Suche nach $\tilde{t}$ in hadronische Endzustände mithilfe von multivariaten Methoden mit dem ATLAS-Detektor

Jonas Graw, betreut von Nicolas Köhler

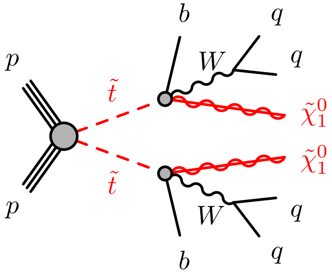
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

27. März 2017

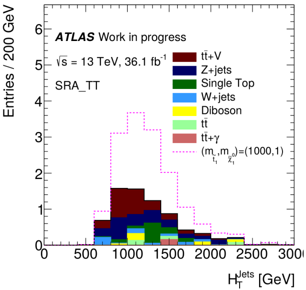
Bundesministerium
für Bildung
und Forschung

FSP 103
ATLAS

MAX-PLANCK-GESELLSCHAFT



$$m_{\tilde{t}} \gg m_{\tilde{\chi}_1^0}$$

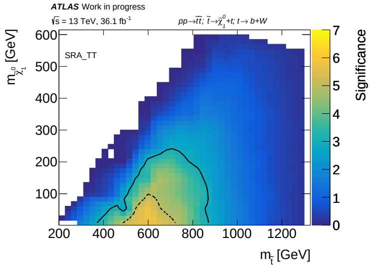


- Keine Leptonen
- min. 6 Jets
- min. 2 b-Jets

- Variablendefinitionen:
  - -- $H_T^{Jets} = \sum_{Jets} p_T$
  - -- $H_{T,sig} = \frac{E_T^{miss}}{\sqrt{H_T^{Jets}}}$

Standard schnittbasierte Methoden

- Verwendung von diskriminierenden Variablen um möglichst viel Untergrund zu eliminieren
- Schnitt auf diskriminierende Variablen um möglichst große Parameterräume mit hoher Signifikanz zu erreichen

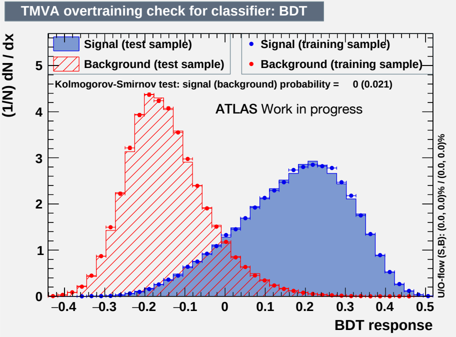


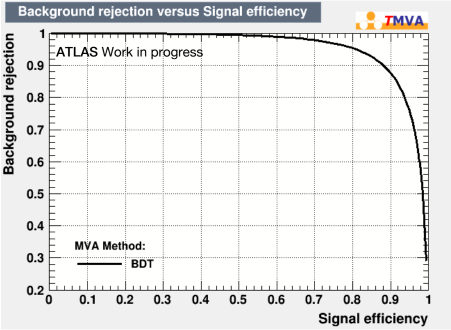
- Optimiert für $m_{\tilde{t}} = 1000$ GeV und $m_{\tilde{\chi}^0} = 1$ GeV
- Wichtige Variablen: m_{T2}, $m_T^{b,min}$, $m_T^{b,max}$, $m^0_{jet,R=1.2}$, $m^1_{jet,R=1.2}$, $m^0_{jet,R=0.8}$

- Wie kann man Parameterräume mit hoher Signifikanz in Zukunft vergrößern?
- Wie kann man Variablen mittels multivariater Methoden besser verwenden?

BDT

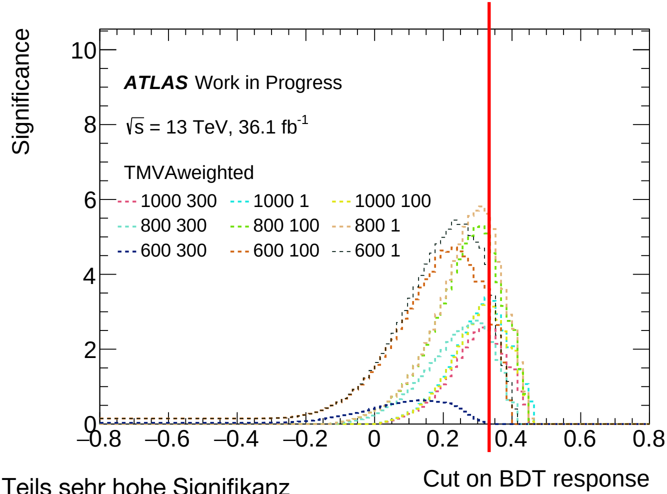
- Boosted Decision Tree (BDT) verwendet um Benutzung der Diskriminanzvariablen zu optimieren
- BDT response ist Indikator, wie wahrscheinlich Ereignis Signal bzw. Untergrund ist





- Test- und Trainingsdaten stimmen sehr gut überein
- Kein Übertraining

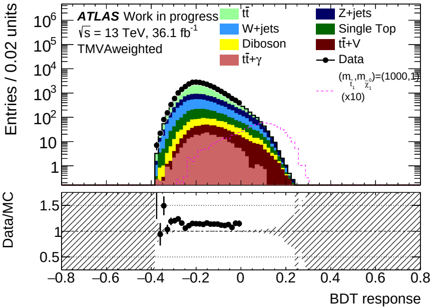
Schnitt auf BDT-respones[1]

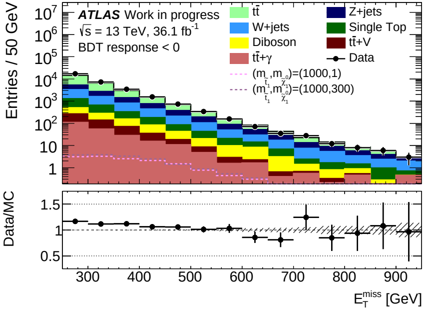


- Teils sehr hohe Signifikanz
- Wichtige Variablen: E_T^{miss}, m_{T2}, $p_T(top)$
- Schneiden bei BDT-response $\geq$ 0,34

[1]Training für Modell mit $m_{\tilde{t}} \geq 1000$ GeV

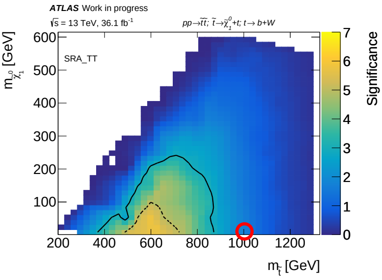
Vergleich zwischen Vorhersage MC und Daten



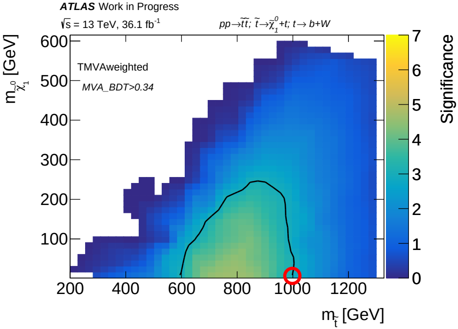


- Blinded für BDT response > 0
- Kleine Abweichung zwischen Daten und Vorhersagen sichtbar
- Gute Übereinstimmung der Daten mit der Untergrundvorhersage

Schnittbasierte Methode



BDT[2]



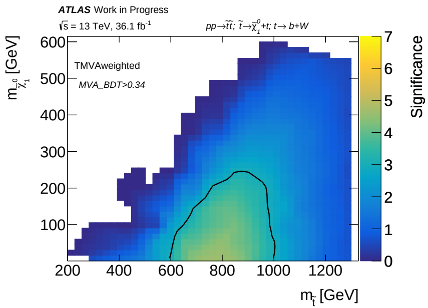
SRATT	BDT
$1,7\sigma$	$2,9\sigma$

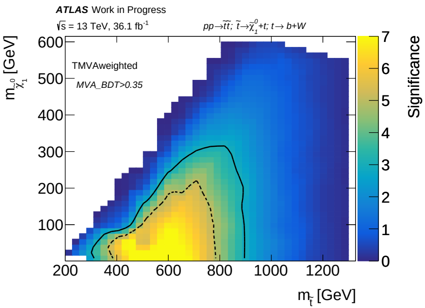
$\rightarrow$ Erwartete Signifikanz von 3σ bis $m_{\tilde{t}} = 1000$ GeV

[2] Training für Modell mit $m_{\tilde{t}} \geq 1$ TeV

Signifikanzen bei unterschiedlichen BDT-Trainings

- Training für Modell mit $m_{\tilde{t}} \geq 1000$ GeV
- Training mit allen $\tilde{t} \to t + \tilde{\chi}_1^0$ Modellen





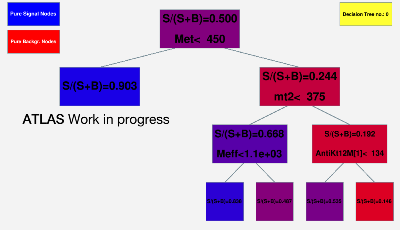
- Deutliche Vergrößerung der signifikanten Fläche, v.a. Richtung Diagonale
- Bei $m_{\tilde{t}}$ =600 GeV und $m_{\tilde{\chi}^0}$ =300 GeV erreicht Training mit allen Modellen Signifikanz von $0,8\sigma$
- Für große $m_{\tilde{t}}$ ist Training mit $m_{\tilde{t}} \geq 1000$ GeV vielversprechender

Zusammenfassung & Pläne

- Multivariate Optimierung von Signalregion für $\tilde{t} \rightarrow 0\ell$-Analyse
- BDT führt zu Erhöhung von Signifikanz
- Training mit Datensätzen mit $m_{\tilde{t}} \geq 1000$ GeV $\rightarrow 3\sigma$ bis 1000 GeV
- Training mit allen Datensätzen $\rightarrow 3\sigma$ für $m_{\tilde{t}} = 600$ GeV, $m_{\tilde{\chi}_1^0} = 250$ GeV
- Wie kann man die Phasenräume vergrößern?
- **Plan:**
  - -- Optimierung entlang der Diagonalen
  - -- Optimierung für hohe $\tilde{t}$-Massen
  - -- Verwendung von anderen multivariaten Analysemethoden

ANHANG

Boosted Decision Tree (BDT)



- Aufteilung Prozesse: Signal und Untergrund
- Abzweigentscheidung wird durch genau eine Diskriminanzvariable getroffen (Schnitt)
- gewählte Diskriminanzvariable gibt bestmög Untergrund an
- **Boosting:** Training eines neuen Baumes, wobei falsch klassifizierte Ereignisse stärker gewichtet werden
- Verwendung eines Entscheidungswaldes anstatt eines Entscheidungsbaums
- BDT response $r(i) \in [-1, 1]$ eines Ereignisses i:
  Klassifizierungsmaß abhänging von den Bäumen, wobei Grenzfälle:

$$\left\{ \begin{array}{l} r(i) = +1 \\ r(i) = -1 \end{array} \right\} \text{: Alle Bäume klassifizieren } i \text{ als} \left\{ \begin{array}{c} \text{Signal} \\ \text{Untergrund} \end{array} \right\}$$

BDT response

- Nach Training: Vergleiche wahren Wert der Testprobe y_i mit Vorhersage s_i. w_i: Wichtungen, wobei $\sum_i w_i = 1$
- Fehlerquotient: $e = \sum_i w_i 1_{s_i \neq y_i}$
- Boostfaktor $\alpha = \beta \cdot ln\left(\frac{1-e}{e}\right)$ mit β konstant, meist $\beta \in [0, 1]$
- Neue Gewichte: $w_i \to w_i \cdot exp\left(\alpha \cdot 1_{s_i \neq y_i}\right)$
- BDT response des Ereignisses i: $r_i = \frac{\sum_m \alpha_m \left(s_i\right)_m}{\sum_m \alpha_m}$, wobei

$$(s_i)_m = \begin{cases} 1 & \text{falls Tree } m \text{ Signal vorhergesagt} \\ -1 & \text{falls Tree } m \text{ Untergrund vorhergesagt} \end{cases}$$

Signifikanzberechnung

- $N(i)$: Anzahl Events mit $BDT - response \leq i$. Dann $N(i) = B(i) + S(i)$
- N_B: Die Anzahl der beobachteten Untergrund-Ereignisse. Dann: $p - value = p(N_B \geq N)$

Korrelationsmatrix