(N)NLO MC as predictive tools

Main focus of $t\bar{t}H$ group: modelling of $t\bar{t}+X$ backgrounds

- MC fitted to data in control regions and extrapolated to signal regions
- crucial MC input for extrapolation: MC shape and its *theoretical* uncertainty

(N)NLO MC are predictive tools and should be used as such

- for a given *input* one (should) get a *well defined output* and *uncertainty*
- input parameters+settings should be carefully chosen and documented
- intrinsic MC uncertainties should be assessed is a coherent and systematic way (requires in-depth studies by theorists)
- different MC tools should agree within the respective uncertainties

Primary criterion to judge MC should be predictivity rather than data

Data driven approach (usual)

(1) select MC that yields best fit to data (ignoring intrinsic uncertainty)

(2) ad-hoc MC uncertainty for extrapolation (e.g. 2-point syst.)

Precision driven approach (preferable)

(1) select MC with well-defined and smallest intrinsic uncertainty

(2) further reduce MC uncertainty by fitting to data

Two-point systematics (i.e. MC$_i$ vs MC$_j$ as uncertainty)

Comparisons between NLOPS MCs

- MC$_i$–MC$_j$ or MC$_i$–NLO differences are often large and can spoil NLO accuracy
- understanding the origin of such differences (e.g. physical or unphysical?) is a first mandatory step towards a meaningful 2-point (MC$_i$–MC$_j$) uncertainties

$\Rightarrow$ $\sigma^{\mathrm{MC}}\left(\vec{X}_{\mathrm{MC}};\vec{Y}\right)$ should be compared with coherent choices+variations of parameters

MC-dependent paramaters $\vec{X}_{\mathrm{MC}}$

- include choice of matching method (e.g. MC@NLO vs Powheg), showers used for 1st and subsequent emissions, all related parameters and settings

MC-independent paramaters $\vec{X}$

- include perturbative input parameters, PDFs, QCD scales and also resummation scale μ_Q below which 1st emission unitarised via Sudakov FF (see next slide)

Parameters should be varied one-by-one around well defined $\vec{X}_0, \vec{Y}_{\mathrm{MC},0}$

- Example: compare Powheg+PY8 vs MC@NLO+PY8 using identical PY8, μ_Q, etc

$\Rightarrow$ difference can be attributed to matching method

MC@NLO vs Powheg matching (how to compare?)

Splitting of radiation: S-events (soft/singular) and H-events (hard/remnant)

$$\mathrm{d}\sigma_S = \mathrm{d}\Phi_B \bar{B}(\Phi_B)\left[\Delta(t_{\mathrm{IR}}) + \Delta(k_T)\frac{R_{\mathrm{soft}}(\Phi_R)}{B(\Phi_B)}\,\Phi_{\mathrm{rad}}\right] \qquad \mathrm{d}\sigma_H = \mathrm{d}\Phi_R\left[R(\Phi_R) - R_{\mathrm{soft}}(\Phi_R)\right]$$

Soft radiation integrated out in $\bar{B}$ $\quad\Rightarrow\quad$ $\bar{B}/B$ = local K-factor

$$\bar{B}(\Phi_B) = B(\Phi_B) + V(\Phi_B) + \int \mathrm{d}\Phi_{\mathrm{rad}} R_{\mathrm{soft}}(\Phi_B, \Phi_{\mathrm{rad}})$$

The only difference between Powheg and MC@NLO is in R_{soft}

Powheg:	$R_{\mathrm{soft}}(\Phi_R) = R(\Phi_R)\, g_{\mathrm{soft}}(\Phi_{\mathrm{rad}}, h_{\mathrm{damp}})$	matrix element
MC@NLO:	$R_{\mathrm{soft}}(\Phi_R) = B(\Phi_B) \otimes K_{\mathrm{shower}}(\Phi_{\mathrm{rad}})\, g_{\mathrm{soft}}(\Phi_{\mathrm{rad}}, \mu_Q)$	parton shower

Soft profile $g_{\mathrm{soft}}(\Phi_{\mathrm{rad}}, \mu_Q)$

- restricts R_{soft} below μ_Q (resummation scale), e.g. $\theta(\mu_Q^2 - k_T^2)$

$\Rightarrow$ **ideal choice for consistent comparison**: $h_{\mathrm{damp}} = \mu_Q$ and same g_{soft} $\quad$ **. . . ?**

Reproducibility issues

Predictivity implies reproducibility

- "for a given *input* one (should) get a *well defined output* and *uncertainty*"

⇒ however reproducibility is often not guaranteed by the highly complex and rapidly evolving MC frameworks

Challenges (especially for complex processes)

- significant differences in MC predictions can arise as a result of new MC features, samples generated by non-expert users, bugs, etc.
- Such large differences should be identified and understood (bugs or improvement?)
- This non-trivial and time consuming task is crucial for precision: if we claim 10% uncertainty today and MC changes by 30% tomorrow we have a problem
- within $t\bar{t}H$-WG to check reproducibility we have established MC benchmarks based on public Rivet analyses and well documented inputs+settings
- very useful to spot bugs, unwanted side-effects of new versions, user oversigths, ...and recommended to validate MC generation within ATLAS/CMS