

Sensitivity studies using multivariate techniques for the search for fully hadronic decays of top squarks with the ATLAS detector

Nicolas Köhler

Max Planck Institute for Physics (Werner-Heisenberg-Institut)

Wednesday 21st March, 2018









The search for hadronically decaying top squarks

 $p = \frac{\tilde{t}}{\tilde{t}} + \frac{\tilde{t}}{W} + \frac{\tilde{t}_1^0}{\tilde{t}_1^0}$ $p = \frac{\tilde{t}}{\tilde{t}} + \frac{W}{W} + \frac{\tilde{t}_1^0}{\tilde{t}_1^0}$

- Introduced by P. Mogg (T 50.9)
- Simplified models with \tilde{t}_1 , $\tilde{\chi}_1^0$ (and eventually $\tilde{\chi}_1^{\pm}$)
- Cut&Count analysis with regions depending on decay kinematics
- Main backgrounds: Z+jets, tt
 , single top quark production





Increasing the sensitivity towards higher top squark masses

- Multivariate techniques of increasing popularity in HEP
- Supervised learning in order to classify signal and background events
- Sensitivity study done with J. Graw (MSc):
 - -- Is MVA really improving compared to Cut&Count?
 - -- Systematic study of machine learning algorithms / settings / inputs etc.
 - ightarrow BDT vs. Neural networks, ...
 - -- How to validate MVA outputs?
 - ightarrow Overtraining, systematics, ...
- Focus on $ilde{t}_1 o t + ilde{\chi}^0_1$ (m $_{ ilde{t}_1} \gg m_{ ilde{\chi}^0_1}$) only



Analysis •••••••



Analysis setup

Use preselection of Cut&Count analysis

- E_T^{miss}-trigger & E_T^{miss} > 250 GeV
- Lepton veto
- \geq 4 jets (p_T > 80,80,40,40 GeV)
- ≥ **1** *b*-jet
- QCD cleaning
- At least 2 reclustered R =1.2 jets (no additional requirement)
- + At least 1 reclustered R =0.8 jet (no additional requirement)



Variables available for training: All SR A+B variables



Training setup for
$$ilde{t}_1 o t + ilde{\chi}_1^0$$
 decays

- Train Boosted decision tree (BDT) with all variables used in Cut&Count analysis
- Which signal sample to use for training?



ightarrow For boosted region, using simulation with $m_{\tilde{t}} >$ 1000 GeV performs best





ightarrow Expected 3 σ line moves from 900GeV to 1TeV

Significance



BDT - settings

Performed trainings scanning through BDT setting parameters

 \rightarrow Look at area under ROC-curve to get measure for MVA performance



 \rightarrow Overtraining in left plot observed when comparing to testing (right) \rightarrow Best BDT parameters would be Maximal depth of 4,

minimal node size of 1%



BDT - settings

Training

Performed trainings scanning through BDT setting parameters

 \rightarrow Look at area under ROC-curve to get measure for MVA performance

Kolmogorov-Smirnov-Score



 \rightarrow Overtraining in left plot observed when comparing to testing (right) \rightarrow Best BDT parameters would be Maximal depth of 4, minimal node size of 1%



BDT - input variables

Train with all but one variable with the optimized BDT settings





BDT - results

Run with optimal BDT settings and skip unimportant variables



ightarrow No real improvement (9%) compared to cut based results

Summary O



Neural networks (NN) - input variables

- The same type of studies were done for neural networks
- Train with all but one variable with the optimized NN settings





ightarrow No real improvement (\sim 5%) compared to cut based results

Nicolas Köhler - MVA studies for fully hadronically decaying top squarks



Summary

- Sensitivity of search for top squarks using multivariate techniques performed by J. Graw
- Looked at BDT and neural networks
- Performed systematic scans on MVA settings and input variables
- Calculated expected exclusion limits with simplified control regions
- $\rightarrow~$ 10% / 5% improvement in expected sensitivity for BDT/NN compared to Cut&Count



BACKUP

03/21/2018

Nicolas Köhler - MVA studies for fully hadronically decaying top squarks



Discriminating variables for SR A+B

Signal Region		TT	TW	Т0
	$m_{\text{jet},R=1.2}^0$	> 120 GeV		
	$m_{\text{jet},R=1.2}^{1}$	$> 120 { m ~GeV}$	[60, 120] GeV	$< 60 { m ~GeV}$
	$m_{\mathrm{T}}^{b,\mathrm{min}}$	> 200 GeV		
	$N_{b-\text{jet}}$	≥ 2		
	τ -veto		yes	
	$\left \Delta\phi\left(\mathrm{jet}^{0,1,2},\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}}\right)\right $	> 0.4		
A	$m_{\text{jet},R=0.8}^{0}$	> 60 GeV		
	$\Delta R\left(b,b\right)$	> 1	-	
	$m_{\mathrm{T2}}^{\chi^2}$	$>400~{\rm GeV}$	$> 400 { m ~GeV}$	$> 500 { m ~GeV}$
	$E_{\mathrm{T}}^{\mathrm{miss}}$	> 400 GeV	$> 500 { m GeV}$	$> 550 { m ~GeV}$
В	$m_{\mathrm{T}}^{b,\mathrm{max}}$	> 200 GeV		
	$\Delta R\left(b,b\right)$	> 1.2		

JHEP 12 (2017) 085



Input distributions: signal normalized to SM integral





Input distributions: signal normalized to SM integral





Input distributions: signal normalized to SM integral





Input distributions: signal normalized to SM integral ATLAS Work in progress ATLAS Work in progress Entries / 50 GeV 13 TeV, 36.1 fb 800 7+iets 13 TeV, 36.1 fb Z+jets 7000 W+iets Single Top W+iets Sinale Top Diboson tt+W/Z Diboson tī+W/Z 6000 (m_,m_)=(1000,1) GeV (m, m)=(1000,1) Ge Data 5000 Data 5000 4000 4000 3000 2000 2000 1000 1000 Data/MC Data/MC 1.5 1 ! 0.5 0 0 100 200 300 400 500 600 900 0 200 800 1000 1200 1400 1600 700 800 1000 400 600 1800 2000 m^{b,min} [GeV] m_r^{b,max} [GeV] ATLAS Work in progress ATLAS Work in progress A14000 0912000 12000 10000 8000 6000 √s = 13 TeV, 36.1 fb⁻¹ Z+iets Entries / 100 Ge/ 13 TeV, 36,1 fb Z+iets 800 W+jets Single Top W+jets Single Top 7000 Diboson tī+W/7 Diboso tī+W/Z 6000 Data (m.m.)=(1000.1) Ge\ .m.)=(1000.1) Ge 5000 8000 4000 3000 4000 2000 2000 1000 n Data/MC Data/MC 1.5 1.50.5 100 200 300 400 500 600 700 800 900 200 400 600 800 1000 1200 1400 1600 1800 2000 1000 m-non-b,max

03/21/2018

- MVA studies for fully hadronically decaying top squarks Nicolas Köhler

IGeVI B/U 5



Input distributions: signal normalized to SM integral ATLAS Work in progress ATLAS Work in progress Entries / 25 GeV 21600 Z+jets 13 TeV, 36,1 fb 7+iets 13 TeV, 36.1 fb 8000 Ö14000 W+iets Single Top W+iets Single Top 7000 12000 tt+W/Z Diboson tt+W/7 Diboson 6000 (m_,m_)=(1000,1) Ge\ (m.,m.)=(1000,1) Ge\ Entries/ 0000 10000 10000 Data 5000 4000 6000 3000 4000 2000 1000 2000 Data/MC Data/MC 1.5 0.5 0 100 150 200 250 300 350 400 0 150 200 250 400 450 500 50 450 500 100 300 350 m⁰_{iet.R=1,2} [GeV] m_{iet R=1.2} [GeV] ATLAS Work in progress ATLAS Work in progress 2000 Entries / 20 GeV 13 TeV, 36.1 fb Z+iets g1800 13 TeV, 36.1 fb Z+iets 6000 Single Top W+jets Single Top W+jets 16000 14000 Diboson 5000 tī+W/7 Diboso tī+W/Z m.m.)=(1000.1) Ge\ Data (m .m .)=(1000.1) Ge\ ജ12000 4000 10000 Hiti 3000 6000 2000 4000 1000 2000 Data/MC Data/MC 1.5 1.3 0.5 100 50 100 200 300 350 400 150 400 m_{jet,R=0.8} [GeV] B/U 6

03/21/2018

MUA-Stubles for fully hadronically decaying top squarks Nicolas Köhler



Input distributions: signal normalized to SM integral ATLAS Work in progress ATLAS Work in progress GeV 012000 Vs = 13 TeV, 36.1 fb 13 TeV, 36.1 fb Z+jets 7+jets tŤ 7000 Entries / 75 G 000 8 0000 0000 0000 W+iets Single Top W+iets Sinale Top Entries / 75 6000 Diboson tt+W/Z tī+W/Z Diboson (m_,m_)=(1000,1) GeV (m_,m_)=(1000,1) Ge\ Data 5000 Data 4000 3000 4000 2000 2000 1000 Data/MC Data/MC 1.5 0.5 0 200 600 0 200 400 600 800 1200 0 400 1400 1400 m⁰₌ [GeV] p⁰_{T iii} [GeV] ATLAS Work in progress ATLAS Work in progress >35000 030000 Vs = 13 TeV, 36.1 fb⁻¹ 25000 13 TeV, 36.1 fb Z+iets Z+iets Single Top Single Top W+jets W+jets ₹ 25000 \$20000 Diboso Diboson tī+W/7 tf+W/Z Data (m.m.)=(1000.1) Ge\ Data (m .m .)=(1000.1) Ge\ Senter 15000 *8*20000 15000 10000 5000 5000 Data/MC Data/MC 1.5 0.5 0 100 700 100 200 600 700 800 300 400 800 400 р¹_{т,іі} [GeV] В/U 7

03/21/2018

m¹ [GeV] MVA studies for fully hadronically decaying top squarks Nicolas Köhler



BDT - overtraining

Train with all but one variable with the optimized BDT settings

Calulate Kolmogrov-Smirnov score comparing testing and training ROC-AUC



ightarrow The bigger the score, the more overtraining



Neural networks - overtraining

Train with all but one variable with the optimized NN settings

Calulate Kolmogrov-Smirnov score comparing testing and training ROC-AUC



ightarrow The bigger the score, the more overtraining