

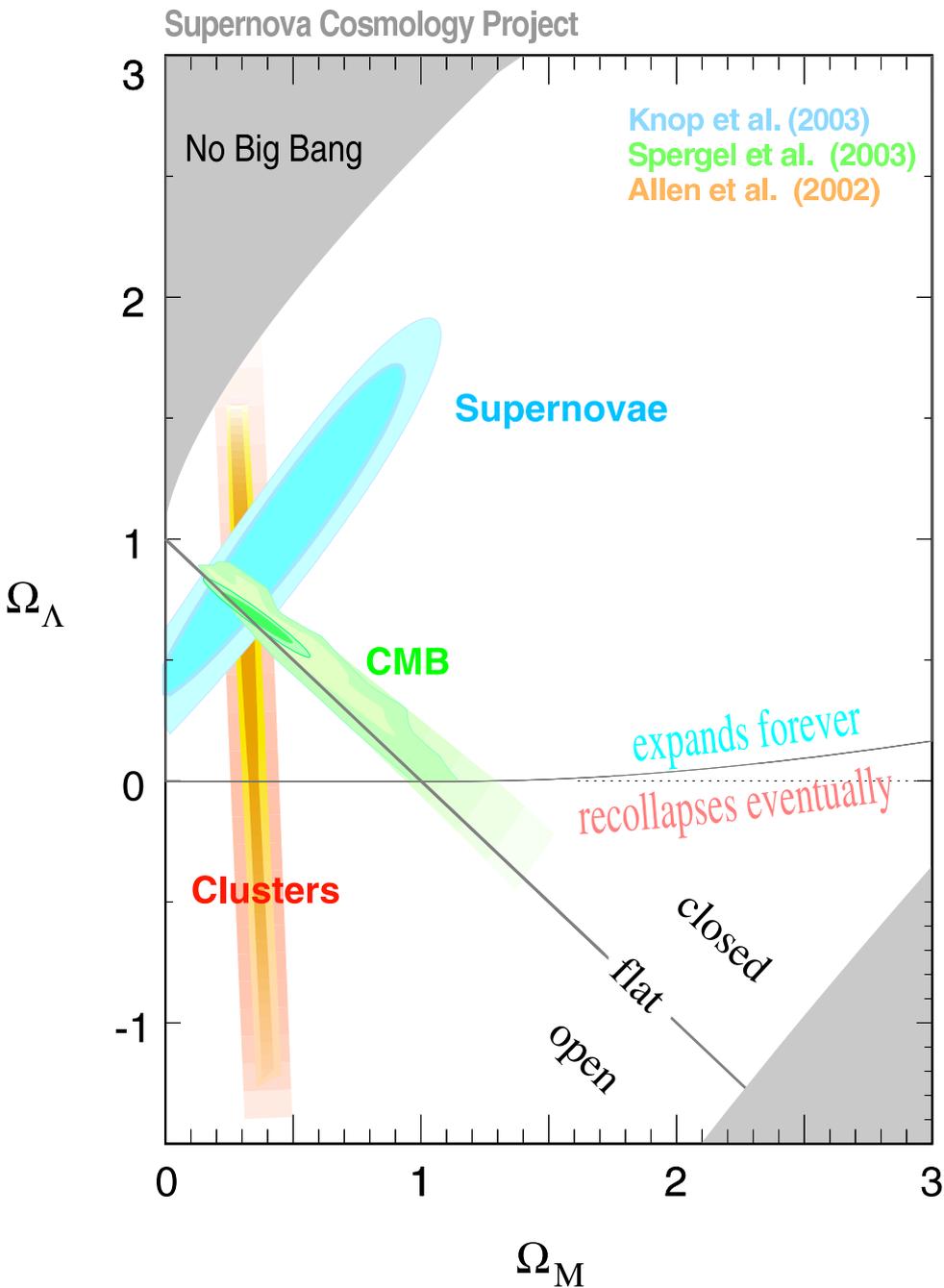
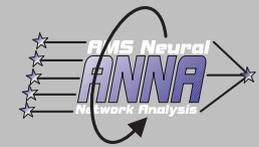


e^+ -Identifikation mit Hilfe von neuronalen Netzen im TRD von AMS-02

Philip von Doetinchem

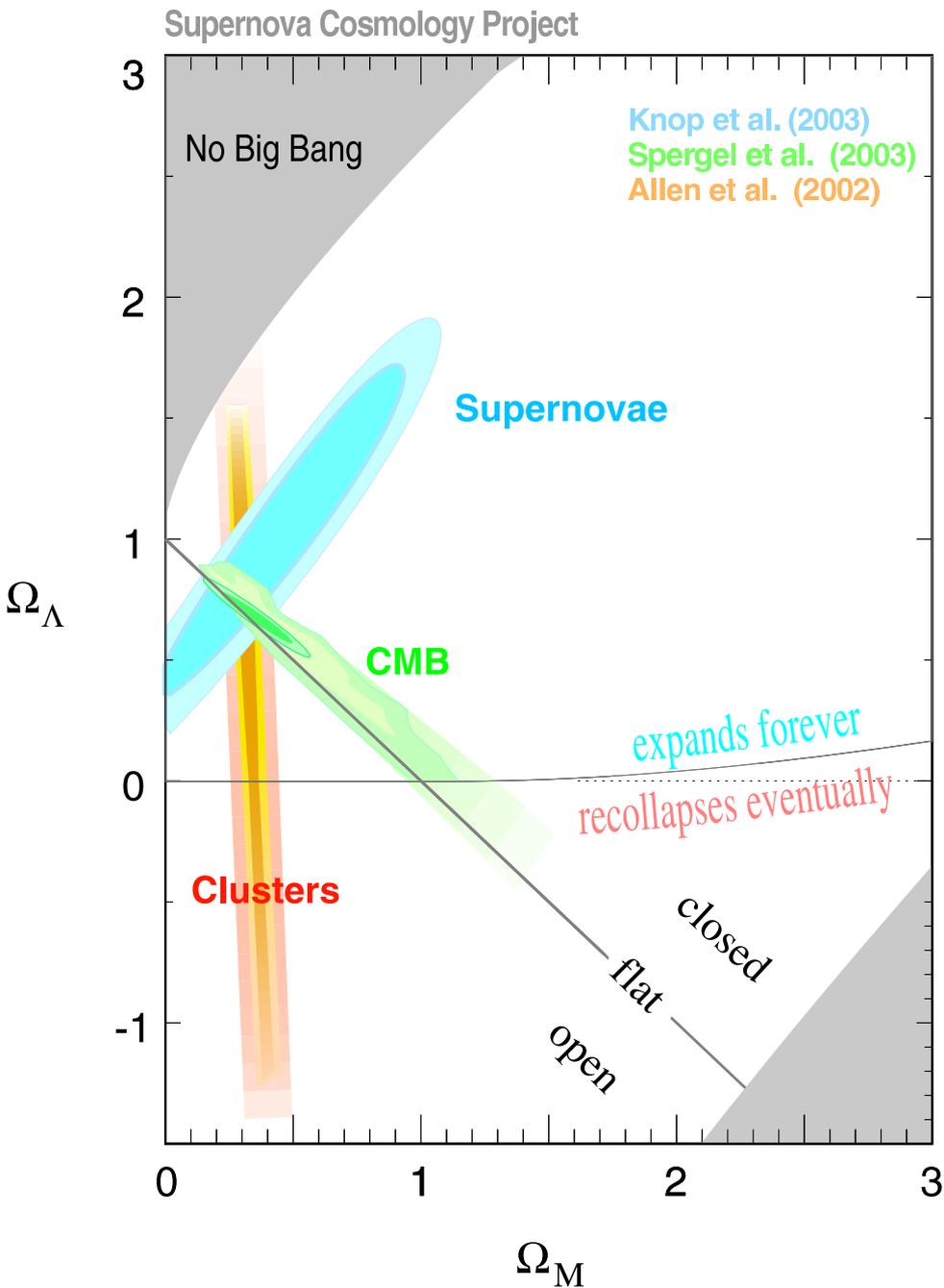
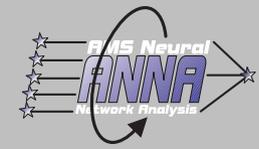
I. Physikalisches Institut B, RWTH Aachen

Motivation für e^+ -Messung



- ☆ verschiedene Messmethoden (Supernovae, Anisotropie des CMB, Galaxie Cluster) liefern Anteil von 22 % am Dichteparameter Ω durch **DUNKLE MATERIE**

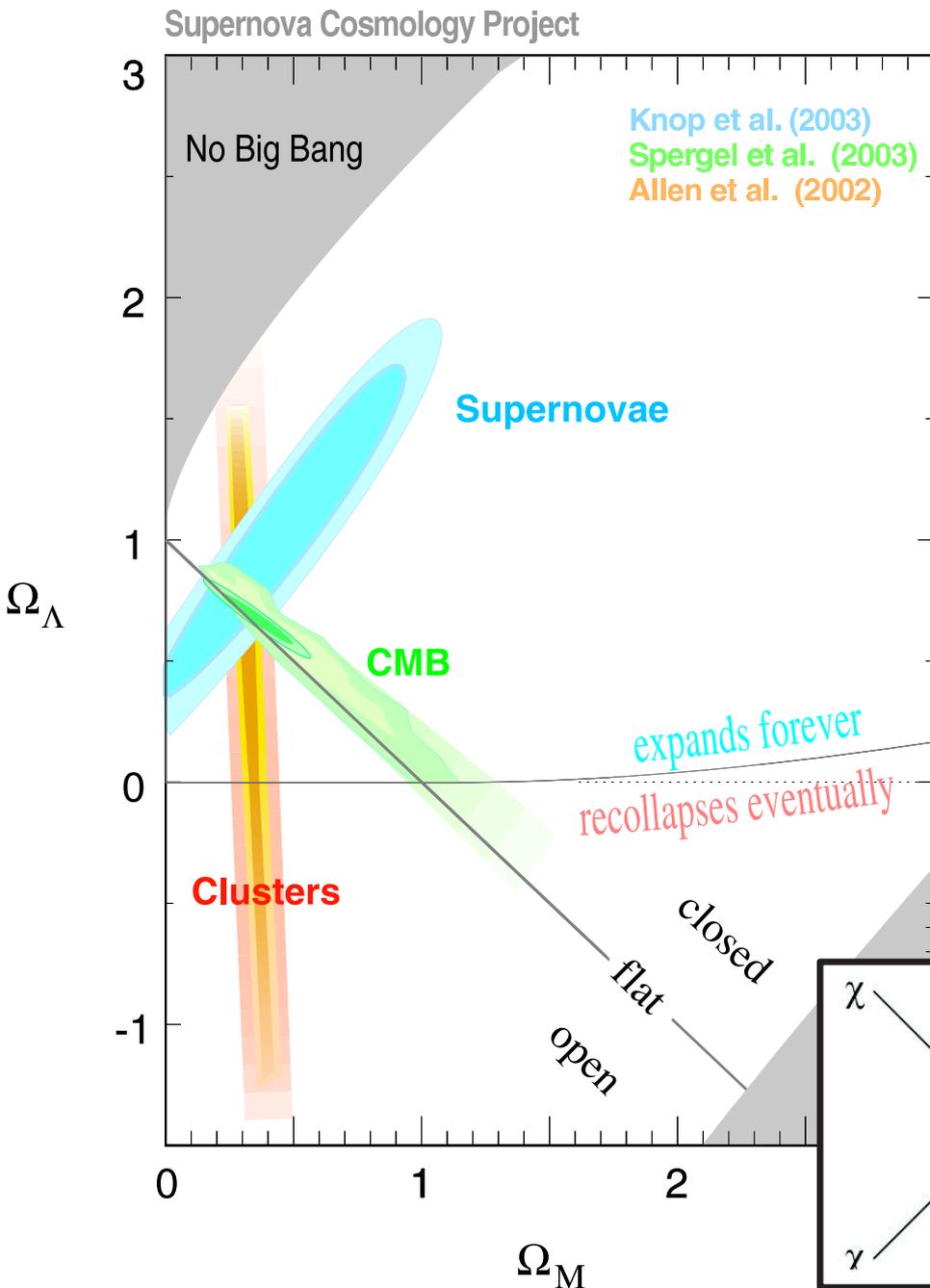
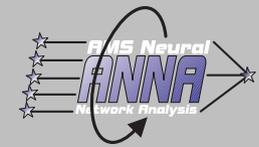
Motivation für e^+ -Messung



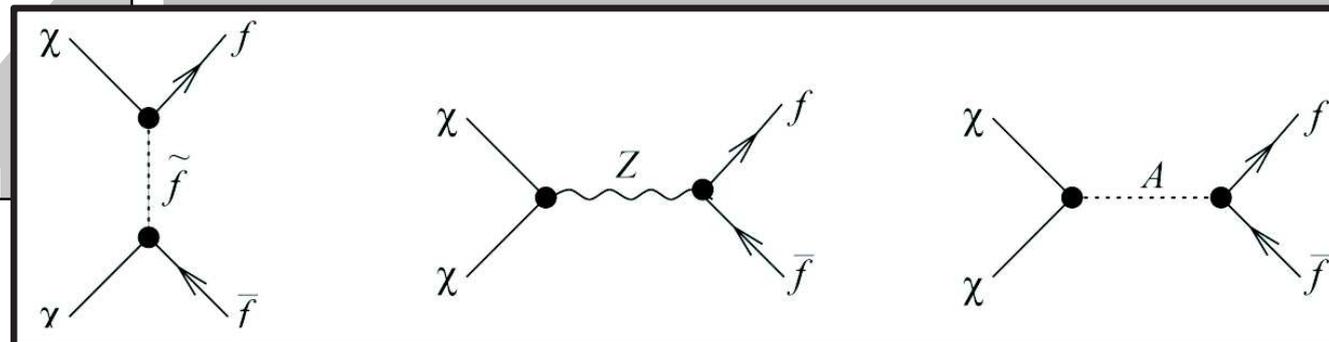
☆ verschiedene Messmethoden (Supernovae, Anisotropie des CMB, Galaxie Cluster) liefern Anteil von 22 % am Dichteparameter Ω durch **DUNKLE MATERIE**

☆ Art der dunklen Materie ist wichtiger Faktor für die Strukturbildung im Universum

Motivation für e^+ -Messung



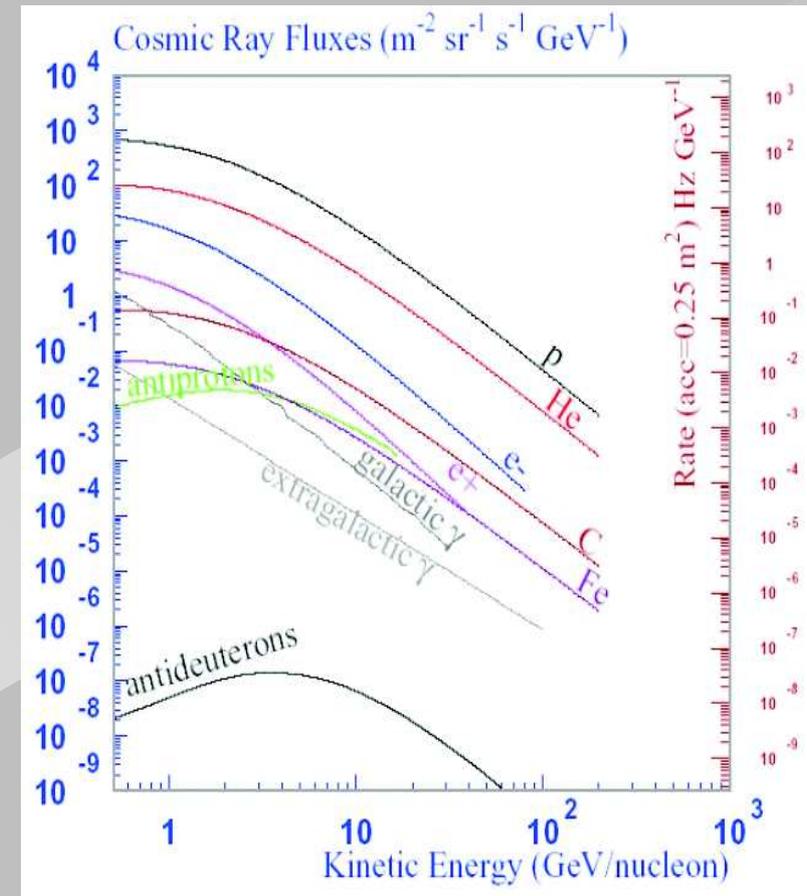
- ★ verschiedene Messmethoden (Supernovae, Anisotropie des CMB, Galaxie Cluster) liefern Anteil von 22 % am Dichteparameter Ω durch **DUNKLE MATERIE**
- ★ Art der dunklen Materie ist wichtiger Faktor für die Strukturbildung im Universum
- ★ Theorien zur Strukturentstehung unseres Universums brauchen sog. **KALTE DUNKLE MATERIE**, guter Kandidat ist das leichteste supersymmetrische Neutralino
- ★ möglicher indirekter Nachweis über Annihilationen in hochenergetische e^+ , da keine anderen Quellen hochenergetischer e^+ im Universum bekannt sind



- ★ Verhältnis von Protonen und Positronen in kosmischer Strahlung:

$$\frac{p^+}{e^+} \approx 10^4$$

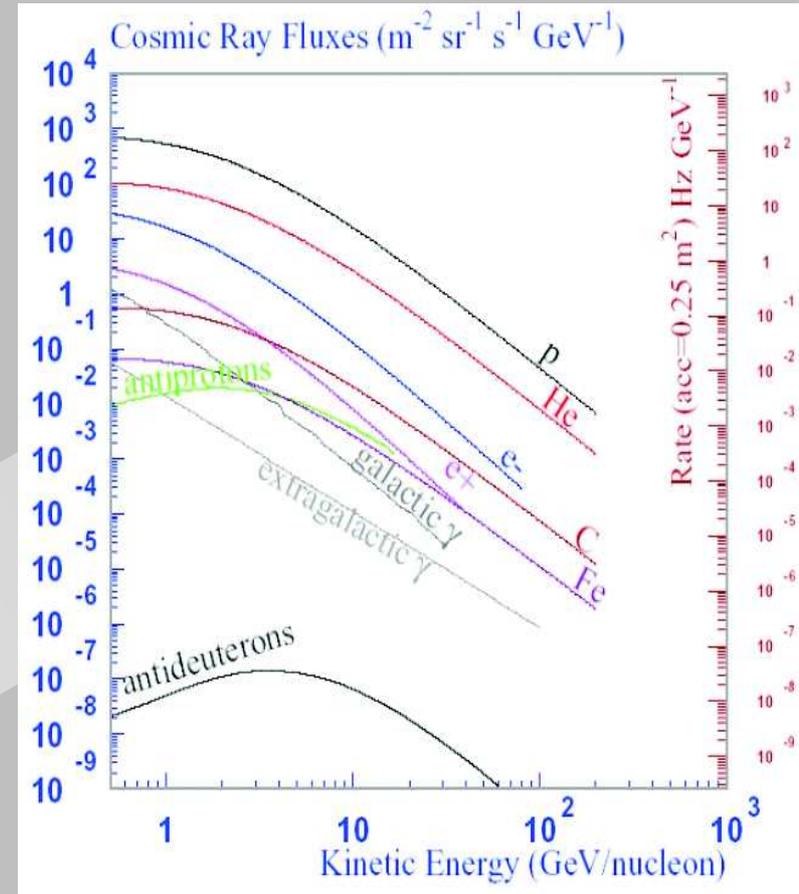
⇒ bei Messung muss eine Trennung von p^+ und e^+ von ca. 10^6 erreicht werden.



- ★ Verhältnis von Protonen und Positronen in kosmischer Strahlung:

$$\frac{p^+}{e^+} \approx 10^4$$

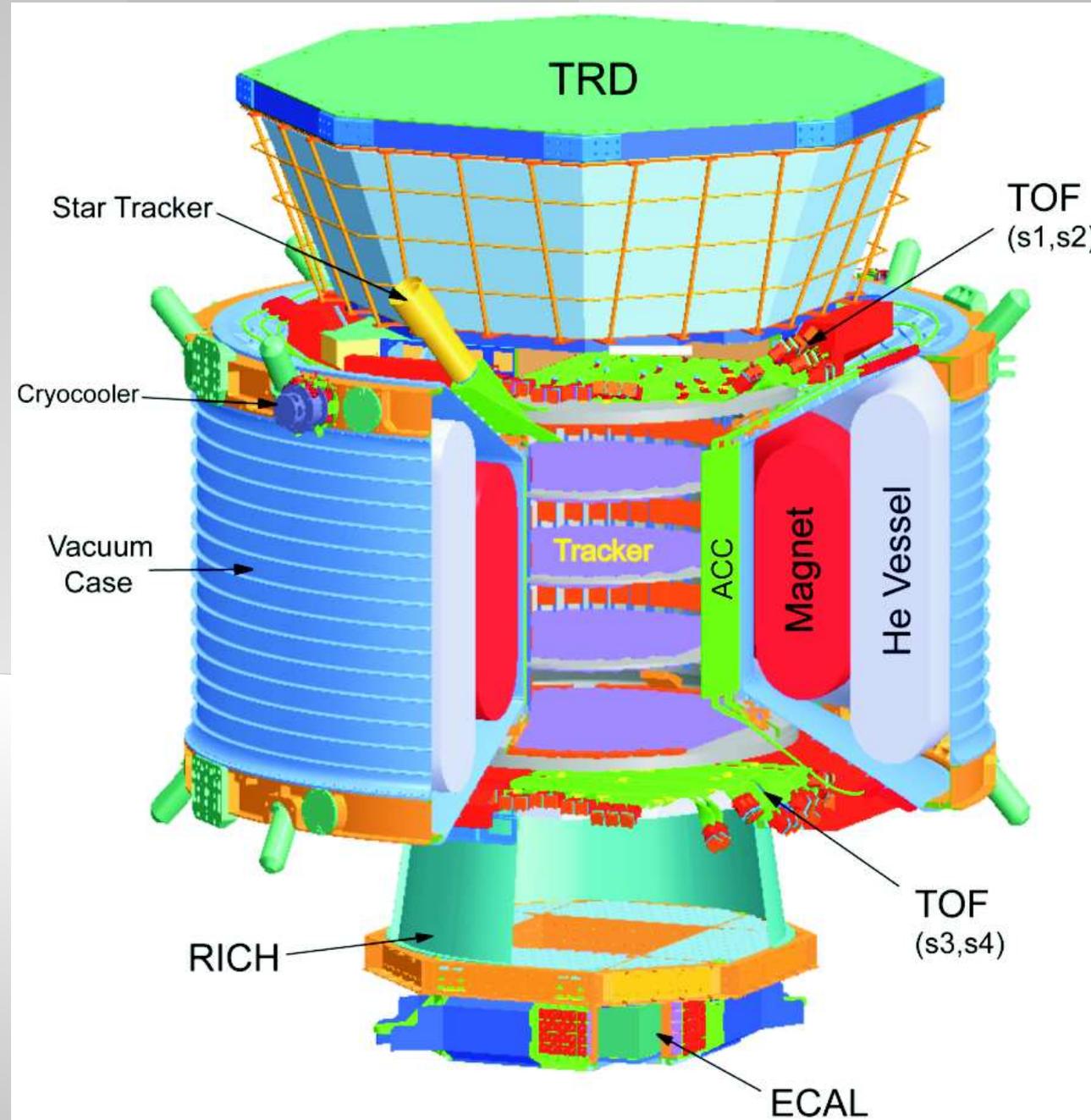
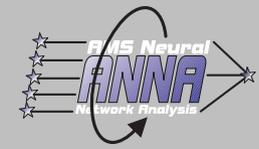
⇒ bei Messung muss eine Trennung von p^+ und e^+ von ca. 10^6 erreicht werden.



- ★ **Effizienz:** Anteil richtig erkannter Teilchen einer Methode (hier: 90 %) ⇒ definiert Cut

- ★ Der Kehrwert des Anteils falsch erkannter Teilchen, die über Cut liegen: **Rejection**

AMS-02 Detektor



☆ **Transition Radiation Detector (TRD)**

Teilchenidentifikation anhand des Lorentzfaktors: $\gamma = E/m$

☆ **Time of Flight (ToF)**

Messung von Geschwindigkeit, Richtung und Ladung (dE/dx)

☆ **Tracker**

Spurrekonstruktion von geladenen Teilchen

☆ **Anti-Coincidence Counter (ACC)**

Teilchenselektion mit ToF

☆ **Ring Image Cherenkov Detector (RICH)**

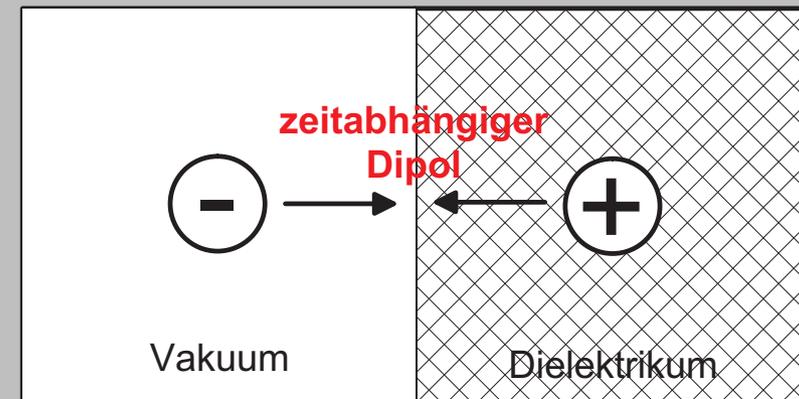
präzise Geschwindigkeitsmessung

☆ **Electromagnetic Calorimeter (ECAL)**

Trennung von hochrelativistischen Leptonen und Hadronen

Trennung durch Übergangsstrahlung (TR)

- ★ Stetigkeit der em-Felder an Grenzflächen sorgt für Strahlung
- ★ mittlere Energie durch Übergangsstrahlung (**ladungsunabhängig**):



$$E \propto \alpha \cdot \gamma = \alpha \cdot \frac{E}{m}$$

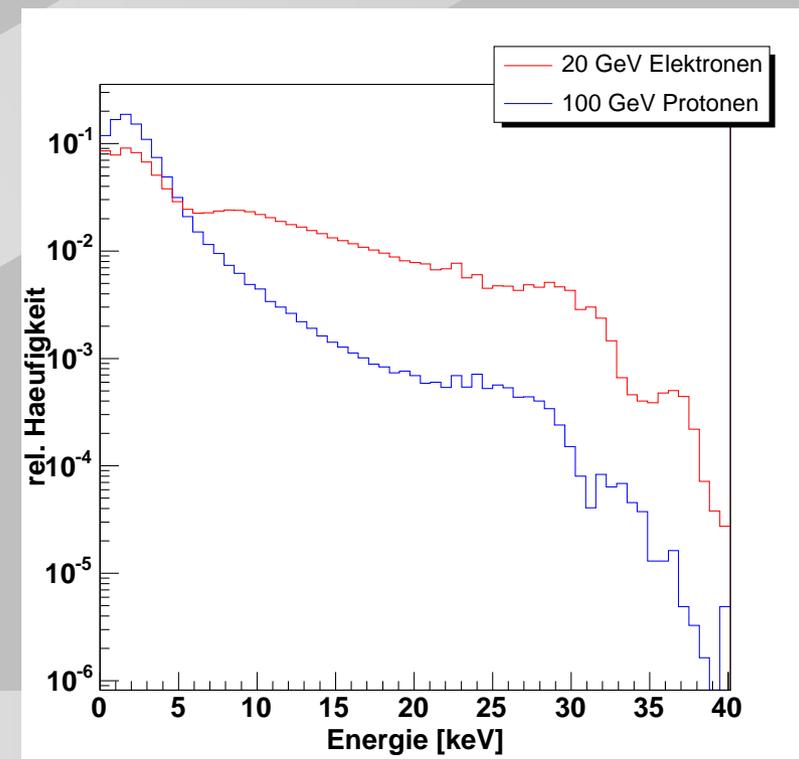
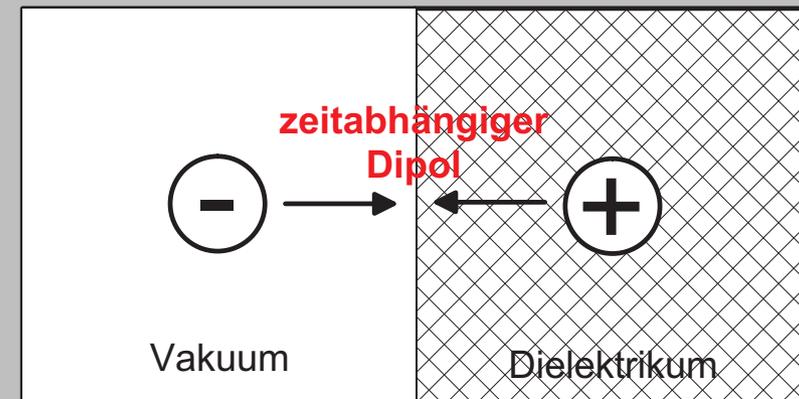
Trennung durch Übergangsstrahlung (TR)

★ Stetigkeit der em-Felder an Grenzflächen sorgt für Strahlung

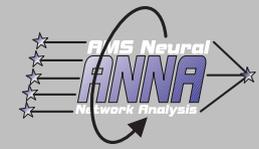
★ mittlere Energie durch Übergangsstrahlung (**ladungsunabhängig**):

$$E \propto \alpha \cdot \gamma = \alpha \cdot \frac{E}{m}$$

★ höhere mittlere Energie der TR für Elektronen dient zur Unterscheidung



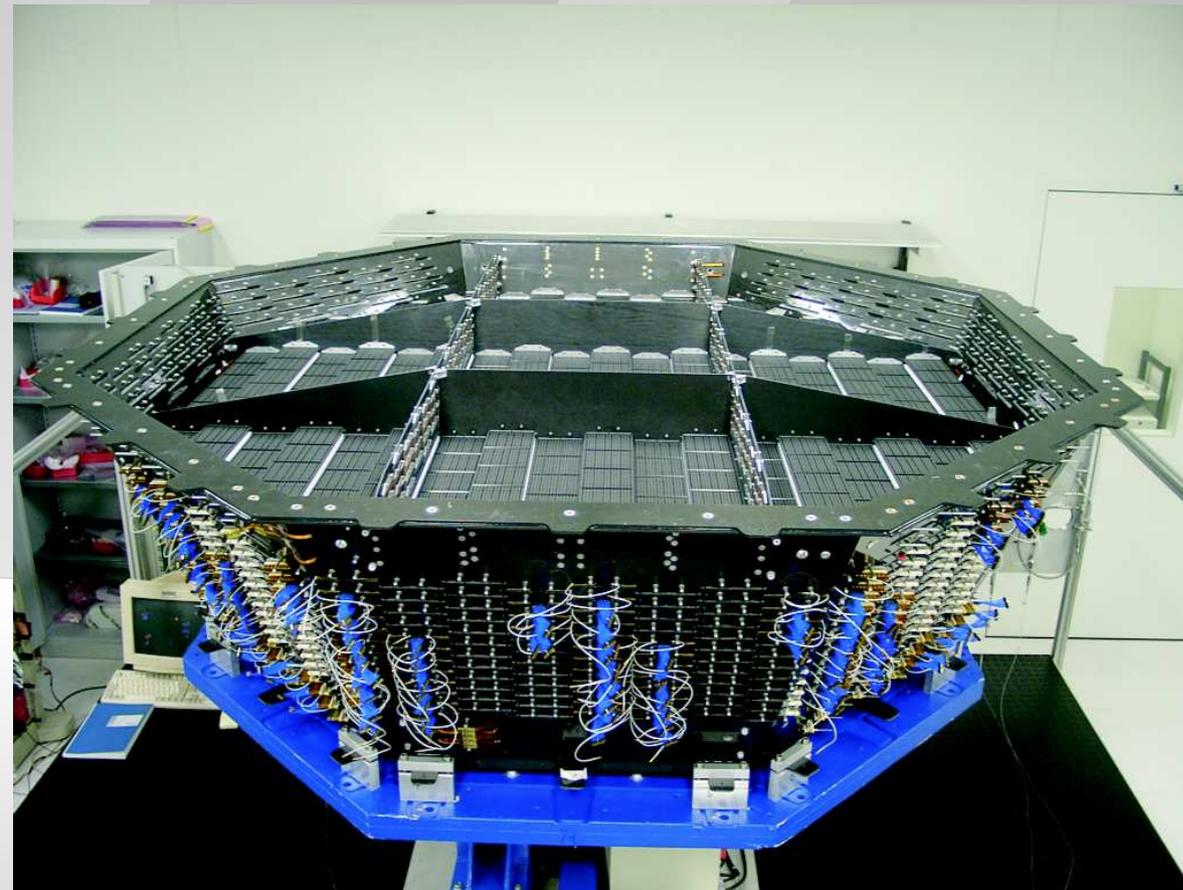
Finaler AMS-02 TRD



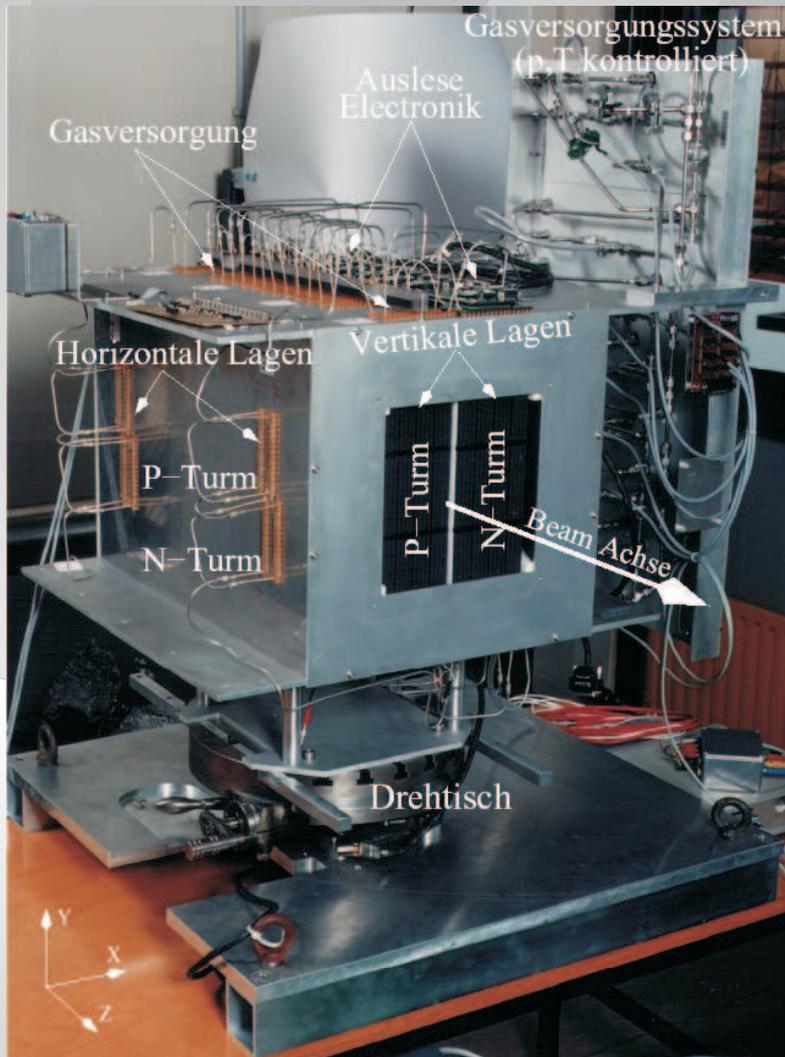
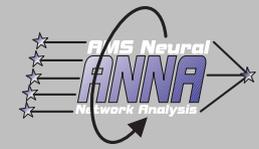
Finales TRD-Design:

- ★ 20 Lagen aus 328 Modulen mit je 16 Proportionalkammern
- ★ gekreuzte Lagen zur 3D-Spurrekonstruktion
- ★ Gas: Xe/CO₂ im Verhältnis 80/20
- ★ zwischen den Lagen ATLAS Radiatorvlies
- ★ besondere Herausforderung: Präzisionsexperiment, was die Belastungen eines Shuttlestarts aushalten muss

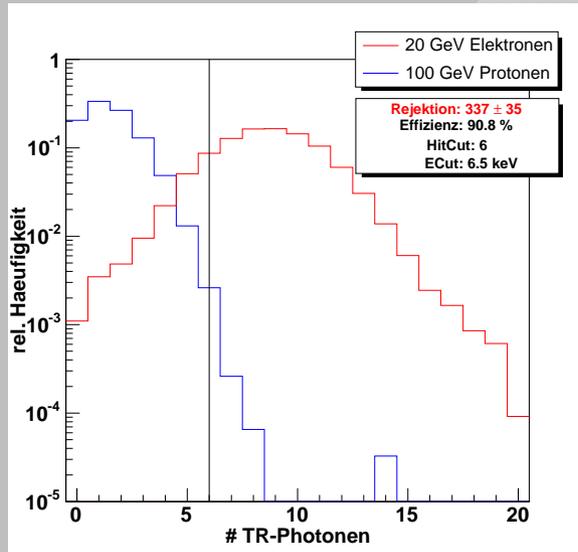
Ziel: Teilchenidentifikation (z.B. Proton Rejection bis zu 10^3 für Positronidentifikation)



20-lagiger TRD-Prototyp



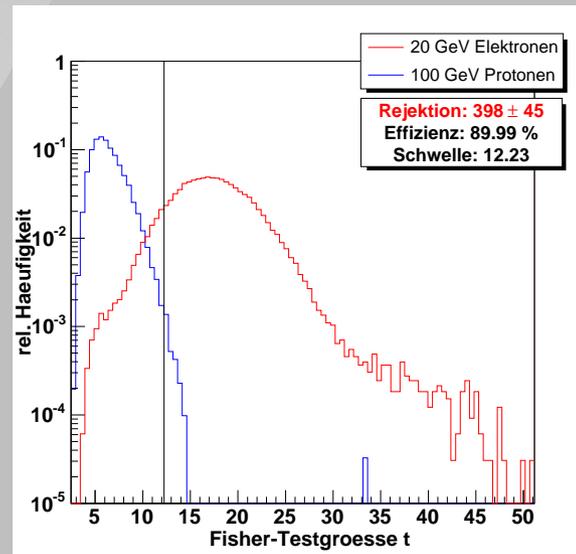
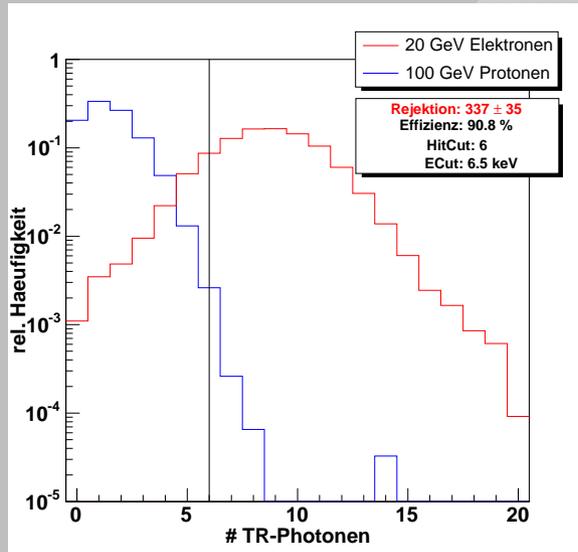
- ★ 20 Lagen, damit Wahrscheinlichkeit der Entstehung und Detektion der Übergangsstrahlung steigt.
- ★ Zwischen den Lagen:
Radiatormaterial
- ★ Auswertung von Beamttests vom SPS am CERN aus dem Jahr 2000
 - e^- : 20 GeV
 - p^+ : 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 250 GeV



Cluster-Counting



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



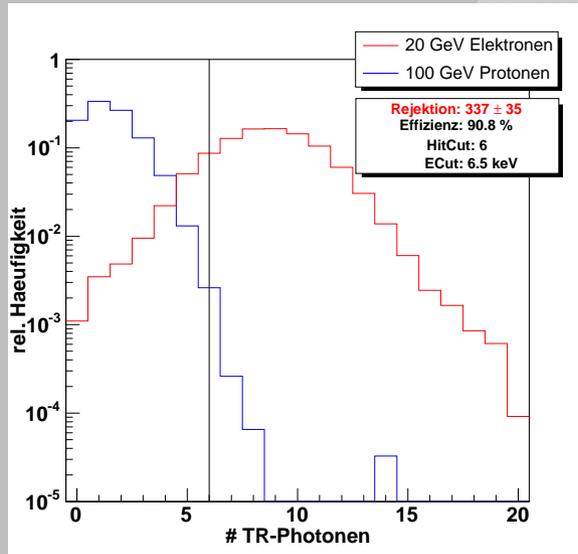
Cluster-Counting

★ **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.

Fisher-Diskriminante



Cluster-Counting

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

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

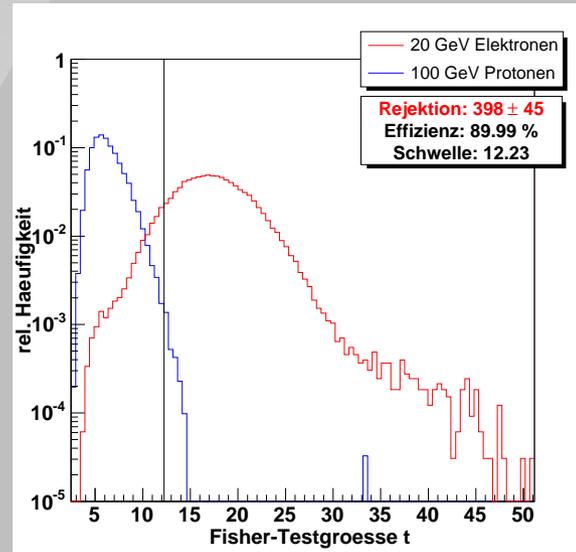
Es kann gezeigt werden, dass der Algorithmus bei Normalverteilung optimal ist.

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

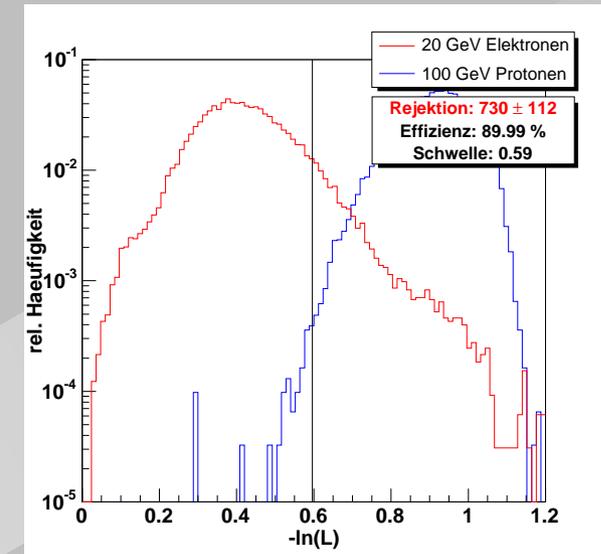
Likelihood-Ratio:
$$L = \frac{\bar{P}_e}{\bar{P}_p + \bar{P}_e}$$

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

e⁺-Identifikation mit Hilfe von neuronalen Netzen im TRD von AMS-02



Fisher-Diskriminante



Likelihood-Analyse

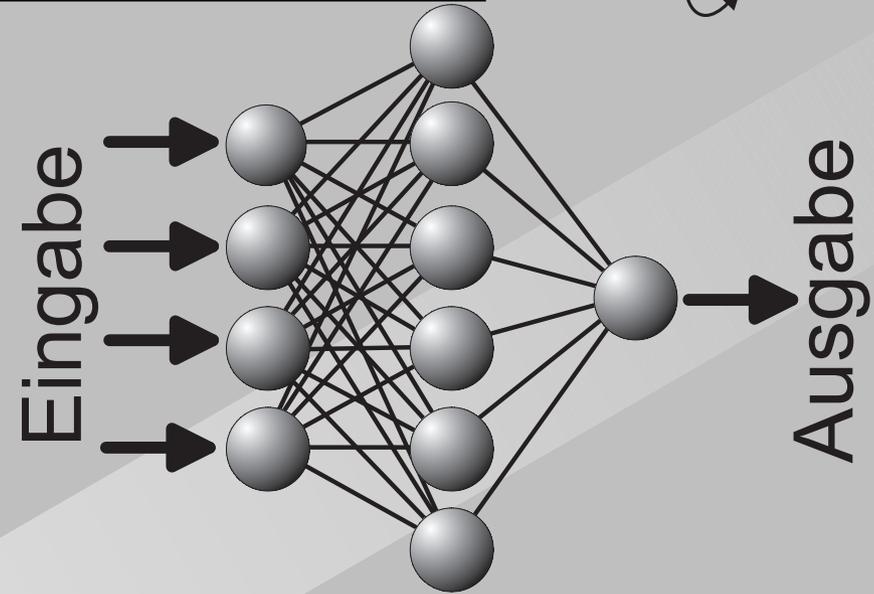
Kurze Einführung in Neuronale Netze (NN)



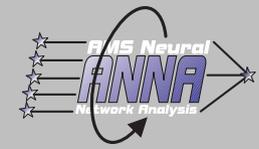
$\bar{\delta}$ - δ -Backpropagation-Lernregel:

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

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



Kurze Einführung in Neuronale Netze (NN)

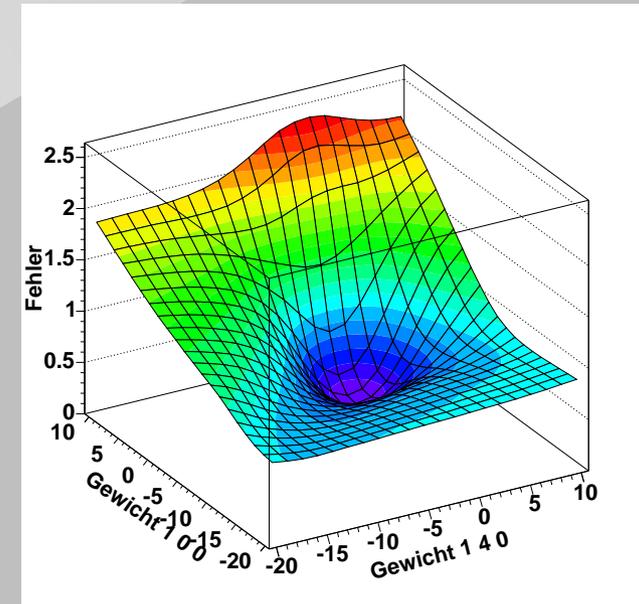
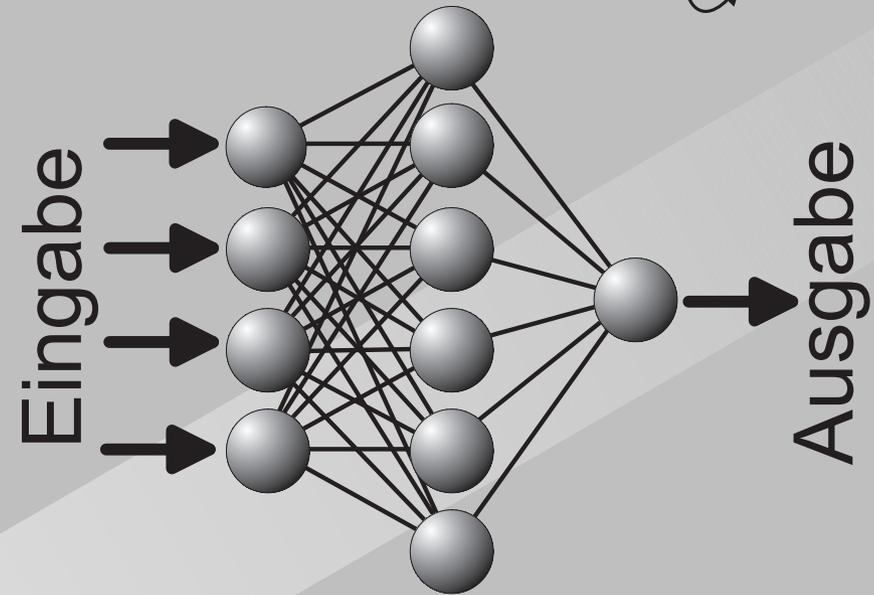


$\bar{\delta}$ - δ -Backpropagation-Lernregel:

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

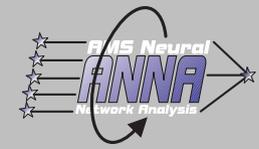
$$\vec{w}(t+1) = \vec{w}(t) - \Delta \vec{w}(t+1)$$

Ziel: Training bis zum minimalen Fehler!



Verlauf des Mean Square Errors ($F(\vec{w})$)
bei Variation eines Gewichtes

Kurze Einführung in Neuronale Netze (NN)



$\bar{\delta}$ - δ -Backpropagation-Lernregel:

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

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

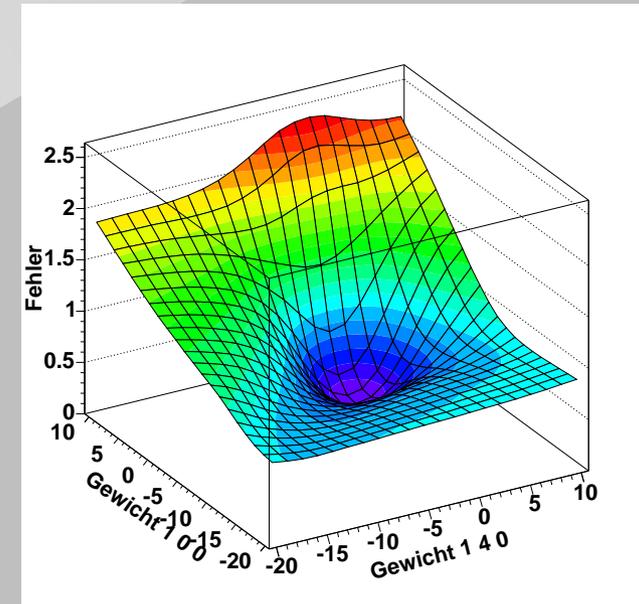
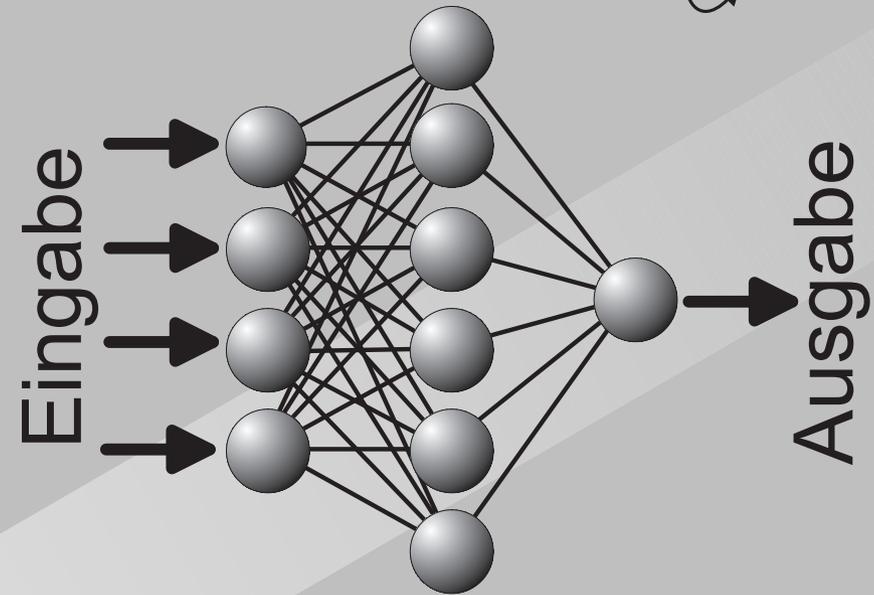
Ziel: Training bis zum minimalen Fehler!

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

η_i : individuelle Lernrate, α : Moment

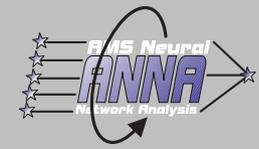
Veränderung der Lernrate:

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



Verlauf des Mean Square Errors ($F(\vec{w})$) bei Variation eines Gewichtes

Teilchenidentifikation mit NN



- ★ Anwendung von NN zur Separation von Teilchensorten mit folgender Abbildungsvorschrift:

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

- ★ Eingabe in das NN:
 - analog zu anderen Analysemethoden: 20 Energien auf Spur (E_{Tube} -NN)
- ★ **Ziel:** nur 1 NN für alle Energien \Rightarrow Verwendung eines gemischten Datensatzes für die Protonen

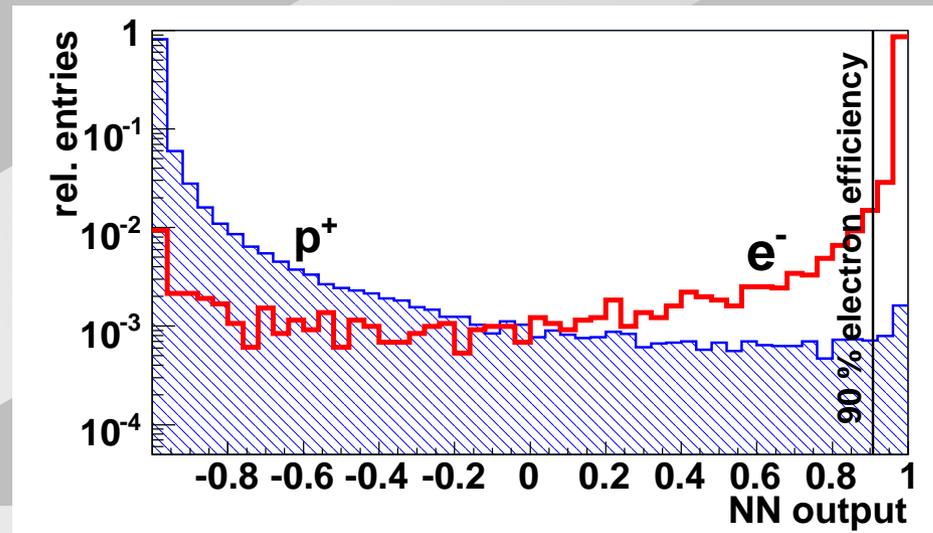
Teilchenidentifikation mit NN



- ★ Anwendung von NN zur Separation von Teilchensorten mit folgender Abbildungsvorschrift:

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

- ★ Eingabe in das NN:
 - analog zu anderen Analysemethoden: 20 Energien auf Spur (E_{Tube} -NN)
- ★ **Ziel:** nur 1 NN für alle Energien \Rightarrow Verwendung eines gemischten Datensatzes für die Protonen



Ausgabe des trainierten NN

- ☆ nach **Neyman-Pearson-Lemma** ist $\frac{P_0(\vec{x})}{P_1(\vec{x}) + P_0(\vec{x})} > c$ bester Hypothesentest
Likelihood arbeitet aber nur mit gemittelten Wahrscheinlichkeiten
⇒ **Verbesserung theoretisch möglich!**

☆ nach **Neyman-Pearson-Lemma** ist $\frac{P_0(\vec{x})}{P_1(\vec{x}) + P_0(\vec{x})} > c$ bester Hypothesentest

Likelihood arbeitet aber nur mit gemittelten Wahrscheinlichkeiten

⇒ **Verbesserung theoretisch möglich!**

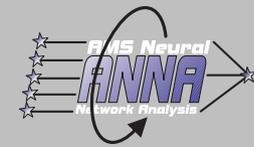
⇒ z.B. Eingabe von bisherigen Klassifikationsmethoden in ein **kombiniertes NN**

☆ Bedingungen für Verbesserung:

- jede Eingabe für sich sorgt schon für eine gute Trennung (*Likelihood, Zahl von Hits über 6.5 keV, E_{Tube} -NN, Fisher Testgröße, mittlere Energie*)

- Eingaben sollen möglichst unkorreliert sein

Korrelationsanalyse, Motivation eines weiteren Netzes



☆ nach **Neyman-Pearson-Lemma** ist $\frac{P_0(\vec{x})}{P_1(\vec{x}) + P_0(\vec{x})} > c$ bester Hypothesentest

Likelihood arbeitet aber nur mit gemittelten Wahrscheinlichkeiten

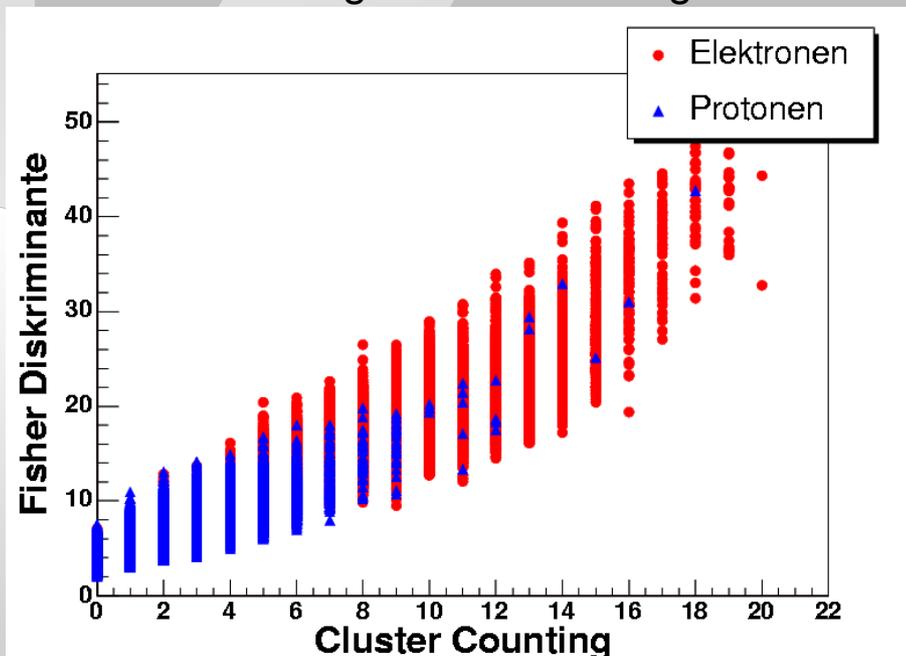
⇒ **Verbesserung theoretisch möglich!**

⇒ z.B. Eingabe von bisherigen Klassifikationsmethoden in ein **kombiniertes NN**

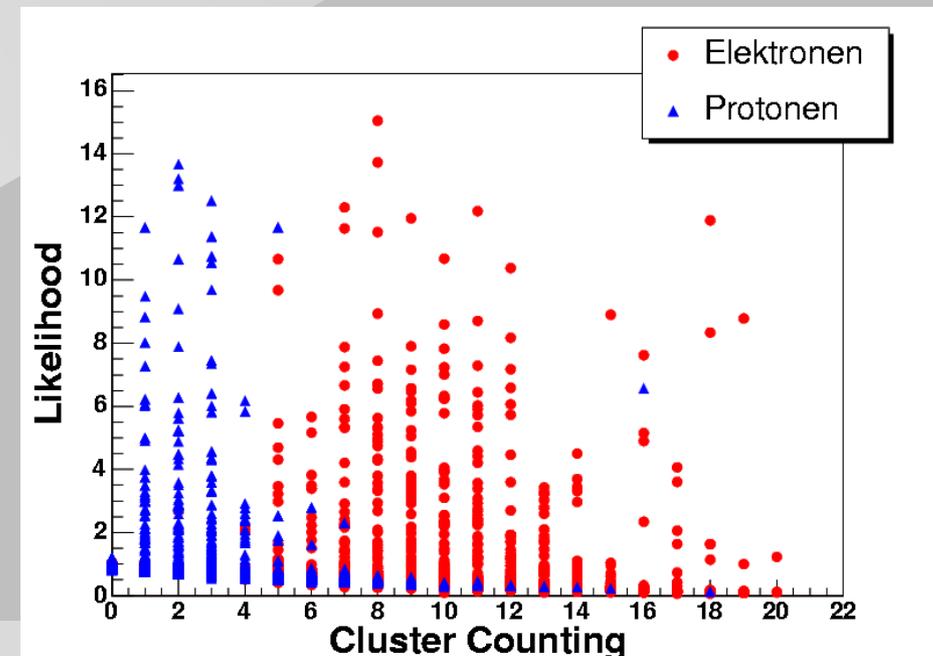
☆ Bedingungen für Verbesserung:

- jede Eingabe für sich sorgt schon für eine gute Trennung (*Likelihood, Zahl von Hits über 6.5 keV, E_{Tube} -NN, Fisher Testgröße, mittlere Energie*)

- Eingaben sollen möglichst unkorreliert sein

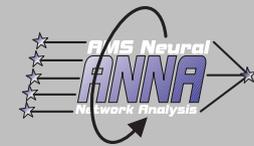


Abweichungen von perfekter Korrelation



geringe Korrelation

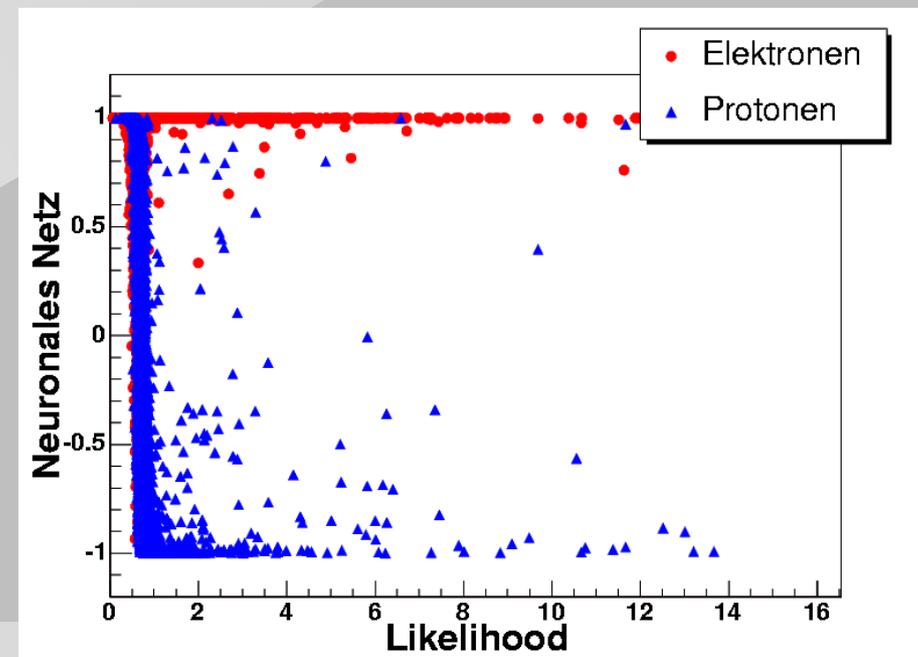
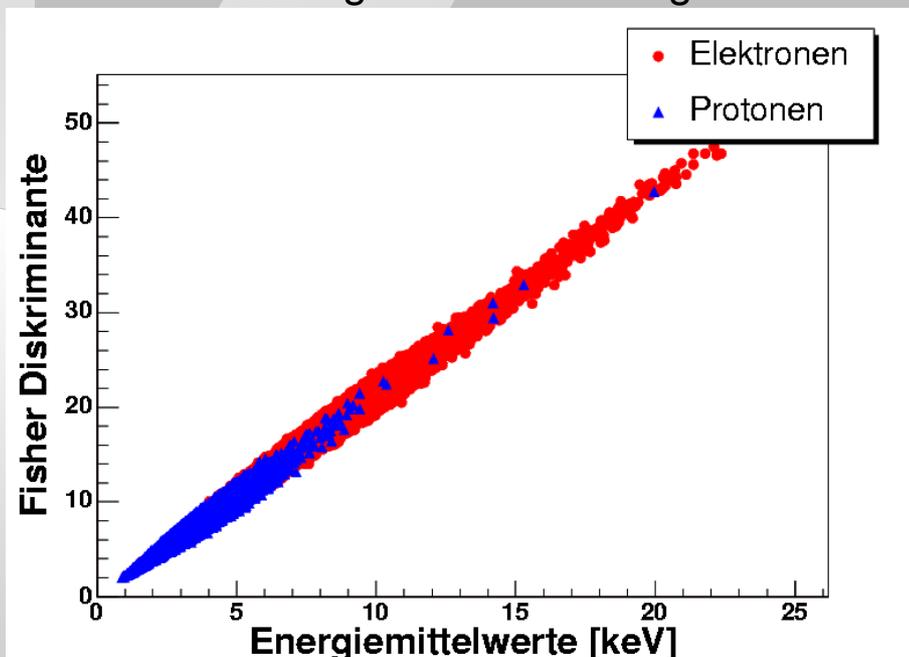
Korrelationsanalyse, Motivation eines weiteren Netzes



- ☆ nach **Neyman-Pearson-Lemma** ist $\frac{P_0(\vec{x})}{P_1(\vec{x}) + P_0(\vec{x})} > c$ bester Hypothesentest
Likelihood arbeitet aber nur mit gemittelten Wahrscheinlichkeiten
⇒ **Verbesserung theoretisch möglich!**

⇒ z.B. Eingabe von bisherigen Klassifikationsmethoden in ein **kombiniertes NN**

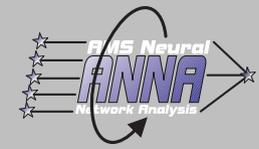
- ☆ Bedingungen für Verbesserung:
 - jede Eingabe für sich sorgt schon für eine gute Trennung (*Likelihood, Zahl von Hits über 6.5 keV, E_{Tube}-NN, Fisher Testgröße, mittlere Energie*)
 - Eingaben sollen möglichst unkorreliert sein



perfekte Korrelation ⇒ mittlere Energie nicht als Eingabe

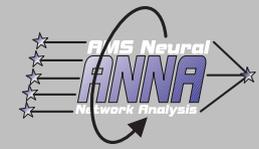
geringe Korrelation

Kombiniertes NN

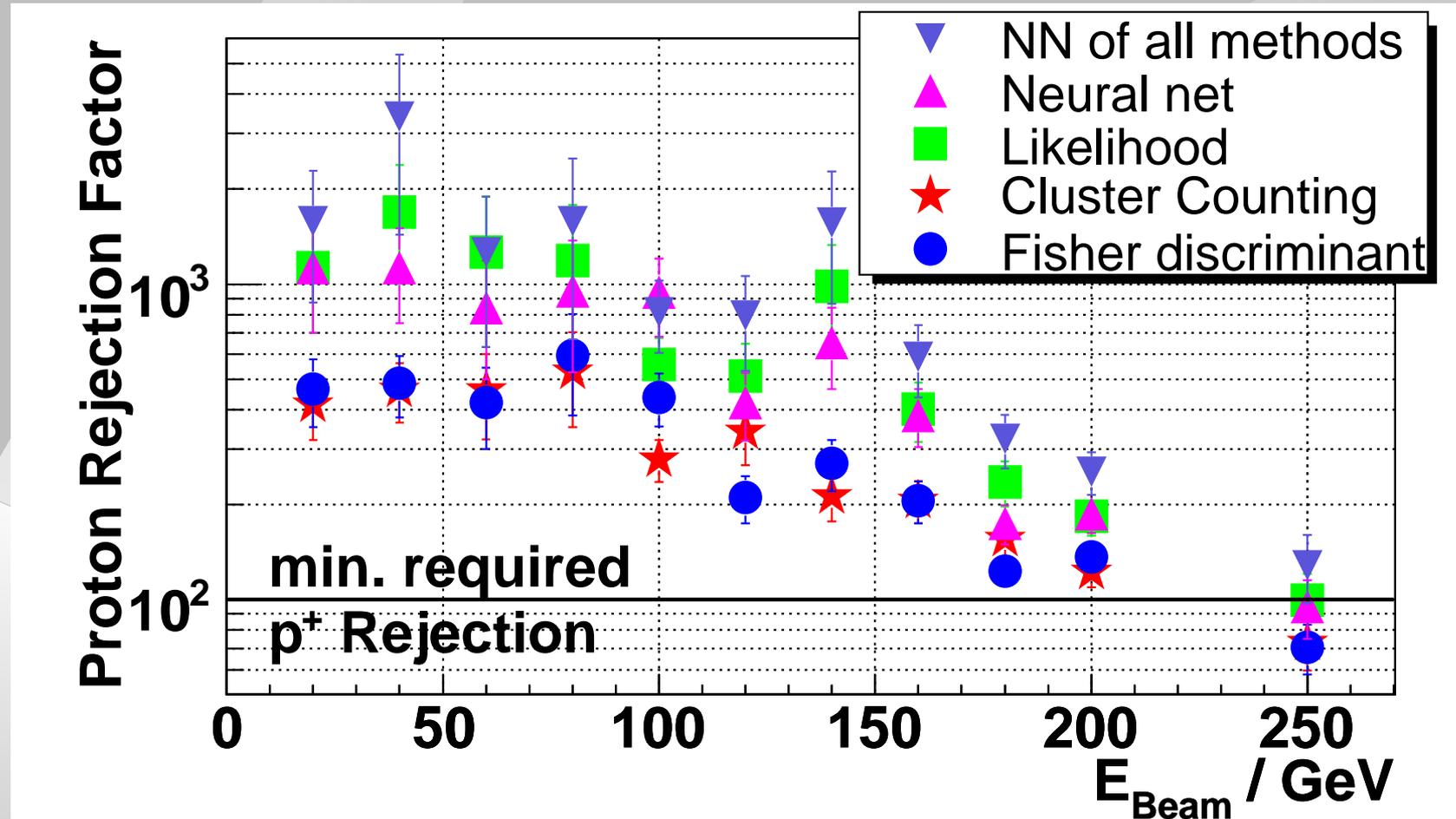


- ☆ Eingabe von *Likelihood, Zahl von Hits über 6.5 keV, E_{Tube} -NN, Fisher Testgröße*
- ☆ haben untereinander eine mittlere Korrelation von $\overline{|r|} \approx 0.7$
⇒ deutliche Abweichung von perfekter Korrelation

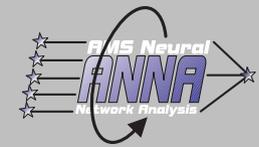
Kombiniertes NN



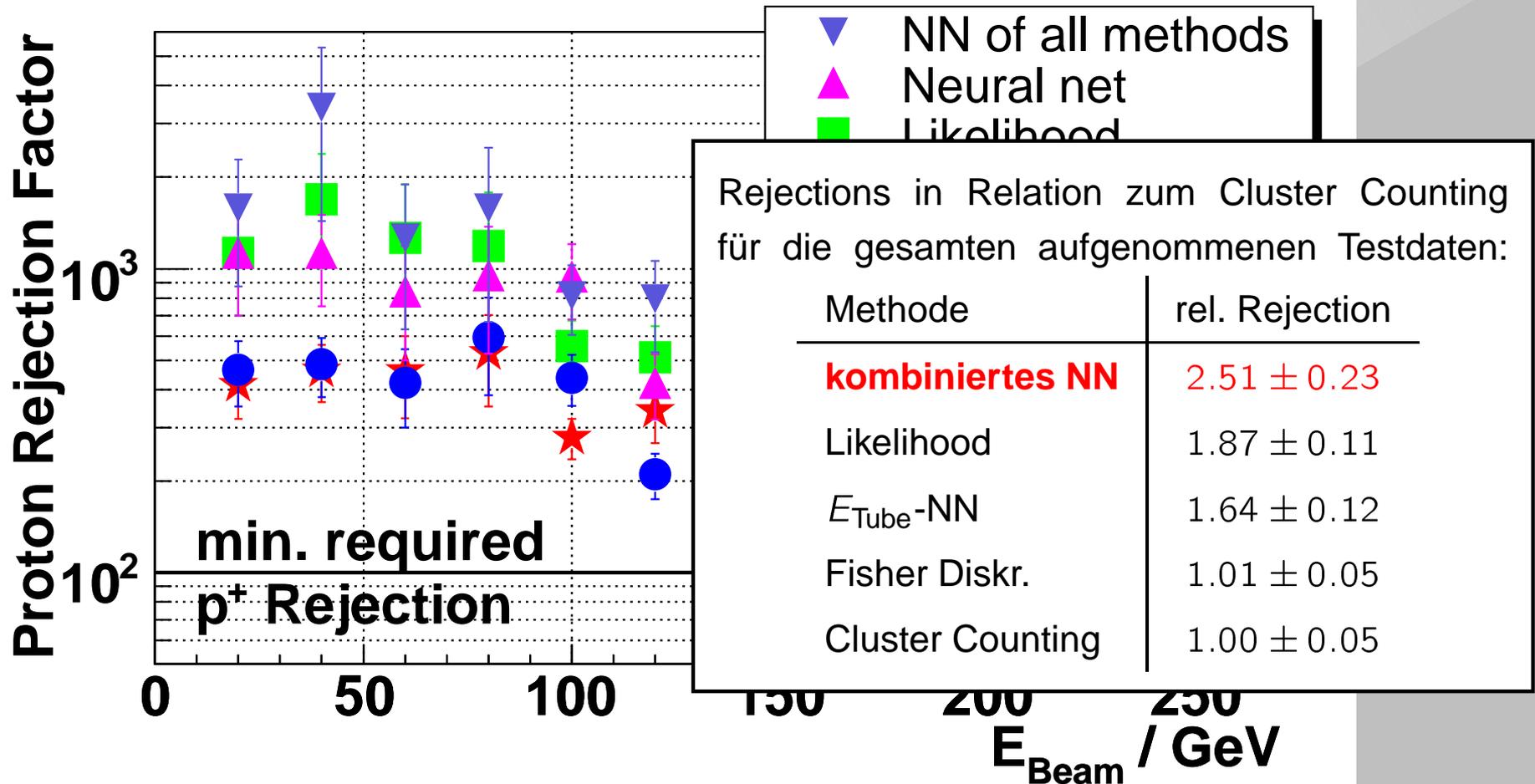
- ☆ Eingabe von *Likelihood, Zahl von Hits über 6.5 keV, E_{Tube}-NN, Fisher Testgröße*
- ☆ haben untereinander eine mittlere Korrelation von $|\overline{r}| \approx 0.7$
 - ⇒ deutliche Abweichung von perfekter Korrelation



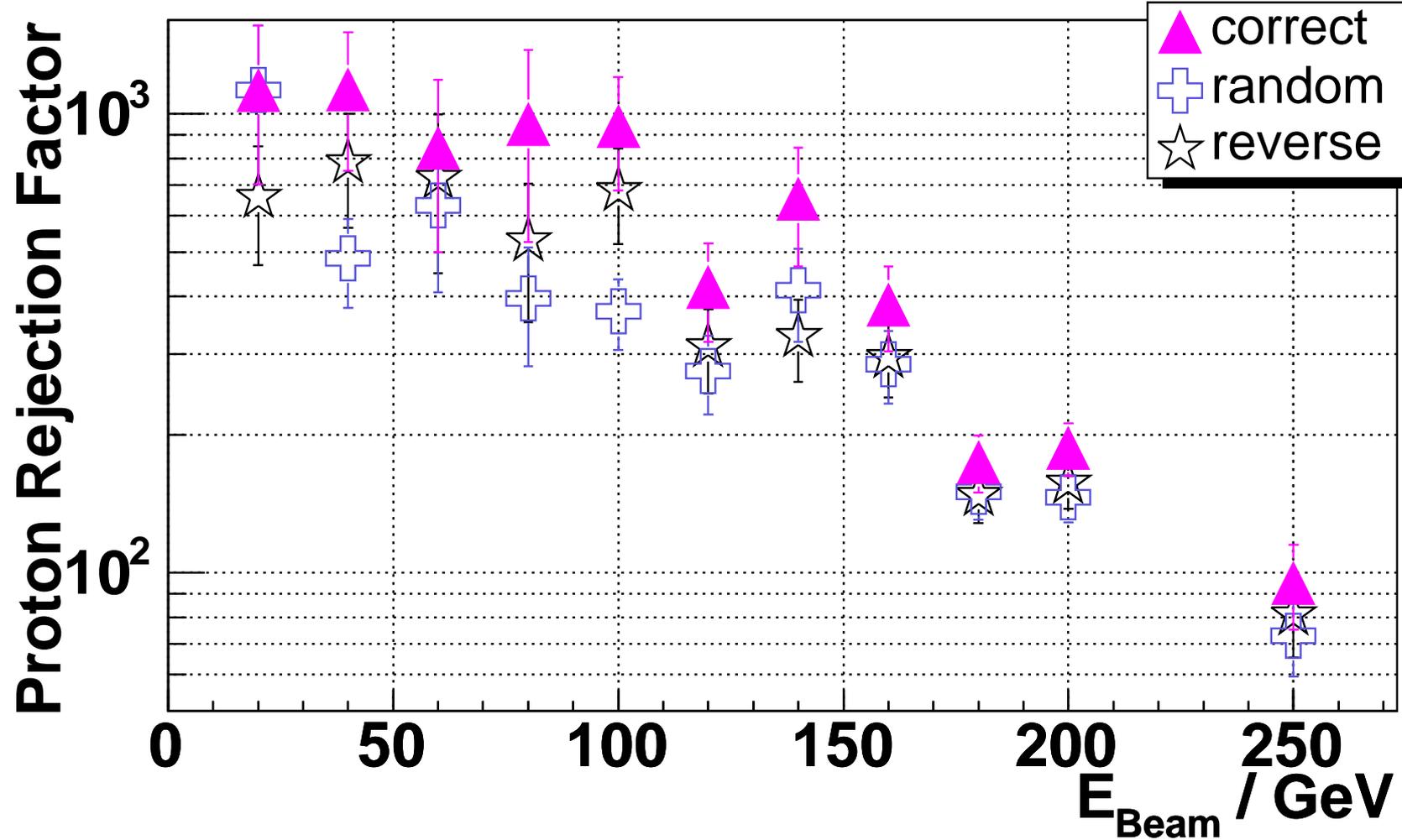
Kombiniertes NN



- ★ Eingabe von *Likelihood, Zahl von Hits über 6.5 keV, E_{Tube} -NN, Fisher Testgröße*
 - ★ haben untereinander eine mittlere Korrelation von $|\overline{r}| \approx 0.7$
- ⇒ deutliche Abweichung von perfekter Korrelation

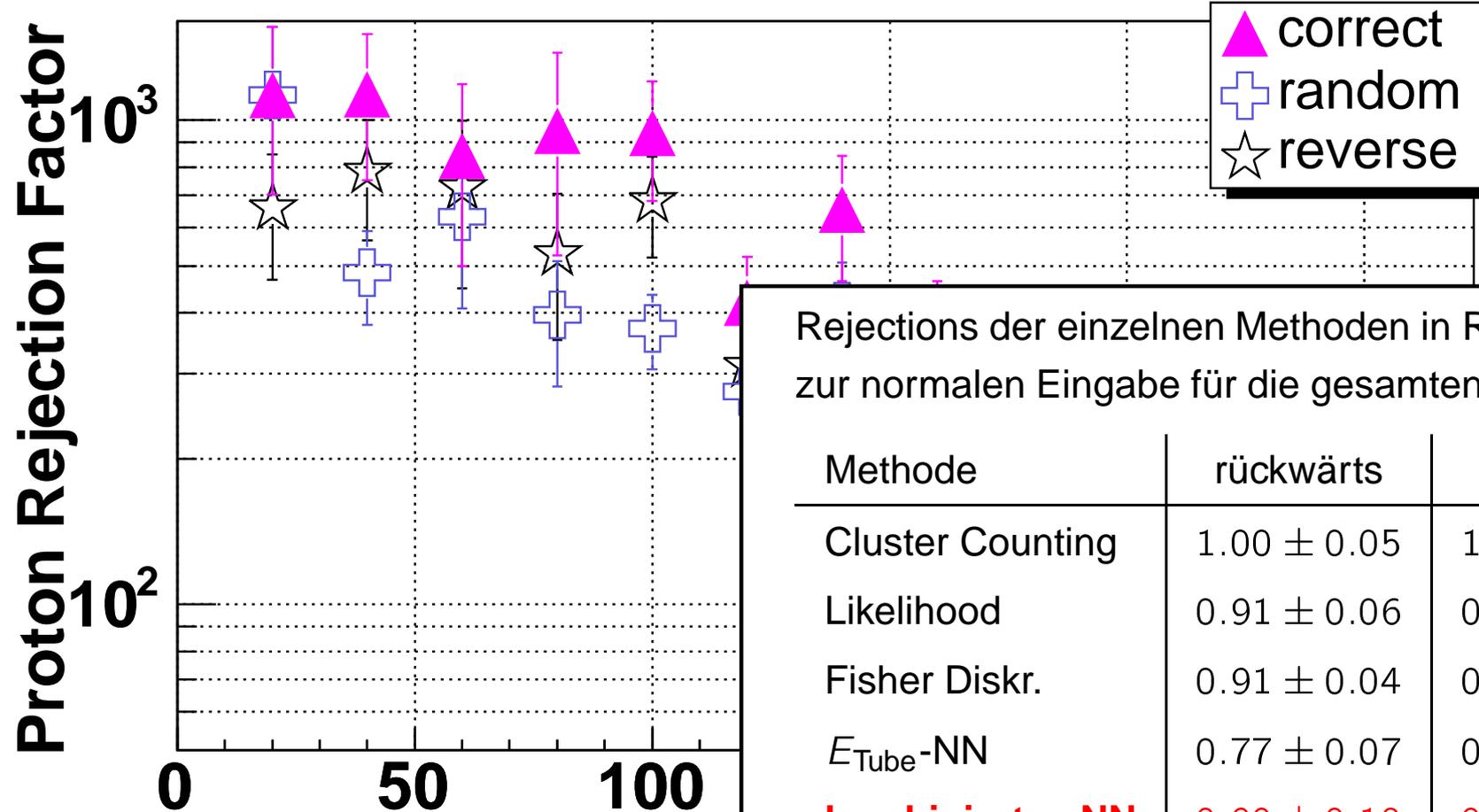
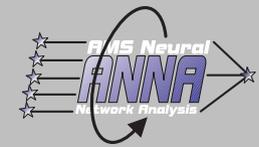


Sensitivität auf Eingabereihenfolge



★ E_{Tube} -NN ist sensitiv auf Eingabereihenfolge der Energien auf der Spur

Sensitivität auf Eingabereihenfolge

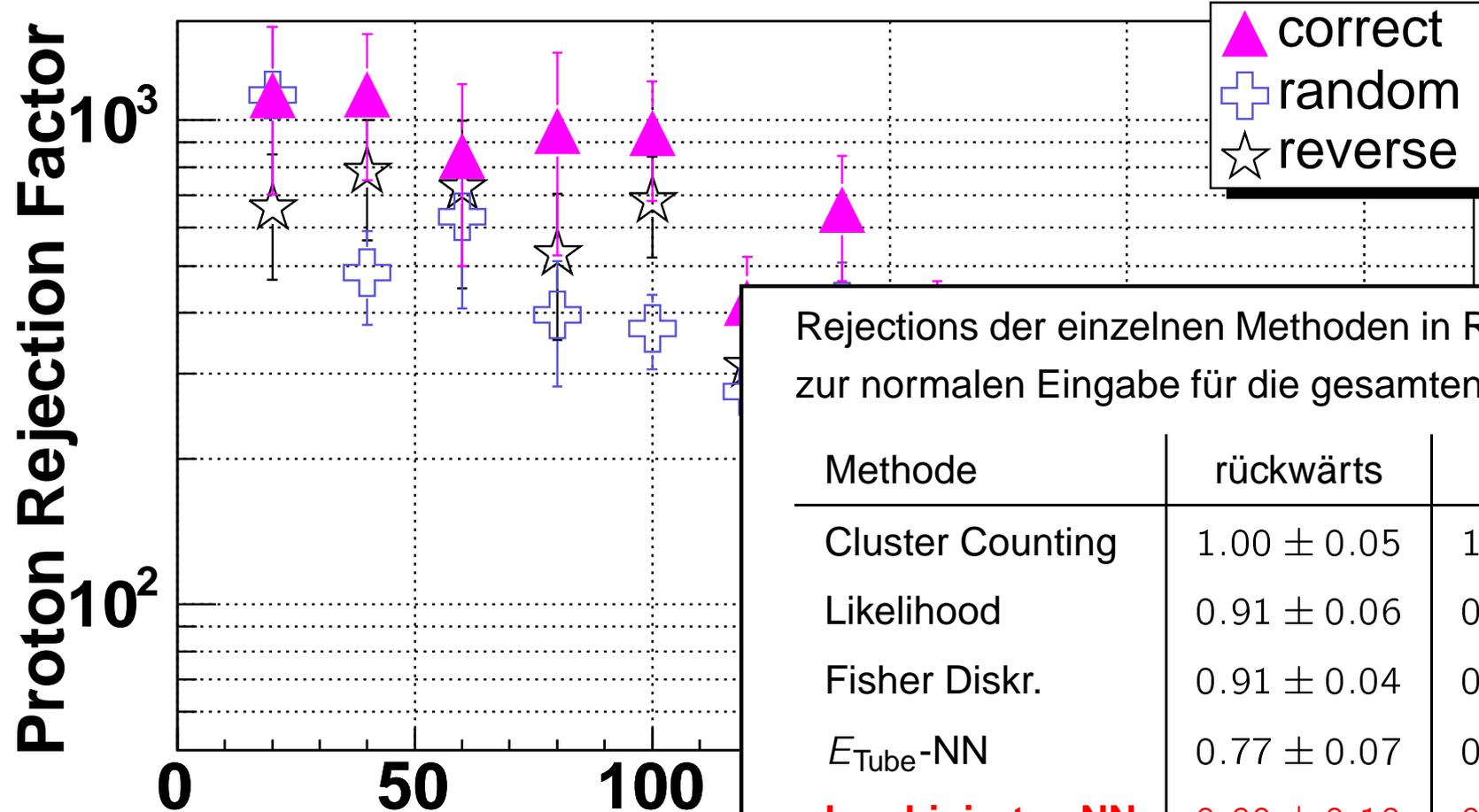


Rejections der einzelnen Methoden in Relation zur normalen Eingabe für die gesamten Testdaten:

Methode	rückwärts	zufällig
Cluster Counting	1.00 ± 0.05	1.00 ± 0.05
Likelihood	0.91 ± 0.06	0.93 ± 0.06
Fisher Diskr.	0.91 ± 0.04	0.82 ± 0.04
E_{Tube} -NN	0.77 ± 0.07	0.70 ± 0.06
kombiniertes NN	0.90 ± 0.10	0.88 ± 0.10

☆ E_{Tube} -NN ist sensitiv auf Eingabereihenfolge

Sensitivität auf Eingabereihenfolge



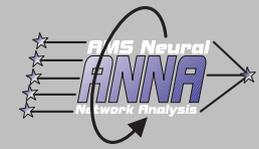
Rejections der einzelnen Methoden in Relation zur normalen Eingabe für die gesamten Testdaten:

Methode	rückwärts	zufällig
Cluster Counting	1.00 ± 0.05	1.00 ± 0.05
Likelihood	0.91 ± 0.06	0.93 ± 0.06
Fisher Diskr.	0.91 ± 0.04	0.82 ± 0.04
E_{Tube} -NN	0.77 ± 0.07	0.70 ± 0.06
kombiniertes NN	0.90 ± 0.10	0.88 ± 0.10

Steigerung der Rejection des kombinierten Netzes gegenüber der Likelihood also durch die **zusätzliche Richtungsinformation** aus E_{Tube} -NN

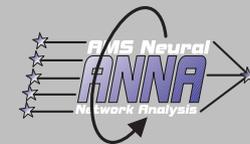
★ E_{Tube} -NN ist sensitiv auf Eingabereihenfolge

Ausblick



- ★ NN gute Möglichkeit zur Teilchenidentifikation
- ★ Fertigstellung der AMS-02-TRD Software: Rekonstruktion und Teilchenidentifikation
- ★ volle Detektorsimulation mit Spektren aus Modellen zur kosmischen Strahlung
- ★ Fertigstellung der Flight Hardware und Test

Ausblick



- ★ NN gute Möglichkeit zur Teilchenidentifikation
- ★ Fertigstellung der AMS-02-TRD Software: Rekonstruktion und Teilchenidentifikation
- ★ volle Detektorsimulation mit Spektren aus Modellen zur kosmischen Strahlung
- ★ Fertigstellung der Flight Hardware und Test

⇒ beten für die Shuttles



DON'T

e^+ -Identifikation mit Hilfe von neuronalen Netzen im TRD von AMS-02



DO