Beyond the SM from Colliders to the Early Universe

May 27 – 30, 2023
U Chicago

1. Configurable Calorimeter simulation for AI – COCOA

2. An Intelligent Bump Hunter or Data Directed Paradigm

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Shikma Bressler
Eilam Gross
Nilotpal Kakati
Amit Shkuri
Etienne Dreyer

Université de Montréal

Bruna Pascual Dias
Fannie Bilodeau
Jean-François Arguin

LPC Particules pLasmas Univers applications
Laboratoire de Physique de Clermont

Samuel Calvet
Julien Noce Donini

TECHNION Israel Institute of Technology

Davide Melini
Enrique Kajomovitz

Tel Aviv - Chicago - Tel Aviv

1. Configurable Calorimeter simulation for AI – COCOA

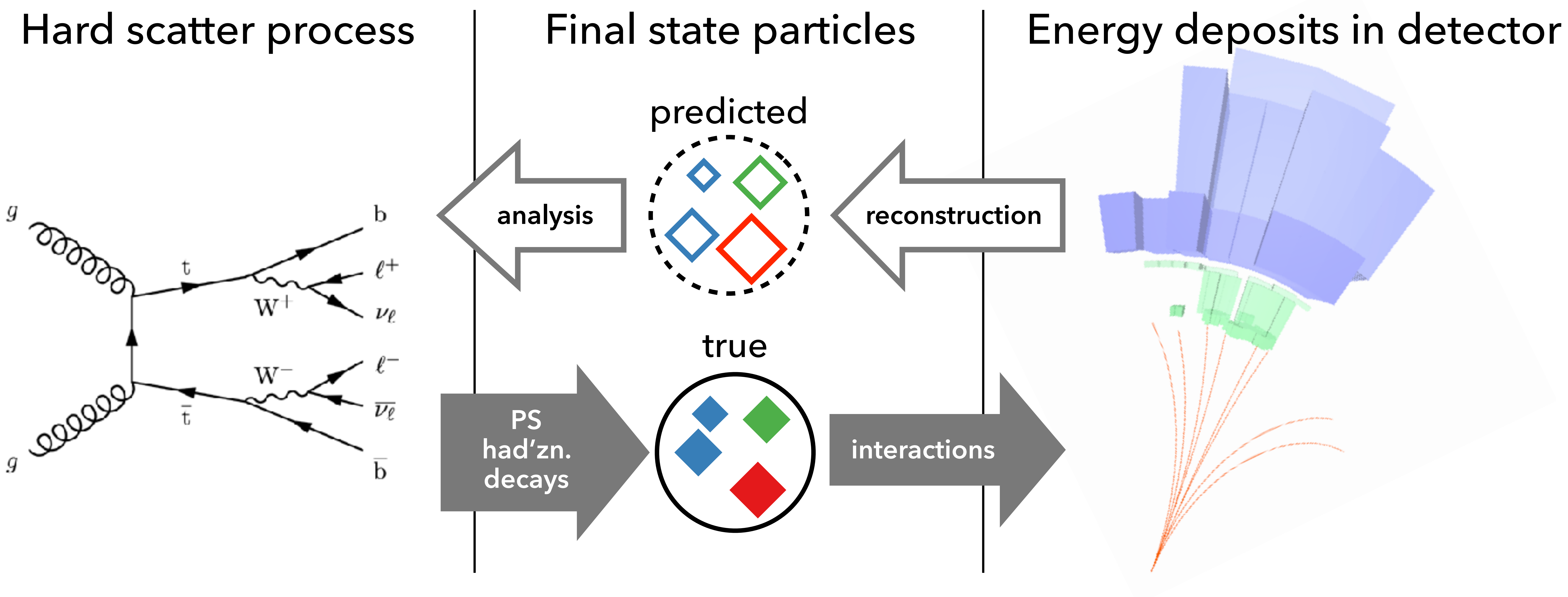
Advertising COCOA

Configurable calorimeter simulation for AI applications

Francesco Armando Di Bello [1], Anton Charkin-Gorbulin [2], Kyle Cranmer [4,5], Etienne Dreyer [3,c], Sanmay Ganguly [6,a], Eilam Gross [3], Lukas Heinrich [7], Lorenzo Santi [9], Marumi Kado [8,9], Nilotpal Kakati [3], Patrick Rieck [4,b], Matteo Tusoni [9]

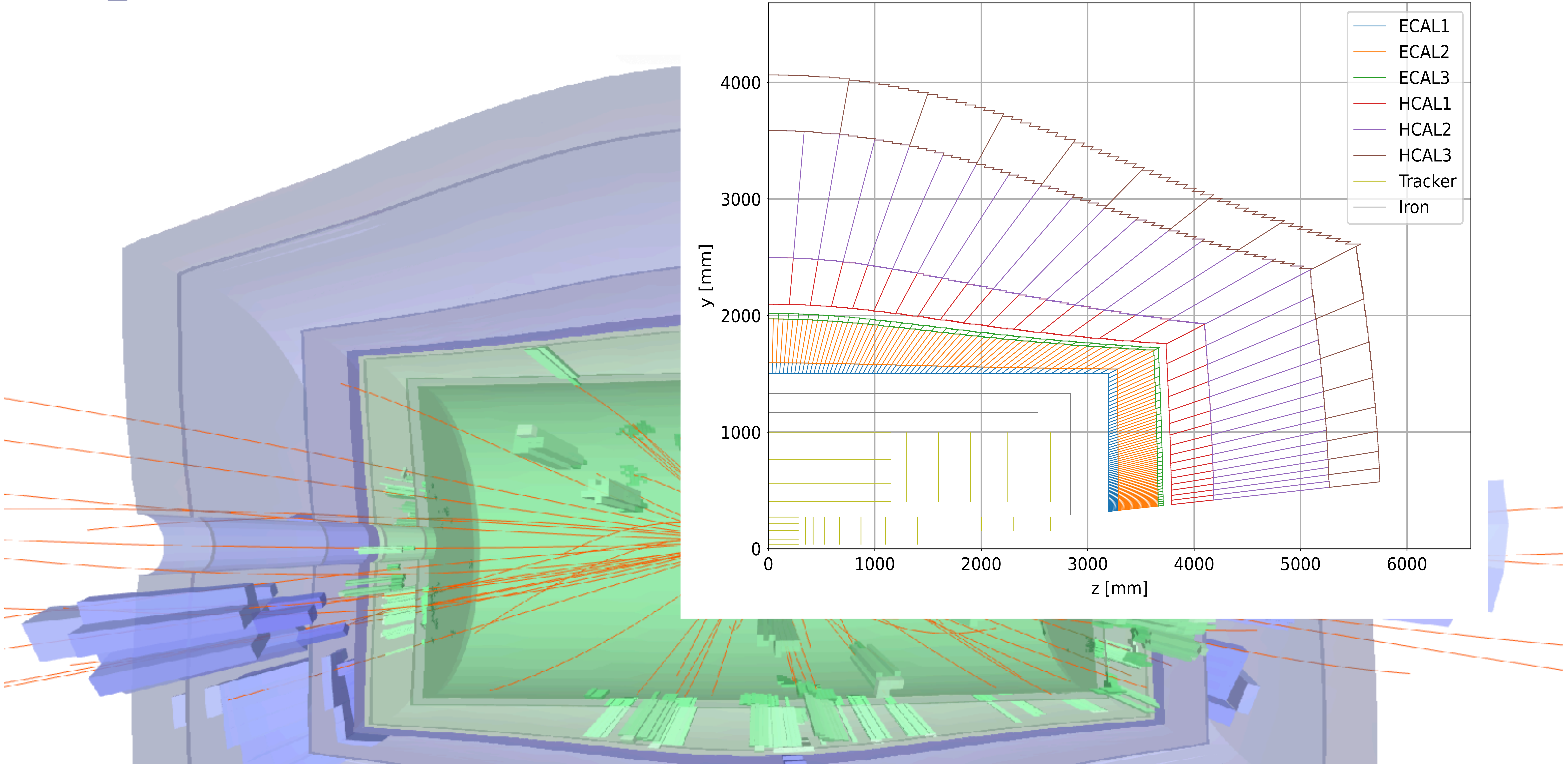
Under review in MLST

Particle reconstruction



Infer the **set** of particles which produced
the **set** of energy deposits in detector

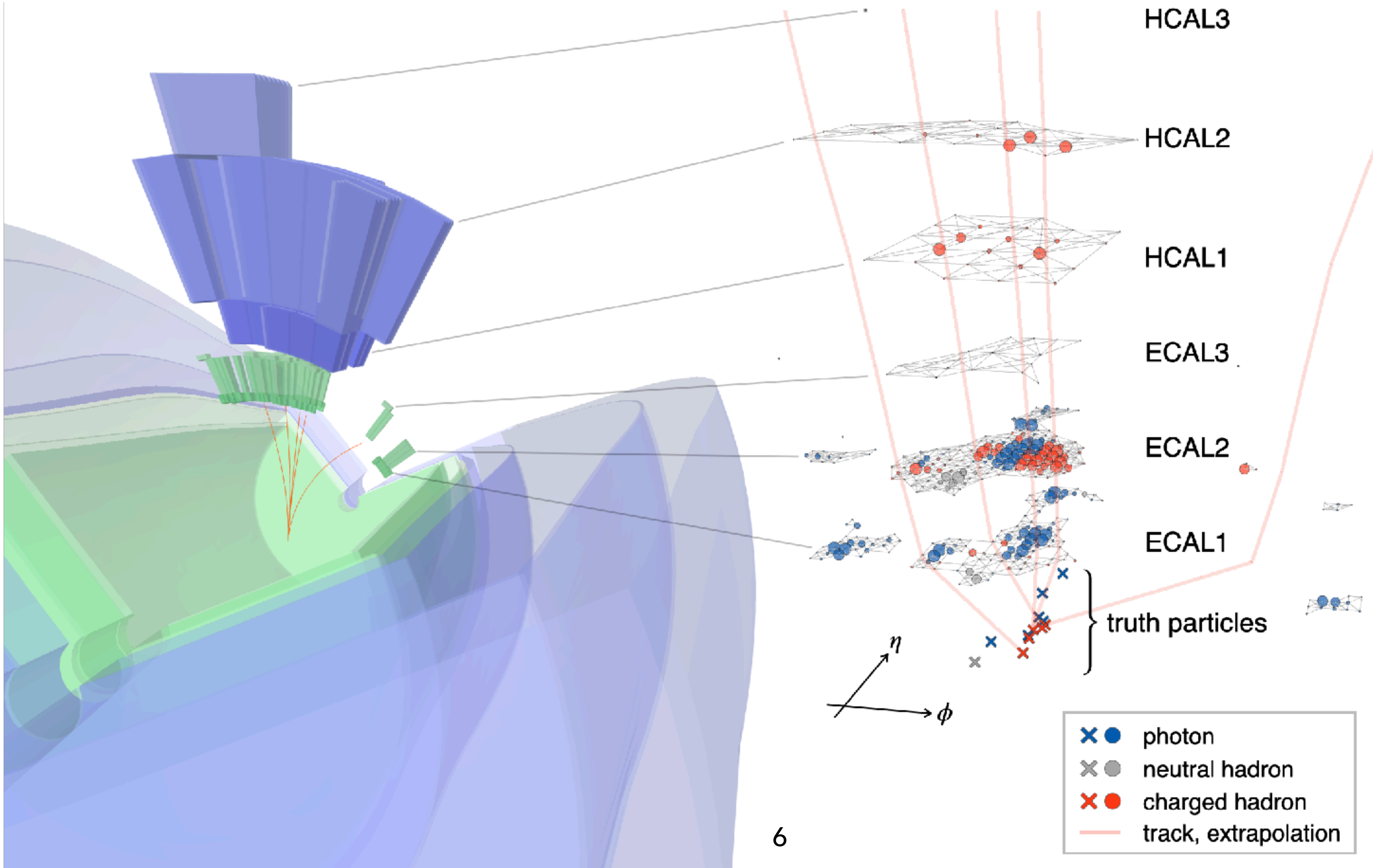
Open calorimeter model



ATLAS-like calorimeter simulation

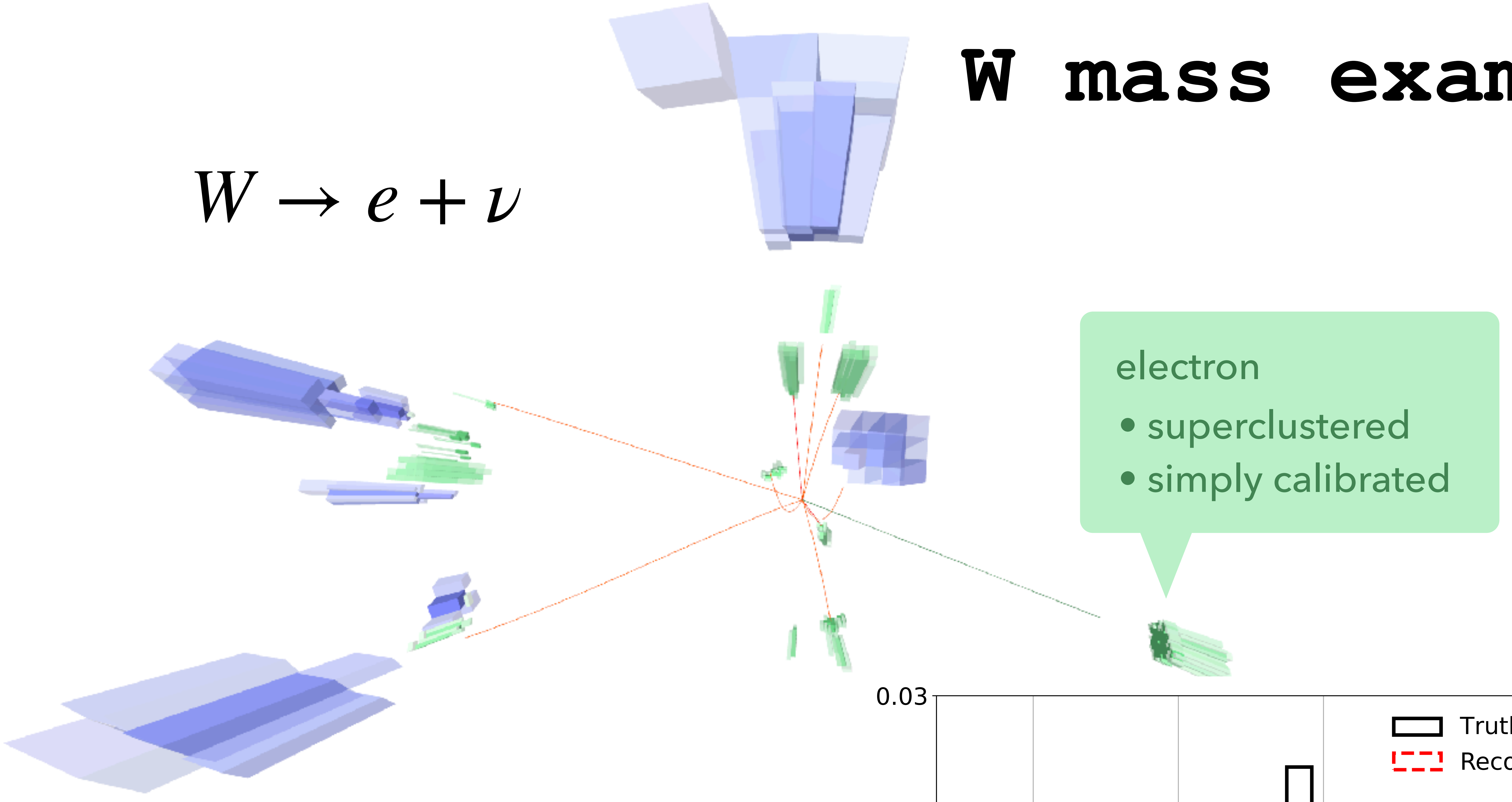
- 3 ECAL + 3 HCAL concentric GEANT4 calorimeter layers
- Interfaced to Pythia8 event generator
- Tracking emulation in 3.8T magnetic field

Graph illustration (single jet)



W mass example

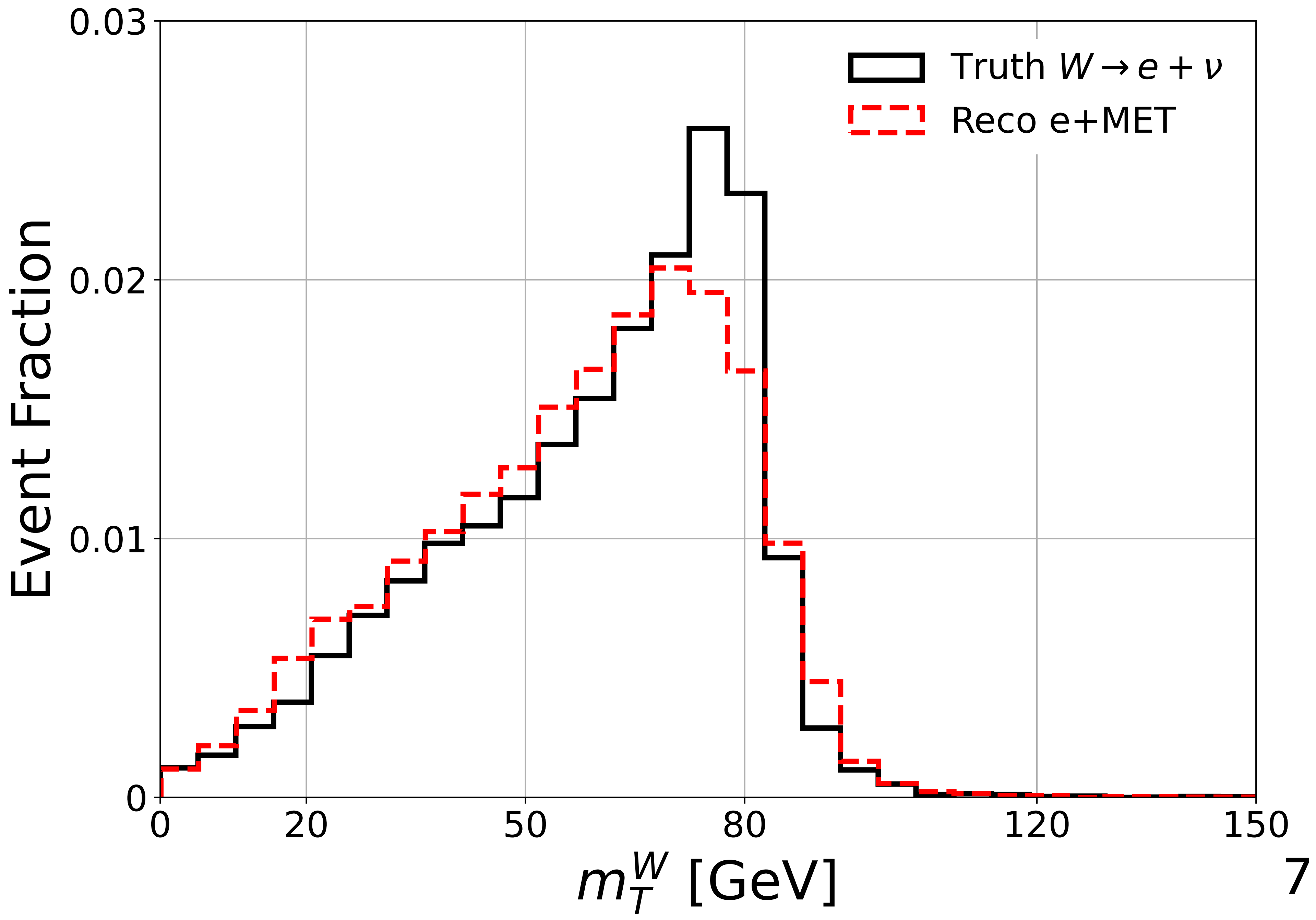
$W \to e + \nu$



Transverse mass:

$$m_{\mathrm{T}}^2 = 2p_{\mathrm{T}}^{(e)} E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\varphi)$$

⇒ demonstrates good overall response at event-level (i.e. $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction)

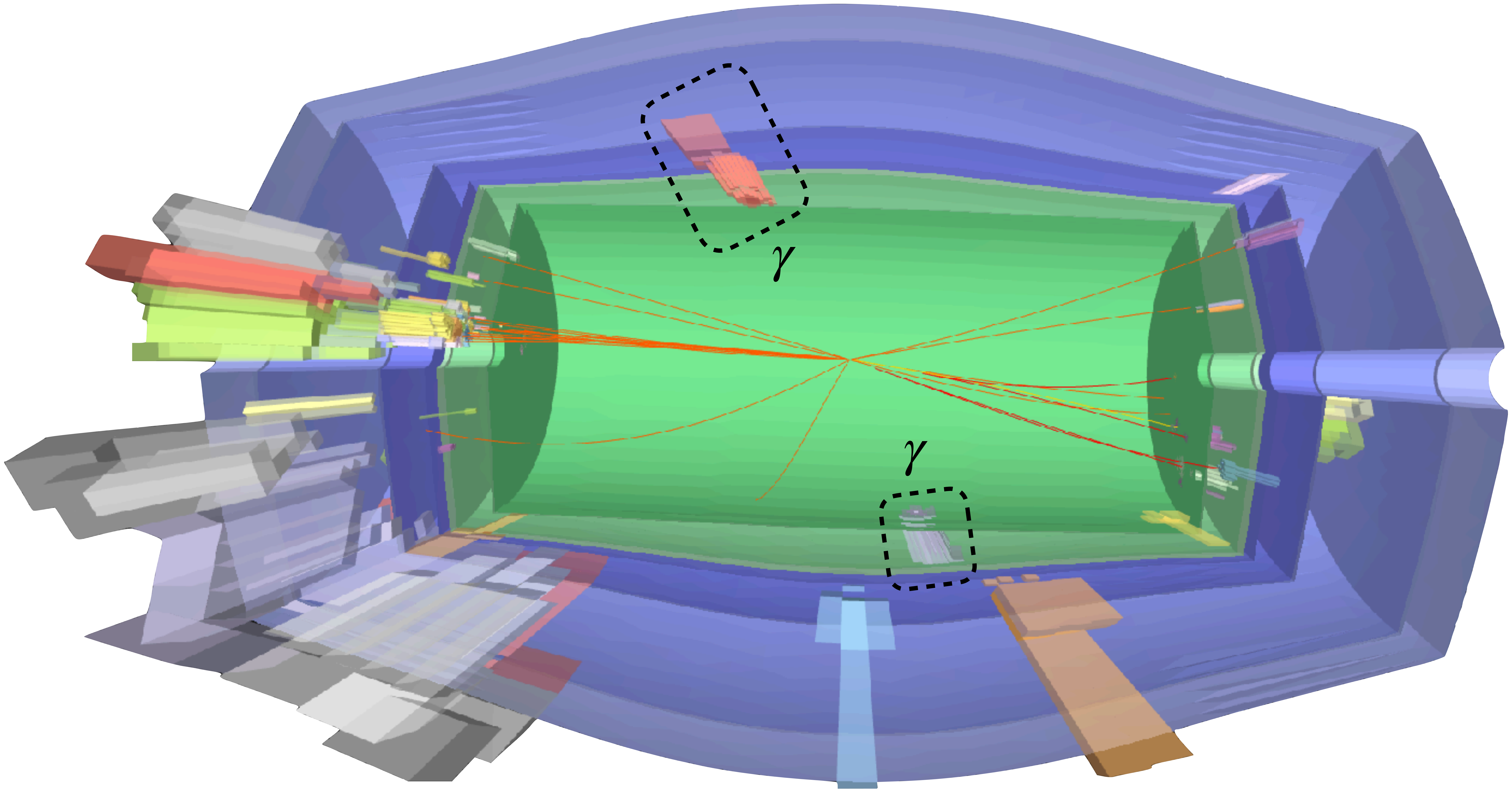


Comparison

	Delphes	COCOA
Goal	**Speed**, realistic treatment of main observables	**Details**, GEANT-based microscopic simulation
Primary use	Smear truth particles to resemble reco objects	Fully-featured datasets for training ML models
• Configurable • Interface to Pythia • Parameterized track smearing • Photon conversions • Jet clustering algo. • Event display	Y	Y
Fully simulated calorimeter shower	N	Y
Specific LHC experiment tunes	Y	Possible
Electron reco	track & cluster	supercluster (Brehm)
Particle flow algorithm	Basic subtraction	Parameterized (ATLAS-like)
Pileup	Y	Possible
Jet substructure	Parameterized	Raw

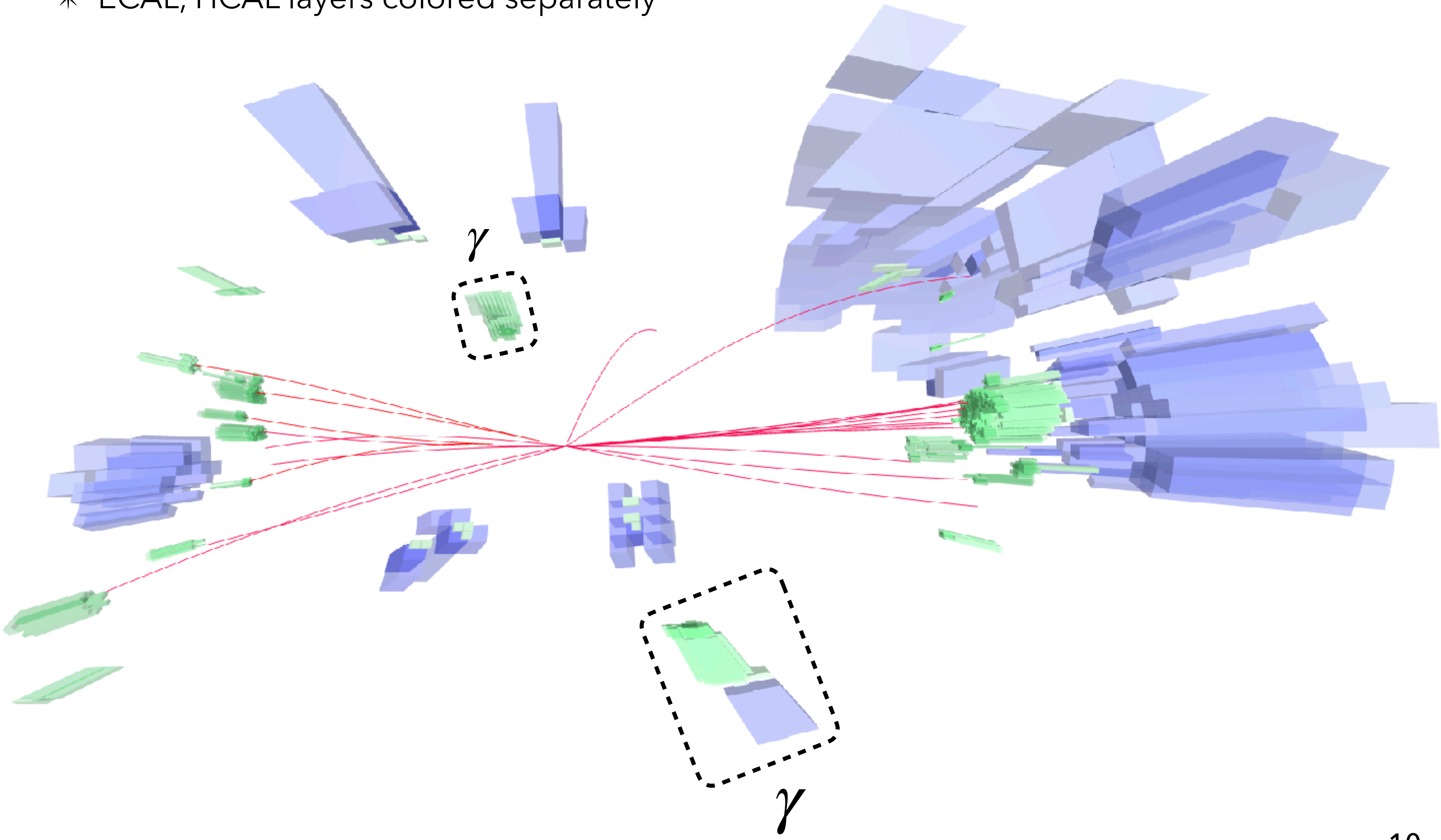
VBF H $\rightarrow\gamma\gamma$

- Forward jets
- Topoclusters colored separately



VBF H → $\gamma\gamma$

- Forward jets
- ECAL, HCAL layers colored separately



Quick start guide

https://cocoa-hep.readthedocs.io/

https://github.com/cocoa-hep/cocoa-hep

```
docker pull ghcr.io/cocoa-hep/cocoa-hep:main
docker image tag \
   $(docker images | grep cocoa-hep | head -n 1 | awk '{print $3}') \
   cocoa-hep
docker run -it cocoa-hep
```

```
git clone git@github.com:cocoa-hep/cocoa-hep.git; cd cocoa-hep/COCOA
source ../setup_cvmfs.sh
mkdir build; cd build
cmake ../
make
cd ..
```

```
./build/COCOA \
       --macro  macro/Pythia8/ttbar.in \
       --config config/config_default.json
```

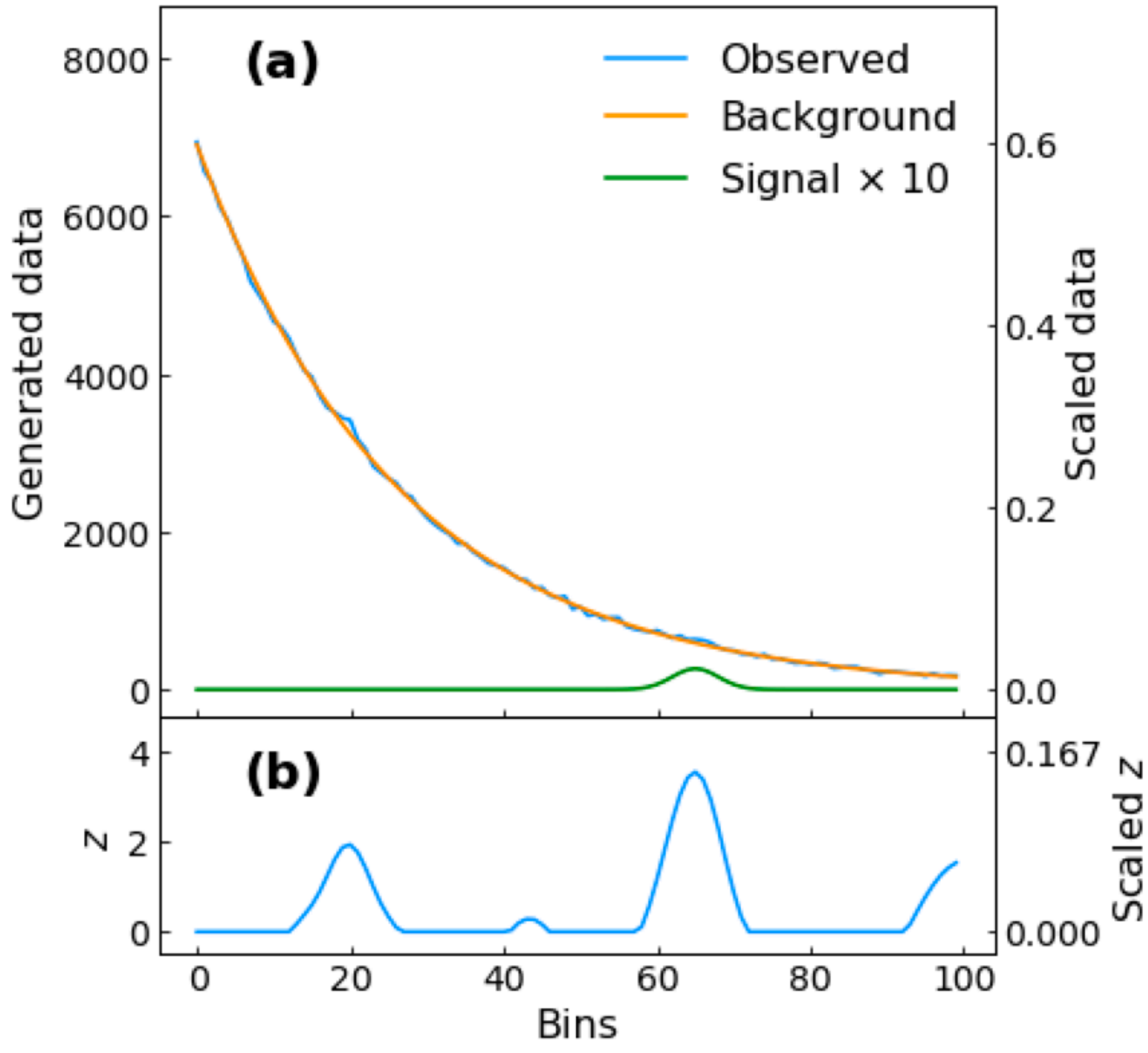
2. An Intelligent Bump Hunter or Data Directed Paradigm

IBH in a Nut Shell

- Our vision is an accessible, fast and reliable bump hunter that is **independent of the expected background knowledge**

- First attempt (proof of concept) Volkovich, De Vito Halevy, Bressler https://arxiv.org/abs/2107.11573, EPJC

- This talk: Introducing a novel improved and Intelligent Deep Learning architecture (WIP).



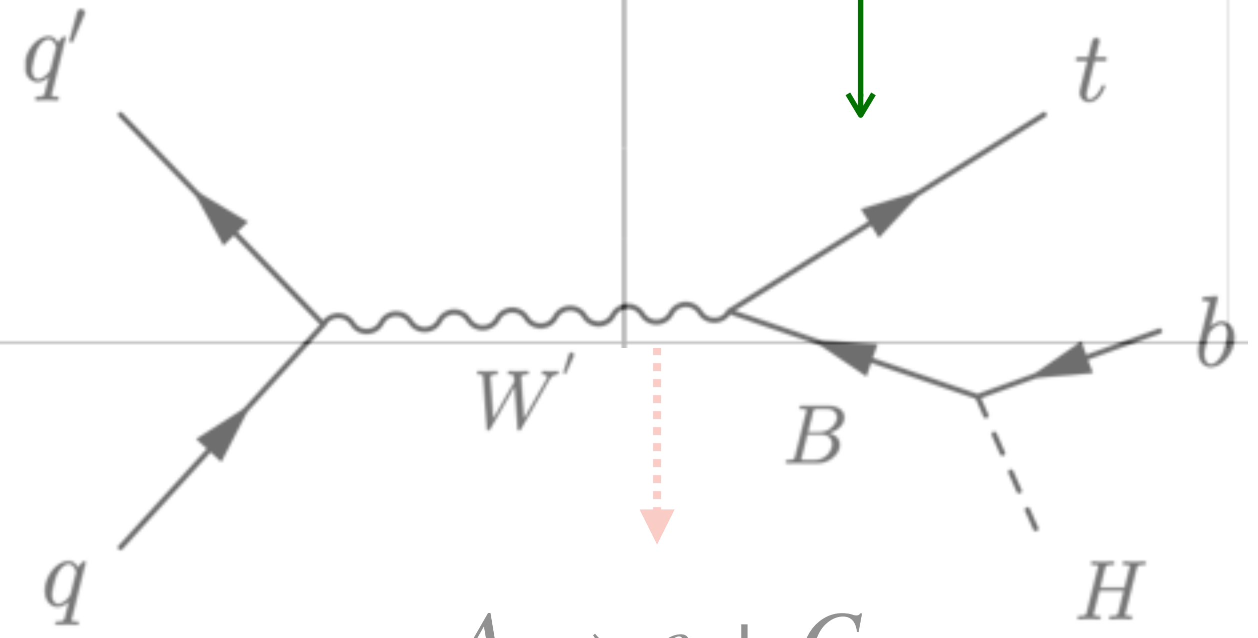
The motivation and status of two-body resonance decays after the LHC Run 2 and beyond

Jeong Han Kim, Kyoungchul Kong, Benjamin Nachman, Daniel Whiteson

Motivation

	e	μ	τ	q/g	b	t	γ	Z/W	H	BSM $\to \mathrm{SM}_1 \times \mathrm{SM}_1$: q/g	γ/π^0's	b	$\cdots$	BSM $\to \mathrm{SM}_1 \times \mathrm{SM}_2$: tZ/H	bH	$\cdots$
e	[37, 38]	[39, 40]	[39]	∅	∅	∅	[41]	[42]	∅	∅	∅	∅		∅	∅	∅
μ		[37, 38]	[39]	∅	∅	∅	[41]	[42]	∅	∅	∅	∅		∅	∅	∅
τ			[45, 46]	∅	[47]	∅	∅	∅	∅	∅	∅	∅		∅	∅	∅
q/g				[29, 30, 50, 51]	[52]	∅	[53, 54]	[55]	∅	∅	∅	∅		∅	∅	∅
b					[29, 52, 56]	[57]	[54]	[58]	[59]	∅	∅	∅		[60]	∅	∅
t						[61]	∅	[62]	[63]	∅	∅	∅		[64]	[60]	∅
γ							[65, 66]	[67–69]	[68, 70]	∅	∅	∅		∅	∅	∅
Z/W								[71]	[71]	∅	∅	∅		∅	∅	∅
H									[72, 73]	[74]	∅	∅		∅	∅	∅
BSM $\to \mathrm{SM}_1 \times \mathrm{SM}_1$: q/g										∅	∅	∅		∅	∅	∅
γ/π^0's											[75]	∅		∅	∅	∅
b												[76, 77]		∅	∅	∅
⋮																
⋮																

- Need a rapid scan of the data in many regions of interest



$A \to e + C$

$C \to q + q$

$A \to e + C$

$C \to t + Z/H$

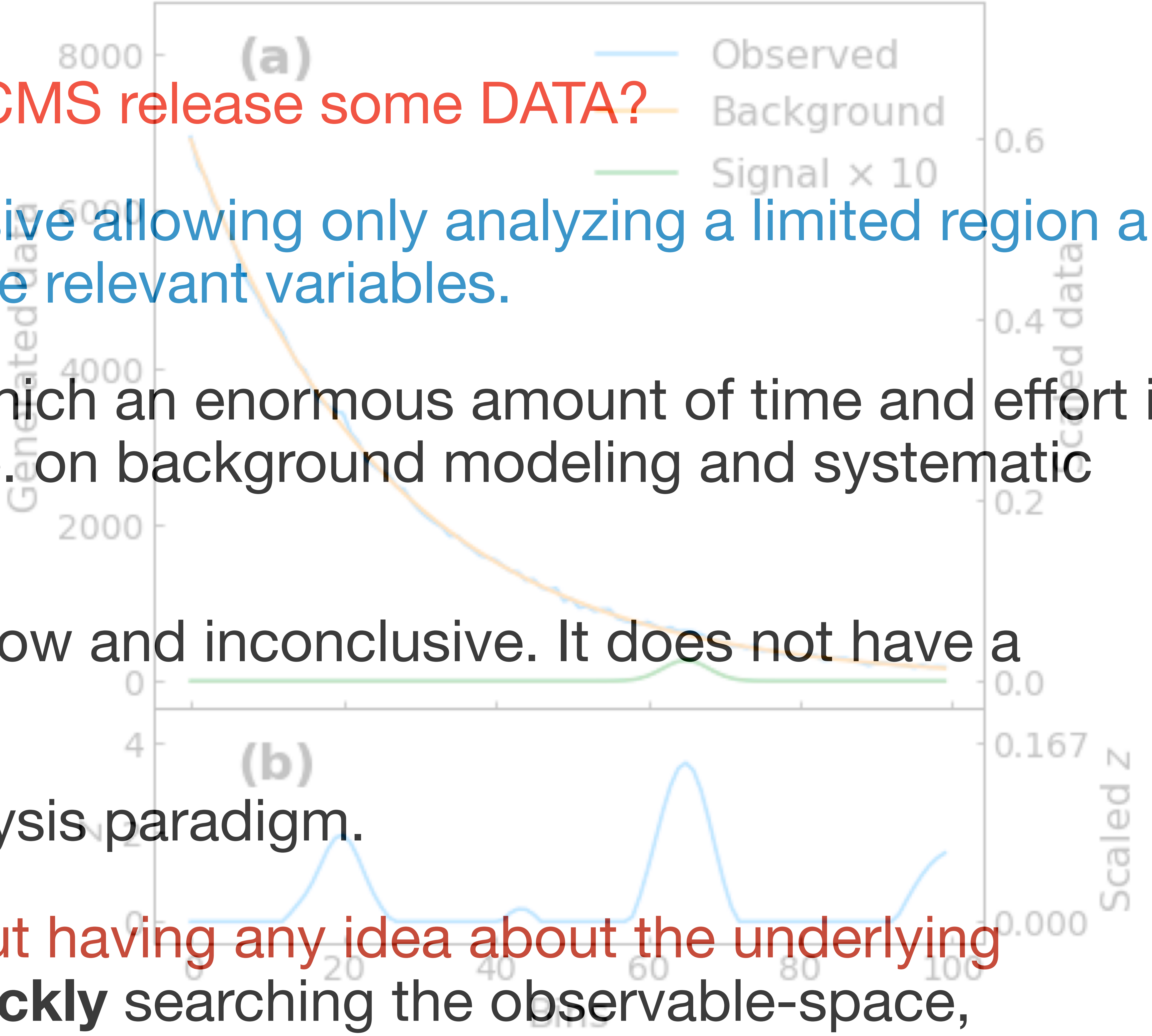
$A \in BSM \to B + C$

1) $B, C \in SM$

2,3) $B \in SM, C \in BSM$

Why DDP with IBH?

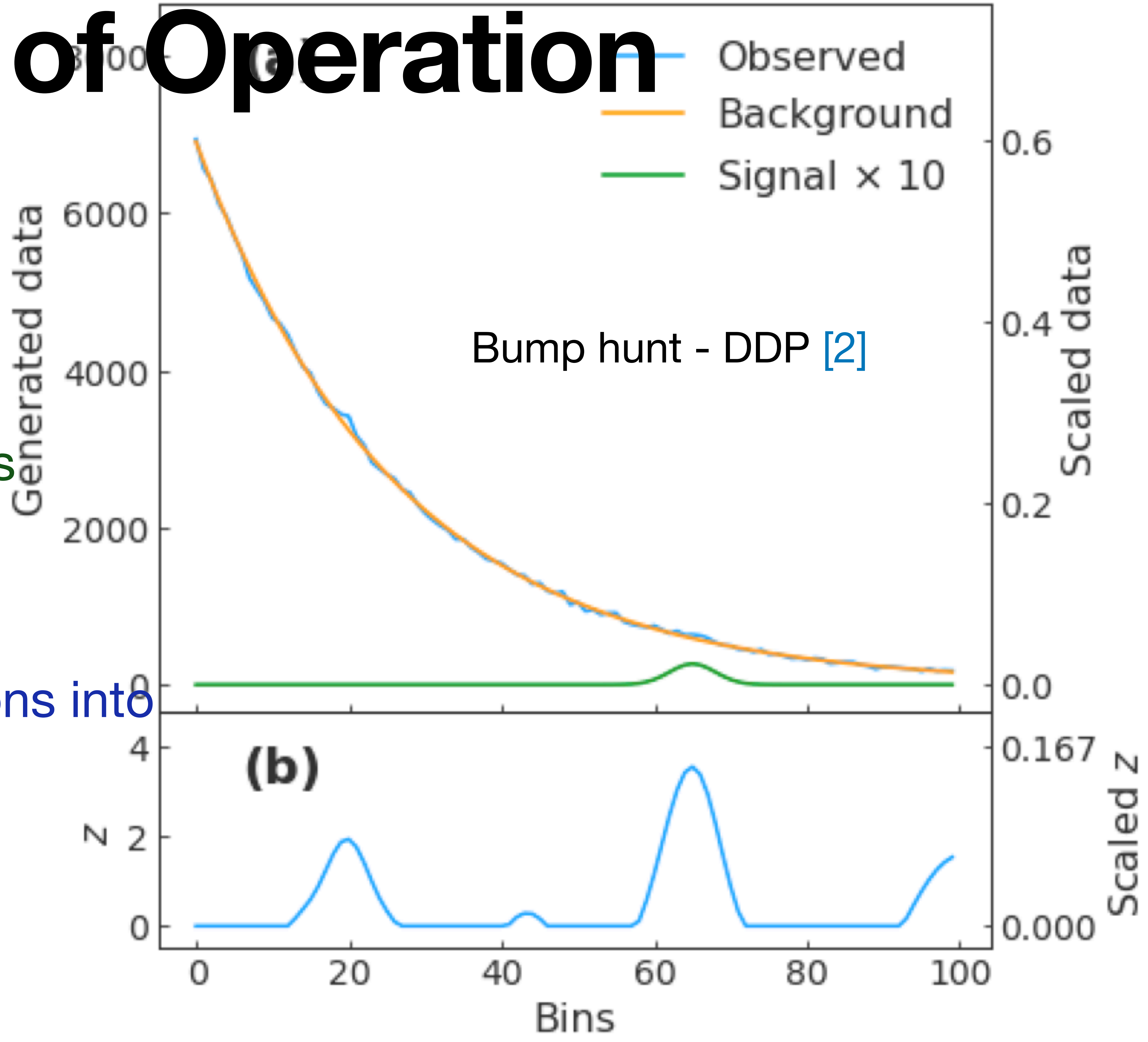
- Isn't that what you do once ATLAS and CMS release some DATA?
- The traditional method is resource intensive allowing only analyzing a limited region a the DATA Phase-Space spanned by some relevant variables.
  - Using a blind analysis paradigm, in which an enormous amount of time and effort is invested before looking at the data, i.e. on background modeling and systematic uncertainty estimation.
- The traditional analysis method is very slow and inconclusive. It does not have a "coverage"of all interesting anomalies
- DDP is complementary to the Blind Analysis paradigm.
- Without using MC simulation, and without having any idea about the underlying background, the strategy consists of **quickly** searching the observable-space, for exclusive regions exhibiting a significant deviation from some fundamental SM property



IBH: Mode of Operation

- Assumption: The invariant mass distributions is smoothly falling in the absence of resonances

- We train a NN to map invariant mass distributions into significance distributions
  WITHOUT A-PRIORI KNOWLEDGE OF THE BACKGROUND



Bump hunt - DDP [2]

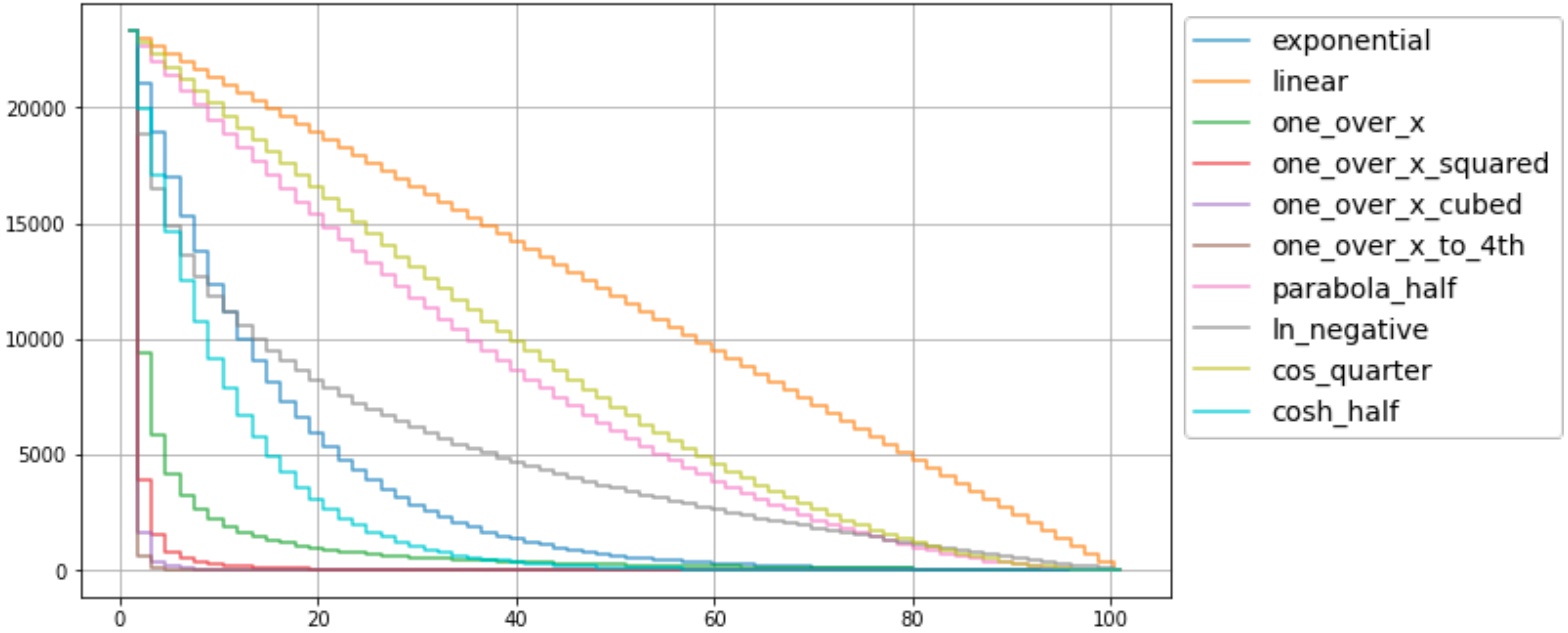
[2] https://arxiv.org/abs/2107.11573

The Bakground Functions Reference

$$be^{-ax},\quad ax+b,\quad \frac{1}{ax}+b,\quad \frac{1}{ax^2}+b,\quad \frac{1}{ax^3}+b,$$
$$\frac{1}{ax^4}+b,\quad a\left(x-x_2\right)^2+y_2,\quad -a\cdot\ln\left(x\right)+b,$$
$$\left(y_1-y_2\right)\cos\left(a\left(x-b\right)\right)+y_2,\quad \cosh\left(a\left(x-x_2\right)\right)+b.$$

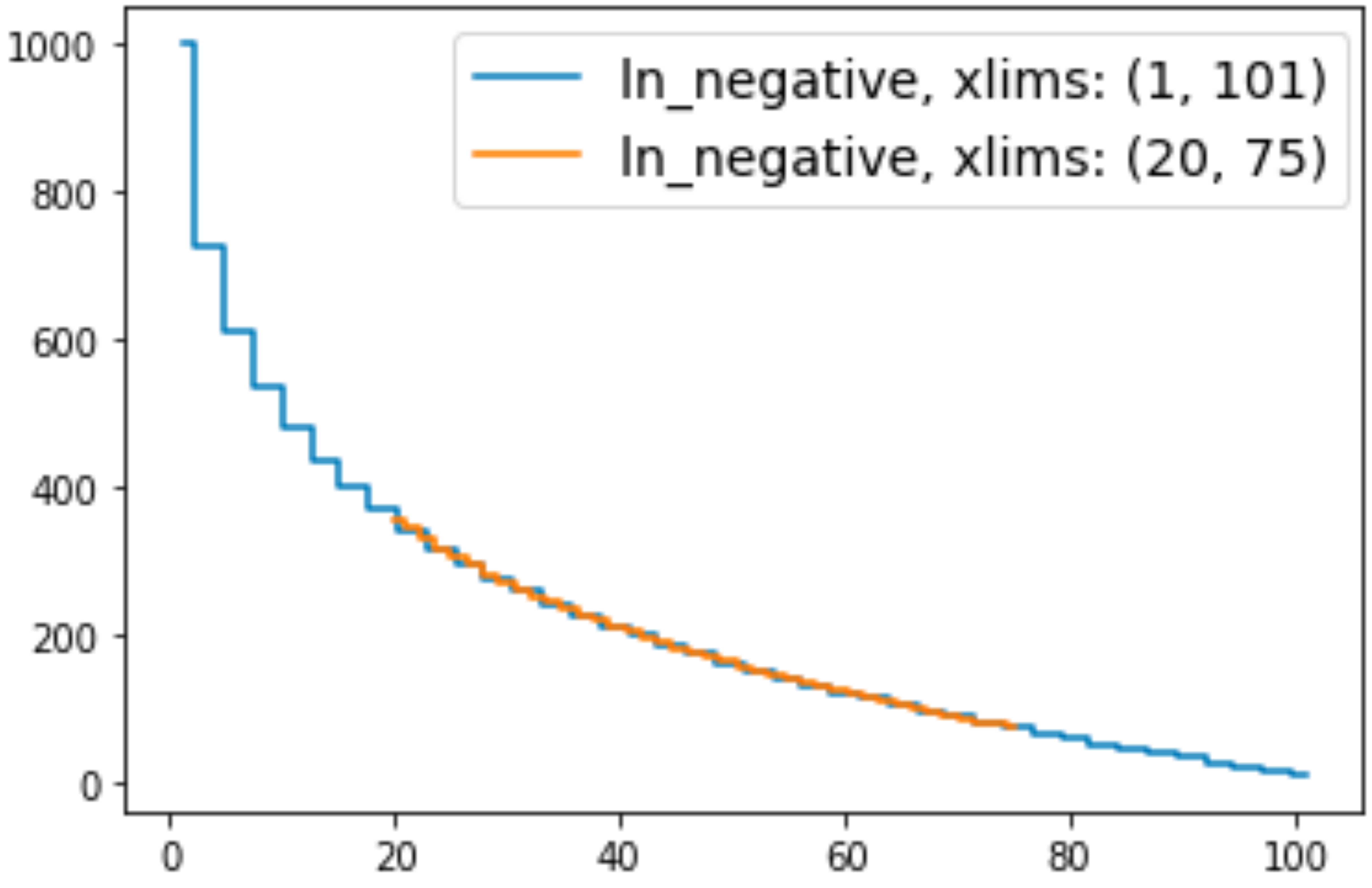
The parameters a and b are defined such that each curve decays between two points, (x1; y1) and (x2; y2), where x1 < x2 are the centers of the extreme bins and y1 > y2 .are randomized from the interval [100,10000]

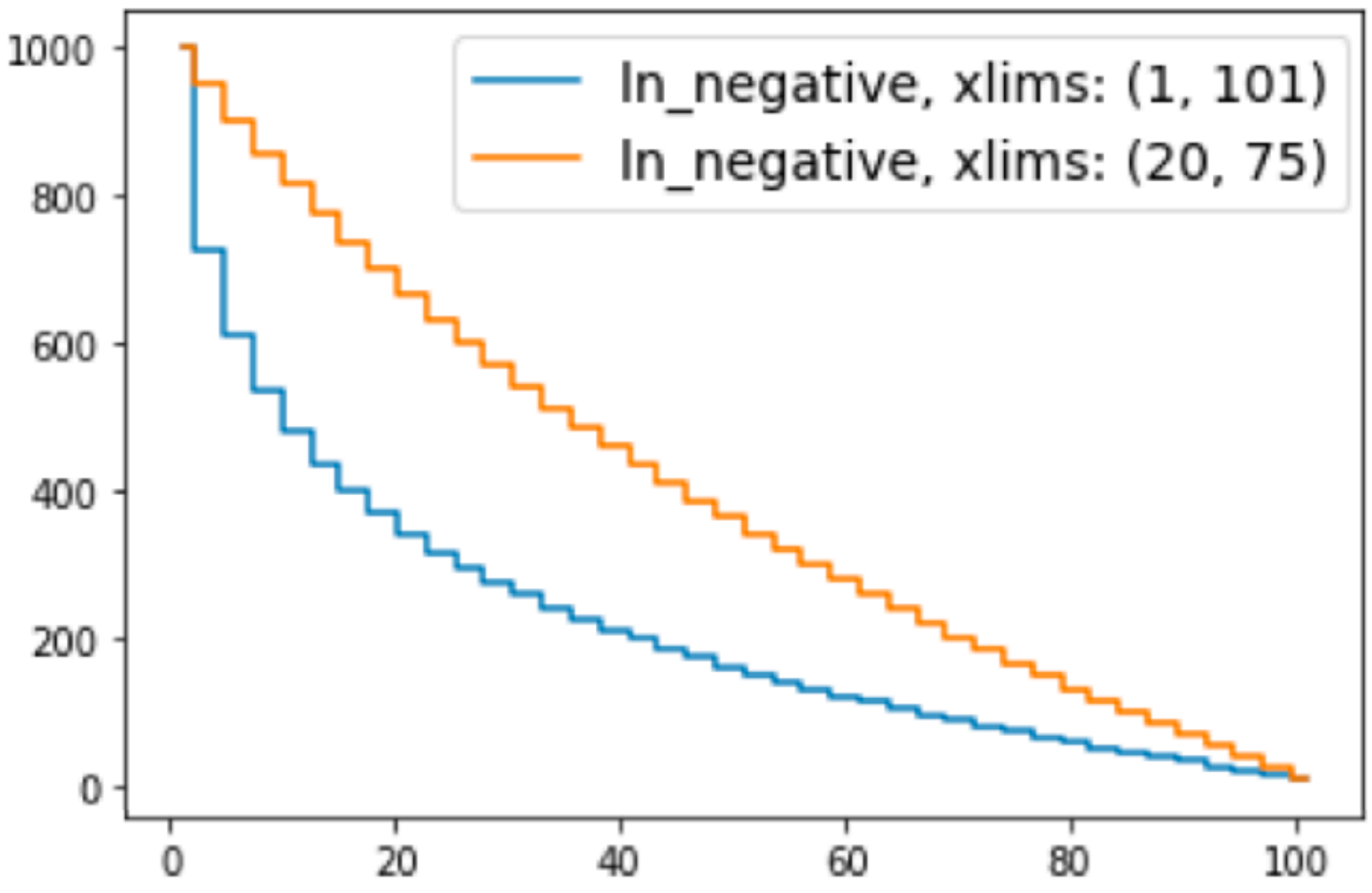
- Smoothly Falling Functions



Generating Functions from the Reference

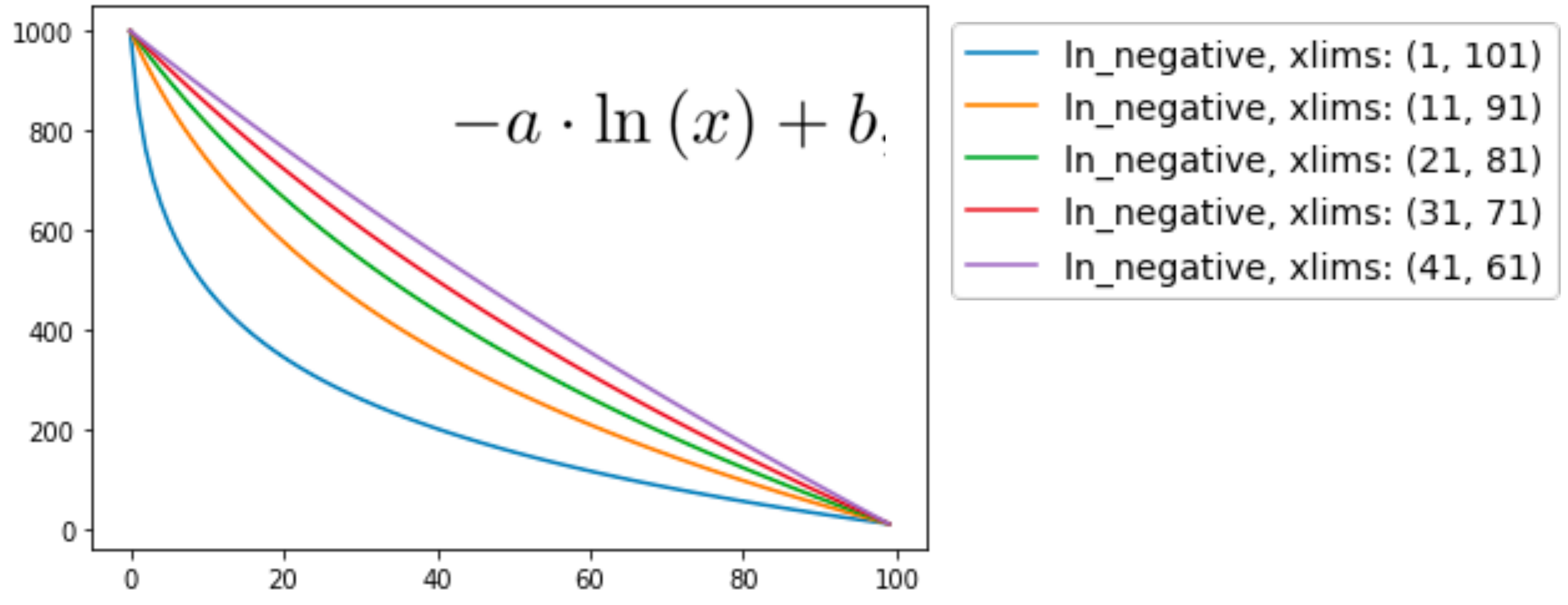
orange: same function is stretched from 20-75—>1-101

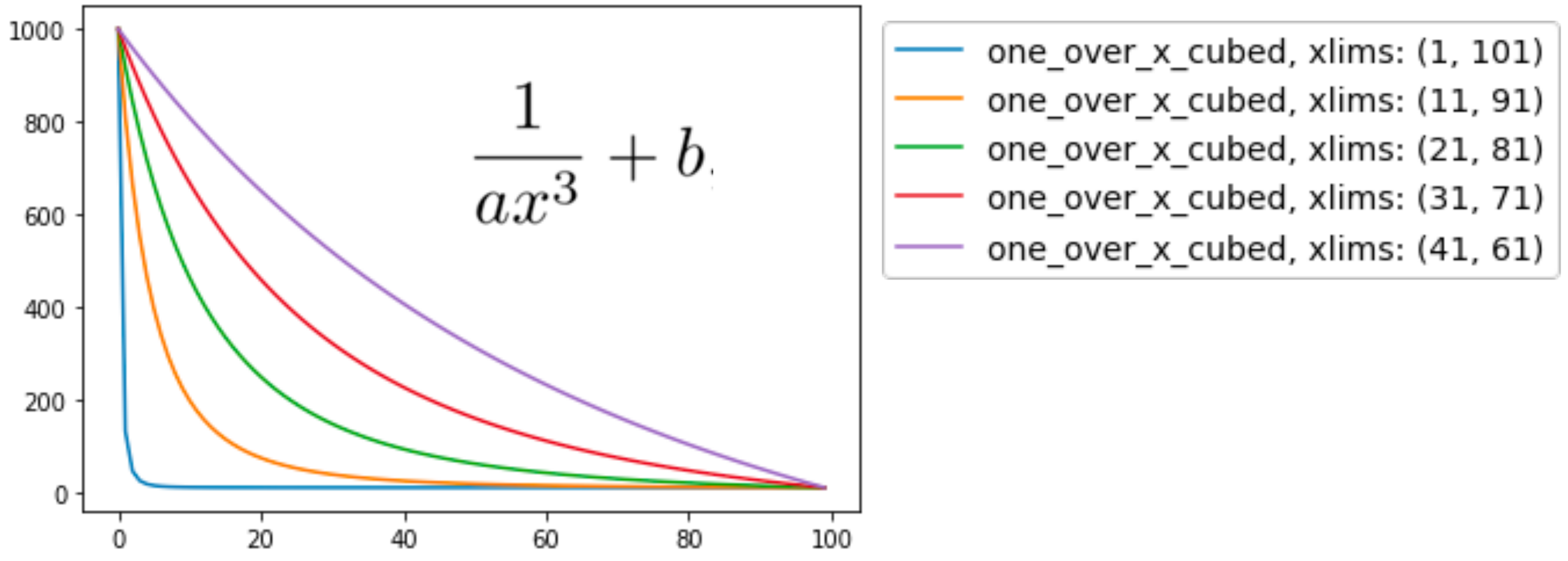


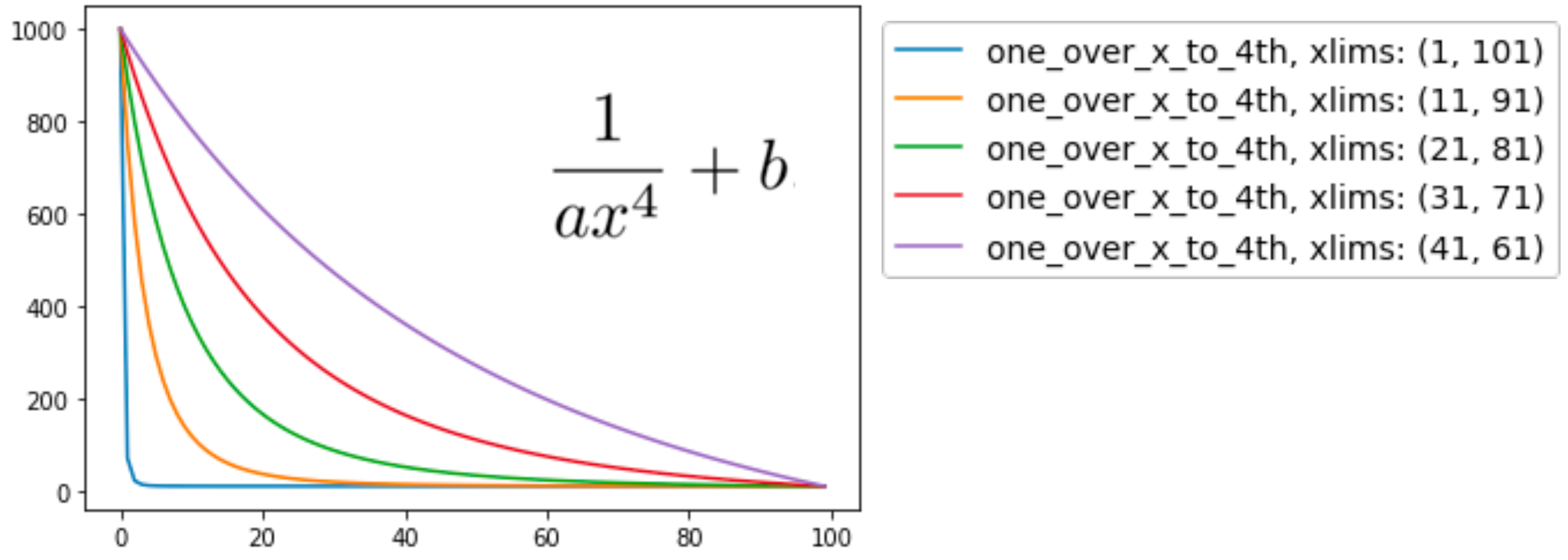


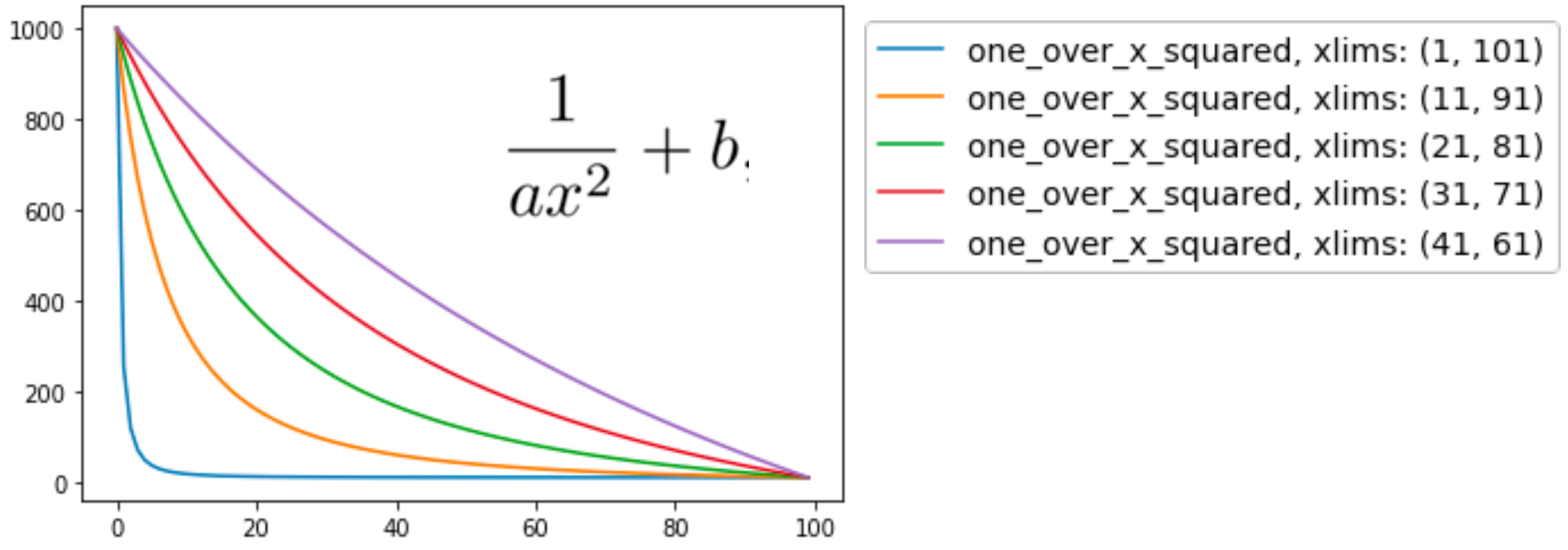
Generating Functions from the Reference

Each plot represents variations of a specific function





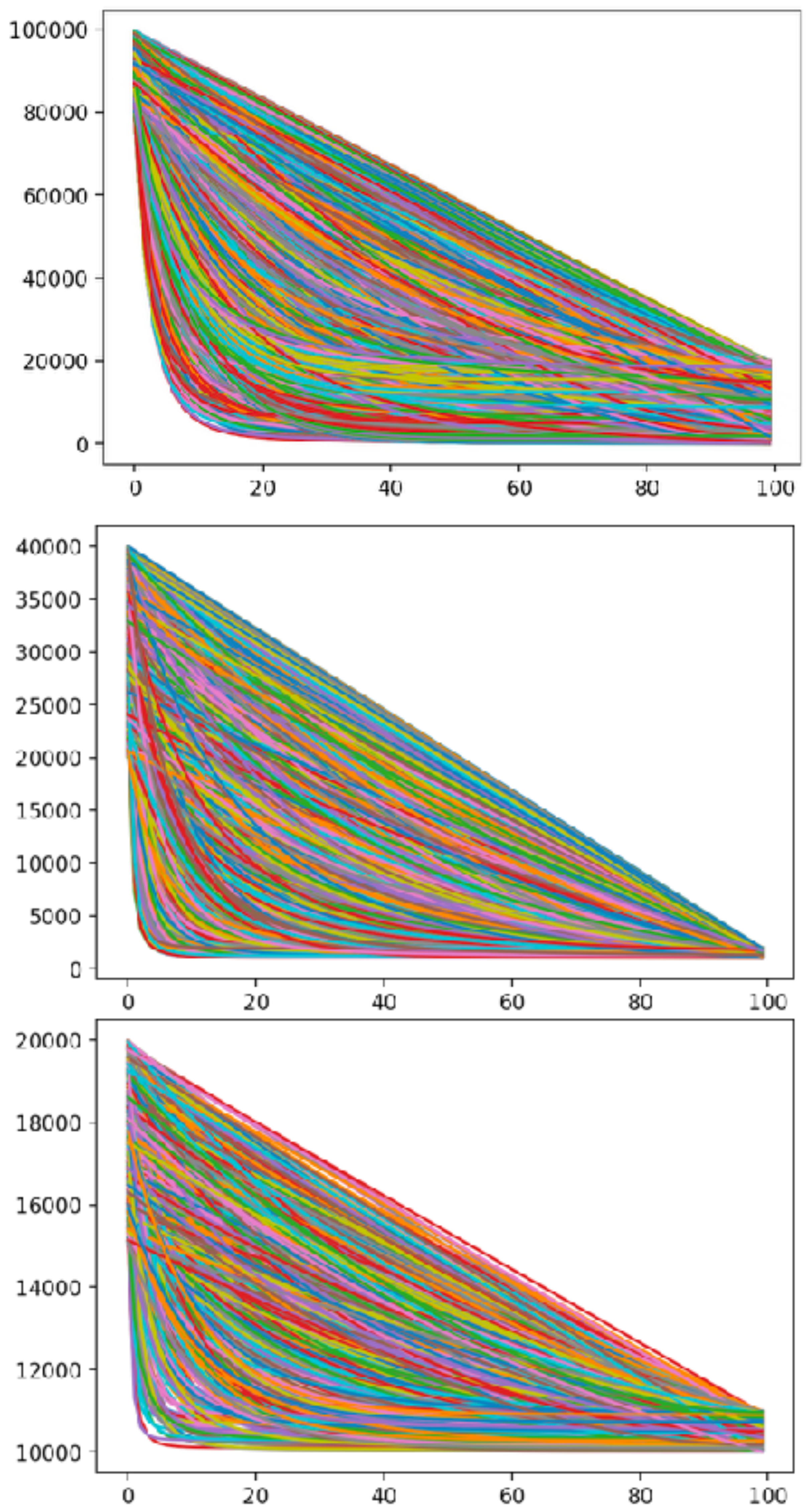




Data Set - Synthetic Data

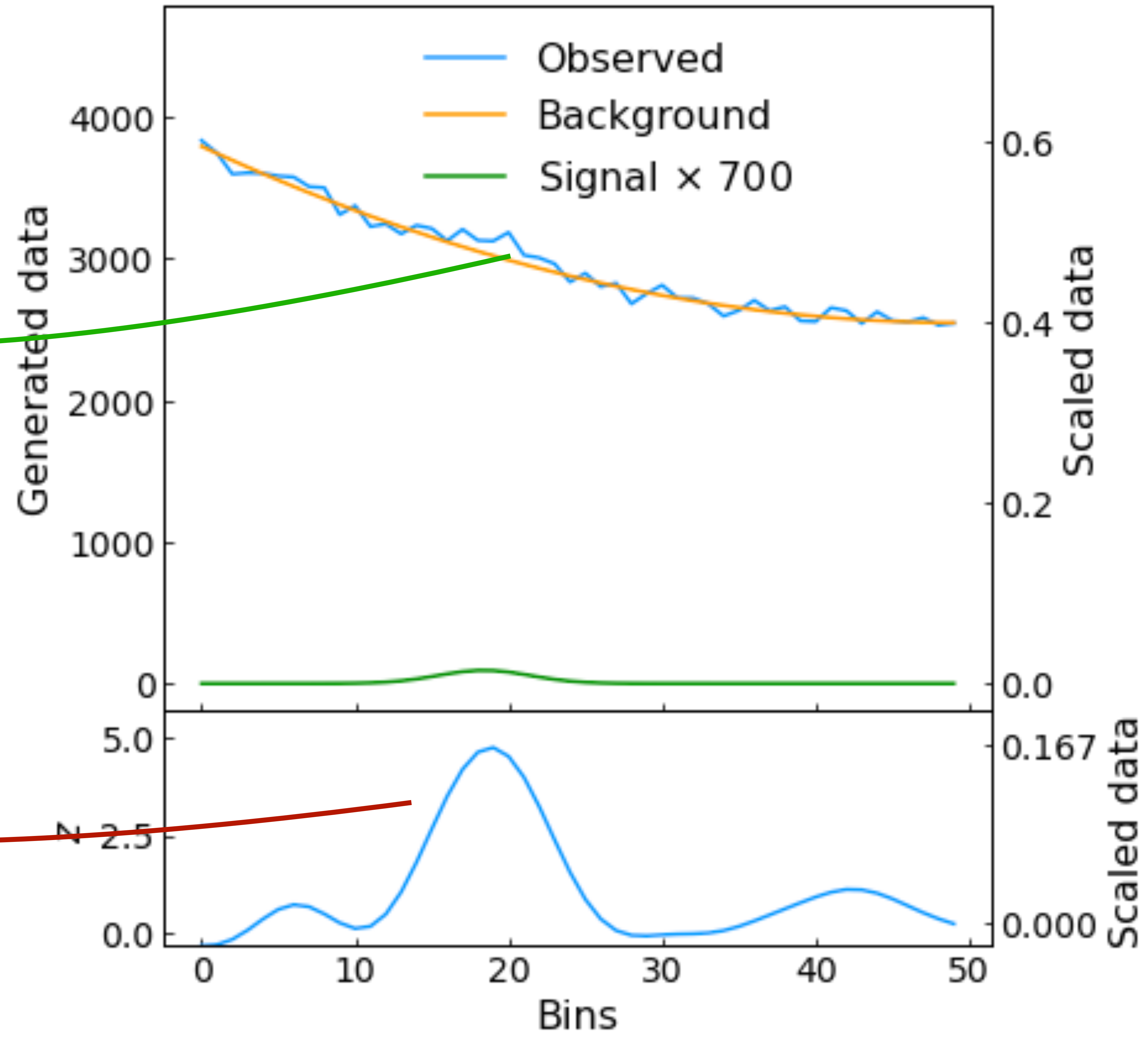
Use the Fnctions Basis to generate a wealth of smoothly falling functions:⁻

1.Toss a dynamic range between 10-100,000
2. Toss a range between 10-100
3.Stretch the function to cover the full range 0-100
4.Bin in 30,40,50,60,….,100 bins



Network Input and Target

Bump hunt - DDP [2]



- Network **input** - Synthetic distributions derived by Poisson fluctuation of smoothy falling functions dressed with some injected Signal
- Network **target** - Bin by bin significance calculated using Profile Likelihood Ratio test statistic relative to the smooth **known** background -> An ideal analysis output

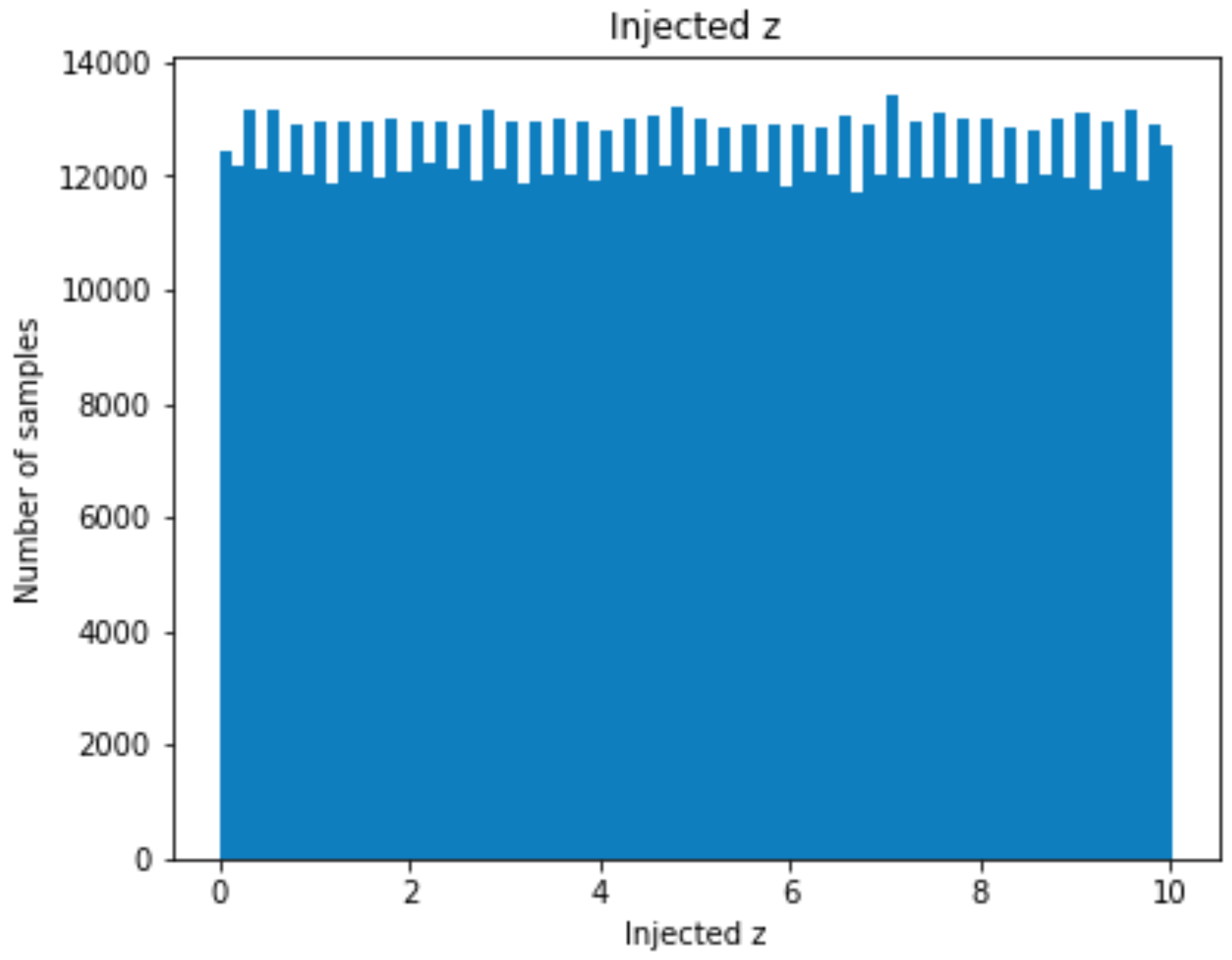
$$Z^2 = -2log\frac{L(B)}{L(\hat{\mu}S + B)}$$

[2] https://arxiv.org/abs/2107.11573

IBH - Injecting Signal Procedure

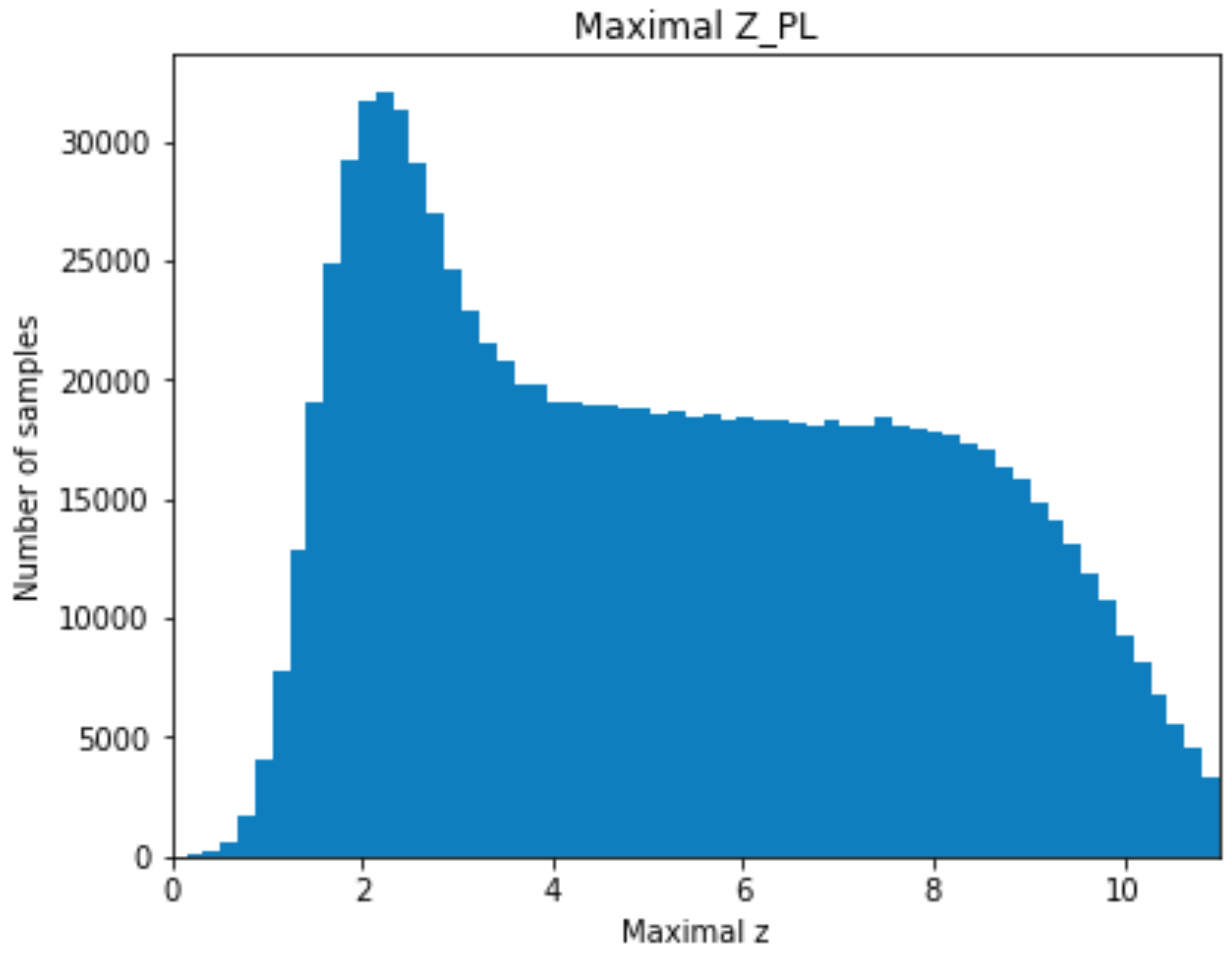
- Data composition:
  - Generating histograms from a Poisson fluctuated 10 analytical falling functions as background
  - Injecting a 6 bin width Gaussian signal on top of the background G(i,3)
  - #bins: 30,40,…,100 mass bins;
    Dynamic range: 10 - 100,000 entries/bin;
    Signal significance range: [0 - 10]
  - Data set: 90% trainning, 10% validation
  - 1M samples for any bin combination,
    900k x 8 = 7.2M training samples

Injected Z



Distribution of the injected signal significance, Zinj

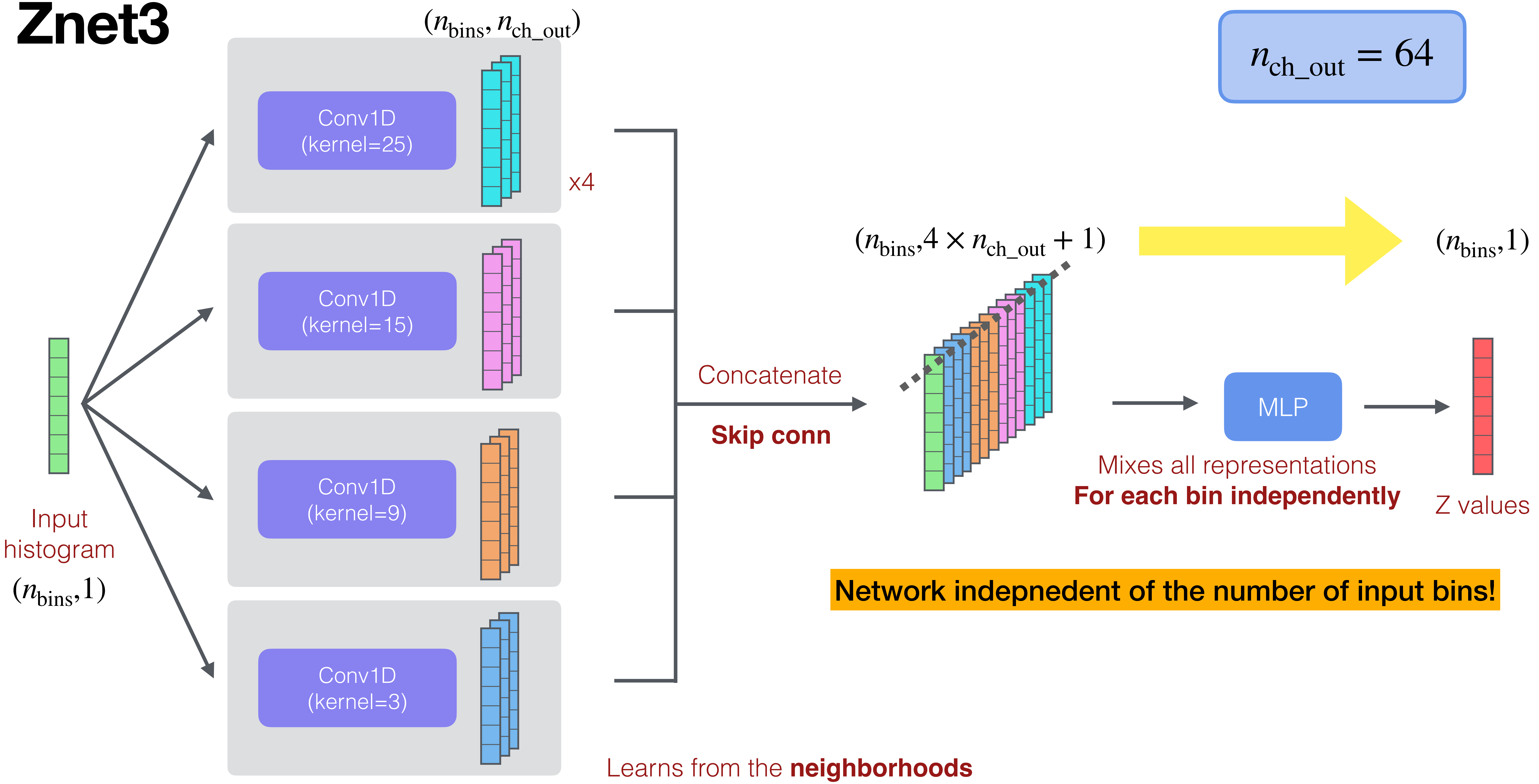
Z_max (PL)



Distribution of the Maximal measured Z

- LEE pushes Z maximal to the right for low Z_inj
- Downward fluctuations push the high Z_inj to the left

Znet3



Network indepnedent of the number of input bins!

Learns from the **neighborhoods**
Of different size from the **raw input**

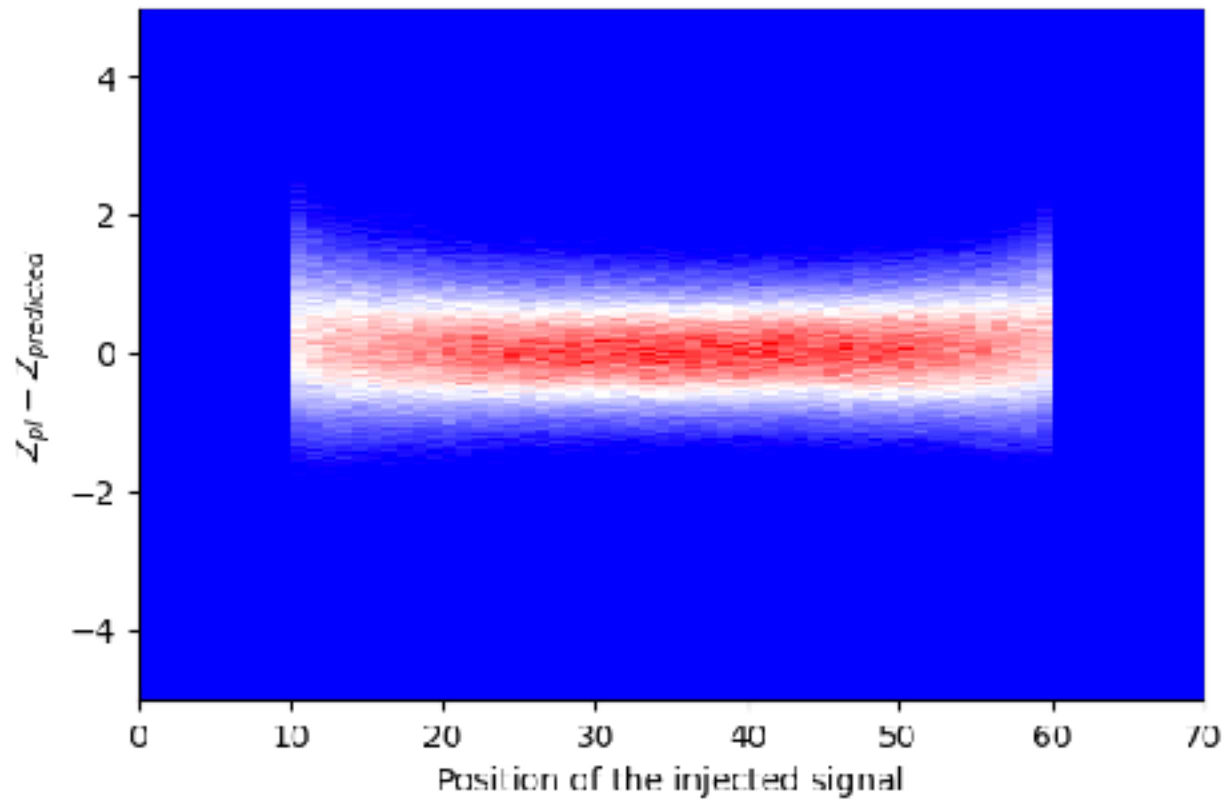
ZNET 3 Performance

$Z_{PL} - Z_{predicted} \quad vs \quad mass$

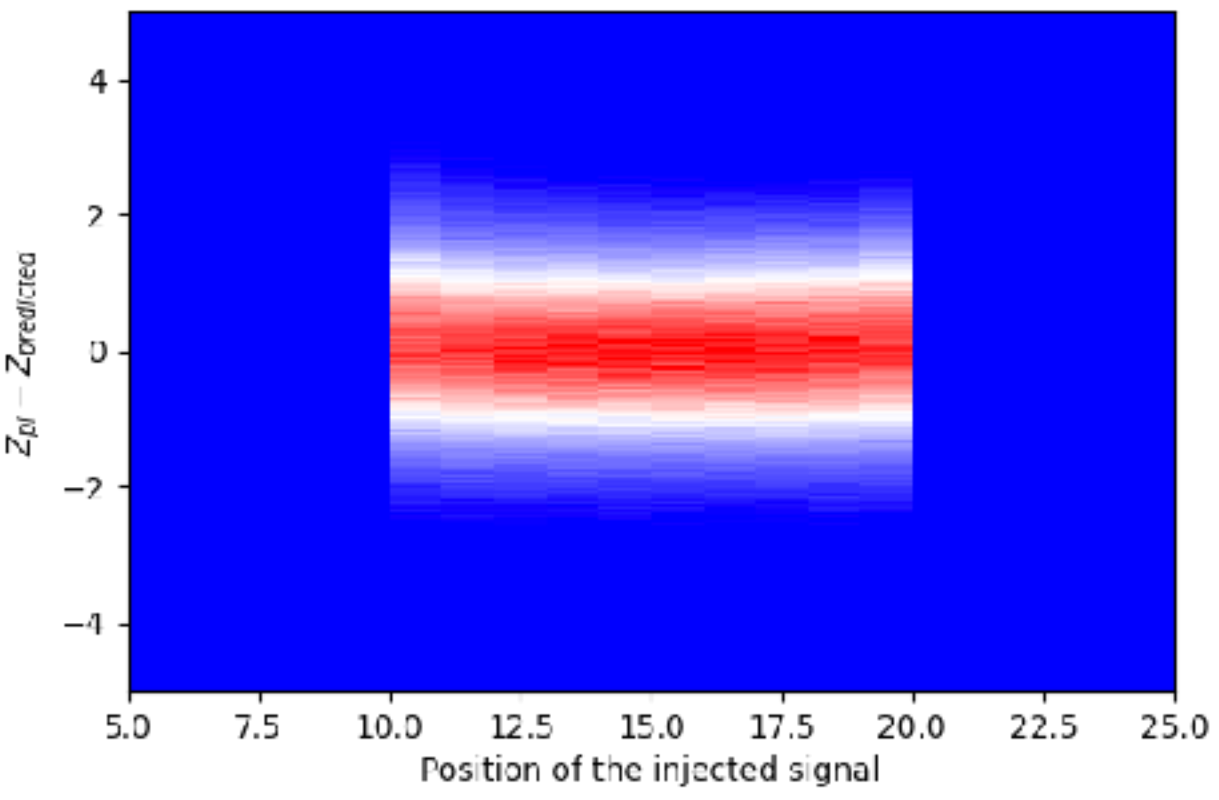
- **Agnostic to bin count**
- **Sensitivity over broad dynamic range (#entries/bin→[10,100K])**
- **Stable**

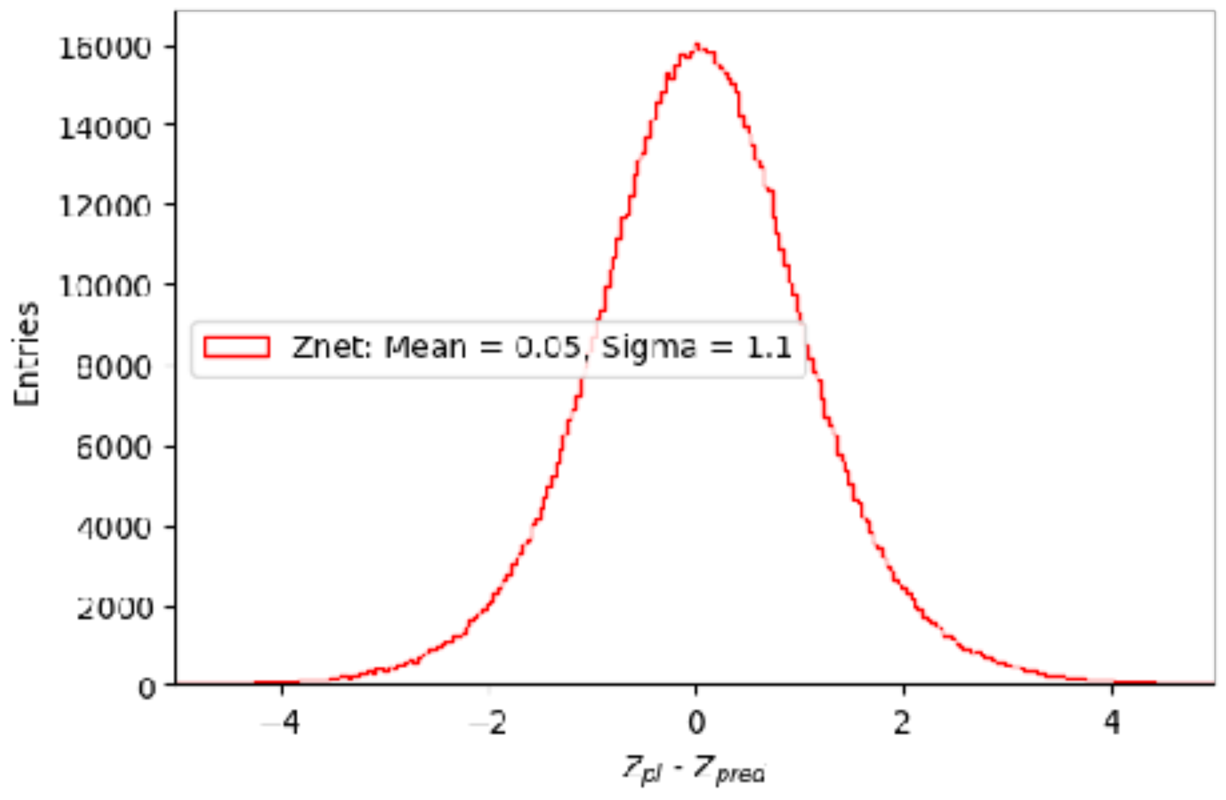
1 network

70 bins

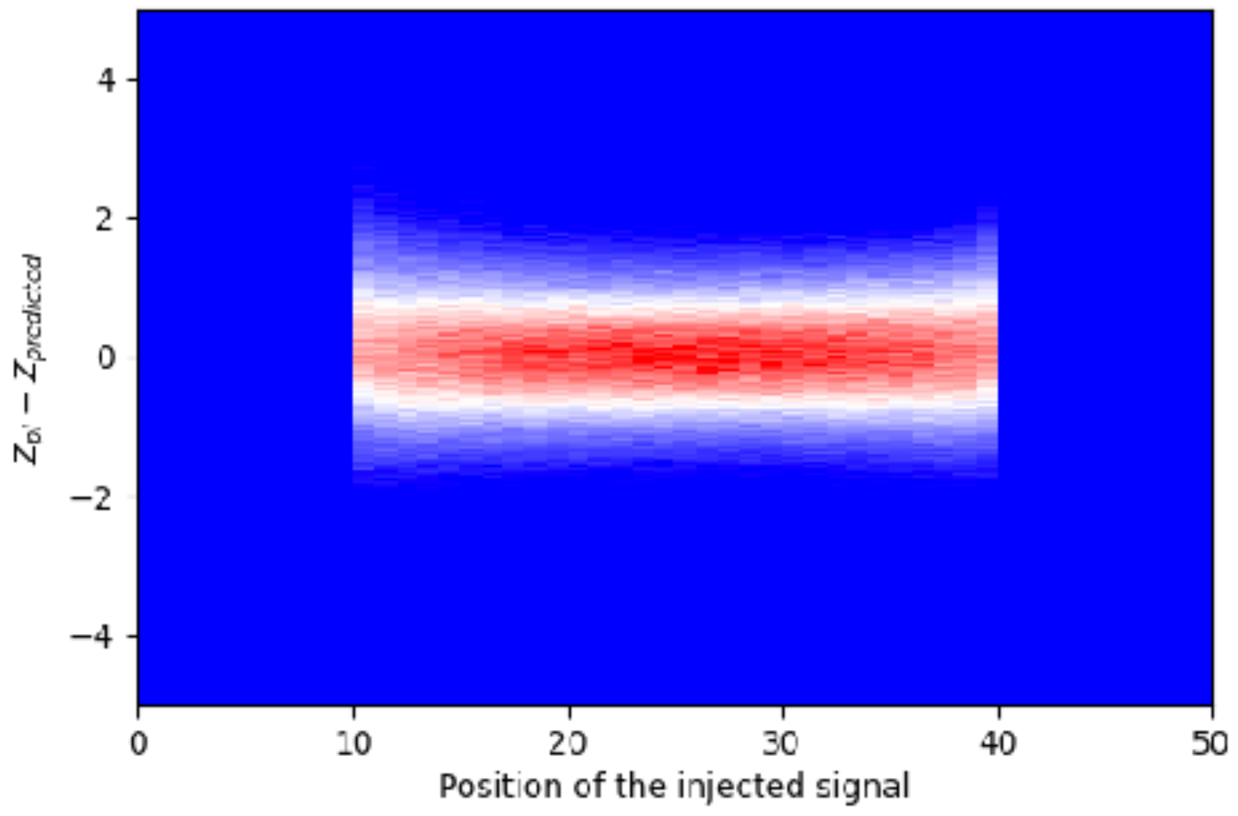


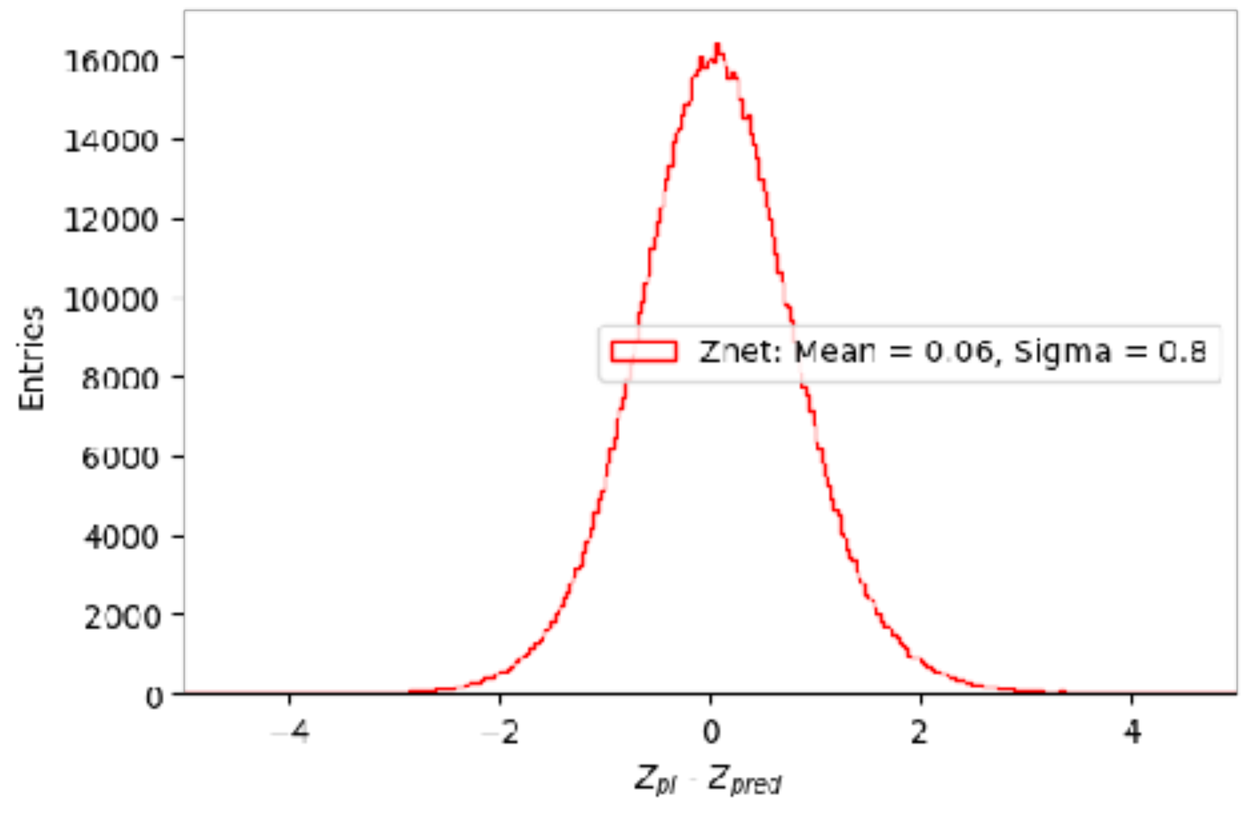
30 bins





50 bins

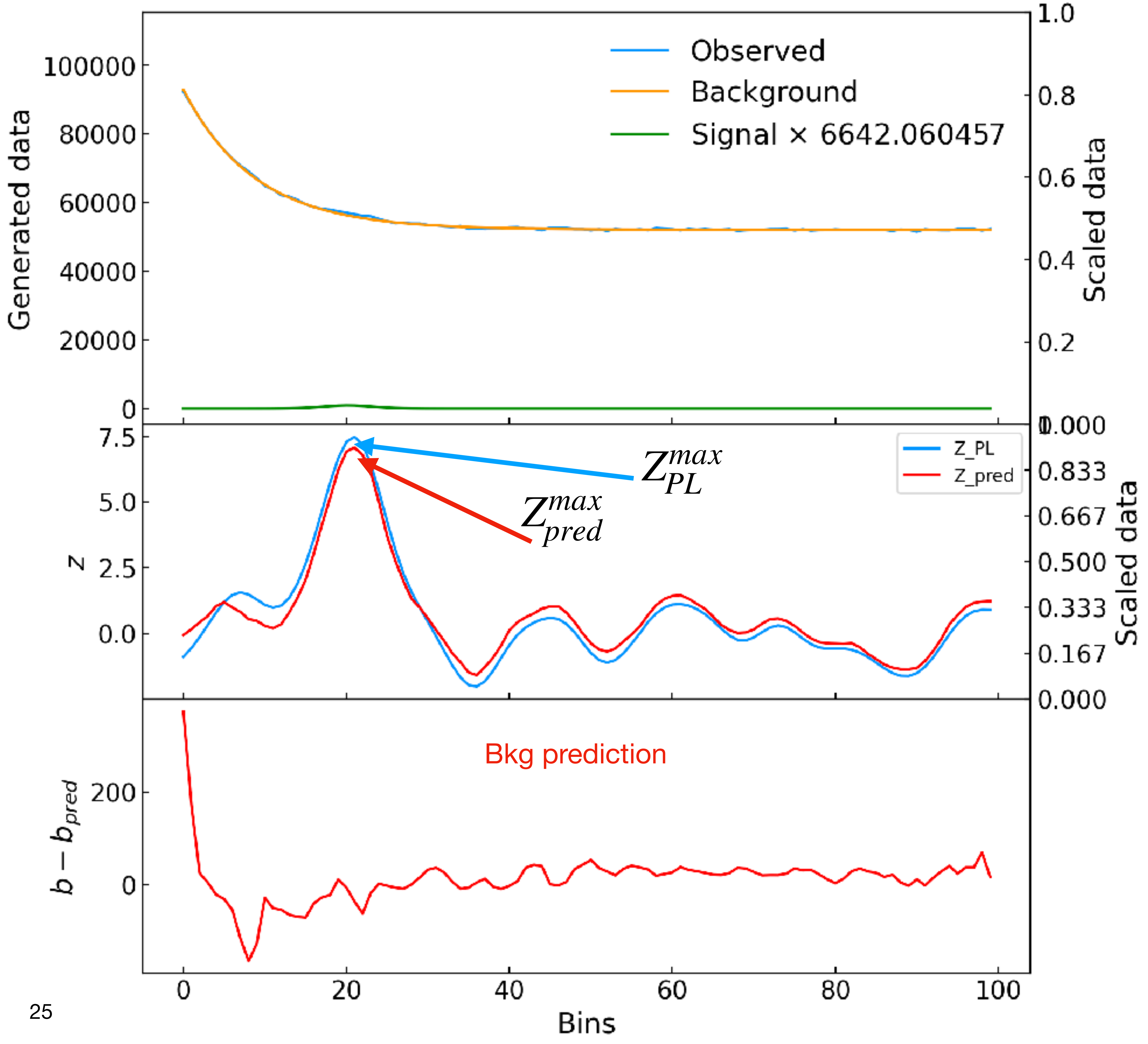




Znet3 Performance

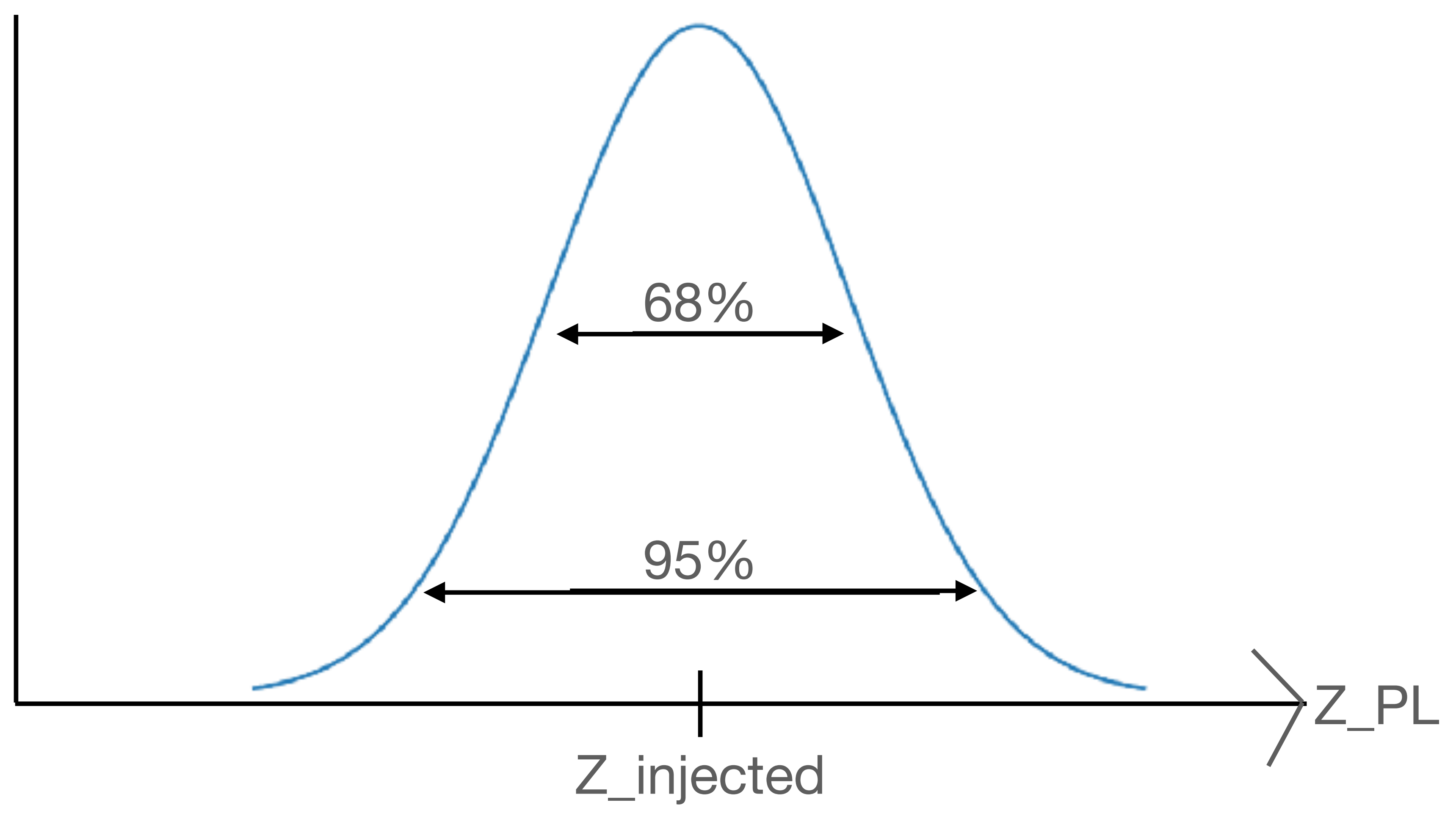
Synthetic Data

- Performance quantified in terms of $Z_{PL}^{max} - Z_{pred}^{max}$
  - Z_{PL}^{max} - maximal significance calculated via the likelihood ratio test
  - Z_{pred}^{max} - The maximal predicted significance



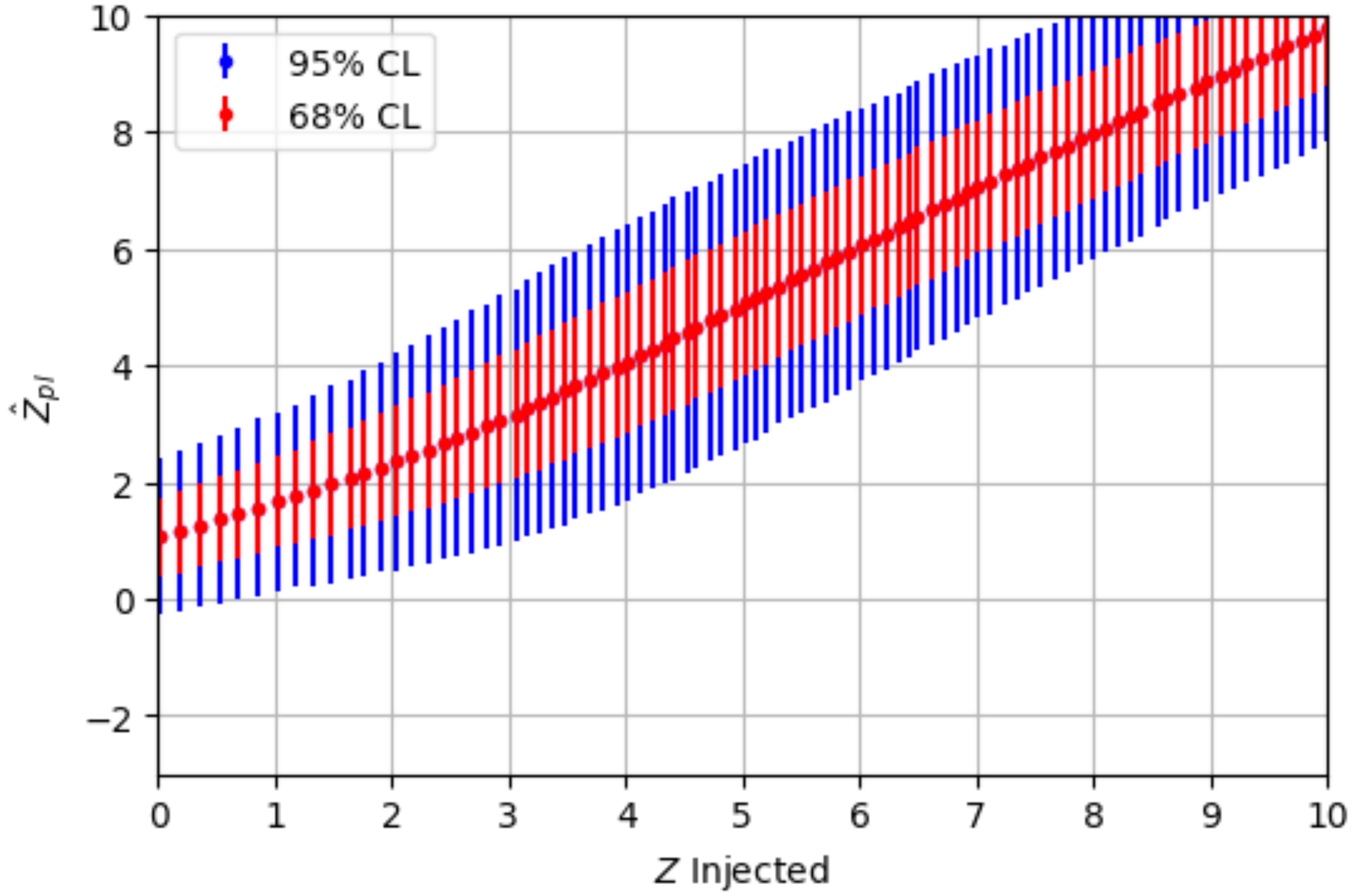
Quantifying NN Performance via a Confidence Belt

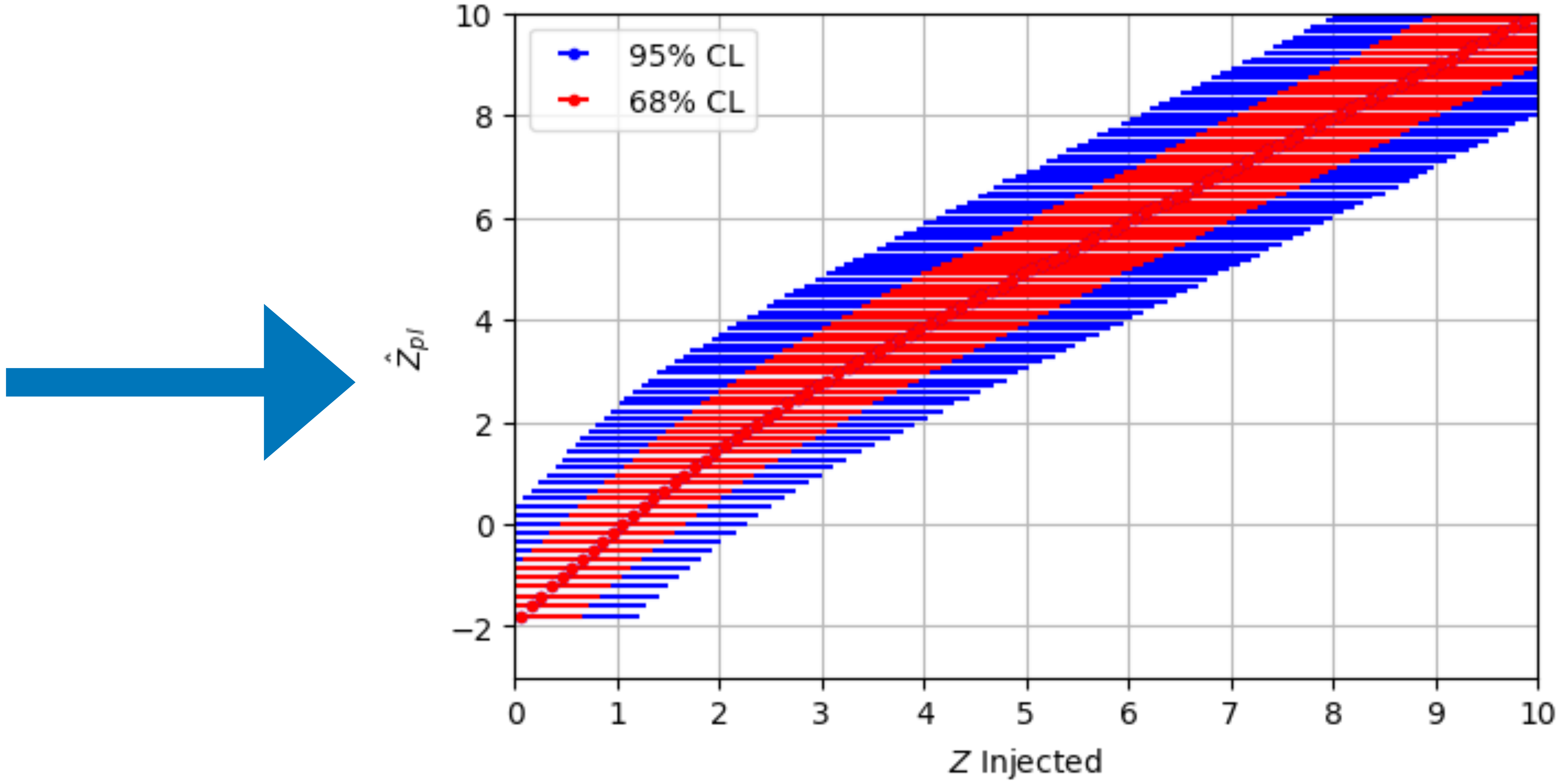
- For a given Z_injected
  - Plotting a histogram of Z_PL
  - Taking $\sigma * 1.96$ which correspond to 95% confidence level
  - Plot it as a function of Z_injected



Quantifying NN Performance

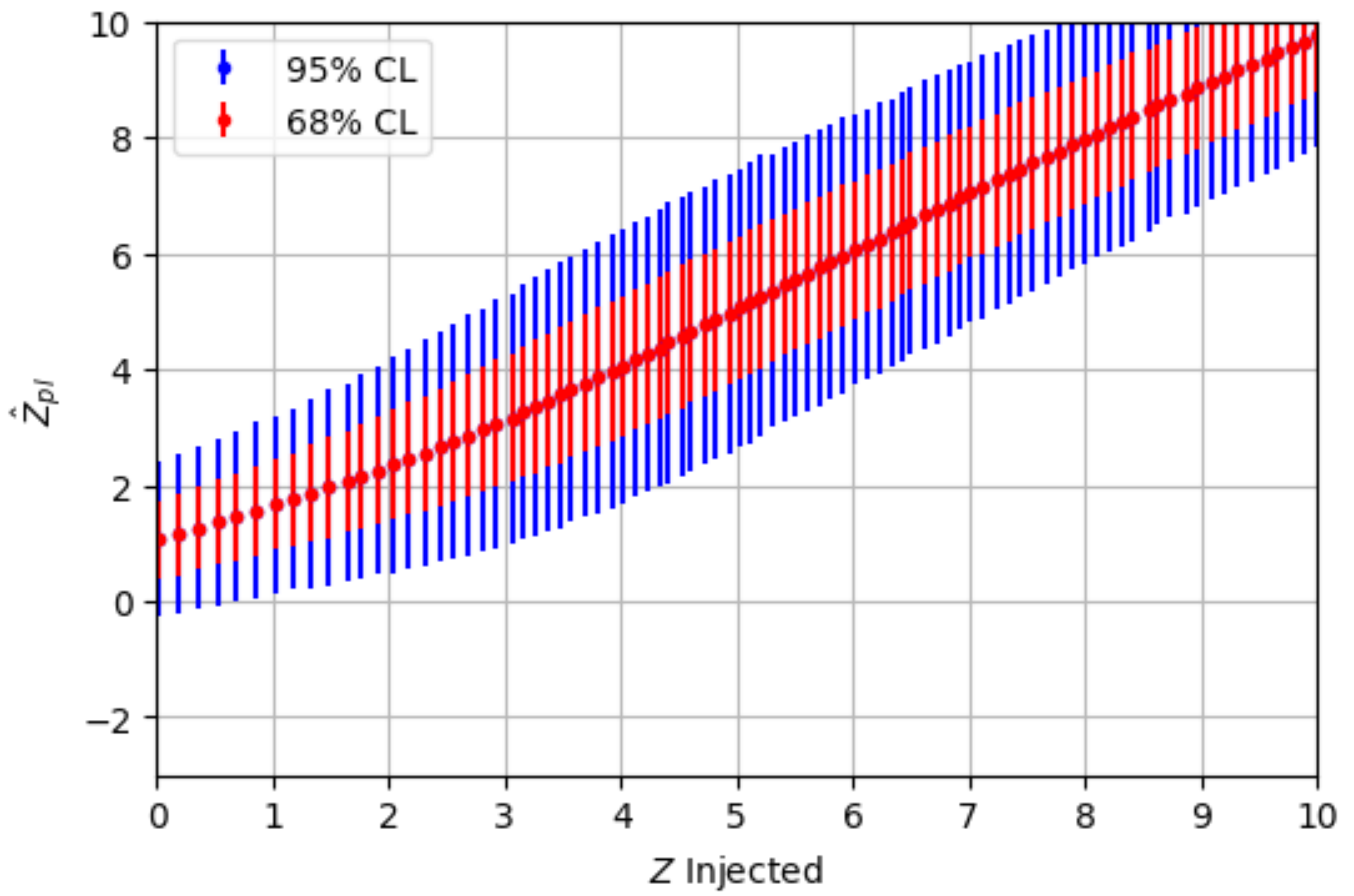
Confidence Interval

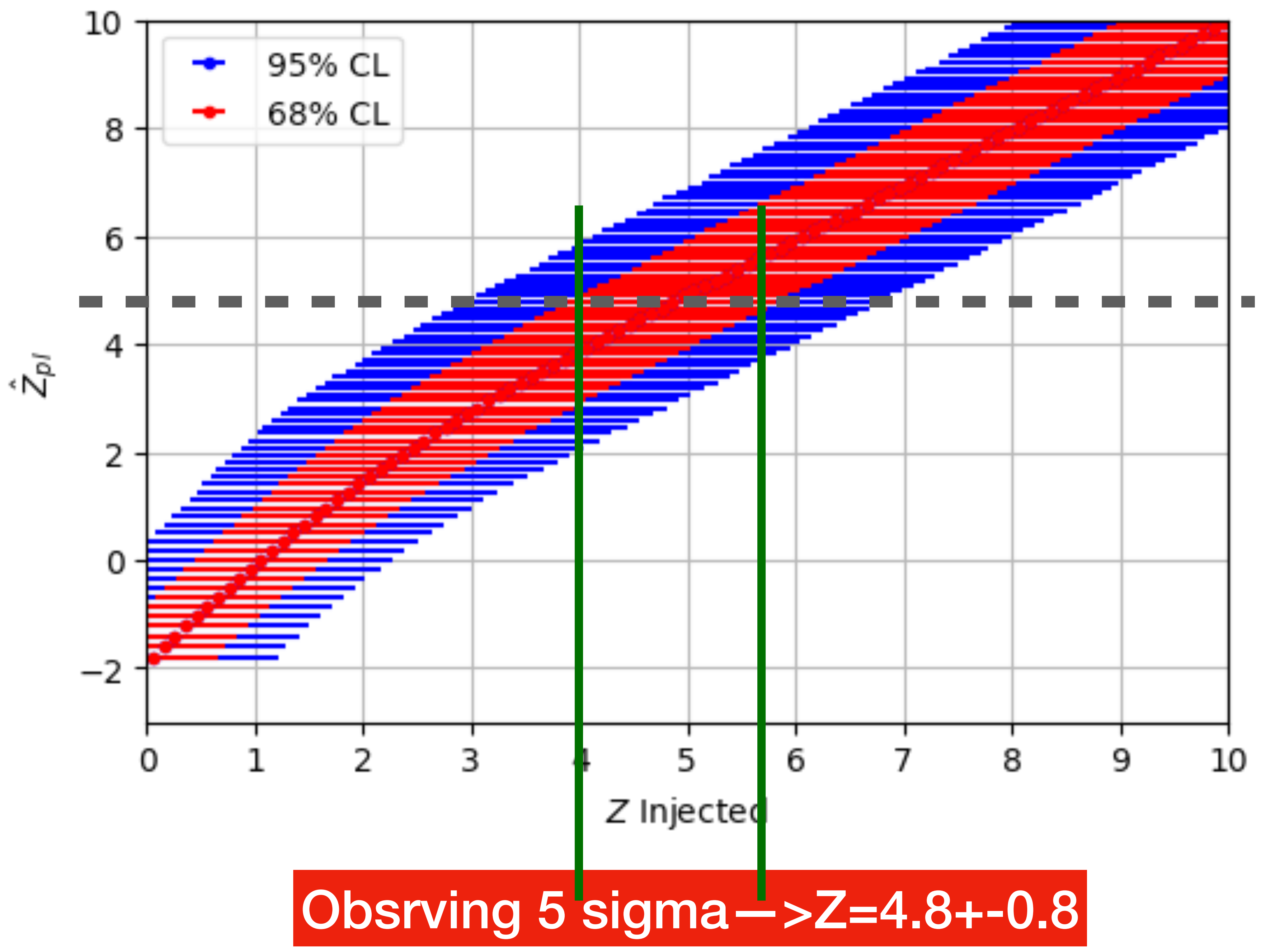




Quantifying NN Performance

Confidence Interval





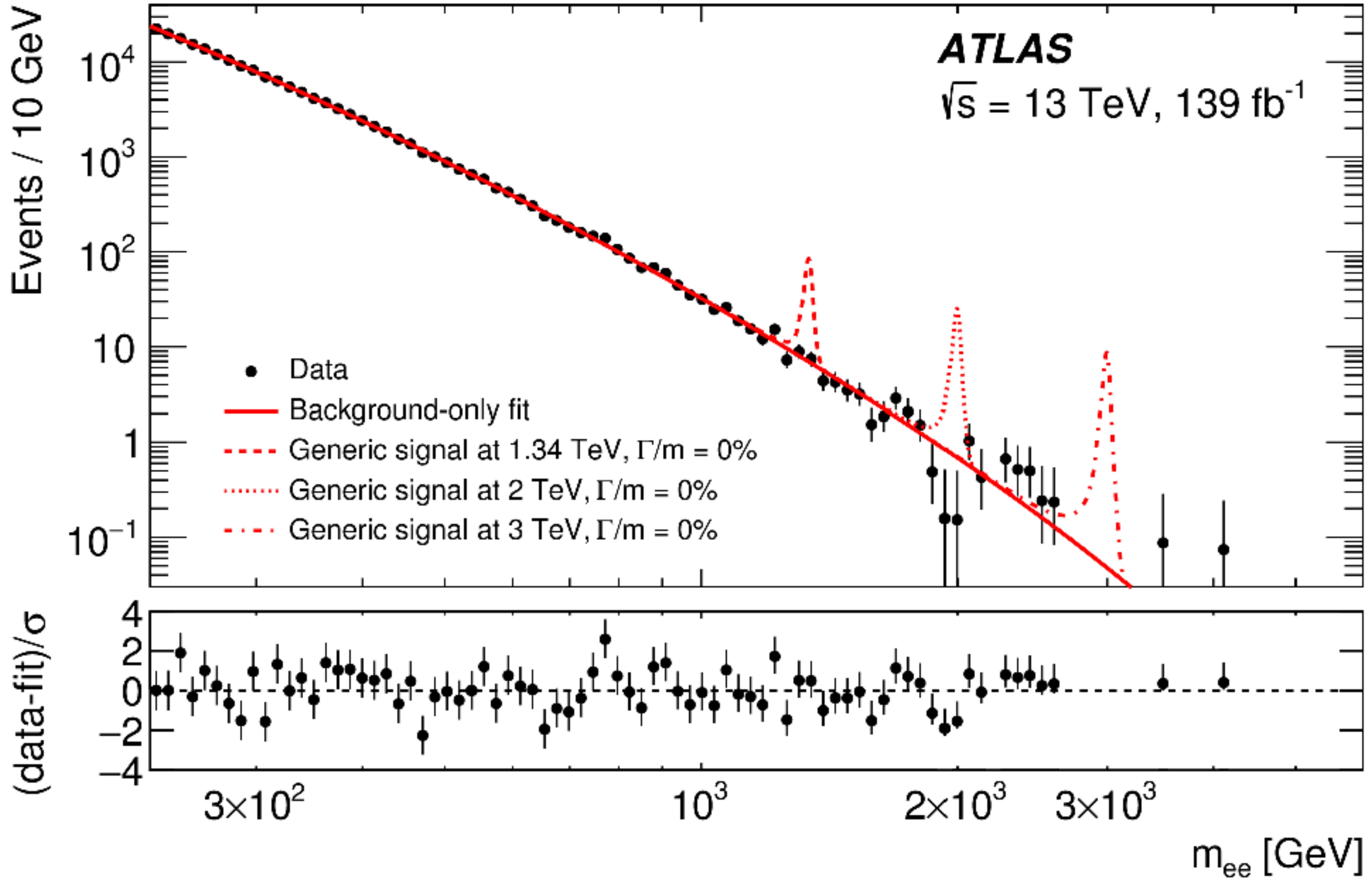
HEP Data

- Search for high-mass dilepton resonances using 13 fb^{-1} of ppcollision data collected at $\sqrt{s} = 13TeV$ with the ATLAS detector

- Generate data from extracted bkg distributions from the HEP paper using fitting functions suggested in the paper:

$$f_{\ell\ell}(m_{\ell\ell}) = f_{\mathrm{BW},Z}(m_{\ell\ell}) \cdot (1 - x^c)^b \cdot x^{\sum_{i=0}^{3} p_i \log(x)^i}$$

Distribution of the dielectron invariant mass for events passing the full selection [3]

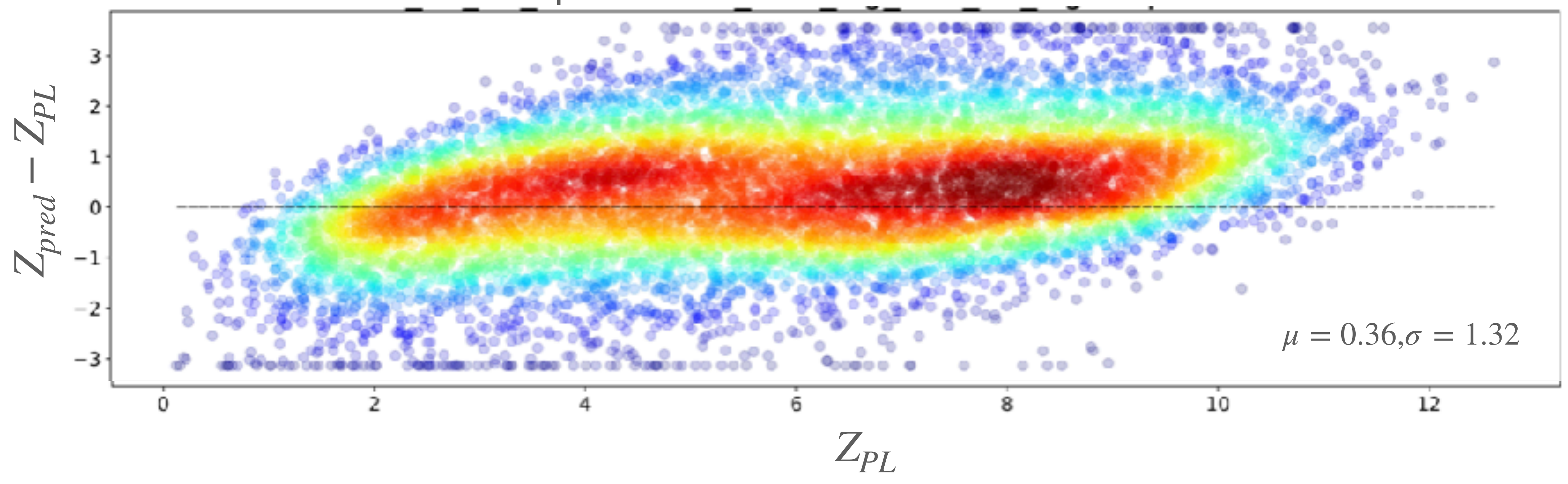


[3] https://doi.org/10.17182/hepdata.88425.v3

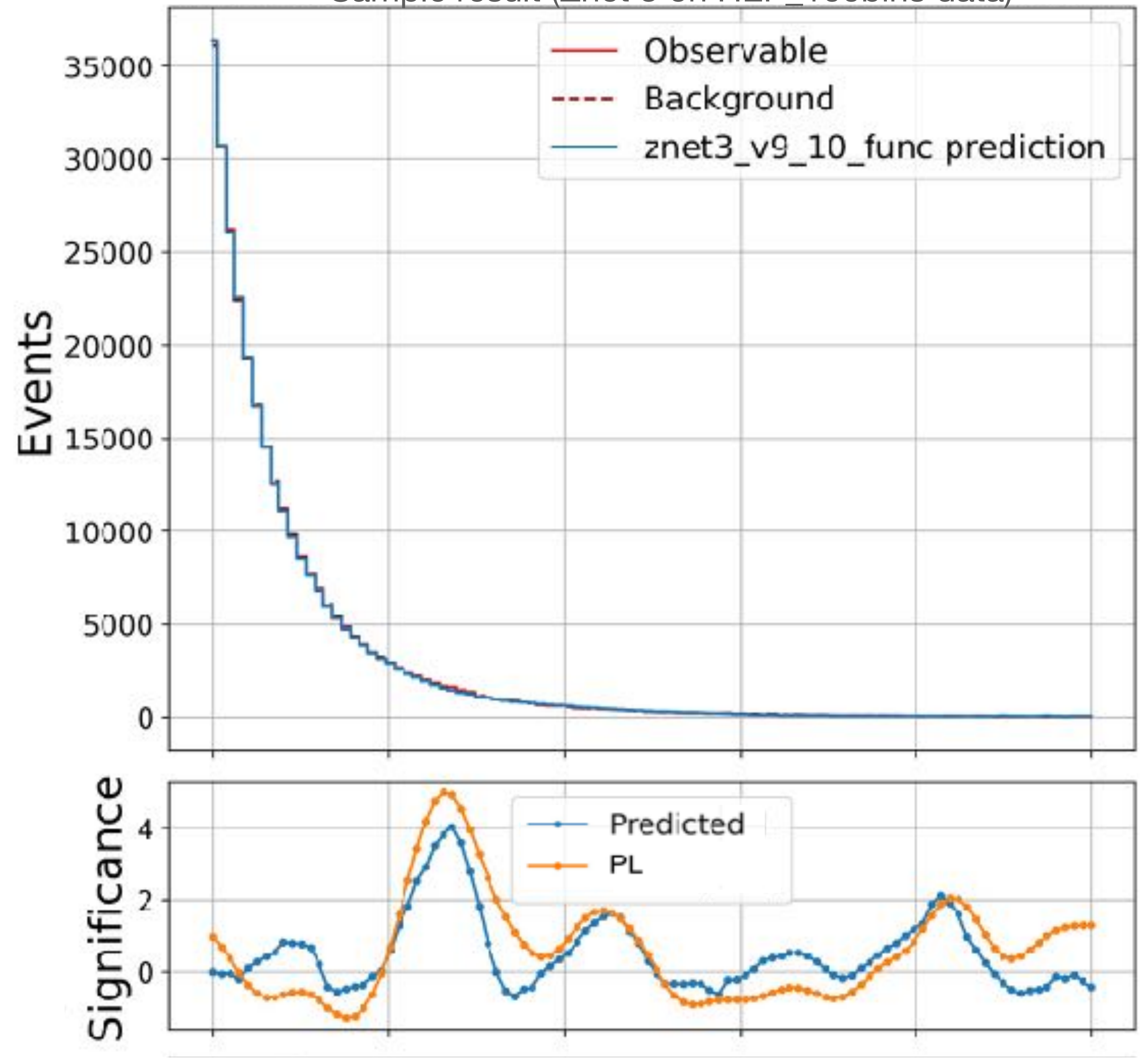
HEP Data

- Data generated using paper fit formula and parameters
- Same generating pipeline:
    - Generate bkg
    - Inject signal
    - Fluctuate
    - Calculate Z_PL
    - Train
    - Evaluate

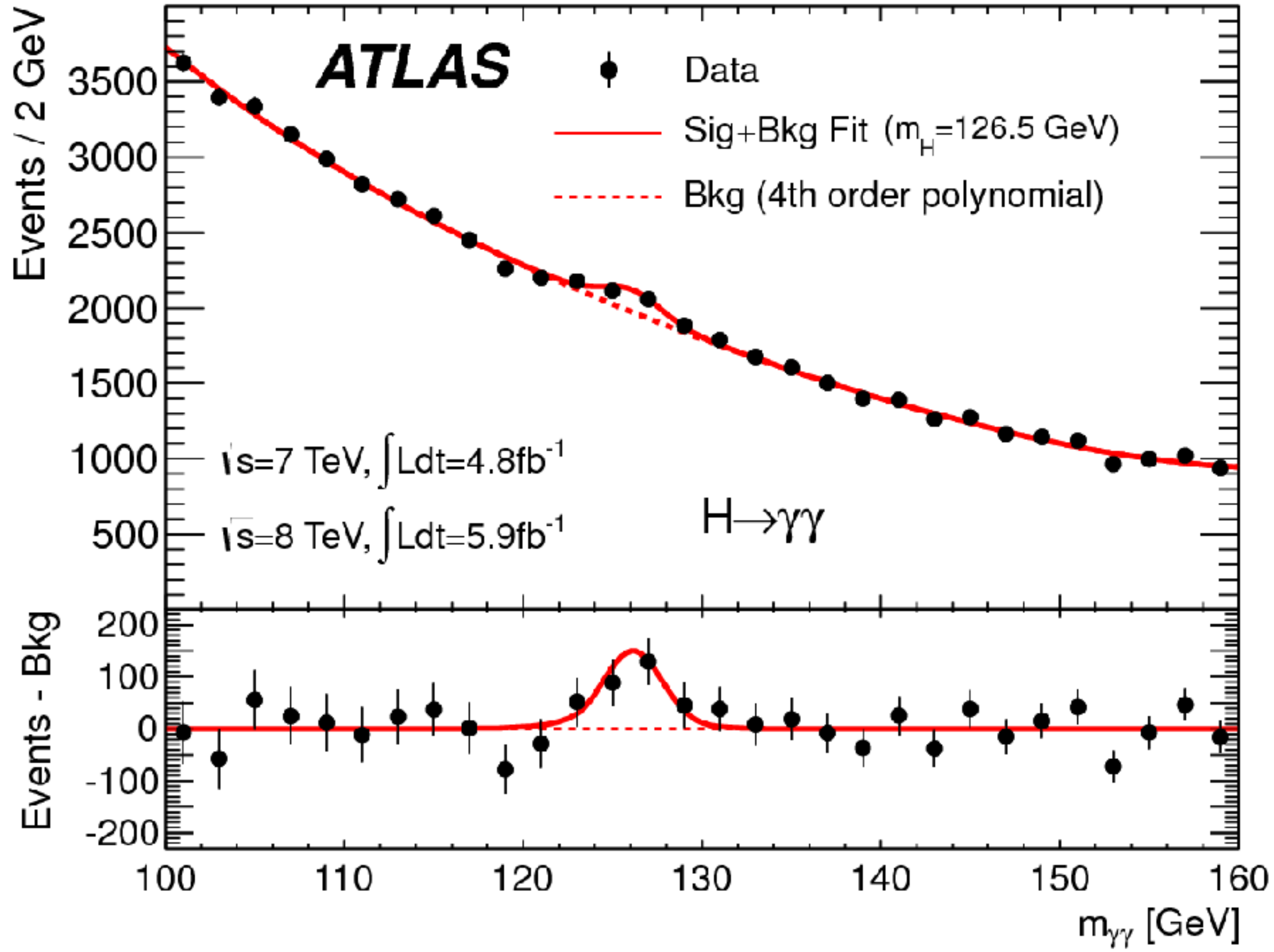
Znet 3 performance: HEP data - all #bins

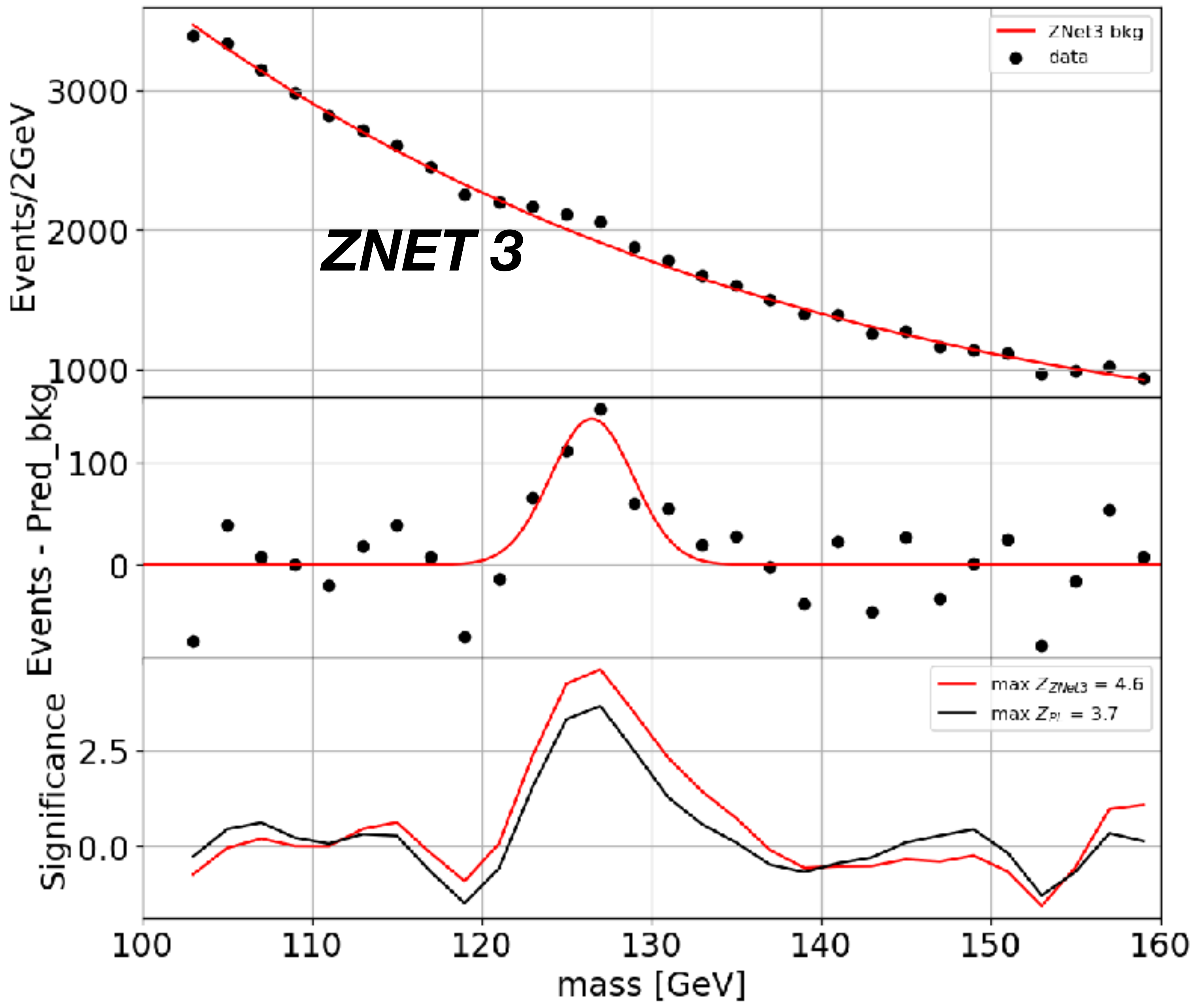


Sample result (Znet 3 on HEP_100bins data)



Real Data example : Higgs - diphoton





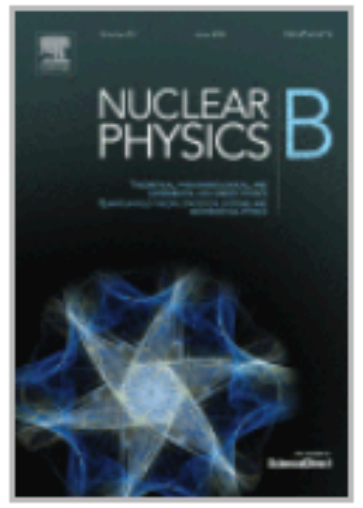
[4] http://arxiv.org/abs/arXiv:1207.7214

Marcela in an Interview - the 90s

At CERN, for the first time in my life, I got immersed deeply into what was going on in the experimental world. Every single day I would discuss with experimental colleagues, postdocs like me, such as Fabiola Gianotti, Patrick Janot, Eilam Gross, Michael Schmitt, Gustavo Wolf, Marta Felcini, ... all obsessed with finding the Higgs and searching for Supersymmetry at LEP. At that time, we heard of the sad news about the

Nuclear Physics B

Volume 580, Issues 1–2, 31 July 2000, Pages 29-57



Reconciling the two-loop diagrammatic and effective field theory computations of the mass of the lightest CP-even Higgs boson in the MSSM

M. Carena [a b], H.E. Haber [c], S. Heinemeyer [d], W. Hollik [e], C.E.M. Wagner [b f g], G. Weiglein [b]

Carlos Wagner

November 21, 2013 · Naperville ·

I found this quote written somewhere. I had notes of its origins, but I misplaced them, and my weak mind no longer has the eidetic memory I used to enjoy in the past, so I car not give proper credit to its author. Let me emphasize that when I first read it, I dismissed it as a great exaggeration, a scientific fallacy and even a plagiarism of Borges' Vindication of the Kabbalah. I am no longer certain of any of these facts, and therefore I decided to share it with you. Here it is, in the way I transcribed it. I apologize to the author, if he/she ever reads it, for the possible mistakes introduced in my transcription :

"There is more energy in the Higgs vacuum than in any other known available energy resource. Although it looks impossible right now, accessing such energy source opens an impressive range of possibilities that are hard to conceive today.

In order to understand the secret message encoded in the Higgs field and find out if this is possible, we need to examine the Higgs properties in great numerical detail, arguably until the absurdity. This will not demand hundreds of holes in the ground, like the extraction of other mineral resources do. A single hole may suffice. The investment is relatively minor and the technology to produce Higgs particles and analyze the properties of the Higgs field, is well understood.

Many economic powers are interested in investing in the analysis of the Higgs properties. I suspect that those investing in this field understand that it may lead to a fundamental key to the future of humanity. Those who don't invest today, will be left behind tomorrow. This has always happened to those who abandoned basic science. There is no reason to think that this time something different will happen."

Carlos Wagner

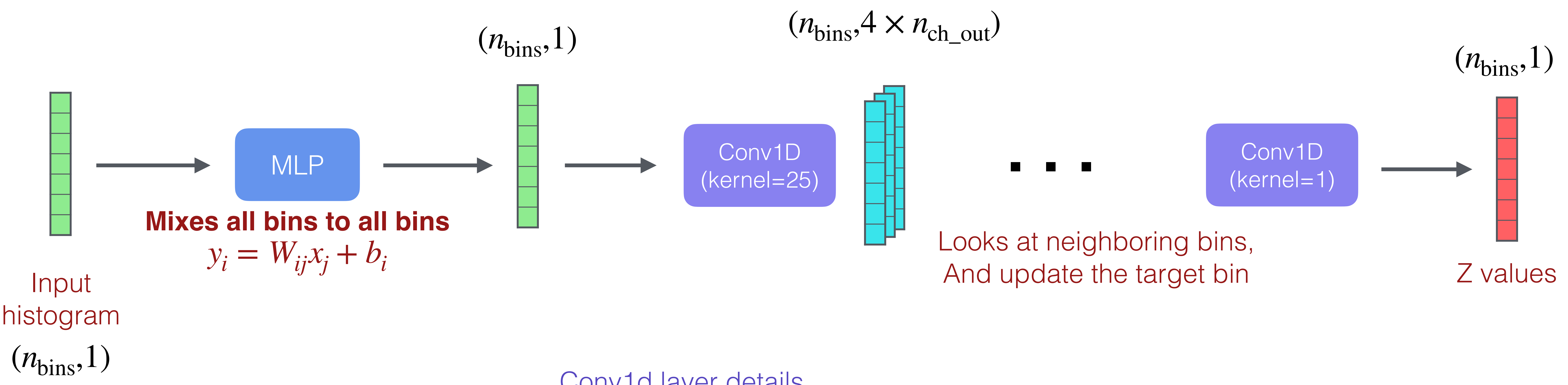
November 21, 2013 · Naperville ·

- There is more energy in the Higgs vacuum than in any other known available energy resource. Although it looks impossible right now, accessing such energy source opens an impressive range of possibilities that are hard to conceive today.
- In order to understand the secret message encoded in the Higgs field and find out if this is possible, we need to examine the Higgs properties in great numerical detail, arguably until the absurdity.
- This will not demand hundreds of holes in the ground, like the extraction of other mineral resources do. A single hole may suffice.
- Many economic powers are interested in investing in the analysis of the Higgs properties. Those who don't invest today, will be left behind tomorrow. This has always happened to those who abandoned **basic science.**


Happy Birthday M&C

BACKUP

Znet - The Vanilla Proof of Context Network



Conv1d layer details

Kernel	25	20	15	10	5	1
$n_{\text{ch\_out}}$	64	32	32	16	8	1

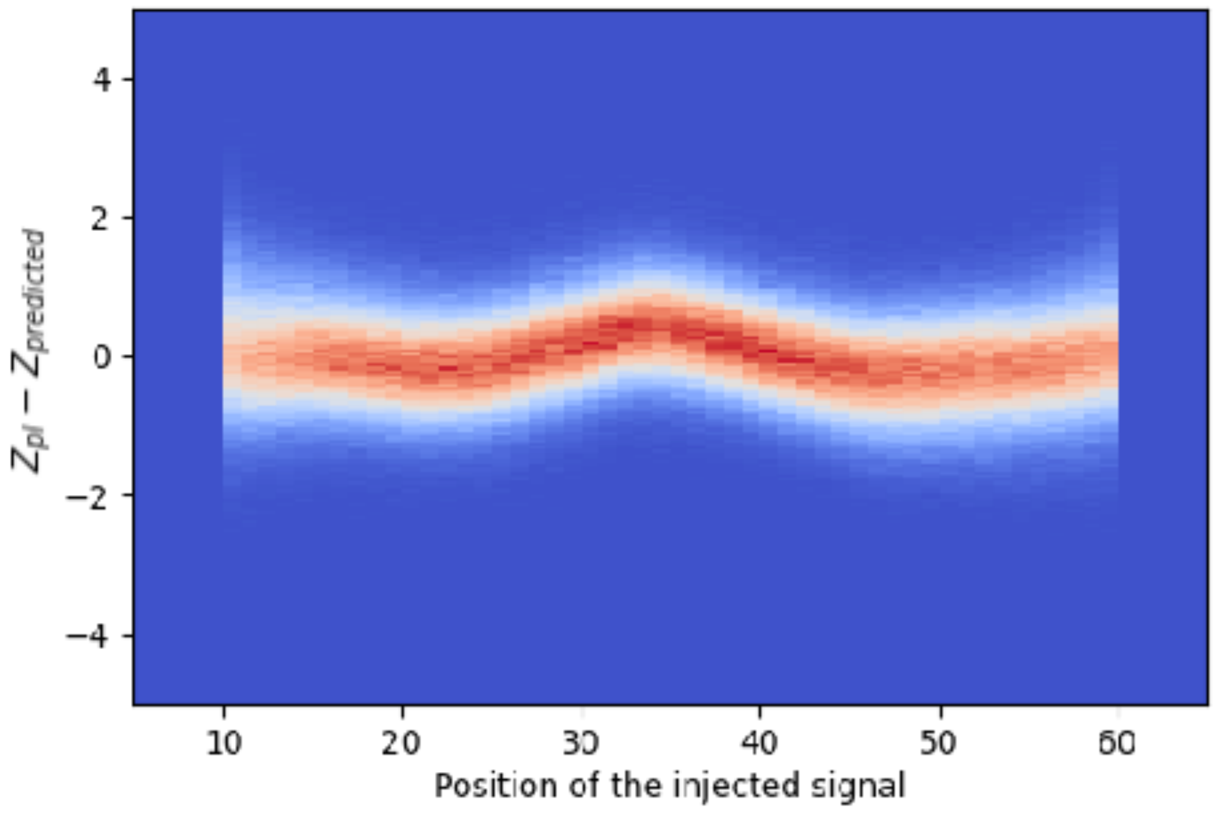
The Vanilla ZNET Performance

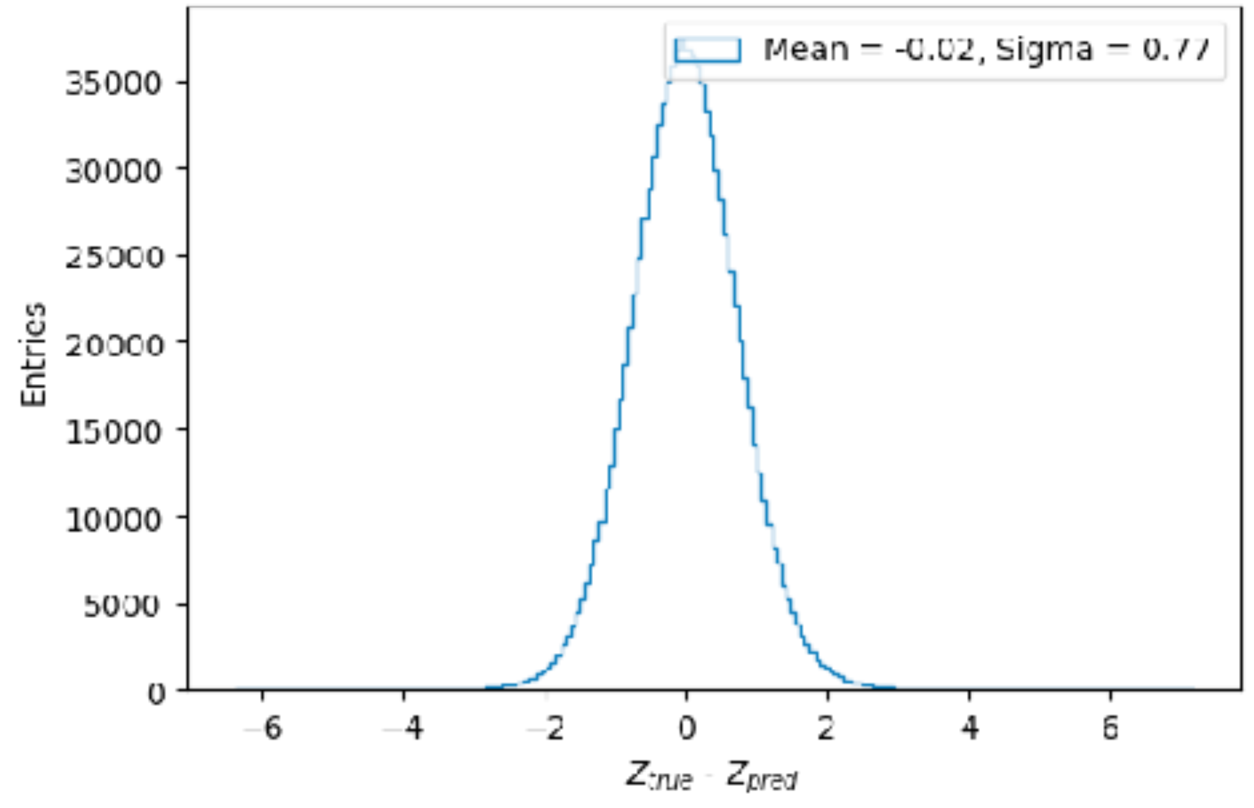
ZNET was built with 24 networks, each for a different dynamic range and number of bins its performance showed dependence on both the injected signal and the mass

- Bin count dependance - Different network for different #bin
  - 30,40,…,100 bins data sets —> 8 combinations
- Dynamic Range dependance
  - [10,1K], [10,10K], [10,100K] —> 3 combinations
- Network is not stable

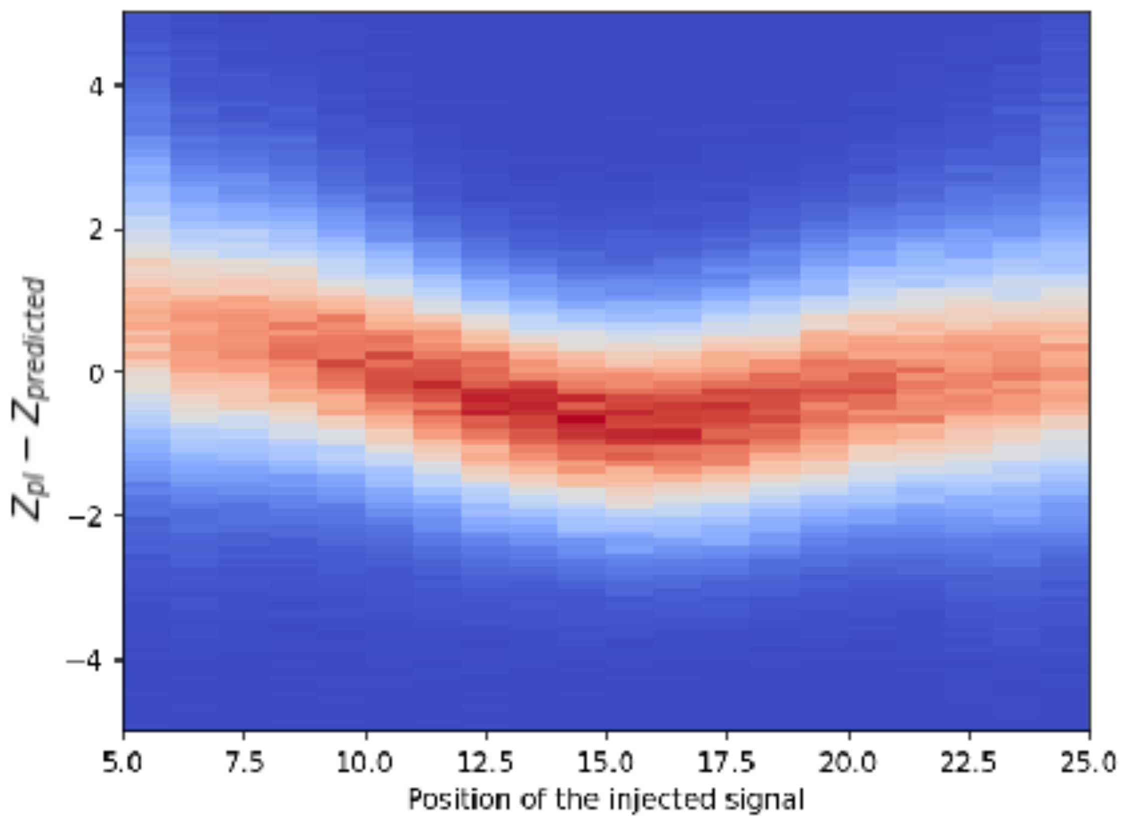
24 networks

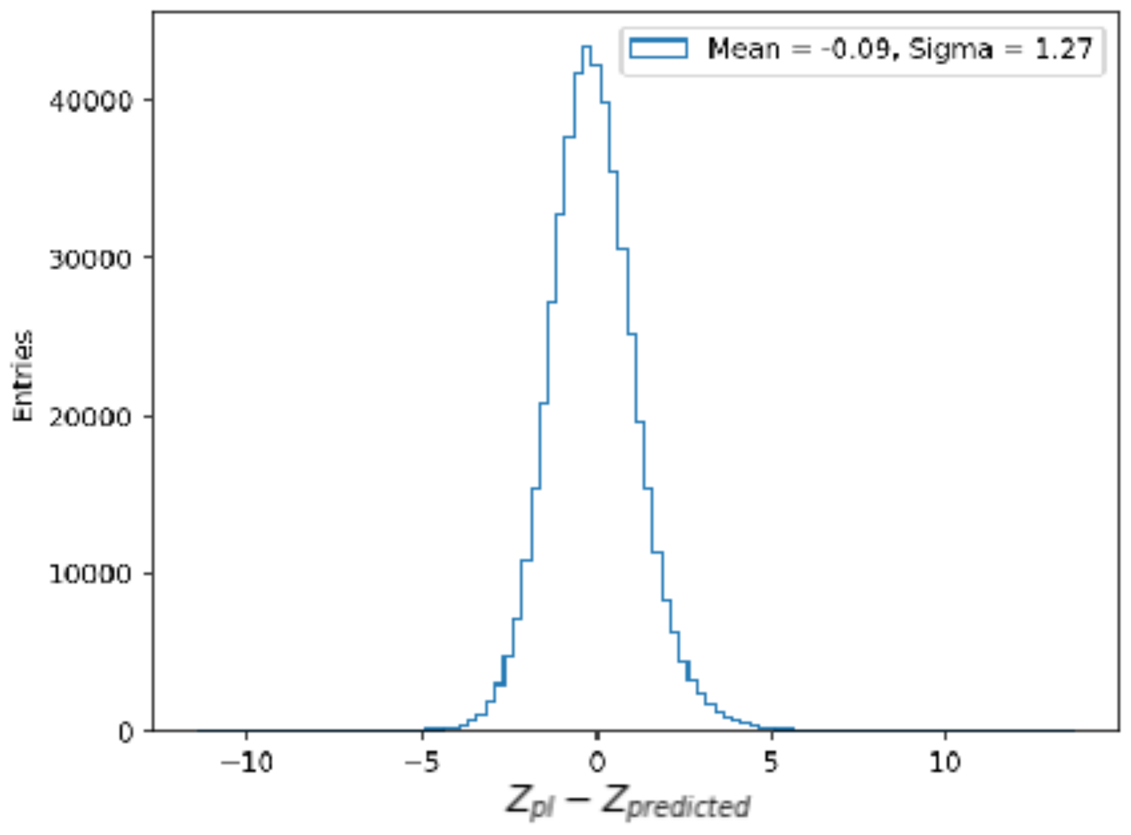
70 bins



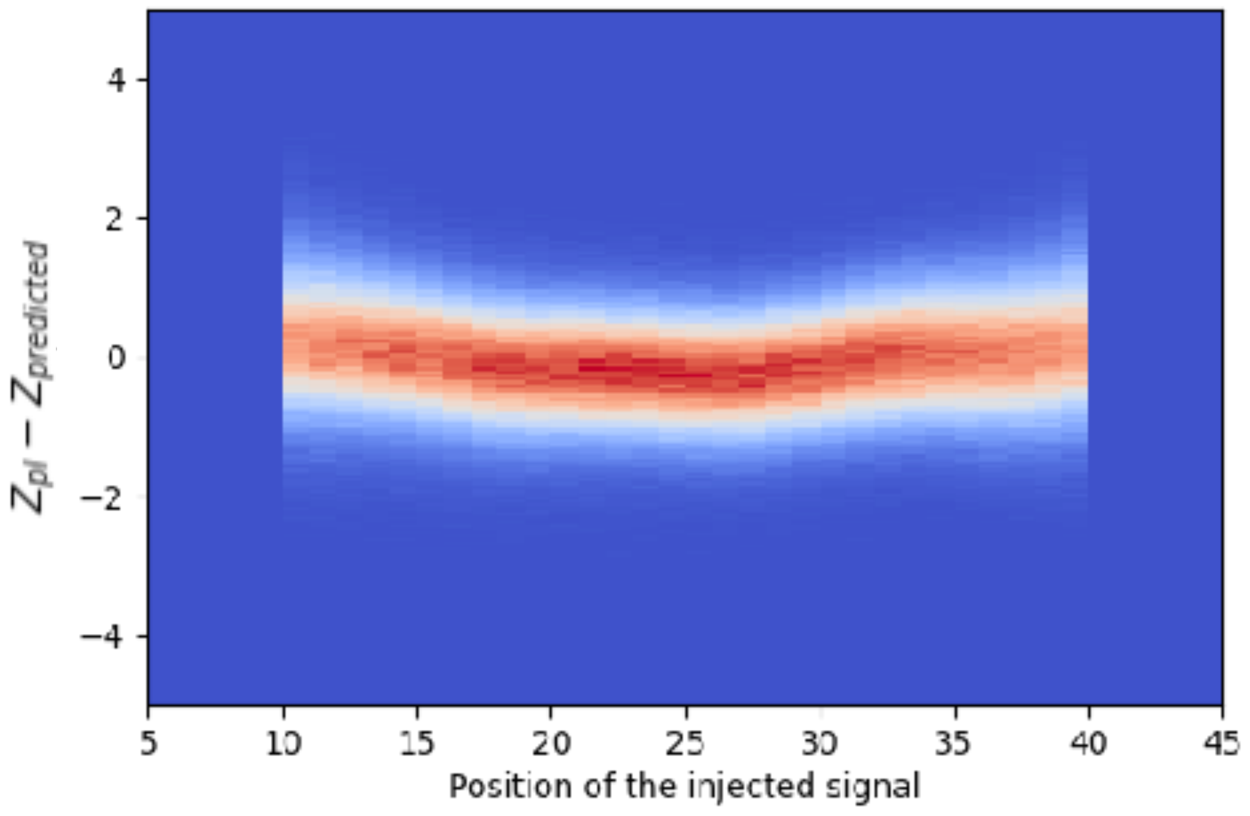


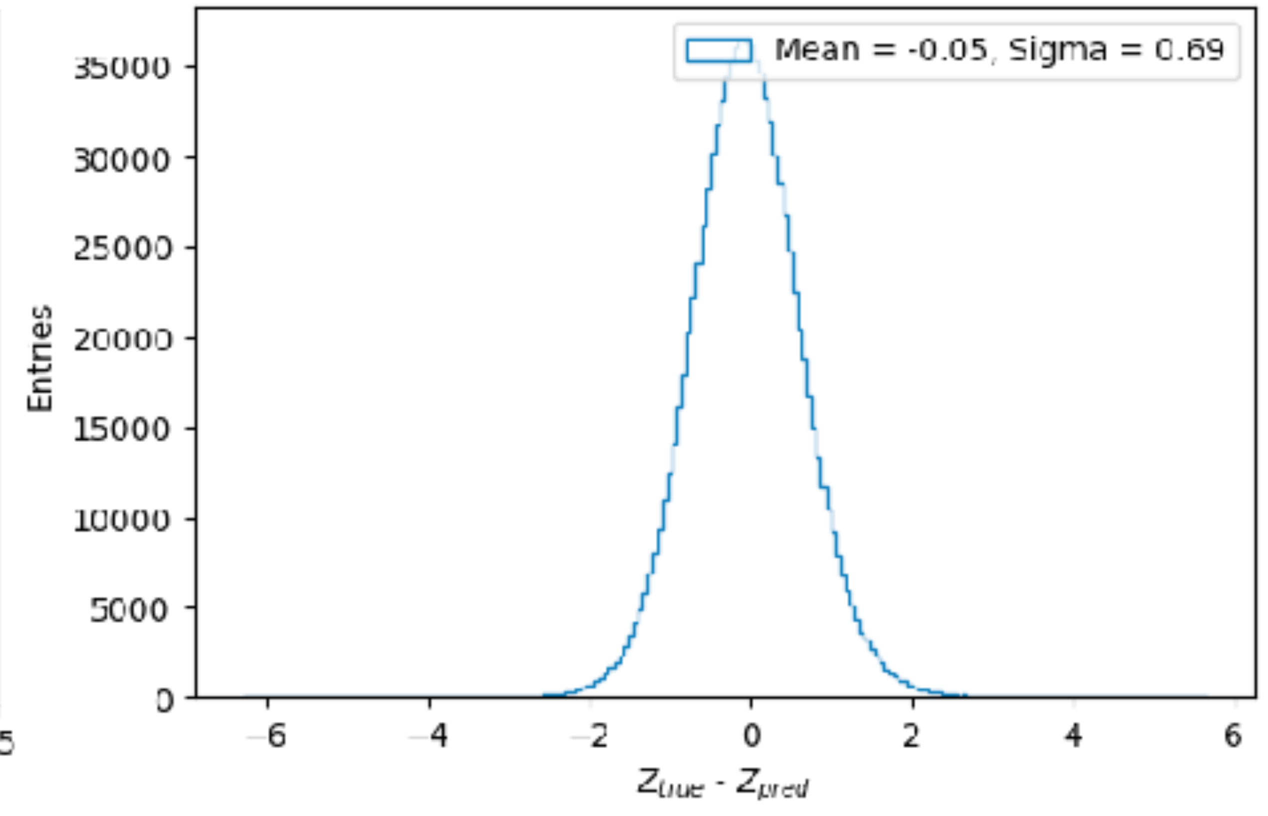
30 bins





50 bins

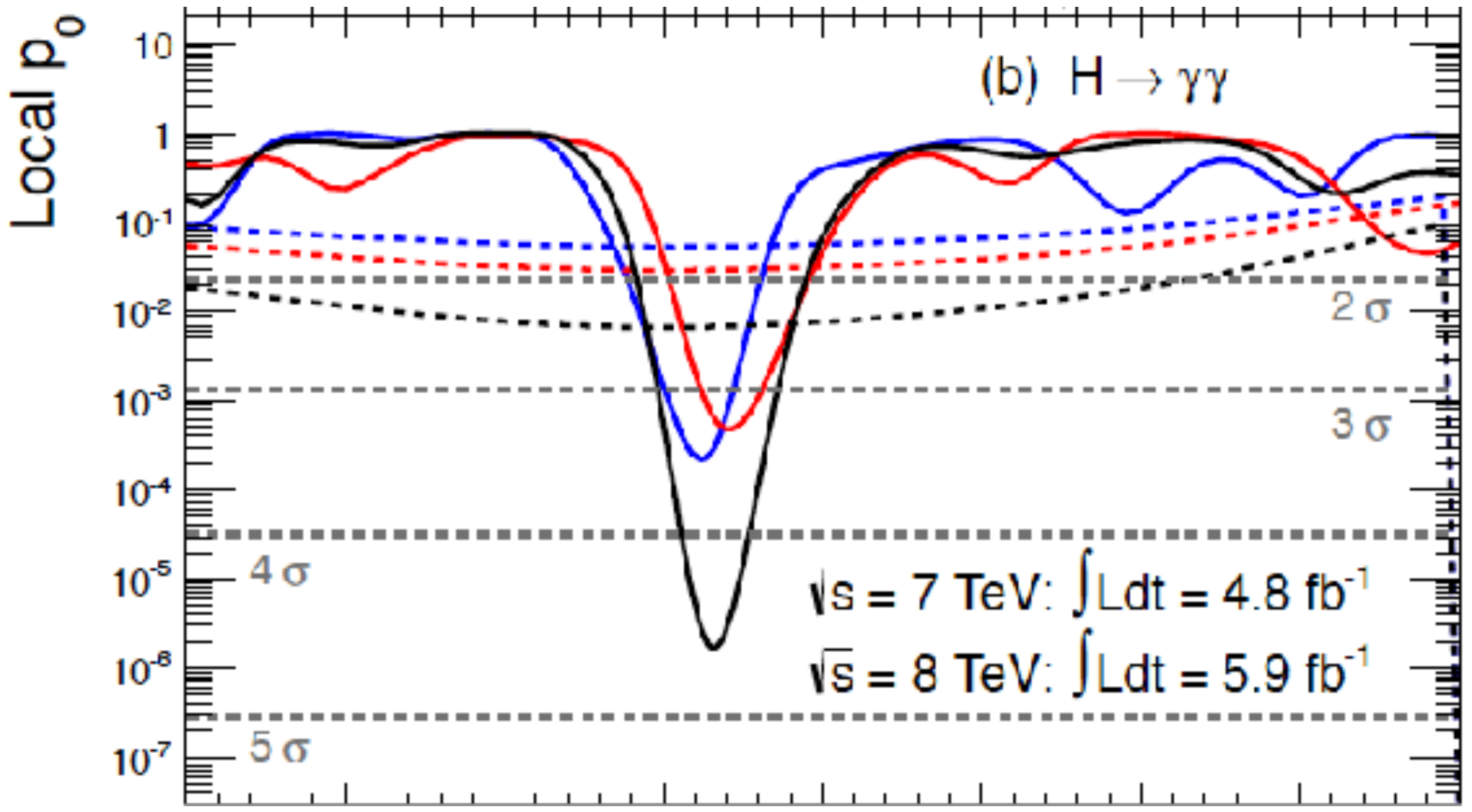


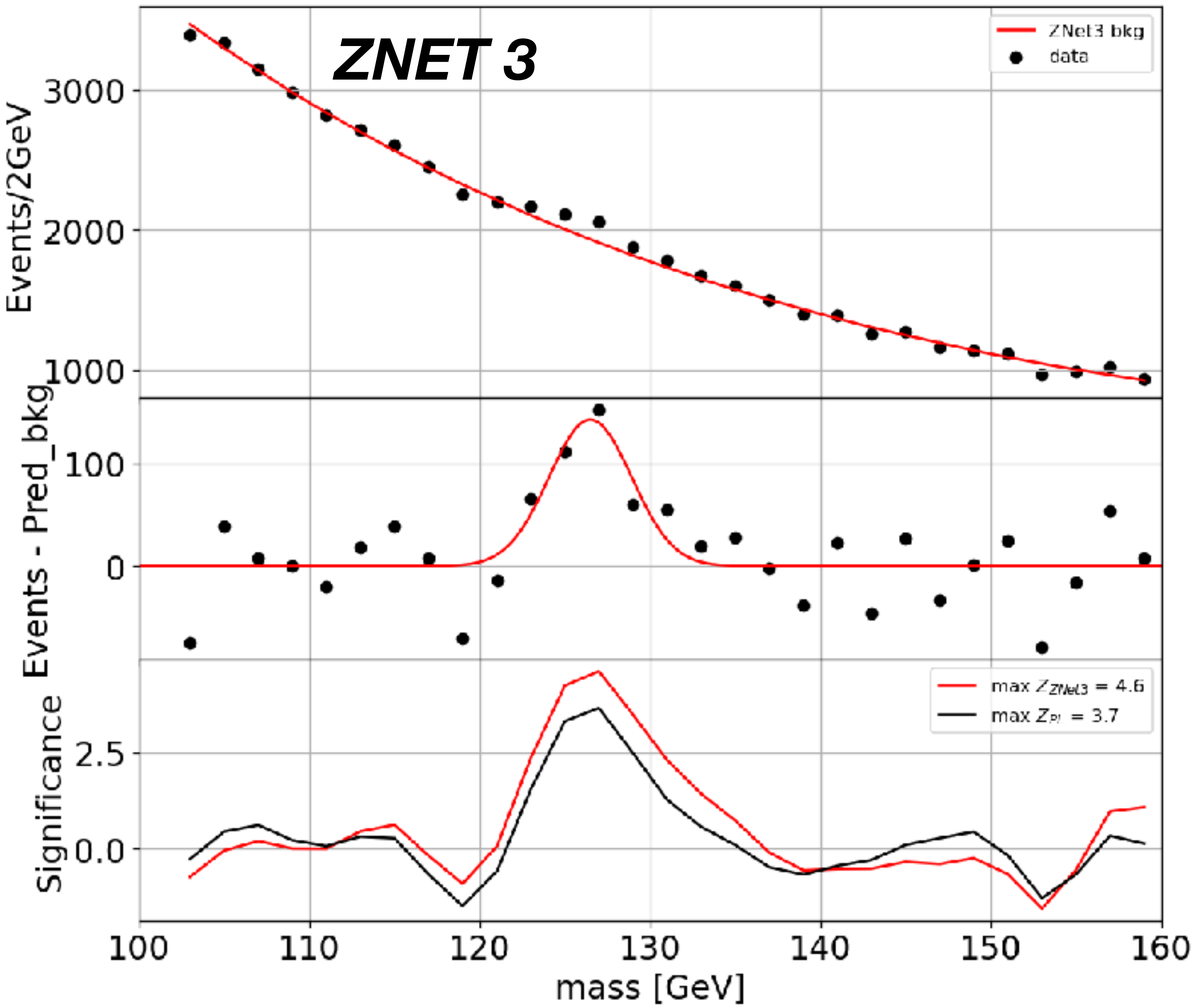


Real Data example : Higgs - diphoton

- ATLAS significance - 4.5σ
- Znet 3 - 4.6σ
- PL - 3.7σ
- In our case - only one sample of observable was counted, which leads to lower significance
- The difference between PL and Znet 3 is expected, within the confidence belt

" To increase the sensitivity to a Higgs boson signal, the events are separated into ten mutually exclusive categories having different mass resolutions and signal-to-background ratios. "





[4] http://arxiv.org/abs/arXiv:1207.7214