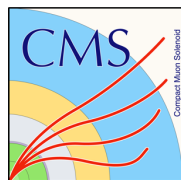


Background modelling in VBS analyses at the LHC

Philipp Pigard

November 16th 2017

VBSCan COST action meeting



Outline

Disclaimer: Experimentalist talking about theory, heavily CMS biased

- Background categories
- Prevalence of MC codes in recent exp. analyses
- Summary of MC codes for VBS
- Case study: Irreducible QCD backgrounds in ZZjj VBS
 - Improving precision and tackling statistics issues
 - Loop-induced backgrounds
- Conclusion and outlook



Talk centers on non-instrumental = irreducible backgrounds

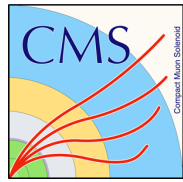
Reducible backgrounds

- Arise from instrumental effects
- At least one of the selected final state objects in the selected event is a fake, non-prompt, or charge is mis-measured
- Examples:
 - Jet fragments faking an electron (DY+jets in 4l)
 - Muon from B hadron decays
 - Charge flip for electrons (DY in ssWW)
- Detector and run dependent,
- Can be reduced by requiring more stringent quality criteria
- Usually estimated in a fully or at least partially data-driven way

Irreducible background

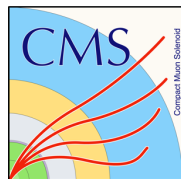
- Independent of specific detector
- The final state objects of the signal are selected in the event, but result from a non-signal scattering process
- Examples:
 - **Final state identical to signal, but mediated by strong instead of weak interaction**
 - Wrongly selected vector bosons (WZjj in ssWW)
- Must be handled by identifying regions of phase-space with reduced contributions, then weighing or removing events
- Modelling relies on SM predictions from MC

Here focus on irreducible backgrounds, particularly from strong interaction (QCD)



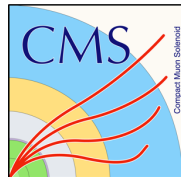
Background importance in recent VBS searches

Channel	Reducible	Irreducible
ssWW	Dominant, contributions from ttbar, WZ+jets	Highly suppressed because very few QCD diagrams satisfy charge requirements, $\sigma_B \approx \sigma_s$
ZZjj (Z γ jj)	Small due to kinematic constraints (on-shell Z), mostly DY+jets	<ul style="list-style-type: none"> ▪ No suppression due to neutral Z bosons (many initial+final state combinations), $\sigma_B \approx 20 \times \sigma_s$ ▪ Sizable contribution from loop-induced processes
WZjj (W γ jj)	Small due to kinematic constraints (on-shell Z), mostly DY+jets and ttbar+X	Large, many initial+final state combinations



Summary of MC used in LHC VBS analyses at 13 TeV

	ssWW 1709.05822, submitted to PRL	ZZjj PLB 774 (2017) 682
Leading QCD	MG5_AMC LO	MG5_AMC FxFx (NLO) with 0,1,2 born partons
Loop-induced QCD	(osWW) MCFM LO, jets pure ISR from PS	MCFM LO, jets pure ISR from PS
Other backgrounds	WZ+jets: MG5_AMC MLM (LO) with 0,1,2,3 born partons	ttZ and WWZ: MG5_AMC NLO



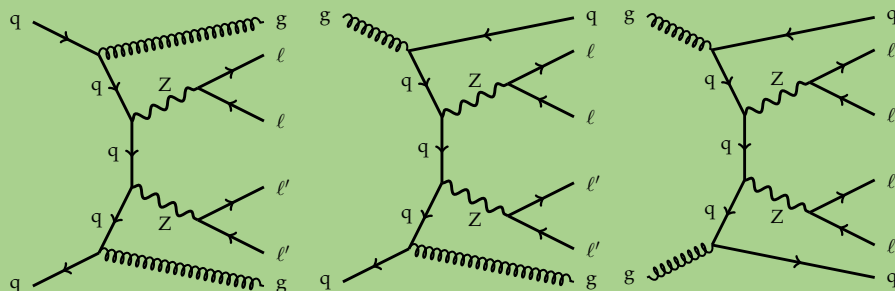
Selection of MC codes and relevant features for dominant QCD backgrounds

Code	Orders	Merging	Loop-induced	Comments
MG5_AMC	LO, NLO	LO (MLM), NLO (FxFx)	✓	
MATRIX	LO, NLO, NNLO	-	At LO	
MCFM	LO, NLO	×	At LO	
PHANTOM	LO	×	×	Parton shower via LHE
POWHEGBOX	NLO	×	LO, NLO (private)	
SHERPA	LO, NLO	LO, NLO, LO+NLO (MEPS@NLO)	LO, 0,1 jets merged	
VBFNLO	LO, NLO	×	At LO	Interface for Herwig PS

For a recent comparison of predictions from major codes see also
ATL-PHYS-PUB-2017-005

Monte Carlo sample development for ZZjj – considerations

A) QCD background dominates event yields



⇒ Accurate predictions are key, want NLO precision

B) Low signal yields demand MVA treatment

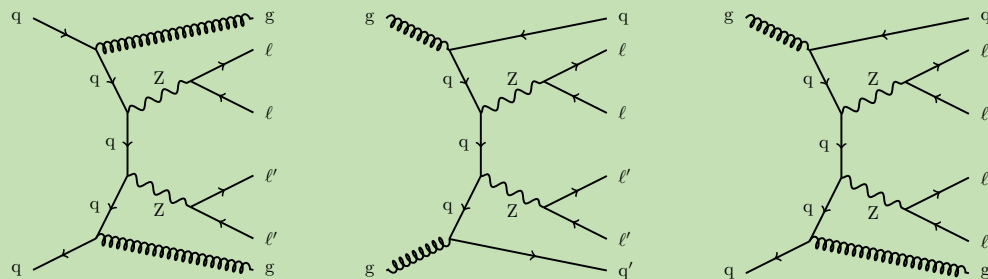
⇒ Need to ensure sufficient statistics

C) Constrain QCD yield from side-band

⇒ Need to be inclusive

MC predictions for the dominant irreducible QCD background in ZZjj

QCD-induced ($\alpha^2\alpha_{EW}^4$)



Starting point:

VBS topology covered by inclusive samples, but only one in 10^5 MC events ends up in ZZjj acceptance; not at NLO in QCD

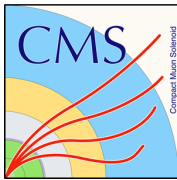
New predictions and validation:

For VBS ZZjj we generated and studied samples from:

- MG5_AMC at LO and MLM
- MG5_AMC at NLO merged (nominal)
- VBFNLO (in VBS phase space)

Challenges:

1. Need statistics in VBS phase, limited MC resources, up to $f = 0.35$ negative weights
2. High-multiplicity final state



Technical implementation using MadGraph package

Separate jet multiplicities

- Allows to generate statistics where needed

Process card I:

```
import model loop_sm-  
no_b_mass  
define p = p b b~  
define j = j b b~  
generate p p > z z [QCD]
```

Process card II:

```
import model loop_sm-  
no_b_mass  
define p = p b b~  
define j = j b b~  
generate p p > z z j [QCD]
```

Process card III:

```
import model loop_sm-  
no_b_mass  
define p = p b b~  
define j = j b b~  
generate p p > z z j j [QCD]
```

MadSpin card:

```
define l+ = e+ mu+  
define l- = e- mu-  
decay z > l+ l-  
decay z > l+ l-  
launch
```

Generating only on-shell bosons

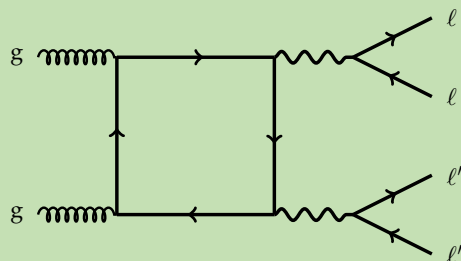
- drastically reduces number of diagrams
⇒ Enables ZZjj at NLO!
- Removes of out-of-acceptance events

Use MadSpin to decay bosons

- Re-creates decay correlations between the fermions
- Drop decays to τ (<1% of total yield)

Challenging territory: the loop-induced QCD background in ZZjj

Inclusive loop-induced ZZ, MCFM

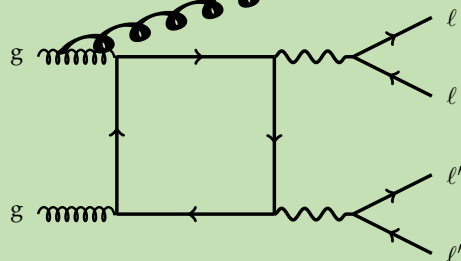


- NNLO contribution to ZZ production
- Used in Higgs and inclusive SM analyses
- Contributes about 10% of inclusive yield

Note: no outgoing partons from matrix element generator

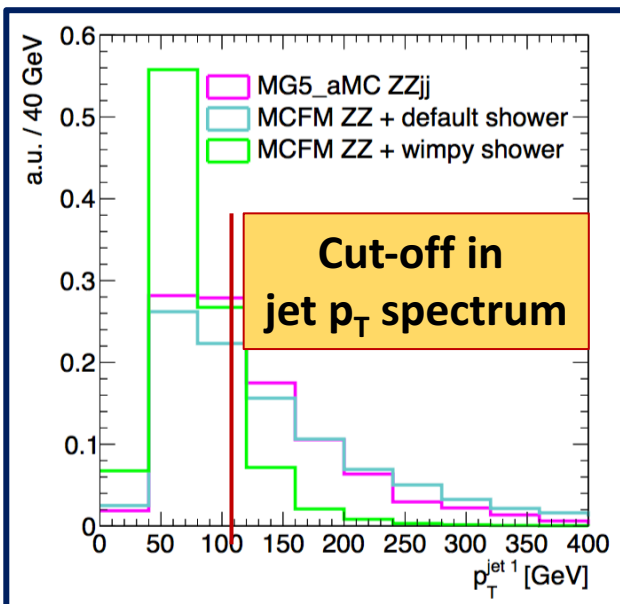
after
parton shower

ISR parton



- Final state is color singlet
⇒ Any final state parton after PS is initial state radiation
- PS does not change hard process
⇒ No radiation from loop

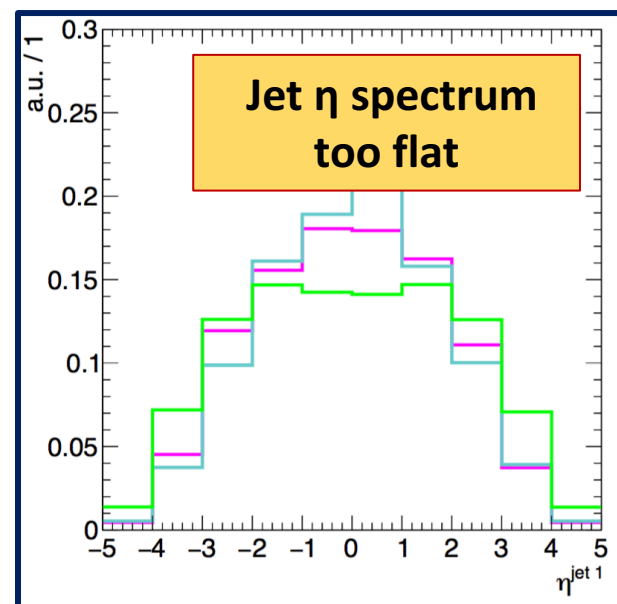
Parton shower configurations in Pythia8



Strange jet distributions ?!

Origin: ISR parton shower was modified to use *wimpy* shower

Done to improve p_T^{ZZ} modelling in HZZ4l analyses



What's a wimpy shower? \Rightarrow **SpaceShower:pTmaxMatch = 1**

From Pythia manual on ISR showers: *“always use the factorization scale [as limit for the hardest emission]. This should avoid double-counting, but may leave out some emissions that ought to have been simulated. ...”*

\Rightarrow Wimpy shower restricts hardest parton emission to the scale written in LHE

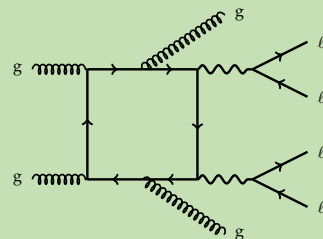
\Rightarrow For ZZ production, usually chosen as $\mu = m_{\text{ZZ}} / 2 = 200$ GeV

\Rightarrow No/few jets with p_T beyond 100 GeV

Need to adapt parton shower for VBS analysis

Matrix element prediction of the loop-induced production of ZZjj

What we want:
Loop-induced ZZjj ($\alpha^4 \alpha_{EW}^4$) from matrix element



+ many more

Based on [discussion with V. Hirschi on MG forum](#):

There are three types of amplitudes:

Tree diagrams

Loop-corrections to tree

Loop-induced diagrams

We want the finite contribution coming from

Loop-induced diagrams

x

Loop-induced diagrams

In MG5_aMC, `noborn=QCD` syntax removes

Tree diagrams

but includes terms from

Loop-corrections to tree

x

Loop-induced diagrams

which are divergent (need double real-emission contributions to cancel)



Technical implementation in MG5_aMC

Solutions:

- A. Exclude quark initial+final states ($g g \rightarrow z z g g$), misses $q \rightarrow qg$ splittings
- B. Better: remove diagrams that are loop-corrections of the tree amplitude

1

Specify process:

Process card:

```
generate p p > z z j j QED=2 QCD=99 [noborn=QCD]
```

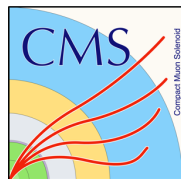
2

Remove loop-corrections, i.e., diagrams where no Z boson is attached to the loop:

```
[<MG_ROOT>/madgraph/loop/loop_diagram_generation.py]
...
if any([abs(pdg) not in range(1,7) for pdg in
diag.get_loop_lines_pdgs()])
or (23 not in diag.get_pdgs_attached_to_loop(structs)):
    valid_diag = False
...
```

3

Decay on-shell bosons. Caveat: no mass smearing nor spin correlations in MadSpin. Pythia won't consider branching ratios when restricting decays.

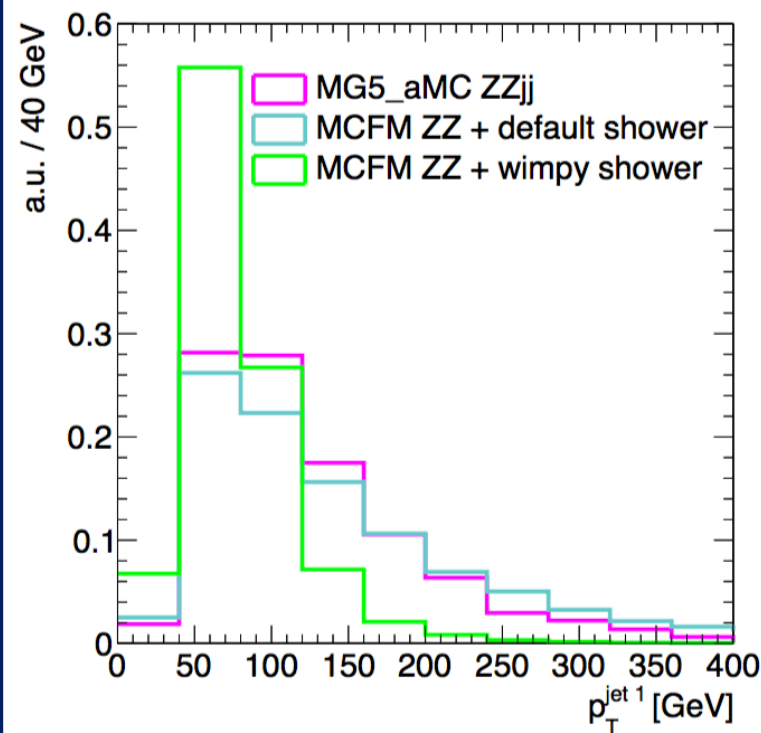


First results on loop-induced ZZjj

**Successfully generated 10k events,
at O(10 min) per event**

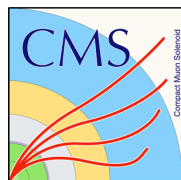
Used to validate modelling provided by nominal
MCFM+Pythia sample

⇒ Good agreement with default parton shower



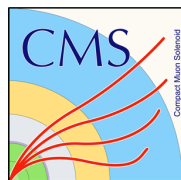
Experimented with MLM merged sample of 0, 1 multiplicities, with the aim to recover an inclusive sample. However, as an (approximately) unitarity-preserving algorithm, merging is unable to model the O(80%) increase in cross section from higher order corrections.

NLO corrections for loop-induced production of Z boson pair has been known for about a year [[JHEP 07\(2016\) 87](#)]. Based on POWHEG Box framework with Pythia as parton shower, AFAIK not yet used in analyses



Conclusions

- Many well-tested tools for on the market, NLO in QCD “standard”
- Comparing codes and predictions is critical to gain confidence in modelling
 - Should share such studies, akin to ATLAS simulation notes
- Nominal MC often determined by practical considerations (proficiency in using tools, existing computing infrastructure, statistics & runtimes)



Looking forward

- Uncertainties from MC predictions already (ssWW) or will soon (ZZjj, WZjj?) be one of the leading systematic uncertainty
 - Effort needed to use latest/greatest tools in analyses
- Tackling syst. Uncertainties will be challenge
 - Will need to include NLO EW and QCD corrections; How to define signal?
 - PDF uncertainties not negligible, requires better understanding (origin of kinematic dependencies)
- If we want to consider VBS measurements in the wider context of SM and Higgs measurements (additional constraints or combinations), will need to understand differences between predictions, ideally harmonize choice of MC
 - Will require adjusting statistics in relevant phase space (“biased sampling”)
- How to deal with polarizations in MC and analyses
- Many exciting developments for MC in loop-induced processes, to be deployed by experiments