QCD Instantons Searches at High Q^2

Status Report

Paweł Sopicki

11.09.2013

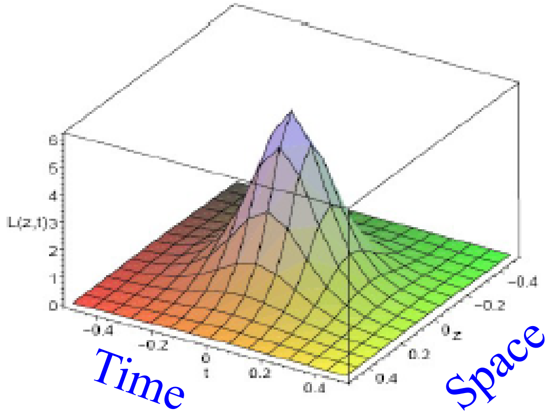
Outline

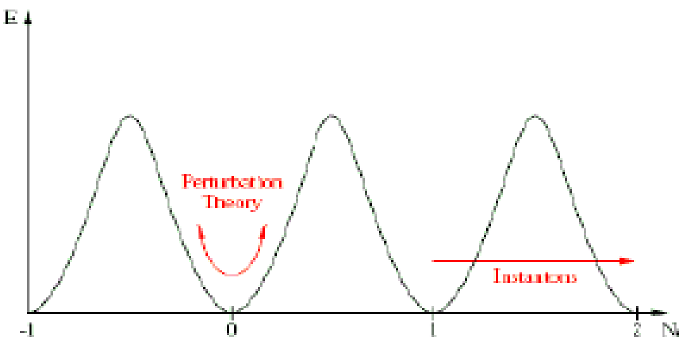
- Introduction & Analysis Strategy
- TMVA analysis
- Upper limits calculations
- Summary

QCD Instantons
& HERA

Instantons

- Solution to Yang-Mills equation of motion in 4d Euclidean spacetime as a longrange fields A_μ with finite action $S(A_\mu) < \infty$
- Physical interpretations: pseudo-particle or tunneling process between topologicaly different vacuum states



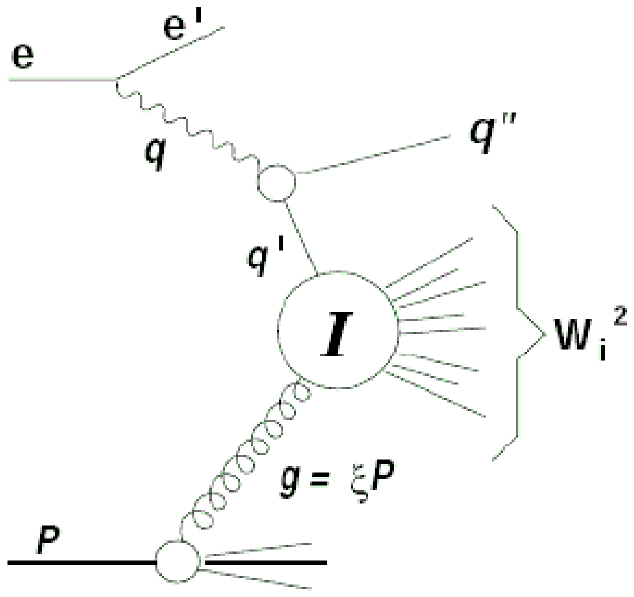


- Lead to violations of baryon-lepton number (in EW) and chirality (QCD)
- Non-perturbative effect with cross section $\sim e^{-4\pi/\alpha}$ (α-coupling constant)

QCD Instantons
& HERA

QCD Instantons

- Instanton-induced events produced in quark-gluon fusion
- Theory and phenomenology worked out by A. Ringwald and F. Schremp
- QCDIns Monte Carlo generator makes full event topology available



$$150 < Q^2 < 15000\,GeV^2, \quad 0.2 < y < 0.7$$

$$\sigma^{(I)}_{HERA} = 10^{+2}_{-2}\,pb$$

Variables of I-subprocess:

$$Q'^2 = -q'^2 = -(q - q'')^2$$

$$x' = Q'^2 \; / \; (2\, g \cdot q')$$

$$W_i^2 = Q'^2 \; (1 - x' \;)/x'$$

Analysis strategy

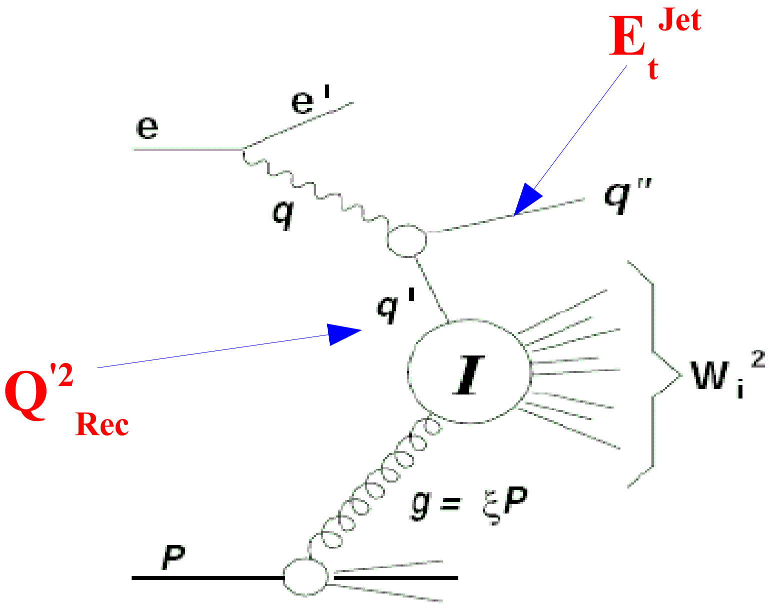
I. DIS Selection

II. Jet level

- Jet selection: kT algorithm
- HCMS: find current jet (**E_t^{Jet}**, **Q'^2_{Rec}**) and remove its objects from HFS.

  Current jet requirements:

  - Maximal E_t and $E_t > 4$GeV
- Calculate **<η>** of HFS and define „instanton band" as objects within <η> +/- 1.1

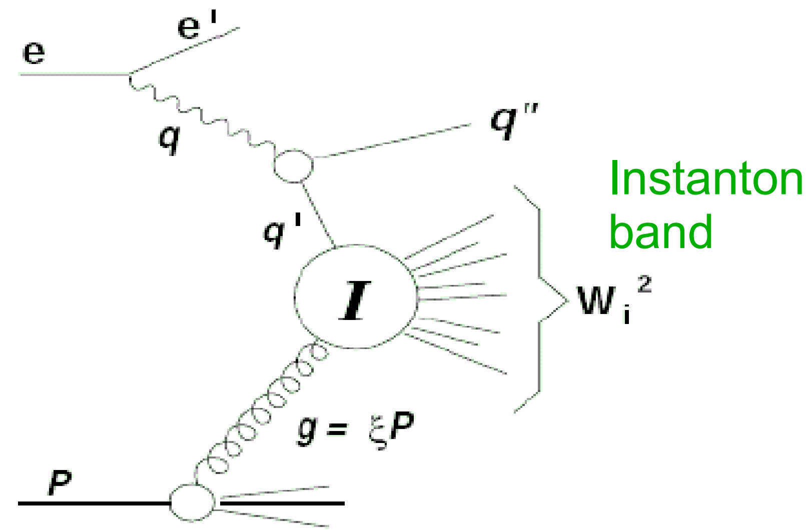


Analysis strategy

I. DIS Selection

II. Jet level

III. **Instanton level**



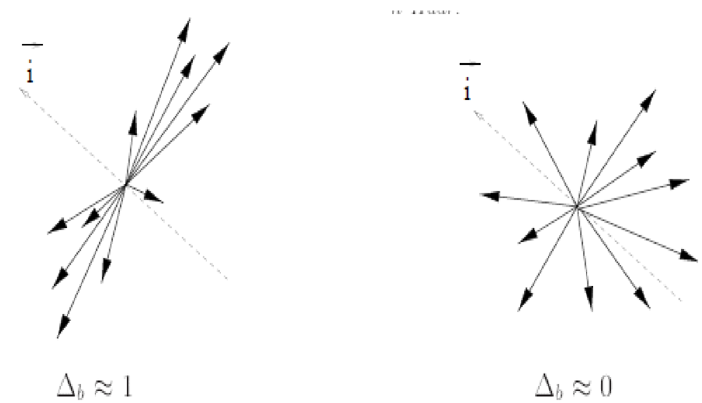
- Boost HFS objects from „instanton band” to „instanton rest frame”

  q'+ξP = 0 , ξ = <ξ> = 0.076

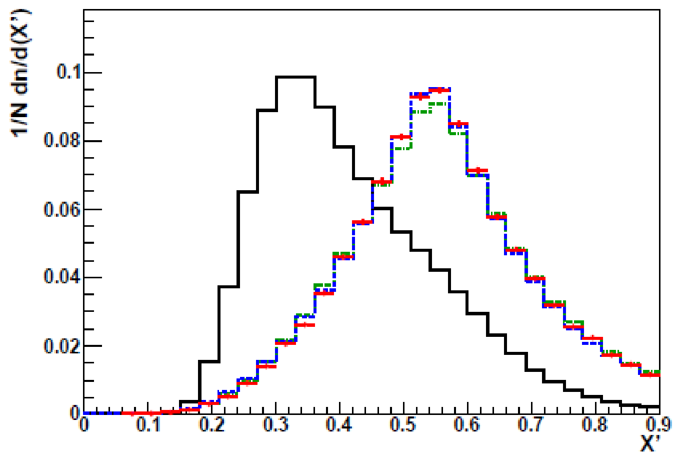
- Calculate observables
  - Transverse energy of the band **Et_{band}**
  - **N_{chr}** – number of charged particles in band
  - Topological observables: **Sphericity**, **Fox-Wolfram moments**, **E_{In}**, **E_{out}**, **Δ_{band}**

$$\Delta_b = (E'_{in,B} - E'_{out,B}) \,/\, E'_{in,B}$$

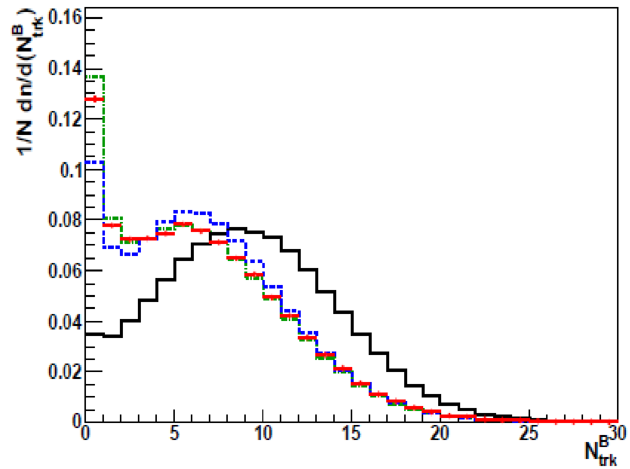
$$E_{out} = \min \sum_{n\ Hadr.} |\vec{p_n} \cdot \vec{i}|$$
$$E_{in} = \max \sum_{n\ Hadr.} |\vec{p_n} \cdot \vec{i}|$$



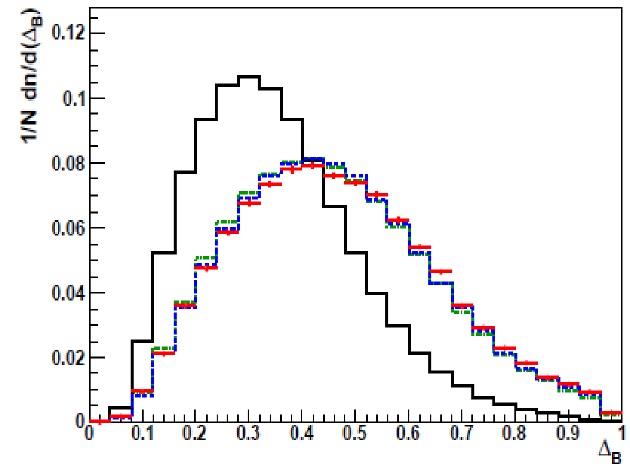
Observables for TMVA



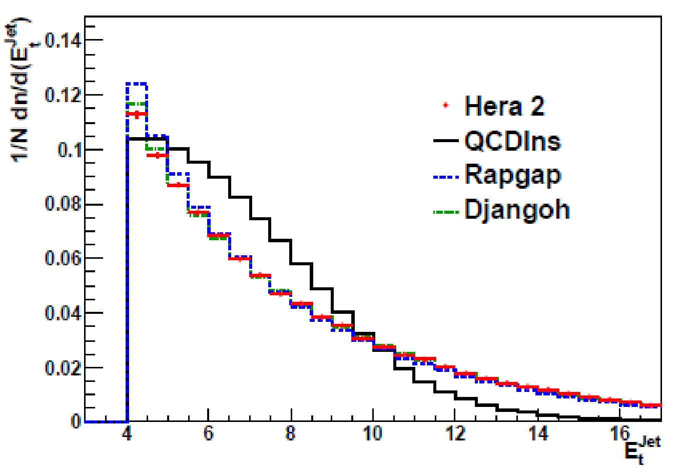
X'_{Rec}



N^{B}_{trk}



Δ_B

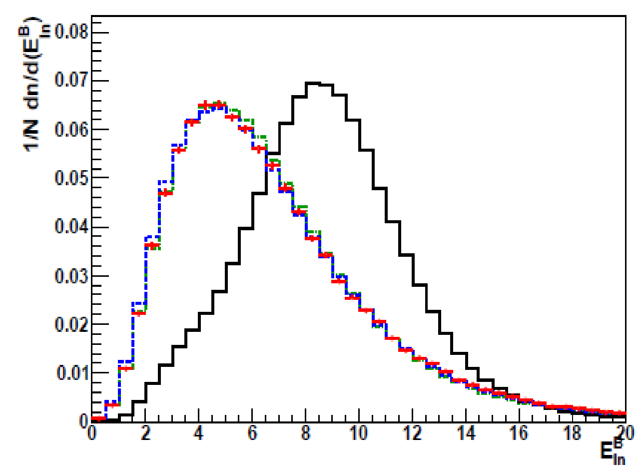


E^{Jet}_{T}

A set of five observables selected:

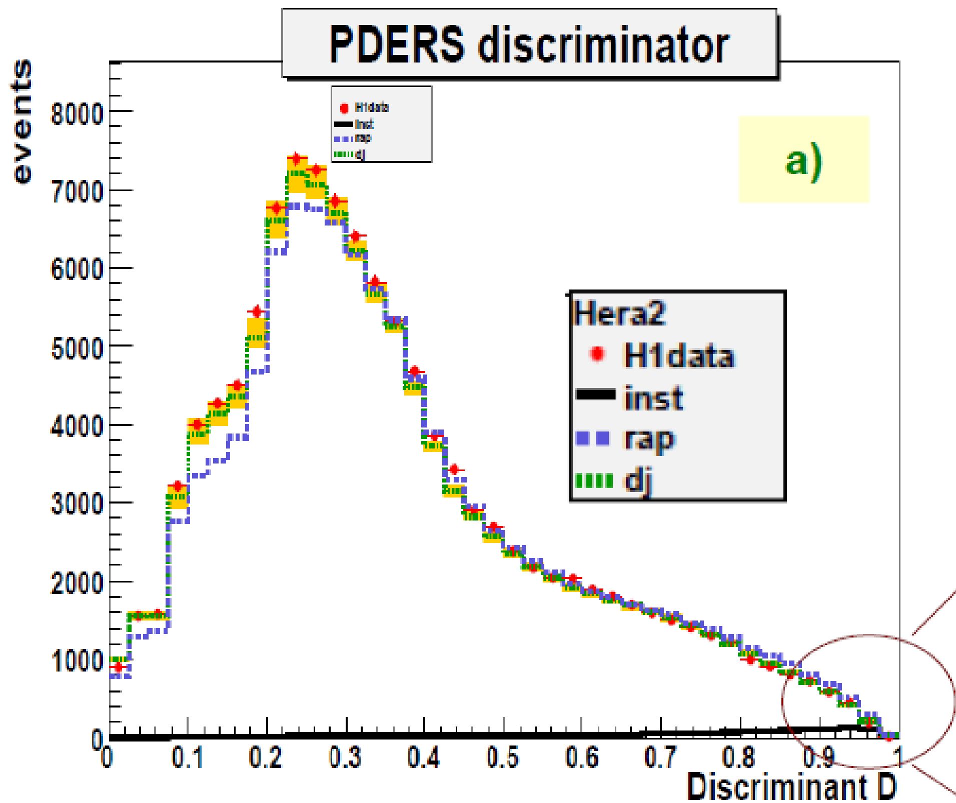
Good S/B separation and relativly good discriminator's background region description

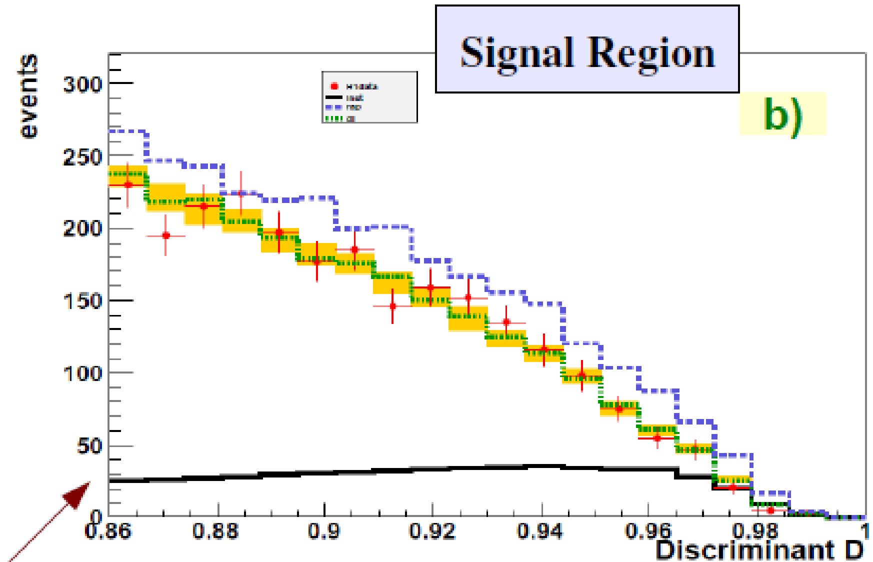
Methods used: PDERS, MLP, BDT and BDTG

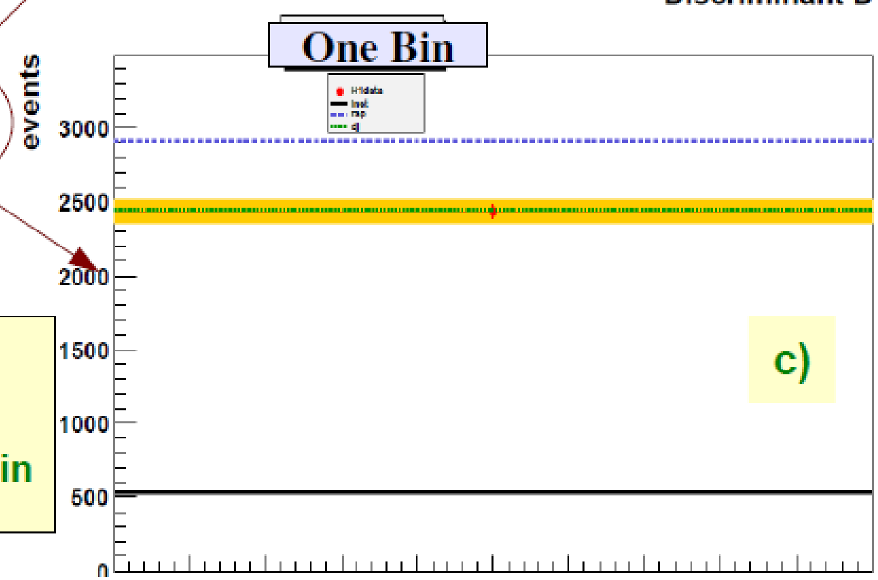


E^{B}_{In}

TMVA: Results from PDERS method







a) Full discriminant range
b) Signal region D>0.86
c) "counting experiment", signal region into one bin

Upper limits

QCDInstanton cross section in the analysis phase space is: 10 +/- 2 pb

For upper limit calculation a CL_s method has been used

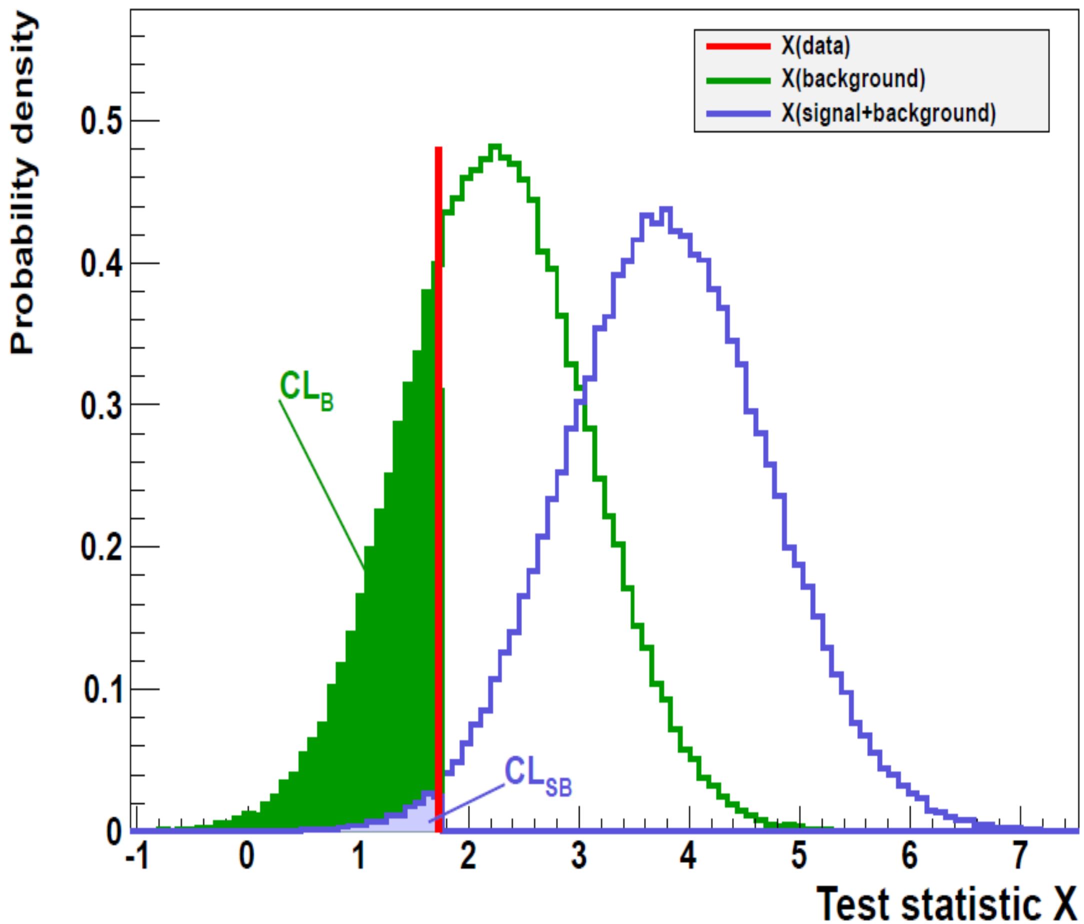
In the CLS method distributions of some variable are used – information from bins is combined into test statistics X:

$$X = \sum_{i=1}^{N_{\mathrm{bin}}} w_i n_i$$

where n_i is the number of D/B/S events in the *i*-bin, w_i – weights that are calculated from a set of linear equations and constructed in such a way that bins with large amount of signal get high (positive) weights

Test statistic distribution

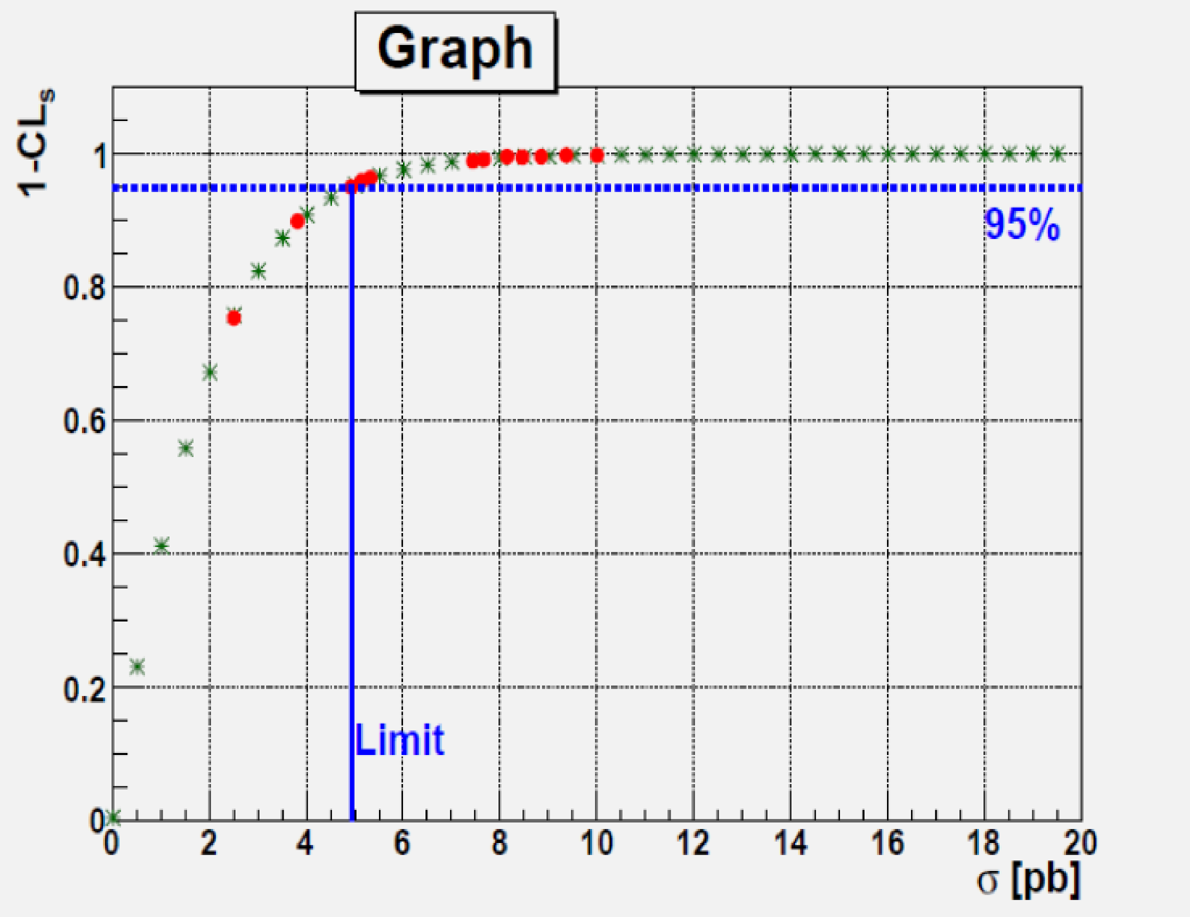
Lets construct test statistics for Data, Background and Backgr+Signal



$$CL_S = \frac{CL_{SB}}{Cl_B} \qquad Confidence\ Level: CL = 1 - CL_S$$

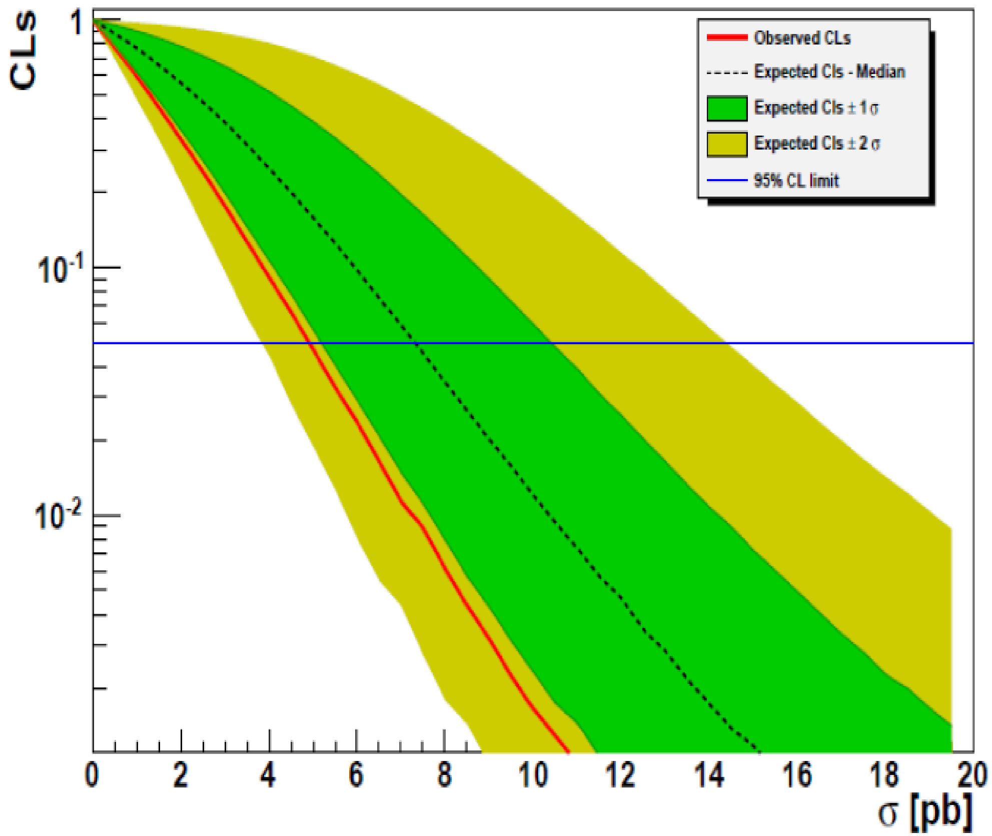
Upper limit scan

Upper limit for 95% confidence level is found by scanning and looking for such a point in which $1\text{-}CL_s = 0.05$

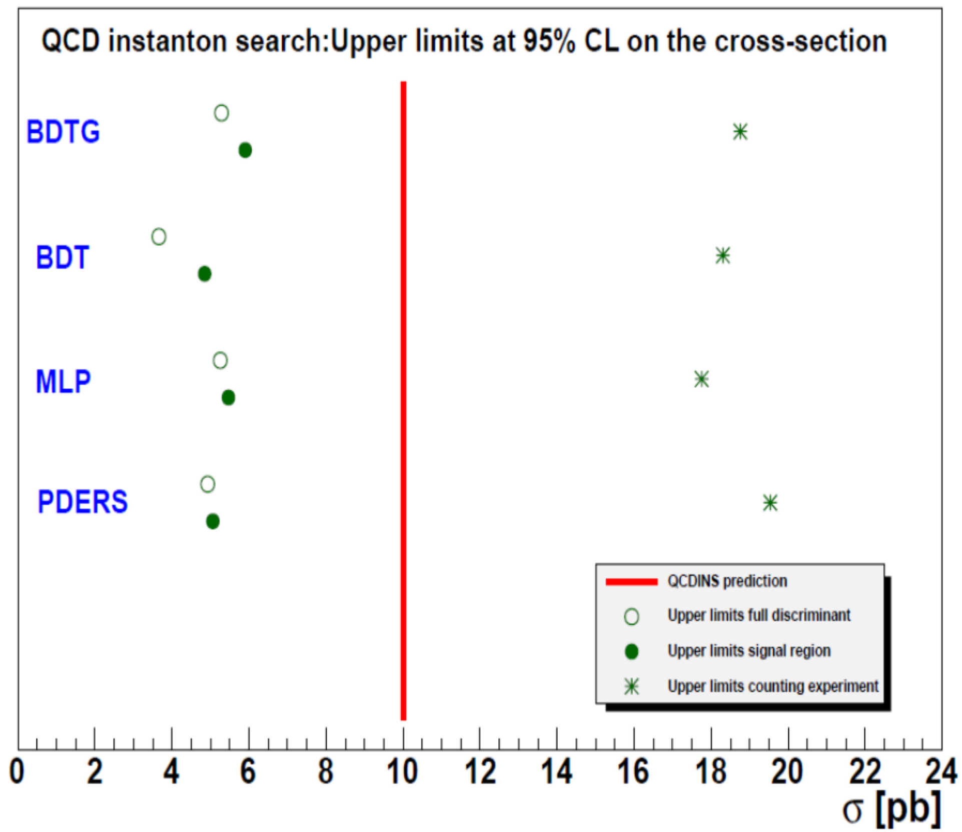


Red points – a quick algorithm results (100k pseudo-exp per point)
Green points – a scan with 2M pseudo-exp per point

Upper limits: previous results



CL observed and expected vs the Instanton Cross Section



Upper limits: issues

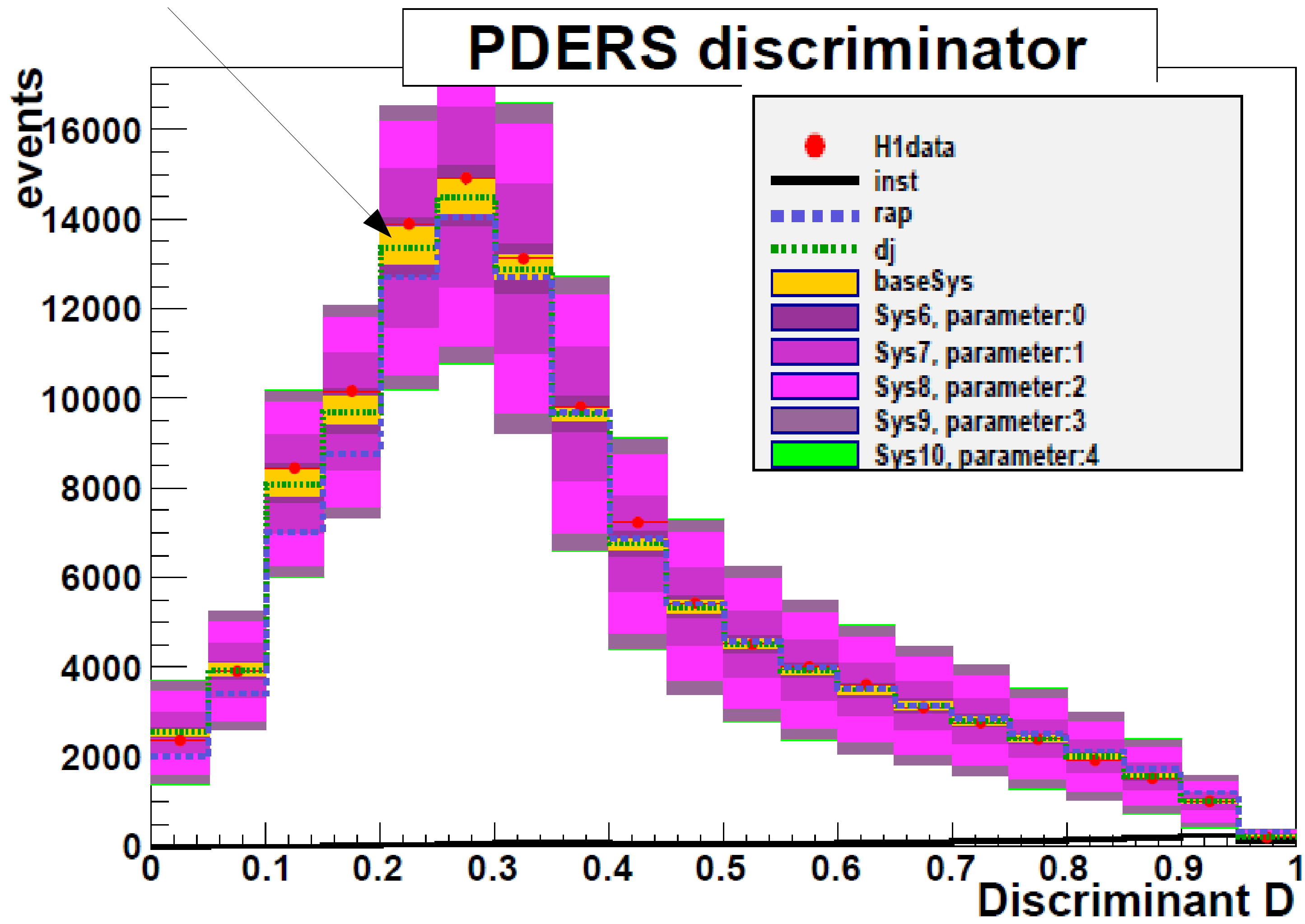
Model uncertainty: difference between background MCs

Uknown errors correlations between bins

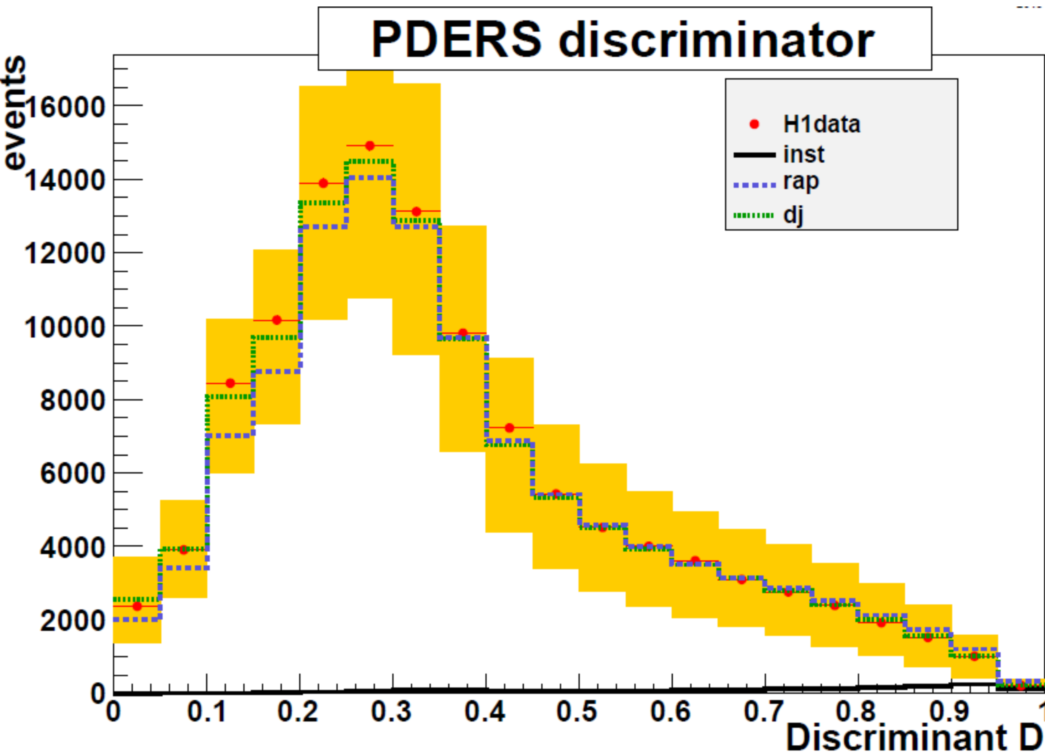
Suggestion: vary parameters of the number of charged particles reweighting function
(new source of systematic error: FitPar)

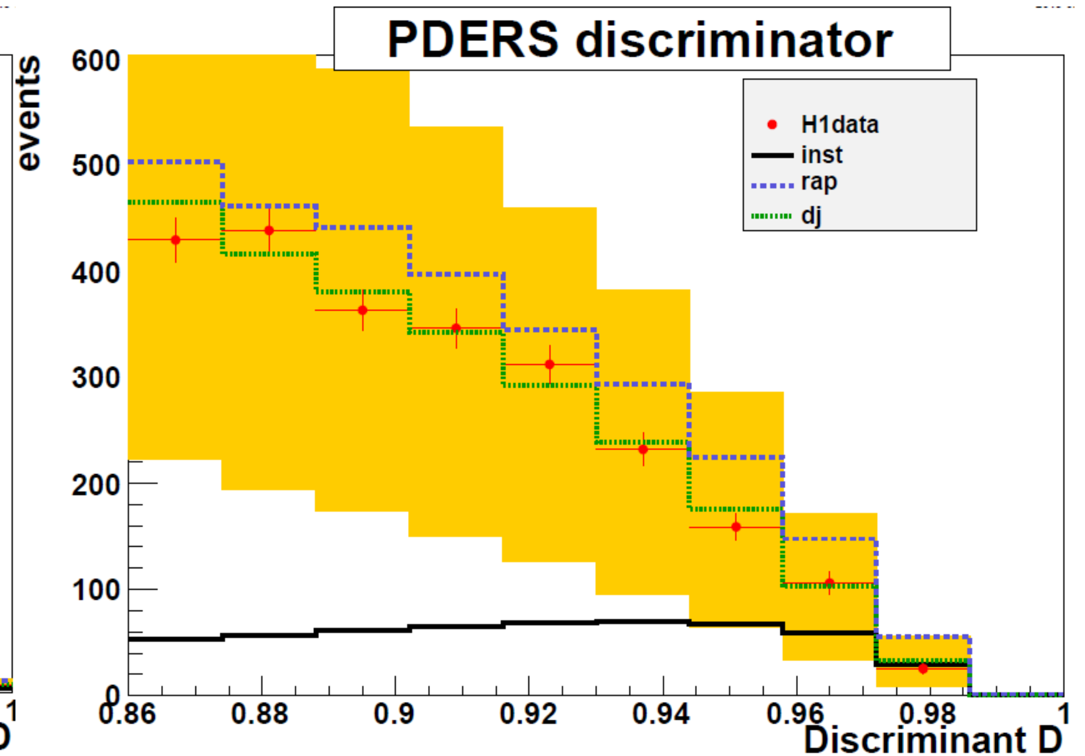
FitPar impact on PDERS

Orange bands represent all systematics for djangoh
except FitPar model uncertainty



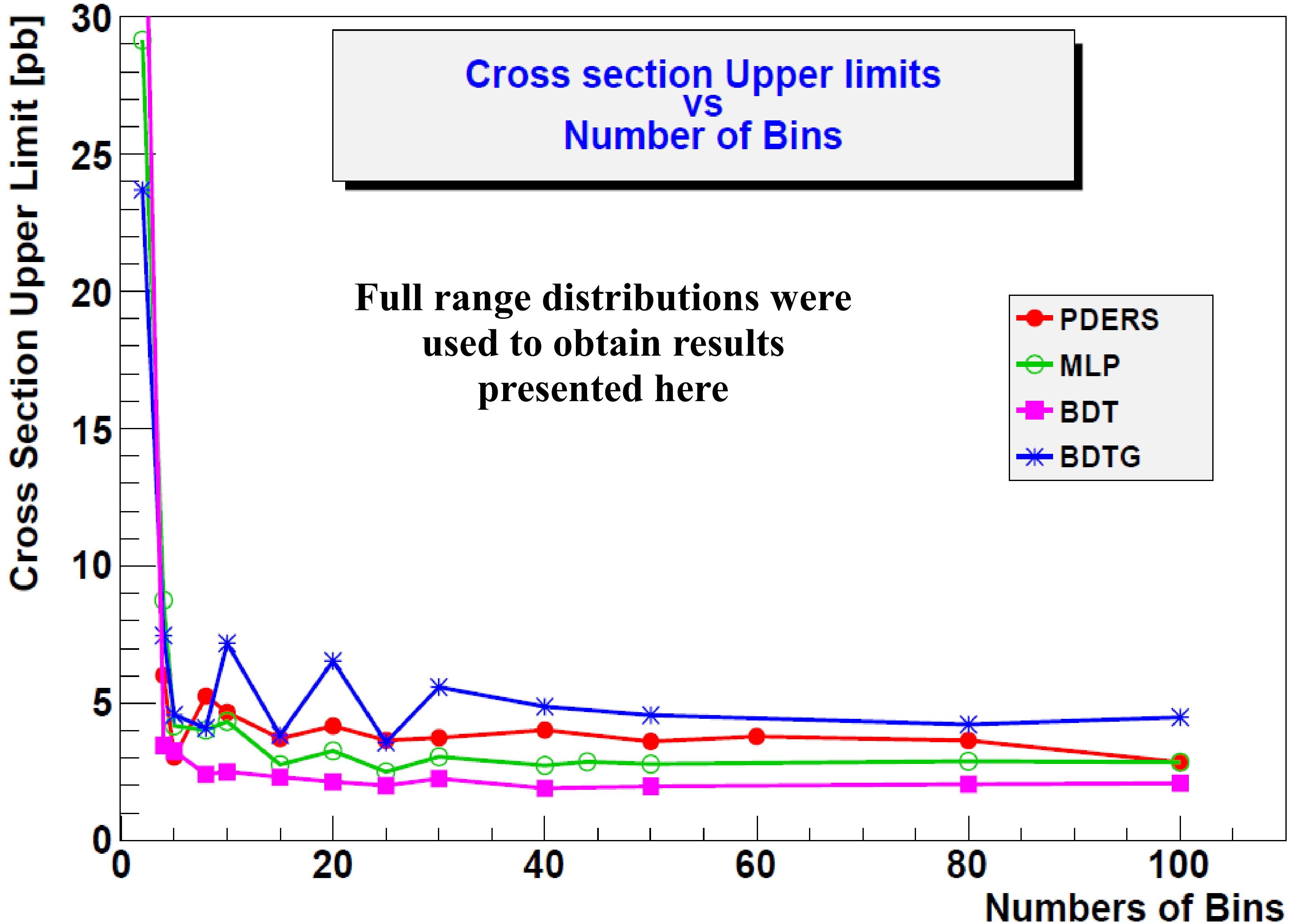
FitPar impact on PDERS





Orange bands represent all systematics for djangoh with FitPar model uncertainty

FitPar impact on upper limits
Number of bins scan



Upper limits: one more issue

Number of tracks reweighting function parameters: are they correlated?

Correlation matrix from fit procedure: some off-diagonal elements close to 1

Files with new variations not yet finished

Summary

Variations of the reweighting function parameters results in huge errors

Upper limits calculations are stable for higher number of bins for all methods used

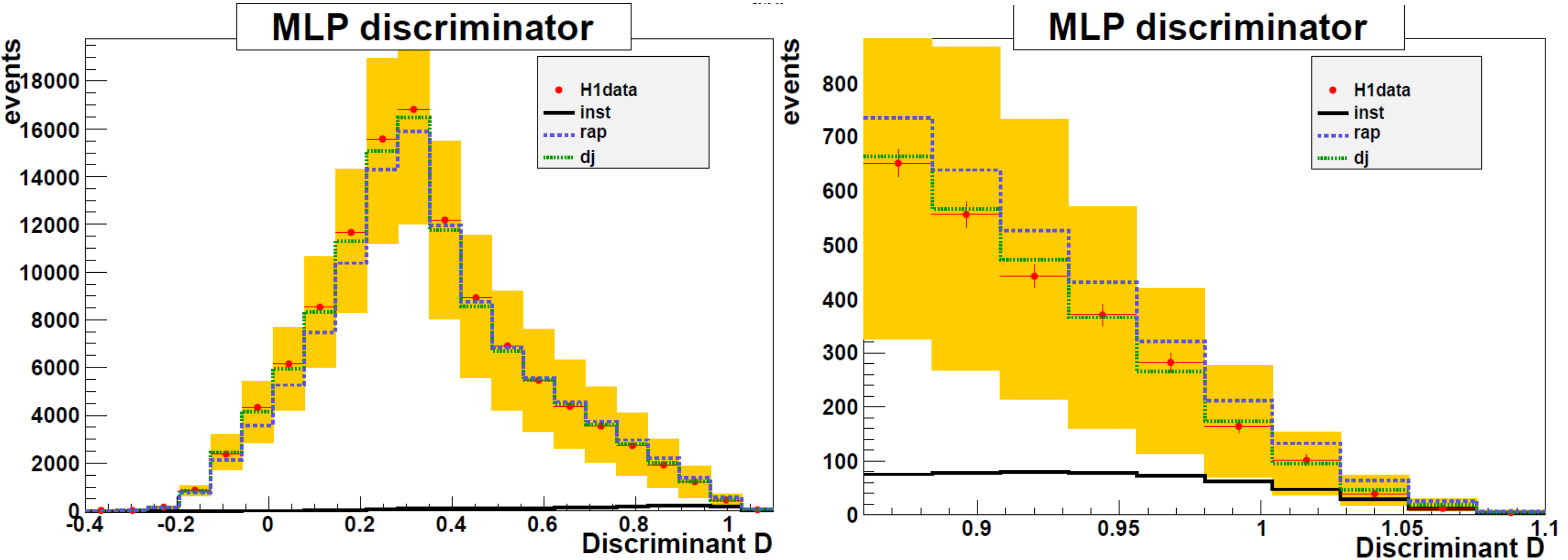
Presented results suggest exclusion of the instanton cross section predicted by Ringwald-Schremmp

(QCDINS MC generator)

Thank you for your attention

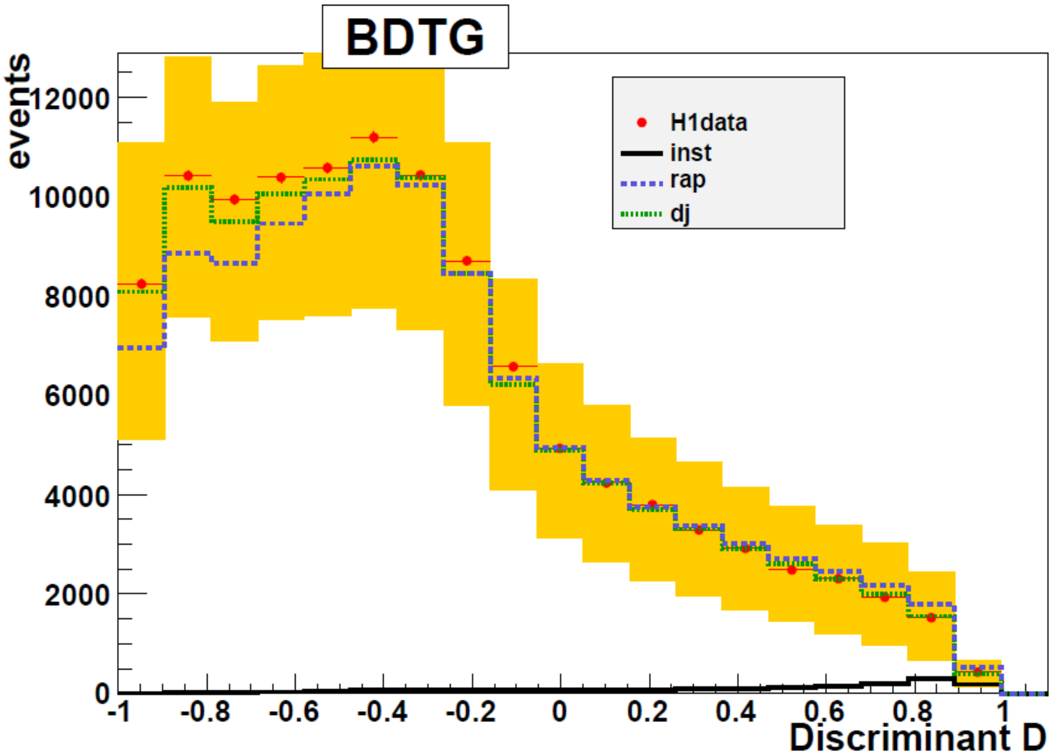
Backup

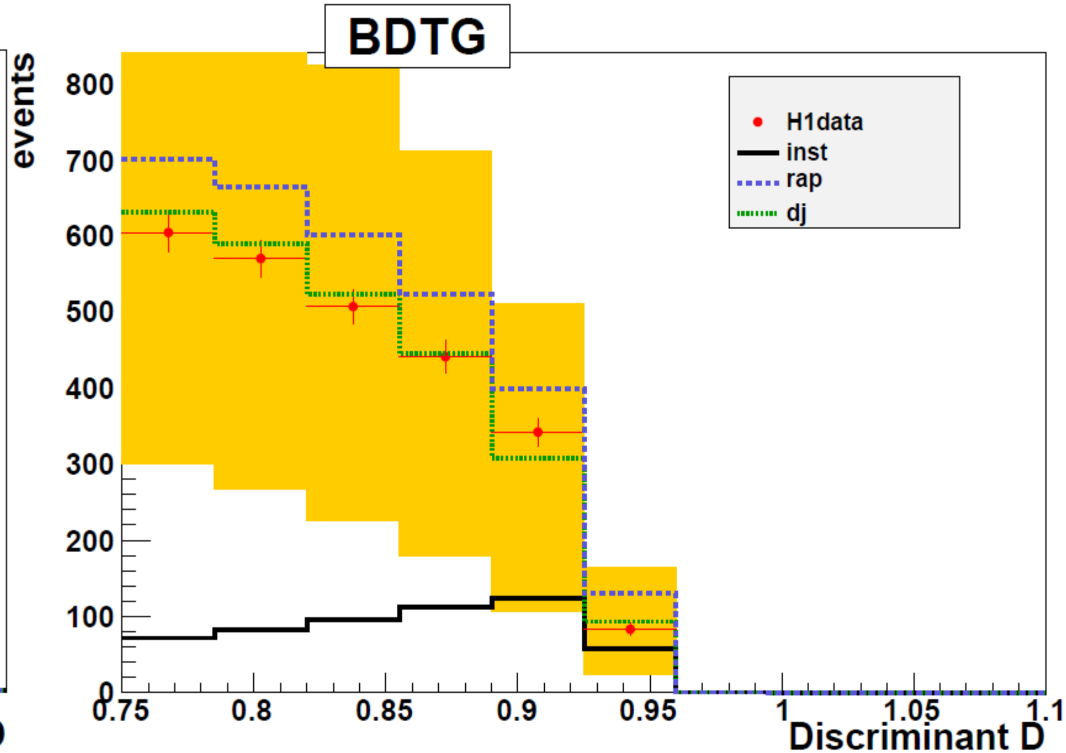
FitPar impact on MLP



Orange bands represent all systematics for djangoh with FitPar model uncertainty

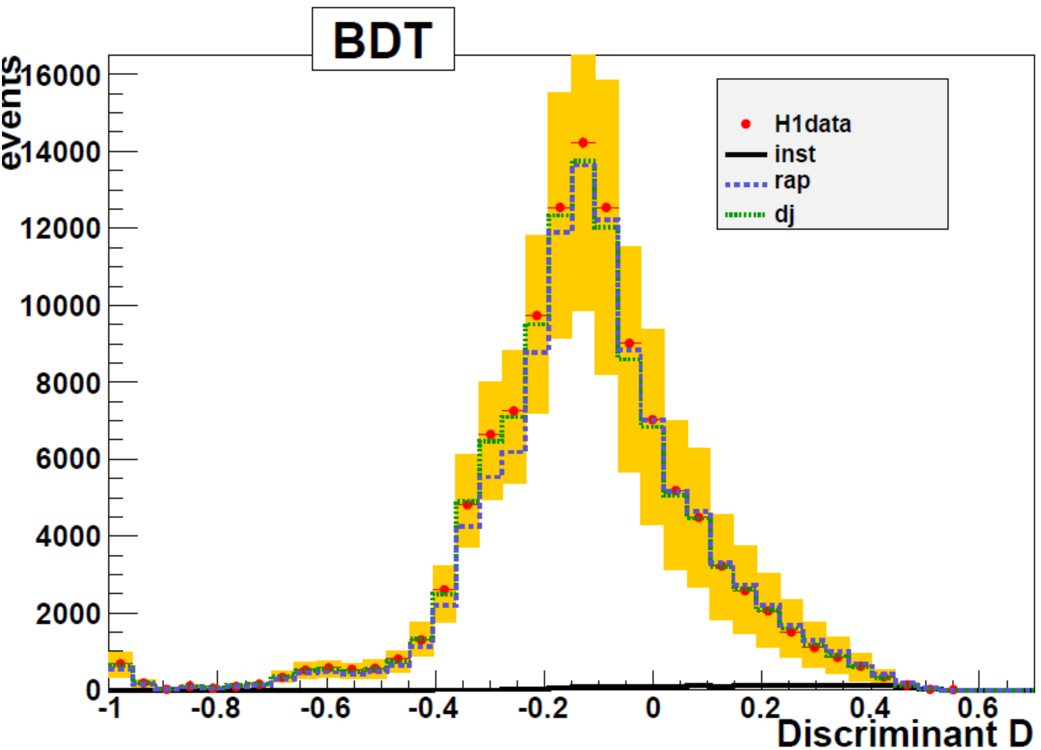
FitPar impact on BDTG

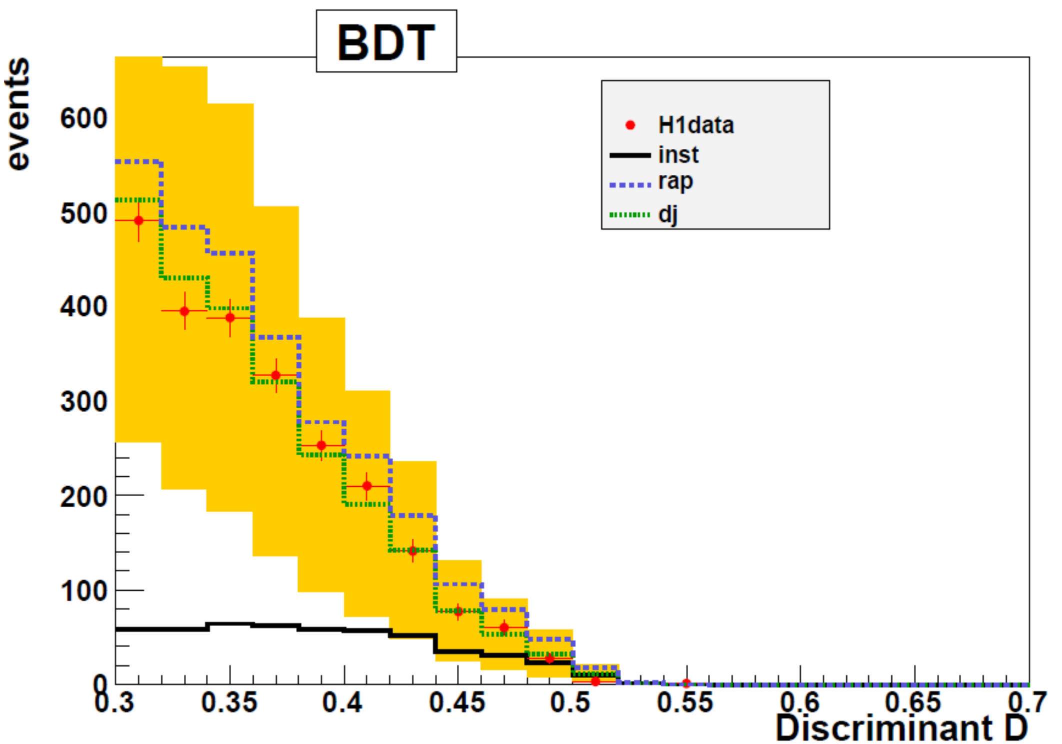




Orange bands represent all systematics for djangoh with FitPar model uncertainty

FitPar impact on BDT





Orange bands represent all systematics for djangoh with FitPar model uncertainty

Event Selection

In such region
QCD Instanton cross
section reduced:
$\sigma \approx 10$pb

DIS selection

Kinematics:

eΣ method

$150 < Q2 < 15000$ GeV

$0.2 < y < 0.7$

Electron Variables

$E_e > 11$ GeV

-190 cm $< Z_{imp} < 15$ cm && $Z_{imp} > 25$

$2° < \varphi \bmod [45°] < 43°$

Technical Cuts & Background

Fiducial Cuts,

Trigger 67

45 GeV $<$ E- $P_z <$ 65 GeV

$|Z_{vrtx}| < 35$ cm

Optimal Vertex with no **CIP-only** vertice

Track-cluster distance 8cm

Background Finders

Jet & Track Selections

Jet Selection

Inclusive kT algorithm in HCMS frame

$Pt > 3$ GeV

$R = 1.35*0.5$

(corresponds to $R_{cone} = 0.5$)

Jets boosted to LAB:

$Pt_{Jet} > 2.5$

$-1 < \eta_{jet} < 2.5$

Track Selection

$Pt > 0.12$

$20^o < \theta$

from Primary Vertex

$|DCA| < 2$ cm

$R_{length} > 10$ cm (for $\theta < 150^o$)

$R_{length} > 5$ cm (for $\theta < 150^o$)

+

must be a Central Track

Current Status

Currently used reweights:

Taken from HighQ2 Jets Analysis (Roman) and added on top of it

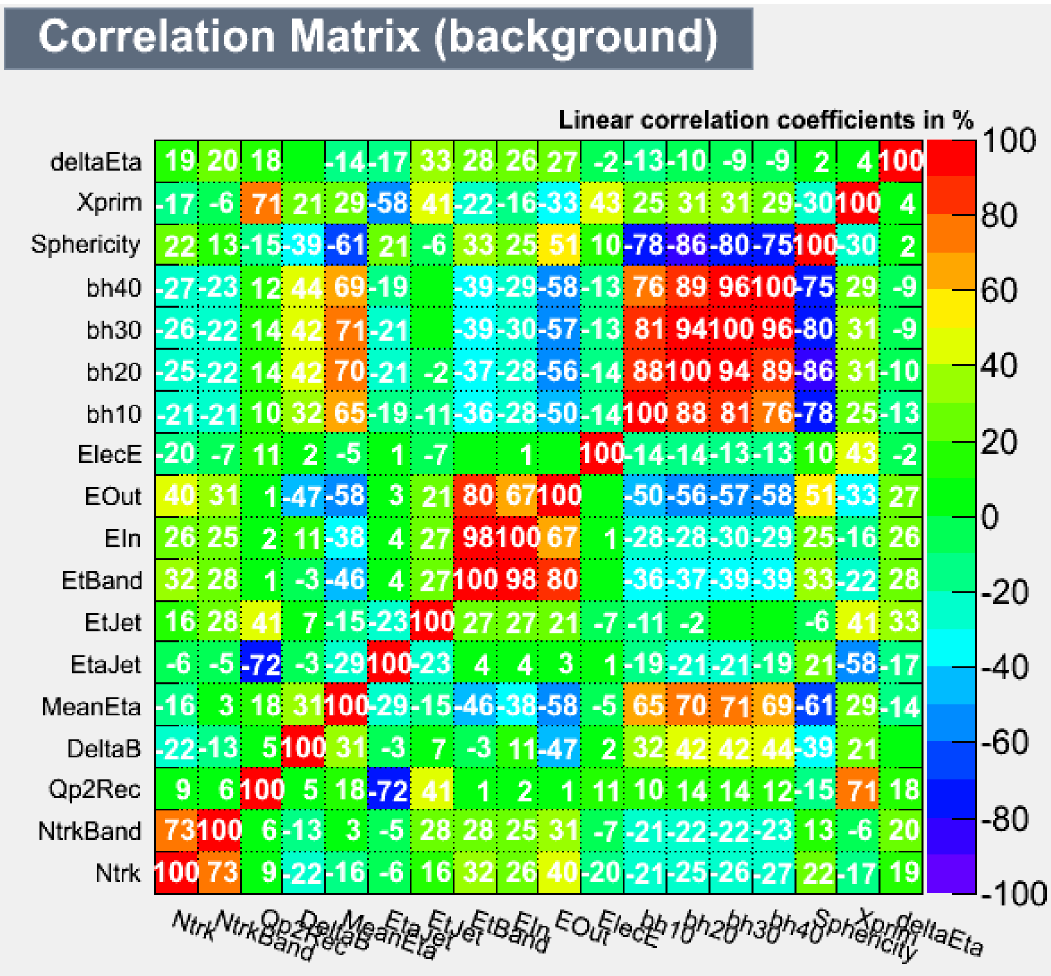
- Gen level: jet multiplicity, jet eta, jet Pt, E-pz of the HFS, MC cross section, jet eta in one-jet and coplanarity of one-jet events + number of tracks

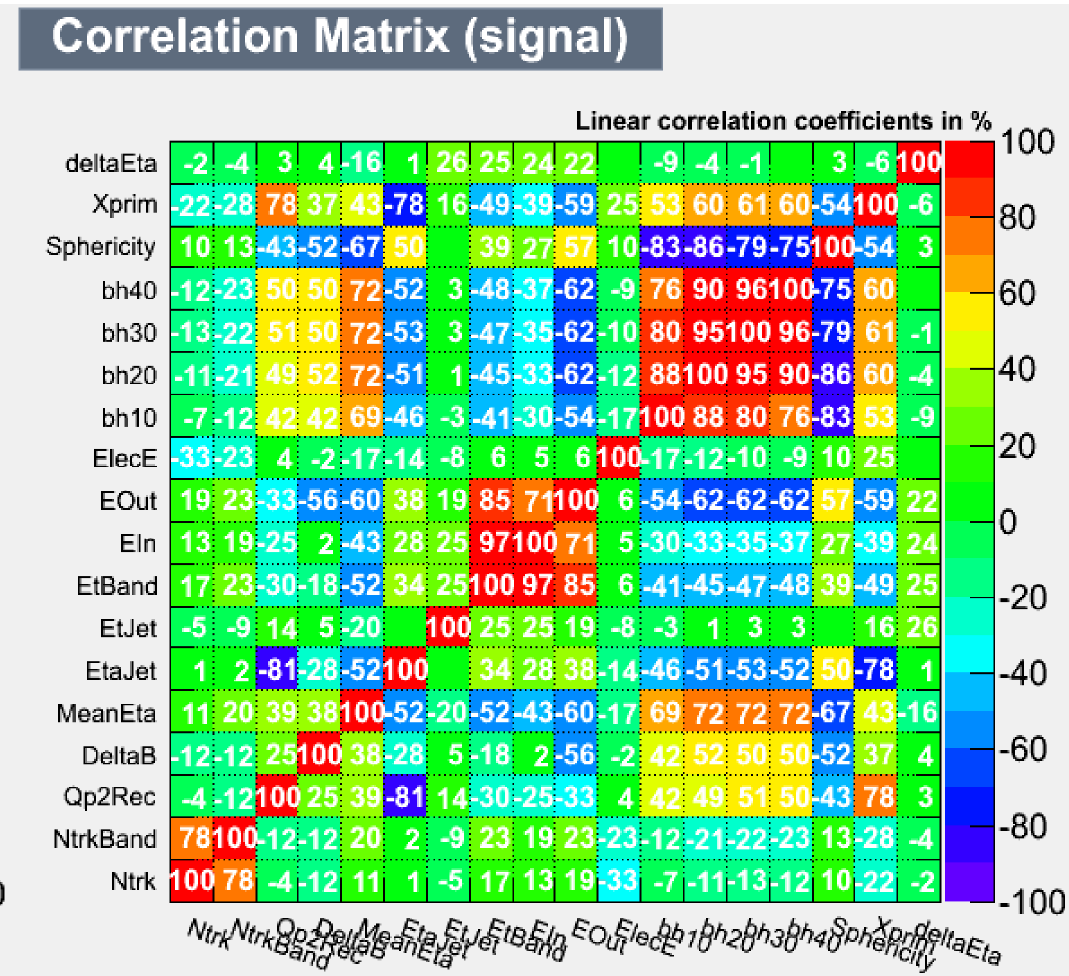
- Rec level: veto inefficiency, trigger (E_e<13GeV), track-cluster link effic. with a 8cm cut

Lumi absolute normalisation after reweights was off for Rapgap (~1%) and Djangoh (~10%). Luminosities used: L_{DATA} = 357.6, L_{Rapgap} = 9133.8, $L_{Djangoh}$ = 10488.15 [pb-1]

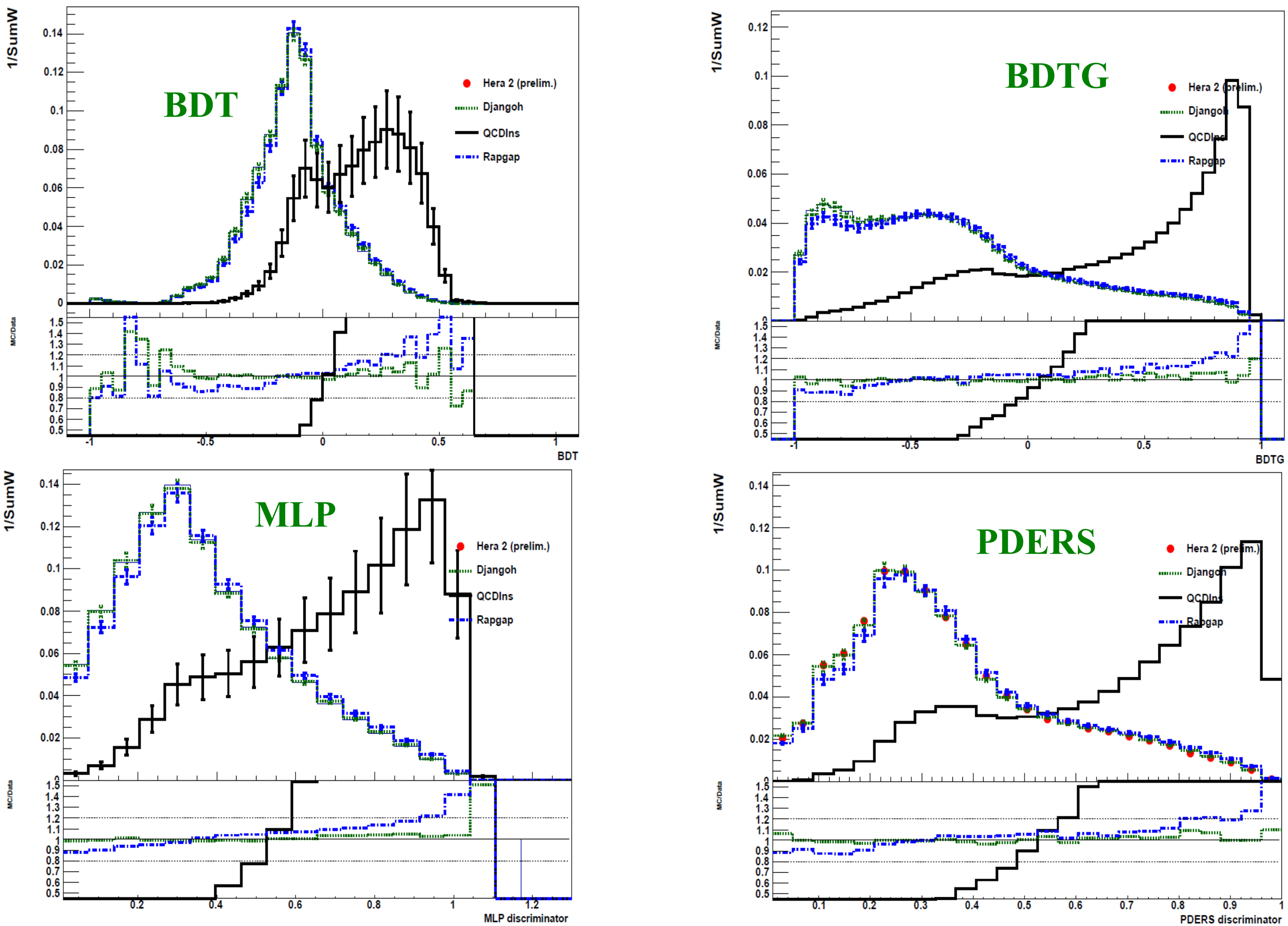
Expected Instanton signal in Data is ~0.8%

Linear Correlation Matrices





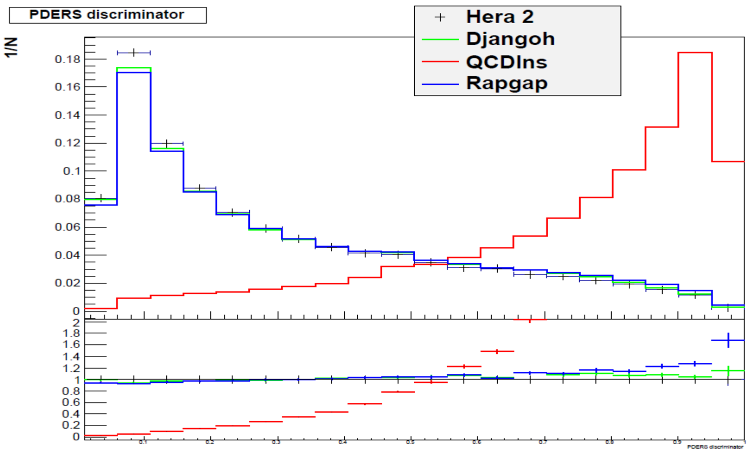
Base set – shape norm.



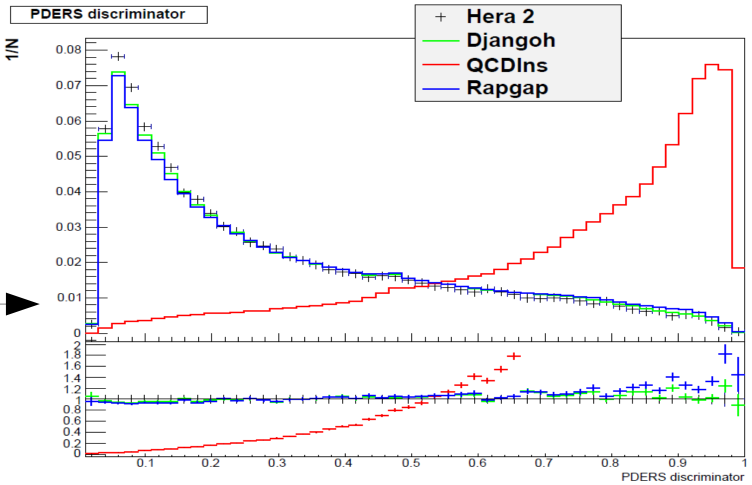
Multi Variate Analysis

Previous HaQ:

'Reminder': PDERS discriminator



Smaller binning

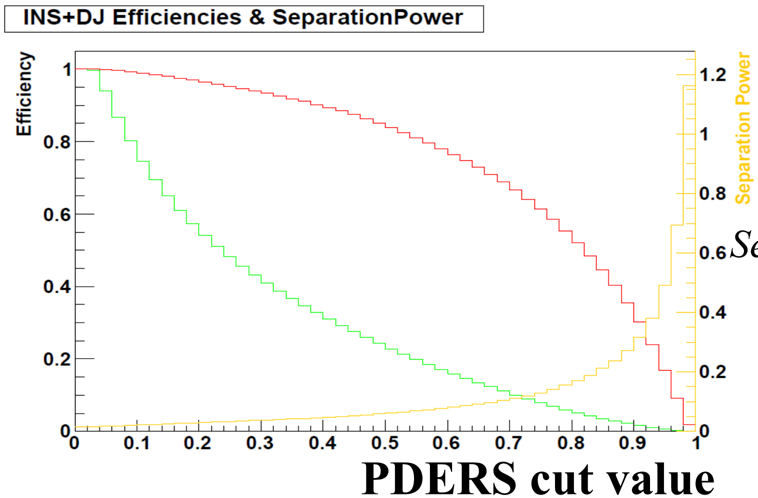


Where to cut on discriminator?

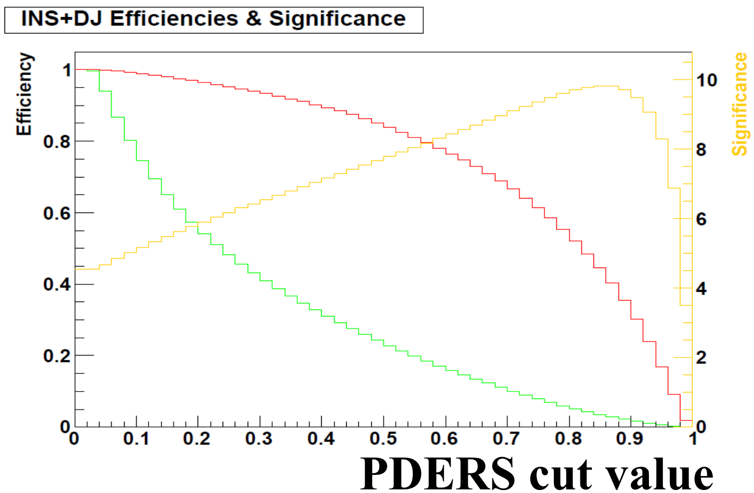
One can compare signal and background efficiencies with Separation Power defined as a ratio of efficiencies

Standard approach uses Significance

$$Significance = \frac{S}{\sqrt{S+B}}$$



$$SepPow = \frac{Sig_{Eff}}{Bkg_{Eff}}$$

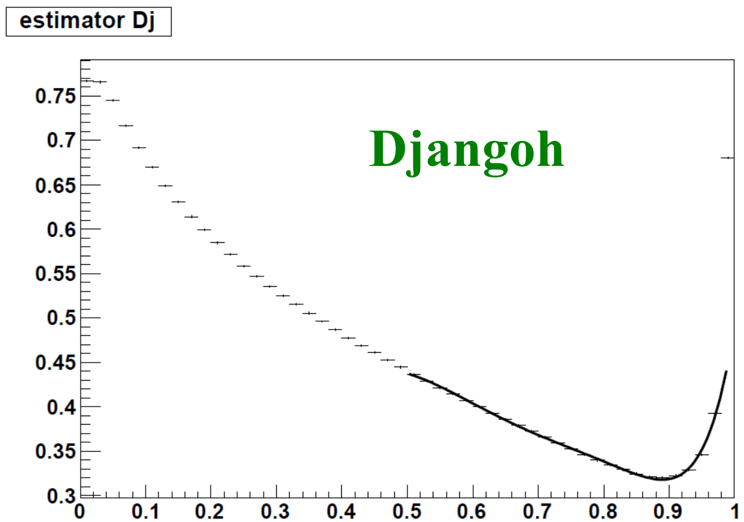


Multi Variate Analysis

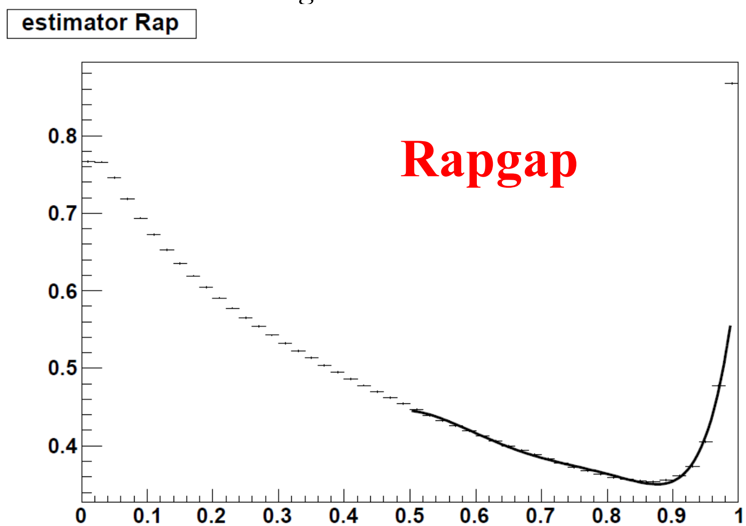
Where to cut on discriminator?

One can use a method (based on Stefan Schmitt's idea) that estimates an error of signal cross section:

$$err(\sigma_{sig}) = \frac{\sqrt{N_{bckg} + \Delta^2_{model}}}{\epsilon * L_{data}} \longrightarrow \frac{\sqrt{N^{cut}_{Bckg} * N^{total}_{Sig}}}{N^{cut}_{Sig} * L_{data}}$$



PDERS cut value



PDERS cut value

Fitted functions' (polynomial) minimums:

minimum in x= 0.889

minimum in x= 0.876

A cut value used so far was 0.9 and the same for all sets&MCs.

Charged particles mutliplicty reweighting

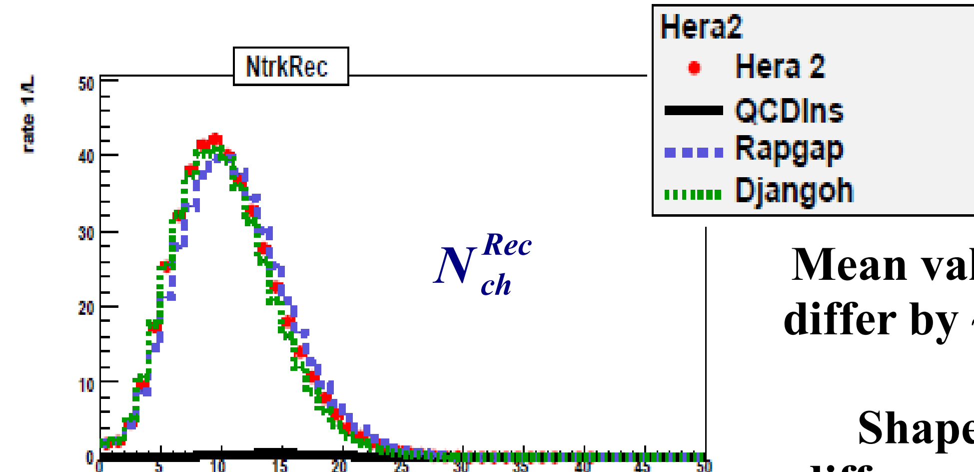
Procedure:

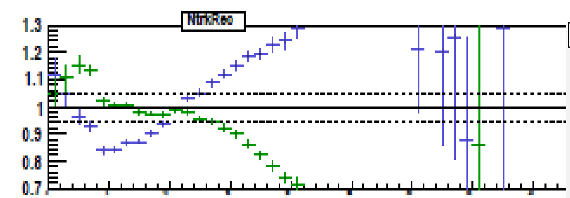
Take number of tracks distributions on Jet level (N_{ch}^{Rec})

Fit polynomial function f_i to Data/MC$_i$ ratio

Calculate new weights using $f_i(N_{ch}^{Gen})$

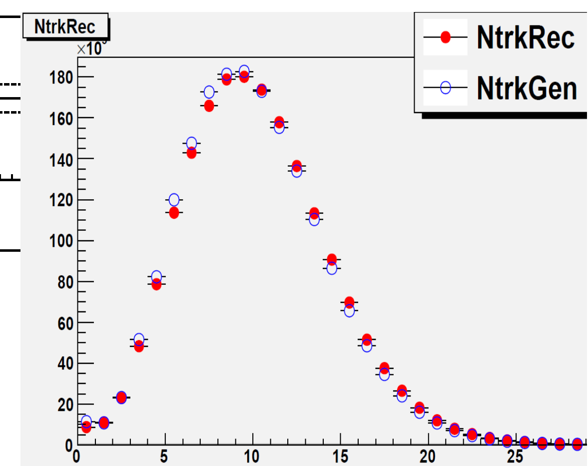
N_{ch}^{Gen} - number of stable, chraged particles from GTR bank

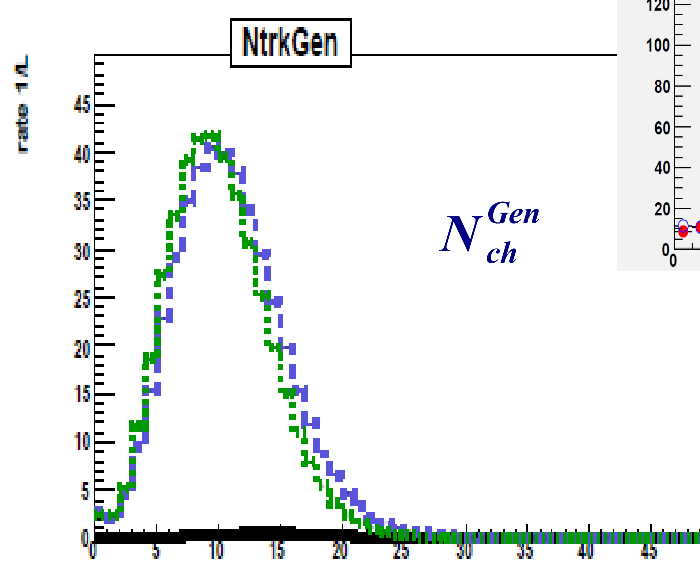




Mean values differ by ~0.2

Shape difference:





New Observable x'_{MEAN}

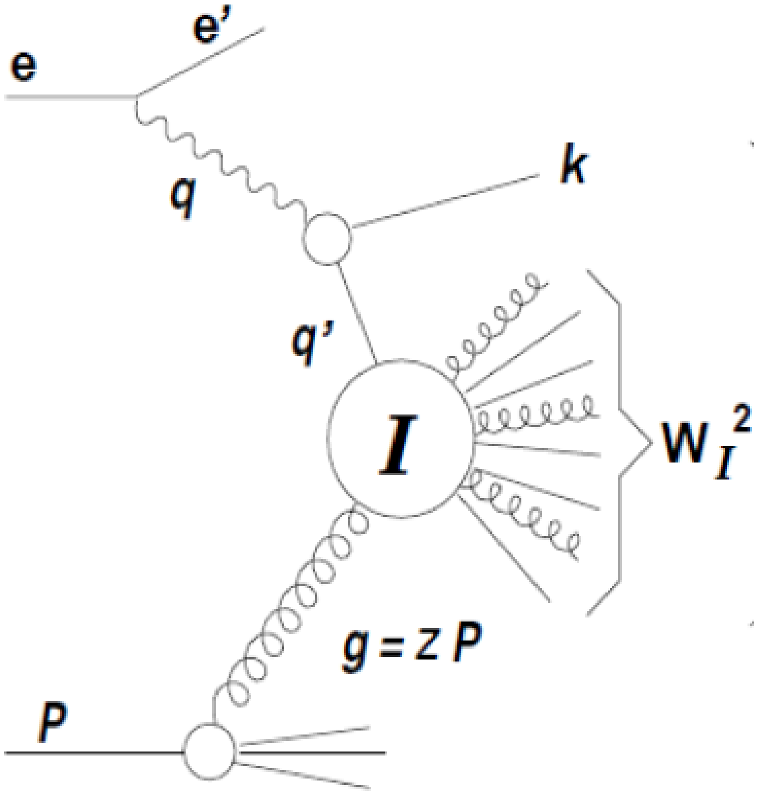
Instanton-subprocess variables:

$$Q'^2 = -q'^2 = -(q-k)^2$$

$$x' = Q'^2 / (2g \cdot q')$$

$$W_I^2 = (q'+g)^2 = Q'^2 (1/x' - 1)$$

$$x' = \frac{Q'^2}{(W_I^2 + Q'^2)}$$



W^2_I and x' "Reconstruction":

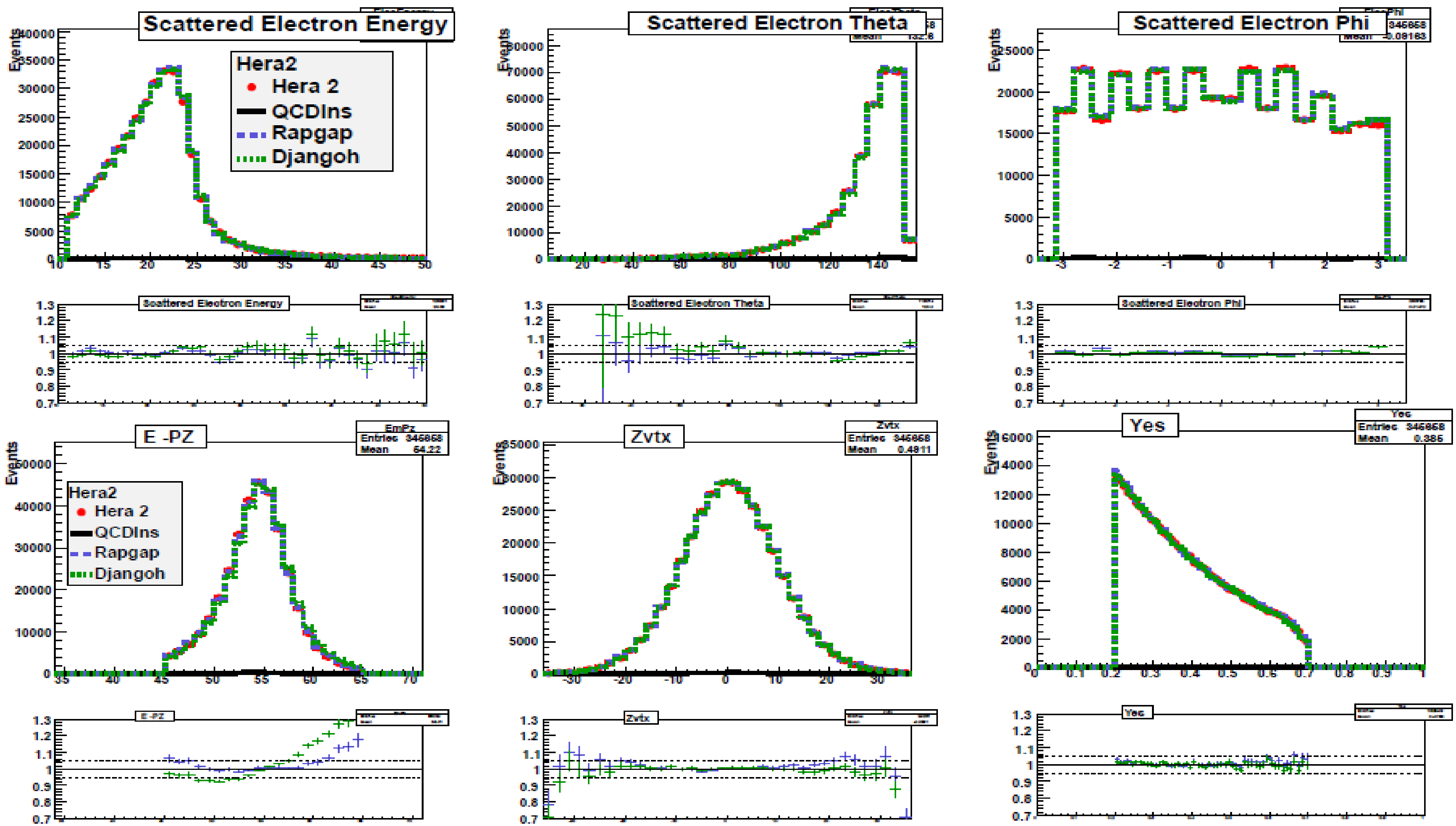
$$W_I^2(Kin) = (q' + zP)^2 \rightarrow (q'_{rec} + \bar{z} P)^2 \rightarrow x'_{Kin}$$

$$W_I^2(Band) = \left(\sum_i v_i\right)_I^2 \rightarrow \left(\sum_i v_i^{rec}\right)_{Band}^2 \rightarrow x'_{Band}$$

$$x'_{MEAN} = \frac{x'_{Kin} + x'_{Band}}{2}$$

Control plots on DIS level (Roman+N_{trk})+Yes rew.

With the reweight on YeΣ both Monte Carlos have a very good data description – Yes 'slope' in Rapgap fixed



Control plots on Instanton level (Roman+N_{trk})+Yes rew.

Absolute normalisation moved down by ~1-2% with respect to DIS level but data description is still very good