Methods for off-shell Higgs boson production simulation used in CMS analyses

Mostafa Mahdavikhorrami,
on behalf of
the CMS Collaboration

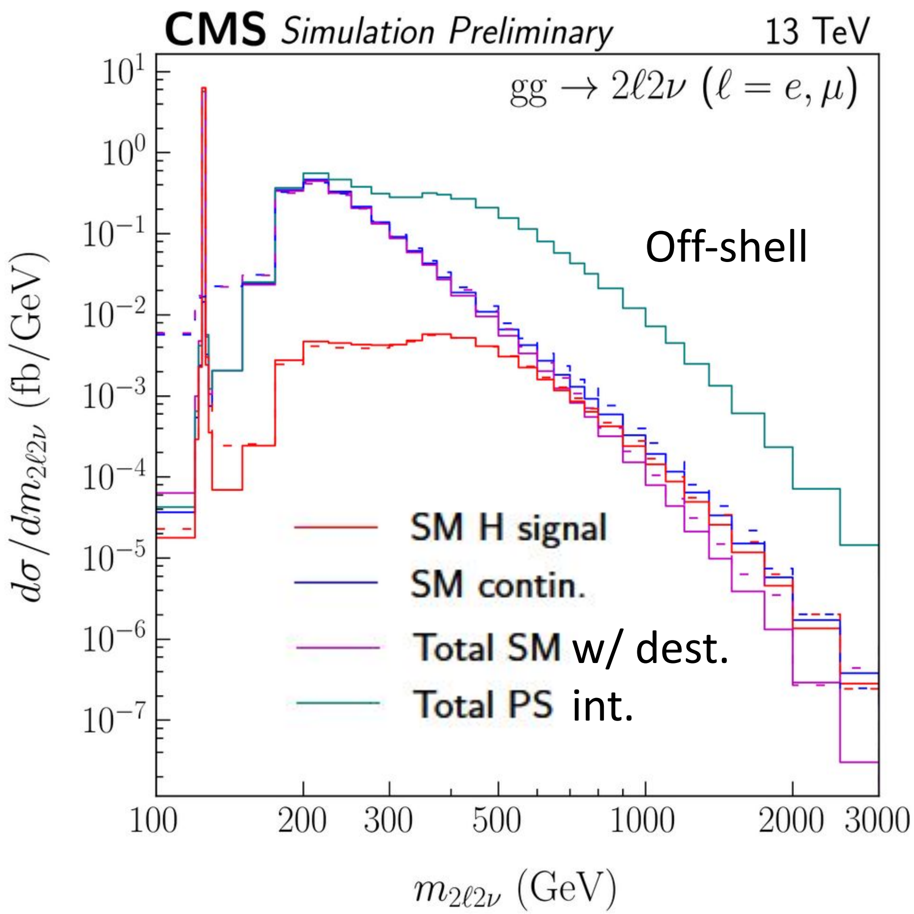
LHCHWG1 Offshell subgroup meeting

Apr. 28, 2022

Introduction

We describe the methods for the signal simulation used in past analyses:

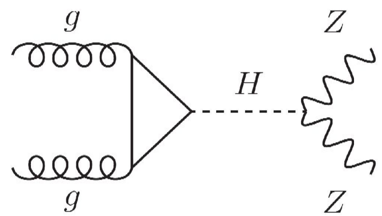
- Off-shell analyses need to deal with various interference effects, so simulation needs to be done consistently and at the same perturbative order for all components.
- The analyses benefit from categorization of events for the dominant Higgs production modes, gluon fusion and EW (mostly VBF in the SM), so they also rely on the information from the two associated jets leading in p_T .



Gluon fusion process

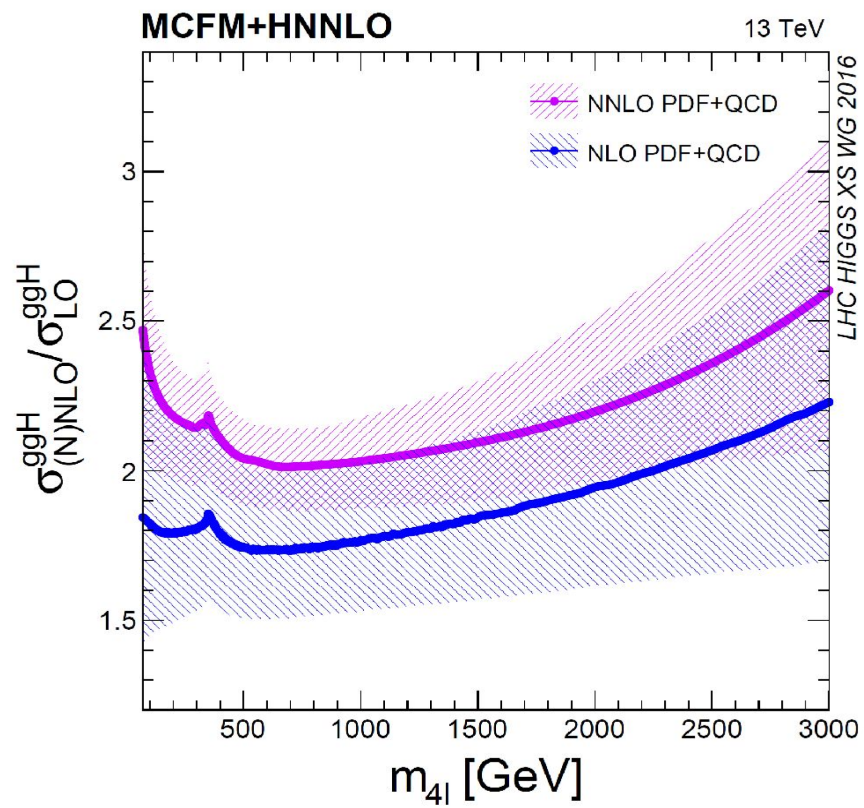
Gluon fusion: Higgs amplitude

$gg \to H \to ZZ$:



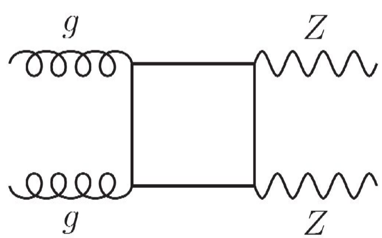
Full cross section calculation is available at different orders for the different components:
→ $gg \to H \to ZZ$: $\mathrm{N^3LO}$ in QCD around $m_H = 125$ GeV, NNLO for the full m_{ZZ} dependence, NLO or LO for event simulation

→ K-factors are
large for NLO/LO (~1.7-1.8),
smaller and flatter for NNLO/NLO (~1.2-1.3),
and the $\mathrm{N^3LO}$/NNLO K-factor is 1.10.



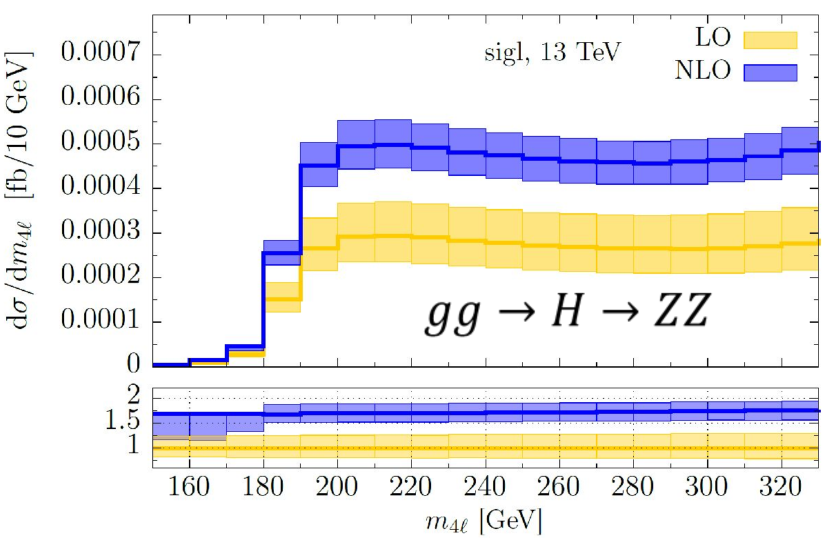
Gluon fusion: Continuum amplitude

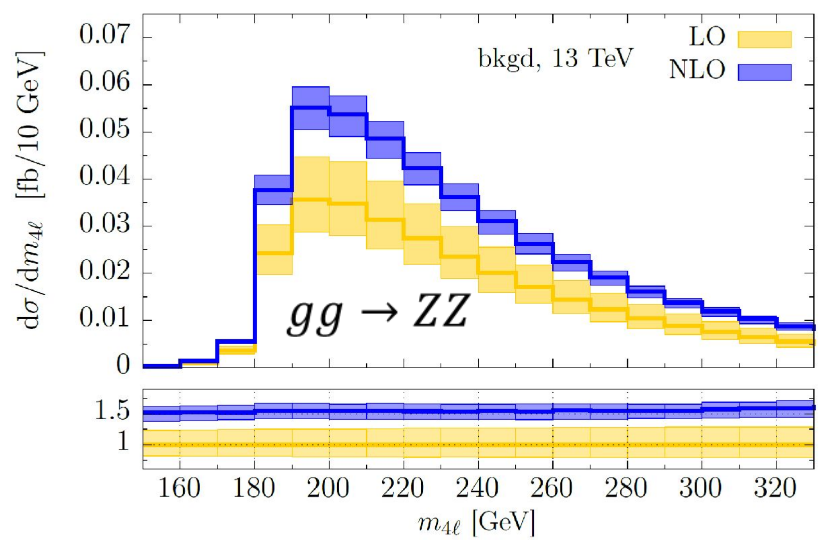
$gg \to ZZ$:



Full cross section calculation is available at different orders for the different components:
gg ⟶ ZZ continuum (and interference): Full calculation and simulation with loop effects are only available at LO in QCD

- Approximate NLO calculations from arXiv:1605.04610 show
  K-factors for gg ⟶ ZZ continuum, gg ⟶ H ⟶ ZZ, and their interference within ~10%
- Current procedure is to use K-factors for gg ⟶ H ⟶ ZZ on all components, and
  unc. $\boldsymbol{\kappa}_{ggzz}$ = 1 ± 0.1 on continuum with related scale $\sqrt{(\boldsymbol{\kappa}_{ggzz})}$ on interference.



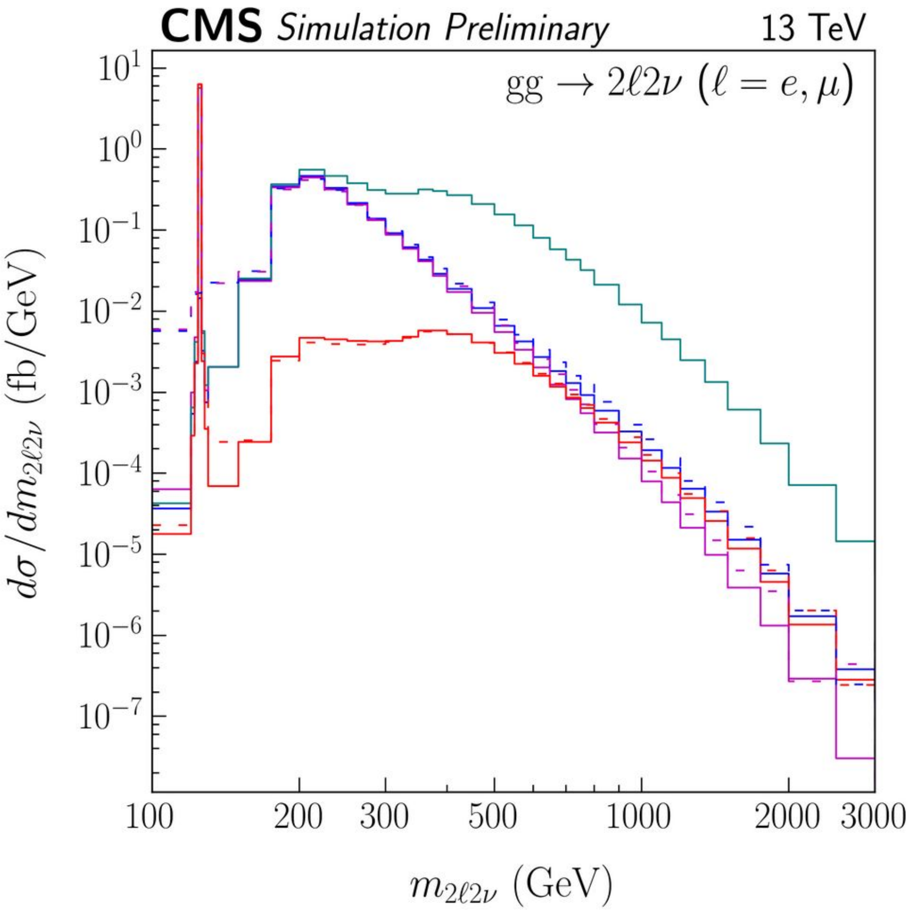


Gluon fusion: Event generation

For the Higgs amplitude contribution, continuum ZZ, or interference, MC event generation can be done in two ways:

- Use JHUGen/MCFM to produce events at LO in QCD, apply NNLO K-factors and N^3LO flat normalization
  - ❖ Relies on Pythia for jet multiplicity and kinematics
- Use POWHEG to produce gg→H, JHUGen for , H→ZZ and the MELA matrix elements from JHUGen/MCFM (instead of event generation) to obtain continuum ZZ and interference.
  - ❖ POWHEG cannot produce off-shell line shape. Instead, produce samples for Higgs samples at m_H=125, 160 ... 200 ... 3000 GeV, which have increasingly larger widths.
  - ❖ hfact = m_H / 10+37.5 GeV to match p_T^H to NNLO+NNLL HRES predictions.
  - ❖ For the gg→H(125)→ZZ amplitude, the only differences in these samples are the propagator and the correction of the m_{ZZ} line shape for the evolution of BR(H→ZZ). The former is just reweighting the propagator to a BW(m_H=125 GeV, Γ_H=4.1 MeV), so it is basically part of the MELA reweighting (arXiv:2002.09888) procedure, and the latter is added as a modification of event weights when running the JHUGen decay step.
  - ❖ The samples are glued together in the end to produce the full spectrum.
  - ❖ We observe this approach produces stable results in jet multiplicity and other kinematics after Pythia.

Gluon fusion: Distributions



→ Jet-inclusive m_{ZZ} distribution comparison exemplified with the $2\ell 2\nu$ final state

→ Almost perfect overlap between POWHEG+JHUGen and JHUGen/MCFM predictions

→ PS refers to the purely-pseudoscalar $H \to ZZ$ coupling with the same on-shell gluon fusion signal strength as SM since we also show an illustration of off-shell effects for a BSM scenario.

SM H signal ($|\mathrm{H}|^2$) — Total SM ($|\mathrm{H}+\mathrm{C}|^2$)

SM contin. ($|\mathrm{C}|^2$) — Total PS ($|\mathrm{PS}+\mathrm{C}|^2$)

Legend is the same in other plots, so I won't repeat it on the next slides.

Constraints:

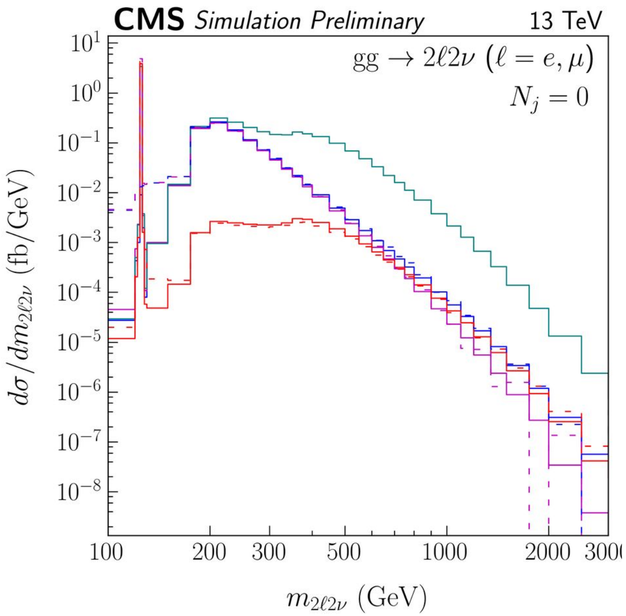
$\mu_{\mathrm{sig}}^{\mathrm{on-shell}}(\mathrm{gg} \to \mathrm{H} \to \mathrm{ZZ} \to 2\ell 2\nu) = 1$

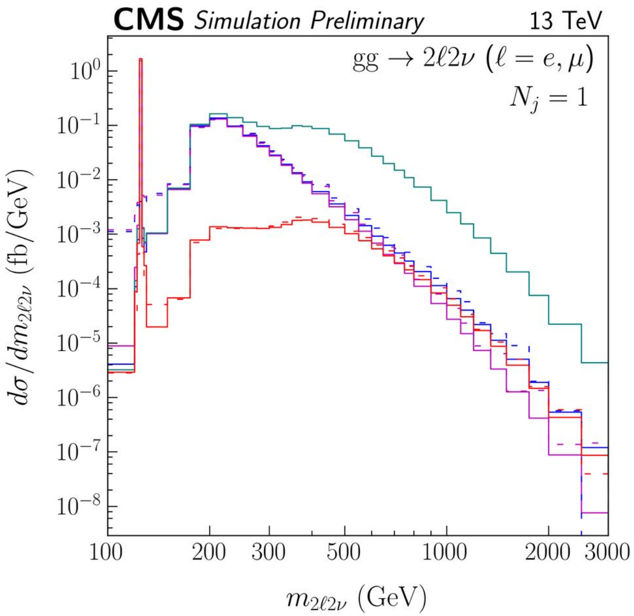
JHUGen/MCFM scale: 1.01

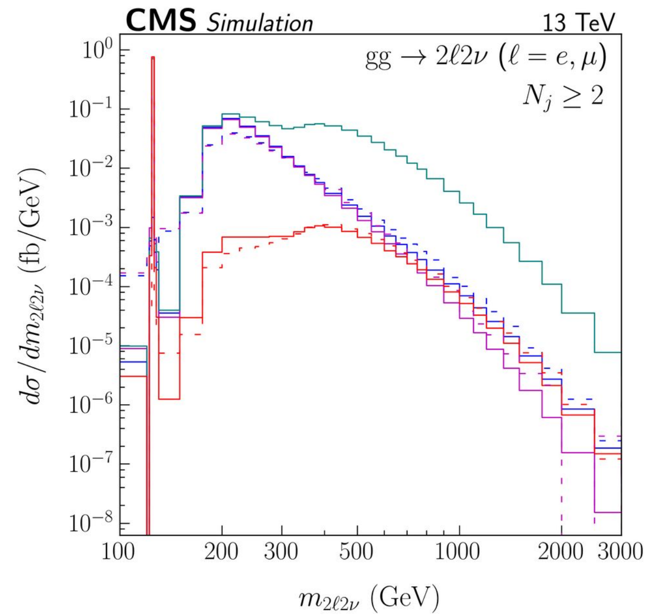
——— POWHEG+JHUGen

- - - JHUGen/MCFM (LO QCD)

Gluon fusion: Jet-exclusive



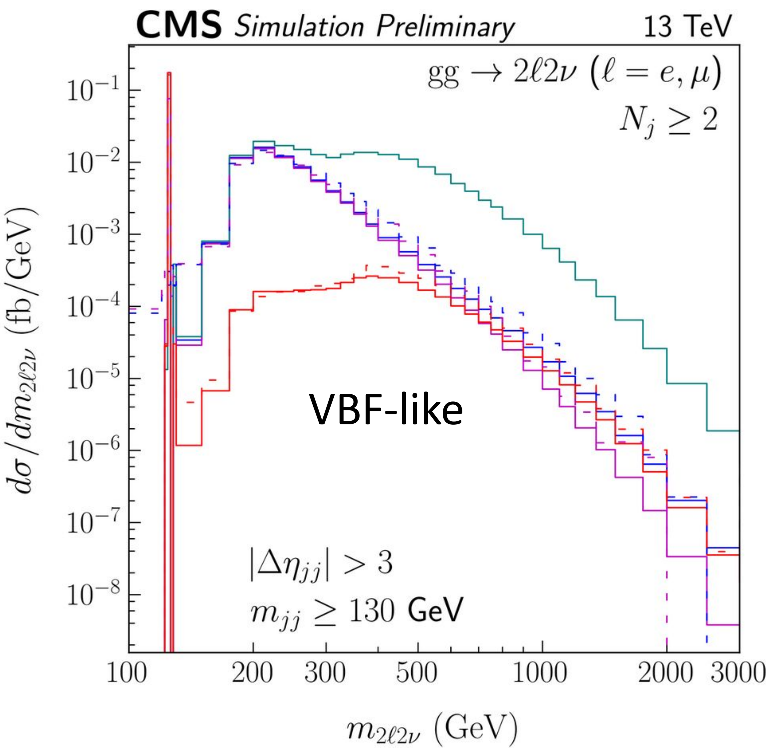


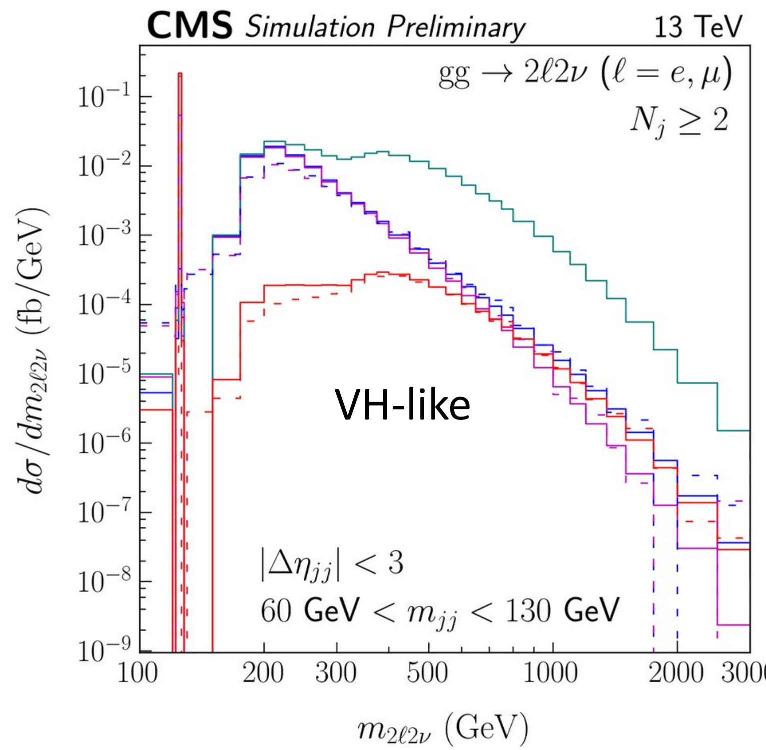


When events are split by jet multiplicity (jets are clustered by anti-k_T algorithm within the radius ΔR = 0.4, with p_T > 30 GeV, $|\eta| < 4.7$ as used in CMS analyses), we find $N_j = 0{,}1$ have similar levels of agreement, but the LO m_{ZZ} distribution destabilizes for $N_j \geq 2$.

→ In the next few slides, we will show that the NLO sample provides more stable dijet kinematics for studies of VH- and VBF-like topologies at off-shell, where gluon fusion would contaminate these regions for the targeted processes.

Gluon fusion: Focus on $N_j \geq 2$



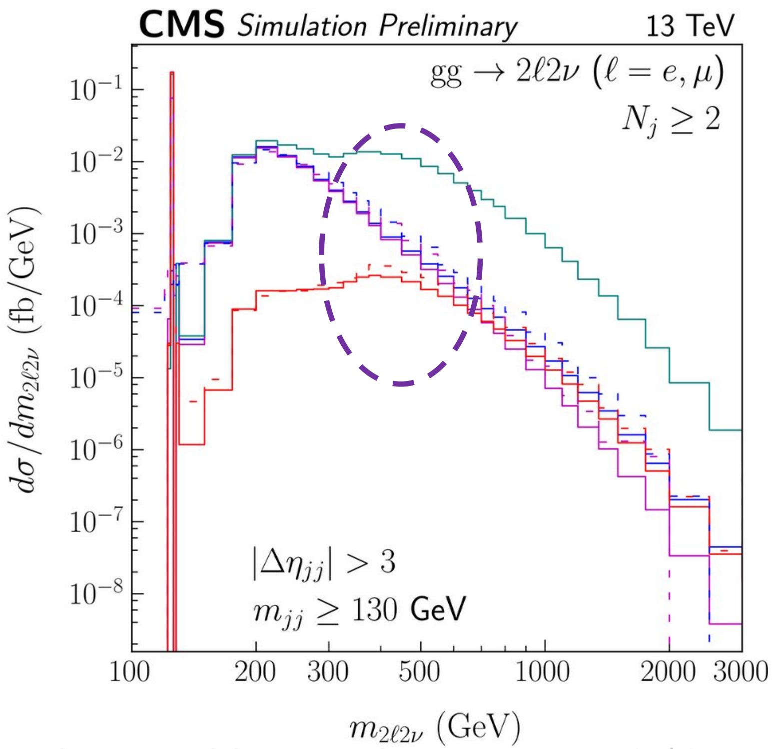


We can categorize $N_j \geq 2$ events further into regions with

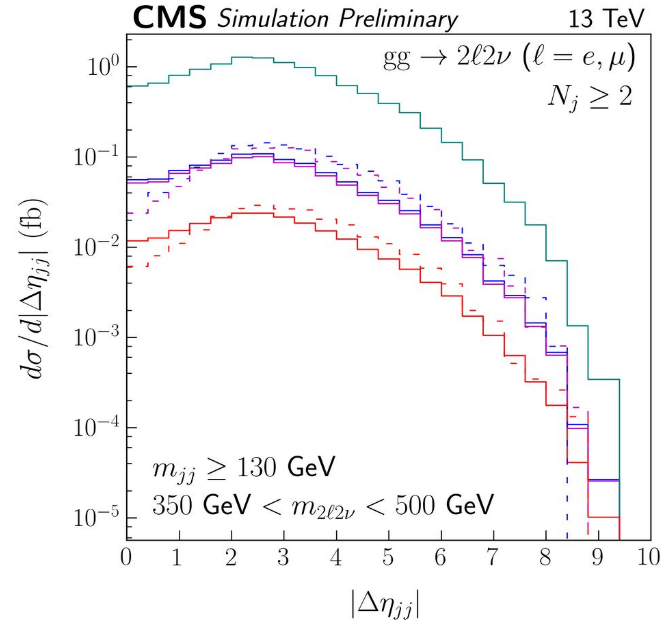
→ high m_{jj} with forward jets ($m_{jj} > 130$ GeV, $|\Delta\eta_{jj}| > 3$, VBF-like topology)

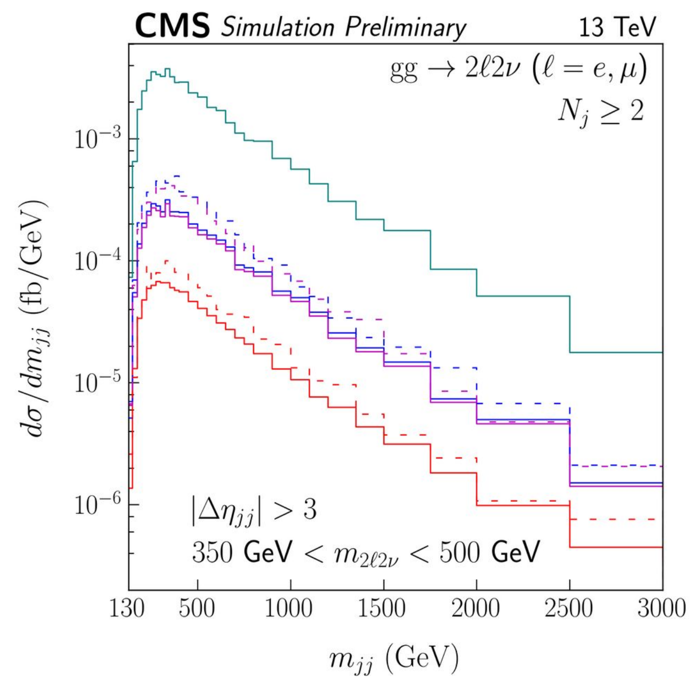
→ low m_{jj} with central jets ($m_{jj} = 60 - 130$ GeV, $|\Delta\eta_{jj}| < 3$, VH-like topology)

Gluon fusion: Focus on $N_j \geq 2$, VBF-like

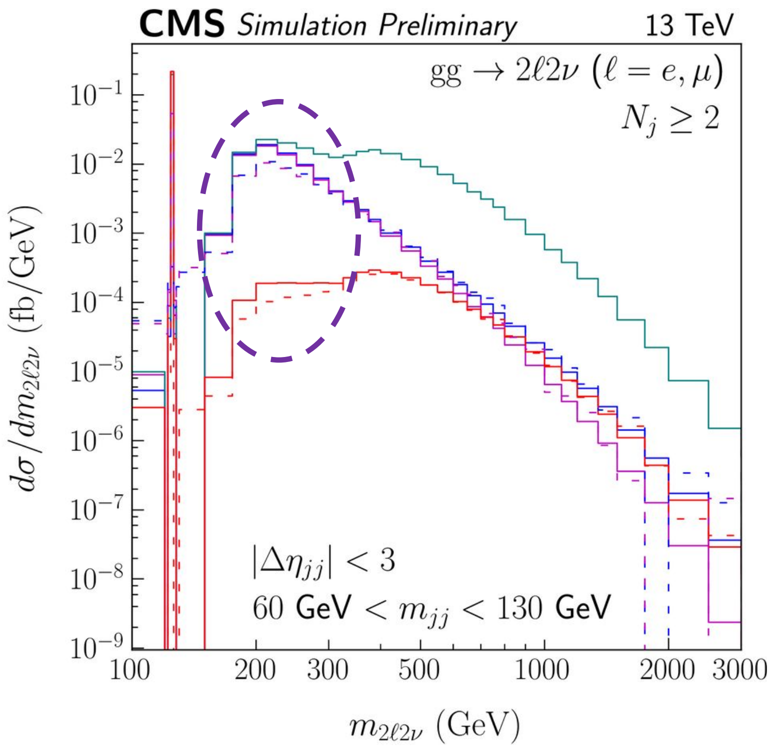


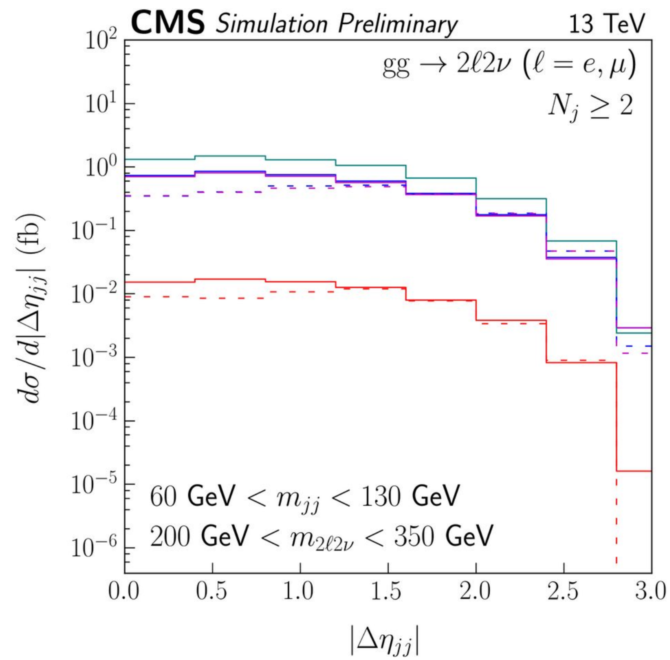
In the VBF-like topology, $m_{ZZ} = 350 - 500$ GeV,
→ $|\Delta\eta_{jj}|$ shape is different between LO and NLO samples for high m_{jj}.
→ Conversely, for the $|\Delta\eta_{jj}| > 3$, we see almost uniformly more events at high m_{jj} in LO sample.
→ The shapes evolve at different m_{ZZ} values and partially cancel out.





Gluon fusion: Focus on $N_j \geq 2$, VH-like

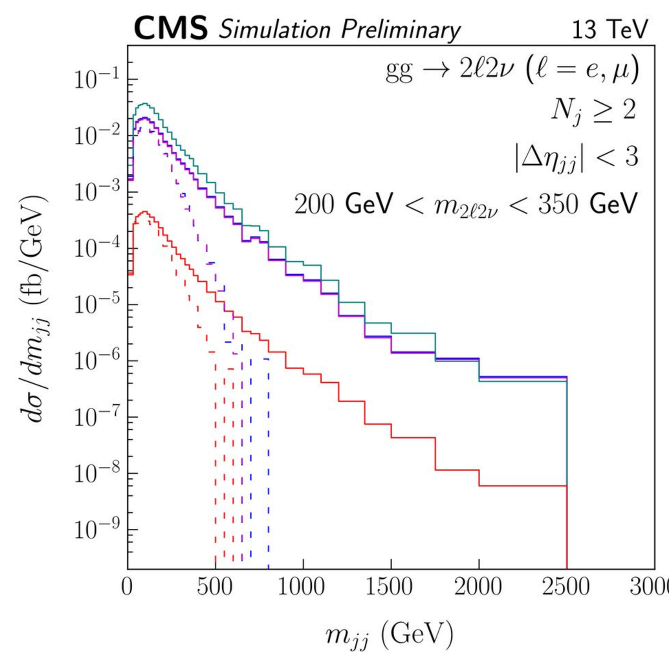




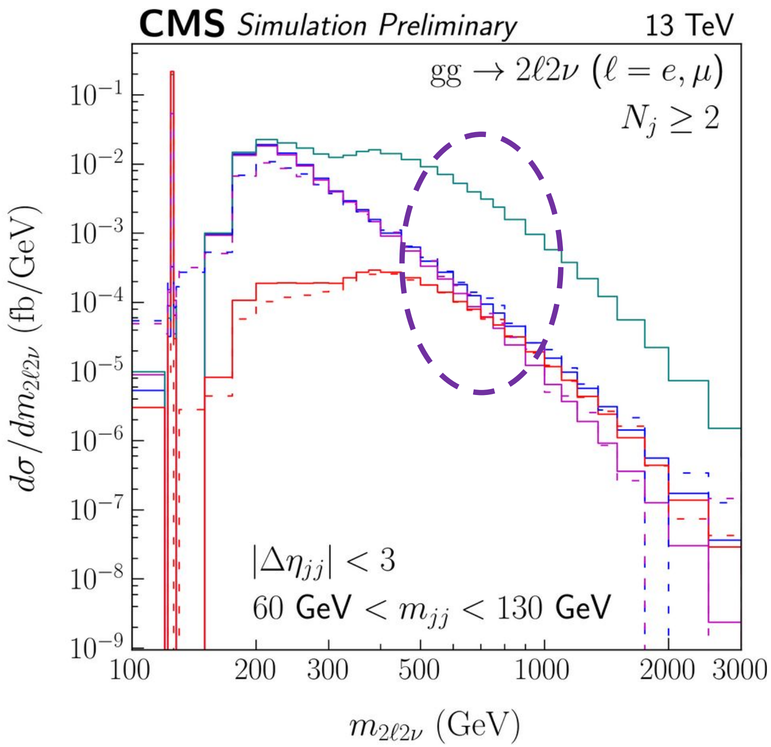
In the VH-like topology, $m_{ZZ} = 200 - 350$ GeV,

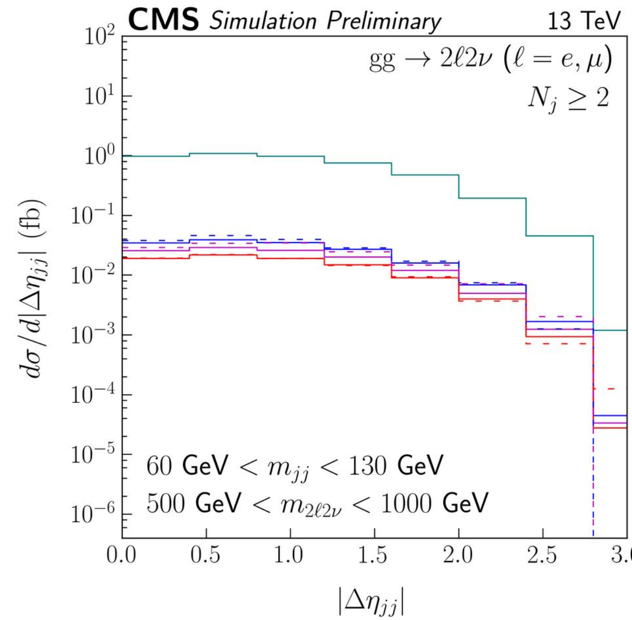
→ Jet radiation on the LO sample shows deficiency near $|\Delta\eta_{jj}| \sim 0$.

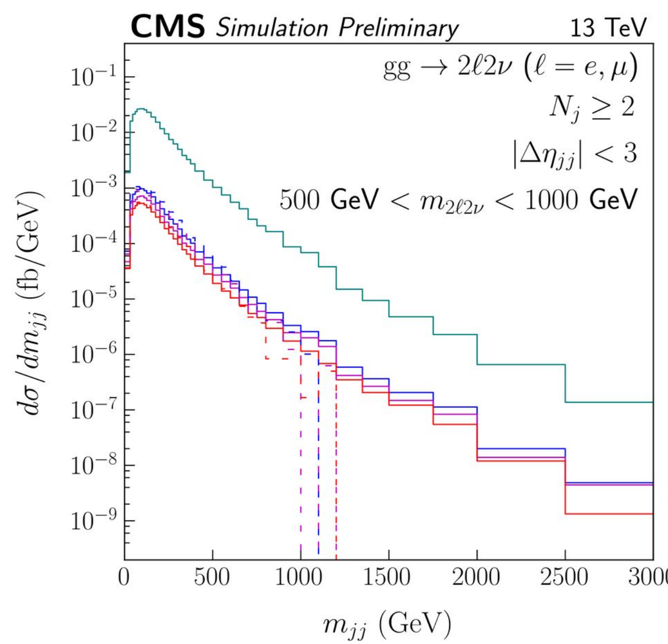
→ Conversely, for the $|\Delta\eta_{jj}| < 3$ events, we see a faster decline in m_{jj} in LO sample after parton shower.



Gluon fusion: Focus on $N_j \geq 2$, VH-like







In the VH-like topology,

→ Both differences are milder (but gradually evolving) at higher m_{ZZ} values.

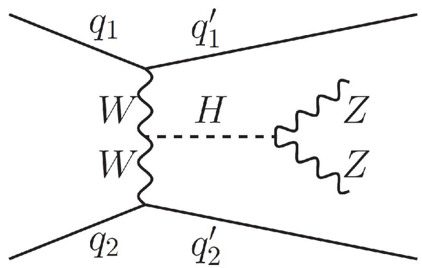
→ We also note $|\Delta\eta_{jj}|$ and m_{jj} shapes are stable in the NLO sample across m_{ZZ}.

EW (VBF+VH/VBS+VZZ) process

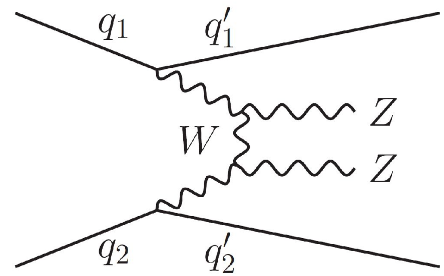
EW process: Available methods

$q_1 q_2 \to q_1' q_2' H \to q_1' q_2' ZZ$:

Example VBF



$q_1 q_2 \to q_1' q_2' ZZ$:



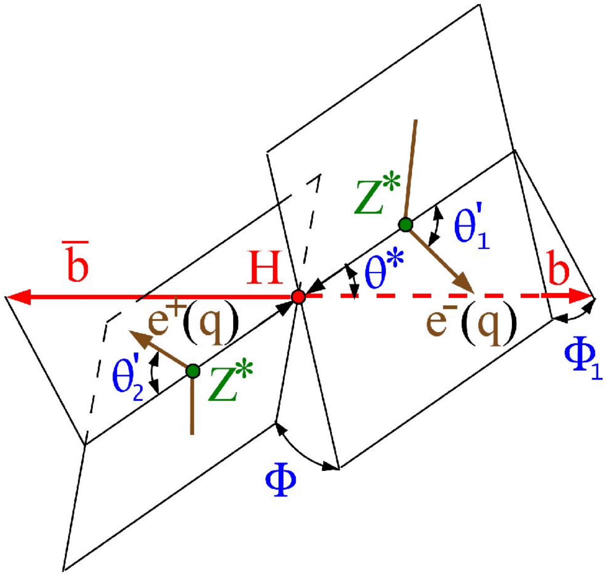
Example VBS

Both matrix element (MELA) and event simulation (MCFM/JHUGen) are available at LO in QCD, for SM or BSM Higgs hypotheses, and continuum background.

- We can improve the event simulation technique for jet kinematics in the off-shell region by starting with POWHEG+JHUGen samples for VBF NLO in QCD, and ZH and WH ~NNLO in QCD (NLO + MiNLO HVJ) as before, and then, apply MELA ME reweighting.
- One has to account for the extra gluons from POWHEG before passing event kinematics to the LO ME calculation.

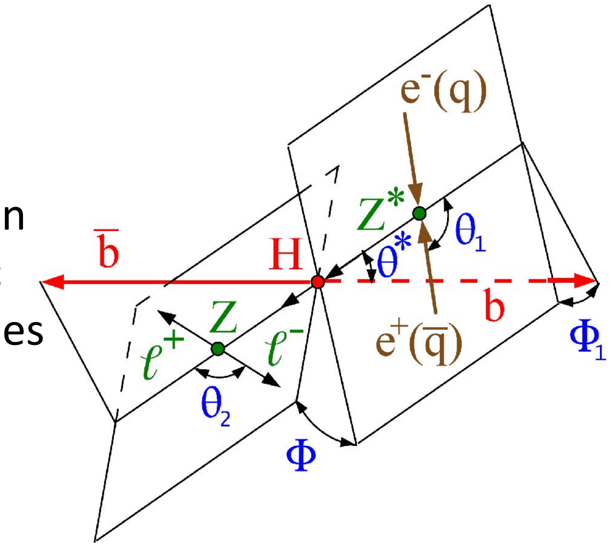
  - We demonstrate in the next few slides that merging gluons to the closest quark in four-momentum dot-product reproduces the LO topology passed to the ME calculation (i.e., gluons from NLO QCD on top of these EW processes are predominantly soft/collinear).

EW process: LO topology kinematics



VBF production kinematic observables

VH production kinematic observables



Can compare the independent kinematic d.o.f.s in the VVH production vertex between NLO-merged vs actual LO distributions from the hard process:

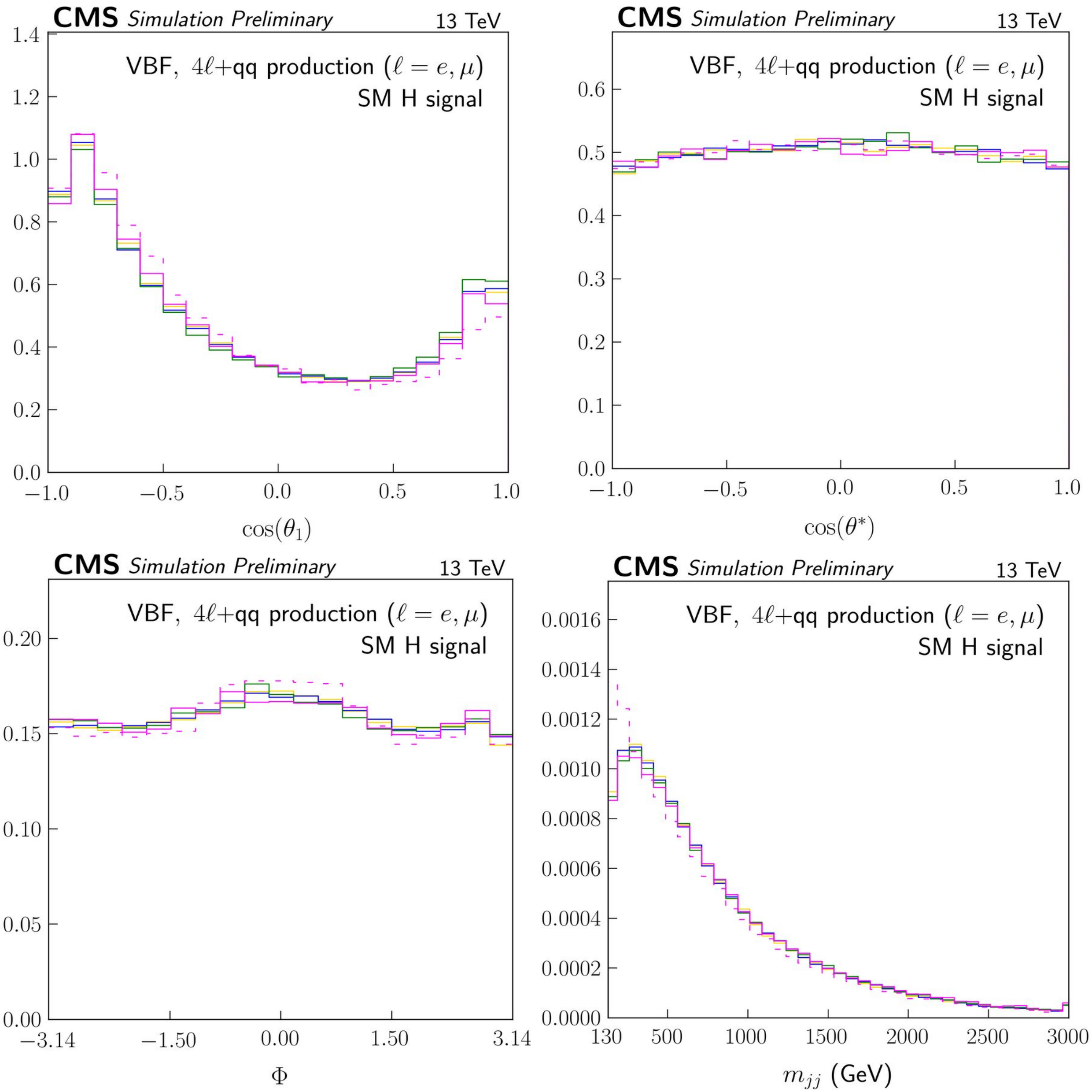
→ This tests the validity of using LO MEs in reweighting.

→ In the ordinary case of VH, all angular variables are calculated in the rest frame of the relevant particles (e.g., θ^* in the Higgs rest frame, θ_1 in the Z^* rest frame)

→ In the case of VBF, $q^2_{V^*} < 0$, so θ'_1 and θ'_2 are computed in the rest frame of the $p_T^{Hjj} = 0$ frame since boosting into the V^* frames is not possible.

→ Next few slides show a few examples.

EW process: VBF LO topology (on-shell)



Constraints:

$m_{jj} > 130$ GeV

124.95 GeV $< m_{4\ell} < 125.05$ GeV

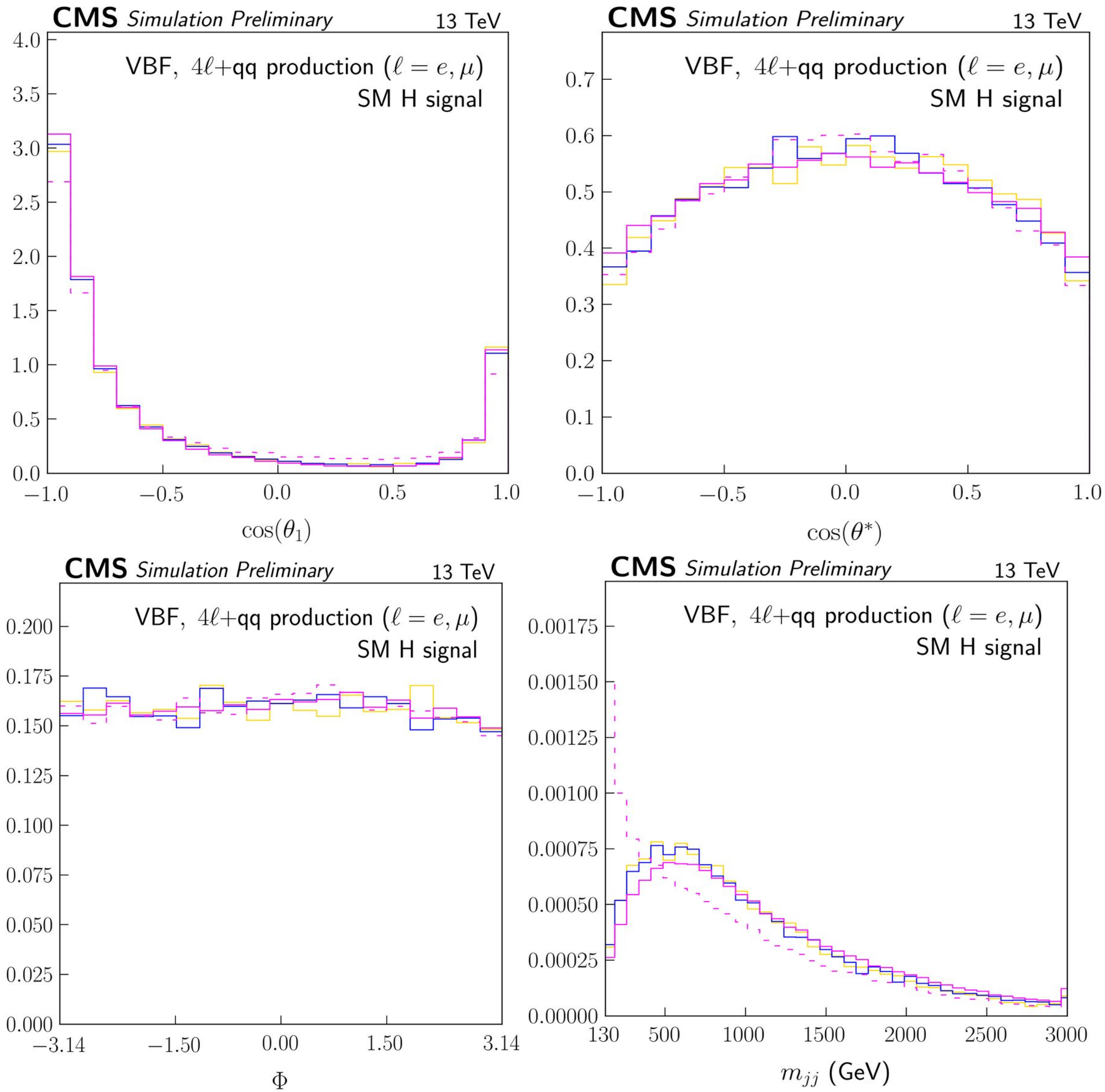
w/ sel. reqs. on ℓ, j

- JHUGen/MCFM (LO QCD)
- Phantom (LO QCD)
- JHUGen on-shell (LO QCD)
- POWHEG+JHUGen merged
- POWHEG+JHUGen unmerged

→ Color legend is the same in all other plots, so will not repeat it.

→ POWHEG unmerged refers to using the two leading-p_T partons instead of deducing the LO topology by merging gluons into quarks for illustration.

EW process: VBF LO topology (off-shell)



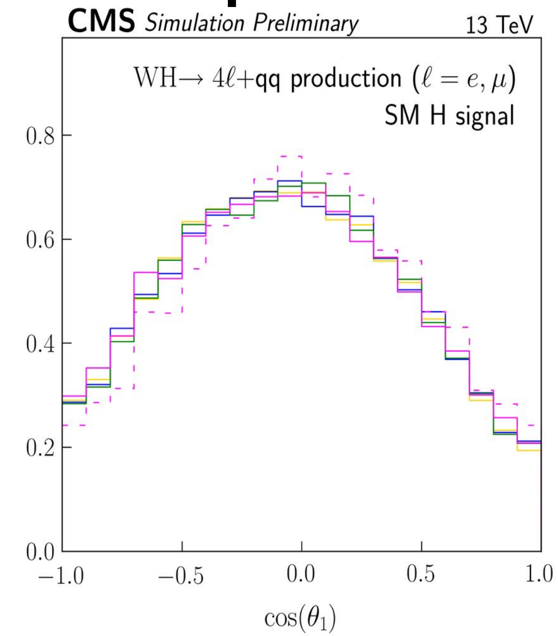
250 GeV $< m_{4\ell} < 500$ GeV

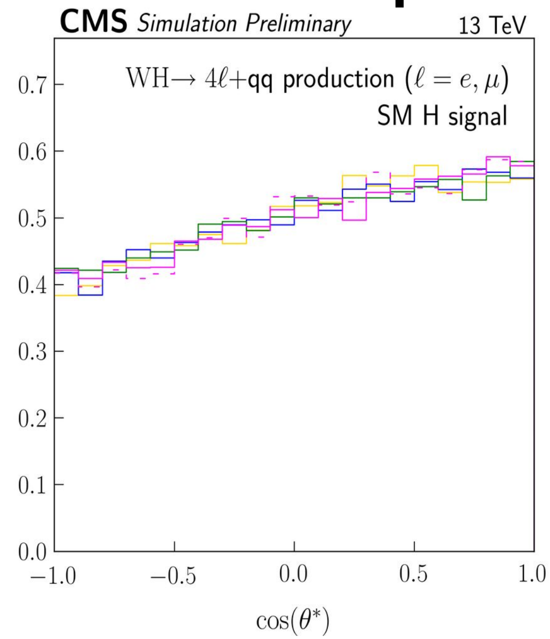
→ The magenta histograms for POWHEG include ME reweighting, which is equivalent to the rewgt. of the BW, so that m_{ZZ} distributions match.

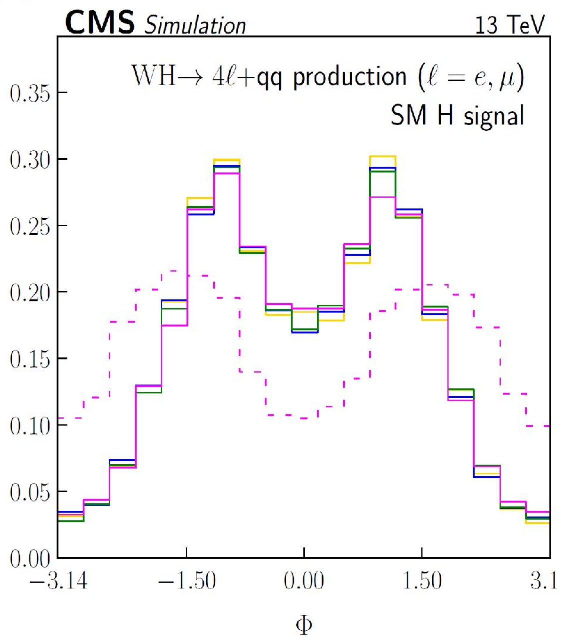
→ Match between different kinematic observables is very good.

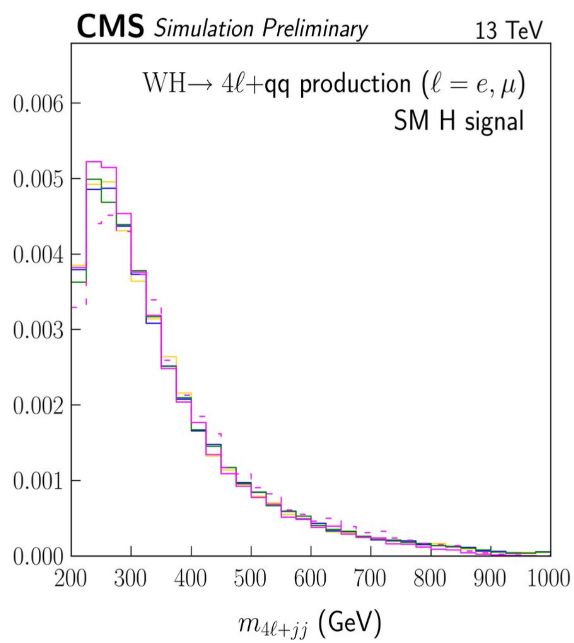
→ As expected, kinematics from H+2 leading-p_T partons without gluon merging is not always a good approximation to the LO topology.

EW process: WH LO topology (on-shell)









Constraints:

$70 \text{ GeV} < m_{jj} < 90 \text{ GeV}$

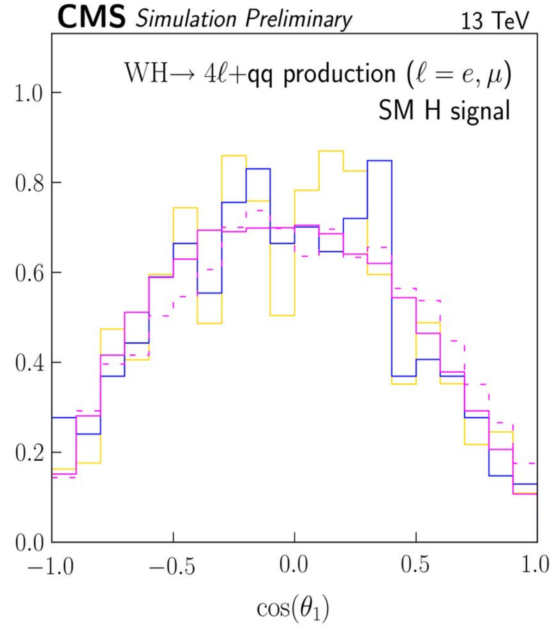
$124.95 \text{ GeV} < m_{4\ell} < 125.05 \text{ GeV}$

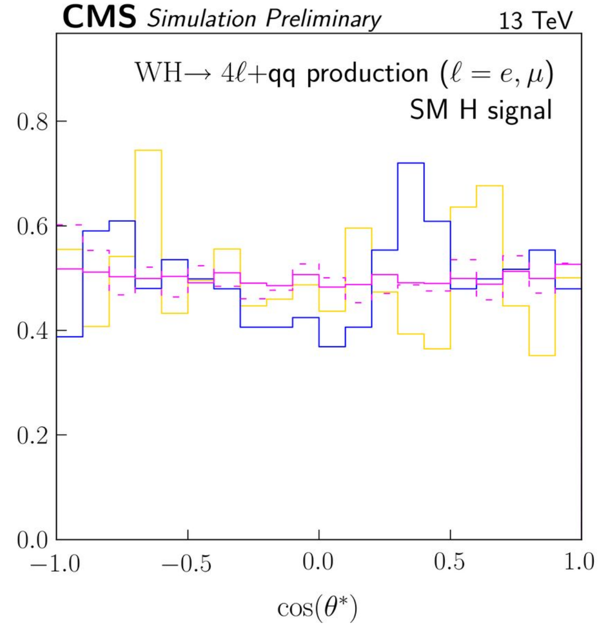
w/ sel. reqs. on ℓ, j

- JHUGen/MCFM (LO QCD)
- Phantom (LO QCD)
- JHUGen on-shell (LO QCD)
- POWHEG+JHUGen merged
- POWHEG+JHUGen unmerged

- Color legend is again the same in all other plots, so will not repeat it.

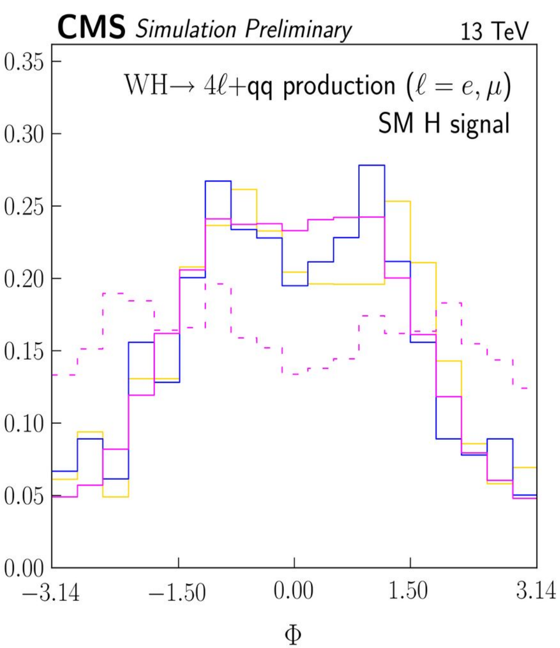
EW process: WH LO topology (off-shell)

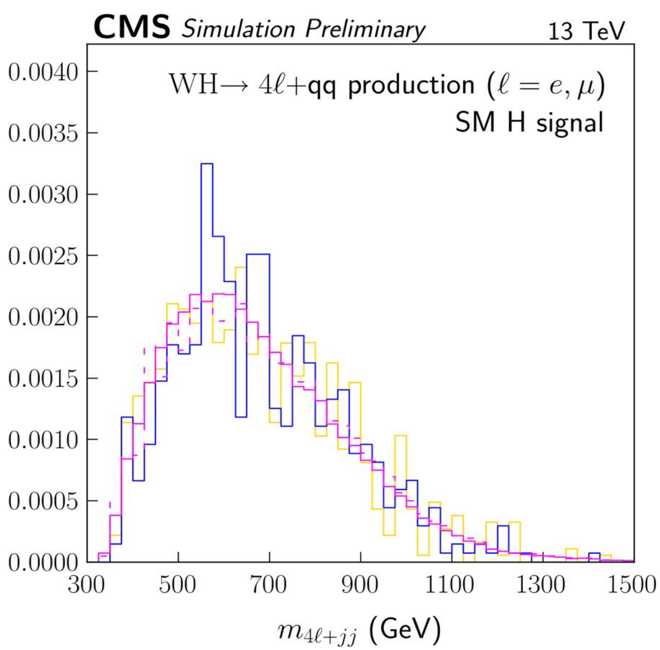




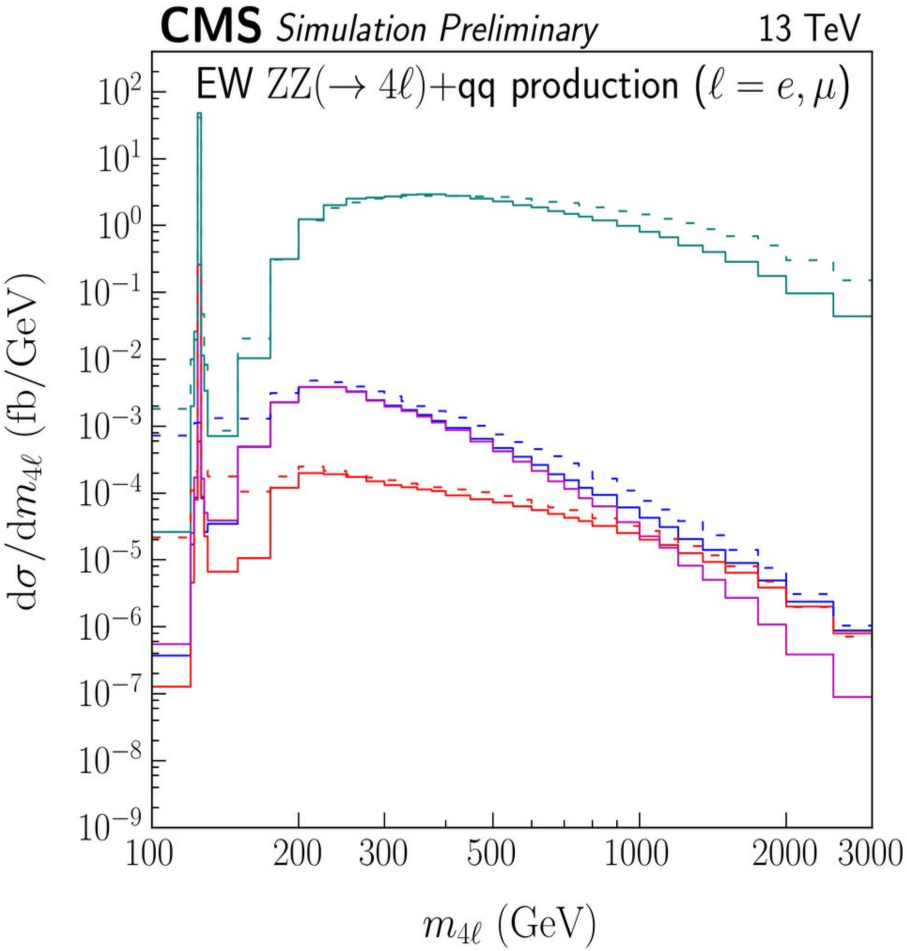
$250\ \text{GeV} < m_{4\ell} < 500\ \text{GeV}$

- ☐ In both on-shell and off-shell regions, we also see adequate match in WH.
- ☐ Rest of the plots are in backup.





EW process: Distributions after Pythia



→ Jet-inclusive m_{ZZ} distribution comparison exemplified with the $4\ell + qq$ final state

→ Prediction from POWHEG+JHUGen for off-shell is less than JHUGen/MCFM, even for the Higgs amplitude. The reason is understood to be from the extra partonic contributions in the initial and final states and their evolution as a fcn of m_{ZZ}.

→ Next will be jet multiplicity and kinematics distributions.

SM H signal ($|\mathrm{H}|^2$)
Total SM ($|\mathrm{H}+\mathrm{C}|^2$)
SM contin. ($|\mathrm{C}|^2$)
Total PS ($|\mathrm{PS}+\mathrm{C}|^2$)

Constraints:

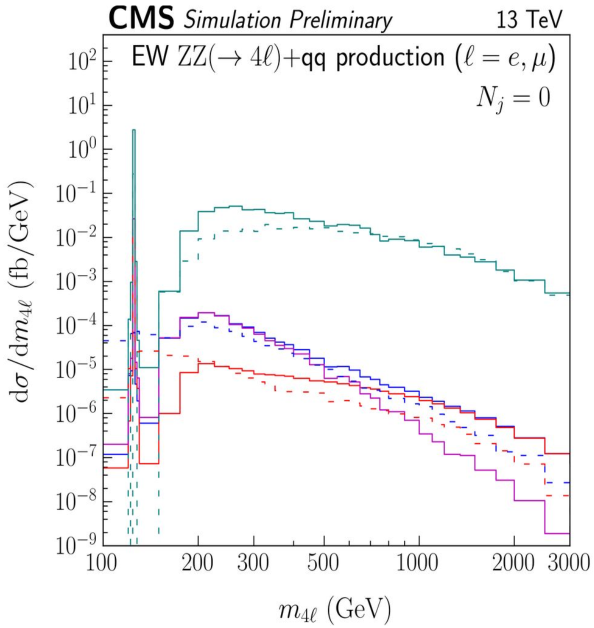
$\mu_{\mathrm{sig}}^{\mathrm{on-shell}}(\mathrm{gg} \to \mathrm{H} \to \mathrm{ZZ} \to 2e2\mu) = 1$

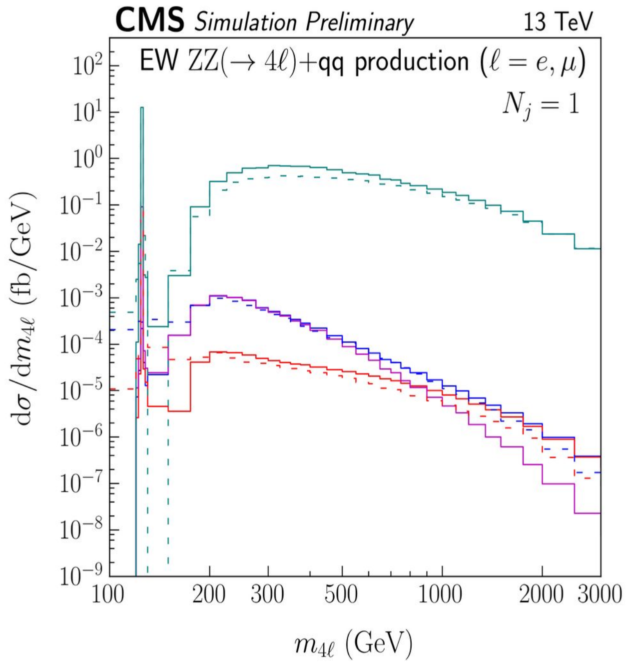
JHUGen/MCFM scale: 0.93

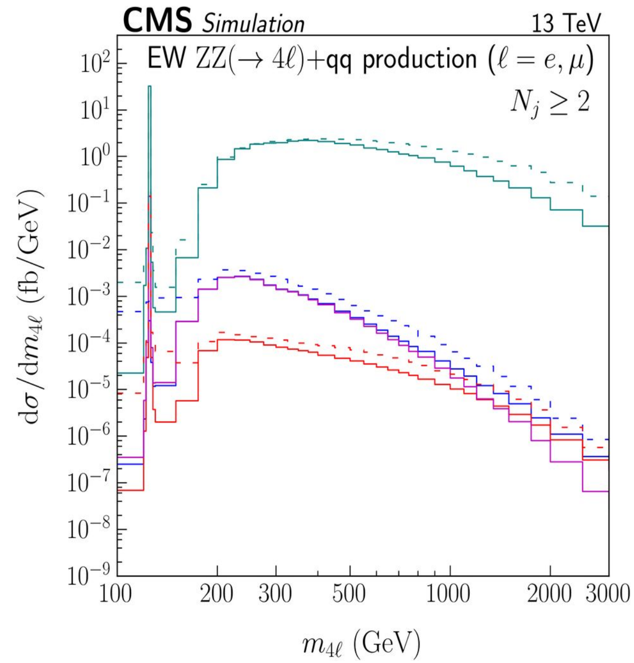
POWHEG+JHUGen

JHUGen/MCFM (LO QCD)

EW process: Jet-exclusive distributions





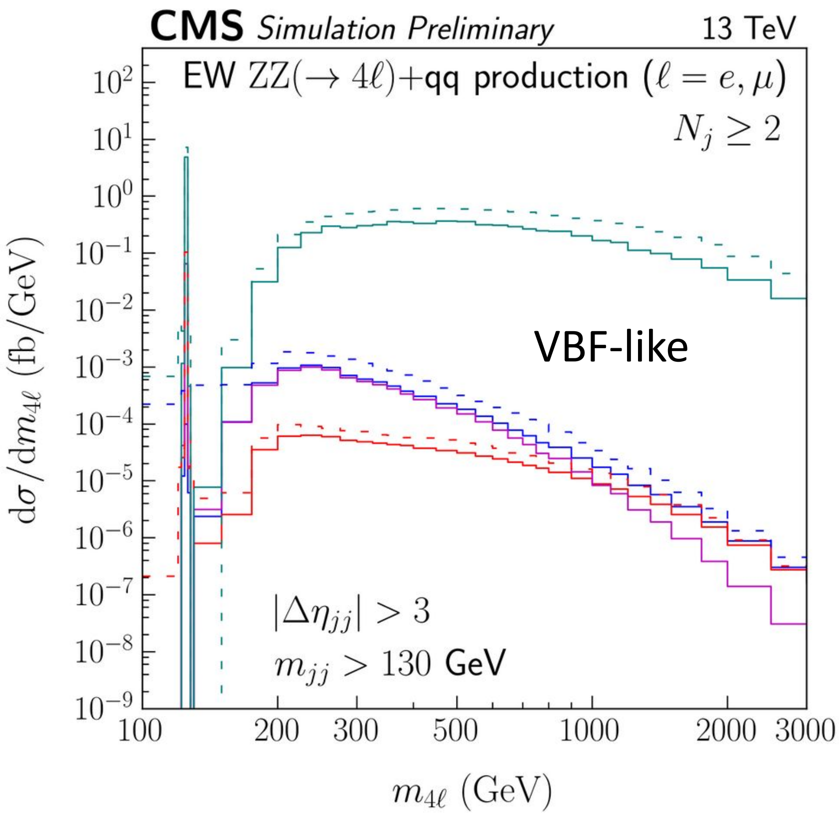


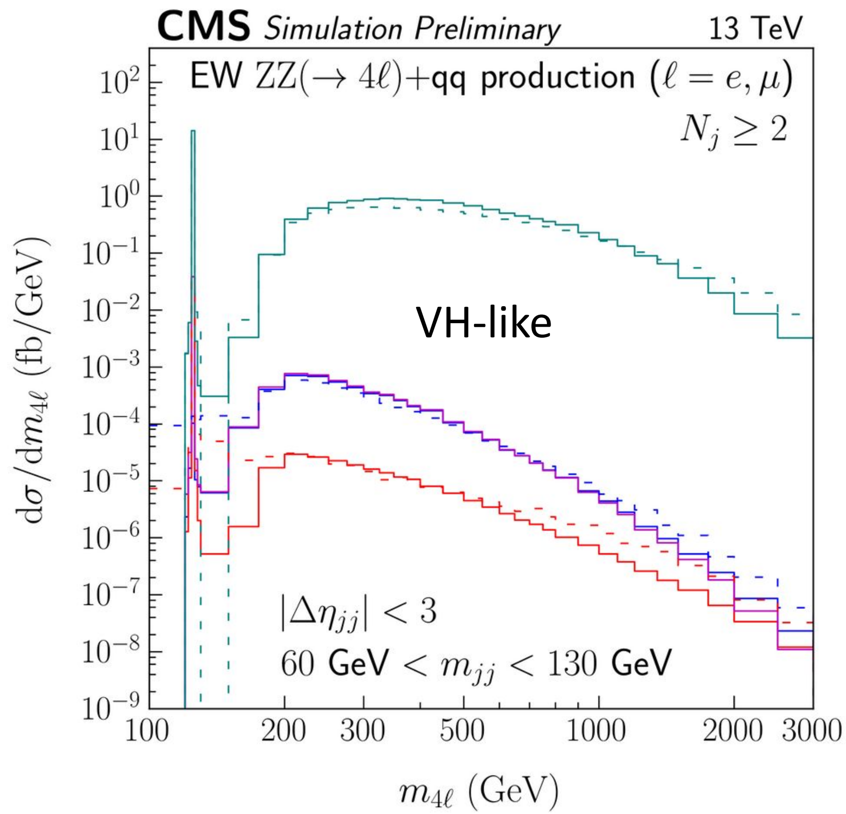
When events are split by jet multiplicity, we find $N_j = 0{,}1$ vs $N_j \geq 2$ differ in opposite directions.

→ The LO prediction for $N_j = 0$ is also distorted at m_{ZZ}~250 GeV significantly.

→ Signal vs. continuum seem to go in similar directions, so the features are not attributable to the LO topology recasting.

EW process: $N_j \geq 2$, VBF- and VH-like

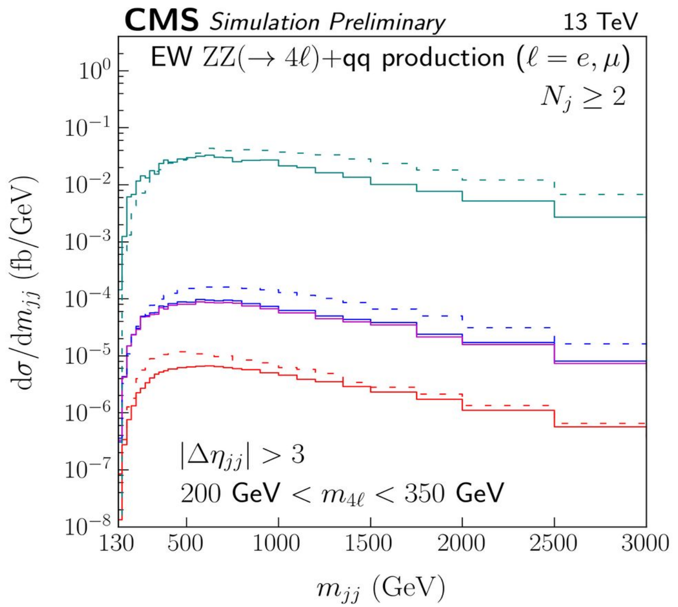


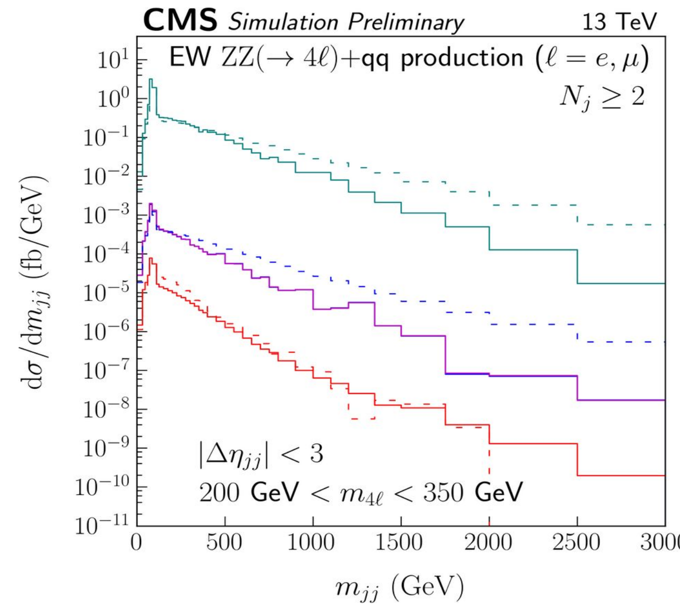


→ Distributions match decently between LO samples and POWHEG prediction for the VH-like topology

→ The differences in $N_j \geq 2$ are observable in the VBF-like region as well.

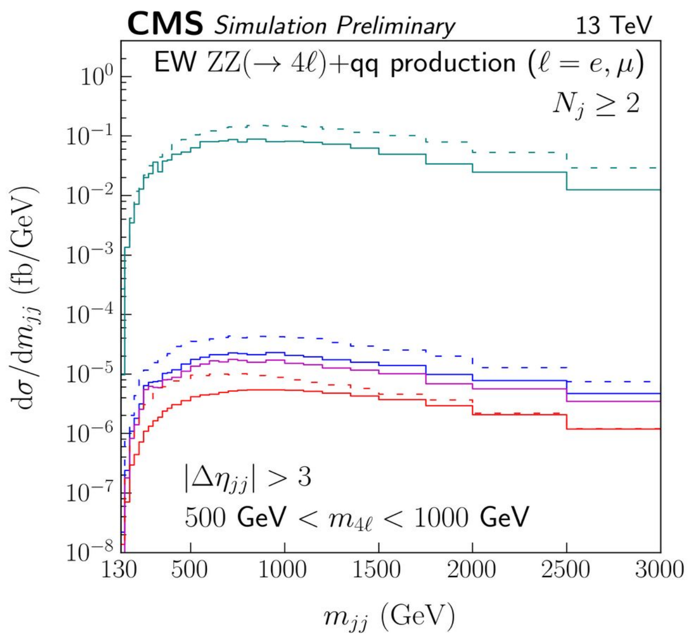
EW process: m_{jj}

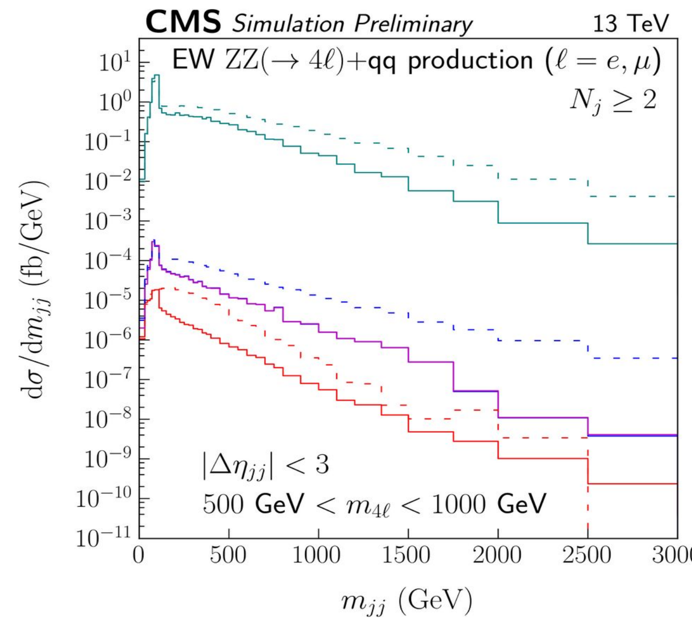




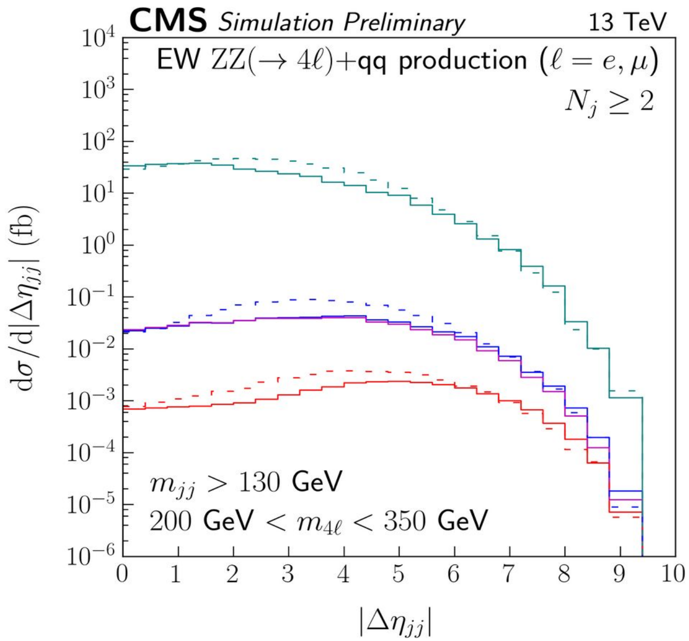
→ Unphysical m_{ZZ} evolution of high-m_{jj}, low $|\Delta\eta_{jj}|$ in LO sample, and discrepancies in m_{jj} shape for high $|\Delta\eta_{jj}|$.

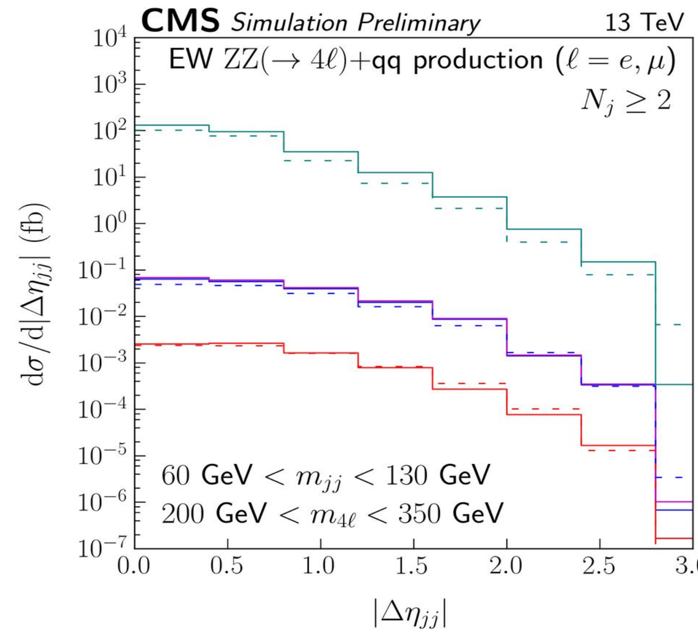
→ m_{ZZ} evolution of POWHEG prediction is stable for both low and high $|\Delta\eta_{jj}|$.



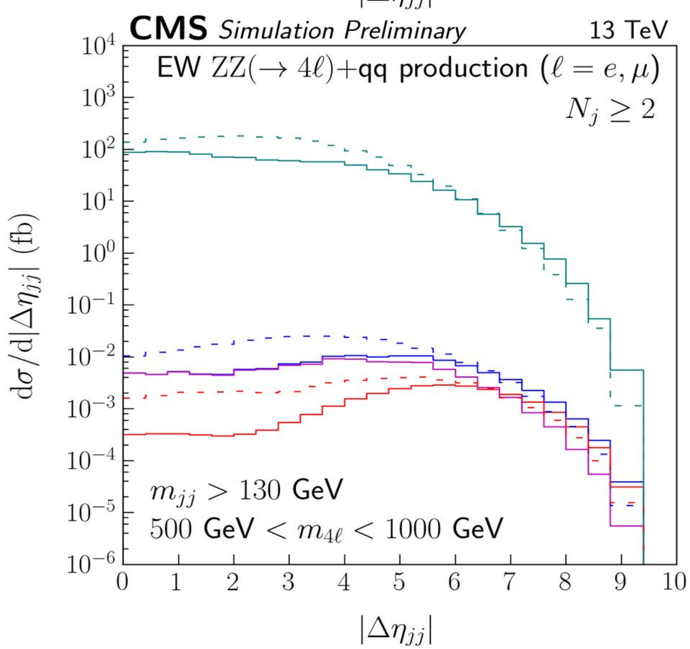


EW process: $|\Delta\eta_{jj}|$

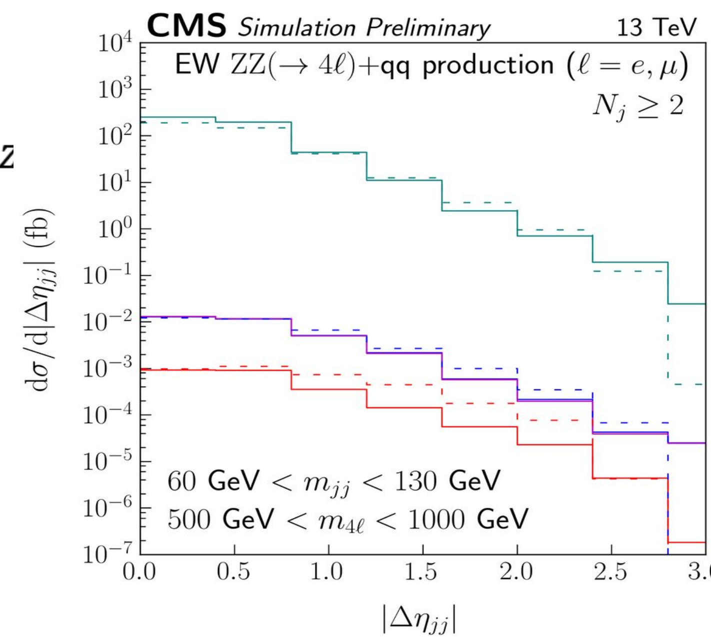




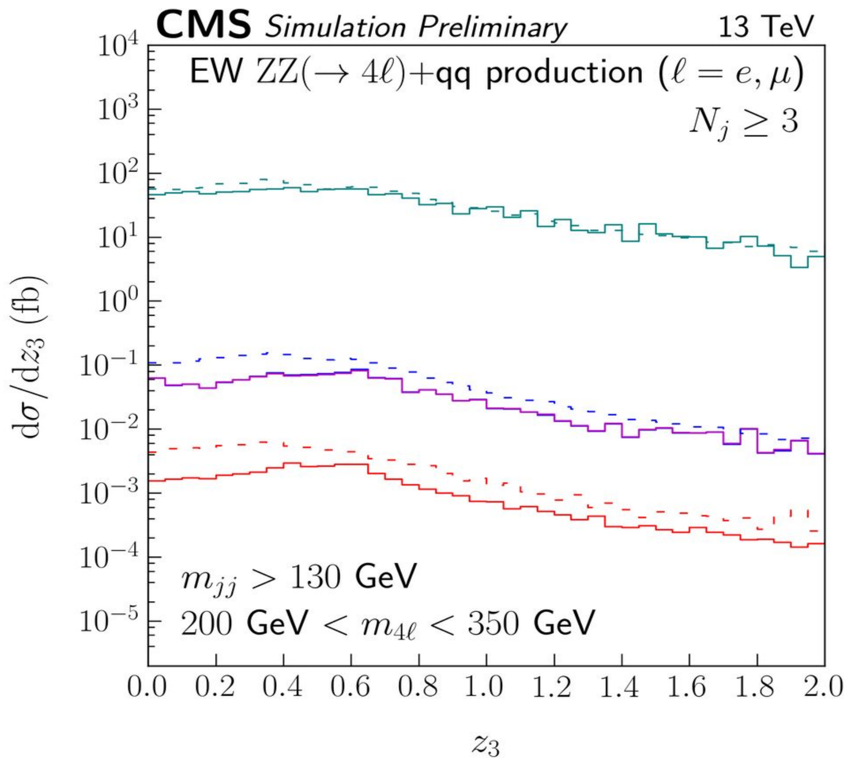
At high-m_{jj}, disagreement in the central region.

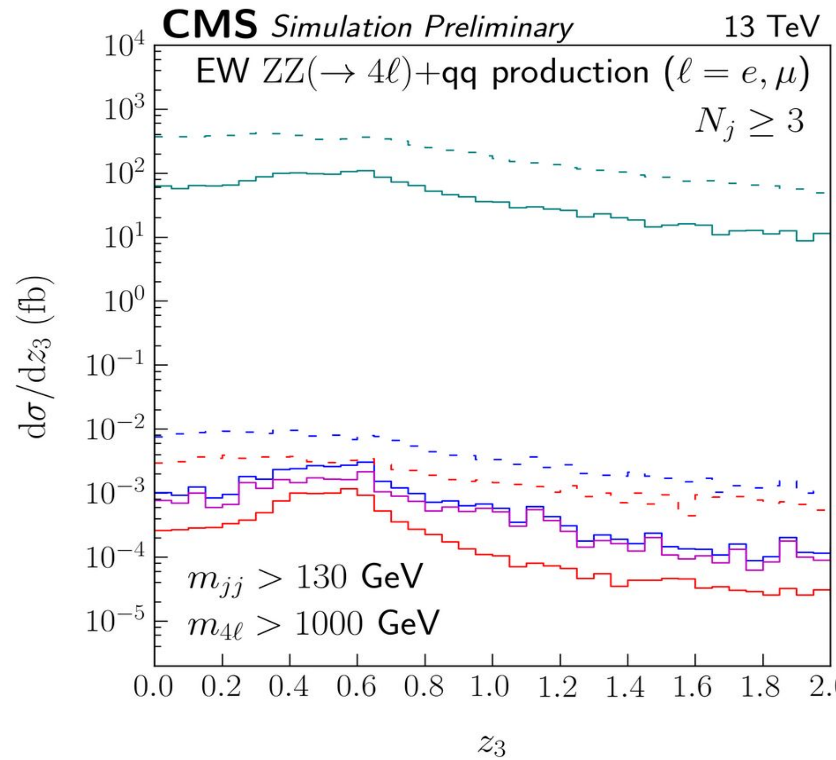


At low-m_{jj}, the disagreement only happens at higher m_{ZZ} values, $|\Delta\eta_{jj}| > 0.5$.



EW process: Centrality of third jet (z_3)





While CMS analyses do not utilize the third jet, which is more sensitive to parton showering schemes, we can still look at its centrality to understand if jets from Pythia are too central.

→ The Zeppenfeld $z_3 = \frac{y_3 - (y_1 + y_2)/2}{|\Delta y_{12}|}$ variable → 0 for central third jets.

→ Distributions from POWHEG are close to best predictions, except for small known discrepancies near $z_3 = 0$. LO prediction, however, has too many jets near $z_3 = 0$. This is consistent with what we already observed with the two leading-p_T jets.

Summary

Presented simulation methods used in the off-shell
→ Main Higgs production modes are gluon fusion, and the EW processes VBF, VH
→ Need to include interference effects with continuum ZZ, or BSM Higgs, at consistent orders in perturbative QCD

Analyses use jet kinematics to distinguish the production mechanism of the Higgs
→ Using matrix element reweighting methods on samples NLO in QCD are found to produce more reliable distributions in jet multiplicity and kinematics.
→ Merging of the extra gluons at NLO needed in the EW process for reweighting to work, validated the recasting of event topology to LO equivalent

Backup

EW processes: Merging algorithm

The MELA and MELAANALYTICS packages impose several rules on the merging procedure for the EW processes in order to make sensible predictions:

- An incoming gluon is never merged into an incoming quark. This rule is invoked implicitly as the q^2 of the incoming partons is always the largest compared to that of any other pair of partons in the event.
- Gluons are never merged into the decay products of the H boson from the JHUGen step as they are produced during the production of the H boson with no prior knowledge about the boson's decay.
- Gluons are also never merged into the decay products of the associated W or Z boson in the VH samples. Doing so distorts the Breit-Wigner nature of these resonances significantly.
- All merging is done in the convention of outgoing particles. This means the four-momenta and charge of incoming particles are reversed in the intermediate steps when those of the two merged particles are summed.
- When an incoming gluon is merged into an outgoing quark, the charge (i.e., PDG id), and the four-momentum of the quark are reversed in the final step of the LO topology construction. This reversion is done so that the event topology ensures having exactly two incoming quarks as expected by the LO matrix elements.
- In VH samples, when extra gluons are encountered, the merging of individual gluons and that of a combined gluon (i.e., from a $\mathrm{g} \to \mathrm{gg}$ process) are all considered separately.
- In VH samples, it is also possible to encounter two extra quarks instead of gluons. These extra quarks are merged into a gluon substitute first, as they are from a $\mathrm{g} \to \mathrm{q\overline{q}}$ branching process, before the merging of this gluon substitute is considered.
- Every merging case is considered, and those that do not produce an incoming-outgoing parton composition that is compatible with the main physics process of the sample (i.e., VBF, ZH, or WH) are skipped.
- A momentum redistribution procedure is applied on the incoming and outgoing particles associated with H boson production so that the resultant topology features massless particles, which is what is required by the use of massless spinors in matrix elements. Denoting the momenta of the two final incoming or outgoing partons as p_1 and p_2, an intermediate four-momentum k is added to p_1 and subtracted from p_2 such that $|p_1 + k|^2 = |p_2 - k|^2 = 0$. This step is common to any matrix element computed using the MELA package. Because event-by-event reweighting is done through a ratio of matrix elements, which are invariant under any arbitrary boost of the event topology, and because factors coming from PDFs cancel in the ratio, the common boost of all particles does not affect reweighting as long as momentum conservation is maintained strictly, and is therefore adjusted arbitrarily.