Data driven W+Jets background estimation using the fake factor method in the WW->lvlv final state

Weimin Song

(on behalf of ATLAS Collaboration)

Weihai, China, Aug.15-19, 2016

Outline

- **Motivation**

- **Fake factor method**

- **Application**

- **Summary**
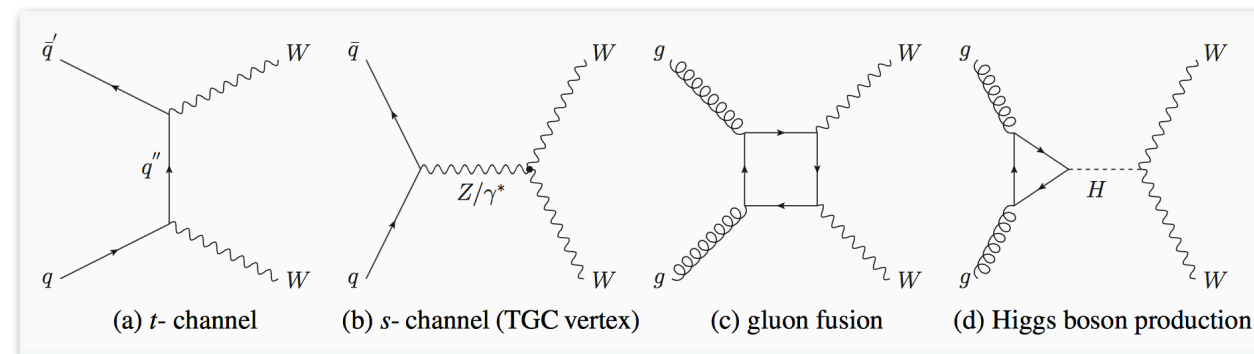
Why WW->lvlv final state?

We are looking for new physics beyond SM.

New physics?

But have not yet discovered any new physics.

W, the massive gauge boson is sensitive to the structure of SM and physics beyond it.

WW final state

(a) t- channel   (b) s- channel (TGC vertex)   (c) gluon fusion   (d) Higgs boson production

WW cross section

h->WW

H->WW

Leptonic decay mode is clean.

WW->lvlv

Leptonic branching ratios of W is larger than that of Z.

Why data driven?

Why we need a data driven method for W+Jets background?

1. The W+Jets contribution is at the same level as the h(125) signal;

2. The W+jets cross section is so large, we can not afford a full MC simulation; (next slide)
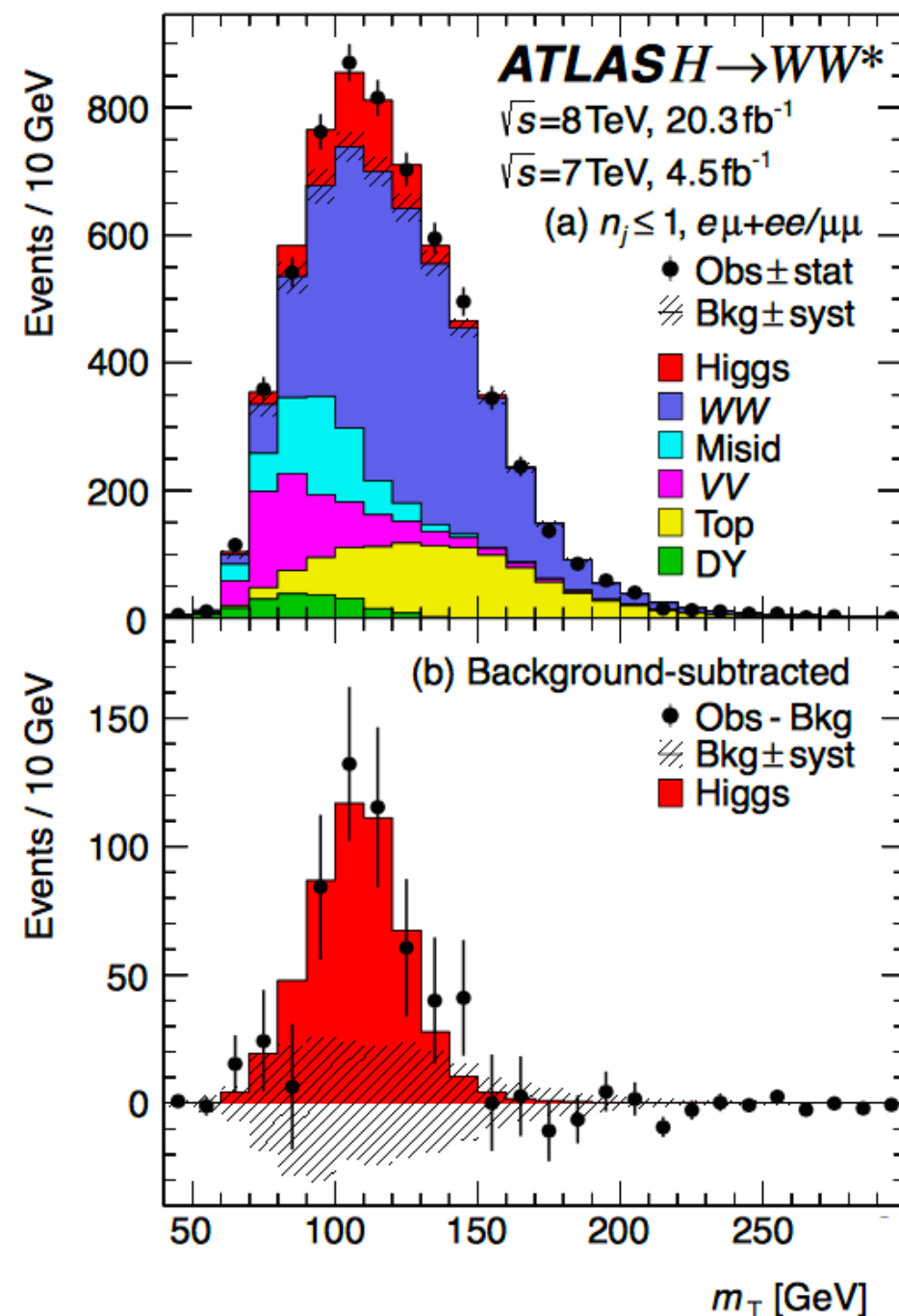
3. The fake mechanism is so complicated to simulate well by MC; (next next slide)

Data driven vs. MC based.

Fake factor method is one of the data-driven methods.

Why data driven?

W+Jets background is at the same level as the H(125) signal.

Why data driven?

Why we need a data driven method for W+Jets background?

1. The W+Jets contribution is at the same level as the h(125) signal;
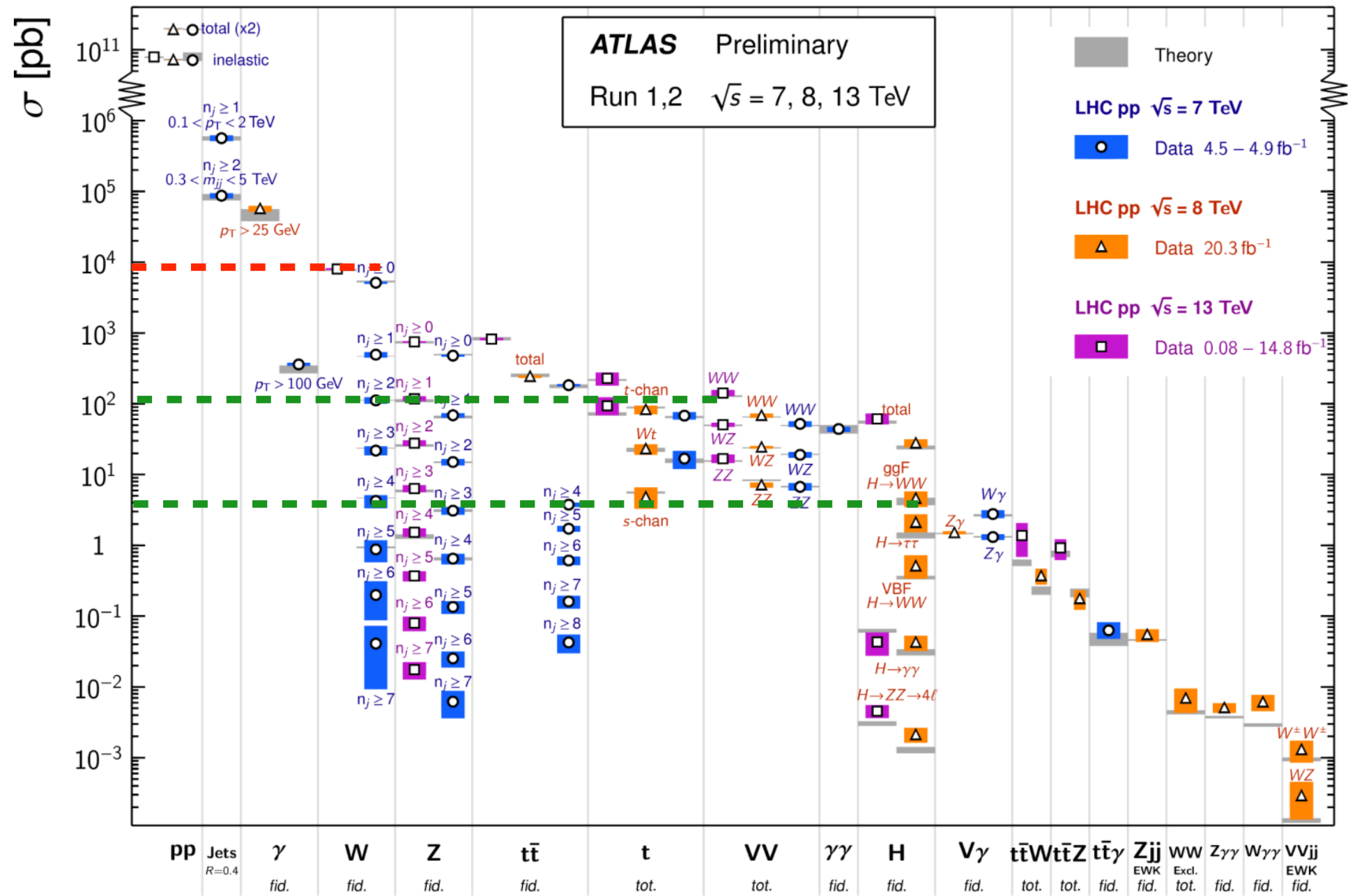
2. The W+jets cross section is so large, we can not afford a full MC simulation; (next slide)

3. The fake mechanism is so complicated to simulate well by MC; (next next slide)

Data driven VS. MC based.

Fake factor method is one of the data-driven methods.

Why data driven?

Cross section of W+Jets is several orders higher than the processes that we are interested in.

Why data driven?

Why we need a data driven method for W+Jets background?

1. The W+Jets contribution is at the same level as the h(125) signal;

2. The W+jets cross section is so large, we can not afford a full MC simulation; (next slide)

Data driven VS. MC based.

3. The fake mechanism is so complicated to simulate well by MC; (next next slide)

Fake factor method is one of the data-driven methods.

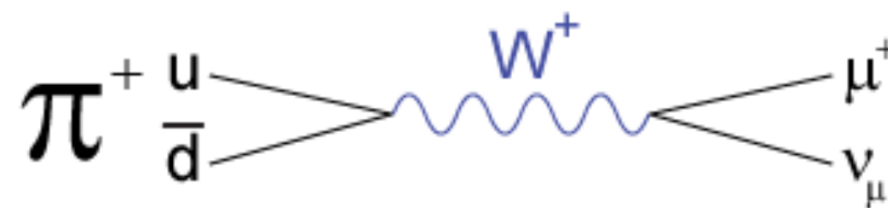# Why data driven?
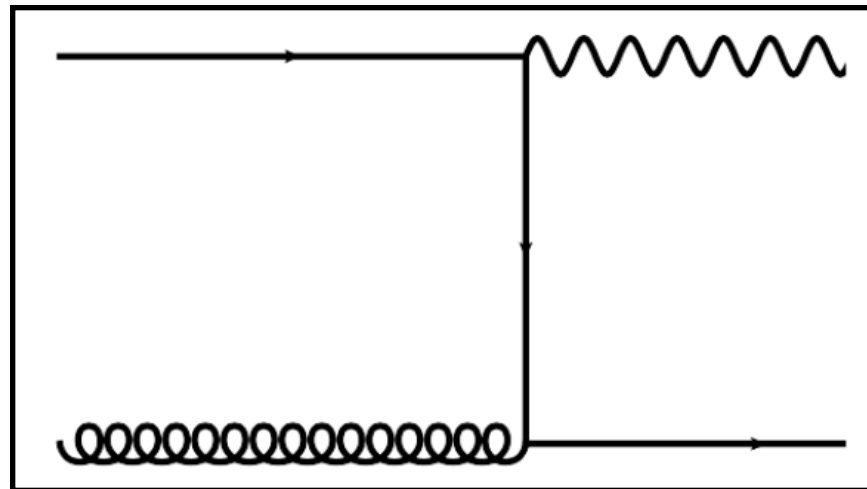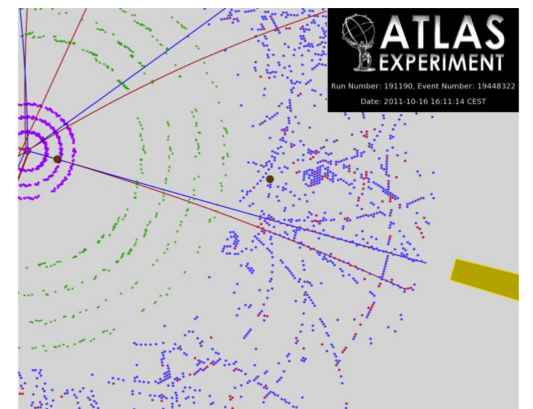
Leading order W+Jets production.
How can it contaminate WW->lνlν?
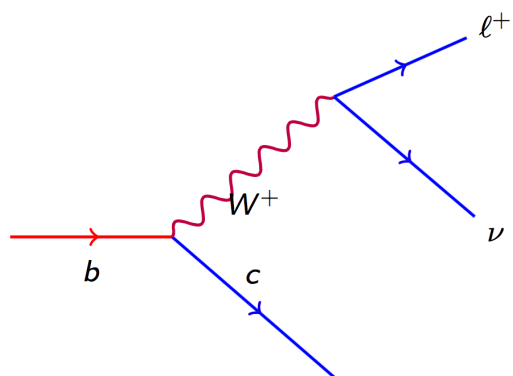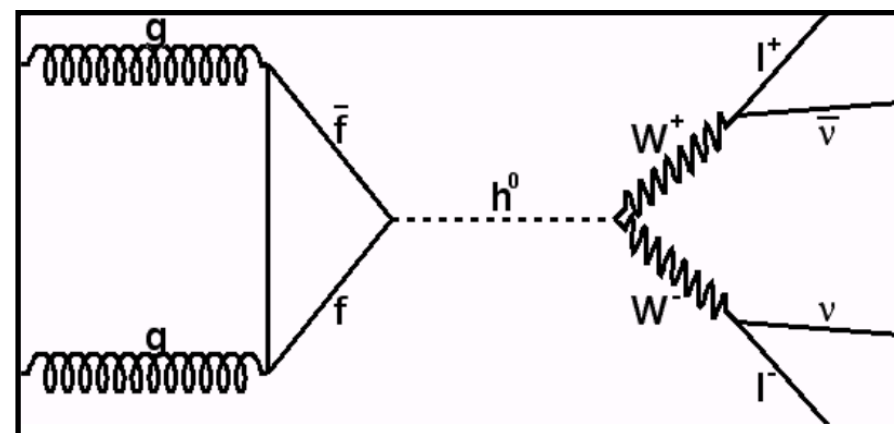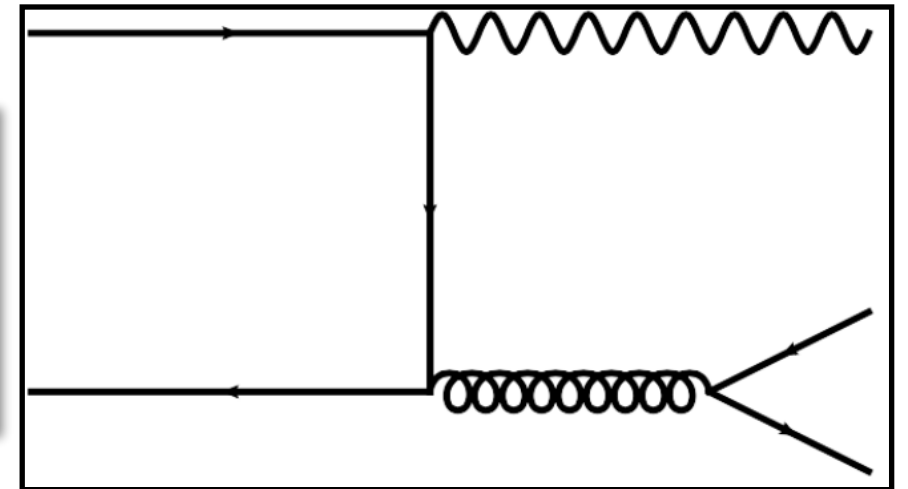
Heavy flavour decay.

Hadron decays in flight.

Gamma conversion

Fake factor method (1)

Two kinds of lepton selection criteria: identified (**id**), inversed identified (**anti-id**)

id: the optimised lepton criteria are required, such as small impact parameter, passing particle identification, good isolation from the other object and so on;
anti-di: inverse at least one of the requirement above;
Fake factor is defined as the ratio of **id** and **anti-id**.

Three kinds of region: signal region (**SR**), W+Jets control region (**CR**), fake control region

| W+Jets in SR (id+id) | ⟷ | W+Jets in CR (id+anti-di) | ★ | Fake factor (id/anti-di) |

Fake factor method (2)

W+Jets control region:

The selection is same as single region, except for one anti-id lepton is required. The background from top and di-boson processes is subtracted with MC simulation.
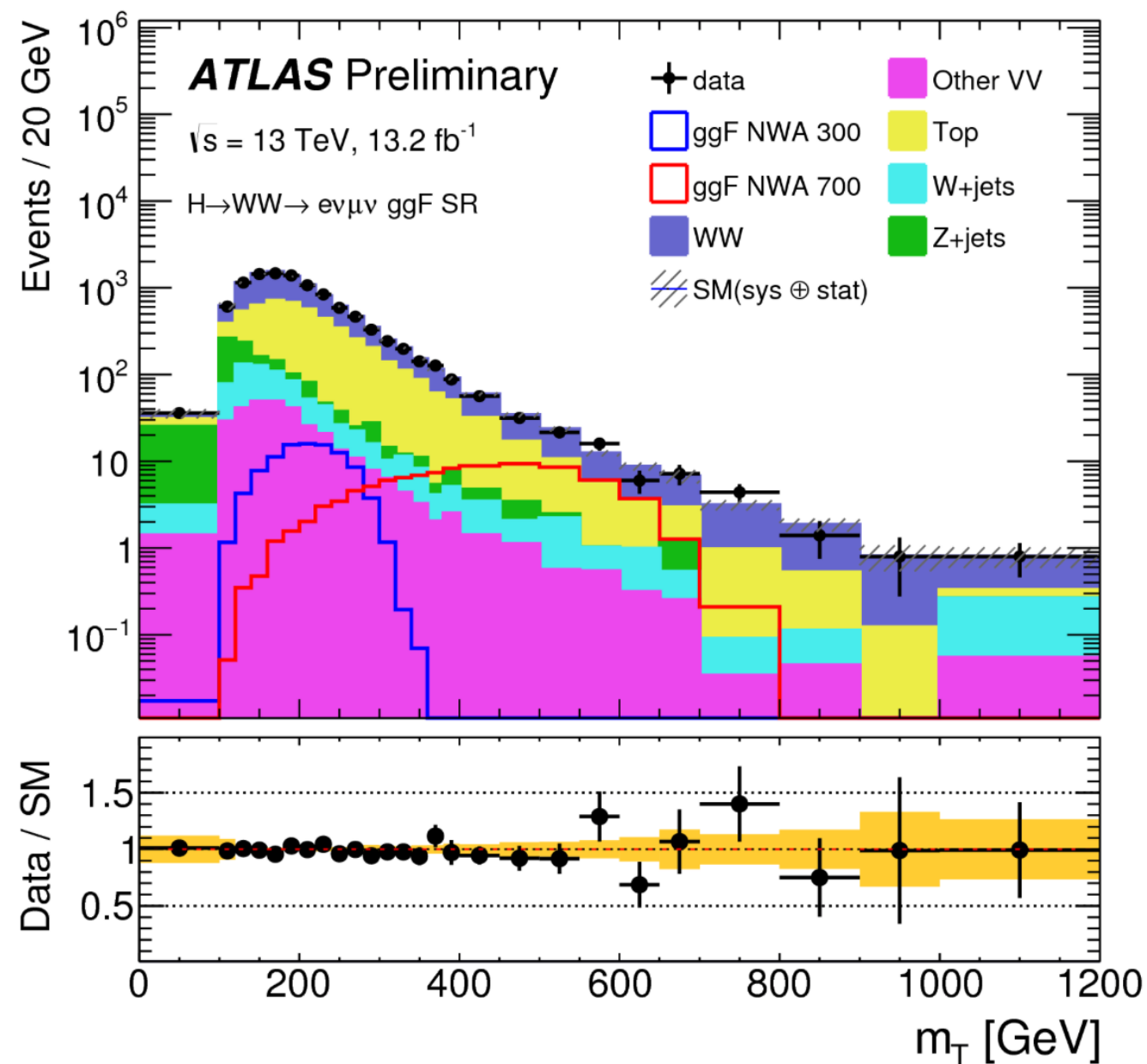
Fake factor:

Fake factor is measured with a jet-rich sample, such as di-Jets or Z+Jets process.

ATLAS $H \rightarrow WW*$
$\sqrt{s}$ = 8 TeV, 20.3fb^{-1}
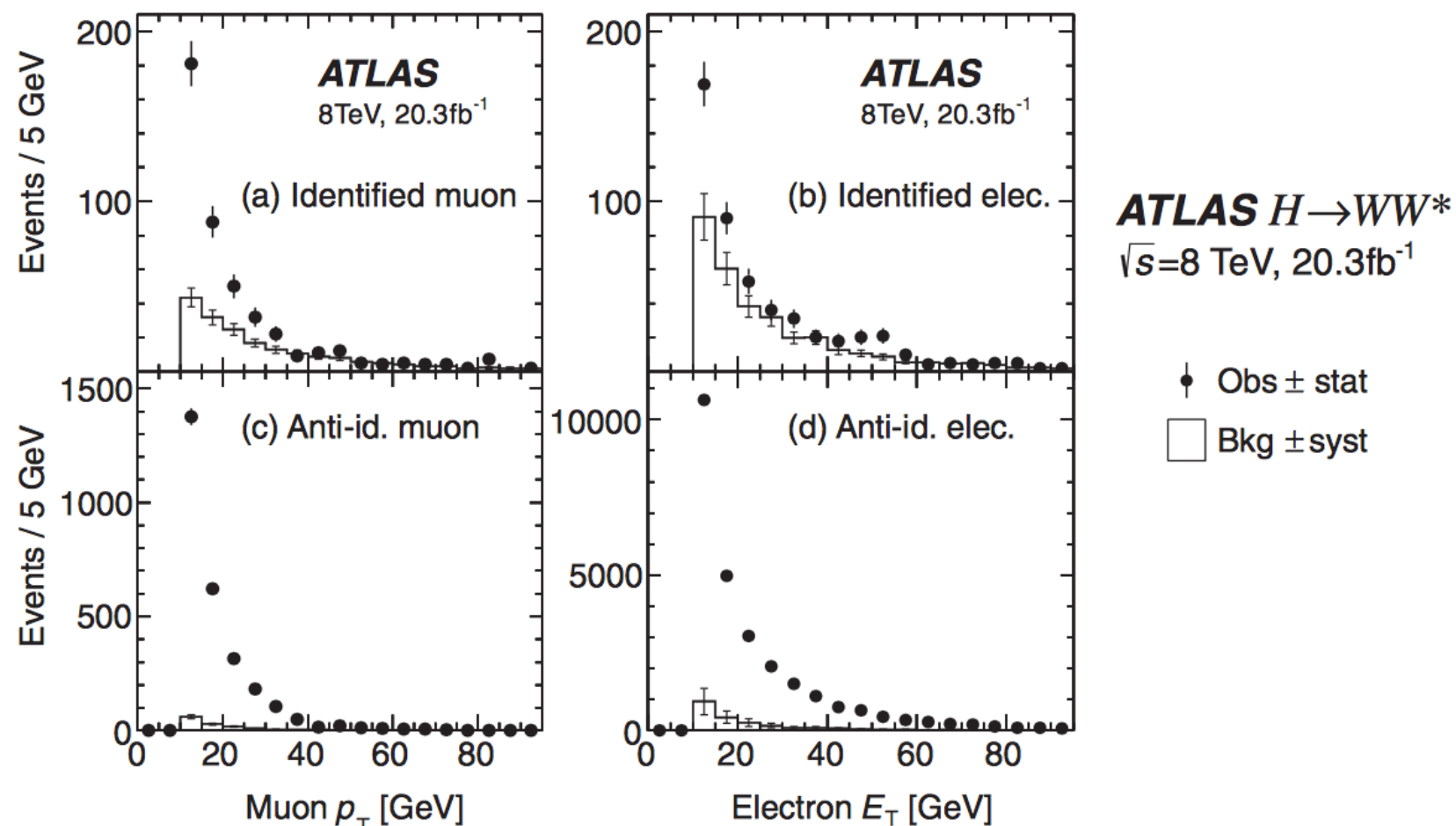
So far so good, but we have to consider the jet flavour component, the trigger bias......

Phys. Rev D. 92. 012006

Application at 13 TeV

With W+Jets control region and fake factor, we can estimate the W+Jets background in signal region.

Search for a high-mass Higgs in W pair. Details can be found in Yongke's talk.
The light-blue component is from data driven W+Jets.

Summary

Background from W+Jets to WW->lvlv is estimated in a data-driven way: the fake-factor method;

The result can be used in all the analyses with WW->lvlv final state. The method can also be used in analyses with related final states.

Thanks!