Analysis Tools

Beautiful_And_Simple_Drawing_Atificer (BASDA)

Yan Wang

05-03-2018

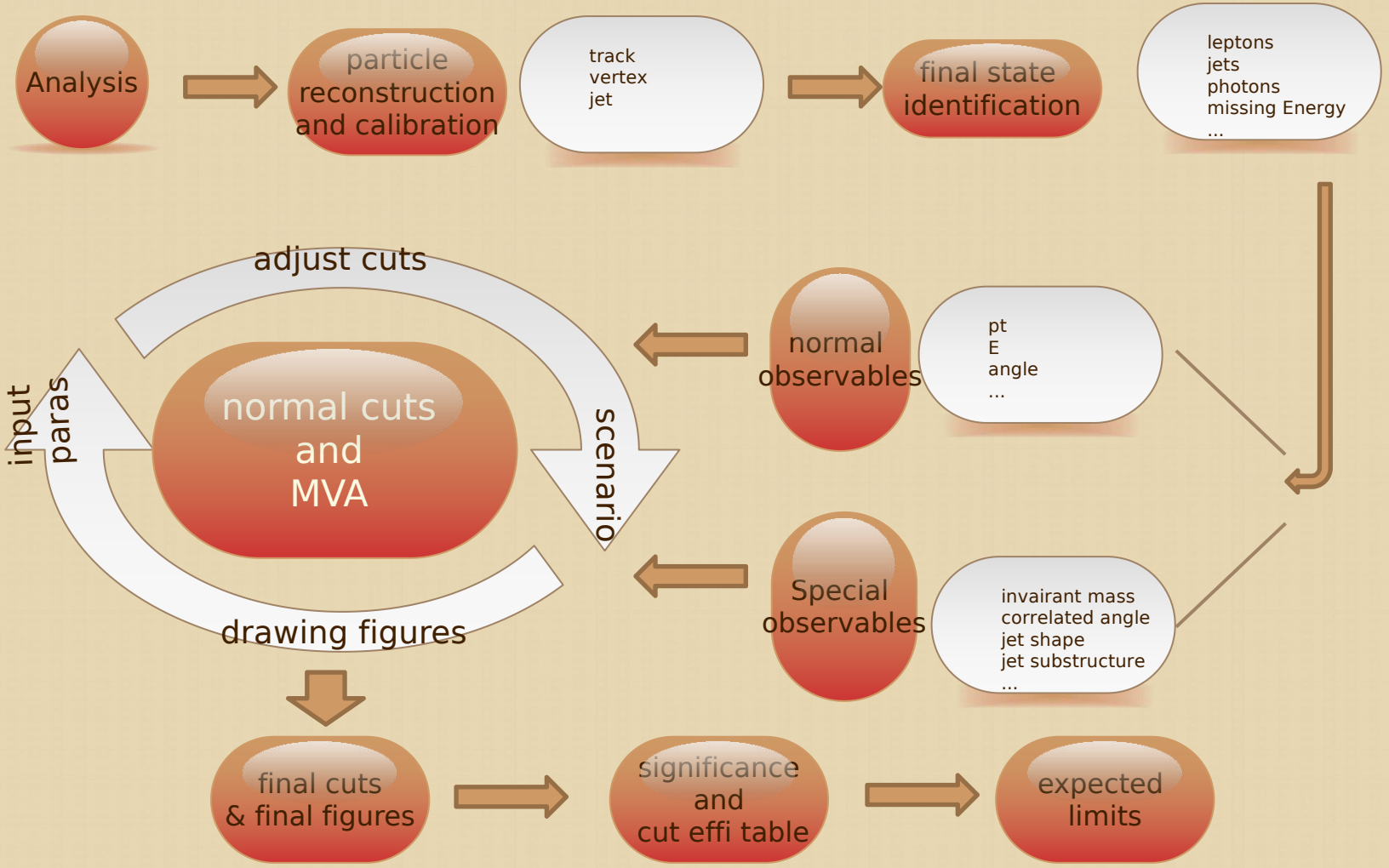
Physics Benchmarks

WG	Process	Physics	Detector	ECM	Who
Higgs & EW	**H->bb/cc/gg**	**BR**	**c-tag, b-tag, JER**	**500 GeV**	**NN + NN**
	H->bb	mass	JER, JES	500 GeV	Ali Ebrahimi (10%) + Junping Tian
	ee->tautau	A_FB, tau-pol, A_LR	tau-reco	500 GeV	Daniel Jeans + **NN**
	H->mumu	BR	momentum resolution	500 GeV	Shin-ichi Kawada + **NN**
	H->invisible	BR limit	JER, hermeticity	500 GeV	Yu Kato + **NN**
	WW->qqlv	MW, TGCs, beam pol.	JES, JER, electron, mu	500 GeV	Kostiantyn Shpak + **NN**
	vvqqqqq	**QGCs**	**JES / JER**	**1 TeV**	**Jakob Beyer + NN**
	gamma Z	**A_LR, sigma_tot, JES**	**photon, JER/JES, e, mu**	**500 GeV**	**NN + NN**
Top, Bottom & QCD	tt->bbqqqq	x-section, AFB	b-tag, vertex charge, PID	500 GeV	Amjad + **NN**
BSM	low deltaM Higgsinos	natural SUSY	low-p tracking, PID, hermeticity	500 GeV	Swathi Sasikumar + **NN**
	mono-photons	**WIMPs / WISPs**	**photon reco, BeamCal**	**500 GeV**	**NN + NN**
	Zh, mh < 125 GeV	limit on ZZh coupling	p res, e reco, JER, hermeticity	500 GeV	Yan Wang + **NN**

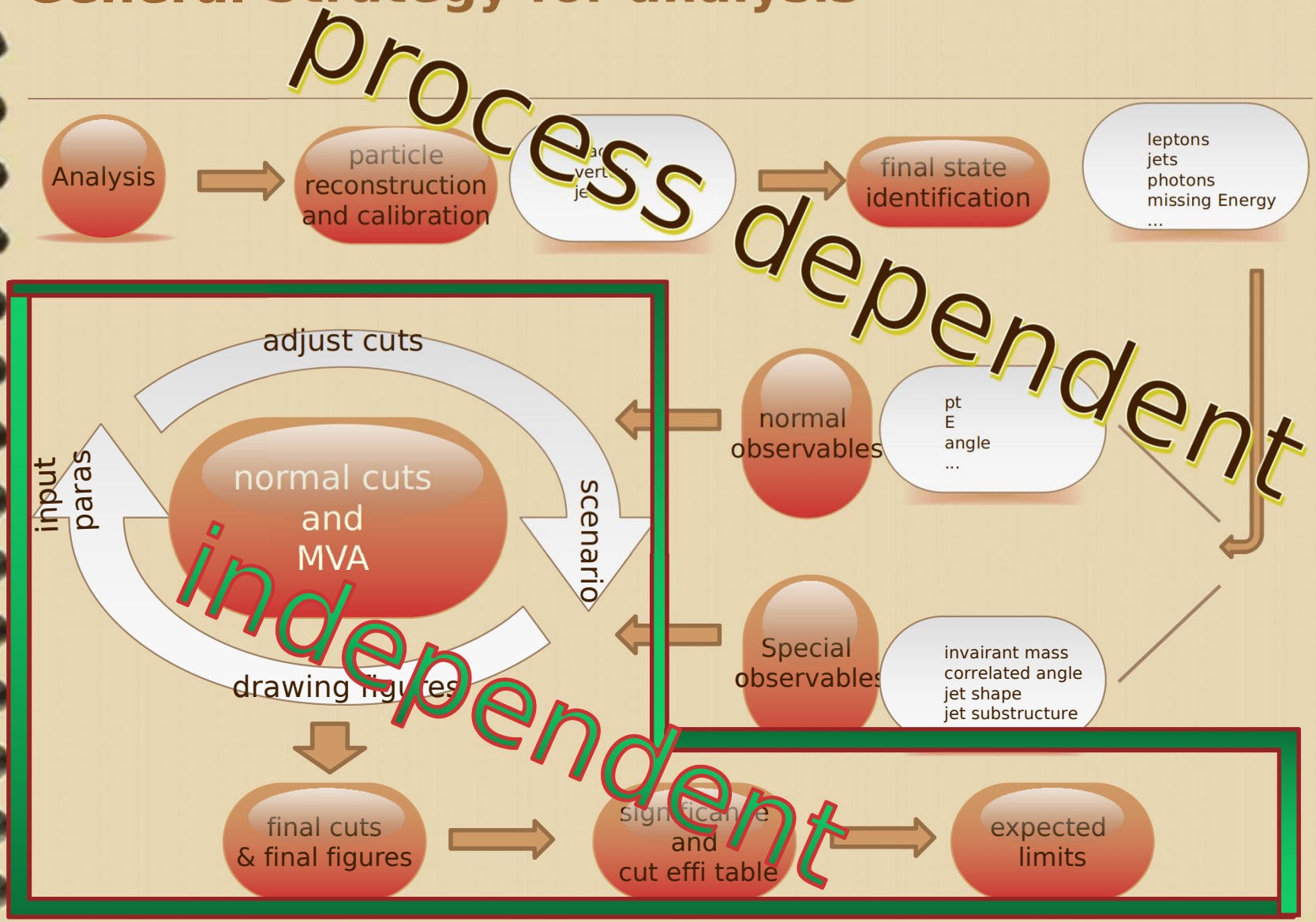
10

a standard analysis code will be helpful!

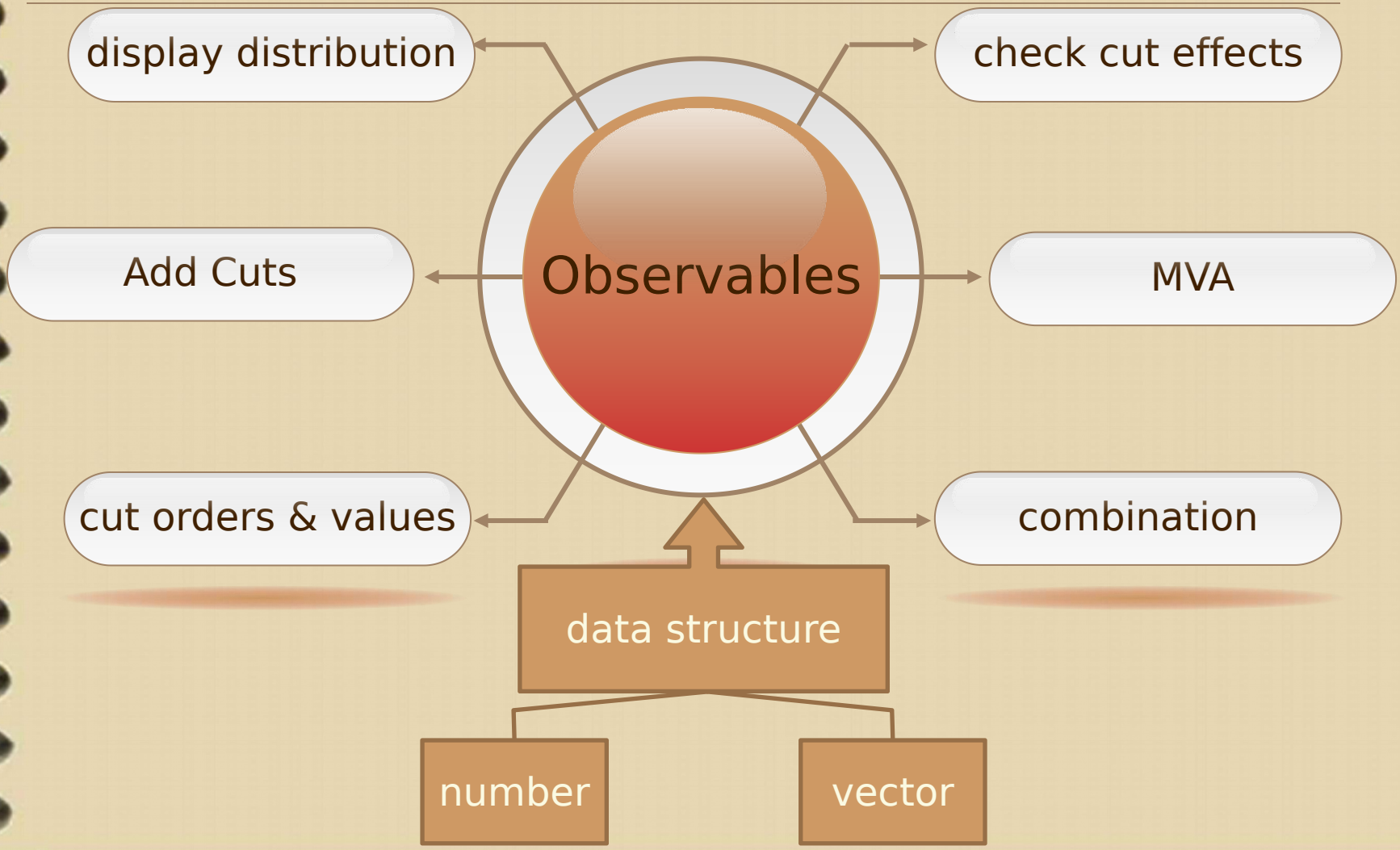
General Strategy for analysis



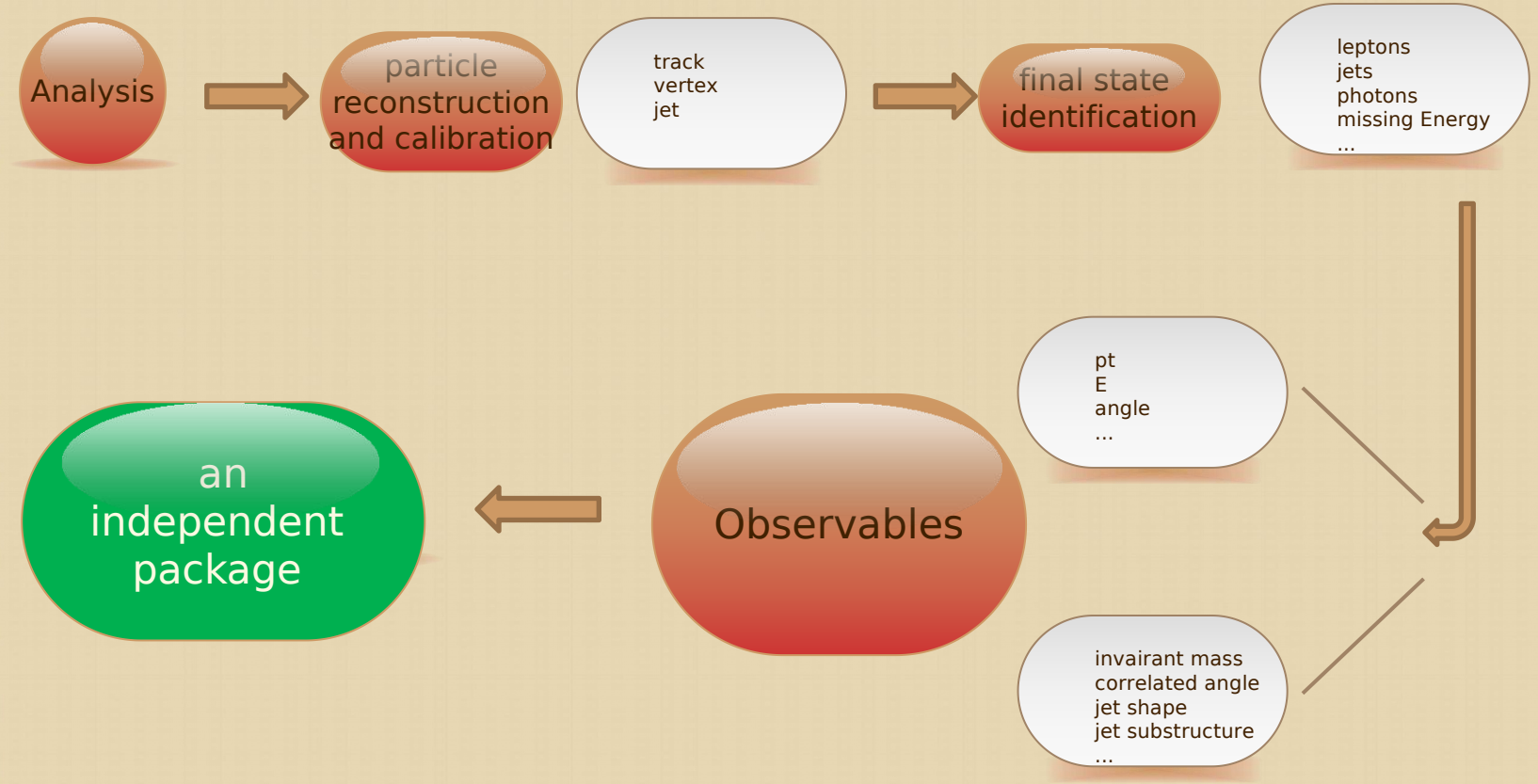
General Strategy for analysis



The key role


display distribution
check cut effects
Add Cuts
Observables
MVA
cut orders & values
combination
data structure
number
vector

General Strategy for analysis



Beautiful_And_Simple_Drawing_Atificer (BASDA)

The beautiful figure

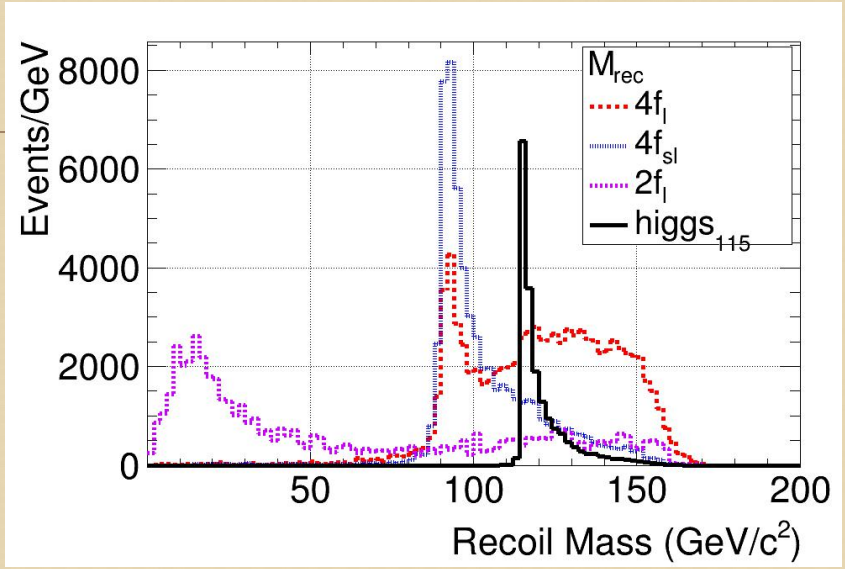
- plot templates
- easy to adjust
- once fixed, used forever

The simple operation

all with control files

Easy to repeat

only need to preserve control file.

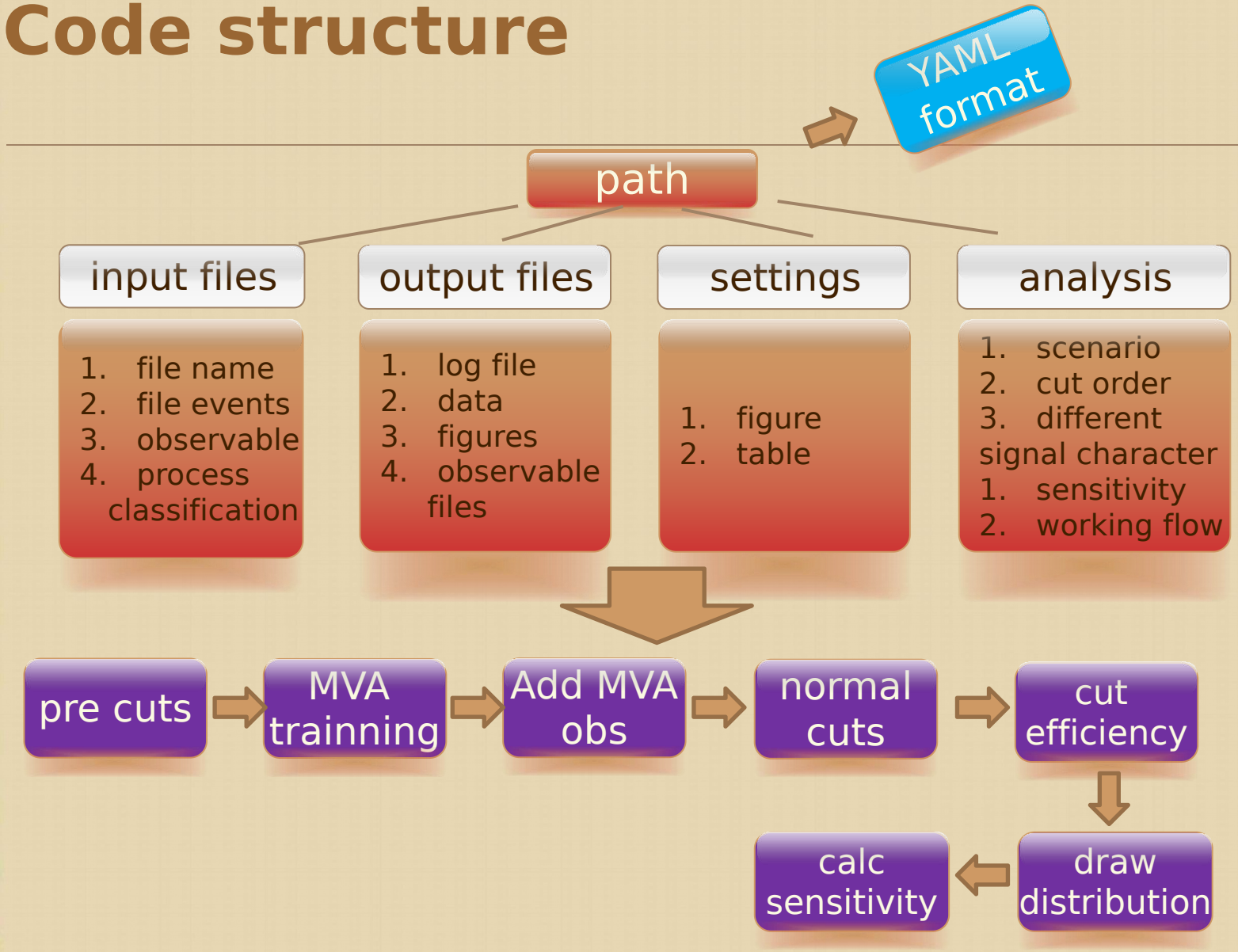


Release time from programing.

Focus on **Physics**!

Cross check

Code structure



YAML format

- YAML---YAML Ain't Markup Language
- a human-readable data serialization language
- in BASDA, only use simplest command

```
key :  value
```

or,

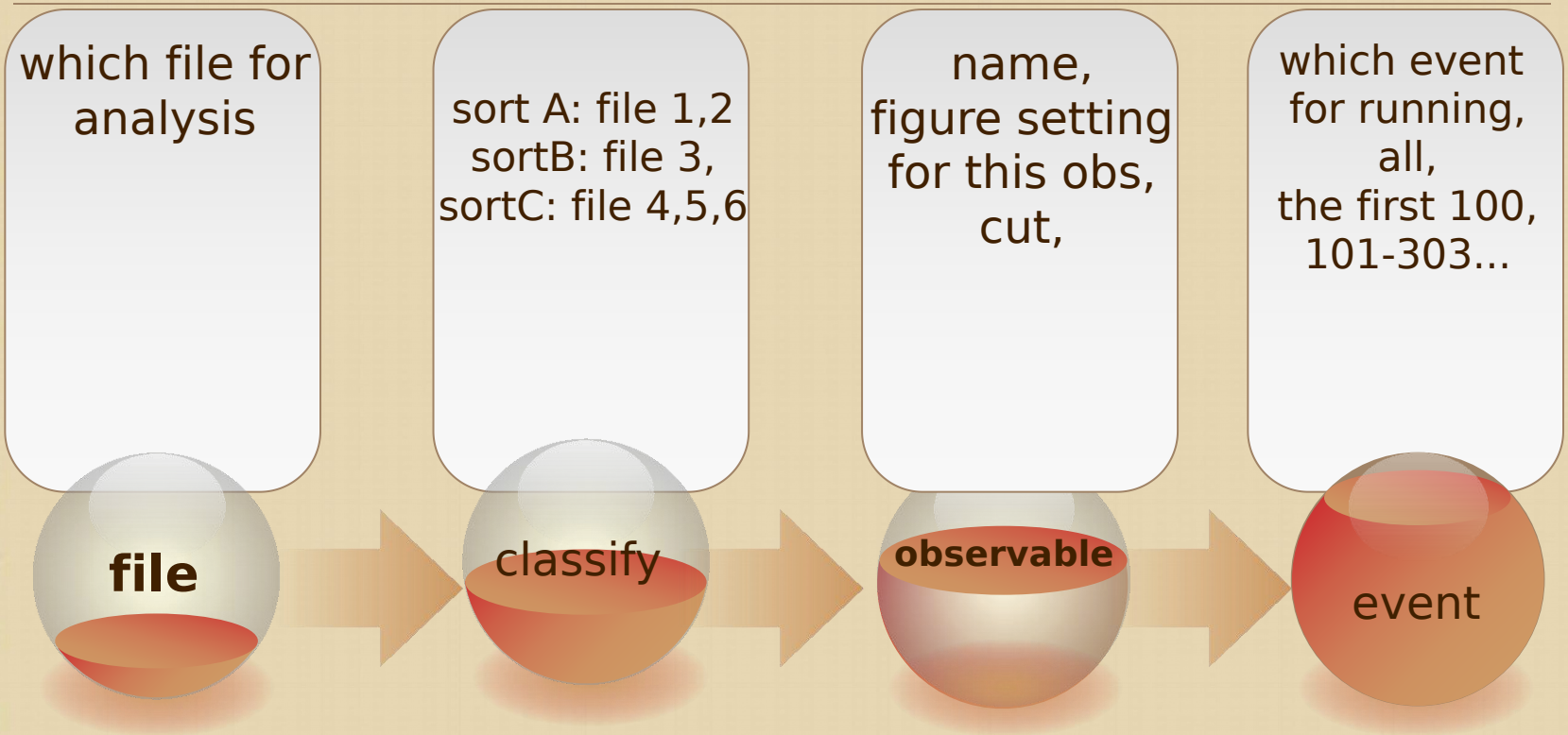
```
key :
    key :  value
    key :
            key :  value
```

or list,

```
key: [value1,value2, value3...]
```

note: tabs is forbiden in YAML. comments with #

input data


which file for analysis
sort A: file 1,2
sortB: file 3,
sortC: file 4,5,6
name, figure setting for this obs, cut,
which event for running, all, the first 100, 101-303...
file
classify
observable
event

file

vim file.dat

```
1  FILE_NUM            : 1
2  FILE_0              : /home/yancy/Code/Event_Light_Higgs/origin_data/xection_aa.dat
3
4  CHANNEL_NUM         : 1
5  FILE_DESCRIP_0      : all
6  FILE_DESCRIP_1      : nh
7  FILE_DESCRIP_2      : 4f_{zz}^{l}
8
9  Root_Head_Name      : datatest
10 Root_Head_BDT_Name : datatrain
```

channel used in this time

root tree name

xection file

vim xection_aa.dat

channel name

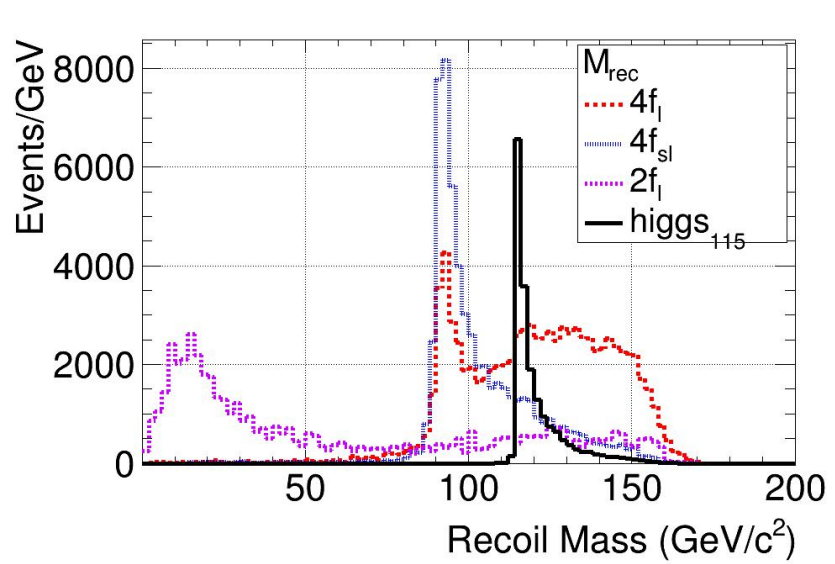
```
1  root file 10:
2      nh_{10}^{10}:
3      |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/nh_10_eL.pR.root   :   79.026852
4      |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/nh_10_eR.pL.root   :   50.415099
5      4f_{ww}^{l}:
6      |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_ww_l_eL.pR.root   :   1564.2091
7      |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_ww_l_eR.pL.root   :   14.691728
8      4f_{ww}^{sl}:
9      |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_ww_sl_eL.pR.root   :   18780.976
10     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_ww_sl_eR.pL.root   :   172.73264
11     4f_{zz}^{l}:
12     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_zz_l_eL.pR.root   :   17.96041
13     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_zz_l_eR.pL.root   :   9.506102
14     4f_{zz}^{sl}:
15     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_zz_sl_eL.pR.root   :   422.1429
16     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_zz_sl_eR.pL.root   :   3.52633
17     4f_{zzww}^{l}:
18     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_zzww_l_eL.pR.root   :   1636.0362
19     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_zzww_l_eR.pL.root   :   53.955513
20     4f_{zzww}^{sl}:
21     |    no such class:  0
22     4f_{szee}^{l}:
23     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_szee_l_eL.pR.root   :   1084.0873
24     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_szee_l_eR.pL.root   :   1019.5228
25     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_szee_l_eL.pL.root   :   1009.6032
26     |    /home/yancy/Code/Data/Event_Light_Higgs/origin_data/results_10/4f_szee_l_eR.pR.root   :   1008.4063
```

signal property

observables in one channel

cross section

classify



vim Bkg_Sort.dat

```
1  sort_num    : 3
2
3  each_sort_num_0 : 1
4  each_sort_num_1 : 2
5  each_sort_num_2 : 2
6  #each_sort_num_3 : 2
7  #each_sort_num_4 : 1
8
9  each_sort_name_0 : nh
10 each_sort_name_1 : 4f
11 each_sort_name_2 : 2f
12 #each_sort_name_3 : aa
13 #each_sort_name_4 : h
14
15 nh_0: 115;higgs_{115}
16
17 4f_0 : l
18 4f_1 : sl
19
20 2f_0 : l
21 2f_1 : bha
22
23 #aa_0 : 2f
24 #aa_1 : minijet
25
26 #h_0 : e2;Higgs_{125}
```

alias

comment

Event

vim Event.dat

```
1 #show many first events want to be analyses
2 first_events_num   : -1
3
4 #which events want to analysse
5 special_events_num : -1
6
```

1. **when first =-1, special=-1 => use all events**
2. **when first = n, special=-1 => use first n events**
3. **when first =-1, special= n => use the n-th event**
4. **when first =n1, special=n2 => use events from n1 to n2**

where if n > total events, n = total events, and n1 < n2

observables

obs name in code

obs name in input file

cut for this obs

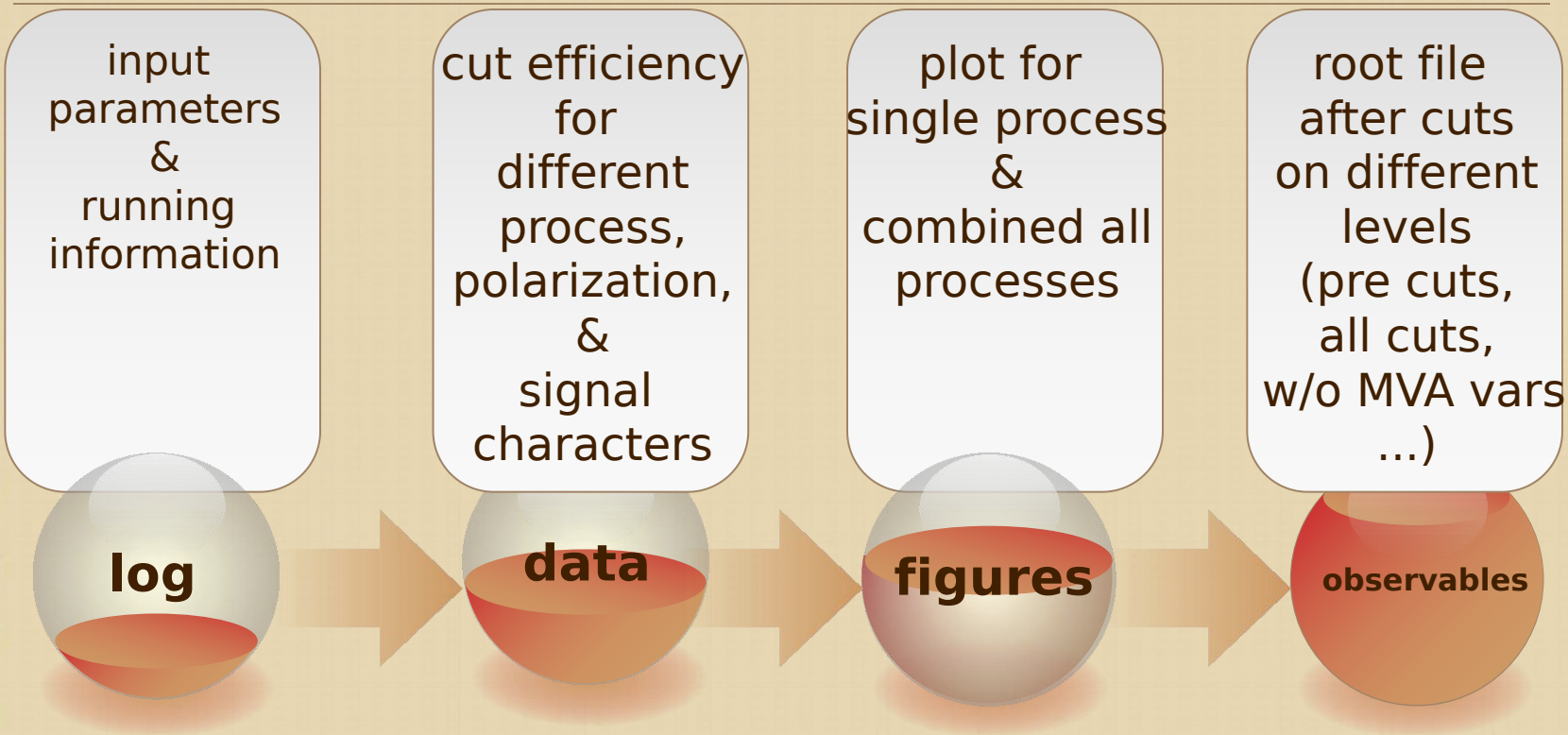
plot para

whether use this obs as MVA input

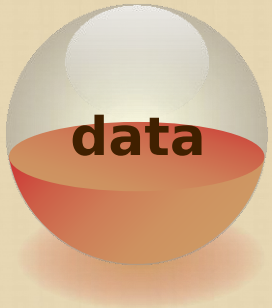
vim Var_nh_30.dat

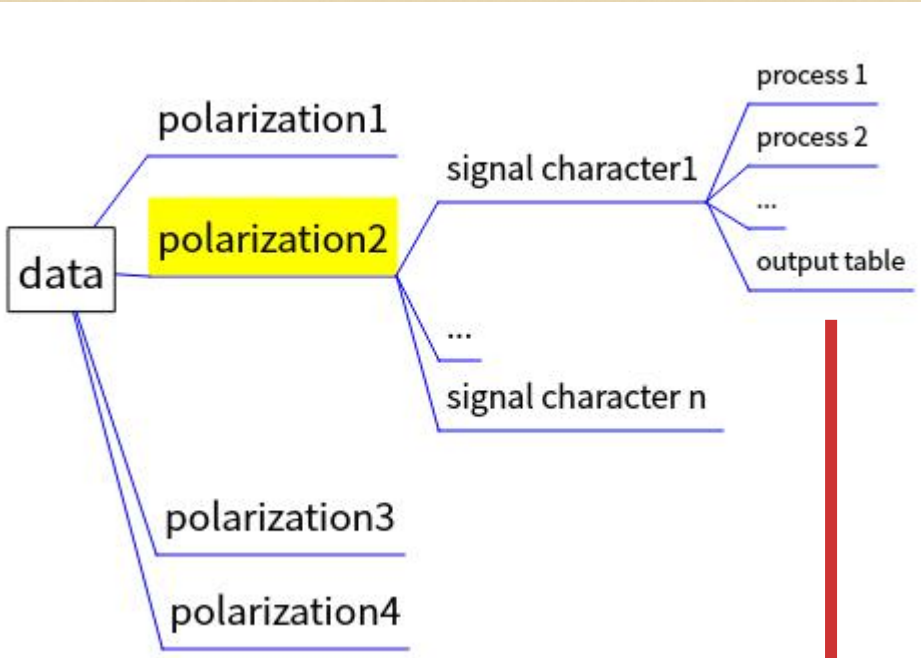
```
variable:
  v1:
    title_name           : "po_muon_kcut_zmass"
    cut_switch           : true
    cut_min              : 73.0
    cut_max              : 120.0
    plot_switch          : true
    latex_name           : "M_{l^{+}l^{-}}"
    Canvas_name          : "c1"
    Canvas_width         : 1000
    Canvas_height        : 700
    leg_left             : 0.75
    leg_up               : 0.7
    leg_right            : 0.9
    leg_down             : 0.9
    leg_header           : "e^{+}e^{-}->#mu^{+}#mu^{-} + X @ 250 GeV"
    xaxis_name           : "M_{Z} (GeV/c^{2})"
    xaxis_bin            : 100
    xaxis_min            : 51
    xaxis_max            : 131
    yaxis_name           : "Events"
    yaxis_bin            : 0
    yaxis_min            : 0
    yaxis_max            : 0
    log_yaxis            : true
    normalization_switch : false
    BDT                  : true
  v2:
    title_name           : "po_muon_kcut_zpt"
    cut_switch           : true
    cut_min              :   10.0
    cut_max              :   120.0
    plot_switch          : true
    latex_name           : "P_{T}^{l^{+}l^{-}}"
    Canvas_name          : "c2"
    Canvas_width         : 1000
    Canvas_height        : 700
    leg_left             : 0.80
    leg_up               : 0.7
    leg_right            : 0.95
    leg_down             : 0.9
    leg_header           : "e^{+}e^{-}->#mu^{+}#mu^{-} + X @ 250 GeV"
    xaxis_name           : "P_{T}^{Z} (GeV/c^{2})"
    xaxis_bin            : 100
    xaxis_min            : 0
    xaxis_max            : 150
    yaxis_name           : "Events"
    yaxis_bin            : 0
    yaxis_min            : 0
    yaxis_max            : 0
    log_yaxis            : true
    normalization_switch : false
    BDT                  : false
  v3:
    title_name           : "po_muon_kcut_invis_costheta"
    cut_switch           : true
```

output data


input parameters & running information
cut efficiency for different process, polarization, & signal characters
plot for single process & combined all processes
root file after cuts on different levels (pre cuts, all cuts, w/o MVA vars ...)
log
data
figures
observables

output data structure



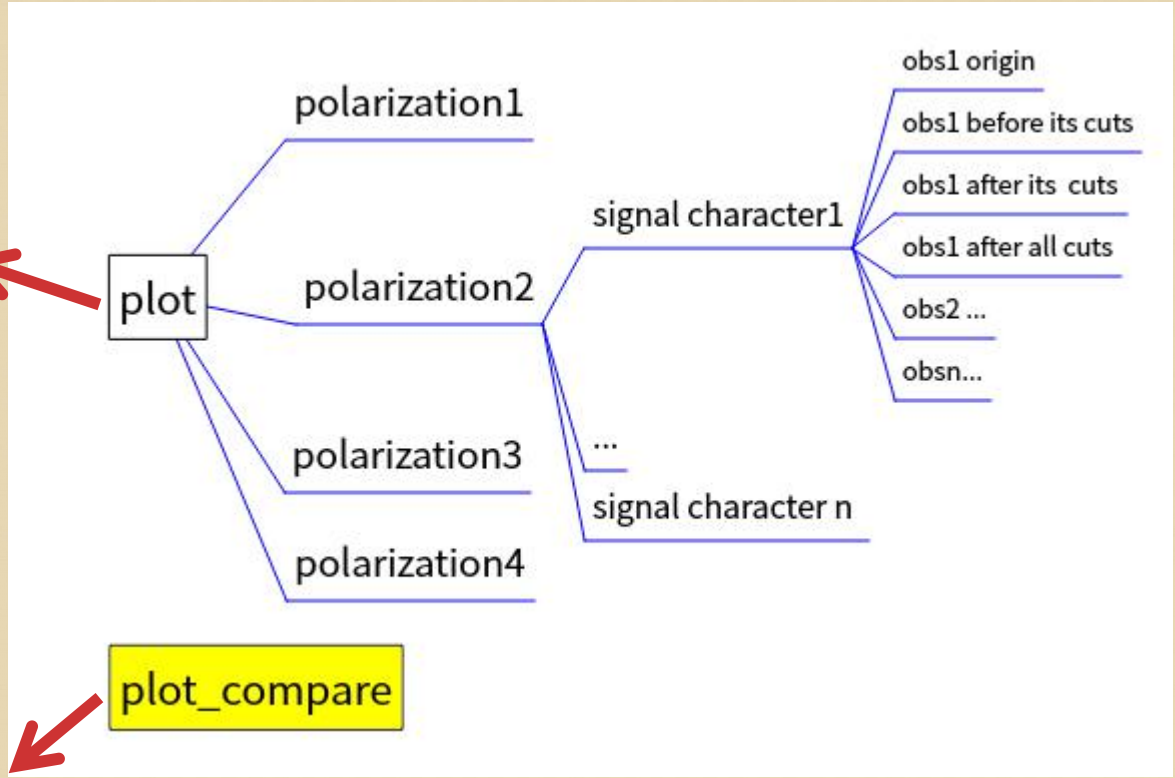


$\int Ldt = 1fb^{-1}$	sig_{ij}^{HA}	$wtata_i^j$	tt_{sl}^l	efficienty	significance
no cut	15.669	409.931	61989.6	1	0.0627186
$N^j \in [2.9, 100]$	12.2375	273.805	52783.2	0.781001	0.0531216
$P_T^{j1} \in [20, 1000]$	12.2375	273.805	52775.7	0.781001	0.0531254
$\eta^{j1} \in [-5, 5]$	12.2375	273.805	52775.7	0.781001	0.0531254
$P_T^{j2} \in [20, 1000]$	12.2375	273.805	52775.7	0.781001	0.0531254
$\eta^{j2} \in [-5, 5]$	12.2375	273.805	52775.7	0.781001	0.0531254
$N^\tau \in [1.9, 100]$	0.91043	17.334	1504.45	0.0581039	0.0233314
$P_T^{\tau 1} \in [20, 1000]$	0.91043	17.334	1504.45	0.0581039	0.0233314
$\eta^{\tau 1} \in [-2.5, 2.5]$	0.91043	17.334	1471.69	0.0581039	0.0235865
$P_T^{\tau 2} \in [20, 1000]$	0.91043	17.334	1471.69	0.0581039	0.0235865
$\eta^{\tau 2} \in [-2.5, 2.5]$	0.91043	17.334	1426.33	0.0581039	0.0239539
$N^{sep} \in [0.9, 100]$	0.239751	4.65097	451.083	0.015301	0.0112277
$P_T^{l1} \in [20, 1000]$	0.222514	4.20652	441.003	0.0142009	0.0105431
$\eta^{l1} \in [-2.5, 2.5]$	0.222514	4.20652	441.003	0.0142009	0.0105431
$M^W \in [70, 110]$	0.20371	2.63503	302.402	0.0130008	0.0116598
$M^A \in [0, 90]$	0.181772	2.0477	105.841	0.0116007	0.0174853
$M^{H^\pm} \in [0, 180]$	0.104989	0.603198	5.04004	0.00670043	0.0437902
$M^t \in [1, 500]$	0.103422	0.603198	5.04004	0.00660042	0.0431425
$M^t \in [0, 4000]$	0.092453	0.444462	2.52002	0.00590038	0.0528784
$\theta_*^l \in [0.8, 1]$	0.034474	0.158736	2.52002	0.00220014	0.020929
$P_T^l \in [0, 30]$	0.029773	0.142863	0	0.00190012	0.0716567
all cut	0.029773	0.142863	0	0.00190012	0.0716567

the output table contains cut efficiency table for all processes with "Tex format", can be directly used in latex file

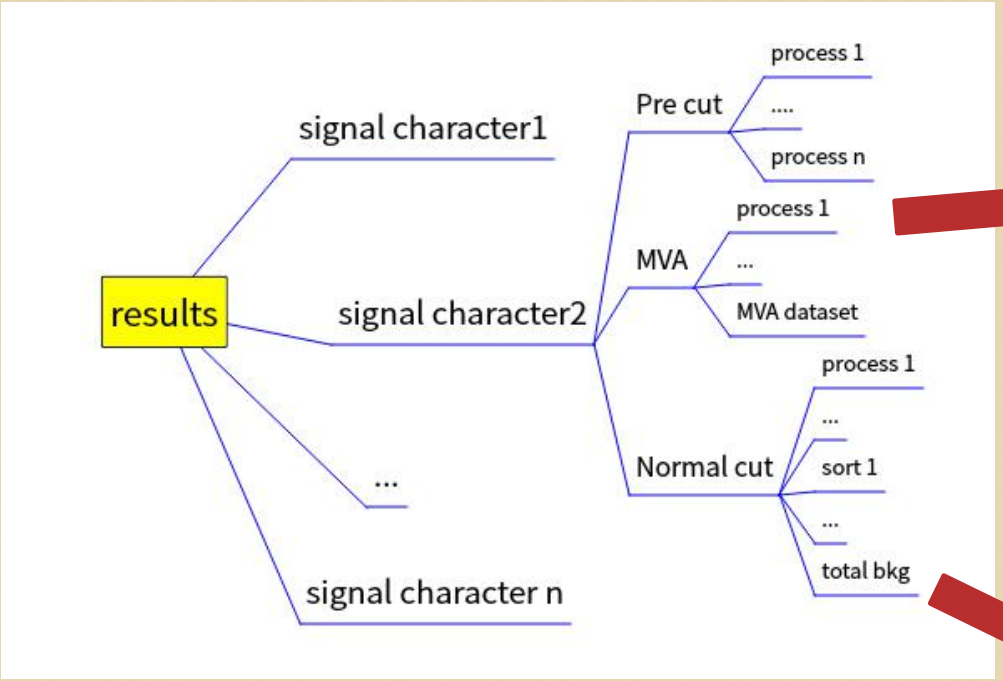
Diagram

for each process



compare all processes in one figure

Root results



1. root format
2. after cuts in different level
3. can be used independently

1. events after all cuts
2. all distributions
3. adjust diagrams directly

Sensitivity

1. control/Sensitivity.dat
2. get data from root results
3. calculate sensitivities according to the control file.

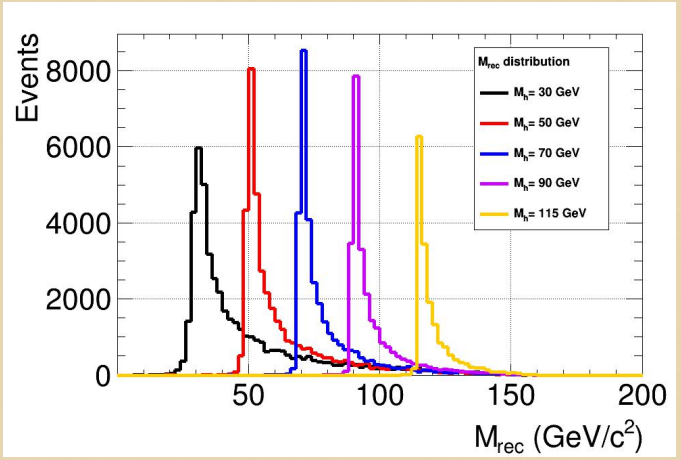
new scenario

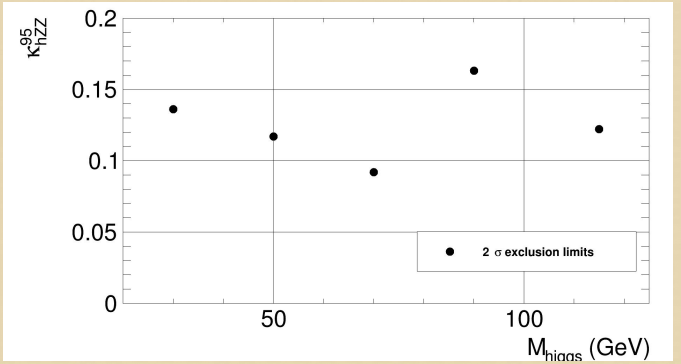
1. control/scenario.dat
2. re-use results in PreCut when changing the polarization. (MVA variable need to be recalculated.)
3. re-use results in NormalCut when changing the luminosity.

compare between different signal characters

after generate results for different signal characters compare them by control/scan.dat

- signal distribution
- cut efficiency table
- sensitivity

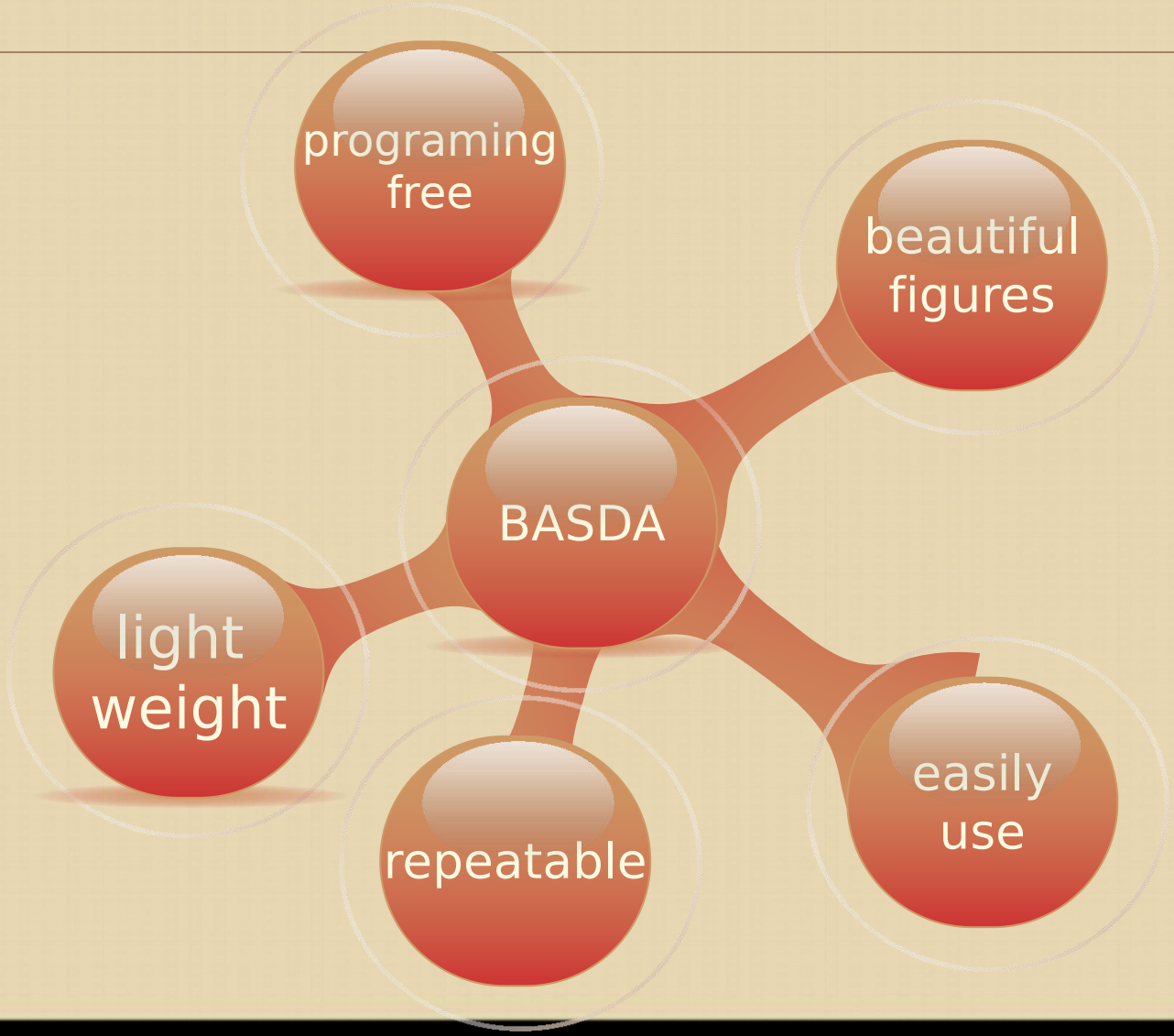




Summary & Future

- https://github.com/YancyW/BASDA
- update with smart pointor to increase steadiblity
- provide more functions, e.g. statistic (wsmaker,nplot...).

Conclusion


programing
free
beautiful
figures
BASDA
light
weight
repeatable
easily
use

Thank you!