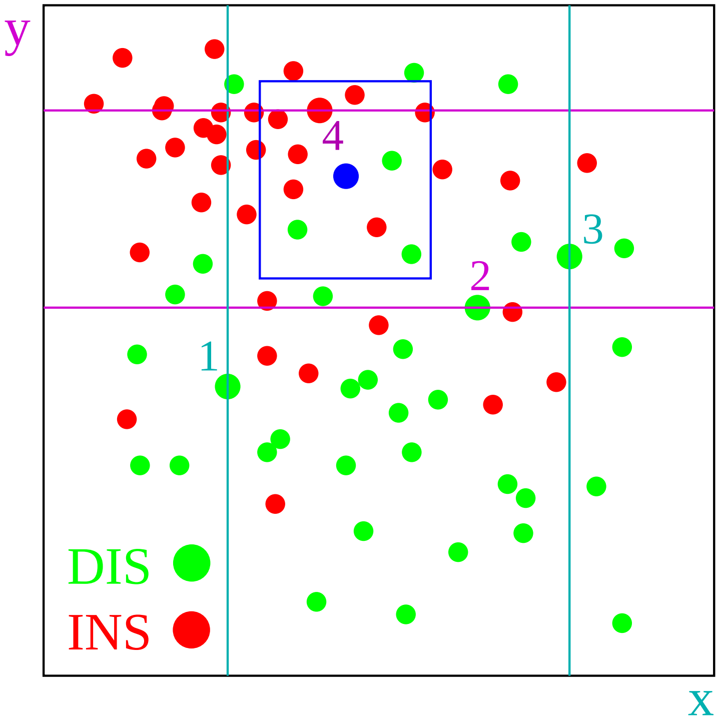
Looking for Instantons using a Multivariate Technique based on Range Searching

B. Koblitz, T. Carli (MPI Munich and University of Hamburg)



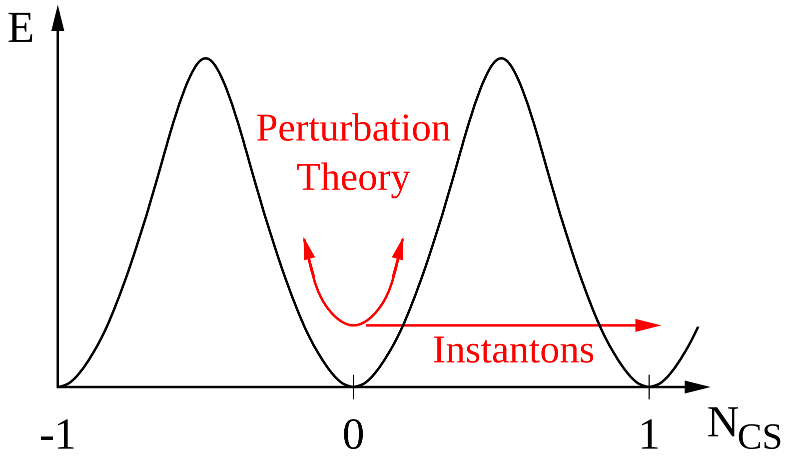
Overview:

- Introducing the Problem: Instantons
- The Method: Range Searching
- Results
- Properties of the Search Method: Comparing with Neural Nets
- Conclusions

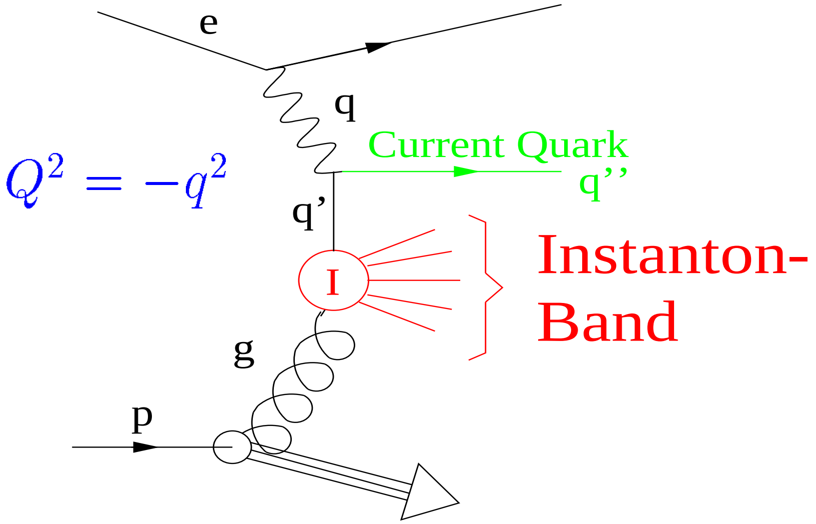
What are Instantons?

Instantons are:

- Tunneling transitions between topologically different ground states
- Quantum number violating (QCD: Chirality, QFD: $B+L$)
- Exponentially suppressed: $\sim e^{-\frac{2\pi}{\alpha}}$
- Part of the Standard Model



Instantons in DIS at HERA:



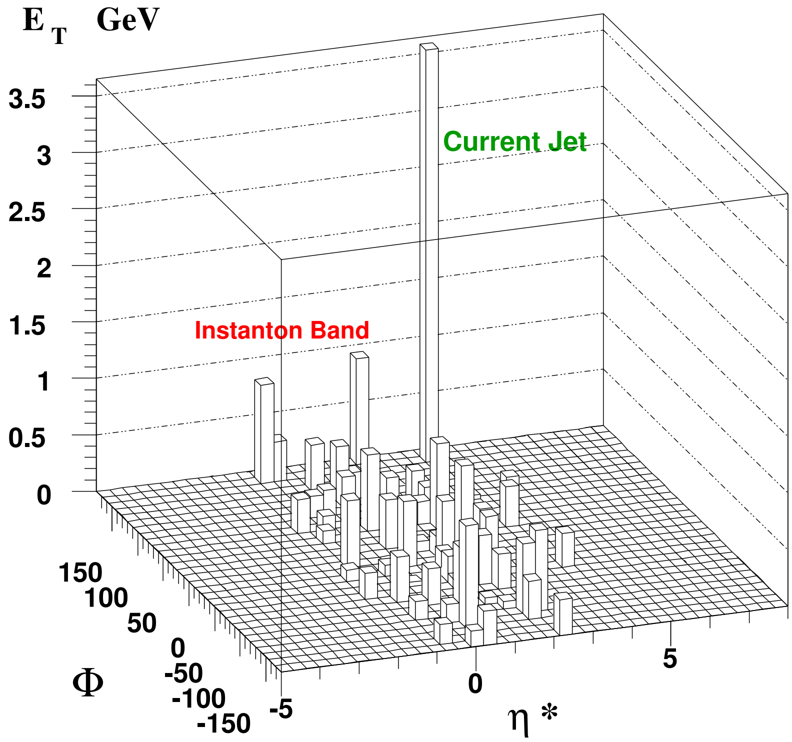
Virtuallity of ingoing quark:

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

An experimental discovery would be a novel manifestation of the non-perturbative properties of QCD.

Properties of Instanton Induced Events

E_T-Map of MC-Event QCDINS [A. Ringwald, F.Schrempp]:



Event properties:

- Hard jet ($\rightarrow Q'^2$)
- Densely populated band in η, with high E_T, isotropic in its rest-frame ($\rightarrow$ Shape variables)
- Enhanced strangeness production in band ($\rightarrow \#K^{\pm}, \#K^0$)

Predicted Cross Section sizeable: [A. Ringwald, F.Schrempp]

$\sigma^{(I)}_{\mathrm{HERA}} = 29.2^{+9.9}_{-8.1}\,\mathrm{pb} \qquad \sigma_{\mathrm{DIS}} \approx 3000\,\mathrm{pb}$

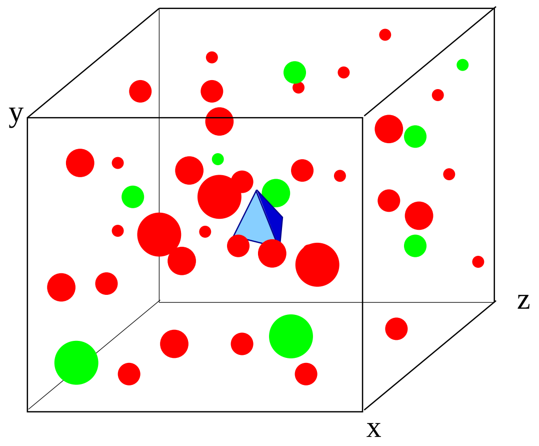
$(x > 10^{-3},\ 0.1 < y < 0.9,\ Q^2 > 113\,\mathrm{GeV}^2\ )$

$\Rightarrow$ Search is difficult: Many variables and need for high separation power!

Classification via Range Searching

Problem: Need a fast and automatic method to scan a large(35) number of variables for best ones (separation, systematics, theory)

Idea: Directly classify an event via counting MC-generated Instanton and DIS events in its vicinity.



$$\ell = \frac{\rho(\mathrm{I})}{\rho(\mathrm{DIS})} = \frac{\#I(V)}{\#DIS(V)}$$

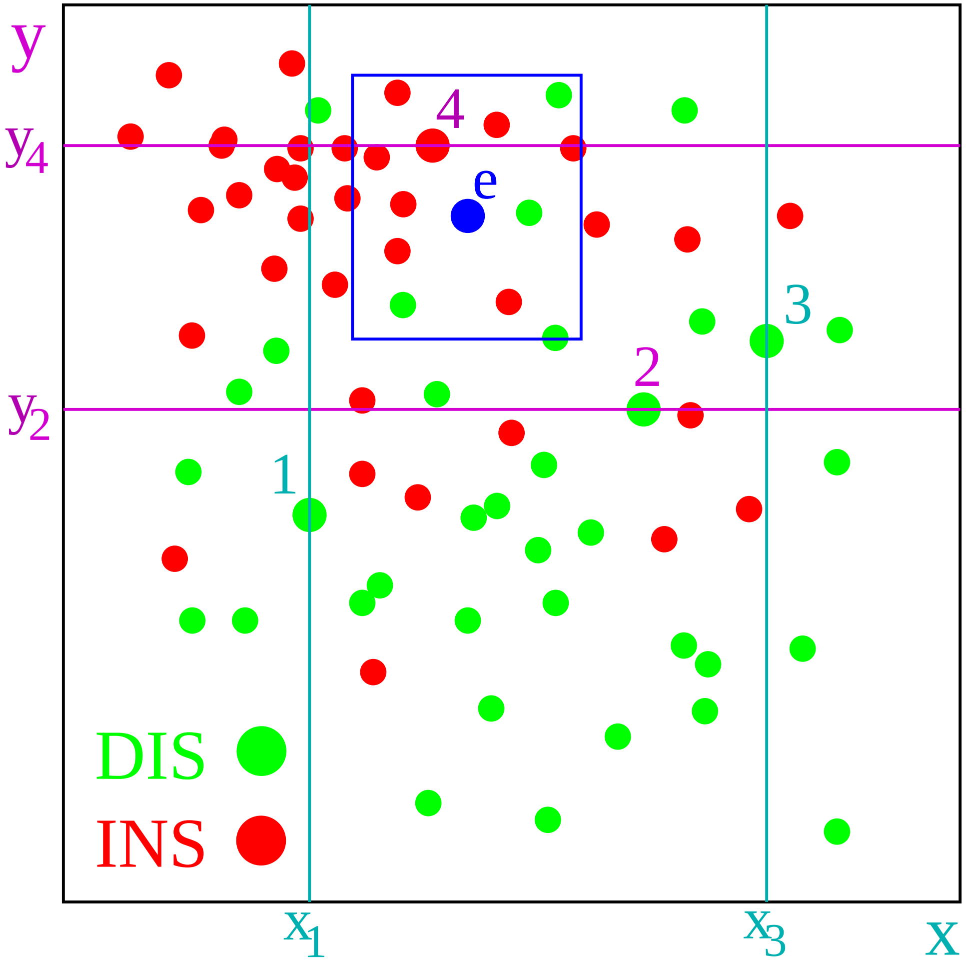
Discriminant: $D = p_{\mathrm{I}} = \frac{\ell}{1+\ell}$

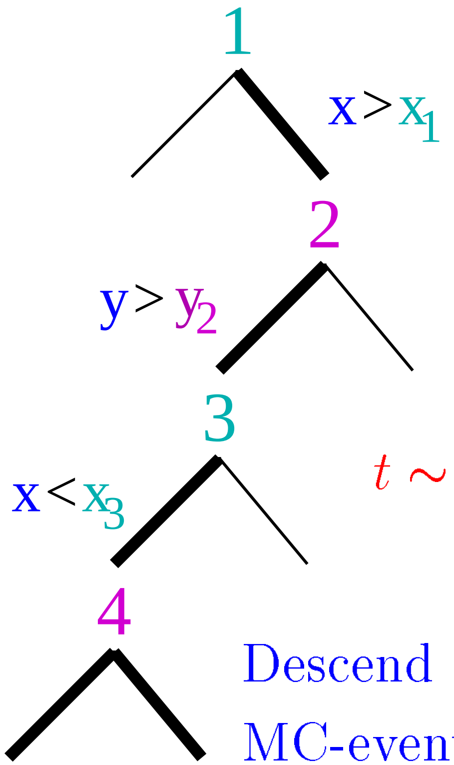
Problem now: How to find events in vicinity in reasonable time

$\Rightarrow$ Range Searching [Sedgewick: Algorithms]

Visualization of Range Searching

Example of Range Searching in a two dimensional phase space in a volume (blue) around an event to be classified $e(x, y)$.

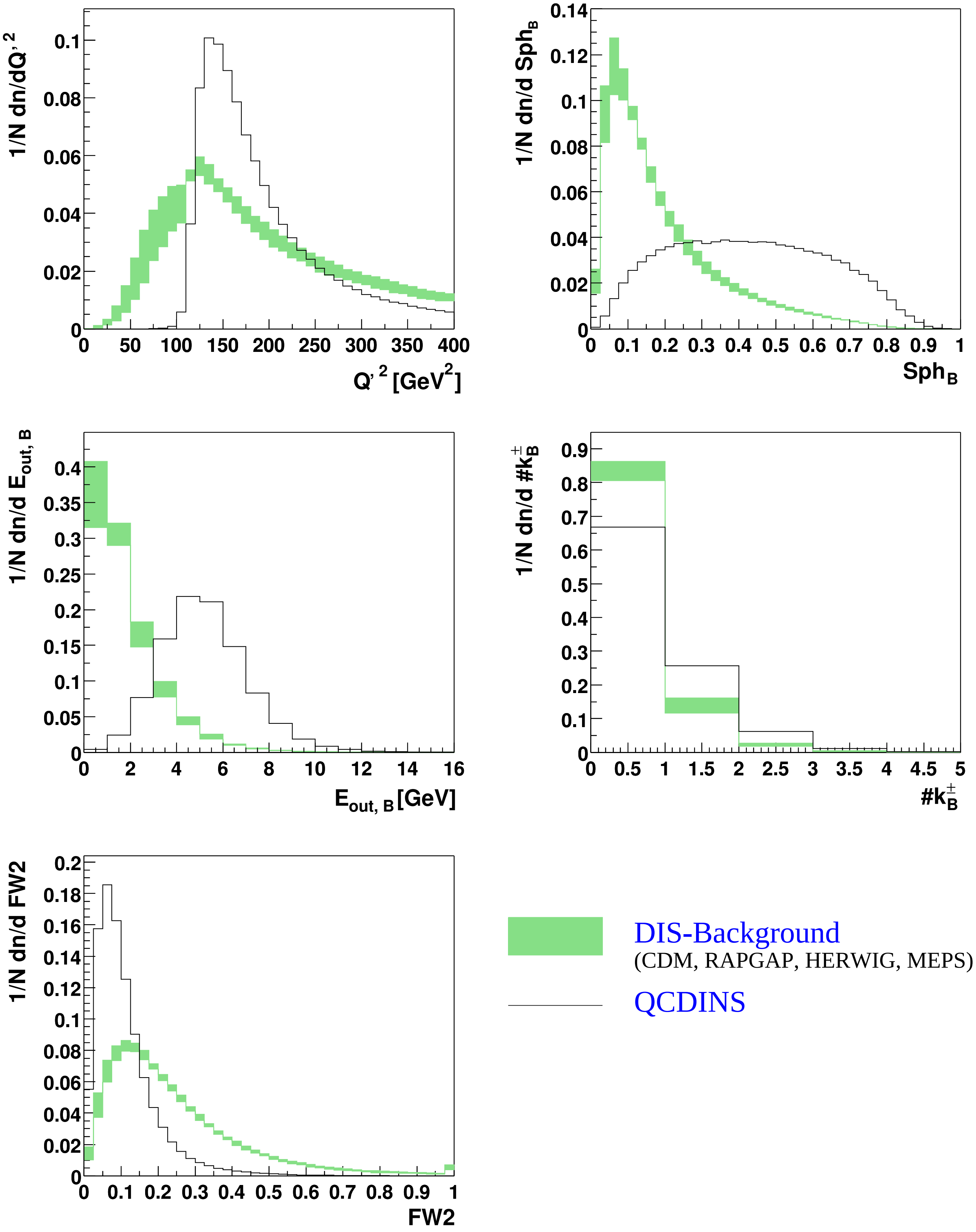




$t \sim \log(n)$

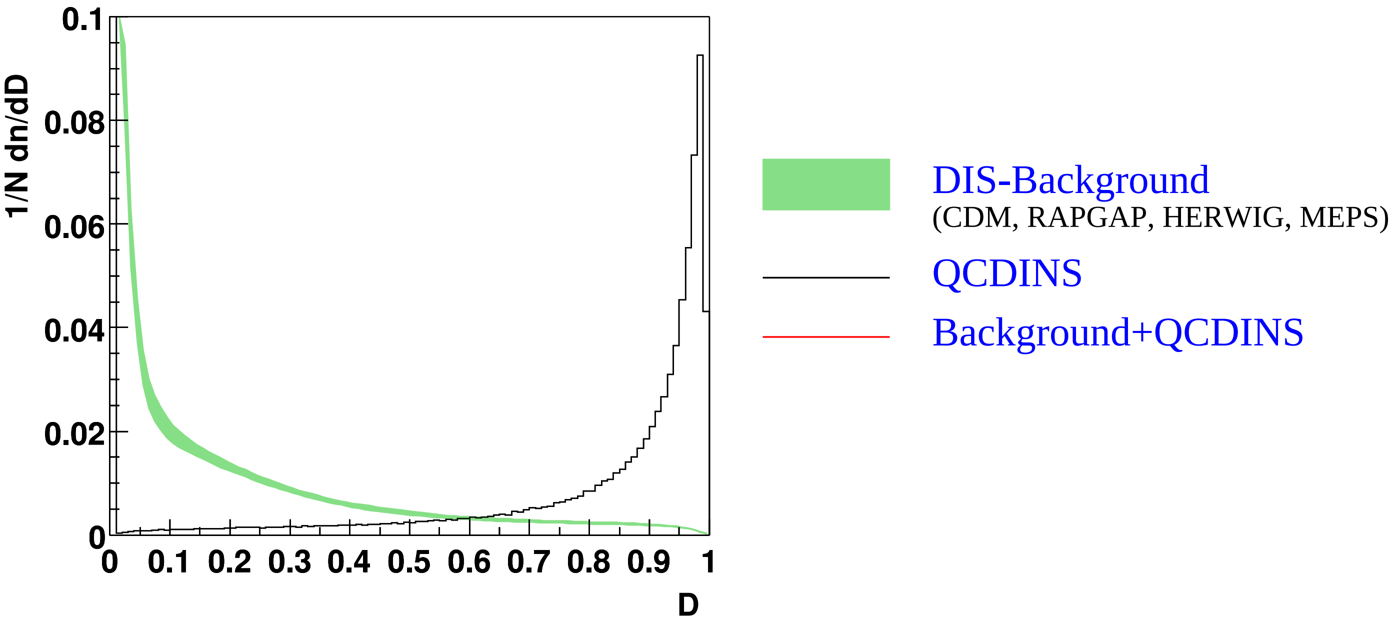
Descend in tree with MC-events according to the relative position of $e(x, y)$

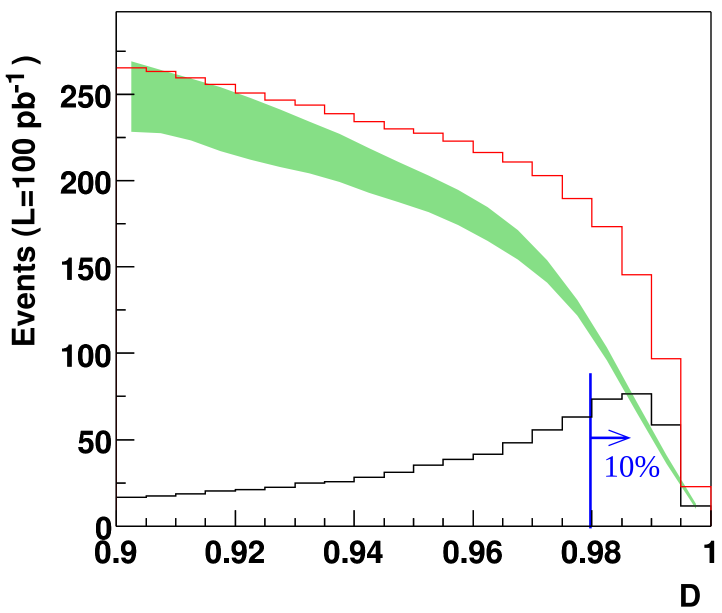
Event Variables



Results

Calculation with 300k events each in Instanton/DIS tree.





$\Rightarrow$ Separation Power $S = \frac{\epsilon_{\mathrm{INS}}}{\epsilon_{\mathrm{DIS}}} = 126$ @ $\epsilon_{\mathrm{INS}} = 10\%$

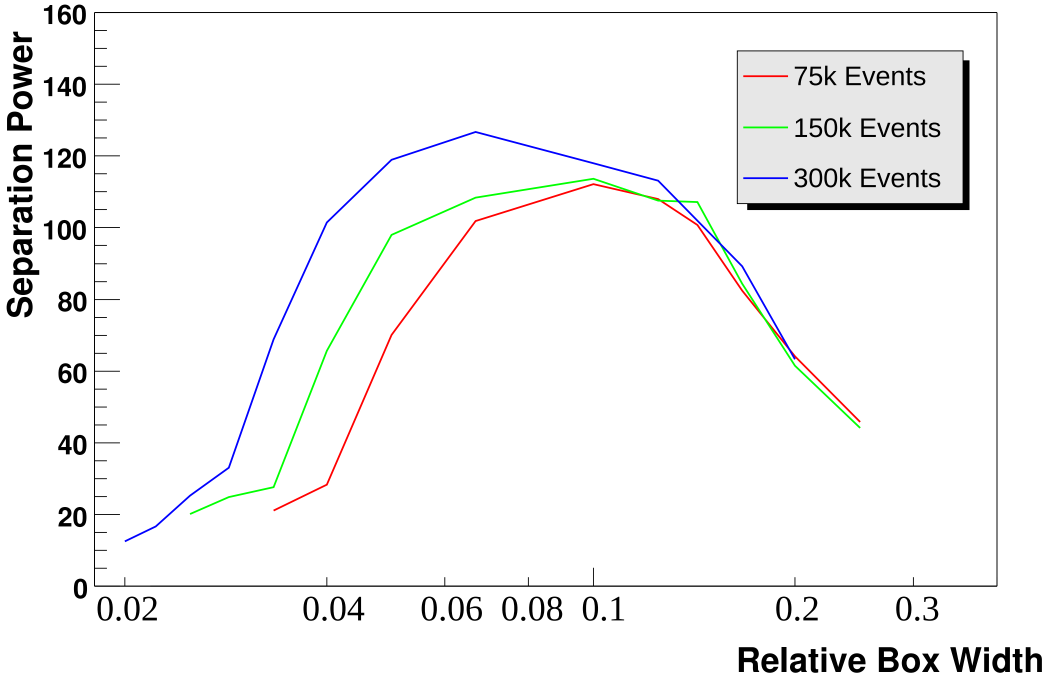
$\Rightarrow$50% Instantons expected in cut region

Properties of the Range Search Method

Parameters of Algorithm:

- Size of neighbourhood
- Number of events in tree
- Minimum events in neighbourhood

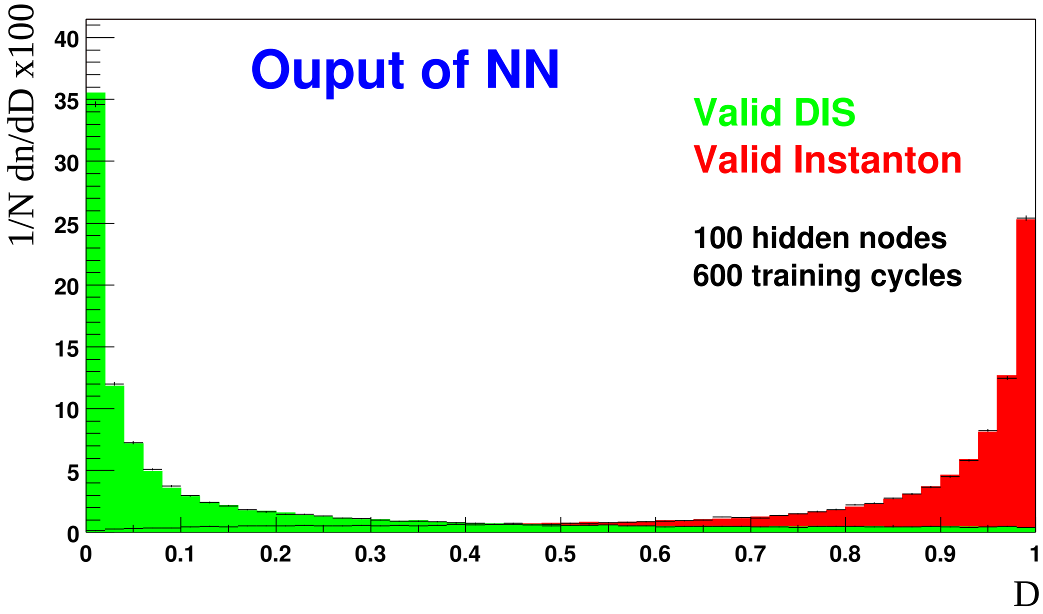
Fix relative scales of variables
⇒ 1 box width parameter

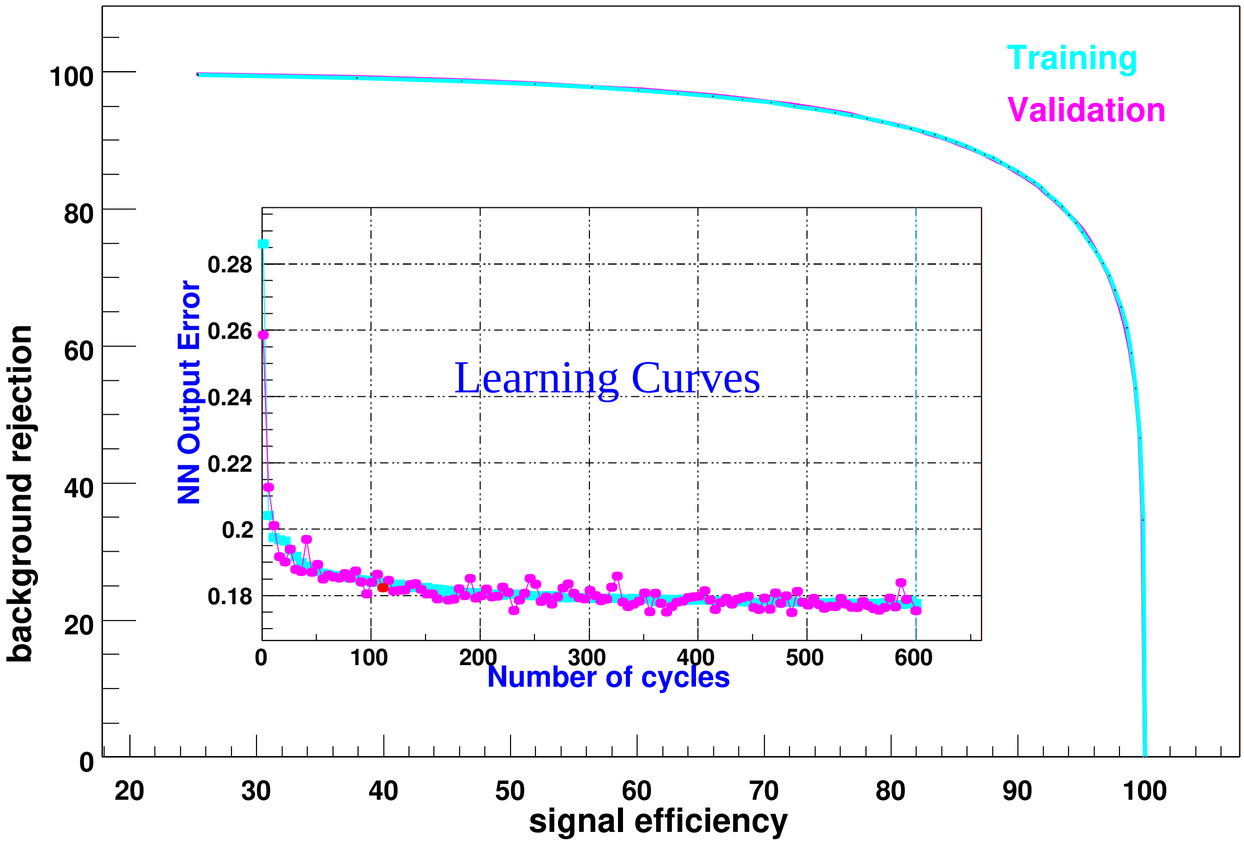


⇒ The more events, the better
⇒ Relatively stable with box width
⇒ No sensitivity for minimum #Events
Average time for calculation ≈ 20 min (Linux PC)

Comparison with a Neural Network

Compare with a single hidden layer feed forward network:





$\Rightarrow$ Separation Power $S = 116 \quad @\, \epsilon_{\mathrm{INS}} = 10\%$

Calculation took $\approx 4\,\mathrm{h}$.

Conclusions

- Searching for QCD-Instantons at HERA is hard...
- A Multivariate Method based on Range Searching allows fast scanning of variables and performs as good as NN's.
- The short processing time allows extensive studies of systematics and helps in finding well performing variables favoured by theorists/experimentalists.
- At a 10% Instanton efficiency the discrimination method provides an expected fraction of 50% Instantons in the data.