Higgs boson production at the Large Hadron Collider: accurate theoretical predictions at higher orders in QCD

Jesús Urtasun Elizari

PhD thesis defense - Milan, February 25th, 2022

UNIVERSITÀ
DEGLI STUDI
DI MILANO

N3PDF
Machine Learning • PDFs • QCD

European
Research
Council

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 740006.

Outline

1. QCD and collider physics
    - The strong interactions
    - Asymptotic freedom and pQCD
    - Factorization in QCD
    - Phenomenology at the LHC
2. All order perturbative resummation
    - Higher order radiative corrections
    - Resummation of large logarithmic corrections
    - Resummed component, asymptotic and fixed-order
3. HTurbo numerical implementation
    - Higgs production at the LHC
    - HTurbo numerical implementation
    - N^3LL implementation
4. Results & Conclusions

Part I
QCD and collider physics

- The Standard Model describes fundamental interactions at the TeV scale
- Particles as local excitations of fields with quantum mechanical behavior
- Lagrangian describing the fundamental objects ot the theory

$$\mathcal{L} = \bar{\psi}_q^i (i\gamma^\mu)(D_\mu)_{ij}\psi_q^j - m_q \bar{\psi}_q^i \psi_{qi} - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu}$$

QCD is the theory of the strong interactions $\longrightarrow$ interactions between quarks and gluons

How to explore proton's inner structure?

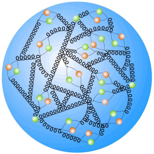
?

- At different scales, hadrons show different behavior
- From point-like objects to complex internal dynamics
- Scattering experiments (DIS) and hadronic physics (LHC)

"A way of describing high energy collisions is to consider any hadron as a composite object of point-like constituents $\longrightarrow$ **partons"** R.Feynman, 1969

QCD and collider physics

Asymptotic freedom and pQCD



- Parton model as LO approximation to QCD
- QCD coupling strength α_s changes with energy
- At high energies the hadron involves extremely complex internal dynamics

QCD is strongly coupled at large distances / low energies $\longrightarrow$ confinement

Non-perturbative physics

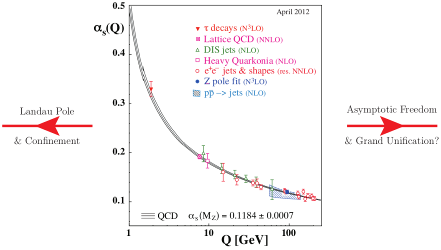
- Running coupling given by Renormalization Group Equation (RGE)

$$\mu \frac{d\alpha_s(\mu)}{d\mu} = \beta(\alpha_s(\mu)) = -\sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{\pi}\right)^{n+1}$$

- Coupling α_s evolves with scale μ as given by RGE $\rightarrow$ LO behavior driven by β_0
- Main difference between QED and QCD
  - $\beta_0^{\mathrm{QED}} < 0 \implies$ strongly coupled at large energies
  - $\beta_0^{\mathrm{QCD}} > 0 \implies$ weakly coupled at large energies

QCD and collider physics

Asymptotic freedom and pQCD



- Running coupling given by Renormalization Group Equation (RGE)

$$\alpha_s(\mu) = \frac{1}{\beta_0 \log\left(\frac{\mu^2}{\Lambda^2_{\mathrm{QCD}}}\right)}$$

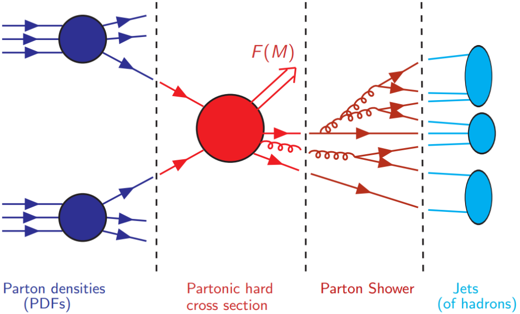
- β_0 LO of the β function, is > 0
- Λ_{QCD}, parameter that defines value of the coupling at large scales

QCD is weakly coupled for $\mu >> \Lambda_{\mathrm{QCD}} \longrightarrow$ asymptotically free

Perturbative Quantum Chromodynamics (pQCD)

QCD and collider physics

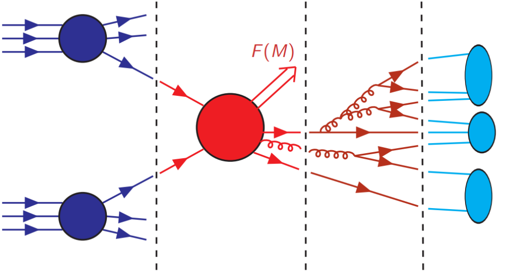
Hadronic processes and factorization



- LHC physics rely on hadronic collisions $\longrightarrow$ pQCD
- Compute **cross section** σ^F $\longrightarrow$ probability for a given process

QCD and collider physics

Hadronic processes and factorization

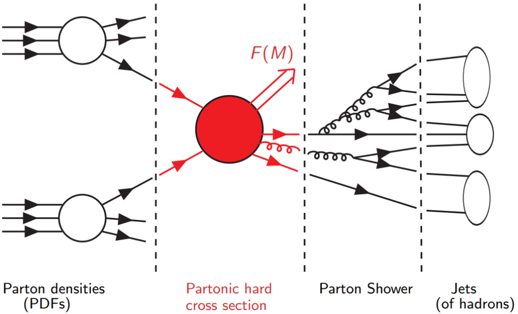


Compute hadronic cross sections is a hard problem $\longrightarrow$ QCD Factorization

$$\sigma^{\mathrm{F}}(p_1,p_2) = \int_0^1 dx_1 dx_2\; f_\alpha(x_1,\mu_F^2) * f_\beta(x_2,\mu_F^2) \;*\; \hat{\sigma}^{\mathrm{F}}_{\alpha\beta}(x_1 p_1, x_2 p_2, \alpha_s(\mu_R^2), \mu_F^2)$$

QCD and collider physics

Hadronic processes and factorization

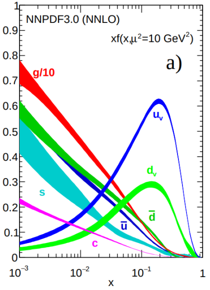


- Parton densities (PDFs) $f_\alpha(x_i, \mu_F^2)$: non perturbative but universal
- Partonic cross section $\hat{\sigma}^{\mathrm{F}}_{\alpha\beta}$: process dependent but computable as perturbative series in α_s

QCD and collider physics

Parton densities

Parton Distribution Functions: probability distribution of finding a particular parton (u, d, ..., g) carrying a fraction x of the proton's momentum



- Each parton has a different PDF $\longrightarrow u(x), d(x), ..., g(x)$
- PDFs can not predicted and yet can not measured $\longrightarrow$ extracted from data (MSTW, CTEQ, NNPDF collaborations)
- The N3PDF project: Machine Learning for PDFs determination

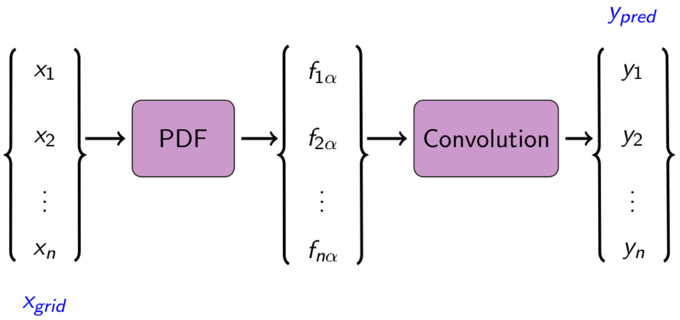
QCD and collider physics

The N3PDF project

- Use `TensorFlow` and `Keras` to determine the PDFs with ML fitting models
- See paper by S.Carrazza - J.Cruz-Martinez
  "Towards a new generation of parton densities with deep learning models", Carrazza et al., https://arxiv.org/abs/1907.05075
- TensorFlow operator implementation $\longrightarrow$ optimize PDF fitting
  "Towards hardware acceleration for parton densities estimation", Urtasun-Elizari et al., https://arxiv.org/abs/1909.10547

QCD and collider physics

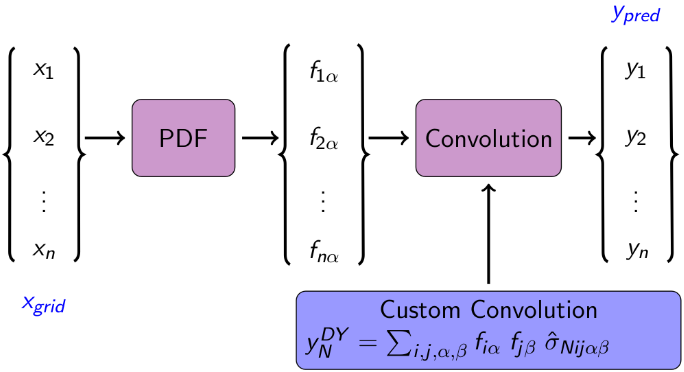
General structure of n3fit



1. Build a NN to compute y_{pred} observables from a grid of momentum fractions x_i
2. Compute χ^2 loss function by comparing with LHC data
3. Update values of PDF through χ^2 minimization $\longrightarrow$ Fit

QCD and collider physics

Operator implementation in TF

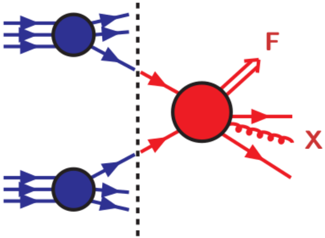


1. TF relies in symbolic computation $\longrightarrow$ High memory usage
2. Implement `C++` operator replacing the convolution
3. [Urtasun-Elizari et al.] ref. at 1910.07049

QCD and collider physics

Partonic cross section and pQCD

- Born cross section is the leading-order (LO) term of the perturbative series
- $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}$ are the NLO, NNLO, N^3LO corrections



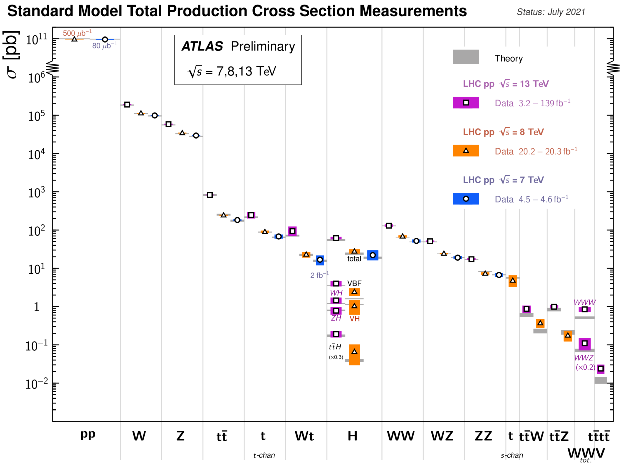
$$\hat{\sigma} = \sigma^{\text{Born}}\left(1 + \alpha_s \sigma^{(1)} + \alpha_s^2 \sigma^{(2)} + \alpha_s^3 \sigma^{(3)} + \ldots\right)$$

Lower order predictions strongly depend on the auxiliary / unphysical scales

Need higher order corrections to increase theoretical accuracy!

QCD and collider physics

LHC phenomenology



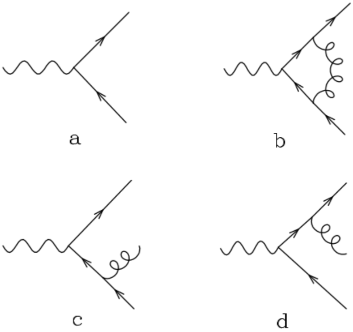
Part II

All order resummation

Resummation in QCD

Higher order corrections - need for resummation

1. Calculation of higher order corrections is not an easy task due to infrared (IR) soft and collinear singularities
2. Final state singularities cancel by combining real and virtual contributions $\longrightarrow$ KLN theorem
3. Initial state collinear singularities factorized inside the PDFs

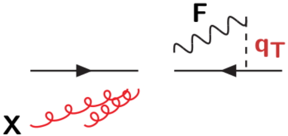


Complete cancellation only works in inclusive final states!

Resumation in QCD

$q_\perp$ resummation

- Describing exclusive final states
- Study the differential $q_\perp$ distribution
  $h_1(p_1) + h_2(p_2) \longrightarrow F(M, q_\perp) + X$



$$\int_0^{Q_\perp^2} dq_\perp^2 \frac{d\hat{\sigma}}{dq_\perp^2} \sim c_0 + \alpha_s(c_{12}L^2 + c_{11}L + c_{10}) + \ldots, \quad \text{where} \quad L = \ln(M^2/q_\perp^2)$$

$\alpha_S L^2$	$\alpha_S L$	$\cdots$	$\mathcal{O}(\alpha_S)$
$\alpha_S^2 L^4$	$\alpha_S^2 L^3$	$\cdots$	$\mathcal{O}(\alpha_S^2)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\alpha_S^n L^{2n}$	$\alpha_S^n L^{2n-1}$	$\cdots$	$\mathcal{O}(\alpha_S^n)$
dominant logs	$\cdots$	$\cdots$	$\cdots$

Truncated fixed-order predictions $\rightarrow$ enhanced $\alpha_s^n \ln^m(M^2/q_\perp^2)$ appear

- Catani - Bozzi - de Florian - Grazzini (CBFG) formalism (*)
  "Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC",
  Bozzi et al., https://arxiv.org/abs/hep-ph/0508068
- Separate partonic $q_\perp$ distribution as follows:

$$\frac{d\hat{\sigma}_{ab}}{dq_\perp^2} = \left[\frac{d\hat{\sigma}_{ab}^{(\text{res.})}}{dq_\perp^2}\right]_{\text{l.a.}} + \left[\frac{d\hat{\sigma}_{ab}^{(\text{fin.})}}{dq_\perp^2}\right]_{\text{f.o.}} \quad , \quad \text{such that}$$

$$\int_0^{q_\perp^2} dq_\perp^2 \, \frac{d\hat{\sigma}_{ab}^{(\text{res.})}}{dq_\perp^2} \sim \sum \alpha_s^n \log^m \left(\frac{M^2}{q_\perp^2}\right) \quad \text{for} \quad q_\perp \to 0$$

$$\lim_{q_\perp \to 0} \int_0^{q_\perp^2} dq_\perp^2 \, \frac{d\hat{\sigma}_{ab}^{(\text{fin.})}}{dq_\perp^2} = 0$$

Resummed and finite components can be matched (LL+LO, NLL+NLO, NNLO+NNLL, ...) to have uniform accuracy in a wide range of $q_\perp$

Resummation in QCD

Resummed component

Resummation holds in impact parameter space b

$$\frac{d\hat{\sigma}^{(\mathrm{res.})}_{ab}}{dq_\perp^2} = \frac{M^2}{\hat{s}} \int db\, \frac{b}{2}\, J_0(bq_\perp)\, \mathcal{W}_{ab}(b, M)$$

with $\mathcal{W}_{ab}$ also expressed in Mellin space (with respect to $z = M^2/\hat{s}$)

$$\mathcal{W}_N(b, M) = \mathcal{H}_N(\alpha_s) \times \exp\{\mathcal{G}_N(\alpha_s, L)\} \quad \text{being} \quad L \equiv \log(M^2 b^2)$$

- Large logarithms exponentiated in the universal Sudakov form factor $\mathcal{G}_N(\alpha_s, L)$
- Constant (b-independent) terms factorized in the process dependent hard factor $\mathcal{H}_N(\alpha_s)$

Resummation in QCD

Extend formalism to N^3LL

Sudakov factor $\mathcal{G}_N$ and hard coefficient $\mathcal{H}_N$ can be expanded as perturbative series in α_s

$$\mathcal{G}_N(\alpha_s, L) = L\, g^{(1)}(\alpha_s L) + g^{(2)}(\alpha_s L) + \frac{\alpha_s}{\pi} g^{(3)}(\alpha_s L) + \left(\frac{\alpha_s}{\pi}\right)^2 g^{(4)}(\alpha_s L) + \dots$$

$$\mathcal{H}_N(\alpha_s) = 1 + \alpha_s \mathcal{H}^{(1)} + \alpha_s^2 \mathcal{H}^{(2)} + \alpha_s^2 \mathcal{H}^{(3)} + \dots$$

For each new order implement a factor of $\mathcal{G}_N$ and Hard $\mathcal{H}_N$

$\mathrm{LL}(\sim \alpha_s^n L^{n+1}) : g^{(1)}, \hat{\sigma}^{(0)}$

$\mathrm{NLL}(\sim \alpha_s^n L^{n}) : g^{(2)}, \mathcal{H}^{(1)}$

$\mathrm{NNLL}(\sim \alpha_s^n L^{n-1}) : g^{(3)}, \mathcal{H}^{(2)}$

$\mathrm{N^3LL}(\sim \alpha_s^n L^{n-2}) : g^{(4)}, \mathcal{H}^{(3)}$

- Implement CBFG resummation in C++ code
- Extend the formalism up to N^3LO+N^3LL accuracy!

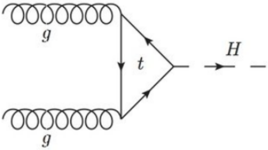
Part III
HTurbo numerical implementation

HTurbo

Resummation for Higgs differential distribution

- Fast and accurate predictions for Higgs boson production cross section
- Predictions for differential cross section $d\sigma^{\mathrm{H}}/dq_{\perp}^{2}$
- Numerical implementation of resummed and finite components

$$d\sigma^{\mathrm{H}}_{\mathrm{(N)NLL+(N)LO}} = d\sigma^{\mathrm{(res.)}}_{\mathrm{(N)NLL}} - d\sigma^{\mathrm{(asy.)}}_{\mathrm{(N)LO}} + d\sigma^{\mathrm{(f.o.)}}_{\mathrm{(N)LO}}$$
$$d\sigma^{\mathrm{(res.)}}_{\mathrm{(N)NLL}} = \hat{\sigma}^{\mathrm{H}}_{\mathrm{LO}} \times \mathcal{H}_{\mathrm{(N)LO}} \times \exp \mathcal{G}_{\mathrm{(N)NLL}}$$
$$d\sigma^{\mathrm{(asy.)}}_{\mathrm{(N)LO}} = \hat{\sigma}^{\mathrm{H}}_{\mathrm{LO}} \times \Sigma_{\mathrm{(N)LO}}$$

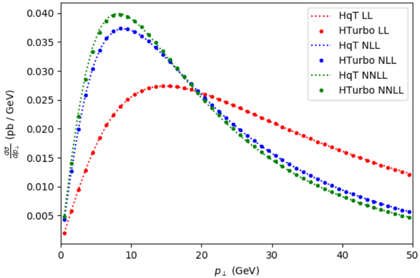


- LO process is just gg $\rightarrow$ H, but NLO and beyond require gg $\rightarrow$ H + jet!

HTurbo

Predictions for Higgs $q_\perp$ distribution

- $q_\perp$ resummation implemented in numerical codes **HqT**, **HRes**, **HNNLO** [Catani, de Florian, Ferrera, Grazzini, Tommasini]
- Higher order accuracy require high computation times
- NNLL predictions can take more than 48h just for 1 PDF set and 1 μ_R, μ_F value $\longrightarrow$ need for fast numerical implementations



Codes producing fast and accurate predictions are needed for precision era of the LHC (High Luminosity LHC, from 80 - 140 fb^{-1} to **2000 fb^{-1}** !)

HTurbo

Starting point: DYTurbo

Numerical code **DYTurbo** [Camarda et al., https://arxiv.org/abs/1910.07049], fast and precise $q_\perp$ resummation and several improvements for Drell-Yan ($h_1 h_2 \to V + X \to l^+ l^- + X$)

First goal: set up a numerical code for Higgs boson production starting from **DYTurbo**

- Set LO amplitude $gg \to H$
- Set Sudakov and Hard coefficients for resummed component
- Set Σ coefficients for asymptotic term
- Implement MC producing the LO and NLO H+jet cross sections
- Compare with **HRes** and **HqT**

Final goal: extend theoretical accuracy up to N^{3}LL+N^{3}LO

HTurbo

Code optimization

Optimized reimplementation of **HqT**, **HRes** and **HNNLO** for q_T-resummation

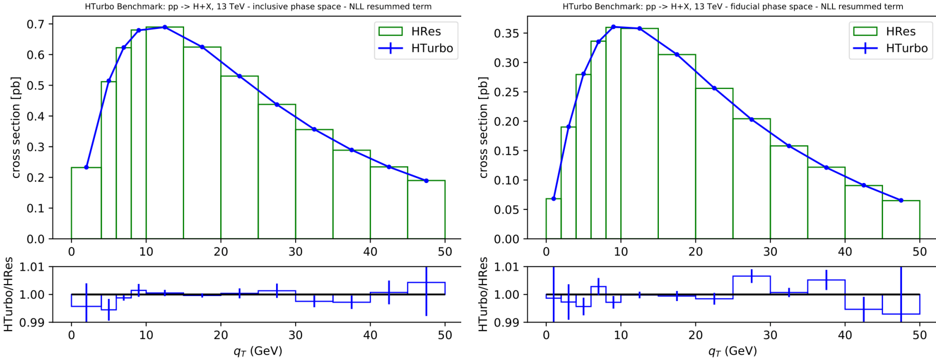
- **C++** structure with **Fortran** interfaces → Multi-threading
- Optimization in the integration routines / integral transforms
  - Factorize boson and decay kinematics
  - Gauss-Legendre quadrature rules (1-dim.)
  - Vegas/Cuhre through **Cuba** (multi-dim.)

Benchmark comparison with **HRes, HNNLO** - numerical and speed performance

- "Higgs boson production at the LHC: fast and precise predictions in QCD at higher orders", Urtasun-Elizari et al., https://arxiv.org/abs/2202.10343

Results

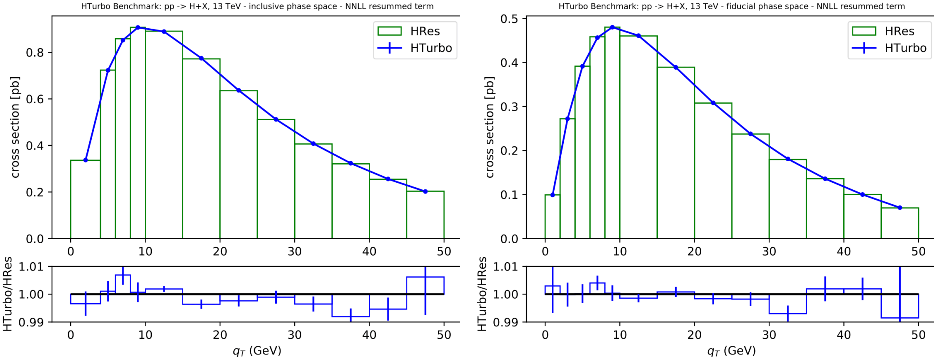
Comparison HTurbo and HRes - NLL resummed



- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- CM energy $\sqrt{s} = 13$ GeV and PDF set `NNPDF31_nlo_as_0118` PDF set

Results

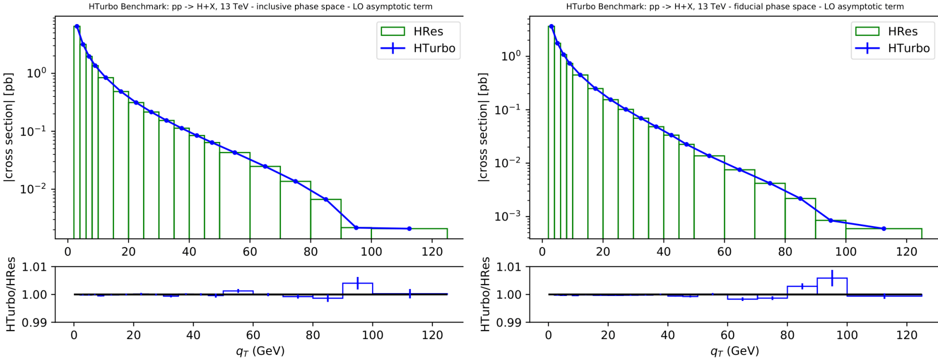
Comparison HTurbo and HRes - NNLL resummed



- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- CM energy $\sqrt{s} = 13$ GeV and PDF set `NNPDF31_nnlo_as_0118` PDF set

Results

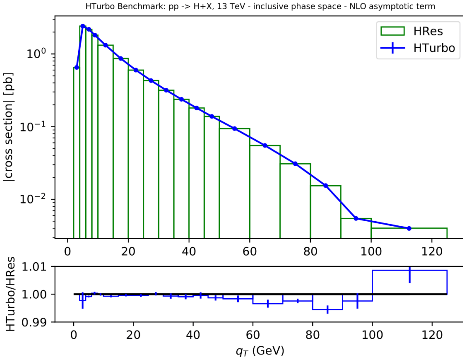
Comparison HTurbo and HRes - LO asymptotic

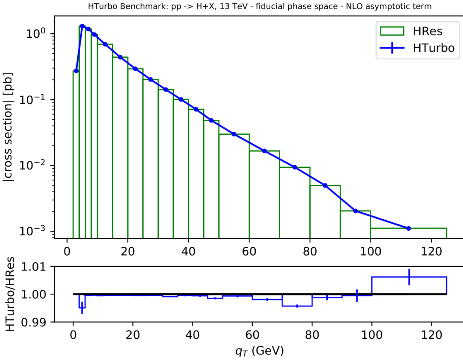


- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- CM energy $\sqrt{s} = 13$ GeV and PDF set `NNPDF31_nlo_as_0118` PDF set

Results

Comparison HTurbo and HRes - NLO asymptotic

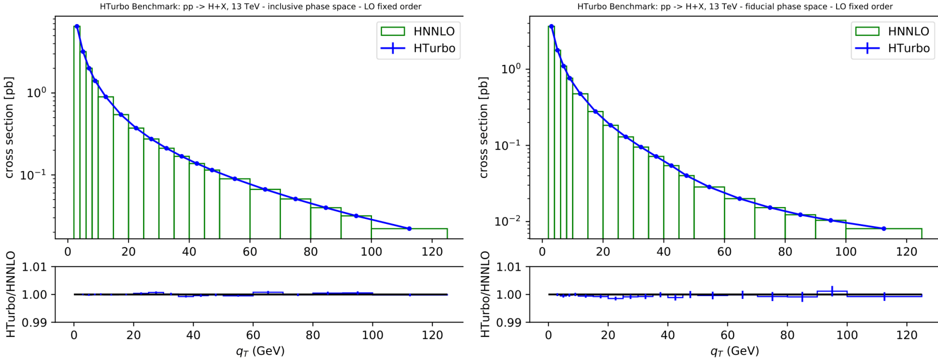




- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- CM energy $\sqrt{s} = 13$ GeV and PDF set `NNPDF31_nnlo_as_0118` PDF set

Results

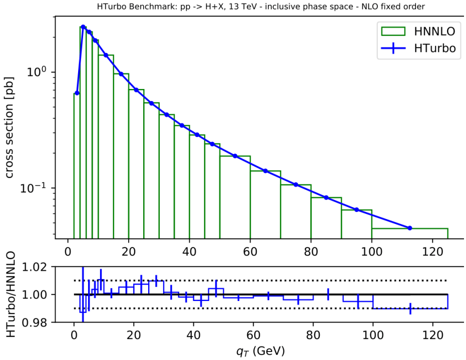
Comparison HTurbo and HRes - LO fixed-order

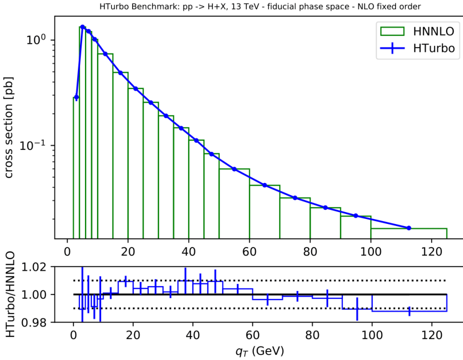


- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- CM energy $\sqrt{s} = 13$ GeV and PDF set `NNPDF31_nlo_as_0118` PDF set

Results

Comparison HTurbo and HRes - NLO fixed-order





- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- CM energy $\sqrt{s} = 13$ GeV and PDF set `NNPDF31_nnlo_as_0118` PDF set

Test time performance in machine with 3.50 GHz Intel Xeon CPUs:

- HRes NLL resummed $\rightarrow$ 0.5h with with 1% uncertainty
- HTurbo NLL resummed (without multi-threading) $\rightarrow$ 20s with 0.001% uncertainty
- HRes NNLL resummed $\rightarrow$ 48h with 1% uncertainty
- HTurbo NNLL resummed (without multi-threading) $\rightarrow$ 5' with 0.01% uncertainty
- Improvement of two orders of magnitude in the time performance
- Improvement of three orders of magnitude in numerical precision

Results

N^{3}LL implementation

Sudakov factor $\mathcal{G}_N$ and hard coefficient $\mathcal{H}_N$ can be expanded as perturbative series in α_s

$$\mathcal{G}_N(\alpha_s, L) = L\, g^{(1)}(\alpha_s L) + g^{(2)}(\alpha_s L) + \frac{\alpha_s}{\pi} g^{(3)}(\alpha_s L) + \left(\frac{\alpha_s}{\pi}\right)^2 g^{(4)}(\alpha_s L) + \ldots$$

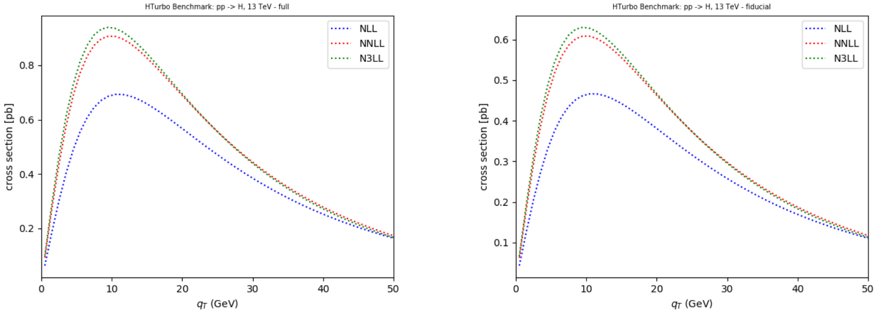
$$\mathcal{H}_N(\alpha_s) = 1 + \alpha_s \mathcal{H}^{(1)} + \alpha_s^2 \mathcal{H}^{(2)} + \alpha_s^2 \mathcal{H}^{(3)} + \ldots$$

For each new order implement a new factor of $\mathcal{G}_N$ and Hard $\mathcal{H}_N$

- Extend the formalism up to N^{3}LO+N^{3}LL accuracy!
- Implementation of N^{3}LL factors following
  - "Anomalous dimension for transverse-momentum resummation", Li - Zhu, https://arxiv.org/abs/1604.01404,
  - "Cusp and collinear anomalous dimensions in four-loop QCD", Von Manteuffel et al., https://arxiv.org/abs/2002.04617

Results

N^{3}LL implementation



- Cross section for fully inclusive (LHS) and fiducial (RHS) phase space ✓
- Implementation of N^{3}LL factors following [Li - Zhu, 1604.01404], [Von Manteuffel et al., 2002.04617]
- First implementation of resummed Higgs cross section at N^{3}LL accuracy!

Summary & Conclusions

1. Accurate predictions are needed towards the precision era of the LHC
2. Resummation is needed for describing differential distributions
3. Fast numerical implementation are needed towards the precision era of the LHC
4. Developing a novel numerical code, **HTurbo**, which implements $q_\perp$ resummation for Higgs boson production
5. HTurbo is faster than any of the existing codes
6. HTurbo contains the first implementation of resummation at N^{3}LL accuracy!
7. Next steps:
   - Add full N^{3}LO+N^{3}LL prediction
   - Perform phenomenological studies comparing with LHC data

Thank you!

This project has received funding from the European Union′s Horizon 2020 research and innovation program under grant agreement No 740006.

Back up

Back up

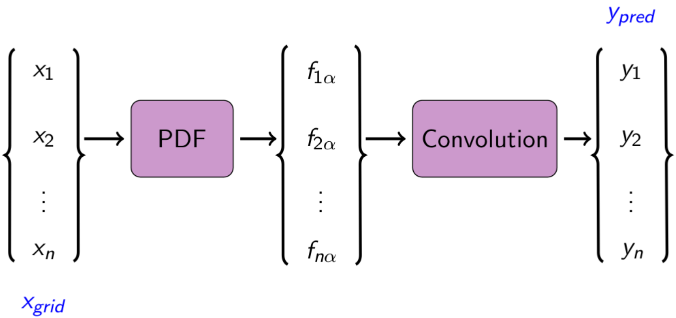
Back up

General structure of n3fit

- Use TensorFlow and Keras to determine the PDFs
- See paper by S.Carrazza - J.Cruz-Martinez
  "Towards a new generation of parton densities with deep learning models",
  https://arxiv.org/abs/1907.05075

Back up

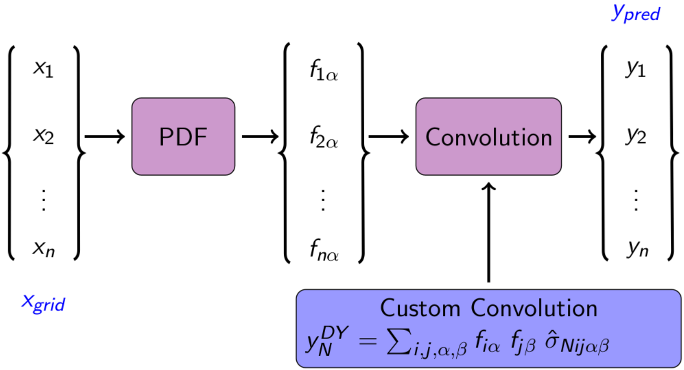
General structure of n3fit



1. Build a NN to compute y_{pred} observables from a grid of momentum fractions x_i
2. Compute loss function by comparing with LHC data
3. Update values of PDF $\longrightarrow$ Fit

Back up

Operator implementation



1. TF relies in symbolic computation $\longrightarrow$ High memory usage
2. Implement C++ operator replacing the convolution

DIS only:

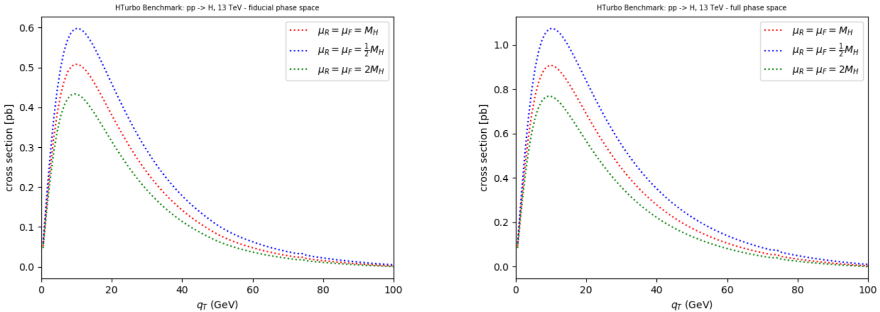
	TensorFlow	Custom	Ratio
Convolution	1.9207904	1.9207904	1.0000000
	2.4611666	2.4611664	0.9999999
	1.3516952	1.3516952	1.0000000
Gradient	1.8794115	1.8794115	1.0000000
	1.505316	1.505316	1.0000000
	2.866085	2.866085	1.0000000

DY-like only:

	TensorFlow	Custom	Ratio
	8.142365	8.142366	1.0000001
Convolution	8.947762	8.947762	1.0000000
	7.4513326	7.4513316	0.9999999
	18.525095	18.525095	1.0000000
Gradient	19.182995	19.182993	0.9999999
	19.551006	19.551004	0.9999999

Results

N^3LL implementation



Estimate theory uncertainty:

- i) by performing scale variations on μ_R, μ_F
- ii) by comparing the last two contributions in the perturbative expansion