Teaching PROFESSOR new math

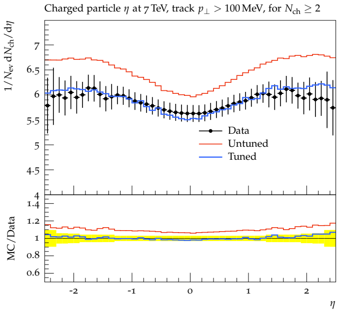
Holger Schulz, Juliane Müller, Sven Leyffer, Anthony Austin,
Stephen Mrenna

CHEP 2018 Sofia, 11 July 2018

U.S. DEPARTMENT OF ENERGY | Office of Science
Fermilab
University of CINCINNATI
Argonne NATIONAL LABORATORY
BERKELEY LAB

Tuning

- Goal: best possible physics prediction of MC generator
- Realistic events contain physics at low scales where perturbation breaks down
- Rely on model assumptions that introduce many parameters
- Need to find “meaningful” settings
- Can be done manually but hard to do on a reasonable time-scale because of MC run-time and dimensionality of problem

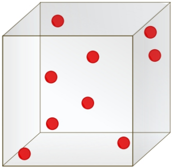


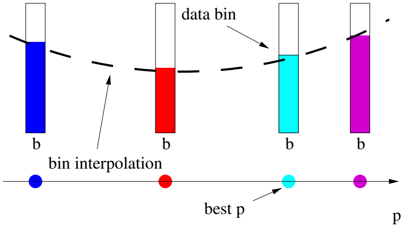
Tuning with Professor in a nutshell

- Random sampling: N parameter points in n-dimensional space
- Run generator and fill histograms (e.g. Rivet) trivial parallel
- Polynomial approximation *per bin*
- Construct goodness-of-fit measure

$$\phi^2(\vec{p}) = \sum_b w_b^2 \cdot \frac{(f_b(\vec{p}) - \mathcal{R}_b)^2}{\Delta^2}$$

- and numerically *minimise* with **`iminuit`**





In the following this will be called the “inner optimisation” problem

Inner and outer optimisation

- Incompatible datasets and mismodelling in MC generator necessitate introduction of tuning weights w_b
- Adjusting the weights has so far been a manual procedure
- The user would iteratively run the "inner optimisation" and look at resulting plots
- We propose an automated procedure for this "outer optimisation":
  - Write goodness-of-fit in terms of histograms/observables
  - The parameter space is now the observable-weight space
  - Inner optimisation yields best fit point, $\hat{p}$, for given

$$\{w_{\mathcal{O}}\}$$

  - $\hat{p}$ is used to evaluate an objective function for the outer optimisation

$$\phi^2(\vec{p}|\{w_{\mathcal{O}}\}) = \sum_{\mathcal{O}=1}^{N} w_{\mathcal{O}}^2 \cdot \sum_{b\in\mathcal{O}} \frac{(f_b(\vec{p}) - \mathcal{R}_b)^2}{\Delta^2}$$

Portfolio objective

- For given $\hat{p}$, we can calculate the per-observable goodness-of fit

$$\nu_{\mathcal{O}}(\hat{p}|\,\{w_{\mathcal{O}}\}) = \frac{1}{N_{\text{bins}}(\mathcal{O})} \sum_{b\in\mathcal{O}} \frac{(f_b(\hat{p}|\,\{w_{\mathcal{O}}\}) - \mathcal{R}_b)^2}{\Delta^2},\ \mathcal{O} = 1,\ldots,N$$

- With N such measures, we can calculate mean and standard deviation

$$\mu(w_{\mathcal{O}},\vec{p}^*) = \frac{1}{N}\sum_{\mathcal{O}=1}^{N} \nu_{\mathcal{O}}(w_{\mathcal{O}},\vec{p}^*)$$

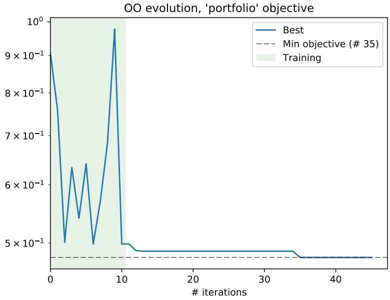
$$\sigma^2(w_{\mathcal{O}},\vec{p}^*) = \frac{1}{N}\sum_{\mathcal{O}=1}^{N} \left[\nu_{\mathcal{O}}(w_{\mathcal{O}},\vec{p}^*) - \mu(w_{\mathcal{O}},\vec{p}^*)\right]^2$$

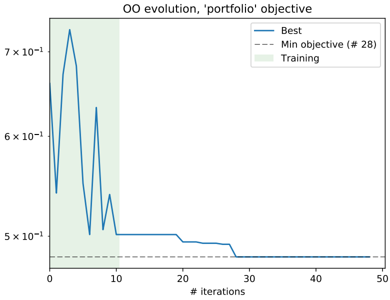
- And construct an objective function to minimise

$$\min_{w_{\mathcal{O}}\in[0,1]} \lambda\mu(w_{\mathcal{O}},\vec{p}^*) + \sigma^2(w_{\mathcal{O}},\vec{p}^*), \quad \text{s.t.} \sum_{\mathcal{O}=1}^{N} w_{\mathcal{O}} = 1.$$

Outer optimisation

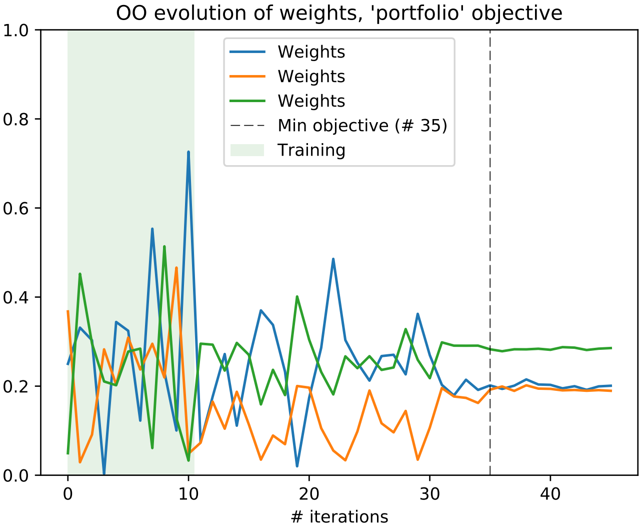
- Minimisation of portfolio objective is iterative
- We train a radial basis function and use it to walk through the weight space
- Convergence is fast but depends on initial guess → multi-start approach (that's ok since inner optimisation is really fast)





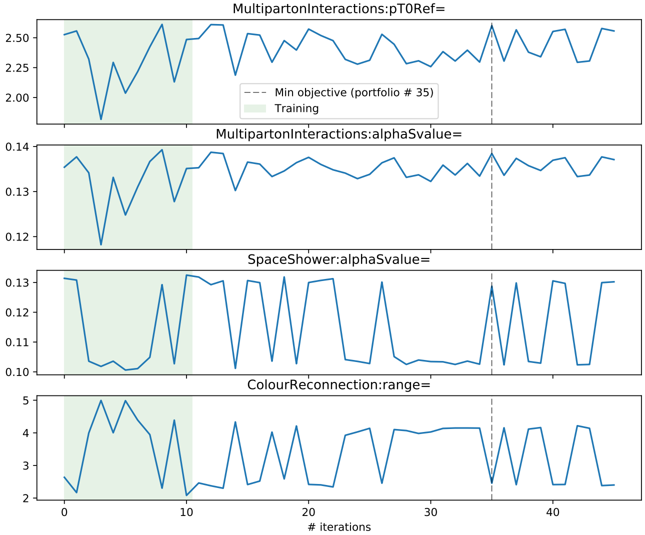
Evolution of weights

- This plot shows the $\{w_{\mathcal{O}}\}$ of the outer optimisation
- Controlplot to check that weight space is reasonably sampled

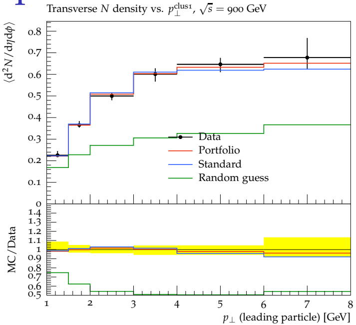


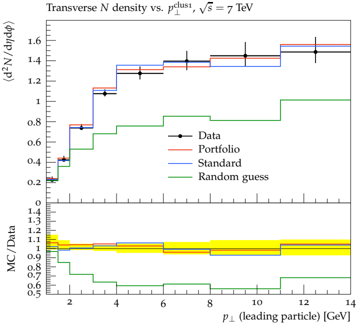
Evolution of inner optimisation

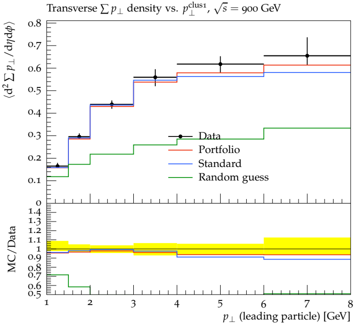
- This plot shows the $\hat{p}$ of the inner optimisation
- Shows the correlation of parameters

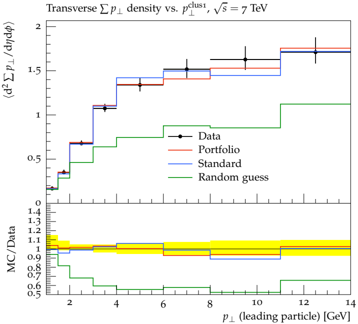


Comparison of results



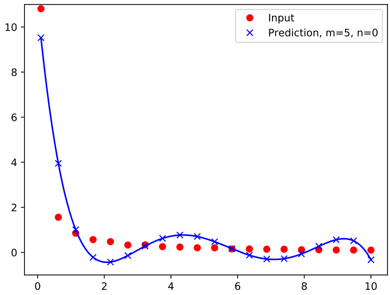


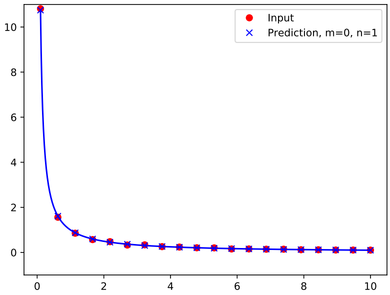




Rational approximation

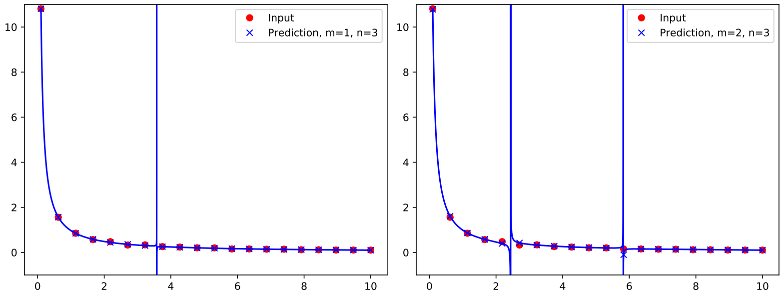
- Polynomial approximation does not capture $1/x$ behaviour well
- E.g. masses in propagators
- $\rightarrow$ Multivariate rational approximation $f(\vec{p}) = g(\vec{p})/h(\vec{p})$
- With g, h being polynomials of order m, n





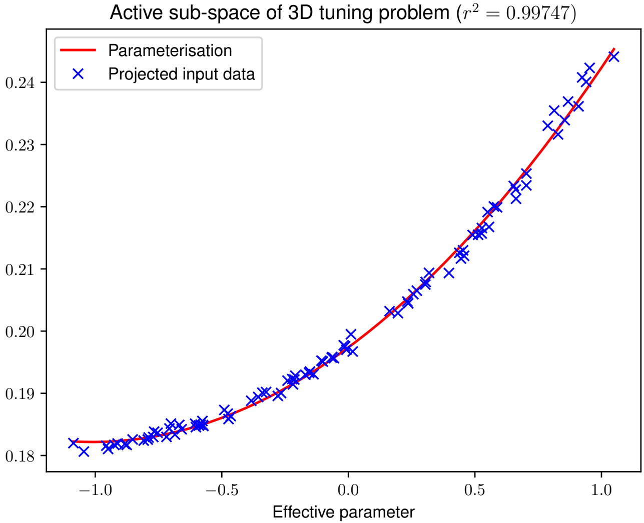
Spurious poles

- Code works in principle but: spurious poles
- For numerical reasons, denominator polynomial can have roots
- Happens if “wrong” m, n are chosen.
- In 1D to easy find roots, much harder in higher dimensions



Active sub-space

- Due to high correlation of parameters, we may be able to reduce the dimensionality by finding active sub-space
- Method is based on gradient sampling and Eigenvalue problem



Summary

- We have a working solution for a more automated tuning:
  - Outer optimisation loop in the weight space
  - Minimisation of portfolio objective function
  - We will assess performance and scaling with dimensions of inner and outer parameter space

- Algorithm for multivariate rational approximation
  - Initial success but spurious poles are a significant problem
  - Active sub-space approach might be a way out