Uncertainty
LHC Event Generators
and Pythia8

Stephen Mrenna
Fermilab*

October 30, 2023

*+ Pythia8 collaboration

Many sciences: highly mathematized
Particle physics: QFT: extremely (most) successful

High precision matching/merging calculations pQCD + PS

- Feynman diagrams (fixed order)
- Parton shower/resummation ("all" orders)

Key:

Know the pQCD expansion of the parton shower (PS)

Don't double count

Landmark: NNLO+PS matching

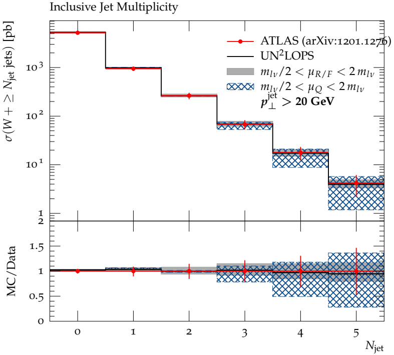
JHEP 1310 (2013) 222, arXiv:1405.3607, arXiv:1603.01620

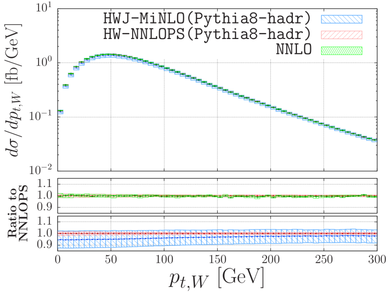
✓NNLO for incl. cross-section

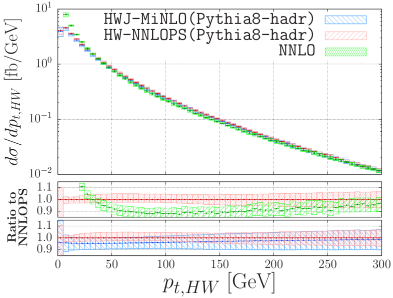
✓NLO+PS for 1-parton observables

✓LO+PS for 2-parton observables

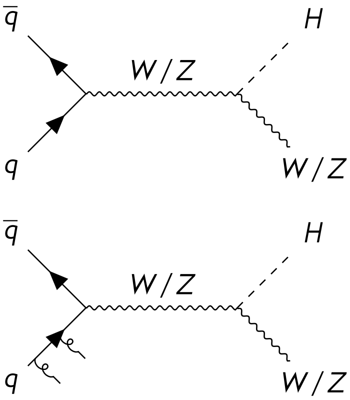
Two schemes: NNLOPS/MINLO and UN^{2}LOPS





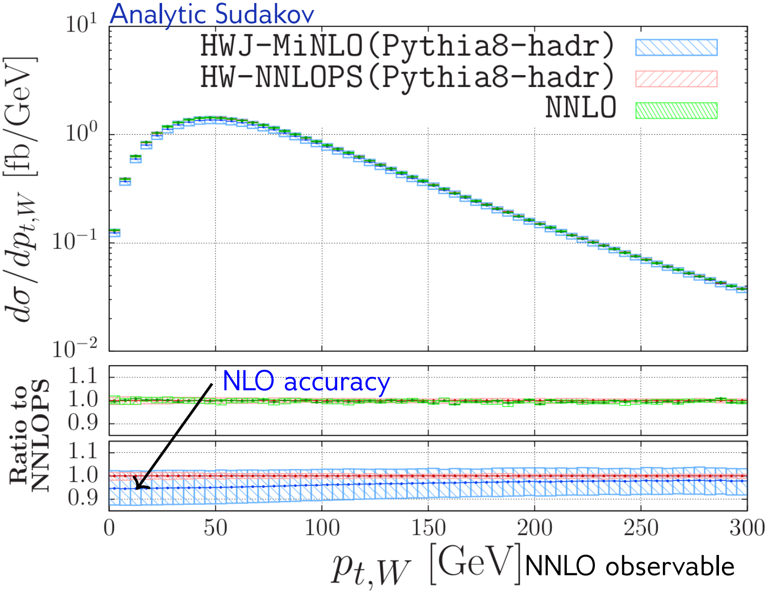


Some of the pieces

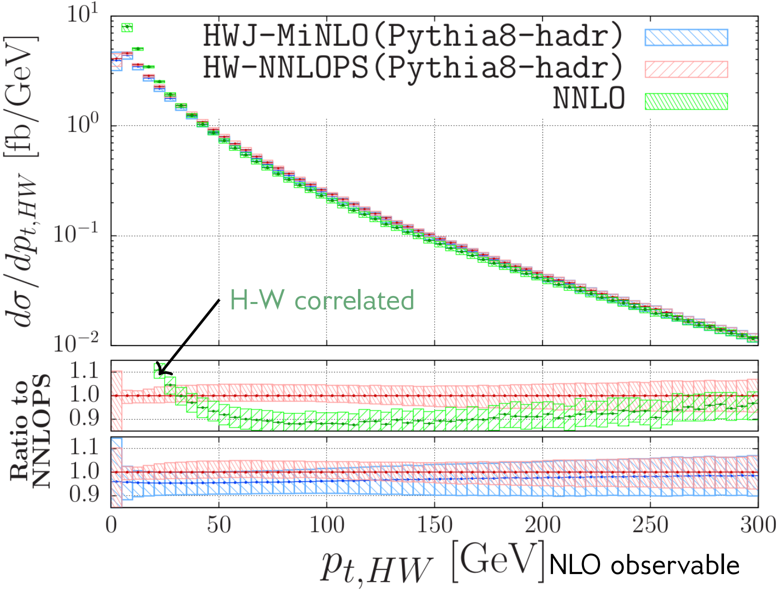


+ Loops

✓NNLO for incl. cross-section

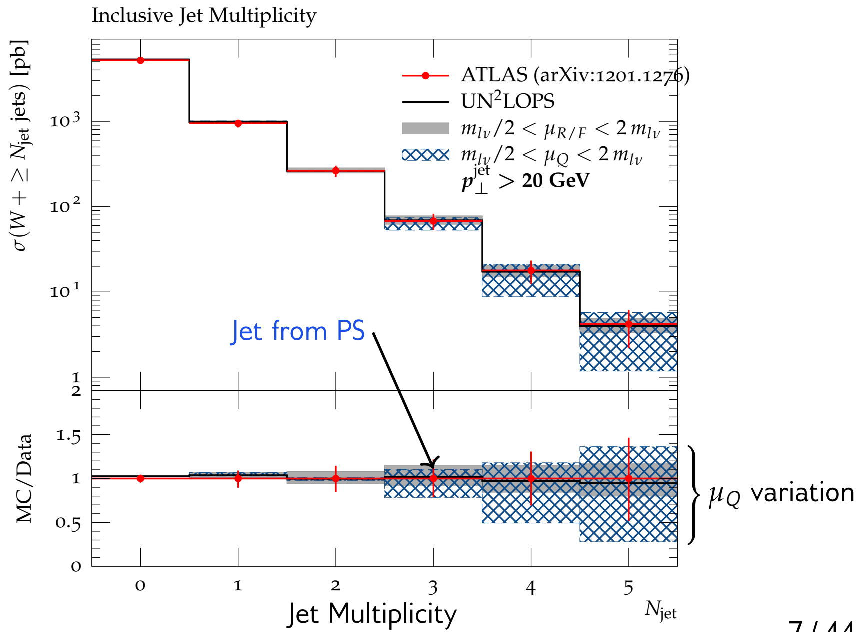


✓NLO+PS for 1-parton observables



NLO observable

✓LO+PS for 2-parton observables (W+jets)



The Bottom Line

High Level of sophistication in short distance, pQCD calculations and partons showers (PS)

More People work on the theoretically clean(er) aspects of perturbation theory

ME and PS are only a part of the story

Nature isn't limited to perturbation theory

Enter event generators

Why must uncertainties be small?

Looking for a chink in the SM armor

- Limits: tails (zero events is bad)
- Discovery: backgrounds after selection
- Measurement: aim for sub-percent in S and B!

Precision physics needs precision in tools!

After sophisticated pQCD+PS, we are left with:

Measurement	Relative error Due to MC	Reference
m_t	65%	JHEP 12 (2021) 161
$VV \to H$	75%	ATLAS-CONF-2020-026
$H \to XX$	50%	JHEP 12 (2020) 085

Uncertainties on What?

Our best "guess"

Focus on models in Pythia (and not pQCD)

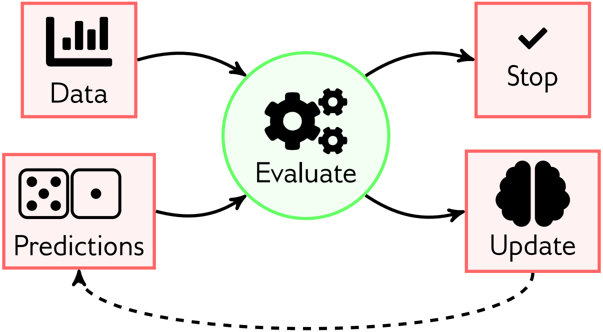
Most models are probabilistic

- Probabilities are about ignorance of causes
- Lack of understanding/inability to calculate

Don't expect perfection!

best $\to$ Tunes

Tuning



What Data?

- Will it have resolving power?
- Detector level? Costly simulations
- Unfolded? Loss of details, correlations?

All modern LHC tunes based on RIVET

RIVET: collection of experimental data, together with matching analysis routines. `10.21468/SciPostPhys.8.2.026`

Can be applied to generator events for comparison with data.

(Distorted) data unfolded/corrected to theory level.

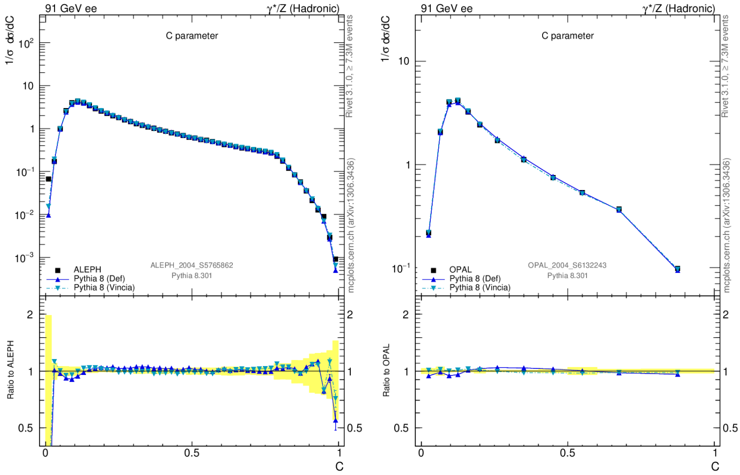
Currently, only (mostly) 1-D distributions

Recent PHYSTAT Seminar: Comparison of Unfolding methods

`https://indico.cern.ch/event/1328619/attachments/2731759/4767080/Unfolding_PhyStat.pdf`

MCPLOTS: `http://mcplots.cern.ch/`

Repository of comparisons between various tunes and data



Part of the LHC@home platform for home computer participation

⚄⚀ : What Predictions? What Parameters?

- Many parameters! (More on this)
- Divide (and conquer) based on physics
    - LEP physics does not depend (mostly) on breakup of proton
        - hadronization, final state radiation (FSR)
    - p_T^Z : initial state radiation (ISR) and intrinsic k_T
        - but, $q\bar{q}$ initial state
    - Underlying Event $\sim$ independent of hard process: MPI models
- Still, sample a DISCRETE set of parameters
    - How to interpolate in high-dim space?
    - Surrogate Model: Professor, Apprentice, or a ML task

Evaluation?

$$\min_{\lambda} \chi^2(\lambda) := \sum w_i \left(\frac{f_i(\lambda) - D_i}{\sqrt{f_i(\lambda)^2 + D_i^2}} \right)^2$$

D_i data

$f_i(\lambda)$ physical model

(many) parameters λ

$w_i > 0$ are fixed to emphasize certain datasets D over others (e.g. to ensure that regions with fewer measurements are matched on average as well as regions with more measurements).

Challenges

No analytic expressions for $f_i(\lambda)$, and each evaluation is computationally expensive

Derivatives of $f_i(\lambda)$ cannot be readily computed by the simulation, though the functions are probably smooth

Several local and global minima may be present (multi-modality)

Different measurement regimes have different observational densities

Most approaches build analytical surrogate models and known gradients to find descent directions for χ^2.

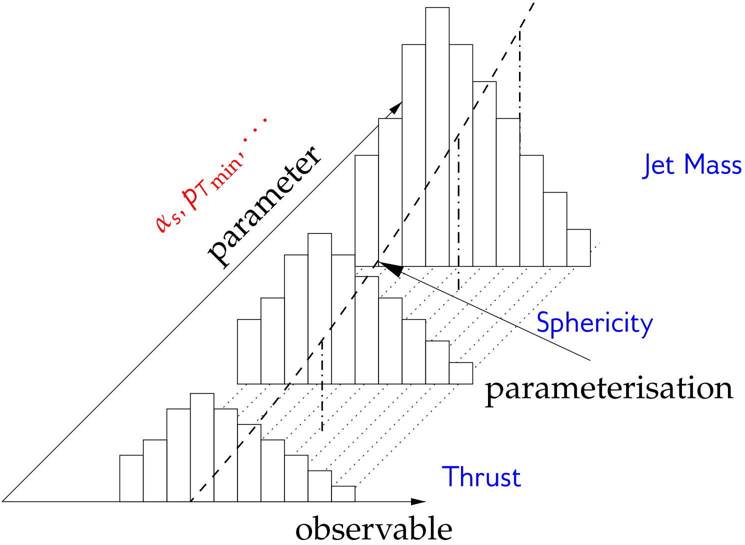
Systematic tuning (vs Brute Force vs Expert)

PROFESSOR: parameter tuning in n-dimensional parameter space

https://professor.hepforge.org

- Generate large event samples at $\mathcal{O}(n^2)$ *random* points. Slow!
- Analyze events (MC truth) and fill relevant histograms.
- Parametrize physics-parameter response in each bin of each histogram
- Minimize χ^2 on parametrized results. Fast!

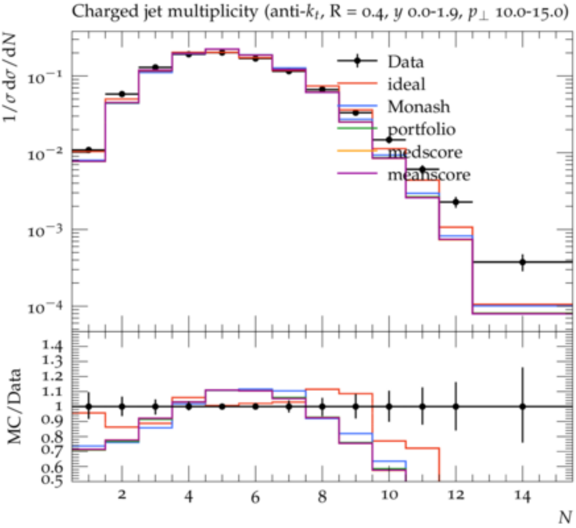
  {Cross check on exact simulation with same parameters}



$$X_{MC} = A_0 + \sum_{i=1}^{n} B_i p_i + \sum_{i=1}^{n} C_i p_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} D_{ij} p_i p_j + \cdots$$

Utopian tunes

Fit one histogram at a time ($N_{\text{bins}} \sim N_{\text{param}}$)



Sometimes: a gap between the prediction and data

Uncertainty Quantification $\neq$ Constraining parameters of models

- Need to challenge models
- models are incomplete (see Kendall's talk)
  "Franken-models"
  "paper over problems"

LHC community might think we are more immune to this that we really are

Systematic Uncertainty on Models

Models good but $\neq$ truth

can affect unfolding

does affect tuning (fit bulk or tail?)

pQCD gives some estimate of the size of neglected orders

What is this for models?

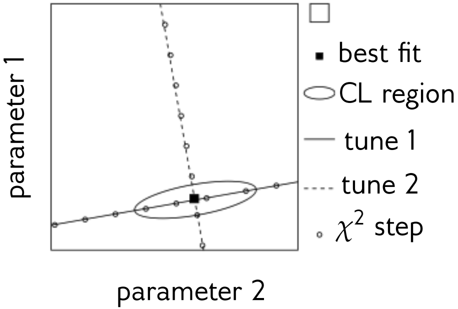
- Alternative 1: Skands flat 5% per bin
- Alternative 2: Add in physics-motivated component
  - constant pieces to parton shower not captured by leading logs
- Alternative 3: Build a model of the gap (ML?)

Uncertainty (finally)

≡ What are the variations around our best tunes such that we can do precision physics *and* not get fooled

Meaningful variation around χ^2_{min}

- Theory-informed groupings
  - e.g. Perugia tunes
- algorithmic definition of parameter groupings
  - Eigentunes (Professor)



Eigentunes: ATLAS A14 Tune LHC

FSR/ISR/MPI tune

10 parameters → 20 independent variations

"too unwieldy"

"by inspection reduced it to a subset"

Param	+ variation	− variation
VAR1: MPI+CR (UE activity and incl jet shapes)		
`BeamRemnants:reconnectRange`	1.73	1.69
`MultipartonInteractions:alphaSvalue`	0.131	0.121
VAR2: ISR/FSR (jet shapes and substructure)		
`SpaceShower:pT0Ref`	1.60	1.50
`SpaceShower:pTdampFudge`	1.04	1.08
`TimeShower:alphaSvalue`	0.139	0.111
VAR3a: ISR/FSR ($t\bar{t}$ gap)		
`MultipartonInteractions:alphaSvalue`	0.125	0.127
`SpaceShower:pT0Ref`	1.67	1.51
`SpaceShower:pTdampFudge`	1.36	0.93
`SpaceShower:pTmaxFudge`	0.98	0.88
`TimeShower:alphaSvalue`	0.136	0.124
VAR3b: ISR/FSR (jet 3/2 ratio)		
`SpaceShower:alphaSvalue`	0.129	0.126
`SpaceShower:pTdampFudge`	1.04	1.07
`SpaceShower:pTmaxFudge`	1.00	0.83
`TimeShower:alphaSvalue`	0.114	0.138
VAR3c: ISR ($t\bar{t}$ gap, dijet decorrelation and Z-boson p_T)		
`SpaceShower:alphaSvalue`	0.140	0.115

A14 and weighting of data

- w_i in χ^2 calculation adjusted "by eye"
- Apprentice developed to automate this
    - Portfolio (payoff vs risk)
    - Variance reduction
    - . . .

10.1051/epjconf/202125103060

SciDAC

Math scientists at ANL, LBNL

The Perugia Tunes

- Perugia 0 (320): Baseline: p_T^Z $\langle p_\perp \rangle \, (N_{\rm ch})$, high-$p_\perp$ tail of charged particle $p_\perp$ spectra.
- Perugia HARD (321): pert $\uparrow$: non-pert $\downarrow$
  less “primordial k_T”, higher IR cutoff for MPI, more active color reconnections :: higher curve for $\langle p_\perp \rangle \, (N_{\rm ch})$,
- Perugia SOFT (322): pert $\downarrow$: non-pert $\uparrow$
  lower IR cutoff for MPI, less color reconnections, :: lower curve for $\langle p_\perp \rangle \, (N_{\rm ch})$.
- Perugia 3 (323): Different balance between MPI and ISR.
  Additional ISR activity balanced by a higher infrared MPI cutoff.
- Perugia NOCR (324): no color reconnections :: acceptable agreement with most distributions, except for the $\langle p_\perp \rangle \, (N_{\rm ch})$.
- Perugia K (328): K factor on the QCD $2 \to 2$ scattering cross sections used for MPIs, UE more “jetty”, higher $p_\perp$ :: not a good fit, but interesting

A large number of parameters need to be fit (tuned) to data to make predictions

Do we just over-fit data?

Actual parameters $<$ # Apparent parameters

Some parameters are just necessary choices (cutoffs, PDF, *etc*)

Sensitive parameters $<$ # Actual parameters

Information in differential data $\gg$ # Sensitive parameters

Different energies, different incoming beams

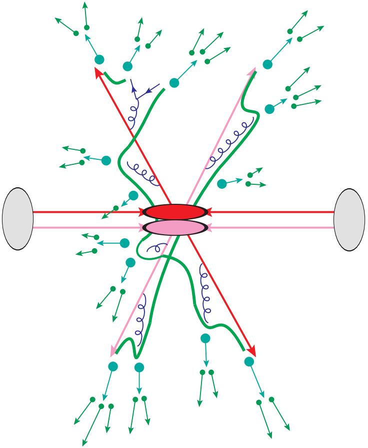
Note: Perturbative, inclusive QCD calculations also have inputs: α_s, μ_R, μ_F, PDFs

My guess: Minimal # parameters for LHC $\sim 15 - 20$

Depends on what question you are asking

What is Pythia8?

A machine that generates realistic, high-energy particle collision event records using all the Standard Model physics we know and inspired models to fill in the rest.



Many diverse phenomenom

Short-distance cross section:
μ_r^H, μ_f^H, PDF^H, α_s^H

Parton shower:
μ_q^{PS}, μ_r^{PS}, μ_f^{PS}, μ_{cut}^{PS}, PDF^{PS}, α_s^{PS}

Multiple interactions:
μ_q^{MPI}, PDF^{MPI}, $\alpha_s^{MPI}\ldots$

String fragmentation:
$f(z)$, string p_T, tension

Beams:
Primordial k_T, remnants

Particle Decays:
BRs, MEs, BE correlations

...

Parameter Tuning for LHC Physics

Divide and Conquer

- Start with data from electron-positron colliders
- Removes complication of the beams
  - No Initial State Radiation
  - No MPI
- Know collision energy E_{cm}
- $\cdots$ but limited range
- Most data @ $E_{\text{cm}} = M_Z \sim 91\,\text{GeV}$

Monash 2013 Tune Parameters

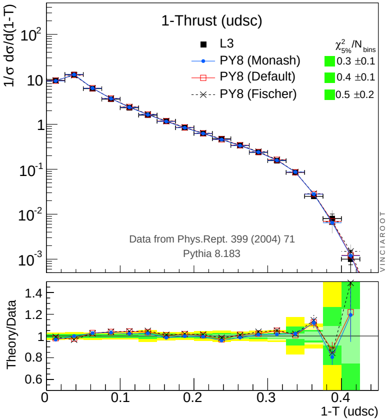
Final-state radiation (FSR) parameters.

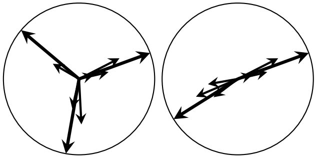
FSR Parameters	**Monash 13**	**(Default)**	**Comment**
TimeShower:alphaSvalue	= 0.1365	= 0.1383	! Effective alphaS(mZ) value
TimeShower:alphaSorder	= 1	= 1	! Running order
TimeShower:alphaSuseCMW	= off	= off	! Translation from $\overline{\text{MS}}$ to CMW
TimeShower:pTmin	= 0.50	= 0.40	! Cutoff for QCD radiation
TimeShower:pTminChgQ	= 0.50	= 0.40	! Cutoff for QED radiation
TimeShower:phiPolAsym	= on	= on	! Asymmetric azimuth distributions

$$\text{FSR:} \quad \mu_R^2 = p_{\perp\text{evol}}^2 = z(1-z)Q^2 \ ,$$

with $Q^2 = p^2 - m_0^2$ the offshellness of the emitting parton (with on-shell mass m_0), and z the energy fraction appearing in the DGLAP splitting kernels, $P(z)$.

pTmin :: avoid Landau pole





$$\alpha_s(M_Z) = 0.1365 \neq 0.13$$

$$\text{CMW} \neq \overline{\text{MS}}$$

$$\Lambda_{\text{CMW}} = 1.6\Lambda_{\overline{\text{MS}}}$$

Hadronic Z decays at $\sqrt{s} = 91.2\,\text{GeV}$. The Thrust distribution in light-flavor tagged events, compared with L3 data

---

Choice of cutoff, hadronization parameters also impact prediction

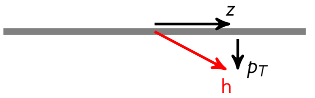
Relies on ME (3j) correction to PS
Could describe T but not jet structure

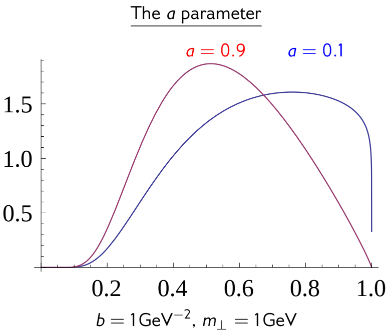
String-breaking parameters.

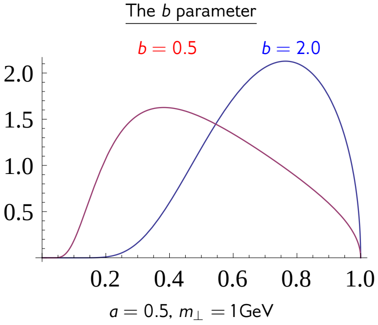
HAD Parameters	Monash 13	(Default)	Comment
# String breaks: pT and z distributions			
StringPT:sigma	= 0.335	= 0.304	! Soft pT in string breaks (in GeV)
StringPT:enhancedFraction	= 0.01	= 0.01	! Fraction of breakups with enhanced pT
StringPT:enhancedWidth	= 2.0	= 2.0	! Enhancement factor
StringZ:aLund	= 0.68	= 0.3	! Lund FF a (hard fragmentation supp)
StringZ:bLund	= 0.98	= 0.8	! Lund FF b (soft fragmentation supp)
StringZ:aExtraSquark	= 0.0	= 0.0	! Extra a when picking up an s quark
StringZ:aExtraDiquark	= 0.97	= 0.50	! Extra a when picking up a diquark
StringZ:rFactC	= 1.32	= 1.00	! Lund-Bowler c-quark parameter
StringZ:rFactB	= 0.855	= 0.67	! Lund-Bowler b-quark parameter
# Flavour composition: mesons			
StringFlav:ProbStoUD	= 0.217	= 0.19	! Strangeness-to-UD ratio
StringFlav:mesonUDvector	= 0.5	= 0.62	! Light-flavour vector suppression
StringFlav:mesonSvector	= 0.55	= 0.725	! Strange vector suppression
StringFlav:mesonCvector	= 0.88	= 1.06	! Charm vector suppression
StringFlav:mesonBvector	= 2.2	= 3.0	! Bottom vector suppression
StringFlav:etaSup	= 0.60	= 0.63	! Suppression of eta mesons
StringFlav:etaPrimeSup	= 0.12	= 0.12	! Suppression of eta' mesons
# Flavour composition: baryons			
StringFlav:probQQtoQ	= 0.081	= 0.09	! Diquark rate (for baryon production)
StringFlav:probSQtoQQ	= 0.915	= 1.000	! Strange-diquark suppression
StringFlav:probQQ1toQQ0	= 0.0275	= 0.027	! Vector diquark suppression
StringFlav:decupletSup	= 1.0	= 1.0	! Spin-3/2 baryon suppression
StringFlav:suppressLeadingB	= off	= off	! Optional leading-baryon suppression
StringFlav:popcornSpair	= 0.9	= 0.5	!
StringFlav:popcornSmeson	= 0.5	= 0.5	!

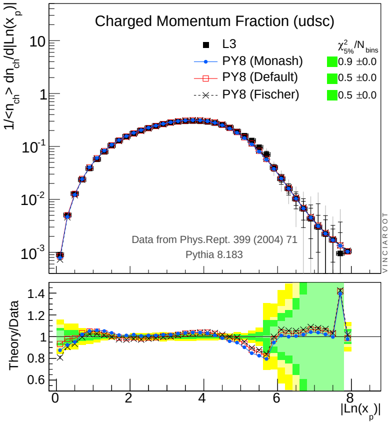
Lund symmetric fragmentation function:

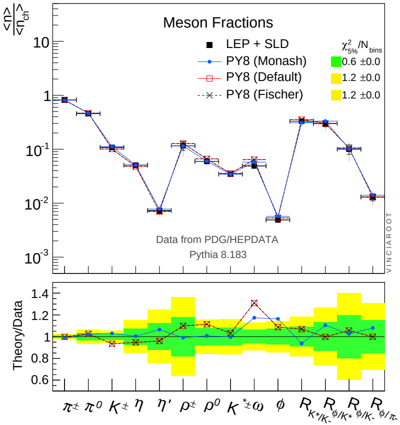
$$f(z) \propto \frac{(1-z)^a}{z} \exp\left(\frac{-bm_\perp^2}{z}\right)$$












```
StringZ:aLund          =  0.68
StringZ:bLund          =  0.98

StringZ:aExtraDiquark =  0.97
StringZ:aExtraSquark  =  0.00
```

```
StringFlav:ProbStoUD      = 0.217
StringFlav:mesonUDvector = 0.5
StringFlav:mesonSvector  = 0.55
StringFlav:etaSup        = 0.60
StringFlav:etaPrimeSup   = 0.12
```

Universality Assumption

Final state effects (FSR, hadronization) are tuned in the relatively clean e^+e^- environment

- Problematic, because kinematics (energy range) is limited, e.g. not much $g \rightarrow Q\overline{Q}$ splitting

Assume: Any differences in hadron collisions come from environment

- Color flow from initial to final state
- High density of colored partons

Must/should test this!

New parameters (models) for proton (nucleon) beams

Initial-state radiation (ISR) and primordial-k_T parameters.

ISR Parameters	Monash 13	(Default)	Comment
SpaceShower:alphaSvalue	= 0.1365	= 0.137	! Effective alphaS(mZ) value
SpaceShower:alphaSorder	= 1	= 1	! Running order
SpaceShower:alphaSuseCMW	= off	= off	! Translation from $\overline{\mathrm{MS}}$ to CMW
SpaceShower:samePTasMPI	= off	= off	! ISR cutoff type
SpaceShower:pT0Ref	= 2.0	= 2.0	! ISR pT0 cutoff
SpaceShower:ecmRef	= 7000.0	= 1800.0	! ISR pT0 reference ECM scale
SpaceShower:ecmPow	= 0.0	= 0.0	! ISR pT0 scaling power
SpaceShower:rapidityOrder	= on	= on	! Approx coherence via y-ordering
SpaceShower:phiPolAsym	= on	= on	! Azimuth asymmetries from gluon pol
SpaceShower:phiIntAsym	= on	= on	! Azimuth asymmetries from interference
TimeShower:dampenBeamRecoil	= on	= on	! Recoil dampening in final-initial dipoles
BeamRemnants:primordialKTsoft	= 0.9	= 0.5	! Primordial kT for soft procs
BeamRemnants:primordialKThard	= 1.8	= 2.0	! Primordial kT for hard procs
BeamRemnants:halfScaleForKT	= 1.5	= 1.0	! Primordial kT soft/hard boundary
BeamRemnants:halfMassForKT	= 1.0	= 1.0	! Primordial kT soft/hard mass boundary

Plus parameters for MPI, partonic structure of proton, $\cdots$

Evolution of Pythia8 Tune in CMS Run2

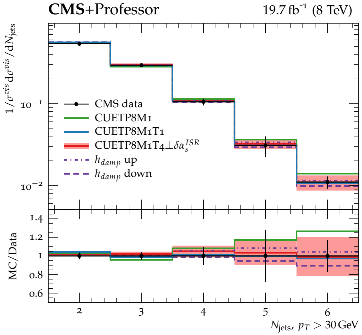
Monash (by-hand) tune was a good starting point, but new data came out afterwards at many different energies

- CMS used Tune CUETP8M1 for most samples (arXiv:1512.00815)
  - Re-tuning of UE parameters on top of Monash, α_S and other shower parameters left untouched
  - In particular this means α_S=0.1365 used for both ISR and FSR in the shower, despite using 0.118 in the ME for NLO samples and 0.130 for LO samples
- Tuning of shower parameters in particular revisited

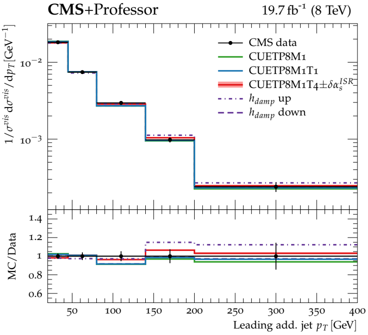
Pythia8 Shower Tuning with $t\bar{t}$ (CMS-PAS-TOP-16-021)

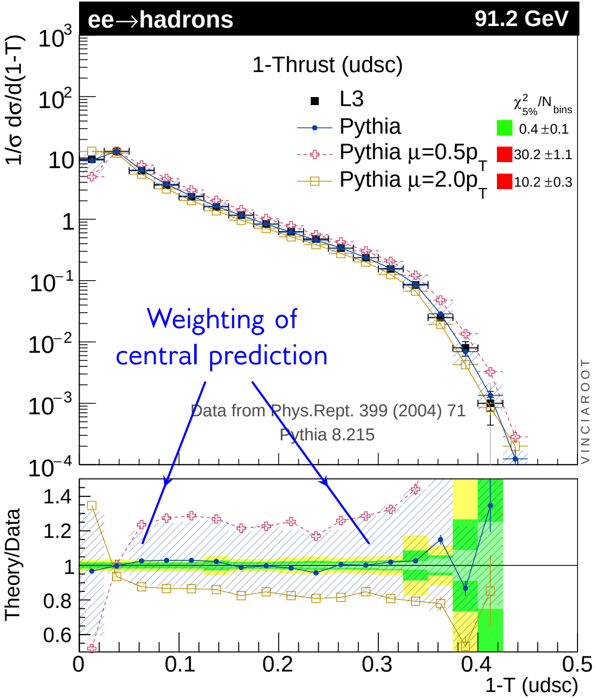
- Differences observed between generators, and sensitivity to shower α_s in $t\bar{t}$ production in kinematics/multiplicity of additional jets
- Retune PS ISR α_s and ME matching parameter using POWHEG+Pythia $t\bar{t}$ vs dilepton+jets data, yields (much) lower value of $\alpha_s^{ISR} = 0.1108$

Njets (p_T >30 GeV)



Lead jet p_T




ee→hadrons
91.2 GeV
1-Thrust (udsc)
L3
Pythia
Pythia μ=0.5p_T
Pythia μ=2.0p_T
χ²_5%/N_bins
0.4 ±0.1
30.2 ±1.1
10.2 ±0.3
1/σ dσ/d(1-T)
Weighting of
central prediction
Data from Phys.Rept. 399 (2004) 71
Pythia 8.215
VINCIAROOT
Theory/Data
1-T (udsc)

In Situ Parton Shower Parameter Variation

- Includes renormalisation-scale and non-singular term variations
- Output = vector of alternative weights for each event
- quick estimate of uncertainties without needing separate runs
- a single sample to run through detector simulation etc.
- (hadronisation etc also only has to be carried out once).
- choose which variations you want, how large, correlated/uncorrelated

Shower is iterative selection of branchings:

$$\mathcal{R}_t \in [0,1] \;=\; \Delta(t_0, t) \;=\; \exp\left(-\int_t^{t_0} dt_1 \int dz_1 P(t_1, z_1)\right)$$

$P(t,z) = \dfrac{\alpha_s(t)}{2\pi}\dfrac{P(z)}{t}$ is complicated :: use veto algorithm

$f(t)$: exact $g(t)$: easy to sample

$$\mathcal{P}_0(t) = \exp\left\{-\int_t^{\bar{t}} g(t')\,\mathrm{d}t'\right\} g(t)\,\underbrace{\frac{f(t)}{g(t)}}_{\mathrm{p_{acc}}} = f(t)\mathrm{e}^{-\int_t^{\bar{t}} g\,\mathrm{d}t'}$$

$$\mathcal{P}_1(t) = \int_t^{\bar{t}} \mathrm{d}t_1 \mathrm{e}^{-\int_{t_1}^{\bar{t}} g\,\mathrm{d}t'} g(t_1)\,\underbrace{\left[1 - \frac{f(t_1)}{g(t_1)}\right]}_{\mathrm{p_{rej}}} \mathrm{e}^{-\int_t^{t_1} g\,\mathrm{d}t'} g(t)\,\underbrace{\frac{f(t)}{g(t)}}_{\mathrm{p_{acc}}}$$

$$\mathcal{P}_1(t) = \mathcal{P}_0(t) \int_t^{\bar{t}} \mathrm{d}t_1 \left[g(t_1) - f(t_1)\right]$$

$$\sum_i \mathcal{P}_i(t) = \mathcal{P}_0(t) \exp\left\{\int_t^{\bar{t}} \left[g(t_1) - f(t_1)\right] \mathrm{d}t_1\right\} = f(t)\mathrm{e}^{-\int_t^{\bar{t}} f\,\mathrm{d}t'}$$

Variations on a theme

$$\underbrace{f \to f'}_{\alpha_s(\mu_R)\to\alpha_s(c\mu_R)} \quad \Rightarrow \mathrm{p}_{\mathrm{acc}}, \mathrm{p}_{\mathrm{rej}} \to \mathrm{p}'_{\mathrm{acc}}, \mathrm{p}'_{\mathrm{rej}}$$

Careful account of weights (ratios of probabilities) in each step of algorithm allows for variation calculations

Similar methodology can be used to bias (say rare) emissions

Even negative weights can be handled (expected at NLL)

- already expanded to include fragmentation, flavor selection
- to do: MPI, rescattering, particle decays
- natural to incorporate into tuning process

Still, can't reweight nothing (or small) to something

$$\text{\# events} = \frac{(\sum w_i)^2}{\sum w_i^2}$$

The Future: ML & Pythia

Reducing uncertainty through better models

ML $\equiv$ very complex parametrization of data

- Identify subtle patterns in complex data
- Automate tedious tasks and perform them rapidly

Lund string model is a probabilistic model

- Probabilities are about ignorance of causes
- Fundamental limitation (field theory + uncertainty principle) or lack of understanding / inability to calculate

Either way, ML can supplement string model to better fit data

Important: Hadronization uncertainties are large for many key measurements at HL-LHC

First Steps

Treat Lund String model predictions as DATA

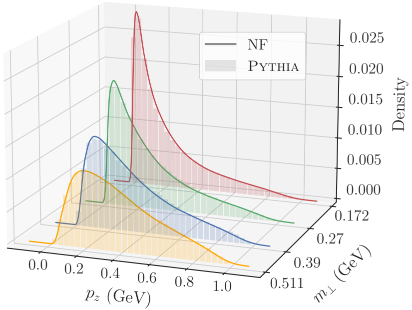


Figure: A comparison between the `Pythia` (histrograms) and `MLHAD` NF generated (solid lines) single emission p_z distributions conditioned over four values of transverse mass $m_\perp$ (left).

Next: Learn on data

- Develop hadronization-dependent observables (IR unsafe)

Message: Uncertainty $\Rightarrow$ Tuning

- Choice of Data, Predictions, and Evaluation is not trivial
  - examples from LEP, TeV, LHC must be adapted to ν-physics
  - divide-and-conquer not necessarily possible
- Need a framework, but it doesn't have to be perfect (or rewritten from scratch)
- Expert or Automatic Tunes??? Both!!!
  - experts use better tools to make tasks easier
- Models $\neq$ truth
  - incorporate systematic uncertainties (know your algorithms!)
  - beware of gaps
- Back-alley statistics

  $\exists$ better algorithms: coverage