Generation comparison for top physics in CMS

- ☞ Top quark pair generation in CMS
- ☞ Standard generation vs six fermions
- ☞ The ME-PS matching
- ☞ Next steps towards data

E. Chabert, R. Chierici, T. Le Grand, S. Perries
CNRS/IPN Lyon

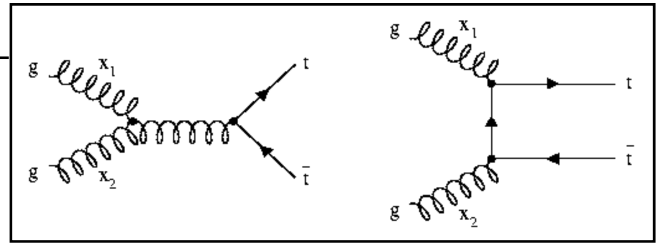
DESY, October 2007

Generation strategies for top (SM) physics in CMS

How did we try and how do we plan to generate top pair physics in CMS?

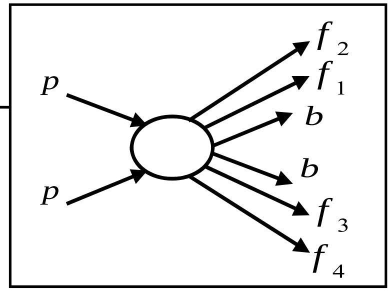
✓ LO "standard": TopRex, MadGraph

- pp→tt × tt→$bf_1f_2bf_3f_4$; pp→(tt→)$bf_1f_2bf_3f_4$
- top spin correlations included
- with or without narrow width approximation



✓ LO 6 fermions: Phantom (MadGraph also)

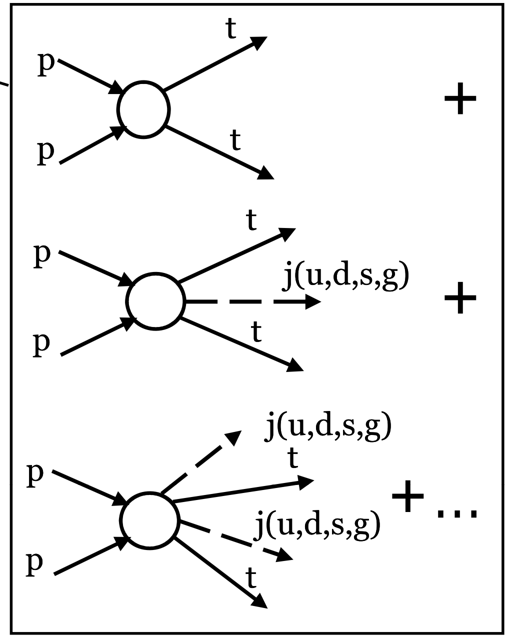
- pp→$bf_1f_2bf_3f_4$
- interferences properly included



✓ HLO with ME-PS matching: ALPGEN, MadGraph (new)

1st choice

- pp→ttNj, N=0,…3+; tt→$bf_1f_2bf_3f_4$; j=u,d,s,g
- matching scheme implemented
- better radiation description up to higher orders



✓ NLO: MC@NLO

- pp→tt at NLO ; tt→$bf_1f_2bf_3f_4$
- radiation correct at first order
- spin correlations added for all top final states

This is not a generator review !

Top pair generation comparison

Aim: test generators and validate their physics content (for top pair physics) in the frame of the CMS software. This will test them in the way we use them (debug !) and help understanding their features and make generation choices.

- Particular focus on the validation of MadGraph.
- MC@NLO still to come, work ongoing
- Other choices of generators also possible…

Input parameter settings as uniform as possible:

Parameter	TopRex	MadGraph	ALPGEN	Phantom
PDFs	CTEQ5L	CTEQ5L	CTEQ5L	CTEQ5L
Renormalization scale	m_T	m_t	m_T	m_T → $\sum_{tops}(m^2+p_T^2)$
Factorization scale	m_T	m_t	m_T	m_T
Top mass (Gev/c^2)	175	174.3	175	175
Λ_{QCD} (PARP(61), PARP(62)) (GeV)	0.25	0.25	0.25	0.25
Q^2_{max} switch (PARP(67))	2.5	2.5	2.5	2.5

Interface to CMS software via Les Houches file standard where possible.
Amount of statistics compatible with the CMS production allocated "bandwidth":

Generator	Production	Type	Interface to CMSSW	Events
TopRex	Offical CMS	inclusive $t\bar{t}$	Direct	$1,5\times10^6$
Alpgen	Official CMS	$t\bar{t}$ +Njets matched, N=0,...,4	Converted ROOT files	2×10^5
Phantom	Private	pp→$b\ell\nu$bqq'	LH files	5×10^5
MadGraph	Private	inclusive $t\bar{t}$	LH files	3×10^6
		$t\bar{t}$ +Njets matched, N=0,...,3	LH files	2×10^5

Conventions and summary

Final state fermions are taken from the generator information in the CMS software

Final state fermion four-momenta are retrieved with an algorithm that does not depends on mother-daughter information (no top quarks or W bosons in a six fermion code)

- exclude all quarks likely to be proton remnants (p_T and η cuts)
- find a b and a bbar (highest p_T ones if ambiguous)
- find a couple of fermions compatible with a W (no mass cut !)
- built two candidate top quarks associating W and b (no mass cut !)

Final state fermions are taken before they radiate FSR

- not too much physical, but factorises FSR contribution
- more sensitive to ISR and topSR

All distributions are normalised to unity

- just interested in shapes and not to absolute cross-section predictions
- look at most relevant variables: p_T, η of b, W and t, $p_T(tt)$, m_t, m_W, m_{tt}, ...

What will be presented here:

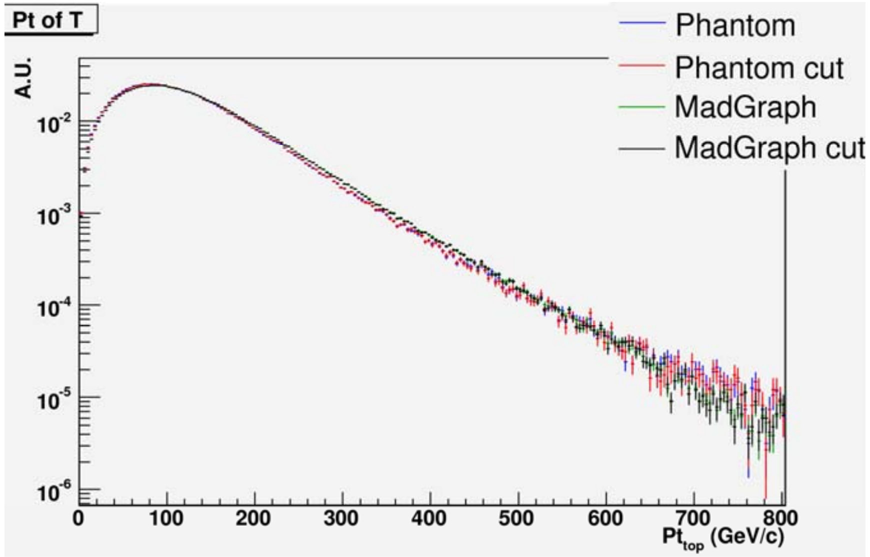
- standard vs six fermions codes
- standard vs matched codes, matched vs matched codes
- two words on new physics in tt
- what we learned and next steps to make good choices

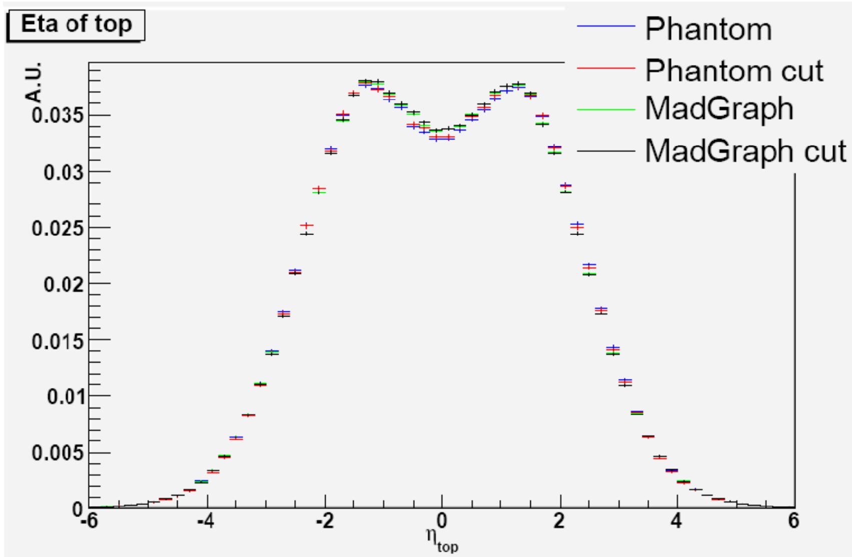
Six fermion effects

- Comparison to standard generations
- A word of caution on 6f and radiation

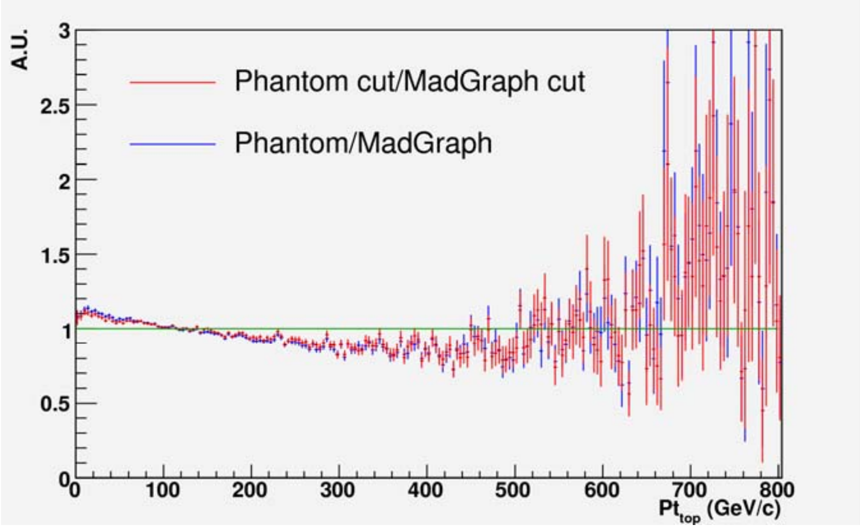
Probing the impact of six fermion codes

Probe distribution comparison on the mass shell (m_t-3Γ_t<m(bℓν,bqq)<m_t+3Γ_t) and outside.
No interference/extra diagram effects visible in the doubly resonant region.

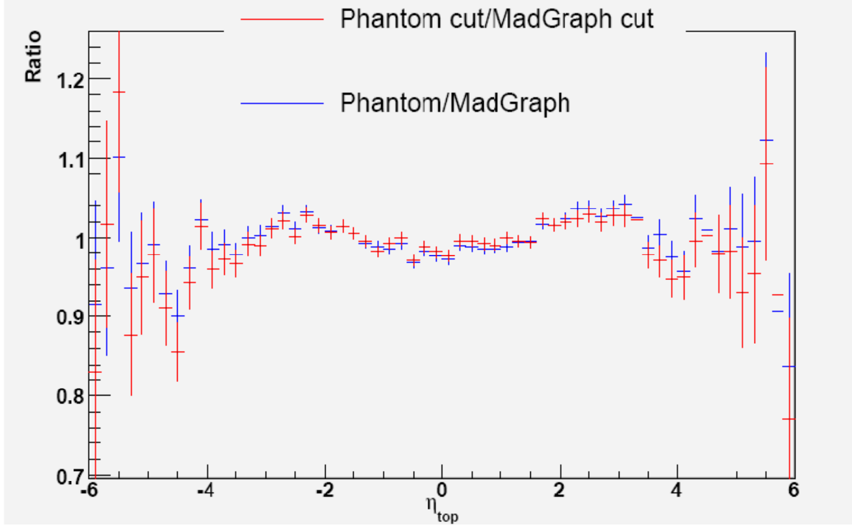




Cut : m_t-3Γ_t<m(bℓν,bqq)<m_t+3Γ_t

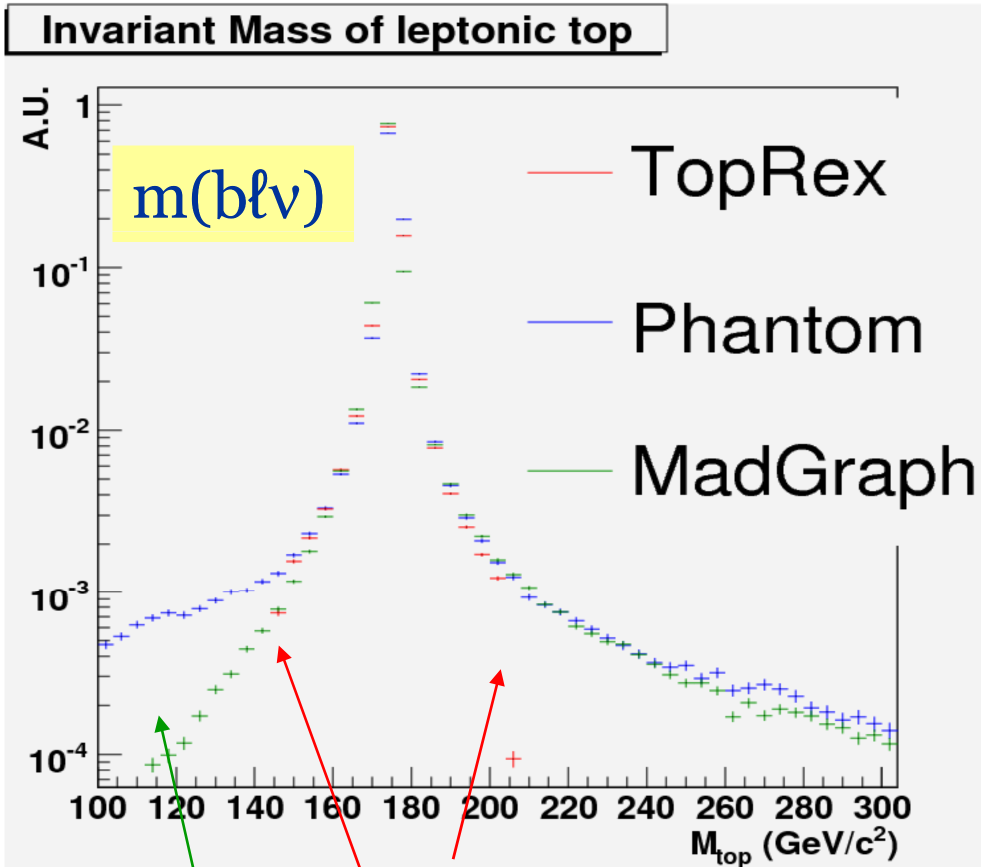


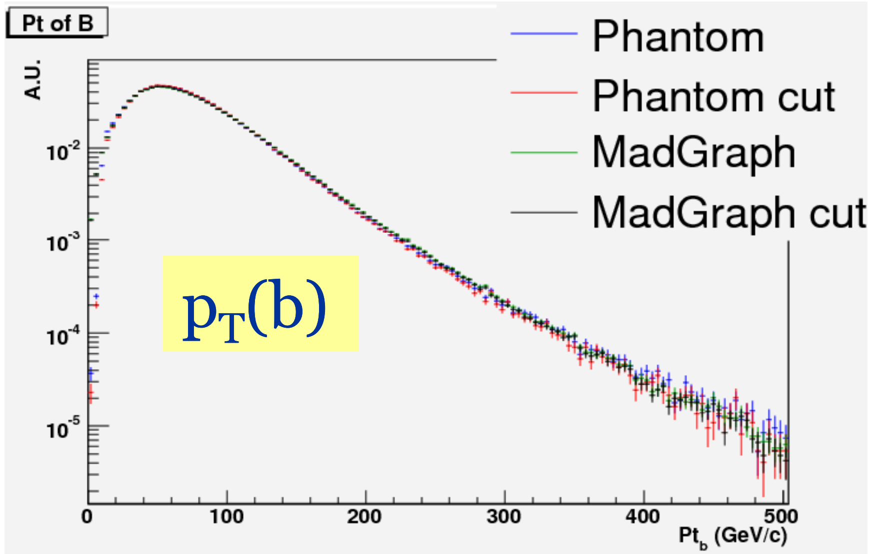
Cut : m_t-3Γ_t<m(bℓν,bqq)<m_t+3Γ_t



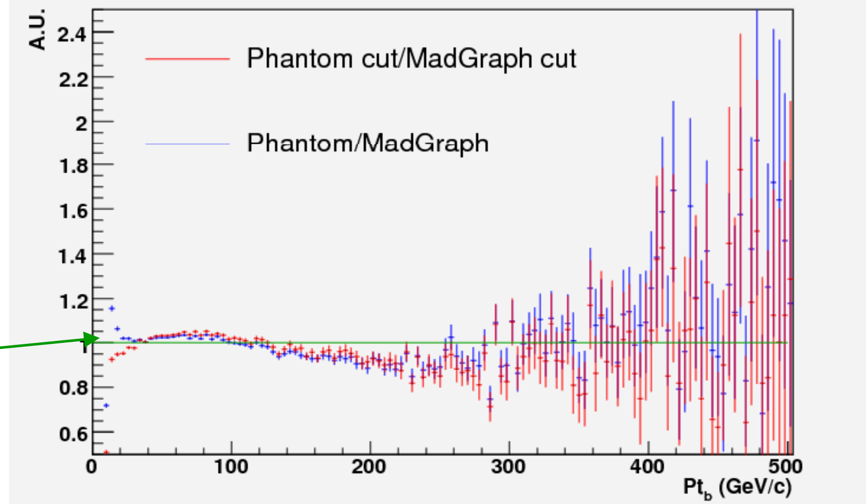
Six fermions effects

Effects visible for low p_T particles and very much outside the mass shell (as expected).
Might be important to use six fermion codes for off-shell analyses (ex: single top)





Cut : $m_t - 3\Gamma_t < m(b\ell\nu, bqq) < m_t + 3\Gamma_t$



cut

Presence of extra diagrams

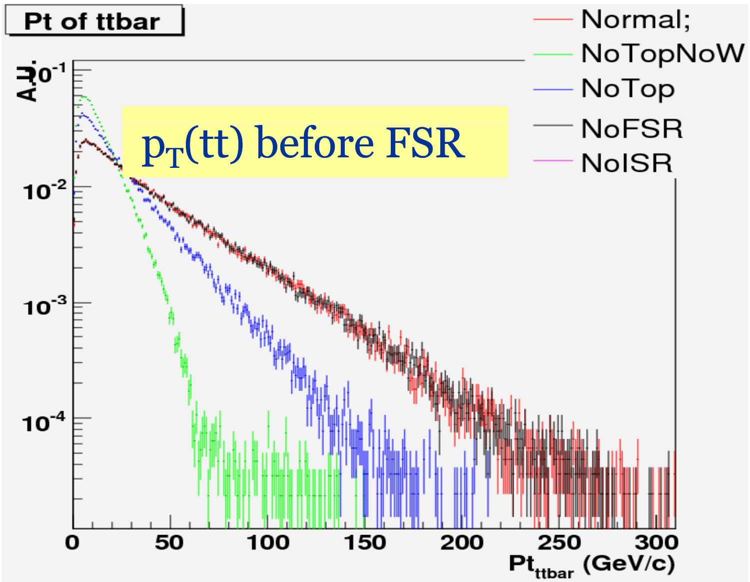
Roberto Chierici

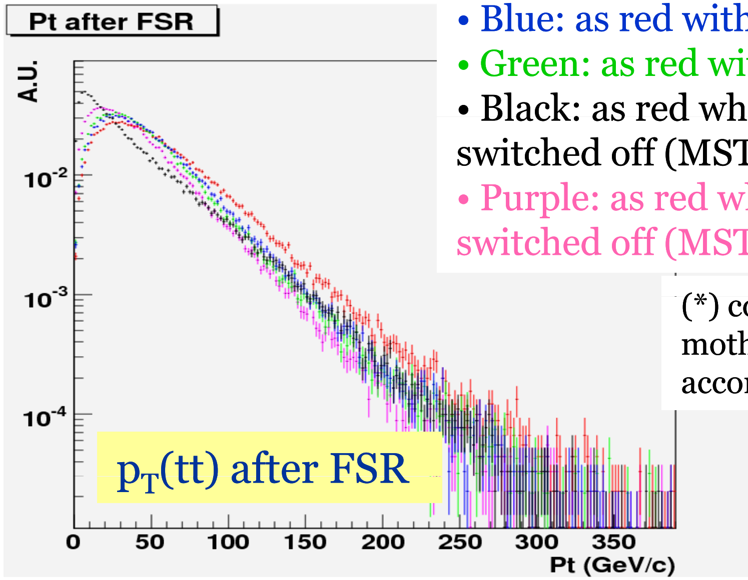
A word of caution

When comparing a 6f generator to standard ones huge discrepancies are seen (LH files +PYTHIA). The effect is explained by the fact that a 6f generator does not have intermediate particles like top and W. After PS radiation this makes huge differences:

1. the top themselves radiate before their decay; 2. the b radiate with the top mass preserved in the process (Ws receive recoils); 3. the quarks from the boosted W radiate, with the W mass preserved in the process; In 2. and 3. the showers match to NLO ME t→bgW and W→qgq'

Test MadGraph LHE files in several conditions before being interfaced to radiation:





- Red: standard LH with full radiation
- Blue: as red without top quarks*
- Green: as red without t and W*
- Black: as red where the FSR has been switched off (MSTP(71)=0)
- Purple: as red where ISR has been switched off (MSTP(61)=0)

(*) colour flow and the links mother-daughter are changed accordingly

- Any intermediate resonance changes the pattern of radiation
- The difference is in good part absorbed after full radiation (most likely because of the gluons in t→bgW and W→qgq' which are formally not FSR and counted systematically as topSR in the study

→ Even after full radiation there are sizeable differences on physics observables.

ME-PS matching

- The importance of matching in tt
- Matched generators vs standard

ME vs PS

No generator adequately reproduces the physics processes for the whole CMS program
Essential to understand which techniques are applicable to which kinematic regime.

- Parton Shower: infinite serie in α_S keeping only singular terms (collinear approx.):
  - ☞ Excellent at low p_T, with emission at any order, simple interface with hadronization
  - ☞ Large uncertainties away from singular regions
  - ☞ To be used for soft (compared to signal scale) jets.
- Fixed order matrix elements: truncated expansion in α_S
  - ☞ Full helicity structure to the given order
  - ☞ To be used for hard (compared to signal scale) jets.

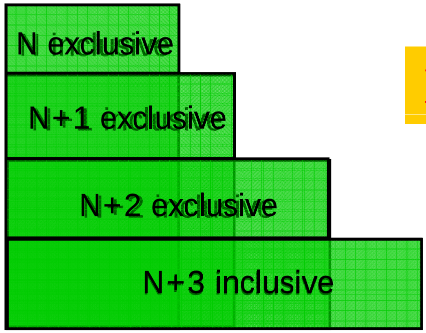
High jet multiplicity events are bound to be better described with ME. For top physics at the LHC this choice is mandatory. Especially for backgrounds.

A long-standing problem in MC generation: how to match PS and ME?

- Cutoff? Where? How to avoid double counting? (ME_N+PS has parts of ME_{N+1}+PS)

Techniques to match up to one additional hard jet exist in PYTHIA, HERWIG, MC@NLO

1. ME_N+PS
2. ME_{N+1}+PS
3. ME_{N+2}+PS
4. ME_{N+3}+PS



MLM

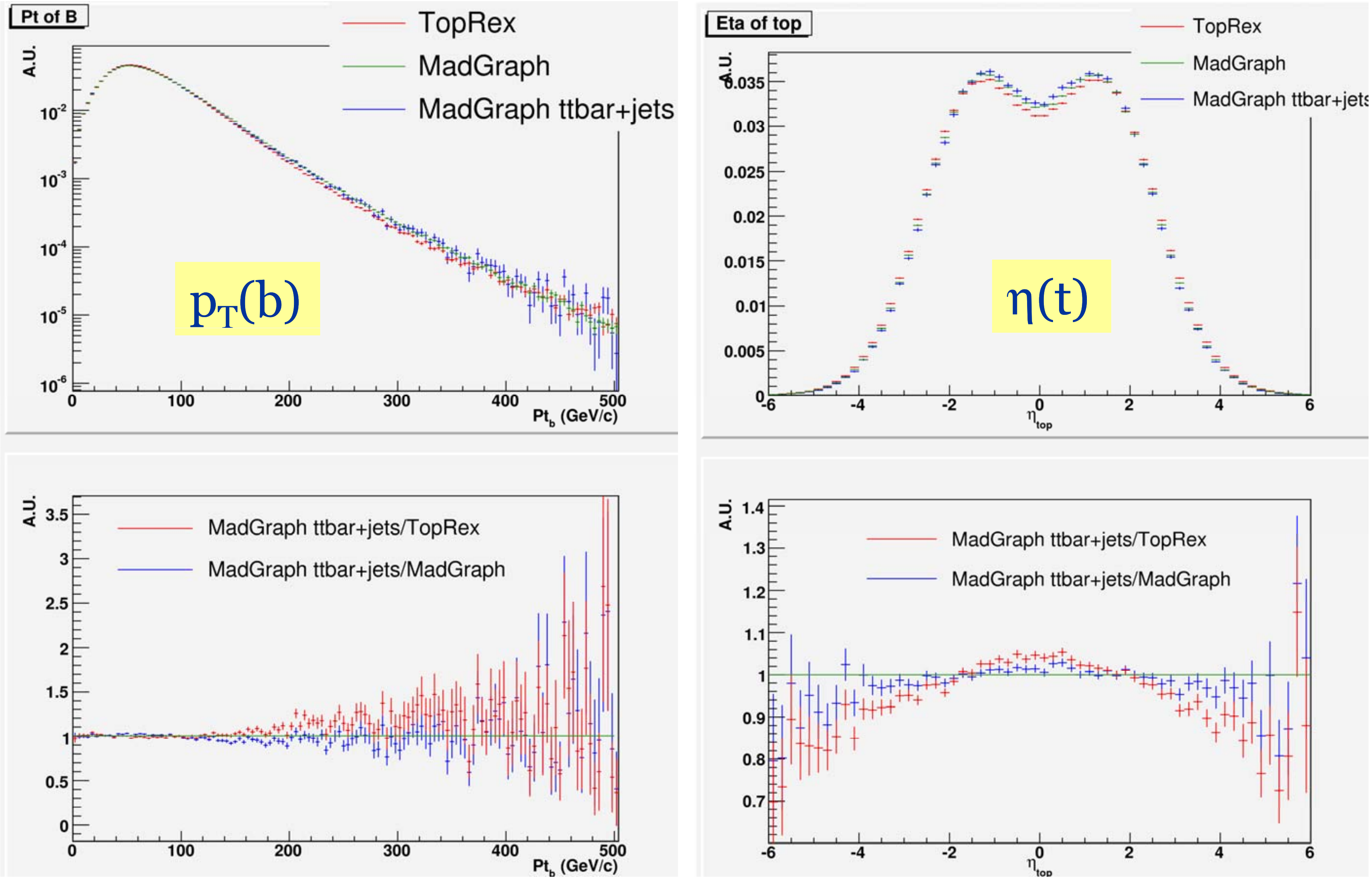
- (cone-)cluster showered event → njets
- match partons from the ME to the clustered jets
- if all partons are matched, keep event. Else discard it.

Works independent of the generation procedure

Beware: PS(today) ≠ PS(yesterday). Tunings need to adapt to the choice of the matching

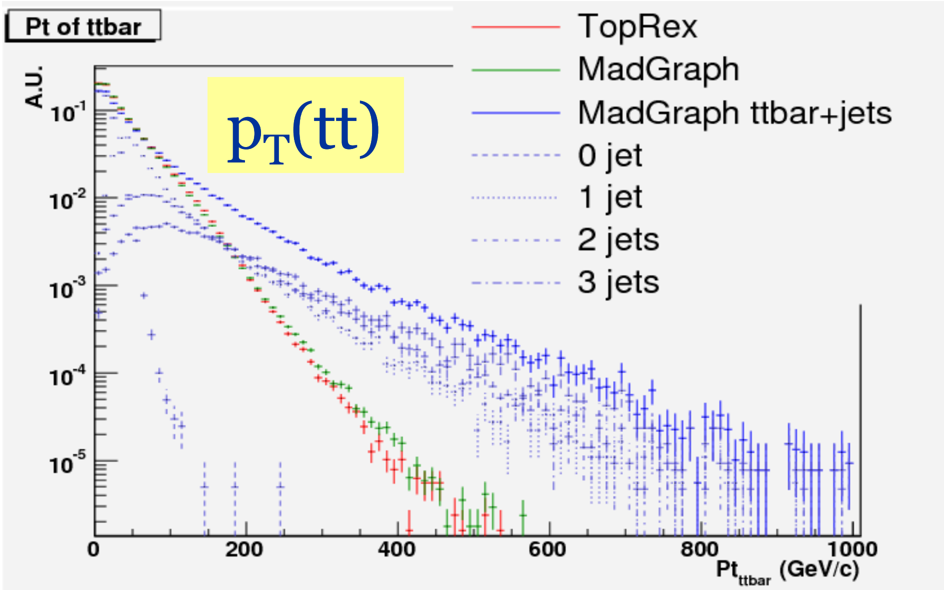
MadGraph matched vs standard generation

A matched tt+Njets generation gives no appreciable/small difference for single fermion quantities.



Matched ME-PS vs standard generation+PS

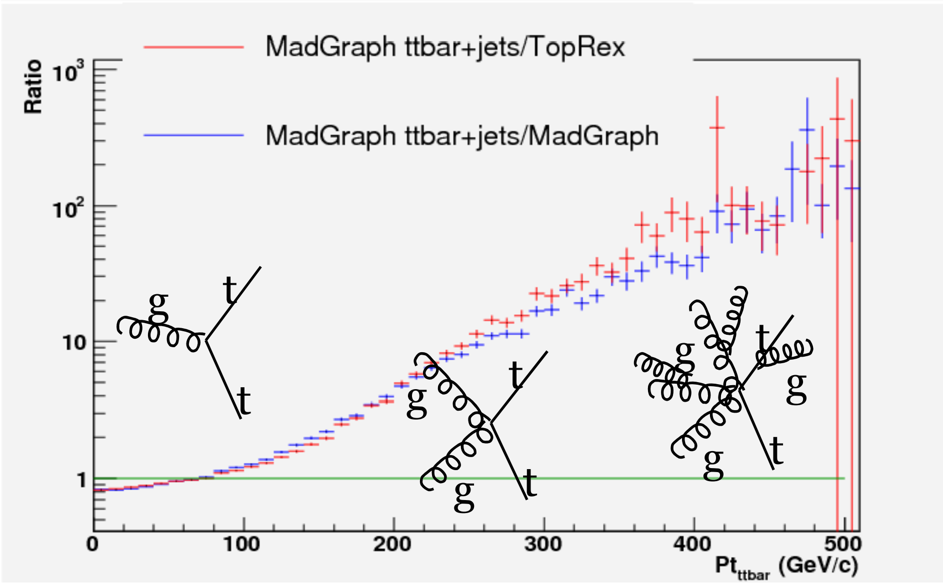
Spectacular differences in transverse variable connected to global radiation

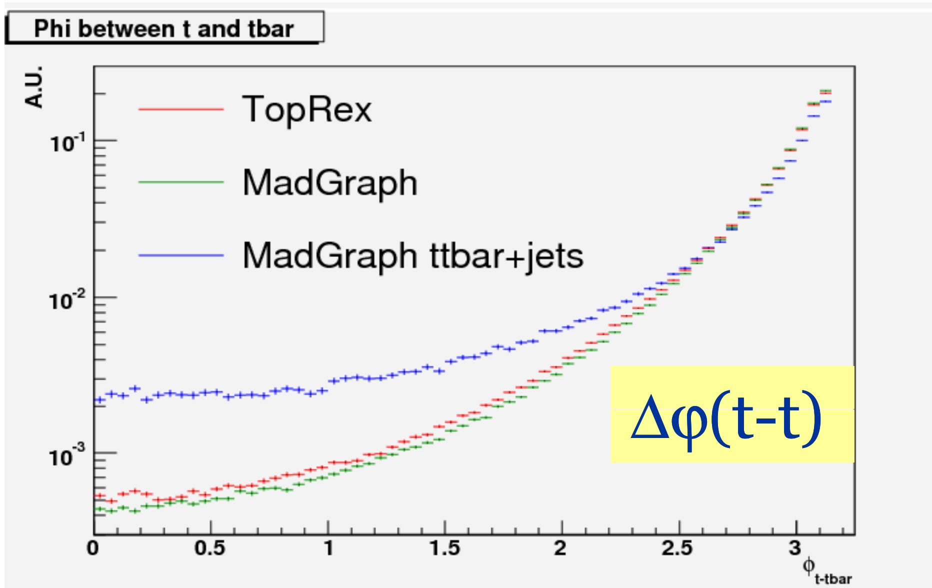


- Large effects at high $p_T(tt)=p_T(\text{radiation})$
- Average $p_T(tt)$~60-70 GeV !
- 40% probability that a tt system recoils against a radiation larger than 50 GeV

→ effect on reconstruction

→ Mandatory to use the same strategies for physics backgrounds like W/Z+Njets



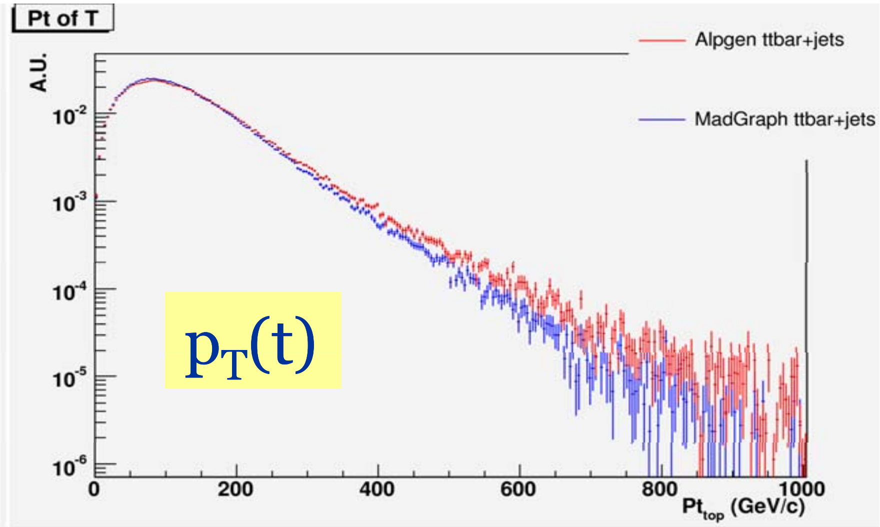


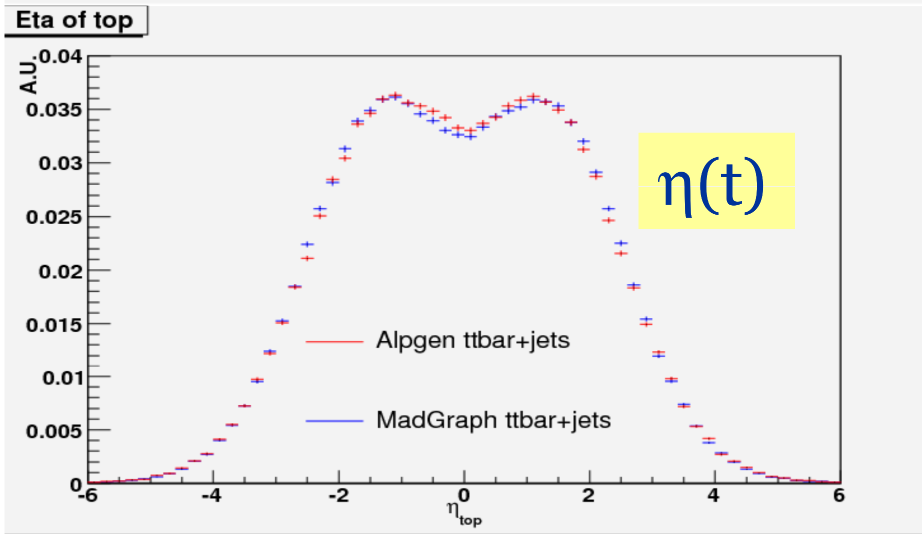
Matched vs matched

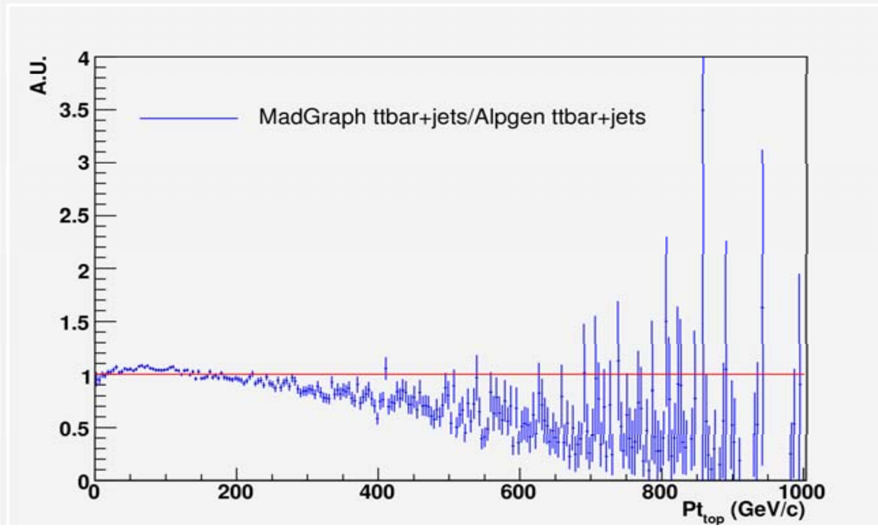
- MadGraph vs ALPGEN in tt
- towards the understanding of theory uncertainties

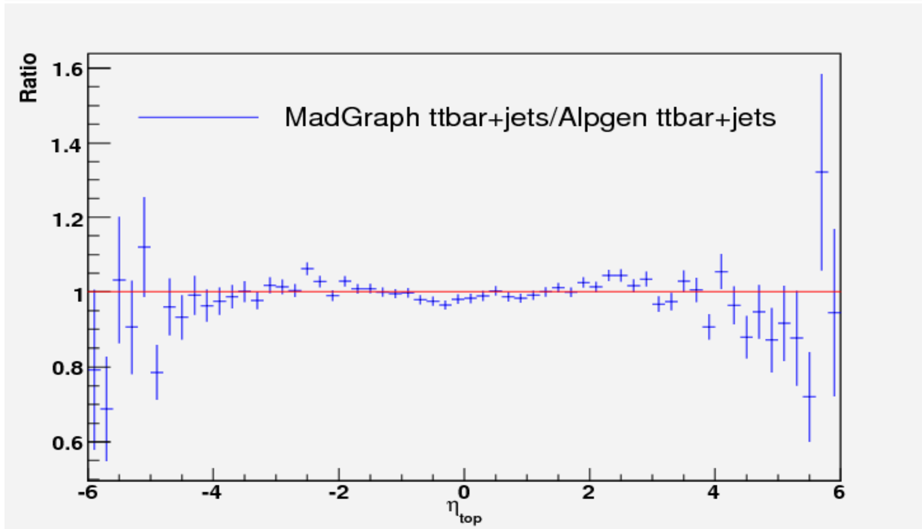
ALPGEN vs MadGraph matched

Good agreement in the description of top/final fermion kinematics
Residual differences due to different generator input settings and scales chosen?

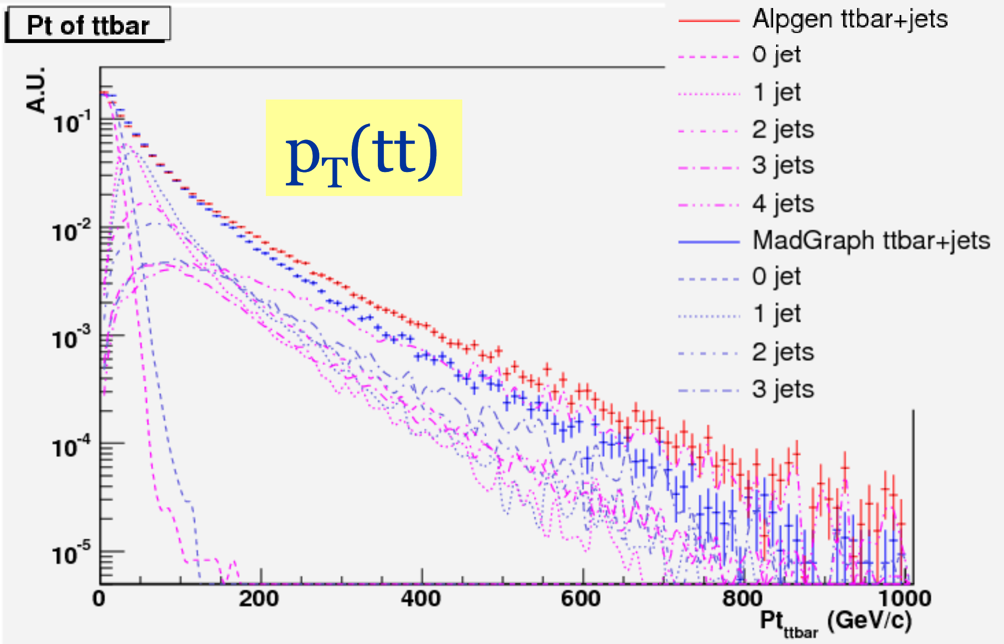








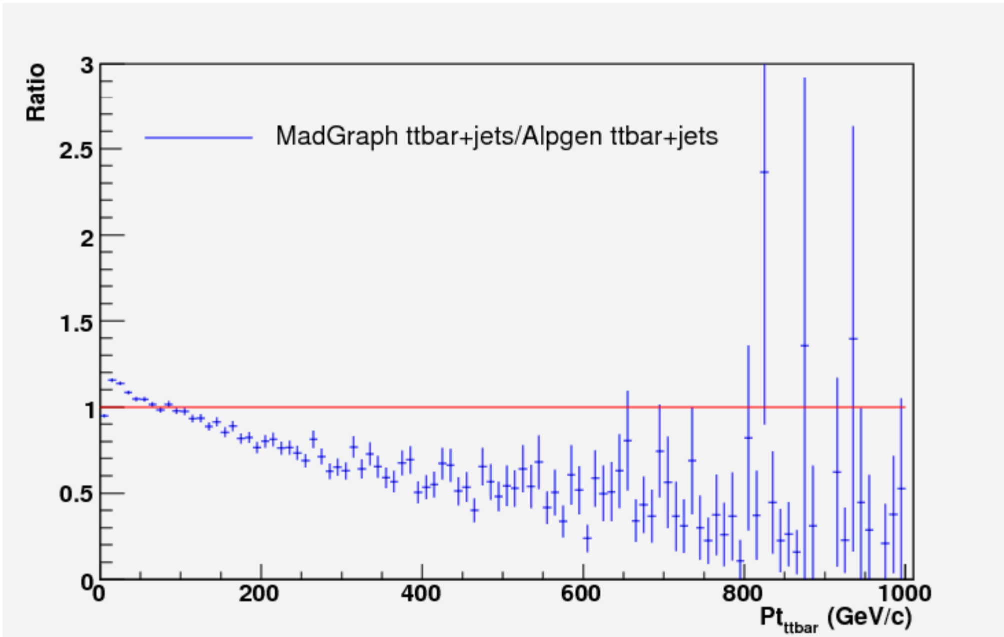
ALPGEN vs MadGraph matched

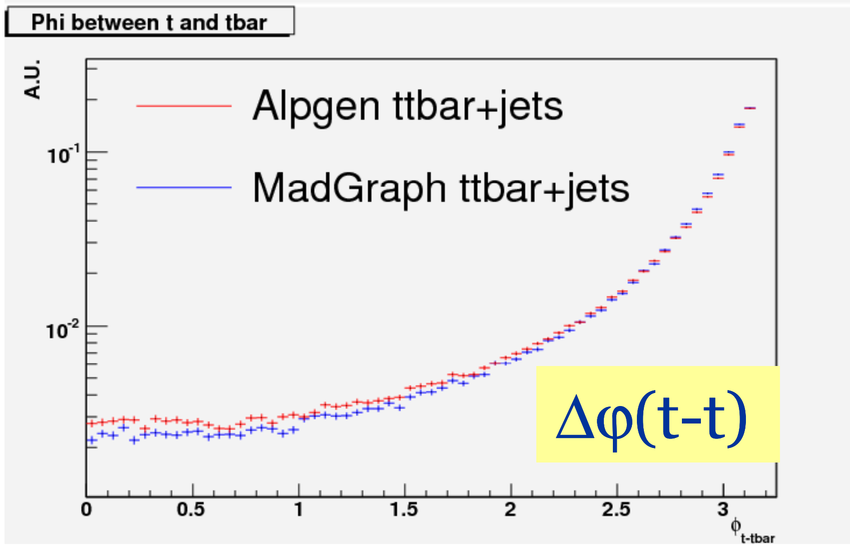


ALPGEN and MadGraph differ by at most 50% on the p_T prediction (several orders of magnitude away from the PS description)

Important to estimate these effects for the analyses:

- Effect of renormalisation and factorisation scales on the predictions
- Effect of the chosen ME-PS matching scale
- Comparison with MC@NLO





Concluding remarks

New physics in tt

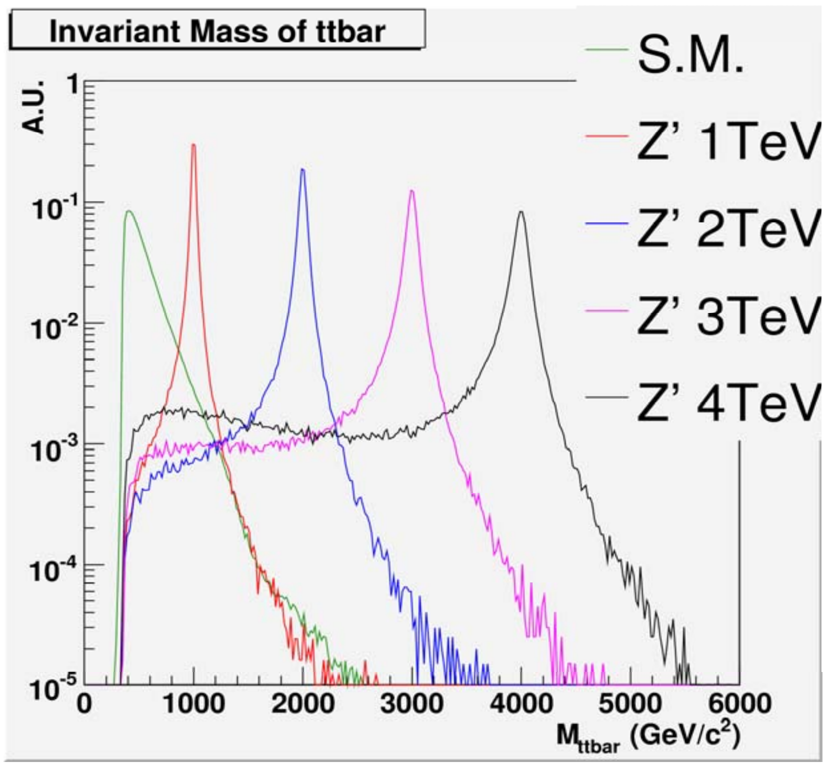
Having the possibility of including new physics signals (in the domain of top production) in the same generator framework used for the SM has big advantages

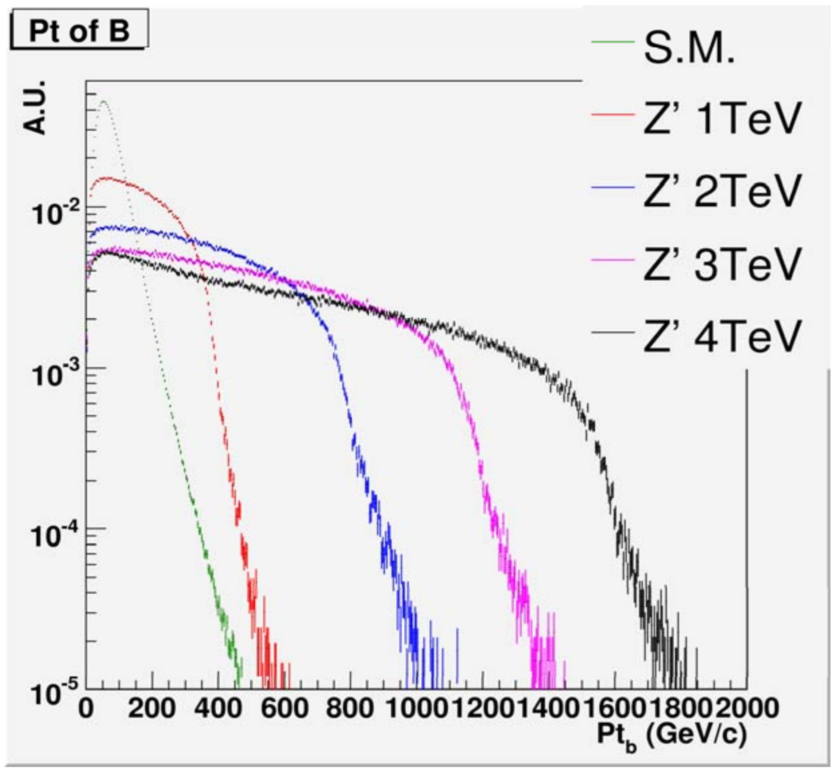
- same framework: cut setup, signal and background described in the same way,...
- interference SM-BSM can be included. Not always negligible !
- appealing perspective of a matching scheme in place also for new physics…

Top physics is an excellent window to new physics in production, association, decay.

Good experience with MadGraph, easy inclusion of any new physics

- Example: high mass Z-like resonances





Lessons learned and outlook

1. Extra diagrams contributions and SM interference effects can be neglected for the description of doubly resonant top physics. It is not so for low p_T, off-shell regions

2. The LH standard must also include equal treatment of intermediate resonances if we want to compare equivalent distributions after radiation. This poses a problem for six fermion generators.

3. Matched HLO generators are mandatory for high p_T physics. Top and not only.

4. ALPGEN and the new MadGraph matching provide a consistent picture of the high p_T top quark pair physics.

5. We are on the good way to take decision about the generation setup for top physics in CMS. Still there is a lot to do:
   - Need to compare with MC@NLO
   - Need to get a feeling of the impact of theory uncertainties
   - Need not to forget new physics
   - Need to attack the serious problem of generators tuning with data