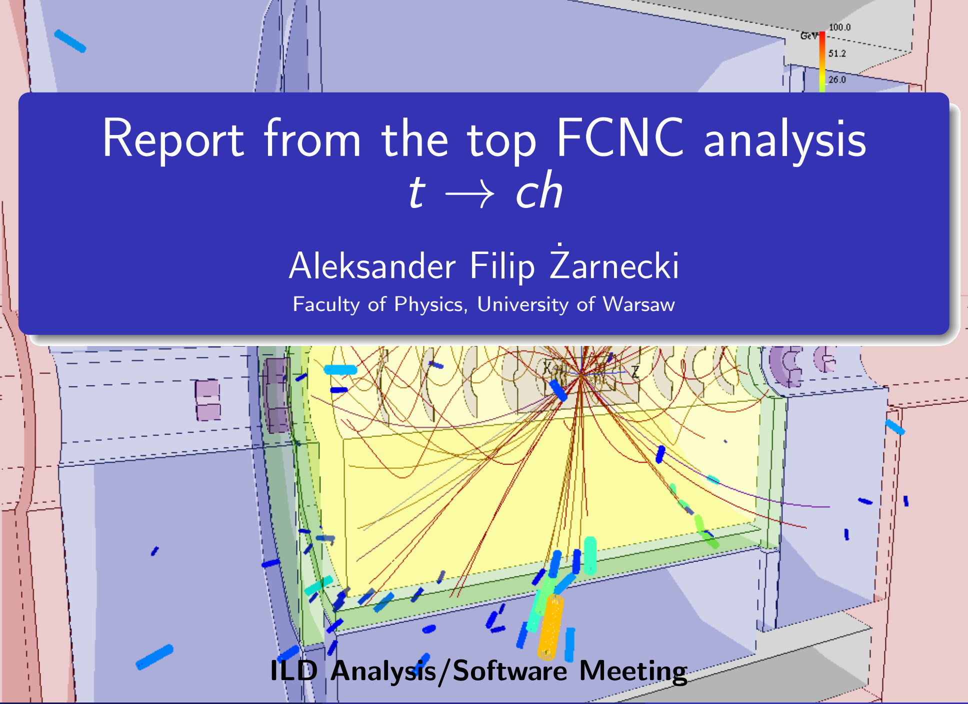
Report from the top FCNC analysis
$t \rightarrow ch$

Aleksander Filip Żarnecki

Faculty of Physics, University of Warsaw



ILD Analysis/Software Meeting

Outline

Motivation

In the Standard Model, FCNC top decays are strongly suppressed (CKM+GIM):

$$BR(t \rightarrow c\,\gamma) \sim 5 \cdot 10^{-14}$$
$$BR(t \rightarrow c\,Z) \sim 1 \cdot 10^{-14}$$
$$BR(t \rightarrow c\,g) \sim 5 \cdot 10^{-12}$$
$$BR(t \rightarrow c\,h) \sim 3 \cdot 10^{-15}$$

Any signal is a direct signature of "new physics"...

Motivation

In the Standard Model, FCNC top decays are strongly suppressed (CKM+GIM):

$$BR(t \to c\,\gamma) \sim 5 \cdot 10^{-14}$$
$$BR(t \to c\,Z) \sim 1 \cdot 10^{-14}$$
$$BR(t \to c\,g) \sim 5 \cdot 10^{-12}$$
$$BR(t \to c\,h) \sim 3 \cdot 10^{-15}$$

Any signal is a direct signature of "new physics"...

Decay $t \to c\ h$ is most interesting:

- well constrained kinematics
- test of Higgs boson couplings
- seems to be most difficult for LHC

Estimated HL-LHC reach:
(Snowmass 2013/ATLAS 2016)

$$BR(t \to qh) \sim 2 \cdot 10^{-4}$$

Two Higgs Doublet Model (2HDM) as a test scenario:

- one of simplest extensions of the SM
- $BR(t \to c\ h)$ up to 10^{-2} (tree level) and 10^{-4} (loop level)

Two Higgs Doublet Model (2HDM) type III used as a test scenario.
Implemented in SARAH $\Rightarrow$ WHIZARD 2 thanks to Florian Staub, many thanks also due to Juergen Reuter and Wolfgang Kilian...

WHIZARD 2.2.5 used to generate signal and background samples.

Test configuration of the model:

- $m_{h_1} = 125$ GeV
- $\mathrm{BR}(t \rightarrow c h_1) = 10^{-3}$
- $\mathrm{BR}(h \rightarrow b\bar{b}) = 100\%$

Generated samples:

- $e^+e^- \longrightarrow t\bar{t}$ (2HDM/SM)
- $e^+e^- \longrightarrow ch_1\bar{t},\ t\bar{c}h_1$ (2HDM)
- $e^+e^- \longrightarrow cb\bar{b}\bar{t},\ t\bar{c}b\bar{b}$ (SM)

Assume that main background to FCNC decays comes from standard decay channels, including $t \rightarrow bW^+$ followed by $W^+ \rightarrow c\bar{b}$

All events generated with CIRCE1 spectra + ISR. No polarization.
Only t, W and h defined to be unstable. No hadronization/decays.
No generator-level cuts imposed.

Very simplified detector description

- detector acceptance for leptons: $|\cos\theta_l| < 0.995$
- detector acceptance for jets: $|\cos\theta_j| < 0.975$
- jet energy smearing:

$$\sigma_E = \begin{cases} \frac{S}{\sqrt{E}} & \text{for} \quad E < 100 GeV \\ \frac{S}{\sqrt{100\ GeV}} & E > 100 GeV \end{cases}$$

with S = 30%, 50% and 80% [GeV$^{1/2}$]

- b tagging (misstagging) efficiencies: (as expected for LCFI+)

Scenario	b	c	uds
Ideal	100%	0%	0%
A	90%	30%	4%
B	80%	8%	0.8%
C	70%	2%	0.2%
D	60%	0.4%	0.08%

Event selection: $t\bar{t}$ final state

"Signal" top: $t \to ch_1$ + higgs decay to $b\bar{b} \Rightarrow$ 2 b tags
"Spectator" top: SM top decay $\Rightarrow$ 1 b tag

Considered final states (resulting from $W^{\pm}$ decay channels):

- semileptonic: 4 jets + lepton + missing p_t
- fully hadronic: 6 jets, no leptons, no missing p_t

Parton Level study

Event selection: $t\bar{t}$ final state

"Signal" top: $t \to ch_1$ + higgs decay to $b\bar{b} \Rightarrow 2\ b$ tags
"Spectator" top: SM top decay $\Rightarrow 1\ b$ tag

Considered final states (resulting from $W^{\pm}$ decay channels):

- semileptonic: 4 jets + lepton + missing p_t
- fully hadronic: 6 jets, no leptons, no missing p_t

Event selection cuts for $\sqrt{s} = 500$ GeV, $50\%/\sqrt{E}$ jet energy resolution

Semileptonic:

- Missing $p_t > 25$ GeV
- Single lepton with $p_t > 15$ GeV
- 4 jets with $p_t > 15$ GeV
- 3 jets b-tagged

Fully hadronic:

- Missing $p_t < 15$ GeV
- No lepton with $p_t > 10$ GeV
- 6 jets with $p_t > 15$ GeV
- 3 jets b-tagged

Signal selection

After pre-selection cuts, compare two hypothesis:

- signal hypothesis   hadronic final state

$$\chi^2_{sig} = \left(\frac{M_{bqq} - m_t}{\sigma_t}\right)^2 + \left(\frac{M_{qq} - m_W}{\sigma_W}\right)^2 + \left(\frac{M_{bbq} - m_t}{\sigma_t}\right)^2 + \left(\frac{M_{bb} - m_h}{\sigma_h}\right)^2$$

- background hypothesis ($t\bar{t}$ hadronic decays)

$$\chi^2_{bg} = \left(\frac{M_{bqq} - m_t}{\sigma_t}\right)^2 + \left(\frac{M_{qq} - m_W}{\sigma_W}\right)^2 + \left(\frac{M_{bbq} - m_t}{\sigma_t}\right)^2 + \left(\frac{M_{bq} - m_W}{\sigma_W}\right)^2$$

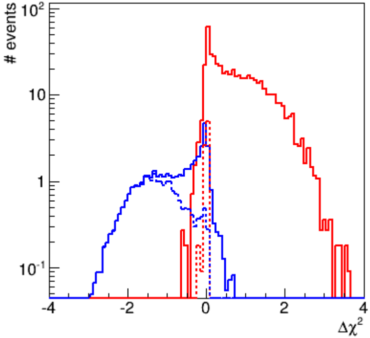
Independent search for best background and signal combinations
Difference in the last term only: h vs W mass discrimination crucial!

Signal selection

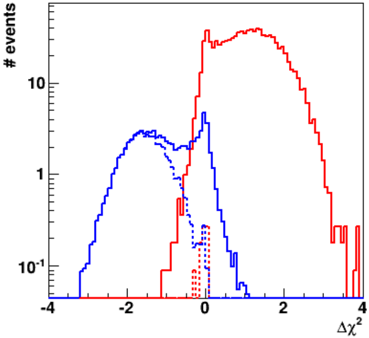
Difference of $\log_{10} \chi^2$ for two hypothesis, for signal and background events

Before (solid) and after (dashed) other selection cuts

Semi-leptonic events



Fully hadronic events



500 GeV, jet energy resolution 30%, 70% *b*-tagging efficiency

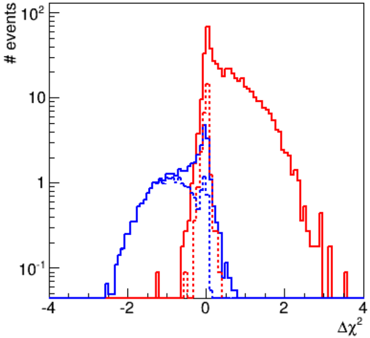
Background rejection strongly depends on the detector performance

Parton Level study

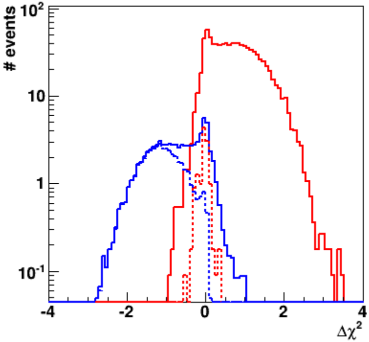
Signal selection

Difference of $\log_{10}\chi^2$ for two hypothesis, for signal and background events
Before (solid) and after (dashed) other selection cuts

Semi-leptonic events



Fully hadronic events



500 GeV, jet energy resolution 50%, 70% *b*-tagging efficiency
Background rejection strongly depends on the detector performance

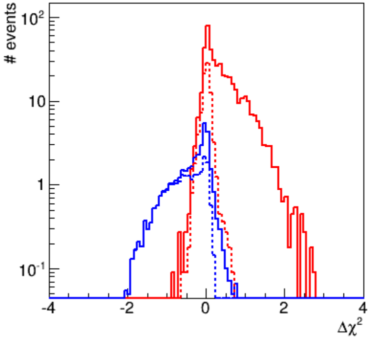
Parton Level study

Signal selection

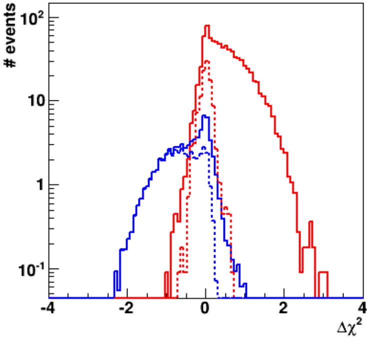
Difference of $\log_{10}\chi^2$ for two hypothesis, for signal and background events

Before (solid) and after (dashed) other selection cuts

Semi-leptonic events



Fully hadronic events



500 GeV, jet energy resolution 80%, 70% *b*-tagging efficiency

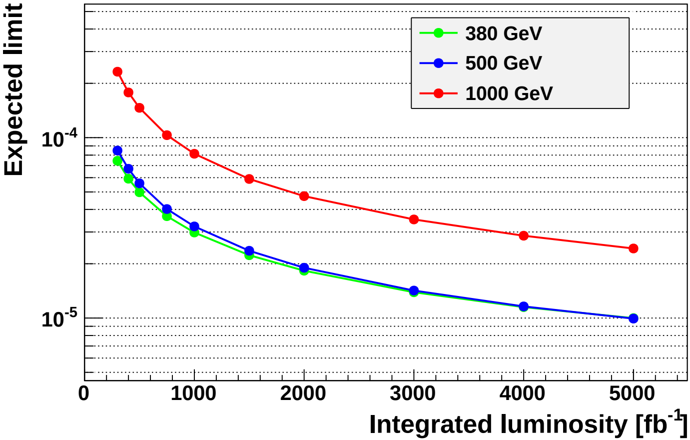
Background rejection strongly depends on the detector performance

Parton level study presented at TopLC'2015 [arXiv:1604.08122]

Promising results on the feasibility of the measurement

Estimated limits on $BR(t \rightarrow ch) \times BR(h \rightarrow b\bar{b})$

Assumed jet energy resolution $\sigma_E = 50\%/\sqrt{E}$ (5% above 100 GeV)

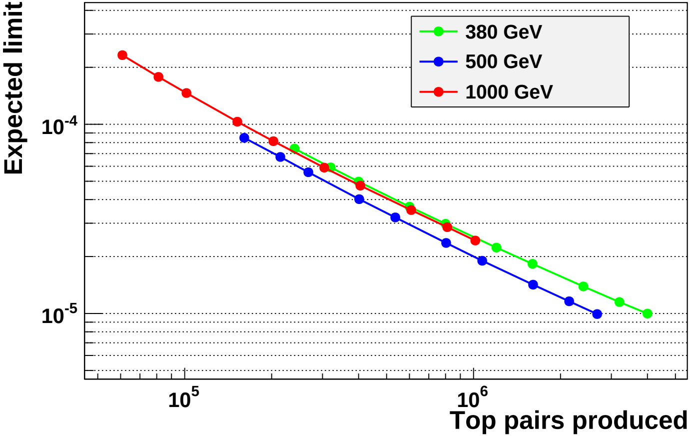


Parton level study presented at TopLC'2015 [arXiv:1604.08122]

Promising results on the feasibility of the measurement

Estimated limits on $BR(t \to ch) \times BR(h \to b\bar{b})$

Assumed jet energy resolution $\sigma_E = 50\%/\sqrt{E}$ (5% above 100 GeV)



Full simulation for CLIC @ 380 GeV

Dedicated samples generated with WHIZARD 2.2.8
Signal: SARAH implementation of 2HDM(III), BR$(t \to ch_1) = 10^{-3}$

Beam spectra for CLIC taken from file (350 GeV scaled to 380 GeV)
Beam polarization of -80%/0% (for e^-/e^+)

Hadronization done in PYTHIA 6.427
quark masses and PYTHIA settings adjusted to CLIC CDR
Standard event processing with CLIC_ILD_CDR500 configuration

Samples considered in the study

- dedicated FCNC signal sample $e^+e^- \longrightarrow ch_1\bar{t},\ t\bar{c}h_1$
- test sample of SM background $e^+e^- \longrightarrow t\bar{t}$ for simulation validation
- full 6-fermion sample as produced for CLIC $t\bar{t}$ studies

Signal and background samples normalised to 500 fb^{-1}
Assumed $t\bar{t}$ cross section at 380 GeV: 820 fb

Event processing

DST files processed with MARLIN, ilcsoft v01-17-09 (ilcDIRAC)

- Using LooseSelectedPandoraPFANewPFOs as input collection
- LCFI+ primary and secondary vertex finder
- LCFI+ jet finding with Valencia algorithm
- LCFI+ vertex corrections and flavour taging
  default weights used (no tuning), but seem to work OK
- root TTree writing

Event processing

DST files processed with MARLIN, ilcsoft v01-17-09 (ilcDIRAC)

- Using LooseSelectedPandoraPFANewPFOs as input collection
- LCFI+ primary and secondary vertex finder
- LCFI+ jet finding with Valencia algorithm
- LCFI+ vertex corrections and flavour taging
  default weights used (no tuning), but seem to work OK
- root TTree writing

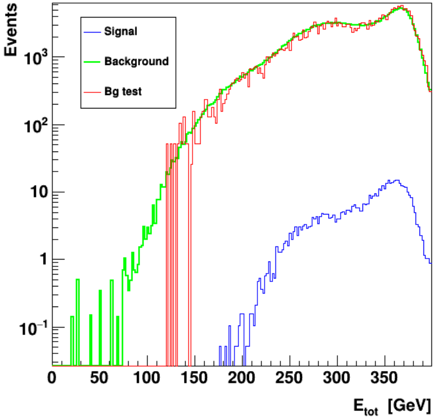
Final analysis in root:

- hadronic decay selection
- pre-selection cuts (loose cuts on flavour tagging)
- kinematic fit
- final selection (cuts or BDT)
  optimised for best BR limit

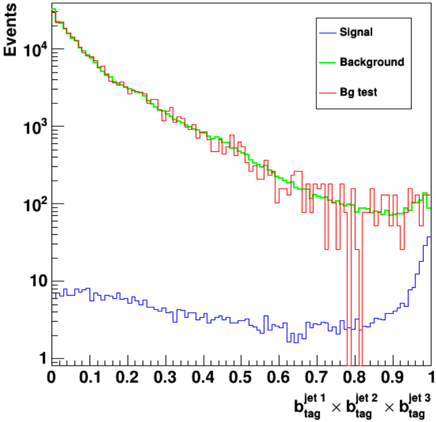
Control plots

Comparing signal sample with full background and test samples.

Total measured energy



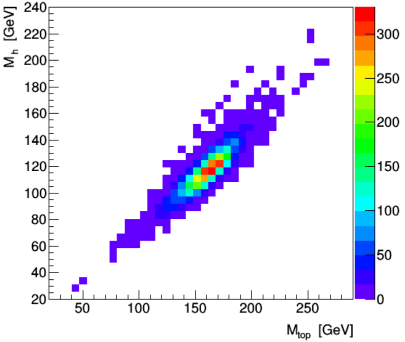
Product of three highest *b*-tag value



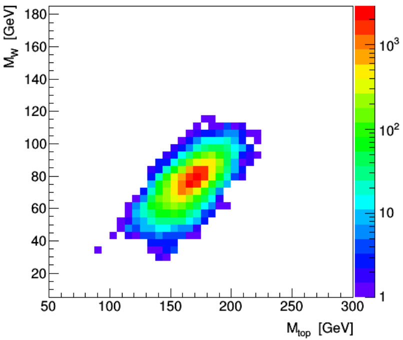
Mass correlation

Significant correlations observed between reconstructed masses of top (3 jets) and its decay product (2 jets)

Higgs and top (signal)



W boson and top (background)



$\Rightarrow$ should be taken into account in event selection

New χ^2 definition

Using mass ratios to reduce influence of mass correlations:

- signal hypothesis

use also top boost as additional constrain

$$\chi^2_{sig} = \left(\frac{M_{bqq} - m_t}{\sigma_t}\right)^2 + \left(\frac{M_{bbc} - m_t}{\sigma_t}\right)^2 + \left(\frac{\frac{E_{bqq}}{M_{bqq}} - \gamma_t}{\sigma_\gamma}\right)^2 + \left(\frac{\frac{E_{bbc}}{M_{bbc}} - \gamma_t}{\sigma_\gamma}\right)^2 + \left(\frac{\frac{M_{qq}}{M_{bqq}} - \frac{m_W}{m_t}}{\sigma_{R_W}}\right)^2 + \left(\frac{\frac{M_{bb}}{M_{bbc}} - \frac{m_h}{m_t}}{\sigma_{R_h}}\right)^2$$

- similar for background hypothesis ($t\bar{t}$ hadronic decays)

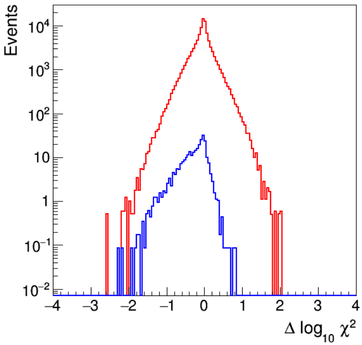
$$\chi^2_{bg} = \quad \ldots \quad + \left(\frac{\frac{M_{qq}}{M_{bqq}} - \frac{m_W}{m_t}}{\sigma_{R_W}}\right)^2 + \left(\frac{\frac{M_{bq}}{M_{bqq}} - \frac{m_W}{m_t}}{\sigma_{R_W}}\right)^2$$

Signal-background discrimination

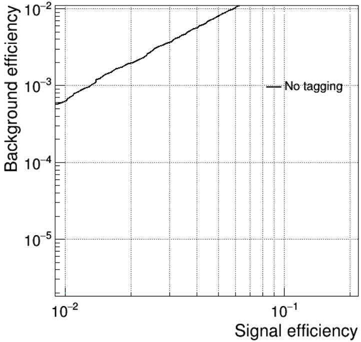
Based on the cut on the difference of $\log_{10}\chi^2$ for two hypothesis

Events with "good" fit of signal hypothesis ($\chi^2_{sig} < 14$, $|\Delta M_{top}| < 45$ GeV)

$\Delta \log_{10}\chi^2$ distribution for signal and background



Background vs signal efficiency after subsequent cuts



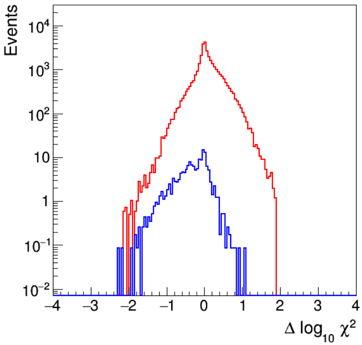
normalized to all decay channels

Signal-background discrimination

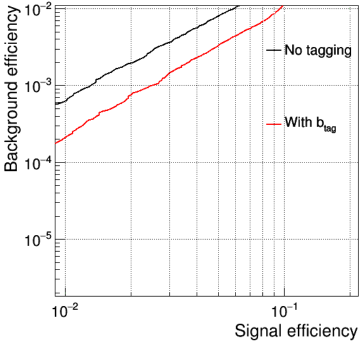
Based on the cut on the difference of $\log_{10}\chi^2$ for two hypothesis

Events with "good" fit of signal hypothesis ($\chi^2_{sig} < 14$, $|\Delta M_{top}| < 45$ GeV)

$\Delta \log_{10}\chi^2$ distribution for signal and background



Background vs signal efficiency after subsequent cuts



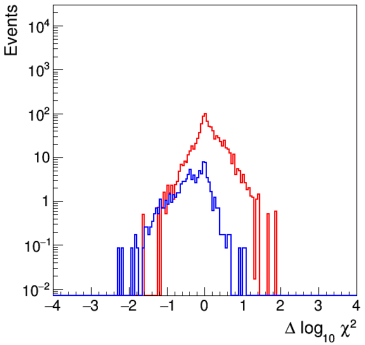
normalized to all decay channels

Signal-background discrimination

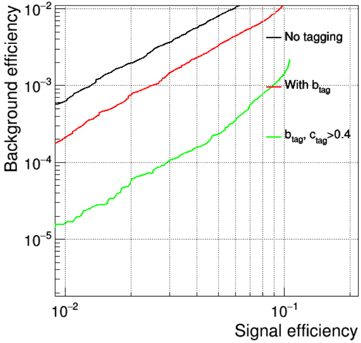
Based on the cut on the difference of $\log_{10}\chi^2$ for two hypothesis

Events with "good" fit of signal hypothesis ($\chi^2_{sig} < 14$, $|\Delta M_{top}| < 45$ GeV)

$\Delta \log_{10}\chi^2$ distribution for signal and background



Background vs signal efficiency after subsequent cuts



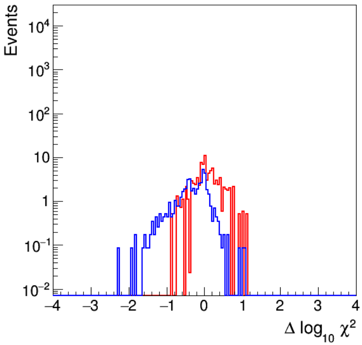
normalized to all decay channels

Signal-background discrimination

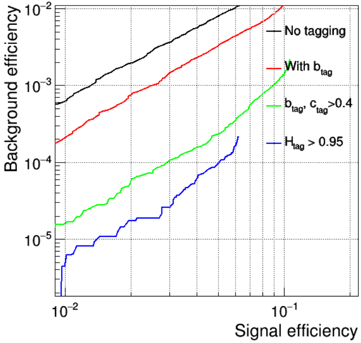
Based on the cut on the difference of $\log_{10}\chi^2$ for two hypothesis

Events with "good" fit of signal hypothesis ($\chi^2_{sig} < 14$, $|\Delta M_{top}| < 45$ GeV)

$\Delta \log_{10}\chi^2$ distribution for signal and background



Background vs signal efficiency after subsequent cuts



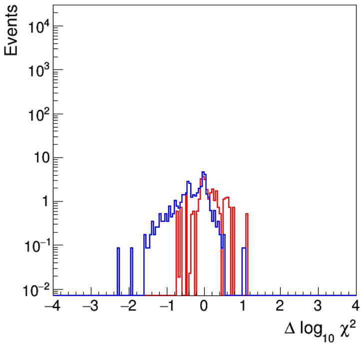
normalized to all decay channels

Signal-background discrimination

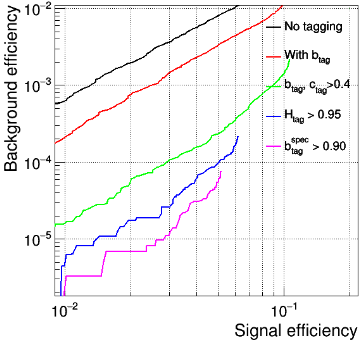
Based on the cut on the difference of $\log_{10}\chi^2$ for two hypothesis

Events with "good" fit of signal hypothesis ($\chi^2_{sig} < 14$, $|\Delta M_{top}| < 45$ GeV)

$\Delta \log_{10}\chi^2$ distribution for signal and background



Background vs signal efficiency after subsequent cuts



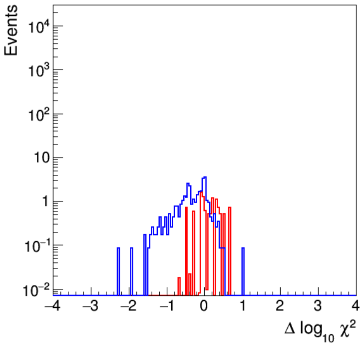
normalized to all decay channels

Signal-background discrimination

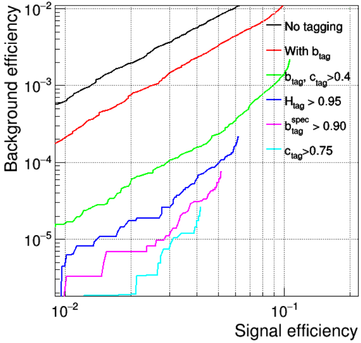
Based on the cut on the difference of $\log_{10}\chi^2$ for two hypothesis

Events with "good" fit of signal hypothesis ($\chi^2_{sig} < 14$, $|\Delta M_{top}| < 45$ GeV)

$\Delta \log_{10}\chi^2$ distribution for signal and background



Background vs signal efficiency after subsequent cuts



normalized to all decay channels

Expected events in six-jet final state

For 500 fb^{-1}, assuming $BR(t \to ch) \times BR(h \to b\bar{b}) = 10^{-3}$ for signal

Analysis level / Selection cut	Expected events		Efficiency	
	$t\bar{t}$ (SM)	Signal	$t\bar{t}$ (SM)	Signal
All events	410'000	819	100%	100%
hadronic events	170'000	543	41.5%	66.3%
Before kinematic fit				
$E_{balance} < 100$ GeV	167'000	499	40.6%	60.9%
3 b jets tagged ($b_{tag} > 0.4$)	13'280	300	3.24%	36.6%
c jet tagged ($b_{tag}+c_{tag} > 0.4$)	9640	276	2.35%	33.8%
After kinematic fit				
Good fit ($\chi^2_{sig}<14$, $\Delta M_t<45$ GeV)	894	87	0.22%	10.7%
b-tag for higgs jets ($b_1 \times b_2 > 0.95$)	89.5	50.8	0.022%	6.2%
b and c tags ($b_3>0.9$, $c_4+b_4>0.75$)	10.7	34.1	$2.6 \cdot 10^{-5}$	4.2%
$\chi^2_{sig}/\chi^2_{bg} < 1.38$ (optimised for limit)	4.89	31.8	$1.2 \cdot 10^{-5}$	3.9%

Expected limits only hadronic channel considered !

Cuts were optimised for the best expected BR limit.

Final signal selection efficiency: 3.9% (5.9% of hadronic decays)
Background suppression: $1.2 \cdot 10^{-5}$

Expected 95% C.L. limit for 500 fb^{-1} at 380 GeV preliminary

$$BR(t \to ch) \times BR(h \to b\bar{b}) \quad < \quad 2.6 \cdot 10^{-4}$$

With luminosity of 1000 fb^{-1} at 380 GeV

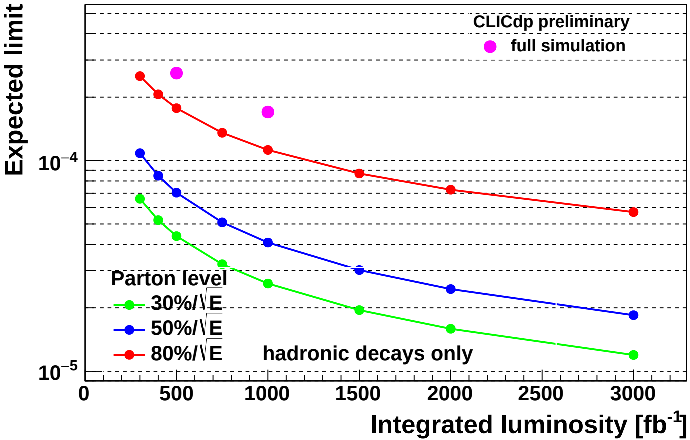
$$BR(t \to ch) \times BR(h \to b\bar{b}) \quad < \quad 1.7 \cdot 10^{-4}$$

assuming $t\bar{t}$ cross section at 380 GeV of 820 fb

see: http://hep.fuw.edu.pl/u/zarnecki/talks/afz_lcws2016.pdf

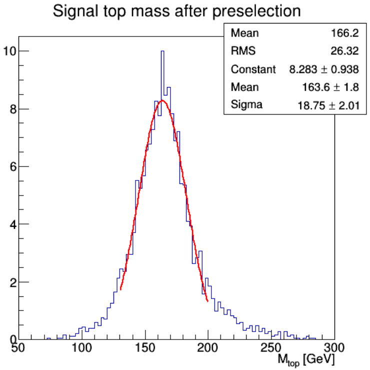
Expected limits on $BR(t \to ch) \times BR(h \to b\bar{b})$

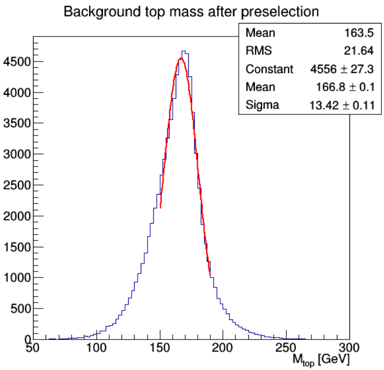
Comparison with parton level results, different jet energy resolutions



Kinematic fit

The main reason for weak limit is poor performance of the kinematic fit.





Mass resolution much worse than expected.

Signal reconstruction much worse than for background events...

Jet matching

To understand top reconstruction better, event kinematics was compared between three levels (hadronic final state considered):

- parton level: six fermion final state (as generated by WHIZARD)
- particle level: result of PYTHIA hadronisation
  MCParticles clustered in six jets (Valencia algorithm)
- jet level: six jet final state, as reconstructed after detector simulation
  LCFIPlus clustering with Valencia algorithm

Jet matching

To understand top reconstruction better, event kinematics was compared between three levels (hadronic final state considered):

- parton level: six fermion final state (as generated by WHIZARD)
- particle level: result of PYTHIA hadronisation
  MCParticles clustered in six jets (Valencia algorithm)
- jet level: six jet final state, as reconstructed after detector simulation
  LCFIPlus clustering with Valencia algorithm

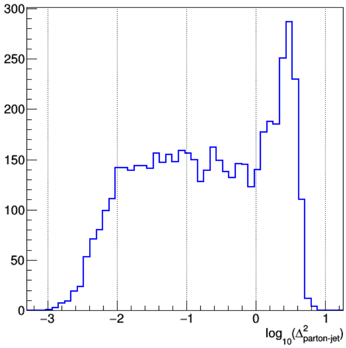
To quantify the level of correspondence between different levels:

$$\Delta^2_{\text{parton-jet}} = \min_{\text{all combinations}} \sum_{partons, jets} \left[\sphericalangle(\vec{p}_{jet}, \vec{p}_{parton})\right]^2$$

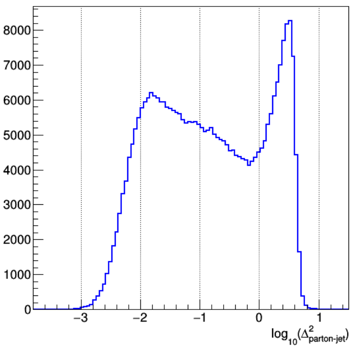
Jet matching

Distance between parton level and detector level jets

Signal events



Background ($t\bar{t}$) events

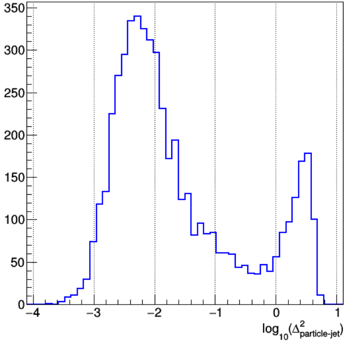


For significant fraction of events reconstructed detector-level jets have nothing to do with the generated fermion configuration!

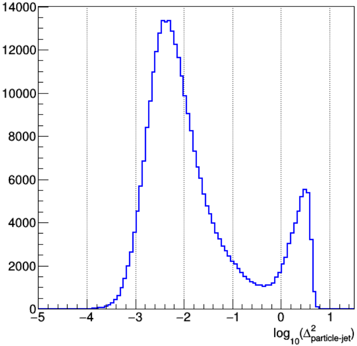
Jet matching

Distance between particle level jets and detector level jets

Signal events



Background ($t\bar{t}$) events

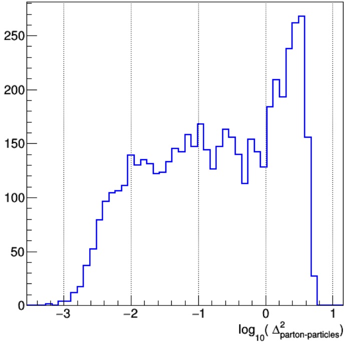


For most events reconstructed detector-level jets follow closely the particle level configuration...

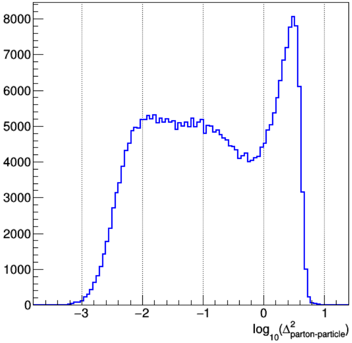
Jet matching

Distance between parton level and particle level jets (no detector involved)

Signal events



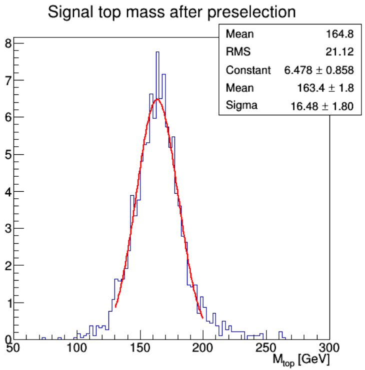
Background ($t\bar{t}$) events

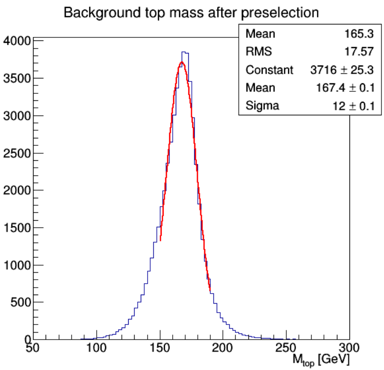


In most cases, information about the partonic final state
is already lost on particle level!

Mass resolution

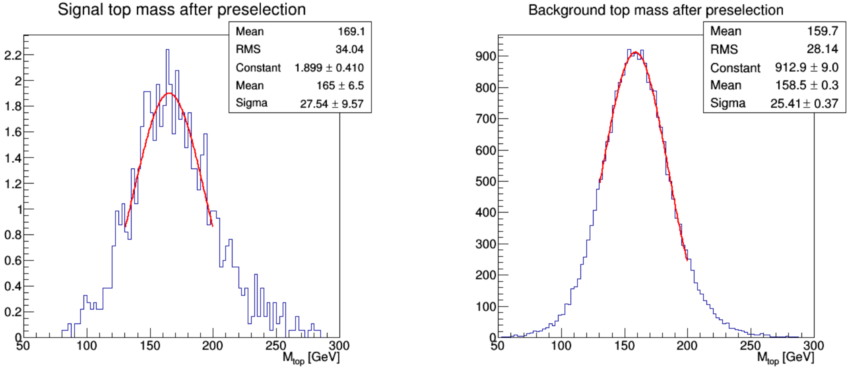
Reconstructed top candidate mass for events with good matching between parton level and detector jet level ($\Delta^2_{\text{parton-jet}} < 0.6$)





Mass resolution

Reconstructed top candidate mass for events with bad matching between parton level and detector jet level ($\Delta^2_{\text{parton-jet}} > 0.6$)

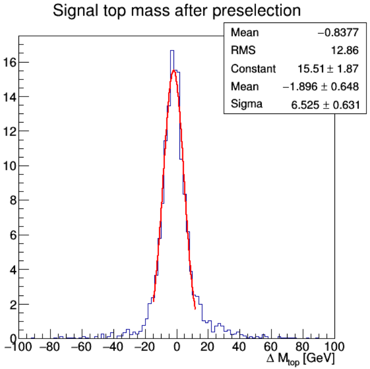


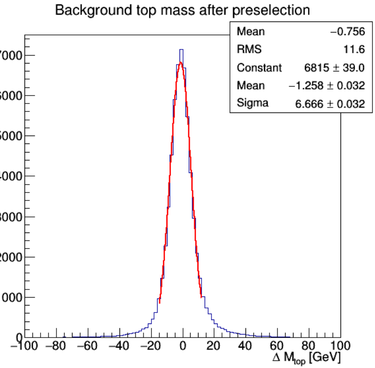
⇒ "unmatched" events give long tails in mass distributions significantly reduce efficiency of background rejection with kinematic fit

Mass reconstruction

Mass resolution

Difference between top candidate mass reconstructed on particle level and detector level (for events with good matching)





$\Rightarrow$ very good detector performance confirmed

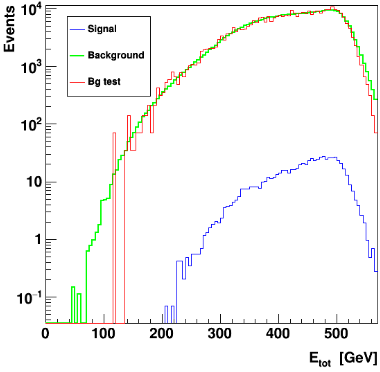
problem is most likely due to particle migrations between jets...

Dedicated signal and background samples generated and processed.
Full 6-fermion sample (negative polarisation) processed

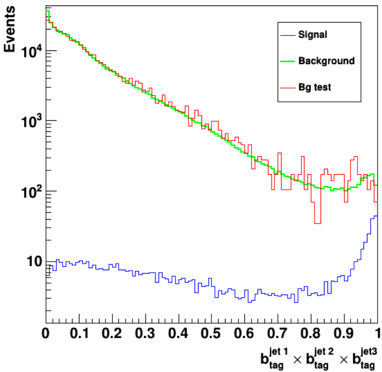
Control plots

Comparing signal sample with full background and test samples.

Total measured energy



Product of three highest *b*-tag value

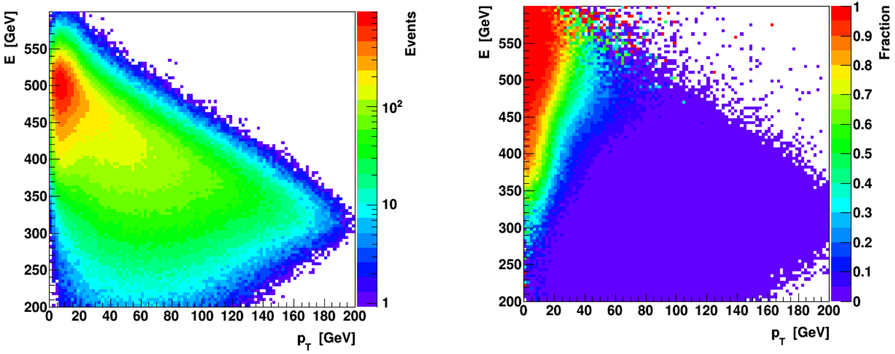


Event pre-selection

Discrimination between hadronic and (semi-)leptonic events by looking at the correlation of transverse momentum and total energy

Background event distribution

Hadronic event fraction



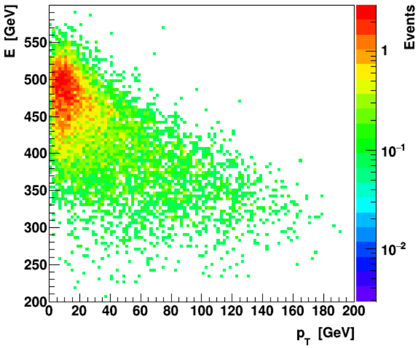
Pre-selection criteria still to be optimised...

First look at $\sqrt{s} = 500$ GeV

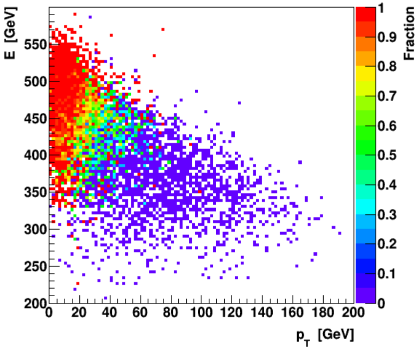
Event pre-selection

Discrimination between hadronic and (semi-)leptonic events by looking at the correlation of transverse momentum and total energy

Signal event distribution



Hadronic event fraction

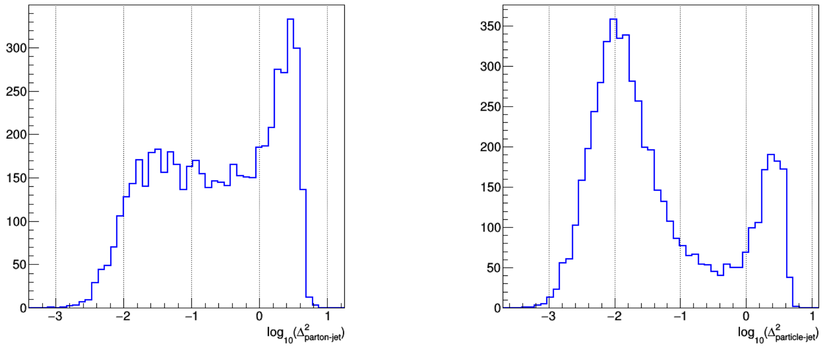


Pre-selection criteria still to be optimised...

First look at $\sqrt{s} = 500$ GeV

Jet matching

Distance between parton level (left) and particle level (right), and detector level jets. Signal events.



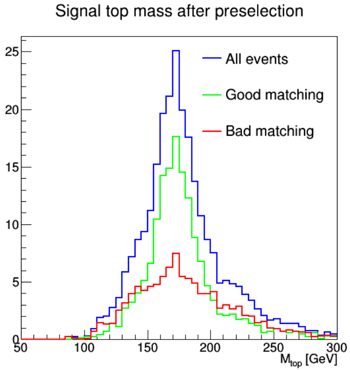
Very similar to 380 GeV case...

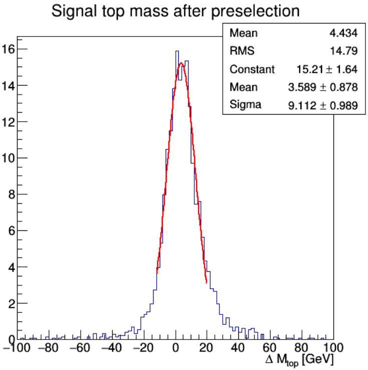
Information about the partonic final state already lost on particle level

First look at $\sqrt{s} = 500$ GeV

Jet matching

Reconstructed top candidate mass (left) and the difference between detector level and particle level reconstruction (right)



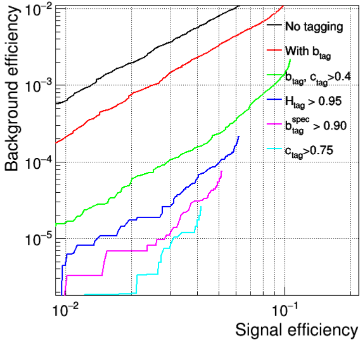


Very similar to 380 GeV case...

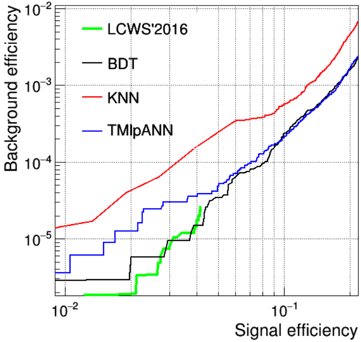
Background vs signal efficiency after the final selection cut

normalized to all decay channels

LCWS'16: cut on $\chi^2_{\rm sig}/\chi^2_{\rm bg}$



MVA: cut on the classifier response



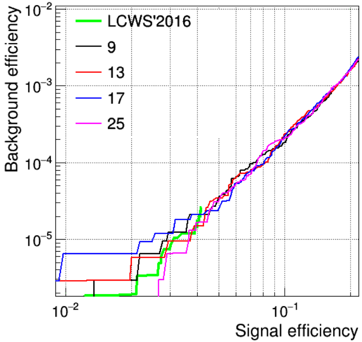
$\Rightarrow$ BDT gives best selection
(compared to other MVA algorithms)

New: using BDT for final selection

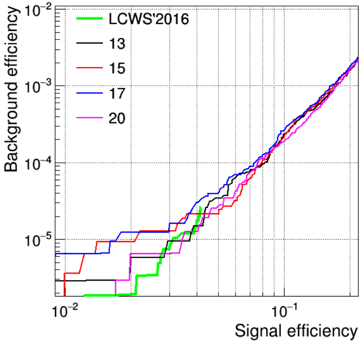
Background vs signal efficiency after the cut on BDT response
For different numbers (and choices) of variables used in BDT

number of variables indicated

Final state variables only



Including global event variables



$\Rightarrow$ best BDT result similar to LCWS'2016 (cut based approach)

Still trying to get improvement...

Conclusions

FCNC top decays $t \to ch$

Preliminary results for 380 GeV presented at LCWS'2016.

Focus on optimizing kinematic reconstruction in the hadronic channel

Expected limit at 500 fb^{-1}

$$BR \quad < \quad 2.6 \cdot 10^{-4}$$

Conclusions

FCNC top decays $t \to ch$

Preliminary results for 380 GeV presented at LCWS'2016.

Focus on optimizing kinematic reconstruction in the hadronic channel

Expected limit at 500 fb^{-1}

$$BR \quad < \quad 2.6 \cdot 10^{-4}$$

First look at 500 GeV, expected limit at 500 fb^{-1}

$$BR \quad < \quad 4.6 \cdot 10^{-4}$$

Conclusions

FCNC top decays $t \to ch$

Preliminary results for 380 GeV presented at LCWS'2016.

Focus on optimizing kinematic reconstruction in the hadronic channel

Expected limit at 500 fb^{-1}

$$BR \quad < \quad 2.6 \cdot 10^{-4}$$

First look at 500 GeV, expected limit at 500 fb^{-1}

$$BR \quad < \quad 4.6 \cdot 10^{-4}$$

Kinematic fit performance much poorer than expected from parton level

Background reduction primarily based on flavour tagging!

Conclusions

FCNC top decays $t \to ch$

Preliminary results for 380 GeV presented at LCWS'2016.

Focus on optimizing kinematic reconstruction in the hadronic channel

Expected limit at 500 fb^{-1}

$$BR \quad < \quad 2.6 \cdot 10^{-4}$$

First look at 500 GeV, expected limit at 500 fb^{-1}

$$BR \quad < \quad 4.6 \cdot 10^{-4}$$

Kinematic fit performance much poorer than expected from parton level

Background reduction primarily based on flavour tagging!

Analysis ongoing with main focus on:

- trying to understand mass reconstruction
- optimising final event selection with BDT
- semi-leptonic channel still waiting...

Thank you!