Jet merging in gg→ZZ production

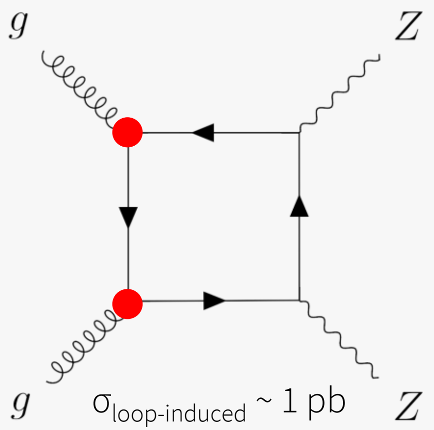
based on work [arxiv:2006.12860]()

Congqiao Li*, Ying An, Claude Charlot, Roberto Covarelli, **Zhe Guan**, Qiang Li

The 6th China LHC Physics Workshop

06 November, 2020

Introduction of loop-induced ZZ process

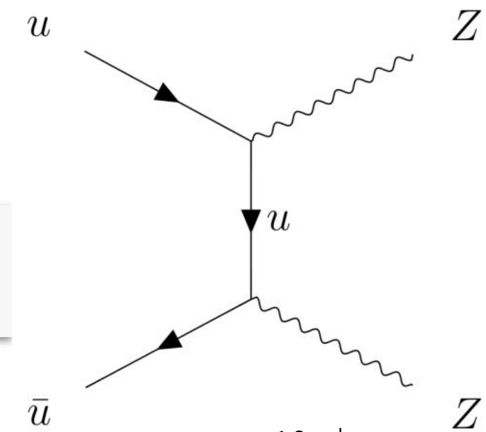


$\sigma_{\text{loop-induced}} \sim 1$ pb

$|\mathcal{M}_{\rm LI}|^2 \sim o(g_s^4) \sim o(\alpha_s^2)$

how much contribution?

loop-induced ZZ process can be enhanced due to large gluon flux at LHC



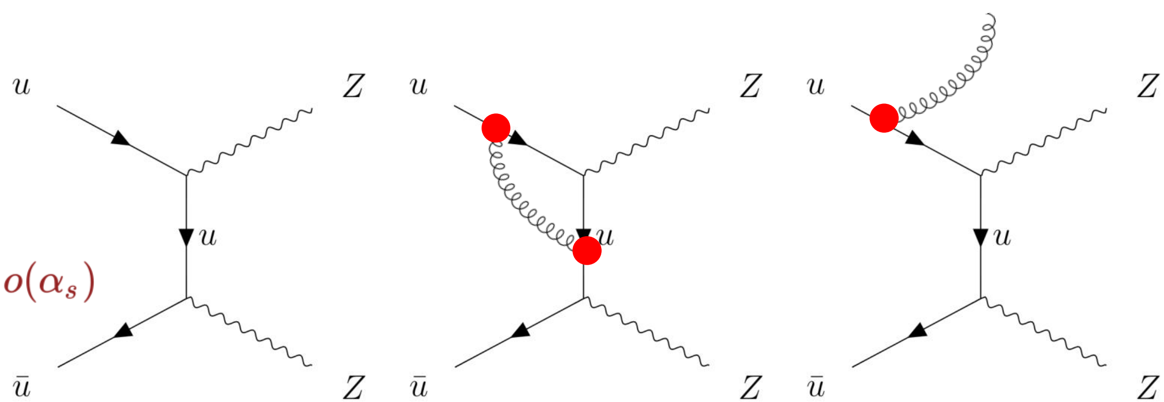
$\sigma_{\text{Born}} \sim 10$ pb

belongs to NNLO of Born-level ZZ process

$\sim o(g_s^2) \sim o(\alpha_s)$

$$|\mathcal{M}|^2 = |\mathcal{M}_{\rm LO}|^2 + 2\mathcal{M}_{\rm LO}^\dagger \mathcal{M}_{\rm virtual} + |\mathcal{M}_{\rm real}|^2$$

+ higher order



$\mathcal{M}_{\rm LO}$ $\mathcal{M}_{\rm virtual} \sim o(g_s^2)$ $\mathcal{M}_{\rm real} \sim o(g_s)$

Why is important?

- **Loop-induced ZZ and Born-level ZZ processes are important background sources in various analyses**
  - Off-shell and on-shell Higgs related measurements ($H \rightarrow ZZ$)
  - Standard model tests through ZZ channel (VBS ZZjj)
  - New physics search ($X \rightarrow ZZ$)...
  - Same situation in any relevant VV analyses
- **Precise simulation comes in two directions**
  - **higher-order calculation** ⇒ provide K-factor in analyses / simulate events at NLO precision
  - **multi-jet simulation** ⇒ provide a better description of jet phase space
- **Loop-induced diboson process**
  - LO contains a quark loop, thus brings growing complexity in both high-order & multi-jet calculation
  - We simulate the loop-induced ZZ process at LO with up to 2 jets
    - meaningful in the VBS ZZjj analysis, since the dijet variables (m_{jj}, $\Delta\eta_{jj}$, ...) are used to control the phase space

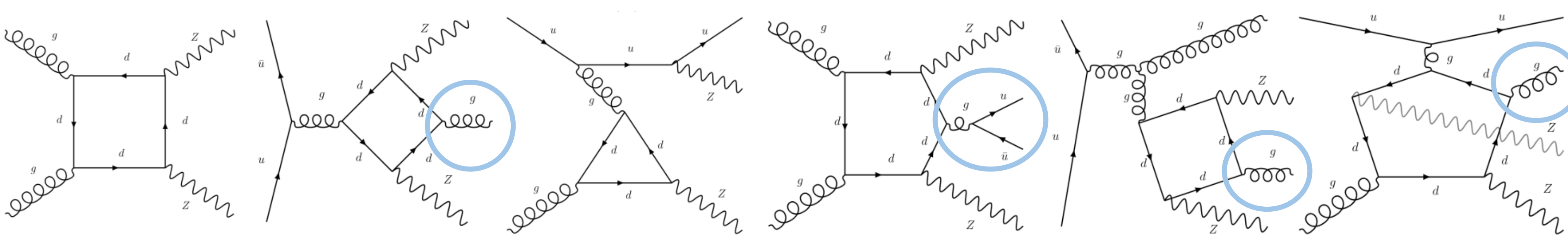
Simulation details

- MadGraph for matrix element simulation
  - N_F = 5 flavor, massive top contribution included
  - simulate ZZ final-state with 0, 1, 2 extra jets, requiring “no-Born” (see below)
  - ISR considered ⇒ include gg, qg, qq initial state
  - produce multi-jet process and match to Pythia parton showers via MLM matching scheme
  - apply a diagram filter to exclude the loop correction diagrams and (temporarily) exclude Higgs-mediated diagrams (more on the next slide)

```
generate g g > z z [noborn=QCD]
add process p p > z z j [noborn=QCD]
add process p p > z z j j [noborn=QCD]
```

Note: some 1-jet, 2-jet diagrams involve final-state jets directly emitted from the loop ⇒ only achievable from ME simulation

Example diagrams



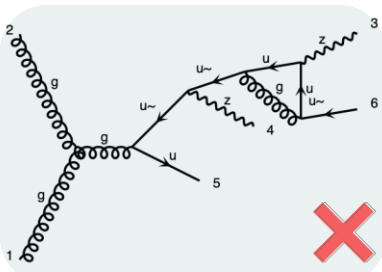
0-jet: 18 diagrams **1-jet: 756 diagrams** **2-jet: 23292 diagrams**

Diagram filter

inspired by discussions:
https://answers.launchpad.net/mg5amcnlo/+question/402723
(see backup for the patch to MadGraph)

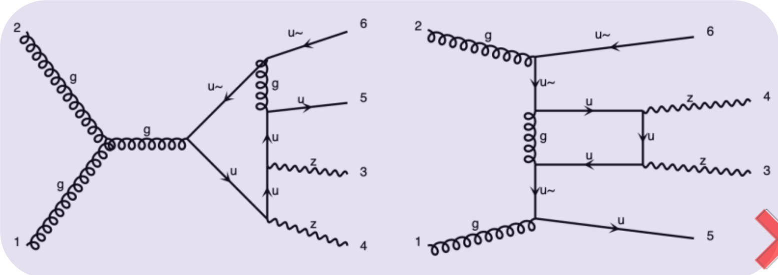
- MadGraph setting **`[noborn=QCD]`** can only identify one-loop diagrams but not loop-induced diagrams
- Some diagrams are loop corrections. If included, might bring unwanted divergence to calculation

- **Rule #1:** the loop in the diagram should not contain any gluon line
  - after removing the gluon line, the remaining diagram is a tree diagram with correct Feynman rules
  - the filtered diagram is a loop correction, including the vertex- or box-corrections

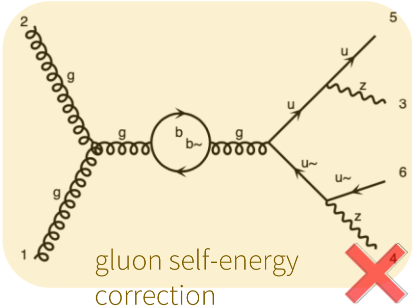


vertex correction

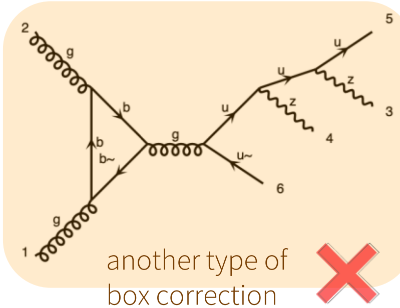
box correction



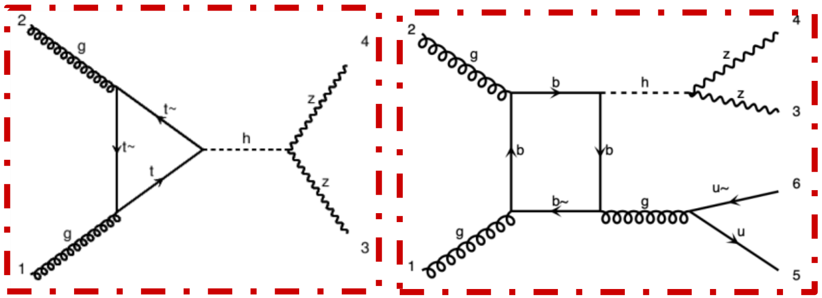
- **Rule #2:** the loop should attach to at least one Z/W/γ particle
  (as a validation: Rule#2 equivalent as “particles attached to the loop should not be all gluons” & “exclude Higgs-mediated diagrams”)



gluon self-energy correction



another type of box correction

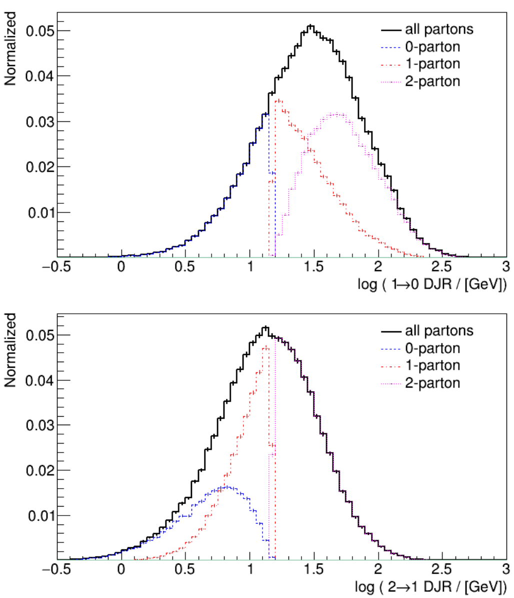


(**Higgs-mediated diagrams are temporarily excluded** to save computing time. Validation shows there is negligeable impact to jet kinematics, details in backup)

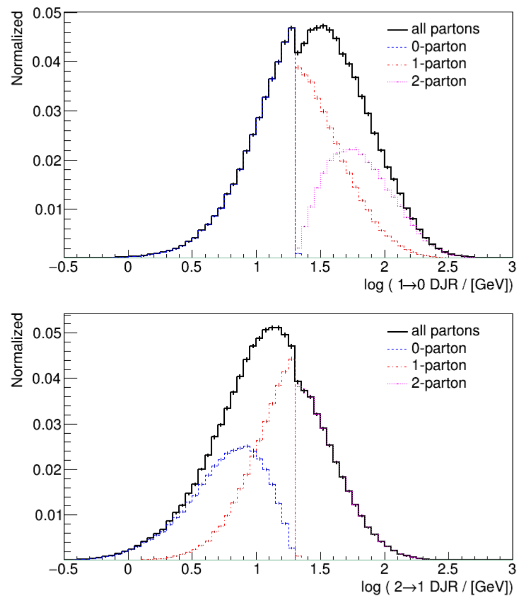
MLM matching to parton showers

- DJR plots validate the goodness of matching (more explanation in backup)
  - Determined by two thresholds: Q_{min}^{ME} (MadGraph parameter `qcut`) applied on LHE final state partons in matrix-element level, and Q_{min}^{jet} (Pythia parameter `QCUT`) applied on generator-level jets
  - matching found to be optimal in smaller threshold: `xqcut` = 5 GeV, `QCUT` = 15 GeV; compared to suggested scale choice `xqcut` = 10 GeV for single Z/W production at LHC

`xqcut` = **5** GeV, `QCUT` = **15** GeV



`xqcut` = **10** GeV, `QCUT` = **20** GeV



process	cross-section [pb]
0-jet	1.041 ± 0.0009
0,1,2-jet: xqcut=5	1.019 ± 0.012
0,1,2-jet: xqcut=10	0.584 ± 0.006

- The matched cross-section also found closer to 0-jet cross-section, compared to sub-optimal scale choice
- Further validation:

  the optimal scale choice holds for the similar 0,1-jet matched simulation

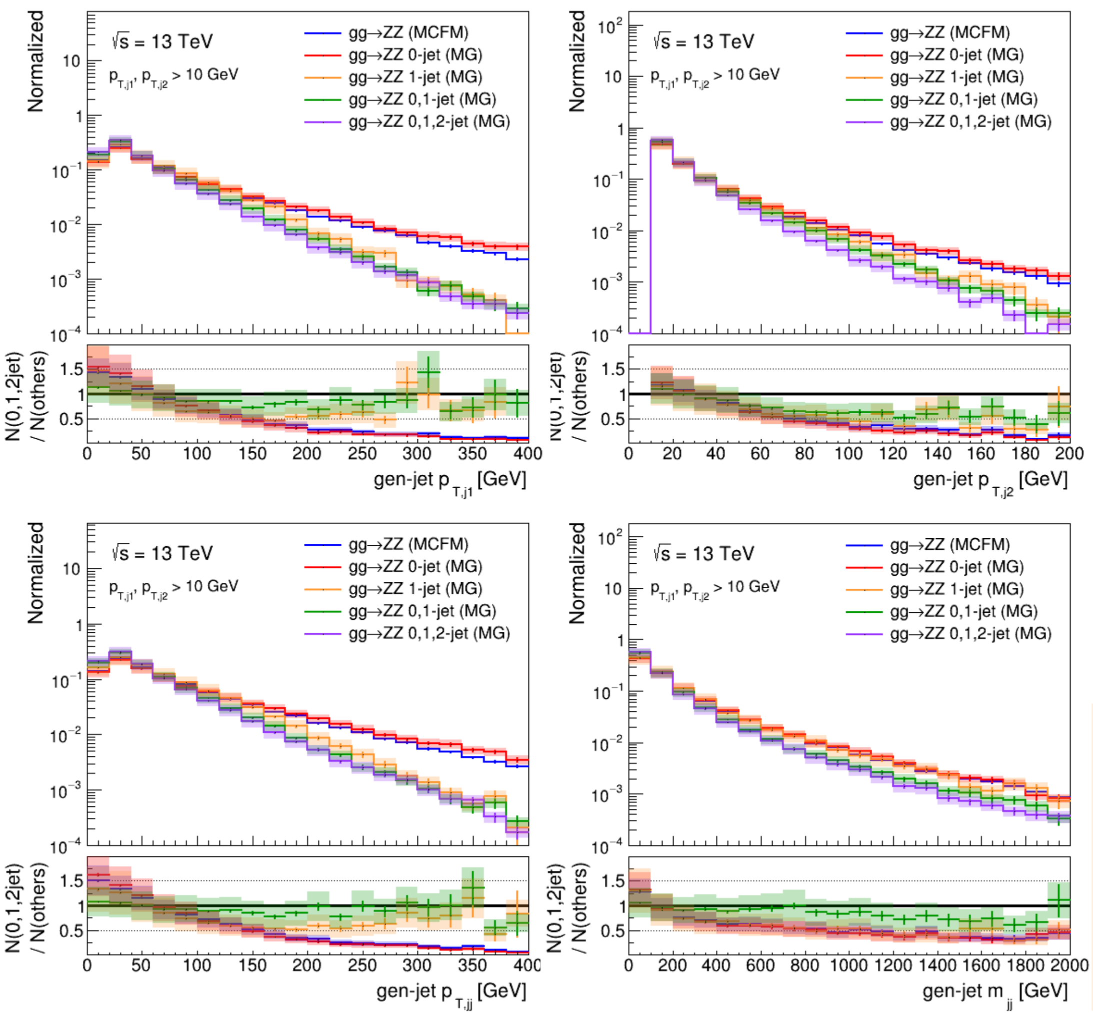
Computation performance

- Use MadGraph "gridpack" mode (`set gridpack true`) for a better handle of phase-space integration & event generation
- Take 24 hr in one core to collect all Feynman diagrams
- Significant increase in computation time:
  - phase-space integration (see table)
  - event generation:
    - 8 min/event for LHE event;
      net production rate 100 min/event considering
      an MLM matching rate of 8%
- Produce 140k events (after matching) in local clusters for the private study; then 960k events are produced on grid for analysis purpose

(*) Run on 2.4 GHz Intel Xeon E5-2680

sub-process	core-hour
0-jet	0.085
1-jet	10.9
2-jet	15300

Validation w.r.t. other ZZ processes



- Comparison over the MCFM ggZZ simulation; MadGraph ggZZ 0-jet; 1-jet; 0,1-jet matched; 0,1,2-jet matched simulation
  - all MadGraph simulation adopts the same definition of parameters, scales, and PDF (NNPDF 3.1)
  - **error bars** show stat. unce. (assuming data stat.), **shaded areas** show the combined PDF & scale uncertainties as syst. unce.
  - histogram normalised to 1

- Starting from 1-jet, **the jet pT & mass gradually turns softer** ⇒ a consequence of ME modeled jets
- 0,1 vs. 0,1,2-jet has **similar first jet kinematics** (as expected), and a **slight discrepancy in the second jet**

Physics impact in VBS analysis

- VBS ZZ(ℓℓℓ'ℓ')jj analysis employes dijet variables to define selections, hence it is crucial to **have a better description of dijet phase space**
  - ZZ 0,1,2-jet matched provides currently the best dijet description on the loop-induced contribution
- Comparison is made among the same samples, requiring a set of generator-level selections based on the **VBS topology**
  - select four gen-leptons, and determine two lepton-pairs as two Z candidates
  - impose on-shell Z selection: **$60 < m_{Z1,2} < 100$ GeV, $m_{ZZ} > 160$ GeV**
  - impose jet selection on leading and subl. jets: **$pT_{j1,2} > 30$ GeV, $m_{jj} > 100$ GeV**
  - further define a tighter VBS-enriched region: **$m_{jj} > 400$ GeV, $|\Delta\eta_{jj}| > 2.4$**

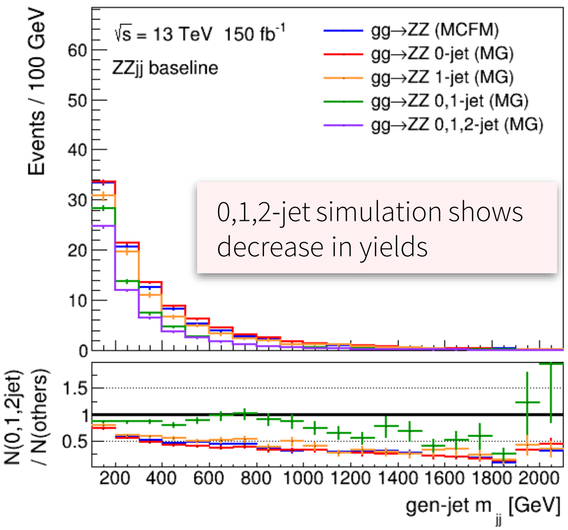
yields ± stat. unce (± syst. unce)

Process	$ZZjj$ baseline	VBS-enriched
MCFM 0-jet	98.0 ± 9.9	26.1 ± 5.1
MG 0-jet	$103.1 \pm 10.1 \pm 18.9$	$27.8 \pm 5.2 \pm 5.1$
MG 1-jet	$88.2 \pm 9.4 \pm 24.5$	$25.0 \pm 5.0 \pm 6.9$
MG 0, 1-jet	$64.3 \pm 8.0 \pm 12.7$	$13.5 \pm 3.6 \pm 2.8$
MG 0, 1, 2-jet	$55.4 \pm 7.4 \pm 12.5$	$11.5 \pm 3.3 \pm 2.9$
MG $pp \to ZZ$	$586.3 \pm 24.2 \pm 32.1$	$65.6 \pm 8.1 \pm 5.2$

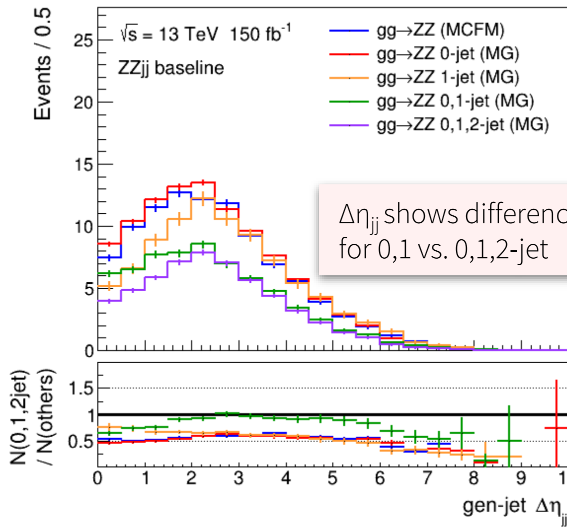
(*) all ggZZ samples normalized to same cross-section (MCFM) after the on-shell Z selection

- **event yields decrease** up to 43% for 0,1,2-jet simulation (34% for 0,1-jet), and to 56% moving to a tighter VBS-enriched region
- Born-level pp→ ZZ is also shown: **gg→ZZ proportion becomes larger** in VBS-enriched region, hence the yield decrease is **more relevant** for this phase-space

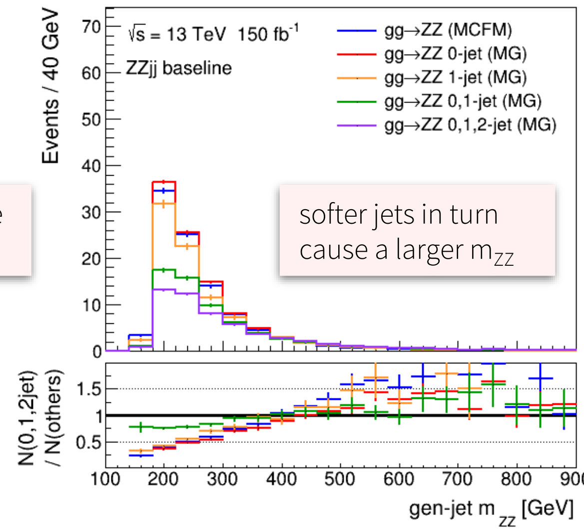
Physics impact in VBS analysis (II)



0,1,2-jet simulation shows decrease in yields



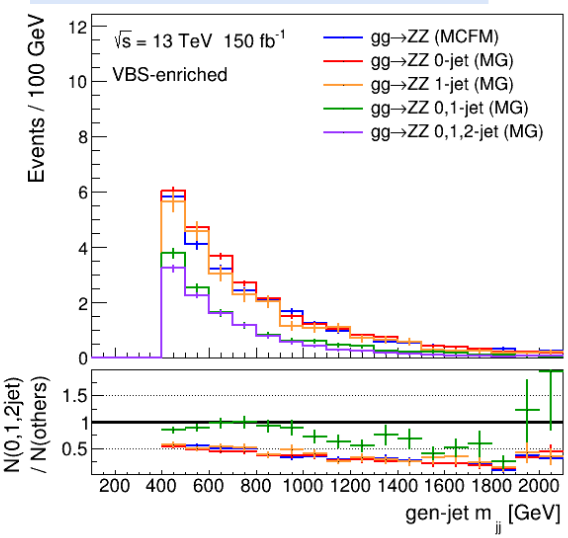
$\Delta\eta_{jj}$ shows difference for 0,1 vs. 0,1,2-jet

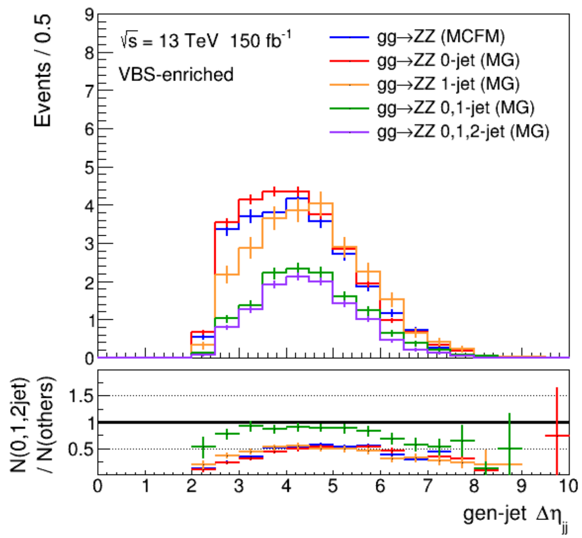


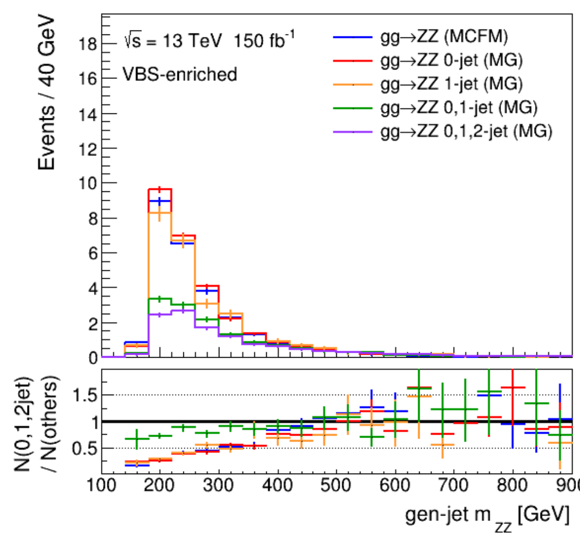
softer jets in turn cause a larger m_{ZZ}

Assuming 150 fb^{-1} stat.

a tighter VBS-enriched region







Summary & outlook

Summary

- Present the first loop-induced ZZ+0,1,2-jet matched simulation at LO, expected to have the best description of dijet phase-space
- Find the optimal matching scale to be smaller ⇒ a consequence of loop-induced jets
- Find the leading and subl. jet softer compare over multiple ggZZ processes, with 0 or 0/1 ME modeled jet
- Discuss the physics impact in VBS ZZjj analysis: event yields decrease up to 40% with the new ggZZ simulation. Bring attention to employ a better description in relevant analyses

Outlook

- Simulation has defects: contributions from the Higgs-mediated diagrams neglected; Z boson width and $Z\to \ell\ell$ spin correlation not simulated
  - **Higgs contribution can be included with an affordable cost**: ~2x in time
  - $Z\to \ell\ell$, if put into ME simulation, **brings a significant burden in integration**.
    Possibilities: ① first try 0,1-jet simulation with $Z\to\ell\ell$ included, then apply the 0,1 vs. 0,1,2-jet discrepancy; ② filter diagrams not directly from $Z\to \ell\ell$ (but should care for unitarity); ③ internal code optimisation? (long term)
- 0,1 vs. 0,1,2-jet differences can be better analysed, and **migrated to similar WW/WZ/Wγ/Zγ loop-induced diboson process in future analysis**, given the impressive time cost for 2-jet simulation

Backup

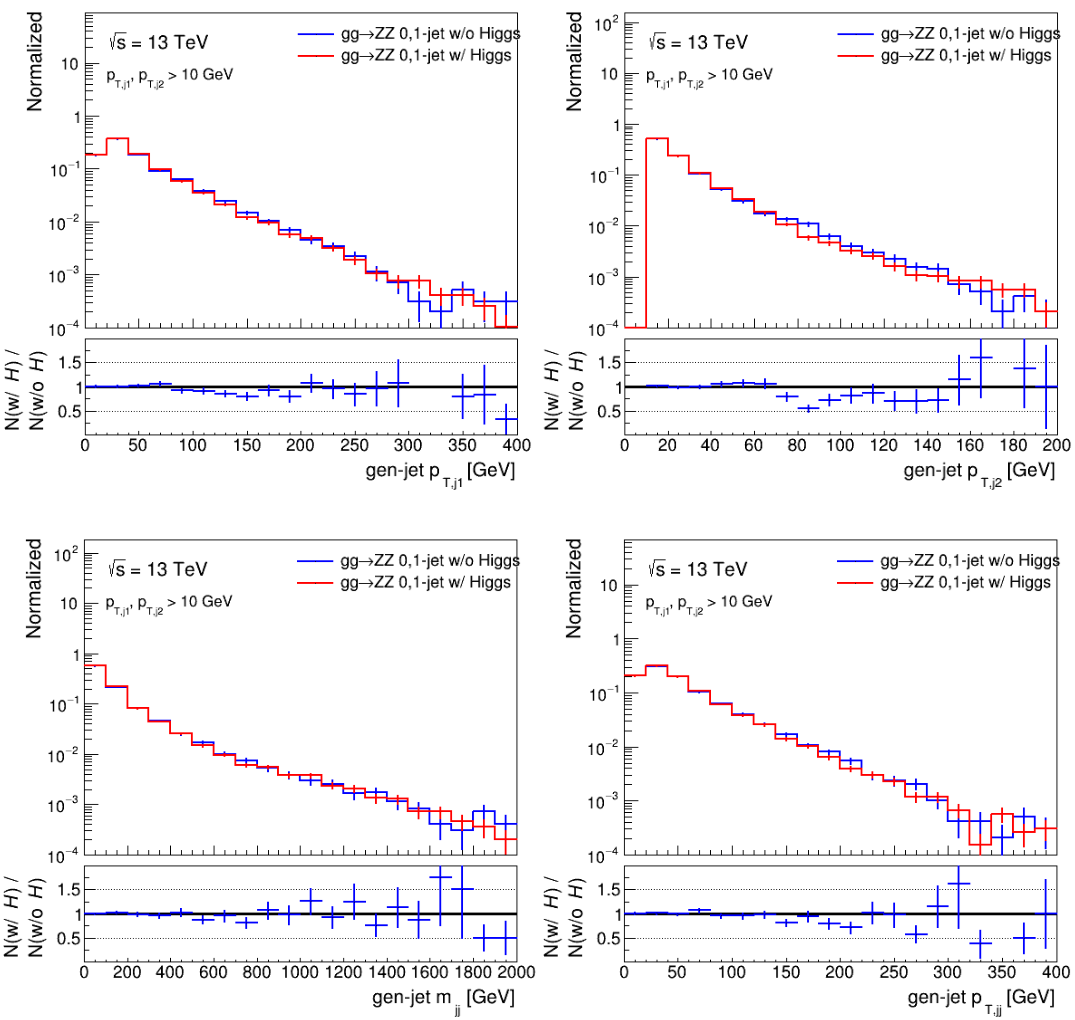
Details on the patch

loop-correction patch:

```
diff --git a/madgraph/loop/loop_diagram_generation.py b/madgraph/loop/loop_diagram_generation.py
--- a/madgraph/loop/loop_diagram_generation.py
+++ b/madgraph/loop/loop_diagram_generation.py
@@ -384,7 +384,7 @@ class LoopAmplitude(diagram_generation.Amplitude):
         # By default the user filter does nothing if filter is not set,
         # if you want to turn it on and edit it by hand, then set the
         # variable edit_filter_manually to True
-        edit_filter_manually = False
+        edit_filter_manually = True
         if not edit_filter_manually and filter in [None,'None']:
             return
         if isinstance(filter,str) and  filter.lower() == 'true':
@@ -415,6 +415,10 @@ class LoopAmplitude(diagram_generation.Amplitude):
                     raise InvalidCmd("The user-defined filter '%s' did not"%filter+
                                  " returned the following error:\n       > %s"%str(e))

+            if any([abs(pdg) not in range(1,7) for pdg in diag.get_loop_lines_pdgs()]) or \
+                 all([pdg in [21] for pdg in diag.get_pdgs_attached_to_loop(structs)]) or (25 in
diag.get_pdgs_attached_to_loop(structs)):
+                valid_diag = False
+
 #            if any([abs(pdg) not in range(1,7) for pdg in diag.get_loop_lines_pdgs()]):
 #                valid_diag = False
```

Higgs-mediated contribution



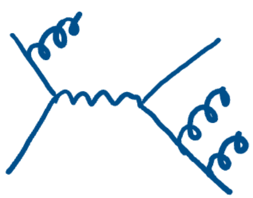
process	cross-section [pb]
w/ Higgs contribution	0.902 ± 0.005
w/o Higgs contribution	0.936 ± 0.009

- Compare a simpler MadGraph gg→ ZZ+0,1-jet simulation with and without Higgs-mediated contribution included
- Cross-sections shows ~3% difference in current phase-space (Z at pole mass and Higgs off-shell)
- Jet kinematics difference is neglectable

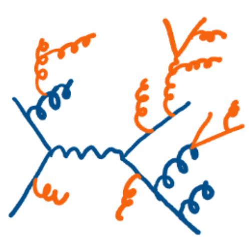
MLM matching -Durham k_T and ME-PS matching

- ➔ Durham k_T measures how soft/collinear a parton is splitted
  - ◆ low kT for soft/collinear emission
- ➔ ME handles the hard/split parton emissions, while PS handles the soft/collinear ones
- ➔ two parameters specified:
  - ◆ `xqcut`: reject soft/collinear (low Durham kT value) in ME level
  - ◆ **QCUT**: reject hard/split (high Durham kT value) in PS level

ME level



PS level



The default clustering scheme used (in MG/Sherpa/AlpGen) to determine the parton shower history is the Durham k_T scheme. For e^+e^-:

$$k_{Tij}^2 = 2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})$$

and for hadron collisions, the minimum of:

$$k_{Ti\mathrm{beam}}^2 = m_i^2 + p_{Ti}^2 = (E_i + p_{zi})(E_i - p_{zi})$$

and

$$k_{Tij}^2 = \max(m_i^2, m_2^2) + \min(p_{Ti}^2, p_{Tj}^2)R_{ij}$$

with

$$R_{ij} = 2[\cosh(y_i - y_j) - \cos(\phi_i - \phi_j)] \simeq (\Delta y)^2 + (\Delta\phi)^2$$

Find the smallest k_{Tij} (or $k_{Ti\mathrm{beam}}$), combine partons i and j (or i and the beam), and continue until you reach a 2 → 2 (or 2 → 1) scattering.

https://indico.cern.ch/event/757167/contributions/3176250/attachments/1733036/2801836/BeijingMGSchool2013-Johan_MLM_lecture.pdf

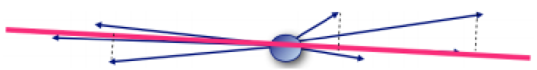
- Durham k_T clustering method:
- ⇒ to retreive the parton shower history of the events
- For every emission vertex, there is a Durham k_T
- min(k_T) must > xqcut

- Same Durham k_T clustering on PS level partons
- Clustering partons into k_T jets (note: not idiomatic jet)
  - ○ stopping line: all kT between jets < QCUT
- all jets after clustering should match with original partons, i.e. k_T(parton, jet) must < QCUT

MLM matching - differential jet rate (DJR) and MLM validation

➔ (Durham) differential jet rate:
- ◆ a variable measures event topology
- ◆ differential n-jet rate: DJR(n+1→n)
  - ● apply Durham clustering method until there are n-jet left
  - ● find mininal $k_T(i, j)$ for any (i, j) within n jet

http://edu.itp.phys.ethz.ch/hs10/ppp1/PPP1_8.pdf, p.167



➔ Validation of a good MLM matching:
- ◆ QCUT ~ (⅙~½)* hard scale
- ◆ Matched xsec (for X+0,1,... jets) shoul be close to unmatched xsec for the 0-jet sample
- ◆ Differential jet rate plots should be smooth
- ◆ When QCUT varies, the matched xsec / DJR should not varies significantly