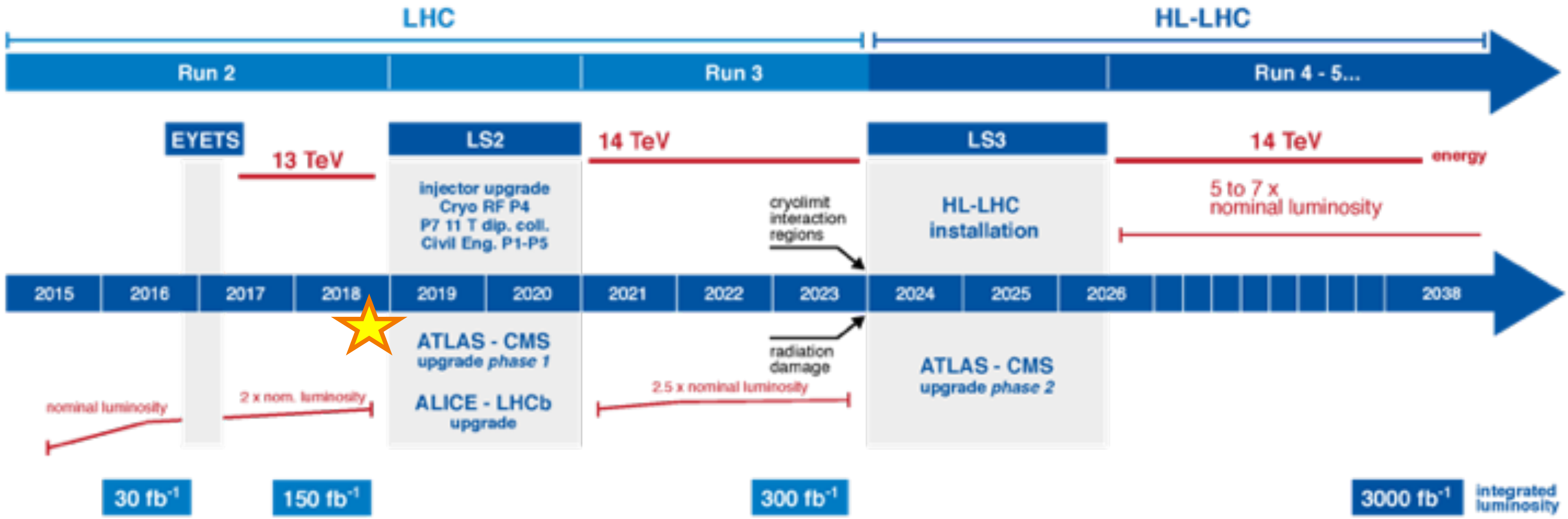
Theory uncertainties at HL-LHC

Osamu Jinnouchi
(TokyoTech)

Beyond the BSM (The 4th Kavli IPMU-Durham IPPP-KEK-KIAS workshop)
2018 Oct 1 - 4

HL-LHC plan & statistical impact



	13 TeV 2016 data	Run2 (– 2018)	Run3 (2021 – 23)	Run4,5... (2026 – 38)
L ($\times 10^{34} s^{-1} cm^{-2}$)	1.5	2.0	2 – 2.5	5 – 7
Int. Lumi (fb^{-1})	36	> 150	> 300	3000 – 4000

Most of the current papers

x 5 stat. available in next spring

x 10 stat. full LHC

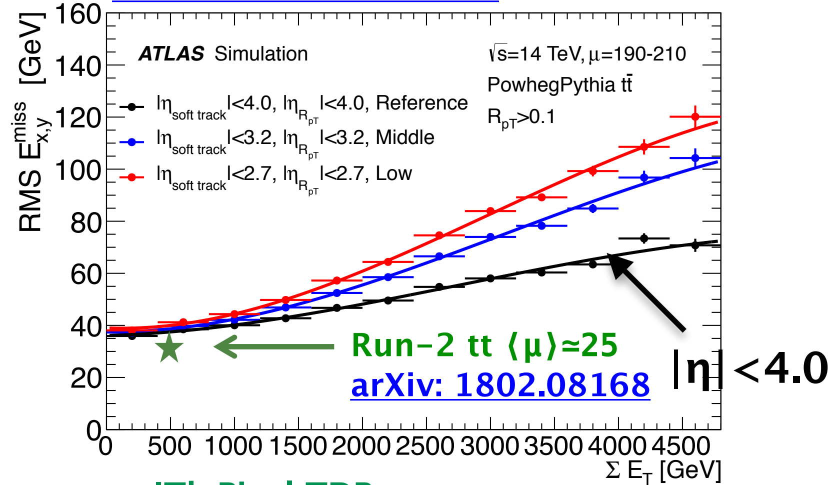
x 100 stat. full HL-LHC

Performance of the key objects

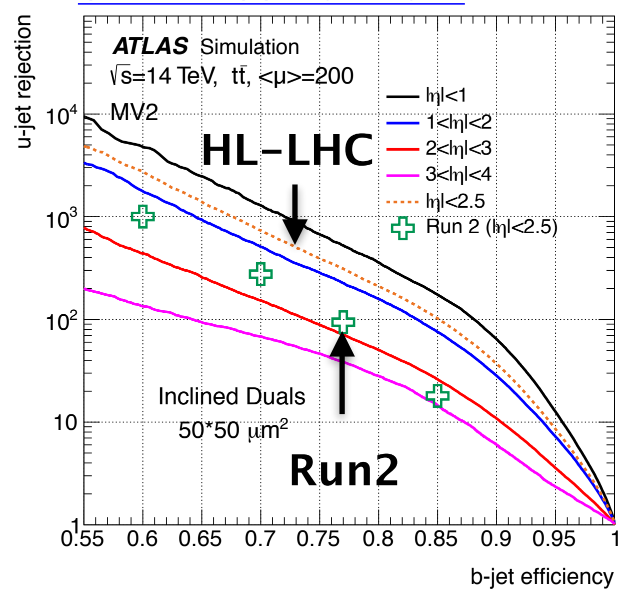
- Pile-up <μ>(2018) ≃50 → <μ>(HL-LHC)≃>200
- E_T^{miss}
  - threshold tuning, algorithm improvement
  - keep the performance on tracking
- b-tagging:
  - owing to the Inner tracker replacement, b-tagging keeps its performance, even better than Run-2
- Trigger
  - increase the overall budget 10kHz (Run-2: 1kHz)
  - improve the BG rejection new hard/software to cope with high rate while keeping the low threshold

major deterioration in performance not foreseen in HL-LHC

Phase-II upgrade scoping document
CERN-LHCC-2015-020



ITk Pixel TDR
CERN-LHCC-2017-021



Introduction

- HL-LHC
  - with 100 times statistics, things look different
    - needs for even higher order corrections
    - can discuss on the detailed shape of the distribution tail
    - can deal with very rare events
  - uncertainties not a big concern so far, become major contributions
- **"theory uncertainties"** could be the one! we should start thinking about reducing them, with the help of huge stats from LHC

disclaimer : I consulted with the following documents

ATL-PHYS-PUB-2018-010: impact of theory uncertainties related to Higgs boson production in the $H \to ZZ^* \to 4l$ and the ttH channels with the ATLAS detector at the HL-LHC

"QCD issues in searches for New Physics with the ATLAS detector" Mattias Saimpert, QCD@LHC2018 presentation

JPS 2018 fall : Yuji Yamazaki (Kobe U.) review talk

Important Physics topics at HL-LHC

[1] Higgs physics at HL-LHC

- Understanding of the Higgs potential
- Discovery & measurement of the **Higgs trilinear self-coupling** to explore the shape of the Higgs potential
- promising channels:

$$HH \rightarrow b\bar{b} + \bar{b}b,\ b\bar{b} + \gamma\gamma,\ b\bar{b} + \tau^{+}\tau^{-}$$

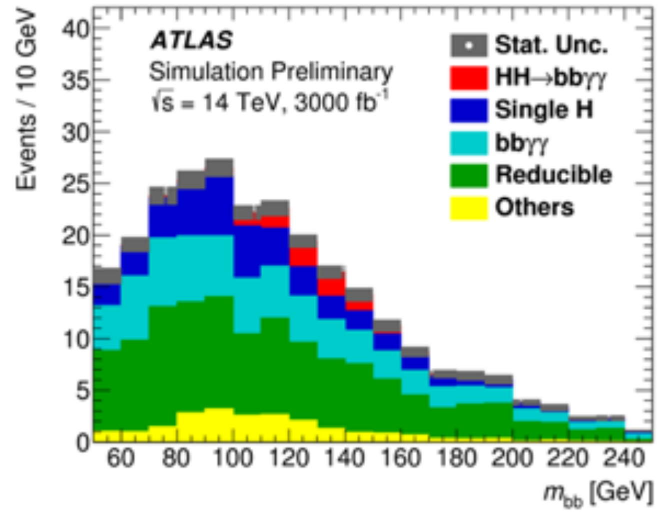
- Extremely **rare**, very challenging
- Yukawa-couplings **for lighter (2nd) generations** need to be covered

[2] Searches of the BSM

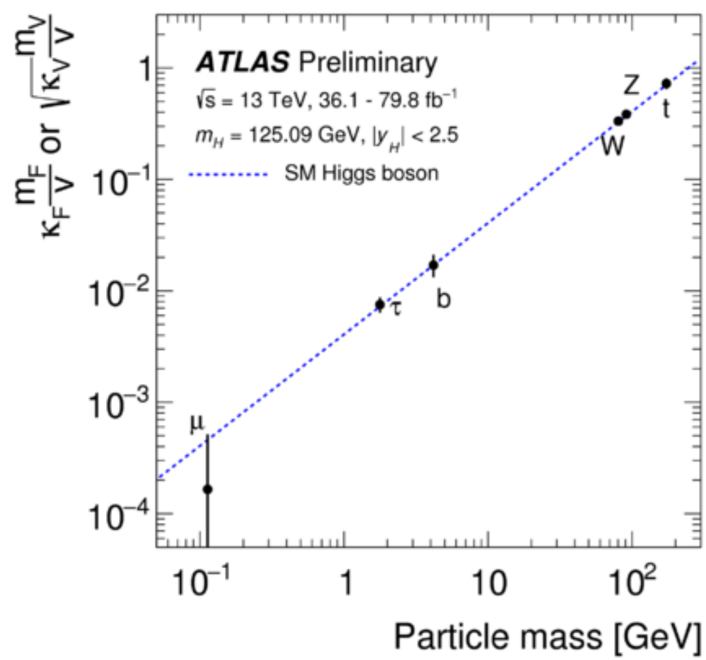
- Precisions on the background & signal estimations are vital
- Severe environment at HL-LHC

Theoretical uncertainties are the key for the success of HL-LHC

ATL-PHYS-PUB-2017-001



ATL-CONF-2018-031



[1] Theory Uncertainties in Higgs physics

Theory uncertainties

- Normalization and modeling
  - both for signal & backgrounds

Ex) Higgs related analysis

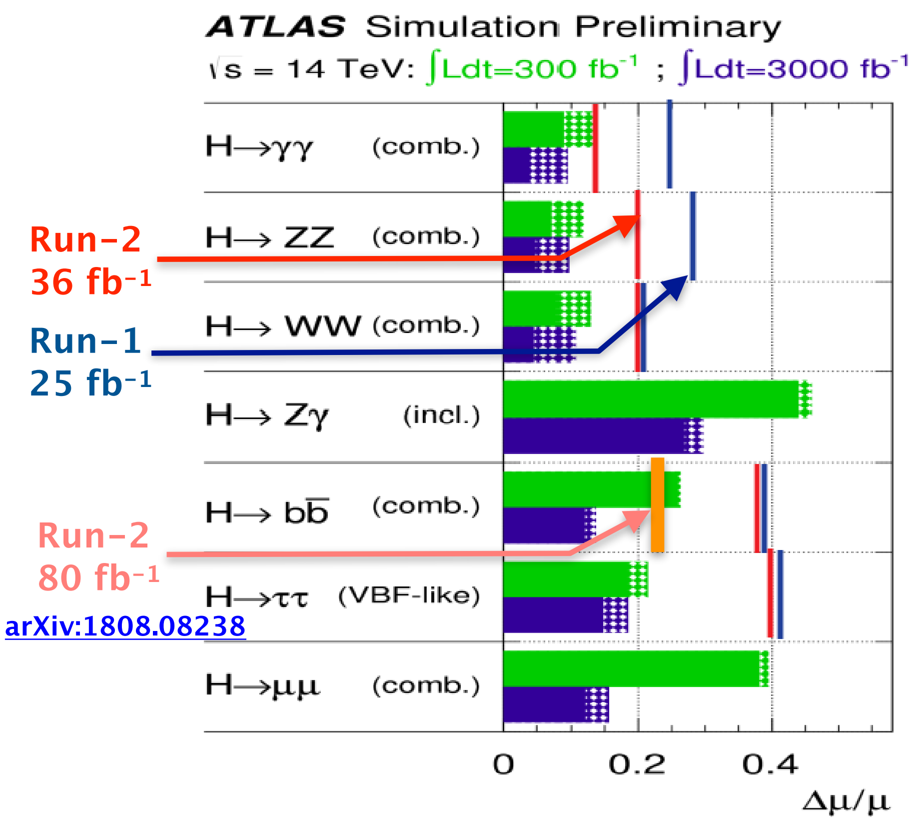
topic	channel	reference	dominant theoretical systematics
BSM	A/H tautau	JHEP 01 (2018) 055	Tau fake estimates, embedding of Z
Combination	Couplings	ATLAS-CONF-2018-031	Signal modelings
DiBoson	hWW	arXiv: 1808.09054	WW modeling
yukawa	VHbb	arXiv:1808.08238	V+jets modeling
yukawa	Htautau	ATLAS-CONF-2018-021	H pT modeling
HH	4b	arXiv: 1804.06174	multi-jet shape
μ yukawa	Hmumu	PRL119(2017)0418002	Drell-yan modeling
top yukawa	ttH	PRD97 (2018) 072003	tt+V modeling

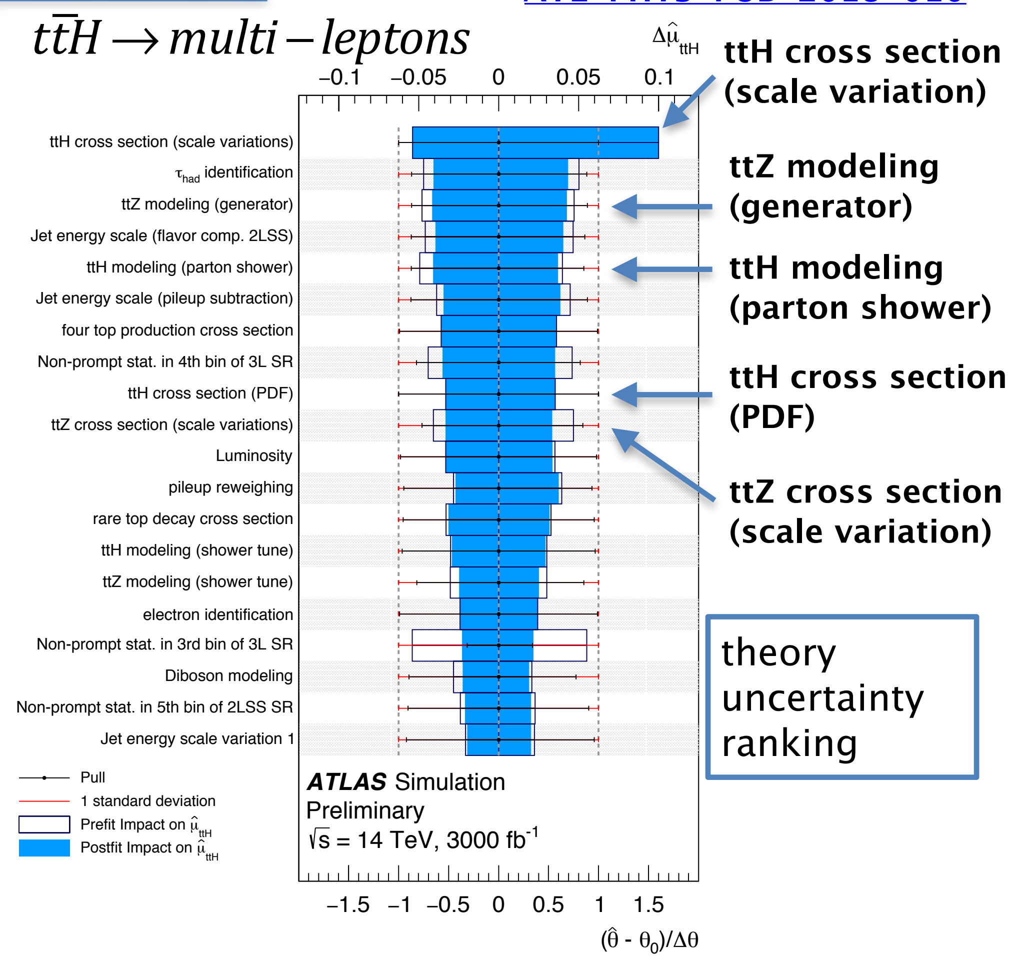
- For the HL-LHC physics projections, **the 1/2 of Run-2 estimation is used**
  - improvements in higher-order corrections, resummation
  - inputs from experiments would improve the modeling
  - PDF reduction expected with the help from theory community (?)

Uncertainties in Yukawa coupling measurements

ATL-PHYS-PUB-2018-010

Projection at 2014 (ATLAS-PHYS-PUB-2014-016)





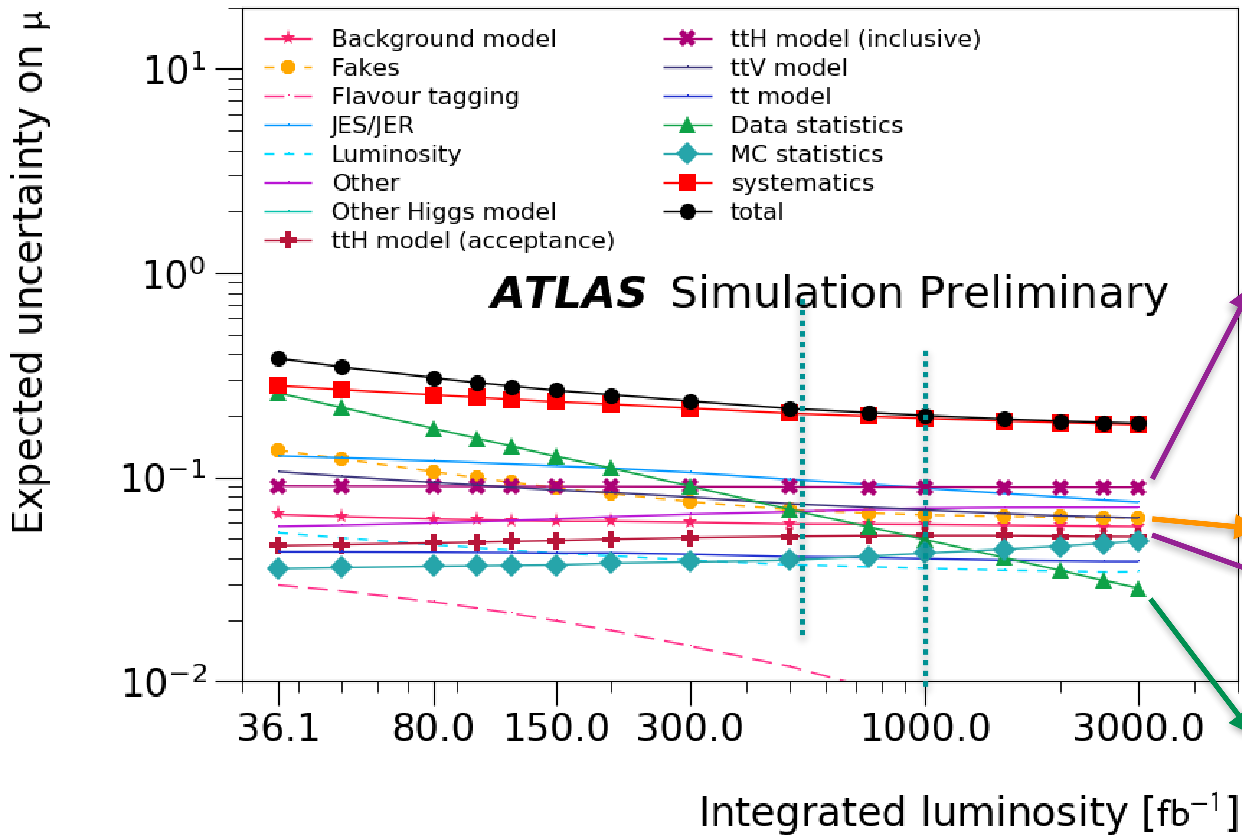
- Due to unexpectedly rapid progress in analyses, e.g. multi-variates, optimizations in analysis/performance, or conservative prospects (?) in 2014, some of them are already reached by the current precisions
- More accurate treatments of systematics will be available, after we obtain good amount of statistics from HL-LHC

σ_{sys} projection in ttH(→multi-leptons)

- signal strength measurement
- σ_{theory} has direct impact on the result

$$\mu = \frac{\sigma_{signal}}{\sigma_{theory}} \propto y_f$$

ATL-PHYS-PUB-2018-010



ttH model (inclusive) affects σ(theory)
- missing HO calc.
- αs
- PDF

fake objects

ttH model (acceptance) affects σ(signal)
- parton shower model

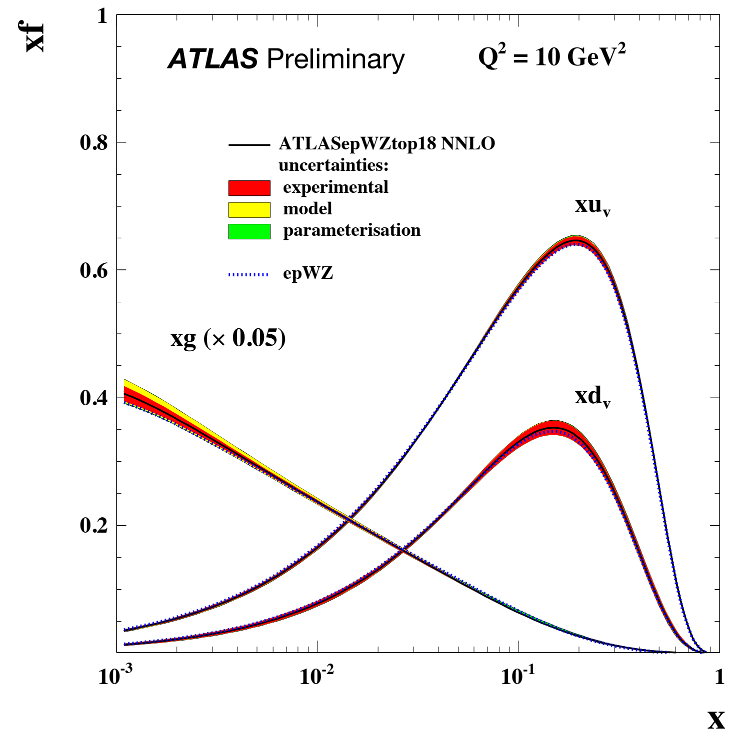
Stats no longer an issue at HL-LHC

LHC PDFs & towards HL-LHC

LHC measurements for PDF fits

Process	Sensitivity to PDFs
W mass measurement	Valence quarks
W and Z production (differential)	Quark flavour separation
W+c production, W and Z (differential)	Strange quark
Drell-Yan (DY): high invariant mass	Sea quarks, high-x, photon PDF
Drell-Yan (DY): low invariant mass	Low-x, resummation effects
W,Z + jets	Gluon medium-x
Inclusive jet and di-jet production	Gluon and $\alpha_s(M_Z)$
Direct photon	Gluon medium-, high-x
$t\bar{t}$, single top production	Gluon and $\alpha_s(M_Z)$

F. Giuli (Oxford)



In HL-LHC, high-x PDF are important

- **high-x gluon PDF**
  - jet and top pair production data are inputs
  - NNLO+NNLL for tt
- **high-x quark PDF**
  - high mass DY: NNLO pQCD+NLO EW

ATL-PHYS-PUB-2018-017

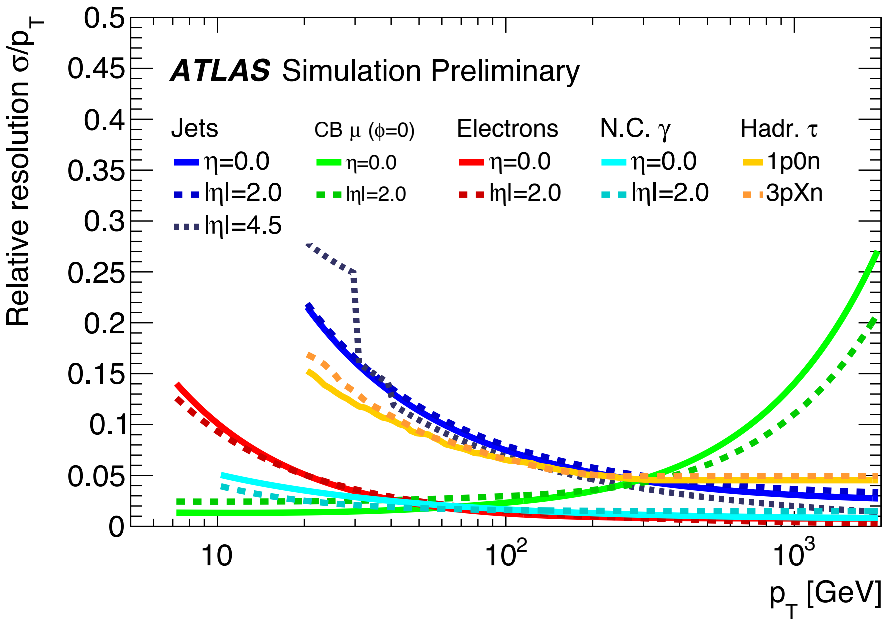
A new ATLAS PDF set:
ATLASepWZTop18

Based on the NNLO analyses of W/Z, top, *ep* data from HERA

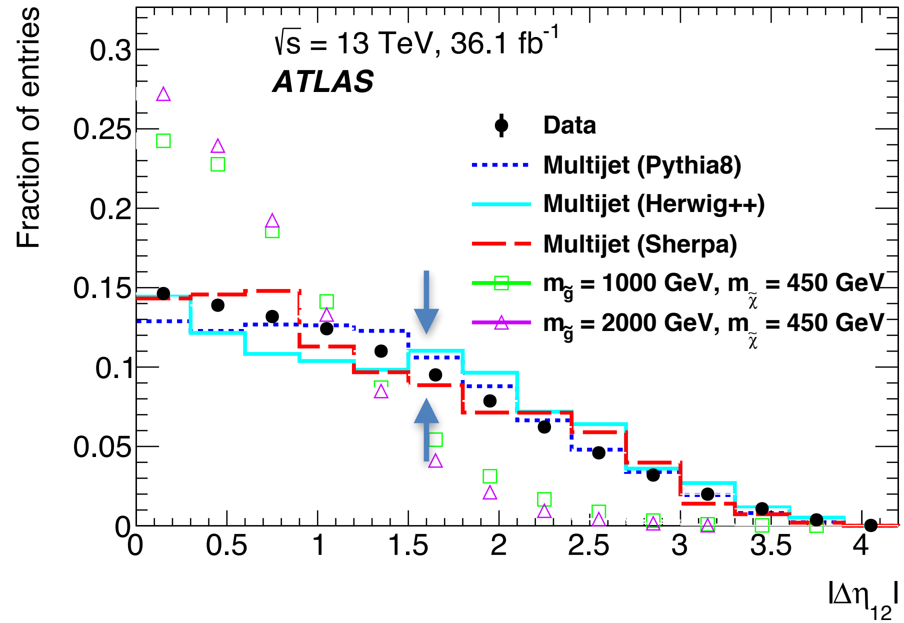
[2] Theory Uncertainties in BSM searches (from Run-2 analyses)

QCD process backgrounds

fake MET through jet resolution effects
ATLAS-CONF-2018-038



fake signal topology from combinatorics
Phys. Lett. B 785 (2018) 136



1) Uncertainties from fakes (above plots)
2) Uncertainties in Signal & background predictions
   - matrix element : QCD order, N_{parton}
   - parton shower model, ISR/FSR tune
   - PDF/α_s variations, flavor scheme, etc

Impact all LHC physics

Background estimation strategy in non-resonant searches

1) Define control regions (CR)

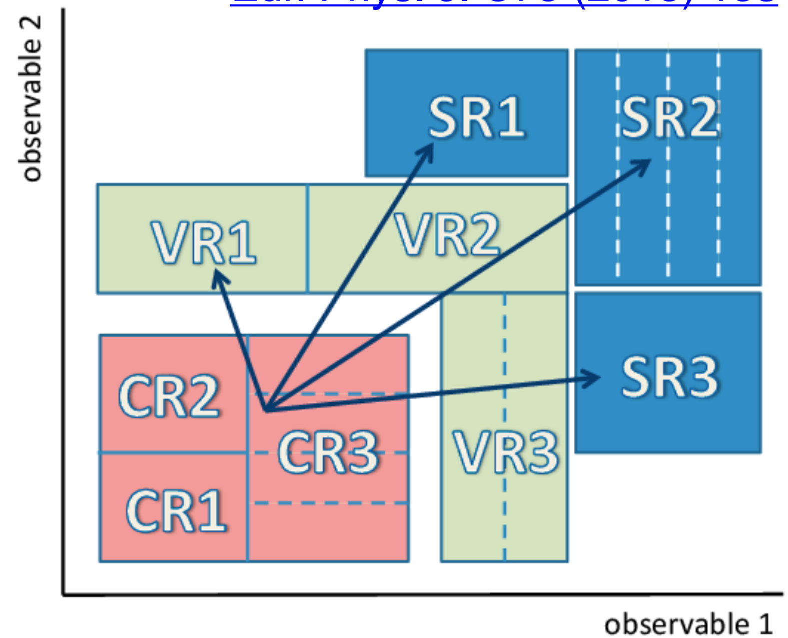
- background dominate regions
- corresponds to background source

2) Normalize MC background prediction in the CR

- simultaneous fit to data in all CR
- normalization uncertainties ~ data stats

3) Cross-check extrapolation quality in validation region (VR)

Eur. Phys. J. C75 (2015) 153

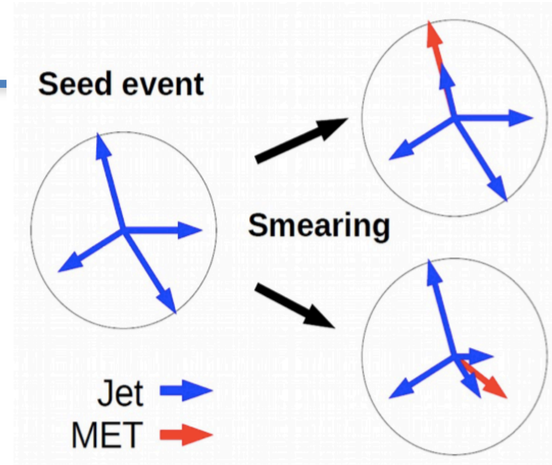


4) Extrapolate to signal region (SR) normalized by Transfer Factor (TF)
Big impact from QCD modeling uncertainties on MC

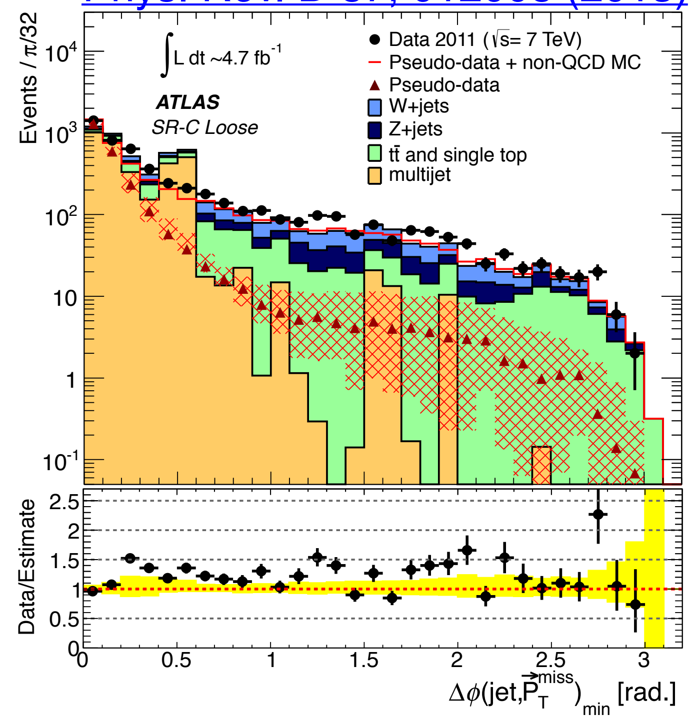
- highest order MC available is used for prediction
- evaluated by varying ME, PS, ISR/FSR tune, PDF/α_s, flavor scheme

(I) Jet smearing method

- [fundamental assumption] high MET arise from fluctuations in the detector response to jets
- Method:
  1) multi-jet event selection with good jets (low MET events)
  2) generate pT response function (RF) based on MC as a starter
  3) jet pT smeared according to expected pT response, generate pseudo-events, compared to data, modify RF, repeat this
  4) generate variables based on final RF

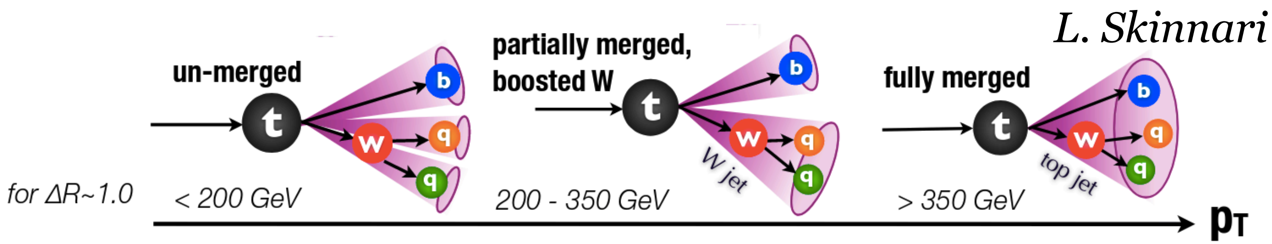


Multijet validation region
SUSY strong 0-lepton analysis
Phys. Rev. D 87, 012008 (2013)



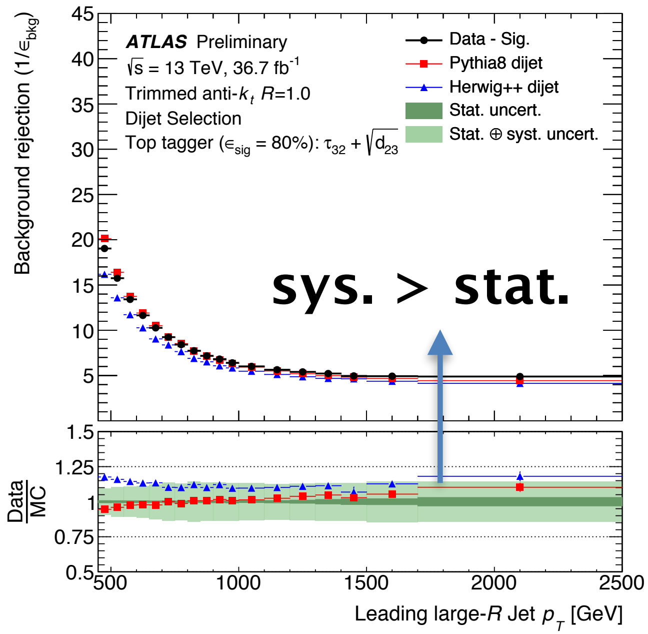
- Method limitations in future use
  - non-gaussian tail (large uncertainties)
  - fake MET from pileup jet removal algorithm inefficiency, not implemented

(II) large-R jet substructure uncertainties



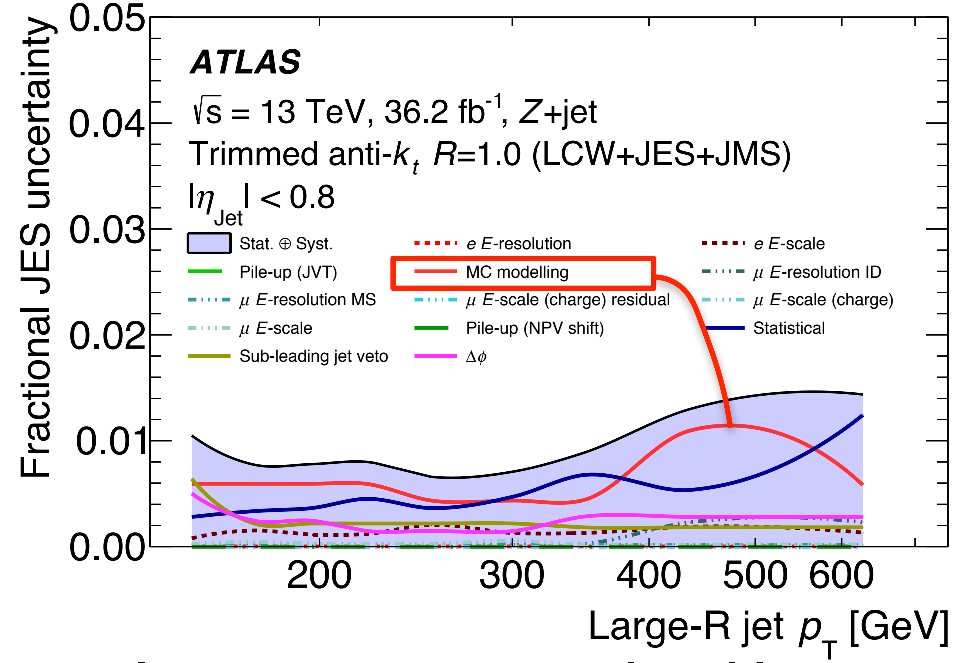
- R=1.0 anti-k_T jet to reconstruct high p_T W, top
- QCD modeling crucial : energy scale, substructure, etc

ATLAS-CONF-2017-064



QCD jet rejection in Top tagger

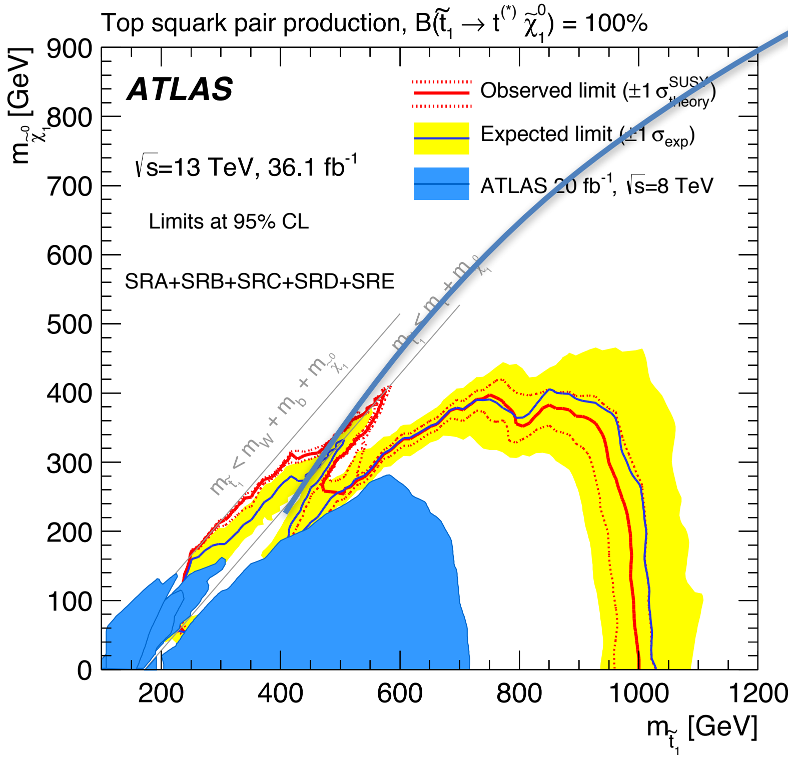
arXiv:1807.09477



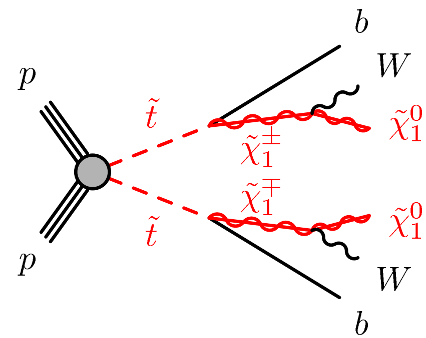
large-R jet Energy scale calibration uncertainties

a) Impact on Stop 0L search (low Δm)

direct stop pair production
JHEP 12 (2017) 085



compressed region:
top mass cannot be reconstructed
hard ISR jet required



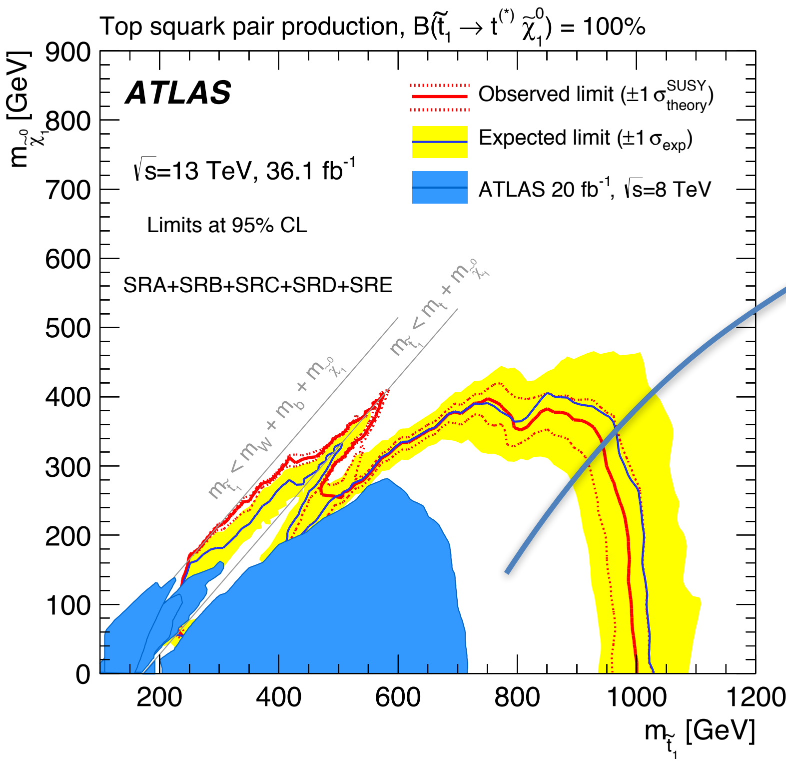
$$R_{ISR} \equiv \frac{E_T^{miss}}{p_T^{ISR}} \sim \frac{m_{\tilde{\chi}_1^0}}{m_{\tilde{t}_1}}.$$

dominant systematics from theory

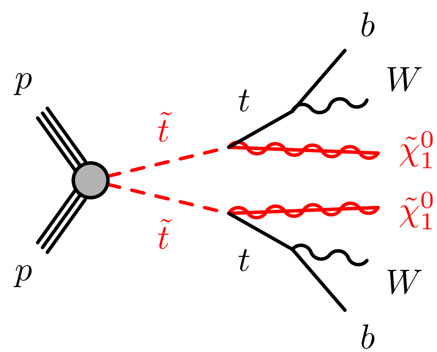
	SRC1	SRC2
Total syst. unc.	31	18
$t\bar{t}$ theory	27	11
JER	4	10
JES	4	5
Multijet estimate	12	3
Single top theory	3	2
R_{ISR} :	**0.3–0.4**	**0.4–0.5**

- dominant systematics from
  - ISR modeling in *tt* events → difference btw MCs
  - multi-jet background

a) Impact on Stop 0L search (high Δm)



- MET> 250 GeV, **0 lepton**
- $\geq$ 4 jets with p_T above (80,80,40,40) GeV
- **higher $\Delta m_{\tilde{t},\tilde{\chi}}$**
  $\geq$ 2 large-R jets at high p_T/m
- min $\Delta\phi$ between jets and MET



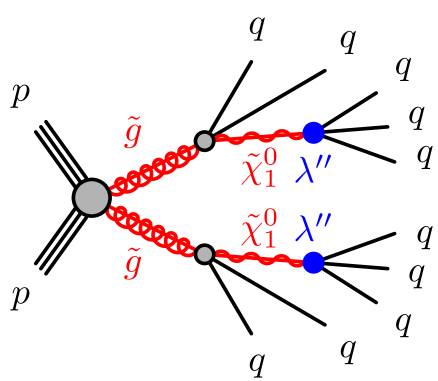
direct stop pair production
JHEP 12 (2017) 085

- dominant systematics from
  - tt events modeling
  - JER, JES, b-tagging

	SRB-TT	SRB-TW	SRB-T0
Total syst. unc.	19	14	15
$t\bar{t}$ theory	10	11	12
JER	3	4	3
JES	7	4	<1
b-tagging	5	4	4
Pileup	8	1	3

b) multi large-R jets

Search for massive supersymmetric particles in multi-jet final states

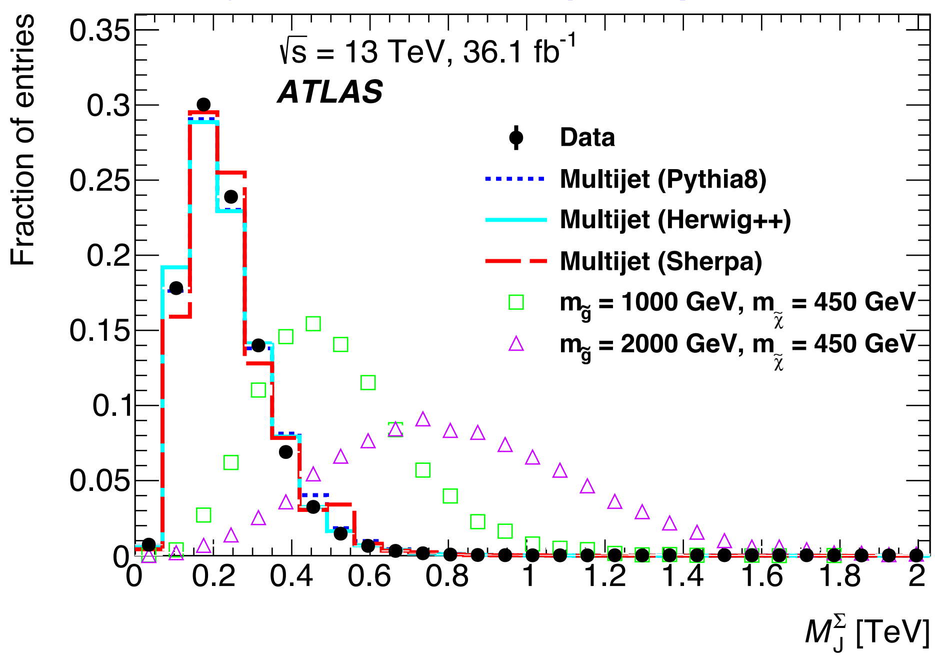


RPV SUSY gluino scenarios

fully hadronic, no MET requirement
many jets (≥ 4/5 large-R jets, b-tag categories)

total jet mass as main discriminant

[Phys. Lett. B 785 (2018) 136](#)



$$M_{\mathrm{J}}^{\Sigma} = \sum_{\substack{p_{\mathrm{T}}>200\,\mathrm{GeV} \\ |\eta|\leq 2.0 \\ j=1-4}} m_{\mathrm{jet}}^{j}$$

main background : multi-jet, combinatorics → huge uncertainties

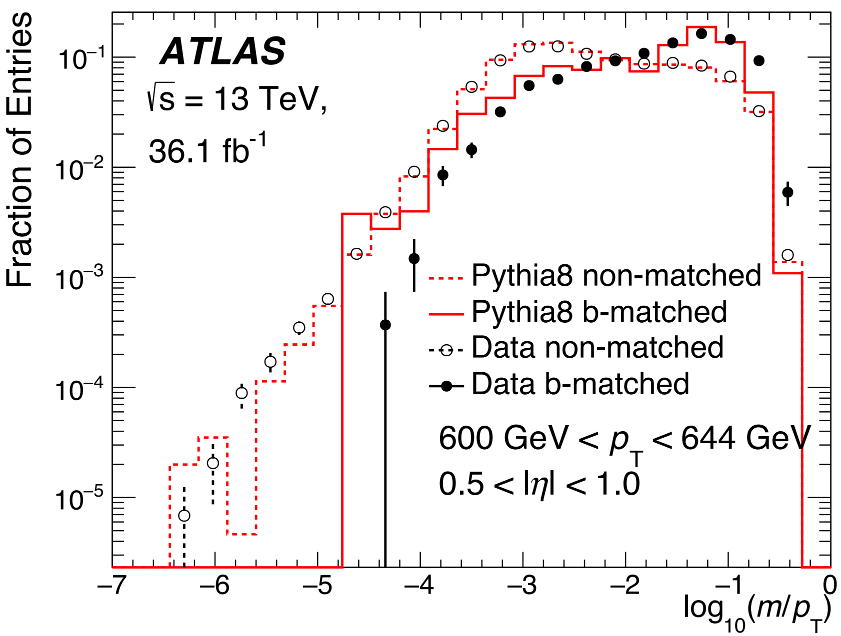
→ fully data-driven estimates (next page)

b) multi large-R jets

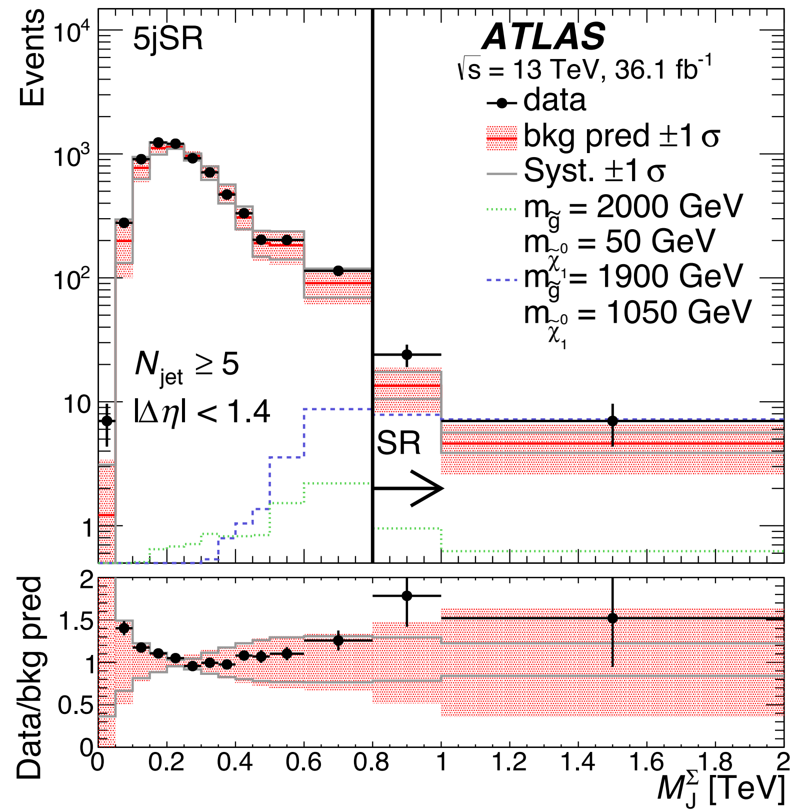
- the SM mass template obtained in CR (JHEP 05 (2014) 005)
- estimate uncertainties by varying CR definition
- background estimate uncertainties, ~20%, have a large impact on the final results

Phys. Lett. B 785 (2018) 136

jet mass template data vs. MC

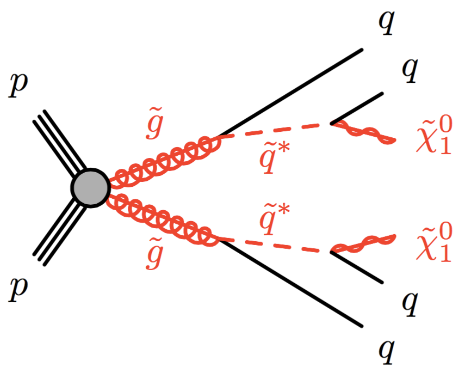


background estimates in SR



c) displaced vertex search

- Search for the long-lived gluinos (Split-SUSY scenarios)

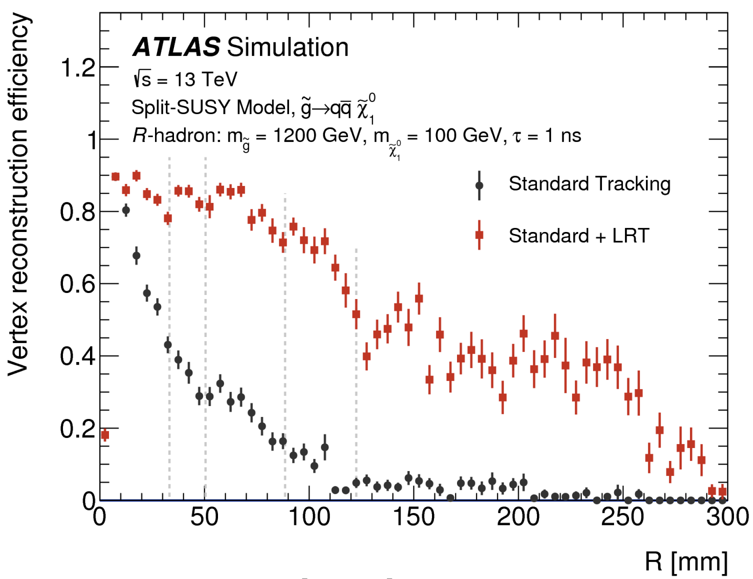


Cut and count analysis

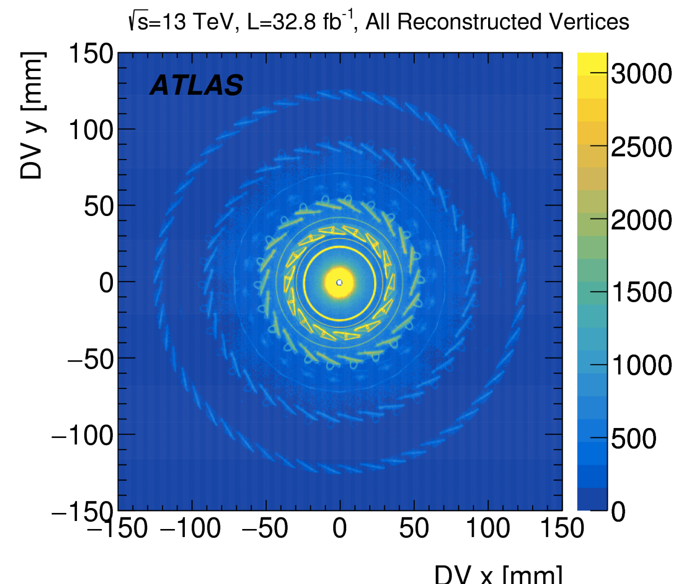
- **Missing ET > 250 GeV**
- **>= 1 Displaced vertex**
  - **mDV>10GeV**
  - **ntrack>=5**

- **No SM irreducible background expected**
- **Experimental (detector) backgrounds**

[Phys. Rev. D97 (2018) 052012](#)



non-standard reconstruction

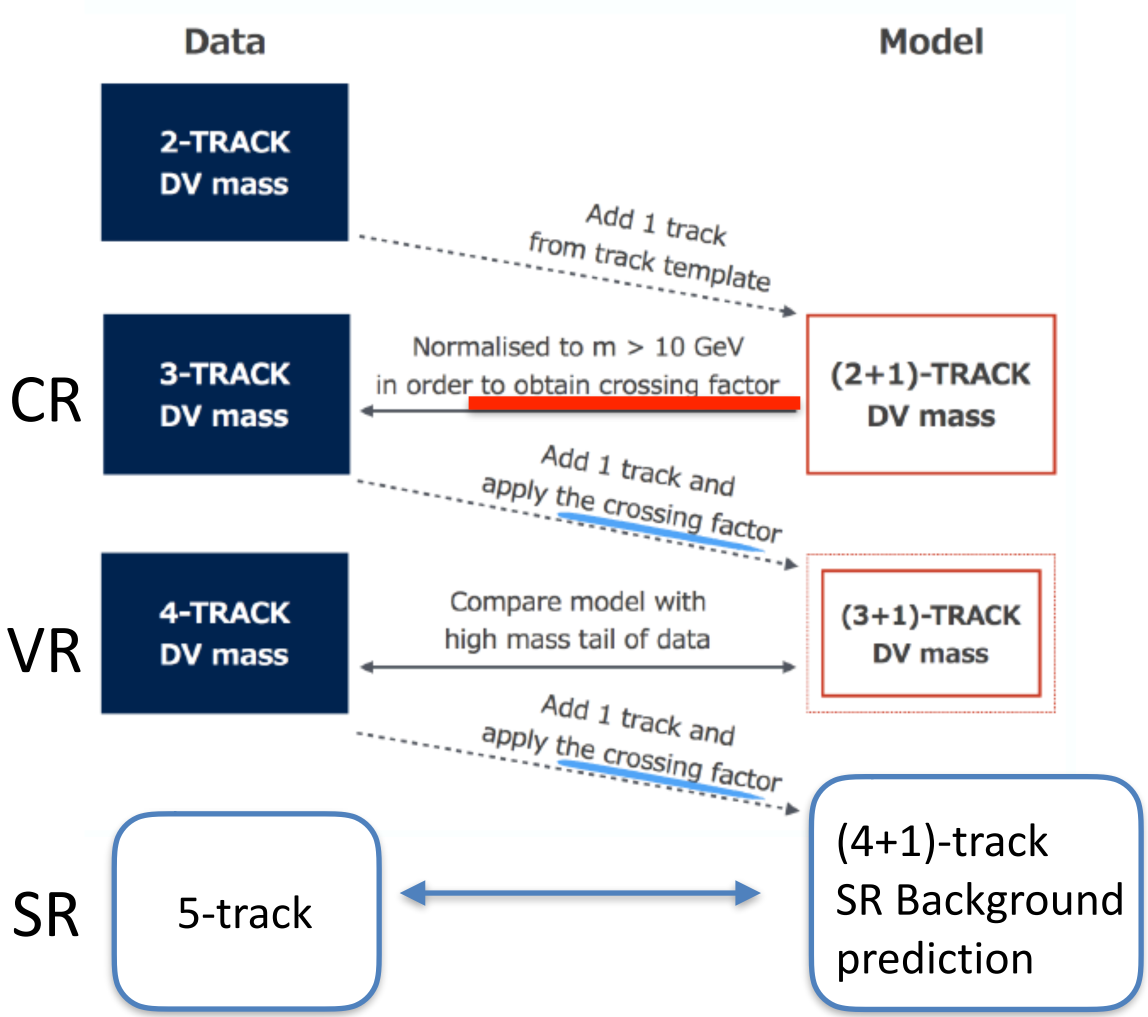


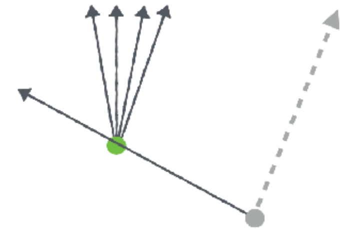
detailed material map needs to be known to suppress hadronic interaction BG

c) displaced vertex search

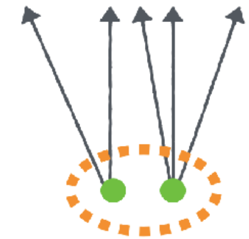
- data driven background estimate

random crossing BG estimation

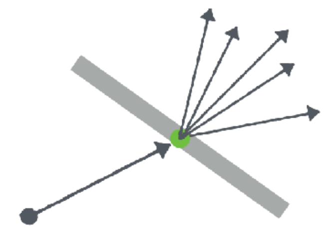




random crossing (high mass region)



merged vertex



hadronic interaction (low mass region)

however dominant bg comes from the hadronic interaction

there are gases in the detector gap regions

prediction in SR is extrapolated from low mass region, expo shape assumed → −100 ~ 300%

need to revisit BG modelings

Personal Remark (request for theory community)

- Towards HL-LHC, people in experimentalists currently work very hard on the hardware upgrades and physics prospects analyses
- Due to the lack of hints/evidences of the new physics yet, we would foresee a hard time in keeping the strong momentum during the lifetime of HL-LHC (budget, attracting younger generation…)
- Precision measurements of Higgs couplings, searches of SUSY in even higher mass regime (>2TeV) are important and just fine, but we need something more
- I think, we (experimentalists) are desperate for new theoretical framework (like "SUSY" before the start of LHC) which guide us to a new paradigm in HL-LHC regime

Summary

- Theory uncertainties will be major systematics in HL-LHC
  - systematics on signal/background predictions
  - estimation of the fake backgrounds

- These affect measurements (eg. Higgs) type analyses, as well as search (eg. BSM) type analyses

- Need to tackle this: looking for the new higher order predictions, new methods/modelings to have better descriptions of data with the simulations

- Cooperation between experiments and theorists is vital