

Continuous signal modelling in a multidimensional space of coupling parameters

Lydia Brenner on behalf of the ATLAS collaboration*

Deutsches Elektronen-Synchrotron (DESY)

E-mail: lydia.brenner@cern.ch

The plans for the second Run of the LHC changes the focus in the Higgs sector from searches to precision measurements. Effective Lagrangians can be used for parameterisation. A signal morphing method is developed to take all parameters into account simultaneously and model interference effects. It provides a continuous description of arbitrary physical signal observables such as cross sections or differential distributions in a multidimensional space of coupling parameters. This method is capable of morphing signal distributions and rates based on a minimal orthogonal set of independent base samples and therefore allows to directly fit the coupling parameters that describe the Standard Model and possible non-Standard Model interactions for, for example, the Higgs boson.

*XIII Quark Confinement and the Hadron Spectrum - Confinement2018
31 July - 6 August 2018
Maynooth University, Ireland*

*Speaker.

1. Introduction

The properties of the newly discovered Higgs boson have been extensively probed by the ATLAS [1] and CMS [2] experiments using LHC Run 1 proton-proton collision data at $\sqrt{s} = 7$ and 8 TeV [3–6]. The studies of the tensor structure of the Higgs boson couplings to gauge bosons were based on signal models including at most one or two Beyond the Standard Model (BSM) coupling parameters at a time, with all remaining BSM parameters set to zero. For Run 2, it is envisioned to have signal models which depend on a larger number of coupling parameters, in order to account for possible correlations among them. Additional coupling parameters in the Higgs boson coupling to the Standard Model (SM) particles change the predicted cross sections, as well as the shape of differential distributions. In this context, it is necessary to revise the existing signal modelling methods and provide alternatives which are better suited for such a multidimensional parameter space.

For this purpose, a morphing method has been developed and implemented. It provides a continuous description of arbitrary physical signal observables such as cross sections or differential distributions in a multidimensional space of coupling parameters. The morphing-based signal model is a linear combination of a minimal set of orthogonal base samples (templates) spanning the full coupling parameter space. The weight of each template is derived from the coupling parameters appearing in the signal matrix element.

Morphing is more than a simple interpolation technique, in that it is not limited to the points in the range spanned by the input samples. In fact, the choice of the input samples is arbitrary, and any set of input samples satisfying the required conditions to build the morphing function will span the entire space, independent of their precise coordinates.

A full explanation and validation of this method is shown in Ref. [7].

2. Morphing principles

The morphing procedure is based on the concepts of the morphing of (possibly multi-dimensional) histograms described in Ref. [8]. It is introduced to describe the dependence of a given physical observable T on an arbitrary configuration of a set of non-SM Higgs boson couplings $\vec{g}_{\text{target}} \equiv \{g_{\text{SM}}, g_{\text{BSM},1}, \dots, g_{\text{BSM},n}\}$ to known particles. This dependence is described by a morphing function

$$T_{\text{out}}(\vec{g}_{\text{target}}) = \sum_i w_i(\vec{g}_{\text{target}}; \vec{g}_i) T_{\text{in}}(\vec{g}_i), \quad (2.1)$$

which linearly combines the values or differential distributions T_{in} at a number of selected discrete coupling configurations $\vec{g}_i = \{\tilde{g}_{\text{SM},i}, \tilde{g}_{\text{BSM},1}, \dots, \tilde{g}_{\text{BSM},n}\}$. The input distributions T_{in} are normalised to their expected cross sections such that T_{out} includes not only the correct shape, but also the correct cross section prediction. Here, g_{SM} denotes the Higgs boson coupling predicted by the Standard Model. Morphing only requires that any differential cross section can be expressed as a polynomial in coupling parameters. For calculation at lowest order and using the narrow-width approximation for a resonance, this yields a second order polynomial each in production and decay.

In practice, the template distributions T_{in} are obtained from the Monte Carlo (MC) simulation of the signal process for a given coupling configuration \vec{g}_i . The minimal number N of Monte Carlo

samples needed to describe the signal at all possible coupling configurations, depends on the number n of studied non-SM coupling parameters. The contribution of each sample T_{in} is weighted by a weight w_i based on the assumption that the value of a physical observable is proportional to the squared matrix element for the studied process

$$T \propto |\mathcal{M}|^2. \quad (2.2)$$

The weights w_i can therefore be expressed as functions of the coupling parameters in the matrix element \mathcal{M} . In this case T can be anything derived from the Matrix element, for example a whole MC sample.

The described procedure allows for a continuous description in an n -dimensional parameter space. A feature-complete implementation has been developed within the RooFit package [9], making use of HistFactory [10]. The provided signal model can therefore be used in commonly used RooFit workspaces in a straightforward, blackbox-like way. A visual representation of the idea for a simple case is shown in Figure 1.

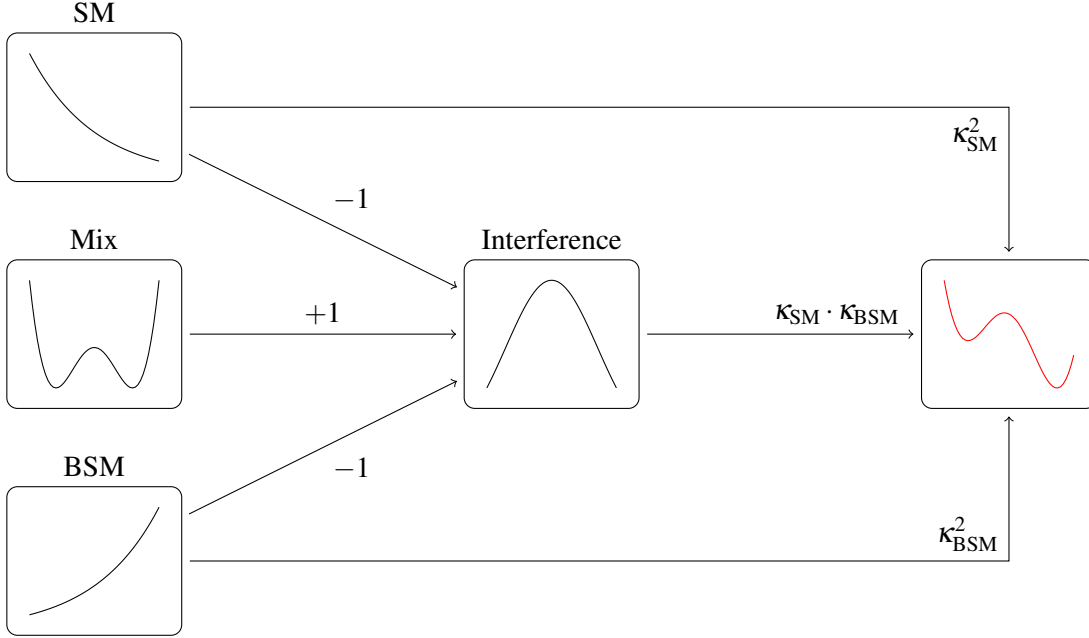


Figure 1: Illustration of the morphing procedure in a simple showcase.

3. General procedure to construct morphing function

A step-by-step explanation on how to construct the morphing function for processes with an arbitrary number of free coupling parameters in two vertices is outlined below.

1. Construct a general matrix element squared

$$|\mathcal{M}(\vec{g})|^2 = \underbrace{\left(\sum_{x \in p,s} g_x \mathcal{O}(g_x) \right)^2}_{\text{production}} \cdot \underbrace{\left(\sum_{x \in d,s} g_x \mathcal{O}(g_x) \right)^2}_{\text{decay}}, \quad (3.1)$$

denoting operators appearing only in the production vertex with p , such only appearing in the decay vertex with d , and such shared between both vertices with s , and assuming that production and decay vertices are uncorrelated, which is the case for a scalar intermediate particle.

2. Expand the matrix element squared to a 4th degree polynomial in the coupling parameters

$$|\mathcal{M}(\vec{g})|^2 = \sum_{i=1}^N X_i \cdot P_i(\vec{g}), \quad (3.2)$$

X_i is a prefactor, which will be represented by an input distribution. In the 4th degree polynomial $P_i(\vec{g}) = g_a g_b g_c g_d$ of the coupling parameters \vec{g} , the same coupling can occur multiple times (e.g. g_{SM}^4 or $g_{\text{BSM},1} g_{\text{BSM},2} g_{\text{BSM},3}^2$). The number of different expressions in the polynomial N is equal to the number of samples needed for the morphing.

3. Next generate input distributions at arbitrary but fixed parameter points \vec{g}_i

$$T_{\text{in},i} \propto |\mathcal{M}(\vec{g}_i)|^2. \quad (3.3)$$

4. Construct the morphing function with an ansatz

$$T_{\text{out}}(\vec{g}) = \sum_{i=1}^N \underbrace{\left(\sum_{j=1}^N A_{ij} P_j(\vec{g}) \right)}_{w_i(\vec{g})} T_{\text{in},i}. \quad (3.4)$$

$$= \vec{P}(\vec{g}) \cdot A \vec{T}, \quad (3.5)$$

where the second line is the first one recast in matrix notation. The matrix A has to be calculated to obtain the full morphing function.

5. Thus, exploit that the output distribution should be equal to the input distribution at the respective input parameters

$$T_{\text{out}}(\vec{g}_i) = T_{\text{in},i} \quad \text{for} \quad i = 1, \dots, N. \quad (3.6)$$

which can also be cast in matrix notation as

$$\begin{aligned} A \cdot (P_j(\vec{g}_i))_{ij} &= \mathbb{1} \\ \Leftrightarrow A \cdot G &= \mathbb{1}. \end{aligned} \quad (3.7)$$

6. The unique solution $A = G^{-1}$ requires the input parameters to fulfil the condition $\det(G) \neq 0$.

When the aim is to perform a likelihood fit on some (pseudo-)data T_d , the minimisation condition is

$$\hat{\vec{g}}(T_d) = \arg \min_{\vec{g}} -2 \ln P \left(T_d \mid \mu = \sum_{i=1}^N \left(\sum_{j=1}^N A_{ij} P_j(\vec{g}) \right) T_{\text{in},i} \right). \quad (3.8)$$

From this it becomes apparent that only the polynomials $P_j(\vec{g})$ need to be recalculated during the minimisation process, while the non-trivial quantities A_{ij} and $T_{in,i}$ stay fixed.

The error propagation of statistical uncertainties to the output T_{out} is conceptually straightforward. Since the \vec{g}_i are free parameters, the matrix A carries no uncertainty besides numerical fluctuations. Thus, uncertainties only propagate via linear combinations. The question of how the input parameters \vec{g}_i need to be chosen such that the expected uncertainty of the output is minimal, within some parameter region of interest, is non-trivial and will be addressed in future studies.

The number N of input base samples depends on how many of coupling parameters enter the production and/or the decay vertex. However, the general morphing principle remains the same and the method can be generalised to a higher-dimensional coupling parameter space.

A general expression for the number of input samples N with n_p couplings appearing only in production, n_d couplings appearing only in decay and n_s couplings shared in production and decay is given by

$$N = \frac{n_p(n_p+1)}{2} \cdot \frac{n_d(n_d+1)}{2} + \binom{4+n_s-1}{4} \quad (3.9)$$

$$+ \left(n_p \cdot n_s + \frac{n_s(n_s+1)}{2} \right) \cdot \frac{n_d(n_d+1)}{2} \quad (3.10)$$

$$+ \left(n_d \cdot n_s + \frac{n_s(n_s+1)}{2} \right) \cdot \frac{n_p(n_p+1)}{2} \quad (3.11)$$

$$+ \frac{n_s(n_s+1)}{2} \cdot n_p \cdot n_d + (n_p+n_d) \binom{3+n_s-1}{3}. \quad (3.12)$$

In this expression the counting is split for (3.9) terms pure in production and decay, or pure in shared, (3.10) terms pure in decay and mixed in production and shared or purely shared, (3.11) terms pure in production and mixed in decay and shared or purely shared, and (3.12) terms mixed in both, and terms mixed in one and purely shared in the other.

This is a general definition of the number of samples N in terms of number of coupling parameters n_p , n_d , and n_s . In case of the gluon fusion process with subsequent decays to vector bosons, the production and decay will have a completely disjoint set of couplings, and the number of input samples will be given by Eq. 3.9 by setting $n_s = 0$. For the VBF Higgs boson production with subsequent decay into vector bosons, when considering the same set of couplings in the production and the decay vertex, the number of samples is given by Eq. 3.9 with $n_p = 0$ and $n_d = 0$.

4. Example for VBF $H \rightarrow WW$

The Effective Lagrangian framework can for example be implemented in Higgs Characterisation model. The Effective Lagrangian for the interaction of scalar and pseudo-scalar states with vector bosons in this model becomes [11],

$$\begin{aligned}
\mathcal{L}_0^V = & \left\{ c_\alpha \kappa_{SM} \left[\frac{1}{2} \tilde{g}_{HZZ} Z_\mu Z^\mu + \tilde{g}_{HWW} W_\mu^+ W^{-\mu} \right] \right. \\
& - \frac{1}{4} \left[c_\alpha \kappa_{H\gamma\gamma} \tilde{g}_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} \tilde{g}_{A\gamma\gamma} A_{\mu\nu} \tilde{A}^{\mu\nu} \right] \\
& - \frac{1}{2} \left[c_\alpha \kappa_{HZ\gamma} \tilde{g}_{HZ\gamma} Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} \tilde{g}_{AZ\gamma} Z_{\mu\nu} \tilde{A}^{\mu\nu} \right] \\
& - \frac{1}{4} \left[c_\alpha \kappa_{Hgg} \tilde{g}_{Hgg} G_{\mu\nu}^a G^{a,\mu\nu} + s_\alpha \kappa_{Agg} \tilde{g}_{Agg} G_{\mu\nu}^a \tilde{G}^{a,\mu\nu} \right] \\
& - \frac{1}{4} \frac{1}{\Lambda} \left[c_\alpha \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ} Z_{\mu\nu} \tilde{Z}^{\mu\nu} \right] \\
& - \frac{1}{2} \frac{1}{\Lambda} \left[c_\alpha \kappa_{HWW} W_{\mu\nu}^+ W^{-\mu\nu} + s_\alpha \kappa_{AWW} W_{\mu\nu}^+ \tilde{W}^{-\mu\nu} \right] \\
& \left. - \frac{1}{\Lambda} c_\alpha \left[\kappa_{H\partial\gamma} Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z} Z_\nu \partial_\mu Z^{\mu\nu} + \kappa_{H\partial W} (W_\nu^+ \partial_\mu W^{-\mu\nu} + h.c.) \right] \right\} \mathcal{X}_0
\end{aligned} \tag{4.1}$$

For this example we consider VBF $H \rightarrow WW$ process with 1 SM (g_{SM}) and 2 BSM operators (g_{HWW} , g_{AWW}). In this case there are 15 template samples needed with different parameters. For each sample 50k events were generated. For the analysis only the signal has been considered and kinematic observable $\Delta\phi_{jj}$ has been used. Since only small deviations from the SM are expected, $g_{SM} = 1$, $\Lambda = 1 \text{ TeV}$ and $\cos \alpha = \frac{1}{\sqrt{2}}$ for all template samples. The BSM parameter limits are chosen such that $\sigma_{pureBSM} \sim \sigma_{SM}$ and all other BSM parameters set to 0. The distribution of the template and validation samples is shown in Figure 2. The VBF $H \rightarrow WW$ template distributions are shown in 3.

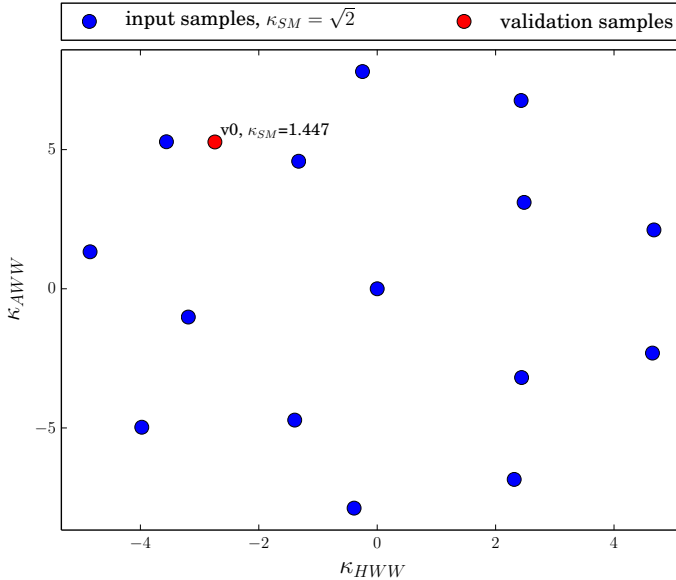


Figure 2: Scatter plot shows blue points in (g_{AWW}, g_{HWW}) space used to generate template samples and the red point is a validation sample produced for cross-checks.

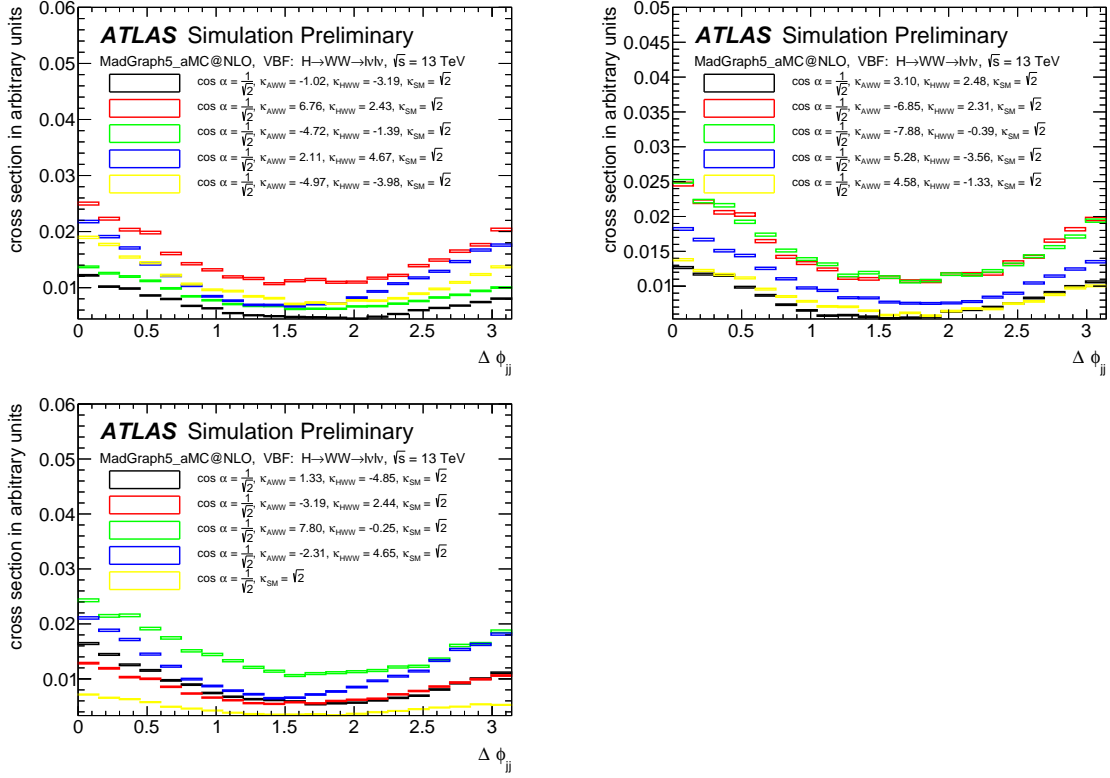


Figure 3: The 15 VBF H to WW template distributions used for the example. Each sample has 50,000 Monte Carlo events. The size of the boxes correspond to the Monte Carlo statistical uncertainties [7].

The distributions are used as pseudo-data for morphing and fitting to the SM template distribution and the validation distribution as shown in Figure 4. The MC statistical uncertainties are used and the template and morphed distributions are statistically independent. For the SM distribution a perfect agreement is found between the morphed distribution and the pseudo-data, as expected. For the validation distribution the morphed distribution agrees with the pseudo-data within the MC statistical uncertainties. The sensitivity on the measured parameters is inferred from the fit uncertainties.

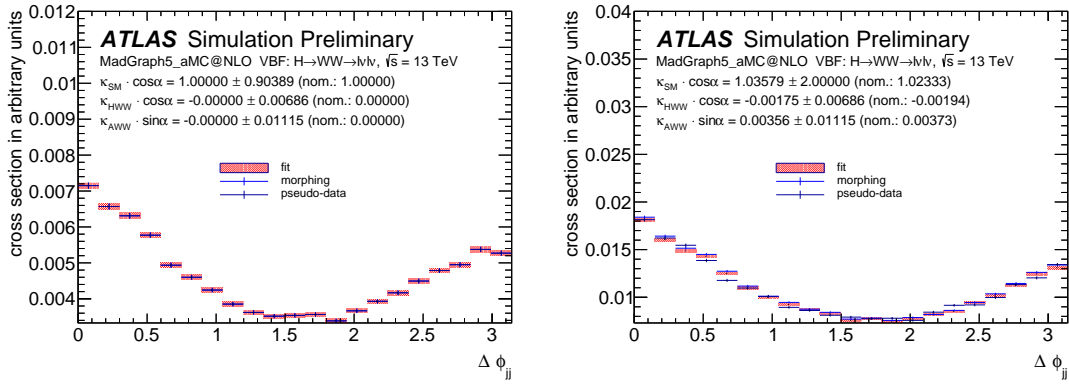


Figure 4: VBF H→WW morphing (blue) and fit (red) to SM template distribution (left) and validation distribution (right) [7].

5. Conclusions

This proceedings describes a method for modelling signal parameters and distributions in a multidimensional space of coupling parameters. This method is capable of continuously morphing signal distributions and rates based on a minimal orthogonal set of independent base samples. Therefore it allows to directly fit for the coupling parameters that describe the SM and possibly non-SM interaction of the Higgs boson with fermions and bosons of the SM.

This method can be utilised to test the properties of the Higgs boson during the LHC Run 2 data-taking period and beyond and has already been tested successfully [7, 12].

References

- [1] ATLAS collaboration, T. A. Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [2] CMS collaboration, T. C. Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [3] ATLAS Collaboration, *Study of the spin and parity of the higgs boson in diboson decays with the atlas detector*, *European Physical Journal C* **75** (2015) 476 [1506.05669].
- [4] CMS Collaboration, *Constraints on the spin-parity and anomalous h $\nu\nu$ couplings of the higgs boson in proton collisions at 7 and 8 TeV*, *Phys.Rev.D* **92** (2015) 012004 [1411.3441].
- [5] ATLAS and CMS Collaborations, *Combined measurement of the higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the atlas and cms experiments*, *Phys.Rev.Lett.* **114** (2015) 191803 [1503.07589].
- [6] ATLAS, CMS collaboration, ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **08** (2016) 045 [1606.02266].
- [7] *A morphing technique for signal modelling in a multidimensional space of coupling parameters*, ATL-PHYS-PUB-2015-047.
- [8] M. Baak, S. Gadatsch, R. Harrington and W. Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, *ArXiv Physics e-prints* (2014) [1410.7388].
- [9] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, *ArXiv Physics e-prints* (2003) [physics/0306116].
- [10] ROOT COLLABORATION collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016.
- [11] P. Artoisenet, P. de Aquino, F. Demartin, R. Frederix, S. Frixione, F. Maltoni et al., *A framework for higgs characterisation*, *ArXiv Physics e-prints* (2013) [1306.6464].
- [12] THE LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, *ArXiv Physics e-prints* (2016) [1610.07922].