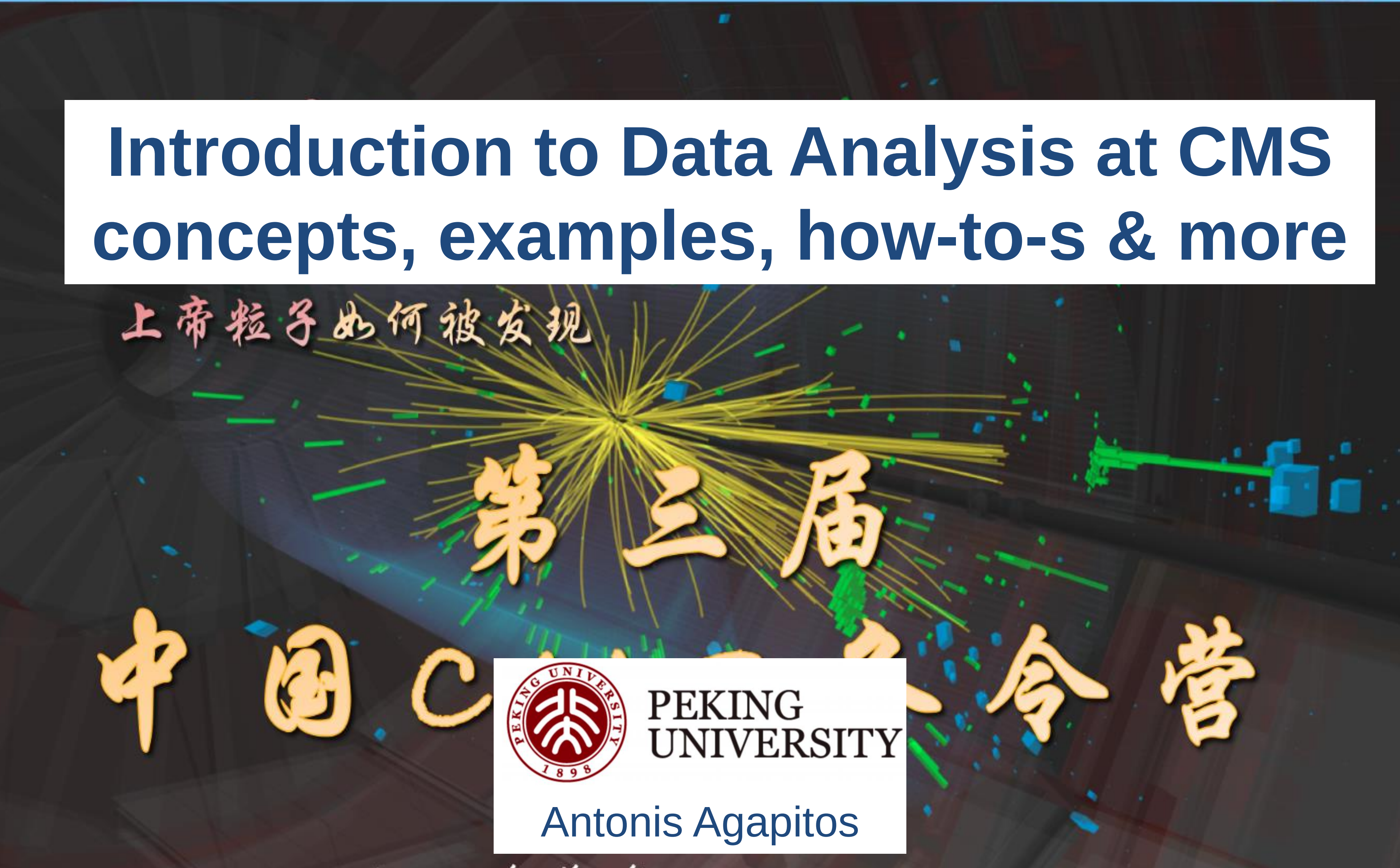
3rd CMS Winter Camp - January 19 2025

BEIHANG UNIVERSITY

Introduction to Data Analysis at CMS concepts, examples, how-to-s & more



PEKING UNIVERSITY

Antonis Agapitos

Outline: in this talk

1. Intro: CERN, LHC, beams, & collisions
2. Detector basics, physics objects, & Trigger
3. The Physics objects we use: e, μ, τ, jets, γ, “MET”
4. LHC collisions’ kinematics & variables
5. CMS PAGs & POGs, overview

6. What is data analysis? The concept of Signal & BKG.
7. Measurements & Searches
8. Resonances: examples
9. Production modes & BRs

Break

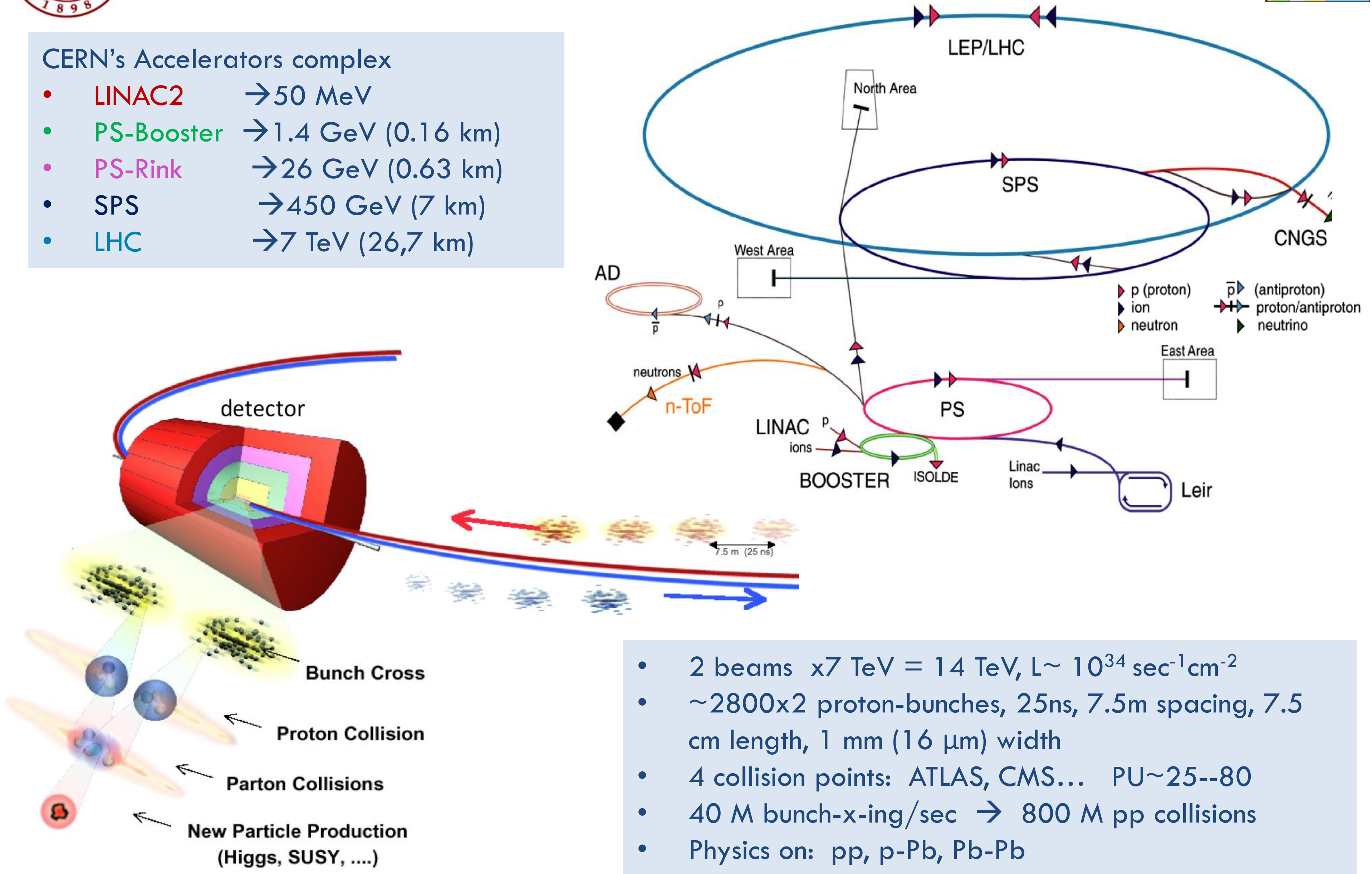
10. How to… (read plots, pulls, limits, ROCs, etc.)
11. Jets reconstruction, clustering, substructure, & tagging
12. The taggers we use (DeepAK8, Particle Net, Part. Transformer)

Analyses Examples:

1. Search with 4 jets in 2 pairs, prediction with parametric fit
2. Search in diphoton spectrum (reso & non-reso cases)
3. Search for resonant X→gWW in fully hadronic mode.
   - Selection, binning, data driven prediction, post-fit results, etc

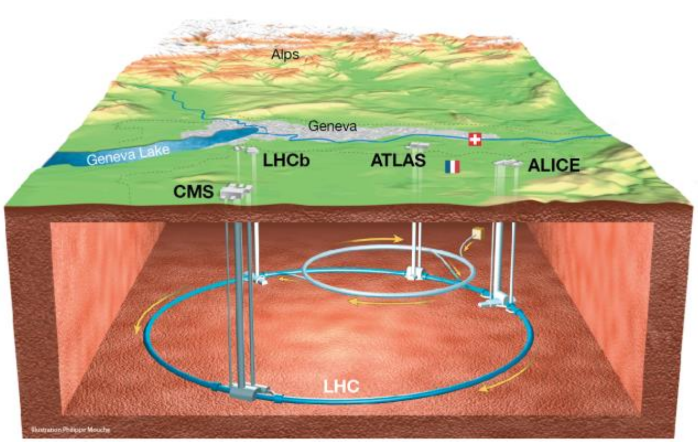
LHC acceleration, beans & collisions

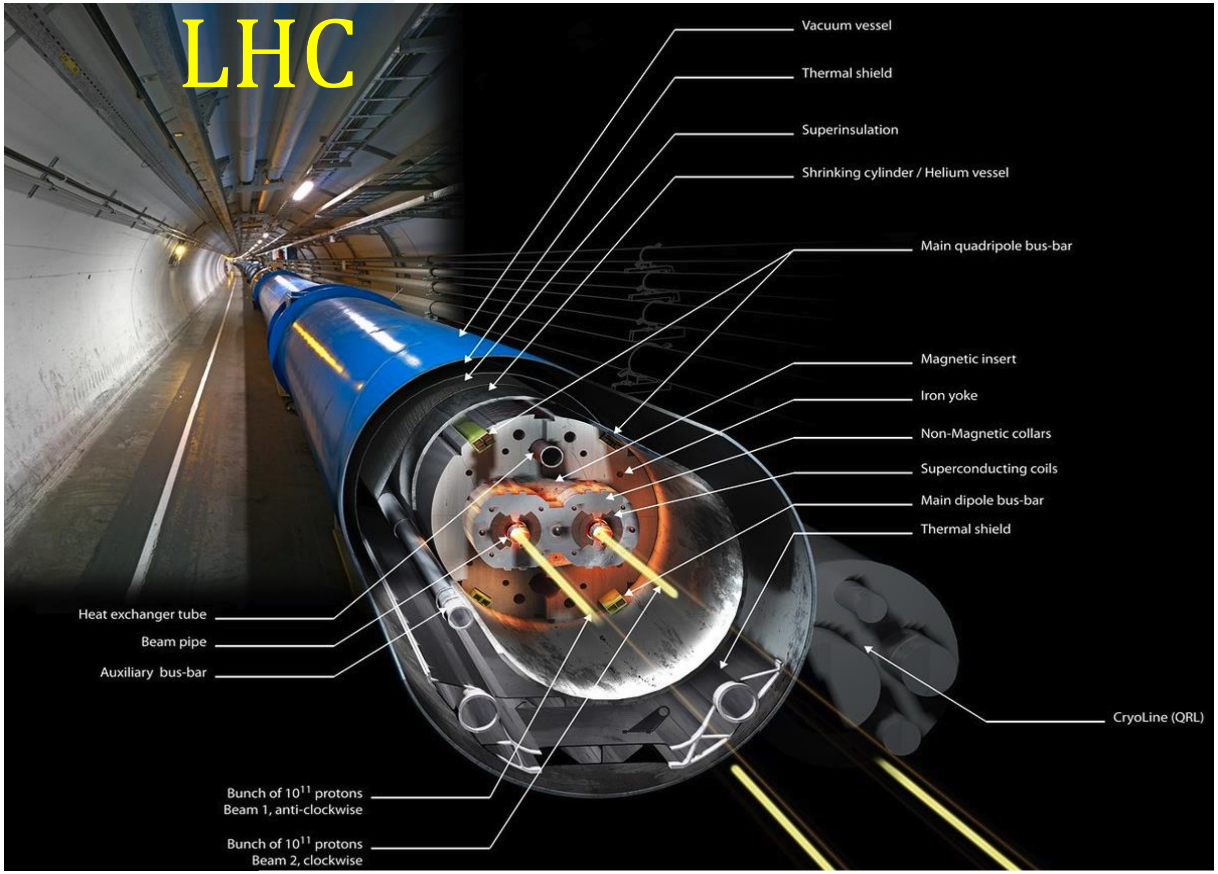
CERN's Accelerators complex
- LINAC2 →50 MeV
- PS-Booster →1.4 GeV (0.16 km)
- PS-Rink →26 GeV (0.63 km)
- SPS →450 GeV (7 km)
- LHC →7 TeV (26,7 km)



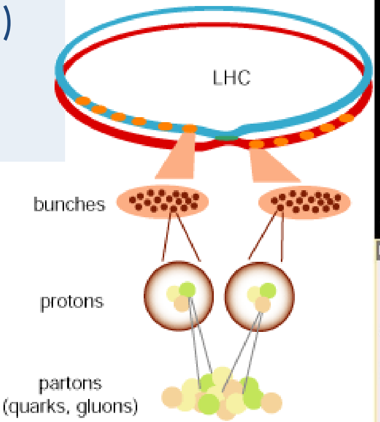
- 2 beams x7 TeV = 14 TeV, L~ 10^{34} $sec^{-1}cm^{-2}$
- ~2800x2 proton-bunches, 25ns, 7.5m spacing, 7.5 cm length, 1 mm (16 μm) width
- 4 collision points: ATLAS, CMS… PU~25--80
- 40 M bunch-x-ing/sec → 800 M pp collisions
- Physics on: pp, p-Pb, Pb-Pb

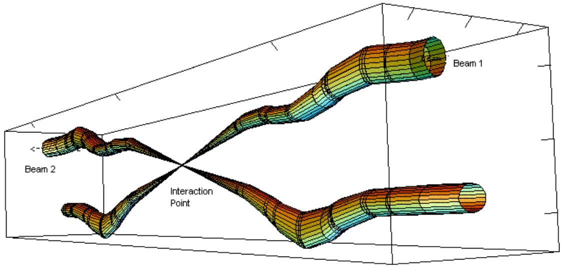
LHC



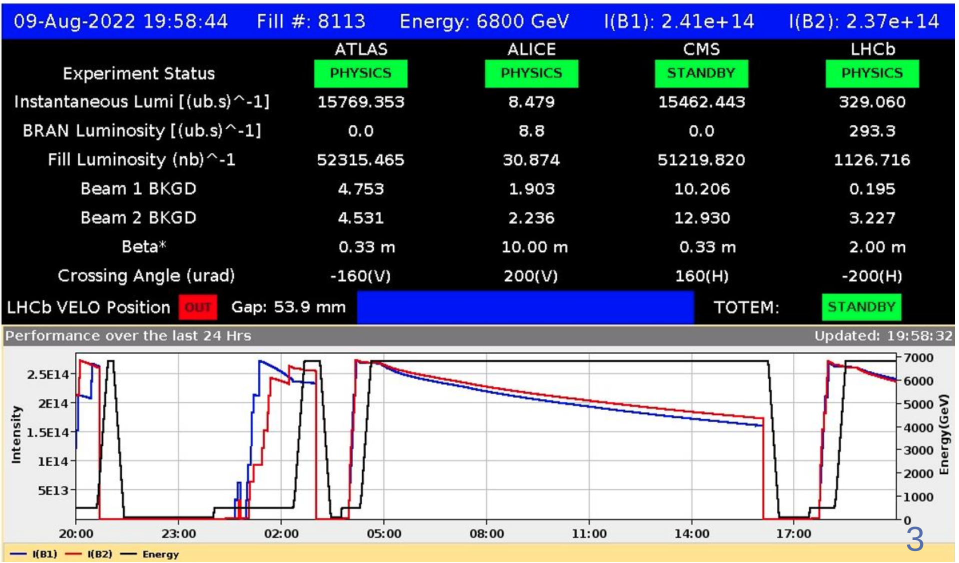


- 26.7 km
- 40-175 m bellow ground
- 1232 dipole magnets
- 392 quadrupoles
- 2x8 RF-cavities
- Revolution freq: 11.3MHz,
- B = 8.33 T ,
- I~12500 Ampere,
- NbTi superconductor
- T = 1.9 K = -271.3 C° (He)
- Vacuum ~100 nPa
- 4 collision p.→ 4 Exprms





Relative beam sizes around IP1 (Atlas) in collision



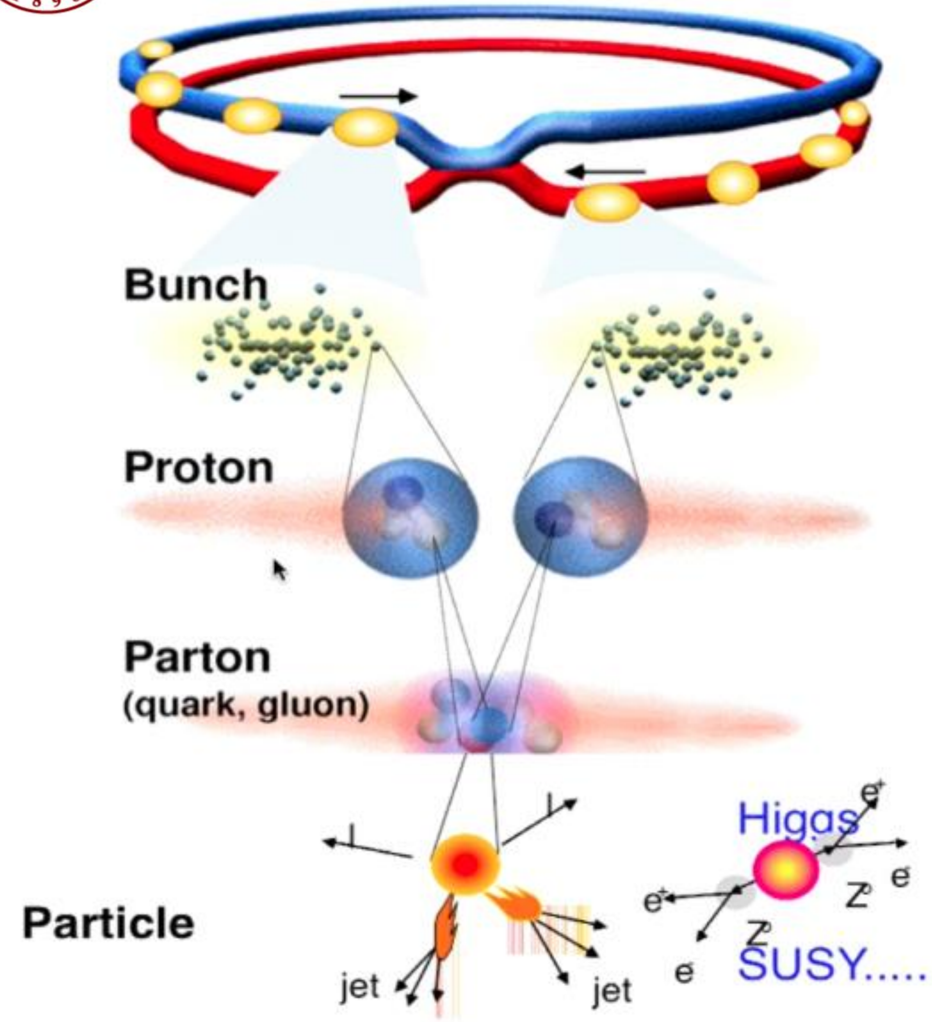
09-Aug-2022 19:58:44 Fill #: 8113 Energy: 6800 GeV I(B1): 2.41e+14 I(B2): 2.37e+14

	ATLAS	ALICE	CMS	LHCb
Experiment Status	PHYSICS	PHYSICS	STANDBY	PHYSICS
Instantaneous Lumi [(ub.s)^-1]	15769.353	8.479	15462.443	329.060
BRAN Luminosity [(ub.s)^-1]	0.0	8.8	0.0	293.3
Fill Luminosity (nb)^-1	52315.465	30.874	51219.820	1126.716
Beam 1 BKGD	4.753	1.903	10.206	0.195
Beam 2 BKGD	4.531	2.236	12.930	3.227
Beta*	0.33 m	10.00 m	0.33 m	2.00 m
Crossing Angle (urad)	-160(V)	200(V)	160(H)	-200(H)

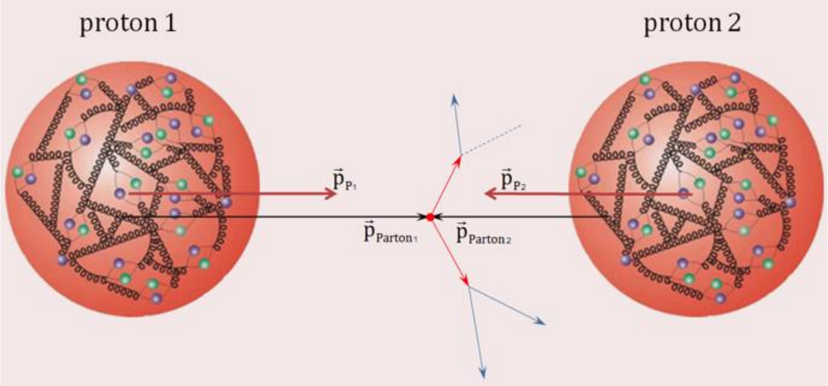
LHCb VELO Position OUT Gap: 53.9 mm TOTEM: STANDBY

Performance over the last 24 Hrs Updated: 19:58:32

Collisions at the LHC: summary



Proton - Proton	2808 bunch/beam
Protons/bunch	10^{11}
Beam energy	7 TeV ($7\mathrm{x}10^{12}$ eV)
Luminosity	$10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
Crossing rate	40 MHz



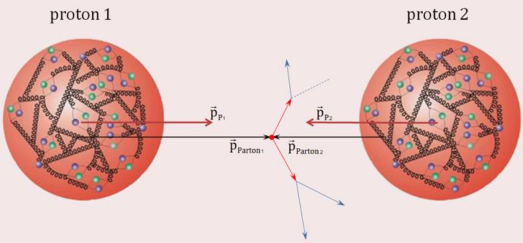
New physics rate ≈ .00001 Hz

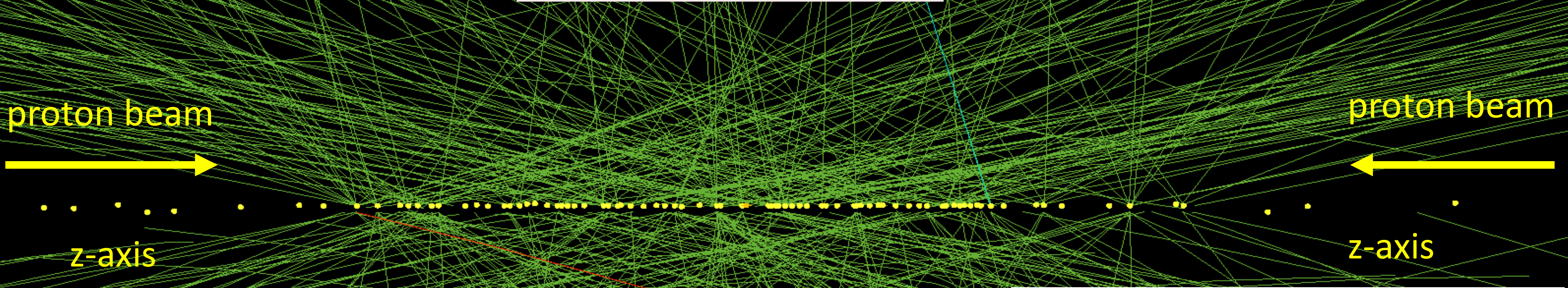
Event selection:
1 in 10,000,000,000,000

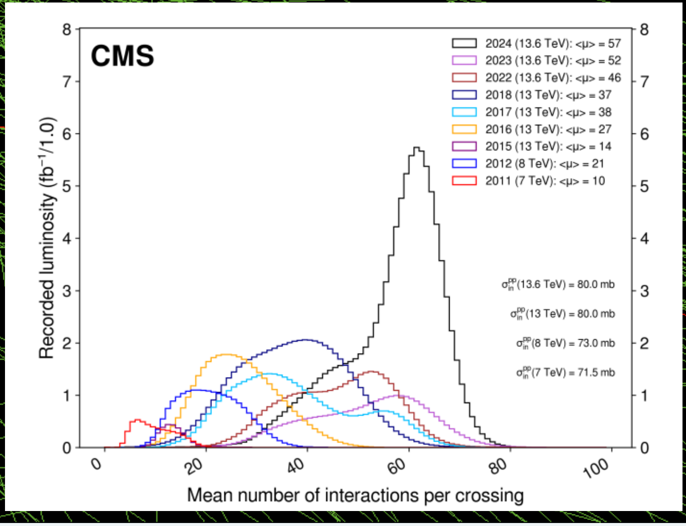
While the protons have 7+7 TeV energy, the partons have a small fraction of that,
→ the parton interaction has energy about an order of magnitude lower than 14 TeV

Realistic “picture” of the collisions at CMS

A “real-life” example of 78 reconstructed collisions in a single event.
HL-LHC will produce ~200 simultaneous pp collisions in a single bunch crossing.

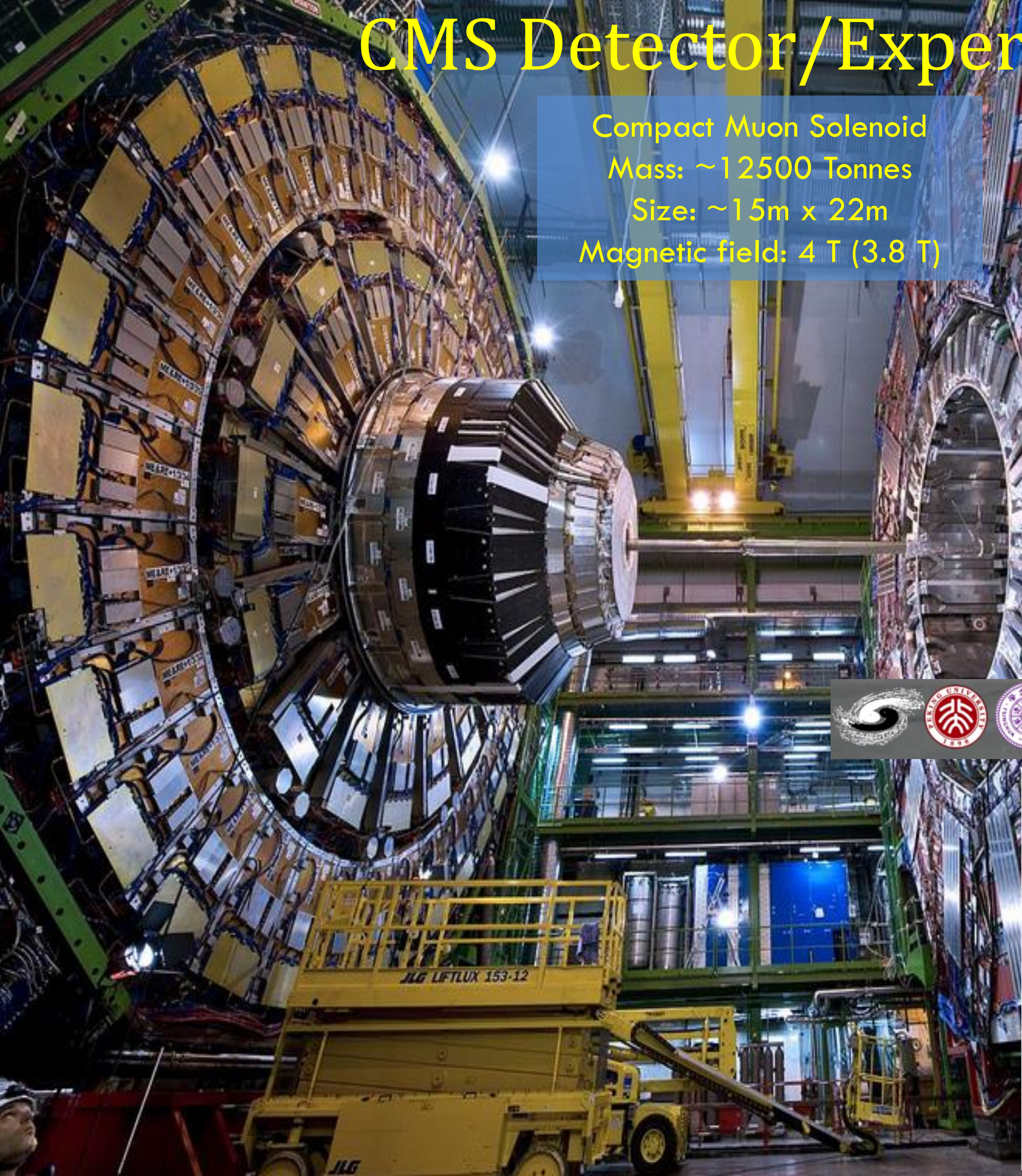






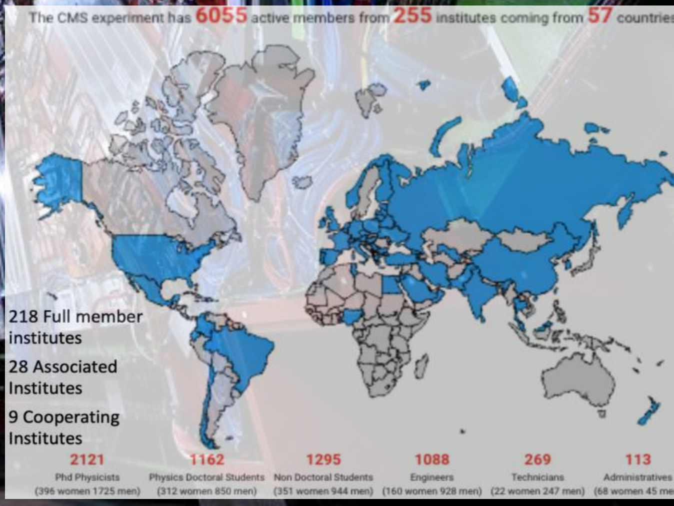
- In each bunch crossing we have many pp interactions
- The vertex with highest pT tracks is selected as primary vertex (PV)
- Only high-pT tracks are shown here
- Hadrons are in green, muon in red

CMS Detector/Experiment



Compact Muon Solenoid
Mass: ~12500 Tonnes
Size: ~15m x 22m
Magnetic field: 4 T (3.8 T)

CMS collaboration is 31 yo
~6100 collaborators
~250 Institutes
~57 countries
here for more



The CMS experiment has 6055 active members from 255 institutes coming from 57 countries

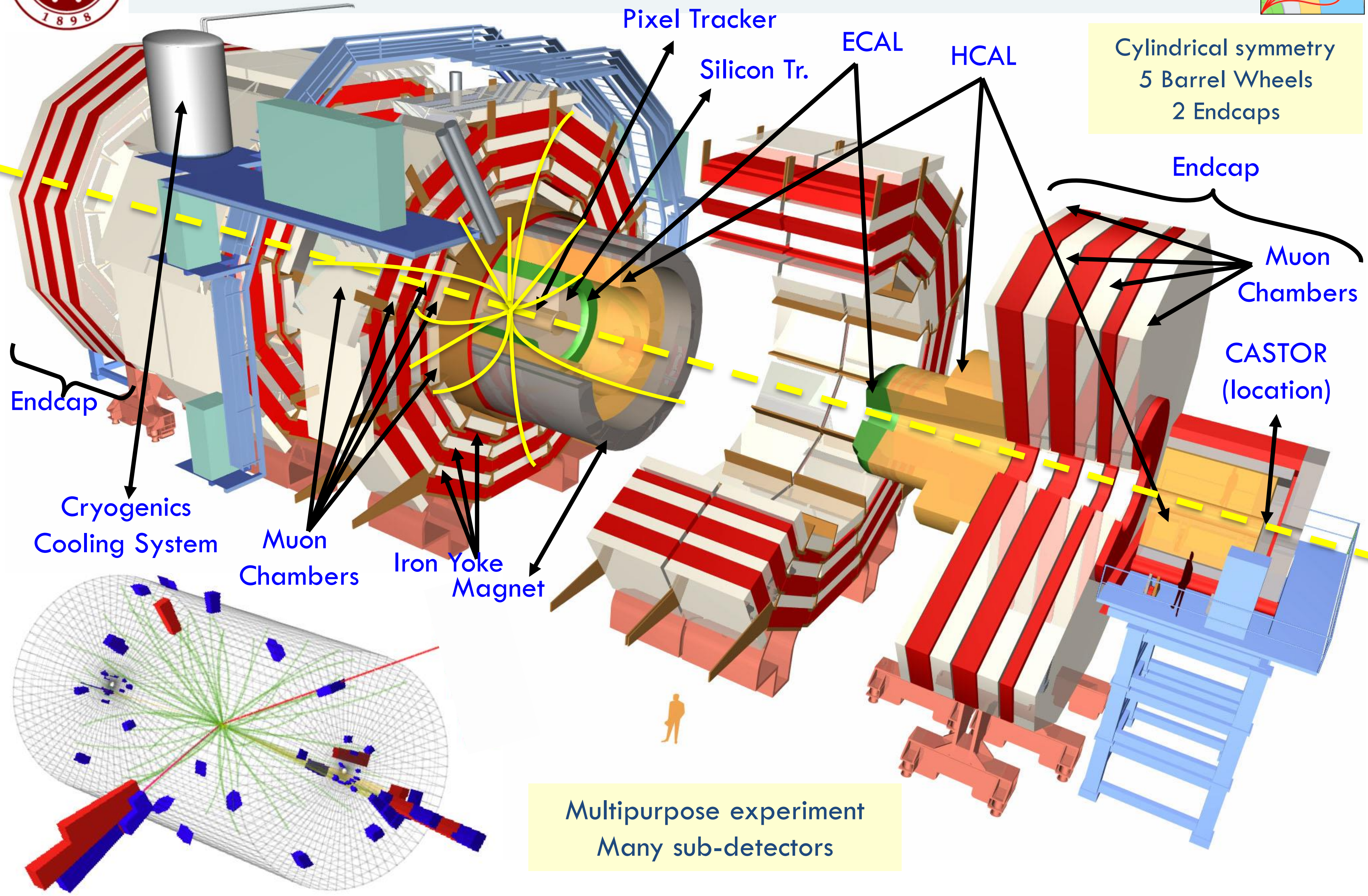
218 Full member institutes

28 Associated Institutes

9 Cooperating Institutes

2121	1162	1295	1088	269	113
Phd Physicists (396 women 1725 men)	Physics Doctoral Students (312 women 850 men)	Non Doctoral Students (351 women 944 men)	Engineers (160 women 928 men)	Technicians (22 women 247 men)	Administratives (68 women 45 me

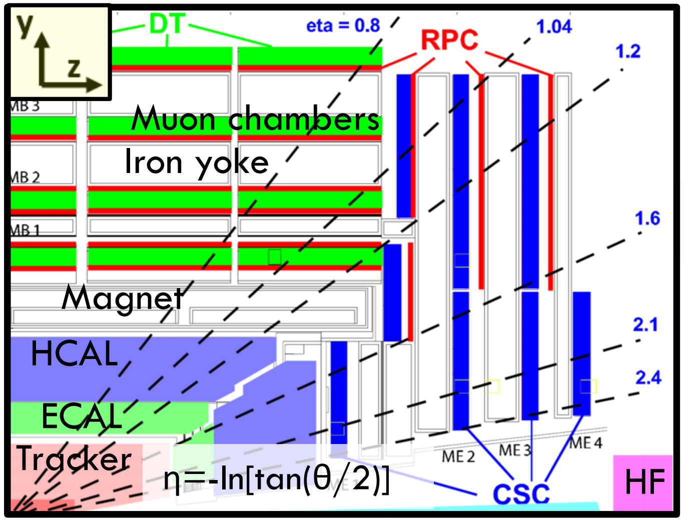
CMS overview

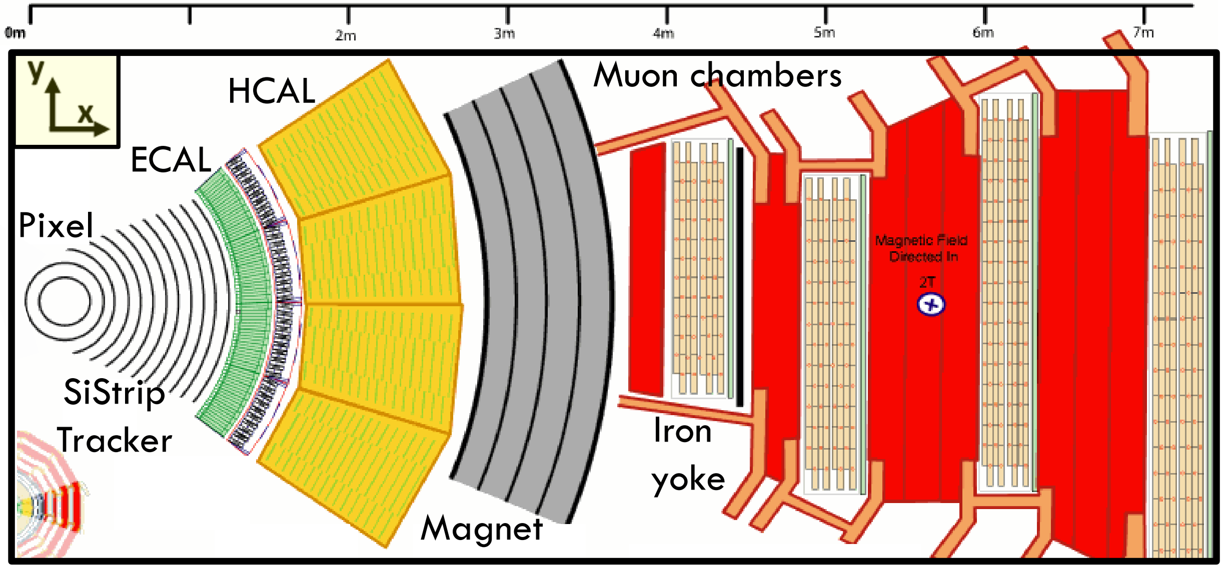


Cylindrical symmetry
5 Barrel Wheels
2 Endcaps

Multipurpose experiment
Many sub-detectors

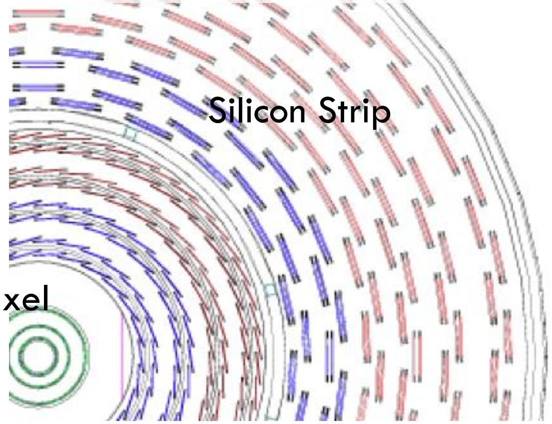
CMS sub-detectors





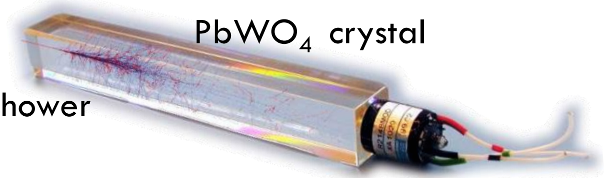
Pixel & Silicon Tracker

- 70M pixels.
- “e-holes” pairs→ signal.
- |η|<2.5, 2.4 (hits:10-13).



ECAL ($e^{\pm}$,γ)

- 76k crystals $PbWO_4$.
- X_0~0.9cm, |η|<3.



HCAL ($p^{\pm}$,n,π,K,Δ…)

- HB, HE(16 leyers), HO, HF.
- Plastic scint.:Quartz fibers.
- Brass(Cu-Zn) absorber
- X_0~1.5cm, |η|<3.

Solenoid Magnet, Iron Yoke

- NiTi, T~1.8K, I~19kA, B~4T.

Muon chambers:

- DTs, CSCs, RPCs, |η|<2.4.
- Argon-based gasses.

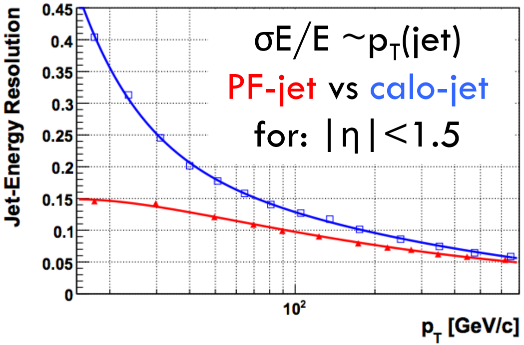
Trigger: L1→HLT→DAQ

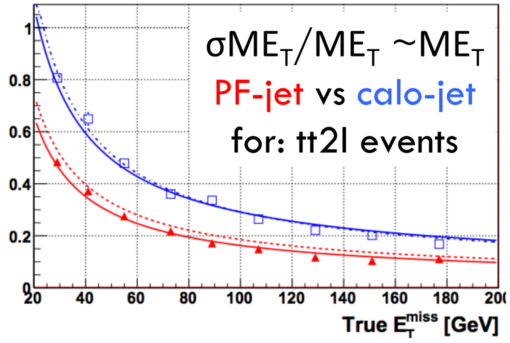
- 40M→40K→~100 ev./s
- Store:Tier-0-1-2→GRID…

From signals → Physics object reco

◆ Main Strategy: "Particle Flow".

- Input: vertexes/tracks/calo-clast.
- Clustered to 5 type of particles:
  γ, $e^{\pm}$, $\mu^{\pm}$, had$^{\pm,0}$.

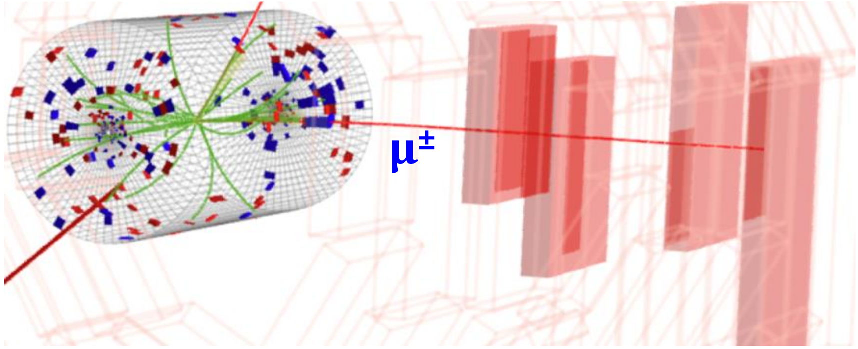




LEPTONS:

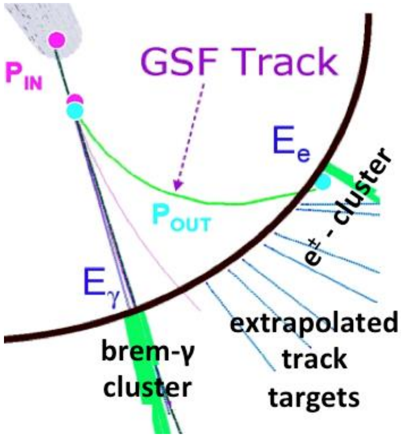
◆ **Muons** [$\mu^{\pm}$] (~stable, τ~2 μs)

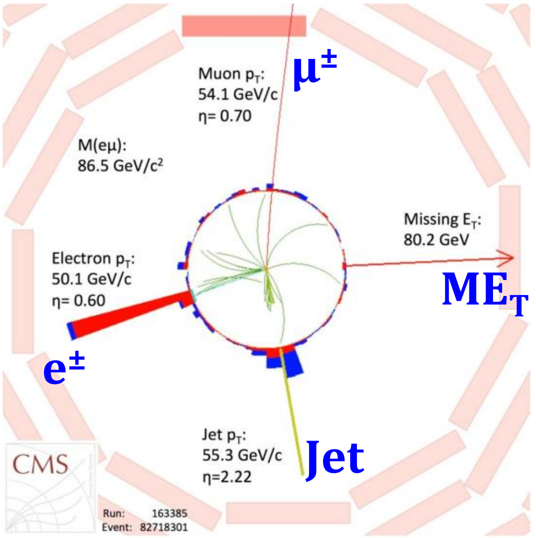
- Reco: Tracker & μ-chambers info.
- Purity-Eff. enhance with cuts:
  hits, Iso, d_z…



◆ **Electrons** [$e^{\pm}$]

- $e^{\pm}$ interact with tracker → radiates "γ".
- Reco: ECAL & Tracker info.
- Correct for brem-γ → fit to get "$e^{\pm}$".
- Reject γ-conversion.
- Purity-Eff. enhance cuts:
  χ^2/ndf, hits, Iso, $d_{z,xy}$, E/p-match, E/H, $\Delta\phi_{in}$,$\Delta\eta_{in}$
  → Correct MC eff-SF.

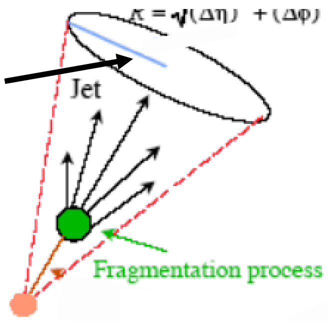


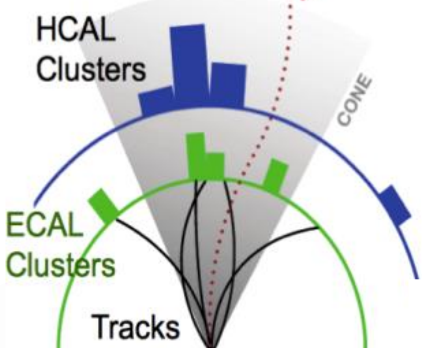


From signals → Physics object reco

COMPOSITE OBJECTS

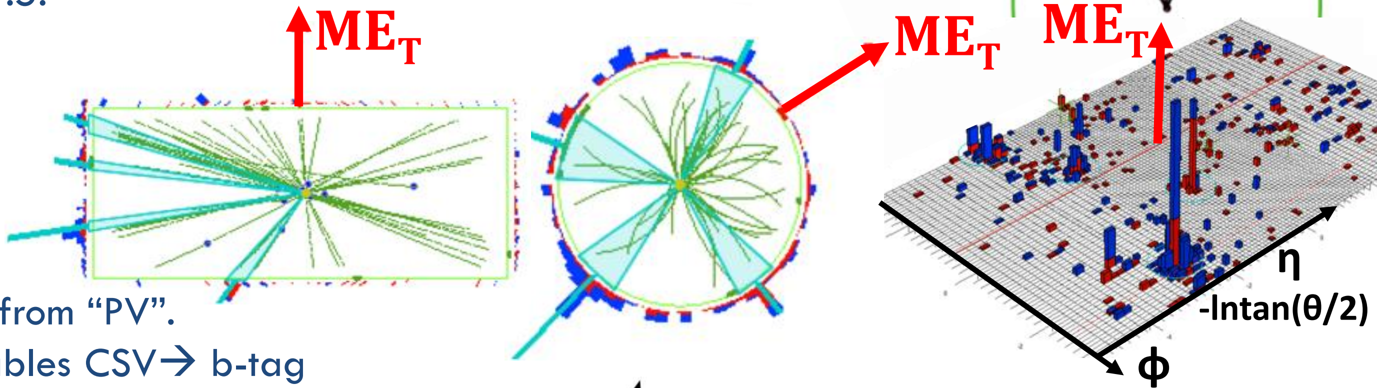
$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$$





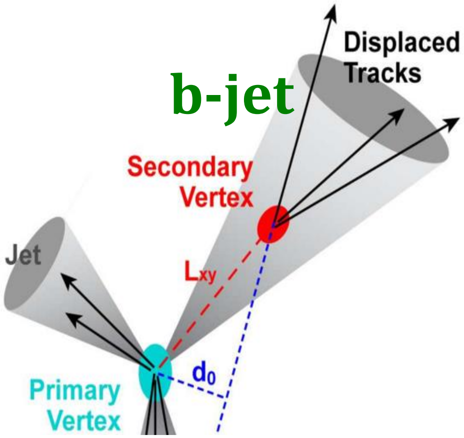
◆ Jets = hadronized q,g

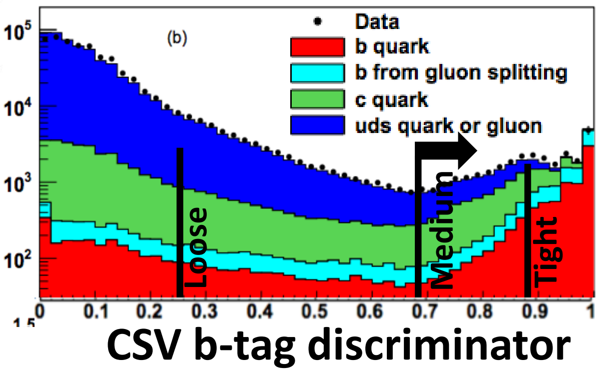
- Flow in cone: ΔR, η-ϕ plane:
- Jet = "PF-objects" + clustering algorithm.
- "anti-k_T", R= 0.4 or 0.8 or 1.5.



◆ b-jets

- b-hadrons: long-lived
- Fly ~mm. b→Wc/u
- Produce "SV" in a distance from "PV".
- Identify "SV" → build variables CSV→ b-tag





◆ ME_T

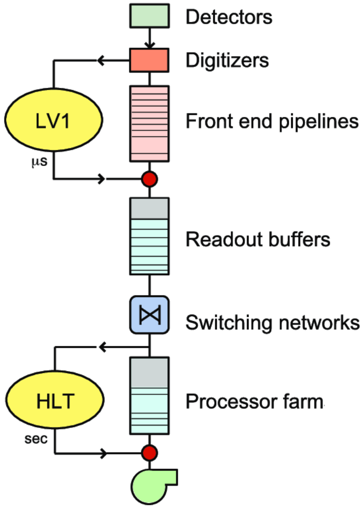
- Momentum imbalance in "xy"-plane using all PF-objects:

$$\vec{\not{E}}_T \equiv -\sum_i \vec{p}_T$$

(I will skip "γ", "τ_h", g-tagger, t-tagger, performance plots & calibr. Technics)

Data online selection & Trigger

- 40 MHz crossing rate.
- We can't record and store all…
- We need to make a wise decision of which events to keep for offline analysis
  → this is done by the "Trigger"



Detectors

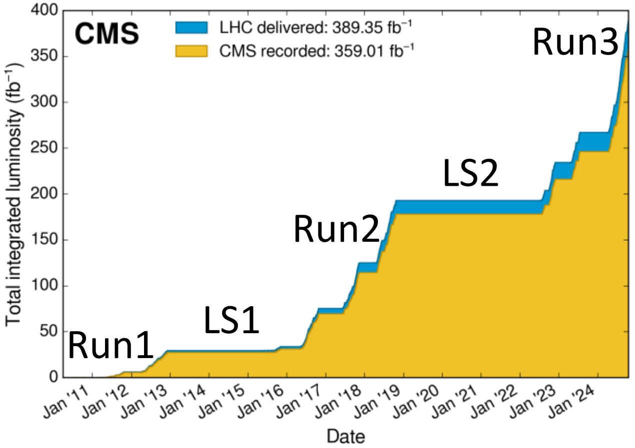
Digitizers

Front end pipelines

Readout buffers

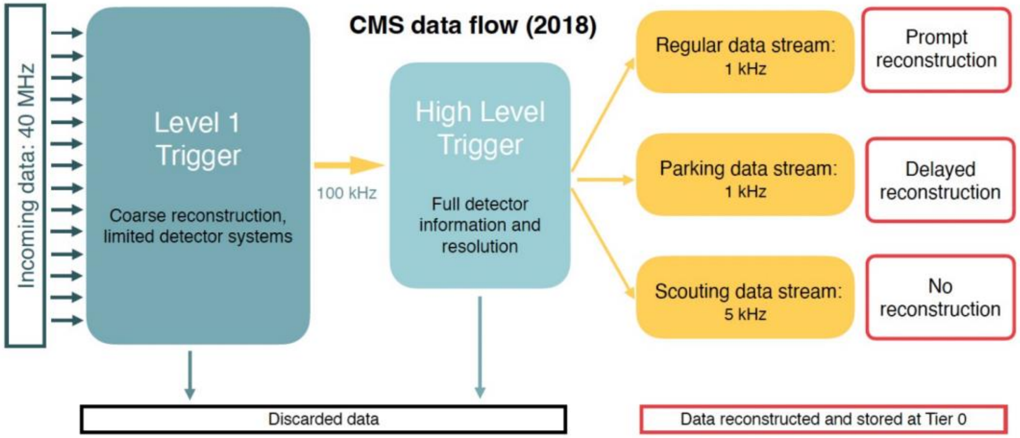
Switching networks

Processor farm



Run 3 pp data certification efficiency 93%

- CMS trigger has two tiers:
  → **Level 1 Trigger**: 40 MHz → 100 kHz (reduction by ~400)
  → **High Level Trigger (HLT)** 100 kHz → 1000 Hz (reduction by ~100)



CMS data flow (2018)

Incoming data: 40 MHz

Level 1 Trigger

Coarse reconstruction, limited detector systems

100 kHz

High Level Trigger

Full detector information and resolution

Regular data stream: 1 kHz

Prompt reconstruction

Parking data stream: 1 kHz

Delayed reconstruction

Scouting data stream: 5 kHz

No reconstruction

Discarded data

Data reconstructed and stored at Tier 0

- Most of events are QCD (not interesting)
- Physics criteria on what to keep, like, high energy leptons, MET, b-jets, H_T, …

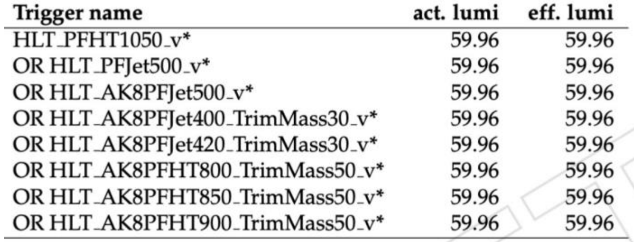
Year	Standard rate [Hz]	Parking rate [Hz]	Scouting rate [Hz]
2011	100	-	400
2012	100	500	500
2016	1000	500	500
2017	1200	500	4000
2018	1350	900	5000

HLT paths & Trigger efficiency

- In the offline analysis, we take data events from a “dataset” depending on what physics we do e.g., single μ/e, photon, JetHT, …

Sample name
/JetHT/Run2018A-UL2018_MiniAODv2-v1/MINIAOD
/JetHT/Run2018B-UL2018_MiniAODv2-v1/MINIAOD
/JetHT/Run2018C-UL2018_MiniAODv2-v1/MINIAOD
/JetHT/Run2018D-UL2018_MiniAODv2-v2/MINIAOD

- Each dataset can have events accepted by several “HLT paths”:



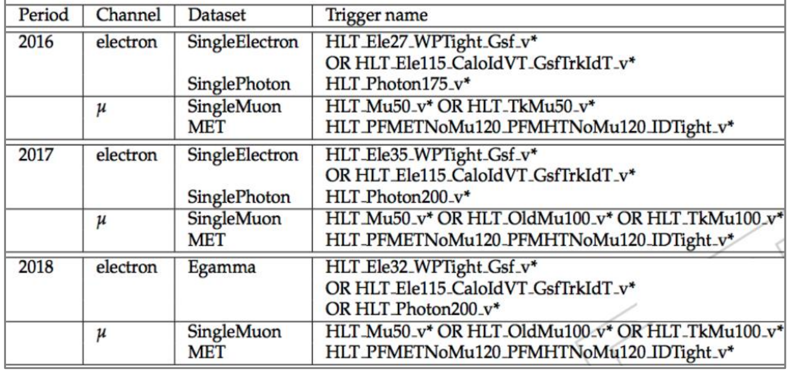
Trigger name	act. lumi	eff. lumi
HLT_PFHT1050_v*	59.96	59.96
OR HLT_PFJet500_v*	59.96	59.96
OR HLT_AK8PFJet500_v*	59.96	59.96
OR HLT_AK8PFJet400_TrimMass30_v*	59.96	59.96
OR HLT_AK8PFJet420_TrimMass30_v*	59.96	59.96
OR HLT_AK8PFHT800_TrimMass50_v*	59.96	59.96
OR HLT_AK8PFHT850_TrimMass50_v*	59.96	59.96
OR HLT_AK8PFHT900_TrimMass50_v*	59.96	59.96

- Each event has HTL val. either accepted (1) or rejected (0).
- We evaluate trigger eff. using another orthogonal dataset

$$eff(var = x, trigger) = \frac{Nevents(var > x \,\&\&\, trigger == True)}{Nevents(var > x)}$$

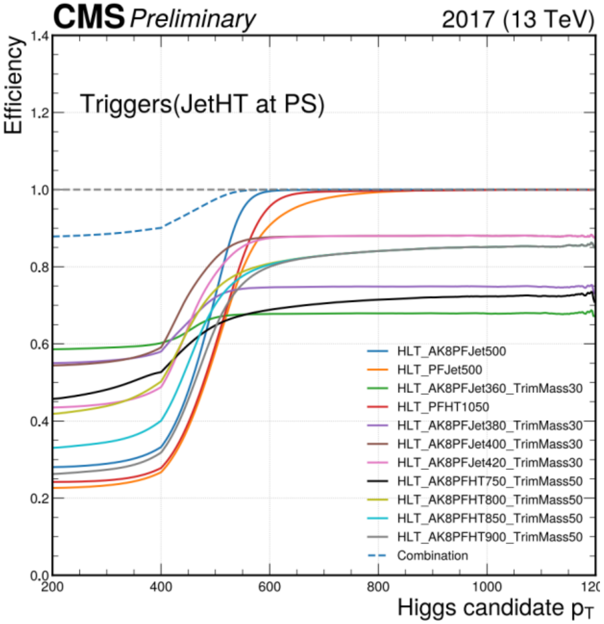
HL-Triggers:

- μ: Single_Muon OR MET
- e: Single_Electron OR Photon



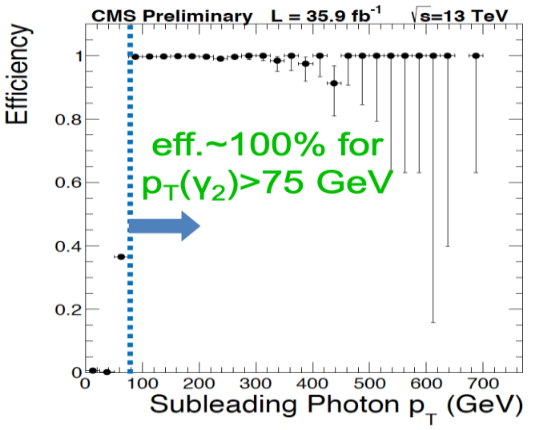
Period	Channel	Dataset	Trigger name
2016	electron	SingleElectron SinglePhoton	HLT_Ele27_WPTight_Gsf_v* OR HLT_Ele115_CaloIdVT_GsfTrkIdT_v* HLT_Photon175_v*
	μ	SingleMuon MET	HLT_Mu50_v* OR HLT_TkMu50_v* HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v*
2017	electron	SingleElectron SinglePhoton	HLT_Ele35_WPTight_Gsf_v* OR HLT_Ele115_CaloIdVT_GsfTrkIdT_v* HLT_Photon200_v*
	μ	SingleMuon MET	HLT_Mu50_v* OR HLT_OldMu100_v* OR HLT_TkMu100_v* HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v*
2018	electron	Egamma	HLT_Ele32_WPTight_Gsf_v* OR HLT_Ele115_CaloIdVT_GsfTrkIdT_v* OR HLT_Photon200_v*
	μ	SingleMuon MET	HLT_Mu50_v* OR HLT_OldMu100_v* OR HLT_TkMu100_v* HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_v*

Trigger “turn-on” curves

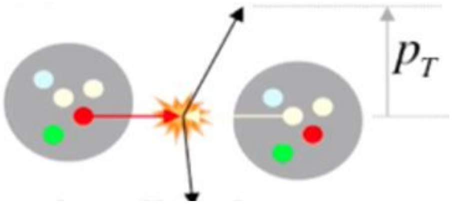


Triggers: [2016 DoubleEG data, 35.9 fb^{-1}]

- Use HLT_DoublePhoton60 ”or” HLT_ECALHT800 for main analysis
- Use HLT_Ele27_WPTight_Gsf for Z → e^+e^- control sample (energy scale and efficiency measurement)

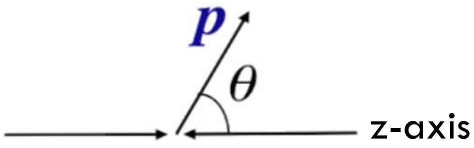


pp collisions' kinematics



- Natural variables would be: p, θ, ϕ,… but…

- Longitudinal momentum & energy: p_Z, E, can NOT be used
  - they are conserved, but are unknown
  - particles close to beam axis escaping detection have large p_Z

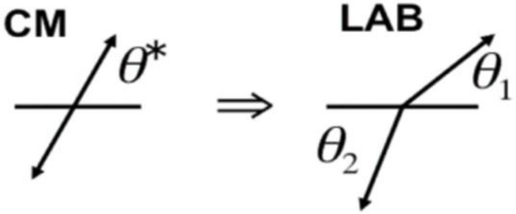


- More useful transverse momentum: p_T

$$p_T \equiv \sqrt{p_x^2 + p_y^2} = |p| \sin\theta$$

Parton CM (energy)^2 → $\hat{s} = x_1 x_2 s$

- Lab frame ≠ parton-parton CM frame
  → additionally, p, E & θ, are NOT Lorentz invariant along z-axis boosts.



- Rapidity "y"

$$y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z} = \frac{1}{2} \ln \frac{1 + \beta\cos\theta}{1 - \beta\cos\theta}$$

- Pseudo-rapidity "η"
  (it is just a function of θ)

$$\beta \rightarrow 1 \ (m << p_T):$$

$$\eta \equiv -\ln\tan\frac{\theta}{2}$$

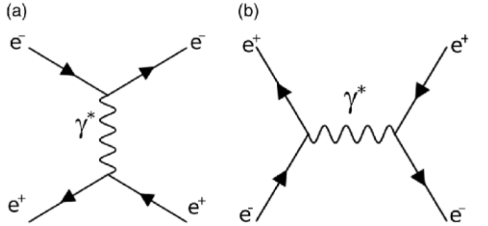
Advantage!!!
Δη is invariant under
z-axis Lorentz boosts

- Distance between 2 particles: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is also an invariant.

→ To take away: particles are described by pT, η, ϕ coordinates

Kinematics & variables used (1/2)

(a) t-channel and (b) s-channel diagrams contributing to Bhabha scattering in the SM.



1. Mandelstam invariant variables:
   $s \equiv (p_1 + p_2)^2 = (p_3 + p_4)^2 = (E_1 + E_2)^2 + (\vec{p}_1 + \vec{p}_2)^2$,
   $t \equiv (p_1 - p_3)^2 = (p_2 - p_4)^2 = (E_1 - E_3)^2 + (\vec{p}_1 - \vec{p}_3)^2$,
   $u \equiv (p_1 - p_4)^2 = (p_2 - p_3)^2 = (E_1 - E_4)^2 + (\vec{p}_1 - \vec{p}_4)^2$,
   where in the relativistic limit $((p_i/E_i) \to 1)$ turn to:
   $s = 2p_1p_2 = 2p_3p_4$
   $t = -2p_1p_3 = -2p_2p_4$
   $u = -2p_1p_4 = -2p_2p_3$
   where momenta $(p_i \equiv p_i^\mu)$ stands for the generic interaction "1+2→3+4"and thus $p_1 + p_2 = p_3 + p_4$. The $\sqrt{s}$ is the total (initial and final) energy into the lab-frame; it is also true that: $s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2$.

2. Transverse momentum: $p_T \equiv \sqrt{p_x^2 + p_y^2} = |p| \sin\theta$.

3. Transverse energy : $E_T \equiv E \sin\theta = p_T(E/|p|)$.

4. Transverse mass (of a single particle): $m_T \equiv \sqrt{p_T^2 + m^2}$.

5. Rapidity: $y \equiv \frac{1}{2} \ln[(E + p_Z)/(E - p_Z)]$.

6. Pseudorapidity: $\eta \equiv \lim_{[(E/p)\to 1]} y = -\ln[\tan(\theta/2)]$.

7. Invariant η-ϕ solid angle: $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

8. Four-momentum of a particle as function of different variables:
   $f(E, p_x, p_y, p_z)$: $p^\mu \equiv (E;\ p_x, p_y, p_z)$,
   $f(m_T, p_x, p_y, y)$: $p^\mu = (m_T \cosh y;\ p_x, p_y, m_T \sinh y)$,
   $f(m_T, p_T, y, \phi)$: $p^\mu = (m_T \cosh y;\ p_T \cos\phi, p_T \sin\phi, m_T \sinh y)$,
   $f(m, p_T, \eta, \phi)$: $p^\mu = (\sqrt{m^2 + p_T^2 \cosh^2\eta};\ p_T \cos\phi, p_T \sin\phi, p_T \sinh\eta)$.

Kinematics & variables used (2/2)

9. Invariant mass of n particles: $M \equiv \sqrt{(p_1^\mu + p_2^\mu + ... + p_n^\mu)^2} = \sqrt{(E_1 + E_2 + ...)^2 - (\vec{p}_1 + \vec{p}_2 + ...)^2}$.

10. Invariant mass of 2 particles: $M = \sqrt{E_1^2 + E_2^2 + 2E_1E_2 - p_1^2 - p_2^2 - 2\vec{p}_1 \cdot \vec{p}_2}$.

11. Invariant mass of 2 particles for $(E_i/p_i) \to 1$: $M = \sqrt{2p_1p_2(1-\cos\theta)}$.

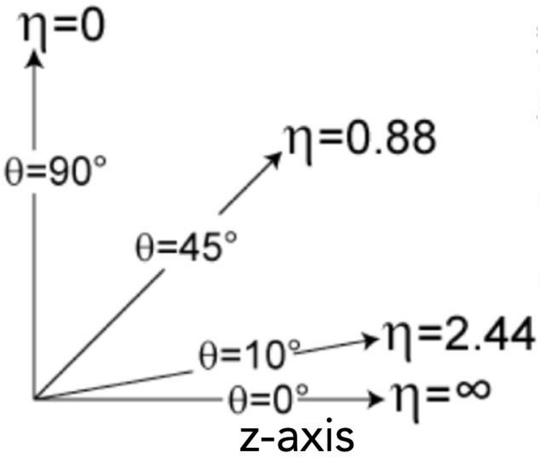
12. Transverse mass of 2 particles: $M_T \equiv \sqrt{(E_1+E_2)^2 - (\vec{p}_{T1} + \vec{p}_{T2})^2}$.

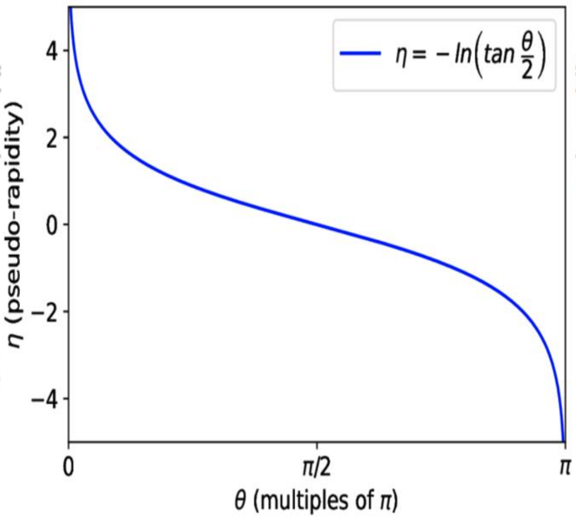
13. Transverse mass of 2 particles for $(E_i/p_i) \to 1$: $M_T = \sqrt{2p_{T1}p_{T2}(1-\cos\theta)}$.

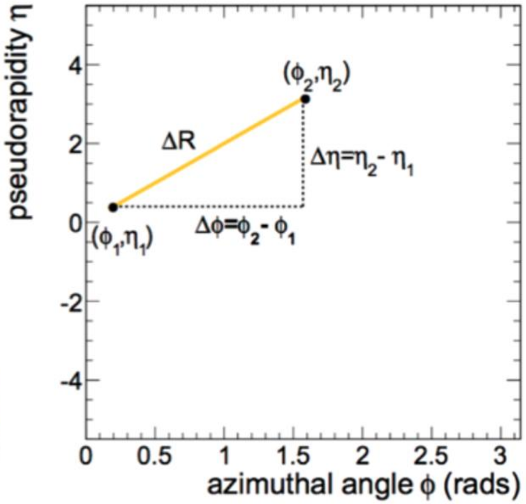
14. Transverse hadronic activity: $H_T \equiv \sum_i p_T(jet[i])$ (where the sum runs over all jets above some thresholds, which in particular analysis are p_T >30 GeV, $|\eta|$ <4.5).

15. Hadronic recoil: $\vec{H}_T^{miss} \equiv -\sum_i \vec{p}_T(jet[i])$ where the sum runs over all jets. (Usually used in Z+jets samples where is equal with the p_T(Z)). It is true that: $|\vec{H}_T^{miss}| \neq H_T$ due to the threshold(s) imposed to selected jets in H_T.

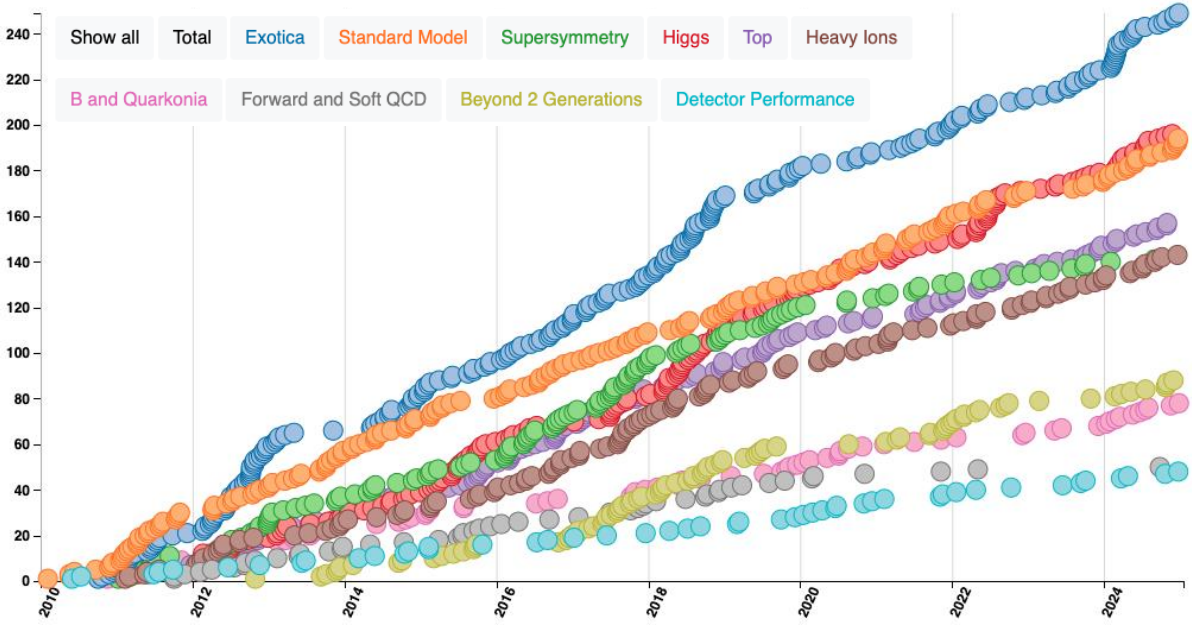
16. Transverse missing energy (or momentum): $\vec{\not{E}}_T \equiv -\sum_i \vec{p}_T(i)$ where the sum runs over, all reconstructed objects: jets, leptons, taus, photons (without cut thresholds).

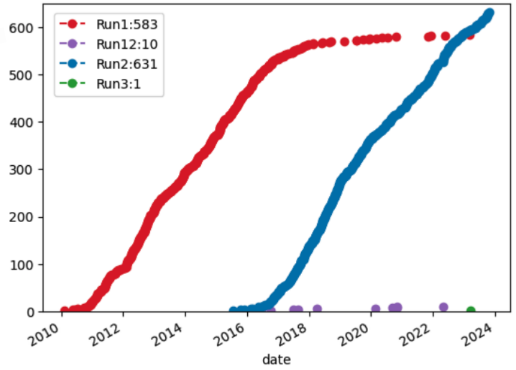


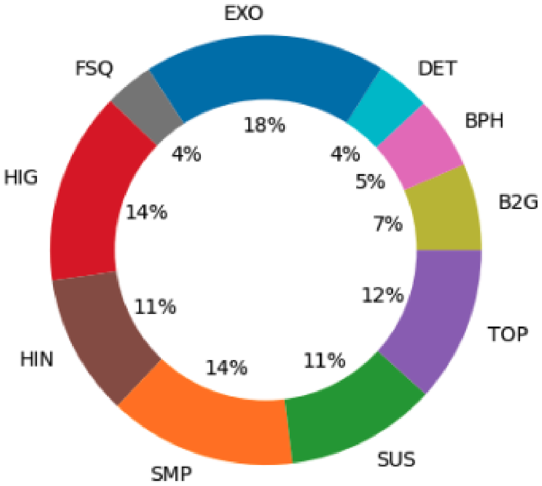




Physics program at CMS







- ***1347 collider data papers*** submitted (as of 2025-1-1)
- Publication statistics Run2 not yet at plateau!
- ~90 analyses in CWR or beyond, ~200 in AWG progress

https://cms-results.web.cern.ch/cms-results/public-results/publications-vs-time/

- Physics Analyses Groups (PAGs)

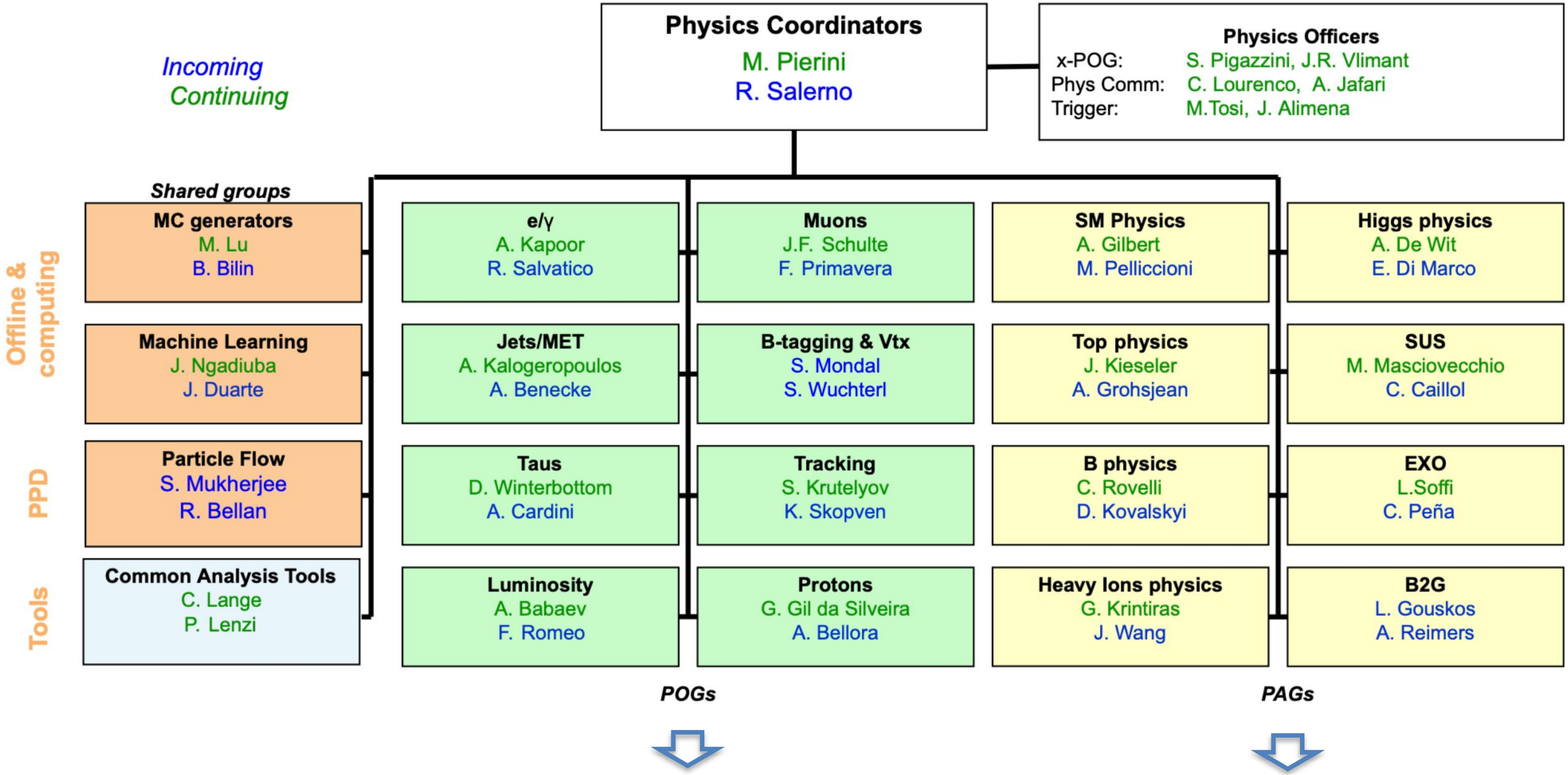
  → (BSM) Searches : **SUS, EXO, B2G**

  → SM Measurements (mostly): **SMP, TOP, BPH, HIG, HIN**

CMS titles
- ❖ 600 “Search”
- ❖ 52 “Observation”
- ❖ 21 “Evidence”
- ❖ 336 “Measurement”
- ❖ 42 “Study”

CMS organigram at Physics coordination areas: POGs & PAGs

How is our community organized?



Incoming
Continuing

Physics Coordinators
M. Pierini
R. Salerno

Physics Officers
x-POG: S. Pigazzini, J.R. Vlimant
Phys Comm: C. Lourenco, A. Jafari
Trigger: M.Tosi, J. Alimena

Shared groups

Offline & computing

MC generators
M. Lu
B. Bilin

Machine Learning
J. Ngadiuba
J. Duarte

PPD

Particle Flow
S. Mukherjee
R. Bellan

Tools

Common Analysis Tools
C. Lange
P. Lenzi

e/γ
A. Kapoor
R. Salvatico

Jets/MET
A. Kalogeropoulos
A. Benecke

Taus
D. Winterbottom
A. Cardini

Luminosity
A. Babaev
F. Romeo

Muons
J.F. Schulte
F. Primavera

B-tagging & Vtx
S. Mondal
S. Wuchterl

Tracking
S. Krutelyov
K. Skopven

Protons
G. Gil da Silveira
A. Bellora

POGs

SM Physics
A. Gilbert
M. Pelliccioni

Top physics
J. Kieseler
A. Grohsjean

B physics
C. Rovelli
D. Kovalskyi

Heavy Ions physics
G. Krintiras
J. Wang

Higgs physics
A. De Wit
E. Di Marco

SUS
M. Masciovecchio
C. Caillol

EXO
L.Soffi
C. Peña

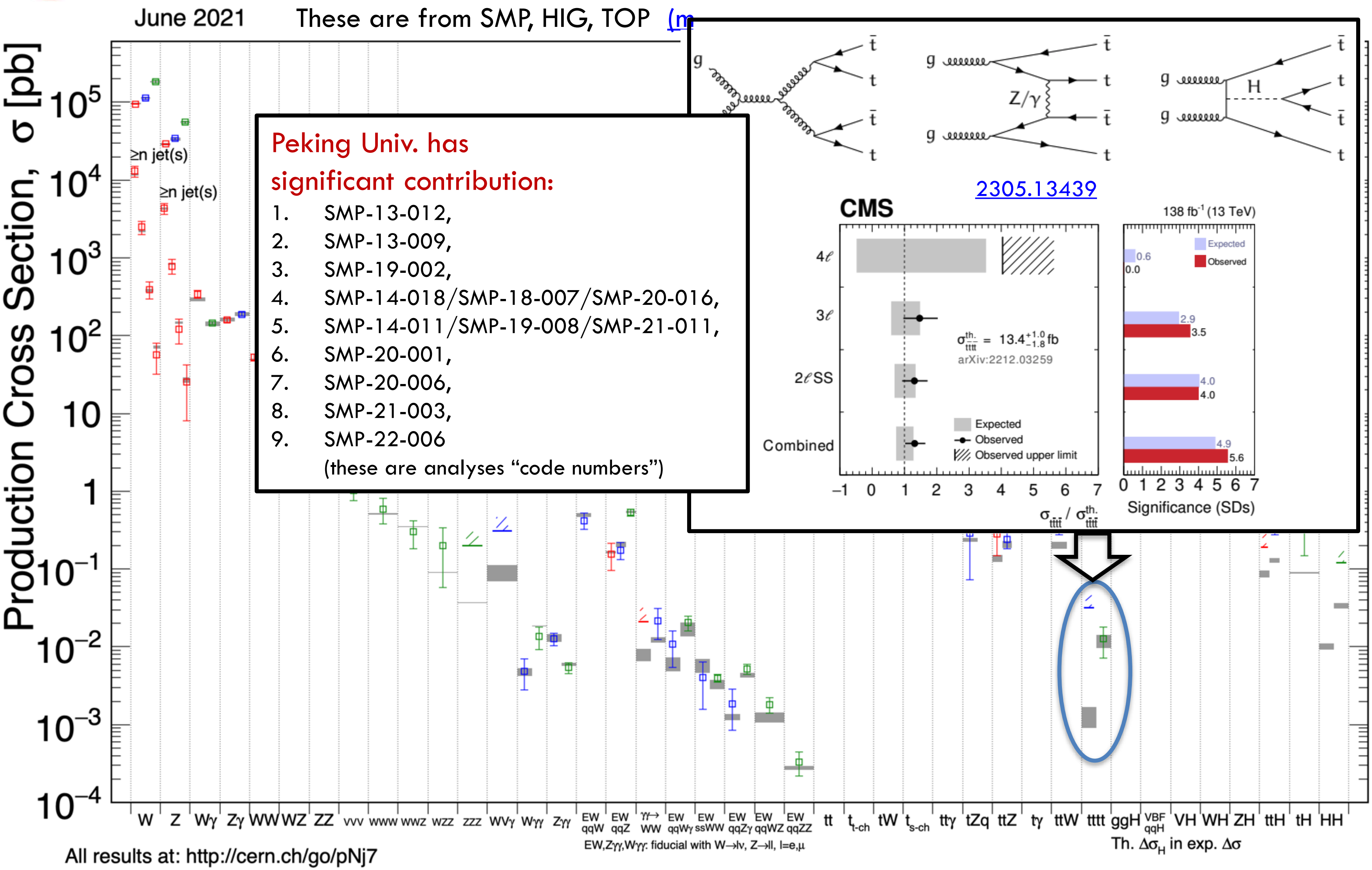
B2G
L. Gouskos
A. Reimers

PAGs

Physics Object Groups (POGs)

Physics Analysis Groups (PAGs)

Several SM measurements (summary)



These are from SMP, HIG, TOP (m

Peking Univ. has significant contribution:

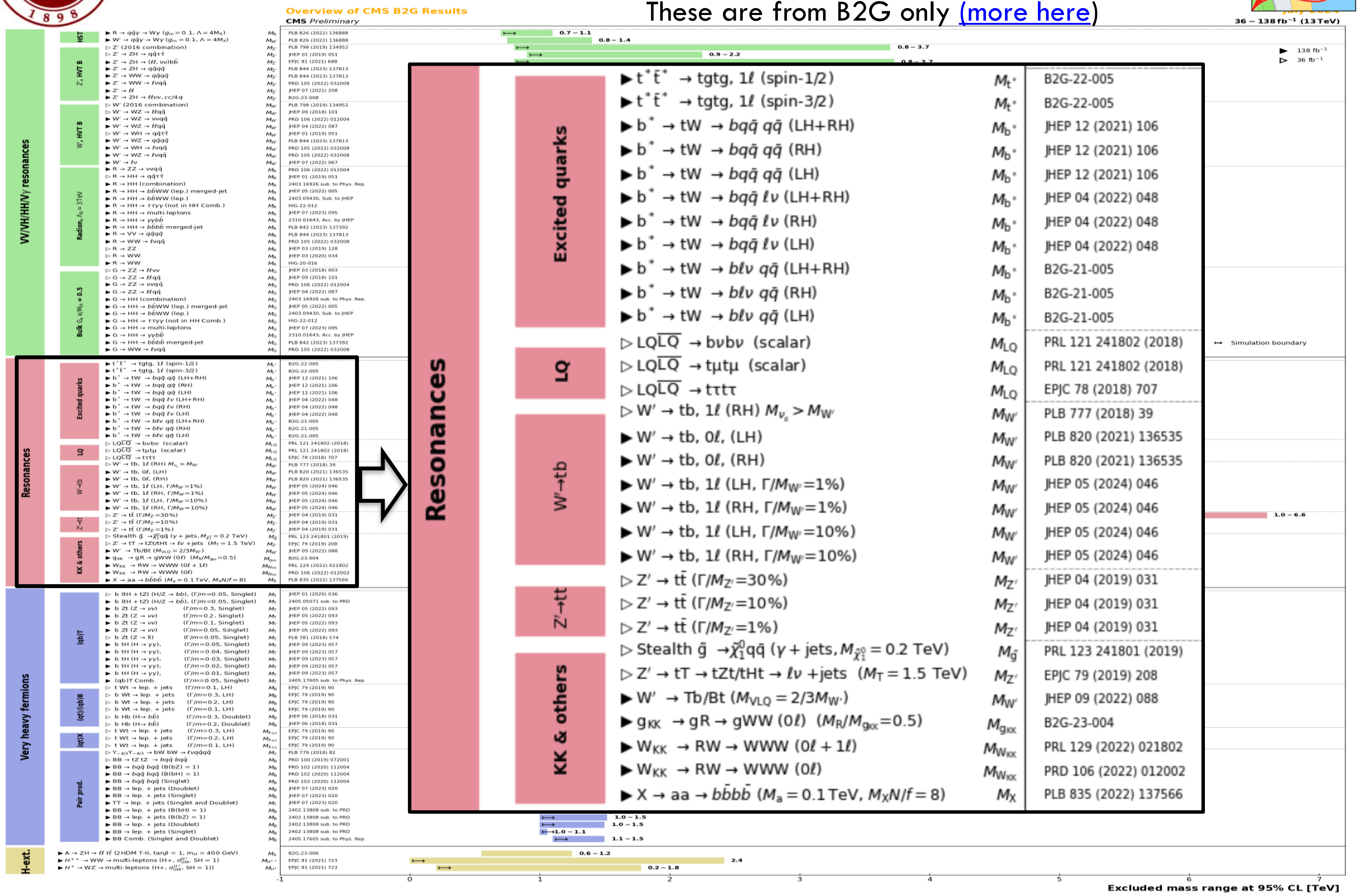
1. SMP-13-012,
2. SMP-13-009,
3. SMP-19-002,
4. SMP-14-018/SMP-18-007/SMP-20-016,
5. SMP-14-011/SMP-19-008/SMP-21-011,
6. SMP-20-001,
7. SMP-20-006,
8. SMP-21-003,
9. SMP-22-006

(these are analyses "code numbers")

2305.13439

Excl. limits from several searches (summary)

These are from B2G only (more here)



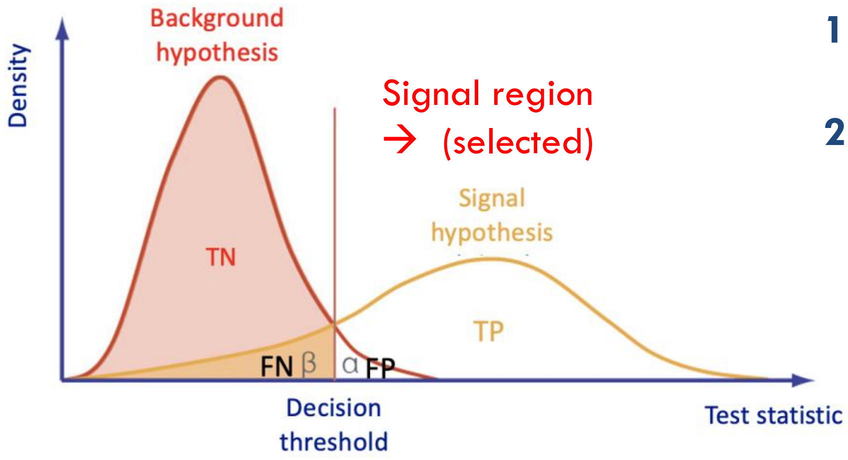
Analysis: Signal & Background

- **What is Data Analysis?**
  → The offline "Data Analysis" refers to the processing of real data (and MC simulated) events (samples) in order to make physics conclusions. Examples will follow.

- **What is Signal (S) and Background (B)?**
  → For each analysis there is a process which we want to study or search for, this is Signal (S). Other SM processes which have similar kinematics and final state appear as Background (B).

- Every analysis has as common goal: to **select Signal and reject Background.**
  This is done by exploiting kinematics and applying cuts on kinematic variables.
- We optimize these cuts based on signal significance e.g. $S/\sqrt{B}$ (or others)

- Finally, we **count S & B events** in some bins and make inferences on physics.

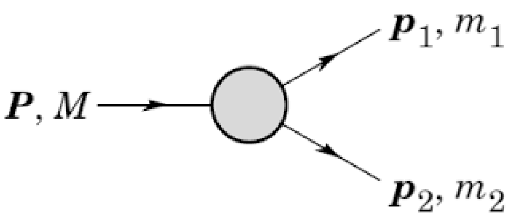


Two general types of analyses in HEP:

1. **SM Measurements:**
   We measure events of an existing physics process.
2. **Searches for BSM physics:**
   We look for a process which is not predicted by SM but from by BSM.
   If such a process exists, it will appear as an excess of events compatible with the BSM model simulation.

Kinematics: Resonances (2 particles)

Invariant mass of a system of 2 particles:



$$\begin{aligned} M^2 &= (E_1 + E_2)^2 - \|\mathbf{p}_1 + \mathbf{p}_2\|^2 \\ &= m_1^2 + m_2^2 + 2\,(E_1 E_2 - \mathbf{p}_1 \cdot \mathbf{p}_2) \end{aligned}$$

$$\left.\begin{aligned} M^2 &= (E_1 + E_2)^2 - \|\mathbf{p}_1 + \mathbf{p}_2\|^2 \\ &= [(p_1, 0, 0, p_1) + (p_2, 0, p_2 \sin\theta, p_2 \cos\theta)]^2 \\ &= (p_1 + p_2)^2 - p_2^2 \sin^2\theta - (p_1 + p_2 \cos\theta)^2 \\ &= 2 p_1 p_2 (1 - \cos\theta). \end{aligned}\right\}$$

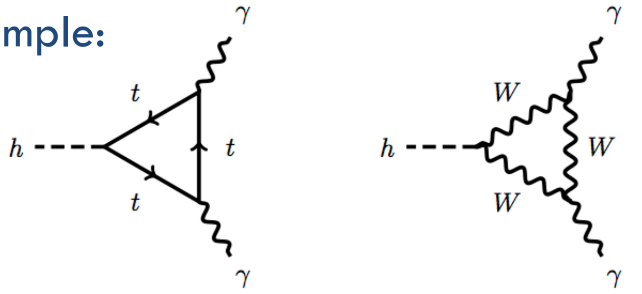
For massless particles

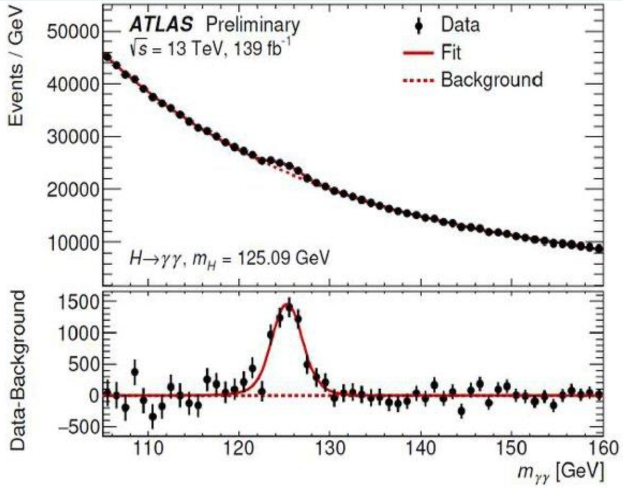
$$M^2 = 2 p_{T1} p_{T2} (\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2))$$

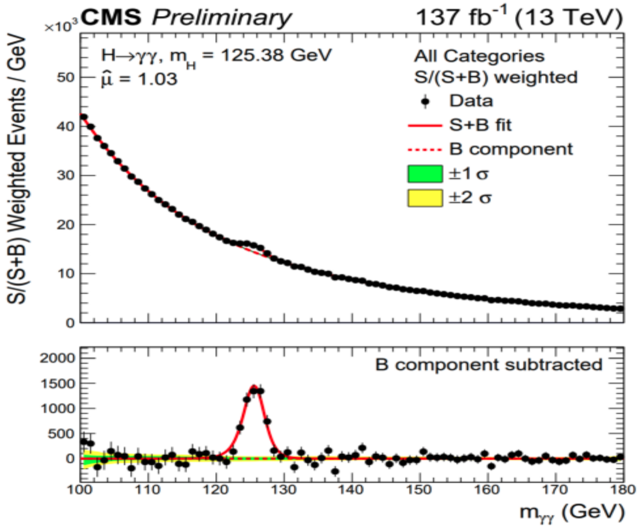
(using η,ϕ)

Diphoton example:

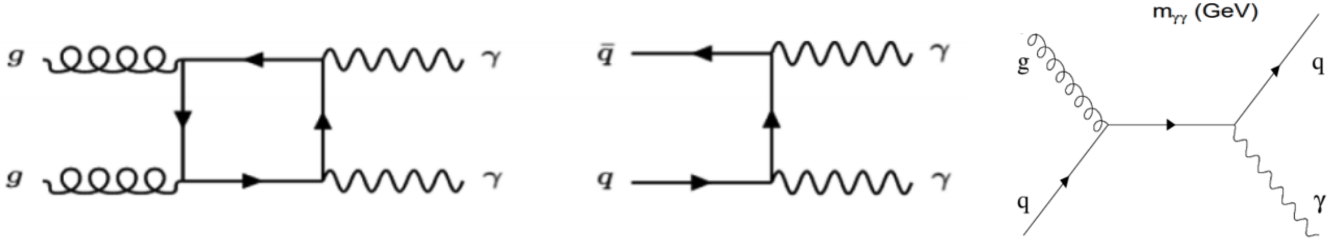
Signal:





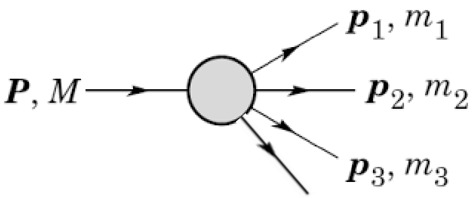


Background (QCD):

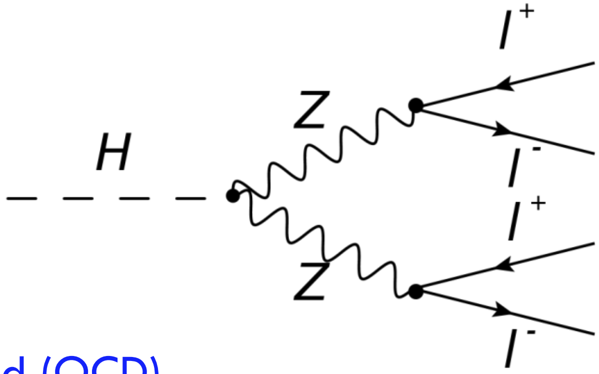


Kinematics: Resonances (4 particles)

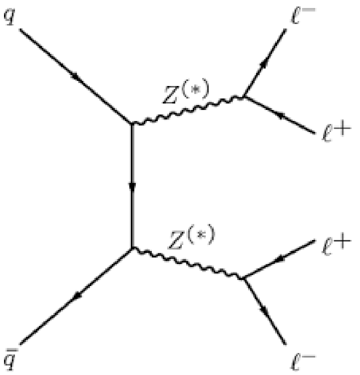
Invariant mass of a system of 3,4... particles:
... Similar formulas

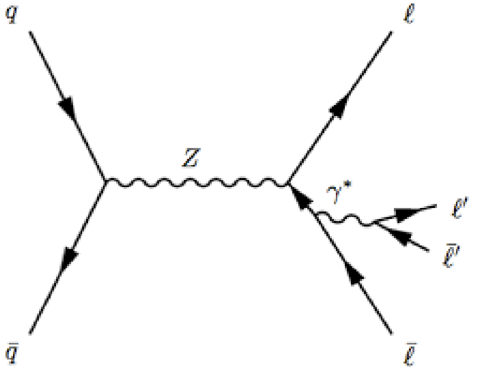


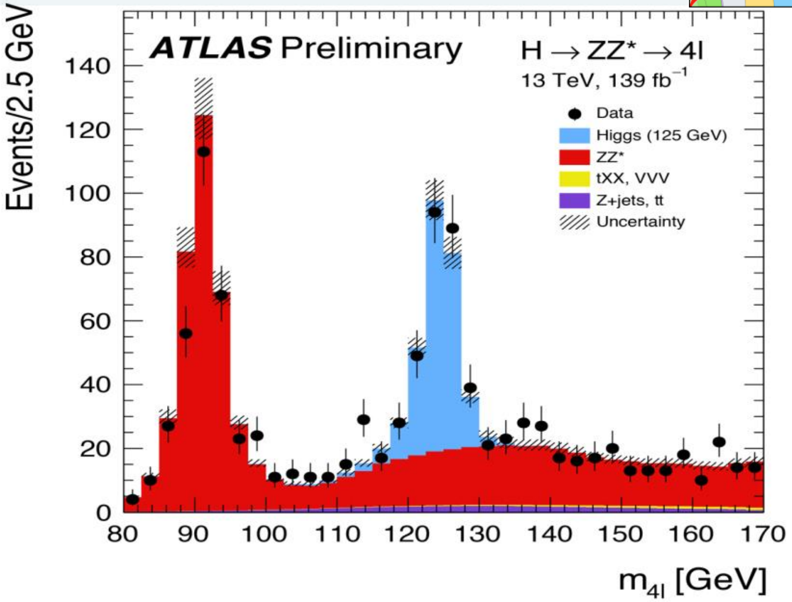
Signal:

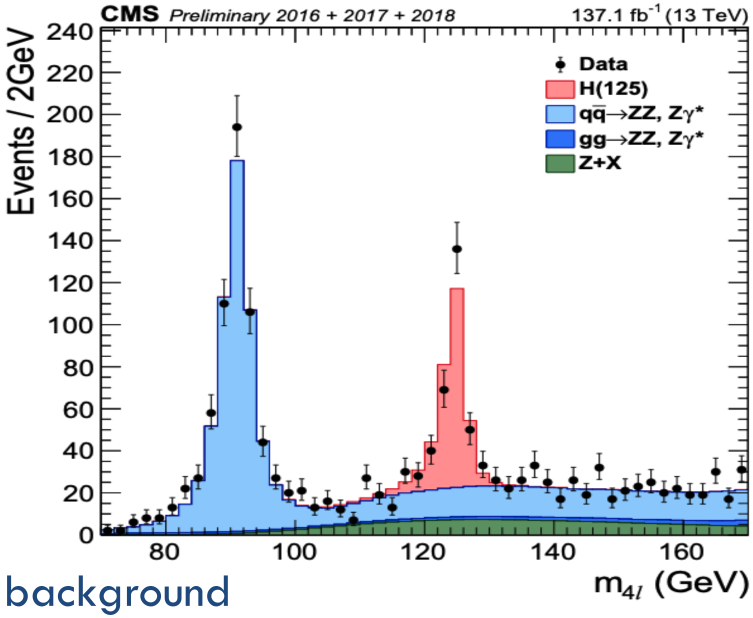


Background (QCD)









MC samples are used to simulate both signal & background

Resonances with neutrinos: Transverse mass (M_T)

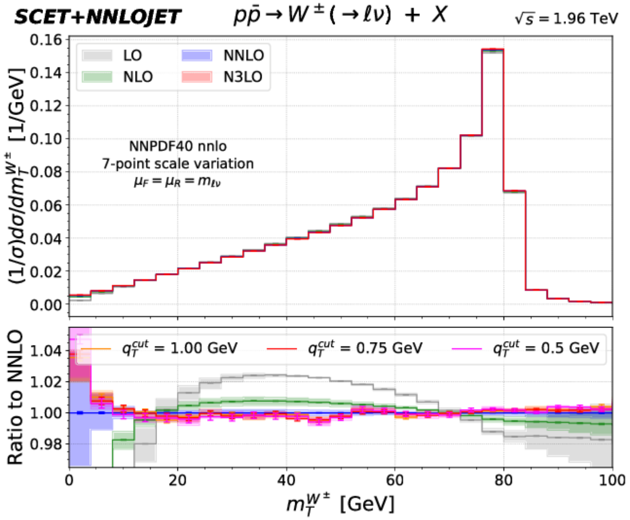
Transverse mass of a system of 2,3,4 particles, with 1 (or more) invisible p (like neutrinos: ν) → MET

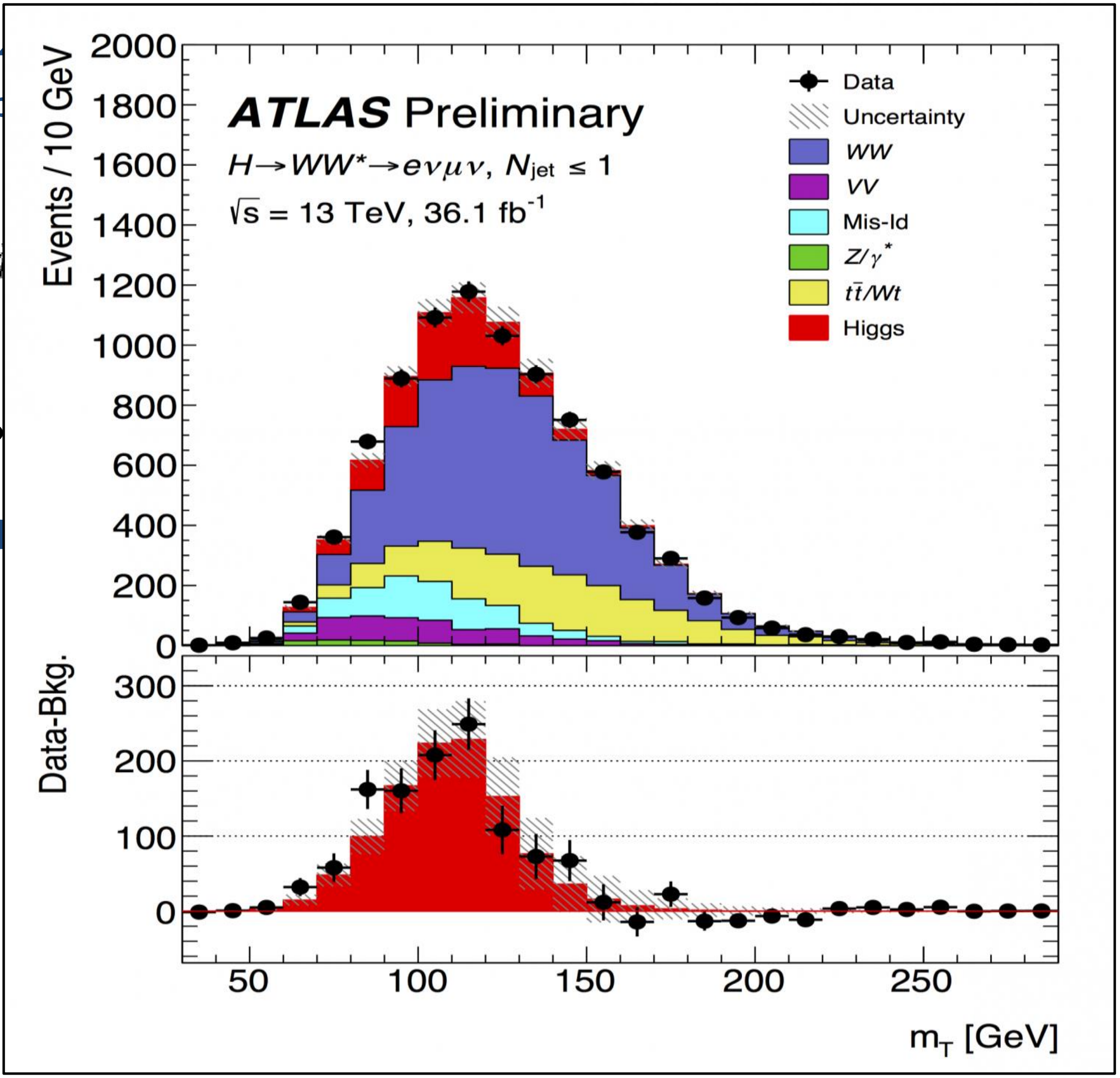
$$M_T^2 = m_1^2 + m_2^2 + 2\left(E_{T,1}E_{T,2} - \vec{p}_{T,1} \cdot\right.$$

For massless particles $$M_T^2 \to 2E_{T,1}E_{T,2}\left(1 -\right.$$

(see [here]() for more, also search for Jacobian p

Asymmetric peak, a bit lower than th nominal masses.



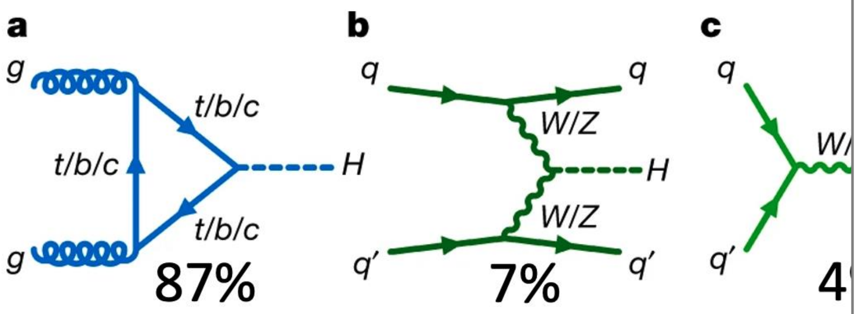


Q: why the eνμν channel has been selected?

m_T [GeV]

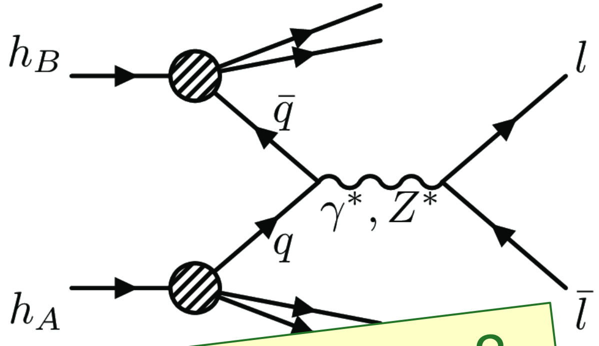
What is the production mechanism?

- The pp collisions are essentially parton collisions where gluons and quarks interact→
- Several different way to produce a particular particle (e.g a Higgs) and a final state.
- Let’s study Higgs production → 4 main mechanisms:

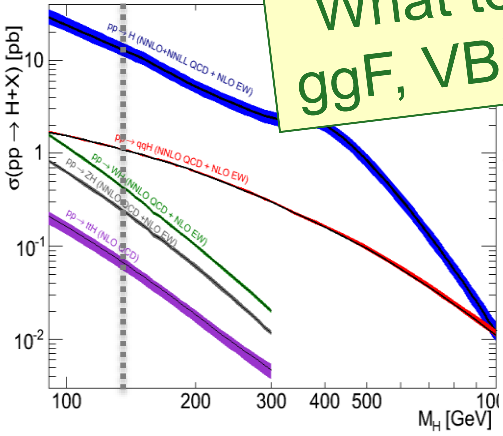


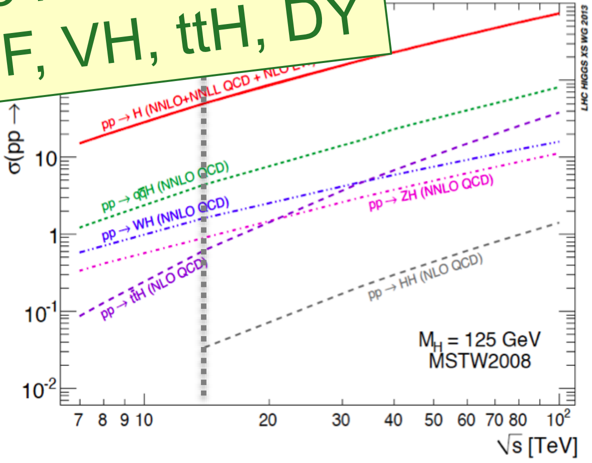
a: gluon-gluon Fusion (**ggF**)
b: Vector Boson Fusion (**VBF**)
c: V-boson association (**VH**)
d: top pair association (**ttH**)
e: single t association (rare)

Another important production mechanism (which does not apply to Higgs production) is the Drell-Yan (DY) production.
This is qqbar-annihilation essentially:



What to remember?
ggF, VBF, VH, ttH, DY



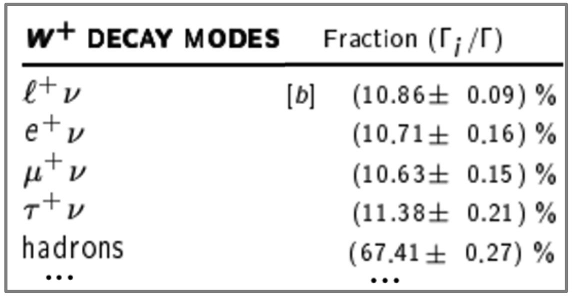


- Final state provides evidences for the production mode, e.g. forward jets in VBF, b-jets in ttH, etc.
- Similar stands for all others W, Z, tt,…
- We rely on MC for the correct productions which is usually inclusive.

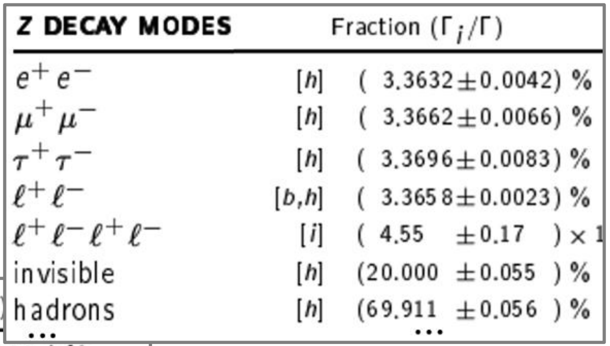
BRs: how to & what to remember

Unstable particles (have decay widths and) decay into two or more particles.
The relative fraction of each decay mode is called Branching Fraction or Ratio (BR or BF).

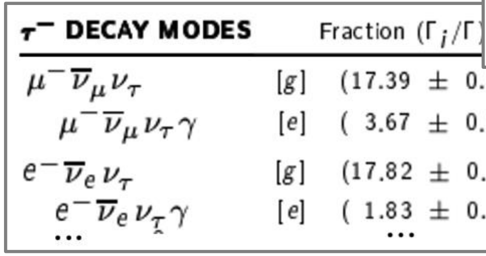
All BRs can be found in PDG
(there is also an app here)
You can find tables like→



W^+ DECAY MODES		Fraction (Γ_i/Γ)
$\ell^+\nu$	[b]	(10.86 ± 0.09) %
$e^+\nu$		(10.71 ± 0.16) %
$\mu^+\nu$		(10.63 ± 0.15) %
$\tau^+\nu$		(11.38 ± 0.21) %
hadrons		(67.41 ± 0.27) %
…		…



Z DECAY MODES		Fraction (Γ_i/Γ)
e^+e^-	[h]	(3.3632 ± 0.0042) %
$\mu^+\mu^-$	[h]	(3.3662 ± 0.0066) %
$\tau^+\tau^-$	[h]	(3.3696 ± 0.0083) %
$\ell^+\ell^-$	[b,h]	(3.3658 ± 0.0023) %
$\ell^+\ell^-\ell^+\ell^-$	[i]	(4.55 ± 0.17) × 1
invisible	[h]	(20.000 ± 0.055) %
hadrons	[h]	(69.911 ± 0.056) %
…		…

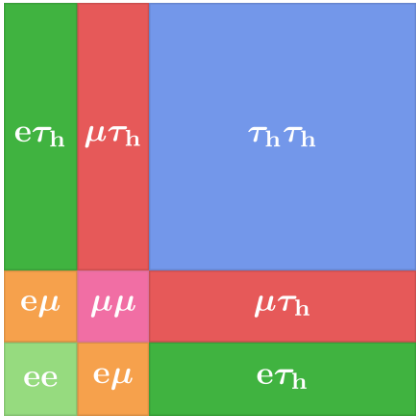


τ^- DECAY MODES		Fraction (Γ_i/Γ)
$\mu^-\overline{\nu}_\mu\nu_\tau$	[g]	(17.39 ± 0.04) %
$\mu^-\overline{\nu}_\mu\nu_\tau\gamma$	[e]	(3.67 ± 0.08) × 10^{-3}
$e^-\overline{\nu}_e\nu_\tau$	[g]	(17.82 ± 0.04) %
$e^-\overline{\nu}_e\nu_\tau\gamma$	[e]	(1.83 ± 0.05) %
…		…

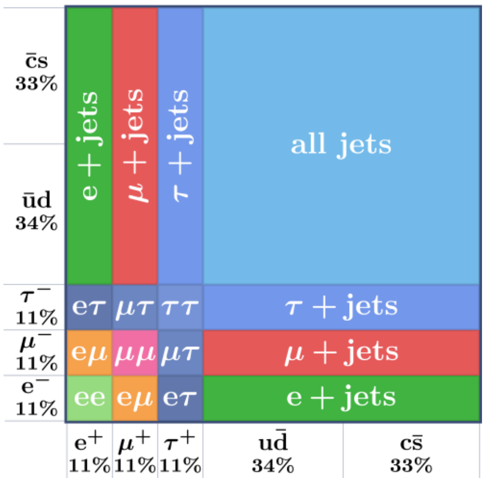
With these BRs, we can estimate the BRs for pairs or particles like tt, ττ, HH, WW, WZ, ZZ, etc.

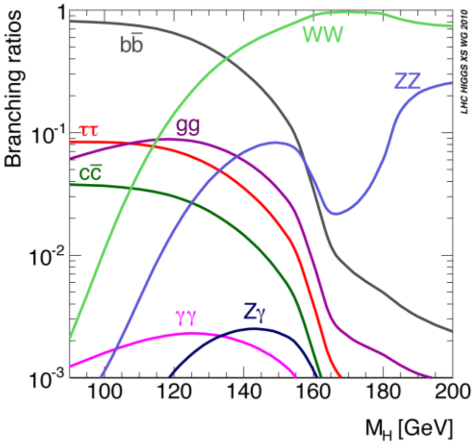
(We can use double-entry matrix like this to help calculations)

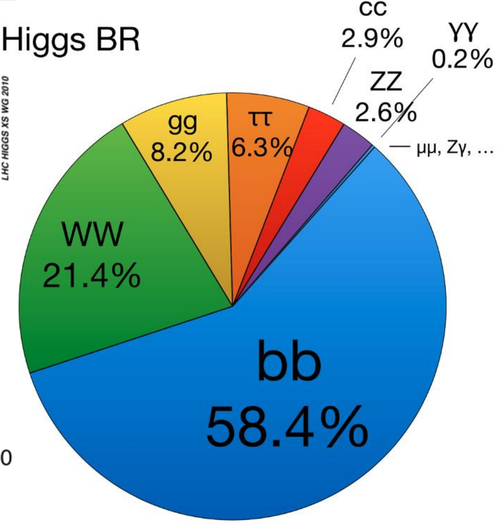
$\tau\tau$ decay modes



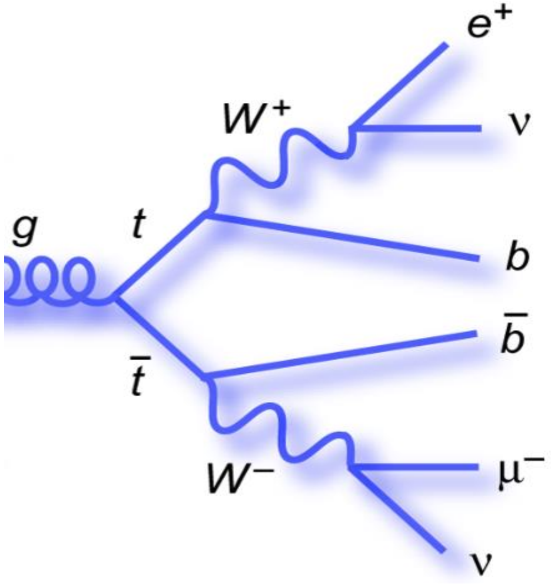
$t\bar{t} \rightarrow b\bar{b}WW$ decay modes







Exercise:
Can you estimate the 0l,1l,2l BRs in tt decays?

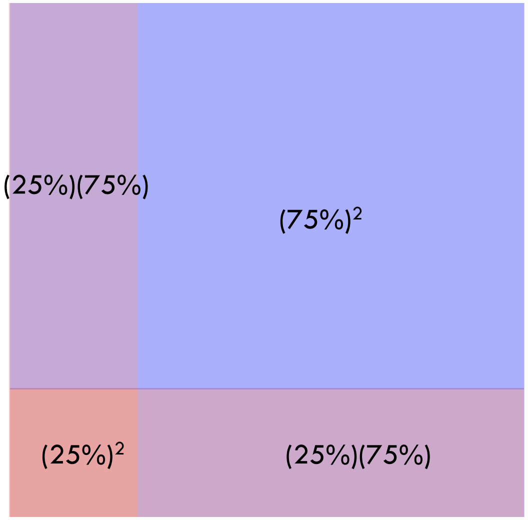


- In CMS we consider “leptons” the e & μ (as are ~stable)
- The τ (unstable) has lep. decays ~35% and hadronic (τ_h) ~65%
- BR(t→bW)~100% (because of CKM matrix…)
- The W has BRs: ~11% μν, ~11% eν, and ~11% x 0.35 $\tau_{lep.}\nu$

→ BR(W→lν) ~ 25%
→ BR(W→jets) ~ 75%

W→lν
25%

W→jets
75%



(25%)(75%)

$(75\%)^2$

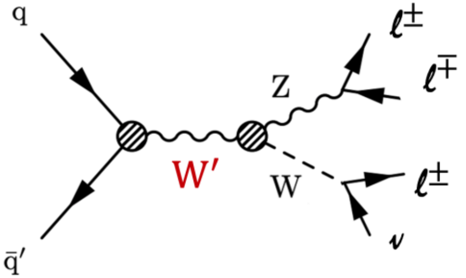
$(25\%)^2$

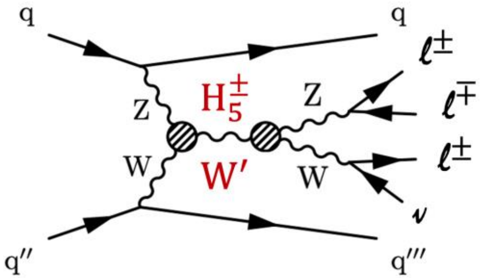
(25%)(75%)

a) BR(WW→0l) ~ $(75\%)^2$ ~ 56%
b) BR(WW→1l) ~ 2 x 75% x 25% ~ 38%
c) BR(WW→2l) ~ $(25\%)^2$ ~ 6%

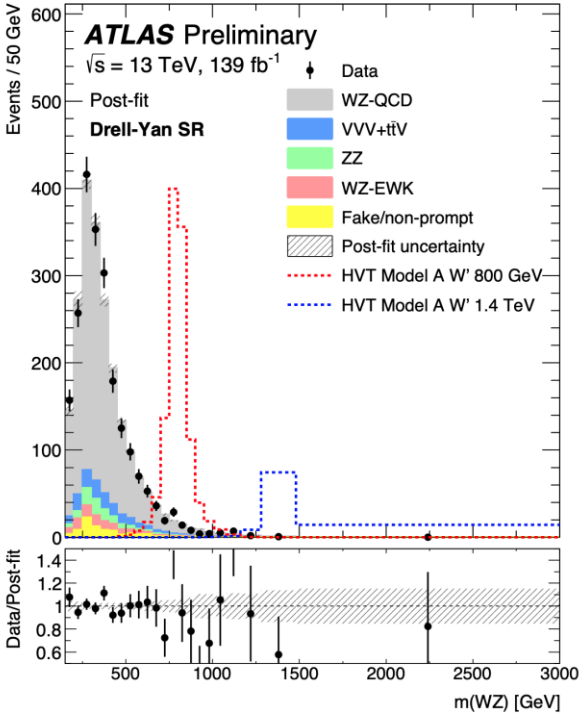
Break

How to read plots, ratios, & limits

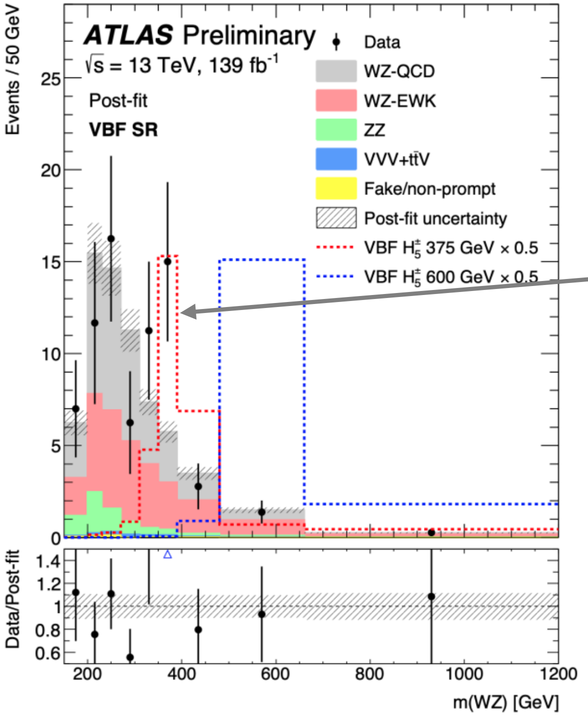




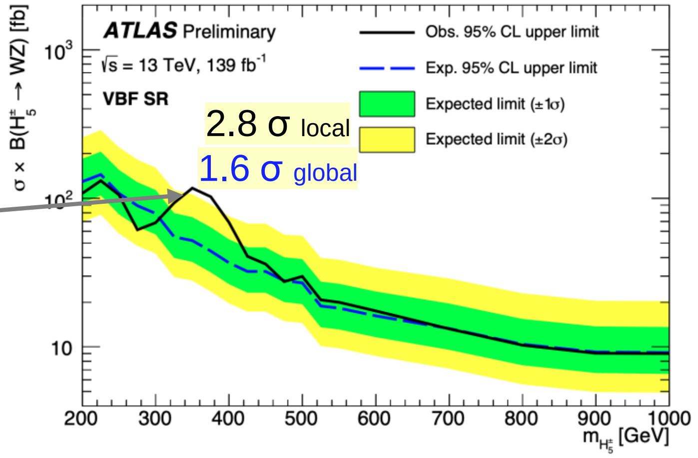
SR for DY



SR for VBF



Limits on $H_5^\pm$ σBR in VBF

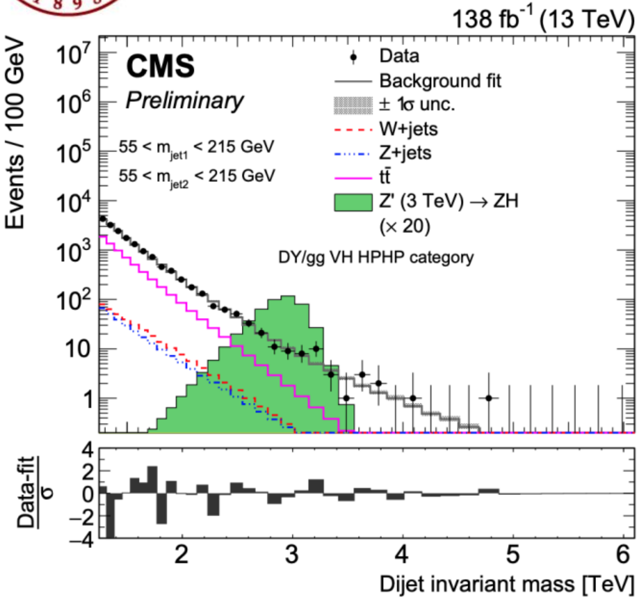


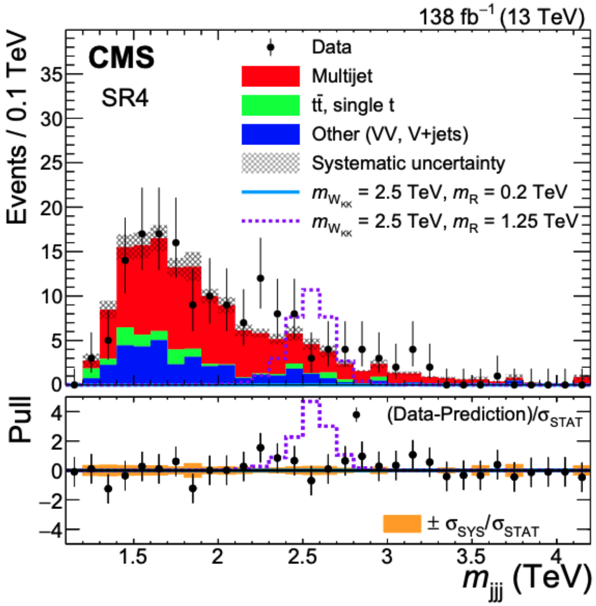
Data & prediction (postfit) are in agreement withing stat.

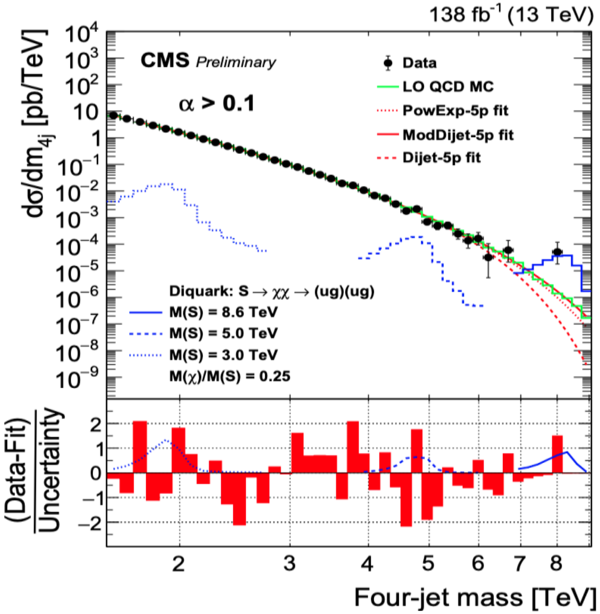
Excess of data events at ~350 GeV

It is important to distinguish the probability to find a fluctuation in some particular location from the probability to find such a fluctuation anywhere. The former is called the **local** significance whereas the latter is referred to as the **global** significance.
LLE is the reason these two are different.

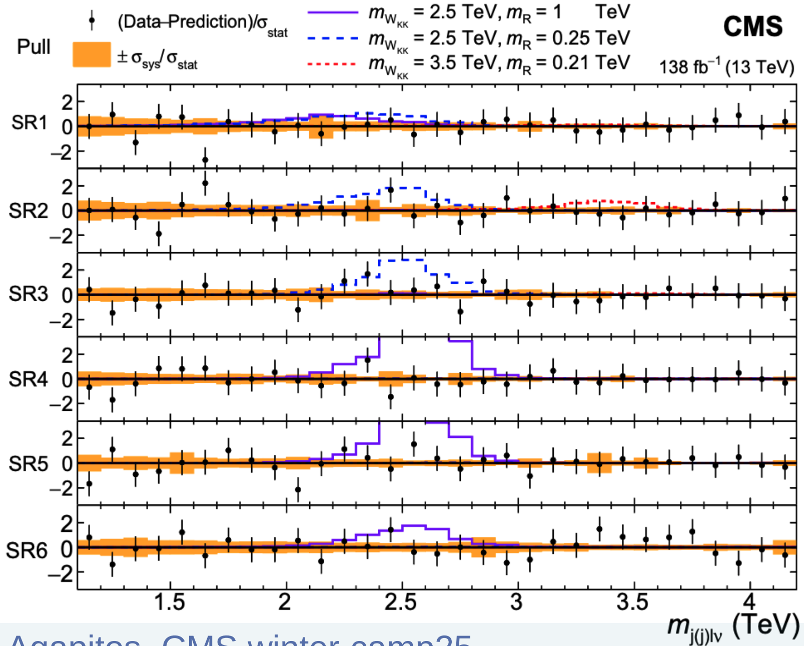
Pulls: (Data – Prediction) / σ





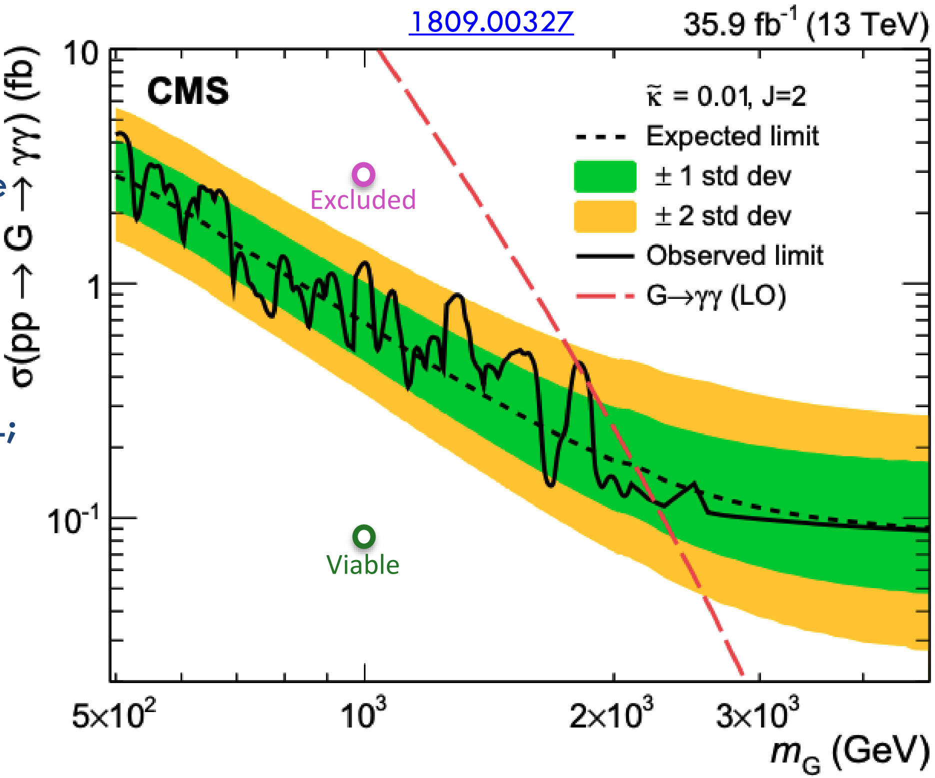


- The **errorbars** are by definition of size 1; this is only to illustrate how far (in s.d.) a point is from the horizontal line of zero.
- If only stat. unc. is present in we expect
  ~68% of points to be withing ±1
  ~95% of points to be withing ±2... etc



How to read limits: on σxB, 1 parameter

- 1 parameter to scan: Graviton (G) mass $m_{G.}$
- We set upper limits on σxBR
  Expected → limit based on prediction without use of data.
  Observed → limit resulted form the data.
- Limits are at "**95% Confidence Level (CL)**".
  → The "signal hypothesis" is excluded at 95% CL; i.e. 1 out of 20 repetitions would result in the opposite (on average).
- Area **above lines** (σxB here) is excluded.
- Area **below lines** is "viable", i.e. might hide the process but statistics is not enough to probe it. → These are **upper limits.**
- The expected limit evaluation comes with ±1σ and ±2σ standard deviation belts.
  (These comes from the likelihood scan…)

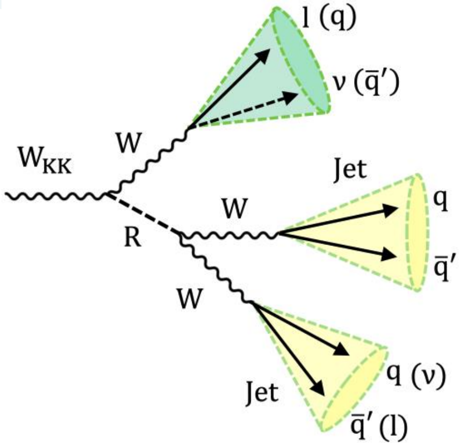


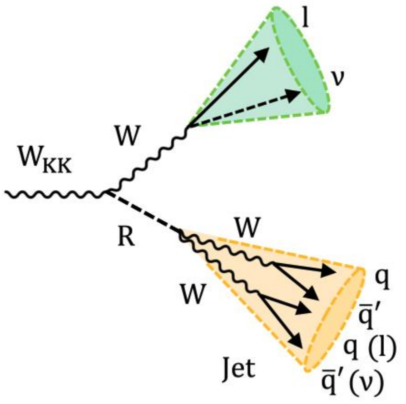
- Spikes (deeps) in observed limit indicate local excesses (deficits) of data events.
- Theory model (red) predict the σxB vs m_G; here m_G below ~2100 GeV are excluded. This is taken from intersection theory&obs.

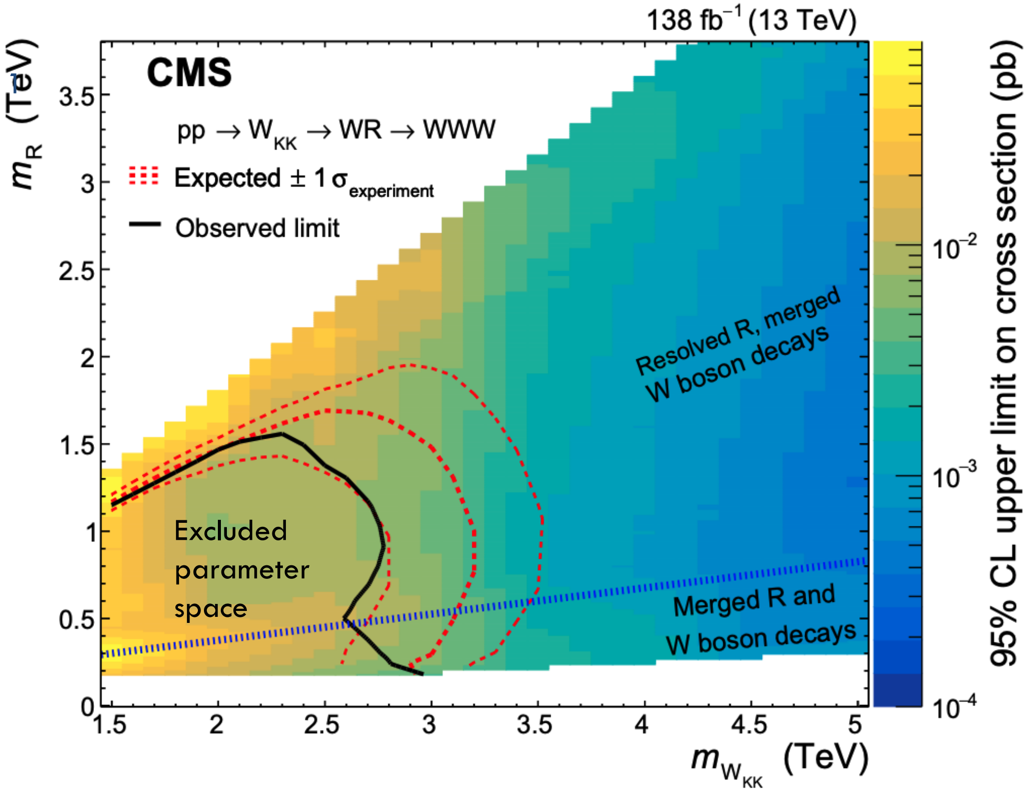
Limit at m_G=500 GeV is 4 fb → we exclude scenarios which predict 4x36xEff. events at 95% CL
Limit at m_G=4 TeV is 0.09 fb → we exclude scenarios which predict 3 events (=0.09x36xEff.)

How to read limits: 2 parameters (masses)

- 2 new particles W_{KK}, R
  → 2 masses to scan for setting limits
- We set **upper limits** on σxB for all masses
  these are model independent.
- We set **lower limits on W_{KK}, R** masses,
  this is shown as a contour line (it was a single point in parameter case).
- Area lower-left of the curve is the excluded parameter space.
  This is model dependent result.
- Observed and expected are shown:
  Expected has only ±1σ band (in red→)
- Here we have observed ~1σ weaker than the expected (due to excesses in data).

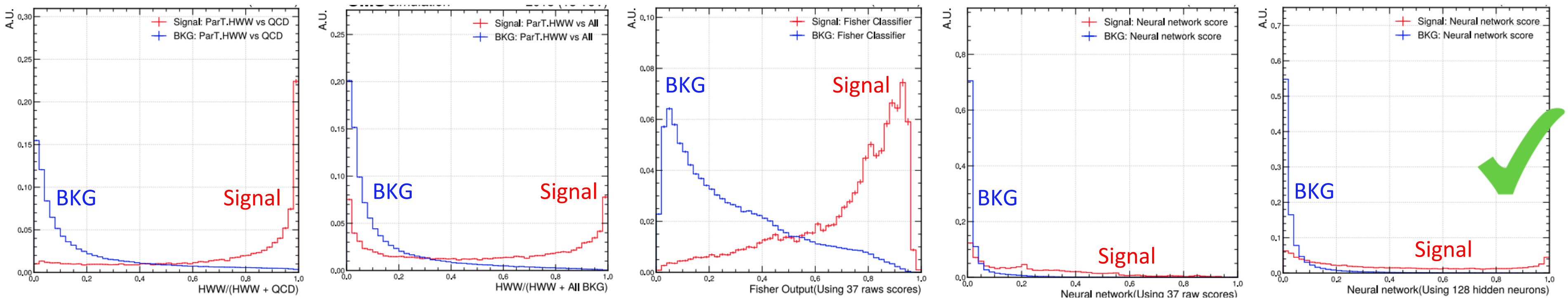


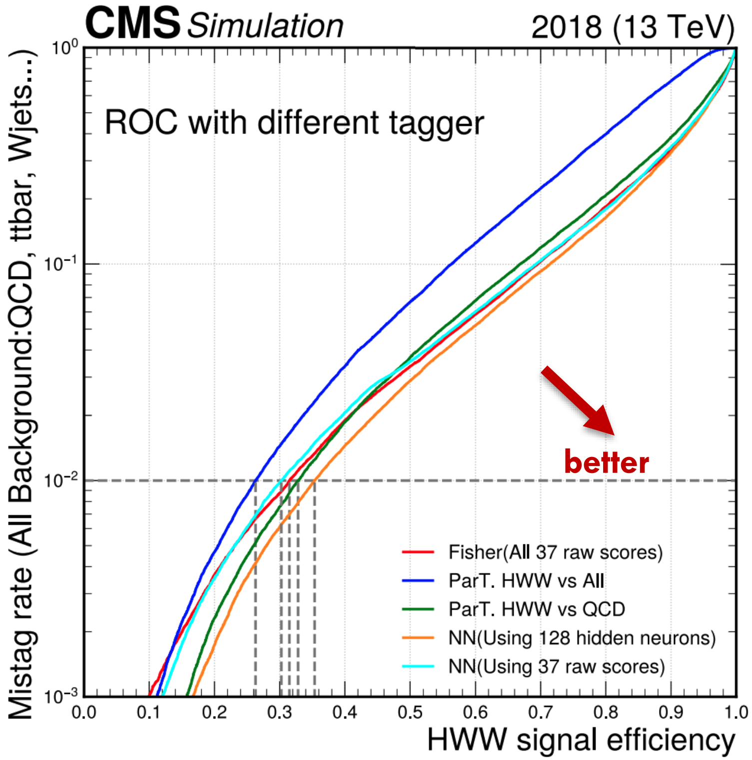




How to read ROCs

Assume we have 5 variables (classifiers) which separate signal from BKG



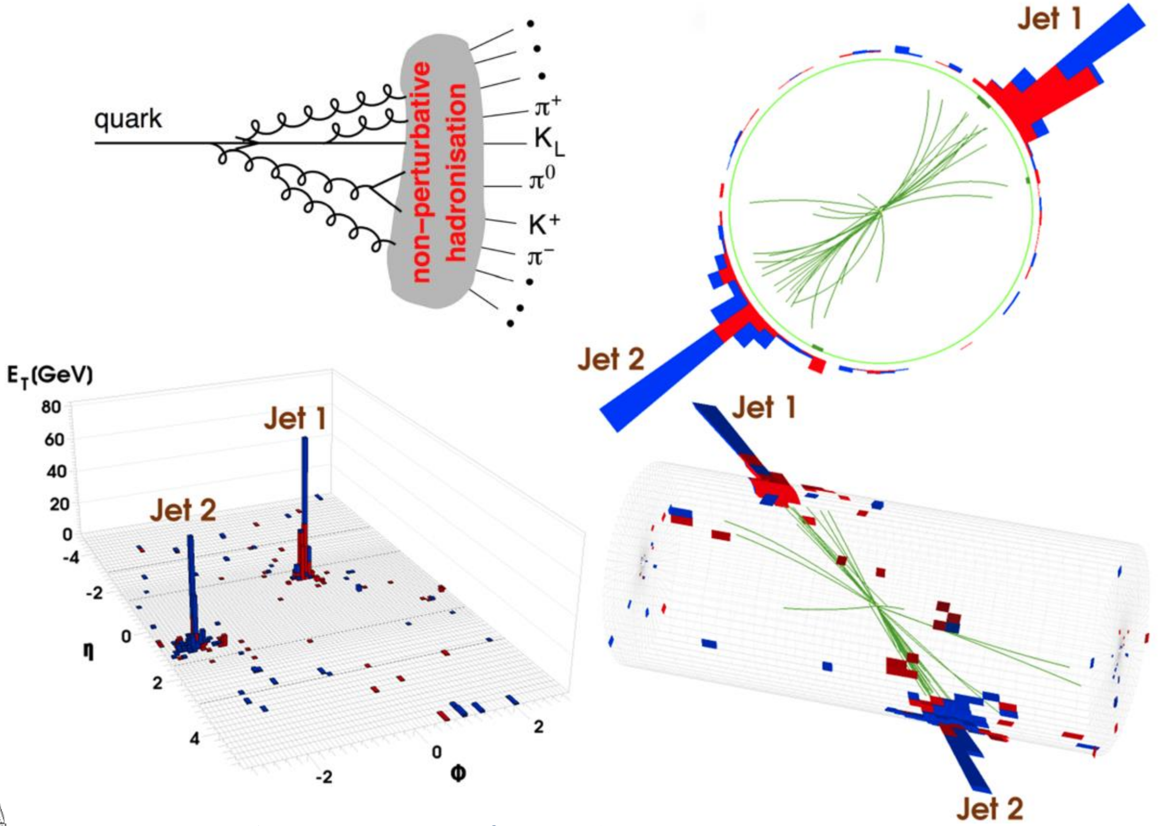


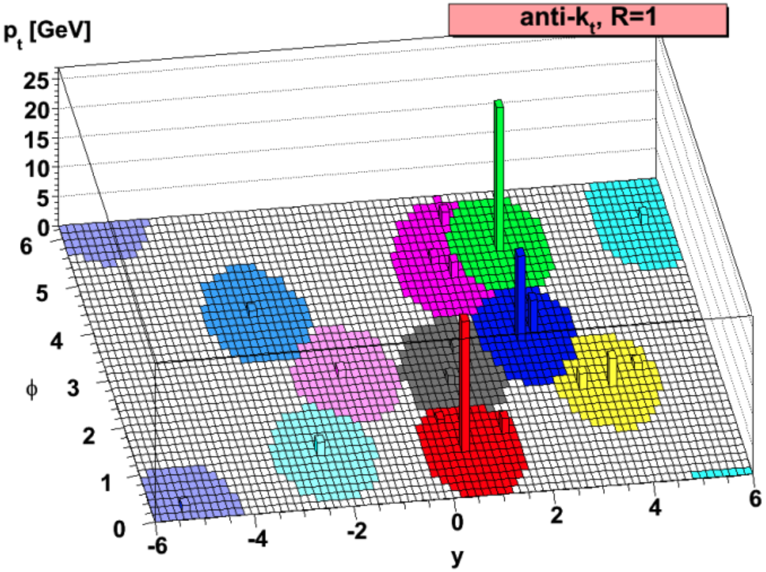
- ROCs: Receiver Operation Characteristics
- Use in binary classification problems to evaluate prediction performance (separation power) of a variable.
- BKG eff. vs. Signal eff. or similarly:
  "1-False positive" vs. "True positive" signal
- The higher the $S/\sqrt{B}$ the better the performance.
- In this example the orange (NN 128), it has the best performance at 10^{-2} BKG rejection rate.

Jets reconstruction & clustering

- Quarks & gluons hadronized into "sprays of long-lived" particles. These propagate in similar directions to the initial q/g forming jets.

- We use the anti-kT algorithm to cluster individual particles (PF candidates) into jets
  → PFJets ***(using clustering par. R)***





- ***Details & illustrations here.***
- ***10k-citations paper!***

- Anti-kT algo. forms "conical" jets,
  → i.e. circles in η-ϕ plane with radius "R".
  (This R is the "clustering parameter").

- Circle's center is the p^μ direction of jet.

- The η-ϕ distance of 2 jets is $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ > R=1.0

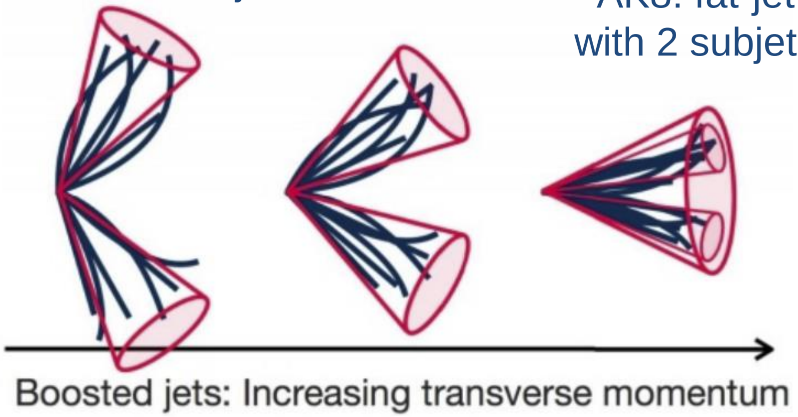
Jets substructure & “tagging”

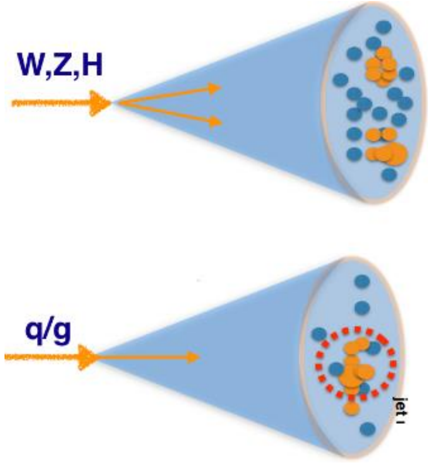
Boosted objects → small angular separation of the products → merged to one jet

Boost should be significant: $\Delta R \sim 2m_j/p_{Tj} < 0.8 \rightarrow p_T(W,...) > 200$ GeV for merging.

AK4: narrow jets

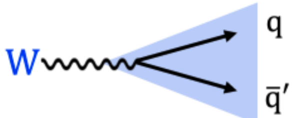
AK8: fat-jet with 2 subjets

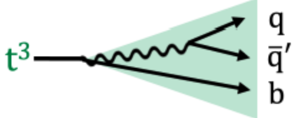


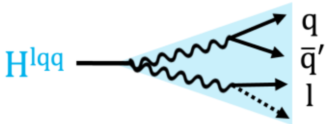


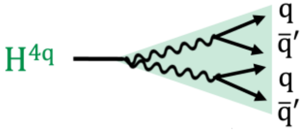
We can make use of the different substructure and separate QCD from W/Z/H/t originated jets

- 2 sub-jets: W/Z → qq, H → bb/cc, H → ττ
- 3 sub-jets: t→ bW → bqq, H → WW → lqq
- 4 sub-jets: H → WW/ZZ → 4q



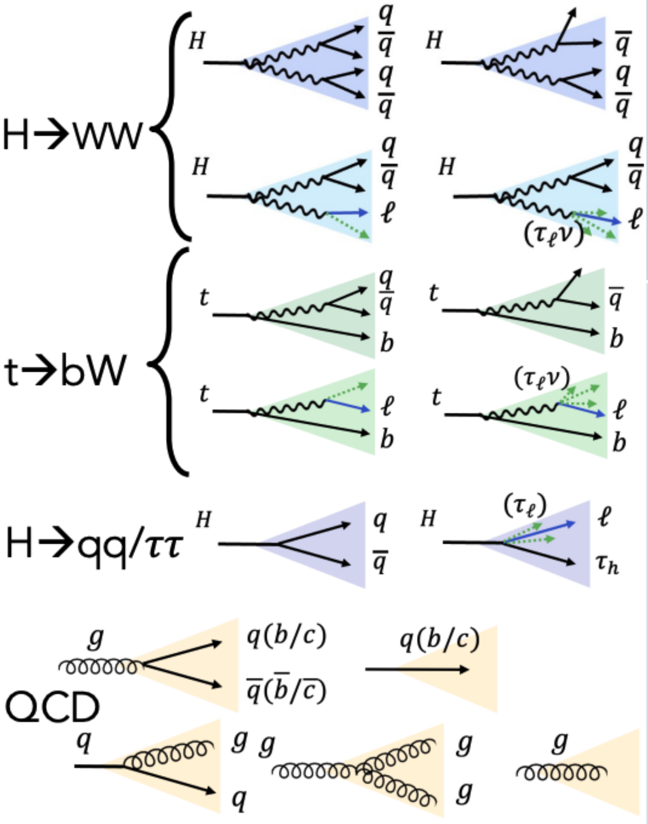






ParticleTransformer Classification

- 37 classes in total:
  16 Signal + 21 Background
- Link of Congqiao's talk in JMAR meeting.



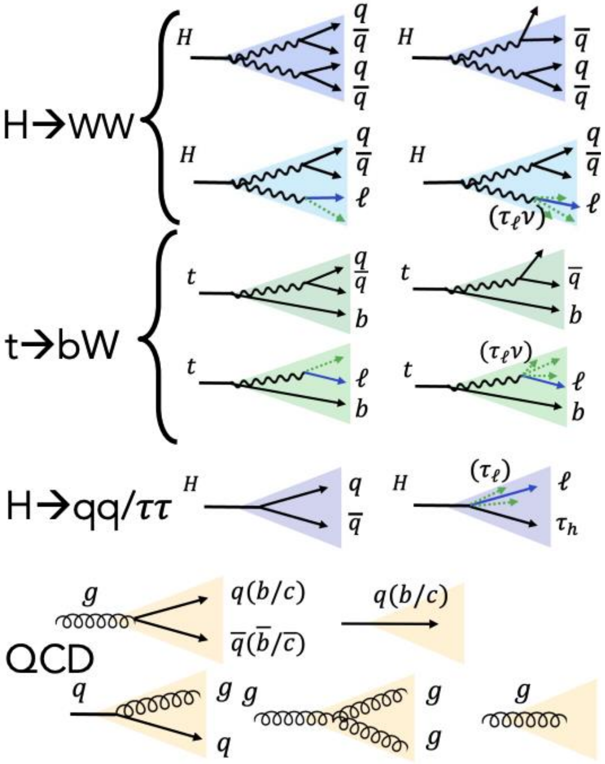
ParticleTransformer(abbr. ParT) 37 raw scores

Signal Category	Label
H → WW full-hadronic	3q(0c)
	3q(1c)
	3q(2c)
	4q(0c)
	4q(1c)
	4q(2c)
H → WW semi-leptonic	eνqq(0c)
	eνqq(1c)
	μνqq(0c)
	μνqq(1c)
	e3νqq(0c)
	e3νqq(1c)
	μ3νqq(0c)
	μ3νqq(0c)
	τ_hνqq(0c)
	τ_hνqq(1c)

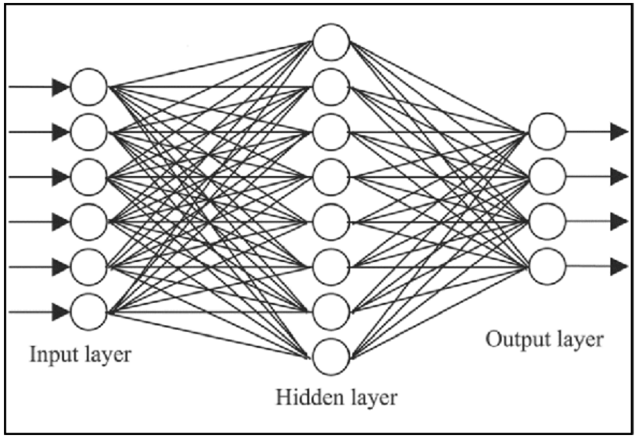
Background Category	Label
t → bW hadronic	bq(0c)
	bq(1c)
	bqq(0c)
	bqq(1c)
t → bW leptonic	beν
	bμν
	bτ_hν
	bτ_eν
	bτ_μν
H → qq	bb
	cc
	qq (q=u/d)
	ss
H → ττ	$\tau_h\tau_h$
	$\tau_h\tau_e$
	$\tau_h\tau_\mu$
QCD	b
	bb
	c
	cc
	others

NN classification & jet identification

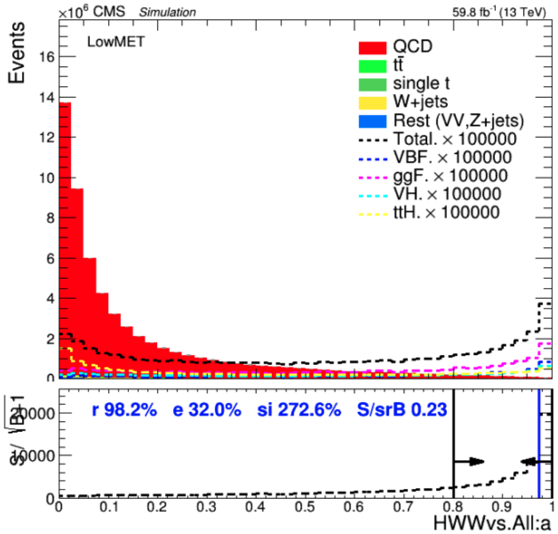
Input: all PF info
from MC simulation

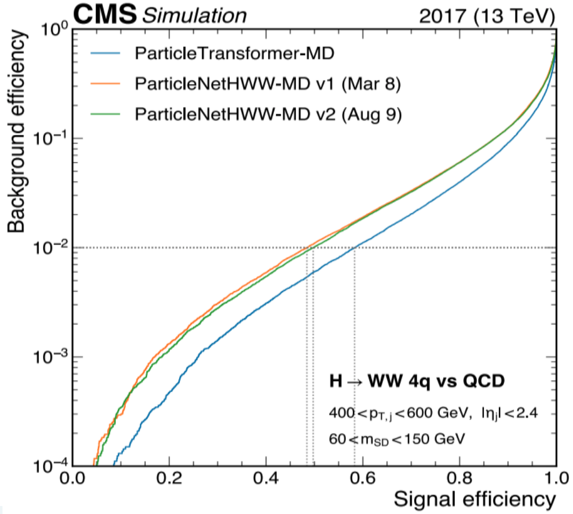


Training on
Graph NN or
attention-based NN
or Deep learning NN
(training with labels)



Output: classes scores
which sum to 1
We form ratios of scores
and use them as taggers!

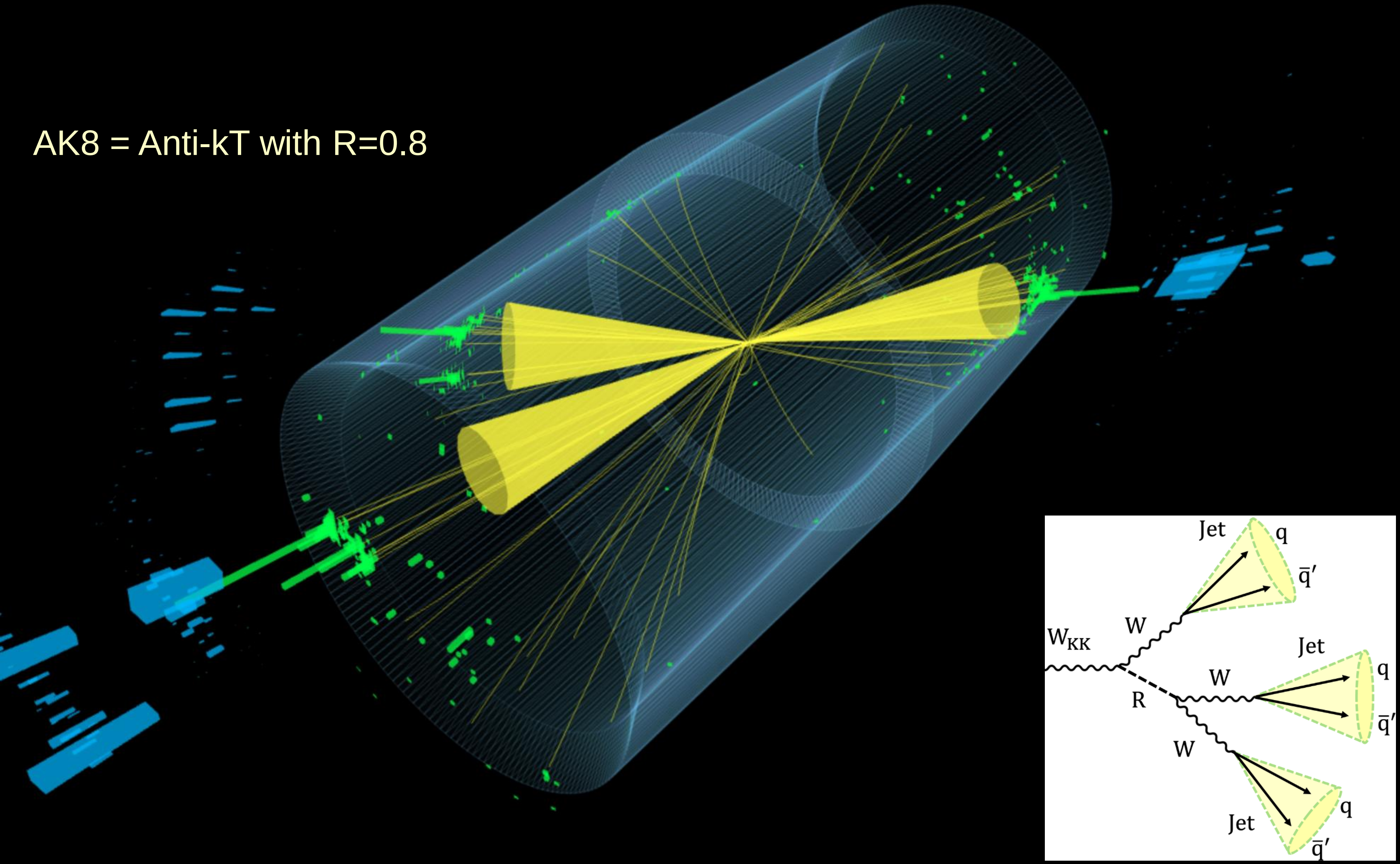




The "full story" in Congqiao's talk earlier today!

Event with 3 AK8 jets, each with 2-prong substructure, all 3 tagged as originated from W→qq with Deep-AK8 algorithm

AK8 = Anti-kT with R=0.8



Let’s examine real analyses examples

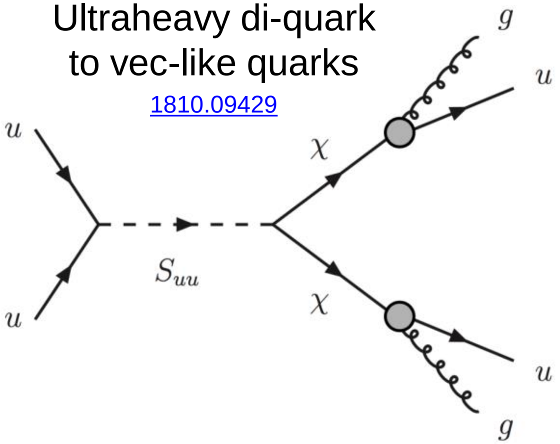
EXO-21-010

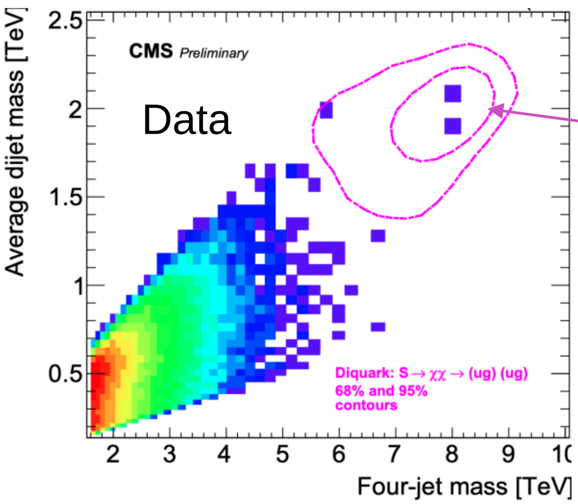
2 pairs of 2 jets: A → XX → (jj)(jj)

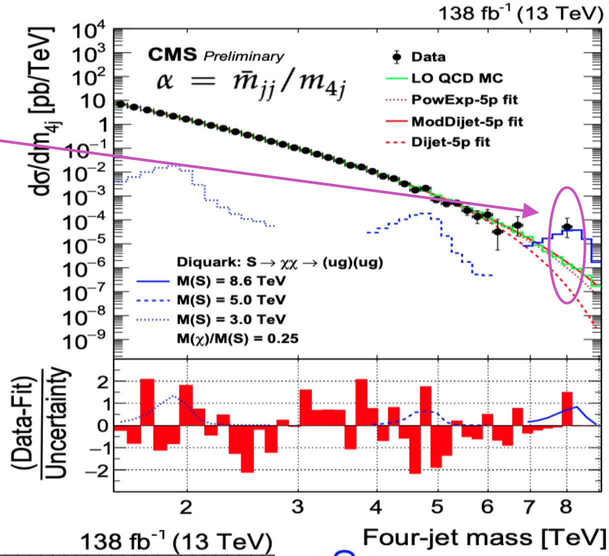
- 4 narrow (AK4) jets → paired to 2 di-jets, symmetrized masses, we select $\frac{|m_{jj1}-m_{jj2}|}{m_{jj1}+m_{jj2}} < 0.1$.
- Search over: m_{4j} and average di-jet mass $\bar{m}_{jj}$; fit param. function to the data in slices of $\frac{\bar{m}_{jj}}{m_{4j}}$.

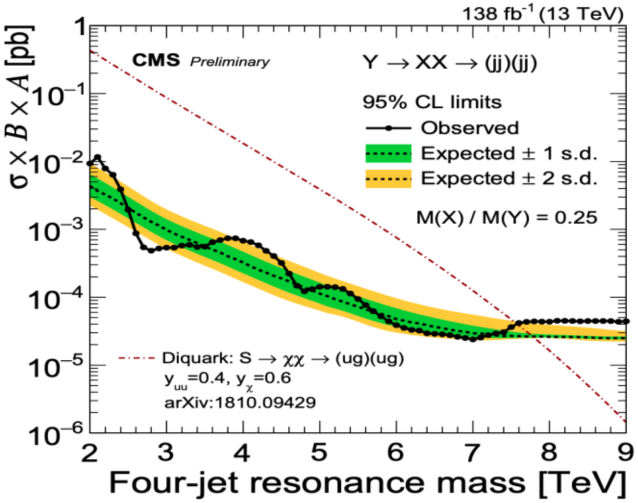
Ultraheavy di-quark to vec-like quarks

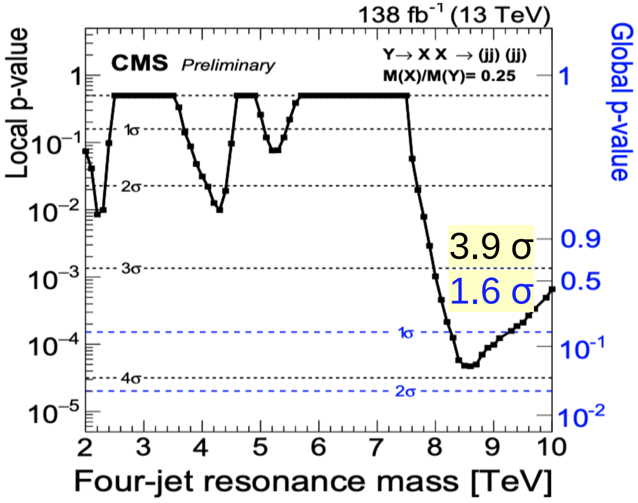
1810.09429







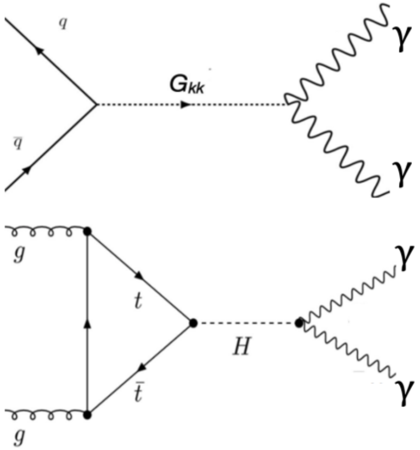




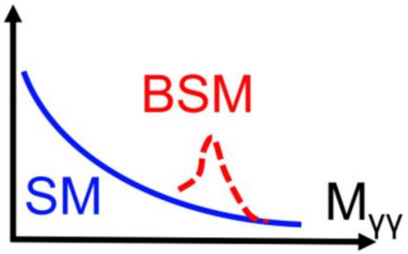
3.9 σ

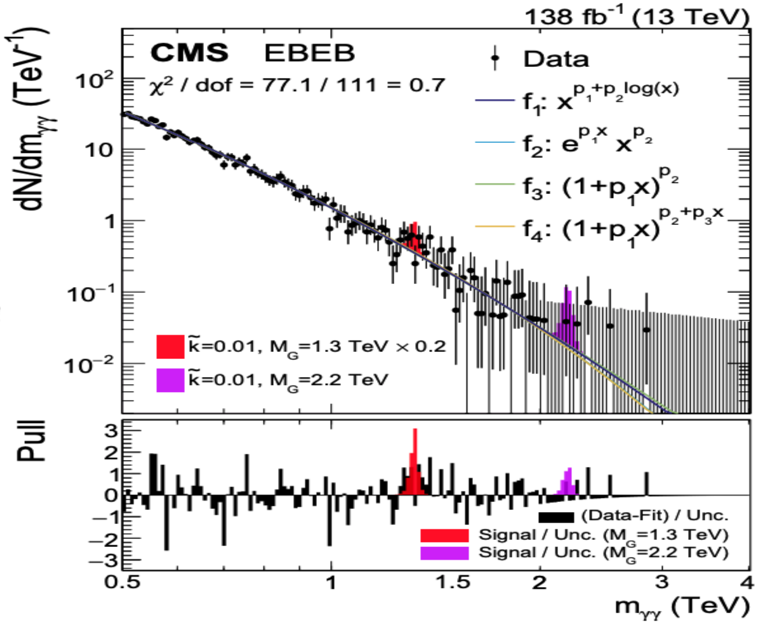
1.6 σ

Search for new physics in diphoton spectrum

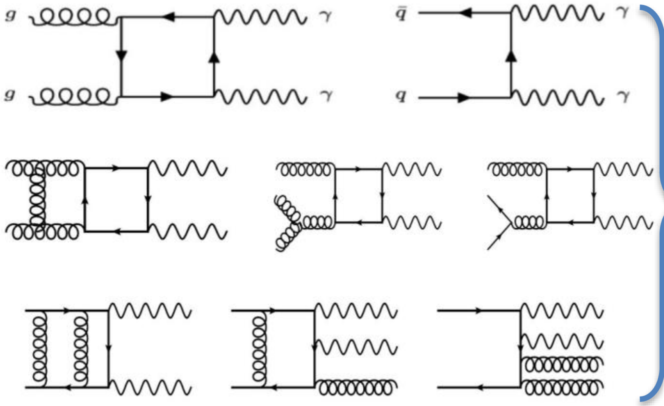


2405.09320

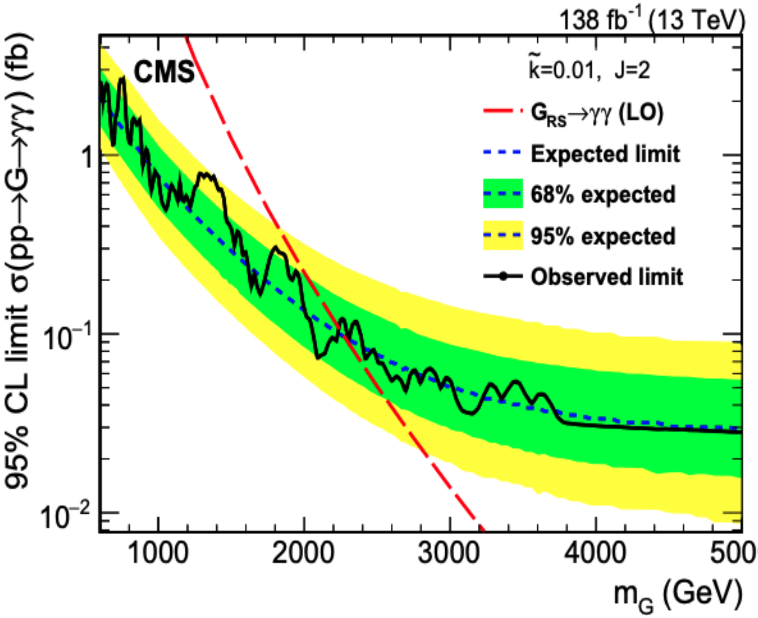




1. Signal from (Bulk RS) Warped Extra Dim. models (Radion or Graviton) or Heavy Higgs.
2. Simple selection: 2γ with p_T>125GeV, at least one central |γ|<1.44, γγ-trigger.
3. Background processes: we do not simulate them, we use parametric functions instead:



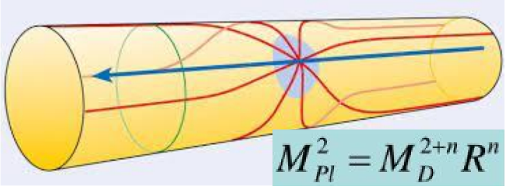
$$
\begin{aligned}
f_1(x) &= p_0 x^{p_1+p_2\log(x)},\\
f_2(x) &= p_0 e^{p_1 x} x^{p_2},\\
f_3(x) &= p_0 (1+x p_1)^{p_2},\\
f_4(x) &= p_0 (1+x p_1)^{p_2+p_3 x},
\end{aligned}
$$

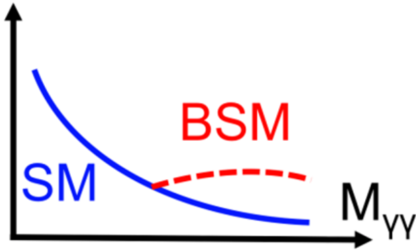


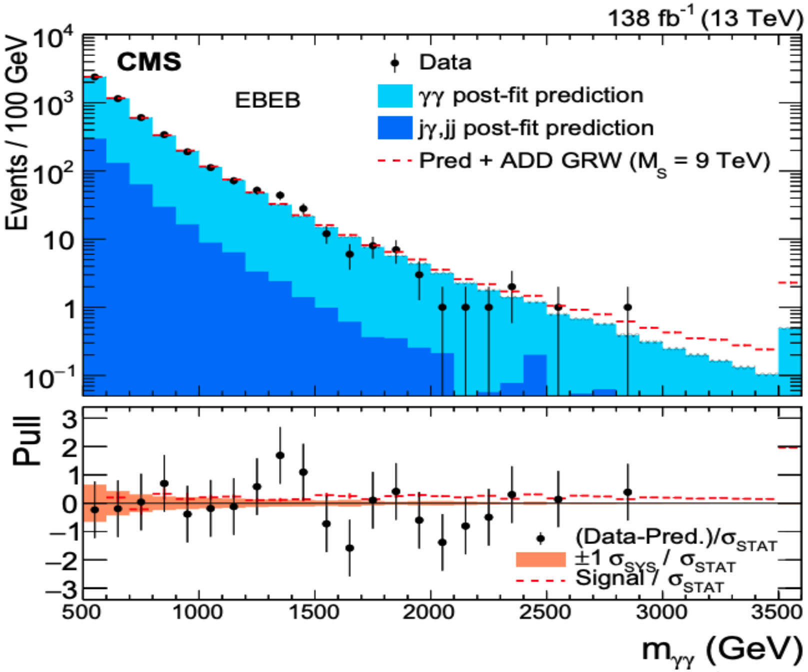
Search for new physics in diphoton spectrum

2405.09320

- ADD large ED model



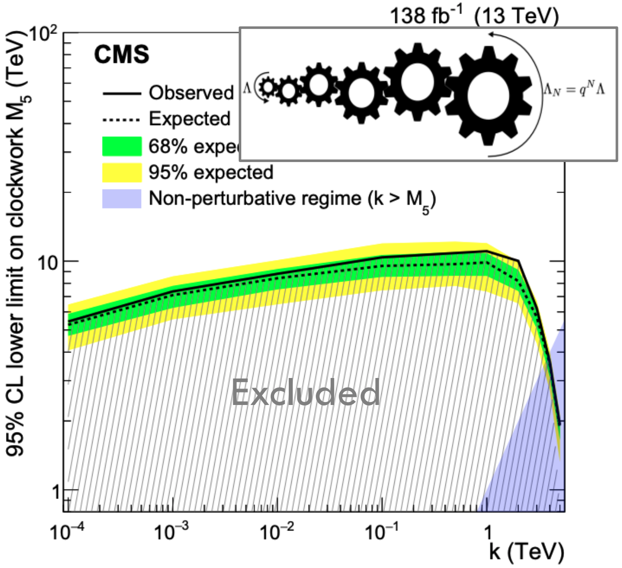




- Select γγ events.
- EBEB (here→) and EBEE categories
- QCD BKG prediction:
  - γγ: Sherpa scaled at NNLO with MCFM.
  - jγ,jj = fakes: 10-30%, data-driven with fake rate.
- Fit the two binned $m_{\gamma\gamma}$ spectra in range 0.5-4 TeV.
- Lower limits on M_S (or Λ_T) scale vs number of ED: (~11 TeV)

Signal:	GRW	Hewett		HLZ				
		negative	positive	$n_{\text{ED}} = 3$	$n_{\text{ED}} = 4$	$n_{\text{ED}} = 5$	$n_{\text{ED}} = 6$	$n_{\text{ED}} = 7$
Expected:	$8.7^{+0.7}_{-0.6}$	$7.3^{+0.3}_{-0.3}$	$7.8^{+0.6}_{-0.5}$	$10.3^{+0.8}_{-0.7}$	$8.7^{+0.7}_{-0.6}$	$7.9^{+0.6}_{-0.5}$	$7.3^{+0.6}_{-0.5}$	$6.9^{+0.6}_{-0.5}$
Observed:	9.3	7.1	8.3	11.1	9.3	8.4	7.8	7.4

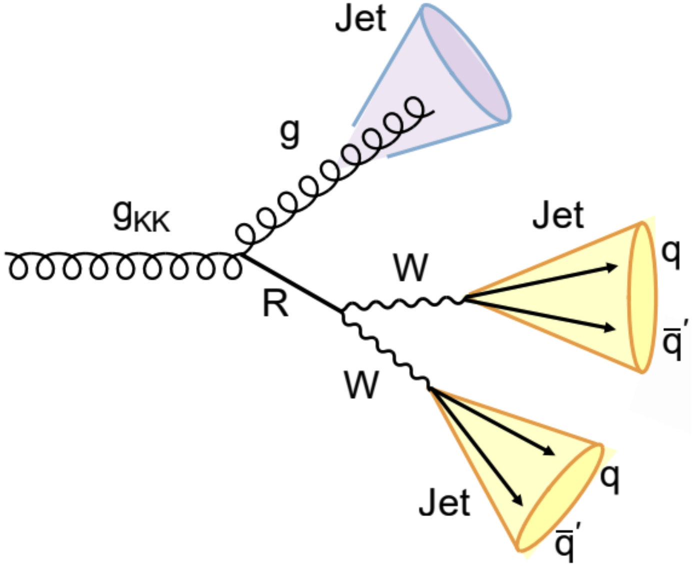
- Interpretation on Continuum Clockwork Mechanism →
  Constrains on M_5 mass vs clockwork spring “k”.

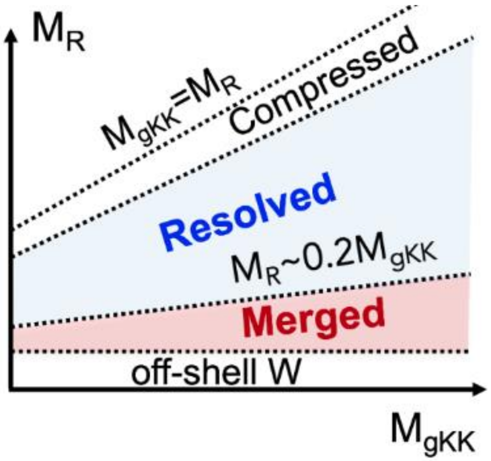


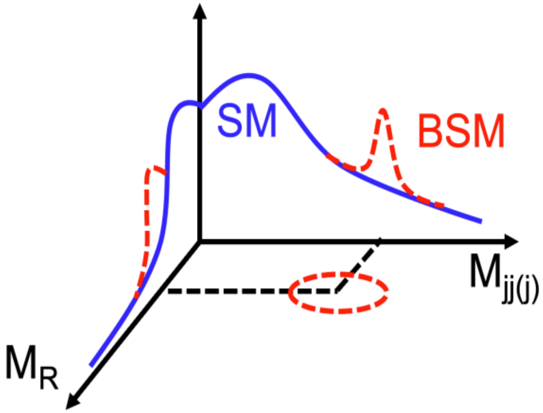
Search for new resonances $g_{KK} \to gR \to gWW$

2410.17303
B2G-23-004

- Use signal from Extended Warped ED model, (2201.08476, 2112.13090) where the process: $g_{KK} \to gR \to gWW$ is dominant.
- We focus on the 0l channel: $g_{KK} \to gR \to gWW \to$ jets (BR~56%).







Strategy:

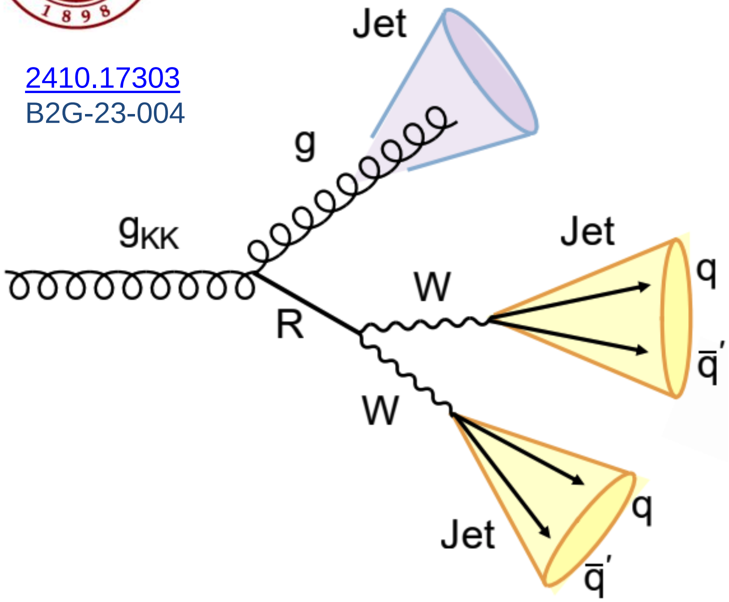
1. Tri-jet selection,
2. identify (tag) 2 jets as W-candidates with PNet,
3. form m_{jj} (R) and m_{jjj}, (g_{KK}),
4. bin over m_{jj}, fit m_{jjj}.

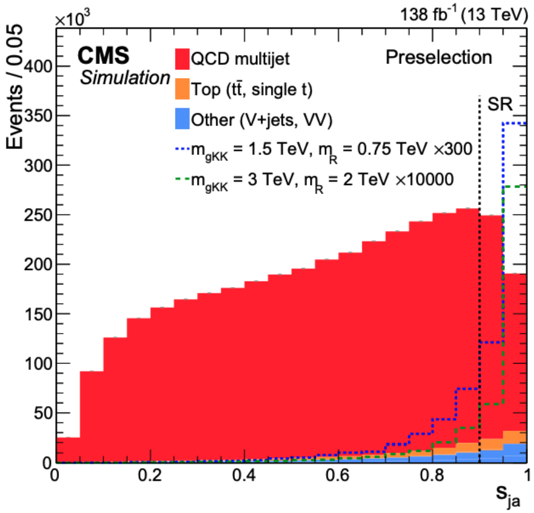
Selection basics:

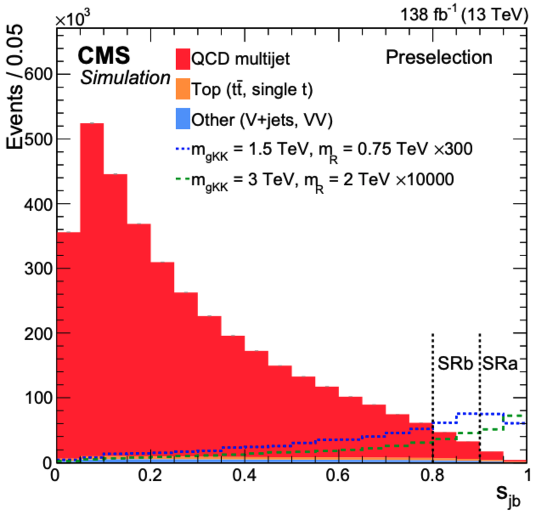
1. $N_{j\text{-}AK8}=3$, $N_{lep}=0$,
2. $p_{Tj1(j2,j3)} > 400\ (200)$ GeV, $|\eta_j| < 2.4$,
3. $m_{ja,jb} > 50$ GeV,
4. $H_T > 1.1$ TeV.

W→qq identification with ParticleNet tagger

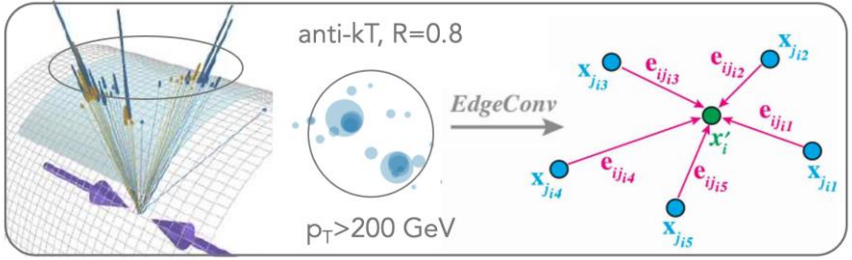
2410.17303
B2G-23-004



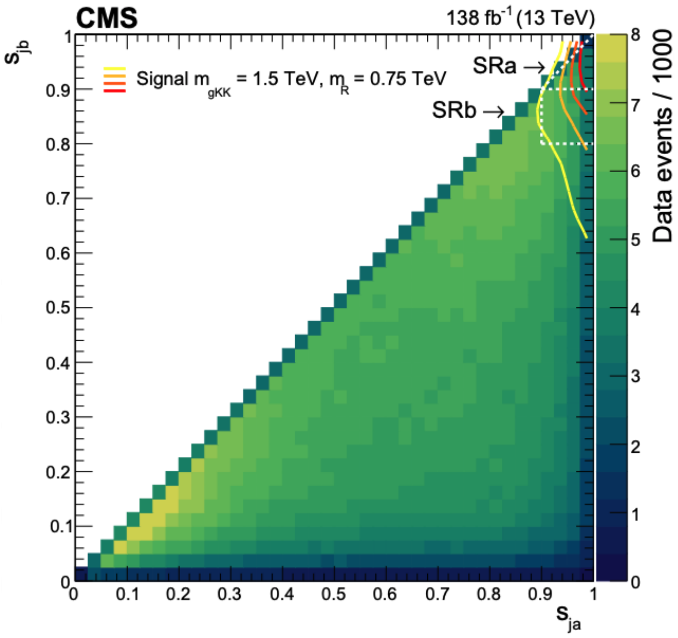




- W→qq identification with ParticleNet tagger 1902.08570



  - Graph NN, treat jets as particle cloud
  - Convolution on point clouds (EdgeConv 1801.07829)
- Tagger score: $s_j = p(W\to qq) / [p(W\to qq) + p(QCD)]$
- Define SRa and SRb based on scores→



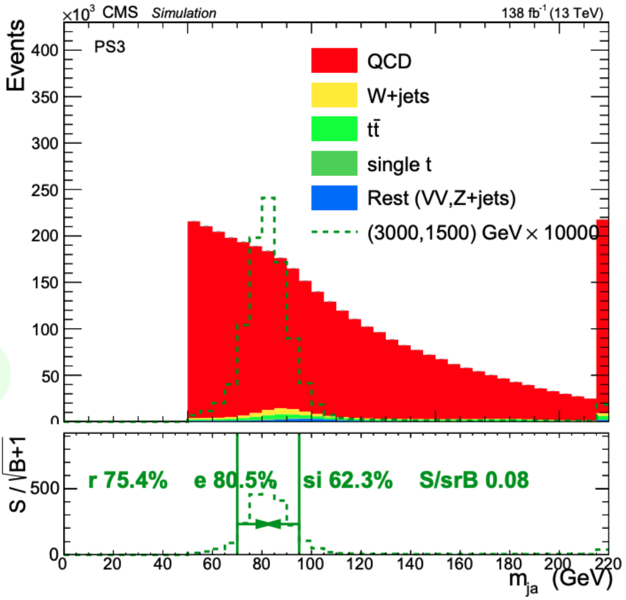
W→qq identification with Soft-drop masses

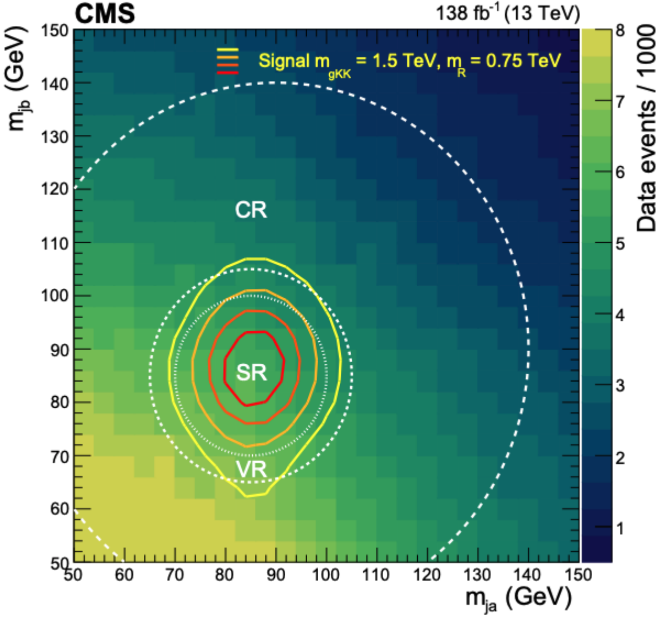
- Two highest PNet score jets: j_a, j_b are assigned as W-cand. (gluon is j_c).
- Demand SD masses: $m_{ja,jb}$ to be on W-mass peak:

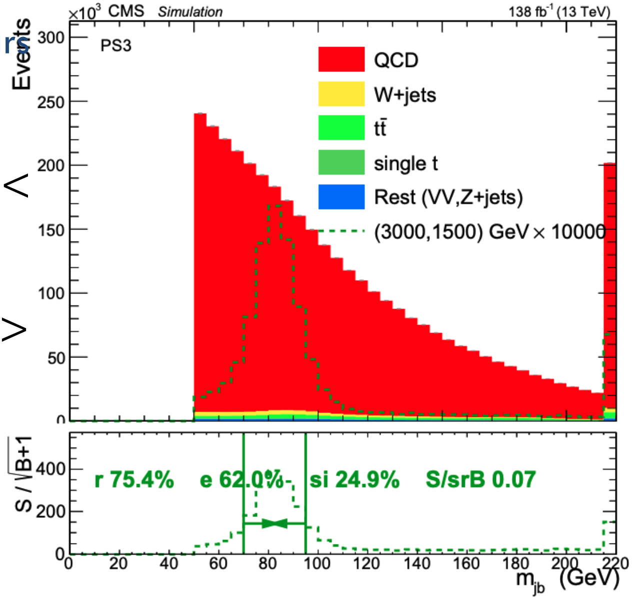
$$m_{85} \equiv \sqrt{(m_{ja}+85)^2+(m_{jb}+85)^2} < 15\ \text{GeV}$$

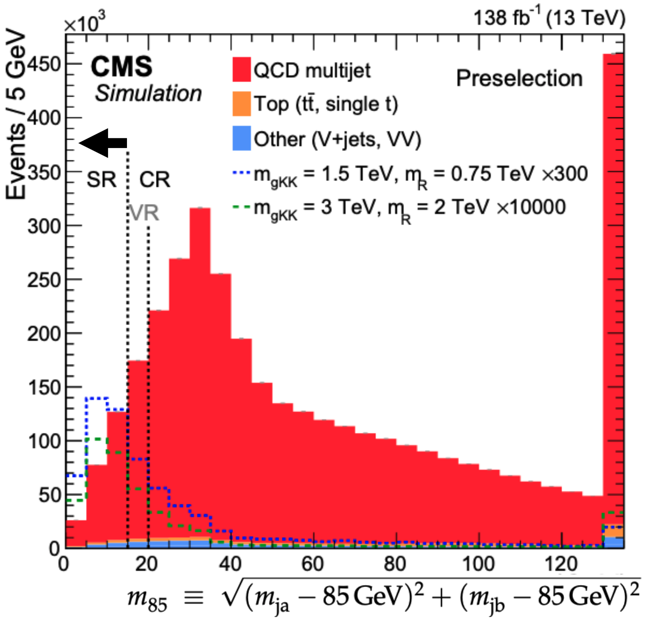
→ Circular area has better performance than rectangular.
→85 GeV values optimal
→ Cut value of 15 GeV appears optimal for all points tested.

- Signal Regions (SRs) have: $m_{85} <$ 15 GeV.
- Control Regions (CRs) are: $m_{85} >$ 15 GeV & $m_{90} < 50$ GeV
- Validation Regions (VRs): $15 < m_{85} < 20$ GeV.









R, gKK reconstruction & SR binning

- M_R reco. from j_a, j_b:

$$m_{jj}* \equiv m_{jj} - mja - m_{jb} + 2(85\ \text{GeV})$$

- M_{gKK} reco. from j_a, j_b, j_c:

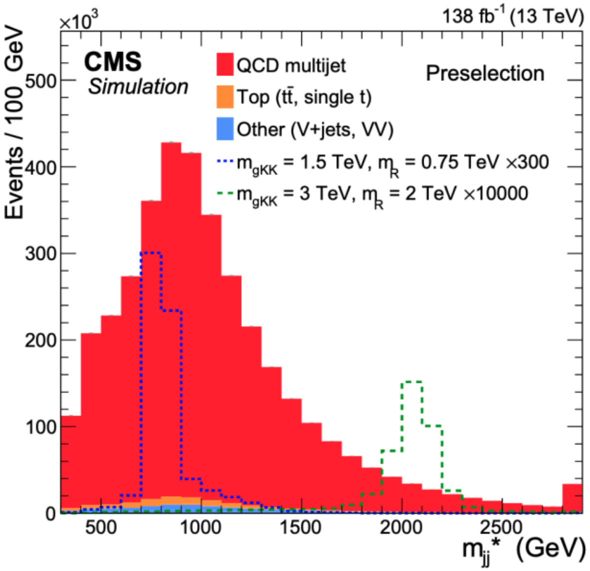
$$m_{jjj}* \equiv m_{jjj} - m_{ja} - m_{jb} + 2(85\ \text{GeV})$$

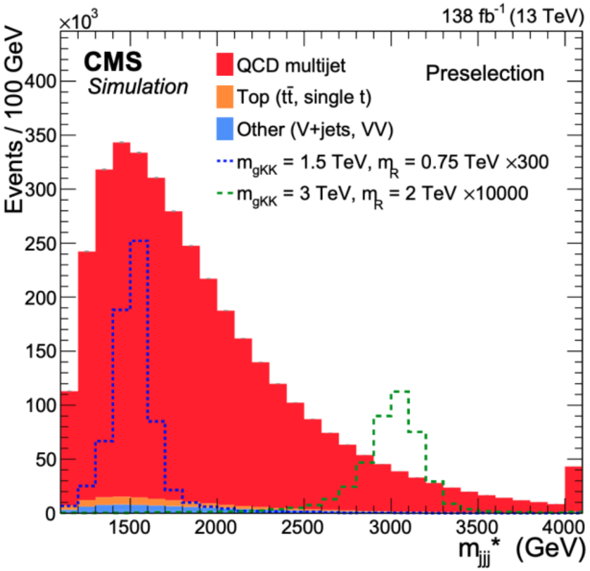
- → i.e. we correct invariant masses to mitigate reso. effect from jet SD masses.
  → sharper peaks (see Fig.4).
  → ~3% significance gain.

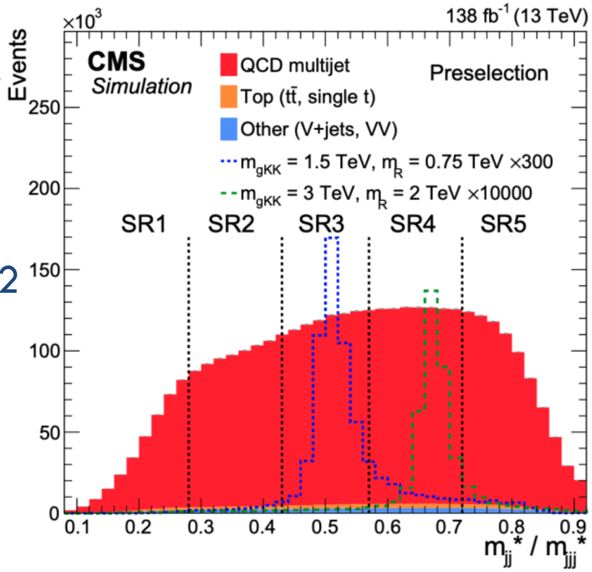
- From ratio $m_{jj}*/m_{jjj}*$ and define 5 bins SR1—5 →
- Effectively binning over m_R.

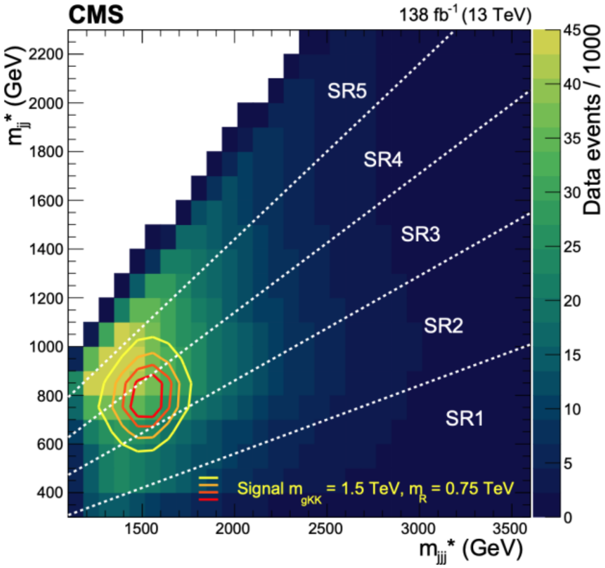
- In each of these 5 SR we have 2 SRs (SRa, SRb) based on PNet scores.

- Thus, we have 10 SRs.
  → We fit the $m_{jjj}*$ spectra.





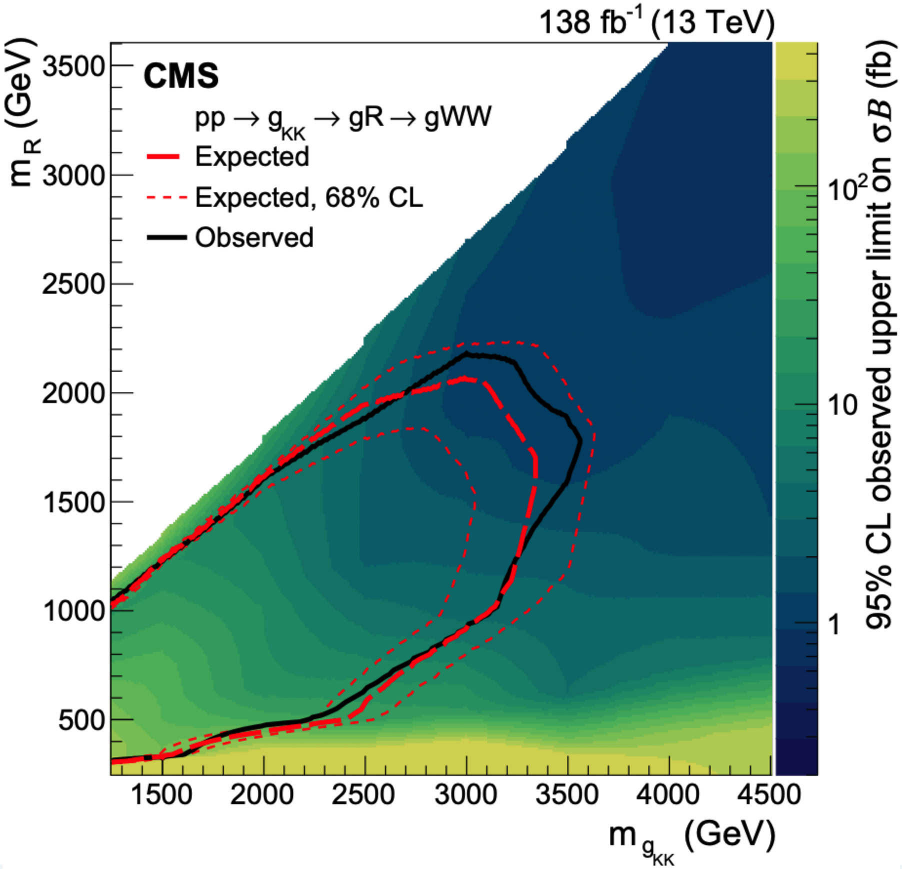


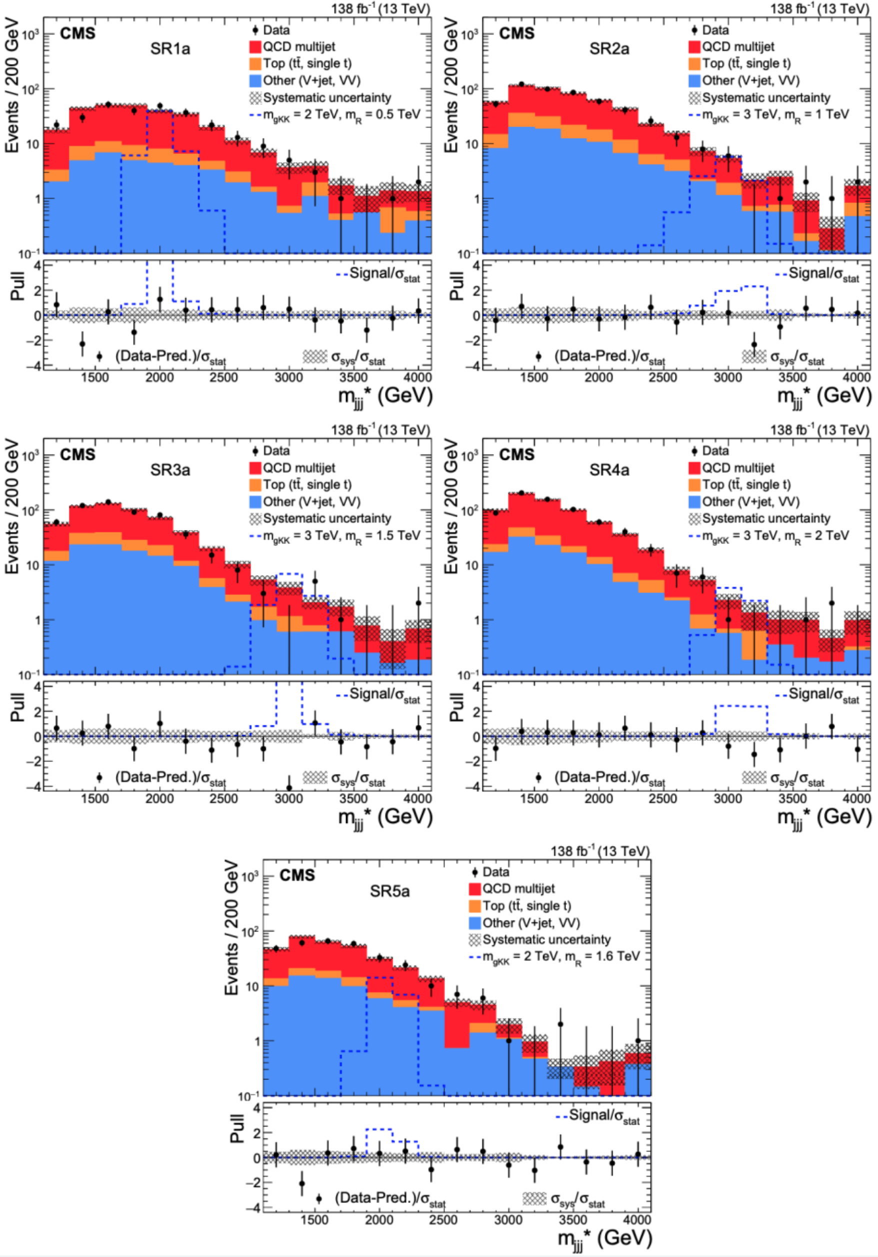


QCD prediction, Results, & Limits

- Define CRs in $m_{ja,jb}$ sideband with:
  $m_{85}>15$ GeV & $m_{90}<50$ GeV.
- Form 10 CRs: CR1—5a, CR1—5b.
- Predict QCD as: $\text{Pred}^{\text{QCD}}_{\text{SRxy}} \equiv [\text{Data}-\text{Rest}]_{\text{CRxy}} \dfrac{\text{QCD}_{\text{SRxy}}}{\text{QCD}_{\text{CRxy}}}$
- MC for the rest processes.

- Exclusion limits on σxB and on masses:





Welcome to CMS!

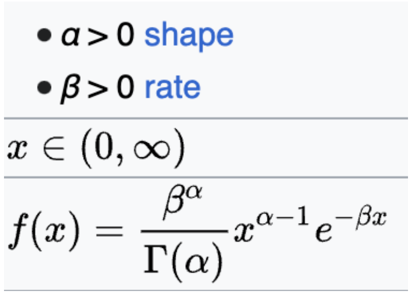
Get ready for long & exciting journey!

An optional hands-on exercise as homework (next)

Hands-on exercise: Signal discrimination optimization, significance evaluations

1. Generate distributions for B & S.
   - get simple python code from here
   - or you can get the coping from lxplus:
     "`cp /afs/cern.ch/work/a/agapitos/public/WWW/gKK/ForSYSUschool/1D-optimization.py .`"
   - L34 we define a BKG distribution (gamma function)
   - L40 we define a signal distribution (gaussian)
   - check the number of events and the parameters set
   - run with "`python 1D-optimization.py &`"

Gamma distribution



- $\alpha > 0$ shape
- $\beta > 0$ rate

$x \in (0, \infty)$

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

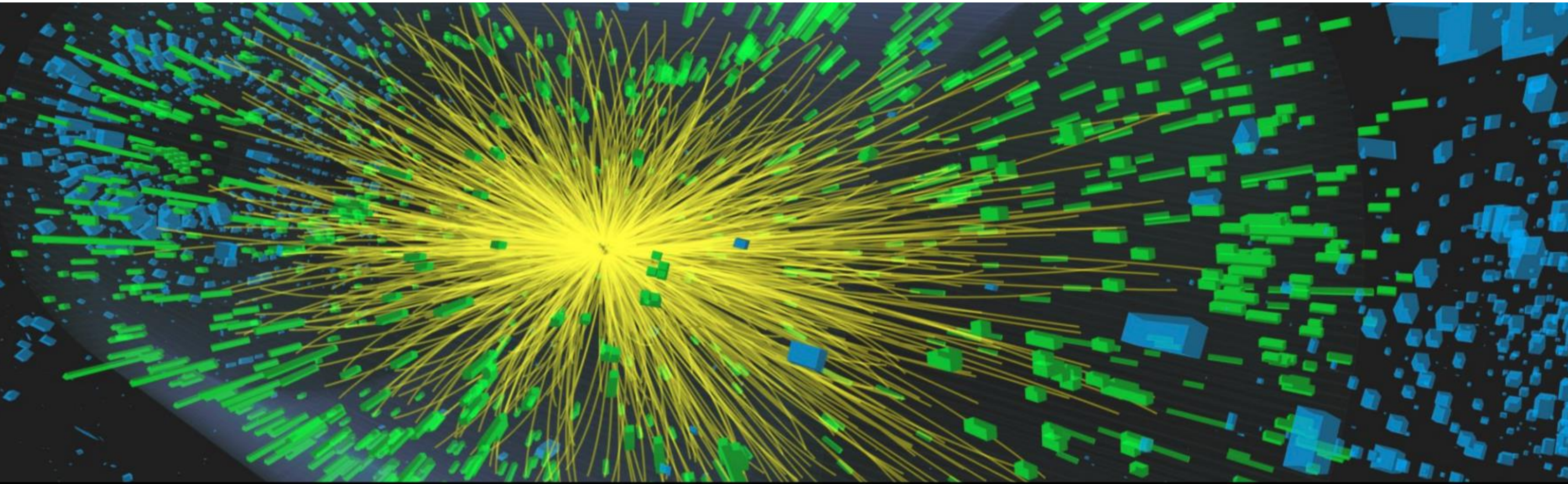
2. Find the optimal cut value of x variable, based on significance $S/\sqrt{B}$,
   - Need to loop over the different bin configurations from left and right
   - comment in the L69-99, and run again. Do you see the two for loops…?

3. Repeat with different signal:
   - comment out/in the L40/41 → use flat distribution for signal

4. Change normalization of S or B and redo optimization.
   - What do you observe/conclude…?

Code designed
by Leyun Gao
(PKU PhD student)

BACK-UP SLIDES

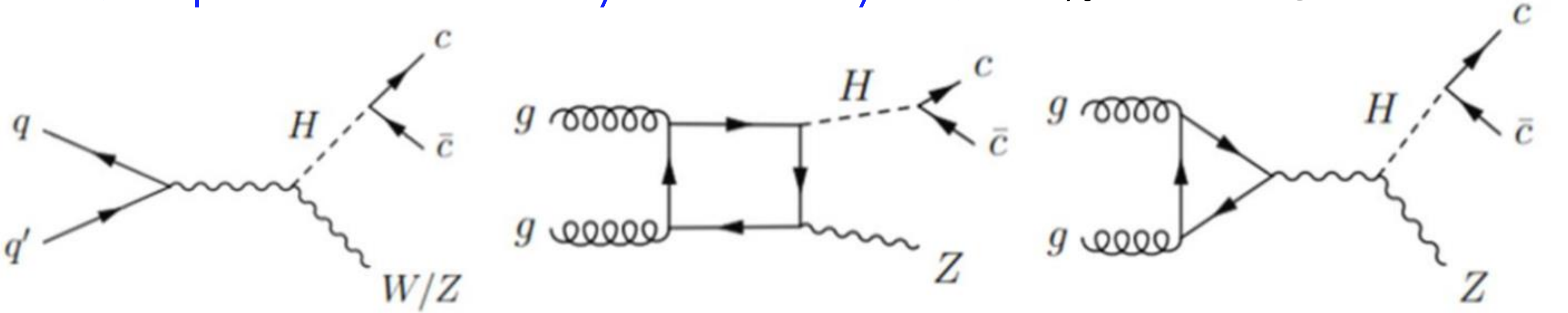


VH production with H→cc probe

We want to measure H→cc (x-sec, couplings etc.) Rare mode BR(cc)~3%

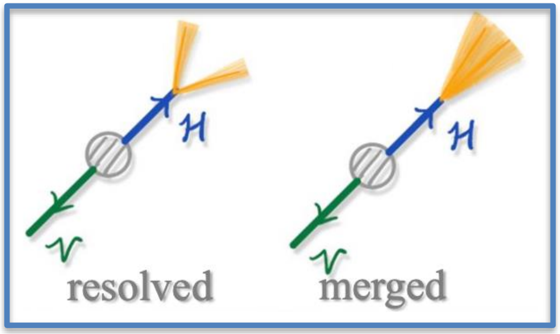
Which production mode would you use and why? →VH ~4% but low BKG!

PRL



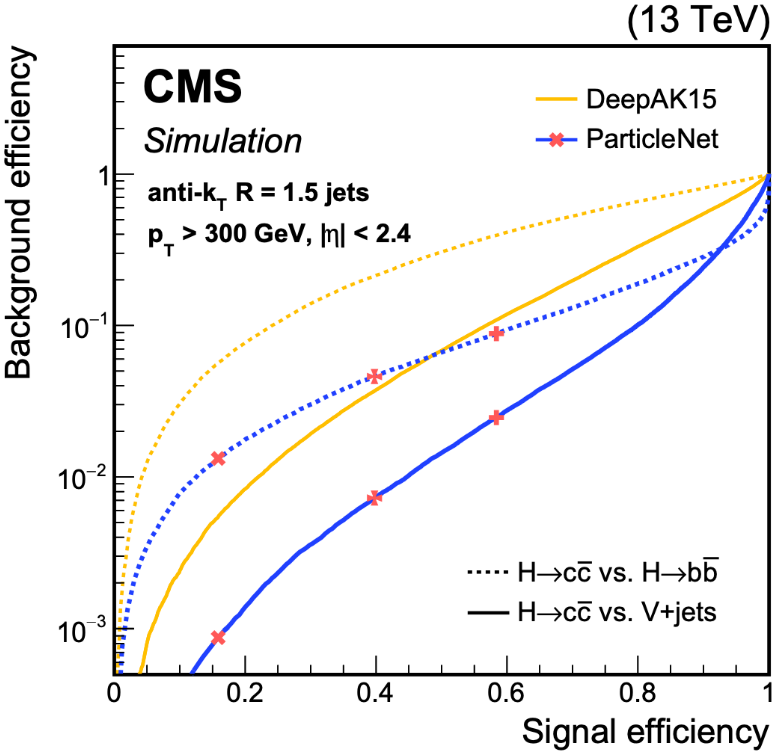
3 channels based on W/Z final states:

- Z → νν
- W → μν/eν
- Z → ee/μμ



Selection basics:

- 0,1,2l
- 2 AK4 / 1 AK15
- c / cc-tagger



VH production, H→cc, measurement

Resolved case:

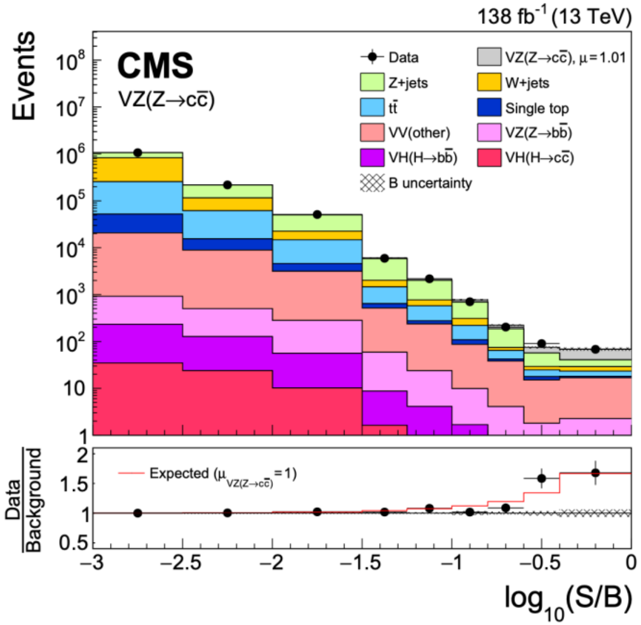
- Use BDTs to discriminate signal
- What is a BDT?
  - LM technique which combines input variables…
  - Advantageous in case where several variables have poor discrim. power.

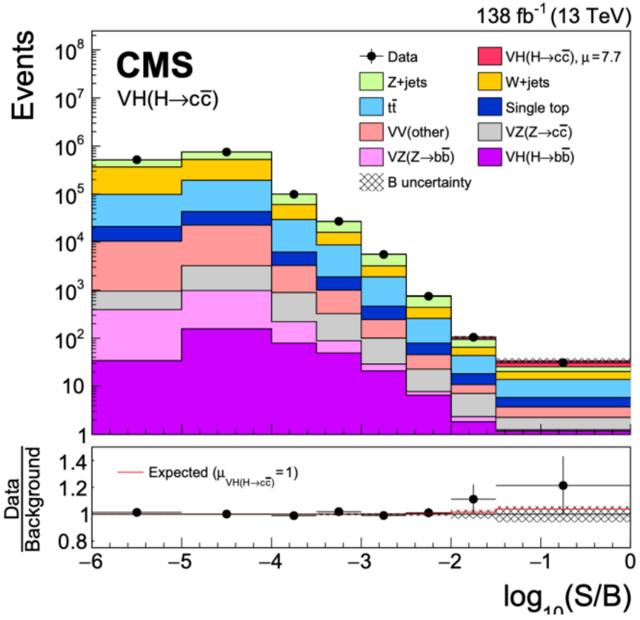
- Calibration with Z→cc
- First observation 5.7σ

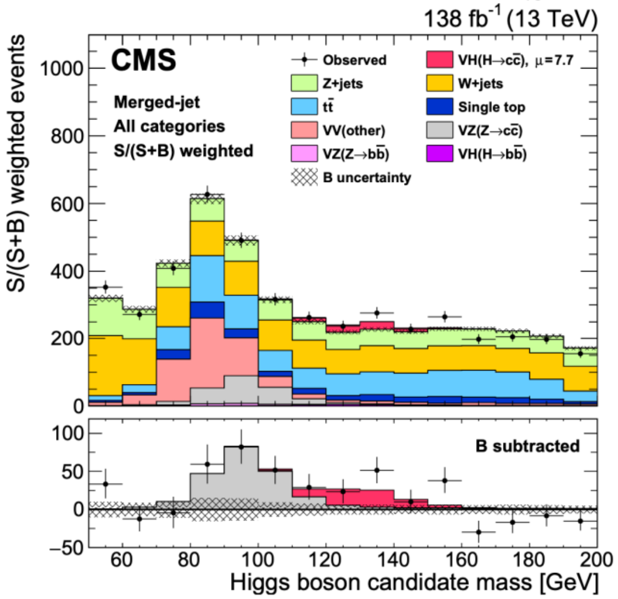
Merged case:

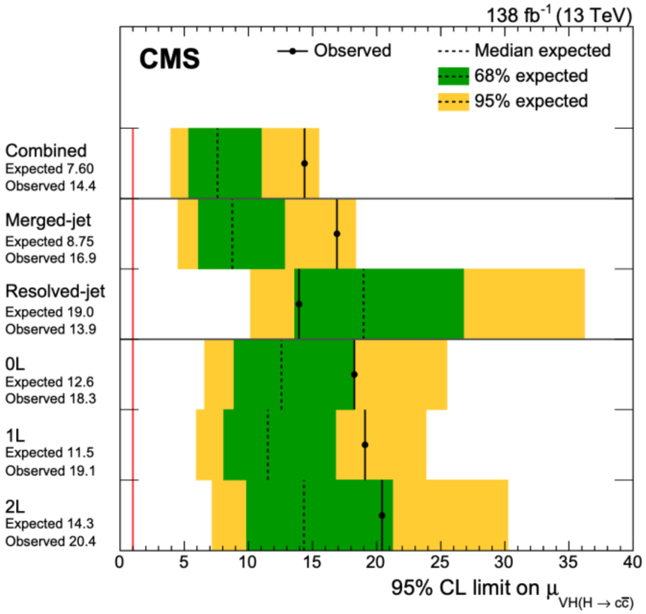
- Use SD mass of AK15 jet
- Fit Z→cc & H→cc

Result: upper limit on σB:
Observed x14 SM prediction
Expected x7 SM prediction



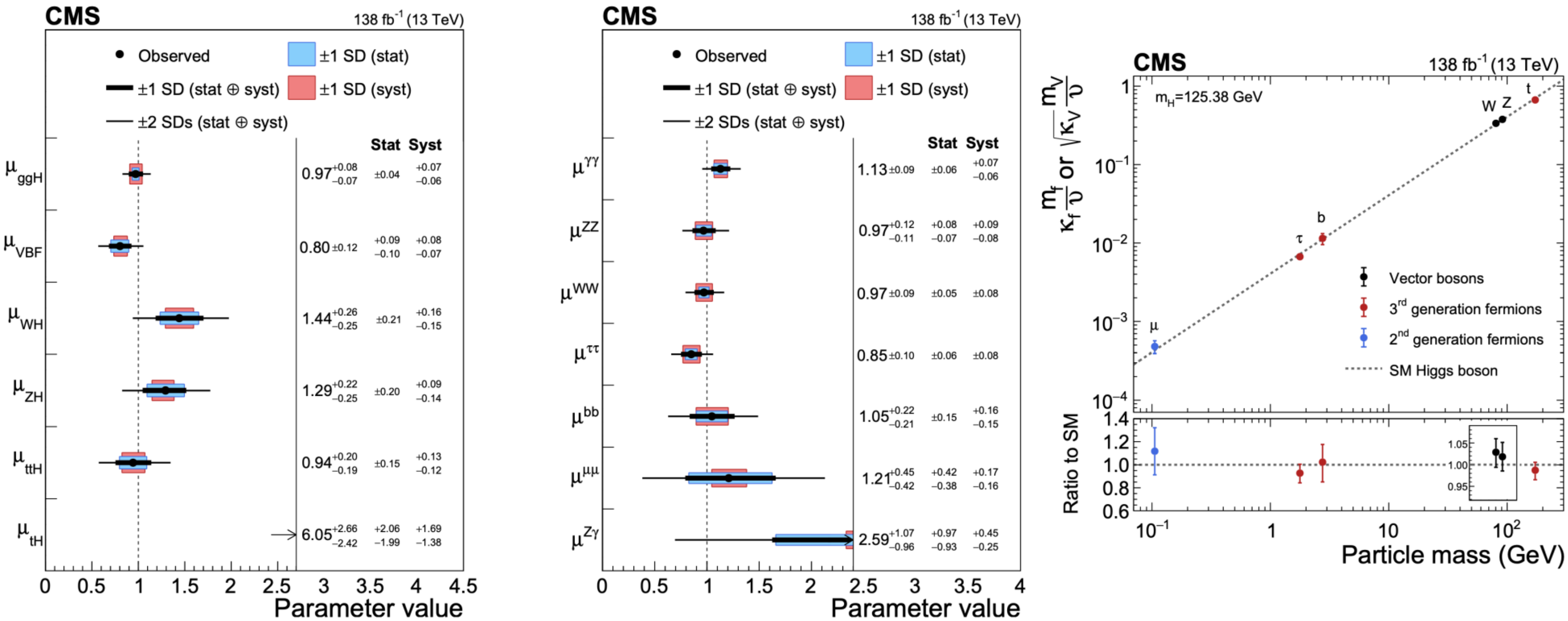






A portrait of Higgs 10y after its discovery

Nature

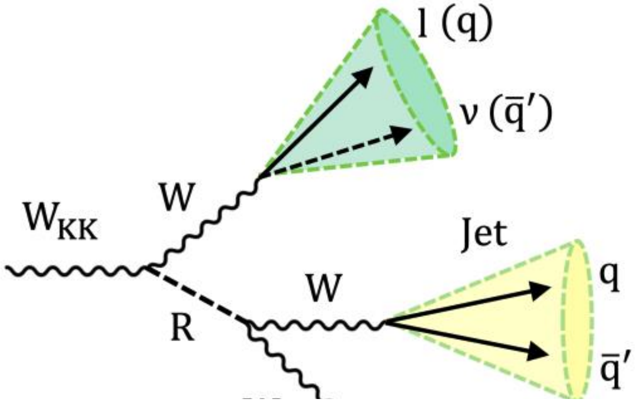


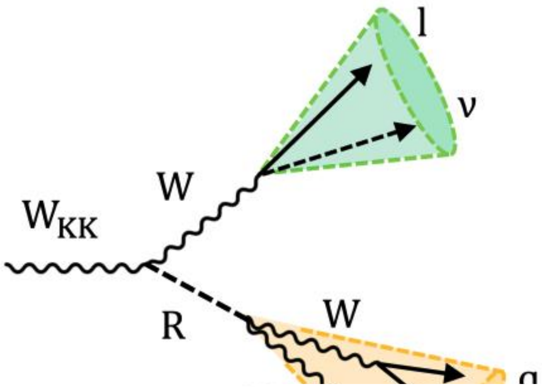
- We are in the era of precision measurements of H sector
- Production modes
- Decay modes
- Measured coupling modifiers to fermions & bosons, as functions of masses

Search for new resonances X→WR→WWW

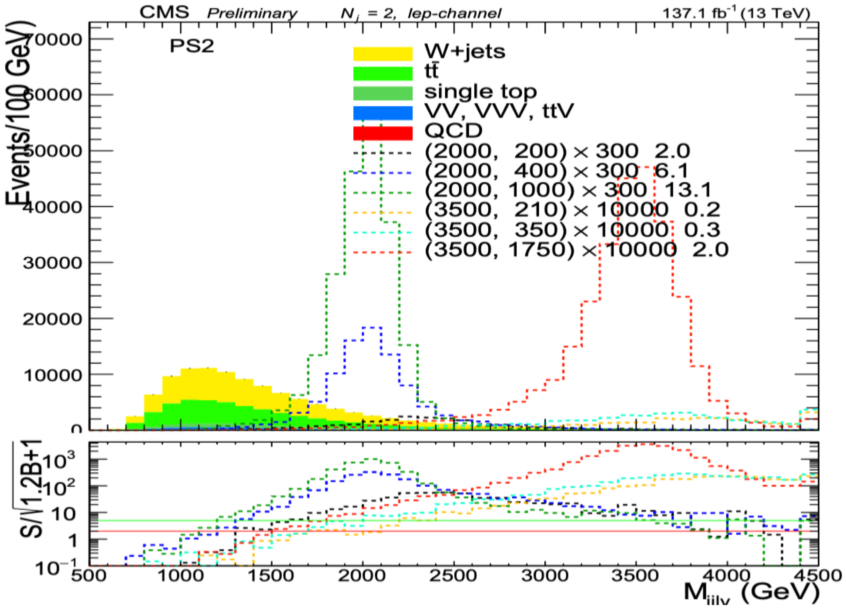
Step 1: Focus on the topology / final state, and based on this make a preselection:

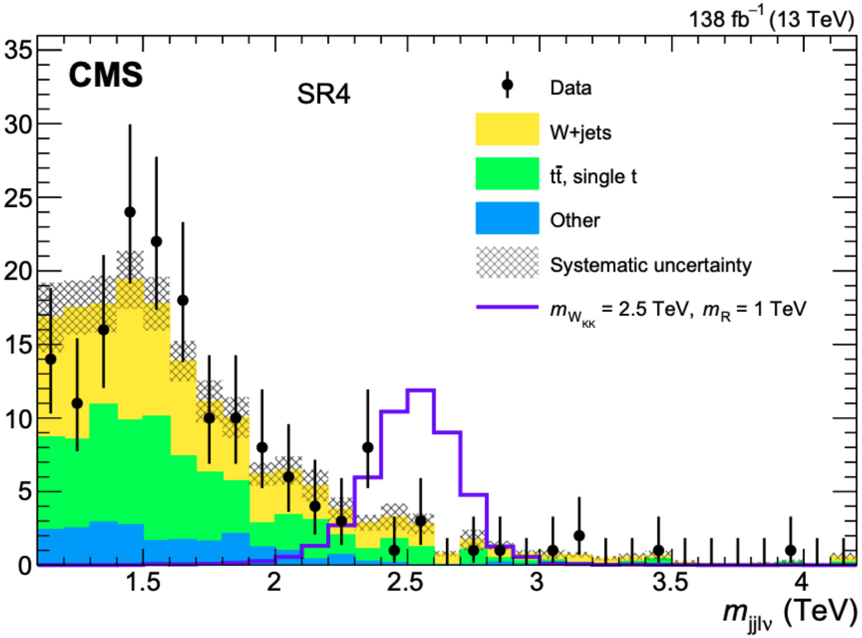
PRL





- Signal comes from the Extended Warped ED model
- 2 Branes, EW propagated into extended bulk
- 2 New particles involved Di-resonant signal

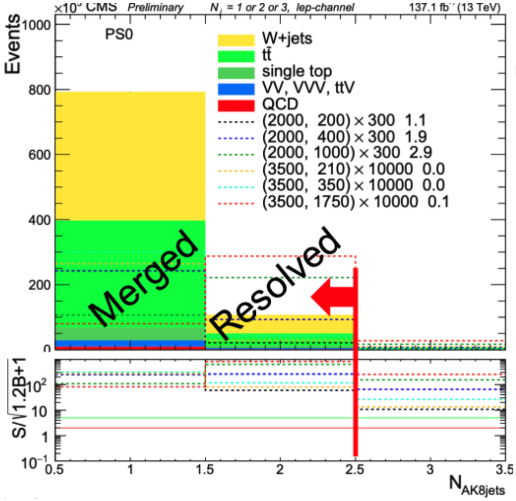




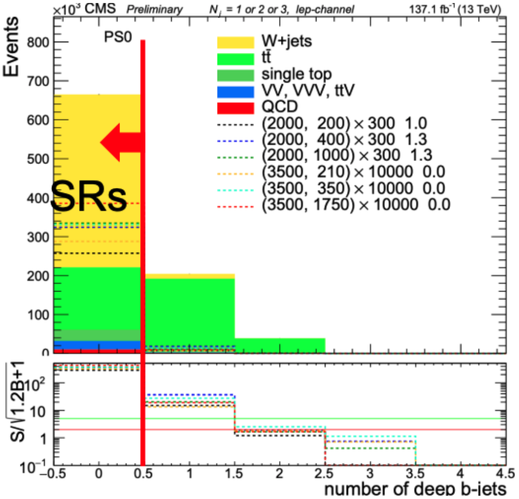
- Q: Which trigger would you use?

X→WR→WWW, Selection

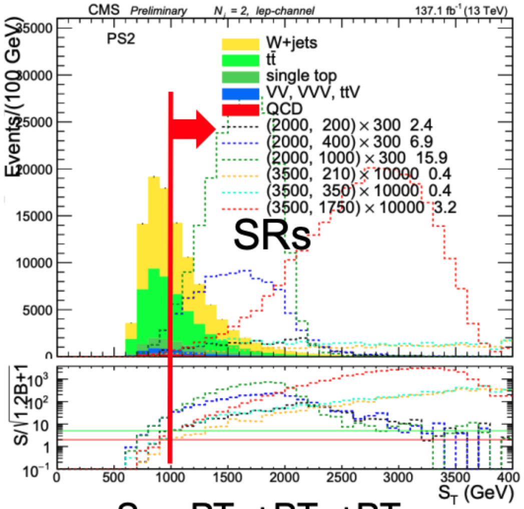
Step 2: we explore kinematics at preselection, i.e. we plot variables and check if further selection cuts can improve **sensitivity** ($S/\sqrt{B}$). Examples:



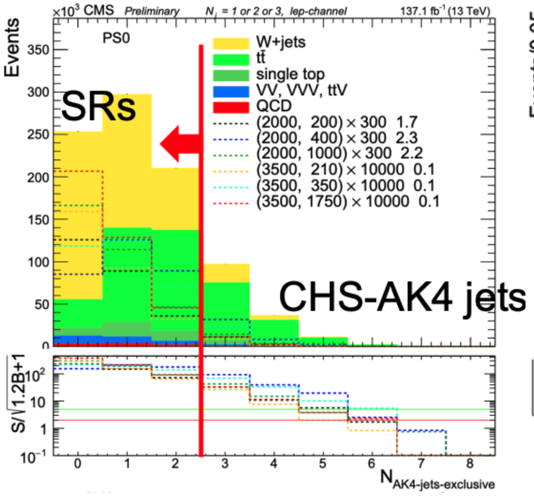
PUPPI-AK8 jets

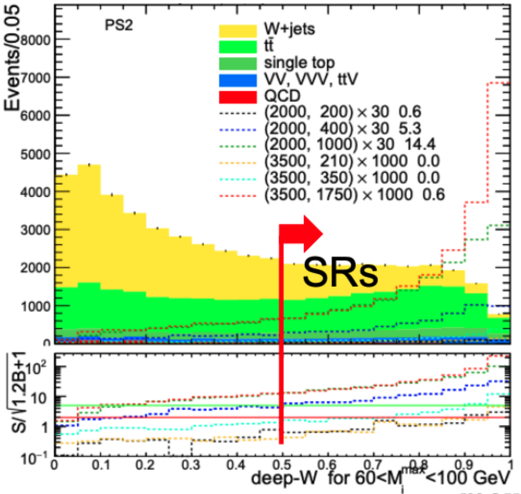


CHS-AK4 jets (deepCSV)

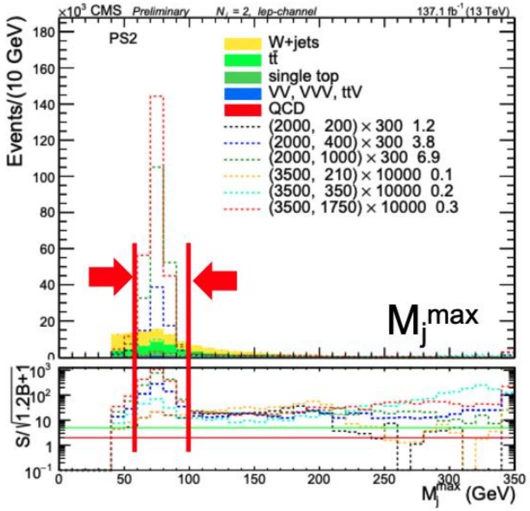


$S_T = PT_{j1}+PT_{j2}+PT_{W(l\nu)}$





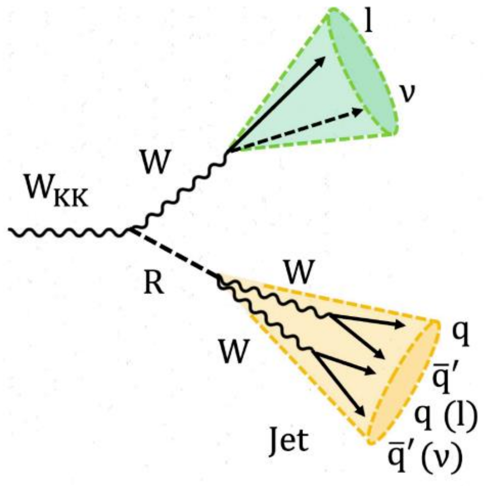
deep−W$^{\max}$

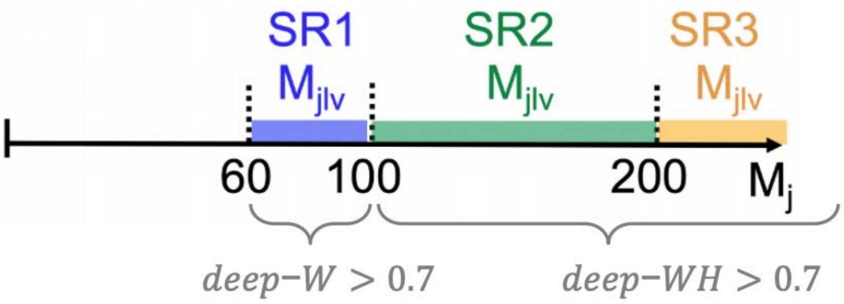


At the end, we set a list of cuts which define signal region (SR)

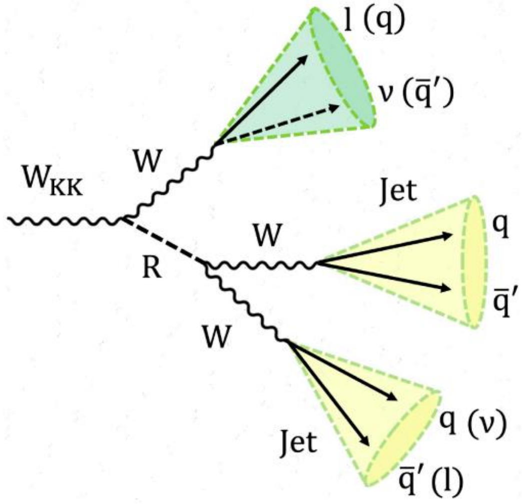
SRs & Binning

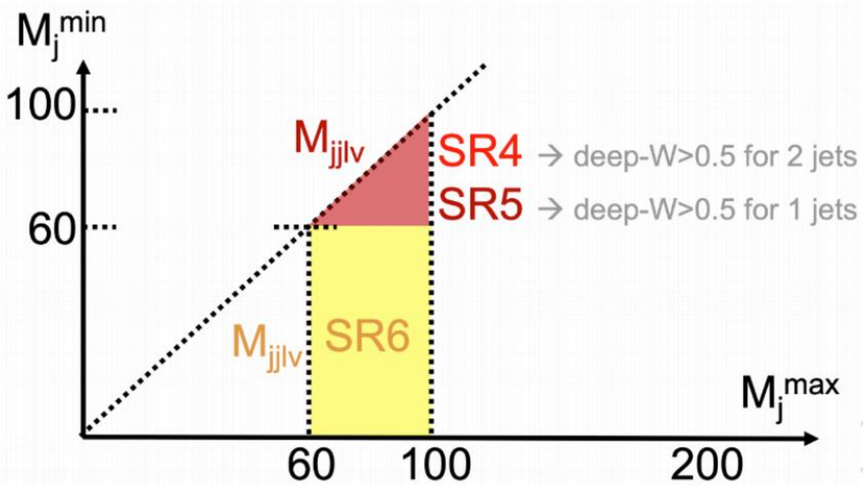
N_j=1





N_j=2





SRs & Binning

The analysis full selection is defined here as preselection + extra cuts
→ This defined the SR

The SR is further “binned” into 6 different categories SR1, SR2,..., SR6.

The m_{jlv} and m_{jjlv} are the variables where signal W_{KK} peaks these are the “observables” used for the fit. to extract potential signal from the data.

- N_j=1,2;
- N_b^M=0, $N_{j\text{-AK4-exc}} \leq 2$;
- $M_j^{\max}$: 60-..., (60-100) GeV for N_j=1(2);
- Taggers deep-W(WH)>0.7(0.7), for M_j:60-100(>100) GeV; while loosen to deep-W>0.5 in N_j=2 with $M_j^{\min}$ >60 GeV case;
- S_T >1 TeV;
- $M_{j\ell\nu}(M_{jj\ell\nu}) > 1.1$ TeV for N_j=1(2);

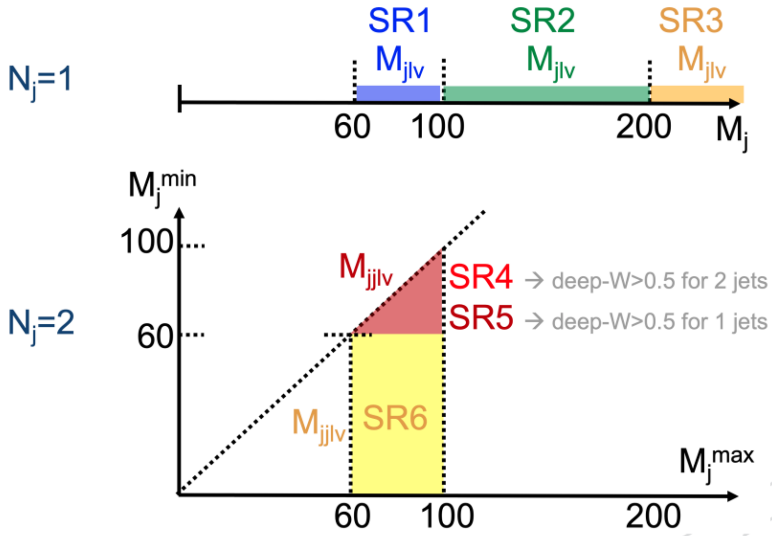


Figure 22: Signal region categorized into six orthogonal regions (SRs): SR1-6

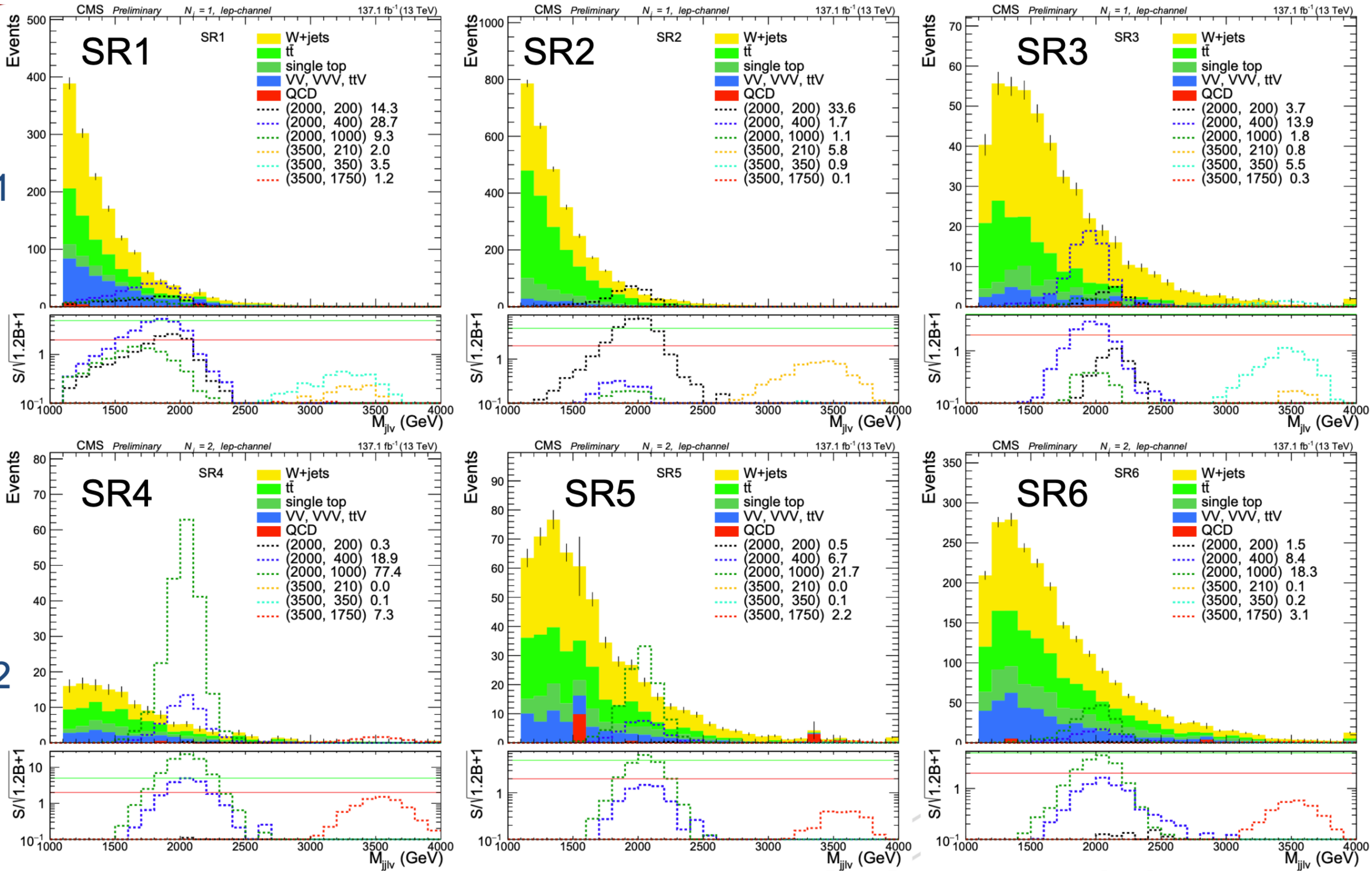
Table 9: Differences in kinematic cuts between the six SRs.

Region	$M_j^{\max}$[GeV]	taggers	$M_j^{\min}$[GeV]	tagger	N_j^{AK8}	N_j^{AK4}	N_b^M
SR1	60-100	deep-W >0.7	-	-	1	≤2	0
SR2	100-200	deep-WH>0.7	-	-	1	≤2	0
SR3	≥200	deep-WH>0.7	-	-	1	≤2	0
SR4	60-100	deep-W >0.5	60-100	deep-W>0.5	2	≤2	0
SR5	60-100	deep-W > (<)0.5	60-100	deep-W(<) >0.5	2	≤2	0
SR6	60-100	deep-W >0.7	0-60	-	2	≤2	0

SRs & Binning

N_j=1

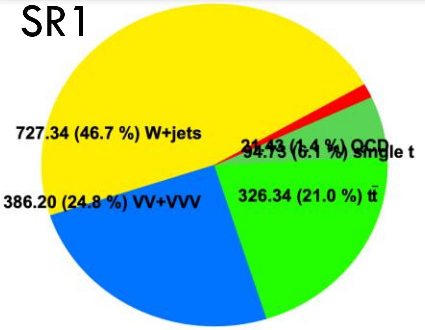
N_j=2



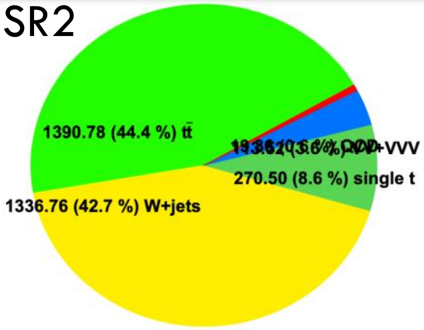
BKG composition in the 6 SRs

BKG composition in these 6 SRs →

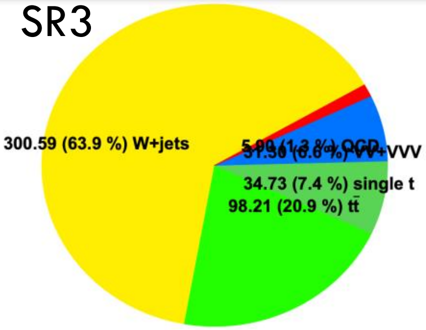
SR1



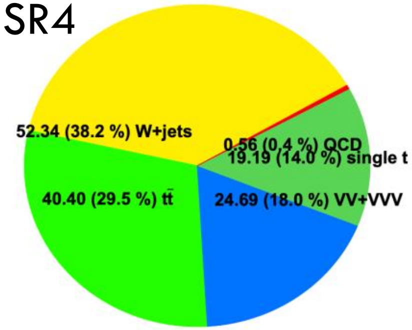
SR2



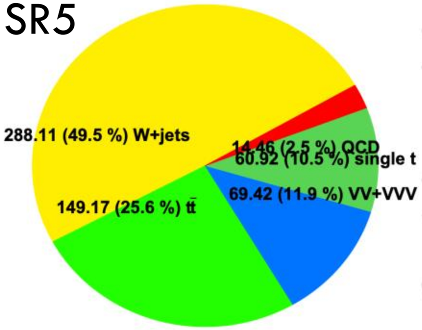
SR3



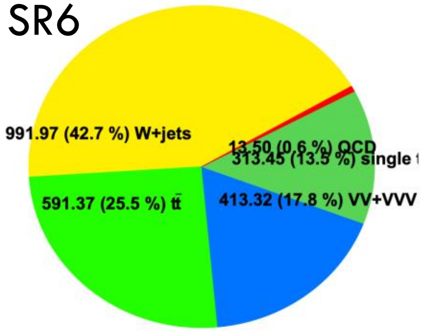
SR4



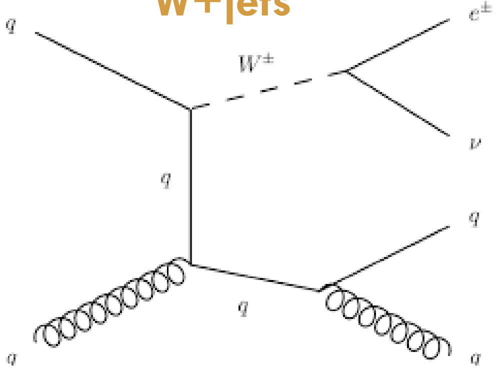
SR5



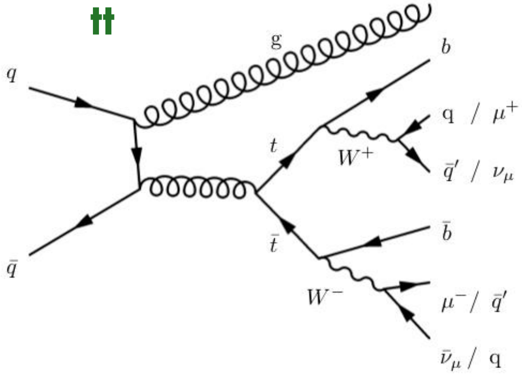
SR6



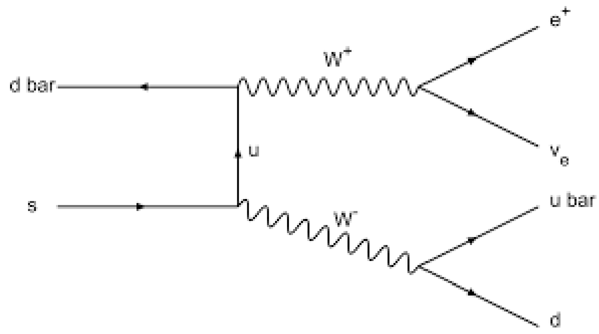
W+jets



tt



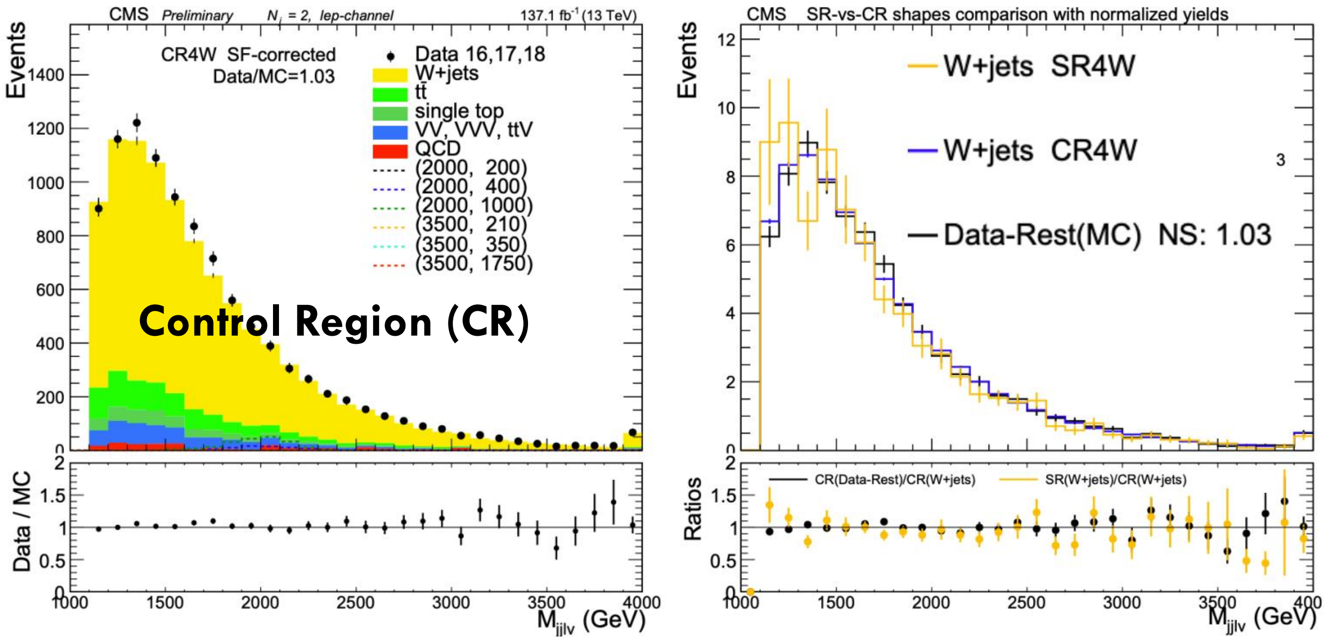
WW



W+jets BKG Data-Driven prediction with CR

- MC very useful but it has limitations; not to be fully trusted.
  → Need to either validate that performs well in data or
  → to be replaced… by a “**Data-Driven prediction**” using real data.
  For this purpose…:

- We define **Control Region (CR),** inverting a cut condition (tagger for this example).
- The resulting CR is pure in W+jets, signal free, has large stat. and
  → the W+jets has similar kinematics in CR and SR.
- Thus we use the data as [Data-rest] in CR to derive W+jets spectrum:



Control Region (CR)

We do the assumption:

$$\frac{\mathrm{Pred}^{\mathrm{SR}}}{\mathrm{MC}^{\mathrm{SR}}} = \frac{[\mathrm{Data} - \mathrm{rest}]^{\mathrm{CR}}}{\mathrm{MC}^{\mathrm{CR}}}$$

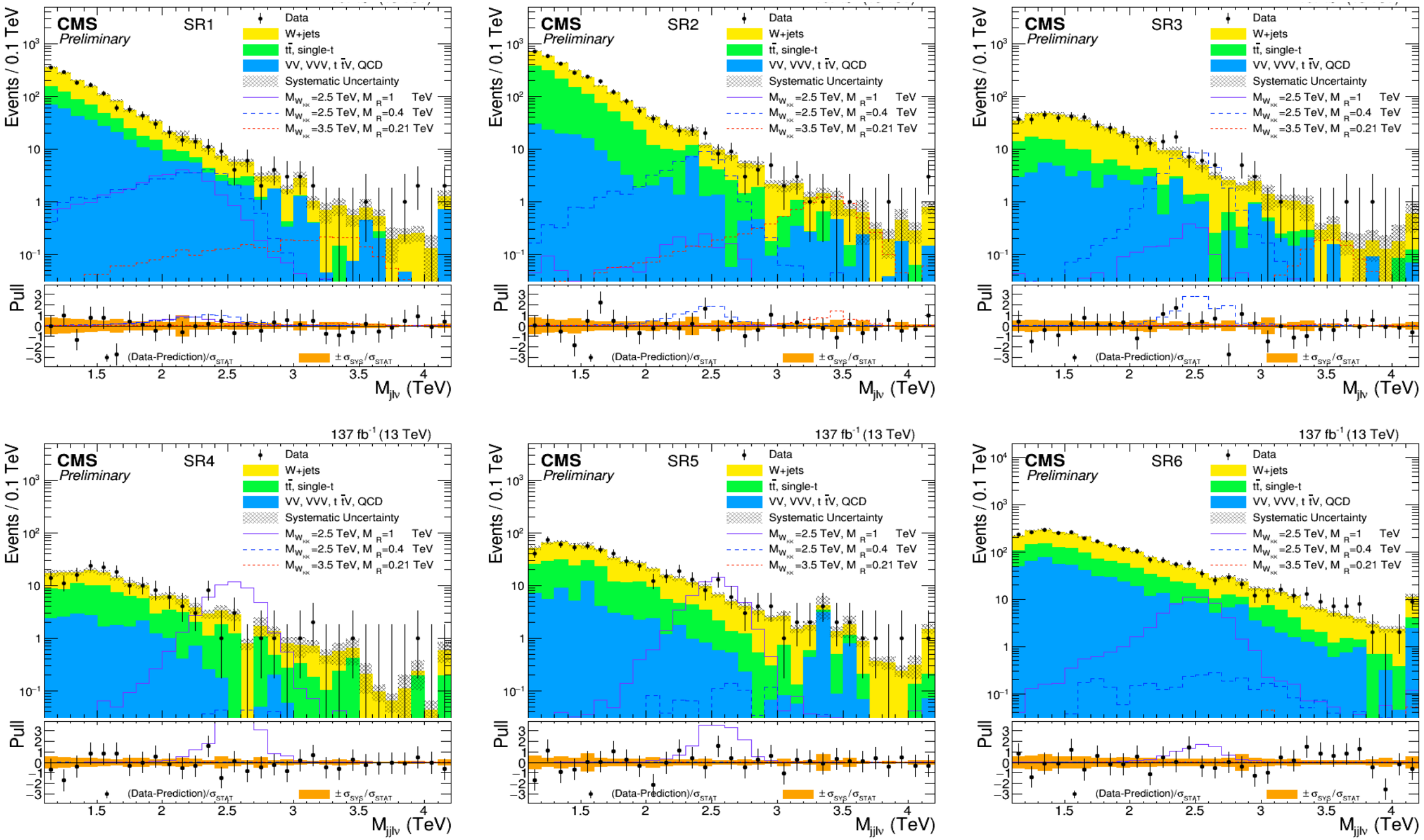
I.e. we trust MC simulates well the SR/CR ratio of the spectra.

We solve for “Pred”.
The resulting spectrum replaces MC as prediction.
(Find more here).

Final “post-fit” prediction

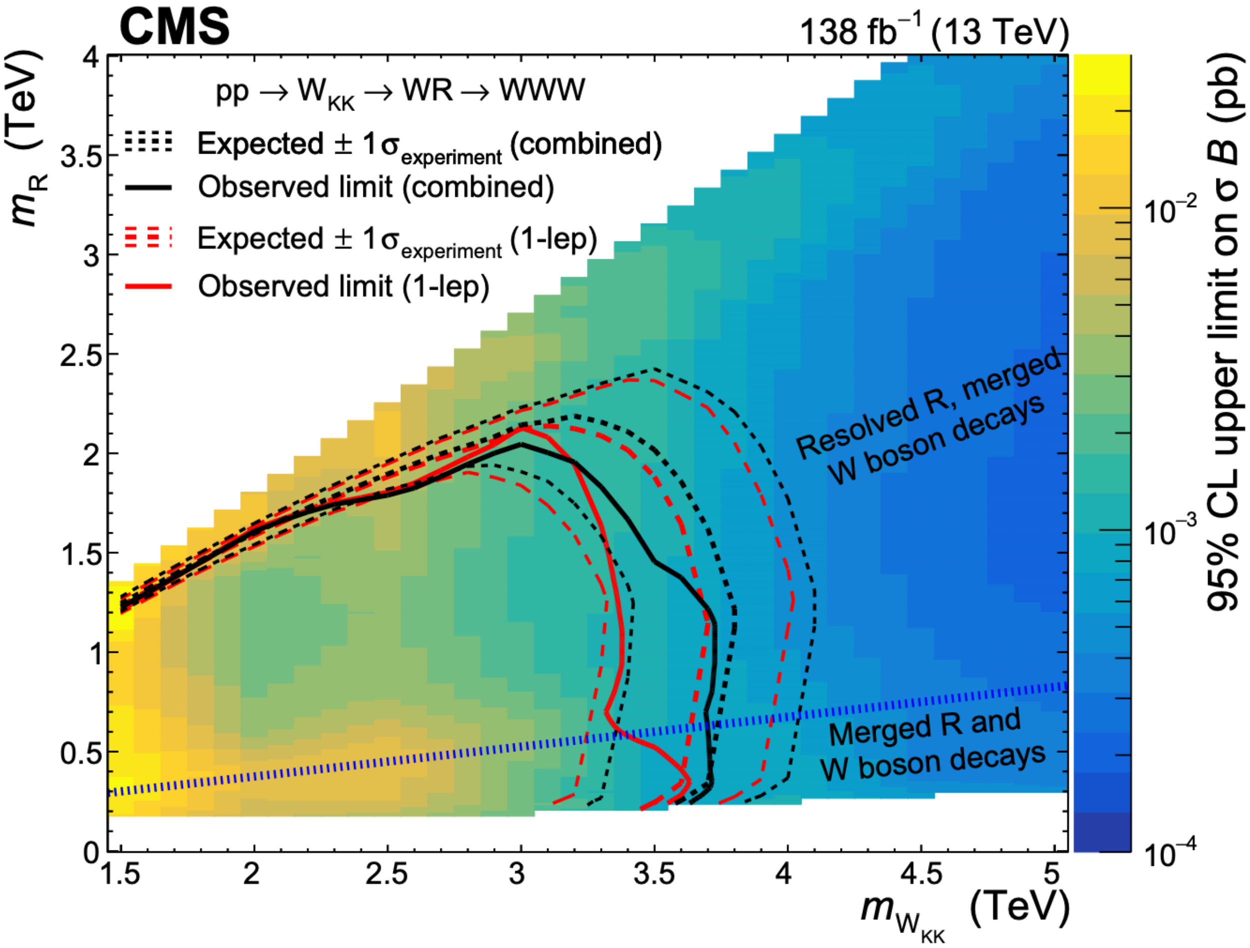
- We fit all 6 SRs simultaneously and get the “post-fit” prediction of the spectra.
- Systematics uncertainties are considered with nuisance parameters in this fit:

(these are for BKG prediction, rates of BKGs, PDFs, JES/R, tagger efficiencies etc.)



Limits on 2D masses plane

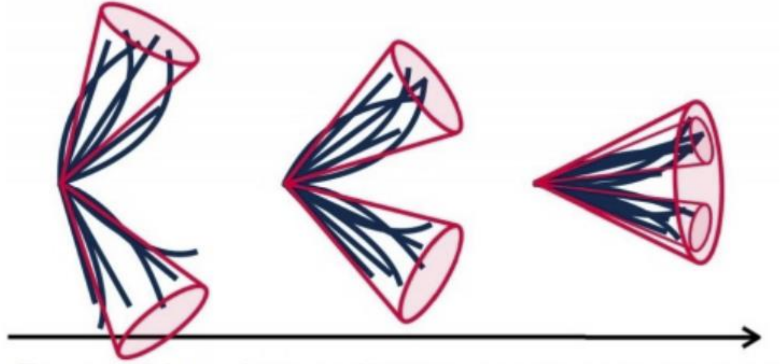
- As the data are in agreement with SM prediction - no evidence of signal / new physics.
- Thus we set limits to the model parameters / masses.



- In this example observed limit is weaker than the expected (why?) this is indicative of an excess of events (present at SR6)

Jets substructure & jet “tagging”

Boosted objects → small angular separation of the products → merged jets with substructure:



Boosted jets: Increasing transverse momentum

W/Z → qq
H → bb or H → qqqq, or H → qqlv

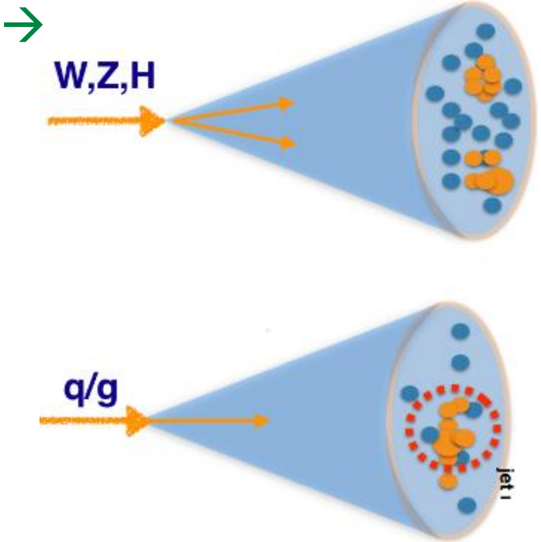
→ anti-kt clustering
→ Large-R jets: $\Delta R = \sqrt{(\Delta\varphi^2 + \Delta\eta^2)} \approx 2m/p_T$
→ ”Groomed” Soft-Drop Masses: $M_J \sim M_V \pm 0.2\ M_V$

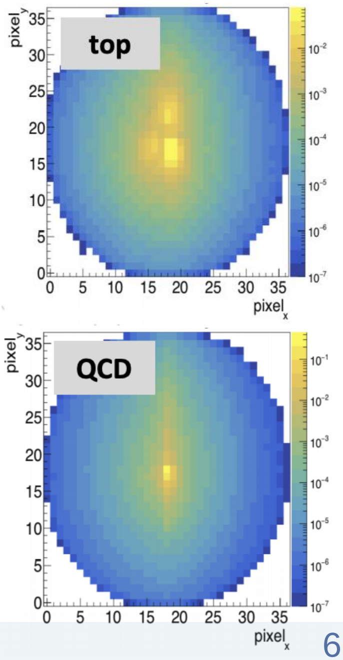
$$\tau_N = \frac{1}{d_0}\sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$$

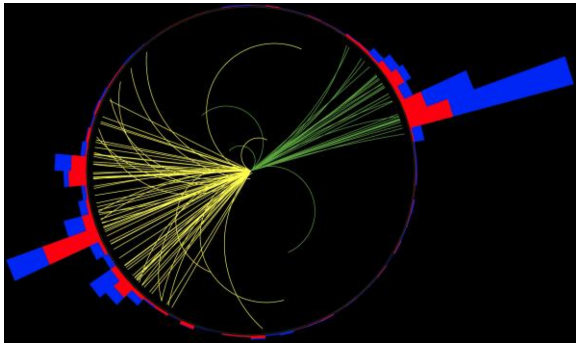
Taggers based on (2-prong) substructure

- τ_N = N-subjettiness → ratios: $\tau_2/\tau_1 = \tau_{21}$ →
- Decorrelated taggers τ_{21}^{DDT}

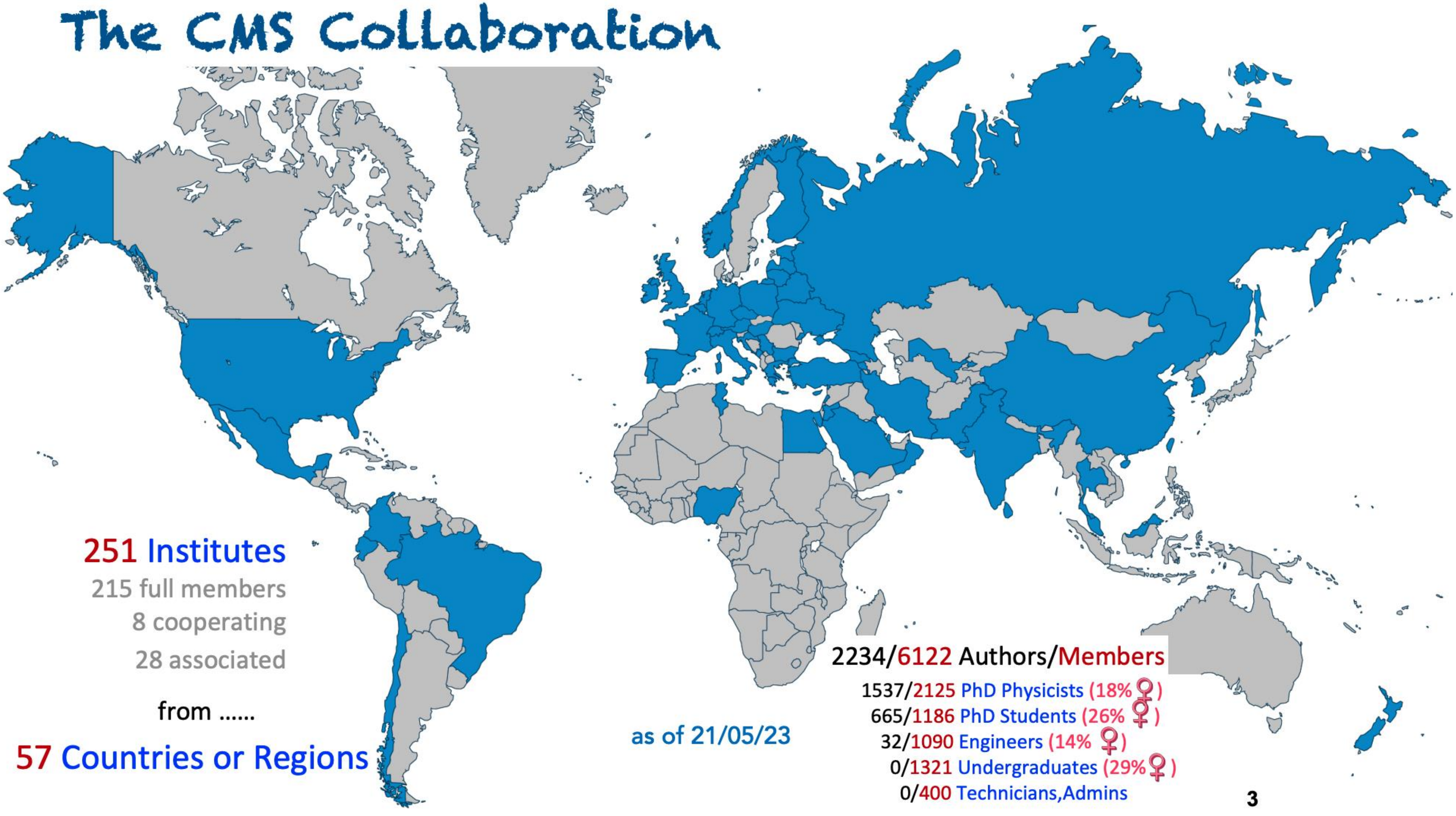
Deep-NN taggers & Image taggers (soon)







The CMS collaboration



The CMS Collaboration

251 Institutes

215 full members

8 cooperating

28 associated

from ……

57 Countries or Regions

as of 21/05/23

2234/6122 Authors/Members

1537/2125 PhD Physicists (18% ♀)

665/1186 PhD Students (26% ♀)

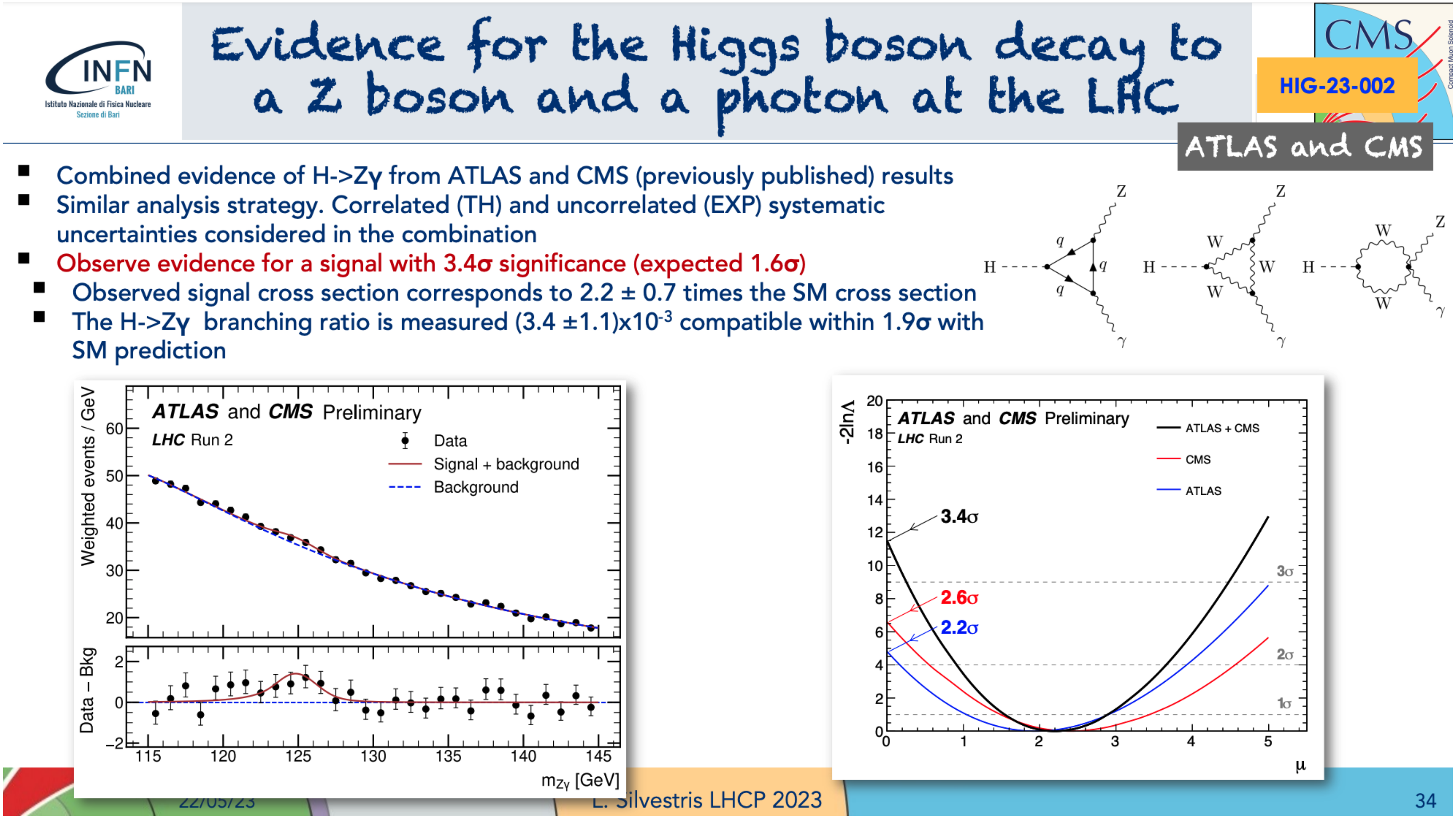
32/1090 Engineers (14% ♀)

0/1321 Undergraduates (29%♀)

0/400 Technicians,Admins

3

The CMS collaboration



Evidence for the Higgs boson decay to a Z boson and a photon at the LHC

HIG-23-002

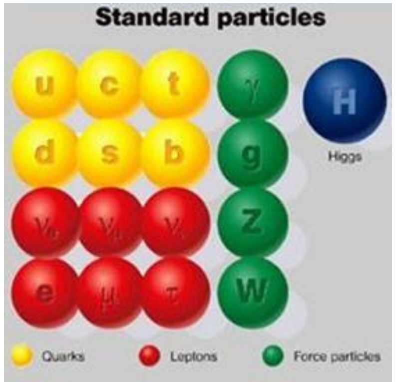
ATLAS and CMS

- Combined evidence of H->Zγ from ATLAS and CMS (previously published) results
- Similar analysis strategy. Correlated (TH) and uncorrelated (EXP) systematic uncertainties considered in the combination
- Observe evidence for a signal with 3.4σ significance (expected 1.6σ)
  - Observed signal cross section corresponds to 2.2 ± 0.7 times the SM cross section
  - The H->Zγ branching ratio is measured (3.4 ±1.1)x10^{-3} compatible within 1.9σ with SM prediction

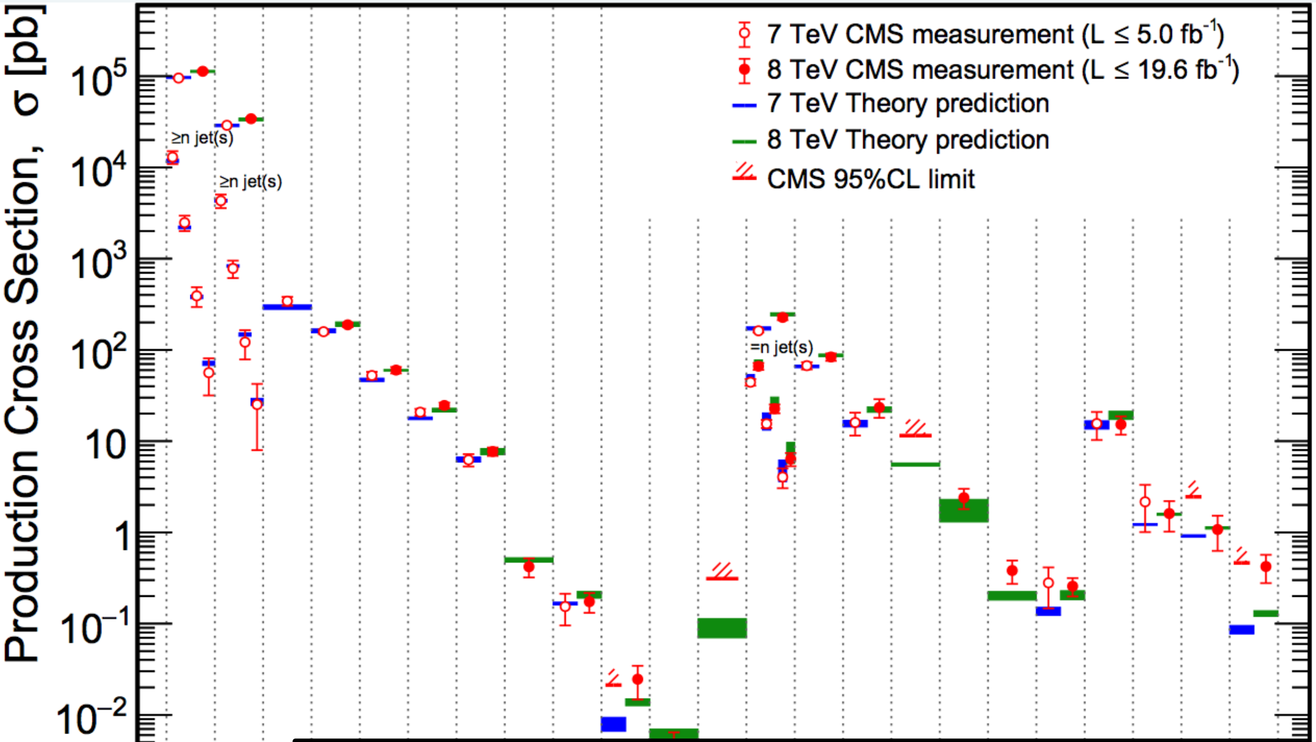
22/05/23 L. Silvestris LHCP 2023 34

Standard Model: Predictions & Measurements

◆ **SM:**

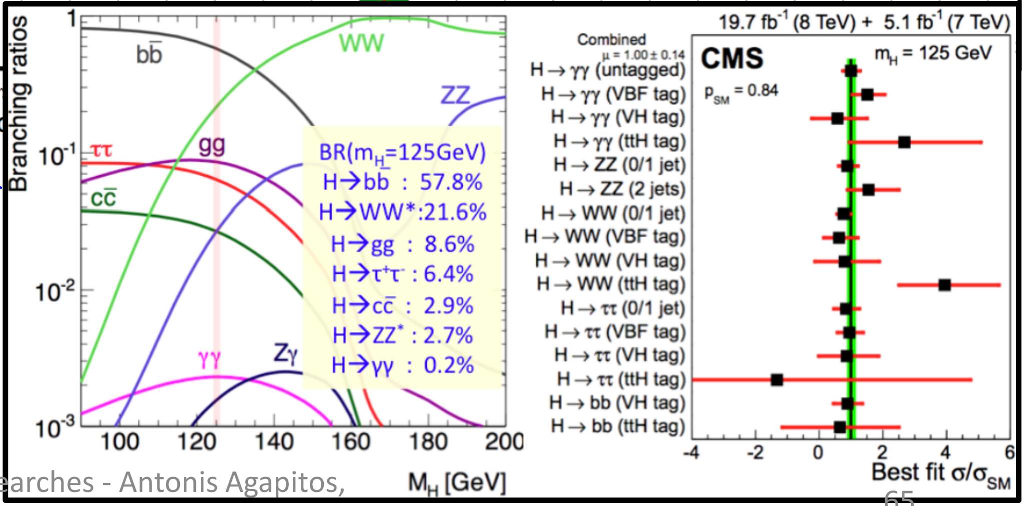


- **Very successful gauge theory.**
- **Precisely predicts processes: σ~[10^{11} - 10^{-3}] pb**



◆ **Along with successful prediction of Higgs BR → measurements…**

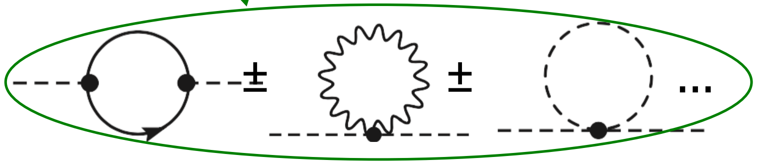
◆ **No evidence for any deviation(s) from the SM.**



SM: Shortcomings in QFT frame

◆ **m_{Higgs}: renormalizable →** $m_{Higgs}^2 = m_{bare}^2 - [\pm\lambda\Lambda_{cut}^2 \pm \ldots] = 125^2\ \mathrm{GeV}^2$

◆ **New Physics at: Λ_{cut}**~10^2-10^{18}GeV ←→ **"Hierarchy/Naturalness" Problem**



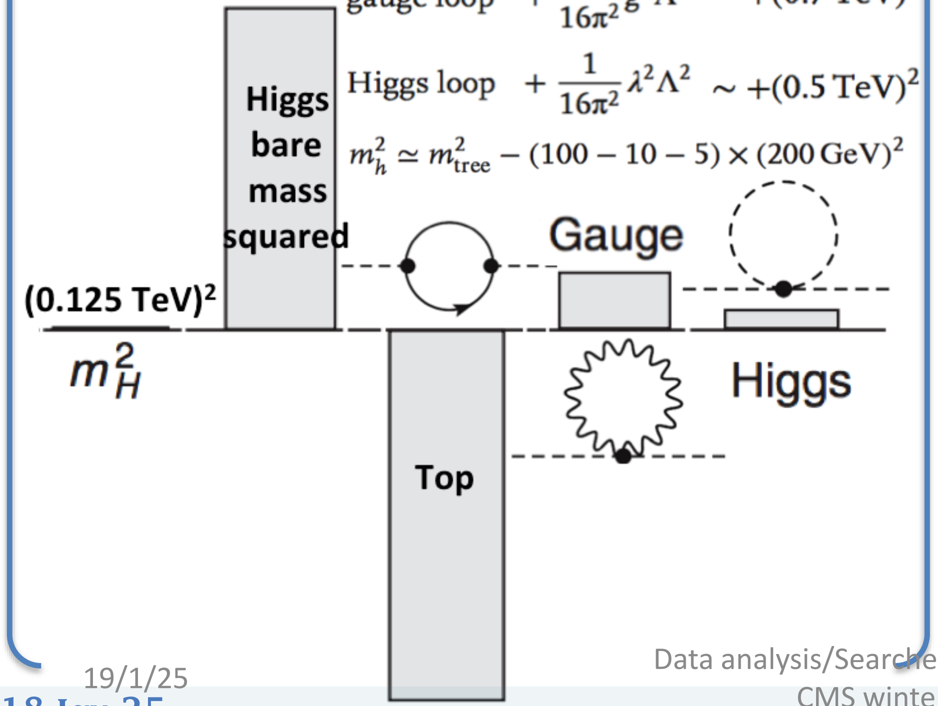
Λ_{cut}~10^{18} GeV (Planck scale):

125^2 = **12345678901234567890123456789012345** – **[12345678901234567890123456788996720]**

Λ_{cut}~10^4 GeV:

top loop	$-\frac{3}{8\pi^2}y_t^2\Lambda^2$	$\sim -(2\,\mathrm{TeV})^2$
gauge loop	$+\frac{1}{16\pi^2}g^2\Lambda^2$	$\sim +(0.7\,\mathrm{TeV})^2$
Higgs loop	$+\frac{1}{16\pi^2}\lambda^2\Lambda^2$	$\sim +(0.5\,\mathrm{TeV})^2$

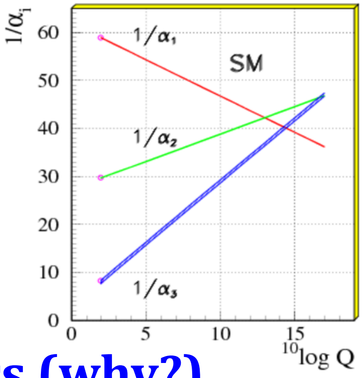
$$m_h^2 \simeq m_{\rm tree}^2 - (100 - 10 - 5) \times (200\,\mathrm{GeV})^2$$



◆ **Problem if : Λ_{cut}** >~TeV
→There is a "missing" mechanism for m_H convergence.

◆ **Maybe new bosons (?) at scales: Λ_{cut}**~10^2–10^{18} GeV.

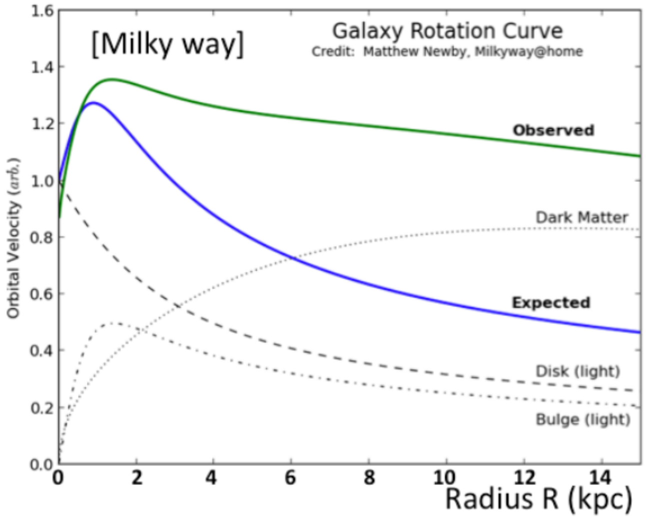
◆ **Open Questions:**

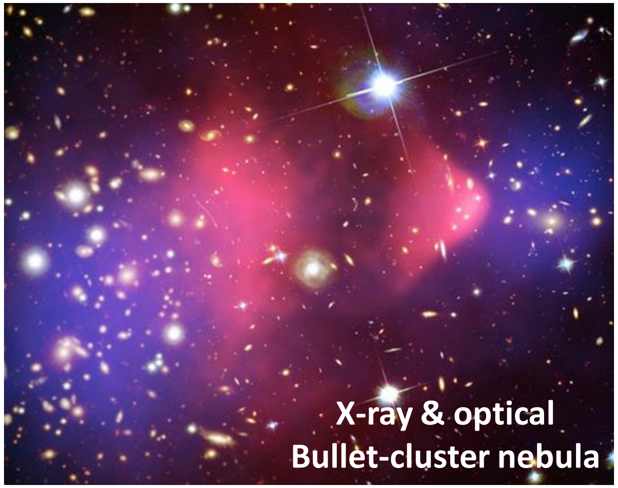


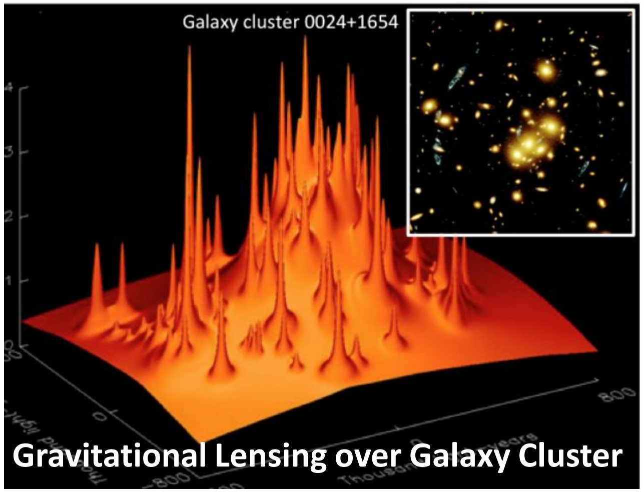
- **Unification "GUT"?**
- **Gravity QFT ?**
- **25 Free Parameters (why?)**
- **3 fermions families (why?)**

SM: Shortcomings in Cosmology

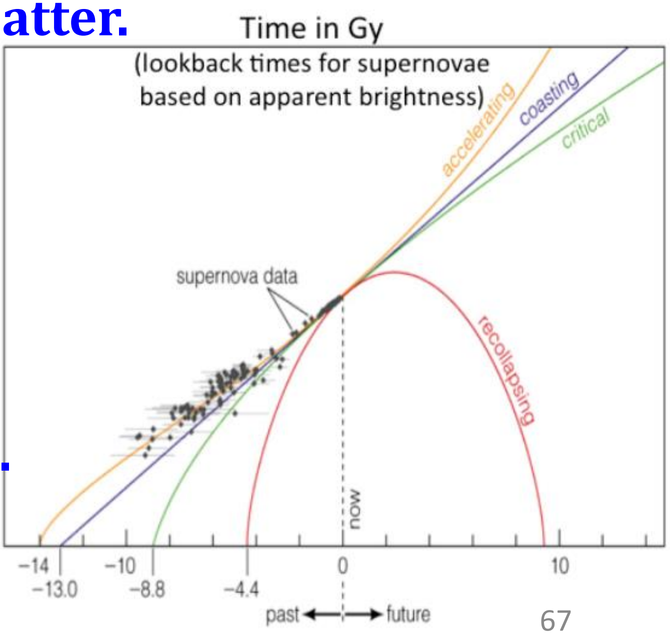
◆ **Dark Matter:**

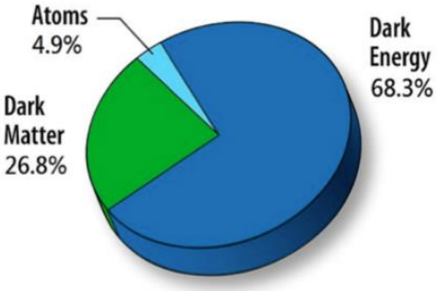






- **~5 times more "Dark" Matter than Visible Matter.**
- **Particle-nature matter → which particle?**

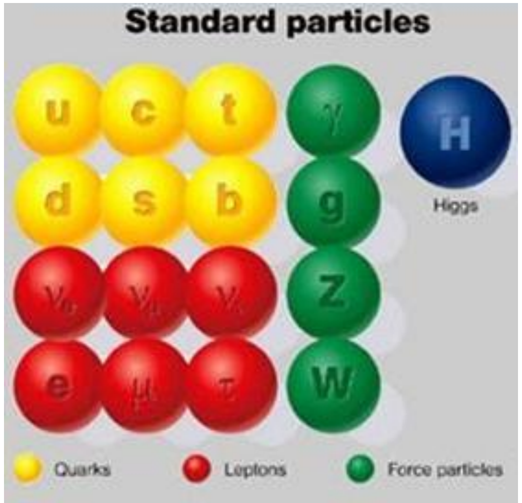




◆ **Dark Energy:**
accelerated expansion
(why? who?)

◆ **SM: cannot explain large scale observations.**
SM → «effective» rather than fundamental.

SUperSYmmetry: SUSY



New Principle of Nature

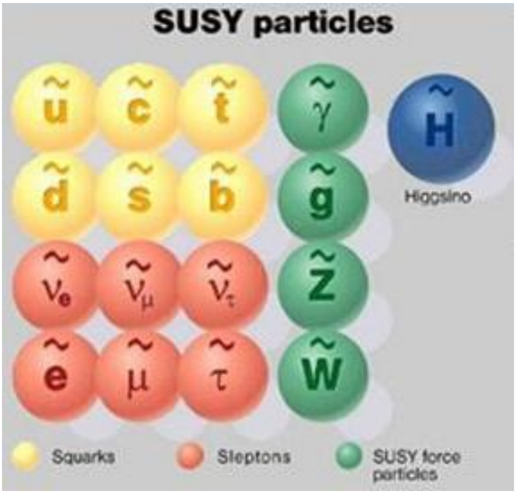
Δspin = ½

Spin Based Symmetry

Fermion ←→ Boson

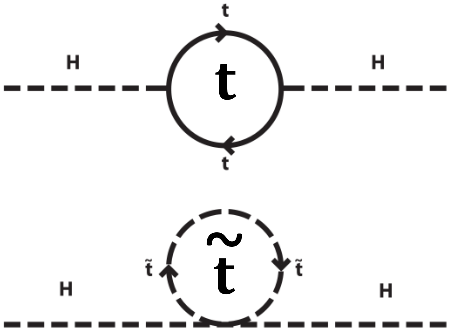
$M_{SM\text{-}particle}$ < or << $M_{SUSY\text{-}Sparticle}$

SUSY: a translation is “Superspace”.



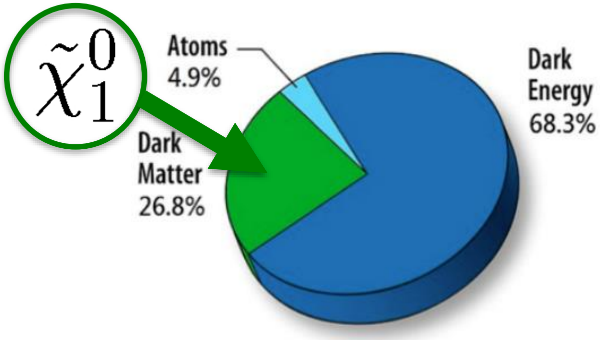
SUSY proposes solutions to SM problems:

◆ Hierarchy Problem



Sparticle loops cancel out corrections (if Λ_{cut}~< 1 TeV)

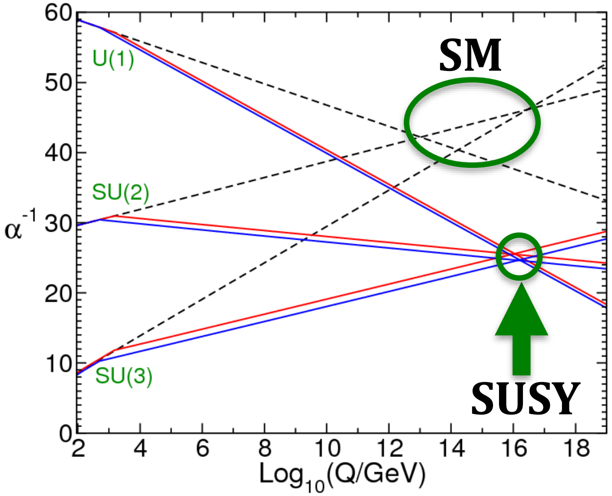
◆ Dark Matter Problem



LSP = Dark Matter (if “R-parity” conserved)

$P_R=(-1)^{2s+3(B+L)}$

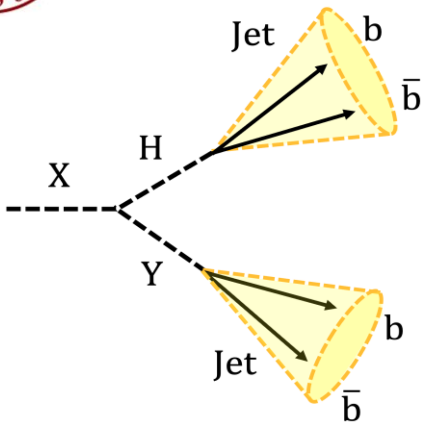
◆ GUT

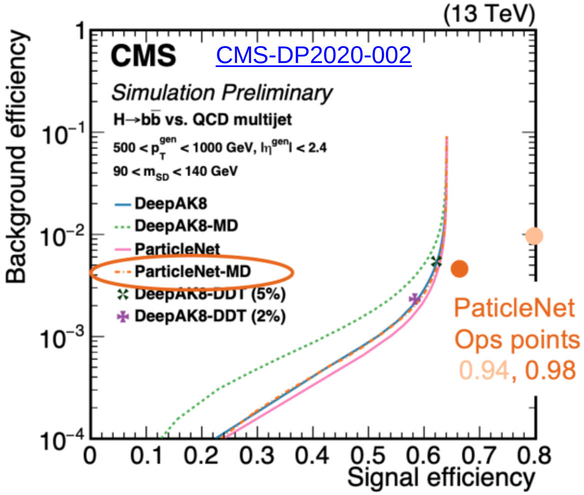


a^{-i} converges @10^{16} GeV

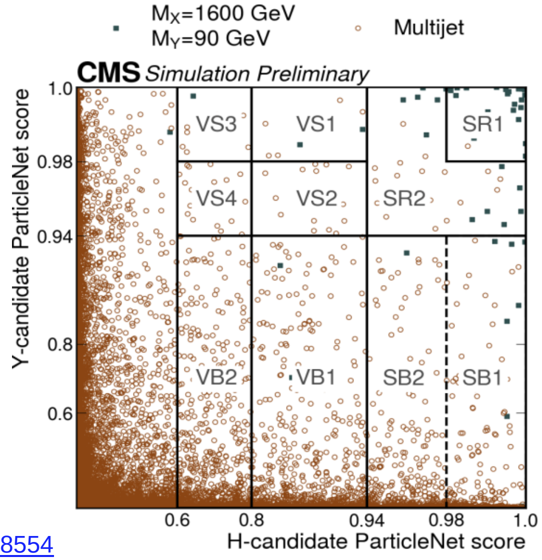
X → YH → 4b boosted

B2G-21-003

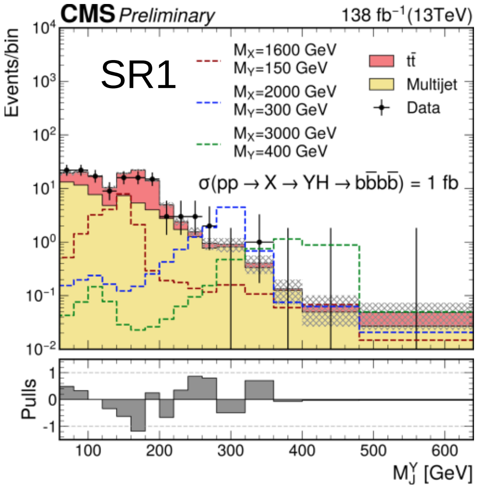


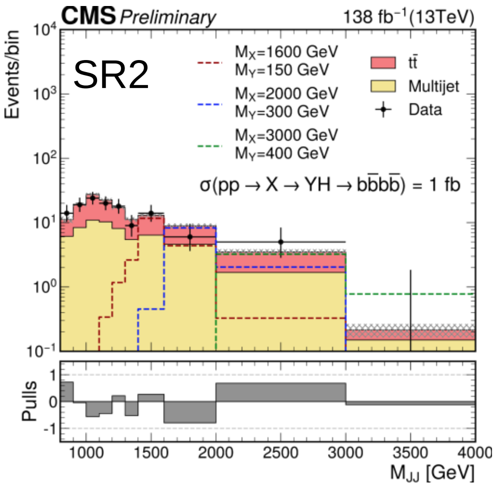


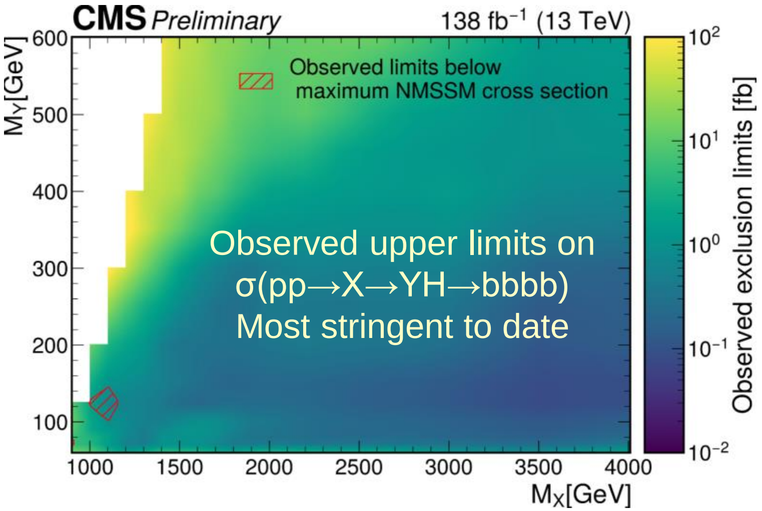
PaticleNet Ops points 0.94, 0.98



- X,Y: scalars, $M_X >> M_{Y(H)}$;
- Models: NMSSM 0910.1785, Two-real-scalar-singlet extension 1908.08554
- 2D search over M_{jj}, M_j^Y variables
- 2 (wide) jets, $m_{H(Y)}$: 110-140(>60) GeV, $|\Delta\eta_{jj}|<1.3$
- Tagging with Graph CNN (ParticleNet), mistag~0.5%, eff~70%, calibration with g→bb jets



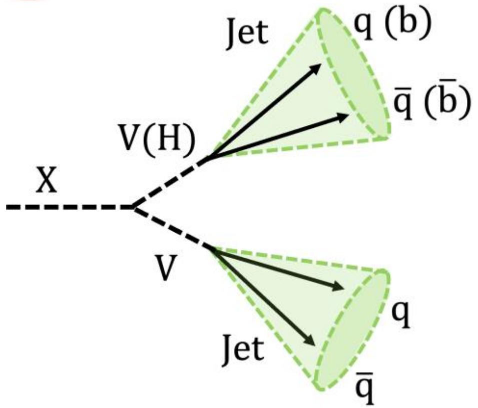




Observed upper limits on σ(pp→X→YH→bbbb)
Most stringent to date

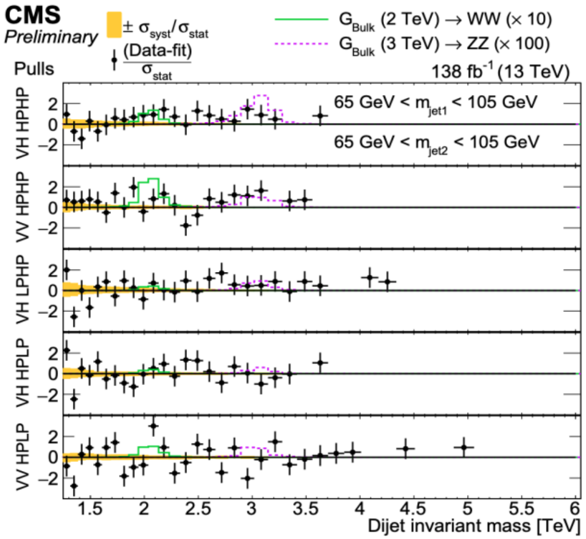
X → VV, VH in DY/gg & VBF

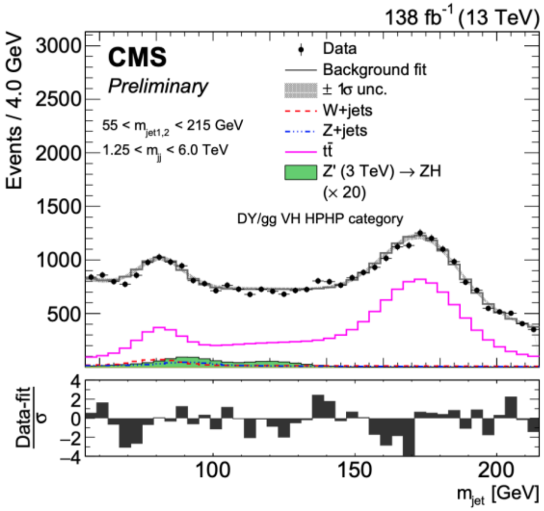
B2G-20-009

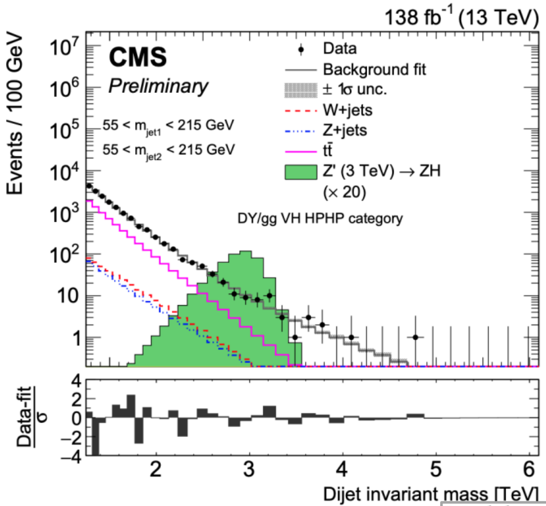


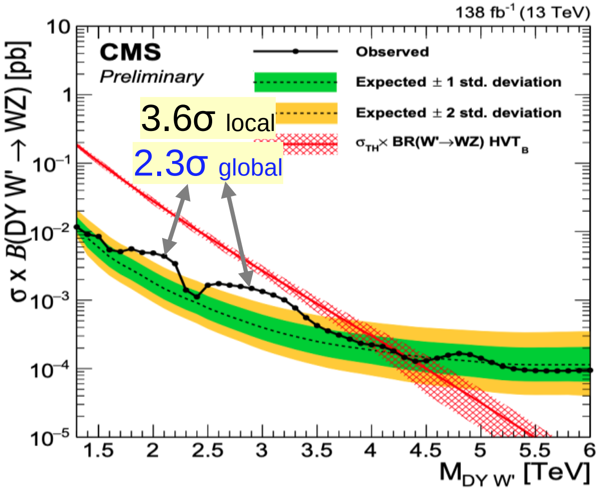
- HVT: W', Z'
- RS Warped ED: R, G_{bulk}

- X → WW, ZZ, WZ, WH, ZH
- Production: DY/ggF & VBF targeted
- 2 wide AK8 jets, $m_{j1,2}$: 55-215 GeV extra 2 narrow fwd jets for VBF
- Tagging with DeepAK8 classifier: W/Z/H → qq/bb, against q/g → 10 categories
- Method: 3D-fit of m_{j1}, m_{j2}, m_{jj}









3.6σ local

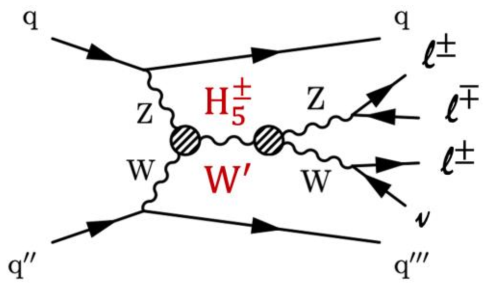
2.3σ global

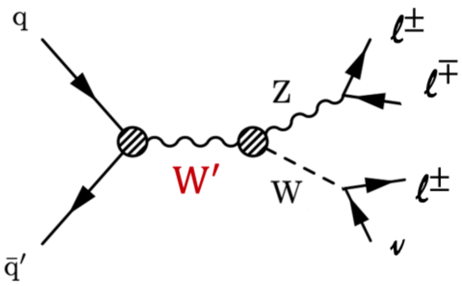
- Best limits to date →
- 2 modest excesses for W'→WZ at ~2.1, 2.9 TeV

Model		Observed limit (TeV)
Radion DY/gg	VV	2.7
HVT model B, W′	WZ / WH	4.4 / 4.0
HVT model B, Z′	WW / ZH	(1.3–3.1, 3.3–3.5) / 3.9
HVT model B, V′	VV+VH / VV / VH	4.8 / 4.5 / 4.2
G_{bulk} ($\tilde{\kappa} = 0.5$) DY/gg	VV	1.4

HDBS-2018-19

$X^{\pm} \rightarrow W^{\pm}Z \rightarrow \ell^{\pm}\nu\,\ell\,\ell$ via VBF & DY





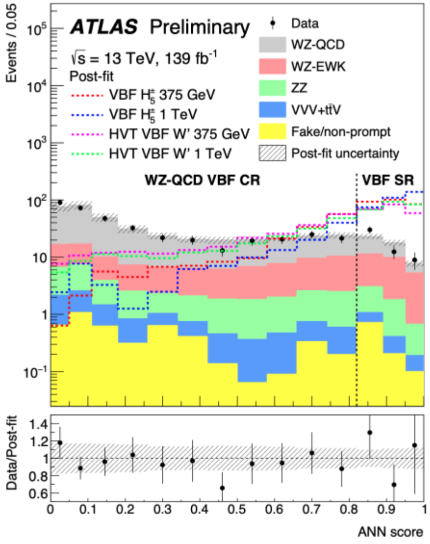
- Georgi-Machacek (GM) model NPB, PLB
  → Fermiophobic 5-plet of scalars: $H_5^{\pm\pm}, H_5^{\pm}, H_5^{0}$
  → $H_5^{\pm}$ is probed here in VBF (WZ-fusion) topology
- HVT model: $W'^{\pm}$ via DY or WZ-fusion

- Fully lep. (ℓ =e/μ)
- Low BKG, but also low BRs
  → best for low m_X

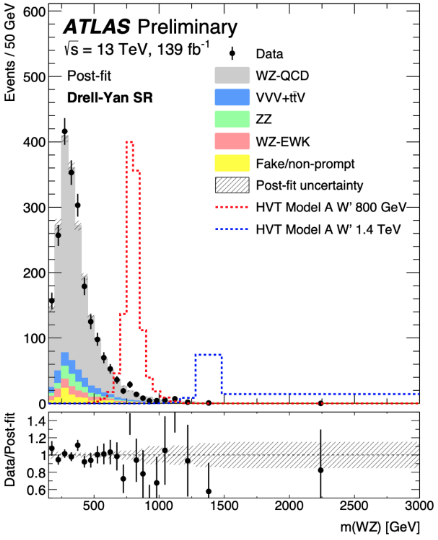
- VBF: 2 jets, m_{jj}>100 GeV, ANN classification
- DY: pT imbalance: $\frac{p_T^V}{m_{WZ}} > 0.35$

- Reconstruct $Z \rightarrow \ell\ell$, $W \rightarrow \nu\ell$, and m_{WZ}
- Prompt-ℓ BKG from MC, fake from data

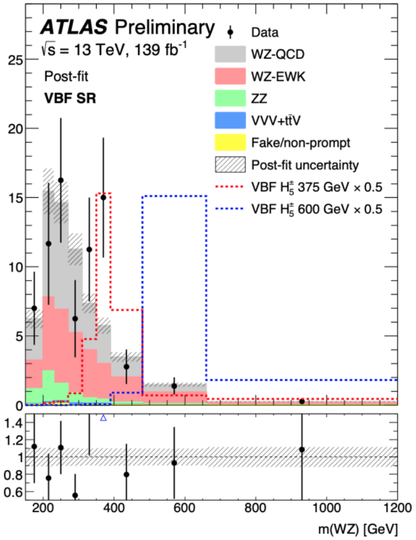
ANN for VBF



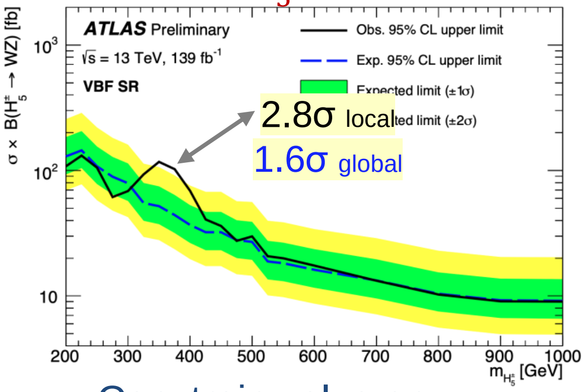
SR for DY



SR for VBF



Limits on $H_5^{\pm}$ x-sec in VBF

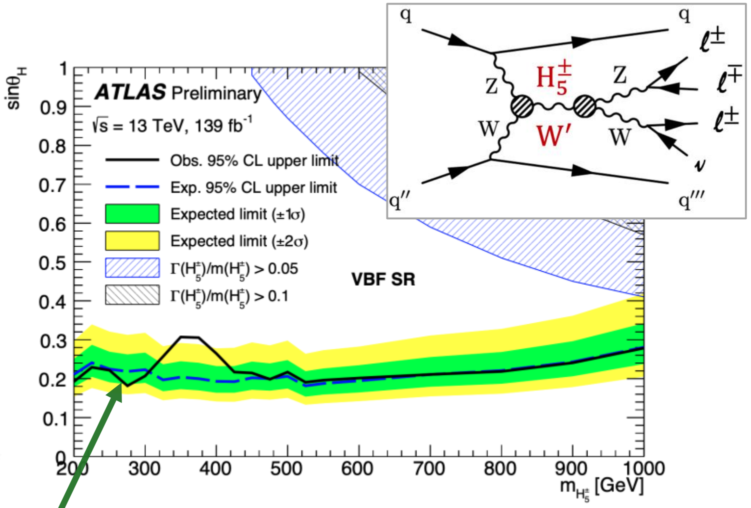


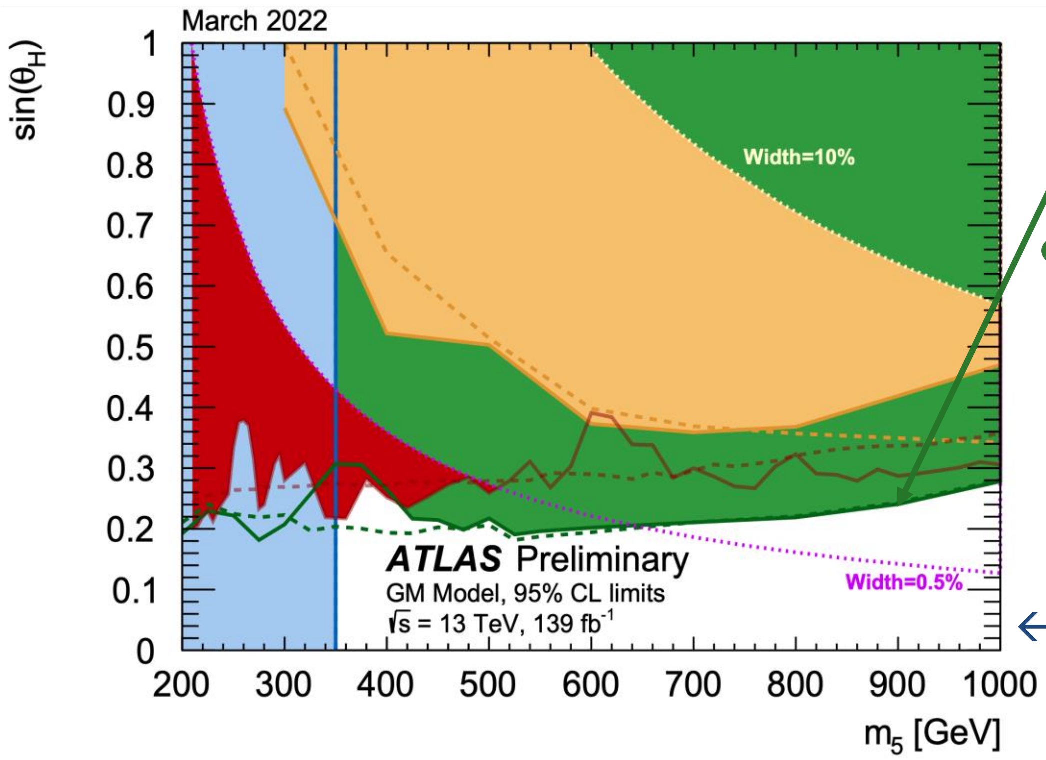
2.8σ local
1.6σ global

Constrains also on:
- param. $\sin\theta_H$ in GM
- HVT models A(B) $m_{W'}$ ~ 2.4(2.5) TeV

Interpretation of BSM Higgs searches in GM model

- 4 BSM Higgs searches are interpreted in Georgi-Machacek model NPB, PLB
- “H5 plane” used, it refers to the 5-plet 1610.07922
  - m_5: mass scale of 5-plet
  - $\sin\theta_H$: mixing parameter (of the VEVs)
- Constrains obtained comparing x-sec limits and theoretical predictions





🟧 $R \to VV$ (semi-leptonic) VBF
Eur. Phys. J. C 80 (2020) 1165

🟩 $H^\pm \to W^\pm Z$ VBF ←Tightest constraint
ATLAS-CONF-2022-005

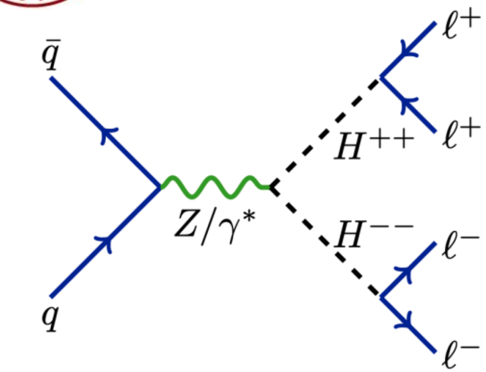
🟥 $H \to ZZ \to 4l + ll\nu\nu$ VBF
Eur. Phys. J. C 81 (2021) 332

🟦 $pp \to H^{\pm\pm}H^{\mp\mp} \to W^\pm W^\pm W^\mp W^\mp$
Via DY → x-sec independent of H5 plane param.
JHEP 06 (2021) 146

←White area is not excluded

$H^{++}H^{--} \rightarrow 4\ell$

ATLAS-CONF-2022-010
EXOT-2018-34

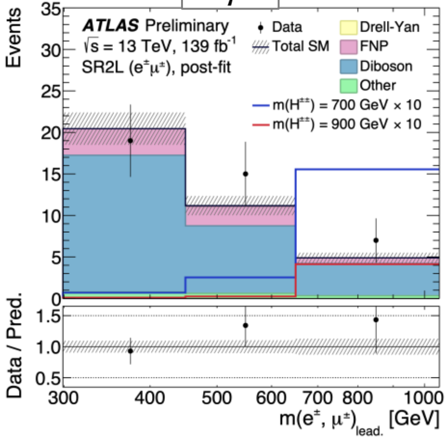


- Left-Right Sym. Models production via DY
- 3-3-1 models 1806.04536
- See-Saw type-II hep-ph/0305288
- Georgi-Machacek NPB, PLB

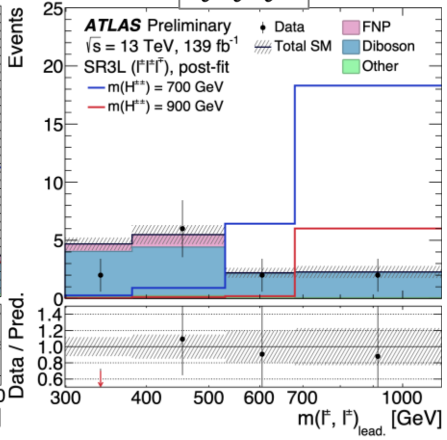
- 2,3,4 ℓ (ℓ=e/μ)
- At least 1 SS ℓ pair (generic probe)
  SS ℓ resonance: BKG-free signature
- 5 SRs: 2ℓ: μμ, ee, eμ →
  3ℓ: all together
  4ℓ: 2 SS pairs
- Observable: $m(\ell^\pm,\ell^\pm)_{\text{lead}} > 300$ GeV
- BKGs: VV, Fake & NonPrompt (FNP)
- CRs & VRs sidebands for prediction

signal regions	SR2L	SR3L	SR4L
Channel	$e^\pm e^\pm$ $e^\pm\mu^\pm$ $\mu^\pm\mu^\pm$	$\ell^\pm\ell^\pm\ell^\mp$	$\ell^+\ell^+\ell^-\ell^-$
Nr. Leptons	2	3	4
$m(\ell^\pm,\ell^\pm)_{\text{lead.}}$ [GeV]	≥ 300	≥ 300	≥ 300
$p_T(\ell^\pm,\ell^\pm)_{\text{lead.}}$ [GeV]	≥ 300	≥ 300	-
$\Delta R(\ell^\pm,\ell^\pm)_{\text{lead.}}$	< 3.5	-	-
$\bar{M}$ [GeV]	-	-	≥ 300
Z-veto	-	✓	✓

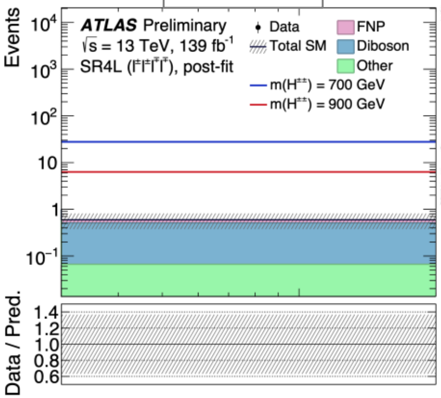
SR2L
$e^\pm\mu^\pm$



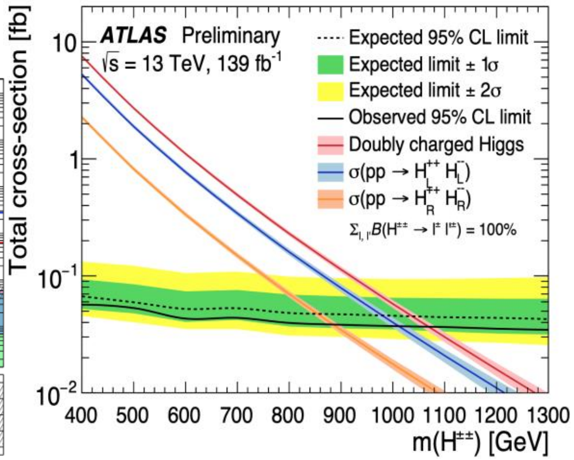
SR3L
$\ell^\pm\ell^\pm\ell^\mp$



SR4L
$\ell^\pm\ell^\pm\ell^\mp\ell^\mp$



Upper limit on $H^{++}H^{--}$ production x-sec assuming $B_{H\rightarrow\mu\mu} = B_{H\rightarrow e\mu} = B_{H\rightarrow ee} = 1/3$



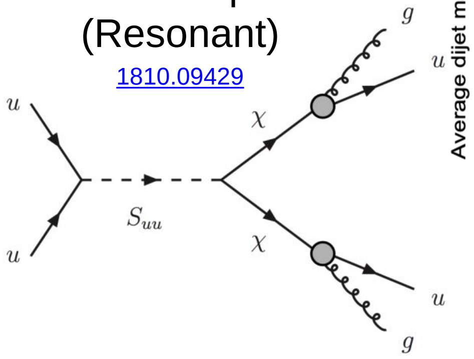
(Y) → XX → (jj)(jj) paired di-jets

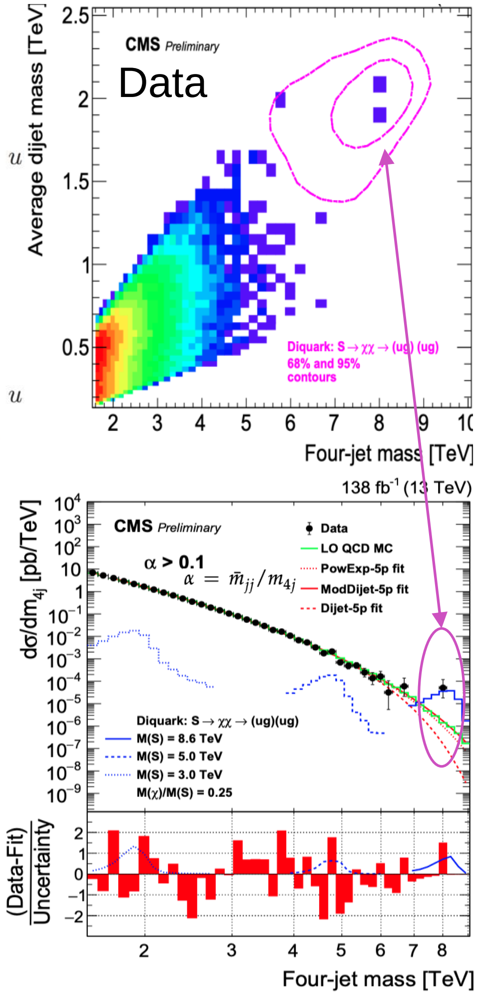
EXO-21-010

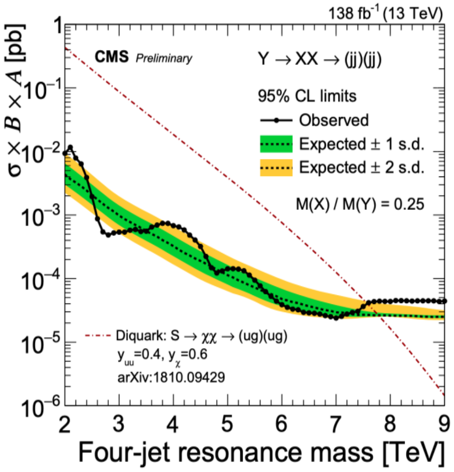
- 4 narrow jets → paired to 2 di-jets, symmetrized masses: $\frac{|m_1 - m_2|}{m_1 + m_2} < 0.1$
- Search over: m_{4j} and average di-jet mass $\bar{m}_{jj}$; fit 3p-function to the data in slices of $\frac{\bar{m}_{jj}}{m_{4j}}$

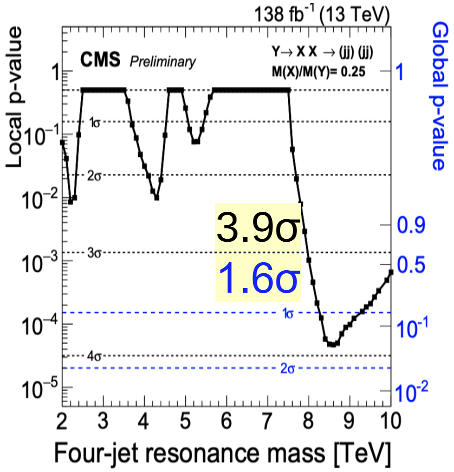
Ultraheavy di-quark to vec-like quarks (Resonant)

1810.09429



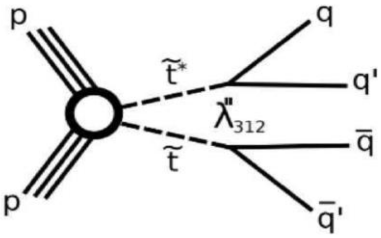


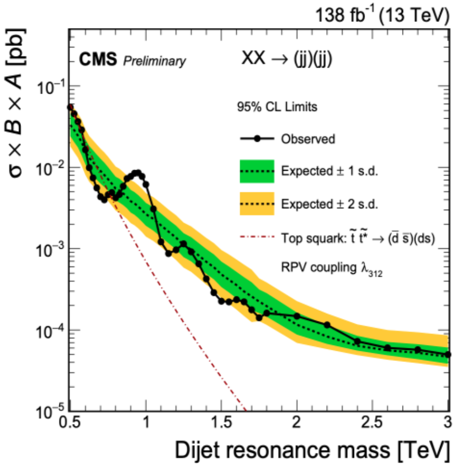


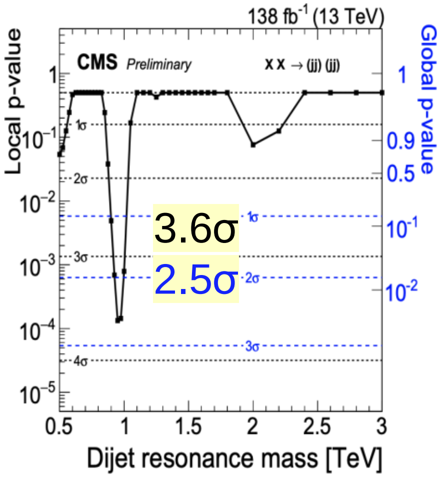


RPV SUSY (non-resonant)

1209.0764







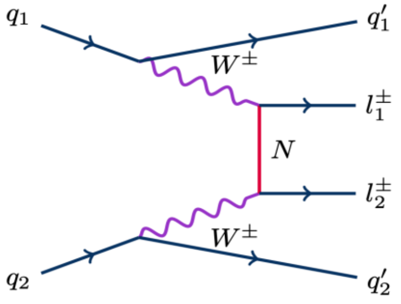
VBF → N → $\mu^{\pm}\mu^{\pm}$ Heavy Majorana N. & W.O. Probe

EXO-21-003

The “neutrinoless double-β decay” version of the LHC

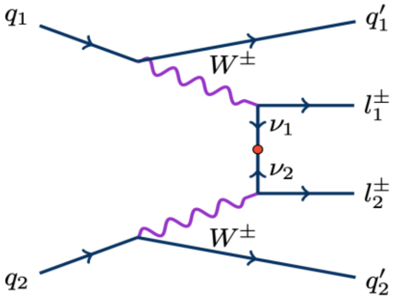
VBF HMN at Seesaw Type-I

mixing element $|V_{\mu N}|^2 = 1$

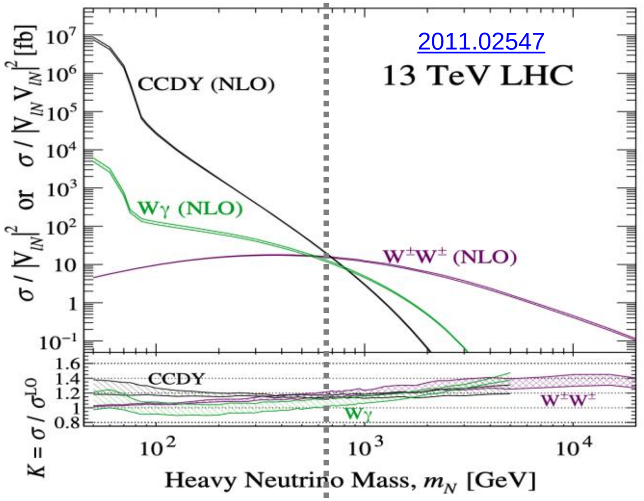


EFT Weinberg Oper. Dim. 5

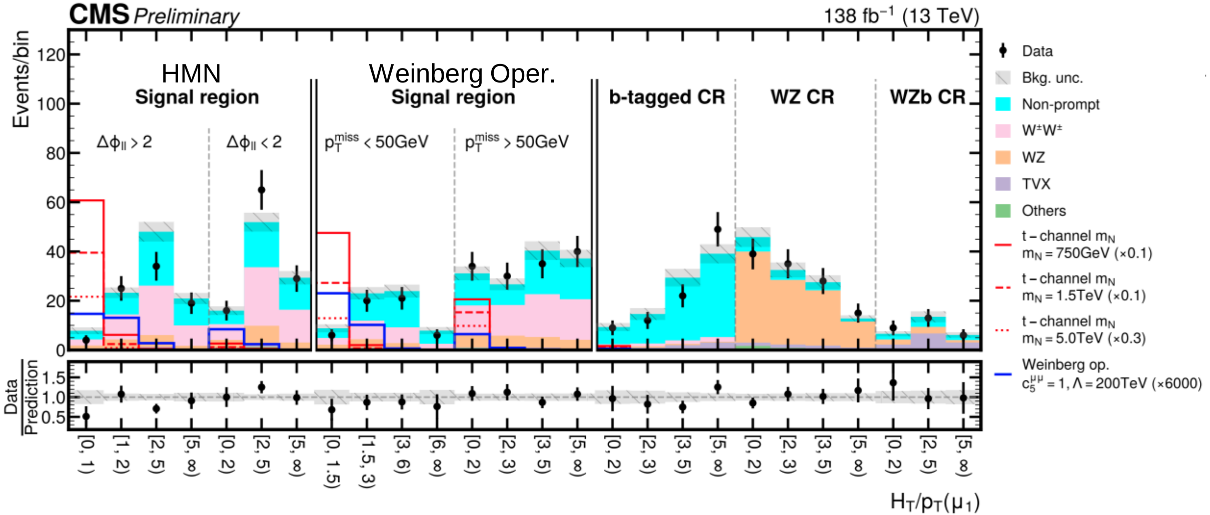
Wilson coef. $C_5^{ll'} = 1$

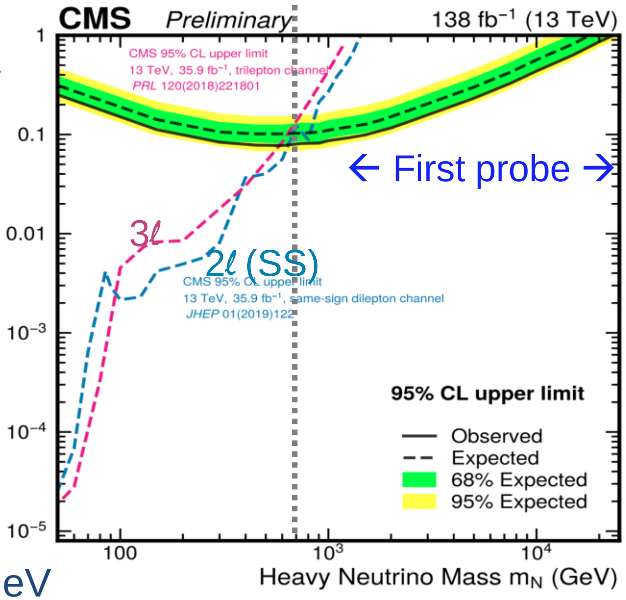


qq→W→ℓN direct ⟺ VBF t-channel



- Use VBF t-channel which dominates high m_N for first time
- 2μ SS + 2 fwd jets; cuts on $\Delta\eta_{jj}$, m_{jj}; fit on $H_T/p_{T\ell_1}$:





← First probe →

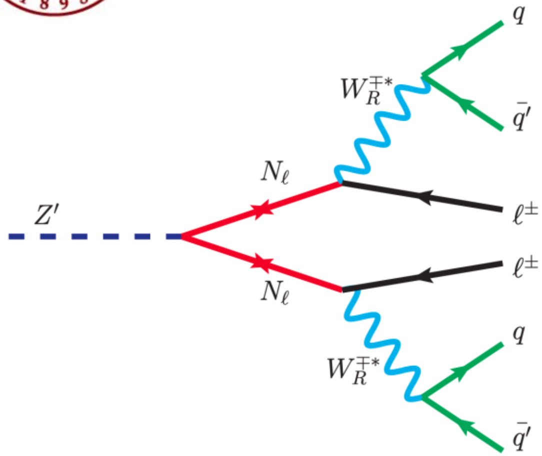
3ℓ

2ℓ (SS)

- HMN: excluded up to m_N~23 TeV
- WO: upper limit on eff. mass $|m_{\mu\mu}| \sim C_5^{ll'}$: obs(exp): 10.8(12.8) GeV

Z’ → NN → ℓ jj ℓ jj Heavy Majorana Neutrino pair

EXO-20-006

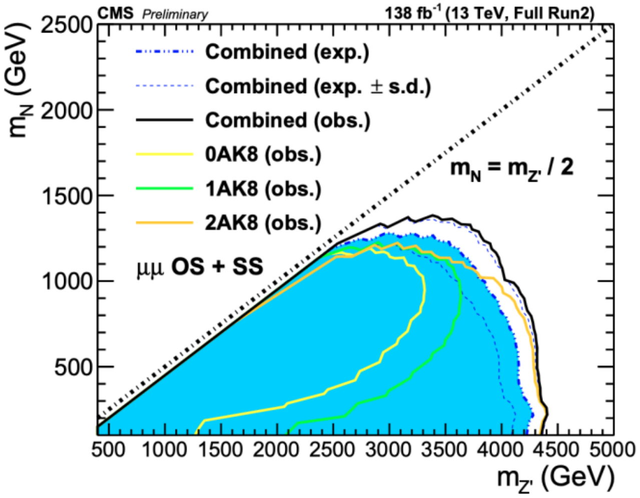


- LRSM: $Z', W_R^\pm, N_{e/\mu/\tau}$
- $m_{N_\ell} < m_{W_R^{\pm *}} = 5$ TeV
- Off-shell $W_R^{\pm *}$, no mixing

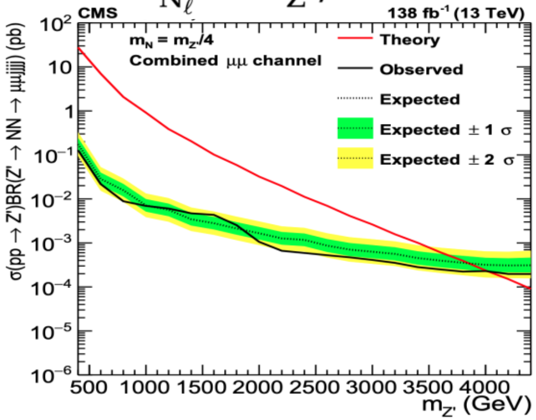
- ee, μμ (OS & SS), $m_{\ell\ell} > 150$ GeV
- Resolved & Boosted probed
- Binning on # of wide AK8 jets:

SR	N(AK8 jet)	N(tight leptons)	N(AK4 jet)
SR1 (0AK8)	$= 0$	$= 2$	≥ 4
SR2 (1AK8)	$= 1$	≥ 1	≥ 2
SR3 (2AK8)	≥ 2	—	—

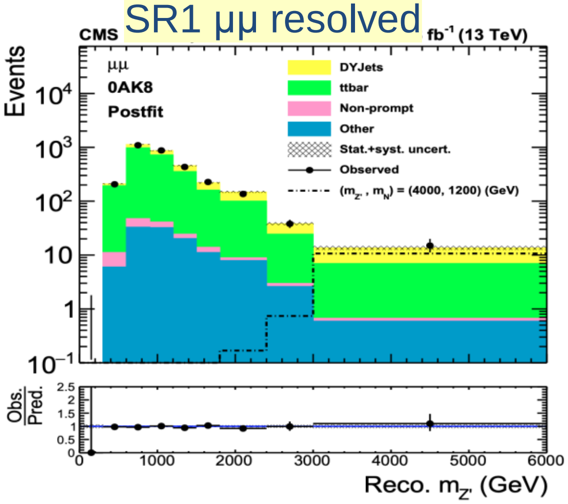
- Reconstruct N_ℓ as “jjℓ” and $m_{Z'}$ minimizing $m(jj\ell)$-asymmetry
- Prediction from eμ, $m_{\ell\ell}$ SBs



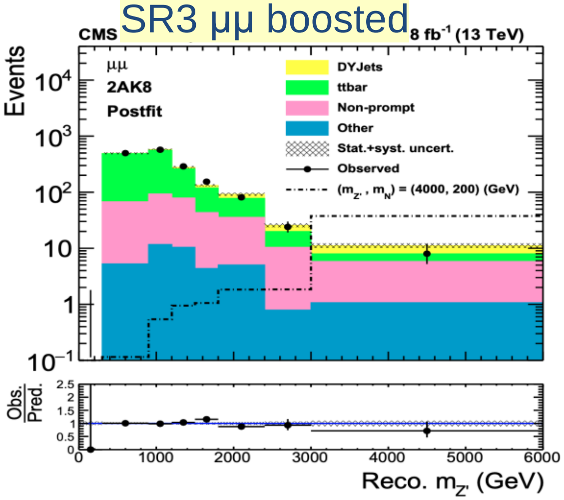
$m_{N_\ell} = m_{Z'}/4$



SR1 μμ resolved



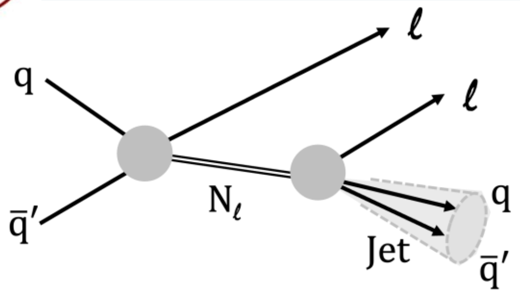
SR3 μμ boosted



- First search of this type for Run2
- Best direct limits on $m_{Z'}, m_N$ plane

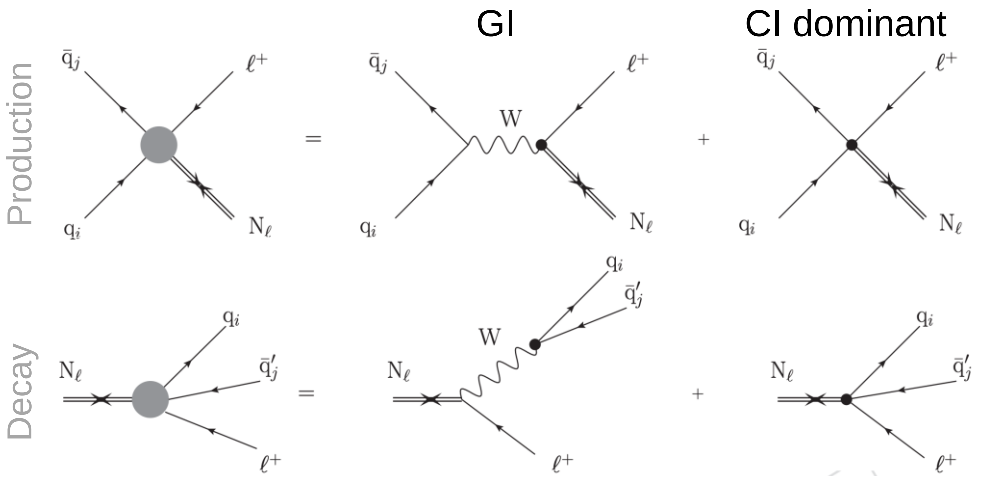
EXO-20-011

Heavy Composite Majorana Neutrino $N\ell \rightarrow \ell\ell(qq)$



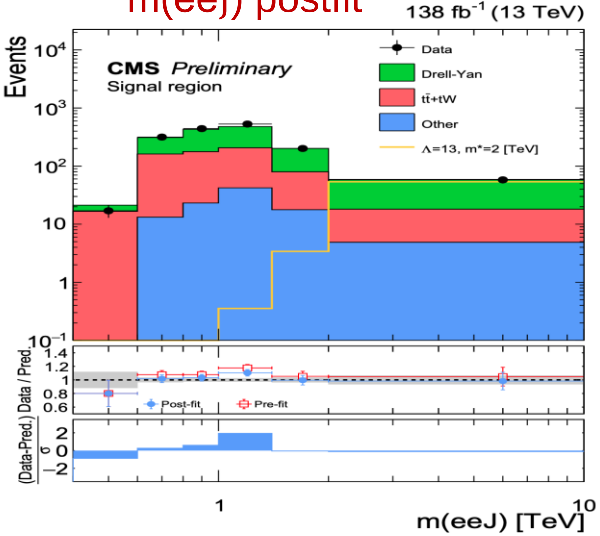
- Composite-fermion models
  1510.07988, 1707.00844, 1810.00374, 1903.12285

  → Excited states of SM fermions
  → Effective interactions: gauge (GI) & contact (CI) between ordinary and excited fermions
  → $m(N_\ell)$: [500 GeV, Λ]



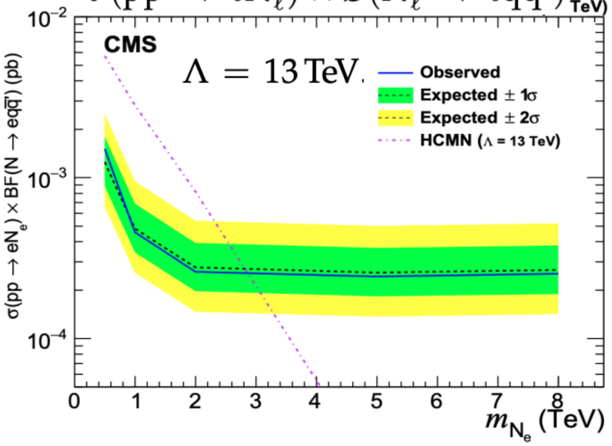
- ee, μμ (SS&OS), $m(\ell\ell)$>300 GeV, ≥1 wide AK8 jet
- Use eμ, $m_{\ell\ell}$: 150-300 GeV as CRs
- Fit: $m(\ell\ell J)$ constrain separately N_μ, N_e masses

m(eeJ) postfit

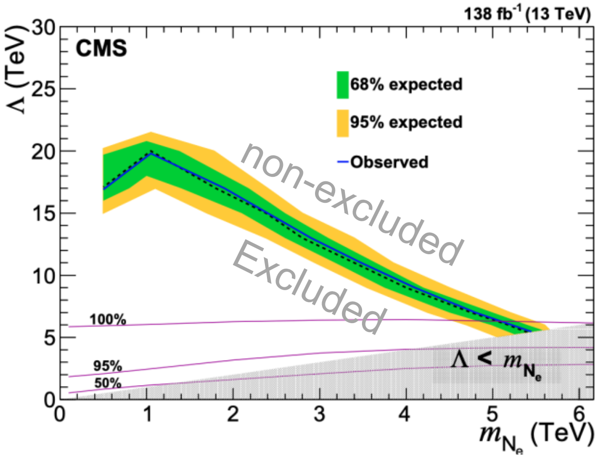


Upper x-sec. limits on:

$\sigma(pp \rightarrow \ell N_\ell) \times \mathcal{B}(N_\ell \rightarrow \ell q\bar{q}')$



Limits on $m(N_\ell)$, Λ plane



A toolkit for BSM searches: Objects

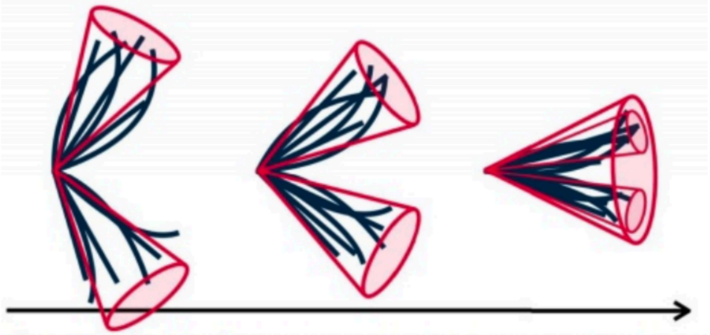
Boosted objects → small angular separation → merged jets

(W/Z→qq; H→bb/qqqq/qqlν)

→ anti-kt clustering

→ Large-R jets: $\Delta R = \sqrt{(\Delta\varphi^2 + \Delta\eta^2)} \approx 2m/p_T$

→ ”groomed” Masses: $M_J \sim M_V \pm 0.2\, M_V$



Boosted jets: Increasing transverse momentum

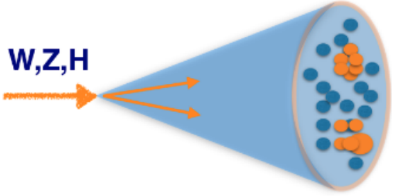
Taggers based on (2-prong) substructure

CMS:

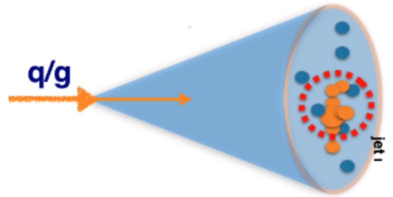
- τ_N = N-subjettiness → ratios: $\tau_2/\tau_1 = \tau_{21}$ → $\tau_N = \frac{1}{d_0}\sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$
- Decorrelated taggers τ_{21}^{DDT}

ATLAS:

- D_2: 2-point E_{cf} / 3-point E_{cf}
- TCC-jets Track-Calo Clusters
  algo unifying tracker & calo info.



$$D_2^{\beta=1} = E_{\mathrm{CF3}} \left(\frac{E_{\mathrm{CF1}}}{E_{\mathrm{CF2}}}\right)^3$$



$$E_{\mathrm{CF1}} = \sum_i p_{\mathrm{T},i}$$

$$E_{\mathrm{CF2}} = \sum_{ij} p_{\mathrm{T},i} p_{\mathrm{T},j} \Delta R_{ij}$$

$$E_{\mathrm{CF3}} = \sum_{ijk} p_{\mathrm{T},i} p_{\mathrm{T},j} p_{\mathrm{T},k} \Delta R_{ij} \Delta R_{jk} \Delta R_{ki}$$

Deep-NN taggers & Image taggers (soon)

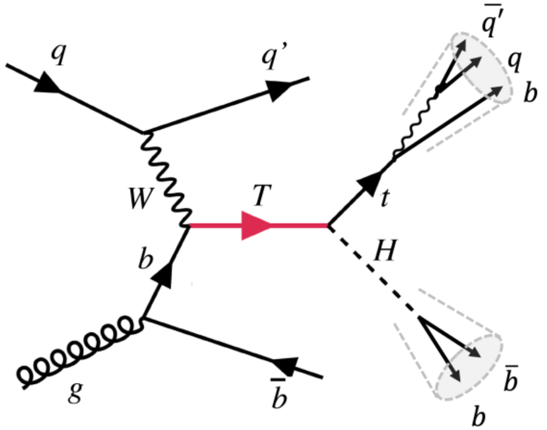
MET + lep from Boson:

→ Reco the W(H) assuming $M_{W(H)}$=80(125) GeV

b-jet tagging based on MVA, DNN

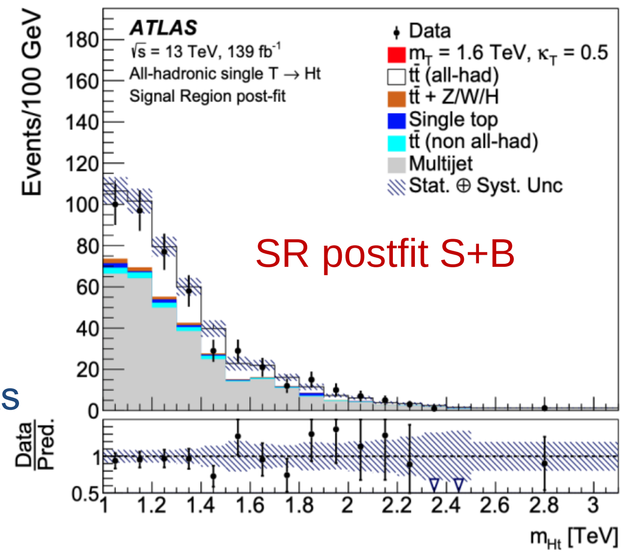
Single Vector-like T → Ht → (bb)(bqq) boosted

CERN-EP-2021-249



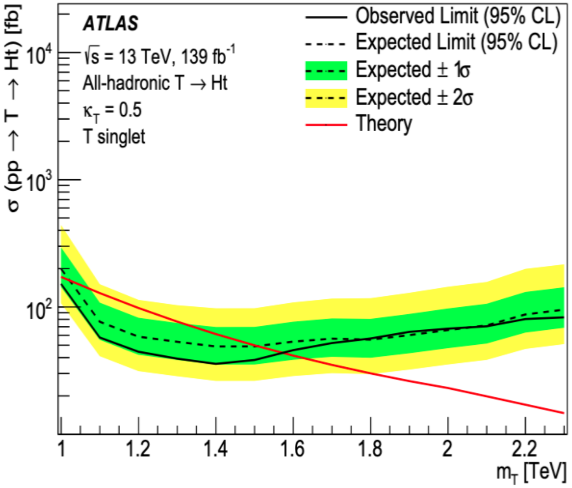
- T assumed to couple to 3rd gen.
- Here: $BR_{T\to Ht} = 1/4$, $m_T > 1$TeV
- κ_T controls production coupling

- 2 wide jets (R=1.0)
- Tagging with:
  - H: m_j: 100-140 GeV, τ_{21}
  - t: m_j: 140-225 GeV, DNN
  - b: RTrack, DL1 (DNN)
- Classification 81 categories based on: "H, t, other" & "0b, 1b, 2b" jets → SR, VR, NR
- Search over m_{jj} → QCD & tt, data-driven pred.

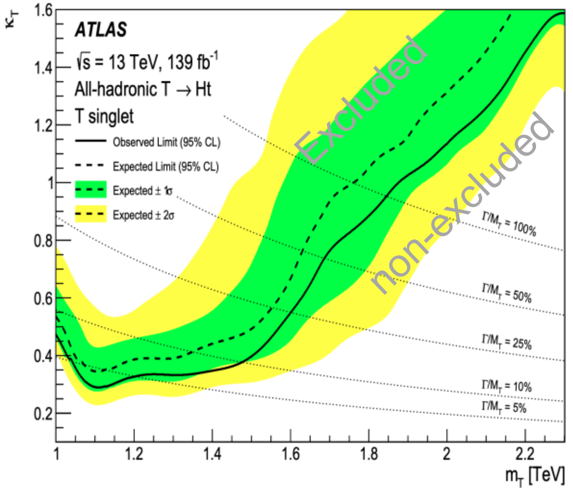


SR postfit S+B

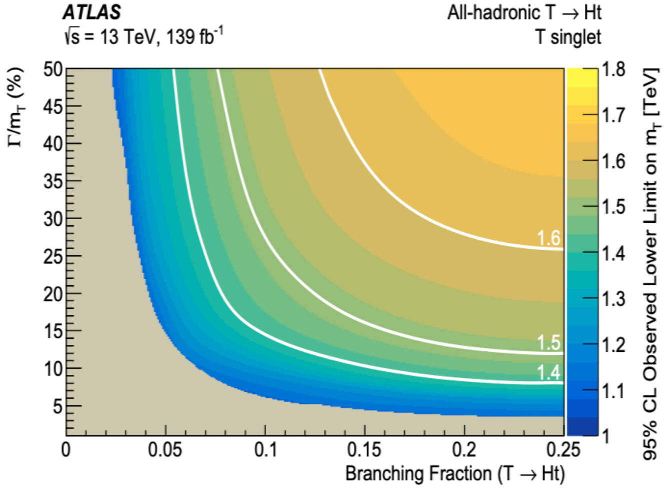
Limits on x-sec. for $\kappa_T = 0.5$



Constraints on coupling κ_T vs m_T



Obs. limit on x-sec. vs BR_{Ht} , Γ/m_T



ATL-PHYS-PUB-2022-012

EXOT-2022-01

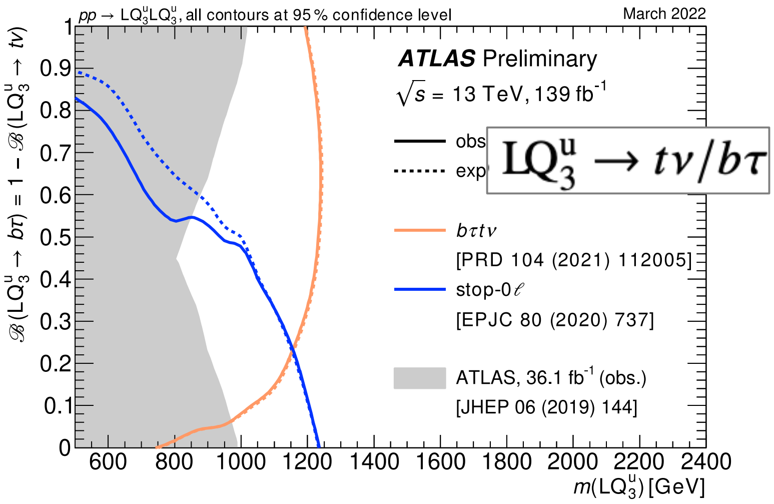
Summary plots for scalar LQ-pairs production

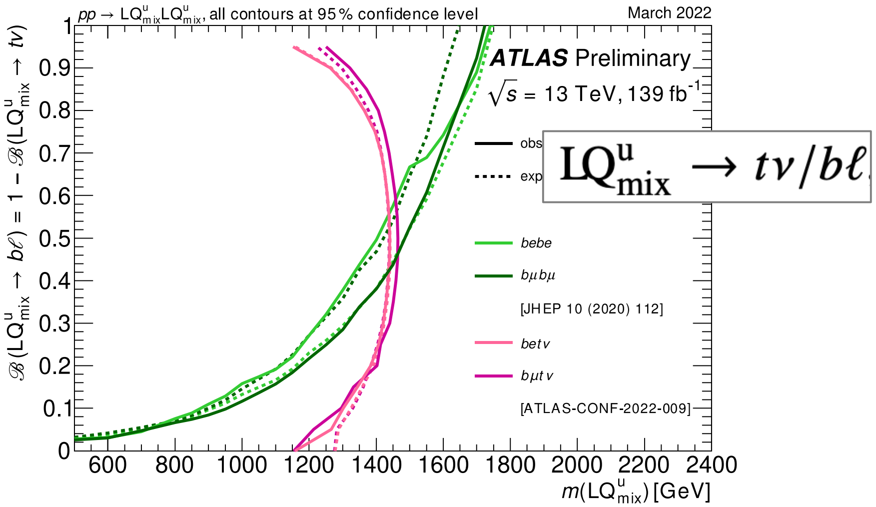
- 4 scenarios; all decays in 3rd gen quarks as final state
- 5 dedicated LQ searched + 2 SUSY searches re-interpretations used → exclusion m_{LQ}~1.2 TeV

LQ → 3rd gen

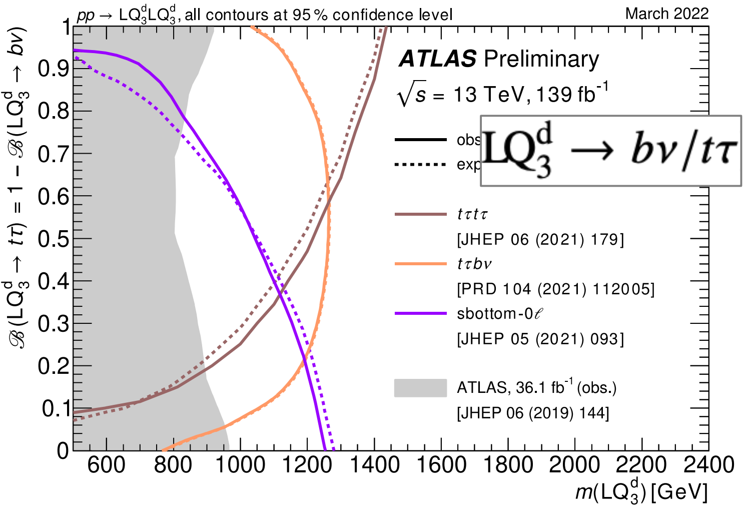
LQ → mixed gen

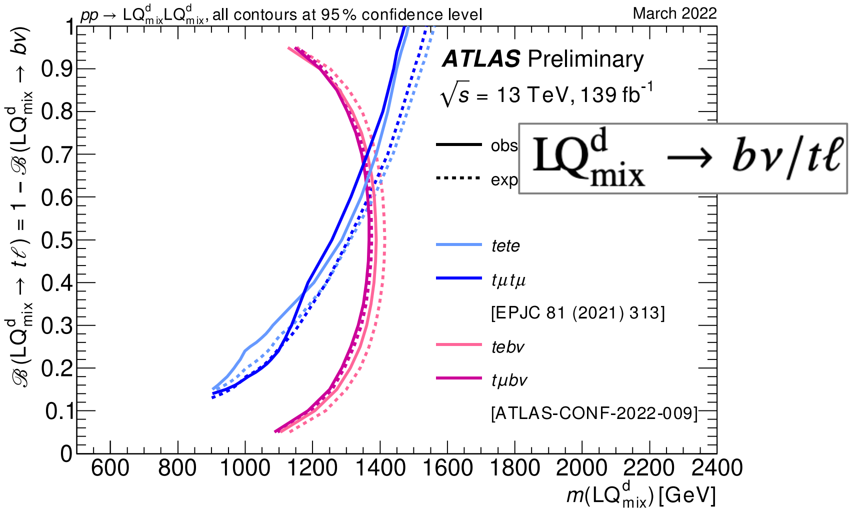
LQ up-type





LQ down-type





Resonant Triboson: X → WR→ WWW

Extended (3 branes) Warped ED model

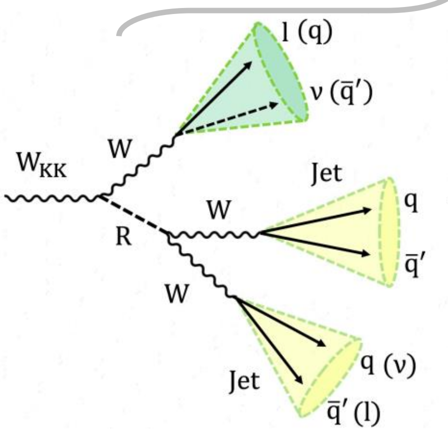
1711.09920, 1612.00047, 1809.07334

B2G-20-001

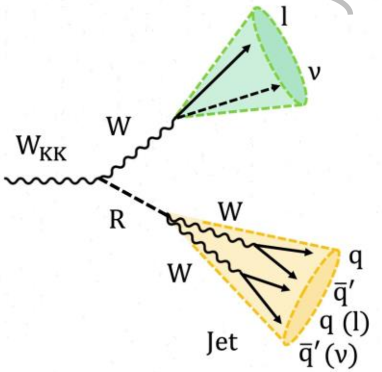
1-lep + jets

Full-hadronic

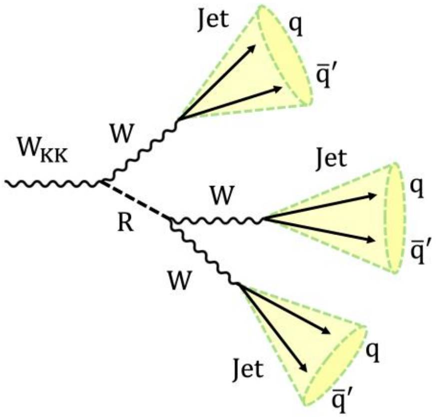
B2G-21-002



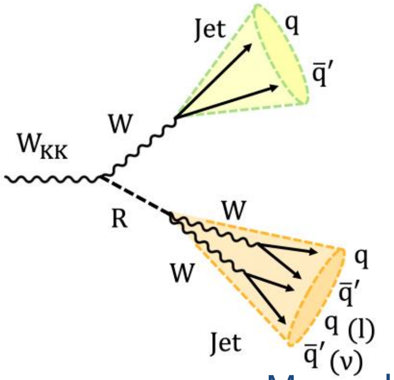
Resolved R



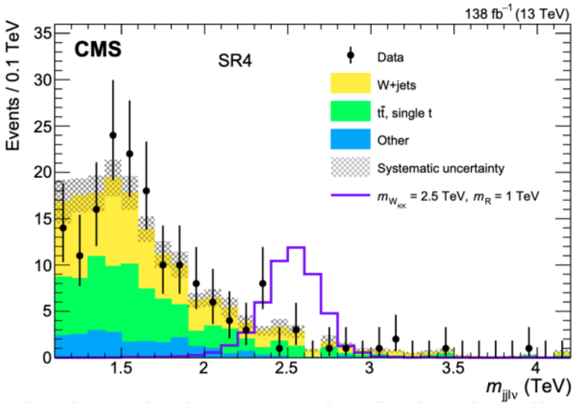
Merged R

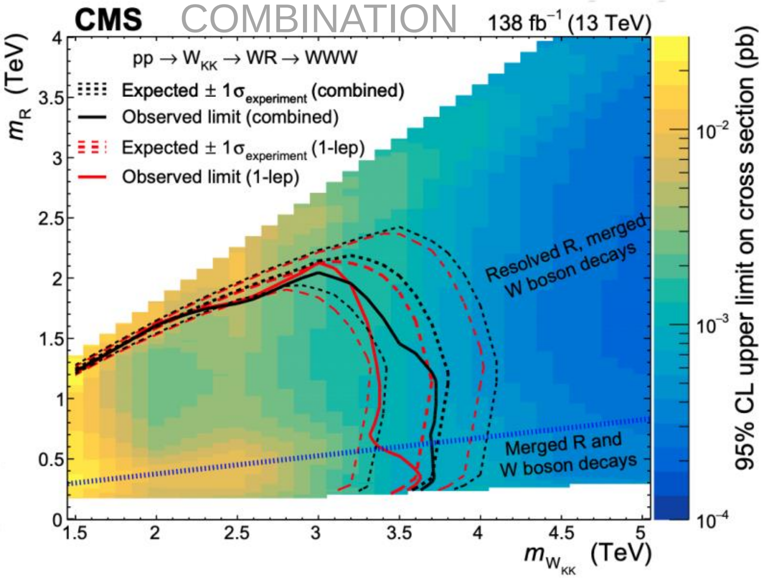


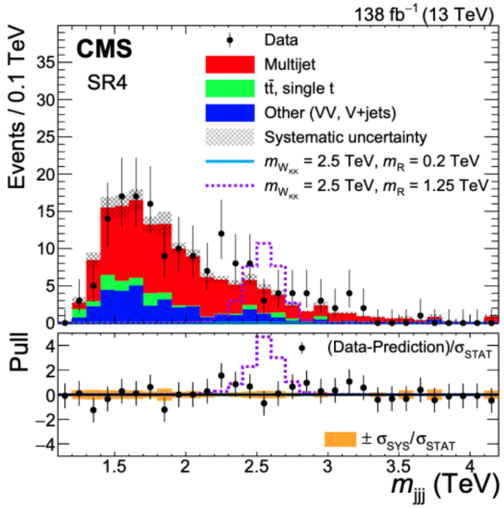
Resolved R



Merged R





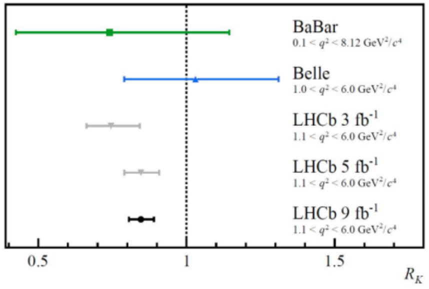


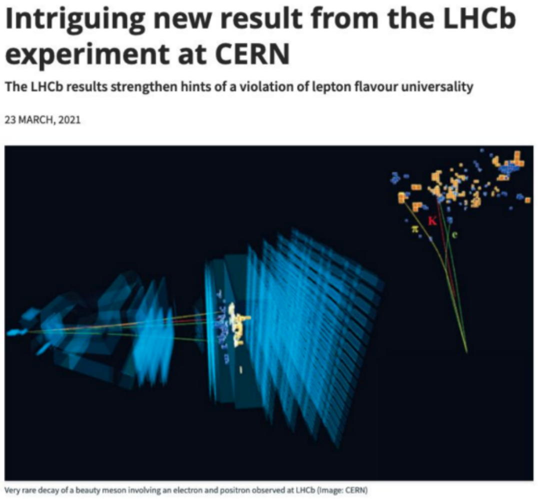
X→LL→ ≥3b,ττ/τν/νν vector-like lep. pair

B2G-21-004

New results from LHCb showing hints of **lepton flavor universality violation**

- Difference between **measurement** and **prediction** now at **3.1σ** for **R_K**
- Interestingly, the **electron** measurement alone is **compatible with SM**, making the **muon** measurement the one that **deviates** in the ratio





Intriguing new result from the LHCb experiment at CERN

The LHCb results strengthen hints of a violation of lepton flavour universality

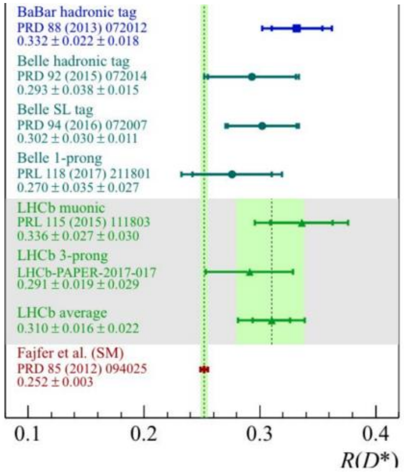
23 MARCH, 2021

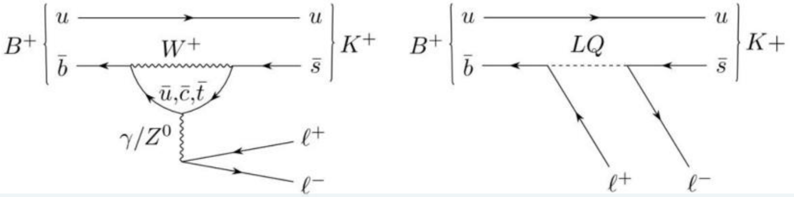
Very rare decay of a beauty meson involving an electron and positron observed at LHCb (Image: CERN)

$$R_K = \frac{\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)} \bigg/ \frac{\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathcal{B}(B^+ \to J/\psi(\to e^+e^-)K^+)} \, .$$

Similar results also present for **R_D^* measurements**

- Deviation from SM at **3.4σ**
- The **tau measurement** is even **further** from SM **compared** to the **muon** channel
- A combined explanation of both anomalies hint to BSM with **yukawa-like structure favoring** to **third generation families**
- Among the many options, one that provides a combined explanation for both anomalies is the **4321** model, extending the SM sector: **SU(4)xSU(3)xSU(2)xU(1)**
- A **Leptoquark** (**U_1**) is then predicted to exist as the source of **LFV**

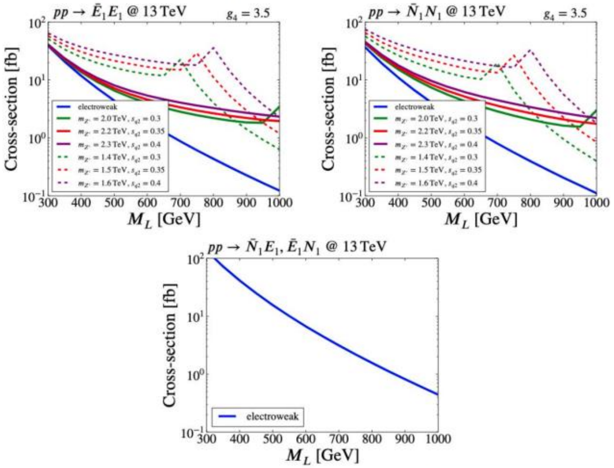




$$\begin{aligned}\mathcal{R}(D^{(*)-}) &\equiv \mathcal{B}(B^0 \to D^{(*)-}\tau^+\nu_\tau)/\mathcal{B}(B^0 \to D^{(*)-}\mu^+\nu_\mu), \\ \mathcal{R}(D^{(*)0}) &\equiv \mathcal{B}(B^- \to D^{(*)0}\tau^-\overline{\nu}_\tau)/\mathcal{B}(B^- \to D^{(*)0}\mu^-\overline{\nu}_\mu),\end{aligned}$$

X→LL→ ≥3b,ττ/τν/νν vector-like lep. pair

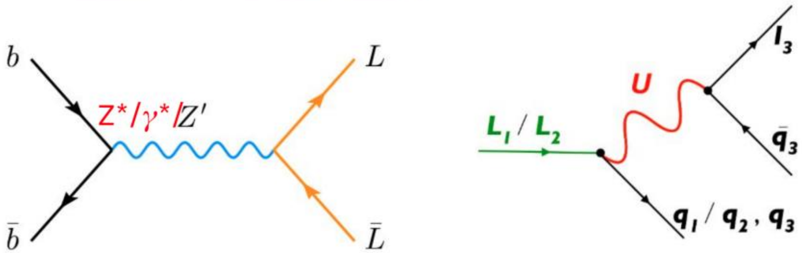
Di Luzio, L., Fuentes-Martin, J., Greljo, A. *et al.* Maximal flavour violation: a Cabibbo mechanism for leptoquarks. *J. High Energ. Phys.* **2018,** 81 (2018).



The SU(4) extension also brings **new families of vector-like fermions**

In this talk: vector-like leptons

- They are produced through **Electroweak** interactions or pair produced with a **Z'**, also required to exist due to **UV completion**
- Coupling to SM fermions through the leptoquark U_1 resulting in a **3-body decay**



The VLL is a multiplet with a neutral (**N**) and charged (**E**) components

- We will focus on **second generation VLLs** coupling to **third generation fermions**
- In the resolved case, all final states contain at least **4 b jets** and a varying number of **taus**, **neutrinos**, and **light jets** from top decays

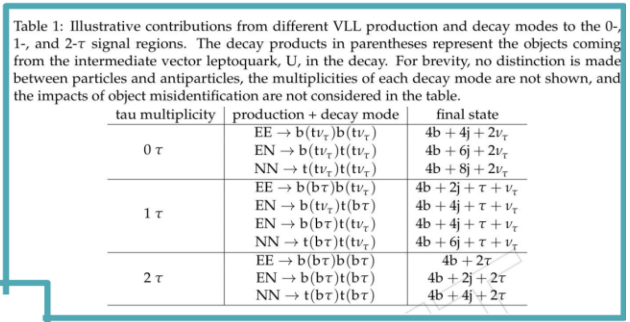


Table 1: Illustrative contributions from different VLL production and decay modes to the 0-, 1-, and 2-τ signal regions. The decay products in parentheses represent the objects coming from the intermediate vector leptoquark, U, in the decay. For brevity, no distinction is made between particles and antiparticles, the multiplicities of each decay mode are not shown, and the impacts of object misidentification are not considered in the table.

tau multiplicity	production + decay mode	final state
0 τ	$EE \to b(t\nu_\tau)b(t\nu_\tau)$	$4b + 4j + 2\nu_\tau$
	$EN \to b(t\nu_\tau)t(t\nu_\tau)$	$4b + 6j + 2\nu_\tau$
	$NN \to t(t\nu_\tau)t(t\nu_\tau)$	$4b + 8j + 2\nu_\tau$
1 τ	$EE \to b(b\tau)b(t\nu_\tau)$	$4b + 2j + \tau + \nu_\tau$
	$EN \to b(t\nu_\tau)t(b\tau)$	$4b + 4j + \tau + \nu_\tau$
	$EN \to b(b\tau)t(t\nu_\tau)$	$4b + 4j + \tau + \nu_\tau$
	$NN \to t(b\tau)t(t\nu_\tau)$	$4b + 6j + \tau + \nu_\tau$
2 τ	$EE \to b(b\tau)b(b\tau)$	$4b + 2\tau$
	$EN \to b(b\tau)t(b\tau)$	$4b + 2j + 2\tau$
	$NN \to t(b\tau)t(b\tau)$	$4b + 4j + 2\tau$

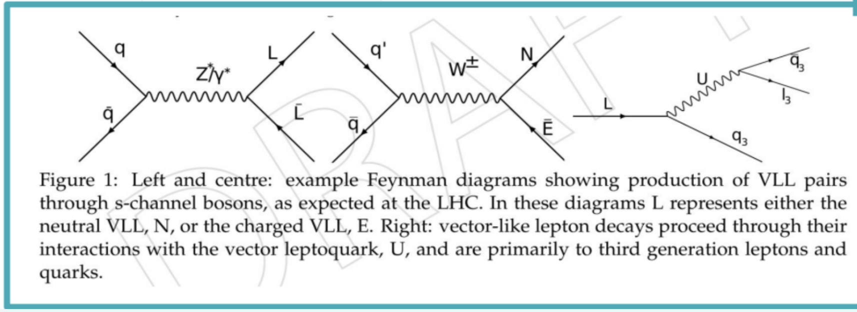
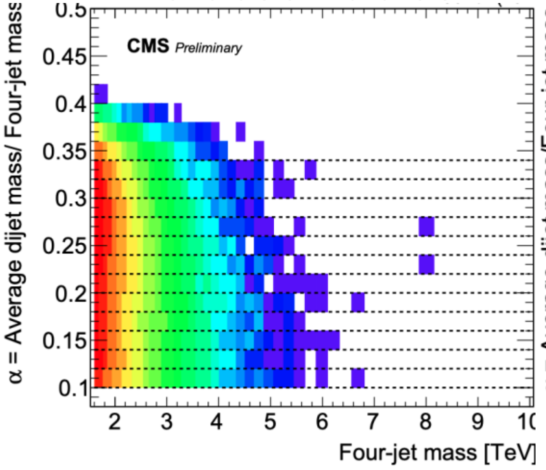


Figure 1: Left and centre: example Feynman diagrams showing production of VLL pairs through s-channel bosons, as expected at the LHC. In these diagrams L represents either the neutral VLL, N, or the charged VLL, E. Right: vector-like lepton decays proceed through their interactions with the vector leptoquark, U, and are primarily to third generation leptons and quarks.

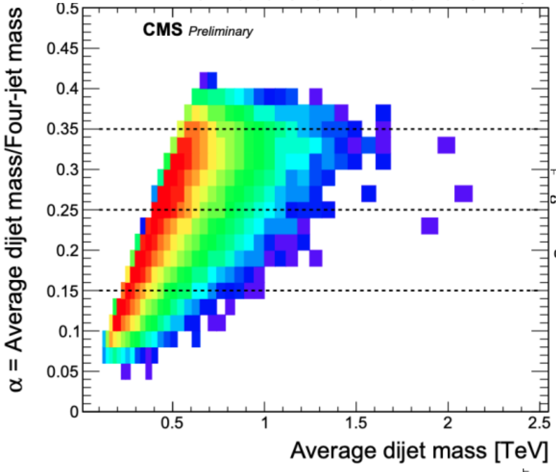
B2G-21-010

(Y) → XX → (jj)(jj) paired di-jets

13 observables



3 observables



$$\Delta R = |(\Delta R_1 - 0.8)| + |(\Delta R_2 - 0.8)|$$

$$\Delta R_1 = \sqrt{(\eta_{j_a} - \eta_{j_b})^2 + (\phi_{j_a} - \phi_{j_b})^2} < 2.0$$

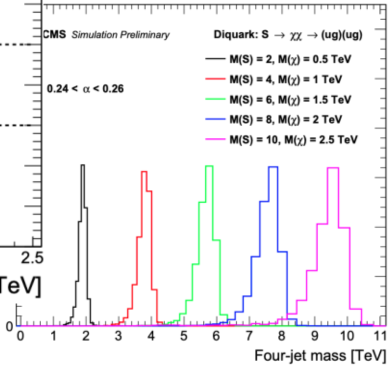
$$\Delta R_2 = \sqrt{(\eta_{j_c} - \eta_{j_d})^2 + (\phi_{j_c} - \phi_{j_d})^2} < 2.0$$

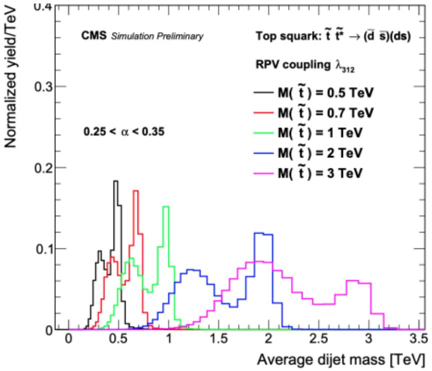
$$\Delta\eta = |\eta_1 - \eta_2| < 1.1$$

$$\text{asymmetry} = \frac{|m_1 - m_2|}{m_1 + m_2} < 0.1$$

$$\bar{m}_{jj} = (m_1 + m_2)/2$$

$$\alpha = \bar{m}_{jj}/m_{4j}$$





Dijet-3p:

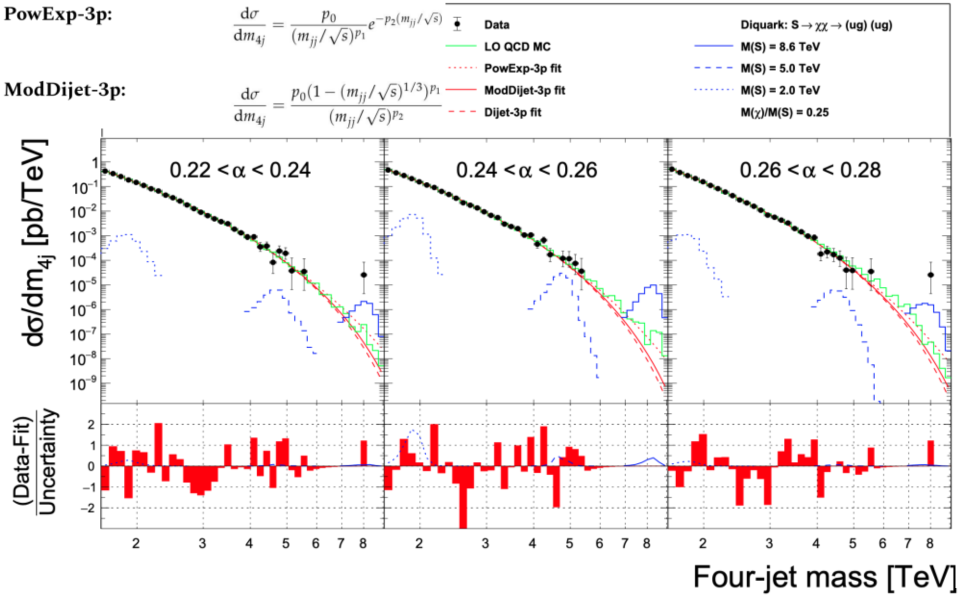
$$\frac{d\sigma}{dm_{4j}} = \frac{p_0(1 - m_{jj}/\sqrt{s})^{p_1}}{(m_{jj}/\sqrt{s})^{p_2}}$$

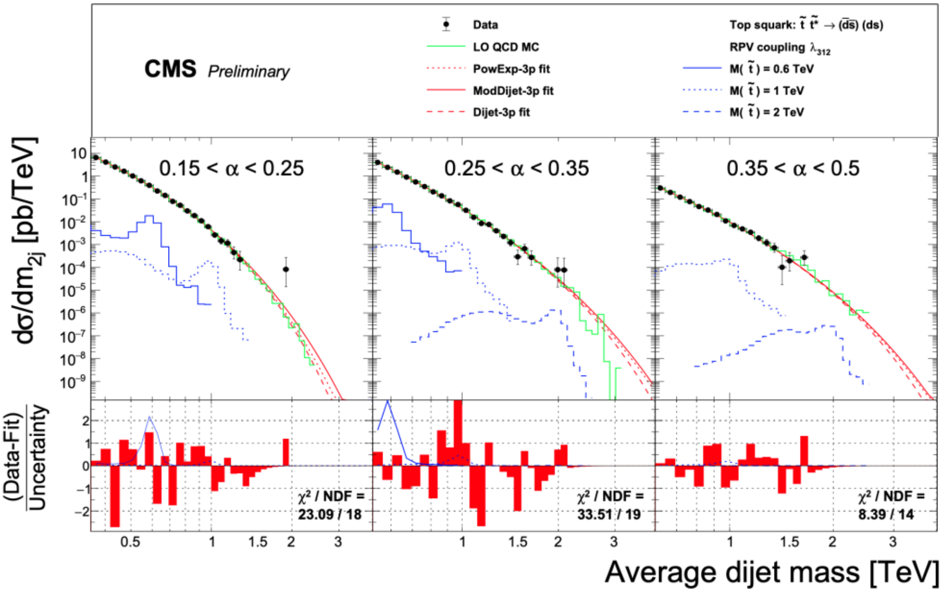
PowExp-3p:

$$\frac{d\sigma}{dm_{4j}} = \frac{p_0}{(m_{jj}/\sqrt{s})^{p_1}} e^{-p_2(m_{jj}/\sqrt{s})}$$

ModDijet-3p:

$$\frac{d\sigma}{dm_{4j}} = \frac{p_0(1 - (m_{jj}/\sqrt{s})^{1/3})^{p_1}}{(m_{jj}/\sqrt{s})^{p_2}}$$



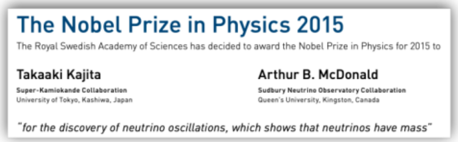


Majorana Neutrinos & Weinberg Op. Probe

EXO-21-003

Introduction: Physics Background

- ➔ Neutrino masses:
  - ❖ Confirmed by neutrino oscillation experiments
  - ❖ Not included in the SM



The Nobel Prize in Physics 2015

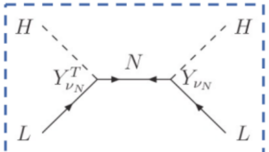
The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Physics for 2015 to

Takaaki Kajita
Super-Kamiokande Collaboration
University of Tokyo, Kashiwa, Japan

Arthur B. McDonald
Sudbury Neutrino Observatory Collaboration
Queen's University, Kingston, Canada

"for the discovery of neutrino oscillations, which shows that neutrinos have mass"

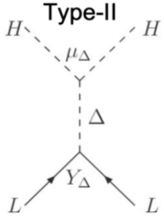
- ➔ Why no neutrino mass mechanism in the SM?
  - ❖ $SU(2)_L \times U(1)_Y$ EW symmetry & only Dimension-4 operators in Lagrangian
  - ❖ Economical particle content:
    - ‣ Only left-hand neutrinos, Dirac mass is thus forbidden.
- ➔ To generate neutrino masses, one must go beyond the SM:
  - ❖ Potential BSM particle solution: Seesaw models

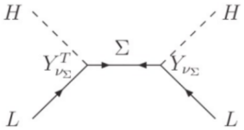


SM + fermion singlet
"Heavy Neutrino"
Type-I

BSM Model for
This analysis

SM + scalar triplet
"Secondary Higgs"
Type-II



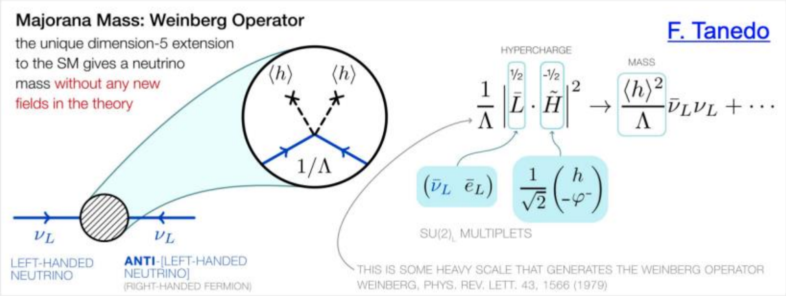


SM + fermion triplet
Type-III

Why named as Seesaw?
The heavier BSM particle is,
The lighter SM neutrino will be!

- ❖ EFT solution: Weinberg Operator

Phys.Rev.Lett. 43 (1979) 1566-1570



Majorana Mass: Weinberg Operator

the unique dimension-5 extension to the SM gives a neutrino mass without any new fields in the theory

F. Tanedo

$$\frac{1}{\Lambda}\left|\bar{L}\cdot\tilde{H}\right|^2 \rightarrow \frac{\langle h\rangle^2}{\Lambda}\bar{\nu}_L\nu_L + \cdots$$

THIS IS SOME HEAVY SCALE THAT GENERATES THE WEINBERG OPERATOR
WEINBERG, PHYS. REV. LETT. 43, 1566 (1979)

$$\mathcal{L}_5 = \frac{C_5^{\ell\ell'}}{\Lambda}\left[\Phi\cdot\overline{L_\ell^c}\right]\left[L_{\ell'}\cdot\Phi\right]$$

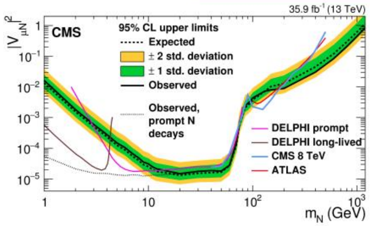
$$\sqrt{2}\Phi \approx v + h$$

$$\mathcal{L}_5 = -\frac{C_5^{\ell\ell'}}{2\Lambda}hh\overline{\nu_\ell^c}\nu_{\ell'} - \frac{C_5^{\ell\ell'}v}{\Lambda}h\overline{\nu_\ell^c}\nu_{\ell'} - \frac{C_5^{\ell\ell'}v^2}{2\Lambda}\overline{\nu_\ell^c}\nu_{\ell'} + \text{H.c.}$$

Experimental perspective of Seesaw Type-I

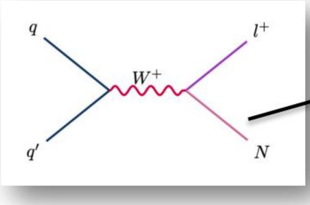
➔ Type-I Seesaw model & Heavy neutrinos has been widely probed in different collider facilities

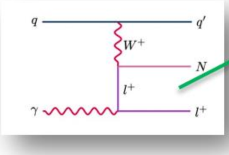
❖ In CMS there are joint efforts in the past ten years. ([Sihyun @ EXO Workshop](#))



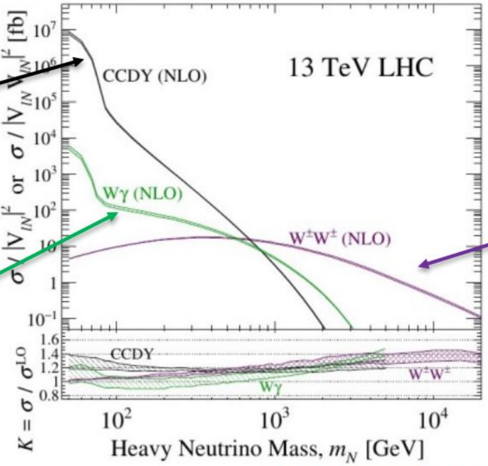
World best limit from CMS 2016 analysis!

Upper: EXO-17-012: Trilepton
down: EXO-17-028: Same-Sign Lepton + 2 jets



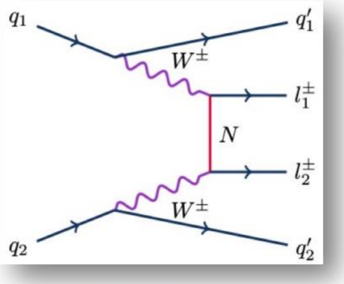


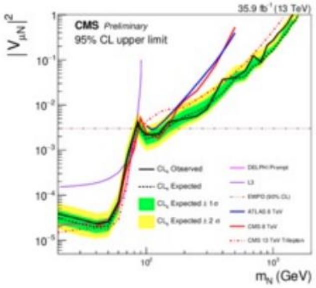
[Phys. Rev. D 103, 055005](#)



Newly proposed approach:

The t-channel production have more sensitivity at higher mass phase space





Traditional heavy neutrino hunting strategy: S-channel production

Pros: Larger cross section in the mass window, high sensitivity
Cons: Cross section drops sharply when **away** from mass window, low sensitivity

Motivation 1:

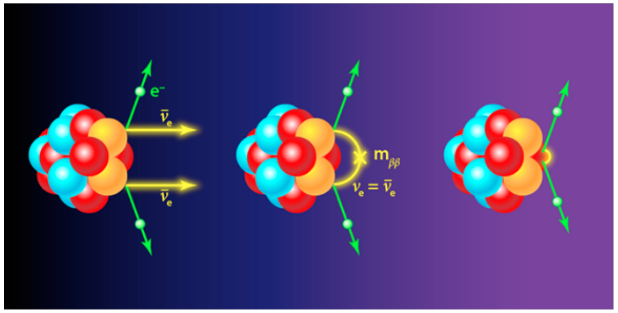
Make use of the VBF t-channel production of heavy neutrinos to enhance the sensitivity at higher mass region.

Majorana Neutrinos & Weinberg Op. Probe

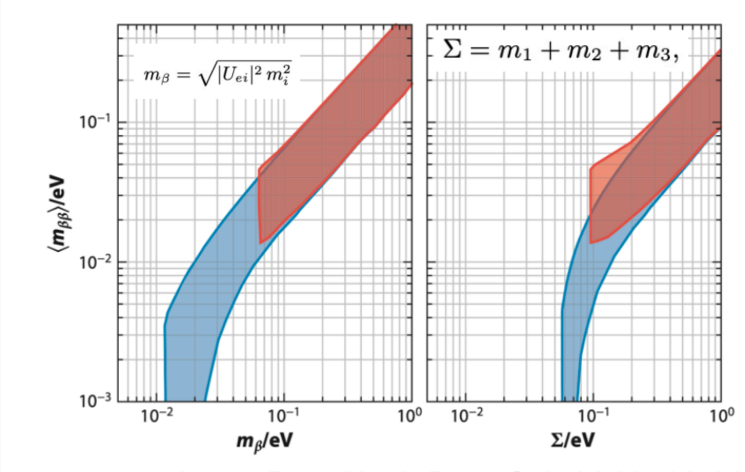
EXO-21-003

Experimental perspective of Weinberg Operator

- ➔ Weinberg Operator (WO) has been applied as the theoretical background of neutrino experiments in the context of nuclear physics.
  - ❖ Weinberg operator offers Majorana mass to SM neutrinos, thus is suitable for neutrino less double beta decay ($0\nu\beta\beta$) experiments
    - ‣ Stringent limits are obtained for electron channel



Red: inverted hierarchy
Blue: normal hierarchy



Annu. Rev. Nucl. Part. Sci. 2019. 69:219–51

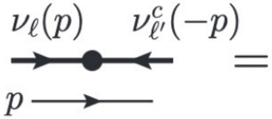
- ❖ However, bounded by typical energy scale, nuclear experiments cannot provide hints for 2 or 3 generations. Must turn to collider experiment for help!

Majorana Neutrinos & Weinberg Op. Probe

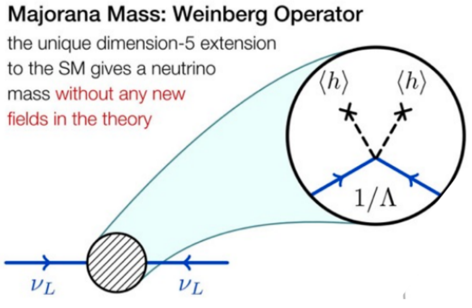
EXO-21-003

Experimental perspective of Weinberg Operator

- ➔ Here lies the abyss: Monte Carlo issue for WO simulation
  - ❖ Head-on neutrino lines appears and makes it hard for Monte Carlo generation;
  - ❖ Newly proposed solution: Phys. Rev. D 103, 115014
    - ‣ Approximate head-on lines with a Majorana fermion line



$$= \frac{i\not{p}}{p^2}\,\frac{-iC_5^{\ell\ell'}v^2}{\Lambda}\,\frac{i\not{p}}{p^2} = \frac{im_{\ell\ell'}}{p^2}$$


Majorana Mass: Weinberg Operator

the unique dimension-5 extension to the SM gives a neutrino mass without any new fields in the theory

- ➔ Thus, collider search for WO becomes achievable
  - ❖ Typical process: VBF same-sign muon production

Wilson Coefficients

SM Higgs Doublet

$$\mathcal{L}_5 = \frac{C_5^{\ell\ell'}}{\Lambda}[\Phi\cdot\overline{L}^c_\ell][L_{\ell'}\cdot\Phi] + \text{H.c.}$$

EFT Scale

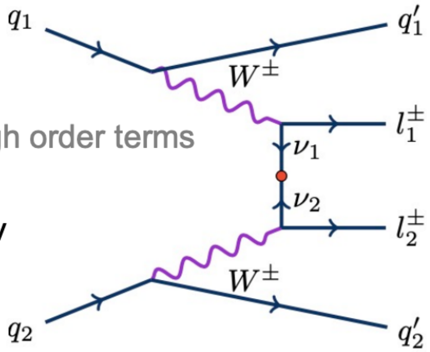
SM Lepton

High order terms

$v = \sqrt{2}\langle\Phi\rangle \approx 246\ \text{GeV}$ Higgs vev

$m_{\ell\ell'} = C_5^{\ell\ell'}v^2/\Lambda$

Effective Majorana Mass



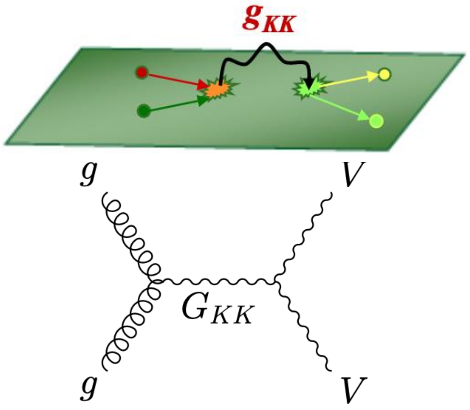
Motivation 2:

Probing Weinberg Operator with VBF same-sign di-muon production
Put limits to EFT Wilson Coefficient and thus effective neutrino mass

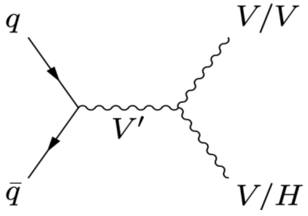
Motivation for a Diboson search

- SM shortcomings indicate some kind of New Physics (Hierarchy, Unific. DM, DE)

1. (Bulk RS) Warped ED,
   spin-0 Radion $(\mathrm{krc}\pi = 35,\ \Lambda_R = 3\ \mathrm{TeV})$
   spin-2 Bulk Graviton $(\sim k = 0.5, 1.0, \dots)$

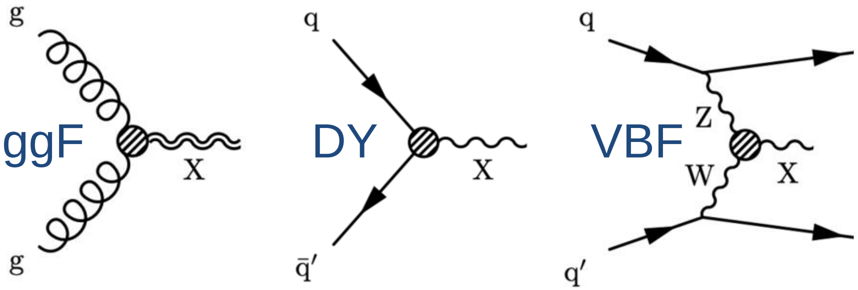


2. Heavy vector triplet (HVT)
spin-1 Z’,W’, coupling with SM
→ Models A, B, C

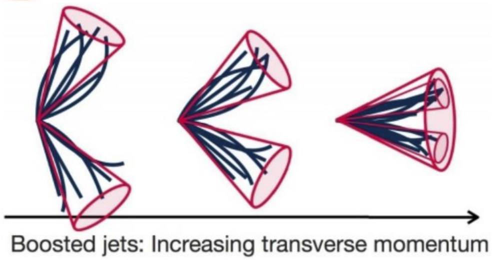


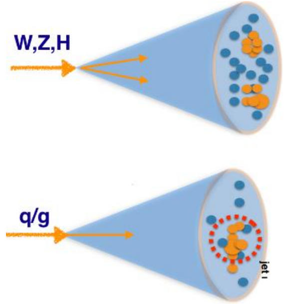
3. Little Higgs models
4. Two Higgs doublets models (MSSM)
5. Extended WED models (V_{KK}→RV)
6. Technicolor models

- Predict heavy bosons at TeV
- 3 production modes:
- Decay modes include VV, VH



Boson tagging as boosted jets





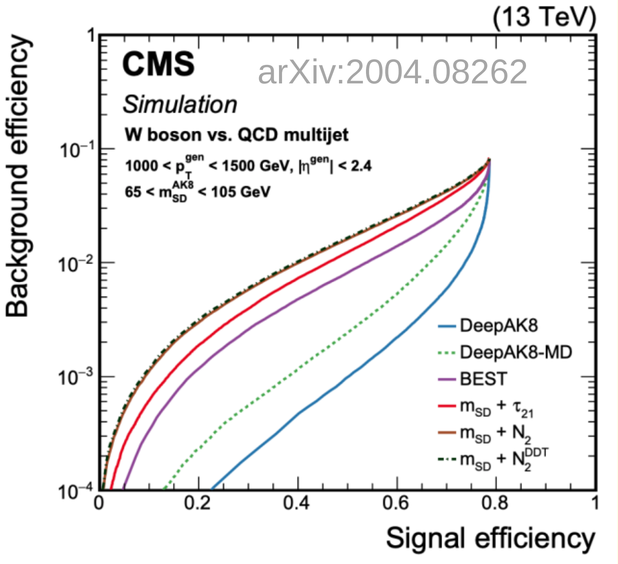
1. Use large-R (radius param.) jets
2. Clustering with anti-k_T;
3. soft-drop or trimmed jet mass
4. substructure, energy-flow, jet shape observables

Various substructure techniques

Most advanced available: deep-AK8

→ classification to 17 jet categories

→ flexibility forming modular taggers



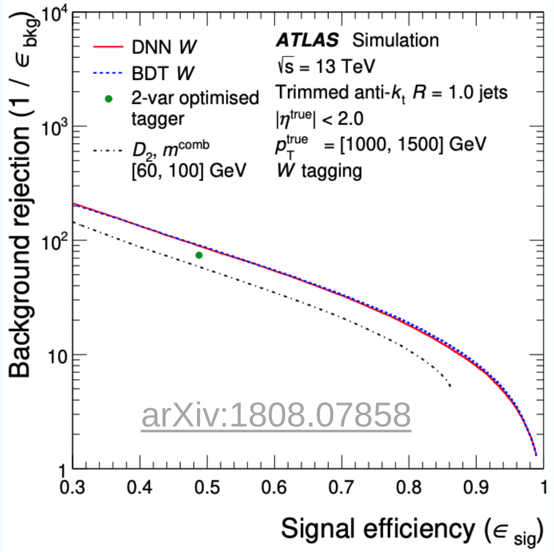
arXiv:2004.08262

N-subjettiness τ_{21} or τ_{21}^{DDT} alternative

Track-Calo Cluster (TCC) jets (utilize tracker & calo info)

→ tagging with BDT and DNN

→ subjets' handling with VR track-jets

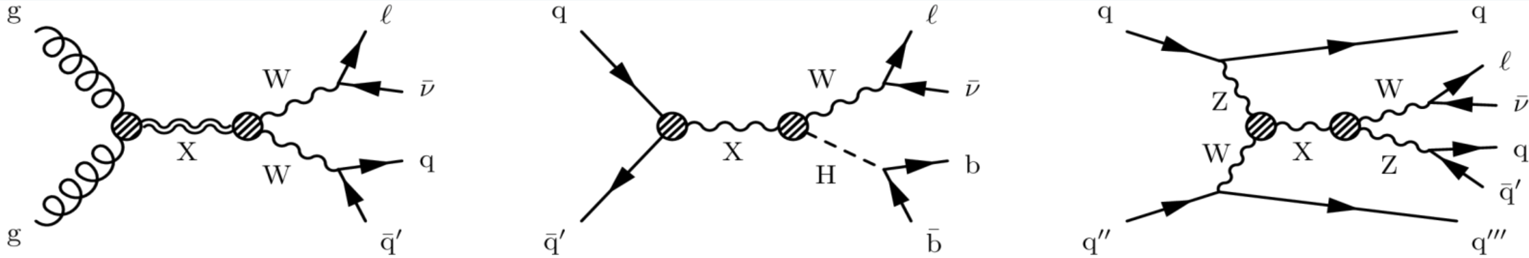


arXiv:1808.07858

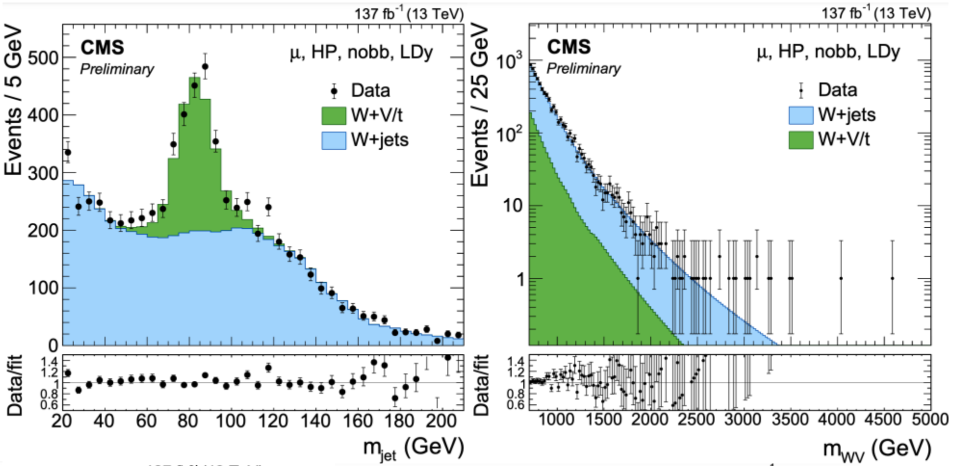
D_2 : 2-point E_{cf} / 3-point E_{cf}

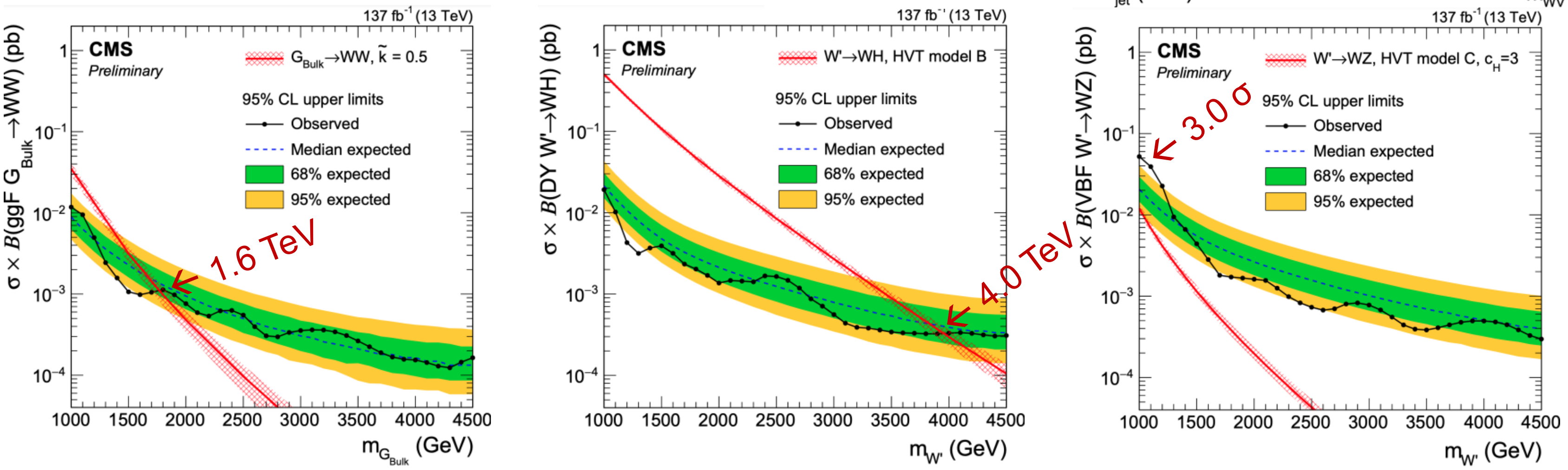
B2G-19-002

X → WV, WH → lν qq/bb



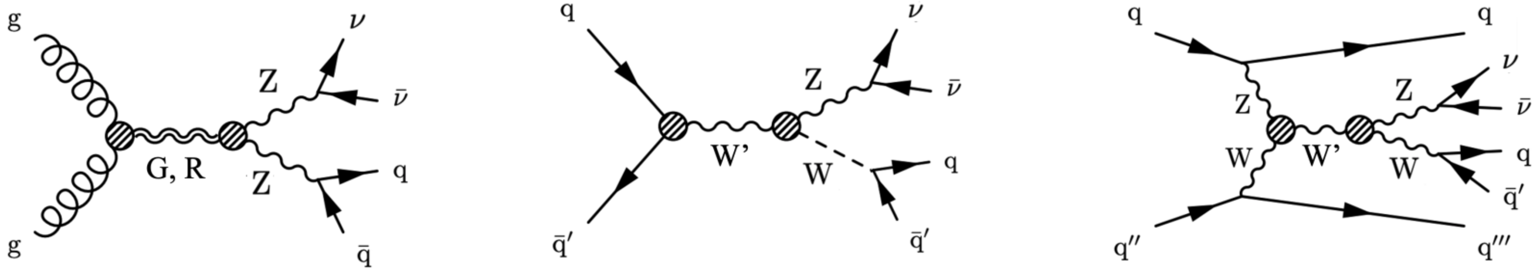
- 2D fit to the m_{Jet}, M_{WV} masses
- V/H-tagging: τ_{21}^{DDT}, double-b tagger
- $W_{l\nu}$, J, back-to-back
- 2 forward jets for VBF, 0 b-jets
- 24 categories based on 4 criteria:
  e/μ, L/H purity, VBF/bb/nobb, L/H $\Delta y_{J,l\nu}$
- BKGs: non-reso (W+jets), reso (tt)
  Prediction with kernel-approach at M_{WV}



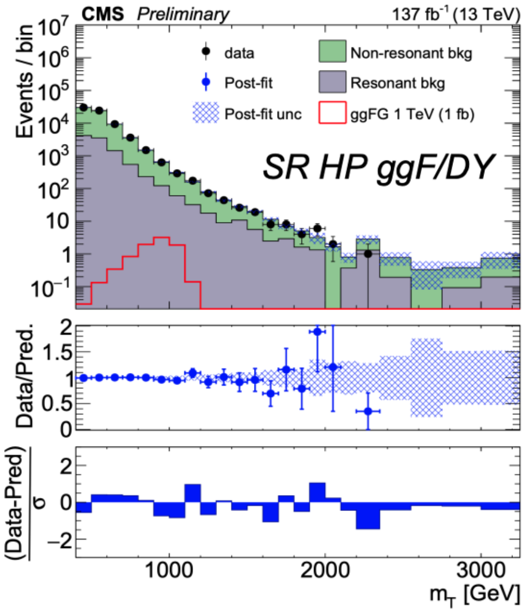


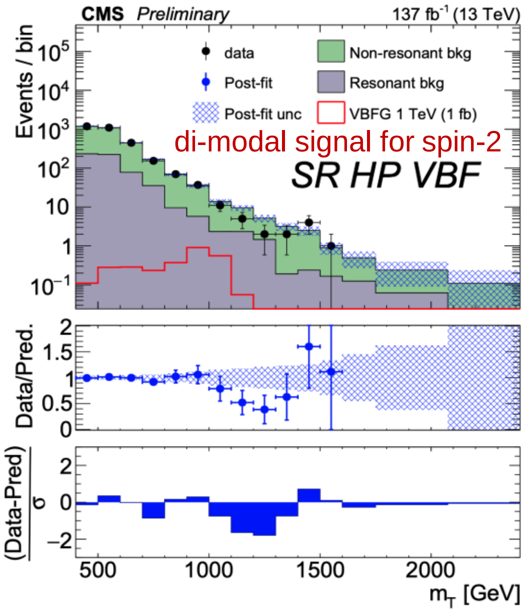
X → ZV → νν qq

B2G-20-008

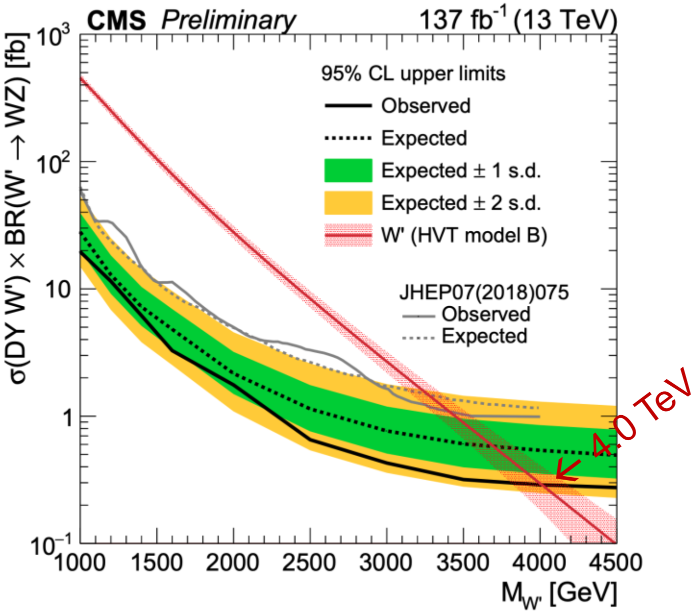


- Use $M_T(J,p_T^{miss})$ as observable; τ_{21} for V-tagging; veto b, l, τ_h, γ, p_T^{miss} ∥ j events
- Categorization to 4 sample: VBF, ggF/DY topology $|\Delta\eta_{jj}|<4$, $\eta_1\eta_2<0$, $m_{jj}>500$ GeV
  τ_{21} High/Low purity $\tau_{21}<0.35$, $0.35<\tau_{21}<0.75$
- SR: $65<m_J<105$ GeV; CR: m_J sideband (m_J: 30-65, 135-300 GeV)
- Dominant BKG: W/Z+jets, estimated from the data in CR per M_T bin





di-modal signal for spin-2

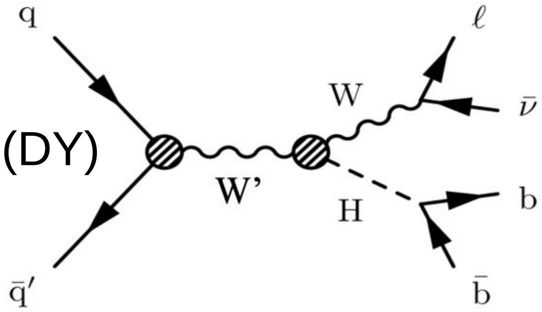


4.0 TeV

CONF-2021-026

W’ → WH → lν bb

qqA (DY)



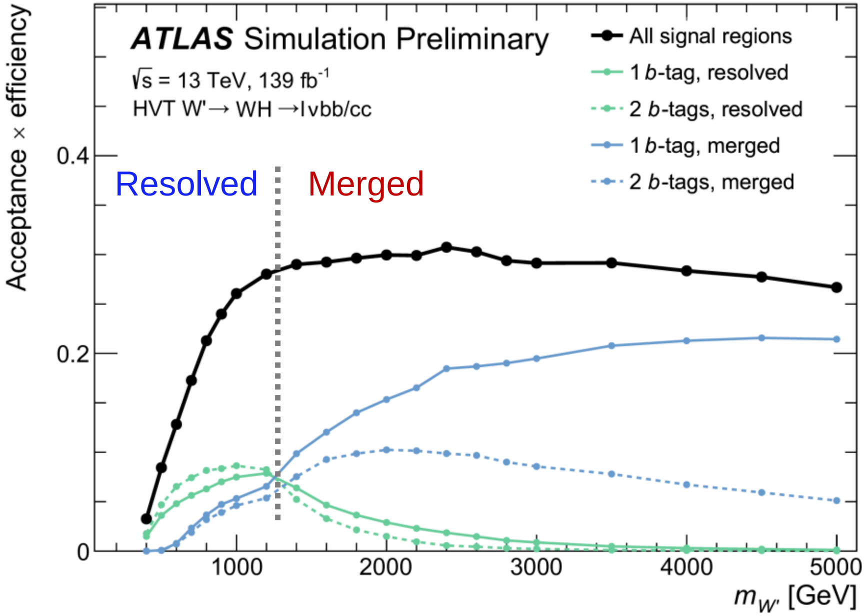
- 1-lep: e/μ, E_T^{miss}
  - W→lν reconstructed by M_W=80 GeV constraint
- Higgs reconstructed:
  2 small-R jets or 1 large-R jet
- Observable: M_{WH}
- BKG rejection with various mass dependent kinematic selections on: E_T^{miss}, M_T, $p_{T,W}$ and more
- BKG:
  (flavor decomposition)

🟧 Top
🟥 W+hf
🟥 W+hl, W+l

4 categories:

- Resolved H: 2 small-R jets (R=0.4)
  1 or 2 b-tagged jets (MV2c10)
  m_{jj}: 110-140 GeV → SR; sidebands → CR
- Merged H: 1 large-R TCC jet (R=1.0)
  1 or 2 b-tagged VR track-jets
  m_J: 75-145 GeV → SR; sidebands → CR
- 4 SRs & 4 CRs in total

Signal acc.×eff. per SR for 1-lep.

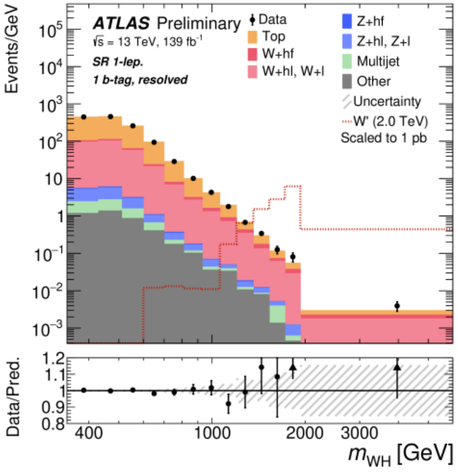


CONF-2021-026

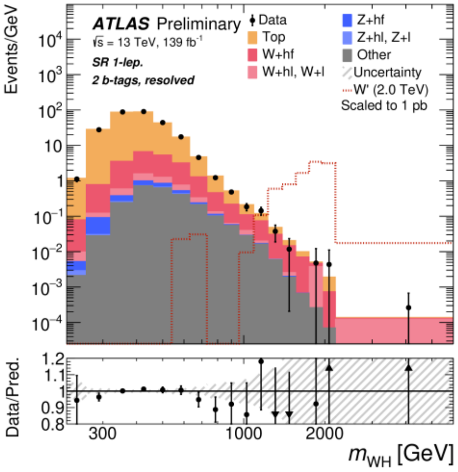
W' → WH → lν bb; Results

Resolved

1b

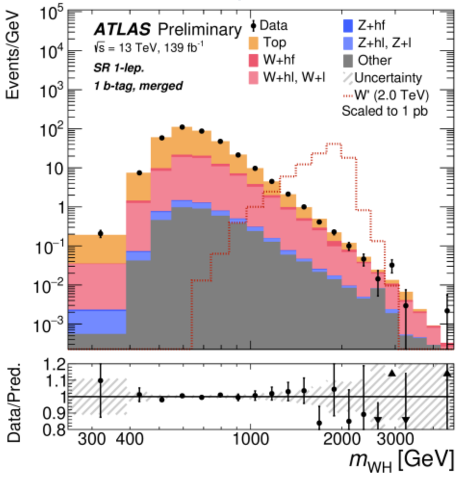


2b

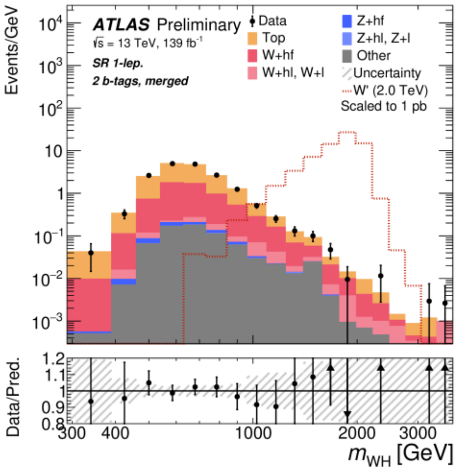


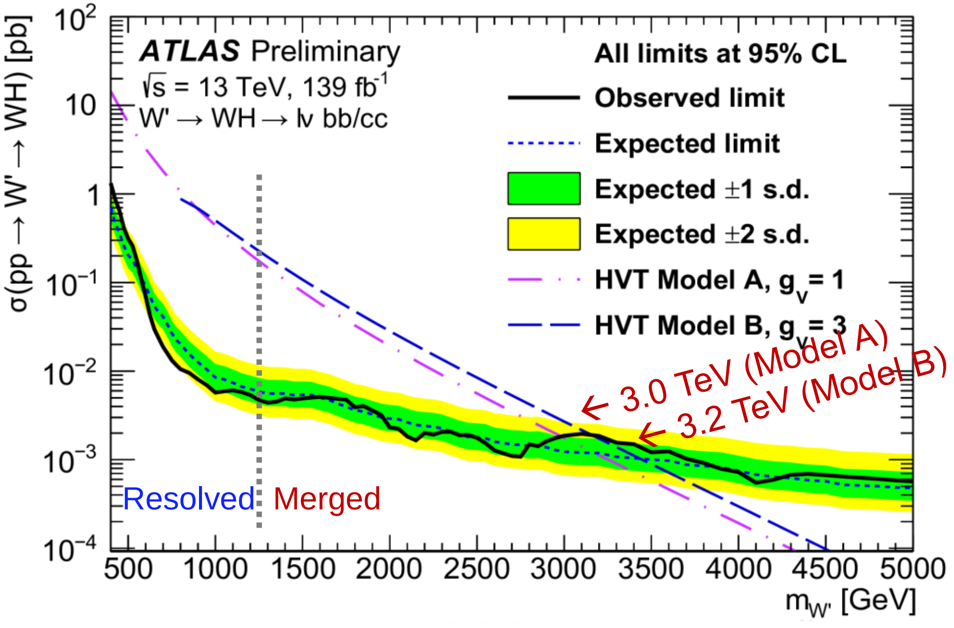
Merged

1b



2b



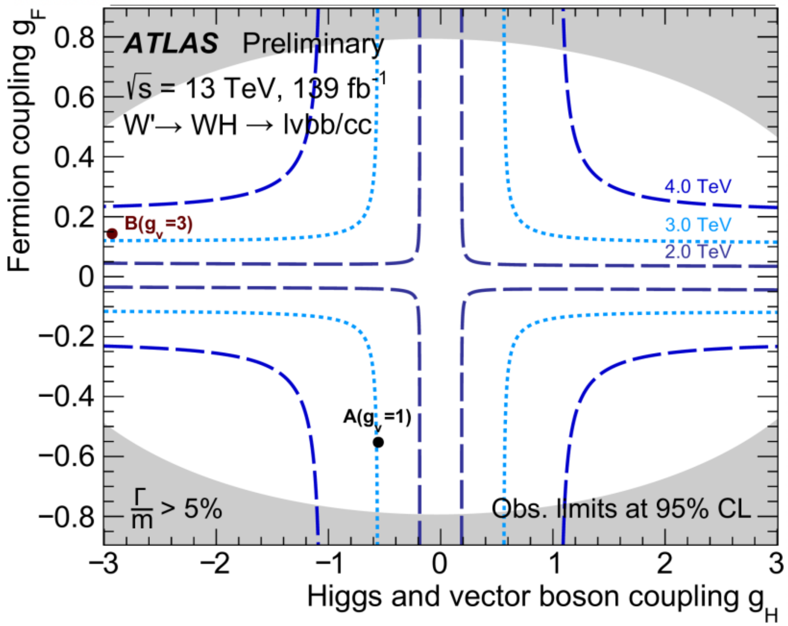


← 3.0 TeV (Model A)

← 3.2 TeV (Model B)

Resolved

Merged

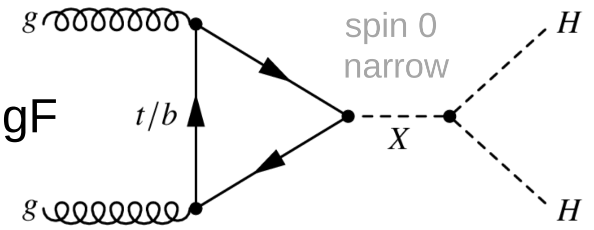


X → HH → bb γγ (Resolved)

CONF-2021-016

MSSM or RS ED
spin 0
narrow

ggF

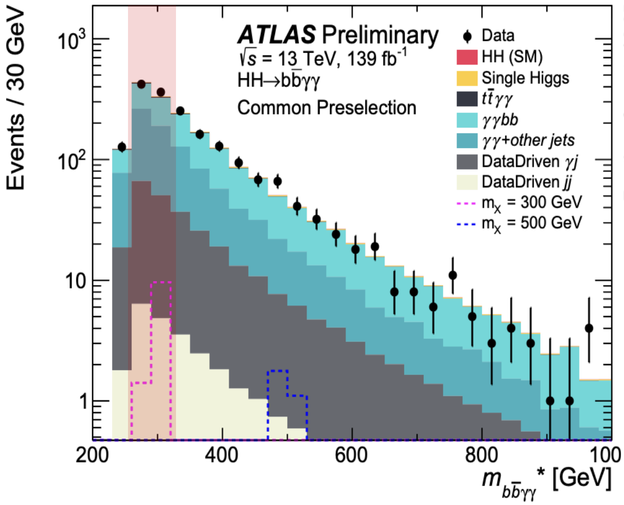


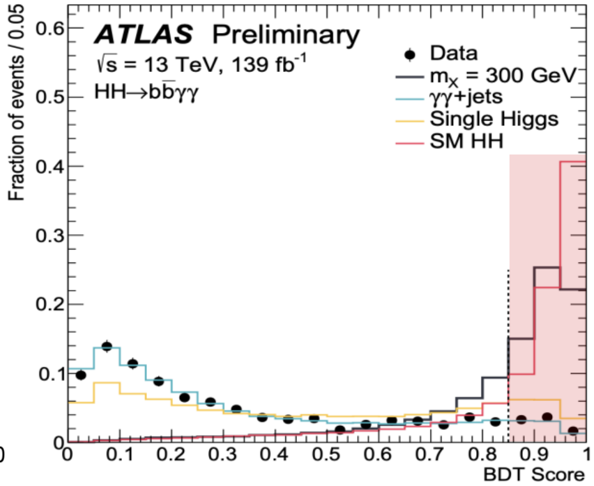
- 2 small-R b-jets (R=0.4)
  (tagger: DL1r, eff.~77%)
- 2γ, $m_{\gamma\gamma}$: 120-130 GeV SR
  sideband → CR for fit
- 4-body mass:

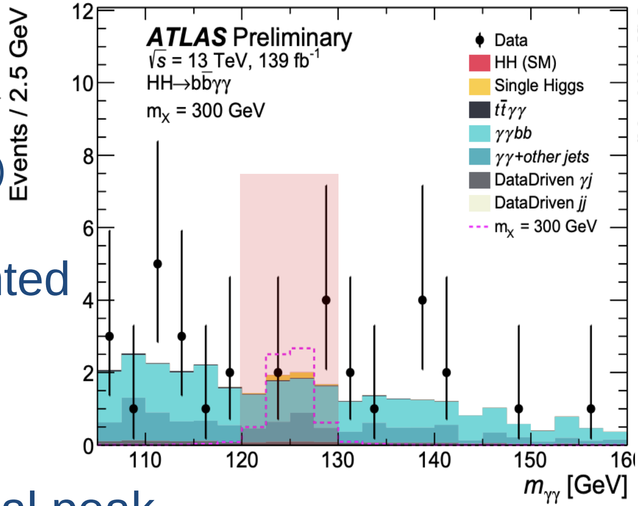
$$m^*_{b\bar{b}\gamma\gamma} = m_{b\bar{b}\gamma\gamma} - m_{b\bar{b}} - m_{\gamma\gamma} + 250\ \text{GeV}$$

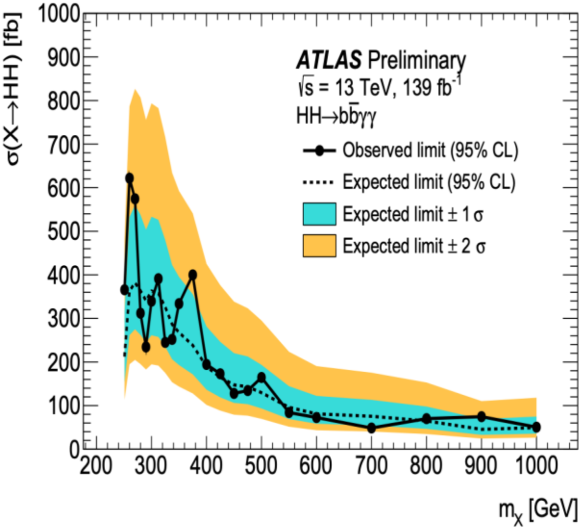
(Canceling detector resolution effects)

- 2 BDTs are combined weighted
  (for γγ+jets, single-H BKGs)
- Search by:
  → slicing $m^*_{bb\gamma\gamma}$ around signal peak
  → fitting analytic function at $m_{\gamma\gamma}$







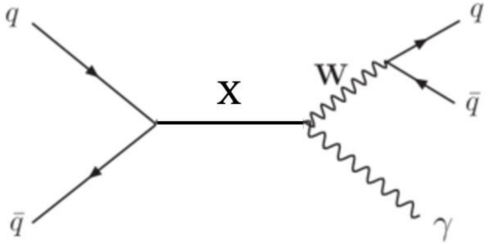


Also, non-resonant signal (κ_λ)
is simultaneously probed

X → Wγ → qq γ

EXO-20-001

Generic V_{qq}+γ search



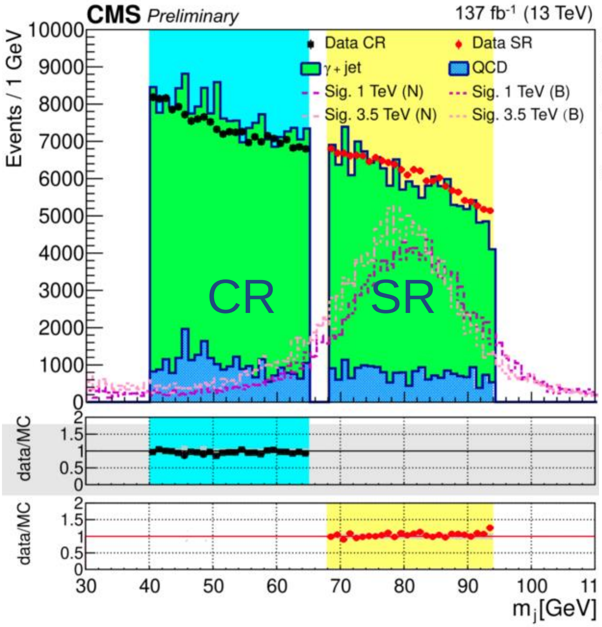
- W→qq merged (R=0.8) jet
- W-tagging with τ_{21}, m_J^{SD}
- Central γ

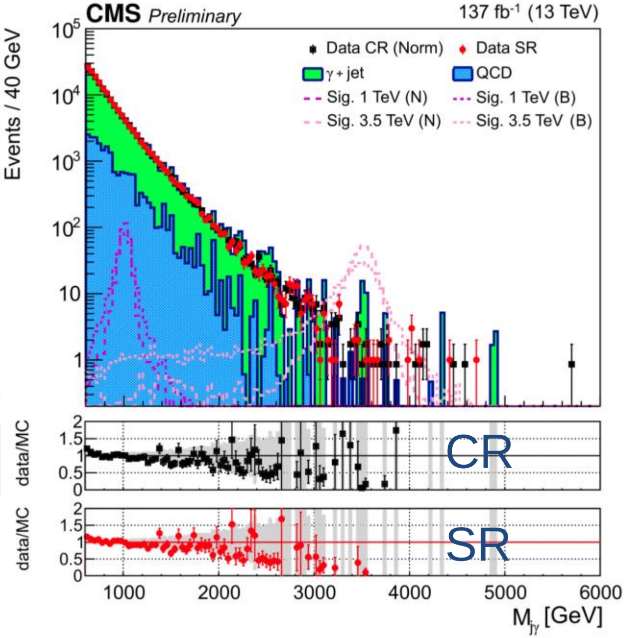
- Main BKG: γ+jets
- Low m_J^{SD} as CR

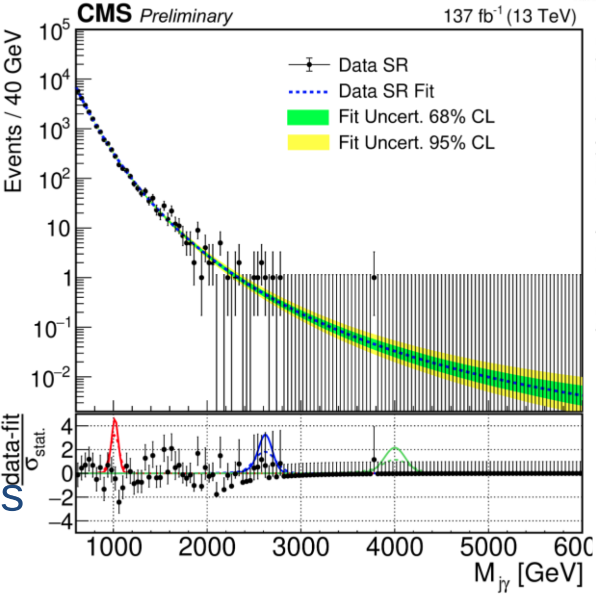
- BKG estimate: fitting analytic function to $M_{J\gamma}$

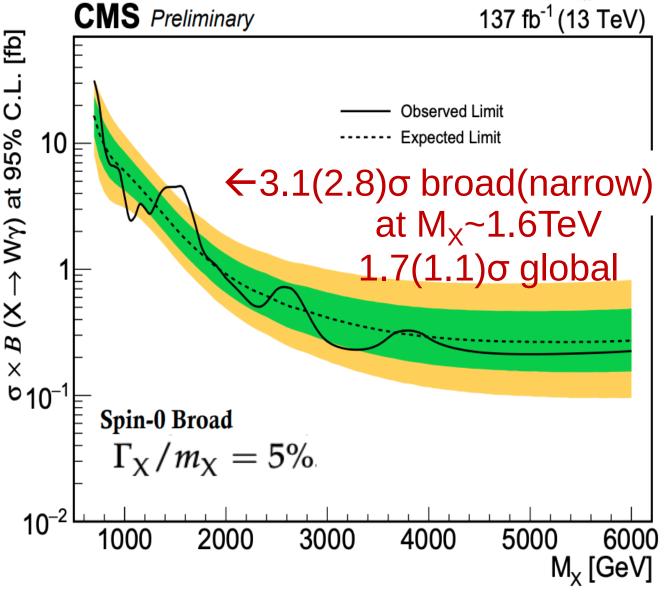
- Best limits to date on: $\sigma_{pp\to X} \times \mathrm{Br}(X \to Wqq\gamma)$

- Model (in)dependent limits spin 0&1, narrow/broad







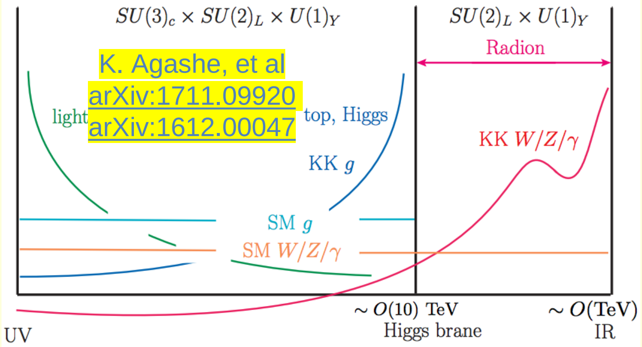


←3.1(2.8)σ broad(narrow) at M_X~1.6TeV
1.7(1.1)σ global

X → RW → WWW → lν jets

B2G-20-001

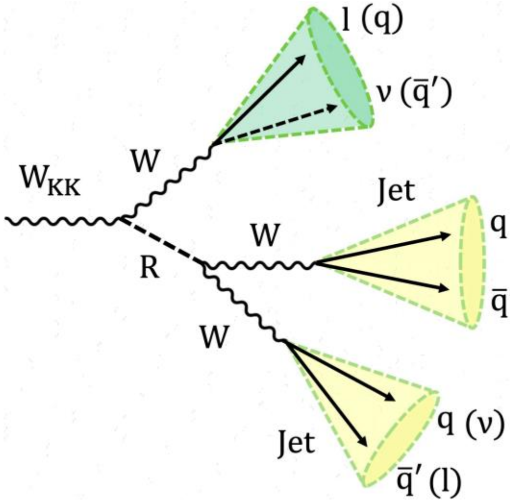
- First tri-boson search
- New model: Extended Warped ED
  → suppressed di-SM processes
  → enhanced tri-SM processes



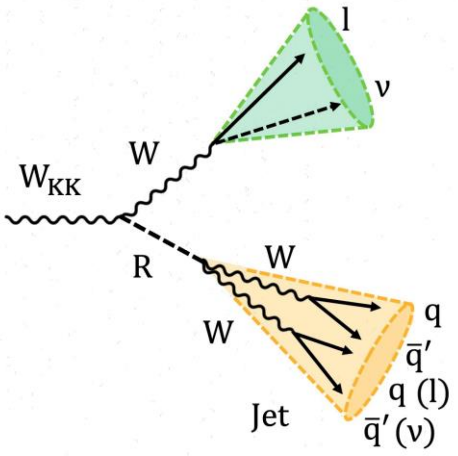
K. Agashe, et al
arXiv:1711.09920
arXiv:1612.00047

- Only EW in extended bulk dominant: $V_{KK} \to R\,V \to VVV$
  Di-resonant
- W→lν: reconstruction
- 1 or 2 AK8 massive jets, 0 b-jet
- deep-AK8 taggers for W & R
- Radion tagging with H_{4q} & W_{qq}
- Calibration with SM-proxy jets: top for $R^{3q,4q}$, W for R^{lqq}

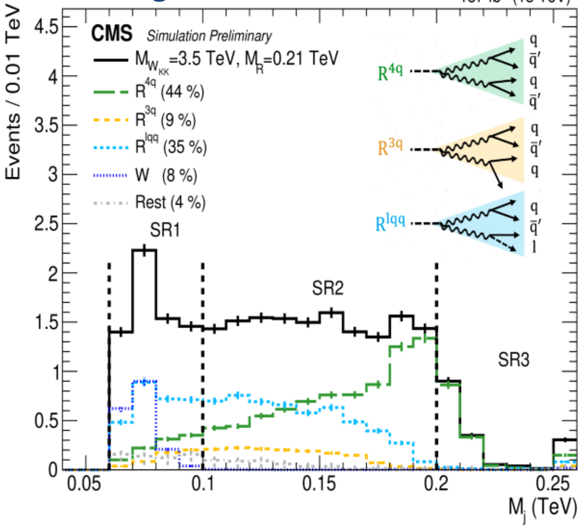
Resolved Radion



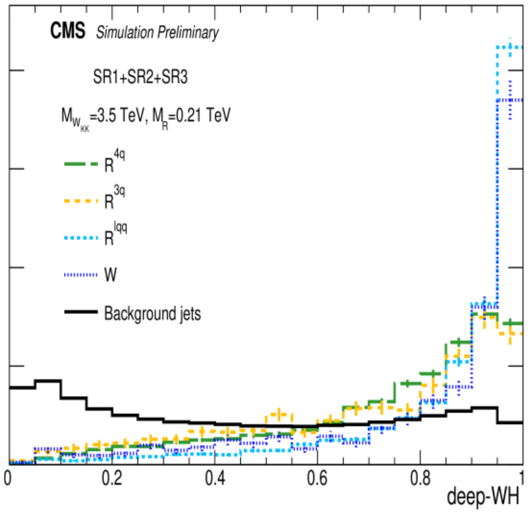
Merged Radion



Merged Radion mass



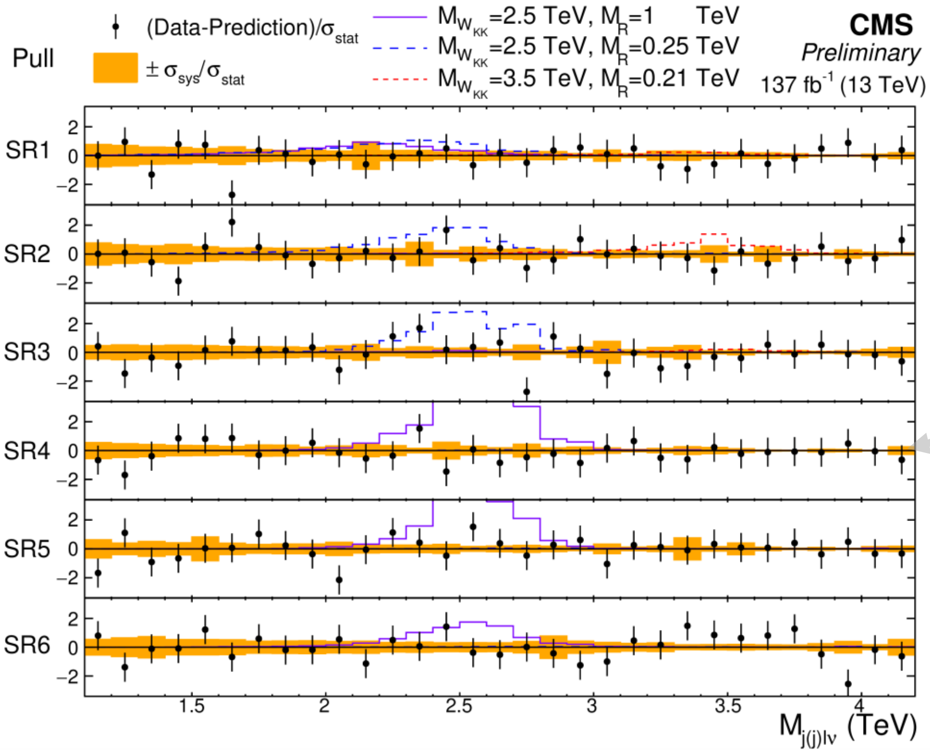
Merged Radion tagger



Triboson results

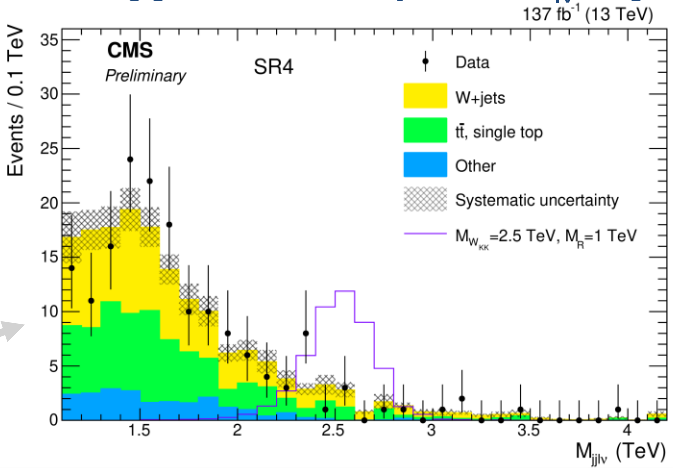
- Probe simultaneously merged & resolved

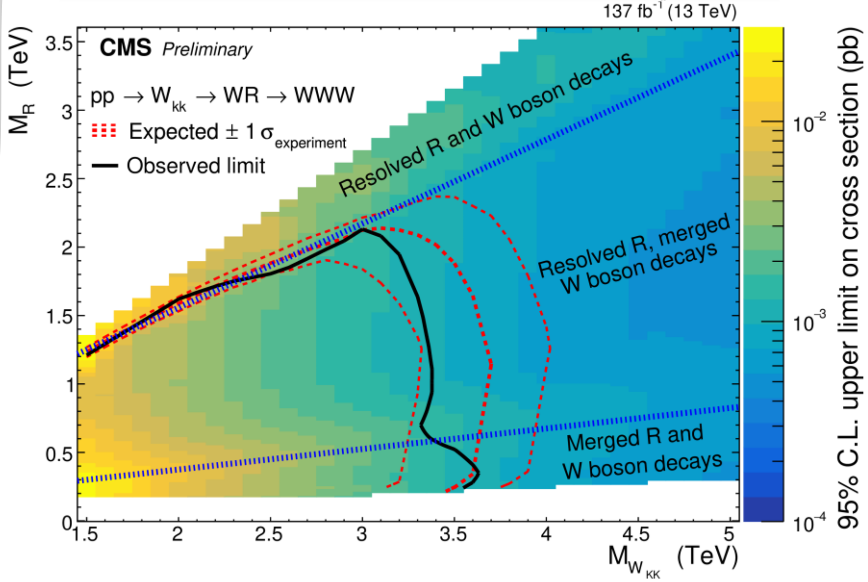
- Categorize to 6 SRs:
  SR1-3 → 1 jets (merged) → M_{lvj}
  SR4-6 → 2 jets (resolved) → M_{lvjj}



- First limits on $\sigma(W_{KK} \to RW \to WWW)$ and on $[M_{WKK}, M_R]$ space

2 W-tagged massive jets + W_{lv} region:

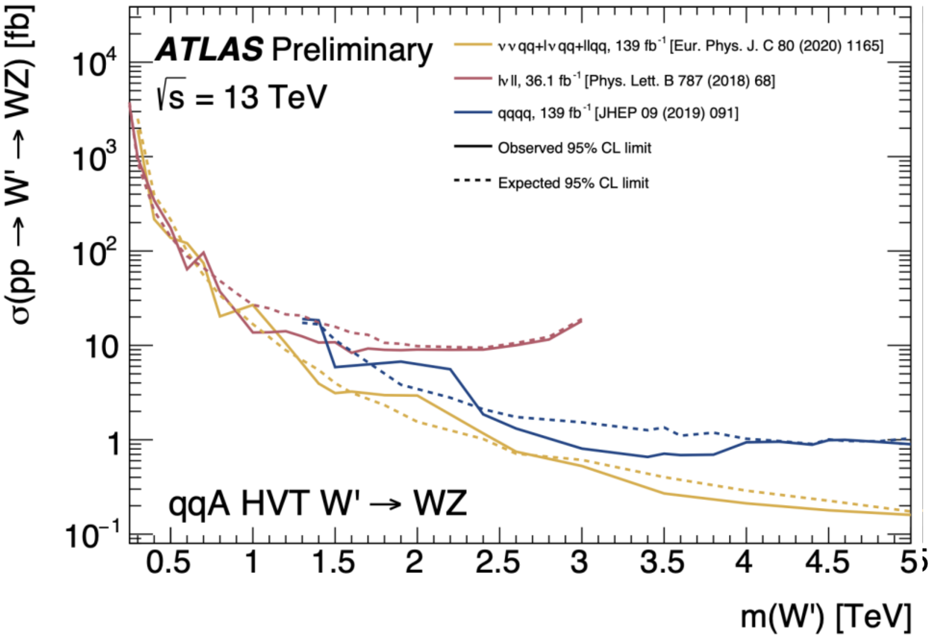


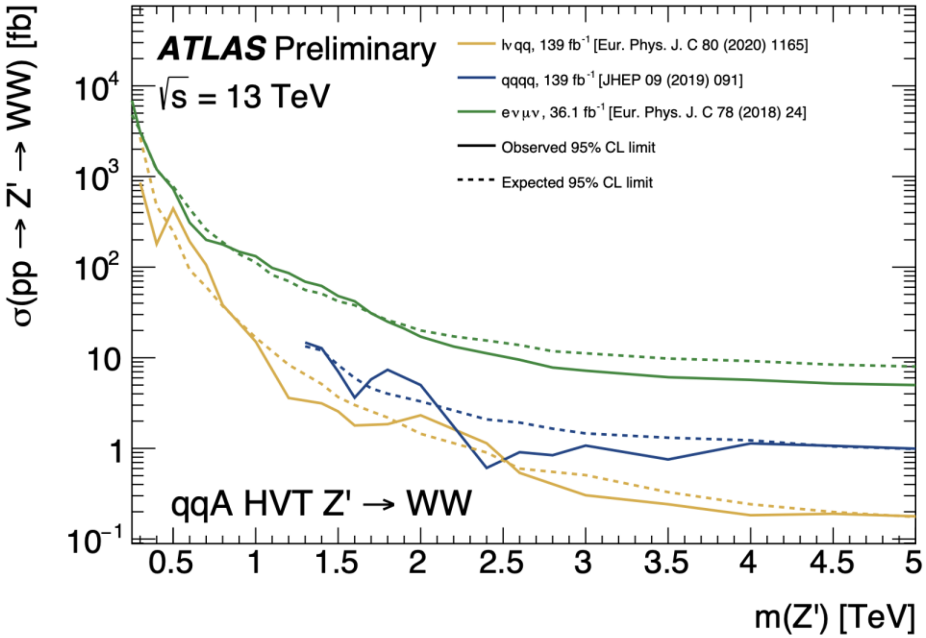


Results interpretation in $Z_{KK} \to ZWW$
Relevant search $g_{KK} \to Rg \to ggg$ at talk

Summary Results & Conclusions

Summary of results (11 ATLAS searches) upper limits on σ×BR (limits superimposed)

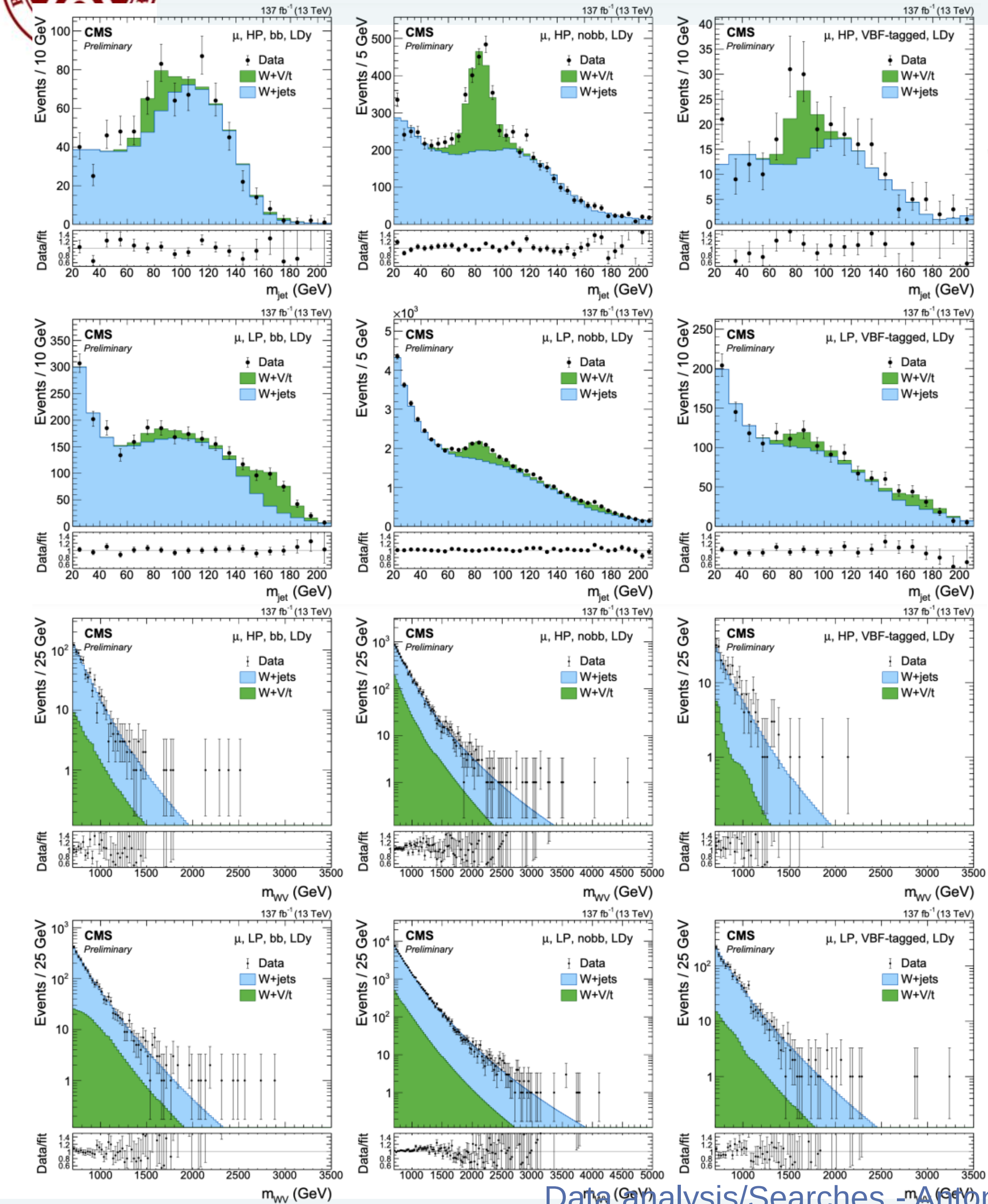




- DNN techniques are exploited to probe very rare events:
  - BKG suppression; Identify V/H in hadronic modes; b-tagging;
- Many analysis/channels are ongoing and more results will come soon
- The TeV-scale exploration is in the beginning (~5% of the LHC lumi. delivered)
- We have long way ahead, with potential surprises and a lot of fun – stay tuned!

Extras from: X → WH,WH → lν qq/bb

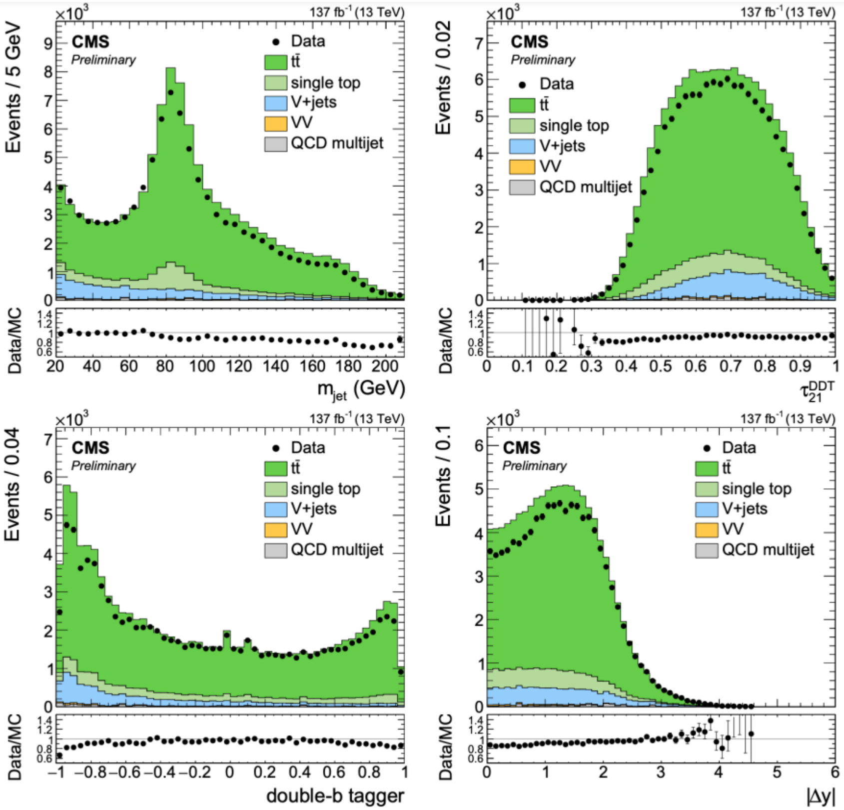
B2G-19-002



$$\tau_{21} = \tau_2/\tau_1$$

$$\tau_{21}^{\mathrm{DDT}} \equiv \tau_{21} - M\log(m_{\mathrm{jet}}^2/(p_{\mathrm{T}}\mu))$$

$$M = -0.08$$
$$\mu = 1\,\mathrm{GeV}$$



$$P_{\mathrm{sig}}(m_{\mathrm{WV}}, m_{\mathrm{jet}}|m_{\mathrm{X}}) = P(m_{\mathrm{WV}}|m_{\mathrm{X}}, \theta_1)\cdot P(m_{\mathrm{jet}}|m_{\mathrm{X}}, \theta_2)$$

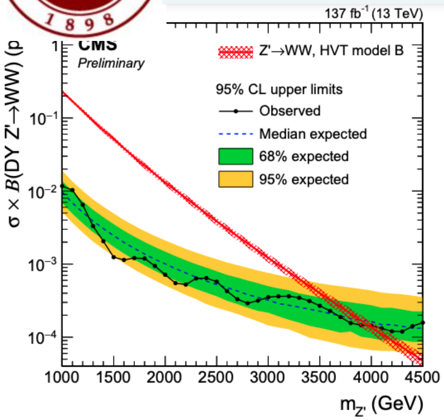
$$P_{\mathrm{W+jets}}(m_{\mathrm{WV}}, m_{\mathrm{jet}}) = P(m_{\mathrm{WV}}|m_{\mathrm{jet}}, \theta_1)\cdot P(m_{\mathrm{jet}}|\theta_2)$$

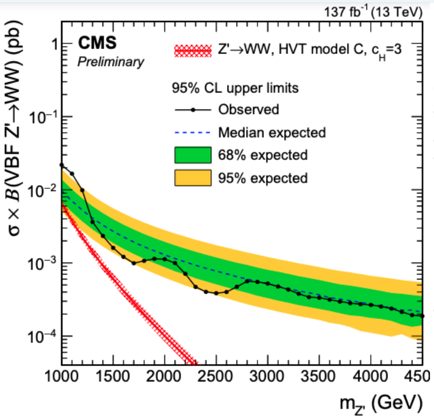
$$P_{\mathrm{W+V/t}}(m_{\mathrm{WV}}, m_{\mathrm{jet}}) = P(m_{\mathrm{WV}}|m_{\mathrm{jet}}, \theta_1)\cdot P(m_{\mathrm{jet}}|\theta_2)$$

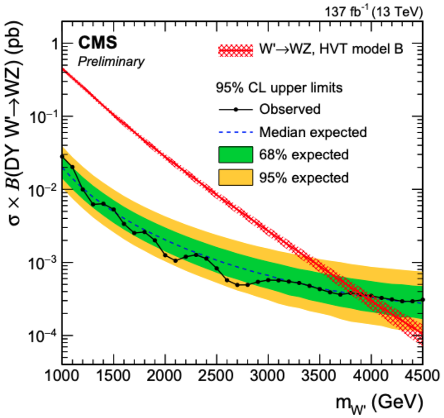
$$P_i(m) = \frac{w_i}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{m - s(p_{\mathrm{T,jet}}^{\mathrm{gen.}})\cdot m_{\mathrm{WV}}^{\mathrm{part.}}}{\sigma(p_{\mathrm{T,jet}}^{\mathrm{gen.}})}\right)^2\right]$$

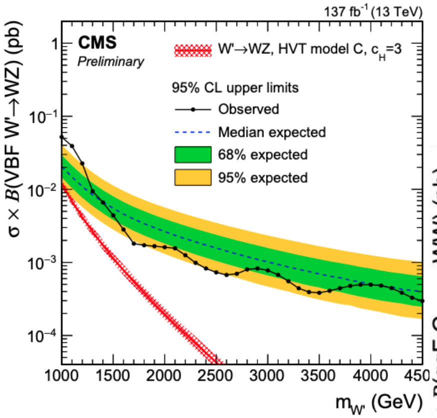
Extras from: X → WH,WH → lν qq/bb

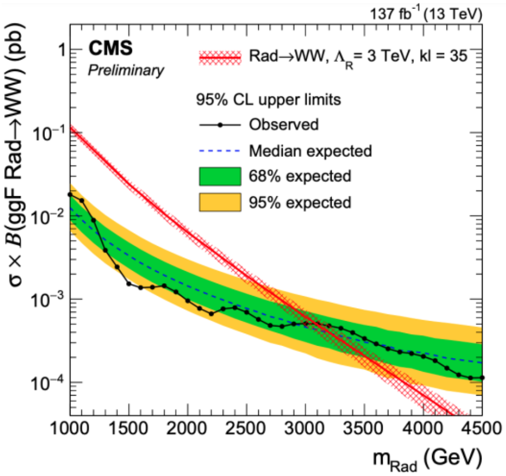
B2G-19-002

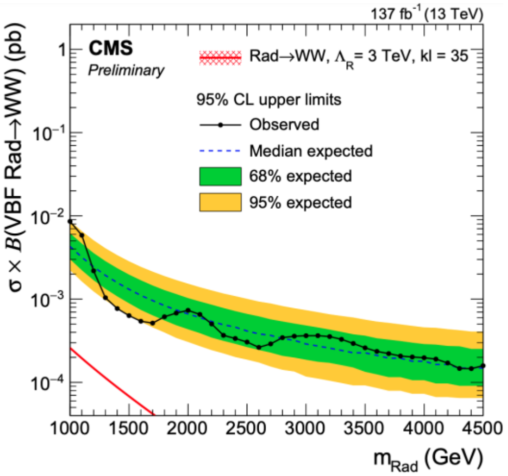


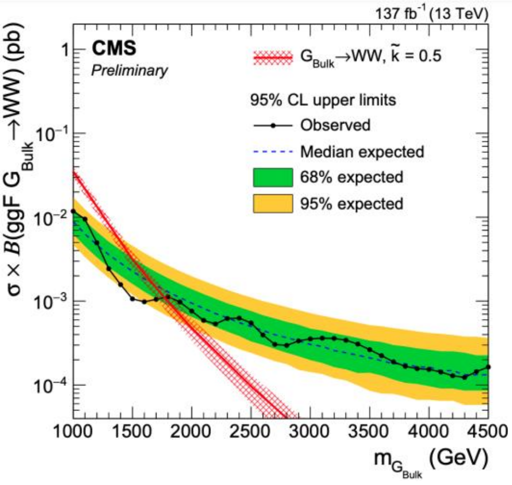


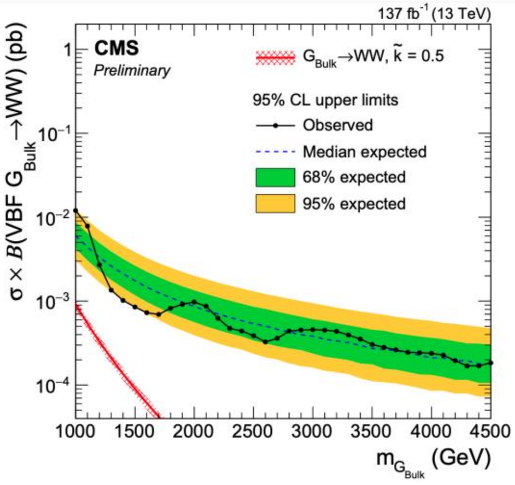


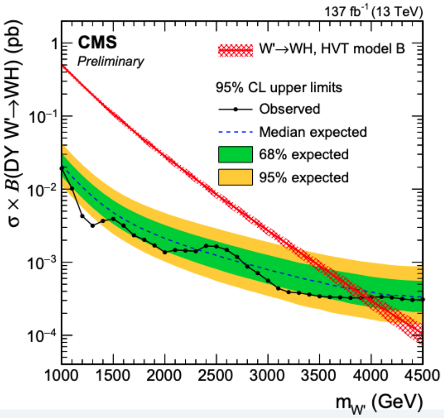






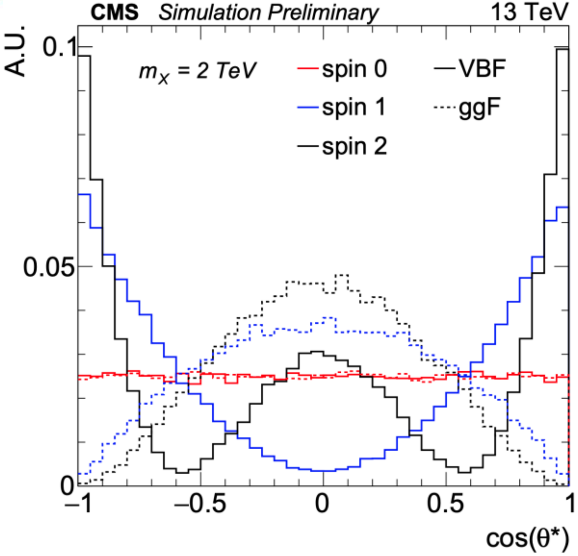


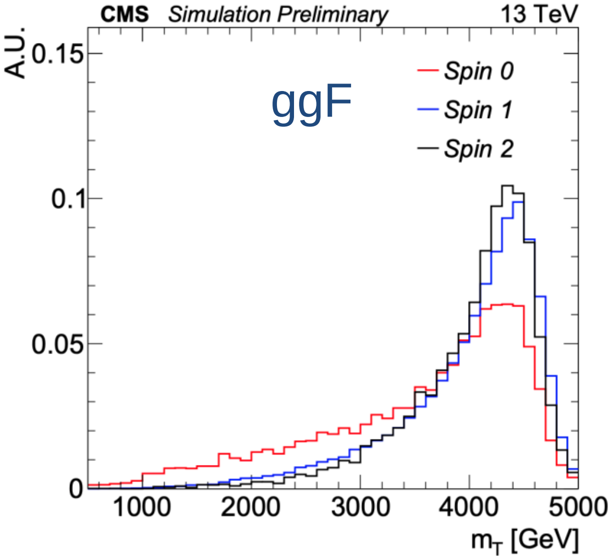


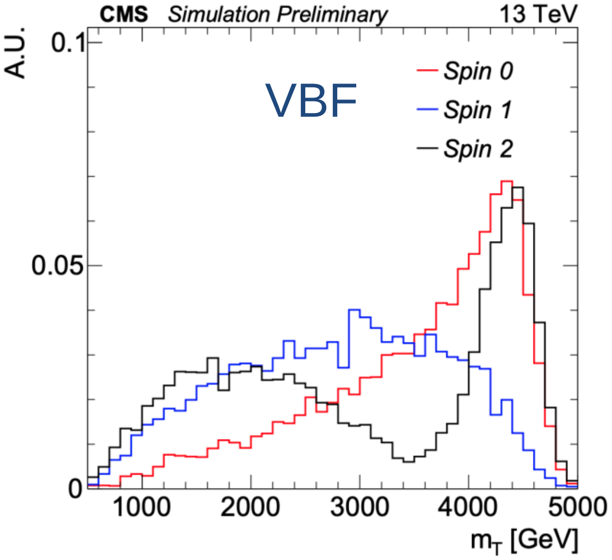


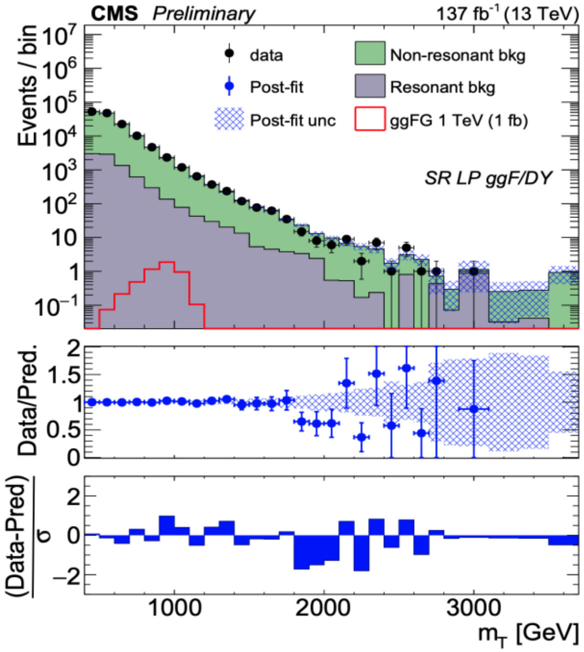
Extras from: X → ZV → νν qq

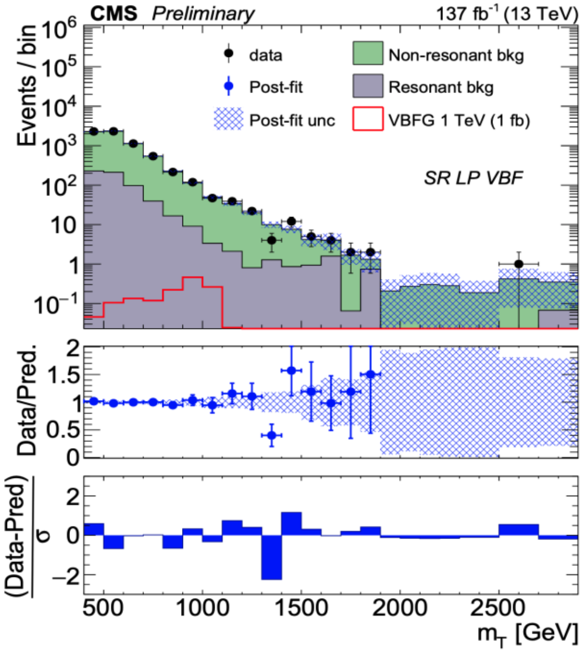
B2G-20-008











MET trigger eff:

200 GeV → 75%

250 GeV → 95%

plateau → 98%

$$m_{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}}^{\mathrm{J}} p_{\mathrm{T}}^{\mathrm{miss}} [1 - \cos\Delta\phi]}$$

$$N^{\mathrm{non-res}}_{\mathrm{pred}} = \alpha \times (N^{\mathrm{obs}}_{\mathrm{CR}} - N^{\mathrm{res}}_{\mathrm{CR}})$$

$$\alpha = \frac{N^{\mathrm{non-res}}_{\mathrm{SR}}}{N^{\mathrm{non-res}}_{\mathrm{CR}}}$$

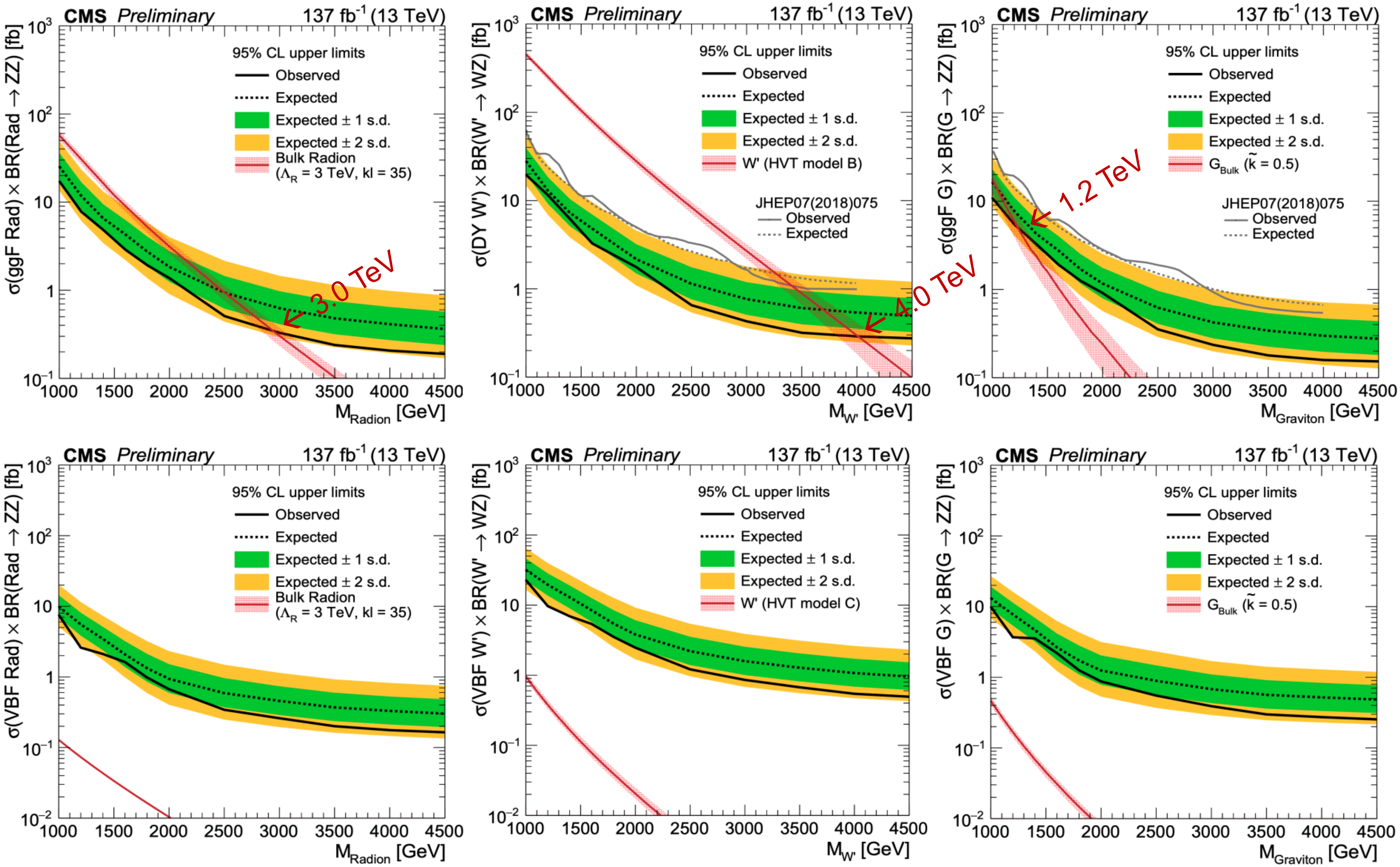
Extras from: X → ZV → νν qq

B2G-20-008

spin-0 spin-1 spin-2

ggF/DY

VBF



← 3.0 TeV

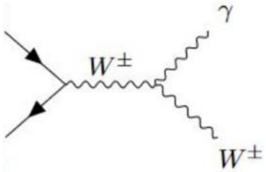
← 4.0 TeV

← 1.2 TeV

X → Wγ → qq γ

EXO-20-001

- Generic search for V_{qq}+γ



- W→qq AK8 jet
- tagging with $\tau_{21}<0.35$
- $p_{Tj(\gamma)}>225$ GeV, $|\eta_{j(\gamma)}|<2(1.44)$

$\Delta R_{J\gamma}>1.1$, $p_{T\gamma}/m_{J\gamma}>0.37$, $\cos\theta^*<0.6$

- Main BKG: γ+jets
- Calibration from low m_j CR

- BKG estimate: fitting analytic function to $M_{J\gamma}$

- Best limits to date on:

$$\sigma_{pp\to X} \times Br(X \to Wqq\gamma)$$

- Model independent limits for spin 0,1, narrow 0.01%, broad 5%

Theory motivation:

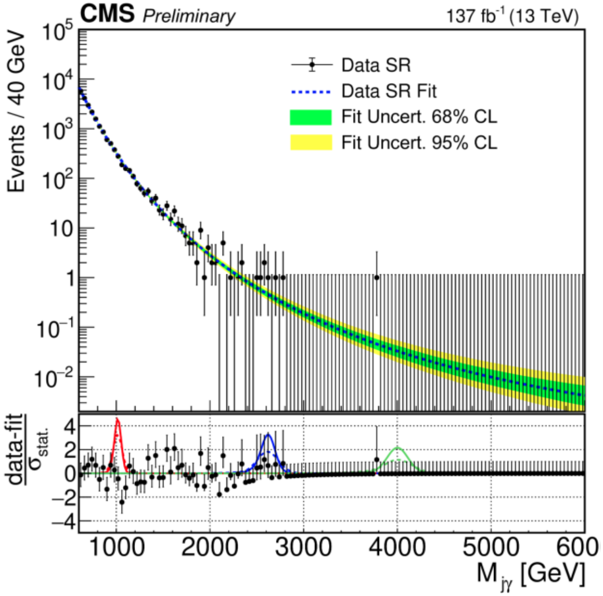
- Triplet pseudo-Goldstone bosons π_3

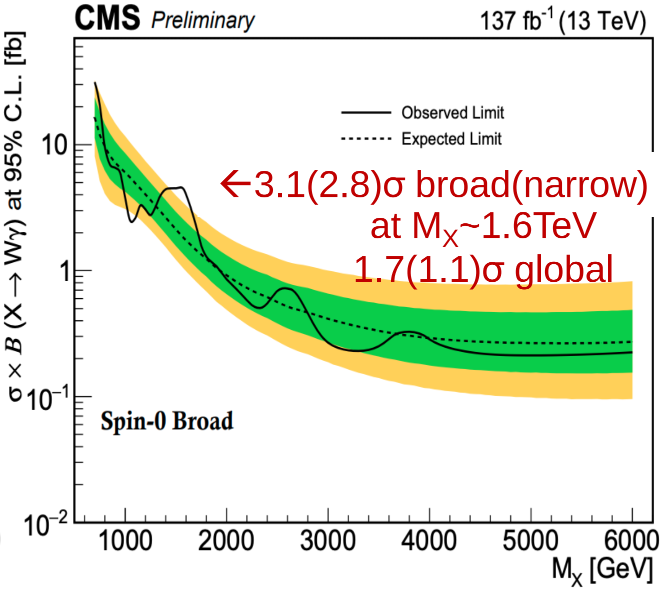
(https://arxiv.org/pdf/1608.01675.pdf)

- Scalar or pseudoscalar $SU(2)_L$ Φ^α coupling via anomaly-induced interaction

- Two Higgs doublet (H+) MSSM
- Technicolor
- HVT

JER: 15%, 8%, 4% for 10, 100, 1000 GeV

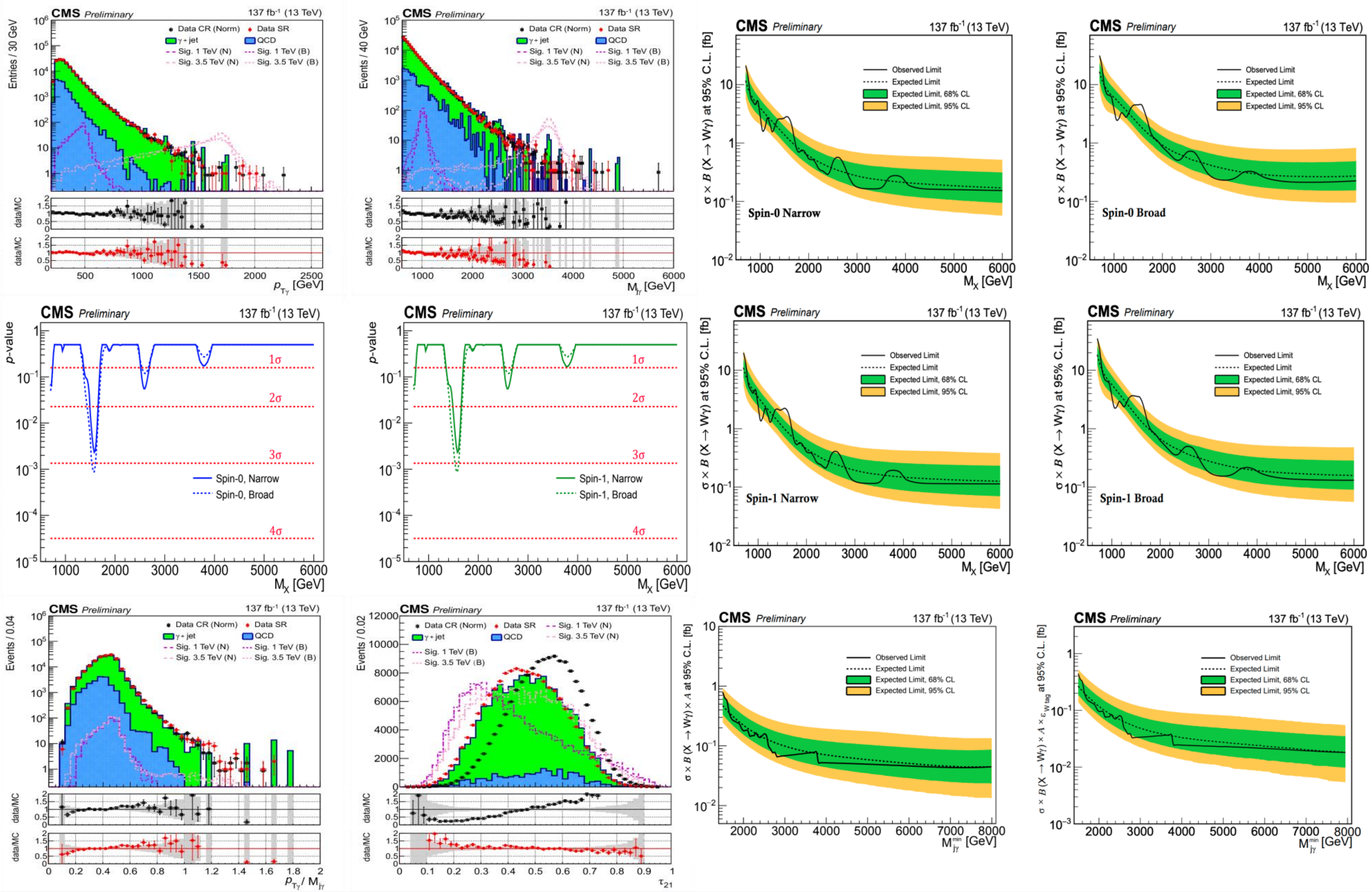




←3.1(2.8)σ broad(narrow) at M_X~1.6TeV
1.7(1.1)σ global

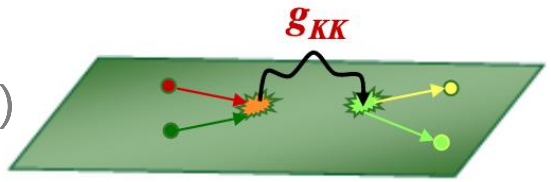
Extras from: X → Wγ → qq γ

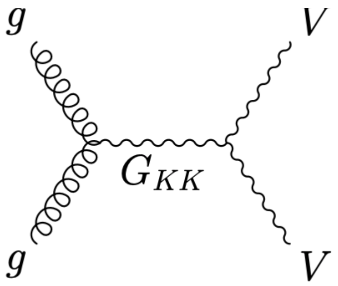
EXO-20-001



Motivation for a Diboson search

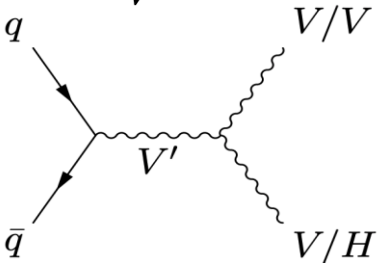
- SM shortcomings indicate some kind of New Physics
  (Hierarchy problem, Unific.of Gravity, Dark Matter/Energy)



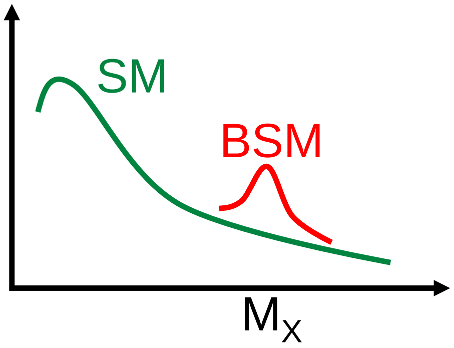


- Many BSM theories have been proposed:
  (Extended Gauge-Symmetry models; RS Warped ED; Two Higgs doublet models; Little/Composite Higgs)
  → Predicting new heavy boson(s) X with
  - spin 0: Radion/Higgs
  - spin 1: W'/Z' (HVT)
  - spin 2: Graviton

  at the TeV scale decaying to a pair of SM bosons→ W, Z, H, γ



- Therefore we can search for BSM Physics in Dibosons FSs



→ HOW TO... search?

Probing Diboson FS at TeV-scale is a challenge to reconstruct boosted & merged V/H reveling substructure

- Selection based on V-like objects suppressing BKG
- Predict in a Data-Driven way the SM BKG
- Look for a peak-structure at M_{VV} tails