

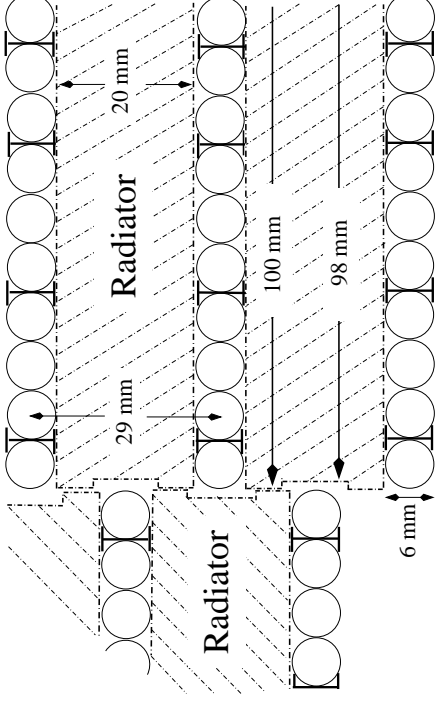
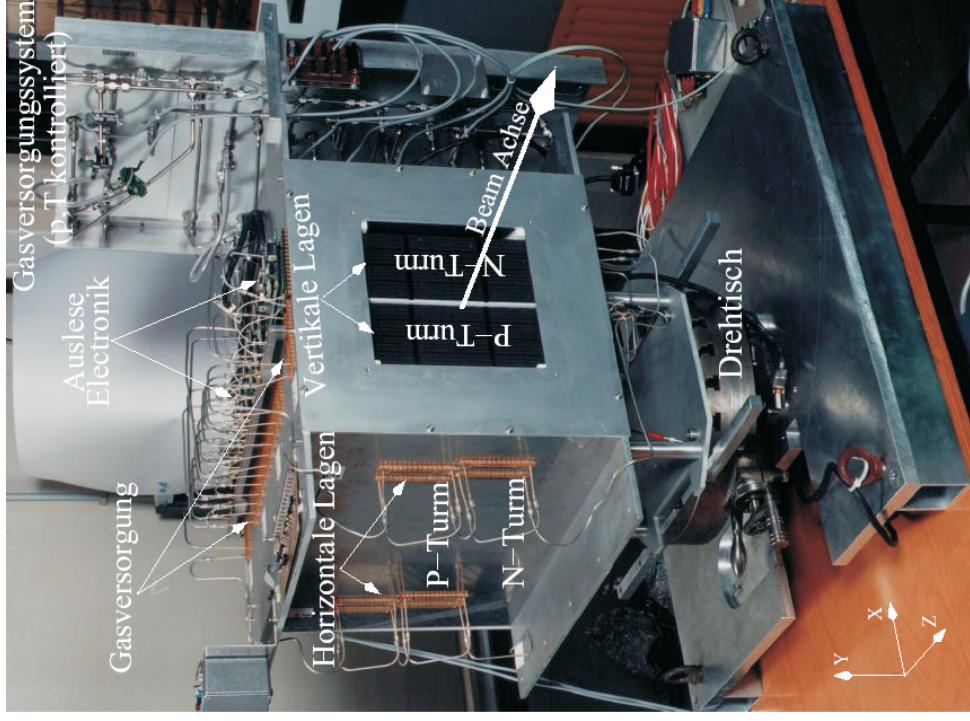


Positron Identifikation mit Hilfe von Neuronalen Netzen mit dem Übergangsstrahlungsdetektor von AMS-02

Philip von Doetinchem

RWTH Aachen, 1. Physikalisches Institut b

AMS und 20-lagiger TRD-Prototyp



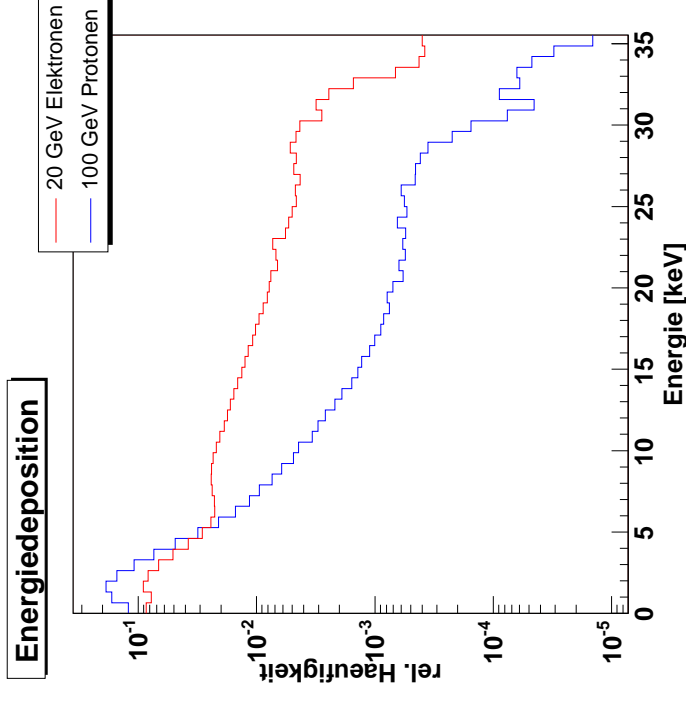
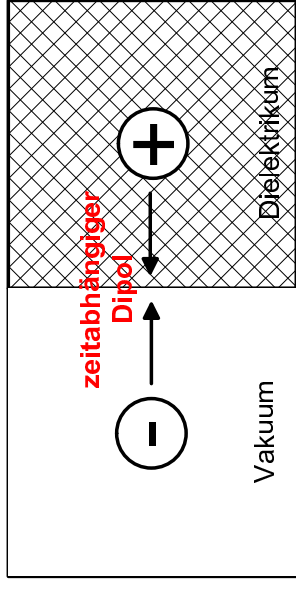
- Auswertung von Beamtests aus dem Jahr 2000
- e^- : 20 GeV
- p^+ : 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250 GeV

Übergangsstrahlung

- Prinzip der Übergangsstrahlung
- Wahrscheinlichkeit der Übergangsstrahlung:

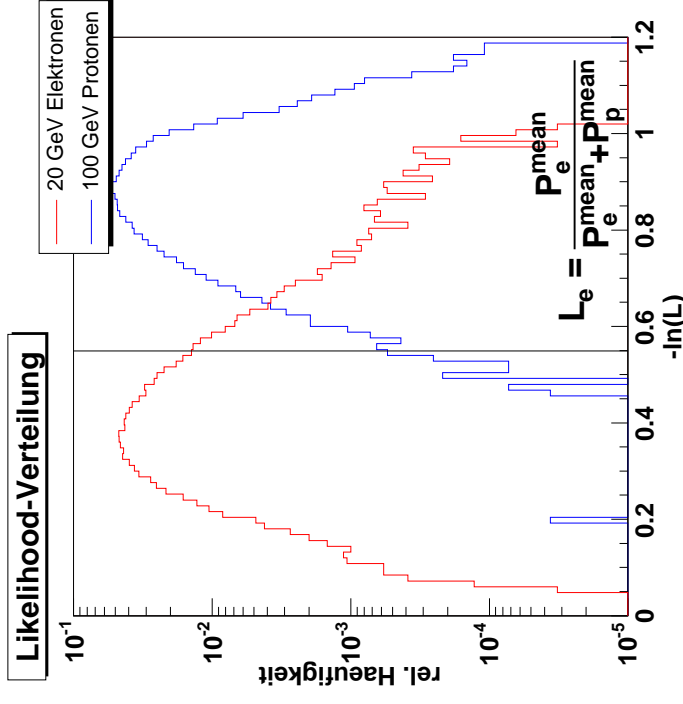
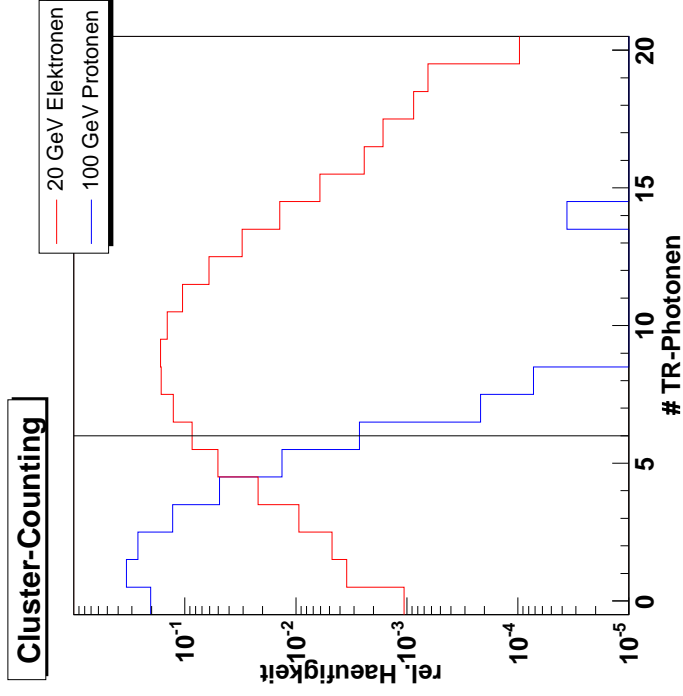
$$W \propto \gamma = \frac{E}{m}$$

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



Bisherige Analysen

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



Kurze Einführung in Neuronale Netze

● Backpropagation-Lernregel:

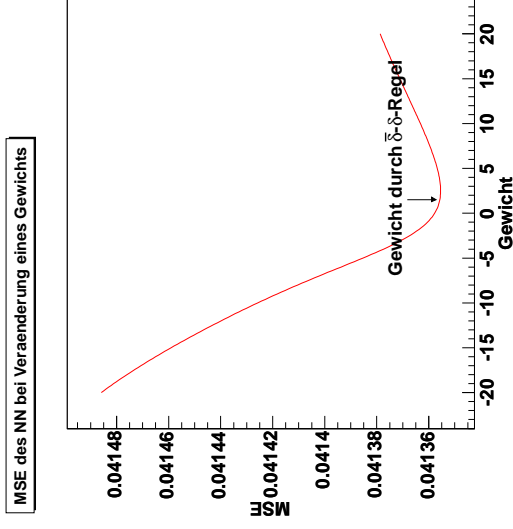
$$F(\vec{w}) = \frac{1}{N} \sum_{k=1}^N \underbrace{(E(\vec{x}_k) - o_{\text{Netz}}(\vec{x}_k, \vec{w}))^2}_{f_k(\vec{w})}$$

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

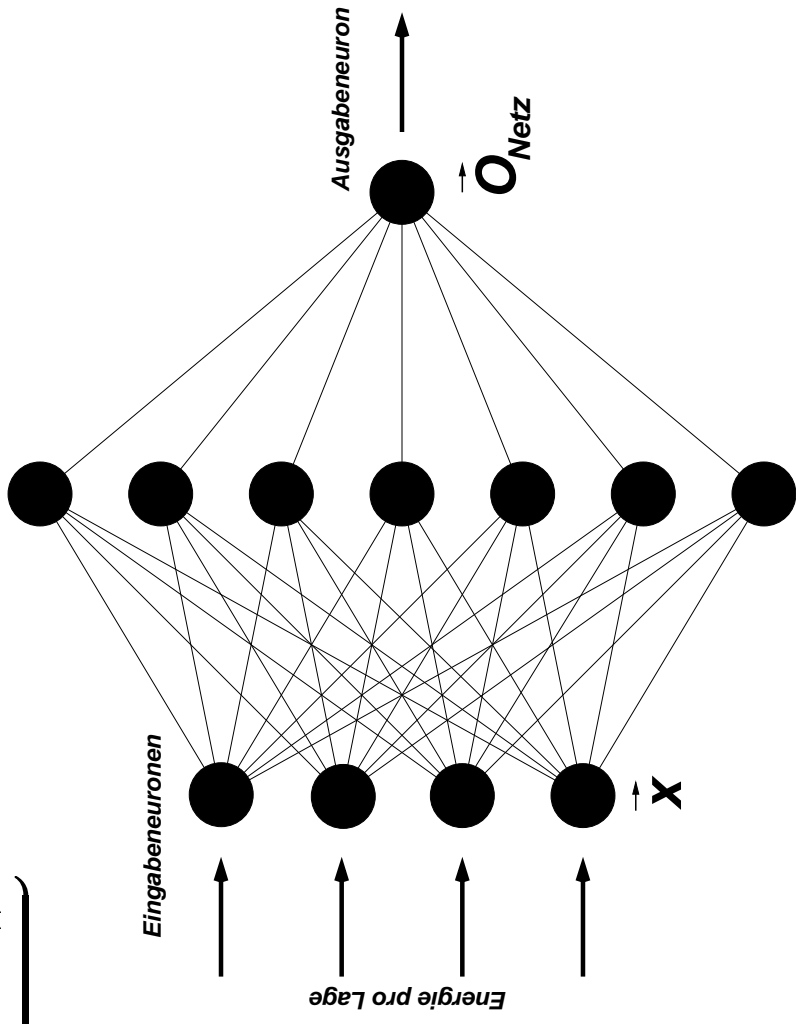
$$\Delta w_i(t+1) = \eta \cdot \frac{1}{N} \sum_{k=1}^N \frac{\partial f_k(\vec{w})}{\partial w_i}$$

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

bei Variation eines Gewichtes:



mittlere Neuronen mit sigmoidem Output



Verbesserung des Backpropagation

- $\bar{\delta}$ - δ -Lernregel, Einführung individueller Lernraten:

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

α ist das sog. Moment. Es verhindert zu schnelle Richtungswechsel.

- Veränderung der Lernrate:

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

Aufgaben des Neuronalen Netzes

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

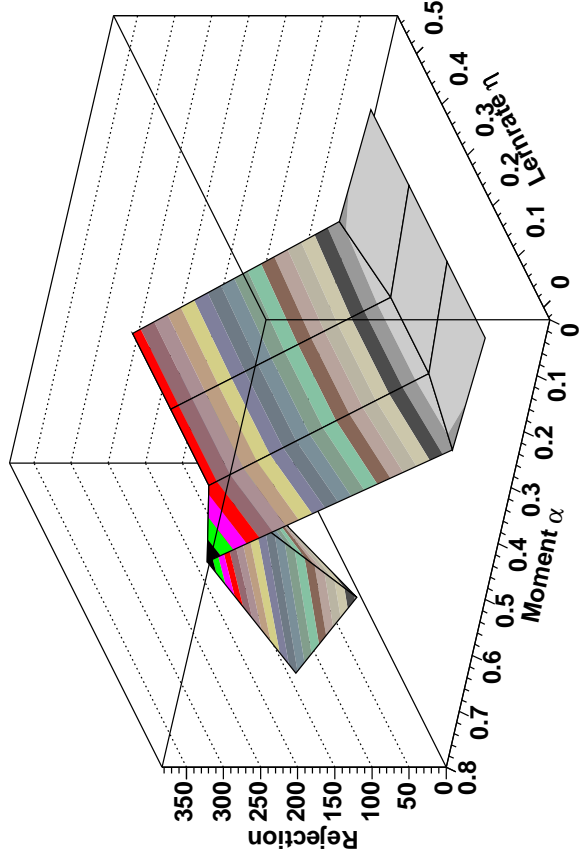
- Finden einer Abbildungsvorschrift:

$$\left. \begin{array}{l} \text{Elektronen} \rightarrow 1 \\ \text{Protonen} \rightarrow 0 \end{array} \right\} \rightarrow \text{Unterscheidungskriterium}$$

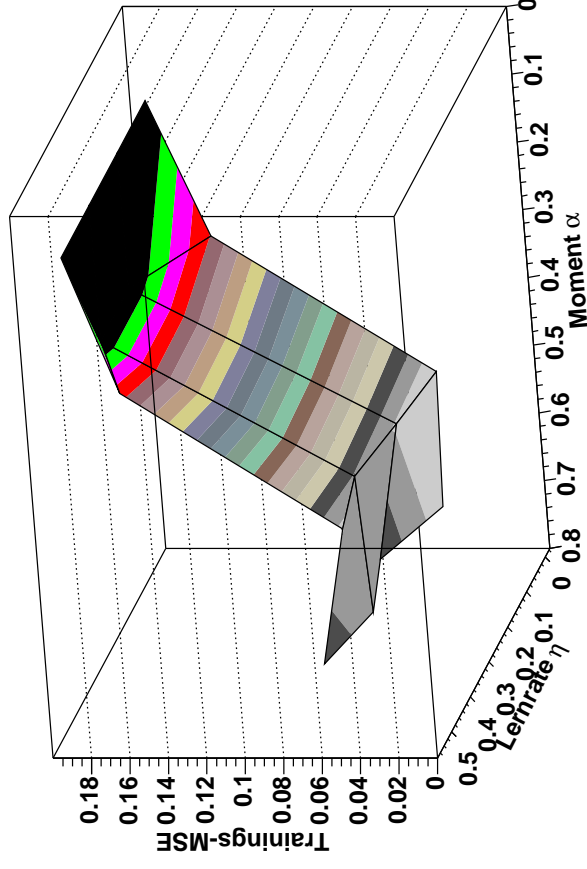
- Rejection soll besser als **100** sein.
- Auffinden der besten Konfiguration des NN durch Variation der Parameter.
- Untersuchung der Energiedepositionen pro Lage auf mögliche Korrelationen.

Architektur und Parameterwahl

Rejection, 100 mittlere Neuronen, Gauss-foermige Gewichtsini., Abbruch nach 1500 Zyklen



Trainings-MSE, 100 mittlere Neuronen, Gauss-foermige Gewichtsini., Abbruch nach 1500 Zyklen

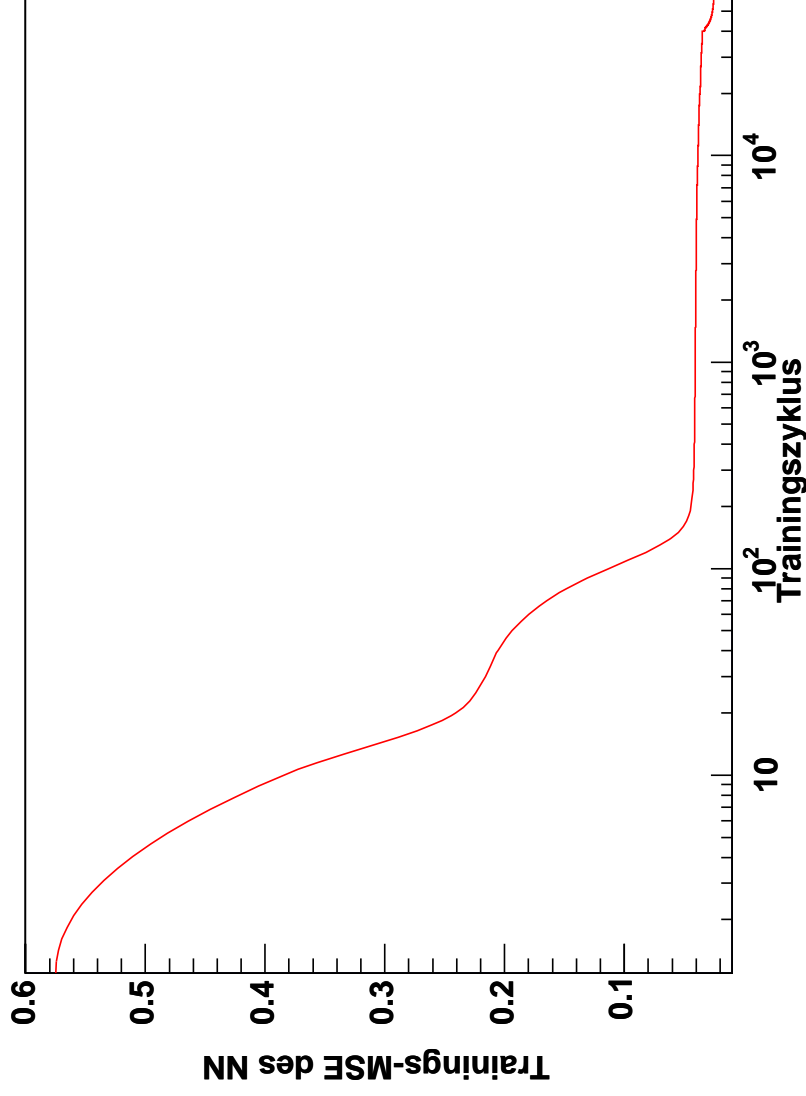


● Auffinden der richtigen Lernparameter ist stark problemabhängig!

Verlauf des MSE

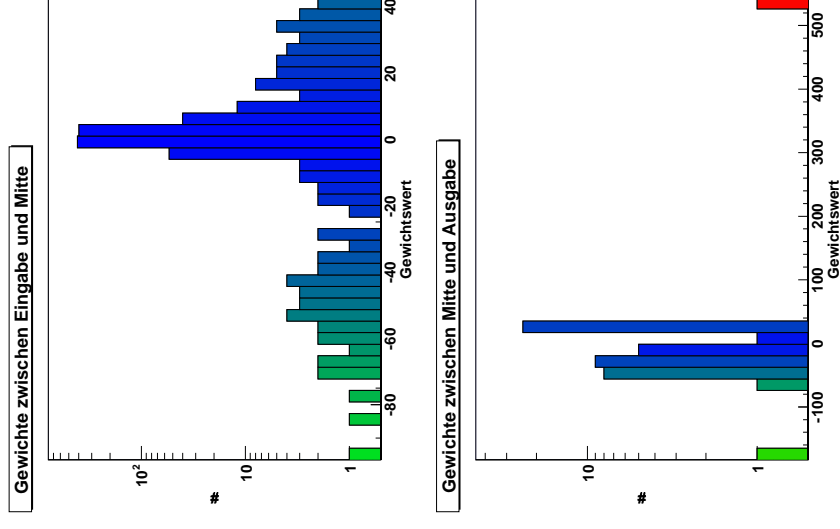
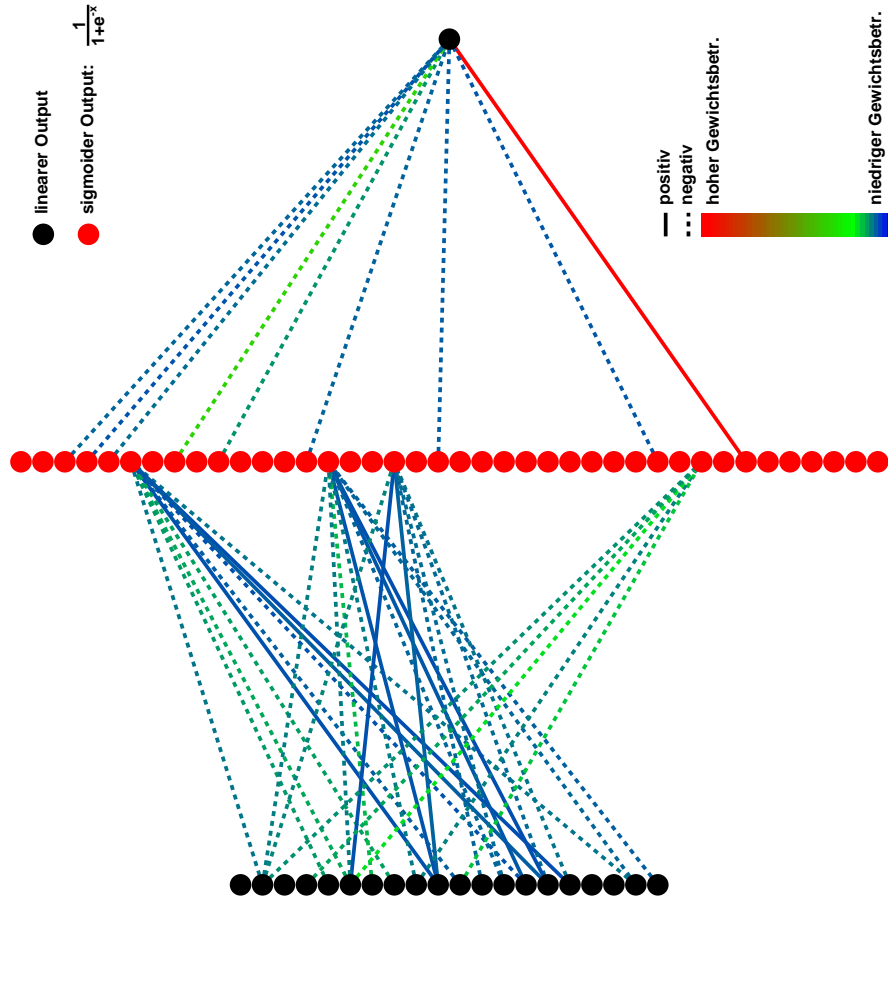
- MSE-Verlauf des bisher besten Netzes

Trainings-MSE des NN gegen Trainingszyklen



Struktur des Netzes

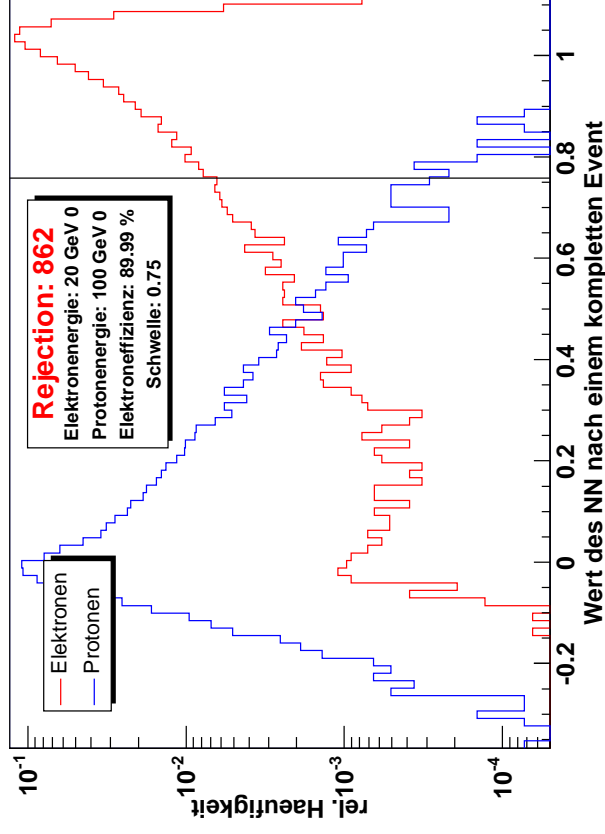
● Gewichte nach dem Training



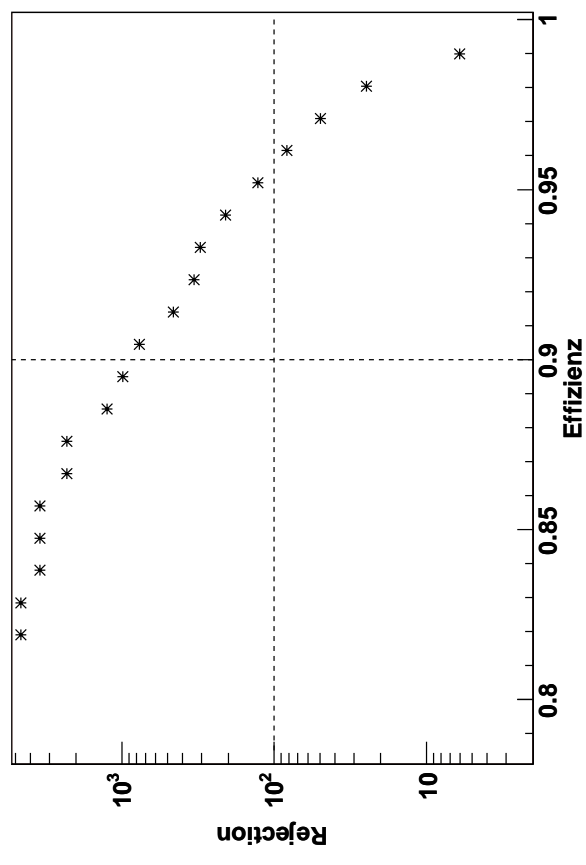
Rejection

Trennung von Elektronen und Protonen nach Durchlauf durch das Netz

Verteilung nach Durchlauf durch trainiertes NN

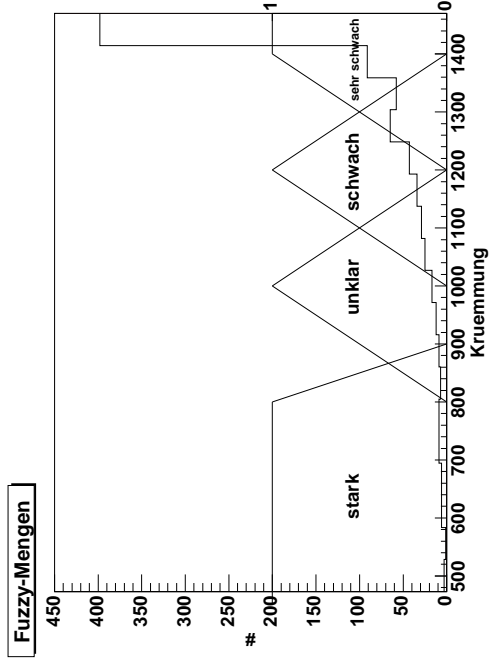


Effizienz gegen Rejection



Weitere Möglichkeiten zur Verbesserung

- Fuzzy-Controller für Vergrößerung der Lernrate und Moment



- Die Krümmung berechnet sich nach der Häufigkeit der Gradientenwechsel
- Lernrate und Moment werden in flachen Bereichen größer und in gekrümmten Gebieten kleiner.

- SVD-LM-Lernregel:

$$F(\vec{w}) = \|\vec{\varphi}(\vec{w})\|^2 \quad \wedge \quad \vec{\varphi}(\vec{w}(t+1)) = \vec{\varphi}(\vec{w}(t)) + \mathbf{J}(\vec{w}(t+1) - \vec{w}(t))$$

$$\Delta \vec{w}(t+1) = 2(\mathbf{J}^T \mathbf{J} + \kappa)^{-1} \mathbf{J}^T \cdot \vec{\varphi}(\vec{w}(t))$$

$$\text{SVD: } \mathbf{J} \rightleftarrows \mathbf{U} \mathbf{S} \mathbf{V}^T \quad \Delta \vec{w}(t+1) = 2\mathbf{V}(\mathbf{S}^2 + \lambda)^{-1} \mathbf{S}^T \mathbf{U}^T \cdot \vec{\varphi}(\vec{w}(t))$$

Ausblick

- Training eines Netzes pro Energie
- Winkeluntersuchungen
- andere Netzstruktur → mehrere mittlere Lagen
- Vergleich mit Monte Carlo-Ereignissen