Theory predictions for $t\bar{t} + b$-jet background to $t\bar{t}H(b\bar{b})$

Stefano Pozzorini

in collaboration with S. Guindon, M. Harrendorf, M. Moreno-Llacer, C. Neu, L. Reina

HXSWG-$t\bar{t}H$ meeting, 6 February 2017

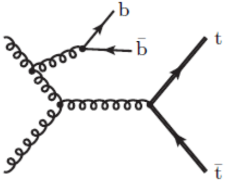
FNSNF
Fonds national suisse
Schweizerischer Nationalfonds
Fondo nazionale svizzero
Swiss National Science Foundation

Universität Zürich[UZH]

Outline

1. NLO+PS $t\bar{t}b\bar{b}$ simulations
2. Combination of $t\bar{t}+X$ and $t\bar{t}b\bar{b}$ simulations
3. $t\bar{t}b\bar{b}$ uncertainties
4. NLO+PS $t\bar{t}b\bar{b}$ tool comparisons

Irreducible $t\bar{t}b\bar{b}$ QCD background at NLO+PS



Nontrivial features of $pp \to t\bar{t}b\bar{b}$

- 6 external coloured partons
- 34 LO diagrams, multiple scales from 5 to 500 GeV
- dominated by topologies with FS $g \to b\bar{b}$ splittings

$\Rightarrow$ collinear regions and m_b important (resummation of IS $g \to b\bar{b}$ splittings not)

NLO+PS $t\bar{t}b\bar{b}$ 5F scheme ($m_b = 0$) with POWHEL [Garzelli et al '13/'14]

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

NLO merging $t\bar{t} + 0, 1, 2$ jets 5F with SHERPA+OPENLOOPS or MG5AMC@NLO

- challenging for $t\bar{t}$+HF and still based on $m_b = 0$ MEs + shower in collinear regions

NLO+PS $t\bar{t}b\bar{b}$ 4F scheme ($m_b > 0$) with SHERPA+OPENLOOPS [Cascioli et al '13] or MG5AMC@NLO

- $t\bar{t}b\bar{b}$ NLO MEs cover full b-quark phase space $\Rightarrow$ **recommended for $t\bar{t}H(b\bar{b})$**

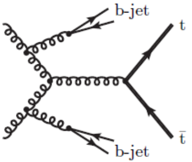
$\Rightarrow$ **NLO accuracy for any inclusive $t\bar{t}$+b-jet observable with ≥ 1 b-jets!**

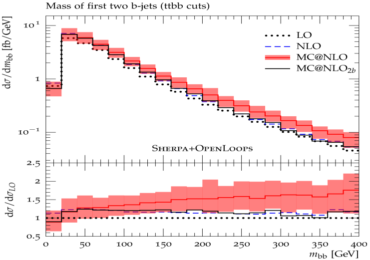
S–MC@NLO $t\bar{t}b\bar{b}$ at 8 TeV in 4F scheme [Cascioli et al '13]

Convergence of 4F scheme: no large $\log(m_b)$ in ttb region!

	ttb	$ttbb$	$ttbb\,(m_{bb} > 100)$
$\sigma_{\rm LO}\,$[fb]	$2644^{+71\%+14\%}_{-38\%-11\%}$	$463.3^{+66\%+15\%}_{-36\%-12\%}$	$123.4^{+63\%+17\%}_{-35\%-13\%}$
$\sigma_{\rm NLO}\,$[fb]	$3296^{+34\%+5.6\%}_{-25\%-4.2\%}$	$560^{+29\%+5.4\%}_{-24\%-4.8\%}$	$141.8^{+26\%+6.5\%}_{-22\%-4.6\%}$
$\sigma_{\rm NLO}/\sigma_{\rm LO}$	1.25	1.21	1.15
$\sigma_{\rm MC@NLO}\,$[fb]	$3313^{+32\%+3.9\%}_{-25\%-2.9\%}$	$600^{+24\%+2.0\%}_{-22\%-2.1\%}$	$181^{+20\%+8.1\%}_{-20\%-6.0\%}$
$\sigma_{\rm MC@NLO}/\sigma_{\rm NLO}$	1.01	1.07	1.28

MC@NLO enhancement in Higgs region from double $g \to b\bar{b}$ splittings





One $g \to b\bar{b}$ splitting from PS

$\Rightarrow$ TH uncertainties related to NLOPS matching and shower crucial!

Outline

Approach proposed in YR4

NLOPS 4F $t\bar{t}b\bar{b}$ sample

- can be applied in its full phase space (no generation cuts)

$\Rightarrow$ inclusive description of $t\bar{t}+ \geq 1b$-quarks

- includes also contributions corresponding to $gb \to t\bar{t}b$ in the 5F scheme

Inclusive $t\bar{t} + X$ sample

- needs to be restricted to $t\bar{t} + 0\,b$-quarks to avoid double counting

$\Rightarrow$ veto events containing b-quarks not arising from showered top decays or MPI or UE

Possible implementations

- $t\bar{t} + X$ and $t\bar{t}b\bar{b}$ samples independent samples
- reweighting of $t\bar{t} + X$ sample through $t\bar{t}b\bar{b}$ in the $t\bar{t}+ \geq 1b$-quarks region

Refinement for region of small $p_{T,b}$

Caveat

- $t\bar{t}b\bar{b}$ sample yields (small) contribution to $t\bar{t}+0\,b$-jet categories of EXP analysis
- $t\bar{t}+0\,b$-jet categories (dominated by $t\bar{t}$+gluons/light-quarks) can bias $t\bar{t}b\bar{b}$ fit

$\Rightarrow$ preferable to restrict $t\bar{t}b\bar{b}$ to $t\bar{t}+b$-jet categories

Proposal: smooth matching of $t\bar{t}+X$ and $t\bar{t}b\bar{b}$ samples

- using smearing function of leading b-jet p_T, such as

$$\xi(p_{T,b}) = \begin{cases} 0 & \equiv \text{ pure } t\bar{t}+0b & \text{for} \quad p_{T,b} < p_{T,\min} \\ \frac{1}{2}\left[1-\cos\left(\pi\frac{p_{T,b}-p_{T,\min}}{p_{T,\max}-p_{T,\min}}\right)\right] & & \text{for} \quad p_{T,\min} < p_{T,b} < p_{T,\max} \\ 1 & \equiv \text{ pure } t\bar{t}+\geq 1b & \text{for} \quad p_{T,b} > p_{T,\max} \end{cases}$$

- with transition region in the vicinity of experimental b-jet threshold, e.g. $[p_{T,\min}, p_{T,\max}] = [15, 25]$ GeV
- same matching procedure should be used in ATLAS and CMS for a transparent comparison and combination of EXP results

Outline

Scale choices (YR4) and uncertainties (no proposal yet)

Factorisation (μ_Q) and resummation (μ_Q) scales

$$E_{T_i} = \sqrt{m_i^2 + p_{T,i}^2}$$

$$\mu_F = \mu_Q = \frac{H_T}{2} = \frac{1}{2} \sum_{i=t,\bar{t},b,\bar{b}} E_{T,i}$$

$\mu_Q \equiv$ shower starting scale is a free paramater in MC@NLO (not in Powheg)

CKKW-like (softer) renormalisation scale

$$\mu_R = \mu_{\text{CKKW}} = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$$

Scale variations (leading uncertainty) ∼20-30%

- factor-2 variations of μ_R and μ_F ⇔ normalisation
- "kinematic" variations of μ_R, μ_F, μ_Q ⇔ shape
- variations of μ_Q in MC@NLO and h_{damp} in Powheg ⇔ NLOPS matching

Other variations

- PDF variations (only few percent)
- shower variations: tune variations, shower recoil scheme, . . .

Correlation of TH uncertainties between categories

Categories

- $t\bar{t}h(b\bar{b})$ analyses based on simultaneous fit of MC to data in various categories with different # of light- and b-jets
- correlations crucial to constrain background in signal region (with multiple b-jets)

Between $t\bar{t}$+light-jet and $t\bar{t}+b$-jet categories

- uncertainties should be uncorrelated

Between sub-categories (e.g. ttb, $ttbb$, ttB)

- uncertainties should be correlated

Motivation: independent shower, matching and ME variations account for different types of uncertainties (e.g. related to collinear $g \to b\bar{b}$ splittings or hard b-production) $\Rightarrow$ no need of separate categories with uncorrelated uncertainties

Outline

1. NLO+PS $t\bar{t}b\bar{b}$ simulations
2. Combination of $t\bar{t}+X$ and $t\bar{t}b\bar{b}$ simulations
3. $t\bar{t}b\bar{b}$ uncertainties
4. NLO+PS $t\bar{t}b\bar{b}$ tool comparisons

Tuned comparison of NLO+PS $t\bar{t}b\bar{b}$ simulations at 13 TeV

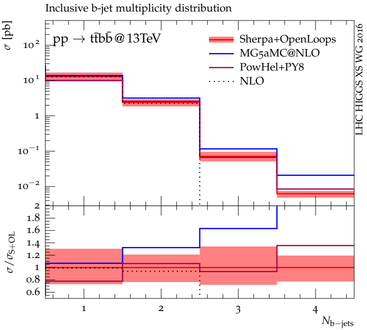
Different NLO+PS methods, showers, and m_b treatments

Tool	Matching	Shower	m_b[GeV]	gencuts
SHERPA2.1+OPENLOOPS	SMC@NLO	Sherpa 2.1	4.75 (4F)	no
MG5_aMC@NLO	MC@NLO	Pythia 8.2	4.75 (4F)	no
POWHEL	Powheg	Pythia 8.2	0 (5F)	$p_{T,b} > 4.75\,\text{GeV}$ $\frac{m_{bb}}{2} > 4.75\,\text{GeV}$

Detailed setup

- HXSWG's Yellow Report 4 [arXiv:1610.07922]
- https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposalTtbb

Inclusive $t\bar{t} + b$-jet multiplicity distribution



- S-MC@NLO (Sherpa+OpenLoops) with $\mu_{R,F}$ variations
- MG5_aMC@NLO+PY8 w.o. variations
- Powhel+PY8 w.o. variations

NLO vs NLO+PS

- decent agreement in NLO accurate bins with ≥ 1 and ≥ 2 b-jets

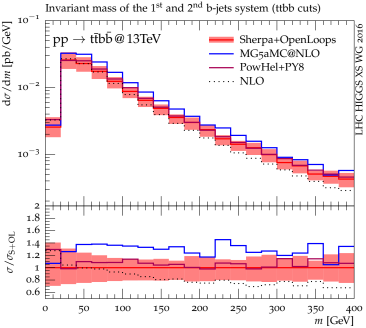
S-MC@NLO vs PowHel+PY8

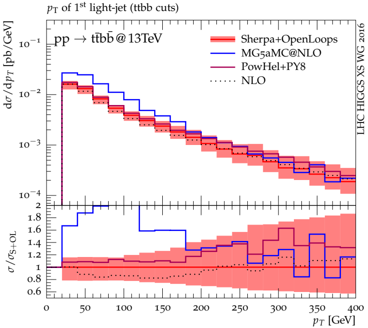
- good overall agreement in spite of differences in matching method, parton shower, N_f-scheme and ad-hoc cuts in Powhel

S-MC@NLO vs MG5aMC@NLO

- good agreement only for $\geq 1\, b$-jets despite similar matching method and same N_f

$t\bar{t}b\bar{b}$ distributions with $\geq 2b$-jets



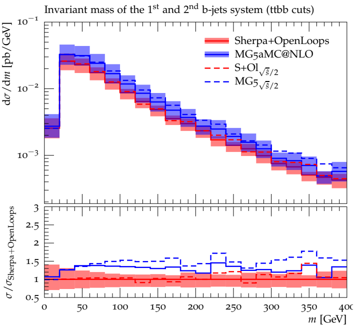


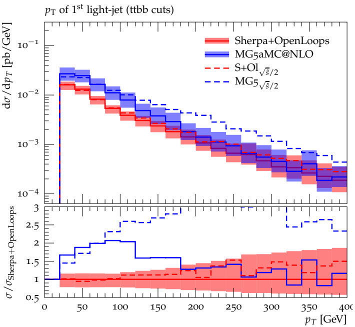
S-MC@NLO vs PowHel+PY8

- well consistent also in observables that receive significant shower corrections
- confirmation of "double-splitting effects" (see e.g. m_{bb})

S-MC@NLO vs MG5aMC@NLO

- 40% enhancement of $t\bar{t}+2\mathrm{b}$ XS & sizable differences in NLO radiation pattern
- related to strong sensitivity to resummation scale (shower starting scale) in MG5 . . .





Nominal MG5_aMC and Sherpa+OpenLoops predictions in YR4

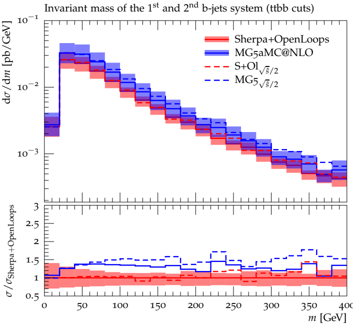
- MG5_aMC supports* only $\mu_Q = f(\xi)\sqrt{\hat{s}}$ $\Rightarrow$ smearing function restricted to $0.1 < f(\xi) < 0.25$ to mimic recommended $\mu_Q = H_T/2$ implemented in Sherpa

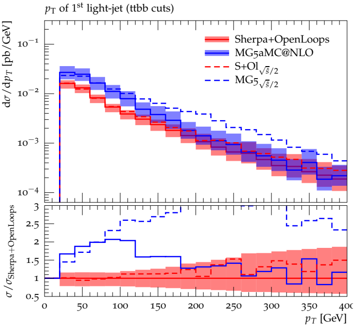
New: μ_Q variations enhance the discrepancy

- $\mu_Q = \sqrt{\hat{s}}/2$ in Sherpa to mimic MG5_aMC default choice $0.1 < f(\xi) < 1$
- strong μ_Q-sensitivity of MG5_aMC $\Rightarrow$ much more pronounced deviations

*New: latest MG5aMC@NLO version implements $\mu_Q = H_T/2$ as default resummation scale

Dependence on resummation scale μ_Q





Interpretation of large Sherpa+OpenLoops vs MG5aMC@NLO differences?

- can we exclude bugs or misusage of tools?
- related to single μ_Q scale in MC@NLO vs multi-scale process $(M_{tt} \sim 100 M_b)$?
- small μ_Q sensitivity and NLOPS/NLO difference in SHERPA+OPENLOOPS suggests MG5AMC@NO-specific issue (?)
- SHERPA+OPENLOOPS $t\bar{t}b\bar{b}$ 4F supported by POWHEL $t\bar{t}b\bar{b}$ 5F, but should be confirmed by POWHEG $t\bar{t}b\bar{b}$ 4F simulation

Conclusions

NLO+PS $t\bar{t}b\bar{b}$ 4F simulations

- recommended for $t\bar{t}$+b-jet backgrounds
- technically automated but physically very tricky (many coloured partons and scales, $g \to b\bar{b}$ splittings)

Matching $t\bar{t} + X$ with $t\bar{t}b\bar{b}$ 4F samples

- veto events with additional b-quarks in $t\bar{t} + X$ sample
- smooth implementation as a function of leading-jet p_T

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

- dominated by μ_R, μ_F, μ_Q scale dependence
- detailed recommendations for shape uncertainties still needed

Discrepancy between Sherpa+OpenLoops and MG5aMC@NLO

- related to μ_Q dependence of MG5aMC@NLO
- calls for thorough investigation/validation

Backup slides

Scale variations for shape (not for normalisation) uncertainties

Consider (aggressive but not fully unreasonable) ***kinematic distortions*** **of** μ_R, μ_F, μ_Q using various combinations of the variables

$$\mu_{\text{CMMPS}} = \prod_{i=t,\bar{t},b\bar{b}} E_{T,i}^{1/4}, \qquad m_{b\bar{b}}, \qquad H_{T,b(t)} = E_{T,b(t)} + E_{T,\bar{b}(\bar{t})}, \qquad H_T = H_{T,t} + H_{T,b}$$

Scale	default	glo-HT	glo-Mt	glo-soft	R-Mbb	R-HTb	R-HTt	Q-CMMPS	Q-Mt
μ_R	μ_{CMMPS}	$H_T/2$	m_t	μ_{CMMPS}	$(m_t m_{b\bar{b}})^{1/2}$	$(m_t H_{T,b}/2)^{1/2}$	$(m_t H_{T,t}/2)^{1/2}$	μ_{CMMPS}	μ_{CMMPS}
μ_F	$H_{T,t}/2$	$H_T/2$	m_t	μ_{CMMPS}	$H_{T,t}/2$	$H_{T,t}/2$	$H_{T,t}/2$	$H_{T,t}/2$	$H_{T,t}/2$
μ_Q	$H_{T,t}/2$	$H_T/2$	m_t	μ_{CMMPS}	$H_{T,t}/2$	$H_{T,t}/2$	$H_{T,t}/2$	μ_{CMMPS}	m_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

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)

Shape variations of differential observables

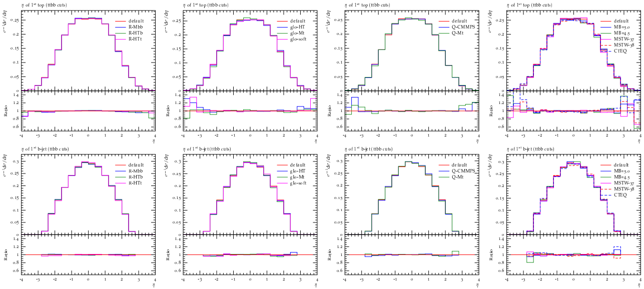
The following plots show a representative selection of shape uncertainties

- normalisation uncertainties removed by normalising all distributions to one
- columns represent (1) R-type (2) glo-type (3) Q-type (4) m_{b}+PDFs variations

Shape uncertainty of top-quark and b-jet rapidities

η_{t_1} $(ttbb)$

η_{b_1} $(ttbb)$



⇒ percent-level variations for $|\eta| < 2.5$; η_{b} very stable

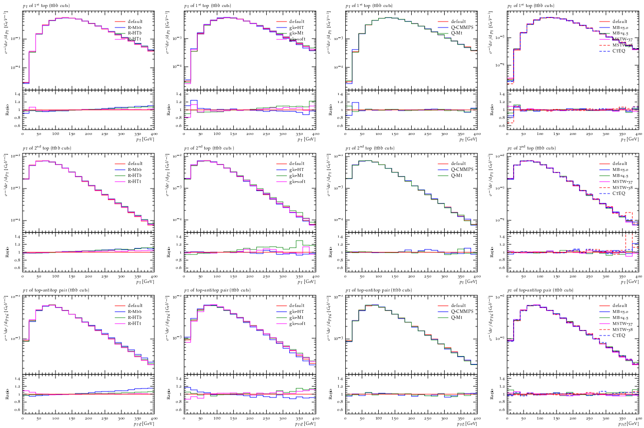
28

Shape uncertainty of top-p_{T}

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

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

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



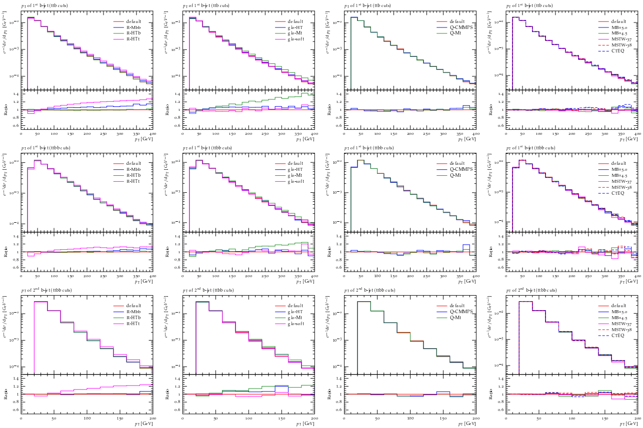
⇒ ~10% variations (20% in the tails) driven by top-dependence of μ_{R}

Shape uncertainty of b-jet p_T

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

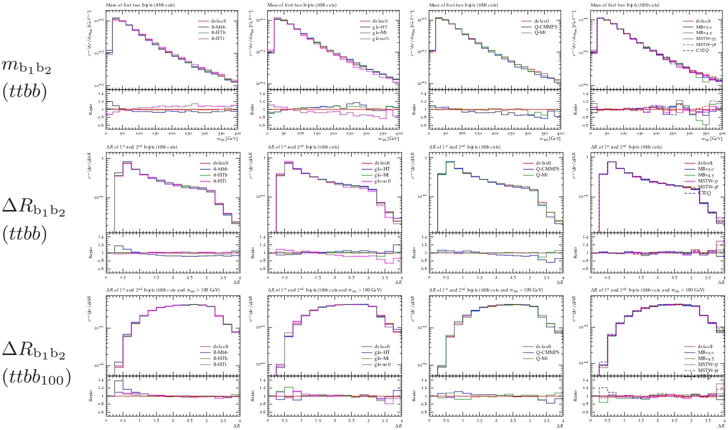
$p_{\mathrm{T},\mathrm{b}_1}$ $(ttbb)$

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



$\Rightarrow$ ~10-20% variations (40% in the tails) driven by b-dependence of μ_R

Shape uncertainty of b-jet correlations

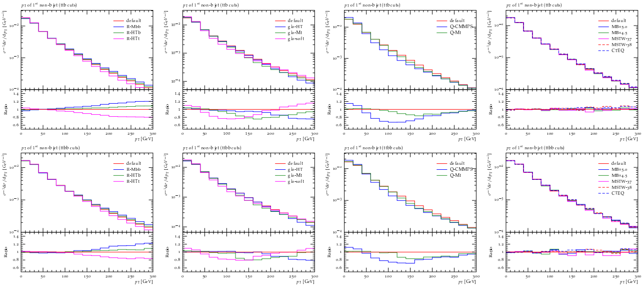


⇒ ~10-20% variations driven by b-dependence of μ_{R} (at small m_{bb} and ΔR) and (agressive) reduction of μ_Q in the tail

Shape uncertainty of 1^{st} light-jet p_{T}

p_{T,j_1} (ttb)

p_{T,j_1} $(ttbb)$



⇒ up to ∼30% variations at intermediate p_{T} values. Indicates that the considered variations (dominated by choice of soft resummation scale) are (probably too) conservative