KRAKOW SCHOOL OF INTERDYSCIPLINARY PHD STUDIES

Nuclear PDF Determination
Using Markov Chain Monte Carlo Methods

Nasim Derakhshanian

Dr. Aleksander Kusina

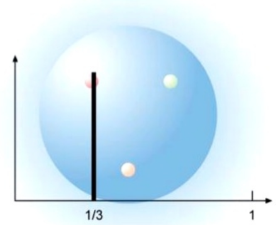
nCTEQ collaboration

IFJ PAN

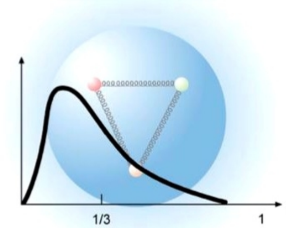
Department of Theoretical Particle Physics

Polish Particle and Nuclear Theory Summit (2PiNTS)

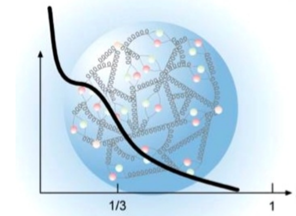
November 2023, Krakow



Free quarks



Bound quarks



Bound quarks + QCD effects

Parton Distribution Function (PDF):

The probability $\mathbf{f_{a/p}(x,\mu)}$ that a parton **a** carries fraction **x** of the proton's momentum

μ: Factorization scale
X: momentum fraction

Factorization Theorem

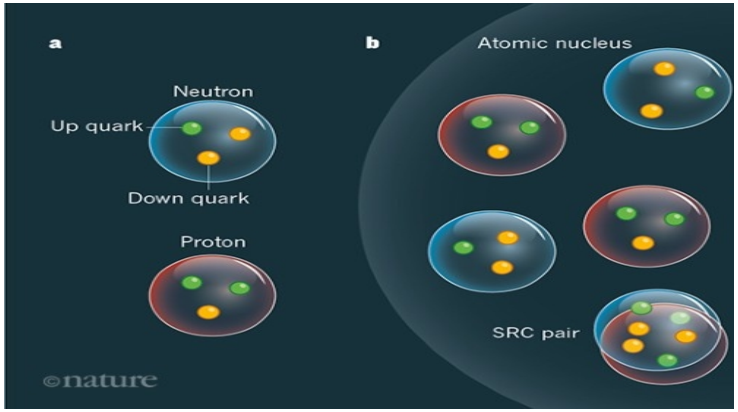
$$\sigma_{P\gamma\to c} = f_{P\to a} \otimes \hat{\sigma}_{a\gamma\to c}$$

Parton densities (long-distance) | Parton interaction (short-distance)

PDF properties:

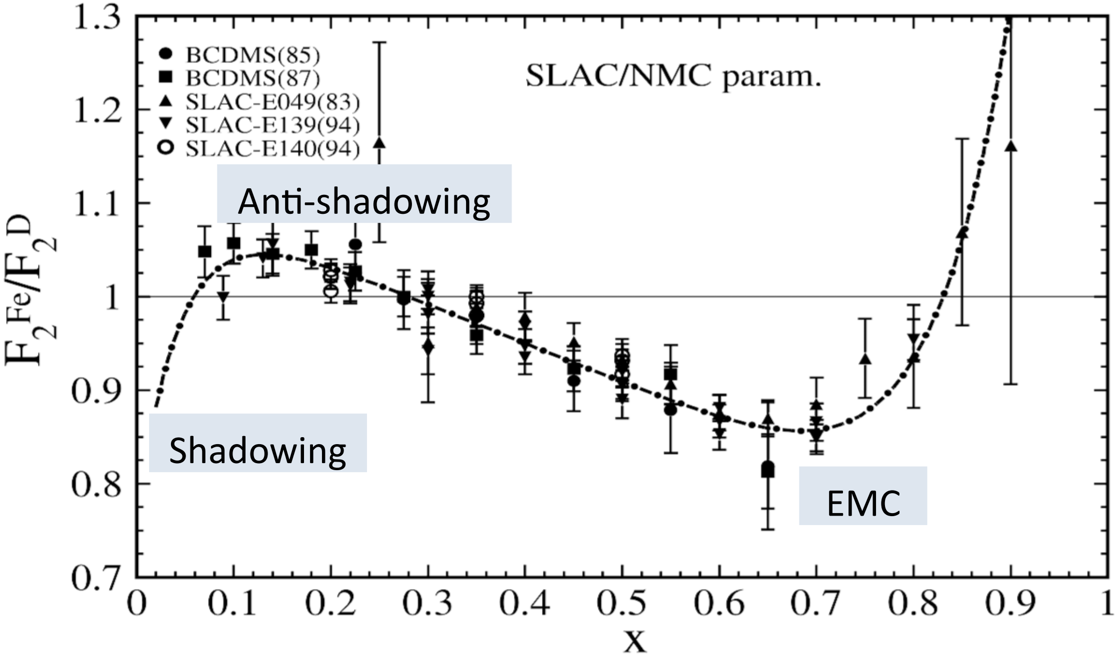
- Universal
- Constrained through momentum and number sum rules
- μ^2-dependence governed by DGLAP evolution equations
- Non-perturbative: x-dependence of PDF is NOT calculable in pQCD

→ **extracting from a fit to experimental data**



Nuclear PDFs(nPDFs):

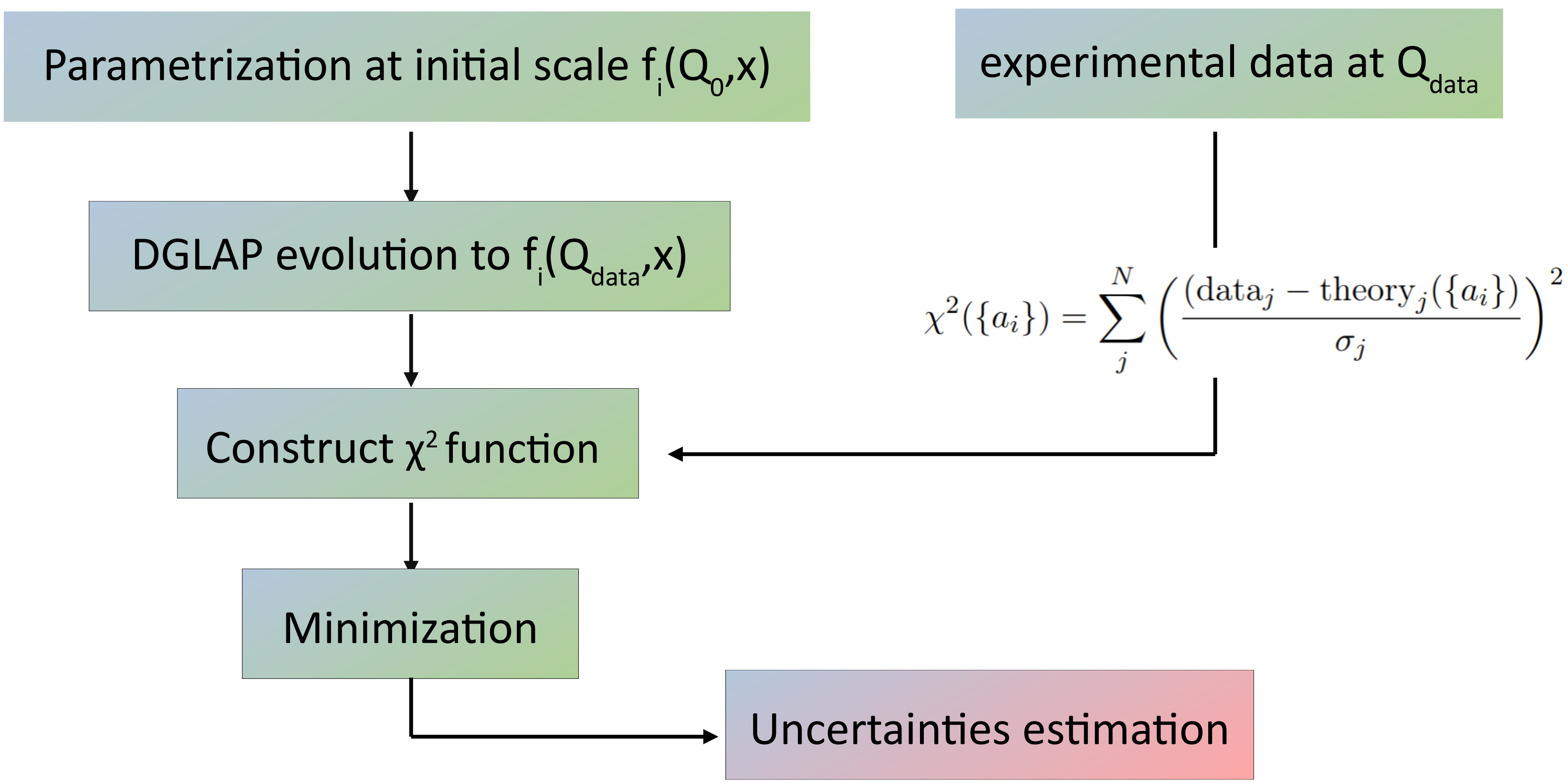
$$F_2^A(x) \neq ZF_2^p(x) + NF_2^n(x)$$



Motivations:

- Interpreting heavy-ion collision data
- Understanding Nuclear Structure

QCD Global analysis:



nCTEQ15 framework for nuclear PDF:

Kovarik et al., arXiv:1509.00792

Functional form for bound proton at Q_0:

$$x f_i^{p/A}(x, Q_0) = c_0\, x^{c_1} (1-x)^{c_2} e^{c_3 x} (1 + e^{c_4} x)^{c_5}$$

$$i = u_v, d_v, g, \bar{u} + \bar{d}, s, \bar{s}$$

Atomic number dependence is characterized in the c_k coefficients as

$$c_k \to c_k(A) \equiv p_k + a_k(1 - A^{-b_k}), \qquad k = \{1, \ldots, 5\}.$$

PDF of a nucleus (A – mass, Z – charge):

$$f_i^{(A,Z)}(x, Q) = \frac{Z}{A} f_i^{p/A}(x, Q) + \frac{A-Z}{A} f_i^{n/A}(x, Q)$$

new nCTEQ global nPDF release: CJ15

Accardi et al., arXiv:1602.03154

Functional form for bound protons at Q_0:

$$x f_i^{p/A}(x, Q_0) \;=\; c_0 x^{c_1} (1-x)^{c_2} (1 + c_3 \sqrt{x} + c_4 x)$$

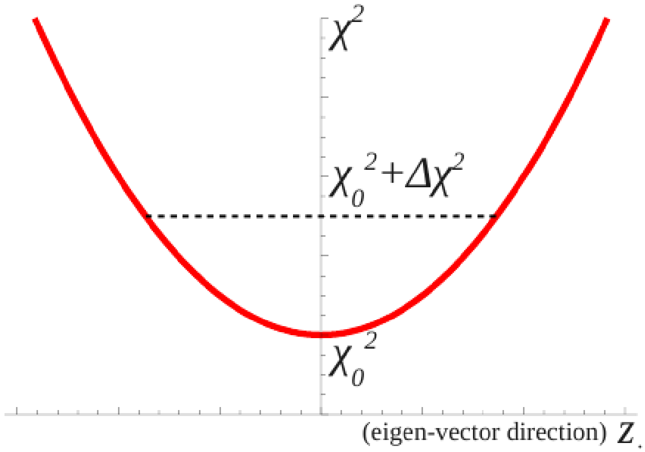
$$i = u_v, d_v, g, \bar{u}+\bar{d}, s, \bar{s}$$

Atomic number dependence is characterized in the c_k coefficients as

$$c_k \;\to\; p_k + a_k \ln(A) + b_k \ln^2(A). \qquad k = \{1, ..., 5\}.$$

PDF of a nucleus (A – mass, Z – charge):

$$f_i^{(A,Z)}(x, Q) = \frac{Z}{A} f_i^{p/A}(x, Q) + \frac{A-Z}{A} f_i^{n/A}(x, Q)$$



PDF uncertainties estimation:

Hessian method: Common method for estimating uncertainties in PDFs.

relying on the Gaussian approximation of $\Delta\chi^2$

Shortcomings:
- Non-gaussian errors
- Global minima judgment

nPDF difficulties
- Lacking data (need low-x & precise data, for several nuclei)
- Complexity and nature of nuclear effects

deeper insight

Markov Chain Monte Carlo method

Markov Chain Monte Carlo (MCMC)

A sequence of random variables where the current value is dependent on the value of the prior variable (Memory-less property)

A technique for randomly sampling a probability distribution and approximating a desired quantity.

Bayes theorem:

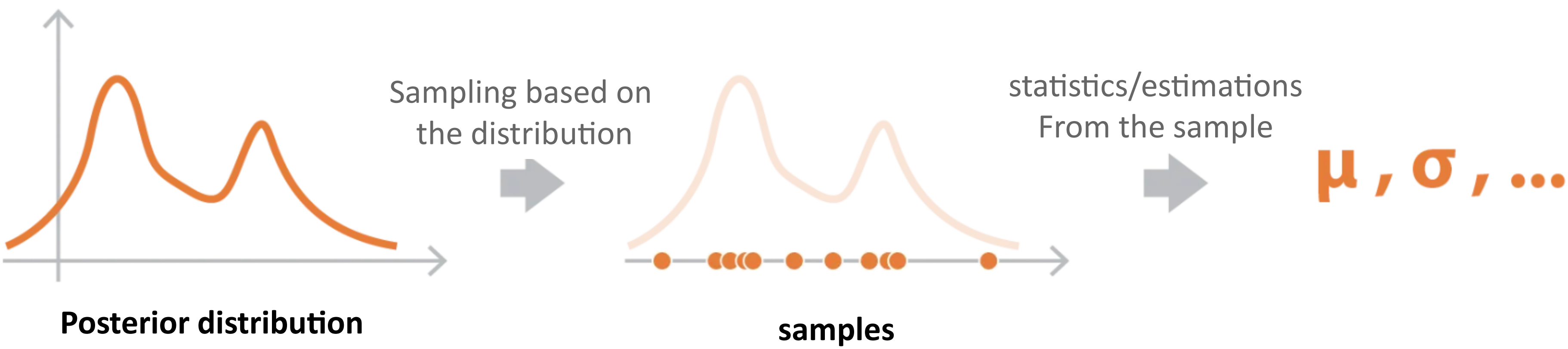
$$p(\theta \mid \text{data}) = \frac{p(\text{data} \mid \theta) \cdot p(\theta)}{p(\text{data})}$$

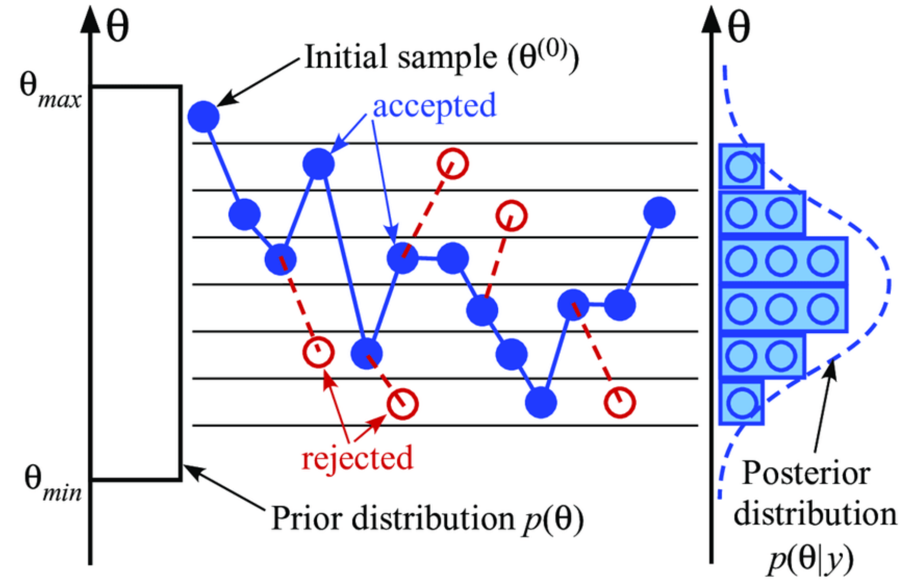
Posterior: $p(\theta \mid \text{data})$
Likelihood: $p(\text{data} \mid \theta)$
Prior: $p(\theta)$
Normalization: $p(\text{data})$

Prior: initial belief about the parameter before considering the data.
Likelihood: probability of observing the data given a specific value of the parameter.
Posterior: updated belief about the parameter given the data.

We aim to find the set of nPDF parameters that maximizes the posterior probability distribution given the experimental data.

Bayesian inference → **MCMC algorithms**





Metropolis algorithm:

- Initialize parameters
- for i=1 to i=N:

Propose a new sample

Generate proposed parameters via proposition function: $\theta^* \sim q(\theta^* \mid \theta_i)$

Sample from uniform distribution: $u \sim U(0,1)$

Compute acceptance ratio: $\alpha = p(\theta^* \mid D) / p(\theta_i \mid D)$

If $u < \min(1, \alpha)$ then $x_{i+1} = x^*$

Judgment of proposed sample

- Else $x_{i+1} = x_i$

Metropolis-Hasting: $x_{t+1} = \mathcal{N}(x_t, C_0)$

Adaptive Metropolis-Hasting: $x_{t+1} = \beta\mathcal{N}(x_t, C_0) + (1-\beta)\mathcal{N}(x_t, \hat{C}_n)$

PDF Global analysis:

Parametrization at initial scale $f_i(Q_0,x)$

DGLAP evolution to $f_i(Q_{data},x)$

Construct χ^2 function

Minimization

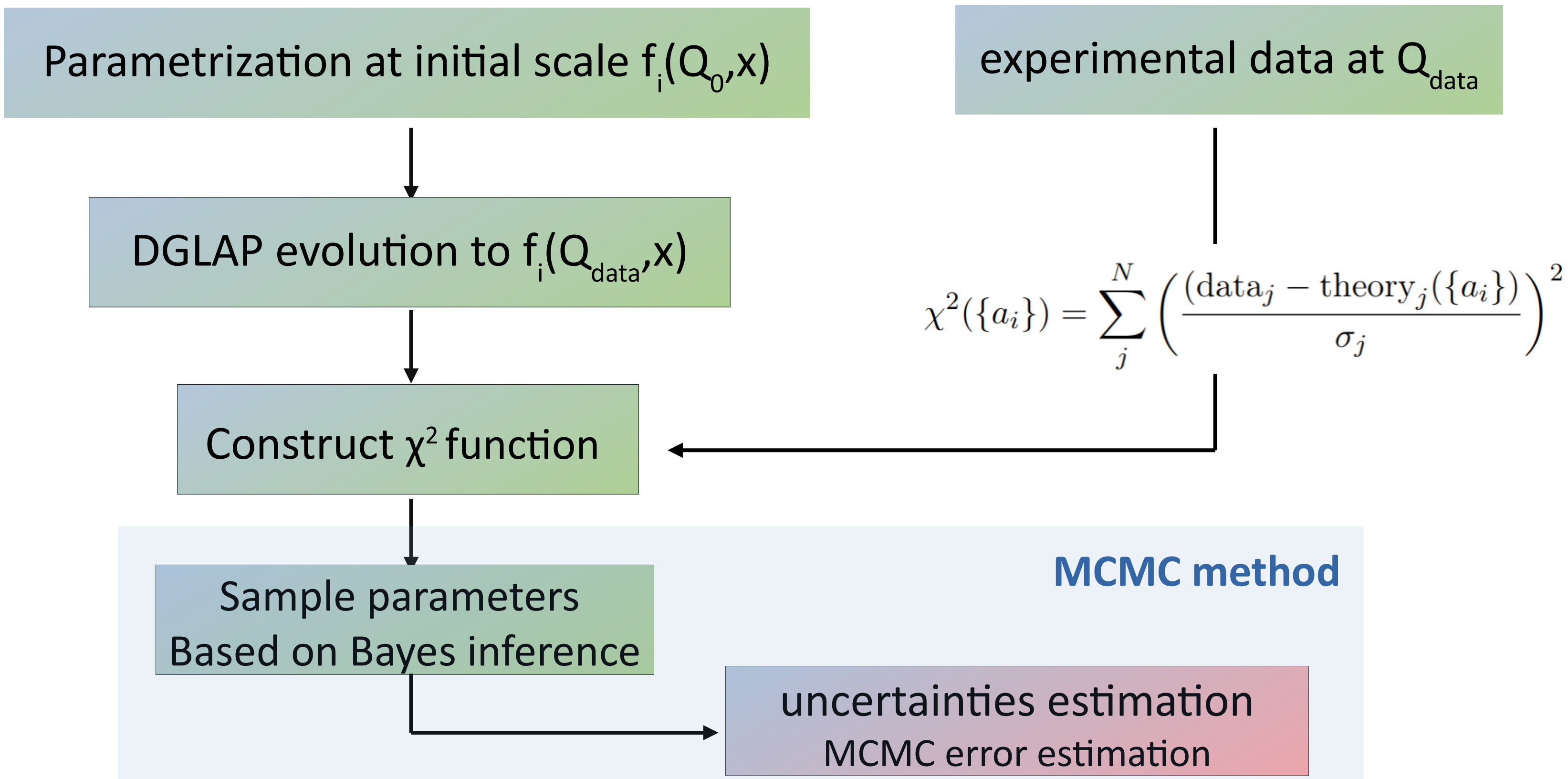
uncertainties estimation
Gaussian error propagation

experimental data at Q_{data}

$$\chi^2(\{a_i\}) = \sum_j^N \left(\frac{(\text{data}_j - \text{theory}_j(\{a_i\})}{\sigma_j} \right)^2$$

Hessian method

PDF Global analysis:



CJ15 nPDF parametrization

$$x f_i^{p/A}(x, Q_0) = c_0 x^{c_1}(1-x)^{c_2}(1 + c_3\sqrt{x} + c_4 x)$$

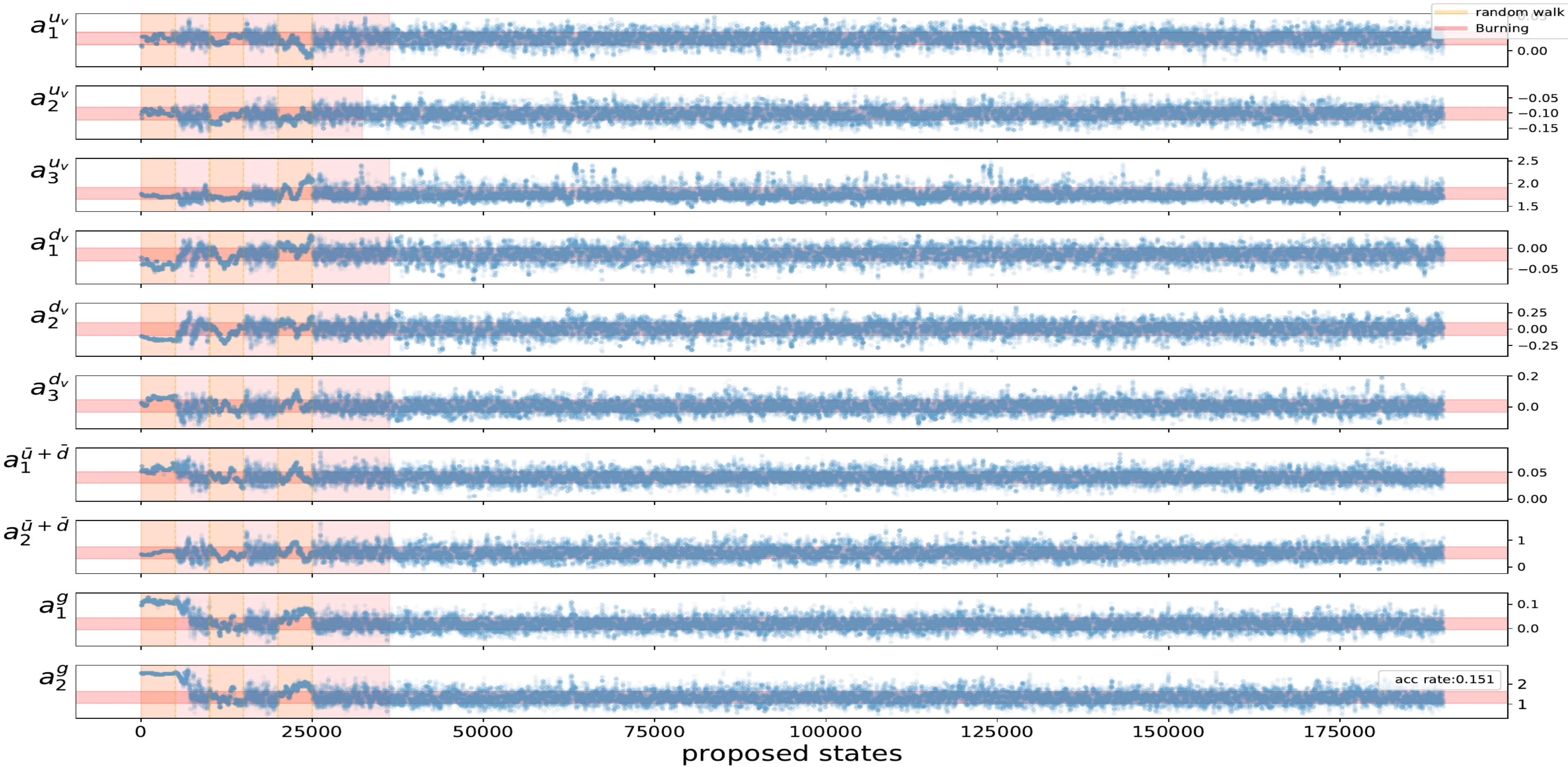
$$c_k \to p_k + a_k \ln(A) + b_k \ln^2(A).$$

Generating the Markov Chain of nPDF parameters:

Each point of the chain is representing a set of nPDF parameters

(6 valance, 2 sea quarks and 2 gloun)

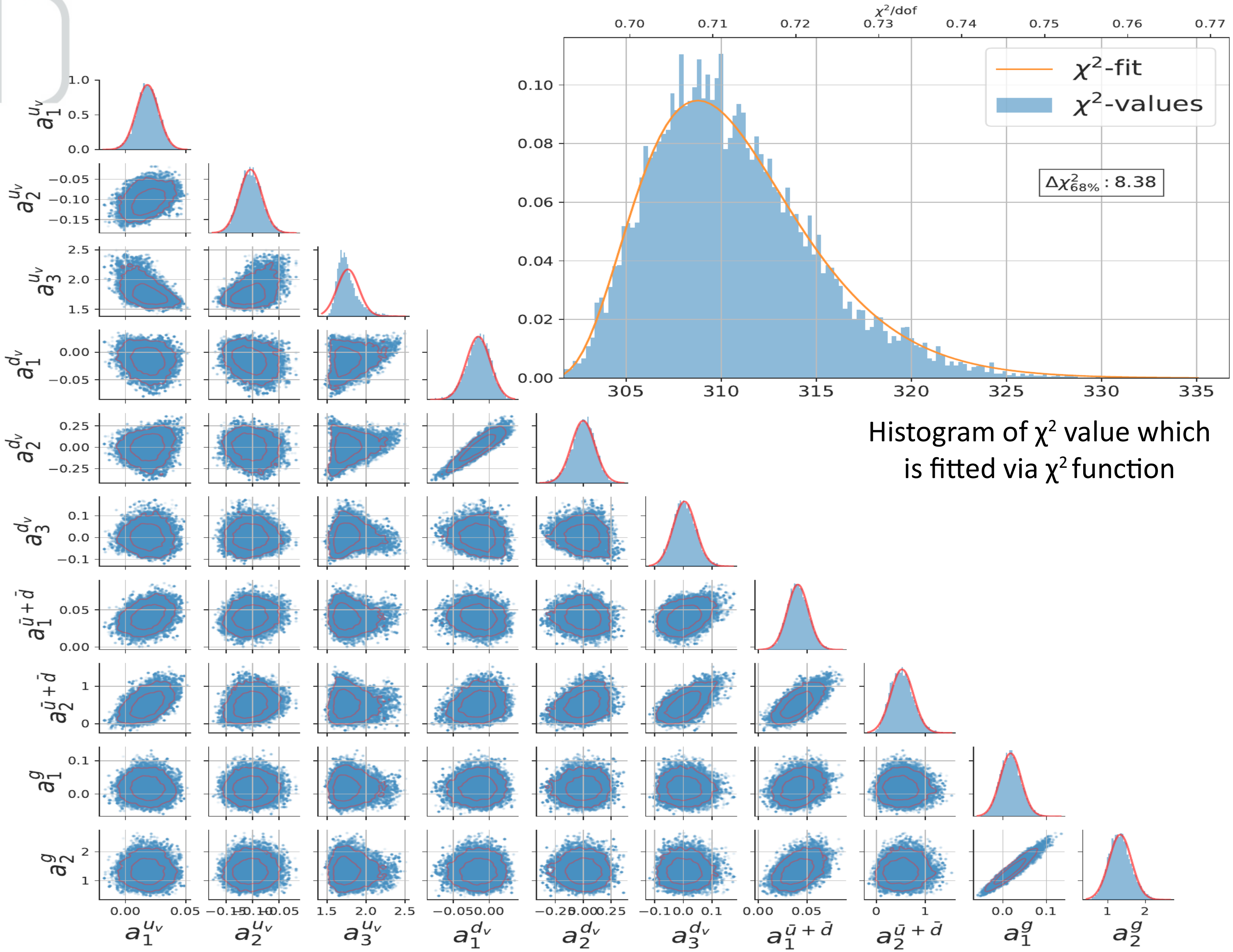
DIS and W/Z boson data:
436 data points



MCMC can reveal non-Gaussian features of the underlying distribution

Pairwise plot

diagonal: histogram of each parameter
off-diagonal: 2D correlation plots between parameters



Histogram of χ^2 value which is fitted via χ^2 function

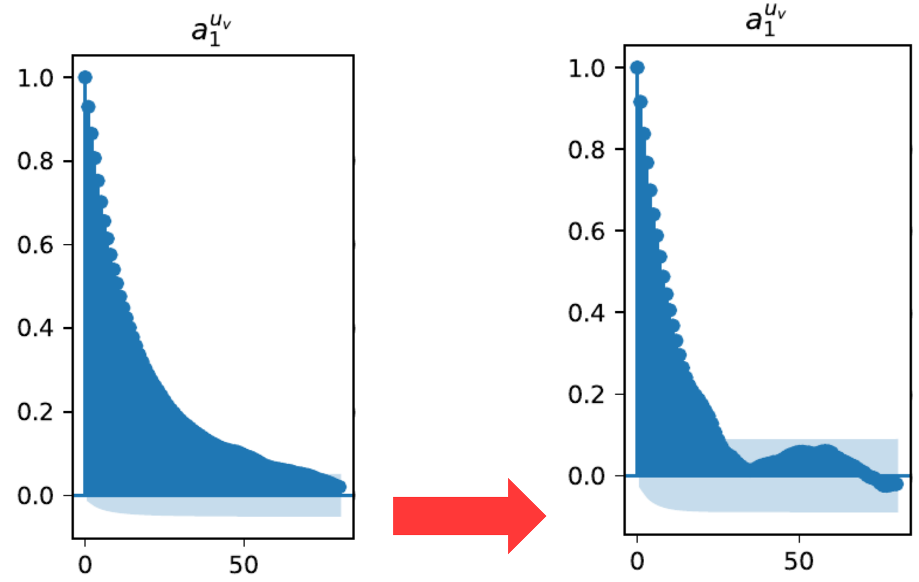
Error estimation:

Autocovariance: $\mathrm{Cov}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(x_{t+k}-\bar{x})(x_t-\bar{x}),$

Autocorrelation: $\rho(k) = \frac{\mathrm{Cov}(k)}{\mathrm{Cov}(0)}$

Autocorrelation time: $\tau_{int} = \frac{1}{2}\sum_{-\infty}^{+\infty}\rho(k)$

Autocorrelation function versus time interval



Thinning by rate 40

discard all except every k-th point of the chain

MCMC error (correlated)

$$\sigma^2_{MCMC} = 2\,\tau_{int}\,\sigma^2_{MC}$$

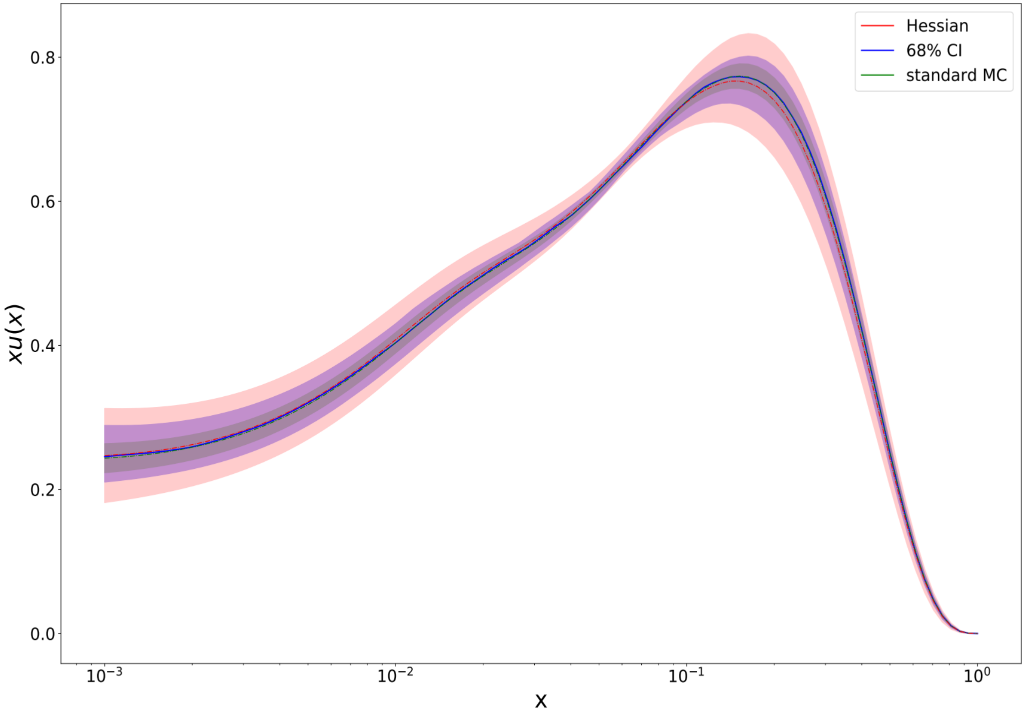
MC error (uncorrelated)

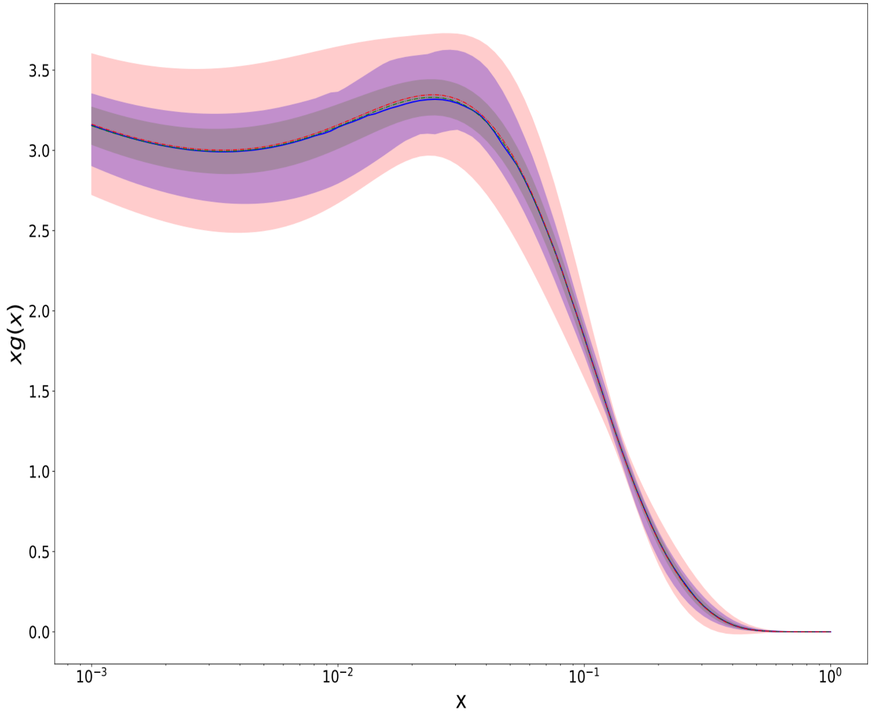
$$\sigma^2_{MC} = \frac{1}{n-1}\sum_{t=1}^{n}(X_t-\hat{\mu})^2$$

MCMC approach:

- Generating the Markov Chain
- Thinning the chain
- Dumping PDF corresponding to each unit of the thinned chain
- Evaluating the error band determined from Monte Carlo error

LHAPDF (set of nPDF grids):





Thank you for your attention

Acknowledgment:
This work was supported by Narodowe Centrum Nauki under grant no.\ 2019/34/E/ST2/00186.