# A determination of $\alpha_s$ from scaling violations using truncated moments

Lorenzo Magnea
Università di Torino
I.N.F.N. Torino
magnea@to.infn.it

Montpellier – July 2, 2002

**Abstract**

The strong coupling, $\alpha_s(M_Z)$, is determined from scaling violations of the nonsinglet DIS structure function, using two novel techniques aimed at controlling and minimizing the theoretical error: a neural network parametrization of BCDMS and NMC data, and QCD evolution by means of truncated Mellin moments.

**Based on:** S. Forte *et al.* : hep-ph/0205286,
hep-ph/0204232.

# Outline of the method

- Our goal: a data–driven determination of $\alpha_s$.
  - minimizing theoretical biases and uncertainties.
  - accurately assessing the effects of errors and correlations.
- Difficulties in extracting $\alpha_s$ from DIS data.
  - Mellin moments: a clean analytic NLO solution, however:
    - Not directly measurable ($x \longrightarrow 0$ implies $\sqrt{s} \longrightarrow \infty$).
  - Momentum space evolution: no extrapolation, however:
    - Integro–differential equation numerically more difficult.
    - Parton parametrization necessary: theoretical bias, difficulties in assessing errors and propagating them to observables.
- Method: truncated moments.
  - Mellin moments over a truncated interval ($x_0 < x < 1$) are observable.
  - They obey a simple evolution equation approximated with arbitrary precision by a matrix equation.
  - In principle: a parametrization is not needed.
- Method: neural network parametrization of $F_2(x, Q^2)$.
  - In practice: data coverage and precision are not sufficient.
  - Neural networks: a bias–free parametrization of $F_2(x, Q^2)$.
  - All errors and correlations correctly taken into account.

# Truncated moments of parton distributions

Truncated Mellin moments

$$q_n(x_0,t) \equiv \int_{x_0}^{1} dx\; x^{n-1} q(x,t)\ .$$

satisfy the AP evolution equation $(t = \log\mu^2)$

$$\frac{d}{dt}\, q_n(x_0,t) = \frac{\alpha_s}{2\pi} \int_{x_0}^{1} dy\; y^{n-1} q(y,t)\; G_n\left(\frac{x_0}{y};\alpha_s\right)\ ,$$

where

$$G_n(x,\alpha_s) = \int_x^1 dz z^{n-1} P(z,\alpha_s)\ .$$

- As $x_0 \to 0$, $G_n$ becomes the anomalous dimension $\gamma_n$. Different moments evolve independently.
- For $x_0 \neq 0$, evolution couples $q_n$ with all $q_k$ with $k > n$. To see it, Taylor expand $G_n(x_0/y)$ around $y = 1$.
- The Taylor expansion of $G_n$ converges in $x_0 < y \leq 1$ ($G_n$ only has integrable singularities due to $+$ distributions at $y = x_0$). Truncating it at the $M$-th term yields the linear system

$$\frac{d}{dt} q_n(x_0,t) = \frac{\alpha_s}{2\pi} \sum_{p=0}^{M}\; c_{p,n}^{(M)}(x_0,\alpha_s)\; q_{n+p}(x_0,t)\ .$$

# Properties of truncated moments

- The matrix of anomalous dimensions governing the evolution of truncated moments is upper triangular.

  $\longrightarrow$ Analytic diagonalization by recursion relation.

- Moments with significantly different indices are weakly coupled for small $x_0$.

  $\longrightarrow$ Legitimate truncation at finite $M$.

- The convergence of the series of approximations for increasing $M$ can be studied systematically.

  $\longrightarrow$ Study effects of remainder of AP $r.h.s.$

  $\longrightarrow$ Evolve sample distributions with different methods.

- The convergence of the approximation as a function of $M$ is good (few percent error for $M \lesssim 20$), except for lowest nonsingular moments (sensitive to singularities at $y = x_0$).

  $\longrightarrow$ Improved version of the method is available

  (see A. Piccione, hep-ph/0107108)

  $\longrightarrow$ For all finite moments $M \lesssim 12$ suffices.

- The method has been extended to singlet and gluon distributions, without new difficulties.

  $\longrightarrow$ NLO analytic solution available in all cases.

  $\longrightarrow$ Threshold logarithms can be included if appropriate.

# Faithful parametrizations: problems

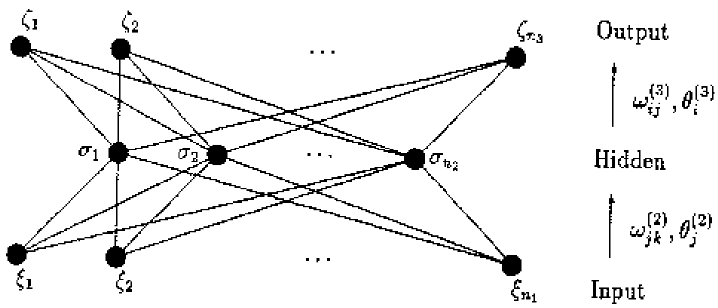
- Standard procedure for fitting PDF's and structure functions.
  - Choose a simple functional form with enough free parameters.
  - Fix parameters by minimizing $\chi^2$.
- Difficulties arise in determining errors on generic observables.
  - Errors and correlations of parameters require at least fully correlated analysis of data errors.
  - Error propagation to observables is difficult/wrong: many observables are nonlinear/nonlocal functionals of parameters.
  - Theoretical bias due to choice of parametrization is difficult to assess (effects can be large if data are not precise, $e.g.$ with polarized distributions).
- Goal: a representation of the probability measure $\mathcal{P}(F_2)$ in the space of structure functions $F_2(x, Q^2)$. Then, for any functional $\mathcal{G}(F_2)$,

$$\left\langle \mathcal{G}\left[F_2(x,Q^2)\right]\right\rangle = \int \mathcal{D}F_2 \; \mathcal{G}\left[F_2(x,Q^2)\right] \; \mathcal{P}(F_2) \; ,$$

  and similarly for higher moments.
- Note: a problem with a long history, active discussions in the context PDF global fits, different proposals available, see $e.g.$ hep-ph/0204316.

# Neural networks: a solution

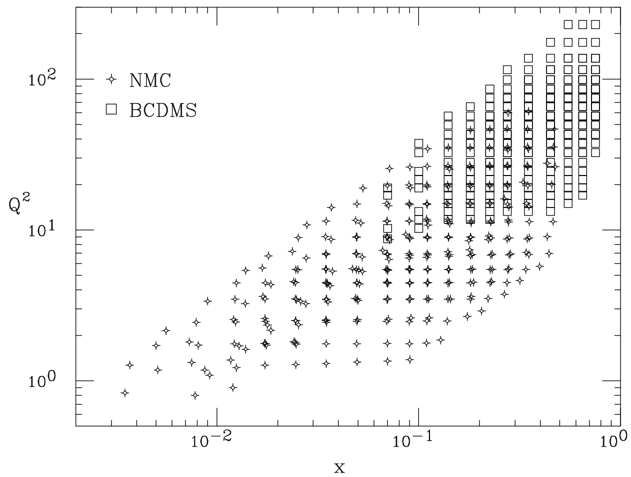
Neural networks are a class of algorithms providing robust, universal, unbiased approximants to incomplete or noisy data.




- Building blocks: neurons, $i.e$ input/output units characterized by activation
$$\xi_i = g\left(\sum_j \omega_{ij}\xi_j - \theta_i\right) \quad ,$$
with (typically) $g(x) \equiv 1/(1+\exp(-\beta x))$.
- Parameters: weights $\omega_{ij}$, thresholds $\theta_i$.
- Architecture: multilayer feed–forward NN. Each neuron receives input from neurons in preceding layer and feeds output to neurons in successive layer.
- Learning: supervised training by back–propagation. Network attempts matching data to output, weights and thresholds varied along steepest descent contours to minimize chosen error function.
- Assumption: smooth function. Size, architecture, learning cycle determined by statistical criteria.

# Neural networks for structure functions

- Data: BCDMS and NMC: 552 data points for the nonsinglet structure function $F_2^{(NS)}(x, Q^2)$.




- Method: Monte Carlo + neural networks.
  - Step 1 Generate an ensemble of $N_{\rm rep}$ pseudo–data sets, with the correct multivariate distribution given by experimental errors, fully correlated.

$$F_i^{(art)\,(k)} = (1 + r_{i,N}^{(k)}\,\sigma_{i,N})\left[F_i^{(exp)} + \frac{\sum_a r_{i,a}^{(k)} f_{i,a}}{100}\,F_i^{(exp)} + r_{i,s}\,\sigma_{i,s}^{(k)}\right].$$

  - Step 2 Train $N_{\rm rep}$ neural networks, each one using one pseudo–data set.
  - Step 3 Evaluate averages, errors, correlations of observables using $N_{\rm rep}$ networks as Monte Carlo representation of probability measure in the space of structure functions.

# Determination of $\alpha_S$: choices

- Truncation point and fitting range.
  Criteria:
  - Data coverage $\longrightarrow$ small error on moments.
  - High $Q^2$ $\longrightarrow$ small power correction.
  - Small $x_0$, few intermediate scales $\longrightarrow$ small correlations between neighboring moments.

  Choices: $x_0 = 0.03; 20\ \mathrm{Gev}^2 < Q^2 < 70\ \mathrm{Gev}^2; n_{\rm sc} = 3.$
- Evolution equation
  - NLO evolution with matching at quark thresholds.
  - Size: $M = 11$ with $n_{\rm min} = 1$.
  - Auxiliary parameter for improved evolution: $N = 6$.

  $\longrightarrow$ Accuracy on evolution: $0.1\%$.
- Fitted moments
  Criteria:
  - Precision of fit requires $n_{\rm fit} > 3$
  - High correlations between neighboring moments may cause *off–diagonal instabilities* $\longrightarrow n_{\rm fit} < 6$.

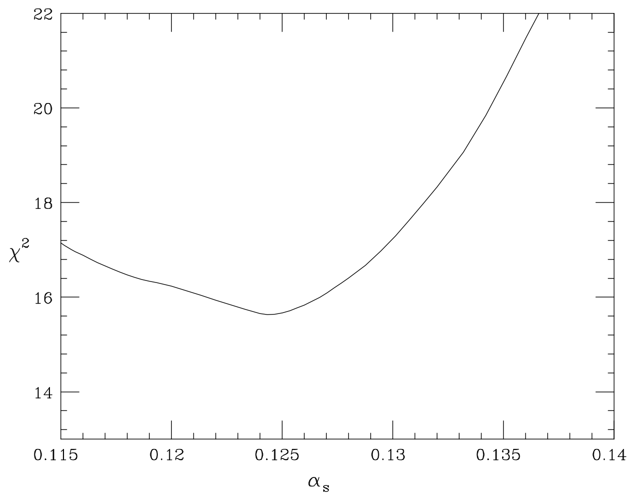
  Choices: Fitted moments: $n = 2, 4, 5, 6, 8$.
- Result with statistical error.

$$\alpha_S(M_Z) = 0.124 \ {}^{+\ 0.004}_{-\ 0.007} \ \text{(stat.)}\ .$$

Note: All fit parameters have been varied in the window of stability with negligible effects on the result.

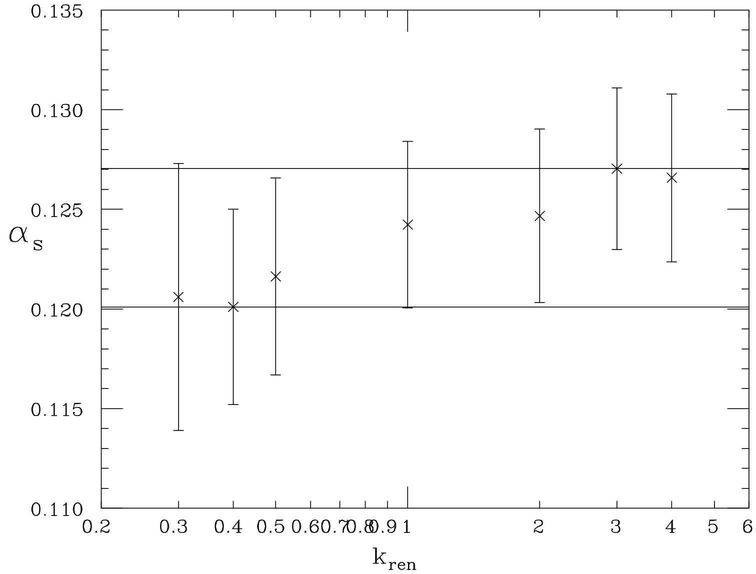
# Determination of $\alpha_S$: errors

- Statistical error: asymmetric $\chi^2$.

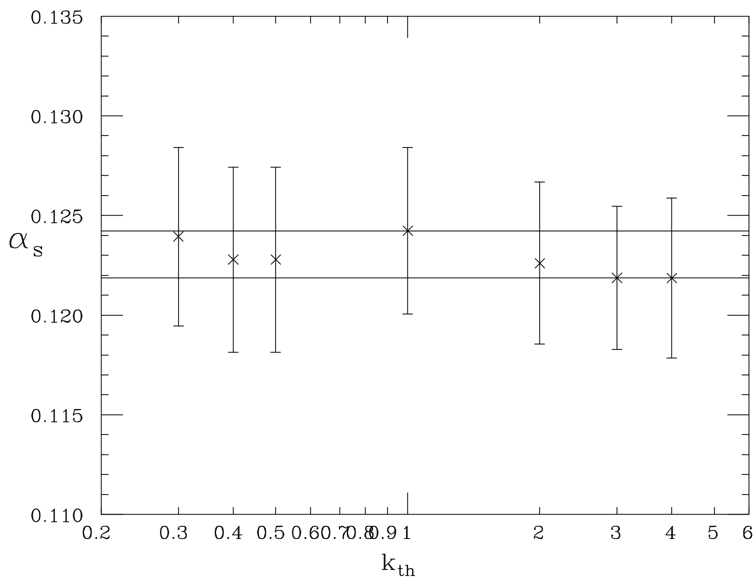



- Theoretical error: power correction.
  Can be: kinematical (target mass corrections); dynamical (higher twist corrections); due to elastic contributions at $x = 1$. All are negligible ($< 1\%$) thanks to our choice of $Q^2$ range.
- Theoretical error: NNLO and higher perturbative evolution.
  Estimated varying renormalization scale (no factorization mass dependence in DIS scheme), $\mu^2_{ren} = k_{ren}Q^2$. Not negligible, indicating sizeable NNLO corrections. $\sigma_{\rm ren} = {}^{+\,0.003}_{-\,0.004}$.
  Note: possible enhancement of threshold logarithm effects.
- Theoretical error: Heavy quark thresholds.
  Estimated varying the threshold position as $Q^2 = k_{th}M_q^2$. Nearly negligible in our $Q^2$ range (only $b$ threshold included for some $k_{th}$). $\sigma_{\rm th} = {}^{+\,0.000}_{-\,0.002}$.

- Varying the renormalization scale, $\mu_{ren}^2 = k_{ren} Q^2$.




- Varying thresold positions, $Q^2 = k_{th} M_q^2$.




- Our final result

$$\alpha_S(M_Z) = 0.124 \,{}^{+\,0.004}_{-\,0.007}\ \text{(exp.)}\ {}^{+\,0.003}_{-\,0.004}\ \text{(th.)}\ .$$

# Summary

- We have determined the strong coupling $\alpha_S(M_Z)$ at NLO from scaling violations of NMC and BCDMS data for the nonsinglet DIS structure function, with the result

$$\alpha_S(M_Z) = 0.124 \ {}^{+\,0.004}_{-\,0.007} \ \text{(exp.)} \ {}^{+\,0.003}_{-\,0.004} \ \text{(th.)} \ .$$

- We have minimized theoretical biases and errors by using
  - evolution with truncated moments: minimize effects of parametrization by cutting out small $x$ range; simple analytical expression for evolved distributions.
  - parametrization with neural networks: no theoretical bias, statistical control over accuracy of interpolation and extrapolation.
  - choice of $(x, Q^2)$ range to minimize effects of nonperturbative and higher order corrections.
- Statistical error is significantly larger than theoretical one: improved data (especially deuteron), wider $Q^2$ range would significantly reduce errors.
- Central value is on high side of world average (though fully compatible); error is asymmetric.
- NOTE: Threshold logarithms $\log\left(Q^2(1-x)\right)$ may affect our determination more significantly than others (see $\mu_{\text{ren}}$ dependence). They can be systematically resummed in Mellin space with truncated moments.