NLO simulation of $t\bar{t}H(b\bar{b})$ background

work in collaboration with
Fabio Cascioli, Philipp Maierhöfer, Stefano Pozzorini, Frank Siegert

Niccoló Moretti
University of Zurich

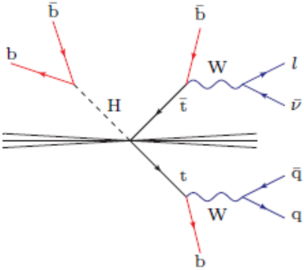
Zurich, September 11, 2014

$\mathrm{t\bar{t}H(b\bar{b})}$ analysis: status and importance of precise theory simulations for signal and background

Status

- $\mathrm{t\bar{t}H(b\bar{b})}$ not observed
- $\mathrm{t\bar{t}}$ + *jets* simulation still based on LO MC tools
- related theory uncertainty is one of the main bottlenecks

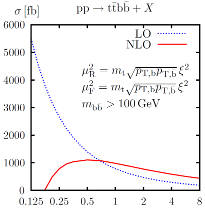
95% exclusion in $\sigma^{\mathrm{SM}}_{\mathrm{t\bar{t}}H}$ units	$H \to \mathrm{b\bar{b}}$	$H \to VV^*$	$H \to \gamma\gamma$
ATLAS	4.1 (2.6)		6.7 (4.9)
CMS	3.3 (2.9)	6.6 (2.4)	7.4 (5.7)



- experimental challenge: highly involved $b\bar{b}b\bar{b}l\nu jj$ signature
- theoretical challenge: NLO corrections to multi-particle backgrounds $pp \to t\bar{t}jj, t\bar{t}c\bar{c}, t\bar{t}b\bar{b}$

Theoretical predictions for ttbb and ttjj: status and needed improvements

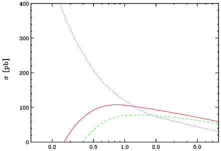
- available: fixed-order NLO for ttbb and tt+jj (at 14 *TeV*)



$\mathrm{t\bar{t}b\bar{b}}$

scale uncertainty 80% → 20%

[A. Bredenstein,A. Denner,S. Dittmaier and S.P.,JHEP 03(2010)021]



$\mathrm{t\bar{t}}jj$

scale uncertainty 70% → 15%

[G. Bevilacqua,M. Czakon,C.G. Papadopoulos,and M. Worek,Phys.Rev.Lett. 104(2010)162002]

- parton-level results not applicable to experimental analysis
  → NLO+PS matching crucial in order to exploit strong reduction of uncertainty

Theoretical predictions for ttbb and ttjj: status and needed improvements

Final goals

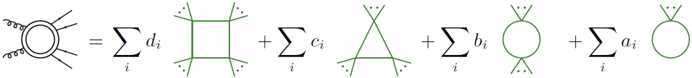
- NLO+PS simulations with different (N_j, N_b, N_c) multiplicity
- consistent merging in a single simulation

! NLO corrections to multi-particle FS processes needed

- $gg \to \mathrm{t\bar{t}b\bar{b}}$ $\mathrm{O}(10^3)$ 1-loop diagrams
- $gg \to \mathrm{t\bar{t}}gg$ $\mathrm{O}(10^4)$ 1-loop diagrams

- ⋆ inconceivable until few years ago, now feasible thanks to new automatic NLO tools
- ⋆ but NLO matching and merging still significantly challenging

Simulation Tools: Sherpa + OpenLoops



OpenLoops [Cascioli, Maierhöfer, S.P., PRL 108(2012)111601]

- fully automated loop-amplitude generator for NLO QCD
- conceived to break multi-particle bottlenecks (fast, numerically stable, flexible)
- now publicly available at *http://openloops.hepforge.org*

Sherpa2.1 [Hoeche, Hoeth, Krauss, Schoenherr, Schumann, Siegert, Zapp]

- Monte Carlo event generator
- fully automated interface to OpenLoops for NLO MEs
- automated matching (S-MC@NLO) and merging of jet multiplicities (MEPS@NLO)

New simulation of ttbb production

(i) 4F scheme: why?

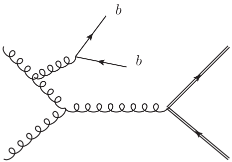
5F scheme ($m_b = 0$):

- $\mathrm{t\bar{t}b\bar{b}}$ MEs cannot describe collinear $g \to \mathrm{b\bar{b}}$ splittings

⇒ *inclusive* $\mathrm{t\bar{t}}$+b-jets requires $\mathrm{t\bar{t}}g$+PS i.e., $\mathrm{t\bar{t}} + \leq 2$ jets NLO merging required

4F scheme ($m_b > 0$):

- no b-quarks in the PDFS (only FS b-quarks from $g \to \mathrm{b\bar{b}}$ splittings)
- full $\mathrm{t\bar{t}b\bar{b}}$ phase space with NLO ME predictions
- IR singularities from (soft and) collinear $g \to \mathrm{b\bar{b}}$ configurations converted into finite $\ln(m_b)$ terms
- can describe ttbb and also ttb final states with unresolved b-quark (important for ttH analysis)



New simulation of ttbb production

(ii) NLO matched to PS

MC@NLO matching to Sherpa PS [S-MC@NLO]

- shower renders NLO calculations applicable to experimental analyses
- multi-parton emissions during evolution from hard scale to $\sim 1\mathrm{GeV}$
  $\Rightarrow$ resummation of large logarithms in exclusive observables
- $\mathrm{t\bar{t}} + N$ b-jets, with $N = 3, 4, \ldots$ via $g \rightarrow \mathrm{b\bar{b}}$ PS splittings

⋆ hadronisation + UE

Sherpa S-MC@NLO matching

[Höche, Krauss, Schönherr, Siegert '11]

Sherpa parton shower kernel based on *CS dipoles* D_{ijk}
($\neq$ Pythia/Herwig)

$$U(t_0,\mu_Q^2) = \Delta(t_0,\mu_Q^2)\mathcal{O}(\Phi_B) + \sum_{ijk}\int_{t_0}^{\mu_Q^2} \mathrm{d}\Phi_{R|B}\frac{D_{ijk}(\Phi_R)}{B(\Phi_B)}\Delta(t,\mu_Q^2)\mathcal{O}(\Phi_R),$$

S-MC@NLO **matching**: NO double counting of first QCD emission (R, D_{ijk})

$$\begin{aligned}\langle\mathcal{O}\rangle &= \int \mathrm{d}\Phi_B\left[B(\Phi_B)+V(\Phi_B)+\sum_{ijk}\int \mathrm{d}\Phi_{R|B}D_{ijk}(\Phi_R)\theta(\mu_Q^2-t)\right]U(t_0,\mu_Q^2)\\ &+ \int \mathrm{d}\Phi_R\left[R(\Phi_R)-\sum_{ijk}D_{ijk}(\Phi_R)\theta(\mu_Q^2-t)\right]\mathcal{O}(\Phi_R).\end{aligned}$$

Setup of S-MC@NLO $\mathrm{t\bar{t}b\bar{b}}$ simulation @ $8\mathrm{TeV}$ in 4F scheme

[Cascioli , Maierhöfer, N.M. ,Pozzorini, Siegert, Phys. Lett. B734 (2014) 210]

Categorisation according to number of b-jets

- anti-k_T jets with $R = 0.4$
- "physical" b-jet definition: any jet containing one *or more* b-quarks is considered b-jet (possible only with $m_b > 0$ matrix elements!)
- classification of events according to the number N_b of QCD b-jets with

$$p_T > 25\mathrm{GeV}, \qquad |\eta| < 2.5$$

Results for the following subsamples

- ttb ($N_b \geq 1$)
- $ttbb$ ($N_b \geq 2$)
- $ttbb_{100}$ ($N_b \geq 2$) in the $\mathrm{t\bar{t}H(b\bar{b})}$ **signal region** $m_{bb} > 100\mathrm{GeV}$

Simulation results

NLO vs S-MC@NLO cross sections

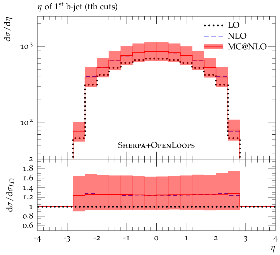
	ttb	ttbb	ttbb($m_{\mathrm{bb}} > 100$)
$\sigma_{\mathrm{LO}}[\mathrm{fb}]$	$2644^{+71\%\ +14\%}_{-38\%\ -11\%}$	$463.3^{+66\%\ +15\%}_{-36\%\ -12\%}$	$123.4^{+63\%\ +17\%}_{-35\%\ -13\%}$
$\sigma_{\mathrm{NLO}}[\mathrm{fb}]$	$3296^{+34\%\ +5.6\%}_{-25\%\ -4.2\%}$	$560^{+29\%\ +5.4\%}_{-24\%\ -4.8\%}$	$141.8^{+26\%\ +6.5\%}_{-22\%\ -4.6\%}$
$\sigma_{\mathrm{NLO}}/\sigma_{\mathrm{LO}}$	1.25	1.21	1.15
$\sigma_{\mathrm{MC}}[\mathrm{fb}]$	$3313^{+32\%\ +3.9\%}_{-25\%\ -2.9\%}$	$600^{+24\%\ +2.0\%}_{-22\%\ -2.1\%}$	$181.0^{+20\%\ +8.1\%}_{-20\%\ -6.0\%}$
$\sigma_{\mathrm{MC}}/\sigma_{\mathrm{NLO}}$	1.01	1.07	1.28

- good perturbative convergence: uncertainty goes from 70% (LO) to 20 – 30% (NLO)
- K-factors small and rather independent of selection ($\mathrm{t\bar{t}b}$ sample free from large $\ln(m_b)$ in 4F scheme)
- surprisingly large S-MC@NLO effect (∼30%) in Higgs-signal region of *ttbb*

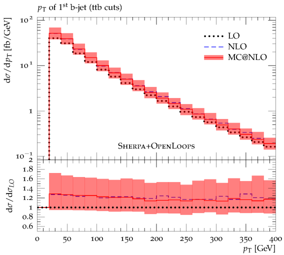
$t\bar{t}b$ analysis ($N_b \geq 1$)

b-jet and top-quark distributions

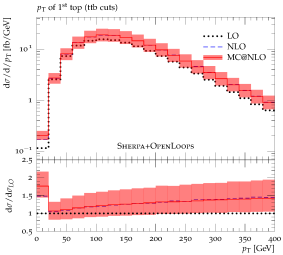
η_{b_1}



p_{T,b_1}



p_{T,t_1}



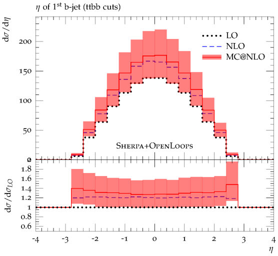
Reliable perturbative prediction

- shape of 1^{st} b-jet very stable wrt NLO corrections
- excellent S-MC@NLO vs NLO agreement (first principle theoretical prediction, small shower dependence)

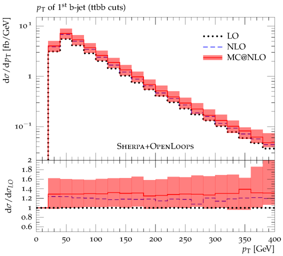
$t\bar{t}b\bar{b}$ analysis ($N_b \geq 2$)

b-jet and top-quark distributions

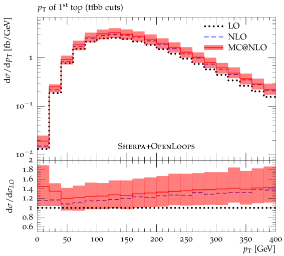
η_{b_1}



p_{T,b_1}



p_{T,t_1}



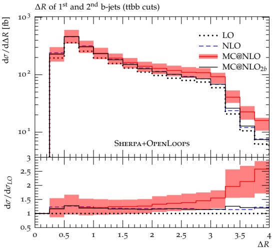
Good stability

- moderate S-MC@NLO excess wrt to NLO
- small distortions in bottom and top p_T

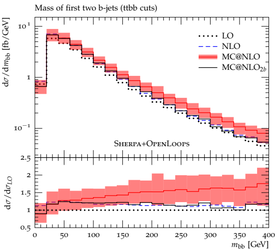
$\mathrm{t\bar{t}b\bar{b}}$ analysis ($N_b \geq 2$)

b-jet correlations

$\Delta R_{b_1 b_2}$



m_{b1b2}



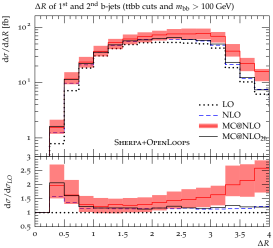
Completely different behaviour

- NLO corrections remain quite flat
- significant S-MC@NLO enhancement at large $\Delta R_{b_1 b_2}$ and large $m_{b_1 b_2}$
- reaches 30% at $m_{b1b2} \sim 125\mathrm{GeV}$
- completely disappears if PS $g \to \mathrm{b\bar{b}}$ splittings switched off! (MC@NLO$_{2b}$ curve)

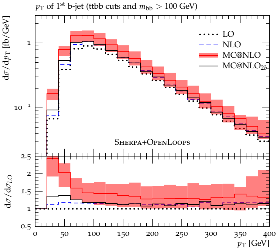
$t\bar{t}b\bar{b}$ analysis ($N_b \geq 2$) with $m_{b_1 b_2} > 100\mathrm{GeV}$

b-jet correlations

$\Delta R_{b_1 b_2}$



p_{T,b_1}

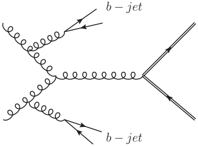


S-MC@NLO **excess at large m_{bb} from back to back soft jets**

- factor-2 enhancement at $\Delta R \sim \pi$
- factor-2 enhancement at small p_T

Consistent with soft-collinear behaviour of gluon-jet emissions

Contribution from double collinear $g \to b\bar{b}$ splittings **(NOT present in $t\bar{t}b\bar{b}$ simulations in 5F scheme)**



Theoretical (scale) uncertainties

Scale choice(s)

Scale choice and variations are essential

- multi-scale problem $m_b^2 \lesssim Q_{ij}^2 \lesssim m_{\mathrm{t\bar{t}}}^2$ (5 to $500\mathrm{GeV}$)
- factor 2 variation → O(80%) $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ variation

Natural scale choice

Idea: factorization of hard $pp \to \mathrm{t\bar{t}}$ (scale $\sim m_t \sim E_{T,t}$) plus b-jet emission ($\sim E_{T,b}$)

$$\Rightarrow \mu_{\mathrm{CMMPS}} = \sqrt[4]{\prod_{i=\mathrm{t,\bar{t},b,\bar{b}}} E_{\mathrm{T},i}}$$

→ small K-factors, no large logs ✓

Theoretical (scale) uncertainties

Data uncertainties

usually NLO uncertainties estimated with factor 2 scale variations ($\mu = \xi \times \mu_{\mathrm{CMMPS}}$, with $\frac{1}{2} \leq \xi \leq 2$)

⇒ Is it enough?

2 kinds of uncertainties:

- **Normalisation uncertainties**
  - can be reduced with data driven approach
- **Shape uncertainties**
  - less constrained by experimental data
  - usual factor 2 scale variations are less appropriate for them

⇒ deep study of shape distortions is needed

Theoretical (scale) uncertainties

Cross sections uncertainties

Consider ***kinematic distortions* of** $\mu_{\rm R}, \mu_{\rm F}, \mu_Q$ using various combinations of the variables

$$\mu_{\rm CMMPS}, \qquad m_{\rm b\bar b},$$

$$H_{T,{\rm b(t)}} = E_{T,{\rm b(t)}} + E_{T,{\rm \bar b(\bar t)}}, \quad H_T = H_{T,{\rm t}} + H_{T,{\rm b}}$$

Scale	default	glo-HT	glo-Mt	glo-soft	R-Mbb	R-HTb	R-HTt	Q-CMMPS	Q-Mt
$\mu_{\rm R}$	$\mu_{\rm CMMPS}$	$H_T/2$	$m_{\rm t}$	$\mu_{\rm CMMPS}$	$(m_{\rm t} m_{\rm b\bar b})^{1/2}$	$(m_{\rm t} H_{T,{\rm b}}/2)^{1/2}$	$(m_{\rm t} H_{T,{\rm t}}/2)^{1/2}$	$\mu_{\rm CMMPS}$	$\mu_{\rm CMMPS}$
$\mu_{\rm F}$	$H_{T,{\rm t}}/2$	$H_T/2$	$m_{\rm t}$	$\mu_{\rm CMMPS}$	$H_{T,{\rm t}}/2$	$H_{T,{\rm t}}/2$	$H_{T,{\rm t}}/2$	$H_{T,{\rm t}}/2$	$H_{T,{\rm t}}/2$
μ_Q	$H_{T,{\rm t}}/2$	$H_T/2$	$m_{\rm t}$	$\mu_{\rm CMMPS}$	$H_{T,{\rm t}}/2$	$H_{T,{\rm t}}/2$	$H_{T,{\rm t}}/2$	$\mu_{\rm CMMPS}$	$m_{\rm t}$
Cuts	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$
ttb	0%	−41%	−27%	+4.7%	+2.3%	1.1%	−32%	−3.5%	−0.3%
$ttbb$	0%	−33%	−17%	−0.7%	+0.2%	3.4%	−22%	−6.4%	−1.1%
$ttbb_{100}$	0%	−29%	−13%	−9.2%	−5.6%	+2.5%	−17%	−14%	−2.9%

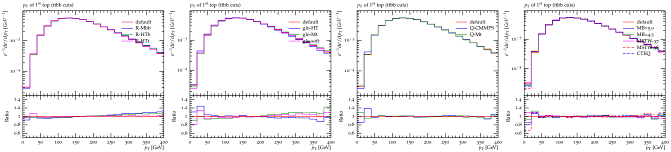
glo	single global scale: hard, fixed and softer
R	renormalisation scale (dominant!): modify or avoid b-jet dependence
Q	resummation-scale (PS uncertainties): softer and fixed

$\Rightarrow$ variation of relative $\rm t\bar t b, t\bar t b\bar b, t\bar t b\bar b_{100}$ rates

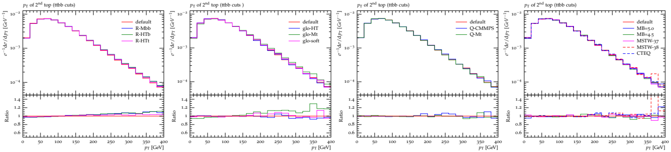
Theoretical (scale) uncertainties

Shape uncertainty of top-p_T

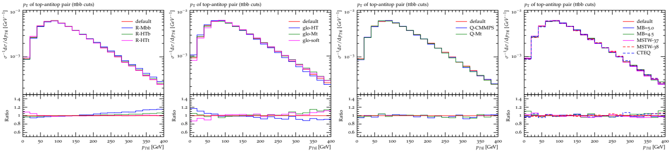
$p_{\mathrm{T},\mathrm{t}_1}$ $(ttbb)$



$p_{\mathrm{T},\mathrm{t}_2}$ $(ttbb)$



$p_{\mathrm{T},\mathrm{t}_1\mathrm{t}_2}$ $(ttbb)$



$\Rightarrow$ $\sim$10% shape variations (20% in the tails) driven by top-dependence of μ_R

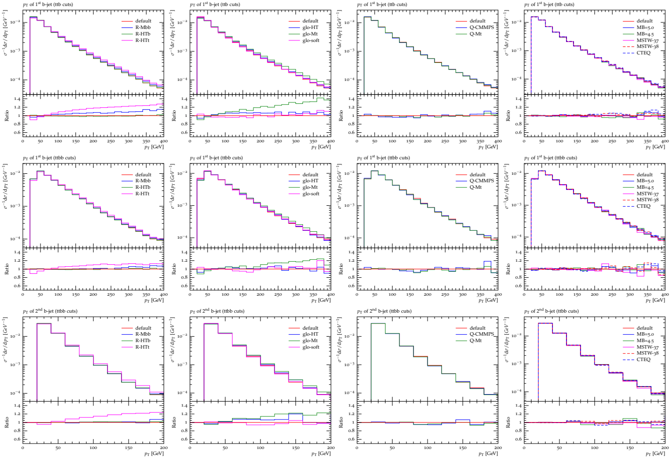
Scale variations

Shape uncertainty of b-jet p_{T}

$p_{\mathrm{T},\mathrm{b}_1}$ (ttb)

$p_{\mathrm{T},\mathrm{b}_1}$ $(ttbb)$

$p_{\mathrm{T},\mathrm{b}_2}$ $(ttbb)$



⇒ ∼10-20% variations (40% in the tails) driven by b-dependence of μ_{R}

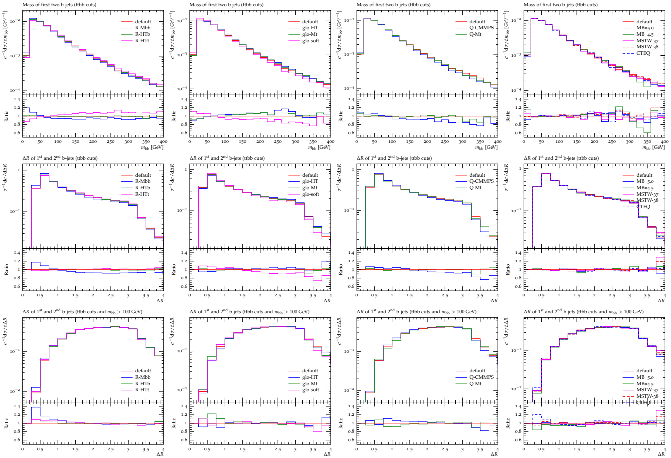
Theoretical (scale) uncertainties

Shape uncertainty of b-jet correlations

$m_{\mathrm{b}_1\mathrm{b}_2}$ $(ttbb)$

$\Delta R_{\mathrm{b}_1\mathrm{b}_2}$ $(ttbb)$

$\Delta R_{\mathrm{b}_1\mathrm{b}_2}$ $(ttbb_{100})$



$\Rightarrow$ $\sim$10-20% variations driven by b-dependence of μ_{R} (at small m_{bb} and ΔR) and (aggressive) reduction of μ_Q in the tail

Scale variations

MC Samples

All predictions available for experimental analyses as a series of MC samples

- central scales
- normalisation + shape variations (scales, PDFs, Shower, $m_b \ldots$)
- top decays
- fully showered
- hadronisation + UE

Summary

S-MC@NLO $t\bar{t}b\bar{b}$ simulation in the 4F scheme

- ⋆ MEs with $m_{\mathrm{b}} > 0$ cover full b-quark phase space → complete $t\bar{t}$+b-jets simulation independent of $t\bar{t}$+light-jets
- ⋆ new b-jets production mechanism: double $g \to b\bar{b}$ splittings **surprisingly important for $t\bar{t}H(b\bar{b})$ analysis**
- ⋆ S-MC@NLO: 20-30% normalisation and ∼10% shape uncertainties

- deep study of shape uncertainties

⇒ 8TeV "all inclusive" samples available (UE , Hadronisation, Decays etc.) ready for experimental analysis

- fundamental step towards complete $t\bar{t}$+multi-jets NLO analysis

[Höche, Krauss, Maierhöfer,Pozzorini, Schönherr, Siegert '14]

BACKUP slides

Warm-up: validation of tools

Validation of tools

- reproduced NLO ttbb tt+0,1,2 jets in the literature
- new 8TeV LO and NLO *ttbb* σ_{tot} results for 3rd HXSWG report

setup	μ_0	$\sigma_{LO}[fb]$	$\bar{\sigma}_{LO}[fb]$	$\sigma_{NLO}[fb]$	K	$\bar{K}$
$S1$	Mt	$503(1)^{+84\%}_{-42\%}$	$342(2)^{+74\%}_{-39\%}$	$671(3)^{+34\%}_{-28\%}$	1.34	1.96
$S1$	μ_{BDDP}	$861(2)^{+96\%}_{-45\%}$	$557(3)^{+83\%}_{-42\%}$	$901(3)^{+23\%}_{-27\%}$	1.04	1.62
$S2$	Mt	$37.21(7)^{+87\%}_{-43\%}$	$25.41(8)^{+75\%}_{-40\%}$	$45.5(1)^{+29\%}_{-26\%}$	1.23	1.79
$S2$	μ_{BDDP}	$54.8(1)^{+95\%}_{-45\%}$	$36.2(2)^{+82\%}_{-42\%}$	$54.3(2)^{+18\%}_{-24\%}$	0.99	1.50

- using "wise" scale essential for convergence of perturbation theory
- NLO can reduce uncertainties up to $\sim 25\%$ at $8\,TeV$

Additional m_{b} and PDF variations with potential impact on shape (and normalisation)

	$M_b = 5.0$	$M_b = 4.5$	CTEQ 4F	MSTW$_{37}$	MSTW$_{38}$
Cuts	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$	$\Delta\sigma/\sigma$
ttb	-3.5%	$+4.4\%$	-10%	-0.1%	$+2.6\%$
$ttbb$	-0.7%	$+2.7\%$	-9.3%	$+0.2\%$	$+4.2\%$
$ttbb_{100}$	-0.1%	$+4.4\%$	-7.8%	-0.7%	$+6.9\%$

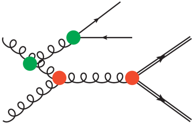
- conservative b-mass variations $m_{\mathrm{b}} = 4.75 \pm 0.25\mathrm{GeV}$ (impact on collinear regions)
- compare central MSTW to central CT10 PDF and MSTW variations with large gluon-shape distortion (MSTW eigenvector 19)

Choice of μ_{R}, μ_{F} and μ_Q

Scale choice in $\alpha_S^4(\mu^2)$ is crucial

- widely separated scales $m_{\mathrm{b}} \leq Q_{ij} \lesssim m_{\mathrm{t\bar{t}b\bar{b}}}$ can generate huge logs

Dynamical "BDDP" scale [Bredenstein, Denner, Dittmaier, S. P. '10] guarantees good convergence by adapting to b-jet p_{T}



$$\alpha_S^4(\mu_{\mathrm{BDDP}}^2) = \alpha_S^4(m_t\sqrt{p_{\mathrm{T},b1}p_{\mathrm{T},b2}}) \simeq \alpha_S^2(m_t^2)\alpha_S(p_{\mathrm{T},b1}^2)\alpha_S(p_{\mathrm{T},b2}^2)$$

Natural generalisation (for $p_{\mathrm{T}} \to 0$ region)

$$\mu_{\mathrm{R}}^4 = \prod_{i=\mathrm{t},\bar{\mathrm{t}},\mathrm{b},\bar{\mathrm{b}}} E_{\mathrm{T},i} = \prod_{i=\mathrm{t},\bar{\mathrm{t}},\mathrm{b},\bar{\mathrm{b}}} \sqrt{m_i^2 + p_{\mathrm{T},i}^2}$$

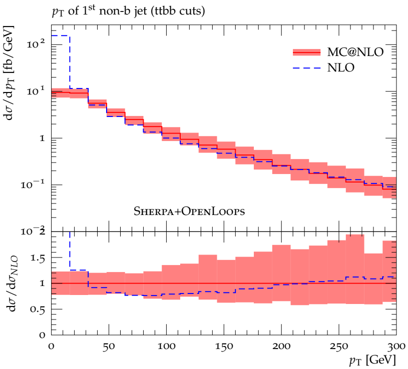
Factorisation and Resummation scales (available phase space for QCD emission)

$$\mu_{\mathrm{F}} = \mu_Q = \frac{1}{2}(E_{\mathrm{T,t}} + E_{\mathrm{T,\bar{t}}})$$

Validation Plot 1:

ttbb **analysis** ($N_b \geq 2$)

ttbb **analysis ($N_b \geq 2$):** 1^{st} light-jet p_{T} distribution



S-MC@NLO vs NLO

- in good (5%) agreement in the tail
- Sudakov damping of NLO IR singularity at $p_{\mathrm{T}} \to 0$
- ∼25% deviation at intermediate p_{T} consistent with expected NNLO effect

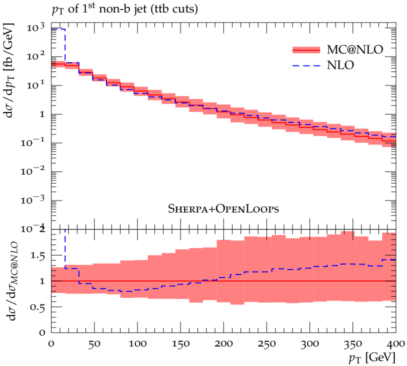
S-MC@NLO **scale uncertainty**

- LO-like uncertainty (∼100%) in the tail irrelevant for $\mathrm{t\bar{t}Hb\bar{b}}$
- NLO-like accuracy (∼25%) up to $100\mathrm{GeV}$

Validation Plot 2:

ttb **analysis ($N_b \geq 1$)**

***ttb* analysis ($N_b \geq 1$):** 1^{st} light-jet p_T distribution (responsible for double splittings)



S-MC@NLO **vs** NLO

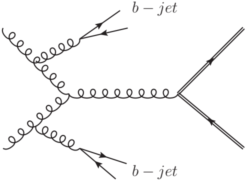
- Sudakov damping of NLO IR singularity at $p_\mathrm{T} \rightarrow 0$
- 25% NLO excess in the hard tail (probably due to dynamic μ_Q, multi-jet final state, unresolved b-quark)

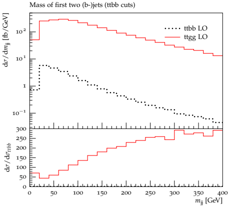
S-MC@NLO **scale uncertainty**

- LO-like uncertainty ($\sim$100%) in the tail irrelevant for $t\bar{t}Hb\bar{b}$
- NLO-like accuracy ($\sim$30%) up to 70GeV

$\Rightarrow$ NLO-like accuracy in the region relevant for $t\bar{t}Hb\bar{b}$

Double $g \to b\bar{b}$ splitting contributions





Consistent with MC enhancement

- $t\bar{t}gg/t\bar{t}b\bar{b}$ ratio grows at same rate of S-MC@NLO excess
- emission of back-to-back small-p_T gluons enhanced by soft-collinear singularity

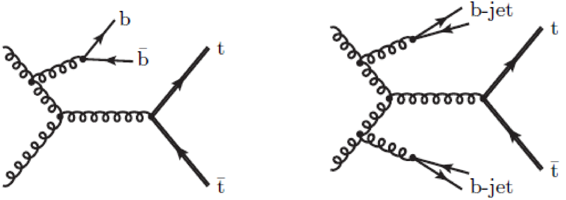
Don't fit into conventional hard-scattering $t\bar{t}b\bar{b}$ picture

- present also in $t\bar{t}$+jets LO merged samples
- but large effect in hard $t\bar{t}Hb\bar{b}$ signal region unexpected

Implications for theory systematics in $t\bar{t}$+HF

- understanding PS systematics crucial (both for 4F $t\bar{t}b\bar{b}$ or 5F $t\bar{t}$+jets)
- in $t\bar{t}Hb\bar{b}$ signal region 4F $t\bar{t}b\bar{b}$ S-MC@NLO provides 1^{st} $g \to b\bar{b}$ splitting at NLO

Why NLO matching for $\mathrm{t\bar{t}b\bar{b}}$ production in 4F scheme



5F scheme ($m_\mathrm{b} = 0$): $\mathrm{t\bar{t}b\bar{b}}$ **MEs cannot describe collinear $\mathrm{g} \to \mathrm{b\bar{b}}$ splittings**

⇒ *inclusive* $\mathrm{t\bar{t}}$+b-jets simulation (quite important for exp. analyses!) requires $\mathrm{t\bar{t}g}$+PS, i.e. $\mathrm{t\bar{t}}+ \leq$2 jets NLO merging

[Höche, Krauss, Maierhöfer,Pozzorini, Schönherr, Siegert '14]

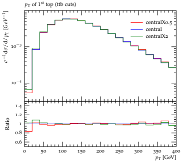
4F scheme ($m_\mathrm{b} > 0$): $\mathrm{t\bar{t}b\bar{b}}$ **MEs cover full b-quark phase space**

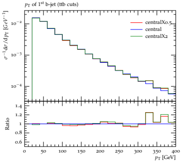
⇒ S-MC@NLO $\mathrm{t\bar{t}b\bar{b}}$ sufficient for inclusive $\mathrm{t\bar{t}}$+b-jets simulation

- access to **new** $\mathrm{t\bar{t}} + 2$**b-jets production mechanism** wrt 5F scheme: double collinear $g \to \mathrm{b\bar{b}}$ splittings (surprisingly important impact on $\mathrm{t\bar{t}Hb\bar{b}}$ analysis!)

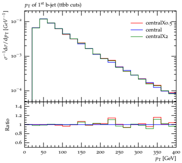
μR standard variations

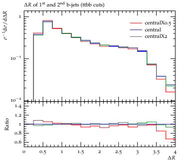
$t\bar{t}b$





$t\bar{t}b\bar{b}$





$t\bar{t}b\bar{b}_{100}$