AutoTunes

1. Motivation
2. Professors
3. Reducing the dimension
4. Results

Work with Leif Gellersen

Motivation

Motivation:
- Retuning the Herwig dipole shower
- Top Tunathon workshop in Manchester
- Is it a solved problem?
- Reduce human interaction

Goals:
- Keep possibility to interpret result
- Set weights for single observables automatically
- Set up flexible framework to test new weights
- Tune many parameters of the model -> high dim.
- Clear, simple and reproducible definition what was done in tune.

The Problem

Tuning is a very physicist dependent procedure:

- What data to tune to?
  - knowledge of data (Do we know some data is not „good"?)
  - Some data in rivet is same data (e.g. differential and integrated, updates…)
  - If many groups at LEP measure the same observable, does this enhance the weight in the next tune? (Only use one b-frag?)

- Correlations in the model?
  - knowledge of model (Should be given… but e.g. influence of 2->3 jet rate on b-fragmentation?)

Professor

Professor:

- Tool to compare and tune theory output to data
- First fit polynomials P for each bin i and parameter vector x to theory prediction T:

$$P_i(\vec{x}) = A_i + B_i^{\mu} x^{\mu} + C_i^{\mu\nu} x^{\mu} x^{\nu} + \ldots \quad \rightarrow \quad \min(\chi^2(T_i(\vec{x}), P_i(\vec{x}))) \;\; \forall \, \vec{x} \, \in \, [\vec{x}_{\min}, \vec{x}_{\max}]$$

- Then minimize $\chi^2(P(\vec{x}), D, \vec{w})$ with D=data points, w=weights for data points.
- So Chi^2 depends on weights assigned to observables
- Polynomial approach good if observables predictions smooth and local -> Taylor expansion.. $[\vec{x}_{\min}, \vec{x}_{\max}]$ not to big!
- Polynomial approach grows fast with number of dimensions. But models have large number of parameters.

Reduce the dimension of the problem

Aim to reduce number of dimensions in a way that maximizes correlations between parameters. Professor can then work on the subspace of x to find best fit values.

How can we get information about the correlation and if a parameter matters for an observable?

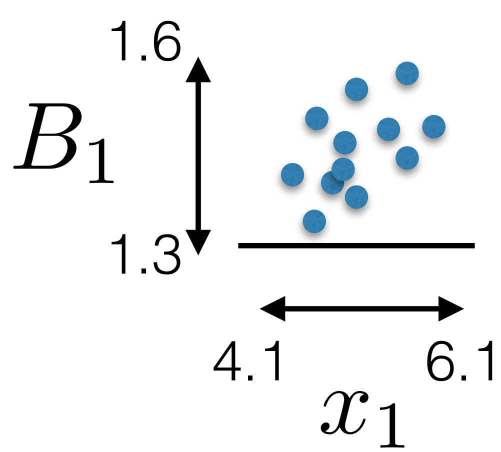
If we find correlations and split the dimensions in subspaces, how can we weight the important bins in this sub-tune.

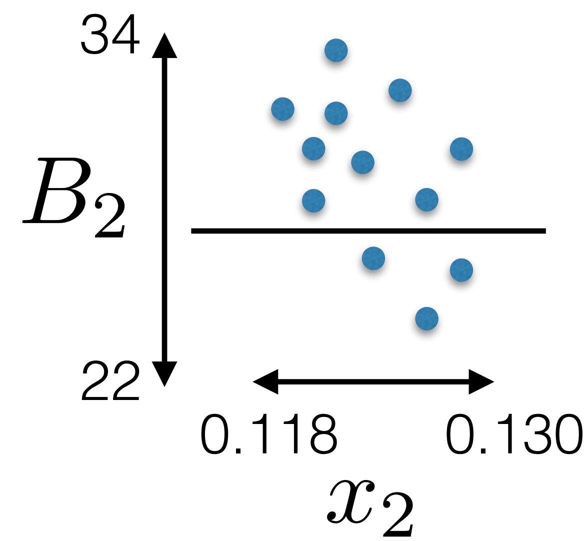
Summarize:

- How to split the d-dim space?
- How to tune sub-space?

Reduce the dimension of the problem

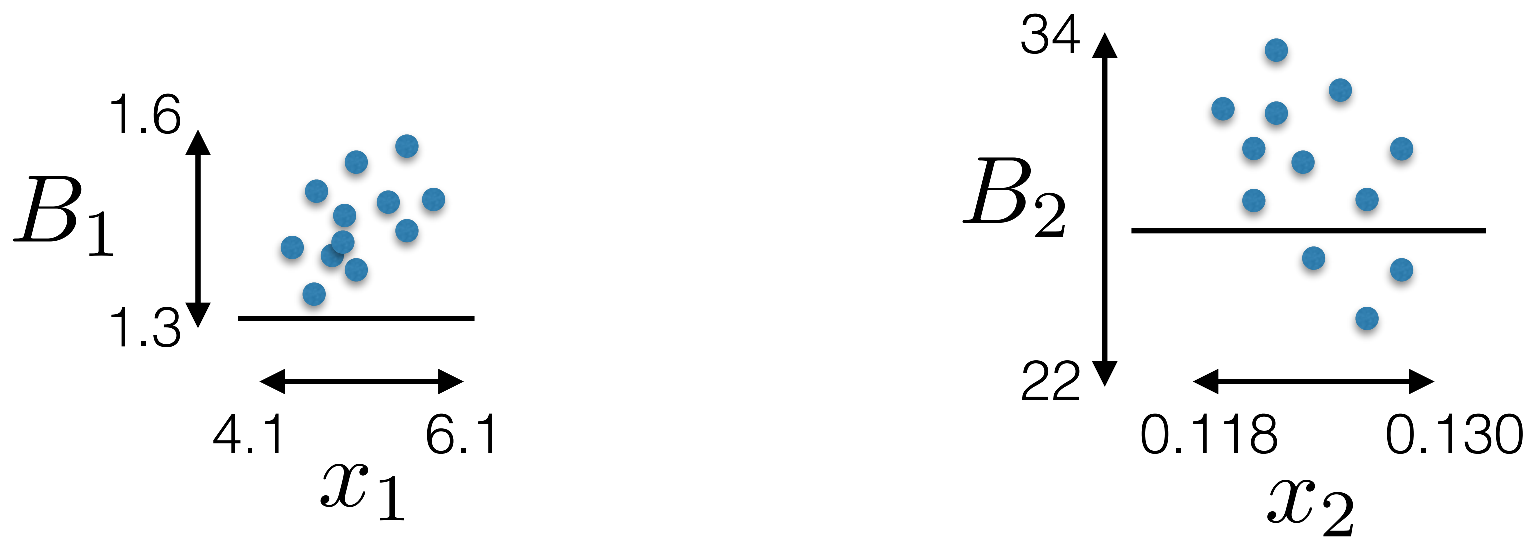
Lets start with observables/bins:





Reduce the dimension of the problem

Lets start with observables/bins:

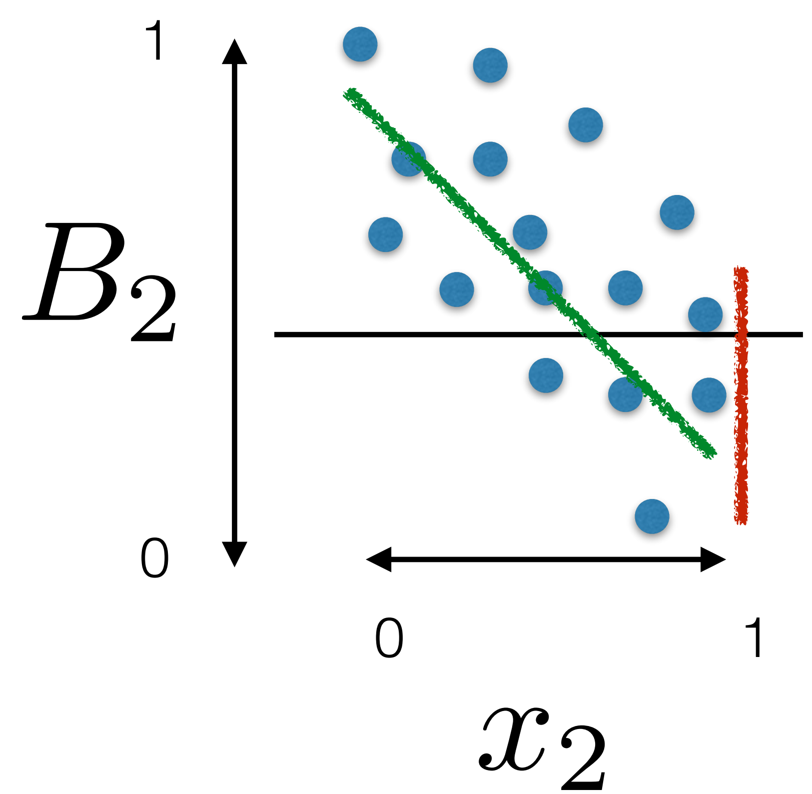


Reduce the dimension of the problem

Slope and scatter:

Slope indicates the dependence on parameter.

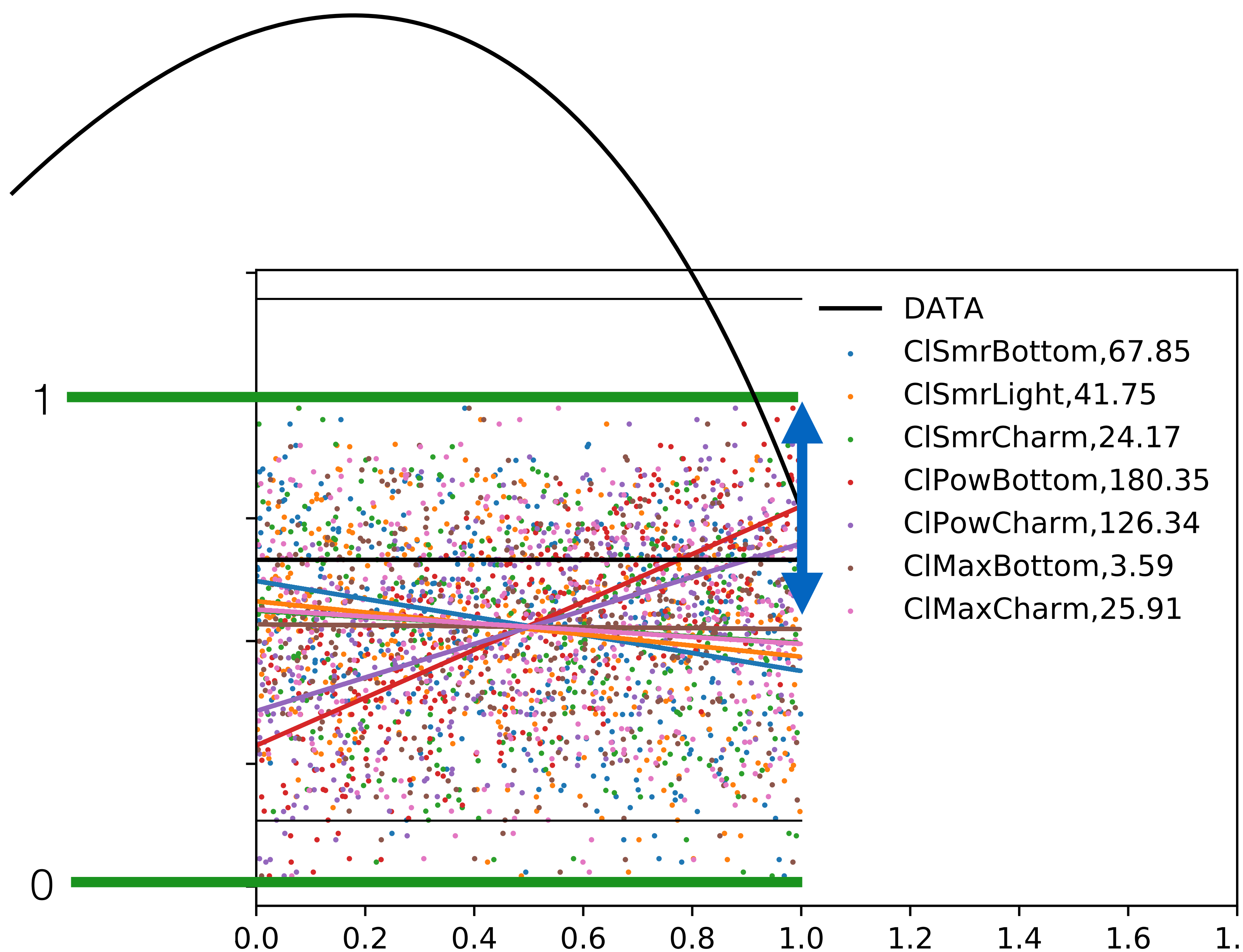
Scatter either by not-linear behavior
and/or
other parameters are important in this.



Reduce the dimension of the problem

Real life Example (arbitrary bin):

- Color of points misleading
- ~1000 random theory points
- height normalization by spread of random points.
- Here only 7 dims.
- If bin depends on more than one parameter, the slope of the a parameters are affected by the others.
- Don't care about data. This is just to get theory dependence.



Another Example:

$$\begin{aligned}
B_0 &= x_1 + x_2 + \lambda x_3 + \lambda x_4 \\
B_1 &= x_1 + x_2 + x_3 + \lambda x_4 \\
B_2 &= \lambda x_1 + x_2 + x_3 + \lambda x_4 \\
B_3 &= \lambda x_1 + \lambda x_2 + x_3 + x_4
\end{aligned}
\qquad
\begin{aligned}
S_0 &= (1, 1, \lambda, \lambda) \\
S_1 &= (1, 1, 1, \lambda) \\
S_2 &= (\lambda, 1, 1, \lambda) \\
S_3 &= (\lambda, \lambda, 1, 1)
\end{aligned}$$

```
B_0&= \p
B_1&= \p
B_2&= \l
B_3&= \l

S_0&= (1
S_1&= (1
S_2&= (\
S_3&= (\
```

Assume 4 bins that are described by 4 parameters.

Same parameter important in various bins!

$$\begin{aligned}
S_0 &= (1, 1, \lambda, \lambda) \\
S_1 &= (1, 1, 1, \lambda) \\
S_2 &= (\lambda, 1, 1, \lambda) \\
S_3 &= (\lambda, \lambda, 1, 1)
\end{aligned}
\qquad
\begin{aligned}
S_0 &= (1, 1, \lambda, \lambda) \\
S_1 &= (1, 1, 1, \lambda) \\
S_2 &= (\lambda, 1, 1, \lambda) \\
S_3 &= (\lambda, \lambda, 1, 1)
\end{aligned}$$

Note:
By normalizing to spread we suppress influence of bins with bad resolution.

High discriminating power if lambda small

Another Example:

$$
\begin{aligned}
B_0 &= x_1 + x_2 + \lambda x_3 + \lambda x_4 \\
B_1 &= x_1 + x_2 + x_3 + \lambda x_4 \\
B_2 &= \lambda x_1 + x_2 + x_3 + \lambda x_4 \\
B_3 &= \lambda x_1 + \lambda x_2 + x_3 + x_4
\end{aligned}
$$

$$
\begin{aligned}
S_0 &= (1, 1, \lambda, \lambda) \\
S_1 &= (1, 1, 1, \lambda) \\
S_2 &= (\lambda, 1, 1, \lambda) \\
S_3 &= (\lambda, \lambda, 1, 1)
\end{aligned}
$$

$$
\begin{aligned}
S_0 &= (1, 1, \lambda, \lambda) \\
S_1 &= (1, 1, 1, \lambda) \\
S_2 &= (\lambda, 1, 1, \lambda) \\
S_3 &= (\lambda, \lambda, 1, 1)
\end{aligned}
$$

$$N_i^\mu = S_i^\mu \circ \Big(\sum_i S_i^\mu\Big)^{-1}$$

Elementwise:

- We normalized the x-range!!
- Parameter that is important in many observables should be tuned last. See tune-step selection

Split/chunk the d-dimensions

$$\begin{aligned} B_0 &= x_1 + x_2 + \lambda x_3 + \lambda x_4 \\ B_1 &= x_1 + x_2 + x_3 + \lambda x_4 \\ B_2 &= \lambda x_1 + x_2 + x_3 + \lambda x_4 \\ B_3 &= \lambda x_1 + \lambda x_2 + x_3 + x_4 \end{aligned}$$

$$\begin{aligned} S_0 &= (1, 1, \lambda, \lambda) \\ S_1 &= (1, 1, 1, \lambda) \\ S_2 &= (\lambda, 1, 1, \lambda) \\ S_3 &= (\lambda, \lambda, 1, 1) \end{aligned}$$

$$N_i^\mu = S_i^\mu \circ (\sum_i S_i^\mu)^{-1}$$

Combination vectors:

$$\begin{aligned} J_0 &= (1, 1, 0, 0) \\ J_1 &= (1, 0, 1, 0) \\ J_2 &= (1, 0, 0, 1) \\ J_3 &= (0, 1, 1, 0) \\ J_4 &= (0, 1, 0, 1) \\ J_5 &= (0, 0, 1, 1) \end{aligned}$$

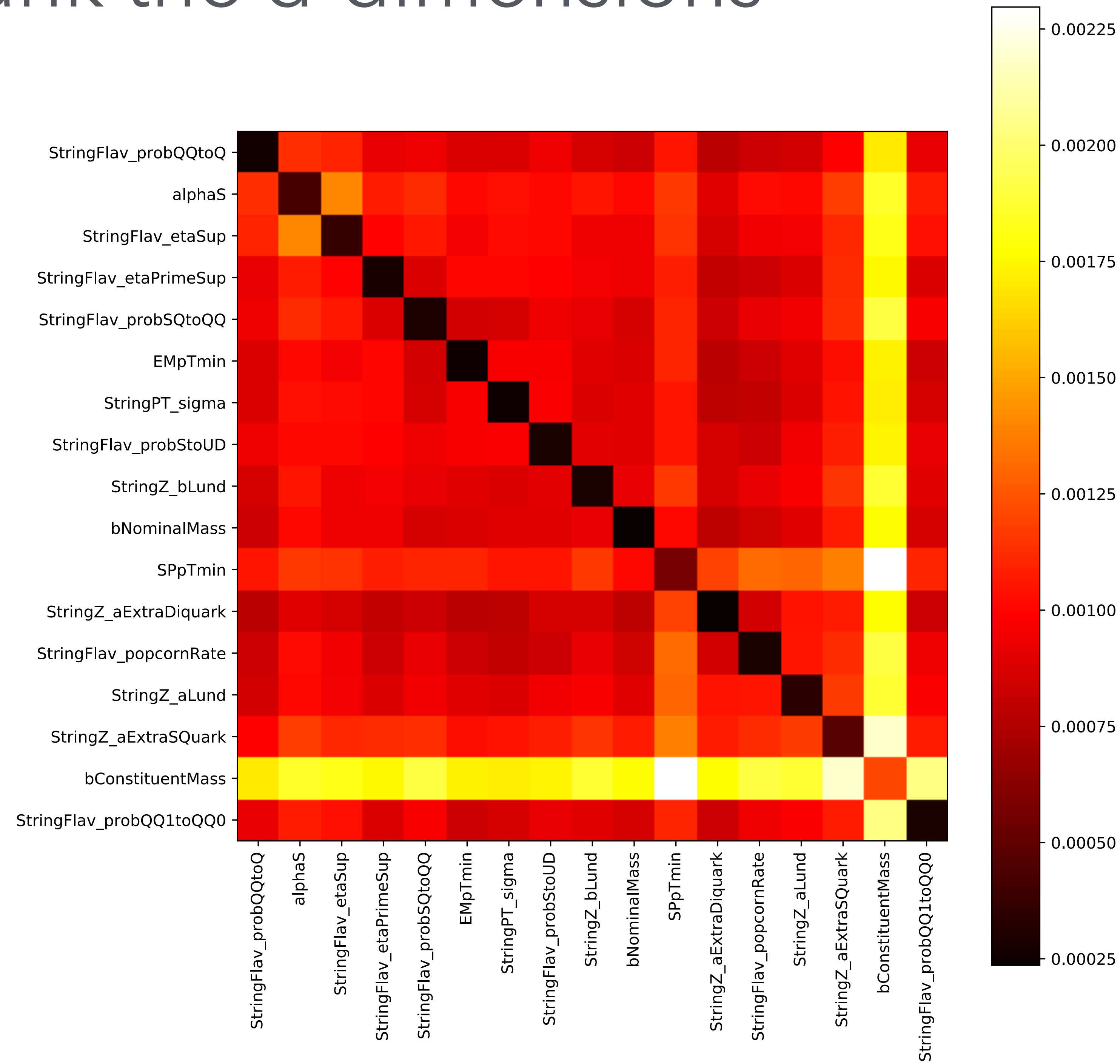
Maximize the projection on subspaces :

$$w_k = \sum_i (N_i J_k)^2$$

Correlates parameters that have influence in same bins.

$\max(w_k)$ defines first sub-tune.
Next sub-tunes orthogonal to J_k.

Split/chunk the d-dimensions



This is the Herwig shower with Lund stings..
Interfaced with modified version of TheP8I interface used for Ariadne.

Split/chunk the d-dimensions

The weight for observables should be large in a sub-tune if it „points" in the direction of the currently tested/tuned subset of directions:

Currently testing various weight combinations (numerically and influence on Chi2):

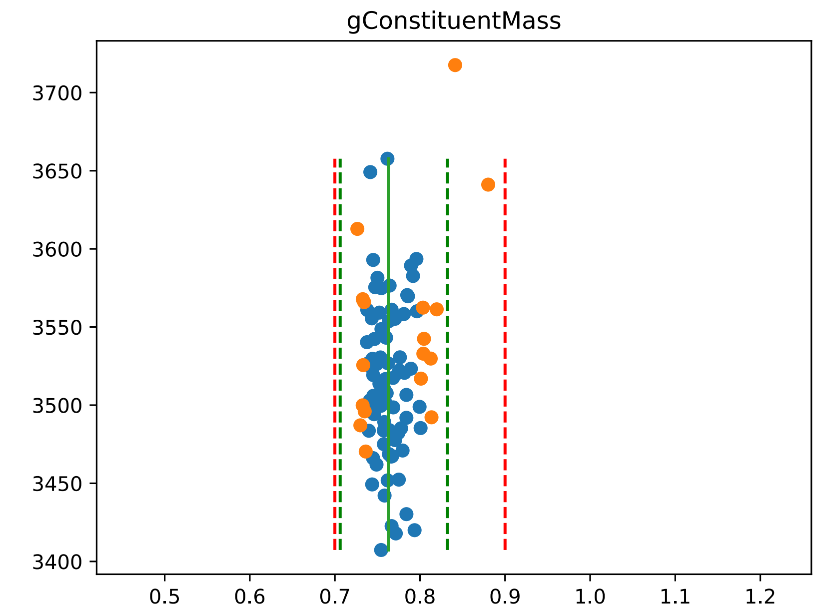
$$W_i = (N_i J_{\mathrm{Step1}})^2$$

$$W_i = (N_i J_{\mathrm{Step1}})$$

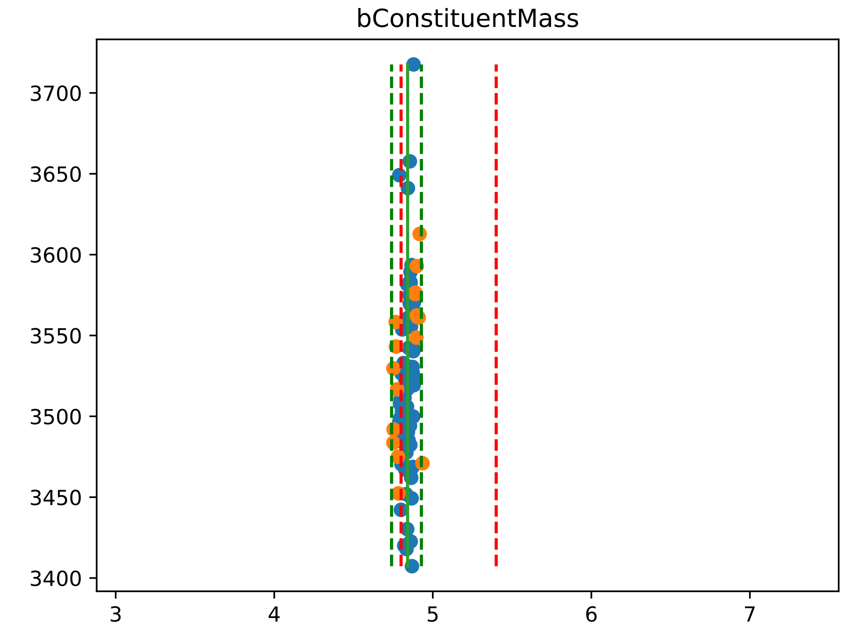
$$W_i = (N_i J_{\mathrm{Step1}})^2 / |N_i|$$

Bins that „point" in same directions can be correlated.
These weights could have influence if we want to reduce impact of similar measurements.

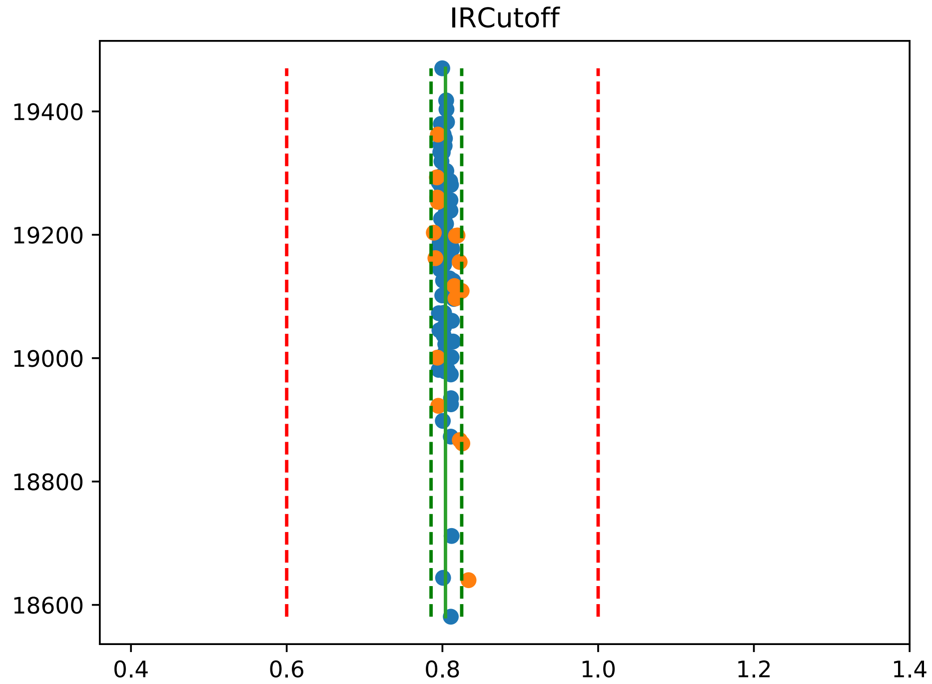
Results



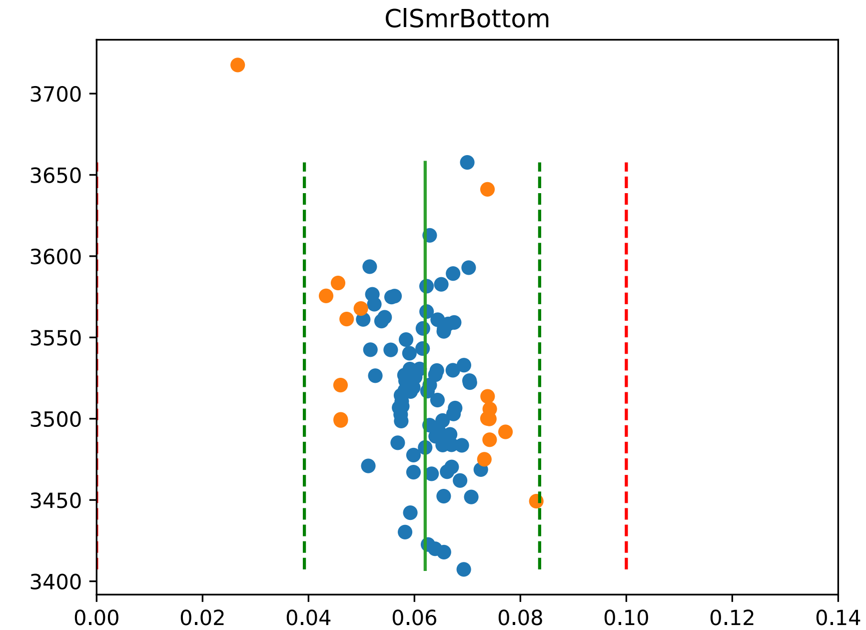
gConstituentMass
3700
3650
3600
3550
3500
3450
3400
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2



bConstituentMass
3700
3650
3600
3550
3500
3450
3400
3
4
5
6
7



IRCutoff
19400
19200
19000
18800
18600
0.4
0.6
0.8
1.0
1.2
1.4



ClSmrBottom
3700
3650
3600
3550
3500
3450
3400
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14

The End

Thank you!