FAST EVALUATION OF THEORY UNCERTAINTIES WITH MCGRID AND SHERPA

- ENRICO BOTHMANN, MAREK SCHÖNHERR, STEFFEN SCHUMANN

- novel internal reweighting of SHERPA
- automated grid production using MCGRID

Georg-August-Universität
Göttingen

2. PhysikalischesInstitut

Motivation

- we have tools that automate predictions for multi-leg processes @ NLO + parton shower, e.g. SHERPA (Gleisberg et al.)

- we want to vary QCD input parameters: $\mu_{\mathrm{F/R}}, \ \alpha_{\mathrm{S}}$, PDFs for theoretical error estimates, PDF and α_{S} fits

- dedicated calculation time-consuming: $\mathcal{O}$(days)

➡ tedious for error estimates

➡ impossible for PDF fits: $\mathcal{O}$(100k) evaluations!

Sherpa: 0811.4622

General idea

- structure of a MC@NLO cross section in the Catani-Seymour subtraction scheme:

$$\sigma_{pp\to X}^{\mathrm{MC@NLO}} = \int \mathrm{d}\sigma^{\mathrm{B}} + \int \mathrm{d}\sigma^{\mathrm{V}} + \int \mathrm{d}\sigma^{\mathrm{I}} + \int \mathrm{d}\sigma^{D_A - D_S} + \int \mathrm{d}\sigma^{R - D_A}$$

- fully understand the dependence on the input parameters for each contribution
- store non-dependent factors as weights in the HEPMC (Dobbs et al.) event record
- apply new choice of input parameters for fast re-evaluation
- fast because the lengthy part of the calculation is stored in weights and re-used

HepMC: CPC 134 (2001) 41–46

Dependency structure: examples

- some are easy, e.g. Born-like event:

$$\int \mathrm{d}\sigma^{\mathrm{B},R-D(A)} = \sum_{e=1}^{N_{\mathrm{evt}}} \left(\frac{\alpha_s(\mu_R^2)}{2\pi}\right)^p f_1(x_1,\mu_F^2) f_2(x_2,\mu_F^2) \cdot w_e$$

- some are intricate, e.g. integrated subtraction part:

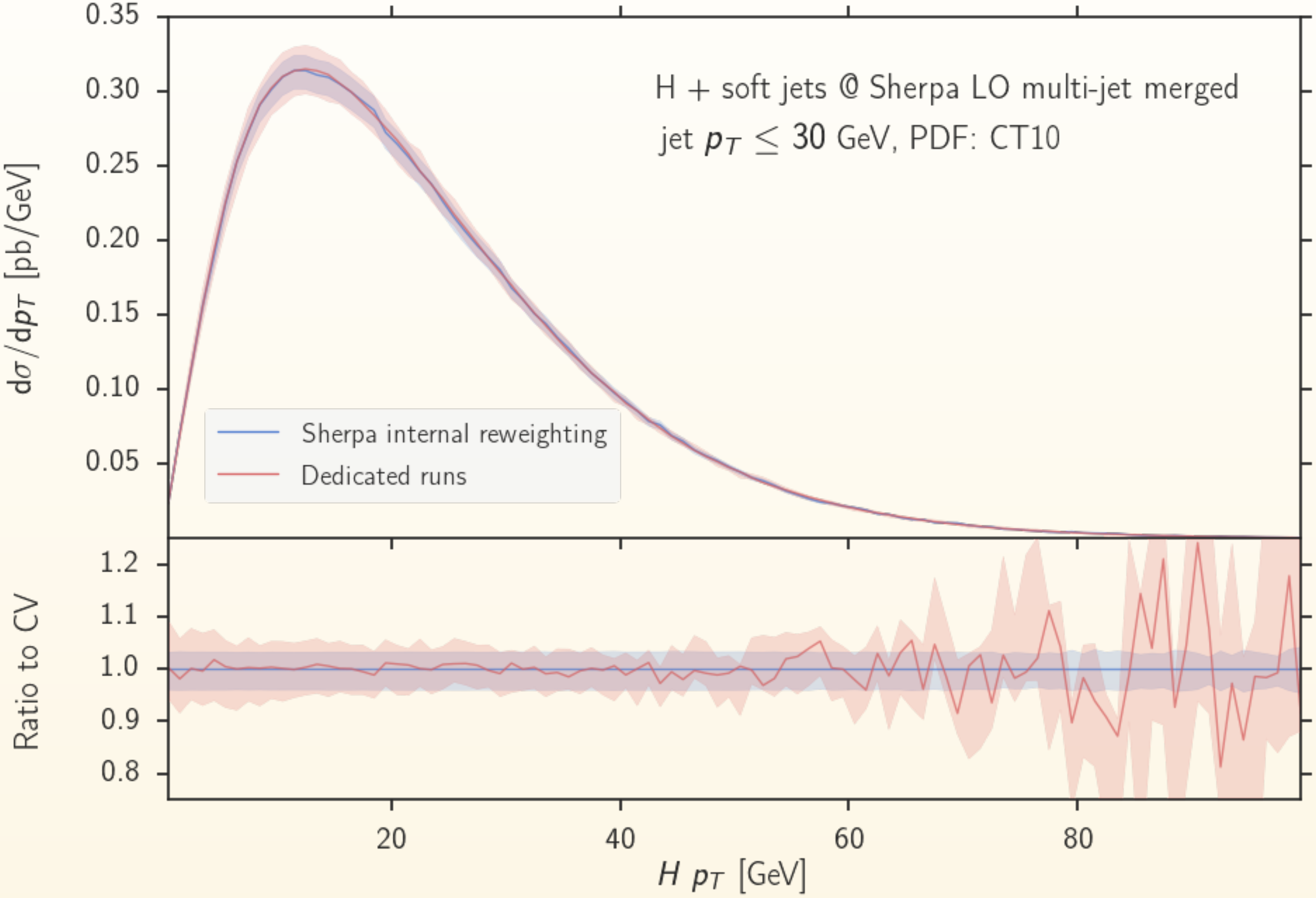
$$\begin{aligned}\int d\sigma^{\mathrm{I}} \quad = \quad & \sum_{e=1}^{N_{\mathrm{evt}}} \left(\frac{\alpha_s(\mu_R^2)}{2\pi}\right)^{p_{\mathrm{NLO}}} \left\{ f_1(x_1,\mu_F^2)\, w_e\, f_2(x_2,\mu_F^2) \right. \\ & + \left(\sum_{k=1}^{4} f_1^{(k)}(x_1,x_1',\mu_F^2)\, w_{e,k}'\right) f_2(x_2,\mu_F^2) \\ & \left. + f_1(x_1,\mu_F^2) \left(\sum_{k=1}^{4} w_{e,k}''\, f_2^{(k)}(x_2,x_2',\mu_F^2)\right) \right\}\end{aligned}$$

Sherpa-internal reweighting

- convolute weights with different choices for the QCD input parameters *on-the-fly* during event generation
- supports: LO, NLO, fixed-order part of MC@NLO, LO&NLO multijet merged
- to be released in upcoming version 2.2.0 of SHERPA (Gleisberg et al.)
- useful for error estimates, config example for PDF and scale errors:
  - `SCALE_VARIATIONS 1.,1.,CT10[all] 0.25,0.25,CT10 4.,4.,CT10`

➡ PDF band & scale errors, only takes factor $\mathcal{O}(1)$ longer

Sherpa-internal reweighting: validation
H + soft jets @ LHC 13 TeV



52 error sets

❖

reweighting run ~4 times slower than single run

❖

smoother reweighting band (same statistics)

QCD NLO interpolation grids

- store slow part of calculation on a (x_1, x_2, Q^2)-grid:
- fill event weights w_e into grid weights $W_{ij\alpha\beta\tau}$ event-by-event
- reconvolute grid with QCD input parameters, only takes $\mathcal{O}$(ms)!:

$$\sigma_{pp\to X} = \sum_{\alpha\beta}^{N_x} \sum_{\tau}^{N_Q} \left(\frac{\alpha_s(Q^2_\tau)}{2\pi} \right)^p \sum_{ij} f_i(x_\alpha, Q^2_\tau) f_j(x_\beta, Q^2_\tau) \cdot W_{ij\alpha\beta\tau}$$

- implementations:
  APPLGRID (Carli et al.), FASTNLO (Wobisch et al.)
- great, but how to generate grids in an automated way?

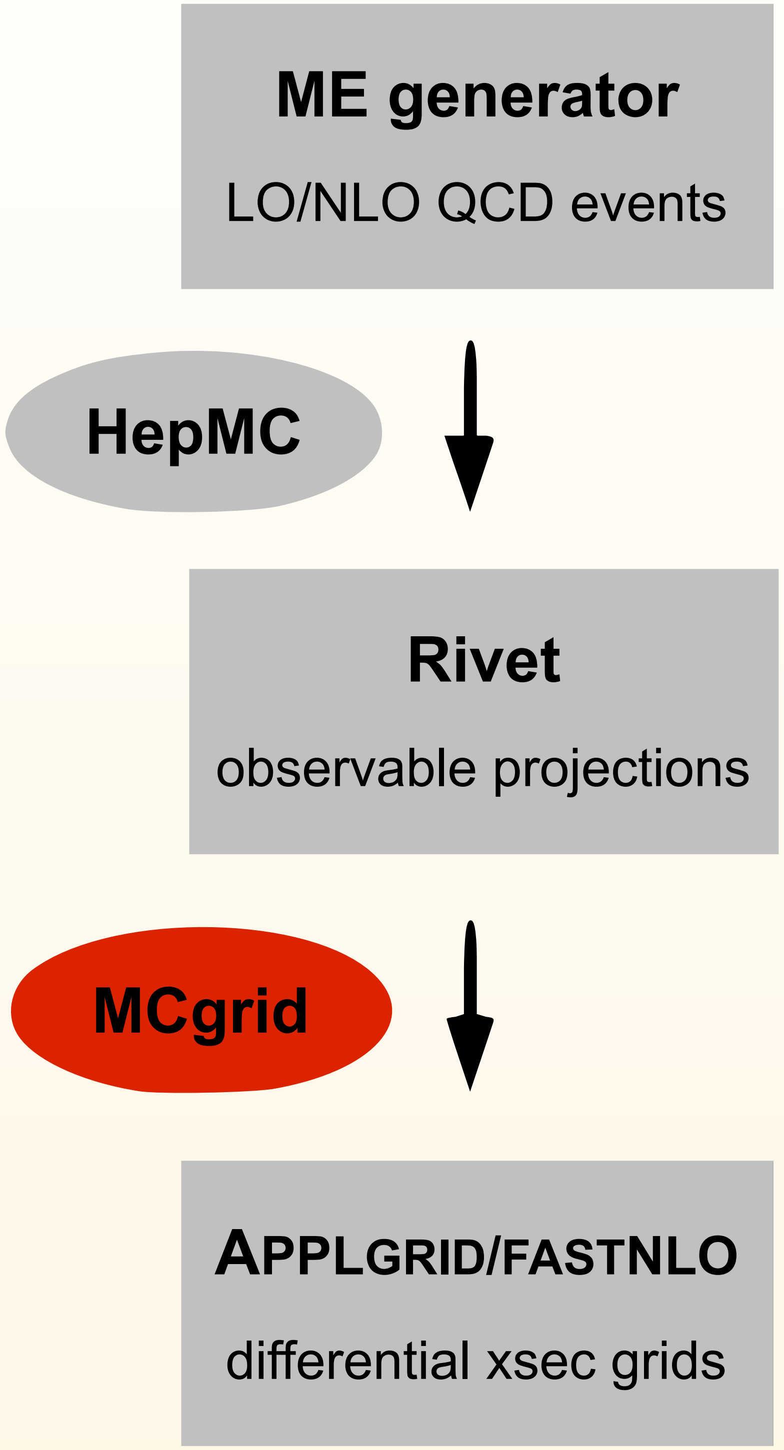
APPLgrid: 0911.2985
fastNLO: 1109.1310

Use MCgrid for automated grid creation

- RIVET (Buckley et al.) already provides analysis codes: setup of histograms, projections, …

➡ idea: build interface between NLO events and dedicated grid implementations as a RIVET plugin

✓ MCGRID (Del Debbio et al.)



ME generator

LO/NLO QCD events

HepMC

Rivet

observable projections

MCgrid

APPLGRID/FASTNLO

differential xsec grids

MCgrid: 1312.4460
Rivet: 1003.0694

MCGRID features

- in principle not generator-specific, because using HEPMC (Dobbs et al.)
    - but relevant event info must be provided: currently only SHERPA
- supports APPLGRID and FASTNLO

new & to be released

- supports: LO, NLO, fixed-order part of MC@NLO
- simultaneous filling of multiple grids (even from multiple analyses)

MCGRID features: Subprocess identification

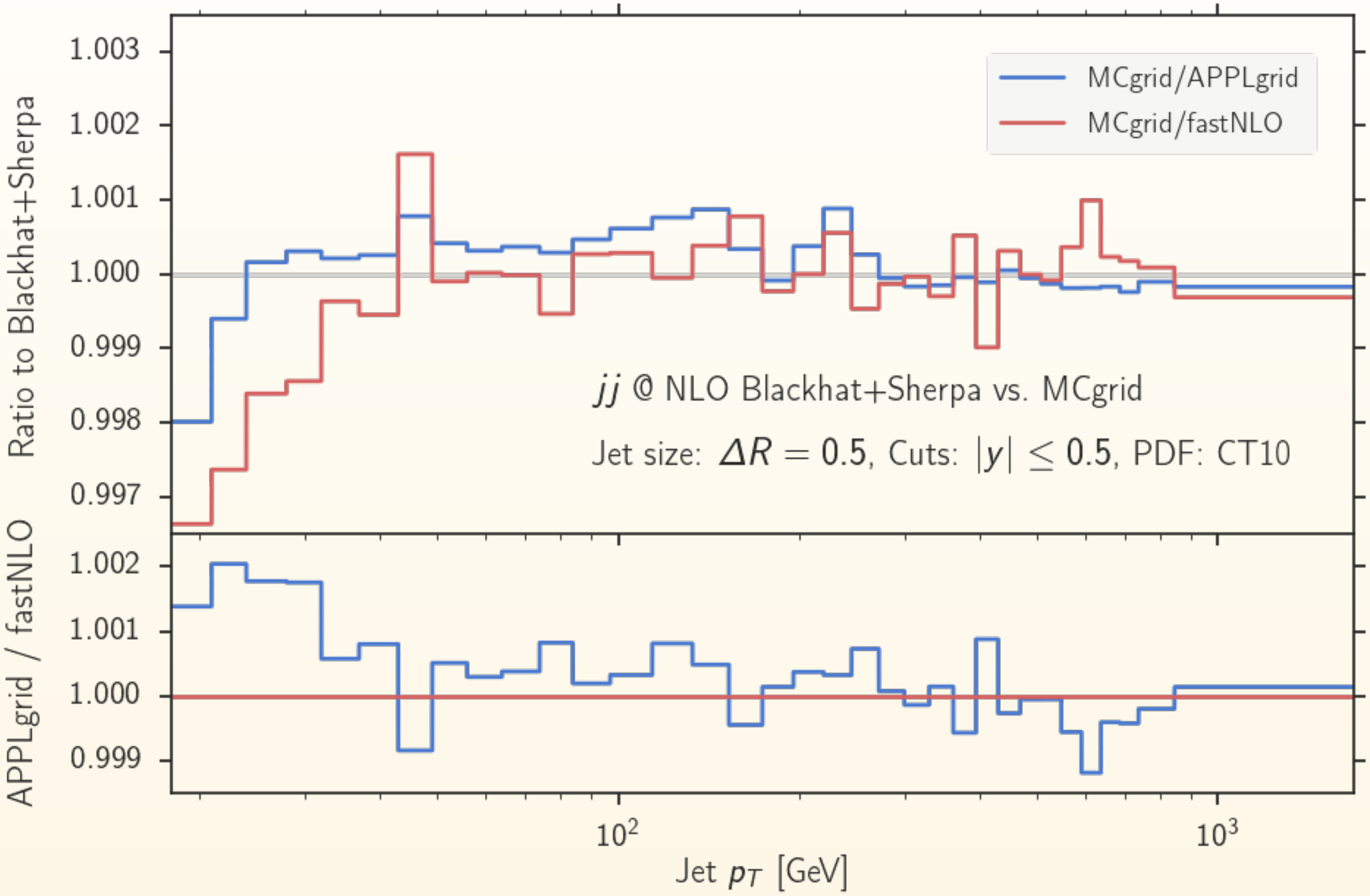
- summing over incoming parton pairs

$$\sigma_{pp\to X} = \sum_{\alpha\beta}^{N_x} \sum_{\tau}^{N_Q} \left(\frac{\alpha_s(Q^2_\tau)}{2\pi}\right)^p \sum_{ij} f_i(x_\alpha, Q^2_\tau) f_j(x_\beta, Q^2_\tau) \cdot W_{ij\alpha\beta\tau}$$

➡ 121 grids, lots of memory, poor performance

- solution: group subprocesses with same $\mathrm{d}\hat{\sigma}_{ij\to X}$
- much better file size & speed
- automated using Sherpa/MCgrid

MCGRID validation: inclusive jets @ LHC 7 TeV



Grid in x and Q^2

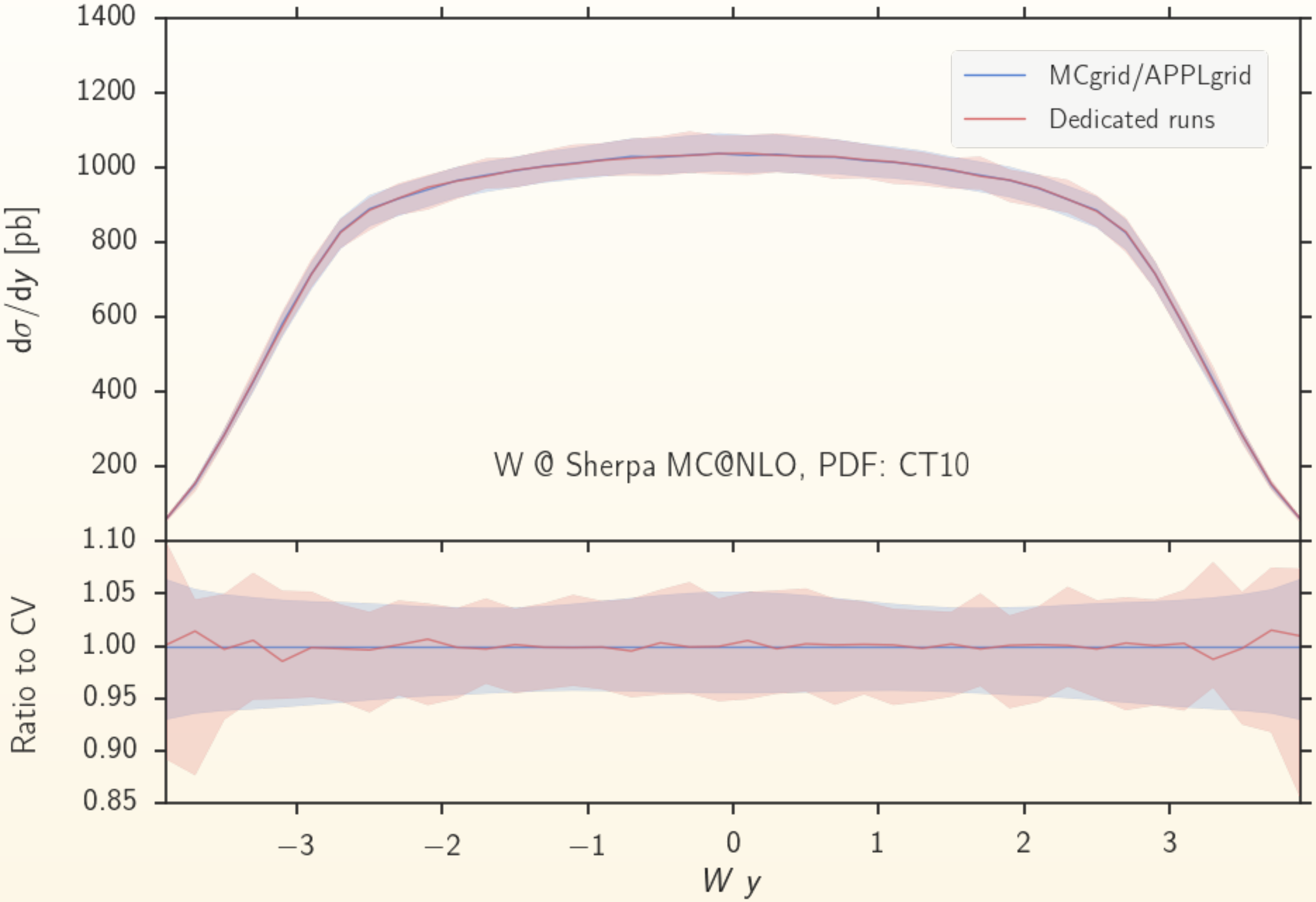
❖

no specific optimizations

❖

promille (interpolation) accuracy

MCGRID validation: W production @ LHC 8 TeV



Grid in x and Q^2

❖

52 error sets

❖

reconvolution time dominated by PDF loading

❖

smoother reweighting band (same statistics)

Residual PDF dependence of the shower

- PDF dependencies inside the parton shower are not yet tracked by reweighting/grid

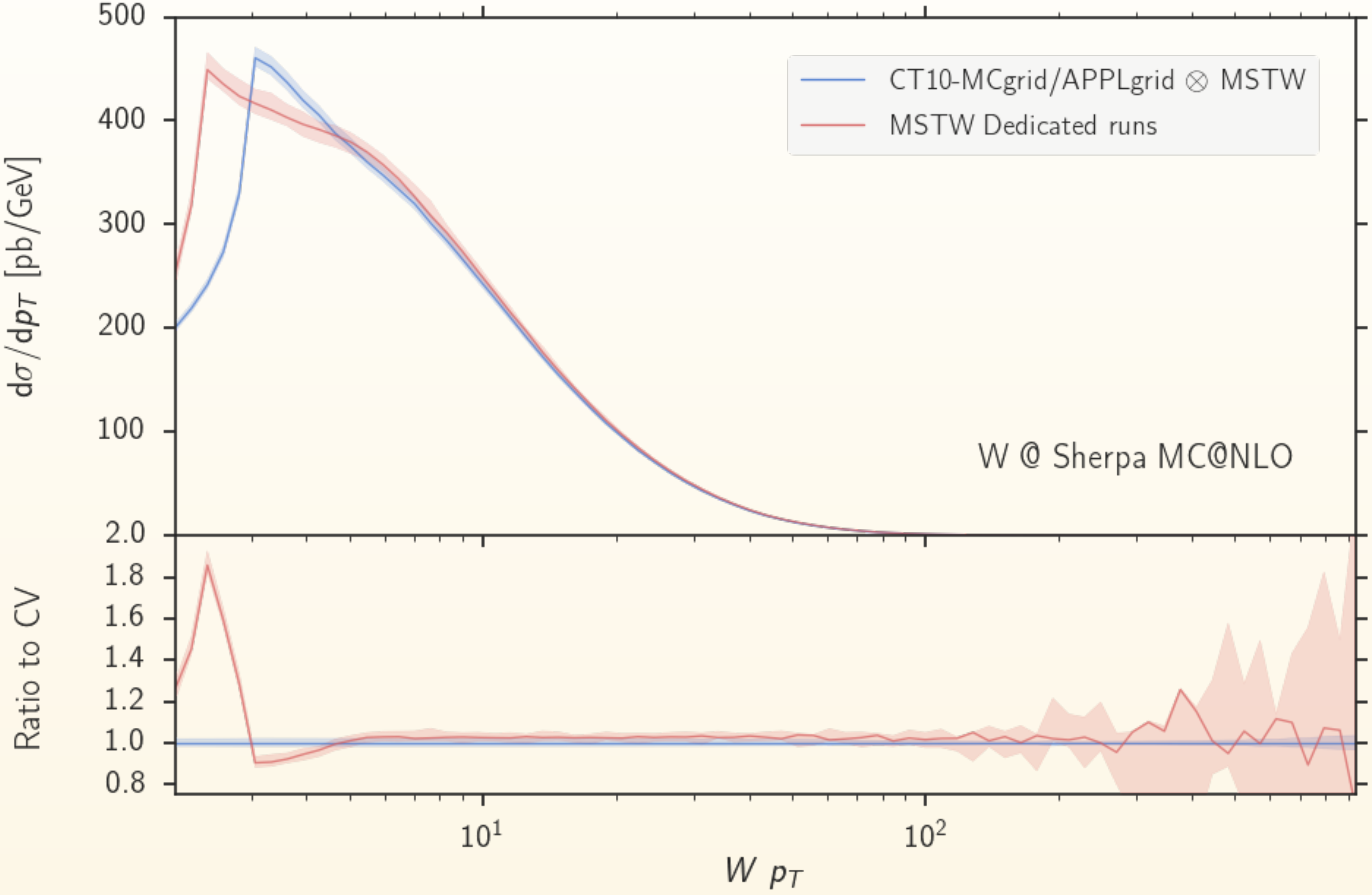
- consider e.g. the Sudakov form factor, which gives the probability for no additional branching:

$$\Delta_a(x, Q_1^2, Q_2^2) =$$

$$\exp\left\{ - \int_{Q_1^2}^{Q_2^2} \frac{\mathrm{d}Q^2}{Q^2} \int_x^{1-\delta} \frac{\mathrm{d}z}{z} \sum_b \hat{P}_{b\to ac}(z) \frac{\alpha_s(Q^2)}{2\pi} \frac{f_b(x/z, Q^2)}{f_a(x, Q^2)} \right\}$$

- if tracked/approximated, could fit PDFs to data that is not covered by fixed-oder calculations

Residual PDF dependence of the shower W production @ LHC 8 TeV



very similar to comparison of dedicated CT10 vs. dedicated MSTW

❖

Large part of deviation already present for single PS emission

Possibilites for varying QCD input parameters

	Rerun	Reweight	Interpolate
Time overhead	**all runs**	**factor of O(1)**	**Negligible**
File sizes	0	0	10 MB
Scale choice	free	free	a priori free / a pasteriori constrained
Tracked constributions	all	LO, NLO, fixed-order part of MC@NLO, LO&NLO multijet	LO, NLO, fixed-order part of MC@NLO
Great for …	avoiding residual PDF dependence of parton shower	getting theory uncertainties on the fly	fits & a posteriori reevaluation

Conclusion & Outlook

- SHERPA 2.2.0 will support internal reweighting for on-the-fly error estimates
  - `SCALE_VARIATIONS 1.,1.,CT10[all] 0.25,0.25,CT10 4.,4.,CT10`
- use NLO QCD interpolation grids for fast a posteriori theoretical error estimates and fits
- MCGRID is an interface for *automated* grid production using event generators and RIVET
- for the future we plan to include into MCGRID:
  - NNLO predictions
  - variable final state multiplicities (merging)
  - (approximate) parton shower effects
    ⇝ beyond fixed order PDF fits

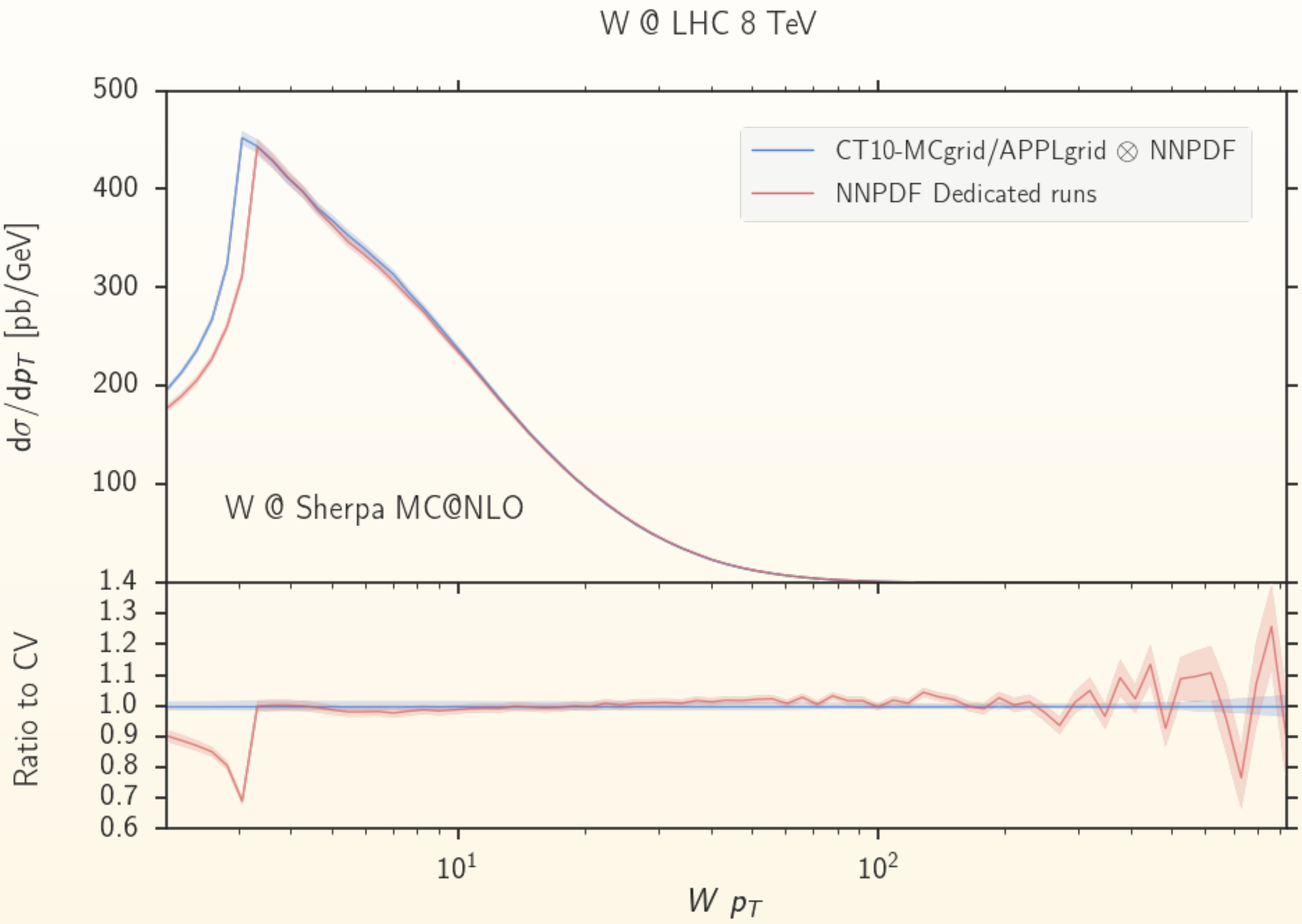
THANKS

- Enrico Bothmann, Marek Schönherr, Steffen Schumann

http://mcgrid.hepforge.org

BACKUP

Residual PDF dependence of the shower W production @ LHC 8 TeV



52 PDF eigenvectors

❖

statistically equivalent

⬇

less fluctuations

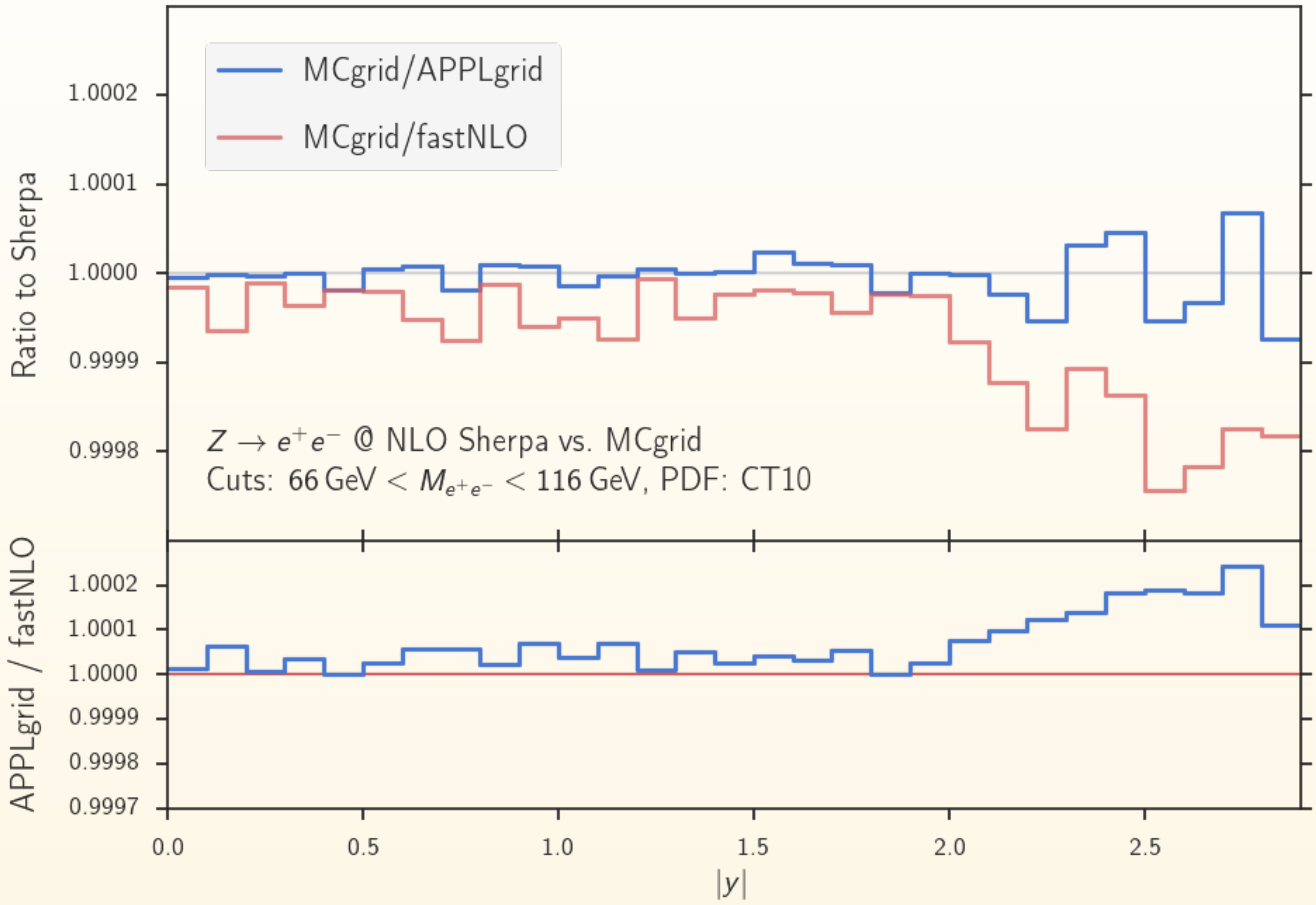
Enable a Rivet analysis for use with MCgrid

```
using namespace MCgrid;
Histo1DPtr _h_yZ; // Rivet histogram
gridPtr _fnlo_yZ; // Corresponding grid

// Init phase
subprocessConfig subproc("DY-ppbar.str", BEAM_PROTON, BEAM_ANTIPROTON);
fastnloGridArch  arch(50, 1, "Lagrange", "OneNode", "sqrtlog10", "linear");
fastnloConfig    config(0, subproc, arch, 1960.0);
_h_yZ     = bookHisto1D(2, 1, 1);                // Book Rivet
_fnlo_yZ = bookGrid(_h_yZ, histoDir(), config);  // Book MCgrid/fastNLO

// Analyse phase
PDFHandler::HandleEvent(event, histoDir());  // Update subprocess statistics
_h_yZ   ->fill(yZ, weight);                  // Fill Rivet
_fnlo_yZ->fill(yZ, event);                   // Fill MCgrid/fastNLO

// Finalise phase
scale(_h_yZ, normalisation);                 // Scale Rivet
_fnlo_yZ->scale(normalisation);              // Scale MCgrid/fastNLO
PDFHandler::CheckOutAnalysis(histoDir());    // Finalise
```

fastNLO specific {

⤳ simple & unified API

Example: Drell-Yan @ Tevatron 1.96 TeV



Grid only in x

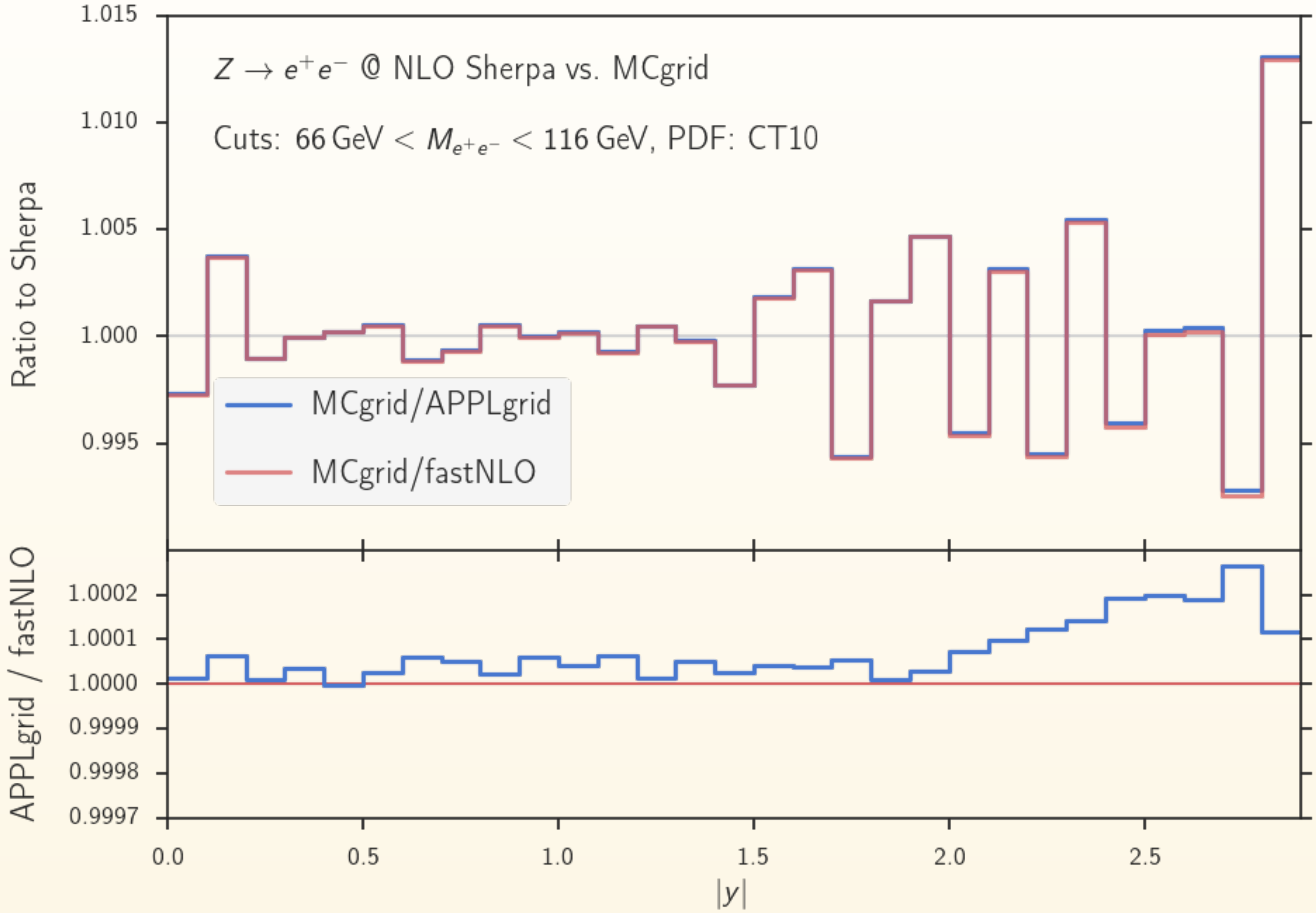
❖

no specific optimizations

❖

sub-promille accuracy

Example: Drell-Yan @ Tevatron 1.96 TeV



Grid only in x

❖

Subprocess ID

❖

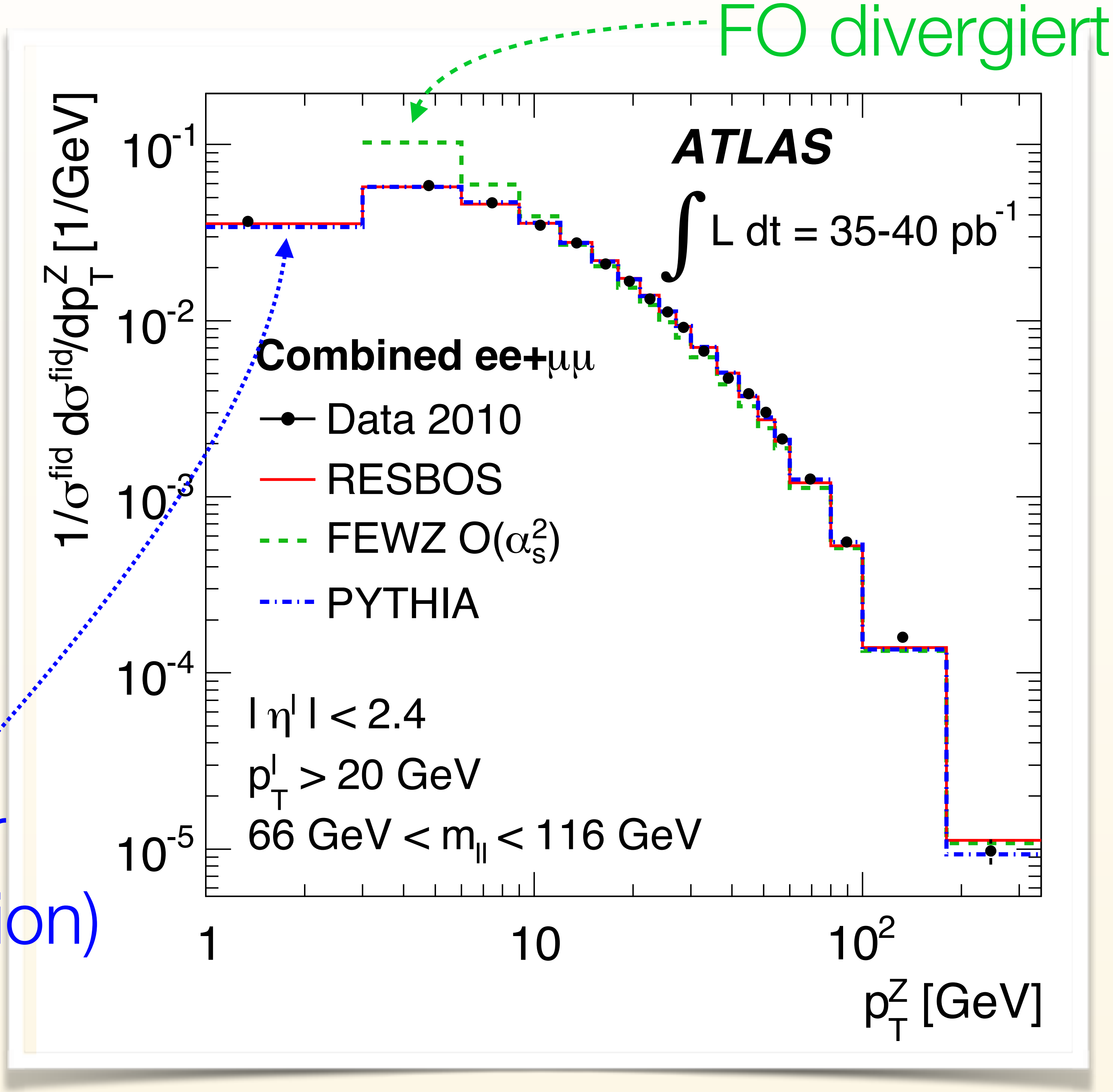
no specific optimizations

❖

sub-percent accuracy

Example for observable that is sensitive to resummation effects

FO divergiert für $p_T^Z \to 0$



All-Order
(Resummation)

⤳ Fixed Order PDF fits cannot use low p_T region

plot taken from some ATLAS paper