Banff International Research Station
for Mathematical Innovation and Discovery

Systematic Effects and Nuisance Parameters
in Particle Physics Data Analyses
April, 23-28 2023

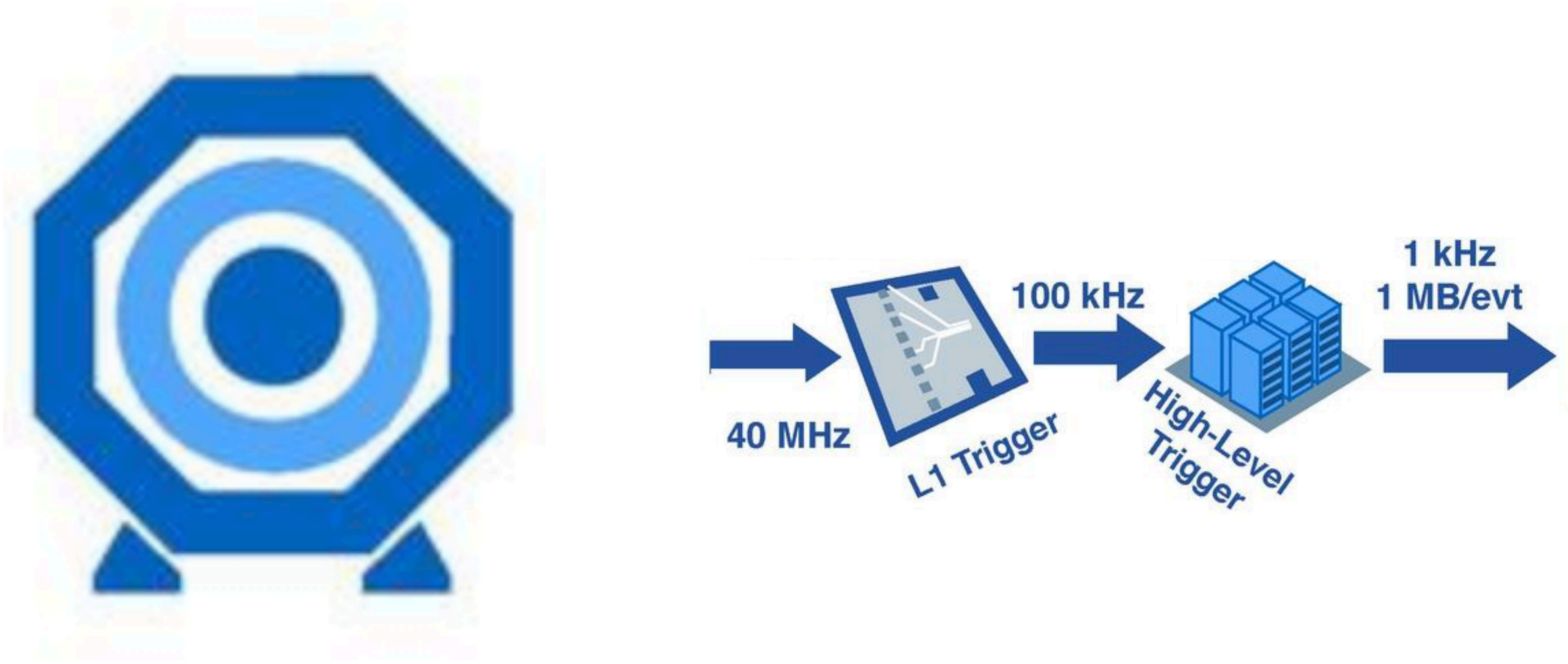
Systematics in model-independent ML-based searches

Gaia Grosso
CERN, University and INFN of Padova

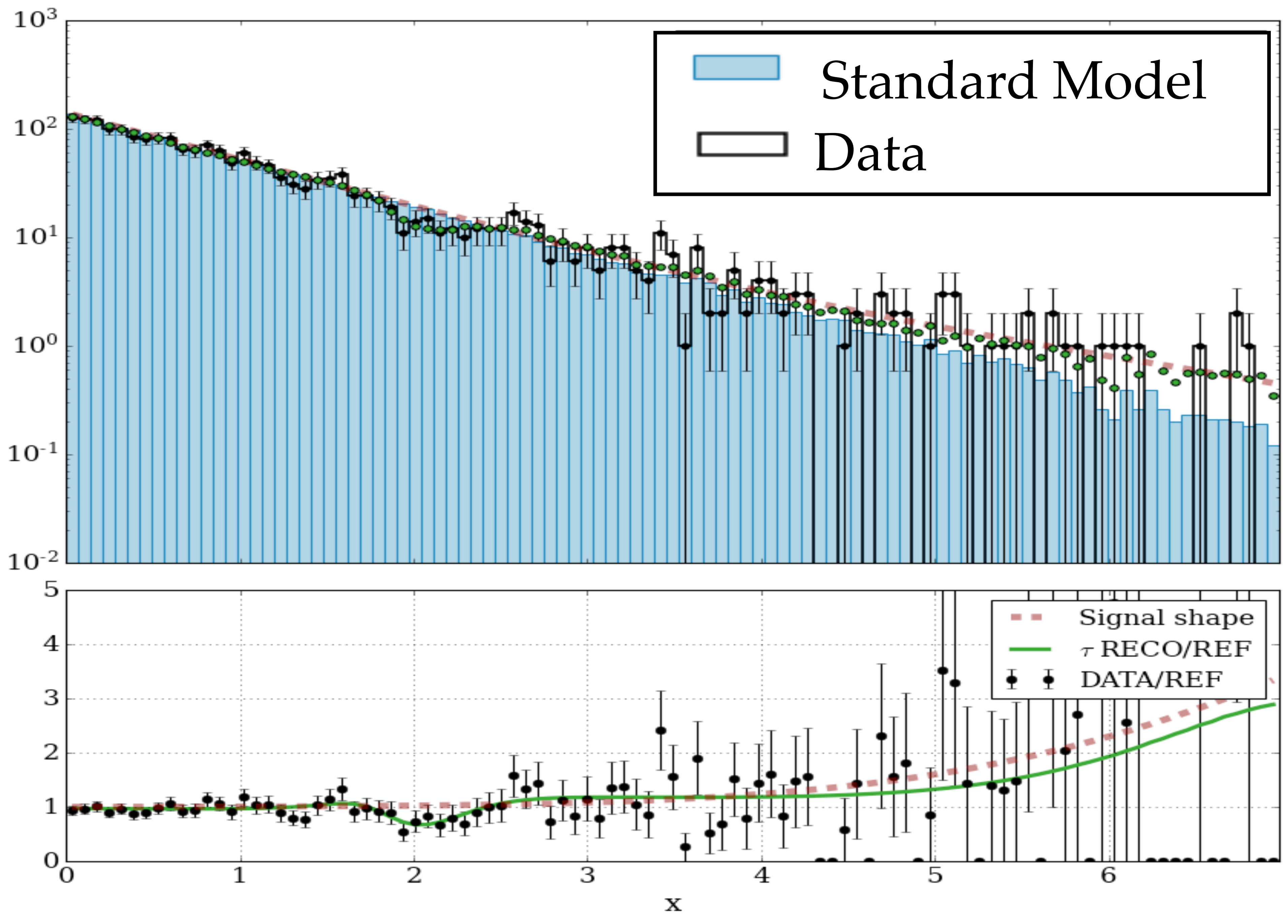
Mainly based on:
Phys. Rev. D 99, 015014 (d'Agnolo, Wulzer)
Eur. Phys. J. C 81, 89 (2021) (d'Agnolo, Grosso, Pierini, Wulzer, Zanetti)
Eur. Phys. J. C 82, 275 (2022) (d'Agnolo, Grosso, Pierini, Wulzer, Zanetti)

How good is the SM description of the data?

experiment



Summary statistics



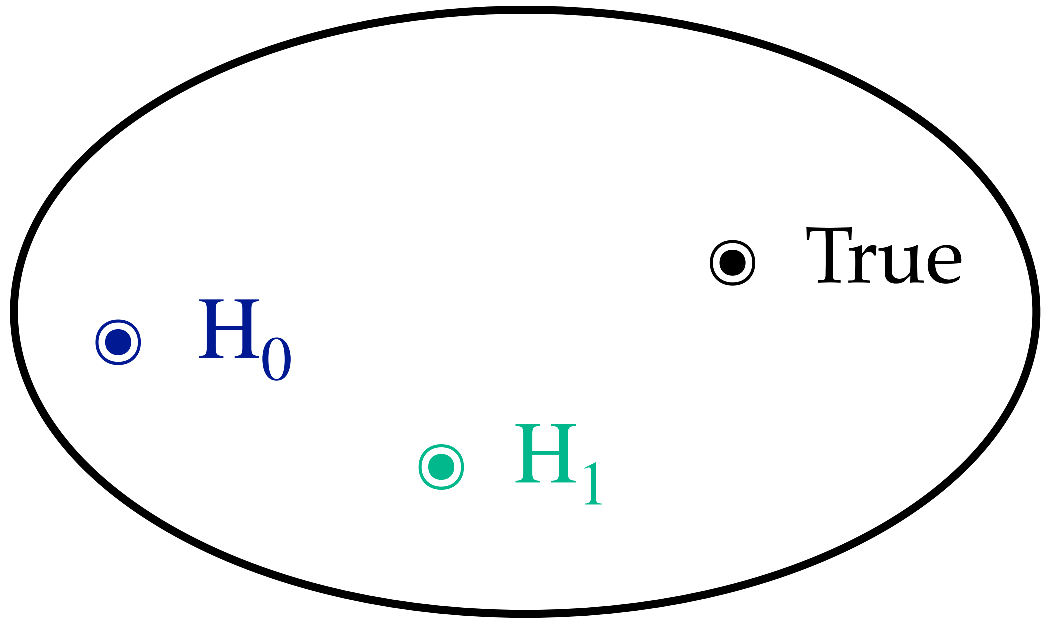
This is a problem of Goodness of Fit of the SM.

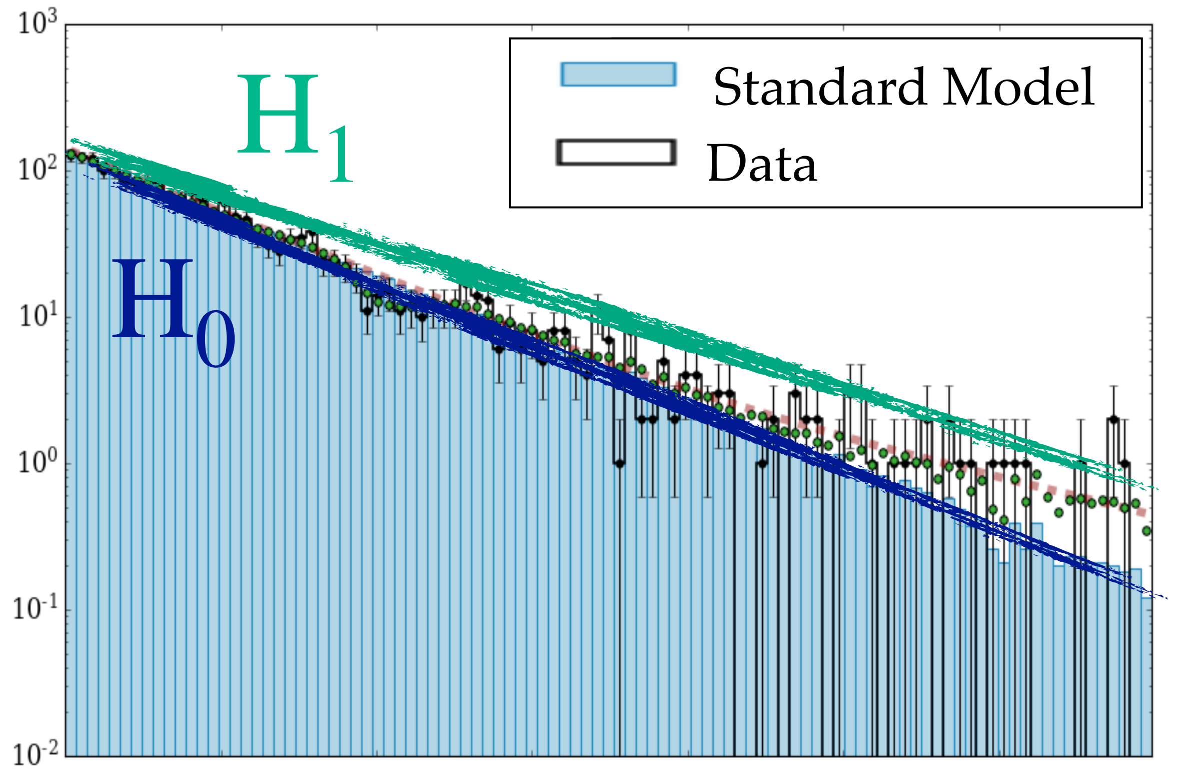
(The need for a "Reference" sample makes it a 2-samples test.)

Likelihood-ratio test

traditionally solved in HEP as a hypothesis test based on likelihoods ratio

→ Model-dependent approaches





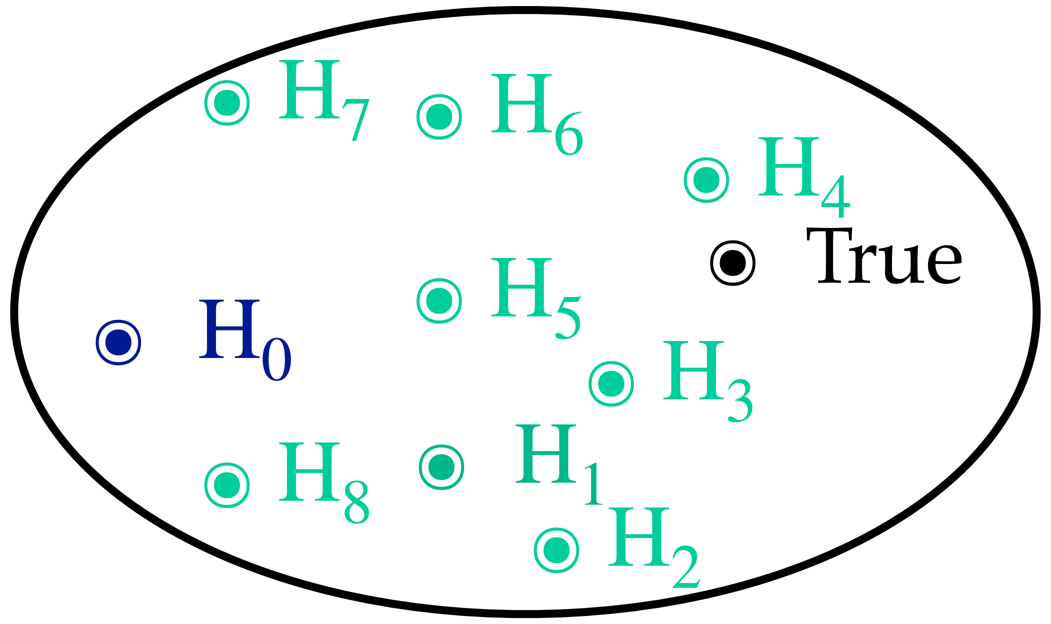
$$t(\mathcal{D}) = 2\log\frac{\mathcal{L}(\mathcal{D}\,|\,\mathrm{H}_1)}{\mathcal{L}(\mathcal{D}\,|\,\mathrm{H}_0)}$$

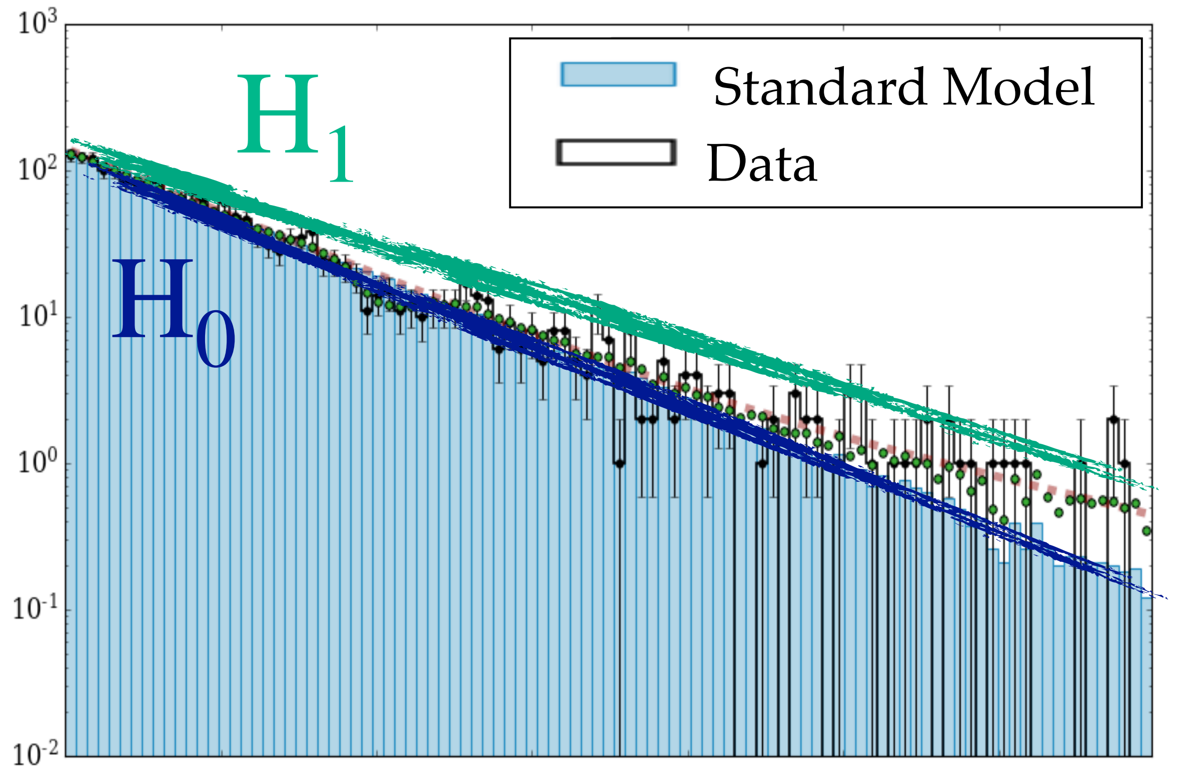
Sensitivity (and optimality) guaranteed (according to Neyman and Pearson) only if the the data do follow the chosen alternative

Likelihood-ratio test

traditionally solved in HEP as a hypothesis test based on likelihoods ratio

→ Model-dependent approaches



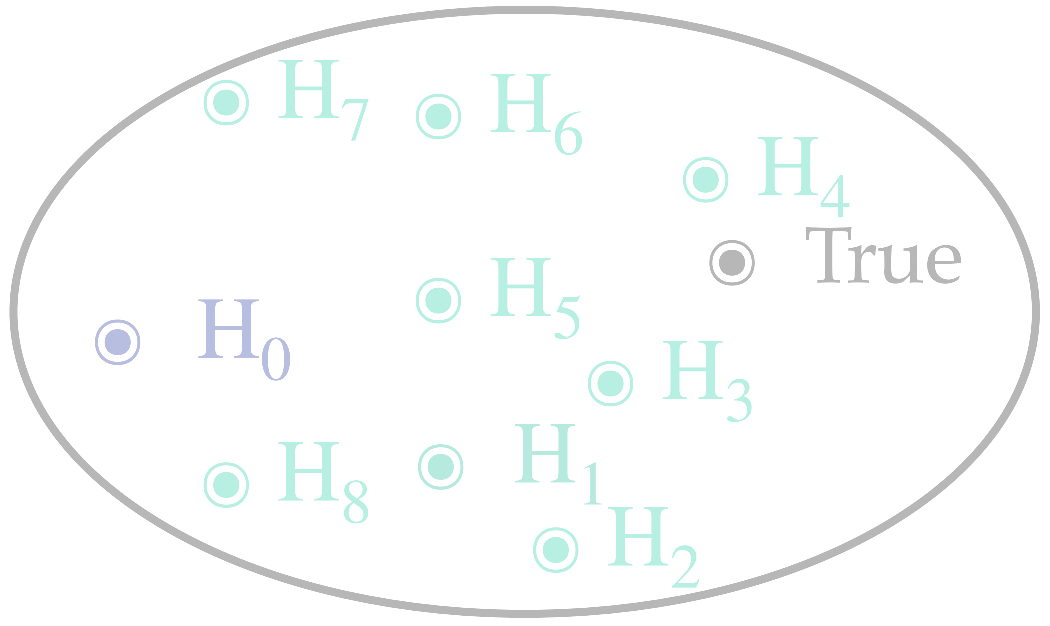


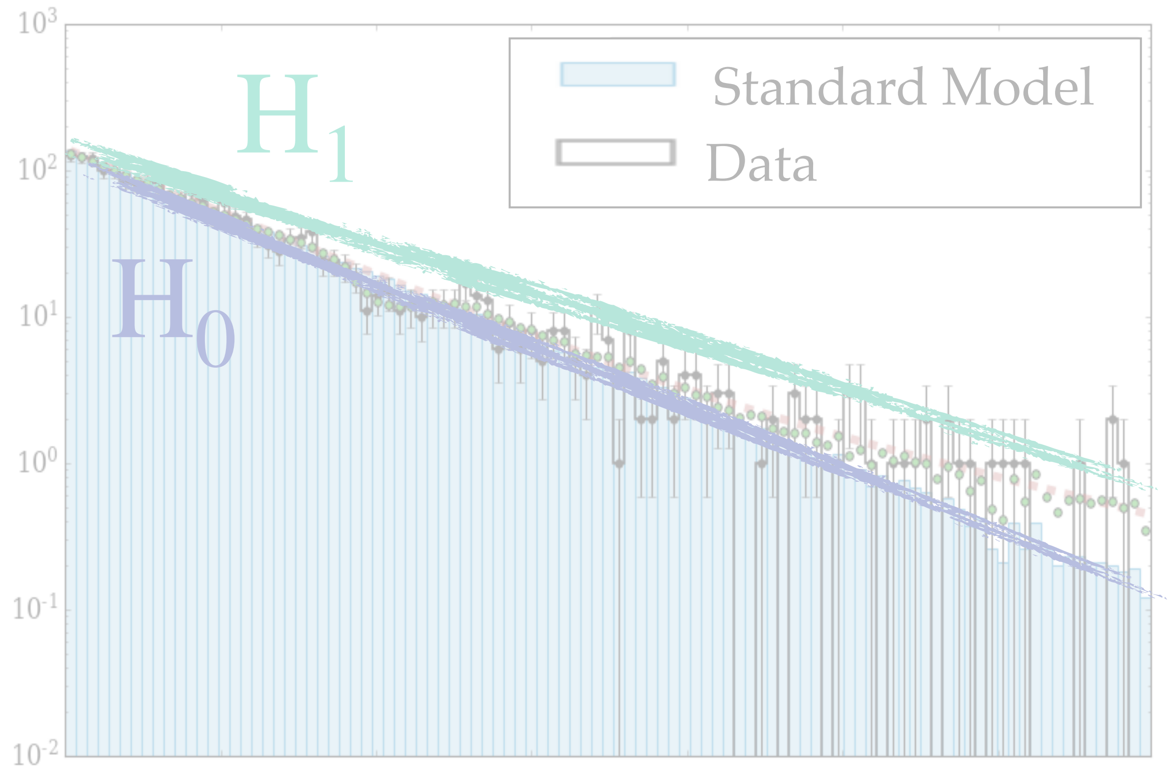
$$t(\mathcal{D}) = 2 \log \frac{\mathcal{L}(\mathcal{D}\,|\mathrm{H}_1)}{\mathcal{L}(\mathcal{D}\,|\mathrm{H}_0)}$$

Sensitivity (and optimality) guaranteed (according to Neyman and Pearson) only if the the data do follow the chosen alternative

Likelihood-ratio test

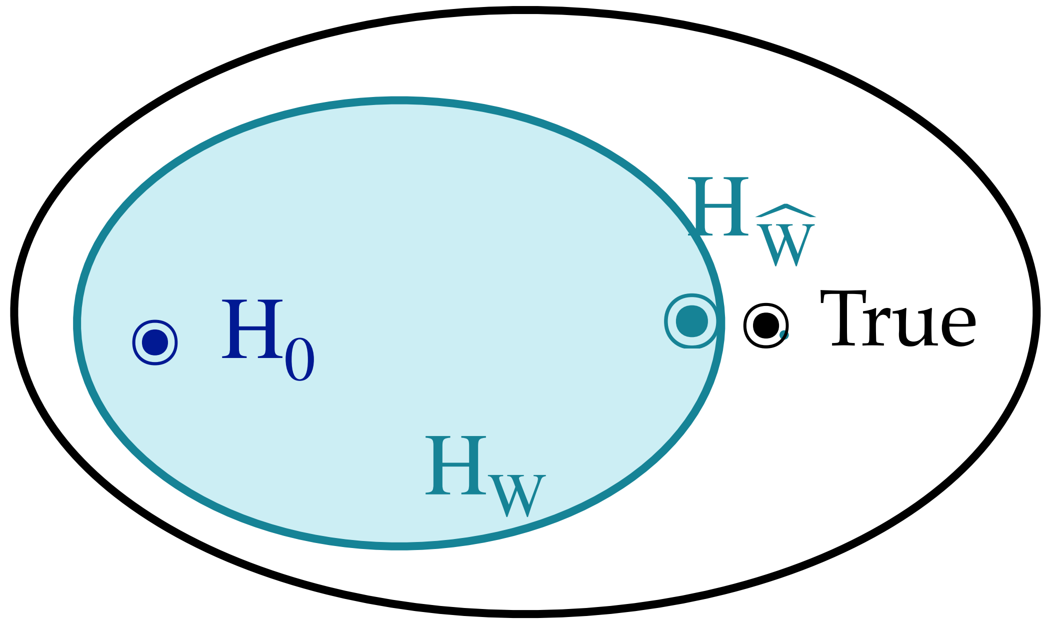
→ Model-dependent approaches

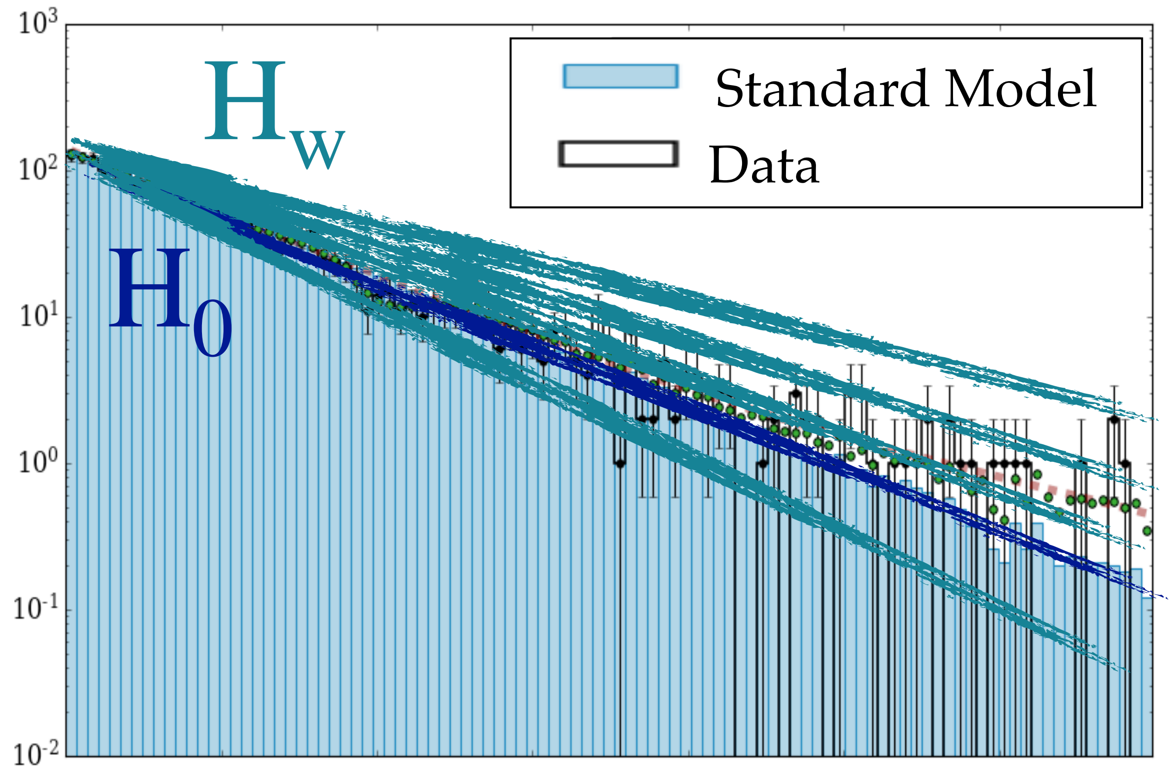




$$t(\mathcal{D}) = 2 \log \frac{\mathcal{L}(\mathcal{D} \,|\, \mathrm{H}_1)}{\mathcal{L}(\mathcal{D} \,|\, \mathrm{H}_0)}$$

→ Model-*in*dependent approaches

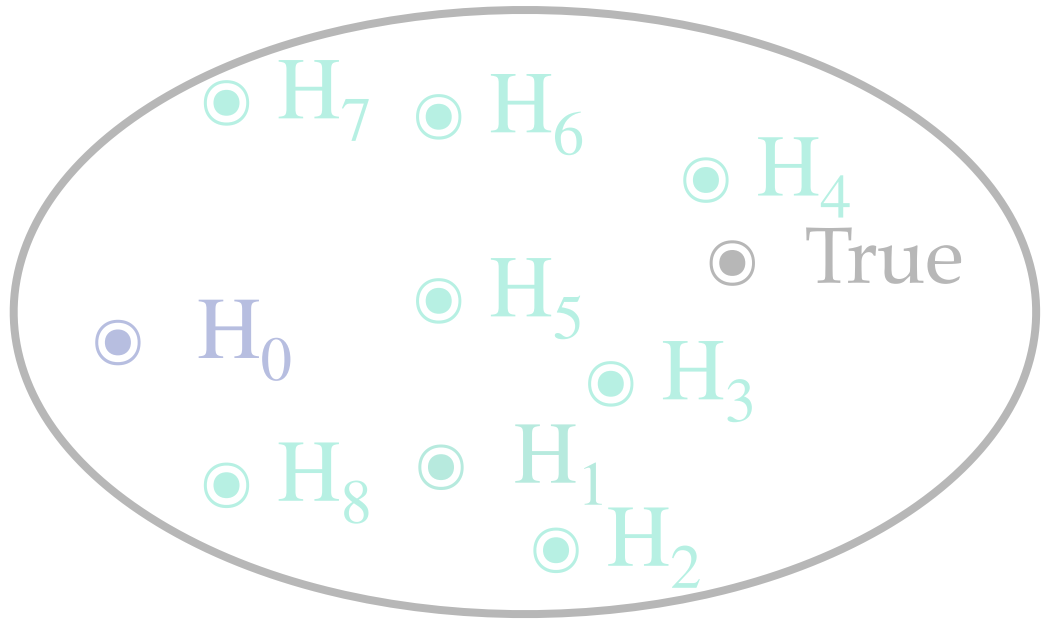


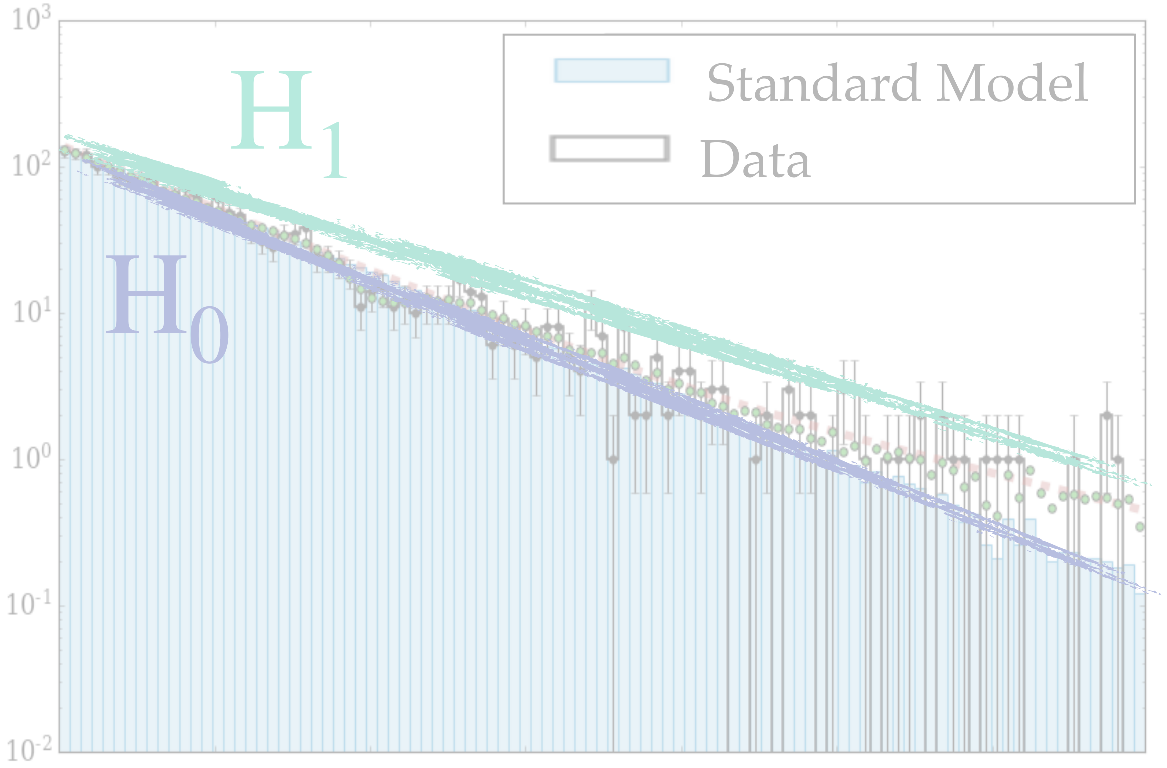


$$t(\mathcal{D}) = \max_{\mathbf{w}} \left[2 \log \frac{\mathcal{L}(\mathcal{D} \,|\, \mathrm{H}_{\mathbf{w}})}{\mathcal{L}(\mathcal{D} \,|\, \mathrm{H}_{\mathbf{0}})} \right]$$

Likelihood-ratio test

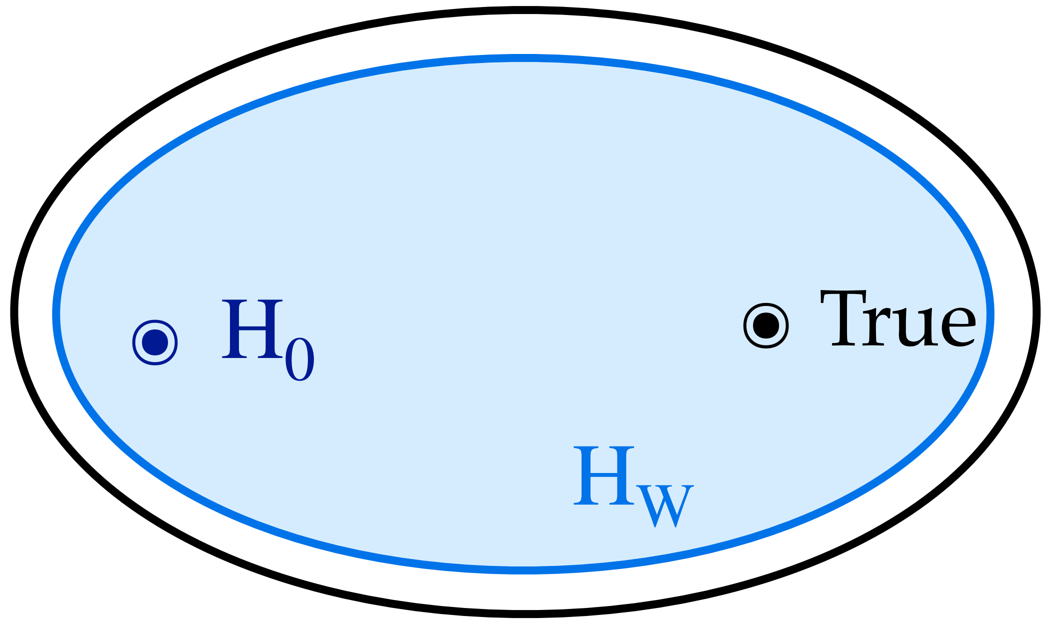
→ Model-dependent approaches

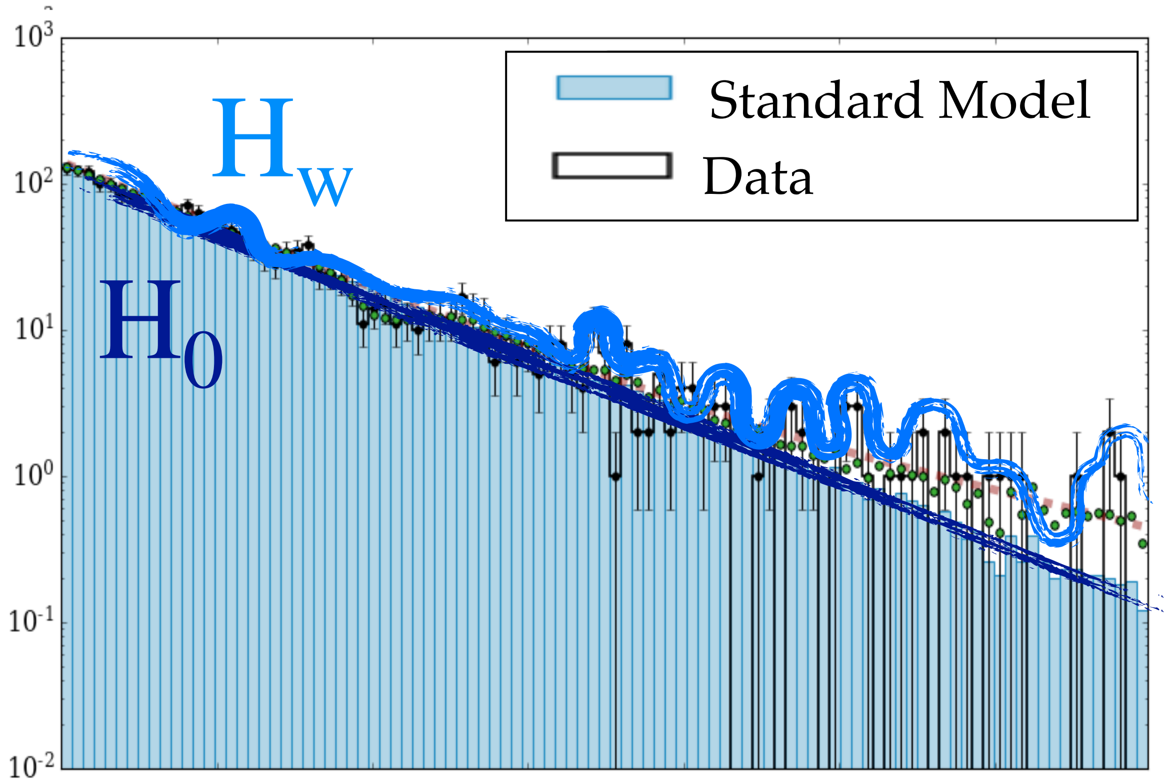




$$t(\mathcal{D}) = 2\ \log \frac{\mathcal{L}(\mathcal{D}\,|\mathrm{H}_1)}{\mathcal{L}(\mathcal{D}\,|\mathrm{H}_0)}$$

→ Model-*in*dependent approaches





$$t(\mathcal{D}) = \max_{\mathbf{w}} \left[2\ \log \frac{\mathcal{L}(\mathcal{D}\,|\mathrm{H}_{\mathbf{w}})}{\mathcal{L}(\mathcal{D}\,|\mathrm{H}_0)} \right]$$

Machine learning based

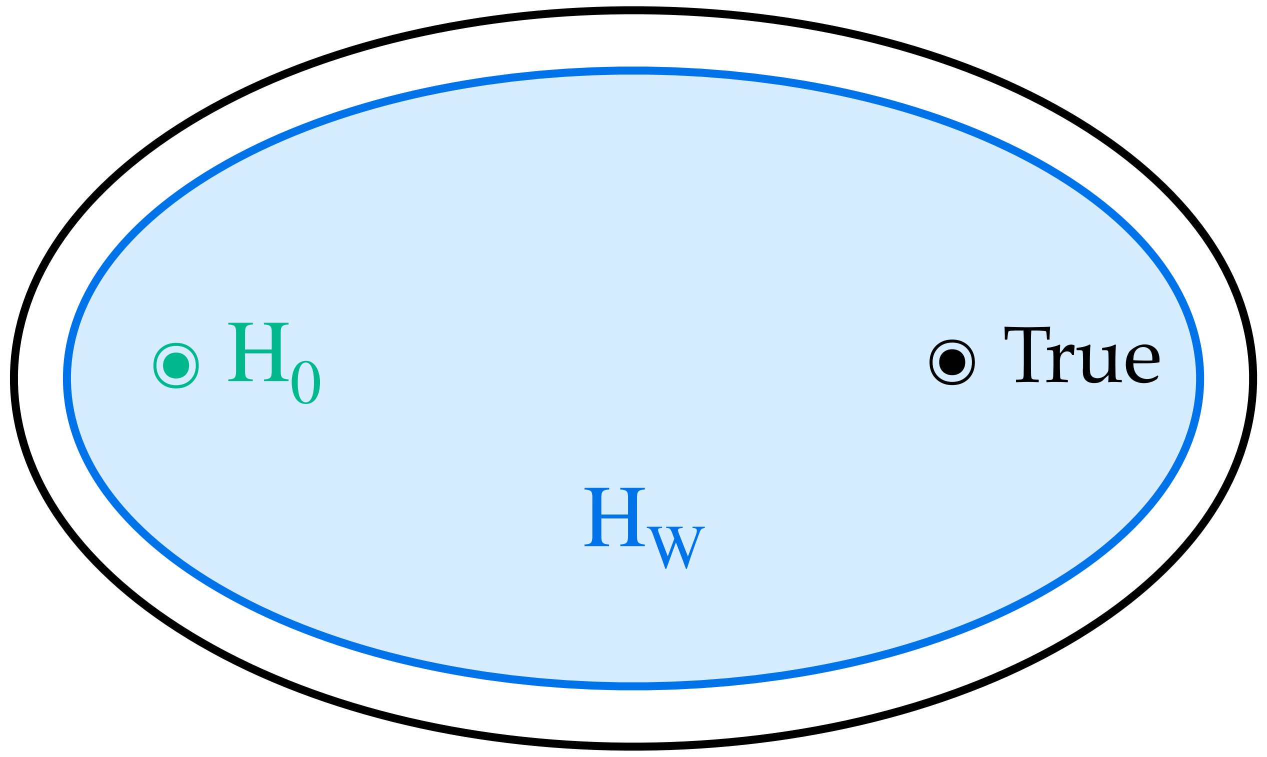
Model-independent approaches
New Physics Learning Machine (NPLM)

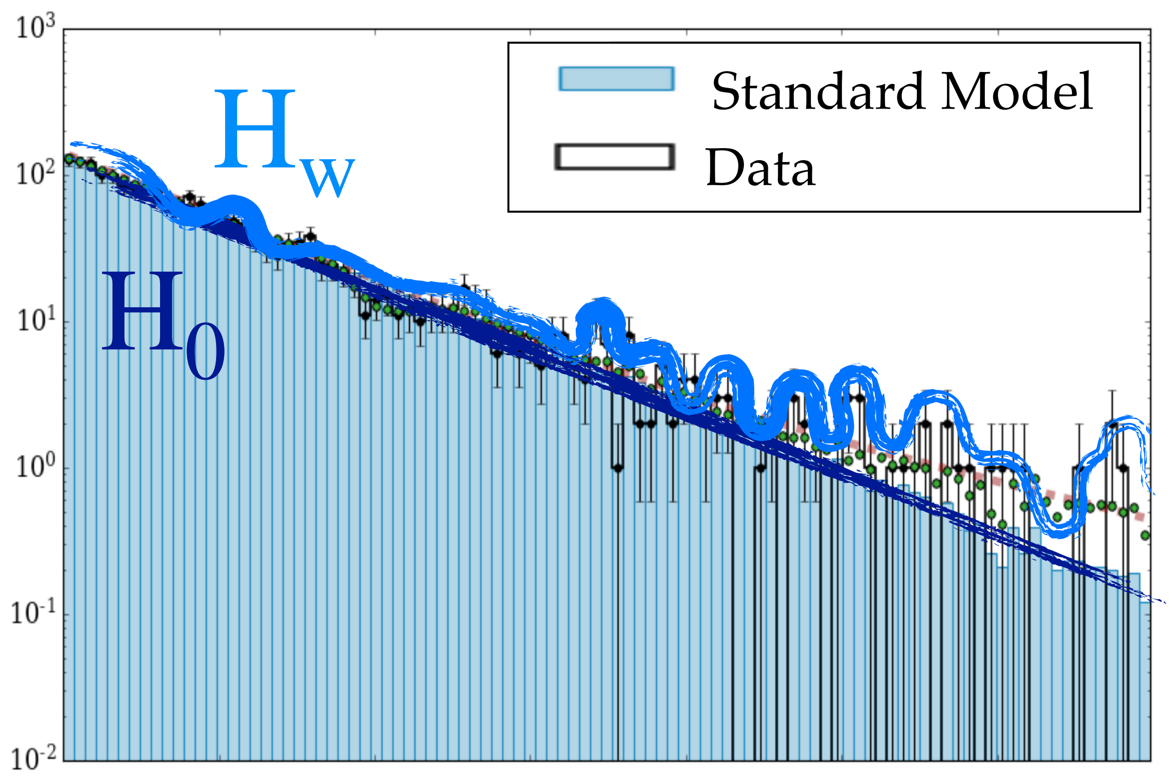
Universal approximator (NN, kernel methods, …)

Expand the family of alternatives to increase the chance of containing the True data distribution

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

"Statistical" anomaly detector





H_0: null hypothesis (SM)
$H_{\mathbf{w}}$: alternative hypothesis (NP)

New Physics Learning Machine (NPLM)

Maximum-Likelihood-ratio test statistic:

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

- $\mathrm{R}_{\boldsymbol{\nu}}$: reference hypothesis (null)
- $\mathrm{H}_{\mathbf{w},\,\boldsymbol{\nu}}$: alternative hypothesis
- $\mathbf{w}$: 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 $\boldsymbol{\nu}$)

Parametrisation of the alternative hypothesis:

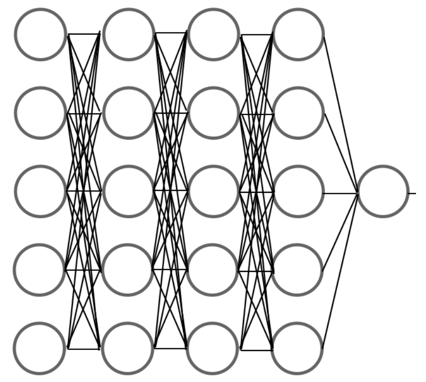
$$n(x\,|\,\mathrm{H}_{\hat{\mathbf{w}},\,\hat{\nu}}) = n(x\,|\,\mathrm{R}_0)\,\frac{n(x\,|\,\mathrm{R}_{\hat{\nu}})}{n(x\,|\,\mathrm{R}_0)}\,e^{f(x;\,\hat{\mathbf{w}})}$$

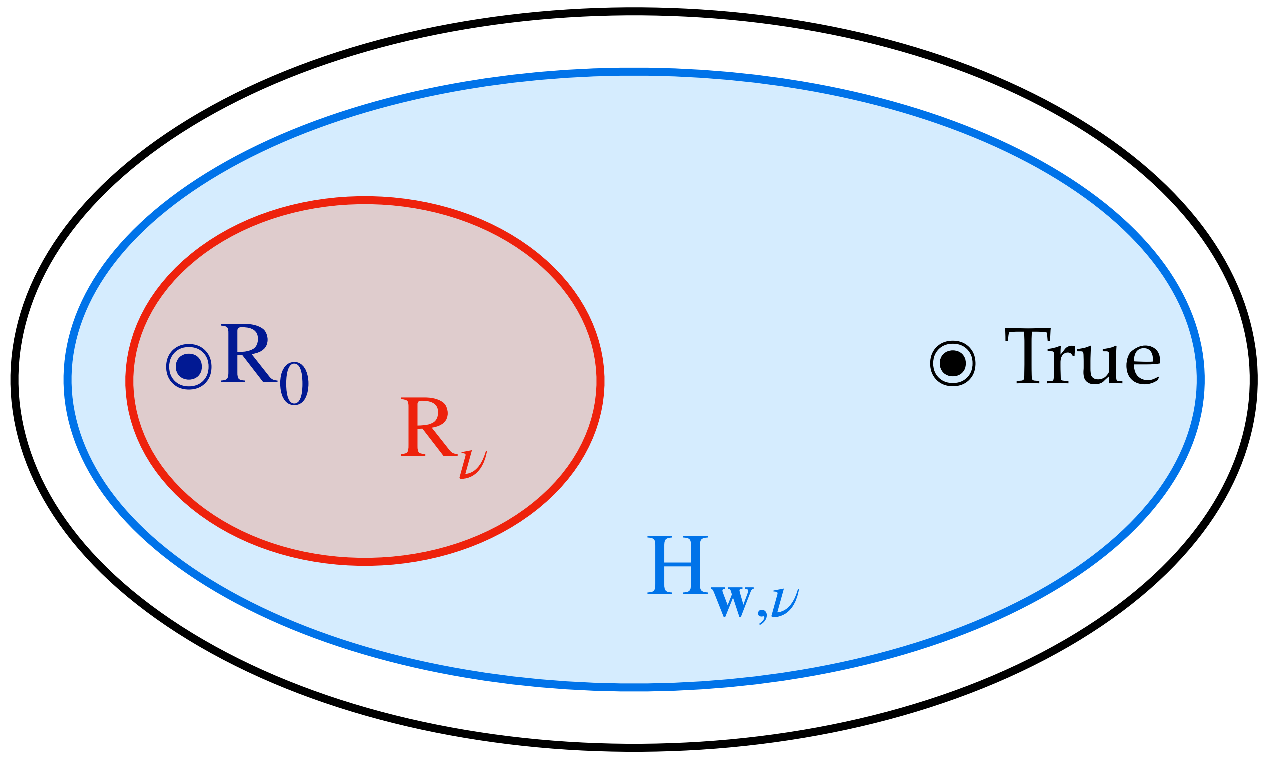
central value SM (ν=0)

Model of systematic uncertainties effects

$$\hat{r}(x;\nu) = \frac{n(x\,|\,\mathrm{R}_\nu)}{n(x\,|\,\mathrm{R}_0)} = \exp\left[\hat{\delta}_1(x)\,\nu + \hat{\delta}_2(x)\,\nu^2 + \dots\right]$$

NN model





Note:
This parametrisation 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)

Maximum-Likelihood-ratio test statistic:

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

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

- $\mathrm{R}_{\boldsymbol{\nu}}$: reference hypothesis (null)
- $\mathrm{H}_{\mathbf{w},\boldsymbol{\nu}}$: alternative hypothesis
- $\mathbf{w}$: 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:

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

Depends on the NN model, sensitive to New Physics

Delta term:

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

Purely SM term, sensitive only to uncertainties related discrepancies

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

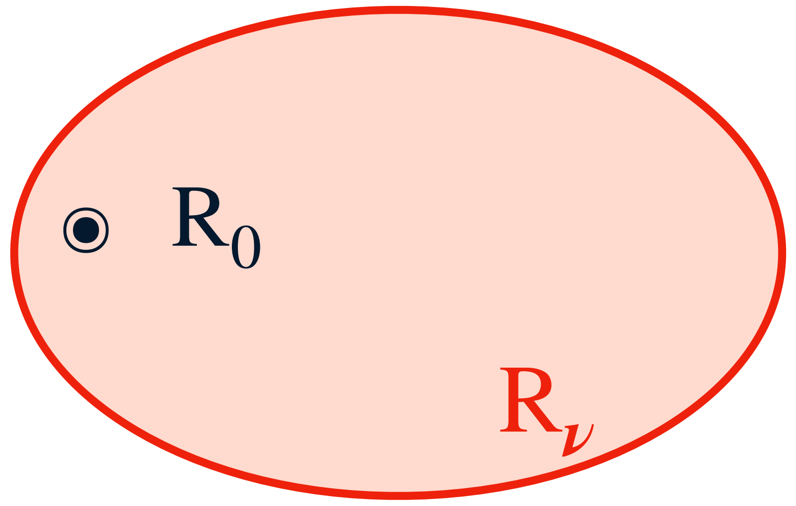
New Physics Learning Machine (NPLM)

Modelling the family of Reference hypotheses

The dependences on nuisance parameters, both in τ and Δ, are modelled via $r(x;\boldsymbol{\nu})$

r modelling

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



Normalization uncertainties:
Analytic description

$$r(x;\boldsymbol{\nu}) \equiv \frac{n(x|\mathrm{R}_{\boldsymbol{\nu}})}{n(x|\mathrm{R}_{\mathbf{0}})} = \exp\left[\sum_{i=1}^{N_\nu} \nu_i\right]$$

Shape uncertainties:
Local Taylor's expansion around the nuisance central value

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

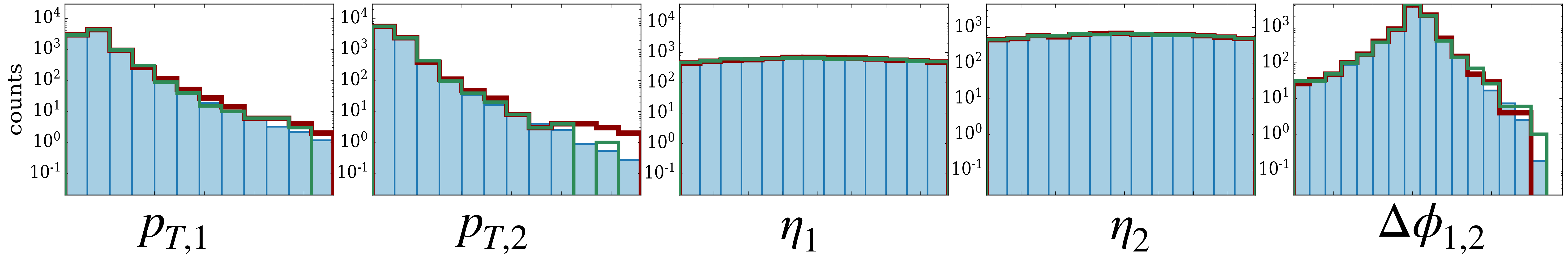
New Physics Learning Machine (NPLM)

Modelling the family of Reference hypotheses: shape effects

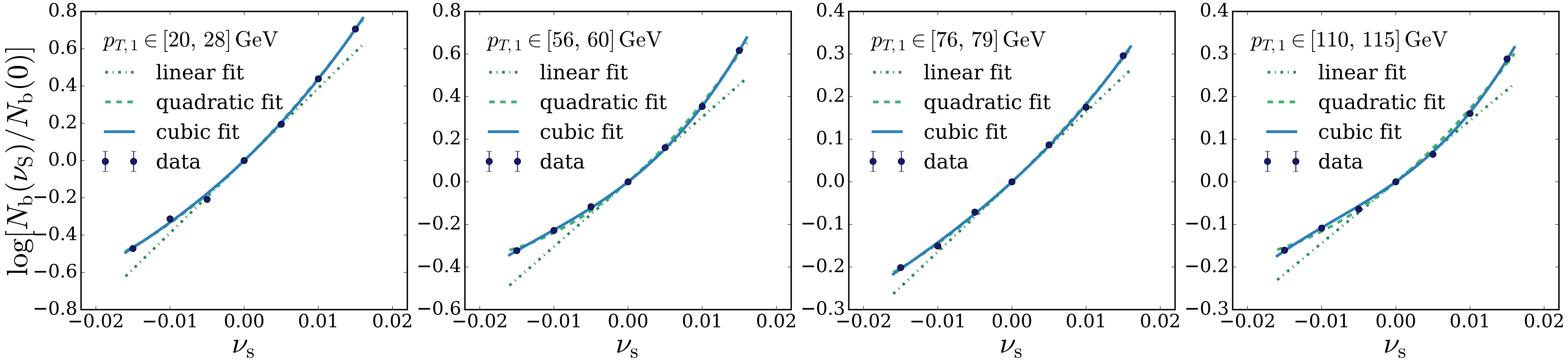
1) Preliminary study: Binned analysis to determine the proper order for the Taylor's expansion

Example:
Two-body final state
(5D analysis)

Input: set of variables describing the kinematic of the final state



Bin-wise scale effects on the electrons p_{T} distribution



New Physics Learning Machine (NPLM)

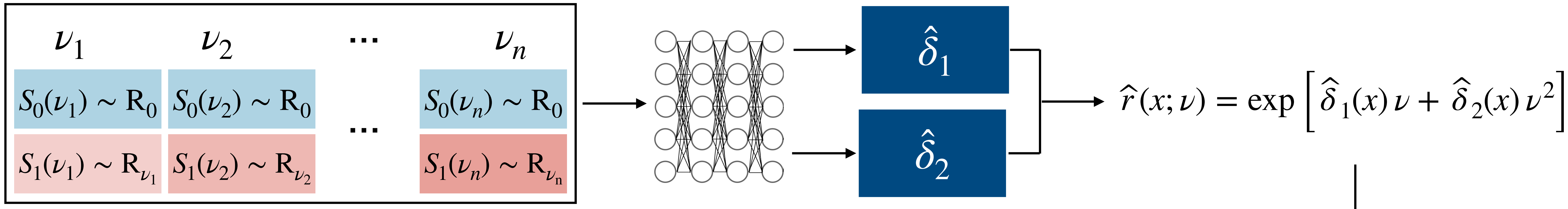
Modelling the family of Reference hypotheses: shape effects

2) Taylor's expansion learning: Training a neural network model to learn each coefficient of the Taylor's expansion of

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

Parametrised classifier

Input samples



Loss function*

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

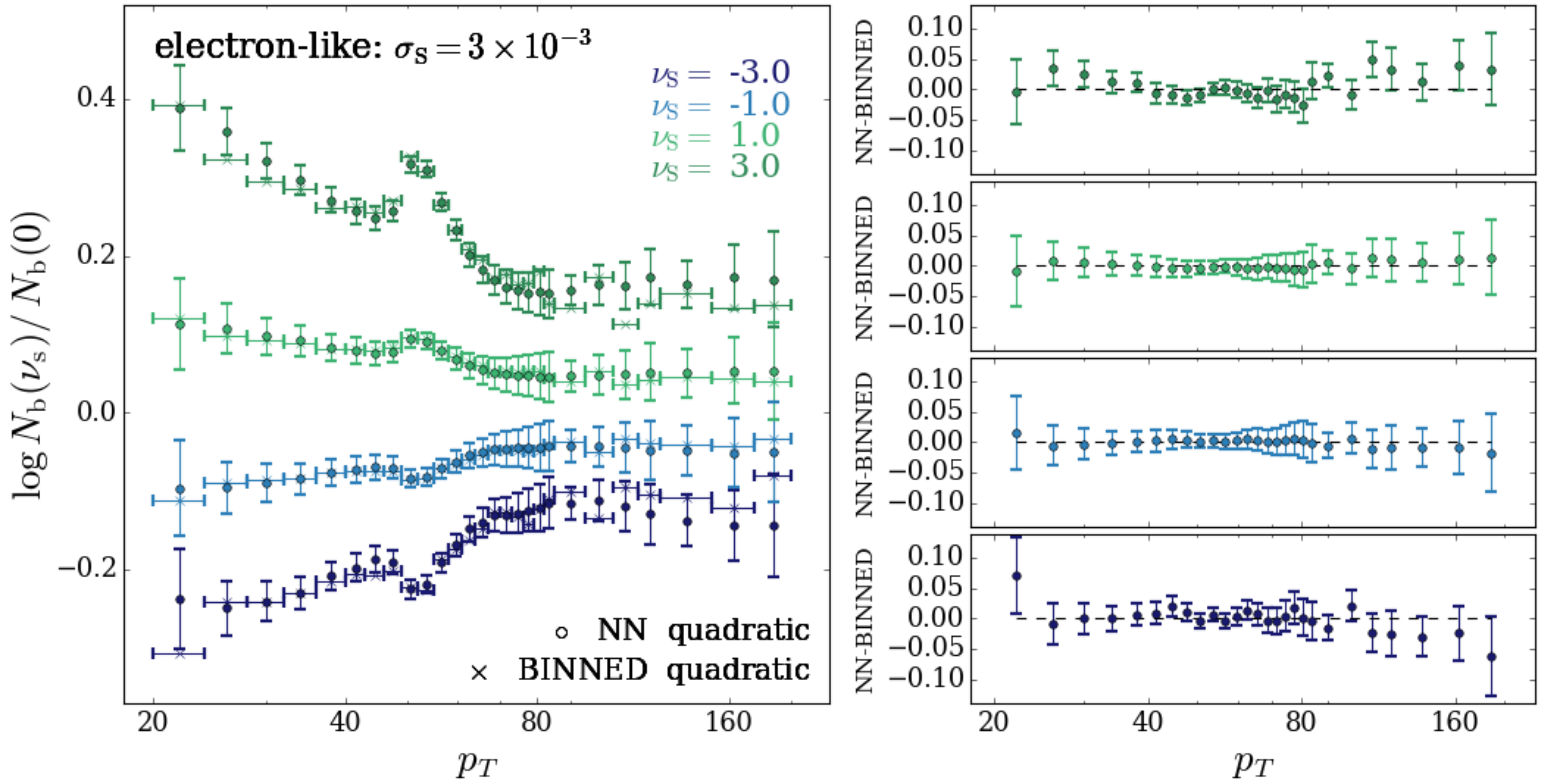
* *Parametrized classifiers for optimal EFT sensitivity* arXiv:2007.10356

New Physics Learning Machine (NPLM)

Modelling the family of Reference hypotheses: shape effects

Taylor's expansion learning

Example: result of the NN training vs. Binned approach



New Physics Learning Machine (NPLM)

The fit on nuisance parameters, both in τ and Δ, is constrained by an auxiliary likelihood term

Auxiliary term

$$\log\left[\frac{\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\mathcal{L}(\mathbf{0}|\mathcal{A})}\right]$$

We postulate that the new physics is absent in the auxiliary data*

$$\mathcal{L}(\mathrm{H}_{\mathbf{w},\boldsymbol{\nu}}|\mathcal{A}) = \mathcal{L}(\mathrm{R}_{\boldsymbol{\nu}}|\mathcal{A}) = \mathcal{L}(\boldsymbol{\nu}|\mathcal{A})$$

We suppose that the nuisance parameters are measured from the auxiliary dataset, independently from each others

$$\mathcal{A} \rightarrow \widehat{\boldsymbol{\nu}}, \text{ random variable } \widehat{\boldsymbol{\nu}} \sim \mathrm{N}(\boldsymbol{\nu}^*, \sigma_{\boldsymbol{\nu}})$$

$$2\ \log\left[\frac{\mathcal{L}(\boldsymbol{\nu}|\mathcal{A})}{\mathcal{L}(\mathbf{0}|\mathcal{A})}\right] = \sum_{\nu_i} -\left(\frac{\widehat{\nu}_i - \nu_i}{\sigma_i}\right)^2 + \left(\frac{\widehat{\nu}_i}{\sigma_i}\right)^2$$

* note: we do see the effects of new physics in the auxiliary data nonetheless, if present (see discussion in section 2.6 of Eur. Phys. J. C)

New Physics Learning Machine (NPLM)

τ term

INPUT
- Reference sample (R) label=0
- Data sample (D) label=1
- Pre-trained modules: $\hat{\delta}_1$, $\hat{\delta}_2$ ... $\hat{\delta}_n$
- Auxiliary measurements $\hat{\nu}(A)$

Model

BSM module

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

δ module: $\hat{\delta}_1$, $\hat{\delta}_2$ ... $\hat{\delta}_n$

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

Loss function

$$L[f(\,\cdot\,;\mathbf{w}),\nu,\hat{\delta}] = \sum_{e\in R} w_e\left[e^{f(x_e,\mathbf{w})+\log\hat{r}(x_e;\nu)} - 1\right] - \sum_{x\in D}\left[f(x,\mathbf{w}) + \log\hat{r}(x;\nu)\right] - \log\left[\frac{\mathscr{L}(\nu|A)}{\mathscr{L}(0|A)}\right]$$

Trainable parameters: $\mathbf{w}, \nu$

Constant terms: $\hat{\delta}$

OUTPUT

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

Δ term

INPUT
- Data sample (D label=1
- Pre-trained modules: $\hat{\delta}_1$, $\hat{\delta}_2$ ... $\hat{\delta}_n$
- Auxiliary measurements $\hat{\nu}(A)$

Model

δ module: $\hat{\delta}_1$, $\hat{\delta}_2$ ... $\hat{\delta}_n$

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

Loss function

$$L[f(\,\cdot\,;\mathbf{w}),\nu,\hat{\delta}] = \mathrm{N}(\mathrm{R}_\nu) - \mathrm{N}(\mathrm{R}_0) - \sum_{x\in D}\left[\log\hat{r}(x;\nu)\right] - \log\left[\frac{\mathscr{L}(\nu|A)}{\mathscr{L}(0|A)}\right]$$

Trainable parameters: ν

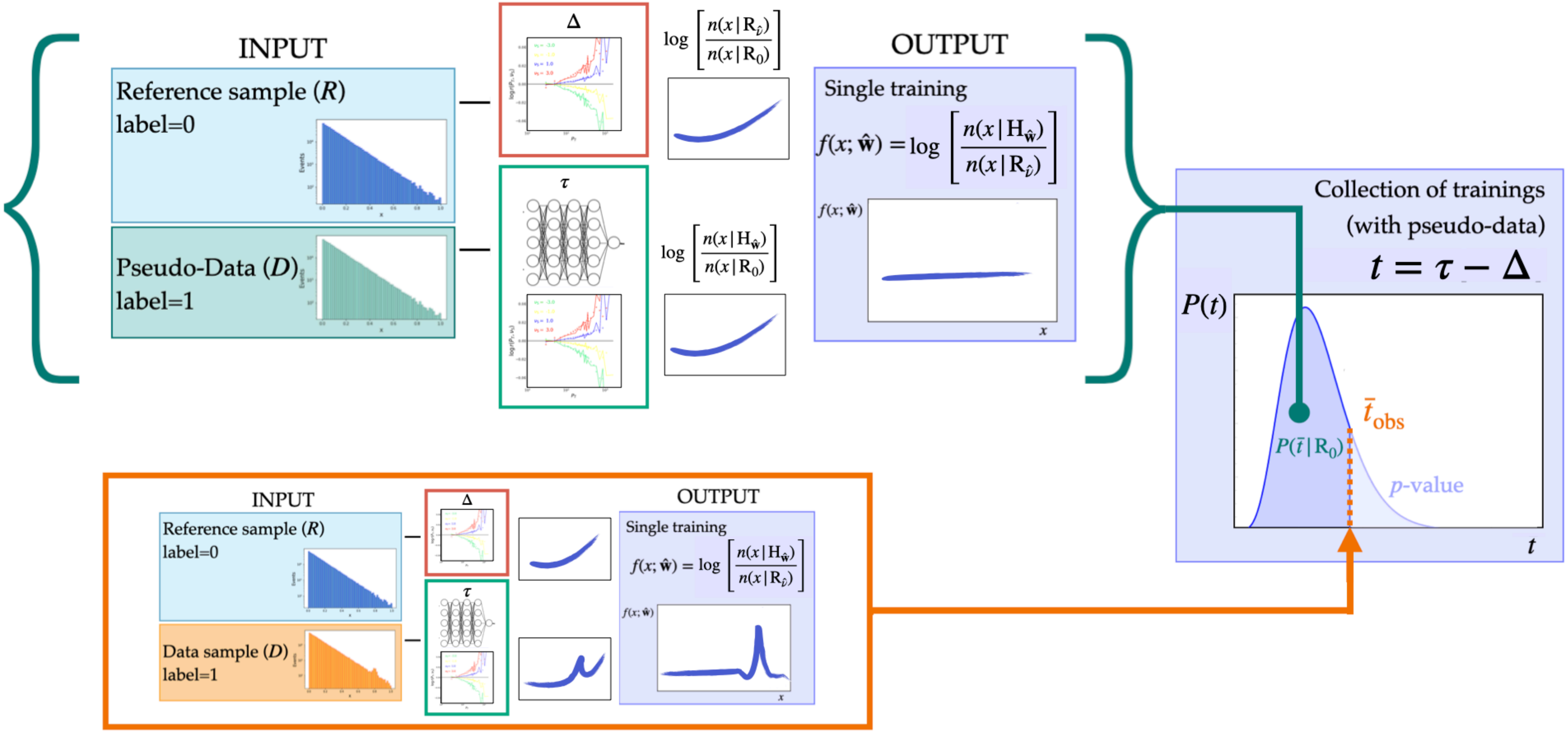
Constant terms: $\hat{\delta}$

OUTPUT

$$\Delta(D,A) = -2\min_{\mathbf{w},\nu} L\left[\nu;\hat{\delta}(\,\cdot\,)\right]$$

New Physics Learning Machine (NPLM)

Calibration



New Physics Learning Machine (NPLM)

Controlling type I errors: NN model regularisation

Due to the **finite size** of the training samples, the **sparsity** of the data (especially in multivariate problems) and the **ill-definition of the loss** (unbounded from below), the distribution of $t(D)$ under R_0 doesn't converge to a stable configuration naturally.

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

New Physics Learning Machine (NPLM)

Controlling type I errors: NN model regularisation

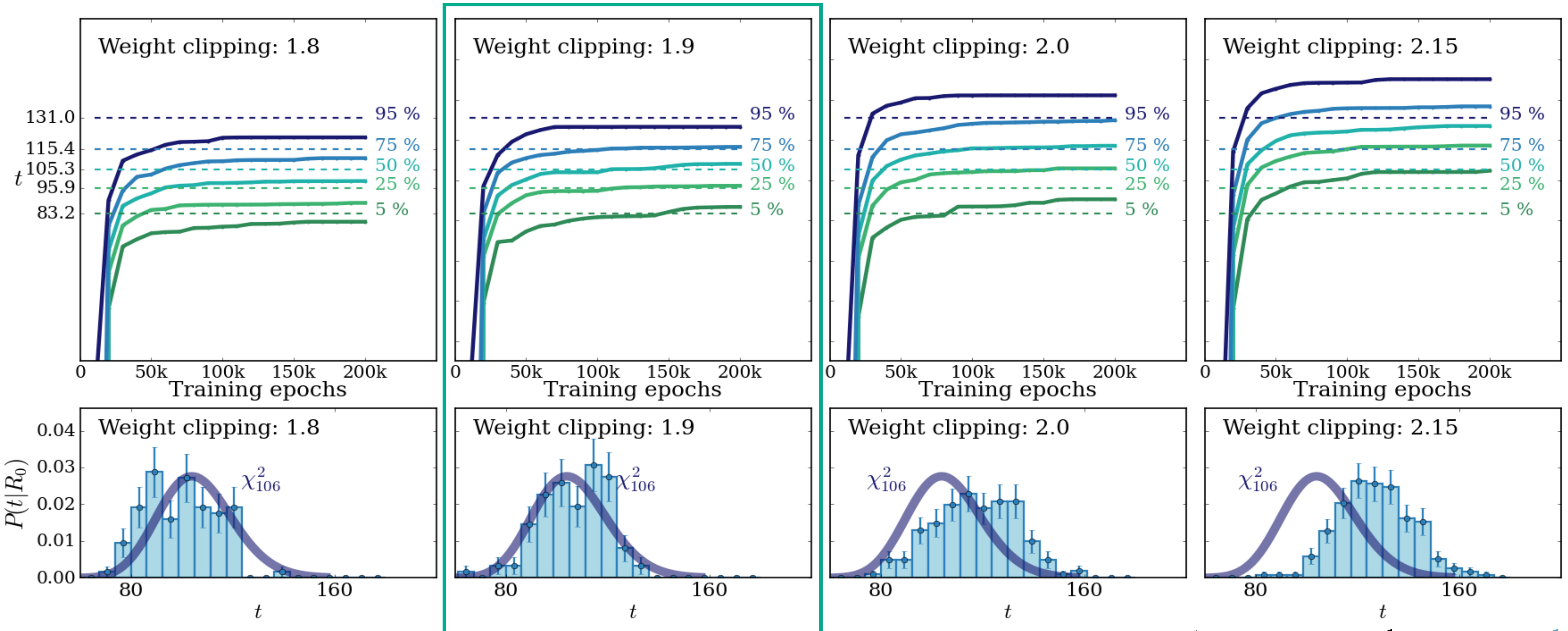
Weight clipping parameter:
Upper boundary to the magnitude that each trainable parameter can assume during the training.

→

For a chosen NN architecture, tuning the weight clipping allows to recover a good agreement of the empirical distribution of t under R_0 with a **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

New Physics Learning Machine (NPLM)

Controlling type I errors: validation of the $(\tau - \Delta)$ procedure

"Toy Data" : test the procedure on simulated toys following the Reference (SM) hypothesis with generation value for the nuisance parameters $\nu^* = \pm\sigma_\nu$:

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

The $t(D)$ distribution under the reference hypothesis $\mathrm{R}_{\boldsymbol{\nu}^*}$ is **compatible with the** $\chi^2_{|\mathbf{w}|}$ (found by regularizing) for values of the true nuisance parameters within the uncertainty ($\boldsymbol{\nu}^* = \pm\sigma_{\boldsymbol{\nu}}$).

t does **not depend** on the true value of the nuisance parameters!

We can build a *frequentist* test statistic targeting the $\chi^2_{|\mathbf{w}|}$.

New Physics Learning Machine (NPLM)

Controlling type I errors: validation of the $(\tau - \Delta)$ procedure

Reference sample: $R \sim \mathrm{R}_0$
Data sample: $D \sim \mathrm{R}_{\nu^*}$

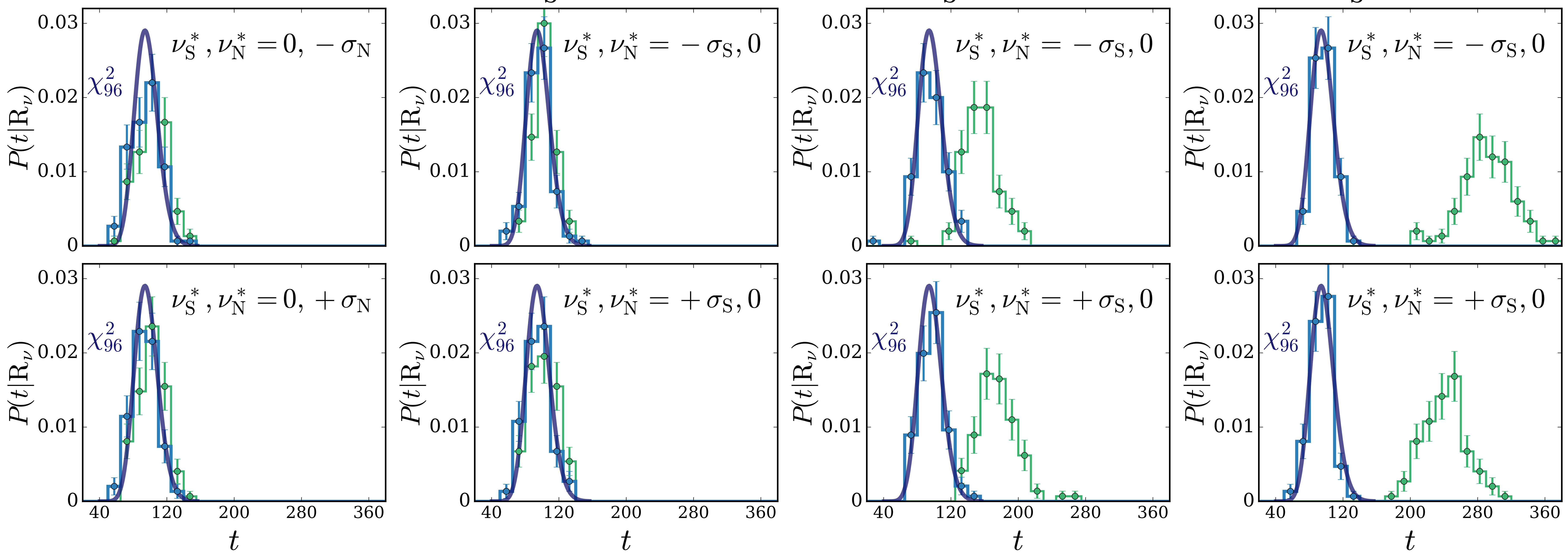
Example:
NN [5-5-5-5-1]
trainable parameters = 96
weight clipping = 2.16

Normalization uncertainty $\sigma_N = 2.5\,\%$

Scale uncertainty Muon-like regime $\sigma_S = 0.05\,\%$

Scale uncertainty Electron-like regime $\sigma_S = 0.3\,\%$

Scale uncertainty Tau-like regime $\sigma_S = 3\,\%$



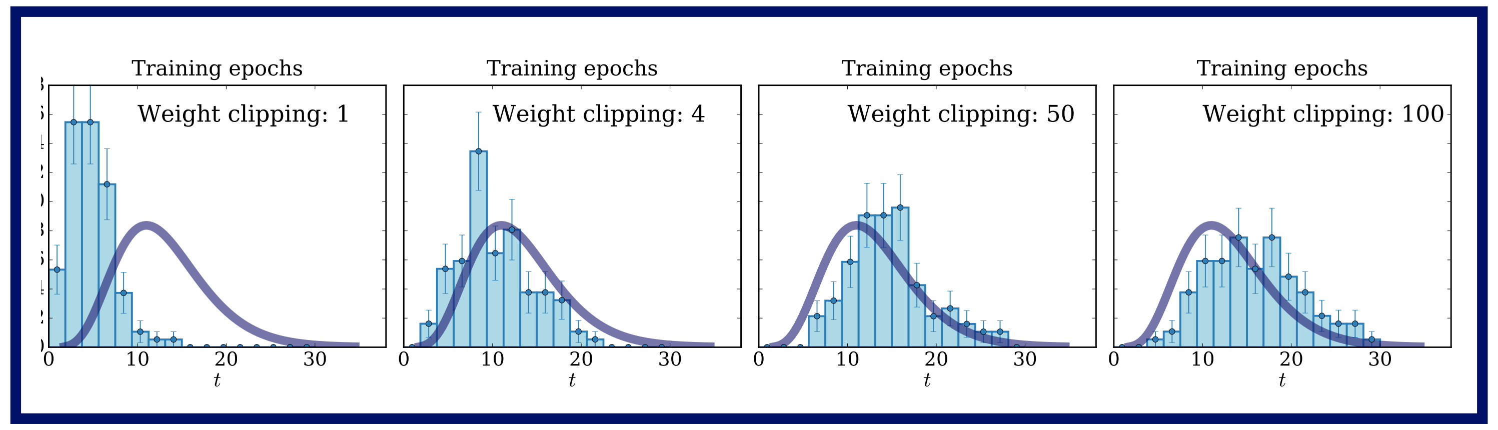
$\tau(D, A)$

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

New Physics Learning Machine (NPLM)

Final procedure in steps:

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

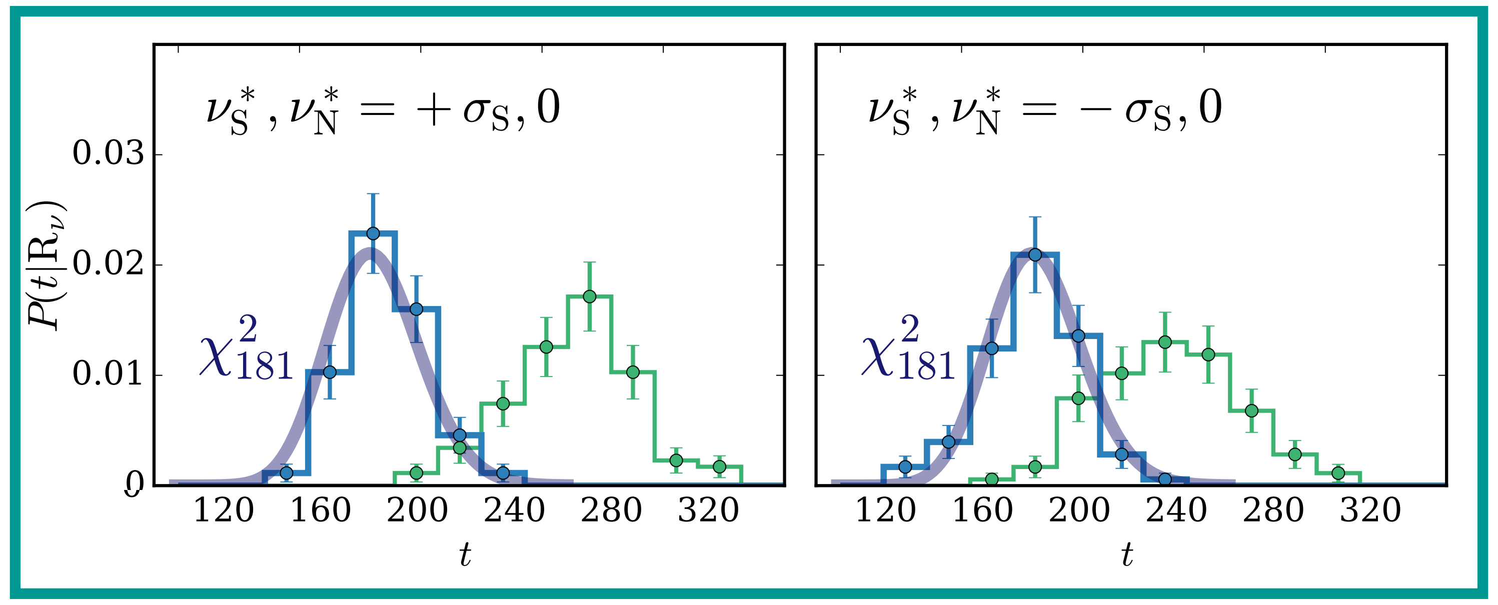


2. **NUISANCE TAYLOR'S EXPANSION LEARNING**:
modelling $\hat{r}(x;\nu) = \exp\left[\hat{\delta}_1(x)\,\nu + \hat{\delta}_2(x)\,\nu^2 + \dots\right]$;

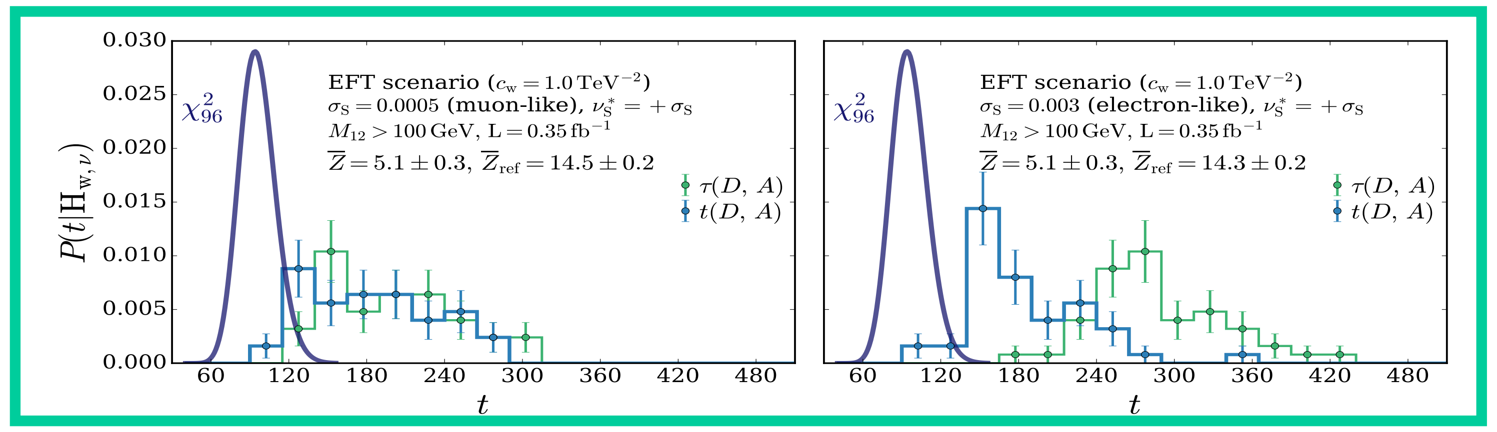
$$\hat{r}(x;\nu) = \exp\left[\hat{\delta}_1(x)\,\nu + \hat{\delta}_2(x)\,\nu^2 + \dots\right]$$

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.



New Physics Learning Machine (NPLM)

Sensitivity to BSM scenarios

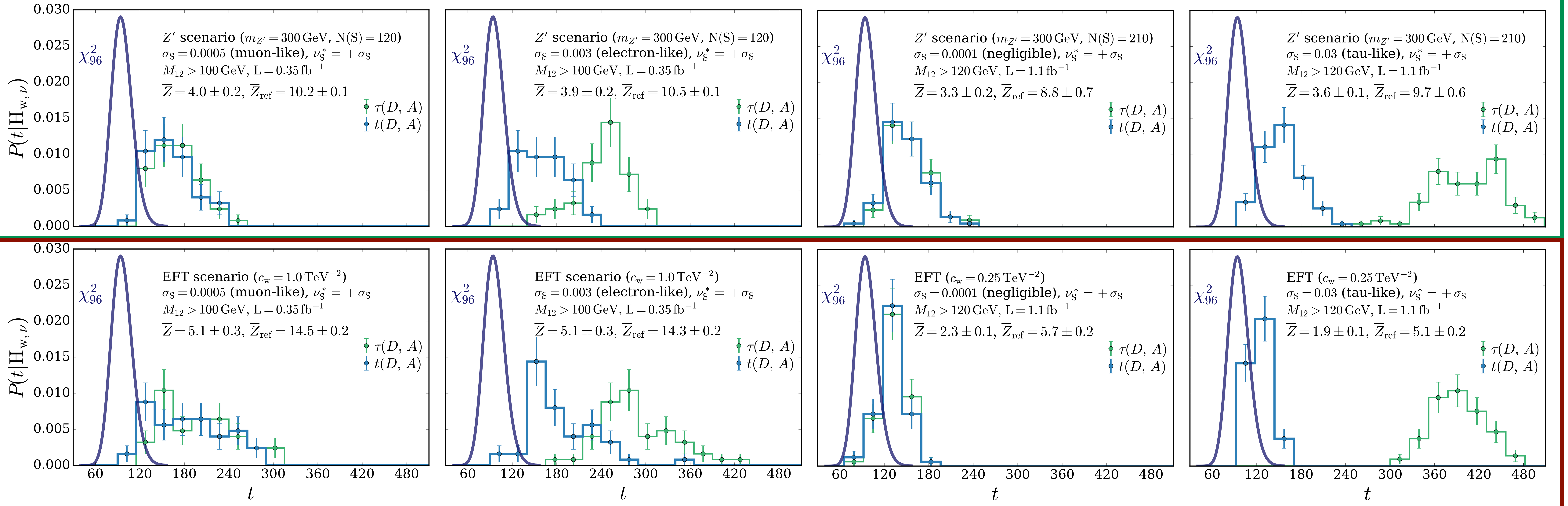
Muon-like regime | Electron-like regime | Negligible systematic uncertainties | Tau-like regime

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

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

$\tau(D, A)$

New Physics Learning Machine (NPLM)

Summary and outlook:

NPLM features:

- End-to-end analysis strategy: from data to a frequentist p-value for discovery
- Multivariate test
- Dealing with systematic uncertainties
- Sensitive to multiple signal patterns at once (global p-value)

NPLM Challenges:

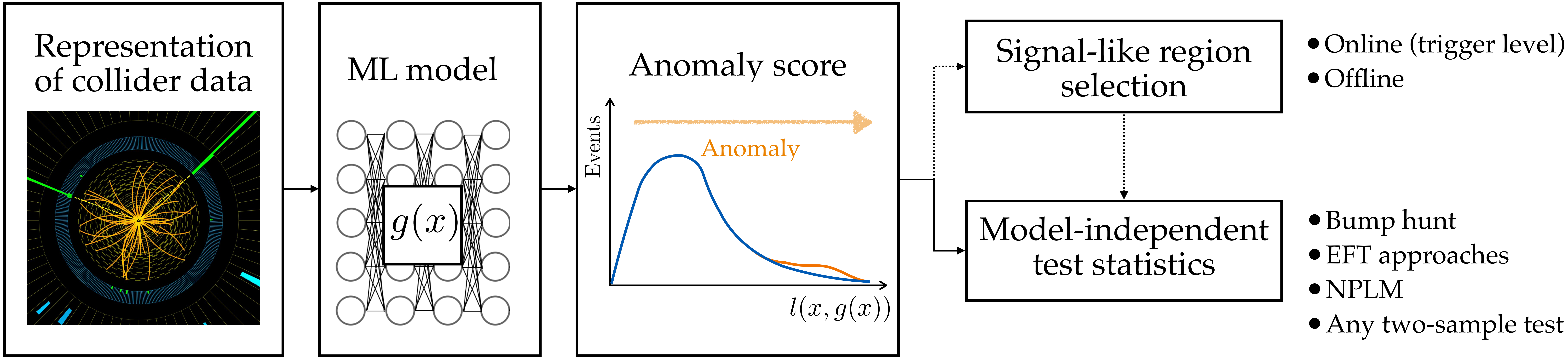
- Execution time (working on a fast kernel methods based implementation Eur. Phys. J. C, 82(10))
- Accurate multidimensional modelling of systematics (accuracy needed grows with the experimental luminosity)

Theoretical aspects to be further investigated:

- The model regularisation selects a portion of the space of the hypotheses, how to avoid performances losses? (exploring smarter alternatives)
- Emergence of a χ^2 for the test statistics distribution in the null hypothesis

ML Tools for model-independent analyses

Anomaly detectors:



Some comments to be further discussed:

- Trained on data (no sys. unc. but need for control regions →less inclusive test)
- Trained on MC (study the response to input domain shifts, like those produced by systematic uncertainties)
- Which model-independent test to run on top?

Backup slides

New Physics Learning Machine (NPLM)

Main Idea: Maximum Likelihood from minimal loss

Parametrizing the alternative distribution

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

R_0 : Reference (null) hypothesis (SM)
$\mathrm{H}_{\mathbf{w}}$: Alternative hypothesis (BSM)

Test statistic

$$\bar{t}(\mathcal{D}) = 2\max_{\mathbf{w}} \log\left[\frac{\mathcal{L}(\mathcal{D}\,|\,\mathrm{H}_{\mathbf{w}})}{\mathcal{L}(\mathcal{D}\,|\,\mathrm{R}_0)}\right] = 2\max_{\mathbf{w}}\left\{\log\left[\frac{e^{-N(\mathbf{w})}}{e^{-N(\mathrm{R})}}\prod_{i=1}^{\mathcal{N}_{\mathcal{D}}}\frac{n(x_i|\mathbf{w})}{n(x_i|\mathrm{R})}\right]\right\}$$

$$= -2\min_{\mathbf{w}}\{\bar{L}[f(\cdot;\,\mathbf{w})]\}$$

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

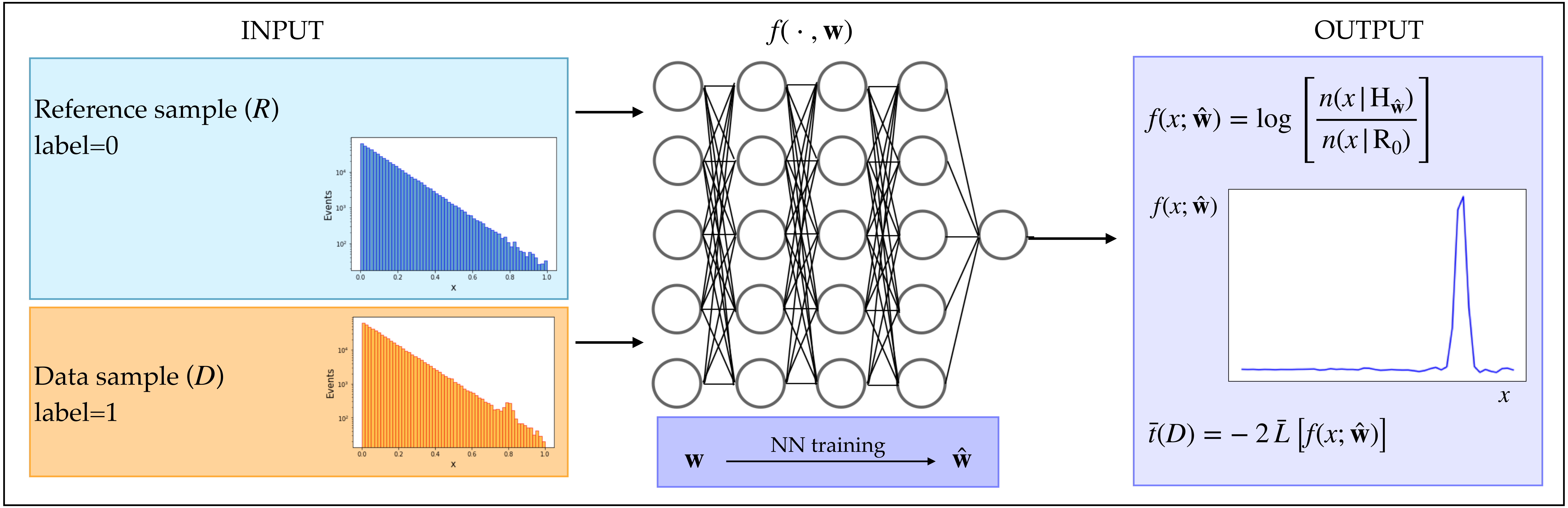
$$\bar{L}\,[f(x;\,\mathbf{w})] = -\sum_{x\in\mathcal{D}}[\,f(x;\,\mathbf{w})] + \sum_{x\in\mathcal{R}} w_x\left[e^{f(x;\,\mathbf{w})} - 1\right]$$

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 $\sum_{x\in R} w_x = N(\mathrm{R}_0)$

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

New Physics Learning Machine (NPLM)

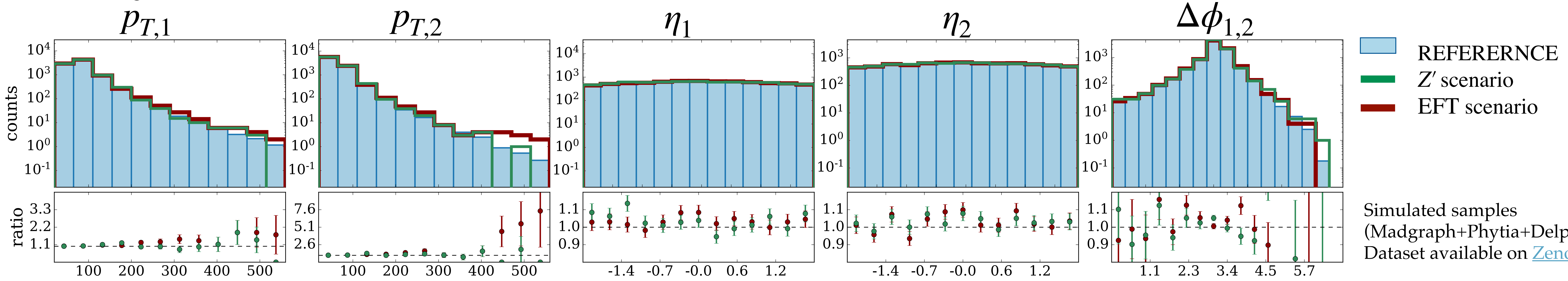
Main Idea: Maximum Likelihood from minimal loss



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

Two-body final state

5D analysis — Input: set of variables fully describing the kinematic of the final state



Simulated samples (Madgraph+Phytia+Delphes)
Dataset available on Zenodo

Uncertainties on the Reference hypothesis (SM):

- Global normalization effect: $\sigma_N = 2.5\,\%$
- Momentum scale effect:

$$p_{T\,1,2}^{(b,e)} = \exp\left[\nu_s \sigma_s^{(b,e)}/\sigma_s^{(b)}\right] p_{T\,1,2}^{(b,e)}$$ (b) barrel region $|\eta| < 1.2$, (e) endcaps region $|\eta| \geq 1.2$

- Muon-like regime: $\sigma_S^{(b)} = 0.05\,\%$, $\sigma_S^{(e)} = 0.15\,\%$
- Electron-like regime: $\sigma_S^{(b)} = 0.3\,\%$, $\sigma_S^{(e)} = 0.9\,\%$
- Tau-like regime: $\sigma_S^{(b)} = \sigma_S^{(e)} = 3\,\%$

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

Two-body final state
BSM scenarios

Resonance in the two-body invariant mass

- **Z' scenario:** new vector boson with the same SM coupling as the Z boson and mass of 300 GeV.
  - Muon-like, electron-like regimes:
    $M_{12} > 100\,\text{GeV}, L = 0.35\ \text{fb}^{-1}, \text{N(S)} = 120$
  - Tau-like regime:
    $M_{12} > 120\,\text{GeV}, L = 1.1\ \text{fb}^{-1}, \text{N(S)} = 210$

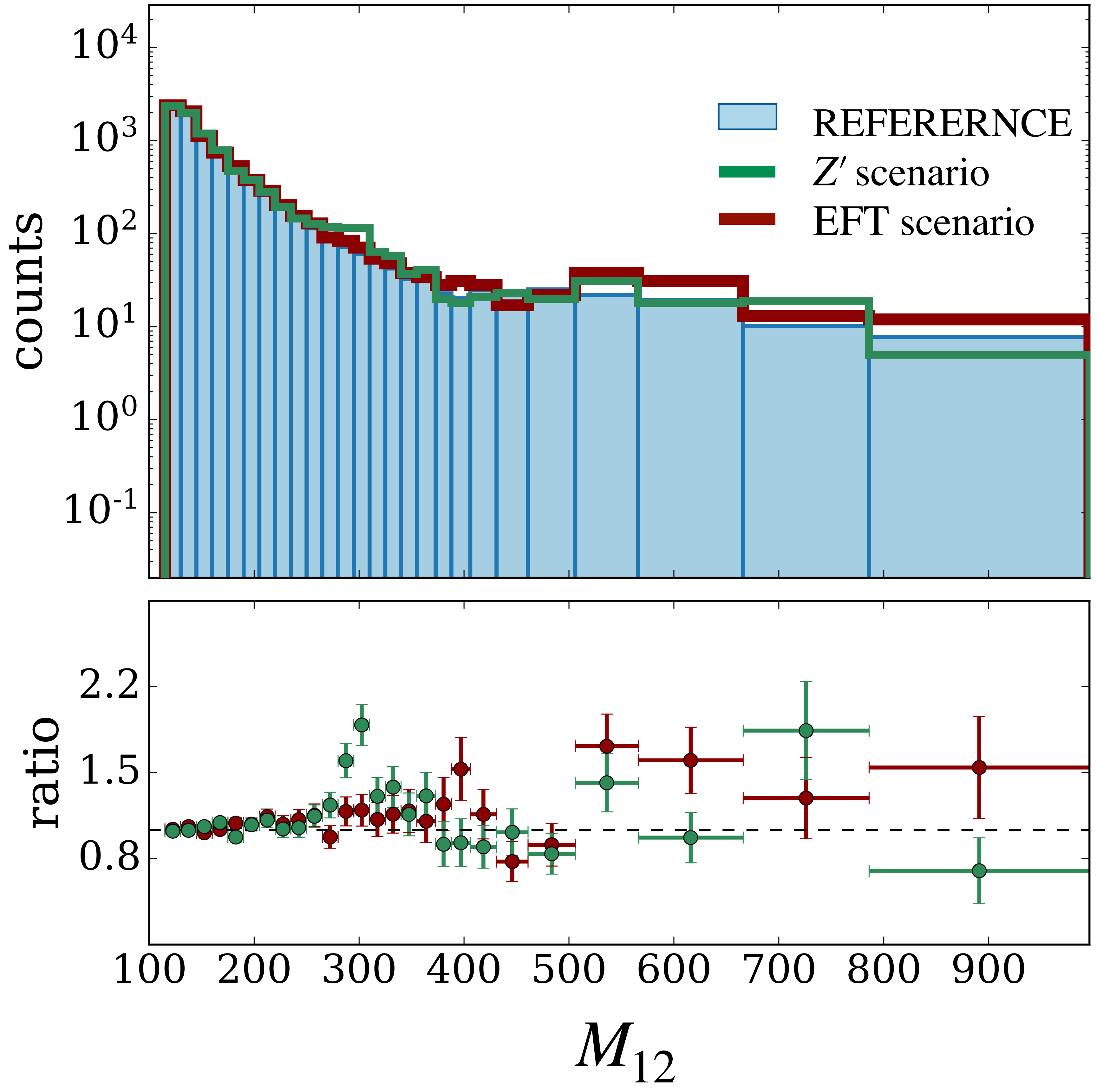
Non resonant excess in the tail of the two-body invariant mass

- **EFT scenario:** dimension-6 4-fermions-contact operator:

$$\frac{c_W}{\Lambda} J^a_{L\mu} J^\mu_{La}.$$

  - Muon-like, electron-like regimes:
    $M_{12} > 100\,\text{GeV}, L = 0.35\ \text{fb}^{-1}, c_W = 1.0\,\text{TeV}^{-2}$
  - Tau-like regime:
    $M_{12} > 120\,\text{GeV}, L = 1.1\ \text{fb}^{-1}, c_W = 0.25\,\text{TeV}^{-2}$

Example:
Tau-like regime



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

Two-body final state

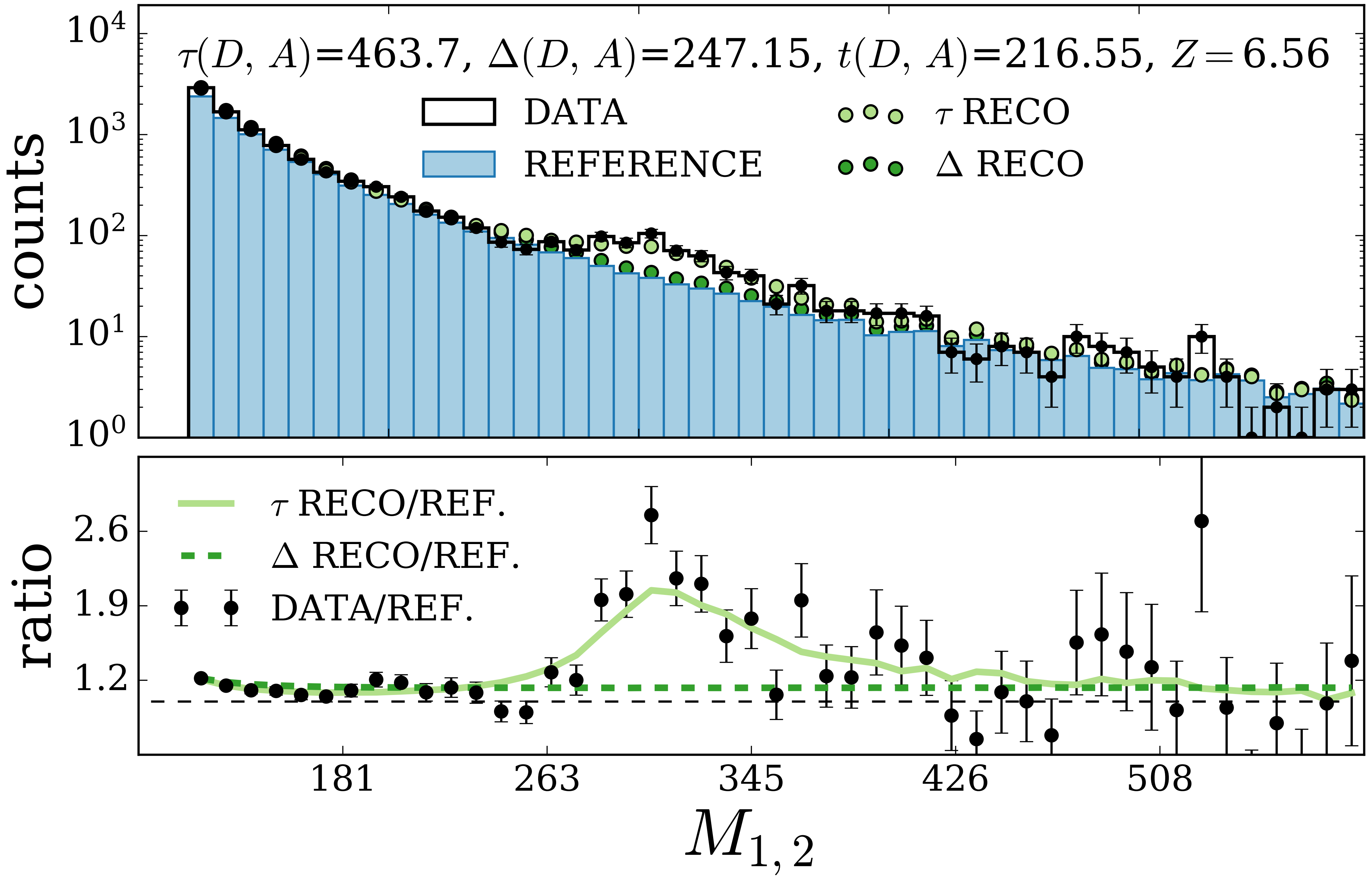
Signal reconstruction with the NN:

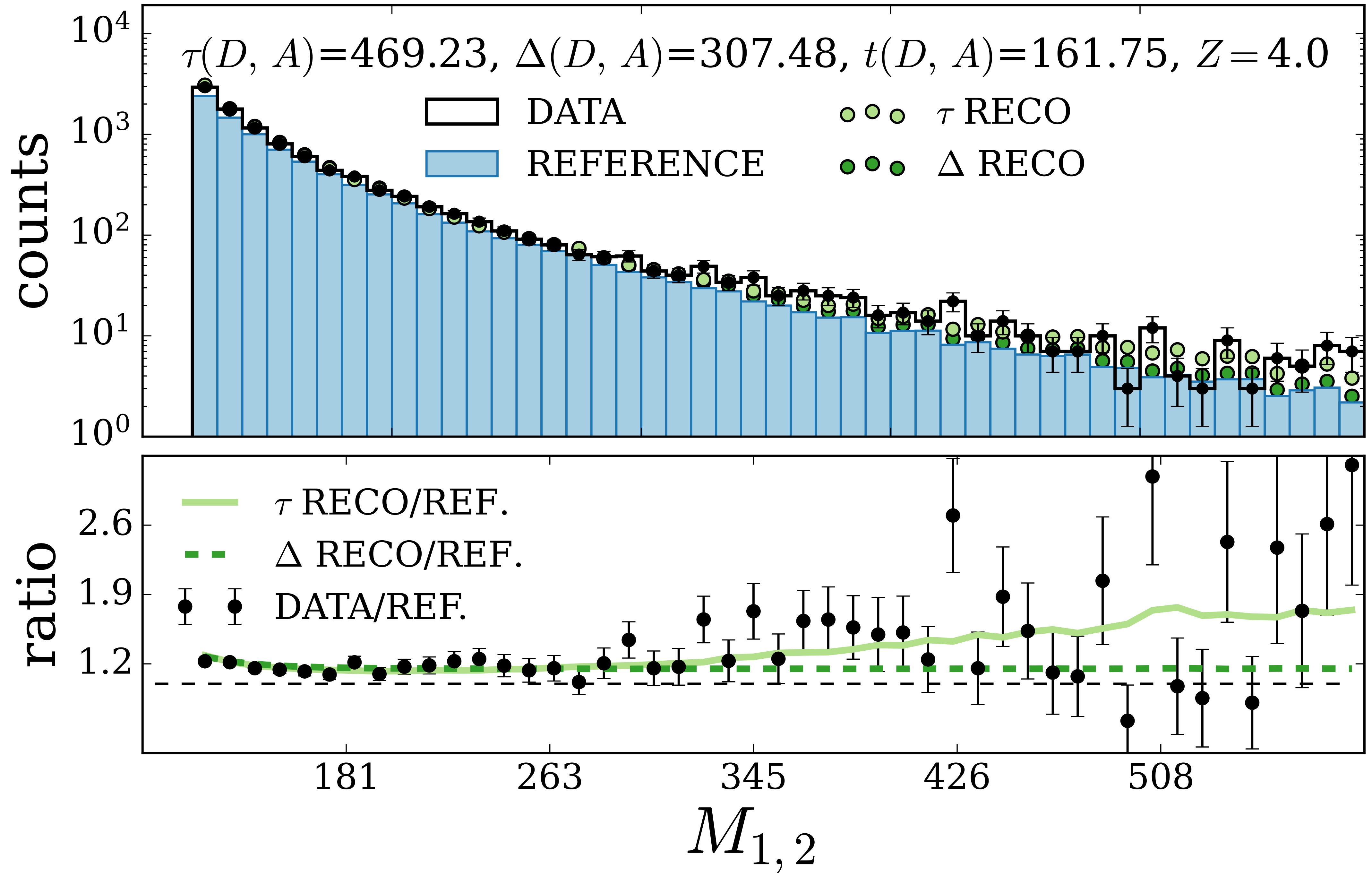
Architecture: [5-5-5-5–1] (96 dof), weigh clipping 2.15, L = 240 fb^{-1}

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

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

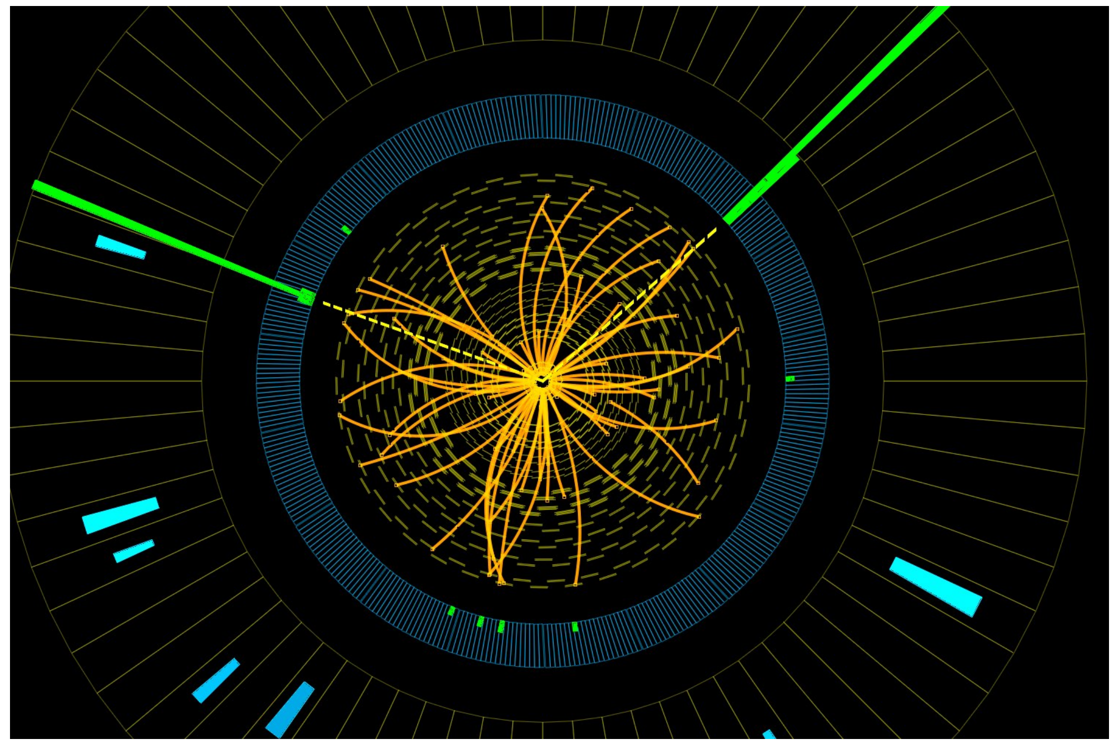
Δ reconstruction: $n(x\,|\,\mathrm{R}_{\hat{\nu}})$



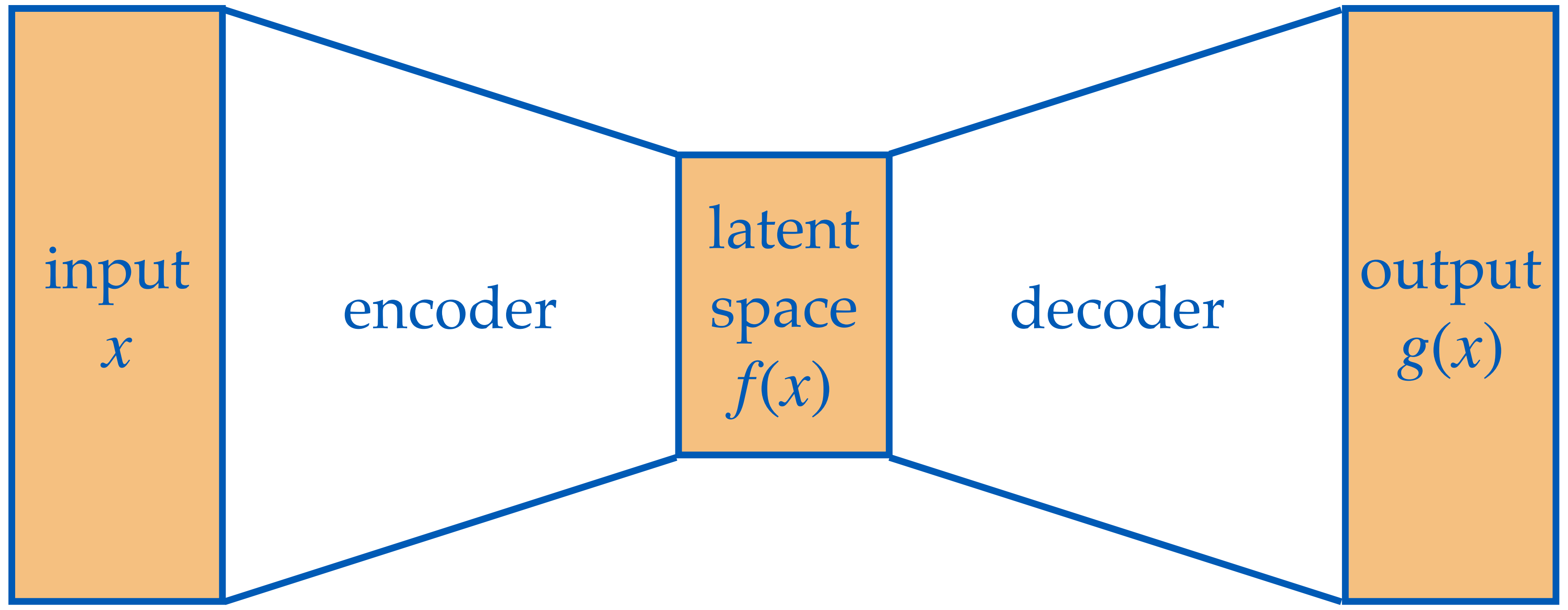


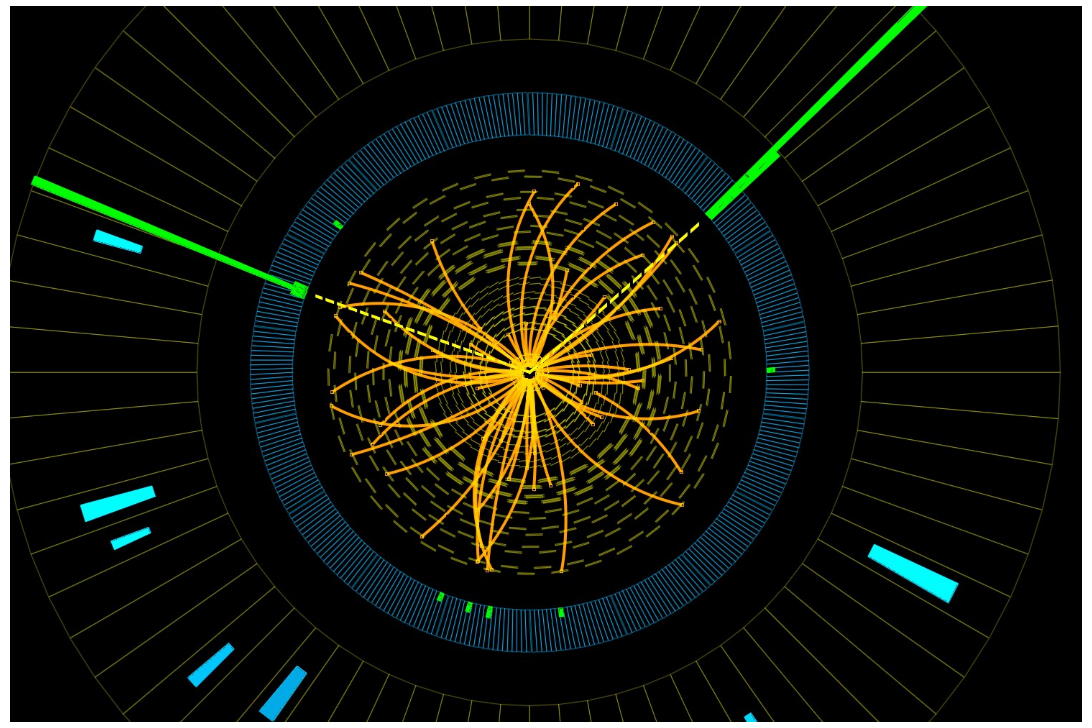
"Learning Multivariate New Physics" Eur. Phys. J. C

Model-independent approaches
Anomaly detectors: autoencoders



Collider data representation





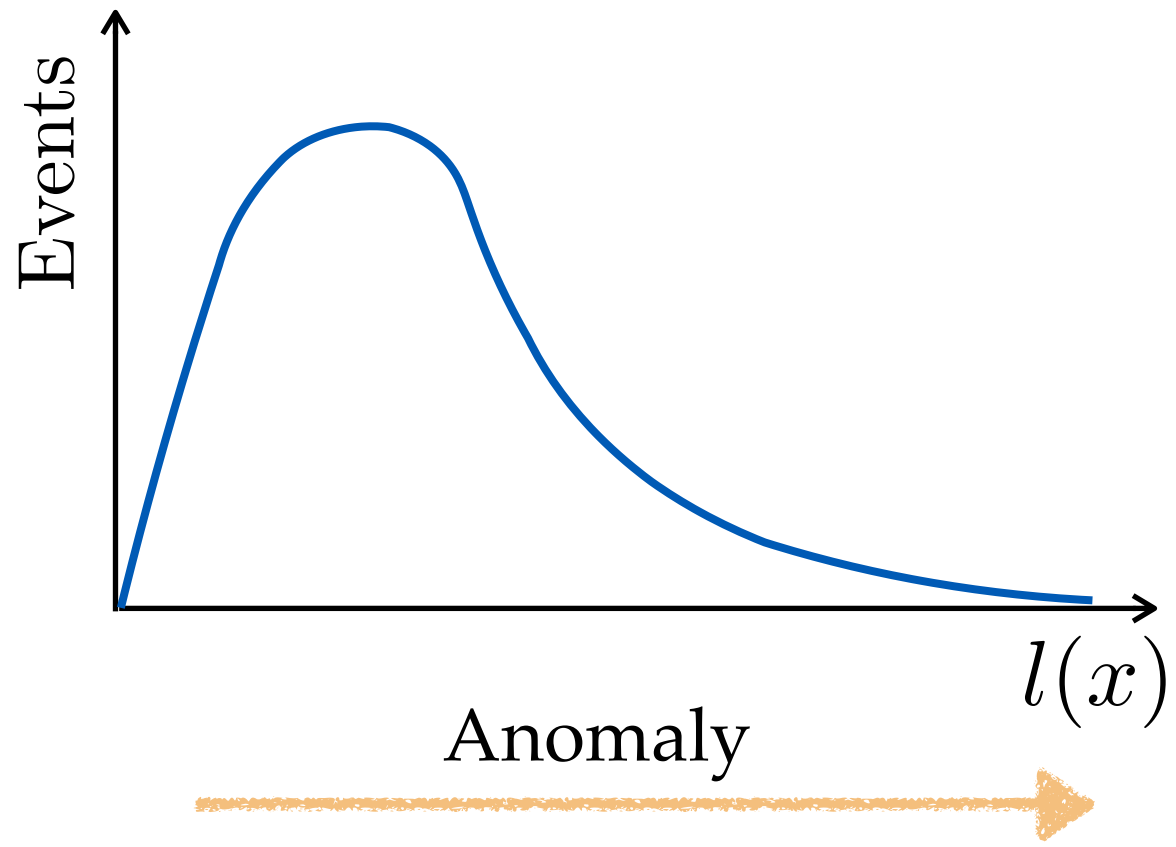
Collider data representation

Reconstruction loss: $L_{\text{reco}} = \sum_x l(x, g(x))$

↓

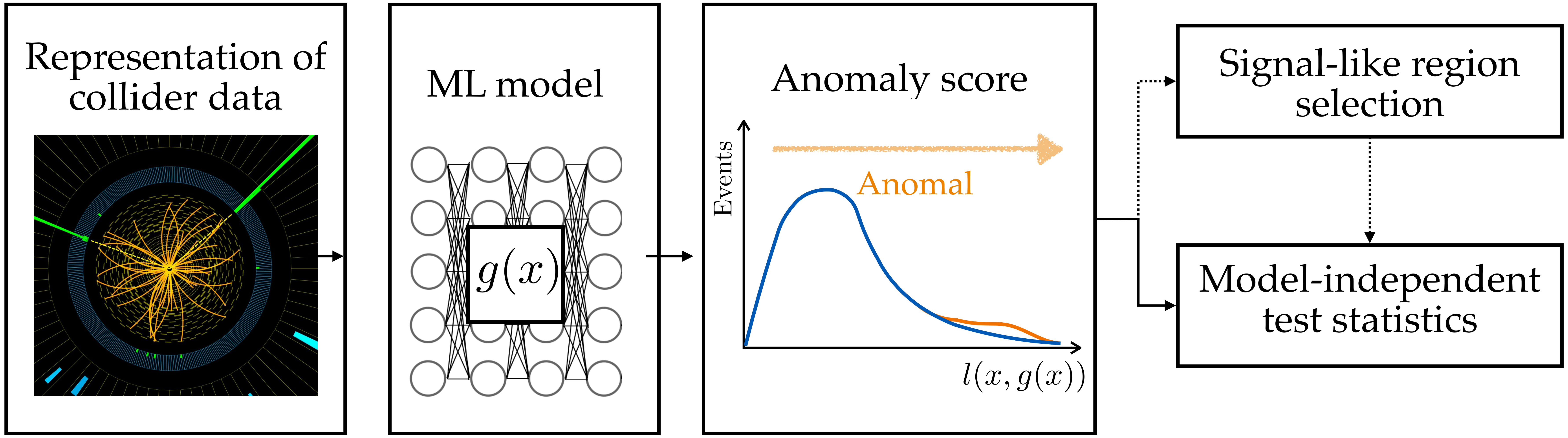
Summary statistics: $l(x, g(x))$

↳ Test statistics



ML Tools for model-independent analyses

Anomaly detectors:



Some examples:

- CWoLa Hunting [Phys. Rev. Lett. 121 (2018), no. 24, 241803 , Phys. Rev. D 99 (2019), no. 1, 014038]
- Tag N' Train (TNT) [JHEP 01 (2021) 153]
- Classifying Anomalies THrough Outer Density Estimation (CATHODE) [arXiv:2109.00546]
- Variational Autoencoders [arXiv:2110.08508]