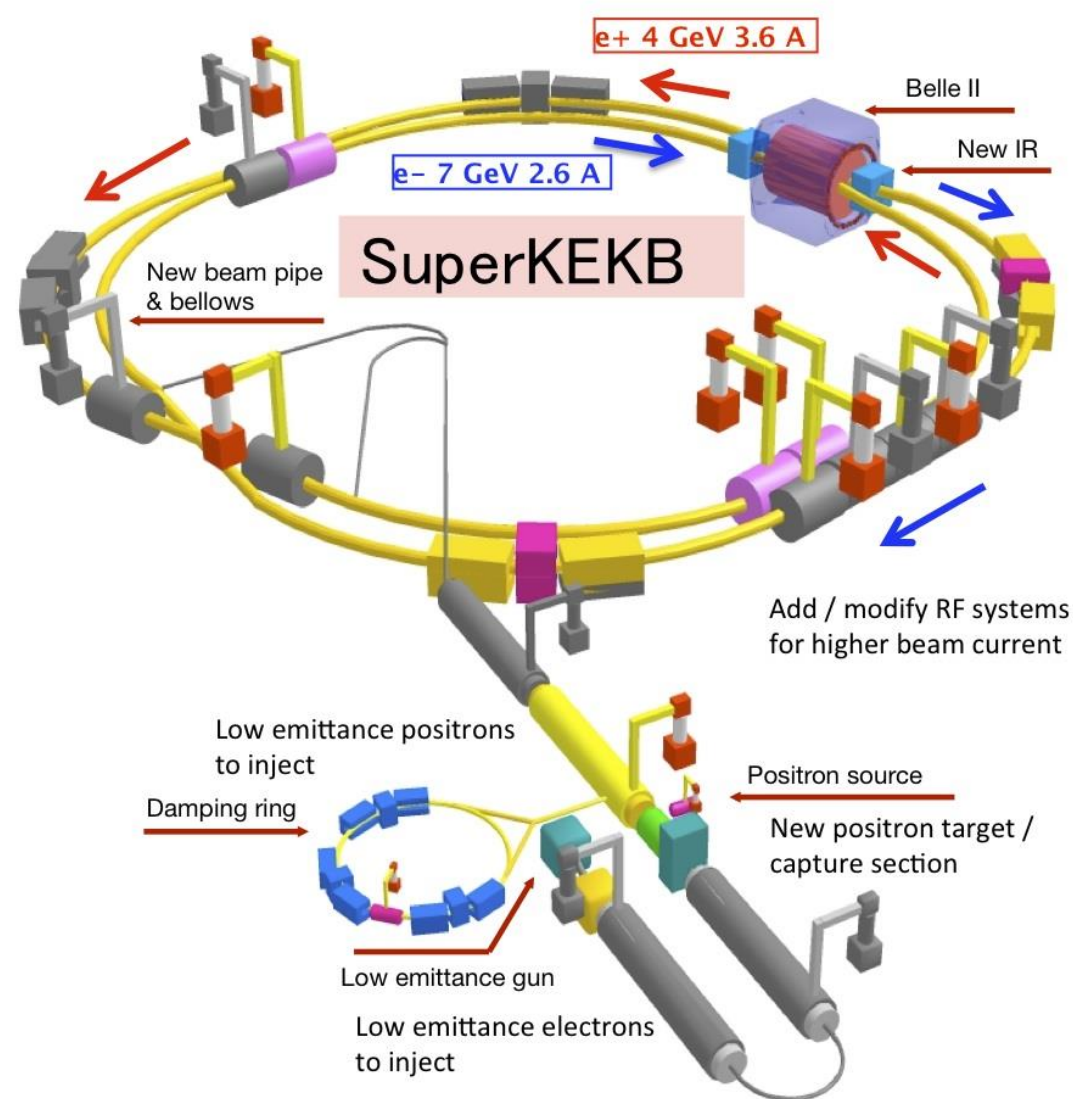


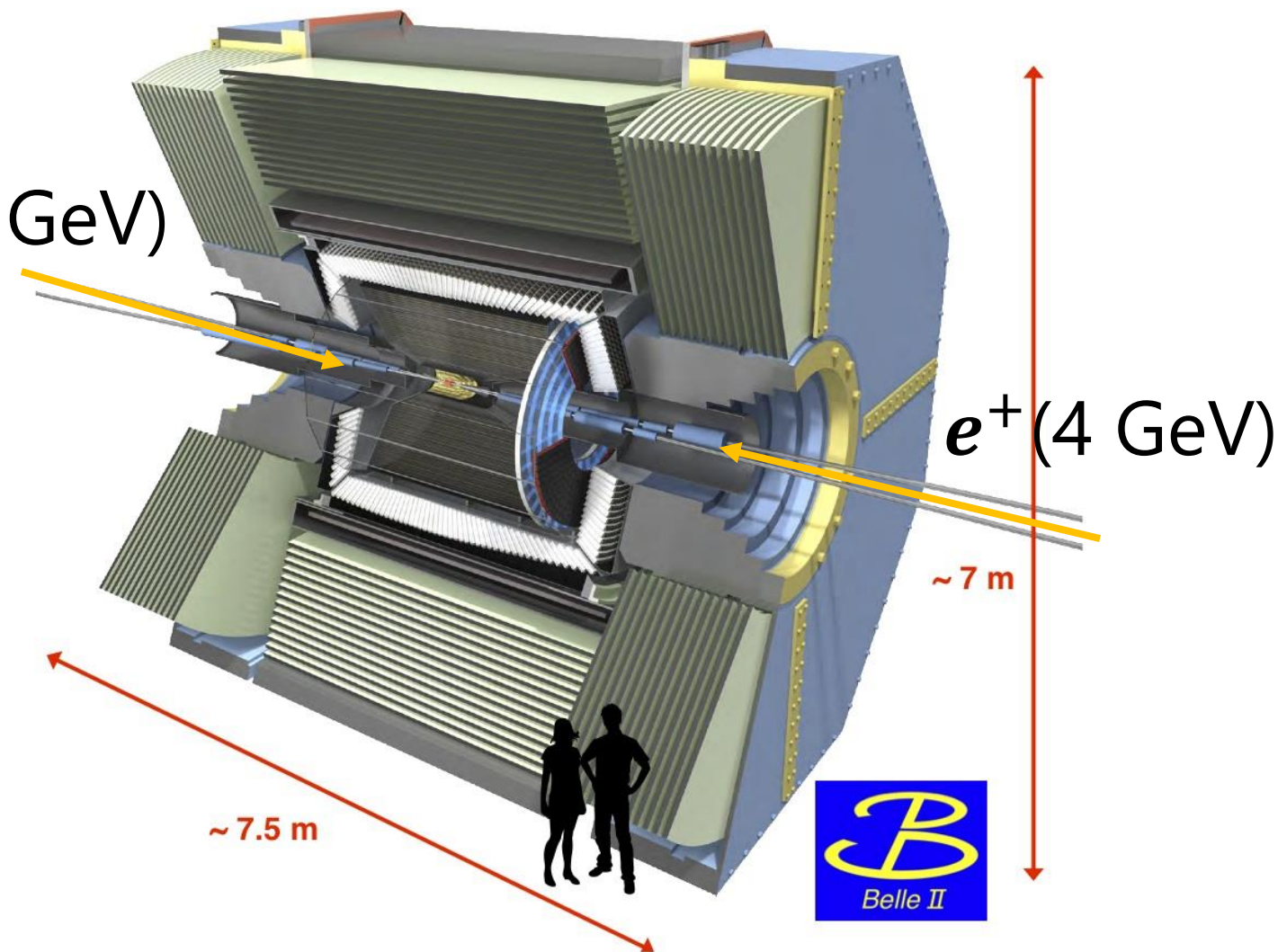
ニューラルネットワークを用いた Belle II 実験TOPカウンター 粒子識別アルゴリズムの開発

Flavor Physics Workshop

原舜基



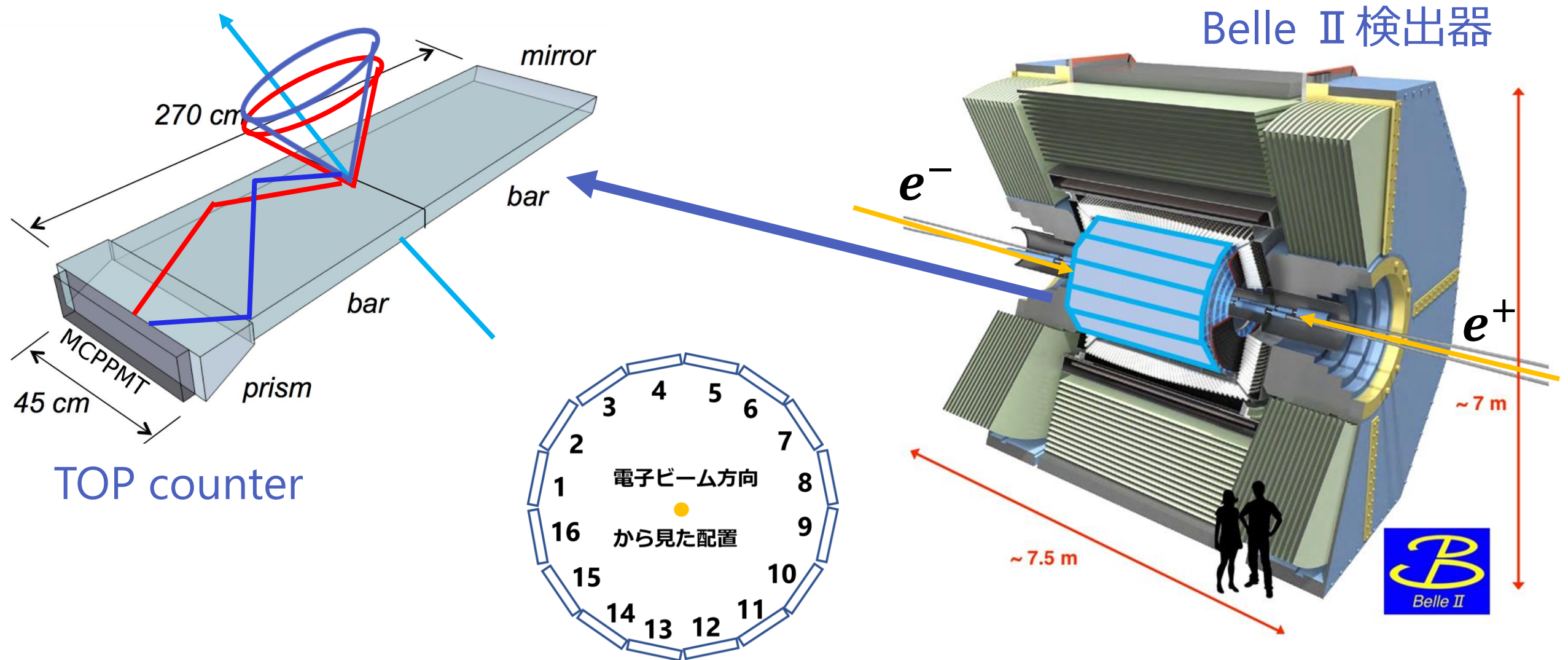
$e^-(7 \text{ GeV})$



- ◆ 重心系エネルギー 10.58 GeV の電子陽電子衝突実験
- ◆ B中間子・ τ 粒子などの崩壊の精密測定による標準理論の精密検証や新物理の探索

Belle II 実験 TOPカウンター

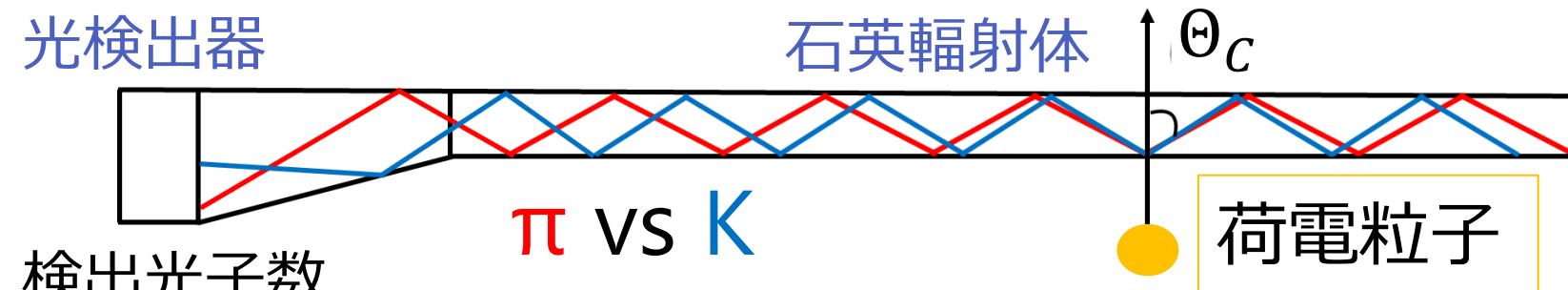
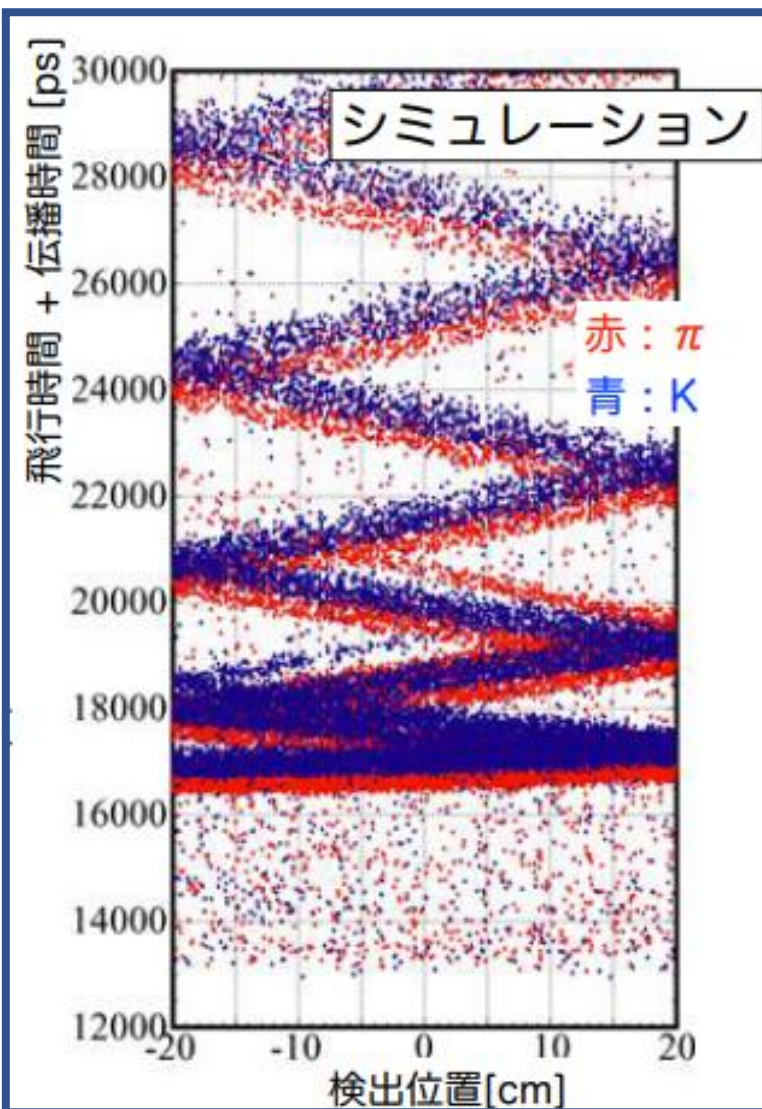
3



- ◆ Time-Of-Propagation (TOP)カウンター{16台}:
Belle II 実験バレル部分の粒子識別装置
チェレンコフ光の伝播経路の差を利用し識別を行う

TOPカウンターの識別原理

4



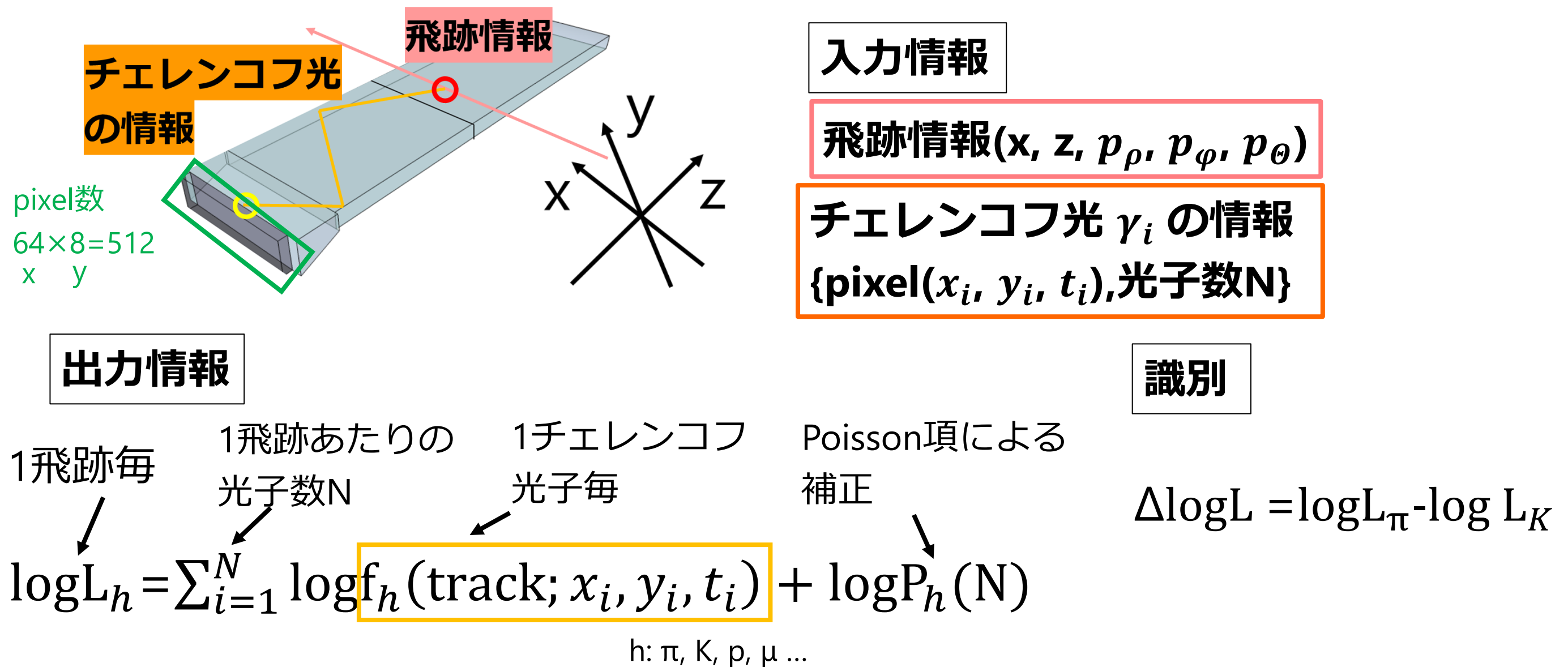
$$\cos\theta_c = \frac{\sqrt{m^2 + p^2}}{np}$$

m:荷電粒子質量
p:荷電粒子運動量
n:石英輻射体の屈折率

- ◆ 荷電粒子が入射した際に発するチェレンコフ光の検出位置・時間はチェレンコフ角 θ_c によって変化する
- ◆ ある運動量におけるチェレンコフ角 θ_c は粒子の種類(質量)によって変化するため、飛跡検出器で得られた運動量と検出位置・時間を組み合わせて粒子識別を行う

現行の識別手法

5



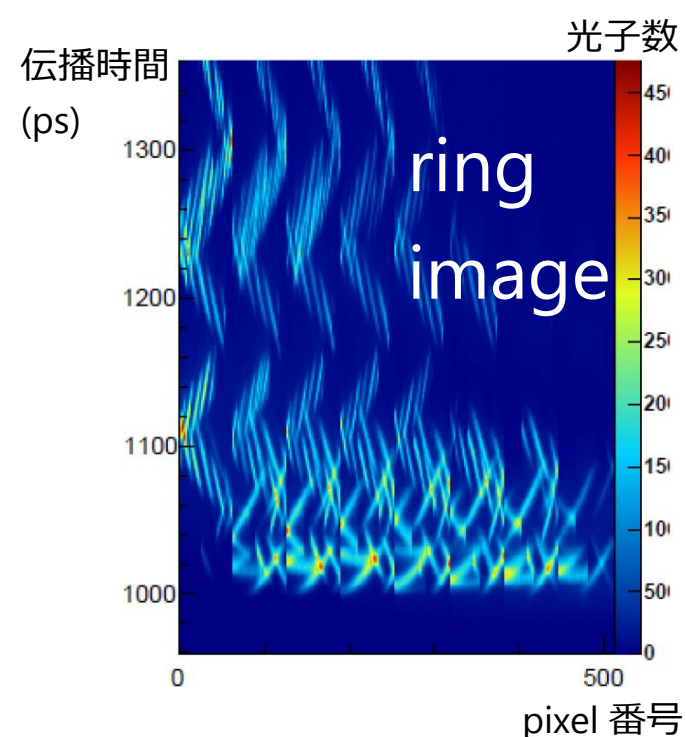
◆ 現行の識別手法では尤度が以下のように計算される

1. 粒子の種類、飛跡情報ごとに解析計算される確率密度関数にチェレンコフ光の情報を当てはめる
2. 1飛跡に相当する光子の分だけ1.を掛け合わせる

現行の手法の課題

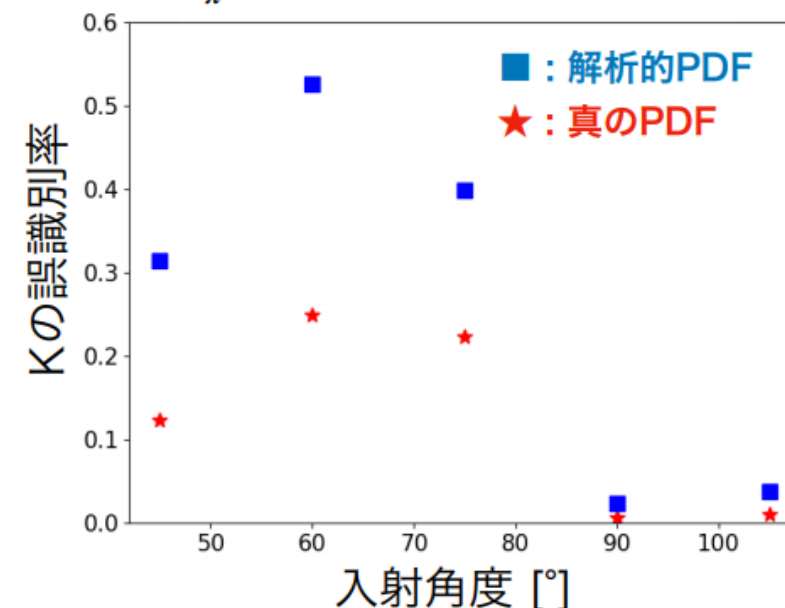
- ◆ 現在は解析的な計算を用いて識別を行っているが、解析的な計算は完璧ではない
 - チェレンコフ光波長の違いの近似
 - 輻射体の中での光子放出点の違いの近似
- ◆ シミュレーションで生成した理想的な確率密度を用いた識別が解析計算した識別性能を上回る
 - しかし、全位相空間で理想的な確率密度を生成することは膨大な粒子数が必要になるため不可能

⇒機械学習を用いて識別を行うことはできないか



$p = 3 \text{ GeV}$ 時

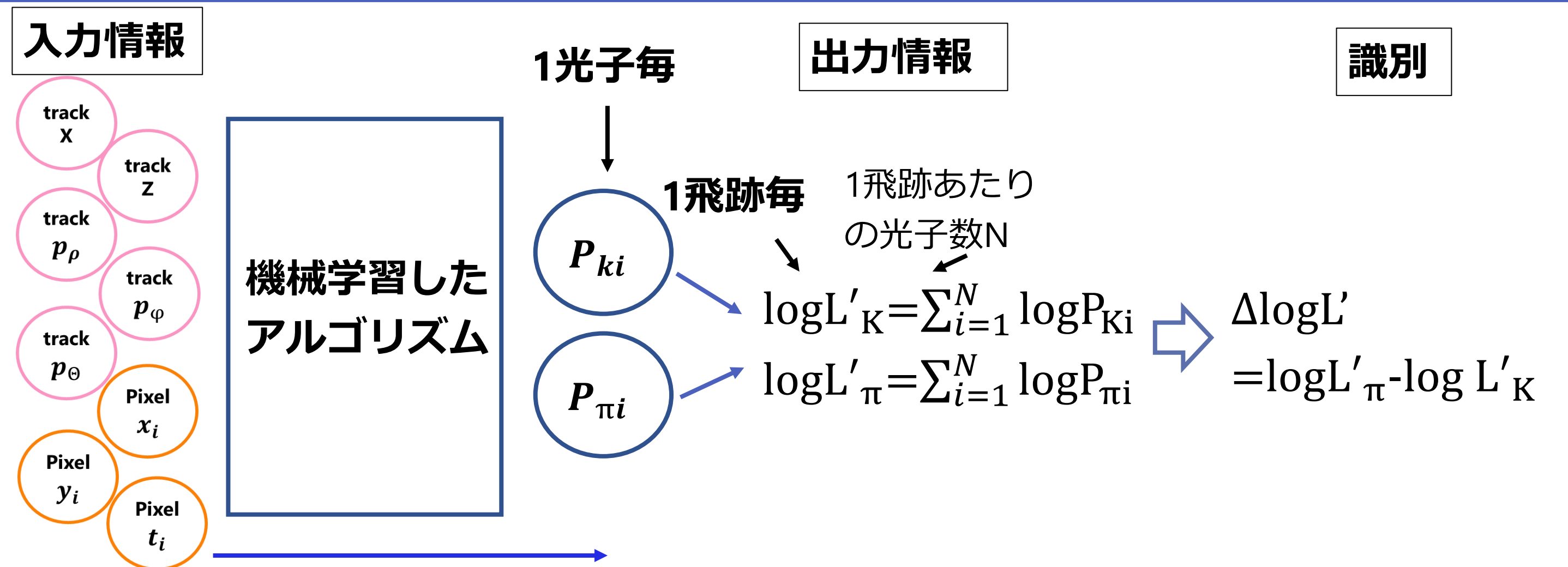
$\epsilon_\pi = 90\%$ 時のKの誤識別率



JPS2020spring
16aH31-7
山之内

新しく考案した識別手法

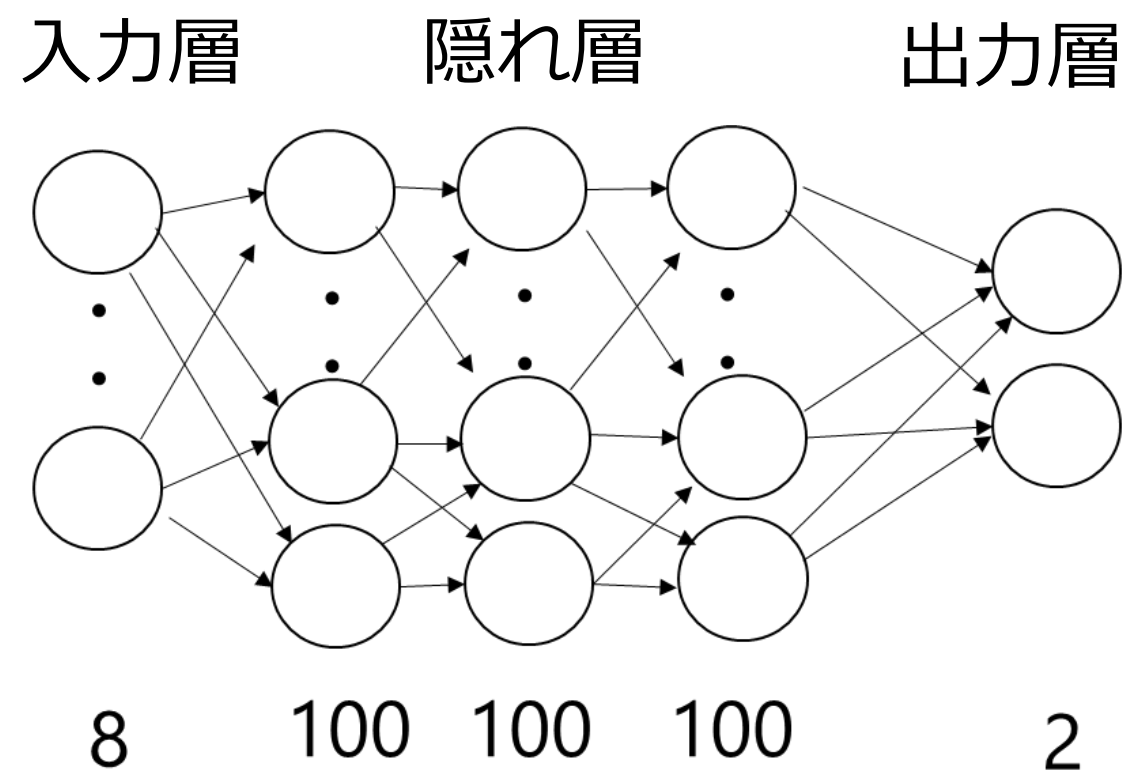
7



- ◆ 新しく機械学習を用いた以下の識別手法を考案した
 1. 飛跡情報(5つ)とチェレンコフ光の情報(3つ)を入力とし、機械学習を用いて1光子ごとに P_{Ki} , $P_{\pi i}$ を求める
 2. 1飛跡に相当する光子の分だけ1.を掛け合わせ L' を求める

本研究で用いる機械学習手法

- ◆ 本研究ではニューラルネットワーク(NN)という機械学習手法を用いる(使用ライブラリ: TensorFlow)



- ◆ 層から層へ信号が伝達される時に必要なパラメータを学習によって最適化する

⇒ **解析計算で表すのが難しい
様々な影響を含めることが可能**

- ◆ K・ π 粒子を衝突点から打ち込んだシミュレーションデータを学習用・評価用に分けて使用

ハイパーパラメータ

隠れ層: 3層 epoch数: 300 Neuron数: 100-100-100

Optimizer: Adam バッチサイズ: 5000 損失関数: 交差エントロピー

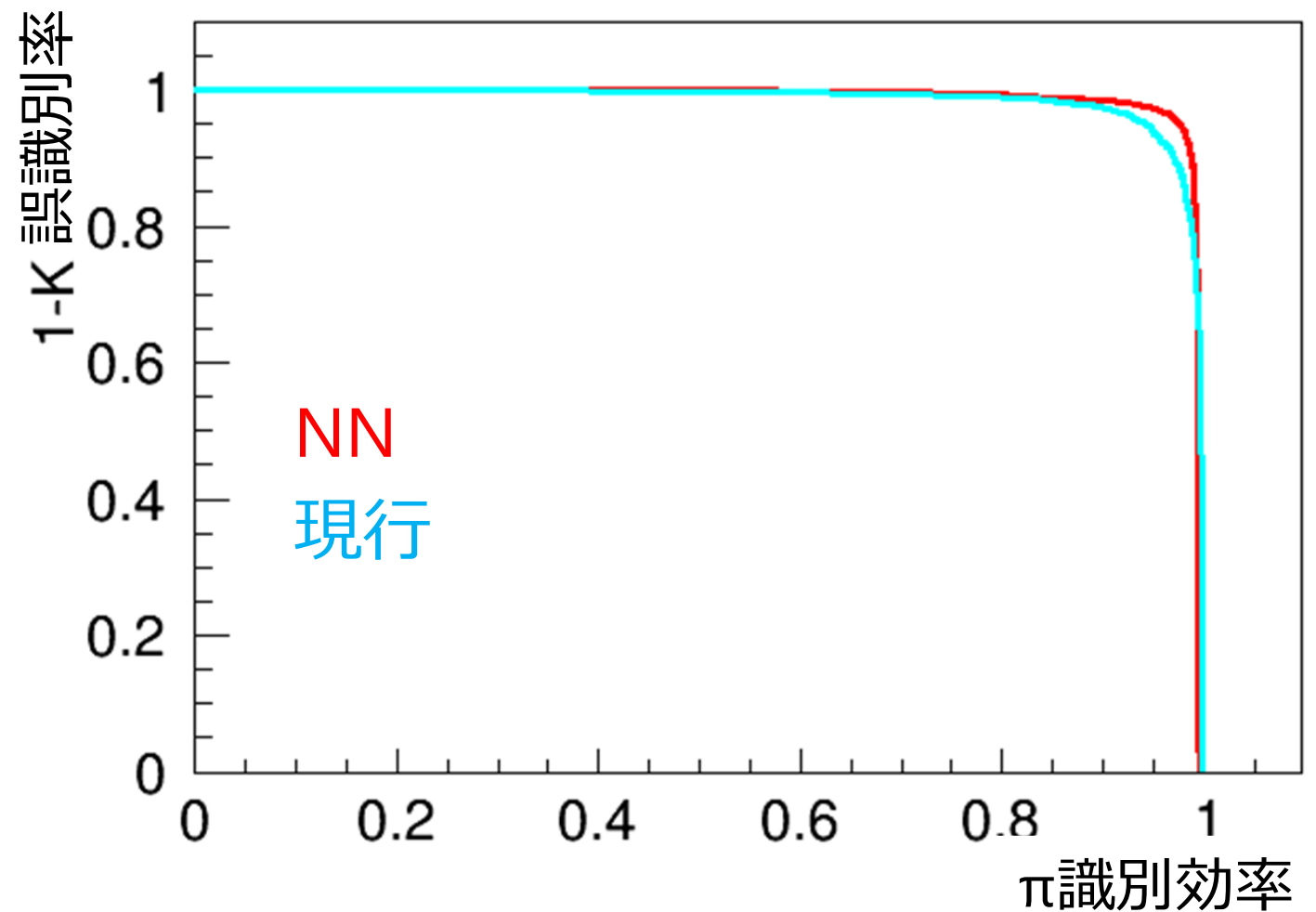
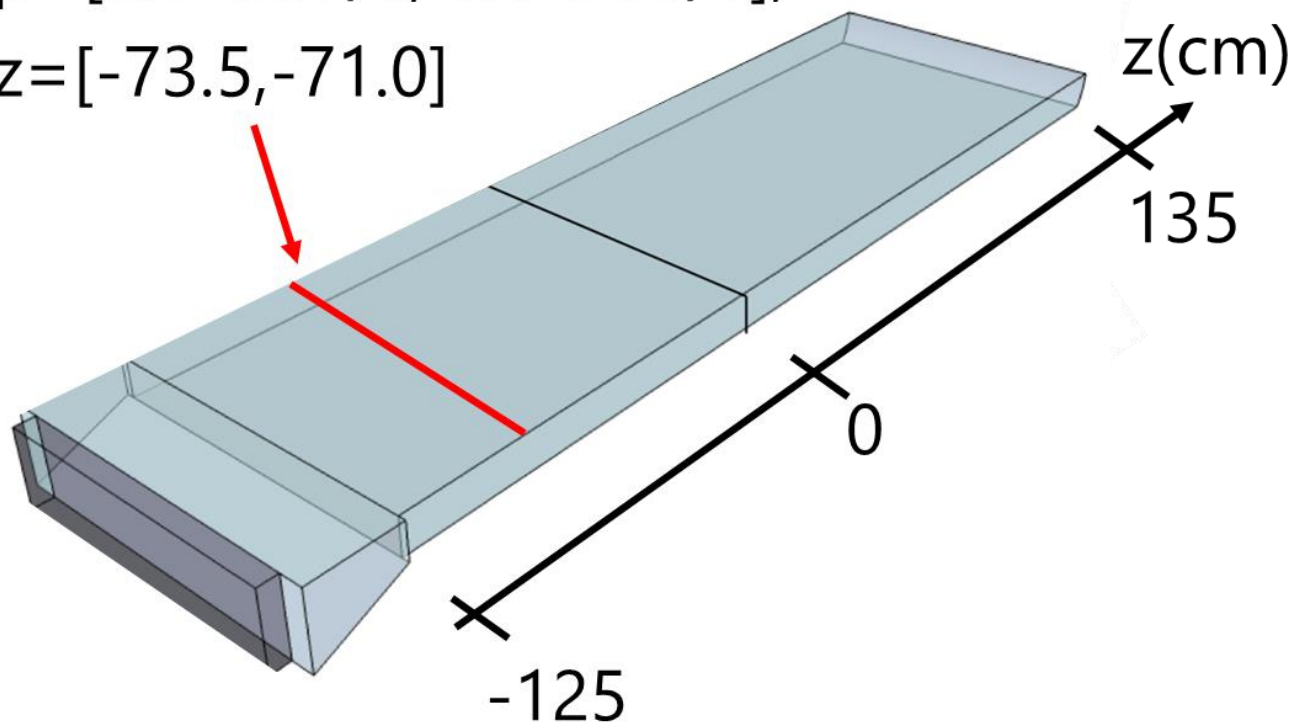
現行の手法との識別性能の比較

9

学習領域(入射飛跡の制限)

$p=[2.9 \text{ GeV}/c, 3.0 \text{ GeV}/c]$,

$z=[-73.5, -71.0]$



- ◆ NNを用いた学習において上図の学習領域を設定し、学習・評価を行った
- ◆ NNを用いた手法はこの学習領域において現行の手法に比べて高い性能を示す

学習領域の大きさ

- ◆ 一定の学習サンプル数においては、 z ・運動量ともに学習領域が大きくなるのに伴い性能が低下する
 - ◆ そのため、複数の学習領域が必要となるが、以下の問題が生じる
 - 学習領域が小さい場合: ネットワーク数増大による実用上のメモリの制限
 - 学習領域が大きい場合: 学習サンプル数増大による学習中のメモリの制限
- ⇒ 学習領域の大きさに対する依存性を調べ、基本設定を決定

学習領域:

$\Delta z = 5$ ($-5 < z < 0$ cm)

$\Delta z = 10$ ($-10 < z < 0$)

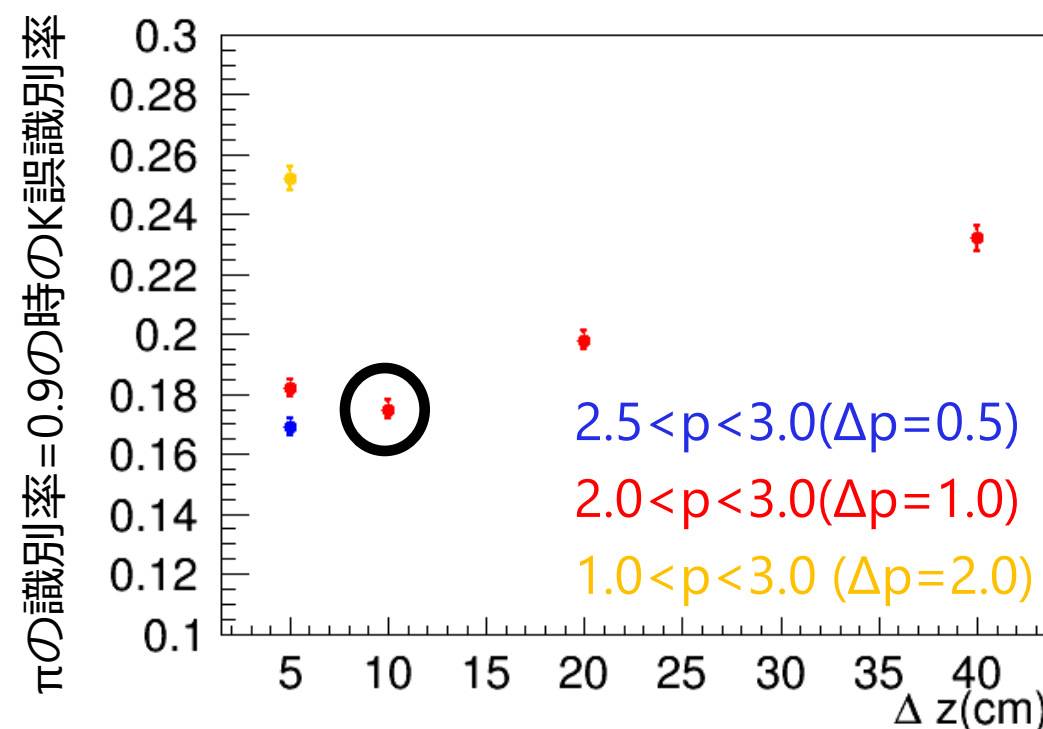
$\Delta z = 20$ ($-20 < z < 0$)

$\Delta z = 40$ ($-40 < z < 0$)

テスト領域:

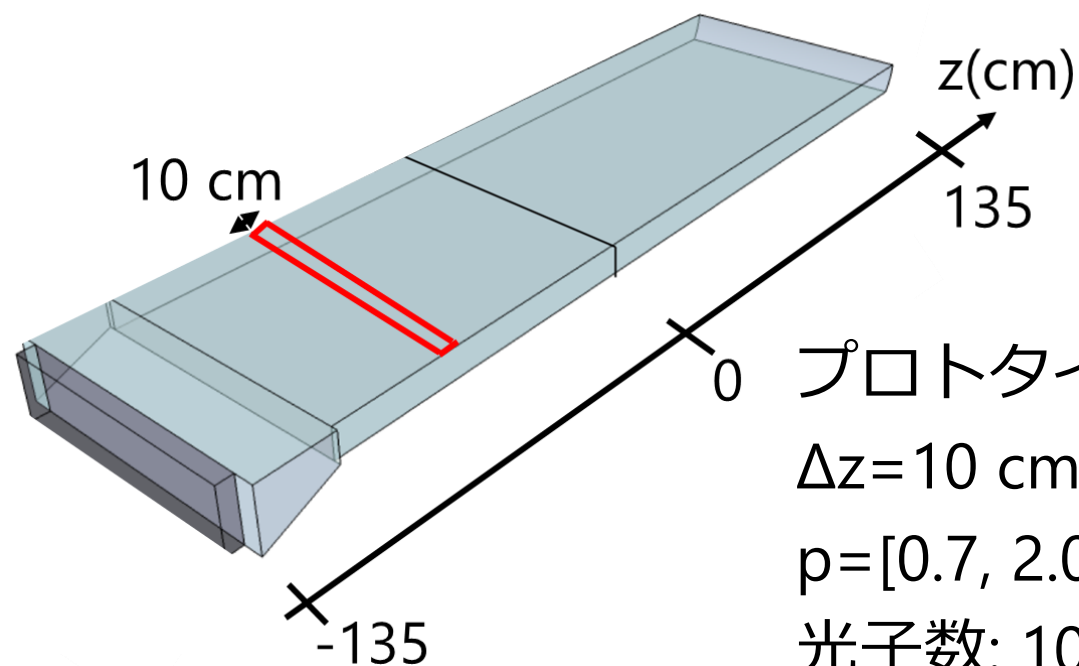
$p = 2.75$ GeV/c,

$x = 5.2$ cm, $z = -2.5$ cm



※運動量
スキャンは
 $\Delta z = 0.5$ のみ

- ◆ 識別性能を考慮すると $(\Delta z, \Delta p) = (5, 0.5)$ or $(10, 1.0)$
 - ⇒ ネットワーク数比 $(5, 0.5) : (10, 1.0) = 4:1$ より $(\Delta z, \Delta p) = (10, 1.0)$ を採用



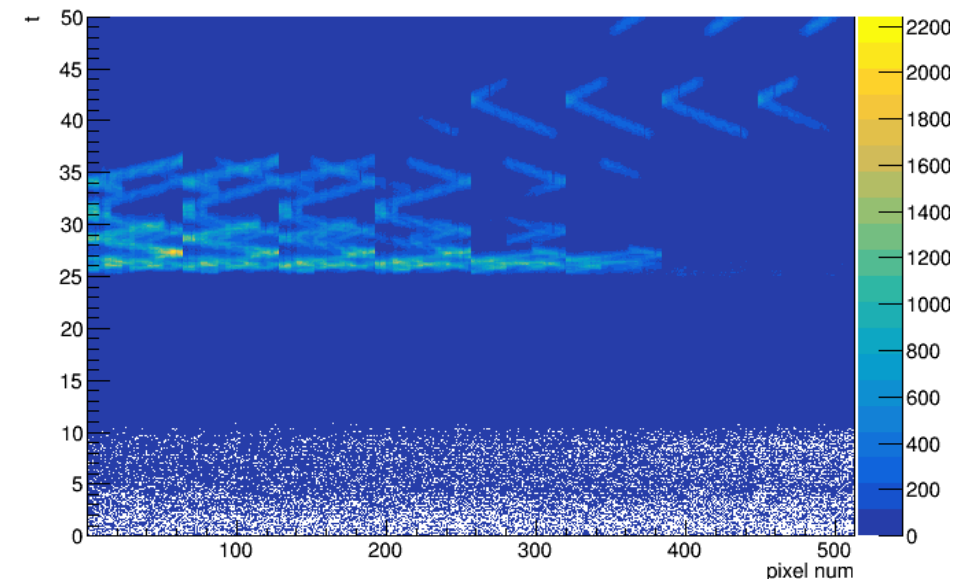
プロトタイプの基本設定:

$\Delta z = 10 \text{ cm}$

$p = [0.7, 2.0 \text{ GeV}/c]$, $p = [2.0, 3.0 \text{ GeV}/c]$

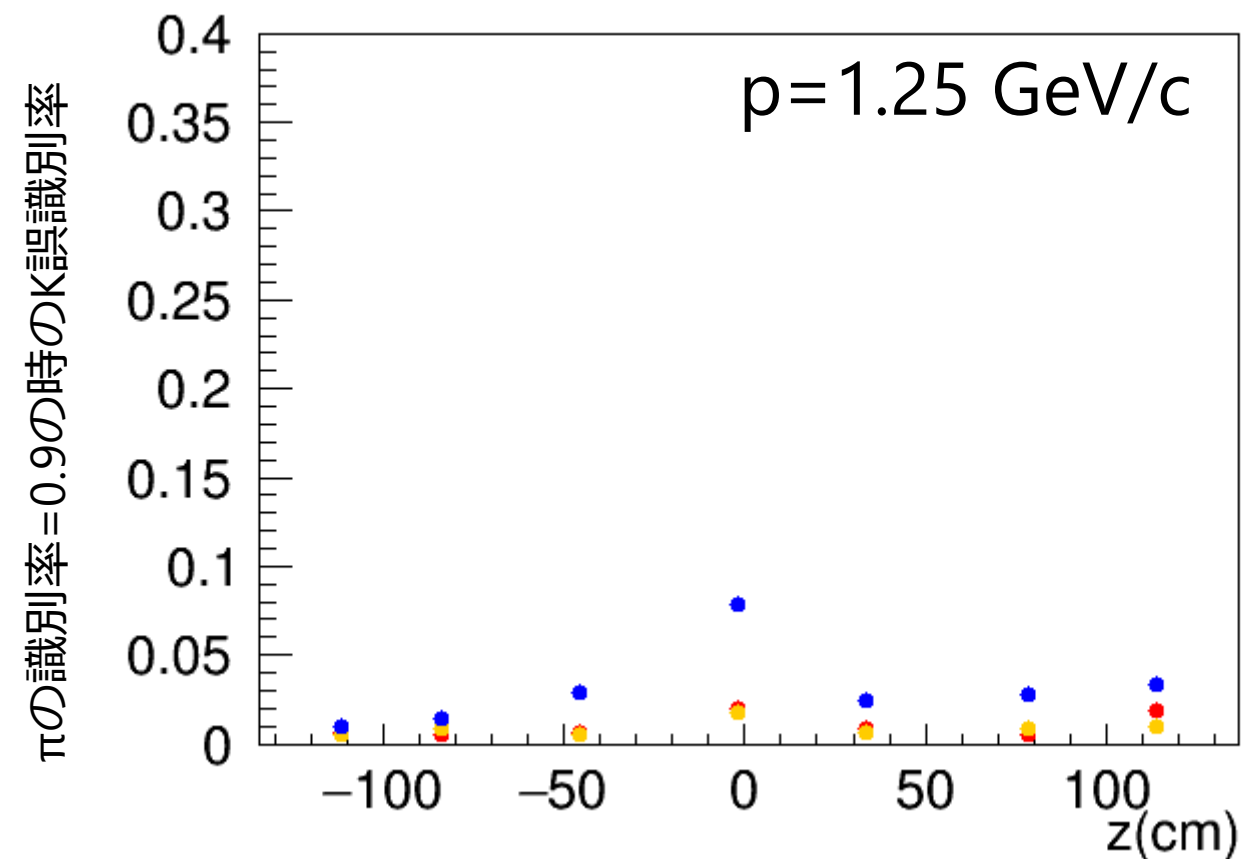
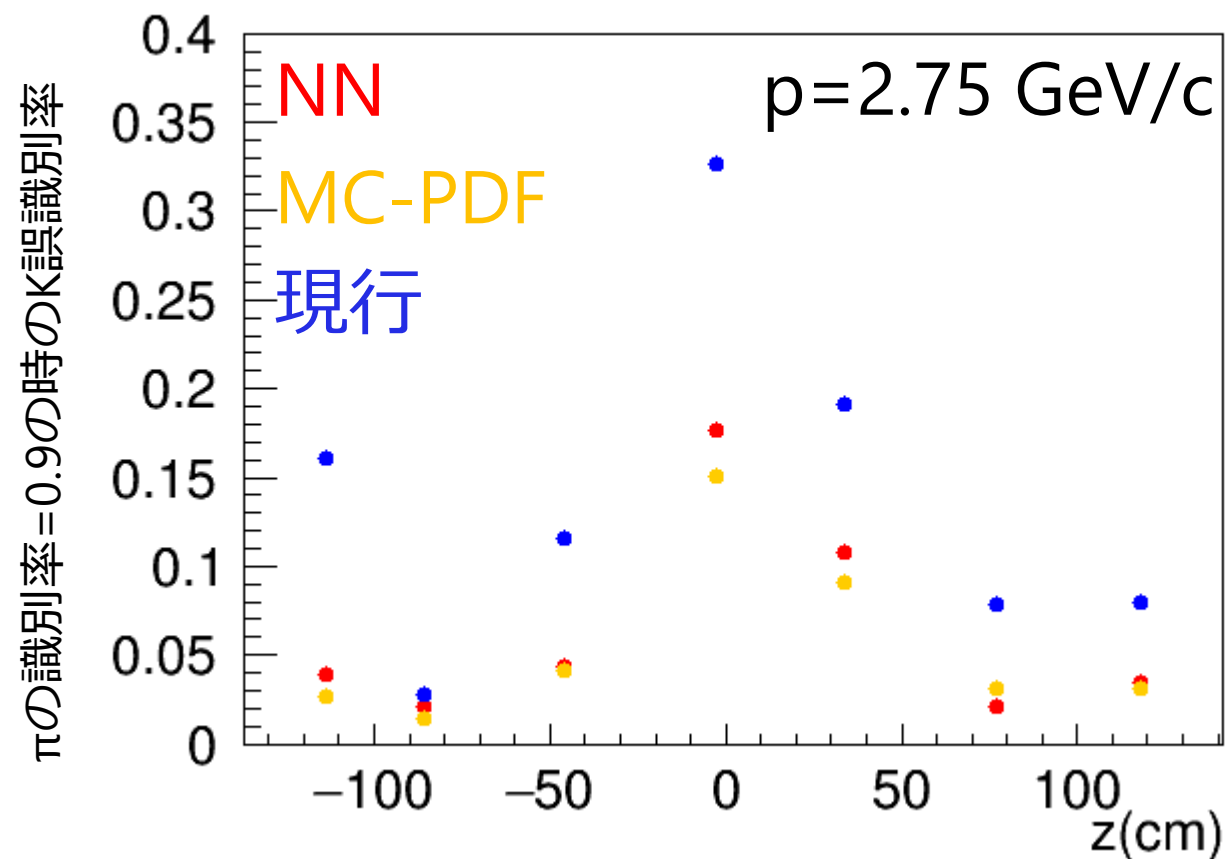
光子数: 10,000,000 (π/K それぞれ)

エポックサイズ: 5000



↓
ヒストグラムを基にした確率密度
から確率密度関数(MC-PDF)を導出

- ◆ 1領域の結果をもとに基本設定を決定したが、飛跡の $z \cdot$ 運動量によって識別性能は大きく変化するため、学習領域の大きさに対する識別性能の依存性も変化する可能性がある
- ◆ そこで、高運動量($p = 2.75 \text{ GeV}/c$)と低運動量($p = 1.25 \text{ GeV}/c$)において z 方向のスキャンを行い性能を評価する
- ◆ その際に現行の手法だけでなく、シミュレーションで生成した理想的な確率密度を用いた手法(MC-PDF)とも性能を比較する



高運動量領域

学習: $p = [2.0, 3.0 \text{ GeV/c}]$, $\Delta z = 10 \text{ cm}$

テスト: $p = 2.75 \text{ GeV/c}$, x, z : 固定点

低運動量領域

学習: $p = [0.7, 2.0 \text{ GeV/c}]$, $\Delta z = 10 \text{ cm}$

テスト: $p = 1.25 \text{ GeV/c}$, x, z : 固定点

- ◆ NNによる粒子識別性能は現行の手法に比べて高い性能を示しており、MC-PDFとも多くの領域で同等の性能を示している
- ◆ 高運動領域の $z = 0 \text{ cm}$ 付近はMC-PDFと差があるため、まだ改善の余地があると考えられる
- ◆ 今後、プロトタイプの全領域をカバーし、現行の手法とより詳細な比較を行っていく

まとめ

- ◆ TOPカウンター におけるニューラルネットワークを用いて粒子識別を行う新たなアルゴリズムを考案した
- ◆ 識別性能への影響を考慮した結果、x方向は全領域、z・運動量領域は制限した複数のネットワークを採用した
- ◆ シミュレーションデータを用いた、運動量2.75 GeV/cと1.25 GeV/c の複数の領域における評価の結果、NNによる識別手法が現行の手法に比べて高い識別性能を示した

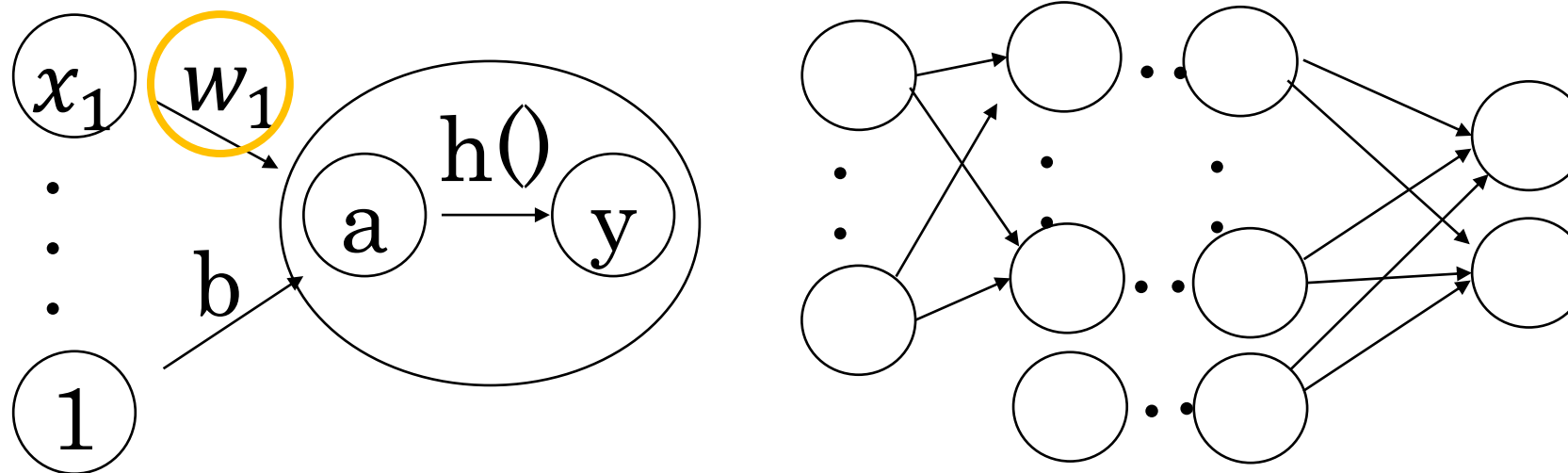
展望

- ◆ プロトタイプ全領域における現行の手法との詳細な比較
- ◆ 衝突データを用いた識別性能の評価

Backup

ハイパーパラメータ

- ◆ 損失関数: 基本的に与えられた正解(タグ)の値に近づくほど小さくなる関数で、これを全体として最小化することで学習を行う
- ◆ バッチサイズ: 損失関数の和を計算する際に使用するデータ数
- ◆ epoch数: バッチサイズ分の学習を行った数
- ◆ 隠れ層の数・ニューロン数
: ネットワーク構造の複雑さを表す
増やすほど複雑な表現が可能になるが、
過学習(学習データへの過度の適応)の原因にもなる
- ◆ Optimizer: 損失関数を最小化する際に用いられる手法。学習の収束の速さや過学習のしやすさに影響。



- ◆ NNのパラメータとして学習結果として保存されるのは重みとバイアスである。
- ◆ それらはhdf5ファイルとして保存され、そのサイズ隠れ層(100×3)の構造において ~ 0.1 MBである。
- ◆ メモリサイズの許容量としては $O(\sim 100 \text{ MB})$ が想定され、その場合 $O(\sim 1000)$ のネットワーク数が制限となる

- ◆ π 識別効率 = $\frac{\pi\text{粒子と識別された}\pi\text{粒子数}}{\text{すべての}\pi\text{粒子数}}$
- ◆ K 誤識別率 = $\frac{\pi\text{粒子と識別された}K\text{粒子数}}{\text{すべての}K\text{粒子数}}$