Search for squarks and gluinos in final states with jets and missing transverse momentum using 139 fb^{-1} of $\sqrt{s}$ =13 TeV pp collision data with the ATLAS detector

09.April.2021

Reinterpretation: Auxiliary Material Presentation (RAMP) kickoff meeting

Kenta Uno

The University of Tokyo → Niigata university

Introduction

Kenta Uno

- 2017 – 2020: Worked on SUSY Strong 0-lepton analysis
  - Main analyzer of BDT analysis
- 2020 – Present: Move to Belle II experiment as post-doc

I use figures in the following Refs.

- Ref1: SUSY Strong 0-lepton paper [link]
- Ref2: My Doctoral thesis [link]

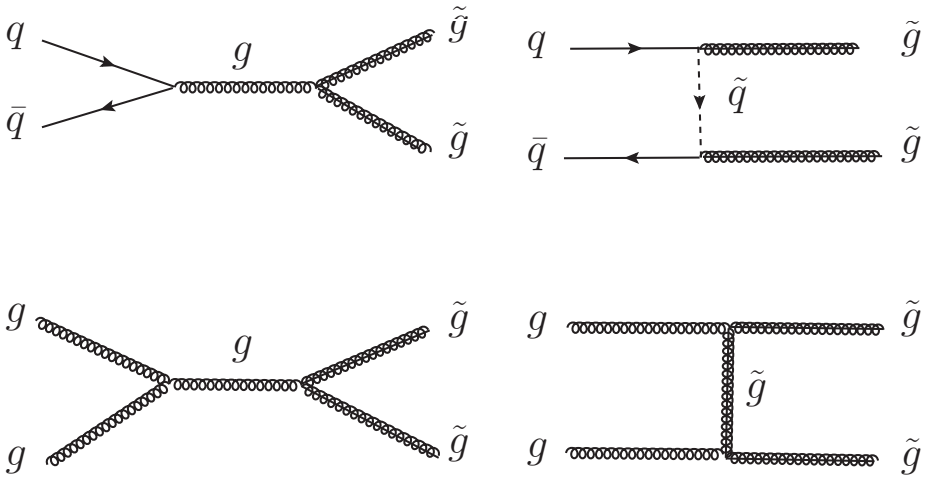
→ If you have question, please contact me, uno@hep.sc.niigata-u.ac.jp

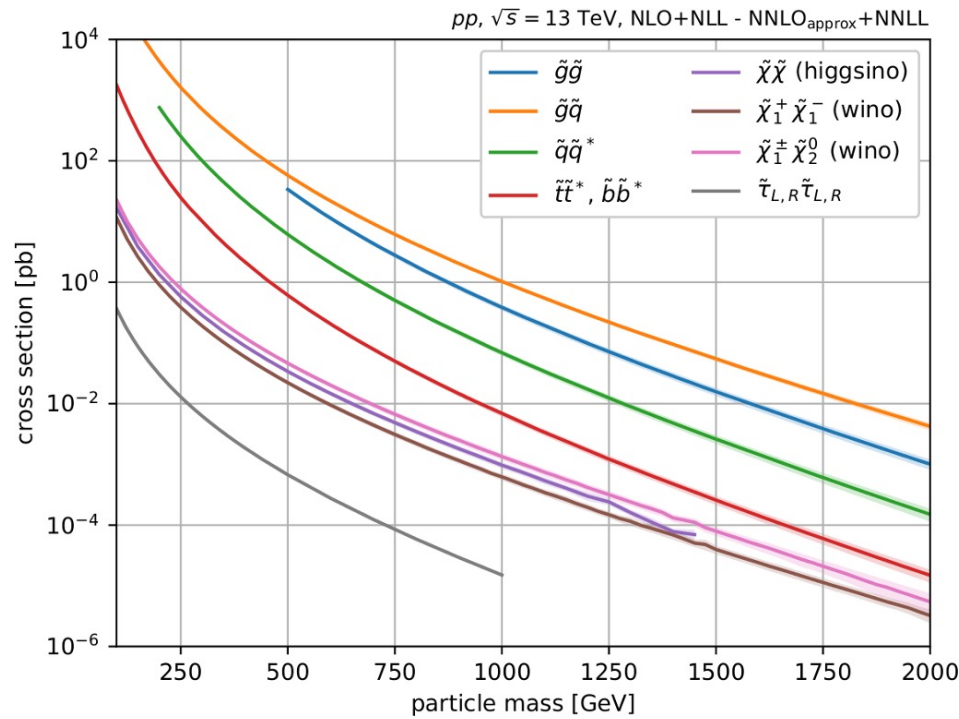
SUSY search at LHC-ATLAS

LHC : Proton-Proton collision at $\sqrt{s} = 13\ \mathrm{TeV}$

- Gluino pair production has large x-section (Strong interaction)
  - Golden channel in LHC SUSY search!

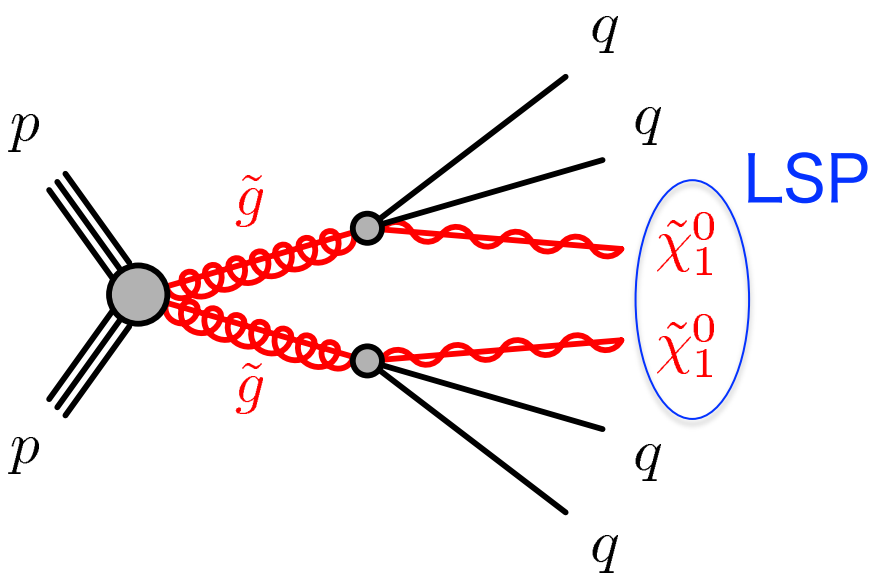
Gluino production at LHC

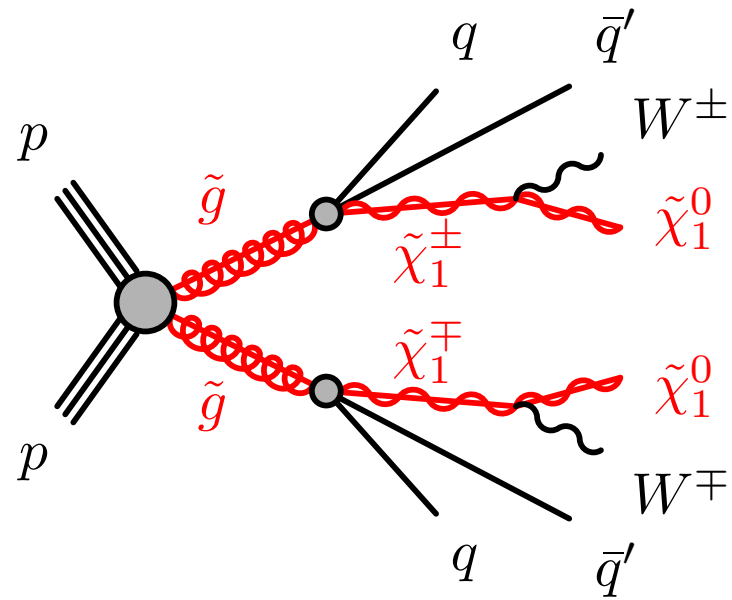




Gluino decay process

- Gluino direct decay ($\tilde{g} \to qq\tilde{\chi}_1^0$): Simplest process
- Gluino one-step decay ($\tilde{g} \to qqW\tilde{\chi}_1^0$):





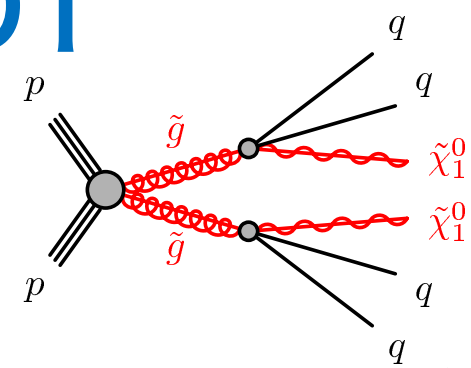
Event Topology

- **Multi-jet + Large missing transverse momentum (MET)**

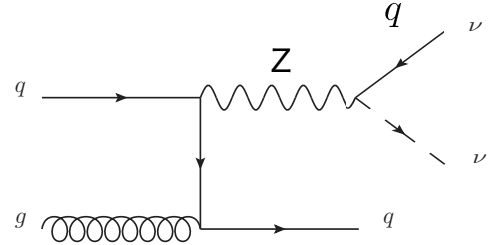
Analysis approach: BDT

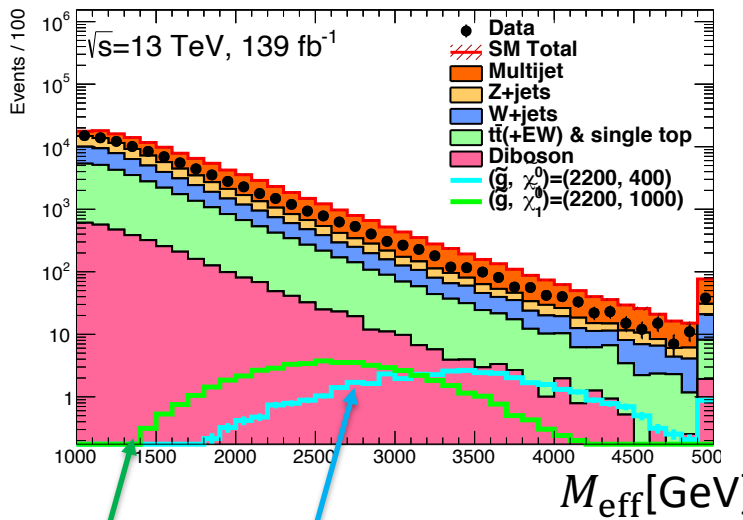
Key variable in the previous analysis: M_{eff}

$$M_{\mathrm{eff}} = \sum_i |p_{\mathrm{T}}(i)| + E_{\mathrm{T}}^{\mathrm{miss}}$$



- However, it is not efficient as $\tilde{\chi}_1^0$ mass is large

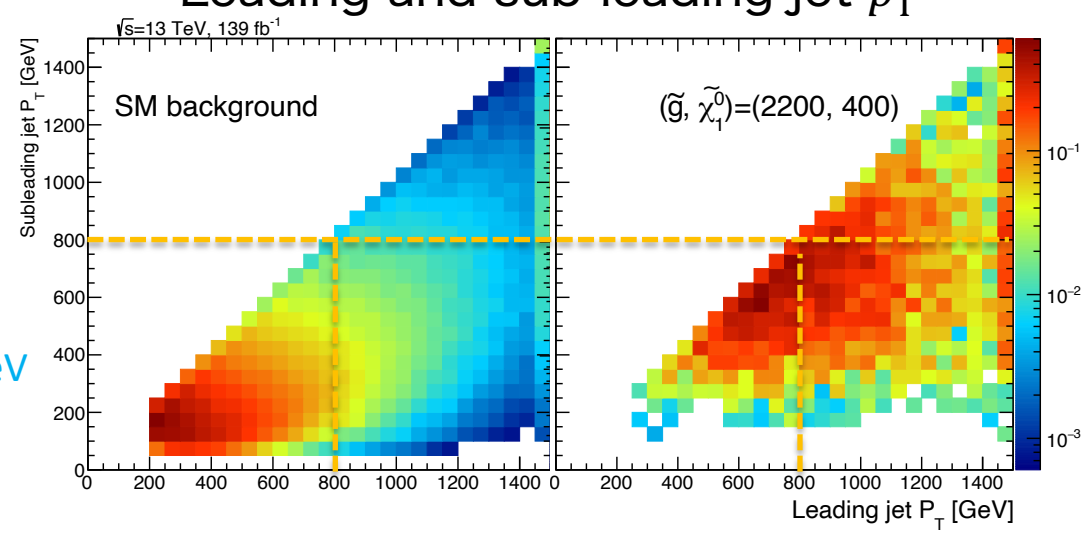




$m(\tilde{g}) = 2200$ GeV, $m(\tilde{\chi}_1^0) = 400$ GeV

$m(\tilde{g}) = 2200$ GeV, $m(\tilde{\chi}_1^0) = 1000$ GeV

Leading and sub-leading jet p_{T}

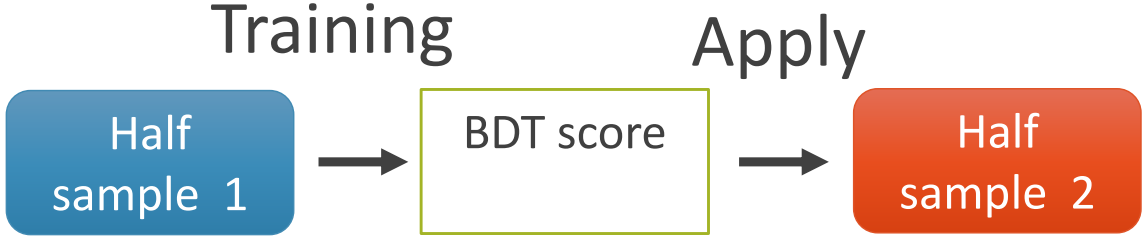


Using correlation between variables is still efficient.

→ BDT analysis is useful for SUSY search

How to train for gluino search?

<u>Training method</u>



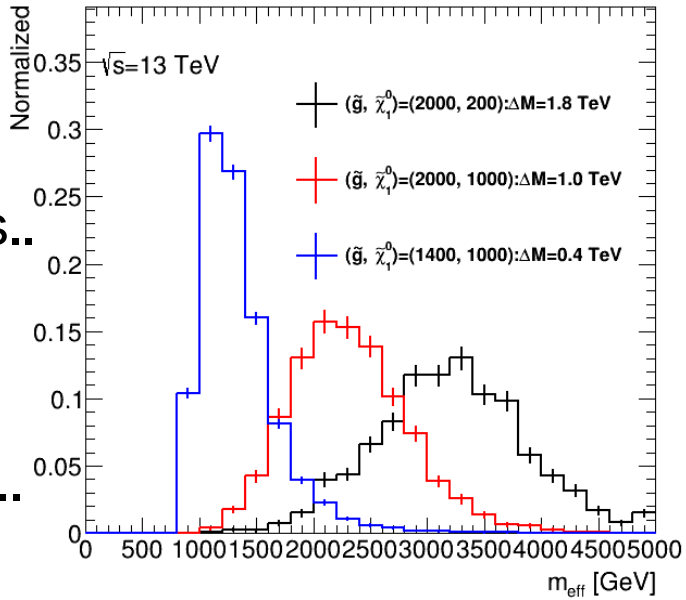
- Use gluino events and SM background events
- Add 10 – 12 input variables in BDT (eg. $p_{\mathrm{T}}, \eta, E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{eff}} \cdots$)

<u>We don't know gluino/neutraino mass.</u>

- The kinematics depends on these masses..

→ We have to prepare BDT score per each gluino/neutralino mass..

- Need many BDT scores : Not possible..

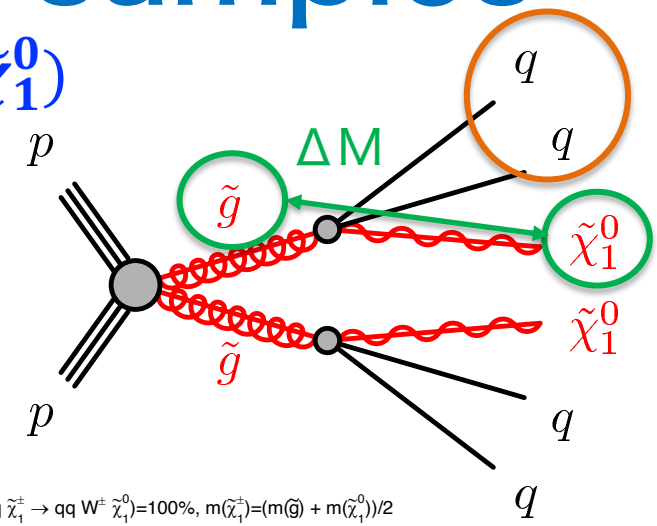


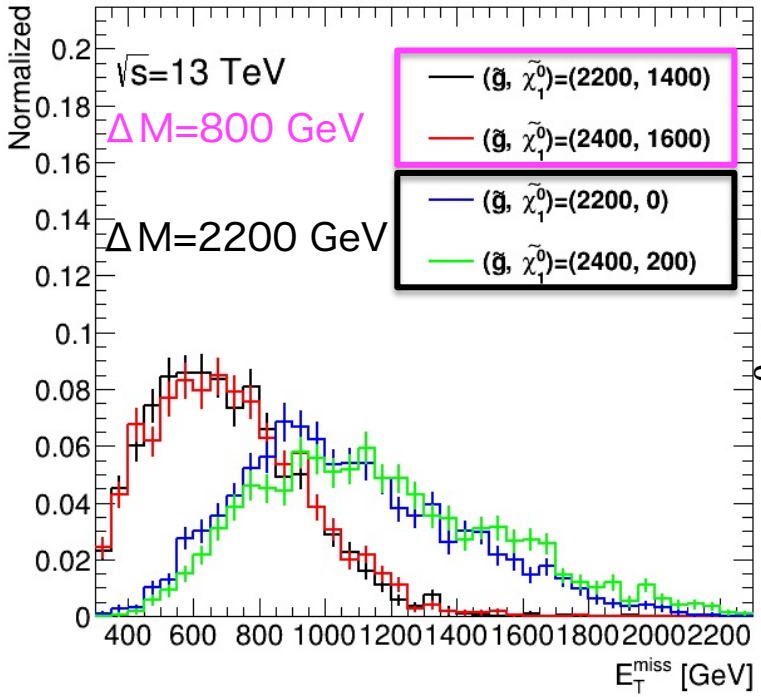
Determine training samples

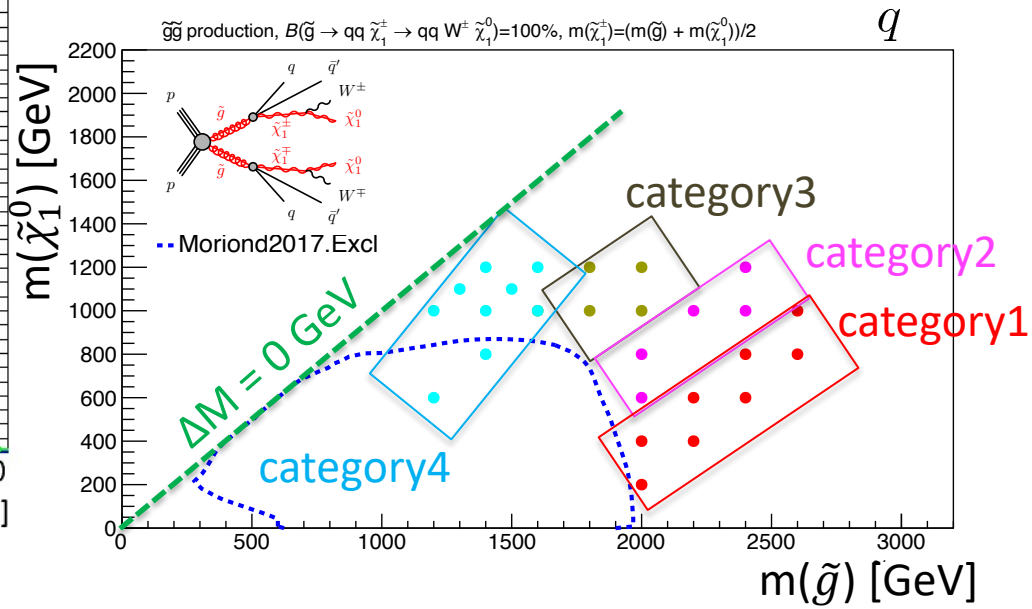
The jet activity depends on $\Delta M(\tilde{g}, \tilde{\chi}_1^0)$

- Similar $\Delta M \leftrightarrow$ Similar kinematics

→ Prepare 4 categories based on ΔM.



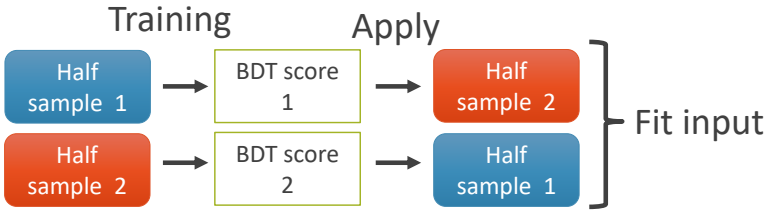




By using category1– 4,

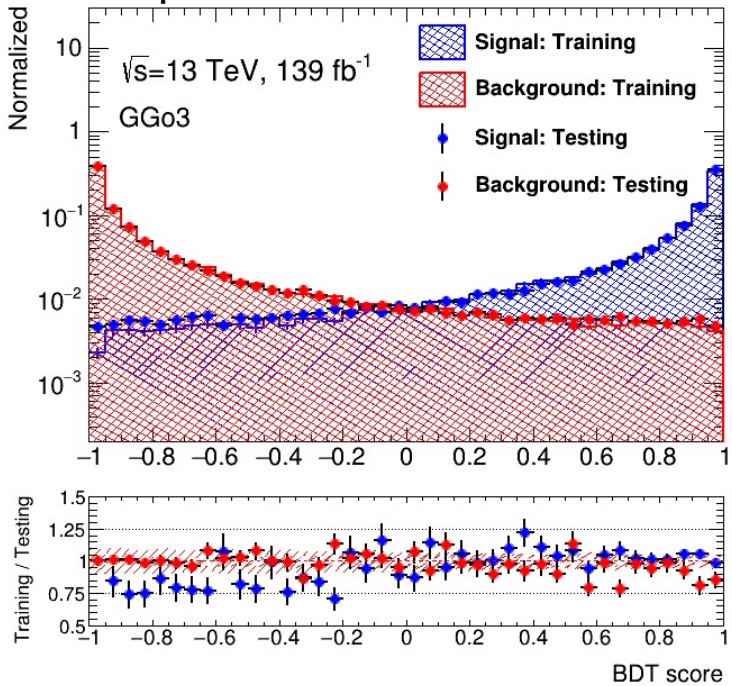
- <u>All signal probing point can be covered.</u>

Overtraining

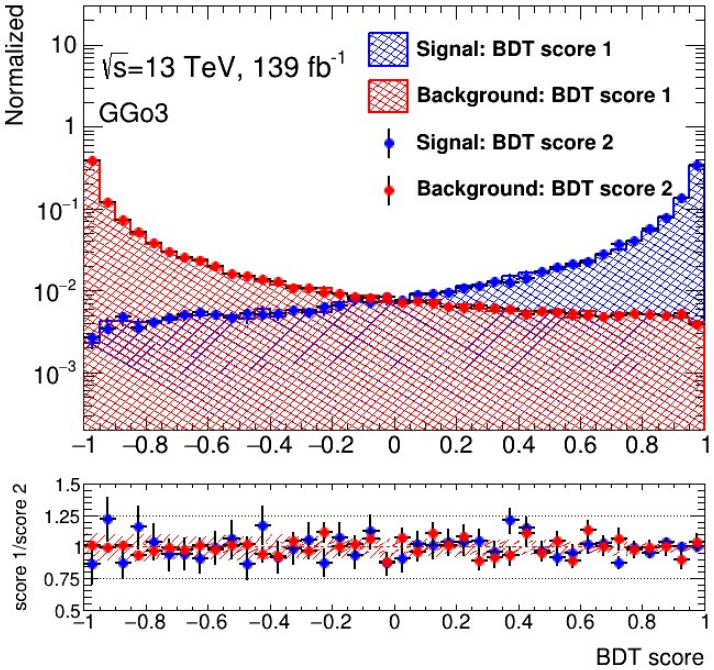


- Samples are divided into 2 subsets
  - If use half sample1 as training, use sample 2 as testing.
- The training is also done to another one.
  - It can keep the statistics in the final fit.

Comparison b.t.w Train and Test



Comparison b.t.w BDT1 and BDT2

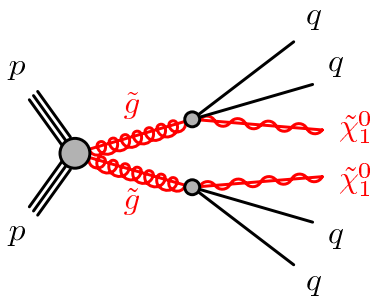


No effect on the bias due to overtraining

Signal region definition

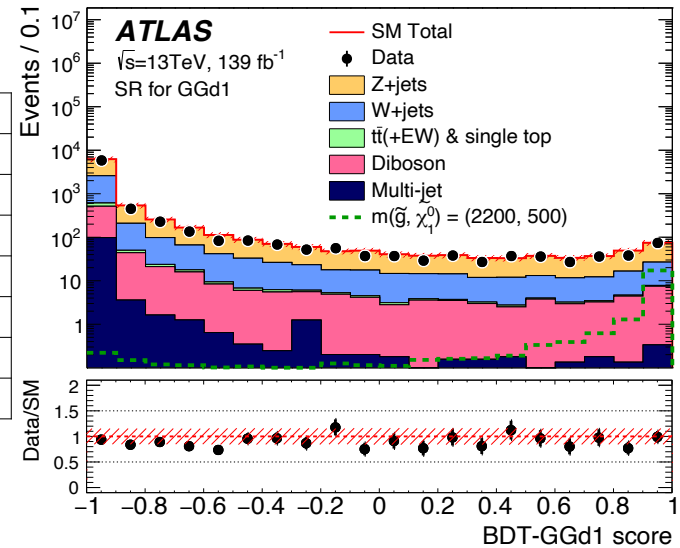
Define SRs per each training category: 8 SRs

Gluino direct decay

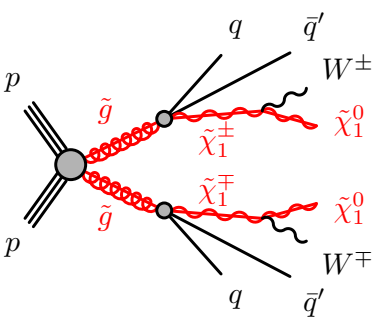


	BDT-GGd1	BDT-GGd2	BDT-GGd3	BDT-GGd4
N_j	≥ 4			
$\Delta\phi(j_{1,2,(3)}, \boldsymbol{p}_T^{miss})$ min	> 0.4			
$\Delta\phi(j_{i>3}, \boldsymbol{p}_T^{miss})$ min	> 0.4			
$E_T^{miss}/m_{eff}(N_j)$	> 0.2			
m_{eff} [GeV]	> 1400		> 800	
BDT score	> 0.97	> 0.94	> 0.94	> 0.87
$\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ [GeV]	1600–1900	1000–1400	600–1000	200–600

Target ΔM

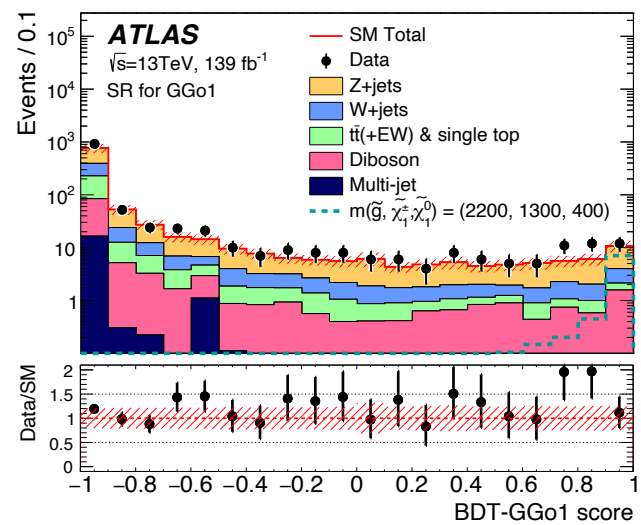


Gluino onestep decay



	BDT-GGo1	BDT-GGo2	BDT-GGo3	BDT-GGo4
N_j	≥ 6		≥ 5	
$\Delta\phi(j_{1,2,(3)}, \boldsymbol{p}_T^{miss})$ min	> 0.4			> 0.2
$\Delta\phi(j_{i>3}, \boldsymbol{p}_T^{miss})$ min	> 0.4			> 0.2
$E_T^{miss}/m_{eff}(N_j)$	> 0.2			
m_{eff} [GeV]	> 1400		> 800	
BDT score	> 0.96	> 0.87	> 0.92	> 0.84
$\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ [GeV]	1400–2000	1200–1400	600–1000	200–400

Target ΔM



A good separation power between sig/bkg

Background estimation

SUSY search: Use tail in the background distribution

- MC simulation should be corrected by using the data

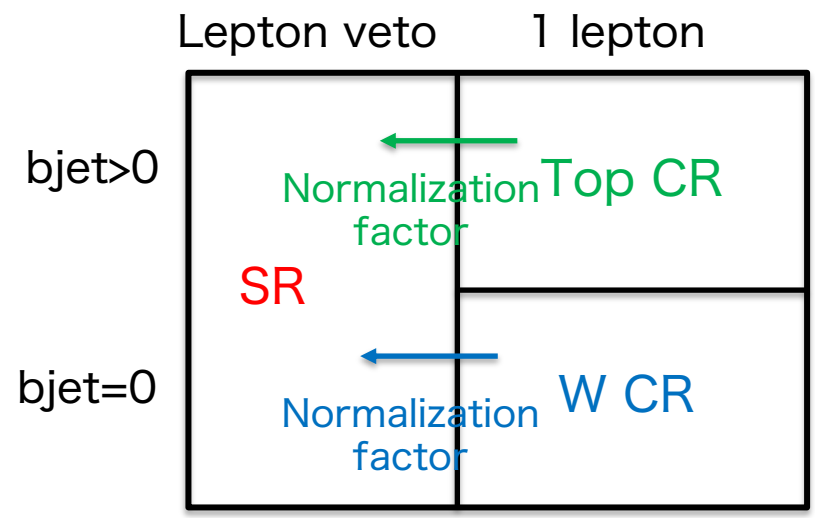
→ Use Control regions (CRs)

Normalization factor

$$N_{SR} = N_{SR}^{MC} \times \left[\frac{N_{CR}^{data}}{N_{CR}^{MC}} \right],$$

N_{SR} : # predicted bkg events in SR
N_{SR}^{MC} : # bkg events in SR (MC simulation)
N_{CR}^{MC} : # bkg events in CR (MC simulation)
N_{CR}^{data}: # bkg events in CR (data)

Signal region VS control region



Assume mismodeling in CRs is same as SRs

- CRs are designed to be kinematically close to corresponding SRs

Fit result

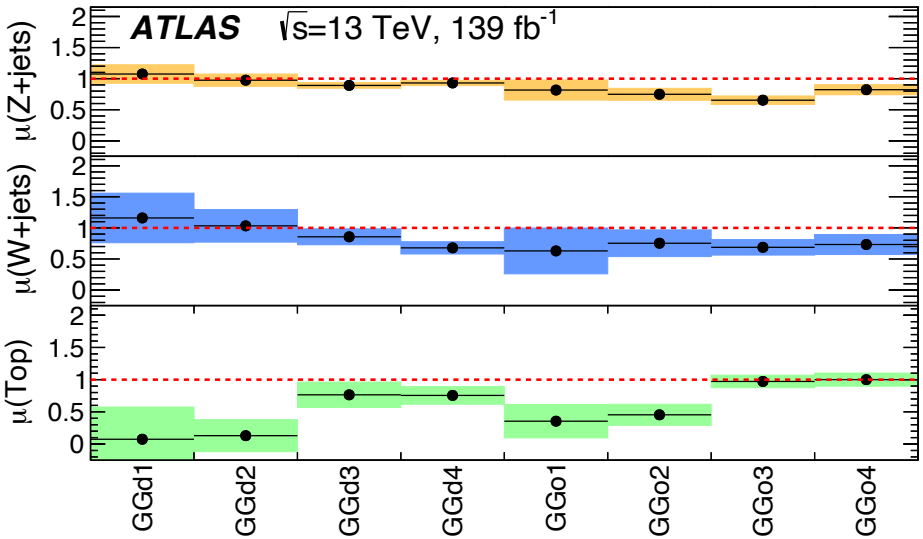
Perform a profile Log Likelihood Ratio (LLR) approach

- Normalization factors obtained from the fit

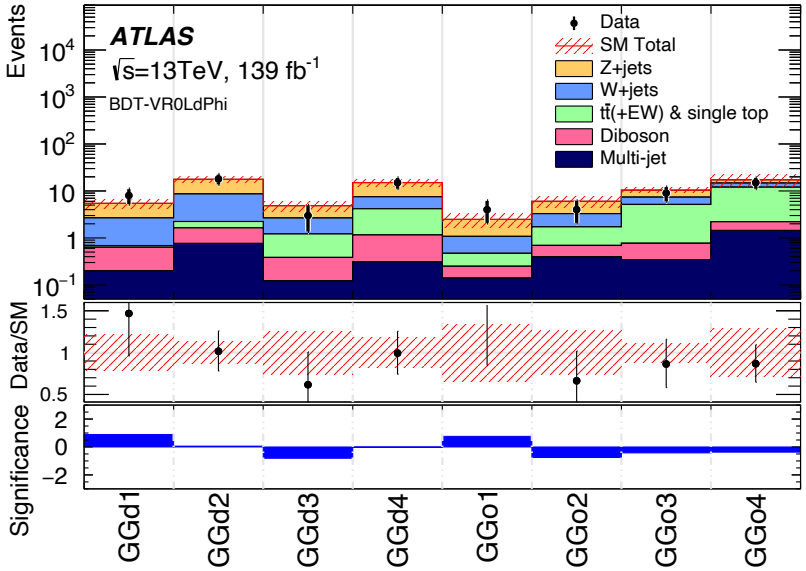
→ Apply them to each MC sample in SR and Validation region (VR)

- VR is to validate background estimation

→ It is kinematically close to SR, but orthogonal SR/CR



Normalization factor

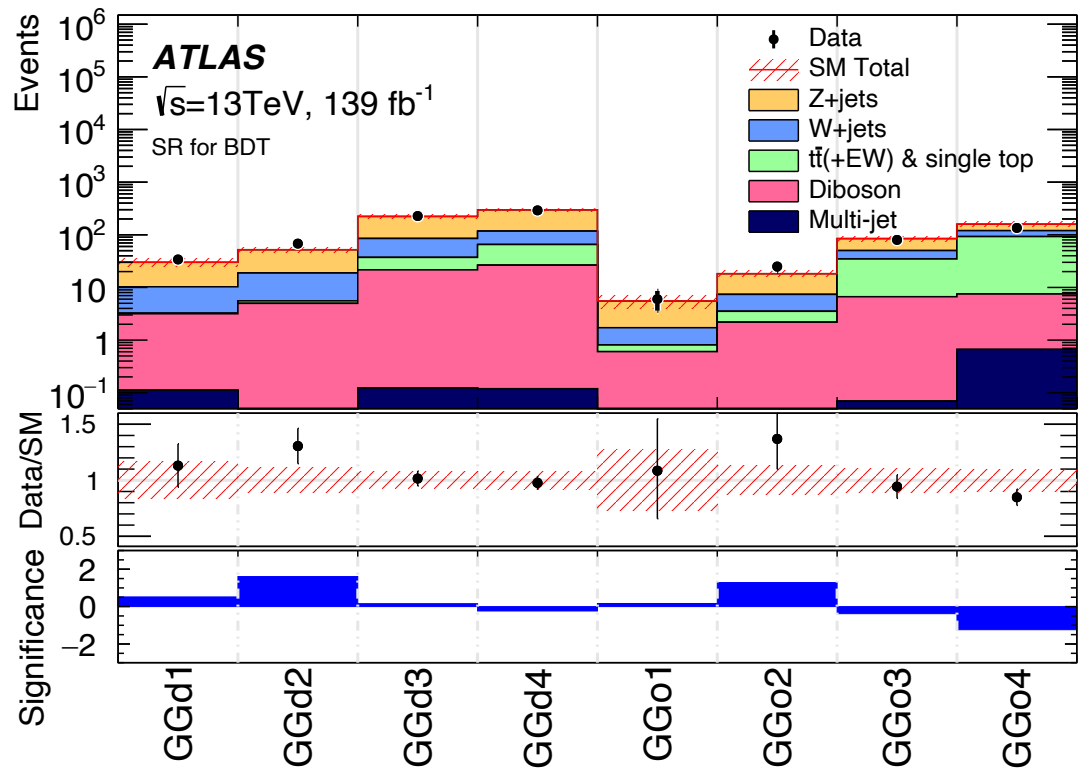


Validation region

Confirmed background estimation in BDT analysis is fine

Result in SRs

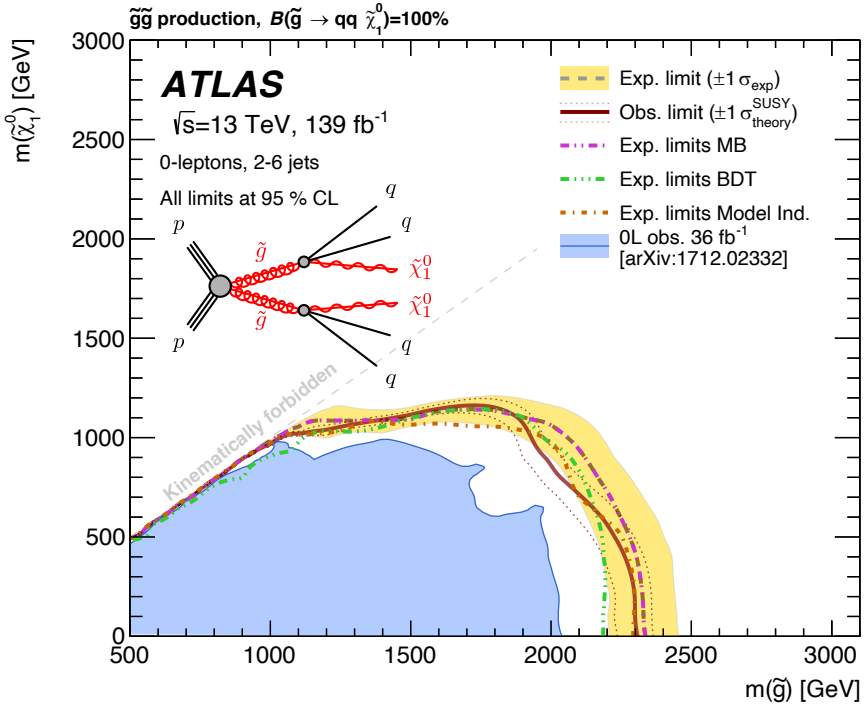
	BDT regions			
Signal Region	**GGd1**	**GGd2**	**GGd3**	**GGd4**
Total bkg pre-fit	29	56	253	348
	Fitted background events			
Diboson	3.0 ± 0.9	4.9 ± 1.4	21 ± 5	26 ± 7
Z/γ^*+jets	20 ± 4	33 ± 5	139 ± 14	180 ± 18
W+jets	7.1 ± 2.6	13 ± 4	48 ± 8	52 ± 9
$t\bar{t}$(+EW) + single top	$0.1^{+0.3}_{-0.1}$	$0.6^{+0.8}_{-0.6}$	16 ± 5	39 ± 11
Multi-jet	$0.1^{+0.1}_{-0.1}$	–	$0.1^{+0.1}_{-0.1}$	$0.1^{+0.1}_{-0.1}$
Total bkg post-fit	30 ± 5	52 ± 6	223 ± 17	298 ± 23
Observed	34	68	227	291
$\langle \epsilon\sigma \rangle^{95}_{\text{obs}}$ [fb]	0.13	0.24	0.33	0.36
S^{95}_{obs}	18	33	46	50
S^{95}_{exp}	15^{+6}_{-4}	20^{+8}_{-6}	44^{+17}_{-12}	54^{+21}_{-15}
p_0 (Z)	0.30 (0.51)	0.05 (1.60)	0.44 (0.15)	0.50 (0.00)
Signal Region	**GGo1**	**GGo2**	**GGo3**	**GGo4**
Total bkg pre-fit	7	25	111	177
	Fitted background events			
Diboson	0.6 ± 0.2	2.2 ± 0.6	6.6 ± 2.2	6.8 ± 2.1
Z/γ^*+jets	3.8 ± 1.3	10.9 ± 1.9	35 ± 6	39 ± 7
W+jets	0.9 ± 0.5	3.9 ± 1.3	16 ± 4	27 ± 6
$t\bar{t}$(+EW) + single top	0.2 ± 0.2	1.3 ± 0.8	28 ± 6	85 ± 14
Multi-jet	–	–	$0.1^{+0.1}_{-0.1}$	$0.7^{+0.7}_{-0.7}$
Total bkg post-fit	5.5 ± 1.5	18 ± 2.4	85 ± 9	159 ± 16
Observed	6	25	80	135
$\langle \epsilon\sigma \rangle^{95}_{\text{obs}}$ [fb]	0.05	0.12	0.16	0.18
S^{95}_{obs}	7.1	17	22	25
S^{95}_{exp}	$6.9^{+2.3}_{-1.6}$	11^{+5}_{-2}	25^{+10}_{-7}	37^{+14}_{-10}
p_0 (Z)	0.49 (0.01)	0.10 (1.28)	0.50 (0.00)	0.50 (0.00)

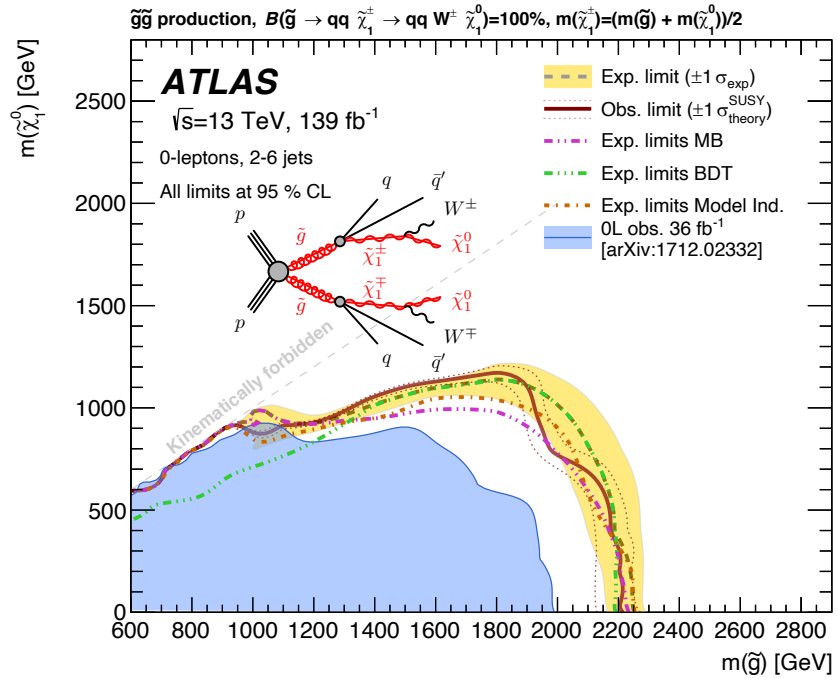


No excess over the background events

Interpretation

Obtained significant improvement





By combining other analysis method, the excluded limits are

- Gluino direct decay: $m_{\tilde{g}} < 2.3$ TeV, $m_{\tilde{\chi}_1^0} < 1.2$ TeV at 95% CL.
- Gluino one-step decay: $m_{\tilde{g}} < 2.2$ TeV, $m_{\tilde{\chi}_1^0} < 1.2$ TeV at 95% CL.

Most stringent gluino limit

Discussion point 1

- The auxiliary material provided for the BDT reinterpretation
  - We published HepData materials [Link]
    - Provided xml files of BDT classifiers
      → ZeroLepton2018-SRBDT-weight.tar.gz
    - Also provided analysis snipped code
      → ZeroLeptonBDT2018.cxx
      → By using this code, you can use same BDT classifiers.
    - Full likelihood is provided in the hepdata material

※ Signal events: fast sim (reco)
※ Background events: full reco

We provide data, SM total and signal yields in each BDT score bin and the acceptance and efficiencies of BDT SRs for simplified GG models.

→ I think it might help everyone..

Discussion point 2

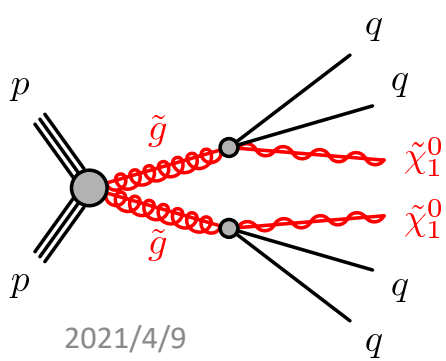
- Discussing for example truth vs reco weights and BDT weight limits of applicability for reinterpretations.
  - Since we use the tail in distribution, “reco” samples are useful.
    - Produced many “reco” signal events for training: ~50k

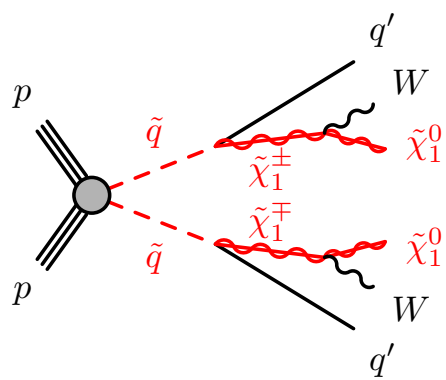
    → We didn’t study “truth” samples : (if they are fine or not)
  - BDT analysis is useful for the signals with similar kinematics

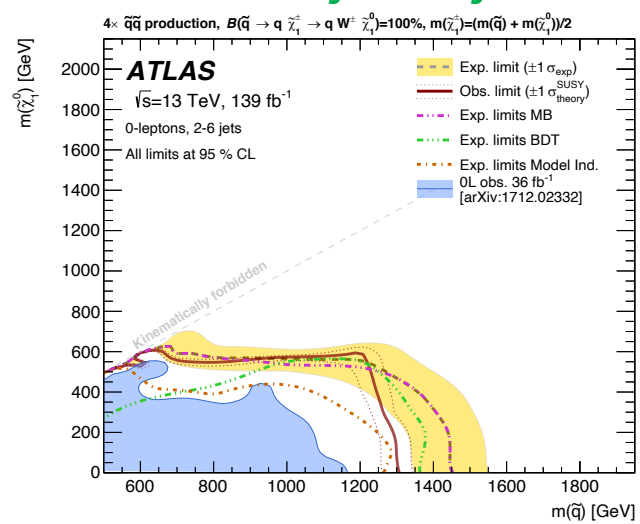
  → Applicable to evaluate other signal events

If jet multiplicity is same, the reinterpretation is fine in my analysis

Eg. Squark onestep: similar gluino direct







Summary

SUSY is one of the promising theory

- New particles are predicted in this theory.

Gluino is one of the primary targets.

- Their pair production has large x-section. (strong interaction)

ML technique (BDT) is key to improve sensitivities.

- Developed training focused on ΔM (SUSY specific feature)
  - It can lead to improve all phase spaces.

No significant excess over the background estimation

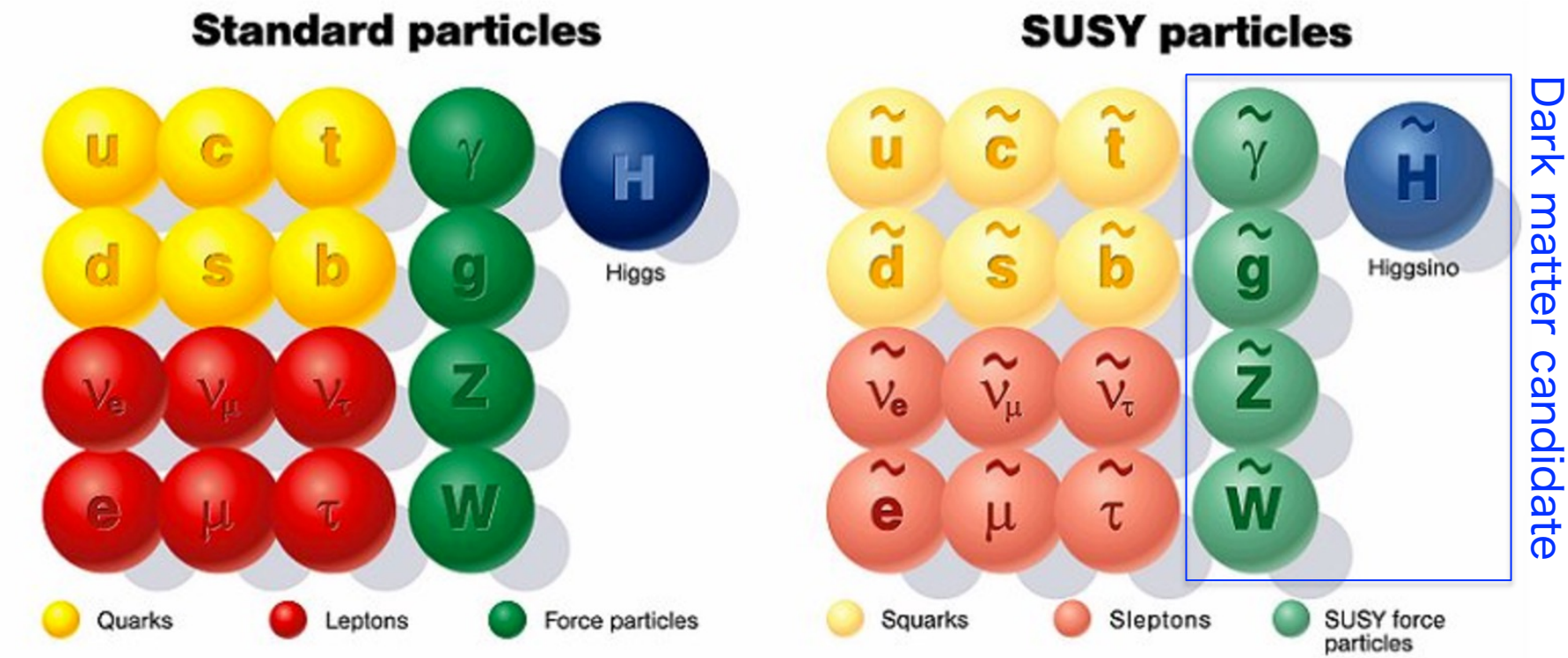
- Gluino mass (< 2.3 TeV) is excluded at 95% CL.

BACK UP

Introduction

Discovered Higgs particle at LHC-ATLAS/CMS experiments

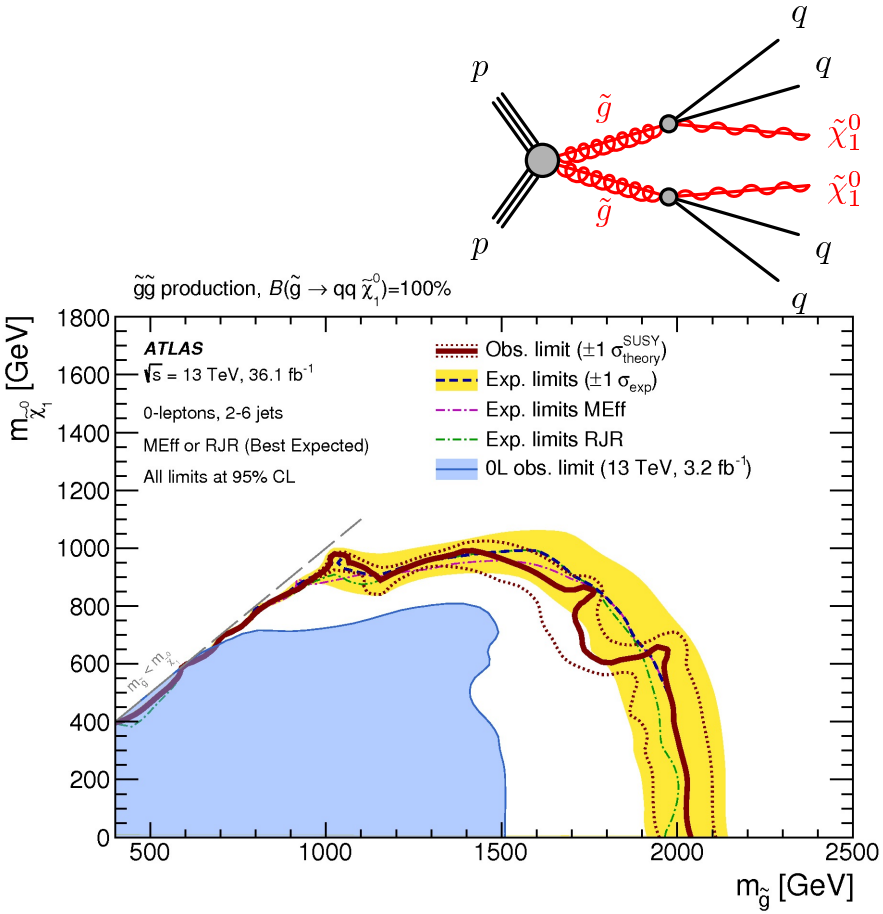
- However, there are still some remaining problems
  - Eg. Hierarchy problem, dark matter …

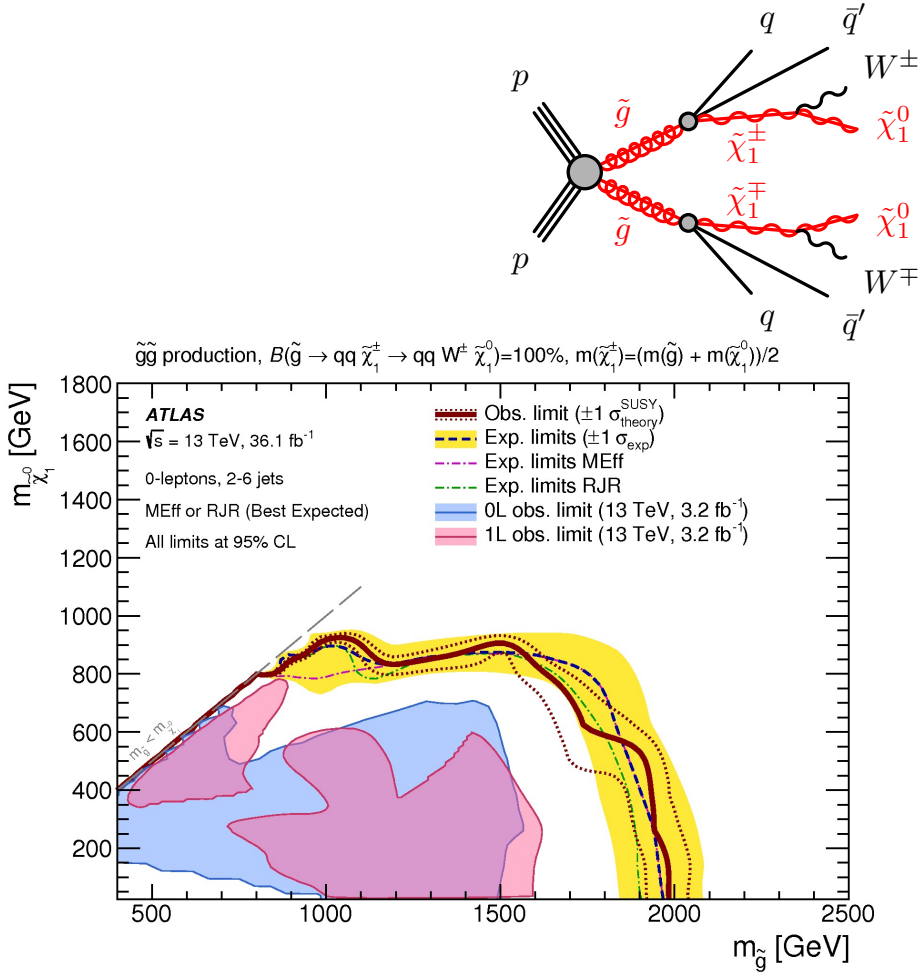


Symmetry between boson and fermion → Super Symmetry

- Possible to solve dark matter, for instance.

Previous result



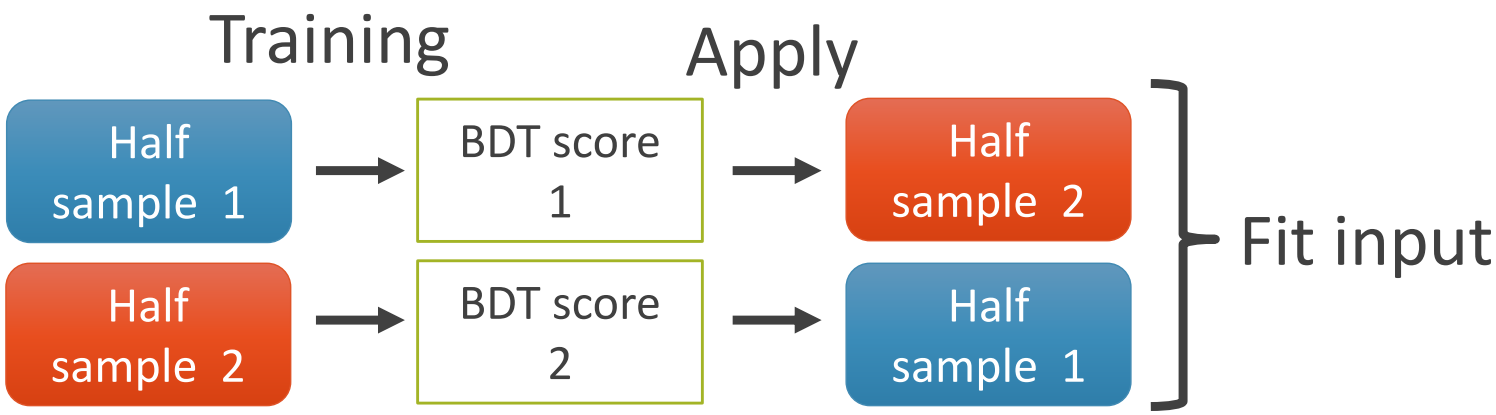


LHC ATLAS 36.1 fb^{-1}: Cut and count analysis

- Gluino direct decay ($\tilde{g} \to qq\tilde{\chi}_1^0$): $m_{\tilde{g}} < 2.0$ TeV, $m_{\tilde{\chi}_1^0} < 1.0$ TeV
- Gluino one-step decay ($\tilde{g} \to qqW\tilde{\chi}_1^0$): $m_{\tilde{g}} < 1.9$ TeV, $m_{\tilde{\chi}_1^0} < 0.90$ TeV

Training method

- Samples are divided into 2 categories : iHalf=0 or 1
  - If the events with “iHalf=0” are used as training, the other events are used as testing.
- The training is also done to another one.
  - It can keep the statistics in the final fit.

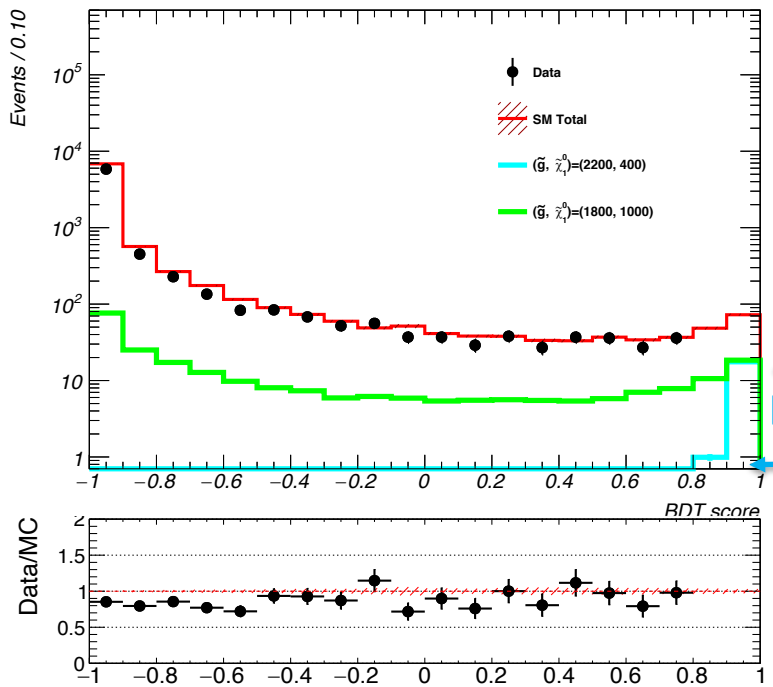


- We don’t use training sample for the analysis itself.
  - Almost no effect on the bias due to overtraining.

BDT score distribution

$m(\tilde{g})$ = 2.2 TeV, $m(\tilde{\chi}_1^0)$ = 400 GeV
ΔM = 1.8 TeV

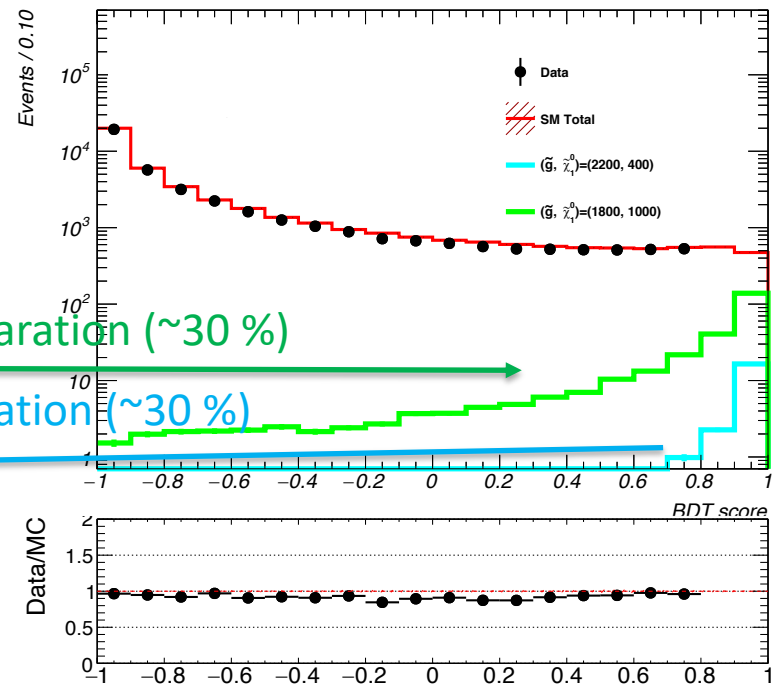
$m(\tilde{g})$ = 1.8 TeV, $m(\tilde{\chi}_1^0)$ = 1000 GeV
ΔM = 800 GeV



Category1 BDT score distribution
(High gluino, Low LSP mass target)

Better separation (~30 %)

Better separation (~30 %)



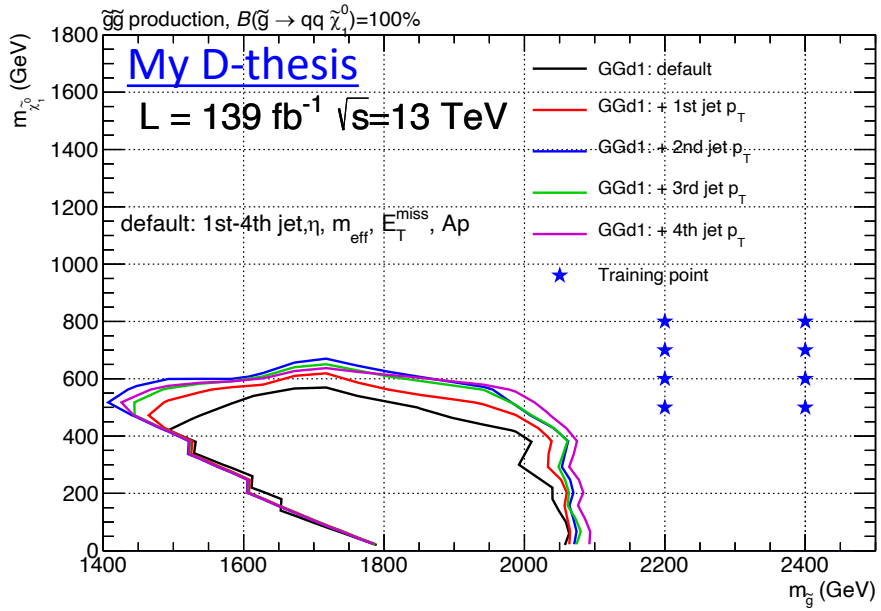
Category3 BDT score distribution
(middle gluino, high LSP mass target)

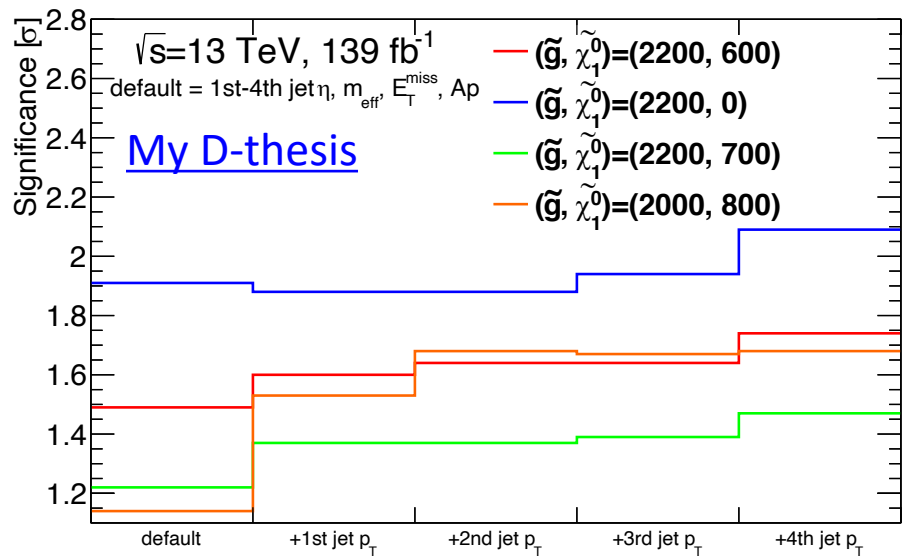
- Good separation power between signal and background.
- A good agreement of data/MC in BDT score distribution.

Input variables

Candidate: Jet p_{T}, η, $\mathrm{m_{eff}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$, Aplanarity ..

- Check the sensitivity by adding the variables
- Try to reduce the number of input variables





By adding Jet P_T, the sensitivity can be improved : ~50% $(m_{\tilde{g}}, m_{\tilde{\chi}_1^0})$=(2000, 800) GeV

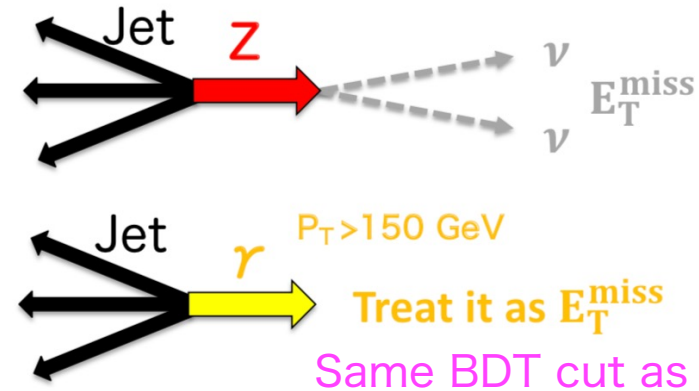
- Correlation among jet p_{T} is effective

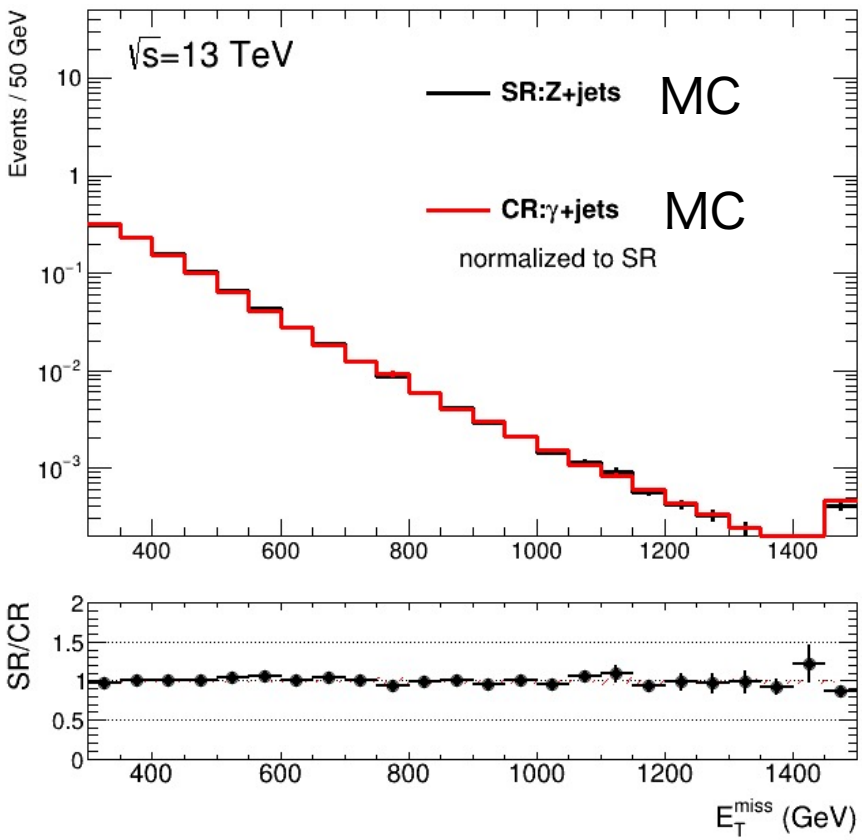
Use 10 -12 input variables: jet $p_{\mathrm{T}}, \eta, E_{\mathrm{T}}^{\mathrm{miss}}, m_{\mathrm{eff}}$, Aplanarity

Z+Jet estimation

Z+Jet events have high MET

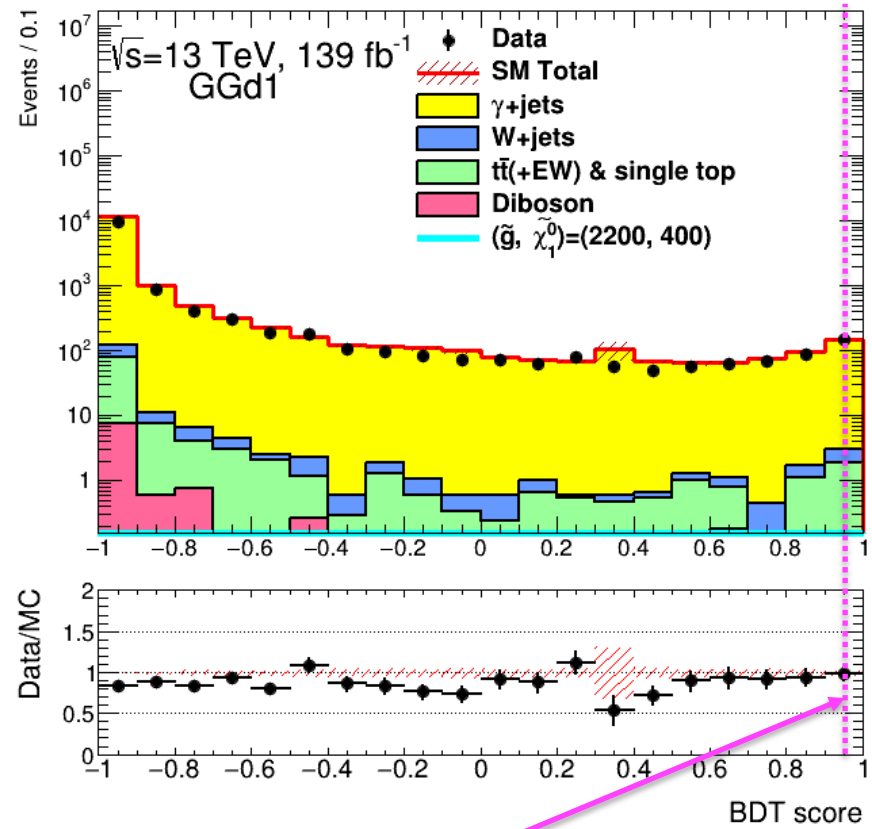
→ Use γ+Jet events as CR





Obtain similar kinematics!

Same BDT cut as SR



Compute normalization factor for Z+Jet events: 1.05

Validation of correlation

Use MC as training → Need to describe correlations by the data

$$\mathrm{corr}(x,y)_i = \frac{(x_i - \langle x \rangle)\cdot(y_i - \langle y \rangle)}{\sigma_x \sigma_y}$$

$$\langle x \rangle = \frac{\sum_i^N x_i}{N}, \quad \langle y \rangle = \frac{\sum_i^N y_i}{N}$$

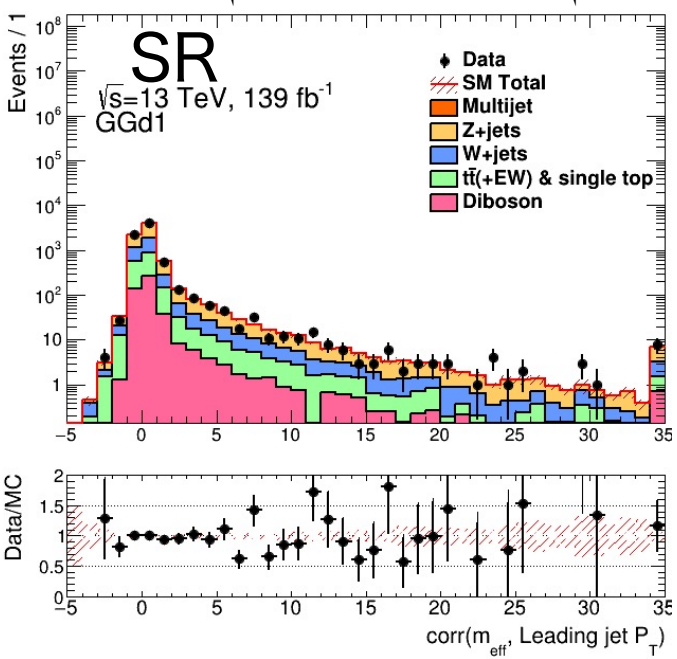
$$\sigma_x = \sqrt{\frac{1}{N}\sum_i^N {x_i}^2}, \quad \sigma_y = \sqrt{\frac{1}{N}\sum_i^N {y_i}^2}$$

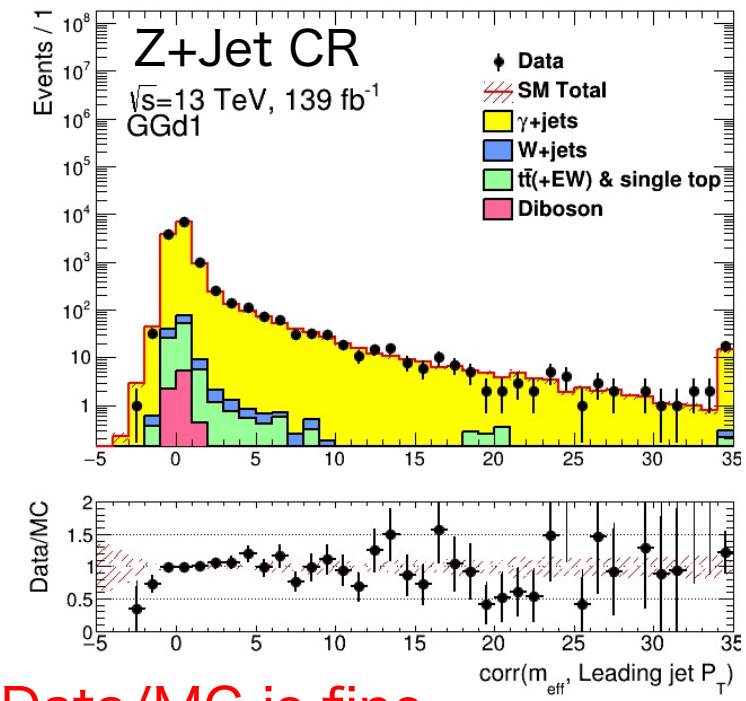
For example, M_{eff} and Leading jet

- Distribution with right tail

→ Has positive correlation

- Agreement with data

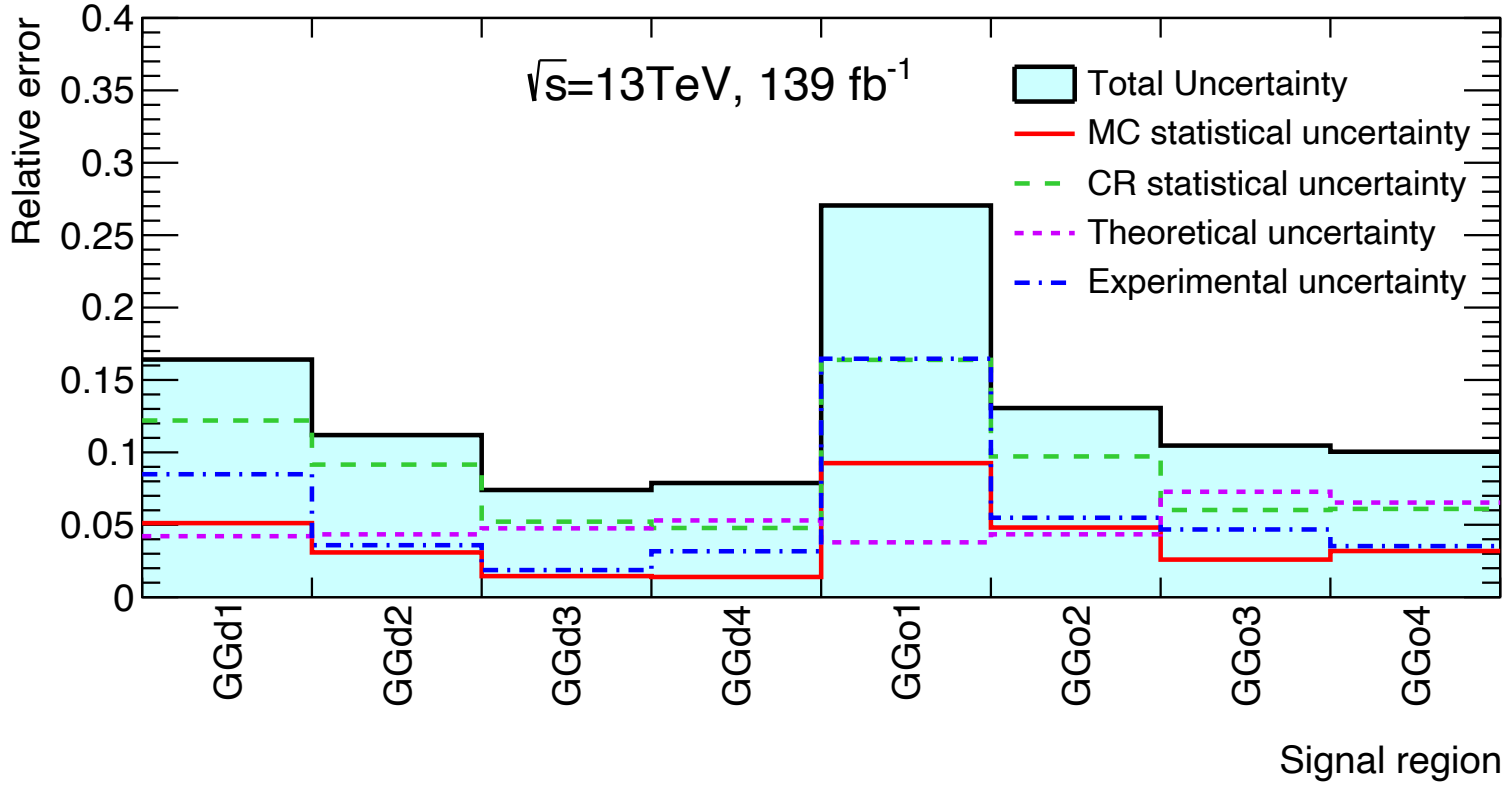




Checked all correlations → Data/MC is fine

Systematic uncertainty

A summary of systematic uncertainty about event yields



Major uncertainty: MC stat, CR stat, Theo and Exp uncertainty

- Low CR stat due to tight BDT score cut → dominant

Simplified Model

We would like to say “Exclude gluino ~XX GeV ”

- Limit MSSM parameter space ? (108 dimension)
  - It is impossible…

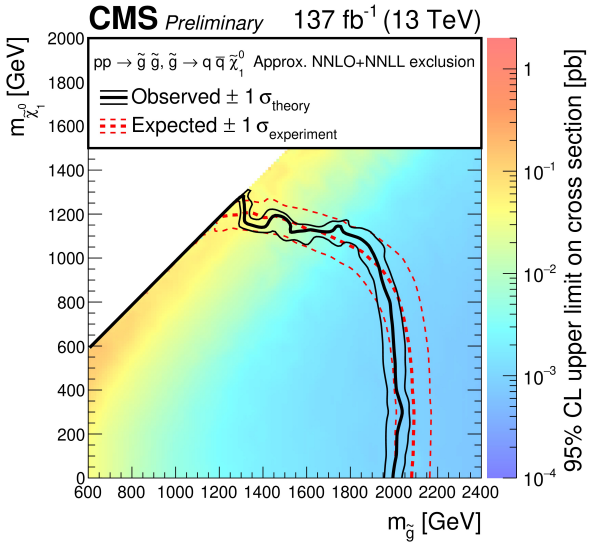
Almost parameters does not depend on sensitivity in exp. side

  - Possible to do the reduction

- Look at the space which only determines the acceptance of experiment
  - For example, SUSY mass * branching ratio

Evaluate excl. limit for all decay chains
~ Limit MSSM 108 dimension

→ Simplified Model approach



BDT configuration

Weight in training

- Each signal is weighted with each cross section.

→ It can avoid the bias of different phase spaces.

- The specific weights are not added in backgrounds

→ All background components are equally treated

Hyper parameter

- Use TMVA
- nTree = 800
- MaxDepth = 3
- nCuts = 20
- MinNodeSize = 2.5
- Shrinkage = 0.20
- Gradient Boost

2D distribution

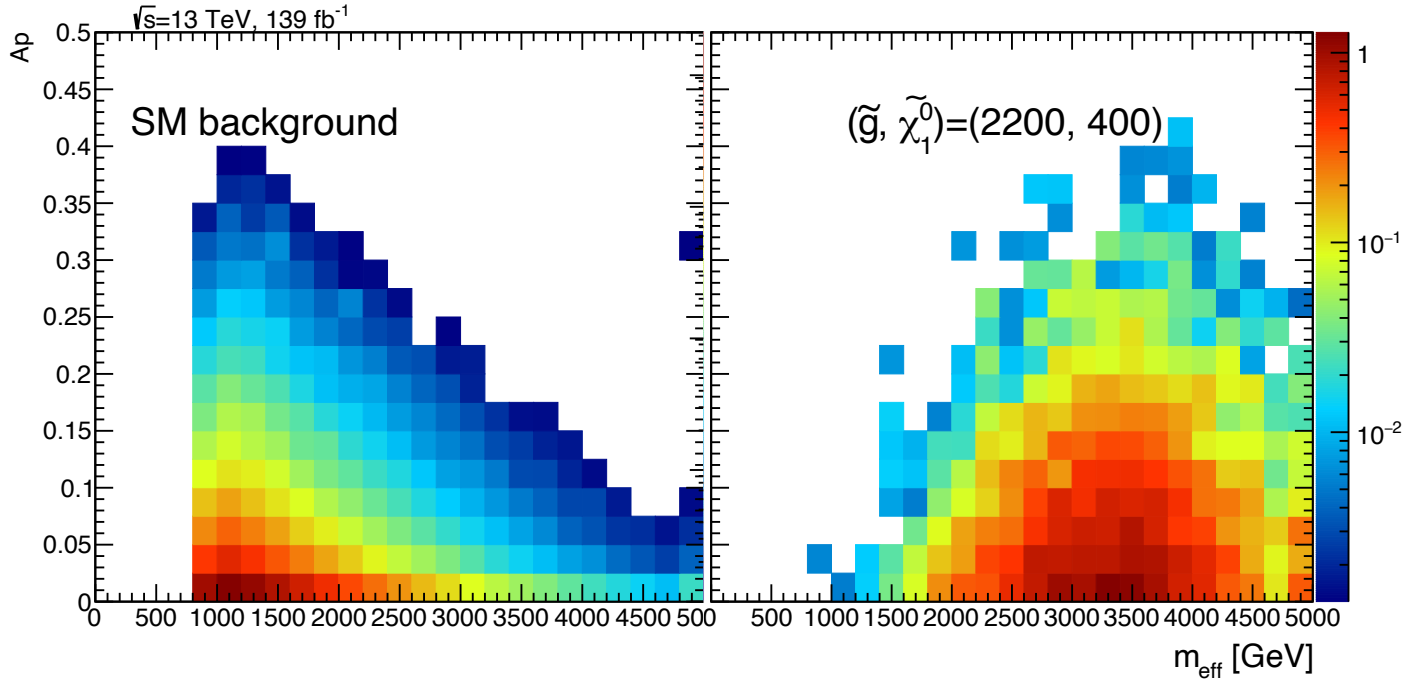
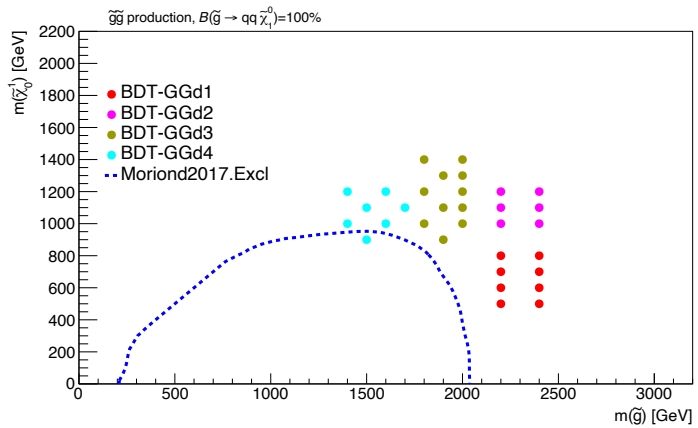


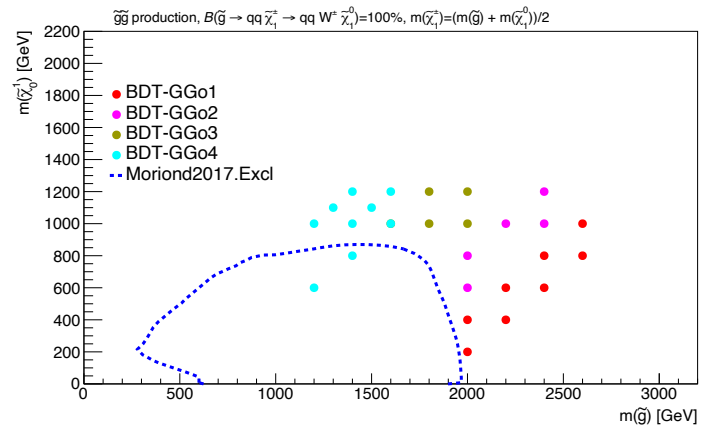
Figure 5.10: Two dimensional distribution of the effective mass and the aplanarity with the preselection. The last bin includes the overflow. The left figure shows the distribution in the SM background and the right figure shows the distribution in the gluino direct decay signal with $(m_{\tilde{g}},\ m_{\tilde{\chi}_0^1})$=(2200, 400) GeV.

If the effective mass is large for the signals, the signal events also have jets from initial state radiations and are close to planar event shape.
Hense, the broad aplanarity distribution is obtained for large missing transverse momentum

Input variable, training sample



(a) GG direct



(b) GG one-step

Figure 5.15: Illustration of training categories for gluino direct decay (a) and gluino one-step decay (b). For gluino direct decay (gluino one-step decay), the red point corresponds to GGd1 (GGo1), the magenta point corresponds to GGd2 (GGo2), the gold point corresponds to GGd3 (GGo3) and the light blue point corresponds to GGd4 (GGo4) training category. The blue dotted line shows the expected exclusion limit on the previous ATLAS result with the data of 36.1 fb^{-1} [34].

Table 5.5: Input variables for the gluino direct decay. Circle (◯) indicates the variable is used in input variables. The hyphen (-) indicates the variable is not used in input variables.

Input variables	GGd1	GGd2	GGd3	GGd4
E_T^{miss} [GeV]	◯	-	◯	◯
p_T(j) [GeV]	1st-4th	1st-4th	1st-4th	1st-4th
η(j)	1st-4th	1st-4th	1st-4th	1st-4th
Aplanarity	◯	◯	◯	-
m_{eff} [GeV]	◯	◯	◯	◯
Total number of input variables	11	10	11	10

Table 5.6: Input variables for the gluino one-step decay. Circle (◯) indicates the variable is used in input variables. The hyphen (-) indicates the variable is not used in input variables.

Input variables	GGo1	GGo2	GGo3	GGo4
E_T^{miss} [GeV]	-	-	-	◯
p_T(j) [GeV]	1st -5th	1st - 4th	1st - 3rd, 5th - 6th	1st - 4th
η(j)	1st - 5th	1st - 4th	1st - 3rd, 5th - 6th	1st - 4th
Aplanarity	◯	◯	◯	-
m_{eff} [GeV]	◯	◯	◯	◯
Total number of input variables	12	10	12	10

How to use the BDT classifier

I think TMVA user-guide can help it

- How to implement xml file into your C++ code
- If needed, I will show my code to github ..