New Physics Learning Machine (NPLM)

Gaia Grosso[1,2]

[1]University and INFN of Padova, [2]CERN

Istituto Nazionale di Fisica Nucleare
SEZIONE DI PADOVA
Servizio Fondi Esterni

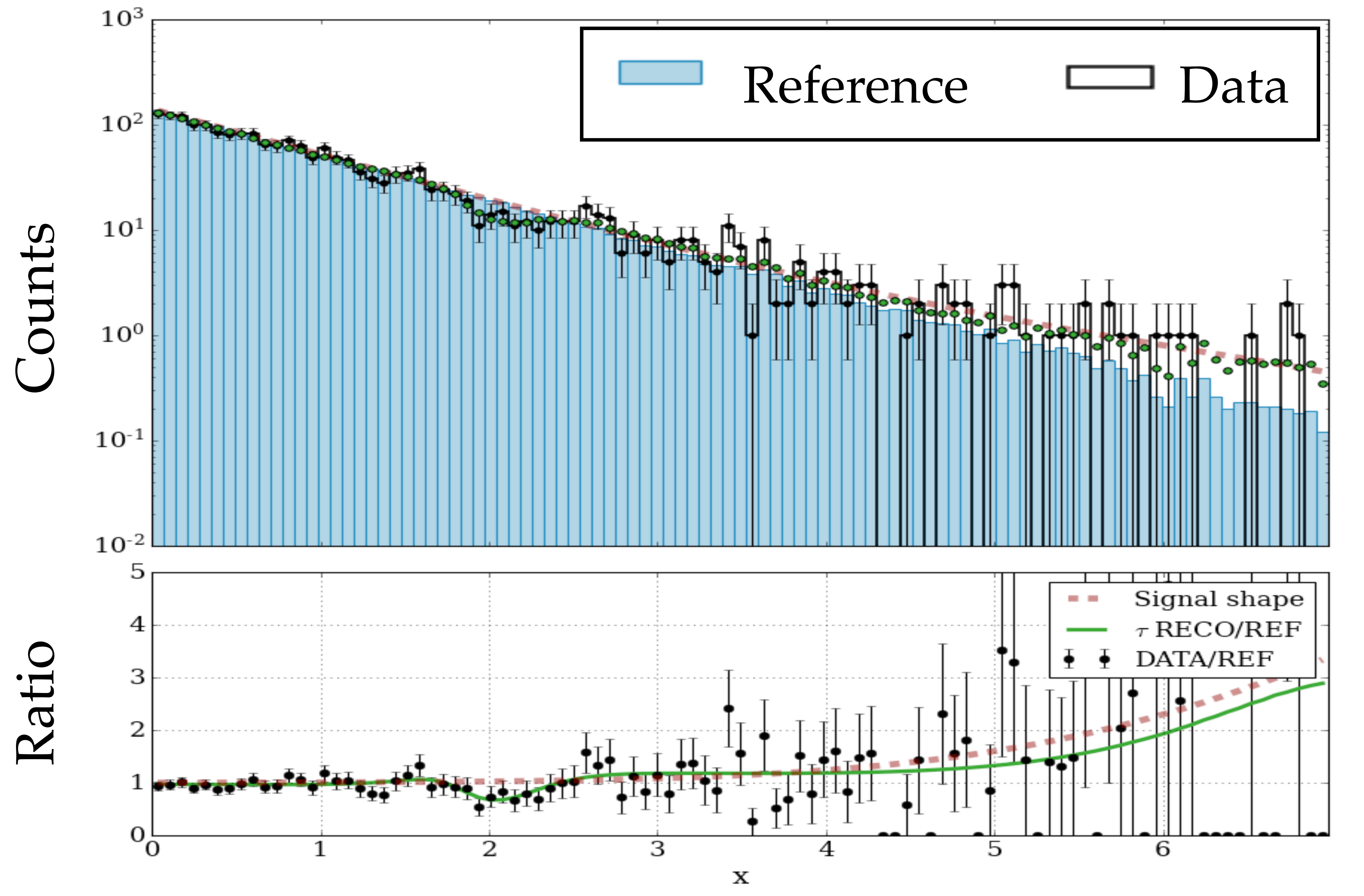
Dipartimento di Fisica e Astronomia Galileo Galilei

In this talk

What is NPLM?

tool for goodness-of-fit tests, based on the principle of maximum-likelihood-ratio.

- Signal agnostic
- Multivariate
- Unbinned



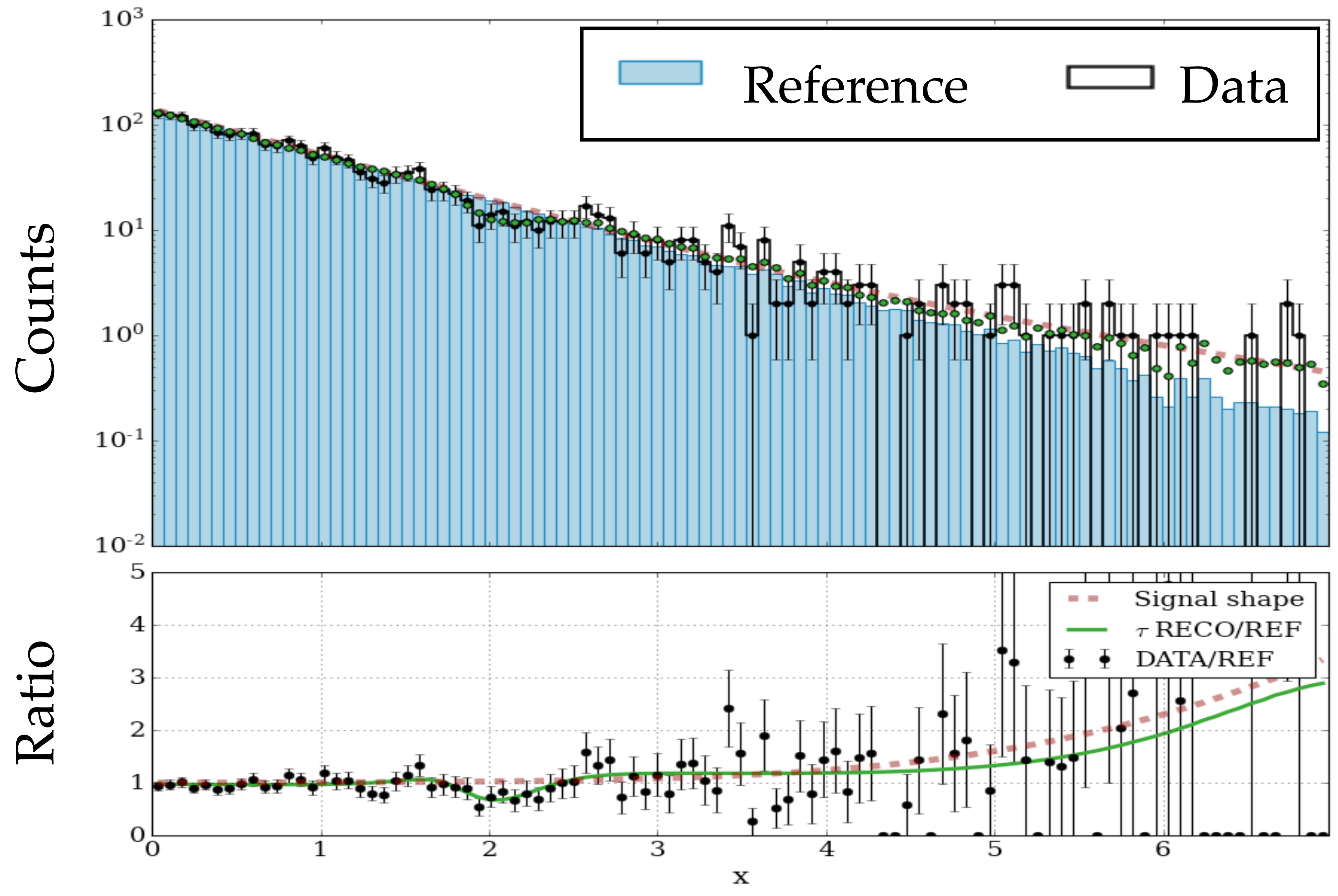
How well does the Reference model describe the data?

In this talk

What is NPLM *for*?

tool for goodness-of-fit tests, based on the principle of maximum-likelihood-ratio.

- Model-independent New Physics searches at collider experiments
- Data quality monitor (DQM)
- Generator validation



How well does the Reference model describe the data?

In this talk

How does it work?

Complete analysis strategy to test the data for departures from a Reference model (from the data to a *p*-value, taking care of *systematic uncertainties* if needed).

- Main concepts and implementation using NN

  "Learning New Physics from a Machine" - d'Agnolo, Wulzer, Phys. Rev. D (2018)

  "Learning Multivariate New Physics" - d'Agnolo, Grosso, Pierini, Wulzer, Zanetti, Eur. Phys. J. C 81, 89 (2021)

- Systematic uncertainties

  "Learning New Physics from an Imperfect Machine" - d'Agnolo, Grosso, Pierini, Wulzer, Zanetti Eur. Phys. J. C 82, 275 (2022)

- Hands on a 1D toy model (material for Q&A)

NPLM
main concepts
and NN implementation

New Physics Learning Machine (NPLM)

Main Concepts (negligible uncertainties)

- Goal: performing a **maximum-likelihood-ratio hypothesis test**
  End-to-end strategy, from the data to a *p*-value for the discovery (frequentist approach)

$$t(\mathcal{D}) = \max_{\mathbf{w}} \left[2 \log \frac{\mathcal{L}(\mathcal{D} \,|\, \mathrm{H}_{\mathbf{w}})}{\mathcal{L}(\mathcal{D} \,|\, \mathrm{R}_{\mathbf{0}})} \right]$$

$\mathrm{R}_{\mathbf{0}}$: null hypothesis
$\mathrm{H}_{\mathbf{w}}$: alternative hypothesis

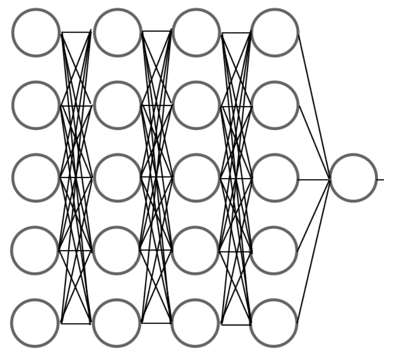
- Exploiting a Neural Network (NN) to **parametrize** the data distribution in terms of a Reference distribution ($\mathrm{R}_{\mathbf{0}}$)

$$n(x \,|\, \mathrm{T}) \approx n(x \,|\, \mathrm{H}_{\hat{\mathbf{w}}}) = n(x \,|\, \mathrm{R}_0)\, e^{f(x,\hat{\mathbf{w}})}$$

True (T) data distribution — Unknown

Data distribution learnt by the NN — Alternative hypothesis

Reference distribution — Null hypothesis (SM)



NN model

- **Signal-model-independent**: reduced assumptions on the signal hypothesis

"Learning New Physics from a Machine" Phys. Rev. D

New Physics Learning Machine (NPLM)

Main Concepts (negligible uncertainties)

Maximum Likelihood from minimal loss:

Test statistic

$$\bar{t}(\mathcal{D}) = 2\max_{\mathbf{w}} \log\left[\frac{\mathcal{L}(\mathcal{D}\,|\,\mathrm{H}_{\mathbf{w}})}{\mathcal{L}(\mathcal{D}\,|\,\mathrm{R}_{\mathbf{0}})}\right] = -2\min_{\mathbf{w}}\{\bar{L}[f(\cdot;\,\mathbf{w})]\}$$

Loss function

$$\bar{L}\,[f(x;\,\mathbf{w})] = -\sum_{x\in\mathcal{D}}[\,f(x;\,\mathbf{w})] + \sum_{x\in\mathcal{R}} w_x\left[e^{f(x;\,\mathbf{w})} - 1\right]$$

w: trainable parameters on the NN model

D: data sample

R: reference sample (built according to the R_0 hypothesis); could be weighted (w)

Assumptions:

- $N_R \gg N_D$ the statistical fluctuations of the reference sample are negligible.
- the weights of the reference sample (w) are such that the reference sample is normalised to match the data sample luminosity $\sum_{x\in R} w_x = N(R_0)$

“Learning New Physics from a Machine” Phys. Rev. D

New Physics Learning Machine (NPLM)

Main Concepts (negligible uncertainties)

Maximum Likelihood from minimal loss:

Test statistic

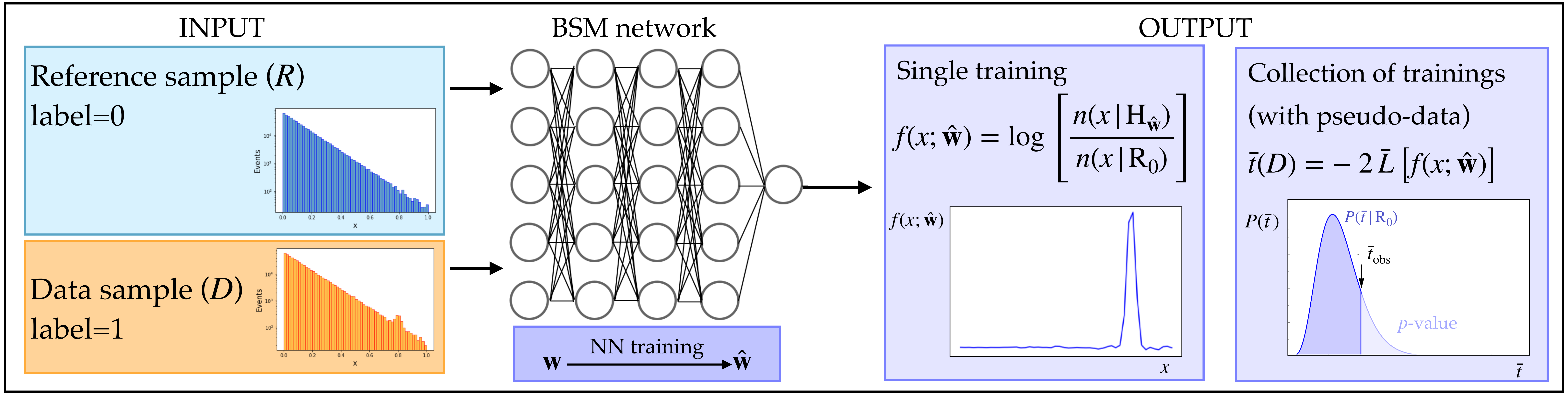
$$\bar{t}(\mathcal{D}) = 2\max_{\mathbf{w}} \log\left[\frac{\mathcal{L}(\mathcal{D}\,|\,\mathrm{H}_{\mathbf{w}})}{\mathcal{L}(\mathcal{D}\,|\,\mathrm{R}_0)}\right] = -2\min_{\mathbf{w}}\{\bar{L}[f(\cdot;\,\mathbf{w})]\}$$

w: trainable parameters on the NN model
D: data sample
R: reference sample (built according to the R_0 hypothesis); could be weighted (w)

Loss function

$$\bar{L}\,[f(x;\,\mathbf{w})] = -\sum_{x\in\mathcal{D}}[\,f(x;\,\mathbf{w})] + \sum_{x\in\mathcal{R}} w_x\left[e^{f(x;\,\mathbf{w})} - 1\right]$$

Assumptions:
- $N_R \gg N_D$ the statistical fluctuations of the reference sample are negligible.
- the weights of the reference sample (w) are such that the reference sample is normalised to match the data sample luminosity $\sum_{x\in R} w_x = N(R_0)$



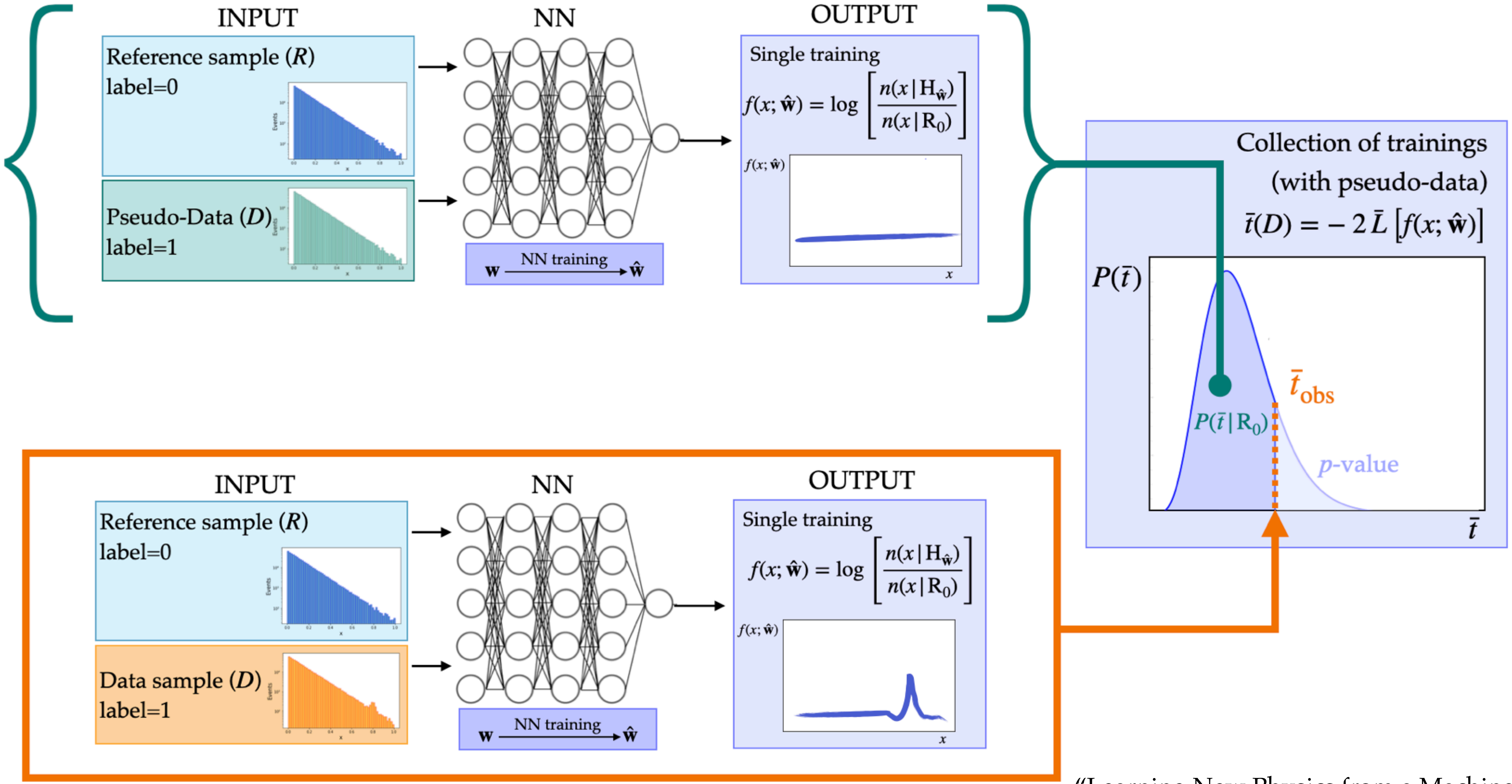
"Learning New Physics from a Machine" Phys. Rev. D

New Physics Learning Machine (NPLM)

Main Concepts (negligible uncertainties)

1. Run NPLM on toy experiments to simulated the response under the null hypothesis

2. Run NPLM the data of interest and check where the test outcome falls to compute an exclusion p-value



"Learning New Physics from a Machine" Phys. Rev. D

New Physics Learning Machine (NPLM)

Main Concepts (negligible uncertainties)

Asymptotic formula for the $\bar{t}$ distribution under R_0:

Wilks-Wald theorem:

Θ_0: set of parameters describing H_0

Θ_1: set of parameters describing H_1

If $H_0 \subseteq H_1$, then under the H_0 hypothesis the test statistic

$$t(\mathcal{D}) = 2\,\log\frac{\mathcal{L}(\mathrm{H_1}|\mathcal{D})}{\mathcal{L}(\mathrm{H_0}|\mathcal{D})}$$

asymptotically follows a χ^2_{df} distribution with $df = |\Theta_1| - |\Theta_0|$

→

If the Wilks' theorem hold, the target distribution for $\bar{t}$ under the R_0 hypothesis is a χ^2_{df} with $df = |\mathbf{w}|$.

Due to the **finite size** of the training samples, the **sparsity** of the data (especially in multivariate problems) and the **approximation** errors, the distribution of $\bar{t}(D)$ under R_0 does not follow the target $\chi^2_{|\mathbf{w}|}$ by default.

→ a **(NN) MODEL REGULARIZATION** procedure can solve this problem!

"Learning Multivariate New Physics" Eur. Phys. J. C

New Physics Learning Machine (NPLM)

Main Concepts (negligible uncertainties)

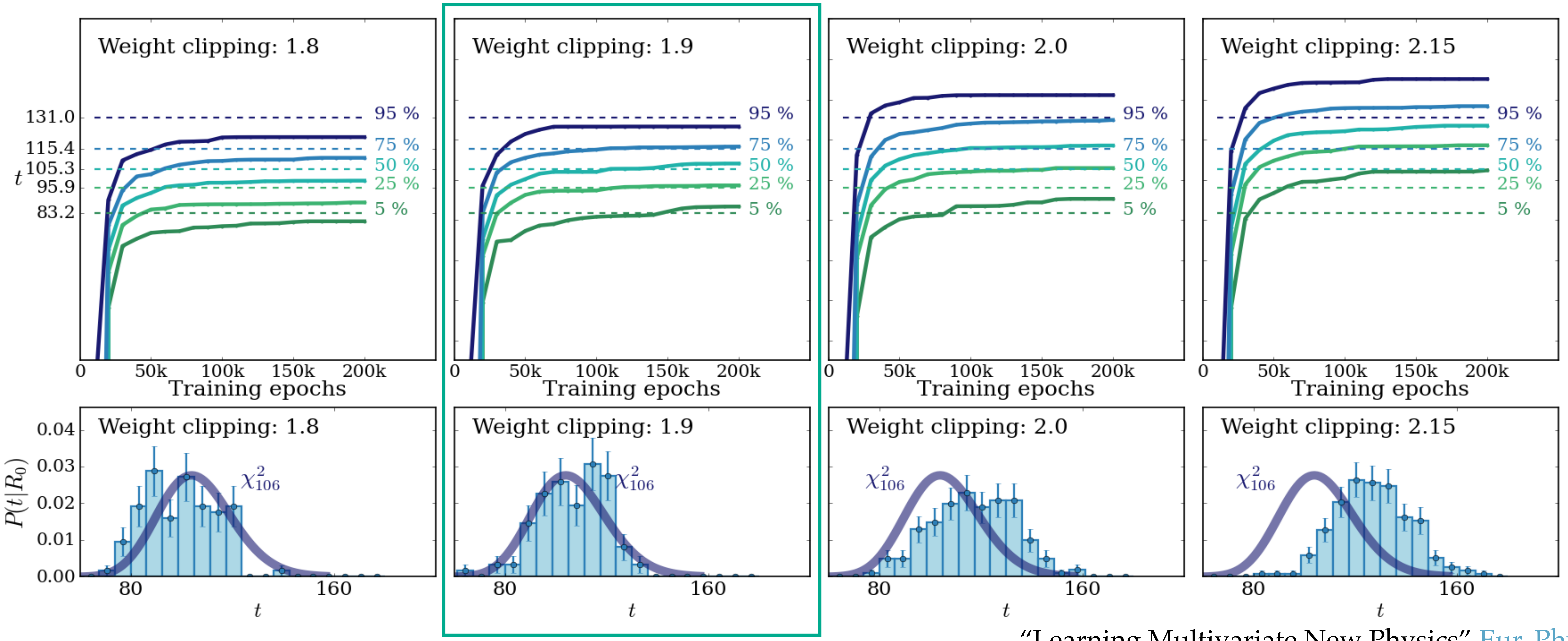
NN Model regularization:

Weight clipping parameter:
Upper boundary to the magnitude that each trainable parameter can assume during the training.

For a chosen NN architecture, tuning the weight clipping allows to recover a good agreement of the empirical distribution of $\bar{t}$ under R_0 with the **target** $\chi^2_{|\mathbf{w}|}$ distribution.

Example:
NN model: 5-7-7-1,
Number of parameters:106

Legend:
- Percentiles of the empirical $\bar{t}$ distribution under R_0
- Percentiles of the target $\chi^2_{|\mathbf{w}|}$
- Empirical $\bar{t}$ distribution under R_0
- Target $\chi^2_{|\mathbf{w}|}$



"Learning Multivariate New Physics" Eur. Phys. J. C

NPLM systematic uncertainties

New Physics Learning Machine (NPLM)

Including systematic uncertainties

Test statistic

$$t(\mathcal{D},\,\mathcal{A}) = 2\log\left[\frac{\max\limits_{\mathbf{w},\,\boldsymbol{\nu}} \mathcal{L}(\mathrm{H}_{\mathbf{w},\,\boldsymbol{\nu}}|\mathcal{D},\,\mathcal{A})}{\max\limits_{\boldsymbol{\nu}} \mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{D},\,\mathcal{A})}\right] = 2\log\left[\frac{\max\limits_{\mathbf{w},\,\boldsymbol{\nu}} \mathcal{L}(\mathrm{H}_{\mathbf{w},\,\boldsymbol{\nu}}|\mathcal{D})\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\max\limits_{\boldsymbol{\nu}} \mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{D})\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}\right]$$

$\mathbf{w}$: trainable parameters on the NN model

$\boldsymbol{\nu}$: set of nuisance parameters modelling the uncertainties effects

$\mathcal{D}$: data sample

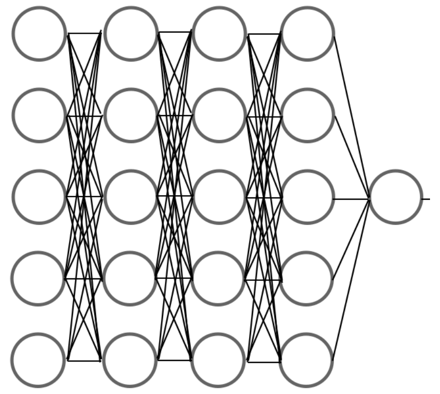
$\mathcal{A}$: auxiliary sample (used to constrain $\boldsymbol{\nu}$)

New parametrization

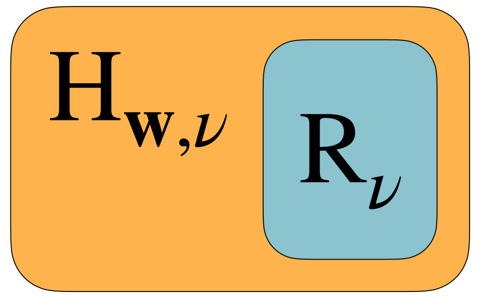
$$n(x\,|\,\mathrm{T}) \approx n(x\,|\,\mathrm{H}_{\hat{\mathbf{w}},\,\hat{\nu}}) = n(x\,|\,\mathrm{R}_0)\,\frac{n(x\,|\,\mathrm{R}_{\hat{\nu}})}{n(x\,|\,\mathrm{R}_0)}\,e^{f(x;\,\hat{\mathbf{w}})}$$

True (T) data distribution	Data distribution learnt by the NN	Reference distribution	New term containing the dependence on ν $r(x;\nu)$
Unknown	Alternative hypothesis	Null hypothesis	

NN model



Note:
This parametrization choice guarantees $\mathrm{R}_\nu \subseteq \mathrm{H}_{\mathbf{w},\nu}$
($\mathrm{R}_\nu = \mathrm{H}_{\mathbf{w},\nu}$ for $f(\,\cdot\,;\,\mathbf{w}) \equiv 0$)



"Learning New Physics from an Imperfect Machine" Eur. Phys. J. C

New Physics Learning Machine (NPLM)

Including systematic uncertainties

Maximum Likelihood from minimal loss:

Test statistic

$$t(\mathcal{D},\,\mathcal{A}) = 2\log\left[\frac{\max_{\mathbf{w},\boldsymbol{\nu}} \mathcal{L}(\mathrm{H}_{\mathbf{w},\boldsymbol{\nu}}|\mathcal{D},\,\mathcal{A})}{\max_{\boldsymbol{\nu}} \mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{D},\,\mathcal{A})}\right] = 2\log\left[\frac{\max_{\mathbf{w},\nu} \mathcal{L}(\mathrm{H}_{\mathbf{w},\boldsymbol{\nu}}|\mathcal{D})\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\max_{\boldsymbol{\nu}} \mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{D})\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}\right]$$

$$= \tau(\mathcal{D},\,\mathcal{A}) - \Delta(\mathcal{D},\,\mathcal{A})$$

$\mathbf{w}$: trainable parameters on the NN model

$\boldsymbol{\nu}$: set of nuisance parameters modelling the uncertainties effects

$\mathcal{D}$: data sample

$\mathcal{A}$: auxiliary sample (used to constrain $\boldsymbol{\nu}$)

Tau term:

$$\tau(\mathcal{D},\,\mathcal{A}) = 2\max_{\mathbf{w},\boldsymbol{\nu}}\log\left[\frac{\mathcal{L}(\mathrm{H}_{\mathbf{w}},\,\boldsymbol{\nu}|\mathcal{D})\,\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\mathcal{L}(\mathrm{R}_{\mathbf{0}}|\mathcal{D})\,\mathcal{L}(\mathbf{0}|\mathcal{A})}\right] = -2\min_{\mathbf{w},\boldsymbol{\nu}} L\left[f(x,\mathbf{w}),\boldsymbol{\nu};\widehat{\delta}(x)\right]$$

Contains the dependence on a NN model

Delta term:

$$\Delta(\mathcal{D},\,\mathcal{A}) = 2\max_{\boldsymbol{\nu}}\log\left[\frac{\mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{D})\,\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\mathcal{L}(\mathrm{R}_{\mathbf{0}}|\mathcal{D})\,\mathcal{L}(\mathbf{0}|\mathcal{A})}\right] = -2\min_{\boldsymbol{\nu}} L\left[\boldsymbol{\nu};\widehat{\delta}(x)\right]$$

Built on the knowledge of the Reference model (purely SM term)

$$r(x;\nu) = \frac{n(x\,|\,\mathrm{R}_\nu)}{n(x\,|\,\mathrm{R}_0)}$$

Taylor's expansion learning:

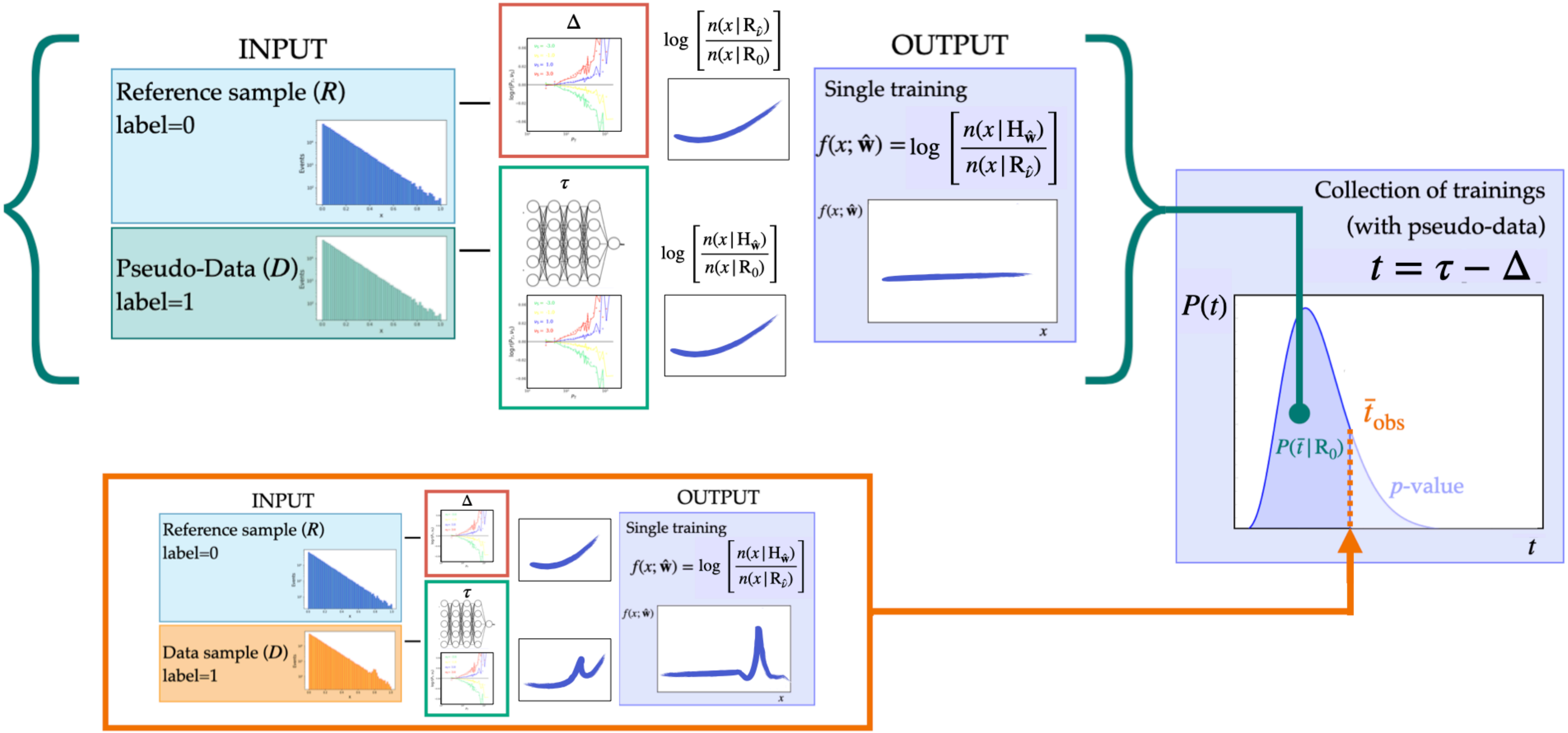
$$\hat{r}(x;\nu) = \exp\left[\widehat{\delta}_1(x)\,\nu + \widehat{\delta}_2(x)\,\nu^2 + \ldots\right]$$

NN 1 NN2 …

"Learning New Physics from an Imperfect Machine" Eur. Phys. J. C

New Physics Learning Machine (NPLM)

Including systematic uncertainties

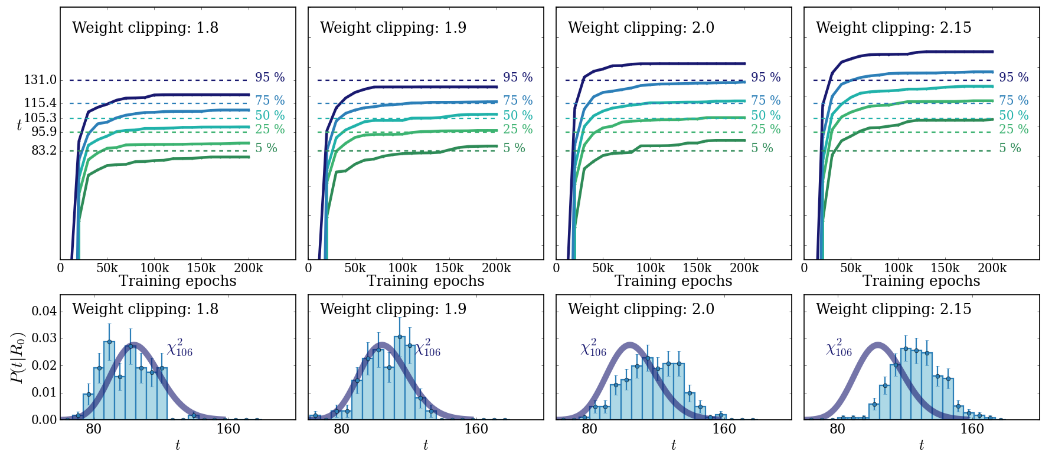


"Learning New Physics from an Imperfect Machine" Eur. Phys. J. C

New Physics Learning Machine (NPLM)

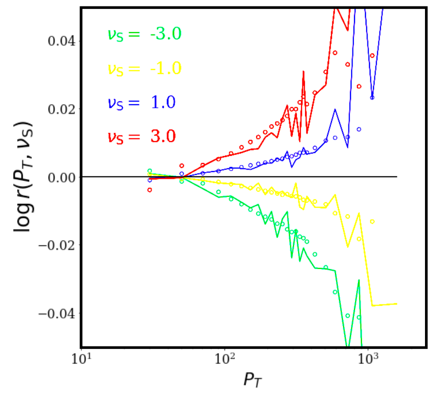
Preparation steps (before looking at the data):

- **NN model selection**

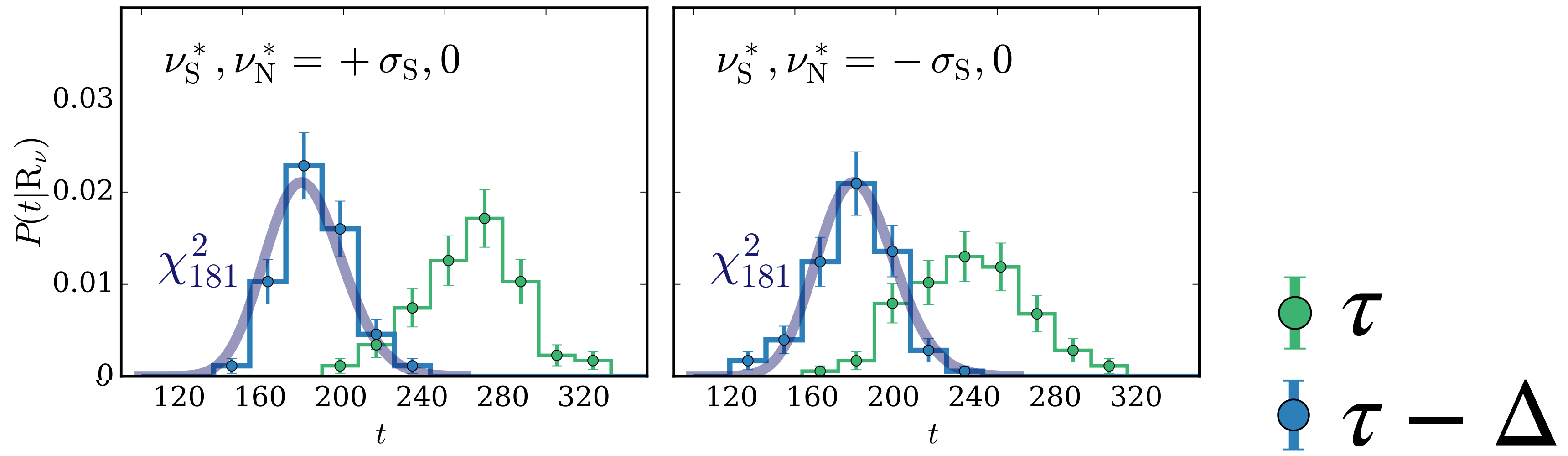


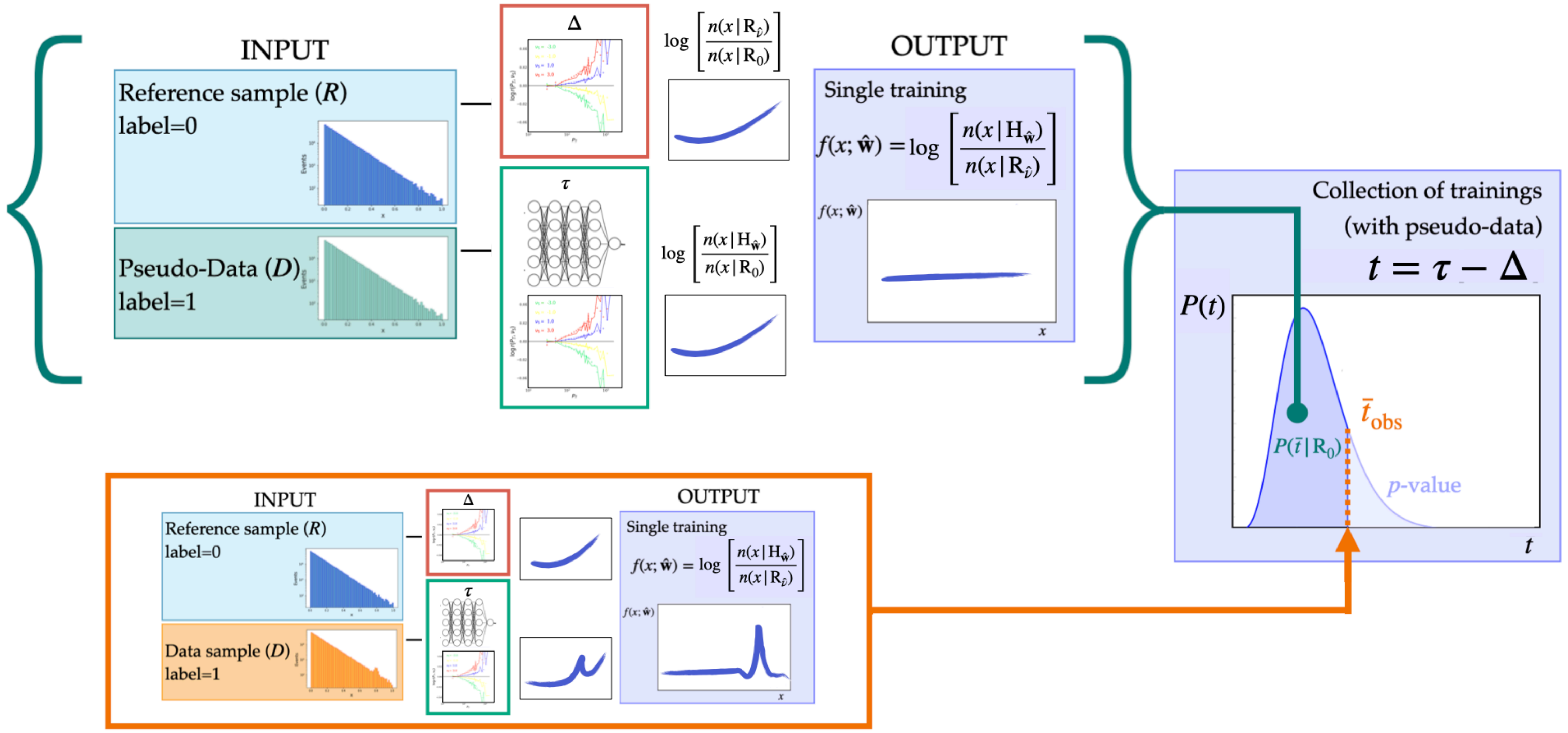
- **Parametric NN for Taylor's expansion on ν**

$$\hat{r}(x;\nu) = \exp\left[\hat{\delta}_1(x)\,\nu + \hat{\delta}_2(x)\,\nu^2 + \ldots\right]$$



- **Validation:** $\mathcal{D} \sim \mathrm{R}_{\boldsymbol{\nu}^*}, \quad \boldsymbol{\nu}^* = \pm\sigma_{\boldsymbol{\nu}}$





"Learning New Physics from an Imperfect Machine" Eur. Phys. J. C

Hands on a 1D toy model

Hands on a 1D toy model

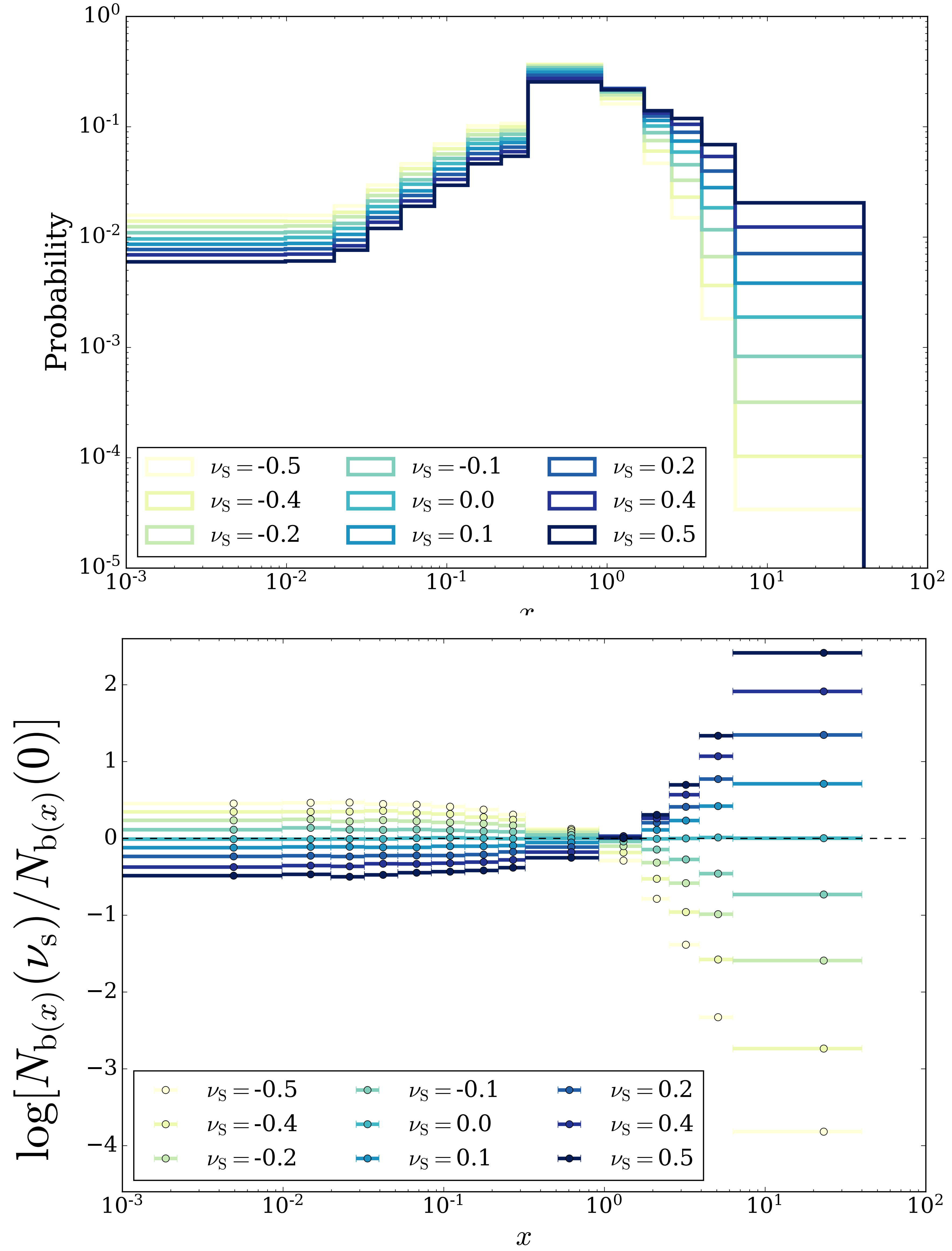
Reference model:

Exponentially falling distribution affected by:

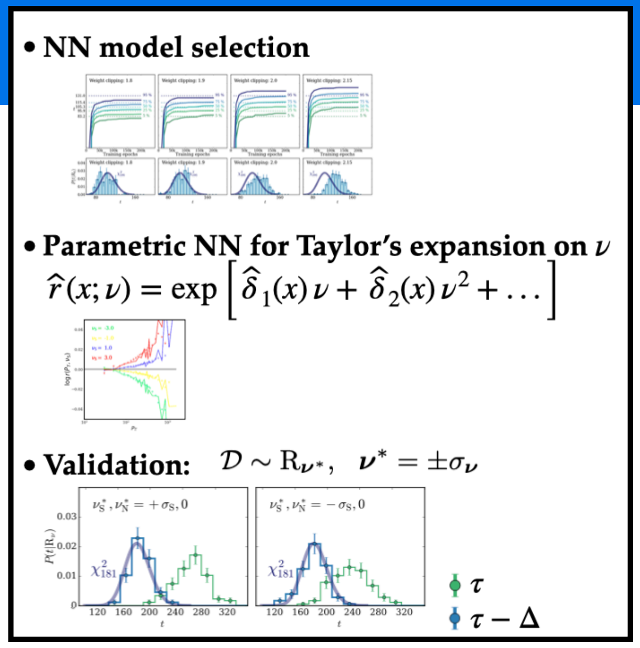
- a scale uncertainty (ν_S)
- a normalization uncertainty (ν_N)

$$n(x|\mathrm{R}_{\boldsymbol{\nu}}) = n(x|\mathrm{R}_{\nu_N,\nu_S}) = \mathrm{N}(\mathrm{R}_{\mathbf{0}})\ \exp\left[-x\,e^{-\nu_S} - \nu_S + \nu_N\right]$$

Scale effect on the distribution



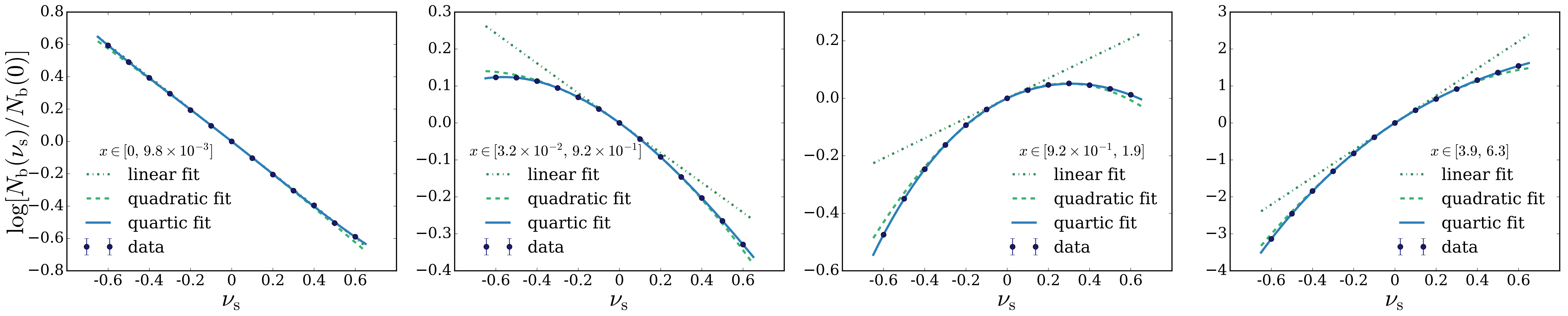
Hands on a 1D toy model



Learning the coefficient of the Taylor Expansion for ν_S:

Preliminary study

Preliminary binned analysis to determine the proper order for the Taylor's expansion



Taylor's expansion learning

Training a neural network model to learn each coefficient of the Taylor's expansion of $r(x;\nu) = \dfrac{n(x|\mathrm{R}_\nu)}{n(x|\mathrm{R}_0)}$

Input samples

ν_1	ν_2	…	ν_n
$S_0(\nu_1) \sim \mathrm{R}_0$	$S_0(\nu_2) \sim \mathrm{R}_0$	…	$S_0(\nu_n) \sim \mathrm{R}_0$
$S_1(\nu_1) \sim \mathrm{R}_{\nu_1}$	$S_1(\nu_2) \sim \mathrm{R}_{\nu_2}$		$S_1(\nu_n) \sim \mathrm{R}_{\nu_n}$

→ $\hat{\delta}_1$, $\hat{\delta}_2$ →

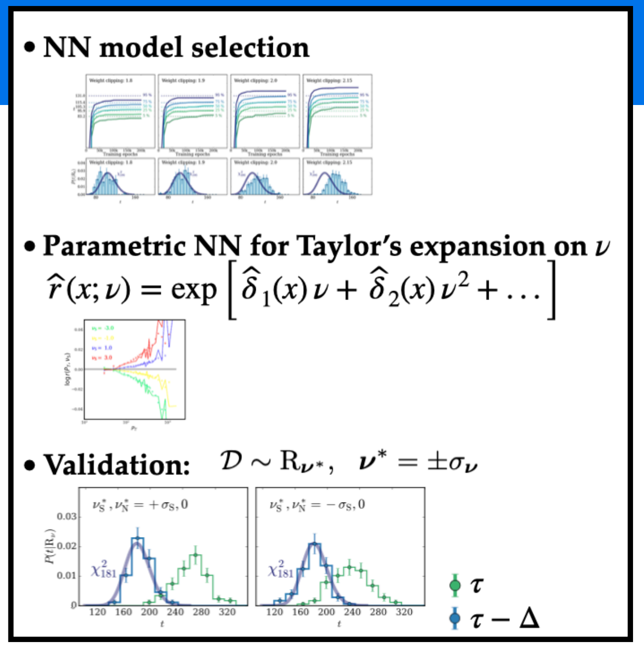
$$\hat{r}(x;\nu) = \exp\left[\hat{\delta}_1(x)\,\nu + \hat{\delta}_2(x)\,\nu^2\right]$$

Loss function*

$$L[\hat{\delta}(\cdot)] = \sum_{\nu_i}\left[\sum_{e\in S_0(\nu_i)} w_e c(x_e)^2 + \sum_{e\in S_1(\nu_i)} w_e\left[1 - c(x_e)\right]^2\right], \qquad c(x) = \frac{1}{1+\hat{r}(x;\nu)}$$

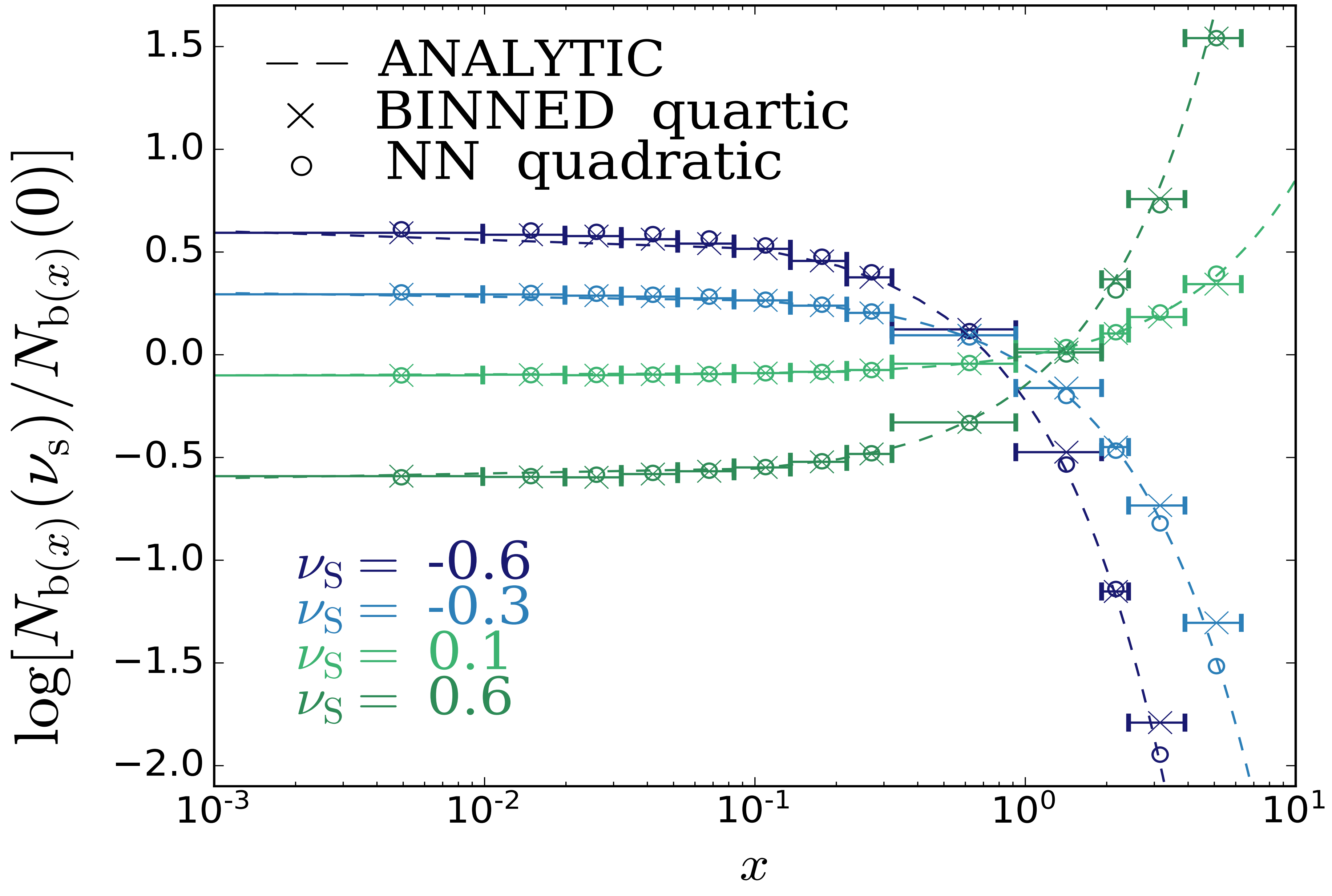
* *Parametrized classifiers for optimal EFT sensitivity* arXiv:2007.10356

Hands on a 1D toy model



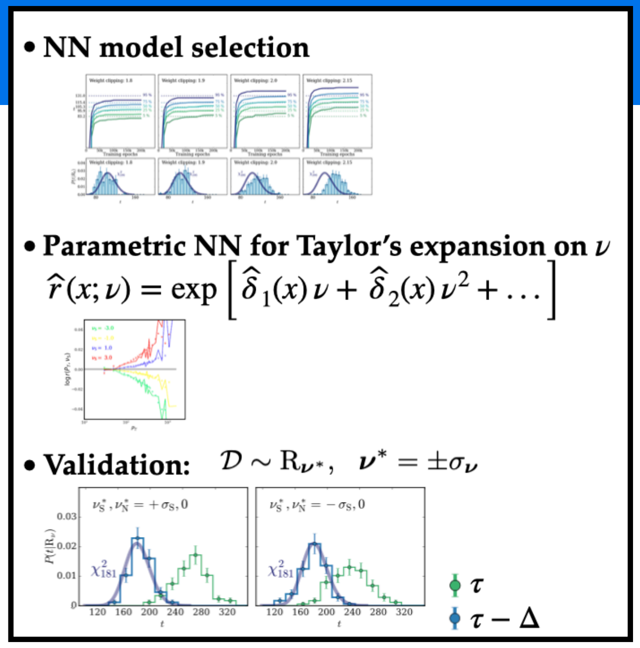
Learning the coefficient of the Taylor Expansion for ν_S:

$$\log r_{b_i}(\boldsymbol{\nu}) = \nu_N + \nu_S\,\widehat{\delta}_{1,b_i} + \frac{1}{2}\nu_S^2\,\widehat{\delta}_{2,b_i} \qquad \forall i \in [1, \dots N_{bins}]$$



"Learning New Physics from an Imperfect Machine" Eur. Phys. J. C

Hands on a 1D toy model



NN model selection:

Weight clipping tuning

(In the R_0 hypothesis)

Here we used:

- 2000 data
- 200 000 reference (R_0)
- Architecture 1-4-1 (9 dof, very simple!)

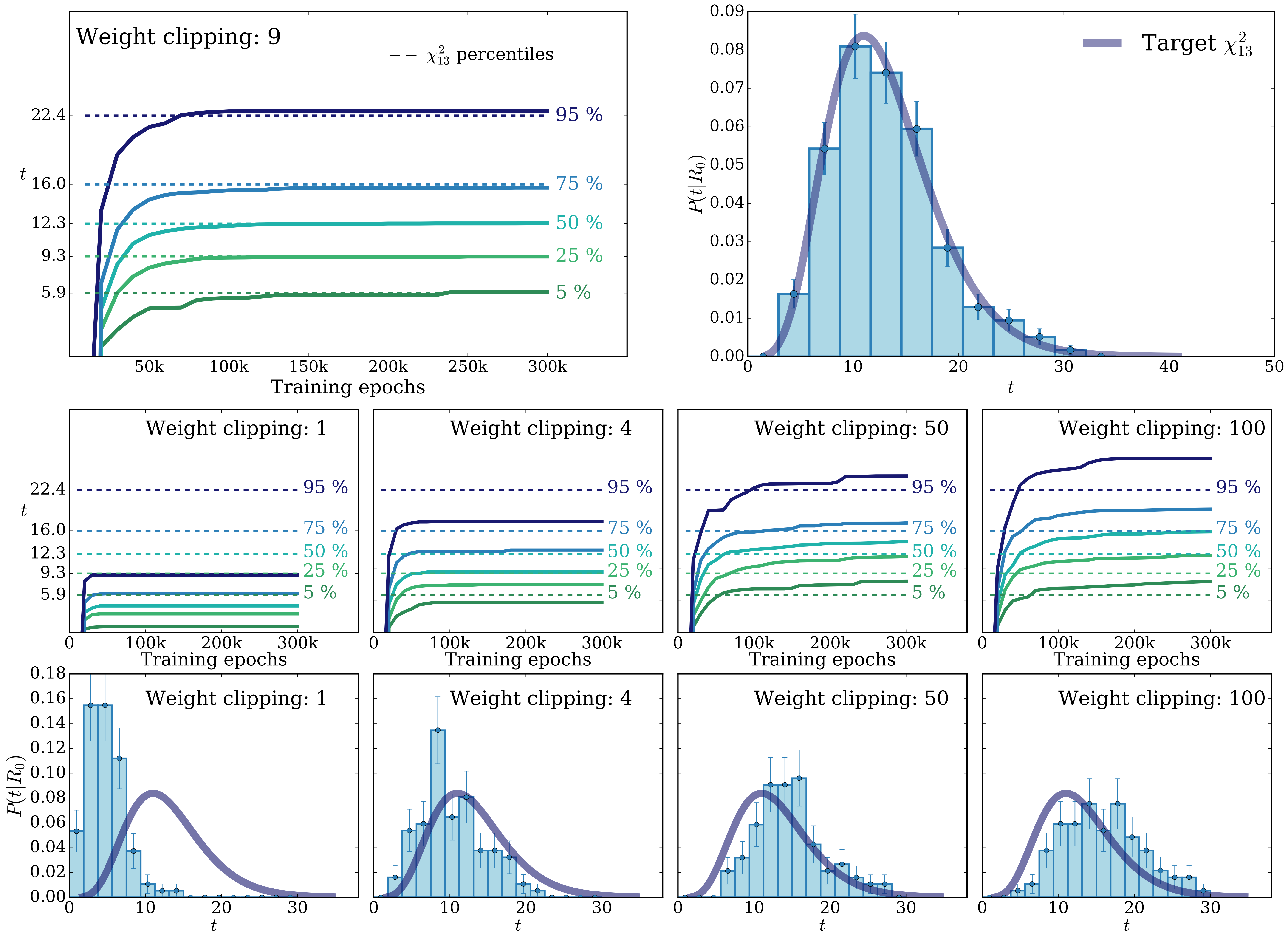
Legend:

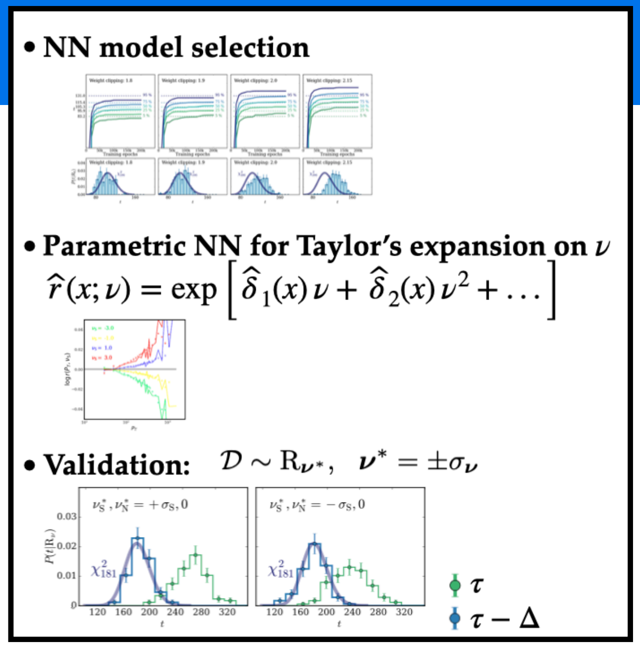
Percentiles of the empirical $\bar{t}$ distribution under R_0

Percentiles of the target $\chi^2_{|\mathbf{w}|}$

Empirical $\bar{t}$ distribution under R_0

Target $\chi^2_{|\mathbf{w}|}$



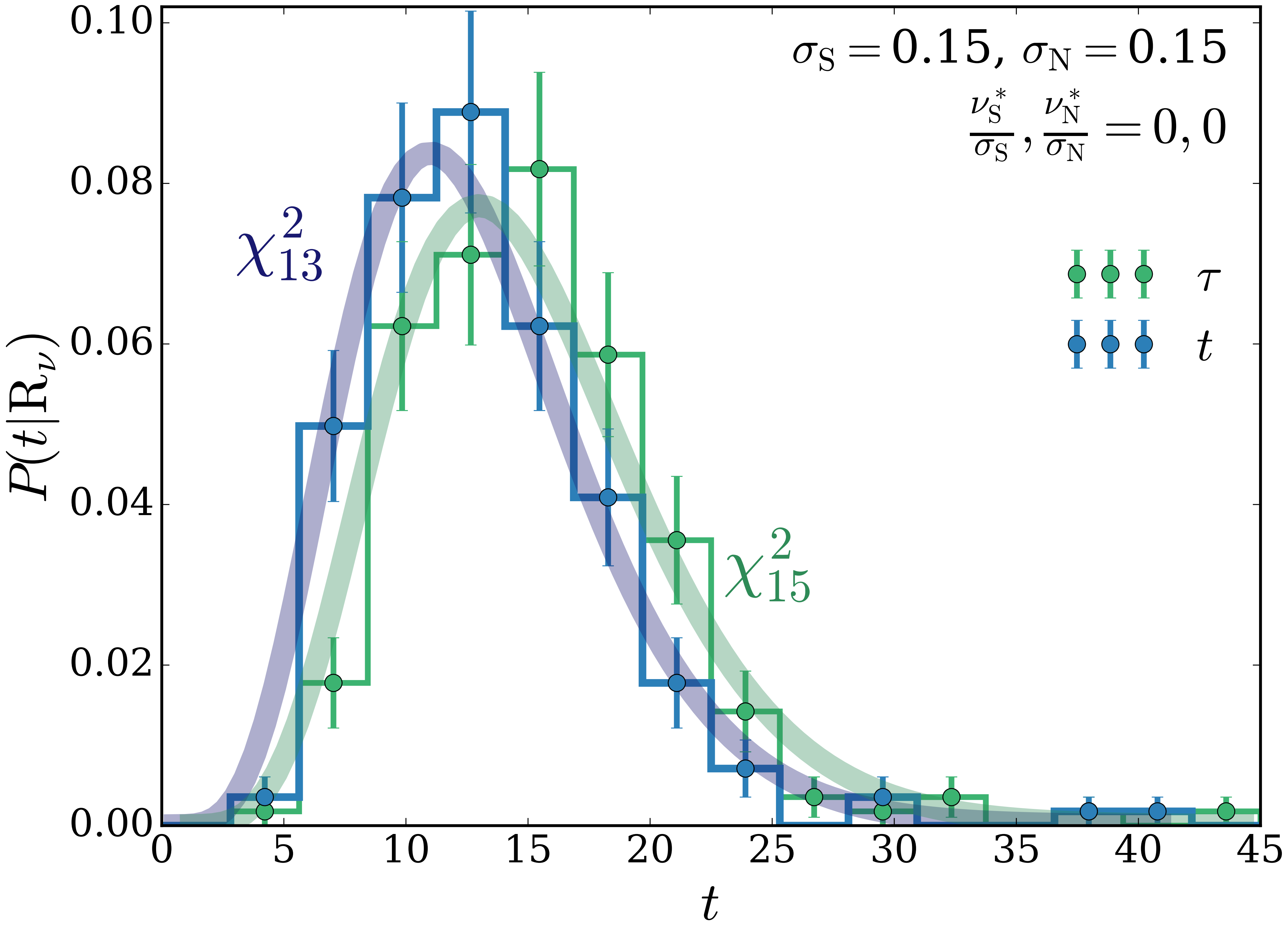


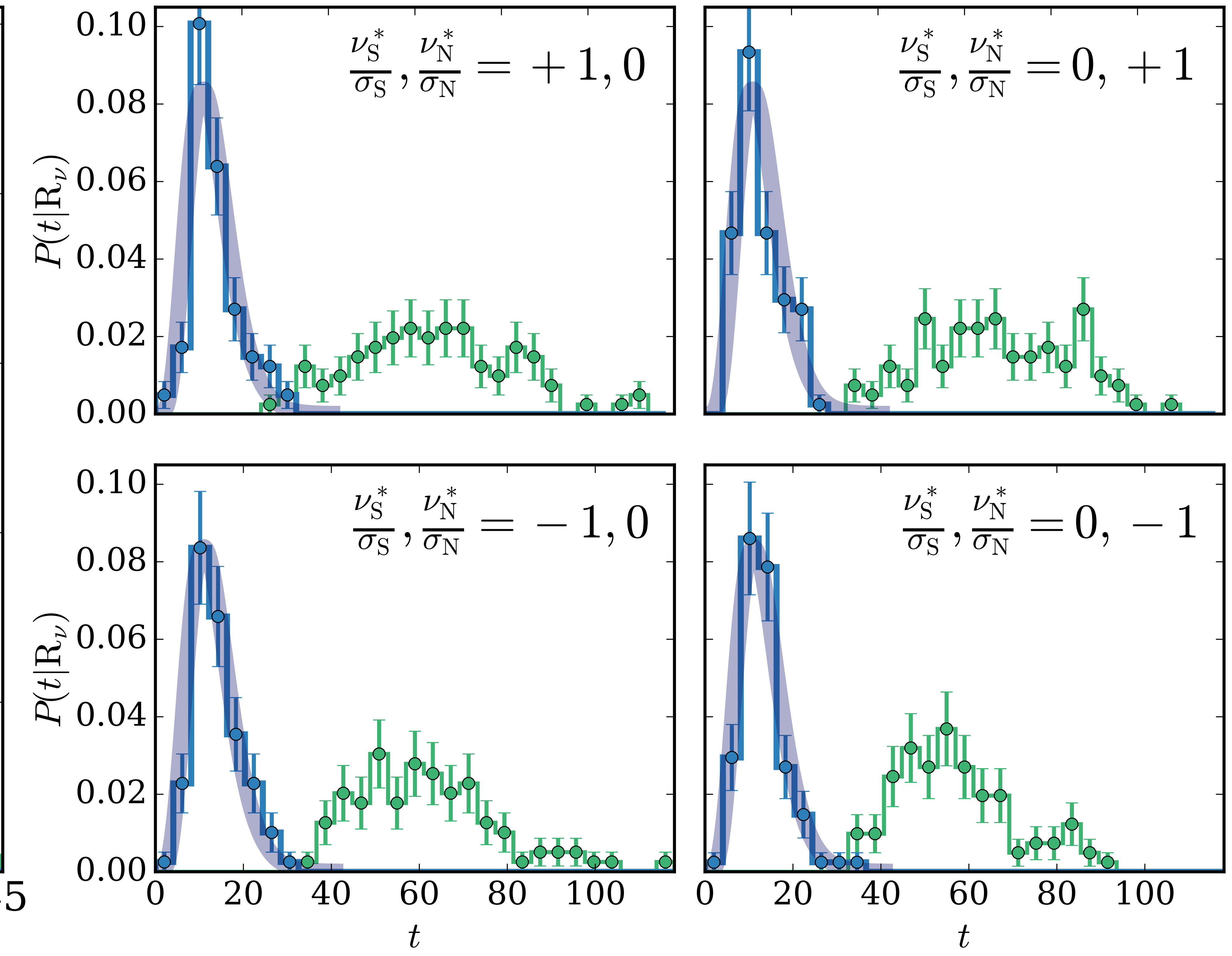
Hands on a 1D toy model

Validation

Verifying that the $\chi^2_{|\mathrm{w}|}$ is recovered

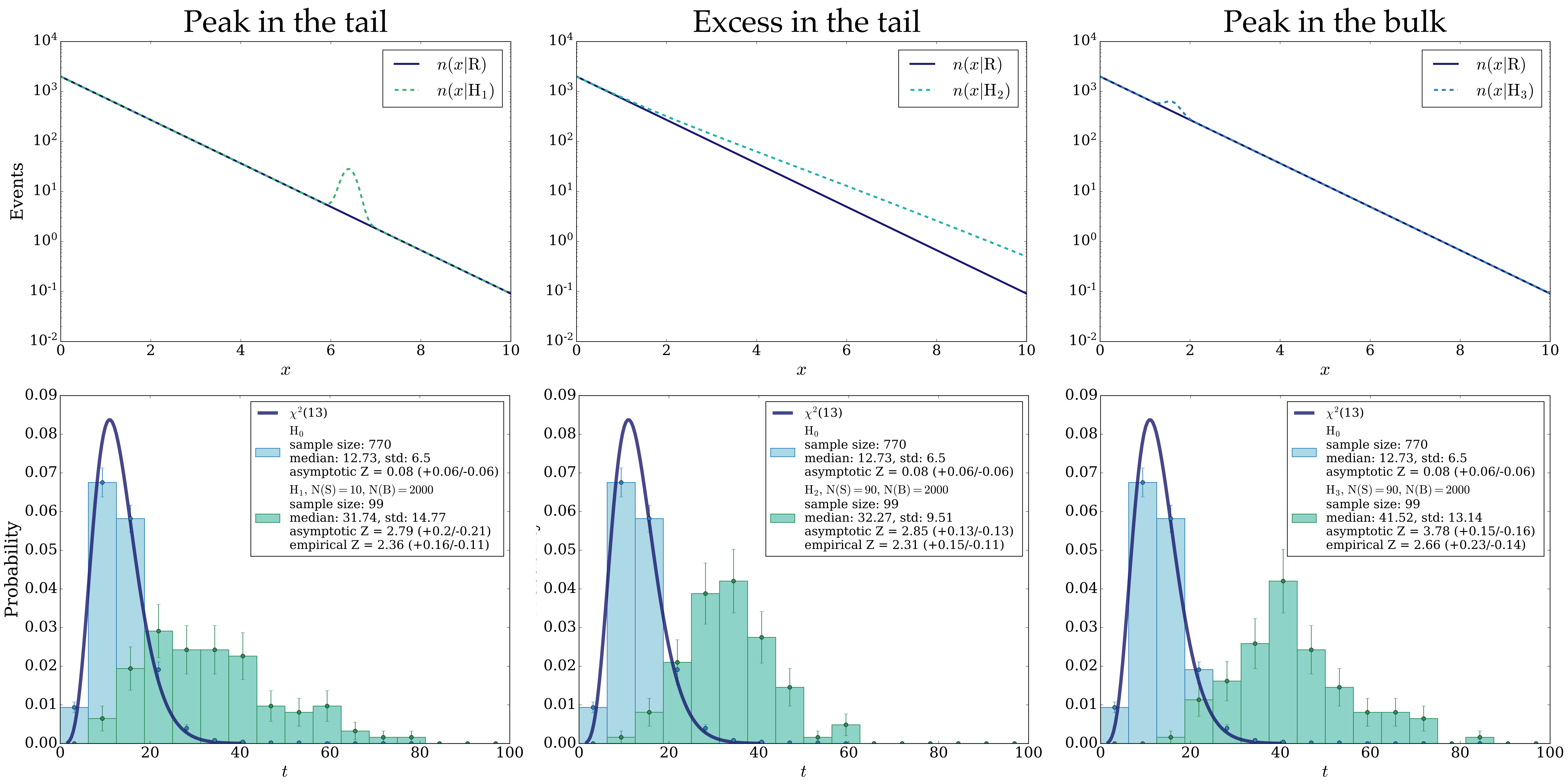
$\mathcal{D} \sim \mathrm{R}_{\boldsymbol{\nu}^*}, \quad \boldsymbol{\nu}^* = \pm\sigma_{\boldsymbol{\nu}}$





Hands on a 1D toy model

Inject signals

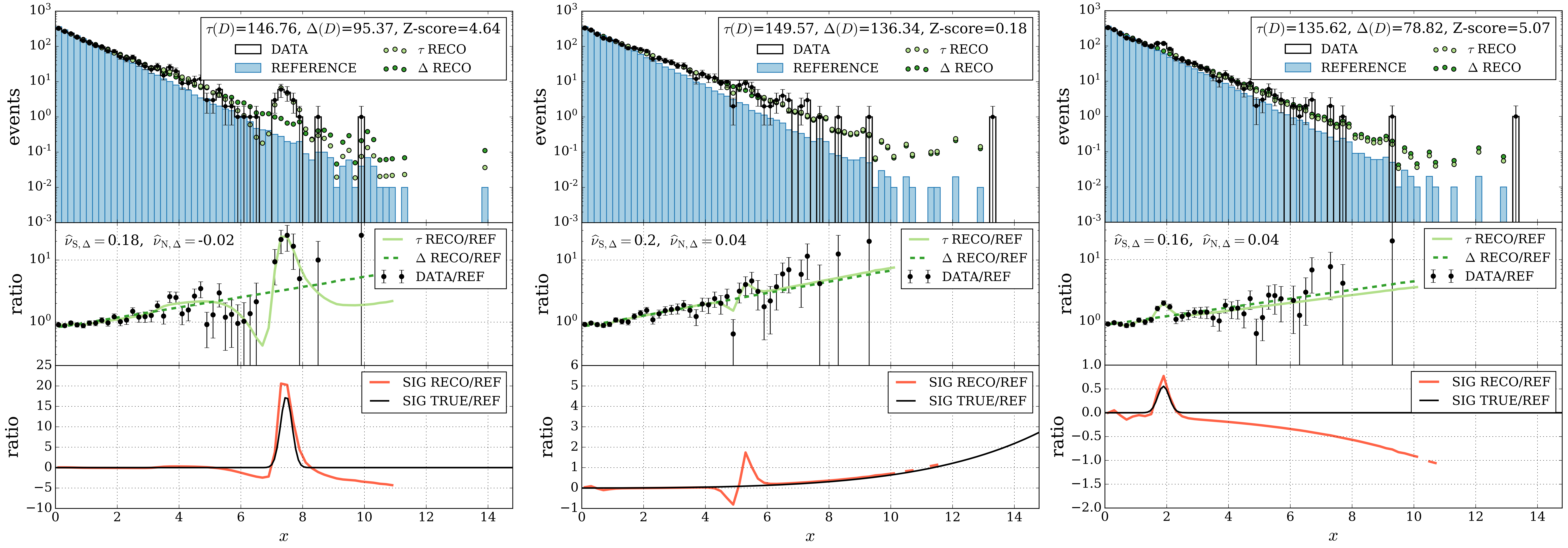


Hands on a 1D toy model

Interpret the result: anomaly characterisation

$$\tau \text{ reconstruction: } n(x\,|\,\mathrm{H}_{\hat{\mathbf{w}},\hat{\nu}}) = n(x\,|\,\mathrm{R}_0)\,\frac{n(x\,|\,\mathrm{R}_{\hat{\nu}})}{n(x\,|\,\mathrm{R}_0)}\,e^{f(x;\,\hat{\mathbf{w}})}$$

$$\Delta \text{ reconstruction: } n(x\,|\,\mathrm{R}_{\hat{\nu}})$$

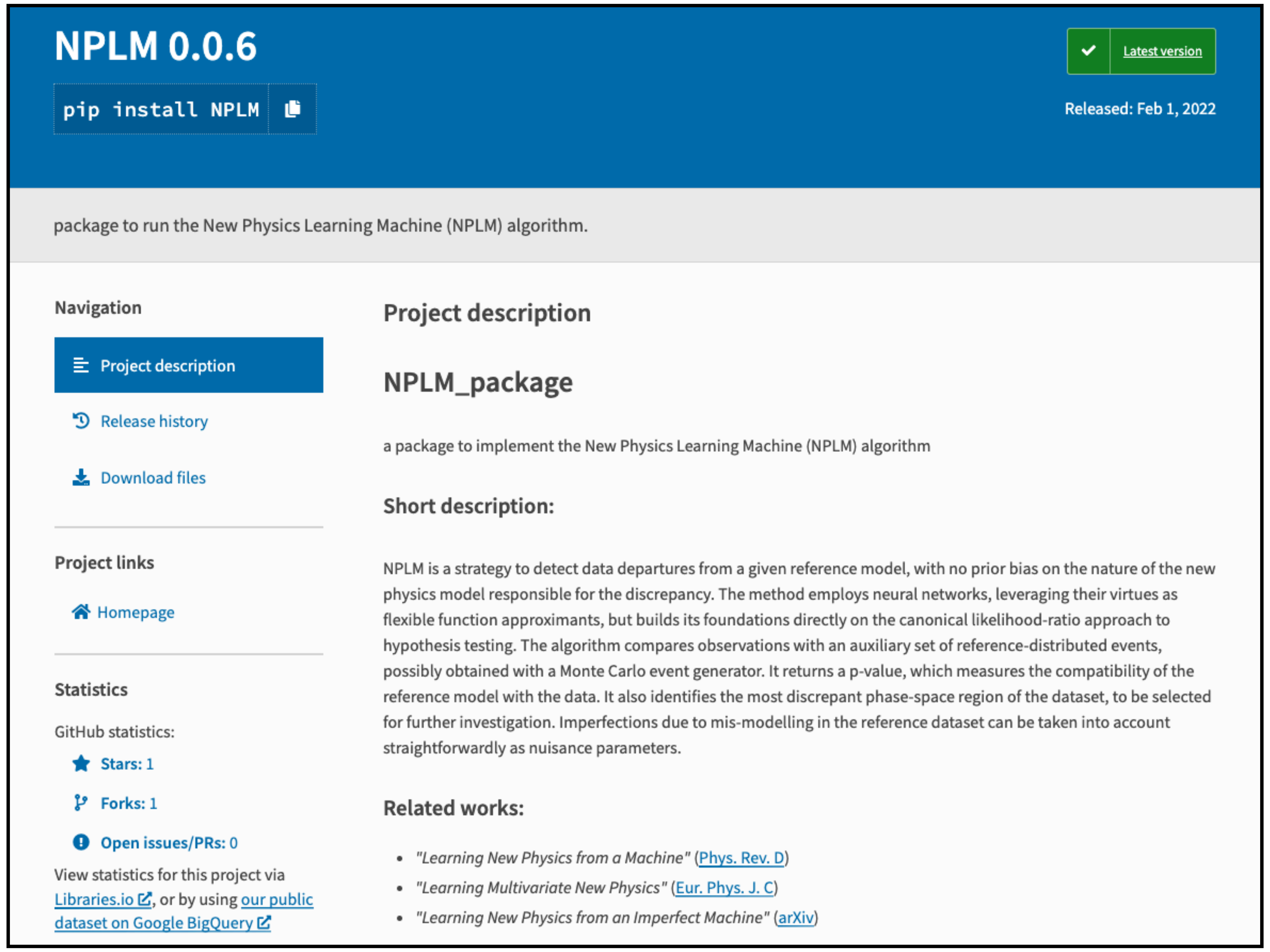


…to conclude

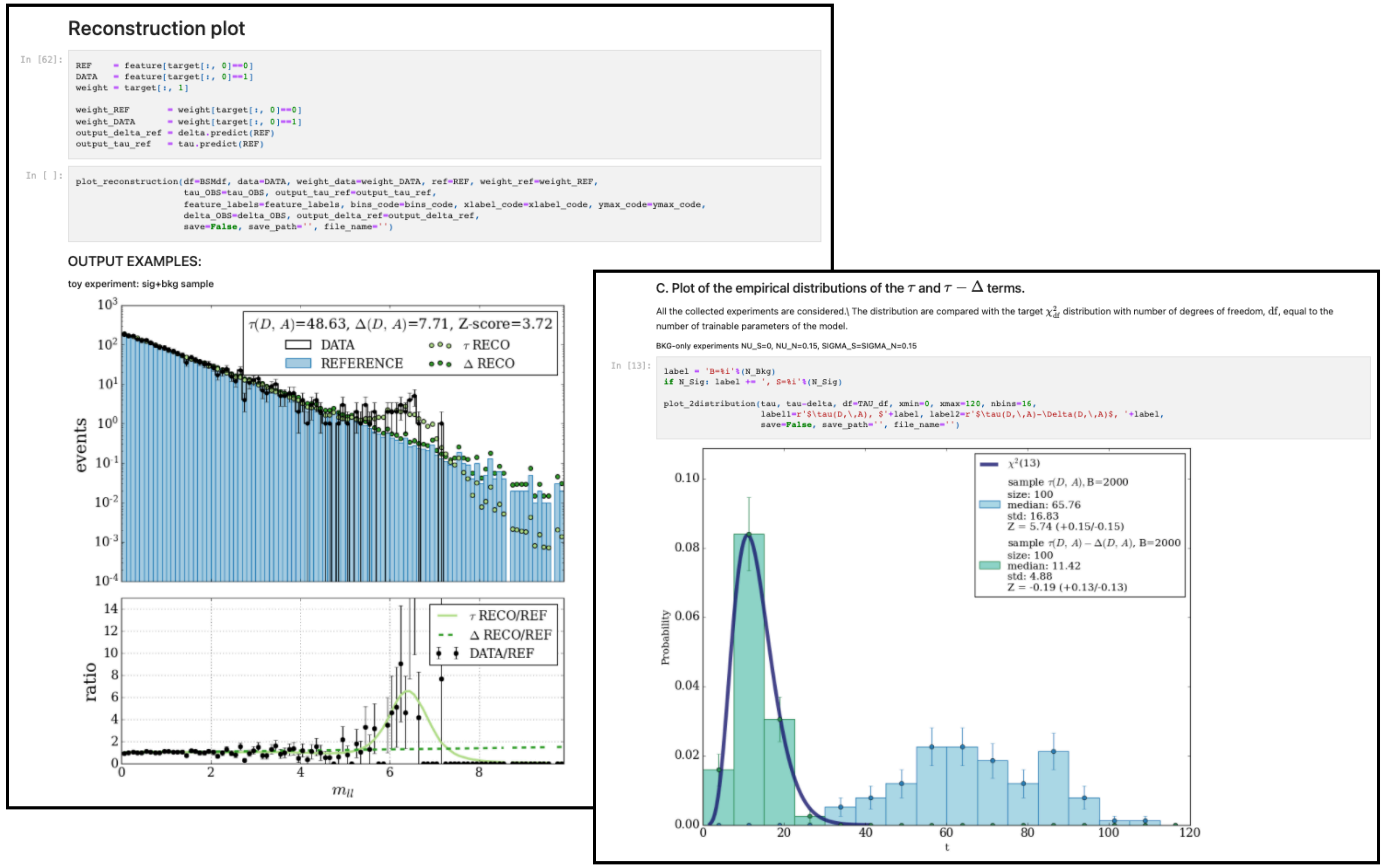
Code and resources

Getting started with NPLM

- NPLM package: python-based package to run the NPLM analysis strategy



- Tutorial on 1D toy model for getting started



Summary

Today's presentation:

- Main concepts and implementation using NN

  "Learning New Physics from a Machine" - d'Agnolo, Wulzer, Phys. Rev. D (2018)

  "Learning Multivariate New Physics" - d'Agnolo, Grosso, Pierini, Wulzer, Zanetti, Eur. Phys. J. C 81, 89 (2021)

- Systematic uncertainties

  "Learning New Physics from an Imperfect Machine" - d'Agnolo, Grosso, Pierini, Wulzer, Zanetti Eur. Phys. J. C 82, 275 (2022)

More about NPLM in a upcoming contribution from Marco Letizia:

- NPLM implementation using kernel methods, DQM application

  "Learning New Physics efficiently with non parametric methods" - Letizia, Grosso et al. Eur. Phys. J. C 82, 879 (2022)

  "Fast kernel methods for Data Quality Monitoring as a goodness-of-fit test" - Letizia, Grosso et al. 2303.05413 (preprint)

Outlook

Ongoing studies and future directions for NPLM

- NPLM as a goodness-of-fit tool: comparison with state of the art, pros and cons (contribution at neurips22 (ML for physical science workshop), paper in preparation)
- Exploring alternative methods of regularisation

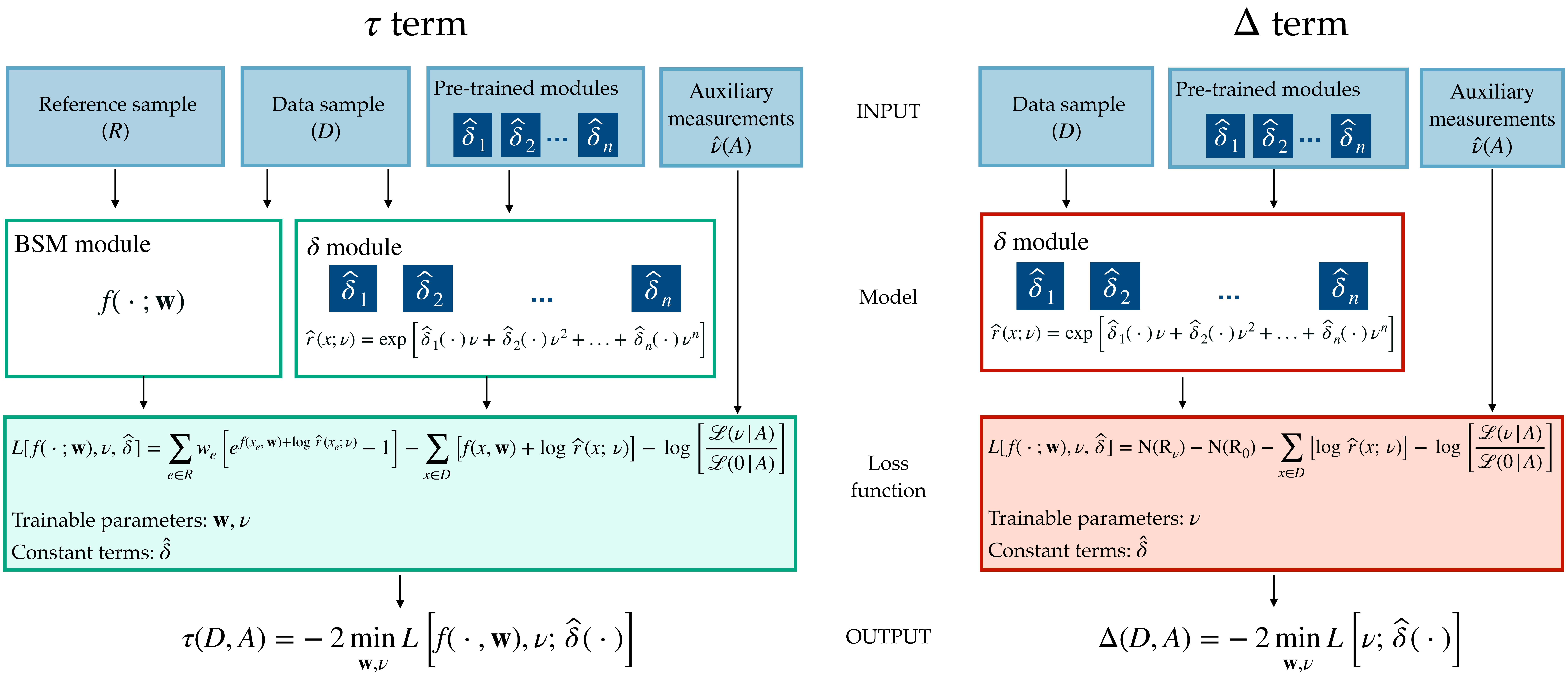
Current applications

- Model-independent search with collider data (Eur. Phys. J. C 82, 275 (2022) + ongoing work)
- Multi-dimensional DQM (preprint 2303.05413)
- Multidimensional validation of MC simulators (ongoing work)
- Bank fraud detection (ongoing work)
- Any idea? :)

Backup slides

New Physics Learning Machine (NPLM)

Including systematic uncertainties



New Physics Learning Machine (NPLM)

Including systematic uncertainties

$$t(\mathcal{D},\mathcal{A}) = \tau(\mathcal{D},\mathcal{A}) - \Delta(\mathcal{D},\mathcal{A})$$

τ term

$$\tau(\mathcal{D},\mathcal{A}) = 2\max_{\mathbf{w},\boldsymbol{\nu}} \log\left[\frac{\mathcal{L}(\mathrm{H}_{\mathbf{w},\boldsymbol{\nu}}|\mathcal{D})}{\mathcal{L}(\mathrm{R}_{\mathbf{0}}|\mathcal{D})}\cdot\frac{\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\mathcal{L}(\mathbf{0}|\mathcal{A})}\right]$$

$$\tau(\mathcal{D},\mathcal{A}) = -2\min_{\mathbf{w},\boldsymbol{\nu}}\left\{L_\tau\left[f(\cdot;\mathbf{w}),\ \boldsymbol{\nu};\ \widehat{\delta}(\cdot)\right]\right\}$$

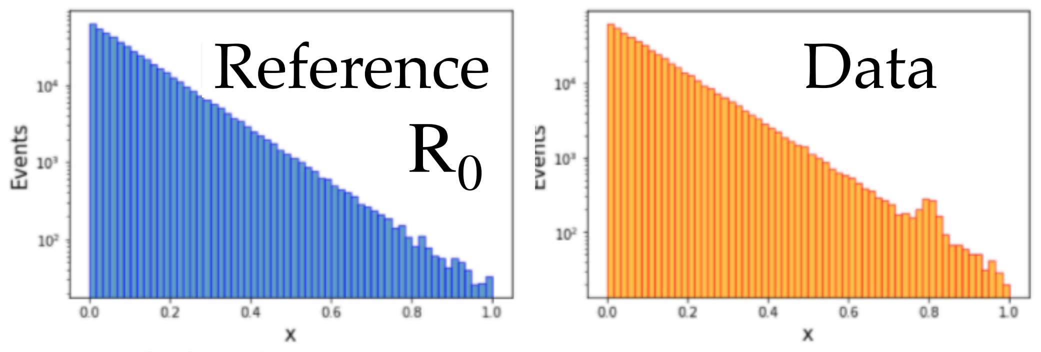
Δ term

$$\Delta(\mathcal{D},\mathcal{A}) = 2\max_{\boldsymbol{\nu}} \log\left[\frac{\mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{D})}{\mathcal{L}(\mathrm{R}_{\mathbf{0}}|\mathcal{D})}\cdot\frac{\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\mathcal{L}(\mathbf{0}|\mathcal{A})}\right]$$

$$\Delta(\mathcal{D},\mathcal{A}) = -2\min_{\boldsymbol{\nu}}\left\{L_\Delta\left[\boldsymbol{\nu};\ \widehat{\delta}(\cdot)\right]\right\}$$

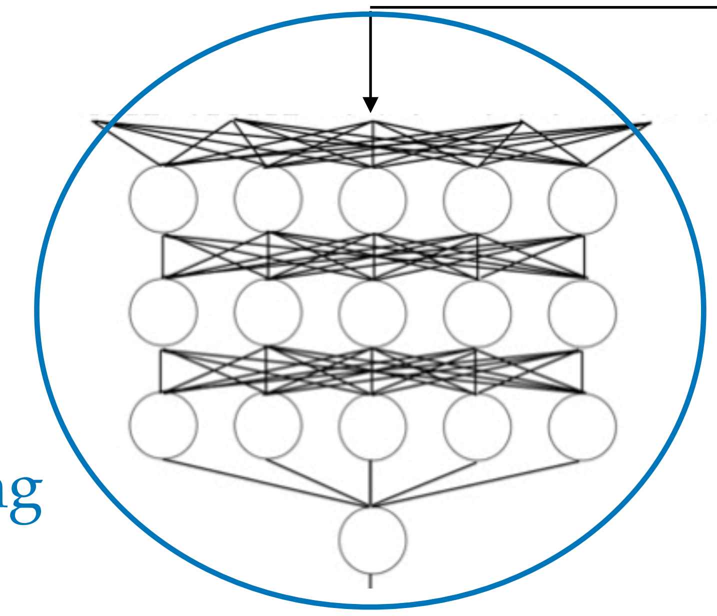
Maximum likelihood from minimal loss

Training samples



$f(\cdot;\mathbf{w})$

NN model
Set up
↓
Weight clipping tuning



$+$

$\boldsymbol{\nu};\ \widehat{\delta}(\cdot)$

$$\frac{n(x|\mathrm{R}_{\boldsymbol{\nu}})}{n(x|\mathrm{R}_{\mathbf{0}})}$$

$\boldsymbol{\nu};\ \widehat{\delta}(\cdot)$

$$\frac{n(x|\mathrm{R}_{\boldsymbol{\nu}})}{n(x|\mathrm{R}_{\mathbf{0}})}$$

r modelling

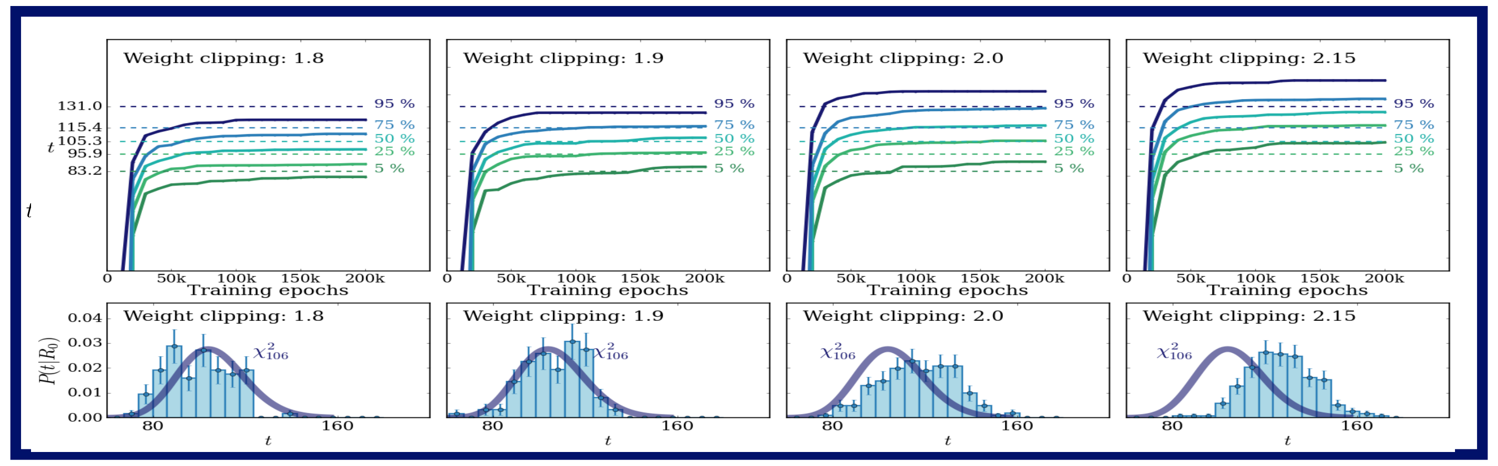
$$r(x;\boldsymbol{\nu}) \equiv \frac{n(x|\mathrm{R}_{\boldsymbol{\nu}})}{n(x|\mathrm{R}_{\mathbf{0}})}$$

New Physics Learning Machine (NPLM)

Including systematic uncertainties

Final procedure in steps:

1. **NN MODEL SELECTION**:
   weight clipping tuning → target $\chi^2_{|\mathbf{w}|}$;



2. **NUISANCE TAYLOR'S EXPANSION LEARNING**:
   modelling $\widehat{r}(x;\nu) = \exp\left[\widehat{\delta}_1(x)\,\nu + \widehat{\delta}_2(x)\,\nu^2 + \ldots\right]$;

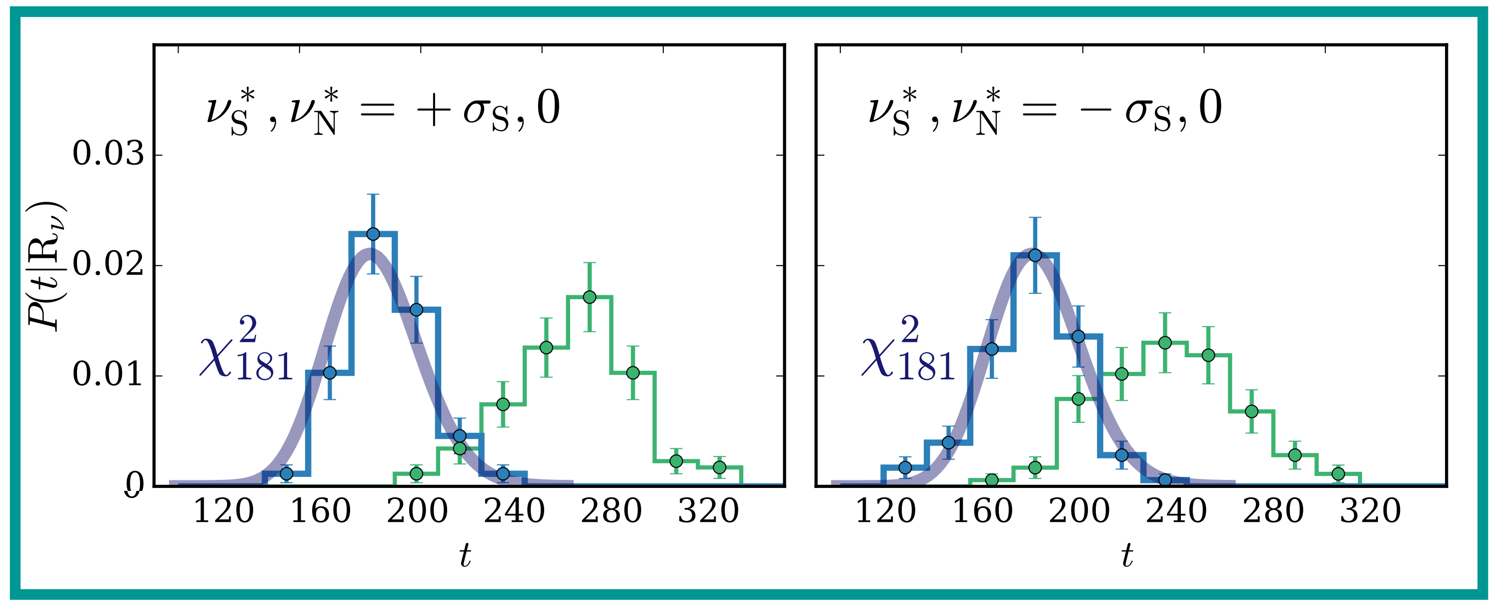
$$\widehat{r}(x;\nu) = \exp\left[\widehat{\delta}_1(x)\,\nu + \widehat{\delta}_2(x)\,\nu^2 + \ldots\right]$$

NN 1 NN2 ...

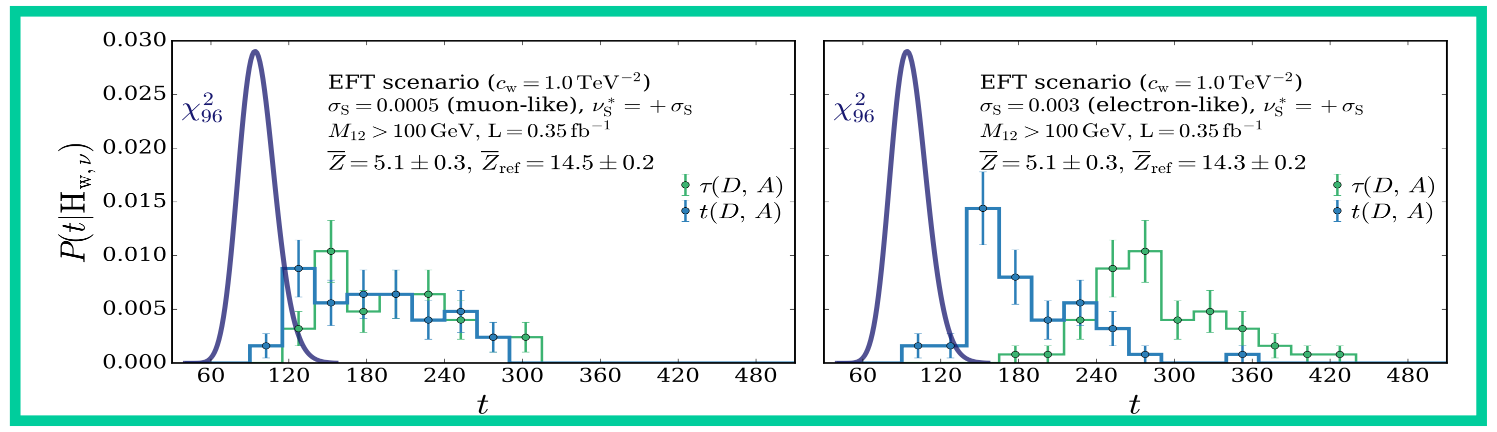
3. **VALIDATION**:

$$\mathcal{D} \sim \mathrm{R}_{\boldsymbol{\nu}^*}, \quad \boldsymbol{\nu}^* = \pm\sigma_{\boldsymbol{\nu}}$$

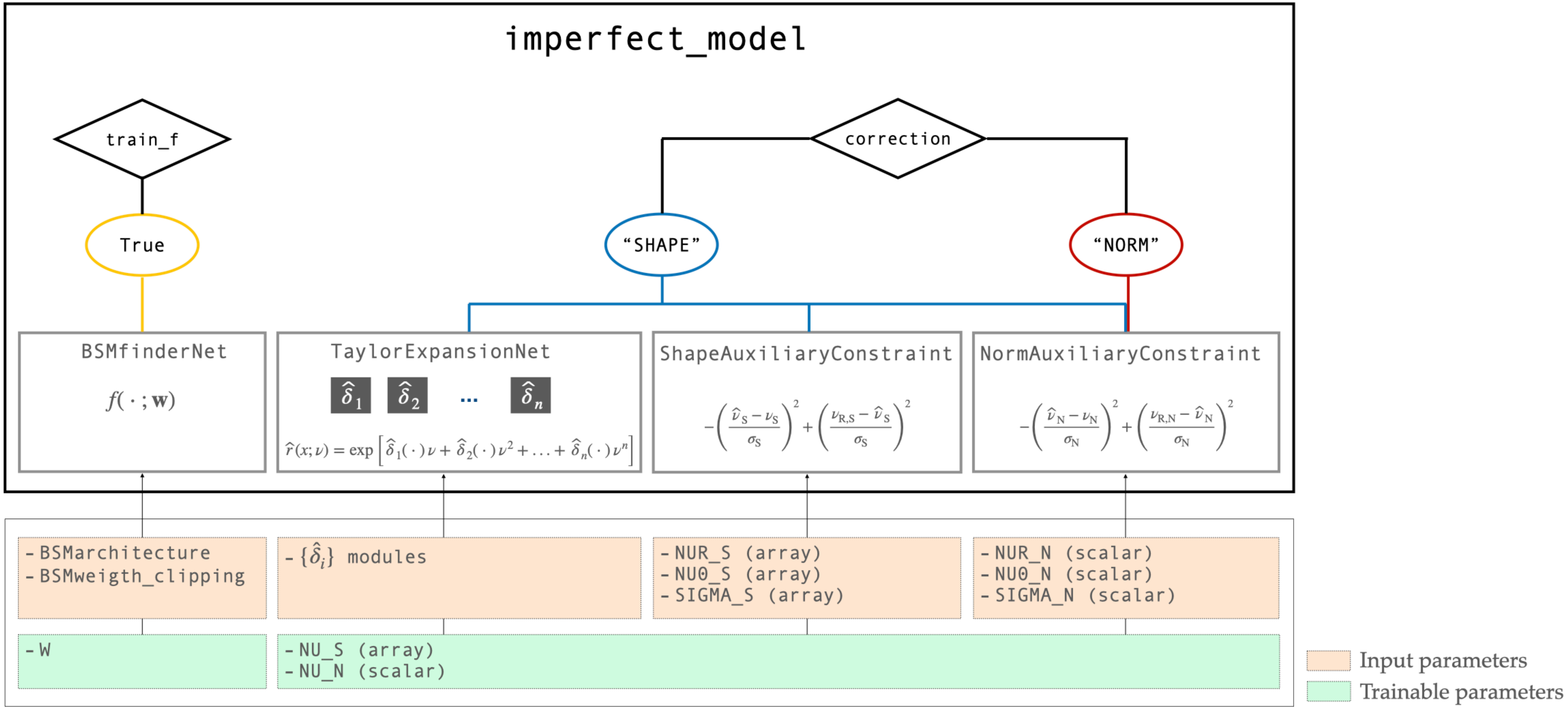
   Verifying that the target $\chi^2_{|\mathbf{w}|}$ is always recovered;



4. **TESTING THE DATA**:
   running the procedure on real data.



The model implementation in Keras + Tensorflow

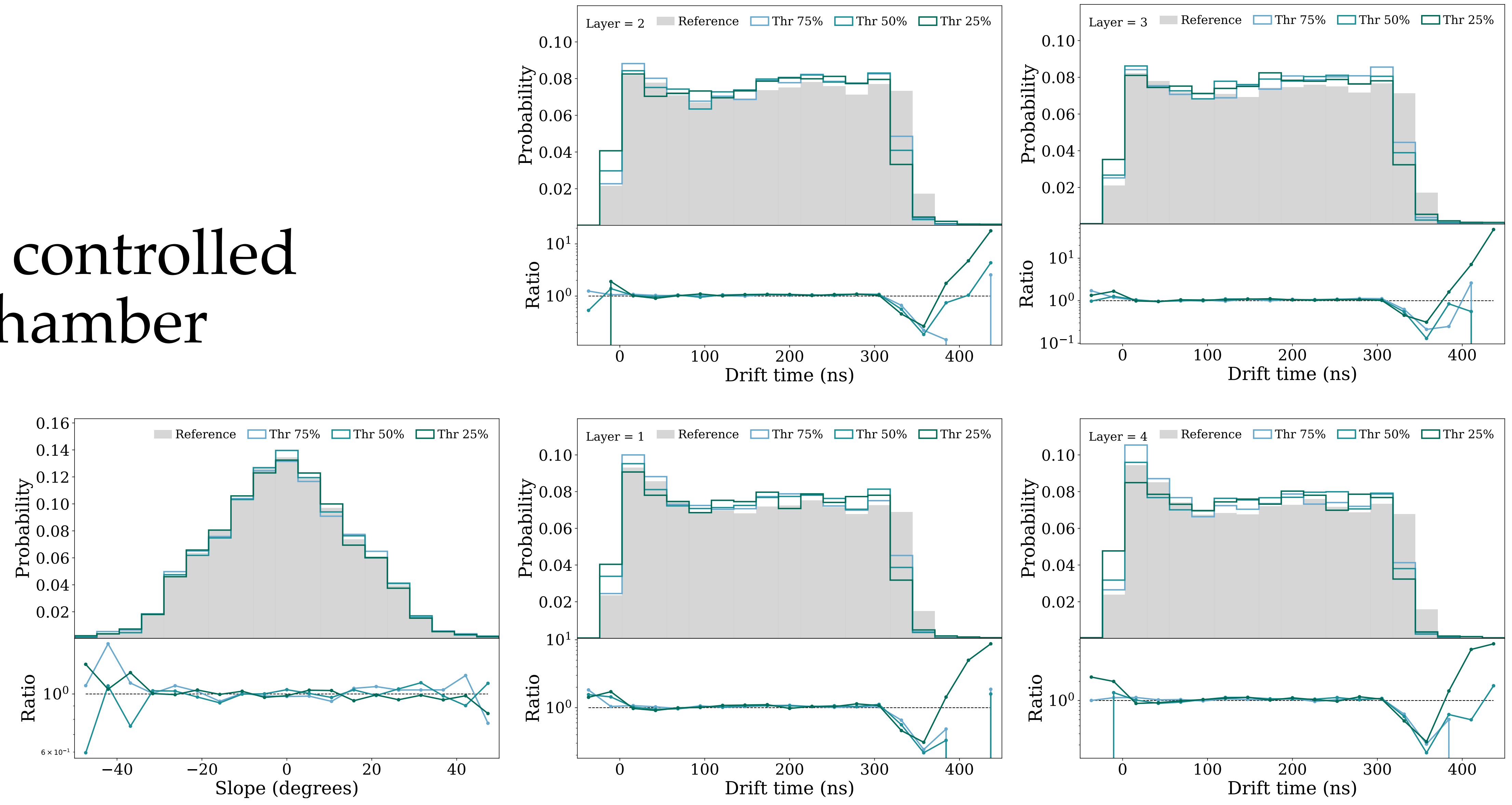


use case :
Validation of generator models
DQM

Example: *n*D DQM

Online monitoring of a DT chamber:

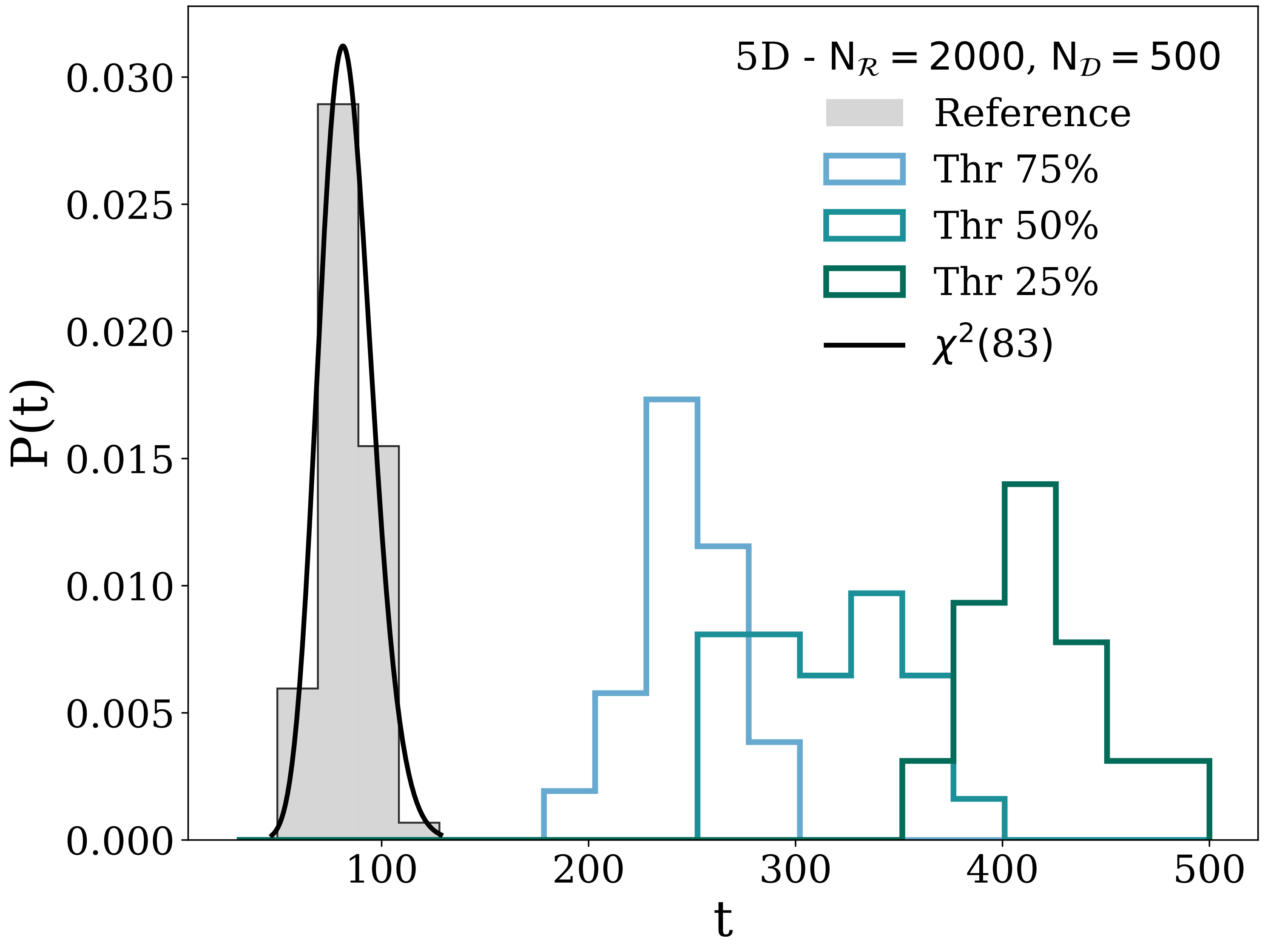
- **Reference sample:** long run in optimal conditions
- **Anomalous samples**: short runs acquired in presence of a controlled anomaly in the value of the **threshold tension** of the DT chamber



Distribution of the observables at different values of the threshold tension

- Result of the test statistics

  Complete separation of the distributions!



NPLM with kernel methods

$M = 50, \sigma = 4.84, \lambda = 10^{-7}$

$N(D) = 5000$

$N_{\text{ref}} = 200\,000$

Execution time: ~ 1.5 s

NEW: arXiv:2303.05413

→ more about this in an upcoming follow up talk!

In collaboration with UniGe | MaLGa

use case :
nD New Physics search

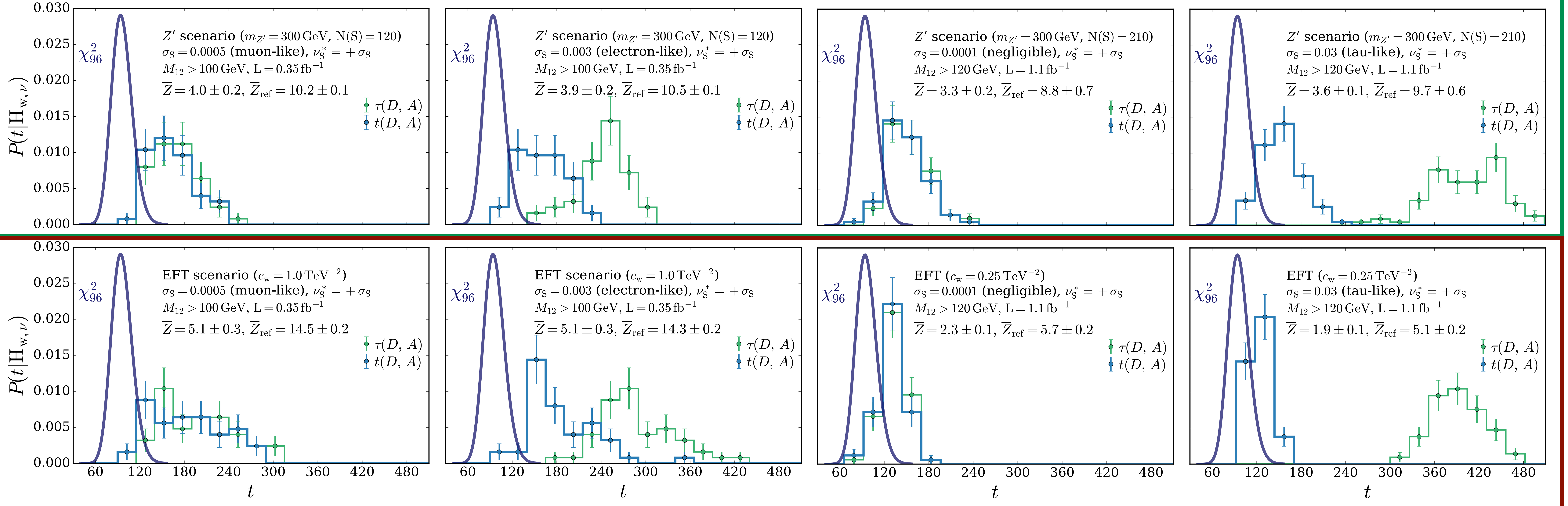
Harder task: nD analysis

Two body final state (5D): sensitivity to NP scenarios

Muon-like regime | Electron-like regime | Negligible systematic uncertainties | Tau-like regime

Z' scenario — $\nu_S^* = +\sigma_S$

EFT scenario — $\nu_S^* = +\sigma_S$



Z-score: $Z = \Phi^{-1}\left[1 - p\right]$

$\overline{Z}$: Z-score from the median of the empirical $t(D, A)$ distribution

$t(D,\ A) = \tau(D,\ A) - \Delta(D,\ A)$

$\tau(D,\ A)$

Harder task: *n*D analysis

Two bodies final state (5D)

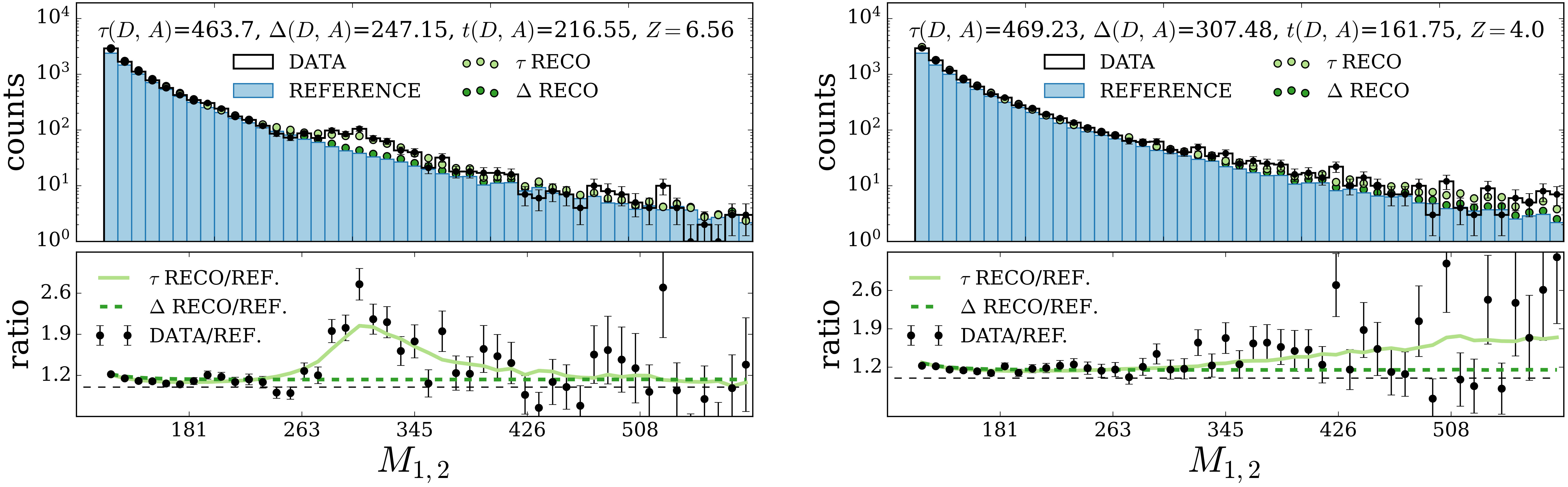
Signal reconstruction with the NN:

Architecture: [5-5-5-5–1] (96 dof), weigh clipping 2.15, $L = 240\ \mathrm{fb}^{-1}$

NOTE:
M_{12} is **not** given as an input to the algorithm!

τ reconstruction: $n(x\,|\,\mathrm{H}_{\hat{\mathbf{w}},\hat{\nu}}) = n(x\,|\,\mathrm{R}_0)\,\dfrac{n(x\,|\,\mathrm{R}_{\hat{\nu}})}{n(x\,|\,\mathrm{R}_0)}\,e^{f(x;\,\hat{\mathbf{w}})}$

Δ reconstruction: $n(x\,|\,\mathrm{R}_{\hat{\nu}})$

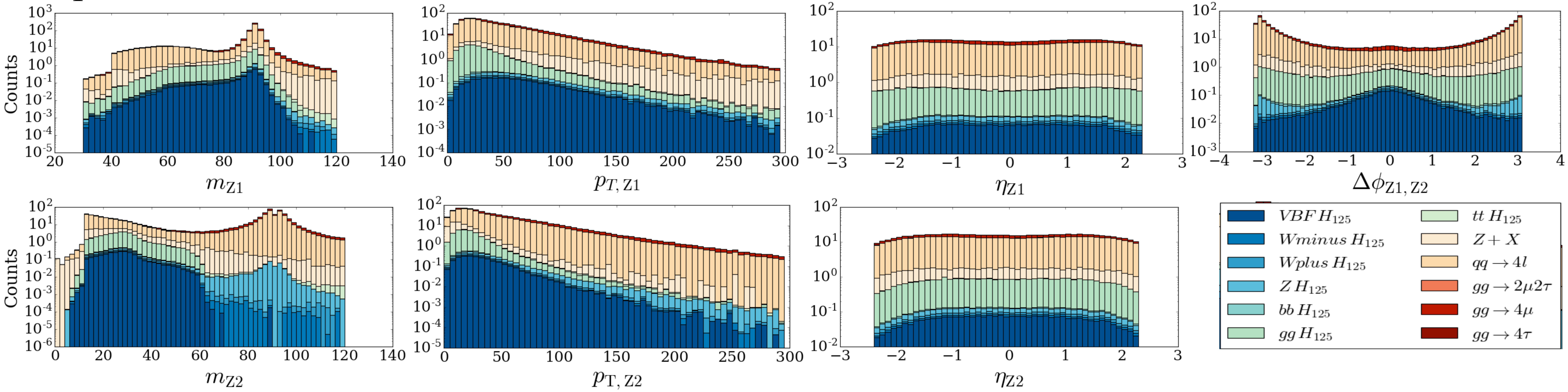


"Learning Multivariate New Physics" Eur. Phys. J. C

Harder task: nD analysis

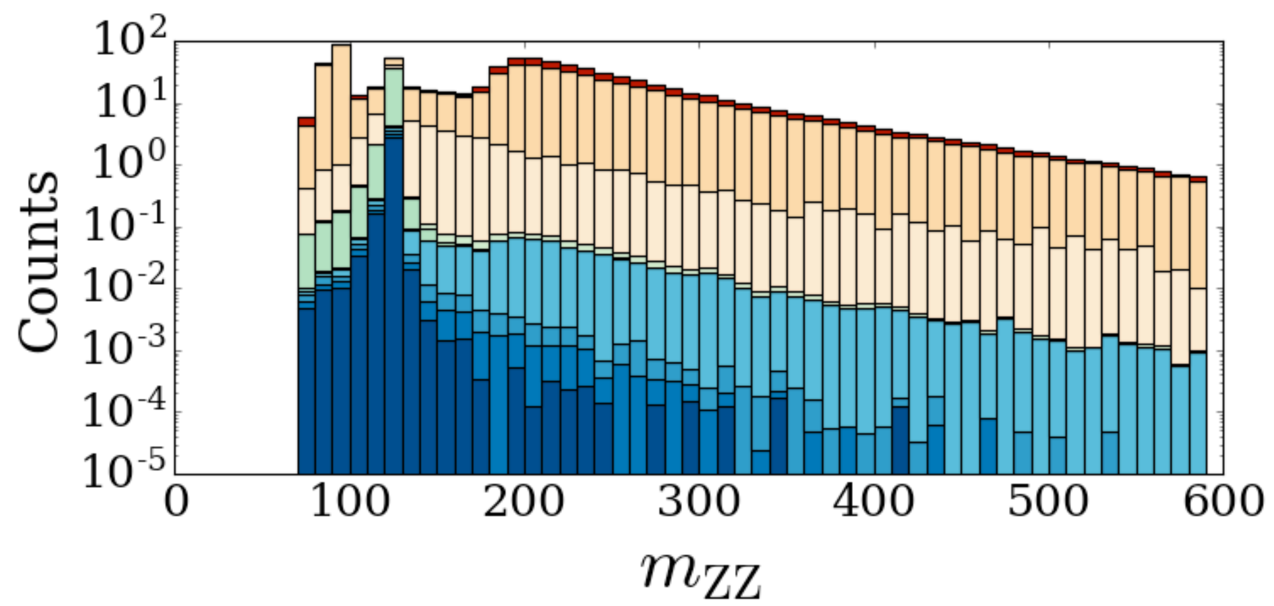
ZZ to 4 muons final state (7D)

Input variables:



Signal benchmarks:

- Higgs boson (as an exercise)



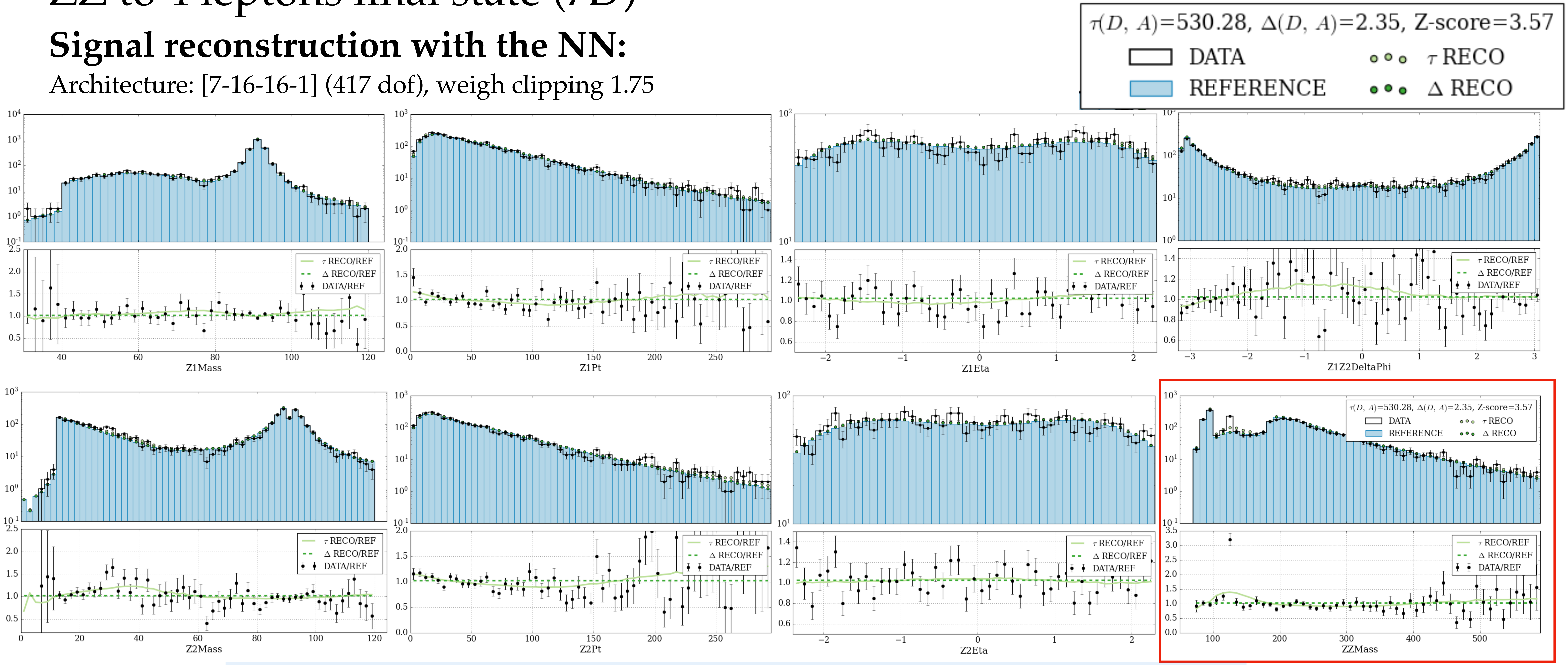
NOTE:
m_{ZZ} is **not** given as an input to the algorithm!

nD analysis

ZZ to 4 leptons final state (7D)

Signal reconstruction with the NN:

Architecture: [7-16-16-1] (417 dof), weigh clipping 1.75



Ongoing work for optmizing NPLM sensitivity in high dimensional problems