Prospects for a Search of Compressed Electroweak Supersymmetry Using Soft Photons with the ATLAS Detector

Oral Qualifying Examination

Nathan Kang

Santa Cruz Institute for Particle Physics
University of California, Santa Cruz

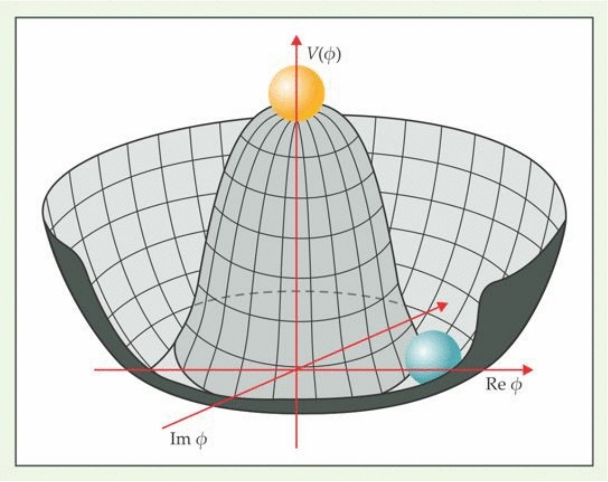
Physics Background and Motivation

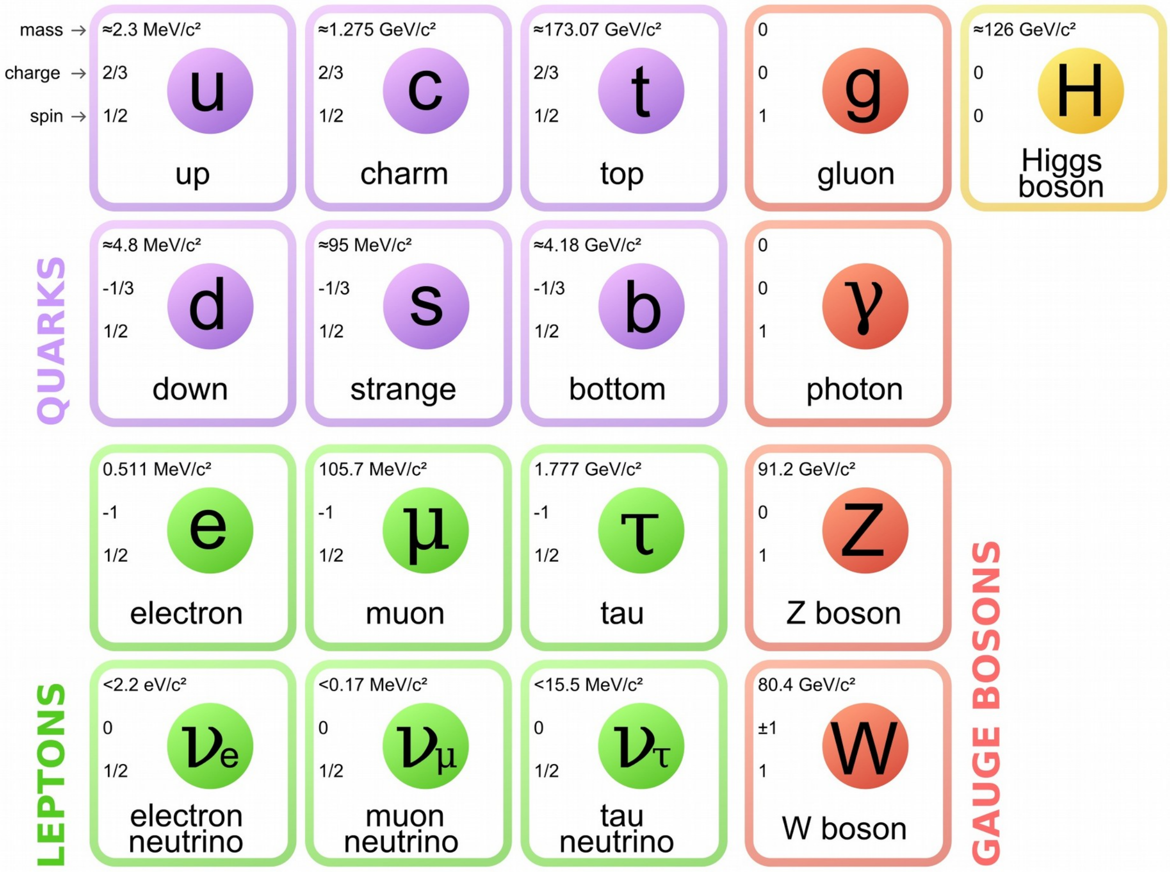
Standard Model

- The Standard Model of particle physics is a quantum field theory with the symmetry group **$SU(3)_C$ x $SU(2)_L$ x $U(1)_Y$**

- Elementary particles are the field quanta

- The Higgs mechanism describes how spontaneous symmetry breaking gives mass to the weak bosons and fermions

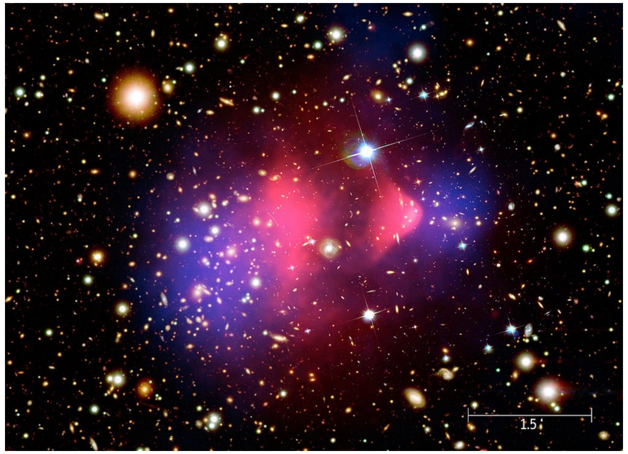


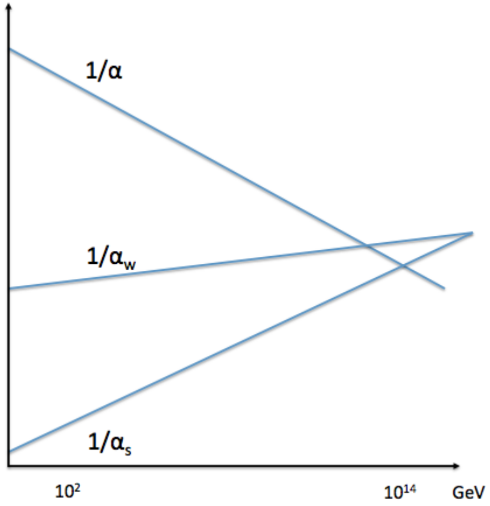


UC SANTA CRUZ ATLAS EXPERIMENT

Standard Model

- The Standard Model is very successful in describing many experimental observations but leaves open some questions/issues
  - Hierarchy problem: Higgs boson mass renormalization includes large quantum corrections that would require fine-tuning
  - Dark matter: absence of particles that fulfill the role of dark matter according to cosmological observations
  - Gauge coupling unification: failure of running couplings to unify at high energy scale in the context of Grand Unified Theories





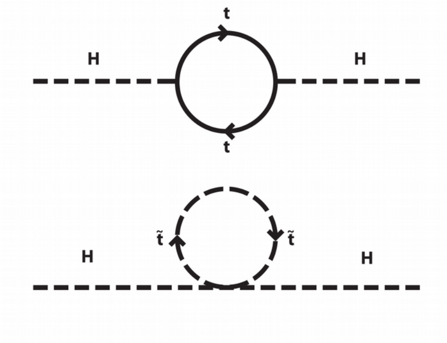
Supersymmetry

- Supersymmetry postulates a symmetry between bosons and fermions such that SM particles have superpartners that have the same quantum numbers except spin (differs by ½) and a new quantum number R-parity

- SUSY should be broken so that the superpartners can differ in mass to match observation

- Welcomed consequences
  - Hierarchy problem: cancellation of large corrections to mass from the inclusion of superpartners
  - Dark matter: the lightest supersymmetric particle is stable if R-parity is conserved thus making it a candidate for dark matter
  - Gauge coupling unification: modification to running couplings allowing for unification

Names	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	$+1$	$H_u^0\ H_d^0\ H_u^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
squarks	0	-1	$\tilde{u}_L\ \tilde{u}_R\ \tilde{d}_L\ \tilde{d}_R$	(same)
			$\tilde{s}_L\ \tilde{s}_R\ \tilde{c}_L\ \tilde{c}_R$	(same)
			$\tilde{t}_L\ \tilde{t}_R\ \tilde{b}_L\ \tilde{b}_R$	$\tilde{t}_1\ \tilde{t}_2\ \tilde{b}_1\ \tilde{b}_2$
sleptons	0	-1	$\tilde{e}_L\ \tilde{e}_R\ \tilde{\nu}_e$	(same)
			$\tilde{\mu}_L\ \tilde{\mu}_R\ \tilde{\nu}_\mu$	(same)
			$\tilde{\tau}_L\ \tilde{\tau}_R\ \tilde{\nu}_\tau$	$\tilde{\tau}_1\ \tilde{\tau}_2\ \tilde{\nu}_\tau$
neutralinos	1/2	-1	$\tilde{B}^0\ \tilde{W}^0\ \tilde{H}_u^0\ \tilde{H}_d^0$	$\tilde{N}_1\ \tilde{N}_2\ \tilde{N}_3\ \tilde{N}_4$
charginos	1/2	-1	$\tilde{W}^\pm\ \tilde{H}_u^+\ \tilde{H}_d^-$	$\tilde{C}_1^\pm\ \tilde{C}_2^\pm$
gluino	1/2	-1	$\tilde{g}$	(same)
goldstino (gravitino)	1/2 (3/2)	-1	$\tilde{G}$	(same)

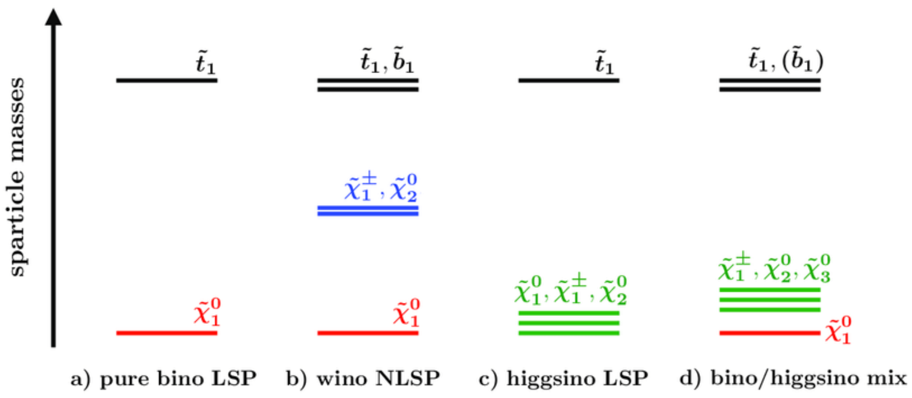


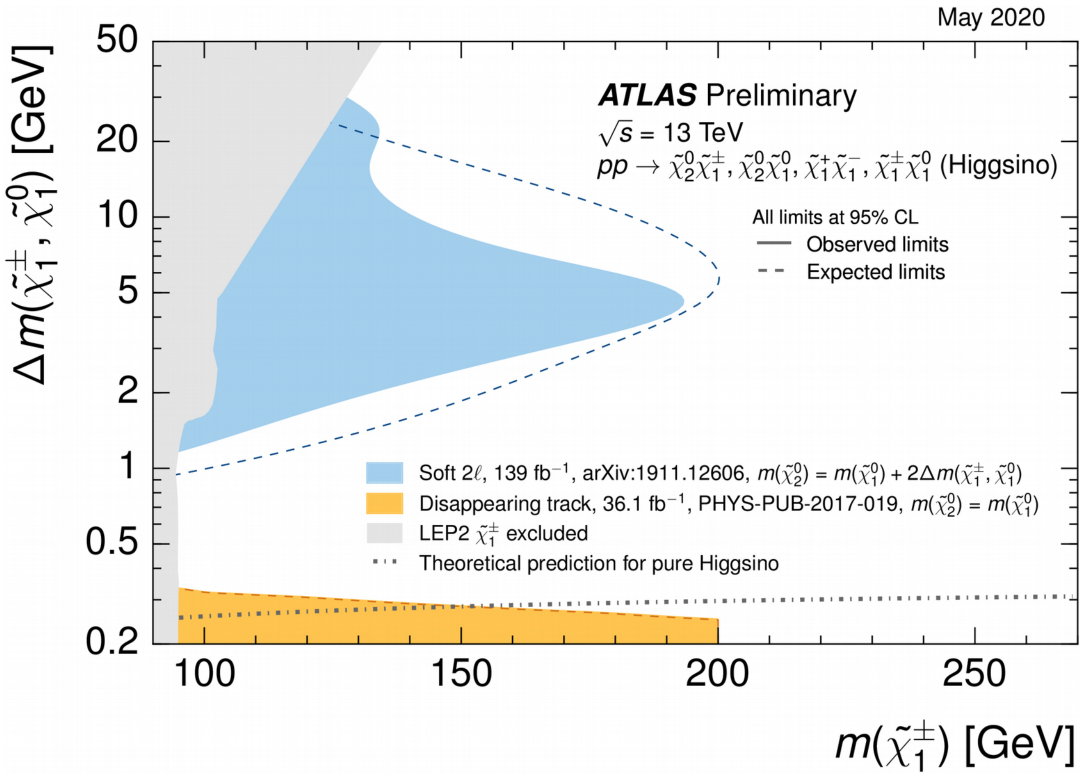
UC SANTA CRUZ
ATLAS EXPERIMENT

Compressed Electroweak SUSY

- We consider models based on the MSSM in which new electroweak states (neutralinos, charginos) are the lightest new particles and nearly mass degenerate from the electroweak symmetry
- For $\Delta m < \sim 300$ MeV the liftetime of the more massive states is long enough for a disappearing track signature
- For $\Delta m >$ a few GeV soft leptons signature from decay
- Existing searches have yet to pass LEP limits on mass splittings between this so we would like to find ways to get sensitivity here

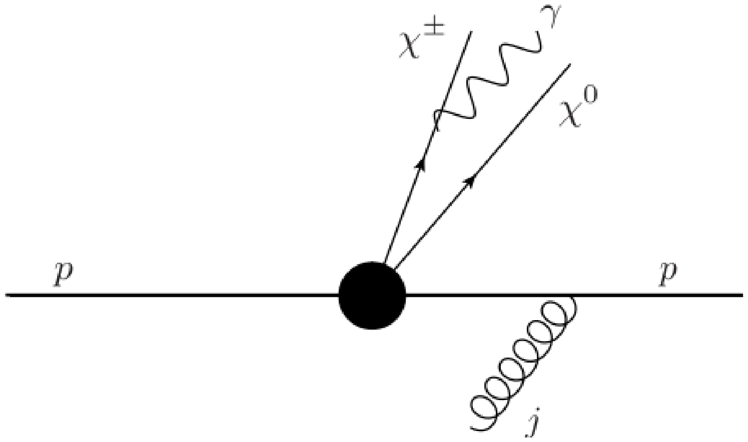
$$\mathbf{M}_{\tilde{N}} = \begin{pmatrix} M_1 & 0 & -g' v_d/\sqrt{2} & g' v_u/\sqrt{2} \\ 0 & M_2 & g v_d/\sqrt{2} & -g v_u/\sqrt{2} \\ -g' v_d/\sqrt{2} & g v_d/\sqrt{2} & 0 & -\mu \\ g' v_u/\sqrt{2} & -g v_u/\sqrt{2} & -\mu & 0 \end{pmatrix}. \qquad \mathbf{N}^* \mathbf{M}_{\tilde{N}} \mathbf{N}^{-1} = \begin{pmatrix} m_{\tilde{N}_1} & 0 & 0 & 0 \\ 0 & m_{\tilde{N}_2} & 0 & 0 \\ 0 & 0 & m_{\tilde{N}_3} & 0 \\ 0 & 0 & 0 & m_{\tilde{N}_4} \end{pmatrix}$$





Search strategy

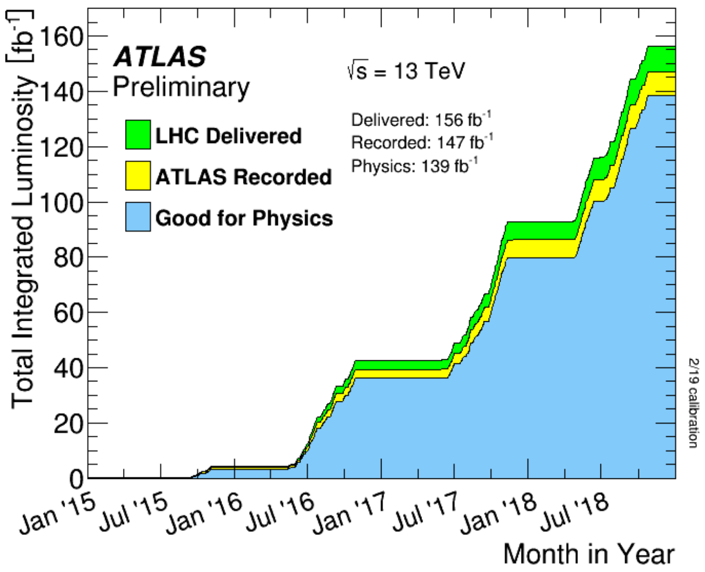
- We investigate the prospects for a search based on arXiv:1605.00658v2 [hep-ph] (Ismail, Izaguirre, Shuve)
- Look for events with the signature of missing transverse energy (E_T^{miss}) recoiling against a hard jet such that photons radiated from charginos prior to their decay to the lightest neutralino would be preferentially aligned with E_t^{miss}
- Although requiring the photon will lead to a smaller signal rate the signal-to-background ratio is expected to increase relative to the E_t^{miss} + monojet search

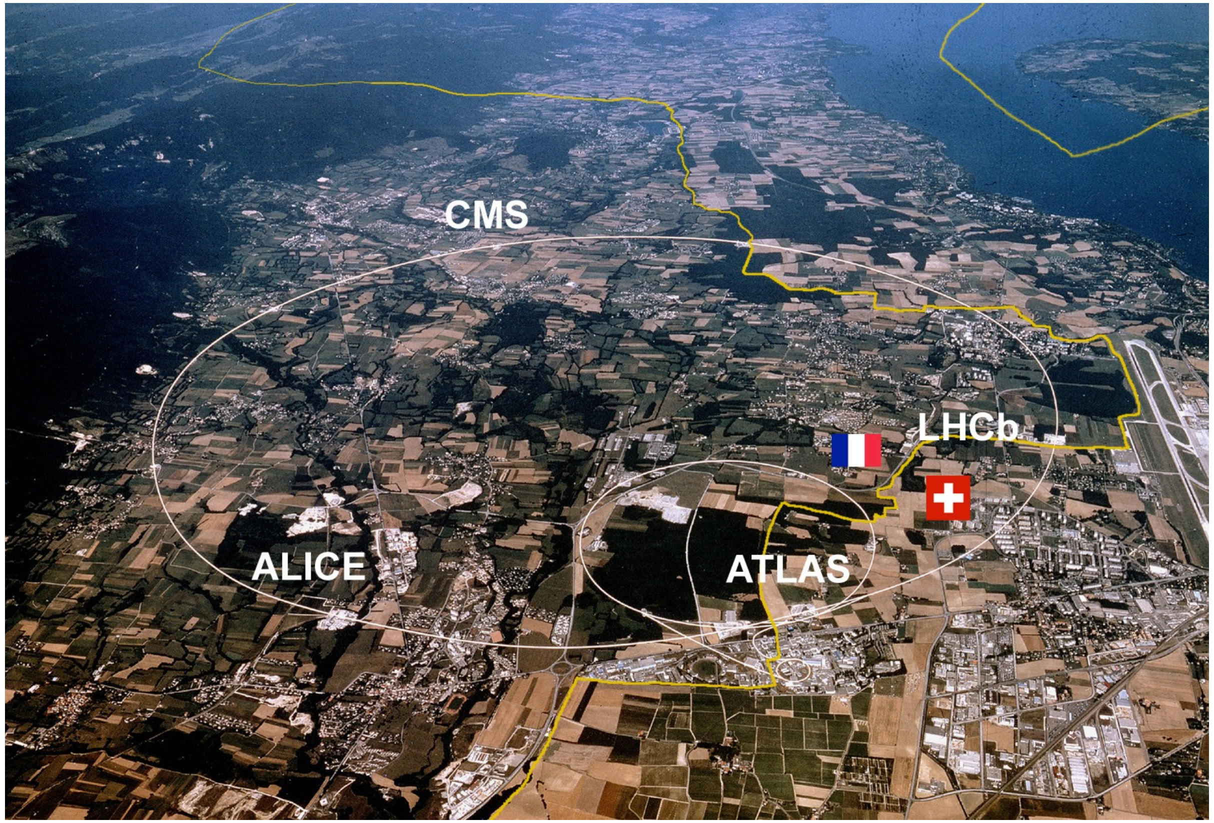


LHC and ATLAS

LHC

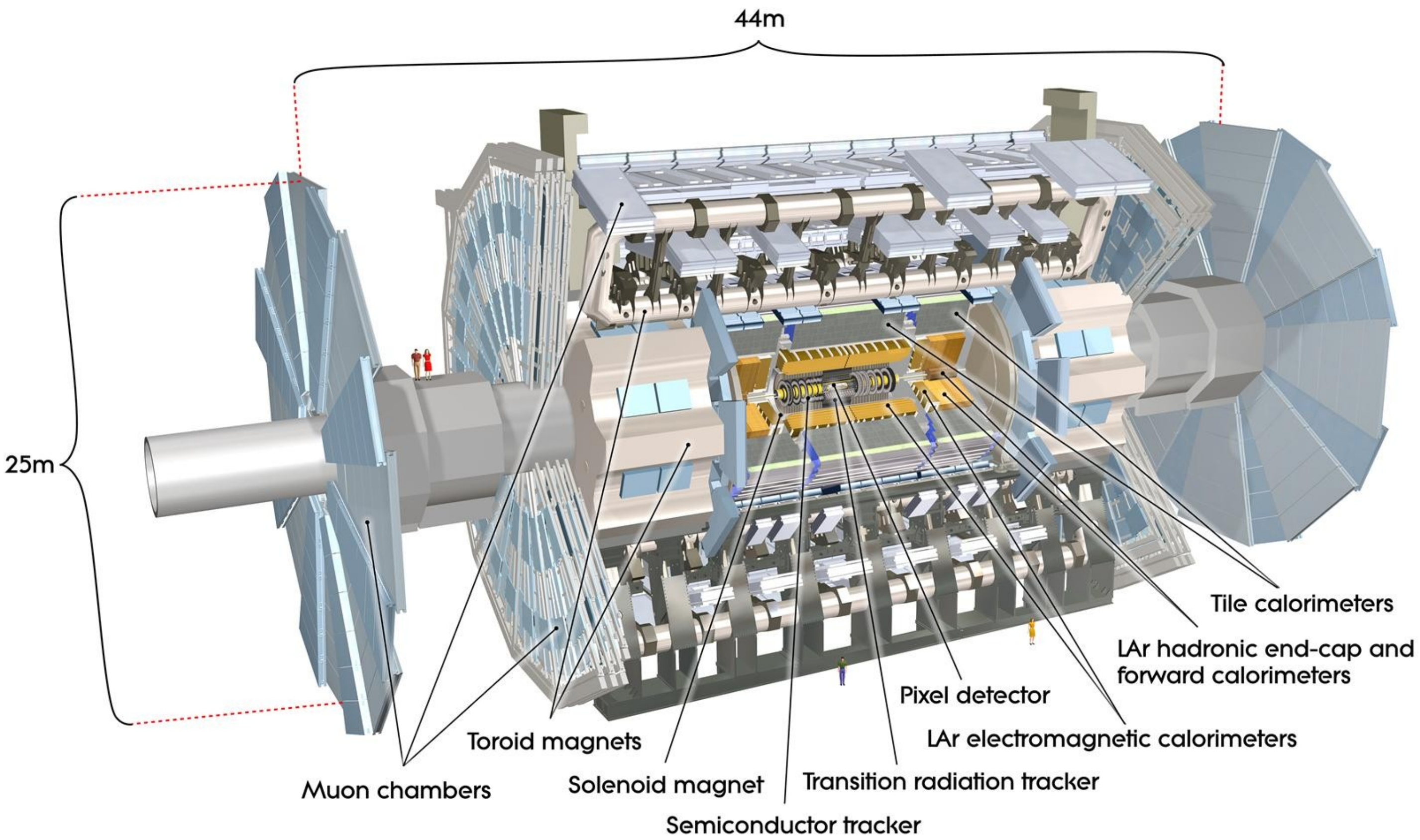
- The Large Hadron Collider accelerates bunches of protons that collide at the beam crossing points every 25 ns
- For Run 2 (2015-2018) the center of mass energy was 13 TeV and integrated luminosity of ~140 fb^{-1} good for physics analysis





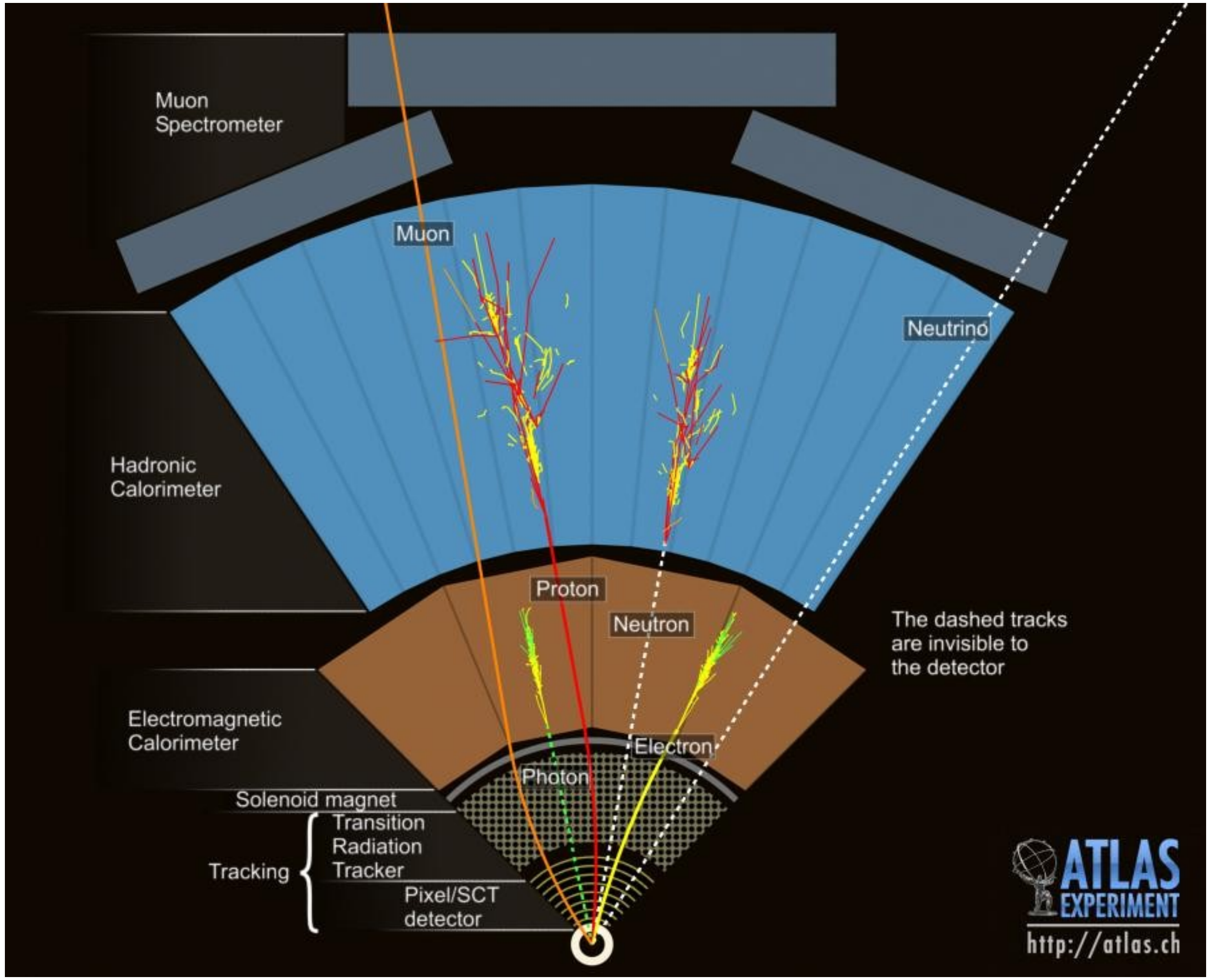
ATLAS

- Main detector components
  - Inner Detector: composed of Pixel Detector, SCT, and TRT in a solenoid field; measures charged particle tracks and momenta
  - Calorimeters: LAr Calorimeter (EM and hadronic end-cap+forward) and Tile Calorimeter (hadronic); measure energy of electrons/photons and hadrons
  - Muon Spectrometer: muons can traverse the previous components unimpeded; muons in a toroidal field to measure tracks and momenta

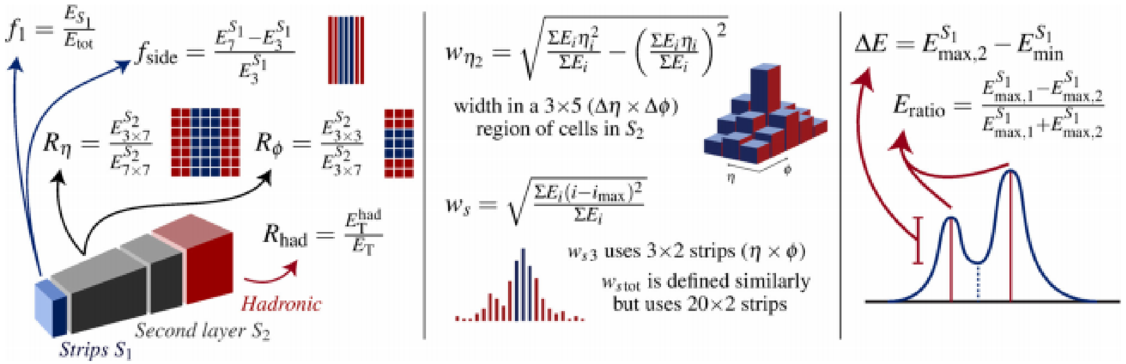


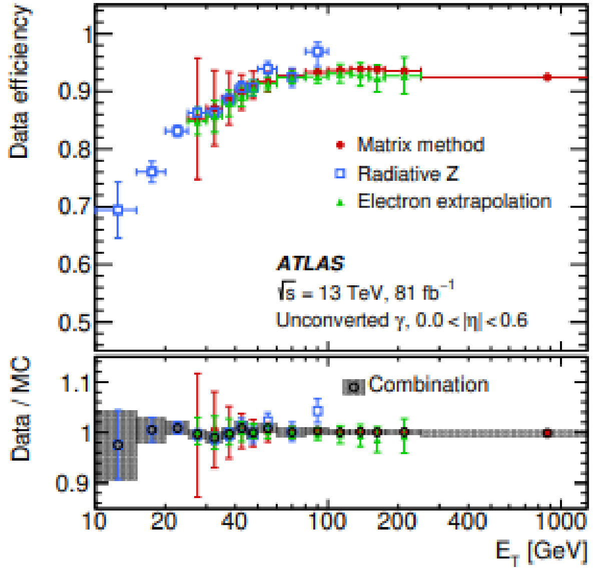
ATLAS

- Detector cross section



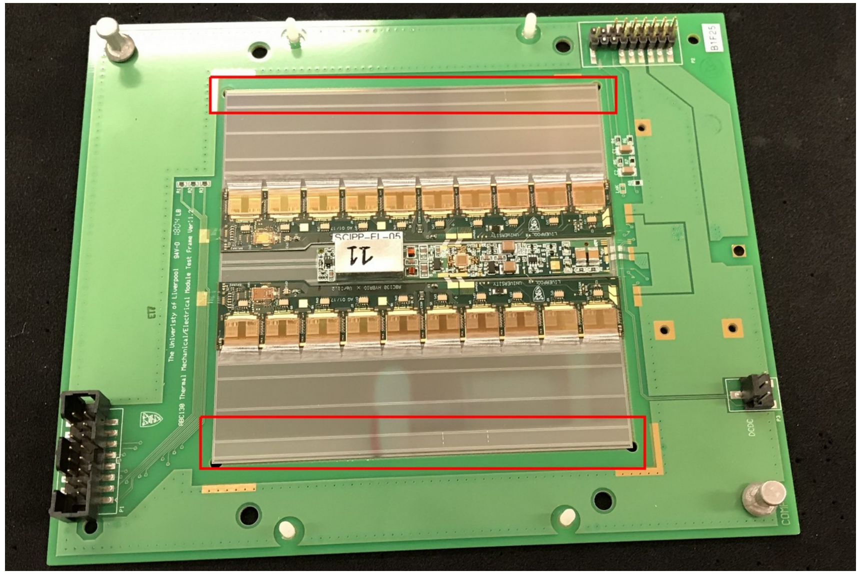
- Photon variables from calorimeters used for identification and an example of efficiency plot for Tight ID

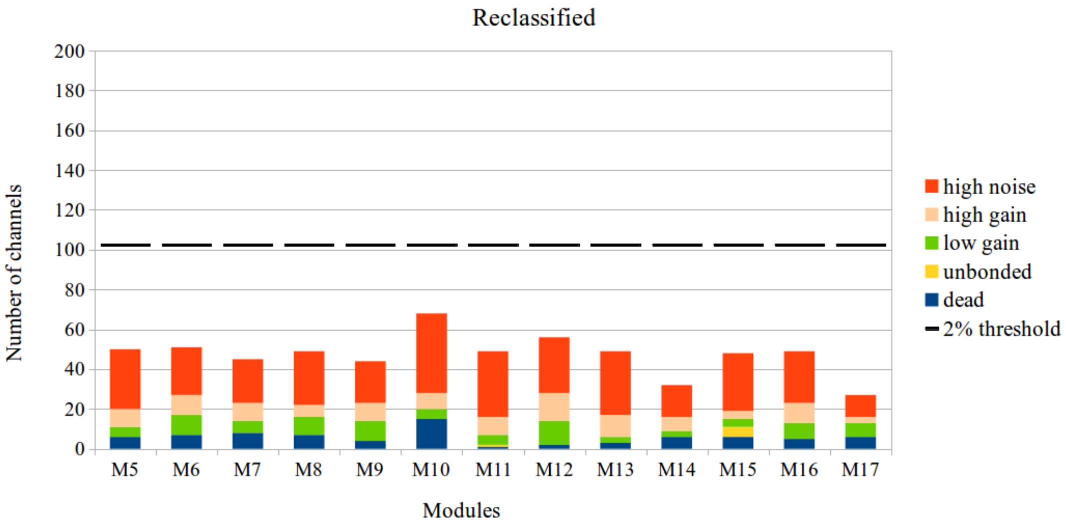




ATLAS Upgrade

- The High Luminosity LHC will be an upgrade to increase the luminosity by factor of 10 and operational in 2027(?)
- As part of the upgrade effort I have been working on the ATLAS Inner Tracker upgrade that will replace the Inner Detector
- Running electrical tests of silicon strip sensors, readout ASICs, and modules
- Development of test-related software and documenting quality control procedures

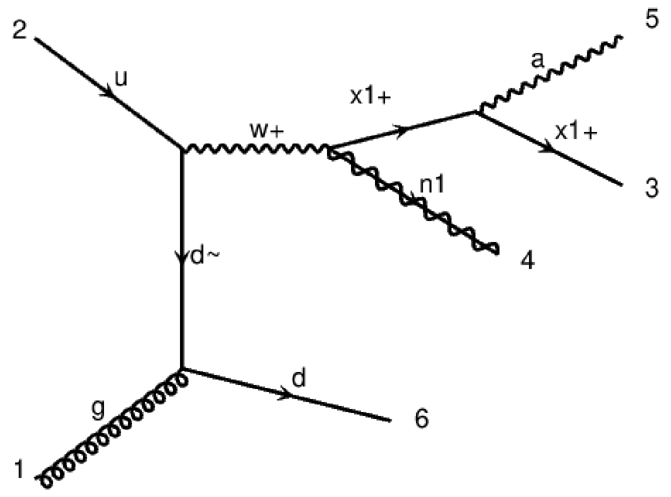


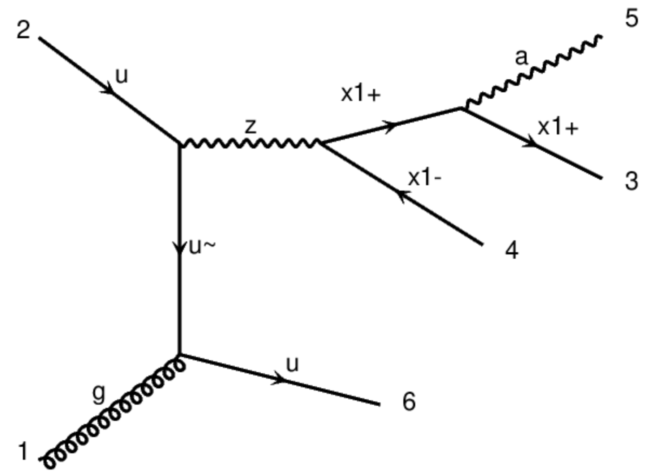


Details and Results of Prospect Study

Signal sample

- Main signal model
  - Simplified higgsino model with other SUSY particles set to much higher masses
  - Masses: N_1 100 GeV, C_1 100.5 GeV, N_2 101 GeV
- Generated 100,000 events for each of the following processes using MadGraph5 for $\sqrt{s}$ = 13 TeV
  - $p\,p \rightarrow C_1 + N_1 + j + \gamma$
  - $p\,p \rightarrow C_1 + N_2 + j + \gamma$
  - $p\,p \rightarrow C_1 + C_1 + j + \gamma$
- Pythia8 with CKKW-L merging for parton shower and hadronization



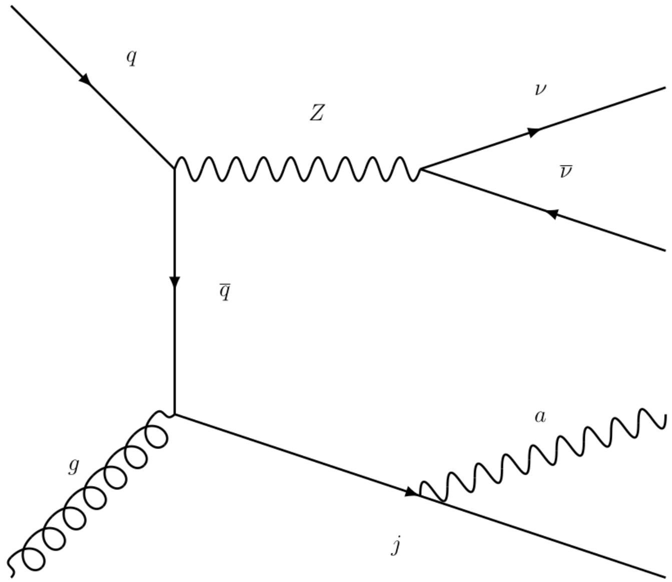


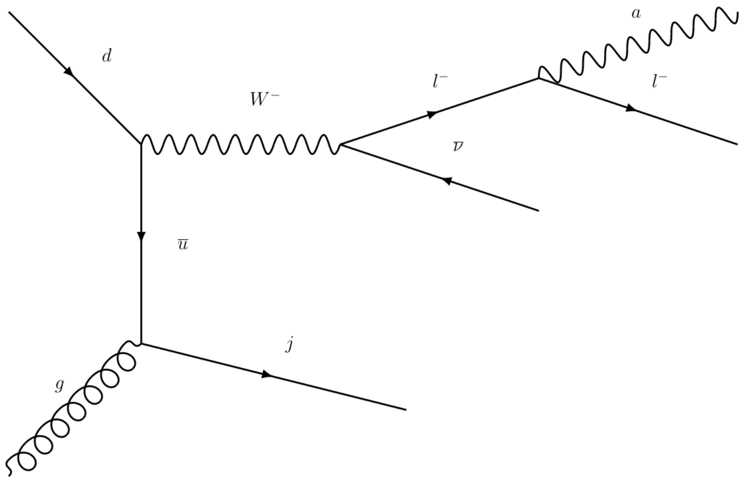
Background samples

- Relevant backgrounds include those with real missing energy (neutrinos) such as Z(νν)+γ+jet and W(lν)+γ+jet
  - Z(νν)+γ+jet mimics the targeted signal but the neutrinos do not radiate photons
  - W(lν)+γ+jet can also mimic the signal if the lepton is missed

- We use Sherpa 2.2.8 V+γ datasets from ATLAS MC production
  - ~96 million events for W(lν)+γ for each generation of lepton e, μ, τ
  - ~16 million events for Z(νν)+γ

- Background contribution from fake photons may also be a significant factor and an estimate using a data driven method will be shown





Preselection and variables

- We make truth samples for the signal and background with various variables after preselection

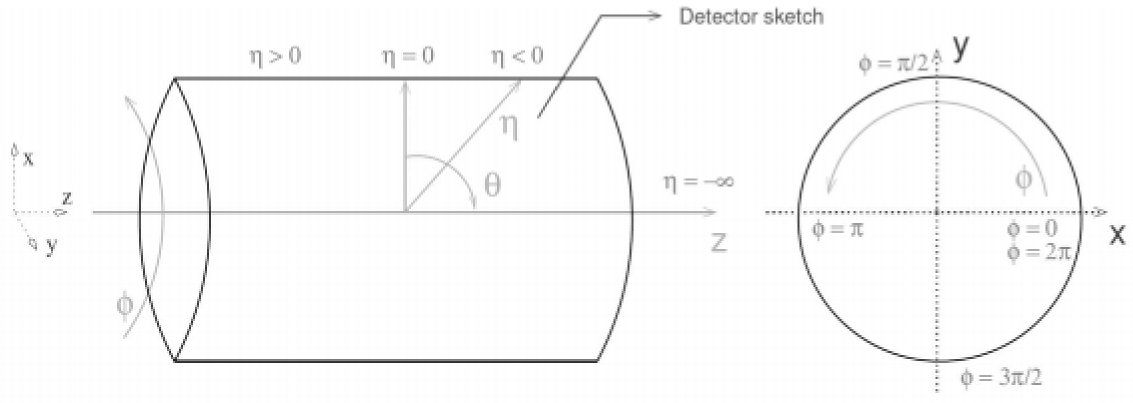
Preselection

Variable	Requirement
E_T^{miss}	> 75 GeV
Number of jets	> 0
Leading jet $\|\eta\|$	< 2.5
Photon p_T	> 7 GeV
Photon $\|\eta\|$	< 2.5
Photon ptcone30	< 2 GeV
Photon etcone20	< 2 GeV
Lepton vetoes	Pass

Lepton vetoes

	Variable	Requirement
electron		
	p_T	> 4.5 GeV
	$\|\eta\|$	< 2.5
	$\frac{ptcone30}{p_T}$	< 0.06
	$\frac{etcone20}{p_T}$	< 0.06
muon		
	p_T	> 3 GeV
	$\|\eta\|$	< 2.5
	$\frac{ptcone30}{p_T}$	< 0.04
	$\frac{etcone20}{p_T}$	< 0.15
tau		
	p_T	> 20 GeV
	$\|\eta\|$	< 2.5

Object	Variable
E_T^{miss}	E_T^{miss}
jets	multiplicity
Leading p_T jet j_1	p_T
	$\|\eta\|$
	energy
	mass
photon	p_T
	$\|\eta\|$
	energy
	ptcone30
	etcone20
(E_T^{miss},photon)	$\|\Delta\phi\|$
(E_T^{miss}, j_1)	ΔR
(j_1,photon)	m_T
	ratios of variables



Results from optimization tool

- We use an optimization tool (https://github.com/kratsg/optimization) to try and maximize the sensitivity
- This is a cut-based optimization in which variables are scanned over a user-specified range of values to determine which combination of cuts maximizes the significance
- Data inputs are the ROOT ntuples with collection of various variables after preselection applied for signal and background
- Events are weighted to 140 fb^{-1} and a relative background uncertainty of 25% is used for calculating the significance from ATL-COM-GEN-2018-026

$$Z = \begin{cases} +\sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1+\frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} & \text{if } n \geqslant b \\ -\sqrt{2\left(n\ln\left[\frac{n(b+\sigma^2)}{b^2+n\sigma^2}\right] - \frac{b^2}{\sigma^2}\ln\left[1+\frac{\sigma^2(n-b)}{b(b+\sigma^2)}\right]\right)} & \text{if } n < b. \end{cases}$$


```
[
    {
        "selections": "br_met_met > {0}",
        "st3": [
            [100, 800, 100]
        ]
    },
    {
        "selections": "br_metphoton_dphi < {0}",
        "st3": [
            [0.2, 1.8, 0.2]
        ]
    },
    {
        "selections": "br_metjet1_dphi > {0}",
        "st3": [
            [3.00, 3.15, 0.025]
        ]
    },
    {
        "selections": "br_jet1_pt > {0}",
        "st3": [
            [100, 800, 100]
        ]
    },
    {
        "selections": "br_metphoton_dr < {0}",
        "st3": [
            [0.2, 2.2, 0.2]
        ]
    },
    {
        "selections": "br_jets_n < {0}",
        "st3": [
            [2, 6, 1]
        ]
    }
]
```

Results from optimization tool

Variable	Requirement
E_T^{miss}	> 600 GeV
Leading jet p_T	> 400 GeV
Number of jets	< 4
$\lvert\Delta\phi\rvert(E_T^{miss}, j_1)$	> 3.025
$\lvert\Delta\phi\rvert(E_T^{miss}, \gamma)$	< 0.8
$\Delta R(E_T^{miss}, \gamma)$	< 1.6

- Combination of variables and cuts that has yielded the highest significance

		Raw Events	Weighted Events
signal		426	3.46
background		3147	19.33
	$W(e\nu)+\gamma$	904	5.17
	$W(\mu\nu)+\gamma$	782	4.32
	$W(t\nu)+\gamma$	1129	4.47
	$Z(\nu\bar{\nu})+\gamma$	332	5.38

- The significance is 0.51 with the event yields shown

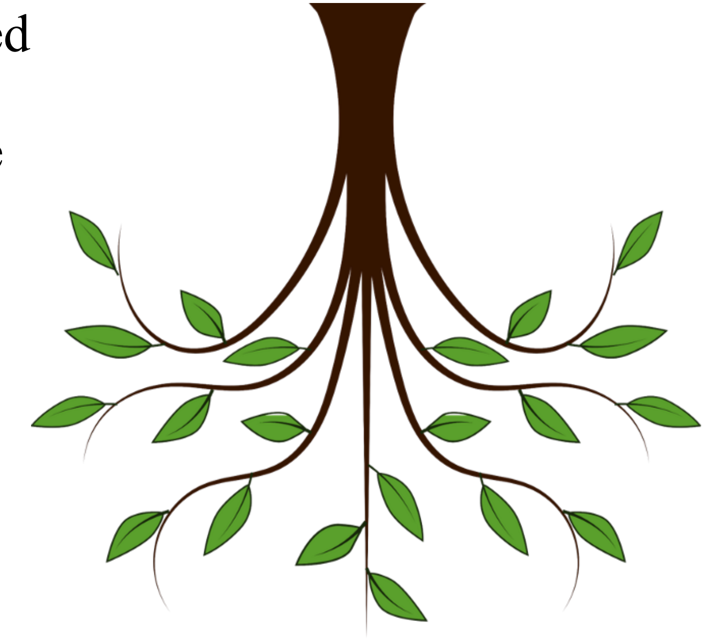
Results from boosted decision trees

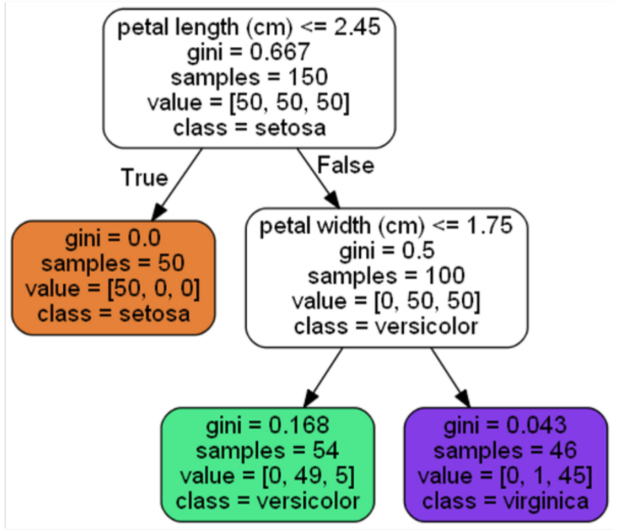
- Attempt to gain more sensitivity than what was obtained with cut-based optimization by trying a machine learning approach using boosted decision trees

- Decision trees take input data in which samples contain features of interest and are assigned different classes

- At the tree nodes the features are checked over a range of values to determine the optimal split according to some measure (ex. Gini impurity) such that the purest subsets are produced

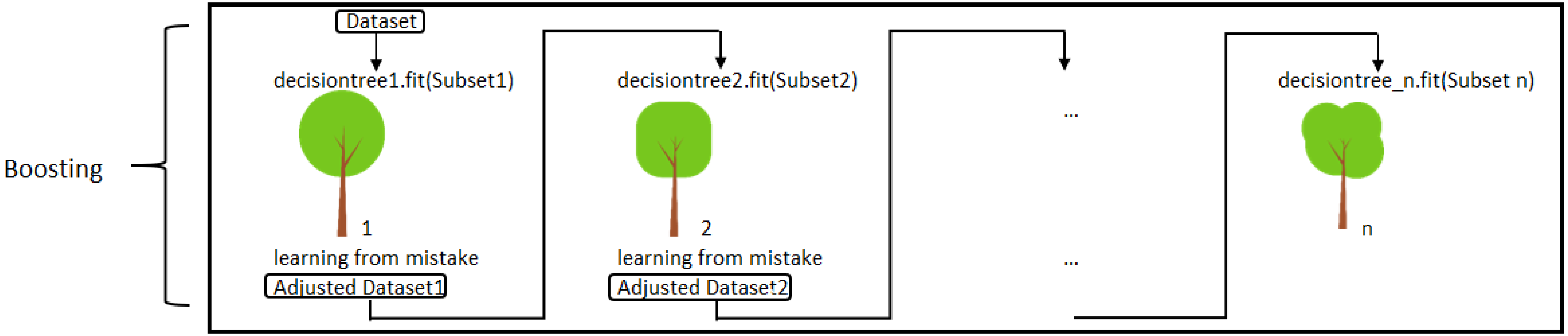
- Leaves are nodes without anything growing from them either because the node is pure or the maximum tree depth has been reached





Results from boosted decision trees

- A single decision tree on its own may not be very good as a classifier
- Boosting algorithms try to combine many weak learners (trees) to make a strong classifier
- Trees are generated sequentially and the tree output is assigned a weight according to the accuracy of its classification; the output of the BDT is a weighted sum of these individual outputs
- The dataset is also weighted so that misclassified samples can be given more importance in the subsequently generated trees

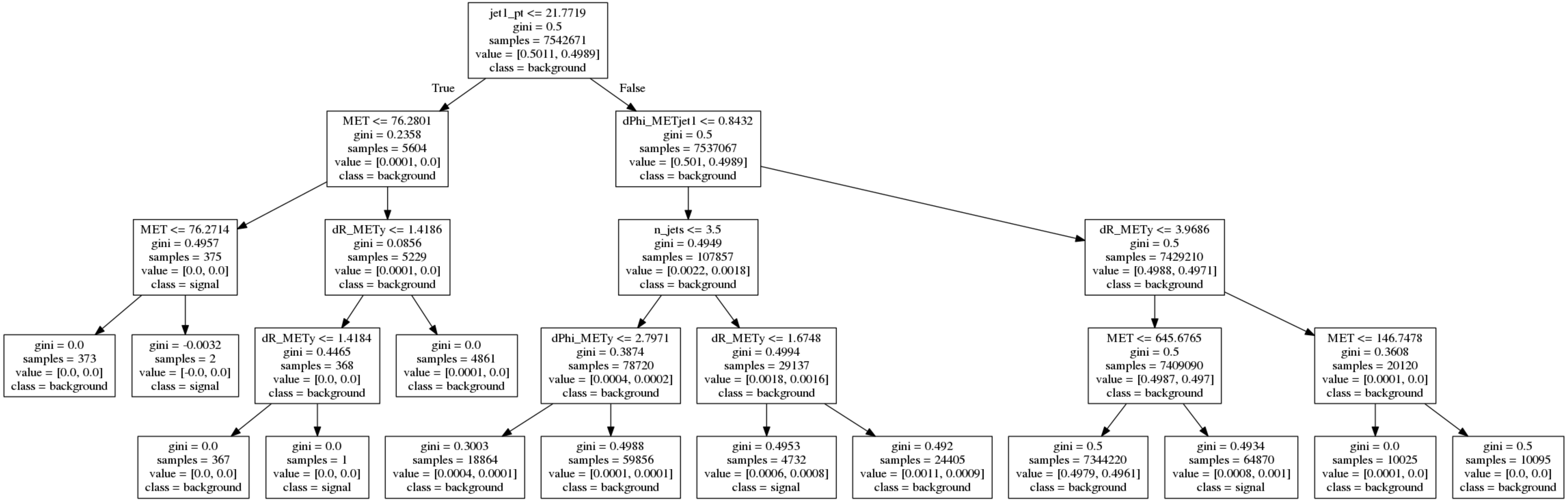


Results from boosted decision trees

- Using the machine learning library scikit-learn with AdaBoost to implement the BDT
- Input datasets from ROOT ntuples with various variables after preselection
- Two classes (signal or background) for samples
- Training with 90% of dataset and testing with 10%; trained with event weights scaled to 140 fb^{-1}
- Decision tree parameters
  - max tree depth: 3 or 4
- AdaBoost parameters
  - algorithm: SAMME (Stagewise Additive Modeling using a Multi-class Exponential loss function)
  - number of estimators (trees): 1000

Results from boosted decision trees

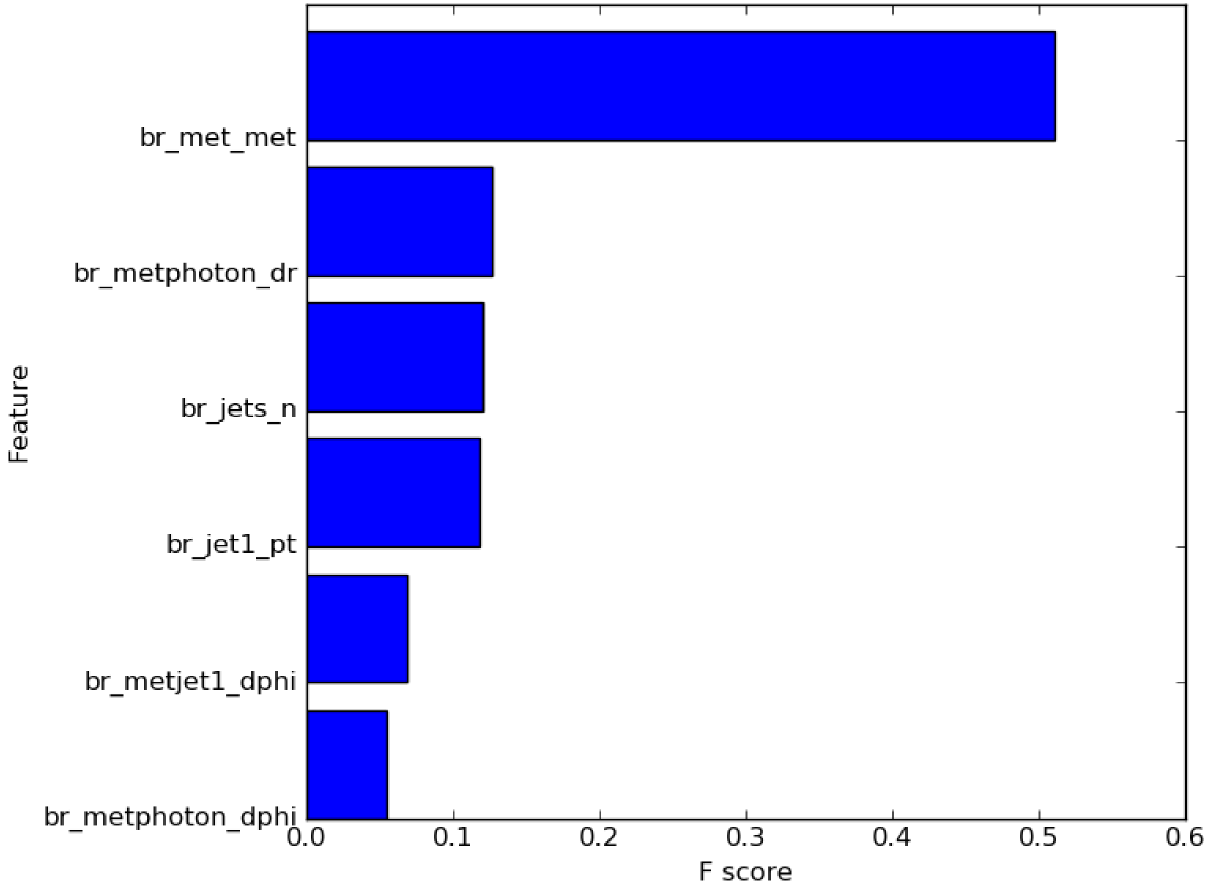
- Example of a single decision tree from BDT



Results from boosted decision trees

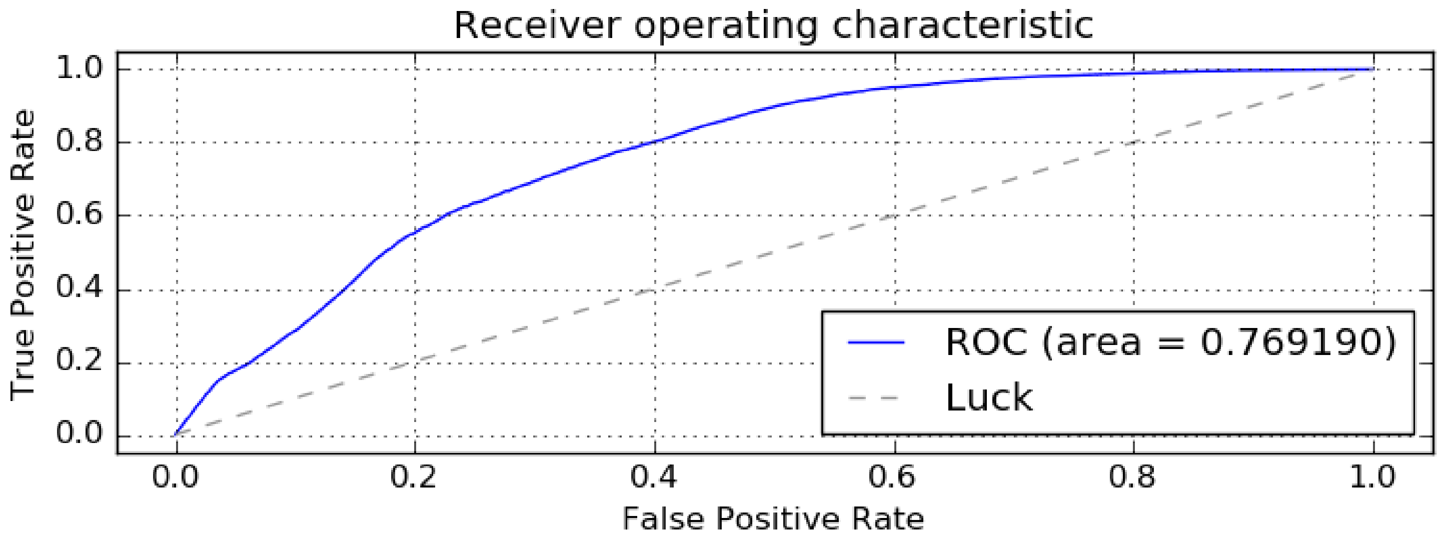
- After going through different combinations of variables we find the following combination gives the best result listed in order of relevance according to BDT

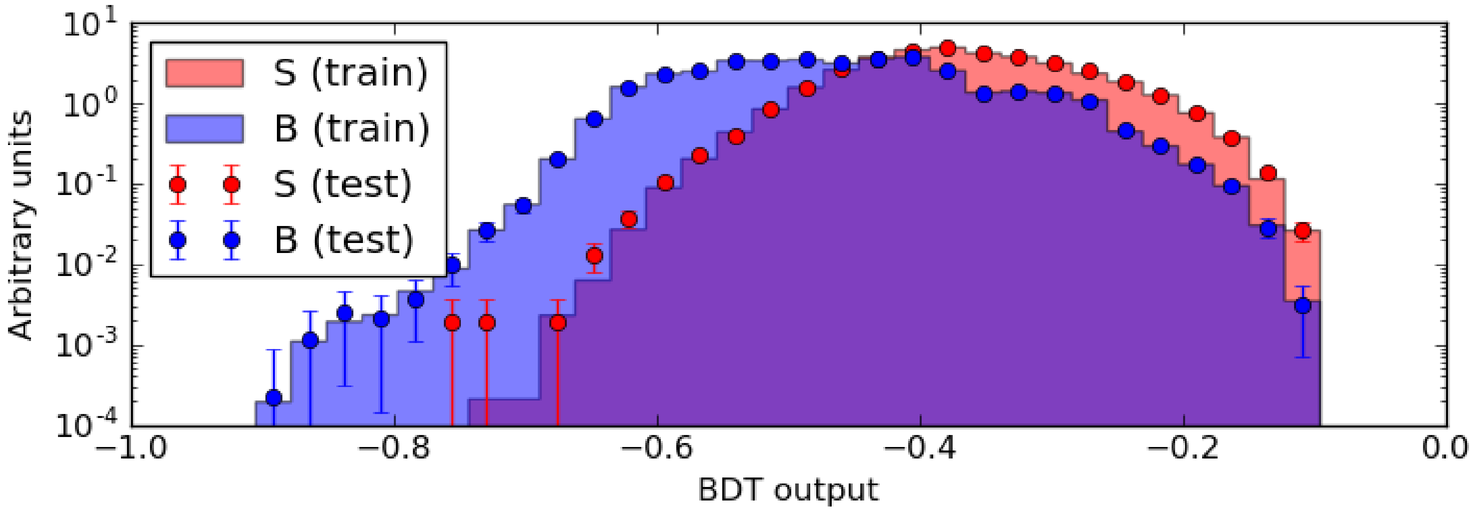
  1) Missing transverse energy
  2) $\Delta R(MET,\gamma)$
  3) number of jets
  4) Leading jet p_T
  5) $|\Delta\varphi|(MET,jet1)$
  6) $|\Delta\varphi|(MET,\gamma)$



Results from boosted decision trees

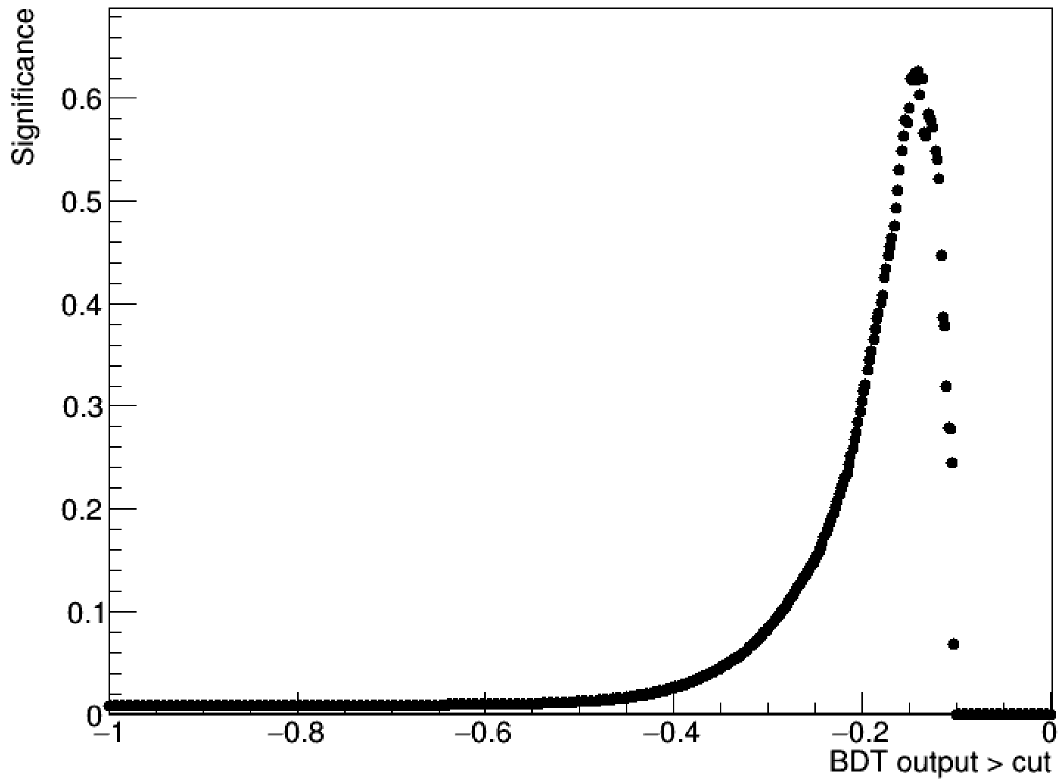
- ROC curve and BDT output scores of signal and background for train and test samples





Results from boosted decision trees

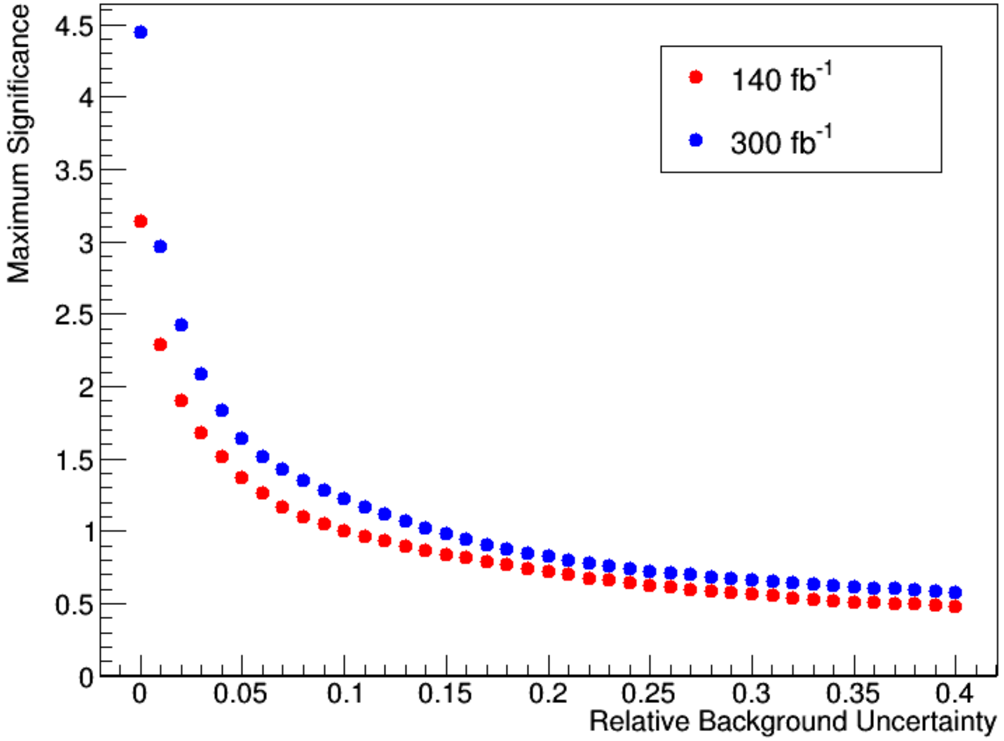
- Significance for 140 fb^{-1} vs cut on the BDT output value (25% relative background uncertainty)



- Peak significance of 0.63 with s = 4.36 and b = 19.57
- Comparing to the cut-based optimization using these same variables yields significance of ~0.5

Results from boosted decision trees

- Maximum significance as a function of relative background uncertainty for 140 fb^{-1} (Run 2) and 300 fb^{-1} (Run 3) for 100 GeV higgsino N_1 with $\Delta m(C_1,N_1) = 0.5$ GeV



- Projected significance for HL-LHC from arXiv:1605.00658v2 [hep-ph]

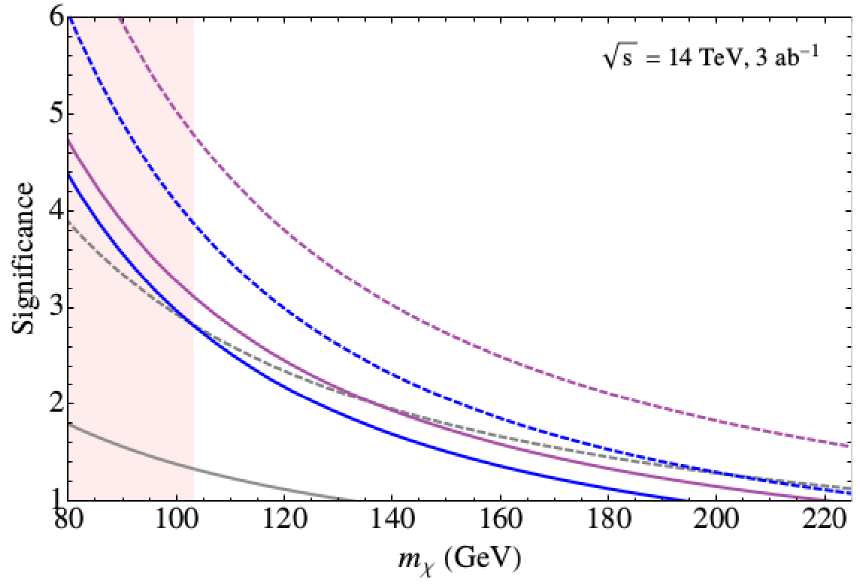


FIG. 4: Projected signal significance for the Higgsino doublet model with 3 ab^{-1} of integrated luminosity at the HL-LHC. Results are shown for the $\gamma + j + \not{E}_T$ search (blue) assuming either 5% (solid) or 2% (dashed) background systematic uncertainties. The estimated $j + \not{E}_T$ (gray) sensitivity is also shown for comparison, along with a naïve combination of $\gamma + j + \not{E}_T$ and $j + \not{E}_T$ sensitivities (purple). The shaded region is excluded by LEP [21].

Fake photon estimation

- Fake photon background was not incorporated in paper
- ABCD method is used to estimate the fake photon contribution to background using data
- Apply the selection on the right to all events
- Events with isolated leptons are vetoed
- ABCD regions are set up using PID and isolation as shown with A being the blinded signal region
- Isolation is defined as topoetcone20 < 0.065 * p_T and ptcone20 < 2 GeV

Variable	Requirement
E_T^{miss}	> 600 GeV
Leading jet p_T	> 400 GeV
Leading jet $\|\eta\|$	< 2.5
Photon p_T	> 7 GeV
Photon $\|\eta\|$	< 2.5
Photon PID	Loose
$\|\Delta\phi\|(E_T^{miss}, j_1)$	> 3.025
$\|\Delta\phi\|(E_T^{miss}, \gamma)$	< 0.8
$\|\Delta\phi\|/\Delta R(E_T^{miss}, \gamma)$	> 0.4

	Pass isolation	Fail isolation
Loose, Tight	A	C
Loose, Not Tight	B	D

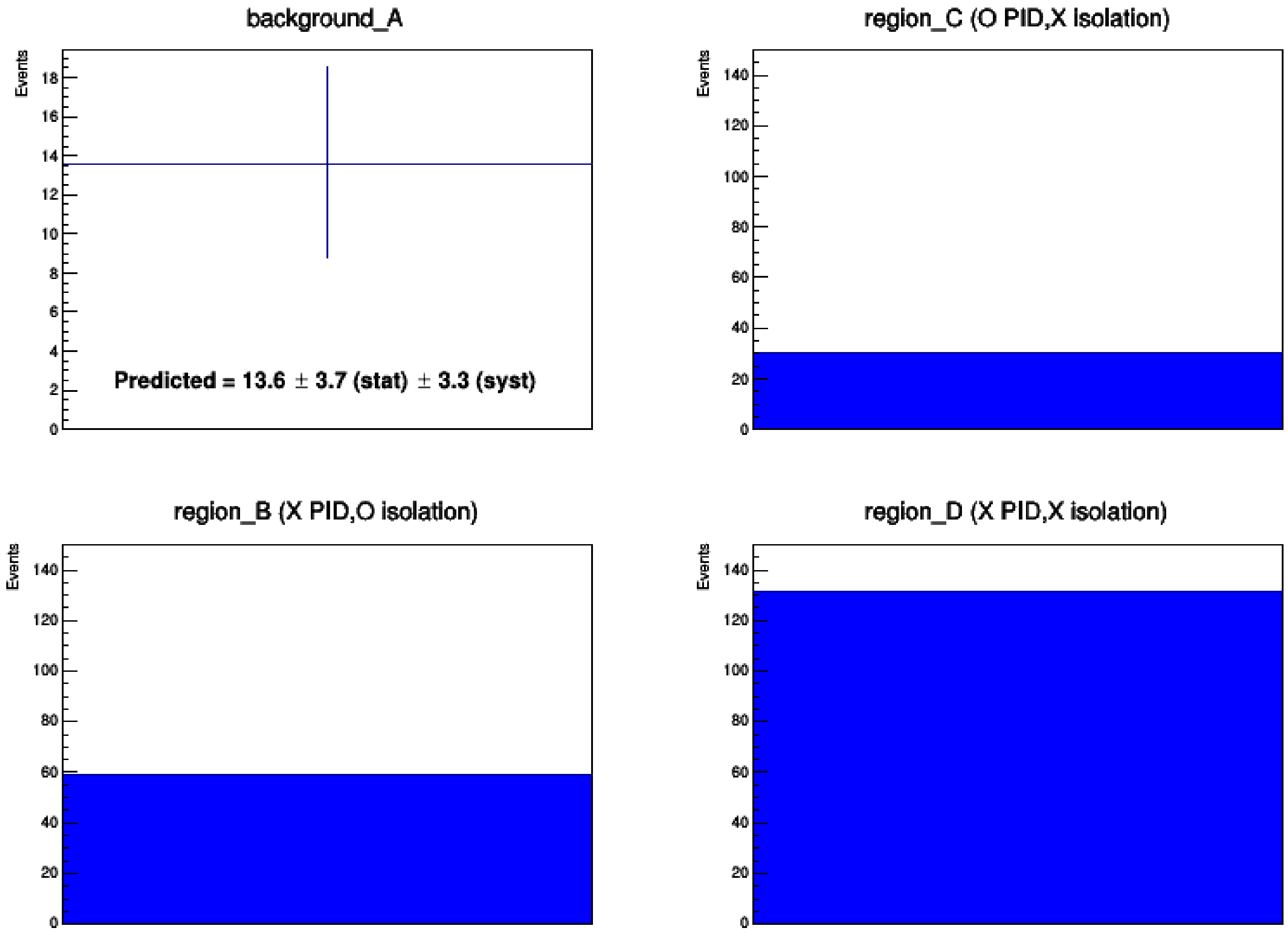
UC SANTA CRUZ ATLAS EXPERIMENT

Fake photon estimation

- 2018 data using EXOT5 derivation is used for the estimation
- The analysis framework xAODAnaHelpers is used to process the data
- Events are weighted to scale from the integrated luminosity for 2018 (58.5 fb-1) to full Run2 (140 fb-1)
- Reconstructed MC background samples for $(W\rightarrow l\nu)+\gamma$ and $(Z\rightarrow \nu\nu)+\gamma$ are also used to subtract off their contribution to the B,C, and D regions for the estimation
- The prediction for the fake photon contribution in the signal region can then be determined by the following equation under the assumption that the two variables used are uncorrelated for the background

$$N_A^{\mathrm{bkg}} = \frac{N_C^{\mathrm{bkg}}}{N_D^{\mathrm{bkg}}} N_B^{\mathrm{bkg}}$$

Fake photon estimation



- Prediction for the number of events from fake photon background is 13.6 ± 3.7(stat) ± 3.3(syst)
- Comparing with the value for V+γ background in the signal region from the cut-based optimization using truth samples b = 29.36 ± 7.34(25% uncertainty), we see in this case that the contribution from fake photons would be quite substantial

Summary and Future Work

- A study of the prospects for a compressed electroweak SUSY search using the ATLAS Run 2 dataset for the signature of E_t^{miss} + hard jet + photon was conducted

- The current results for the sensitivity of this channel for the models described are not as promising as we had hoped

- Pursuing this type of search does not seem worthwhile pending significant changes to the search strategy that could boost the sensitivity substantially

- Improvements to study that could be made
  - better optimizing BDT parameters
  - switching to another machine learning algorithm that may be more powerful
  - include more of the event information as data inputs
  - multi-bin fit for higher significance

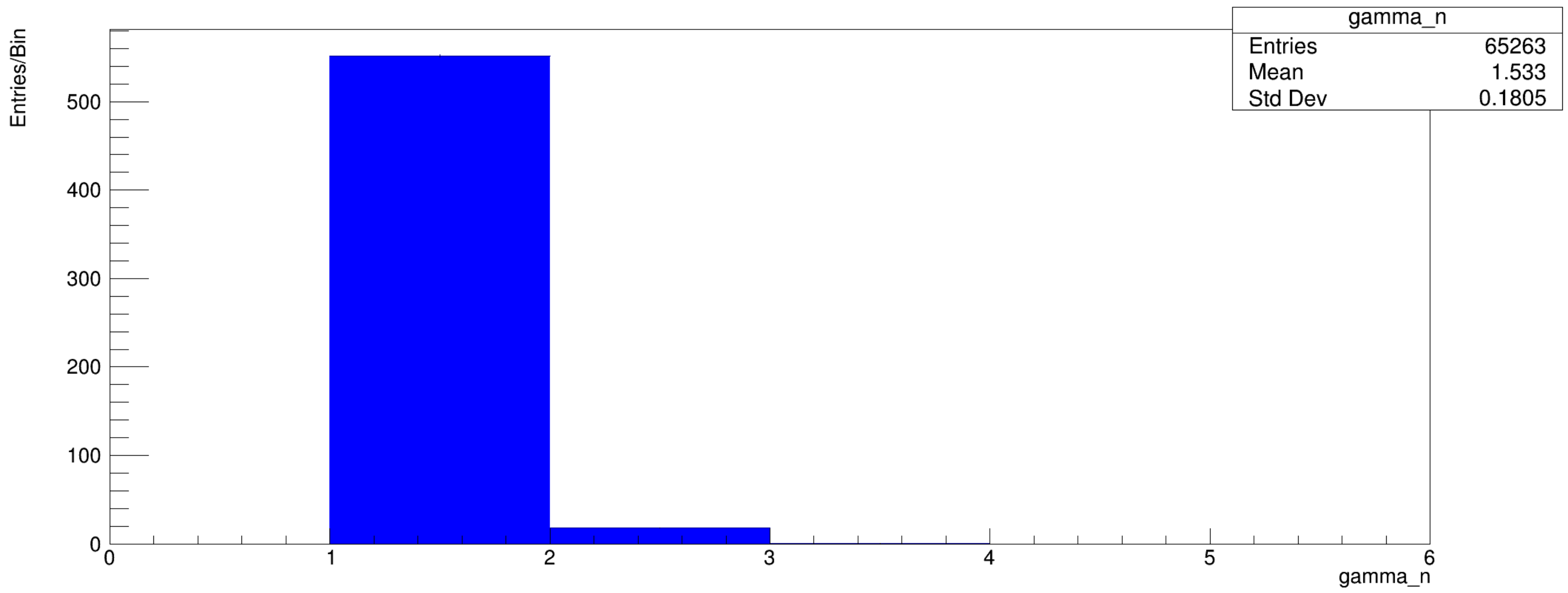
- The fake photon background appears to be non-negligible and needs consideration if pursuing an actual search

Acknowledgments

Many thanks to

- Faculty Advisor: Mike Hance
- Exam Committee: Jason Nielsen, Bruce Schumm, Anyes Taffard
- SCIPP ATLAS postdocs and grad students
- SCIPP ITk Strips group
- Physics faculty, staff, and fellow grad students

Backup



gamma_n
Entries 65263
Mean 1.533
Std Dev 0.1805
Entries/Bin
500
400
300
200
100
0
0 1 2 3 4 5 6
gamma_n

Backup

- xAODAnaHelpers configuration file

```
from xAODAnaHelpers import Config
c = Config()

c.algorithm("BasicEventSelection", {
    "m_applyGRLCut": True,
    "m_GRLxml": "GoodRunsLists/data18_13TeV/20190708/data18_13TeV.periodAllYear_DetStatus-v105-pro22-
13_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml",
    "m_applyEventCleaningCut": True,
    "m_applyJetCleaningEventFlag": True,
    "m_applyTriggerCut": True,
    "m_triggerSelection": "HLT_xe110_pufit_xe(65|70)_L1XE50|HLT_xe120_pufit_L1XE50",
    "m_doPUreweighting": True,
    "m_autoconfigPRW": True,
    "m_lumiCalcFileNames": "GoodRunsLists/data18_13TeV/20190708/ilumicalc_histograms_None_348885-364292_OflLumi-
13TeV-010.root",
    "m_name": "myBasicEventSelection"
    })

c.algorithm("JetCalibrator", {
    "m_inContainerName": "AntiKt4EMTopoJets",
    "m_outContainerName": "AntiKt4EMTopoJetsCalibrated",
    "m_jetAlgo": "AntiKt4EMTopo",
    "m_name": "myJetCalibrator"
    })

c.algorithm("PhotonCalibrator", {
    "m_inContainerName": "Photons",
    "m_outContainerName": "PhotonsCalibrated",
    "m_name": "myPhotonCalibrator"
    })

c.algorithm("ElectronCalibrator", {
    "m_inContainerName": "Electrons",
    "m_outContainerName": "ElectronsCalibrated",
    "m_esModel": "es2016PRE",
    "m_decorrelationModel": "FULL_v1",
    "m_name": "myElectronCalibrator"
    })
```

Backup

- xAODAnaHelpers configuration file

```
c.algorithm("MuonCalibrator", {
    "m_inContainerName": "Muons",
    "m_outContainerName": "MuonsCalibrated",
    "m_name": "myMuonCalibrator"
    })

c.algorithm("TauCalibrator", {
    "m_inContainerName": "TauJets",
    "m_outContainerName": "TausCalibrated",
    "m_name": "myTauCalibrator"
    })

c.algorithm("PhotonSelector", {
    "m_inContainerName": "PhotonsCalibrated",
    "m_outContainerName": "PhotonsSelected",
    "m_vetoCrack": True,
    "m_doOQCut": True,
    "m_name": "myPhotonSelector"
    })

c.algorithm("TreeAlgo", {
    "m_jetContainerName": "AntiKt4EMTopoJetsCalibrated",
    # "m_jetContainerName": "AntiKt4EMTopoJetsSelected",
    "m_jetDetailStr": "kinematic rapidity energy scales",
    # "m_photonContainerName": "PhotonsCalibrated",
    "m_photonContainerName": "PhotonsSelected",
    "m_photonDetailStr": "kinematic isolation PID purity",
    "m_elContainerName": "ElectronsCalibrated",
    # "m_elContainerName": "ElectronsSelected",
    "m_elDetailStr": "kinematic isolation isolationKinematics PID",
    "m_muContainerName": "MuonsCalibrated",
    # "m_muContainerName": "MuonsSelected",
    "m_muDetailStr": "kinematic isolation isolationKinematics quality",
    "m_tauContainerName": "TausCalibrated",
    # "m_tauContainerName": "TausSelected",
    "m_tauDetailStr": "kinematic",
    "m_METReferenceContainerName": "MET_Reference_AntiKt4EMTopo",
    "m_METReferenceDetailStr": "metClus metTrk",
    "m_name": "myTreeAlgo"
    })
```