Continuous signal modelling in a multidimensional space of coupling parameters

Lydia Brenner on behalf of the ATLAS collaboration

DESY

1 August 2018

Introduction

Plans for Run 2 and beyond

- Perform combined studies of **many (all) parameters** in the matrix element
- Take **all correlations** between different operators into account
- Use constraining power from **rate & shape information**
- Combine results from different channels

→ Challenge: **large parameter space**

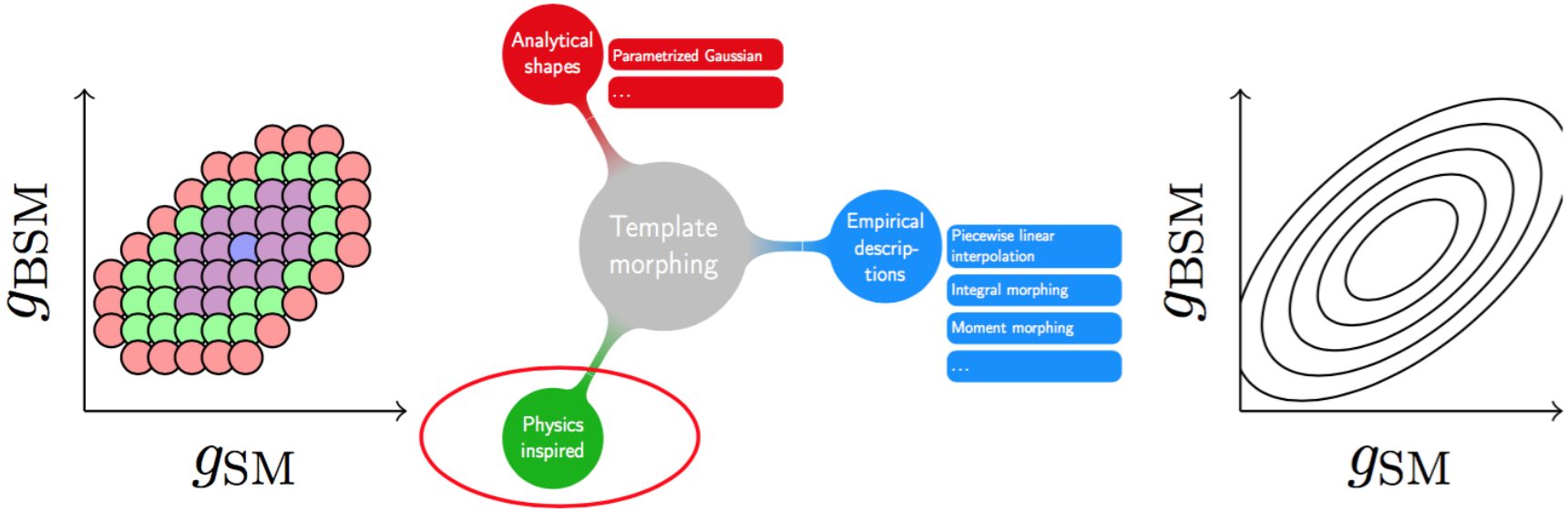
→ For properties necessary to build a signal model taking **all parameters** into account **simultaneously** & modelling all interference effects → **Morphing** PubNote Link: **atl-phys-pub-2015-047**

Introduction

- Start with the formulation of a Likelihood $L(\vec{x}|\vec{\mu},\vec{\theta})$
- Predict observable distribution from a composite model:
  (B)SM HEP model $\times$ soft physics model $\times$ detector response $\times$ detector reconstruction

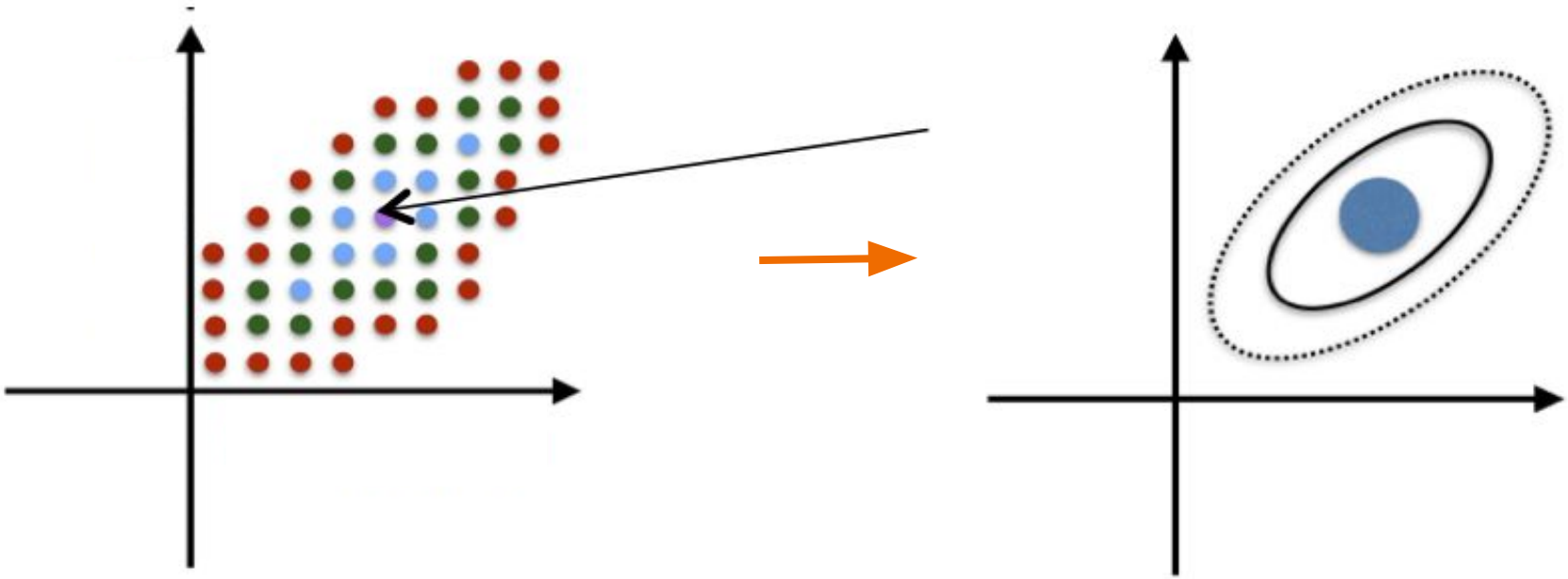
Problem: We do not have a continues description of $L(\vec{x}|\vec{\mu},\vec{\theta})$

- Can only calculate $L(x)$ for each point $\vec{\mu},\vec{\theta}$



Introduction

Morphing The procedure to turn a collection of points into a continuous function

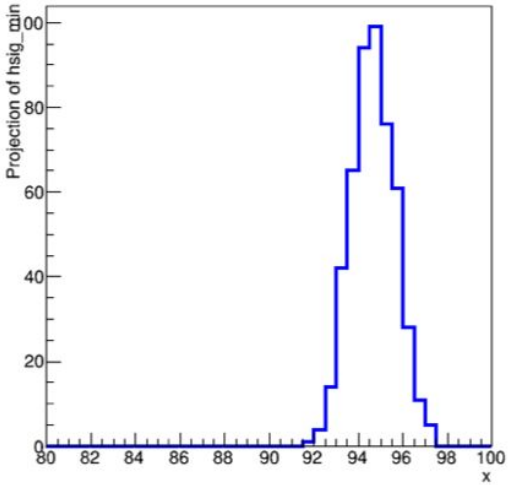


Interpolating between models

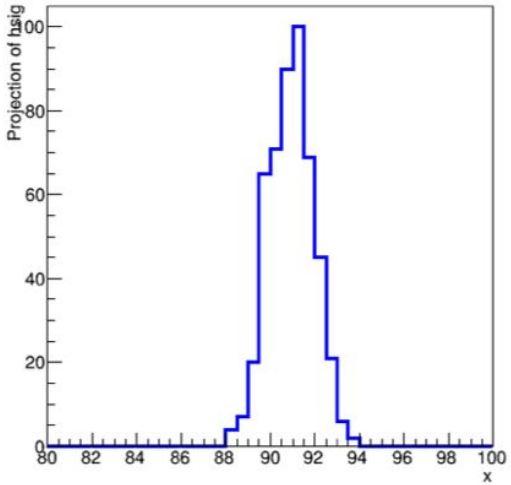
Need to define a morphing algorithm to define the number of signal events $s(x)$ for any value of another parameter a

We only know $s(x)$ for $a = -1, 0, 1$

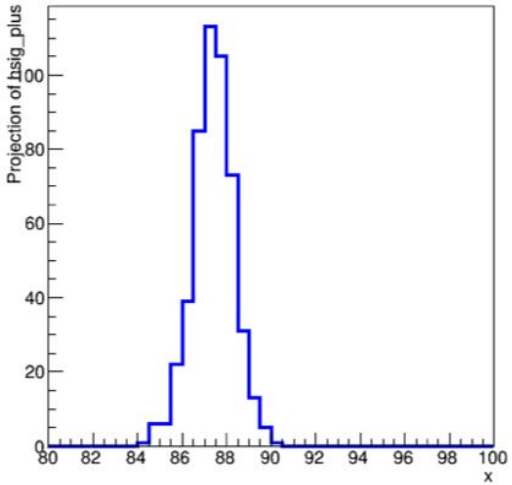
s(x)|a=-1



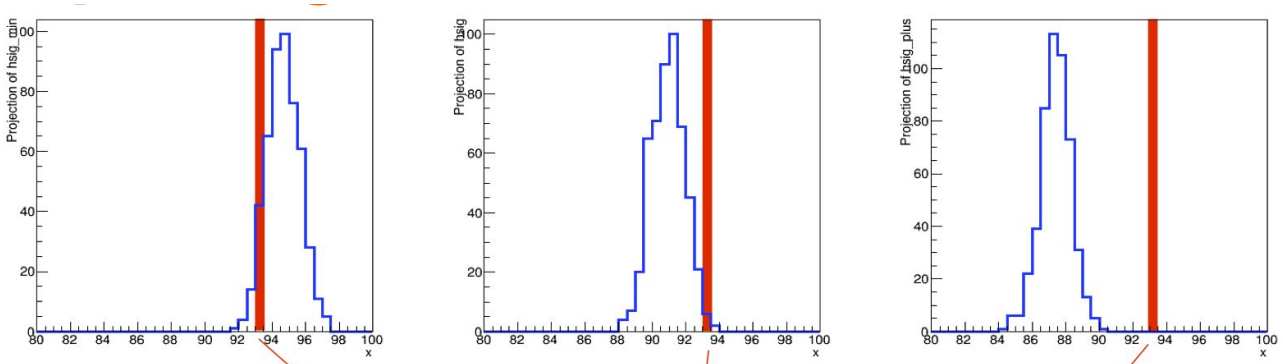
s(x)|a=0



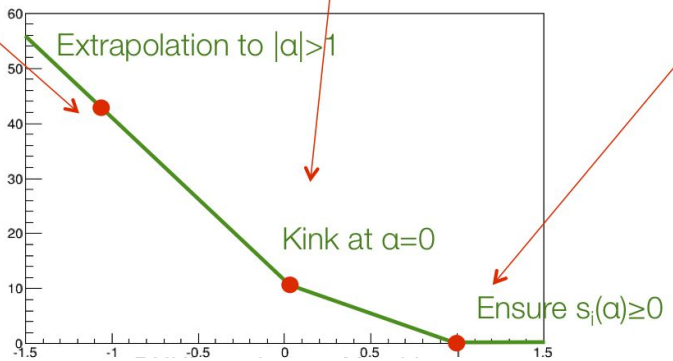
s(x)|a=1



Interpolating between models



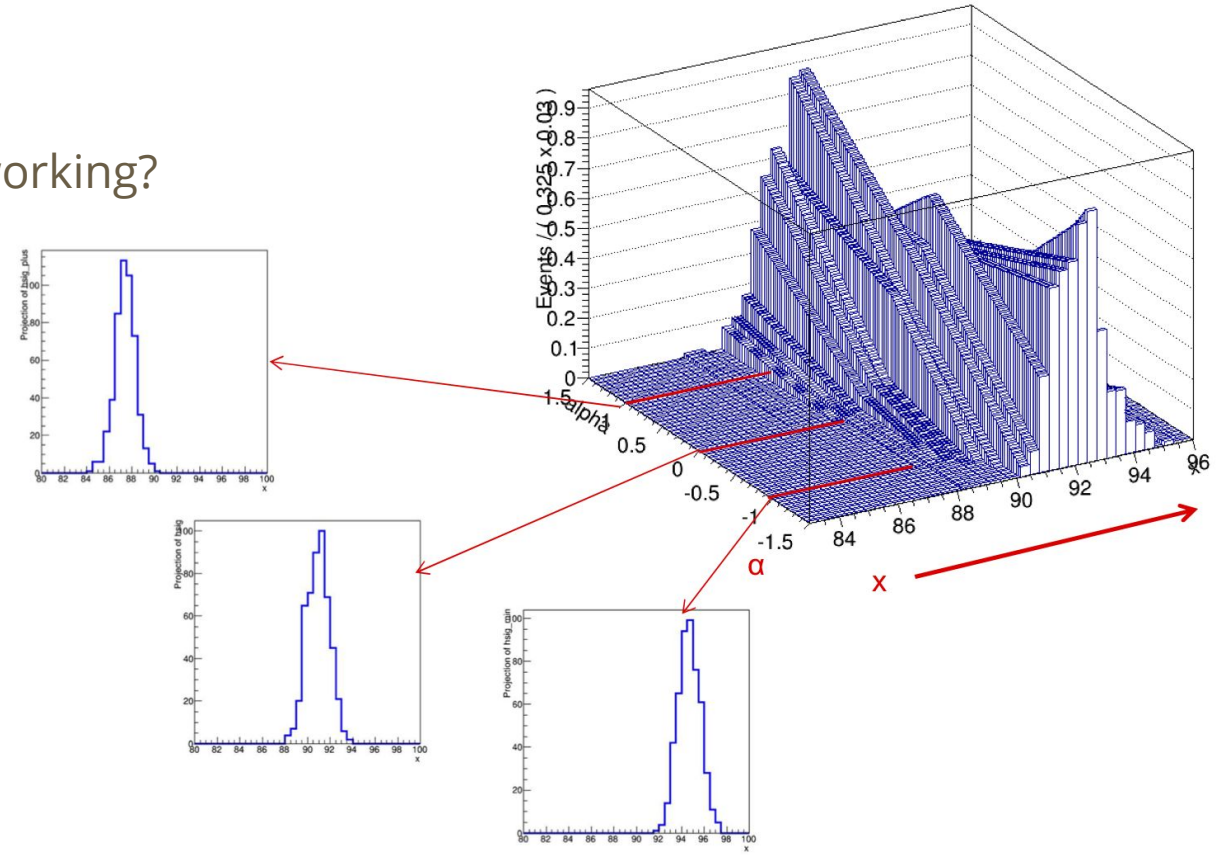
Simplest solution is piecewise linear interpolation



Interpolates response model bin by bin

Linear interpolation

When does this stop working?

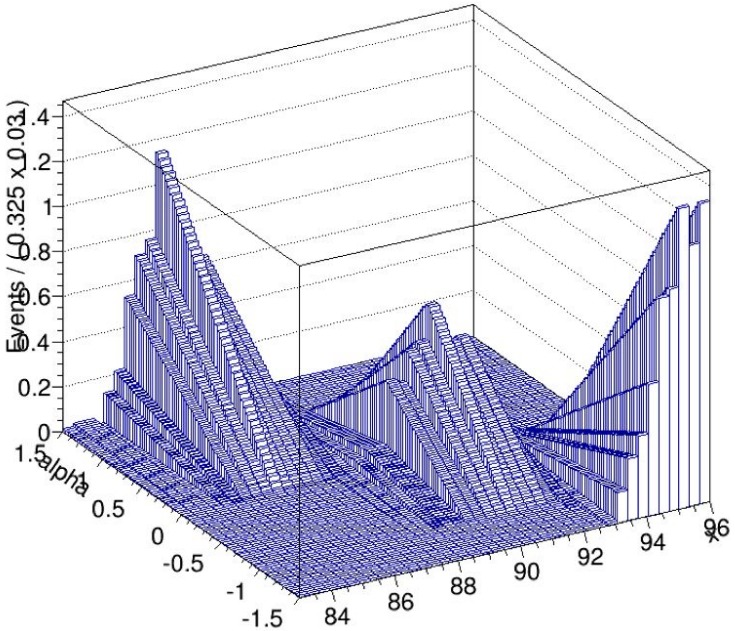


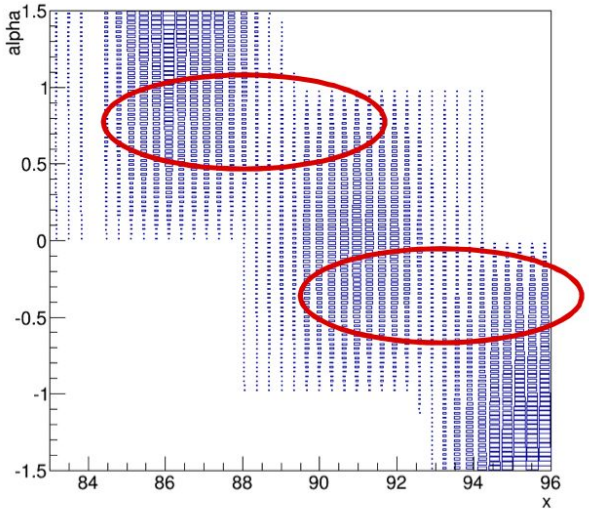
Linear interpolation

When does this stop working?

Example:

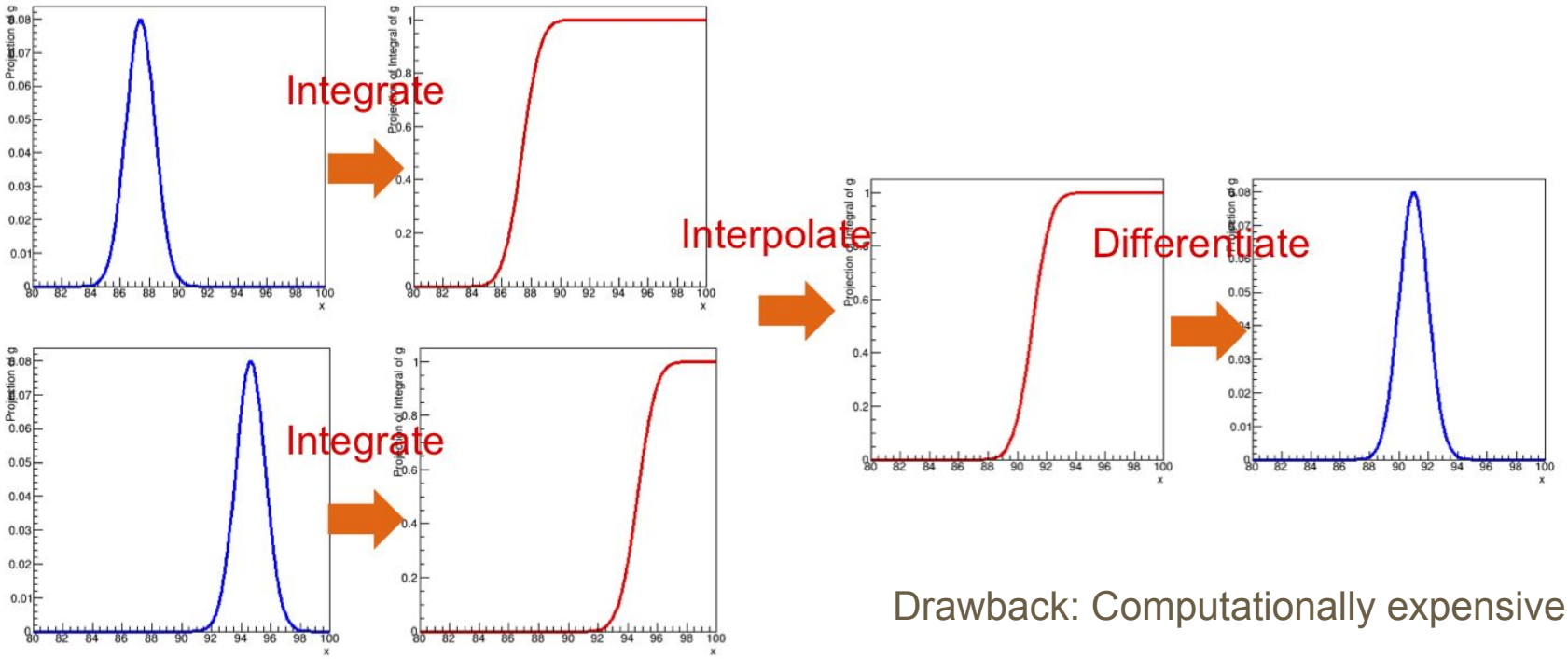
Large shift





Horizontal interpolation

Interpolate the cumulative distribution function



Drawback: Computationally expensive

Moment morphing

Constructs a morphed interpolated function that has linearly interpolated moments

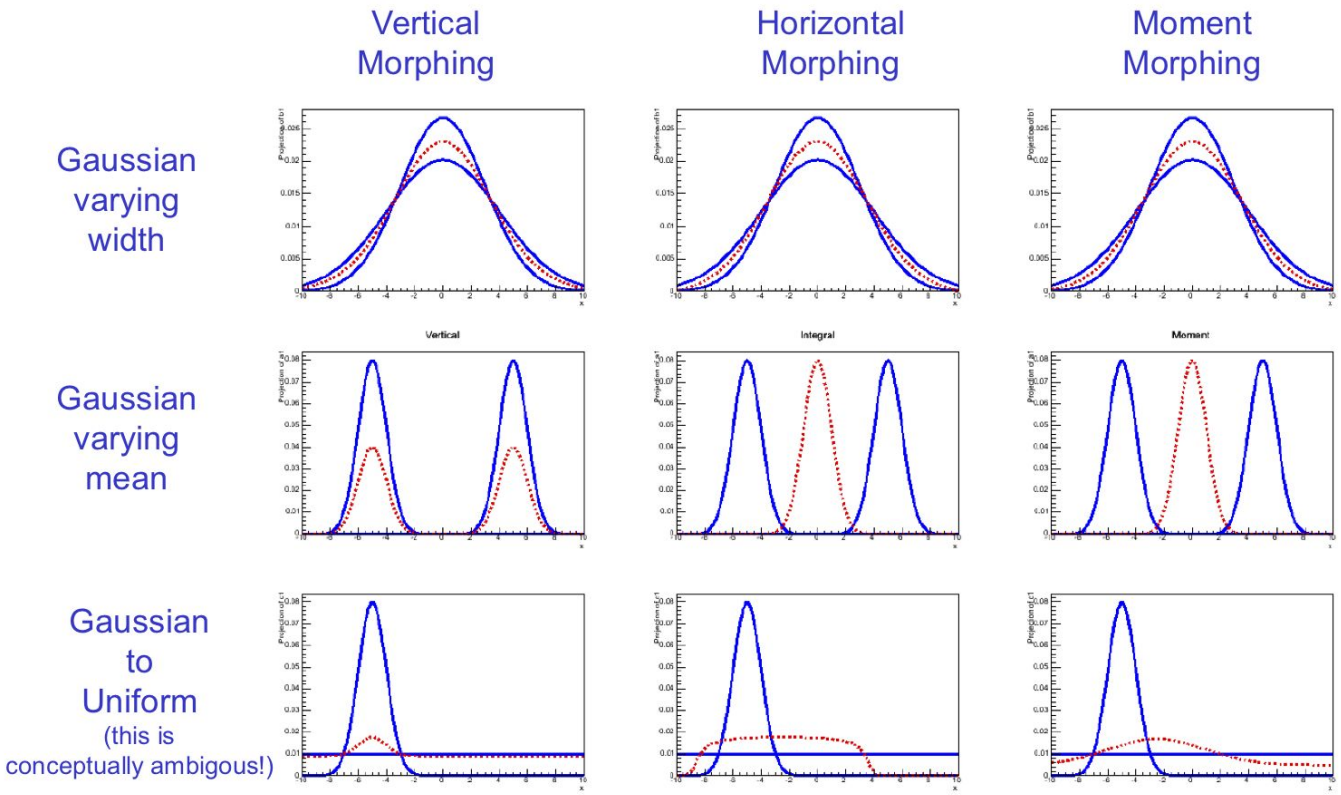
- First two moments of template models are the mean and variance

Multidimensional interpolation option
Computationally expensive, but only once

ARXIV:1410.7388

Comparison of methods

Different ways to create a continuous distribution of the likelihood



Can we use physics instead of an empirical procedure for signal morphing? Yes!
The Lagrangian of the physics model includes the dependence on signal parameters
Templates correspond to Matrix Element squared
→ Effective Lagrangian Morphing

Effective Lagrangian framework implemented in Higgs Characterisation model

- Effective Lagrangian for the interaction of scalar and pseudo-scalar states with vector bosons

$$
\begin{aligned}
\mathcal{L}_0^V = \Big\{ & c_\alpha \kappa_{SM} \left[\frac{1}{2} \tilde{g}_{HZZ} Z_\mu Z^\mu + \tilde{g}_{HWW} W^+_\mu W^{-\mu} \right] \\
& - \frac{1}{4} \left[c_\alpha \kappa_{H\gamma\gamma} \tilde{g}_{H\gamma\gamma} A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} \tilde{g}_{A\gamma\gamma} A_{\mu\nu} \tilde{A}^{\mu\nu} \right] \\
& - \frac{1}{2} \left[c_\alpha \kappa_{HZ\gamma} \tilde{g}_{HZ\gamma} Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} \tilde{g}_{AZ\gamma} Z_{\mu\nu} \tilde{A}^{\mu\nu} \right] \\
& - \frac{1}{4} \left[c_\alpha \kappa_{Hgg} \tilde{g}_{Hgg} G^a_{\mu\nu} G^{a,\mu\nu} + s_\alpha \kappa_{Agg} \tilde{g}_{Agg} G^a_{\mu\nu} \tilde{G}^{a,\mu\nu} \right] \\
& - \frac{1}{4} \frac{1}{\Lambda} \left[c_\alpha \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ} Z_{\mu\nu} \tilde{Z}^{\mu\nu} \right] \\
& - \frac{1}{2} \frac{1}{\Lambda} \left[c_\alpha \kappa_{HWW} W^+_{\mu\nu} W^{-\mu\nu} + s_\alpha \kappa_{AWW} W^+_{\mu\nu} \tilde{W}^{-\mu\nu} \right] \\
& - \frac{1}{\Lambda} c_\alpha \left[\kappa_{H\partial\gamma} Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z} Z_\nu \partial_\mu Z^{\mu\nu} + \kappa_{H\partial W} (W^+_\nu \partial_\mu W^{-\mu\nu} + h.c.) \right] \Big\} \chi_0
\end{aligned}
$$

Used in Run 1

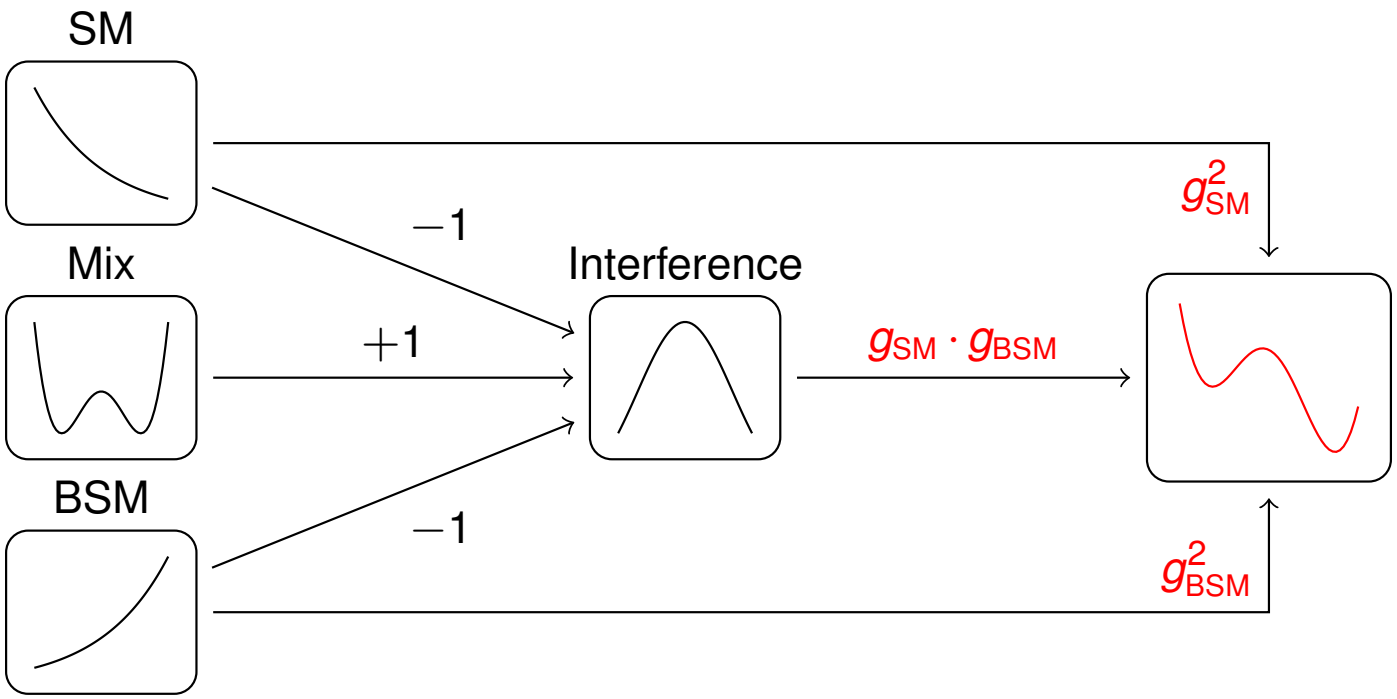
Plan Run 2

- **ARXIV:1306.6464**
- Implemented in MADGRAPH5_AMC@NLO

Signal model construction: Morphing

- **Morphing function** for an observable T_{out} at any coupling point $\vec{g}_{target}$ constructed from weighted sum of input samples T_{in} at fixed coupling points $\vec{g}_i$

$$T_{out}(\vec{g}_{target}) = \sum_{i=1}^{N_{input}} w_i(\vec{g}_{target}; \vec{g}_i) \cdot T_{in}(\vec{g}_i) \qquad \text{e.g. } T = \Delta\phi_{jj}$$



atl-phys-pub-2015-047

Example for 2 free parameters in one vertex

- Process with **two parameters** applied in **one vertex**: g_{SM} and g_{BSM}
- Matrix element can be **factorized**:

$$\mathcal{M}(g_{\mathrm{SM}}, g_{\mathrm{BSM}}) = g_{\mathrm{SM}}\mathcal{O}_{\mathrm{SM}} + g_{\mathrm{BSM}}\mathcal{O}_{\mathrm{BSM}}$$
$$|\mathcal{M}(g_{\mathrm{SM}}, g_{\mathrm{BSM}})|^2 = g_{\mathrm{SM}}^2|\mathcal{O}_{\mathrm{SM}}|^2 + g_{\mathrm{BSM}}^2|\mathcal{O}_{\mathrm{BSM}}|^2 + 2g_{\mathrm{SM}}g_{\mathrm{BSM}}\mathcal{R}(\mathcal{O}_{\mathrm{SM}}^*\mathcal{O}_{\mathrm{BSM}})$$

- **Distribution** of a kinematic observable **proportional to the matrix element squared**

$$T(g_{\mathrm{SM}}, g_{\mathrm{BSM}}) \propto |\mathcal{M}(g_{\mathrm{SM}}, g_{\mathrm{BSM}})|^2$$

- **3 generated distributions** needed to obtain distribution with arbitrary parameters
- E.g. generate MC events for $T(1,0)$, $T(0,1)$, $T(1,1)$

$$T_{in}(1,0) \propto |\mathcal{O}_{\mathrm{SM}}|^2$$
$$T_{in}(0,1) \propto |\mathcal{O}_{\mathrm{BSM}}|^2$$
$$T_{in}(1,1) \propto |\mathcal{O}_{\mathrm{SM}}|^2 + |\mathcal{O}_{\mathrm{BSM}}|^2 + 2\mathcal{R}(\mathcal{O}_{\mathrm{SM}}^*\mathcal{O}_{\mathrm{BSM}})$$

- Distribution with **arbitrary parameters** $(g_{\mathrm{SM}}, g_{\mathrm{BSM}})$

$$T_{out}(g_{\mathrm{SM}}, g_{\mathrm{BSM}}) = \underbrace{({g_{\mathrm{SM}}}^2 - g_{\mathrm{SM}}g_{\mathrm{BSM}})}_{=w_1} T_{in}(1,0) + \underbrace{({g_{\mathrm{BSM}}}^2 - g_{\mathrm{SM}}g_{\mathrm{BSM}})}_{=w_2} T_{in}(0,1) + \underbrace{g_{\mathrm{SM}}g_{\mathrm{BSM}}}_{=w_3} T_{in}(1,1)$$

Example for 2 free parameters in one vertex: generalisation of input parameter

- Generalize to **arbitrary input parameters** $\vec{g}_i$ used to generate input distributions $T_{in}(\vec{g}_i)$

$$T_{in}(g_{\mathrm{SM},i}, g_{\mathrm{BSM},i}) \propto {g_{\mathrm{SM},i}}^2 |\mathcal{O}_{\mathrm{SM}}|^2 + {g_{\mathrm{BSM},i}}^2 |\mathcal{O}_{\mathrm{BSM}}|^2 + 2 g_{\mathrm{SM},i} g_{\mathrm{BSM},i} \mathcal{R}(\mathcal{O}^*_{\mathrm{SM}} \mathcal{O}_{\mathrm{BSM}}), \qquad i = 1, \ldots 3$$

- Ansatz for **output distribution**

$$\begin{aligned} T_{out}(g_{\mathrm{SM}}, g_{\mathrm{BSM}}) = & \underbrace{\left(a_{11} {g_{\mathrm{SM}}}^2 + a_{12} {g_{\mathrm{BSM}}}^2 + a_{13} g_{\mathrm{SM}} g_{\mathrm{BSM}}\right)}_{w_1} T_{in}(g_{\mathrm{SM},1}, g_{\mathrm{BSM},1}) \\ & + \underbrace{\left(a_{21} {g_{\mathrm{SM}}}^2 + a_{22} {g_{\mathrm{BSM}}}^2 + a_{23} g_{\mathrm{SM}} g_{\mathrm{BSM}}\right)}_{w_2} T_{in}(g_{\mathrm{SM},2}, g_{\mathrm{BSM},2}) \\ & + \underbrace{\left(a_{31} {g_{\mathrm{SM}}}^2 + a_{32} {g_{\mathrm{BSM}}}^2 + a_{33} g_{\mathrm{SM}} g_{\mathrm{BSM}}\right)}_{w_3} T_{in}(g_{\mathrm{SM},3}, g_{\mathrm{BSM},3}) \end{aligned}$$

Example for 2 operators in one vertex

- T_{out} should be equal to T_{in} for $\vec{g}_{target} = \vec{g}_i$

$$1 = a_{11} {g_{\mathrm{SM},1}}^2 + a_{12} {g_{\mathrm{BSM},1}}^2 + a_{13} g_{\mathrm{SM},1} g_{\mathrm{BSM},1}$$
$$0 = a_{21} {g_{\mathrm{SM},1}}^2 + a_{22} {g_{\mathrm{BSM},1}}^2 + a_{23} g_{\mathrm{SM},1} g_{\mathrm{BSM},1}$$
$$\dots$$

- Constraints in **matrix form**

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} {g_{\mathrm{SM},1}}^2 & {g_{\mathrm{SM},2}}^2 & {g_{\mathrm{SM},3}}^2 \\ {g_{\mathrm{BSM},1}}^2 & {g_{\mathrm{BSM},2}}^2 & {g_{\mathrm{BSM},3}}^2 \\ g_{\mathrm{SM},1} g_{\mathrm{BSM},1} & g_{\mathrm{SM},2} g_{\mathrm{BSM},2} & g_{\mathrm{SM},3} g_{\mathrm{BSM},3} \end{pmatrix} = \mathbb{1}$$
$$\Leftrightarrow \qquad A \cdot G = \mathbb{1}$$

- **Definite solution** $A = G^{-1}$ requires the samples to have parameters such that $\det(G) \neq 0$
- Very flexible in choosing the parameters for the input distributions

→ Can be chosen to **reduce statistical uncertainty** in considered parameter space

General morphing and number of input templates

- More complicated when processes share amplitudes between **production and decay**, for example VBF $H \to VV$
- General matrix element squared at **LO** & assuming **narrow-width-approximation** (ignoring the effect on the total width)
  $\Rightarrow$ **polynomials** of 2nd order in production and 2nd order in decay

$$T(\vec{g}) \propto |\mathcal{M}(\vec{g})|^2 = \underbrace{\left(\sum_{i=1}^{n_p+n_s} g_i \mathcal{O}_i\right)^2}_{\text{production vertex}} \cdot \underbrace{\left(\sum_{j=1}^{n_d+n_s} g_j \mathcal{O}_j\right)^2}_{\text{decay vertex}}$$

with number of parameters in **production vertex** (n_p), **decay vertex** (n_d) and **shared in vertices** (n_s)

- **Number of required input templates** equal to number of different terms in expanded matrix element squared
  $\to$ dependent on process and considered parameters
  $\to$ N_{input} function of n_p, n_d and n_s
- Example: 13 free parameters for VBF H→ZZ process:
  - $n_p = 4$ operators in production: $g_{\mathrm{HWW}}, g_{\mathrm{AWW}}, g_{\mathrm{H}\partial\mathrm{W}}, g^*_{\mathrm{H}\partial\mathrm{W}}$
  - $n_s = 9$ operators in both vertices: $g_{\mathrm{SM}}, g_{\mathrm{HZZ}}, g_{\mathrm{AZZ}}, g_{\mathrm{H}\partial\mathrm{Z}}, g_{\mathrm{H}\gamma\gamma}, g_{\mathrm{A}\gamma\gamma}, g_{\mathrm{HZ}\gamma}, g_{\mathrm{AZ}\gamma}, g_{\mathrm{H}\partial\gamma}$
  - $n_d = 0$, no operators only in decay

$\to$ **1605 samples** needed!

Generality of the method

- Morphing only requires that any differential cross section can be expressed as **polynomial in BSM couplings**
- Method can be used on **any generator** that allows one to vary input couplings
- Works on **truth** and **reco-level** distributions
- **Independent of physics process**
- Works on distributions and cross sections

Comparison of methods

- **Needed:** MC samples covering wide range of values for coupling parameters
- Run 1 HWW and HZZ analyses: **Matrix Element Reweighting**
  (Event by event matrix element reweighting of one source MC sample with large statistics)

$$w(\vec{g}_{target}) = w(\vec{g}_i)\frac{|\mathcal{M}(\vec{g}_{target})|^2}{|\mathcal{M}(\vec{g}_{source})|^2}$$

ME Reweighting

For every configuration point

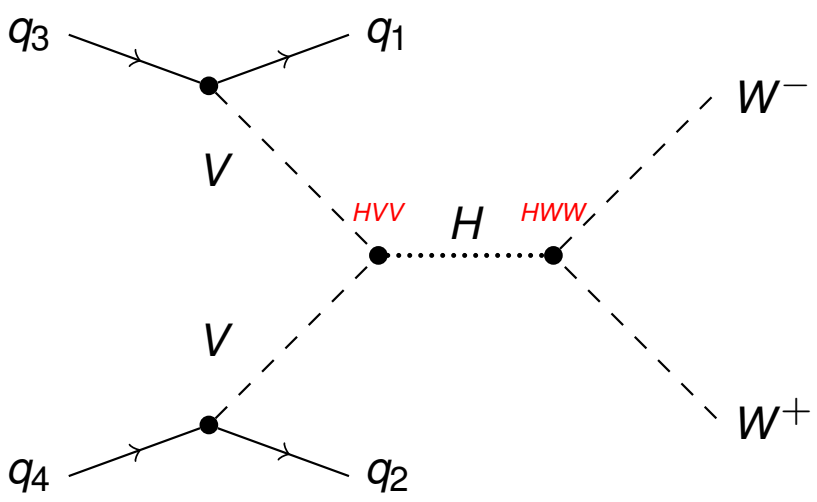
- rerun analysis
- write event weights to disk
- additional interpolation

Morphing

- only calculates linear sums of coefficients
- all other inputs are pre-computed once
- computationally fast & convenient tool

- **Morphing function:** Instead of "matrix element reweighting" use morphing to obtain a distribution with arbitrary coupling parameters
- Can be applied directly and without change to
  - Cross sections
  - Distributions (before or after detector simulation)
  - MC events
- **Exact continuous analytical description of rates and shapes**
- Even possible to **fit** coupling parameters to data & derive limits

VBF H→WW example



- VBF H→WW process with **SM** (g_{SM}) and **2 BSM** operators (g_{HWW}, g_{AWW})

→ **15 samples** with different parameters needed

- 50k events generated for each sample
- Kinematic observable used: $\Delta\phi_{jj}$
- Only signal considered

VBF H→WW example: Samples

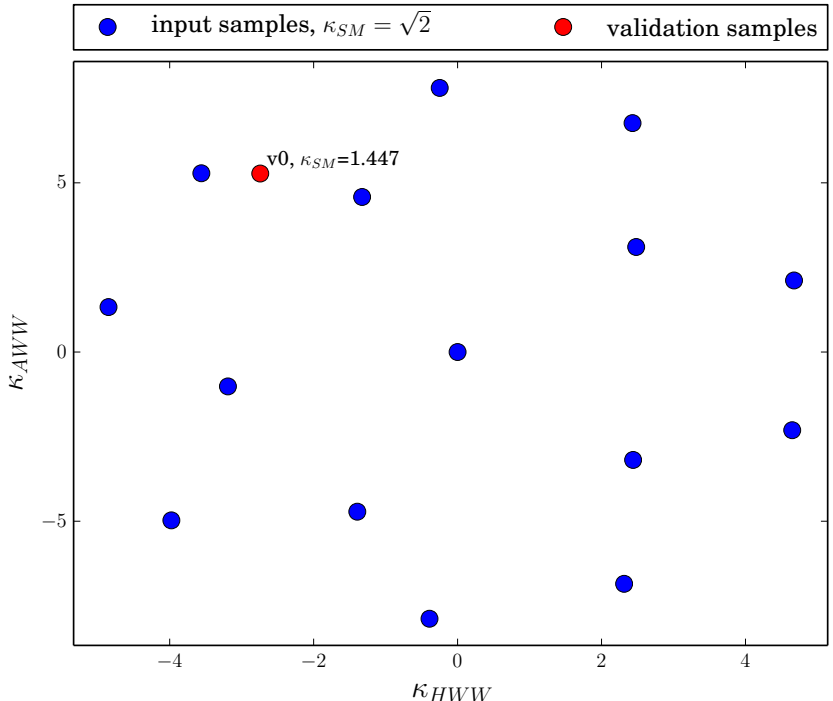
- Expect only **small deviations from SM**

→ $g_{\text{SM}} = 1$ for all input samples ($\Lambda = 1\,\text{TeV}$, $\cos\alpha = \frac{1}{\sqrt{2}}$)

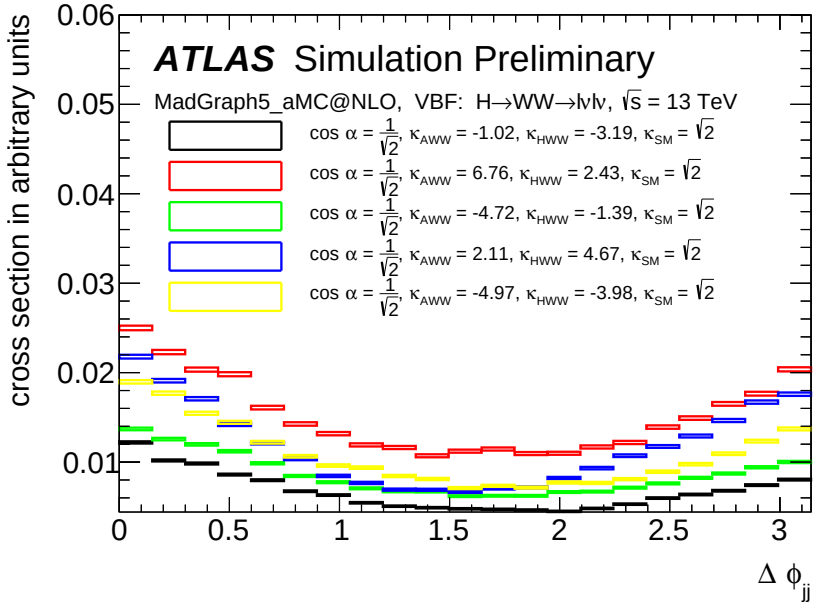
→ BSM parameter limits chosen such that $\sigma_{\text{pure BSM}} \sim \sigma_{\text{SM}}$

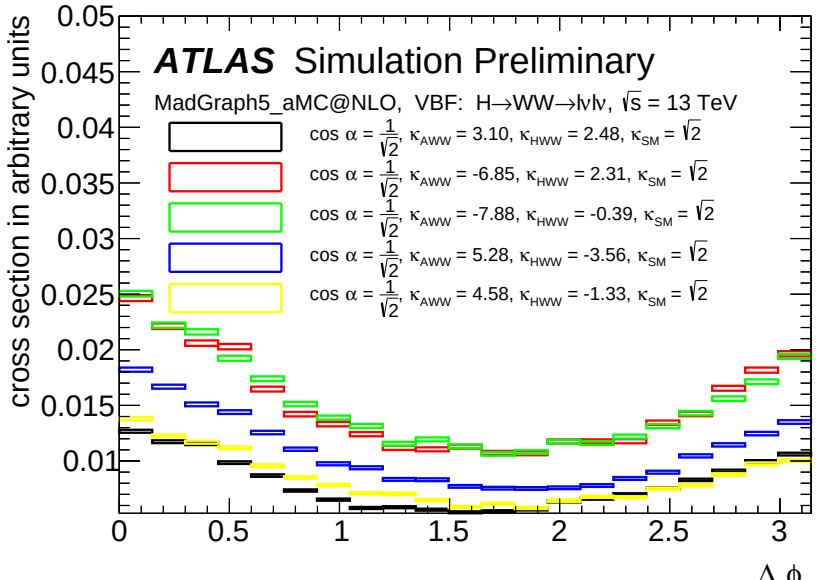
→ all other BSM parameters set to 0

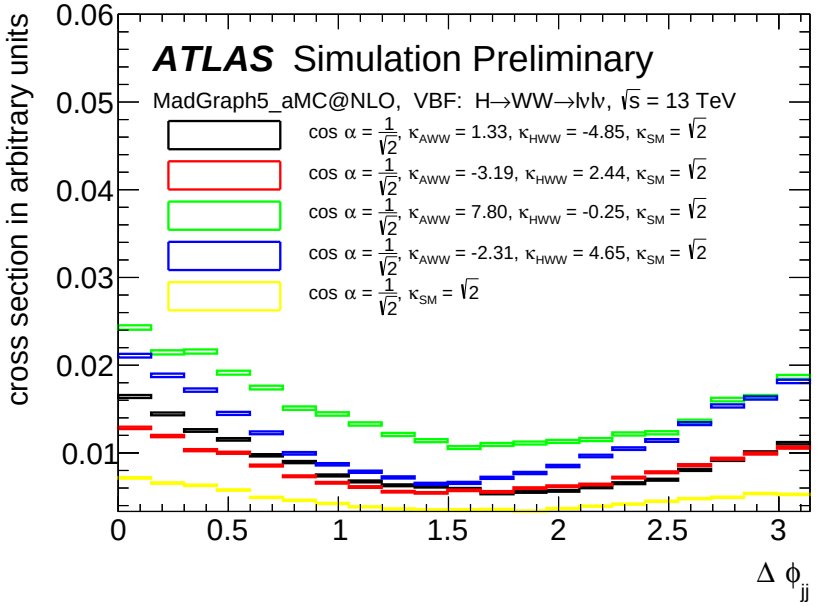
- Scatter plot shows blue points in (g_{AWW},g_{HWW}) space used to generate **input samples**
- A **validation sample** is produced at the red point for cross-check
  - → **morphing** can reproduce the distribution there
  - → **fit** can reproduce the parameters from the validation sample
- Use $g_i = \kappa_i \times g_i^{SM}$ to quickly see difference from the SM prediction



VBF H→WW example: Input distributions







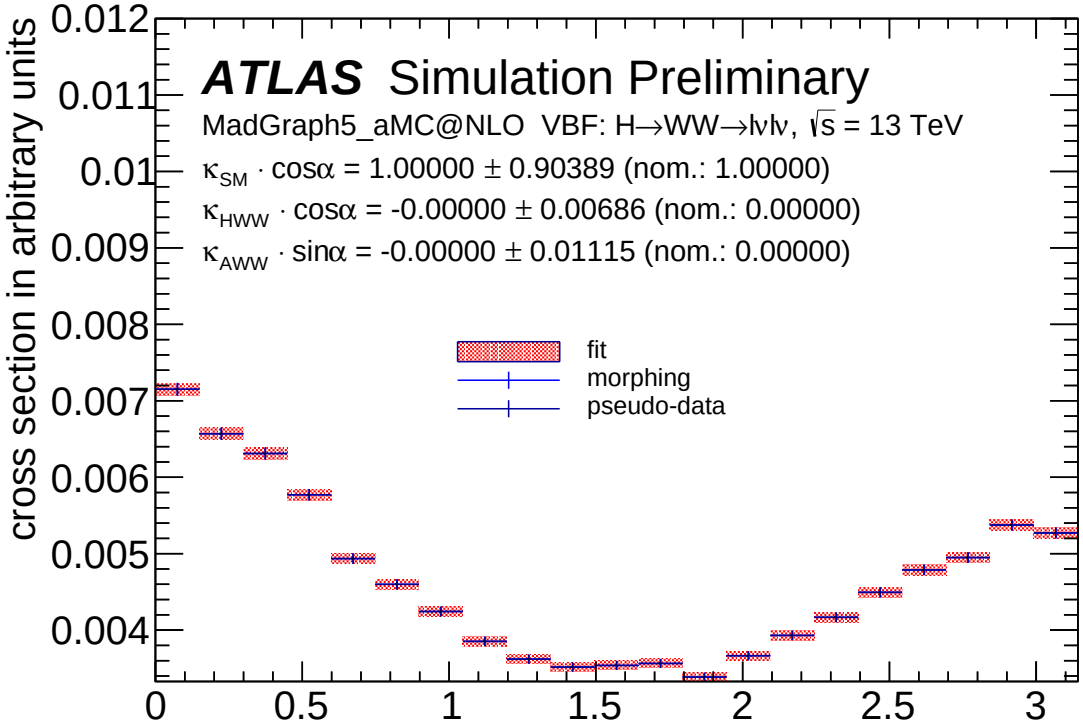
atl-phys-pub-2015-047

VBF H→WW example: Morphing and fit to SM input sample

- Morphing and fit to SM input distribution (pseudo-data)
- MC stat. uncertainty used
- Input and morphed distribution stat. dependent
  - perfect agreement in morphing
  - Post-fit parameters match exact nominal values
- **Sensitivity** on parameters shown in fit uncert.
- **Correlations** at SM point in table

	κ_{SM}	κ_{HWW}	κ_{AWW}
κ_{SM}	1.00	0.15	-0.23
κ_{HWW}	0.15	1.00	0.36
κ_{AWW}	-0.23	0.36	1.00

atl-phys-pub-2015-047

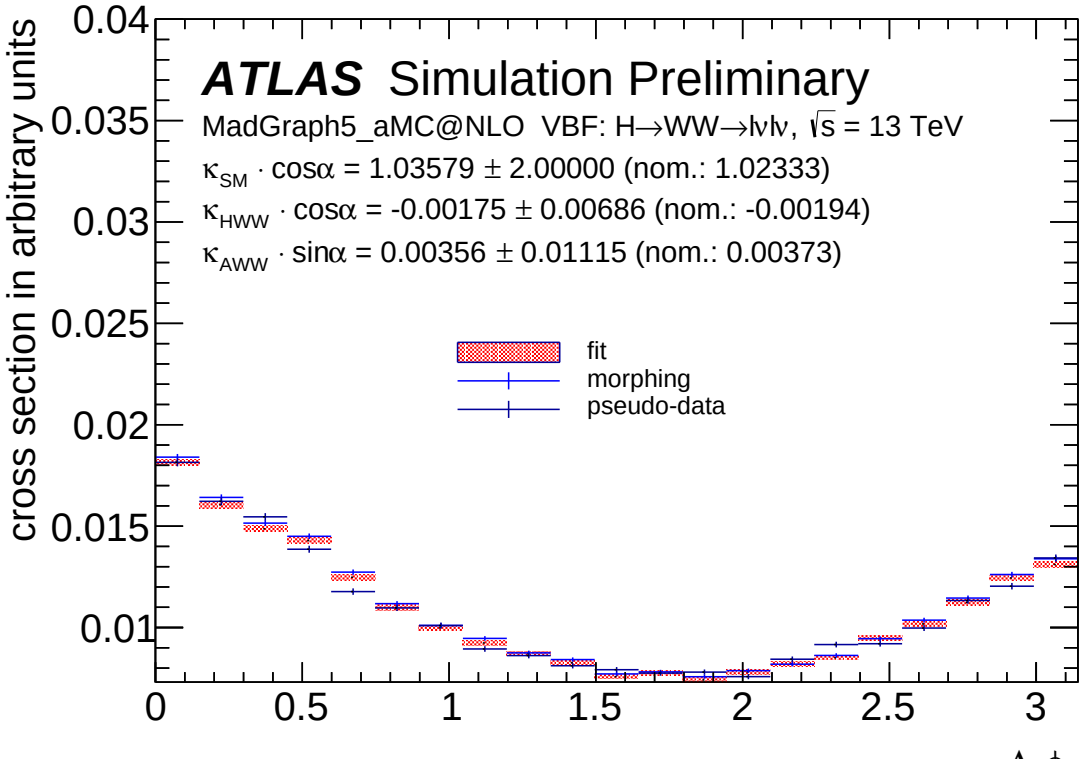


VBF H→WW example: Morphing and fit to validation sample

- Morphing and fit to validation distr. (pseudo-data)
- Validation and morphed distribution stat. independent
  - Agreement in morphing within MC stat. uncertainty
  - Fit results match nominal values within fit uncertainties
- **Sensitivity** on parameters shown in fit uncert.
- **Correlations** vary at different parameter point

	κ_{SM}	κ_{HWW}	κ_{AWW}
κ_{SM}	1.00	0.20	-0.95
κ_{HWW}	0.20	1.00	0.09
κ_{AWW}	-0.95	0.09	1.00

atl-phys-pub-2015-047



Summary

- Plan for Run 2: **Coupling and properties measurements**
- Combine **rate and shape information**, possibly within effective Lagrangian framework
- New method for modelling BSM effects
  - continuous
  - analytical
  - fast
- Used in several Higgs (example: **JHEP 03 (2018) 095**) and Exotics analyses in ATLAS and CMS

Backup

Number of input distributions

$$
\begin{aligned}
N_{input} = \quad & \frac{n_p\,(n_p+1)}{2} \cdot \frac{n_d\,(n_d+1)}{2} + \binom{4+n_s-1}{4} \\
& + \left(n_p \cdot n_s + \frac{n_s\,(n_s+1)}{2} \right) \cdot \frac{n_d\,(n_d+1)}{2} \\
& + \left(n_d \cdot n_s + \frac{n_s\,(n_s+1)}{2} \right) \cdot \frac{n_p\,(n_p+1)}{2} \\
& + \frac{n_s\,(n_s+1)}{2} \cdot n_p \cdot n_d + (n_p+n_d) \binom{3+n_s-1}{3}
\end{aligned}
$$

with number of parameters in **production vertex** (n_p), **decay vertex** (n_d) and **shared in vertices** (n_s)

Propagation of statistical uncertainties

- Morphing function for a bin in distribution

$$T_{out}^{bin}(\vec{g}_{target}) = \sum_i w_i(\vec{g}_{target};\vec{g}_i)\, T_{in}^{bin}(\vec{g}_i)$$

- For one input distribution, the bin content is calculated as follows

$$T_{in}^{bin}(\vec{g}_i) = N_{MC,in}^{bin}(\vec{g}_i) \cdot \sigma_{in}(\vec{g}_i)\mathcal{L}/N_{\mathrm{MC,in}}$$

- The uncertainty on that bin is $\sqrt{N_{MC,in}^{bin}(\vec{g}_i)}$
- The propagated statistical uncertainty is

$$\Delta T_{out}^{bin} = \sqrt{\sum_i w_i^2(\vec{g}_{target};\vec{g}_i) N_{MC,in}^{bin}(\vec{g}_i) \cdot (\sigma_{in}(\vec{g}_i)\mathcal{L}/N_{\mathrm{MC,in}})^2}$$

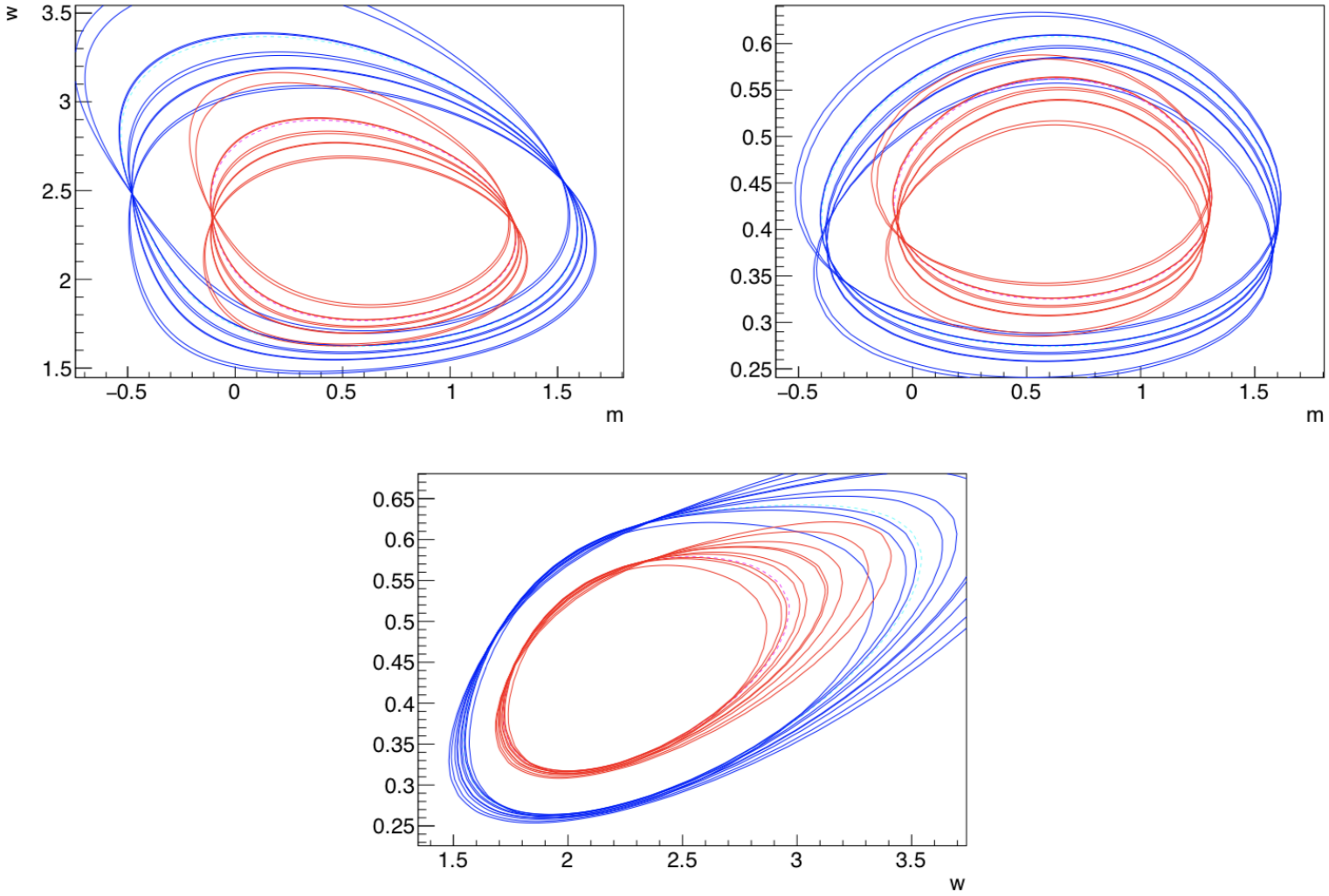
- Highly **dependent** on
  - **input parameters** $\vec{g}_i$
  - desired **target parameters** $\vec{g}_{target}$

Better method needed for n-dimensional Likelihood representation

BlurRing Full package available: https://tinyurl.com/BlrRng

Paper on Arxiv; ArXiv:1805.07213

Can clearly see deviations from hessian (elliptical) shape

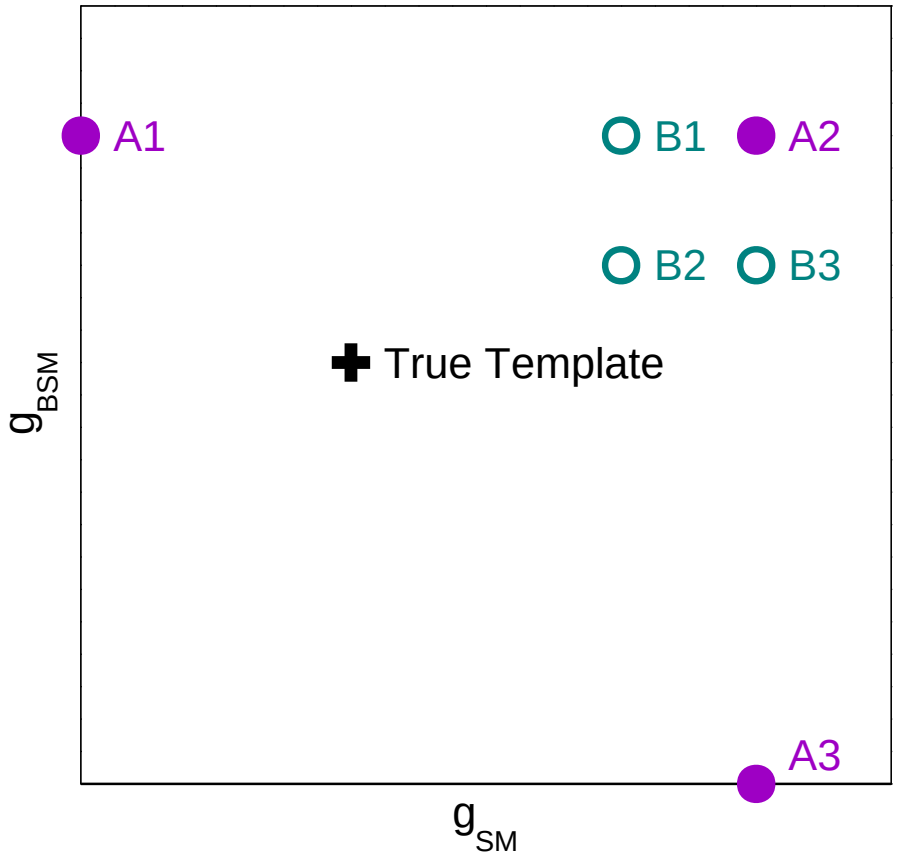


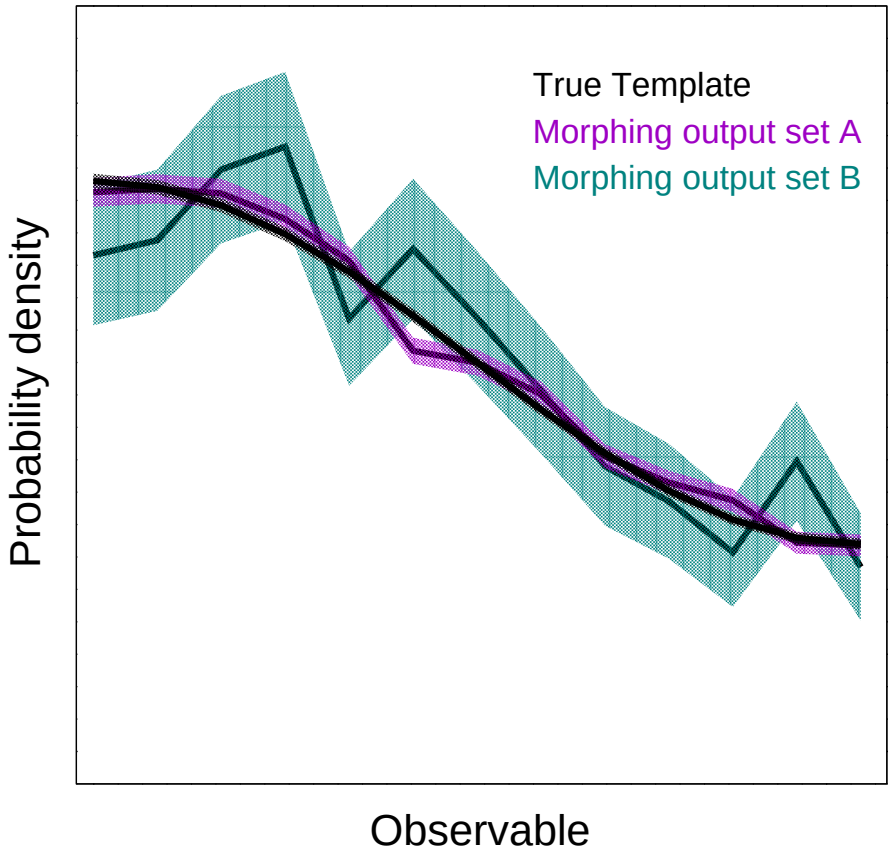
Choice of input parameters

[noframenumbering]

- Aim to **generalise morphing** to have arbitrary g_i

→ Can be chosen to **reduce statistical uncertainty**





VBF H→WW example: Rel. uncertainty on number of expected events

- Dependence of **stat. uncertainty** propagated in morphing function on generated input parameter grid
- Distribution of samples in parameter space **reduces** stat. uncertainty