Model independent measurements of Standard Model cross sections with domain adaptation

B. Camaiani, R. Seidita, L. Anderlini, R. Ceccarelli, V. Ciulli, P. Lenzi, M. Lizzo, L. Viliani

(Re)interpretation of the LHC results for new physics
December 15, 2022

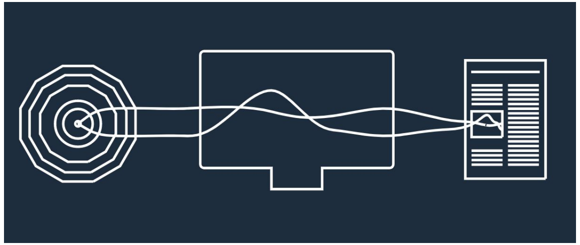
Introduction

- HEPData is an open-access repository for High Energy Physics (HEP) data, that comprises results related to several thousand publications, including those from the Large Hadron Collider (LHC)

- Let's imagine someone in the future who wants to re-interpret a result from HEPData and compare it with a new physics model

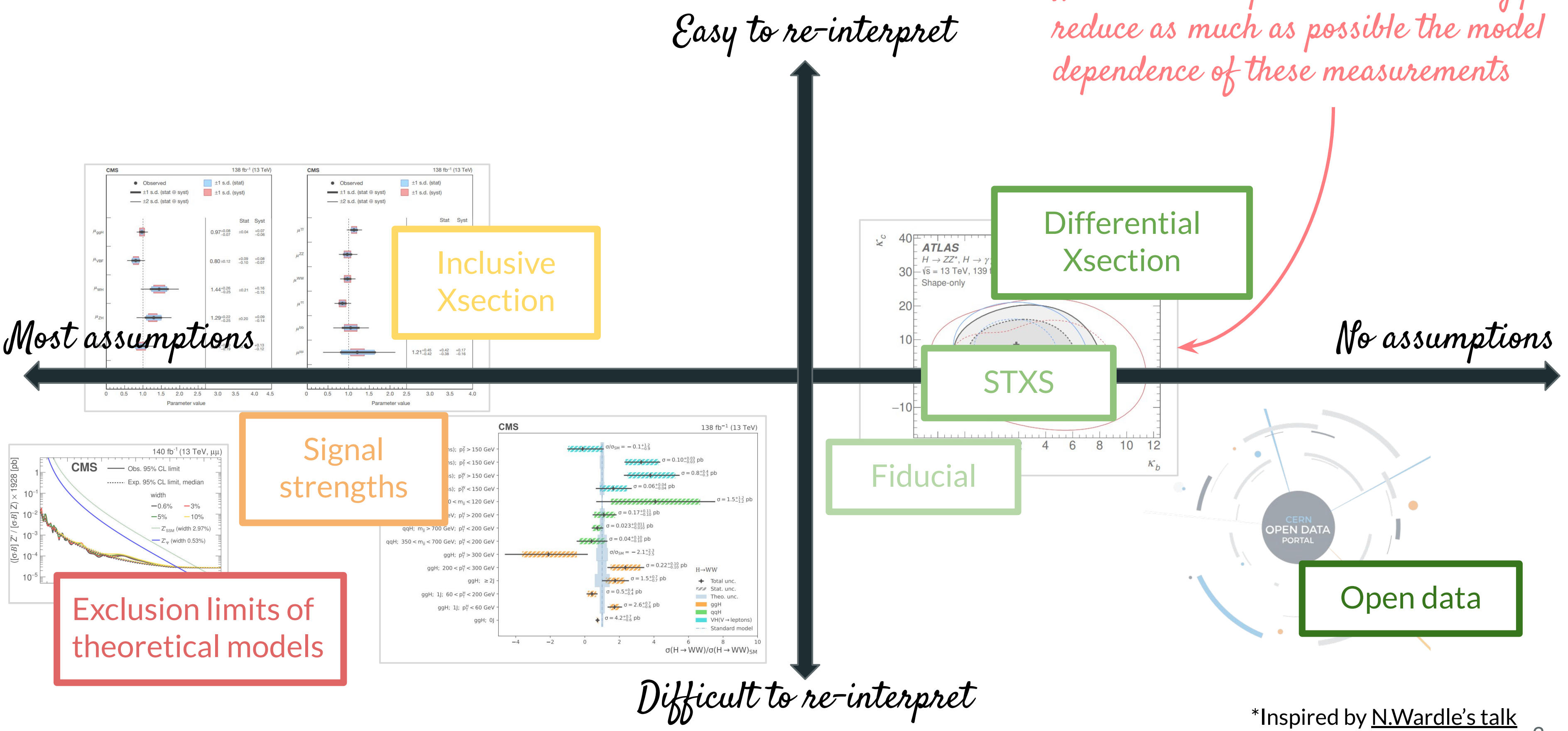
- If the analysis that provided the result was performed with some physics model assumptions - which is often the case - the result will be biased towards that model, spoiling the re-interpretability of the measurement and leaving our friend confused

HEPData



Our goal is reflect on how design a measurement that can be easily re-interpreted, and thus limiting the model dependence as much as possible

LHC measurements



*Inspired by N.Wardle's talk

Sources of model dependence

- **Unfolding:** the response matrix quantifying the smearing between the reconstructed- and particle-level phase space can depend on the physics model

In the case the response matrix is quite diagonal, the related model dependence is expected to be small and may be neglected

- **Definition of the phase space:** event selection efficiency and acceptance factors may be different according on the considered signal model

Selection efficiency depends on the detector smearing, whereas the acceptance factor can be estimated a posteriori and be accounted for

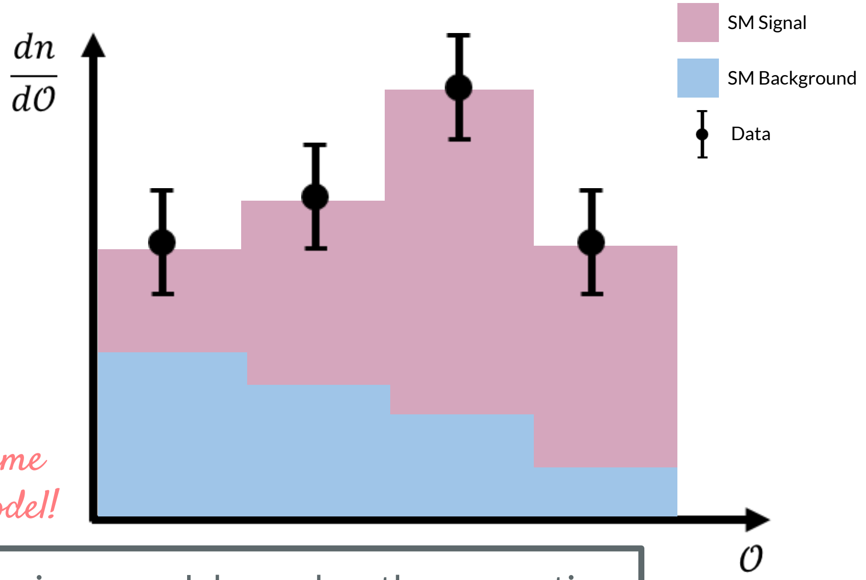
- **Signal extraction procedure:** performed via a template fit ("shape effect")

Main source of model dependence
(see next slide)

Template fit

- Let $\mathcal{O}$ be an observable with good signal to background discriminating power
- The probability density function (PDF) is usually not known a priori and it is replaced by templates
- Monte Carlo simulations are produced for both signal and background process
- MC simulation of the signal is fitted to experimental data

Need to assume a physics model!



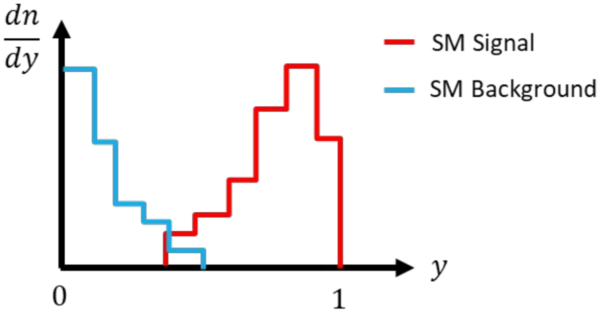
The shape of the observable distribution may in general depend on the properties of the physics theory governing the signal process under consideration

The fit result has a bias toward the prediction of the physics model used to generate the templates

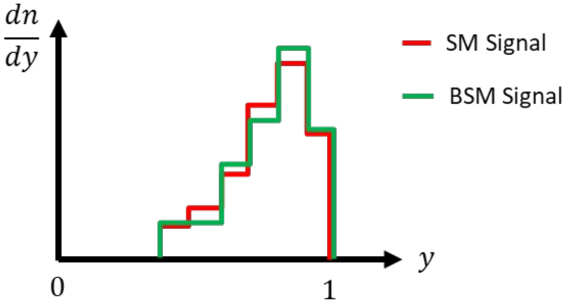
Main Objective

The main goal is to implement a new fit variable y that is agnostic with respect to the signal hypothesis:

- y must be able to discriminate signal from background events



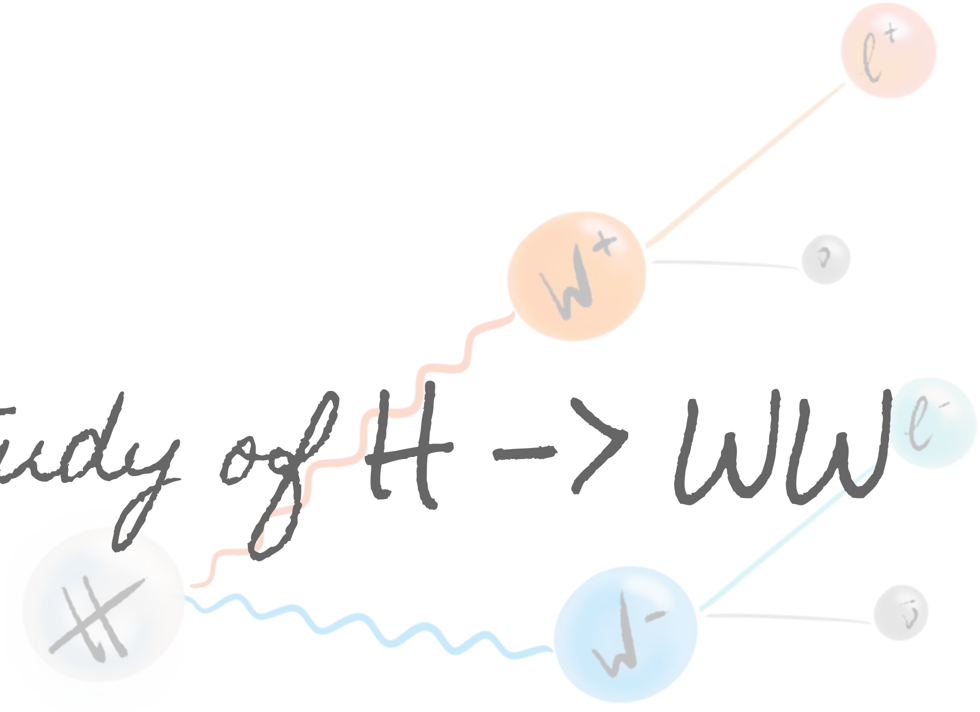
- y must not be able to distinguish the physics model of signal events



y does not introduce a bias in the fit result since the shape of its distribution is roughly the same regardless of the theoretical model describing the data

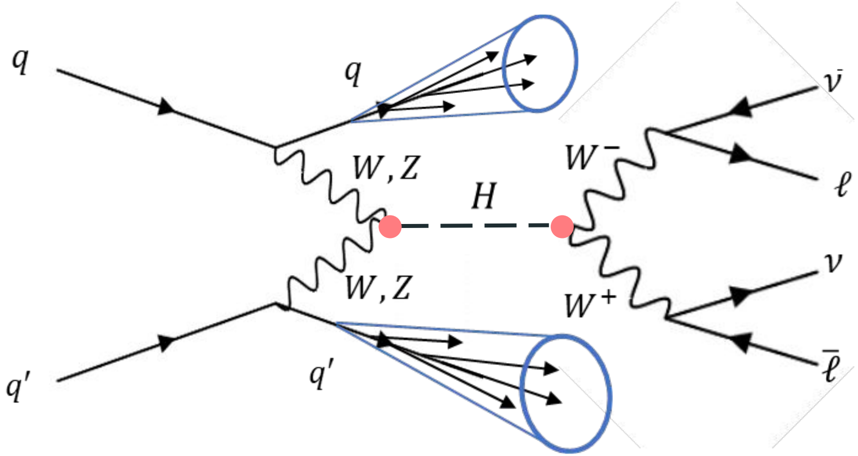
Domain Adaptation *focus on an* ***Adversarial Deep Neural Network***

The case study of H -> WW



EPJC 82 921 (2022)

Signal process



Vector boson fusion (VBF)

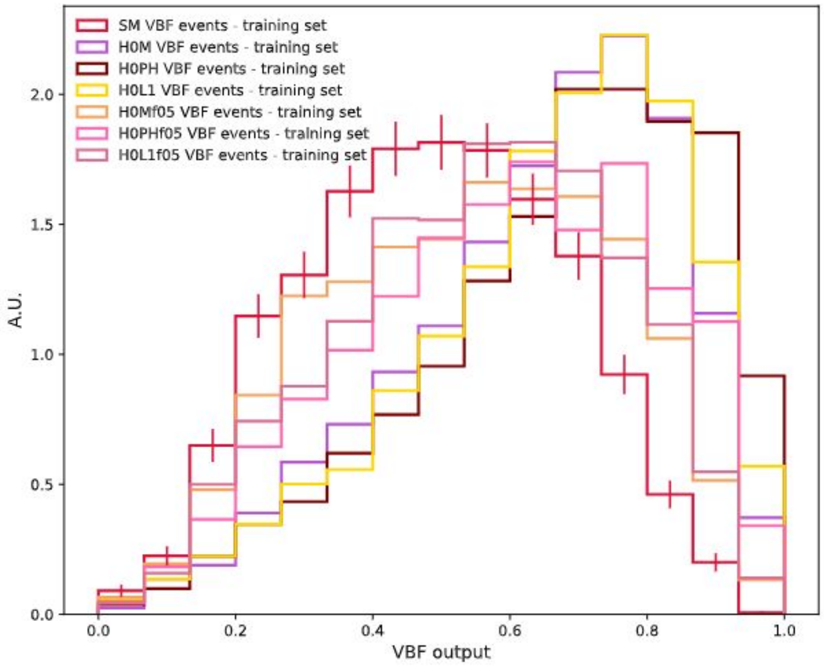
- Rare process
- In H→WW→2l2v, Higgs boson invariant mass can not be reconstructed
- Main backgrounds: ggH, non resonant WW and top pair production

Some BSM theories predict Anomalous Couplings (AC) in the HVV vertex
(more details in backup)

Objective:
Define a deep neural network that discriminate signal from backgrounds **in a way that is independent on the physics model describing the coupling in the HVV vertex**

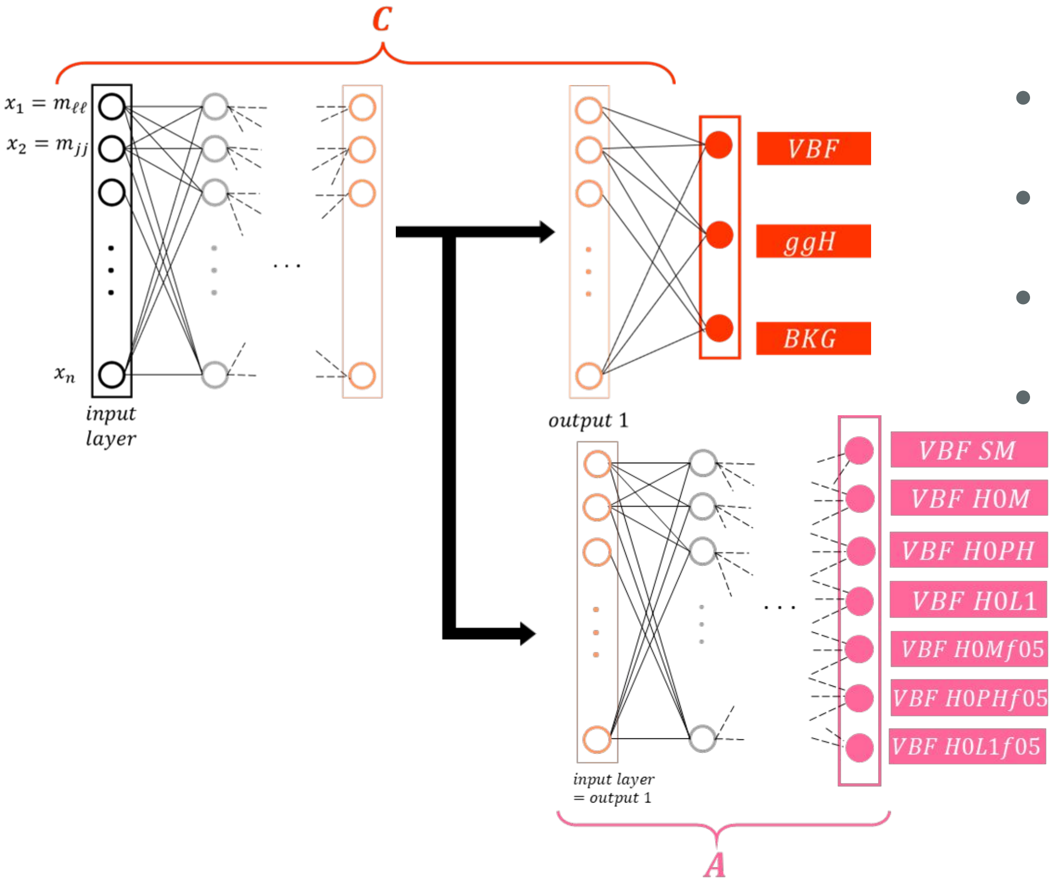
Training performed in a STXS-like phase space region, defined by $350 < m_{\mathrm{jj}} \leq 700 \text{ and } p_{\mathrm{T}}^{\mathrm{H}} < 200 \text{ GeV}$

First attempt: the simplest approach



- Standard feed-forward **DNN with classification tasks trained on:**
  - **SM VBF events**
  - **BSM VBF events**
  - SM gluon-gluon fusion and background events
- Three outputs: VBF, ggH and BKG
- Shapes of the VBF-output differ significantly, showing a **residual degree of dependence on the physics model assumption for the signal process**
- Next step: Domain Adaptation

Adversarial deep neural network (ADNN)



Classifier

- Takes as input the measurable kinematic variables of an event
- Aims to determine if the event is signal- or background-like
- Each output represents the probability that an event belongs to the corresponding class
- Is trained on a data sample including events coming from different "domains", i.e. different signal models

Adversary

- Is trained only on signal events (SM + 6 BSM theories)
- Tries to guess the physics model of signal events, regressing the domain from the second-to-last layer of C

Competitive learning

- The classifier is penalized if its output contains too much information on the domain of origin of signal events

- If C manages to prevent A from identifying the signal model, then the classification is independent of the domains of origin of the events

Compute first the gradient of $\mathcal{L}$ with respect to the C weights.
A weights frozen in this step.

The parameter α regulates the interplay between A and C

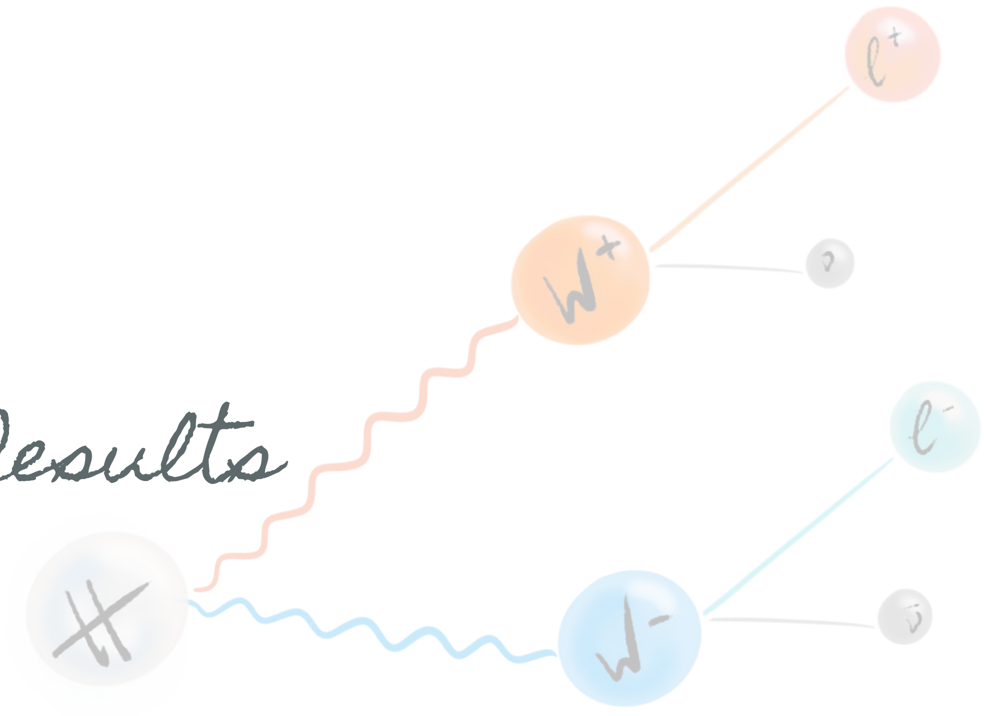
Two-step training procedure

1. $\mathcal{L}oss = \mathcal{L}oss(C) - \alpha \cdot \mathcal{L}oss(A)$

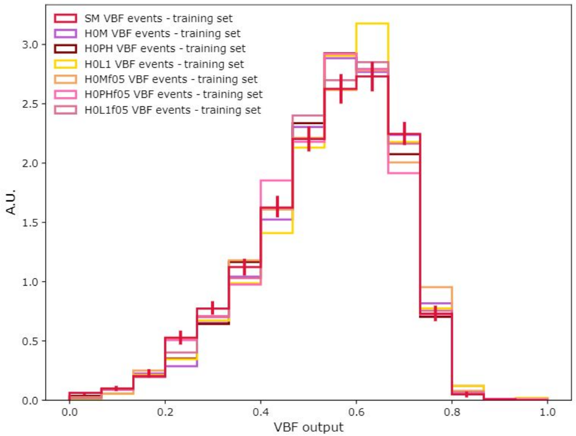
2. $\mathcal{L}oss(A)$

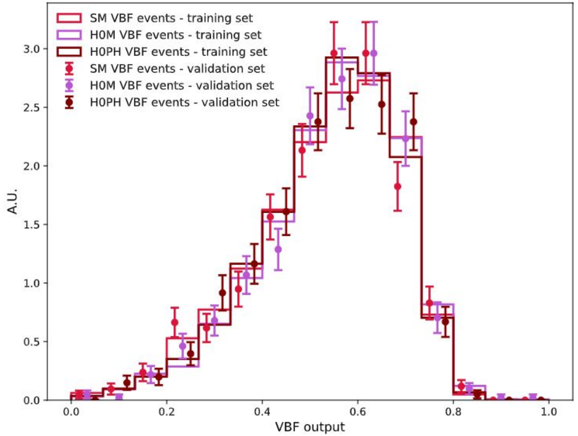
Compute the gradient of $\mathcal{L}(A)$ with respect to the A weights

Results



Training results

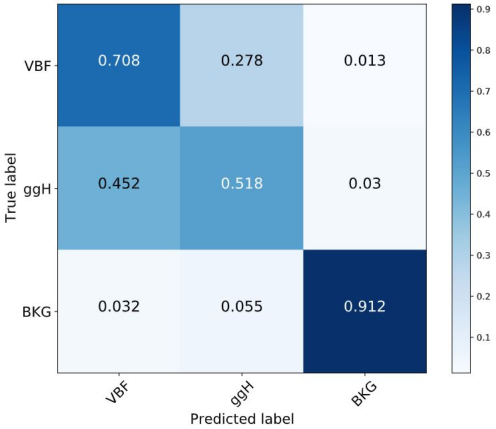




VBF-output distribution

✔ Shapes in fair agreement with each other

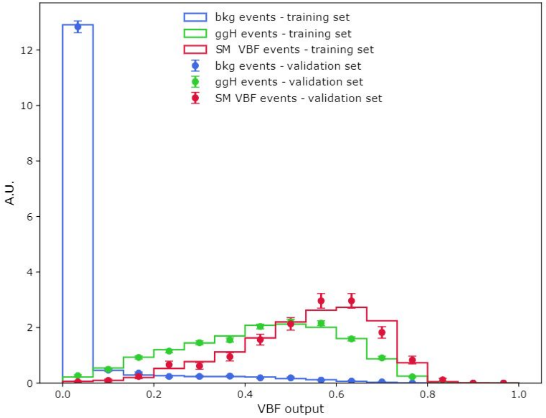
✔ The classifier is unable to recognize the domain of origin of signal events



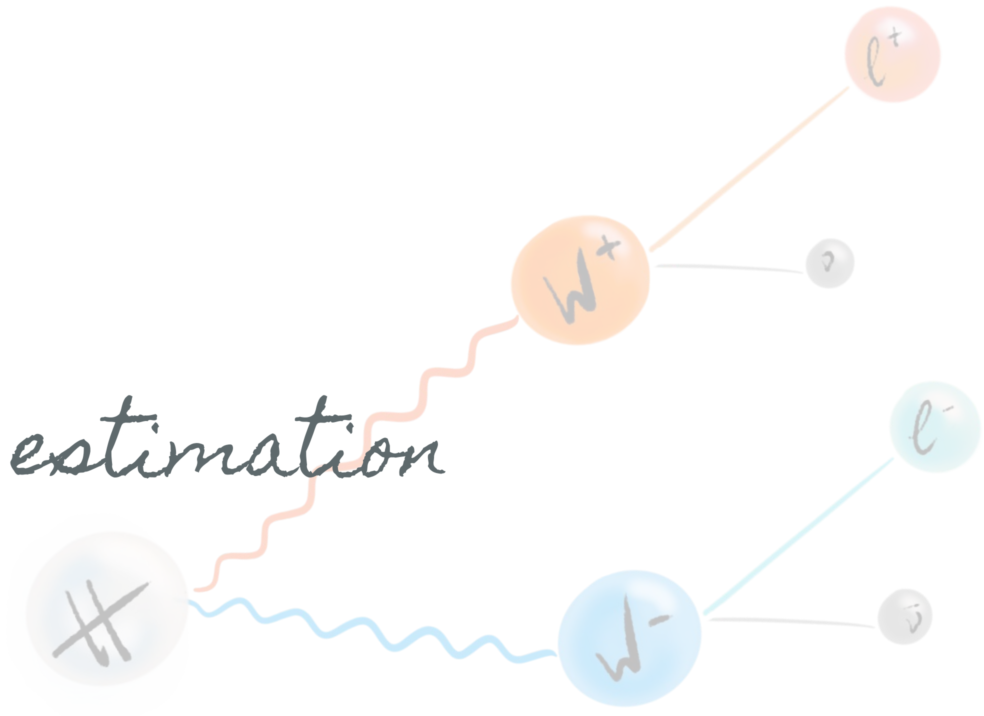
Discriminating power

✔ Good between bkg and VBF

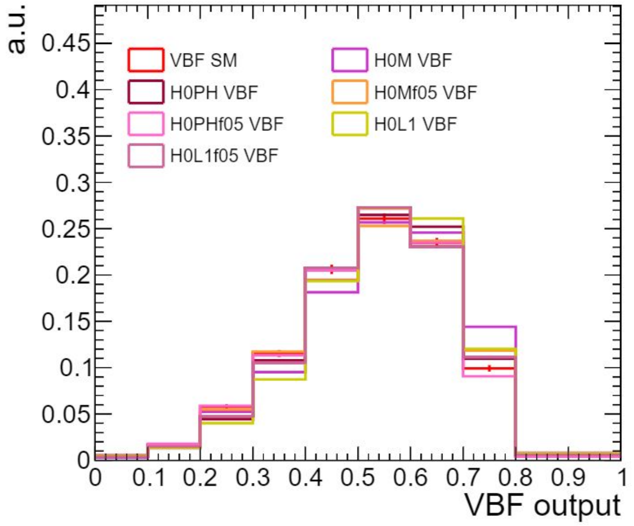
~ Poor between ggH and VBF



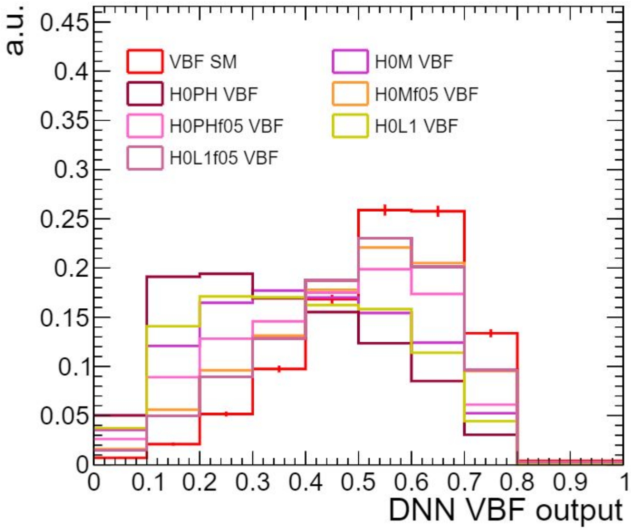
Bias estimation



ADNN vs DNN



accuracy ~ 69 %



accuracy ~ 72 %

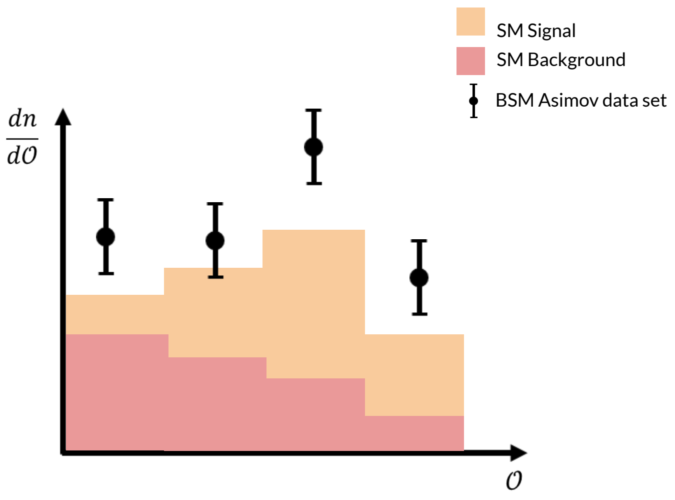
- In order to quantify the bias reduction made possible by the ADNN, a standard feed-forward DNN has been employed
- **Trained using only SM VBF events**
- The DNN structure has been optimized by maximizing the categorical accuracy

Best trial

- $\# \, hidden \ layers \ = 1$
- $\# \, nodes \ in \ each \ layer \ = 60$
- $learning \ rate \ = 0.0006$

Bias estimation

- VBF-output used as fit variable
- **Asimov data set considering a BSM signal hypothesis** and all SM background contributions
- **Fit using the SM signal** and background templates



Fit result

$$\hat{\mu} = \frac{s_{BSM}^{fit}}{s_{SM}^{exp}}$$

- number of measured BSM signal events
- number of expected SM signal events

Cross sections ratio

$$\tilde{\mu} = \frac{\sigma_{BSM}}{\sigma_{SM}}$$

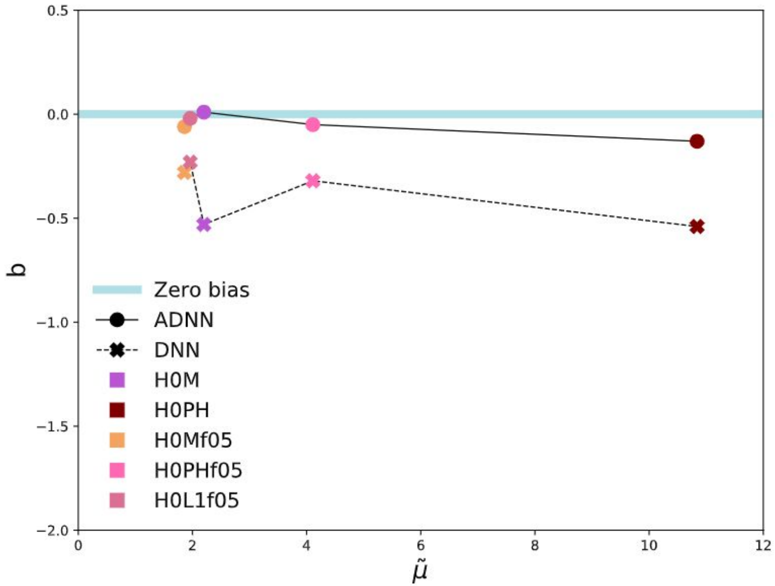
expected value if there is no bias due to the shape effect

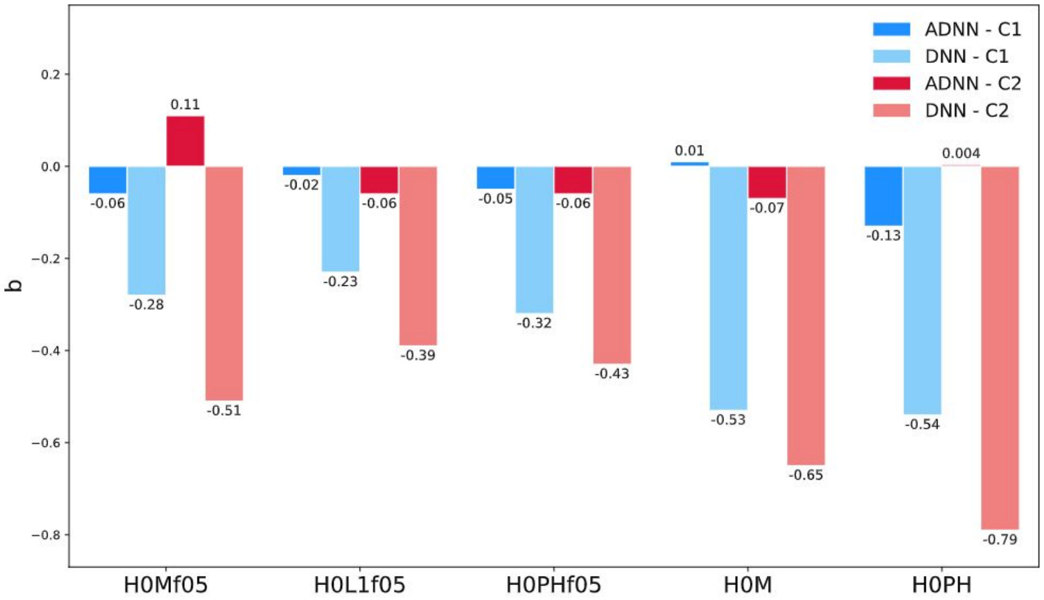
Total bias

$$b = \frac{\hat{\mu} - \tilde{\mu}}{\tilde{\mu}}$$

Bias results

- Strongly reduction of the biases when the ADNN
  - C1 is $350 < m_{\mathrm{jj}} \leq 700$ and $p_{\mathrm{T}}^{\mathrm{H}} < 200$ GeV
  - C2 is a different phase space region





- DNN bias ~20-80% of the expected value
- ADNN bias <10% of the expected value

Conclusions

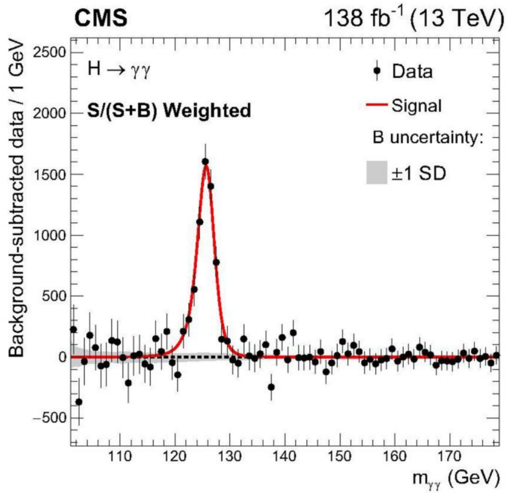
- **Performing model independent measurements is crucial to easily re-interpret results**
- **Domain Adaptation technique is a possible solution** to this problem, and we propose an implementation based on a two-networks system called **Adversarial Deep Neural Network** (ADNN)
- The **H→WW case study** has been presented, where the DA allowed to **significantly reduce the measurement bias due to the signal modeling assumptions**
- Due to its general character, this approach can be useful not only for Higgs boson measurements but also for all SM measurements

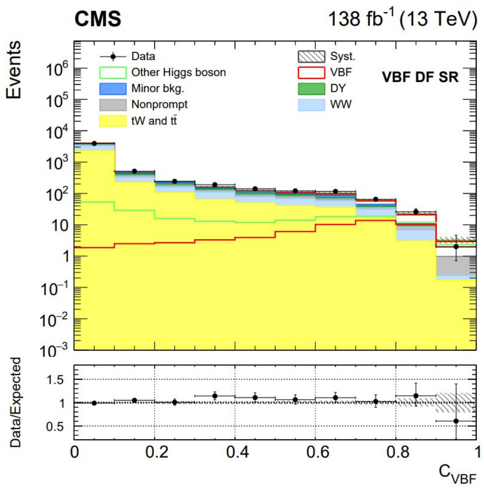
Thanks for your attention

Backup

H→γγ vs H→WW

- The Higgs boson invariant mass can be reconstructed in H→γγ
- **The discriminating variable is model independent**
- No bias in the fit procedure





- Due to neutrinos, the Higgs invariant mass can not be reconstructed in H→WW→2l2ν
- Deep Neural Network discriminant as fit variable
- Training set from Monte Carlo simulation
- **The shape of the discriminant depends on the physics hypothesis used to generate the training set**

Introduction to domain adaptation

- **Domain Adaptation** (DA) is a particular case of transfer learning (TL) whose goal is to apply an algorithm trained in one or more *source domains* to a different *target domain*.

Some definitions

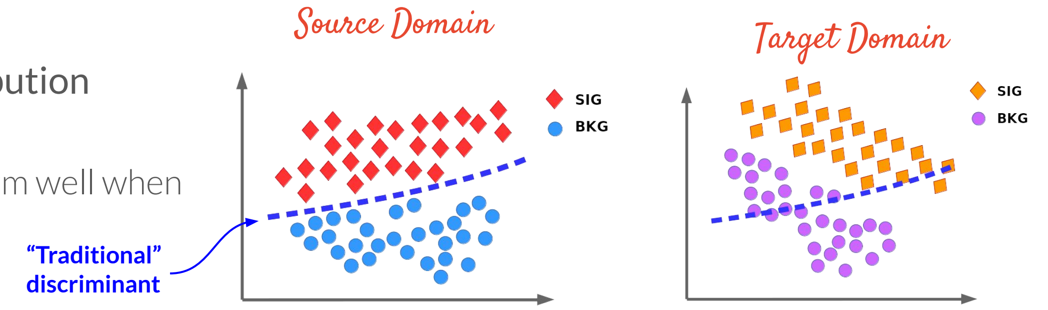
★ **Domain**: a feature space $\boldsymbol{X}$ + marginal probability distribution P(x), with $x = \{x_1, ..., x_n\}$ in $\boldsymbol{X}$

★ **Task**: a label space Y and a function f: $\boldsymbol{X}$->Y used to predicted the label y given the input x

- *Source* and *Target* domains are assumed to be related → distributions of source and target data are not completely different

★ **Domain shift**: change in the data distribution between training and deployment

  - Common AI algorithms do not perform well when domain shifts are present



- Different ways to achieve DA, unsupervised or (semi-)supervised

Supervised DA

- Supervised learning:
  - X input space, Y output space (or label space)
  - $(x_i, y_i) \in S$ i.i.d. from a distribution D_S (unknown and fixed) of support $X \times Y$
  - objective: learn $h: X \rightarrow Y$ from S such that it commits the least error possible for labelling new examples coming from D_S
- **Supervised domain adaptation**:
  - two different (but related) distributions D_S (source domain) and D_T (target domain) on $X \times Y$
  - objective: learn h from the two domains such that it commits as little error as possible on the target domain D_T

- main idea : find a representation space that is common to source and target domains
- focus on an adversarial deep learning approach (**Adversarial Deep Neural Network**)

Anomalous couplings in HVV

Scattering amplitude of one spin-0 Higgs boson (H) and two spin-1 gauge bosons (V_1 V_2)

$$A(HVV) \sim \left[\boxed{a_1^{VV}} + \underbrace{\boxed{\frac{k_1^{VV} q_{V1}^2 + k_2^{VV} q_{V2}^2}{(\Lambda_1^{VV})^2}}}_{\mathbf{L_1}} \right] m_{V1}^2 \epsilon_{V1}^* \epsilon_{V2}^* + \boxed{a_2^{VV} f_{\mu\nu}^{*(1)} f^{*(2)\mu\nu}} + \boxed{a_3^{VV} f_{\mu\nu}^{*(1)} \tilde{f}^{*(2)\mu\nu}}$$

This general structure has 4 couplings:

- $a_1^{VV} \neq 0$ SM couplings $J^{CP}=0^{++}$

AC:

- $L_1 \neq 0$ $H-Vff$ or $H-ffff$ couplings predicted by **H0L1** model
- $a_2^{VV} \neq 0$ loop-induced ($HZ\gamma$, $H\gamma\gamma$, Hgg) CP-even coupling predicted by **H0PH** model
- $a_3^{VV} \neq 0$ three loop induced CP-odd coupling predicted by **H0M** model

+ **H0L1f05**, **H0PHf05** and **H0Mf05** theories which are mixtures between the SM and one of the previous model

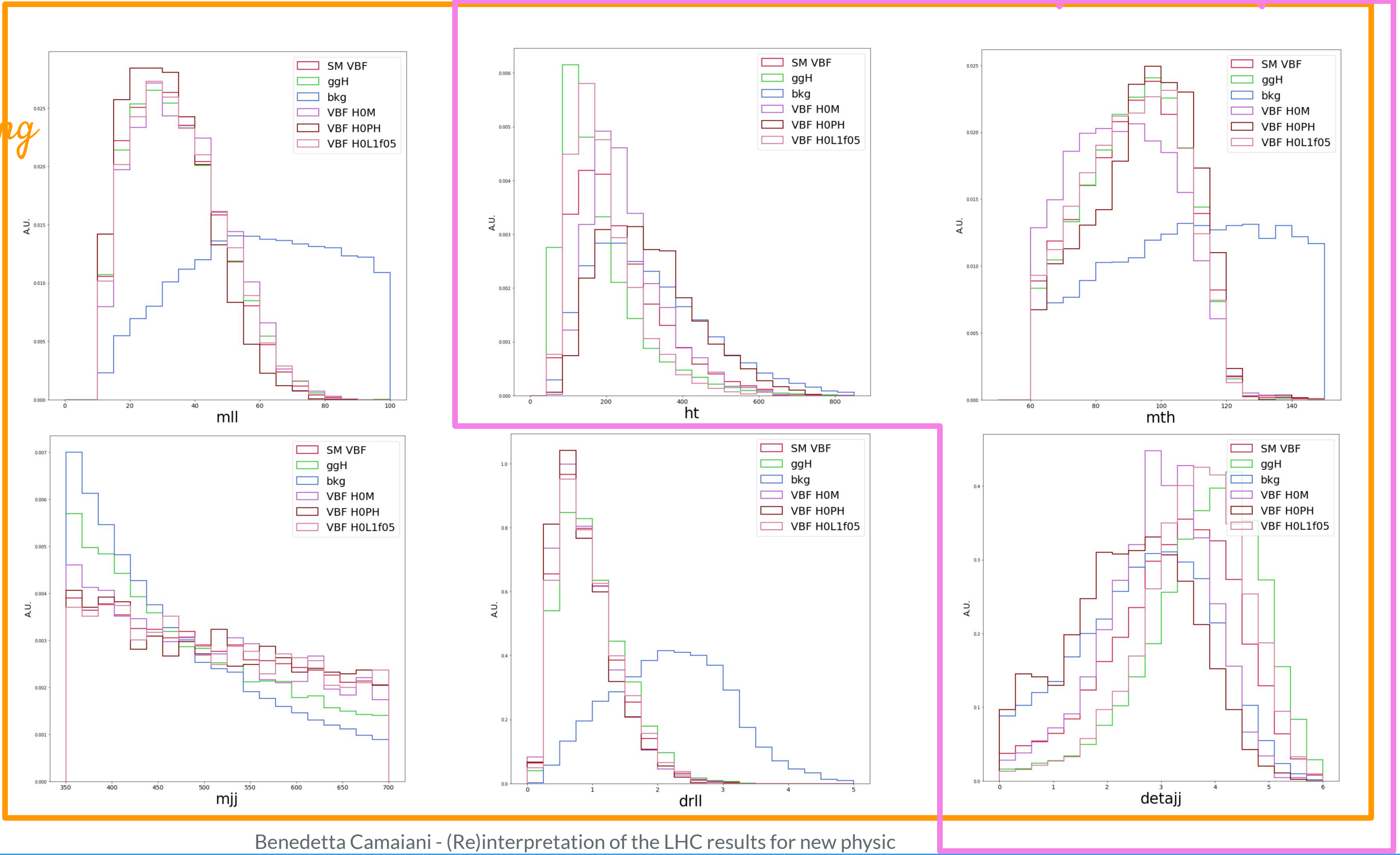
Analysis strategy

Global selection	
oppositely-charged eμ final state, at least two jets with $p_T > 30$ GeV, $p_T^{\ell_1} > 25$ GeV, $p_T^{\ell_2} > 13$ GeV, $p_T^{\ell_3} < 10$ GeV, $m_{\ell\ell} > 12$ GeV, $p_T^{\ell\ell} > 30$ GeV, $E_T^{miss} > 20$ GeV, $m_T^H > 60$ GeV, $m_T^{\ell_2} > 30$ GeV, $\lvert\eta_{j_1}\rvert < 4.7$, $\lvert\eta_{j_2}\rvert < 4.7$	
C1	**C2**
$350 < m_{jj} \leq 700$ GeV, $p_T^H < 200$ GeV, $\lvert y_H\rvert < 2.5$	$m_{jj} > 700$ GeV, $p_T^H < 200$ GeV, $\lvert y_H\rvert < 2.5$ or $m_{jj} > 350$ GeV, $p_T^H > 200$ GeV, $\lvert y_H\rvert < 2.5$

Some input features

strongly dependent on theoretical signal modeling

highly discriminating between signal and backgrounds



Optimization

- Two-step optimization using *Optuna* by:
  - Maximizing the **categorical accuracy** of the classifier
  - Minimizing the average of the **two-sample Kolmogorov-Smirnov** (K-S) **test statistic** between the distributions of the VBF-output shapes of signal events simulated under the SM and each of the considered BSM hypotheses (“average K-S test statistic”)
- First step: optimization of n^C_{layers}, η^C, n^A_{layers}, η^A, n_{nodes} and α
- Second step: optimization of η^C, η^A and α, keeping n^C_{layers}, n^A_{layers}, n_{nodes} fixed to their best value determinated in the first step
- Each step composed of 100 training trials of 800 epochs each
- Hyperparameters values varied according to a Bayesian optimization approach

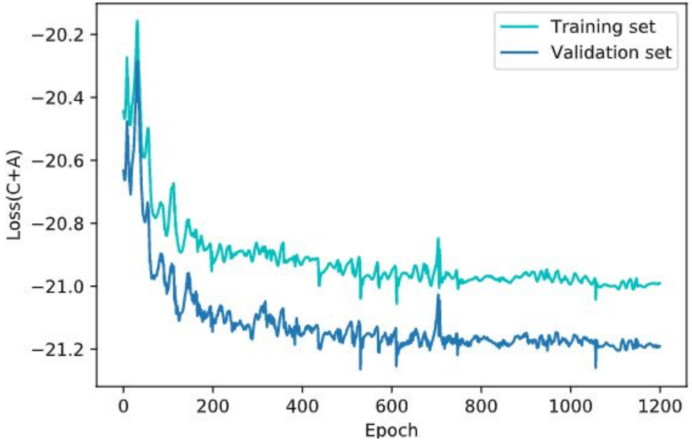
Optimization - result of the second step

Hyperpatameter	C1
α	100
η^{C}	0.00098
η^{A}	0.006
n_{nodes}	48
n_1^{C}	4
n_1^{A}	9

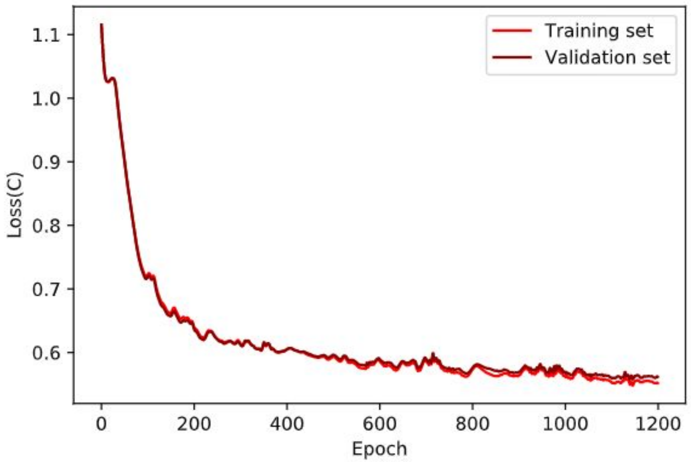
Objective function	
categorical accuracy	69%
average K-S test statistic	0.05

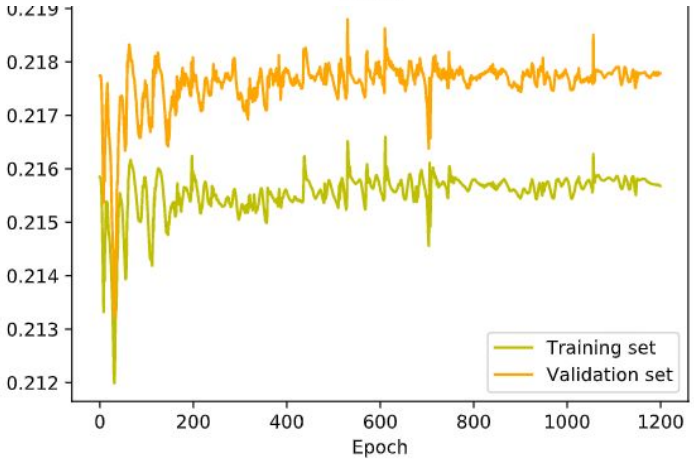
Training results (1)

- Loss function



$$\mathcal{L}oss = \mathcal{L}oss(C) - \alpha \cdot \mathcal{L}oss(A)$$

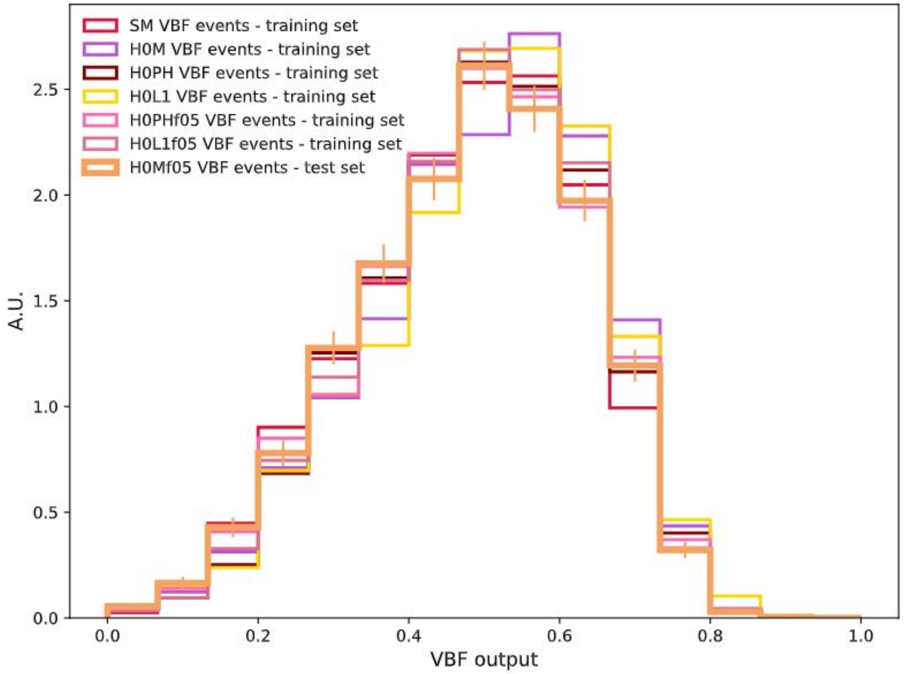




- L(C) decreases as usual

- L(A) saturates to a constant value, meaning that the performance of A is equivalent to random guessing

Agnosticism against unseen model



- The model chosen by Nature is unknown and in general may be different from the ones the discriminator has been trained on (unknown mixture between CP-even and CP-odd couplings)
- **An ADNN has been trained excluding one of the mixed models** (ticker line) and compared the discriminator shape between the models used in the training and the excluded one
- Having trained against the pure models is sufficient for the discriminator to be agnostic against the mixed model

Performance

- **Asimov data set corresponding to the assumption of SM cross sections**
- **Fit using the SM signal and background templates**

- μ is equal to 1 by construction
- Not significant deterioration on the total total errors of μ when the ADNN is used
- DNN and ADNN have almost the same discriminating power

	DNN	ADNN
Category	Total uncertainty	Total uncertainty
C1	$-0.38/+0.55$	$-0.39/+0.56$

Bias values

- $\sim$ expected value
- ^ fit result

H0PHf05					
		DNN		ADNN	
Category	$\tilde{\mu}$	$\hat{\mu}$	b	$\hat{\mu}$	b
C1	4.11	1.34	−0.32	3.91	−0.05
C2	5.73	1.77	−0.43	5.39	−0.06

H0Mf05					
		DNN		ADNN	
Category	$\tilde{\mu}$	$\hat{\mu}$	b	$\hat{\mu}$	b
C1	1.86	1.34	−0.28	1.75	−0.06
C2	3.58	2.77	−0.51	3.17	0.11

H0M					
		DNN		ADNN	
Category	$\tilde{\mu}$	$\hat{\mu}$	b	$\hat{\mu}$	b
C1	2.20	1.04	−0.53	2.22	0.01
C2	7.71	2.69	−0.65	7.13	−0.07

H0L1f05					
		DNN		ADNN	
Category	$\tilde{\mu}$	$\hat{\mu}$	b	$\hat{\mu}$	b
C1	1.96	1.52	−0.23	1.92	−0.02
C2	4.02	2.46	−0.39	3.76	−0.06

H0PH					
		DNN		ADNN	
Category	$\tilde{\mu}$	$\hat{\mu}$	b	$\hat{\mu}$	b
C1	10.84	4.98	−0.54	9.40	−0.13
C2	27.64	5.82	−0.79	27.76	0.004