Analysis Tools

Beautiful_And_Simple_Drawing_Atificer (BASDA)

Yan Wang

11-04-2018

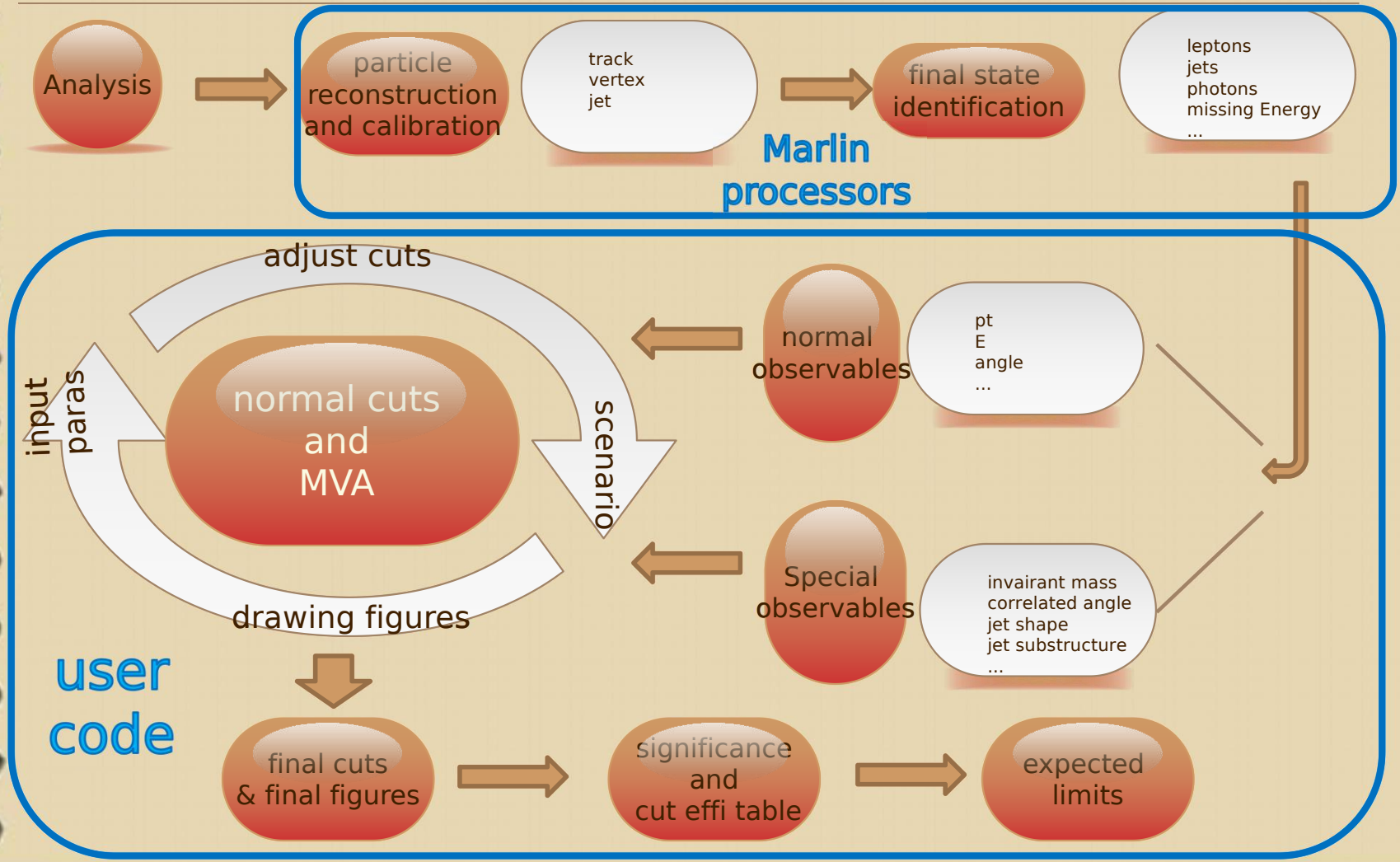
Motivation

Motivation: a standard analysis code will be helpful!

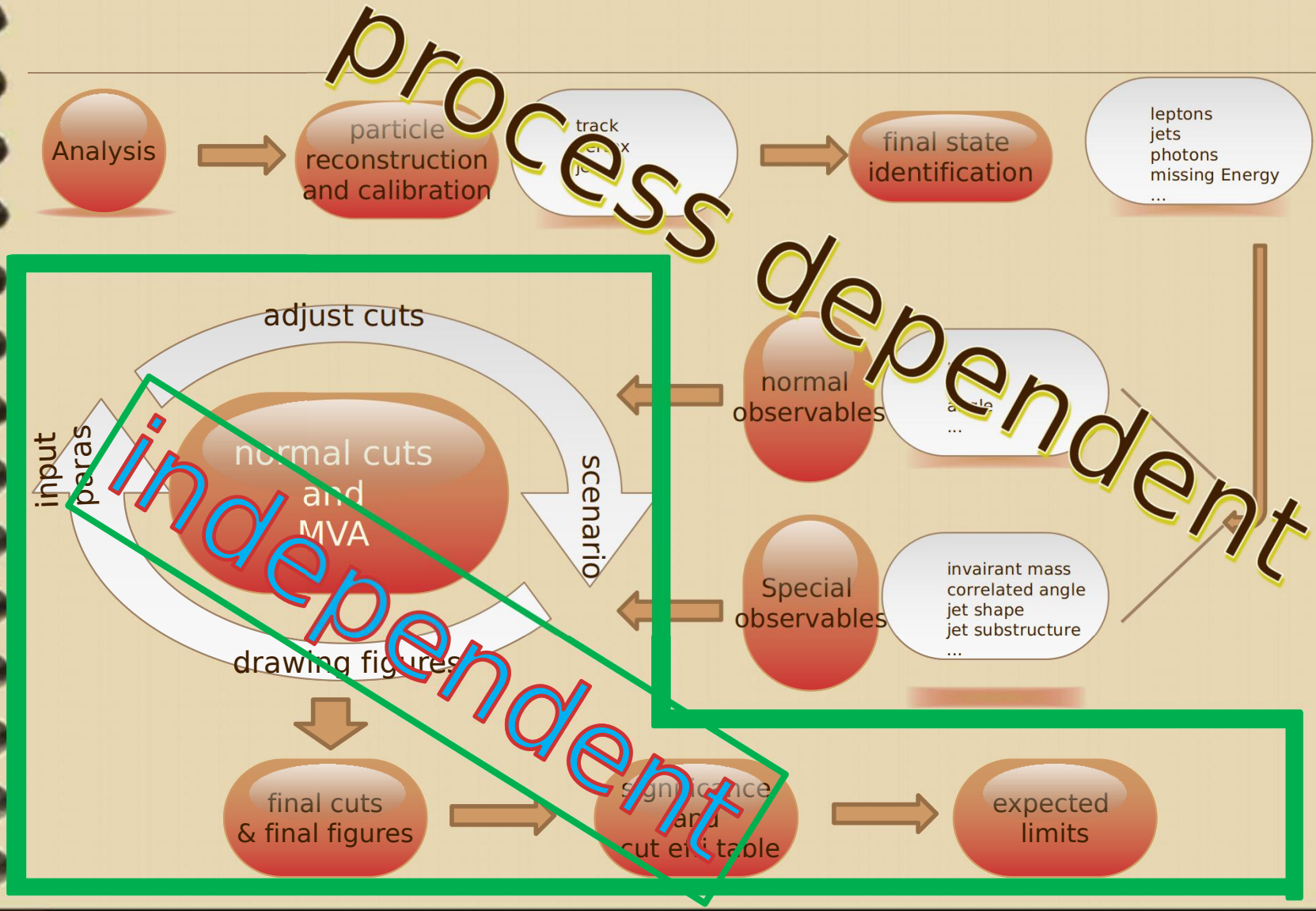
Physics Benchmarks

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**

General Strategy for analysis

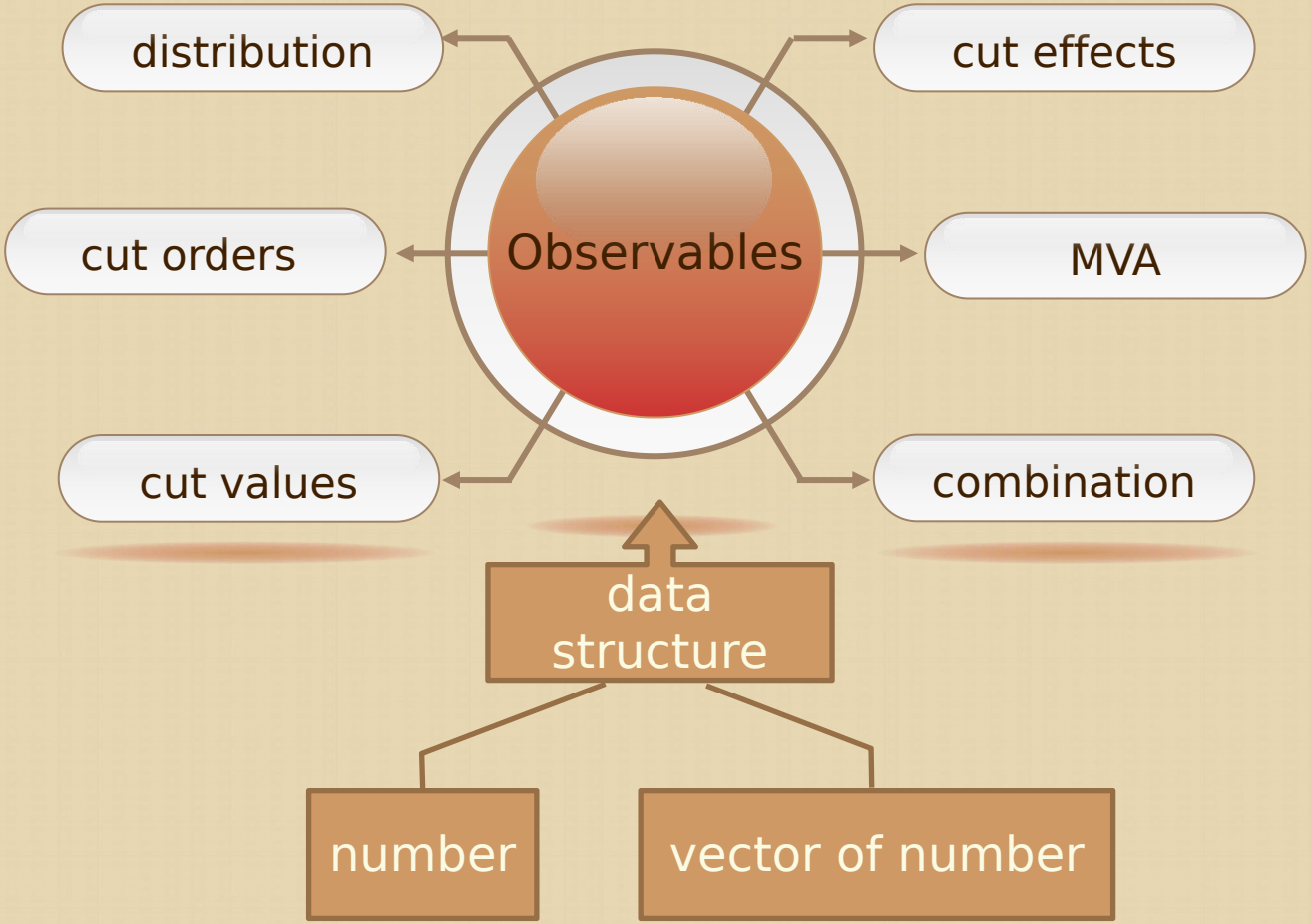


General Strategy for analysis

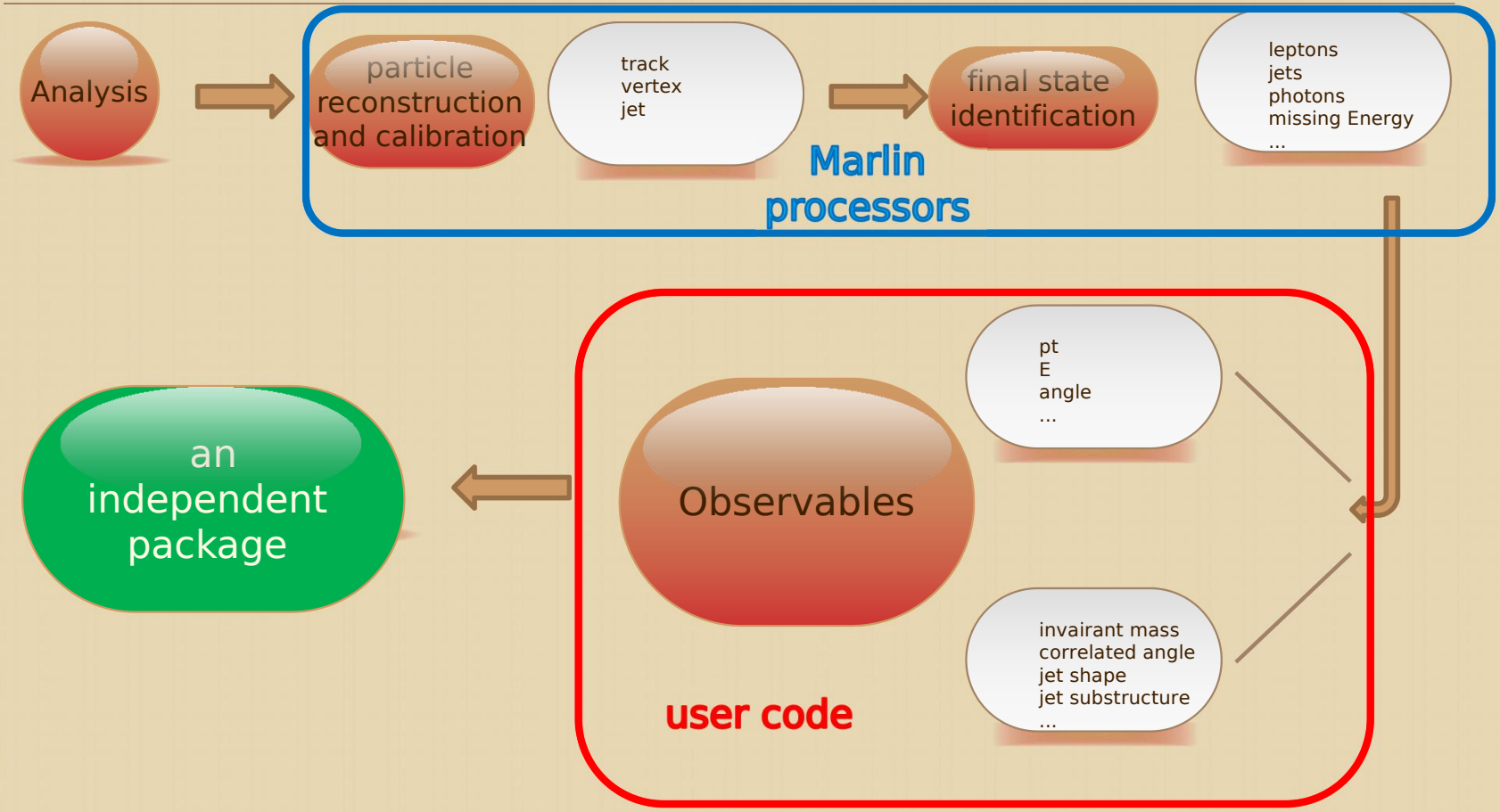


The key role

Bad programmers worry about the code.
Good programmers worry about data structures and their relationships.
--- Linus Torvalds



General Strategy for analysis

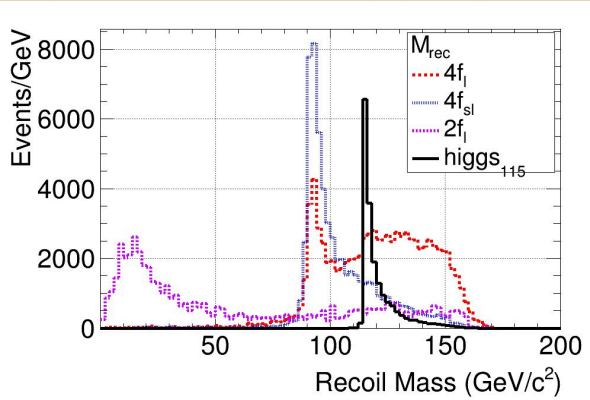


Beautiful_And_Simple_Drawing_Atificer
(BASDA)

BASDA

The beautiful figure

- plot templates --- easy to set
- a cohesive, unified style for the whole group



The simple operation

all with control files

Release time from programing.

Focus on **Physics**!

Easy to repeat

only need to preserve control files.

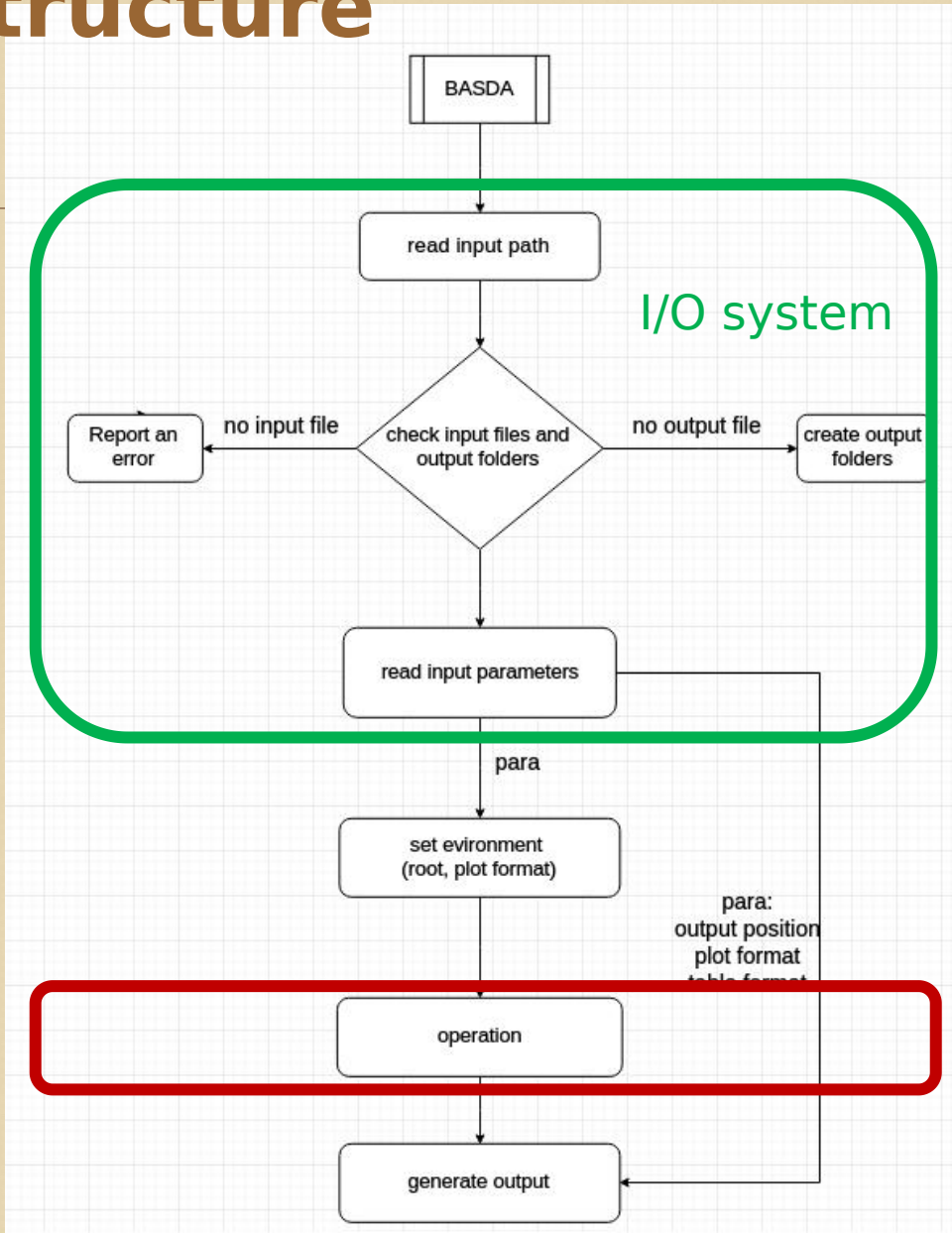
Easy for comparing and finding the differences

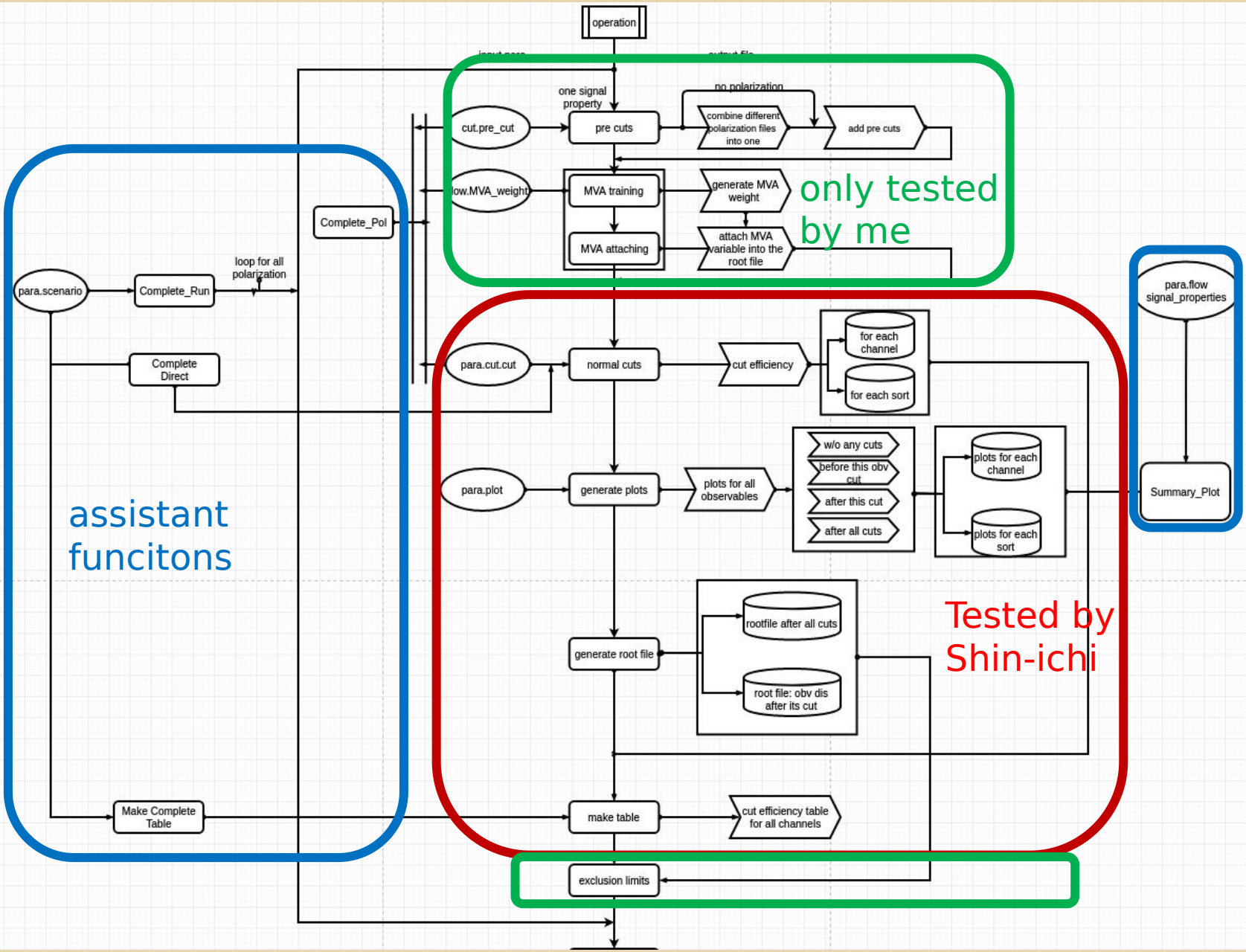
Complete tool sets

provide many analysis tools, fulfill most of requirements

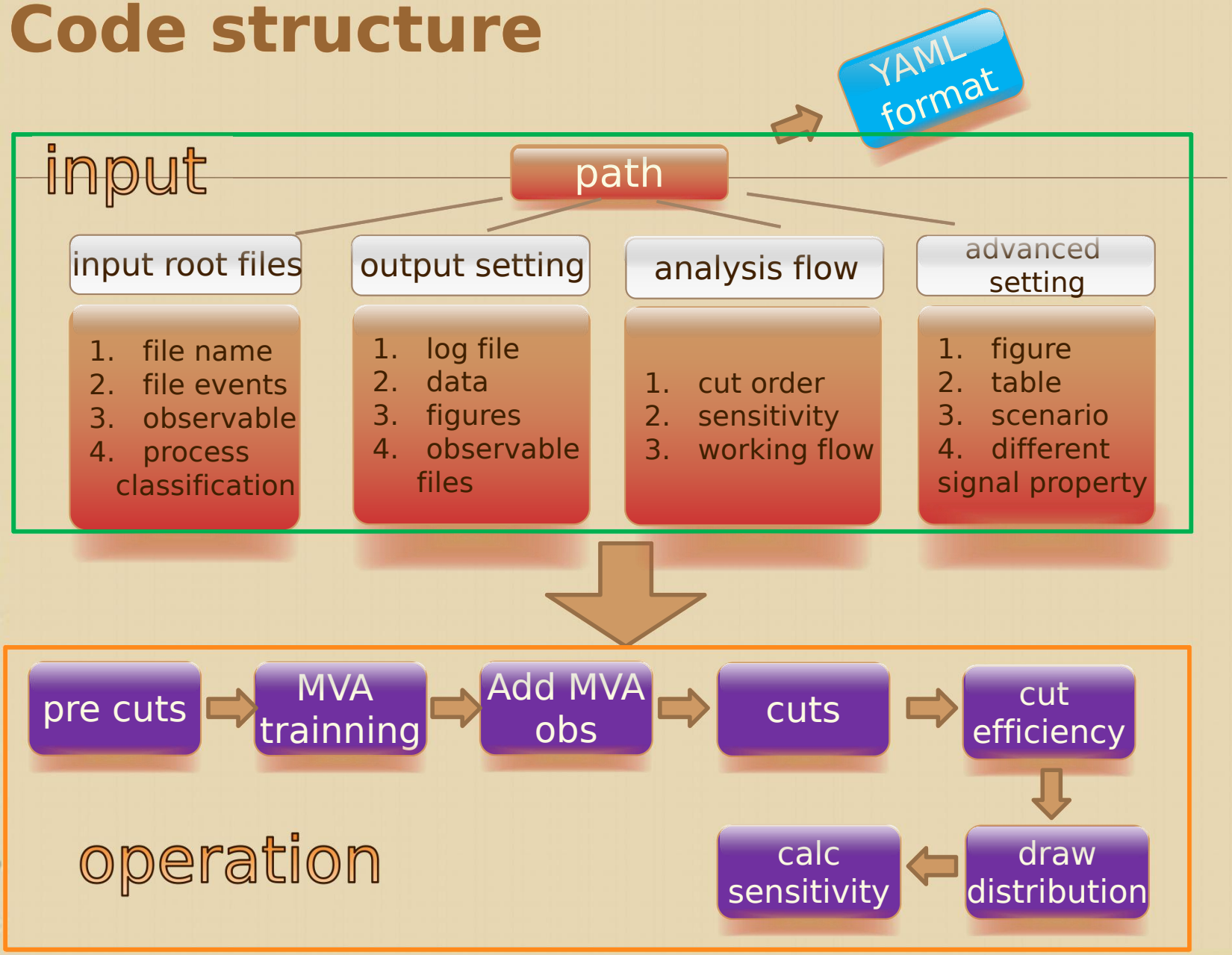
one-click for all analysis and the final results

Code structure

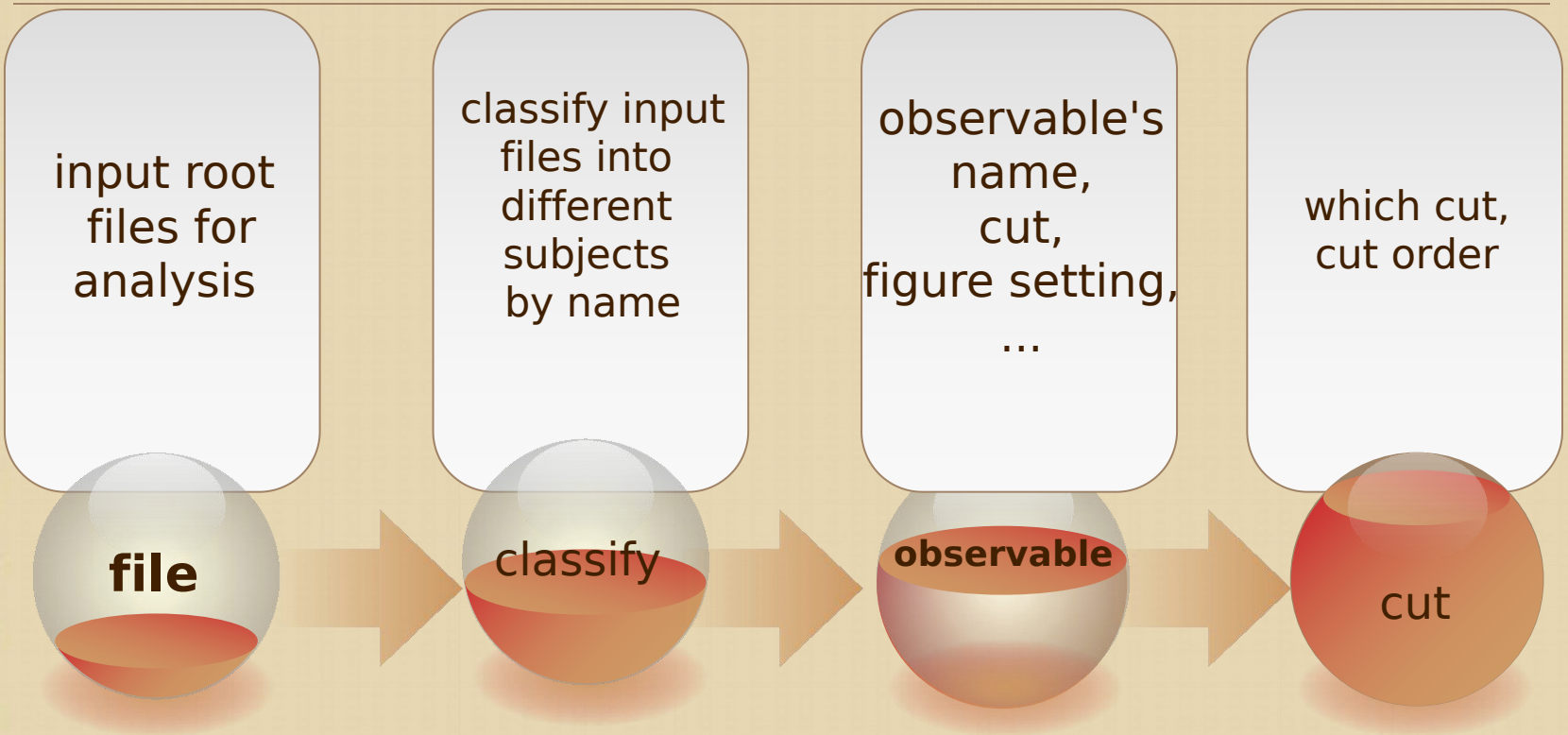



operation
one signal property
no polarization
cut.pre_cut
pre cuts
combine different polarization files into one
add pre cuts
low.MVA_weight
MVA training
generate MVA weight
MVA attaching
attach MVA variable into the root file
only tested by me
Complete_Pol
loop for all polarization
para.scenario
Complete_Run
Complete Direct
para.cut.cut
normal cuts
cut efficiency
for each channel
for each sort
para.plot
generate plots
plots for all observables
w/o any cuts
before this obv cut
after this cut
after all cuts
plots for each channel
plots for each sort
para.flow signal_properties
Summary_Plot
assistant funcitons
generate root file
rootfile after all cuts
root file: obv dis after its cut
Tested by Shin-ichi
Make Complete Table
make table
cut efficiency table for all channels
exclusion limits

Code structure

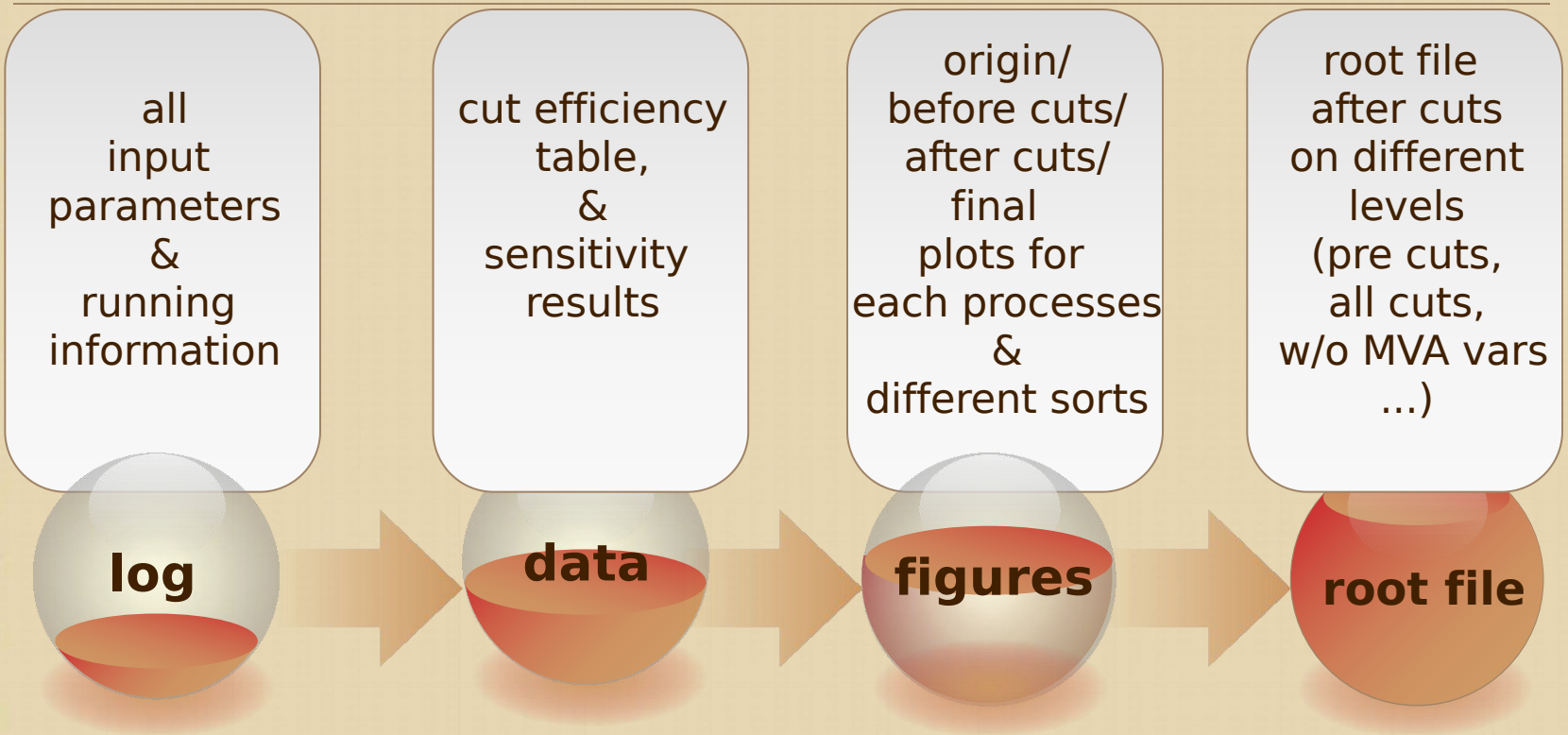


input data


input root files for analysis
classify input files into different subjects by name
observable's name, cut, figure setting, ...
which cut, cut order
file
classify
observable
cut

output

output data

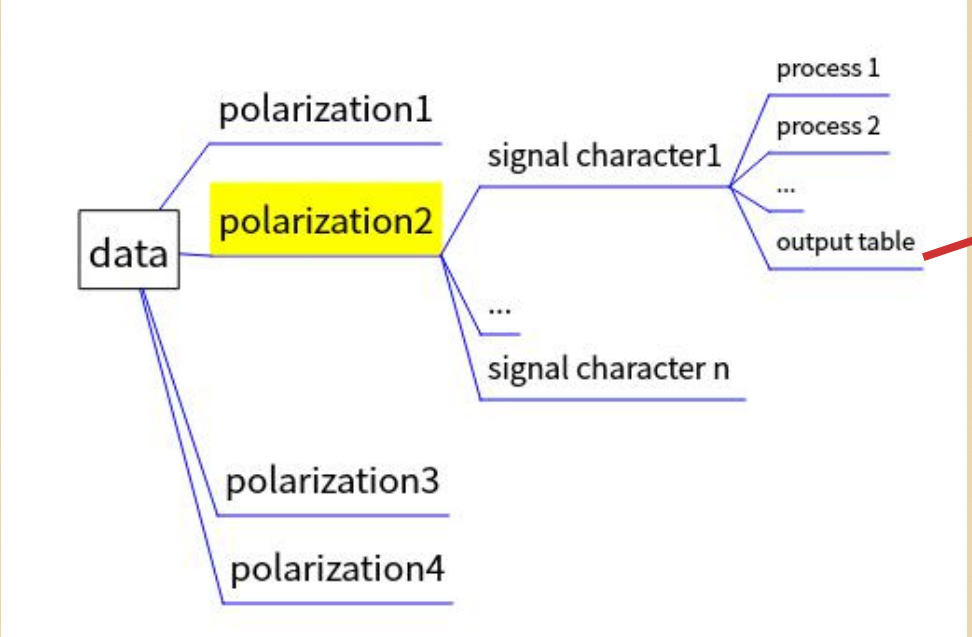


all temporary data are preserved for reuse.

for my light-higgs analysis:

- 23 higgs mass banchmark points, with four polarized results (--,-+,+-,++),
- total output takes 8.3G hard-disk space

data

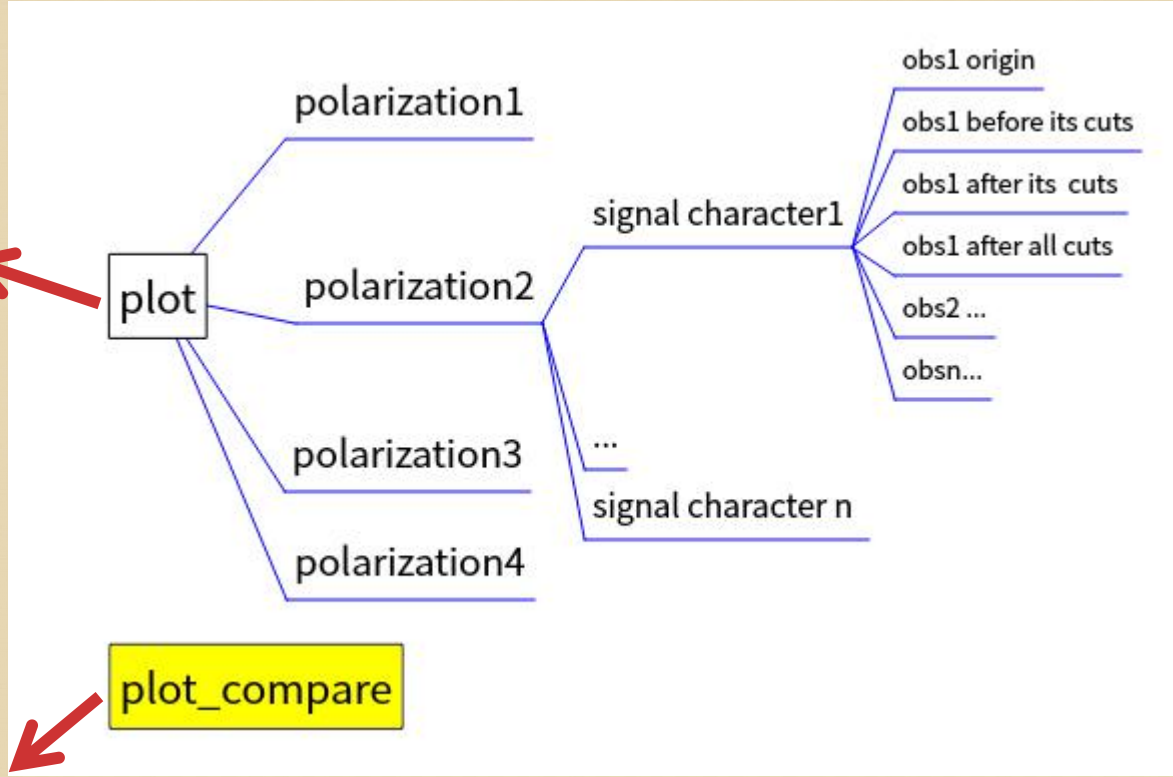


cut efficiency table with "Tex format",
directly used in latex file

$\int Ldt = 1fb^{-1}$	sig_{G}^{HA}	$wtata_{l}^{t}$	tt_{sl}^{t}	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
$N^{lep} \in [0.9, 100]$	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
$\theta_{*}^{t} \in [0.8, 1]$	0.034474	0.158736	2.52002	0.00220014	0.020929
$P_{T}^{t} \in [0, 30]$	0.029773	0.142863	0	0.00190012	0.0716567
all cut	0.029773	0.142863	0	0.00190012	0.0716567

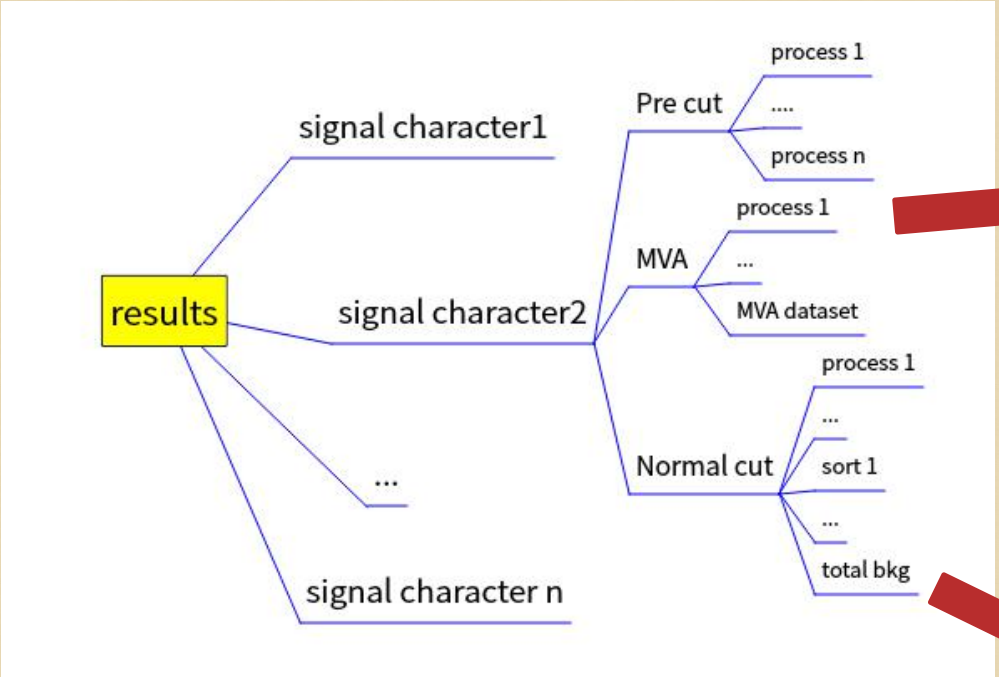
figures

all channel



all sorts in one figure (png format)

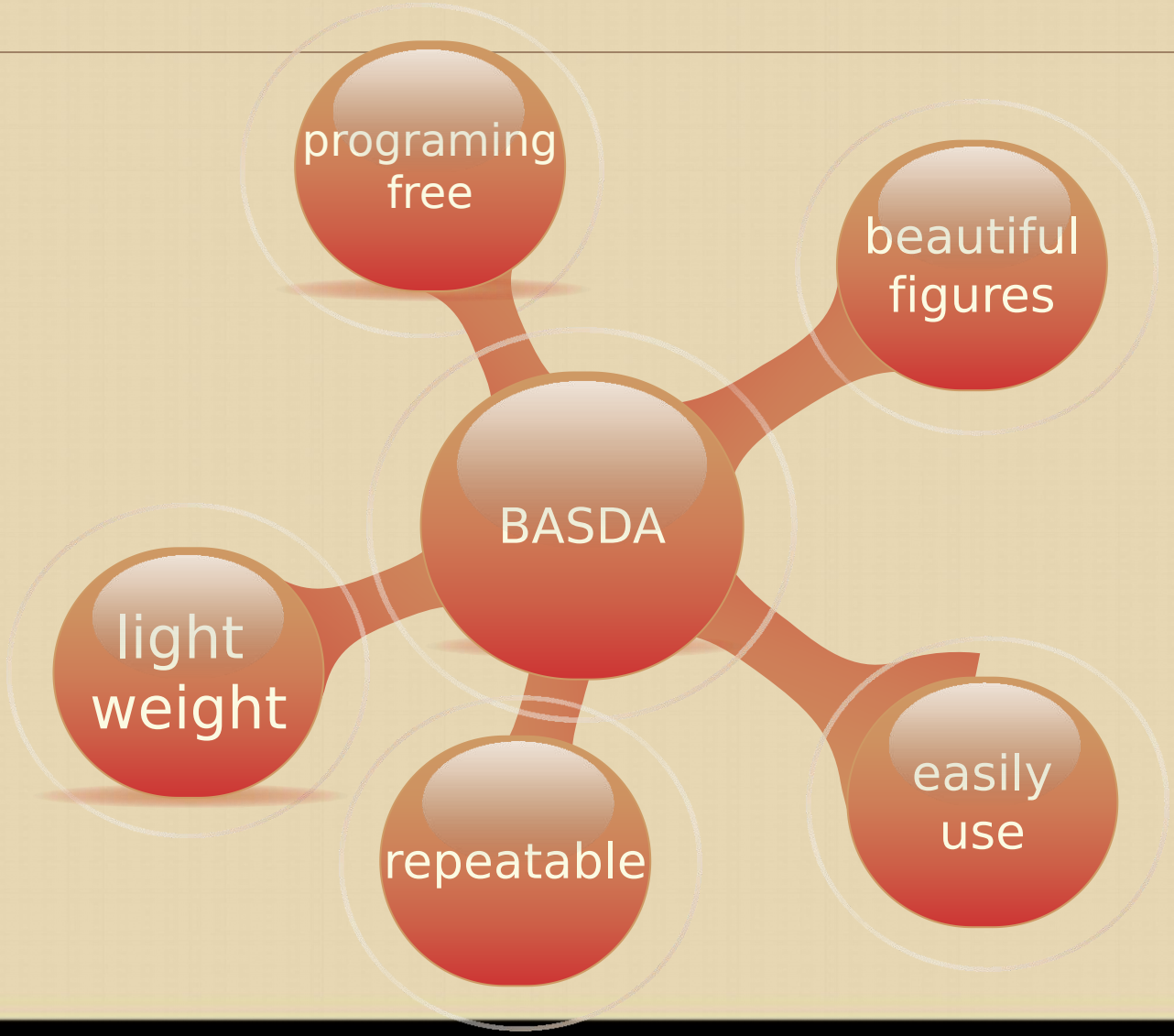
root file



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

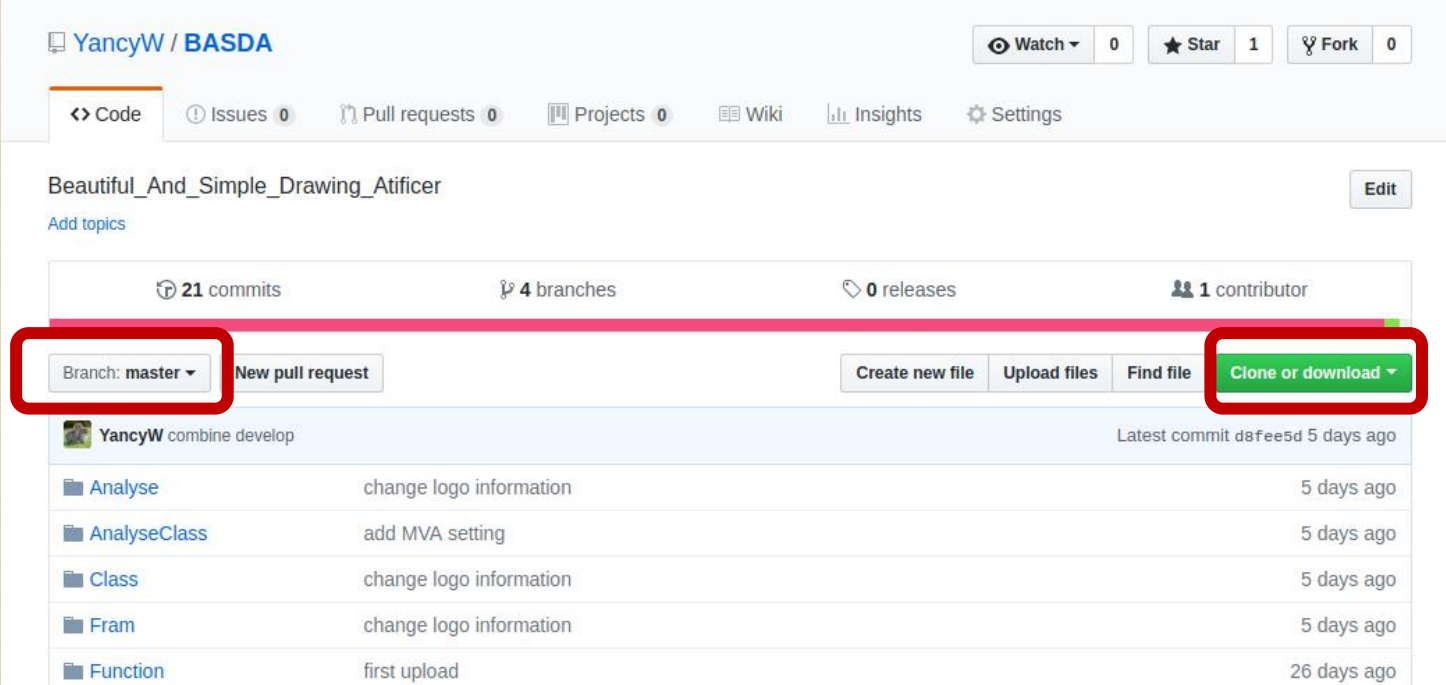
1. events after all cuts
2. all distributions
3. adjust diagrams directly

Conclusion


programing free
beautiful figures
BASDA
light weight
repeatable
easily use

Installation

- https://github.com/YancyW/BASDA
  - four branches: master, release, develop, feature.



Summary & Future

- Manuals/Examples:
  - mygithub/doc/BASDA_Usage_Example.pdf
  - mygithub/doc/Usage.md
- More tests:
  - MVA
  - Code consistency
- More functions:
  - 2-D plots
  - statistic (like wsmaker,nplot ... in ATLAS group).
- GUI for basic users

Thank you!