HIGHER-ORDER QCD CALCULATIONS FOR THE LHC

Alexander Huss

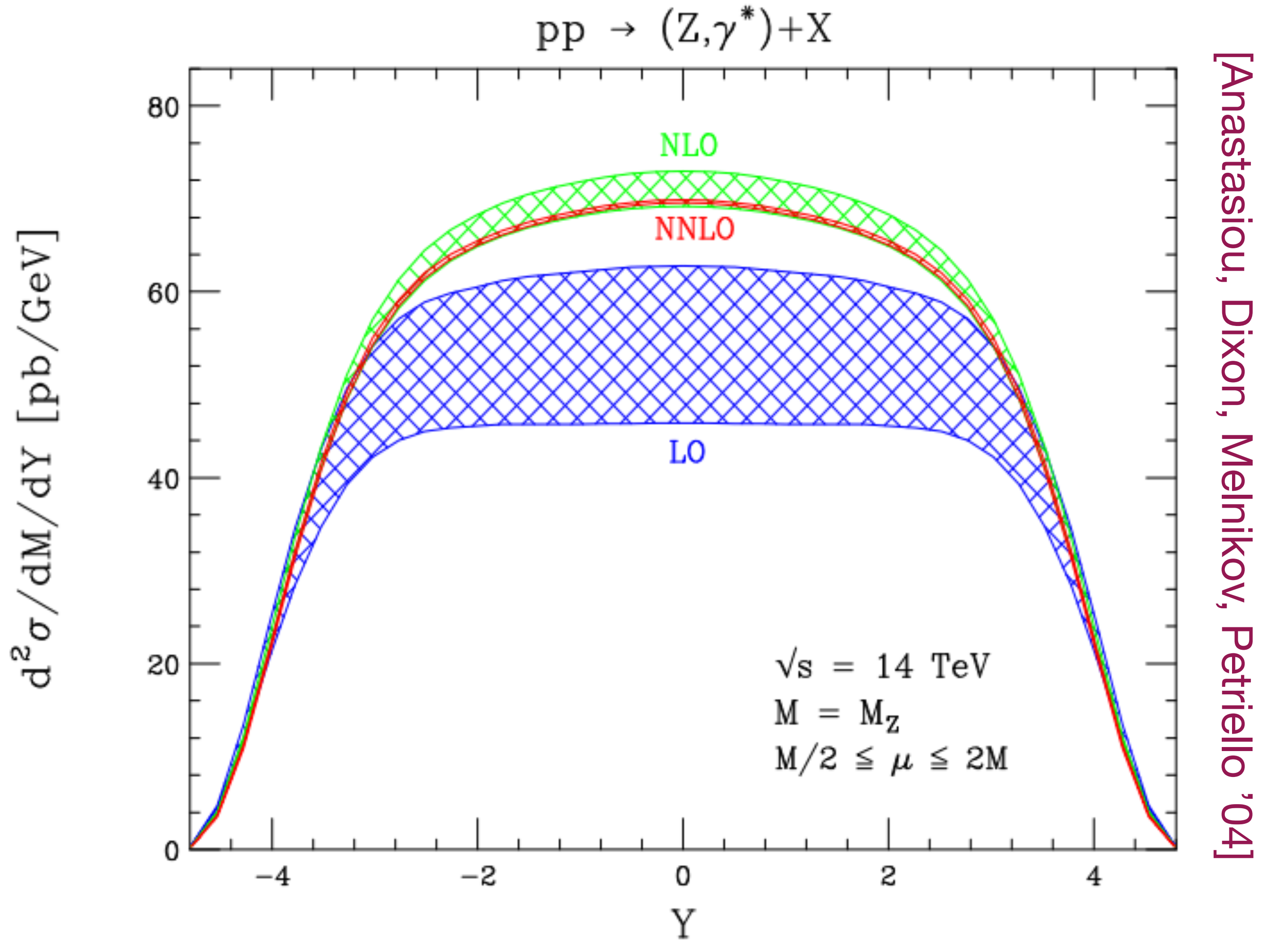
CERN

NNPDF Collaboration & N3PDF Meeting — August 29th 2022

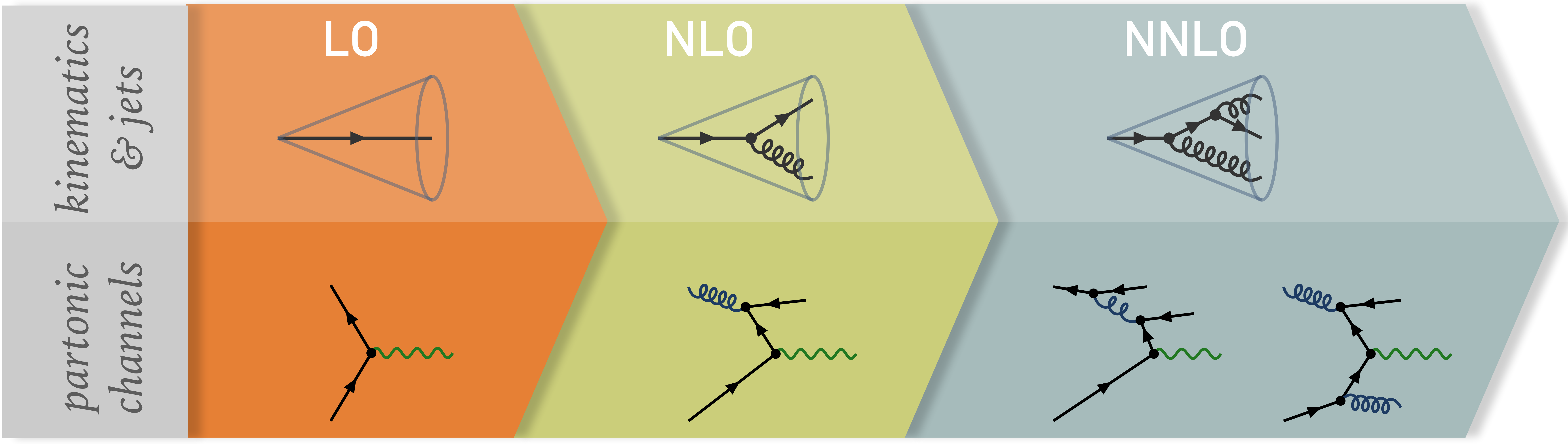
The Plan.

1. *NNLO predictions for the LHC*
    - *jets & interpolations grids*
    - *identified photons & fragmentation*
2. *Differential N^3LO*
    - *Higgs & fiducial power corrections*
    - *Drell-Yan & PDFs*
3. *Bayesian approach to MHO*
    - *the abc model & correlations*
4. *Summary & Outlook*

WHAT WE HOPE NNLO WILL GIVE US

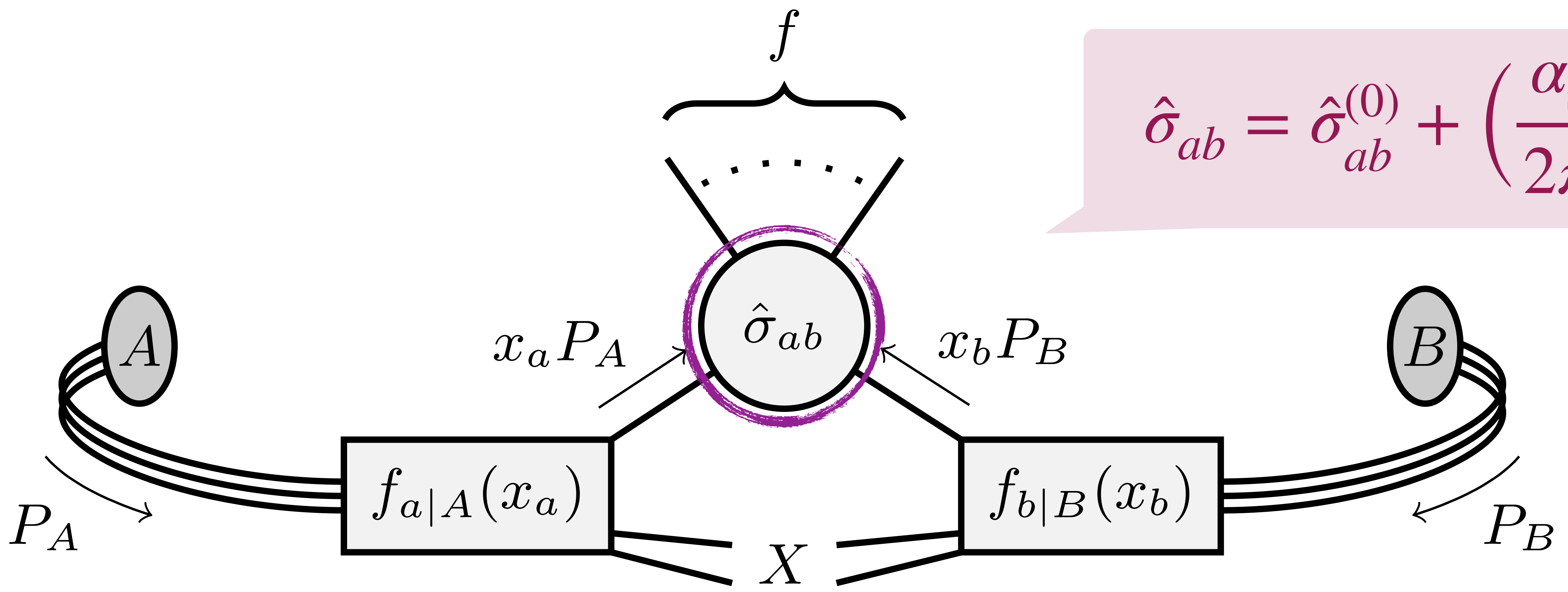


- reduced uncertainties (↭ missing higher orders)
- guaranteed that all partonic channels open up at NNLO
- better modelling of final-state kinematics & jets



The Master Formula

$$\hat{\sigma}_{ab} = \hat{\sigma}^{(0)}_{ab} + \left(\frac{\alpha_s}{2\pi}\right) \hat{\sigma}^{(1)}_{ab} + \left(\frac{\alpha_s}{2\pi}\right)^2 \hat{\sigma}^{(2)}_{ab} + \cdots$$



$$\sigma_{AB} = \sum_{ab} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \; f_{a|A}(x_a)\; f_{b|B}(x_b)\; \hat{\sigma}_{ab}(x_a, x_b)\; \left(1 + \mathcal{O}(\Lambda_{\mathrm{QCD}}/Q)\right)$$

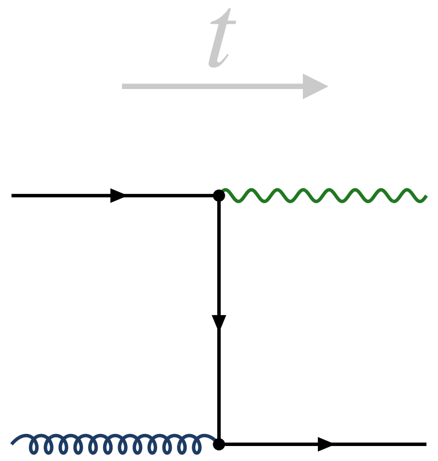
parton distribution functions
(non-perturbative, universal)
in principle, improvable

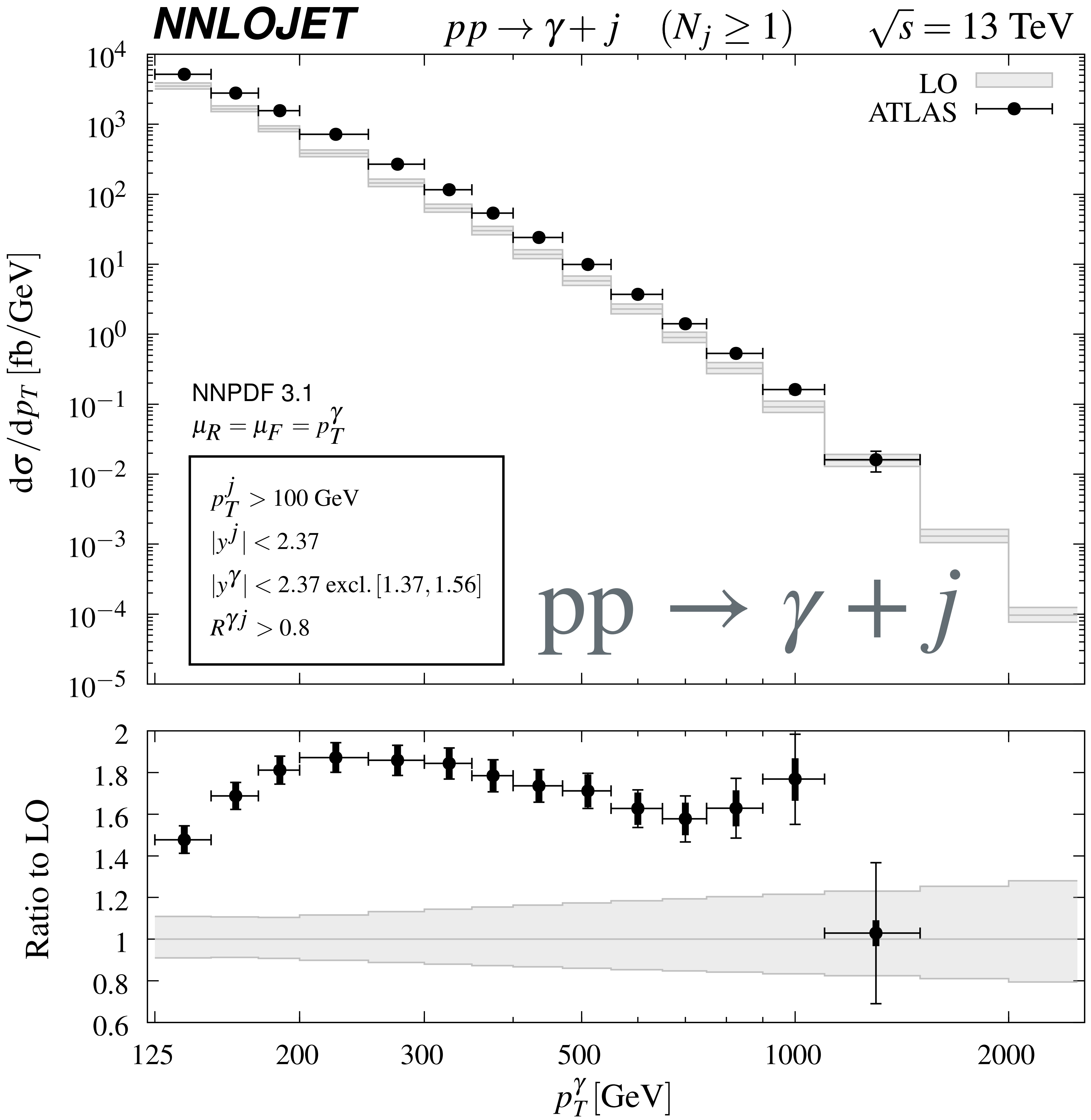
hard scattering
(perturbation theory)
systematically improvable

non-perturbative effects
(power suppressed)
ultimately, limiting factor?

Perturbation Theory @ Leading Order

LO



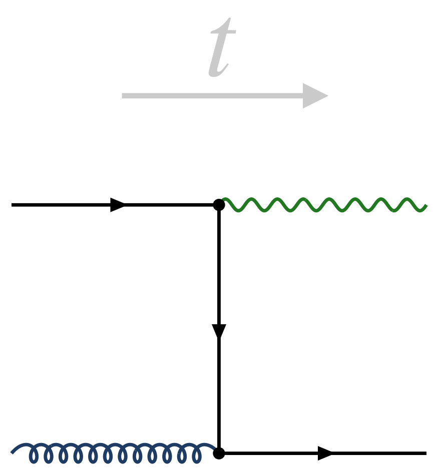


[Chen, Gehrmann, Glover, Höfer, **A.Huss** '19]

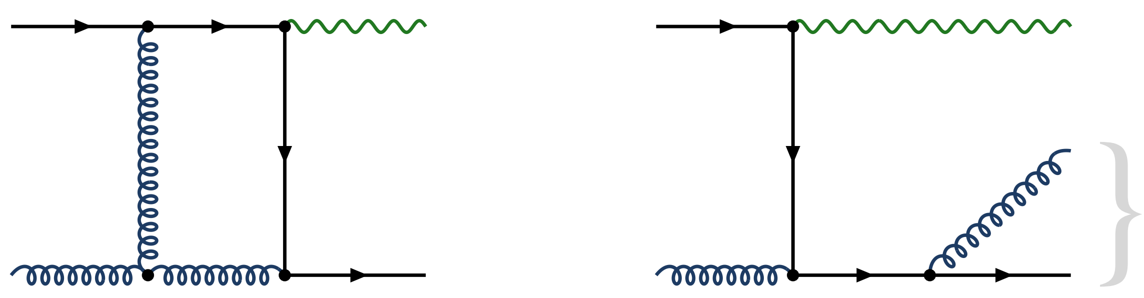
Only captures gross features & unreliable uncertainty estimates

PERTURBATION THEORY @ NEXT-TO-LEADING ORDER

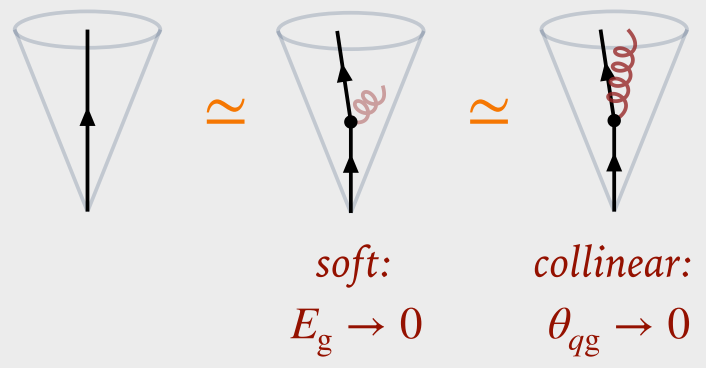
LO



NLO

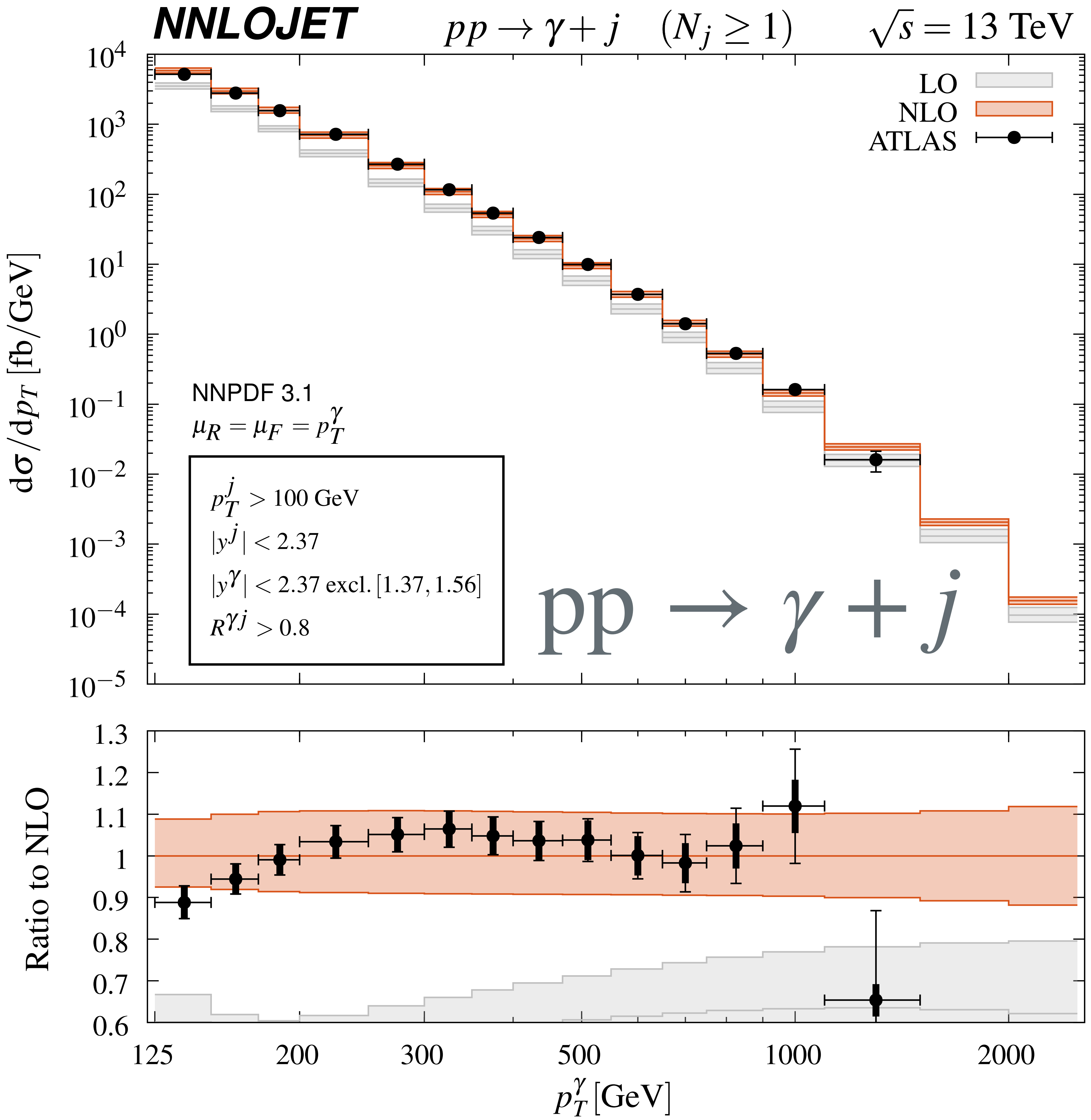


can become unresolved



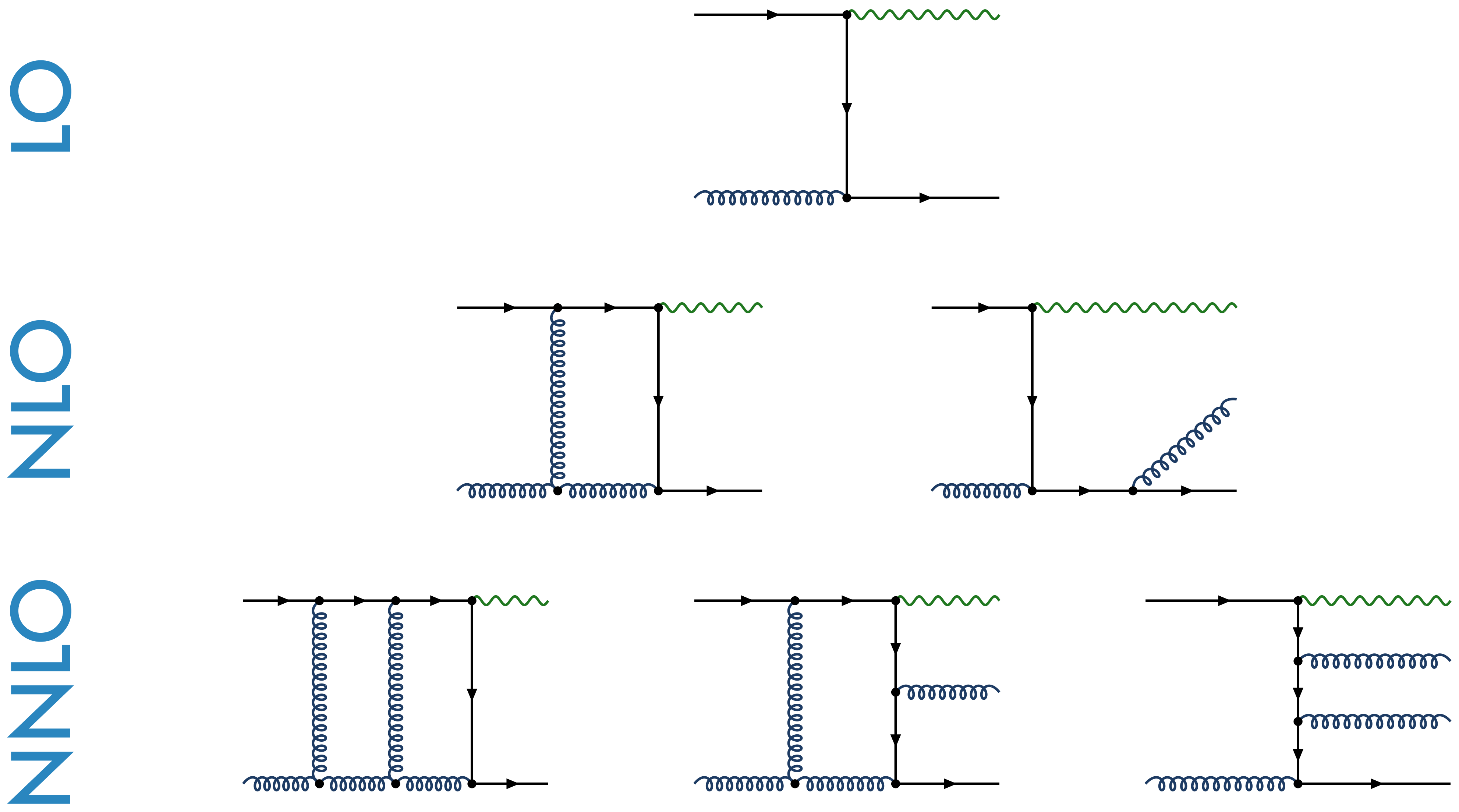
higher order: more loops & legs

typically $\mathcal{O}(10-30\%)$ precision
here: limits the *interpretation* of data!

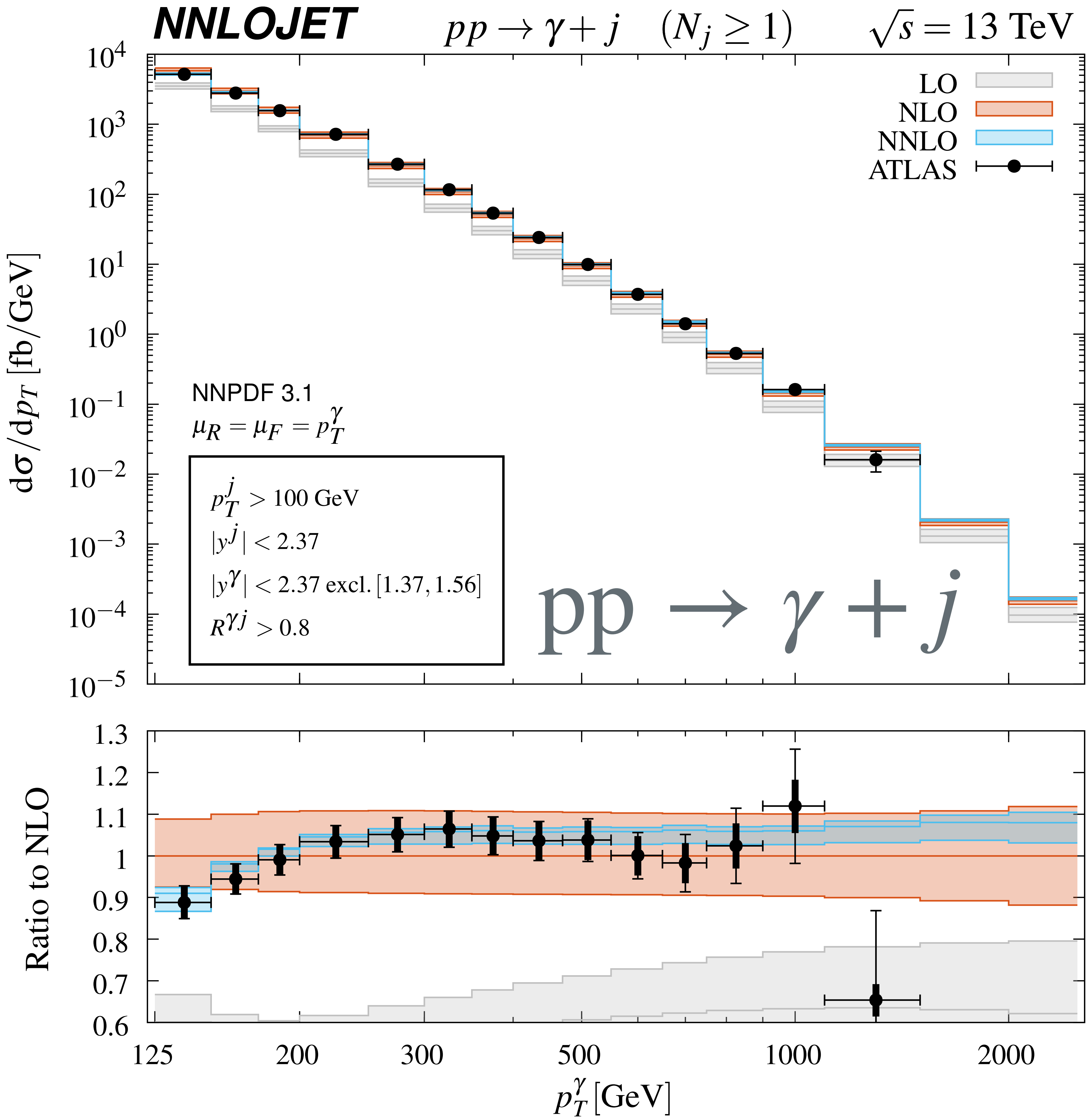


[Chen, Gehrmann, Glover, Höfer, **A.Huss** '19]

Perturbation Theory @ Next-to-Next-to-Leading Order

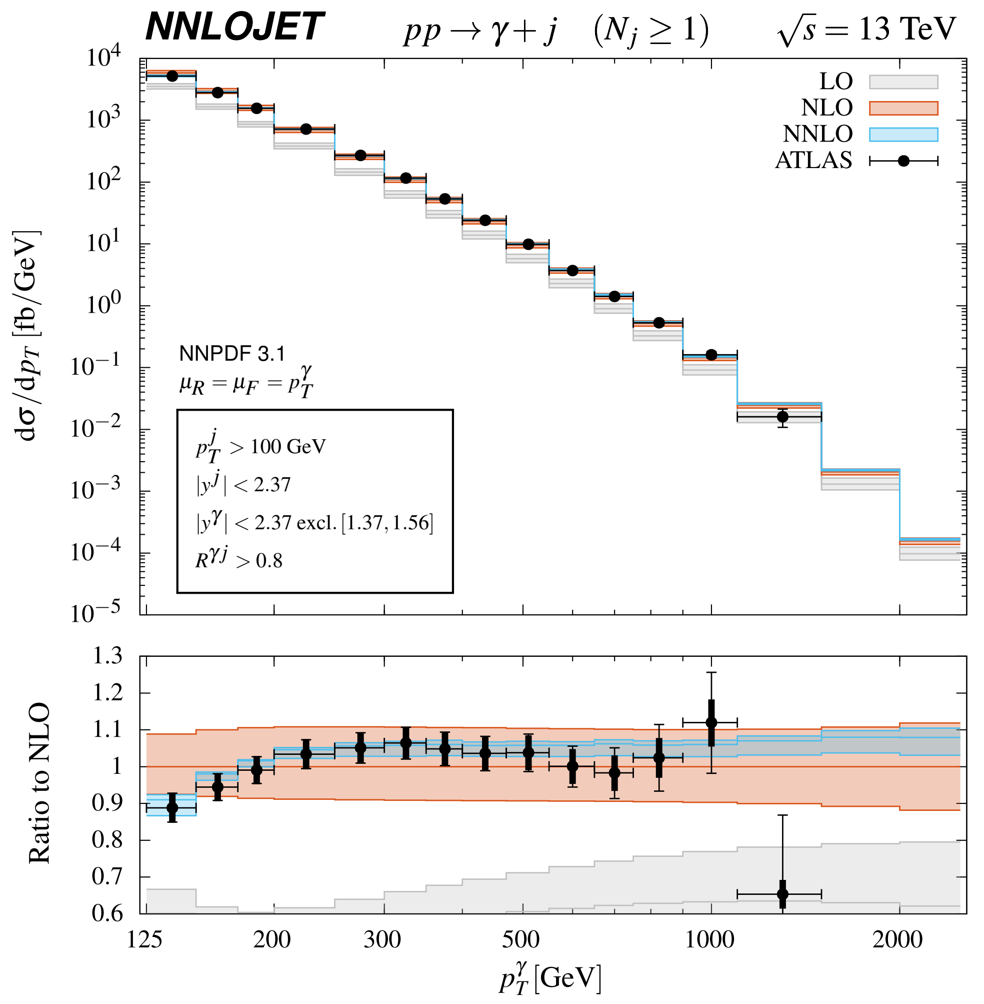


mandatory to achieve *single digit* of relative precision



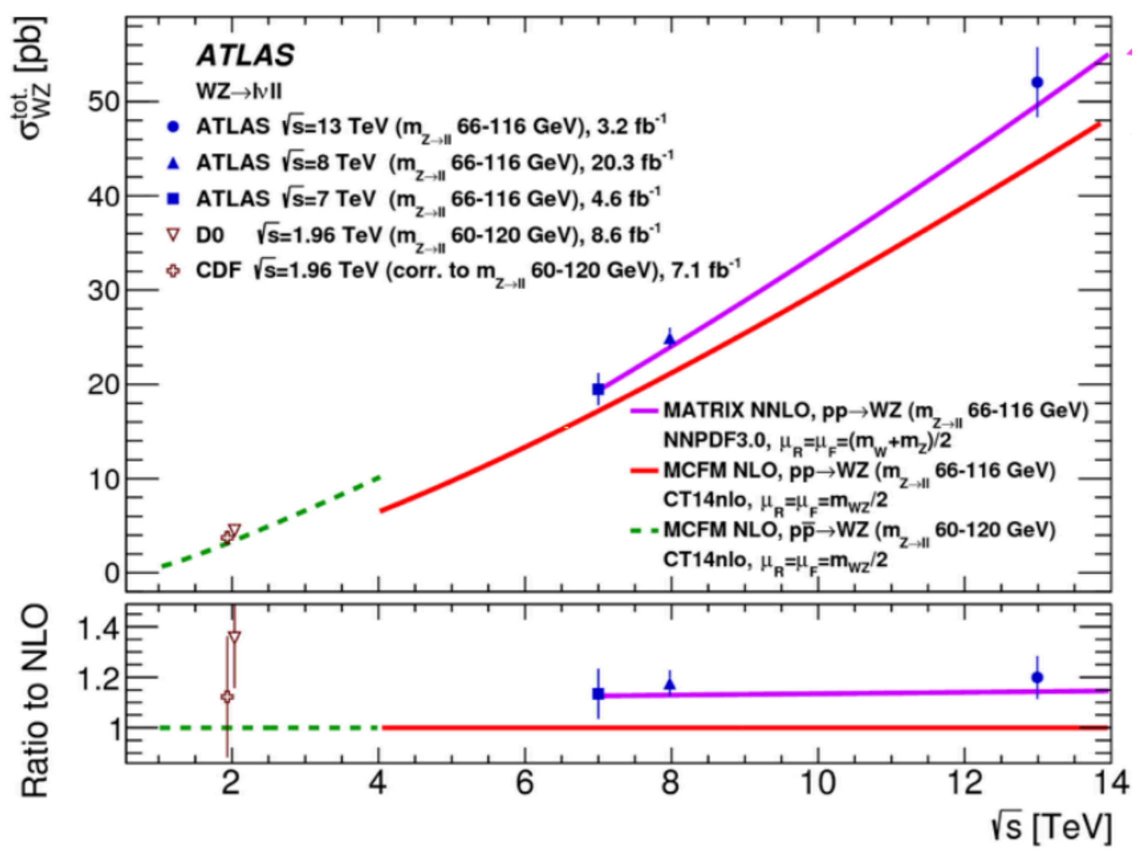
[Chen, Gehrmann, Glover, Höfer, **A.Huss** '19]

What we ~~hope~~ EXPECT NNLO will give us

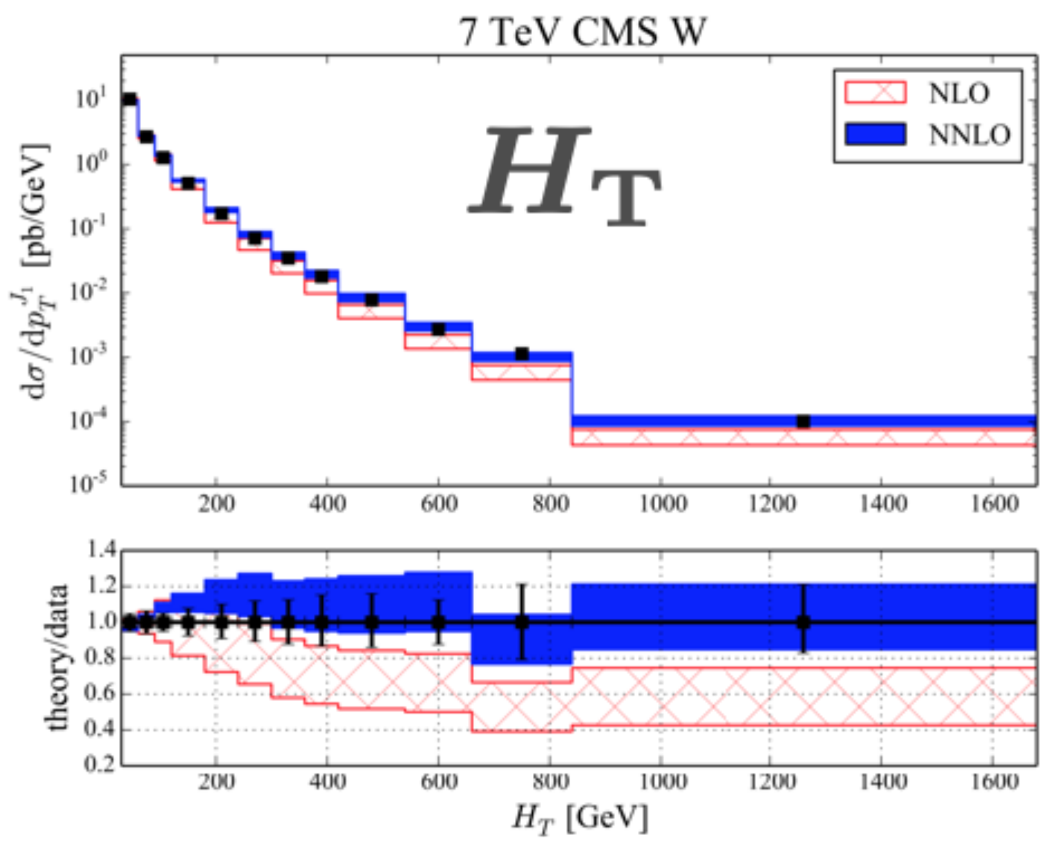


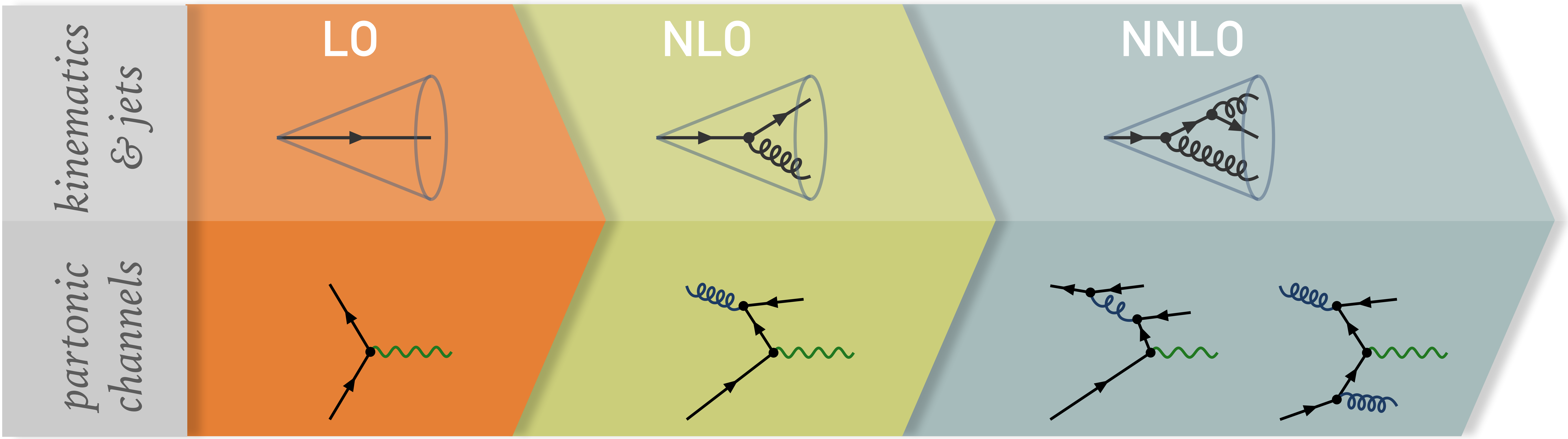
[Chen, Gehrmann, Glover, Höfer, AH '19]

[MATRIX] *WZ*



W + jet [Boughezal, Liu, Petriello '16]

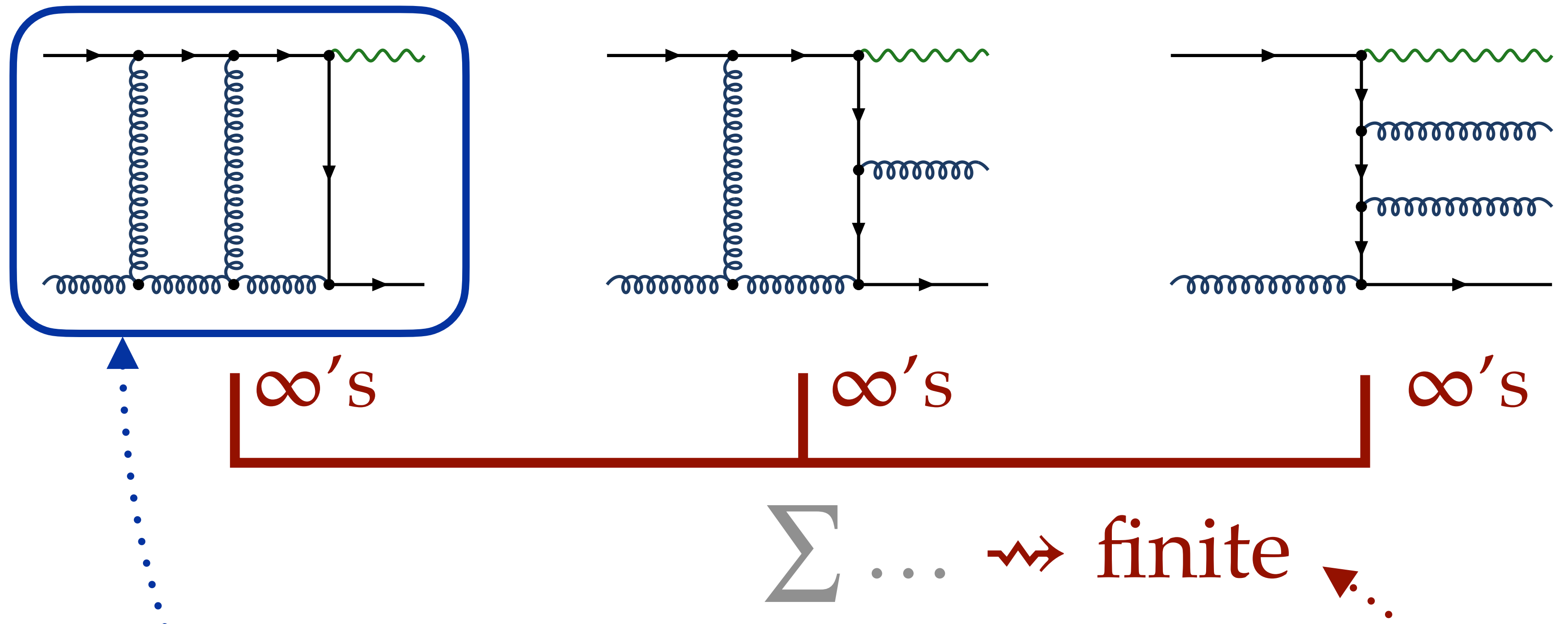


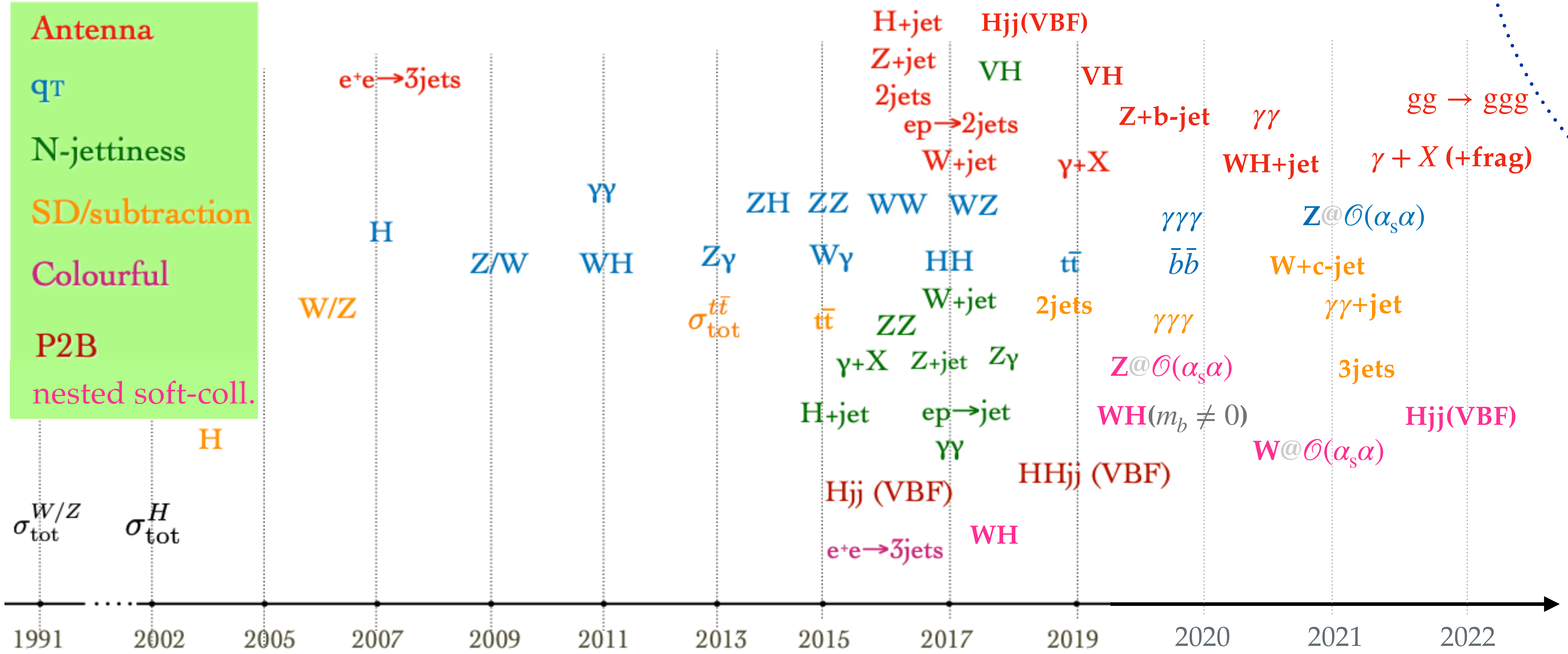


What can we do Today? — The NNLO Timeline

Tremendous progress in the past ~ 5–10 years!

↪ $2 \to 2$ under good control; $2 \to 3$ next frontier





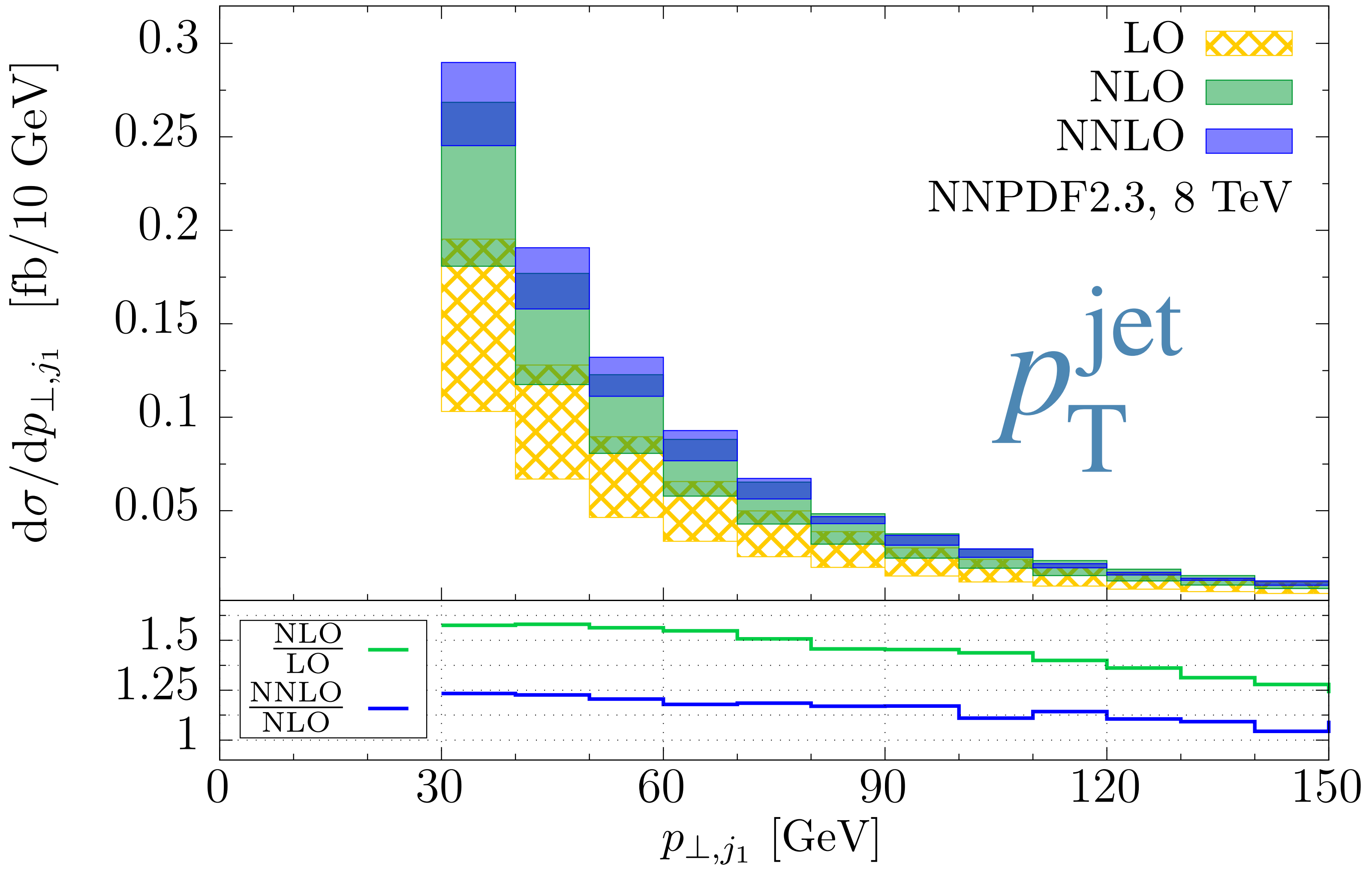
Main challenges:

- amplitudes & multi-loop integrals
- infrared subtractions

Independent Calculations – H + jet ×3!

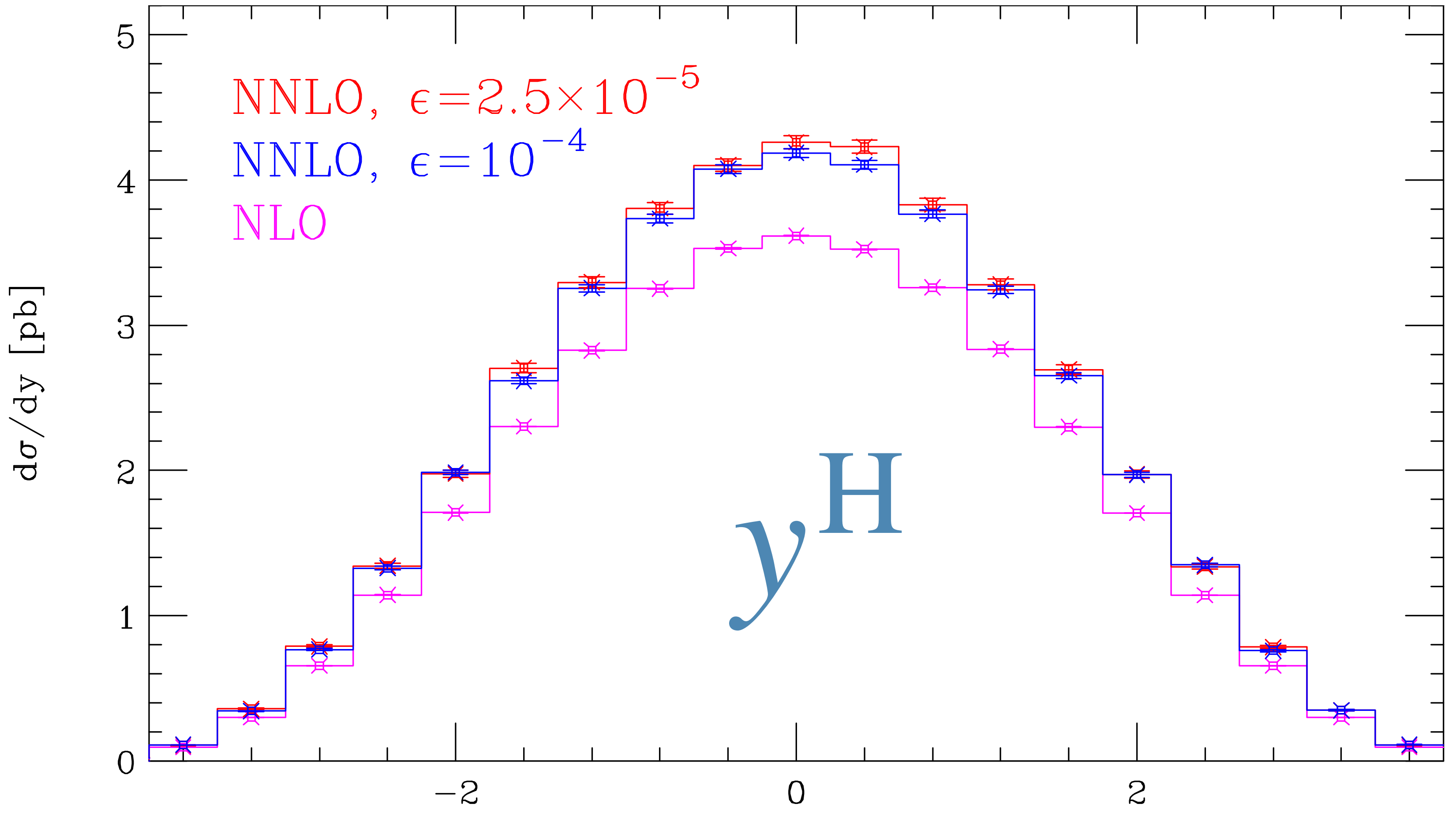
residue subtraction

[Caola, Melnikov, Schulze ’15]



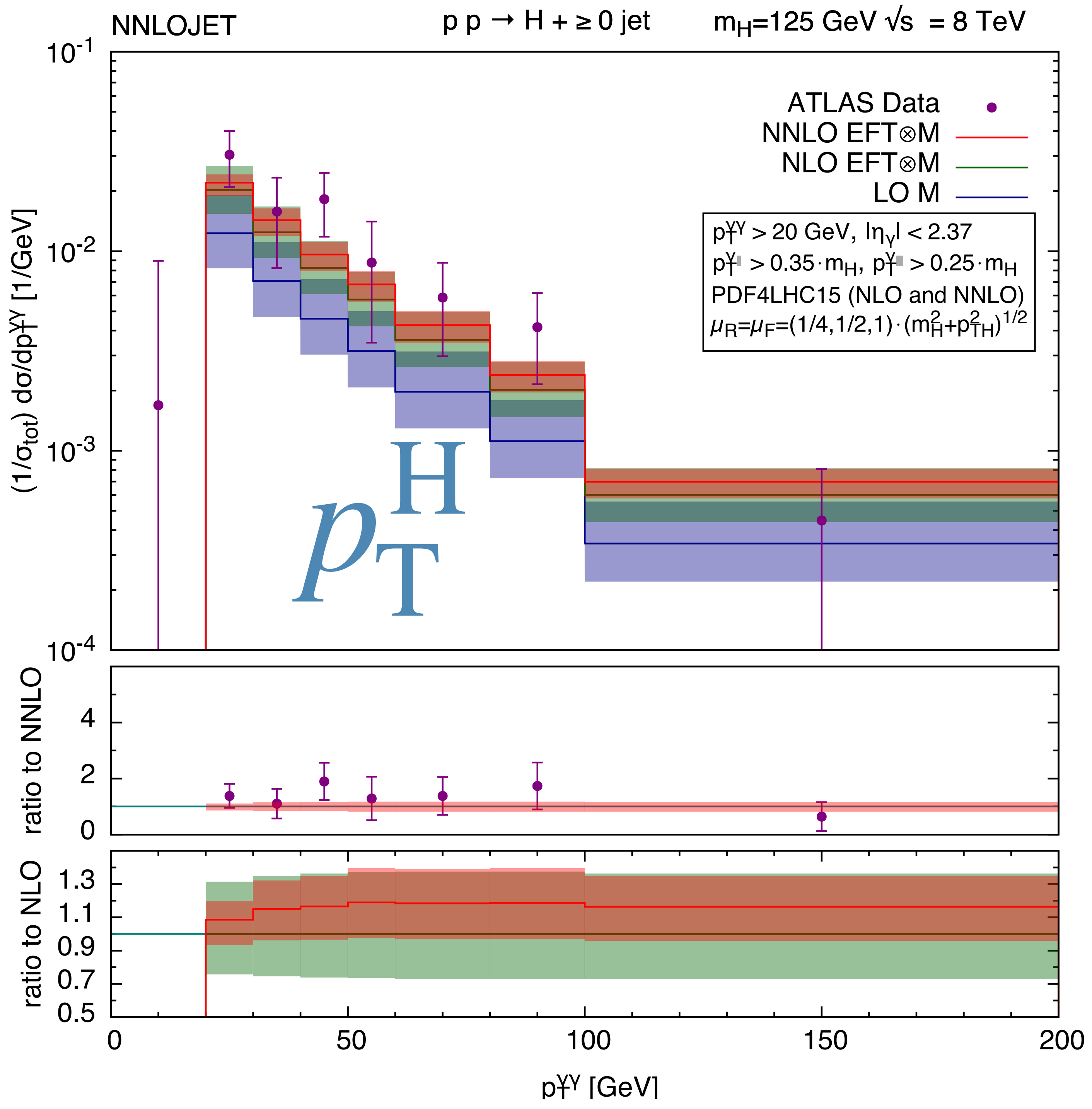
τ_1 jettiness subtraction

[Boughezal, Focke, Giele, Liu, Petriello ’15]
[Campbell, Ellis, Seth ’19]



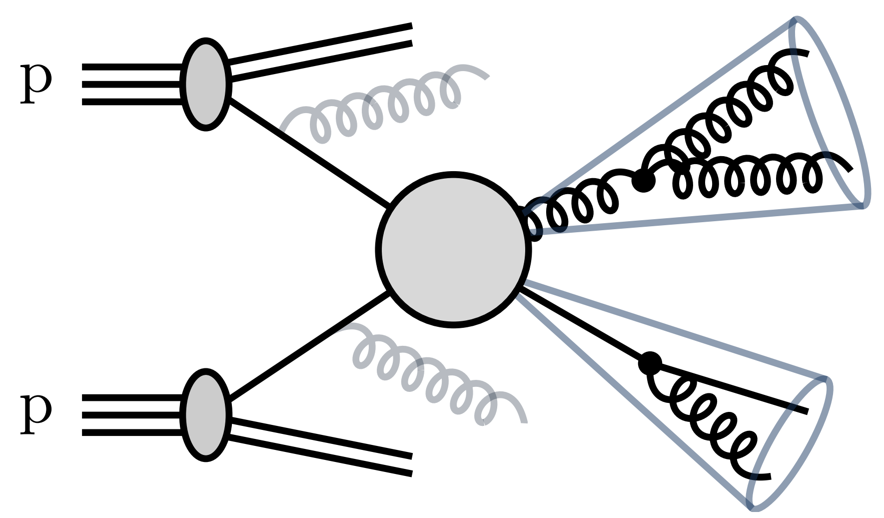
antenna subtraction

[Chen, Cruz-Martinez, Gehrmann, Glover, Jaquier ’16]



- very complex calculations ⇝ validation!
  - long-standing [~’15] discrepancy in H + jet
    ↪ only resolved in [’19]
- benchmark approaches

Jest are…

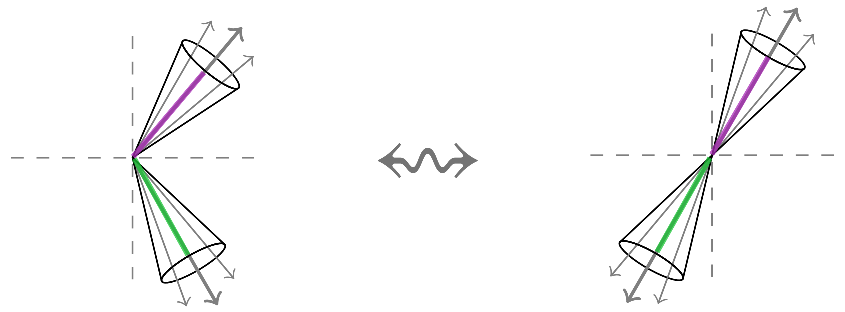


the ideal pQCD laboratory
simple 2 → 2 parton scattering

Constrain PDFs
(high-x gluon)

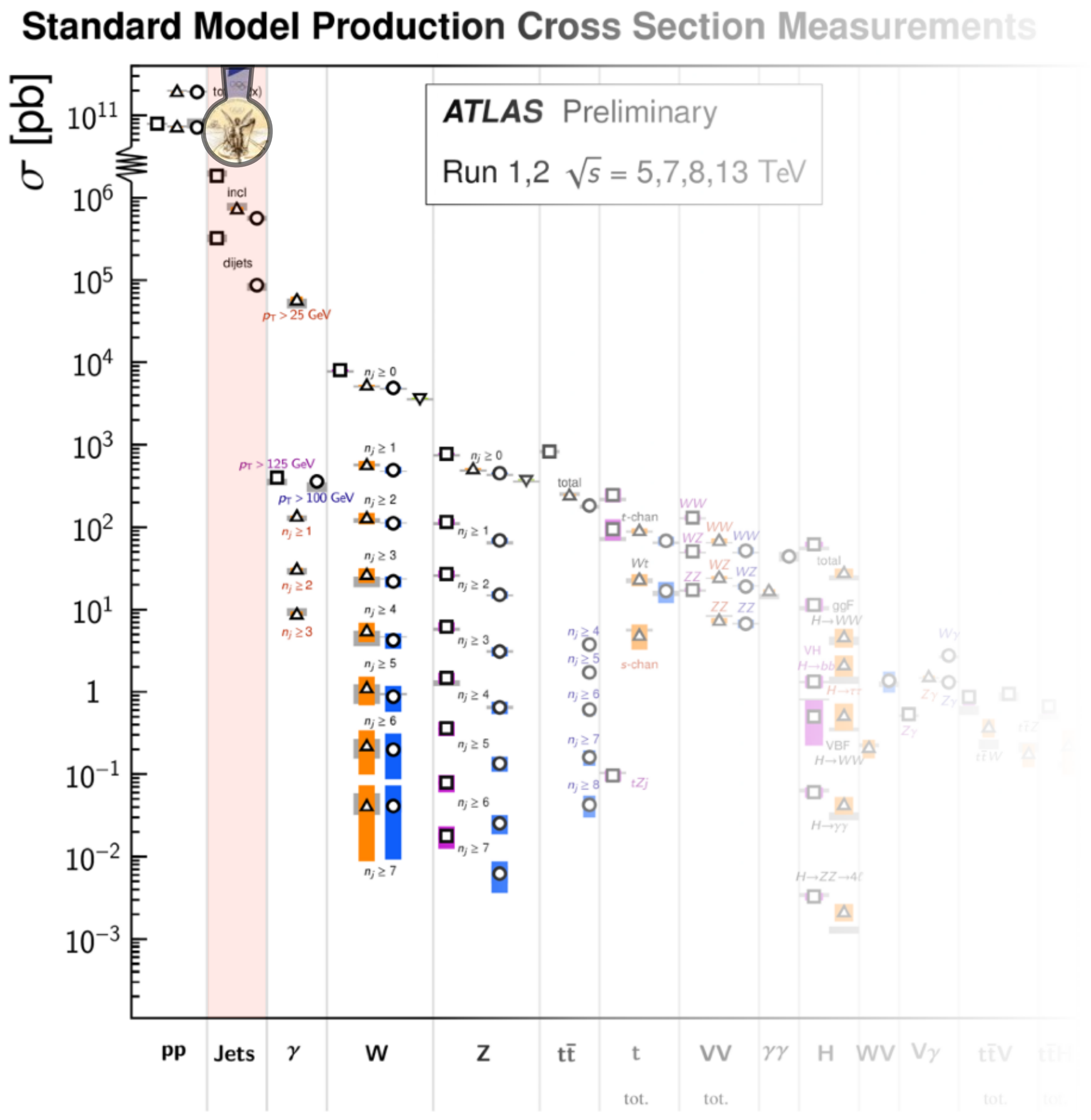
$$x = \frac{p_\mathrm{T}}{\sqrt{s}}\left(e^{\pm y_j} + e^{\pm y_{j'}}\right)$$

- @ LO: 3 variables $(p_\mathrm{T}, y_j, y_{j'})$
- inclusive jet[2] (some smearing)

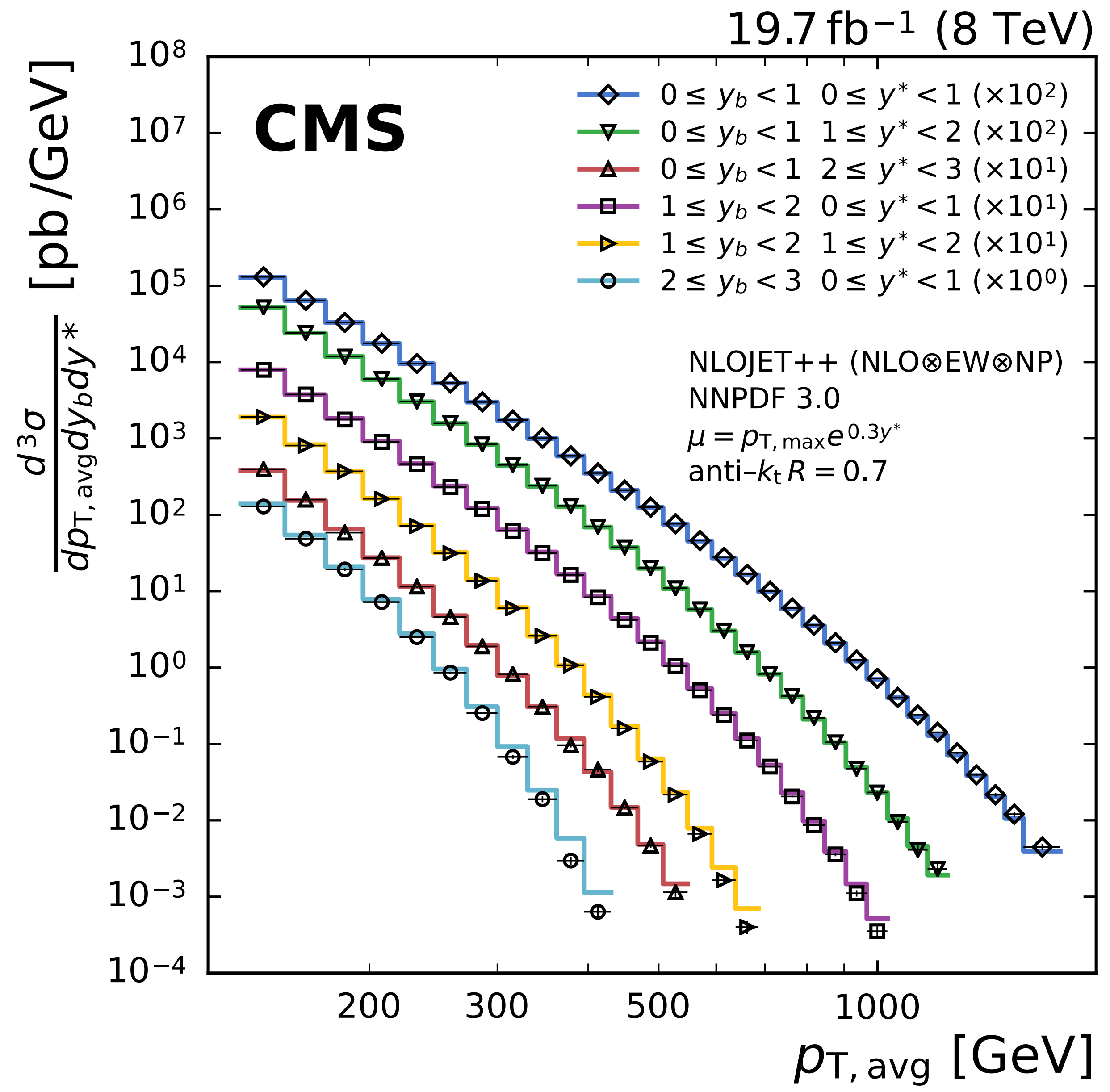


- di-jet[3] (reconstructible: 3-D)

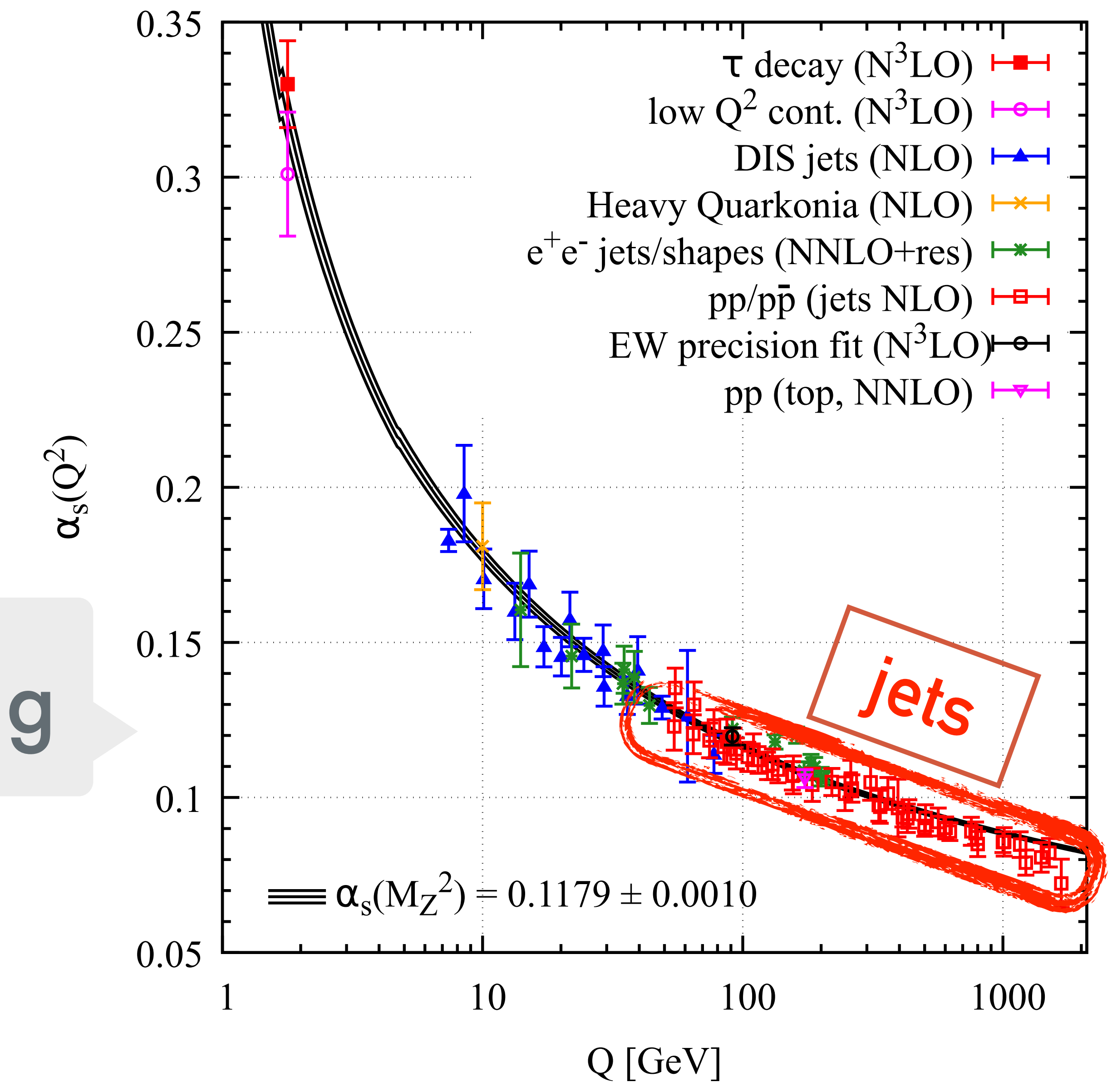
produced in *abundance*



wide kinematical range



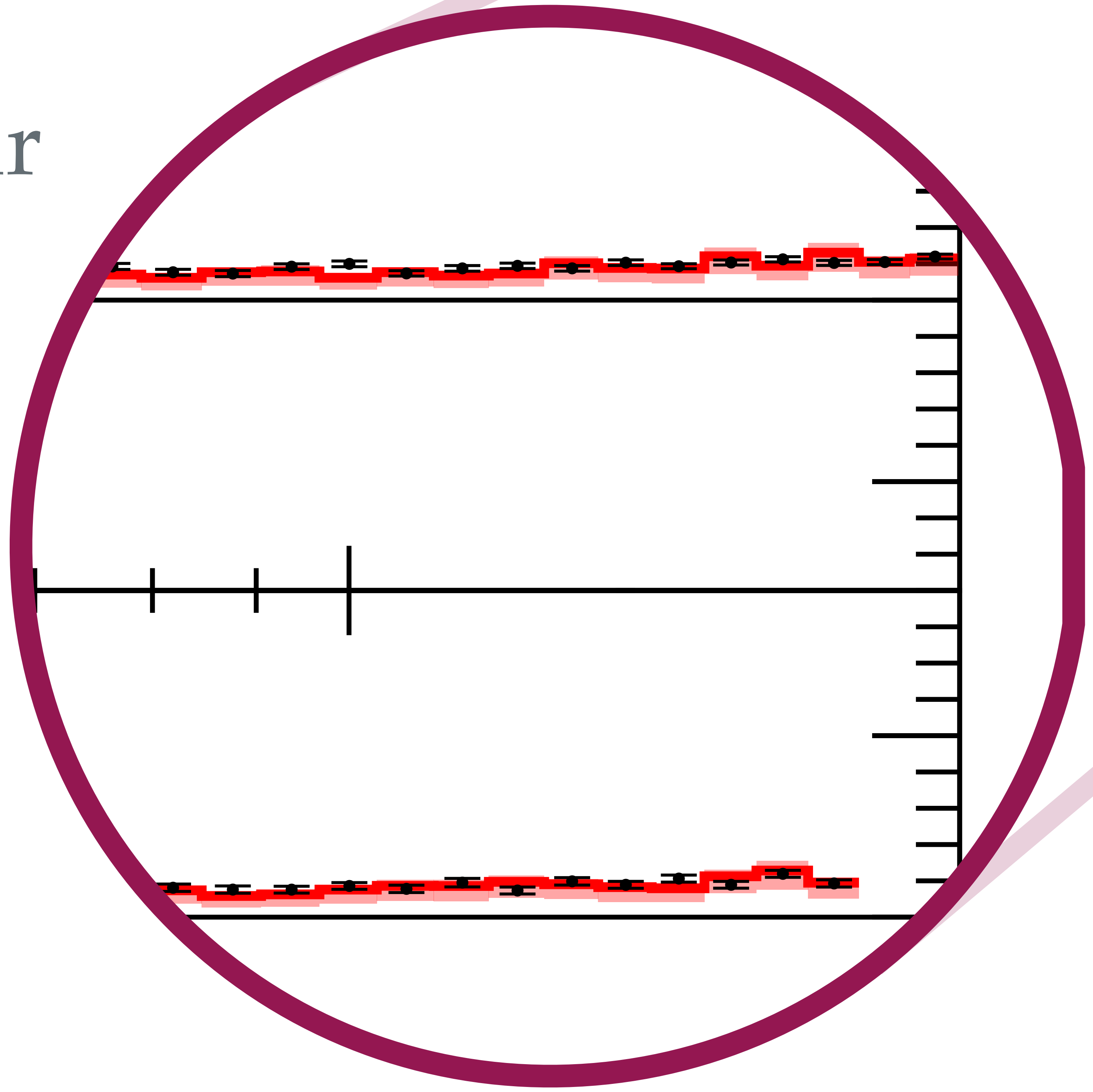
α_s & running

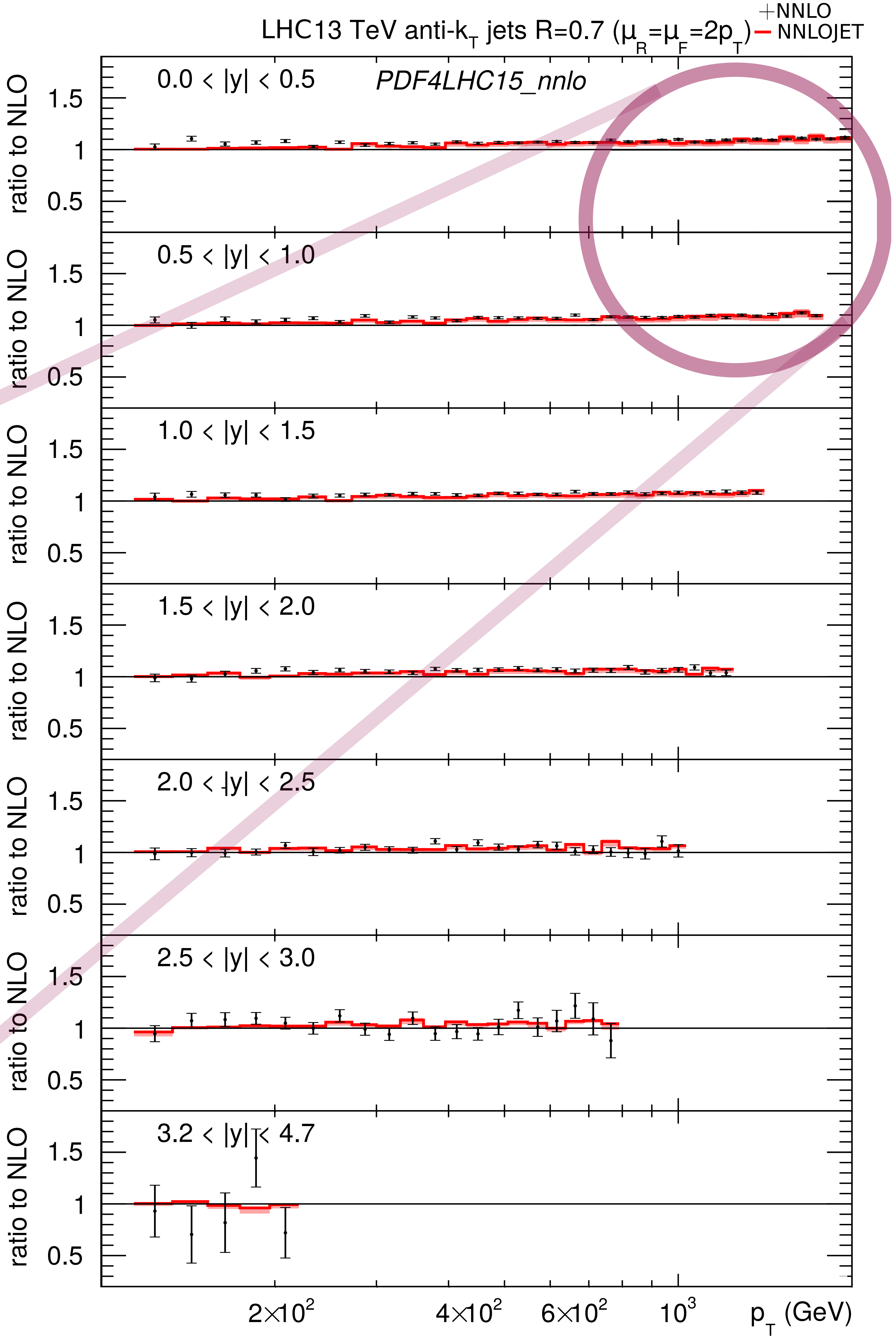


Inclusive Jets – 2 Calculations!

- NNLOJET [Currie, Glover, Pires '16]
- STRIPPER [Czakon, van Hameren, Mitov, Poncelet '19]

- in very good agreement!
- sub-leading colour negligible!(?)



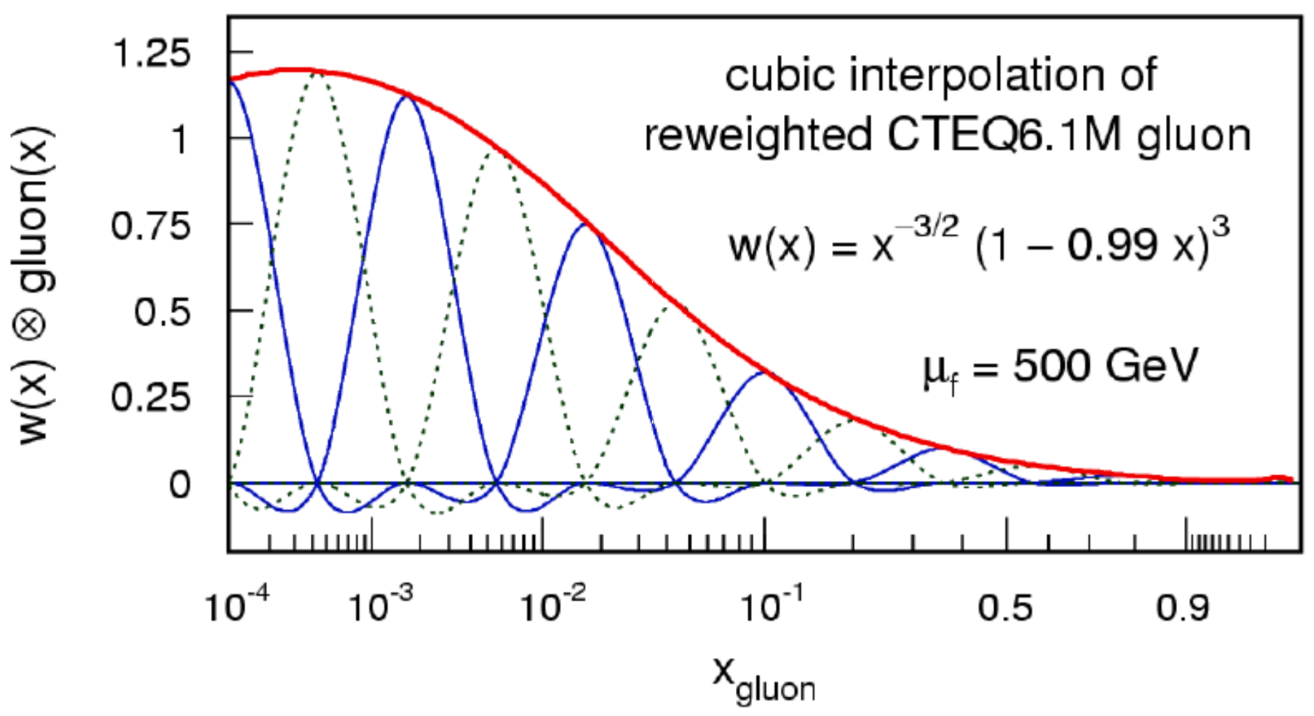


[APPLgrid, fastNLO, NNLOJET `19, `22]

- NNLO calculations $\mathcal{O}$(100k) CPU hours ⇝ prohibitive in PDF & α_s fits!
  - ↪ approximate the costly convolution using a grid:

sum: cheap!

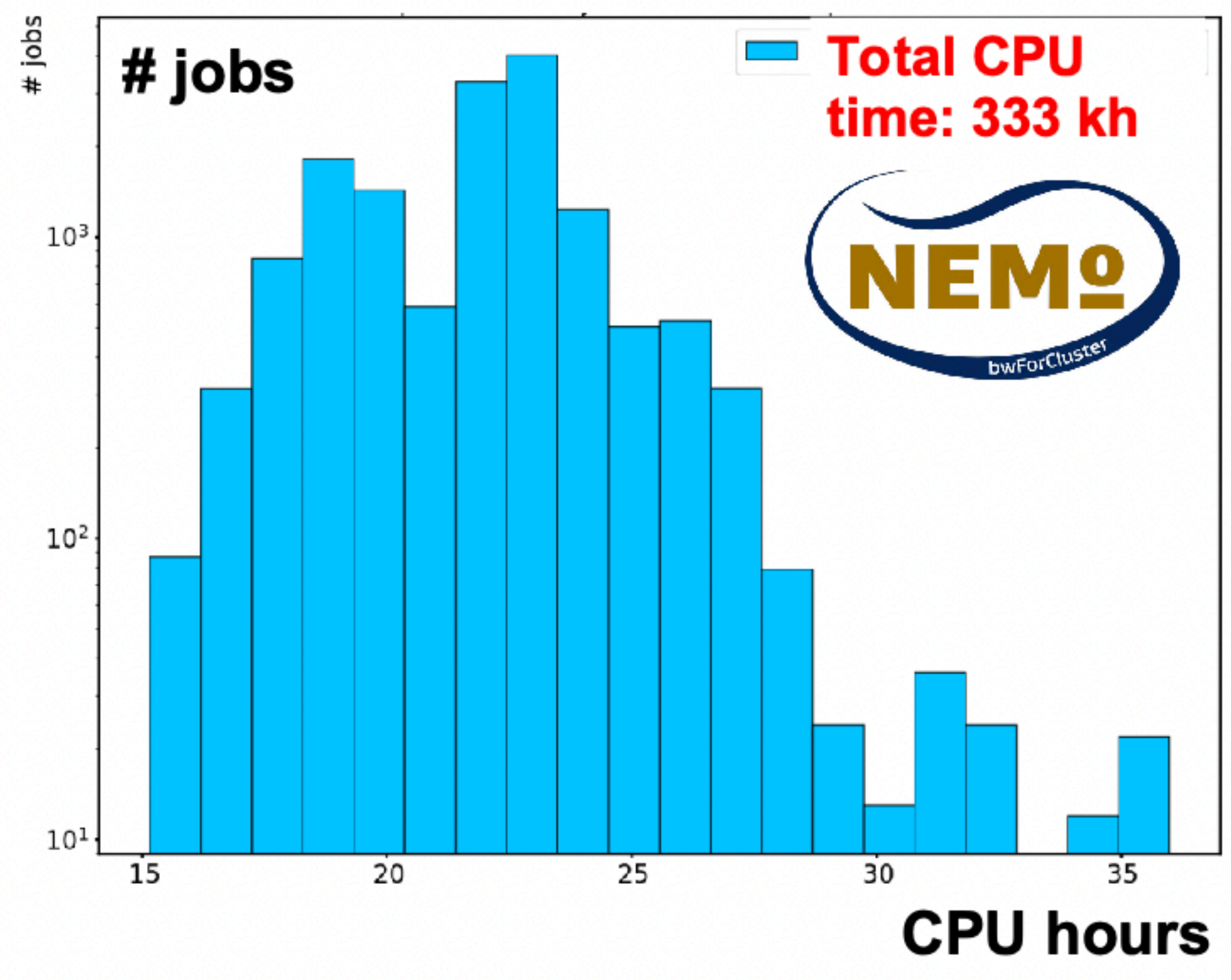
$$\sigma = \int_0^1 \mathrm{d}x\ f_a(x)\ \alpha_s^n\ \hat{\sigma}_a(x) \simeq \sum_i f_a(x^{(i)})\ \alpha_s^n \underbrace{\left[\int_0^1 \mathrm{d}x\ \hat{\sigma}_a(x)\ E^{(i)}(x)\right]}_{\hat{\sigma}_a^{(i)}}$$



$$f_a(x) \simeq \sum_i f_a(x^{(i)})\ E^{(i)}(x)$$

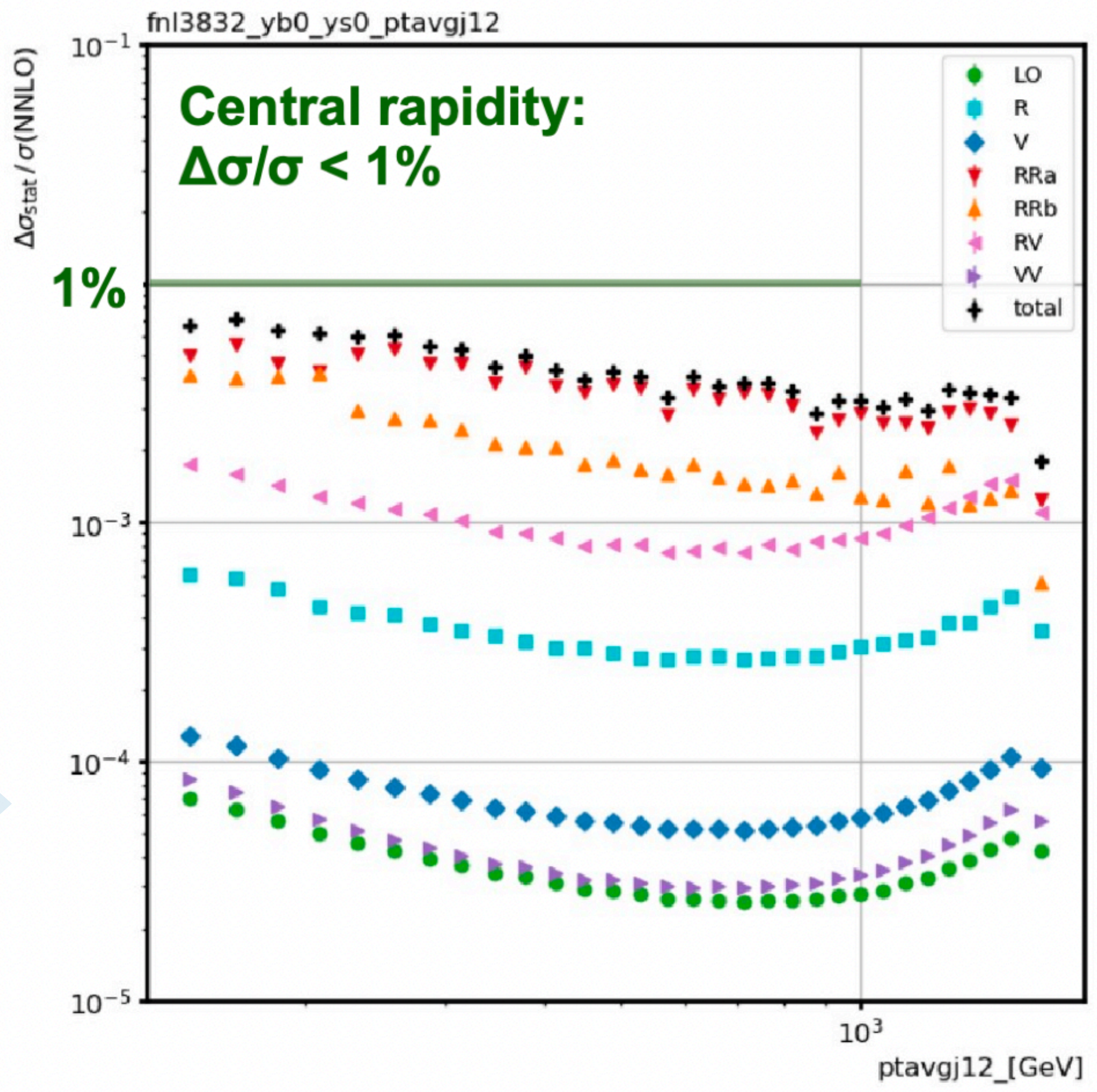
THE INVESTMENT

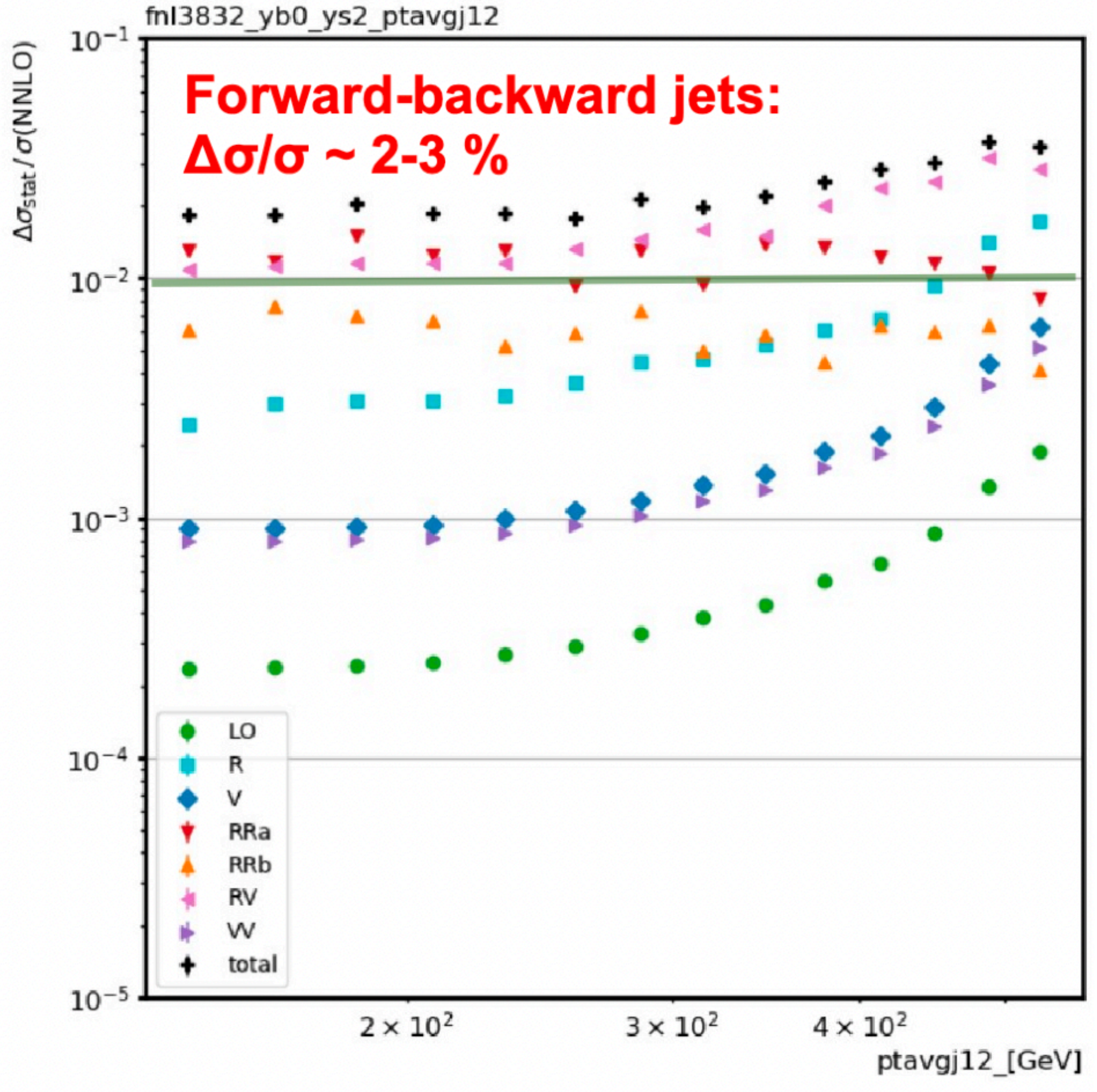
[APPLgrid, fastNLO, NNLOJET `22]

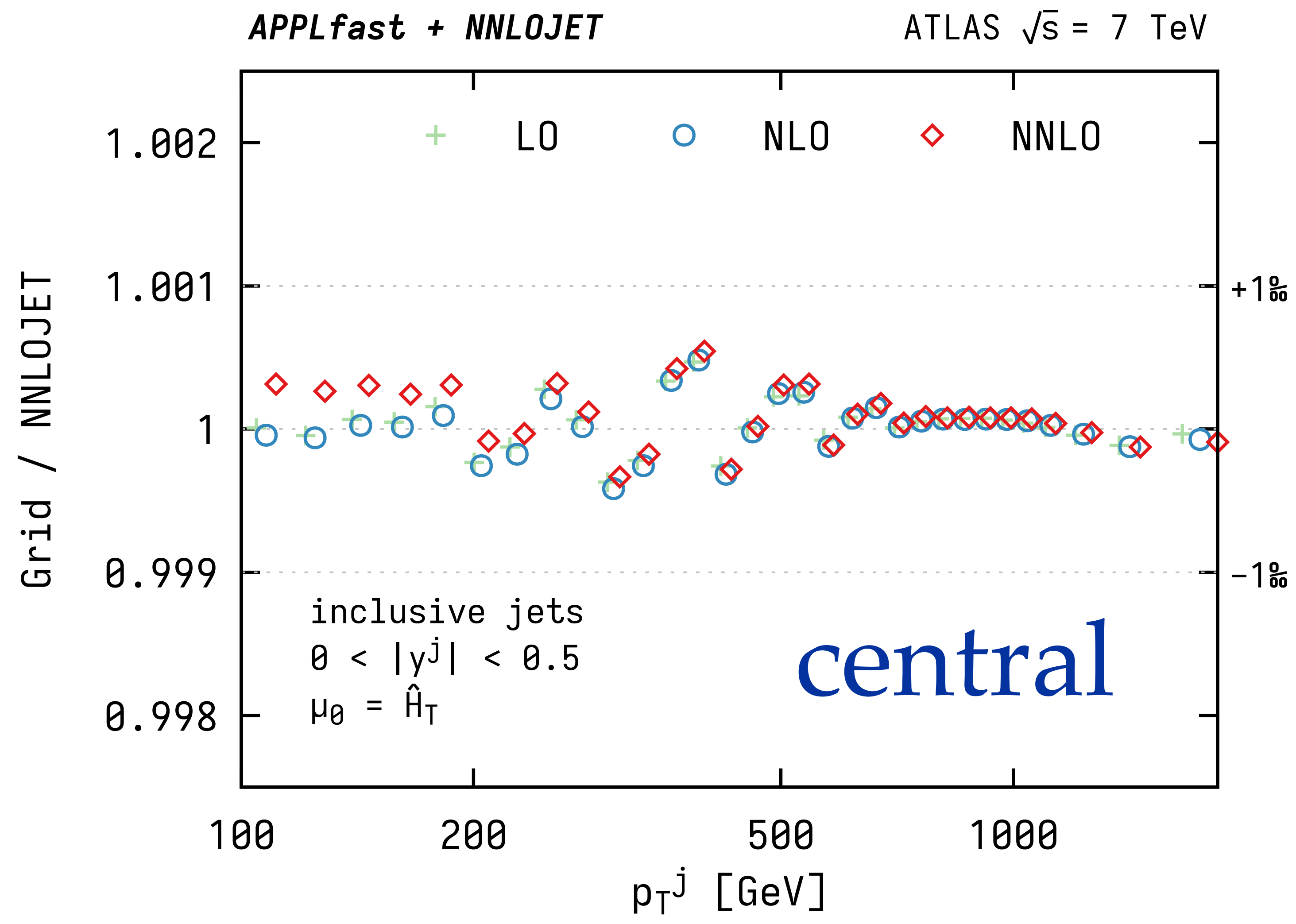


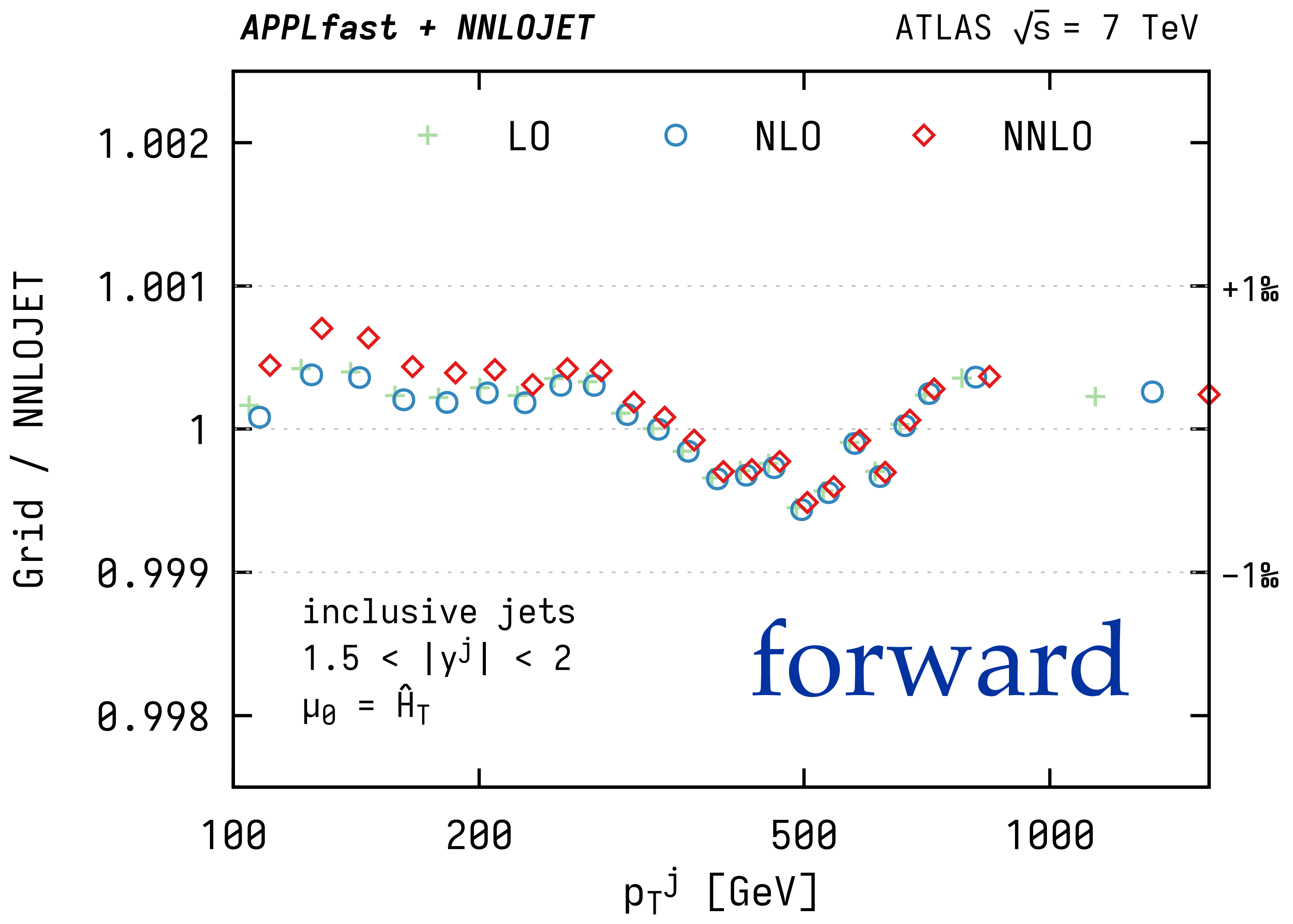
grid generation
⇝ runtime ~ ×2–3 (fastNLO)
⇝ storage ;-(

MC uncert. $\lesssim$ few %





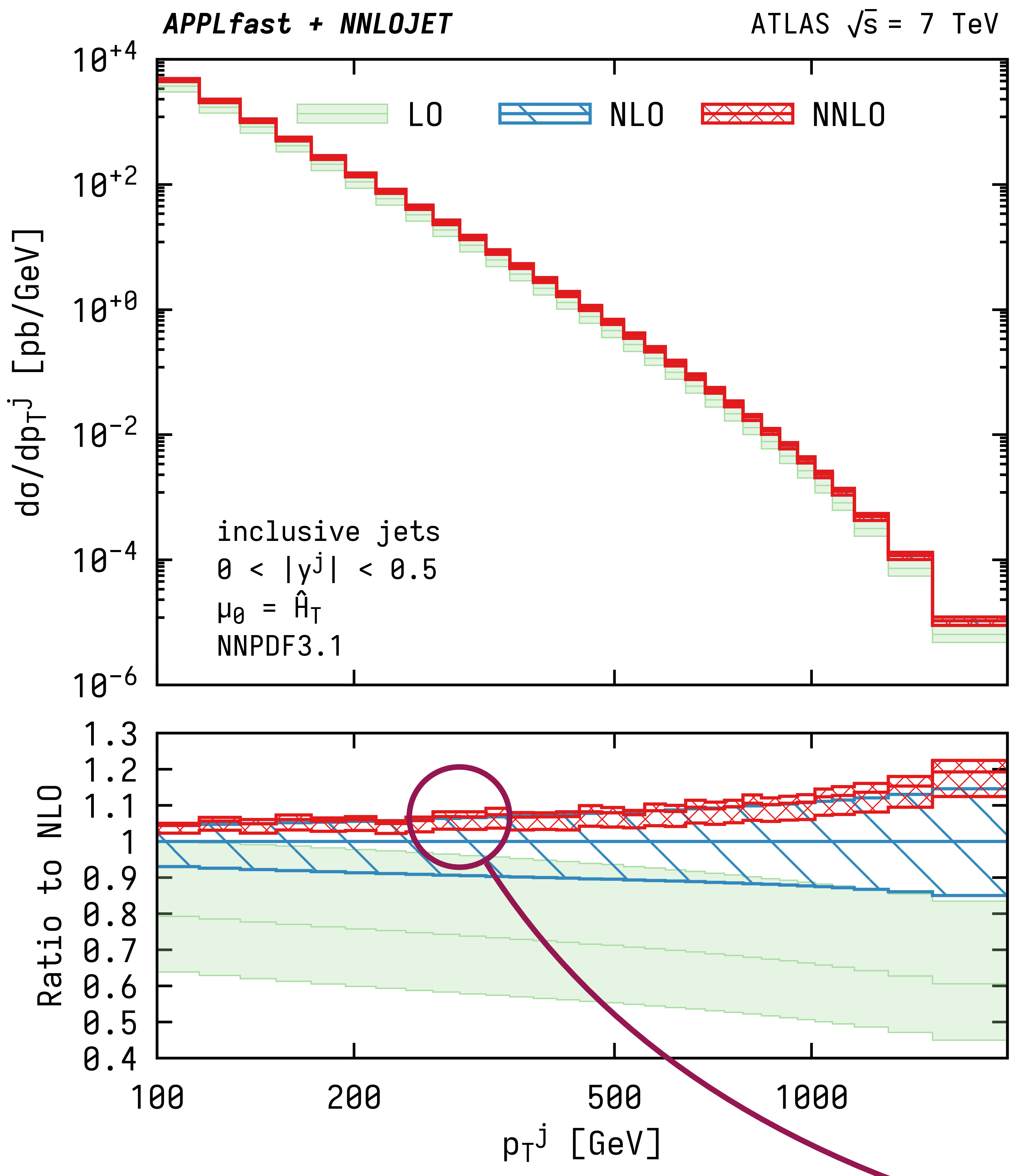




- closure tests
  grid evaluation vs.
  "vanilla" NNLOJET
  ↪ interp. bias $\lesssim 0.1\,\%$

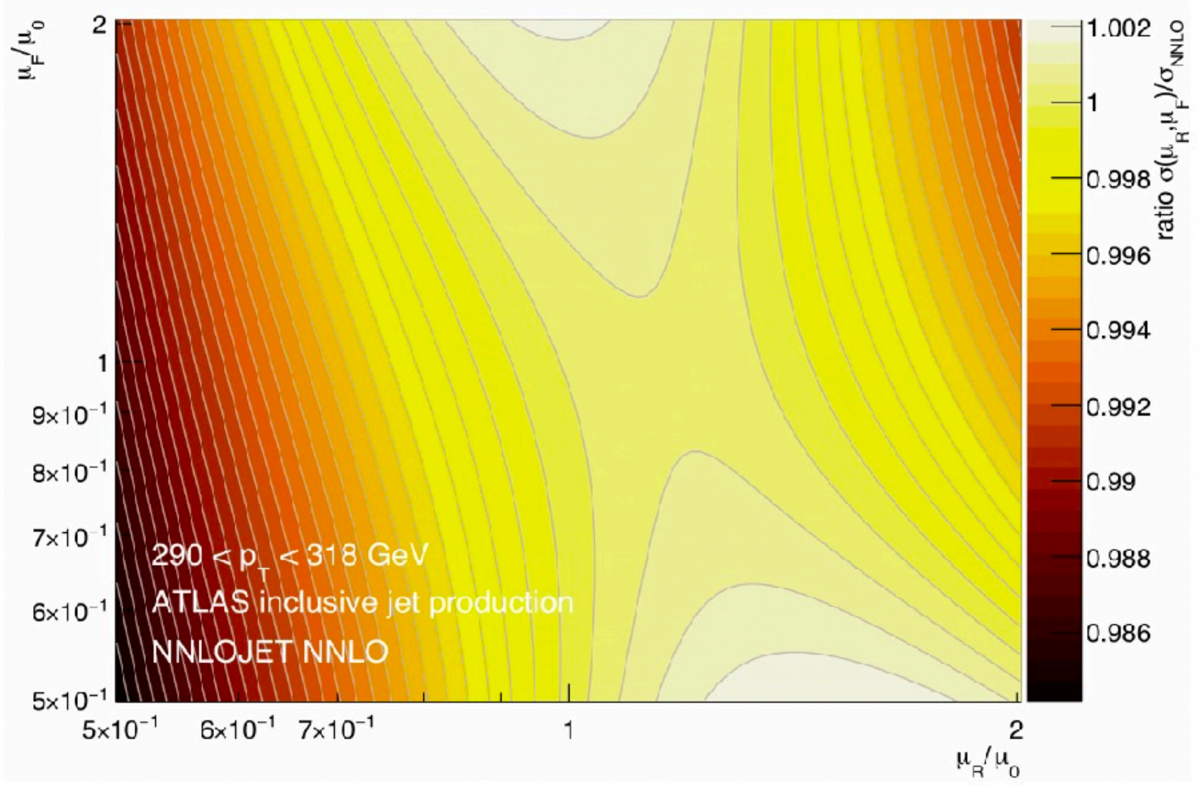
THE INVESTMENT & RETURN

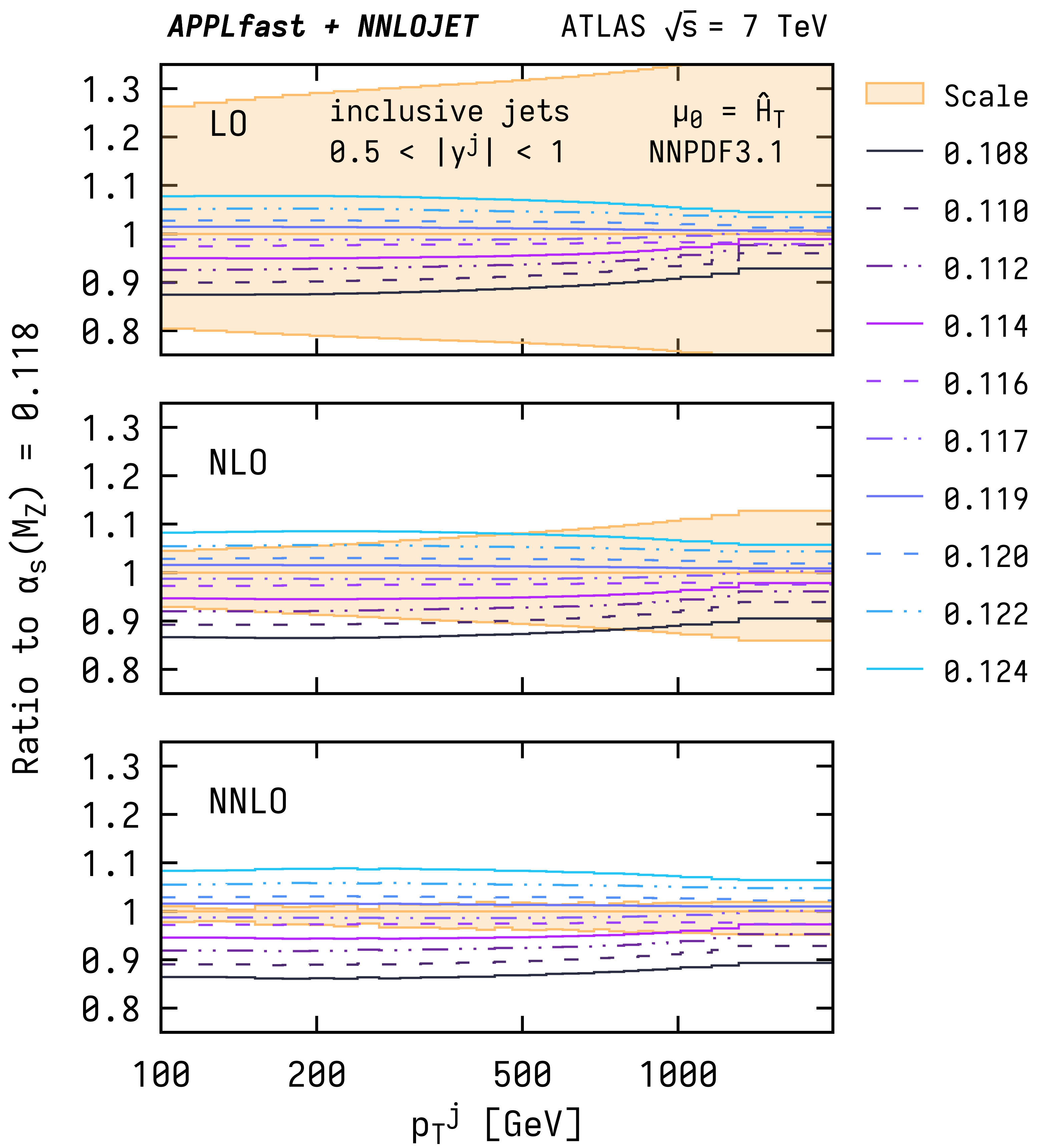
[APPLgrid, fastNLO, NNLOJET `22]



scale dependence
vs.
$\alpha_s \in [0.108, 0.124]$

2D (μ_R, μ_F) dependence

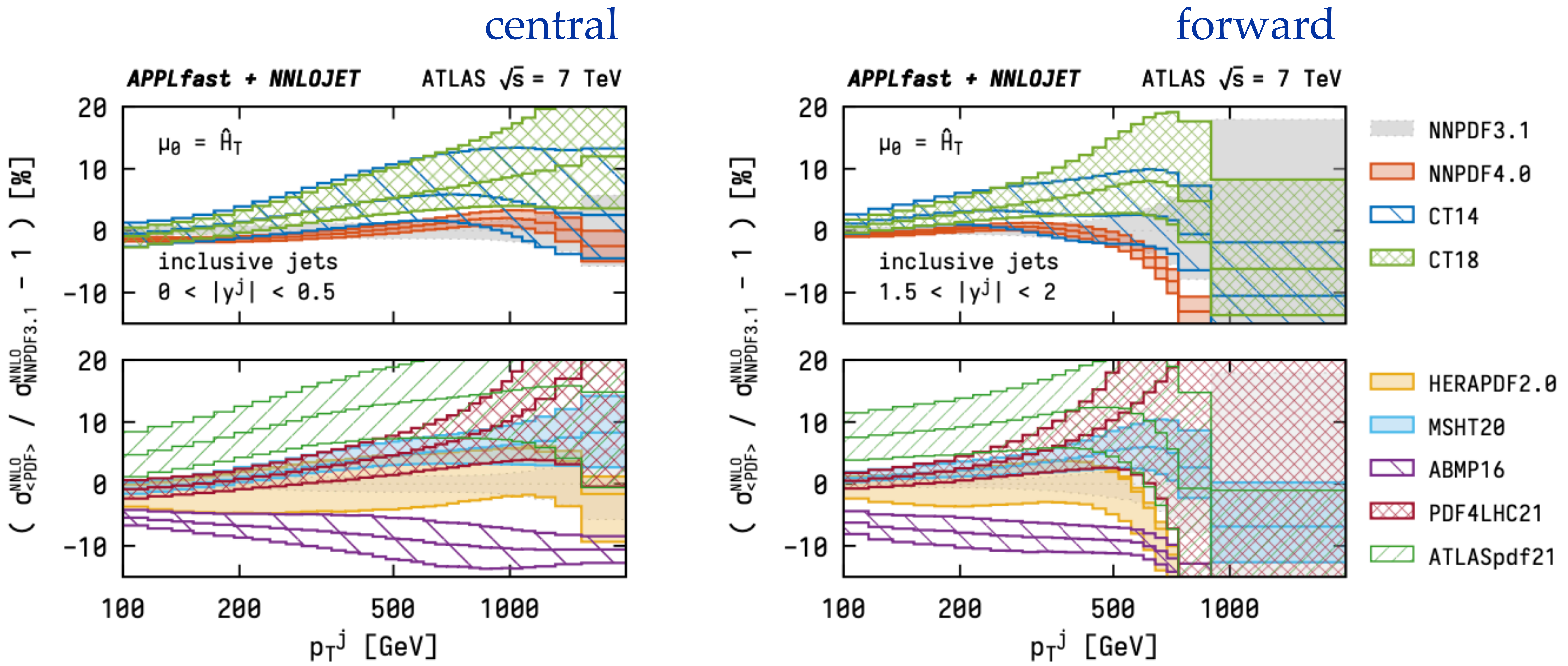




- predictions obtained in seconds!

PDF Dependence & Uncertainties

[APPLgrid, fastNLO, NNLOJET `22]



- ABMP16 & ATLASpdf21 largest excursion from the rest of the "pack"
- extremely small NNPDF4.0 PDF errors

VALIDITY OF K-FACTORS

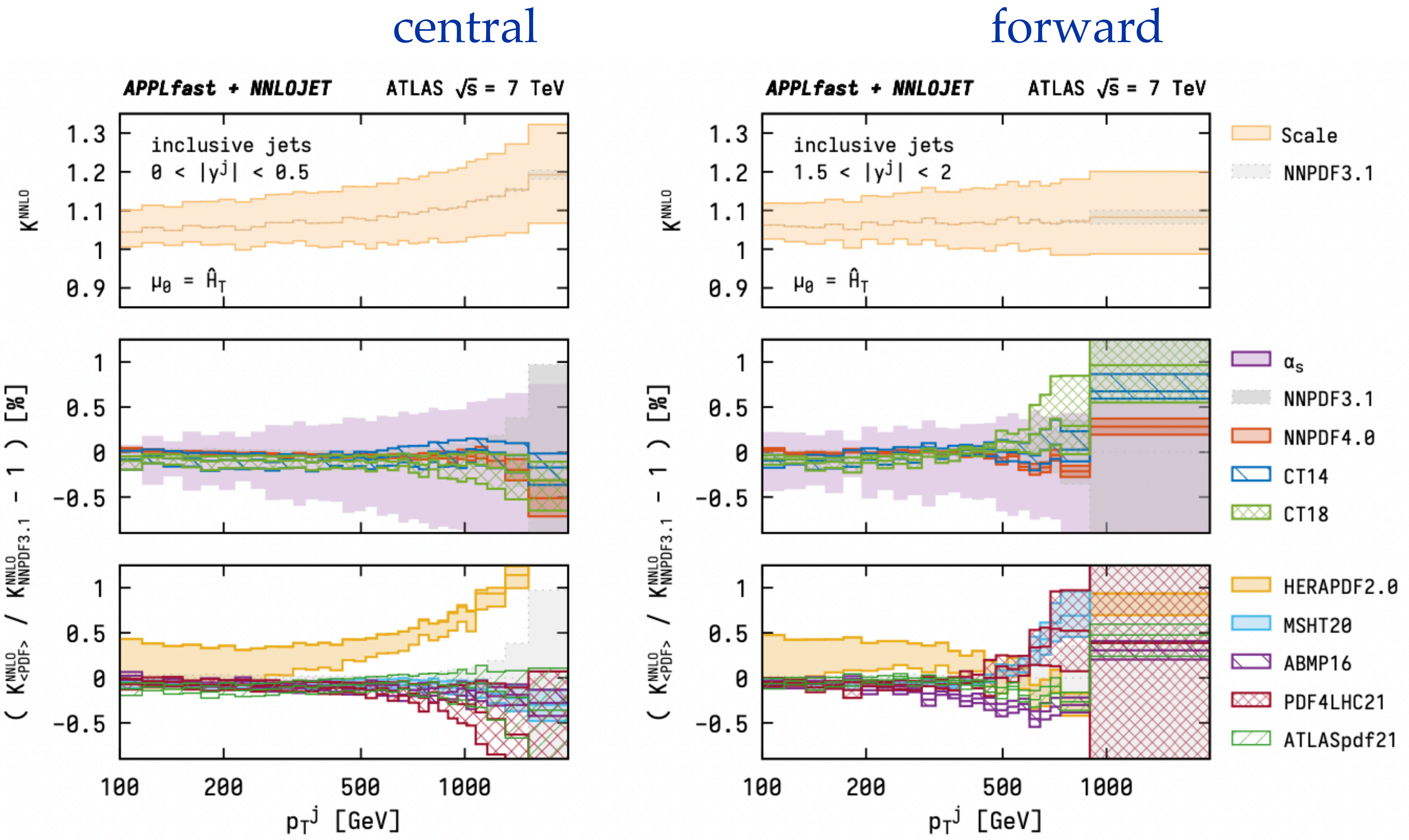
[APPLgrid, fastNLO, NNLOJET `22]

$$K^{\mathrm{NNLO}}(\mu) \equiv \frac{\mathrm{d}\sigma^{\mathrm{NNLO}}(\mu)/\mathrm{d}p_{\mathrm{T}}}{\mathrm{d}\sigma^{\mathrm{NLO}}(\mu)/\mathrm{d}p_{\mathrm{T}}}$$

$$\sigma^{\mathrm{NNLO}}_{\mathrm{approx.\,1}}(\mu) = \sigma^{\mathrm{NLO}}(\mu) \times K^{\mathrm{NNLO}}(\mu_{\mathrm{ref}})$$
$$\sigma^{\mathrm{NNLO}}_{\mathrm{approx.\,2}}(\mu) = \sigma^{\mathrm{NLO}}(\mu) \times K^{\mathrm{NNLO}}(\mu)\,,$$

- K-factor must be applied with correlated scales to avoid $\mathcal{O}(10\%)$ scale unc.
- extremely robust ($\lesssim 0.5\,\%$) w.r.t. PDF choice!

(exception: HERAPDF2.0)



Available Grids Tables

[APPLgrid, fastNLO, NNLOJET `22]

inclusive jets

Data	$\sqrt{s}$ [TeV]	$\mathcal{L}$ [fb^{-1}]	no. of points	anti-k_T R	kinematic range [GeV]	fiducial cuts	$\mu_{R/F}$-choice
CMS [30]	2.76	0.00543	81	0.7	$p_T^{jet} \in [74, 592]$	$\|y\| < 3.0$	p_T^{jet}, $\hat{H}_T$
ATLAS [28]	7.0	4.5	140	0.6	$p_T^{jet} \in [100, 1992]$	$\|y\| < 3.0$	p_T^{jet}, $\hat{H}_T$
CMS [31]	7.0	5.0	133	0.7	$p_T^{jet} \in [114, 2116]$	$\|y\| < 3.0$	p_T^{jet}, $\hat{H}_T$
ATLAS [32]	8.0	20.3	171	0.6	$p_T^{jet} \in [70, 2500]$	$\|y\| < 3.0$	p_T^{jet}, $\hat{H}_T$
CMS [33]	8.0	5.6 19.7	248	0.7	$p_T^{jet} \in [21, 74]$ $p_T^{jet} \in [74, 2500]$	$\|y\| < 4.7$	p_T^{jet}, $\hat{H}_T$
ATLAS [34]	13.0	3.2	177	0.4	$p_T^{jet} \in [100, 3937]$	$\|y\| < 3.0$	p_T^{jet}, $\hat{H}_T$
CMS [35]	13.0	36.3 33.5	2×78	0.4 0.7	$p_T^{jet} \in [97, 3103]$	$\|y\| < 2.0$	p_T^{jet}, $\hat{H}_T$

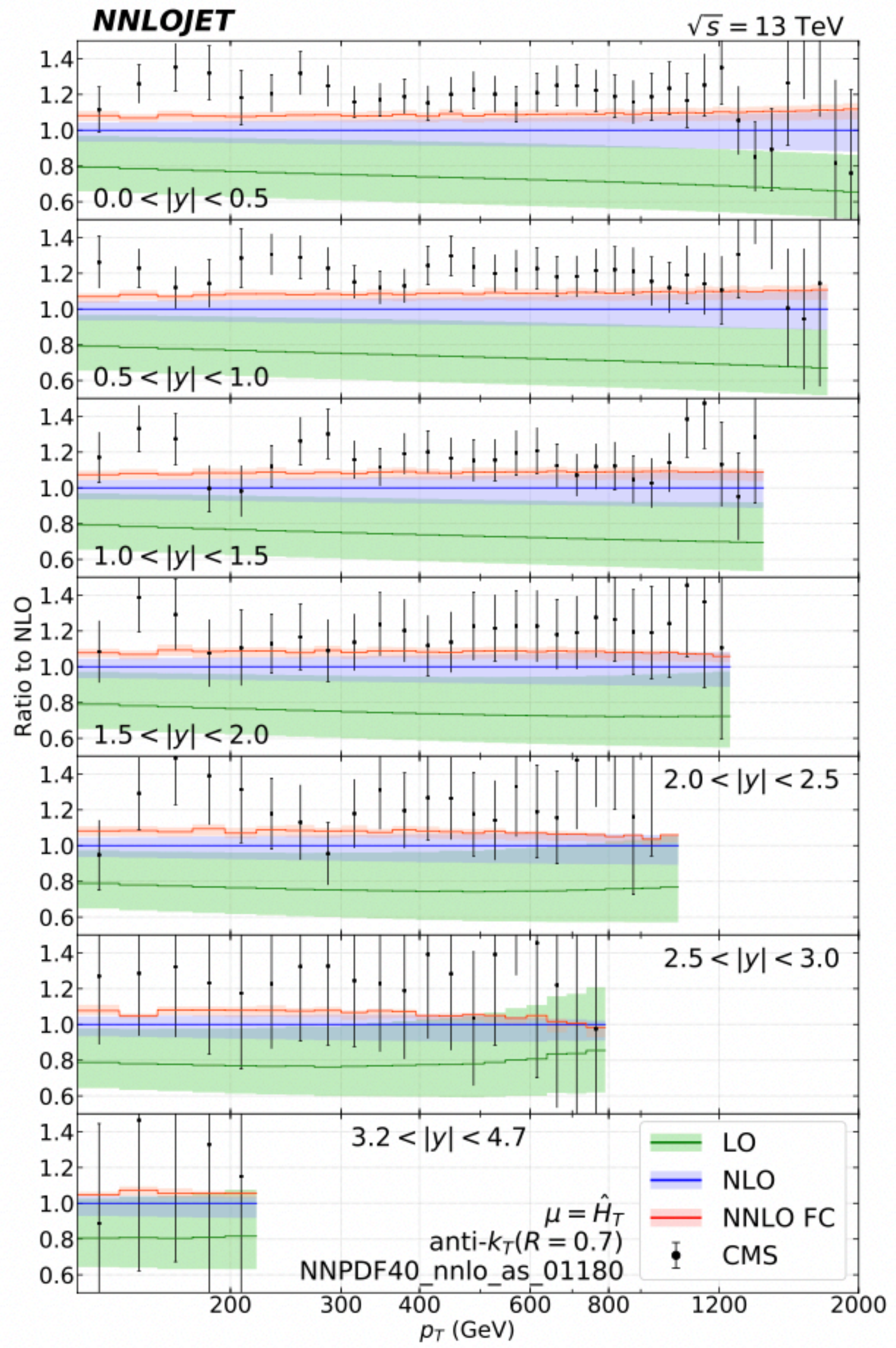
di-jets

Data	$\sqrt{s}$ [TeV]	$\mathcal{L}$ [fb^{-1}]	no. of points	anti-k_T R	kinematic range [GeV]	fiducial cuts	$\mu_{R/F}$-choice
ATLAS [55]	7.0	4.5	90	0.6	$m_{12} \in [260, 5040]$	$\|y_1\|, \|y_2\| < 3.0$ $[p_{T,1}, p_{T,2}] > [100, 50]$GeV $y^* < 3.0$	m_{12}
CMS [31]	7.0	5.0	54	0.7	$m_{12} \in [197, 5058]$	$\|y\| < 5.0$ $[p_{T,1}, p_{T,2}] > [60, 30]$GeV $\|y_{max}\| < 2.5$	m_{12}
CMS [49]	8.0	19.7	122	0.7	$\langle p_{T1,2} \rangle \in [133, 1784]$	$\|y\| < 5.0$ $p_{T,1}, p_{T,2} > 50$GeV $\|y_1\|, \|y_2\| < 3.0$	$p_{T,1} \exp(0.3\, y^*)$ m_{12}
ATLAS [34]	13.0	3.2	136	0.4	$m_{12} \in [260, 9066]$	$\|y_1\|, \|y_2\| < 3.0$ $p_{T,1}, p_{T,2} > 75$GeV $\langle p_{T1,2} \rangle > 100$GeV $y^* < 3.0$	m_{12}

- all grids available on: `ploughshare.web.cern.ch`
- *caveat:* calculation based on leading-colour approximation in NNLO parts
  ⇝ leading: N_c^2, $N_c n_f$, n_f^2 (sub-leading: $\times 1/N_c^2$)

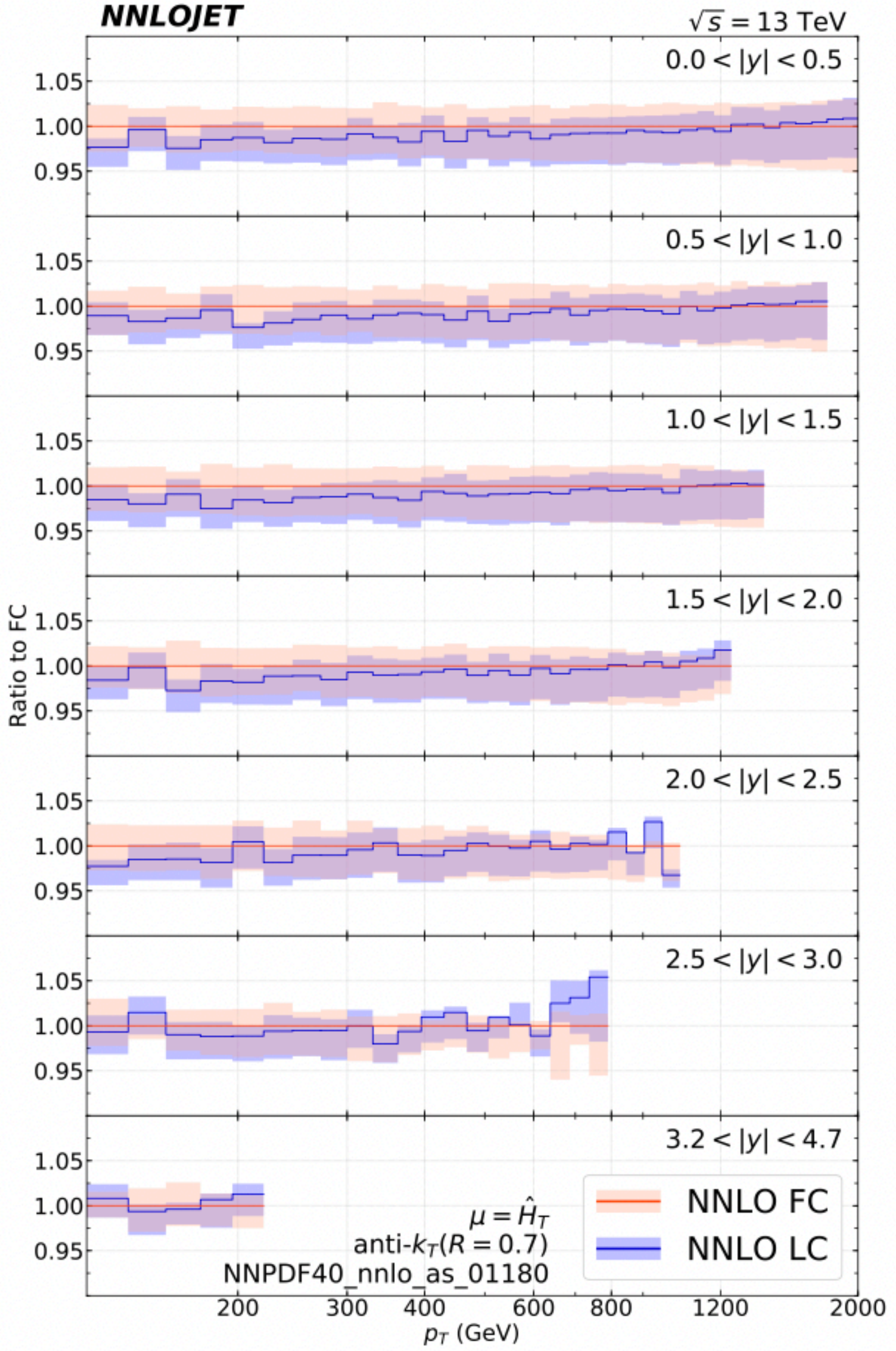
How Good is LC?

sub-leading colour: SLC
leading colour: LC
full colour (LC+SLC): FC



- improved agreement with the data
- +ve SLC contribution
  - ↪ up to 20% on $\delta\sigma^{\mathrm{NNLO}}$
  - ↪ largest @ low-p_{T}
  - ↪ diminishes @ high-p_{T}
- impact on NNLO:
  - ⇝ $\lesssim 2\,\%$
  - ⇝ within Δ_{scl}

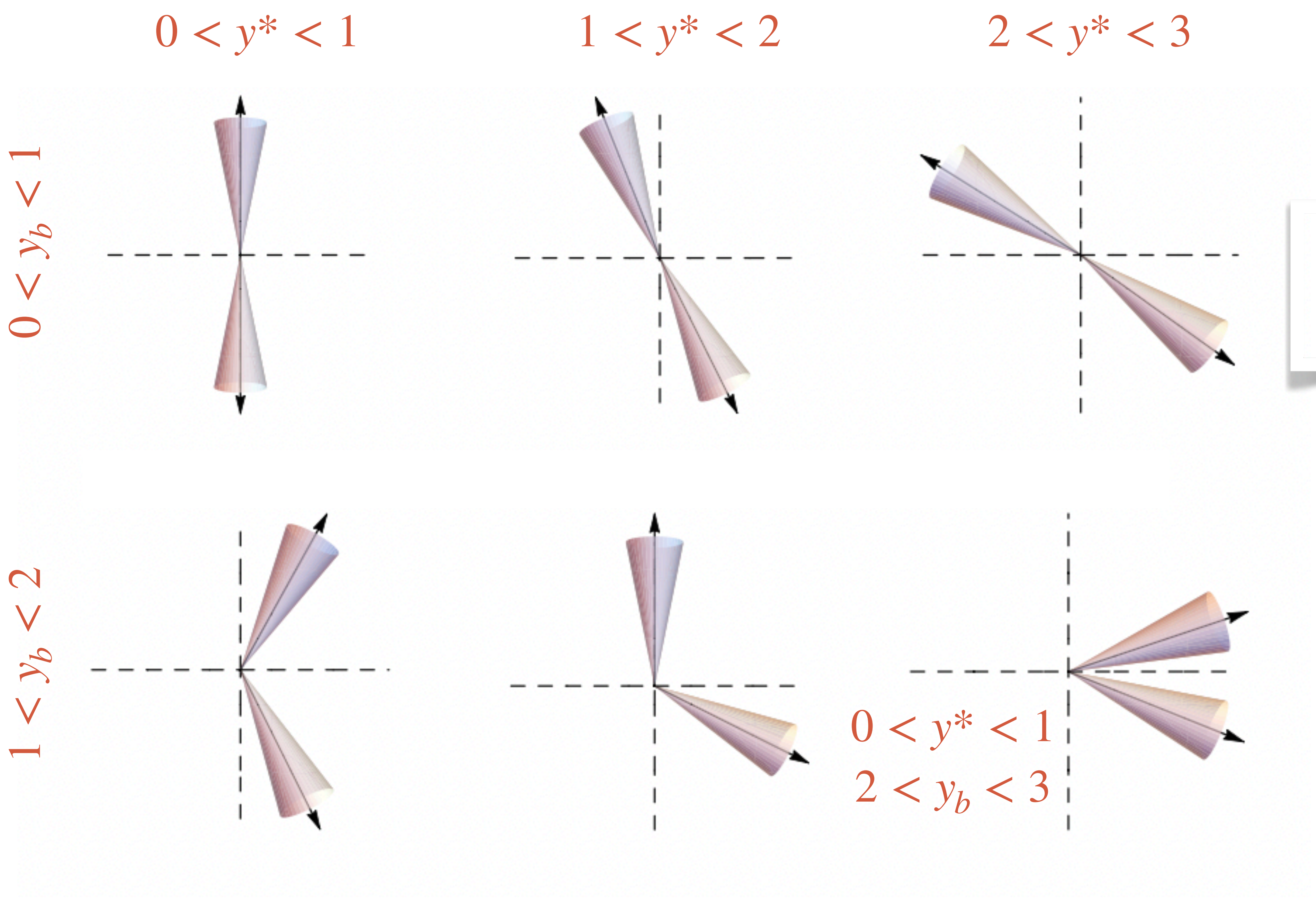
SLC small in incl. jets (R=0.4, 0.7)
still small on di-jet $\mathrm{d}\sigma/\mathrm{d}m_{jj}$ (R=0.4)
substantial in 3D di-jet (R=0.7)



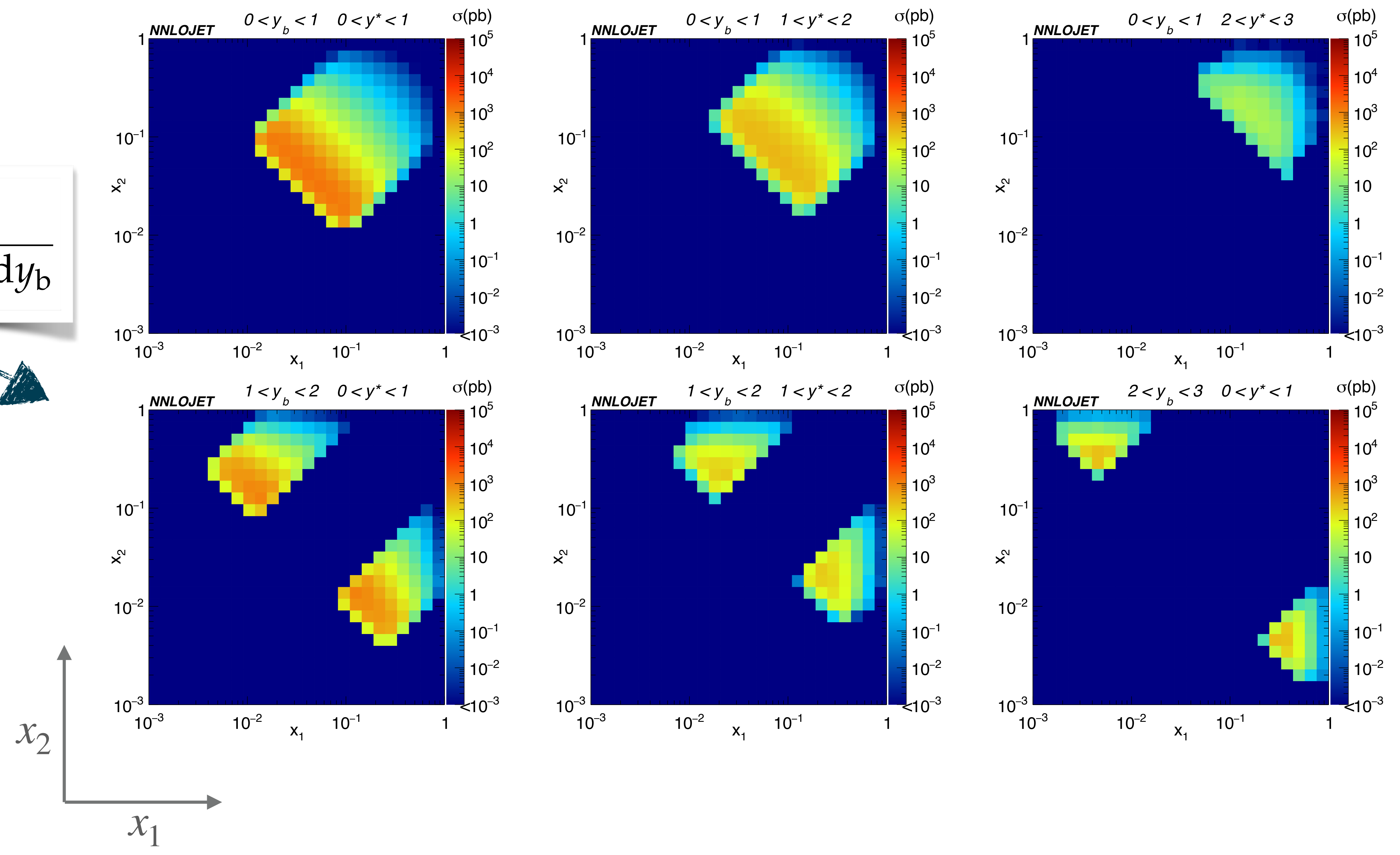
TRIPLY-DIFFERENTIAL DI-JET PRODUCTION

- different **event topologies**
  - ⇝ disentangle mom. fractions (x_1, x_2)

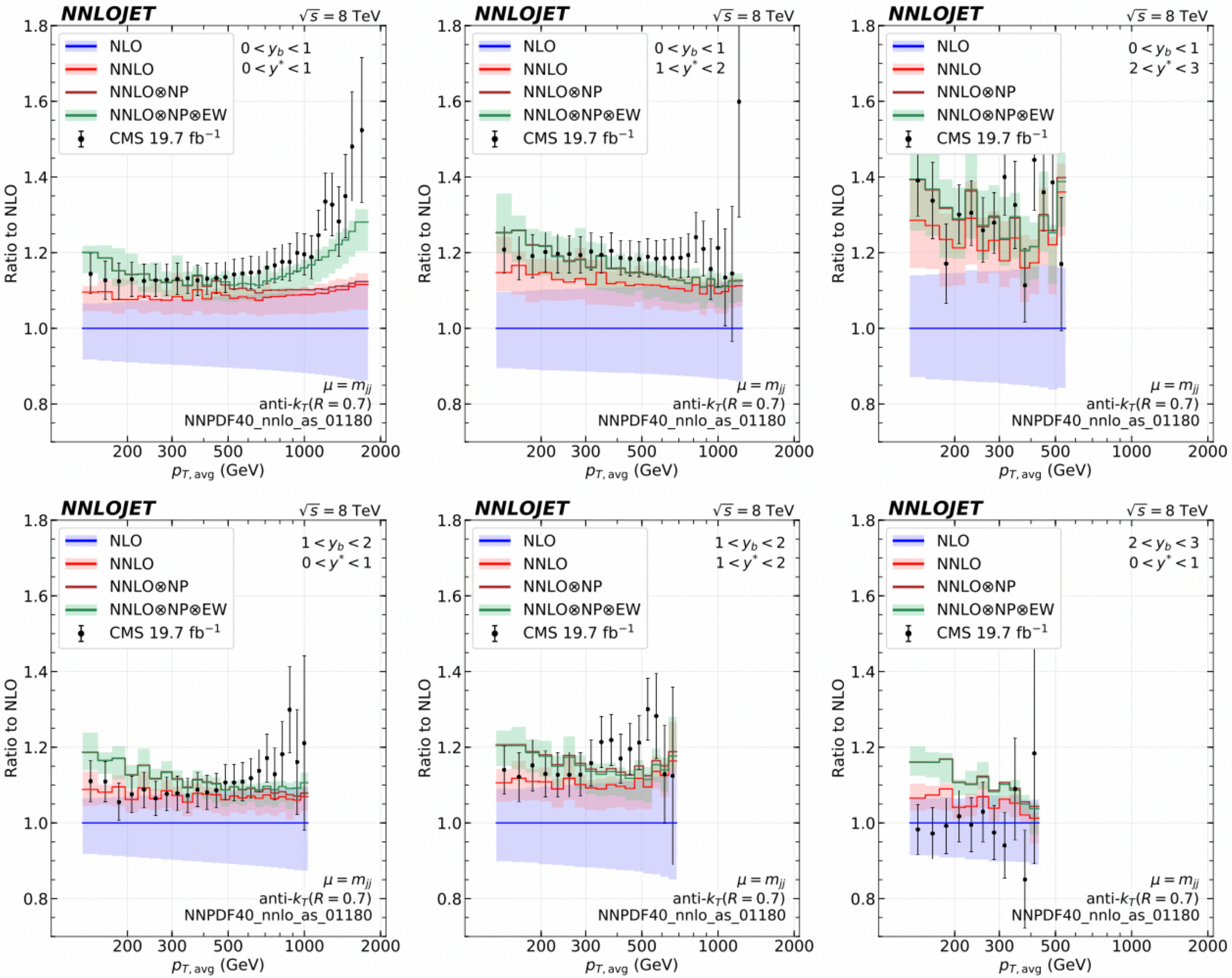
$$x_{1,2} = \frac{2p_{\mathrm{T,avg}}}{\sqrt{s}} e^{\pm y_b} \cosh(y^*)$$



$$\frac{\mathrm{d}^3\sigma}{\mathrm{d}p_{\mathrm{T,avg}}\mathrm{d}y^*\mathrm{d}y_{\mathrm{b}}}$$



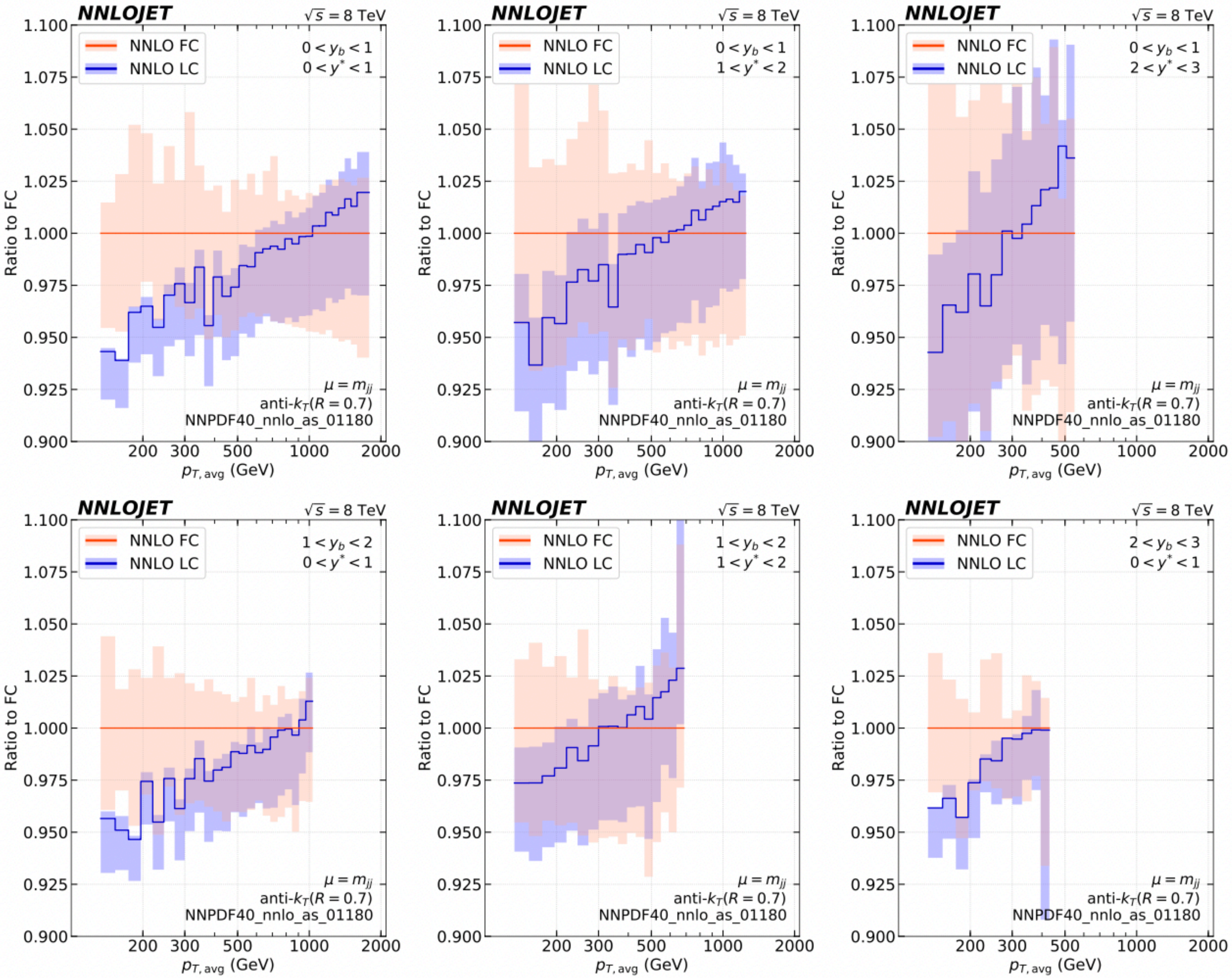
TRIPLY-DIFFERENTIAL DI-JET PRODUCTION — TH VS. DATA



[Chen, Gehrmann, Glover, AH, Mo '22]

- large NP corrections
  @ low-$p_{\mathrm{T,avg}}$
- EW corrections only impacts
  ⇝ high-$p_{\mathrm{T,avg}}$
  & $y_b, y^* < 1$
- improved description of data
  & reduced uncertainties

Triply-Differential Di-jet Production — FC vs. LC

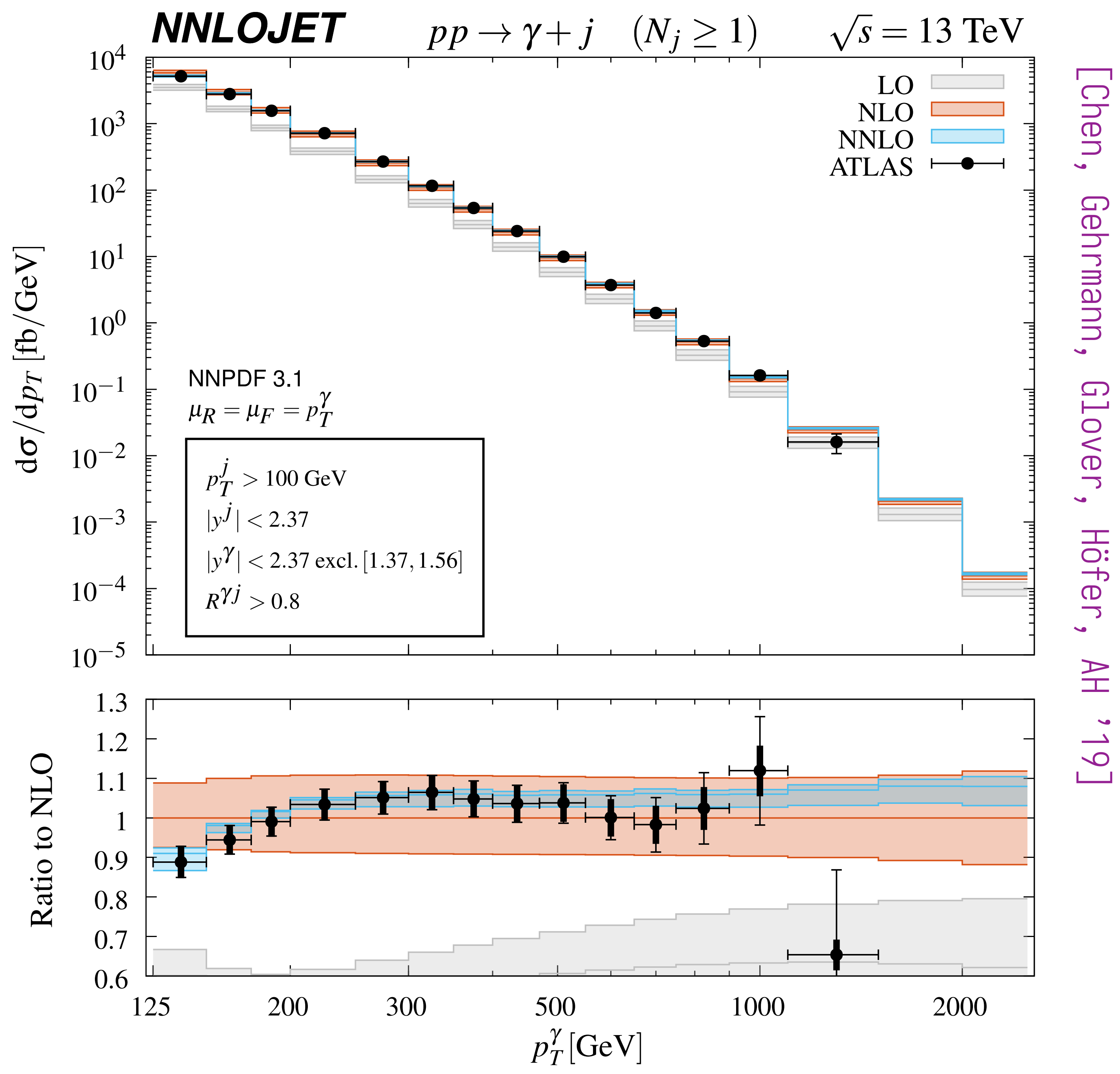


[Chen, Gehrmann, Glover, AH, Mo '22]

- large SLC contributions
  - ↪ low-$p_{T,avg}$ ⇝ 30–60%
  - ↪ med-$p_{T,avg}$ ⇝ small $|\cdot|$
  - ↪ high-$p_{T,avg}$ ⇝ $-20\,\%$
- LC → FC
  - ↪ $+5\,\%$ enhancement
- grids with FC very desirable!
  - ↪ resolve tension with other datasets? [NNPDF4.0]

Identified Objects — Challenges in TH vs. Exp

Isolated Photons $\gamma + \text{jet}$

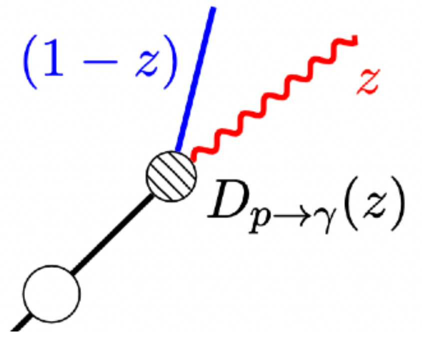


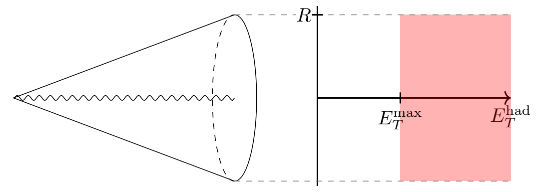
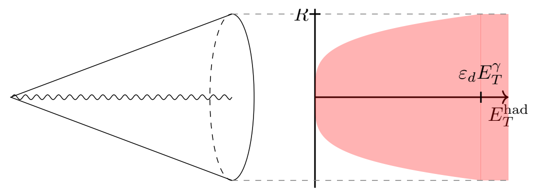
[Chen, Gehrmann, Glover, Höfer, AH '19]

"ok" for tight isol.

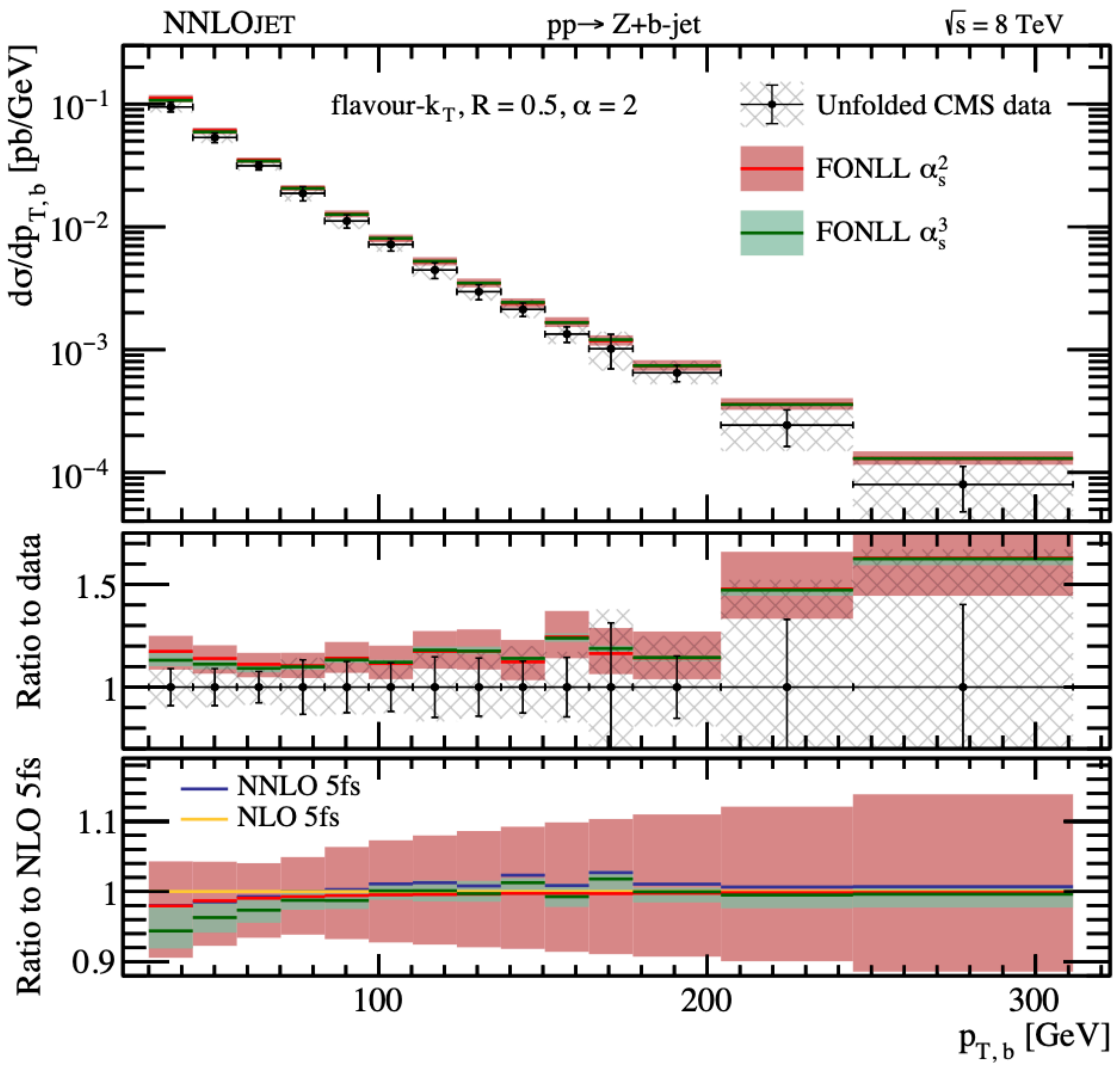
tune @ NLO?

per-cent mismatch

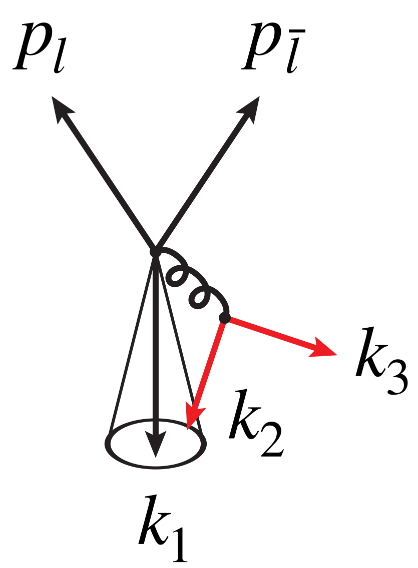


- **Exp:**  *fixed cone*
- **TH:**  *smooth cone*

Jet Flavour $Z + \text{b-jet}$



[Gauld, Gehrmann-De Ridder, Glover, AH, Majer '20]



- **Exp:** *anti-k_T*
- **TH:** *flavour-k_T*

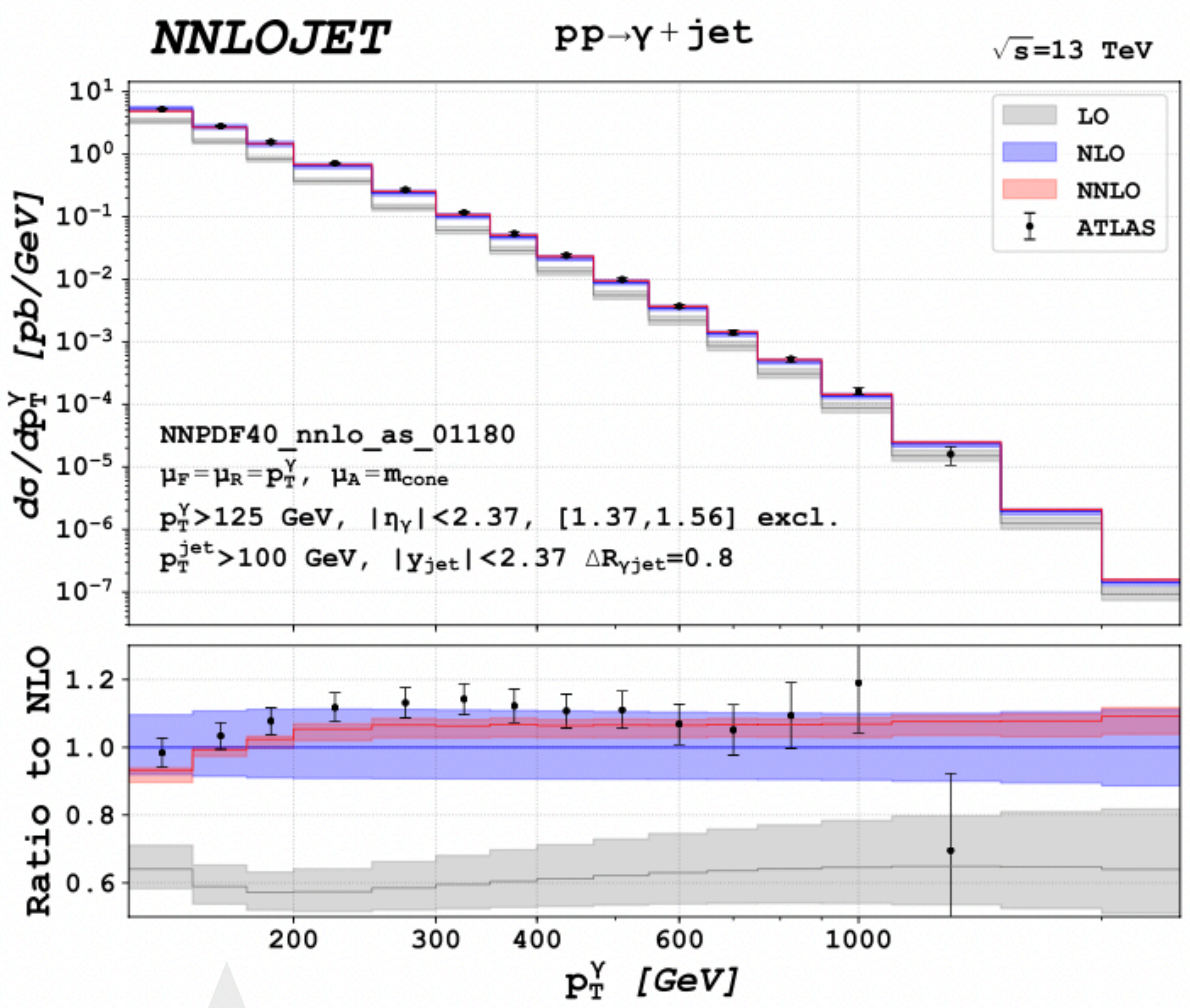
unfolding

$\mathcal{O}(10\%)$

⇝ Hadrons

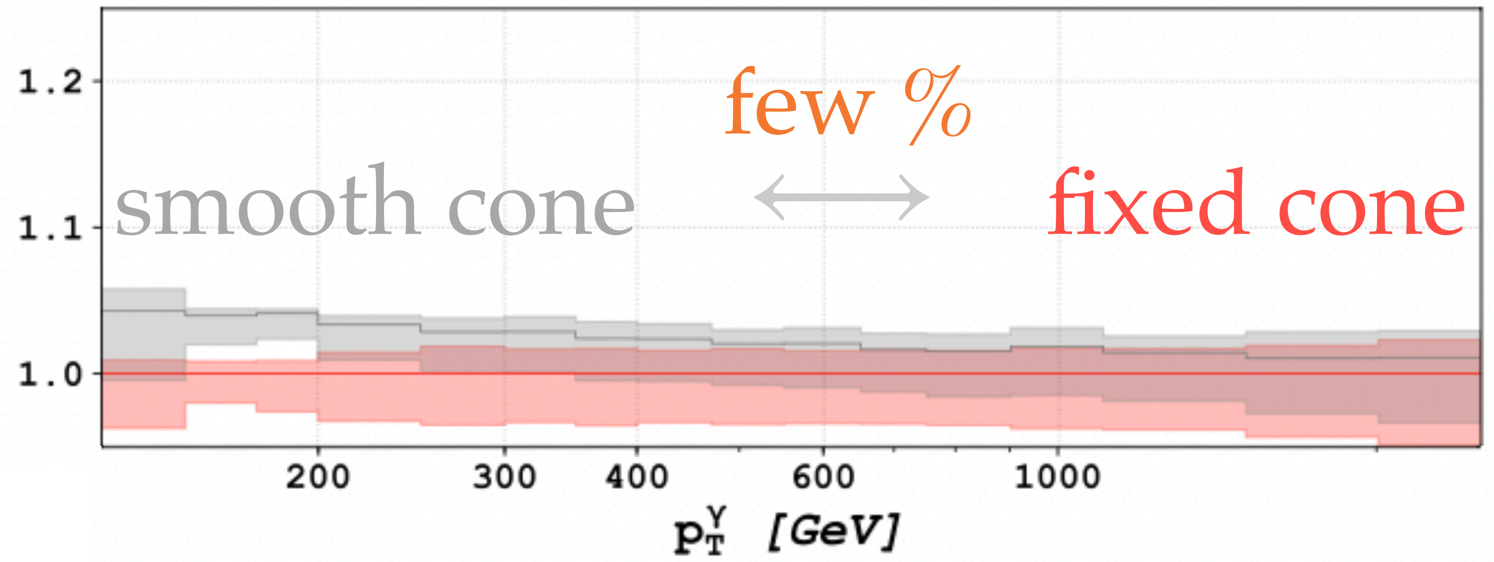
- B-hadron Stripper [Czakon, Generet, Mitov, Poncelet '21]
- antenna final-final [Gehrmann, Stagnitto '22]
- …

γ + jet @ NNLO with Fragmentation



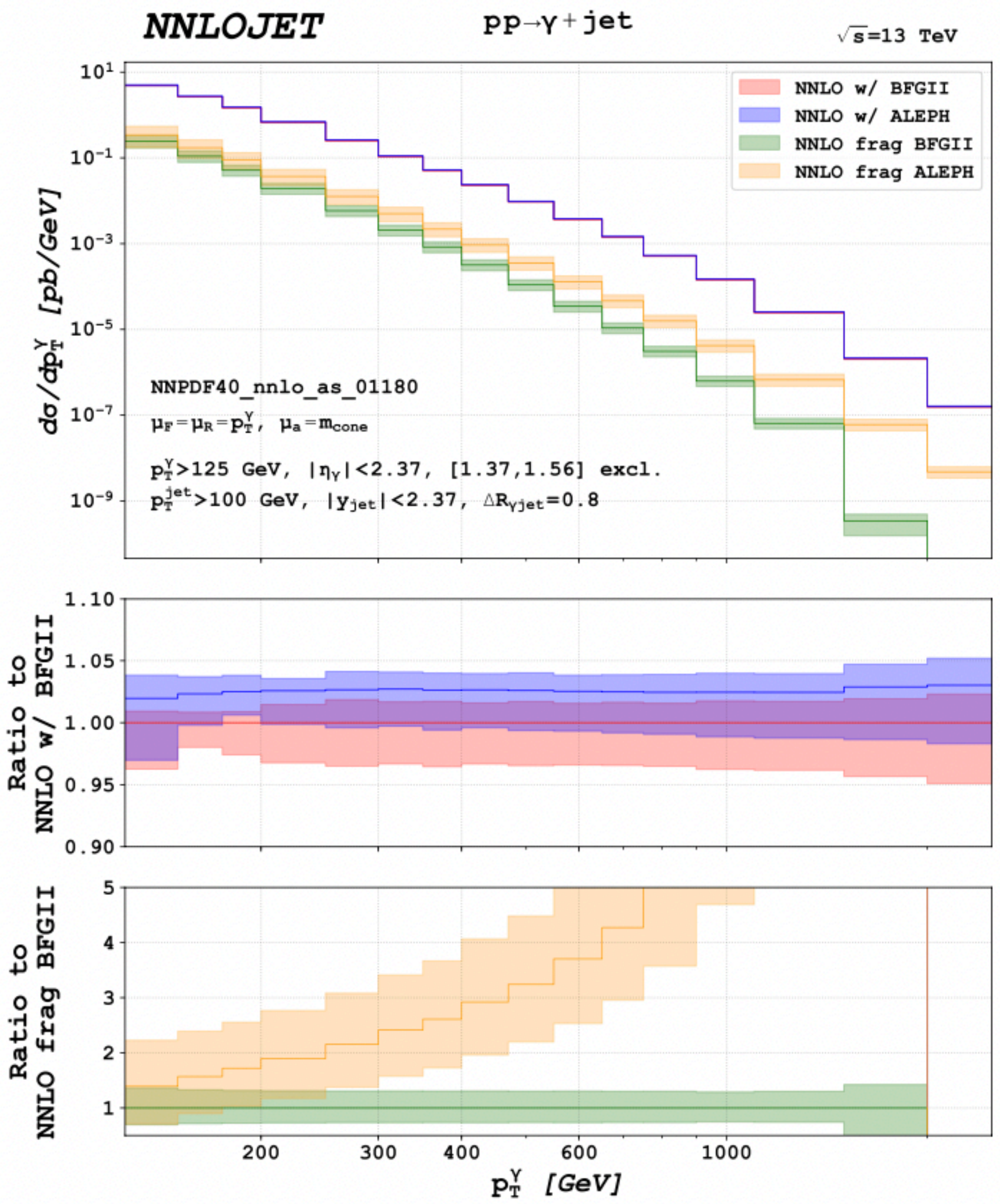
slight deterioration of agreement

maybe ok at NLO; **<u>matters</u>** at NNLO



Dependence on $D_{a\to\gamma}$

- BFG II vs. ALEPH
  [Bourhis, Fontannaz, Guillet '98]
  [ALEPH collab. '96]
- differences on $d\sigma/dp_T^\gamma$
  ⇝ 2 – 4%
- frag. contrib. $\times 10^{-1}$
  ⇝ $\mathcal{O}(1)$ differences
- access to $D_{a\to\gamma}$ @ LHC
  ↪ new observables?
  ↪ NNFrag?
  ↪ …



Conclusions & Outlook Part 1

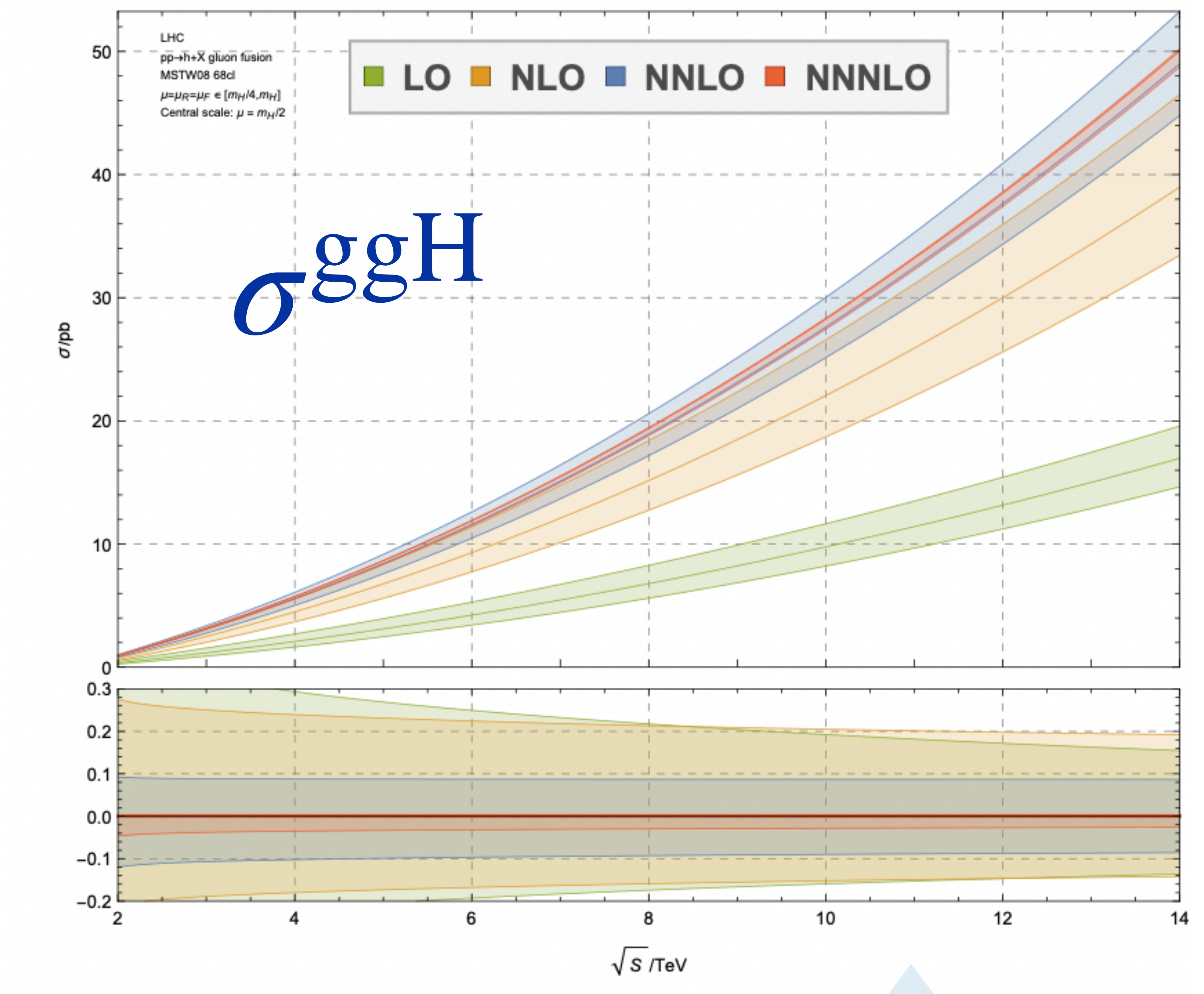
- NNLO QCD calculations in good shape
  - $2 \to 2$ essentially solved
  - $2 \to 3$ new frontier ⇝ methods reaching maturity
  - *loop amplitudes* becoming a bottleneck again
  - in the quest for percent-level theory ⇝ mixed QCD×EW important
- dissemination of results
  - public codes (MCFM, Matrix), nTuples, …
  - fast interpolation grids ⇝ `APPLgrid` `fastNLO` `PineAPPL` (anyway needed in fitting)
- identified objects ⇝ mismatch in TH vs. Exp/NNLO
  - photon isolation, flavour tagging, hadron fragmentation, …

The Plan.

1. *NNLO predictions for the LHC*
    - *jets & interpolations grids*
    - *identified photons & fragmentation*
2. *Differential N^3LO*
    - *Higgs & fiducial power corrections*
    - *Drell-Yan & PDFs*
3. *Bayesian approach to MHO*
    - *the abc model & correlations*
4. *Summary & Outlook*

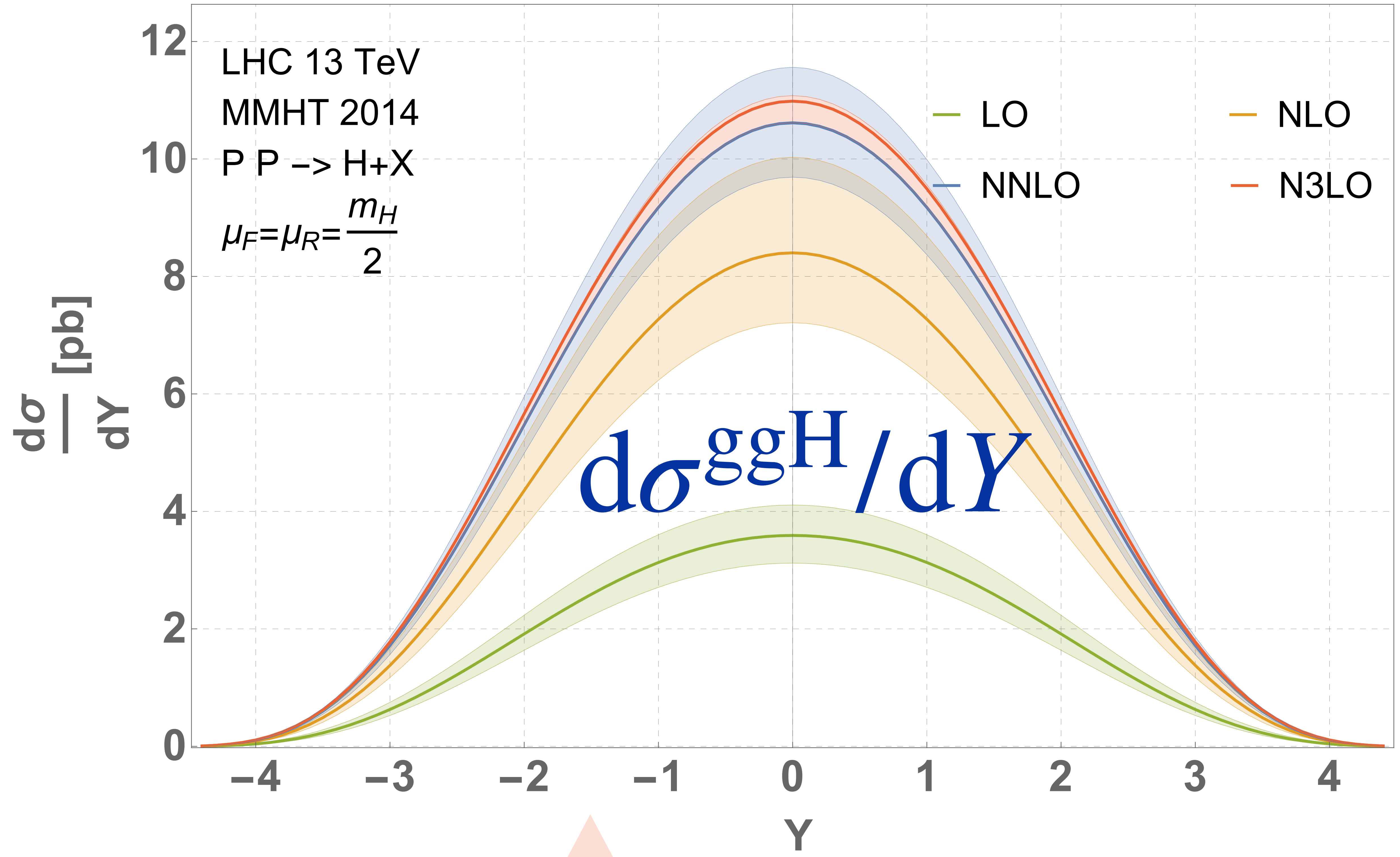
Higgs ggH @ N3LO — Inclusive* Predictions

[Anastasiou, Duhr, Dulat, Herzog, Mistlberger '15]



nice convergence of perturbative expansion

[Dulat, Mistlberger, Pelloni '18]



differential info lost:
↪ Higgs kinematics, QCD radiation, ...

* analytically integrated over emissions: ⊕ extremely fast; ⊖ idealised setup

Fully Differential ggH @ N3LO

[Chen, Gehrmann, Glover, AH, Mistlberger, Pelloni '21]

- Projection-to-Born:

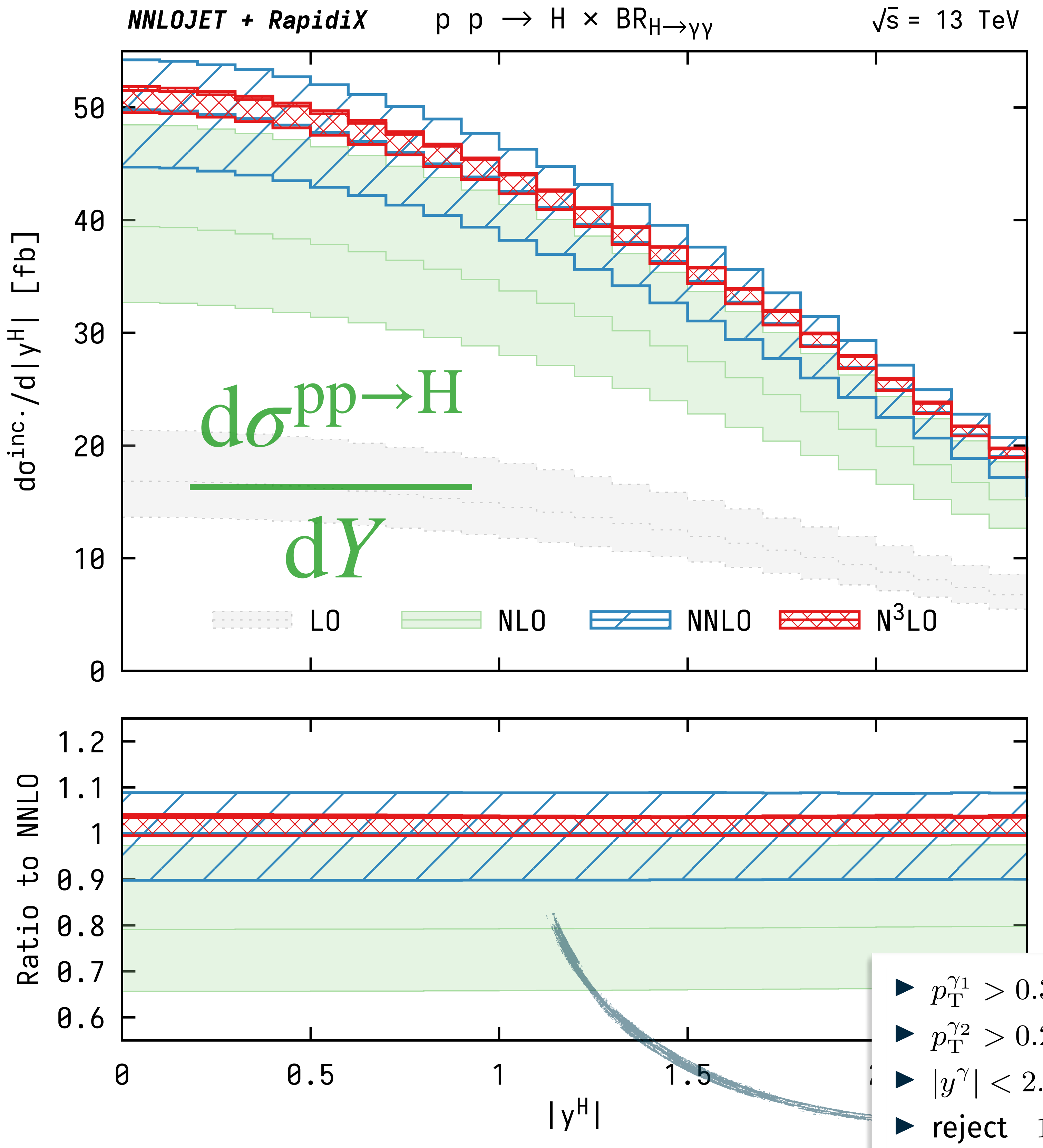
[VBF @ NNLO: Cacciari et al. '15]

$$\frac{\mathrm{d}\sigma_F^{\mathrm{N}^k\mathrm{LO}}}{\mathrm{d}\mathcal{O}} = \frac{\mathrm{d}\sigma_{F,\,\mathrm{inc.}}^{\mathrm{N}^k\mathrm{LO}}}{\mathrm{d}\mathcal{O}_B} + \left\{ \frac{\mathrm{d}\sigma_{F+\mathrm{jet}}^{\mathrm{N}^{k-1}\mathrm{LO}}}{\mathrm{d}\mathcal{O}} - \frac{\mathrm{d}\sigma_{F+\mathrm{jet}}^{\mathrm{N}^{k-1}\mathrm{LO}}}{\mathrm{d}\mathcal{O}}\bigg|_{\mathcal{O}\to\mathcal{O}_B} \right\}$$

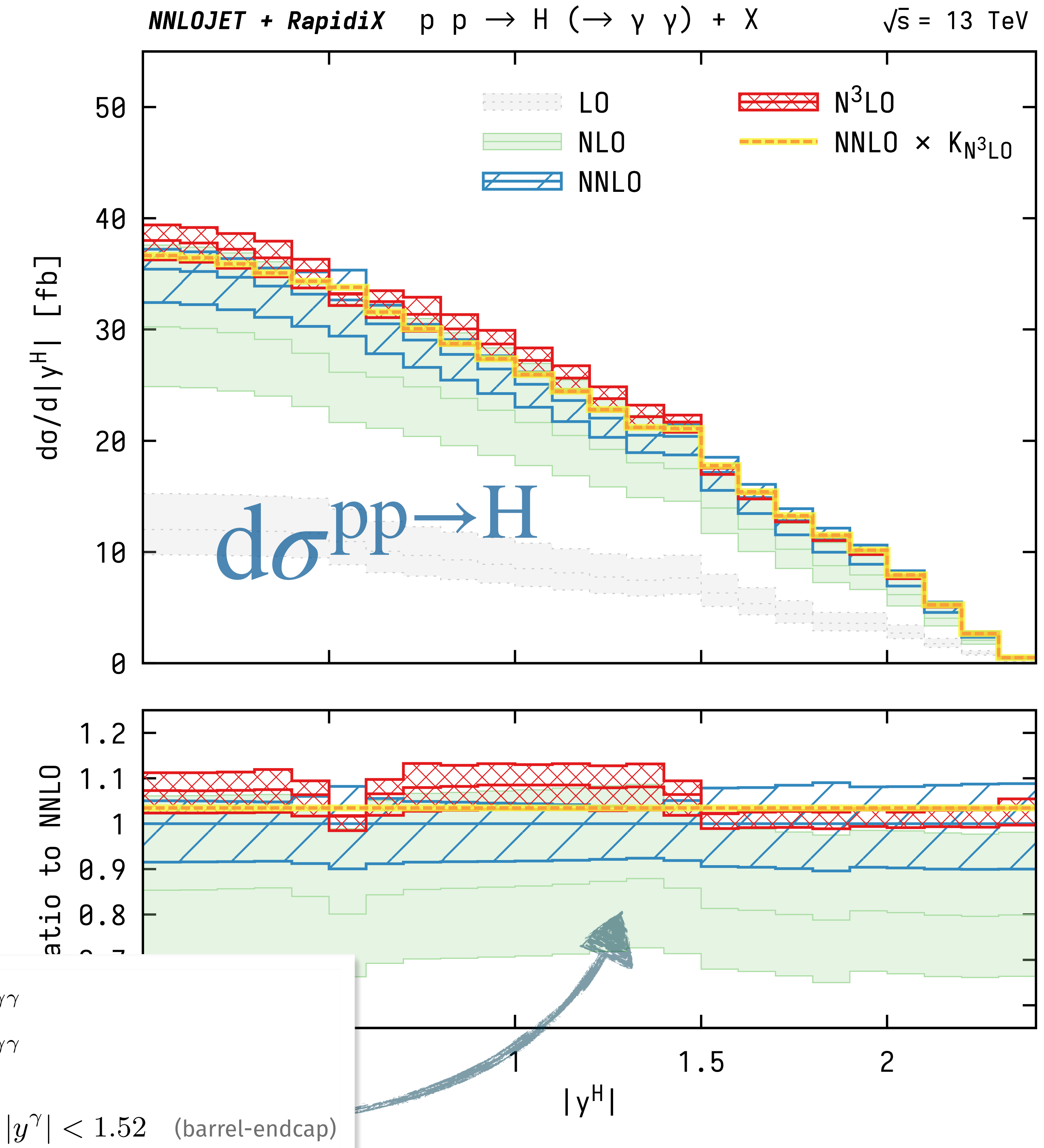
- idea:

restore *differential* info of an *inclusive* calculation

Inclusive

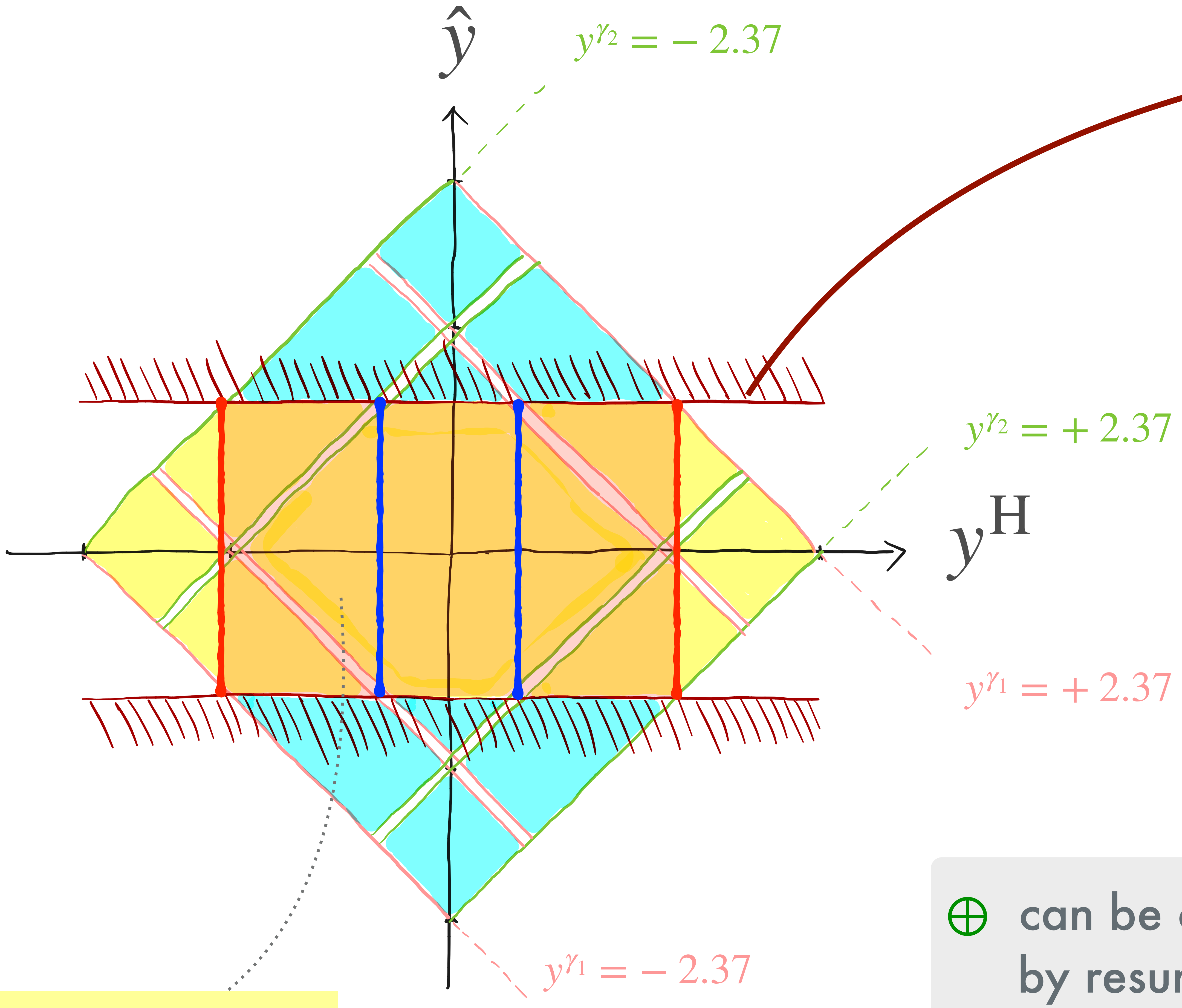


Fully Differential



- $p_T^{\gamma_1} > 0.35 \cdot m_{\gamma\gamma}$
- $p_T^{\gamma_2} > 0.25 \cdot m_{\gamma\gamma}$
- $|y^\gamma| < 2.37$
- reject $1.37 < |y^\gamma| < 1.52$ (barrel-endcap)
- *photon isolation* in $\Delta R < 0.2$
  $\hookrightarrow \sum_{\Delta R_{i\gamma}<0.2} p_{T,i} < 0.05 \cdot E_T^\gamma$

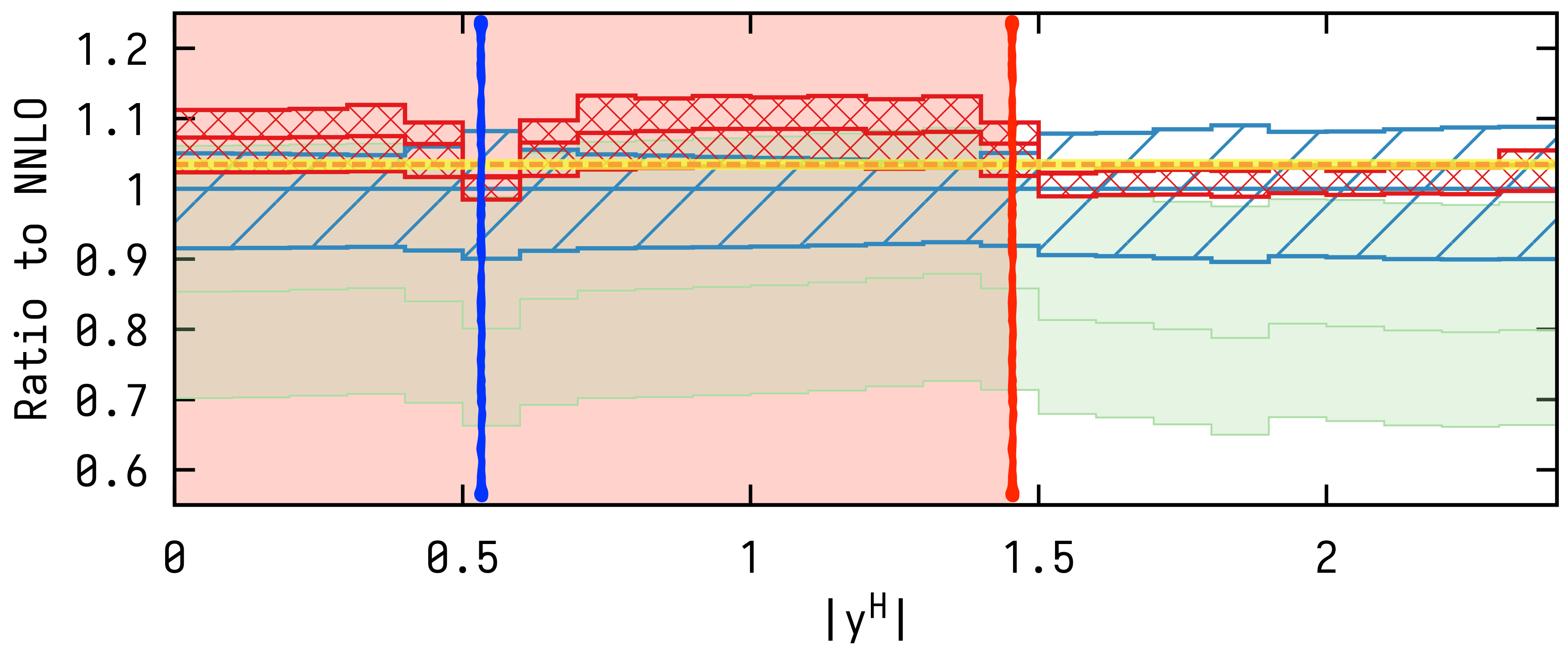
Fiducial Acceptances & y_H



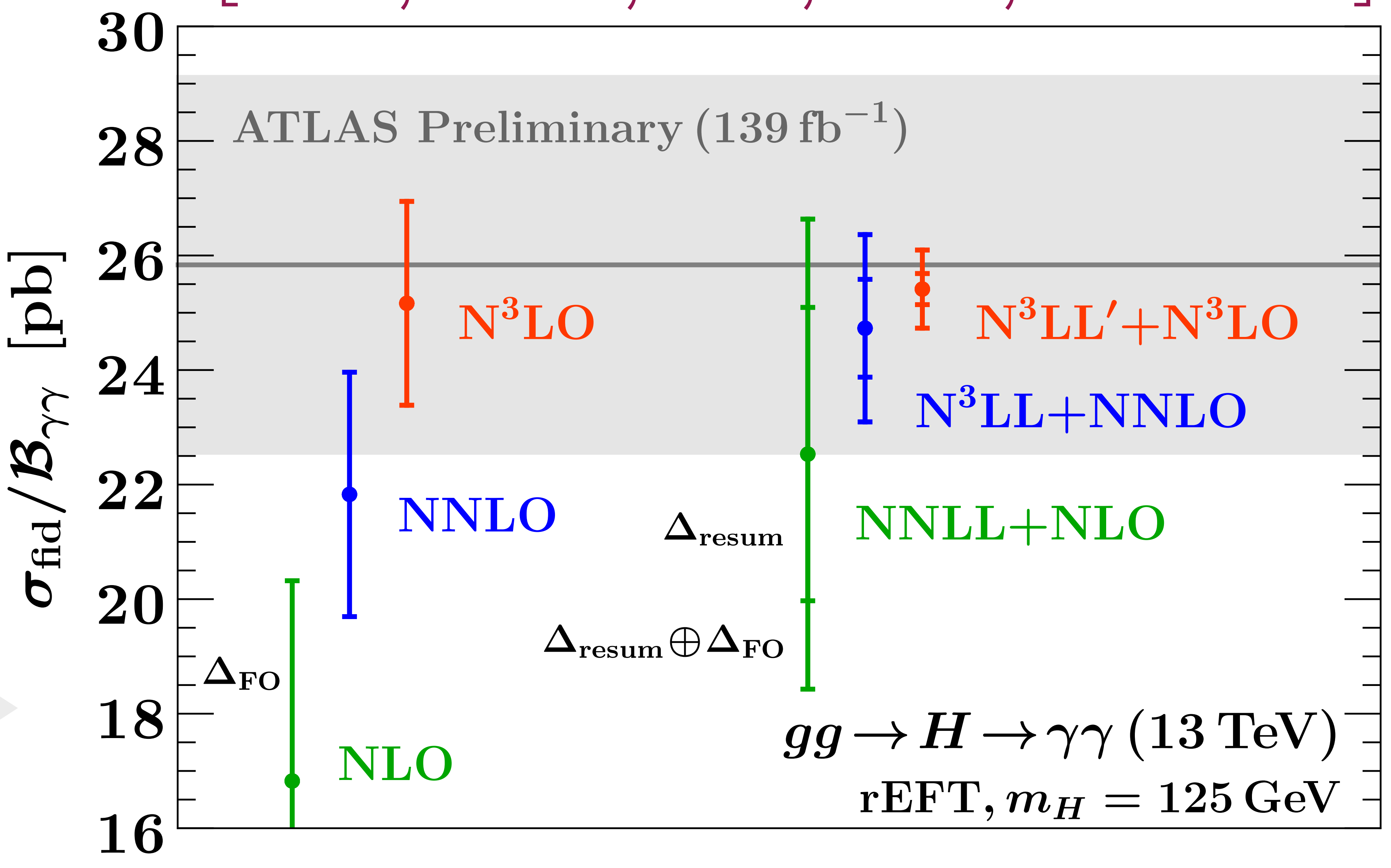
Born acceptance

$$\hat{y} \sim \frac{1}{2}|\Delta y(\gamma_1,\gamma_2)| \leq \cosh^{-1}\left(\frac{M_\mathrm{H}}{2p_\mathrm{T}^{\min}}\right) \approx 0.9$$

- linear fiducial power corrections!



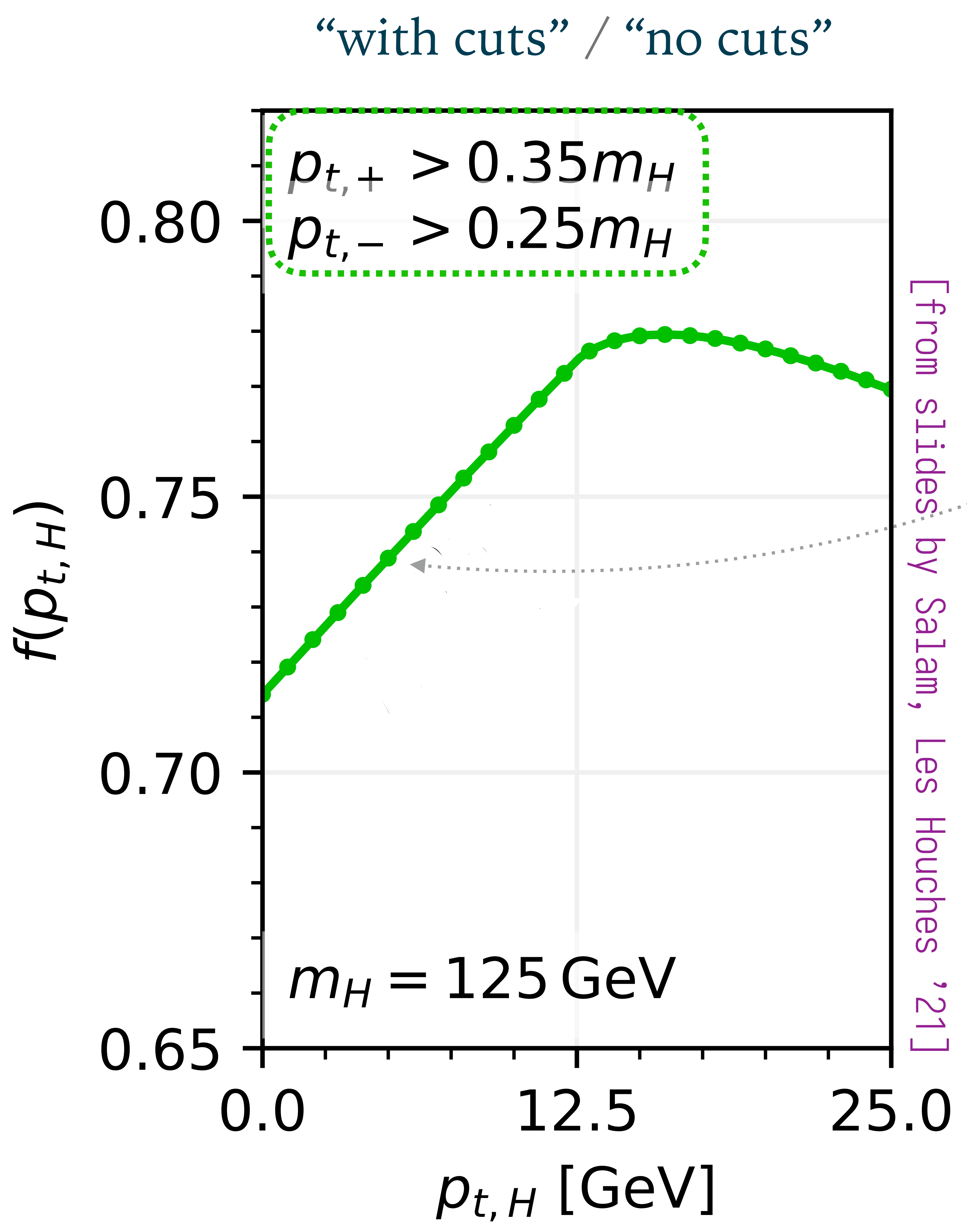
[Billis, Dehnadi, Ebert, Michel, Tackmann '21]



⊕ can be cured by resummation

⊖ hard $\sigma^{\mathrm{fid.}}$ should not need resummation

ACCEPTANCE $f(p_T^H)$



$$f(p_T^H) = f_0 + f_1 \cdot p_T^H + \mathcal{O}\left((p_T^H)^2\right)$$

[Frixione, Ridolfi ’97; Ebert, Tackmann ’19 + Michel, Stewart ‘21; Alekhin et al. ’21]

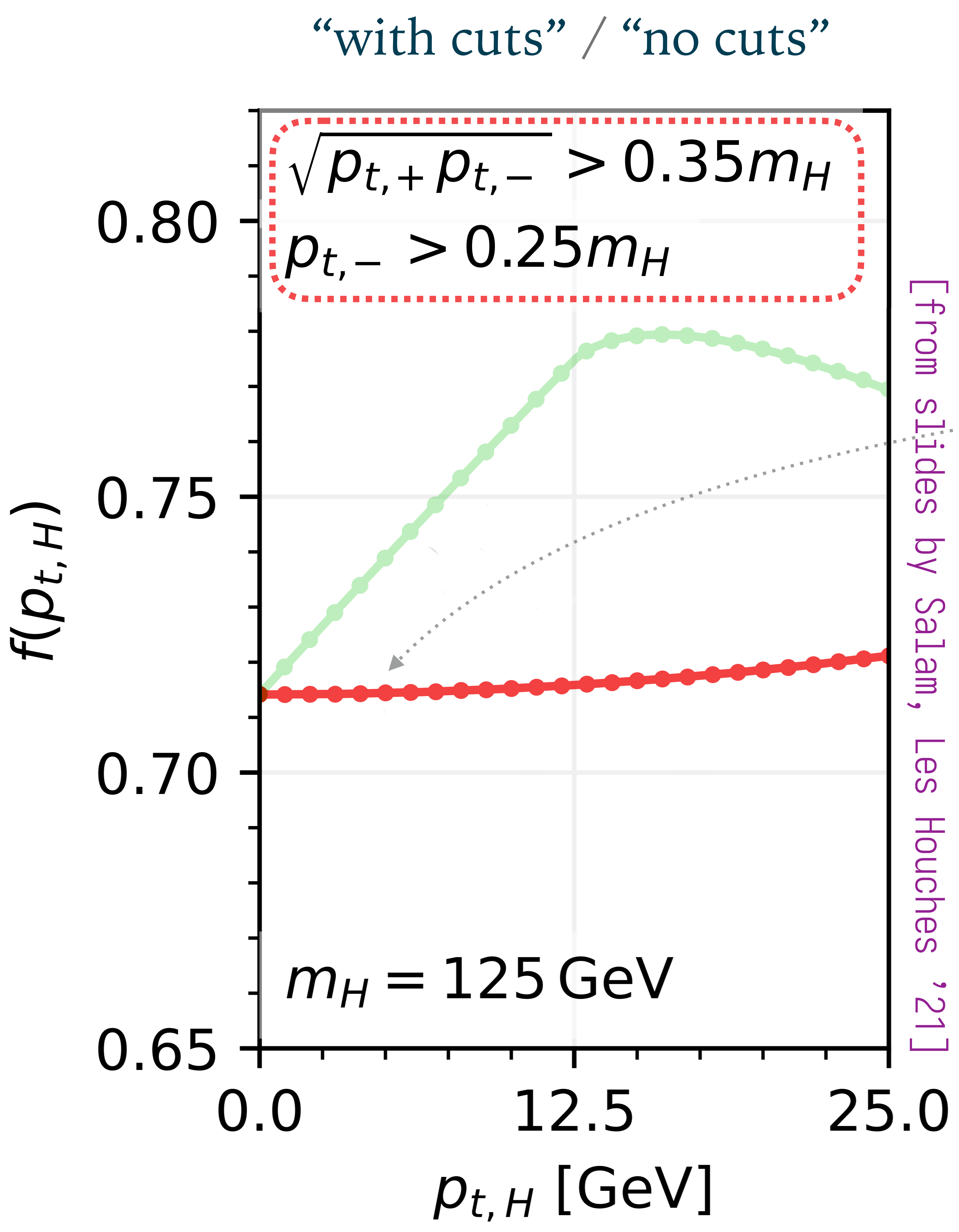
- Linear p_T^H dependence
  - **factorial growth** for fixed-order
  - ***sensitivity*** to very low p_T^H

$$\frac{\sigma_{\text{asym}} - f_0\, \sigma_{\text{inc.}}}{\sigma_0\, f_0} \simeq 0.18_{\alpha_s} - 0.15_{\alpha_s^2} + 0.31_{\alpha_s^3} + \ldots$$

$$\simeq 0.12 \text{ @ N}^3\text{LL}$$

[Salam, Slade '21]

Acceptance $f(p_T^H)$



$$f(p_T^H) = f_0 + f_1 \cdot p_T^H + f_2 \cdot (p_T^H)^2 + \mathcal{O}\left((p_T^H)^3\right)$$

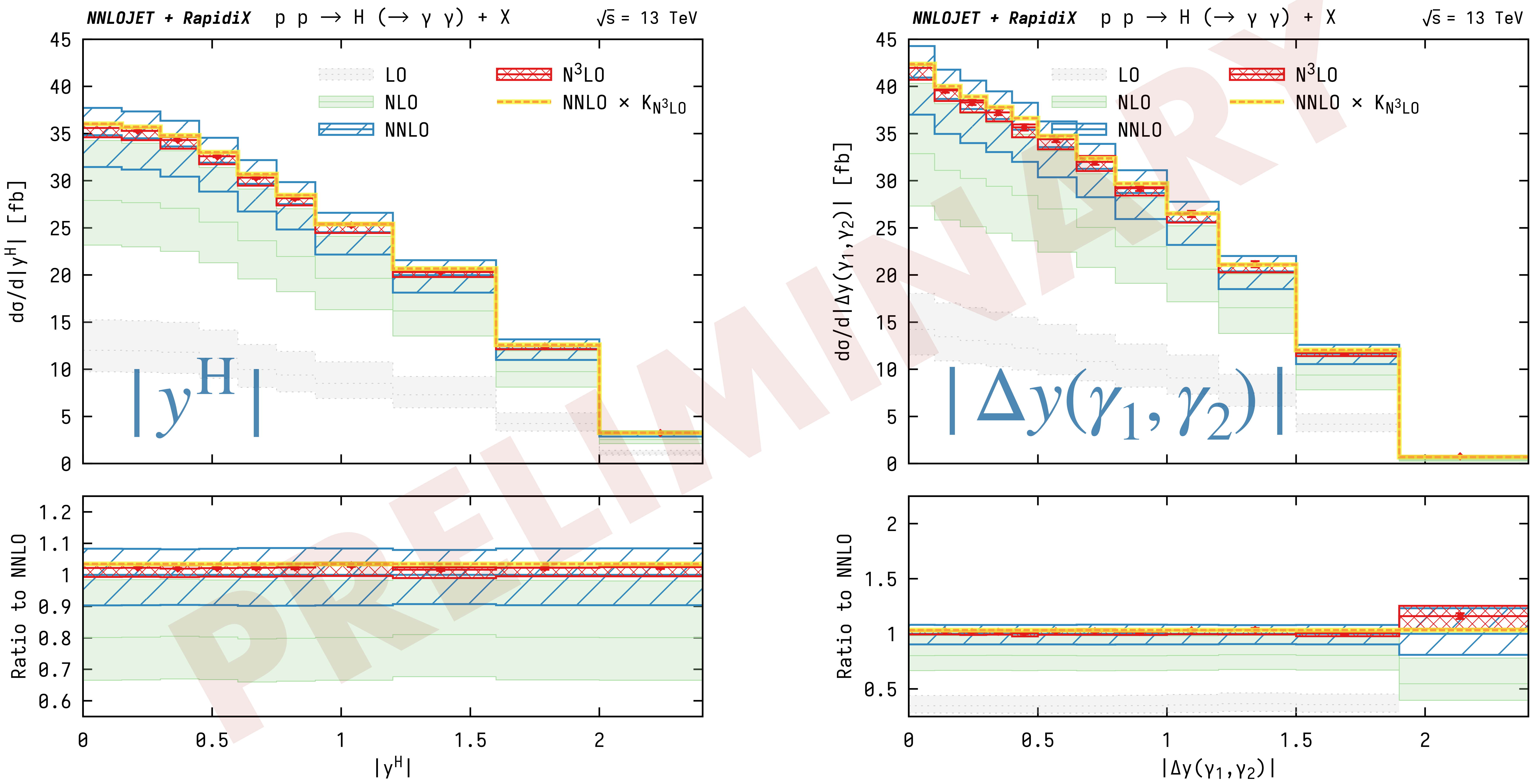
- Quadratic p_T^H dependence
  - *suppress* factorial growth
  - fixed order $\simeq$ resummation ✓

$$\frac{\sigma_{\text{prod}} - f_0\,\sigma_{\text{inc.}}}{\sigma_0\, f_0} \simeq 0.005_{\alpha_s} + 0.002_{\alpha_s^2} - 0.001_{\alpha_s^3} + \ldots$$

$$\simeq 0.006 \text{ @ N}^3\text{LL}$$

[Salam, Slade '21]

Higgs @ N3LO with Product Cuts



$$\sqrt{p_T^{\gamma_1}\, p_T^{\gamma_2}} \geq 0.35 \cdot M_H$$

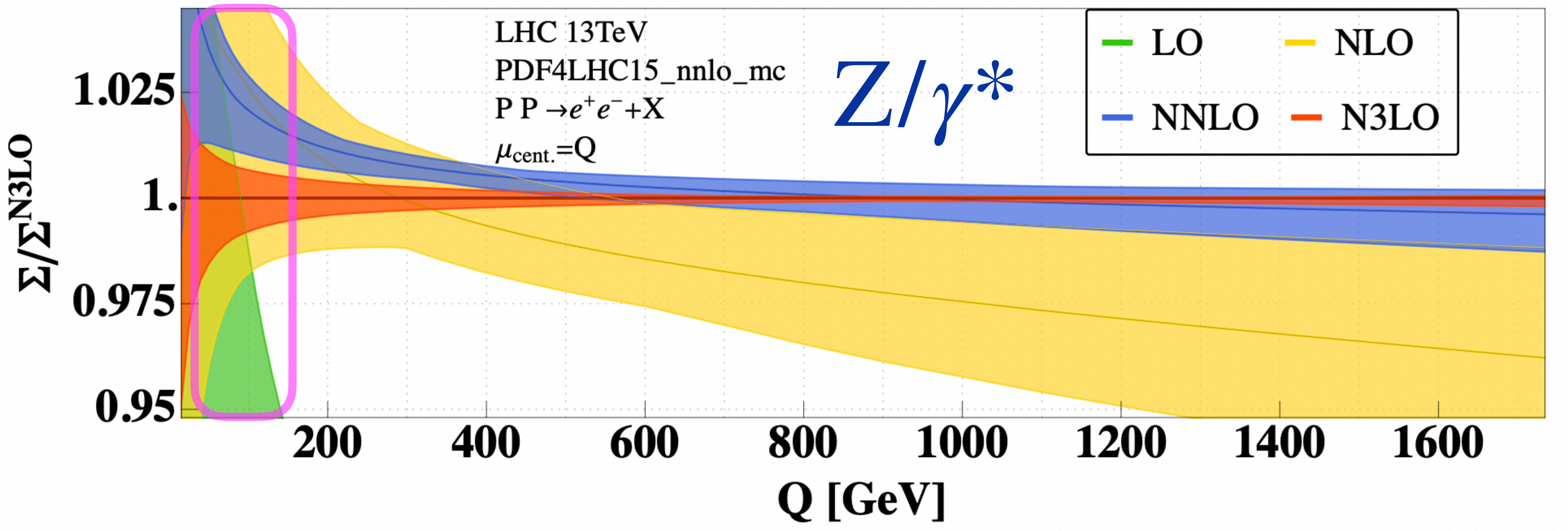
$$p_T^{\gamma_2} \geq 0.25 \cdot M_H$$

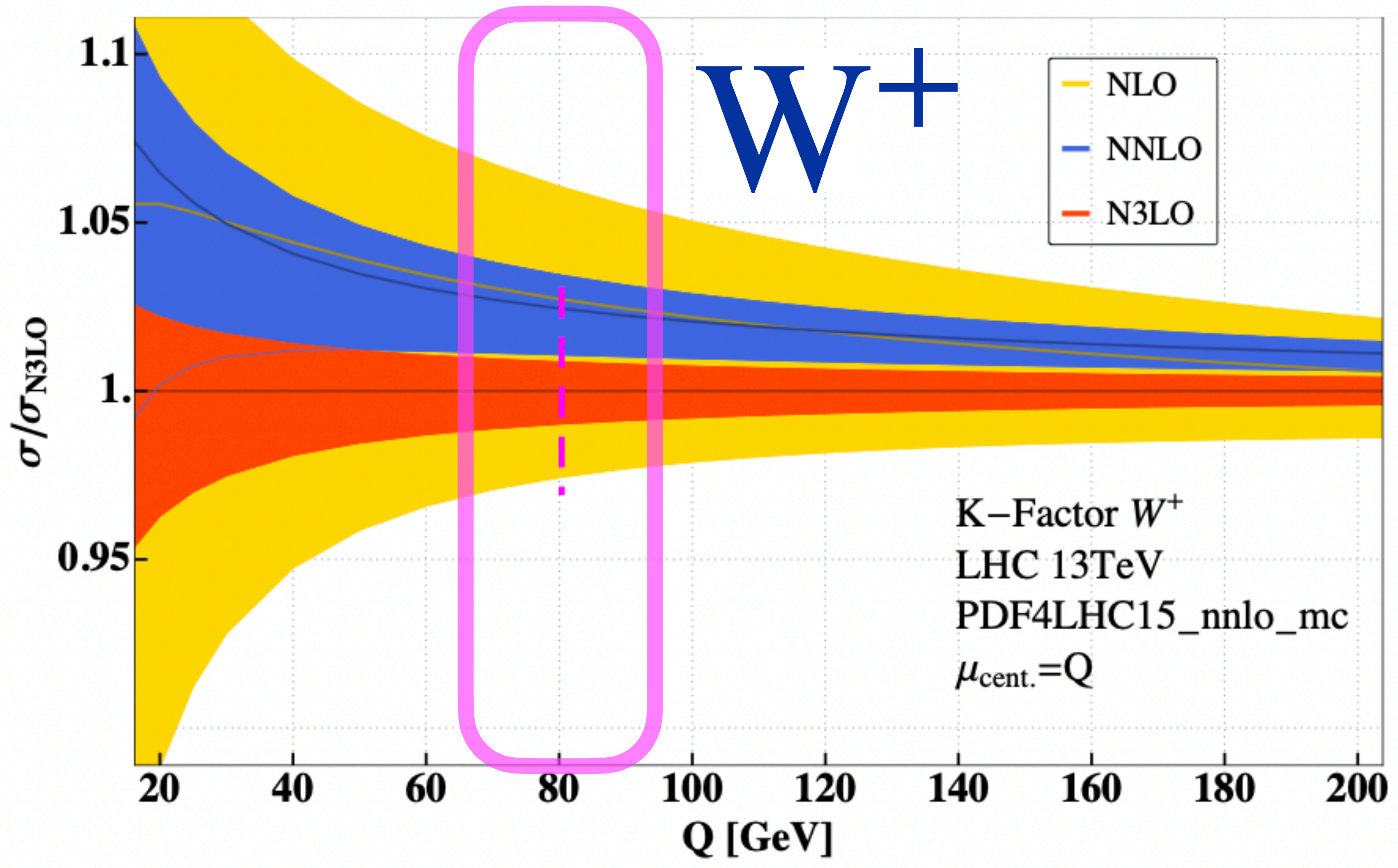
- no visible instabilities ⇝ flat K-factor
- $N^3LO \simeq NNLO \times K_{N^3LO}$

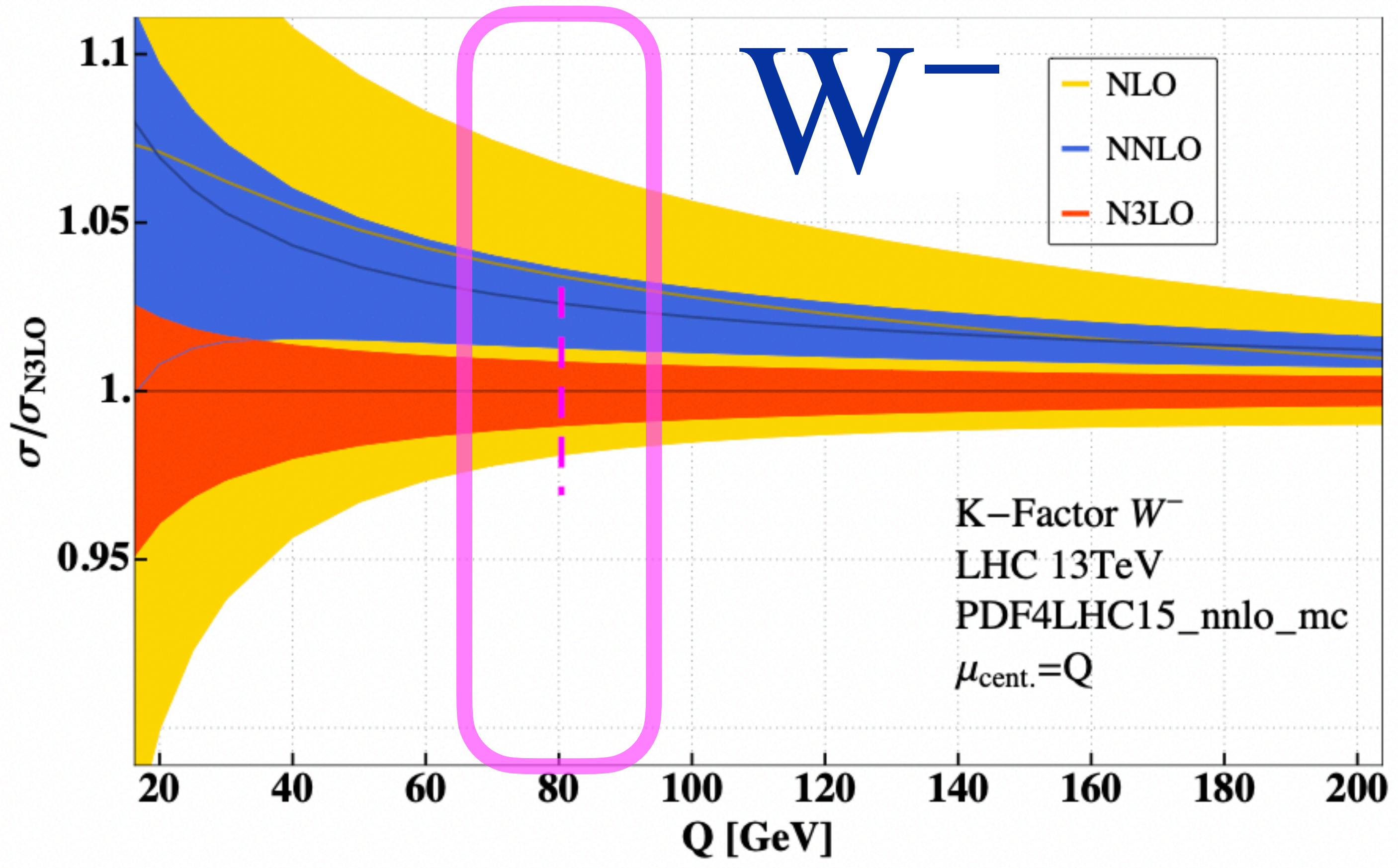
- no visible instabilities; no “features” in the corrections; very flat K-factors
- $N^3LO \sim NNLO \times K_{N^3LO}$

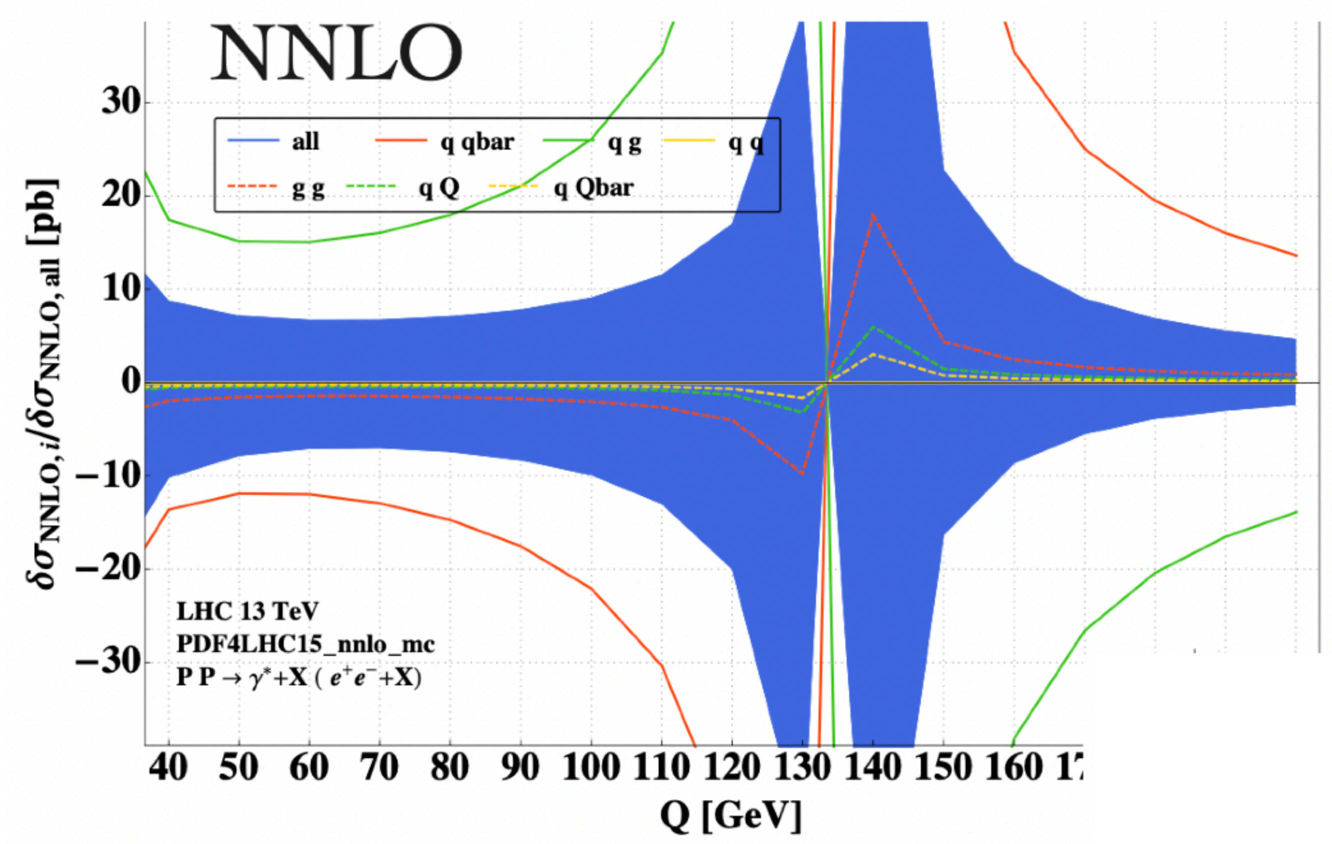
Drell–Yan @ N3LO – Q Dependence

[Dulat, Duhr, Mistlberger '20 '21]









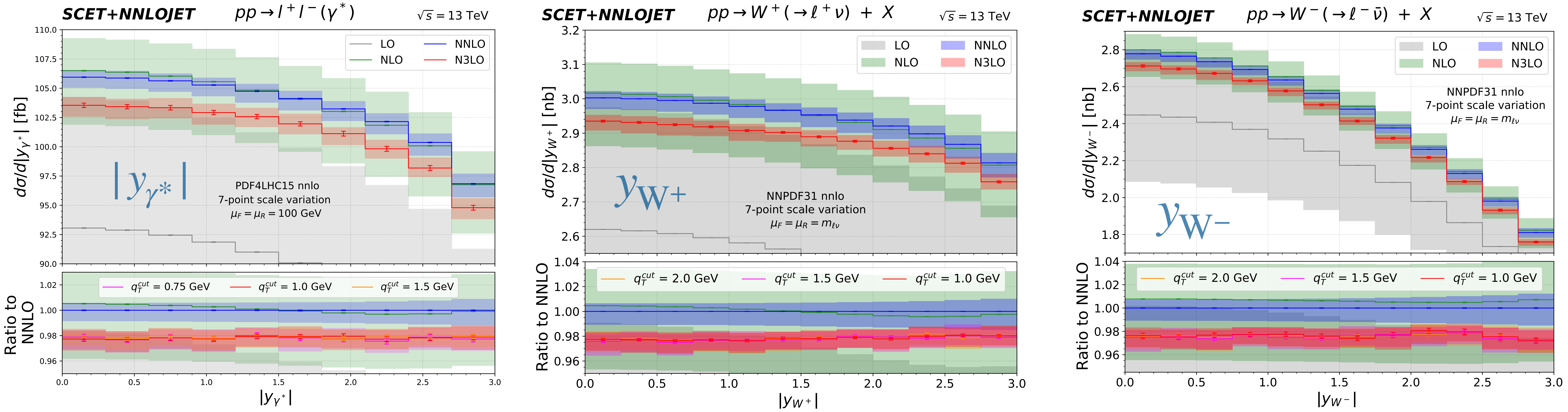
NNLO: $1 \sim \pm 20$
(large cancellations)
↪ artificially small?

N3LO: $1 \sim \pm 2$

resonance region ⇝ non-overlapping bands; $\Delta_{\rm scl}^{\rm NNLO} \simeq \Delta_{\rm scl}^{\rm N^3LO}$?!

Drell–Yan @ N3LO – Y_V Distributions

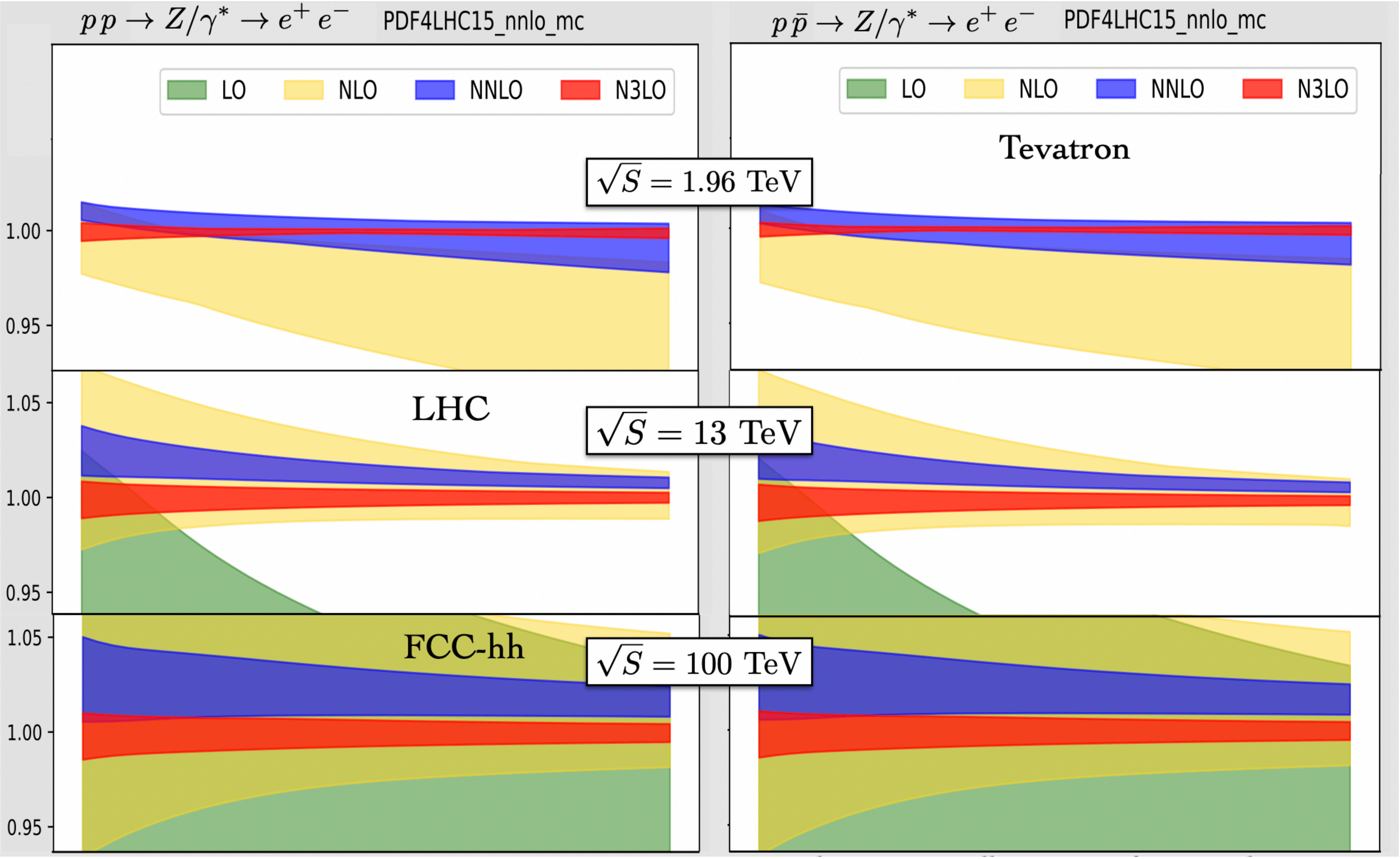
[Chen, Gehrmann, Glover, AH, Yang, Zhu '21,'22]



- same collider @ 13 TeV ⇝ almost universal NNLO → N³LO corrections!
- NC & CC$^{\pm}$ processes probe different parton content across Y_V (valence u vs. d, …)

DRELL–YAN @ N3LO — COLLIDER DEPENDENCE

[slides from C. Duhr: TH colloquium ’22]



no “odd” scale behaviour @ 2 TeV

main **difference** from: collider energy

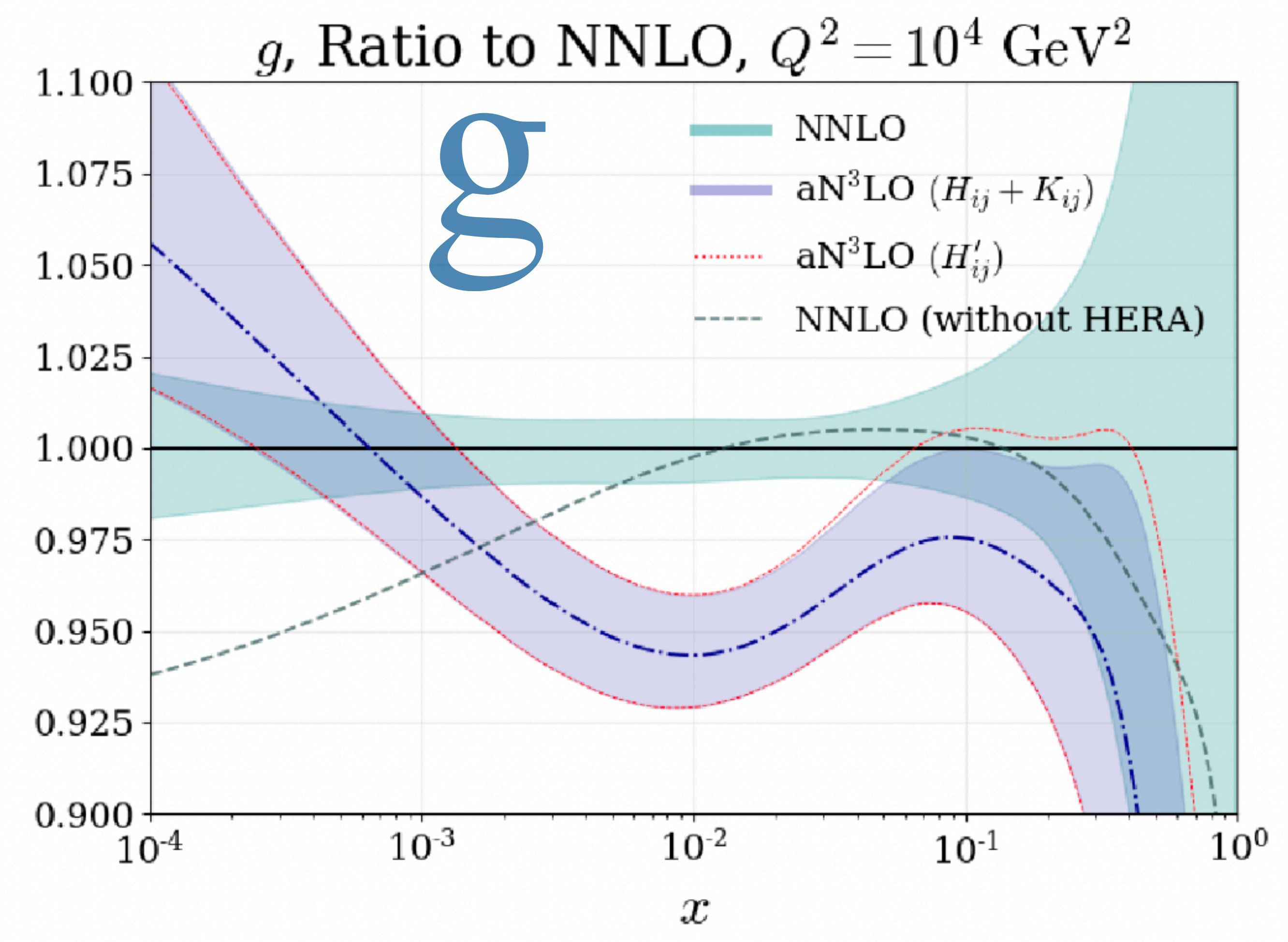
very **similar** between: pp vs. $p\bar{p}$

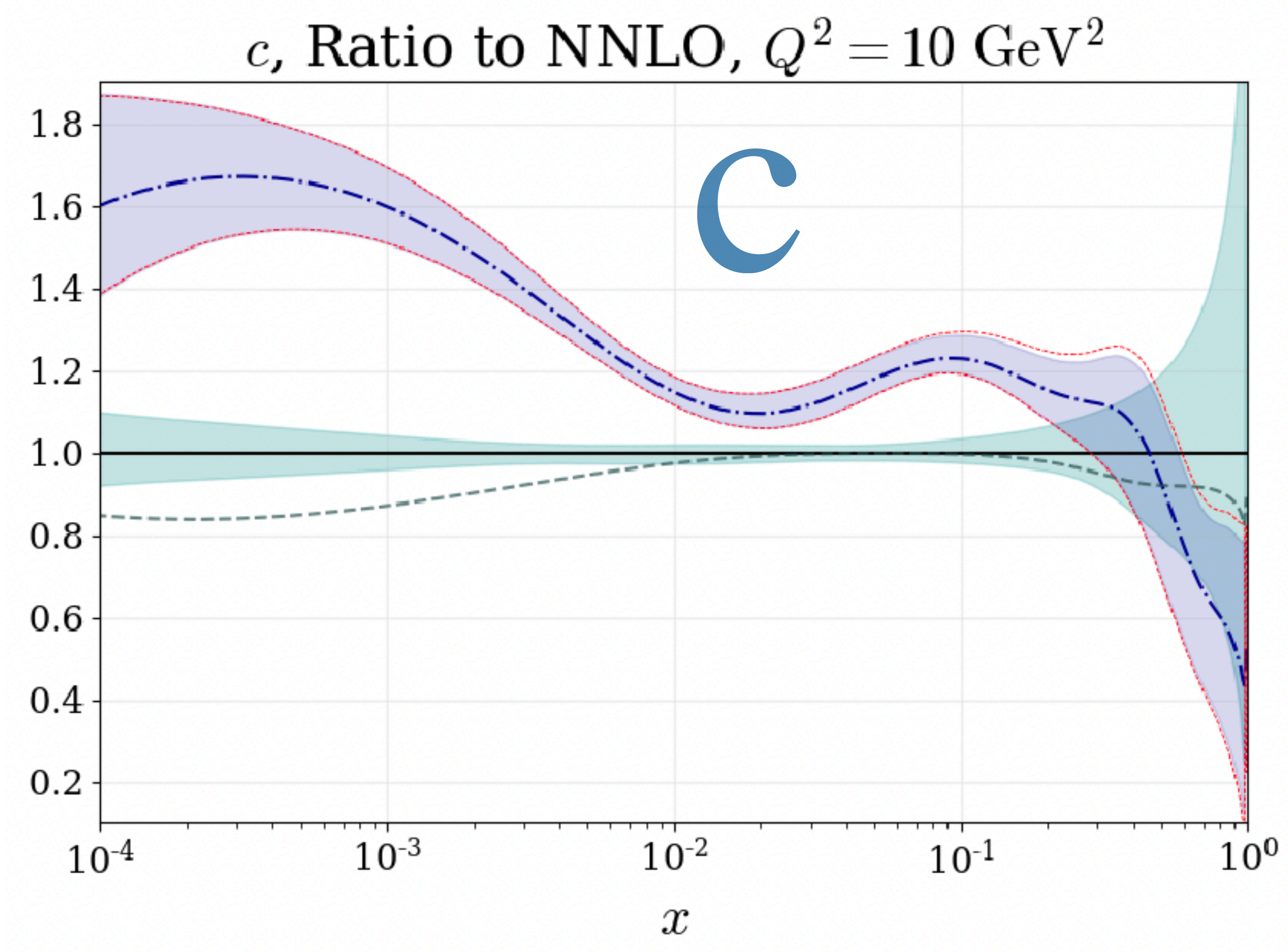
N3LO Parton Distribution Functions

- N3LO evolution
  - ⇝ 4-loop splitting functions
    [Moch, Ruijl, Ueda, Vermaseren, Vogt '17,'18,'22, in progress]

- aN3LO PDFs (MSHT)

[McGowan, Cridge, Harland-Lang, Thorne '22]

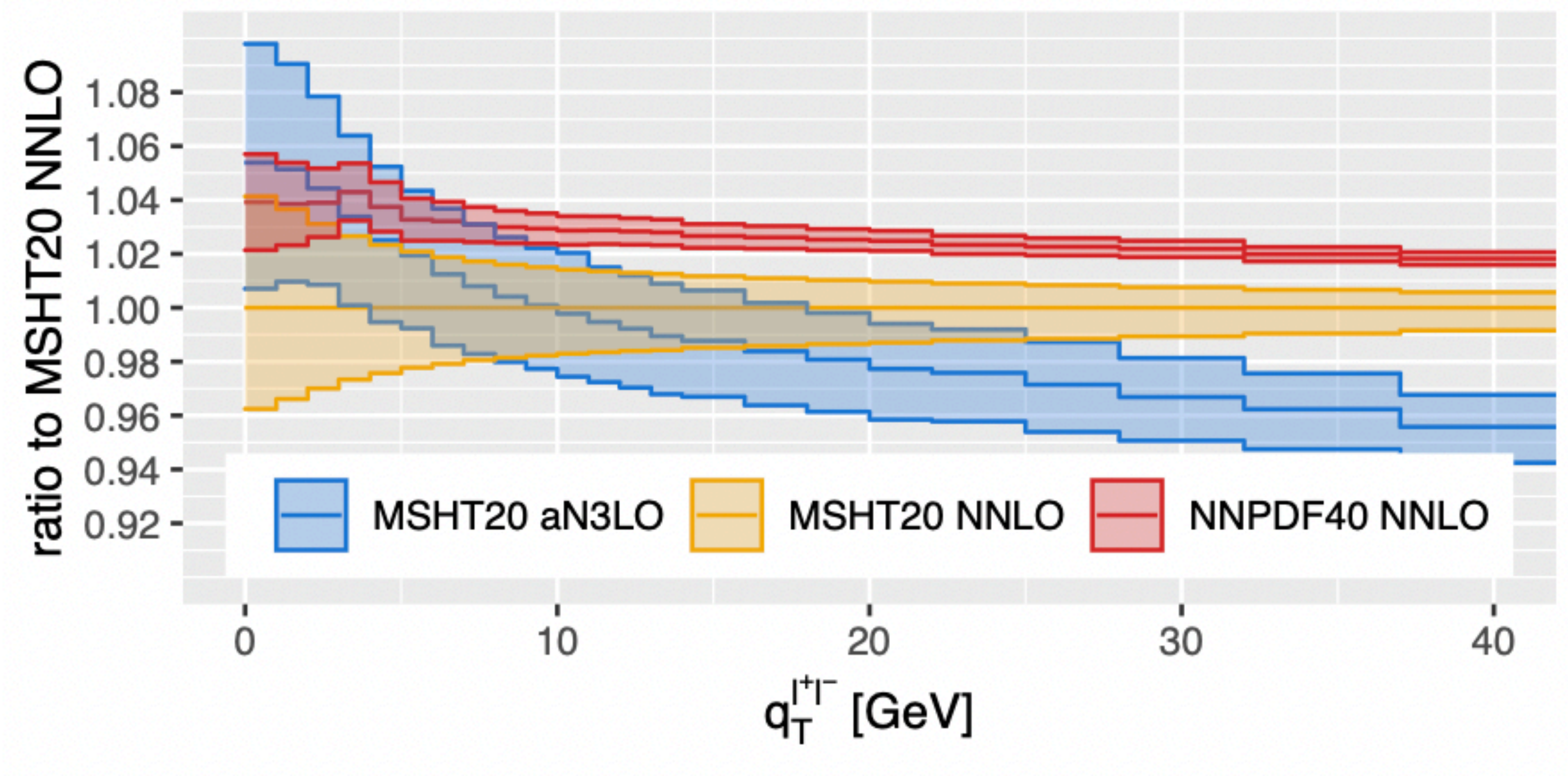




ggH: $\delta\sigma^{\mathrm{N^3LO}}$ ↘ VBF: $\delta\sigma^{\mathrm{N^3LO}}$ ↗

- purely resummed p_T^Z spectrum
  - ⇝ PDF uncertainties

[Neumann, Campbell '22]



syst. differences between PDFs

PDF (NNLO → N³LO) $\delta\sigma^{\mathrm{N^3LO}}$ ↗ (?)

Conclusions & Outlook Part 2

- N^3LO predictions are key to reach percent-level accuracy
  - computation of *inclusive* $2 \to 1$ processes very mature ⇜⇝ ggH, DY, VBF, VH, …
  - differential predictions for $pp \to$ "colour neutral" appearing
    ⇜⇝ relies on very stable NNLO "+jet" calculation
  - *but:* performance of slicing methods very poor ⇜⇝ $\mathcal{O}$(10M) CPU core hours
- Fiducial cuts ⇜⇝ linear power corrections (other processes?)
  ↪ crucial for practicability of slicing approaches
- Inadequacies in traditional scale variations ⇜⇝ DY @ N3LO
  ↪ effect from missing N^3LO PDFs?
  ↪ more robust TH uncertainties desirable
  (Padé approximant, Bayesian models, PMC, series transforms, …)

THE PLAN.

1. *NNLO predictions for the LHC*
    - *jets & interpolations grids*
    - *identified photons & fragmentation*
2. *Differential N^3LO*
    - *Higgs & fiducial power corrections*
    - *Drell-Yan & PDFs*
3. *Bayesian approach to MHO*
    - *the abc model & correlations*
4. *Summary & Outlook*

WHAT IS THE UNCERTAINTY Δ_{TH} OF MY RESULT?

- increasingly urgent to address with $\Delta_{\mathrm{EXP}} \searrow$ ($\leftrightsquigarrow$ HL-LHC)
  - what does 5σ mean if Δ_{TH} non-negligible?
  - interpretation of data in need for robust Δ_{TH}: PDF fits, χ^2 in ATLAS jets, …
- various sources that contribute to Δ_{TH}:
  - Δ_{α_s}, Δ_{param}: parametric uncertainties $\leftrightsquigarrow$ exp. extraction
  - Δ_{PDF}: parton distribution functions (PDFs) $\leftrightsquigarrow$ fits
  - $\Delta_{\mathrm{non\ pert.}}$: hadronisation, UE, … $\leftrightsquigarrow$ parton showers [e.g. HERWIG vs. PYTHIA]
  - Δ_{MHO}: *missing higher-order (MHO)* corrections

Focus here

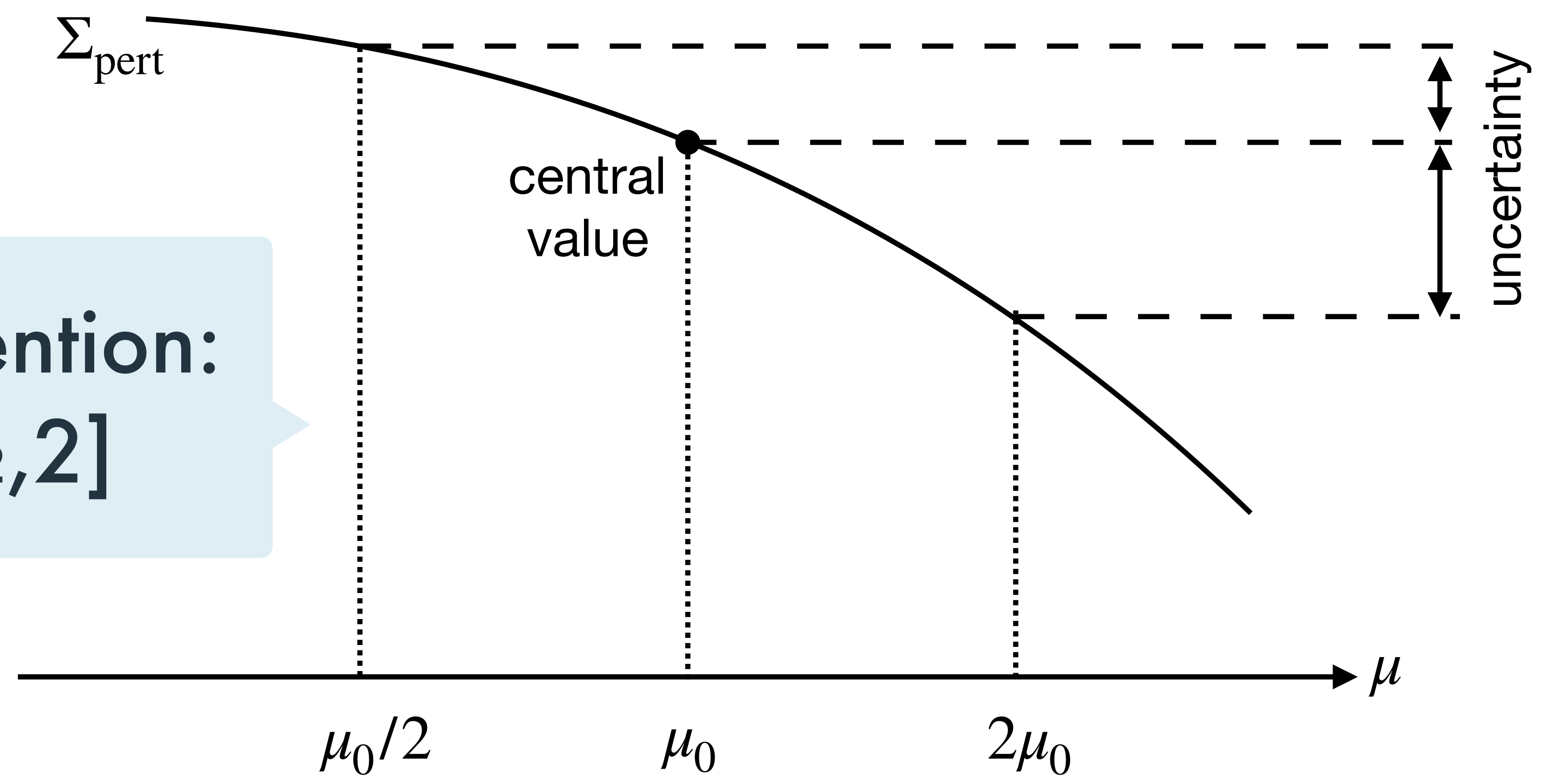
Conventional Approach for Δ_{MHO} — Scale Variation

- approximation for an observable @ (next-to-)^n leading order:
  - N^{n}LO: $$\Sigma \simeq \Sigma_n(\mu) = \sum_{k=0}^{n} \Sigma^{(k)}(\mu)$$ $\propto \alpha_s^{n_0+k}$

electroweak (EW):
↪ scheme dependence
↪ $\alpha \ll \alpha_s$

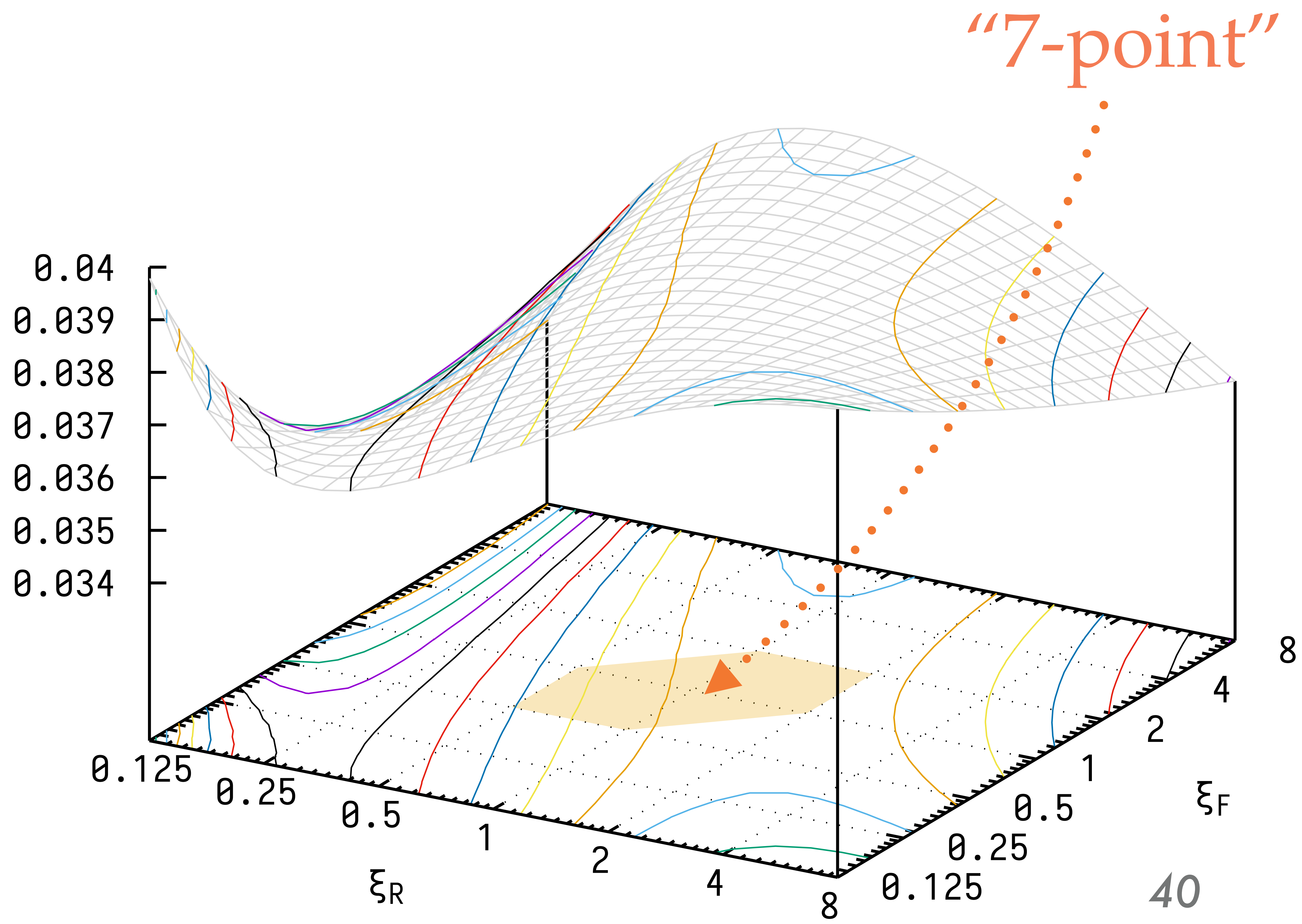
- truncation of series induces a sensitivity to terms of the next order

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu}\Sigma_n(\mu) = \mathcal{O}(\alpha_s^{n_0+n+1}) = \mathcal{O}(\Delta_{\text{MHO}})$$



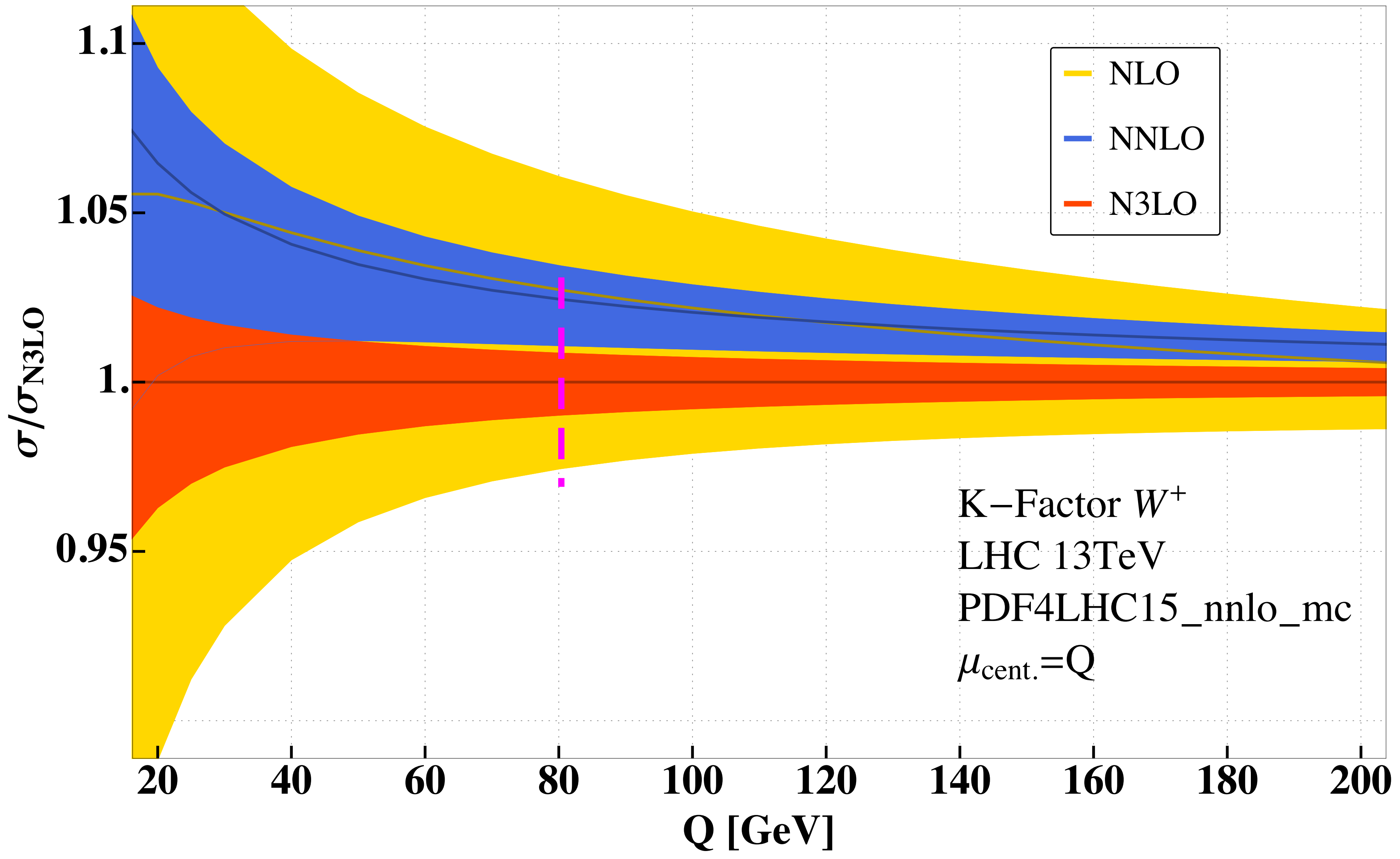
convention:
[½,2]

(μ_R, μ_F)



Issues with Standard Scale Variations

- known to be insufficient:
  - exclusive jet(s) (veto)
  - ratios (correlation?)
  - cancellations (e.g. $q\bar{q}$ vs. qg in DY)



[Duhr, Dulat, Mistlberger '20]

- choice of the central scale
  - fastest apparent convergence (FAC)
    ↪ $\Sigma^{(n)}(\mu_{\text{FAC}}) = 0$
  - principle of minimal sensitivity (PMS)
    ↪ $\frac{\partial}{\partial\mu}\Sigma^{(n)}(\mu)\Big|_{\mu_{\text{PMS}}} = 0$
  - BLM / PMC
    [Brodsky, Lepage, Mackenzie '83]; [Brodsky, Di Giustino '12]
  - …
- crucially: *no statistical interpretation!*
  ⇝ need to do better

Probability Distributions for Δ_{MHO}

[Cacciari, Houdeau '11]

- Sequence of perturbative *corrections* δ_k normalised w.r.t. LO (dimensionless)

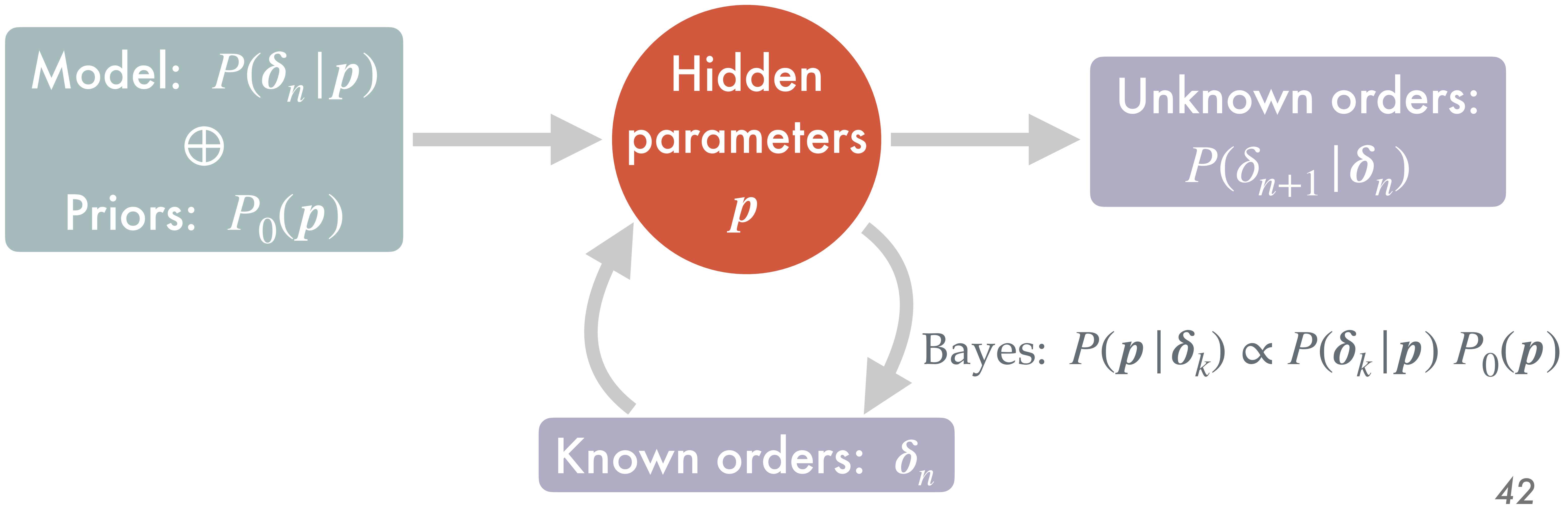
$$\Sigma_n = \Sigma^{(0)} \left(1 + \delta_1 + \ldots + \delta_n\right) \quad \rightsquigarrow \quad \delta_k = \mathcal{O}(\alpha_s^k)$$

- Probability distribution for δ_{n+1}, given $\boldsymbol{\delta}_n = (\delta_0, \delta_1, \ldots, \delta_n)$

$$P(\delta_{n+1} \,|\, \boldsymbol{\delta}_n) = \frac{P(\boldsymbol{\delta}_{n+1})}{P(\boldsymbol{\delta}_n)} = \frac{\int \mathrm{d}^m\boldsymbol{p}\; P(\boldsymbol{\delta}_{n+1} \,|\, \boldsymbol{p})\; P_0(\boldsymbol{p})}{\int \mathrm{d}^m\boldsymbol{p}\; P(\boldsymbol{\delta}_n \,|\, \boldsymbol{p})\; P_0(\boldsymbol{p})}$$

$$P(A, B) = P(A \,|\, B)\; P(B)$$

$$P(A) = \int \mathrm{d}B\; P(A, B)$$



The CH Model

[Cacciari, Houdeau '11]

- perturbative expansion $\delta_k = c_k\,\alpha_s^k$ bounded by a geometric series: $|c_k| \leq \bar{c} \quad \forall k$

$$\Big|\sum_k \delta_k\Big| \leq \sum_k |c_k|\,\alpha_s^k \leq \sum_k \bar{c}\,\alpha_s^k$$

  - one hidden parameter: $\bar{c}$
  - constrain upper bound $\bar{c}$ from known orders
    - ⇝ constraint on unknown coefficients c_{n+1}
- limitations:
  - α_s at what scale? why not: $\dfrac{\alpha_s}{\pi}$, $\dfrac{\alpha_s}{2\pi}$, $\alpha_s \ln^2(v)$, $\alpha_s \ln(v)$, … ?
- why not let the model figure out the expansion parameter itself?

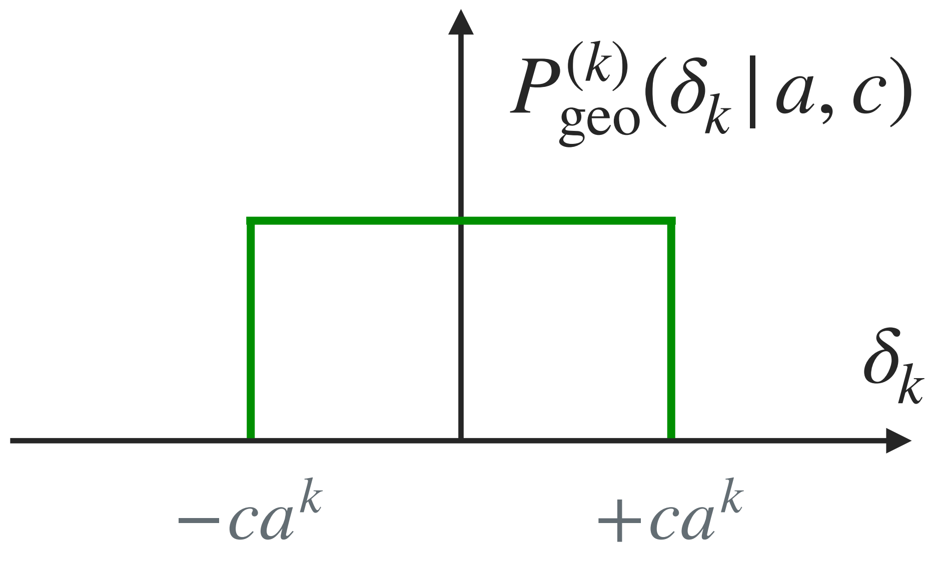
The Geometric Model

[Bonvini '20]

- bounded by a geometric series with expansion parameter a:

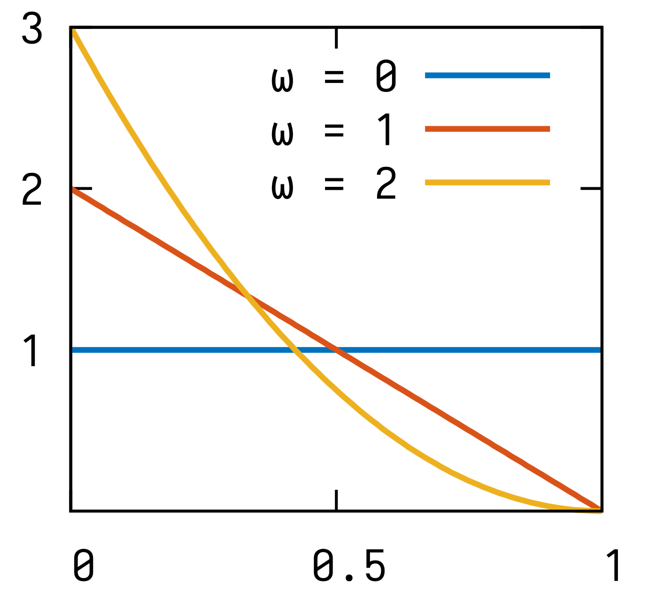
$$|\delta_k| \leq c\, a^k \quad \forall k$$ ⇜⇝ two model parameters: $a, \ c$

- **model:** $$P^{(k)}_{\text{geo}}(\delta_k | a, c) = \frac{1}{2c\, a^k}\, \Theta\!\left(c - \frac{|\delta_k|}{a^k}\right)$$



- **priors:** $$P_0(a, c) = P_0(a)\, P_0(c)$$

$$P_0(a) = (1+\omega)\, (1-a)^\omega\, \Theta(a)\, \Theta(1-a)$$



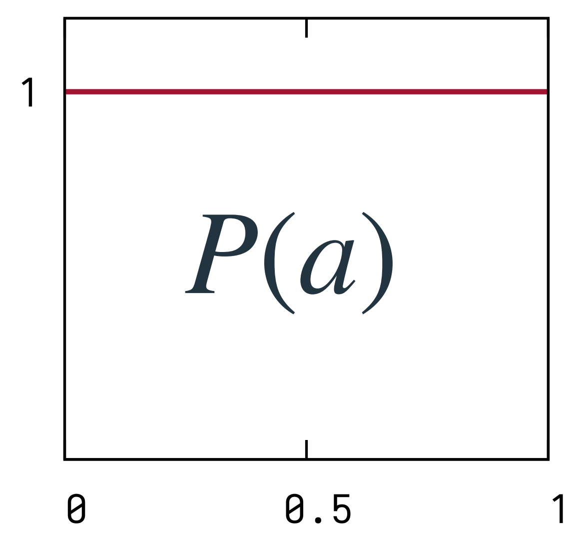
$$P_0(c) = \frac{\varepsilon}{c^{1+\varepsilon}}\, \Theta(c-1)$$ ⇜⇝ $\mathrm{d}c/c \sim \mathrm{d}\ln(c)$ (ε: regulator)

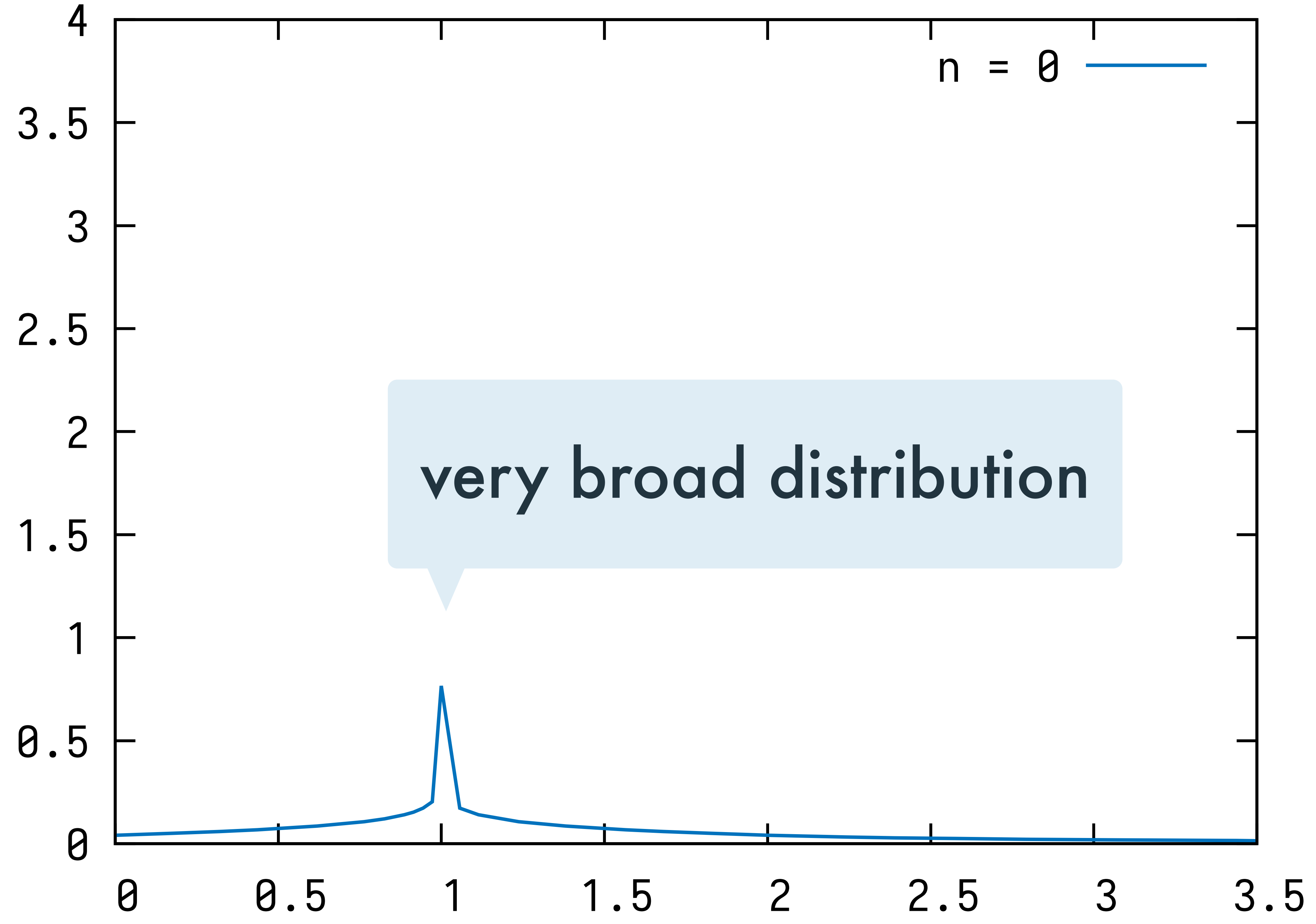
THE INFERENCE STEP — GEOMETRIC SERIES: $\delta_k = (0.7)^k$

- LO > $\delta_0 \equiv 1$

$$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

chose $\omega = 0$ for flat prior in a





very broad distribution

no inference yet!
$P(\delta_1)$ entirely determined by the *model & priors*

$$P(\delta_1) = \int \mathrm{d}a \int \mathrm{d}c\; P^{(1)}_{\text{geo}}(\delta_1\,|\,a,c)\; P_0(a,c)$$

The Inference Step — Geometric series: $\delta_k = (0.7)^k$

- **LO** > $\delta_0 \equiv 1$

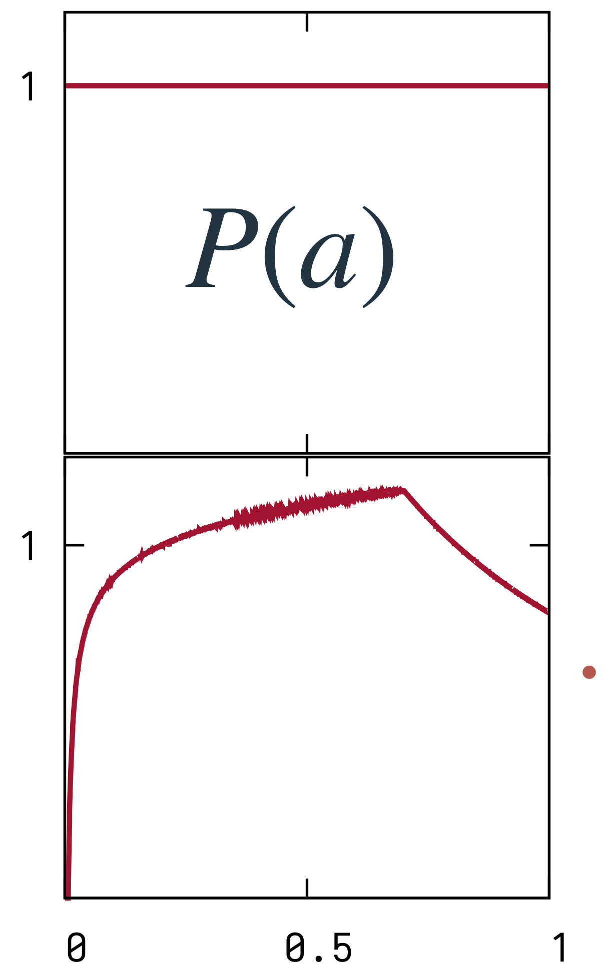
  $$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

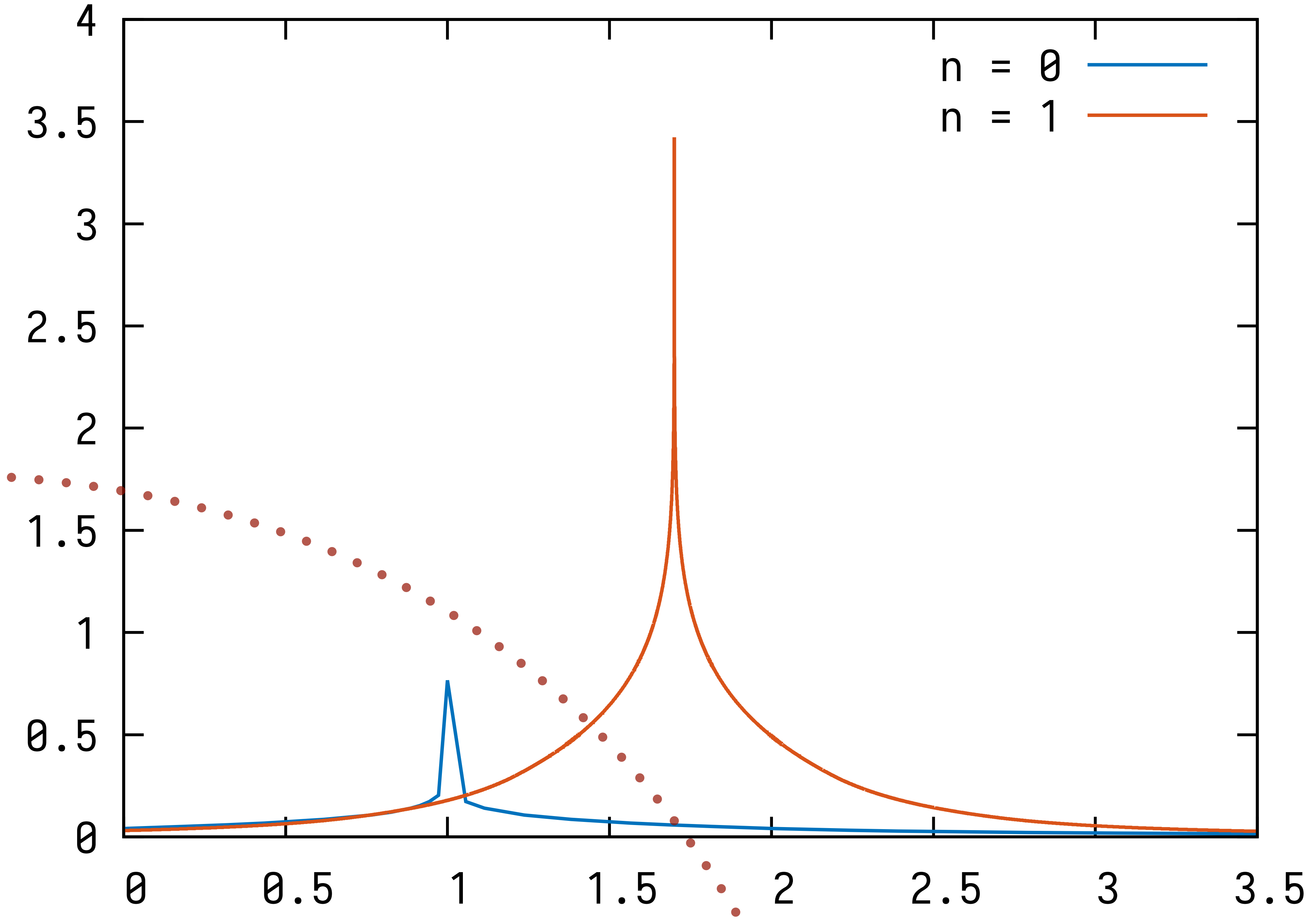
- **NLO** > $\delta_1 = 0.7$

  $$\underbrace{P(a,c\,|\,\delta_1)}_{\text{posterior}} \propto \underbrace{P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)}_{\text{likelihood}}\ \underbrace{P_0(a,c)}_{\text{prior}}$$

Bayes' theorem:

$$P(A\,|\,B) = \frac{P(B\,|\,A)\,P(A)}{P(B)}$$





$$P(\delta_2\,|\,\delta_1) = \int \mathrm{d}a \int \mathrm{d}c\ P(\delta_2\,|\,\delta_1,a,c)\ P(a,c\,|\,\delta_1)$$

$$\propto \int \mathrm{d}a \int \mathrm{d}c\ P^{(2)}_{\rm geo}(\delta_2\,|\,a,c)\ P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\ P_0(a,c)$$

δ_k independent:

$$P(\delta_2\,|\,\delta_1) = P(\delta_2)$$

The Inference Step — Geometric series: $\delta_k = (0.7)^k$

- **LO** > $\delta_0 \equiv 1$

$$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

- **NLO** > $\delta_1 = 0.7$

$$P(a,c\,|\,\delta_1) \propto P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\;P_0(a,c)$$

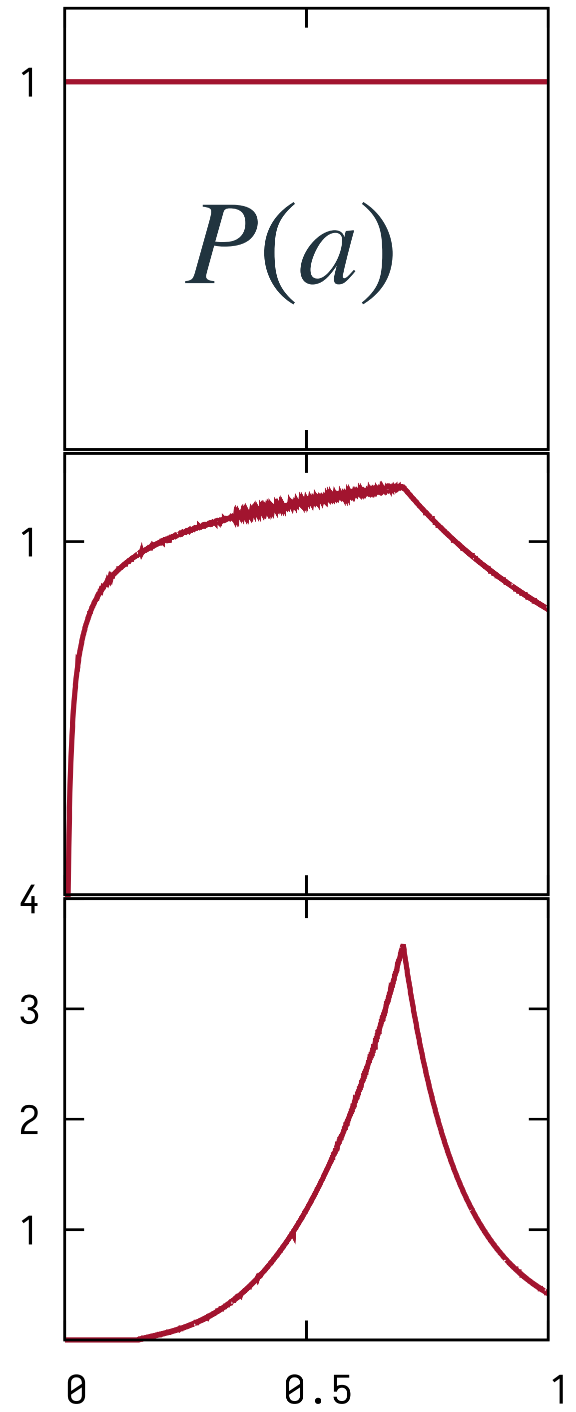
- **N²LO** > $\delta_2 = 0.7^2$

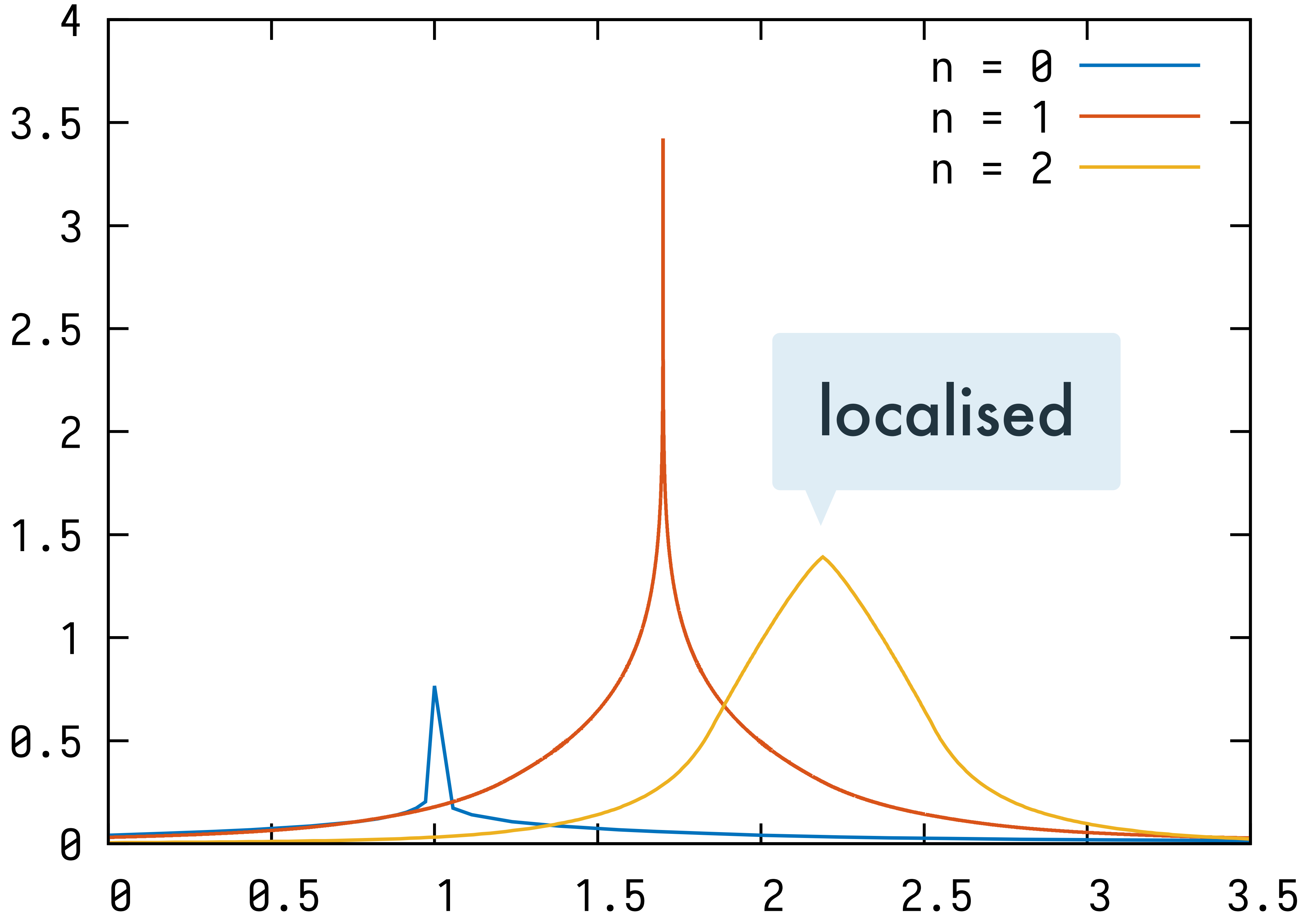
$$P(a,c\,|\,\delta_1,\delta_2) \propto P(\delta_2\,|\,\delta_1,a,c)\;P(a,c\,|\,\delta_1)$$

$$\propto P^{(2)}_{\rm geo}(\delta_2\,|\,a,c)\;P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\;P_0(a,c)$$

Bayes' theorem & independence

$a \sim 0.7$
also: $c \sim 1$





$$P(\delta_3\,|\,\delta_1,\delta_2) \propto \int \mathrm{d}a \int \mathrm{d}c\; \prod_{k=1}^{3}\left[P^{(k)}_{\rm geo}(\delta_k\,|\,a,c)\right] P_0(a,c)$$

The Inference Step — Geometric series: $\delta_k = (0.7)^k$

- LO > $\delta_0 \equiv 1$

$$P_0(a,c) = \Theta(a)\,\Theta(1-a)\,P_0(c)$$

- NLO > $\delta_1 = 0.7$

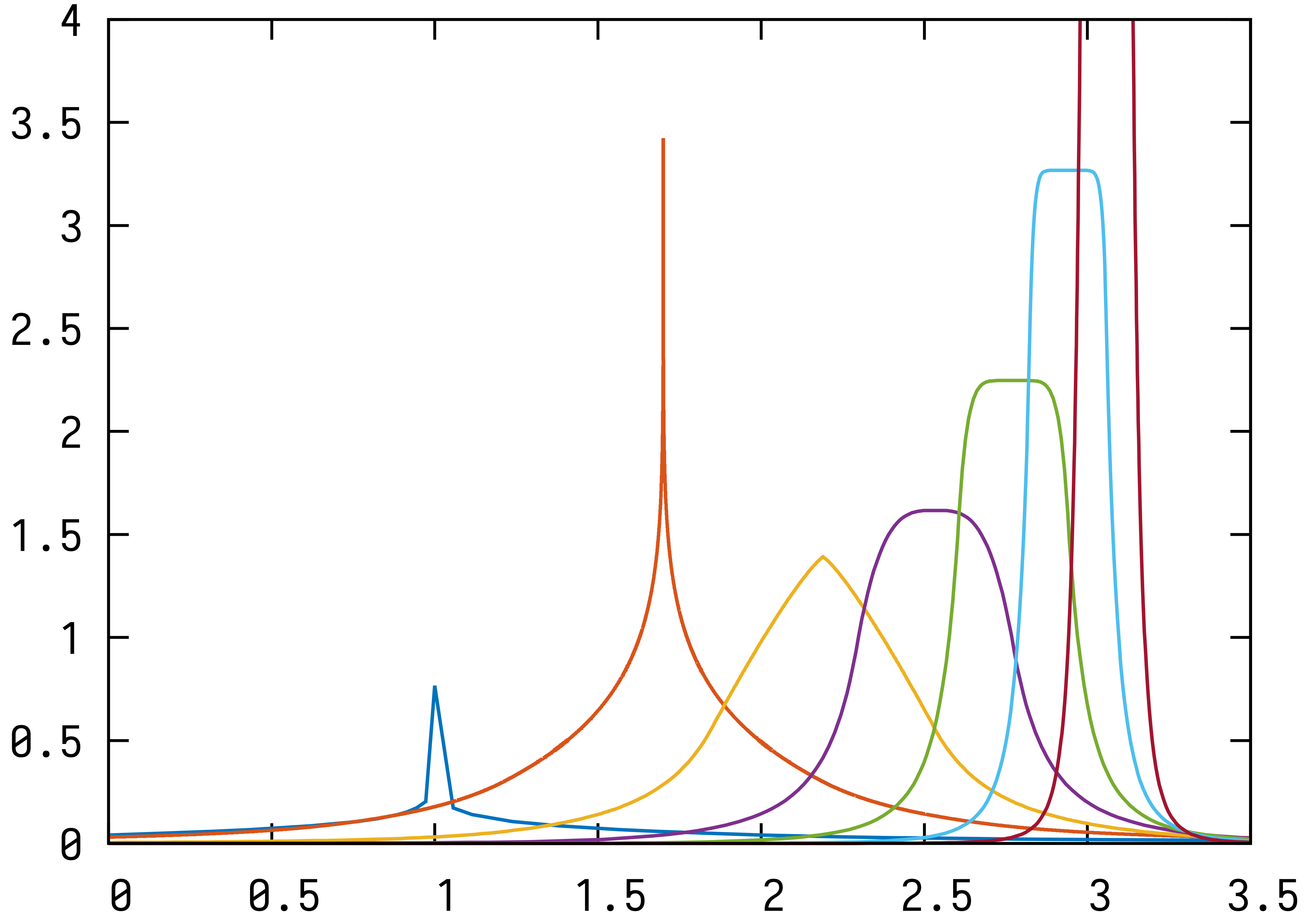
$$P(a,c\,|\,\delta_1) \propto P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\,P_0(a,c)$$

- N²LO > $\delta_2 = 0.7^2$

$$P(a,c\,|\,\delta_1,\delta_2) \propto P(\delta_2\,|\,\delta_1,a,c)\,P(a,c\,|\,\delta_1)$$

$$\propto P^{(2)}_{\rm geo}(\delta_2\,|\,a,c)\,P^{(1)}_{\rm geo}(\delta_1\,|\,a,c)\,P_0(a,c)$$

- …



$$P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n) \propto \int {\rm d}a \int {\rm d}c\, \prod_{k=1}^{n}\left[P^{(k)}_{\rm geo}(\delta_k\,|\,a,c)\right] P_0(a,c)$$

can be solved analytically

The *abc* Model — Asymmetric Geometric Model

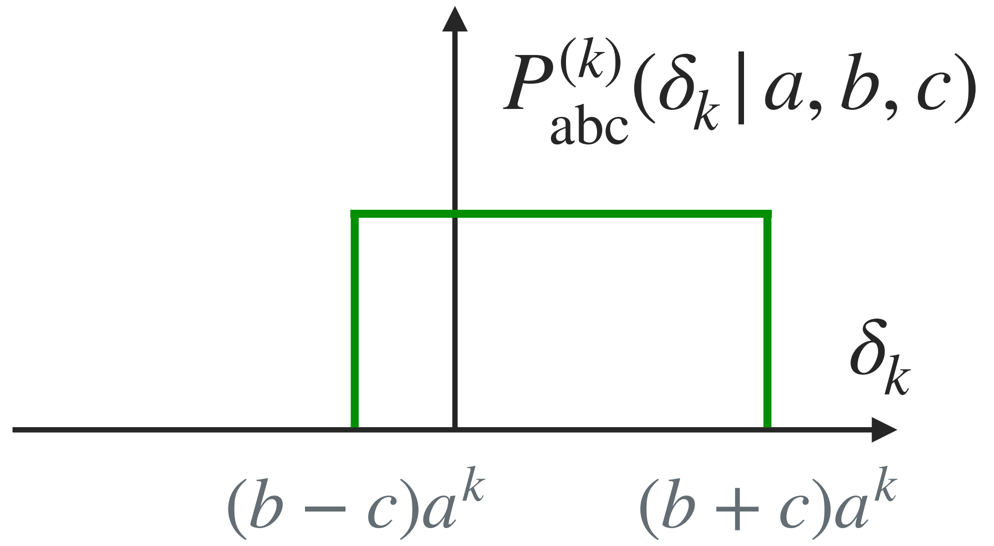
[Duhr, AH, Mazeliauskas, Szafron '21]

- geometric model is symmetric: $P(\delta_0, \ldots, \delta_n) = P(|\delta_0|, \ldots, |\delta_n|)$ ⇝ $\langle \delta_{n+1} \rangle_{\text{geo}} = 0$
- allow for different lower & upper bound:

$$b - c \leq \frac{\delta_k}{a^k} \leq b + c \quad \forall k$$

⇜⇝ three model parameters: a, b, c (bias / offset: b)

- **model:** $$P^{(k)}_{abc}(\delta_k \,|\, a, b, c) = \frac{1}{2c\,|a|^k}\, \Theta\left(c - \left| \frac{\delta_k}{a^k} - b \right| \right)$$



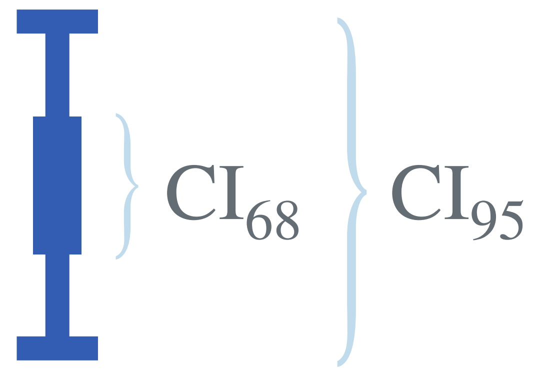
- **priors:** $$P_0(a, b, c) = P_0(a)\, P_0(b, c)$$

$$P_0(a) = \frac{1}{2}\,(1+\omega)\,(1-|a|)^\omega\, \Theta(1-|a|)$$

⇜⇝ support: [-1,+1] (alternating ✔)

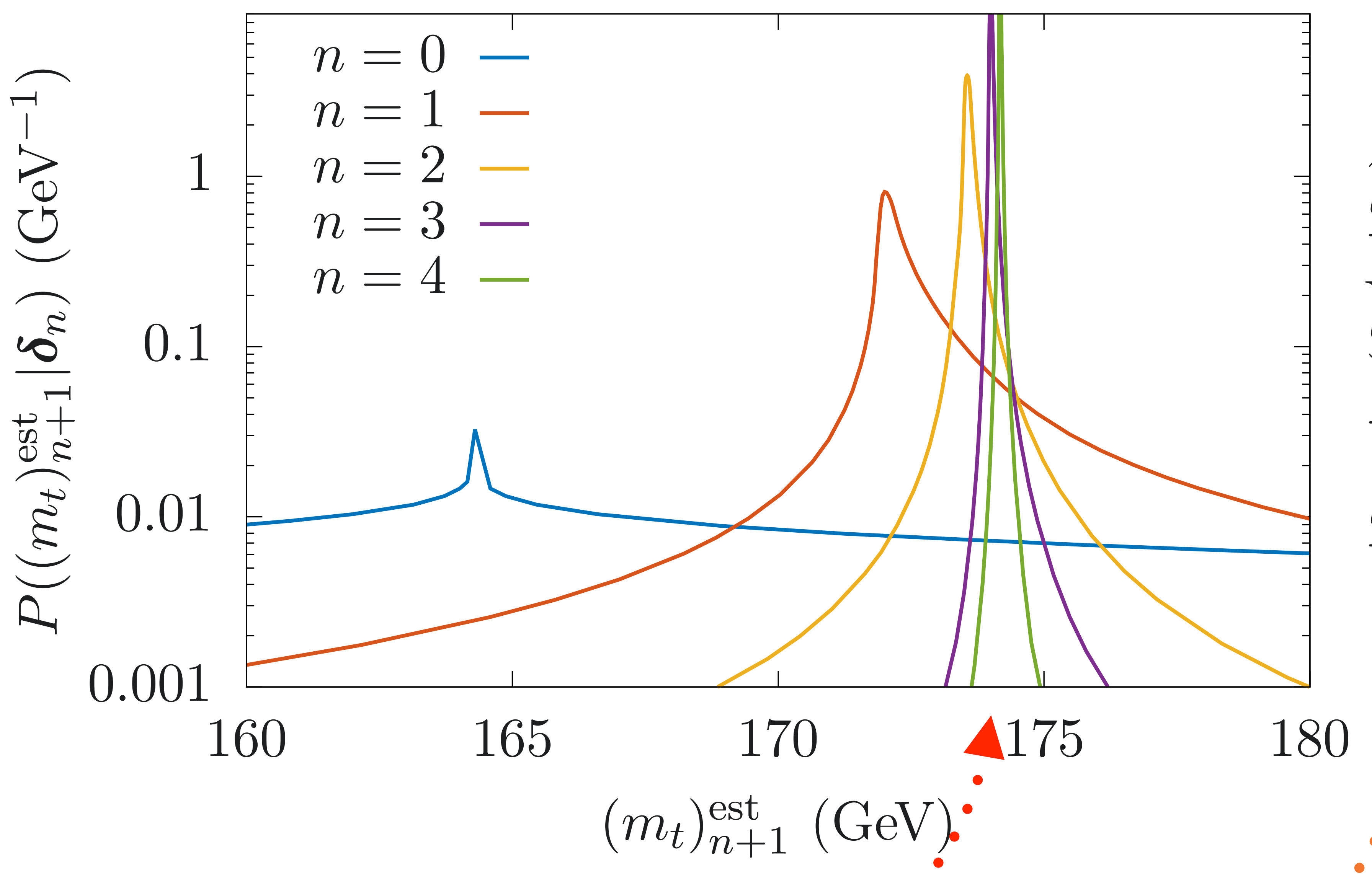
$$P_0(b, c) = \frac{\varepsilon \eta^\varepsilon}{c^{1+\varepsilon}}\, \Theta(c - \eta)\, \frac{1}{2\xi c}\, \Theta(\xi c - b)$$

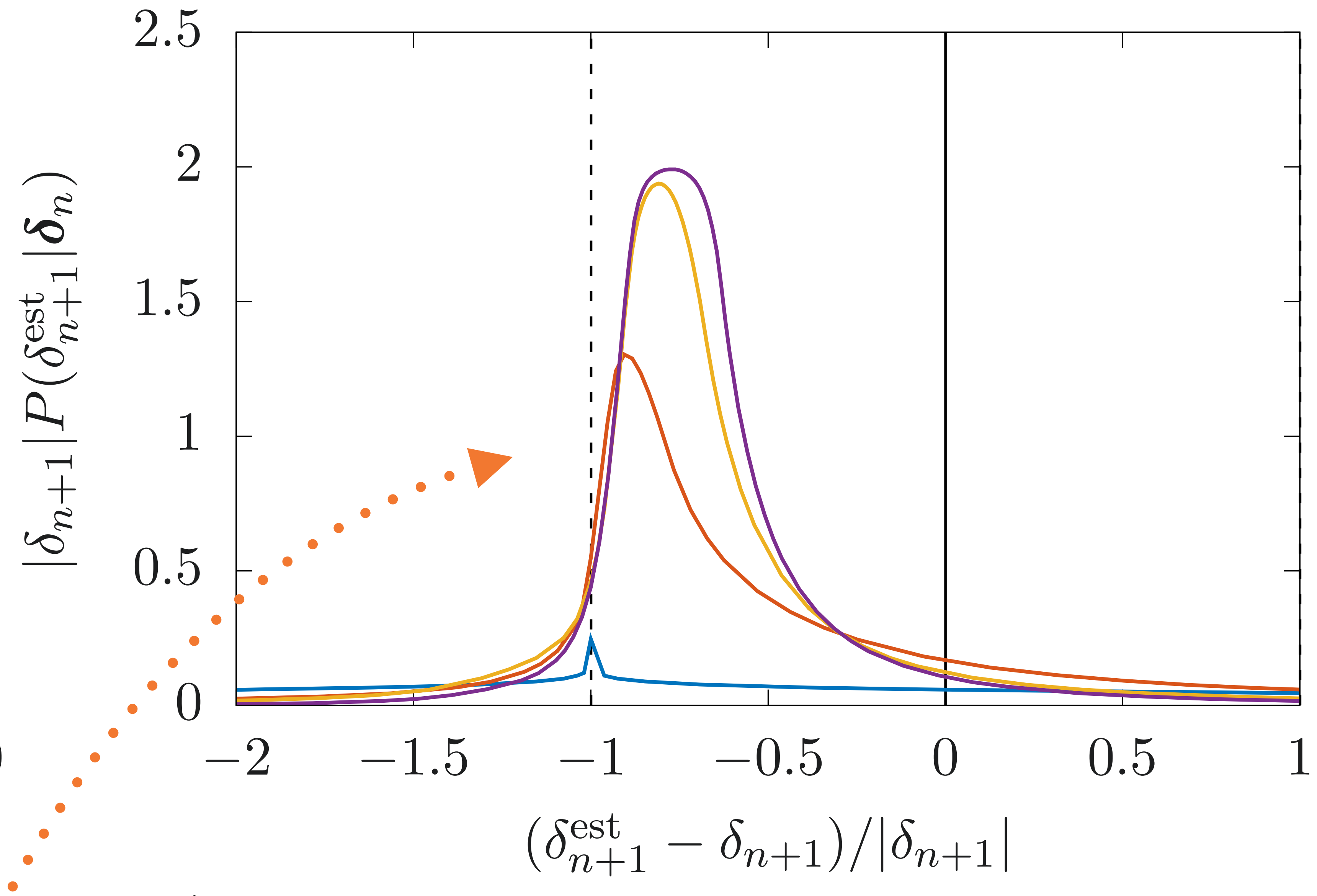
A Real-World Example — m_t (OS $\leftrightsquigarrow$ $\overline{\mathrm{MS}}$)



$$m_\mathrm{t} = \frac{Z_m^{\mathrm{OS}}}{Z_m^{\overline{\mathrm{MS}}}(\mu_R)}\,\overline{m}_\mathrm{t}(\mu_R)$$

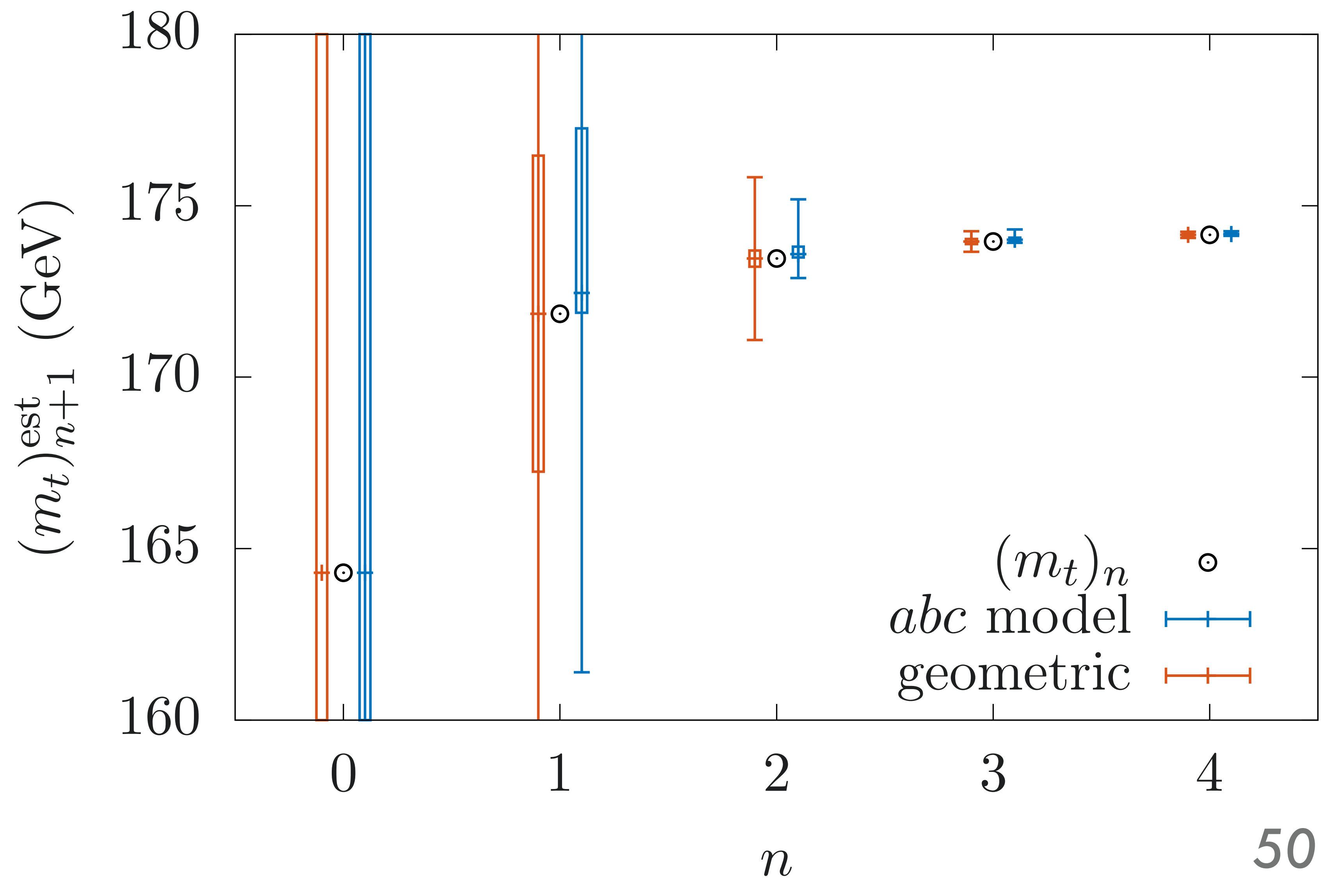
$$= \sum_k m_\mathrm{t}^{(k)}$$





strongly peaked $n \nearrow$

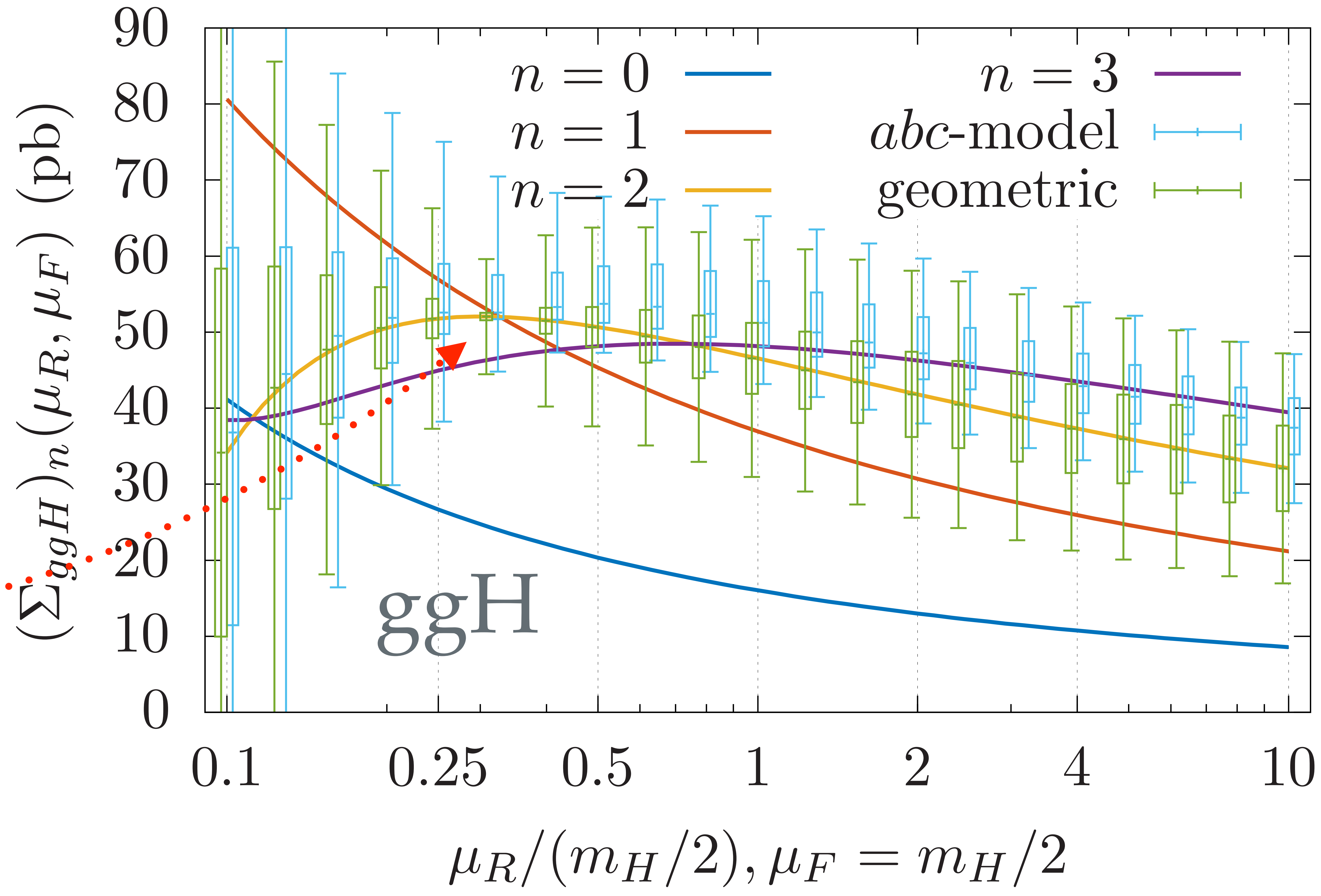
positive corrections anticipated



- estimate for $m_\mathrm{t} - (m_\mathrm{t})_4$
  - $\mathrm{CI}_{68} = [0.008, 0.046]$ GeV
  - $\mathrm{CI}_{95} = [-0.027, 0.112]$ GeV

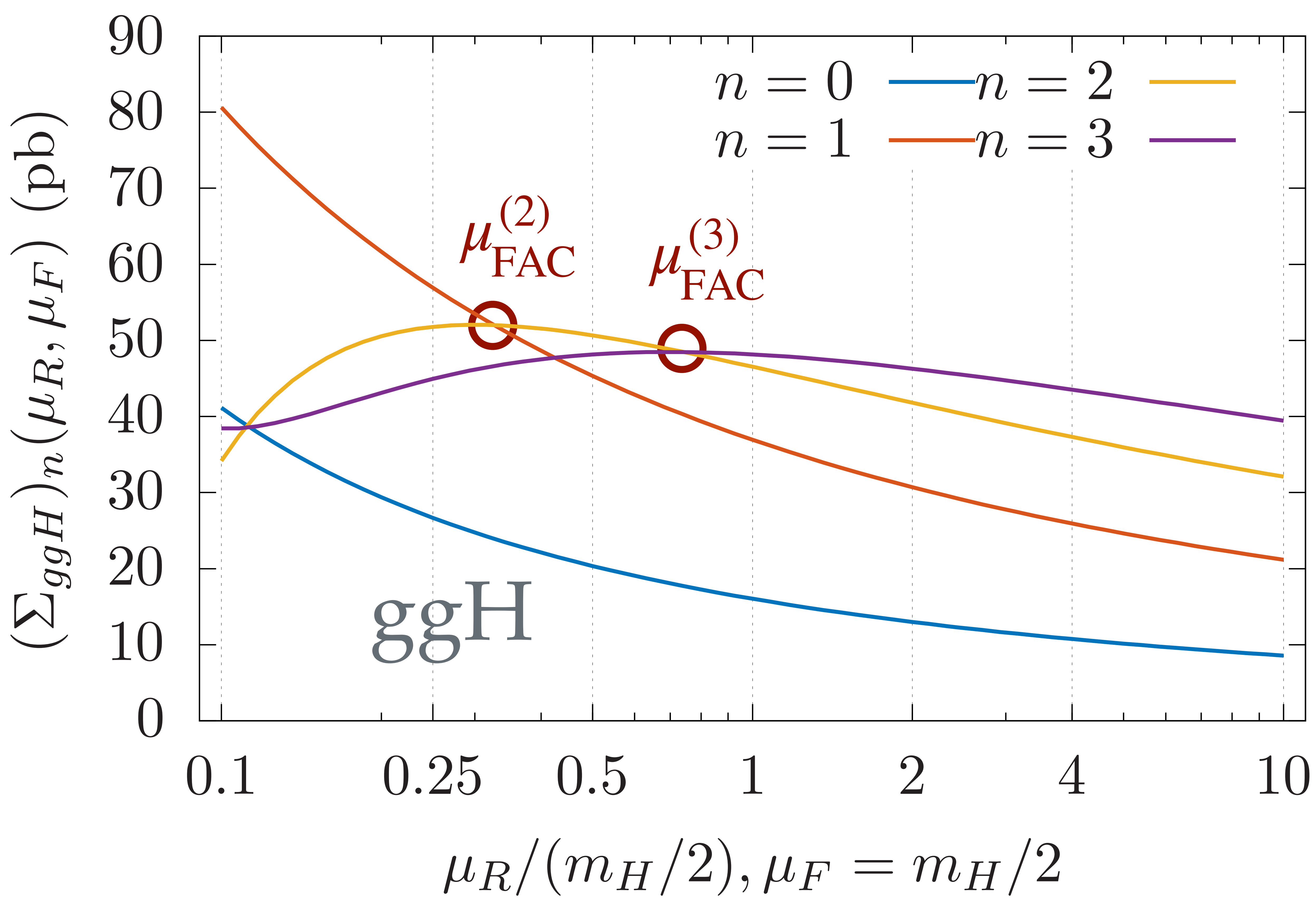
What to do with the the scale μ?

- $\forall\mu \rightsquigarrow P(\delta_3 \,|\, \delta_0, \delta_1, \delta_2; \mu)$
  - $\text{CI}_{68/95}$ (geo) (abc)
- geo
  - always entered around NNLO
  - very narrow peak
- abc
  - $\mu/\mu_0 \gtrsim 1 \rightsquigarrow$ anticipate pos. N3LO
  - $\mu/\mu_0 \lesssim 1 \rightsquigarrow$ bias slowly disappears



WHAT TO DO WITH THE THE SCALE μ?

- $\forall\mu \rightsquigarrow P(\delta_3 \mid \delta_0, \delta_1, \delta_2; \mu)$
  - $\mathrm{CI}_{68/95}$ (geo) (abc)
- two options:
  1. invoke some *principle* to pick the *"optimal"* scale
     - FAC, PMS, PMC, …

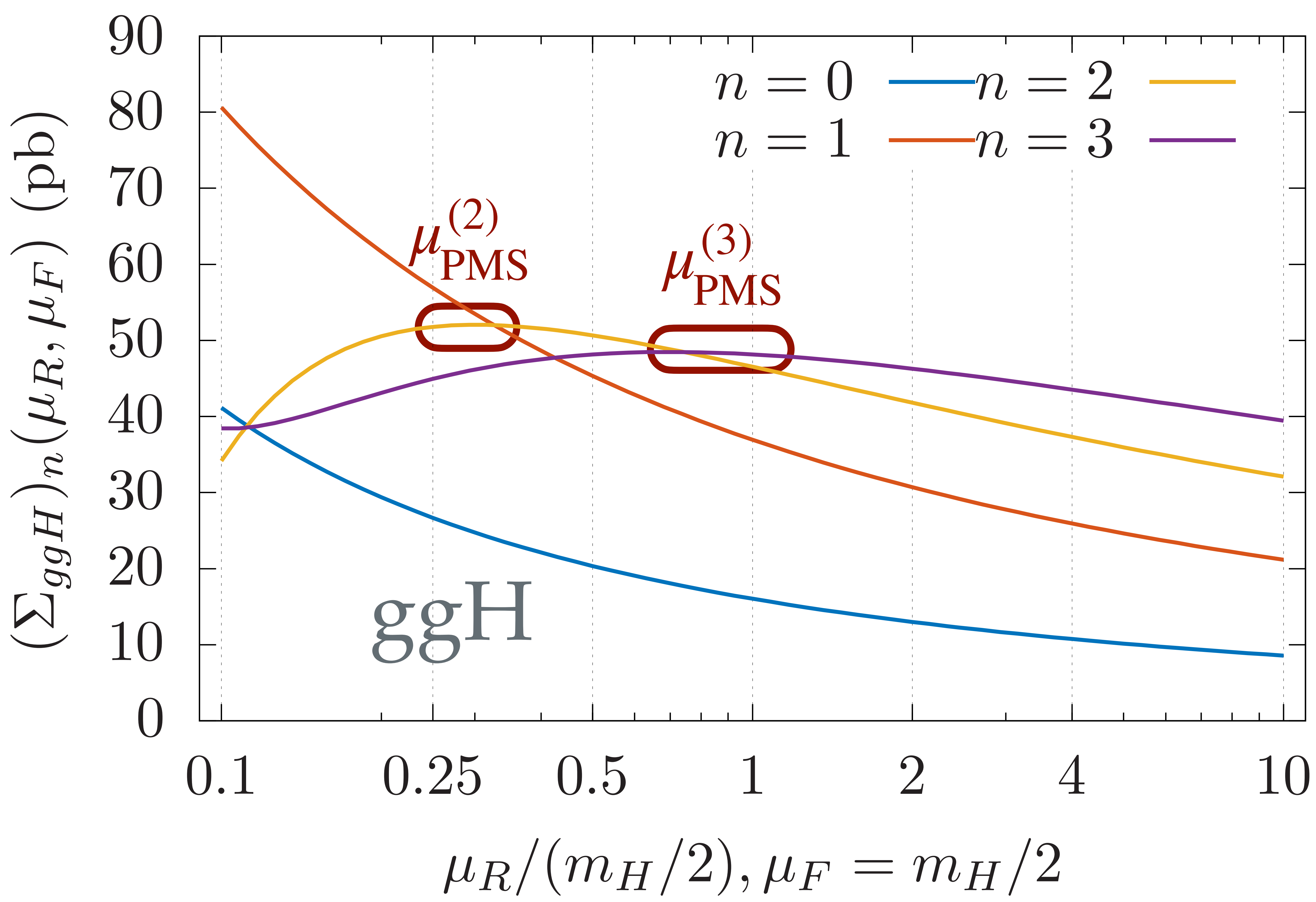


Fastest Apparent Convergence
$\Sigma_n(\mu_{\mathrm{FAC}}) = \Sigma_{n-1}(\mu_{\mathrm{FAC}})$

depends on order
might not be unique

What to do with the the scale μ?

- $\forall\mu \rightsquigarrow P(\delta_3\,|\,\delta_0, \delta_1, \delta_2; \mu)$
  - $\mathrm{CI}_{68/95}$ (geo) (abc)
- two options:
  1. invoke some *principle* to pick the *"optimal"* scale
     - FAC, PMS, PMC, …



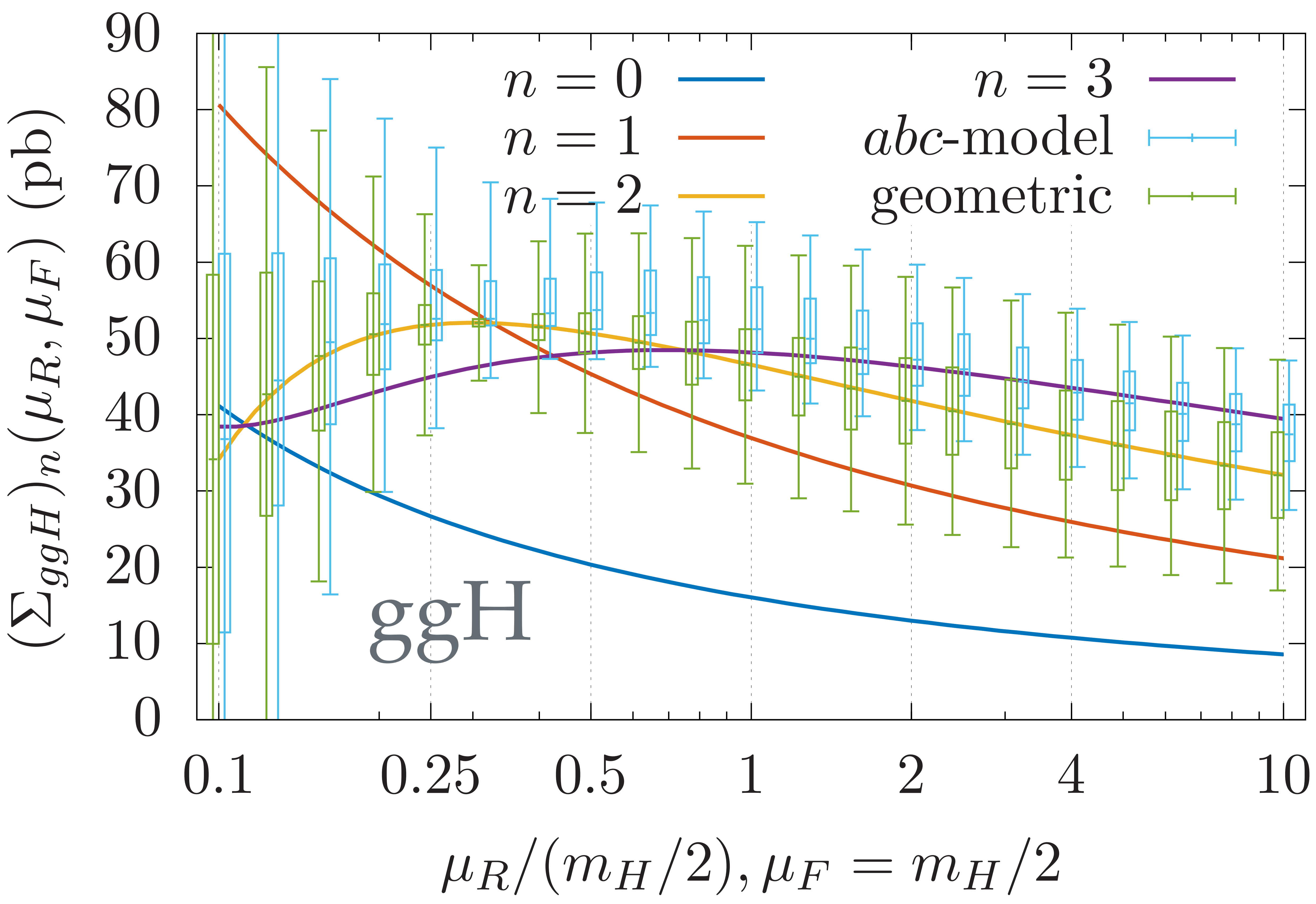
Principle of Minimal Sensitivity

$$\frac{\partial}{\partial\mu}\Sigma_n(\mu)\,\big|_{\mu_{\mathrm{PMS}}} = 0$$

depends on order
might not be unique

What to do with the the scale μ?

- $\forall\mu \rightsquigarrow P(\delta_3 \,|\, \delta_0, \delta_1, \delta_2; \mu)$
  - $\mathrm{CI}_{68/95}$ (geo) (abc)



- two options:
  1. invoke some principle to pick the "optimal" scale
     - FAC, PMS, PMC, …
  2. combine different $P(\delta_{n+1} \,|\, \boldsymbol{\delta}_n; \mu)$

pursued in the following

Prescriptions for Scales

Scale Marginalisation (sm):

[Bonvini '20]

- treat μ as a hidden model parameter
  & *marginalise* over it:

$$P_{\rm sm}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n) = \int \mathrm{d}\mu\; P(\delta_{n+1}, \mu\,|\,\boldsymbol{\delta}_n)$$

$$= \int \mathrm{d}\mu\; P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n;\mu)\; P(\mu\,|\,\boldsymbol{\delta}_n)$$

- $P(\mu\,|\,\boldsymbol{\delta}_n) \propto P(\boldsymbol{\delta}_n;\mu)\; P_0(\mu)$ with prior:

$$P_0(\mu) = \frac{1}{2\mu \ln F}\,\Theta\!\left(\ln F - \left|\ln\!\left(\frac{\mu}{\mu_0}\right)\right|\right)$$

Scale Average (sa):

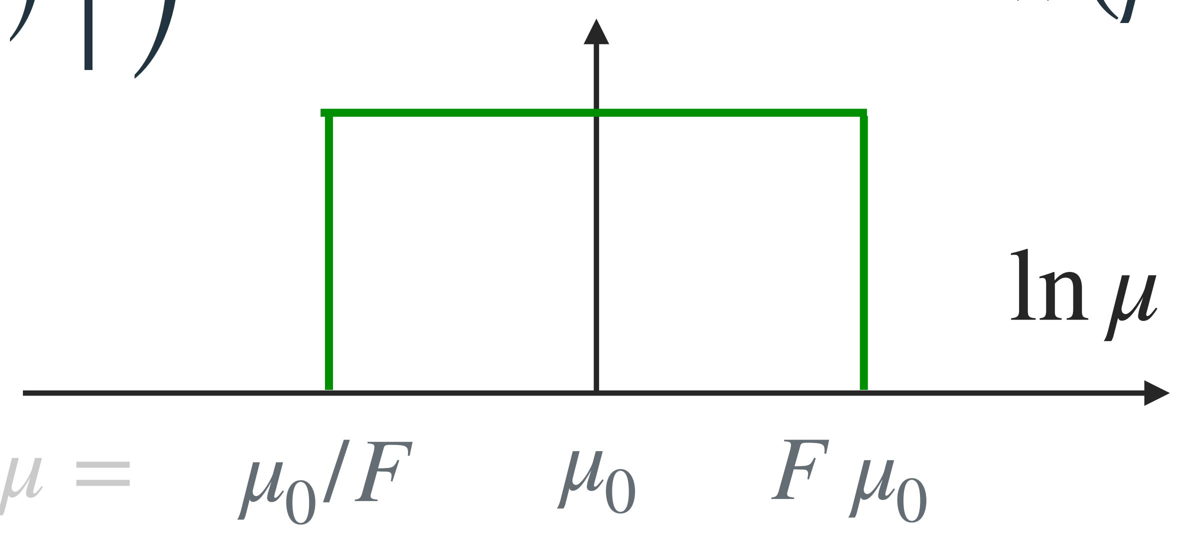
[Duhr, AH, Mazeliauskas, Szafron '21]

- μ has no probabilistic interpretation
  ⇝ *average* over it:

$$P_{\rm sa}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n) = \int \mathrm{d}\mu\; w(\mu)\; P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n;\mu)$$

- weight function:

$$w(\mu) = \frac{1}{2\mu \ln F}\,\Theta\!\left(\ln F - \left|\ln\!\left(\frac{\mu}{\mu_0}\right)\right|\right)$$



Peak of the Distributions*

[Duhr, AH, Mazeliauskas, Szafron '21]

Scale Marginalisation (sm):

- if $\mu_{\mathrm{FAC}} \in [\mu_0/F, F\,\mu_0]$ then $P_{\mathrm{sm}}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n)$ peaks at $\Sigma_n(\mu_{\mathrm{FAC}})$
  - $P(\boldsymbol{\delta}_n\,|\,\mu)$ dominated by $(k=n)$ term
  - symmetric model
    ⇝ $\delta_n(\mu) = 0$ enhanced

Scale Average (sa):

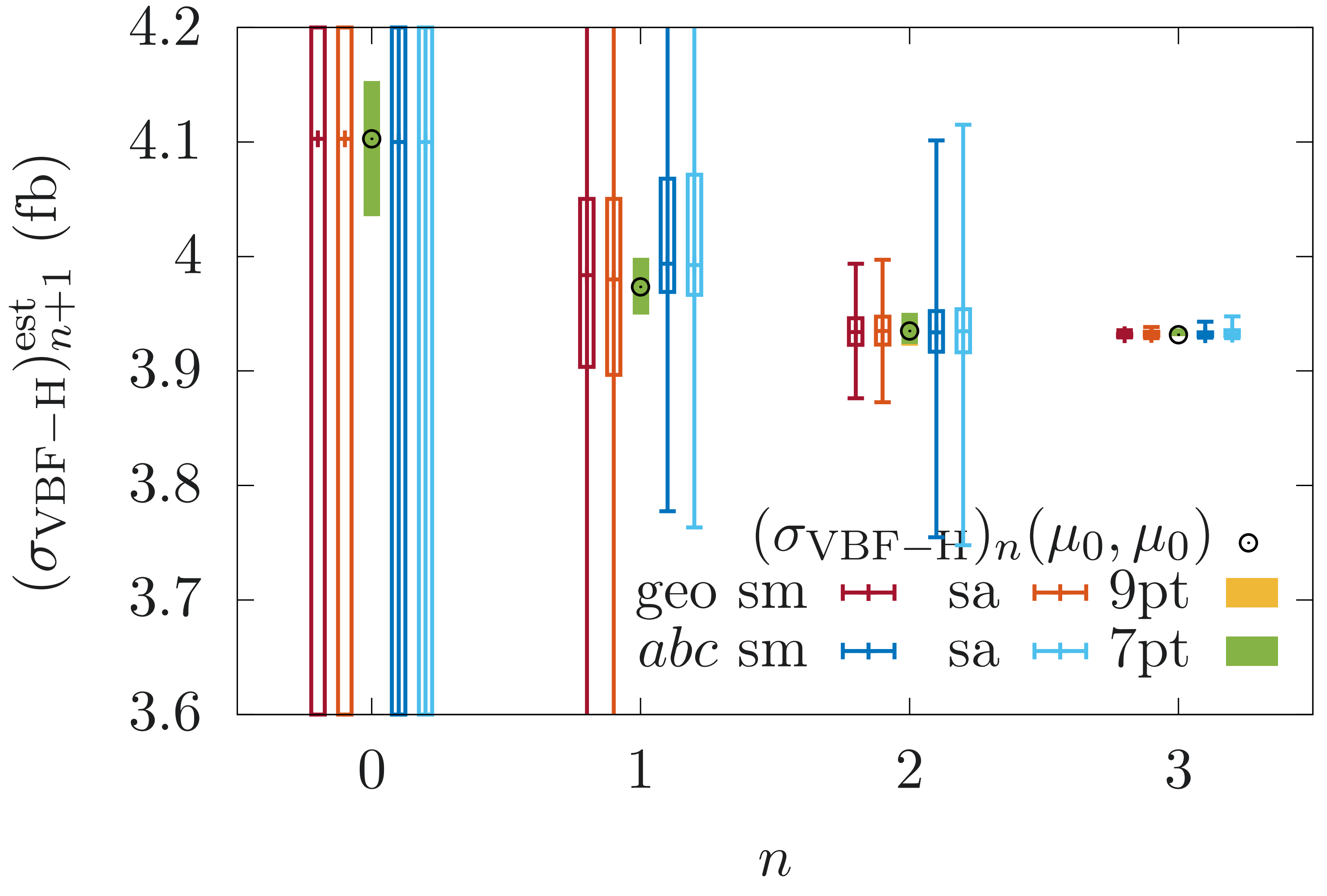
- if $\mu_{\mathrm{PMS}} \in [\mu_0/F, F\,\mu_0]$ then $P_{\mathrm{sa}}(\delta_{n+1}\,|\,\boldsymbol{\delta}_n)$ peaks at $\Sigma_n(\mu_{\mathrm{PMS}})$
  - overlap between $P(\delta_{n+1}\,|\,\boldsymbol{\delta}_n;\mu)$ enhanced at stationary point
    ⇝ $\Sigma_n'(\mu_{\mathrm{PMS}}) \approx 0$

Choice of how to interpret the scale has consequences for predictions!

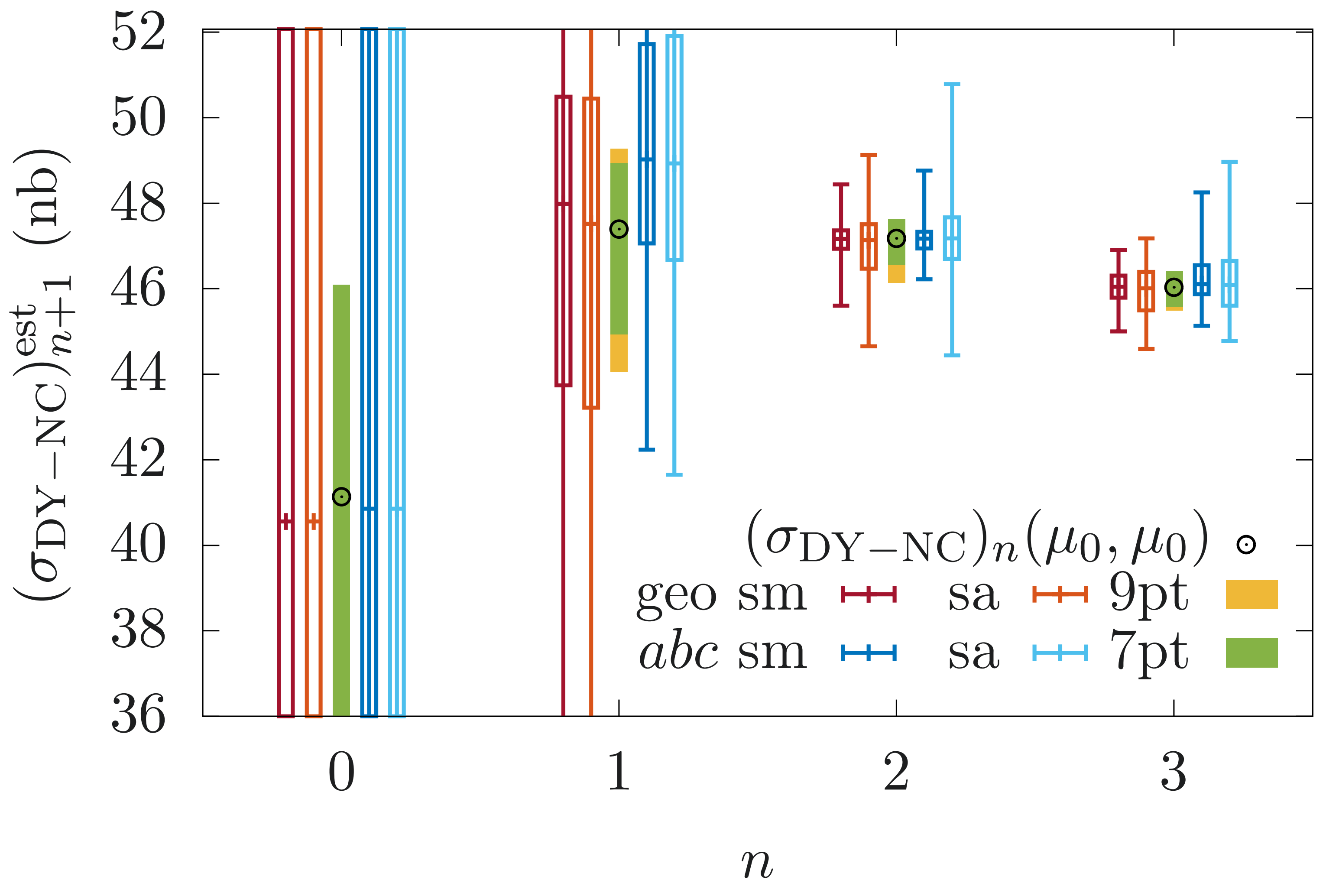
* for symmetric models, a convergent series, and reasonable assumptions

Inclusive Cross Sections up to N3LO

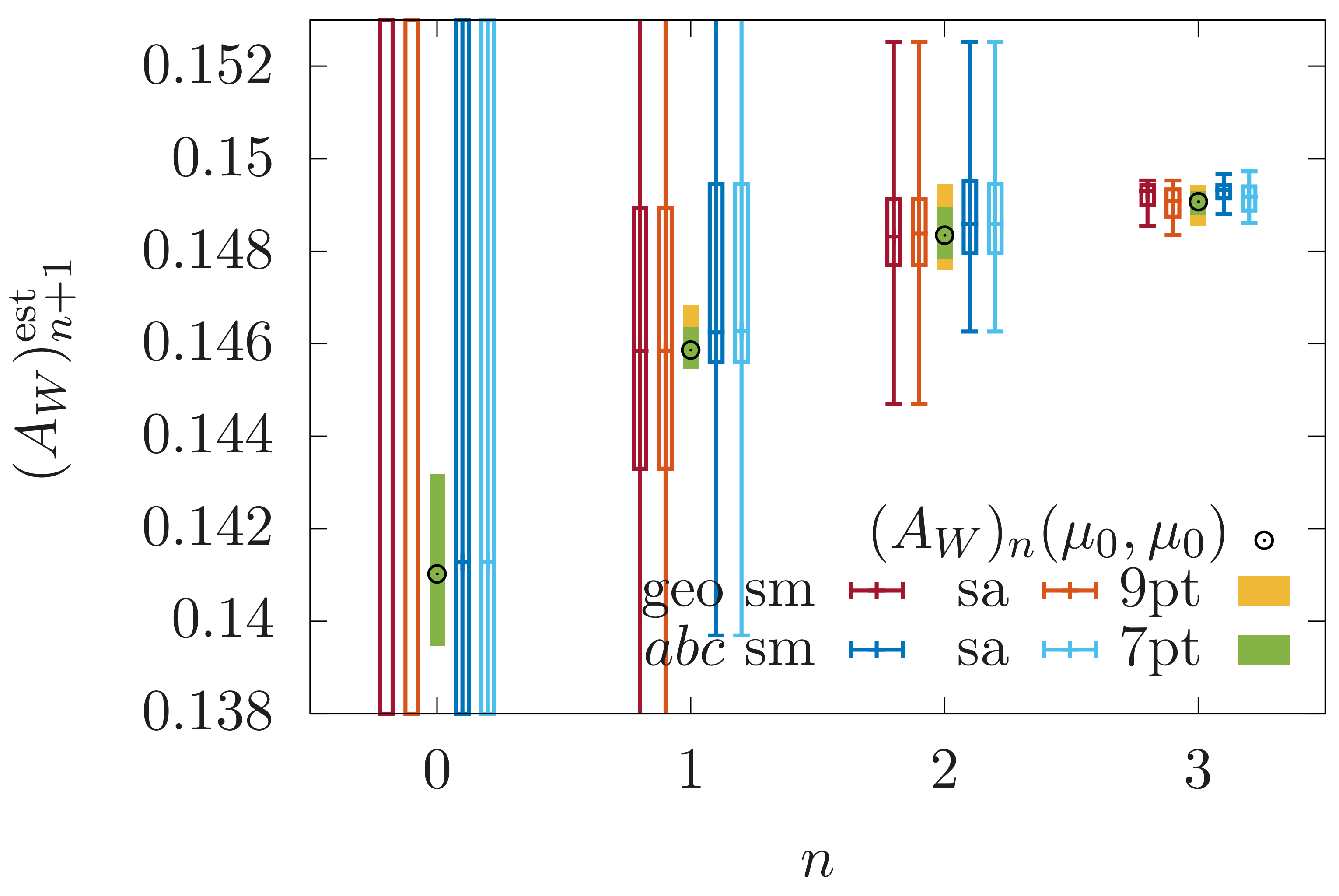
VBF–H



DY–NC



$A_W = \dfrac{W^+ - W^-}{W^+ + W^-}$



- $n < 2$: CI_{68} bigger than 9pt
- $\delta_1 < 0 \rightsquigarrow$ *abc* alternating
- $n > 2$: all prescriptions similar

- δ_3 is large and outside of 9pt!
- similar unc.: sa $\simeq$ 9pt
- $n = 2$: sm $\ll$ others (μ_{FAC})
- $n = 3$: all prescriptions similar

- large cancellations in the ratio
- $n < 2$: 9pt performs poorly
- $(A_W)_n \nearrow$ (anticipated by *abc*)
- size: *abc* $\lesssim$ others

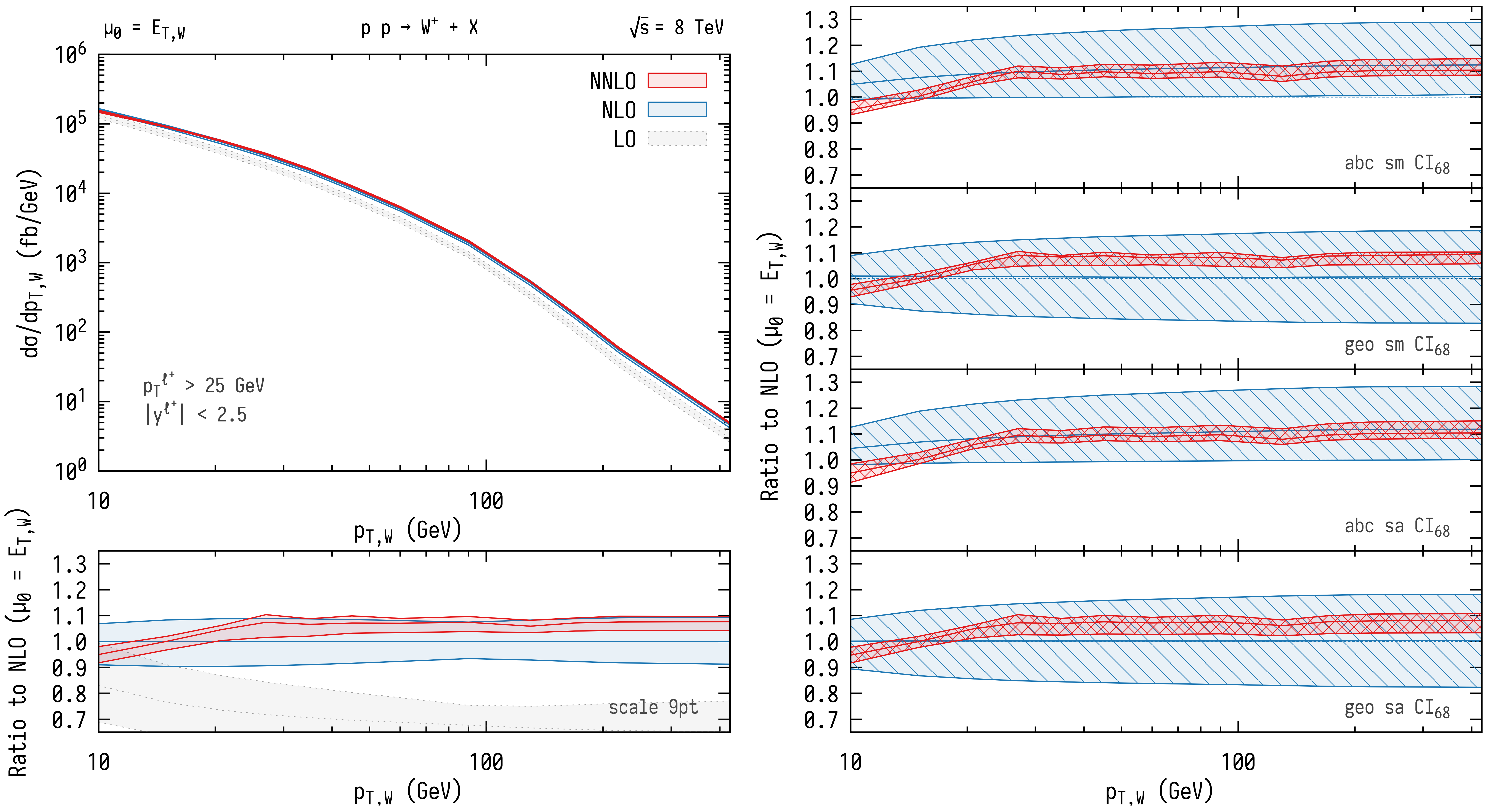
overall: not radically different estimates for Δ_{MHO}

Differential Distributions

- Bayesian approach also applicable to distributions
  - ⇝ treat each bin individually ⇜⇝ will not include correlations!

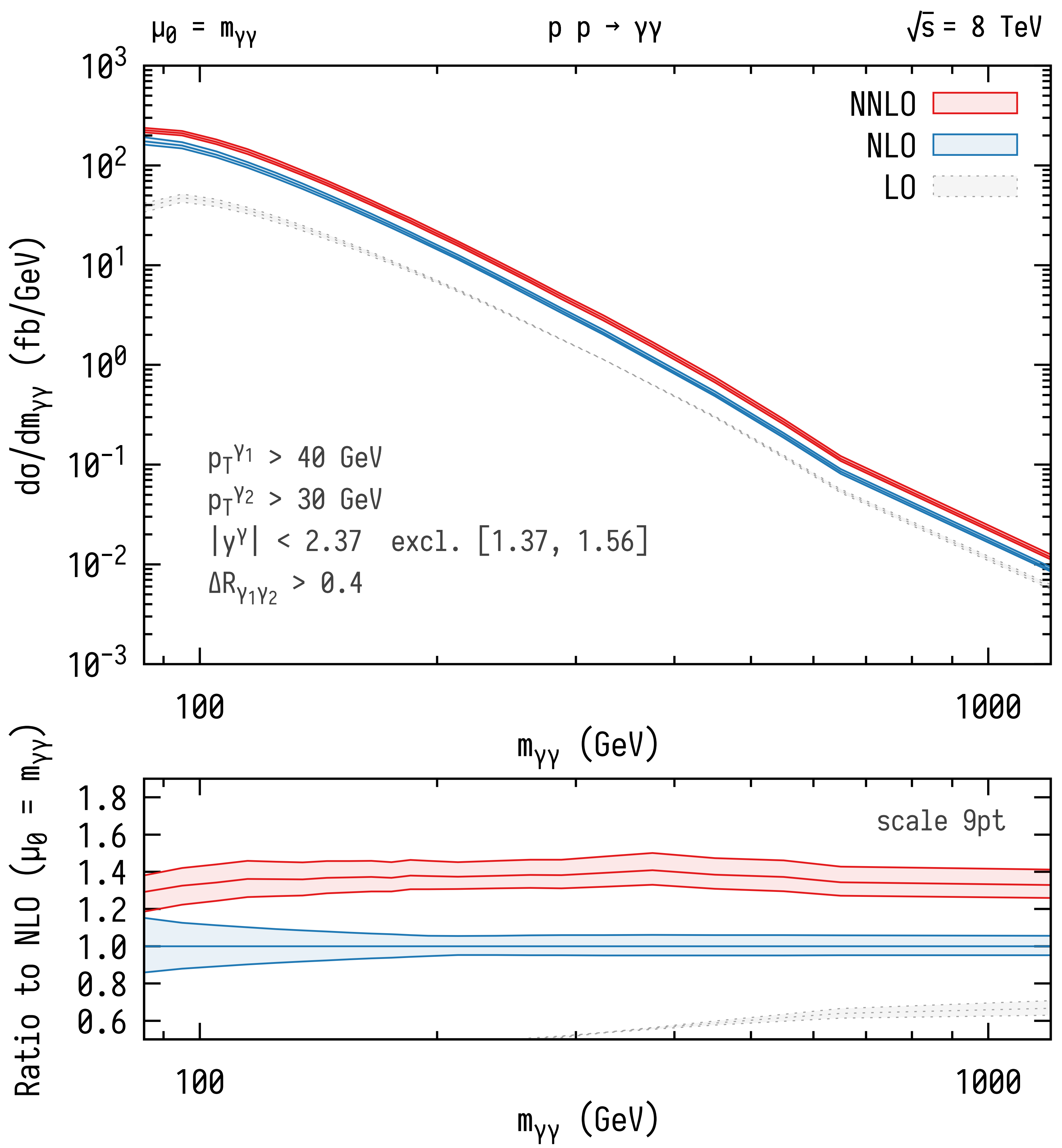
- new challenges
  - no longer "easy" to identify an appropriate hard scale μ_0 (up to rescaling)
    - ⇝ inclusive ggH: M_H vs. $\frac{1}{2} M_\mathrm{H}$? Just let the model figure it out.
  - differential distributions can probe different kinematic regimes
    - ⇝ dynamical scale choice ⇜⇝ *many choices!*
    - ⇝ e.g. in jet production: $p_\mathrm{T}^j\,,\ p_\mathrm{T}^{j_1}\,,\ \langle p_\mathrm{T}^j \rangle_\mathrm{avg}\,,\ H_\mathrm{T} \equiv \sum_{i \in \mathrm{jets}} p_\mathrm{T}^i\,,\ \hat{H}_\mathrm{T} \equiv \sum_{i \in \mathrm{partons}} p_\mathrm{T}^i\,, \ldots$
  - re-cycling via quadrature limited ⇝ ideally interpolation grids

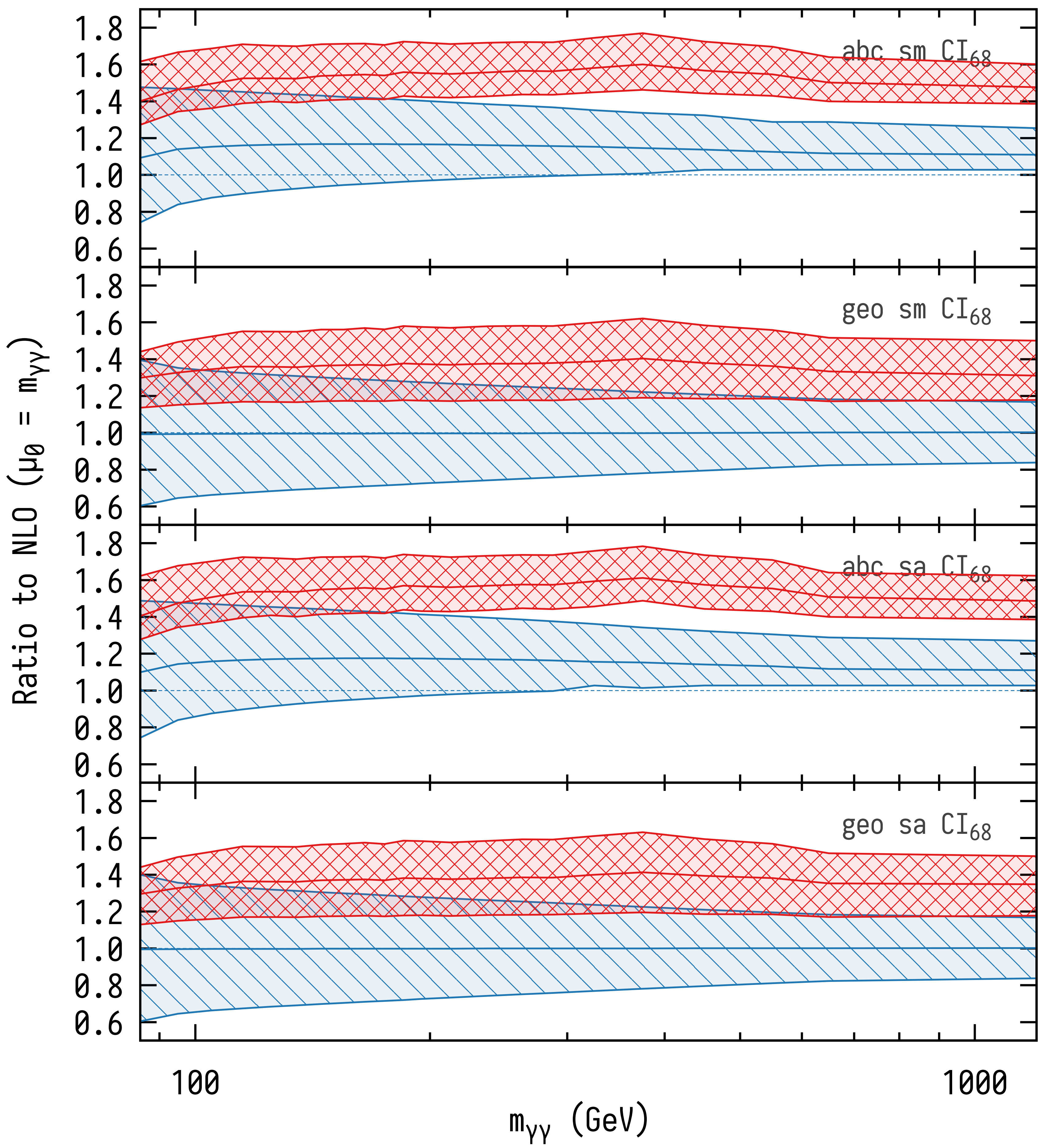
W-Boson + Jet Production



- $n < 2$:
  - CI_{68} bigger than 9pt
  - *abc* captures pos. shift
- $n = 2$:
  - almost identical bands
  - Δ_{MHO} very robust
- sm vs. sa
  - almost identical CI

Di-Photon Production

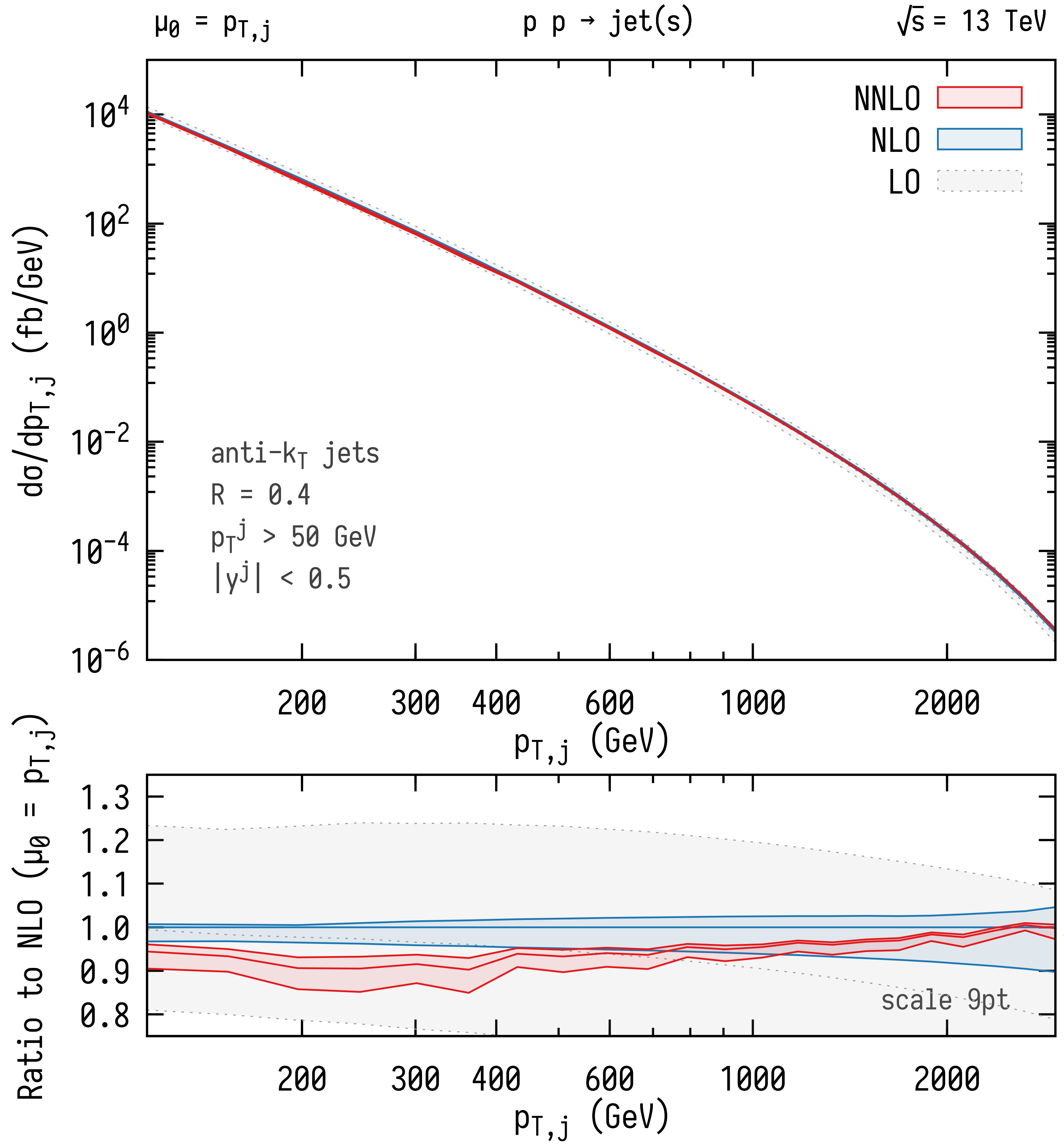




- example where 9pt fails
  - large corrections
  - $\Delta_{\text{MHO}}^{\text{NNLO}} \gtrsim \Delta_{\text{MHO}}^{\text{NLO}}$
  - no sign of convergence
- $n < 2$:
  - CI_{68} ~ 2-3 × 9pt
- $n = 2$:
  - marginal overlap for geo
  - differences in *size* & *position*
  - ideally N3LO for robust Δ_{MHO}
- sm ≃ sa
  - large corrections prohibit FAC points

The Problem with Jets...

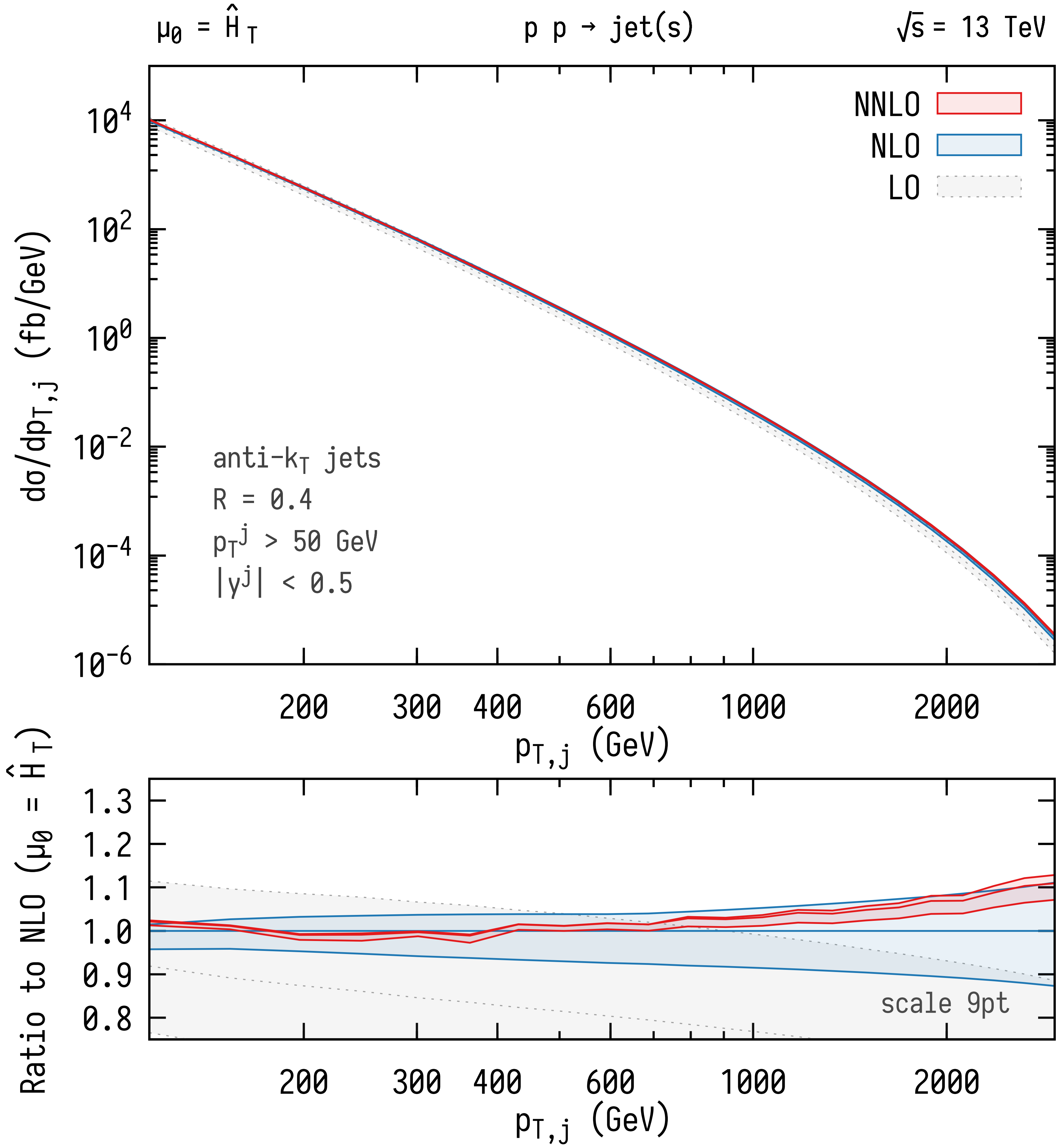
$\mu_0 = p_T^j$: infrared sensitivity



?

instability at low p_T

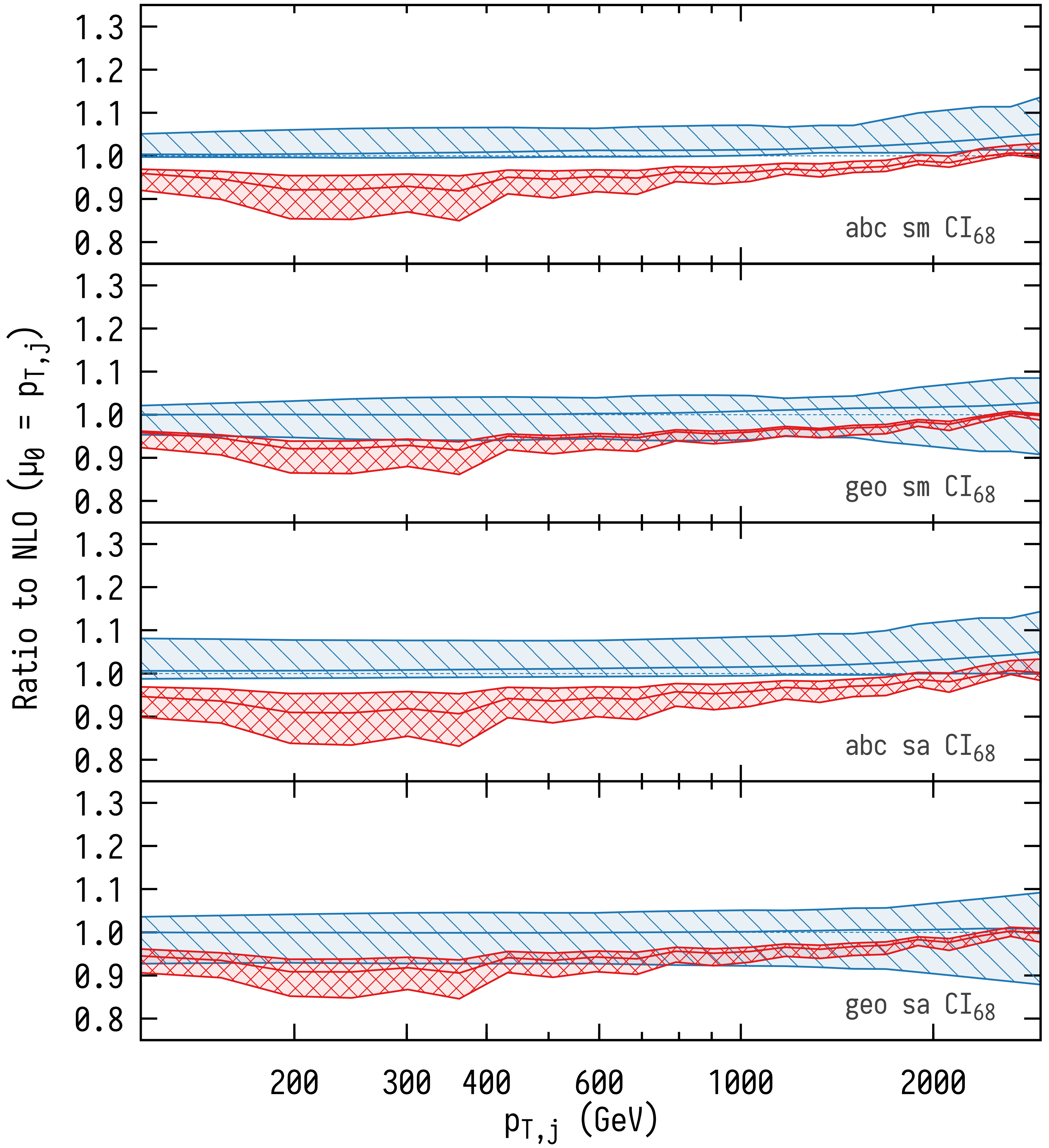
$\mu_0 = \hat{H}_T$: recommendation [Currie et al. '18]



The Problem with Jets... Persists

non-trivial change of dynamical scales
cannot be captured by a simple re-scaling

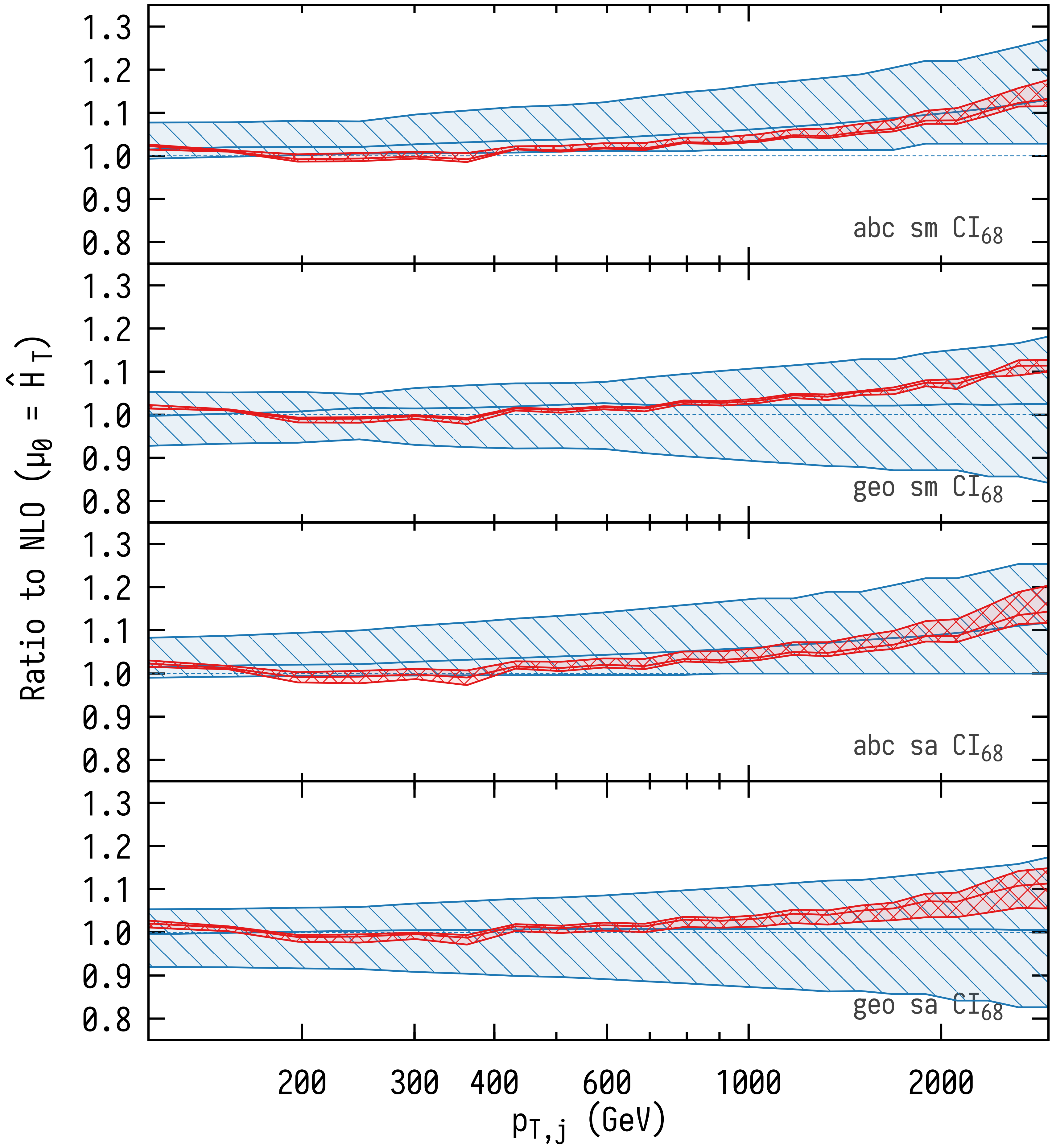
$\mu_0 = p_T^j$: infrared sensitivity



larger NLO Δ_{MHO}

barely any difference at NNLO

$\mu_0 = \hat{H}_T$: recommendation [Currie et al. '18]



abc captures positive corrections

Work in Progress — Correlations

- idea: if two bins show similar (opposite) perturbative behaviour
  ↪ two bins should be partially (anti-)correlated.
- we want: joint probability distribution $P(x, y)$ for two bins x & y
  ↪ preserve projections for compatibility:

$$P(x) = \int \mathrm{d}y\, P(x, y) = \int \mathrm{d}z\, P(x, z)$$

  ↪ hidden parameter $-1 < c < +1$ to smoothly implements the correlation
- possibilities: algorithmic "earth movers distance"; map $P(x)$ onto $P(y)$, …
  ↪ can be done much simpler

Work in Progress – Correlation Model in miho

- projections of multi-dim. Gaussians (+ correlation matrix) are again Gaussian
  ↪ map P_i onto Gaussians, implement correlations, map back

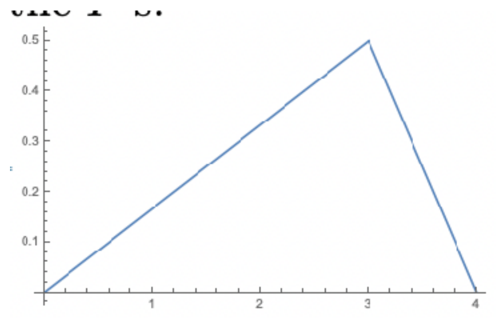
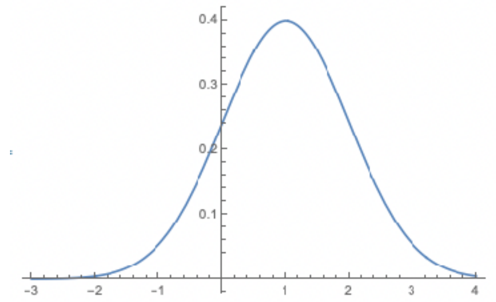
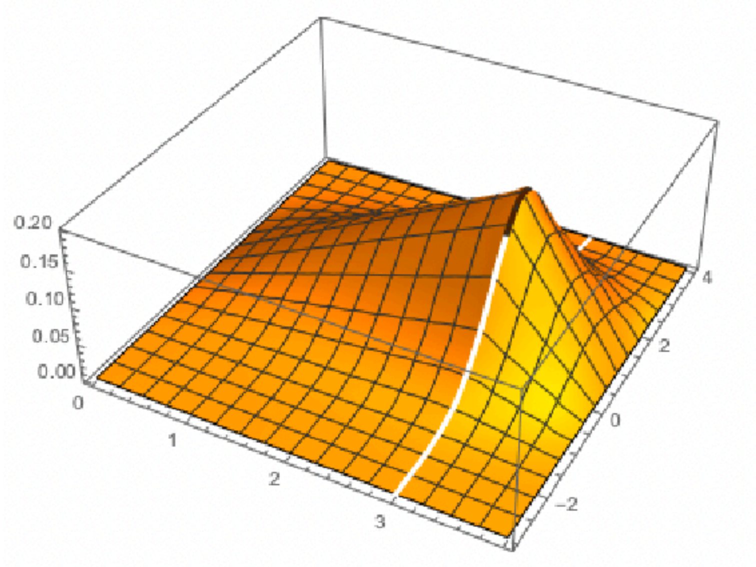
$$P(x,y) = P_1(x)P_2(y) \times \left.\frac{d\Phi^{-1}(\alpha)}{d\alpha}\right|_{\alpha=\Sigma_1(x)} \left.\frac{d\Phi^{-1}(\beta)}{d\beta}\right|_{\beta=\Sigma_2(y)} \times \frac{1}{2\pi\sqrt{1-c^2}} \exp\left(-\frac{1}{2(1-c^2)}\left[\xi(x)^2+\eta(y)^2-c2\xi(x)\eta(y)\right]\right)$$

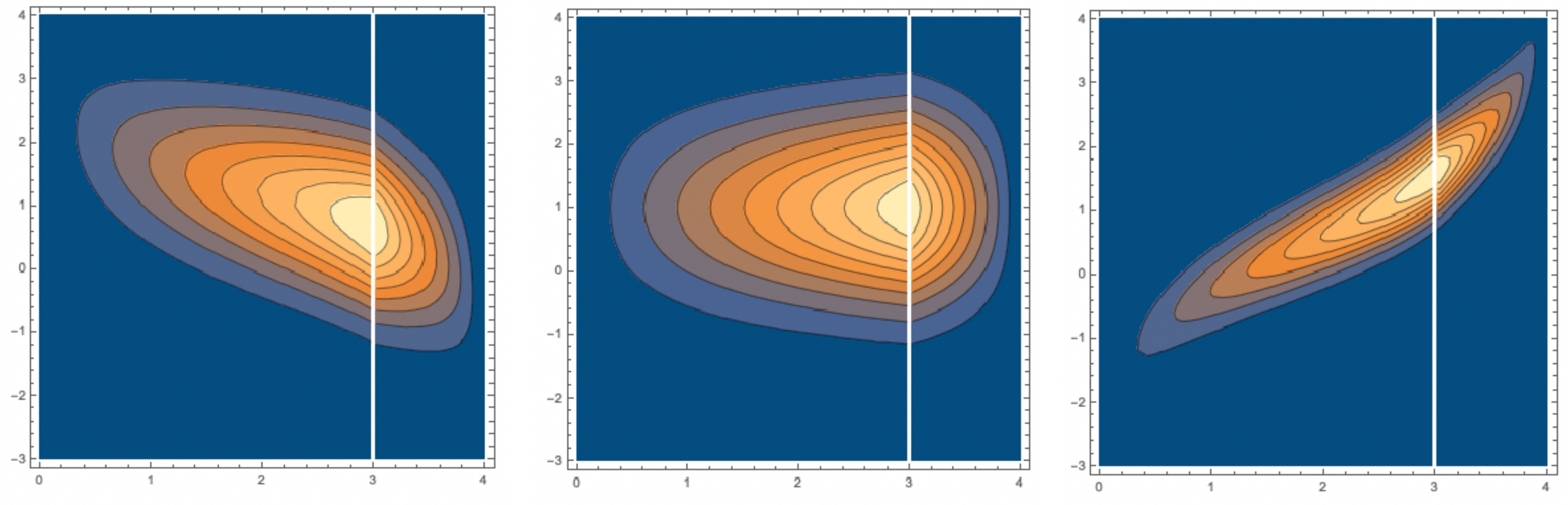
$$\Sigma_i(x) = \int_{-\infty}^{x} dx' P_i(x')$$

$$\Phi^{-1}(p) = \sqrt{2}\mathrm{Erf}^{-1}(-1+2p)$$

$$\xi(x) = \Phi^{-1}\left(\Sigma_1(x)\right)$$

$$\eta(y) = \Phi^{-1}\left(\Sigma_2(y)\right)$$


×

=




$c = -0.5$ $c = 0$ $c = 0.9$

use inference to constrain c

Conclusions & Outlook Part 3

- Bayesian inference is a powerful framework to estimate Δ_{MHO}
  - statistical interpretation ⇝ $P(\delta_{n+1} \,|\, \boldsymbol{\delta}_n)$
  - exposes our *assumptions & biases* clearly ⇝ model & priors
  - *but:* it is not more reliable than scale variation ⇝ careful analysis required
- typically for $n < 2$: $\mathrm{CI}_{68} > 9\mathrm{pt}$; $n \geq 2$: $\mathrm{CI}_{68} \simeq 9\mathrm{pt}$
- *public code:* ミホ (miho) ⇝ `https://github.com/aykhuss/miho`
- future directions
  - correlations (PDF fits & data interpretation)
  - marginalisation over models, …

Conclusions & Outlook Part 3

- Bayesian inference is a powerful framework to estimate Δ_{MHO}
  - statistical interpretation ⇝ $P(\delta_{n+1} \,|\, \boldsymbol{\delta}_n)$
  - exposes our *assumptions & biases* clearly ⇝ model & priors
  - *but:* it is not more reliable than scale variation ⇝ careful analysis required
- typically for $n < 2$: $\text{CI}_{68} > 9\text{pt}$; $n \geq 2$: $\text{CI}_{68} \simeq 9\text{pt}$
- *public code:* ミホ (miho) ⇝ `https://github.com/aykhuss/miho`
- future directions
  - correlations (PDF fits & data interpretation)
  - marginalisation over models, …

Thank you!