La Thuile YSF: March 9, 2023

Search for new physics in top quark production with additional leptons using the framework of effective field theory

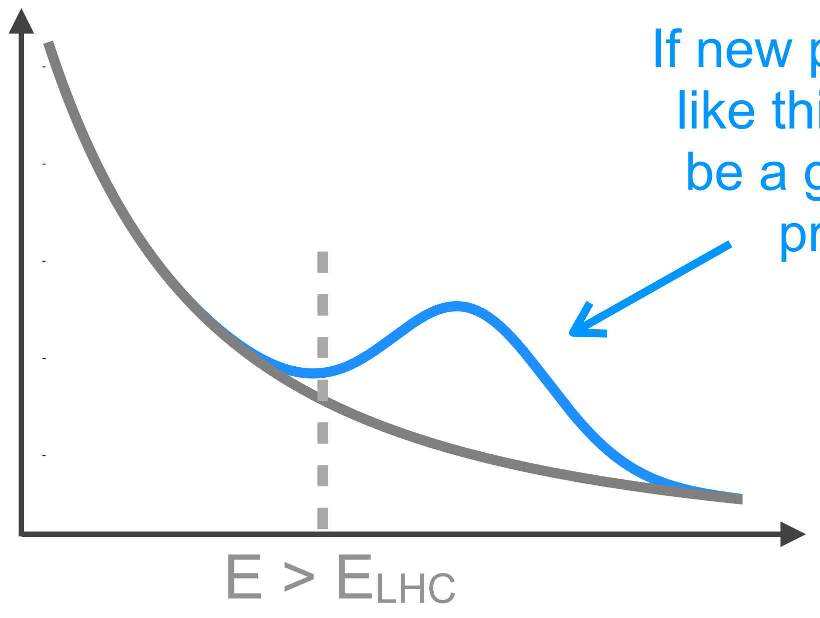
Kelci Mohrman
On behalf of the CMS collaboration

Motivation for indirect searches for new physics

- There are strong indications that the SM is not the complete description of nature, but there's no guarantee that the new particles would be light enough to be produced on shell at the LHC

- Indirect methods of probing higher mass scales are thus becoming increasingly interesting in the search for new physics at the energy frontier

- Effective field theory (EFT) is an example of such an indirect probe, and offers a model independent method of extending the discovery reach of the LHC



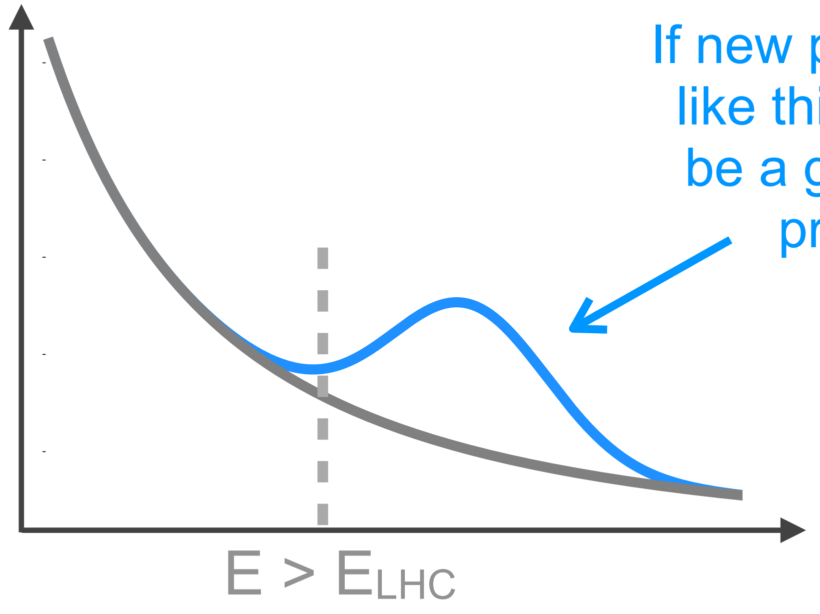
If new physics looks like this, EFT may be a good way of probing it

Motivation for indirect searches for new physics

- There are strong indications that the SM is not the complete description of nature, but there's no guarantee that the new particles would be light enough to be produced on shell at the LHC

- Indirect methods of probing higher mass scales are thus becoming increasingly interesting in the search for new physics at the energy frontier

- Effective field theory (EFT) is an example of such an indirect probe, and offers a model independent method of extending the discovery reach of the LHC



If new physics looks like this, EFT may be a good way of probing it

Brief introduction to EFT

- EFT treats the SM as the lowest order term in an expansion of higher-dimensional operators, that describe physics at a scale Λ, interacting with a strength determined by a dimensionless parameter called a Wilson coefficient, *c*
- If all Wilson coefficients (WCs) are 0, the SM Lagrangian is recovered
  −> a non-zero WC would indicate new physics

Wilson Coefficient (strength of interaction)

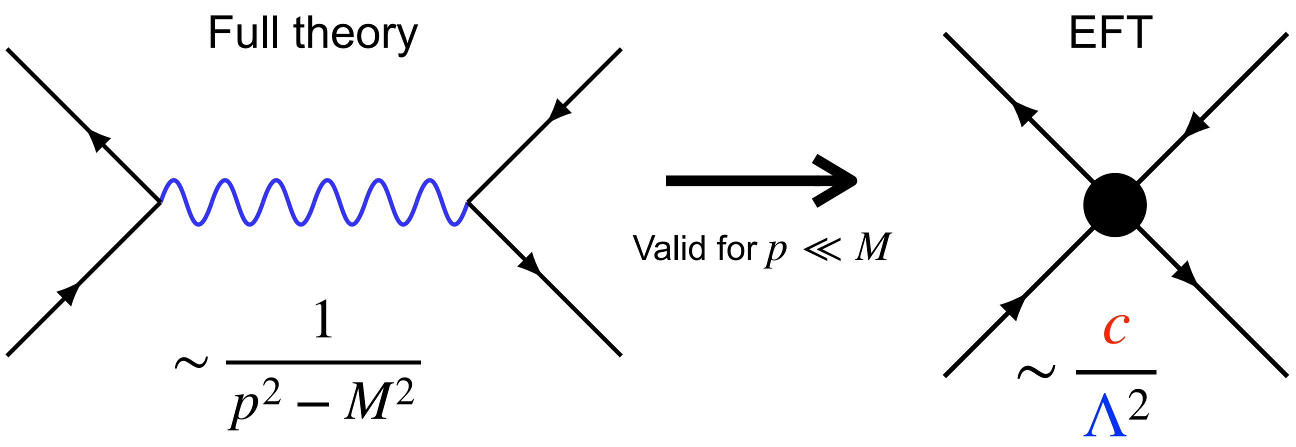
Energy scale of the new physics

Operators are built of products of SM fields and their derivatives

$$\mathscr{L}_{EFT} = \mathscr{L}_{SM} + \sum_i \frac{c_i}{\Lambda}\mathcal{O}_i^{(5)} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \ldots \longleftarrow$$

We focus on the **dim 6** terms, as they are the lowest order terms that contribute

- Example: If a heavy particle can't be produced on-shell at the LHC, would be hard to find via a direct search, but EFT can describe the interaction with a dim6 EFT operator, where the strength of the interaction is determined by the WC *c*



Brief introduction to EFT

- EFT treats the SM as the lowest order term in an expansion of higher-dimensional operators, that describe physics at a scale Λ, interacting with a strength determined by a dimensionless parameter called a Wilson coefficient, *c*
- If all Wilson coefficients (WCs) are 0, the SM Lagrangian is recovered
  −> a non-zero WC would indicate new physics

Wilson Coefficient (strength of interaction)

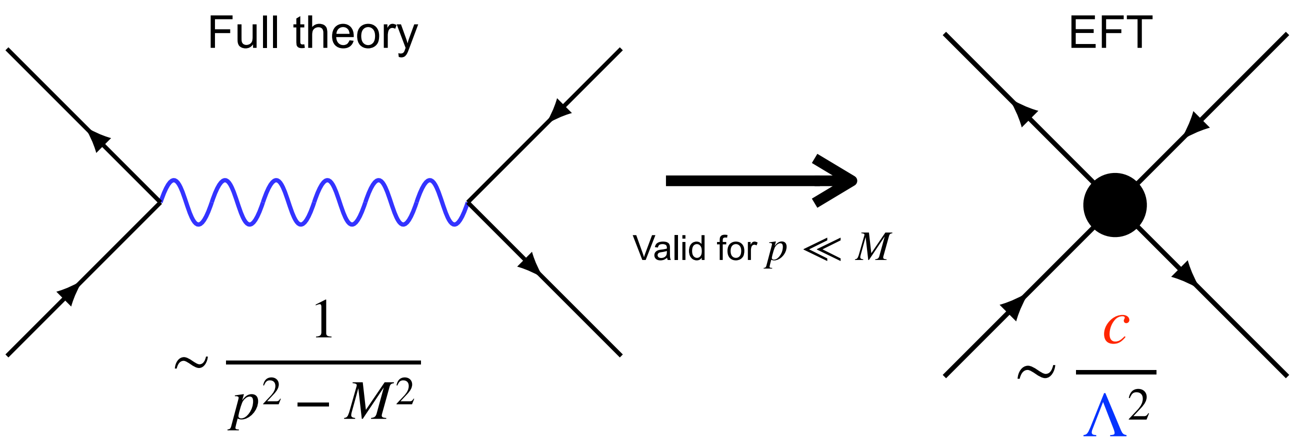
Energy scale of the new physics

Operators are built of products of SM fields and their derivatives

$$\mathscr{L}_{EFT} = \mathscr{L}_{SM} + \sum_i \frac{c_i}{\Lambda}\mathcal{O}_i^{(5)} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \dots \longleftarrow$$

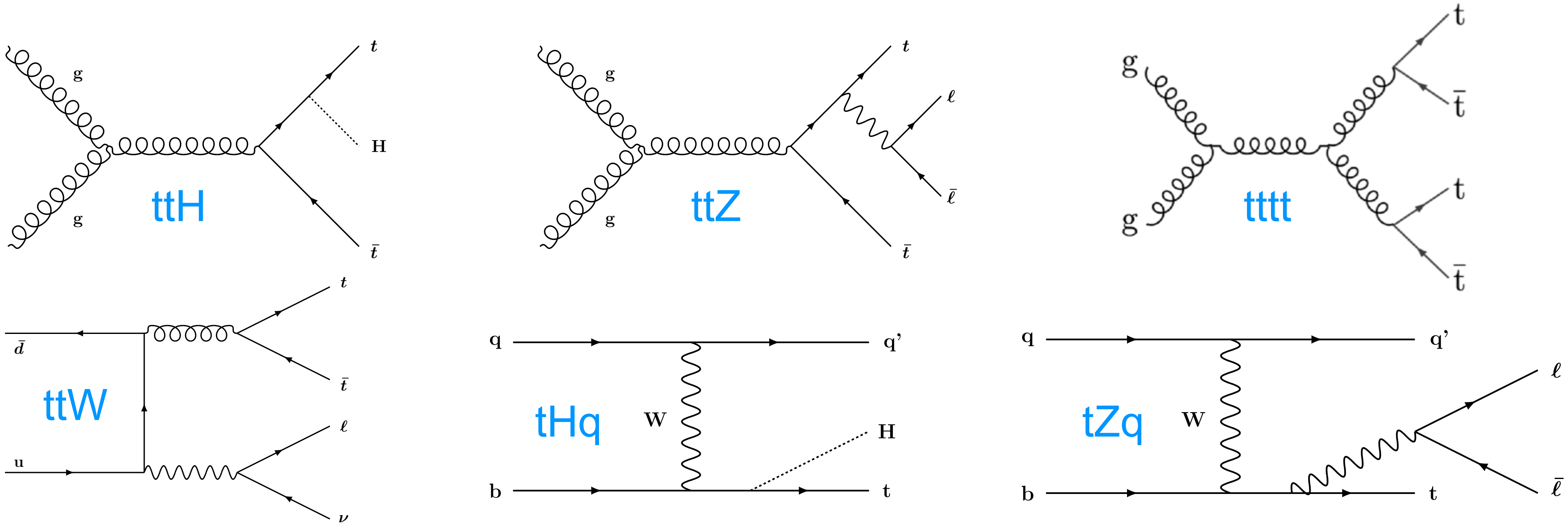
We focus on the **dim 6** terms, as they are the lowest order terms that contribute

- Example: If a heavy particle can't be produced on-shell at the LHC, would be hard to find via a direct search, but EFT can describe the interaction with a dim6 EFT operator, where the strength of the interaction is determined by the WC *c*



Overview of analysis goals

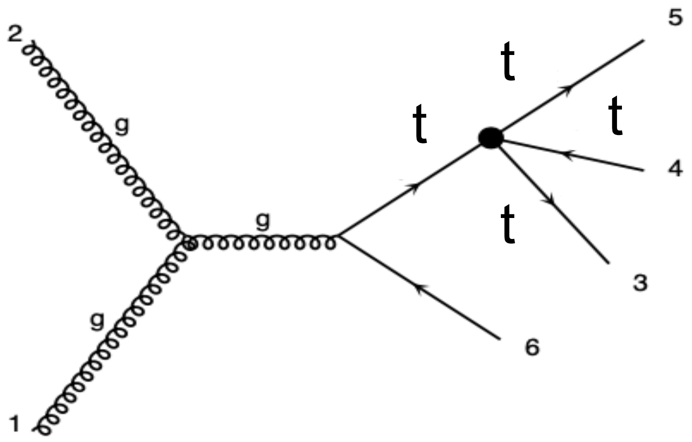
- This analysis uses EFT to probe new physics impacting associated top production
  - These processes are relatively rare, involve heavy particles, and may be an interesting region for new physics to be hiding
  - Signal processes: ttH, ttlnu, ttll, tllq, tHq, tttt
  - Global approach, probe all effects of dimension-6 EFT operators (involving top quarks) that can impact these processes



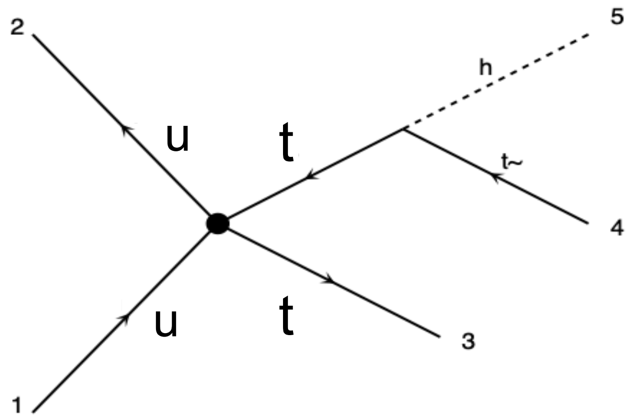
A few example associated top production diagrams

EFT operators impacting associated top processes

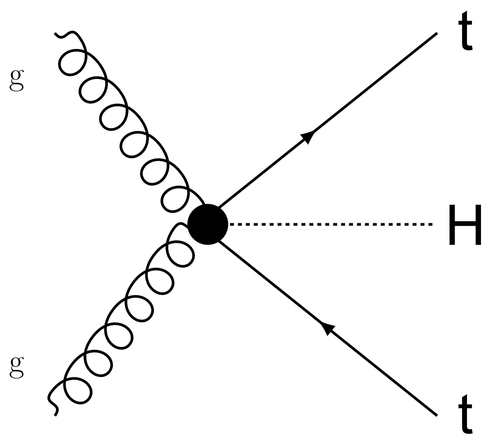
- We study 26 Wilson Coefficients (WCs) that significantly impact associated top processes, the operators fall into 4 main categories:



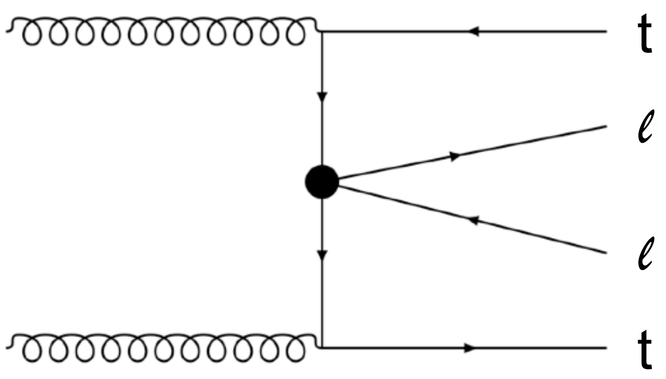
Operators involving 4 heavy quarks



Operators involving two heavy quarks and 2 light quarks



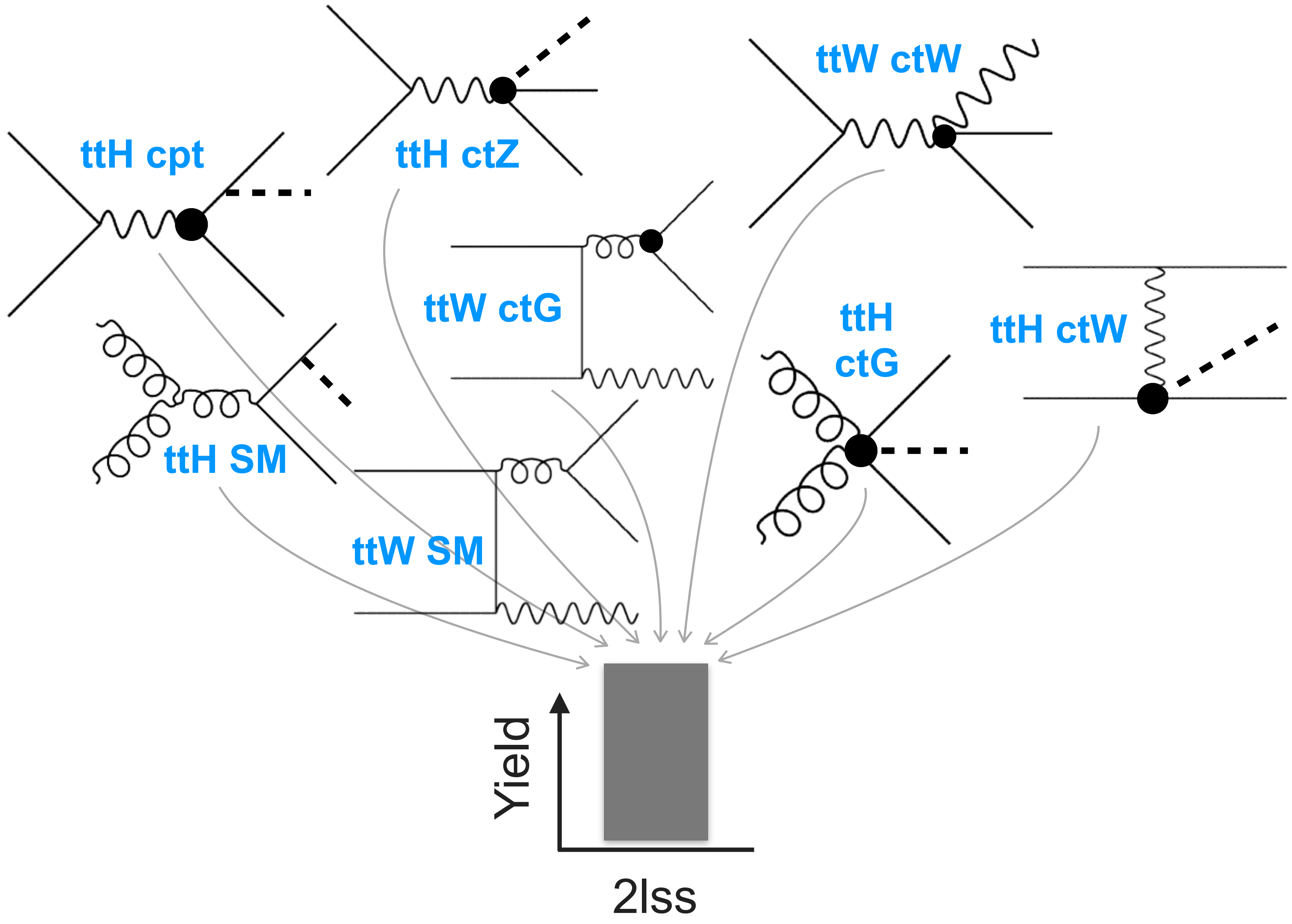
Operators involving heavy quarks and bosons



Operators involving 2 heavy quarks and 2 leptons

Strategy for multilepton EFT analysis

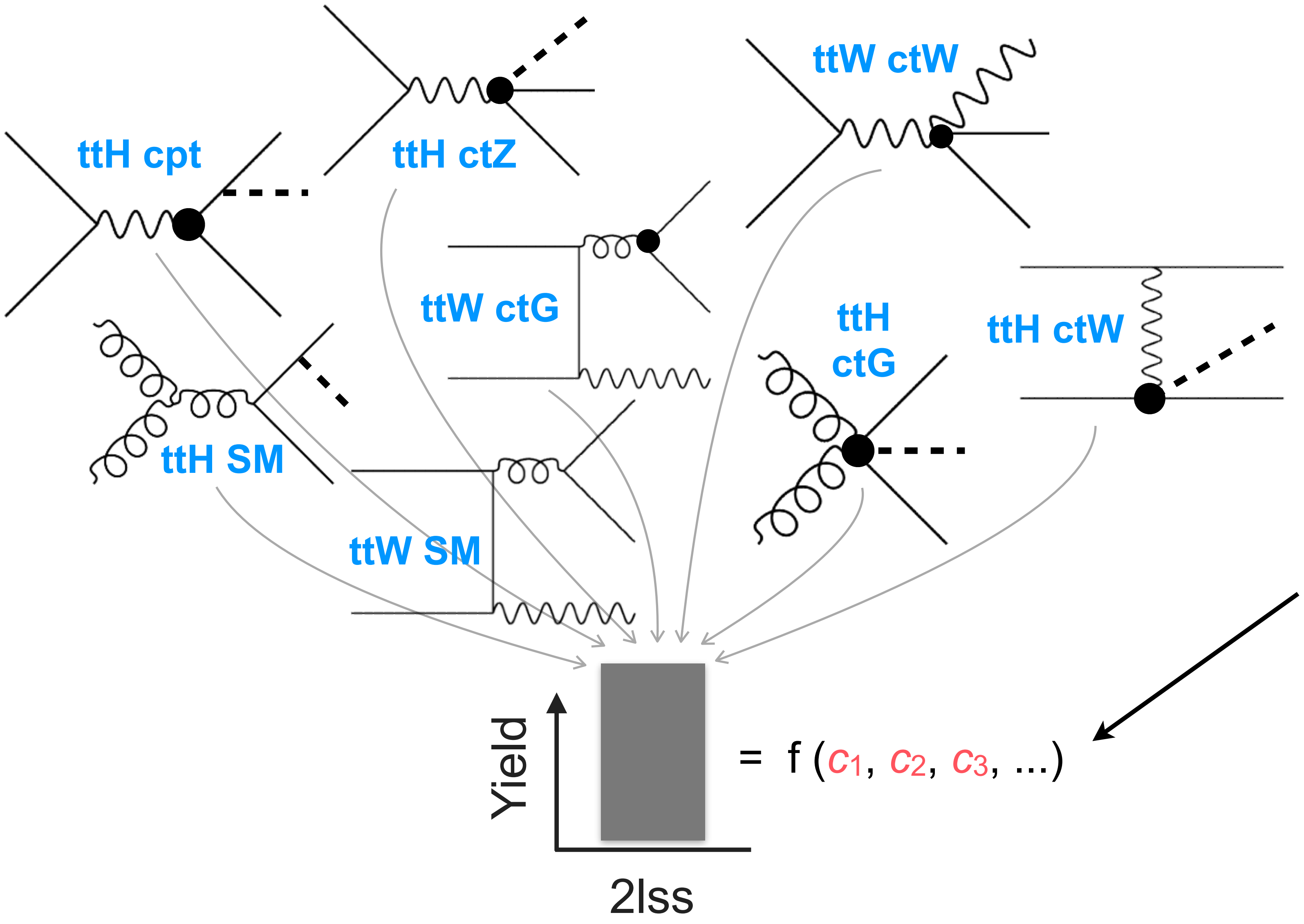
- We focus on multilepton signatures, many advantages but also challenges:
  - Multiple signal processes contribute to the same final state signatures
  - Many WCs can affect the processes, interfere with each other and the SM



Important to analyze all operators across all channels simultaneously

Strategy for multilepton EFT analysis

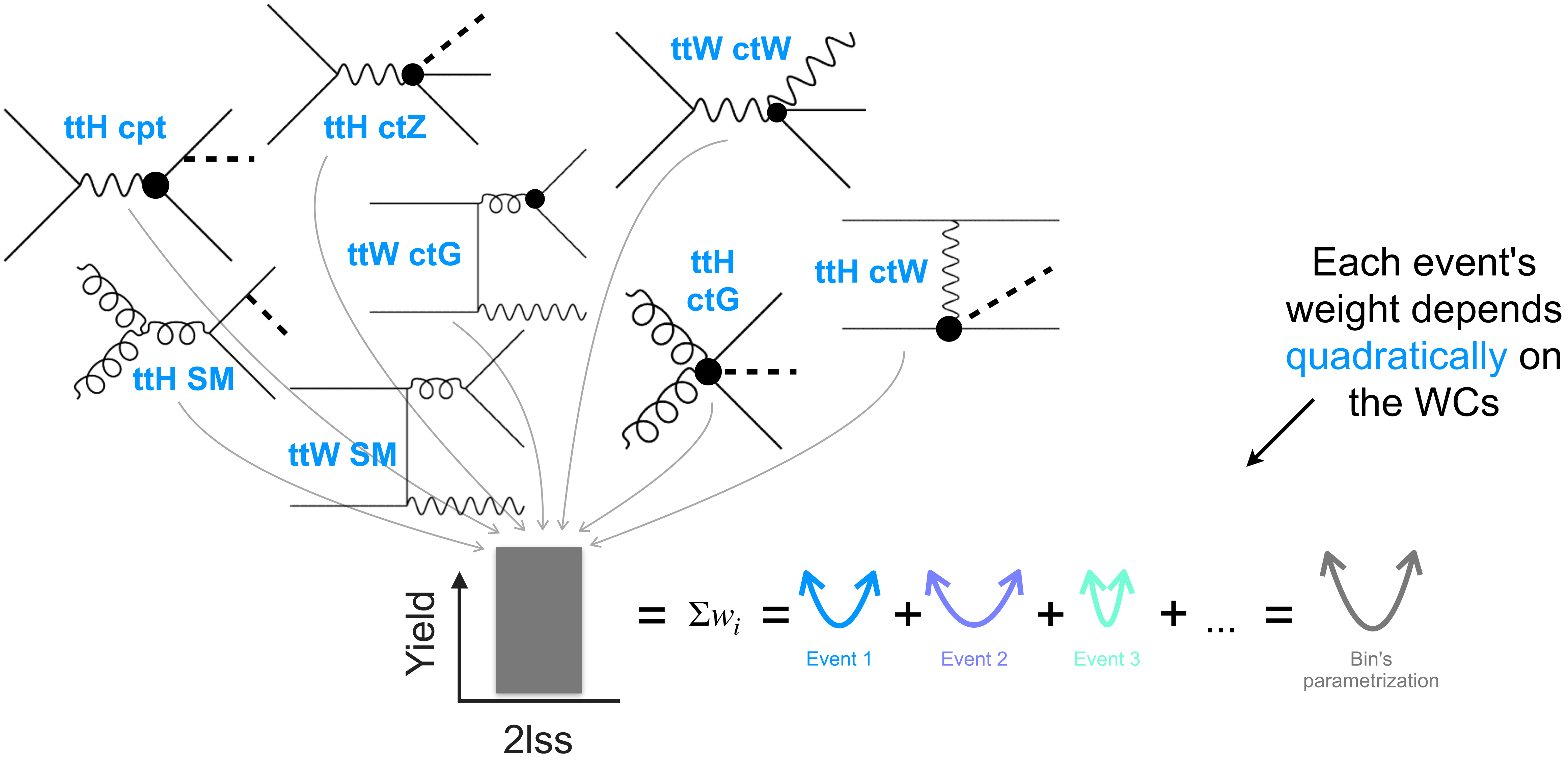
- We focus on multilepton signatures, many advantages but also challenges:
  - Multiple signal processes contribute to the same final state signatures
  - Many WCs can affect the processes, interfere with each other and the SM



Want to know number of predicted events in a given detector-level observable bin, as a function all of the WCs, i.e. Yield = Yield(c_1, c_2, c_3, ...)

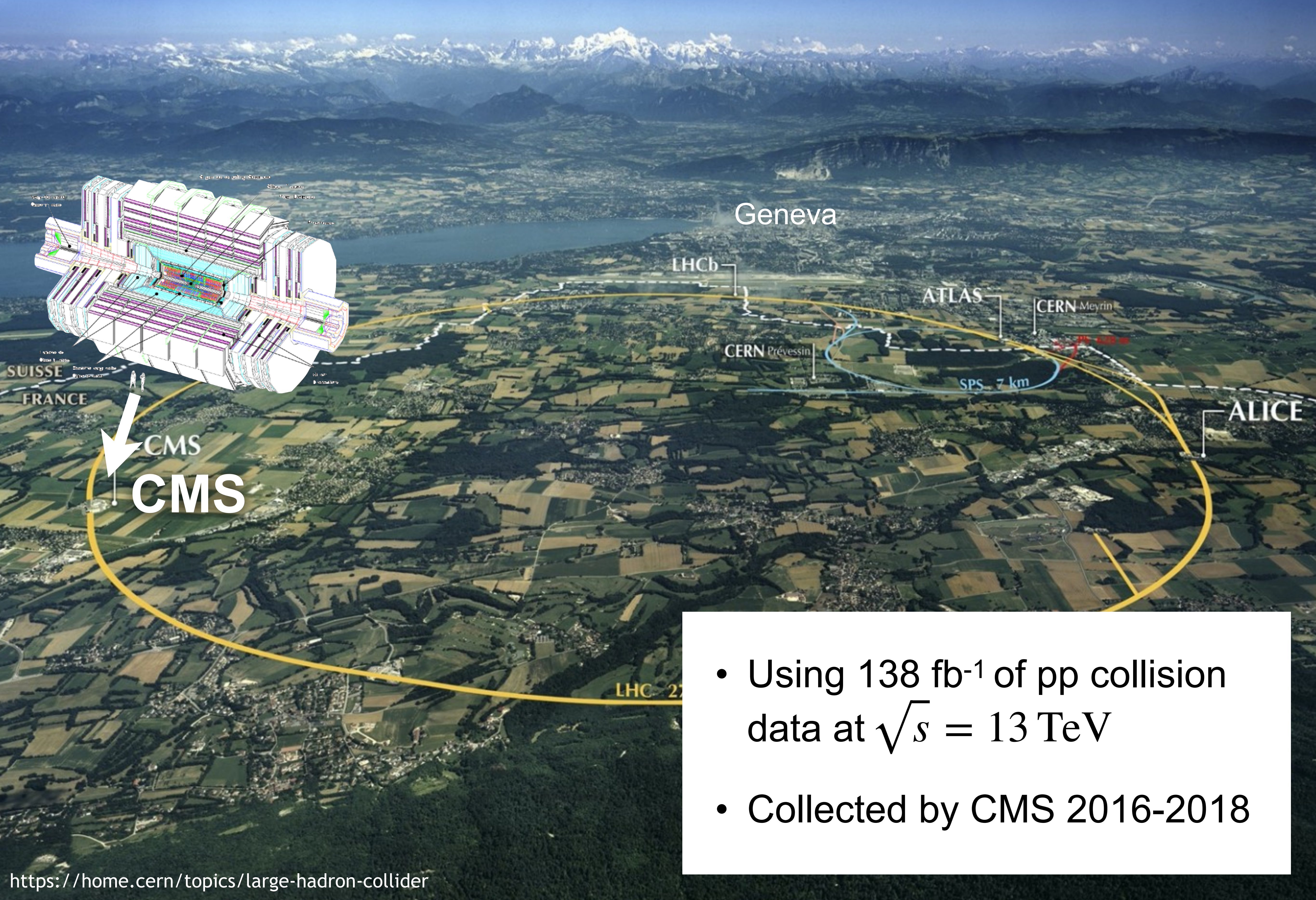
Strategy for multilepton EFT analysis

- We focus on multilepton signatures, many advantages but also challenges:
    - Multiple signal processes contribute to the same final state signatures
    - Many WCs can affect the processes, interfere with each other and the SM



Each event's weight depends quadratically on the WCs

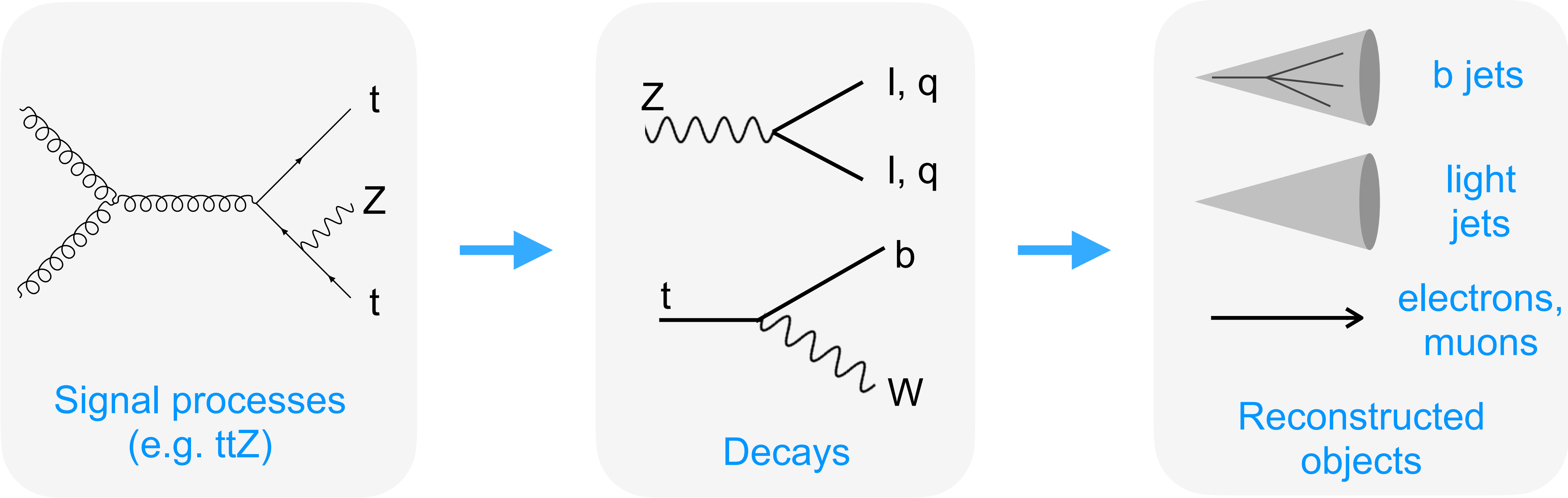
The LHC and CMS



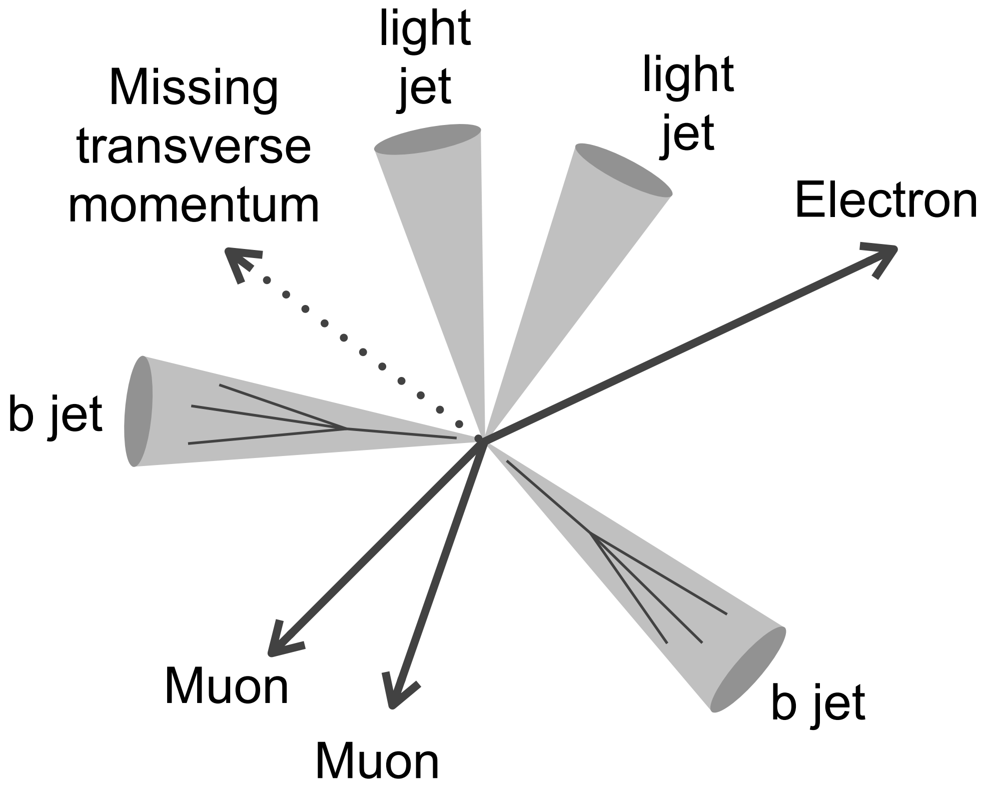
- Using 138 fb^{-1} of pp collision data at $\sqrt{s} = 13\,\mathrm{TeV}$
- Collected by CMS 2016-2018

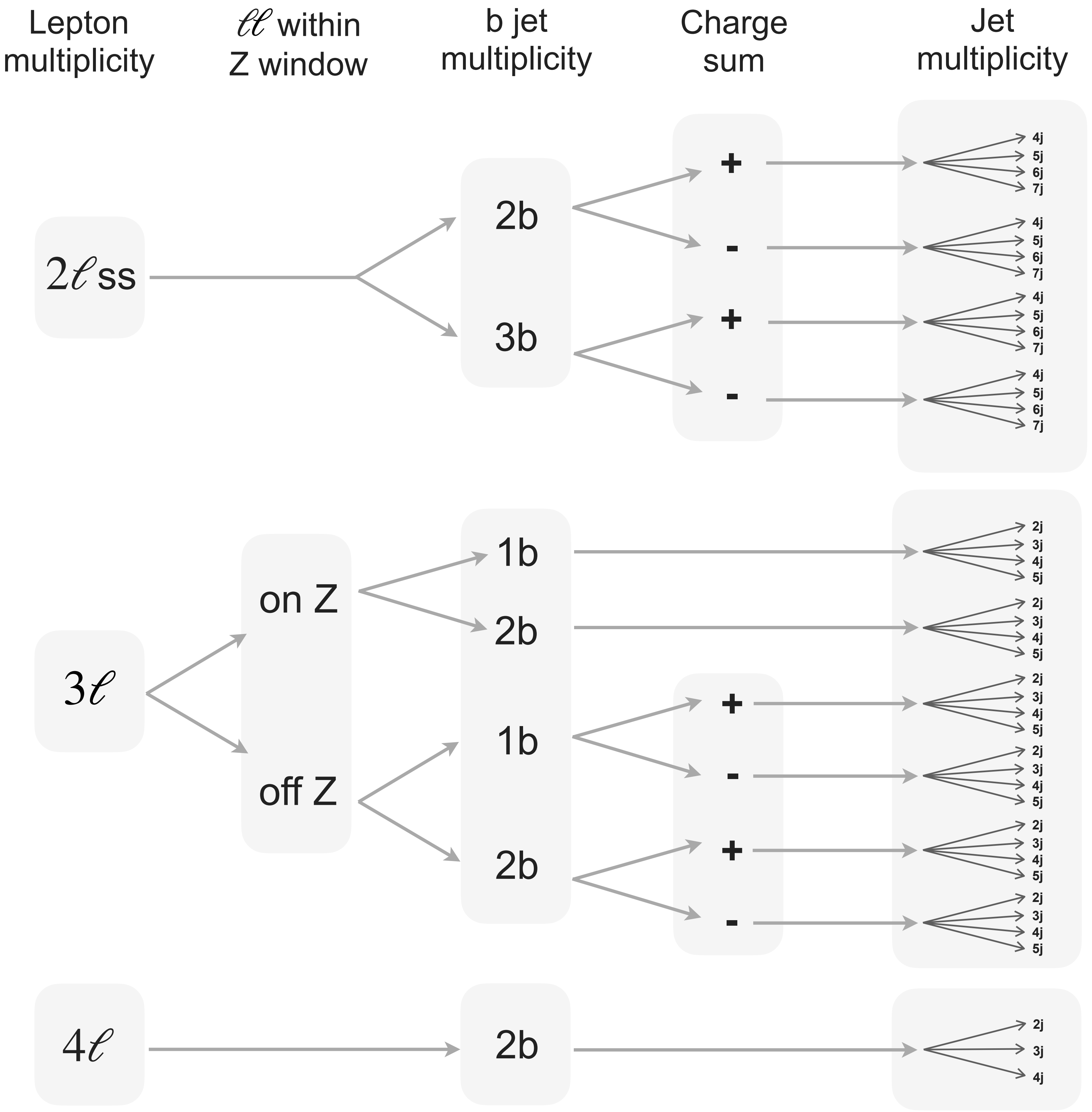
https://home.cern/topics/large-hadron-collider

Experimental signatures



- We're interested in leptonic decays of associated top processes
- These lead to signatures of leptons, jets, and b jets





Lepton multiplicity

$\ell\ell$ within Z window

b jet multiplicity

Charge sum

Jet multiplicity

2ℓss

2b

+

4j
5j
6j
7j

-

4j
5j
6j
7j

3b

+

4j
5j
6j
7j

-

4j
5j
6j
7j

3ℓ

on Z

1b

2j
3j
4j
5j

2b

2j
3j
4j
5j

off Z

1b

+

2j
3j
4j
5j

-

2j
3j
4j
5j

2b

+

2j
3j
4j
5j

-

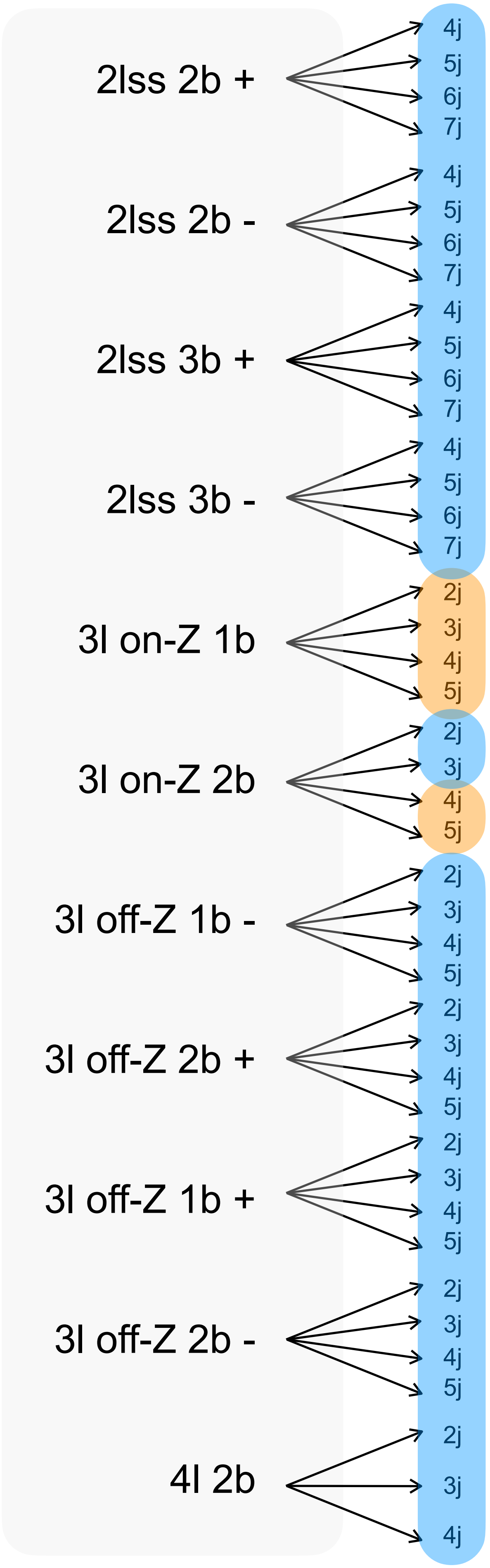
2j
3j
4j
5j

4ℓ

2b

2j
3j
4j

Improving sensitivity with differential distributions



2lss 2b + → 4j, 5j, 6j, 7j

2lss 2b - → 4j, 5j, 6j, 7j

2lss 3b + → 4j, 5j, 6j, 7j

2lss 3b - → 4j, 5j, 6j, 7j

3l on-Z 1b → 2j, 3j, 4j, 5j

3l on-Z 2b → 2j, 3j, 4j, 5j

3l off-Z 1b - → 2j, 3j, 4j, 5j

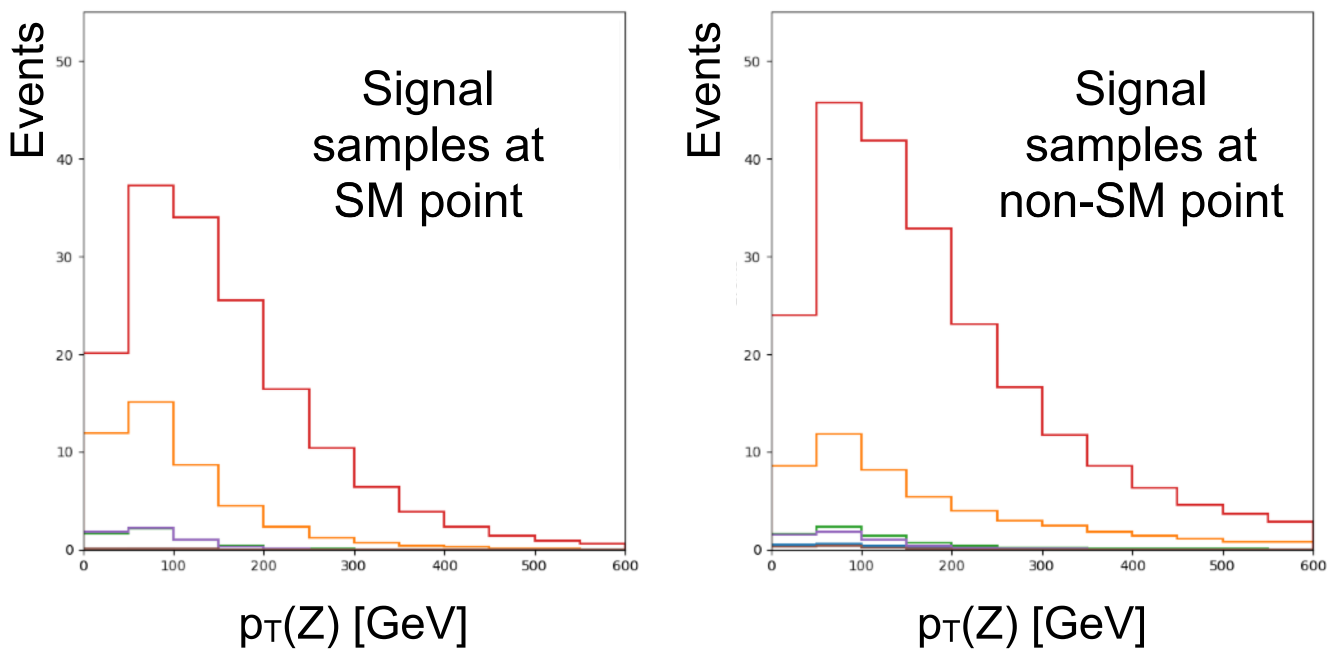
3l off-Z 2b + → 2j, 3j, 4j, 5j

3l off-Z 1b + → 2j, 3j, 4j, 5j

3l off-Z 2b - → 2j, 3j, 4j, 5j

4l 2b → 2j, 3j, 4j

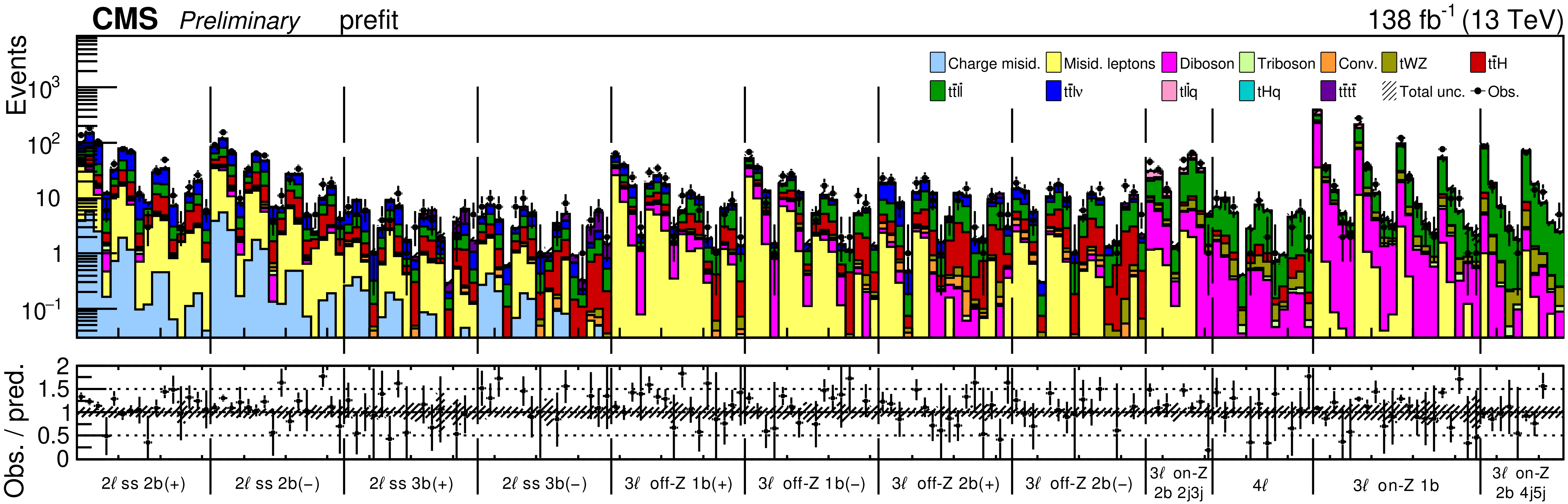
- In order to improve sensitivity, we fit a differential kinematic distribution for each of the 43 categories
- Use different variables (**p_T(lj0)**, **p_T(Z)**) in different regions to optimize sensitivity to EFT effects



When we reweight to a non-SM point, we can see the shape and normalization of the distribution changes

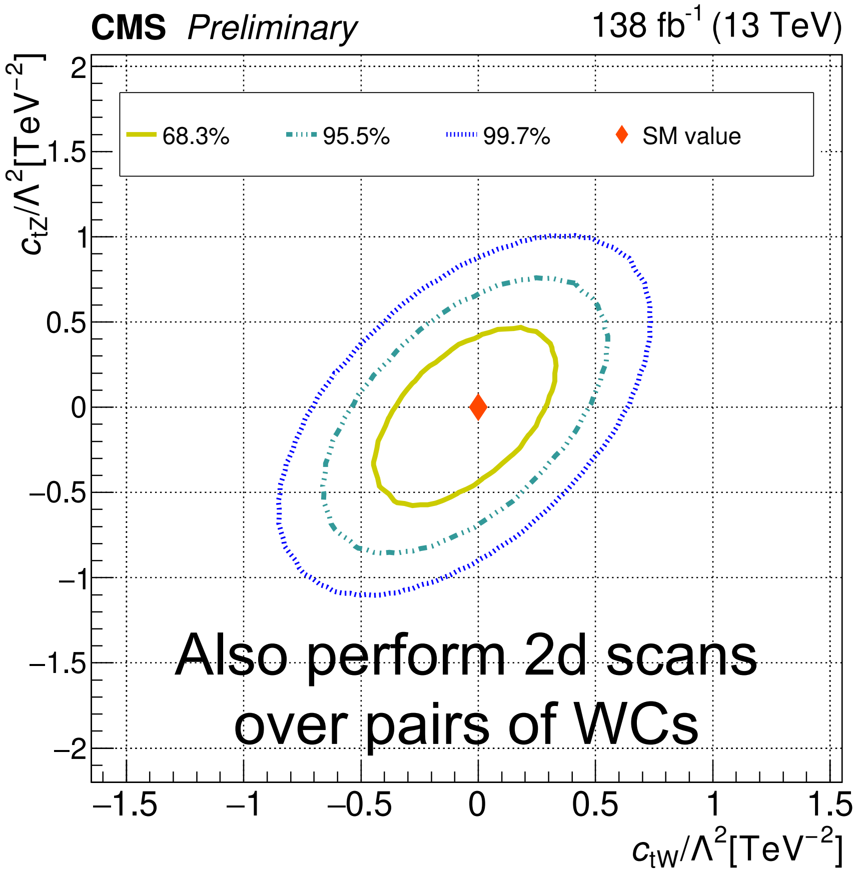
Summary of event selection and categorization

- Binning the 43 categories with these kinematical distributions → 178 total bins
- The predicted yield in each bin depends quadratically on the 26 WCs
- Perform a likelihood fit (where the WCs are the POIs) to extract the confidence intervals for WCs
    - Backgrounds (mainly dibosons and misidentified leptons) also contribute
    - Systematic uncertainties also must be accounted for

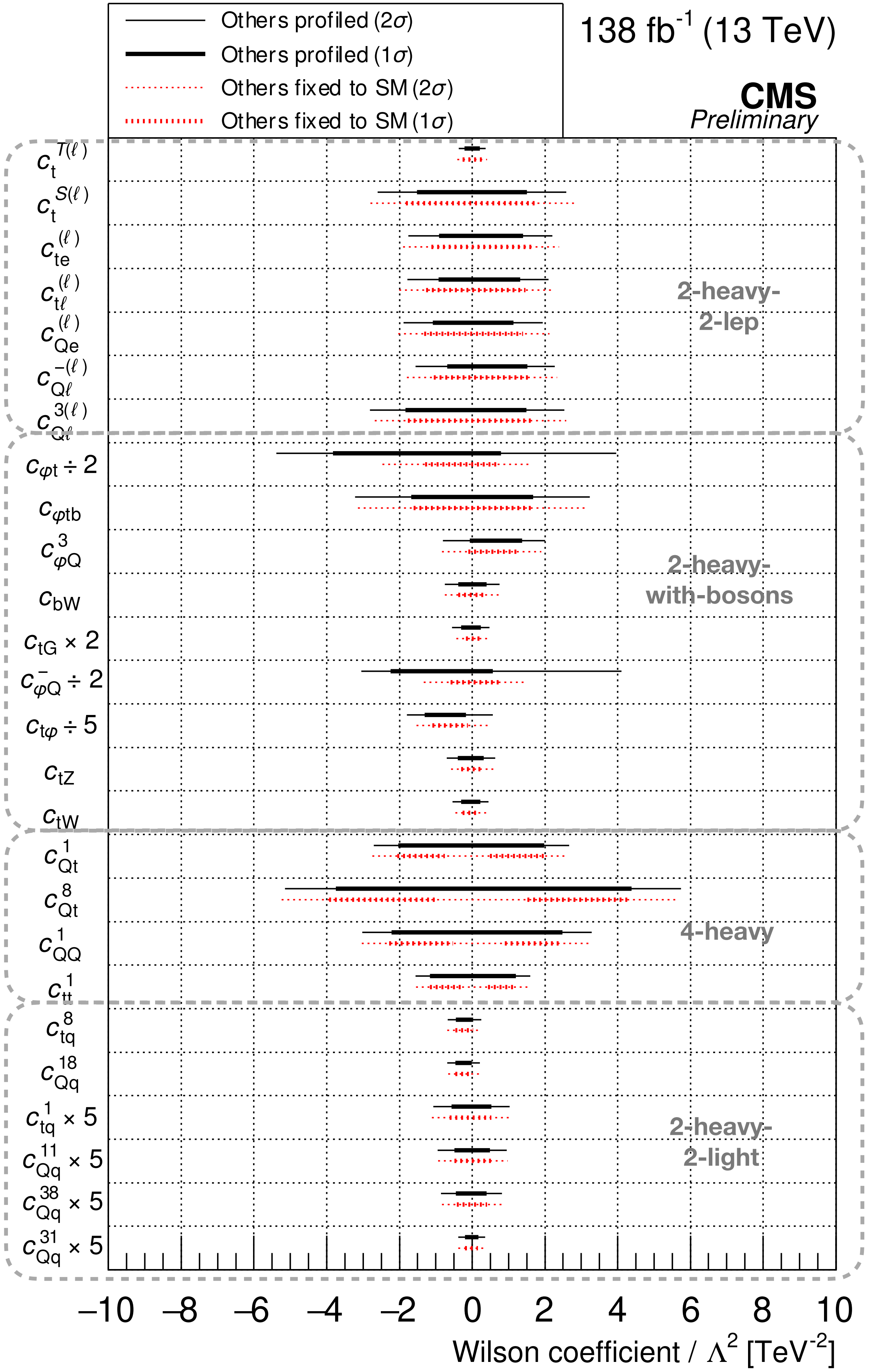


Results

- Extract the 1σ and 2σ confidence intervals for the WCs where other WCs are frozen and **profiled**
- Results are consistent with SM
- For most of the WCs, sensitivity is limited by statistics



Also perform 2d scans over pairs of WCs



Summary and future directions

- SM EFT provides a systematic framework in which to characterize the effects of heavy new physics and potentially extend the discovery reach of the LHC

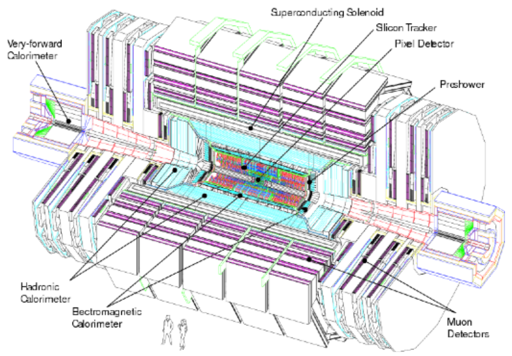
- This analysis has searched for new physics impacting associated top production in multilepton final states within the context of EFT
    - Used 138 fb^{-1} of pp collisions collected by CMS during 2016-2018
    - Studied 26 dimension-six WCs
    - Performed simultaneous fit to extracted confidence intervals for the WCs
    - All results are consistent with the SM
    - Details in CMS PAS TOP-22-006

- There are many directions in which to improve and expanded the analysis:
    - More data
    - Improvements in EFT modeling
    - Optimizations of categorizations and kinematic variables
    - Targeting more signal processes and other final states
    - Combinations with other analyses

Backup

CMS trigger

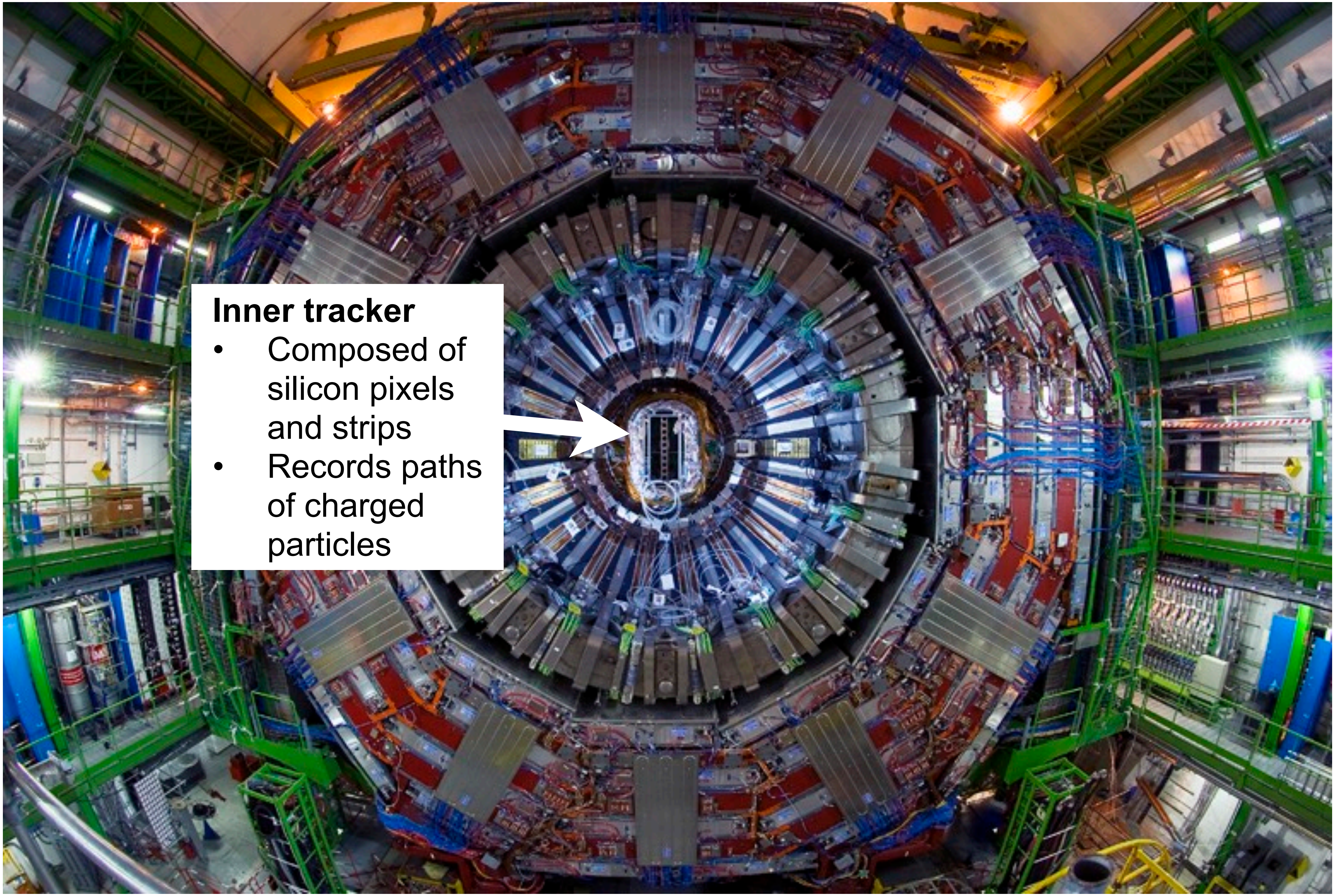
- Bunch crossings ~40MHz, → too much data to record/store
- Purpose of trigger: Reduce event rate to manageable ~1kHz while keeping as many potentially interesting events as possible



Bunch crossings in CMS → 40 MHz → L1 trigger → 100 kHz → High level trigger → 1 kHz → Offline reconstruction and analysis

Important to monitor the rates of the L1 and HLT paths, unexpected rates can be a useful early warning sign of issues in various parts of the detector

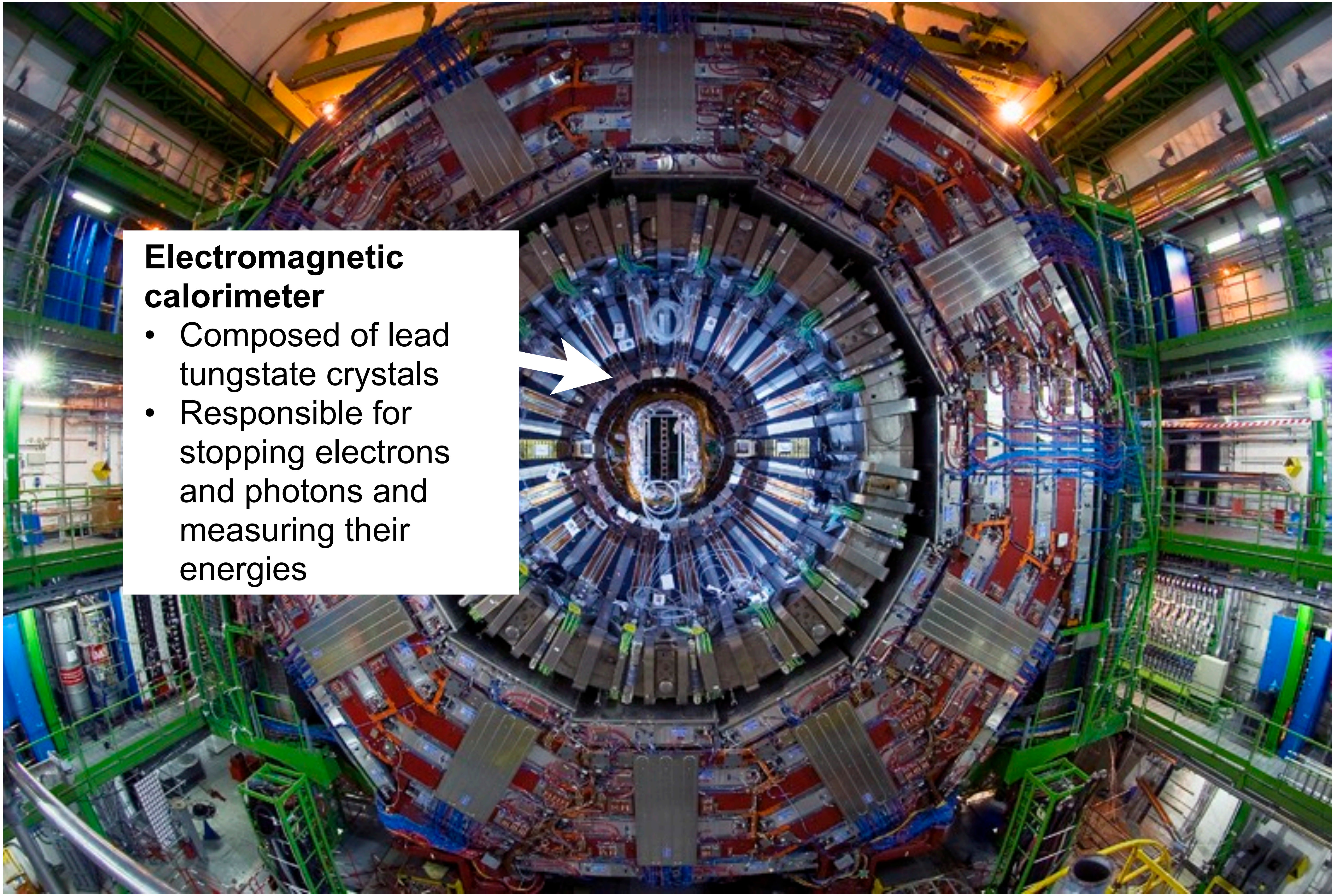
The subdetectors of CMS



Inner tracker

- Composed of silicon pixels and strips
- Records paths of charged particles

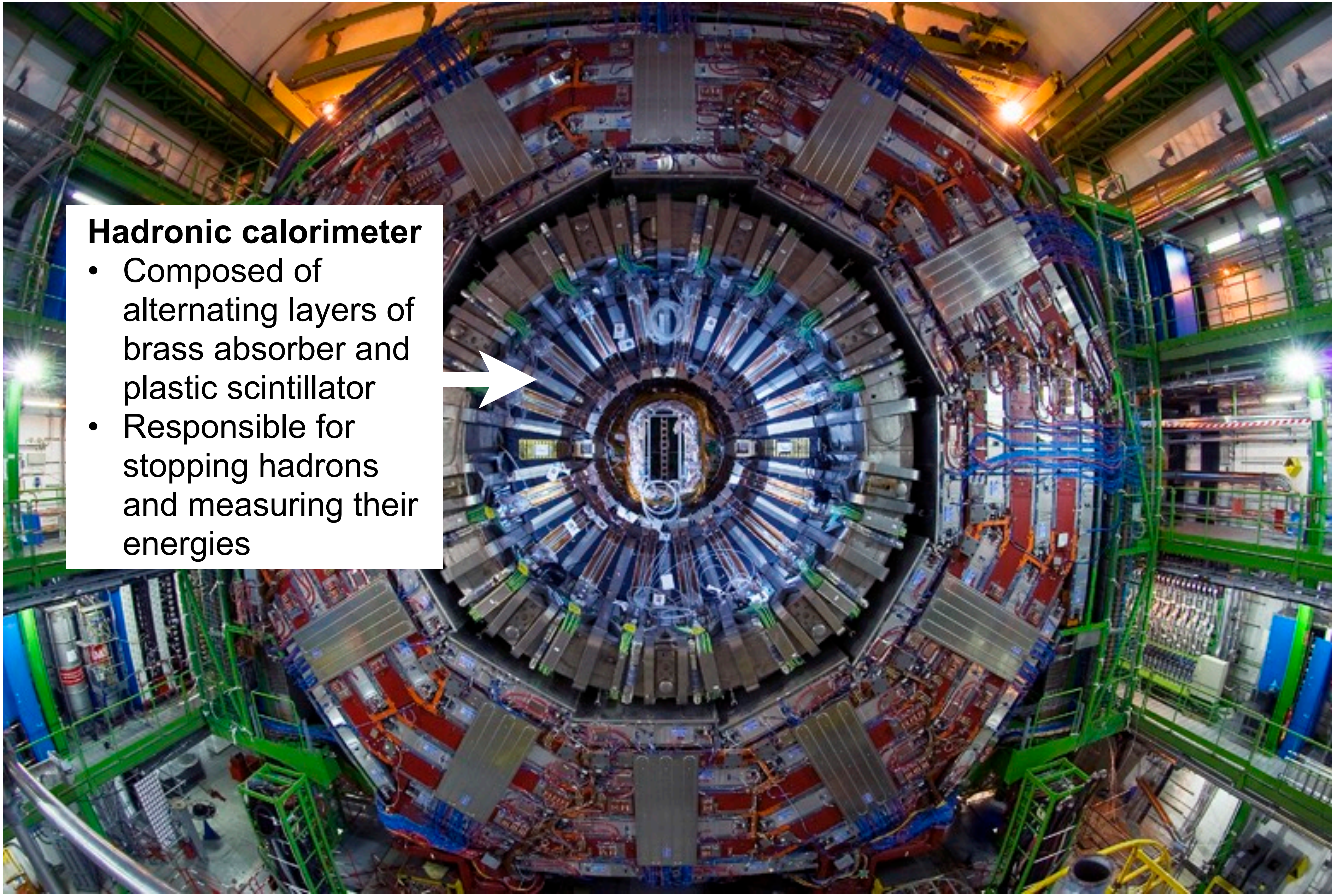
The subdetectors of CMS



Electromagnetic calorimeter

- Composed of lead tungstate crystals
- Responsible for stopping electrons and photons and measuring their energies

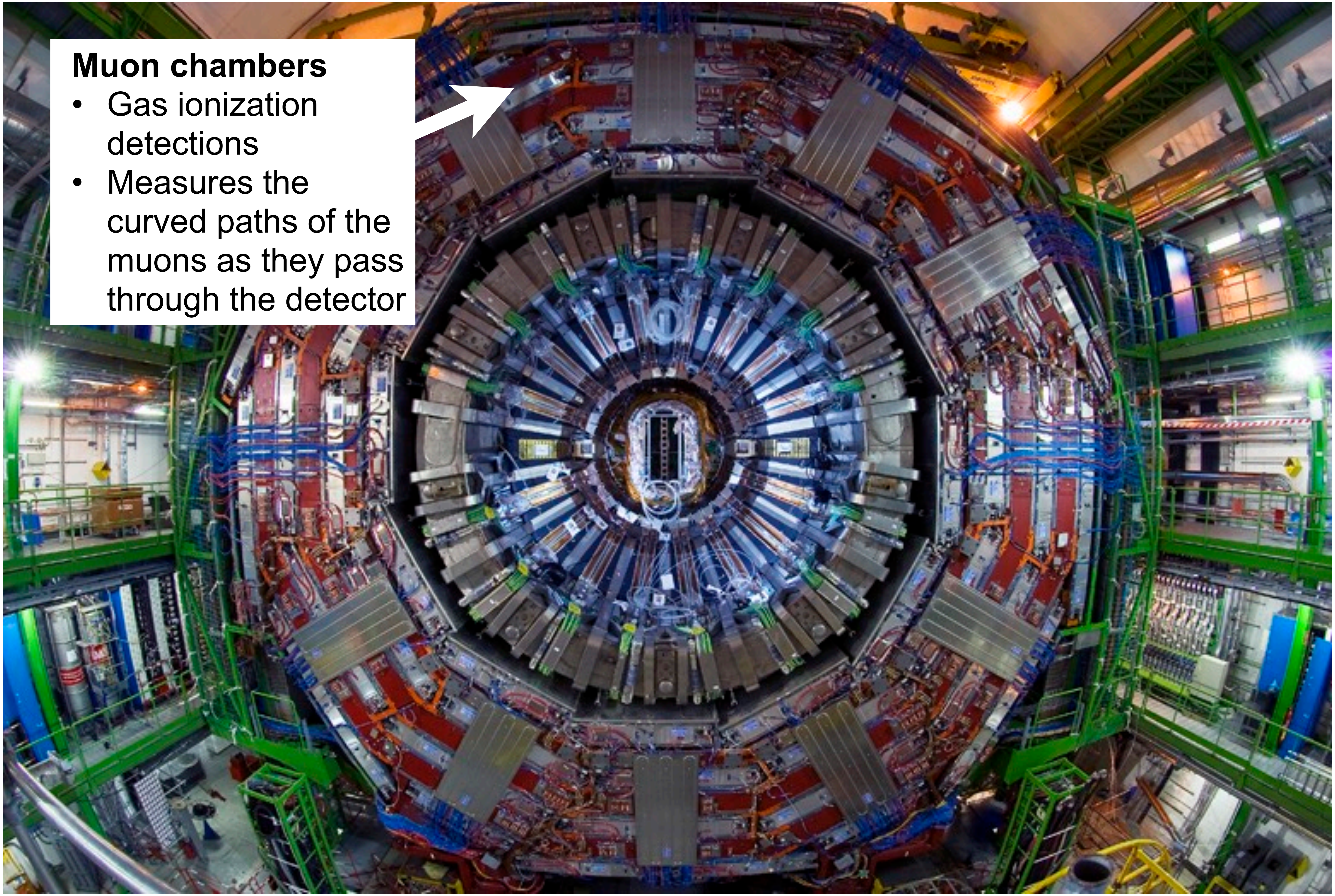
The subdetectors of CMS



Hadronic calorimeter

- Composed of alternating layers of brass absorber and plastic scintillator
- Responsible for stopping hadrons and measuring their energies

The subdetectors of CMS

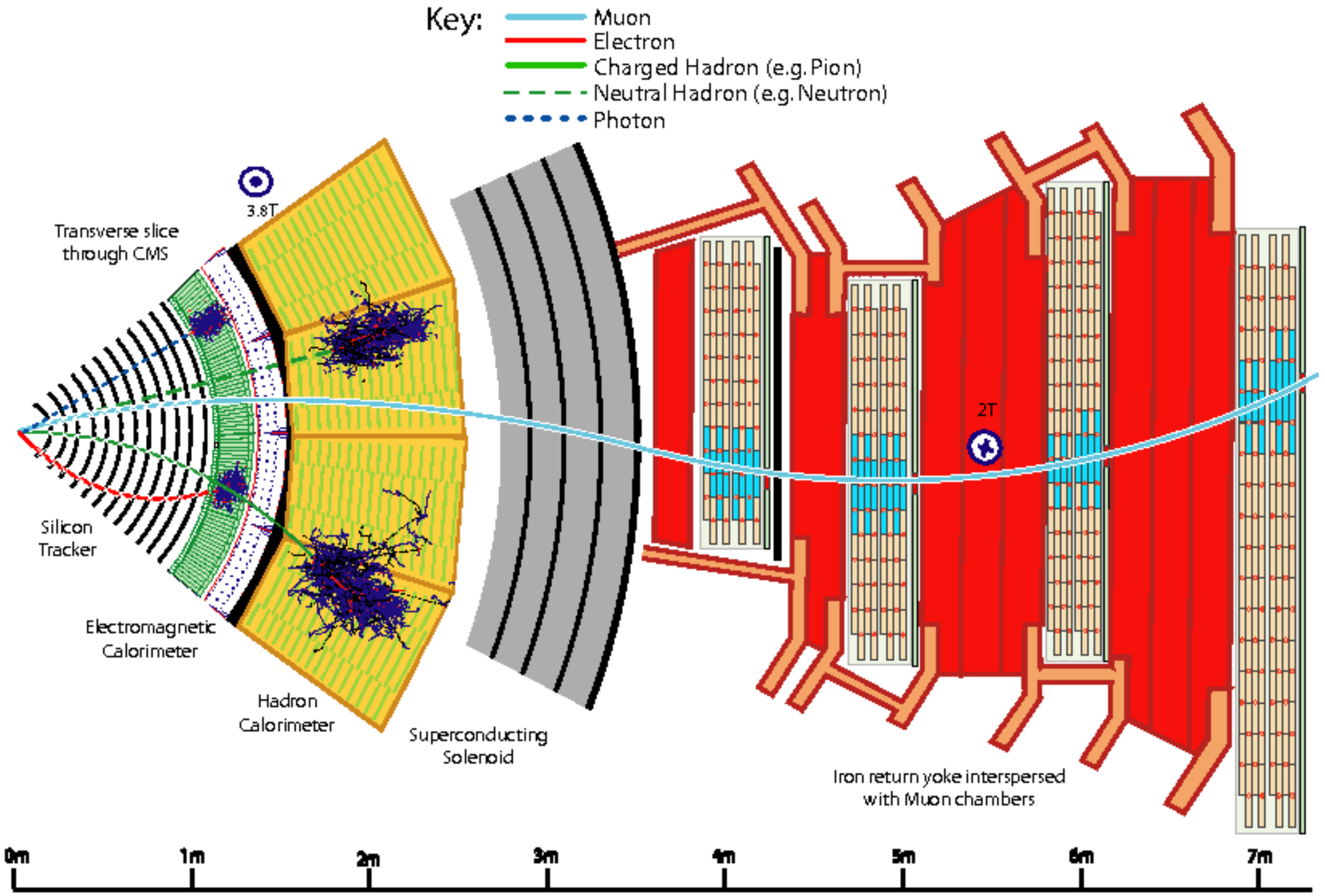


Muon chambers

- Gas ionization detections
- Measures the curved paths of the muons as they pass through the detector

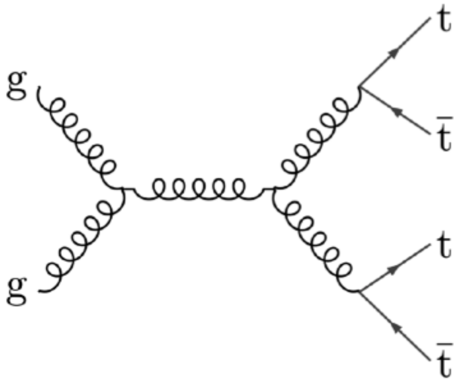
Object reconstruction

- CMS uses a holistic reconstruction technique called **particle flow** to correlate the elements from each subdetector and construct a global picture of each event
- First identifies **muons**, then **electrons** and **isolated photons**, finally **charged hadrons**, **neutral hadrons**, **non-isolated photons**

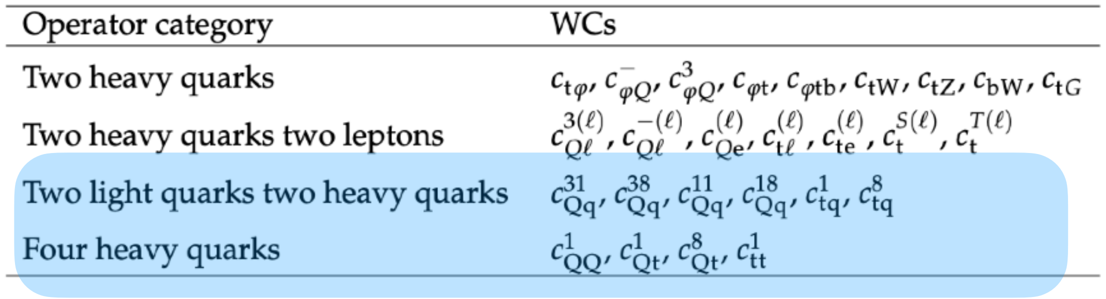


Comparison to TOP-19-001

- TOP-22-006 builds on the techniques and tools developed in the TOP-19-001 analysis
- Some of the important improvements over TOP-19-001 include:
    - Including an additional signal process (tttt) in addition to the 5 already included in TOP-19-001
    - Studying 10 additional Wilson coefficients in addition to the 16 probed in TOP-19-001, for a total of 26
    - Using the full Run 2 data set (TOP-19-001 only used 2017)
    - Fitting differential distributions, which allows us to gain additional sensitivity (TOP-19-001 followed a more inclusive approach, fitting 35 categories based primarily on object multiplicities)

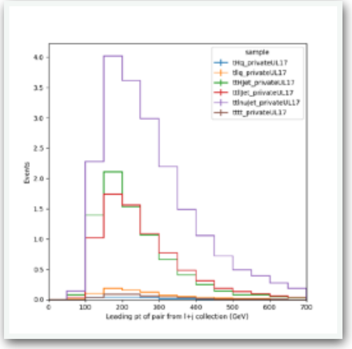


More process



Operator category	WCs
Two heavy quarks	$c_{t\varphi}, c^{-}_{\varphi Q}, c^{3}_{\varphi Q}, c_{\varphi t}, c_{\varphi tb}, c_{tW}, c_{tZ}, c_{bW}, c_{tG}$
Two heavy quarks two leptons	$c^{3(\ell)}_{Q\ell}, c^{-(\ell)}_{Q\ell}, c^{(\ell)}_{Qe}, c^{(\ell)}_{t\ell}, c^{(\ell)}_{te}, c^{S(\ell)}_{t}, c^{T(\ell)}_{t}$
Two light quarks two heavy quarks	$c^{31}_{Qq}, c^{38}_{Qq}, c^{11}_{Qq}, c^{18}_{Qq}, c^{1}_{tq}, c^{8}_{tq}$
Four heavy quarks	$c^{1}_{QQ}, c^{1}_{Qt}, c^{8}_{Qt}, c^{1}_{tt}$

More WCs



More differential

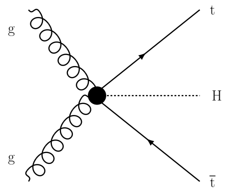
Modeling the signal contribution

- We generate MC samples for our six signal processes (ttH, ttlnu, ttll, tllq, tHq, tttt) using MG with the dim6top model (1802.07237) to incorporate the relevant EFT effects in the MC
    - The dim6top model uses Warsaw basis of dimension-6 operators
    - The dim6top model is LO, so we include an extra jet in the matrix element (when possible) to improve the modeling at high jet-multiplicities and to capture relevant EFT dependence that enters with an extra parton
    - We include 26 WCs (all WCs from dim6top that significantly impact the processes contributing to our data samples)

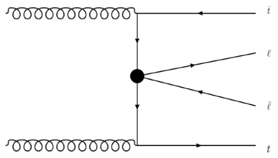
- We generated ~300M events in total

Operator category	WCs
Two heavy quarks	$c_{\mathrm{t}\varphi}$, $c^{-}_{\varphi Q}$, $c^{3}_{\varphi Q}$, $c_{\varphi \mathrm{t}}$, $c_{\varphi \mathrm{tb}}$, c_{tW}, c_{tZ}, c_{bW}, $c_{\mathrm{t}G}$
Two heavy quarks two leptons	$c^{3(\ell)}_{Q\ell}$, $c^{-(\ell)}_{Q\ell}$, $c^{(\ell)}_{Q\mathrm{e}}$, $c^{(\ell)}_{\mathrm{t}\ell}$, $c^{(\ell)}_{\mathrm{te}}$, $c^{S(\ell)}_{\mathrm{t}}$, $c^{T(\ell)}_{\mathrm{t}}$
Two light quarks two heavy quarks	c^{31}_{Qq}, c^{38}_{Qq}, c^{11}_{Qq}, c^{18}_{Qq}, c^{1}_{tq}, c^{8}_{tq}
Four heavy quarks	c^{1}_{QQ}, c^{1}_{Qt}, c^{8}_{Qt}, c^{1}_{tt}

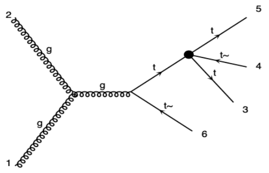
Operators involving heavy quarks and bosons



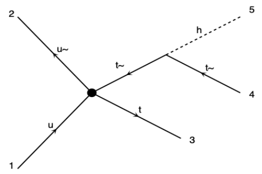
Operators involving 2 heavy quarks and 2 leptons



Operators involving 4 heavy quarks



Operators involving two light quarks and 2 heavy quarks



Quadratic dependence of the weights on the WCs

- Matrix element can be written as sum of SM and new physics:

$$\mathscr{M} = \mathscr{M}_{SM} + \sum_i c_i \mathscr{M}_i \longleftarrow$$

c_i are the Wilson coefficients

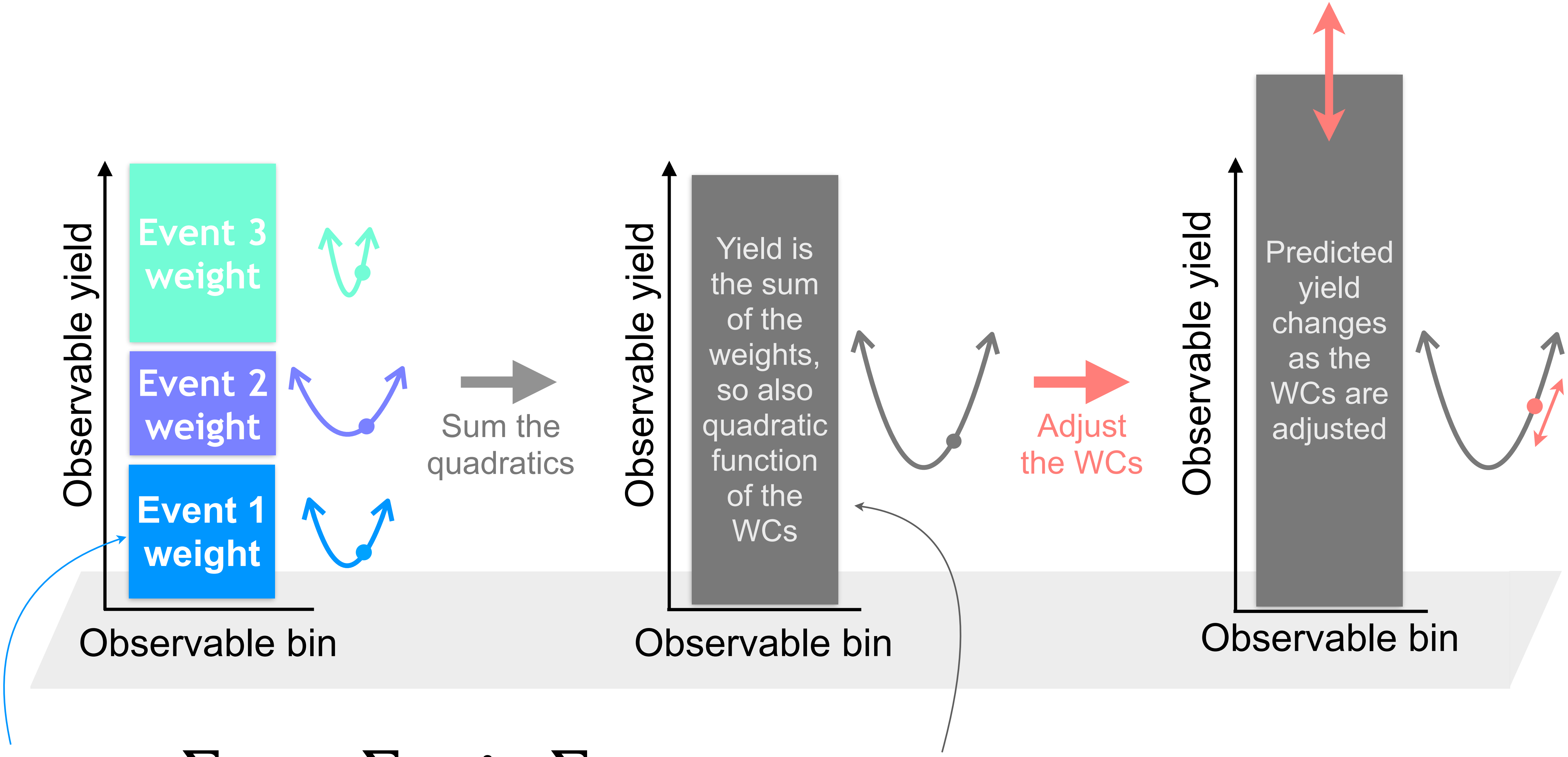
- Since, $\sigma \propto \mathscr{M}^2$, cross sections depend quadratically on the Wilson coefficients c_i
- E.g. for just one c_1:

$$d\sigma(c_1) \propto |\mathscr{M}_{SM} + c_1\mathscr{M}_1|^2 \propto \underbrace{s_0}_{\text{SM}} + \underbrace{s_1}_{\text{Interference with SM}} c_1 + \underbrace{s_2}_{\text{Pure NP}} c_1^2 \longleftarrow$$

In general this is an *n*-dimensional quadratic (*n* = number of WCs)

- Each event's weight will also depend quadratically on the WCs, which we can find via a reweighting procedure

Summary of how the parameterized weights are used to find the yield in a given bin



Sum the quadratics

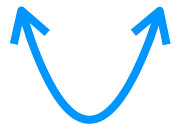
Yield is the sum of the weights, so also quadratic function of the WCs

Adjust the WCs

Predicted yield changes as the WCs are adjusted

$$w_i = s_{0i} + \sum_j s_{1ij}\, c_j + \sum_j s_{2ij}\, c_j^2 + \sum_{j \neq k} s_{3ijk}\, c_j c_k$$

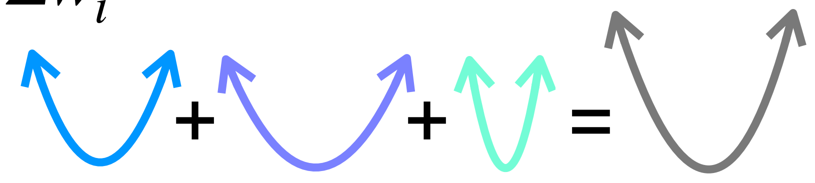
=



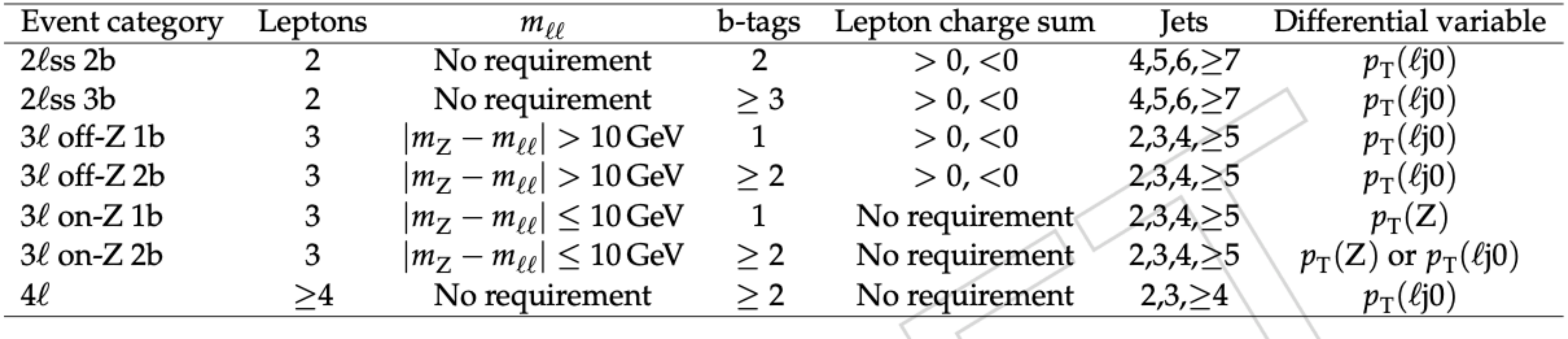
The structure constants of the n-dimensional quadric are determined with MG reweighting

$$\text{Yield} = \Sigma w_i$$

=



Event selection details



Event category	Leptons	$m_{\ell\ell}$	b-tags	Lepton charge sum	Jets	Differential variable
2ℓss 2b	2	No requirement	2	>0, <0	4,5,6,$\geq$7	$p_T(\ell j0)$
2ℓss 3b	2	No requirement	≥ 3	>0, <0	4,5,6,$\geq$7	$p_T(\ell j0)$
3ℓ off-Z 1b	3	$\lvert m_Z - m_{\ell\ell}\rvert > 10\,\mathrm{GeV}$	1	>0, <0	2,3,4,$\geq$5	$p_T(\ell j0)$
3ℓ off-Z 2b	3	$\lvert m_Z - m_{\ell\ell}\rvert > 10\,\mathrm{GeV}$	≥ 2	>0, <0	2,3,4,$\geq$5	$p_T(\ell j0)$
3ℓ on-Z 1b	3	$\lvert m_Z - m_{\ell\ell}\rvert \leq 10\,\mathrm{GeV}$	1	No requirement	2,3,4,$\geq$5	$p_T(Z)$
3ℓ on-Z 2b	3	$\lvert m_Z - m_{\ell\ell}\rvert \leq 10\,\mathrm{GeV}$	≥ 2	No requirement	2,3,4,$\geq$5	$p_T(Z)$ or $p_T(\ell j0)$
4ℓ	$\geq$4	No requirement	≥ 2	No requirement	2,3,$\geq$4	$p_T(\ell j0)$

- All jets required to have $|\eta| < 2.4$ and $p_T > 30$ GeV, all electrons require $|\eta| < 2.5$, all μ require $|\eta| < 2.4$, with lepton p_T cuts (in GeV):
  - 2lss 1st and 2nd: $p_T > 25$, $p_T > 15$
  - 3l 1st, 2nd, and 3rd: $p_T > 25$, $p_T > 15$, $p_T > 15$ (10) for e (μ)
  - 4l 1st, 2nd, 3rd, and 4th: $p_T > 25$, $p_T > 15$, $p_T > 15$ (10) for e (μ), $p_T > 15$ (10) for e (μ)

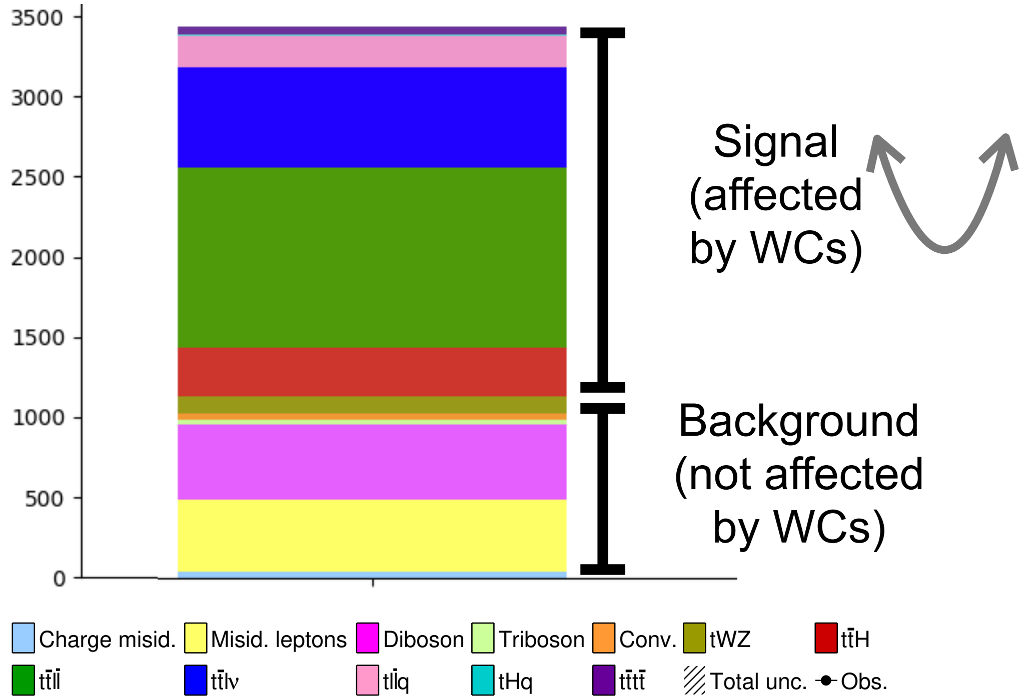
Backgrounds and systematics

- Signal processes (impacted by the EFT) are not the only contributions to our signal regions → about 1/3 of yield is background

- Main backgrounds: From processes that lead to the same final states as our signal processes, and from misidentified leptons

- Model backgrounds with combination of MC and data-driven approaches

Yield integrated over all signal regions



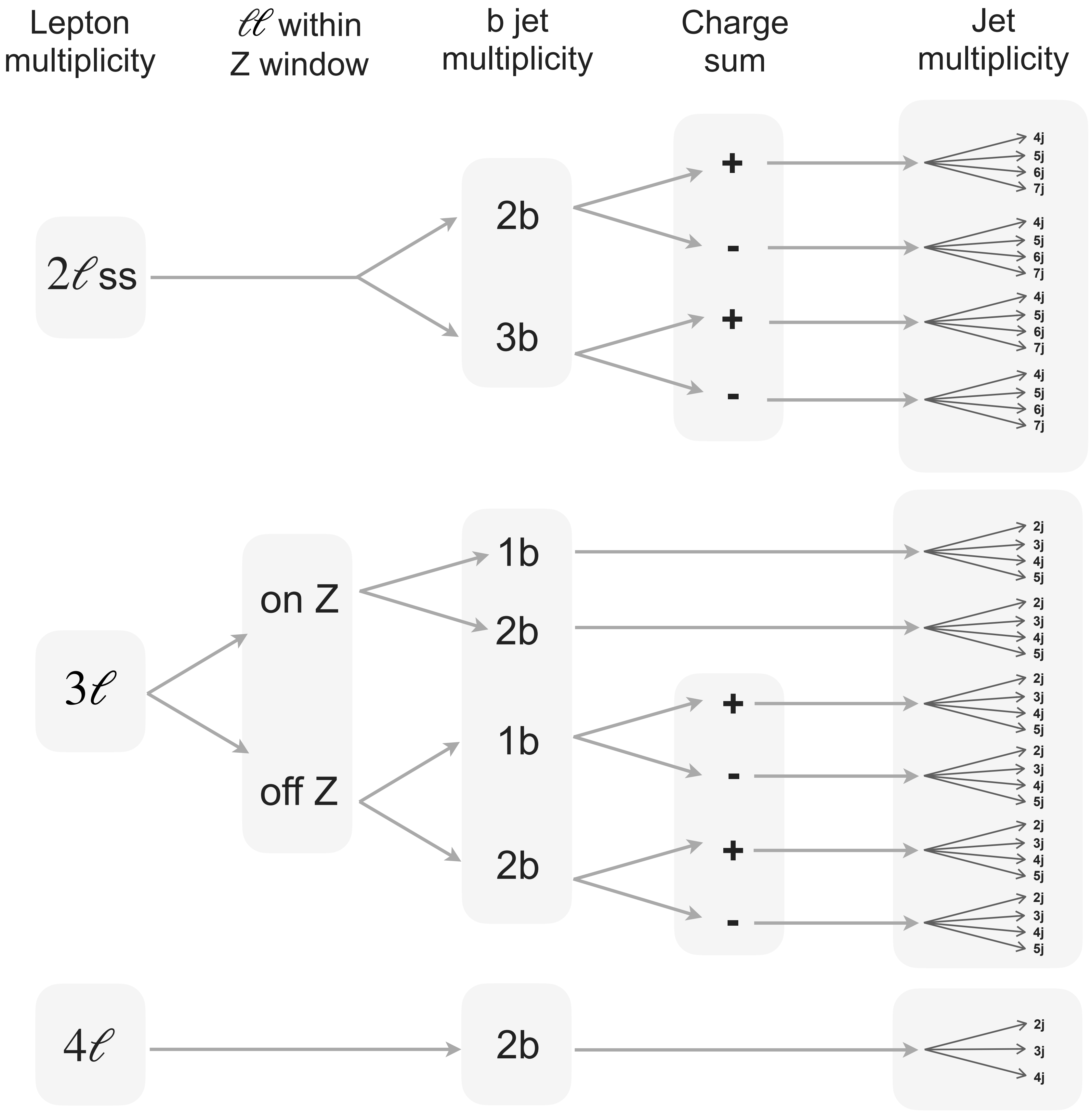
Signal (affected by WCs)

Background (not affected by WCs)

- Various systematic uncertainties (impacting signal and background) also must be accounted for in the fit

Event selection summary

- Cannot fully isolate signal processes
- But EFT impacts each process differently and goal is to gain sensitivity to the EFT effects, so the purpose of event selection categorization is to differentiate as much as possible between the admixture



Lepton multiplicity

$\ell\ell$ within Z window

b jet multiplicity

Charge sum

Jet multiplicity

2ℓss

2b
+
4j 5j 6j 7j
-
4j 5j 6j 7j

3b
+
4j 5j 6j 7j
-
4j 5j 6j 7j

3ℓ

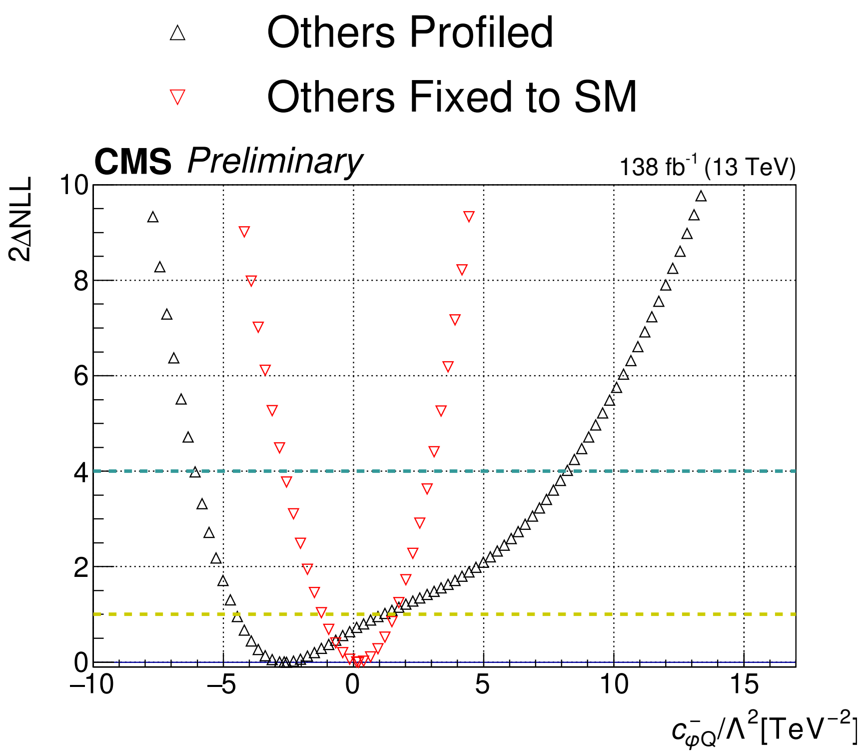
on Z
1b
2j 3j 4j 5j
2b
2j 3j 4j 5j

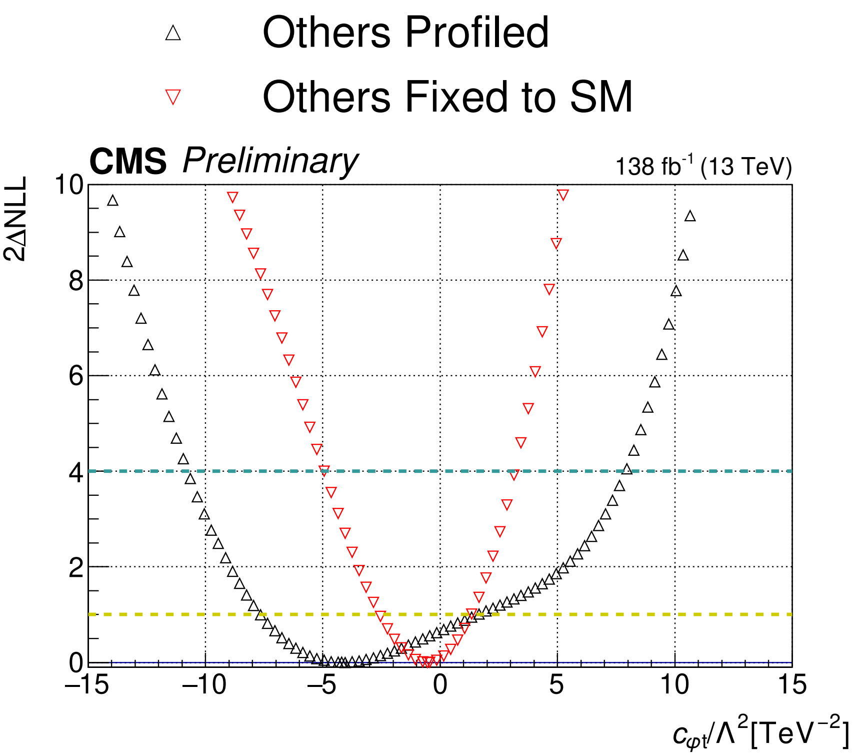
off Z
1b
+
2j 3j 4j 5j
-
2j 3j 4j 5j
2b
+
2j 3j 4j 5j
-
2j 3j 4j 5j

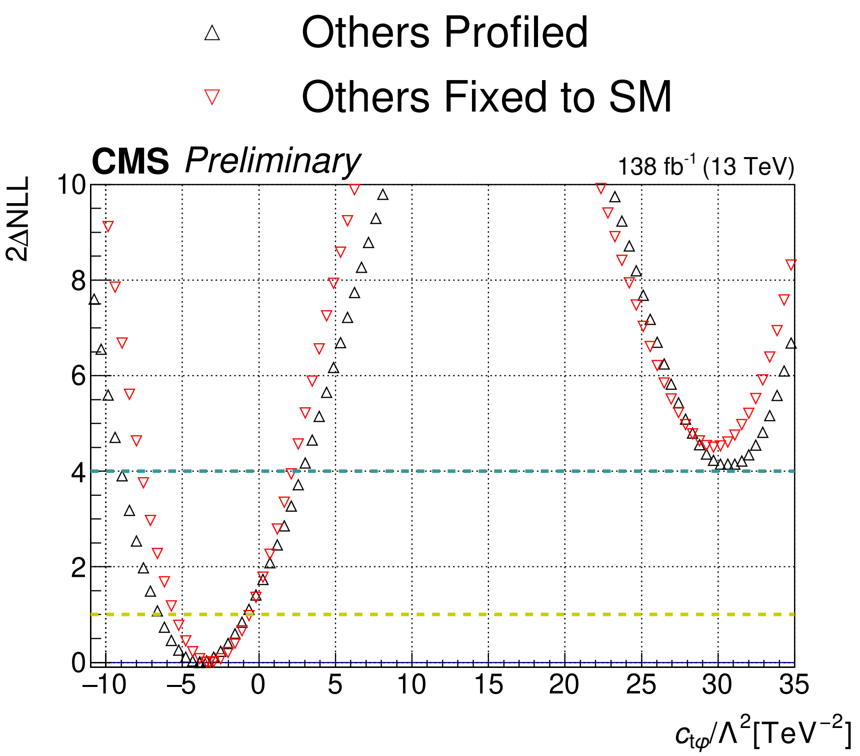
4ℓ

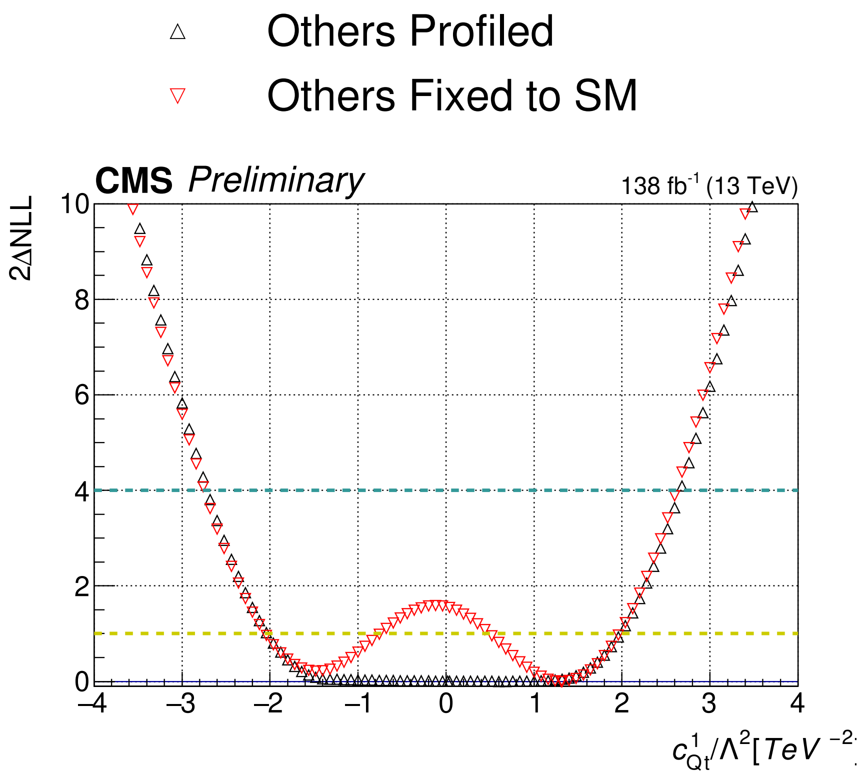
2b
2j 3j 4j

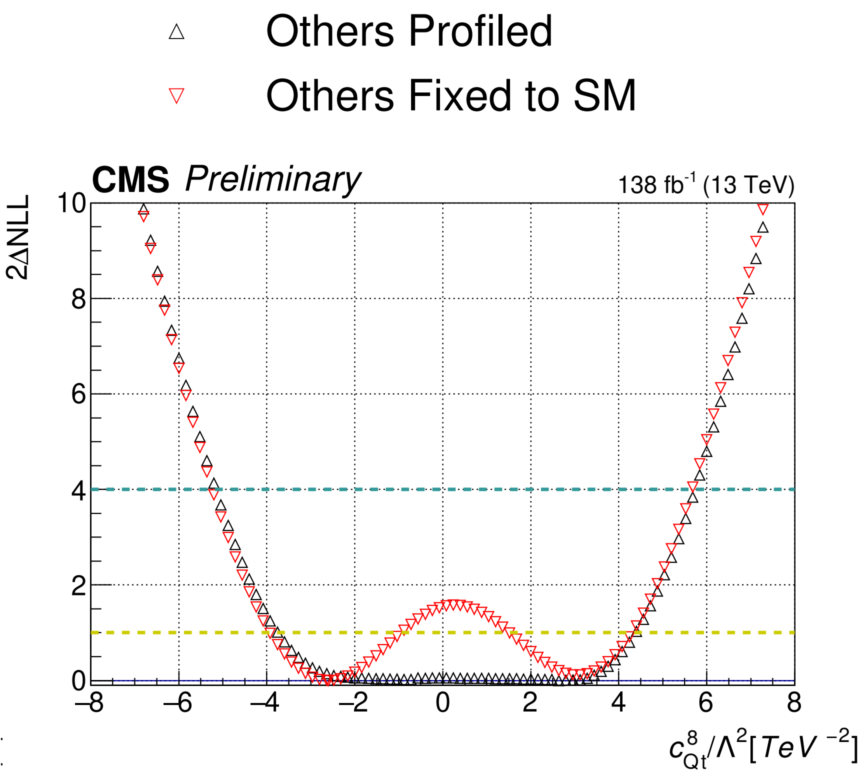
Example one-dimensional scans

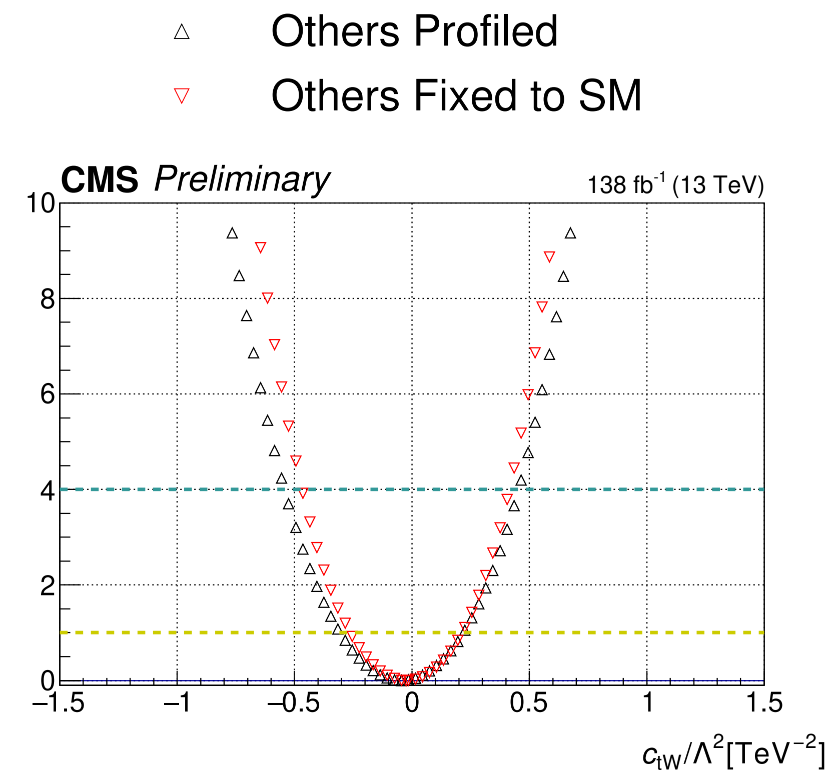






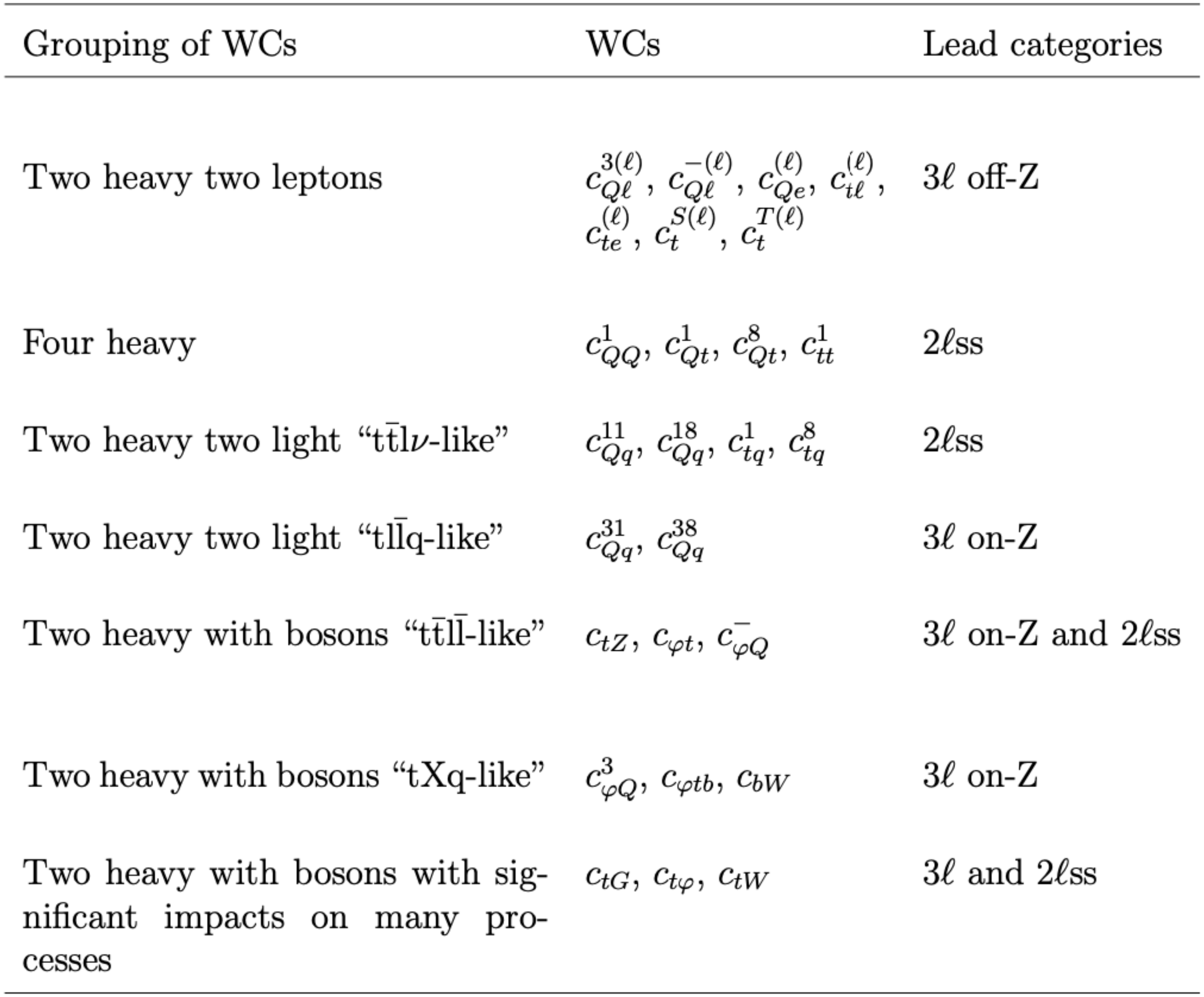






Discussion of results: Interpretation of sensitivity

- The sensitivity to most of the WCs comes from a wide range of bins across all selection categories



Grouping of WCs	WCs	Lead categories
Two heavy two leptons	$c_{Q\ell}^{3(\ell)}$, $c_{Q\ell}^{-(\ell)}$, $c_{Qe}^{(\ell)}$, $c_{t\ell}^{(\ell)}$, $c_{te}^{(\ell)}$, $c_{t}^{S(\ell)}$, $c_{t}^{T(\ell)}$	3ℓ off-Z
Four heavy	c_{QQ}^{1}, c_{Qt}^{1}, c_{Qt}^{8}, c_{tt}^{1}	2ℓss
Two heavy two light “t$\bar{\text{t}}$lν-like”	c_{Qq}^{11}, c_{Qq}^{18}, c_{tq}^{1}, c_{tq}^{8}	2ℓss
Two heavy two light “tl$\bar{\text{l}}$q-like”	c_{Qq}^{31}, c_{Qq}^{38}	3ℓ on-Z
Two heavy with bosons “t$\bar{\text{t}}$l$\bar{\text{l}}$-like”	c_{tZ}, $c_{\varphi t}$, $c_{\varphi Q}^{-}$	3ℓ on-Z and 2ℓss
Two heavy with bosons “tXq-like”	$c_{\varphi Q}^{3}$, $c_{\varphi tb}$, c_{bW}	3ℓ on-Z
Two heavy with bosons with significant impacts on many processes	c_{tG}, $c_{t\varphi}$, c_{tW}	3ℓ and 2ℓss