Globale Mustererkennung in den HERA-B-Spurkammern

Diplomarbeit
von
Christoph Borgmeier

Institut für Physik
Mathematisch-Naturwissenschaftliche Fakultät
Humboldt-Universität zu Berlin

September 1996

Inhaltsverzeichnis

Abbildungsverzeichnis

Einleitung

Ein wesentliches Merkmal der elementaren Wechselwirkungen ist ihr Verhalten unter diskreten Symmetrieoperationen. Es ist seit langem gesichert, daß sich die schwache Wechselwirkung weder unter der Ladungskonjugation C noch unter der Raumspiegelung P (Paritätstransformation) invariant verhält. Diese Asymmetrie manifestiert sich in der Händigkeit der Neutrinos. Die Naturgesetze scheinen dagegen weitgehend invariant gegenüber der Kombination CP zu sein. Eine geringe Asymmetrie wurde bisher nur im K^0-System nachgewiesen. Sie wird im Standardmodell durch eine Phase in der Quark-Mischungsmatrix beschrieben.

Im B^0-System wird eine weitaus größere CP-Verletzung erwartet. Ihre Messung ist ein wichtiger Test des Standardmodells. Eine experimentell und theoretisch eindeutig zu interpretierendes Ergebnis erreicht man mit der Messung des *goldenen B-Zerfalls*

$$B^0 \longrightarrow J/\psi\ K^0_S \longrightarrow \ell^+\ell^-\pi^+\pi^- \quad .$$

Da dieser Kanal nur bei 1/100000 der B-Zerfälle auftritt, benötigt man zur Messung der Asymmetrie eine große Anzahl von B-Mesonen. Um diese zu erzeugen und nachzuweisen, wird zur Zeit (1996) das HERA-B-Experiment an der Hadron-Elektron-Ringanlage (HERA) am Deutschen Elektronen-Synchrotron (DESY) aufgebaut. HERA-B ist ein Festtarget-Experiment am HERA-Protonenstrahl, welches Proton-Nukleon-Wechselwirkungen bei einer Schwerpunktsenergie von etwa 40 GeV untersuchen wird.

Bei dieser Energie entsteht etwa bei jeder 10^6ten Wechselwirkung ein B-Meson. Um in wenigen Jahren Laufzeit genügend goldene Zerfälle messen zu können, ist eine hohe Wechselwirkungsrate nötig, was dazu führt, daß bei jedem Strahlpaket-Durchgang im Mittel vier Reaktionen stattfinden müssen, die oft zusammen über 100 geladene Teilchen erzeugen, welche als Spuren in den Spurkammern nachgewiesen werden. Zusätzlich entstehen Spuren durch Wechselwirkungen mit dem Detektormaterial, so daß die Rekonstruktion der ursprünglichen Ereignisse weiter erschwert wird.

Die Spurrekonstruktion spielt deshalb eine zentrale Rolle im HERA-B-Experiment. Bisher wurde lediglich das klassische Verfahren der Spurverfolgung untersucht und die Ereignisrekonstrukktion auf einen *Kalman-Filter*-Algorithmus aufgebaut. Die Schwierigkeit der Aufgabe legt es jedoch nahe, auch alternative Spurfindungsverfahren zu untersuchen. In den letzten Jahren sind insbesondere mehrere Verfahren vorgeschlagen worden, die sich am Funktionsprinzip des Gehirns orientieren. Diese neuronalen Methoden können konventionellen in einigen Gebieten aufgrund ihrer globalen Vorgehensweise, das heißt, die gleichzeitige Optimierung der Zuordnung aller Signale zu Spuren, im Gegensatz zu lokalen Suchalgoritmen überlegen sein. Außerdem arbeiten die biologischen Systeme in einem hohen Grad parallel, was dazu führt, daß auch komplexe Aufgaben in kurzer Zeit gelöst werden können.

In dieser Arbeit werden zwei unterschiedliche globale Algorithmen auf ihre Anwendbarkeit zur Spurfindung in den HERA-B-Spurkammern untersucht. Der erste davon ist das Verfahren nach Denby und Peterson, welches mit primitiven Schaltelementen arbeitet und nur geringste Annahmen über die Form des Detektors oder der Spuren macht. Der Algorithmus kann so erweitert werden, daß er auf Ereignisse in einem HERA-B-ähnlichen Modelldetektor angewendet werden kann.

Als zweite Methode wird der *Elastic-Arms-Algorithmus* nach Ohlsson und Peterson untersucht. Dieser arbeitet mit einem Spurmodell und liefert die Parameter der gefundenen Spuren. Studium und Optimierung dieses Algorithmus erfolgen zunächst ebenfalls mit Hilfe des Modelldetektors. Anschließend werden mit dem Elastic-Arms-Algorithmus Spuren im komplett simulierten HERA-B-Detektor gesucht.

1 Das HERA-B-Experiment

1.1 Der „goldene" Zerfallskanal

Im Standardmodell können die schwachen Übergänge zwischen Quarks durch die unitäre Cabibbo-Kobayashi-Maskawa-Matrix (CKM-Matrix) mit den Matrixelementen V_{ij} für den Übergang eines Quarks i in ein Quark j beschrieben werden [1]. Eine Näherung wurde von Wolfenstein [2] vorgeschlagen, mit den reellen Parametern λ, A, ρ und η. Sie ist in Anhang A.1 näher erläutert. Eine für die Physik der schweren Quarks besonders wichtige Unitaritätsbedingung ist

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad ,$$

was sich durch ein Dreieck in der ρ-η-Ebene der Wolfenstein-Parametrisierung darstellen läßt (Abbildung 1.1).

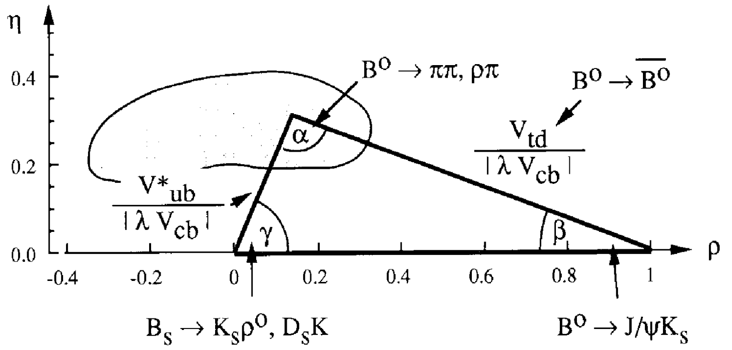


Abbildung 1.1: Unitaritätsdreieck in der ρ-η Ebene der Wolfenstein-Parametrisierung der CKM-Matrix.

Eine nicht-verschwindende Fläche des Dreiecks bedeutet, daß es im Standardmodell eine CP-Verletzung im B-System gibt. Der Parameter β des Standardmodells ist ein Winkel des Unitaritätsdreiecks in der ρ-η-Ebene, der sich im „goldenen B-Zerfall"

$$B^0/_{\bar{B}^0} \longrightarrow J/\psi\ K_S^0 \longrightarrow \ell^+\ell^-\pi^+\pi^- \tag{1.1}$$

messen läßt.[1] Dabei zerfällt das B^0 oder das $\bar{B}^0$ in einen CP-Eigenzustand. Die zu beobachtende Asymmetrie A_{int} der zeitintegrierten Raten ist

$$A_{\text{int}} = \frac{N - \bar{N}}{N + \bar{N}} \quad , \tag{1.2}$$

[1]In dieser Arbeit wird nur der B-Zerfall in $J/\psi \longrightarrow e^+e^-, \mu^+\mu^-$ und $K_S^0 \longrightarrow \pi^+\pi^-$ als „golden" bezeichnet, da hauptsächlich dieser mit HERA-B eindeutig zu messen sein wird.

wobei N die Zahl der erzeugten B^0 und $\bar{N}$ die der $\bar{B}^0$ ist. Da die B-Mesonen in einem starken Prozeß flavour-erhaltend erzeugt werden, kann man das nach Gleichung 1.1 zerfallende B durch die Analyse des gleichzeitig erzeugten Teilchens mit entgegengesetzter Beauty-Quantenzahl identifizieren. Diesen Vorgang bezeichnet man als *tagging*. Das sogenannte *Mixing* überführt einen gewissen Anteil der B-Mesonen in $\bar{B}$ und umgekehrt. Daher kann das Tagging nur statistische Aussagen machen.

Die Asymmetrie hängt mit dem Winkel β des Unitaritätsdreiecks wie folgt zusammen:

$$A = \sin 2\beta \sin xt \quad , \tag{1.3}$$

wobei wegen der B^0-$\bar{B}^0$-Oszillation mit dem *Mixing-Parameter* x auch die Flugzeit t eingeht. Aufgrund der bisherigen Messungen der CKM-Matrixelemente erwartet man $0.3 < \sin 2\beta \lesssim 0.5$ [3]. Um eine Genauigkeit von $\Delta \sin 2\beta \approx 0.1$ zu erreichen, werden $\mathcal{O}(1000)$ vollständig rekonstruierte und getaggte goldene Zerfälle benötigt.

Bei HERA-B-Schwerpunktsenergien von etwa 30 GeV gilt für den bb-Wirkungsquerschnitt im Vergleich zum Wirkungsquerschnitt für alle inelastischen Proton-Nukleon-Wechselwirkungen:

$$\frac{\sigma(pN \to b\bar{b} + X)}{\sigma_{\mathrm{inelastisch}}} \approx 10^{-6} \quad . \tag{1.4}$$

Das Verzweigungsverhältnis des gesuchten Zerfalls ist

$$\frac{\Gamma(B^0 \to J/\psi\, K_S^0 \to \ell^+\ell^-\pi^+\pi^-)}{\Gamma_{B,\mathrm{total}}} \approx 10^{-5} \quad . \tag{1.5}$$

Die gesuchten B-Zerfälle treten also nur bei etwa jeder 10^{11}ten Wechselwirkung auf. Um etwa 1000 goldene Zerfälle im Jahr zu beobachten, braucht man bei Berücksichtigung der Trigger- und Rekonstruktionseffizienz eine Wechselwirkungsrate von 40 MHz.

1.2 Der HERA-B-Detektor

HERA-B ist ein Fixed-Target-Experiment am HERA-Protonenstrahl, welches sich derzeit (1996) in der Aufbauphase befindet. Der Detektor ist in Abbildung 1.2 dargestellt. Die HERA-Protonpaket-Rate ist 10 MHz, was bedeutet, daß im Mittel vier Wechselwirkungen pro Strahldurchgang stattfinden müssen, um die nötige Ereignisrate zu gewährleisten.

Das Draht-Target besteht aus vier vertikalen und vier horizontalen Drähten, die den Protonen-Strahl als zwei Rechtecke umschließen. Sie befinden sich im Halo des Strahls, so daß nur die weit außen im Strahl laufenden Protonen benutzt werden und der Strahl weitgehend ungestört bleibt.

Der HERA-B-Detektor ist vor allem darauf optimiert, die goldenen Zerfälle zu finden. Er deckt den Polarwinkel-Bereich von 10 mrad bis etwa 220 mrad ab, was etwa 90% des Raumwinkels im Schwerpunktsystem bedeutet.

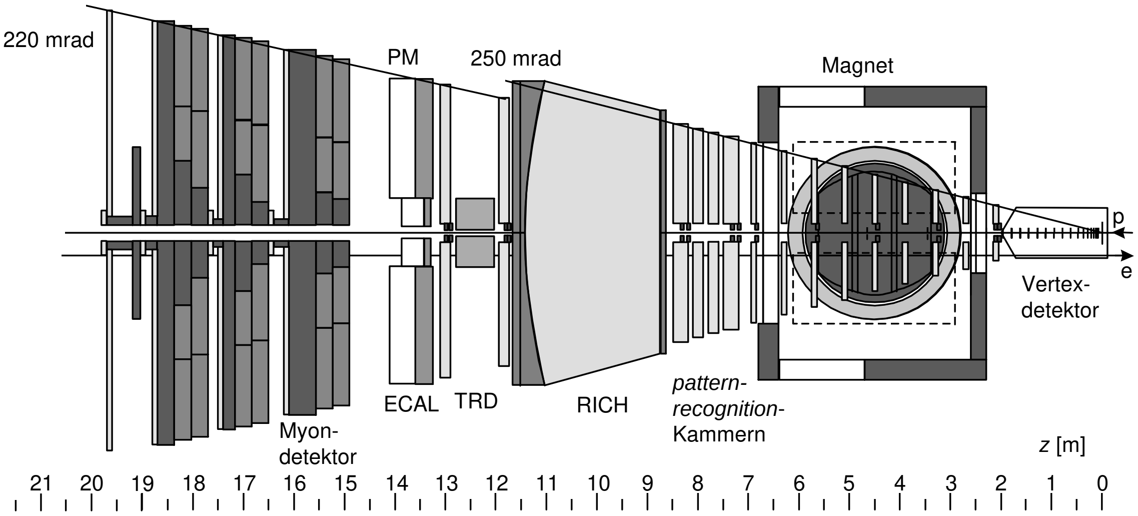


Abbildung 1.2: Der HERA-B-Detektor in der Aufsicht

Der Koordinatenursprung liegt beim Target, die Protonen bewegen sich entlang der z-Achse. Die y-Achse verläuft vertikal, die x-Achse horizontal, in Abbildung 1.2 weist sie aus der Zeichenebene heraus.

Für einen goldenen Zerfall wird sowohl das J/ψ als auch das K_S^0 rekonstruiert, ersteres über die Leptonen, letzteres über die geladenen Pionen. Durch Schnitte auf die invarianten Massen läßt sich ein sehr großer Teil des Untergrundes unterdrücken. Eine weitere Unterdrückung bekommt man durch die Rekonstruktion der Vertizes, zum Beispiel durch die Bedingung, daß das J/ψ und das K_S^0 einen gemeinsamen Erzeugungsvertex haben.

Insgesamt müssen Elektronen, Myonen und Pionen identifiziert und ihre Impulse gemessen werden. Für das Tagging muß man auch geladene Kaonen rekonstruieren können. Die Zerfallslängenmessung benötigt außerdem eine genaue Rekonstruktion des B-Zerfallsvertex. Der Erzeugungsvertex ist durch die Drahtesposition gegeben.

Die Myonen werden durch Spuren im Myondetektor identifiziert, die Elektronen durch elektromagnetische Schauer im Kalorimeter (ECAL). Die mit hohem Impuls im Bereich kleiner Winkel verlaufenden Elektronen können zusätzlich mit dem Übergangsstrahlungsdetektor (TRD) identifiziert werden. Die Unterscheidung der Kaonen und Pionen geschieht mit dem *Ring-Image-Cherenkov*-Zähler (RICH).

1.3 Ereignisrekonstruktion bei HERA-B

Die Ereignisrekonstruktion beginnt in den Spurkammern hinter dem Magneten, die als *Pattern-Recognition*-Kammern bezeichnet werden. Die dort gefundenen geraden Spuren werden durch den Magneten verfolgt, wobei das inhomogene Feld berücksichtigt werden muß. Unabhängig davon wird nach Spuren im Silizium-Vertexdetektor gesucht. Weiterhin fließen die Informationen aus ECAL, RICH, TRD und dem Myondetektor ein.

Da die Spurdichte zu kleinen Winkeln hin stark anwächst, kommen für das innere und das äußere Spurkammersystem verschiedene Detektortypen zum Einsatz. Die äußeren Spurkammern, die eine geringere Granularität benötigen, aber eine große Fläche abdecken müssen, setzen sich aus *Wabendriftkammern (honeycomb drift chambers, HDC)*, mit einem Zelldurchmesser von 5 mm oder 10 mm und einer Auflösung von etwa 200 μm zusammen. Es gibt Kammern in drei verschiedenen Richtungen. Die Signaldrähte der sogenannten 0°-Lagen liegen in Richtung der y-Achse. Zur Bestimmung der y-Komponenten der Teilchendurchgänge gibt es Stereolagen mit einem Winkel von $+5°$ und $-5°$. Die Lagen sind nicht äquidistant angeordnet, sondern in einer Superlagenstruktur gegliedert. Das bedeutet, daß mehrere Driftkammerlagen direkt aneinander grenzen, so daß der Abstand der Signaldrähte in z-Richtung in der Größenordnung von 1 cm liegt. Diese Struktur erlaubt der Mustererkennung, lokal nach kurzen Spursegmenten zu suchen, so daß das Problem der Kombinatorik bei Zuordnung der Detektorsignale zu den Spuren auf niedriger Ebene verkleinert werden kann. Im Pattern-Recognition-Detektor wird es vier Superlagen mit insgesamt 28 HDC-Lagen geben.

Die inneren Kammern bestehen aus *Mikrostreifen-Gaszählern (microstrip gas chambers,* MSGC*)* mit einer Auflösung von etwa 80 μm bei einem Streifenabstand von 300 μm. Die Superlagenstruktur und die Anordnung der Stereolagen entspricht den äußeren Kammern. Es gibt allerdings weniger Lagen, im Pattern-Recognition-Bereich sind es lediglich 18.

Die Struktur des Pattern-Recognition-Detektors mit seinen vier Superlagen ist in der Detaildarstellung, Abbildung 1.3 (unten), zu erkennen. Der besseren Übersicht wegen sind nur die 0°-Lagen gezeigt. Im Bereich $x < 20$ cm sind zehn MSGC-Lagen dargestellt. Da jeweils zwei Module mit kleinem Versatz in z-Richtung vertikal übereinander angeordnet sind, erscheint eine komplette Lage als zwei Linien. In der Projektion sind die Signale der Teilchen punktförmig dargestellt, da es sich um einen Streifendetektor handelt.

Ein Driftkammersignal besteht aus einer Driftzeitinformation, die auf die Entfernung der Teilchentrajektorie vom Signaldraht schließen läßt. In der Projektion senkrecht zum Draht berührt die Spur also eine kreisförmige *Isochrone*. Diese Isochronen sind in Abbildung 1.3 vor allem in den 10 mm-Kammern (bei $x > 60$ cm) gut zu erkennen. Die Module der 5 mm-Kammern sind in z-Richtung versetzt angeordnet, so daß die jeweils ersten drei von sechs Linien einer Superlage das eine Modul repräsentieren, die letzten drei das andere (vgl. auch Abbildung 4.11).

Die Driftkammersignale liefern keine eindeutige Ortskoordinate. In den meisten Fällen wird man die Richtung der Spuren θ_x in etwa kennen, so daß bei gemessenem Driftweg r und der Projektion auf die x-Achse $x_{\text{drift}} = r/\cos\theta_x$ die Spur bei $x = x_{\text{draht}} \pm x_{\text{drift}}$ liegt. Diese Zweideutigkeit bezeichnet man als *Driftkammer-Ambiguität*.

1.4 Parameterdarstellung der Teilchentrajektorien in den HERA-B-Spurkammern

Die geraden Spuren im Bereich der HERA-B-Spurkammern bei $710\,\mathrm{cm} < z < 840\,\mathrm{cm}$ können wie folgt beschrieben werden:

$$\vec{x}(z) = \begin{pmatrix} x_0 \\ y_0 \\ z_{\mathrm{ref}} \end{pmatrix} + (z - z_{\mathrm{ref}}) \begin{pmatrix} \tan\theta_x \\ \tan\theta_y \\ 1 \end{pmatrix} \quad . \tag{1.6}$$

Die vier Parameter einer Spur sind also die Achsenabschnitte in der Referenzebene $z = z_{\mathrm{ref}}$, x_0 und y_0, sowie die Richtungstangens in der xz- und der yz-Ebene, $\tan\theta_x$ und $\tan\theta_y$. Der Parametersatz einer Spur wird im folgenden mit $\mathbf{t}$ bezeichnet.

1.5 Problem der Spurfindung in HERA-B

Die Ereignisstruktur bei HERA-B stellt hohe Anforderungen an die Spurerkennung. Die meisten der 100 bis 200 Spuren, welche gleichzeitig in den Pattern-Recognition-Spurkammern zu finden sind, erzeugen jeweils etwa 18 bis über 30 Hits[2] . Die höchste Belegungsdichte liegt im Bereich der inneren Wabendriftkammern und beträgt etwa 20%.

Da der größte Teil der Spuren einen Erzeugungsvertex nahe am Target hat und die Ablenkung des Magnetfeldes umgekehrt proportional zum Impuls ist, gibt es für die höherenergetischen Spuren eine deutliche Korrelation zwischen Richtungstangens und Achsenabschnitt

$$\begin{aligned} \tan\theta_x &\approx k_x x_0 \quad , \\ \tan\theta_y &\approx k_y y_0 \quad . \end{aligned}$$

Diese Korrelation ist in Abbildung 4.13 zu erkennen, in der die Spuren eines HERA-B-Ereignisses als Kreise im Parameterraum dargestellt sind. Die niederenergetischen Spuren werden vom Magneten stärker abgelenkt und haben deshalb zumeist eine große Steigung $|\tan\theta_x|$ und sorgen so für eine große Anzahl von Spurkreuzungen in der xz-Projektion. In der yz-Projektion bewirkt das Magnetfeld fast keine Ablenkung der Spuren, so daß die Korrelation $\tan\theta_y$ - y_0 deutlich größer ist.

1.6 Die klassische Spurerkennung

Die Spurrekonstruktion in den HERA-B-Spurkammern benutzt die klassische Methode der Spurverfolgung (*track following*) mittels eines *Kalman-Filter*-Algorithmus [4]. Sie besteht aus der Suche nach Anfangsspursegmenten *(seeds)* und der *Propagation* durch die Detektorlagen. Die Spuren werden zunächst in den 0°-Lagen gesucht, erst danach fließen die Informationen der Stereo-Lagen ein.

[2]Die Treffer der Teilchen in den Detektorelementen werden im folgenden *Hits* genannt.

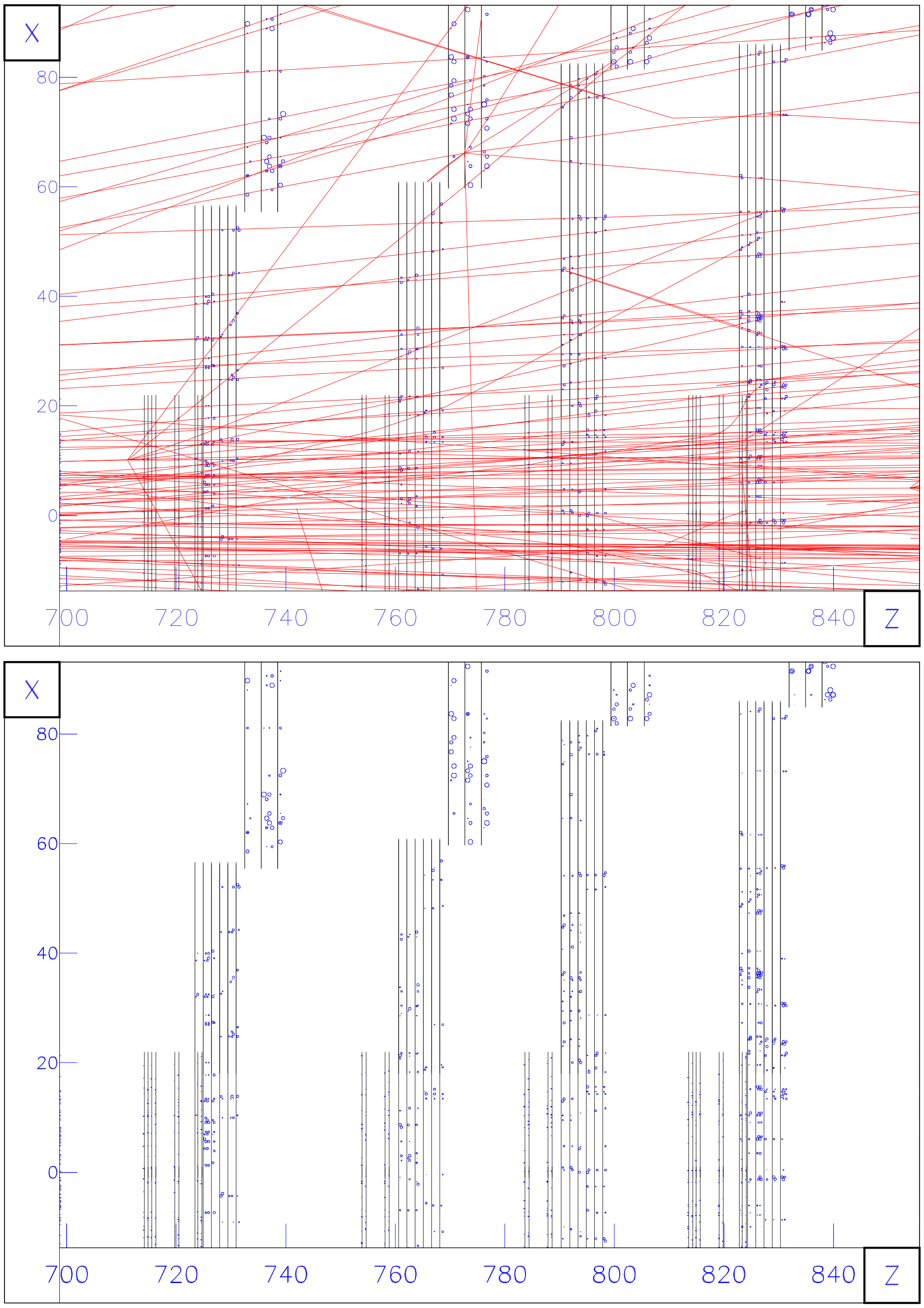


Abbildung 1.3: Detail eines HERA-B-Ereignisses mit 379 simulierten Spuren, oben die Monte-Carlo-Wahrheit, unten die Hits im Detektor (Aufgabenstellung der Spurerkennung).

Mögliche Spursegmente werden innerhalb der Superlagen gesucht. Zur Propagation wählt der Kalman-Filter in den zu betrachtenden benachbarten Detektorlagen kompatible Hits aus und benutzt sie zur sukzessiven Verbesserung der Spurparameter und der Kovarianzmatrix. Auf diese Weise wird mit der Anzahl der zugeordneten Hits das Suchfenster in nachfolgenden Lagen kleiner. Wenn sich ein Spuranfang nicht durch eine minimale Zahl von Detektorlagen verfolgen läßt, kann davon ausgegangen werden, daß der angenommene Spuranfang eine falsche Hitkombination ist, und die Spur wird verworfen.

Durch Vielfachstreuung hervorgerufene Spurkrümmungen können in der Kovarianzmatrix berücksichtigt werden. Der Algorithmus liefert sowohl die den Spuren zugeordneten Hits als auch die Spurparameter. Die Spurfindungseffizienz bei simulierten HERA-B-Ereignissen beträgt über 95%, wobei die Rechenzeit für ein Ereignis im Sekundenbereich liegt.

Ein Nachteil des Verfahrens ist, daß Spuren, die in den Bereichen, in denen nach Anfängen gesucht wird, nicht genügend passende Hits erzeugt haben, nicht verfolgt werden. Dieser Effekt kann nur durch ausreichende Redundanz kompensiert werden. Eine solche *lokale* Methode führt daher meist einen gewissen *Bias*[3] ein.

Ein *globales* Verfahren hat dagegen generell die Eigenschaft, alle zur Verfügung stehenden Informationen gleichberechtigt zu nutzen. Das Ergebnis sollte damit weder von der Reihenfolge der betrachteten Hits noch von einer irgendwie gearteten Reihenfolge der betrachteten Spuren abhängen. Außerdem können die globalen Algorithmen besonders günstige Eigenschaften bezüglich des Einsatzes auf Parallelrechnern aufweisen.

[3] Der Begriff „Bias“ bezeichnet eine Verzerrung des Ergebnisses durch gewählte Voreinstellungen.

2 Neuronale Modelle

Das menschliche Gehirn kann kognitive Aufgaben —wie Mustererkennung— in vielen Fällen besser lösen als eine konventionelle Rechenmaschine. Es hat die Fähigkeit, ein Muster global zu erfassen und zu bewerten. Es ist selbstorganisierend, damit anpassungsfähig und fehlertolerant. Der Grund dafür ist der hohe Grad an paralleler Informationsverarbeitung, weiterhin die Kombination von Daten und Algorithmen in denselben Schaltelementen.

Die attraktiven Eigenschaften des biologischen Systems könnten künstlich erreicht werden, indem man die primitiven Schaltelemente simuliert. Solch ein neuronaler Ansatz bietet sich auch zur Lösung der Spurerkennungsaufgaben für HERA-B an.

2.1 Das biologische Vorbild

Die Nervenzellen des Gehirns sind einfache Schaltelemente und werden *Neuronen* genannt. Ein Neuron besteht aus einem *Zellkörper*, den *Dendriten* und einem *Axon*. Diese drei Elemente sind in Abbildung 2.1 zu erkennen. Die Dendriten sind Signaleingänge. Sie sammeln die Signale über chemische Kontakte von anderen Neuronen und verändern das elektrische Potential des Zellkörpers. Das Überschreiten einer Potentialschwelle löst einen Impuls auf dem Axon aus. Das Axon verzweigt sich am Ende in Dendriten, die das Signal auf andere Neuronen übertragen.

2.2 Künstliche neuronale Netze

Für die meisten künstlichen neuronalen Netze kann man eine vereinfachende Parametrisierung des biologischen Neurons benutzen. Ihr Schaltbild ist in Abbildung 2.2 dargestellt. Das in dieser Arbeit benutzte Modell entspricht der in [5, 1. Kapitel] beschriebenen Erweiterung des McCulloch-Pitts-Neurons [6] mit folgenden Eigenschaften:

- Es verarbeitet die Eingangssignale x_j.
- Die Gewichte w_{ij} beeinflussen die Wirkung der Eingänge auf das Potential des Neurons i. Wenn $w_{ij} > 0$ gilt, erhöht ein Signal am Eingang j das Potential, andernfalls wirkt x_j hemmend.
- Die Aktivierungsschwelle heißt s_i,
- die Aktivitätsfunktion ist damit $z_i = \sum_j w_{ij}x_j - s_i$.

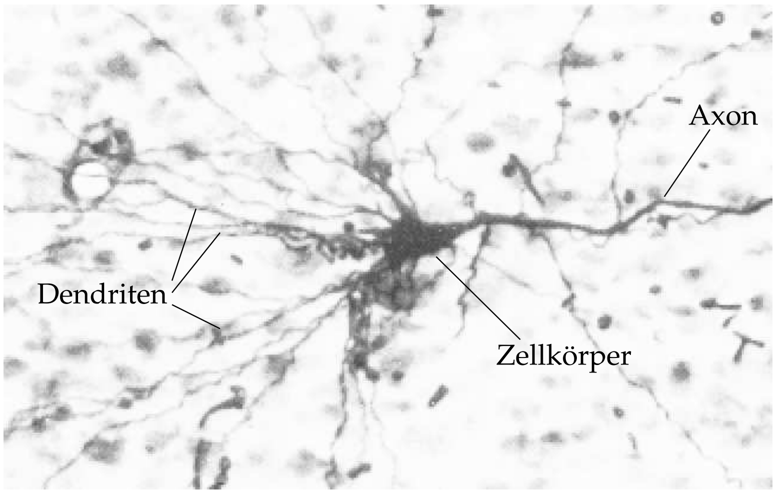


Abbildung 2.1: Gehirnzelle

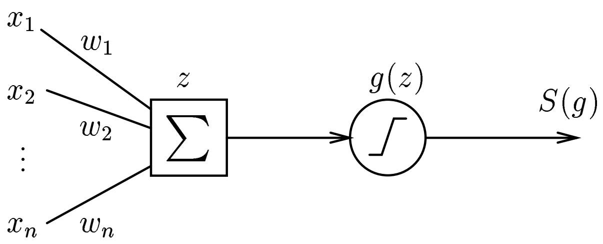


Abbildung 2.2: Neuron als logisches Schaltelement

- Die Übertragungsfunktion g besitzt ein Schwellenverhalten mit der Schwelle bei 0.
- Die *Aktivität* S des Neurons ist das Ergebnis der Übertragungsfunktion, $S = g(z)$.
- Die Berechnung der Aktivität benötigt eine konstante Zeit.

Ein Schaltschritt wird durch die Aktualisierungsgleichung

$$S_i = g\Big(\sum_i w_{ij} x_j - s_i\Big) \tag{2.1}$$

beschrieben.

2.3 Parallele Informationsverarbeitung

Das menschliche Gehirn ist ein System aus etwa 10^{11} Neuronen. Diese verhalten sich im Prinzip wie das oben beschriebene Modell, indem sie die Daten von anderen Neuronen gewichtet addieren und ein Ausgangssignal über eine nichtlineare Übertragungsfunktion erzeugen. Dieses Signal wird wiederum an andere Neuronen weitergegeben.

Damit unterscheidet sich das biologische System grundsätzlich von konventionellen künstlichen Informationsverarbeitungsanlagen, welche zumeist nach dem von-Neumann-Schema aufgebaut sind. Der von-Neumann-Rechner benutzt nur eine Schalteinheit, die sehr komplizierte Programme verarbeiten kann. Ein neuronales Netz hingegen besteht aus sehr vielen Prozessoren, welche jeweils ein einfaches Programm (Gleichung 2.1) ausführen.

Ein künstliches neuronales Netz wird eine einfachere Struktur haben als das biologische Vorbild. Die Verbindungen der Neuronen sind zumeist fest vorgegeben, und das Verhalten des Netzes wird durch die Gewichte w_{ij} festgelegt. Es gibt Netze, in denen keine Rückkopplung vorkommt, sondern der Informationsfluß nur in einer Richtung durch mehrere Lagen von Neuronen läuft. Diese nennt man *Feed-Forward-Netze*. Für die Spurerkennung sind jedoch Netze mit Rückkopplung von größerem Interesse. Eine Möglichkeit, Rückkopplungen einzuführen, wird im folgenden Abschnitt beschrieben.

2.4 Das Hopfield-Modell

Im *Hopfield*-Modell [7] ist im allgemeinen jedes Neuron mit jedem anderen verknüpft. Die Rückführung der Ausgänge an die Eingänge kann man durch $x_i = S_i$ beschreiben. Die Gewichte w_{ij} werden durch das Problem gegeben, das System wird zufällig oder trivial initialisiert. Im folgenden gilt:

- Alle Kopplungen zwischen den Neuronen sind symmetrisch:

  $$w_{ij} = w_{ji} \tag{2.2}$$

  Die Gewichte, die die Neuronen an sich selbst koppeln, sind null.

- Die Aktivität kann nur zwei Werte annehmen, und zwar

  $$S_i = \begin{cases} 1 & \text{Neuron } \textit{aktiv} \\ 0 & \text{Neuron } \textit{inaktiv} \end{cases} \quad . \tag{2.3}$$

  Als Übertragungsfunktion kann die Heaviside-Funktion Θ benutzt werden.

Die Aktualisierung des Neurons i sieht damit wie folgt aus:

$$S_i := \Theta\Big(\sum_i w_{ij} S_j - s_i\Big) \quad . \tag{2.4}$$

Diese Wertzuweisung kann für alle Neuronen synchron erfolgen. Das entspricht der Multiplikation eines Vektors $\mathbf{S}$ mit einer Matrix $\mathbf{w}$, der Addition des Schwellenvektors $\mathbf{s}$ und der Übertragung der Elemente durch die Heaviside-Funktion:

$$\mathbf{S} = \Theta(\mathbf{w} \cdot \mathbf{S} - \mathbf{s}) \quad . \tag{2.5}$$

Das Ergebnis der Vektor-Heaviside-Funktion wird erst nach der vollständigen Berechnung in $\mathbf{S}$ gespeichert.

Alternativ kann die Aktualisierung (2.4) asynchron erfolgen. Dann benutzt nur das zuerst aktualisierte Neuron i_1 den Anfangszustand $\mathbf{S}_{\mathsf{start}}$, alle später aktualisierten finden einen Eingabevektor $\mathbf{S}'$ mit einem geänderten S_{i_1} vor. Die asynchrone Aktualisierung kann entweder in einer immer gleichen oder in einer zufälligen Reihenfolge stattfinden.

2.5 Die Energiefunktion

Die Bedingung 2.2 garantiert, daß das Netz nach einigen Aktualisierungsschritten in einen stabilen Zustand läuft [7, 8]. Dieser stellt die Lösung des durch die Wahl der Gewichte $\mathbf{w}$ gegebenen Problems dar. Zur Darstellung der Aufgabe und zur Analyse des Verhaltens des Netzes kann man ihm eine sogenannte *Energie-* oder *Kostenfunktion* zuordnen:

$$E = -\frac{1}{2}\left(\sum_{ij} w_{ij} S_i S_j - 2\sum_i s_i S_i\right) \quad . \tag{2.6}$$

Der stabile Zustand, in den ein Hopfield-Netz läuft, entspricht einem Minimum der Energiefunktion. Das läßt sich wie folgt erklären:

Zur Minimumsuche kann man ein Gradientenabstiegsverfahren benutzen. Dieses verschiebt den Zustand des Systems $\mathbf{S}$ in kleinen Schritten in Richtung des negativen Gradienten. Der Gradient h_i am Ort des Neurons i („lokales Feld") schaltet in jedem Schritt die Aktivität S_i in Richtung kleiner werdender Energie.

$$\begin{aligned} h_i &= -\frac{\partial E(S)}{\partial S_i} \\ &= \sum_{j=1}^{n} w_{ij} S_j - s_i \end{aligned} \tag{2.7}$$

Die Iterationsgleichung für jede Aktivität S_i lautet damit bei der n-ten Iteration:

$$S_i^{(n)} = \Theta\left(h_i^{(n-1)}\right) \quad , \tag{2.8}$$

was genau der Aktualisierungsgleichung des Hopfield-Netzes (2.4) entspricht.

Es lassen sich viele Probleme in der Form der Energiefunktion (2.6) beschreiben, wobei die Lösung zummeist in einem *globalen* Minimum liegt.

2.6 Stochastische Modelle

Da die Energiefunktion oft sehr viele lokale Minima hat, führt ein einfaches Gradientenabstiegsverfahren nicht immer zum gesuchten globalen Minimum. Deshalb wird eine thermische Bewegung des Systems simuliert, welche es bei einer höheren Temperatur aus flachen Minima stochastisch herausschieben kann.

2.6.1 Erweiterung des Hopfield-Modells

Das Neuron i werde in jedem Aktualisierungsschritt mit einer Wahrscheinlichkeit von P_i^+ aktiviert und mit einer Wahrscheinlichkeit von P_i^- nicht aktiviert, es gilt also $P_i^+ + P_i^- = 1$. Die Heaviside-Funktion aus der Aktualisierungsgleichung (2.8) wird dazu durch die Fermifunktion ersetzt und die Temperatur $T = 1/\beta$ eingeführt.

$$P_i^{\pm} = \frac{1}{1 + \exp\left(\mp 2\beta h_i\right)} \tag{2.9}$$

Bei niedrigen Temperaturen ($\beta \to \infty$) geht $P_i^{\pm}$ gegen 0 oder 1, damit geht die stochastische Aktualisierung in (2.4) über. Bei höheren Temperaturen konvergiert das System langsam; die Wahrscheinlichkeit, das globale Minimum zu finden, steigt aber. Ein Verfahren, den Temperaturparameter zu variieren, wird in Abschnitt 2.7 beschrieben.

2.6.2 Die Mean-Field-Approximation

Die stochastische Erweiterung des Hopfield-Modells läßt sich durch die *Mean-Field*-Approximation annähern, die die diskreten Aktivierungen durch kontinuierliche Größen ersetzt. Man berechnet dazu die Erwartungswerte V_i der Aktivierungen S_i, sowie die Erwartungswerte u_i der lokalen Felder h_i:

$$V_i \equiv \langle S_i \rangle \quad \text{und} \quad u_i \equiv \langle h_i \rangle \tag{2.10}$$

Diese Mittelwerte stellen nach der aus der statistischen Physik bekannten *Mean-Field-Theory (MFT)* eine Approximation der dynamischen Variablen des Systems dar [9]. Mit den Wahrscheinlichkeiten P_{i+}, ein Neuron i zu aktivieren, und P_{i-}, ein Neuron in den inaktiven Zustand zu versetzen, gilt:

$$\begin{aligned} V_i &= \frac{1}{2}(P_{i+} - P_{i-} + 1) \\ &= \frac{1}{2}\Big(\frac{1}{1+\exp(-2\beta u_i)} - \frac{1}{1+\exp(+2\beta u_i)} + 1\Big) \\ &= \frac{1}{2}\Big(\tanh(\beta u_i) + 1\Big) \quad . \end{aligned} \tag{2.11}$$

Mit $u_i = -\partial E / \partial V_i$ ergibt sich die MFT-Aktualisierungsgleichung

$$V_i = \frac{1}{2}\left(1 + \tanh\Big(-\beta \frac{\partial E}{\partial V_i}\Big)\right) \quad . \tag{2.12}$$

Das bedeutet, daß ein Hopfield-Netz mit der Übertragungsfunktion

$$g(z) = \frac{1}{2}\Big(1 + \tanh(\beta z)\Big)$$

näherungsweise die Eigenschaften des stochastischen Modells aus Abschnitt 2.6.1 hat.

2.7 Das Simulated-Annealing-Verfahren

Die Temperatur T bewirkt ein Abflachen der Energiefunktion, so daß nebeneinanderliegende Minima verschmelzen können. Bei einer hohen Temperatur ist nur ein Energieminimum vorhanden, welches leicht gefunden werden kann. Wenn man die Temperatur senkt, spaltet sich das Minimum auf. Diese Auftrennung entspricht einem Phasenübergang in der Thermodynamik [9]. Das bisherige Minimum wird dabei zu einem Sattelpunkt, und ein neues absolutes Minimum befindet sich in nächster Nähe, sofern die Temperatur sehr wenig unter der Phasenübergangstemperatur liegt.

Beim langsamen Senken der Temperatur bis $T \to 0$ („einfrieren") kann man sich so bis zum absoluten Minimum der ursprünglichen Energiefunktion (2.6) vortasten. Man nennt diesen Prozeß *Simulated Annealing* [10].

3 Die Denby-Peterson-Methode

3.1 Ein neuronaler Algorithmus zur Spurfindung

Denby [11] und Peterson [12] schlugen 1987 und 1988 eine Möglichkeit der Spurerkennung in Detektoren mit Hopfield-Netzen vor. Sie wurde bereits erfolgreich auf Daten existierender Experimente angewendet. In [13] ist zum Beispiel die Rekonstruktion von Daten der *ALEPH TPC*[1] mit dem Algorithmus nach Denby und Peterson beschrieben.

Bei der Denby-Peterson-Methode werden alle Verbindungen zweier Hits, die benachbarte Punkte einer Spur sein können, mit jeweils einem Neuron identifiziert. Im Gleichgewichtszustand sollen genau die Neuronen aktiviert sein, die auf Spuren liegen. Das Netz hat nur einfache Kriterien zur Beurteilung der Verbindungen. Der Algorithmus setzt voraus, daß die zu suchenden Spuren in etwa gerade sind. Sie sollen, wenn sie sich an einem Hit treffen, einen Winkel möglichst nahe bei π einschließen. Kurze Segmente werden bevorzugt. Weiterhin werden Verzweigungen explizit unterdrückt und eine ungefähre Anzahl zu aktivierender Segmente vorgegeben.

Die Hopfield-Energiefunktion aus Gleichung 2.6 wird zu

$$\begin{aligned} E &= -\frac{1}{2}\sum \delta_{jk}\frac{-\cos^m \theta_{ijl}}{d_{ij}+d_{jl}} S_{ij}S_{kl} \\ &\quad + \frac{1}{2}\alpha\left(\sum_{l\neq j} S_{ij}S_{il} + \sum_{k\neq i} S_{ij}S_{kj}\right) + \frac{1}{2}\delta\left(\sum S_{kl} - N\right)^2 \quad . \end{aligned} \tag{3.1}$$

Hier werden die Indizes i, j, k und l benutzt, um die Detektorhits zu kennzeichnen, die Aktivität S_{ij} beschreibt den Status des Segments, welches von Hit i auf Hit j zeigt. Der Winkel zwischen zwei Segmenten (ij) und (jl) ist[2] $\theta_{ij,jl} \equiv \theta_{ijl}$, die Länge des Segments (ij) ist d_{ij}. Die Terme der zweiten Zeile sind Zwangsbedingungen mit den Lagrange-Multiplikatoren α und δ. Sie sind so gewählt, daß die Terme ungefähr null sind, falls die Bedingung erfüllt ist, und positiv sind, falls die Bedingung nicht erfüllt ist. Der Beitrag zweier Segmente, die sich einen Hit teilen, ist ohne die Zwangsbedingungen in Abbildung 3.5 dargestellt.

Der α-Term unterdrückt Verzweigungen anhand der Orientierung aneinanderstossender Neuronen, wie in Abbildung 3.1 dargestellt. Die Orientierung der Neuronen

[1] TPC = *Time Projection Chamber* (Typ einer Driftkammer, der dreidimensionale Koordinaten liefert)

[2] Der Winkel θ in dieser Arbeit ergänzt sich mit dem in [12] und [14] benutzten zu π, daher bekommt der cos-Term ein zusätzliches Minus-Zeichen. Das Vorzeichen des cos-Terms sollte in der Energiefunktion vorkommen [14], der Exponent m wird deshalb als ungerade gewählt.

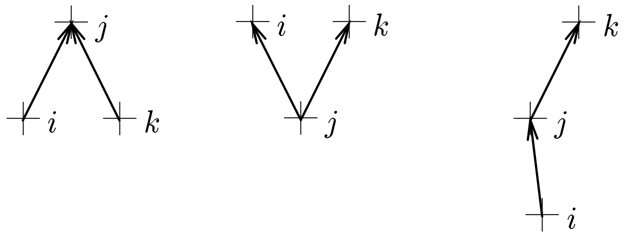


Abbildung 3.1: Zu unterdrückende Verzweigungen $S_{ij}S_{kj}$ und $S_{ji}S_{jk}$ und die gesuchte Form $S_{ij}S_{jk}$. Die Neuronen wurden so generiert, daß sie eine positive vertikale Komponente haben.

muß aus einer Ordnung der Hits folgen, beispielsweise könnte man für ein Fixed-Target-Experiment mit einem Koordinatensystem wie bei HERA-B nur Neuronen mit $\Delta z > 0$ erzeugen. Verzweigungen kann man daran feststellen, daß die ersten Indizes oder die zweiten Indizes eines $S_{ij}S_{kl}$-Terms gleich sind. Der gewünschte Zusammenhang ist dagegen $S_{ij}S_{jk}$ mit $i \neq k$. Der Parameter N im δ-Term muß etwa bei der Anzahl der Hits abzüglich des erwarteten Detektorrauschens liegen.

3.2 Erweiterungen des Denby-Peterson-Algorithmus für die Anwendung bei HERA-B

Es zeigt sich, daß die Energiefunktion 3.1 für schwierige Anwendungen Probleme bereitet:

- Die Zwangsbedingungen zur Spurverzweigung erfordern eine sonst nicht notwendige Orientierung der Segmente.

- Es gibt keine Möglichkeit, Fehlschaltungen wie in Abbildung 3.2 systematisch aufzulösen. Das System befindet sich in einem lokalen Energieminimum, und ein zusätzliches richtiges Segment verletzt die Zwangsbedingung unabhängig vom Winkel der beteiligten Segmente. Das Anpassen der „Energieschwelle“ zum globalen Minimum ist nur schwer möglich.

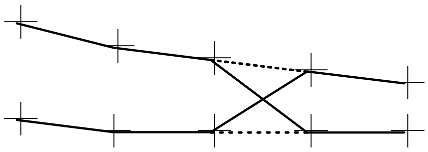


Abbildung 3.2: Schema falscher Aktivierungen bei nahe benachbarten Spuren

- Die Abstandsfunktion bewirkt eine ungleichmäßige Berücksichtigung der verschiedenen Detektorlagen einer Superlagenstruktur.

- Bei sehr ungünstigen Winkeln unterhalb von $\pi/2$ werden die Energie-Beiträge aus dem $S_{ij}S_{kl}$-Term längerer Segmente günstiger als die der eigentlich zu bevorzugenden kurzen Segmente. In Abbildung 3.5 wird diese Eigenschaft durch die vom linken Bildrand nach unten weisenden Höhenlinien verdeutlicht.

3.2.1 Modifizierte Energiefunktion

Man kann auf beide Zwangsbedingungen verzichten, wenn man die Energiefunktion wie folgt abändert:

- Die Kosten für ungünstige Winkel und die für die Segmentlänge werden linear superponiert:

$$\frac{\cos^m \theta_{ijl}}{d_{ij} + d_{jl}} \quad \longrightarrow \quad f(\cos^m \theta_{ij,kl}) + \alpha \left(\delta - \frac{1}{d_{ij}}\right) \quad .$$

- Der Beitrag eines Segments zur Energiefunktion soll positiv sein für alle Winkel außer den gewünschten nahe bei π (Abbildung 3.3).

- Eine typische inverse Segmentlänge δ wird eingeführt. Die Abhängigkeit der Energiefunktion ist in Abbildung 3.4 für $\delta = 3\,\mathrm{m}^{-1}$ gezeigt.

Der Verzicht auf die Zwangsbedingungen erlaubt es, die Segmente als nicht-orientiert anzusehen. Da die Gewichte symmetrisch sind (Hopfield-Bedingung 2.2), reicht es, nur die untere Dreiecksmatrix zu berechnen und den fehlenden Faktor 2 auszugleichen. Die Energiefunktion lautet damit:

$$E = \sum_{ij\,|i<j} \left(\frac{1}{2} \sum_{kl\,|k<l} t_{ij,kl} S_{kl} + \alpha \left(\delta - \frac{1}{d_{ij}}\right) \right) S_{ij} \tag{3.2}$$

$$t_{ij,kl} = \begin{cases} \gamma + |\cos\theta_{ij,kl}|^m \times \begin{cases} \eta & \text{für } \cos\theta_{ij,kl} \geq 0 \\ -1 & \text{für } \cos\theta_{ij,kl} < 0 \end{cases} & \text{für } \begin{array}{l} j=k \wedge i \neq l \\ j=l \wedge i \neq k \\ i=k \wedge j \neq l \\ i=l \wedge j \neq k \end{array} \\ 0 & \text{sonst} \end{cases}$$

Daraus ergibt sich das „lokale Feld“:

$$h_{ij} = -\frac{\partial E}{\partial S_{ij}} = -\left(\sum_{kl\,|k<l} t_{ij,kl} S_{kl} + \alpha \left(\delta - \frac{1}{d_{ij}}\right) \right) \quad . \tag{3.3}$$

Der Energiebeitrag zweier Segmente, die einen gemeinsamen Hit haben, ist in Abhängigkeit vom Winkel und der Segmentlänge, die für beide gleich groß sein soll, in Abbildung 3.6 dargestellt.

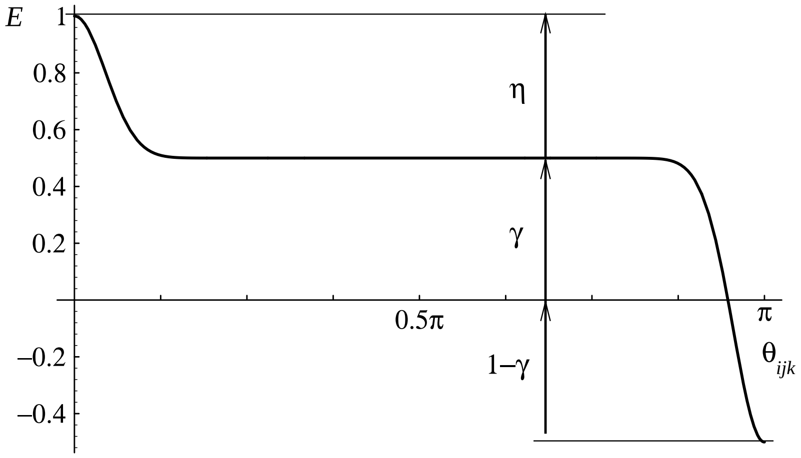


Abbildung 3.3: Modifizierte Energiefunktion $E = E(\theta_{ijk})$. Die Parameter sind der Anschaulichkeit wegen nicht die später (Tabelle 3.1) benutzten.

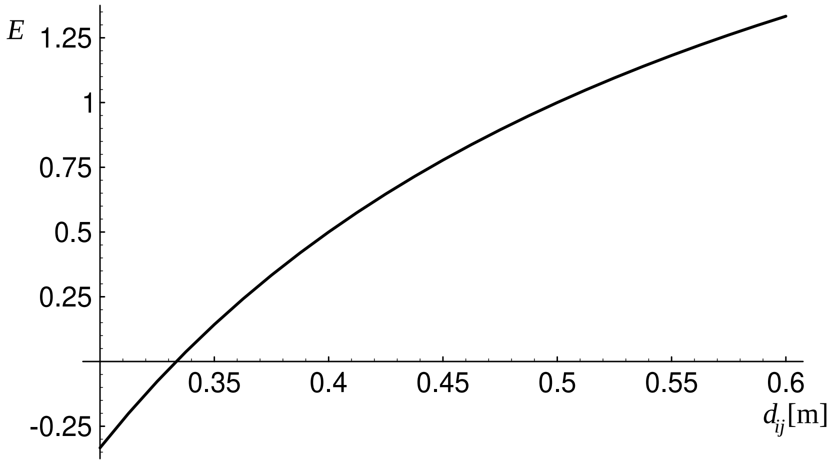


Abbildung 3.4: Modifizierte Energiefunktion $E = E(d_{ij})$

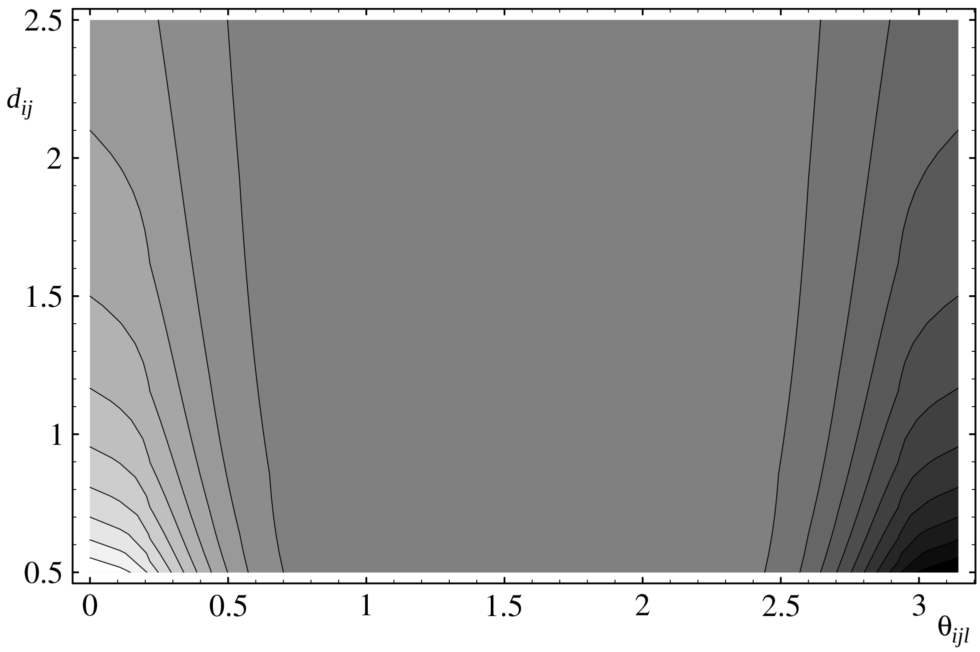


Abbildung 3.5: Petersons Potential $E(\theta_{ijl}, d_{ij} = d_{jl})$

$E_{\max} \rightarrow$ weiß
$E_{\min} \rightarrow$ schwarz

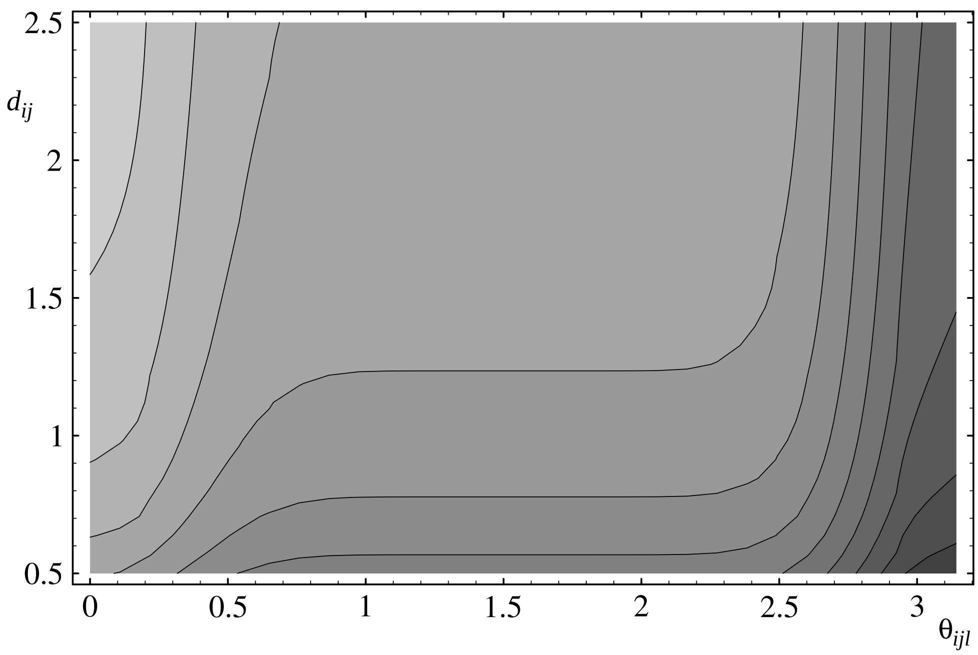


Abbildung 3.6: Modifizierte Energiefunktion $E(\theta_{ijl}, d_{ij} = d_{jl})$

3.2.2 Erweiterte Abstandsdefinition

Der Parameter δ der Energiefunktion gibt die typische inverse Segmentlänge an. Bei einer Superlagenstruktur haben manche Segmentlängen aber einen Unterschied von etwa einer Größenordnung. Es wird keine der gesuchten Segmentketten gefunden, solange $\vec{d}_{ij}$ der euklidische Abstandsvektor ist. Abhilfe schafft hier das gedachte Zusammenrücken der Superlagen.

Für einen aus einer komplexen Struktur verschiedenartiger Module aufgebauten Detektor ist ein einfaches gedachtes Verschieben der Lagen nicht ausreichend. Hier muß man sich verdeutlichen, daß der Term $1/|\vec{d}_{ij}|$ verhindern soll, daß passende Hits einer Detektorlage ausgelassen werden und sollte diese Wirkung in einer erweiterten Abstandsdefinition beibehalten.

Die zu den Detektorlagen parallele Komponente des Abstandsvektors bewirkt außerdem, daß Segmente, die senkrecht auf die Detektorlagen zeigen, bei gleicher Anzahl durchquerter Lagen bevorzugt werden. Dieser Effekt ist zumeist erwünscht und sollte auch bei einer modifizierten Abstandsdefinition vorkommen. In unserem Beispiel ist es die x-Komponente.

Die für die folgenden Untersuchungen benutzte Abstandsfunktion ist deshalb definiert als:

$$d_{ij} := \sqrt{\Big((n_{ij}+1)\, d_{\text{typ}}\Big)^2 + (\vec{d}_{ij})_x^2} \quad . \tag{3.4}$$

Die Anzahl der Detektorlagen zwischen Hit i und Hit j bezeichnen wir mit n_{ij}, den Lagenabstand innerhalb einer Superlage mit d_{typ} und den geometrischen Abstandsvektor mit $\vec{d}_{ij}$. Die x-Komponente des Abstands ist $(\vec{d}_{ij})_x$.

3.3 Beschreibung des Algorithmus

Zu Beginn der Rekonstruktion liegen die Koordinaten der Detektorhits vor. Zunächst werden die Neuronen (ij) mit $S_{ij} = 0$ als Objekte im Rechnerspeicher erzeugt und mit den Hits i und j verknüpft. Danach werden die Gewichte $t_{ij,kl}$ und die Schwellen $\alpha(\delta - 1/d_{ij})$ berechnet.

Wenn das Netz auf diese Weise initialisiert ist, kann die eigentliche Spurensuche beginnen. In unserem Fall werden die Neuronen nacheinander und immer in derselben Reihenfolge aktualisiert.[3] Dabei wird die stochastische Aktualisierung nach Abschnitt 2.6.1 benutzt. Während des Iterierens sollte überprüft werden, ob ein stabiler Zustand eingetreten ist. Wenn sich die Aktivitäten S des Netzes bei mehreren Iterationen nicht ändern, ist zumeist[4] der gewünschte Zustand erreicht.

Zur Identifizierung der Spuren muß der stabile Zustand, in den das Hopfield-Netz läuft, untersucht werden. Im einfachsten Fall gibt es keine Kreuzungen oder Verzwei-

[3] Eine zufällige Reihenfolge kann zu schnellerer Konvergenz führen, ist aber auch etwas aufwendiger zu implementieren [13].

[4] sofern $T \neq 0$ kann wegen des stochastischen Einflusses genaugenommen kein konstanter stabiler Zustand erreicht werden, in der Praxis läßt sich aber ein ausreichend kleines T finden.

gungen der Neuronen. Dann kann jede Kette von Verbindungen einer Spur zugeordnet werden.

Bei vielen nahezu parallelen Spuren wie in HERA-B kommt es allerdings oft vor, daß sich Spuren ein Segment teilen. Dieses wird bei einem einfachen Aufsammeln der Segmentkette einer der Spuren zugeordnet, wobei die andere Kette zerstört wird. Außerdem kann auch die eine Spur aus den Teilketten von zwei verschiedenen Spuren zusammengesetzt und damit nicht mehr korrekt rekonstruiert werden. Dieses Problem wird in Abschnitt 3.5.2 näher erörtert.

3.4 Der Modelldetektor

Zur Beurteilung des Algorithmus soll ein Modellexperiment dienen. Dieses besteht aus folgenden Komponenten:

Der Prozeß erzeugt Teilchen der Masse $0.137\,\mathrm{GeV}/c^2$ am Ort $\vec{x} = (0,0,0)$ mit einer Energie, die zwischen $1\,\mathrm{GeV}$ und $10\,\mathrm{GeV}$ gleichverteilt ist. Der Azimuth- (ϕ) und der Polarwinkel (θ) sind im Schwerpunktsystem jeweils gleichverteilt, was eine nicht-isotrope Verteilung bedeutet. Die Ladung ist $\pm e$. Die Teilchen werden mit $\gamma = 20$ in z-Richtung geboostet.

Der Magnet biegt die Spuren im Bereich $3.5\,\mathrm{m} \leq z \leq 5.5\,\mathrm{m}$ entsprechend einer Feldstärke von $B = 2\,\mathrm{T}$ in der xz-Ebene ab.

Die Detektorkammern sind wahlweise äquidistant oder in einer Superlagenstruktur angelegt. Sie verhalten sich wie Driftkammern, welche eine eindeutige Ortsinformation liefern, also keine Rechts-Links-Ambiguitäten besitzen, und in y-Richtung verlaufen. Ihre Eigenschaften werden durch die Parameter Hit-Effizienz, Rausch-Rate, Auflösung und Größe einer Kammer beschrieben. Die Rauschrate wird wie in [15] als Anteil der echten Hits angegeben.

Die Liste der Hits wird an das Mustererkennungsprogramm weitergegeben. Dieses soll sie den Spuren entsprechend zusammenfassen. Zum Vergleich mit dem HERA-B-Detektor wird die Belegungsdichte in dem Teil des Detektors, der zumeist die maximale Dichte hat, gemessen. Die maximale Belegungsdichte liegt bei HERA-B bei bis zu 20%. Die Detektorparameter werden, wie in Tabelle 3.1 aufgelistet, ähnlich denen von HERA-B gewählt.

Der Modelldetektor mißt nur die zweidimensionale Projektion des Ereignisses, entsprechend den 0°-Lagen des HERA-B-Systems. In Abbildung 3.7 ist ein einfaches Ereignis mit 38 Spuren und einer maximalen Belegungsdichte von 8 % dargestellt. Der Erzeugungsvertex und der Magnet liegen links außerhalb des Bildes. Abbildung 3.9 zeigt den Zustand des Netzes nach einer Iteration, Abbildung 3.8 zeigt den stabilen Zustand des Netzes nach 36 Iterationen.

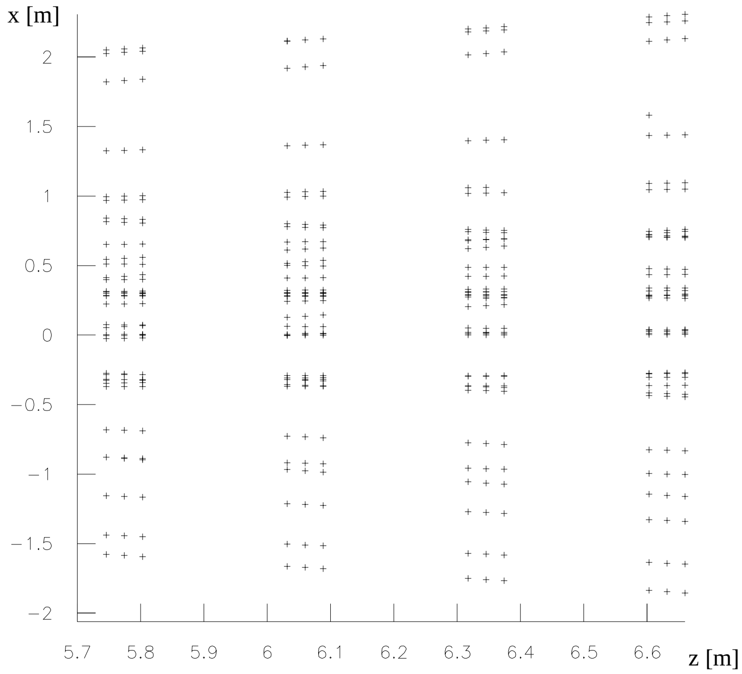


Abbildung 3.7: Hits im Modelldetektor

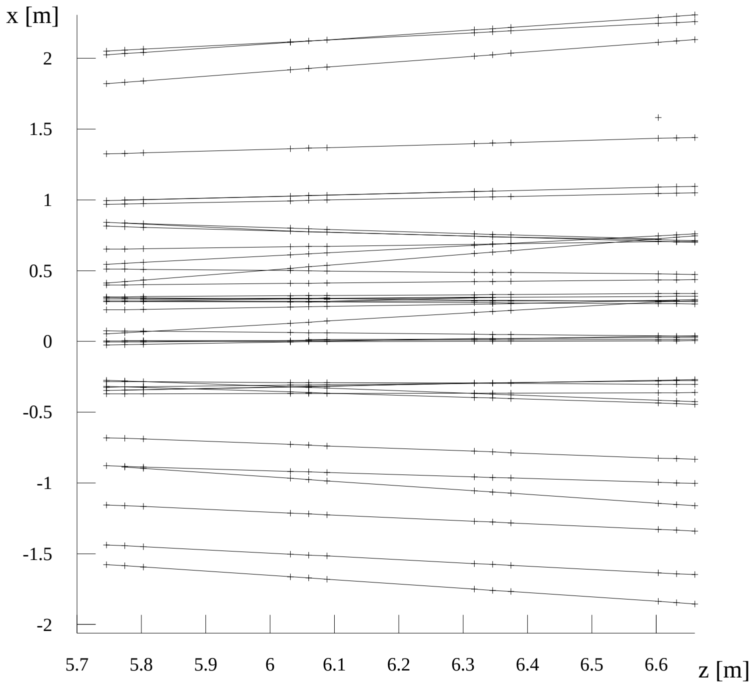


Abbildung 3.8: Lösung des Problems aus Abbildung 3.7 mit dem Algorithmus nach Denby-Peterson

Anzahl Superlagen:	4
Detektorlagen pro Superlage:	3
Abstand der Superlagen	0.3 m
Abstand der Detektorlagen	0.03 m
Auflösung:	200 μm
Zelldurchmesser:	5 mm
Treffer-Effizienz:	98 %
Detektor-Rauschen:	1 %

Tabelle 3.1: Parameter des Modelldetektors

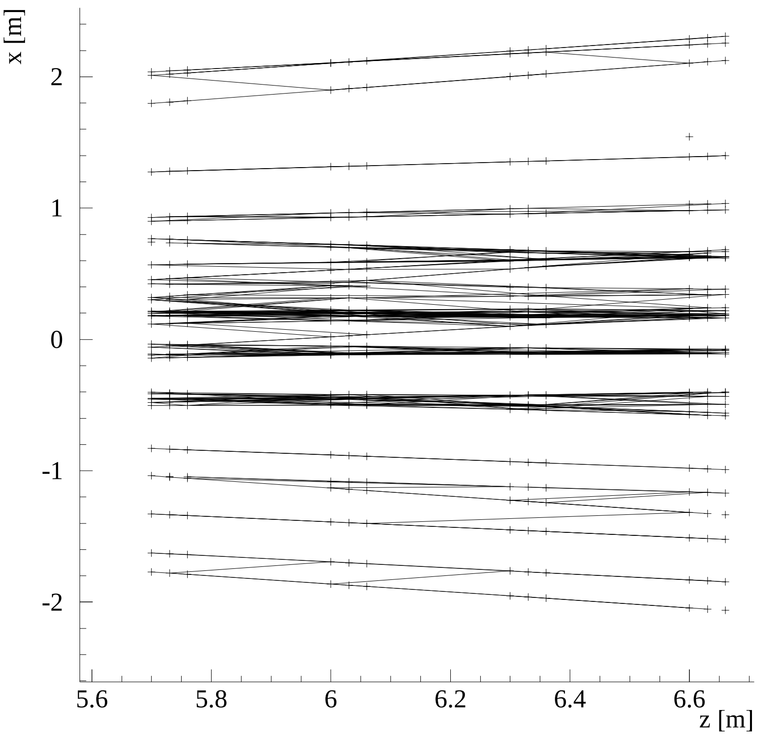


Abbildung 3.9: Zustand des Netzes nach einer Iteration.

3.5 Ergebnis für den Denby-Peterson-Algorithmus

3.5.1 Definition der Spurfindungseffizienz

Die Definition einer gefundenen Spur orientiert sich hier an der Zuordung der Hits zu den Spuren. Eine alternative Definition ist das sogenannte *Parameter-Matching*, welches die Parameter der gefundenen Spuren mit denen der zu suchenden Spuren vergleicht. Dazu müssen sorgfältige Genauigkeitsabschätzungen gemacht werden. Nachfolgende Ausgleichsrechnungen, wie etwa die Vertexrekonstruktion, benötigen die Kovarianzmatrizen der rekonstruierten Spuren. In eine Parameter-basierte Effizienzberechnung müssen also auch die Kovarianzmatrizen eingehen.

Wegen kürzeren Beschreibung wird im vorliegenden Fall eine Hit-orientierte Effizienzdefinition benutzt, und wie bei der konventionellen Spurensuche in den HERA-B-Spurkammern gilt eine Teilchenspur als gefunden, wenn

- sie eine Mindestanzahl zugeordneter Hits hat, im Fall der zweidimensionalen Spurensuche sind es acht,
- und von den zugeordneten Hits 70% vom selben simulierten Teilchen stammen.

3.5.2 Messung der Spurfindungseffizienz

Zur Messung der Spurfindungseffizienz des Algorithmus nach Denby und Peterson wurde das Hopfield-Netz mit den in Tabelle 3.2 zusammengefaßten Parametern für Gleichung 3.3 benutzt. Die Parameter wurden durch systematisches Probieren gefunden. Für die Aktualisierung der Neuronen wurde das stochastische Verfahren aus Abschnitt 2.6.1 in Verbindung mit der Simulated-Annealing-Methode benutzt. Der Temperaturparameter β_j für die Iteration j wurde wie folgt gewählt:

$$\beta_j = \beta_{\text{start}} \left(\frac{\beta_{\text{end}}}{\beta_{\text{start}}} \right)^{j/n} \quad . \tag{3.5}$$

Falls sich die Aktivierung keines der Neuronen über drei Iterationszyklen änderte, wurde angenommen, daß das Netz einen stabilen Zustand erreicht hatte und das Iterieren abgebrochen.

Zur Auswertung wurde nach Ketten von aktivierten Neuronen gesucht. Dabei wurde bei den Hits mit maximalem z begonnen. An Kreuzungen oder Verzweigungen wurde in möglichst gerader Richtung fortgefahren. Einmal benutzte Aktivierungen wurden gelöscht.

Die Spurfindungseffizienz ist in Abbildung 3.10 dargestellt. Dazu wurden 300 Ereignisse mit einer maximalen Belegungsdichte zwischen 0 und 30% benutzt. Es zeigt sich, daß die Effizienz bis zur einer maximalen Belegungsdichte von 20% bei über 85% liegt. Die unteren Meßpunkte geben die relative Anzahl falsch gefundener Spuren („Geister“) an.

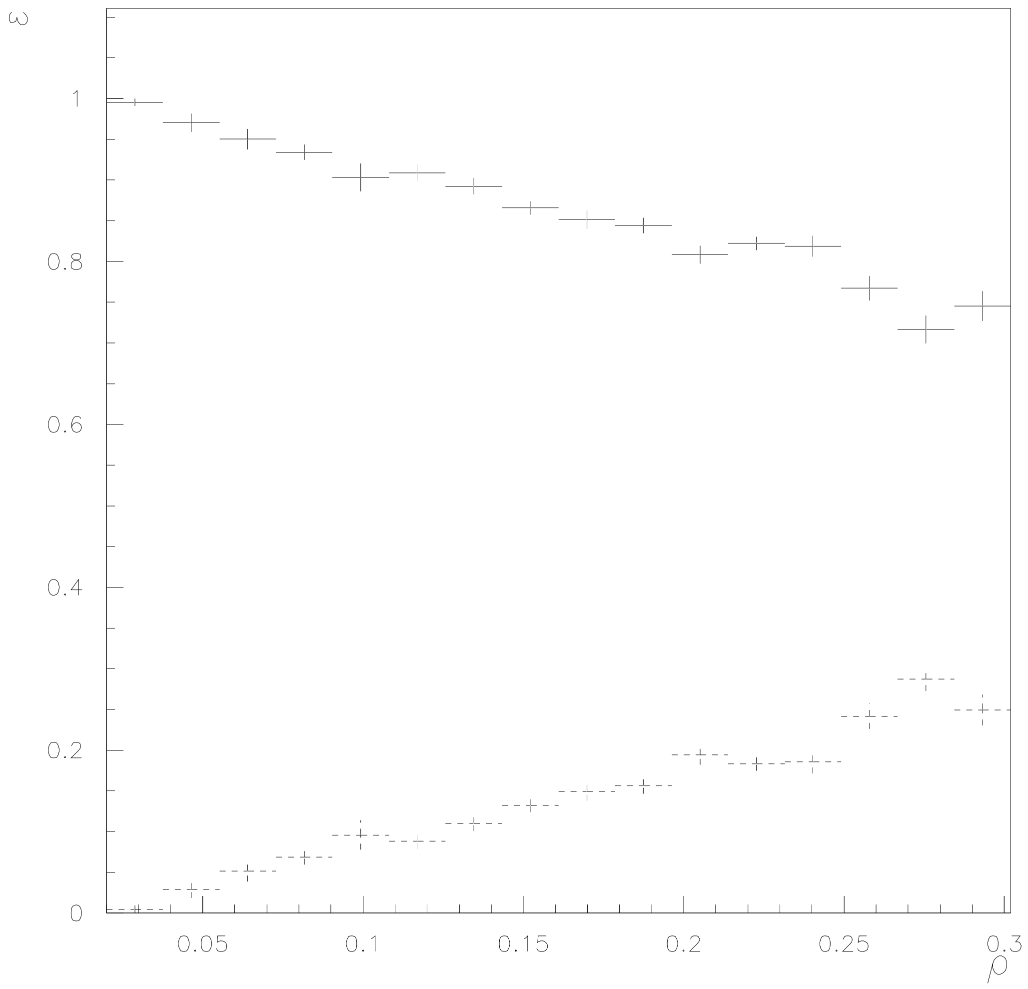


Abbildung 3.10: Spurfindungseffizienz ε in Abhängigkeit von der maximalen Belegungsdichte ρ des Denby-Peterson Algorithmus im Modelldetektor, untere Einträge: Geisterrate

Koeffizient des Abstands-Terms	α	$= 0.3$
Exponent	m	$= 71$
Offset	γ	$= 0.5$
Asymmetry	η	$= 0.5$
erwarteter Hit-Abstand	δ	$= 0.9 \ (d_{\text{typ}})^{-1}$
typischer Abstand der Lagen	d_{typ}	$= 0.01\,\text{m}$
inverse Anfangstemperatur	β_{start}	$= 0.5$
inverse Endtemperatur	β_{end}	$= 300$
maximale Anzahl von Iterationen	n	$= 50$

Tabelle 3.2: Parameter für den Denby-Peterson-Algorithmus

Eigenschaften des Denby-Peterson-Algorithmus

Die Effizienz läßt sich nicht direkt mit anderen Mustererkennungsalgorithmen für HERA-B vergleichen, da der Modelldetektor nur einen Teil der Eigenschaften von HERA-B simuliert. Man kann jedoch grundsätzliche Eigenschaften des Denby-Peterson-Algorithmus festhalten:

- Es wird kein parametrisierbares Spurmodell benutzt. Man ignoriert damit Informationen über das System, die für die Spurfindung nützlich sein können. Zum Beispiel wird eine Spur, deren Steigung bei jedem Hit systematisch zunimmt, nicht von einer geraden Spur unterschieden, deren Hits stochastisch abweichen. Im HERA-B-Pattern-Recognition-Detektor könnte man den ersten Fall als Spurkandidaten ausschließen.

- Die endliche Auflösung des Detektors wird nicht direkt berücksichtigt, da sie als lokale Spurkrümmung beschrieben wird. Für diese gilt die Bemerkung des vorhergehenden Punktes.

Rechenzeit und Speicherplatzbedarf

Die zur Spurfindung benötigte Zeit teilt sich beim Denby-Peterson-Algorithmus auf in Initialisierungszeit und Iterationszeit. Es zeigt sich, daß letztere stärker mit der Spurmultiplizität steigt. Deshalb soll sie genauer diskutiert werden. Die Zahl der erzeugten Neuronen steigt quadratisch mit der Spurdichte. Die Zahl der von null verschiedenen Matrixelemente jeder Zeile der Gewichtsmatrix steigt mit der Anzahl der Neuronen, die an das entsprechende Segment angrenzen.[5] Ihre Zahl steigt ungefähr proportional zur Spurdichte. Damit müßte näherungsweise gelten:

$$\text{Bearbeitungszeit } T(\rho) \propto \rho^3 \quad .$$

[5]In der hier verwendeten sequentiellen Implementierung des Algorithmus werden gezielt nur die Gewichte $w_{ij} \neq 0$ benutzt.

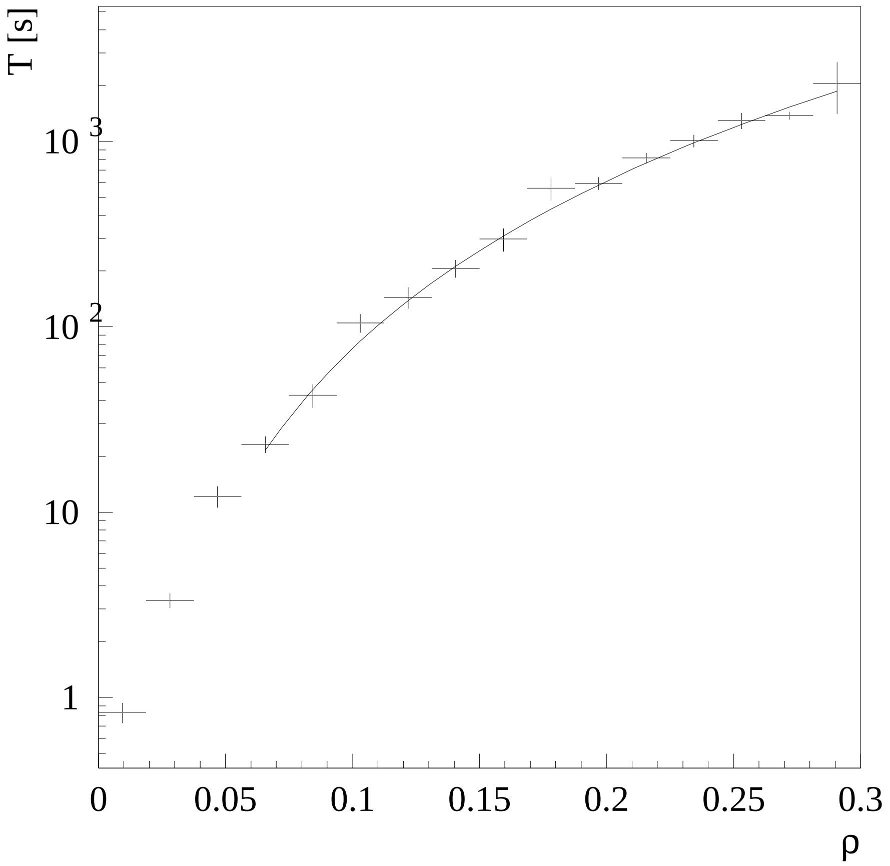


Abbildung 3.11: Rechenzeit T des Denby-Peterson-Algorithmus in Abhängigkeit von der maximalen Belegungsdichte ρ im Vergleich zur Ausgleichskurve $T(\rho) = 7.63 \cdot 10^4 \rho^3$.

Die Bearbeitungszeit für ein Ereignis in Abhängigkeit von der maximalen Belegungsdichte ρ auf einer *SGI/Challenge* (MIPS R4400-Prozessoren) ist in Abbildung 3.11 dargestellt. Dort ist die maximale Anzahl von Iterationen $n_{\mathrm{max}} = 50$, die bei Ereignissen mit $\rho_{\mathrm{max}} \gtrsim 0.1$ auch durchgeführt werden, da das Netz vorher keinen konstanten Zustand erreicht. Die Ausgleichskurve, die die oben geschätze Abhängigkeit benutzt, beschreibt offenbar tatsächlich die Daten.

Mit steigender Zahl von null verschiedener Gewichte wird auch die Energiefunktion komplizierter, so daß der Annealing-Prozeß in kleineren Schritten ausgeführt werden muß und insgesamt mehr Iterationen nötig sind. Man kann auch bei Belegungsdichten, bei denen die Spurfindungseffizienz nach Abbildung 3.10 absinkt, eine höhere Effizienz erzielen, wenn man die Iterationszahl in Abhängigkeit von der maximalen Belegungsdichte erhöht. Das würde allerdings zu einer noch stärkeren Abhängigkeit der Rechenzeit führen.

Die Iterationsgleichung eines synchron aktualisierten Hopfieldnetzes (2.5) hat eine

Struktur, die im Prinzip eine parallele Implementierung auf einem Rechner mit sogenannter *Single-Instruction-Multiple-Data*-Architektur (*SIMD*) erlaubt. In diesem Fall müssen soviele Prozessoren wie Neuronen vorhanden sein, da jeder Prozessor eine Zeile der Gewichtsmatrix **w** speichern muß. Die Elemente des Eingabevektors **S** werden nacheinander an alle Prozessoren gleichzeitig geliefert, so daß jeder Prozessor eine Multiplikation des Elementes durchführen kann und die Ergebnisse summieren kann. Danach hat jeder Prozessor ein Element des Ergebnisvektors errechnet.

Da beim HERA-B-Spurfindungsproblem mehr als 10^5 mögliche Verbindungen zwischen Hits entstehen können, bräuchte man für diese Implementierung größere Rechner, als heute existieren. Andererseits sind nur sehr wenige Elemente einer Zeile der Gewichtsmatrix nicht null, so daß eine sequentielle Berechnung des Ergebnisvektors kaum langsamer sein wird. Eine parallele Implementierung scheint daher mit heute existierender Hardware nicht sinnvoll. Ein ähnlicher Schluß wird auch in [13] gezogen.

3.5.3 Ausblick

Abbildung 3.10 zeigt, wie eine weitere Verbesserung des Algorithmus erzielt werden kann. Der Anteil der falschen Spuren („Geister“) ergänzt sich mit dem Anteil der gefundenen Spuren ziemlich genau zu eins. Das ist ein Hinweis darauf, daß das Hopfield-Netz zumeist die richtigen Segmente ermittelt hat, allerdings noch einige Kreuzungspunkte oder geteilte Segmente übrig sind. Diese werden bei der Auswertung oftmals falsch verfolgt und ergeben deshalb falsche Hitkombinationen. Die Verbesserung sollte folglich hier ansetzen, etwa mit einer Ausgleichsrechnung, die die günstigsten Zuordnungen bei Kreuzungen feststellt, und einer Möglichkeit, Segmente, die zu zwei Ketten gleichzeitig gehören, zu berücksichtigen.

Eine Anwendung auf die HERA-B-Simulation bedeutet, daß Driftkammersignale ausgewertet werden müssen. Die Energiefunktion läßt sich dazu so erweitern, daß beide möglichen Spurdurchgänge als Hits angesehen werden und eine gegenseitig hemmende Wechselwirkung der beteiligten Neuronen eingeführt wird.

Eine naheliegende Möglichkeit, die dreidimensionalen Informationen des HERA-B-Detektors zu verarbeiten, ist das Zusammenfassen der Hit-Informationen aus verschieden ausgerichteten Detektorlagen zu kartesischen Koordinaten. Diese können mit dem Denby-Peterson-Algorithmus wie beschrieben über die Winkel und Segmentlängenbedingungen ausgewertet werden [15].

4 Der Elastic-Arms-Algorithmus

4.1 Neuronale Klassifikation und Parameter-Ausgleichsrechnung

Der Elastic-Arms-Algorithmus wurde 1992 von M. Ohlsson, C. Peterson und A. L. Yuille vorgestellt [16] und erfolgreich auf simulierte Daten der *DELPHI*-TPC angewendet. Er paßt vorgegebene Parameterdarstellungen der Spuren den Detektorhits an. Eine solche Parameterdarstellung wird im folgenden als *Template* bezeichnet.[1]

Das Ereignis bestehe aus N Hits, welche N geometrische Objekte definieren, die die Teilchenspuren berühren. Diese N geometrischen Objekte werden durch $\{\xi_1, \ldots, \xi_N\}$ dargestellt. Für Streifendetektoren sind sie linienförmig, für Driftkammern haben sie die Form der Isochronen, welche zumeist durch Zylindermäntel darstellbar sind. Für Pixeldetektoren oder Zeitprojektionskammern (TPCs) sind die ξ punktförmig. Zu jedem Ereignis gibt es einen Satz $\mathbf{t} = \{t_1, \ldots, t_M\}$ von M gesuchten Teilchentrajektorien, deren Parameterdarstellungen im einfachsten Fall denen von Geraden oder Helizes entsprechen.

Der quadratische Abstand eines Templates a zum Hit i sei $M_{ia}(\xi_i, t_a)$, die Zuordnung eines Hits i zu einem Template a sei durch $S_{ia} \in \{0, 1\}$ gegeben. Dem Template $a = 0$ sollen alle Rauschsignale zugeordnet werden. Die Berechnung der Abstände ist in Anhang A.2 zusammengefaßt. Gesucht werden die Parameter $\mathbf{t}$, die den Energieterm

$$E'(\mathbf{S}, \xi, \mathbf{t}) = \sum_{i=1}^{N} \sum_{a=1}^{M} S_{ia} M_{ia}(\xi_i, t_a) \tag{4.1}$$

minimal werden lassen. Als Zwangsbedingung wird weiterhin die Zuordnung jedes Hits zu genau einem Template gefordert:

$$\sum_{a=0}^{M} S_{ia} = 1 \quad \forall\, i \quad . \tag{4.2}$$

Diese Zuordnung heißt *Potts-Bedingung* [9]. Sie fließt in die Verteilungsfunktion bei der Mean-Field-Technik ein. Wir ersetzen dazu $\mathbf{S}$ durch sein thermisches Mittel:

$$V_{ia} = \langle S_{ia} \rangle_T \quad . \tag{4.3}$$

[1] Ohlsson und Peterson benutzen die Bezeichnung „Deformable Templates“ und „Elastic Arms“ gleichwertig. „Deformierbar“ ist ein Template insbesondere dann, wenn seine Form von seinen Parametern abhängt, z. B. vom Impuls bei Anwesenheit eines Magnetfeldes. Dies ist jedoch in unserem Beispiel nicht der Fall.

Mit dem lokalen Feld[2]

$$
\begin{aligned}
u_{ia} &= -\frac{\partial E'}{\partial V_{ia}} \\
&= -M_{ia} \quad .
\end{aligned}
\tag{4.4}
$$

ergibt sich der *Potts-Faktor* [9]

$$
\begin{aligned}
V_{ia} &= \frac{e^{-\beta u_{ia}}}{\sum_{b=0}^{M} e^{-\beta u_{ib}}} \\
&= \frac{e^{-\beta M_{ia}}}{\sum_{b=0}^{M} e^{-\beta M_{ib}}} \\
&\overset{\text{def}}{=} \frac{e^{-\beta M_{ia}}}{e^{-\beta\lambda} + \sum_{b=1}^{M} e^{-\beta M_{ib}}} \quad .
\end{aligned}
\tag{4.5}
$$

Er gibt den Grad der Zuordnung des Hits i zum Template a an und es gilt $\sum_a V_{ia} = 1$ für alle i. Der Parameter λ erlaubt nicht zuzuordnende Detektorhits, zum Beispiel Rauschsignale.

Das Energieminimum sucht man über ein Gradientenabstiegsverfahren mit einem Parameterschritt von

$$
\begin{aligned}
\Delta t_a^{(k)} &= -\eta_a^{(k)} \frac{\partial E'}{\partial t_a^{(k)}} \bigg|_{\sum_b S_{ib} = 1} \\
&= -\eta_a^{(k)} \sum_i S_{ia} \frac{\partial M_{ia}}{\partial t_a^{(k)}} \bigg|_{\sum_b S_{ib} = 1} \\
&\approx -\eta_a^{(k)} \sum_i V_{ia} \frac{\partial M_{ia}}{\partial t_a^{(k)}} \quad .
\end{aligned}
\tag{4.6}
$$

Der Parameter $\eta_a^{(k)}$ ist die Schrittweite des Parameters k des a-ten Templates. Die Schrittweite wird zumeist für alle Templates gleich sein, so daß im folgenden nur noch der Ausdruck $\eta^{(k)}$ benutzt wird. Mit Gleichung 4.6 und 4.5 kann man über Iterieren eine globale Lösung des Spurerkennungsproblems finden, sofern der Startzustand $\mathbf{t}_{\text{start}}$ nicht zu weit vom gesuchten Energieminimum entfernt ist. Die Energiefunktion (4.1) wird mit den Potts-Faktoren zu

$$
E(\mathbf{S}, \xi, \mathbf{t}) = \sum_{i=1}^{N} \sum_{a=1}^{M} V_{ia} M_{ia}(\xi_i, t_a) \quad .
\tag{4.7}
$$

Die Temperatur $T = 1/\beta$ bestimmt die Größe des Einflußbereiches der Hits. Ihre Einheit ist das Quadrat einer Länge. Bei $T \to 0$ wird jeder Hit i nur dem nächstliegenden Template eindeutig zugeordnet (kleinstes M_{ia}), bei höherer Temperatur werden die Templates mit Abständen bis etwa $\sqrt{T}$ anteilig zugeordnet. Der Faktor $e^{-\beta M_{ia}}$ fällt bei $T = M_{ia}$ auf $1/e$ ab.

[2]Das lokale Feld ist in [18] etwas anders definiert, an dieser Stelle soll aber nicht von der in Abschnitt 2.6.2 gegebenen Definition abgewichen werden.

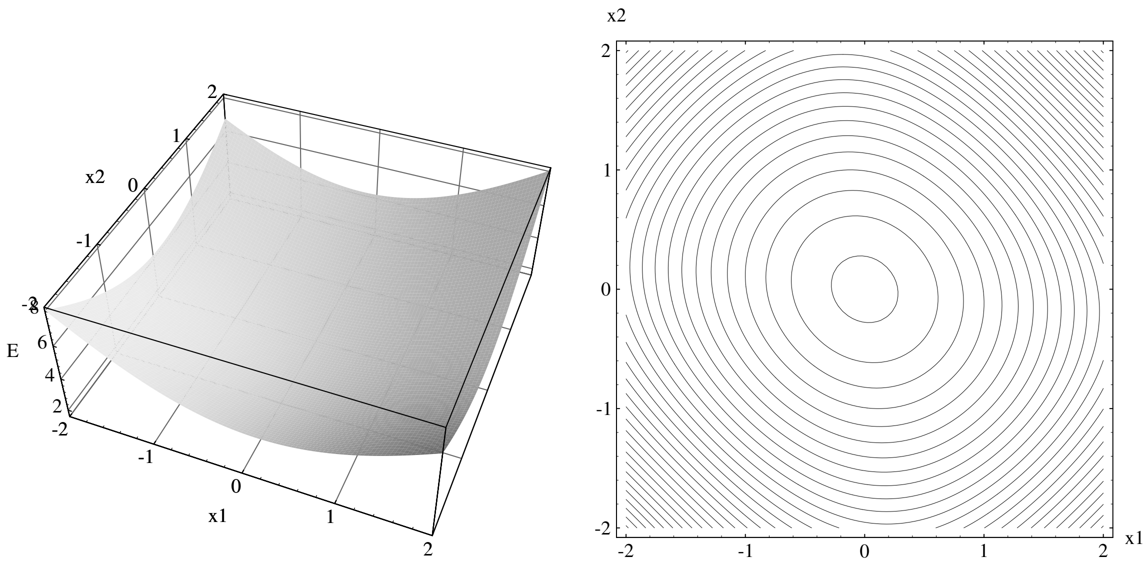


Abbildung 4.1: Räumliche Darstellung und Höhenliniendiagramm der Energiefunktion des eindimensionalen Detektors in Abhängigkeit von den Positionen der Templates x_1 und x_2 bei hoher Temperatur. Die Hits liegen bei $x_a = -1$ und $x_b = 1$, das Energieminimum bei $(x_1 = 0, x_2 = 0)$.

4.2 Der Simulated-Annealing-Prozeß

4.2.1 Untersuchungen am eindimensionalen Detektor

Der Einfluß des Temperaturparameters und des Parameters λ auf die Energiefunktion soll anhand eines extrem einfachen Beispiels erläutert werden. Gegeben sei ein eindimensionaler Detektor, die Lage zweier Templates werde durch die skalaren Parameter x_1 und x_2 beschrieben. Die Energiefunktion (4.7) ist so für zwei Spuren räumlich darstellbar, bei zwei Hits x_a und x_b lautet sie für den Fall $\lambda \to \infty$

$$\begin{aligned} E(x_1, x_2) &= \frac{e^{-\beta(x_1-x_a)^2}(x_1-x_a)^2 + e^{-\beta(x_1-x_b)^2}(x_1-x_b)^2}{e^{-\beta(x_1-x_a)^2} + e^{-\beta(x_1-x_b)^2}} \\ &\quad + \frac{e^{-\beta(x_2-x_a)^2}(x_2-x_a)^2 + e^{-\beta(x_2-x_b)^2}(x_2-x_b)^2}{e^{-\beta(x_2-x_a)^2} + e^{-\beta(x_2-x_b)^2}} \quad . \end{aligned} \tag{4.8}$$

Da sich die Energiefunktion bei Vertauschung der Templates nicht ändert, ist die Energiefunktion symmetrisch zur Geraden $x_1 = x_2$. Wenn im folgenden ein bestimmter Punkt im Phasenraum beschrieben wird, gilt für den gegenüberliegenden Punkt dasselbe, auch wenn er nicht explizit erwähnt wird.

Zunächst gebe es zwei Hits bei $x_a = -1$ und $x_b = 1$. Es werde kein Rauschen erwartet, das bedeutet $\lambda \to \infty$. Die Energiefunktion bei hoher Temperatur $\beta = 0.1$ ist in Abbildung 4.1 dargestellt. Es gibt ein Minimum, welches im Schwerpunkt der Hits liegt, in diesem Fall bei null.

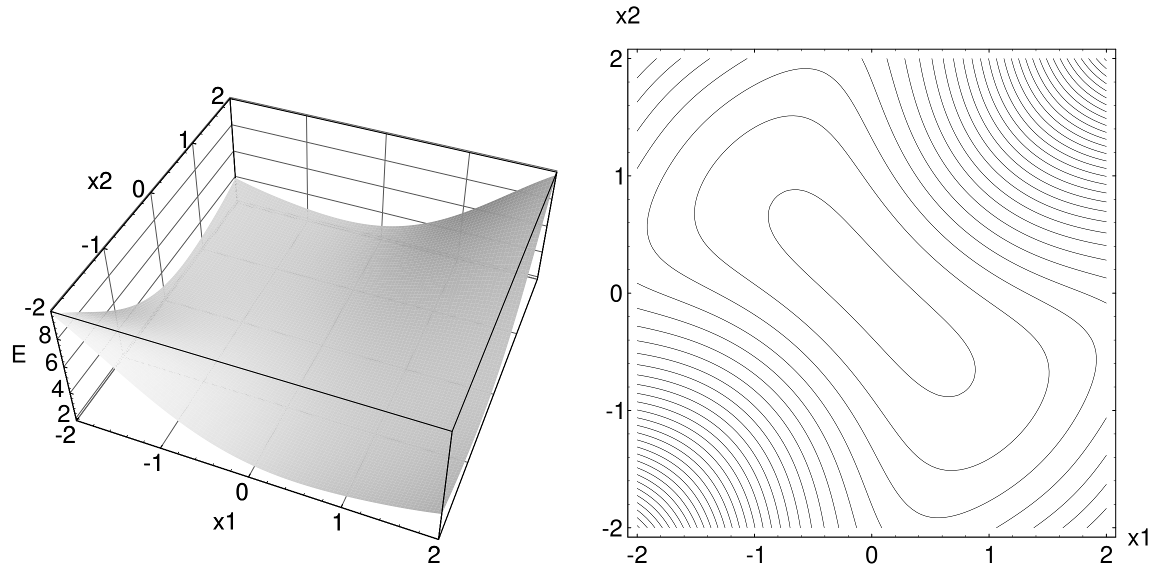


Abbildung 4.2: Energiefunktion des eindimensionalen Detektors bei der kritischen Temperatur.

Beim Absenken der Temperatur wird das Minimum zu einem Sattelpunkt (Abbildung 4.2), und es entstehen zwei neue Minima, die mit sinkender Temperatur zu den endgültigen und gleichwertigen Lösungen $(x_1 = x_a, x_2 = x_b)$ und $(x_1 = x_b, x_2 = x_a)$ konvergieren (Abbildung 4.3). Diesen Vorgang kann man mit einem Phasenübergang identifizieren [9, 16], wobei die Phasenübergangstemperatur vom Abstand der Spuren abhängt. Im eindimensionalen Fall läßt sich die Übergangstemperatur ausrechnen, indem man die Temperatur findet, die die Sattelpunktsbedingung der Energiefunktion (4.8) erfüllt:

$$\frac{\partial^2 E(0,0)}{\partial (x_2 - x_1)^2} = 0 \quad . \tag{4.9}$$

Die kritische Temperatur ergibt sich daraus zu

$$T_c = \frac{1}{\beta_c} = \left(\frac{x_a - x_b}{2}\right)^2 \quad . \tag{4.10}$$

Es fällt außerdem auf, daß die Energie für $x_1 = x_2$ im ganzen x-Bereich höher ist als in der jeweiligen Umgebung mit $x_1 \neq x_2$. Das bedeutet eine *effektive abstoßende Kraft* zwischen den Templates. Die Herleitung dieser Kraft findet sich in [16].

Als nächstes wird ein dritter Hit bei $x_c = 0.236$ eingeführt.[3] Die zwei Templates sollen nun je einem Hit zugeordnet werden, so daß genau ein Hit als Rausch-Hit behandelt wird. Um welchen der drei es sich dabei handelt, ist zunächst nicht von Belang.

Die Energiefunktion unterscheidet sich bei hohen Temperaturen kaum von der in Abbildung 4.1 dargestellten. Unterhalb der kritischen Temperatur bilden sich dagegen

[3] Der Wert $x_c = x_a + (x_b - x_a) \cdot 2/(1 + \sqrt{5})$ ist willkürlich gewählt, eine symmetrische Hit-Konfiguration und ganzzahlige Abstandsverhältnisse sollten vermieden werden.

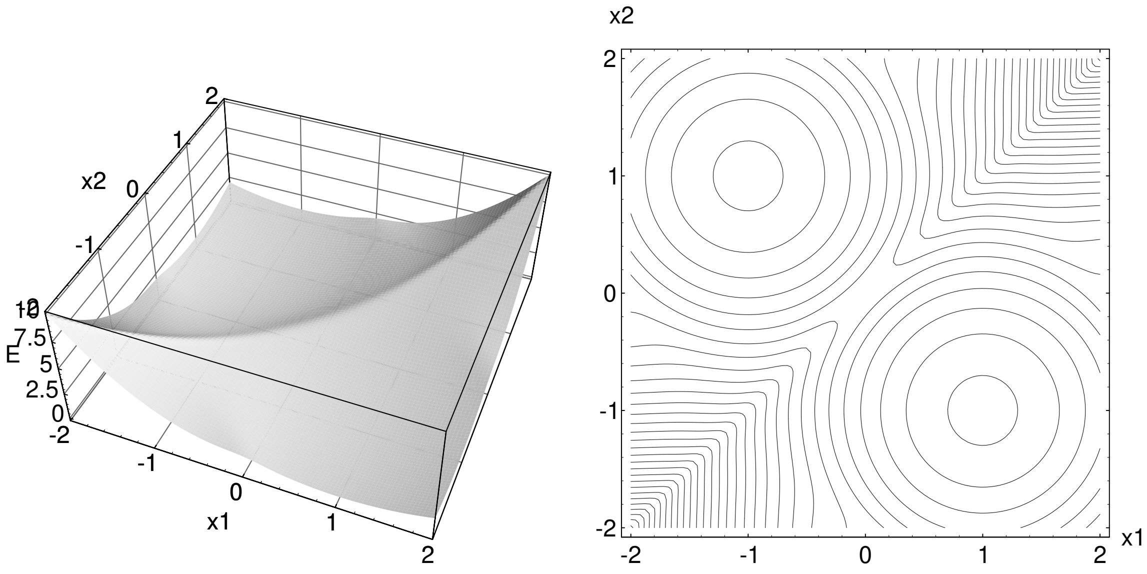


Abbildung 4.3: Energiefunktion bei niedriger Temperatur

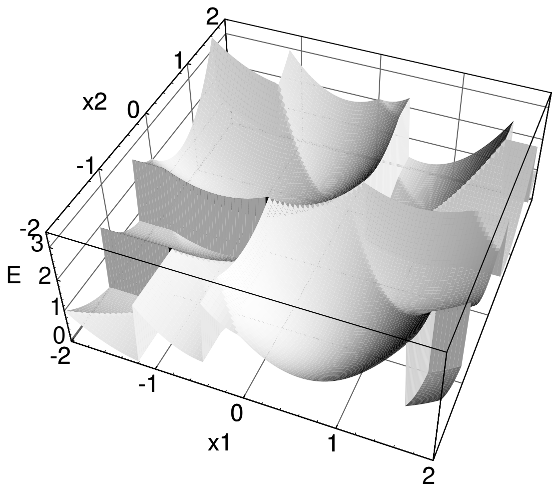
verschiedene Minima aus. Abbildung 4.4 zeigt die Energiefunktion bei sehr niedriger Temperatur $\beta \to \infty$ und $\lambda = 2$. Die beiden Minima aus Abbildung 4.3 sind zu vier Minima aufgespalten, bei denen jeweils ein Template zentrisch zwischen zwei Hits liegt. Die niedrigere Energie liegt vor, wenn ein Template zwischen x_b und x_c liegt.

Der Rauschparameter macht sich bei $\beta \to \infty$ in Form unstetig angeschlossener neuer lokaler Minima bemerkbar. Diese Stufen treten an den Stellen auf, wo der Abstand eines Templates von einem Hit größer als $\sqrt{\lambda}$ wird. Die Stufe bei $x_1 = -0.41$ und $x_2 < -0.41$ bedeutet die Zuordnung von Hit $x_b = 1$ als Rausch-Hit. Das Energieminimum liegt dort bei $(x_1 = -0.41, x_2 = x_a = -1)$, das heißt, nur ein Template wird richtig zugewiesen. Der Parameter λ gibt damit das maximale Entfernungsquadrat von zuzuordnenden Hits an, was dazu führt, daß Abbildung 4.4 kein Energieminimum bei der Zuordnung der beiden Templates zu je genau einem Hit zeigt. Um alle nicht in das Spurmodell passenden Hits auszulassen, sollte λ also das Quadrat eines kleinen Vielfachen der Hit-Auflösung sein.

Wenn im eindimensionalen Detektor die Hits bei $x > 0$ getrennt werden sollen, muß die Auflösung deutlich besser als $(x_b - x_c)$ sein. Wir nehmen willkürlich eine Auflösung von $\sigma = 0.12$ an und setzen $\lambda = (5\sigma)^2$.

Die Energiefunktion am Ende des Annealing-Vorganges mit $\lambda = 0.4$ ist in Abbildung 4.5 gezeigt. Bei allen sechs möglichen Lösungen, von denen jeweils zwei gleich sind, ist ein lokales Energieminimum vorhanden. Ein weiteres besteht bei $x1 = (x_c + x_b)/2$. Außerdem gibt es niedrige Energien an den Rändern des betrachteten Bereiches, wie schon bei der $\lambda = 2$-Konfiguration beschrieben. Es ist deutlich erkennbar, daß bei $x_1 = -1$ und $x_2 < -0.4$ die Lage von x_2 gleichgültig wird, die Energiefunktion bildet dort eine „Rinne".

Diese Untersuchung am eindimensionalen Detektor zeigt, daß die Energiefunktion



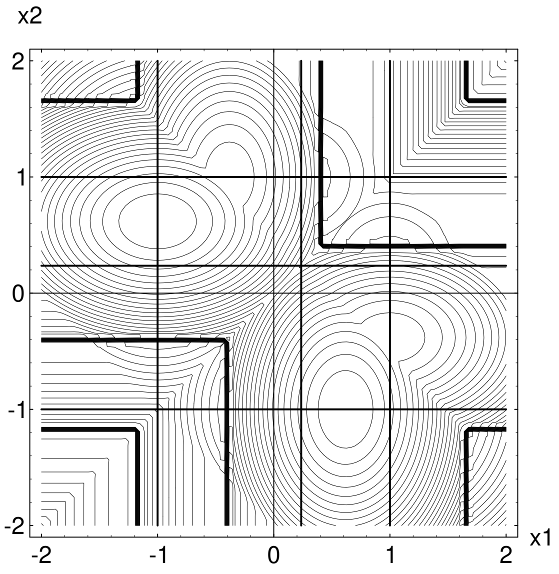


Abbildung 4.4: Energiefunktion bei drei Hits (niedrige Temperatur) und $\lambda = 2$. Die Hits sind im Höhenliniendiagramm durch Geraden gekennzeichnet.

schon bei einfachen Problemen eine sehr komplexe Gestalt annimmt und man schon hier den Annealing-Prozeß langsam ausführen muß, um nicht das zu verfolgende Energieminimum zu verlieren.

4.2.2 Die Entwicklung der Potts-Faktoren

Es ist interessant, die Entwicklung der Potts-Faktoren V_{ia} während des Annealing-Prozesses zu beobachten. Dazu wird wieder der Modelldetektor aus 3.4 verwendet. Die Parameterdarstellung der Spuren ist die zweidimensionale Projektion von Gleichung 1.6 auf die xz-Ebene, und der Parameterraum wird durch den Richtungstangens $\tan\theta_x$ und den Achsenabschnitt x_0 in der Referenzebene bei $z = z_{\text{ref}}$ aufgespannt. Damit die Korrelation der Parameter verschwindet, wird die Referenzebene in der Mitte des Detektors gewählt. Um die Rechnungen zu verkürzen, wird für den Fall des Modelldetektors der Ursprung des Koordinatensystems in die Referenzebene verschoben, das heißt $z_{\text{ref}} = 0$.

Die Summe der Potts-Faktoren eines Templates gibt die Anzahl der zugeordneten Hits an, wobei es auch anteilige Zuordnungen gibt. Das bedeutet, für $T \longrightarrow \infty$ ist $\sum_i V_{ia}$ für alle a gleich N/M. Bei endlicher Temperatur werden nur Hits zwischen eng benachbart liegenden Templates geteilt. Diese Aufteilung, bei denen zum Beispiel alle nicht verschwindenden Potts-Faktoren zweier Templates gleich 0.5 sind, lassen die Terme $\sum_i V_{ia}$ für beide Templates gleich werden. Die Hits werden erst bei einer bestimmten niedrigeren Temperatur eindeutig zugeordnet, dabei laufen im allgemeinen die Größen $\sum_i V_{ia}$ für beide Templates auseinander.

Wir betrachten ein Ereignis mit 16 Spuren, die in zwölf Detektorlagen 185 Hits erzeugt haben. Mit einer einfachen Initialisierung (Abschnitt 4.3) werden 20 Templates

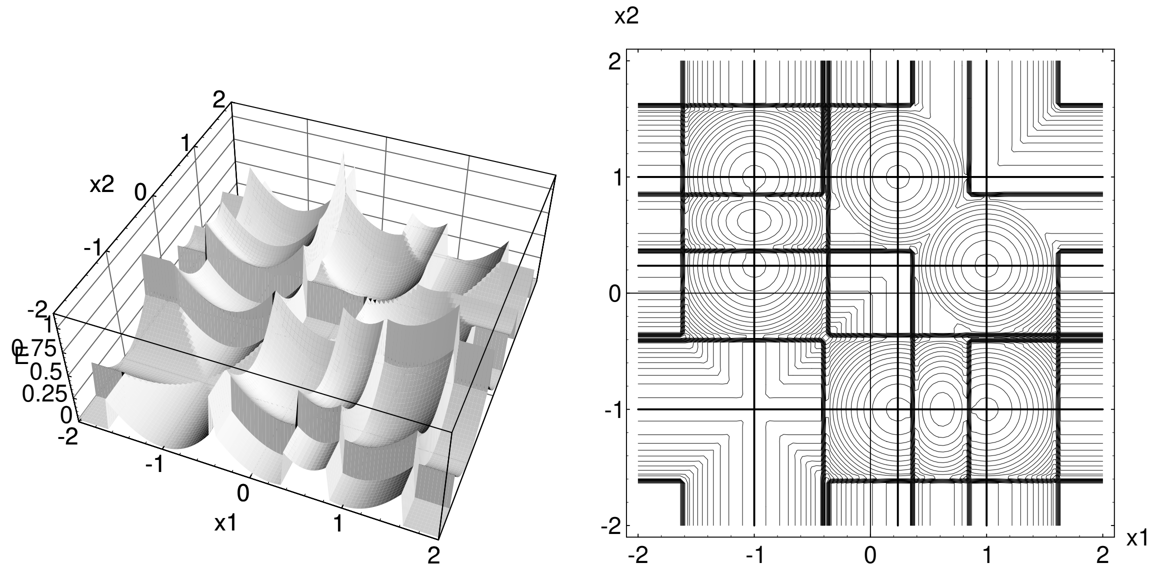


Abbildung 4.5: Wie Abbildung 4.4 mit $\lambda = 0.4$.

erzeugt, die im Parameterraum grob in der Nähe der Spuren liegen. Die Initialisierung ist in Abbildung 4.7 dargestellt.

In Abbildung 4.8 ist der Verlauf von $\sum_i V_{ia}$ über einen Annealing-Prozeß mit 3000 Iterationen und einer Schrittweite von $\eta_{x_0} = \eta_{\tan\theta x} = 0.02$ dargestellt. Dabei wurde β_j nach Gleichung 3.5 berechnet.

Von der Initialisierung bei $\beta = 0.1\,\mathrm{m}^{-2} = 1/T \approx (3.1\,\mathrm{m})^{-2}$ laufen alle Templates in den Schwerpunkt, und es gilt $\sum_i V_{ia} \approx N/M = 185/20 \approx 9.2$ für alle a. Die der Hitauflösung von $200\,\mu$m entsprechende Temperatur wird bei $\beta = 2.5 \cdot 10^7\,\mathrm{m}$ erreicht. Hier ist die bestmögliche Zuordnung der Hits erreicht. Bei höheren β gibt es weitere Phasenübergänge, die von Templates herrühren, welche sich eine Spur geteilt haben. Diese Auftrennungen führen dazu, daß kein Template die geforderte Anzahl von Zuordnungen behält und die Spur nicht rekonstruiert werden kann.

Die Stufen ab $\beta = 10^7$ resultieren aus numerischen Effekten. Bei der Berechnung von Potts-Faktoren von Hits, deren quadratischer Abstand zum nächstliegenden Template viel größer als $T = 1/\beta$ ist, wird im Rahmen der numerischen Auflösung sowohl der Zähler als auch der Nenner null. Die Potts-Faktoren können daher nicht mehr berechnet werden und werden willkürlich auf 0 gesetzt. In dem Fall $\beta_{\text{Start}} = 0.1\,\mathrm{m}^{-2}$ werden wegen des Zusammenlaufens der Spuren am Anfang ohnehin die äußeren Spuren verloren, was man nur durch einen viel langsameren Annealing-Prozeß wieder ausgleichen könnte.

Ein Annealing-Prozeß mit der gleichen Anzahl von Iterationen, aber einer größeren Zahl gefundener Spuren ist in Abbildung 4.9 gezeigt. Dort ist der Startwert $\beta_{\text{start}} = 800\,\mathrm{m}^{-2}$. Das Diagramm zeigt, daß mindestens zwölf Templates bei $\beta_{\text{end}} = 10^7\mathrm{m}^{-2}$ zwölf Hits bekommen haben, mindestens zwei haben sechs Hits, was bedeutet, daß jeweils zwei Templates auf dieselbe Spur konvergiert sind. In der Auswertung kann man diese Templates als übereinanderliegend erkennen, zu einem Template vereinigen und so die

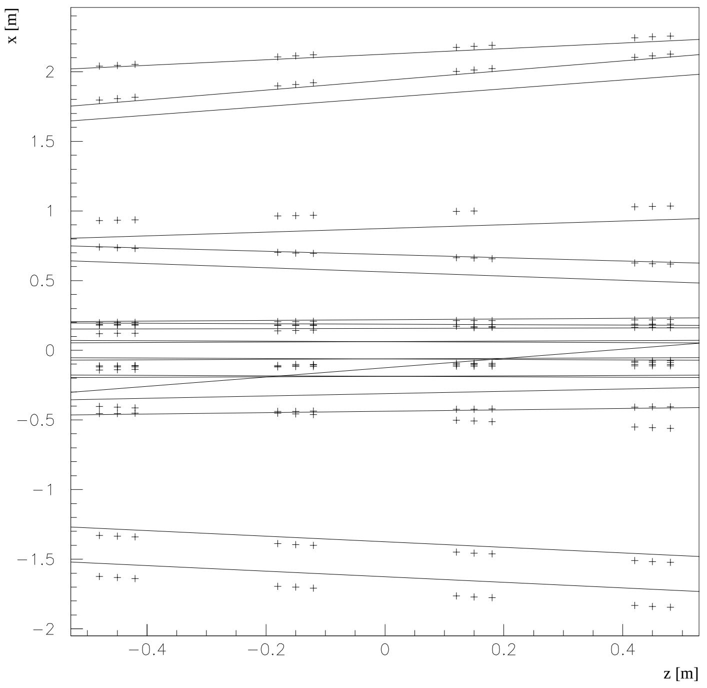


Abbildung 4.6: Initialisierung eines 16 Spur-Ereignisses mit 20 Templates im Modelldetektor

Spur richtig rekonstruieren.

4.3 Initialisierung über eine Hough-Transformation

Der Elastic Arms Algorithmus benötigt eine Initialisierung. Um die Konvergenz schneller zu erreichen, sollte die Anfangsverteilung der Templates bereits grob der tatsächlichen Spurverteilung im Parameterraum entsprechen, wobei überschüssige Templates in bestimmten Phasenraumregionen weniger schaden als fehlende.

4.3.1 Die lokale Hough-Transformation

Die Spurdichte im Parameterraum kann man mit einer *lokalen Hough-Transformation* abschätzen [16]. Für den zweidimensionalen Modelldektor berechnet man zu Paaren von Hits die Parameter $(x, \tan\theta_x)$ der Verbindungssegmente und trägt die Werte in ein zweidimensionales Histogramm ein. Das Histogramm für das Ereignis aus Abbildung 4.6 ist in Abbildung 4.10 dargestellt. Dabei wurden alle Verbindungen von Hits innerhalb jeweils einer Superlage eingetragen.

Es gibt verschiedenen Möglichkeiten, die Hitpaare auszuwählen, zum Beispiel über Nachbarschaftsbeziehungen der Detektorlagen oder über den maximalen Abstand der Hits. Das Histogramm deckt dabei nur den Parameterbereich ab, in dem man nach Spuren sucht. Man wird in den meisten Fällen sowohl „richtige“ als auch „falsche“

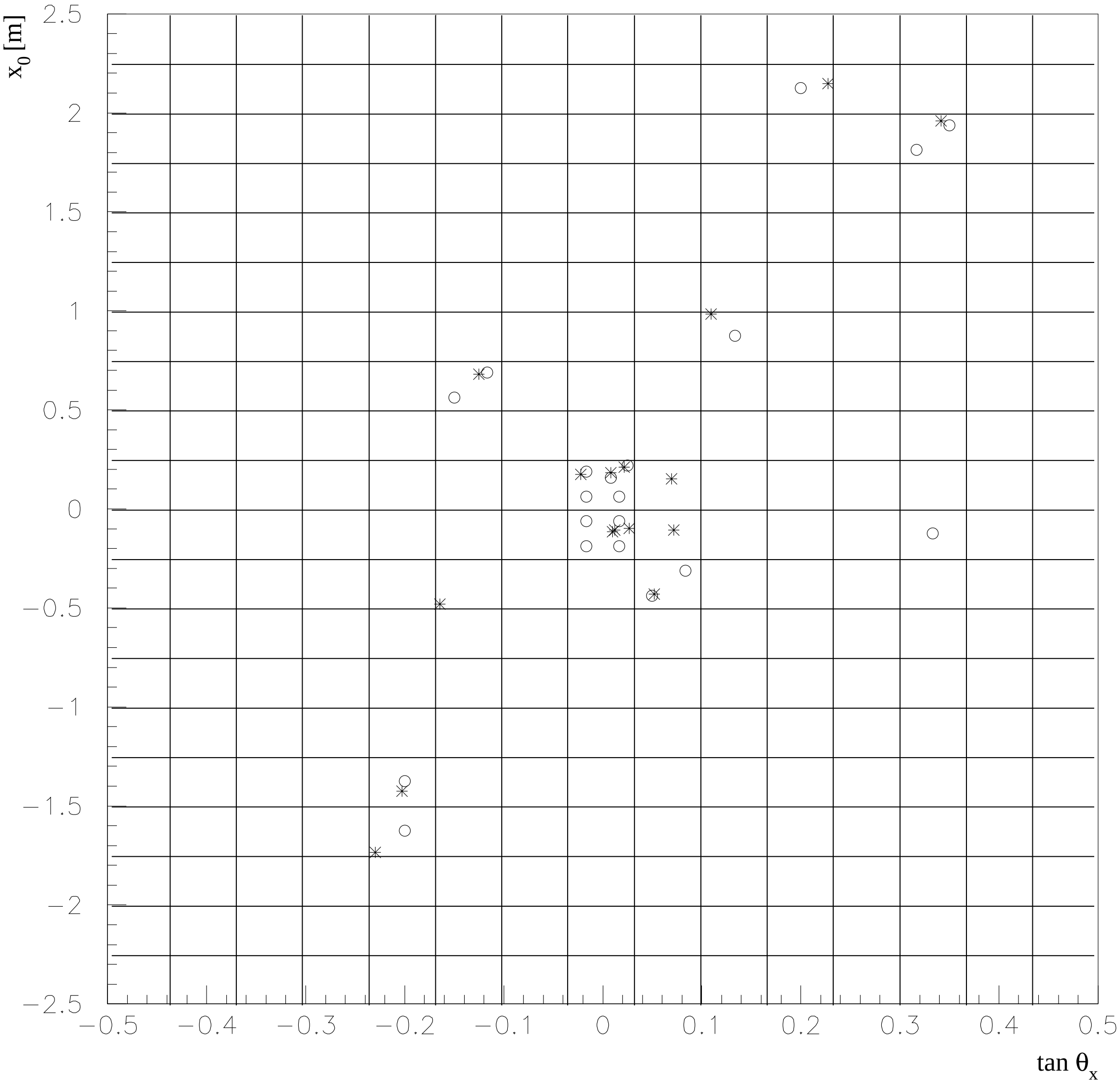


Abbildung 4.7: Modelldetektor: Ereignis aus Abbildung 4.6 im Parameter-Raum

Monte-Carlo-Spuren: $*$
Template-Initialisierungen: $\circ$

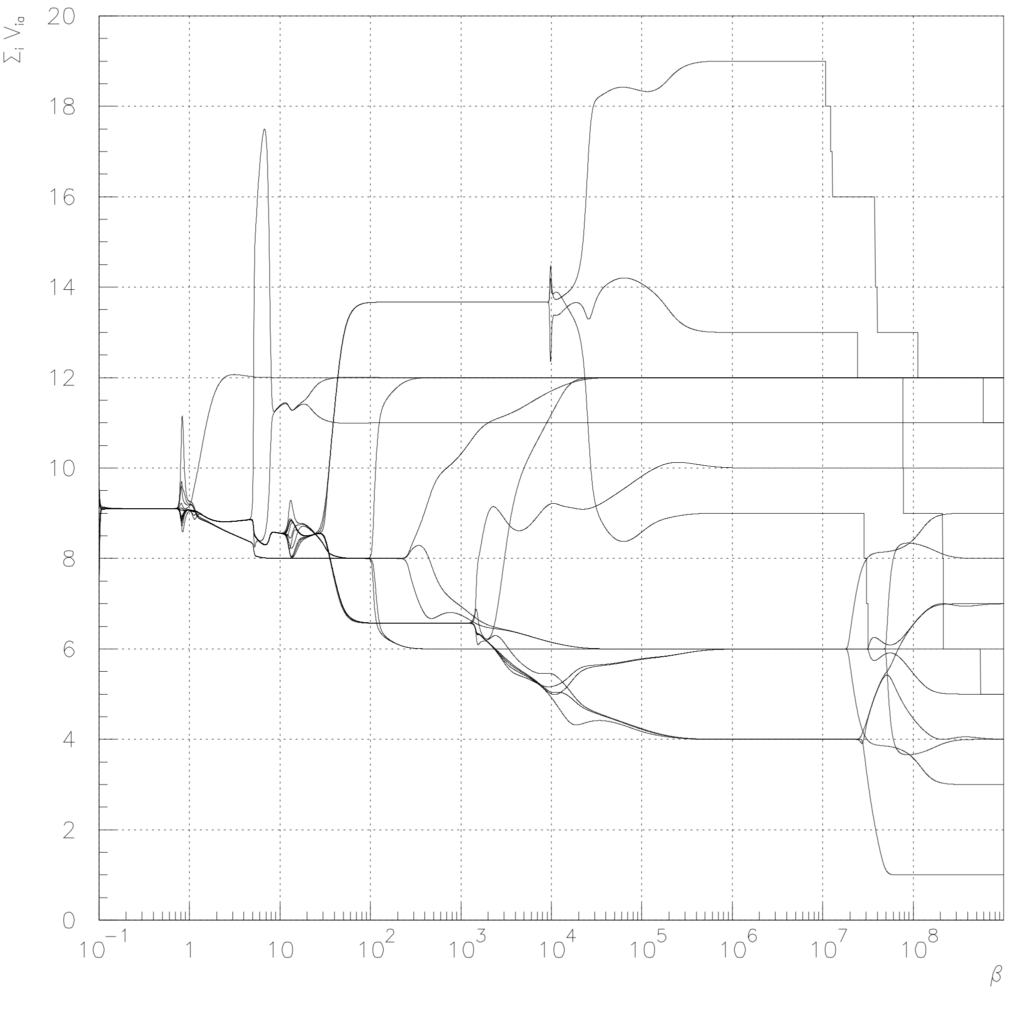


Abbildung 4.8: Entwicklung der $\sum_i V_{ia}$, Start bei sehr hoher Temperatur

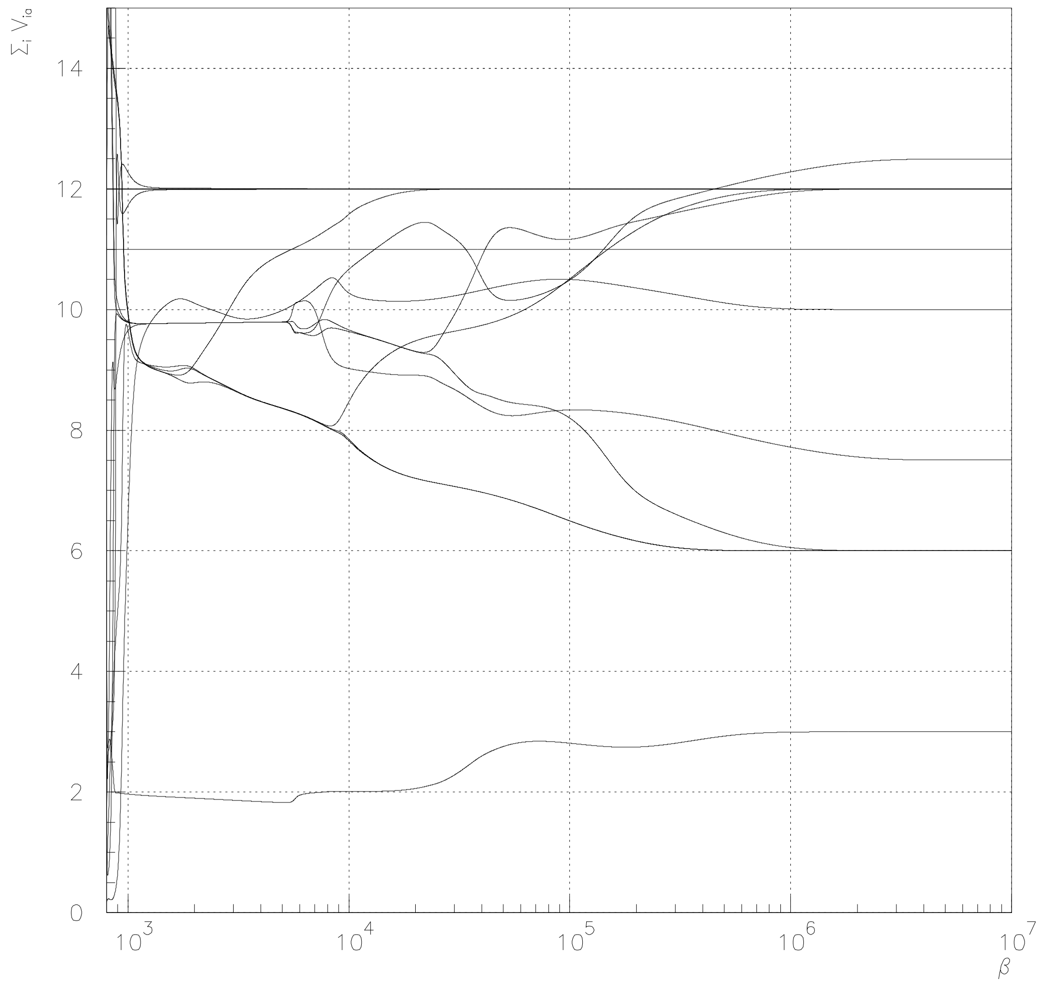


Abbildung 4.9: Entwicklung der $\sum_i V_{ia}$, Start bei einer der den typischen Abständen der Template-Initialisierungen von den Hits entsprechenden Temperatur

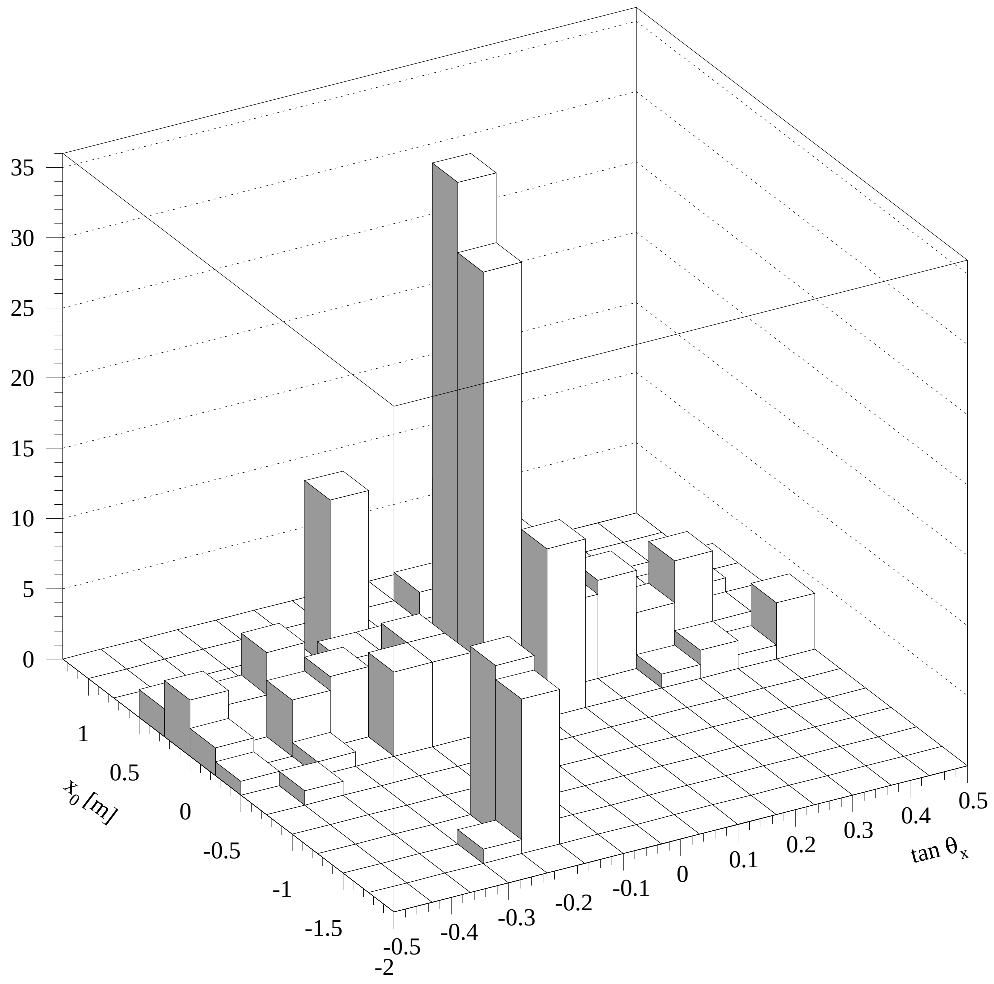


Abbildung 4.10: Hough-Transformation des Ereignisses aus Abbildung 4.6

Einträge haben, das heißt sowohl aus Hits einer Spur als auch aus Hits von zwei verschiedenen Spuren bestehende Segmente. Die Auswertung des Histogramms ist daher nicht einfach. T. Schober hat dieses Thema ausführlich in [17] behandelt.

4.3.2 Auswertung der Transformation

Aus den Einträgen schätzt man die Spurdichte im Parameterraum ab. Im Fall des Modelldetektors reicht es aus, eine typische Anzahl von Einträgen einer Spur anzunehmen. Die Zahl hängt von der maximalen Länge der betrachteten Verbindungssegmente, von der Detektorgeometrie und vom Verhältnis der Intervallgröße des Histogramms zur Detektorauflösung ab. Kanäle mit weniger Einträgen werden nicht berücksichtigt. Auf diese Weise wird ein großer Teil der nicht zu Spuren gehörenden Verbindungen übergangen. In der Initialisierung aus Abbildung 4.7 wurde beispielsweise ein Template für jeweils 5 Einträge generiert.

Die Einträge einer Spur streuen zumeist in mehrere benachbarte Kanäle des Histogramms, wobei der zentrale Kanal oberhalb der Schwelle, und die umliegenden Kanäle unterhalb liegen. In den dicht belegten Bereichen des Parameterraums summieren sich diese Randeinträge verschiedener Spuren zu Werten, welche oberhalb der Schwelle liegen. Deshalb werden mit diesem Verfahren im Bereich hoher Spurdichten zuviele Templates erzeugt. Dieser Effekt spielt im Modelldetektor nur eine untergeordnete Rolle, da die überzähligen Templates nach dem in Abschnitt 4.5 beschriebenen Verfahren gelöscht werden.

4.4 Initialisierung mittels direkter Suche nach Spursegmenten in den Superlagen

Die Superlagenstruktur der HERA-B-Spurkammern erlaubt eine lokale Suche nach kurzen Spursegmenten. Bei der Kombination der Spursegmente der einzelnen Superlagen kann man eine Auswahl möglicher Spuren bekommen. Deren Benutzung als Ausgangspunkt für den Elastic-Arms-Algorithmus wird im folgenden als *Segment-Initialisierung* bezeichnet.

4.4.1 Bestimmung der Anfangssegmente

Da die HERA-B-Spurkammern nur in der xz-Projektion für diesen Zweck ausreichend viele Detektorlagen haben, wurde die Segment-Initialisierung nur in der Projektion implementiert. Ein Segment wird wie die zweidimensionale Spur aus Abschnitt 4.2.2 durch die Parameter $(\tan\theta_x, x_0)$ dargestellt.

Drei Hits (z_1, x_1), (z_2, x_2) und (z_3, x_3) können beispielsweise durch eine Ausgleichsrechnung zu einem Segment $(\tan\theta_x, x_0)$ zusammengefaßt werden. Ein solches wird im folgenden *Triplett* genannt. Die Größe

$$\chi^2 = \sum_{i=1}^{3} \left(\frac{x_0 + (z_i - z_{\mathrm{ref}}) \tan\theta_x - x_i}{\sigma_{x_i}} \right)^2 \qquad (4.11)$$

folgt der χ^2-Verteilung bei einem Freiheitsgrad. Im vorliegenden Beispiel sollten 95% der richtigen Tripletts gesucht werden. Dazu muß man alle Tripletts mit $\chi^2 < 3.8$ auswählen.

Damit die spätere Kombination von Segmenten verschiedener Superlagen möglichst eindeutig zu erreichen ist, sollte im Idealfall genau ein Segment für jede Spur in jeder Superlage erzeugt werden. Da eine Spur aber bis zu acht Hits[4] in einer Superlage erzeugen kann, sollten die daraus erzeugten bis zu über 50 Tripletts[5] zusammengefaßt werden. Das geschieht durch den Vergleich von $\tan\theta_x$ und $x_{\bar{z}}$ an der Stelle des mittleren z der zu vergleichenden Segmente. Im vorliegenden Beispiel wurden Segmente bei Abweichungen von $\Delta\tan\theta_x \leq 0.02$ und $\Delta x_{\bar{z}} \leq 0.02\,\mathrm{cm}$ zusammengefaßt.

Ein Beispiel von auf diese Weise erzeugten Anfangssegmenten ist in Abbildung 4.11 gezeigt. Dort gehören die linke und die rechte Bildhälfte zu unterschiedlichen Modulen, die sich in y-Richtung nicht überschneiden. Es ist unwahrscheinlich, daß eine Spur in beiden Hälften Hits erzeugt, deshalb werden keine Segmente gebildet, die beide Module verbinden.

Es zeigt sich, daß die Anzahl der erzeugten Segmente, die nicht zu Spuren gehören, meistens eine Größenordnung über der Zahl der korrekten Segmente liegt.

4.4.2 Kombination der Segmente

Die Kovarianzmatrix eines Segments, dessen Zentrum in der Referenzebene liegt, ist

$$\mathbf{C} = \begin{pmatrix} \sigma^2_{\tan\theta_x} & 0 \\ 0 & \sigma^2_{x_0} \end{pmatrix} \quad .$$

Die Verschiebung der Referenzebene um Δz ergibt die folgende Kovarianzmatrix:

$$\mathbf{C}' = \mathbf{TCT}^{\mathsf{T}} \quad \text{mit} \quad \mathbf{T} = \begin{pmatrix} 1 & 0 \\ \Delta z & 1 \end{pmatrix} \quad .$$

Zwei beliebige Segmente $\mathbf{s}_1$ und $\mathbf{s}_2$ mit den Kovarianzmatrizen $\mathbf{C}_1$ und $\mathbf{C}_2$ können wie folgt zu einem Segment $\mathbf{s}$ mit der Kovarianzmatrix $\mathbf{C}$ kombiniert werden:

$$\mathbf{s} = \begin{pmatrix} \tan\theta_x \\ x_0 \end{pmatrix} = \underbrace{\left(\mathbf{C}_1^{-1} + \mathbf{C}_2^{-1}\right)^{-1}}_{\mathbf{C}} \left(\mathbf{C}_1^{-1}\mathbf{t}_1 + \mathbf{C}_2^{-1}\mathbf{t}_2\right) \quad . \qquad (4.12)$$

Dabei gilt

$$chi^2 = (\mathbf{s} - \mathbf{s}_1)^{\mathsf{T}} \mathbf{C}_1^{-1} (\mathbf{s} - \mathbf{s}_1) + (\mathbf{s} - \mathbf{s}_2)^{\mathsf{T}} \mathbf{C}_2^{-1} (\mathbf{s} - \mathbf{s}_2) \quad . \qquad (4.13)$$

[4] Da sich die Zellen einer einzelnen Lage gegenseitig überlappen, kann eine Spur dort jeweils zwei Hits erzeugen. Bei vier Lagen innerhalb einer Superlage ergibt sich eine maximale Anzahl von acht Hits.

[5] Die Zahl der Tripletts ist bei n Hits $\binom{n}{3}$.

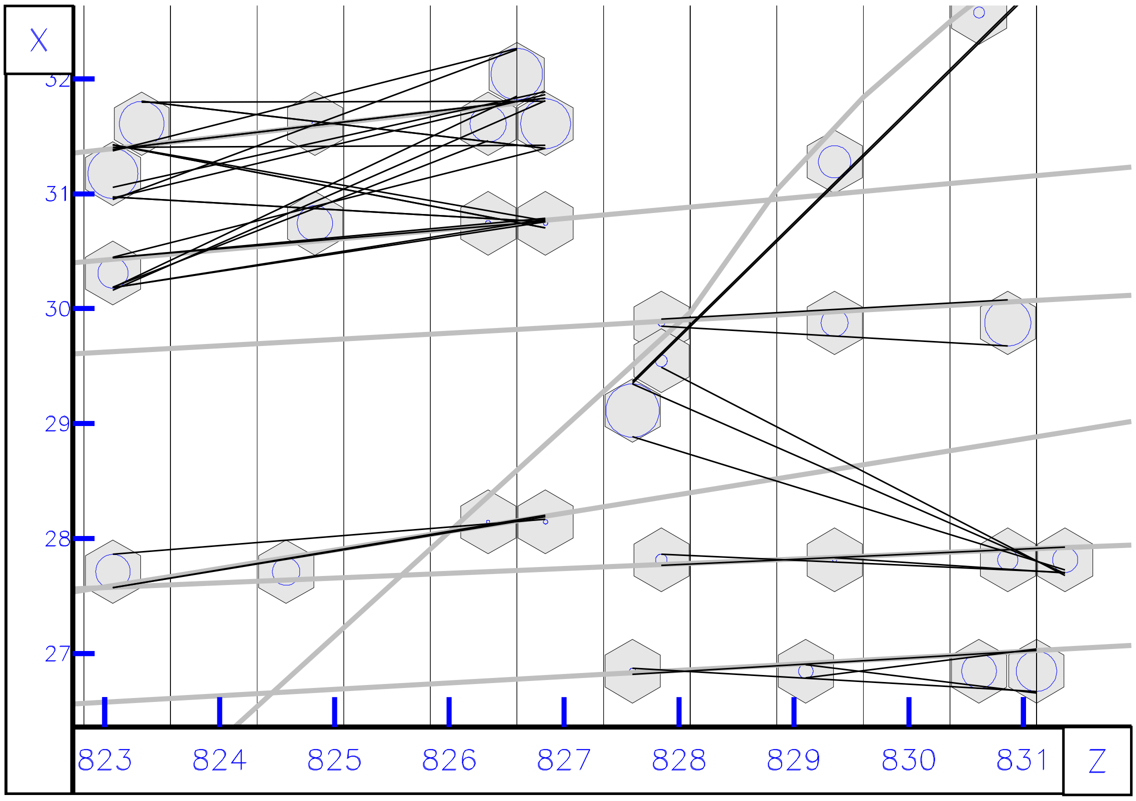


Abbildung 4.11: Mögliche Spursegmente innerhalb einer Superlage. Die Monte-Carlo-Spuren sind grau dargestellt.

Das χ^2 (4.13) folgt einer χ^2-Verteilung für zwei Freiheitsgrade. Wenn ein solches Kombinationssegment aus Segmenten zweier unterschiedlicher Superlagen ein günstiges χ^2 aufweist, ist die Wahrscheinlichkeit hoch, daß es eine Spur repräsentiert. Es werden die günstigsten Segmentkombinationen zwischen verschiedenen Superlagen gesucht. Der Algorithmus dazu läuft wie folgt ab:

Zunächst werden in benachbarten Superlagen alle Segmentkombinationen bis zu einem χ^2_{grenz} gesammelt, wobei jedes Anfangssegment höchstens einmal benutzt werden darf und in Konfliktfällen die Kombination mit dem besten χ^2 gewählt wird. Dann werden zunächst diese Kombinationen aus weiter auseinanderliegenden Superlagen auf dieselbe Weise kombiniert und danach die übrigbleibenden Kombinationen mit den noch nicht benutzten Anfangssegmenten.

Nachdem Anfangssegmente-Kombinationen aller vier Superlagen in einer einzigen Liste von Segmentkombinationen aufgegangen sind, wird diese als Initialisierung des Elastic-Arms-Algorithmus benutzt.

4.5 Erweiterungen

Der Term $\sum_i V_{ia}$ eines Templates a gibt offenbar Aufschluß darüber, ob zuviel oder zuwenig Templates in einem Bereich des Parameterraums liegen. Ist $\sum_i V_{ia}$ deutlich größer als die zu erwartende Anzahl von Hits pro Spur, kann man ein zusätzliches Template a' mit gegenüber a leicht veränderten Parametern generieren; ist der Term deutlich zu klein, kann man das Template a löschen.

Dieses Verfahren wurde zur Messung der Effizienz im zwölflagigen und 24-lagigen Modelldetektor angewendet (Abschnitt 4.6.1).

4.6 Ergebnisse

4.6.1 Messung der Effizienz im Modelldetektor

Zur Messung der Spurfindungseffizienz des Elastic-Arms-Algorithmus wurde zunächst wieder der Modelldetektor aus Abschnitt 3.4 benutzt. Dieser wurde sowohl als 4×3- und als 8×3-Lagen-Detektor untersucht. Mit den Erweiterungen aus Abschnitt 4.5 wurde die Effizienz in Abhängigkeit von der Belegungsdichte gemessen. Die benutzten Parameter sind in Tabelle 4.1 zusammengefaßt. Die Bedingungen zum Löschen und Erzeugen von Templates beziehen sich auf den zwölflagigen Detektor, beim 24-lagigen sind sie um den Faktor 2 größer. Für die lokale Hough-Transformation wurden alle im betrachteten Parameterraum liegenden Verbindungssegmente zwischen Hits jeweils einer Superlage in das Histogramm eingetragen.

Das Ergebnis ist in Diagramm 4.12 dargestellt, wobei für den 4×3-Lagen-Detektor die unteren Werte gelten. Die Effizienz liegt beim Acht-Superlagen-Detektor bei einer maximalen Belegungsdichte ρ von 0.03 bei 97% und sinkt gleichmäßig bis zu $\rho = 0.17$ auf 83%. Der Vier-Superlagen-Detektor liegt 5% bis 10% darunter.

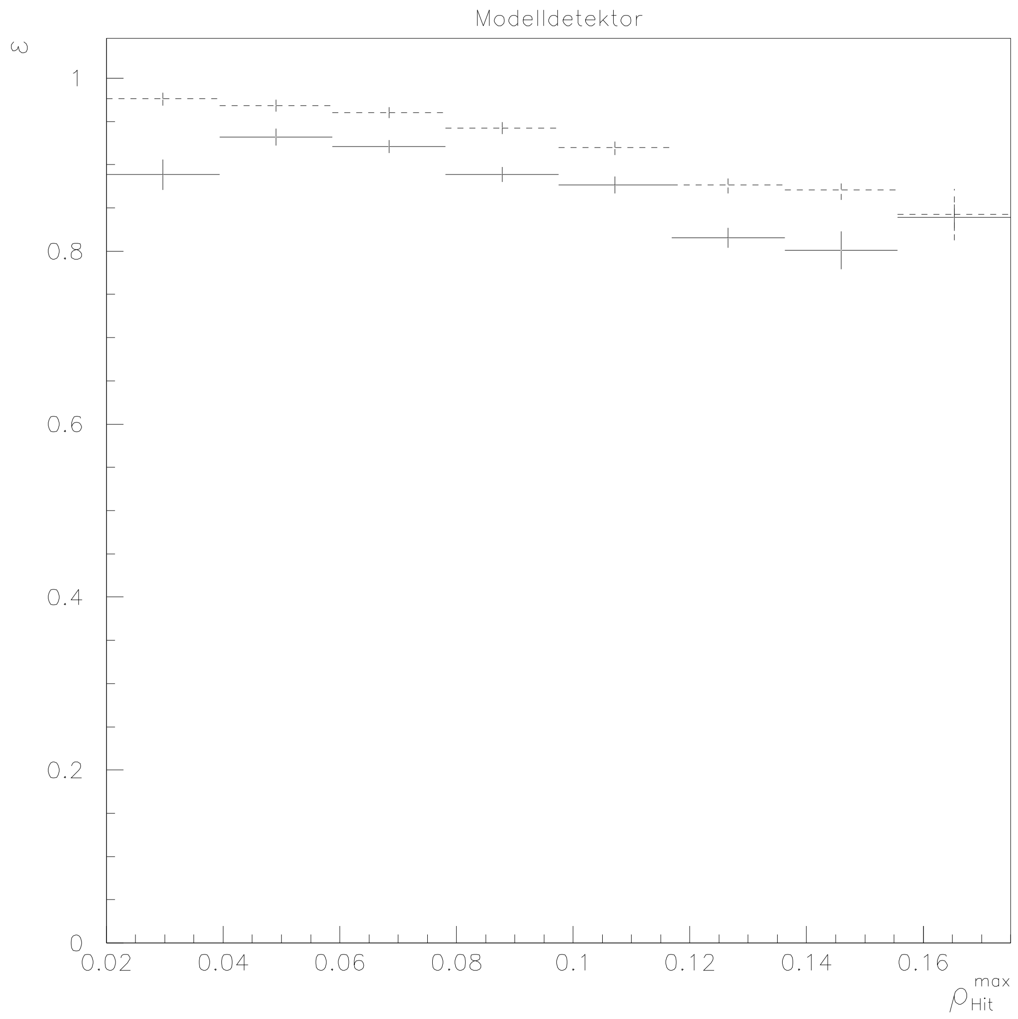


Abbildung 4.12: Spurfindungseffizienz ε im Modelldetektor in Abhängigkeit von der maximalen Belegungsdichte $\rho_{\mathbf{Hit}}^{\mathbf{max}}$. Die unteren Werte gelten für den 4×3-Lagen-Detektor, die oberen für den 8×3-Lagen-Detektor.

Hough-Transformation	
Parameterraumgröße	$\tan\theta_x{}^{(\max)} = 0.5$ $x_0{}^{(\max)} = 2.5$
Anzahl der Histogramm-Intervalle:	15×20
geforderte Einträge pro Template	5

Elastic-Arms	
inverse Starttemperatur	$800\,\mathrm{m}^{-2}$
inverse Endtemperatur	$10^6\,\mathrm{m}^{-2}$
Anzahl Iterationen	3000
λ	$(10^{-3}\,\mathrm{m})^2$
Schrittweite $(\eta^{(\tan\theta_x)}, \eta^{(x_0)})$	(0.02,0.02)
Löschbedingung für a	$\sum_i V_{ia} < 8$
Erzeugungsbedingung für a'	$\sum_i V_{ia} > 18$

Tabelle 4.1: Parameter für den Elastic-Arms-Algorithmus im zwölflagigen Modelldetektor.

Die Effizienz konnte nicht wie beim Denby-Peterson-Algorithmus bis zu einer Maximal-Belegunsdichte von $\rho = 0.3$ gemessen werden, da die Rechenzeit auf demselben Rechner für ein Ereignis schon bei $\rho = 0.17$ bei etwa 3000 s liegt.

4.6.2 Anwendung auf HERA-B

Der Übergang vom Modelldetektor auf HERA-B bringt folgende zusätzliche Anforderungen mit sich:

- Die Ambiguitäten der Driftkammersignale müssen vom Algorithmus aufgelöst werden.

- Die zweidimensionalen punktförmigen Hits werden durch linien- oder zylinderförmige dreidimensionale Hits ersetzt. Damit wird die Abstandsfunktion aufwendiger (Anhang A.2), und die Energiefunktion wird komplexer. Das begrenzt die mögliche Schrittweite, so daß die Konvergenz später erreicht wird.

- Die Modulstruktur des Detektors beeinträchtigt die Vorhersage der Anzahl der Hits einer Spur. Die erwartete Zahl liegt zwischen etwa 18 Hits in den MSGC-Modulen und 28 bis über 40 in den Driftkammern. Spuren in unmittelbarer Nachbarschaft können deutlich verschieden viele Hits haben, da die einzelnen Driftkammerzellen einer Lage sich zum Teil überdecken. Damit ist die Identifizierung überschüssiger Templates oder die Feststellung fehlender Templates nach Abschnitt 4.5 nicht ohne weiteres möglich.

- Die Hough-Transformation liefert kein einfach zu interpretierendes Bild des Ereignisses im Parameterraum (Abschnitt 4.6.3).

- Die komplexen physikalischen Vorgänge bei der Proton-Nukleon-Streuung bewirken eine ungleichmäßigere Verteilung der Spuren im Detektor. Es gibt daher lokal höhere Spurdichten als im Modelldetektor, was vor allem die Auswertung der Hough-Transformierten erschwert.

4.6.3 Wahl der Initialisierung

Die Hough-Transformation

Die lokale Hough-Transformierte der Simulation eines HERA-B-Ereignisses ist in Abbildung 4.13 dargestellt. Dabei wird die Zahl der Einträge der einzelnen Intervalle durch Graustufen angedeutet, wobei die Felder mit wachsender Eintragszahl dunkler werden. Es wurden nur die $0°$-Lagen, also die xz-Projektion, benutzt. Der Parameterraum ist deshalb wie in Abschnitt 4.2.2 zweidimensional. In das Histogramm wurden die Verbindungen aller Hitpaare mit einem maximalen Abstand der z-Koordinaten von 5 cm eingetragen.

Es zeigen sich Effekte, die im Modelldetektor nicht oder nur in geringem Maße auftreten und die Auswertung der Hough-Transformation erschweren. Es gibt eine große Zahl von Segmenten, die nicht Teil von Spuren sind. Sie liegen nicht gleichmäßig im Parameterraum verteilt, sondern bilden streifenförmige Muster, welche sich wie folgt erklären lassen:

Ein einzelner Hit (z, x) mit dem Abstand Δz von der Referenzebene $z_{\mathbf{ref}}$ ist kompatibel mit der Schar von Spuren $(\tan\theta_x, x_0)$, für die

$$\Delta z \tan\theta_x + x_0 = x \tag{4.14}$$

gilt. Diese Gleichung ist für Spursegmente in der Nähe von (z, x) annähernd erfüllt. Beim Eintragen solcher Segmente verschiedener Steigung in ein Histogramm erzeugt man folglich näherungsweise einen Graphen der Geradengleichung 4.14. Dieser Effekt wird in Abbildung 4.14 demonstriert. Dort werden drei Lagen bei $z_i = (0.50, 0.53, 0.56)\,\mathrm{m}$ des Modelldetektors aus Abschnitt 3.4 im Bereich $0.50\,\mathrm{m} < x < 0.52\,\mathrm{m}$ von vier eng beieinanderliegenden Spuren durchquert (linke Abbildung). Diese erzeugen zwölf Hits, wobei ein zusätzlicher Hit vom Detektorrauschen herrührt. Die lokale Hough-Transformation erzeugt das Bild auf der rechten Seite. Mit $\langle\bar{x}\rangle = 0.51\,\mathrm{m}$ und $\langle\bar{z}\rangle = 0.53\,\mathrm{m}$ läßt sich Gleichung 4.14 auflösen zu

$$x_0 = -0.53\,\mathrm{m} \cdot \tan\theta_x + 0.51\mathrm{m} \quad .$$

Die Histogramm-Einträge in Abbildung 4.14 liegen ungefähr auf dieser Geraden.

Bei mehreren dieser dichten Hit-Ansammlungen sind deren Auswirkungen auf die lokale Hough-Transformierte wie in Abbildung 4.13 nicht ohne weiteres zu trennen.

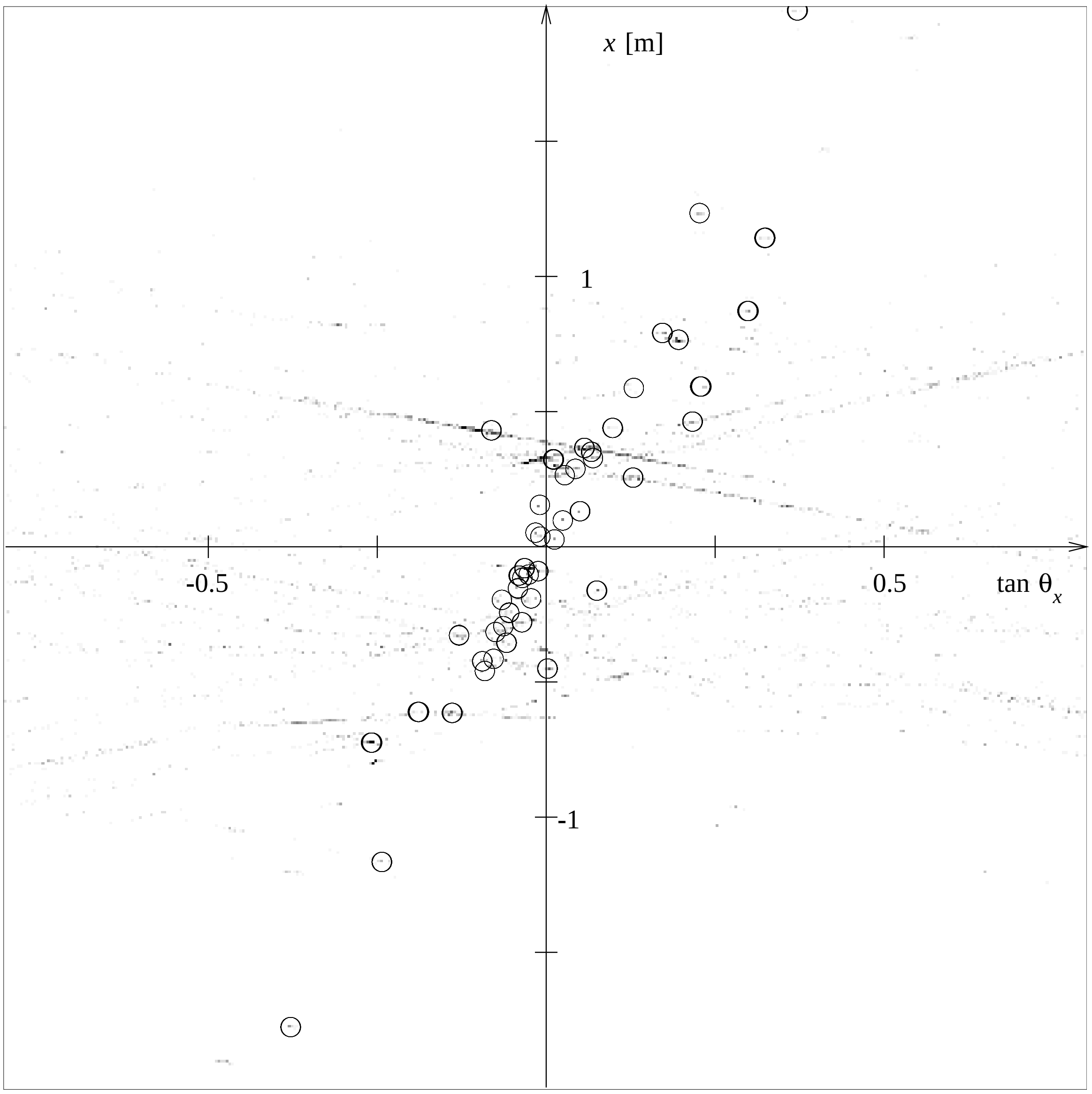


Abbildung 4.13: Hough-Transformation eines simulierten HERA-B-Ereignisses. Die Parameter der Monte-Carlo-Spuren sind durch Kreise gekennzeichnet.

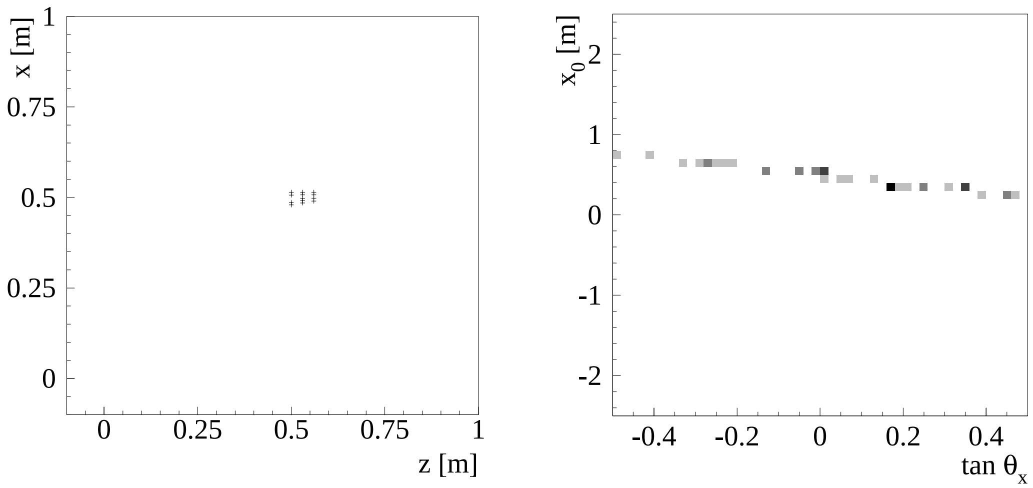


Abbildung 4.14: 13 Hits von vier eng zusammenliegenden Spuren in einer Superlage und das Resultat der lokalen Hough-Transformation.

Die Segment-Initialisierung

Es ist daher nicht problemlos von dem Histogramm auf die Spurdichte im Parameterraum zu schließen. Aus diesem Grund wurde für die HERA-B-Simulationsdaten die Segment-Initialisierung verwendet.

Diese ist ausreichend schnell — ihre Rechenzeit liegt im Sekundenbereich — und hat eine recht gute Segmentfindungseffizienz. Bei Zuordnung der Hits der 0°-Lagen zu den Segmenten kann die auf Hit-Zuordnungen basierende Effizienzdefinition aus Abschnitt 3.5.1 benutzt werden. Die resultierende zweidimensionale Spurfindungseffizienz ist in Abbildung 4.15 dargestellt. Der maximale Hit-Segment-Abstand ist $5\sigma_{\mathrm{Hit}}$, weiter entfernte Hits werden nicht zugeordnet.

Da die HERA-B-Spurkammern aus Modulen in verschiedenen y-Bereichen bestehen, kann man bei der Segmentsuche in den 0°-Lagen Informationen über die grobe Lage der Segmente in y-Richtung bekommen. Es werden keine Segmente gebildet, deren Modulzugehörigkeit auf eine extreme Steigung in y-Richtung schließen lassen würde. Die Templates können in der mittleren y-Koordinate der beteiligten Module initialisiert werden.

Mit dieser im wesentlichen zweidimensionalen Initialisierung werden nun mit der Elastic-Arms-Methode dreidimensionale Spuren gesucht.

4.6.4 Spurfindungseffizienz

Die konventionelle Mustererkennung für die HERA-B-Spurkammern benutzt folgende Definition der Spurfindungseffizienz:

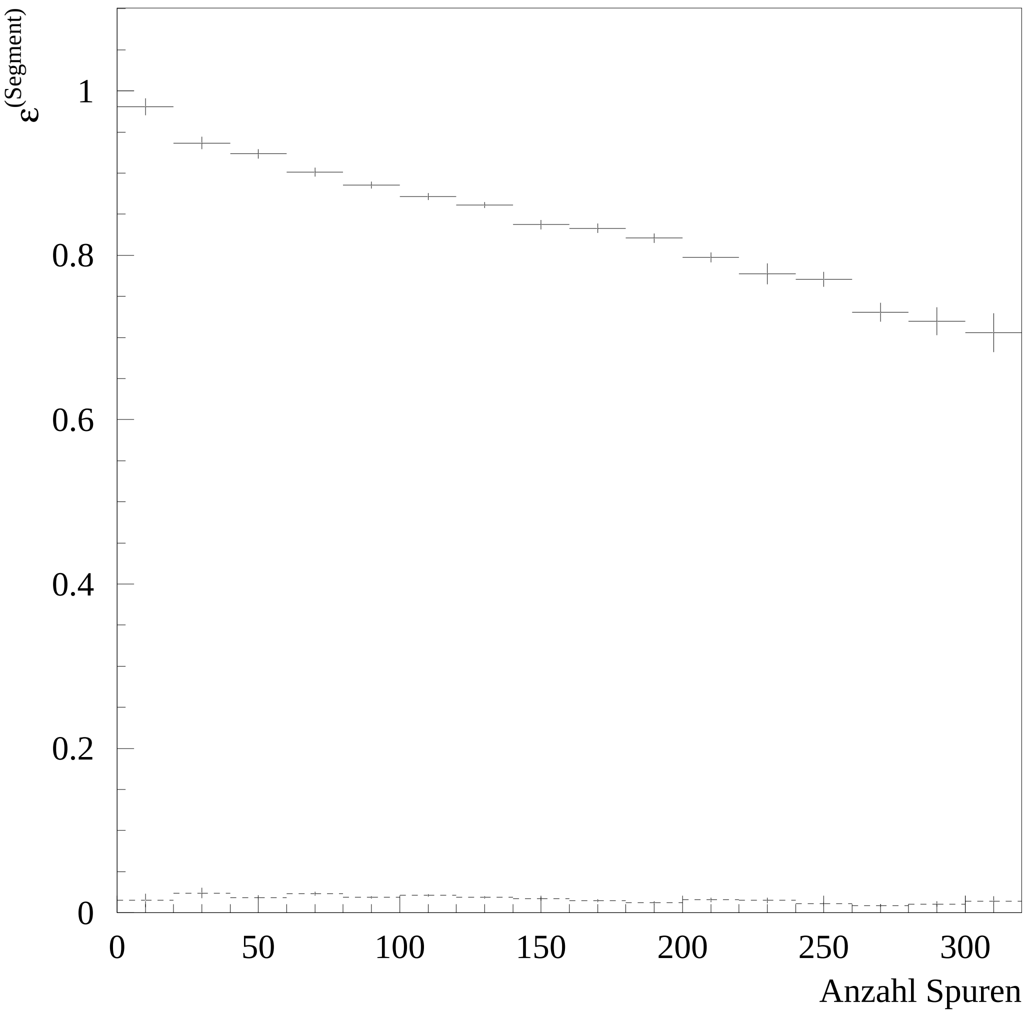


Abbildung 4.15: Zweidimensionale Spurfindungseffizienz $\epsilon^{(\mathrm{Segment})}$ der Segmentsuche in der xz-Projektion und Geisterrate

- Es werden *Referenzspuren* definiert, die von Teilchen mit einem minimalen Impuls von 1 GeV/c herrühren und mindestens 15 Hits in den Spurkammern haben. Alle anderen Spuren werden bei der Effizienzberechnung ignoriert.

- Rekonstruierte Spuren müssen ein minimale Anzahl von Hits haben, diese Anzahl liegt bei etwa 13 bis 15.

- 70% der zugeordneten Hits müssen von derselben simulierten Spur stammen. Es wird vorausgesetzt, daß die Hits von anderen Spuren so nahe liegen, daß die Ausgleichsrechnung („*Spurfit*") nur wenig beeinflußt wird.

Da der Elastic-Arms-Algorithmus auch weit entfernte Punkte einem Template zuordnen kann, muß hier eine weitere Anforderung gestellt werden:

- die maximal 30% Hits von anderen Spuren dürfen nicht weiter als $5\,\sigma_{\mathrm{Hit}}$ von der zuzuordnenden Monte-Carlo-Spur entfernt liegen.

Die Zuordnung weiter entfernter Hits würde eine Ausgleichsrechnung extrem verfälschen. Aus den in Abschnitt 3.5.1 genannten Gründen soll an einer Hit-orientierten Effizienzdefinition festgehalten werden.

4.6.5 Wahl der Parameter

Die Anfangstemperatur muß etwa bei der zu erwartenden Anfangsdistanz der Templates und der Hits liegen. Die längsten zusammenhängend ausgelesenen Detektordrahtstücke sind etwa 2 m lang. Wenn das Template einer am Ende vorbeiführenden Spur in der Höhe der Mitte des Drahtes initialisiert wird, liegt es bei 5° Stereowinkel etwa 3 cm entfernt von den Stereodrähten. Der Parameter β_{start} ist damit[6]

$$\beta_{\mathrm{start}} \approx \frac{1}{(3\,\mathrm{cm})^2} \approx 0.05\,\mathrm{cm}^{-2} \quad . \tag{4.15}$$

Für β_{end}:

$$\beta_{\mathrm{end}} \approx \frac{1}{(200\,\mu\mathrm{m})^2} = 2500\,\mathrm{cm}^{-2} \quad . \tag{4.16}$$

Die Schrittweite η sollte so gewählt werden, daß in keinem Fall das Energieminimum übersprungen wird, andererseits muß η möglichst groß sein. Es zeigt sich, daß folgende Werte diese Kriterien erfüllen:

$$\begin{aligned} \eta^{(\tan\theta)} &= 10^{-7} && \text{für } \tan\theta_x \text{ und } \tan\theta_y \\ \eta^{(0)} &= 0.005 && \text{für } x_0,\, y_0 \quad . \end{aligned} \tag{4.17}$$

[6] In der HERA-B-Simulation ist die Basis-Längeneinheit 1 cm, im Modelldetektor dagegen 1 m.

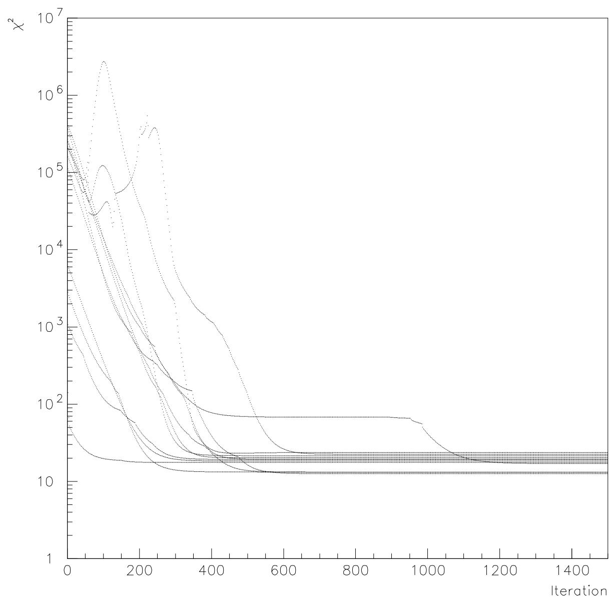


Abbildung 4.16: Konvergenzverhalten zwölf einzelner Templates

4.6.6 Das Konvergenzverhalten der Templates im HERA-B-Detektor

Das Verhalten der Templates in der HERA-B-Simulation wird schrittweise mit vereinfachten Ereignissen untersucht. Vor jeder Elastic-Arms-Rechnung wird die Segment-Initialisierung durchgeführt. Im folgenden werden die Templates a gelöscht, für die $\sum_i V_{ia} < 8$ gilt. Diese Grenze ist sehr niedrig gewählt, um das Verhalten überschüssiger Templates aufzuzeigen.

Einzelne Spuren

Das Konvergenzverhalten läßt sich mit einer einzigen Spur beobachten. Das χ^2 des Templates sollte gegen das χ^2 der Ausgleichsrechnung konvergieren. In Abbildung 4.16 sind die χ^2 von 12 Ereignissen mit je einer Spur im Verlauf von 1500 Iterationen aufgetragen. Dazu ist kein Annealing-Prozeß nötig, da für alle Potts-Faktoren bei jeder Temperatur $V_{ia} = 1$ gilt. Die Konvergenz wird bei allen zwölf Ereignissen spätestens nach 1000 Iterationen erreicht und liegt bei $10 < \chi^2 < 25$, was gleich der Zahl Freiheitsgrade bei 14 bis 29 Hits und damit der Erwartungswert ist.

Isotrop verteilte Myonen

Als nächster Schritt soll das Konvergenzverhalten für einen Annealing-Vorgang untersucht werden. Dazu werden im Akzeptanzbereich des Detektors isotrop zehn Myonspuren mit einer Energie von $E = 10\,\mathrm{GeV}$ erzeugt. Diese Spuren zeigen keine merklichen

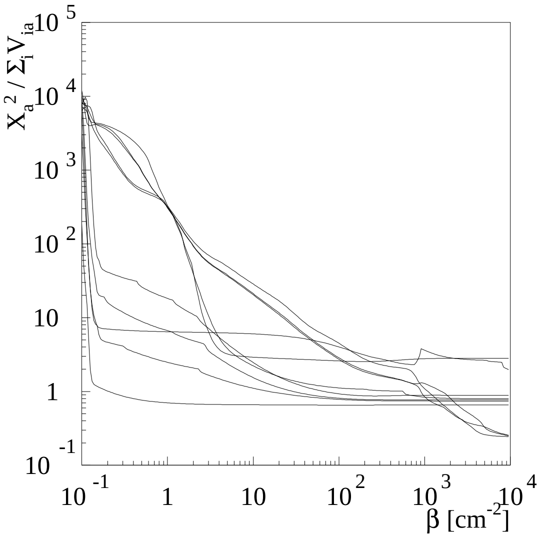


Abbildung 4.17: Konvergenzverhalten zehn isotrop verteilter 10 GeV-Myonspuren

Einflüsse von Vielfachstreuung, so daß das Ergebnis des Rekonstruktion eindeutig zu beurteilen sein sollte.

Während des Annealing-Prozesses wird β während 50000 Iterationen von $\beta_{\text{start}} = 0.05\,\text{cm}^{-2}$ auf $\beta_{\text{end}} = 10000\,\text{cm}^{-2}$ erhöht. Die weitestgehende Rekonstruktion sollte bei $\beta = 1/\sigma_{\text{hit}}^2 \approx 2500\,\text{m}^{-1}$ erreicht werden. Die höhere Auflösung der inneren Spurkammern kann hier aufgrund der insgesamt geringen Spurdichte vernachlässigt werden. Zur Beschreibung der Entfernung der Templates von den Hits wird die Größe

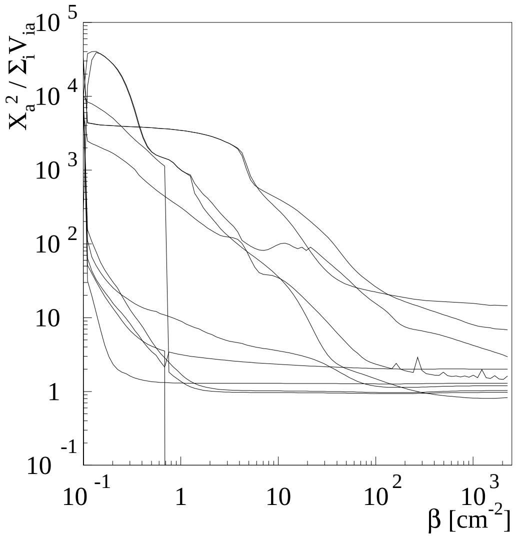
$$X_a^2 = \sum_i V_{ia} M_{ia} \tag{4.18}$$

eingeführt. Sie konvergiert bei $\beta \to \infty$ gegen das χ^2 eines Templates. In Abbildung 4.17 ist der Ausdruck $X_a^2 / \sum_i V_{ia}$ für die zehn Templates während des Annealing-Prozesses dargestellt. Bei $\beta \geq 2500\,\text{m}^{-2}$ sind die meisten Templates auf einen Wert von $\chi^2 / \sum_i V_{ia} \lesssim 1$ konvergiert. Der zu erwartende Wert liegt abhängig von der Zahl der Hits geringfügig[7] unter 1. Das weist darauf hin, daß die meisten Templates auf den simulierten Spuren liegen. Nach der Effizienzdefinition aus Abschnitt 4.6.4 werden acht der zehn Spuren gefunden.

Gefilterte HERA-B-Ereignisse

Als nächstes soll der Elastic-Arms-Algorithmus auf eine HERA-B-Ereignis-Simulation angewendet werden. Zur besseren Beurteilung der Konvergenz werden alle nicht zu

[7] Der Erwartungswert einer χ^2-Verteilung ist die Zahl der Freiheitsgrade n_F, was für ein konvergiertes Template bei vier Spurparametern gleich $\sum_i V_{ia} - 4$ ist.



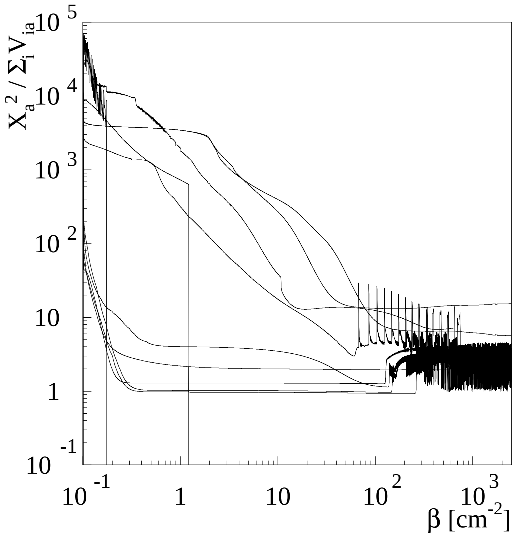


Abbildung 4.18: Konvergenzverhalten für ein gefiltertes HERA-B-Ereignis mit zehn Referenzspuren,
links $\eta^{(\tan\theta)} = 0.005$, $\eta^{(0)} = 10^{-7}$, rechts $\eta^{(\tan\theta)} = 0.01$, $\eta^{(0)} = 10^{-6}$.

Referenzspuren gehörenden Hits vorher gelöscht. Im Beispielereignis bleiben damit zehn Spuren.

Der Verlauf der Größe X_a^2 ist in Abbildung 4.18 gezeigt. Dort sind vertikale Linien eingetragen, die das Löschen von Templates anzeigen.[8] Die linke Abbildung zeigt das Konvergenzverhalten bei den oben beschriebenen Schrittweiten, die rechte das Verhalten bei einer größeren Schrittweite von ($\eta^{(\tan\theta)} = 0.01, \eta^{(0)} = 10^{-6}$. Während links eine stetige Verkleinerung von X^2 zu erkennen ist, gibt es rechts sowohl am Anfang als auch am Ende des Annealing-Prozesses Oszillationserscheinungen, die eine Konvergenz verhindern. Die Templates, welche bei beiden Schrittweiten früh konvergieren, erreichen ihre endgültige Lage schneller. Die Spurfindungseffizienz für den Annealing-Prozeß auf der linken Seite ist 80%.

4.6.7 Rechenzeit des Elastic-Arms-Algorithmus

Ein Problem des Elastic-Arms-Algorithmus für die Anwendung bei HERA-B liegt im Rechenzeitbedarf. Die Zehn-Spur-Ereignisse aus Abschnitt 4.6.6 benötigen in der vorliegenden Implementierung für einen Annealing-Prozeß etwa 4000 s. Die Rechenzeit hängt in der ursprünglichen Formulierung der Methode mindestens quadratisch mit der Anzahl der Spuren ≈ Anzahl der Templates M zusammen:

$$\begin{aligned} \text{Zeitbedarf } T &\propto \mathcal{O}(M \cdot N \cdot n_{\text{Iterationen}}) \\ &\propto \mathcal{O}(M^2 \cdot n_{\text{Iterationen}}) \quad . \end{aligned} \tag{4.19}$$

[8]Beim Löschen eines Templates wird die Numerierung vertauscht, was in den Kurven zu erkennen ist.

Da auch die Zahl der nötigen Iterationen mit der Spuranzahl steigt, wird der Zeitbedarf für vollständige HERA-B-Ereignisse mit $N \gtrsim 100$ verglichen mit den vorliegenden zehn Spur-Beispielen sehr groß.

4.7 Ausblick

Es wurde gezeigt, daß mit dem Elastic-Arms-Algorithmus Spuren in der vollständigen HERA-B-Geometrie gefunden werden können.

Eine Verbesserung des Algorithmus sollte beim Rechenzeitbedarf ansetzen. Die Potts-Faktoren jedes Templates sind in Abhängigkeit von β nur in der näheren Umgebung des Templates ungleich null. Eine Vorauswahl der überhaupt zu berechnenden Potts-Faktoren kann die Rechenzeit deshalb deutlich verkleinern. Der Organisationsaufwand für diese Erweiterung ist allerdings recht hoch.

Weiterhin kann die Suche des Energieminimums beschleunigt werden. Es gibt verschiedene schnellere Verfahren, eine Optimierung in vieldimensionalen Räumen zu erreichen. Eine Variante ist, den Vektor des letzten Schrittes in Gradientenrichtung zu speichern und über einen *Impuls-Term* in die folgenden Schritte mit einzubeziehen. Eine andere Möglichkeit ist es, eine Näherung der zweiten Ableitung der Energiefunktion zu berechnen und sie als Schrittweite zu benutzen, was eine deutlich bessere Konvergenz verspricht. Diese Möglichkeit wird in [18] vorgestellt.

Es gibt verschiedene Ansätze, eine sehr flexible Spurrekonstruktion mit der Elastic-Arms-Methode zu erreichen. Die Potts-Faktoren des Templates $a = 0$ können anzeigen, daß Hits zu einer Spur gehören, die nicht in die Parameterdarstellung paßt. Damit kann man zum Beispiel Sekundärvertizes suchen [18]. Ein anderer Ansatz ist die Einführung von Parametrisierungen anderer Effekte wie zum Beispiel δ-Elektronen [19].

5 Zusammenfassung

In dieser Arbeit wurden die Möglichkeiten neuronaler Algorithmen zur Spurerkennung im HERA-B-Detektor untersucht.

Dazu wurde zunächst ein zweidimensionaler Modelldetektor entworfen, der im wesentlichen die Kenngrößen des HERA-B-Spursystems aufweist. Für diesen Modelldetektor wurde der Spurfindungsalgorithmus nach Denby und Peterson implementiert. Es zeigte sich, daß die Struktur der Ereignisse im Detektor hohe Anforderungen an den Algorithmus stellt. Deshalb wurde die Energiefunktion modifiziert.

Mit dieser Änderung findet der Algorithmus bei einer Belegungsdichte bis zu 20% mehr als 85% der Spuren des Modellexperiments. Die Effizienz ließe sich durch Parameteroptimierung und eine geschicktere Auswertung des Endzustands des neuronalen Netzes verbessern.

Weitere Schritte zur Untersuchung des Algorithmus müßten eine Erweiterung auf die Behandlung von Driftkammersignalen und die dreidimensionale Spurerkennung sein. Ersteres ist leicht über einen weiteren Term in die Energiefunktion zu integrieren, letzteres erfordert weitere Überlegungen. Die benötigte Rechenzeit liegt in der benutzten Implementierung deutlich über der von konventionellen Spurverfolgungsmethoden.

Als zweites Beispiel neuronaler Mustererkennung wurde der Elastic-Arms- oder Deformable-Template-Algorithmus nach Ohlsson und Peterson untersucht. Dieser paßt Parameterdarstellungen der zu suchenden Spuren, sogenannte Templates, den Detektorhits an. Die Templates müssen vorher initialisiert werden, was über eine lokale Hough-Transformation oder eine lokale Spursegmentsuche geschehen kann.

Auch dieser Algorithmus wurde für einen Einsatz in einem HERA-B-ähnlichen Modelldetektor modifiziert. Es wurde eine Möglichkeit geschaffen, Templates dynamisch zu erzeugen und zu löschen. Im Modelldetektor wird so eine Spurfindungseffizienz von über 80% erreicht.

Die Erweiterung auf die HERA-B-Simulation bereitet kaum Schwierigkeiten, da der Algorithmus auf eine naheliegende Weise beliebig positionierte Detektorsignale verarbeitet. Die Templates werden in der xz-Projektion über die Methode der Segmentkombination initialisiert.

Die Musterereignisse mit zehn Spuren konnten rekonstruiert werden, wobei allerdings noch eine Optimierung der Parameter nötig ist. Die Spurfindungseffizienz liegt mit den benutzten Parametern bei etwa 80%. Die Rechenzeit ist in der urspünglichen Formulierung des Algorithmus zu groß, um realistische HERA-B-Ereignisse zu rekonstruieren. Die Zahl der Rechenschritte könnte aber stark verringert werden, indem man viele Terme, die vorhersehbar null ergeben, nicht berechnet.

Eine weitere naheliegende Verbesserung ist die Suche nach einem günstigeren Weg zur

Minimierung der Energiefunktion, wie zum Beispiel die Einführung eines sogenannten Impuls-Terms oder die Berücksichtigung der zweiten Ableitung.

Insgesamt konnte gezeigt werden, daß neuronale Algorithmen auch zur Spurerkennung bei modernen Experimenten der Hochenergiephysik prinzipiell geeignet sind. Ihre Untersuchung hat aber gerade erst angefangen.

A Formeln und Tabellen

A.1 Die Wolfenstein-Parametrisierung

Die Wolfenstein-Parametrisierung der CKM-Matrix lautet:

$$\begin{aligned} V &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\ &= \begin{pmatrix} 1-\lambda^2/2 & \lambda & \lambda^3 A(\rho - i\eta) \\ -\lambda & 1-\lambda/2 & \lambda^2 A \\ \lambda^3 A(1-\rho-i\eta) & -\lambda^2 A & 1 \end{pmatrix} \end{aligned} \tag{A.1}$$

Es ist eine Näherung, die die Größenverhältnisse der Quark-Übergänge beschreibt. Der Parameter λ entspricht dem Sinus des Cabibbo-Winkels und ist damit etwa 0.221. Er beschreibt die empirisch gewonnenen Größenverhältnisse der Übergangsmatrixelemente zwischen den verschiedenen Quark-Familien (1,2 und 3):

$$\begin{aligned} |V_{11}| \approx |V_{22}| \approx |V_{33}| &\approx 1 \\ |V_{12}| \approx |V_{21}| &\sim \lambda \\ |V_{23}| \approx |V_{32}| &\sim \lambda^2 \\ |V_{13}|, |V_{31}| &\sim \lambda^3 \end{aligned}$$

A.2 Abstand einer geraden Spur von einem linienförmigen Hit

In diesem Abschnitt werden die Berechnungen für den Elastic-Arms-Algorithmus zusammengefaßt (Skizze in Abbildung A.1). Die Parametrisierung der Spuren ist

$$\vec{x}_t(r) = \vec{t}_0 + r\vec{t} \quad . \tag{A.2}$$

Mit $t_0 = (x_0, y_0, z_{\mathrm{ref}})^\mathsf{T}$, $\vec{t} = (\tan\theta_x, \tan\theta_y, 1)^\mathsf{T}$ und $r \equiv z - z_{\mathrm{ref}}$ ist sie identisch mit Gleichung 1.6. Der linienförmige Hit mit der Länge $|\vec{h}|$, dessen unteres Ende durch die Koordinaten $\vec{h}_0$ beschrieben wird, wird wie folgt parametrisiert:

$$\vec{x}_h(s) = \vec{h}_0 + s\vec{h} \qquad 0 \leq s \leq 1 \quad . \tag{A.3}$$

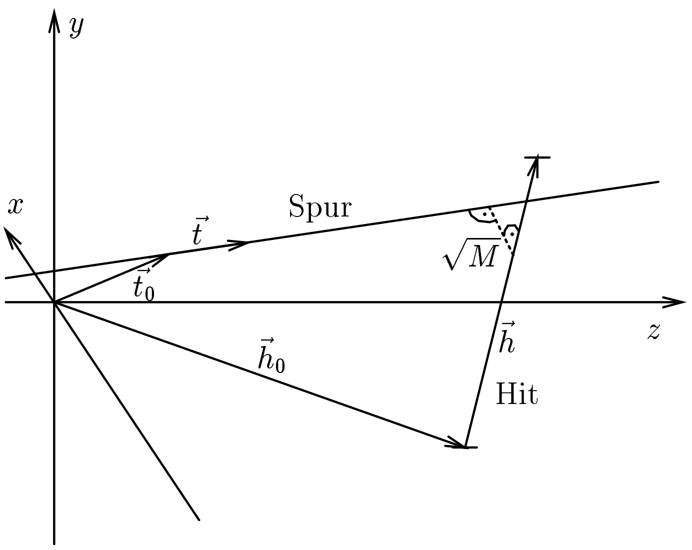


Abbildung A.1: Parameterdarstellung eines Hits

Damit sind die quadrierten Abstände M zwischen einem Hit und einem Template:

$$M = \begin{cases} \dfrac{\Big((\vec{t}_0 - \vec{h}_0)\cdot(\vec{t}\times\vec{h})\Big)^2}{(\vec{t}\times\vec{h})^2} & 0 \leq s_0 \leq 1 \\[2ex] \dfrac{\Big((\vec{t}_0 - \vec{h}_0)\times\vec{t}\Big)^2}{|\vec{t}|^2} & s_0 < 0 \\[2ex] \dfrac{\Big((\vec{t}_0 - \vec{h}_0 + \vec{h})\times\vec{t}\Big)^2}{|\vec{t}|^2} & s_0 > 1 \end{cases} \tag{A.4}$$

$$\text{mit} \qquad s_0 = \frac{(\vec{t}_0 - \vec{h}_0)\cdot\Big(\vec{t}\times(\vec{h}\times\vec{t})\Big)}{(\vec{t}\times\vec{h})^2} \quad . \tag{A.5}$$

Deren Ableitungen der Abstände nach den Geradenparametern sind

$$\frac{\partial M}{\partial t_{0,i}} = -2S\frac{(\vec{t}\times\vec{h})_i}{|\vec{t}\times\vec{h}|} \tag{A.6}$$

$$\frac{\partial M}{\partial t_i} = -2S\left(\frac{\vec{t}\times\vec{h}}{|\vec{t}\times\vec{h}|} - S\frac{\vec{h}\times(\vec{t}\times\vec{h})}{(\vec{t}\times\vec{h})^2}\right)_i \tag{A.7}$$

$$\text{mit} \quad S = \frac{(\vec{t}_0-\vec{h}_0)\cdot(\vec{t}\times\vec{h})}{|\vec{t}\times\vec{h}|} \quad . \tag{A.8}$$

Bei Driftkammersignalen, deren Hits die Form eines Zylindermantels haben, wird vor der Abstandsberechnung die Linie gesucht, die Element des Zylindermantels ist und der Spur am nächsten liegt.

B Implementierung des Elastic-Arms-Algorithmus

Die beiden Spurerkennungsalgorithmen wurden in C++ unter Verwendung der Klassenbibliothek *CLHEP* implementiert. Als Beispiel eines objektorientierten Ansatzes folgt ein Programm-Auszug des Elastic-Arms-Algorithmus.

Die Grundbausteine des Programms sind Listen, die Objekte vom Typ T speichern (`HepAList`<T>), Iteratoren, die die Listen durchlaufen (`HepAListIterator`<T>), Vektoren (`HepVector`<T>) und geometrische Objekt, wie etwa `Template` oder `Hit`.

Die zentrale Objektklasse heißt `ElasticArms`. Sie definiert die *Methode (Memberfunction)* `Iterate(...)`, welche einen Iterationsschritt ausführt. Als Parameter übernimmt sie eine Liste von Template-Objekten (`templates`) und eine Liste von Detektor-Hits (`hits`).

Die Funktion berechnet zunächst die Schritte für alle Templates und ändert die Templates erst im Anschluß. Die Beiträge der einzelnen Hits zu den Schritten werden im Vektor `tr` akkumuliert. Die geometrischen Funktionen werden aus `ElasticArms::fillExpField(...)` aufgerufen.

```
// Elastic Arms Iteration.

void ElasticArms::Iterate( const HepAList<HitBase>& hits,
                           HepCList<Template>& templates )
{
  HepVector<Template> tr[templates.length()];          // Memory for track steps
  for( HepAListIterator<HitBase> hIt( hits );           // Loop through hit list.
       hIt.current(); hIt.next() ) {
    HepVector<double>
      m( templates.length() );                  // Vector for exp(-beta*distSquared.)
    double ma = fillExpField( m, *hIt.current(), templates );
                  // Get all exp-terms and their sum.
    for( HepCListIterator<Template> tIt( templates );  // Loop through templ. list.
         tIt.current(); tIt.next() ) {
      double
        pottsFactor = m[tIt.index()] / (exp(-beta*lambda) + ma);
      if(!isnan(potts)) {                          // Test if Potts factor is valid.
        InfiniteLine l = hIt.current()->derivedDistSquared( *trackP, beta );
                  // Get derivative of squared distance and store it in line object.
        tr[tIt.index()] += potts*l;                                  // Accumulate.
```

```
            }
        }
    }
    for( HepCListIterator<Track> tIt( tracks );          // Loop through track list.
            tIt.current(); tIt.next ) {
        *tIt.current() += derTr[tIt.index()];                                   // Step.
    }
}
```

Der Vektor `m` wird von der Hilfsfunktion `ElasticArms::fillExpField(...)` gefüllt. Diese übernimmt einen Hit und die Template-Liste sowie den Vektor und füllt ihn mit den Ausdrücken $\exp(-\beta M_{ia})$. Außerdem wird die Summe der exp-Funktionen zurückgegeben. Der Abstand M_{ia} wird von der Member-Function `distSquared(const Template& t)` der Hit-Klasse berechnet. Diese führt die Objekte je nach Typ des Hits (MSGC oder HDC) auf Linien im Raum zurück und berechnet die Abstände nach den Gleichungen aus A.2.

```
// Calculate exp(-beta*distSquare) for one hit and all templates.

inline double ElasticArms::fillExpField( HepVector<double>& m,
                                         const HitBase& hitP,
                                         const HepCList<Track> & tracks
                                         ) const
{
    double ma = 0;
    for( HepCListIterator<Track> tIt( tracks ); tIt.current(); tIt.next() ) {
        double distSquare = hitP->distSquared( *tIt.current() );
        ma += (m[tIt.index()] = exp(-beta*distSquare ) );
    }
    return ma;
}
```

Literaturverzeichnis

[1] J. L. Rosner, *The Cabibbo-Kobayashi-Maskawa Matrix*, in S. Stone (Ed.), *B decays*, World Scientific, Singapore (1992) 470–519.

[2] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix* Phys. Rev. Lett. 51 (1983), 1945.

[3] HERA-B-Collaboration, *HERA-B Proposal* DESY-PRC 94/02 (1994)

[4] R. Mankel, *Application of the Kalman Filter Technique in the HERA-B Track Reconstruction*, HERA-B-Note 95-239, `http://www-hera-b.desy.de/subgroup/software/kalman-filter.ps`, 1995.

[5] J. Hertz, A. Krogh, R. G. Palmer, *Introduction to the Theory of Neural Computation*, Addison-Wesley Publishing Company, 1991.

[6] W. S. McCulloch, W. Pitts, *A Logical Calculus of Ideas Immanent in Nervous Activity* Bull. Math. Biophys. 5 (1943) 115–133. Nachgedruckt in [21].

[7] J. J. Hopfield, *Neural Networks and Physical Systems with Emergent Collective Computational Abilities*, Proceedings of the National Academy of Science, USA, 79 (1982) 2554. Nachgedruckt in [21].

[8] Y. Shrivastava, *Guaranteed Convergence in a Class of Hopfield Networks*, IEEE Transactions on Neural Networks, Vol. 3, No. 6 (1992) 951–961.

[9] C. Peterson, B. Söderberg, *A New Method for Mapping Optimization Problems onto Neural Networks*, International Journal of Neural Systems 1 (1989) 3 537–549.

[10] S. Kirkpatrick, C. D. Gelatt Jr, M. P. Vecchi, *Optimization by Simulated Annealing* Science 220 (1983) 671–680 Nachgedruckt in [21].

[11] B. Denby, *Neural Networks and Cellular Automata in Experimental High Energy Physics*,
Comput. Phys. Commun. 49 (1988) 429–448.

[12] C. Peterson, *Track Finding with Neural Networks*,
Nuclear Instruments and Methods in Physics Research A279 (1989) 537–549.

[13] G. Stimpfl-Abele, Lluís Garrido, *Fast track finding with neural networks*,
Comput. Phys. Commun. 64 (1991) 46–56.

[14] H. Kolanoski, *Anwendungen neuronaler Netze in der Teilchenphysik*,
Vorlesungsskript Dortmund 1993.

[15] M. Gyulassy, M. Harlander, *Elastic tracking and neural network algorithms for complex pattern recognition*,
Comput. Phys. Commun. 66 (1991) 31–46.

[16] M. Ohlsson, C. Peterson, A. L. Yuille, *Track finding with deformable templates – the elastic arms approach*,
Comput. Phys. Commun. 71 (1992) 77–98.

[17] T. Schober, Diplomarbeit, Humboldt-Universität zu Berlin, 1996.

[18] M. Ohlsson, *Extensions and Explorations of the Elastic Arms Algorithm*,
Comput. Phys. Commun. 77 (1993) 19–32.

[19] R. Blankenbecler, *Deformable Templates – Revised and Extended – With an OOP Implementation*,
Comput. Phys. Commun. 81 (1994) 318–334.

[20] R. Blankenbecler, *A Unified Treatment of Track Reconstruction and Particle Identification*,
Comput. Phys. Commun. 81 (1994) 335–342.

[21] J. A. Anderson, E. Rosenfeld, *Neurocomputing: Foundations of Research*,
MIT Press, Cambridge (1988).

Danksagungen

Mein Dank geht an Prof. Dr. Kolanoski dafür, daß er mir die Möglichkeit gegeben hat, nach Zeuthen zu kommen, für die interessante Aufgabenstellung, für vielfache Hilfen zum neuronalen Thema und für detaillierte Verbesserungsvorschläge;

an Prof. Dr. Lohse für interessante Anregungen und echtes Interesse an meiner Arbeit;

an Dr. Rainer Mankel für die geduldige Betreuung meiner Diplomarbeit, für unermüdliche Hilfe bei der Suche nach guten Wörten und für die vielfältige moralische Unterstützung;

an Axel Köhler, der eine Engelsgeduld bewiesen hat bei der Installation von einigen C++-Compilern und verschiedenen LaTeX-Versionen sowie bei der umständlichen Konfigurierung verschiedener Programme

und an meine Eltern, die mir durch moralische und finanzielle Unterstützung das Studium ermöglicht haben.