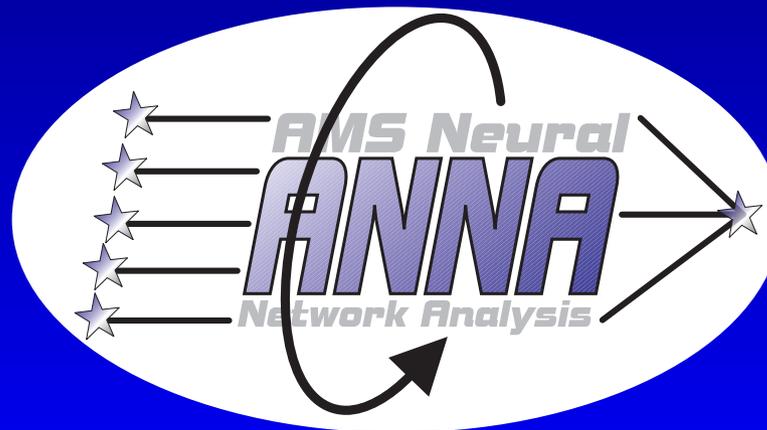


Separation von e^- und p^+ mit Hilfe von Neuronalen Netzen im TRD von AMS-02

Philip von Doetinchem

Phys. I B, RWTH-Aachen



Dunkle Materie

★ Dunkle Materie: Warum?

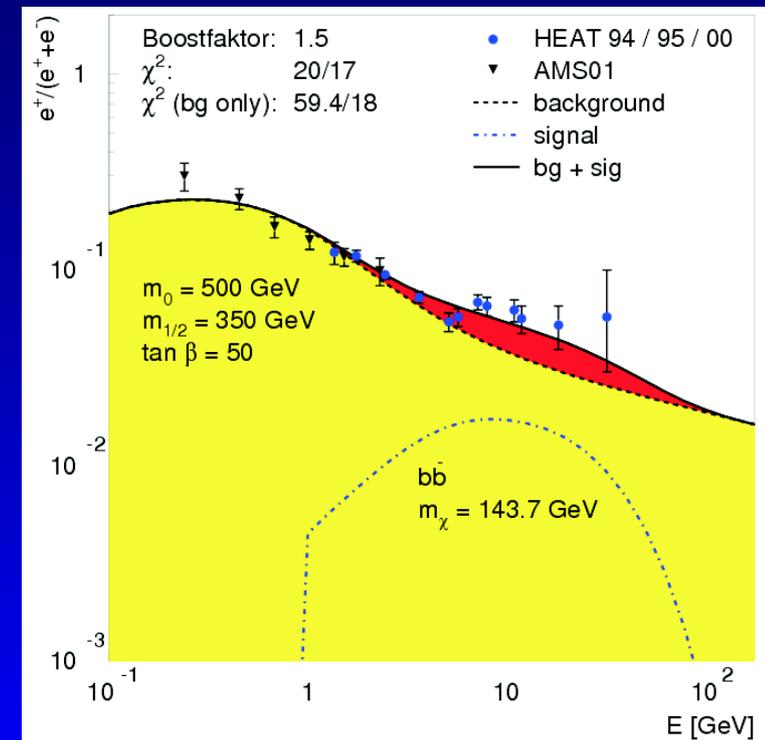
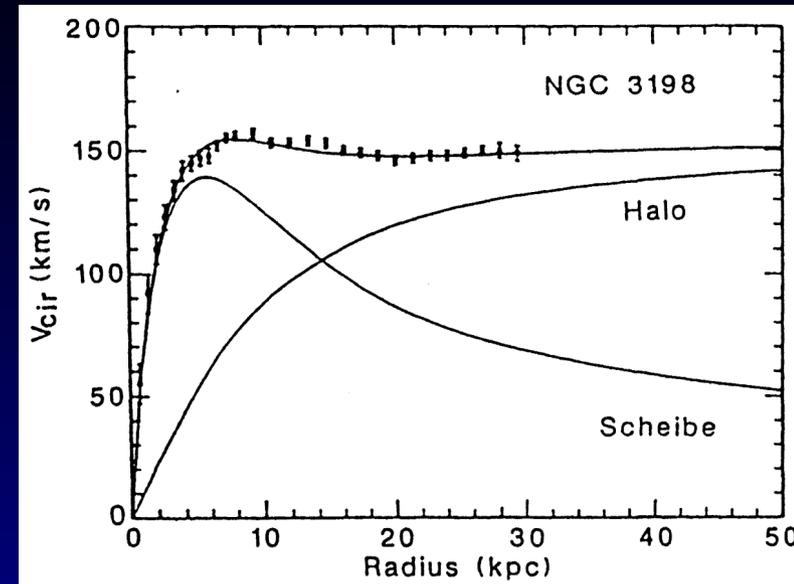
$$v \propto \begin{cases} r & r < R \\ r^{-\frac{1}{2}} & r > R \end{cases} \quad (1)$$

ABER:

$$v(r) \approx \text{const. für } r > R \quad (2)$$

★ guter Kandidat für kalte Dunkle Materie:
supersymmetrisches Neutralino

★ Messmöglichkeit über
Annihilationen in Positronen



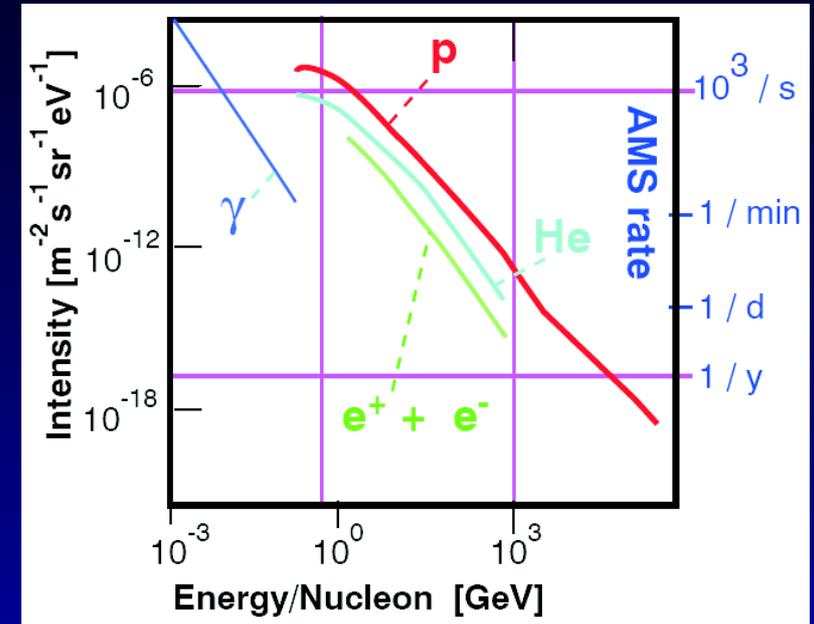
Anforderungen

- ★ Unterschied der Flüsse von Protonen und Positronen im Weltall:

$$\frac{p^+}{e^+} = 10^4 \quad (3)$$

⇒ Trennung von p^+ und e^+ muss ca. 10^6 sein.

- ★ Positroneneffizienz: richtig erkannte Positronen (Hier: 90 %)
- ★ Der Kehrwert des Anteils der Protonen, die über dieser Schwelle liegen: **Rejection**

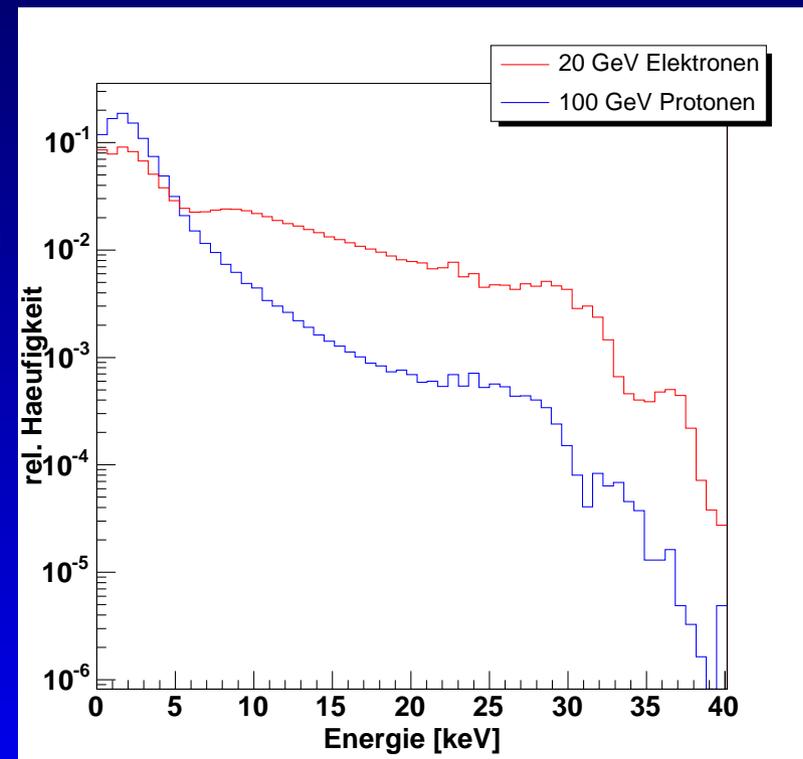
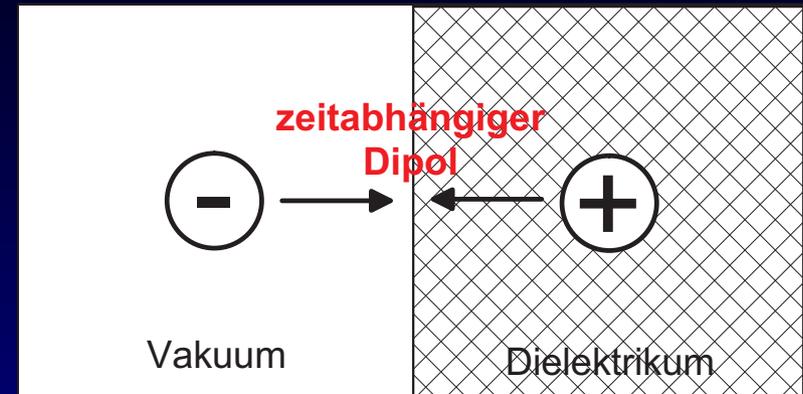


Übergangsstrahlung

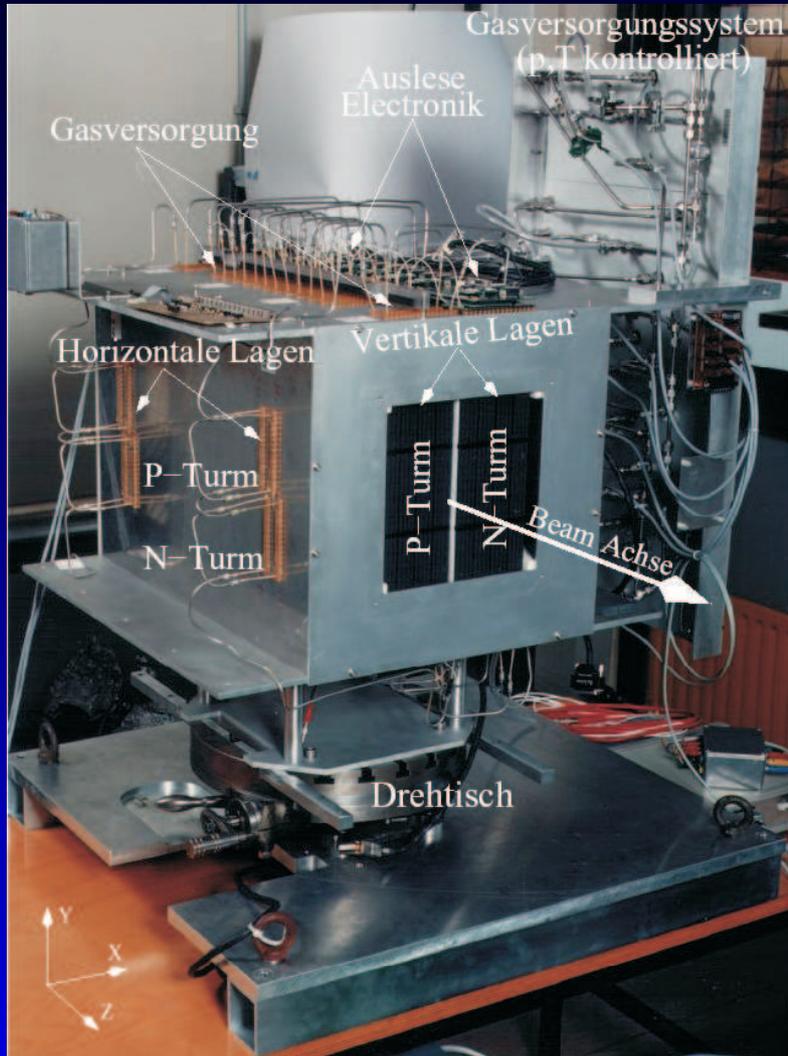
- ★ Stetigkeit der em-Felder an Grenzflächen sorgt für Strahlung
- ★ Wahrscheinlichkeit der Übergangsstrahlung (ladungsunabhängig):

$$W \propto \gamma = \frac{E}{m} \quad (4)$$

- ★ unterschiedlicher Verlauf der Energiedepositionen für Elektronen und Protonen dient zur Unterscheidung



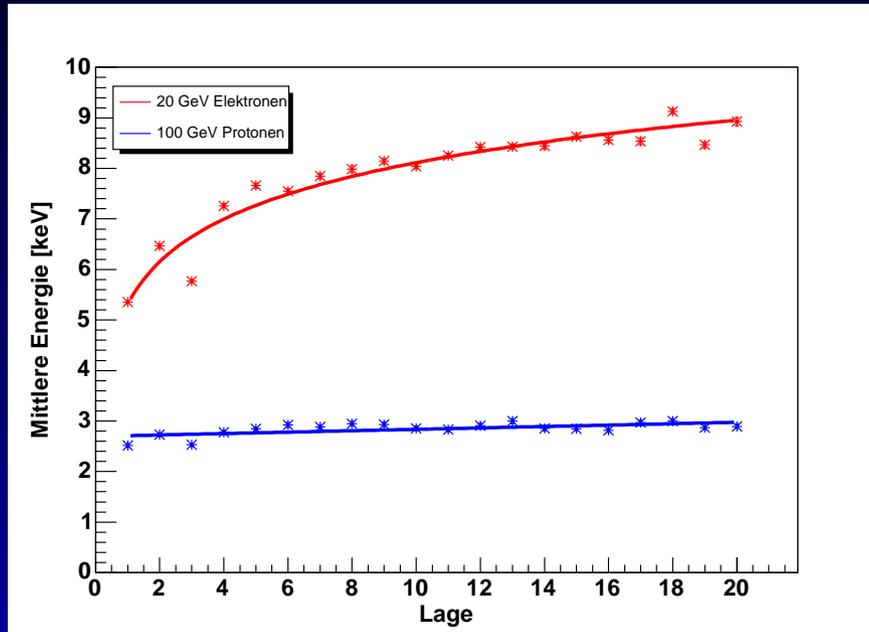
20-lagiger TRD-Prototyp



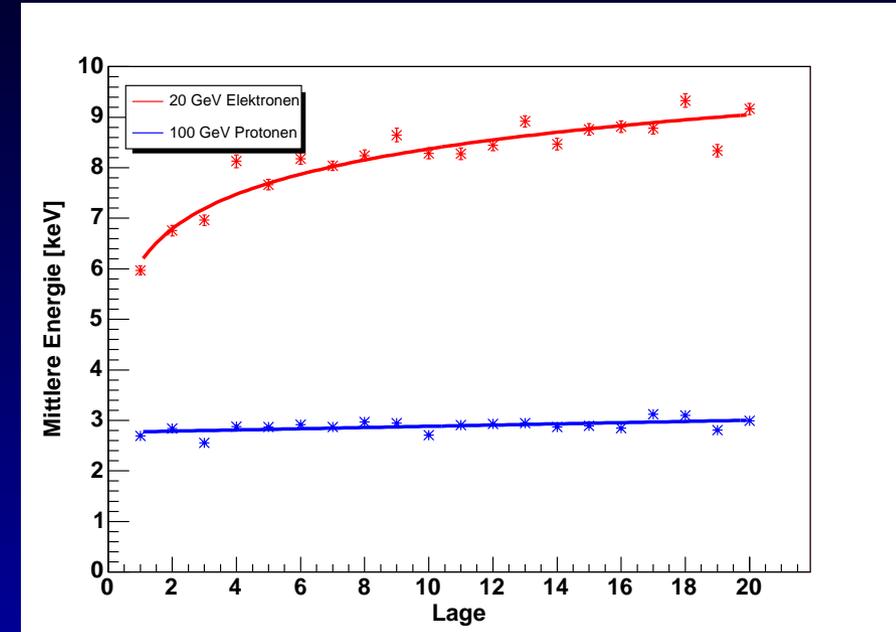
- ★ 20 Lagen, damit die Wahrscheinlichkeit der Übergangsstrahlung steigt.
- ★ Zwischen den Lagen: Radiatormaterial
- ★ Auswertung von Beamttests aus dem Jahr 2000
 - e^- : 20 GeV
 - p^+ : 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250 GeV
- ★ Verwendung des Spurfits aus der Diss. von J. Orboeck



Mittlere Energie pro Lage



Winkel: -1.5°

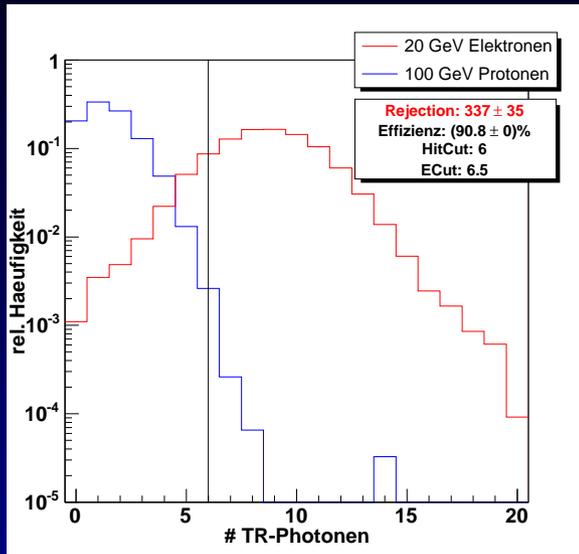


Winkel: $9.3^\circ / 8.3^\circ$

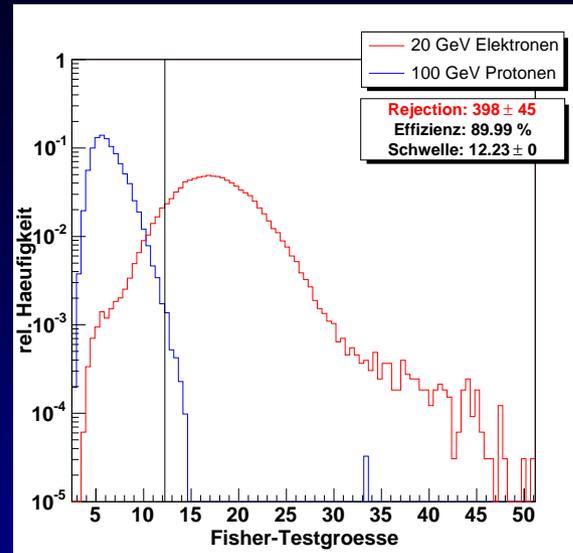
- ★ Mittlere Energie wächst, da der WQ klein ist für die Detektion der hochenergetischen TR-Photonen.
- ★ Teilchen durchquert je nach Winkel mehr oder weniger Radiatormaterial vor den gedrehten Lagen.



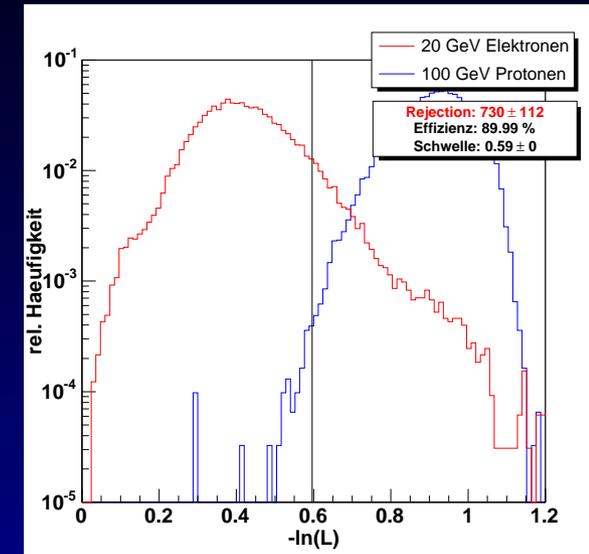
Klassische Analysemethoden



Cluster-Counting



Fisher-Diskriminante



Likelihood-Analyse

- ★ **Cluster-Counting:** Es wird die Anzahl der hohen Energieeinträge pro Event bestimmt und als Kriterium zur Klassifikation verwendet.

- ★ **Fisher-Diskriminante:**
$$t = \sum_{i=1}^n f_i X_i - \frac{1}{2} \sum_{i=1}^n f_i (\bar{X}_i^{(1)} + \bar{X}_i^{(2)})$$

Es kann gezeigt werden, dass der Algorithmus bei Normalverteilung optimal ist.

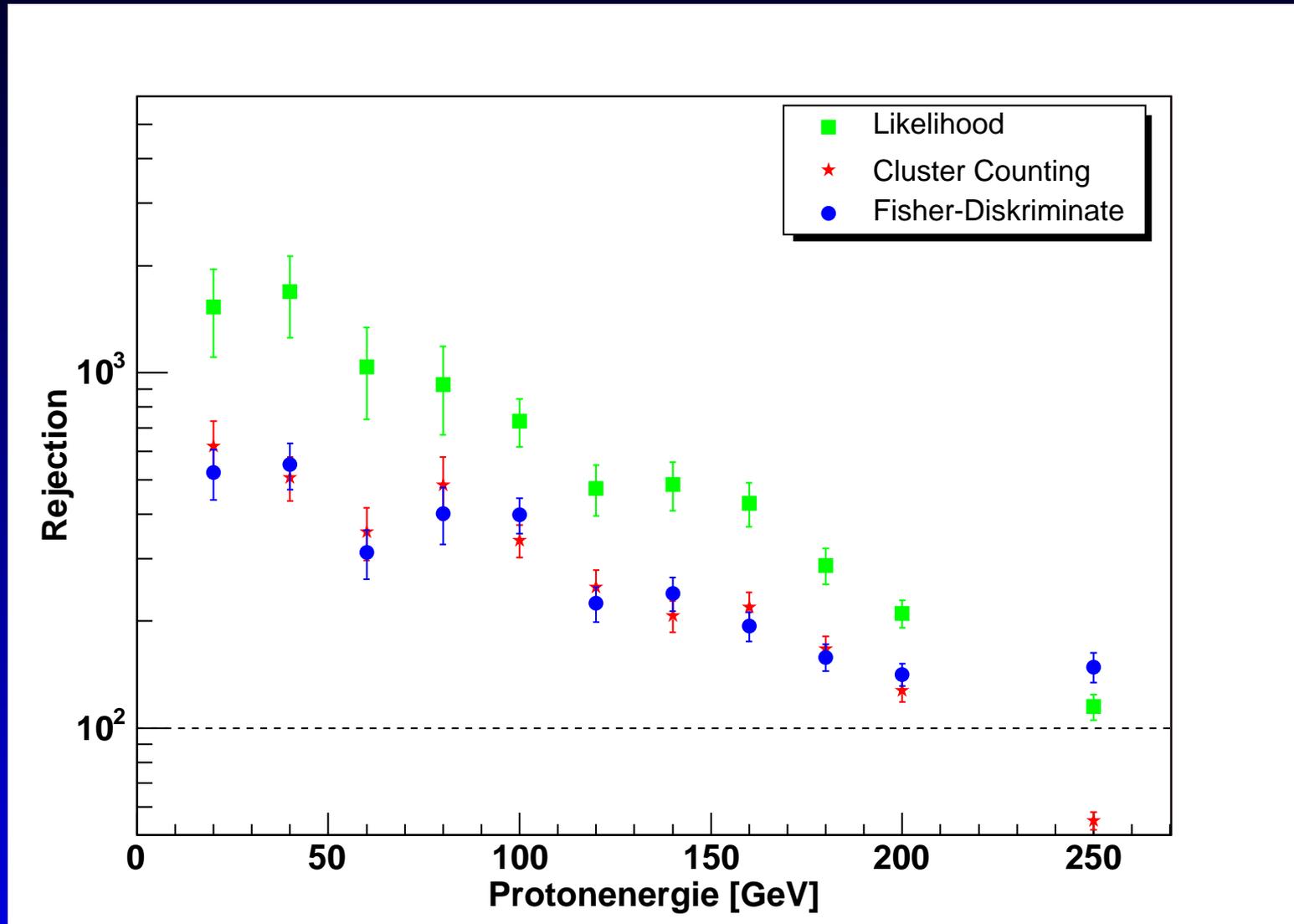
- ★ **Likelihood:** Ermittlung des mittleren LH-Wertes pro Event:
$$\bar{P}_{e/p} = \sqrt[n]{\prod_i P_{e/p}^{(i)}(E)}$$

Likelihood-Ratio:
$$L = \frac{\bar{P}_e}{\bar{P}_p + \bar{P}_e}$$

Verwendung analytischer Funktionen für $P_{e/p}(E)$, die an die Daten angepasst wurden.



Vergleich der Methoden



Kurze Einführung in Neuronale Netze

★ Backpropagation-Lernregel:

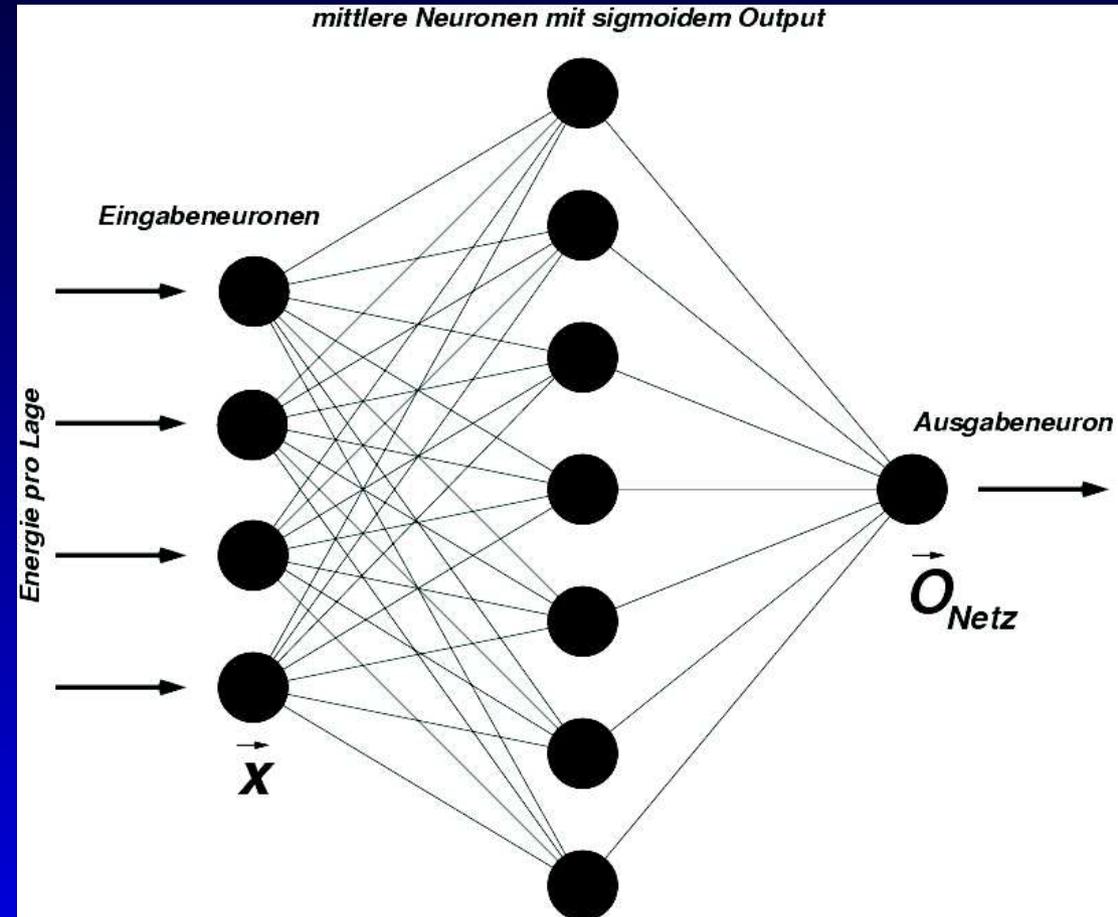
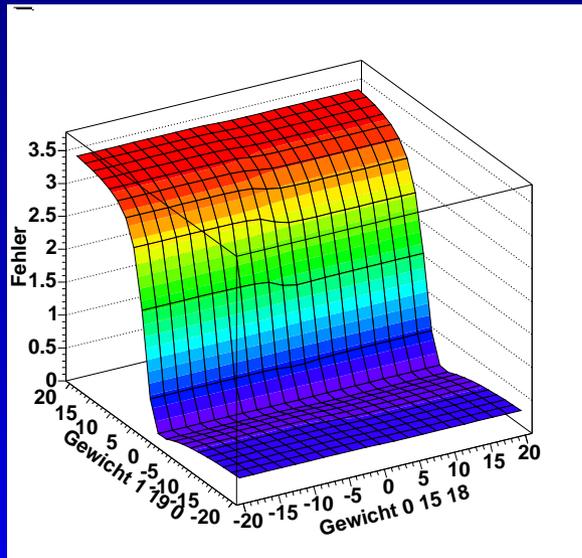
$$F(\vec{w}) = \frac{1}{N} \sum_{k=1}^N \underbrace{(E(\vec{x}_k) - o_{\text{Netz}}(\vec{x}_k, \vec{w}))^2}_{f_k(\vec{w})}$$

$$\vec{w}(t+1) = \vec{w}(t) - \Delta \vec{w}(t+1)$$

$$\Delta w_i(t+1) = \eta \cdot \frac{1}{N} \sum_{k=1}^N \frac{\partial f_k(\vec{w})}{\partial w_i}$$

★ Verlauf des Mean Square Errors ($F(\vec{w})$)

bei Variation eines Gewichtes:



NN: Lernregeln

- ★ $\bar{\delta}$ - δ -Lernregel: Anpassung der Lernrate je nach Produkt aufeinanderfolgender Gradienten:

$$\Delta w_i(t+1) = \eta_i(t+1) \cdot \frac{1}{N} \sum_{k=1}^N \frac{\partial f_k(\vec{w})}{\partial w_i}(t+1) + \alpha \cdot \Delta w_i(t) \quad (5)$$

η_i : individuelle Lernrate, α : Moment

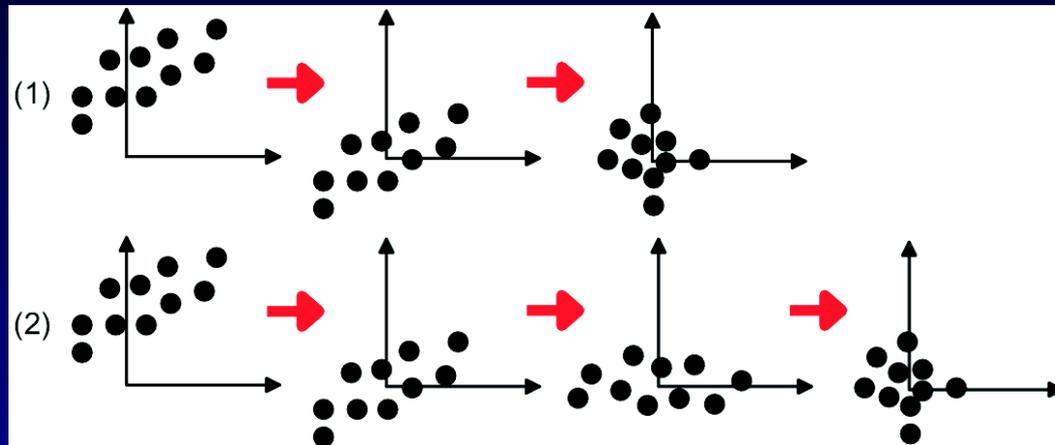
Veränderung der Lernrate:

$$\eta_i(t+1) = \begin{cases} \eta_i(t) \cdot \eta^- & \frac{\partial f_k(\vec{w})}{\partial w_i}(t+1) \frac{\partial f_k(\vec{w})}{\partial w_i}(t) < 0 \\ \eta_i(t) + \eta^+ & \frac{\partial f_k(\vec{w})}{\partial w_i}(t+1) \frac{\partial f_k(\vec{w})}{\partial w_i}(t) > 0 \\ \eta_i(t) & \text{sonst} \end{cases} \quad (6)$$

- ★ **RProp-Lernregel:**
 - ★ nur Verwendung des Vorzeichens vom Fehlergradienten
 - ★ Zurücknahme von Schritten, die über ein Minimum hinausgehen.
 - ★ Anpassung der Lernraten wie bei $\bar{\delta}$ - δ -Lernregel



NN: Datentransformation



★ Transformation der Input-Daten, um alle Eingaben gleichwertig zu berücksichtigen.

★ pro Input-Neuron: Mittelwert 0 mit Breite 1

1. Abzug des Mittelwertes und Normierung auf Breite:

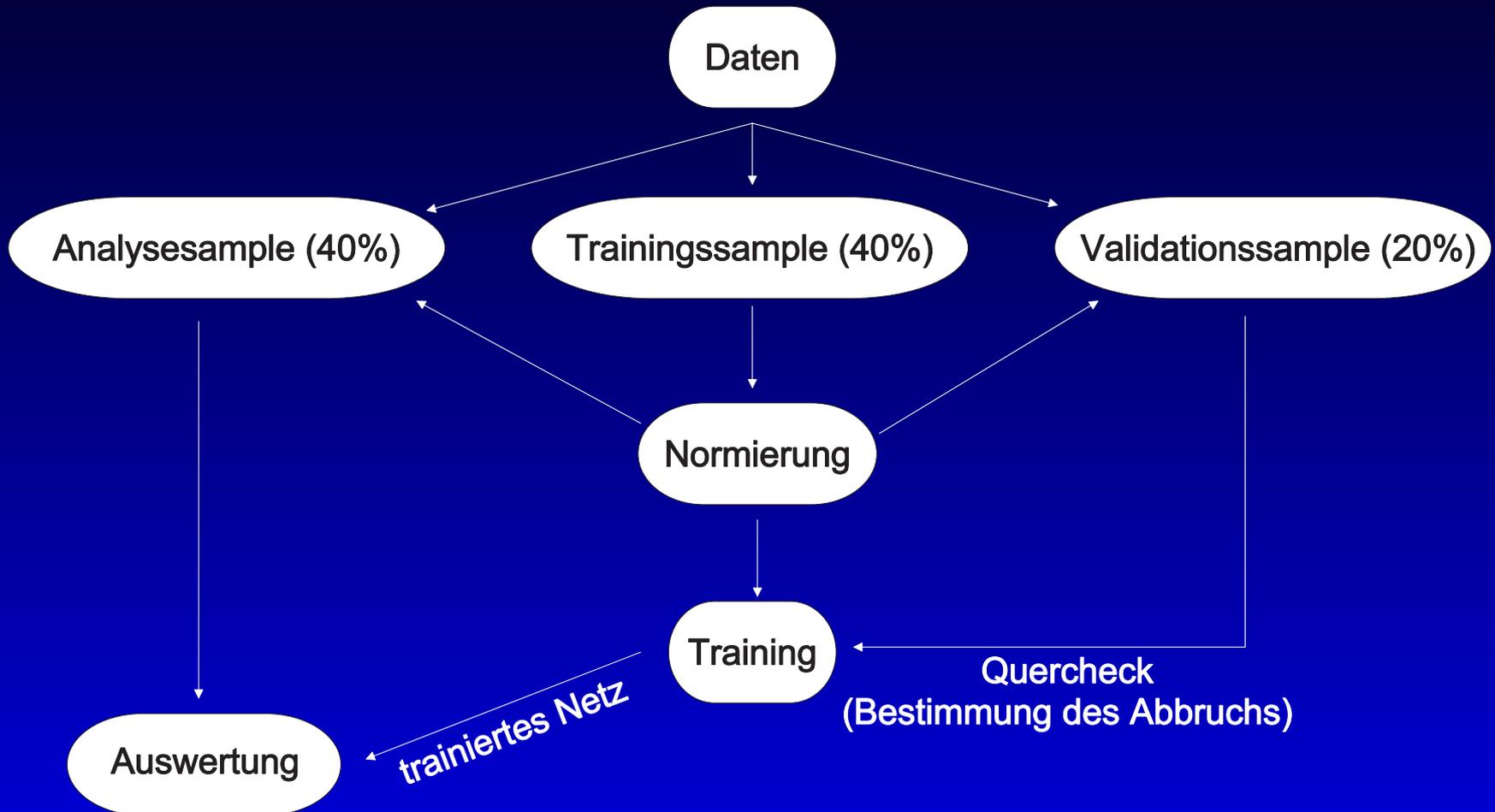
$$x_i^{\text{trafo}} = \frac{x_i - \bar{x}}{\sigma} \quad (7)$$

2. Transformation der Kovarianzmatrix auf die Einheitsmatrix mit Eigenrotation:

$$V = \sum_n (\vec{x}_n - \vec{\bar{x}})(\vec{x}_n - \vec{\bar{x}})^T \quad (8)$$



Analyseprozess



Aufgaben des Neuronales Netzes

- ★ Neuronales Netz (NN) erhält als Input die Energiedepositionen der Hits auf der Spur pro Lage. Also: 20 Input Neuronen
- ★ Finden einer Abbildungsvorschrift:

$$\left. \begin{array}{l} \text{Elektronen} \rightarrow 1 \\ \text{Protonen} \rightarrow -1 \end{array} \right\} \rightarrow \text{Unterscheidungskriterium}$$

- ★ Rejection soll besser als 100 sein.
- ★ Auffinden der besten Konfiguration des NN durch Variation der Parameter.
- ★ Untersuchung der Energiedepositionen pro Lage auf mögliche Korrelationen.



Auffinden der besten Lernmethode

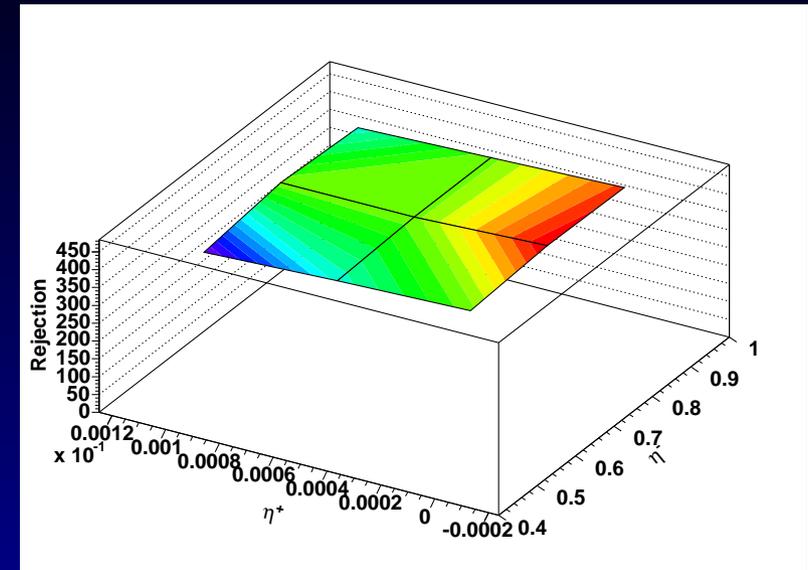
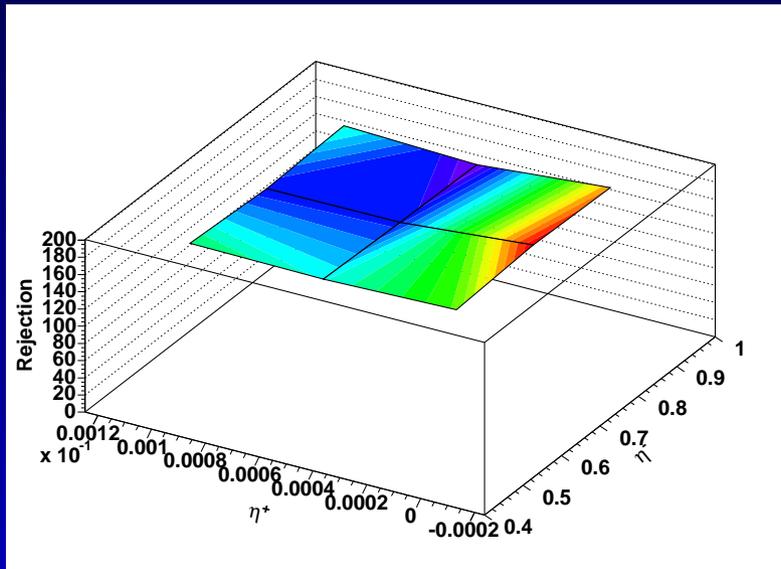
1. Verwenden eines dreilagigen Netzes mit 20 (21)-Eingaben und variabler Zahl mittlerer Neuronen.
2. Training mit RProp- oder $\bar{\delta}$ - δ -Lernregel
3. Eingabe von Energien oder von Likelihood-Ratios
4. Datentransformation:
 - (a) Abzug des Mittelwertes und Normierung auf Breite
 - (b) Transformation der Kovarianzmatrix auf Einheitsmatrix



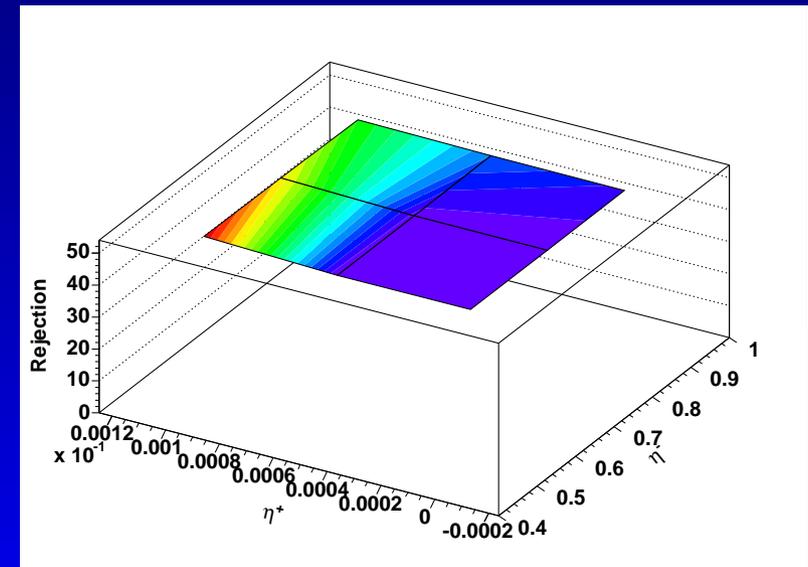
RProp-Auswertung

Datentransformation (a), Energieeingabe \rightarrow

Datentransformation (b), Energieeingabe \downarrow



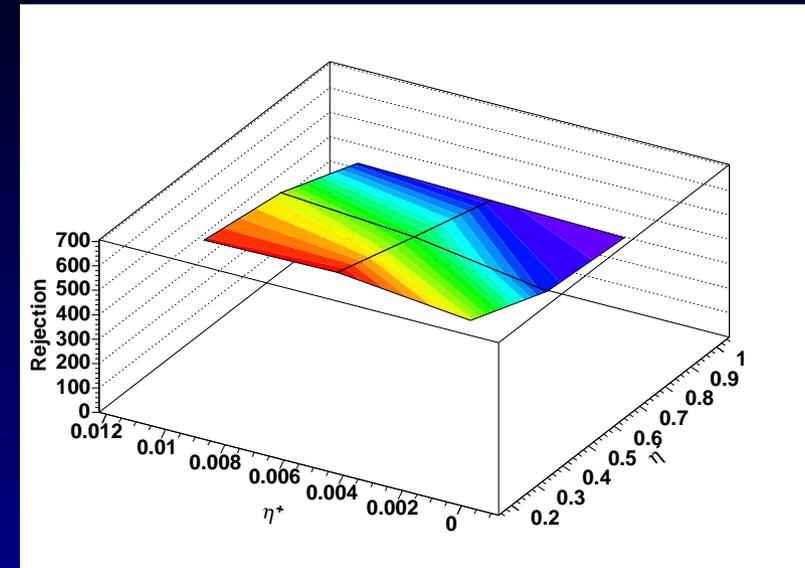
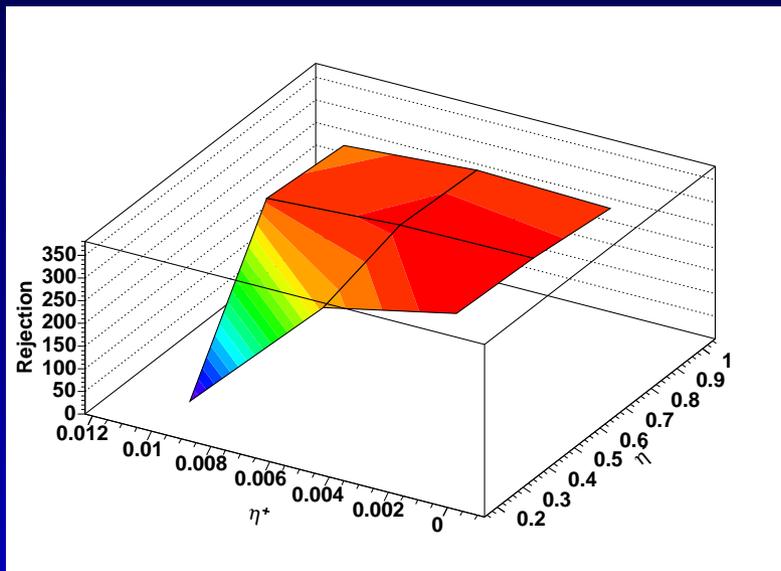
Datentransformation (a), LH-Ratio-Eingabe \rightarrow



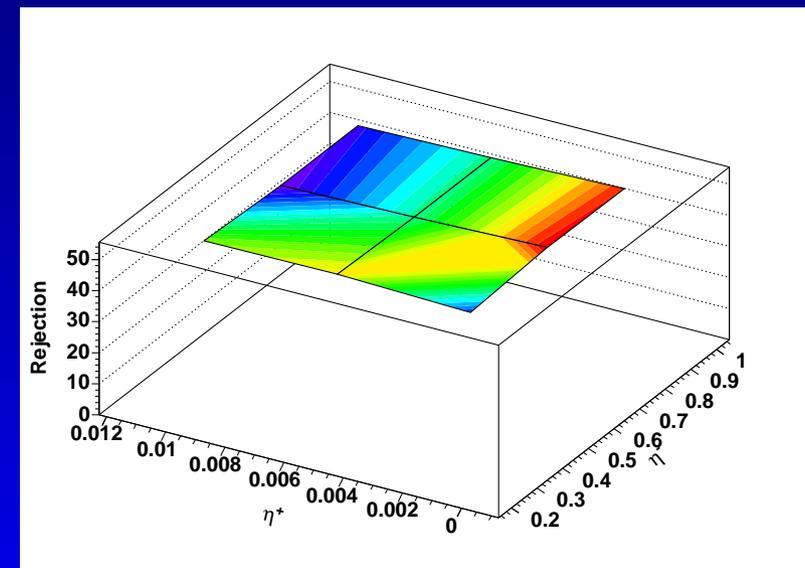
$\bar{\delta}$ - δ -Auswertung

Datentransformation (a), Energieeingabe \rightarrow

Datentransformation (b), Energieeingabe \downarrow

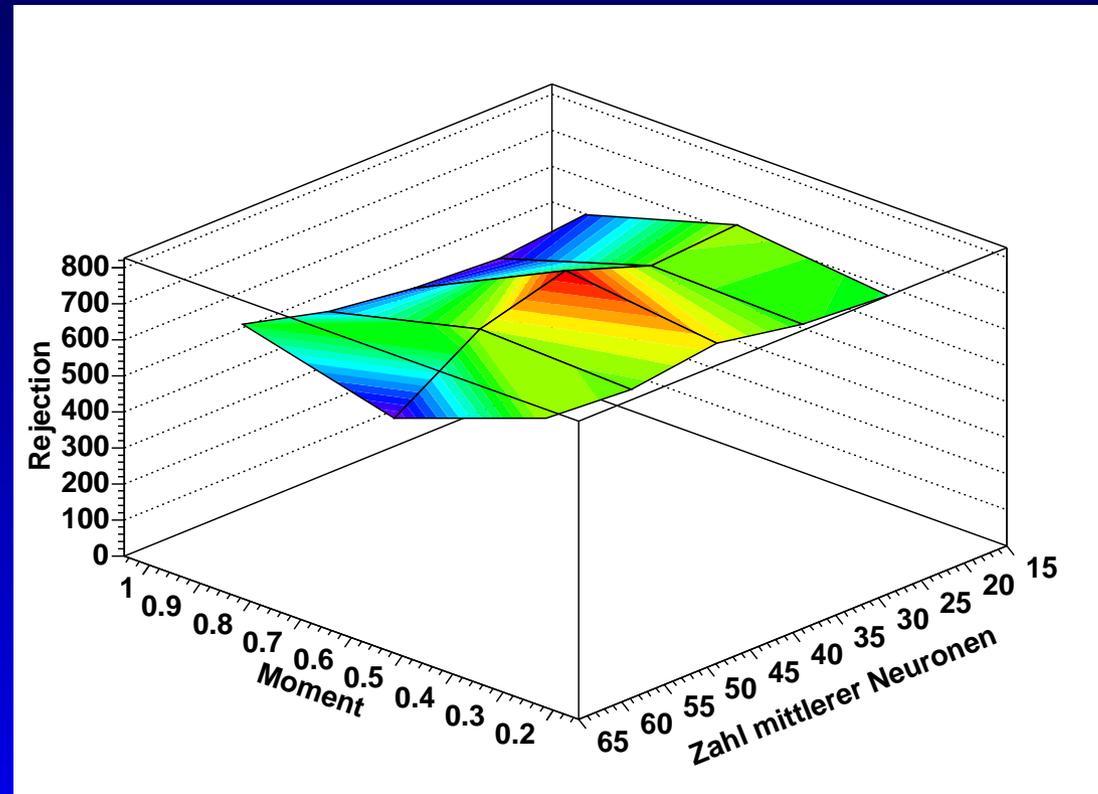


Datentransformation (a), LH-Ratio-Eingabe \rightarrow

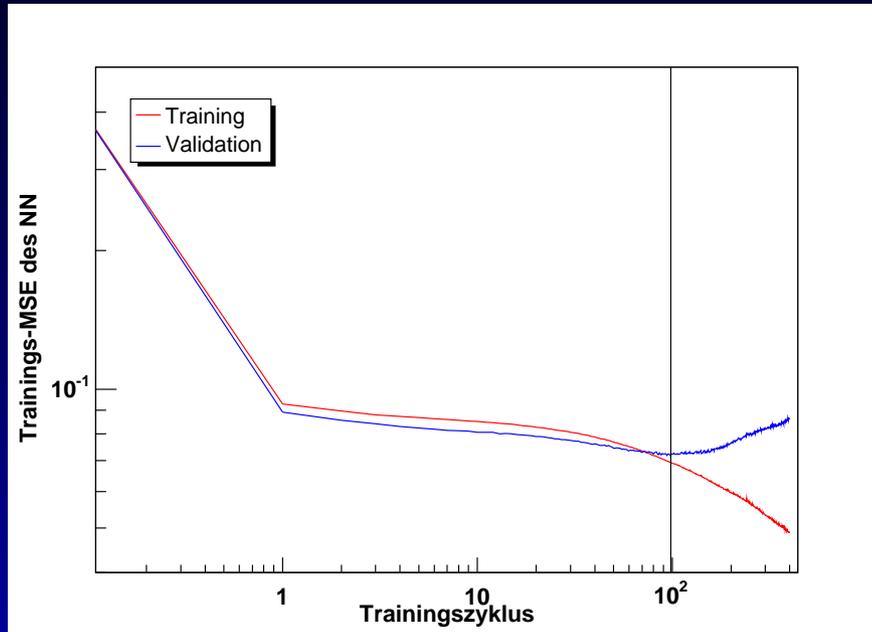


Weitere Parametervariation

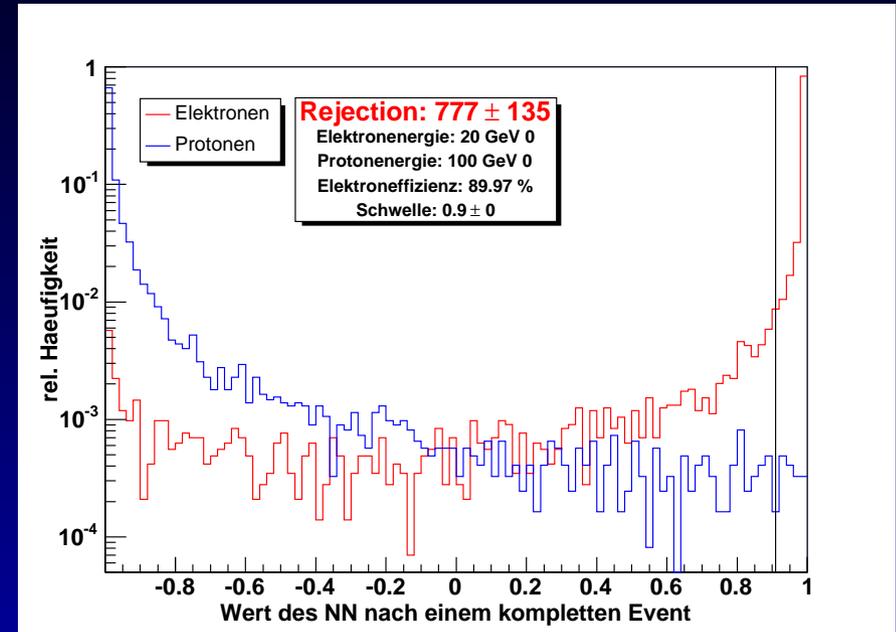
- ★ bisher bestes Training mit
 - ★ Eingabe der normierten Energien
 - ★ $\bar{\delta}$ - δ -Lernregel
- ★ genauere Variation von mittleren Neuronen und Moment α



Verlauf des Fehlers und Rejection



Fehler

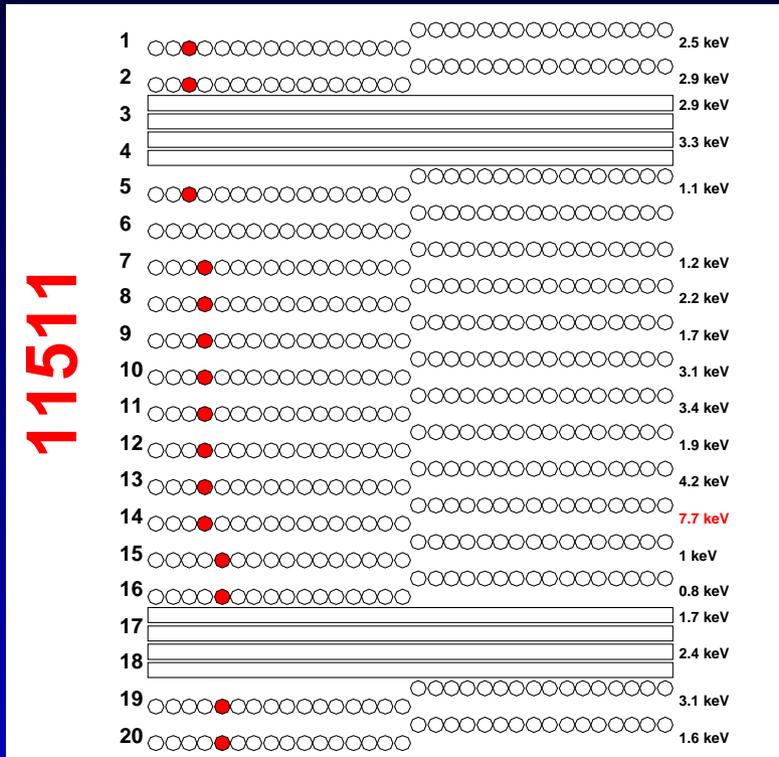


Ausgabe

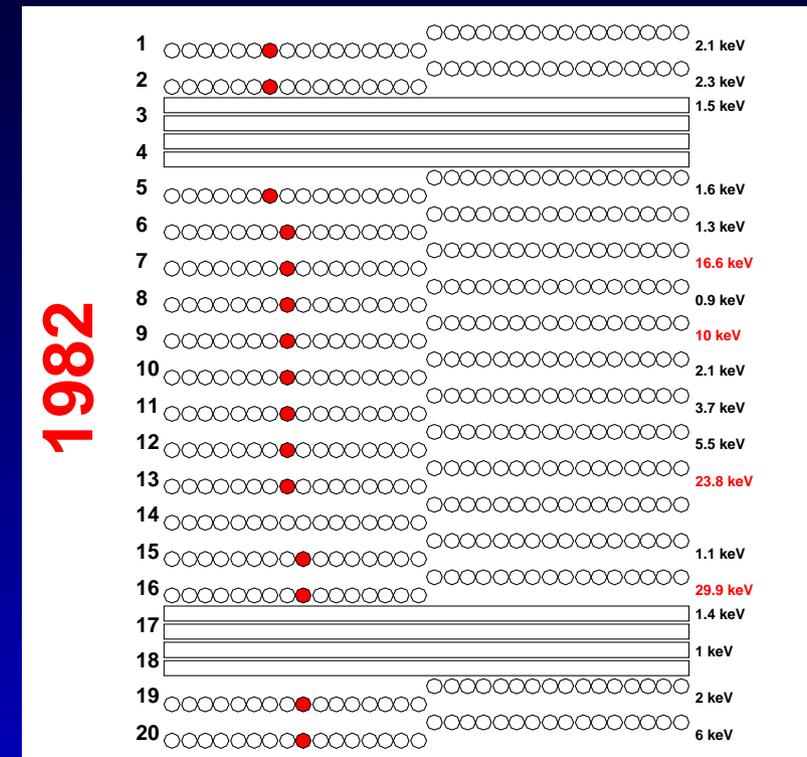
- ★ Training wird an der Stelle mit minimalem Fehler des Validationssamples abgebrochen.
- ★ Rejection liegt also weit über den geforderten 100!



Beispiele für falscherkannte Events



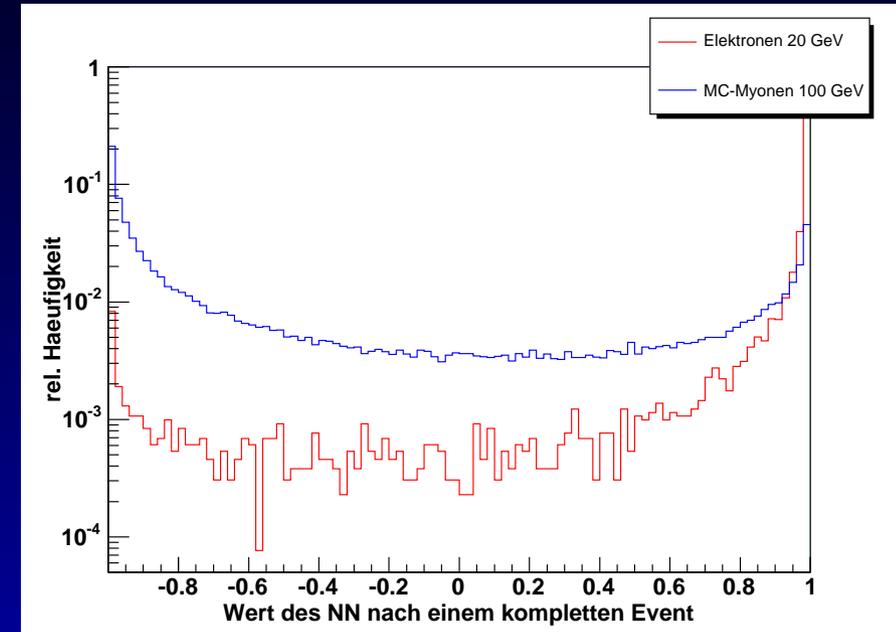
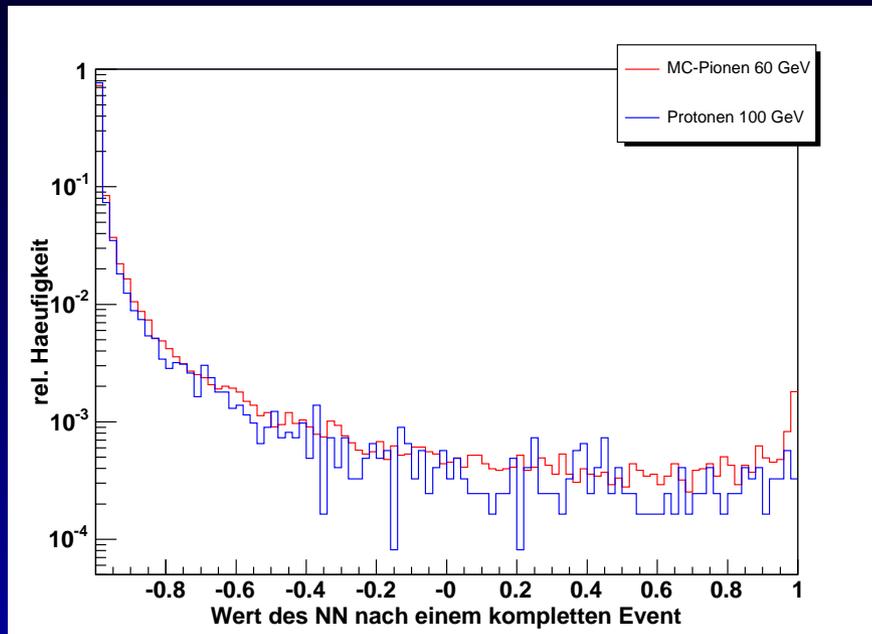
falscherkanntes Elektron:
keine hohen Einträge



falscherkanntes Proton:
wenige sehr hohe Einträge



Mögliche Erklärung mit Monte Carlos



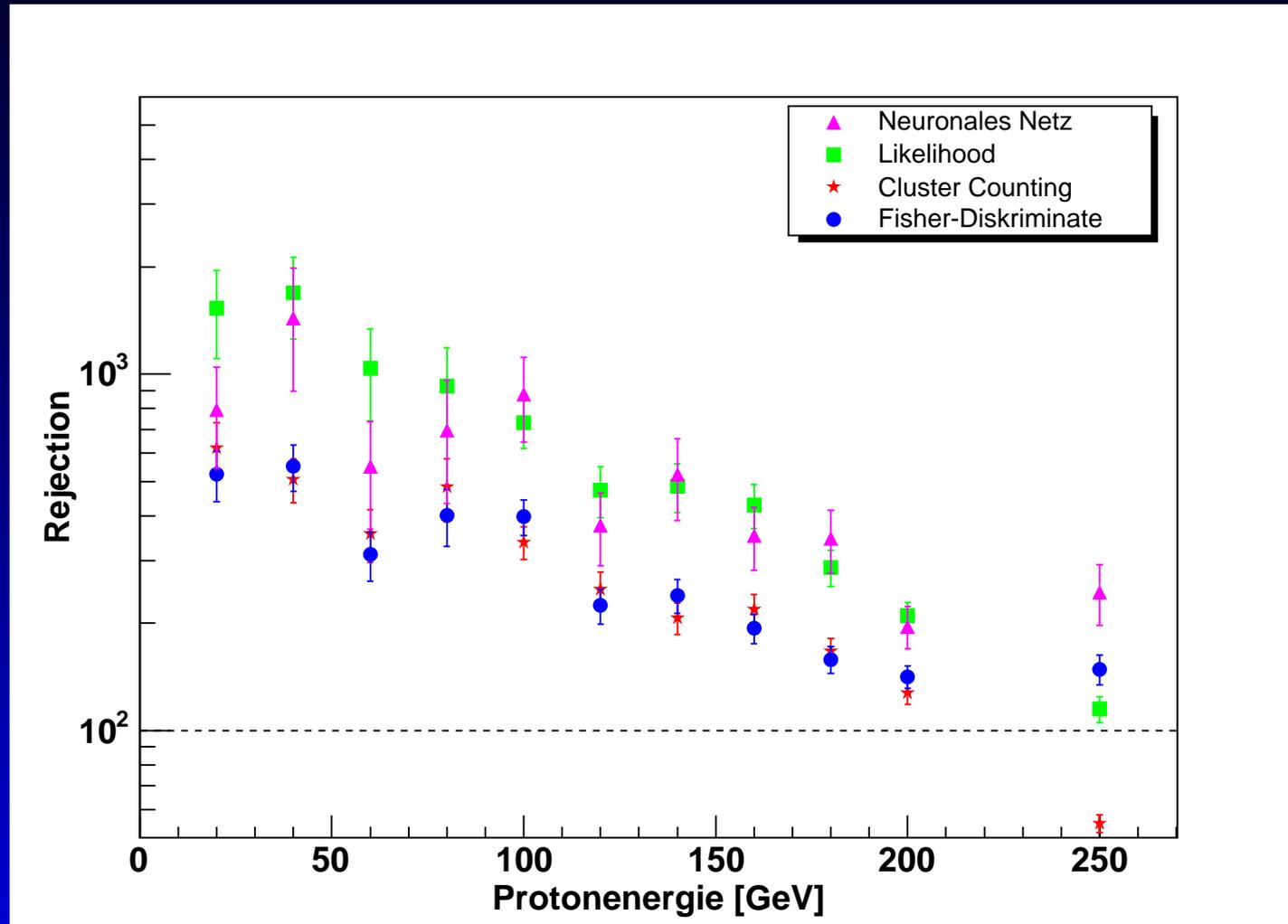
Pionen 60 GeV

Myonen 100 GeV

- ★ Falscherkannte Events sind evtl. Kontamination der Teilchenstrahlen.
- ★ bei Elektronenstrahl evtl. Myon-Verunreinigung
- bei Protonenstrahl evtl. Pion-Verunreinigung



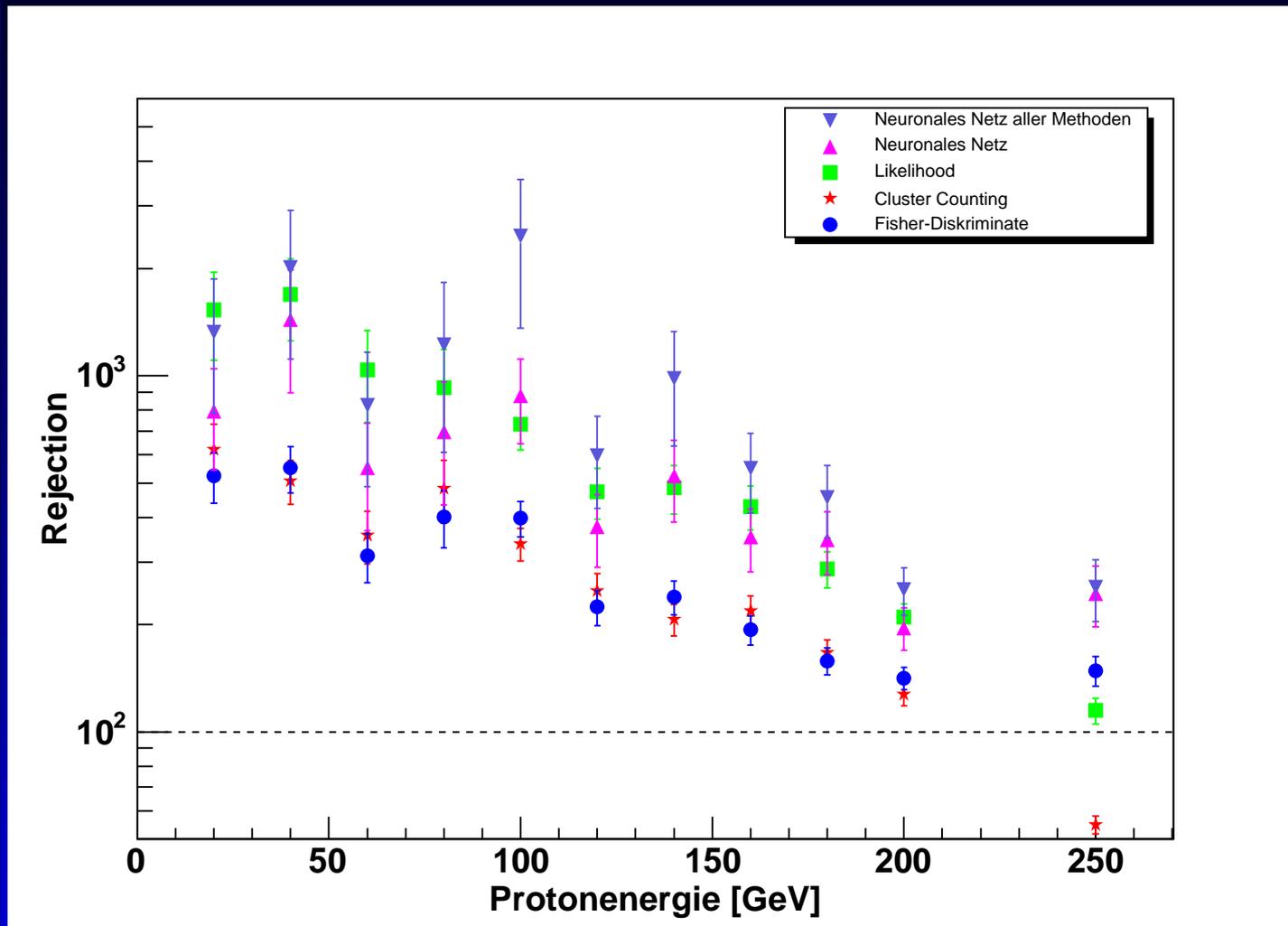
Training eines NN für alle Energien



- ★ Ein einziges trainiertes Netz liefert bei allen Energien einen Wert von über 100 für die Rejection!



Weiteres NN aller Analysemethoden

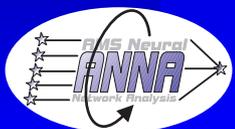


★ Eingabe von nur 5 Größen:

- ★ mittlere Energie
- ★ Anzahl von Hits über bestimmter Energie
- ★ NN-Wert
- ★ LH-Wert
- ★ Fisher-Wert

★ Ziel: Ausnutzen von allen individuellen Vorteilen der Methoden.

⇒ Weitere Verbesserung der Rejection!



Zusammenfassung und Ausblick

- ★ Klassische Analyse:
Cluster-Counting, Fisher-Diskriminante und Likelihood-Methode
- ★ Neuronale Netze sind weitere gute Analysemöglichkeit.
- ★ NN-Analyse hat bei allen Energien eine Rejection über 100!
- ★ Eingabe von MCs um die Statistik zu erhöhen.

