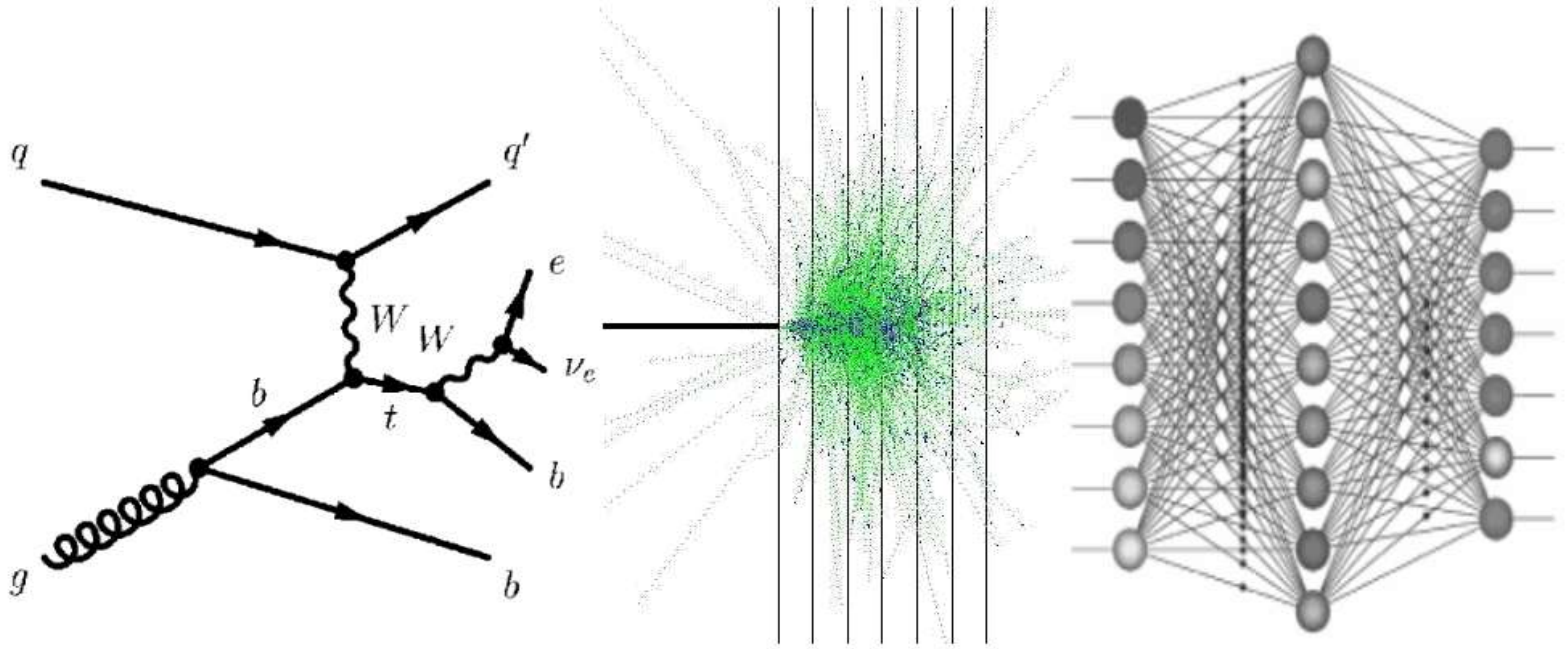
An Artificial Neural Network for Electron Identification with CDF

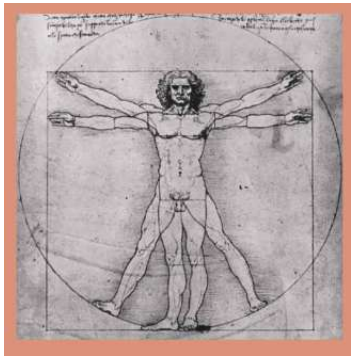


Yves Kemp, Thomas Müller,
Hartmut Stadie, Wolfgang Wagner
Universität Karlsruhe

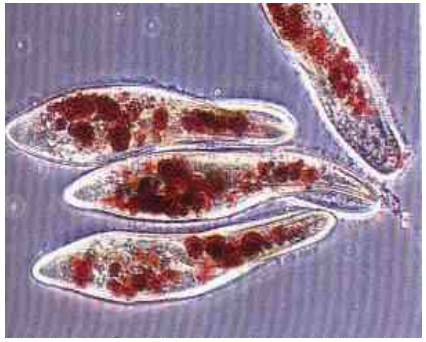
From Universe to Quark:

10^{26} m

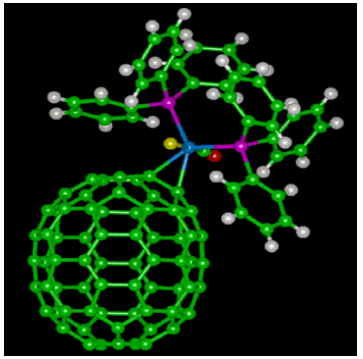
10^{7} m



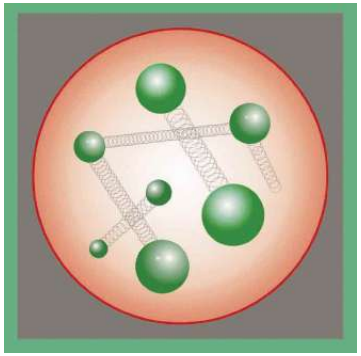
10^{1} m



10^{-4} m



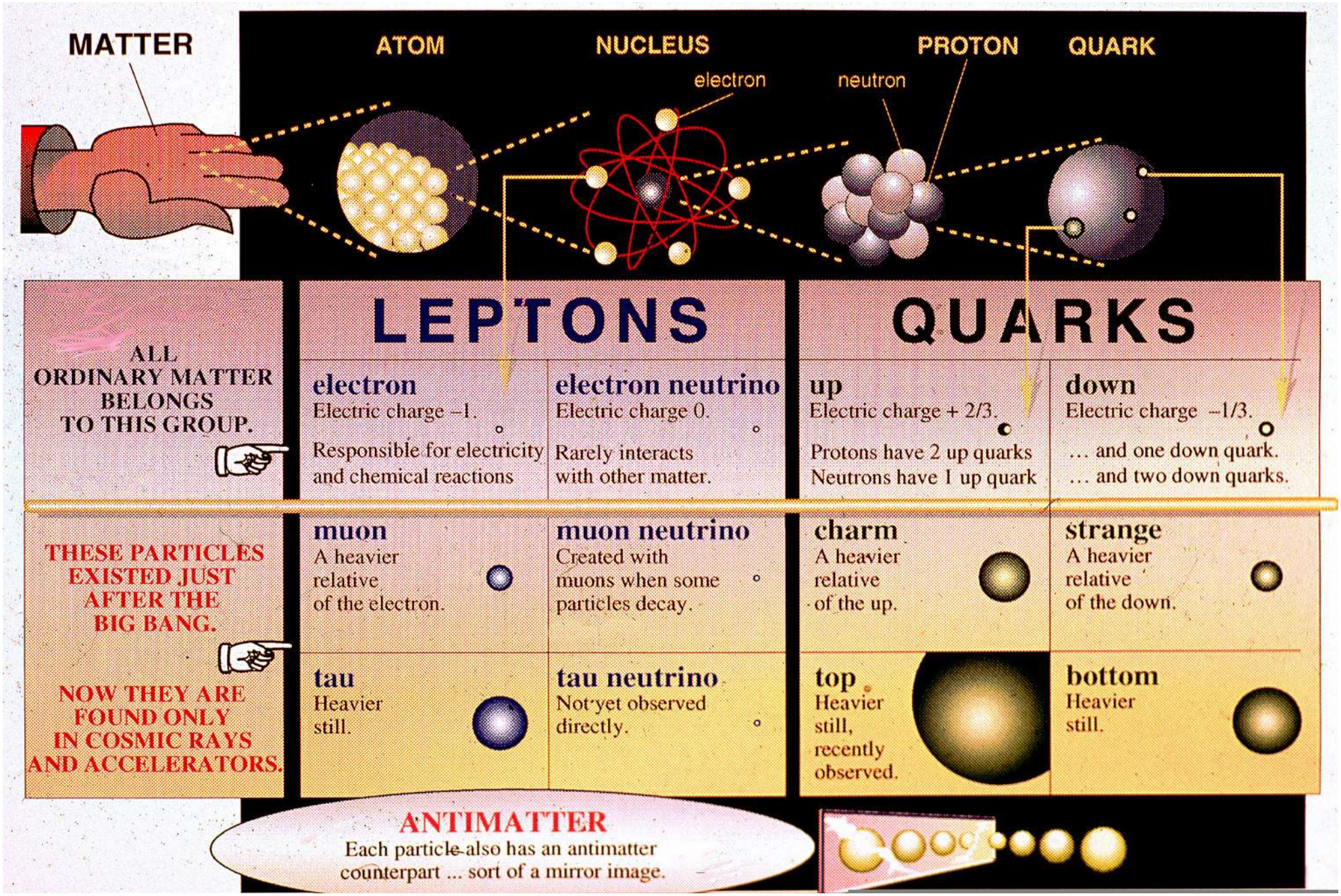
10^{-8} m



10^{-15} m

Yves Kemp
Universität Karlsruhe

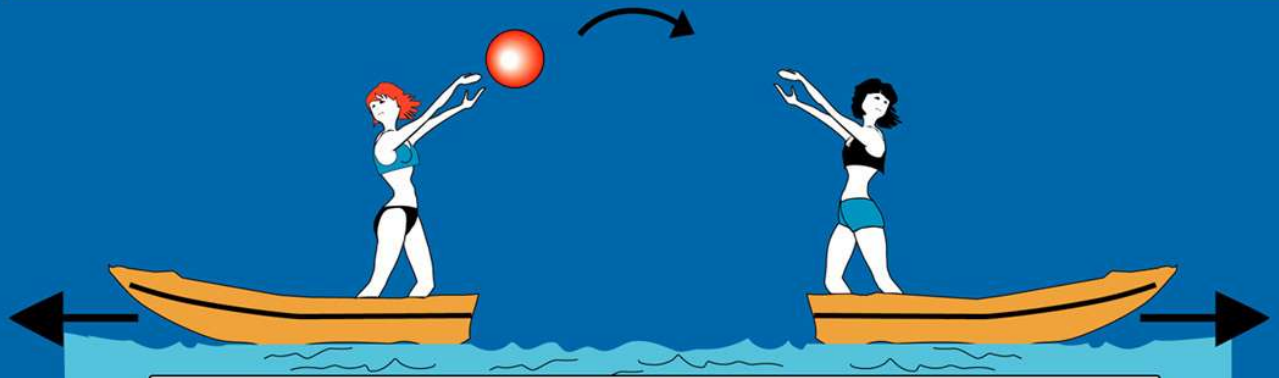
Constituents of Matter:



Forces and Interactions

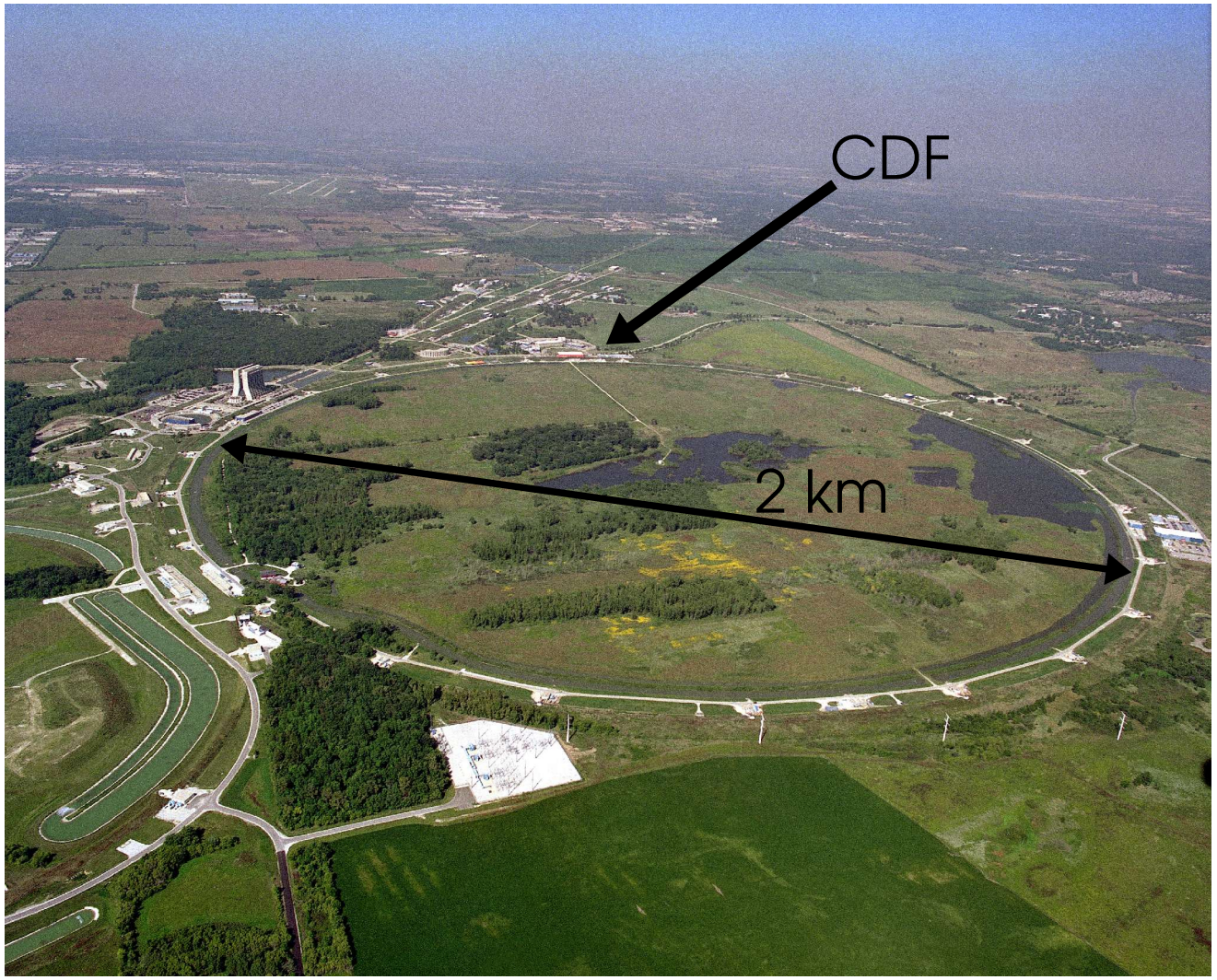
The forces in Nature

TYPE	INTENSITY OF FORCES (DECREASING ORDER)	BINDING PARTICLE (FIELD QUANTUM)	OCCURS IN :
STRONG NUCLEAR FORCE	~ 1	GLUONS (NO MASS)	ATOMIC NUCLEUS
ELECTRO -MAGNETIC FORCE	~ 10^{-3}	PHOTONS (NO MASS)	ATOMIC SHELL ELECTROTECHNIQUE
WEAK NUCLEAR FORCE	~ 10^{-5}	BOSONS Zº , W+ , W- (HEAVY)	RADIOACTIVE BETA DESINTEGRATION
GRAVITATION	~ 10^{-38}	GRAVITONS (?)	HEAVENLY BODIES



CERN AC _Z04_ V25/8/1992

Tevatron Accelerator Chain



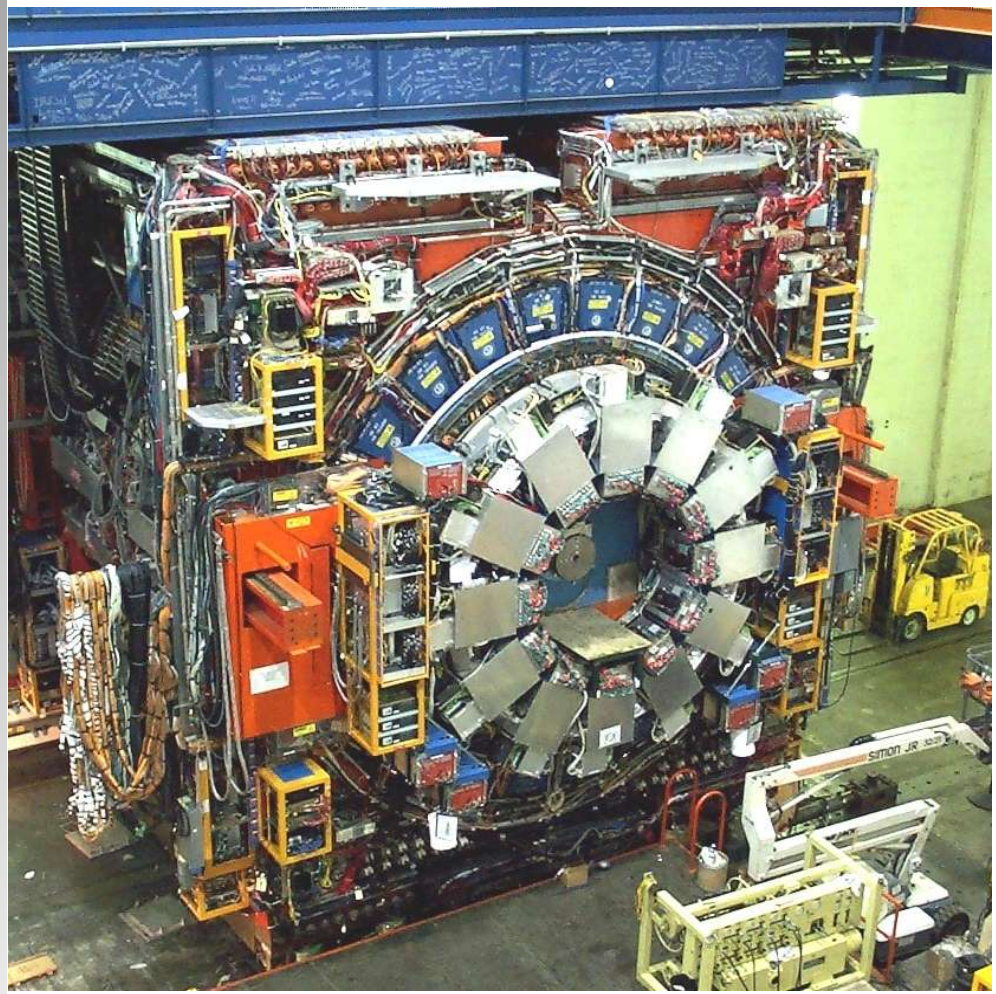
Tevatron: World largest accelarator. Located at Fermilab near Chicago

Protons and anti-protons circulating with 1 TeV.

99.99995 % of **c**

2 collision points: CDF and D0.

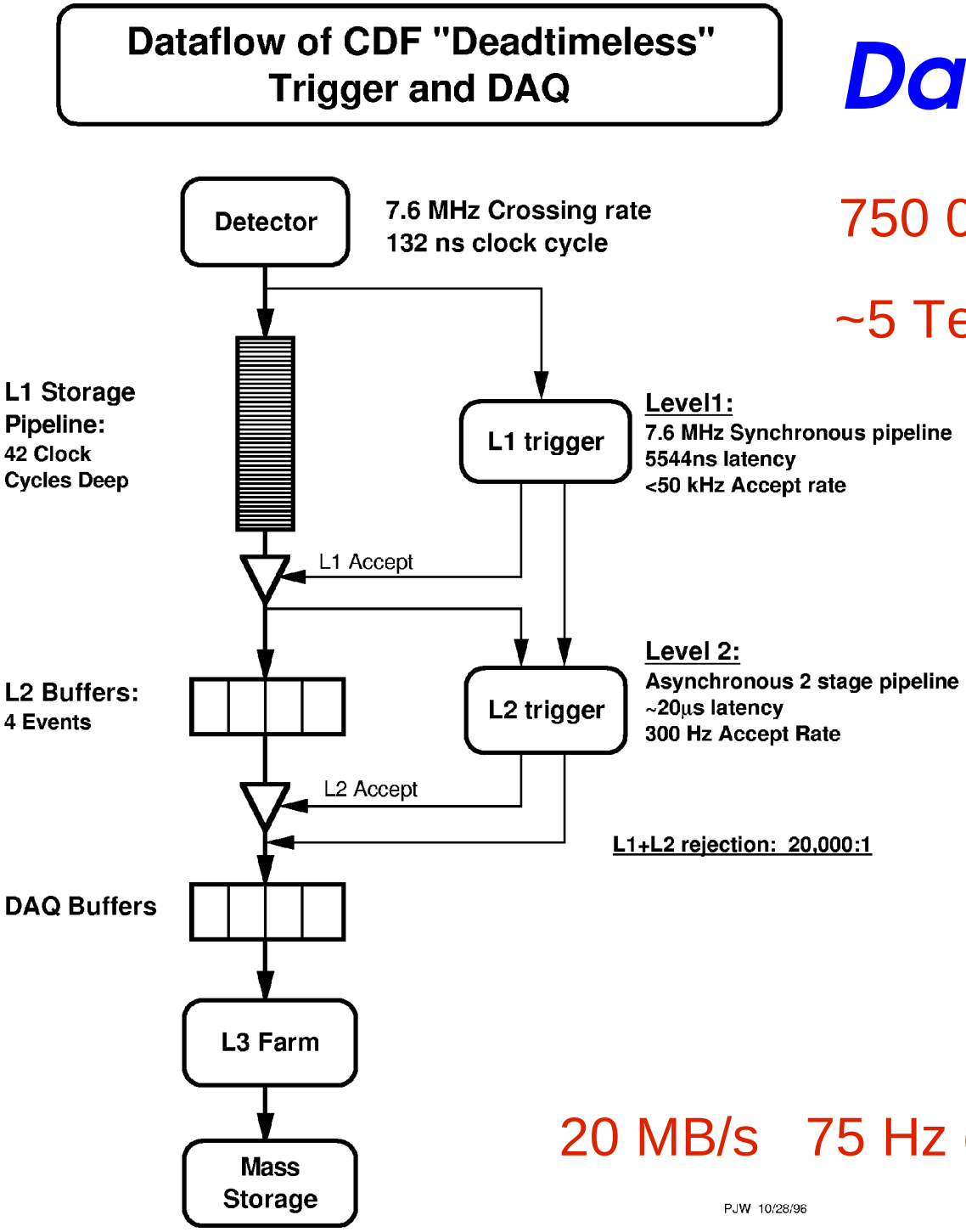
Central Detector at Fermilab (CDF)



- 2000 tons, 16m long
  10m height
- Data rate 20 MB/s
- World largest Silicon Vertex Detector: 7m^2 of silicon sensors
- 1.4 Tesla magnetic coil
- ~500 Million $US
- ~600 active collaborators in 58 institutions

Yves Kemp
Universität Karlsruhe

Data acquisition

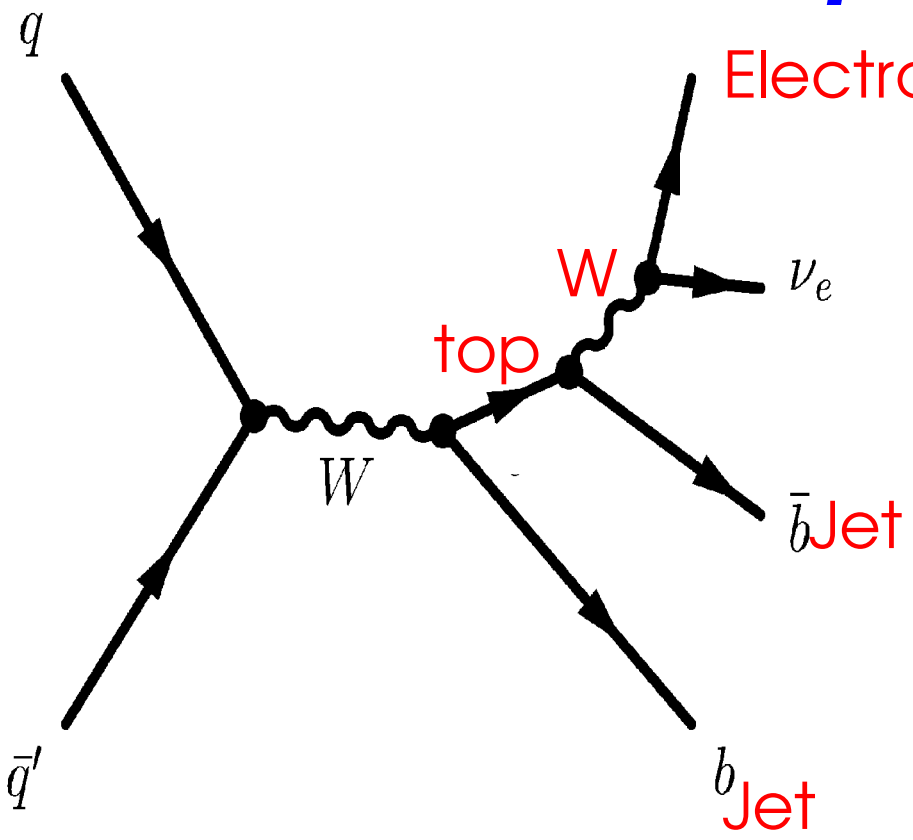


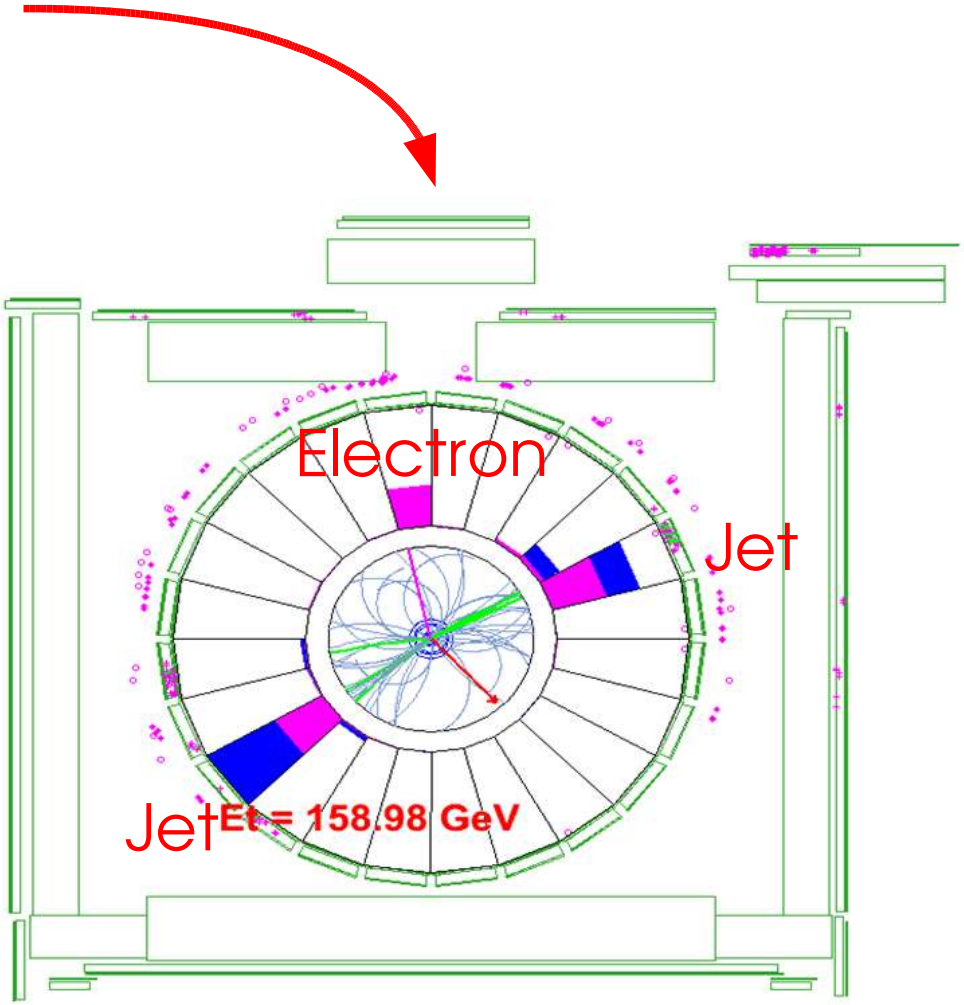
750 000 readout channels

~5 TeraByte/s

20 MB/s 75 Hz data written to tape

Electroweak Top-Quark production

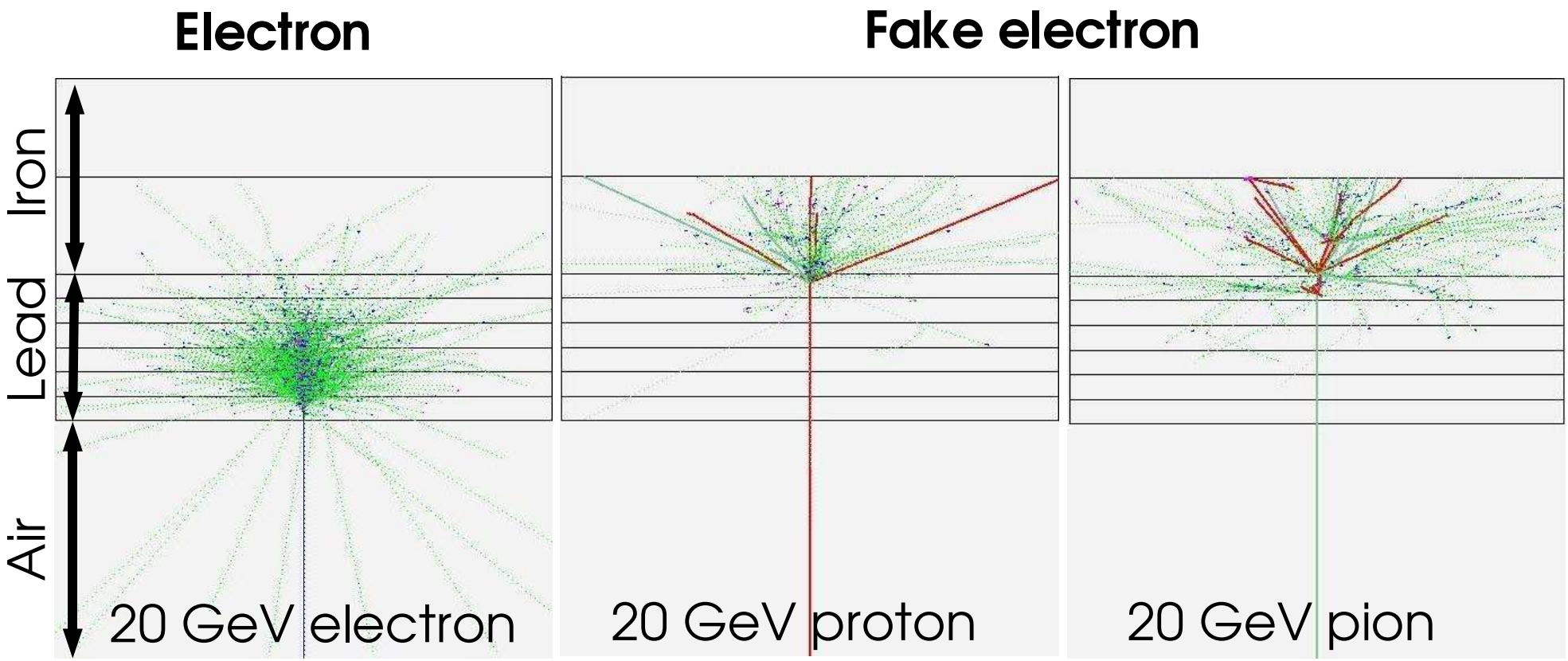




Extremely rare:
Occurs about 50 times per year
(remember: 7.6 MHz crossing rate...)

Yves Kemp
Universität Karlsruhe

Electron and Hadron Interaction



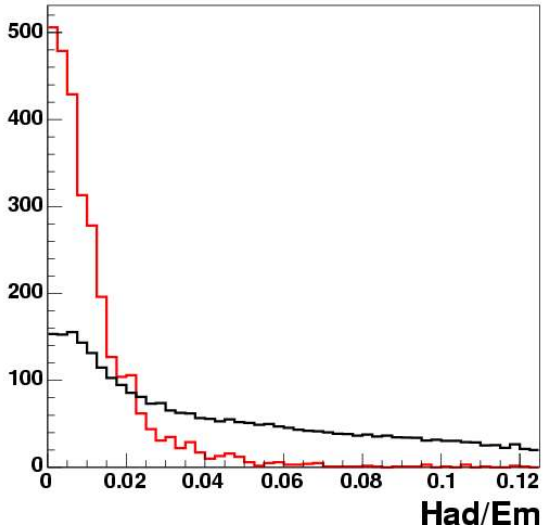
- Pair production (interaction with Pb nuclei)
- Bremsstrahlung
- Inelastic hadronic interactions
- Excitation of absorber nuclei

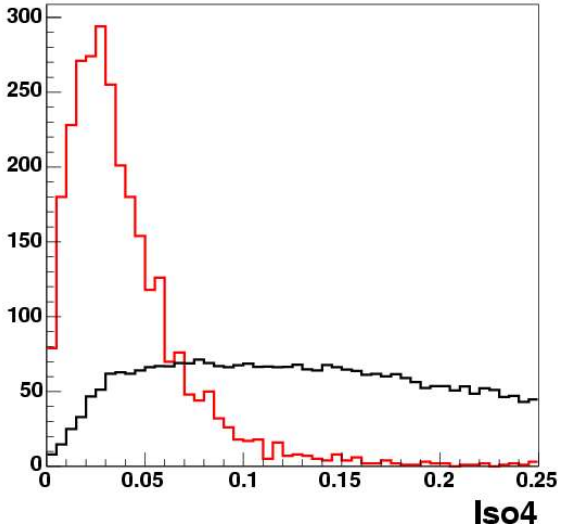
Variables for forward electron ID

- Electrons in instrumented forward region
- Em E_T (Energy in EM calorimeter)
- HadE/EmE (Ratio of energy in Hadronic over energy in EM calorimeter)
- Isolation Ratio (Energy outside main cluster region)
- PEM χ^2 (comparison with test beam data)
- PES 5by9 u/v (Shower profile in PES in u and v direction)

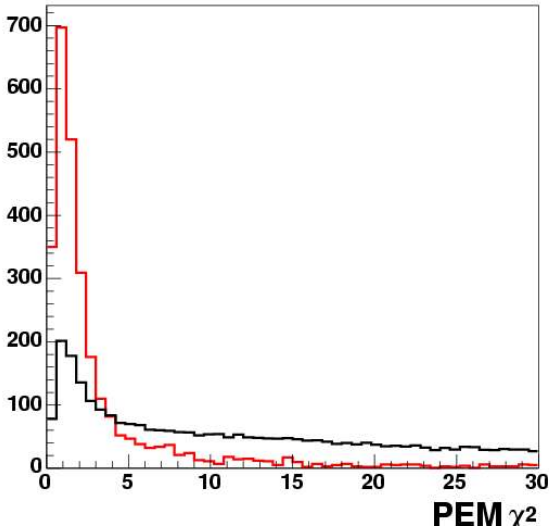
Yves Kemp
Universität Karlsruhe

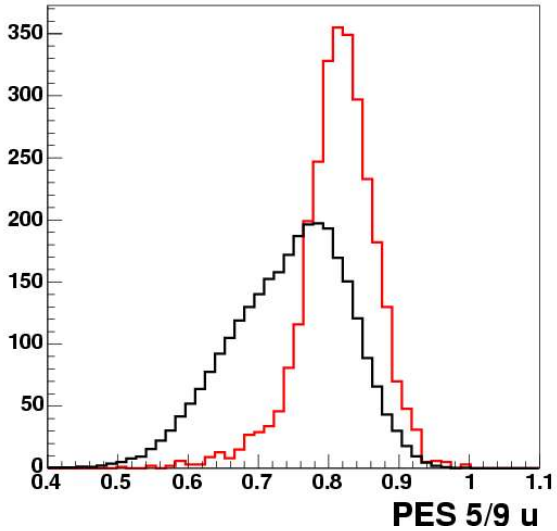
Selection variables

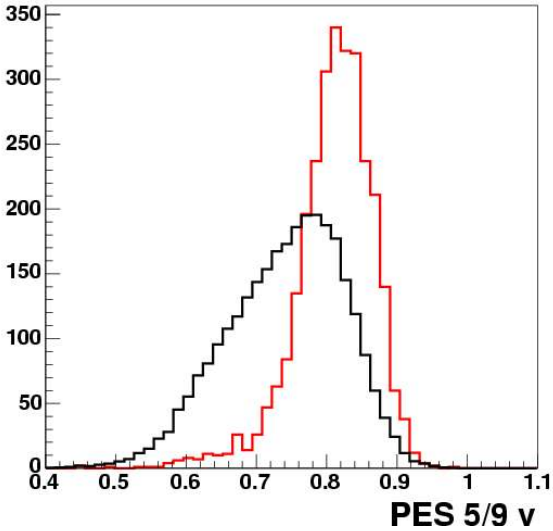




----Signal ----Background (fake)



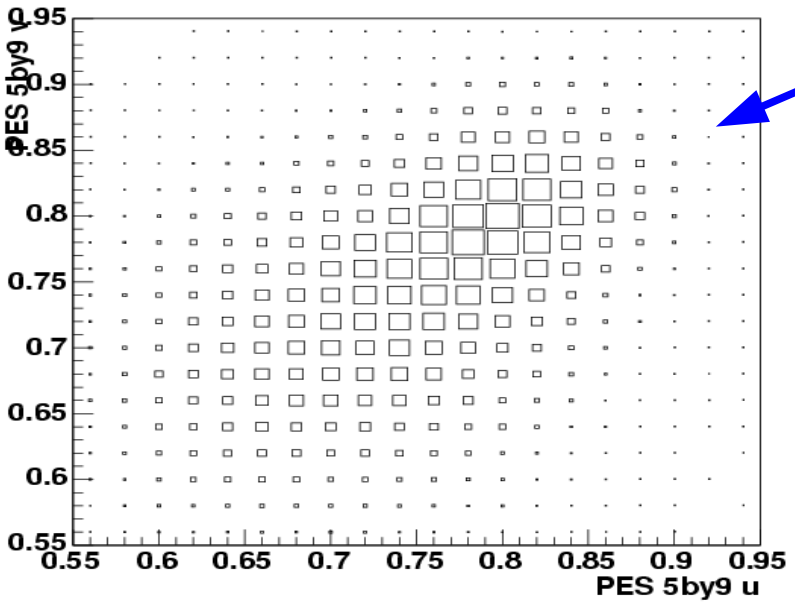




Correlation matrix

	Target	Had/Em	Iso4	PEM2	PESu	PESv
Target	100	-49.4	-66.6	-64.5	42.9	43.2
Had/Em		100	52.8	44.8	-24.8	-24.2
Iso4			100	71	-38.9	-38.5
PEM2				100	-42.8	-43.3
PESu					100	**45.9**
PESv						100

Target is -1 for background, 1 for signal



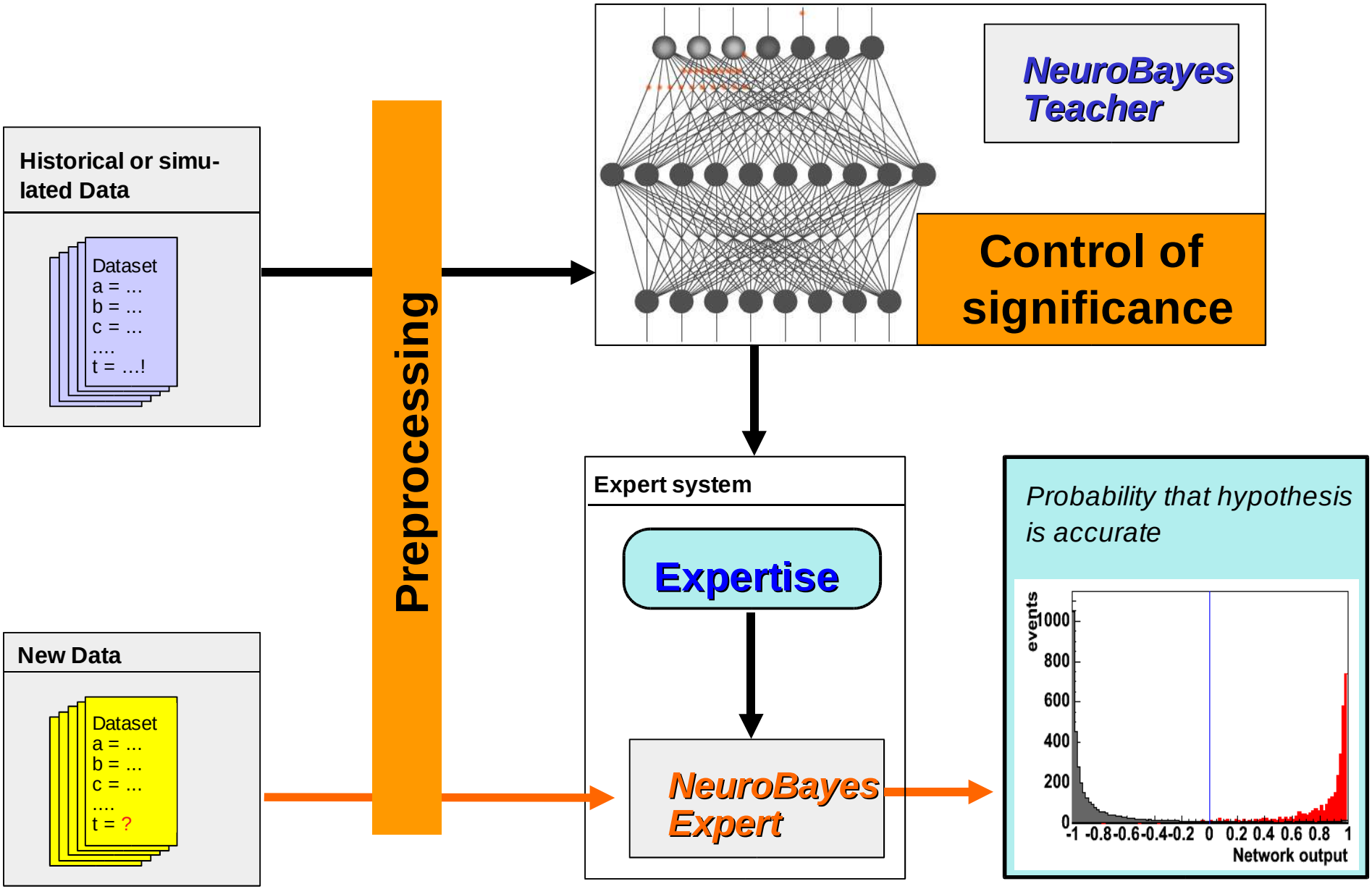
Correlation between the two PES variables due to cross talk and geometry

Yves Kemp
Universität Karlsruhe

<phi-t> **NeuroBayes®**

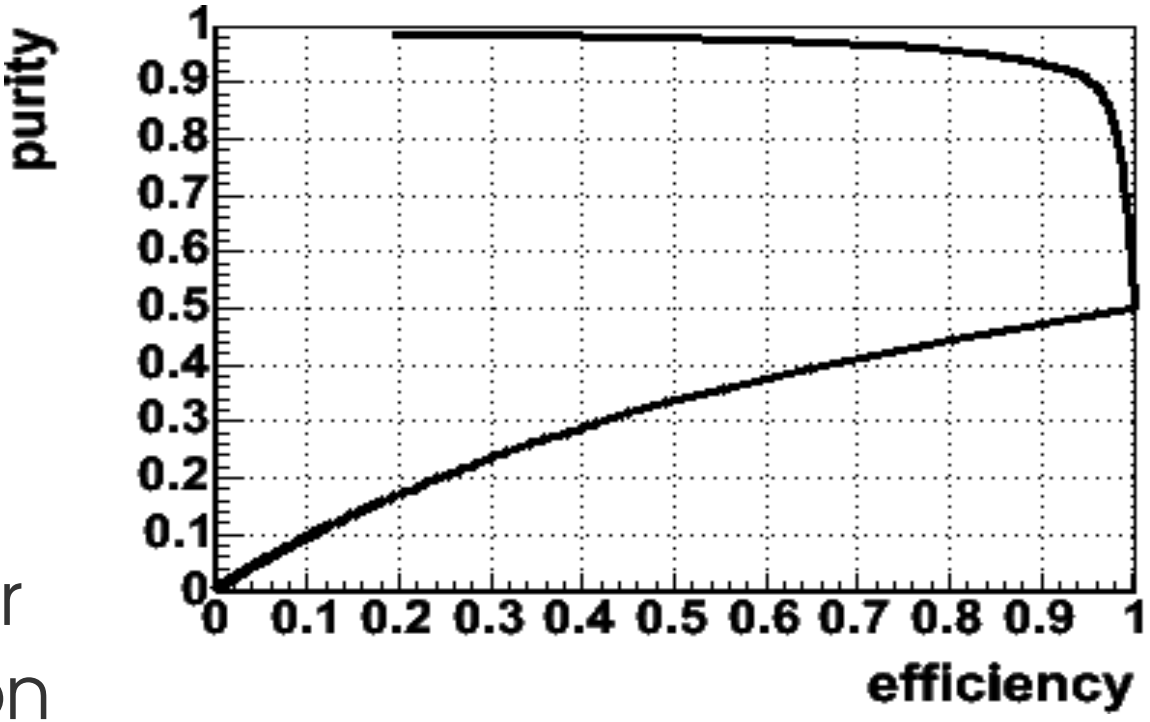
- Artificial Neural Network software package
- 2nd generation neural algorithms
  - Bayesian regularisation
  - Optimized preprocessing with transformation and decorrelation of the input variables and linear correlation to output
- Method of 2nd order allow for fast training
- Good treatment of outliers
- Does not learn by heart statistical noise
- Can make binary decisions (classification)
- Can predict uncertainties
- Can compute probability densities

ANN: Working principle



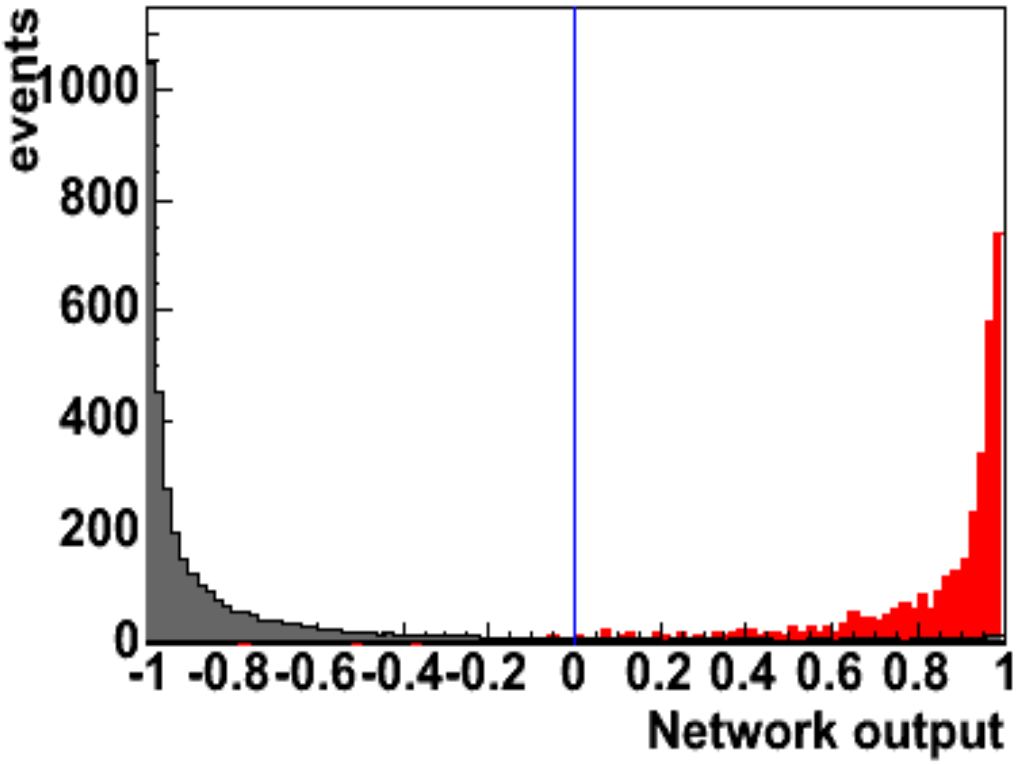
Artificial Neural Network

- 5 variables
  - Had/Em
  - Isolation
  - PEM chi2
  - PES 5/9 u
  - PES 5/9 v
- Preprocessing of the variables
- 10 nodes in intermediate layer
- Binary classification (-1 background, 1 signal)
- 200 iterations

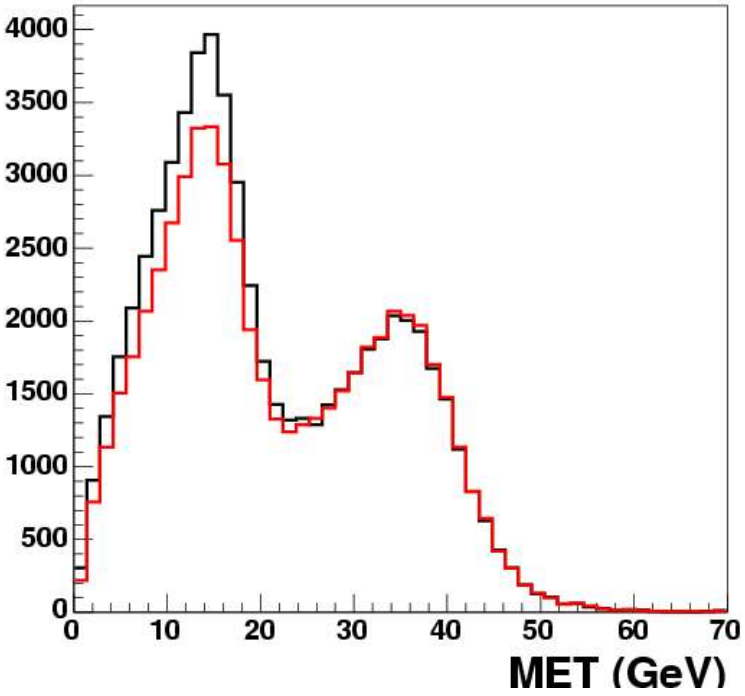


Artificial Neural Network results

- Standard Cuts:
  - Signal efficiency 84%
  - Background eff. 5.3%
- ANN cut >0.23:
  - Same signal efficiency
  - Background eff. 4.6%
  - 15% less background
- ANN cut >0.16:
  - Same background eff.
  - Signal efficiency 91%
  - 8% more signal

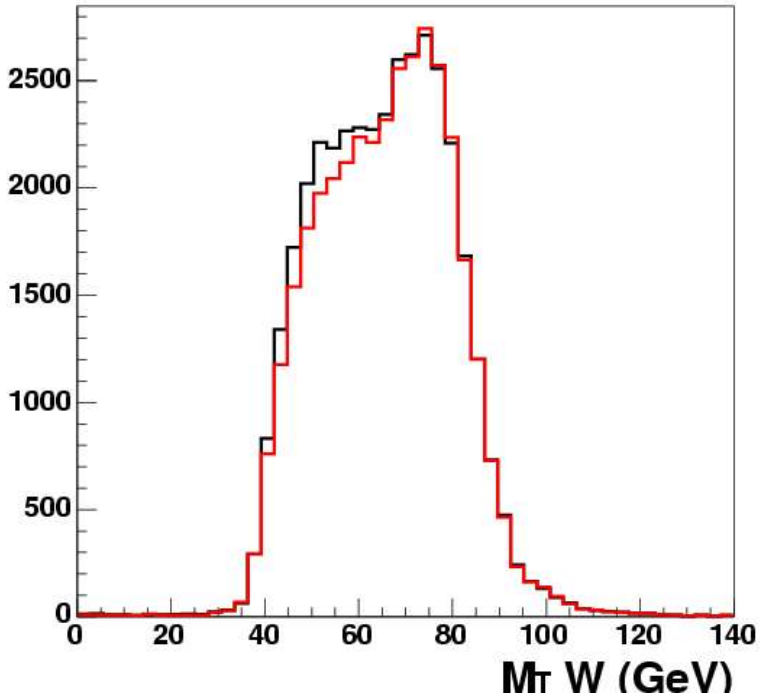


Independent tests



--cuts

--NN



Missing ET

NN cut: 40291 ev. < 25 GeV

cutbased: 34949 ev. < 25 GeV

Transverse W-Mass (MET>15 GEV)

NN cut: 36355 events

cutbased: 37687 events

Conclusion, outlook

- Experimental challenges require use of advanced statistical methods
- Artificial Neural Networks are used to combine correlated variables
  $\rightarrow$maximal exploitation of data
- Increase electron identification efficiency by 8%
- Generic tool for electron identification

Selection of the samples

- Signal sample:
  - 1 clearly identified electron in central region
  - Another electron candidate in forward region (Z-Candidate)
  - ~3000 events remain
- Background (fake) sample:
  - 2 hadronic jets (1 central, 1 forward)
  - Momentum in r-φ-plane are opposite
  - Several preselection cuts
  - ~70,000 events remain

Both samples taken from data!

Yves Kemp
Universität Karlsruhe