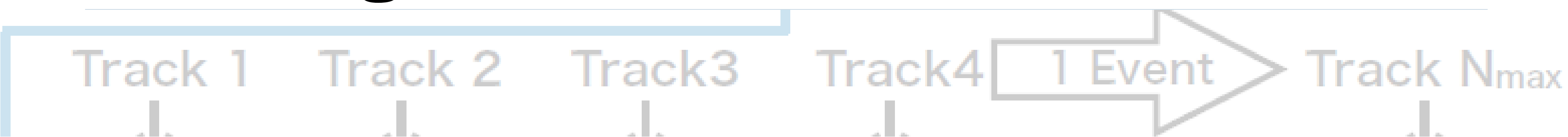


Development of a Vertex Finding Algorithm using Recurrent Neural Network

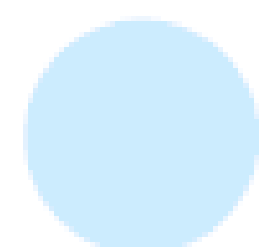


Kiichi Goto, Taikan Suehara, Tamaki Yoshioka (Kyushu U)

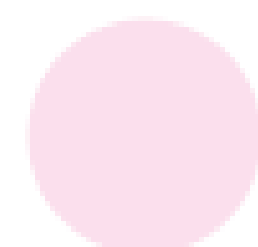
Masakazu Kurata (Kyushu → Tokyo → KEK)

Hajime Nagahara, Yuta Nakashima, Noriko Takemura (IDS, Osaka-U)

Masako Iwasaki (Osaka CU / Osaka U)



Custom LSTM



Custom Attention LSTM

Contents

1. Motivation – LCFIPlus and flavor tagging
2. Network structure for vertex finding
3. Performance evaluation
 - Accuracy of the network
 - Performance of vertex finding – comparison with LCFIPlus
 - Evaluation of the network within Marlin framework
 - Performance of flavor tagging – comparison with LCFIPlus
4. Summary and Prospects

Source codes:

<https://github.com/Goto-K/VertexFinderwithDL> (python part)

<https://github.com/Goto-K/LCFIPlus> (adaptation to LCFIPlus)

LCFIPlus and flavor tagging

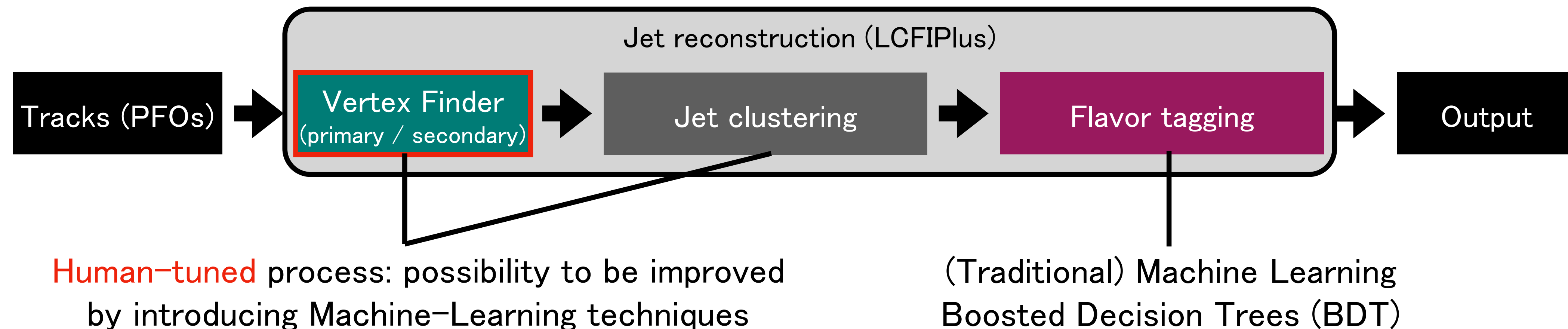
Structure of LCFIPlus

LCFIPlus paper: <https://doi.org/10.1016/j.nima.2015.11.054>

LCFIPlus: Standard flavor tagging software for ILD (also used in SiD, CLICdp, ...)

Modular structure to accommodate various algorithms for jet reconstruction

- ▶ Vertex finder (primary: tear-down / secondary: build-up)
- ▶ Jet clustering (Durham / Valencia-like / K_T) using vertex information; beam-jet rejection incorporated
- ▶ Jet vertex refiner (Tuning vertices with jet information; association of vertices to jet when external jet clustering used)
- ▶ Flavor tagging (b/c/uds)



This work: **replace Vertex Finder with Deep-Learning (DL) networks** as a first step for replacing all jet reconstruction with DL technologies

Vertex finding and simulation conditions

Vertex finding in LCFIPlus

- Build-up method (used for secondary vertices after removing primary/ V_0 tracks)
 1. Produce a vertex by each track pair (from all tracks, $O(n^2)$ combinations)
 2. Select vertices with good quality (cut on χ^2 , mass, direction, etc.)
 3. Associate additional tracks to the selected vertices (with χ^2 criteria)
 4. Associate primary tracks with comparison of χ^2 with primary and selected vertices

For DL-based algorithm,
build-up like method is considered for the network structure

Two neural networks for the DL-based vertex finder

1. Network for selecting track pairs as “vertex seeds”

A Simple feed-forward network currently used

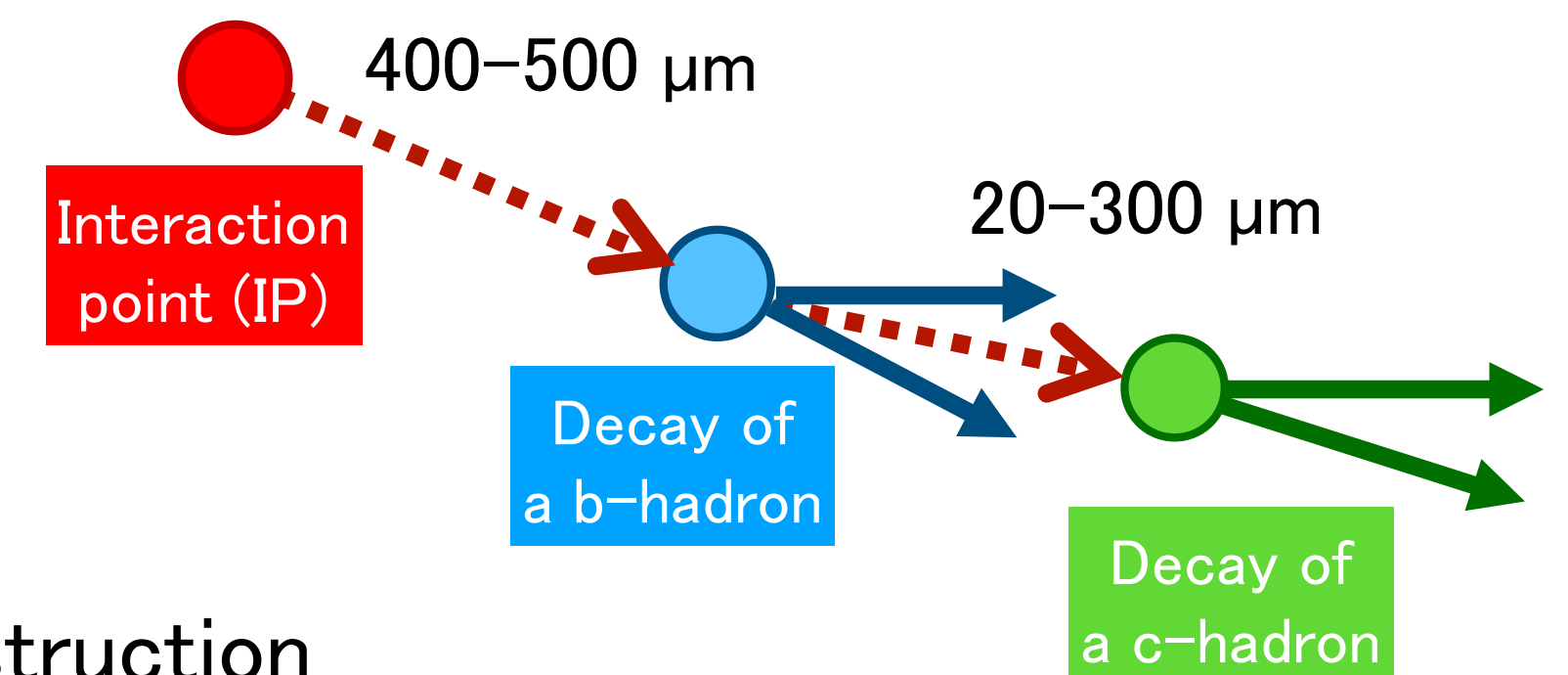
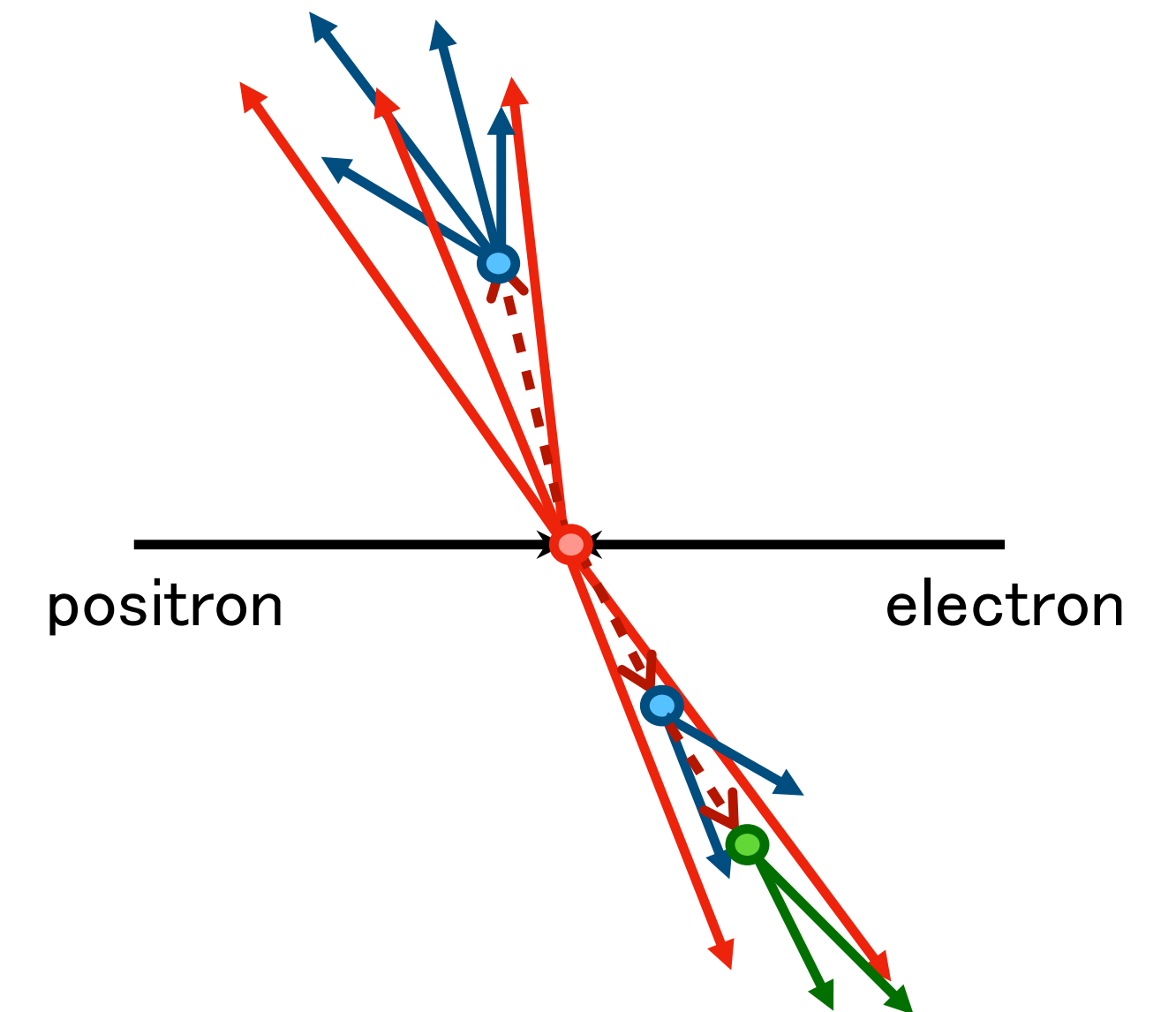
2. Associate tracks to vertex seeds

Recurrent-type neural network is employed

Simulation conditions of this study

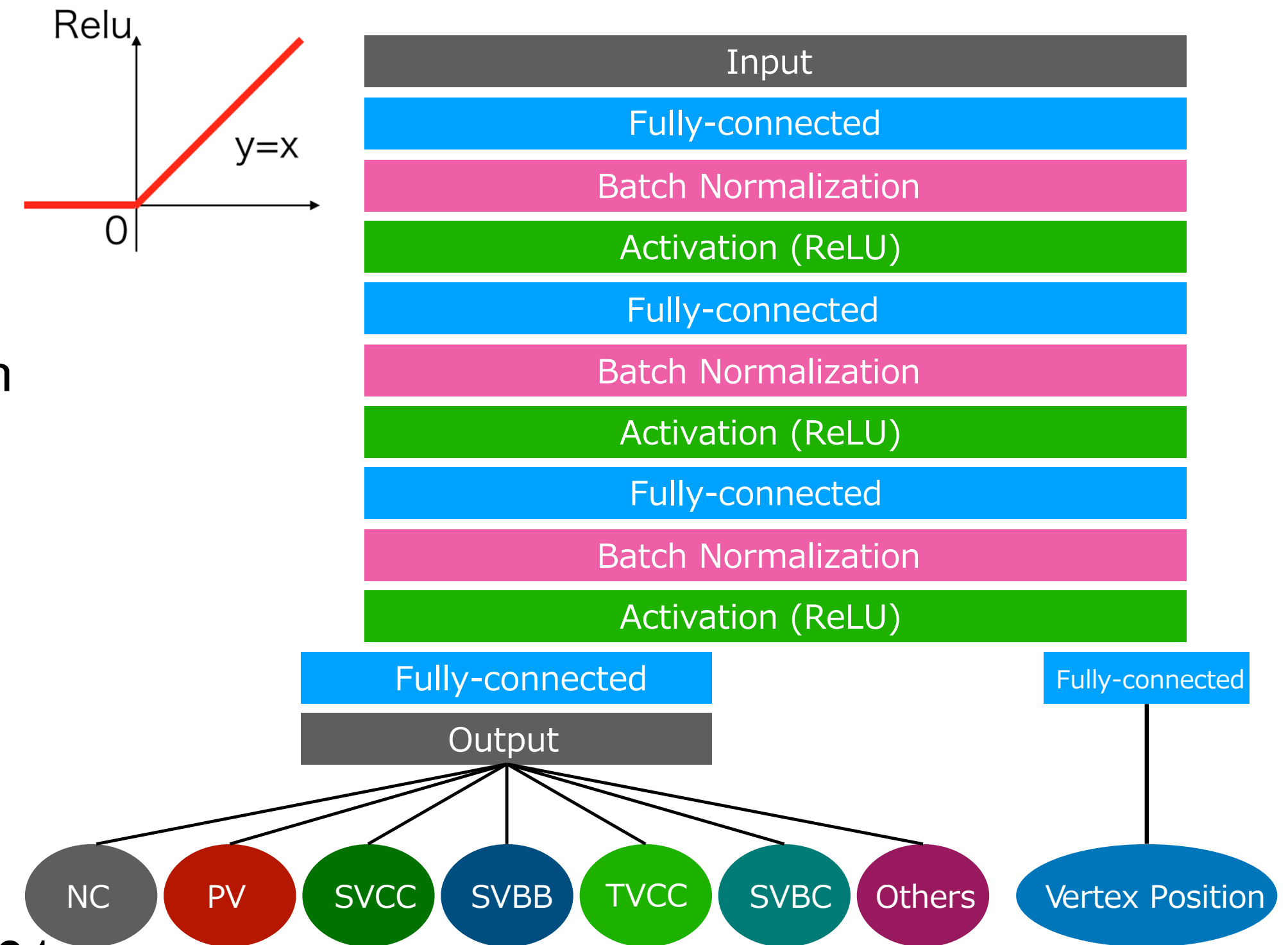
- ILD DBD simulation (for comparison with LCFIPlus) / DBD standard reconstruction
- $e^+e^- \rightarrow qq$ ($q = b, c, uds$) at 91 GeV CM energy, $\sim 500k$ events each (events divided to be used in training and evaluation)

Decay vertices in a typical event

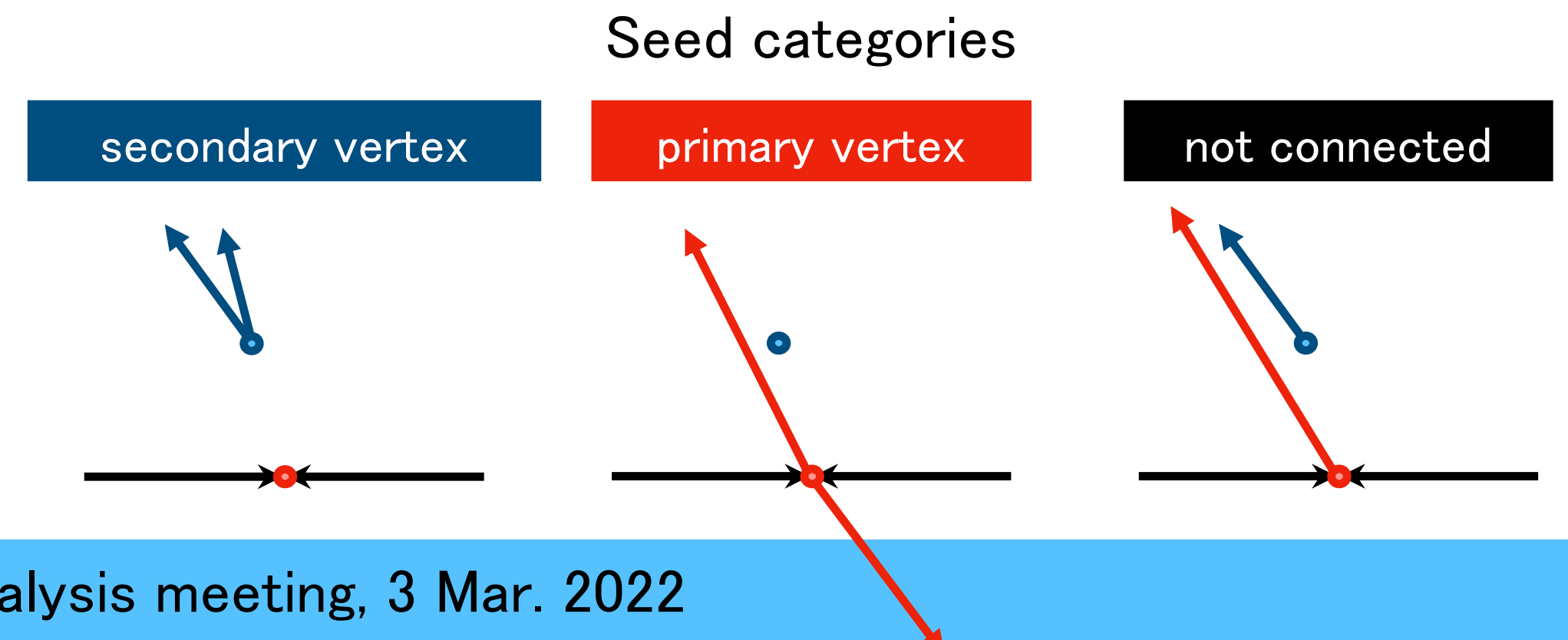


Network design 1: selecting “vertex seeds”

- Simple feed-forward fully-connected network
- Input: parameters of 2 tracks (total 44 params)
 - Helix parameters ($d_0, z_0, \phi, \tan\lambda, \Omega$)
 - Covariance matrix (15 params)
 - Charge and energy
- 3 fully-connected layer with batch normalization and ReLU activation
- 7 categories for output after final fully-connected layer
 - NC, PV, SVCC, SVBB, TVCC, SVBC, others
- Regression of vertex position with separate fully-connected layer
 - To build “position recognition” algorithm inside the main layers
- Loss function tuned to train categorization and position network
 - 1st step: $w(\text{cat}, \text{pos}) = (0.1, 0.9)$, 1000 epoch, learning rate = 0.001
 - 2nd step: $w(\text{cat}, \text{pos}) = (0.9, 0.1)$, 1500 epoch, learning rate = 0.001
 - 3rd step: $w(\text{cat}, \text{pos}) = (0.95, 0.05)$, 500 epoch, learning rate = 0.0001
- PV and SV/TVxx categories are used for “vertex seeds”



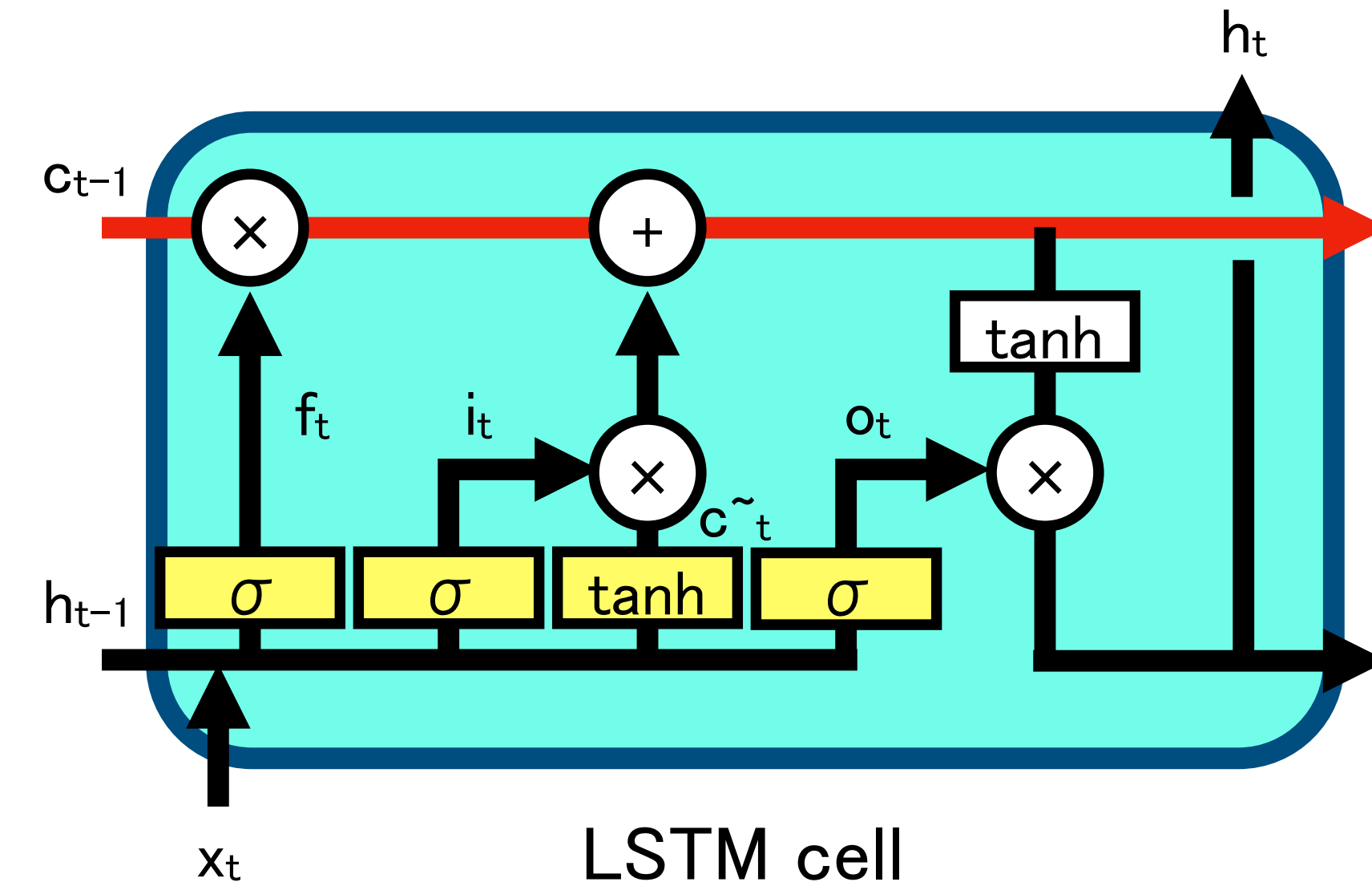
PV: both tracks from primary vertex
 NC: track coming from different vertex
 SVBB: both tracks from b hadrons in $e^+e^- \rightarrow bb$ samples
 TVCC: both tracks from c hadrons in $e^+e^- \rightarrow bb$ samples
 SVBC: one track from b, the other from c in bb samples
 TVCC: both track from c hadrons in $e^+e^- \rightarrow cc$ samples
 Others: tracks coming from V_0 or other vertices



Recurrent Neural Network (RNN) and variance

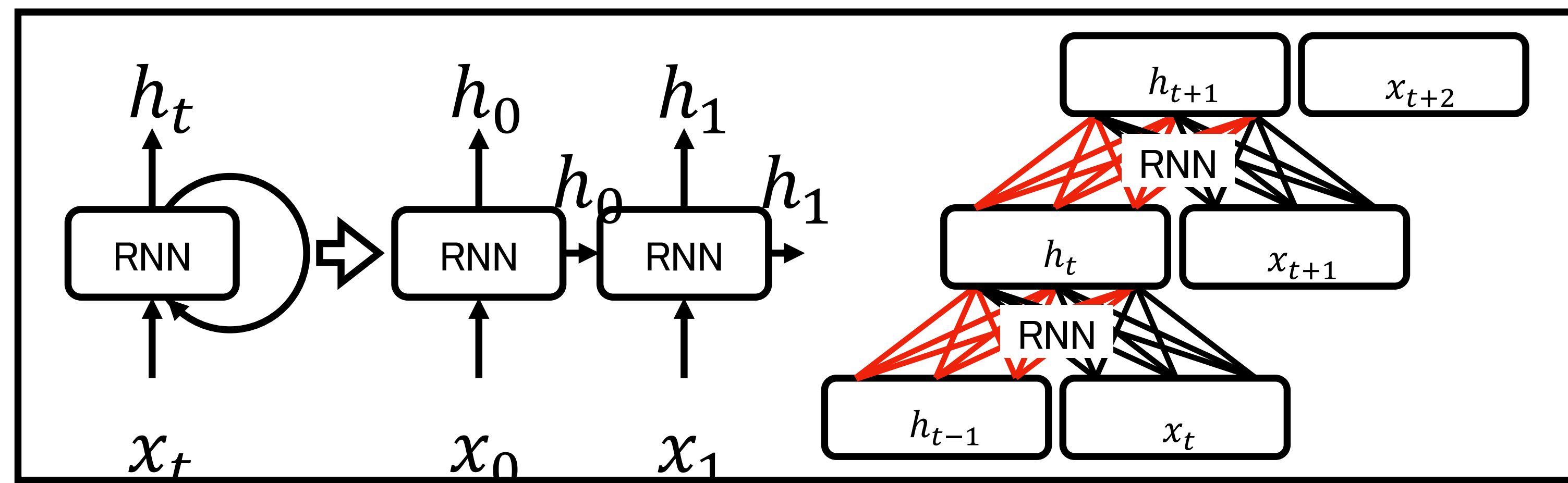
➤ Recurrent Neural Network

- Neural network designed for variable input length
- Main application: natural language processing (translation etc.)
- “RNN cell” defines a unit of network structure (with learning weights)
- Each input (x_t) is processed sequentially with the same RNN cell (and same weights)
- “Hidden states h_t ” are also inputs of the next cell
- Problem on “gradient loss / gradient explosion”
→ various RNN cell structures are proposed



➤ LSTM (long short term memory)

- One of the RNN cell structure practically used
- “Gate” structures to avoid gradient loss/explosion
 - forget gate
 - input/output gate
- Short-term memory to retain relations to neighbor inputs

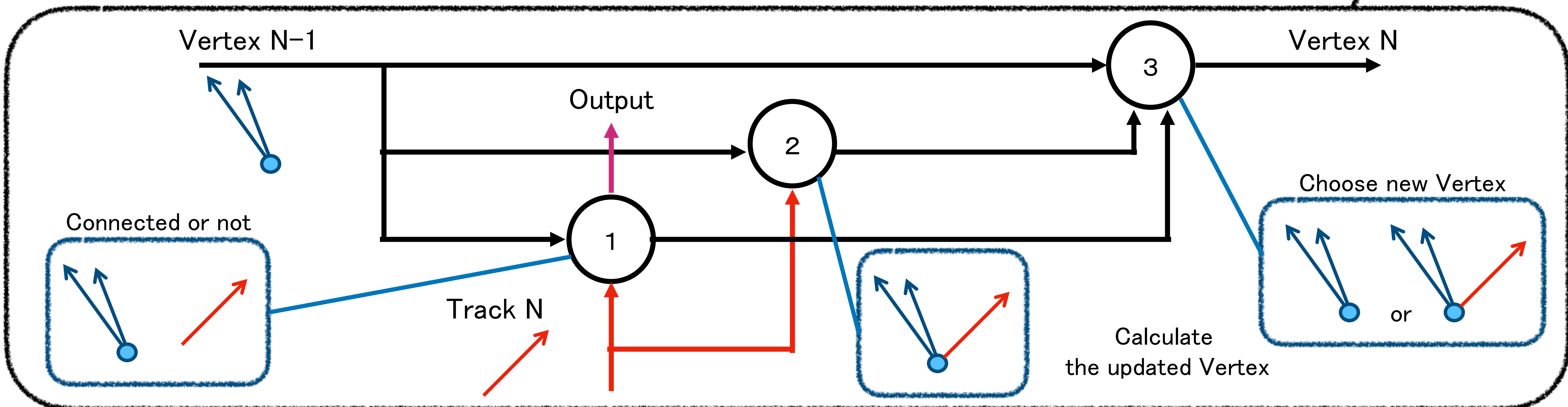
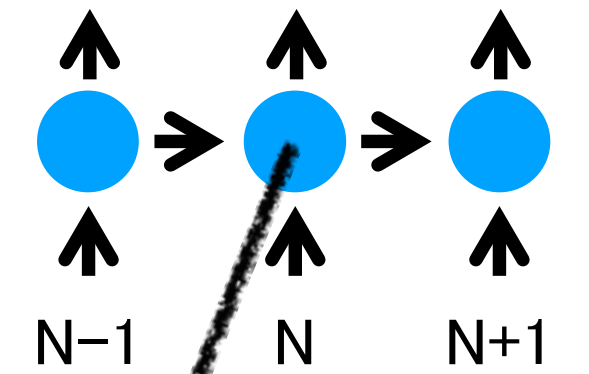


Network design 2: associate tracks to vertices

Custom RNN structure modified from plain LSTM

- No importance in the order of tracks in the track association \rightarrow “short-term memory” is not necessary
- The custom cell only propagates “long-term memory” which is recognized as “vertex information” to the next cell
- “forget, input and output gates” structures are kept to avoid gradient loss / explosion
- Procedures:
 1. Evaluate if the n-th track should be associated or not: $h_N = \sigma(D_h[\sigma(W_o t_N + R_o V_{N-1}) \cdot \tanh(V_{N-1})])$
 2. Combine the n-th track and the n-1 th vertex to produce a new vertex:

$$U_N = \sigma(W_i t_N + R_i V_{N-1}) \cdot \tanh(W_z t_N + R_z V_{N-1}) + \sigma(W_f t_N + R_f V_{N-1}) \cdot V_{N-1}$$
 3. Combine the old and new vertex according to the evaluation in the 1st step: $V_N = (1 - h_N)V_{N-1} + h_N U_N$



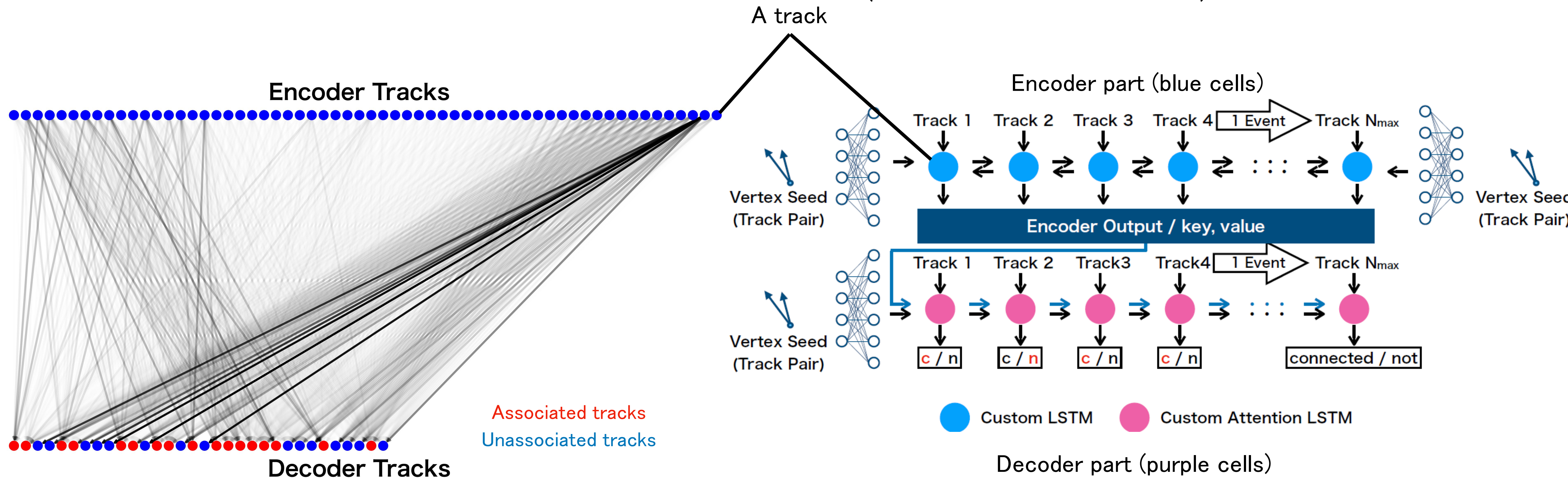
Network design 2: additional features

Attention mechanism

- State-of-the-art scheme of machine learning to specify “attention” to certain elements of the network
- Usually used in encoder-decoder models

Encoder-decoder model

- Encoder: making array of “hidden states” forming storage of information (“vertex” in our case) at the output
- Decoder: derive the essential information from the encoder output (in our case “associated” or not)

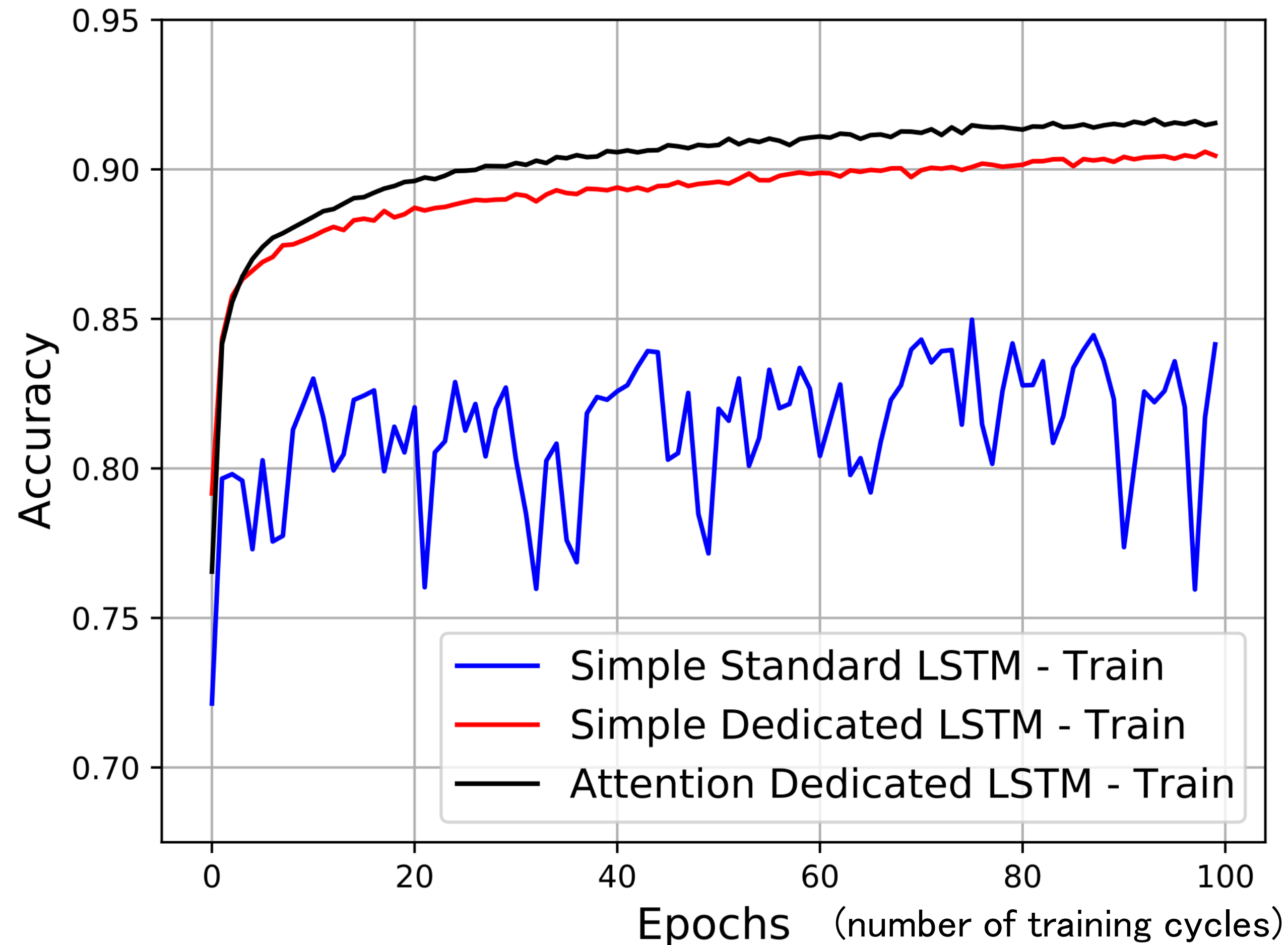


Performance of the custom network

Comparison to standard structure

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP: true positive
TN: true negative
FP: false positive
FN: false negative



Improvement seen by using custom LSTM structure (red) and attention (black)

Vertex Finder

Algorithm for Vertex Finder

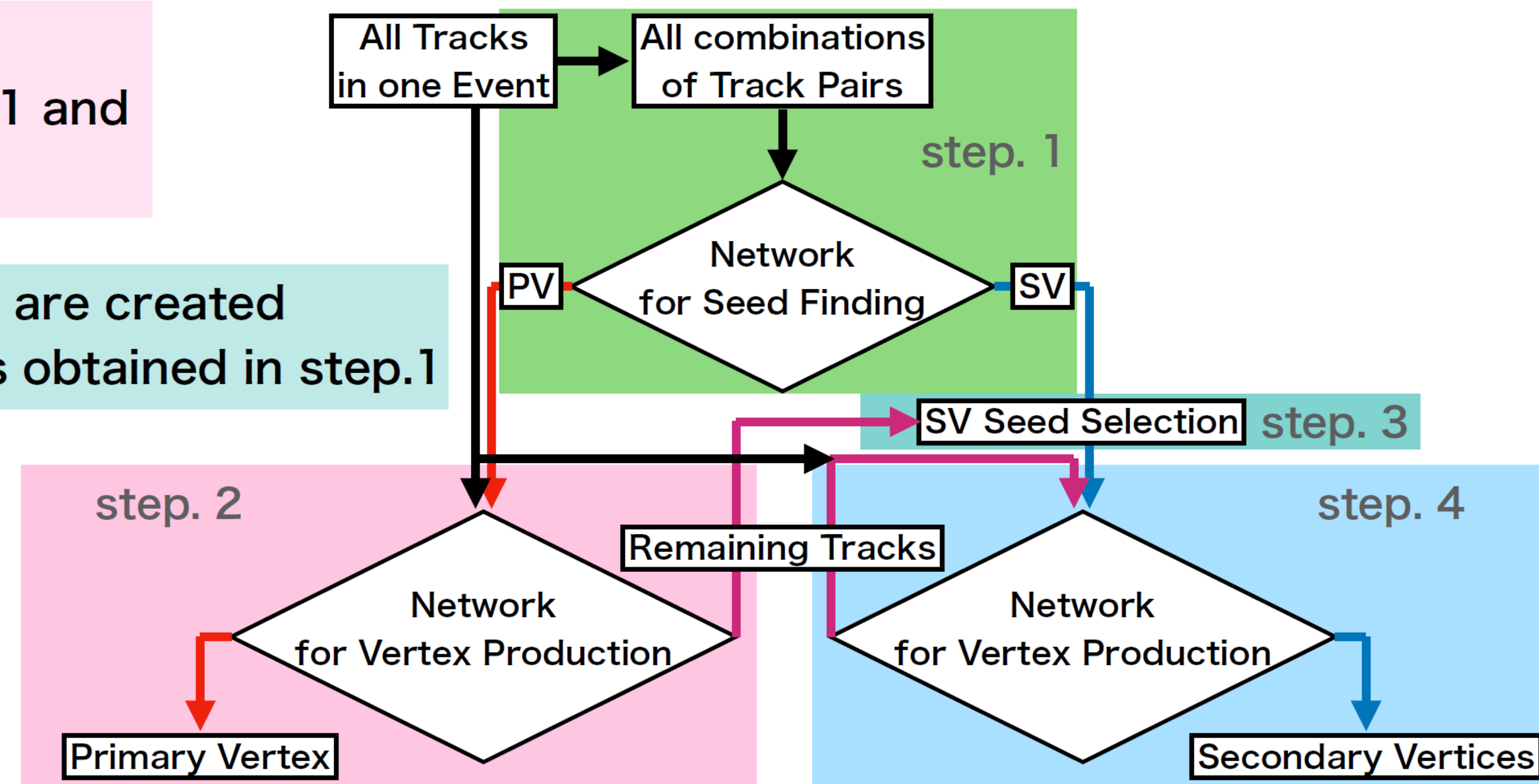
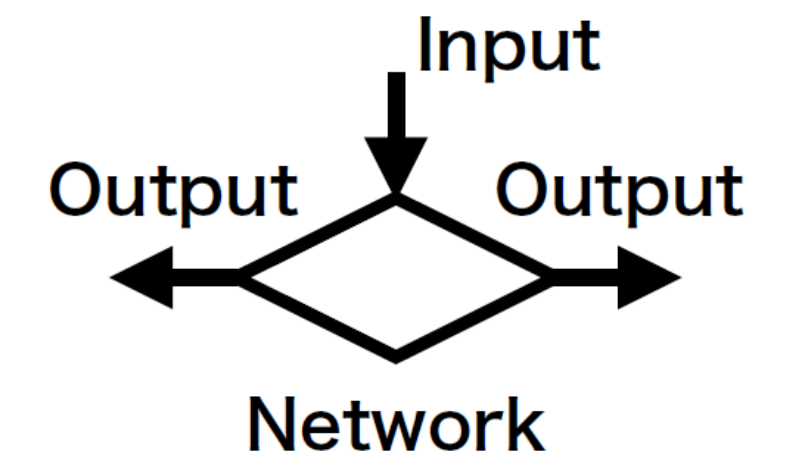
- Finding the vertices using following steps

1. Considering all combinations of two tracks in a event, the vertex seeds are searched by “network for seed finding”

2. The primary vertex is created using the seeds of primary vertex obtained in step.1 and “network for vertex production”

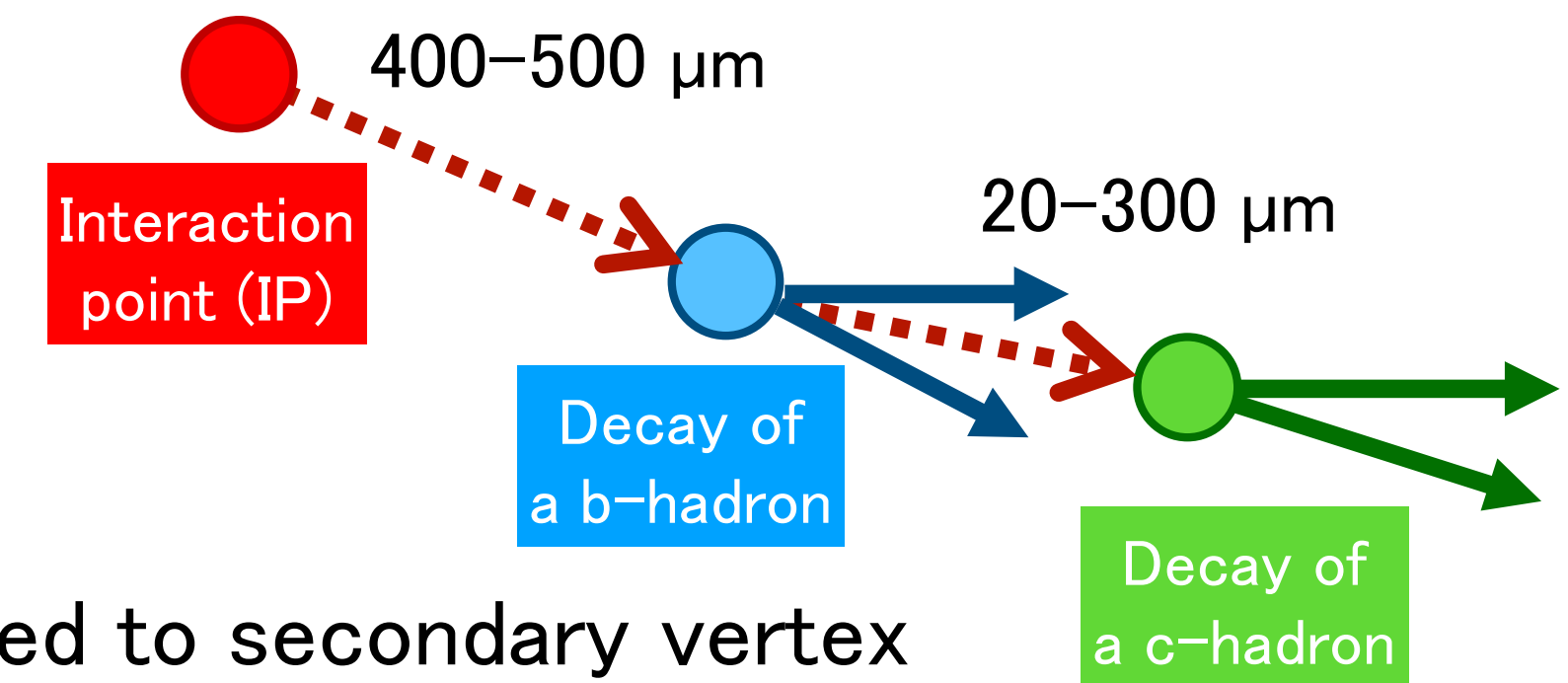
3. The purer set of seeds of secondary vertices are created by screening the seeds of secondary vertices obtained in step.1

4. The secondary vertices are created using the seeds of secondary vertices selected in step.3 and “network for vertex production”



Performance of the DL-based vertex finder

Comparison with LCFIPlus (track-by-track criteria) with **bb** samples



- True label

1. **Primary** – tracks with no (semi)stable parents
2. **Bottom** – tracks originated from b-hadrons
3. **Charm** – tracks originated from c-hadrons
4. **Others** – other tracks (mainly V_0 tracks)

- Criteria

1. In secondary vertex – associated to secondary vertex
2. – of same decay chain – **all tracks in the vertex** from **the same b parent**
3. – of same parent – **all tracks in the vertex** from the same immediate parent (ie. **success of b-c separation**)

Performance of DL-based vertex finder

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	2.2%	63.3%	68.4%	9.5%
– of same decay chain		62.3%	67.2%	
– of same parent		38.1%	36.2%	6.4%

Performance of LCFIPlus vertex finder

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	0.2%	57.9%	60.3%	0.5%
– of same decay chain		57.5%	59.9%	
– of same parent		34.0%	37.2%	0.3%

- 5–10% higher efficiency on the reconstruction of secondary vertices
- More contamination of primary and other tracks (need additional selection on track quality etc.)

Summary and Prospects

Summary

Source codes:

<https://github.com/Goto-K/VertexFinderwithDL> (python part)

<https://github.com/Goto-K/LCFIPlus> (adaptation to LCFIPlus)

- We developed a vertex finding algorithm based on modern deep neural networks.
- Track association done with customized RNN-type network with attention mechanism.
- Efficiency of the reconstruction of secondary vertices is improved, while mis-reconstruction of primary / other tracks to secondary vertices is increased.

DL-based vertex finder

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	2.2%	63.3%	68.4%	9.5%
– of same decay chain		62.3%	67.2%	
– of same parent		38.1%	36.2%	6.4%

LCFIPlus vertex finder

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	0.2%	57.9%	60.3%	0.5%
– of same decay chain		57.5%	59.9%	
– of same parent		34.0%	37.2%	0.3%

Prospects

- Improvement of the vertex finder
 - More tuning of the “seed finding” network, using more appropriate network to use “crossing point” of two tracks
 - Including more physical properties to RNN network as well as improving structure
- Development of DNN-based flavor tagging
 - Including low-level information (tracks)
 - **Combining vertex finder and flavor tagging by Graph Neural Network – under development**

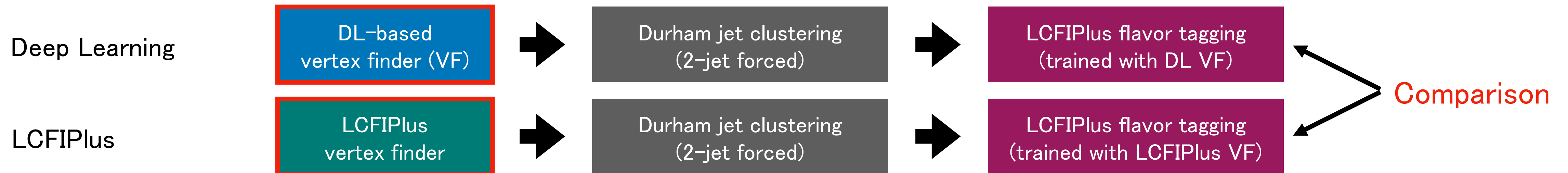
This work is supported by RCNP (Osaka U.) Project
“Application of deep learning to accelerator experiments”.

Backup

Performance of the flavor tagging (FT)

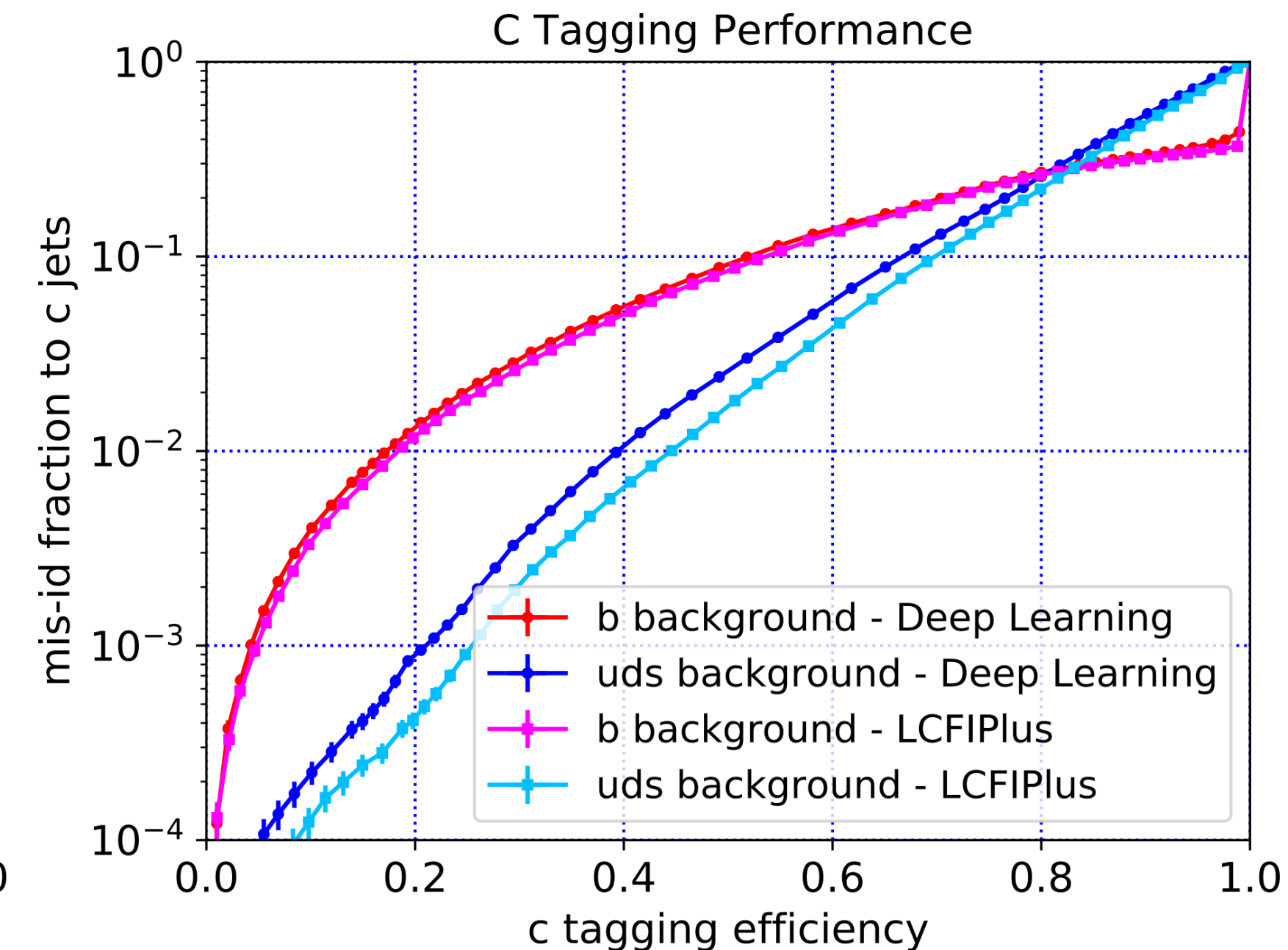
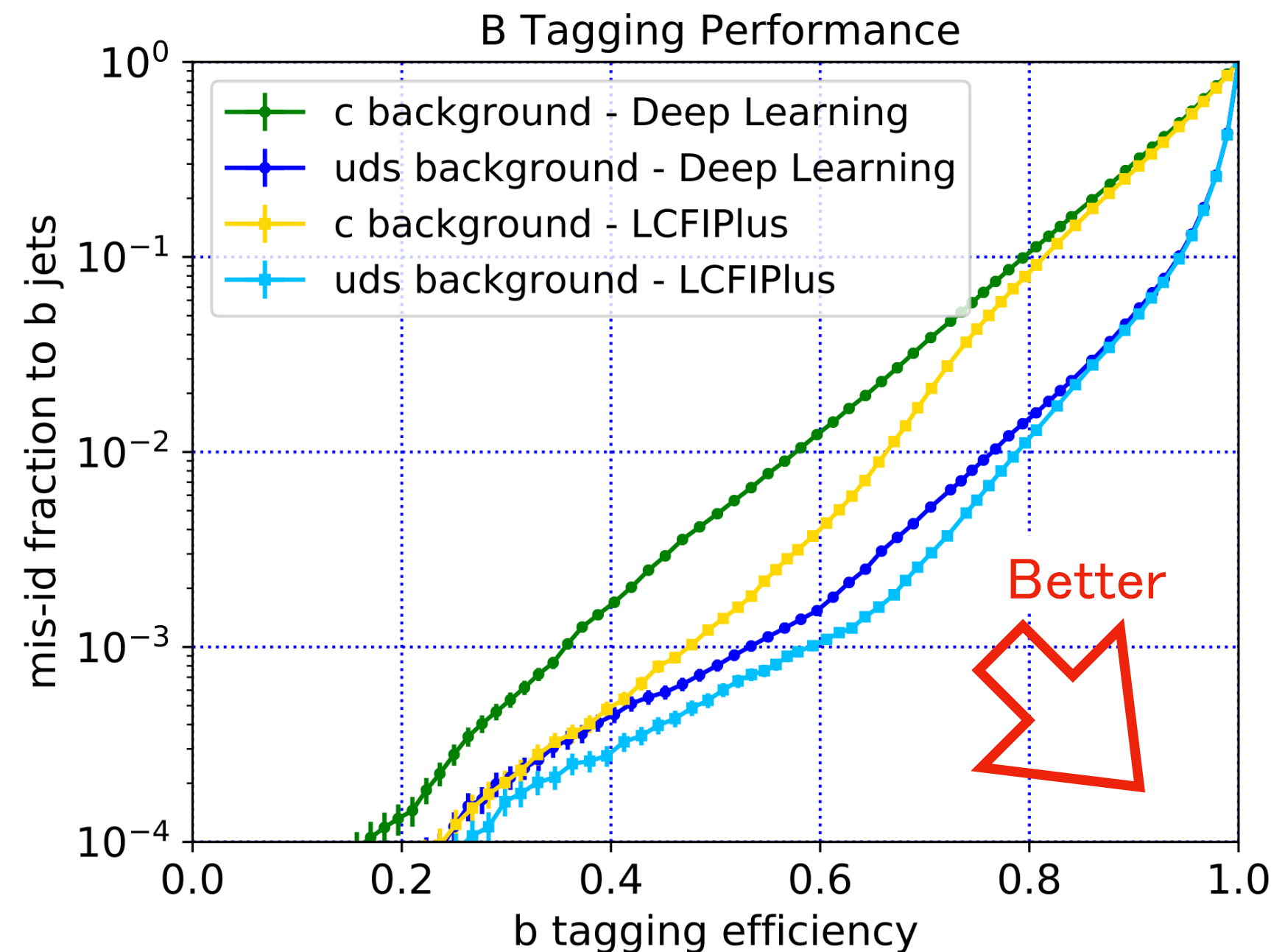
Procedure for flavor tagging

- Replace vertex finder and use other algorithm as same as LCFIPlus' case (with same parameter)



- The performance of **b-tagging** of LCFIPlus cannot be reproduced with DL vertex finder
 - Similar performance in **c-tagging**
 - Probably due to contamination of primary tracks
 - Tuning of parameters / input variables are highly optimized with LCFIPlus → some bias on LCFIPlus
- DL-based vertex finder has an advantage of possibility of closer connection to flavor tagging algorithm
 - “organic” connection of networks possible if FT fully written in DNN

→ FT algorithm to be rewritten with DNN



Adaptation to C++ / LCFIPlus / ILCSoft (Marlin)

Method for inference (evaluation) in C++

- Tensorflow/Keras is used for building/training the network
 - Fully python (version 3)
 - We obtain input with LCIO → ROOT tree → NumPy conversion
- LCFIPlus is running as a Marlin processor, fully C++
 - For comparison of the flavor tagging, the output vertices should be in LCIO or LCFIPlus format.
- Keras is provided only in python, but Tensorflow has official C++ implementation
 - This is one reason we chose Tensorflow as the framework (while PyTorch only has beta implementation on C++ port).
 - Inference (evaluation of the network) can be done without Keras, thus possible to run in C++.
- We introduce VertexFinderwithDL algorithm inside LCFIPlus (thus possible to be called from Marlin)
 - Tensorflow and bazel (as dependency) are needed to be installed
 - Can run both with GPU and without GPU (cuda / cuDNN necessary for GPU run)
 - Results have compared with python version; identical result obtained
 - Output vertices are compatible with LCFIPlus output
 - Vertex fitting (to obtain position, χ^2 etc.) is done using LCFIPlus functions after selecting tracks with DL networks.

Training in python
Inference in C++

CMake results

```
-- Found LCFIVertex: /gluster/data/ilc/ilcsoft/v02-02/LCFIVertex/v00-08
-- Found Tensorflow: /home/goto/local/include/tf ◀
-- Found Protobuf: ◀
-- Found Eigen3: /home/goto/local/include/eigen3 (Required is at least version "2.91.0") ◀
-- Check for ROOT_CINT_EXECUTABLE: /gluster/data/ilc/ilcsoft/v02-02/root/6.18.04/bin/rootcint
```

Adaptation to C++ / LCFIPlus / ILCSoft

```
2020-11-14 21:33:38.248302: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcudart.so.10.1
2020-11-14 21:33:38.248381: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcublas.so.10
2020-11-14 21:33:38.248490: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcufft.so.10
2020-11-14 21:33:38.248562: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcurand.so.10
2020-11-14 21:33:38.248660: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcusolver.so.10
2020-11-14 21:33:38.248731: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcusparsesolver.so.10
2020-11-14 21:33:38.248835: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcudnn.so.7
```

Libraries
for GPU

```
2020-11-14 21:33:38.251605: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1697] Adding visible gpu devices: 0, 1
2020-11-14 21:33:38.251752: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1096] Device interconnect StreamExecutor with strength 1 edge matrix:
2020-11-14 21:33:38.251874: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1102]      0 1
2020-11-14 21:33:38.251999: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1115] 0:   N Y
2020-11-14 21:33:38.252111: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1115] 1:   Y N
2020-11-14 21:33:38.254091: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1241] Created TensorFlow device (/job:localhost/replica:0/task:0/devi
0, name: TITAN RTX, pci bus id: 0000:81:00.0, compute capability: 7.5)
2020-11-14 21:33:38.255119: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1241] Created TensorFlow device (/job:localhost/replica:0/task:0/devi
1, name: TITAN RTX, pci bus id: 0000:c1:00.0, compute capability: 7.5)
```

Setup GPUs

```
2020-11-14 21:33:38.334339: I tensorflow/cc/saved_model/loader.cc:203] Restoring SavedModel bundle.
2020-11-14 21:33:38.446862: I tensorflow/cc/saved_model/loader.cc:152] Running initialization op on SavedModel bundle at path: /home/goto/ILC/Deep_L
6_50000samples_100epochs_ps_100epochs_s
2020-11-14 21:33:38.509773: I tensorflow/cc/saved_model/loader.cc:333] SavedModel load for tags { serve }; Status: success: OK. Took 292443 microsec
```

Restoring
model

```
2020-11-14 21:33:44.005873: I tensorflow/stream_executor/platform/default/dso_loader.cc:44] Successfully opened dynamic library libcublas.so.10
```

My Processor “Vertex Finder with DL” is running

```
[ MESSAGE "Marlin" ] ---- no GEAR XML file given -----
[ MESSAGE "VertexFindingwithDL" ]
[ MESSAGE "VertexFindingwithDL" ] ---- VertexFindingwithDL - parameters:
[ MESSAGE "VertexFindingwithDL" ]     Algorithms: VertexFindingwithDL
[ MESSAGE "VertexFindingwithDL" ]     IgnoreLackOfVertexRP: 0
[ MESSAGE "VertexFindingwithDL" ]     MCPCollection: MCParticlesSkimmed
[ MESSAGE "VertexFindingwithDL" ]     MCPFORelation: RecoMCTruthLink
[ MESSAGE "VertexFindingwithDL" ]     MagneticField: 3.5
[ MESSAGE "VertexFindingwithDL" ]     PFOCollection: PandoraPFOs
[ MESSAGE "VertexFindingwithDL" ]     PIDAlgorithmName: LikelihoodPID
[ MESSAGE "VertexFindingwithDL" ]     PrintEventNumber: 0
[ MESSAGE "VertexFindingwithDL" ]     ReadSubdetectorEnergies: 1
[ MESSAGE "VertexFindingwithDL" ]     TrackHitOrdering: 1
[ MESSAGE "VertexFindingwithDL" ]     UpdateVertexRPDaughters: 0
[ MESSAGE "VertexFindingwithDL" ]     UseMCP: 1
[ MESSAGE "VertexFindingwithDL" ] -----
```


3. Inference with C++

For Evaluation in LCFIPlus

- I want to show the performance of Flavor Tagging with my Vertex Finder
 - ➔ I need to run these networks in LCFIPlus
- I completed the implementation the Vertex Finder with DL (Tensorflow 2.1.0) to the LCFIPlus in iLCSoft (v02-02)
 - Some cmake files are required and some find packages are added to the CMakeLists
 - Also I have to use the shared libraries of tensorflow C++ API built by “bazel”

CMake results

```
-- Found LCFIVertex: /gluster/data/ilc/ilcsoft/v02-02/LCFIVertex/v00-08
-- Found Tensorflow: /home/goto/local/include/tf ◀
-- Found Protobuf: ◀
-- Found Eigen3: /home/goto/local/include/eigen3 (Required is at least version "2.91.0") ◀
-- Check for ROOT_CINT_EXECUTABLE: /gluster/data/ilc/ilcsoft/v02-02/root/6.18.04/bin/rootcint
-- Check for ROOT_DICT_OUTPUT_DIR: /home/goto/ILC/LCFIPlus/build/rootdict
-- Check for ROOT_DICT_CINT_DEFINITIONS:
-- Found Doxygen: /usr/bin/doxygen (found version "1.8.14") found components: doxygen dot
--
-- -----
```

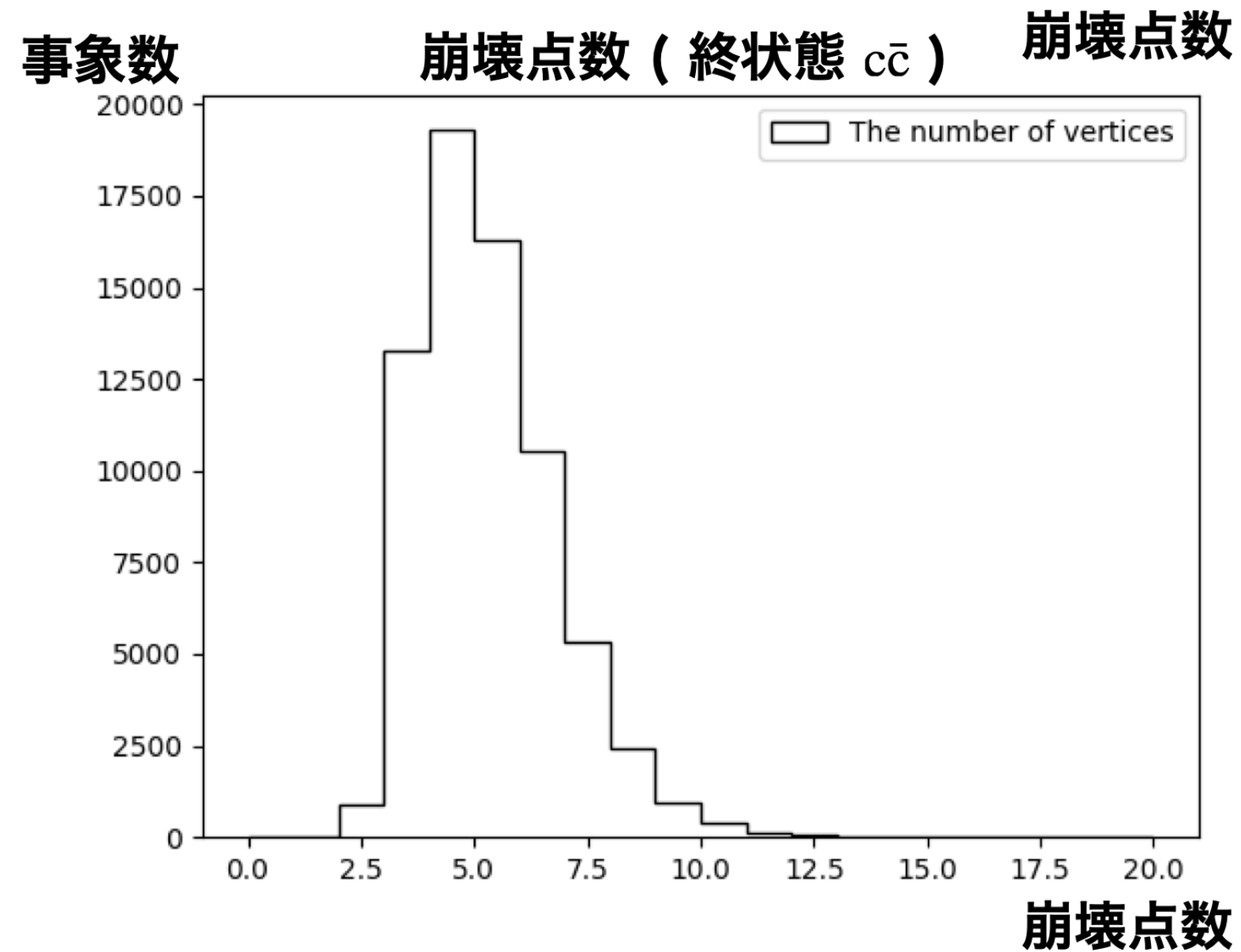
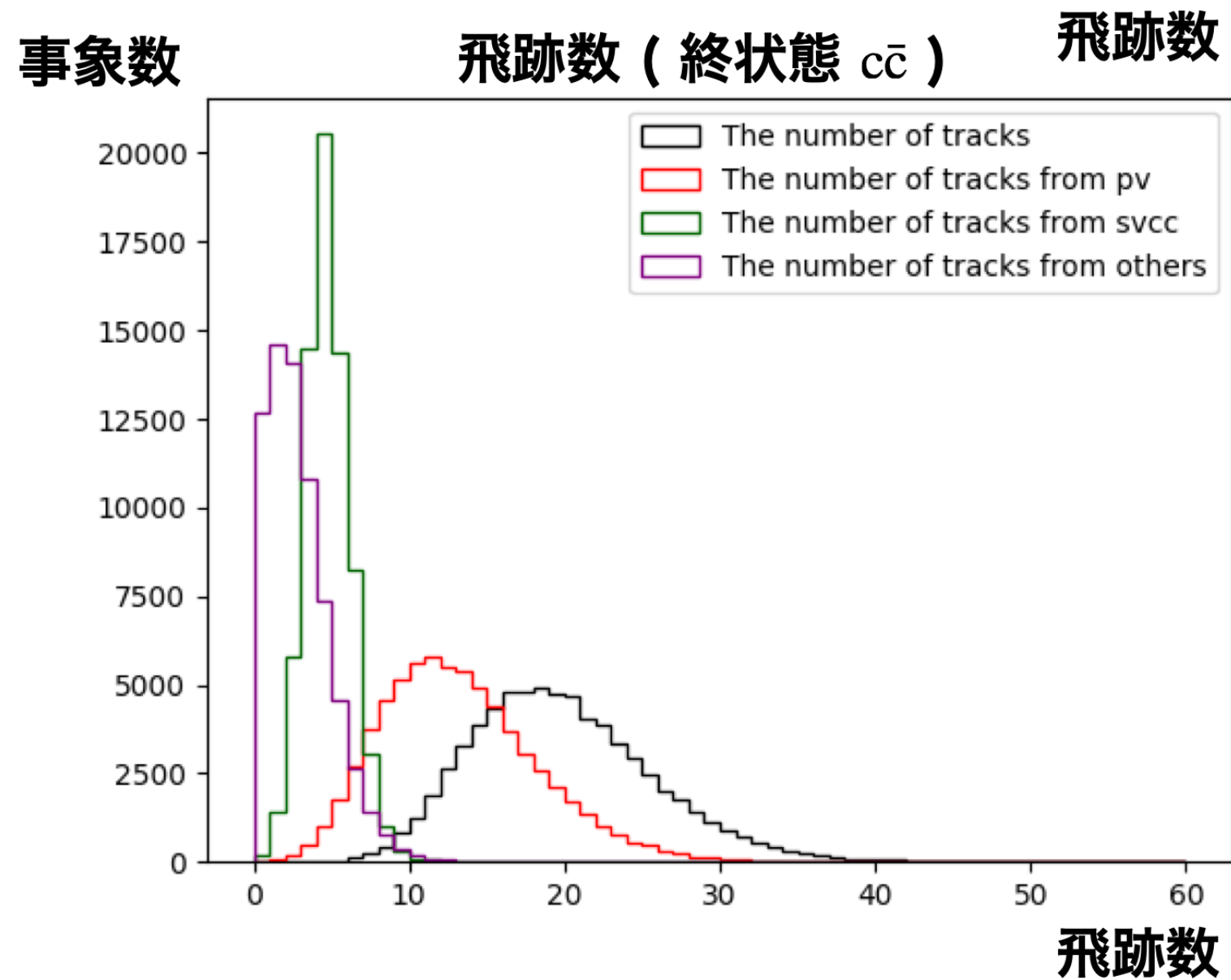
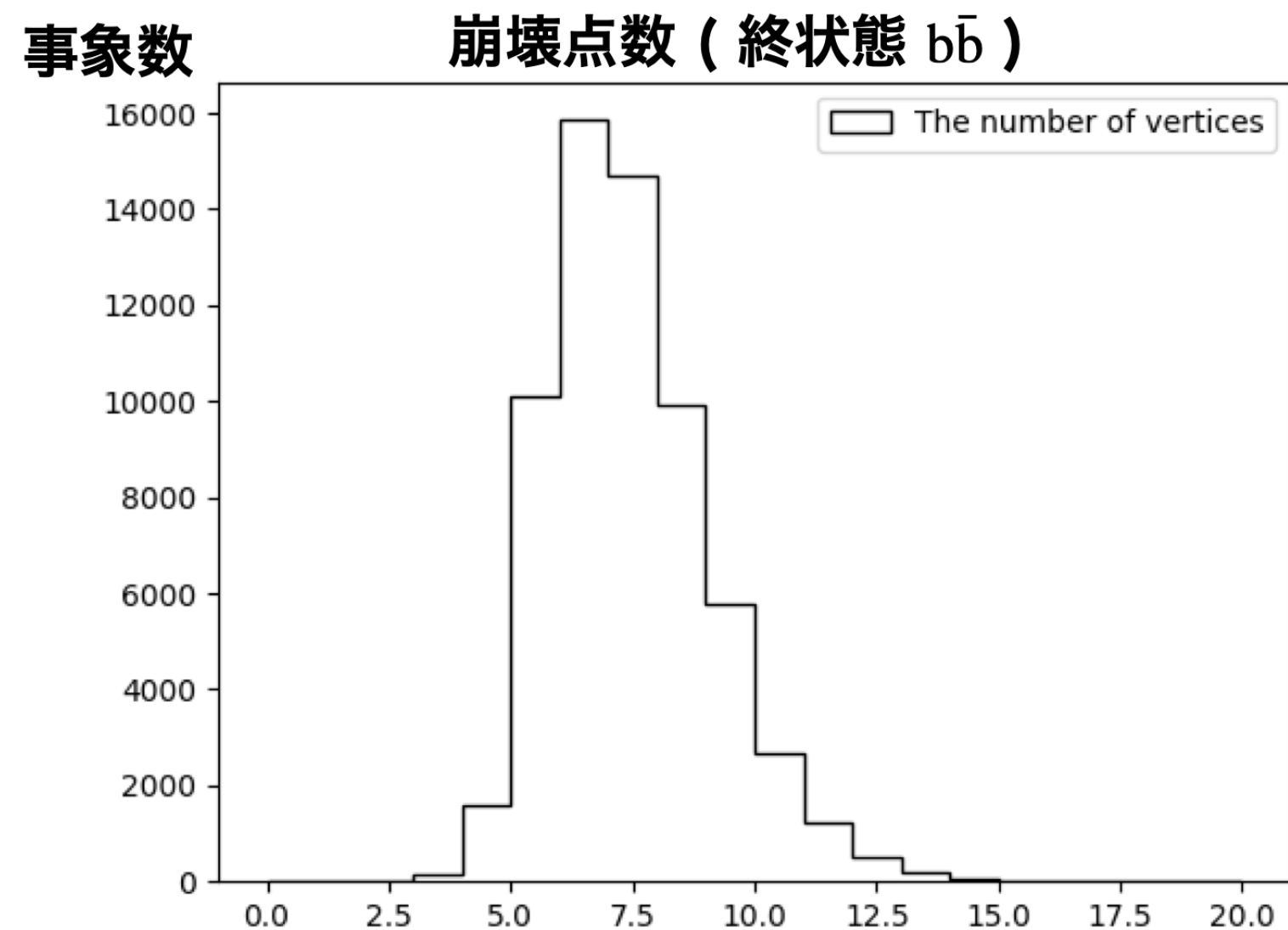
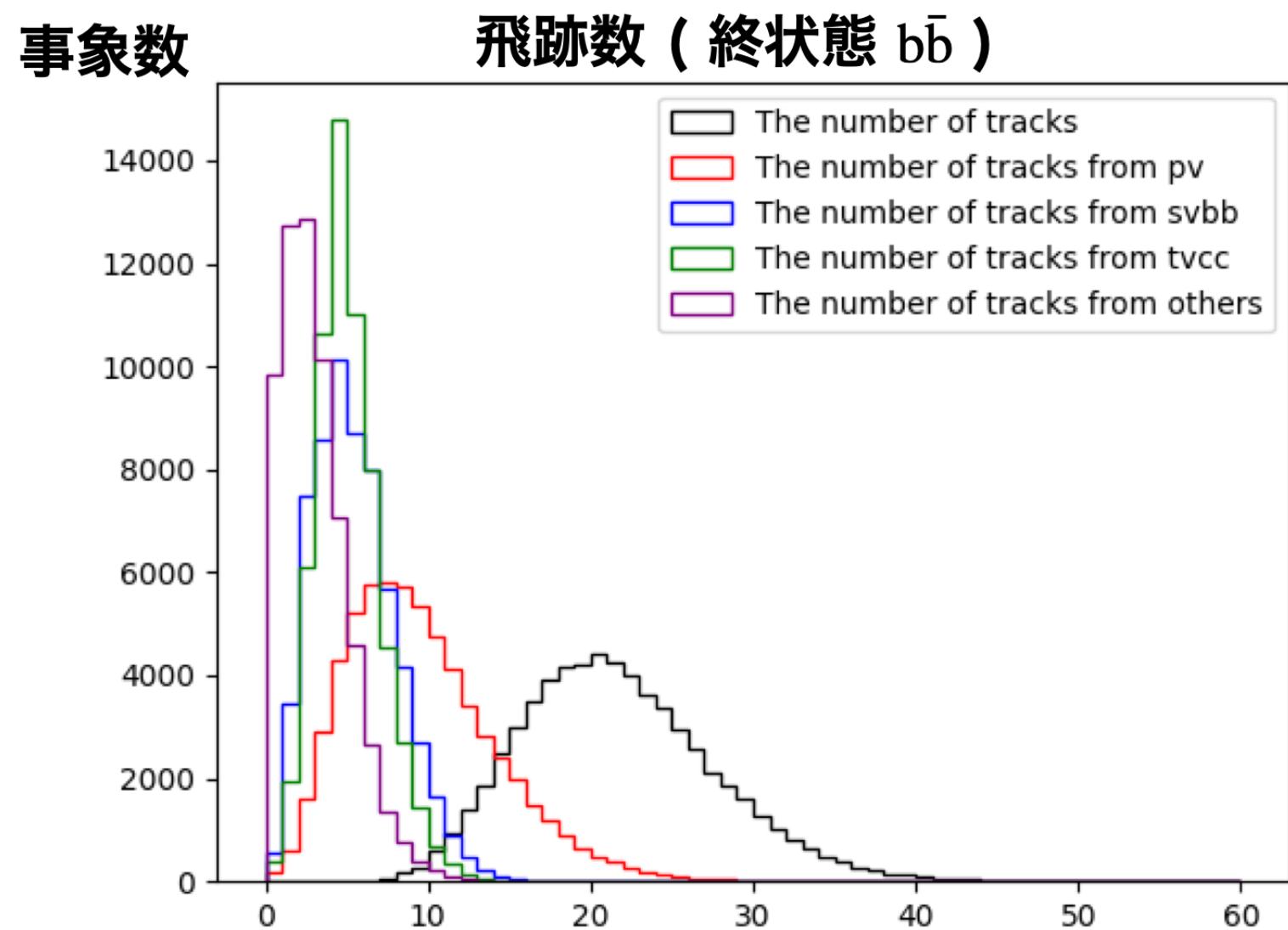
3. Inference with C++

Software setups @ “beep-gpu” server in Kyushu Univ

- For use Tensorflow in LCFIPlus (iLCSoft)
 - Download Tensorflow from GitHub
 - Install Bazel v0.29.1
 - Build Tensorflow C++ API & make shared library (libtensorflow_cc.so, libtensorflow_framework.so)
 - Tensorflow v2.1.0 / CUDA v10.1 / cuDNN v7 / Eigen v3.3.90 / Protobuf v3.8 /
(g++ v8.4.0 / C++11, 14)
 - Move header files and libraries to the /usr/local/include/tf and .../lib/ or your own local
 - Also need to put the eigen3/unsupported, google/protobuf, tf/absl in the /usr/local/include/
 - Make cmake file (FindTensorflow, Eigen3, Protobuf) and write find_package in CMakeLists.txt
 - Include/eigen3/unsupported and libtensorflow_framework.so are not available in this way
We have to use the absolute path to these files
- Install iLCSoft v02-02 (please give attention to cmake version)

1. イントロダクション

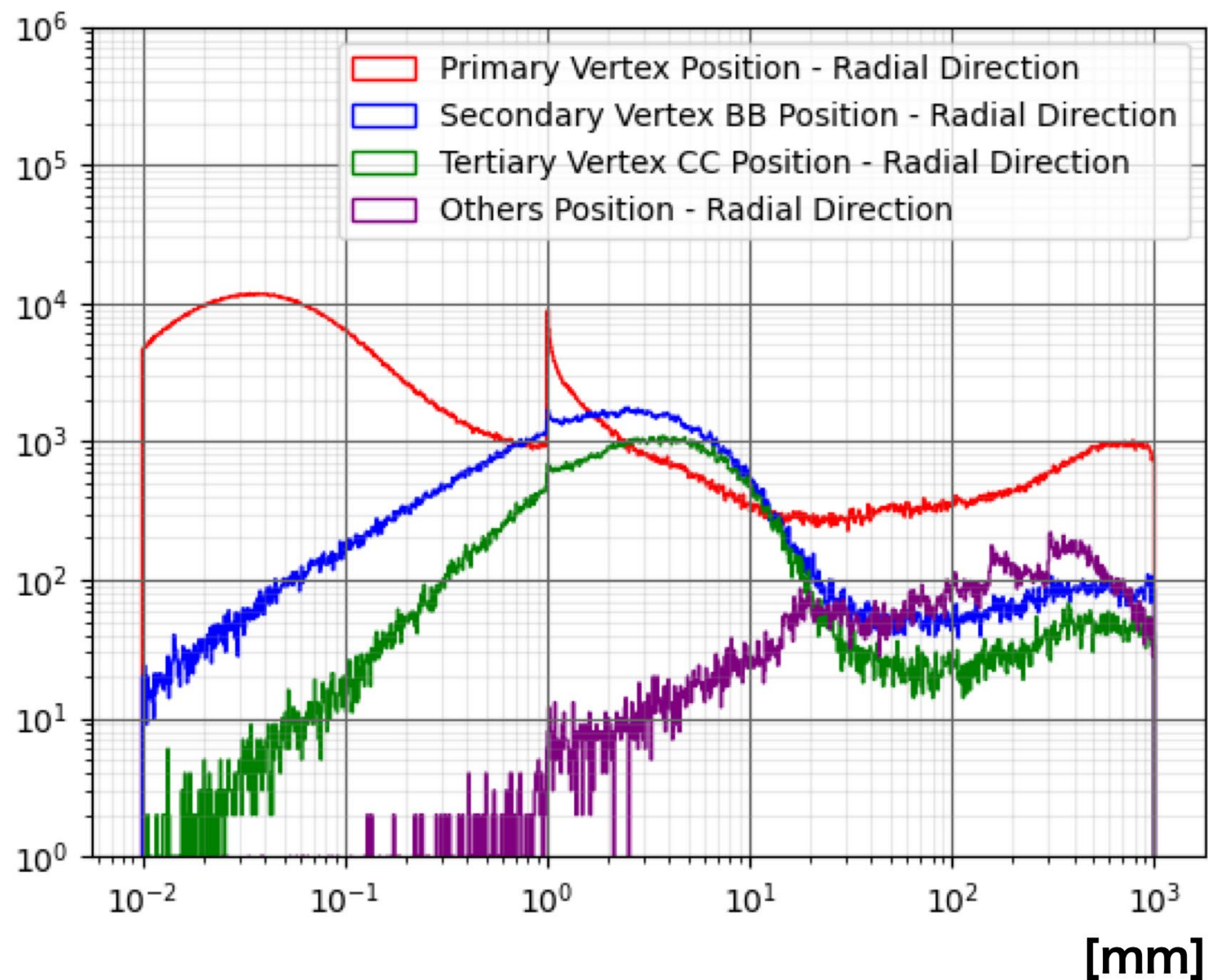
データの性質



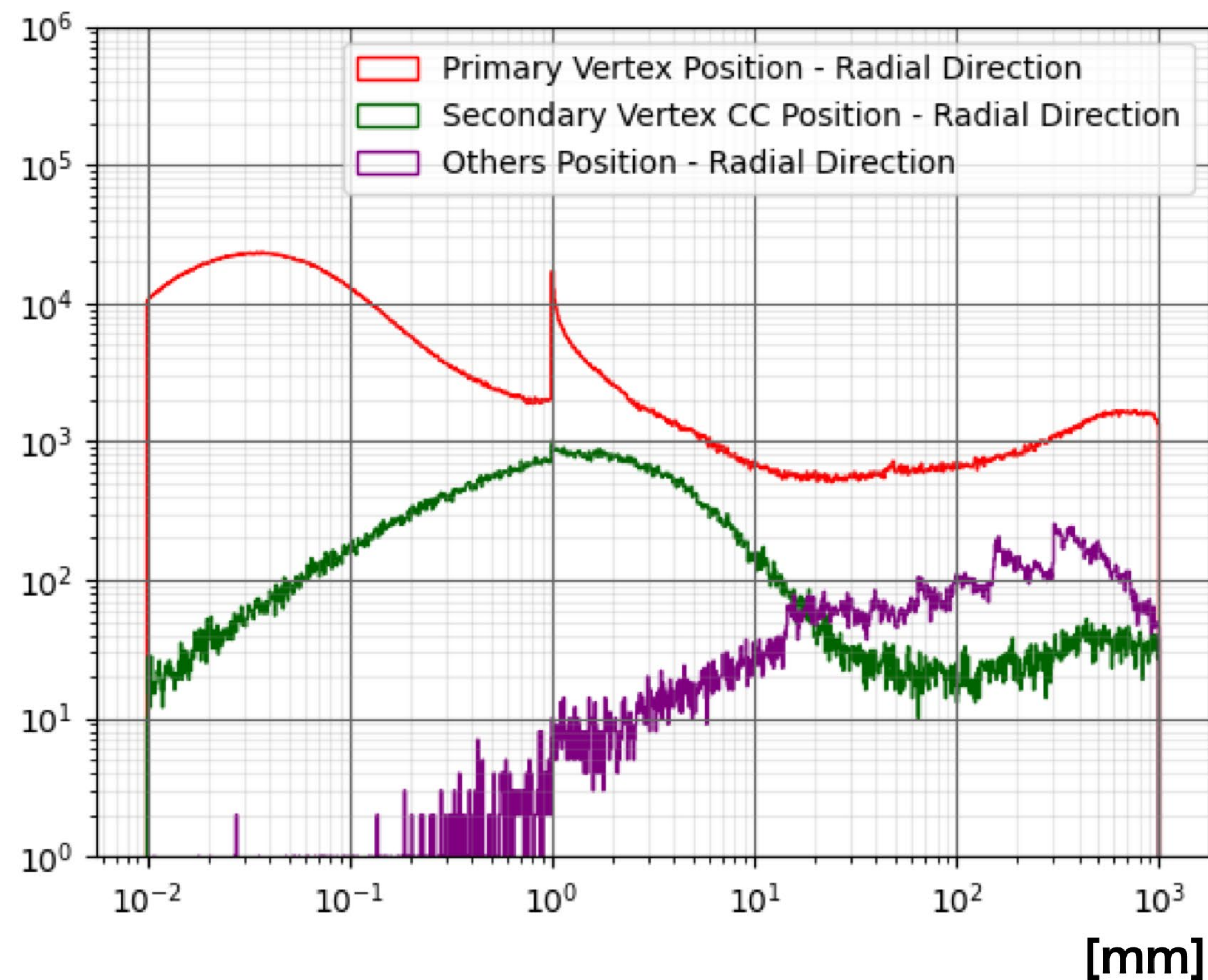
1. イントロダクション

データの性質

終状態 $b\bar{b}$



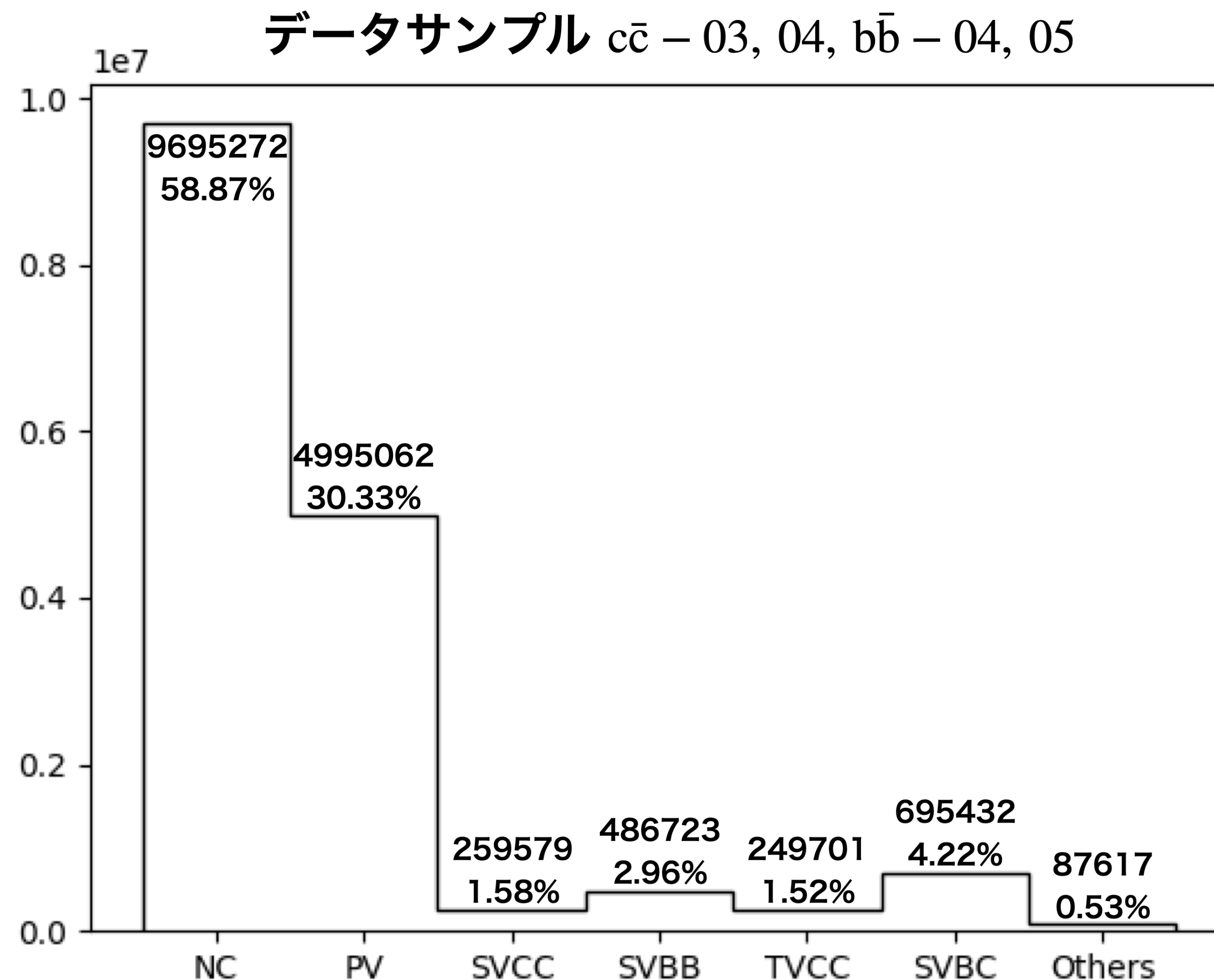
終状態 $c\bar{c}$



1. イントロダクション

データの性質

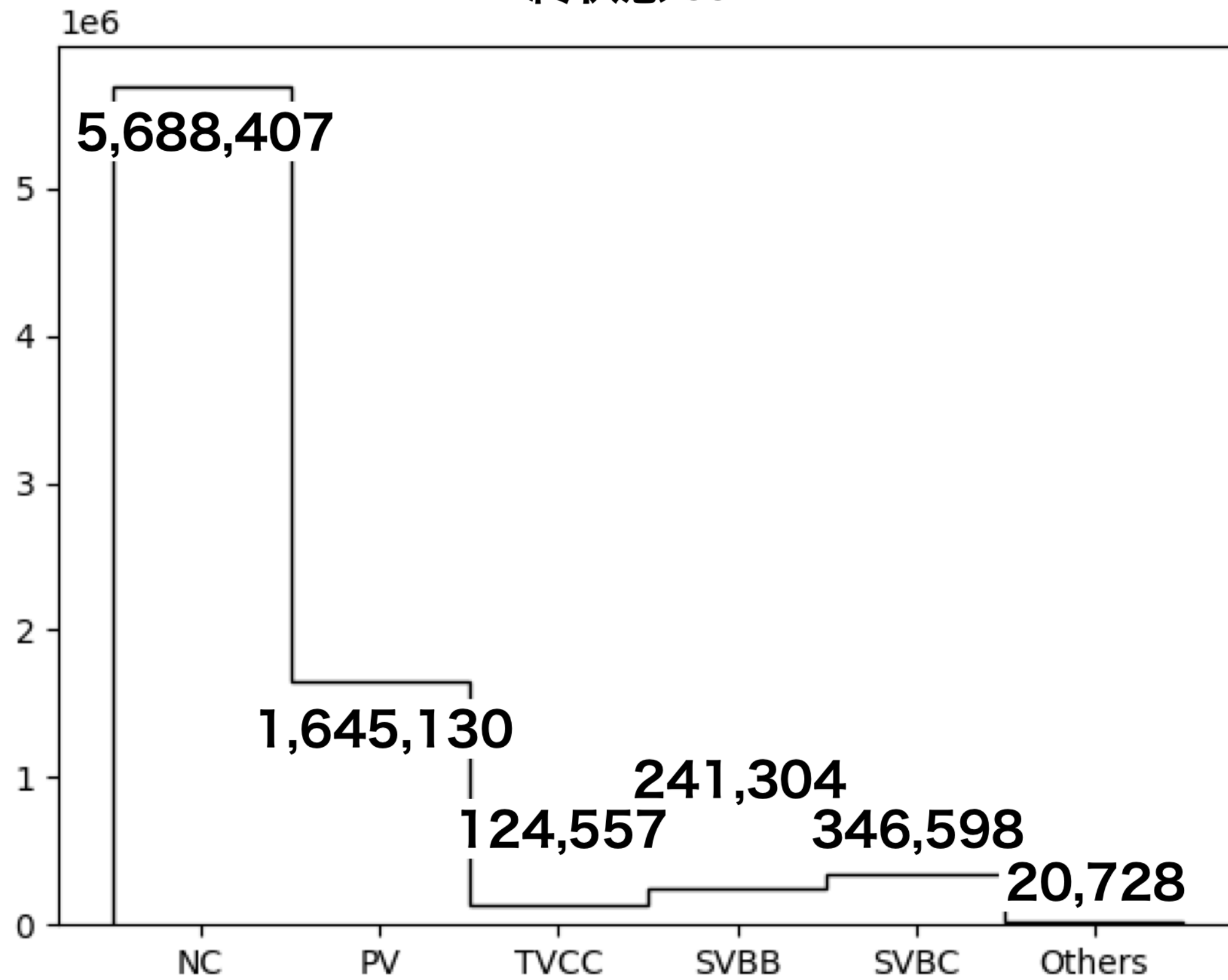
- 飛跡対 (崩壊点の種) の種類
 - NC : 結合していない飛跡対
 - PV : primary vertex 由来
 - SVCC : 終状態 $c\bar{c}$ での secondary vertex 由来
 - SVBB : 終状態 $b\bar{b}$ での secondary vertex 由来
 - TVCC : 終状態 $b\bar{b}$ での tertiary vertex 由来
 - SVBC : 終状態 $b\bar{b}$ での secondary vertex から1本 tertiary vertex から1本で構成された飛跡対
 - Others : その他の崩壊点由来の飛跡対
 - V^0 の崩壊, 光子変換など



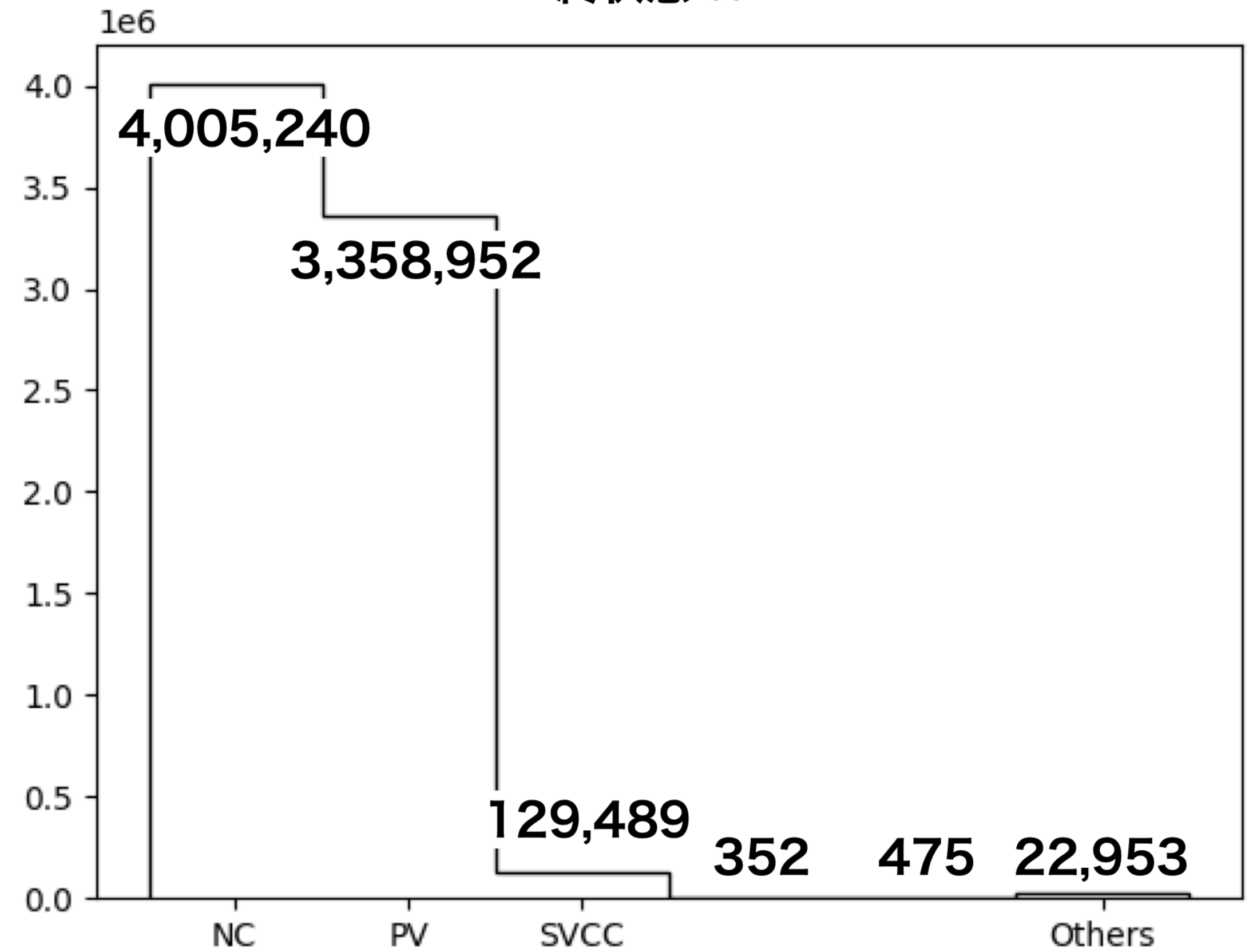
1. イントロダクション

データの性質

終状態 $b\bar{b}$



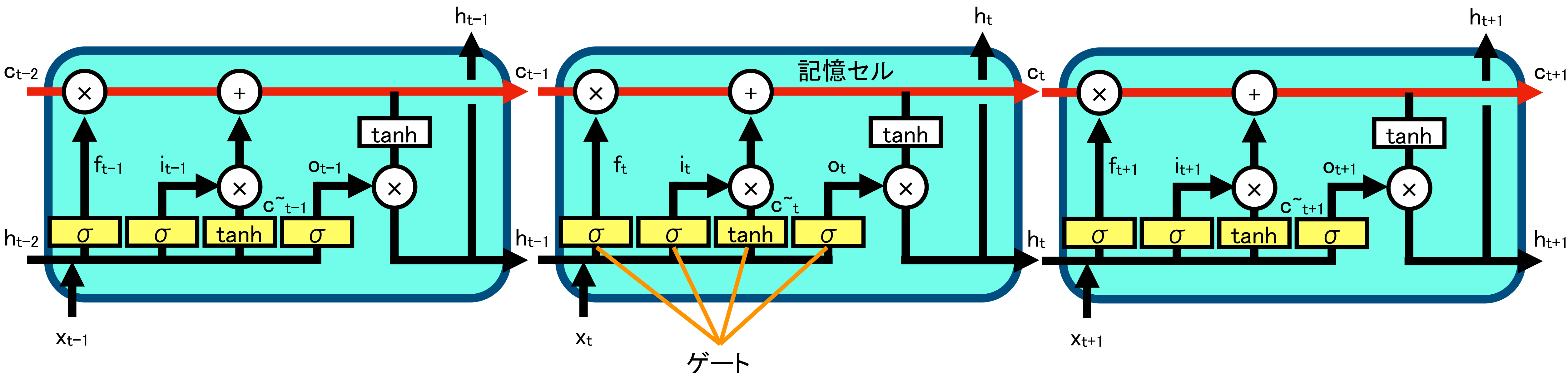
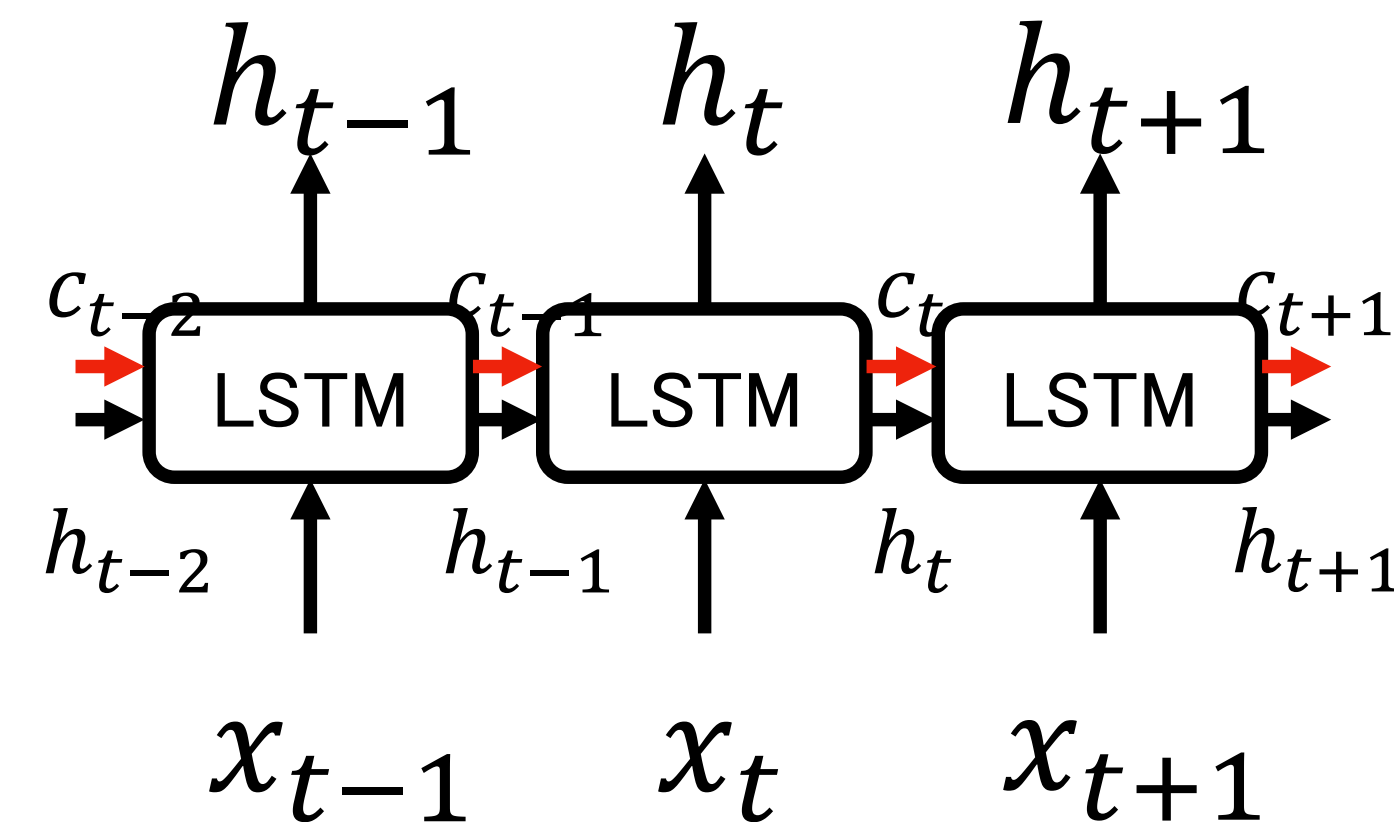
終状態 $c\bar{c}$



1. イントロダクション

LSTM (Long short-term memory)

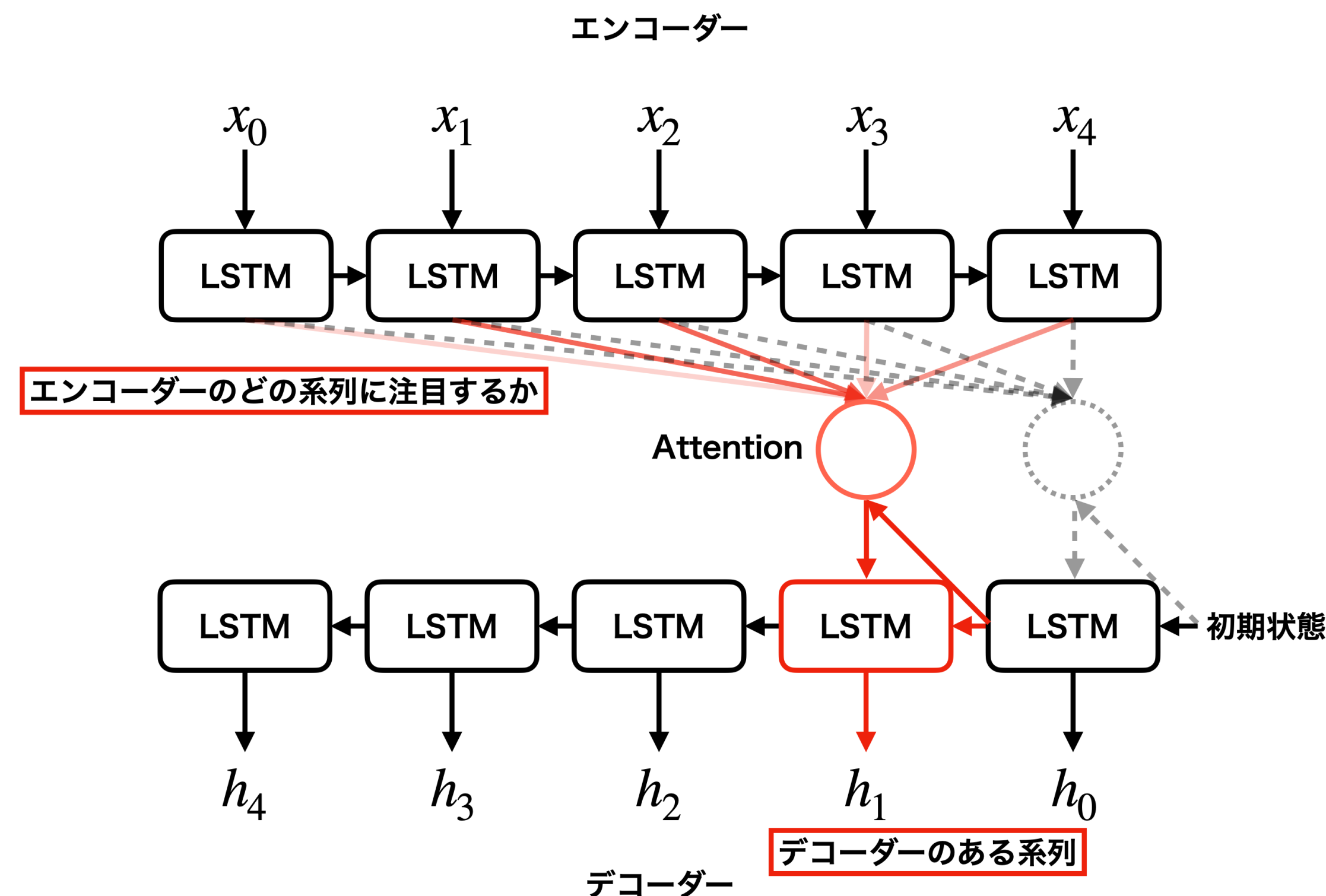
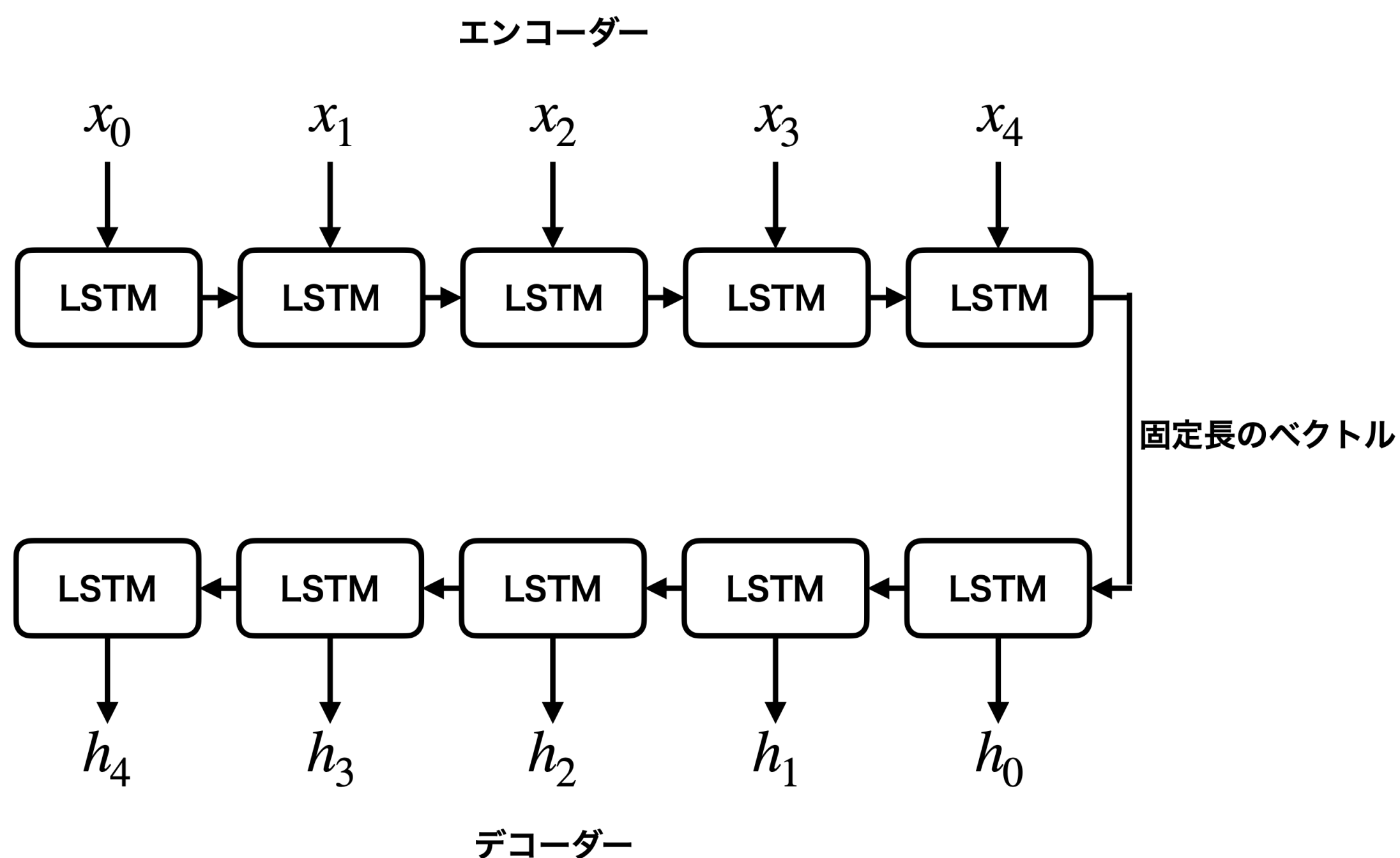
- 日本語では「長短期記憶」
- リカレントニューラルネットワークの問題を解決する為に開発されたネットワーク
 - リカレントニューラルネットワークは**長期的な情報を保持出来ない**
- 4つのゲートと記憶セルを持っている
 - ゲートは重み更新についての問題を解決するためのテクニック
 - 記憶セル c** は長期的な情報保持のためのテクニック



1. イントロダクション

Attention

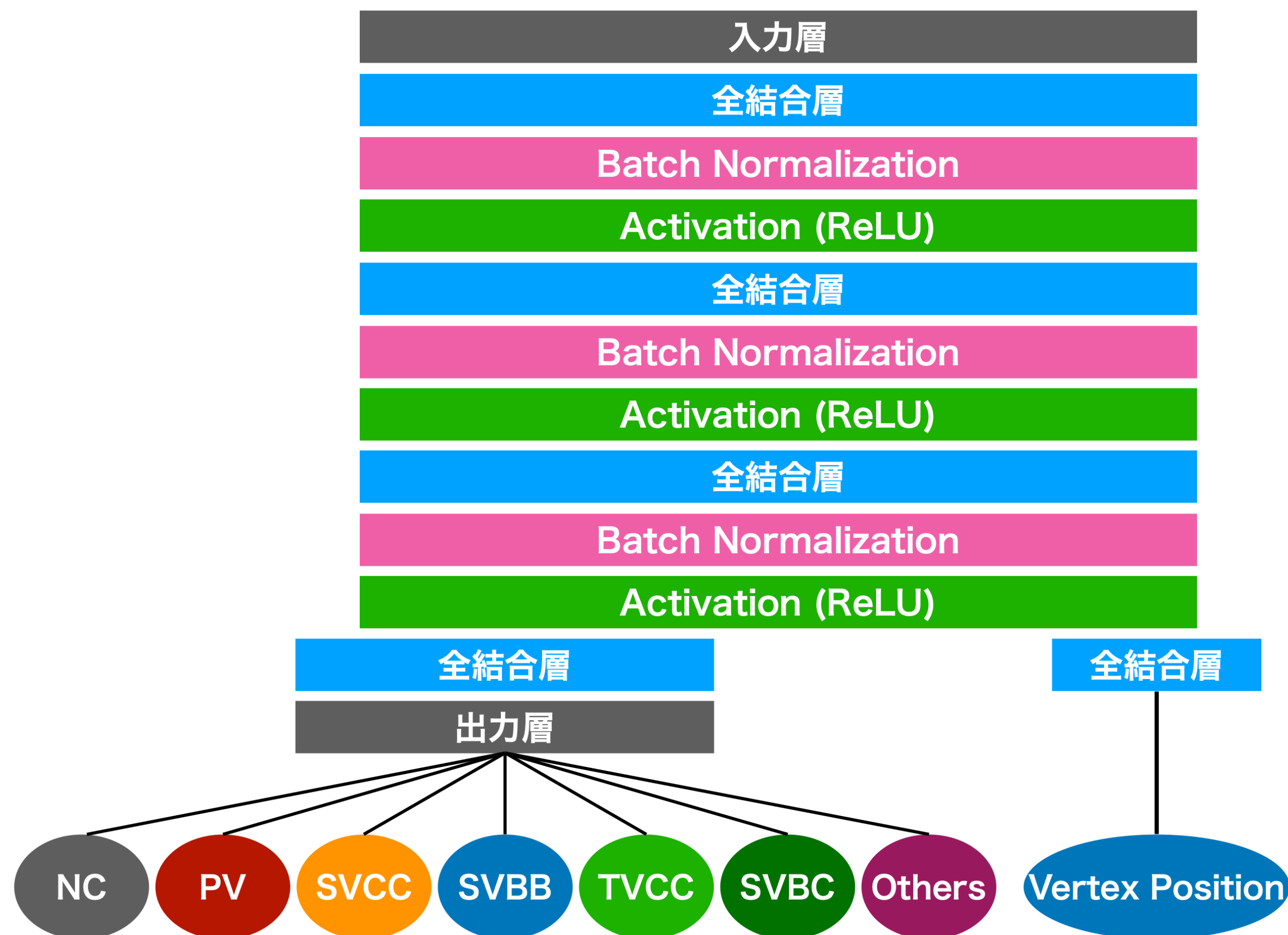
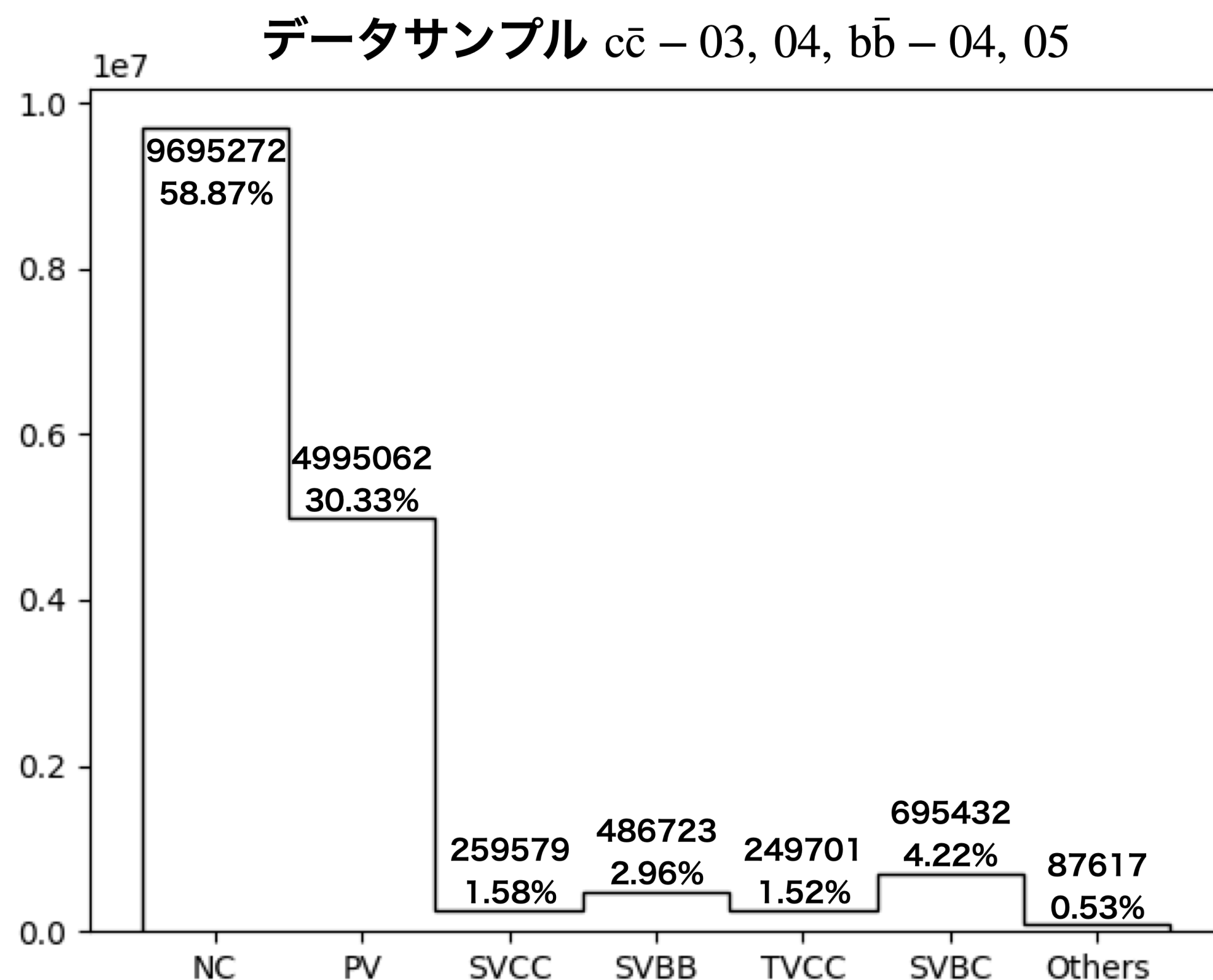
- 情報のある部分に注目させるための技術
 - 代名詞が何を意味しているか
 - 質問に対する答えの位置 など...
- 単純なエンコーダー・デコーダーでは系列の長さに関わらず、同じ大きさの情報で伝達してしまう
- Attentionはエンコーダーに応じて、情報を確保できる



2. 崩壊点検出のためのニューラルネットワーク

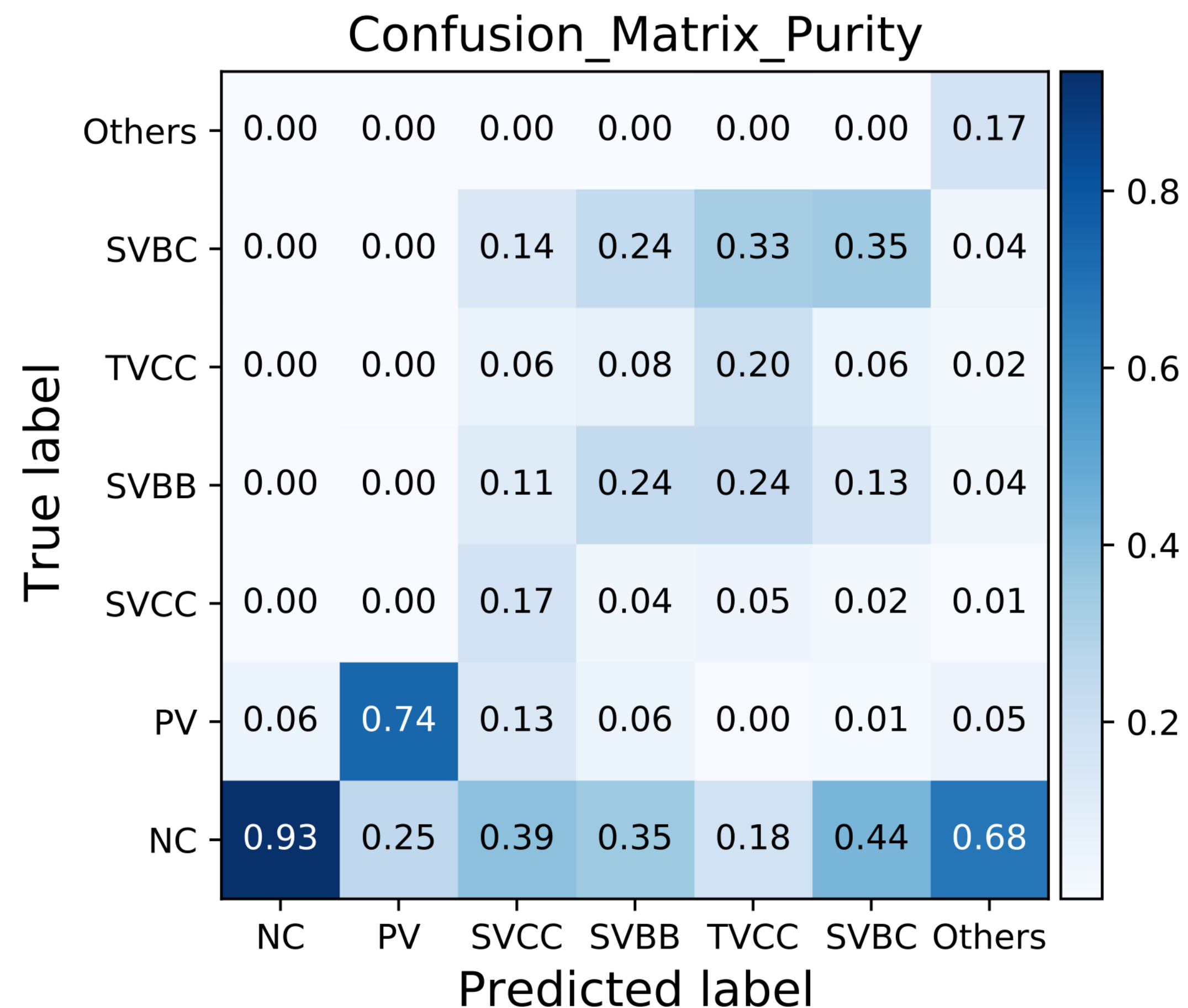
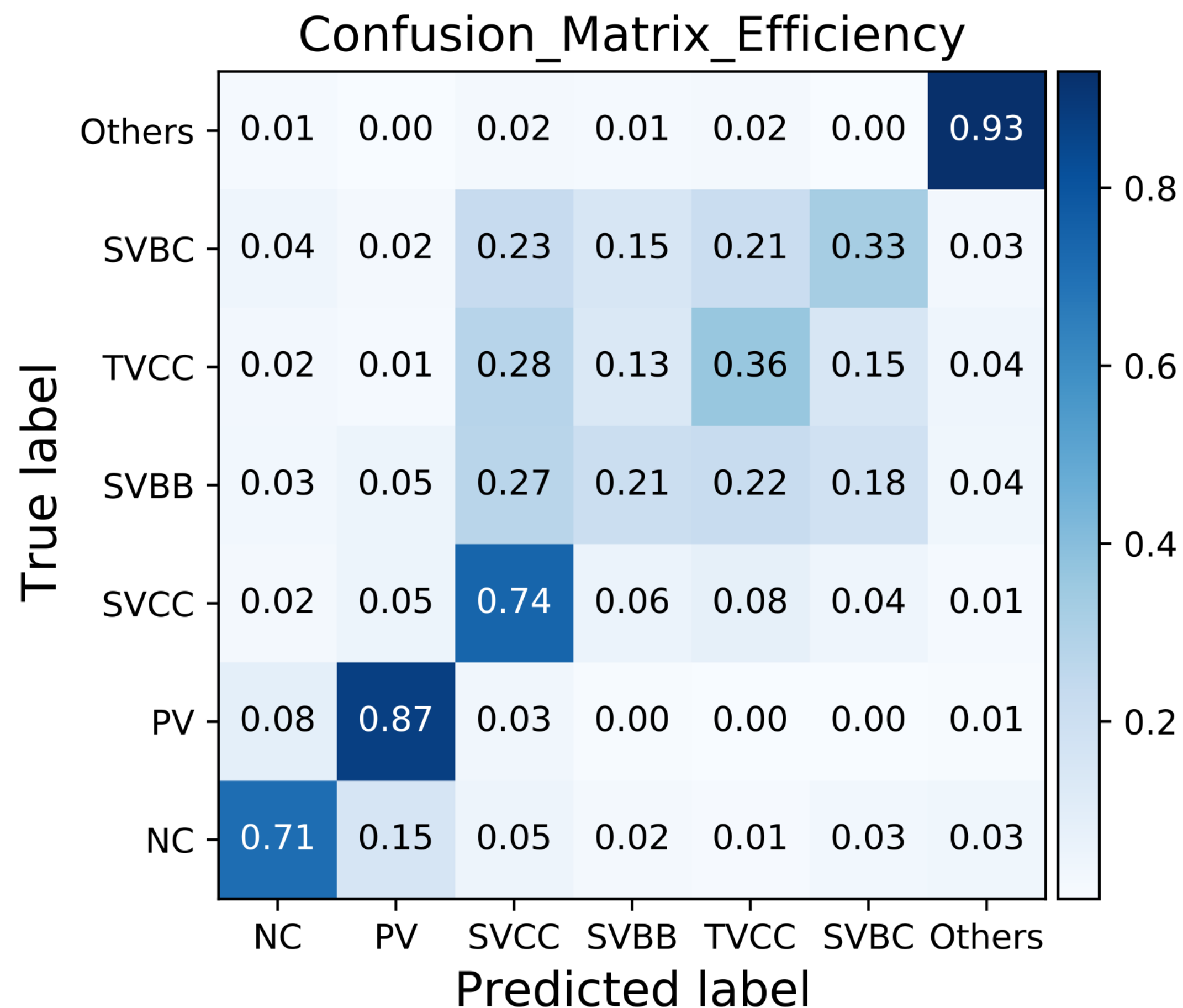
飛跡対についてのネットワーク -構造と性能-

- ・ シンプルなネットワークを使用
- ・ **不均衡データ**である為、損失関数に**重み**を付ける
 - ▶ 損失関数：学習に使用する評価関数
(最小化するように学習が進む)



2. 崩壊点検出の為にニューラルネットワーク

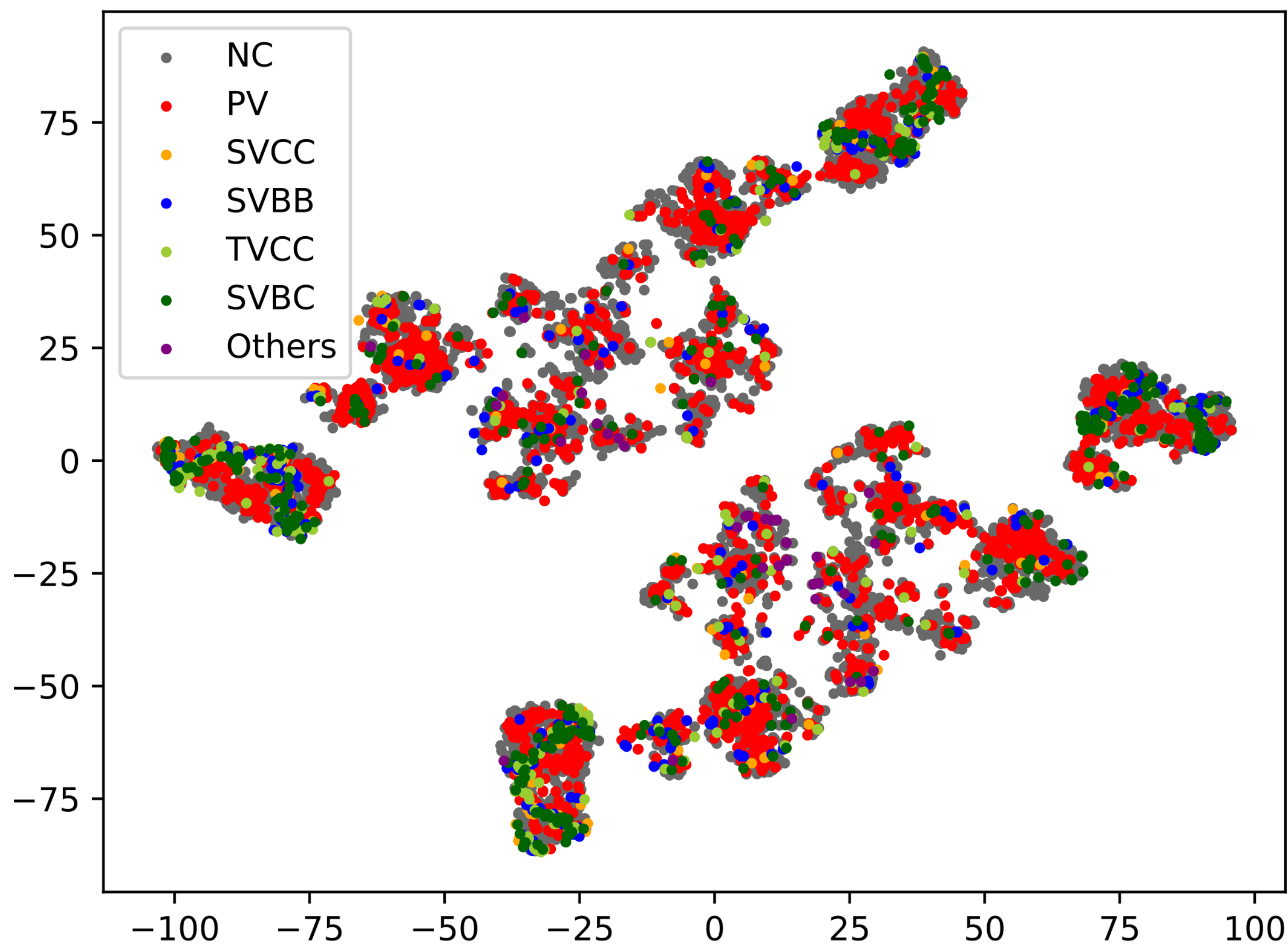
飛跡対についてのネットワーク -構造と性能-



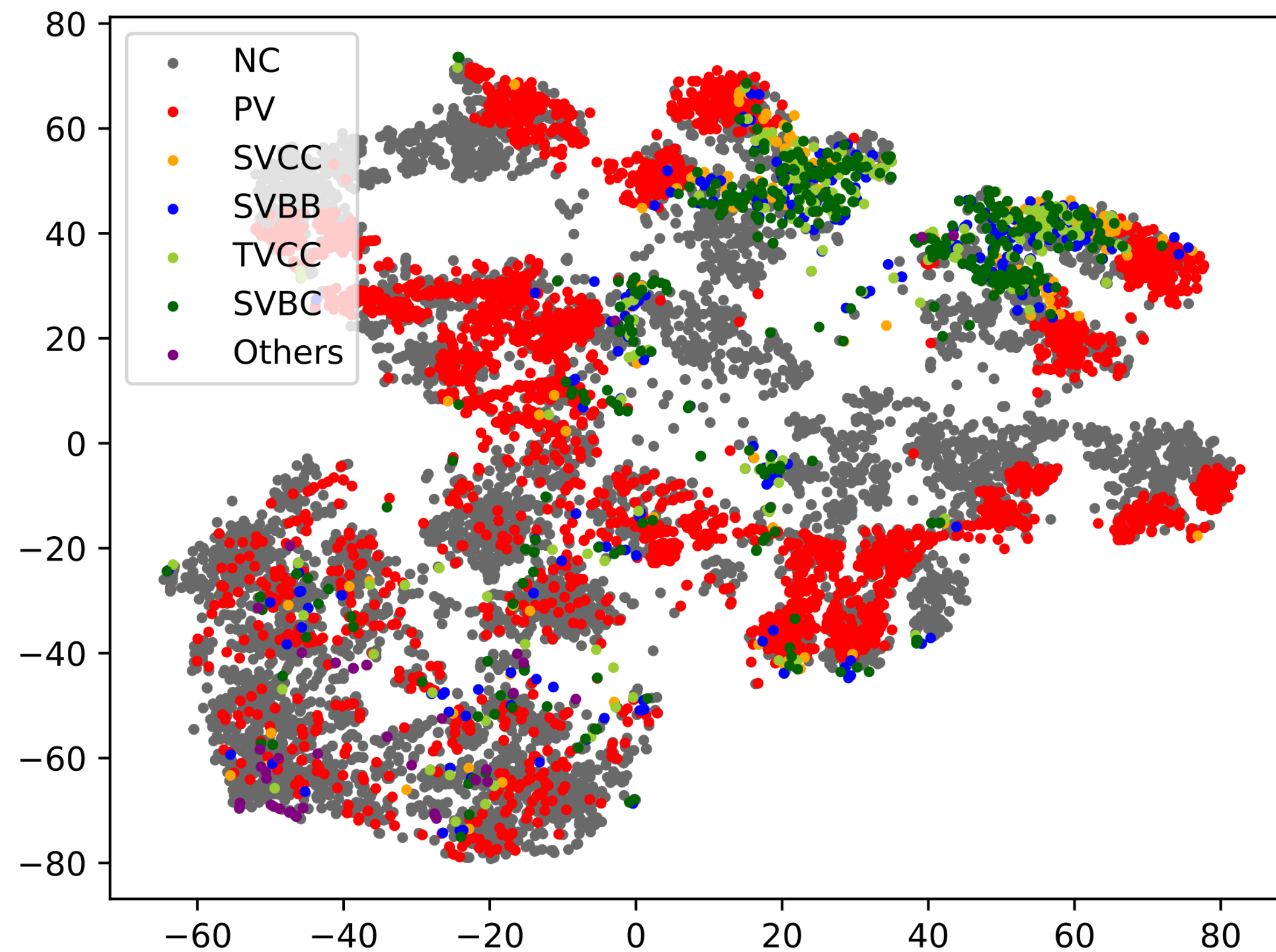
2. 崩壊点検出の為にニューラルネットワーク

飛跡対についてのネットワーク -構造と性能-

入力変数



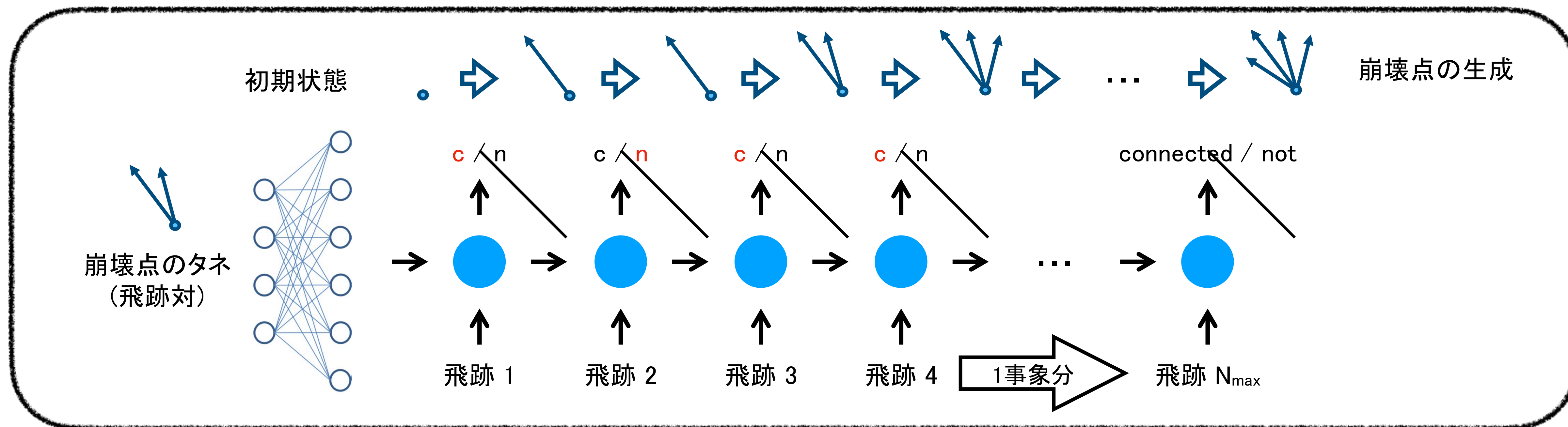
出力の直前の全結合層



2. 崩壊点検出の為にニューラルネットワーク

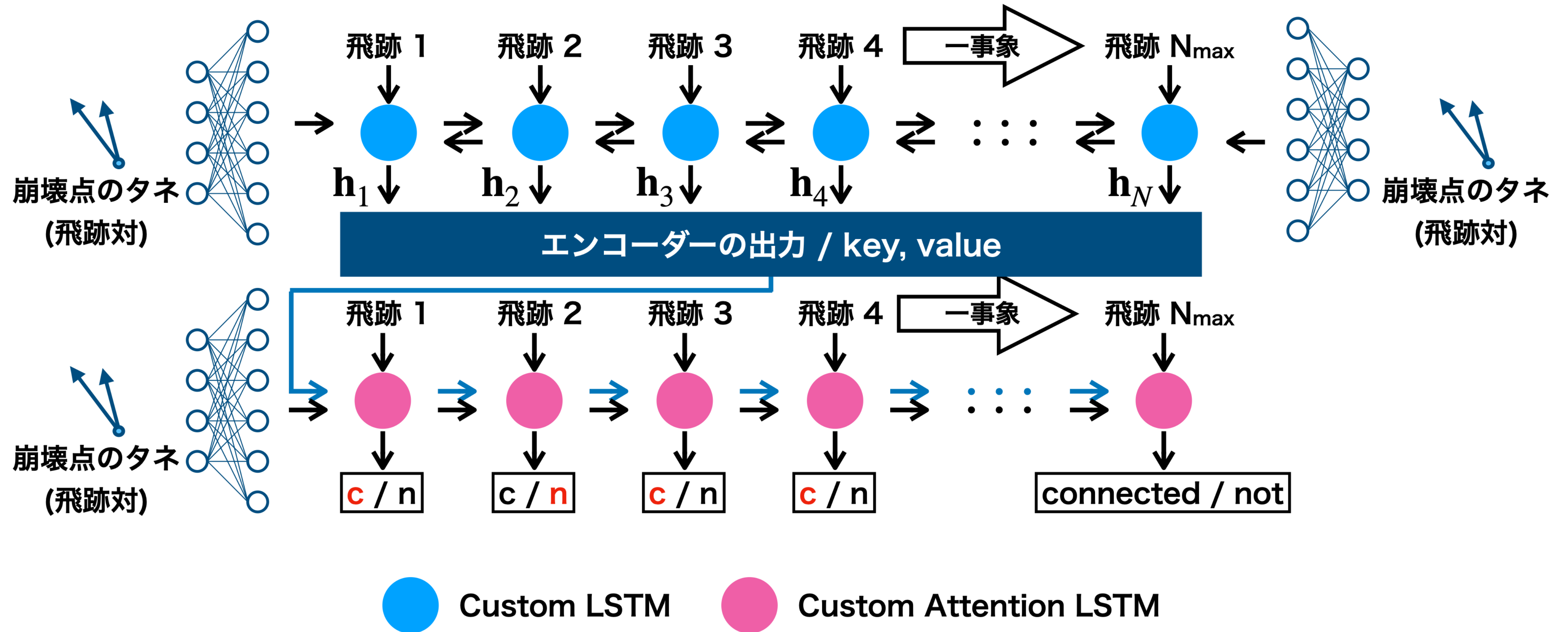
LSTMを用いたアプローチ

- 二本以上の飛跡を処理できるネットワークを構築したい
- 問題点
 - 含まれる**飛跡の数**が事象毎に異なる
 - 含まれる**崩壊点の数**が事象毎に異なる
 - ➡ 可変長なネットワーク(リカレントニューラルネットワーク)
- 初期状態(崩壊点のタネ)に対して、飛跡が繋がっているかどうか(初期状態を学習する)



2. 崩壊点検出のためのニューラルネットワーク

任意の数の飛跡についてのネットワーク - 構造 -

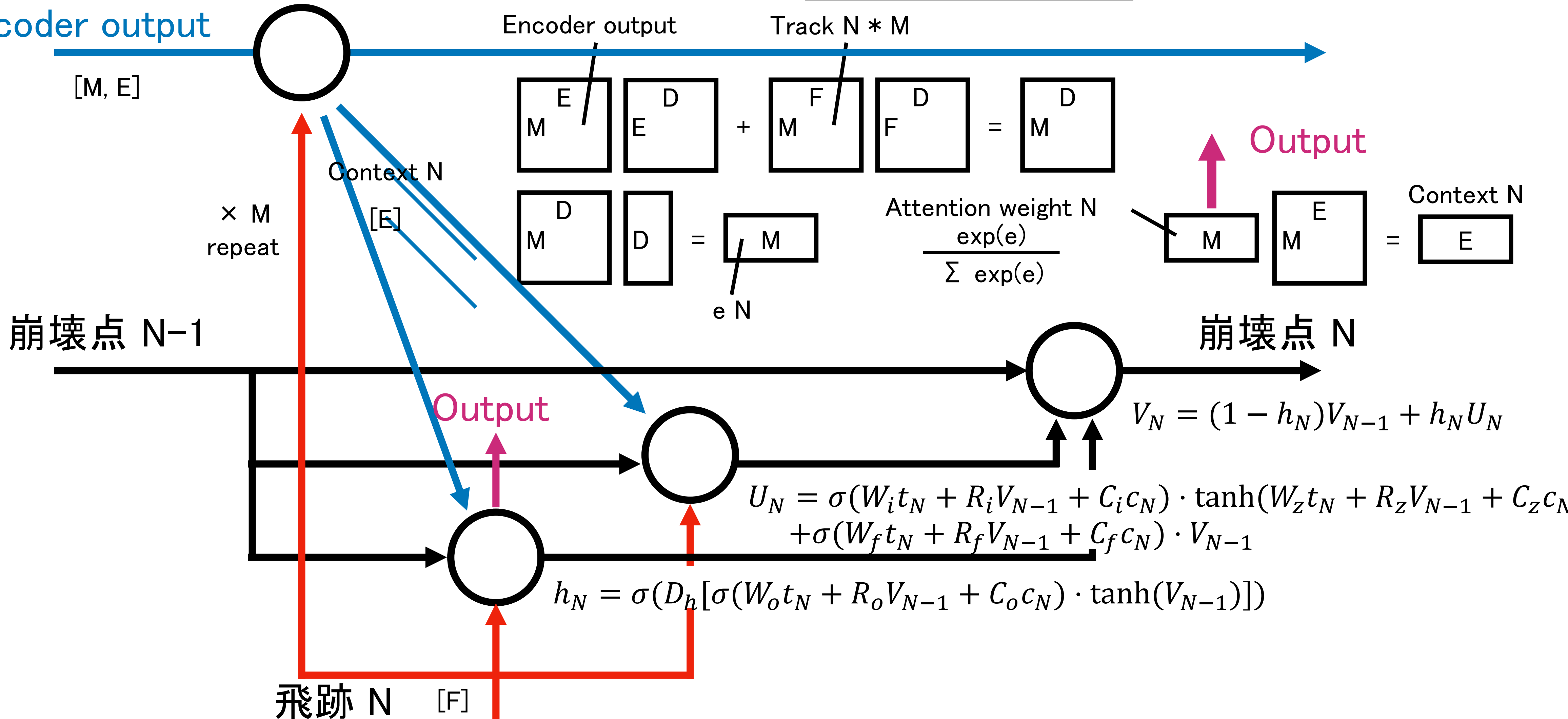


2. 崩壊点検出の為にニューラルネットワーク

崩壊点検出の為にAttention LSTM

Encoder output

Additive Attention



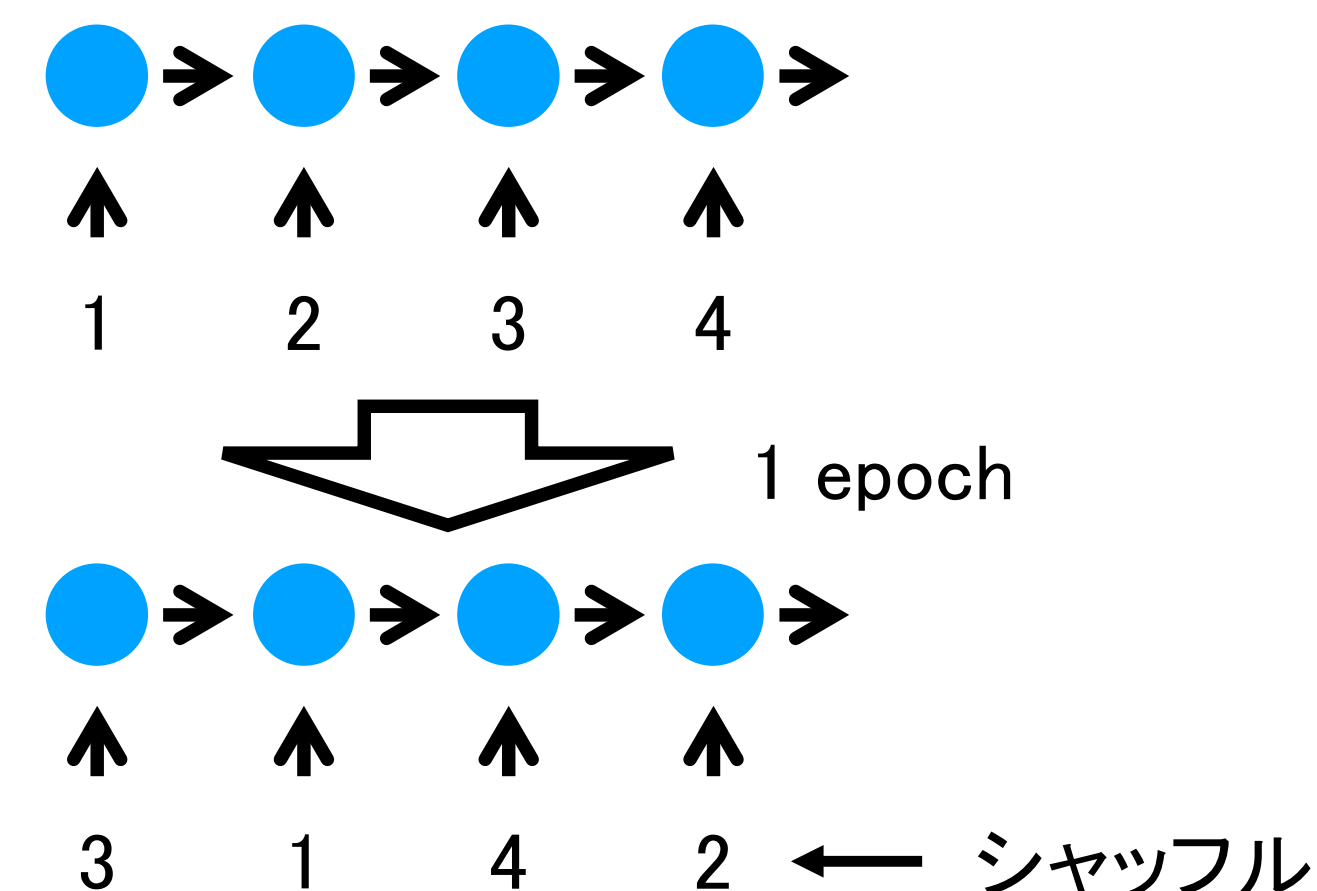
2. 崩壊点検出の為にニューラルネットワーク

任意の数の飛跡についてのネットワーク -学習と性能-

- 損失関数 : binary cross entropy
- 最適化/学習率 : Adam/0.001
 - ▶ 重み更新の手法とステップ幅
- 学習回数(Epoch) : 100 epochs
- バッチサイズ : 32
 - ▶ 重み更新毎のサンプル数
- Framework / Hardware : Tensorflow, Keras / TITAN RTX

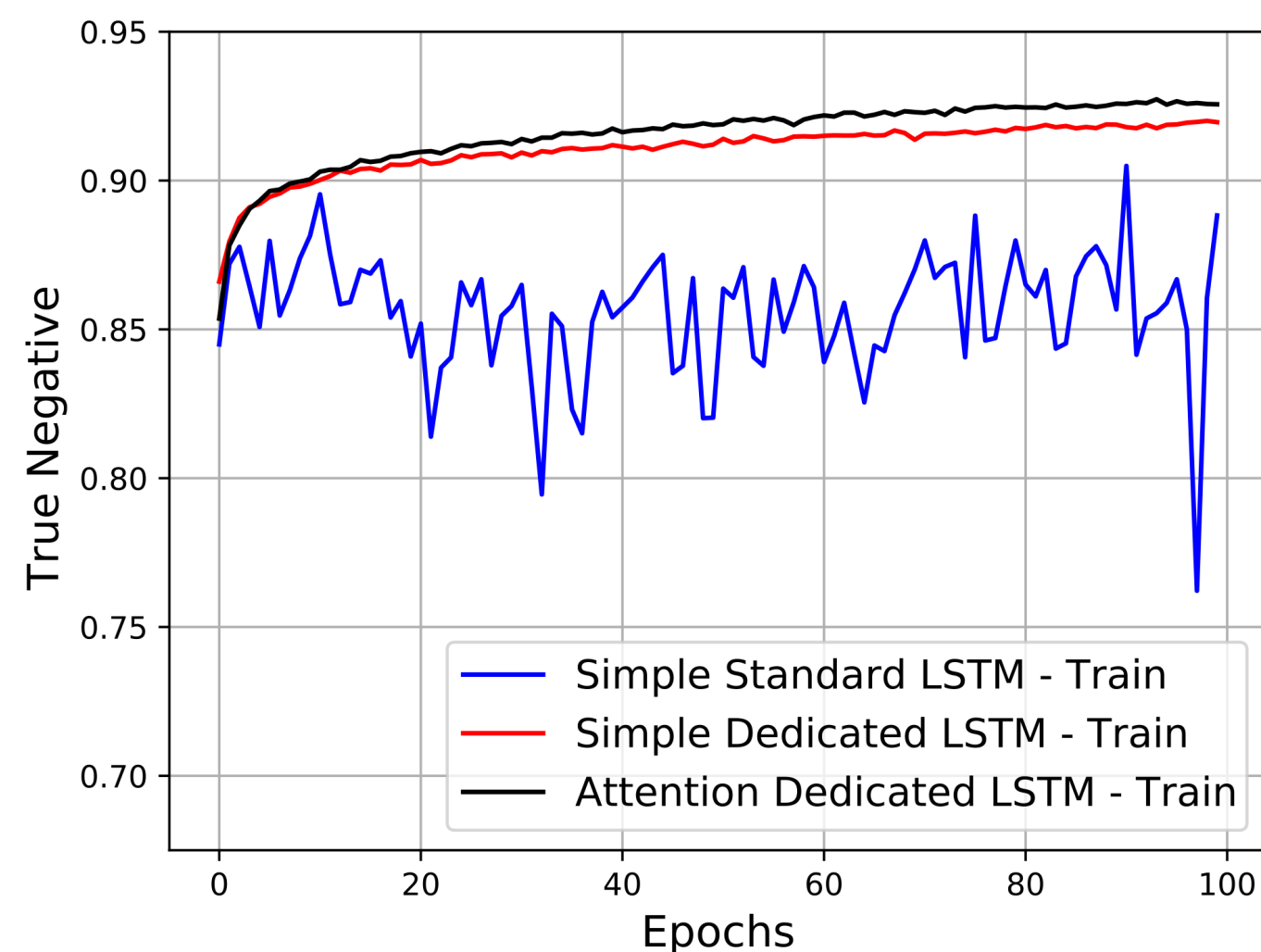
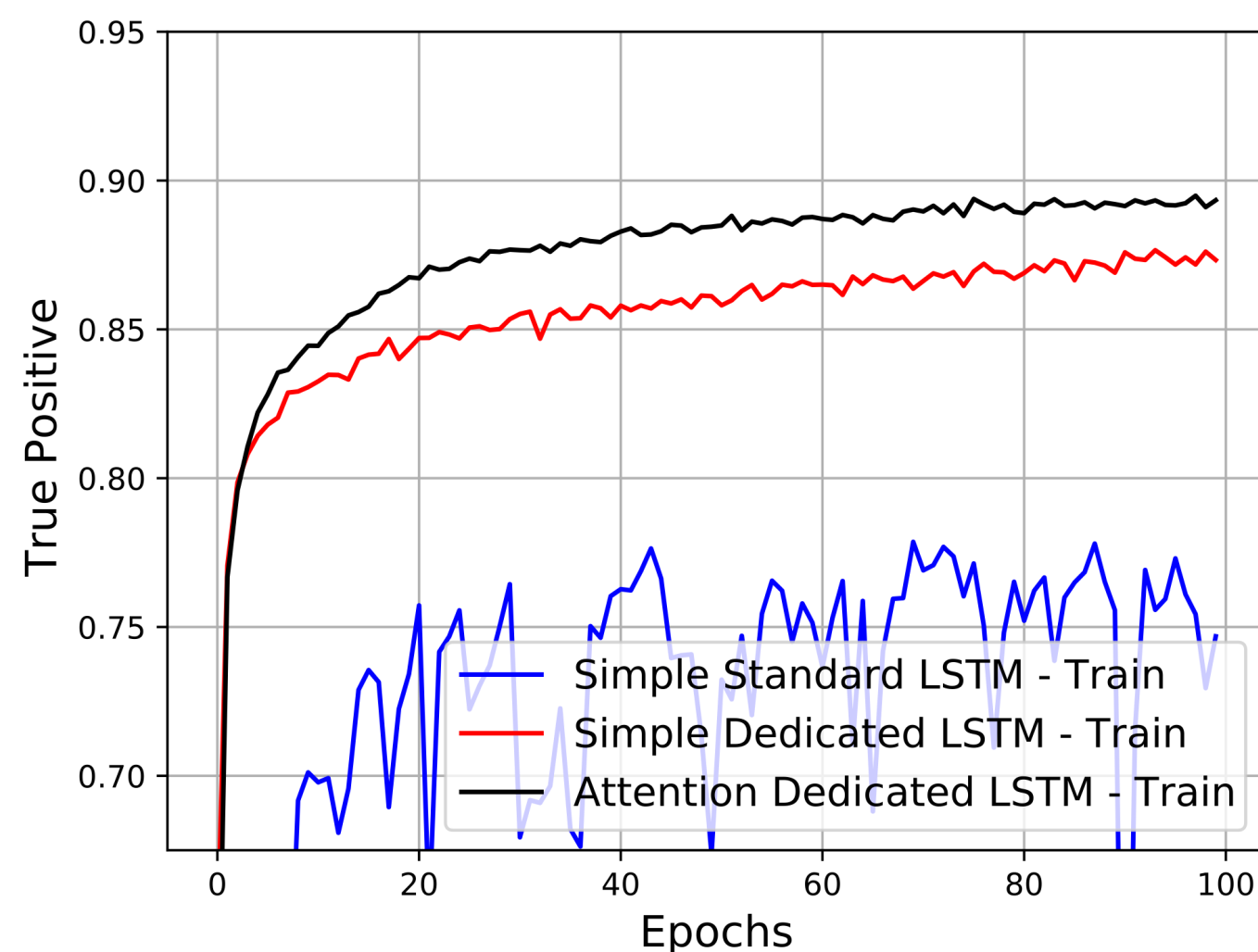
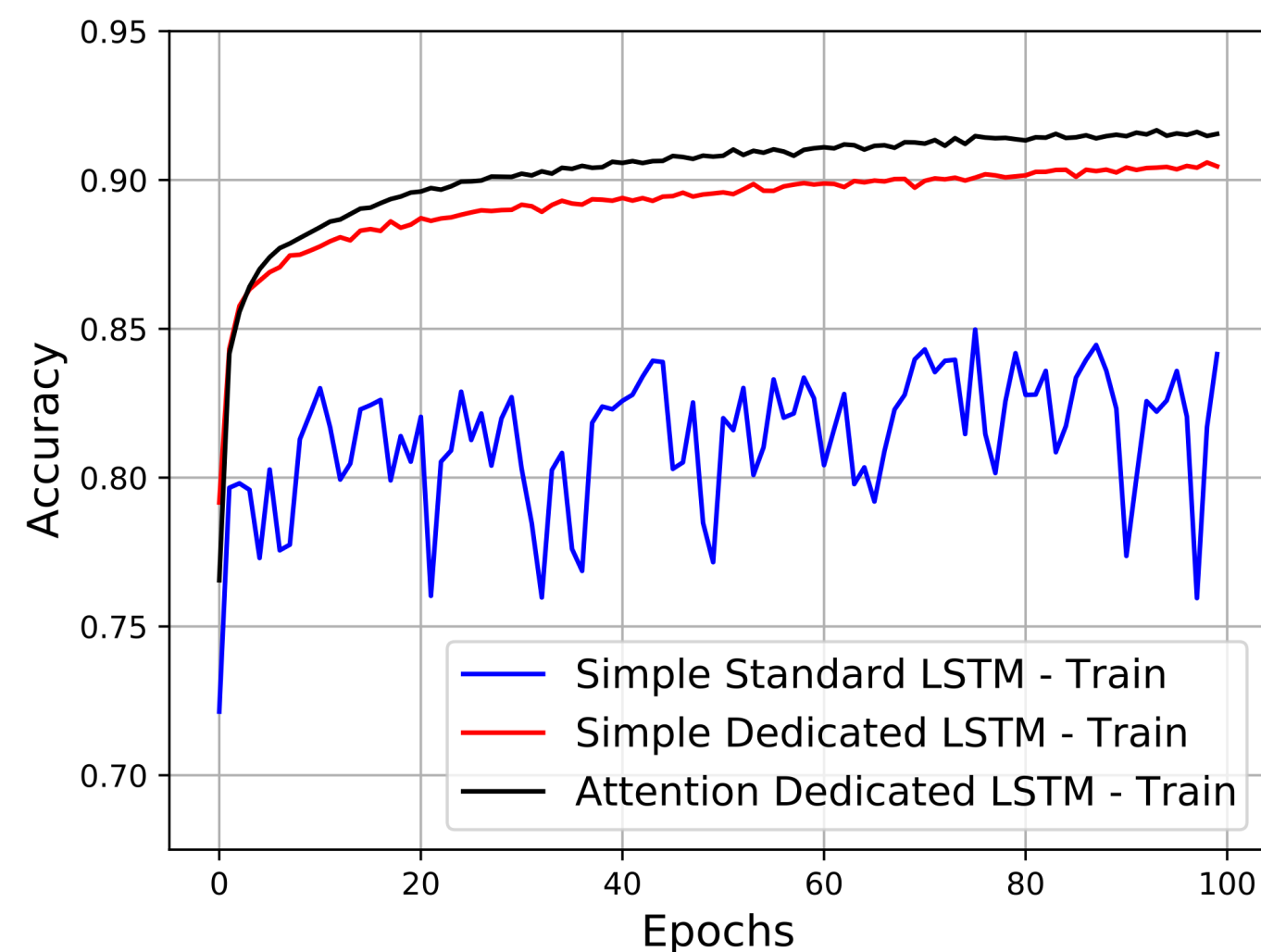
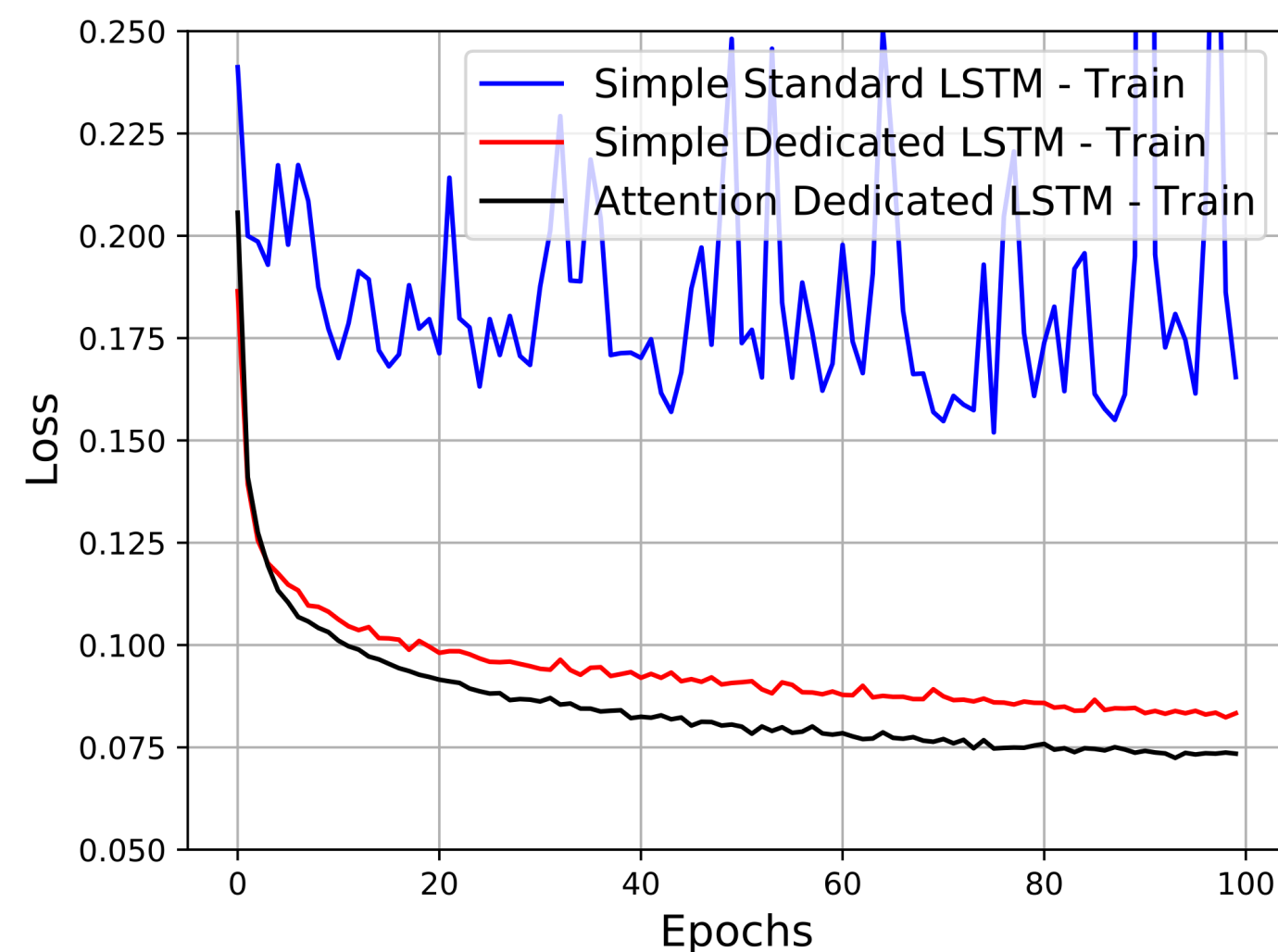
- 20000 事象 (1159547 samples) ➡ □ 1 epoch毎にランダムに50000 sampleを選び学習
 - ▶ 崩壊点毎に教師データが1 sample生成される

- ゼロパディングとマスク
 - 学習時は全事象の飛跡の最大数で「**ゼロ埋め (パディング)**」し、学習に影響が出ないように「**マスク**」している
- 飛跡順のシャッフル
 - 本来、飛跡に順序はない為、学習においても出来る限り系列に依存しないよう 1 epoch毎に**飛跡の順序をシャッフル**している



2. 崩壊点検出のためのニューラルネットワーク

任意の数の飛跡についてのネットワーク - 学習と性能 -

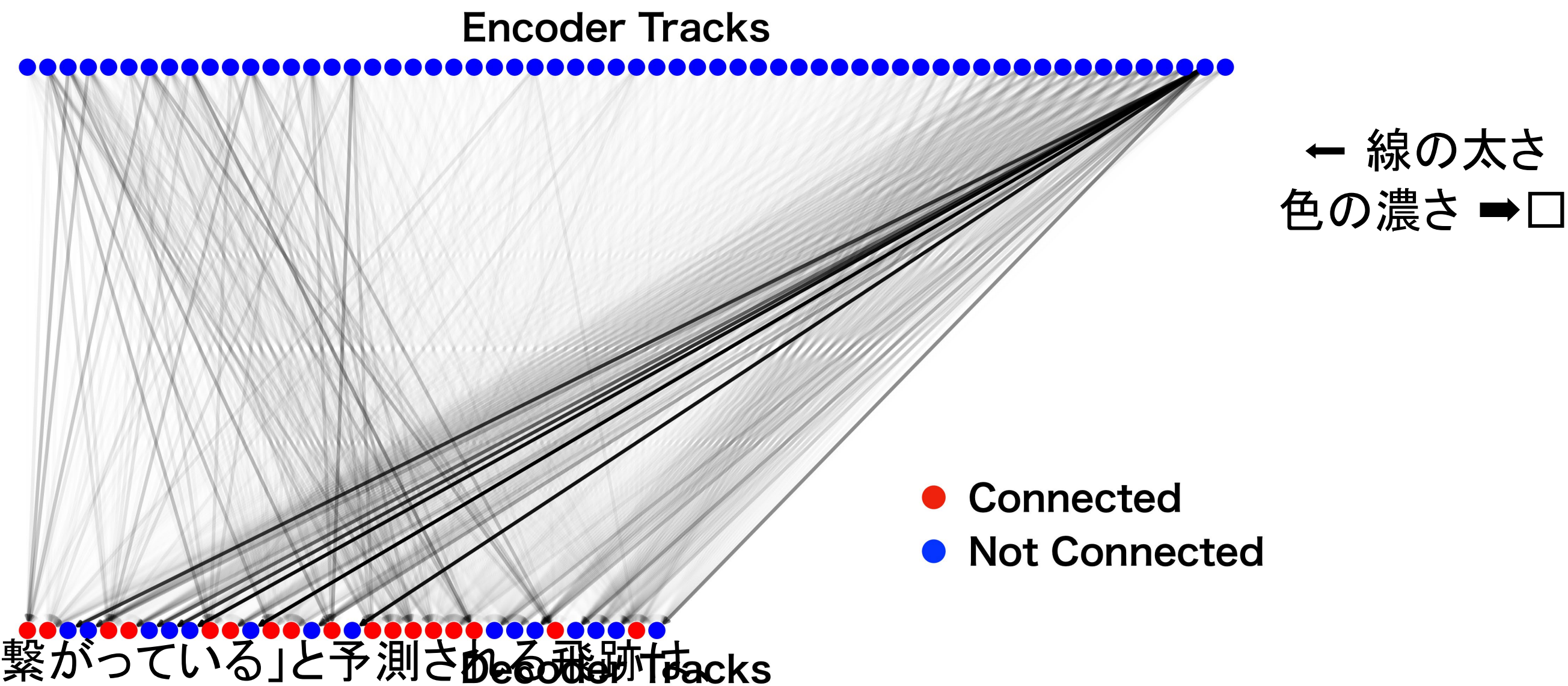


2. 崩壊点検出の為にニューラルネットワーク

Attention

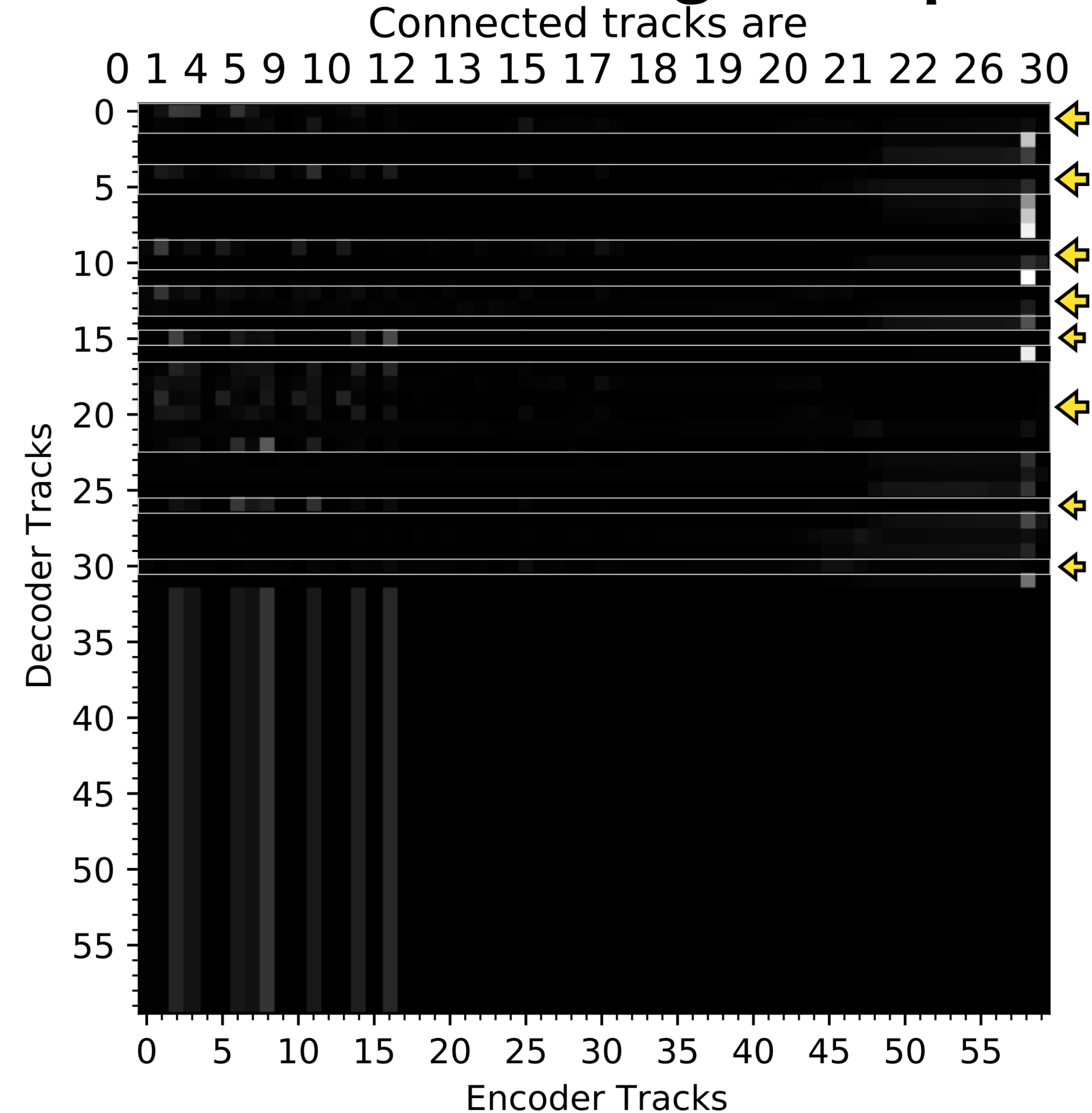
- 各飛跡が事象内の全飛跡に「注意 (Attention)」して欲しい
「どの飛跡」が「どの飛跡」に注意しているか

Attention Weight Graph



- 「繋がっている」と予測される飛跡は
飛跡全体から他の飛跡の情報を受け取れている
- 「どの飛跡」から情報を受け取っているかの調査が必要

Attention Weight Map



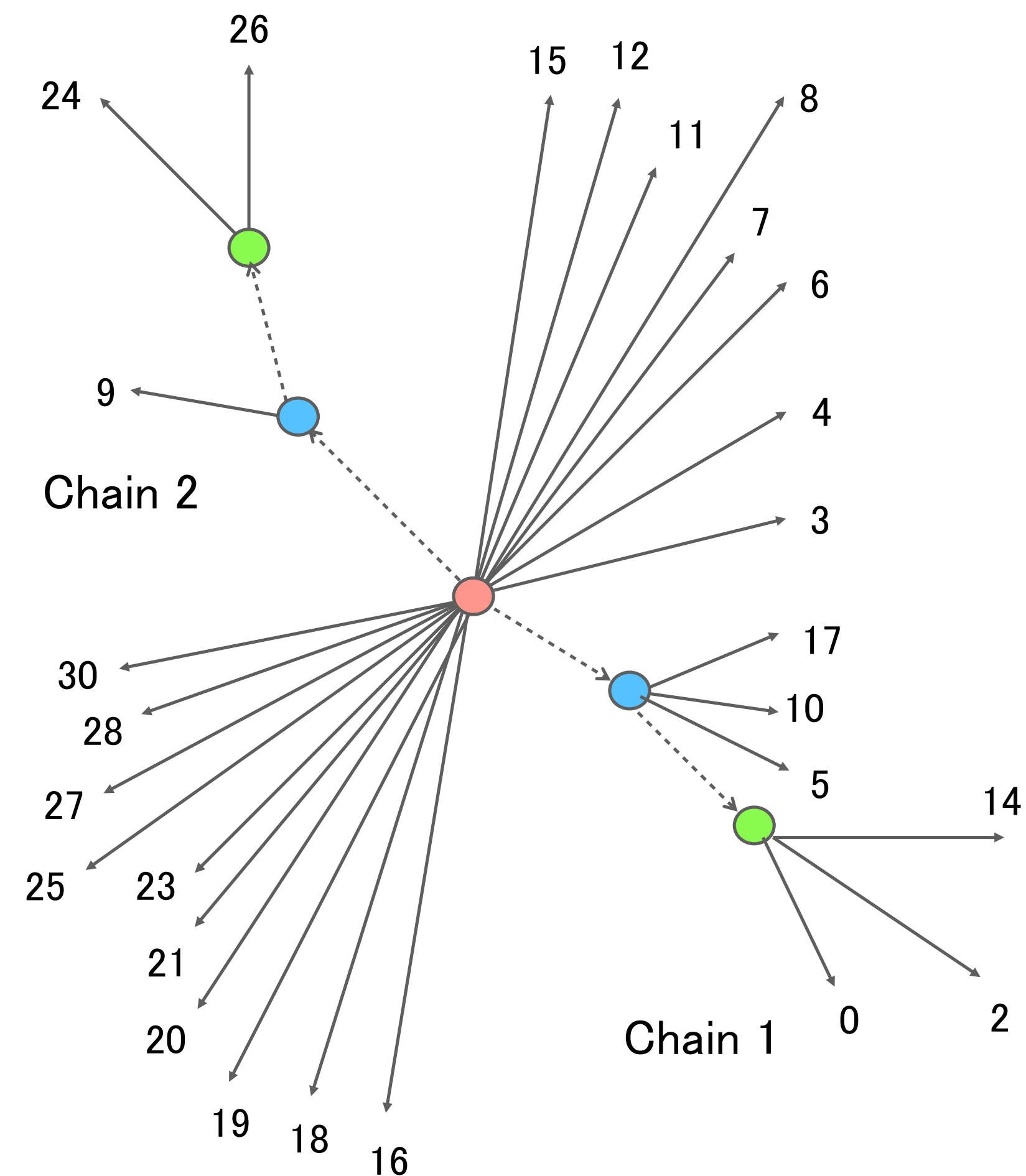
3. 深層学習を用いた崩壊点検出

全飛跡 (31 本)

[0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 27 28 29 30]

```
True Primary Vertex  
[3, 4, 6, 7, 8, 11, 12, 15, 16, 18, 19, 20, 21, 23, 25, 27, 28, 30]  
Predict Primary Vertex  
[3, 4, 6, 7, 8, 11, 12, 15, 16, 18, 19, 20, 21, 23, 25, 27, 28, 29, 30]  
True Secondary Vertex Chain 1  
cc : [0, 2, 14]  
bb : [5, 10, 17]  
one track : []  
True Secondary Vertex Chain 2  
cc : [24, 26]  
bb : []  
one track : [9]  
Predict Secondary Vertex 0  
[24, 26]  
Predict Secondary Vertex 1  
[2, 10]  
Predict Secondary Vertex 2  
[5, 17]  
Predict Secondary Vertex 3  
[0, 14]  
-----  
MC Primary / Reco SV : 0.0  
MC Others / Reco SV : 0.0  
MC Bottom / Reco SV : 1.0 Same Chain : 1.0 Same Particle : 0.6666666666666666  
MC Charm / Reco SV : 1.0 Same Chain : 1.0 Same Particle : 0.8  
-----
```

True Others
[1, 13, 22, 29]



3. 深層学習を用いた崩壊点検出

現行の手法 (LCFIPlus) との比較 - フレーバー識別

