Template morphing

Continuous modelling in a multidimentional space of parameters

Lydia Brenner

Introduction

Start with the formulation of a likelihood: $L(\vec{x} \mid \vec{\mu}, \vec{\theta})$

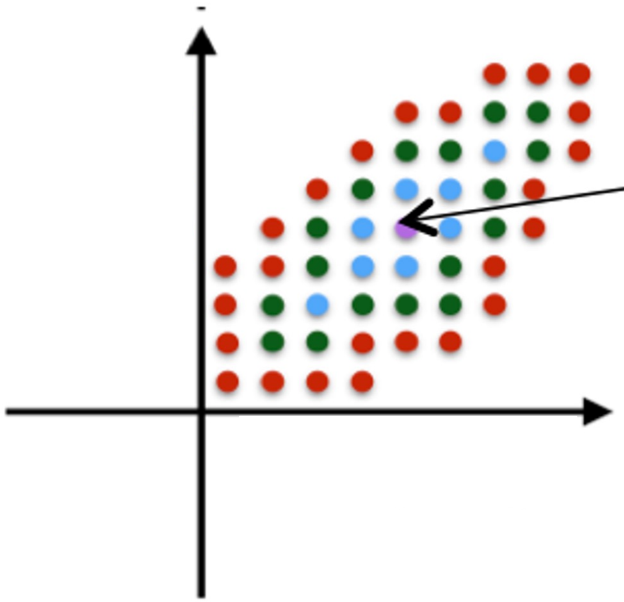
- (B)SM physics model * Soft physics model * ATLAS detector description * ATLAS analysis reconstruction

Problem: We don't have a continuous description of $L(\vec{x} \mid \vec{\mu}, \vec{\theta})$

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

Introduction

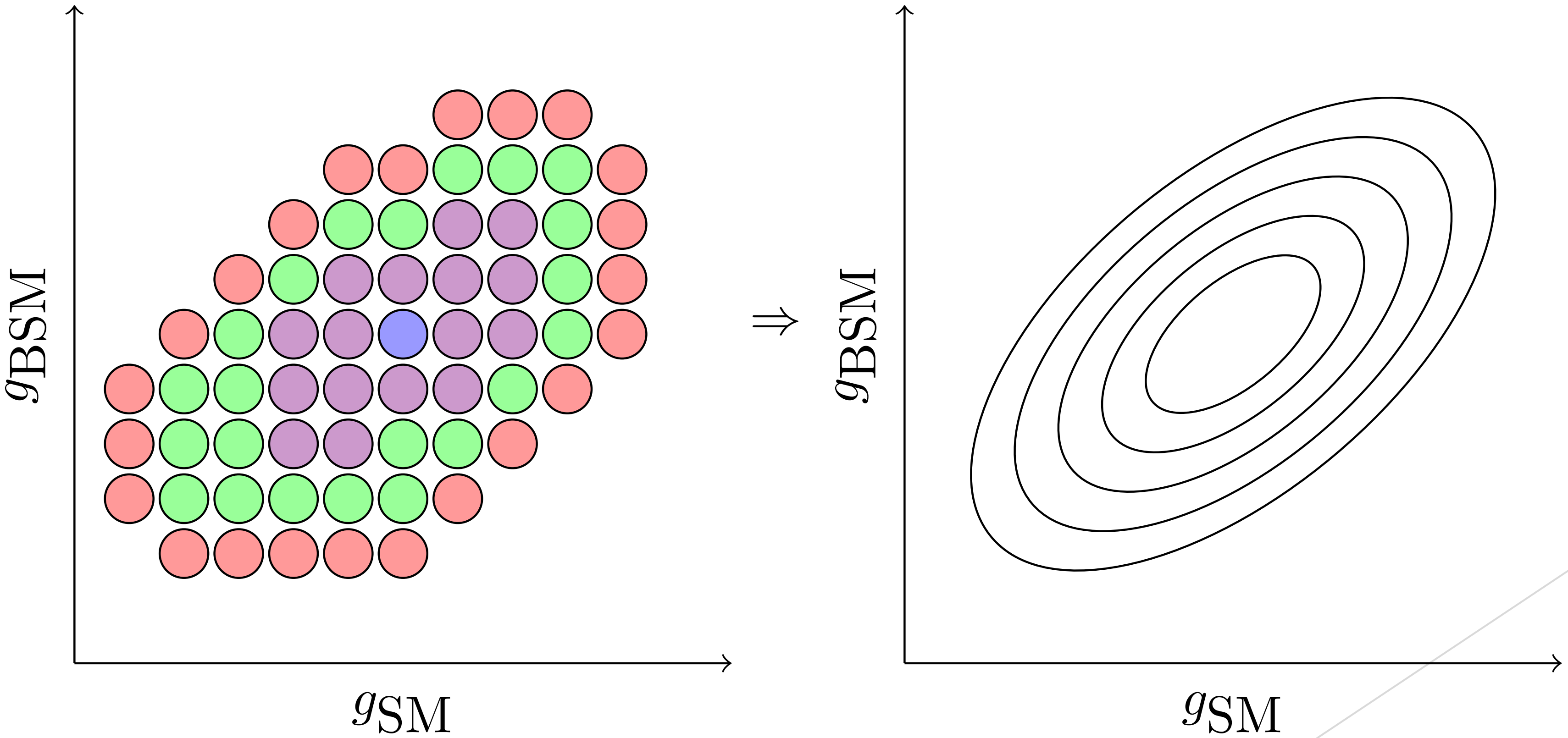
Can approximate statistical procedure with grid scan



$$L_{ij}(\vec{\theta}) = S_{ij}(x \mid \tilde{R}_i, \tilde{I}_j, \vec{\theta}) + B(x \mid \vec{\theta})$$

Introduction

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

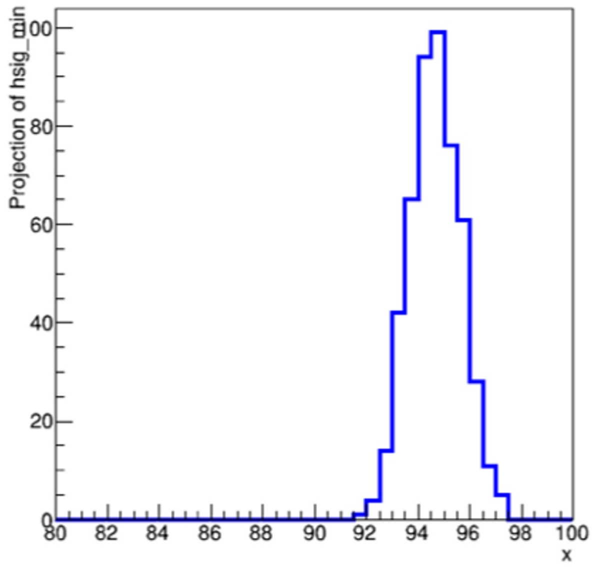


Interpolating between models

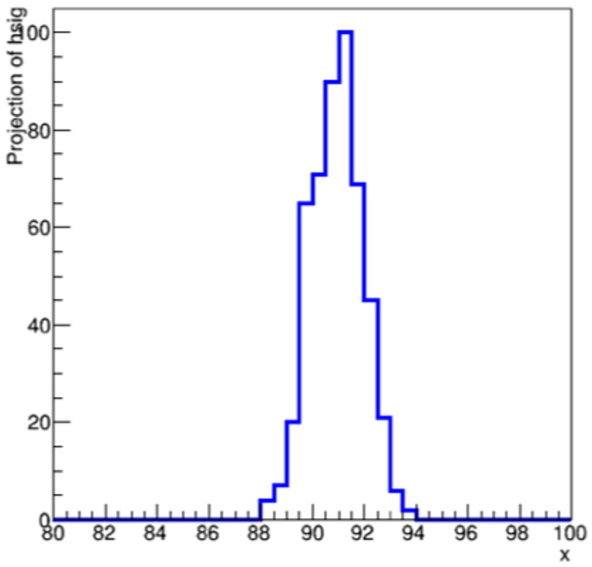
- Need to define a morphing algorithm to define *s(x)* for any value of *a*

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

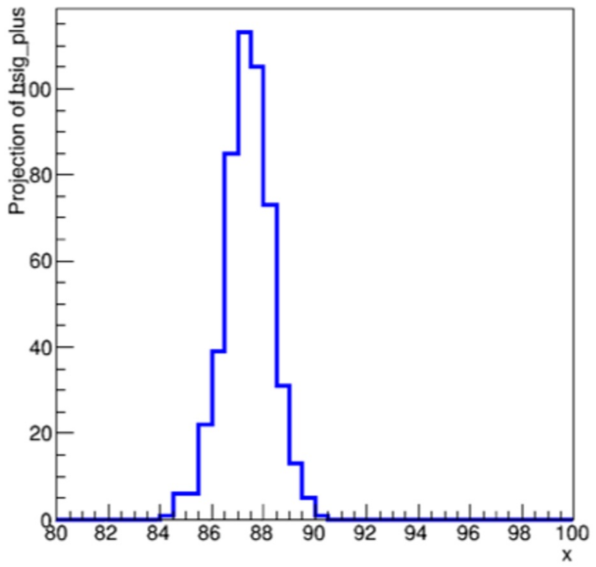
s(x)|a=-1



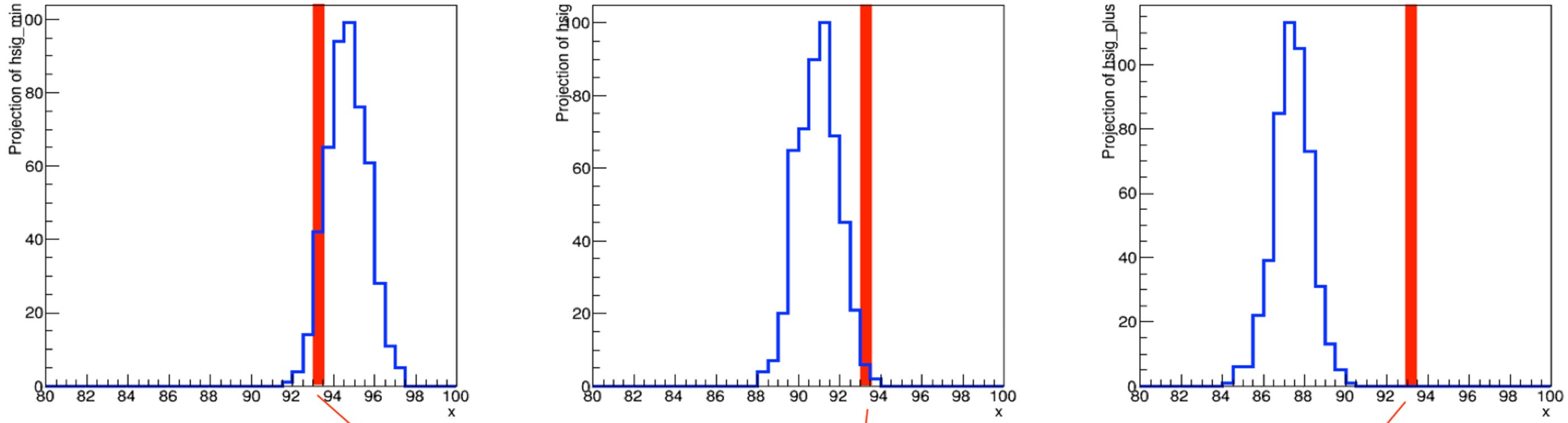
s(x)|a=0



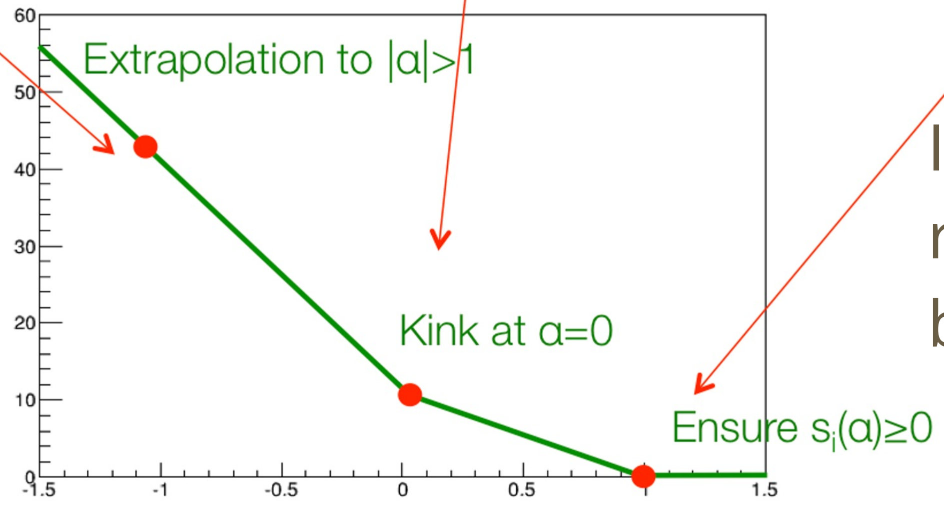
s(x)|a=1



Interpolating between models



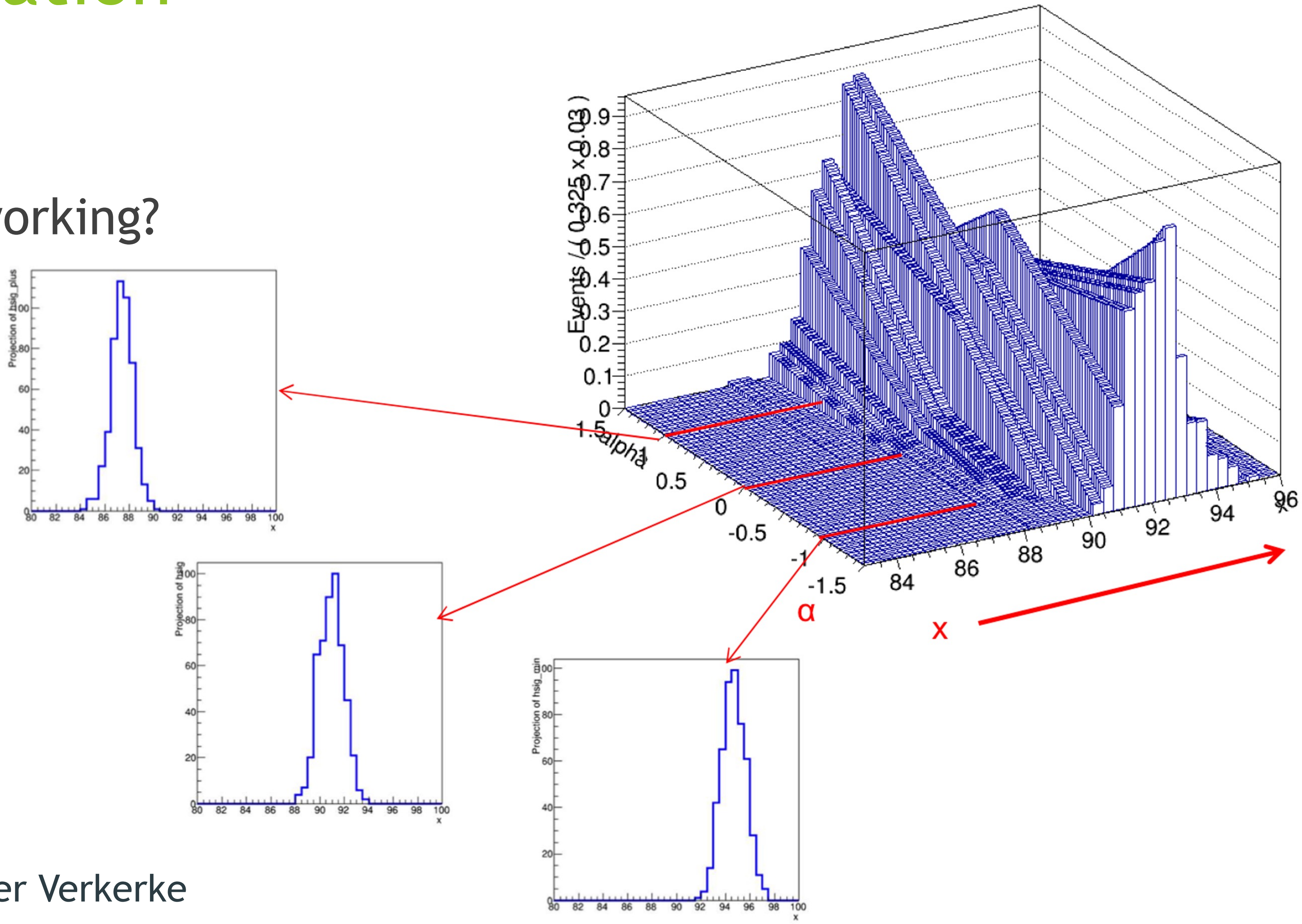
Simplest solution is piecewise linear interpolation



Interpolates response model bin by bin

Linear interpolation

When does this stop working?



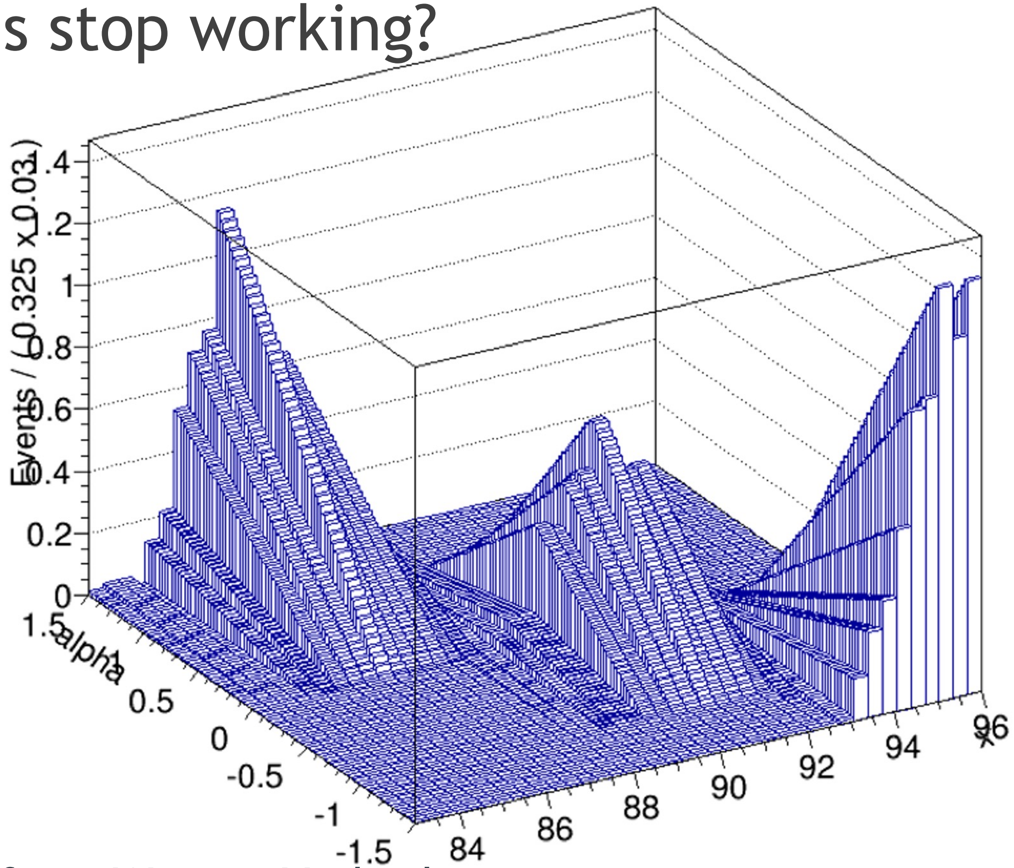
Visualisation from Wouter Verkerke

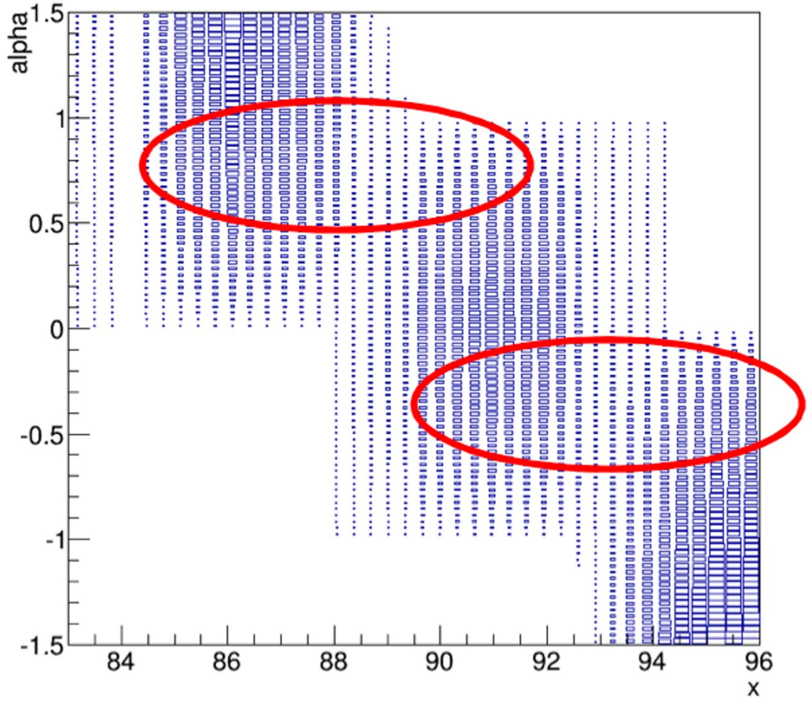
Linear interpolation

When does this stop working?

Example:

Large shift

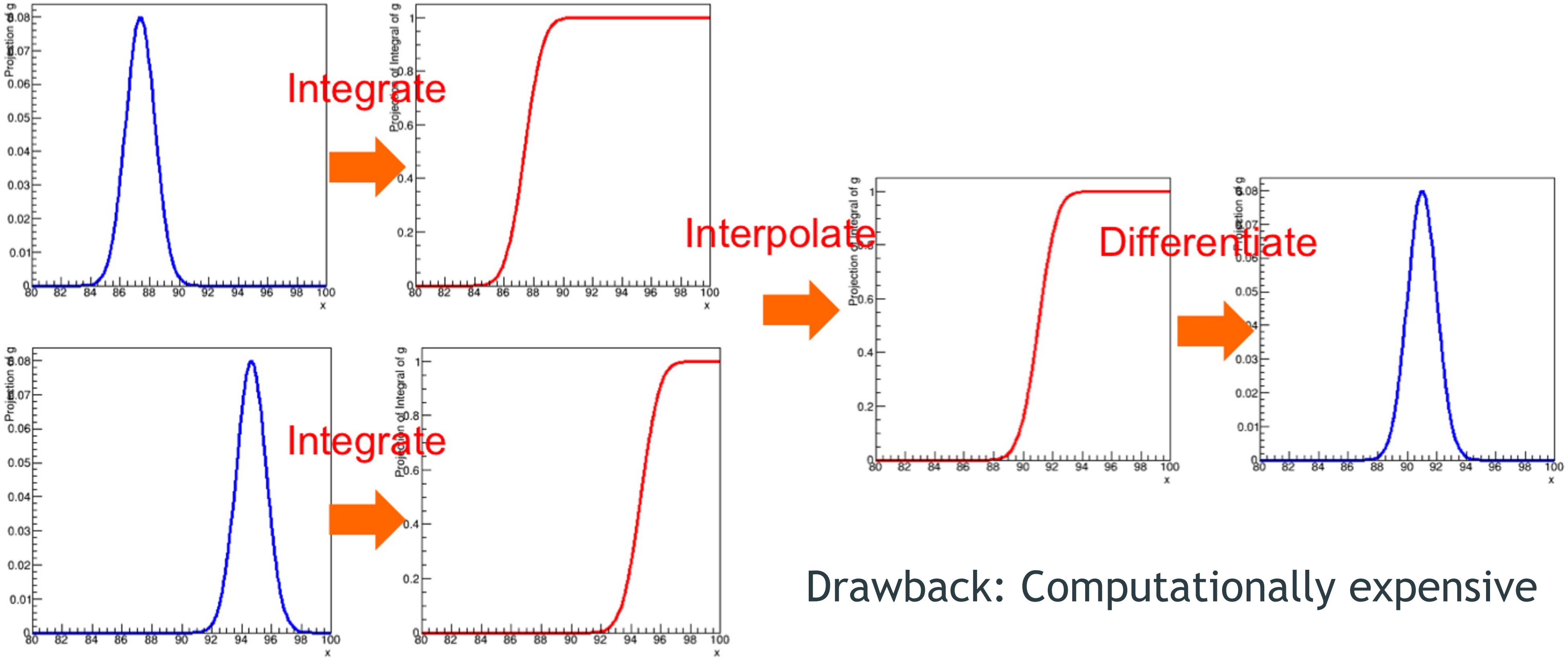




Visualisation from Wouter Verkerke

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

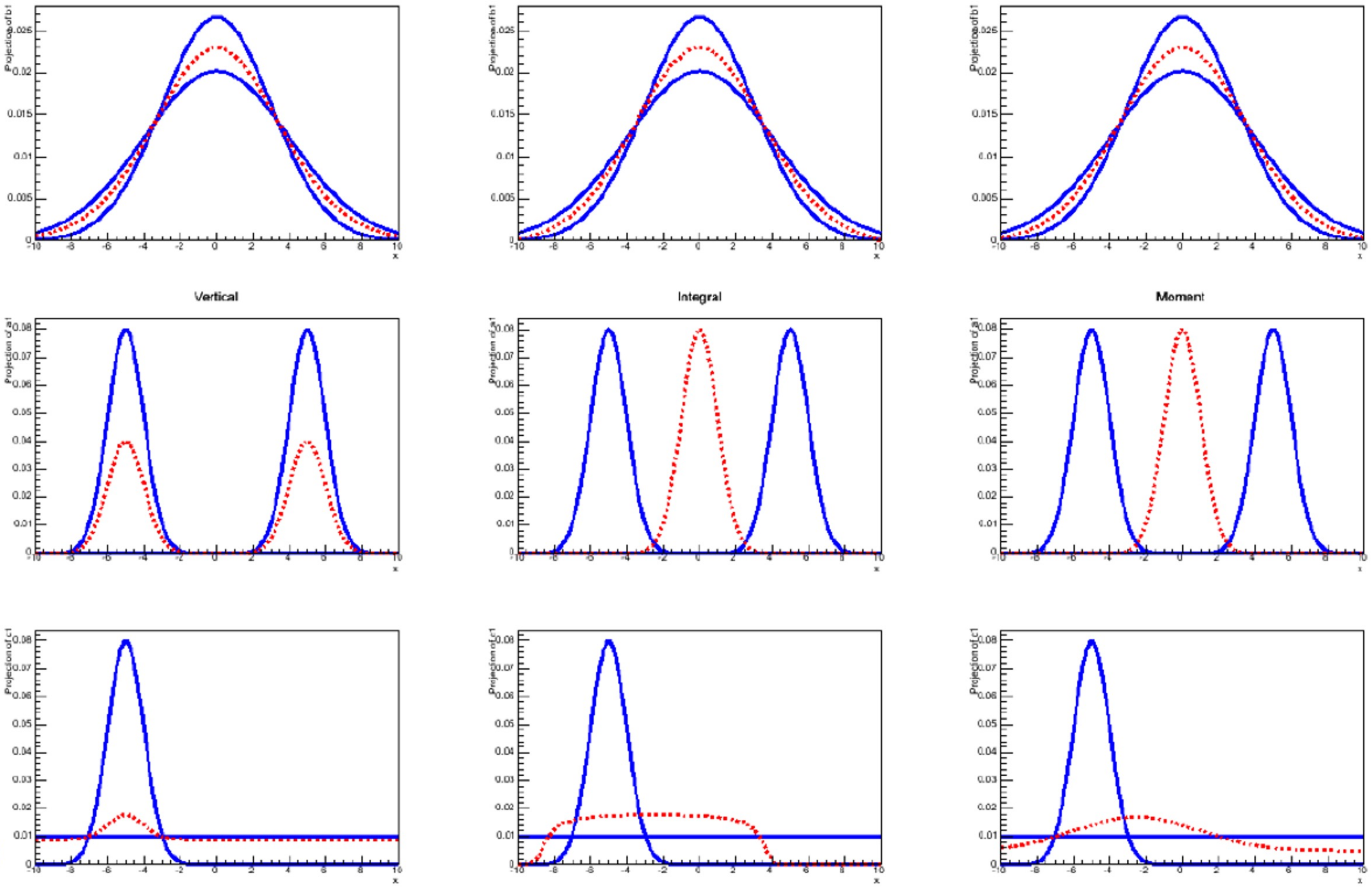
Multidimensional interpolation option

Computationally expensive, but only once

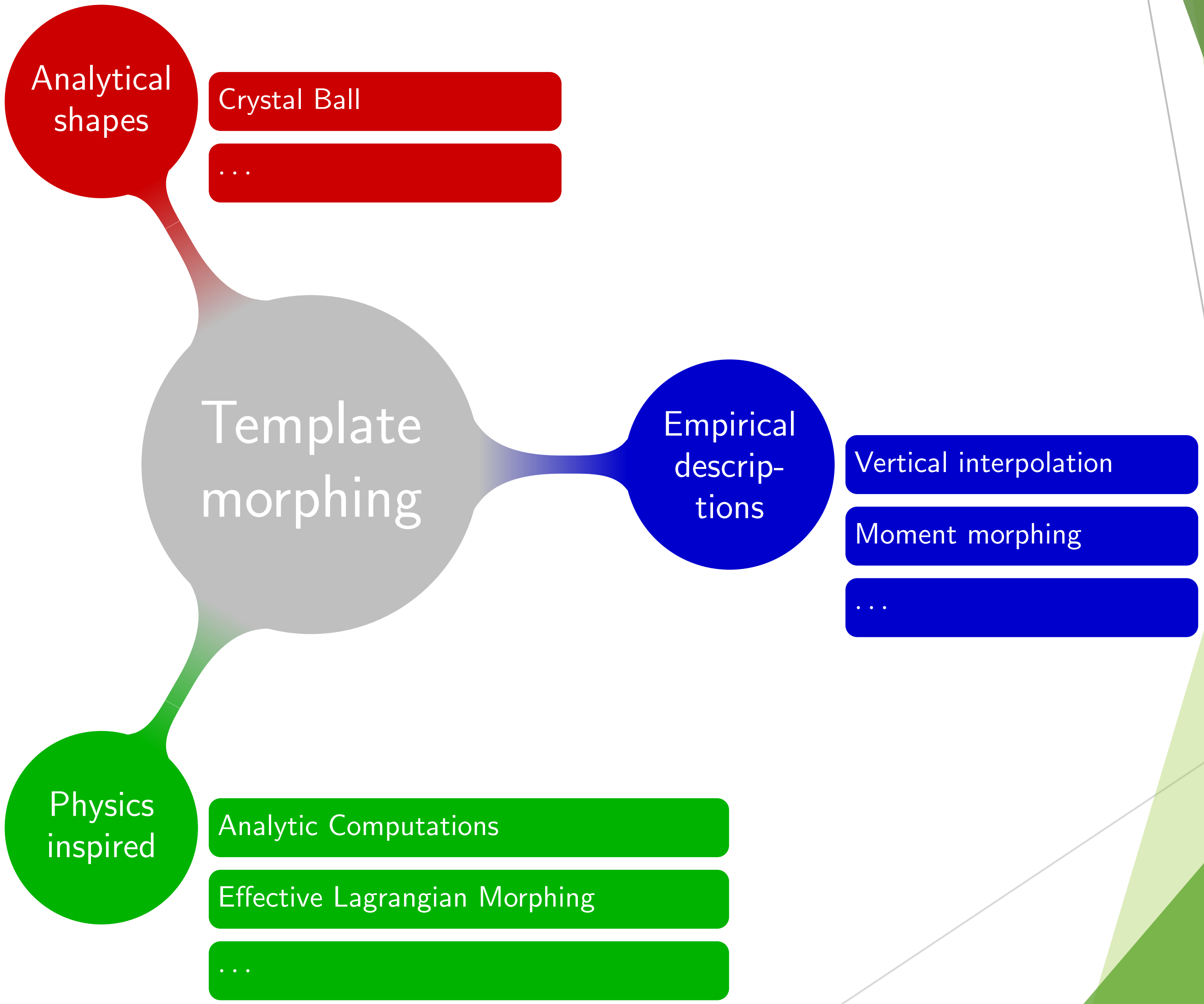
Comparing the methods

Different ways to create a continuous distribution of the likelihood

	Vertical Morphing	Horizontal Morphing	Moment Morphing
Gaussian varying width			
Gaussian varying mean			
Gaussian to Uniform (this is conceptually ambigous!)			



Overview of methods



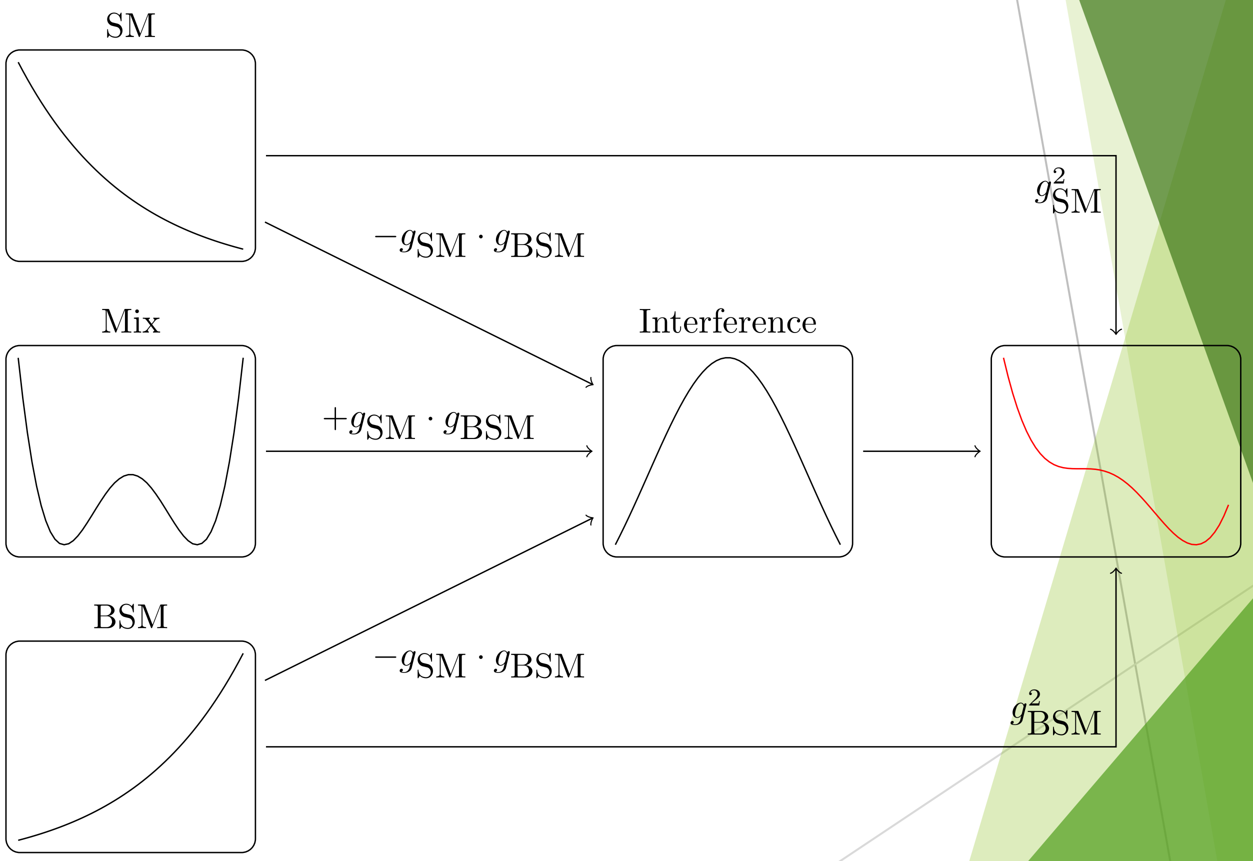
Effective lagrangian morphing

This method allows for more complicated distributions

- Continuous
- Analytic
- Fast

Combines rate and shape information simultaneously

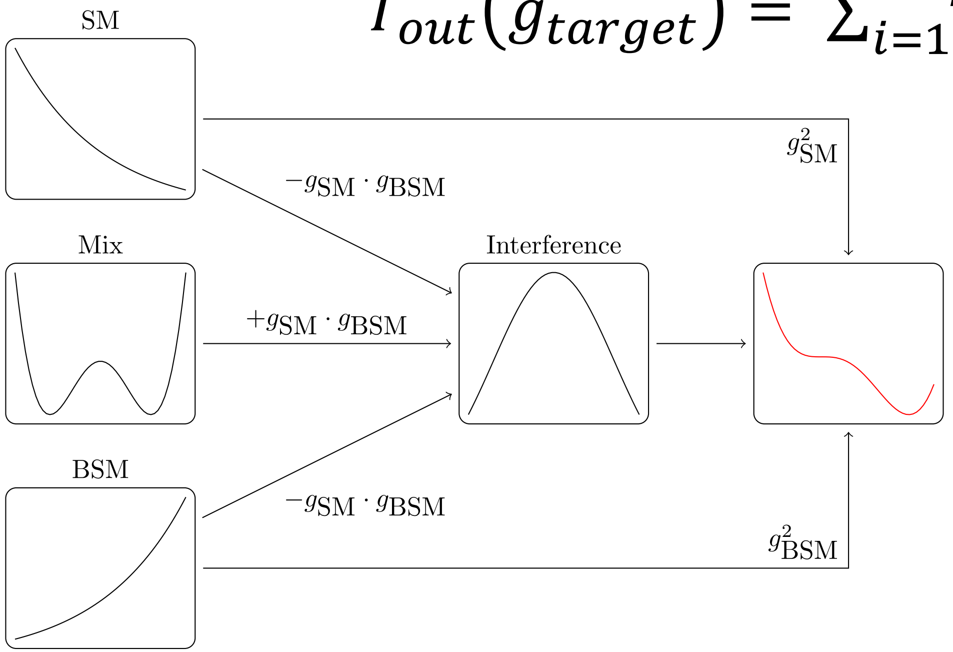
Can use any Lagrangian as starting point, I will use effective models in my examples

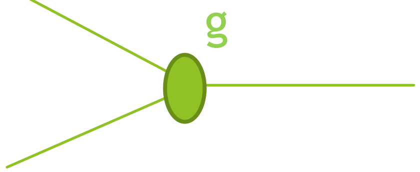


Model prametrisation

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 $\vec{g}_i$

$$T_{out}(\vec{g}_{target}) = \sum_{i=1}^{N_{input}} w_i(\vec{g}_{target}; \vec{g}_i)\, T_{in}(\vec{g}_i)$$



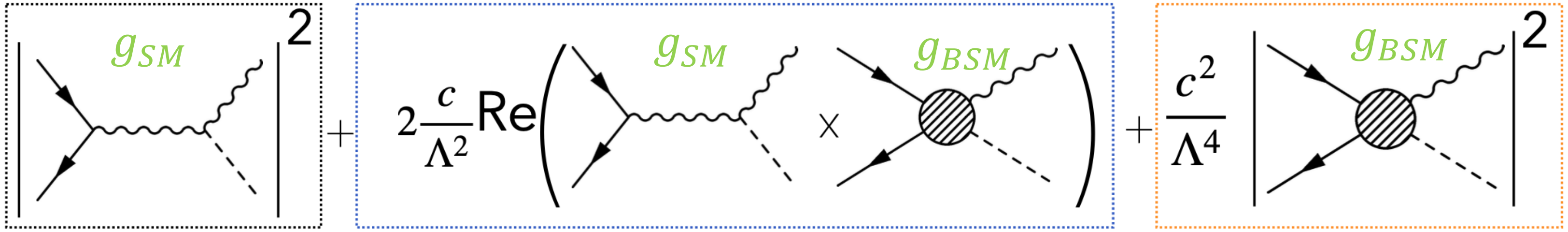


Example with two free parameters in one vertex

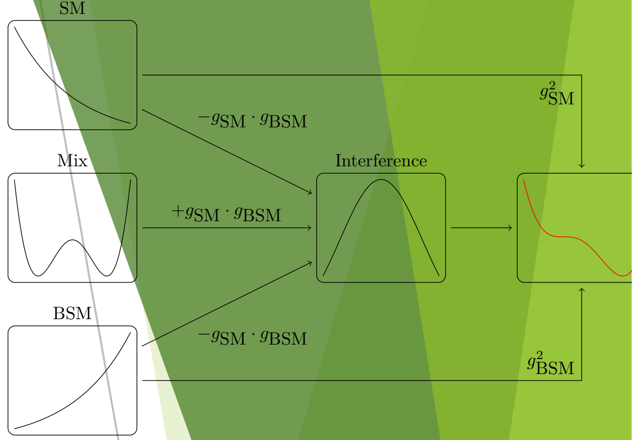
Distribution of a kinematic observable proportional to the matrix element squared

$\mathcal{M}(g_{SM}, g_{SM}) = g_{SM}\ \mathcal{O}_{SM} + g_{BSM}\ \mathcal{O}_{BSM}$

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



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



Example with two free parameters in one vertex

Three generated distributions $T_{in}\ (g_{SM}, g_{BSM})$ needed to obtain distribution with arbitrary parameters

$T_{in}(1,0)\ =\ |\mathcal{O}_{SM}|^2$

$T_{in}(0,1)\ \ |\mathcal{O}_{BSM}|^2$

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

Going now to arbitrary parameters (g_{SM}, g_{BSM}) using

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

We get

$T_{out}(g_{SM}, g_{BSM}) = (g_{SM}{}^2 - g_{SM}\ g_{BSM})T_{in}(1,0) + (g_{BSM}{}^2 - g_{SM}\ g_{BSM})\ T_{in}(0,1) + g_{SM}\ g_{BSM}\ \ T_{in}(1,1)$

Generalisation to n dimentions

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

Where n_p is the number of parameters in the production vertex, n_d the number in the decay vertex, and n_s the number shared in both vertices

So the number of input parameters needed is

$$n_{input} = \frac{n_p(n_p+1)}{2}\frac{n_d(n_d+1)}{2} + \binom{4+n_s-1}{4} + \left(n_p n_s + \frac{n_s(n_s+1)}{2}\right)\frac{n_d(n_d+1)}{2} + \left(n_d n_s + \frac{n_s(n_s+1)}{2}\right)\frac{n_p(n_p+1)}{2} + \frac{n_s(n_s+1)}{2} n_p n_d + (n_p+n_d)\binom{3+n_s-1}{3}$$

Propegation of sample uncertainties

Reminder: the morhping function for a bin in the distribution is

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

For one input distribution, the bin content is calculated as

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

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

The propegated statistical uncertainty is

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

Dependent on chosen input paramters points $\vec{g}_i$ as well as desired output parameter point $\vec{g}_{target}$

Input parameter point $\vec{g}_i$, or input distribtions T_{in}, can be chosen to reduce MC statistical uncertainties

Interpolation of systematic uncertainties

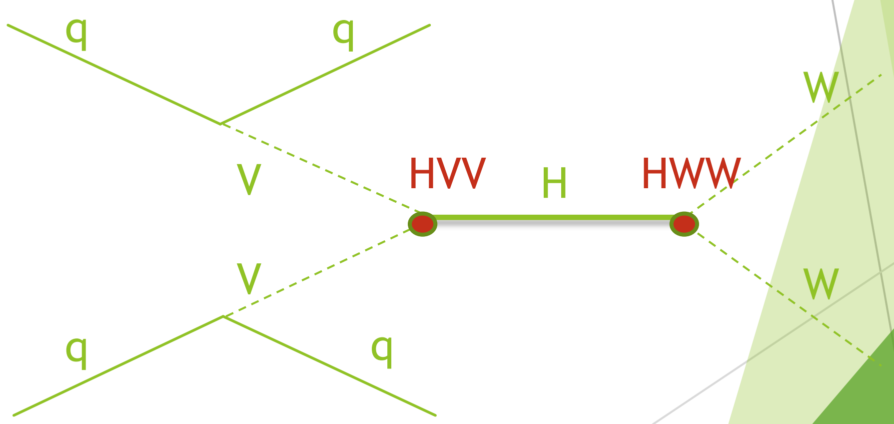
Following same method of template morphing for (total) uncertainty

- Try taking into account possible changing in uncertainties in the multi-dimentional space
  - Estimating all uncertainties in all input sample points can be too expensive or complicated
- Try taking into account possible changes in correlations between uncertainties in the multi-dimentional space
  - Physics of the uncertainties does not follow the same physics as the signal model

VBF H→ WW example

VBF H→ WW process with one SM (g_{SM}) and two BSM (g_{HWW} ,g_{AWW}) parameters

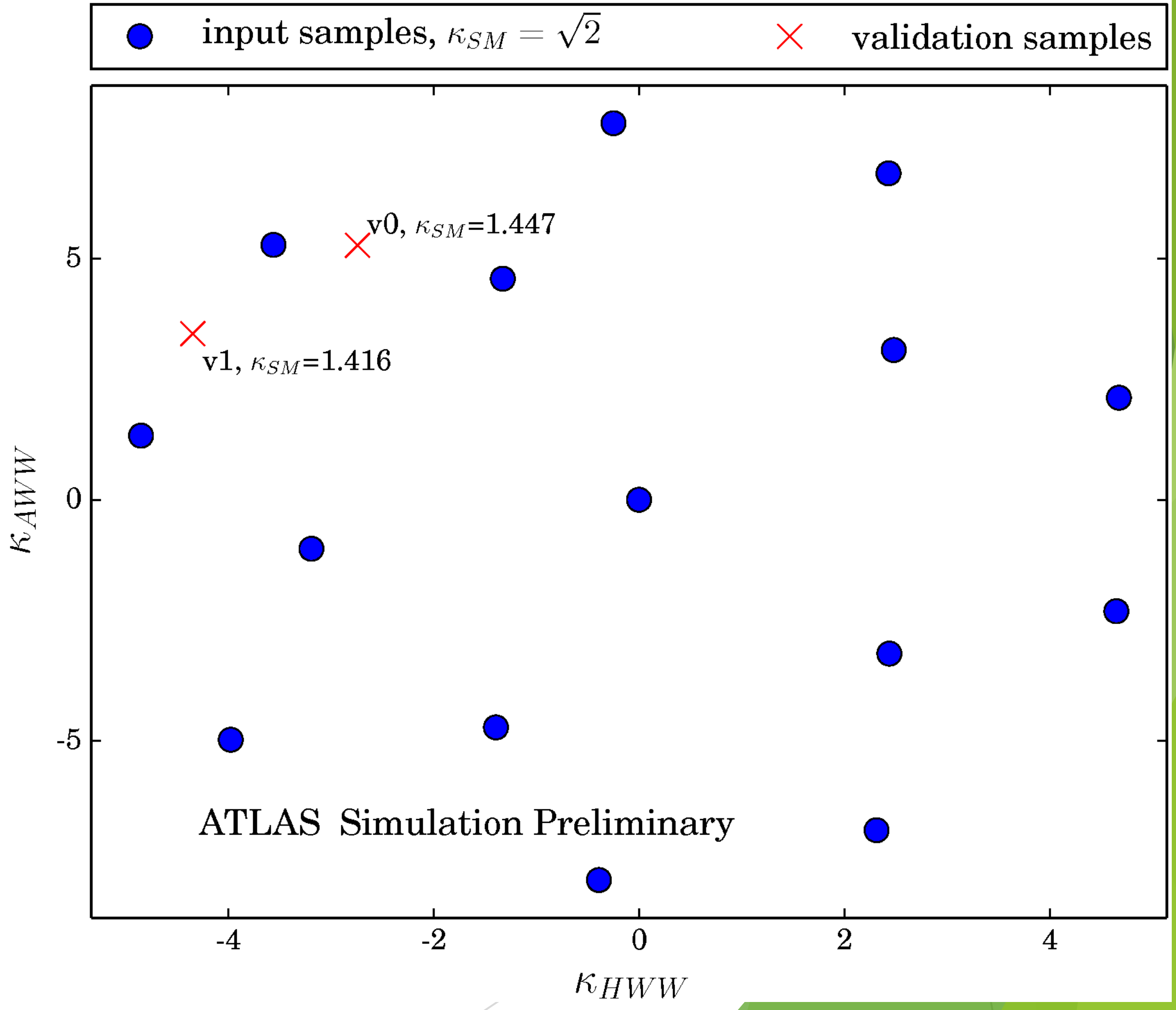
- 15 samples needed as inputs
- Each sample with a 50k sample size
- Consider signals only, background free
- Look at one kinematic observable $\Delta\phi_{jj}$



VBF H→ WW example : Input samples

Expact only small deviation from SM

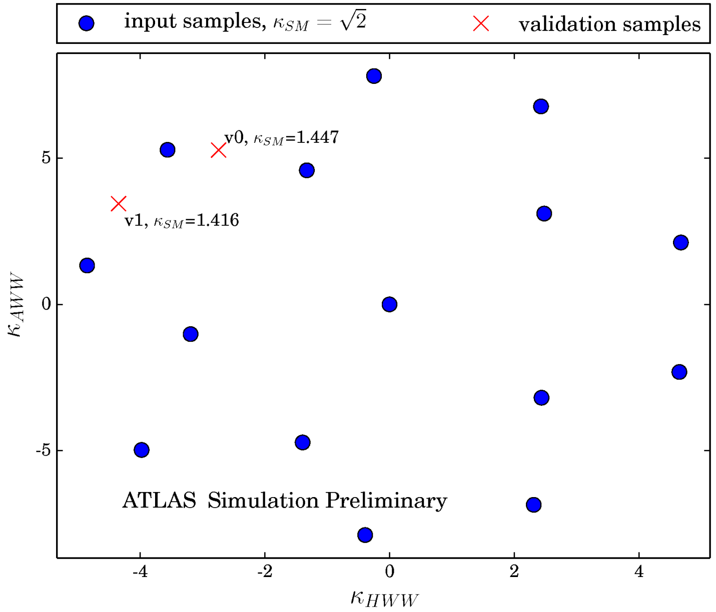
- $g_{SM} = 1$ for all input samples
- BSM parameter limits chosen such that $\sigma_{pure\ BSM} \sim \sigma_{SM}$

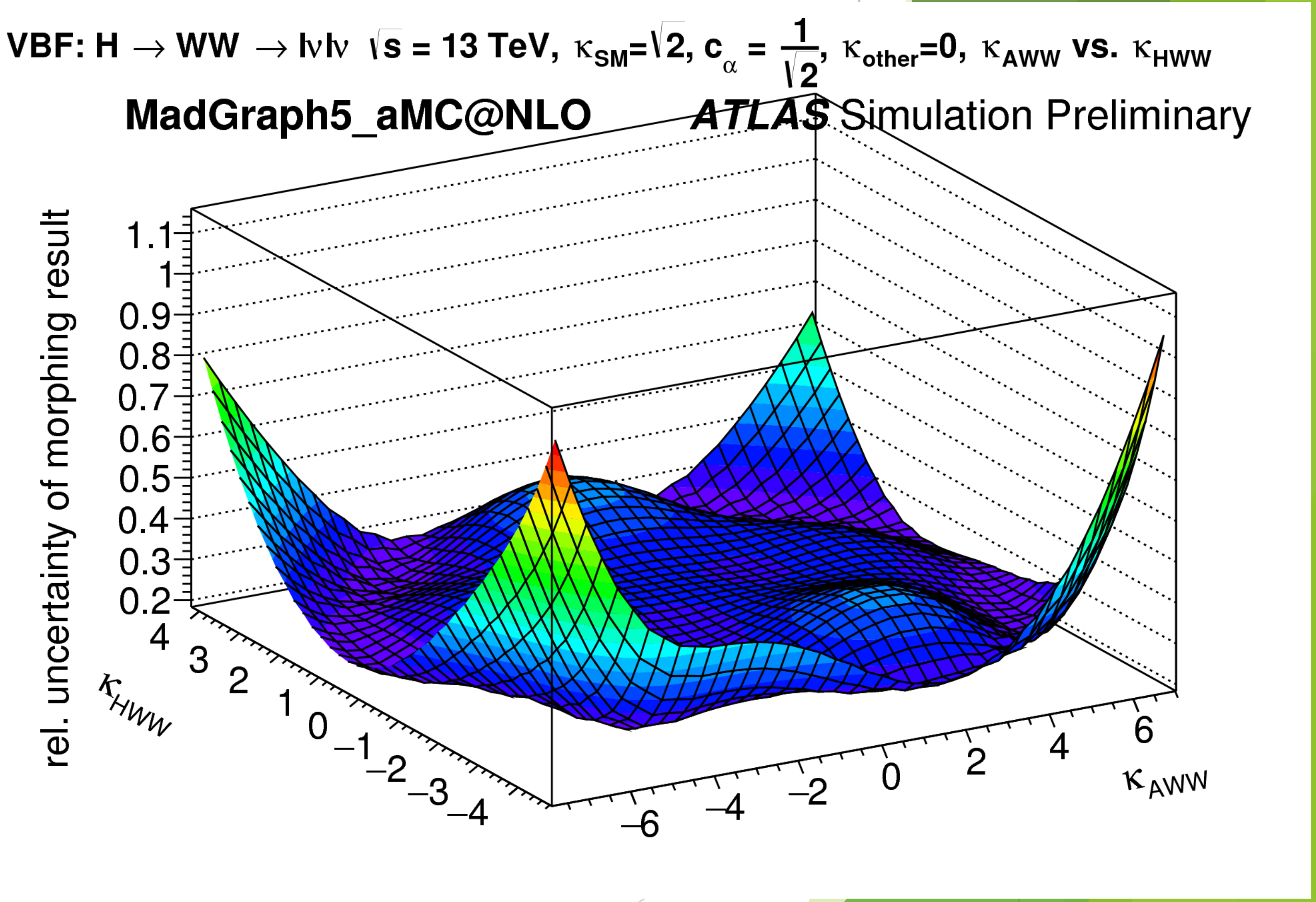


VBF H→ WW example : Input samples

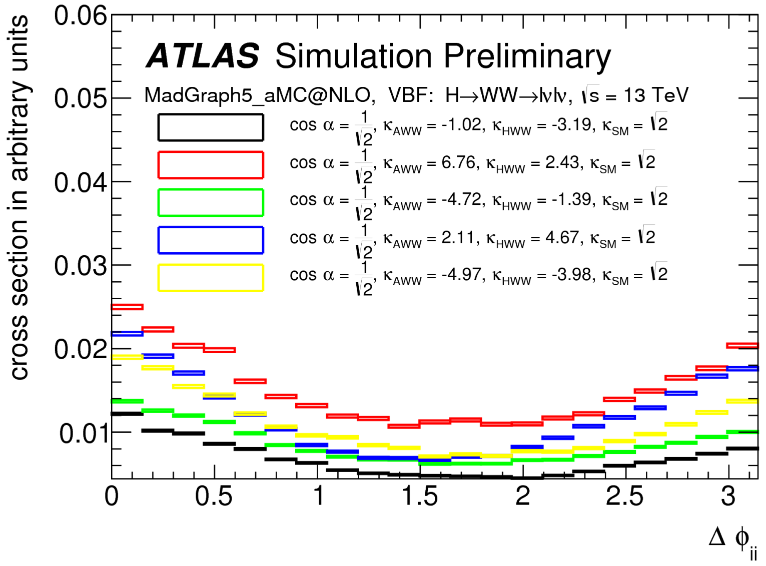
Expact only small deviation from SM

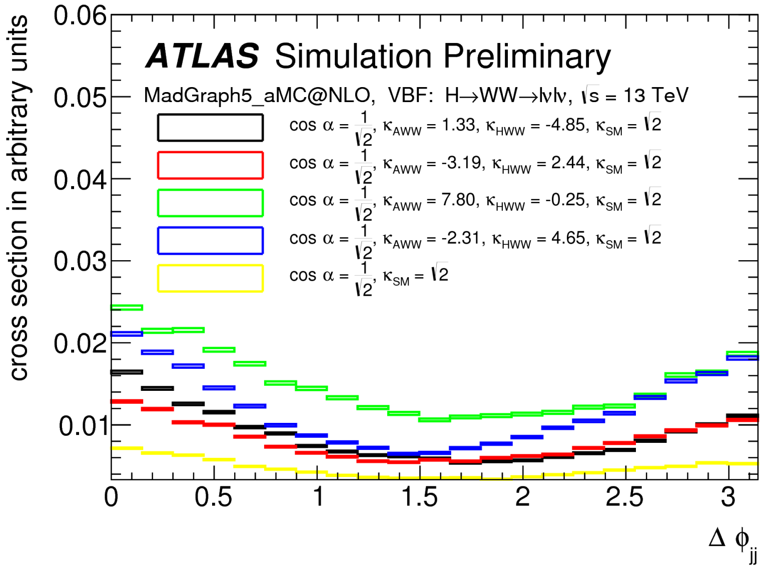
- $g_{SM} = 1$ for all input samples
- BSM parameter limits chosen such that $\sigma_{pure\ BSM} \sim \sigma_{SM}$

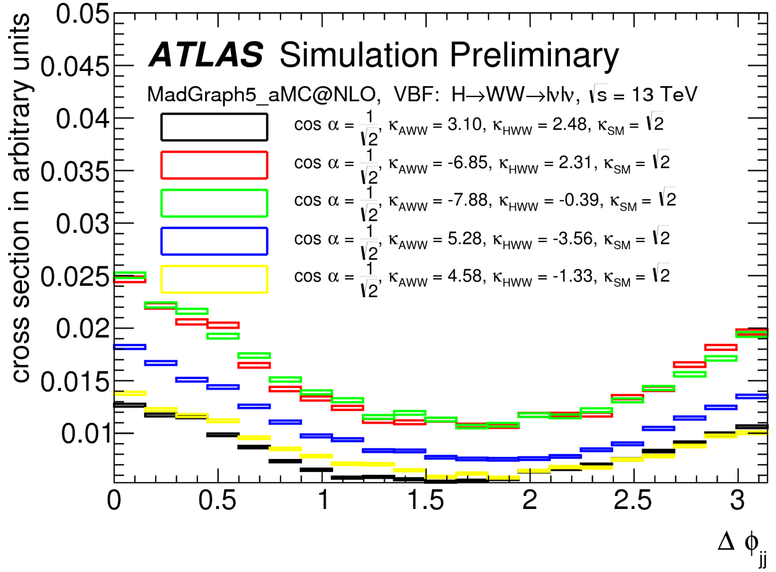


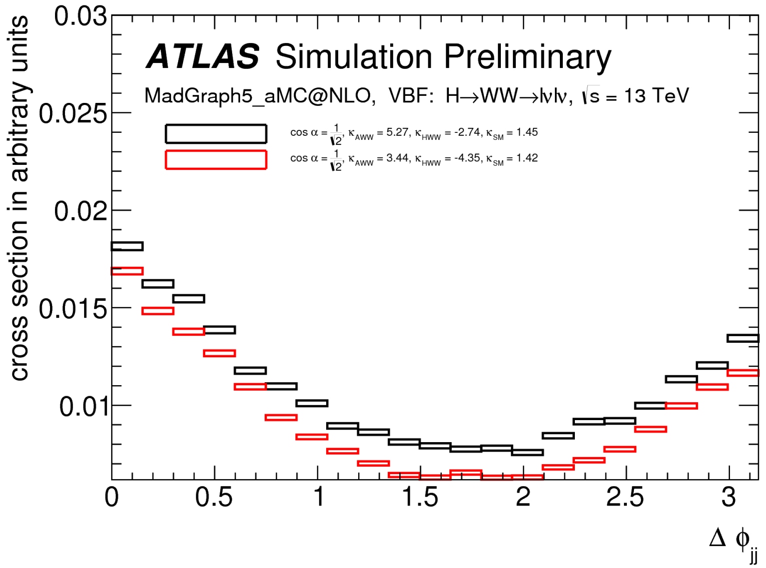


VBF H→ WW example : Distributions

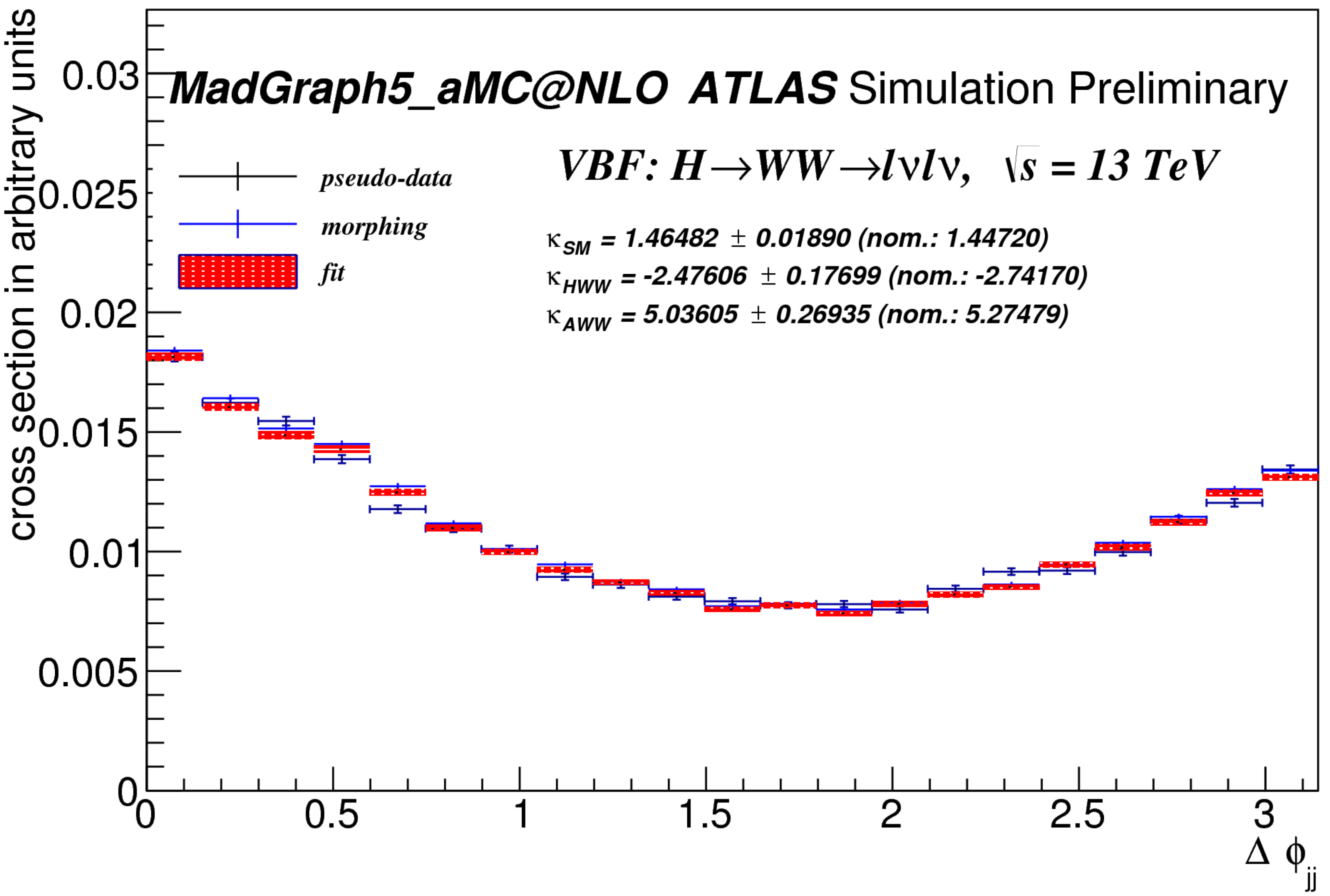








VBF H→ WW example : Fit



Summary

The morphing techniques provide a powerful way to model the distributions in combined likelihoods

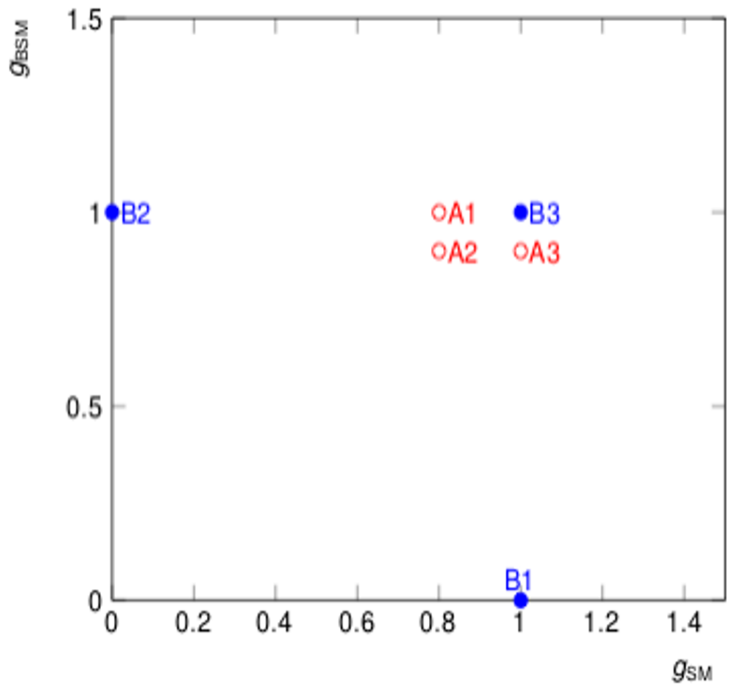
- All available with ROOT release v6.26.00

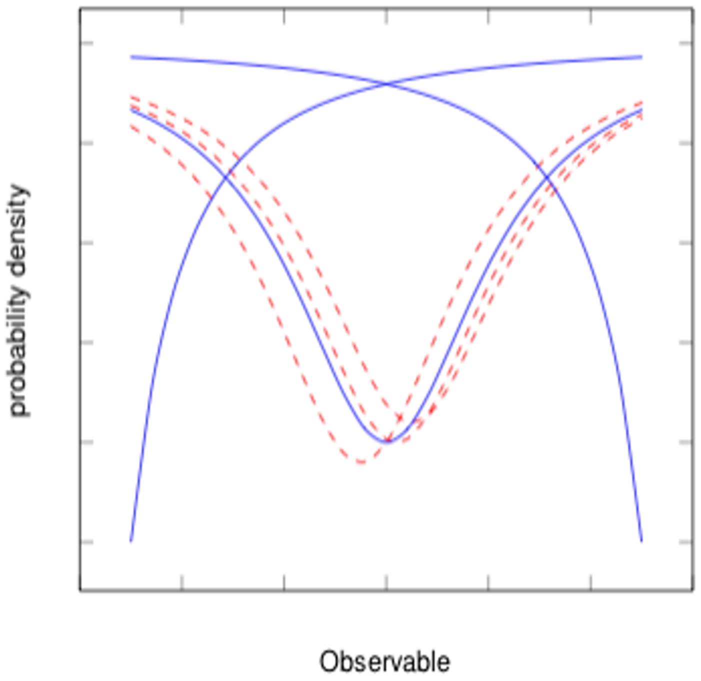
Different methods are correct in different situation

- Consider computational costs
- Uncertainty propagation of systematics non-trivial

back up

Parameters for input distribution





Choose to reduce statistical uncertainties

Generalisation to n dimentions

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

Where n_p is the number of parameters in the production vertex, n_d the number in the decay vertex, and n_s the number shared in both vertices

So the number of input parameters needed is

$$n_{input} = \frac{n_p(n_p+1)}{2}\frac{n_d(n_d+1)}{2} + \binom{4+n_s-1}{4} + \left(n_p n_s + \frac{n_s(n_s+1)}{2}\right)\frac{n_d(n_d+1)}{2} + \left(n_d n_s + \frac{n_s(n_s+1)}{2}\right)\frac{n_p(n_p+1)}{2} + \frac{n_s(n_s+1)}{2} n_p n_d + (n_p+n_d)\binom{3+n_s-1}{3}$$

So for example with 13 free parameters in VBF H → VV you need 1605 input parameters

- Lots of input samples creation can be computationaly expensive
- Interpolation computationally cheap

Higgs characterisation model

$$
\begin{aligned}
\mathcal{L}_0^V = \Big\{ & c_\alpha \kappa_{\rm SM} \big[\tfrac{1}{2} g_{HZZ}\, Z_\mu Z^\mu + g_{HWW}\, W_\mu^+ W^{-\mu} \big] \\
& - \frac{1}{4} \big[c_\alpha \kappa_{H\gamma\gamma} g_{H\gamma\gamma}\, A_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{A\gamma\gamma} g_{A\gamma\gamma}\, A_{\mu\nu} \widetilde{A}^{\mu\nu} \big] \\
& - \frac{1}{2} \big[c_\alpha \kappa_{HZ\gamma} g_{HZ\gamma}\, Z_{\mu\nu} A^{\mu\nu} + s_\alpha \kappa_{AZ\gamma} g_{AZ\gamma}\, Z_{\mu\nu} \widetilde{A}^{\mu\nu} \big] \\
& - \frac{1}{4} \big[c_\alpha \kappa_{Hgg} g_{Hgg}\, G_{\mu\nu}^a G^{a,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg}\, G_{\mu\nu}^a \widetilde{G}^{a,\mu\nu} \big] \\
& - \frac{1}{4} \frac{1}{\Lambda} \big[c_\alpha \kappa_{HZZ}\, Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ}\, Z_{\mu\nu} \widetilde{Z}^{\mu\nu} \big] \\
& - \frac{1}{2} \frac{1}{\Lambda} \big[c_\alpha \kappa_{HWW}\, W_{\mu\nu}^+ W^{-\mu\nu} + s_\alpha \kappa_{AWW}\, W_{\mu\nu}^+ \widetilde{W}^{-\mu\nu} \big] \\
& - \frac{1}{\Lambda} c_\alpha \big[\kappa_{H\partial\gamma}\, Z_\nu \partial_\mu A^{\mu\nu} + \kappa_{H\partial Z}\, Z_\nu \partial_\mu Z^{\mu\nu} + \big(\kappa_{H\partial W}\, W_\nu^+ \partial_\mu W^{-\mu\nu} + h.c. \big) \big] \Big\} X_0
\end{aligned}
$$