepton**.




| Decay Mode | BR [%] |
|---|---|
| $e\nu_e\nu_\tau$ | 17.83 |
| $\mu\nu_\mu\nu_\tau$ | 17.41 |
| 1-prong $\nu_\tau$ | 37.10 |
| 3-prong $\nu_\tau$ | 15.20 |

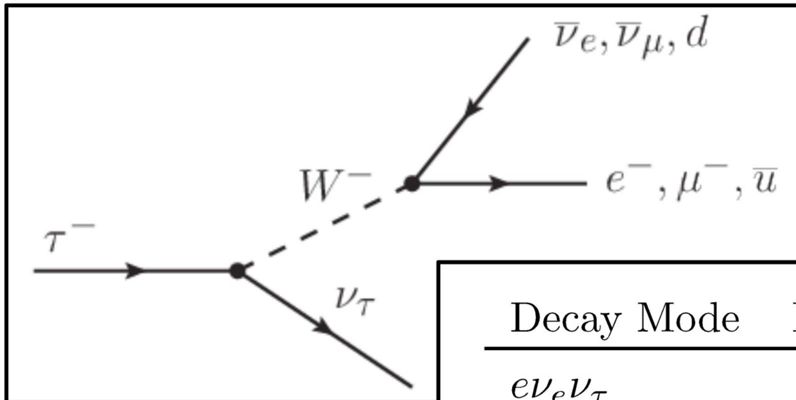



- One of the **big five** in the investigation of the Higgs sector @ low mass.

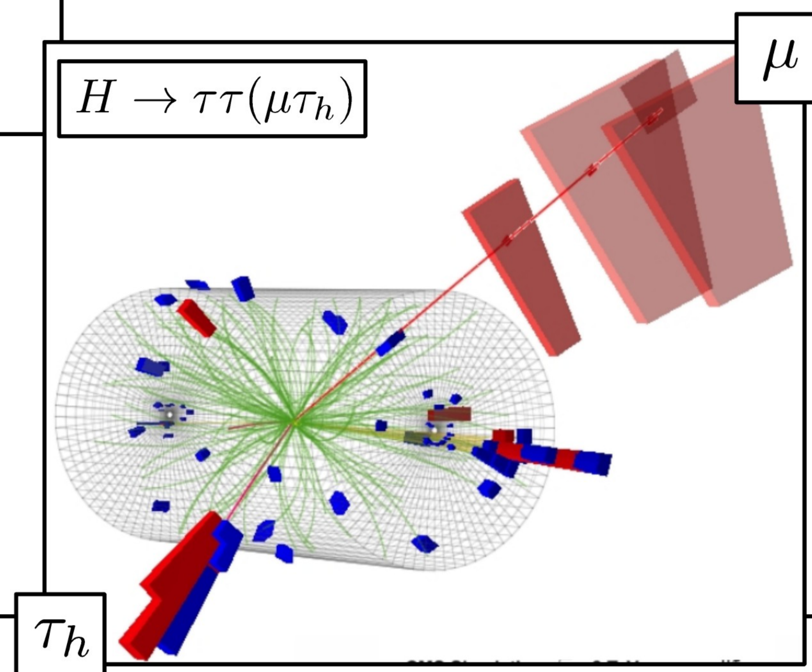
| | Channel | Resolution | S/B |
|---|---|---|---|
| $\kappa_{HVV} = \frac{2m_V^2}{v}$ | $H \to \gamma\gamma$ | 1-2% | $\mathcal{O}(0.1)$ |
| | $H \to ZZ$ | 1-2% | $\mathcal{O}(>1)$ |
| | $H \to WW$ | 20% | $\mathcal{O}(1)$ |
| $\kappa_{Hff} = \frac{m_f}{v}$ | $H \to b\bar{b}$ | 10% | $\mathcal{O}(0.1)$ |
| | $H \to \tau\tau$ | 15% | $\mathcal{O}(0.1)$ |

# Di-$\tau$ final state

- High mass allows for **decays into hadrons**:

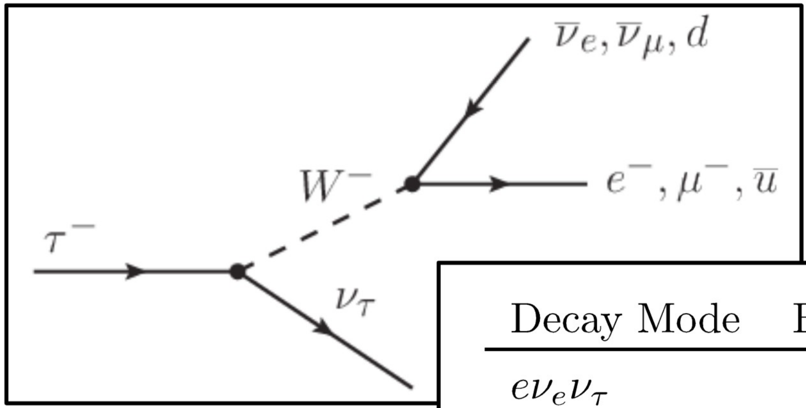



| Decay Mode | BR [%] |
|---|---|
| $e\nu_e\nu_\tau$ | 17.83 |
| $\mu\nu_\mu\nu_\tau$ | 17.41 |
| 1-prong $\nu_\tau$ | 37.10 |
| 3-prong $\nu_\tau$ | 15.20 |



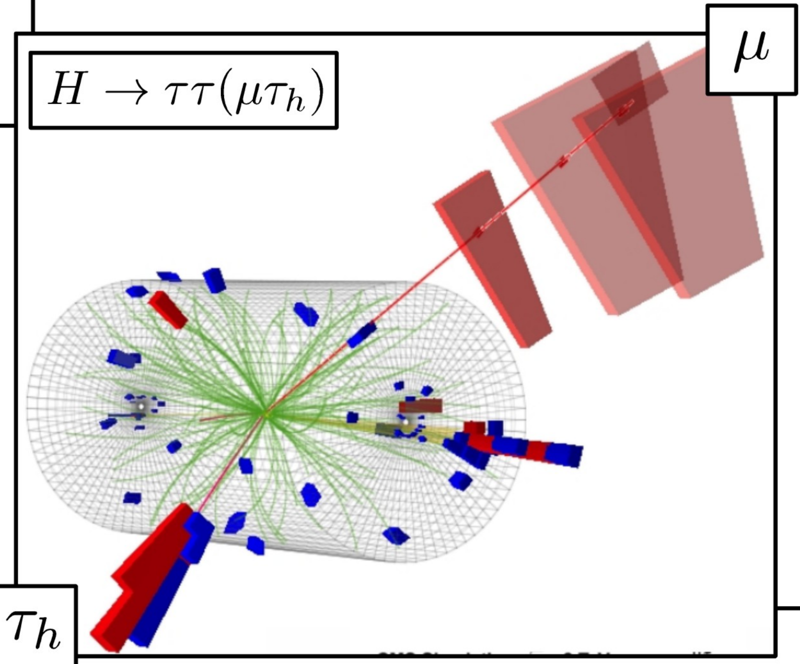
# Di-$\tau$ final state

- High mass allows for **decays into hadrons**:



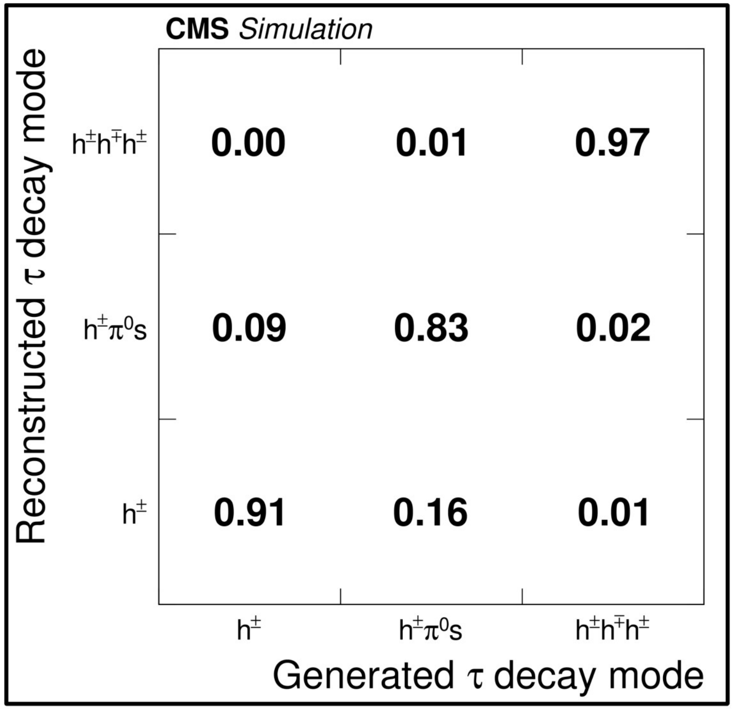

| Decay Mode | BR [%] |
|---|---|
| $e\nu_e\nu_\tau$ | 17.83 |
| $\mu\nu_\mu\nu_\tau$ | 17.41 |
| 1-prong $\nu_\tau$ | 37.10 |
| 3-prong $\nu_\tau$ | 15.20 |

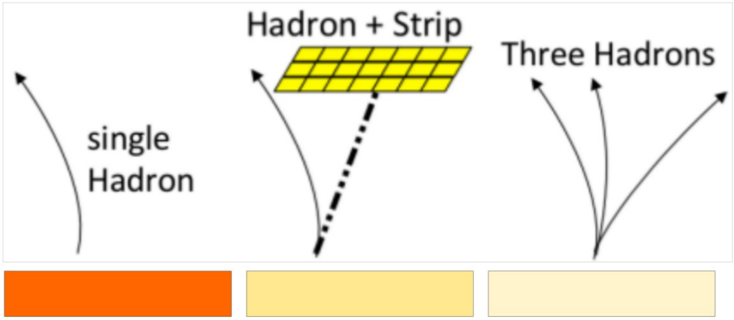
**~90%** of all di-$\tau$ final states contain at least one $\tau_h$.

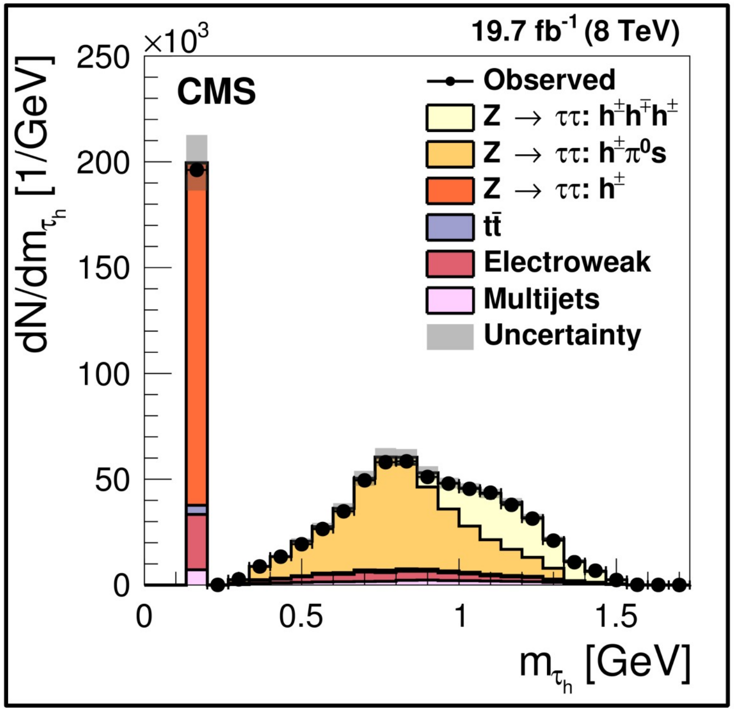


# Hadronic $\tau$-decays

- Start from anti-$k_{\mathrm{T}}$ clustered jets of particle flow objects with opening parameter of 0.4.
- Require **one or three high $p_{\mathrm{T}}$ charged hadrons** ($\rightarrow$ prongs).



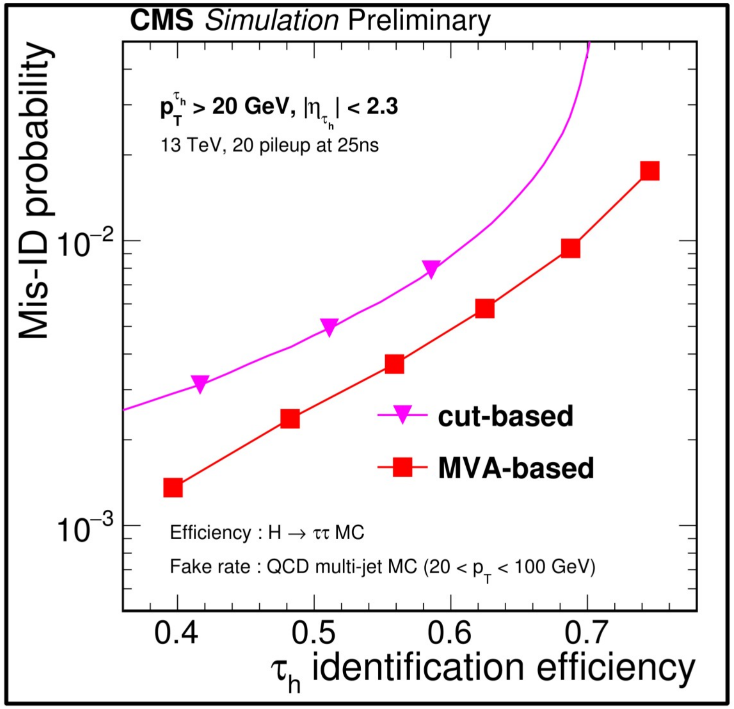

arxiv:1510.07488

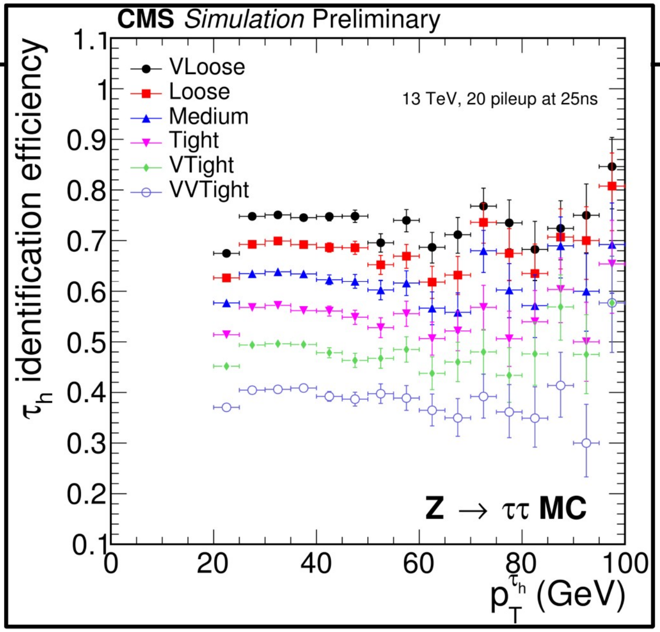
- Apply ID criteria to increase purity.

# $\tau_{\rm h}$ -Identification

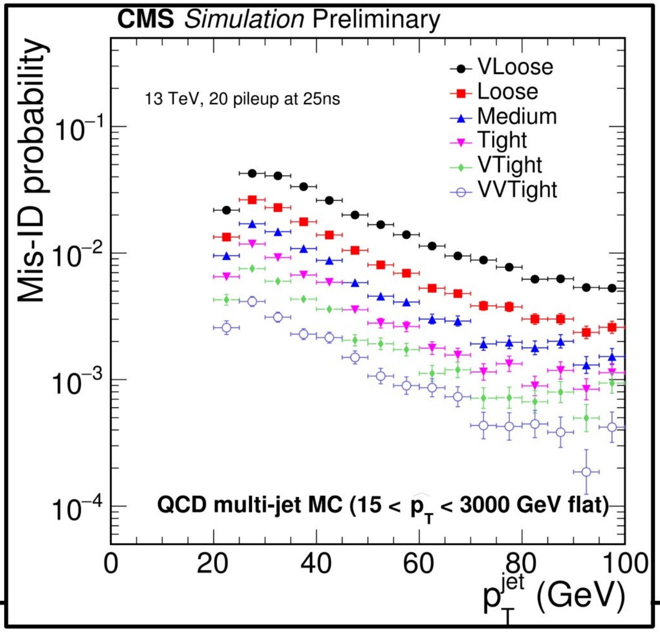
- **MVA based $\tau_h$-identification**: energy deposits close to $\tau$-candidate + impact parameter information on prongs.

- Discrimination against muons and electrons.



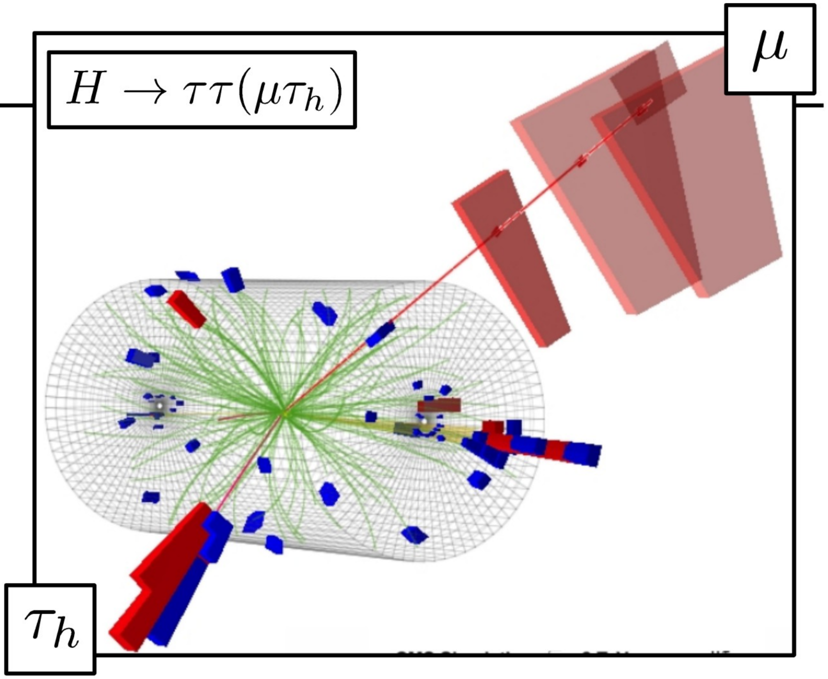

CMS-PAS-TAU-16-002

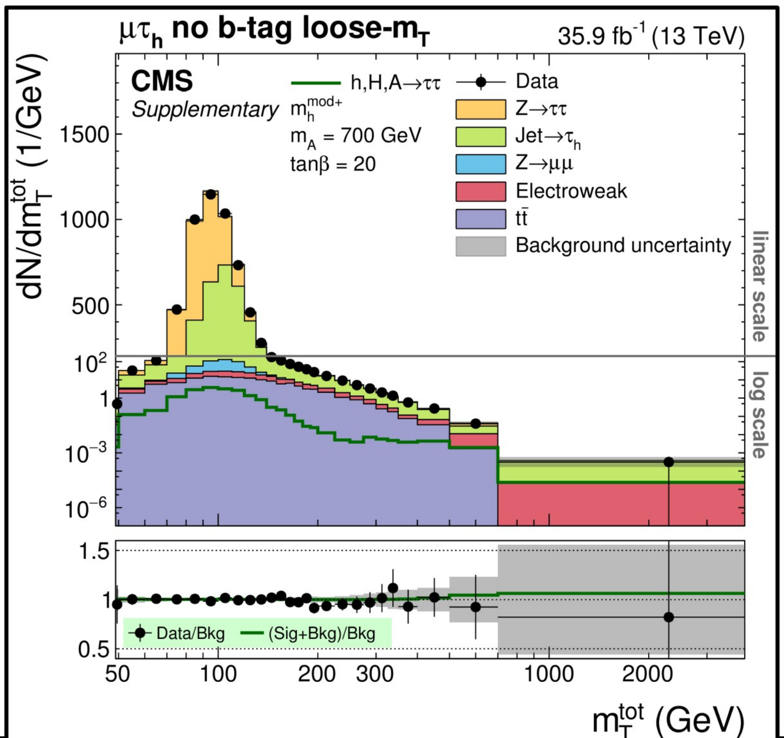



- Predefined working points used in analyses.

# Di-$\tau$ final state

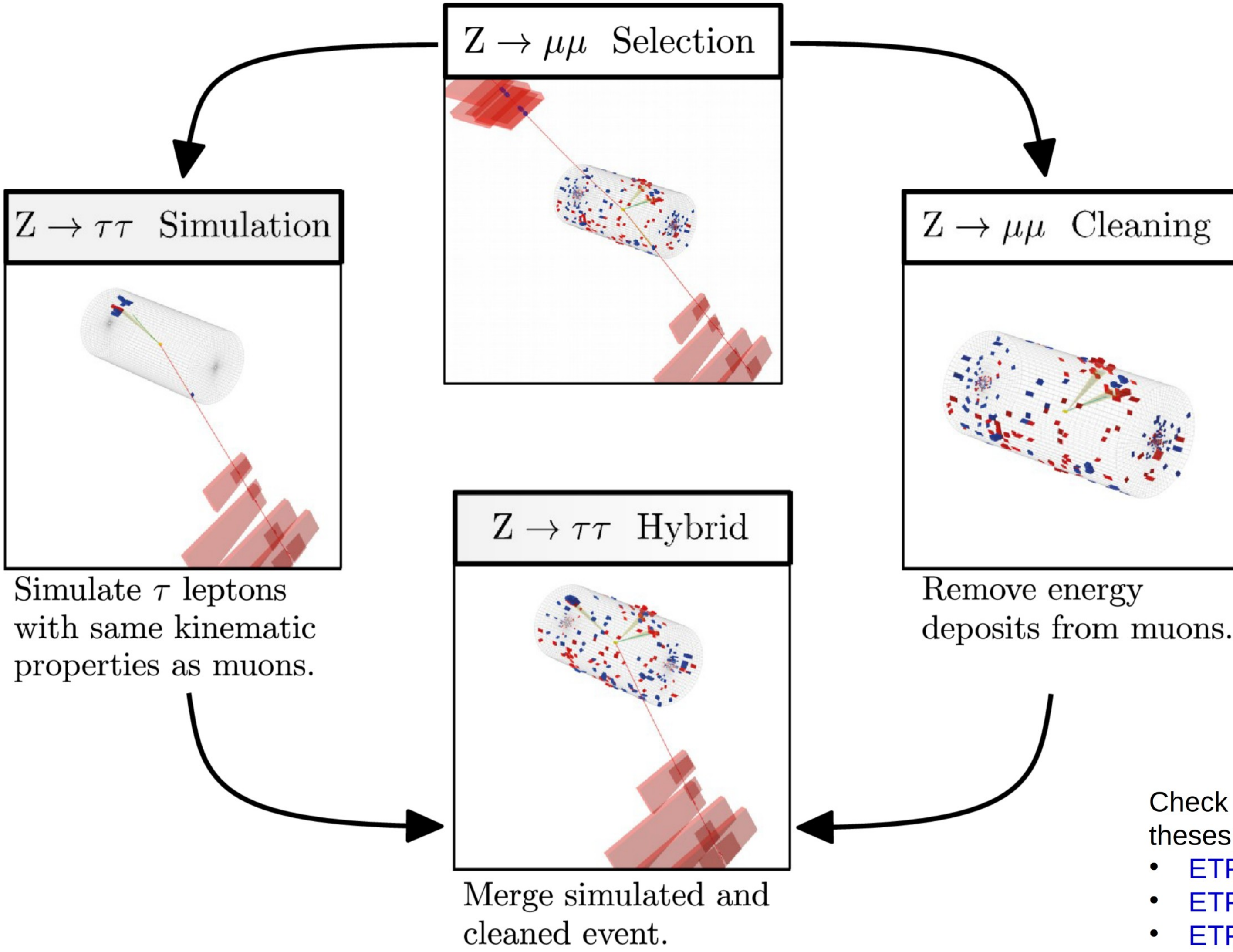
- Search for **2 isolated high $p_T$ leptons** ($e$, $\mu$, $\tau_h$).
- Reduce obvious backgrounds, control what can't be reduced.
- Reconstruct discriminating variable, related to di-$\tau$ final state.



# $\tau$-embedding

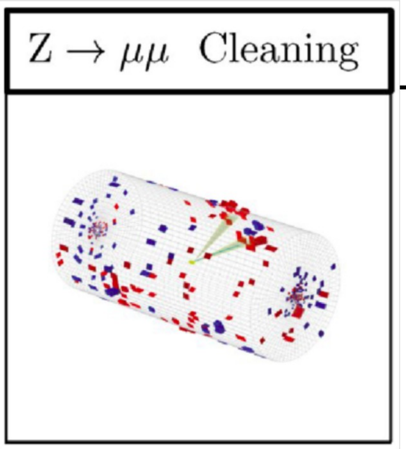
arxiv:1903.01216



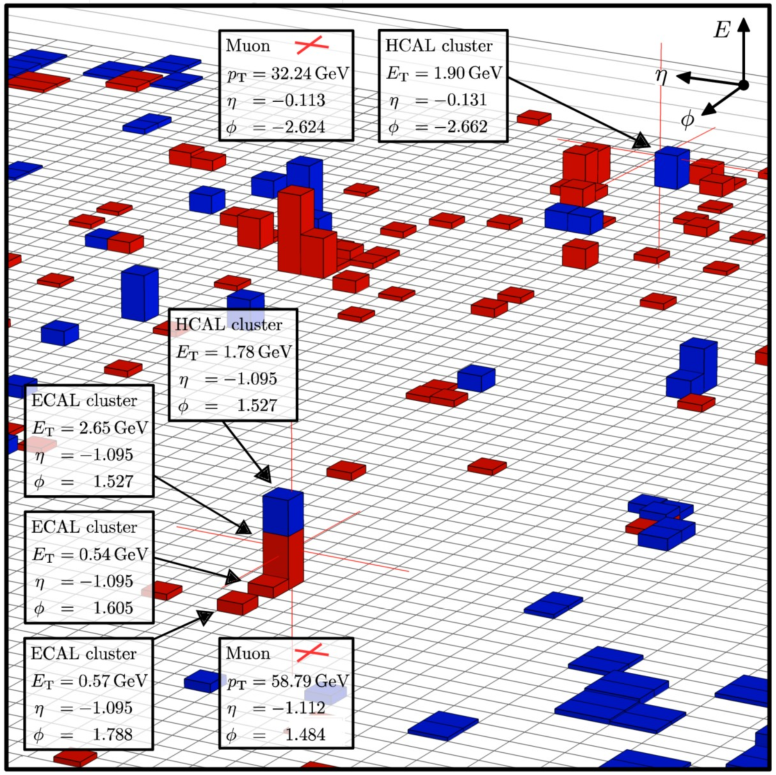

Check the following theses for more details:

- ETP-KA/2016-23
- ETP-KA/2017-31
- ETP-KA/2018-11
- ETP-KA/2019-05

# $\tau$-embedding

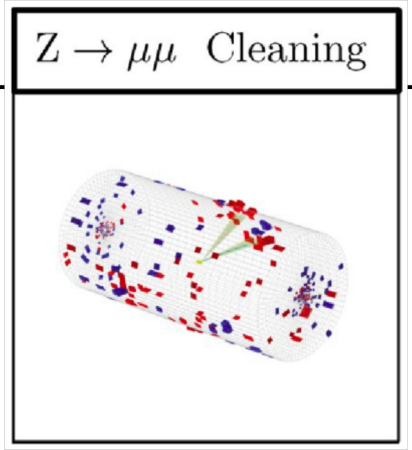



Before:

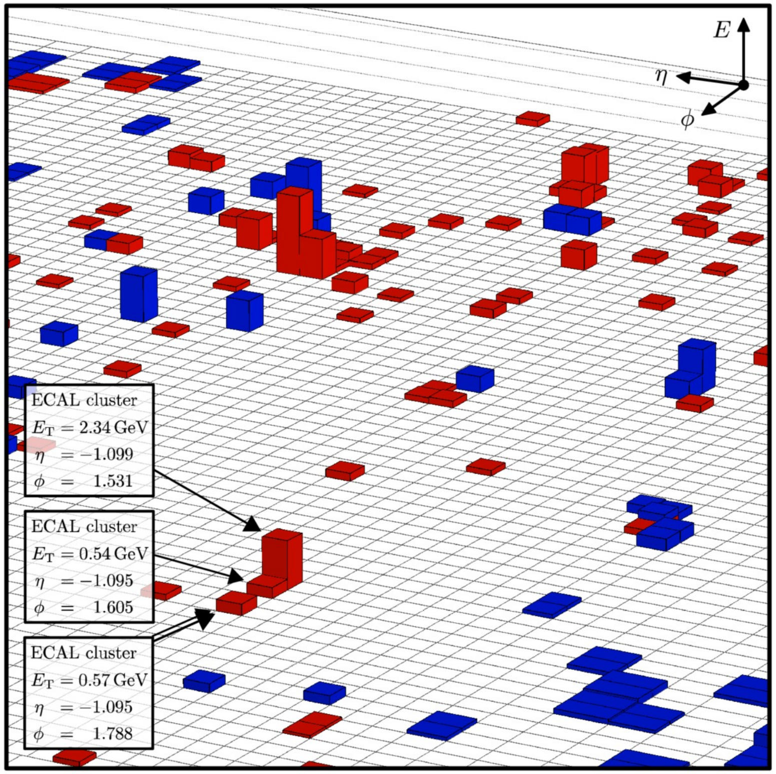



arxiv:1903.01216

# $\tau$-embedding

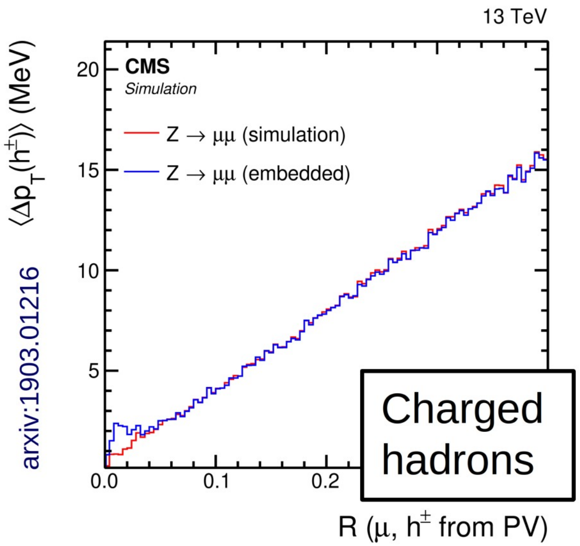


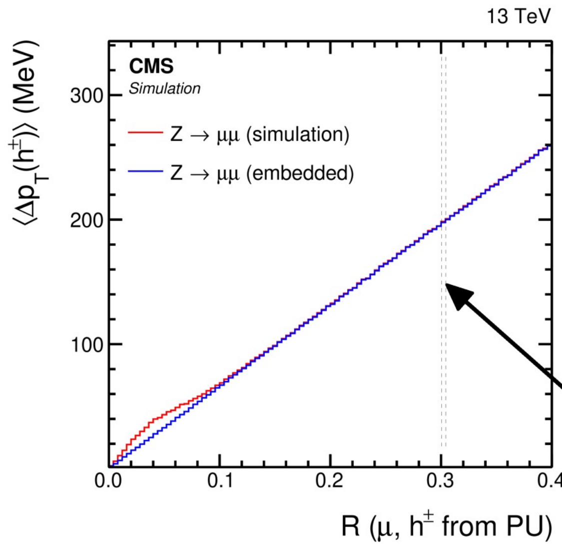

After:



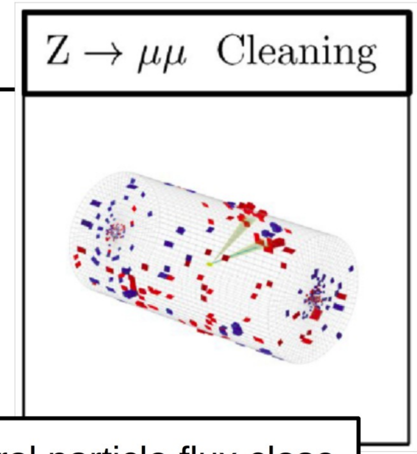

arxiv:1903.01216

# $\tau$-embedding

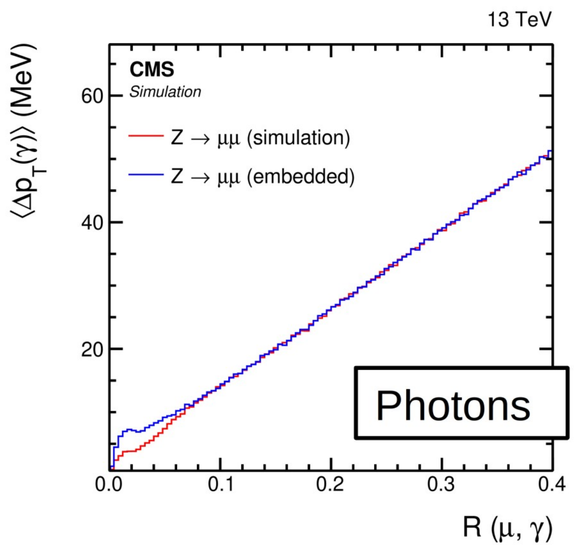


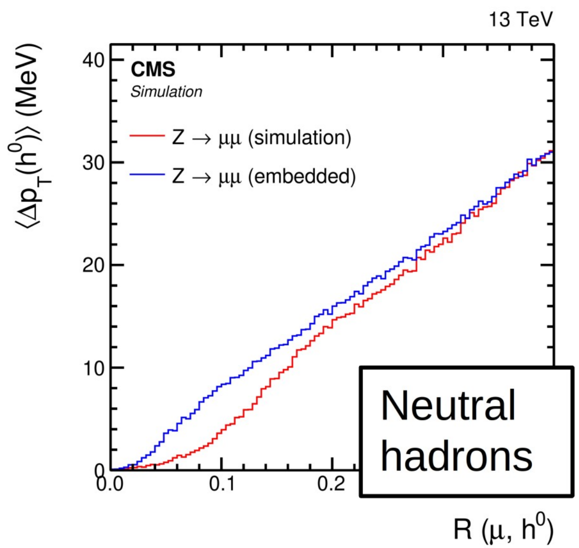



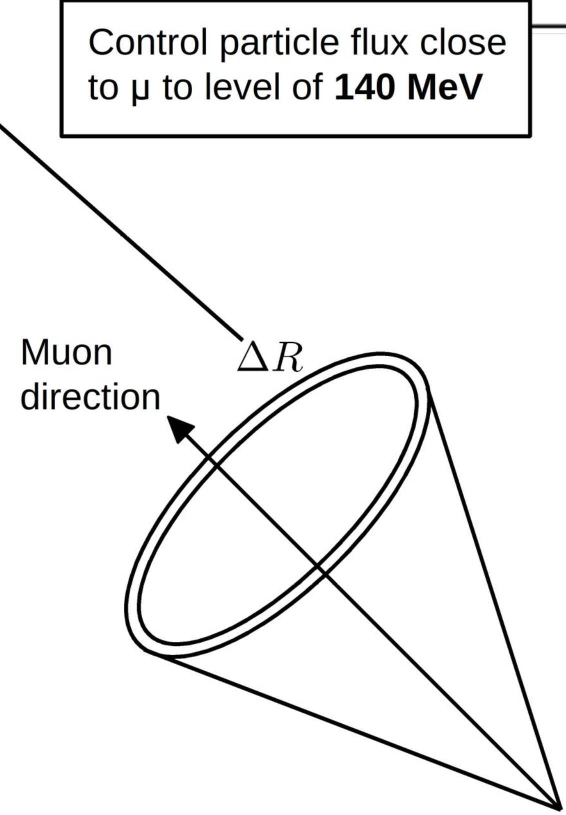

arxiv:1903.01216



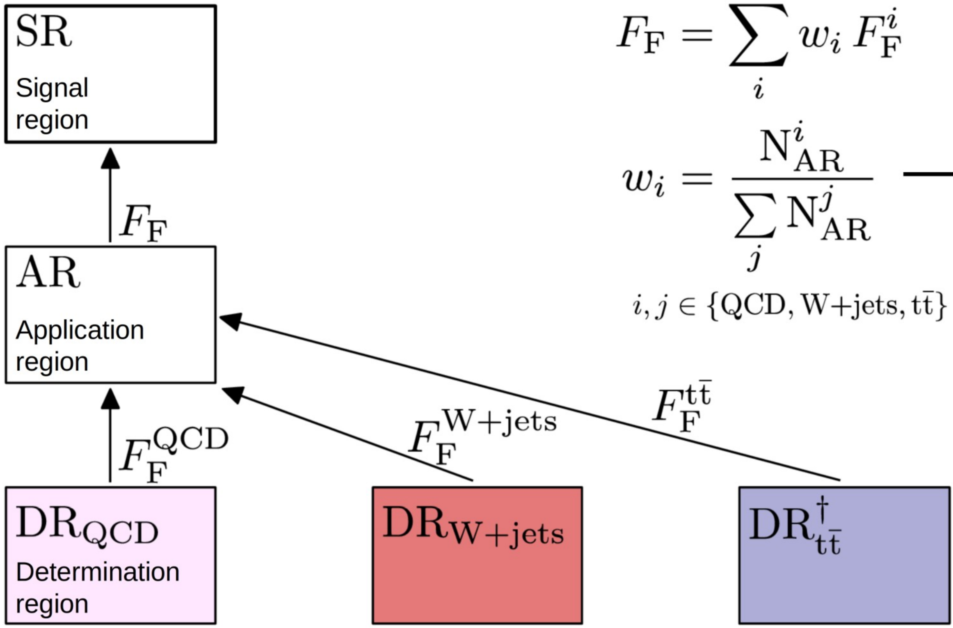

Control particle flux close to μ to level of **140 MeV**



# Fake factor (FF) method

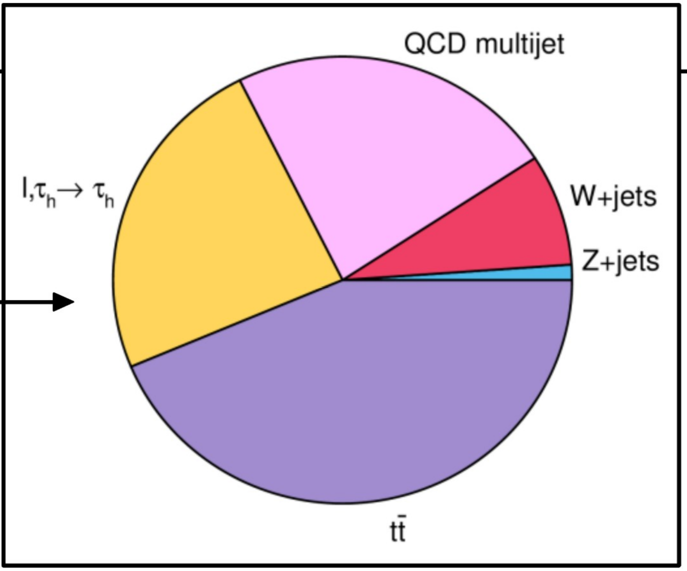


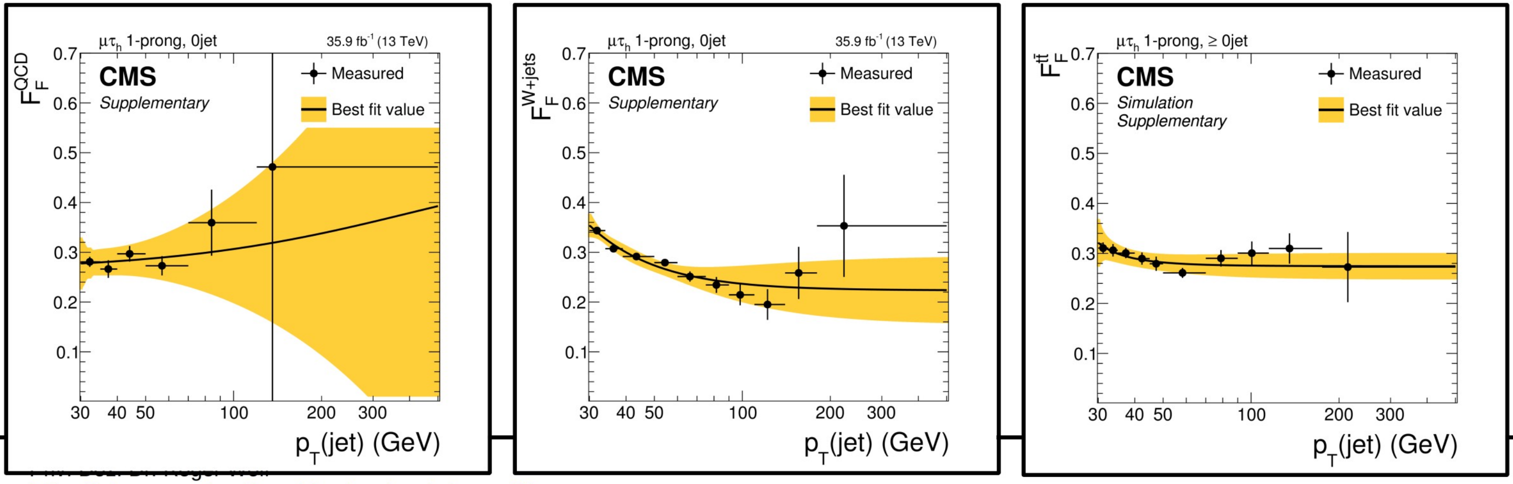

$$F_F = \sum_i w_i \, F_F^i$$

$$w_i = \frac{N^i_{AR}}{\sum_j N^j_{AR}}$$

$$i, j \in \{QCD, W{+}jets, t\bar{t}\}$$

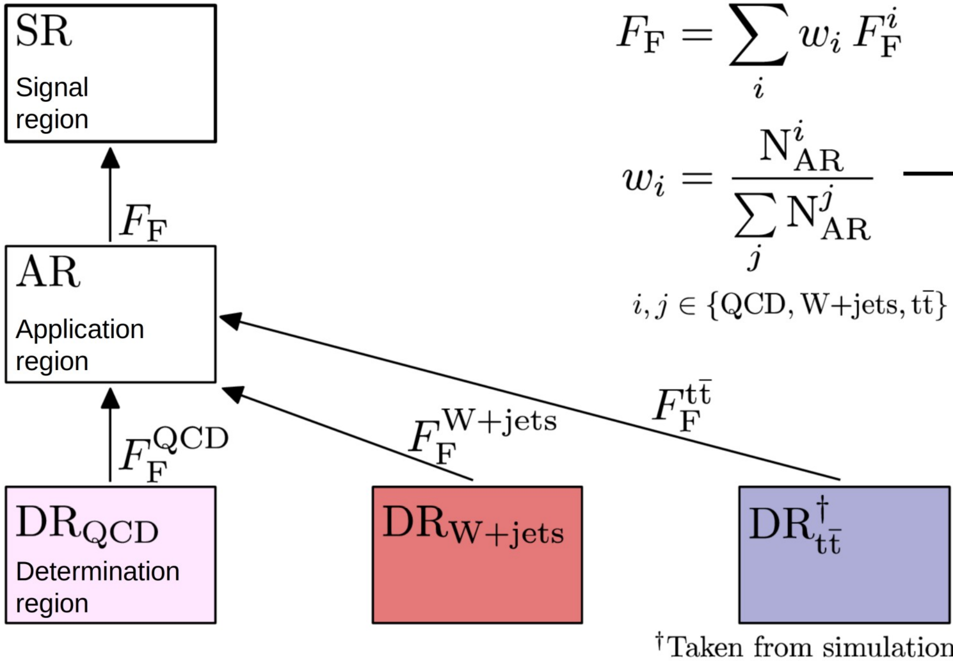
$^{\dagger}$Taken from simulation




JHEP 09 (2018) 007

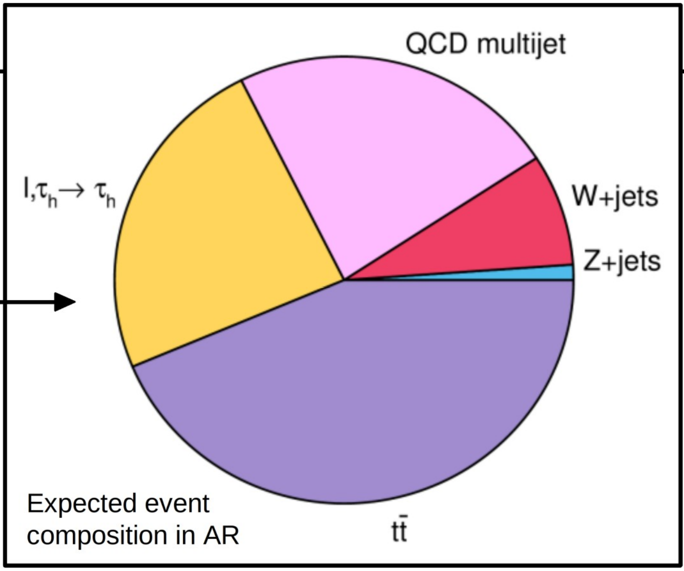
# Fake factor (FF) method

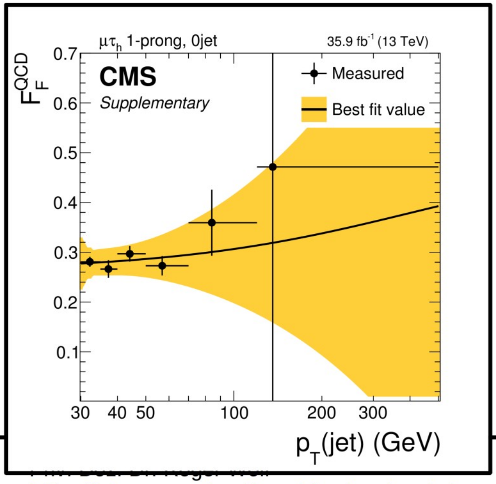


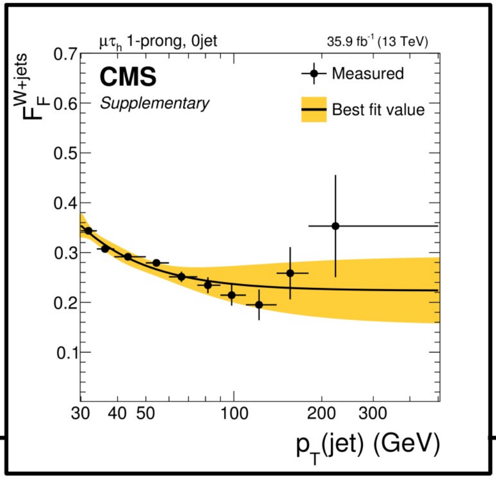

$$F_{\mathrm{F}} = \sum_i w_i \, F_{\mathrm{F}}^{i}$$

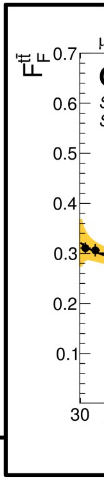
$$w_i = \frac{\mathrm{N}_{\mathrm{AR}}^{i}}{\sum_j \mathrm{N}_{\mathrm{AR}}^{j}}$$

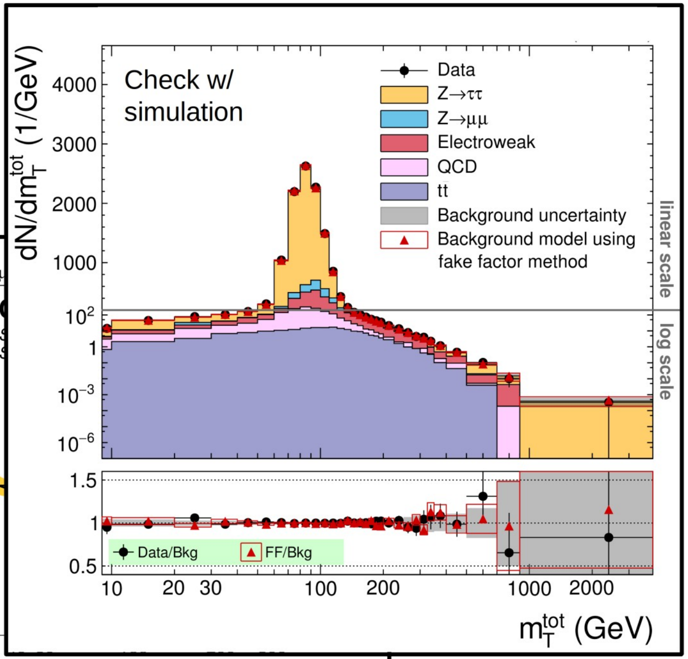
$$i, j \in \{\mathrm{QCD}, \mathrm{W+jets}, t\bar{t}\}$$



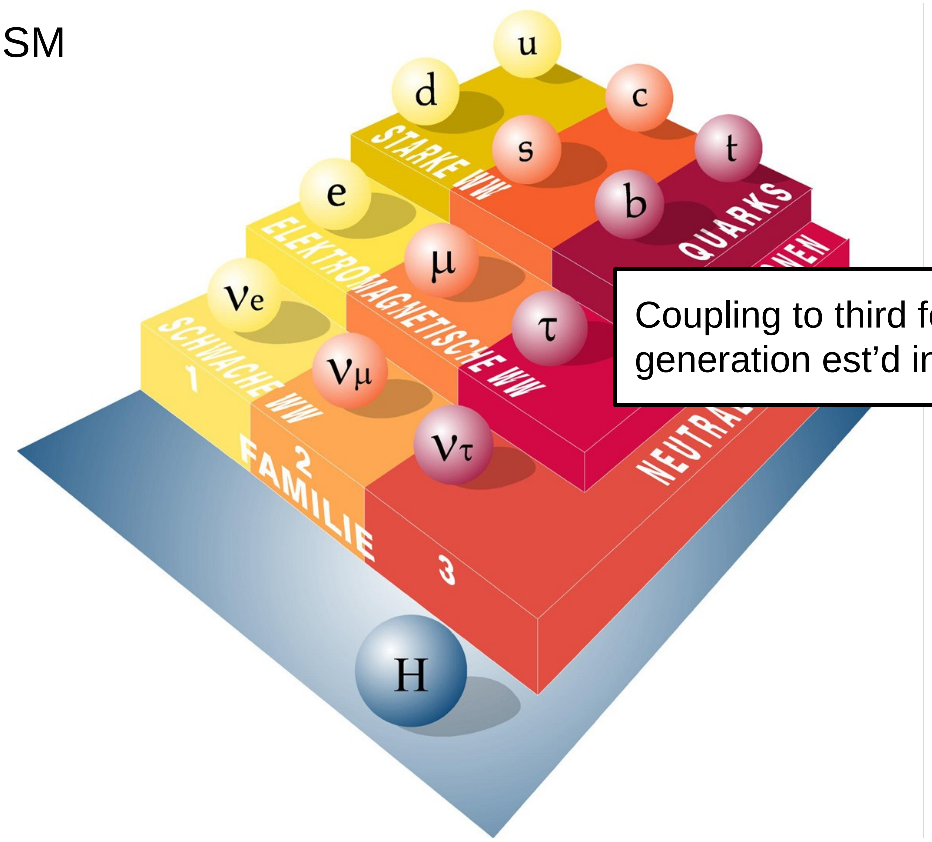



# SM $\mathrm{H} \to \tau\tau$ analysis

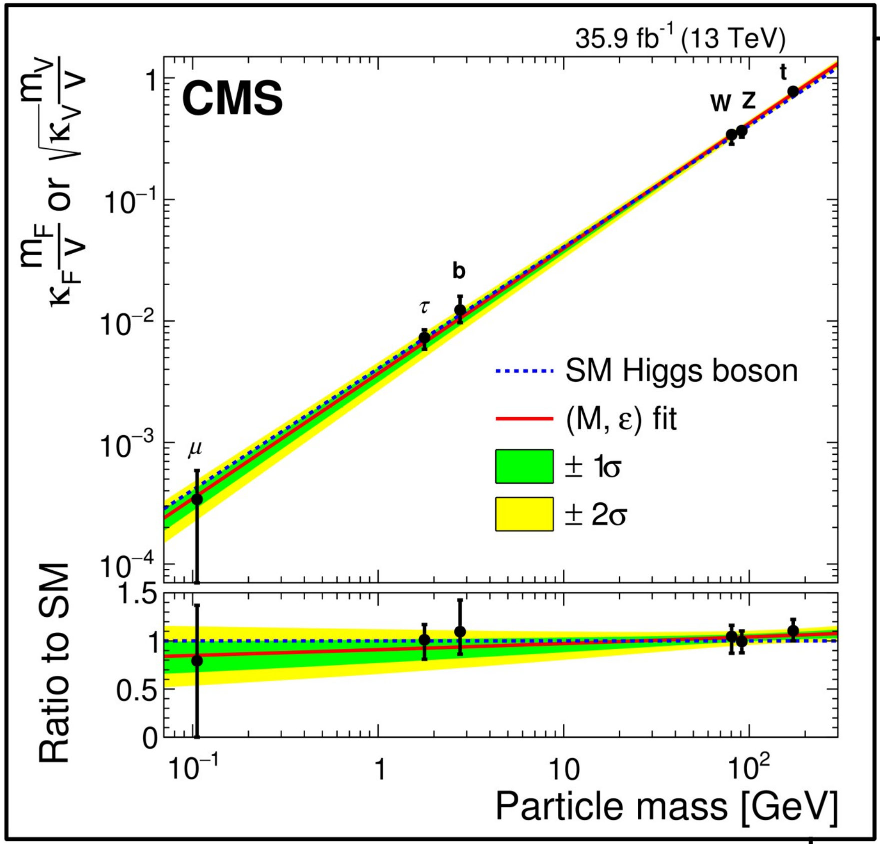
- Undoubted that what we observe at 125 GeV is a Higgs boson.
- Measurement scope:
  - Investigate **coupling** structure.
  - Check for **deviations** from the SM expectation.

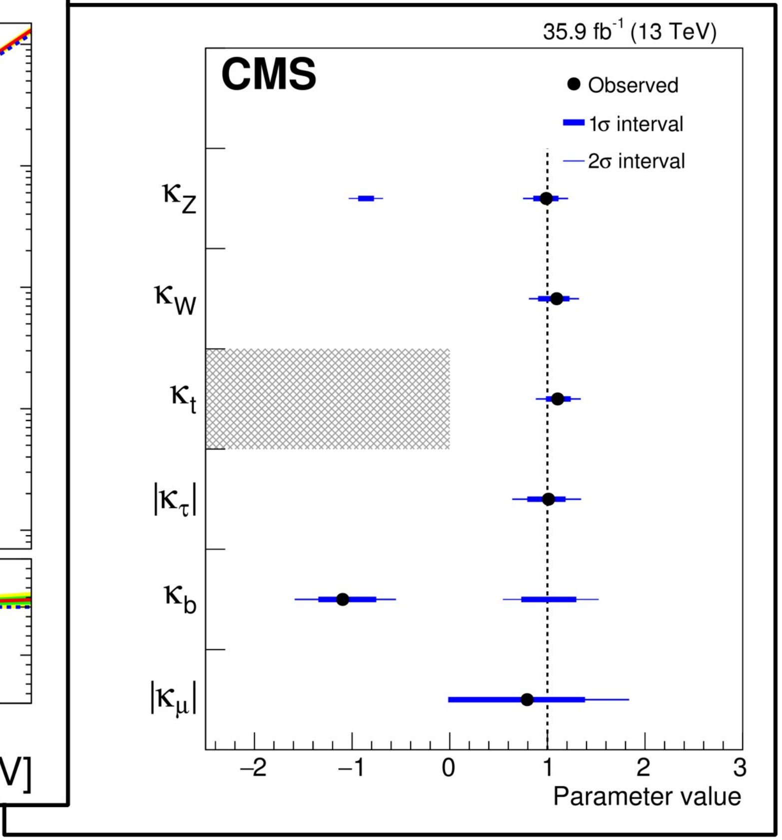



Coupling to third fermion generation est'd in 2018.

# Higgs coupling structure

- Part of classic analysis of **rate measurements** in production modes & final states.




arxiv:1809.10733

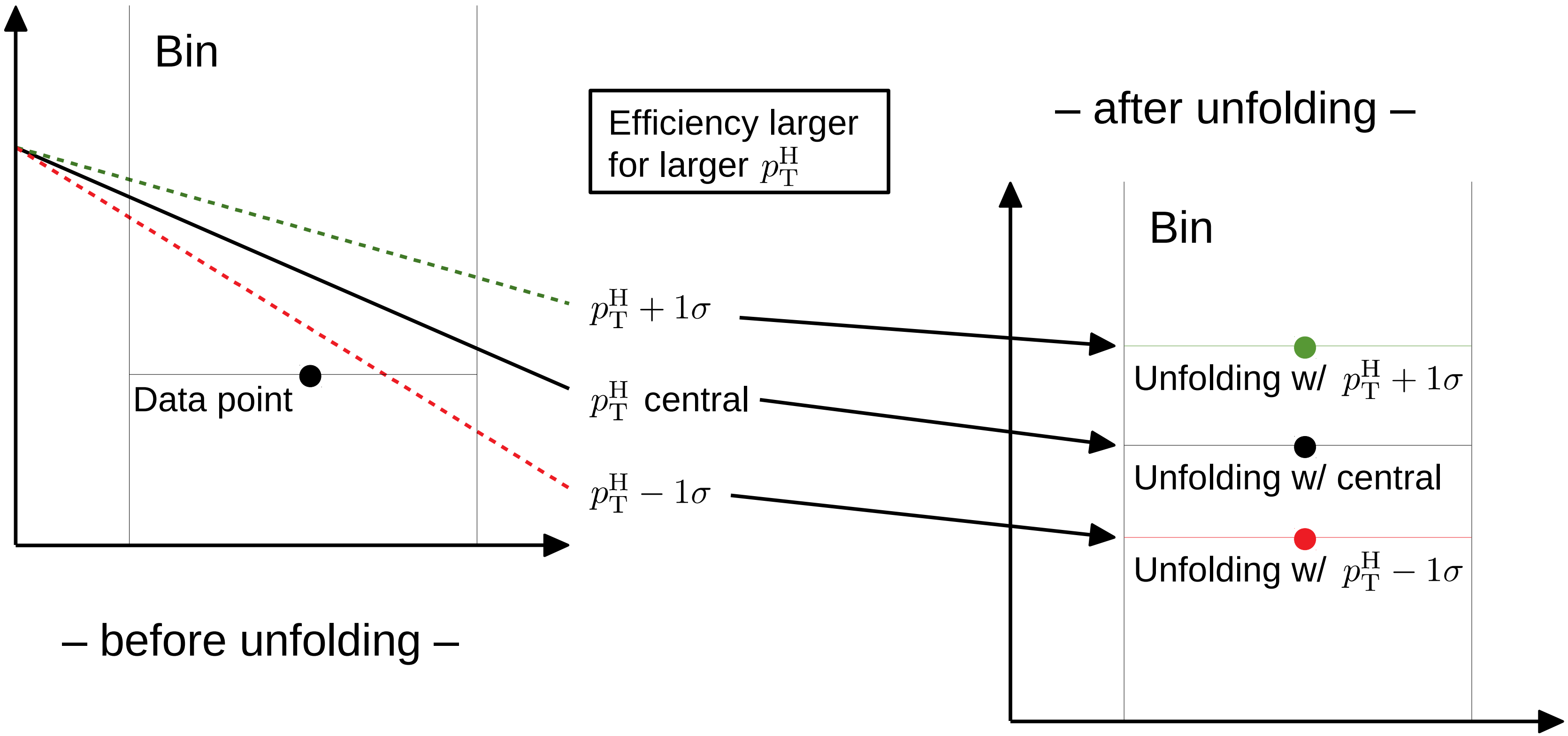
**BUT**: new physics has influence on kinematic distributions.

# Simplified template cross section (STXS)

- Define common phasespace regions based on pseudo-observable objects and quantities:
  - Convention to allow for **combination** of final states and across experiments.
  - Kinematic bins help to reduce **influence of theory uncertainties** (e.g. in $p_{\mathrm{T}}^{\mathrm{H}}$ or $N_{\mathrm{Jet}}$) on measurement.

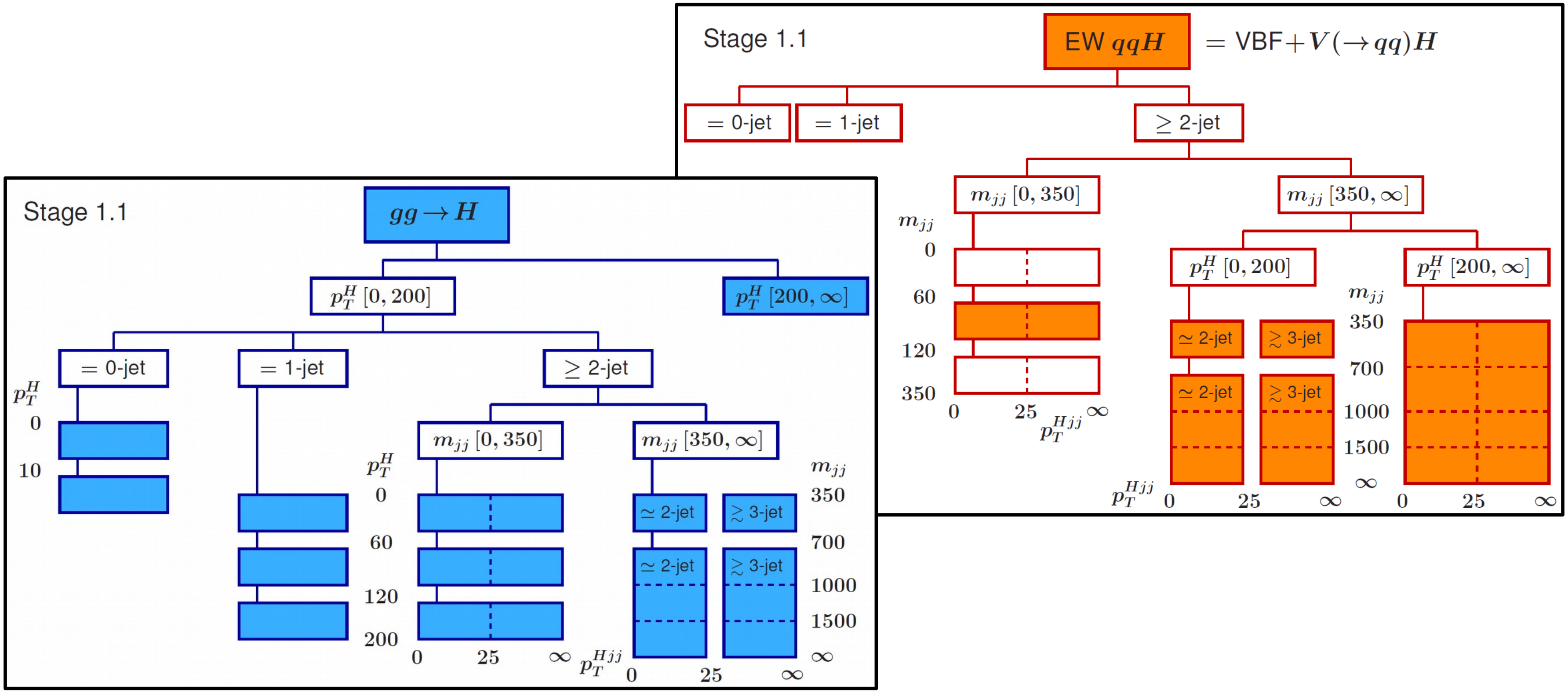
# Simplified template cross section (STXS)

- Define common phasespace regions based on pseudo-observable objects and quantities:
  - Convention to allow for **combination** of final states and across experiments.
  - Kinematic bins help to reduce **influence of theory uncertainties** (e.g. in $p_{\mathrm{T}}^{\mathrm{H}}$ or $N_{\mathrm{Jet}}$) on measurement.



# Simplified template cross section (STXS)

- Defined for analysis of **LHC Run-2** data by LHC HXSWG:



# Template vs. fiducial cross section

**Simplified template cross section (STXS):**

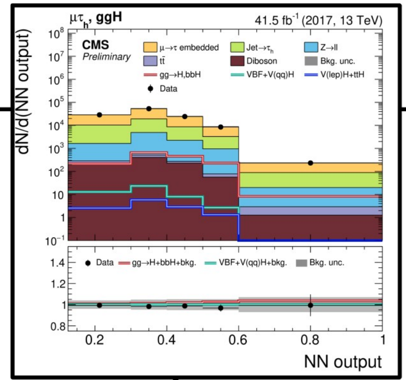
- E.g. Higgs production in VBF & gluon fusion.

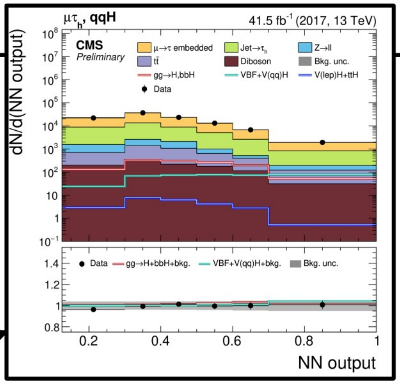
↔

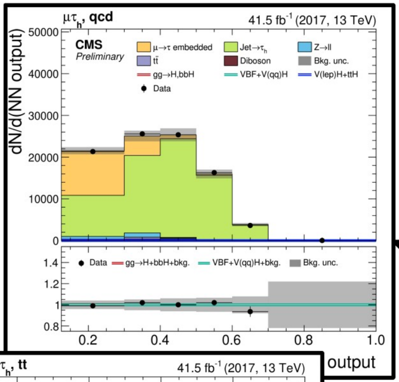
**Fiducial cross section:**

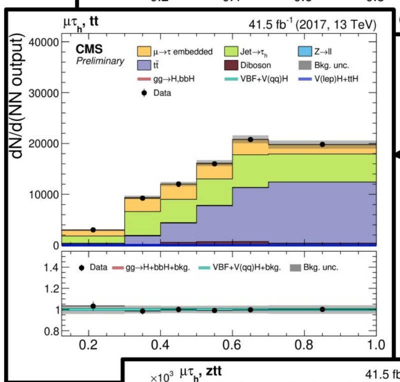
- Obey detector acceptance and stick to measurable quantities.
- E.g. Higgs production in association w/ two jets w/ $m_{\mathrm{jj}} > 350\ \mathrm{GeV}$.

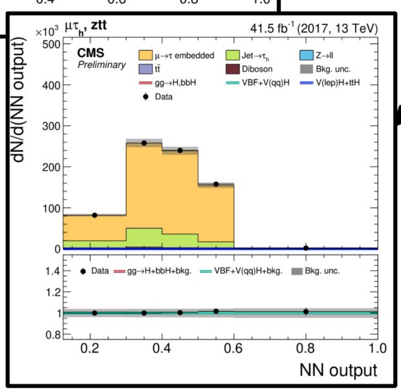
# Event classification




qcd 0.02

ggH 0.15

qqH 0.08

tt 0.05

ztt 0.70

...

Event in data

Modern multi-class NN

CMS-PAS-HIG-18-032

Check ETP-KA/2017-26 for more details.

- Trained to differentiate btw. all signal & background processes.
- Output → tuple of scores (~Bayesian probabilities) for data event to belong to a given process.
- Highest score defines the event category that the event is associated to.

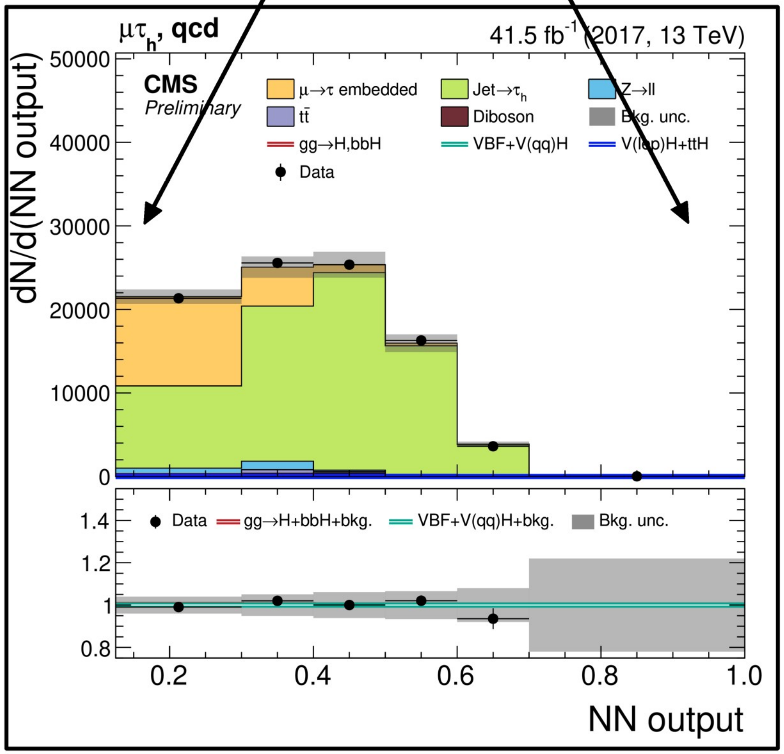
# Signal extraction

- Signal derived from **maximum likelihood fit** to NN output of each event category.
- Pure background categories help to constrain backgrounds in signal categories.

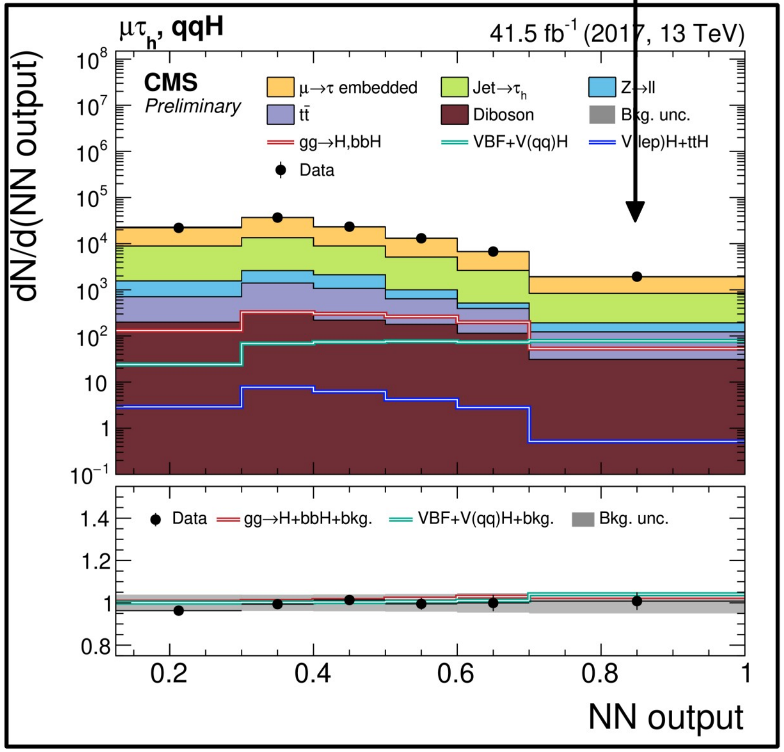
Within plot transition btw.
*ambiguous* & *clear* signature

Find most-signal
like events here

**qcd**




**qqH**




CMS-PAS-HIG-18-032

# NN inputs

- Use one NN for each final state and separated btw. 2016 & 2017 (→ **8 NNs**):

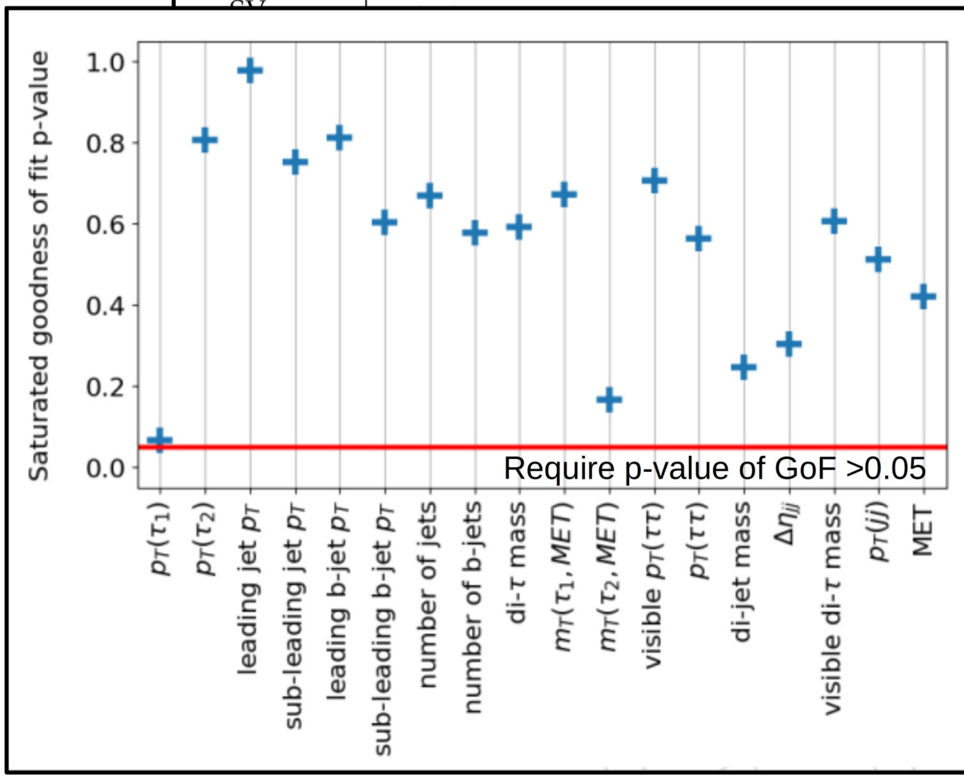
| Variable | $e\mu$ | $e\tau_h$ | $\mu\tau_h$ | $\tau_h\tau_h$ | Variable | $e\mu$ | $e\tau_h$ | $\mu\tau_h$ | $\tau_h\tau_h$ |
|---|---|---|---|---|---|---|---|---|---|
| $m_{\tau\tau}^{SV}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ | $p_T^{jj}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ |
| $m_{T\tau\tau}^{SV}$ | ✓✓ | —— | —— | —— | $p_T(\text{b jet}_1)$ | —— | ✓✓ | ✓✓ | ✓✓ |
| $p_{T\tau\tau}^{SV}$ | ✓✓ | —— | —— | —— | $p_T(\text{b jet}_2)$ | —— | ✓✓ | ✓✓ | ✓✓ |
| $m_{vis}$ | ✓— | ✓— | ✓— | ✓✓ | $p_T^{miss}$ | —✓ | ✓✓ | ✓— | ✓— |
| $p_T^{vis}$ | ✓✓ | ✓✓ | ✓— | ✓— | $D_\zeta$ | ✓✓ | —— | —— | —— |
| $p_T^{\tau_1}$ | —— | —— | ✓— | ✓✓ | $m_T^{e}$ | —— | ✓✓ | —— | —— |
| $p_T^{\tau_2}$ | ✓— | ✓✓ | ✓✓ | ✓— | $m_T^{\mu}$ | ✓✓ | —— | ✓✓ | —— |
| $\Delta R^{e\mu}$ | ✓✓ | —— | —— | —— | $m_T^{e+\mu}$ | ✓— | —— | —— | —— |
| $p_T(\text{jet}_1)$ | ✓✓ | ✓✓ | ✓✓ | ✓— | $\max(m_T^{\mu}, m_T^{e})$ | ✓✓ | —— | —— | —— |
| $\eta\ (\text{jet}_1)$ | ✓— | —— | —— | —— | $m_T^{\tau_h}$ | —— | ✓✓ | ✓— | ✓✓ |
| $p_T(\text{jet}_2)$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ | $p_T^{\tau\tau+miss}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ |
| $\eta\ (\text{jet}_2)$ | ✓— | —— | —— | —— | $p_T^{\tau\tau jj+miss}$ | —✓ | —— | —— | —— |
| $m_{jj}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ | $N_{\text{b jet}}$ | —— | ✓✓ | ✓✓ | ✓✓ |
| $\Delta\eta_{jj}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ | $N_{jet}$ | ✓✓ | ✓✓ | ✓✓ | —✓ |
| | | | | | Variables in use: | 21(19) | 18(13) | 17(16) | 16(12) |

# NN inputs

- Use one NN for each final state and seperated btw. 2016 & 2017 (→ **8 NNs**):

| Variable | eμ | e$\tau_h$ | μ$\tau_h$ | $\tau_h\tau_h$ |
|---|---|---|---|---|
| $m_{\tau\tau}^{SV}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ |

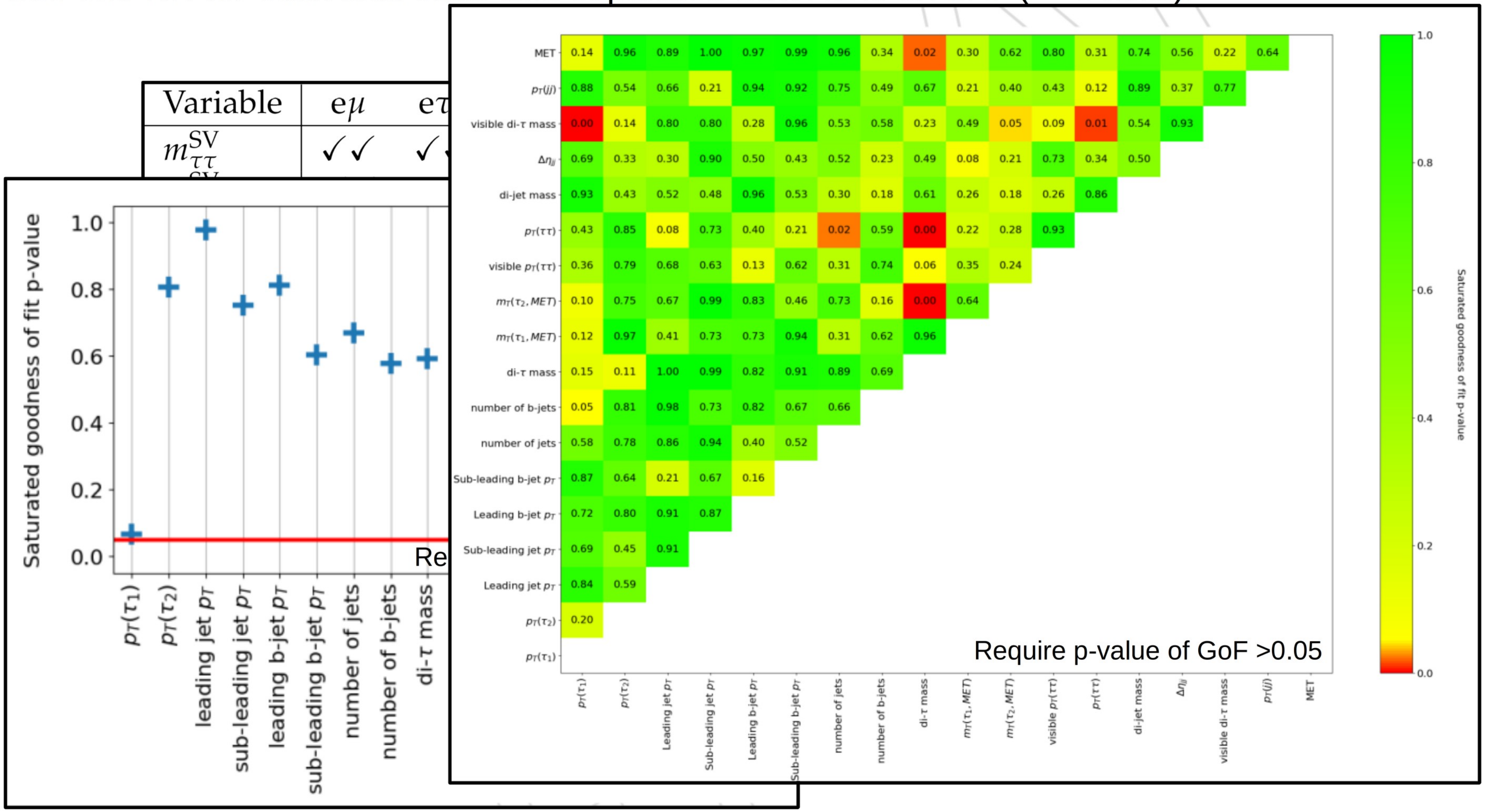
| Variable | eμ | e$\tau_h$ | μ$\tau_h$ | $\tau_h\tau_h$ |
|---|---|---|---|---|
| $p_T^{jj}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ |
| $p_T(\text{b jet}_1)$ | –– | ✓✓ | ✓✓ | ✓✓ |
| $p_T(\text{b jet}_2)$ | –– | ✓✓ | ✓✓ | ✓✓ |
| $p_T^{miss}$ | –✓ | ✓✓ | ✓– | ✓– |
| $D_\zeta$ | ✓✓ | –– | –– | –– |
| $m_T^e$ | –– | ✓✓ | –– | –– |
| $m_T^\mu$ | ✓✓ | –– | ✓✓ | –– |
| $m_T^{e+\mu}$ | ✓– | –– | –– | –– |
| $\max(m_T^\mu, m_T^e)$ | ✓✓ | –– | –– | –– |
| $m_T^{\tau_h}$ | –– | ✓✓ | ✓– | ✓✓ |
| $p_T^{\tau\tau+miss}$ | ✓✓ | ✓✓ | ✓✓ | ✓✓ |
| $p_T^{\tau\tau jj+miss}$ | –✓ | –– | –– | –– |
| $N_{\text{b jet}}$ | –– | ✓✓ | ✓✓ | ✓✓ |
| $N_{\text{jet}}$ | ✓✓ | ✓✓ | ✓✓ | –✓ |
| Variables in use: | 21(19) | 18(13) | 17(16) | 16(12) |




Making sure that input variables are well described by our model exploiting goodness-of-fit (GoF) test in 1d…

# NN inputs

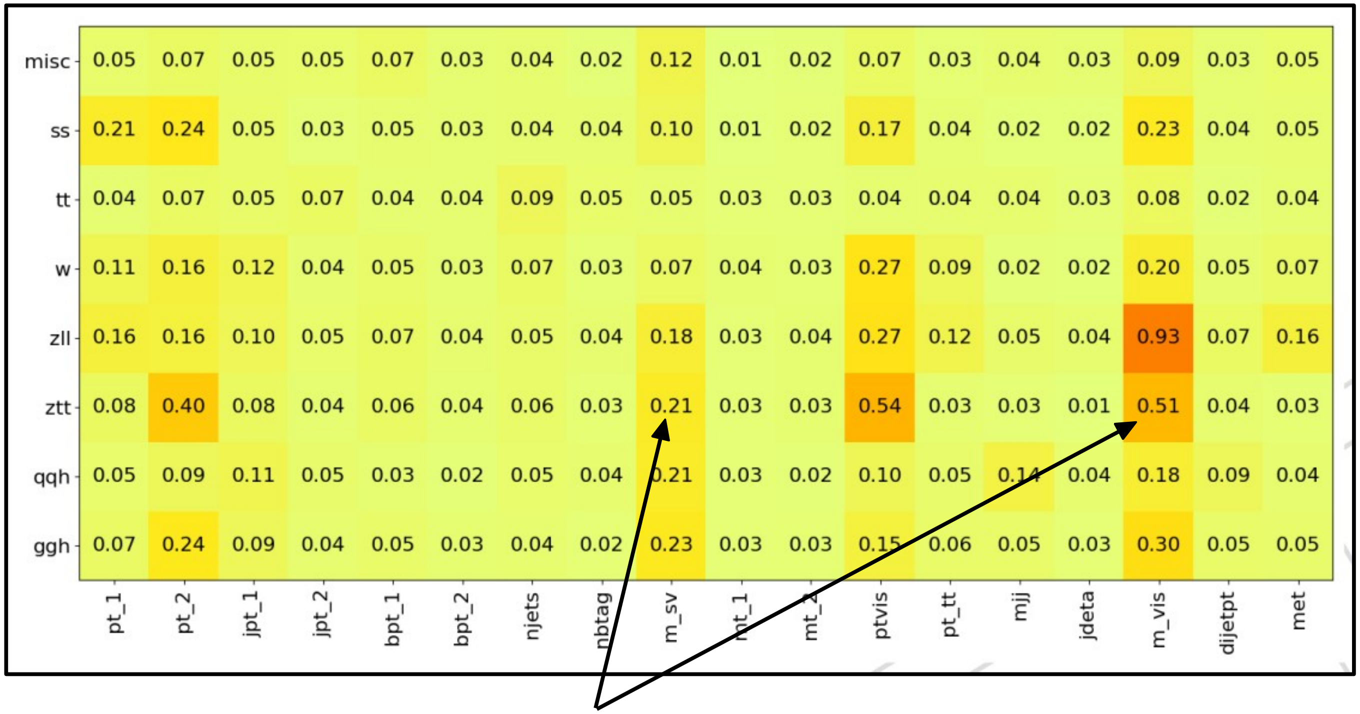
- Use one NN for each final state and seperated btw. 2016 & 2017 (→ **8 NNs**):




Making sure that input variables are well described by our model exploiting goodness-of-fit (GoF) test in 1d… & 2d.

# “Unboxing” the NN

- Decipher what the NN is doing using a Taylor expansion of the full NN output function. **Impact analysis** like on LEP likelihood, but here on NN output function.



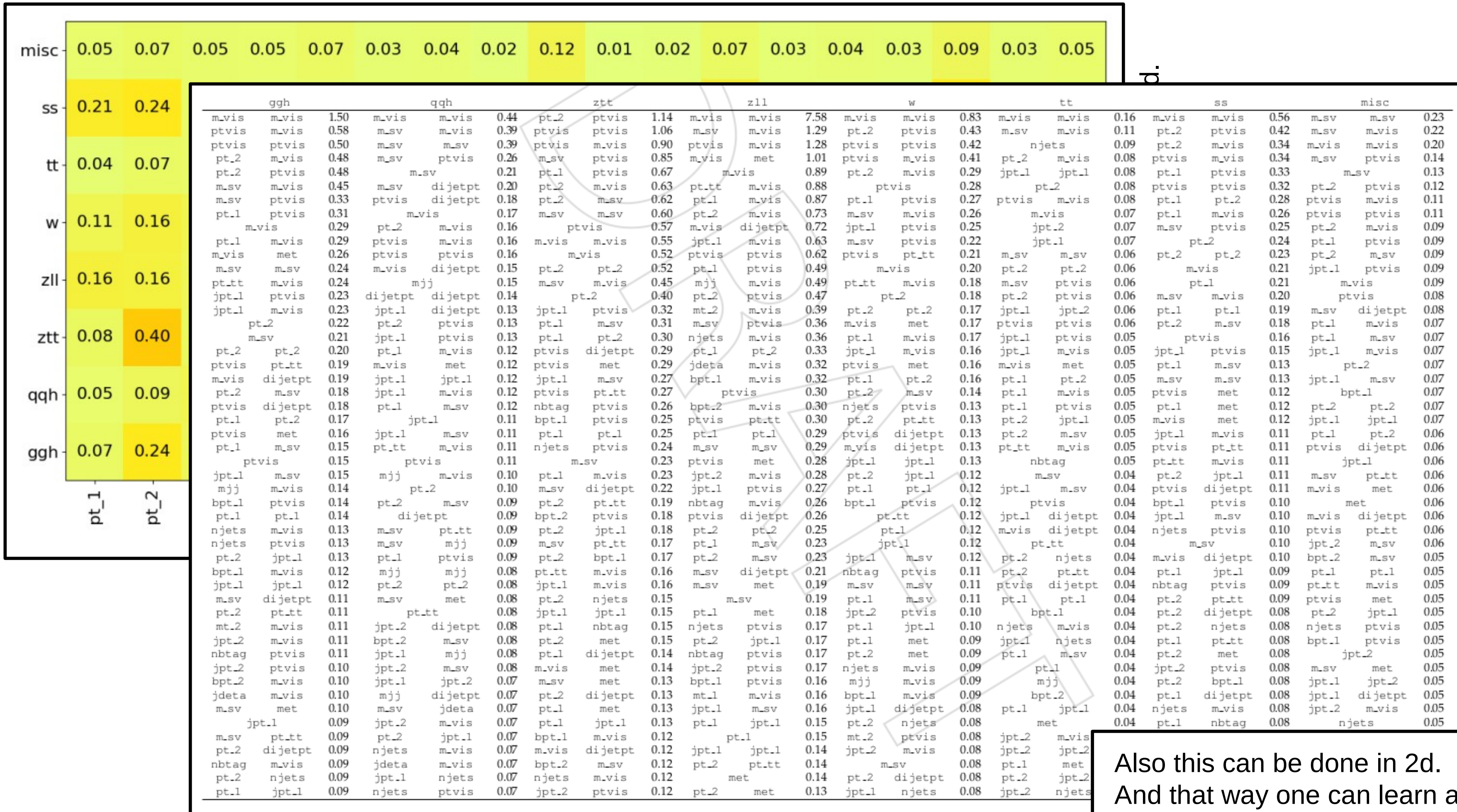

Note that all values >0 are allowed.

Relative size of number indicates how sensitive the NN output is on the given input.

Comput. Softw. Big Sci. 2 (2018) 5

# “Unboxing” the NN

- Decipher what the NN is doing using a Taylor expansion of the full NN output function. **Impact analysis** like on LEP likelihood, but here on NN output function.

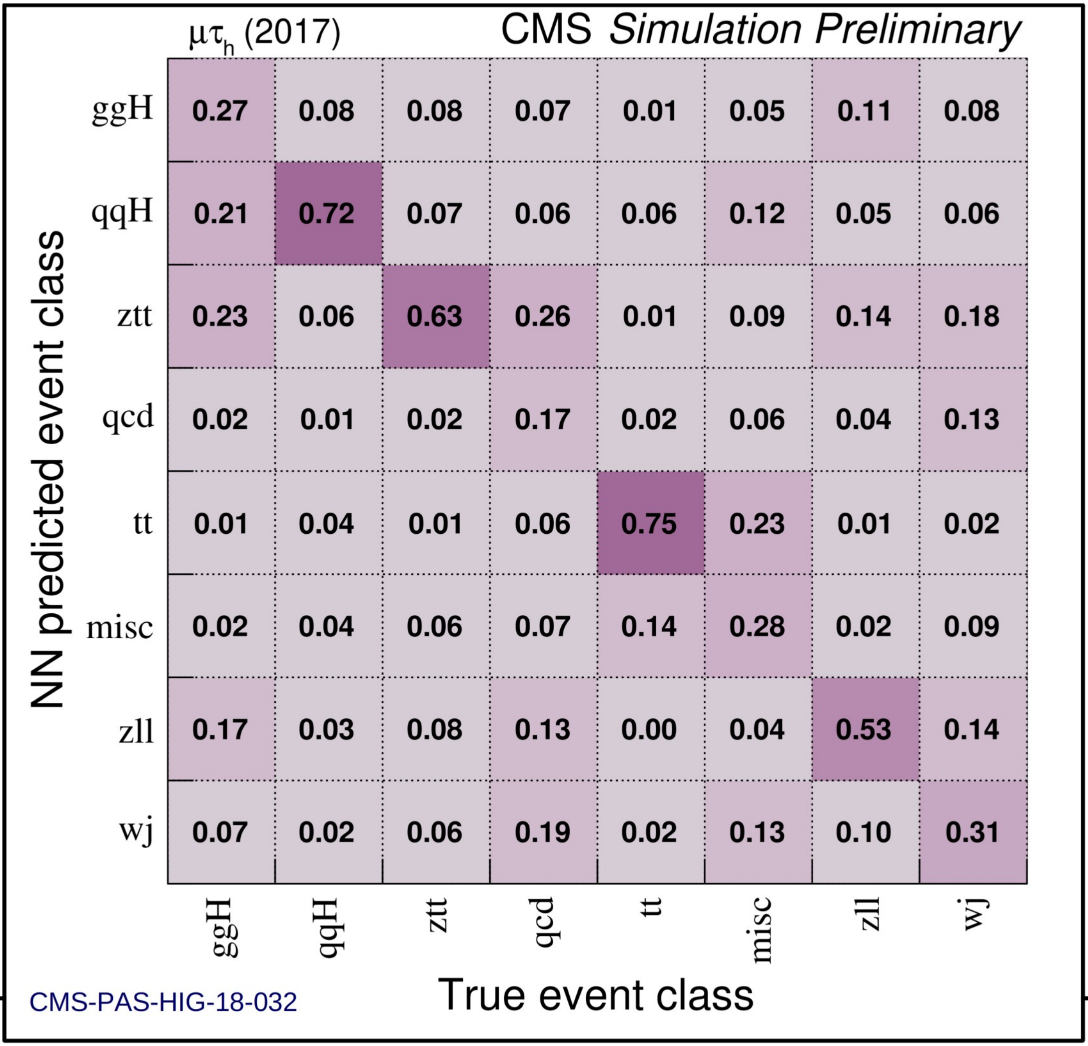



Also this can be done in 2d. And that way one can learn a lot about the NN task and how it is solved.

# How well can the NN do?

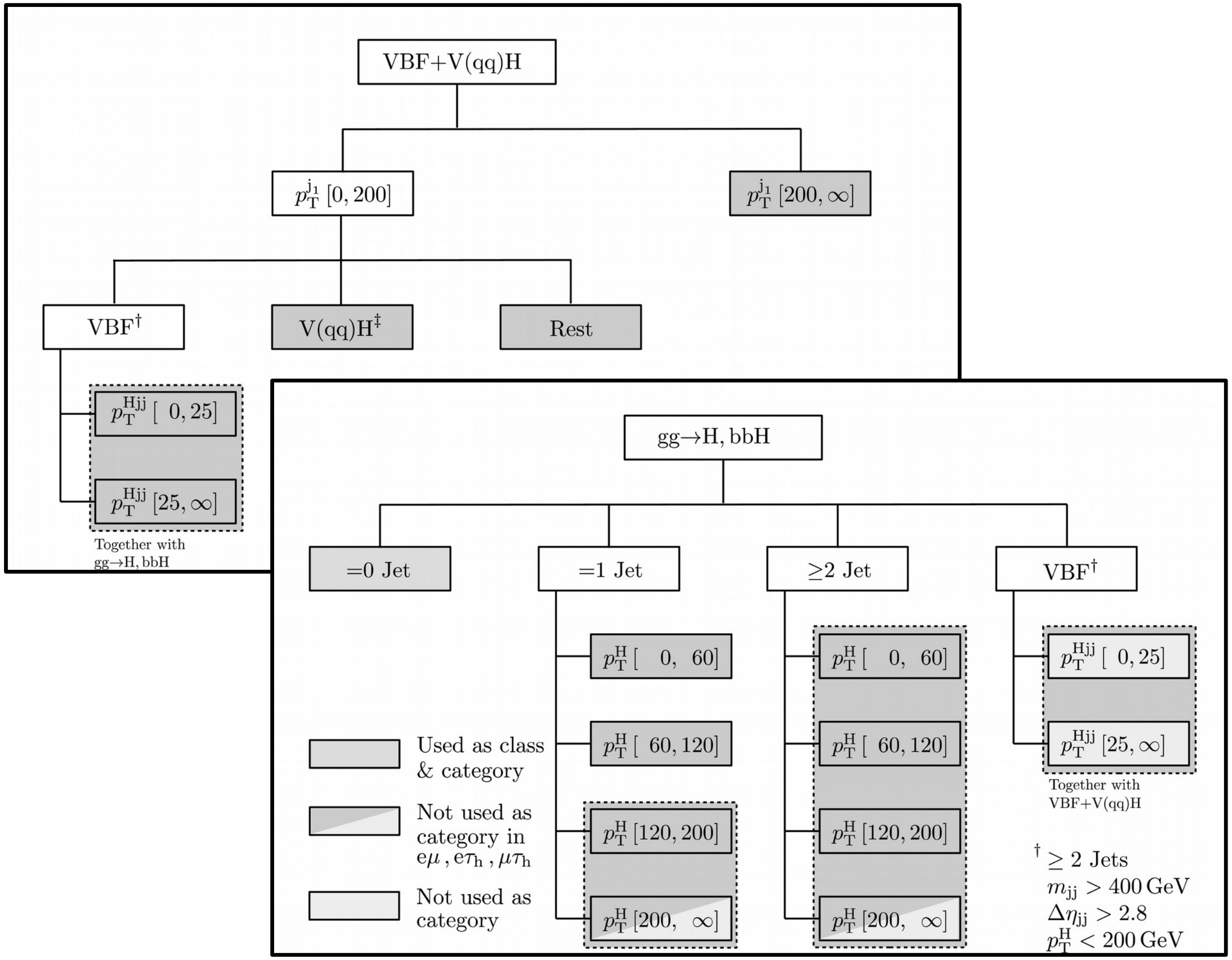
- Confusion matrix tells how well the NN can **identify each individual process**:

- In this representation: all columns normalized to unity.
- 72% of all **`qqH`** events can be identified as such.
- Assess success of NN by comparison to random association (prob. 1/8=12.5%).



# STXS classification

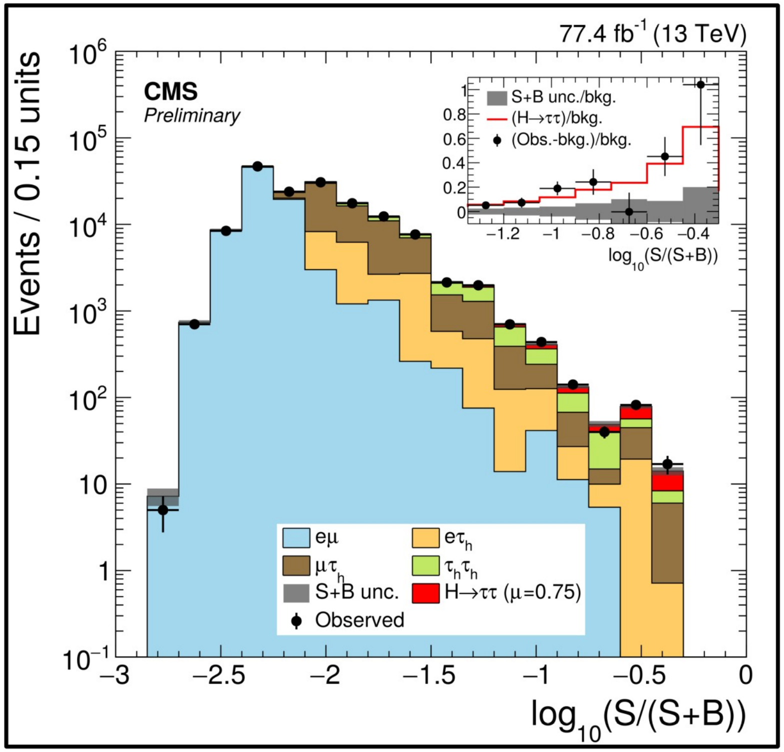
- After classification of **ggH** and **qqH** events are split into STXS bins, based on selection requirements on theory-related quantities after reconstruction:




CMS-PAS-HIG-18-032

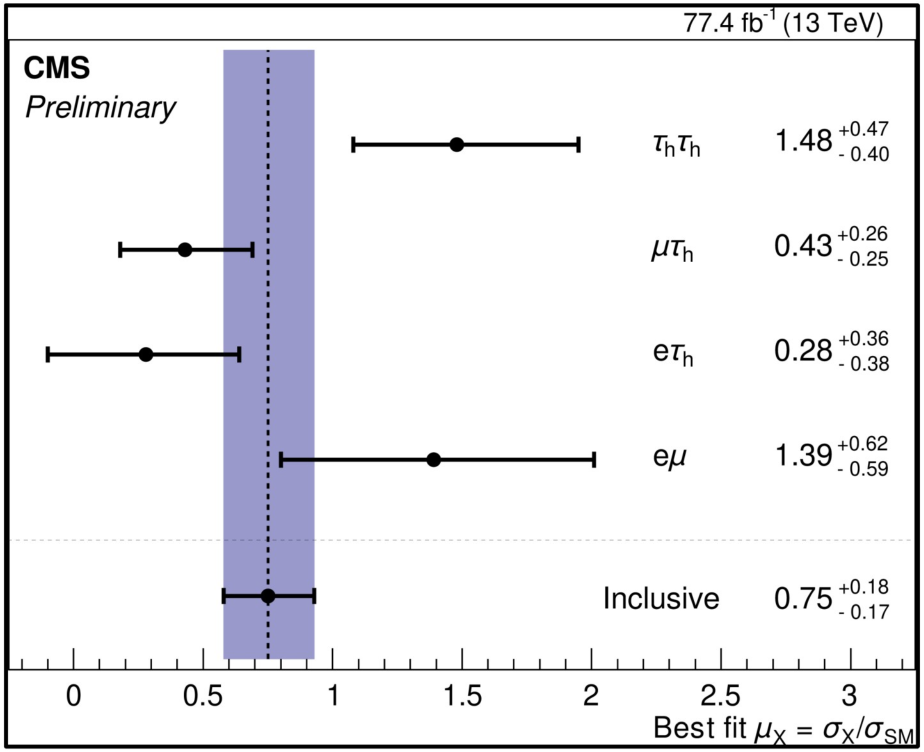
# Results (inclusive)

Inclusive signal (sorted by log(S/(S+B)))




CMS-PAS-HIG-18-032

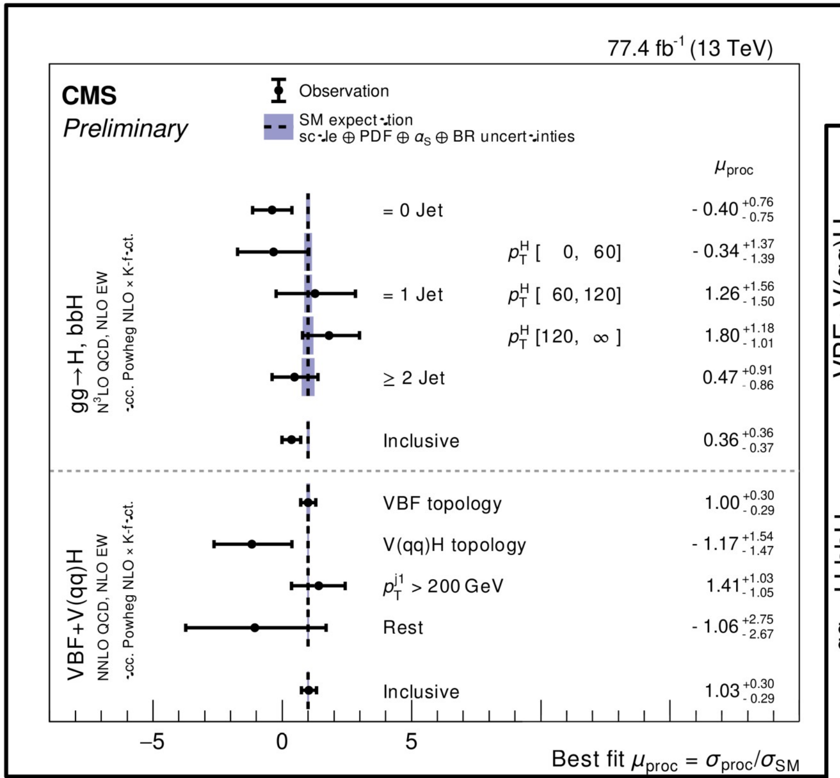
Signal strength: (top) split by final state and (bottom) inclusive




- **Clear signal** seen, though a bit on the low side, compared to other Higgs decay modes.

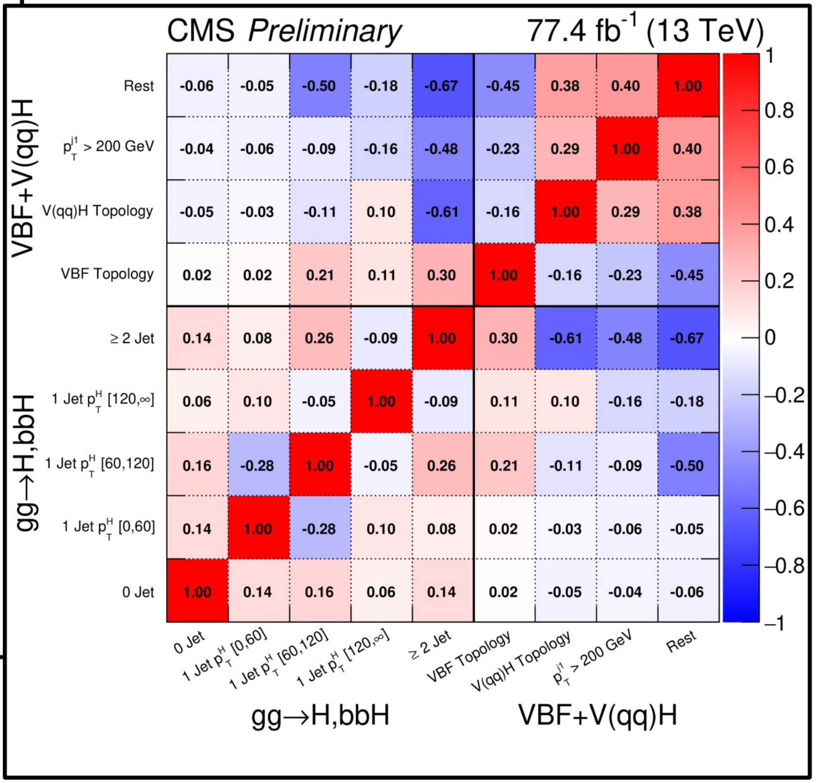
# Results (STXS)

- More **differential** measurement in 9 predefined STXS bins:



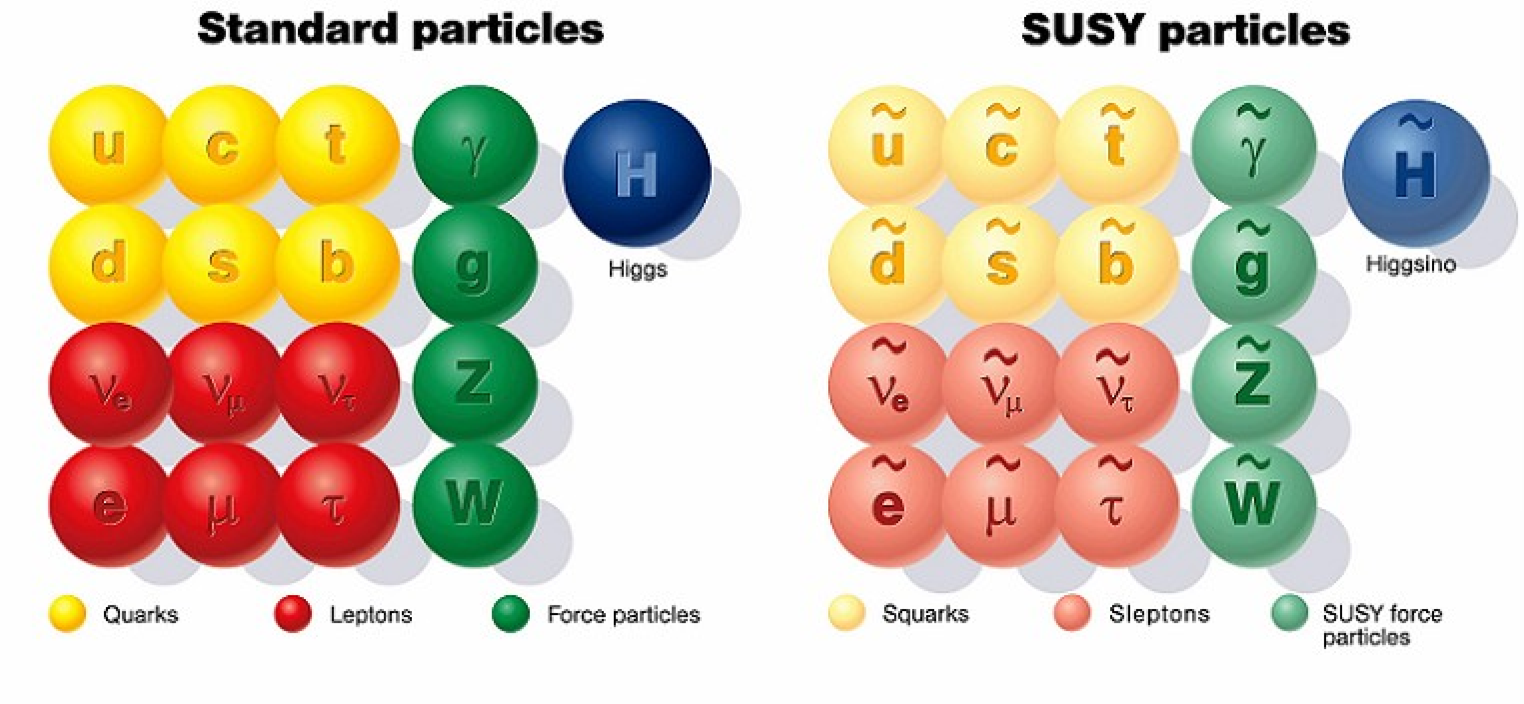

STXS bins

CMS-PAS-HIG-18-032




Correlation matrix

# MSSM $H \rightarrow \tau\tau$ analysis (*)




(*) as proxy for a well motivated Two Higgs Doublet Model (2HDM) extension of the SM.

# Higgs sector in SUSY

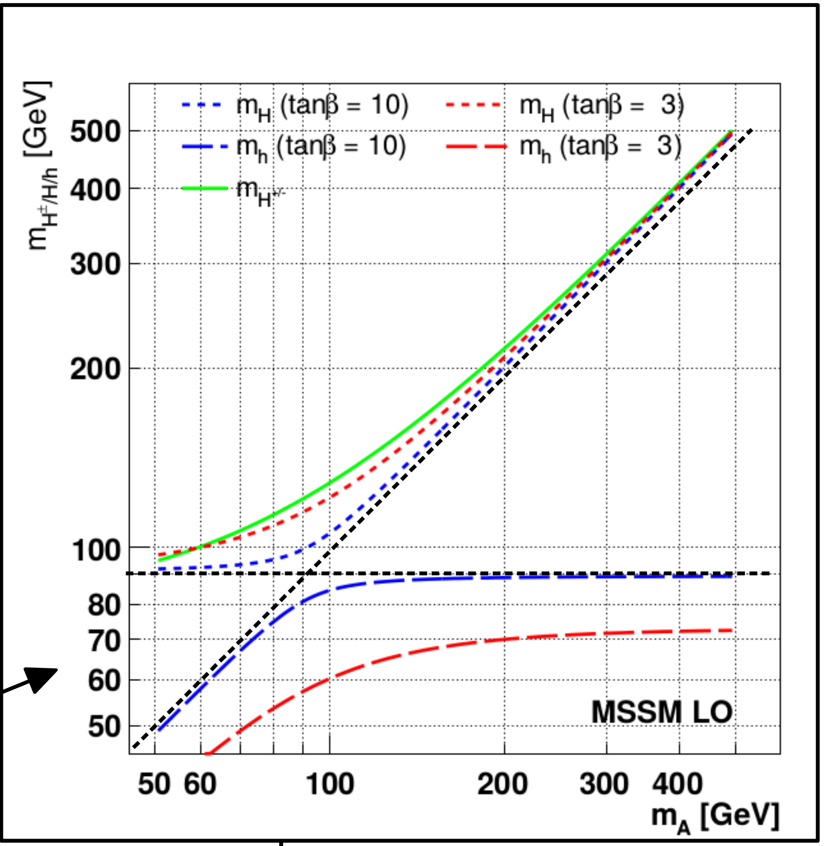
**NB**: w/o CP-violation in the SUSY Higgs sector.

- SUSY requires @ least 2 Higgs doublets (2HDM type-II) → **five Higgs bosons**:

$$\phi_u = \begin{pmatrix} \phi_u^+ \\ \phi_u^0 \end{pmatrix}, \quad Y_{\phi_u} = +1, \quad v_u : \mathrm{VEV}_u$$

$$\phi_d = \begin{pmatrix} \phi_d^0 \\ \phi_d^- \end{pmatrix}, \quad Y_{\phi_d} = -1, \quad v_d : \mathrm{VEV}_d$$

$$N_{\mathrm{ndof}} = 8 \; - \underbrace{3}_{W,\; Z} = \underbrace{5}_{H^{\pm},\; H,\; h,\; A}$$




- Strict mass requirements imposed by symmetry

- At tree level two free parameters: $m_A$, $\tan\beta = {}^{v_u}/{v_d}$.

$$m_{H^{\pm}}^2 = m_A^2 + m_W^2$$

$$m_{H,\,h}^2 = \frac{1}{2}\left(m_A^2 + m_Z^2 \pm \sqrt{(m_A^2+m_Z^2)^2 - 4m_A^2 m_Z^2 \cos^2 2\beta}\right)$$

$$\tan\alpha = \frac{-(m_A^2+m_Z^2)\sin 2\beta}{(m_Z^2-m_A^2)\cos 2\beta + \sqrt{(m_A^2+m_Z^2)^2 - 4m_A^2 m_Z^2\cos^2 2\beta}}$$

$\alpha$ : angle between $H$ and $h$ in mass matrix

# $m_h$ and $\tan\beta$ in the MSSM

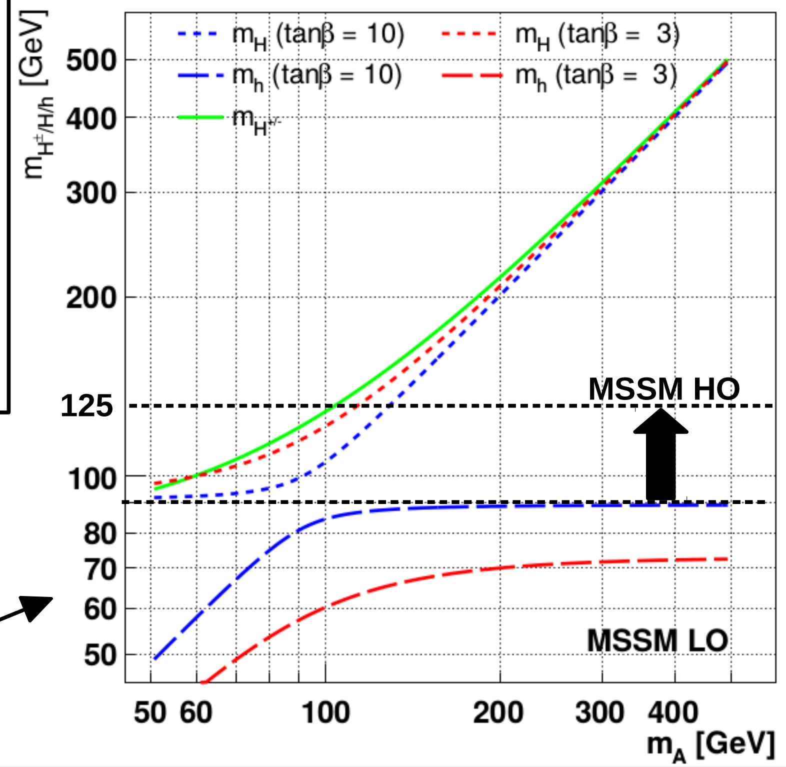
**NB**: w/o CP-violation in the SUSY Higgs sector.

$$m_h^2 \approx m_Z^2 \cos^2 2\beta + \Delta_{\mathrm{rad}}$$

$$\Delta_{\mathrm{rad}} = \frac{3}{(4\pi)^2}\frac{m_t^4}{v^2}\left(\ln\left(\frac{m_{\tilde{t}}^2}{m_t^2}\right) + \frac{X_t^2}{m_{\tilde{t}}^2}\left(1 - \frac{X_t^2}{12 m_{\tilde{t}}^2}\right)\right)$$

- +30% of $m_h$ due to higher order corrections.
- Following factors help to increase $m_h$: large $m_t$, large $m_{\tilde{t}}$ , large $X_t$ , large $\tan\beta$.

$$X_t = m_t\,(A_t - \mu\cot\beta)$$




$$m_{H^\pm}^2 = m_A^2 + m_W^2$$

$$m_{H,\,h}^2 = \tfrac{1}{2}\left(m_A^2 + m_Z^2 \pm \sqrt{(m_A^2+m_Z^2)^2 - 4m_A^2 m_Z^2\cos^2 2\beta}\right)$$

$$\tan\alpha = \frac{-(m_A^2+m_Z^2)\sin 2\beta}{(m_Z^2-m_A^2)\cos 2\beta + \sqrt{(m_A^2+m_Z^2)^2 - 4m_A^2 m_Z^2\cos^2 2\beta}}$$

$\alpha$ : angle between $H$ and $h$ in mass matrix

Priv. Doz. Dr. Roger Wolf
http://ekpwww.physik.uni-karlsruhe.de/~rwolf/

# Down-type fermions in the MSSM

**NB**: w/o CP-violation in the SUSY Higgs sector.

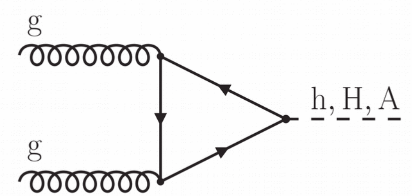
CP-odd part of coupling.

| | $g_{VV}$ | $g_{uu}$ | $g_{dd}$ |
|---|---|---|---|
| $A$ | $-$ | $\gamma_5 \cot\beta$ | $\gamma_5 \tan\beta$ |
| $H$ | $\cos(\beta-\alpha) \;\to 0$ | $\sin\alpha/\sin\beta \;\to \cot\beta$ | $\cos\alpha/\cos\beta \to \tan\beta$ |
| $h$ | $\sin(\beta-\alpha) \;\to 1$ | $\cos\alpha/\sin\beta \;\to \;1$ | $-\sin\alpha/\cos\beta \to \;1$ |

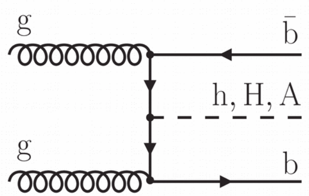
Relative to corresponding couplings to a SM Higgs boson.

For $m_A \gg m_Z$: $\alpha \to \beta - \pi/2$ (coupling $A/H$ to down-type fermions enhanced by $\tan\beta$).

## Production modes:



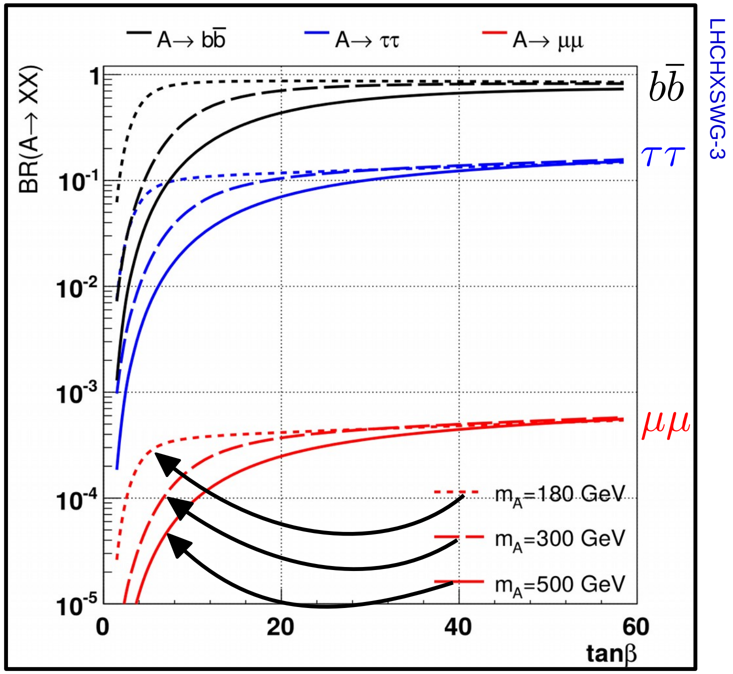

$gg \to \phi$ ("$gg\phi$")




$gg \to \phi b\bar{b}$ ("$bb\phi$")

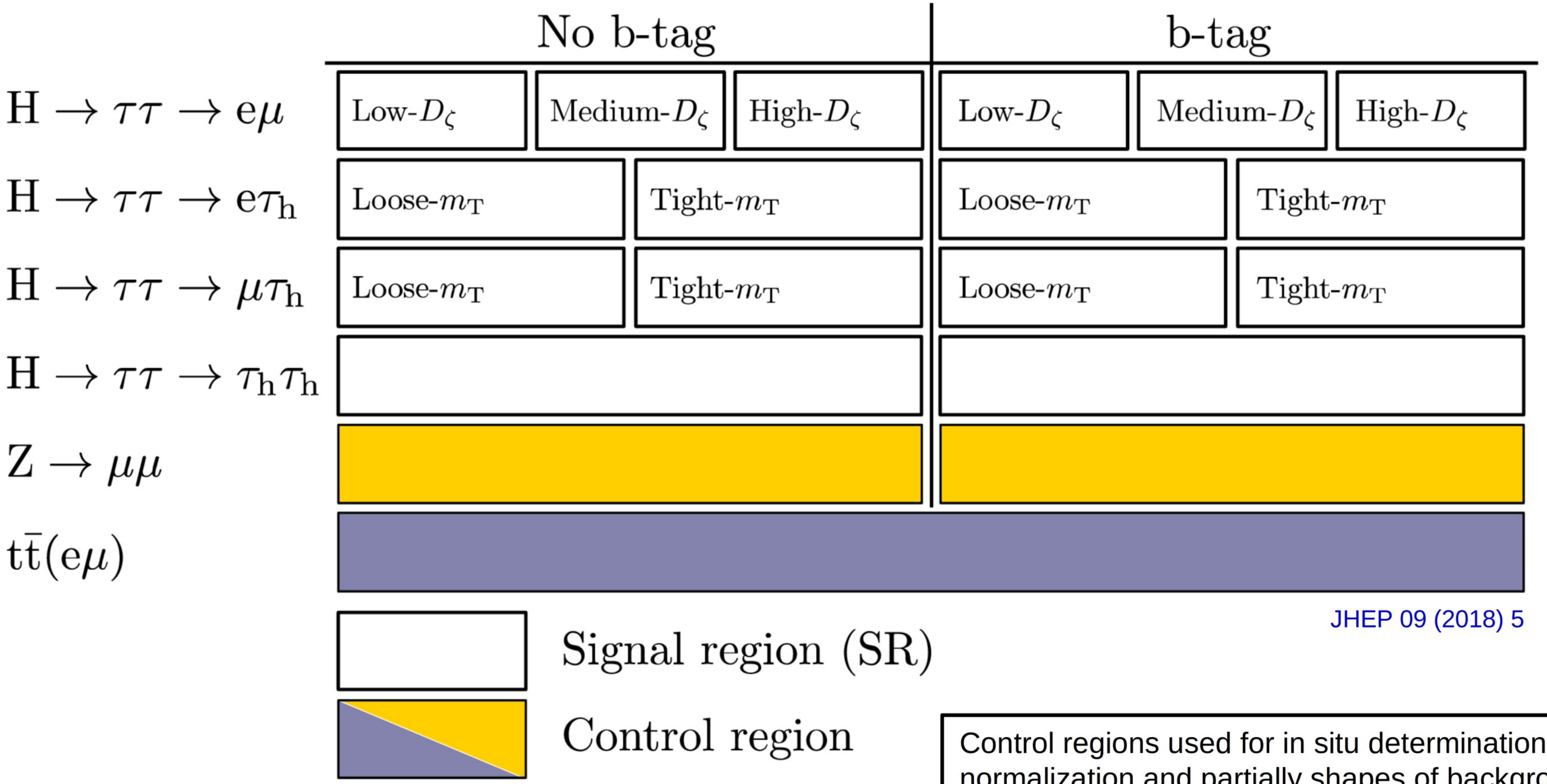
**X**

## Decay channels: $m_h^{\mathrm{mod+}}$



# Additional event information/categorization

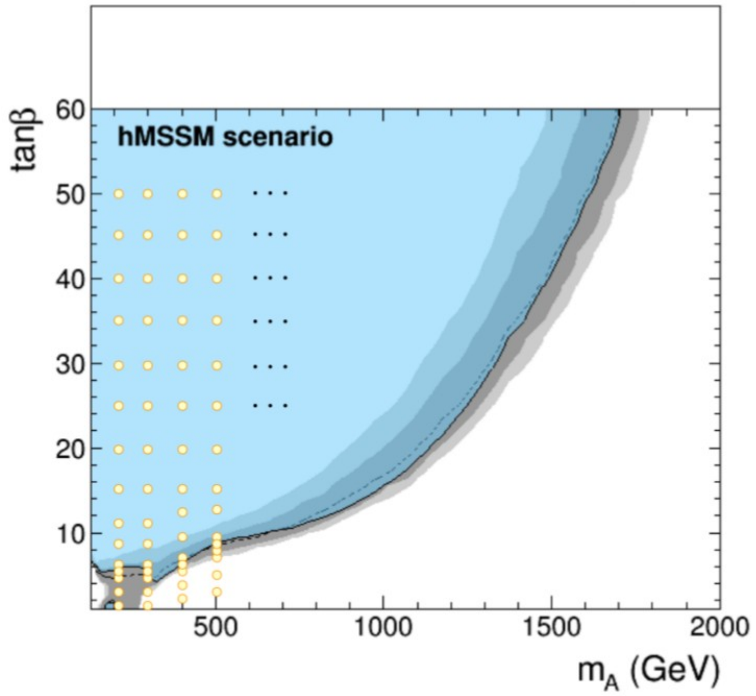
- Exploiting high mass of di-τ final state (via $m_\mathrm{T}$) and increased coupling to b quarks (via b-tag).
- Apart from this stay **more simplistic** with event categorization w.r.t SM analysis.




Control regions used for in situ determination of normalization and partially shapes of backgrounds in ML fit used for statistical inference of the signal.

# Signal modeling

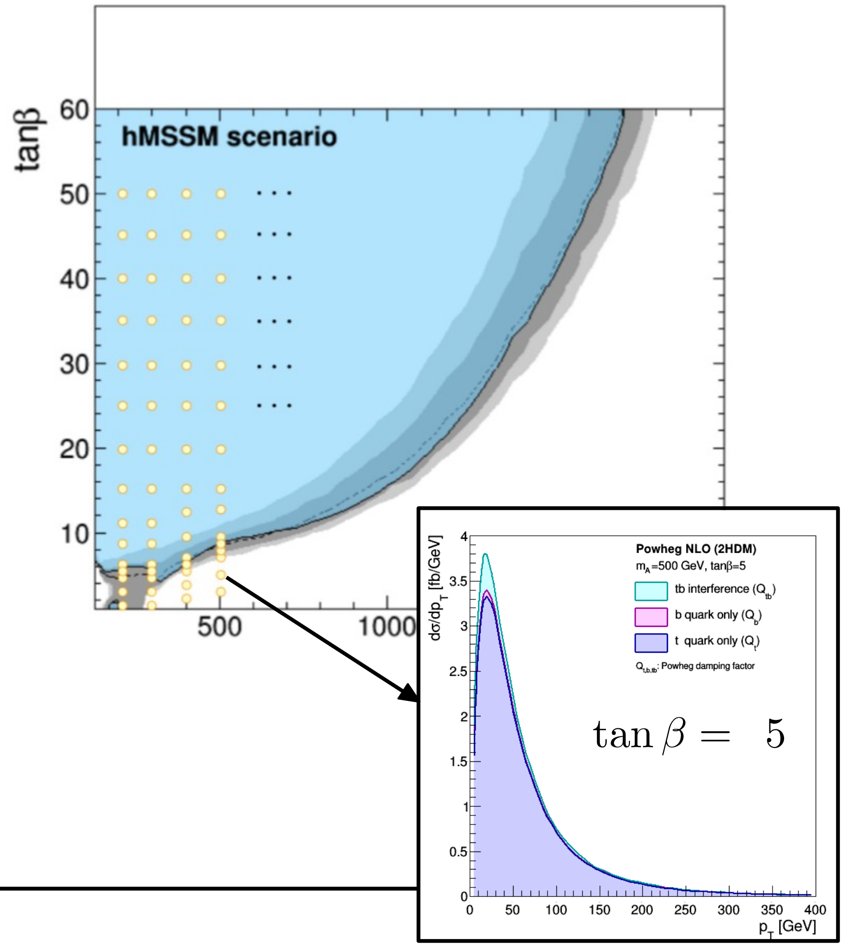
**Test MSSM vs SM hypothesis:** allows for well defined statistical problem, even when reaching sensitivity to the 125 GeV Higgs boson.




- Typical scan to determine exclusion contours in specific models.
- Determine CLs in each point in parameter space to obtain limit at significance level $\alpha$.

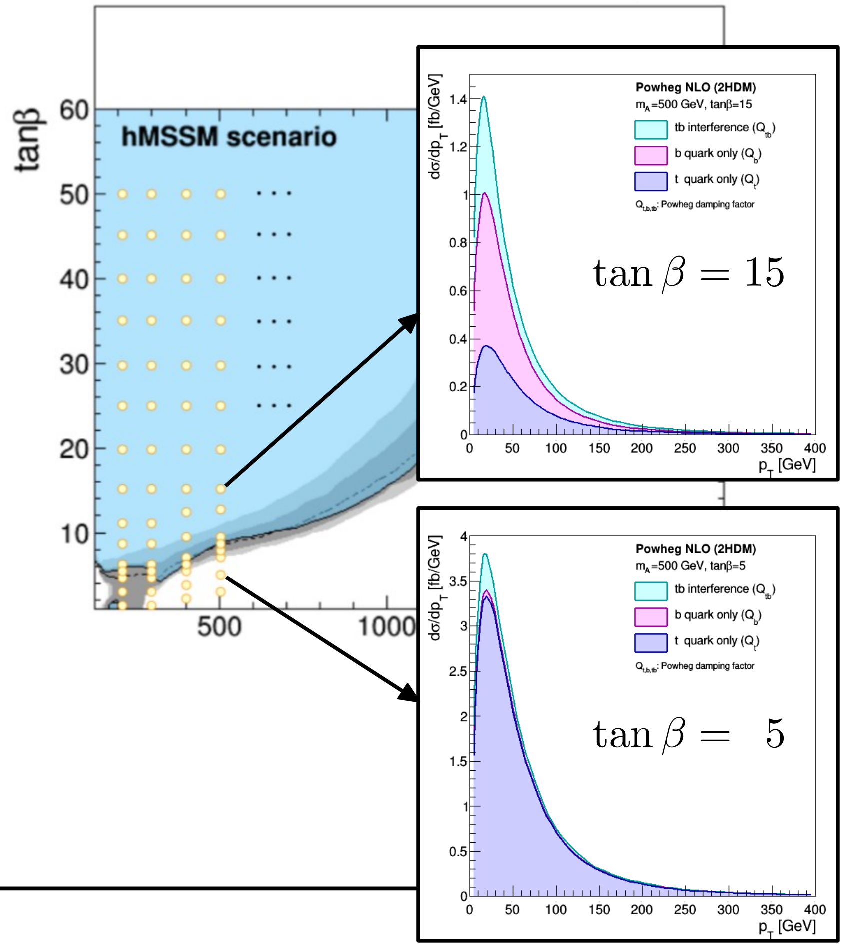
# Signal modeling

- $p_T(A,H,h)$@ NLO QCD + PS → **multiscale problem**.
- Plus: b contribution varies as a function of $\tan\beta$.



# Signal modeling

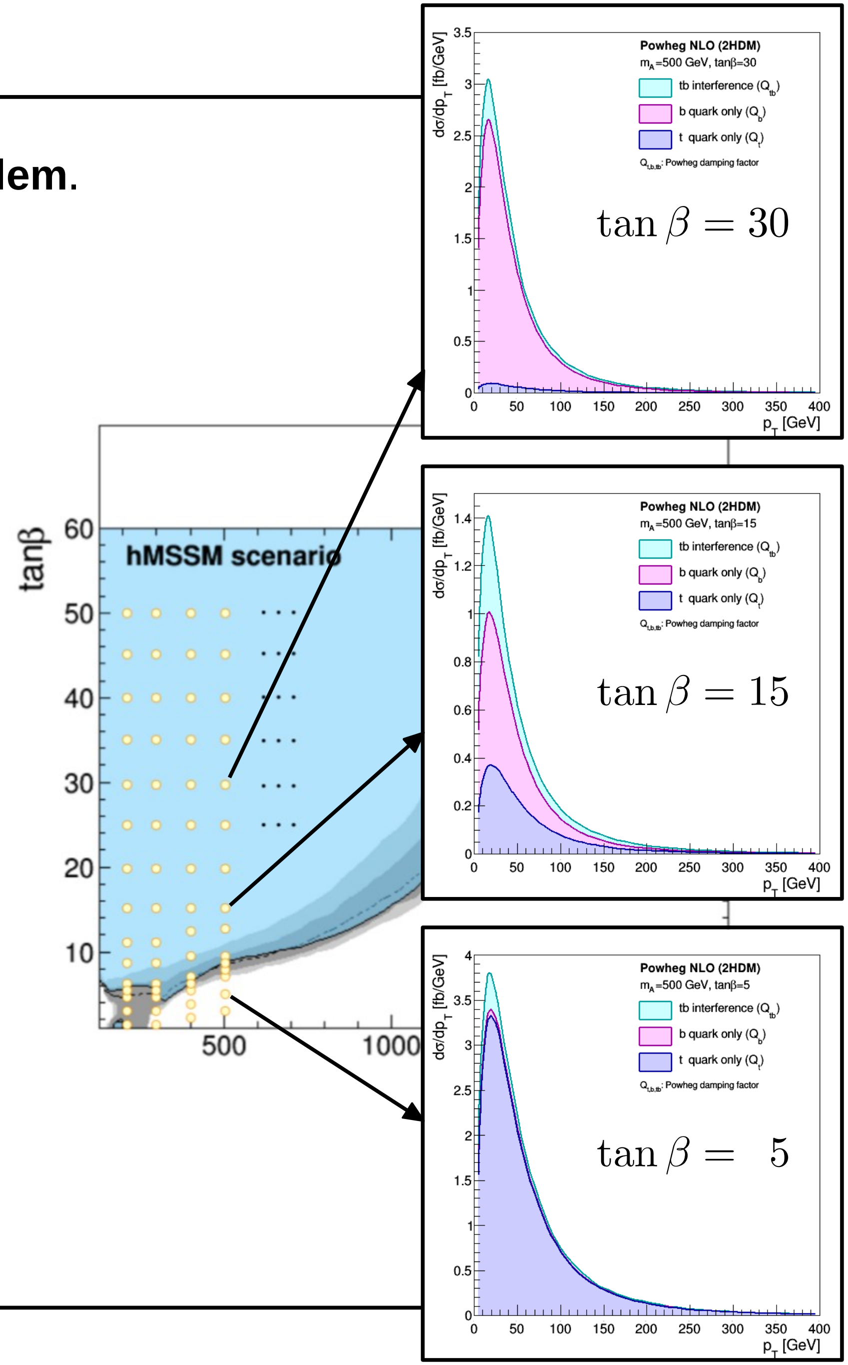
- $p_T(A, H, h)$@ NLO QCD + PS → **multiscale problem**.
- Plus: b contribution varies as a function of $\tan\beta$.



# Signal modeling

- $p_T(A,H,h)$@ NLO QCD + PS → **multiscale problem**.
- Plus: b contribution varies as a function of $\tan\beta$.

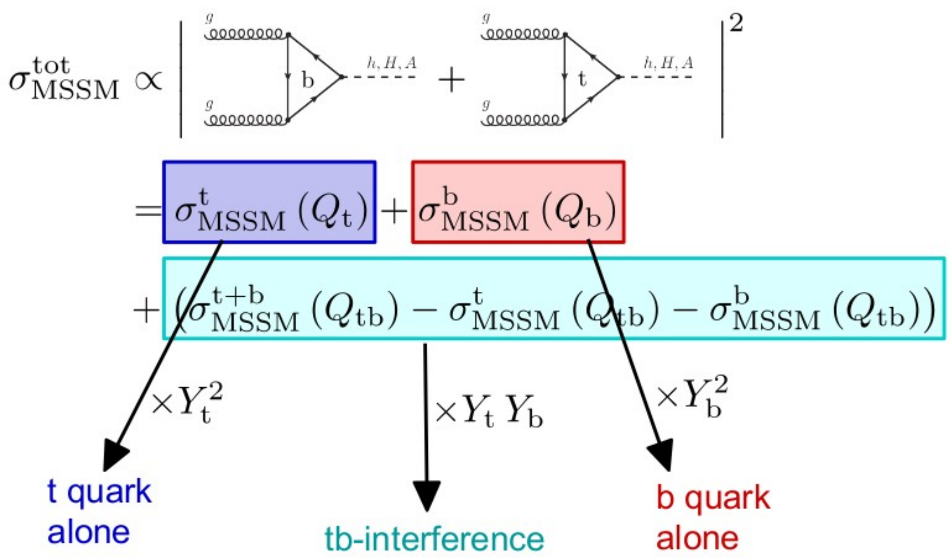
Change in $p_T(A,H,h)$ implies change in signal acceptance.

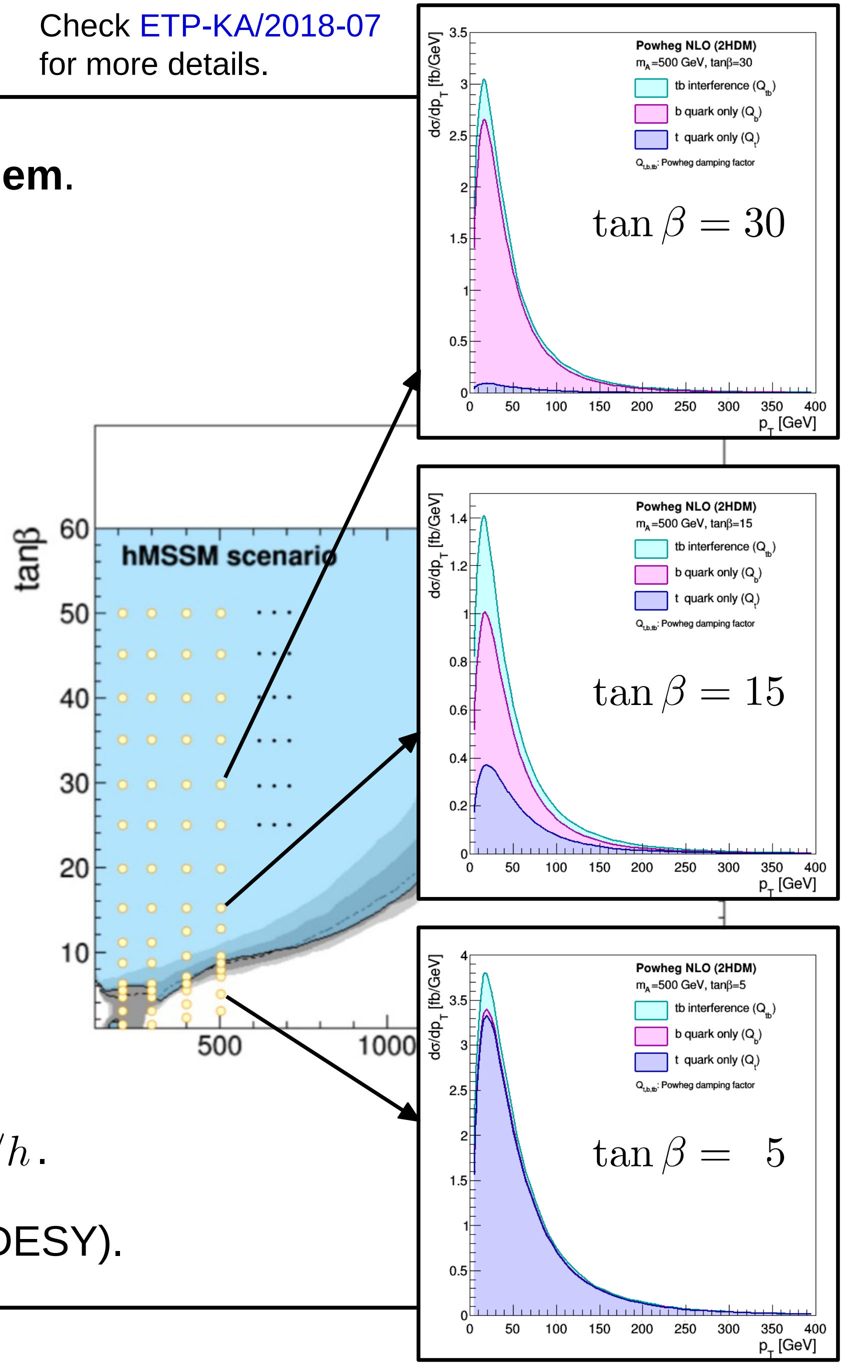


# Signal modeling

Check ETP-KA/2018-07 for more details.

- $p_T(A, H, h)$@ NLO QCD + PS → **multiscale problem**.
- Plus: b contribution varies as a function of $\tan\beta$.



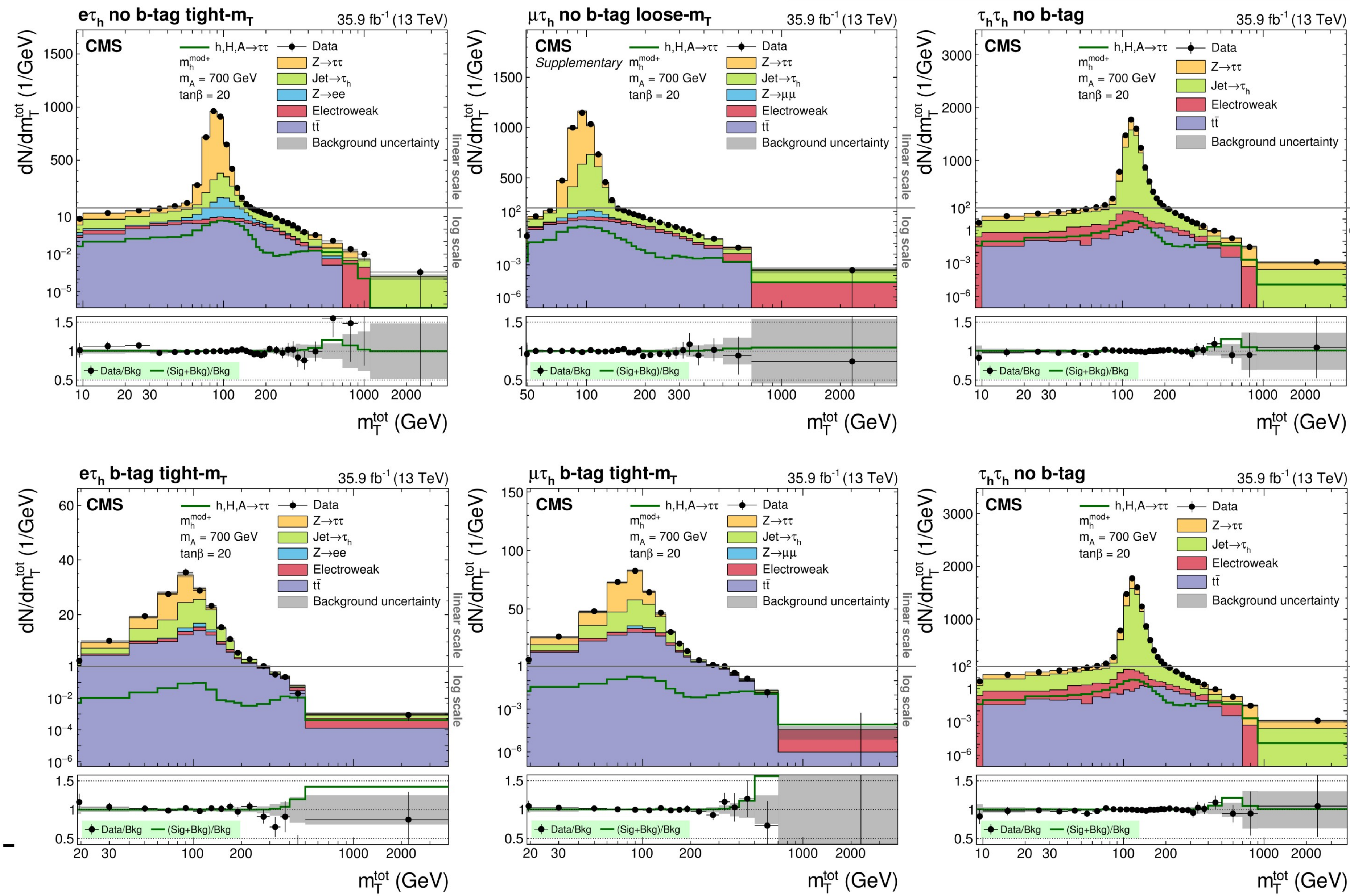




- Taking into account all $\tan\beta$ enhanced SUSY corrections and non-trivial $\tan\alpha$ dependency for $H/h$.
- Developed with S. Liebler (KIT) and E. Bagnashi (DESY).

# Observation

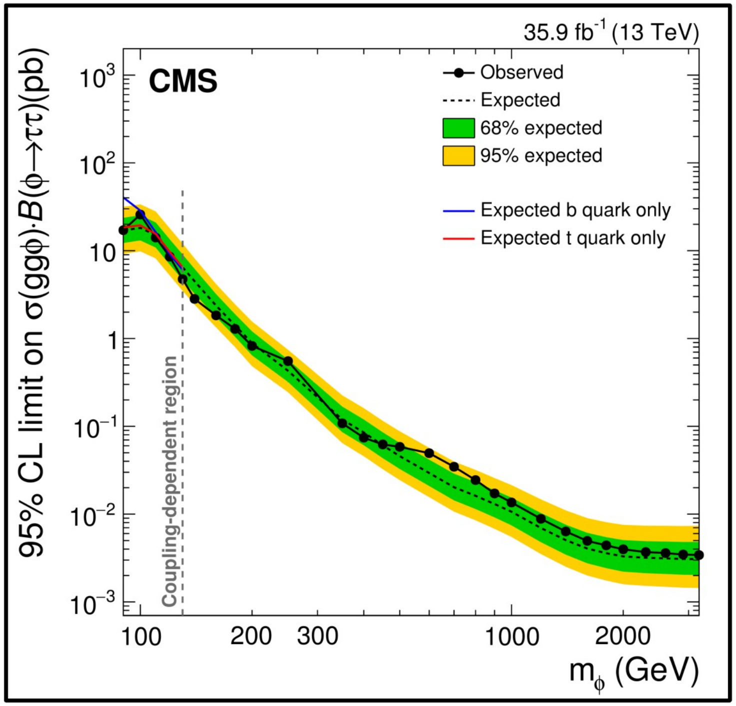
JHEP 09 (2018) 5

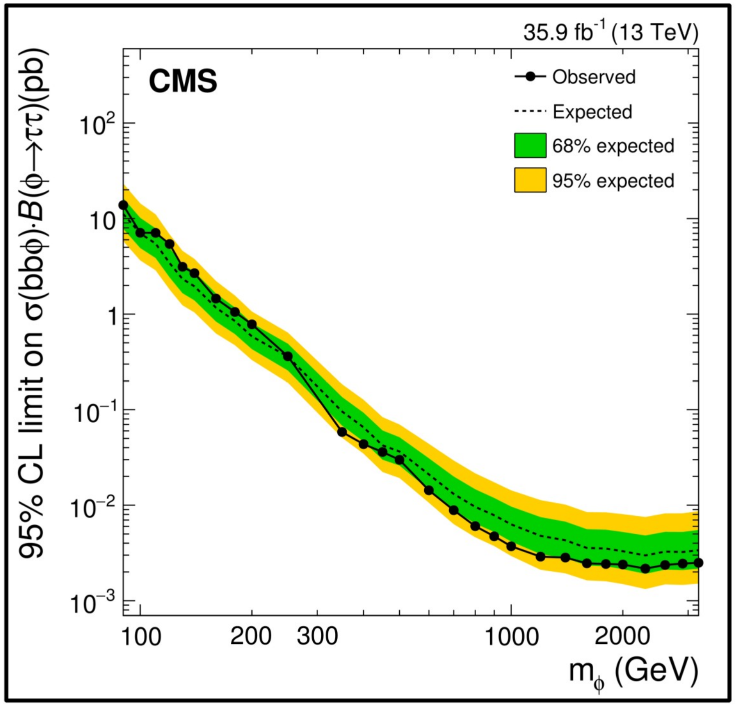
Shown are the most sensitive categories with an MSSM $m_h^{\mathrm{mod+}}$ hypothesis w/ $m_A = 700\ \mathrm{GeV}$ and $\tan\beta = 20$ fitted to the data.



# Model independent limits

- **Narrow width approximation**, two parameters of interest, $\mu_{gg\phi}$ and $\mu_{bb\phi}$ .



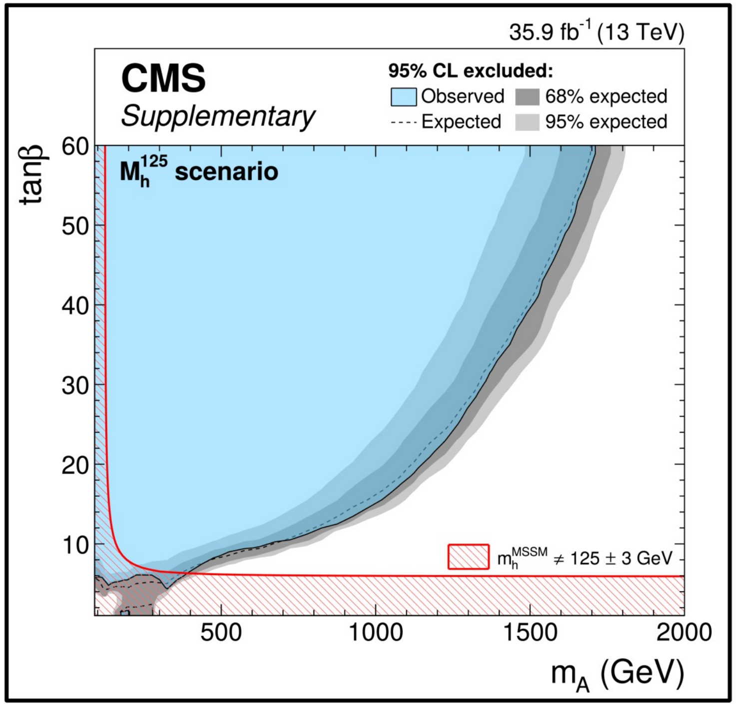




JHEP 09 (2018) 5

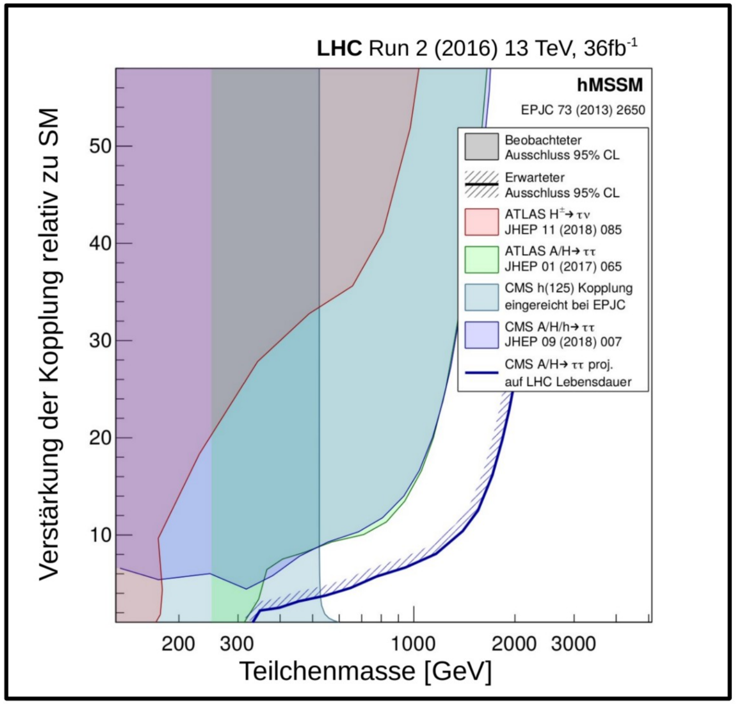
- No deviation beyond $2\sigma$ found.
- Cross checks discussed e.g. in ETP-KA/2017-21 and ETP-KA/2017-31.

# Model dependent exclusion contours

- In predefined **benchmark models**:




JHEP 09 (2018) 5




- In general parameter space is explored down to $\tan\beta \gtrsim 6$ for $\mathrm{m}_A \lesssim 250\ \mathrm{GeV}$ and up to $\mathrm{m}_A \leq 1600\ \mathrm{GeV}$.

# Summary

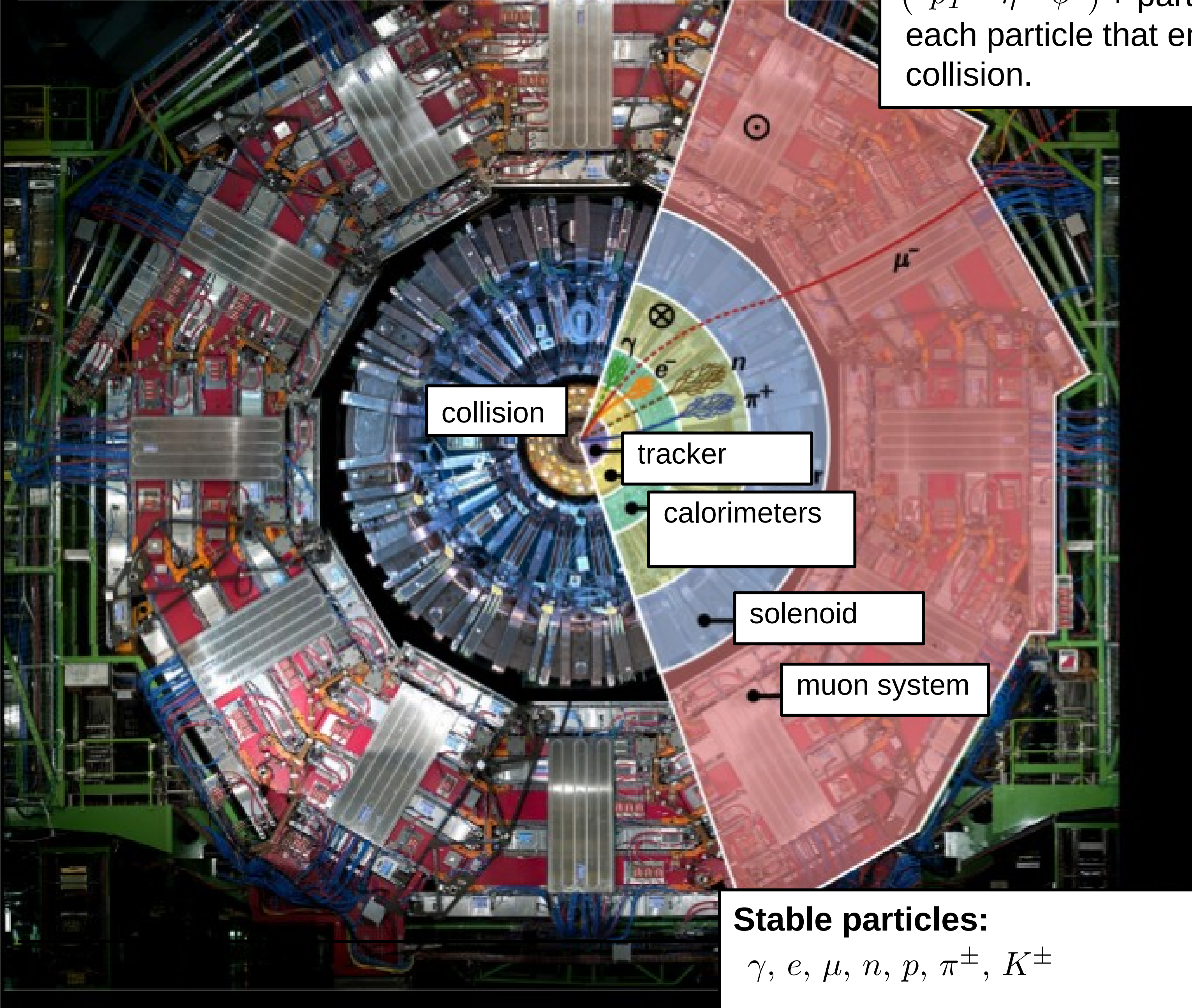
- Di-$\tau$ is one of the **most interesting final states** in the Higgs physics program of the LHC.
  - Best access to Higgs boson **couplings to fermions**.
  - **Large event yields**, reasonably well accessible (e.g. for studies of specific production modes, like VBF).
  - Most interesting final state to **search for extensions** of the SM Higgs sector.
- CMS had a very successful start in analyzing the LHC Run-2 data.
- KIT has a significant contribution to everything that has been shown.
- We are looking forward to analyze the full LHC Run-2 data.

# Backup

# Reminder of CMS

**What we want to know:**

$(\, p_T \quad \eta \quad \phi \,)$ + particle type $(\mathrm{m})$ from each particle that emerges the collision.

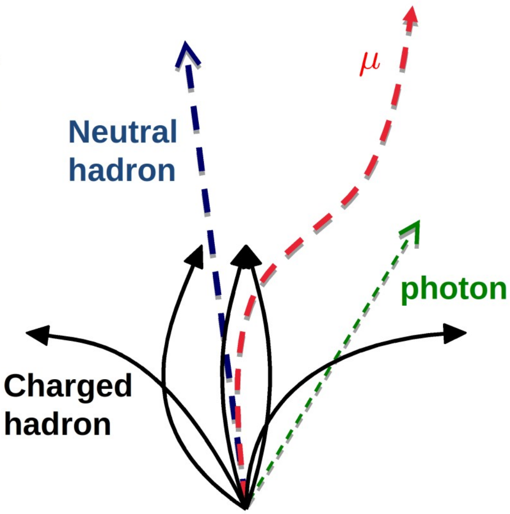


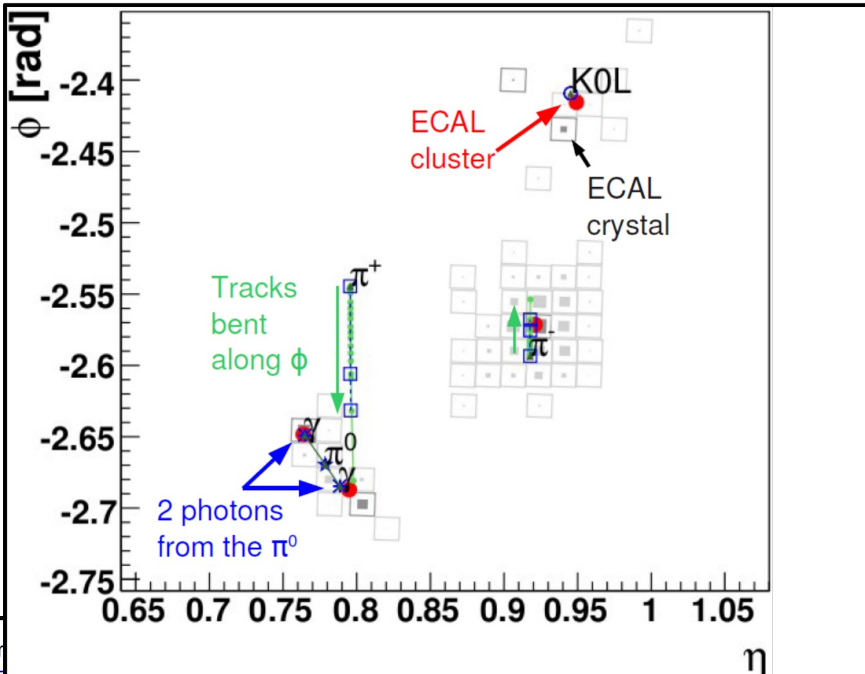

**Stable particles:**

$\gamma,\ e,\ \mu,\ n,\ p,\ \pi^{\pm},\ K^{\pm}$

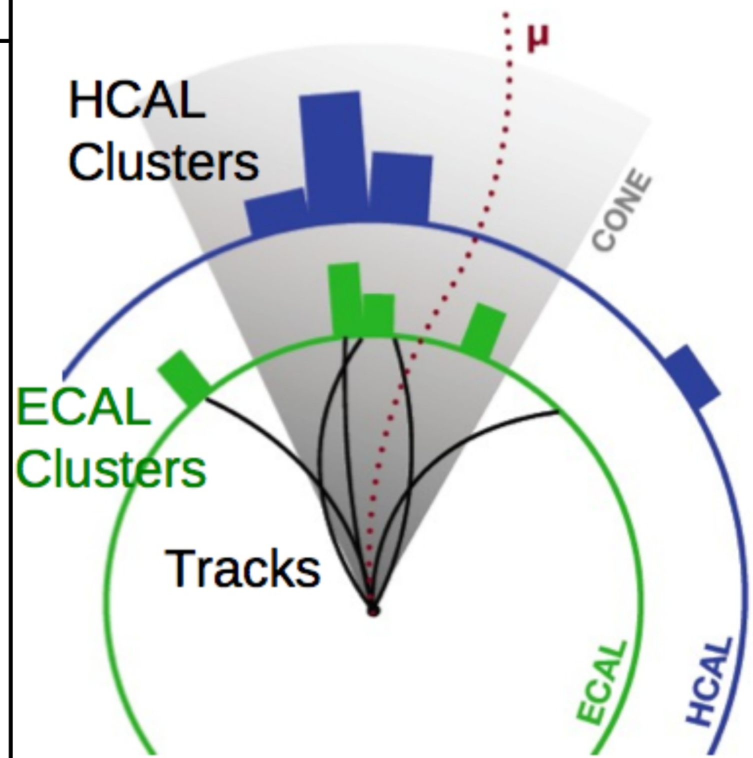
# Particle reconstruction @ CMS

- **Combine all energy deposits** in the detector to obtain a unique event description.




**Particle Flow:**

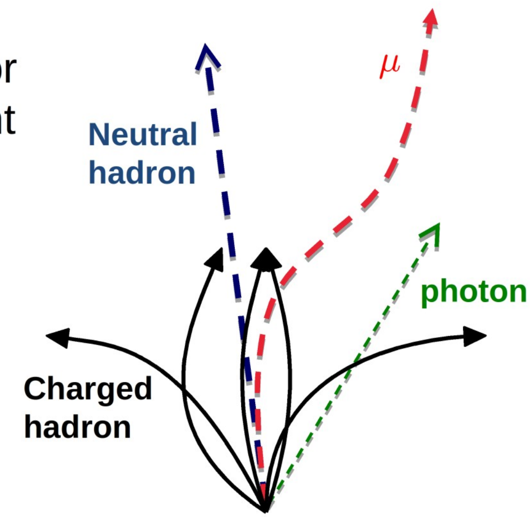


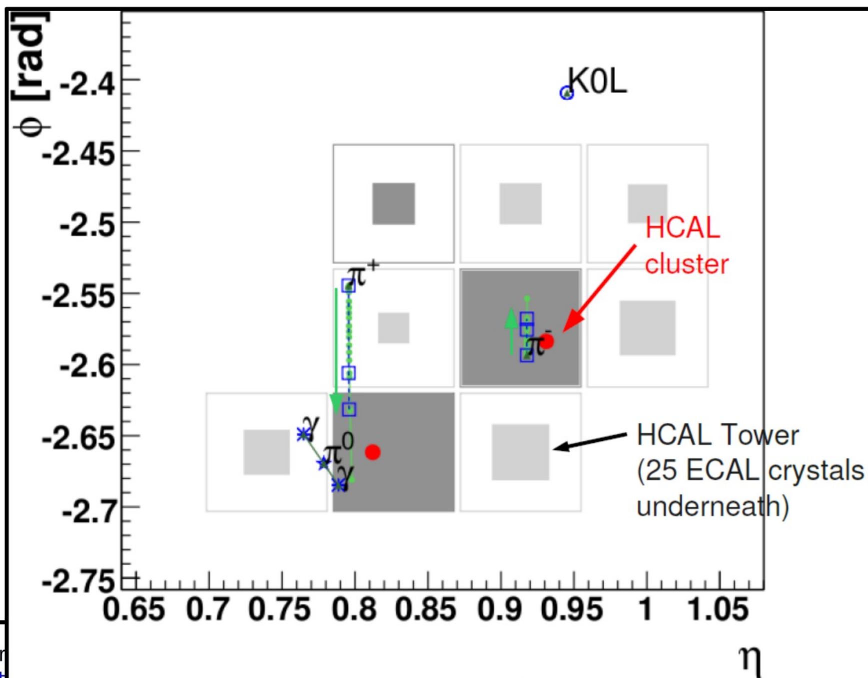

arxiv:1706.04965

- Unambiguous list of stable particles: muons, electrons, photons, charged & neutral hadrons.

# Particle reconstruction @ CMS

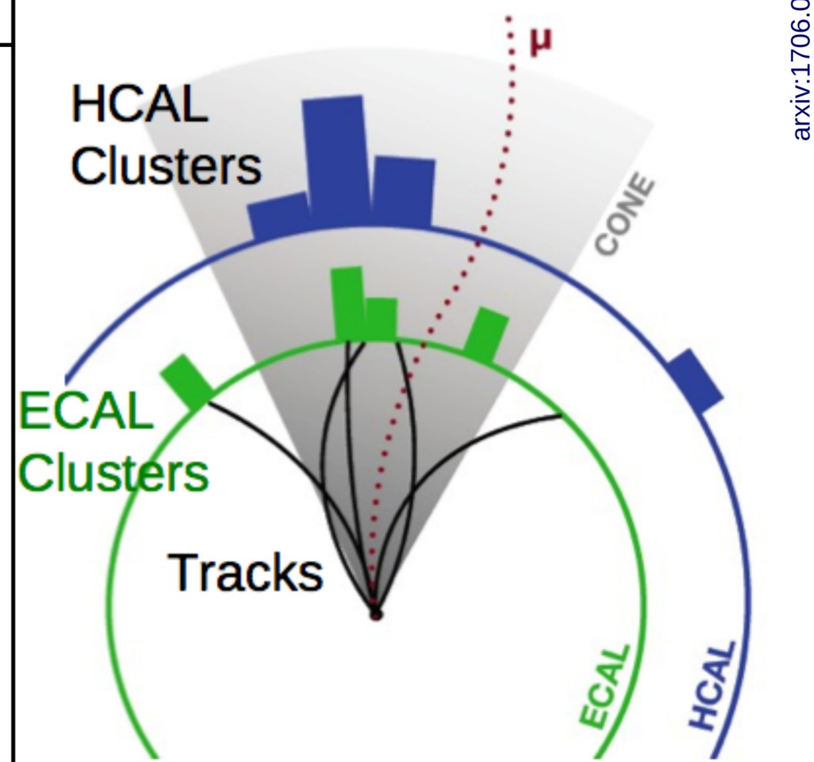
- **Combine all energy deposits** in the detector to obtain a unique event description.




**Particle Flow:**

arxiv:1706.04965




- Unambiguous list of stable particles: muons, electrons, photons, charged & neutral hadrons.