fastNLO

Thomas Kluge, Klaus Rabbertz, Markus Wobisch
DESY University Karlsruhe Fermilab

FRONTIERS IN CONTEMPORARY PHYSICS
Vanderbilt University Nashville, Tennessee 23-28, 2005

- Motivation
- Concept
- The Product
- Status / Outlook

fastNLO

Motivation

- Computations of higher-order pQCD predictions for hadronic-final state observables are time-consuming
- Often need repeated computations of the same cross section for different PDFs and/or $\alpha_s(M_z)$ values
- Examples for a specific analysis:
  - use various PDFs (CTEQ, MRST, Alekhin, Botje, H1, ZEUS, ...)
  - determine PDF uncertainties (PDF error sets)
  - use data set in fit of PDFs and/or α_s
- For some observables NLO predictions can be computed extremely fast (e.g.: DIS structure functions)
- ... but some are extremely slow: Drell-Yan and Jet Cross Sections

$\Longrightarrow$ need new procedure for very fast repeated computations of NLO cross sections

New Concept

- Can be used for **any** observable in hadron-induced processes (hadron-hadron / DIS / photoproduction)
- Although labeled "fastNLO" $\rightarrow$ can be used in any order $\Rightarrow$ fastN^{n}LO
- Our concept does not include the theoretical calculation itself (leave this to theorists) $\rightarrow$ it requires existing flexible computer code — here: NLOJET++ (Zoltan Nagy)
- During the first computation no time is saved
  need full time of the original code: hours, days, weeks, months, ...
  to achieve high statistical precision
- This concept involves one single approximation (see later)
  But: precision of approximation can be quantified & arbitrarily improved
- Any further computation takes **one second** (independent of statistical precision)

$\Rightarrow$ **here:** example for inclusive jet production in hadron-hadron collisions

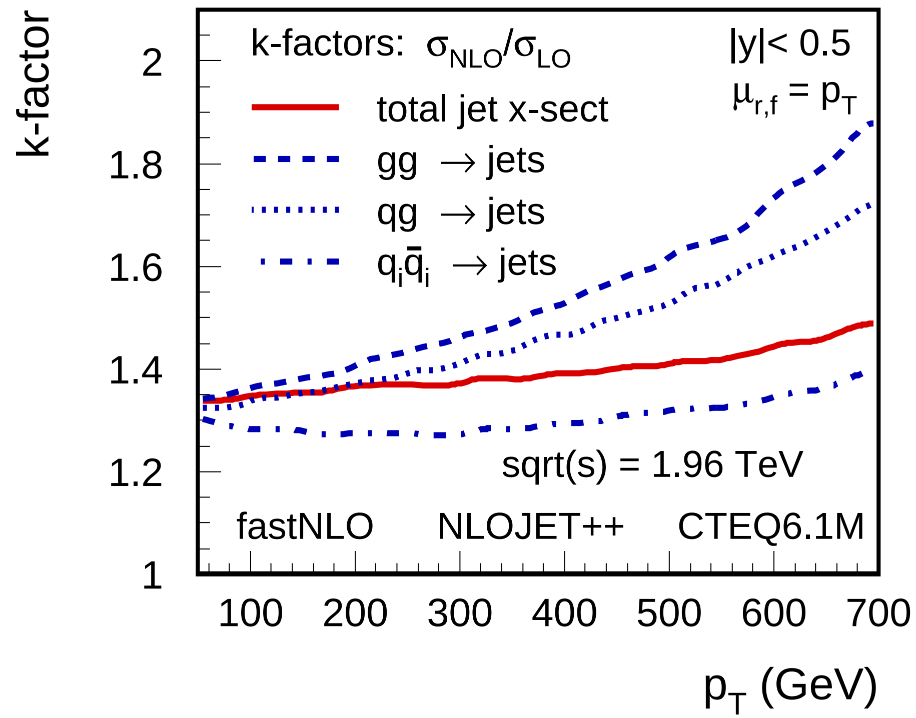
Current CTEQ Procedure

k-factor approximation:

- for a given PDF $\longrightarrow$ compute k-factor for each bin: k = sigma(NLO)/sigma(LO)
- “relatively fast”: compute LO cross section for arbitrary PDF
- multiply sigma(LO) with k-factor $\rightarrow$ get “NLO” prediction

problem:

- k-factor itself depends on the PDFs $\longrightarrow\longrightarrow\longrightarrow$
- higher for gluon induced subprocesses



reason:

- different x-coverage in LO and NLO
- different k-factors for different subprocesses

limitations:

- even the LO computation is slow
- computing time depends on statistical precision

fastNLO

- as exact as you like
- much, much faster

Jet Cross Section in hadron-hadron

General cross section formula for hadron-hadron collisions:

$$\sigma_{\rm hh} \;=\; \sum_n \alpha_s^n(\mu_r) \sum_{\text{PDFflavors } i} \;\sum_{\text{PDFflavors } j} c_{i,j,n}(\mu_r,\mu_f) \otimes f_i(x_1,\mu_f) \otimes f_j(x_2,\mu_f)\,.$$

- strong coupling constant α_s in order n
- perturbative coefficient $c_{i,j,n}$
- parton density functions (PDFs) of the hadrons $f_i(x),\; f_j(x)$
- renormalization scale μ_r, factorization scale μ_f, (ignore in the following $\Rightarrow \mu_{r,f} = p_T$)
- momentum fraction x

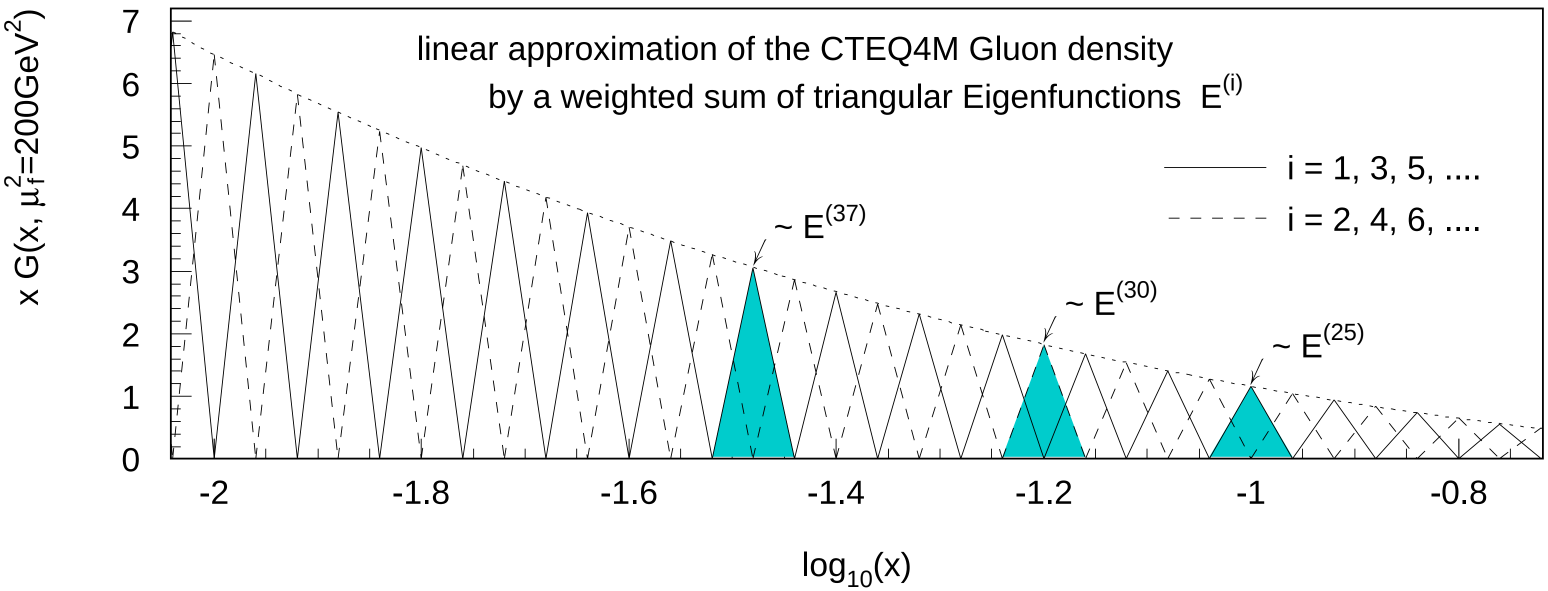
Standard procedure:

- integration over whole phase space (x_1, x_2) (usually Monte-Carlo method)
- at each MC integration point:
  - computation of observable (e.g. run jet algorithm, determine p_T, $|y|$ bin)
  - compute perturbative coefficient
  - get α_s and PDFs values

  $\Rightarrow$ add contribution to bin

goal: try to separate the PDFs from the integral

PDF Approximation



- introduce a set of discrete x-values labeled $x^{(i)}$ $(i = 0, 1, 2, \cdots, n)$
- with $x^{(n)} < x^{(n-1)} < x^{(n-2)} < \cdots < x^{(0)} = 1$
- around each $x^{(i)}$, define an eigenfunction $E^{(i)}(x)$
- with $E^{(i)}(x^{(i)}) = 1$, $E^{(i)}(x^{(j)}) = 0$ for $i \neq j$ and $\sum_i E^{(i)}(x) = 1$ for all x

- express a single PDF $f(x)$ by a linear combination of eigenfunctions $E^{(i)}(x)$ with coefficients given by the PDF values $f(x^{(i)})$ at the discrete points $x^{(i)}$

$$f(x) = \sum_i f(x^{(i)})\, E^{(i)}(x)$$

PDF Approximation (2)

processes with two hadrons – need Eigenfunctions in 2d-space (x_1, x_2)

- define $E^{(i,j)}(x_1, x_2) \equiv E^{(i)}(x_1) E^{(j)}(x_2)$
- product of two PDFs $f(x_1, x_2) \equiv f_1(x_1)\, f_2(x_2)$ is given by

$$f(x_1, x_2) = \sum_{i,j} f(x_1^{(i)}, x_2^{(j)})\; E^{(i,j)}(x_1, x_2)$$

note: this is an **approximation**!!
choice of triangular Eigenfunctions $\Longrightarrow$ linear interpolation of PDFs between adjacent $x^{(i)}$
this is the **only** approximation in fastNLO — precision can be arbitrarily improved!!
precision depends on:

- choice of set of $x^{(i)}$ — e.g. on $\log_{10}(1/x)$ or $\sqrt{\log_{10}(1/x)}$ (needs clever choice)
- number of x-bins (brute force) $\longrightarrow$ memory $\propto n^2$

$\Rightarrow$ **goal:** precision of 0.3% for all bins

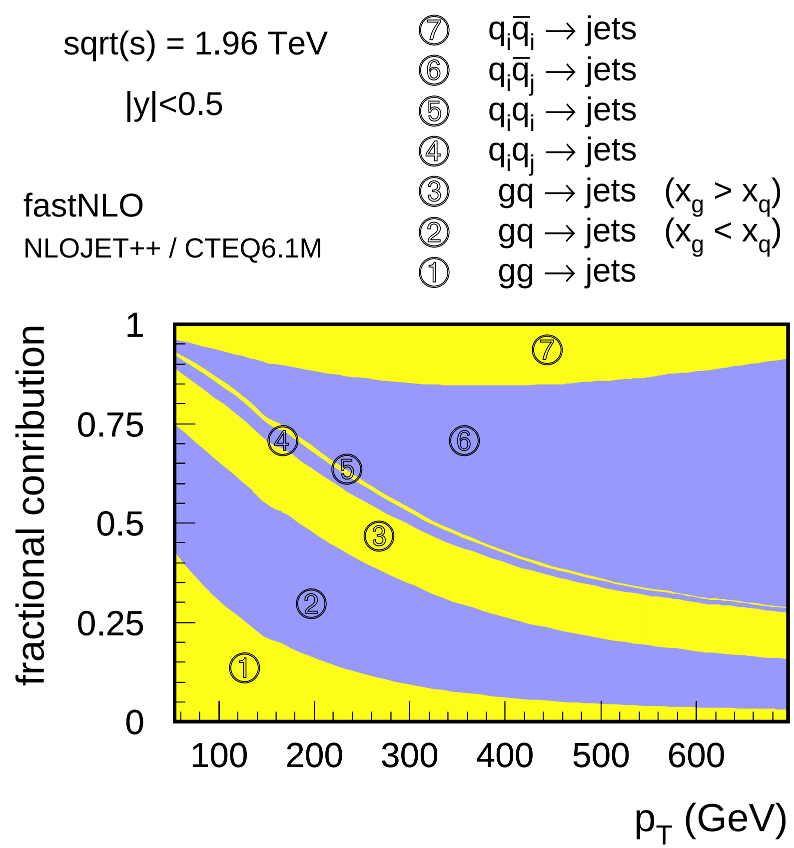
Partonic Subprocesses in h-h

now: don't want to deal with 13×13 PDFs!!

For hadron-hadron $\longrightarrow$ jets there are **seven** relevant partonic subprocesses:

$$
\begin{array}{rlcrlcl}
gg & \longrightarrow \text{jets} & & & & \propto & H_1(x_1,x_2) \\
gg & \longrightarrow \text{jets} & \text{plus} & \bar{q}g & \longrightarrow \text{jets} & \propto & H_2(x_1,x_2) \\
gq & \longrightarrow \text{jets} & \text{plus} & g\bar{q} & \longrightarrow \text{jets} & \propto & H_3(x_1,x_2) \\
q_iq_j & \longrightarrow \text{jets} & \text{plus} & \bar{q}_i\bar{q}_j & \longrightarrow \text{jets} & \propto & H_4(x_1,x_2) \\
q_iq_i & \longrightarrow \text{jets} & \text{plus} & \bar{q}_i\bar{q}_i & \longrightarrow \text{jets} & \propto & H_5(x_1,x_2) \\
q_i\bar{q}_i & \longrightarrow \text{jets} & \text{plus} & \bar{q}_iq_i & \longrightarrow \text{jets} & \propto & H_6(x_1,x_2) \\
q_i\bar{q}_j & \longrightarrow \text{jets} & \text{plus} & \bar{q}_iq_j & \longrightarrow \text{jets} & \propto & H_7(x_1,x_2)
\end{array}
$$

partonic subprocesses for $p\bar{p} \rightarrow$ jets



The H_i are linear combinations of PDFs
$\longrightarrow$ reduced from 13×13 to seven!!

detail:
for hadron - anti-hadron collisions:
PDFs of the anti-hadron are expressed by the PDFs of the hadron (quarks $\leftrightarrow$ anti-quarks)
here: swap $H_4 \leftrightarrow H_7$ and $H_5 \leftrightarrow H_6$

PDFs — Definitions

$$
\begin{aligned}
G(x,\mu_f) &= g(x,\mu_f) \\
Q(x,\mu_f) &= \sum_i q_i(x,\mu_f) \\
\bar{Q}(x,\mu_f) &= \sum_i \bar{q}_i(x,\mu_f) \\
S(x_1,x_2,\mu_f) &= \sum_i \left(q_i(x_1,\mu_f)\, q_i(x_2,\mu_f) + \bar{q}_i(x_1,\mu_f)\, \bar{q}_i(x_2,\mu_f)\right) \\
A(x_1,x_2,\mu_f) &= \sum_i \left(q_i(x_1,\mu_f)\, \bar{q}_i(x_2,\mu_f) + \bar{q}_i(x_1,\mu_f)\, q_i(x_2,\mu_f)\right)
\end{aligned}
$$

- $q_i(x)$ ($\bar{q}_i(x)$) — quark (anti-quark) density of flavor i
- $i = 1, \ldots, n_f$ — No. of flavors
- $G(x)$ — gluon density

Relevant Combinations of PDFs

$$
\begin{aligned}
H_1(x_1,x_2) &= G(x_1)\,G(x_2)\,, \\
H_2(x_1,x_2) &= \left(Q(x_1)+\bar{Q}(x_1)\right)\; G(x_2)\,, \\
H_3(x_1,x_2) &= G(x_1)\; \left(Q(x_2)+\bar{Q}(x_2)\right)\;, \\
H_4(x_1,x_2) &= Q(x_1)Q(x_2)+\bar{Q}(x_1)\bar{Q}(x_2)-S(x_1,x_2)\,, \\
H_5(x_1,x_2) &= S(x_1,x_2)\,, \\
H_6(x_1,x_2) &= A(x_1,x_2)\,, \\
H_7(x_1,x_2) &= Q(x_1)\bar{Q}(x_2)+\bar{Q}(x_1)Q(x_2)-A(x_1,x_2)\,.
\end{aligned}
$$

These are the seven combinations of PDFs, corresponding to the seven subprocesses

symmetries:
$H_n(x_1,x_2) = H_n(x_2,x_1)$ for $n = 1,4,5,6,7$ and $H_2(x_1,x_2) = H_3(x_2,x_1)$

$$H_k(x_1,x_2) = \sum_{(i,j)} H_k(x^{(i)},x^{(j)})\, E^{(i,j)}(x_1,x_2)$$

where $H_k(x^{(i)},x^{(j)})$ is a number $\leftrightarrow$ PDF information

Jet Cross Section in hadron-hadron

With these definitions of the seven $\boldsymbol{H_i}$ the cross section reads:

$$\boldsymbol{\sigma}_{\mathrm{hh}} = \sum_{n} \boldsymbol{\alpha}_s^n(\boldsymbol{\mu}_r) \sum_{k=1}^{7} \boldsymbol{c}_{k,n}(\boldsymbol{\mu}_r, \boldsymbol{\mu}_f) \otimes \boldsymbol{H}_k(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{\mu}_f)$$

Now: express $\boldsymbol{H_k}$ by linear combinations of the $\boldsymbol{E}^{(i,j)}(\boldsymbol{x}_1, \boldsymbol{x}_2)$

$$\boldsymbol{\sigma}_{\mathrm{hh}} = \sum_{n} \boldsymbol{\alpha}_s^n(\boldsymbol{\mu}_r) \sum_{k=1}^{7} \boldsymbol{c}_{k,n}(\boldsymbol{\mu}_r, \boldsymbol{\mu}_f) \otimes \left(\sum_{i,j} \boldsymbol{H}_k(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)}) \cdot \boldsymbol{E}^{(i,j)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \right)$$

or, better:

$$\boldsymbol{\sigma}_{\mathrm{hh}} = \sum_{n} \boldsymbol{\alpha}_s^n(\boldsymbol{\mu}_r) \sum_{k=1}^{7} \sum_{i,j} \boldsymbol{H}_k(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)}) \left(\boldsymbol{c}_{k,n}(\boldsymbol{\mu}_r, \boldsymbol{\mu}_f) \otimes \boldsymbol{E}^{(i,j)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \right)$$

important: integral is independent of PDFs!
the numbers $\boldsymbol{H}_k(\boldsymbol{x}^{(i)}, \boldsymbol{x}^{(j)})$ contain all information on the PDFs

$\Rightarrow$ **exactly what we wanted!!**

Last Step ...

define:

$$\tilde{\boldsymbol{\sigma}}_{k,n}^{(i,j)} \equiv \boldsymbol{c}_{k,n}(\boldsymbol{\mu}_r, \boldsymbol{\mu}_f) \otimes \boldsymbol{E}^{(i,j)}(\boldsymbol{x}_1, \boldsymbol{x}_2)$$

$\Rightarrow$ the $\tilde{\boldsymbol{\sigma}}_{k,n}^{(i,j)}$ contain all information on the observable

(the perturbative coefficients, the jet definition, and the phase space restrictions).

but: $\tilde{\boldsymbol{\sigma}}_{k,n}^{(i,j)}$ is independent of the PDFs and $\boldsymbol{\alpha}_s$ – needs to be computed only once!

The cross section is then given by the simple product **($\longrightarrow$ Master Formula!)**

$$\boldsymbol{\sigma}_{\mathrm{hh}} = \sum_{i,j,k,n} \boldsymbol{\alpha}_s^n(\boldsymbol{\mu}_r)\ \boldsymbol{H}_k(\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(j)})\ \tilde{\boldsymbol{\sigma}}_{k,n}^{(i,j)}$$

can be reevaluated **very** quickly for different PDFs and $\boldsymbol{\alpha}_s$ values,

as e.g. required in the determination of PDF uncertainties or in global fits of PDFs

Implementation Steps

to implement a new observable in fastNLO:

- find theorist to provide flexible computer code
- identify elementary subprocesses & relevant PDF linear combinations
- define analysis bins (e.g. p_T, $|y|$)
- define Eigenfunctions $E(x), E(x_1, x_2)$ (e.g. triangular) & the set of $x^{(i)}$
- to optimize x-range: find lower x-limit ($x_{\text{limit}} < x < 1$) (for each analysis bin)

<u>example:</u> DØRun I measurement of Incl. Jet Cross Section, Phys. Rev. Lett.86, 1707 (2001)

- 90 analysis bins in (E_T, η)
- 2 orders of $\alpha_s(p_T)$ (LO & NLO)
- 7 partonic subprocesses
- No. of x-intervals for each bin: 50 (100?) $\longleftarrow$ (study precision of PDF approximation)
  $\Rightarrow (n^2 + n)/2 =$ 1275 (5050?) Eigenfunctions $E^{(i,j)}(x_1, x_2)$

- compute 1.6M (6.4M?) variables $\tilde{\sigma}_{k,n}^{(i,j)}$ (times three, if scale variations are included)
  $\Rightarrow$ stored in huge table!!!

compute VERY long to achieve very high precision — (after all: needs to be done only once!)

The Product

Everything will be downloadable from the fastNLO Webpage

Package for a single observable includes:

- Tables of $\tilde{\sigma}_{k,n}^{(i,j)}$ in different orders - for different scales
- Stand-Alone Code to:
  - read tables
  - loop over PDFs (LHAPDFlib interface - or custom user interface for global fitters)
  - output cross section numbers as: array, ASCII, ROOT/HBOOK histograms
- Examples

Code computes NLO Predictions for a whole set of data points in the order of seconds (depends on speed of PDF interface)

Can easily be included into user-specific analysis framework fastNLO

Summary / Outlook

Status:

- concept for fastNLO is fully developed
- implementation of code for hadron-hadron jet cross section finished
- currently: studying precision / x-binning / “tweaking”

Outlook:
(start with inclusive jet production)

- first: provide tables and user code for published Run I results from CDF and DØ at 630 GeV and 1800 GeV — in analysis specific bins
(⟶ data can easily be included in all PDF fits)
- next: provide tables and user code for Run II and LHC energies – flexible in p_T, y
- need to know: reasonable (p_T, y) binning for LHC (?)
- for different jet algorithms – which jet algorithm(s) will be used at the LHC (?)
- later: extend to dijet production / Drell-Yan@NNLO / $\cdots$???

⇒ first results by summer

Computer Code

NLOJET++ best program
ee, ep, pp / ep,pp: 2- and 3-jet at NLO
subtraction method (no dependence on phase space slicing parameter)
full flexible ren, fact scales ($\boldsymbol{\mu_r = p_{T\,\mathrm{jet}}}$)
can compute multiple scales in single job
disadvantage: slow — for large Nbin: CPU time proportional to Nbin

JETRAD: iterative optimization of PS (high statistic in regions of small x-sect)

- not sure if relevant in this talk???????????