A Search for SUSY in Multi-b Jet Events at ATLAS

Matthew Epland

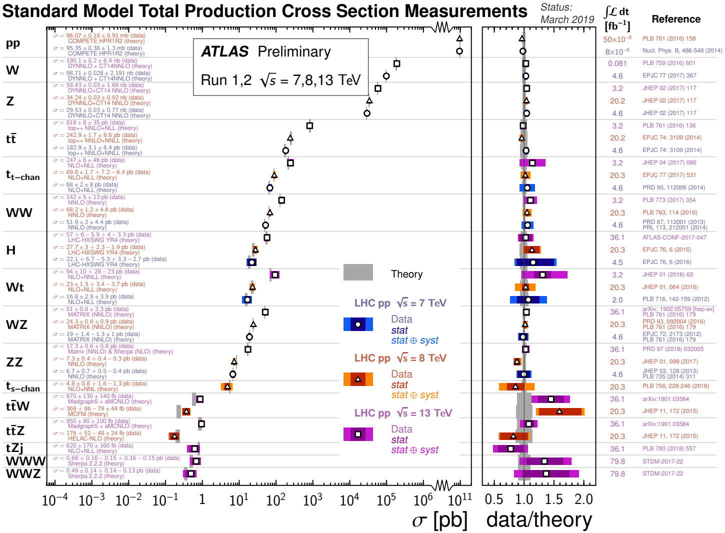
Advisor: Prof. Ayana Arce

Duke University

March 25th, 2019

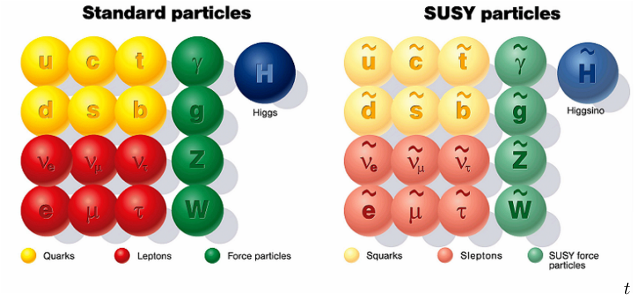
Our Standard Problems

- The SM works amazingly well over 12 orders of magnitude
- But, there are known issues: Dark Matter, Hierarchy Problem

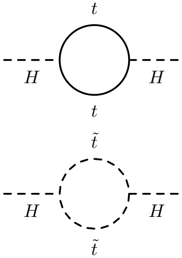


[1]

Enter Supersymmetry (SUSY)



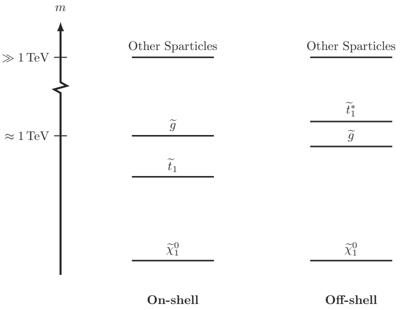
- What if we send Bosons ↔ Fermions?
- New supersymmetric loops of opposite sign cleanly fix the hierarchy problem
- Lightest supersymmetric particle (LSP) is a natural dark matter candidate

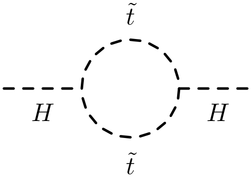


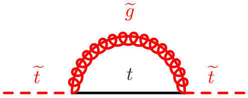
SUSY solves both issues in a "beautiful" way, and we should be able to see it in many channels at the LHC!

SUSY at LHC: Gluino Searches

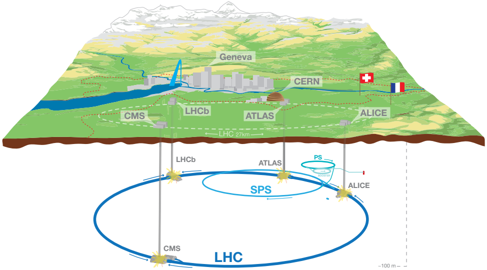
- Scalar Higgs renormalized by $\widetilde{t} \rightarrow$ Naturalness conditions on $m_{\widetilde{t}}$
- Scalar $\widetilde{t}$ renormalized by $\widetilde{g} \rightarrow$ Naturalness conditions on $m_{\widetilde{g}}$
- Results in light $\widetilde{g}$ and $\widetilde{t}$, (hopefully) can be produced at the LHC!
- $\widetilde{g}$ have large color charge $\rightarrow$ High production cross section!
- R-parity forces $\widetilde{g}$ pair production, final state LSPs $\rightarrow E_{\mathrm{T}}^{\mathrm{miss}}$

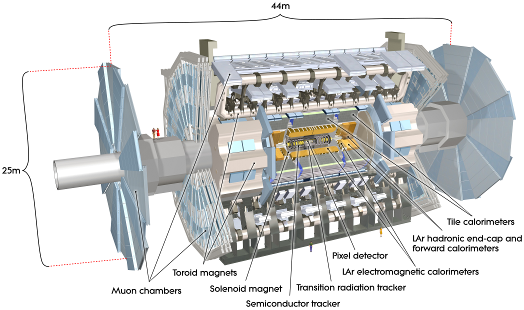






The LHC and ATLAS

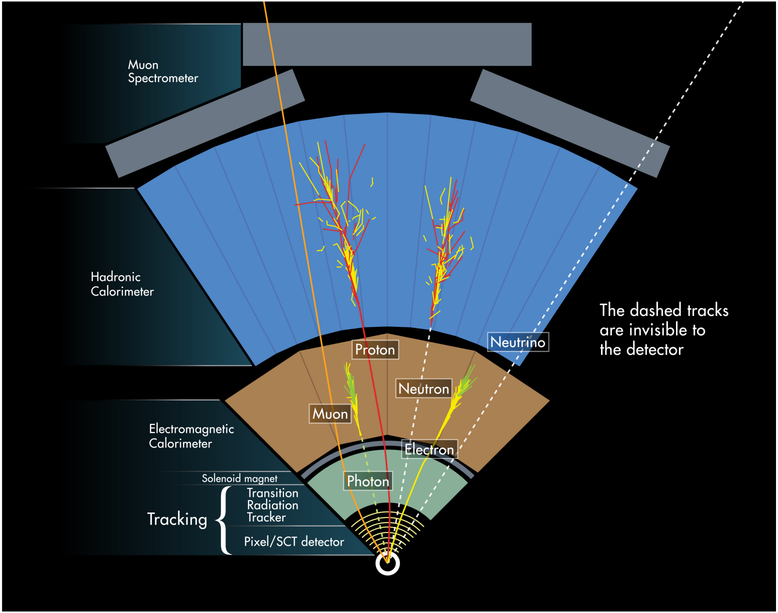




- Large Hadron Collider:
  - $\sqrt{s} = 13$ TeV pp collisions
  - 27 km circumference
  - ~ 1 GJ of stored energy
  - $\sim 10^{11}$ p cross / 25 ns
  - Total $L = 148.5\,\text{fb}^{-1}$
- ATLAS:
  - Hermetic general purpose particle physics detector
  - 2012 discovery of H
  - 2 T axial, 0.2 T–3.5 T toroidal magnetic fields
  - $\sim 10^8$ readout channels
  - Triggered down to 1 kHz
  - $\sim 2\,\text{GB}\,\text{s}^{-1}$, ~ 1 PB/year

[2, 3]

The ATLAS Detector

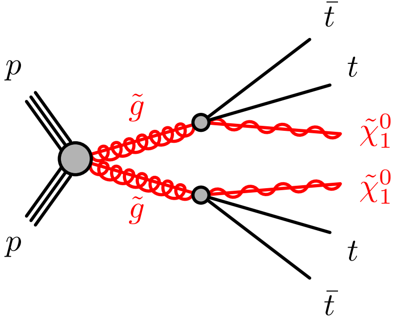


[4]

The Multi-b Search

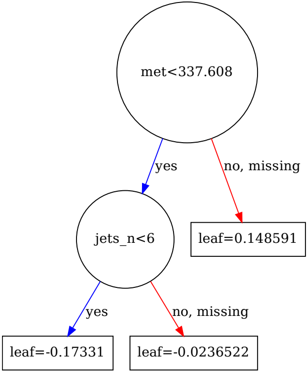
Multi-b Introduction

- pp collisions pair produce $\widetilde{g}$, decay to SM particles + LSPs ($\widetilde{\chi}_1^0$)
  - Assume off-shell $\widetilde{t}$ to simplify to two parameters; $m_{\widetilde{g}}$, $m_{\widetilde{\chi}_1^0}$
  - Assume 100 % BR, all other SUSY particles decouple
- Search for final states of ≥ 4 b-jets + $E_{\mathrm{T}}^{\mathrm{miss}}$ in 0L & 1L channels
- 79.8 fb^{-1} analyzed with a traditional search in July [5]
- Re-analyzed with BDT based approach for dissertation and Run 2 R&D



Boosted Decision Trees (BDT)

- BDT: Ensemble of trees of branching cuts
- Trees iteratively grown by minimizing the objective function
  - Includes training loss and regularization
  - 2nd order Taylor expansion
    $\rightarrow$ Gradient descent
- Each new tree complements existing trees
- Assigns each event to a leaf with weight w
  - Background-like $w < 0$, signal-like $0 < w$
- Sum weights from trees, take logistic function, get signal score $0 < \hat{y} < 1$
- Using the XGBoost library [6]
  github.com/dmlc/xgboost



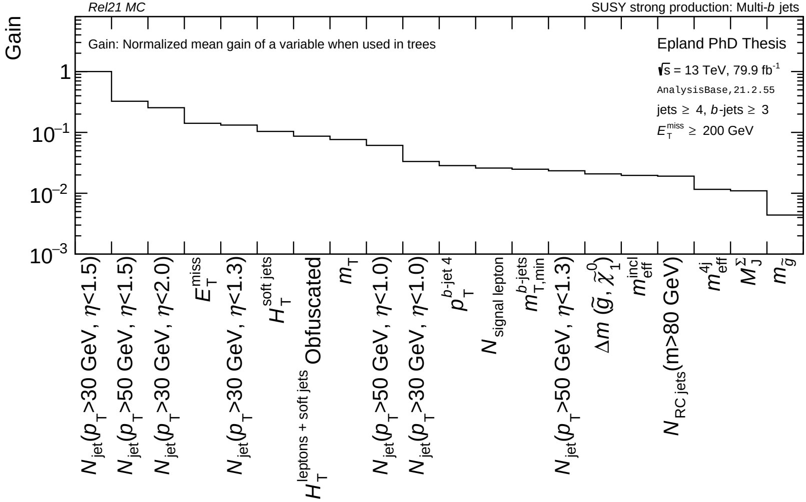
Data Samples and Preselections

- Gtt signal and SM background simulations:
  - $t\bar{t}$, single-top, $t\bar{t}+X$, W+jets, Z+jets, diboson
- Preselections:
  - Pass lowest unprescaled $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger, $E_{\mathrm{T}}^{\mathrm{miss}} \geq 200\,\mathrm{GeV}$
  - $N_{\mathrm{jet}} \geq 4$, $N_{b\text{-jet}} \geq 3$, $p_{\mathrm{T}}^{\mathrm{jet}\,1} > 30\,\mathrm{GeV}$
  - If $N_{\mathrm{sig\,lep}} = 0$ (0L), $\Delta\phi_{\mathrm{min}}^{4j} > 0.4$
- Events (73k) Breakdown: 53.6 % Train, 13.3 % Validation, 33 % Test
- Gtt signal and the BDT are parameterized by $m_{\tilde{g}}$, $m_{\tilde{\chi}_1^0}$
- For training uniformly assign background events $m_{\tilde{g}}$, $m_{\tilde{\chi}_1^0}$ values
  - Reweight so each mass point has equal sig, bkg ($w_{\mathrm{all}}^{\mathrm{bkg}}/N_{\mathrm{mass\ points}}^{\mathrm{sig}}$)

Training Variables

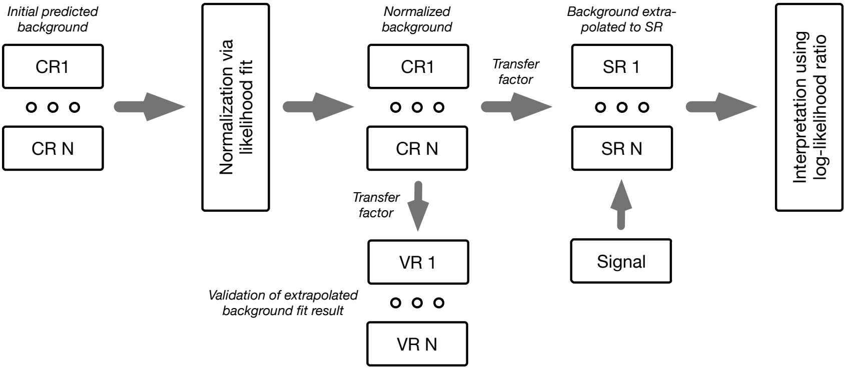
- Core set of 7 variables provide much of the performance:
  - $N_{\text{jet}}(p_{\text{T}} > 30\,\text{GeV}, \eta < 1.3)$, $N_{\text{jet}}(p_{\text{T}} > 30\,\text{GeV}, \eta < 1.5)$
  - $N_{\text{jet}}(p_{\text{T}} > 30\,\text{GeV}, \eta < 2.0)$, $N_{\text{jet}}(p_{\text{T}} > 50\,\text{GeV}, \eta < 1.5)$
  - $H_{\text{T}}^{\text{leptons + soft jets}}$ Obfuscated, m_{T}, $E_{\text{T}}^{\text{miss}}$
- Additional 11 provide the rest:
  - $N_{\text{sig lep}}$, $N_{\text{RC jet}}(m > 80\,\text{GeV})$, $N_{\text{jet}}(p_{\text{T}} > 30\,\text{GeV}, \eta < 1.0)$
  - $N_{\text{jet}}(p_{\text{T}} > 50\,\text{GeV}, \eta < 1.0)$, $N_{\text{jet}}(p_{\text{T}} > 50\,\text{GeV}, \eta < 1.3)$
  - $H_{\text{T}}^{\text{soft jets}}$, $m_{\text{eff}}^{\text{incl}}$, $m_{\text{T,min}}^{b\text{-jets}}$, M_J^{Σ}, m_{eff}^{4j}, $p_{\text{T}}^{b\text{-jet}\,4}$
- Plus $m_{\tilde{g}}$ and $\Delta m = m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ to parameterize the signal (20 total)

Variable Importance: Gain



Statistical Framework Overview

- To find SUSY and normalize backgrounds we construct signal regions and (low signal) control & validation regions



- ATLAS has standardized code (HistFitter [7]) to fit the regions, produce PDFs & transfer factors, handle systematics, make profile likelihood ratios, and run the hypothesis testing

Signal Region Selection: Parameter Points

Defining Parameter Points

- Need to define CRs, VRs, and SRs in $\hat{y}_{\text{SUSY, BDT}}$ before fitting
- BDT$(m_{\tilde{g}}, \Delta m)$: Selecting parameters effectively returns a particular BDT and associated output score $\hat{y}(m_{\tilde{g}}, \Delta m)$
  - Will also divide events into 0L and 1L lepton channels
- 115 signal mass points for training, can choose any for predictions
  - Don't want to have unwieldy fits or be impacted by the LEE
  - BDT interpolation not assured! Just use training points to be safe
- Select mass points to create a few $\hat{y}$, and thus SRs, which together have good coverage and performance across all masses
- How can we choose these points to target regions in mass space?

Defining Parameter Points

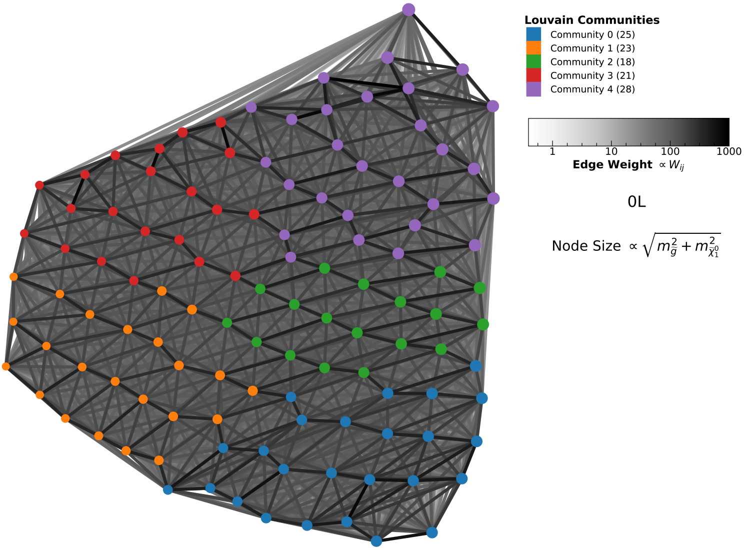
- Need metrics to describe how similar two points are:
  - RMSD: Compare BDT at different points via predicted $\hat{y}$ values
    - $\text{RMSD}(p_i, p_j) = \sqrt{\sum_k w_k \left(\hat{y}_k^{p_i} - \hat{y}_k^{p_j}\right)^2 / \sum_k w_k}$
  - Radius in mass space: Keeps regions compact
    - $R_m\left(p_i, p_j\right) = \sqrt{\left(\Delta m_{\tilde{g}}\right)^2 + \left(\Delta m_{\tilde{\chi}_1^0}\right)^2}$
- Need to invert: $1/(\text{RMSD} + \text{RMSD}_{\text{min}>0}) + 1/R_m$

- Relative measure, can't k-means cluster, but can graph!
- Find communities within resulting graph via Louvain method [8, 9]

- Find representative points by optimizing the significance
  - Loop over points in community, use each as $m_{\tilde{g}}$, Δm
  - Evaluate[1] Z at other points, compute $Z_{\text{metric}} = \langle \text{min}(5.0, Z_{\text{B}}) \rangle$
    - Re-weight signal at each point to have equal production cross section
  - Pick point which maximizes Z_{metric} across the community

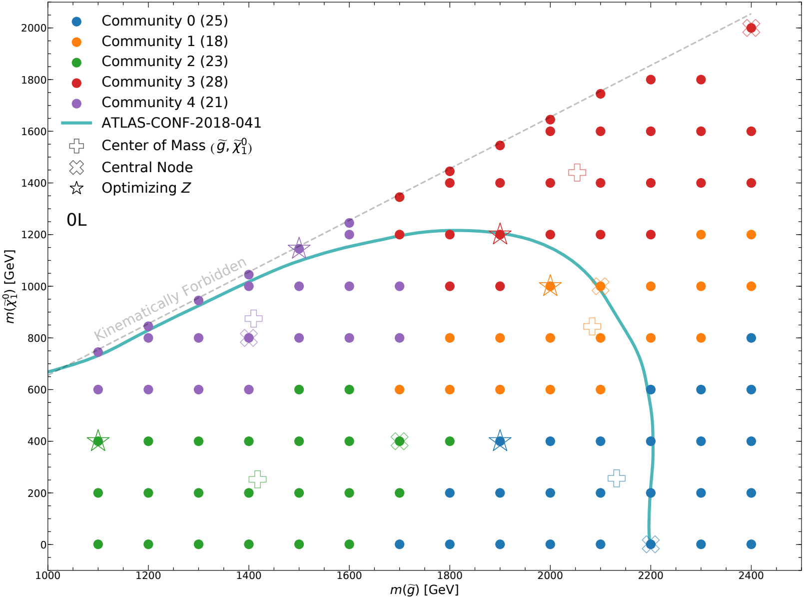
[1] Z_{B}: Approximate significance via BinomialExpZ, with bkg uncertainty of 50 %

Graph 0L

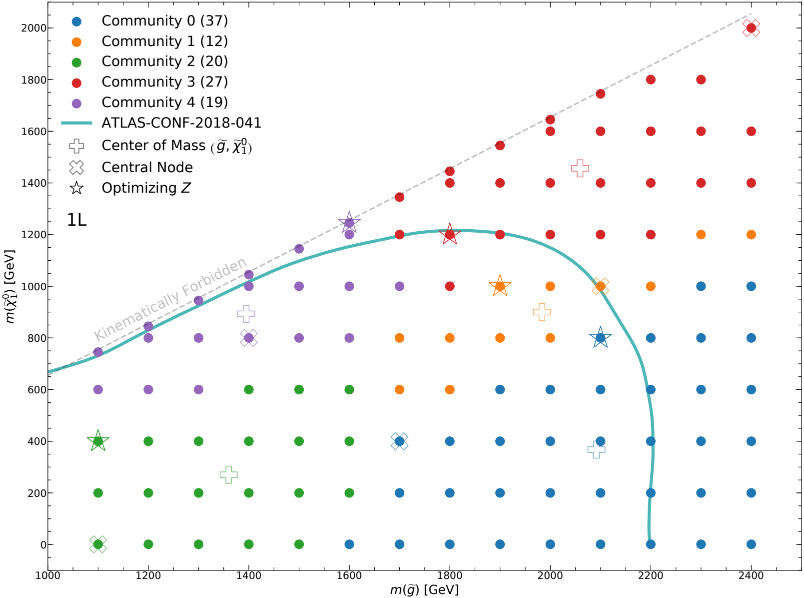


Louvain Communities
Community 0 (25)
Community 1 (23)
Community 2 (18)
Community 3 (21)
Community 4 (28)
1
10
100
1000
Edge Weight $\propto W_{ij}$
0L
Node Size $\propto \sqrt{m_{\tilde{g}}^2 + m_{\tilde{\chi}_1^0}^2}$

Parameter Points 0L



Parameter Points 1L



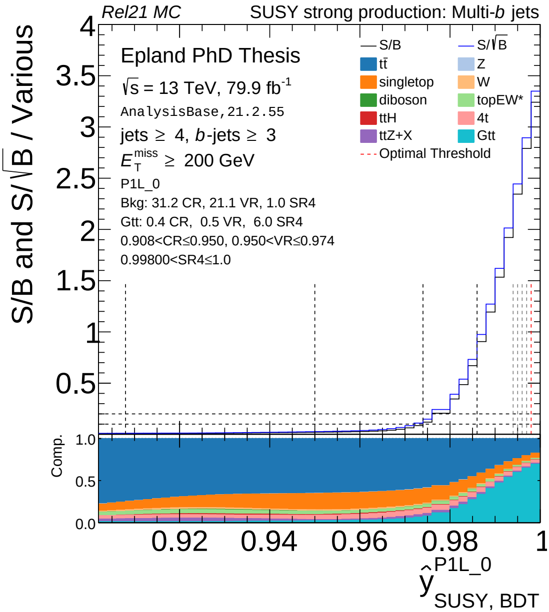
Signal Region Selection: $\hat{y}$ Regions

$\hat{y}$ Region Selection

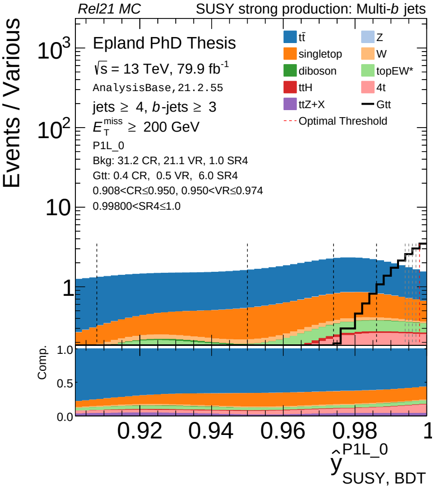
- Now we need to define the associated CR, VR, SR per parameter point
- Keep things simple and just use regions in $\hat{y}_{\text{SUSY, BDT}}$
- Would like to keep regions as close to 1.0 (signal-like) as possible, while having “enough” statistics in each
- This will make CRs & VRs with the most signal-like background events
- Can further sub-divide and shape fit within the resulting SR range
- Apply prior limits to Gtt production cross section where possible

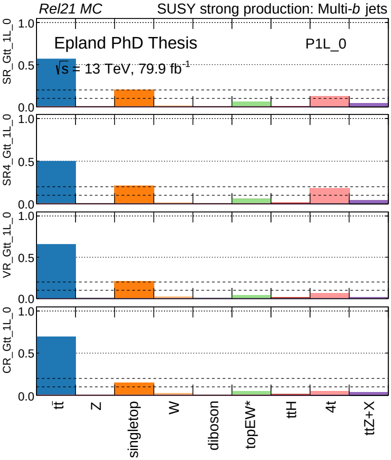
$\hat{y}$ Region Selection

- Plot S/B vs $\hat{y}$
  - Smoothed with Gaussian kernel
- Select regions:
  - SR: $S/B \geq 1.0$, $W_{\text{sig}} \geq 4.0$, $W_{\text{bkg}} \geq 1.0$
  - VR: $S/B \leq 0.2$, $S/\sqrt{B} \leq 3$, $W_{\text{bkg}} \geq 20$
  - CR: $S/B \leq 0.1$, $W_{\text{bkg}} \geq 30$
- Shape fit within SR:
  - Use Z_B optimal threshold for top bin
  - $+4$ bins at lower $\hat{y}$
  - Keep $W_{\text{bkg}} \approx 1.0$



$\hat{y}_{\text{SUSY, BDT}}$ & Bkg Composition





Fit Construction

Profile Likelihood Fit

- Likelihood function L for observed $\boldsymbol{n}$, nuisance parameters $\boldsymbol{\theta}^0$, $\boldsymbol{\theta}$
  - $P =$ Poisson, $\lambda\,$(MC expectations, $\boldsymbol{\theta}$); $G =$ standard Gaussian

$$L\left(\boldsymbol{n}, \boldsymbol{\theta}^0 \,|\, \boldsymbol{\theta}\right) = \prod_{i \in \mathrm{SR}} P\left(n_i,\, \lambda_i\right) \times \prod_{j \in \mathrm{CR}} P\left(n_j,\, \lambda_j\right) \times \prod_{k \in \mathcal{S}} G\left(\theta_k^0 - \theta_k\right)$$

- Make test statistic q from the log-likelihood ratio
  - $\mu_{\mathrm{sig}} = 0$ (1) for background only (signal + background) expectation
  - Find $\hat{\mu}_{\mathrm{sig}}$, $\hat{\boldsymbol{\theta}}$ which maximizes L absolutely, $\hat{\hat{\boldsymbol{\theta}}}$ for the chosen μ_{sig}

$$q_{\mu_{\mathrm{sig}}} = -2\log\left(L(\mu_{\mathrm{sig}}, \hat{\hat{\boldsymbol{\theta}}})/L(\hat{\mu}_{\mathrm{sig}}, \hat{\boldsymbol{\theta}})\right)$$

- Take asymptotic limit to get integral of the PDF, *i.e.* p-value
- "Profile" systematics / NPs to find most conservative $\boldsymbol{\theta}^0$ from data
- Create CL_s to better handle downward background fluctuations

$$CL_s = p_{s+b}\,/\,(1 - p_b)$$

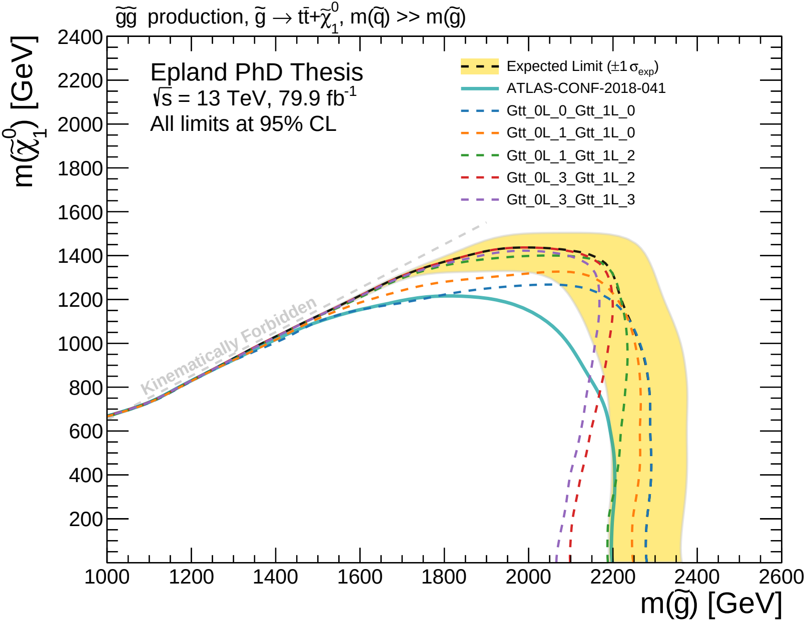
Combining Signal Regions

- First drop individual regions as needed; those which have:
  - Large VR pulls (Gtt_1L_1)
  - Large theory systematics, per SR bin (removes Gtt_0L_4)
- Take each combination of orthogonal 0L & 1L regions as multi-bin fit
  - Include all associated regions in one likelihood function
- Treat each 0L & 1L combination as non-orthogonal single-bins
  - Take maximum of independent CL_s contours
- Drop combinations which do not push the exclusion limit
  - Removes Gtt_0L_2, Gtt_1L_4

Region Definitions

Region	$m_{\tilde{g}}$	$m_{\tilde{\chi}_1^0}$	$N_{\text{sig lep}}$	Type	$\hat{y}$ Selection
Gtt_0L_0	1900 GeV	400 GeV	0	CR	$0.88400 \leq \hat{y}_{0L\_0} < 0.94200$
				VR	$0.94200 \leq \hat{y}_{0L\_0} < 0.97200$
				SR4	$0.99836 \leq \hat{y}_{0L\_0} \leq 1.00000$
Gtt_0L_1	2000 GeV	1000 GeV	0	CR	$0.91800 \leq \hat{y}_{0L\_1} < 0.95000$
				VR	$0.95000 \leq \hat{y}_{0L\_1} < 0.97200$
				SR4	$0.99717 \leq \hat{y}_{0L\_1} \leq 1.00000$
Gtt_0L_3	1900 GeV	1200 GeV	0	CR	$0.93000 \leq \hat{y}_{0L\_3} < 0.95400$
				VR	$0.95400 \leq \hat{y}_{0L\_3} < 0.97000$
				SR4	$0.99621 \leq \hat{y}_{0L\_3} \leq 1.00000$
Gtt_1L_0	2100 GeV	800 GeV	≥ 1	CR	$0.90800 \leq \hat{y}_{1L\_0} < 0.95000$
				VR	$0.95000 \leq \hat{y}_{1L\_0} < 0.97400$
				SR0	$0.99400 \leq \hat{y}_{1L\_0} < 0.99500$
				SR1	$0.99500 \leq \hat{y}_{1L\_0} < 0.99600$
				SR3	$0.99700 \leq \hat{y}_{1L\_0} < 0.99800$
				SR4	$0.99800 \leq \hat{y}_{1L\_0} \leq 1.00000$
Gtt_1L_2	1100 GeV	400 GeV	≥ 1	CR	$0.93400 \leq \hat{y}_{1L\_2} < 0.95200$
				VR	$0.95200 \leq \hat{y}_{1L\_2} < 0.96600$
				SR0	$0.99300 \leq \hat{y}_{1L\_2} < 0.99400$
				SR1	$0.99400 \leq \hat{y}_{1L\_2} < 0.99500$
				SR3	$0.99600 \leq \hat{y}_{1L\_2} < 0.99706$
				SR4	$0.99706 \leq \hat{y}_{1L\_2} \leq 1.00000$
Gtt_1L_3	1800 GeV	1200 GeV	≥ 1	CR	$0.91800 \leq \hat{y}_{1L\_3} < 0.94400$
				VR	$0.94400 \leq \hat{y}_{1L\_3} < 0.96200$
				SR4	$0.99562 \leq \hat{y}_{1L\_3} \leq 1.00000$

Individual Expected Exclusion Limits

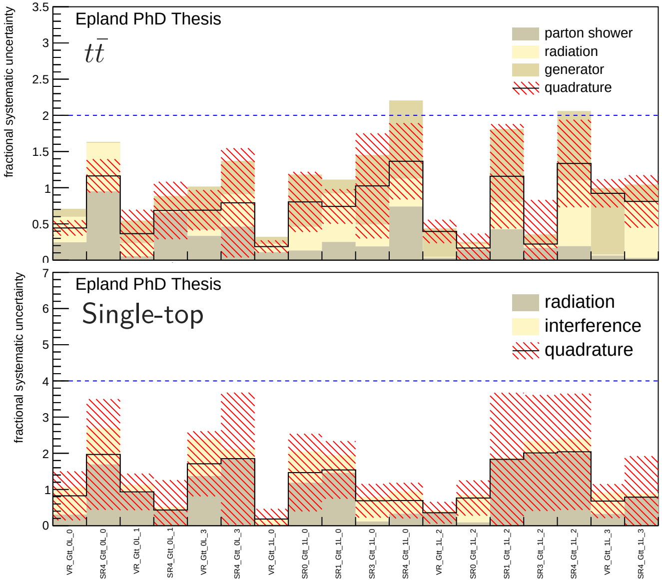


Systematics

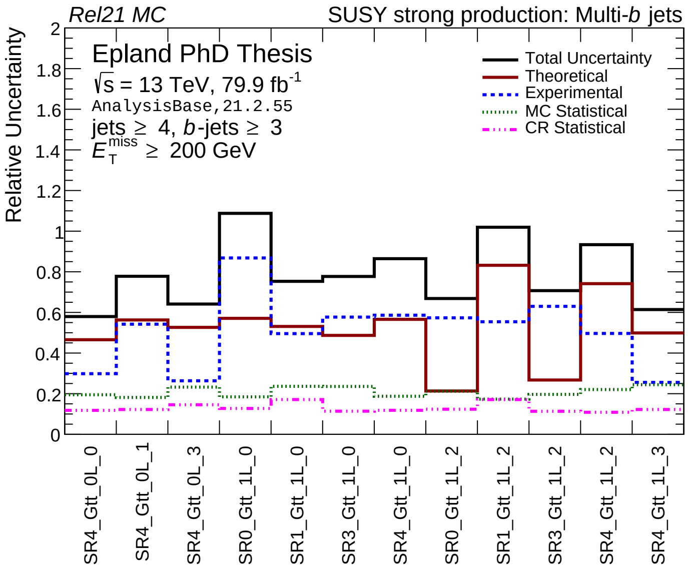
Systematic Uncertainties

- Follow the prior multi-b search approach [5]
- Experimental systematics include: JER, JES, b-tagging, JVT...
- Main theory systematics are estimated from truth samples
  - $t\bar{t}$: Generator, Parton Shower, Radiation
  - Single-top: Radiation, Interference between $t\bar{t}$ and Wt (via $WWbb$)
  - Also have theory systematics on Gtt production cross section, W/Z+jets, $t\bar{t}+X$, diboson backgrounds...
- To avoid poor nuisance parameter constraints in fit, remove SR bins where $t\bar{t}$ (single-top) theory uncertainty + error is $\geq 200\,\%$ ($\geq 400\,\%$)
  - $\approx \pm 1$ event for $\approx$ 0.5 ($\approx$ 0.25) expected $t\bar{t}$ (single-top) per SR bin
  - Truth samples are not reconstructed, nominal distributions are different
  - Some SRs have poor MC statistics
  - Done before SR unblinding, no bias between regions based on data

Theory Systematics



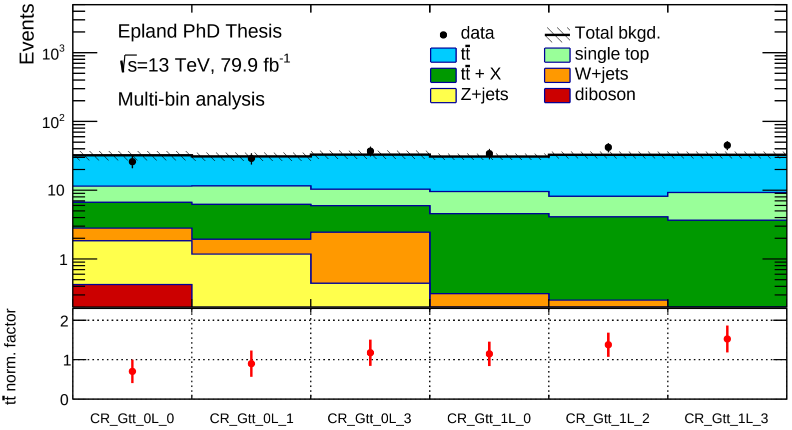
Relative Uncertainties[2]



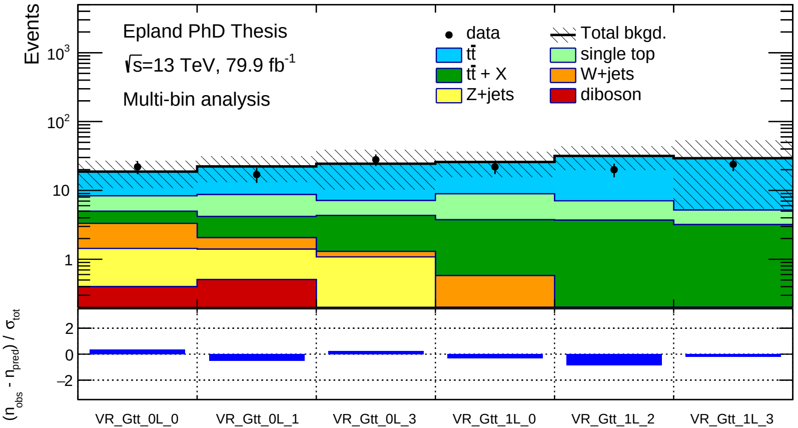
[2]Uncorrelated sums of quadrature here; correlations properly treated in HistFitter

Background Fits

CR $t\bar{t}$ Normalization Factors

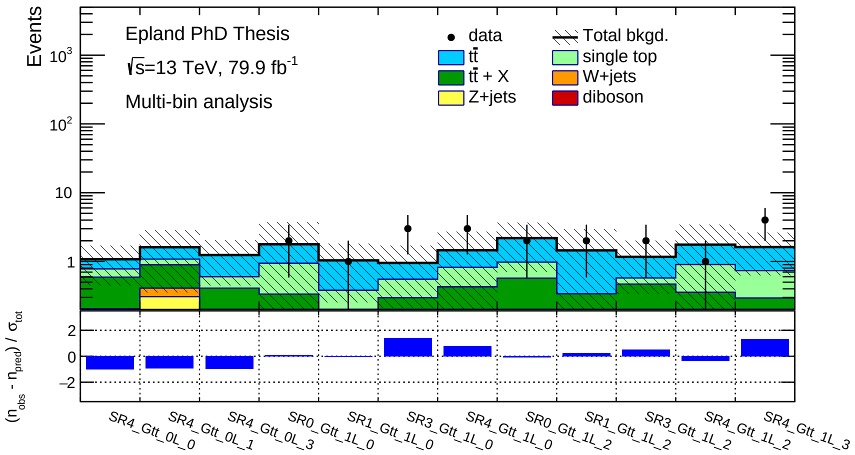


VR Pulls


Epland PhD Thesis
$\sqrt{s}$=13 TeV, 79.9 fb^{-1}
Multi-bin analysis
data
Total bkgd.
$t\bar{t}$
single top
$t\bar{t}$ + X
W+jets
Z+jets
diboson
Events
10^3
10^2
10
1
2
0
−2
$(n_{obs} - n_{pred}) / \sigma_{tot}$
VR_Gtt_0L_0
VR_Gtt_0L_1
VR_Gtt_0L_3
VR_Gtt_1L_0
VR_Gtt_1L_2
VR_Gtt_1L_3

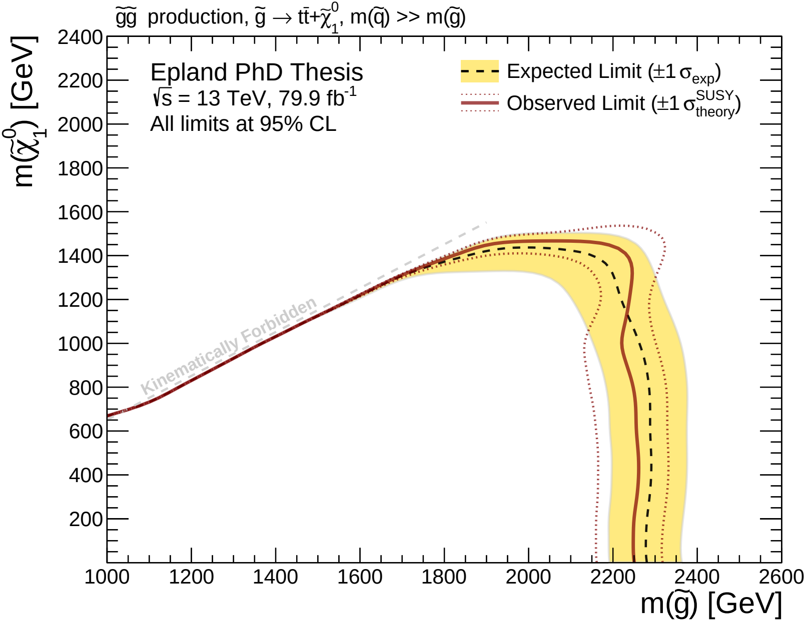
Results

SR Pulls

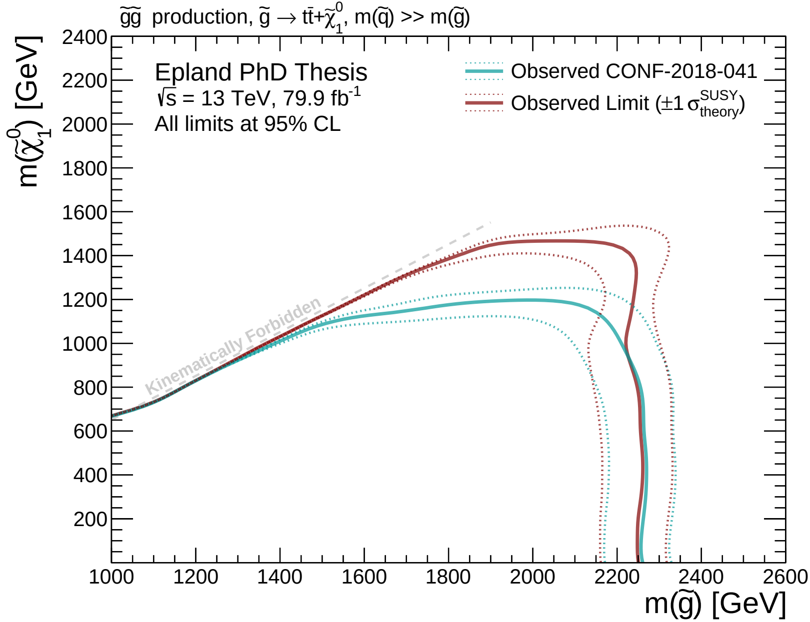


Events
Epland PhD Thesis
√s=13 TeV, 79.9 fb-1
Multi-bin analysis
data
tt̄
tt̄ + X
Z+jets
Total bkgd.
single top
W+jets
diboson
(n_obs - n_pred) / σ_tot
SR4_Gtt_0L_0
SR4_Gtt_0L_1
SR4_Gtt_0L_3
SR0_Gtt_1L_0
SR1_Gtt_1L_0
SR3_Gtt_1L_0
SR4_Gtt_1L_0
SR0_Gtt_1L_2
SR1_Gtt_1L_2
SR3_Gtt_1L_2
SR4_Gtt_1L_2
SR4_Gtt_1L_3

Exclusion Limit



Exclusion Limit Comparison

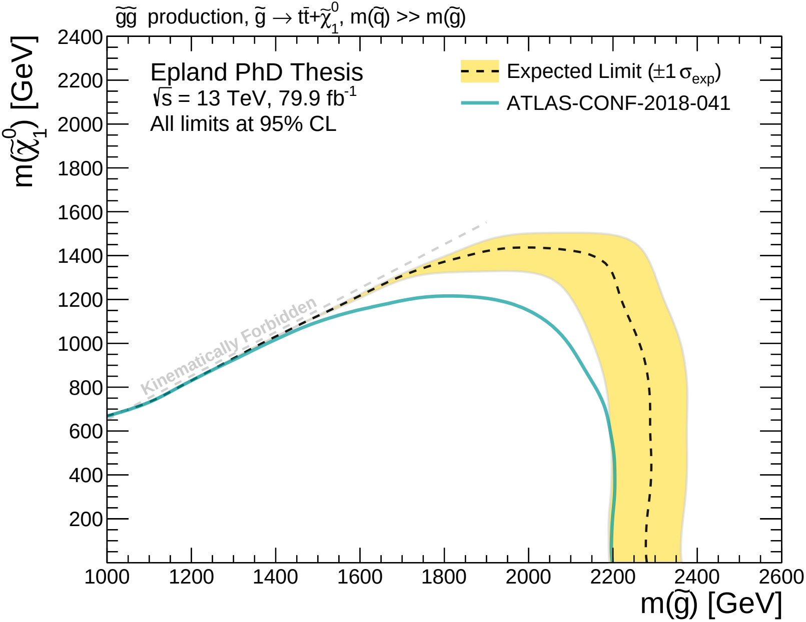


Conclusions

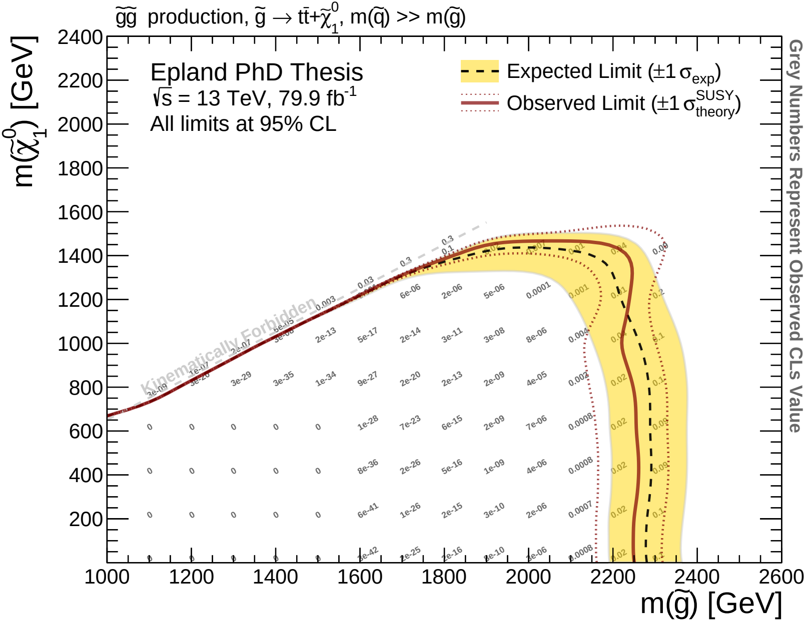
- BDT re-analysis of the 79.9 fb^{-1} dataset complete!
  - No major performance differences between the train and test sets
  - Satisfactory data / MC agreement for all input variables
  - VR pulls are small in both sets
  - Nuisance parameter pulls and constraints look good
- No SUSY, but observed limit matches expected limit fairly well!
  - Expected limit improved by 100 GeV–200 GeV in $m_{\tilde{g}}$, $m_{\tilde{\chi}_1^0}$
  - Observed limit expanded by 250 GeV to $\approx$ 1.4 TeV in $m_{\tilde{\chi}_1^0}$
- Contributed BDT approach to multi-b R&D for the full Run 2 search
  - Parameterizing the BDT was successful
  - Found new useful kinematic variables
  - Developed new methods of creating SRs from the BDT

Backup

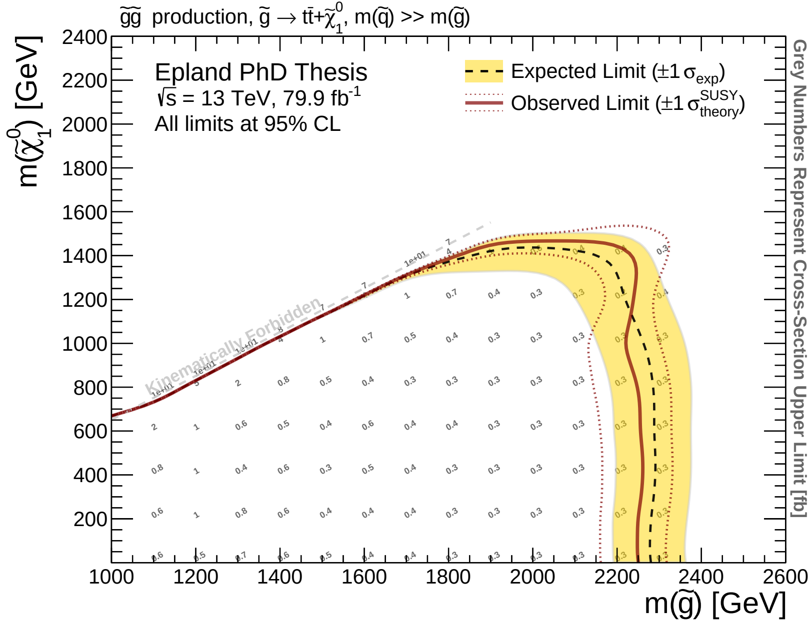
Exclusion Limit Comparison (Expected)



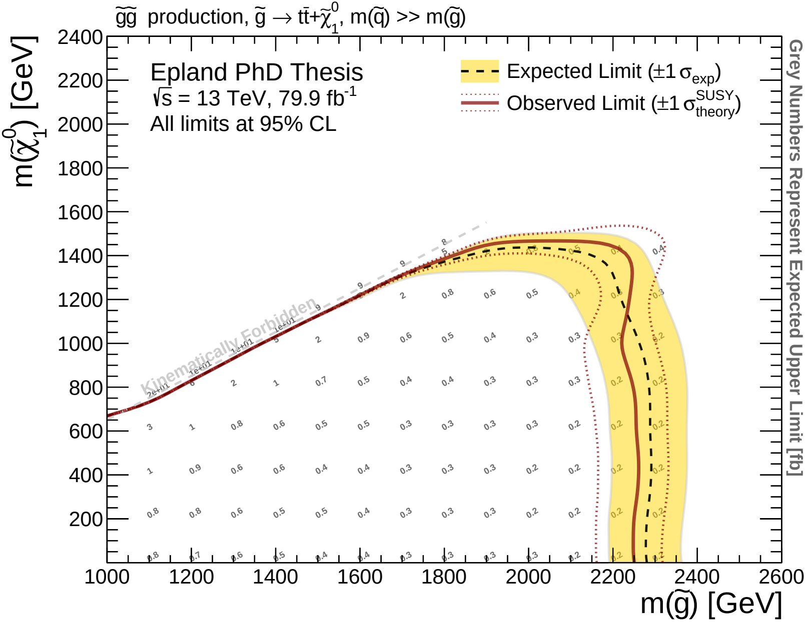
Exclusion Limit (CL_s Values)



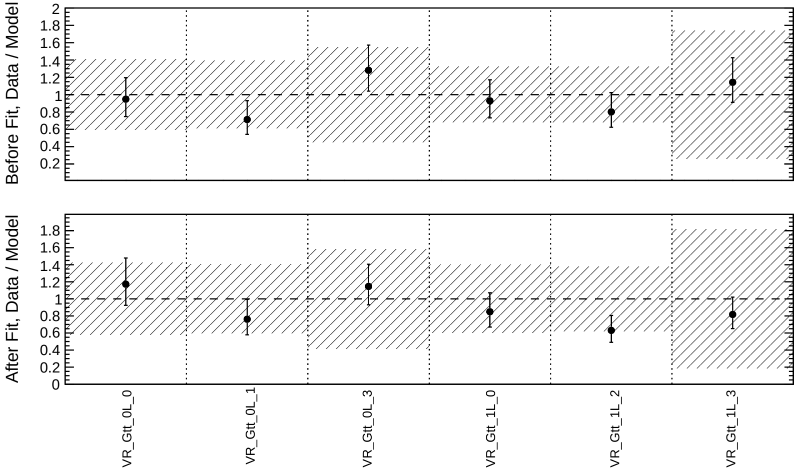
Exclusion Limit (Cross Sections)



Exclusion Limit (Expected Cross Sections)

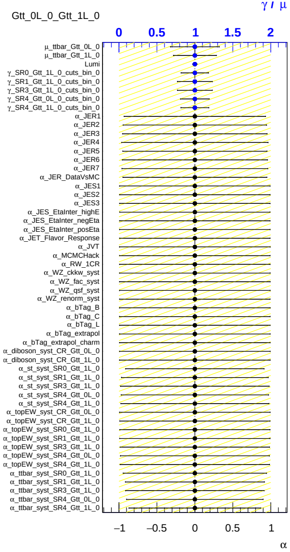


VR Data / MC Before and After Fit

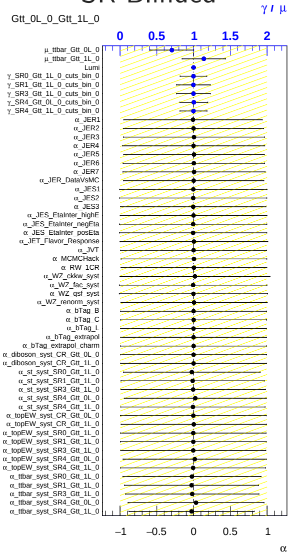


Fit Nuisance Parameters

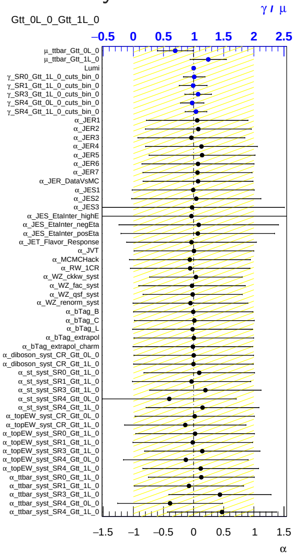
Fully Blinded



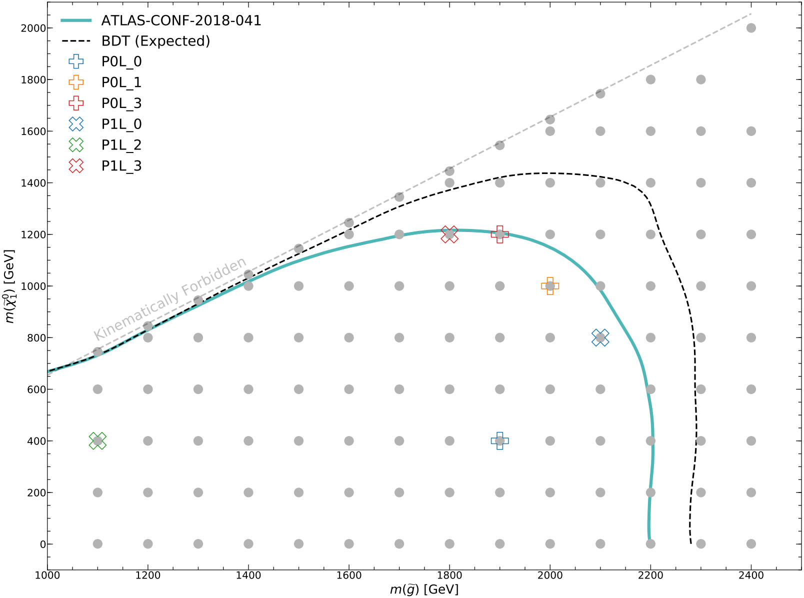
SR Blinded



Fully Unbinded



Location of Selected Parameter Points



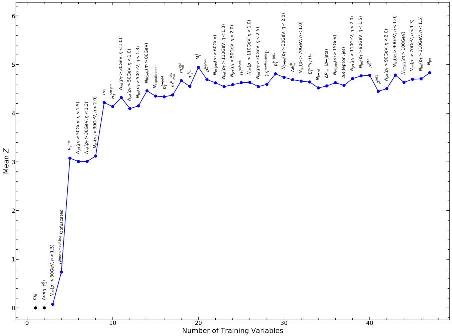
Observing Zero 0L Events

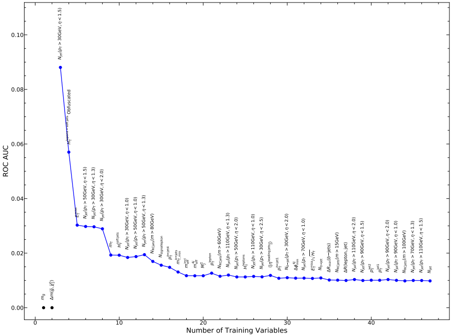
- No events were observed for any of the three 0L SR bins
- Looked at MC samples to find the expected number of bkg events
  - Careful not to double count since they are non-orthogonal bins
- Poisson probability is then $p\,(n = 0\,|\,\lambda) = e^{-\lambda}$
- $p\,(0\,|\,\lambda) = 0.15$ on the train set
  - With max (min) of 0.16 (0.07) on the systematics
- Probability of observing zero 0L events by chance is not insignificant
- Re-examination of differences between 0L and 1L in future work may be prudent

Training Variables and Hyperparameters

Training Variables Selection

- Iteratively:
  - Train BDT with all 70 potential variables
  - Evaluate performance: Z_{B}, mean Z_{B}, ROC AUC
  - Remove lowest gain variable (that's not a parameter), repeat
- Review results and decide on the best set of training variables
  - Can get into local mins, requires some babysitting / judgment calls





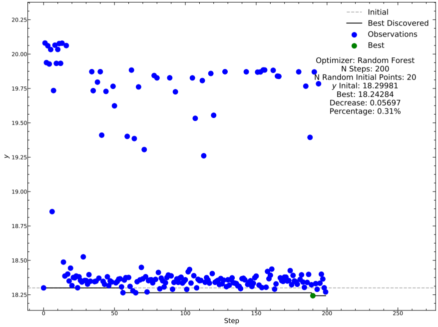
Hyperparameters

- Mostly using XGBoost defaults:
  - Objective function: binary logistic
  - Learning rate[1] $\eta = 0.0722758514998$
  - Max number of trees / boosting rounds $K_{\text{max}} = 200$
  - Number of early stopping rounds $= 10$
    - Validation threshold[1] $= 0.769402992287$
    - Best number of rounds $K = 197$
    - Trains in ≈ 2 min on 4 CPU cores
  - Max depth of tree[1] $= 7$

[1] Optimized

Hyperparameter Bayesian Optimization

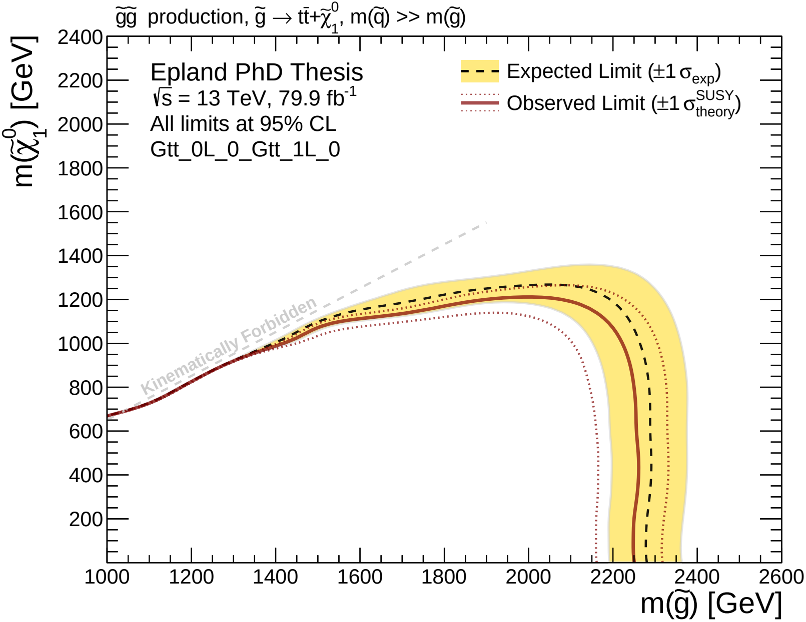
- Use BO when a function f is expensive & can't compute the gradient
  - $f(\text{hyperparameters}) =$ train BDT, evaluate, return mean Z
- Sample prior distribution, infer posterior, iterate many times (slow!)
- Random Forest regressor or Gaussian Process
  - GP is a maximum likelihood method
  - Start from a kernel: RBF, Matern, white noise, Gaussian noise
- Done in Scikit-Optimize, see MLHEP 2018 slides & example below



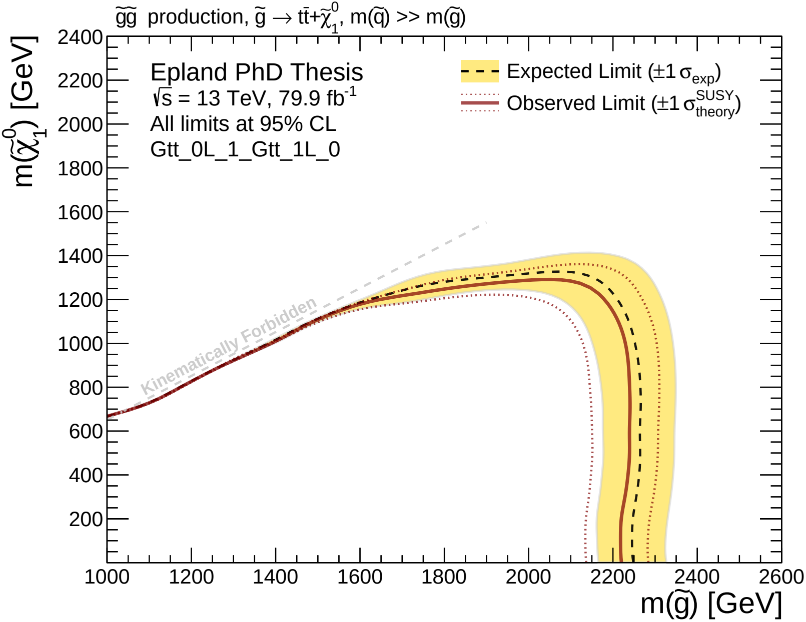

```python
from skopt import Optimizer

bo = Optimizer(dimensions=(3,10), # BDT depth
  n_initial_points=10, acq_func='EI',
  base_estimator = GaussianProcessRegressor(
  kernel = RBF(length_scale_bounds=[1e-3, 1e3])
    + WhiteKernel(noise_level=1e-5,
      noise_level_bounds=[1e-6,1e-2])))

for i in range(n_steps):
  x = bo.ask() # get next test point
  y = target_function(x) # evaluate function
  bo.tell(x, y) # update with the result

best_result_index = np.argmin(bo.yi)
x_best = bo.Xi[best_result_index]
```

Individual Exclusion Limits

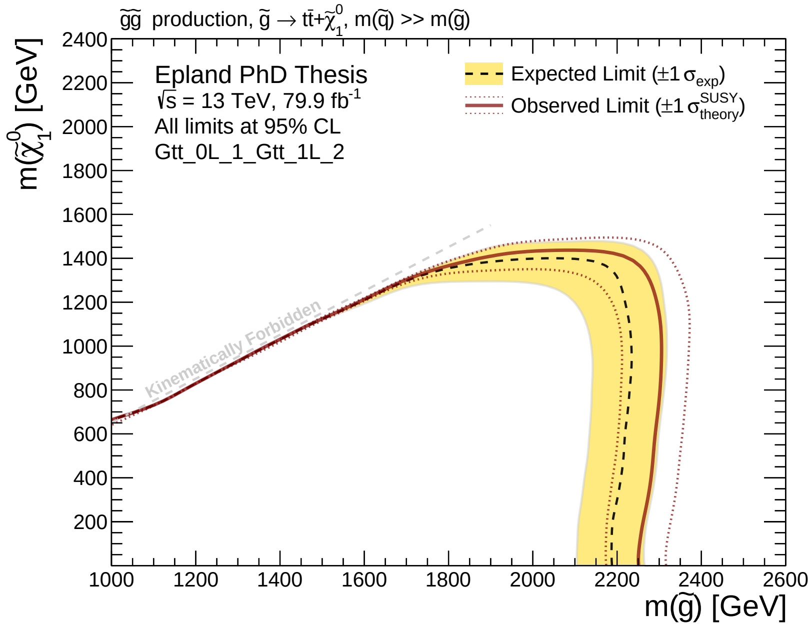
Gtt_0L_0_Gtt_1L_0 Exclusion Limit



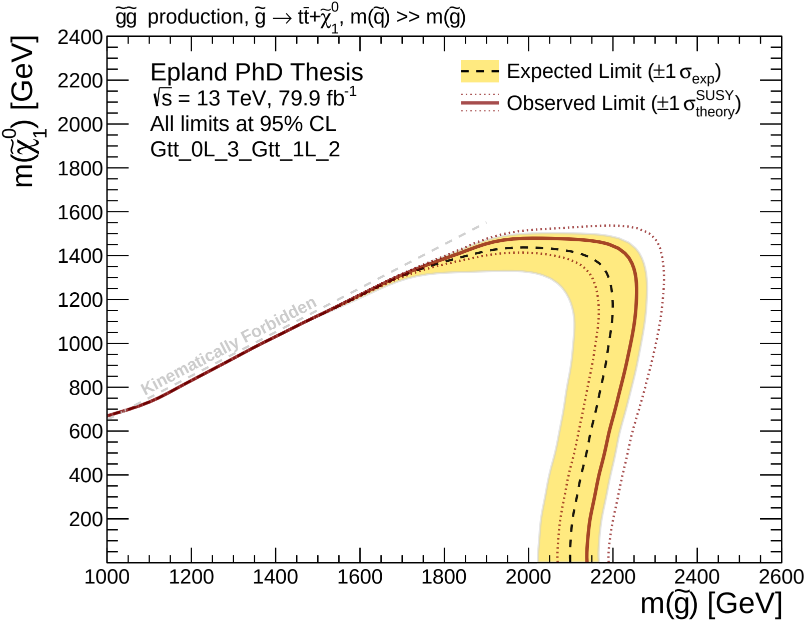
Gtt_0L_1_Gtt_1L_0 Exclusion Limit



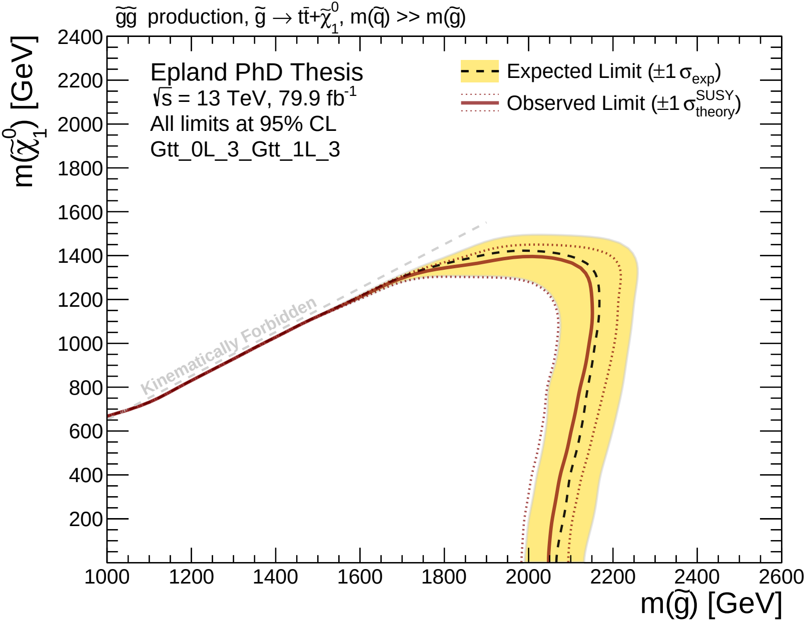
Gtt_0L_1_Gtt_1L_2 Exclusion Limit



Gtt_0L_3_Gtt_1L_2 Exclusion Limit



Gtt_0L_3_Gtt_1L_3 Exclusion Limit

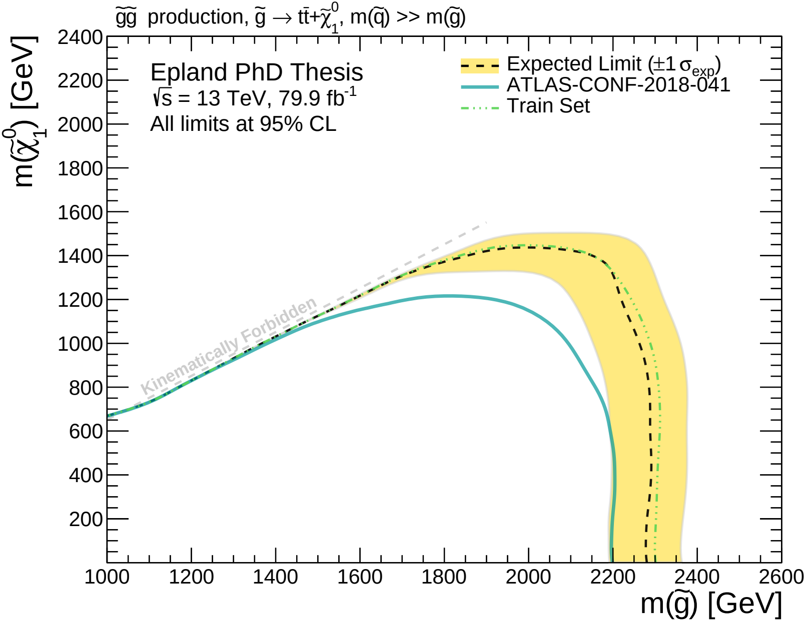


Overfitting

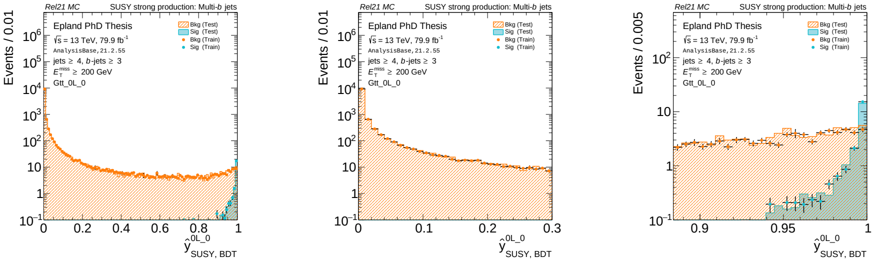
Overfitting

- Used a test/train split, only looked at train set for input variable, hyperparameter, signal region optimization, and HF R&D
    - Compared test/train $\hat{y}$ distributions, no red flags
    - Did 5-fold cross-validation to see how accuracy varied, no red flags
- Run on test set after choosing regions, before unblinding
    - Minor drop in exclusion limit performance but not concerning ($\sim$20 GeV)
- Only using the test set in final presented results
    - Unless otherwise noted

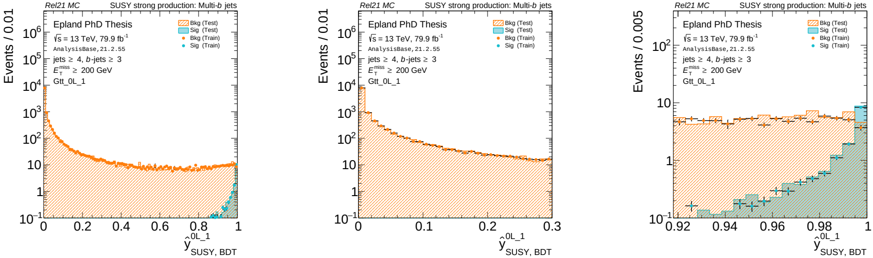
Train Set Exclusion Limit



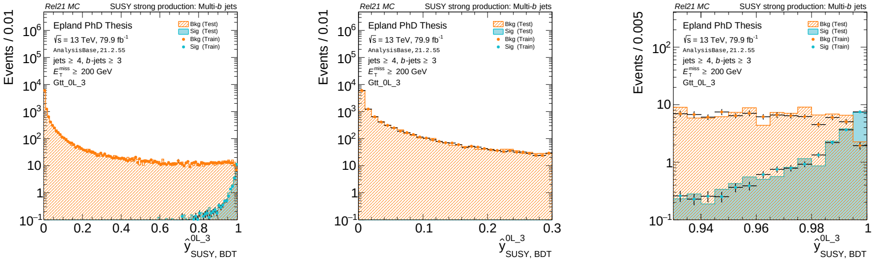
Test and Train $\hat{y}$ Gtt_0L_0



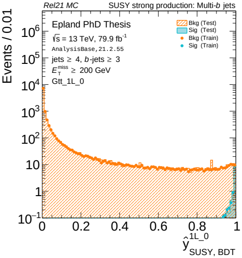
Test and Train $\hat{y}$ Gtt_0L_1

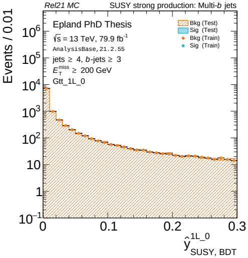


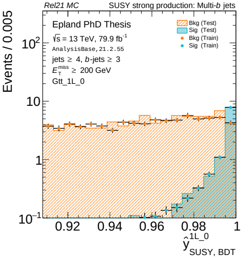
Test and Train $\hat{y}$ Gtt_0L_3



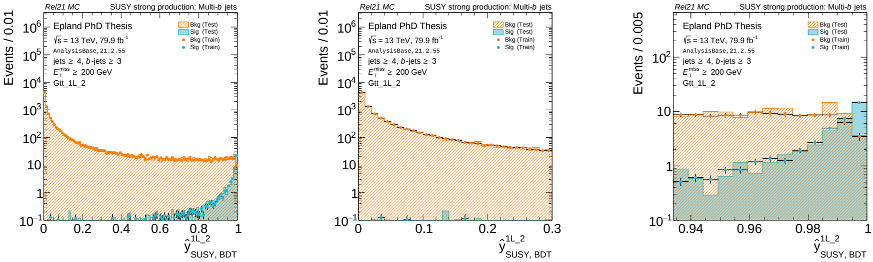
Test and Train $\hat{y}$ Gtt_1L_0



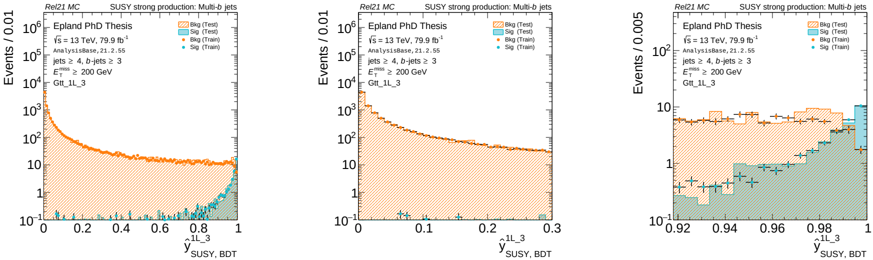




Test and Train $\hat{y}$ Gtt_1L_2

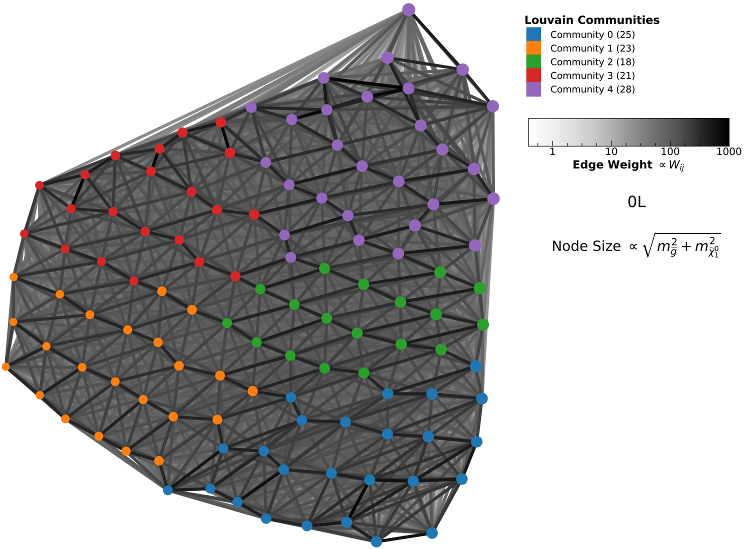


Test and Train $\hat{y}$ Gtt_1L_3



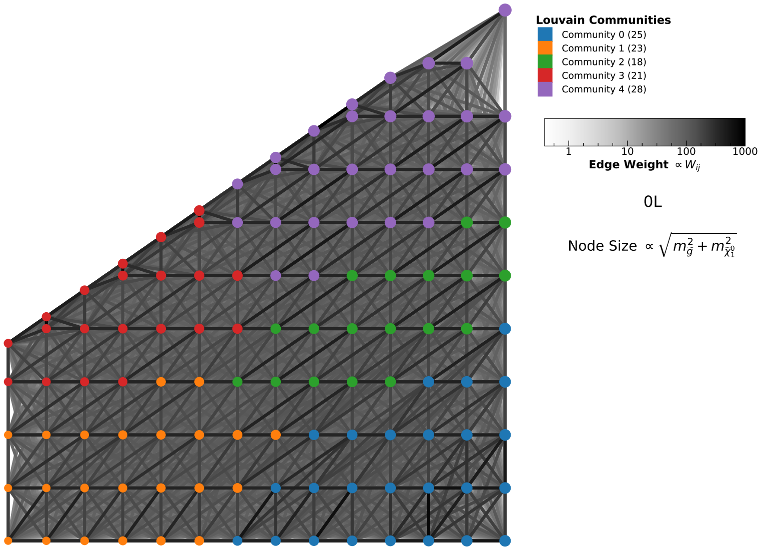
Parameter Point Graphs and Components

Graph 0L

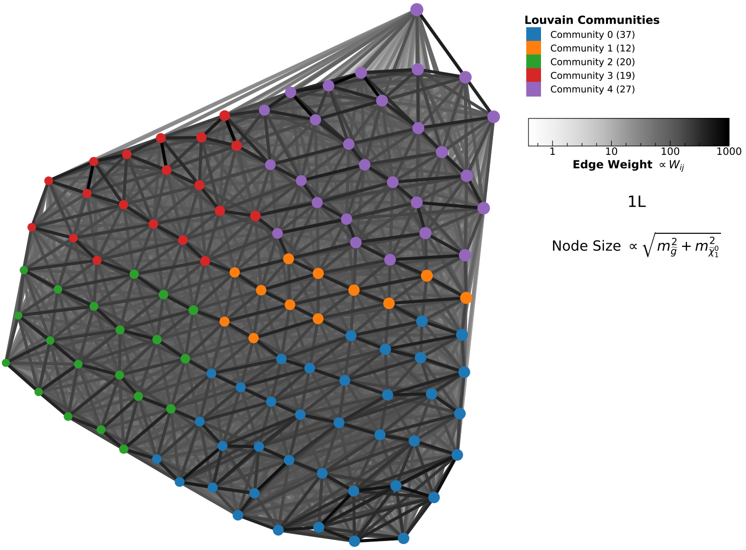


Louvain Communities
Community 0 (25)
Community 1 (23)
Community 2 (18)
Community 3 (21)
Community 4 (28)
1
10
100
1000
Edge Weight $\propto W_{ij}$
0L
Node Size $\propto \sqrt{m_{\tilde{g}}^2 + m_{\tilde{\chi}_1^0}^2}$

Graph (Mass Grid) 0L

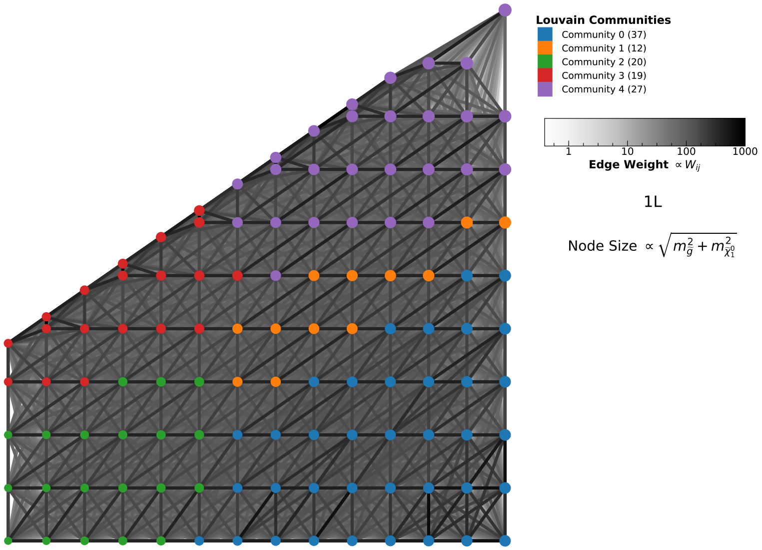


Graph 1L

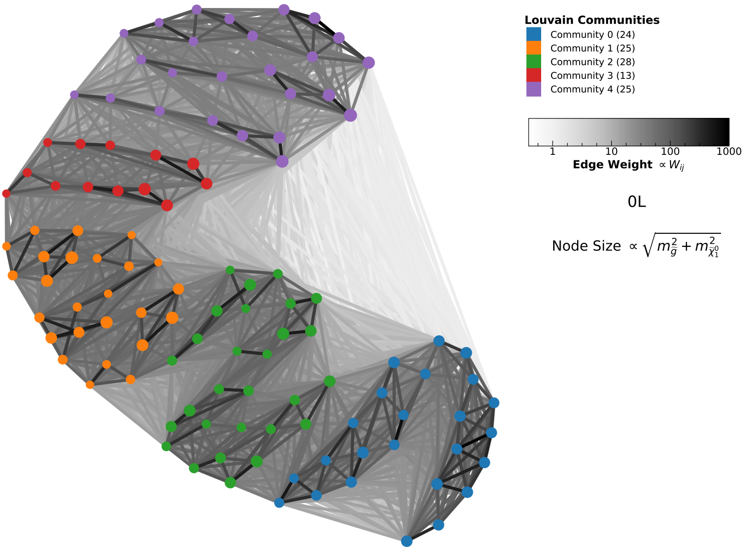


Louvain Communities
Community 0 (37)
Community 1 (12)
Community 2 (20)
Community 3 (19)
Community 4 (27)
1
10
100
1000
Edge Weight $\propto W_{ij}$
1L
Node Size $\propto \sqrt{m_{\tilde{g}}^2 + m_{\tilde{\chi}_1^0}^2}$

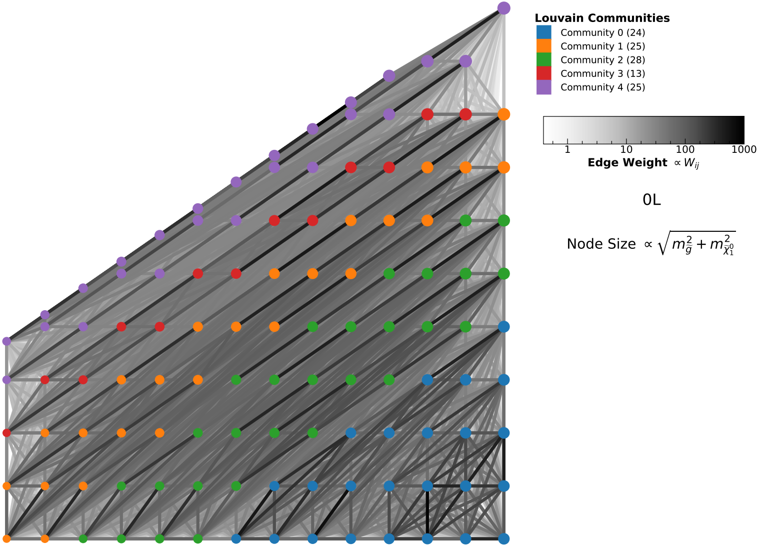
Graph (Mass Grid) 1L



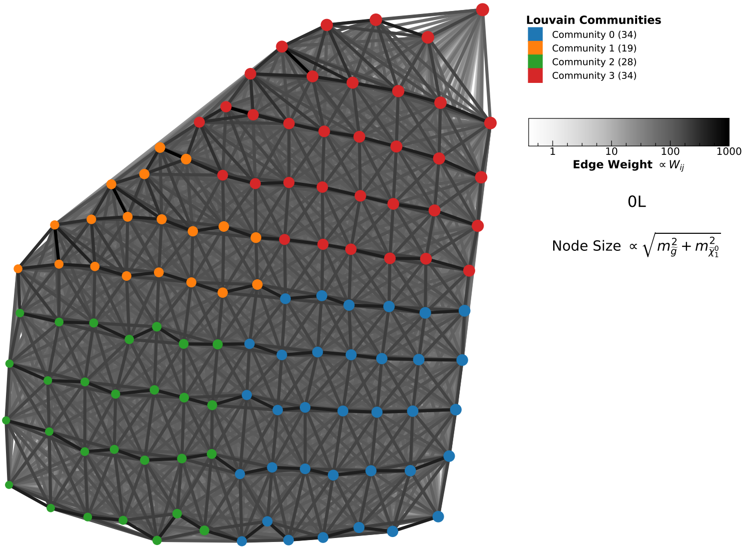
RMSD Alone: Graph 0L



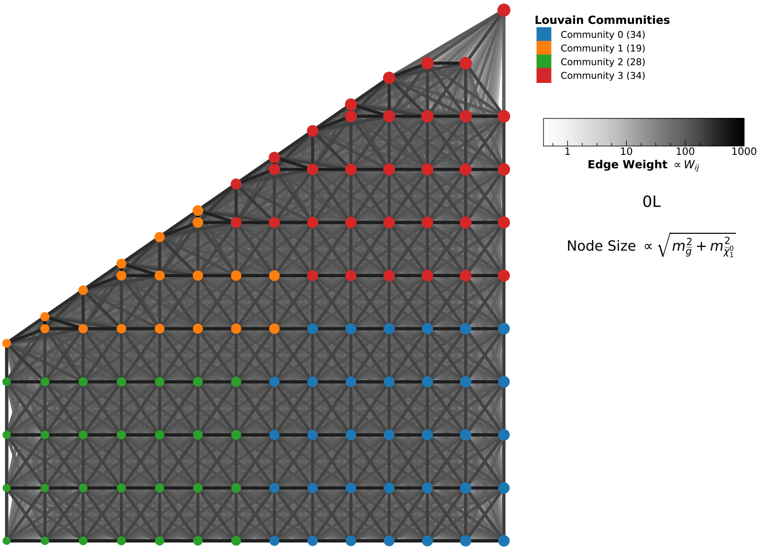
RMSD Alone: Graph 0L, Mass Grid



Mass Radius Alone: Graph 0L

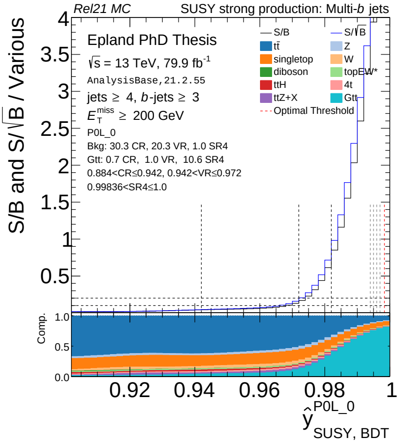


Mass Radius Alone: Graph 0L, Mass Grid

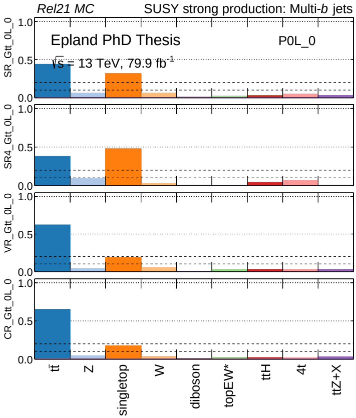


S/B & Background Compositions

S/B & Bkgs: P0L_0

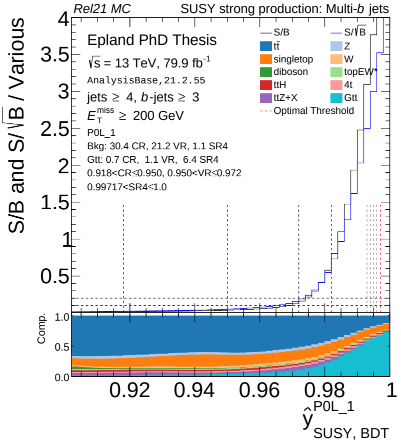


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
P0L_0
Bkg: 30.3 CR, 20.3 VR, 1.0 SR4
Gtt: 0.7 CR, 1.0 VR, 10.6 SR4
0.884<CR≤0.942, 0.942<VR≤0.972
0.99836<SR4≤1.0
S/B
S/√B
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Optimal Threshold
S/B and S/√B / Various
Comp.
ŷ P0L_0 SUSY, BDT

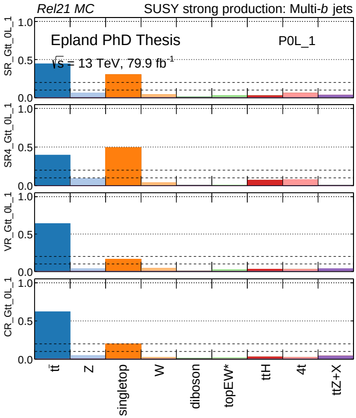


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
P0L_0
√s = 13 TeV, 79.9 fb-1
SR_Gtt_0L_0
SR4_Gtt_0L_0
VR_Gtt_0L_0
CR_Gtt_0L_0
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X

S/B & Bkgs: P0L_1

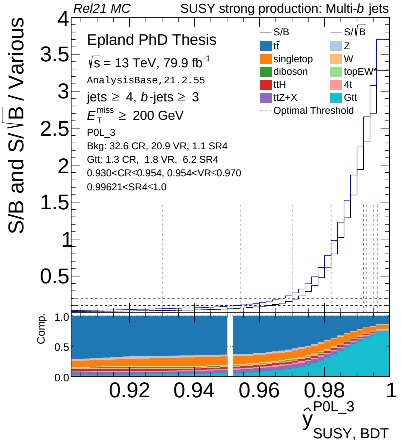


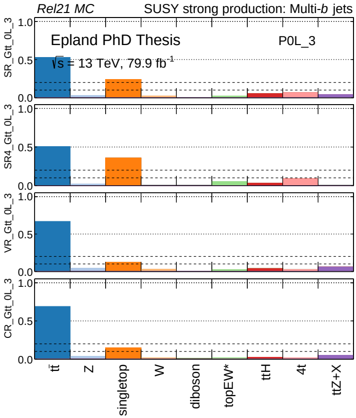
Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
P0L_1
Bkg: 30.4 CR, 21.2 VR, 1.1 SR4
Gtt: 0.7 CR, 1.1 VR, 6.4 SR4
0.918<CR≤0.950, 0.950<VR≤0.972
0.99717<SR4≤1.0
S/B
S/√B
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Optimal Threshold
S/B and S/√B / Various
Comp.
ŷ^P0L_1 SUSY, BDT



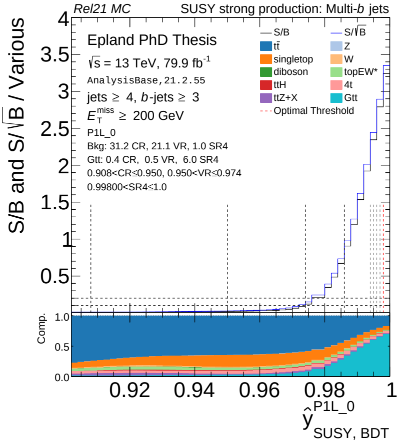
Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
P0L_1
√s = 13 TeV, 79.9 fb-1
SR_Gtt_0L_1
SR4_Gtt_0L_1
VR_Gtt_0L_1
CR_Gtt_0L_1
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X

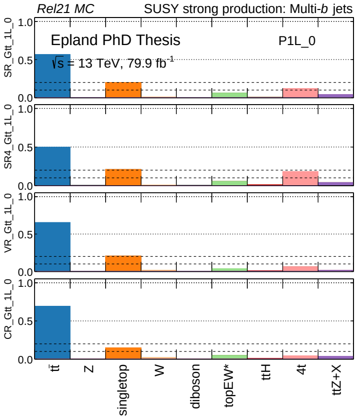
S/B & Bkgs: P0L_3


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
P0L_3
Bkg: 32.6 CR, 20.9 VR, 1.1 SR4
Gtt: 1.3 CR, 1.8 VR, 6.2 SR4
0.930<CR≤0.954, 0.954<VR≤0.970
0.99621<SR4≤1.0
S/B
S/√B
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Optimal Threshold
S/B and S/√B / Various
Comp.
ŷ SUSY, BDT P0L_3

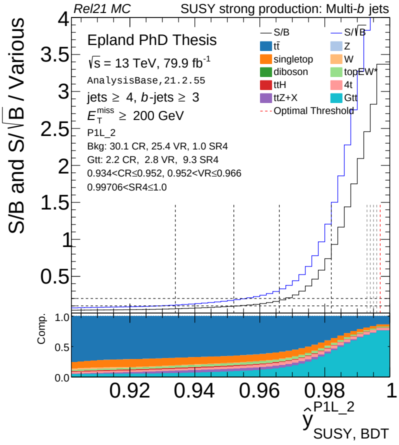

Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
P0L_3
√s = 13 TeV, 79.9 fb-1
SR_Gtt_0L_3
SR4_Gtt_0L_3
VR_Gtt_0L_3
CR_Gtt_0L_3
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X

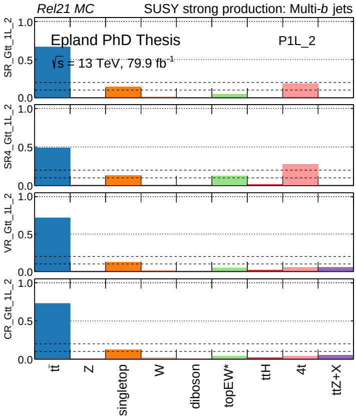
S/B & Bkgs: P1L_0


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
P1L_0
Bkg: 31.2 CR, 21.1 VR, 1.0 SR4
Gtt: 0.4 CR, 0.5 VR, 6.0 SR4
0.908<CR≤0.950, 0.950<VR≤0.974
0.99800<SR4≤1.0
S/B
S/√B
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Optimal Threshold
S/B and S/√B / Various
Comp.
ŷ P1L_0 SUSY, BDT

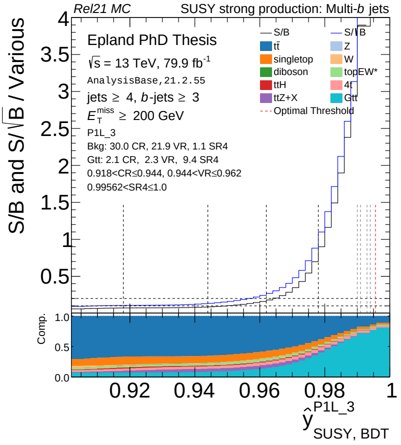

Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
P1L_0
√s = 13 TeV, 79.9 fb-1
SR_Gtt_1L_0
SR4_Gtt_1L_0
VR_Gtt_1L_0
CR_Gtt_1L_0
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X

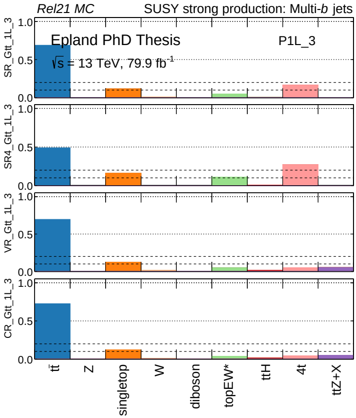
S/B & Bkgs: P1L_2


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
P1L_2
Bkg: 30.1 CR, 25.4 VR, 1.0 SR4
Gtt: 2.2 CR, 2.8 VR, 9.3 SR4
0.934<CR≤0.952, 0.952<VR≤0.966
0.99706<SR4≤1.0
S/B
S/√B
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Optimal Threshold
S/B and S/√B / Various
Comp.
ŷ^P1L_2 SUSY, BDT


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
P1L_2
√s = 13 TeV, 79.9 fb-1
SR_Gtt_1L_2
SR4_Gtt_1L_2
VR_Gtt_1L_2
CR_Gtt_1L_2
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X

S/B & Bkgs: P1L_3


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
P1L_3
Bkg: 30.0 CR, 21.9 VR, 1.1 SR4
Gtt: 2.1 CR, 2.3 VR, 9.4 SR4
0.918<CR≤0.944, 0.944<VR≤0.962
0.99562<SR4≤1.0
S/B
S/√B
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Optimal Threshold
S/B and S/√B / Various
Comp.
ŷ^P1L_3 SUSY, BDT


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
P1L_3
√s = 13 TeV, 79.9 fb-1
SR_Gtt_1L_3
SR4_Gtt_1L_3
VR_Gtt_1L_3
CR_Gtt_1L_3
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X

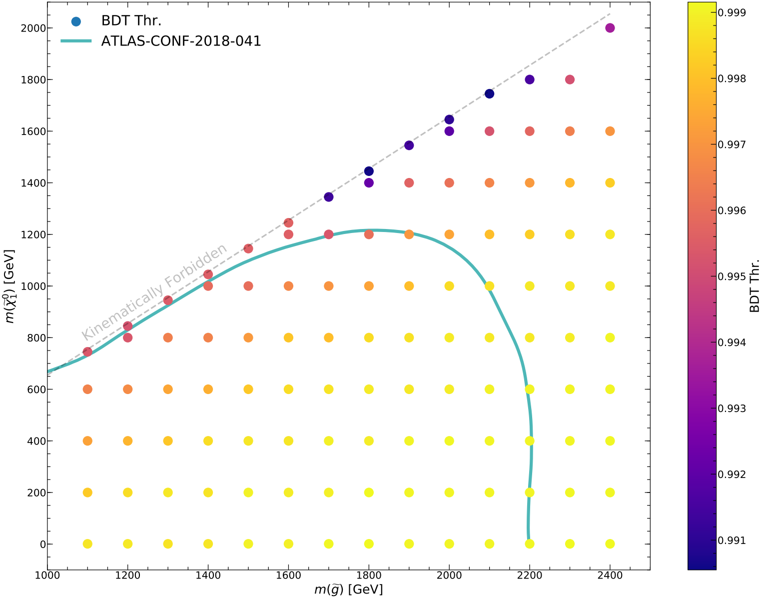
Development Studies

Estimating Significance Z

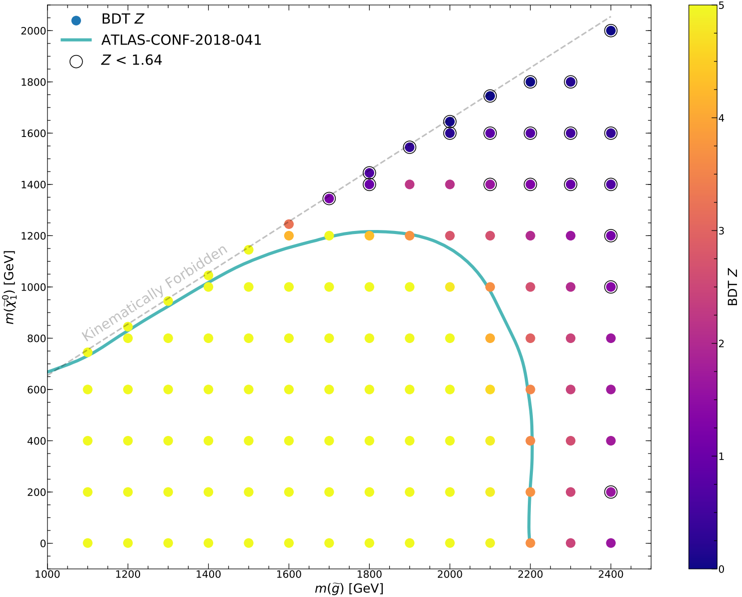
- Ultimately will be fitting output $\hat{y}$ in HistFitter, quite involved...
- To begin, use `BinomialExpZ`[3] to make an estimate of Z, Z_{B}
  - Works off of expected sig and bkg yields
  - Apply conservative bkg uncertainty of 50 %
- Optimize sig decision threshold on $\hat{y}_{\text{SUSY, BDT}}$ with respect to Z_{B}
  - Subject to keeping the bkg yield > 0.5, and $t\bar{t}$ statistical uncertainty $\sqrt{\sum w_i^2}/w < 0.3$
- The resulting optimal $\hat{y}_{\text{SUSY, BDT}}$ thresholds are always > 0.99

[3] Converted to python

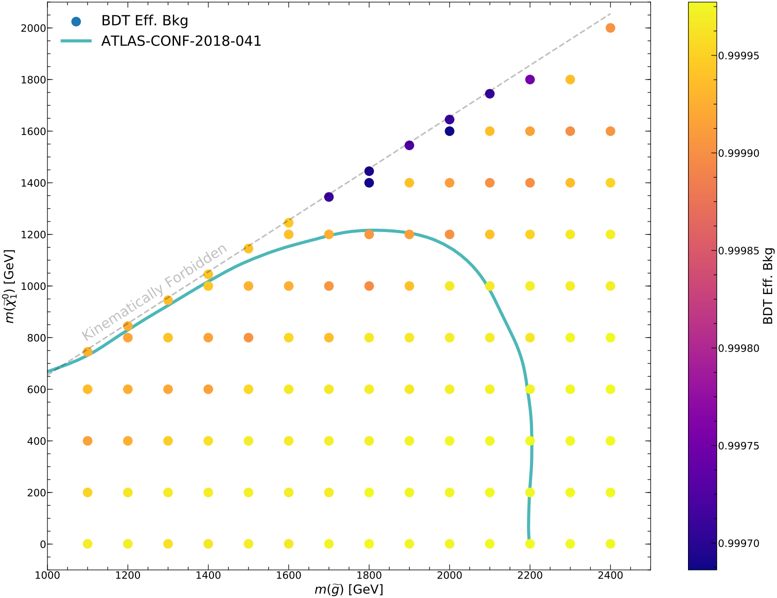
Optimal $\hat{y}_{\mathbf{SUSY,BDT}}$ Thresholds vs Mass Point



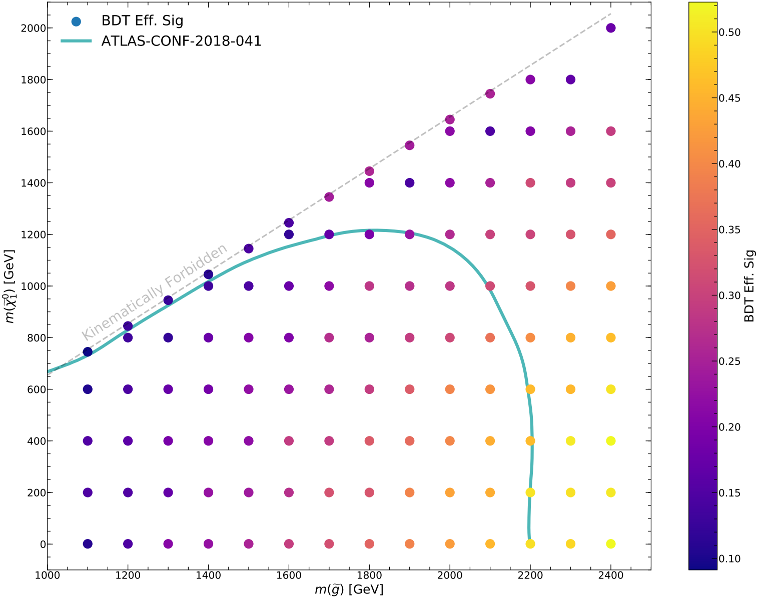
$Z_{\mathbf{B}}$ vs Mass Point



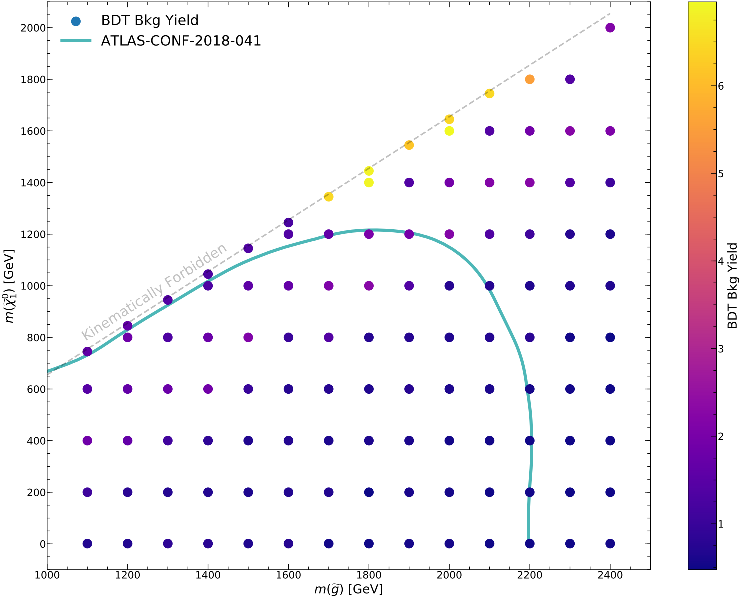
BDT Bkg Efficiency vs Mass Point



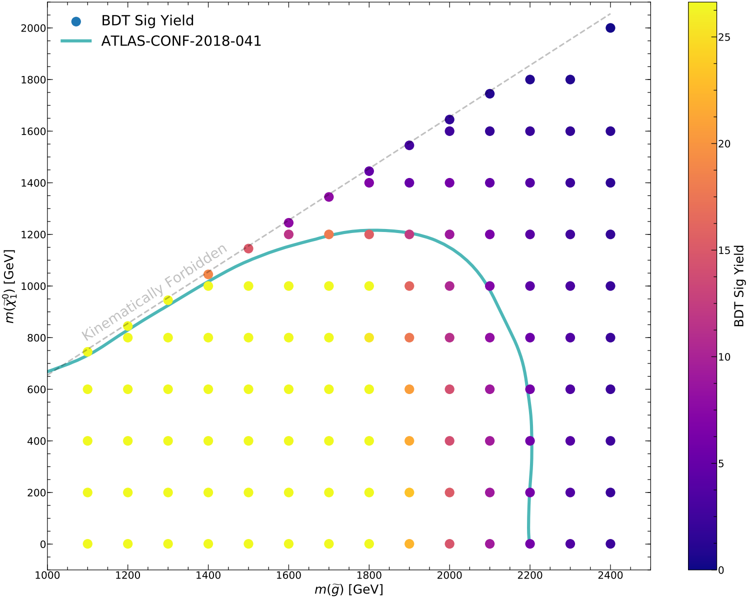
BDT Sig Efficiency vs Mass Point



Bkg Yield vs Mass Point

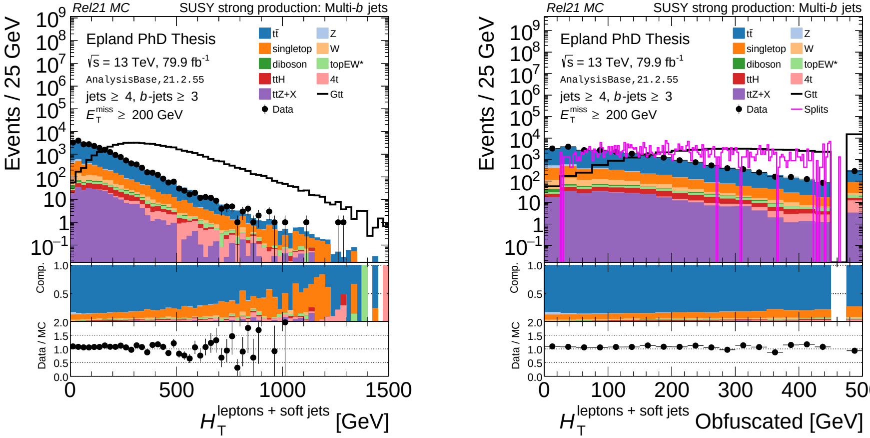


Sig Yield vs Mass Point



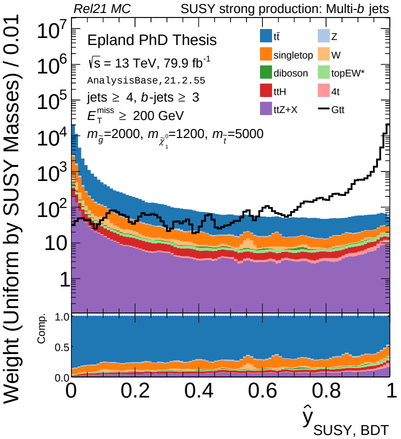
$H_T^{\text{leptons + soft jets}}$ Obfuscated

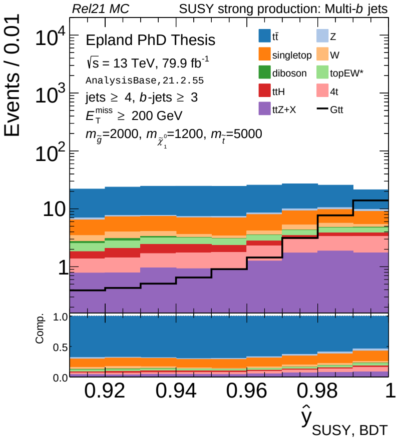
- Have to address the data / MC kink at ≈ 500 GeV
- Make $H_T^{\text{leptons + soft jets}}$ Obfuscated by setting > 450 GeV to 480 GeV



Compressed: BDT Output $\hat{y}_{\text{SUSY, BDT}}$

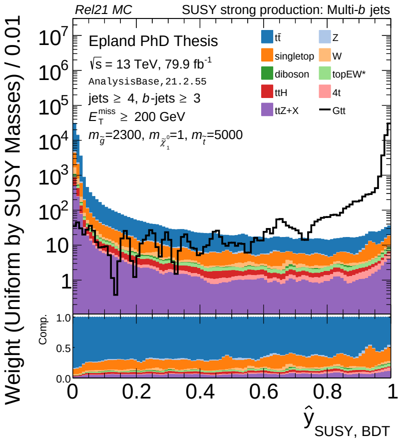
- Small $\Delta m\left(\tilde{g}, \tilde{\chi}_1^0\right) \rightarrow$ soft $\tilde{g}$ decay products, low $E_{\mathrm{T}}^{\text{miss}}$
  - $m_{\tilde{g}} = 2\,\text{TeV}$, $m_{\tilde{\chi}_1^0} = 1.2\,\text{TeV}$, $\Delta m = 800\,\text{GeV}$
- With physical weights have to go to $\hat{y}_{\text{SUSY, BDT}} > 0.99$ to find sig

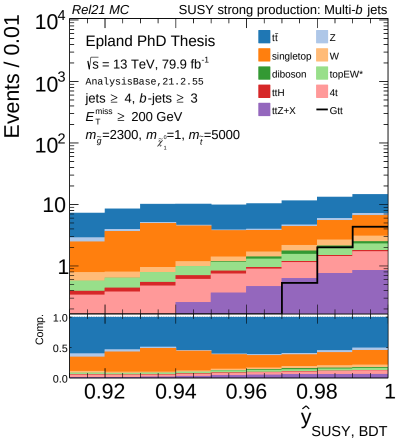




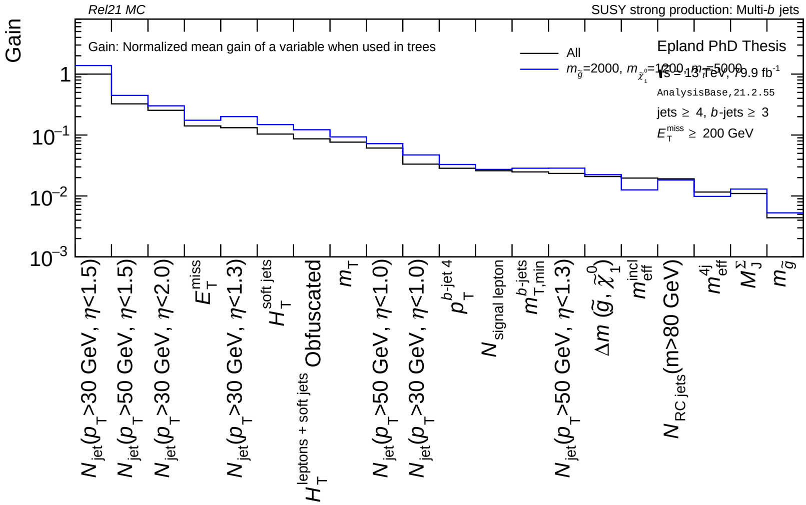
Boosted: BDT Output $\hat{y}_{\mathsf{SUSY,\,BDT}}$

- Large $\Delta m\left(\tilde{g}, \tilde{\chi}_1^0\right) \rightarrow$ highly boosted final state, high $E_{\mathrm{T}}^{\mathrm{miss}}$
  - $m_{\tilde{g}} = 2.3$ TeV, $m_{\tilde{\chi}_1^0} = 1$ GeV, $\Delta m \sim 2.3$ TeV
- Similar story at high $\hat{y}_{\mathsf{SUSY,\,BDT}}$

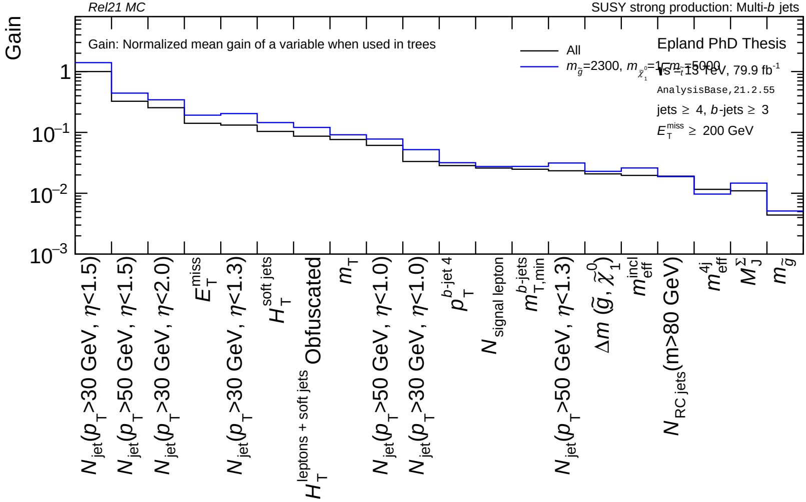




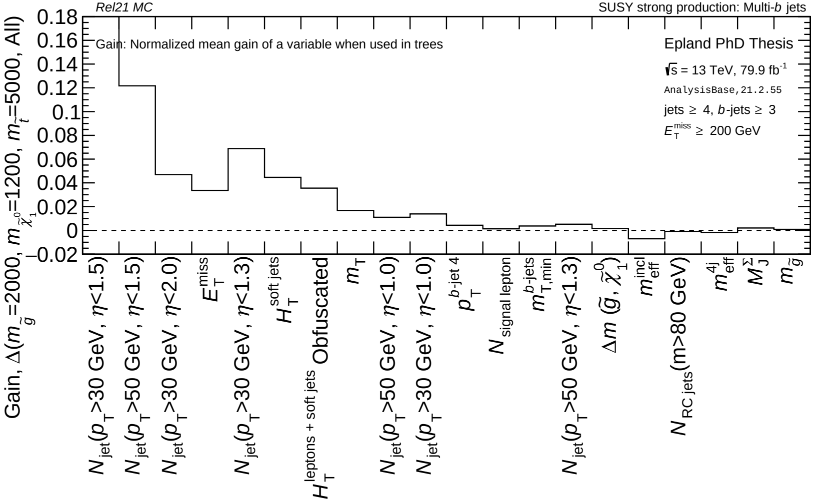
Compressed: Gain



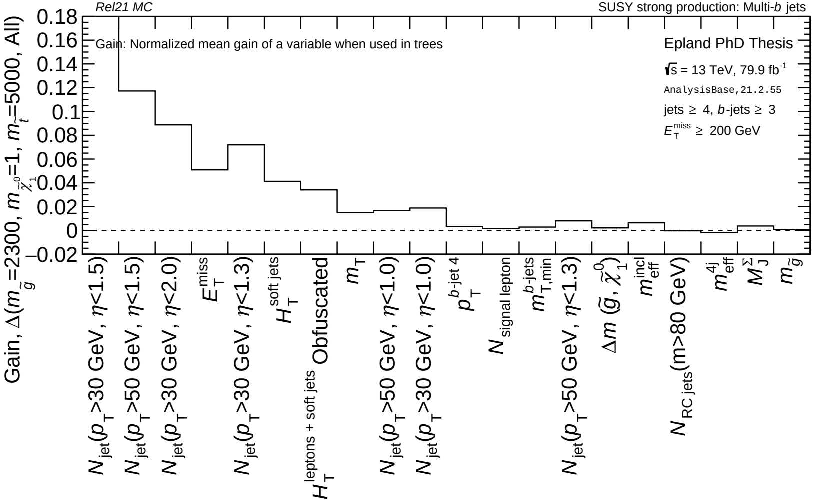
Boosted: Gain



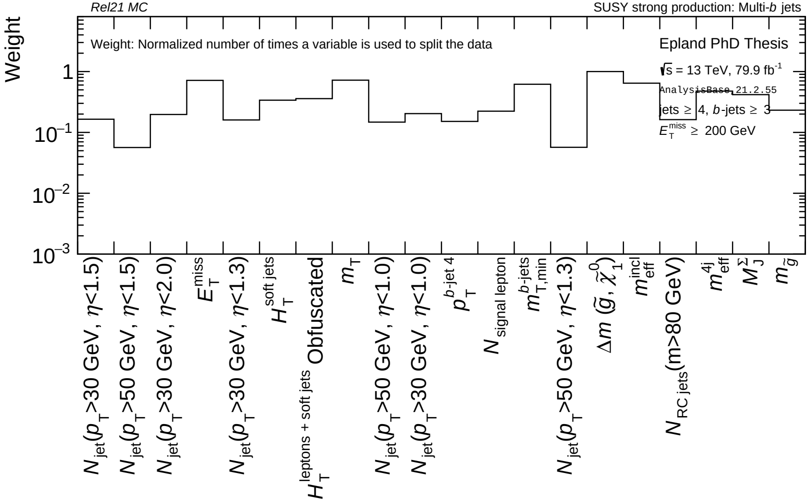
Compressed: Δ Gain



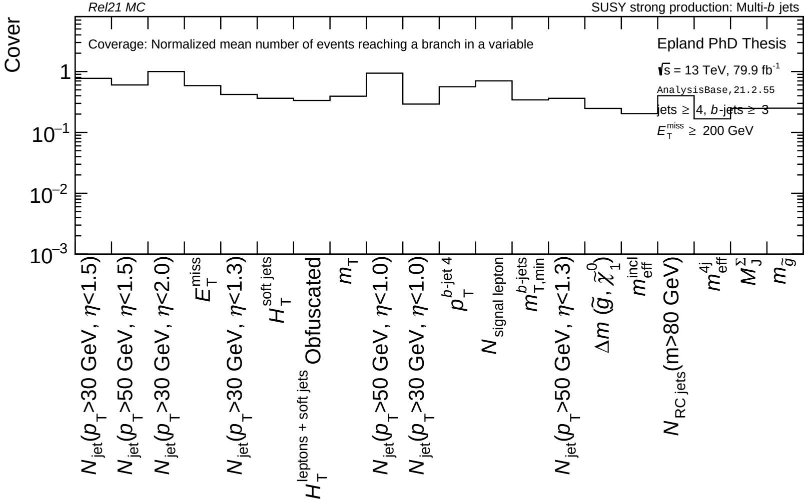
Boosted: Δ Gain



Variable Importance: Weight

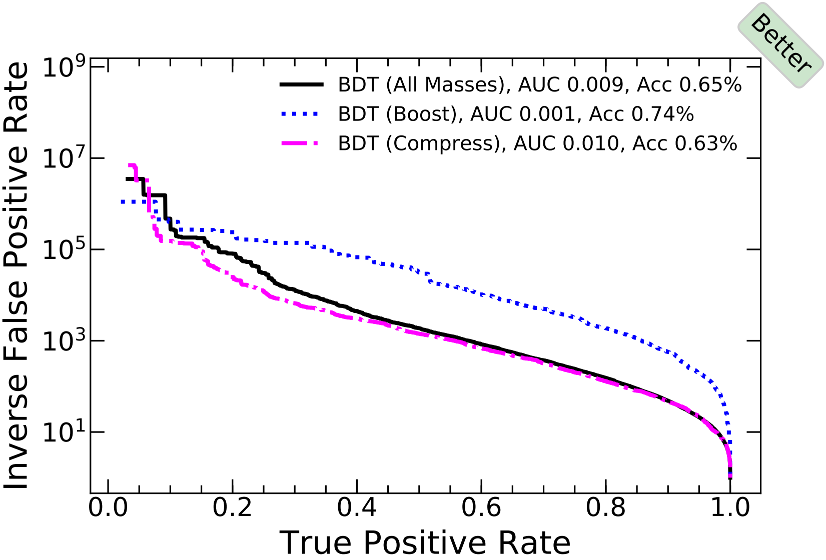


Variable Importance: Coverage



ROC Curves

- Better performance for boosted point/regime, as expected



Significance Optimization Pseudocode

```
# Numpy/Scipy port of the RooStats function 'BinomialExpZ' by Louis-Guillaume Gagnon
def significance(signalExp, backgroundExp, relativeBkgUncert):
  tau = 1.0 / (backgroundExp * relativeBkgUncert*relativeBkgUncert)
  x = 1.0 / (1.0 + tau)
  y = signalExp + backgroundExp
  z = 1 + backgroundExp * tau
  P_B = scipy.special.betainc(y, z, x)
  return -scipy.special.ndtri(P_B)
```

```
def find_best_thr(y, y_pred, W, relativeBkgUncert=0.5, bkg_cut_threshold = 0.5):
  fpr, tpr, thr = roc_curve(y, y_pred, sample_weight=W)
  # separate sig / bkg with numpy masks
  sigs = tpr*np.sum(W_sig)
  bkgs = fpr*np.sum(W_bkg)
  Zs = significance(sigs, bkgs, relativeBkgUncert)
  max_Z = -float('inf'); yield_sig_best = -float('inf'); yield_bkg_best = -float('inf')
  i_best = None
  for i in range(Zs.shape[0]):
    if Zs[i] > max_Z:
      W_bkg_selected = W_bkg[np.where(y_pred_bkg >= thr[i])]
      if W_bkg_selected_sum <= bkg_cut_threshold:
        limiting_constraint = 1; continue;

      W_ttbar=W_bkg[np.where((y_pred_bkg>thr[i])&(B_bkg==bkg_type_ttbar))]
      ttbar_stat_uncert=np.sqrt(np.sum(np.square(W_ttbar)))/W_ttbar_sum
      if not (W_ttbar_sum > 0 and ttbar_stat_uncert < ttbar_stat_cut_threshold):
        limiting_constraint = 2; continue

      max_Z = Zs[i]; yield_sig_best = sigs[i]; yield_bkg_best = bkgs[i]; i_best = i

  return thr[i_best], max_Z, yield_sig_best, yield_bkg_best
```

(Partial) Example Tree

```
booster[0]:
0:[jets_n_pt30_eta15 <5.5] yes=1,no=2,missing=1,gain=37636.6953,cover=20332.127
   1:[met<431.58255] yes=3,no=4,missing=3,gain=7176.16992,cover=9852.65234
      3:[mT<140.011749] yes=7,no=8,missing=7,gain=2296.76367,cover=8852.43164
         7:[mTb_min<186.546021] yes=15,no=16,missing=15,gain=166.470703,cover=8030.64453
            15:[signal_leptons_n <1.5] yes=31,no=32,missing=31,gain=99.6738281,cover
     =7373.80518
               31:[jets_n_pt30_eta20 <5.5] yes=63,no=64,missing=63,gain=56.8789062,cover
     =7229.31689
                  63:[meff_4j <2219.40161] yes=123,no=124,missing=123,gain=16.296875,cover
     =5966.33008
                     123:leaf=-0.141397834,cover=5958.4873
                     124:leaf=-0.0323069319,cover=7.84283161
                  64:[m_diff <650] yes=125,no=126,missing=125,gain=31.7756348,cover
     =1262.98694
                     125:leaf=-0.105186649,cover=370.599792
                     126:leaf=-0.131471351,cover=892.387207
               32:[jets_n_pt30_eta13 <4.5] yes=65,no=66,missing=65,gain=48.8466034,cover
     =144.488342
                  65:[signal_leptons_n <2.5] yes=127,no=128,missing=127,gain=30.9985657,
     cover=112.536476
                     127:leaf=-0.105914511,cover=109.036499
                     128:leaf=0.0933374166,cover=3.49997544
                  66:[m_diff <450] yes=129,no=130,missing=129,gain=11.7550745,cover
     =31.9518681
                     129:leaf=0.0635010377,cover=10.0030622
                     130:leaf=-0.0273681637,cover=21.9488068
            16:[jets_n_pt30_eta15 <4.5] yes=33,no=34,missing=33,gain=171.120239,cover
     =656.838989
```

Multi-b

Multi-b Variables (1/2)

- Transverse mass between $E_{\mathrm{T}}^{\mathrm{miss}}$ and leading lepton
  - Has kinematic endpoint near m_W for leptonic W decays in $t\bar{t}$ & W+jets

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\mathrm{lepton}} E_{\mathrm{T}}^{\mathrm{miss}} \left(1 - \cos\left(\Delta\phi\left(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,\mathrm{lepton}}\right)\right)\right)}$$

- Min transverse mass between $E_{\mathrm{T}}^{\mathrm{miss}}$ and three leading b-jets
  - Has kinematic endpoint near m_t for $t\bar{t}$ background
  - Larger for SUSY as $\widetilde{\chi}_1^0$ $E_{\mathrm{T}}^{\mathrm{miss}}$ is largely independent of b-jets

$$m_{\mathrm{T,min}}^{b\text{-jets}} = \min_{i\leq 3}\left(\sqrt{2p_{\mathrm{T}}^{b\text{-jet}\,i} E_{\mathrm{T}}^{\mathrm{miss}} \left(1 - \cos\left(\Delta\phi\left(\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}},\, \vec{p}_{\mathrm{T}}^{\,b\text{-jet}\,i}\right)\right)\right)}\right)$$

- Sum p_{T} from soft components of the event (new!)
  - *Capped at 450 GeV, "Obfuscated"

$$H_{\mathrm{T}}^{\mathrm{soft\ jets}} = \sum_{5\leq i} p_{\mathrm{T}}^{\mathrm{jet}\,i}$$

$$H_{\mathrm{T}}^{\mathrm{leptons\,+\,soft\ jets}*} = H_{\mathrm{T}}^{\mathrm{soft\ jets}} + H_{\mathrm{T}}^{\mathrm{leptons}}$$

Multi-b Variables (2/2)

- Min $\Delta\phi$ between $E_{\mathrm{T}}^{\mathrm{miss}}$ and any of the four leading jets
  - Helps reduce multi-jet background in 0L channel

$$\Delta\phi_{\mathrm{min}}^{4j} = \min_{i \leq 4}\left(\left|\phi_{\mathrm{jet}\,i} - \phi_{\vec{p}_{\mathrm{T}}^{\mathrm{miss}}}\right|\right)$$

- Mass of leading four jets

$$M_J^{\Sigma} = \sum_{i \leq 4} m_{\mathrm{RC\,jet}\,i}$$

- Effective mass of $E_{\mathrm{T}}^{\mathrm{miss}}$ plus all signal leptons & jets[4]

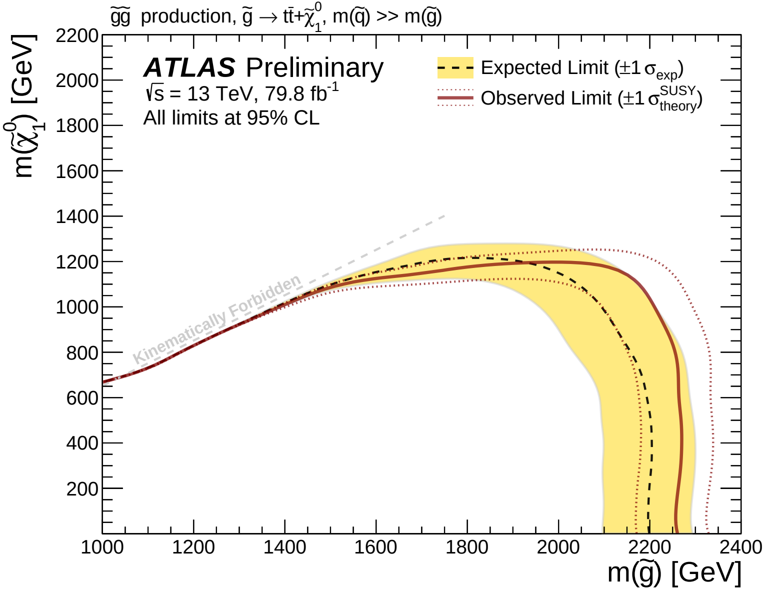
$$m_{\mathrm{eff}}^{\mathrm{incl}} = E_{\mathrm{T}}^{\mathrm{miss}} + \sum_i p_{\mathrm{T}}^{\mathrm{jet}\,i} + \sum_j p_{\mathrm{T}}^{\mathrm{lep}\,j}$$

- And for just the first 4 jets

$$m_{\mathrm{eff}}^{4j} = E_{\mathrm{T}}^{\mathrm{miss}} + \sum_{i=1}^{4} p_{\mathrm{T}}^{\mathrm{jet}\,i}$$

[4] With $p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $|\eta| < 2.8$

ATLAS-CONF-2018-041 Exclusion Limits



[5]

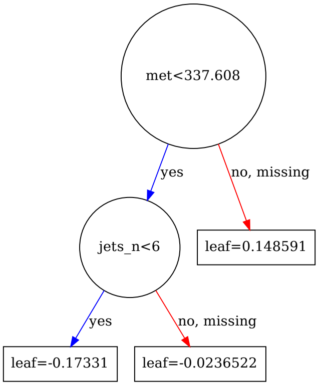
Machine Learning

Machine Learning Basics

- Supervised learning: Train model on many known examples $\vec{x}_i, y_i$
- Model consists of θ_j parameters, *e.g.* linear $\hat{y}_i = \sum_j \theta_j x_{ij}$
- Minimize two part objective function: $\text{obj}\left(\theta\right) = L\left(\theta\right) + \Omega\left(\theta\right)$
- Training Loss: $L\left(\theta\right)$ measures model performance on training set
  - MSE: $L\left(\theta\right) = \sum_i \left(y_i - \hat{y}_i\right)^2$
  - Logistic: $L\left(\theta\right) = \sum_i \left[y_i \ln\left(1 + \exp(-\hat{y}_i)\right) + \left(1 - y_i\right) \ln\left(1 + \exp(\hat{y}_i)\right)\right]$
- Regularization: $\Omega\left(\theta\right)$ measures model complexity, prevents over fitting
  - L1 regularization: $\Omega\left(\theta\right) \sim \lambda\|\theta\|$
  - L2 regularization: $\Omega\left(\theta\right) \sim \lambda\|\theta\|^2$

Classification and Regression Trees (CARTs)

- Tree with branches of cuts chosen when training the model
  - Are just regular cuts, so we can (try to) understand what is happening, and we don't need parameter scaling
- Model's prediction assigns each event to a leaf, gets weight w_j
  - Background-like $w_j < 0$, signal-like $0 < w_j$



Gradient Boosting

- However individual CARTs are poor & limited models $\rightarrow$
- Use an ensemble (**boosting**) of K trees, sum the individual weights
  - Take logistic function of output to get probability $0 < \hat{y}_{\text{BDT}} < 1$
- Iteratively add each new tree $f_k(x_i)$, complementing the existing trees

$$\begin{aligned}
\hat{y}_i^{(0)} &= 0 \\
\hat{y}_i^{(1)} &= f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\
\hat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\
\hat{y}_i^{(t)} &= \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i)
\end{aligned}$$

- "Choose" each $f_k(x_i)$ by minimizing obj (θ)
  - In practice, grow f_k from 0 branches as there are ∞ possible trees
- Lots of math (2nd order Taylor expansion. . .) $\rightarrow$ **gradient descent**

XGBoost

- eXtreme Gradient Boosting: github.com/dmlc/xgboost
- Open source library for gradient boosted trees [6]
- High performance, used in many winning ML challenge solutions
    - Including by the devs in the Higgs challenge
- Very versatile (CPU, GPU, Hadoop, Spark, Python, R, Scala, C++...)
- Uses L1 + L2 regularization: T = number of leaves, w_j = leaf weights

$$\Omega\left(f\right) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$$

SUSY

General Properties of SUSY

- Generate SUSY transform with operator Q:

$$Q\,|\text{Boson}\rangle = |\text{Fermion}\rangle \qquad\qquad Q\,|\text{Fermion}\rangle = |\text{Boson}\rangle$$

- Q must be spin $1/2$ (*i.e.* spacetime symmetry) for $Q\,|\text{B}\rangle = |\text{F}\rangle$
- Spacetime symmetry + SM fermion chirality + Haag-Łopuszański-Sohnius extension of Coleman-Mandula theorem $\rightarrow$

$$\{Q, Q^\dagger\} = P^\mu$$
$$\{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0$$
$$[P^\mu, Q] = [P^\mu, Q^\dagger] = 0$$

- $[P^\mu, Q] = 0$ implies $-P^2$ (m^2 operator) commutes with Q, $Q^\dagger$ $\rightarrow$ In unbroken SUSY, particles and superpartners have identical m
- Q also commutes with gauge generators $\rightarrow$ Share electric charge, weak isospin, and color degrees of freedom

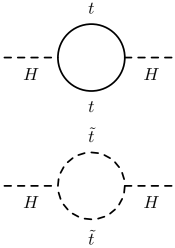
[10]

Solving the Hierarchy Problem

- In practice each SM fermion partners with two complex scalar fields
  - One for each of its left and right-handed 2-component spinor elements
  - Results in 4 fermion d.o.f. pairing with 4 boson d.o.f.

- When Higgs interactions are added the two partner complex scalar fields exactly cancel the fermion correction!
  - For $\lambda_S = |\lambda_f|^2 = \lambda$

$$\begin{aligned}\Delta m_H^2 = &- \frac{\lambda}{8\pi^2}\Lambda_{\mathsf{UV}}^2 \\ &+ 2 \times \frac{\lambda}{16\pi^2}\Lambda_{\mathsf{UV}}^2 \\ &+ \mathcal{O}\Big(\Big(m^2 \log \Lambda_{\mathsf{UV}}\Big)\Big)\end{aligned}$$



[10]

Minimal Supersymmetric SM (MSSM) (1/2)

- Add the minimum number of SUSY fields to the SM $\rightarrow$ MSSM
- Form supermultiplets of SM particles and their superpartners
  - Fermions $\leftrightarrow$ scalar fermions (sfermions: squarks, sleptons)
  - Gauge bosons $\leftrightarrow$ fermion gauginos (gluinos, winos, bino, photino)
- Results in the following fields / particles
  - Plus all fermion generations and anti-particles

Super-multiplets	Super-fields	Bosonic Fields	Fermionic Partners	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
gluon/gluino	$\widehat{V}_8$	g	$\widetilde{g}$	8	1	0
gauge boson/	$\widehat{V}$	$W^{\pm}, W^0$	$\widetilde{W}^{\pm}, \widetilde{W}^0$	1	3	0
gaugino	$\widehat{V}'$	B	$\widetilde{B}$	1	1	0
slepton/	$\widehat{L}$	$(\widetilde{\nu}_L, \widetilde{e}_L)$	$(\nu, e^-)_L$	1	2	-1
lepton	$\widehat{E}^c$	$\widetilde{e}^*_R$	e^c_L	1	1	2
squark/	$\widehat{Q}$	$(\widetilde{u}_L, \widetilde{d}_L)$	$(u, d)_L$	3	2	$\frac{1}{3}$
quark	$\widehat{U}^c$	$\widetilde{u}^*_R$	u^c_L	$\bar{3}$	1	$-\frac{4}{3}$
	$\widehat{D}^c$	$\widetilde{d}^*_R$	d^c_L	$\bar{3}$	1	$\frac{2}{3}$
Higgs/	$\widehat{H}_u$	(H^+_u, H^0_u)	$(\widetilde{H}^+_u, \widetilde{H}^0_u)$	1	2	1
Higgsino	$\widehat{H}_d$	(H^0_d, H^-_d)	$(\widetilde{H}^0_d, \widetilde{H}^-_d)$	1	2	-1

[11]

Minimal Supersymmetric SM (MSSM) (2/2)

- The Higgs supermultiplet is a bit more complicated...
  - Two complex Higgs doublets $\leftrightarrow$ fermion Higgsino doublet
  - Plus an anti-particle copy of the supermultiplet

- Need the Higgs doublet to cancel Higgsino generated gauge anomalies
- Also for up & down-type mass generation to be consistent with SUSY
  - Superpotential $W = \lambda_d \widehat{H}_d \widehat{Q} \widehat{D}^c - \lambda_u \widehat{H}_u \widehat{Q} \widehat{U}^c + \lambda_e \widehat{H}_d \widehat{L} \widehat{E}^c + \mu \widehat{H}_u \widehat{H}_d$

- Neutral (charged) gauginos and Higgsinos mix $\rightarrow$ physical neutralinos $\widetilde{\chi}^0$ (charginos $\widetilde{\chi}^{\pm}$) mass states
- Results in 5 Higgs particles: $H^{\pm}$, CP-even h^0 & H^0, and CP-odd A^0
  - 125 GeV Higgs is h^0 (by construction $m(h^0) < m(H^0)$)

End result is a model with **124 free parameters!**

[11]

R-Parity and the LSP

- In SM baryon & lepton numbers (B, L) are conserved $\rightarrow p$ is stable
  - $p \rightarrow e^+\pi^0$ has a mean lifetime of $> 8.2 \times 10^{33}$ years [11]
- In SUSY particle-superpartner operators can violate B–L conservation
- To fit experiment, force $\mathcal{L}_{\text{MSSM}}$ to obey R-parity, where S is spin
  - This restores (renormalizable) B–L conservation

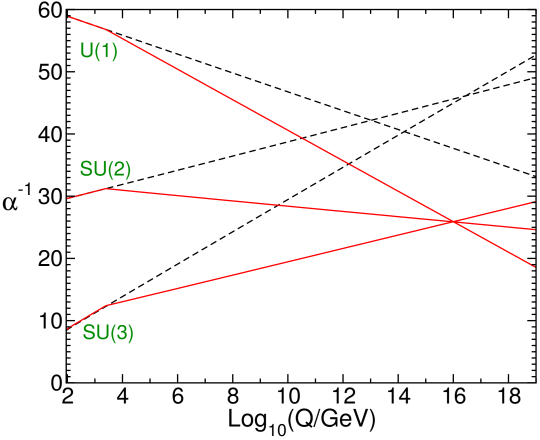
$$R = (-1)^{3(B-L)+2S}$$

- All SM particles have $R = +1$; SUSY $R = -1 \;\; \rightarrow$
  - SM colliders produce even numbers of SUSY particles
  - SUSY states can't fully decay to SM
  - The lightest supersymmetric particle (LSP) is absolutely stable
- If the LSP is uncharged, it only weakly interacts with SM particles $\rightarrow$
- LSP can be an excellent DM candidate, appears as $E_{\text{T}}^{\text{miss}}$ in colliders

[10]

Unification

- SM (dashed) $SU(2)_L$ & $U(1)_Y$ gauge couplings converge (electroweak unification) at a high energy scale, but $SU(3)_C$ does not
- MSSM (solid) adds the right amount of new particles to the loop corrections for all to converge around $M_U \approx 1.5 \times 10^{16}$ GeV!



[10]

Broken SUSY & Naturalness Considerations

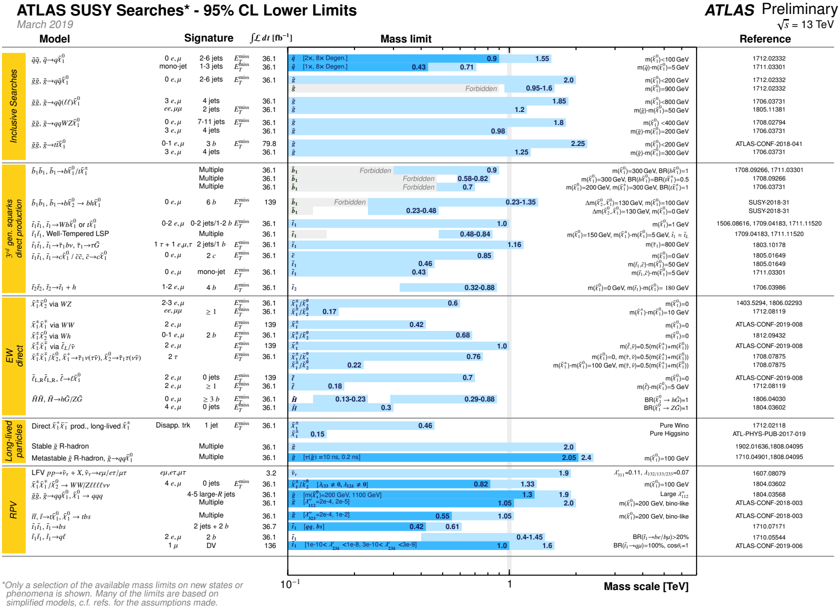
- Would have noticed a $511\,\text{keV}$ $\widetilde{e}$ long ago $\rightarrow$ SUSY must be broken
- To continue to cancel the Higgs mass corrections, would like to keep $m_S \approx m_f$ & $\lambda_S \approx |\lambda_f|^2$ $\rightarrow$ Soft SUSY breaking
- Sets some conditions but still arbitrary; $105/124$ parameters control it

- Soft SUSY breaking introduces a "little" hierarchy problem for m_Z
- To avoid fine-tuning, naturalness leads to upper limit on M_{SUSY}
- $M_{\text{SUSY}} \sim 1\,\text{TeV}$ requires $\sim 1\%$ fine-tuning to get the correct m_Z

- LHC SUSY searches are really starting to push against this limit, expect to see even more conflict after Run 2

[10, 11]

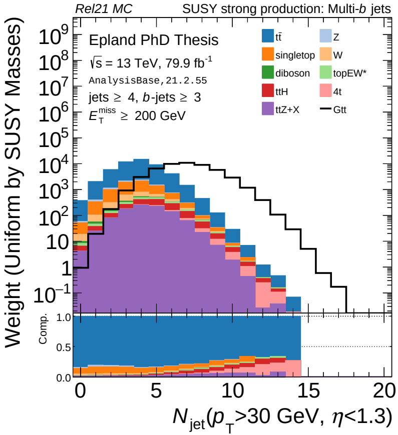
ATLAS SUSY Mass Limits

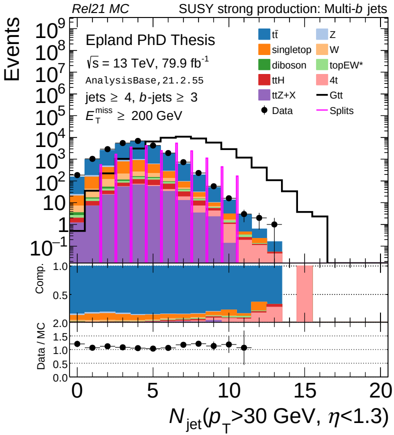


[12]

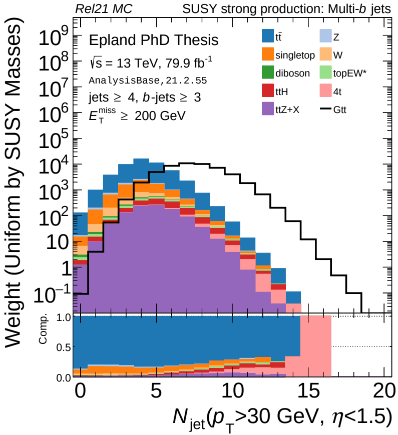
Training Variables (with BDT Splits)

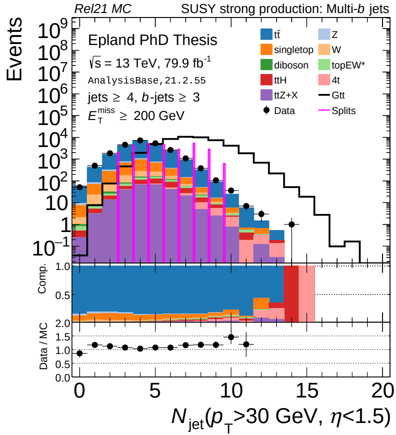
$N_{\text{jet}}(p_{\text{T}} > 30\text{ GeV}, \eta < 1.3)$



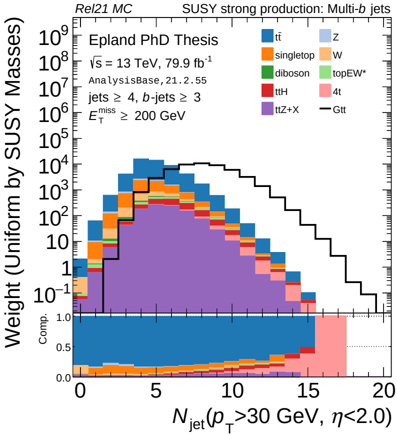


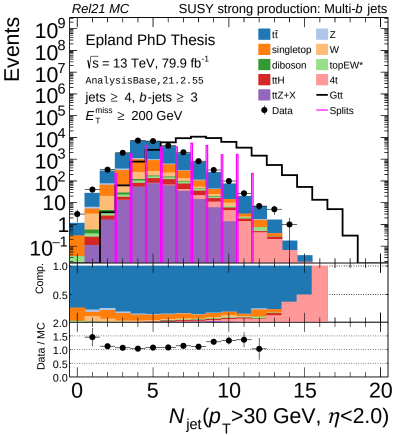
$N_{\mathbf{jet}}(p_{\mathbf{T}} > 30\ \text{GeV}, \eta < 1.5)$



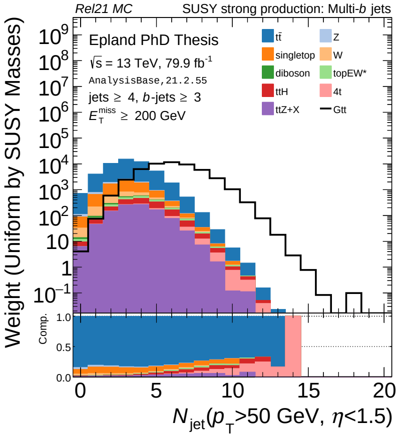


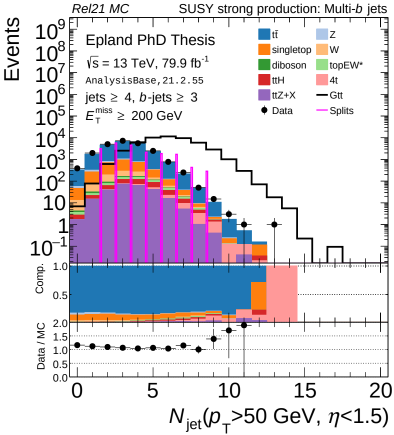
$N_{\mathbf{jet}}(p_{\mathbf{T}} > 30\ \mathrm{GeV}, \eta < 2.0)$



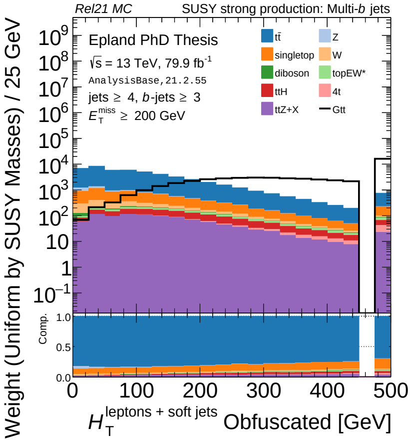


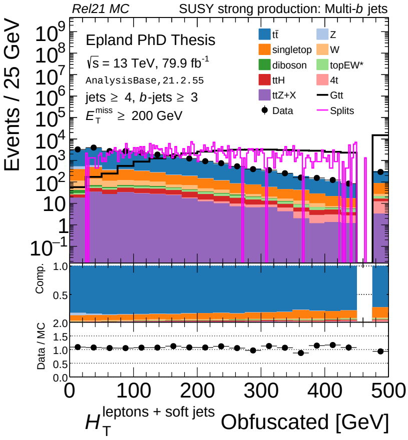
$N_{\mathbf{jet}}(p_{\mathbf{T}} > 50\ \text{GeV}, \eta < 1.5)$



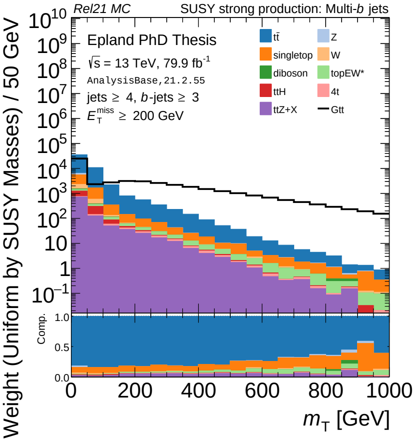


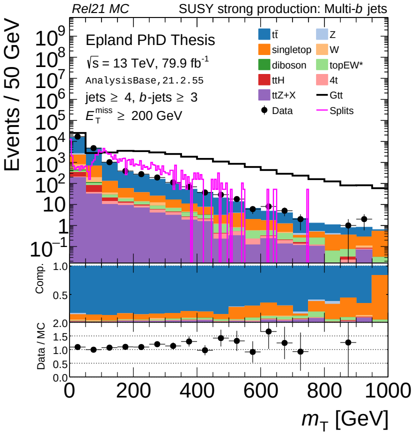
$H_T^{\text{leptons + soft jets}}$ Obfuscated



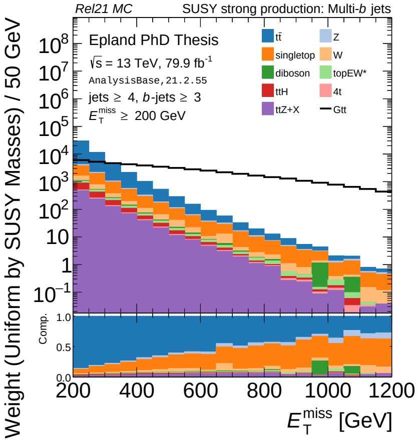


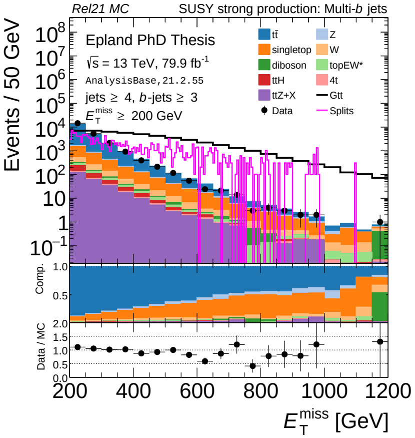
m_{T}


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Weight (Uniform by SUSY Masses) / 50 GeV
Comp.
m_T [GeV]


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Data
Splits
Events / 50 GeV
Comp.
Data / MC
m_T [GeV]

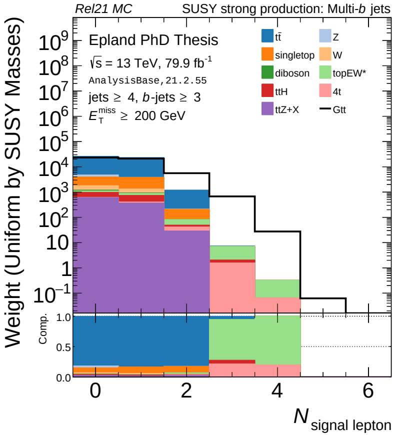
$E_{\mathrm{T}}^{\mathrm{miss}}$

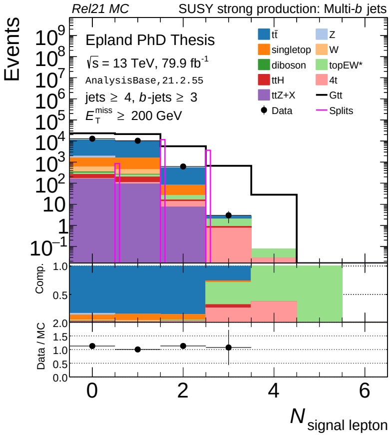




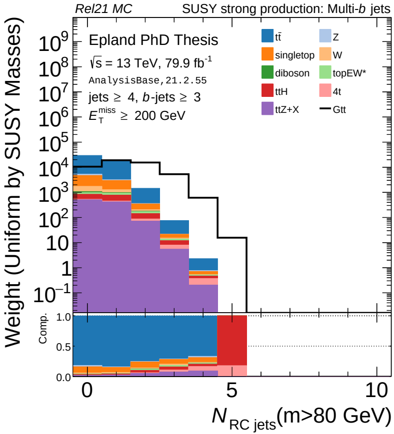
Second Tier Training Variables

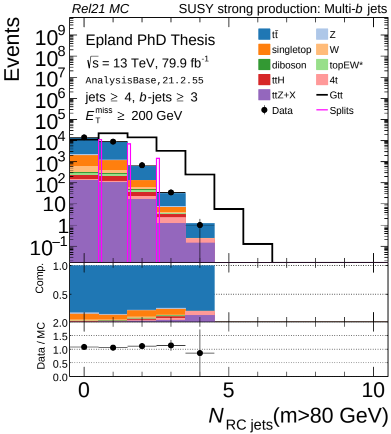
$N_{\textbf{sig lep}}$



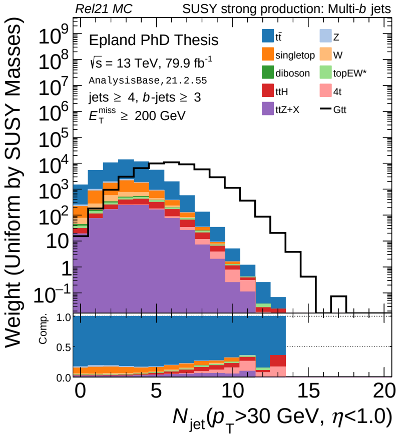


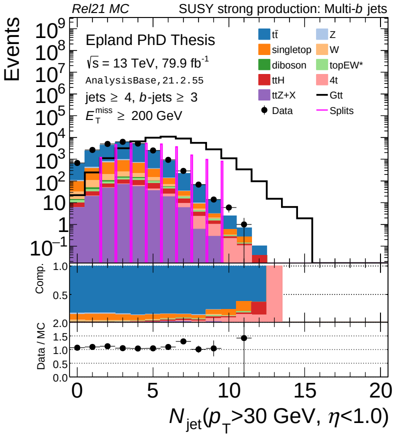
$N_{\mathbf{RC\,jet}}(m > 80\ \mathrm{GeV})$



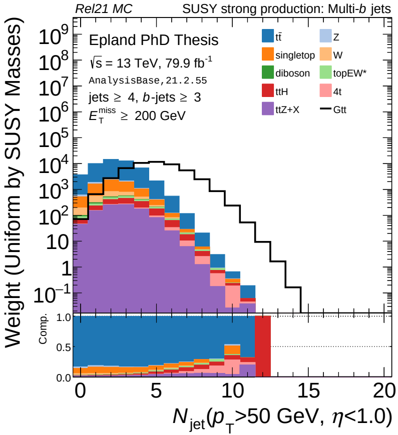


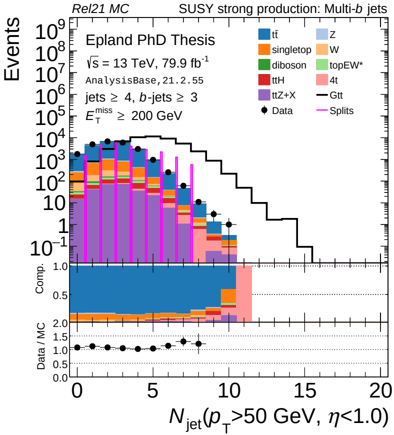
$N_{\mathbf{jet}}(p_{\mathbf{T}} > 30 \text{ GeV}, \eta < 1.0)$



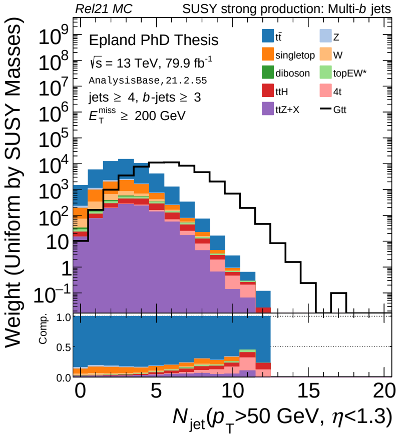


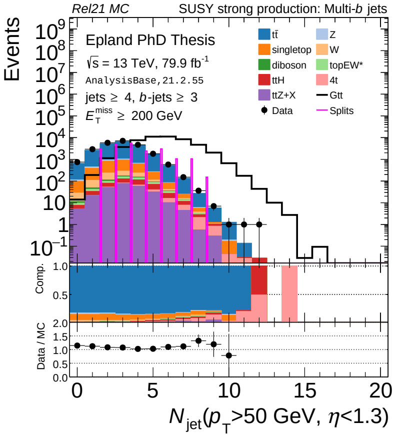
$N_{\mathbf{jet}}(p_{\mathbf{T}} > 50\ \mathrm{GeV}, \eta < 1.0)$



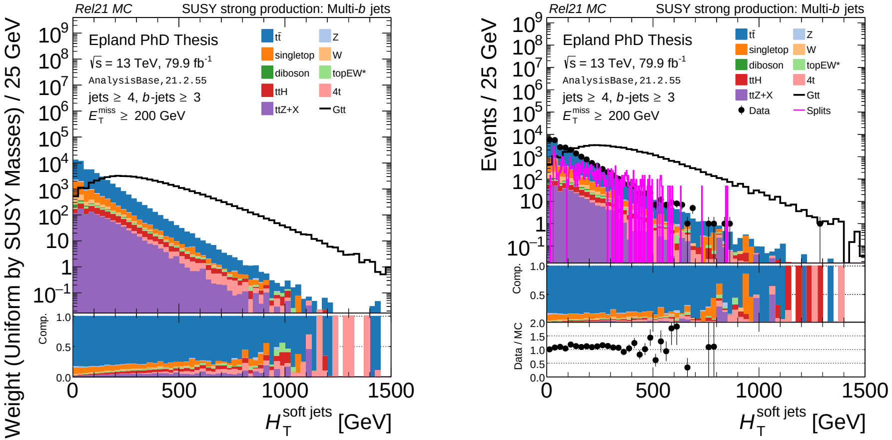


$N_{\mathbf{jet}}(p_{\mathbf{T}} > 50\ \mathrm{GeV}, \eta < 1.3)$

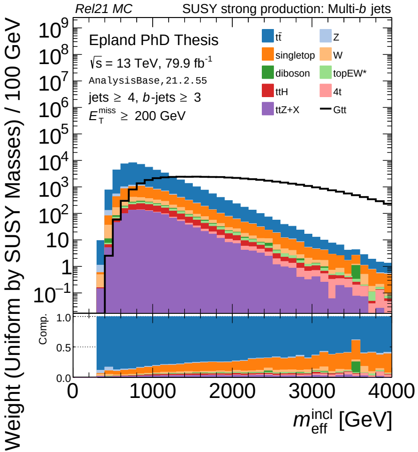


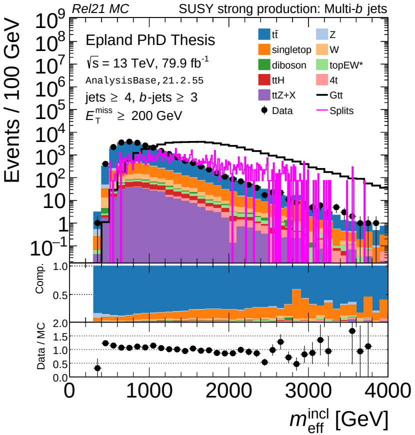


$H_{\mathrm{T}}^{\textbf{soft jets}}$

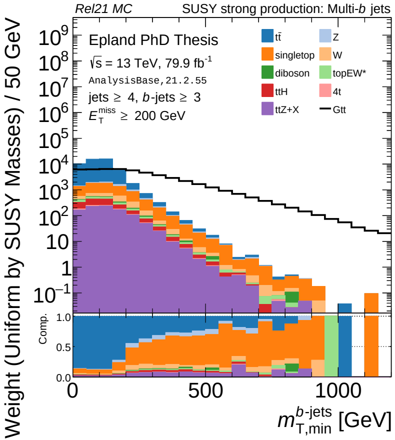


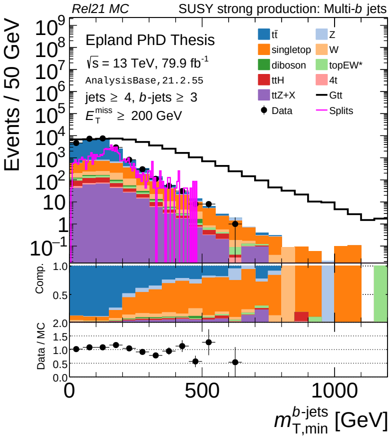
$m_{\text{eff}}^{\text{incl}}$



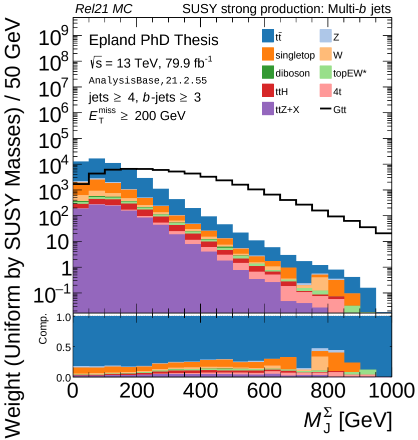


$m_{\mathrm{T,min}}^{b\text{-jets}}$

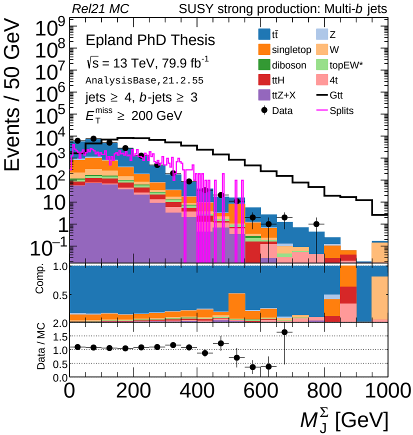




M_J^Σ

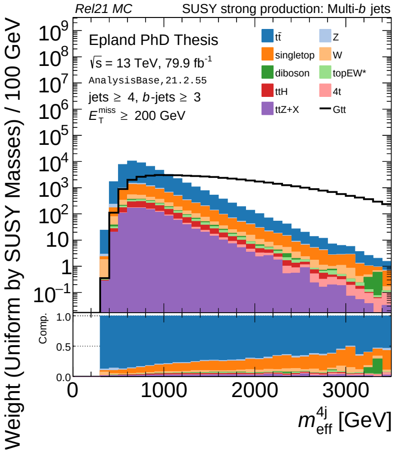


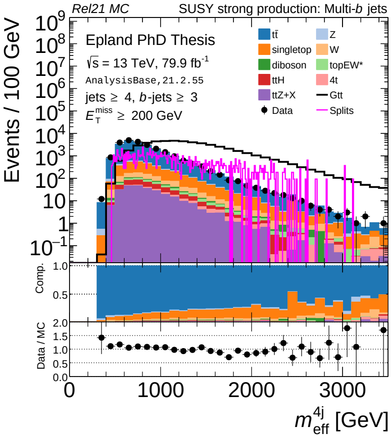
Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Weight (Uniform by SUSY Masses) / 50 GeV
Comp.
M_J^Σ [GeV]



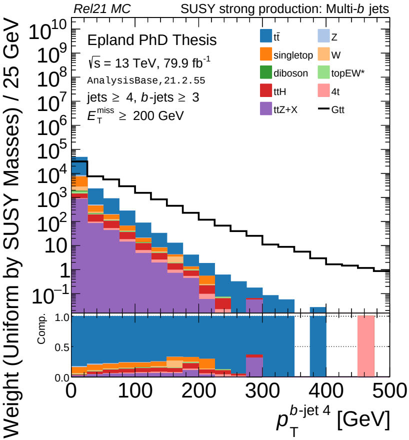
Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Data
Splits
Events / 50 GeV
Comp.
Data / MC
M_J^Σ [GeV]

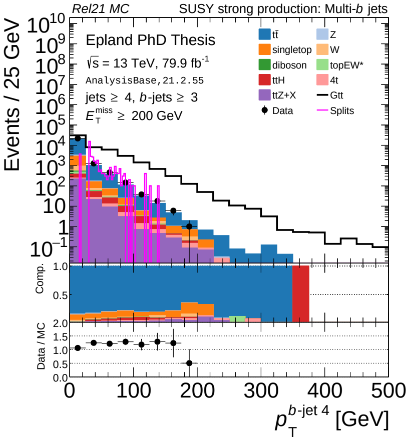
m_{eff}^{4j}


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Weight (Uniform by SUSY Masses) / 100 GeV
Comp.
m_eff^4j [GeV]


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Data
Splits
Events / 100 GeV
Comp.
Data / MC
m_eff^4j [GeV]

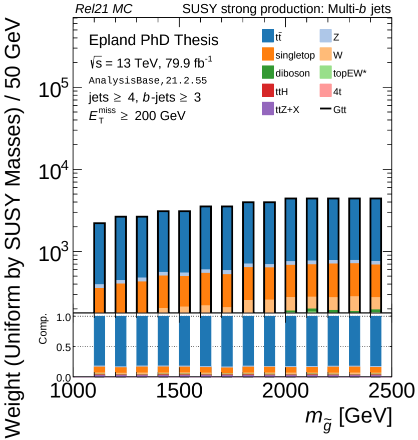
$p_{\mathbf{T}}^{b\text{-jet}\,4}$

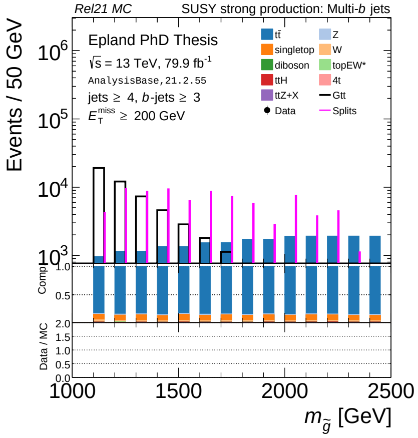




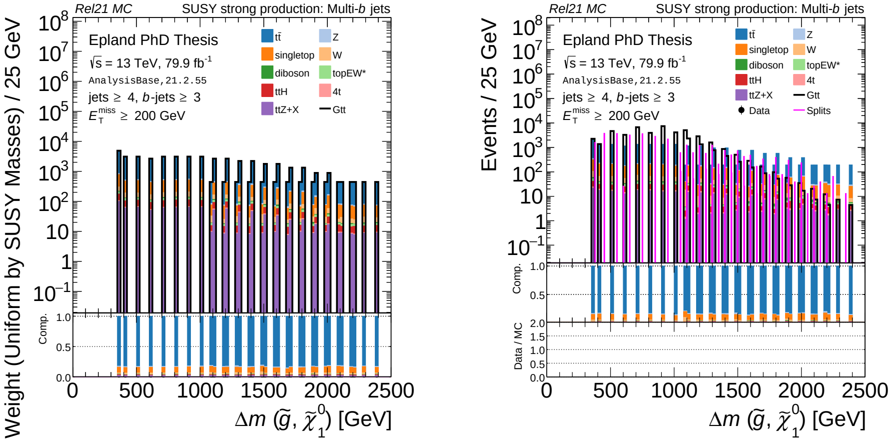
Parameters

$m_{\tilde{g}}$


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Weight (Uniform by SUSY Masses) / 50 GeV
Comp.
m_g̃ [GeV]

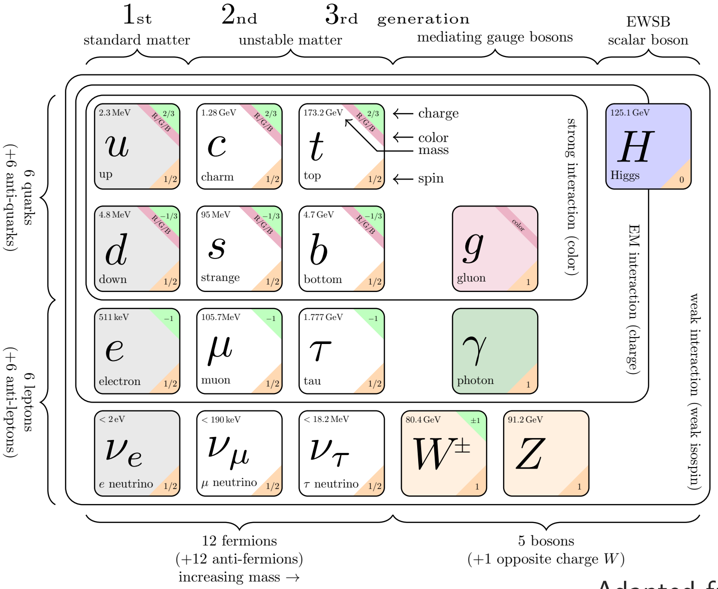

Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Data
Splits
Events / 50 GeV
Comp.
Data / MC
m_g̃ [GeV]

Δm


Rel21 MC
SUSY strong production: Multi-b jets
Epland PhD Thesis
√s = 13 TeV, 79.9 fb-1
AnalysisBase,21.2.55
jets ≥ 4, b-jets ≥ 3
E_T^miss ≥ 200 GeV
tt̄
Z
singletop
W
diboson
topEW*
ttH
4t
ttZ+X
Gtt
Data
Splits
Weight (Uniform by SUSY Masses) / 25 GeV
Events / 25 GeV
Comp.
Data / MC
Δm (g̃, χ̃⁰₁) [GeV]

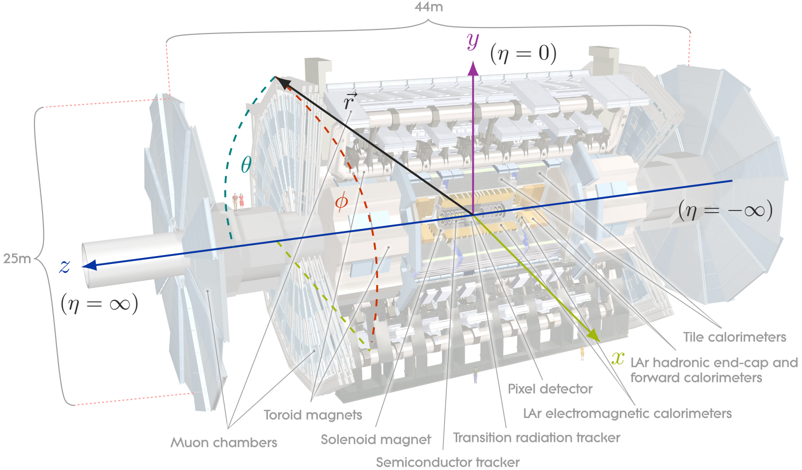
Misc.

Particles of the SM



Adapted from [13]

The ATLAS Coordinate System



[3]

References

[1] ATLAS Collaboration, Summary plots from the ATLAS Standard Model physics group, July, 2018. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM.

[2] Service graphique, CERN, https://cds.cern.ch/record/1708849.

[3] J. Pequenao, Computer generated image of the whole ATLAS detector, Mar, 2008. https://cds.cern.ch/record/1095924.

[4] J. Pequenao and P. Schaffner, An computer generated image representing how ATLAS detects particles, Jan, 2013. https://cds.cern.ch/record/1505342.

[5] ATLAS Collaboration, ATLAS-CONF-2018-041, 2018, https://cds.cern.ch/record/2632347.

[6] T. Chen and C. Guestrin, CoRR **abs/1603.02754** (2016), arXiv:1603.02754 [cs.LG].

[7] M. Baak, G. J. Besjes, D. Côte, A. Koutsman, J. Lorenz, and D. Short, Eur. Phys. J. **C75** (2015) 153, arXiv:1410.1280 [hep-ex].

[8] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, Journal of Statistical Mechanics: Theory and Experiment **2008** (2008) P10008, http://stacks.iop.org/1742-5468/2008/i=10/a=P10008.

[9] T. Aynaud, python-louvain, Louvain Community Detection, 2011–. https://github.com/taynaud/python-louvain.

[10] S. P. Martin, Adv. Ser. Direct. High Energy Phys. **18** (1998) 1, arXiv:hep-ph/9709356 [hep-ph].

[11] Particle Data Group Collaboration, C. Patrignani et al., Chin. Phys. **C40** (2016) 100001.

[12] ATLAS Collaboration, Summary plots from the ATLAS Supersymmetry physics group, July, 2018. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SUSY.

[13] D. Galbraith and C. Burgard, Standard model of the standard model, CERN Webfest 2012. http://www.texample.net/tikz/examples/model-physics.