high tea

for your freshly brewed analysis

The HighTEA collaboration

Michal Czakon, Zahari Kassabov, Alexander Mitov, **Rene Poncelet** and Andrei Popescu

What is HighTEA in a nutshell?

A tool to make state-of-the-art collider phenomenology ...

- … available to everyone
- … accessible to everyone
- … sustainable

A tool to make state-of-the-art collider phenomenology ...

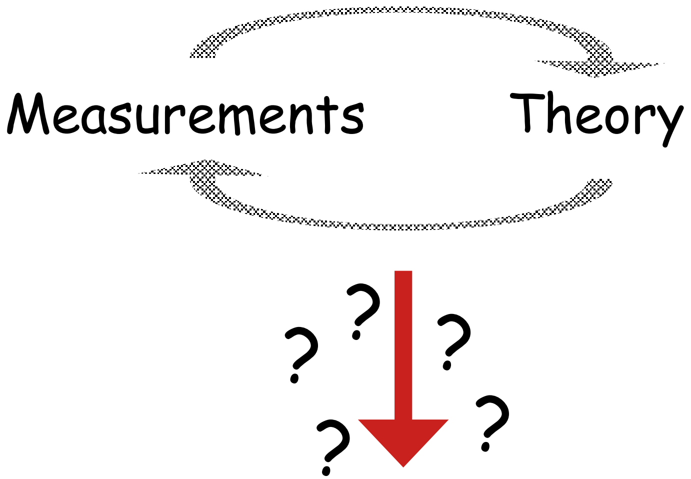
- ... available to everyone*
  - No computing resources needed
  - No access to complicated codes required
- ... accessible to everyone
  - No specific programming skills required
  - No expertise in theory or HEP tools needed
- ... sustainable
  - Only a fraction of computing cost to conventional computations

*who has at least a smartphone

Why is this needed?

Fact of Life:

"We are getting most out of collider experiments by comparing measurements to the 'best' available predictions"



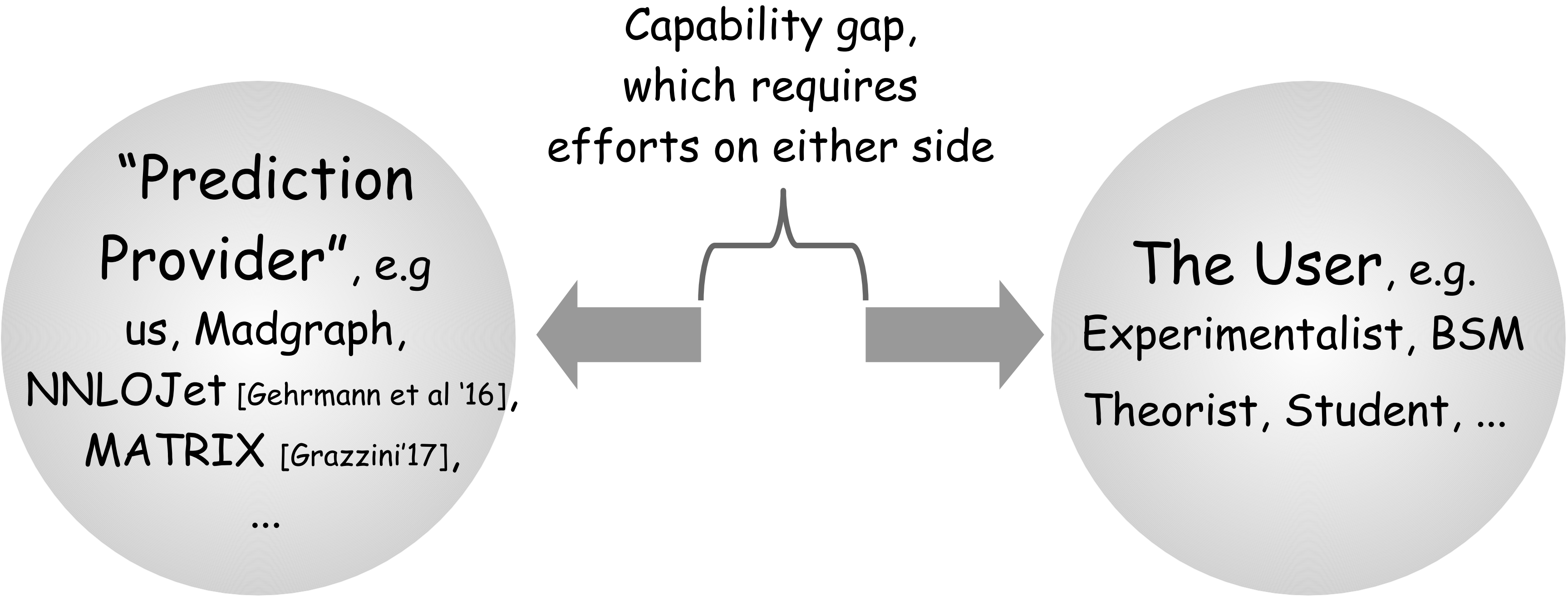
Where do those who do the comparisons get hold of the "best" predictions?

The present situation

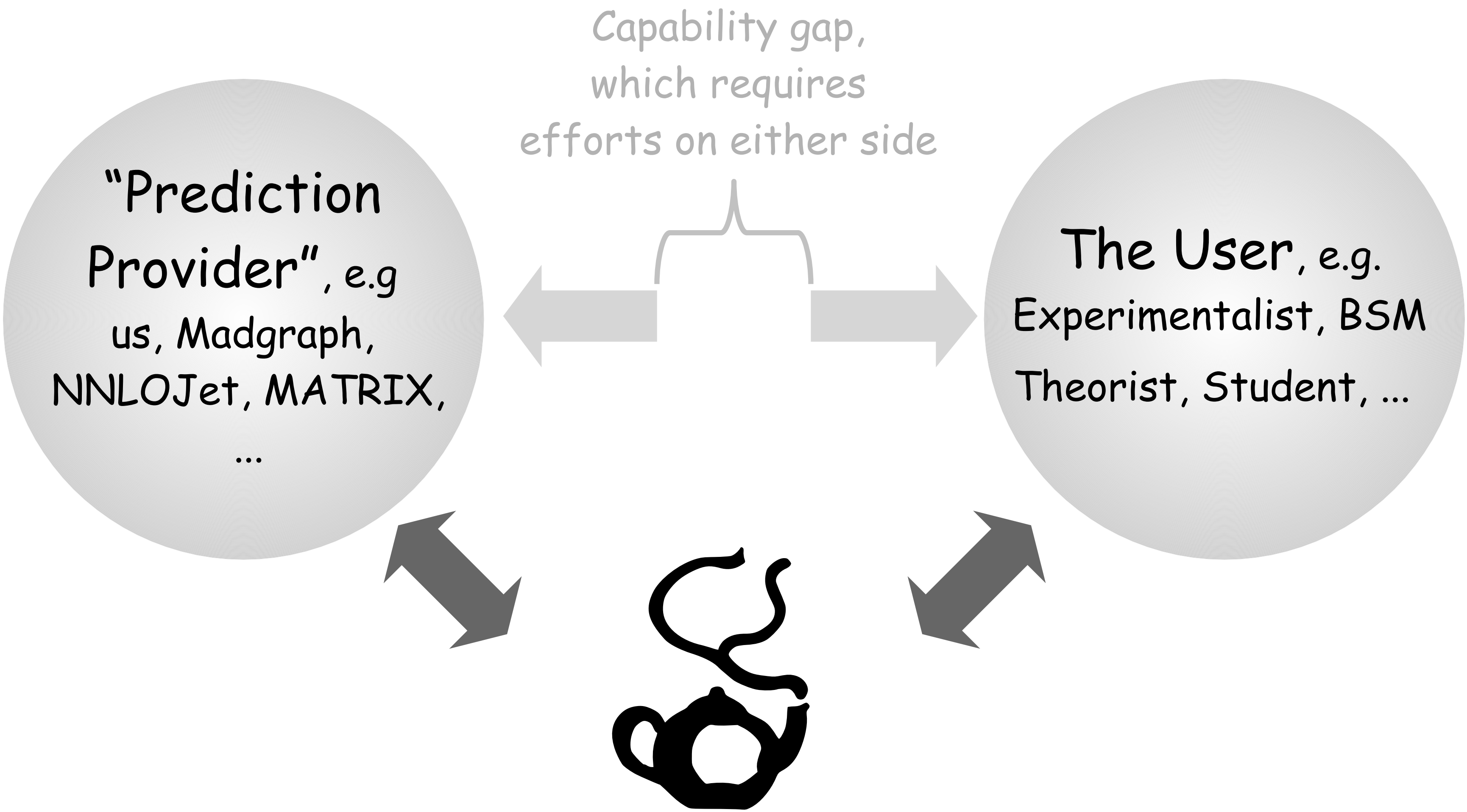
"I'm interested in observable X for process Y"

- Run a public general purpose code (like MadGraph [Alwall'14], Sherpa [Gleisberg'08],...)
  - Implementation of X
  - Expertise to get 'sensible' results (many technical parameters)
  - Computational resources
  - Basically restricted to NLO QCD (+EW)
- Ask the authors of paper "(N)NLO QCD corrections for Y" for a prediction of your particular observable.
  - Inflexible ("But what is about binning Z?")
  - Time consuming (human and computing time)

The role of HighTEA



The role of HighTEA



How does HighTEA work?

- A database of pre-computed “Theory Events”
  (Parton level / Particle level / MC Truth)
  - Currently: partonic fixed order events
  - Extensions to included showered/hadronised events is feasible
  - Equivalent to a full fledged computation while factorising scale and PDF dependence

There are attempts in the literature:
LHE [Alwall et al ‘06],
Ntuple [BlackHat ‘08’13],

- Analysis of the data through an user interface
  - Easy-to-use
  - Flexible
  - Fast

Public HighTEA resources

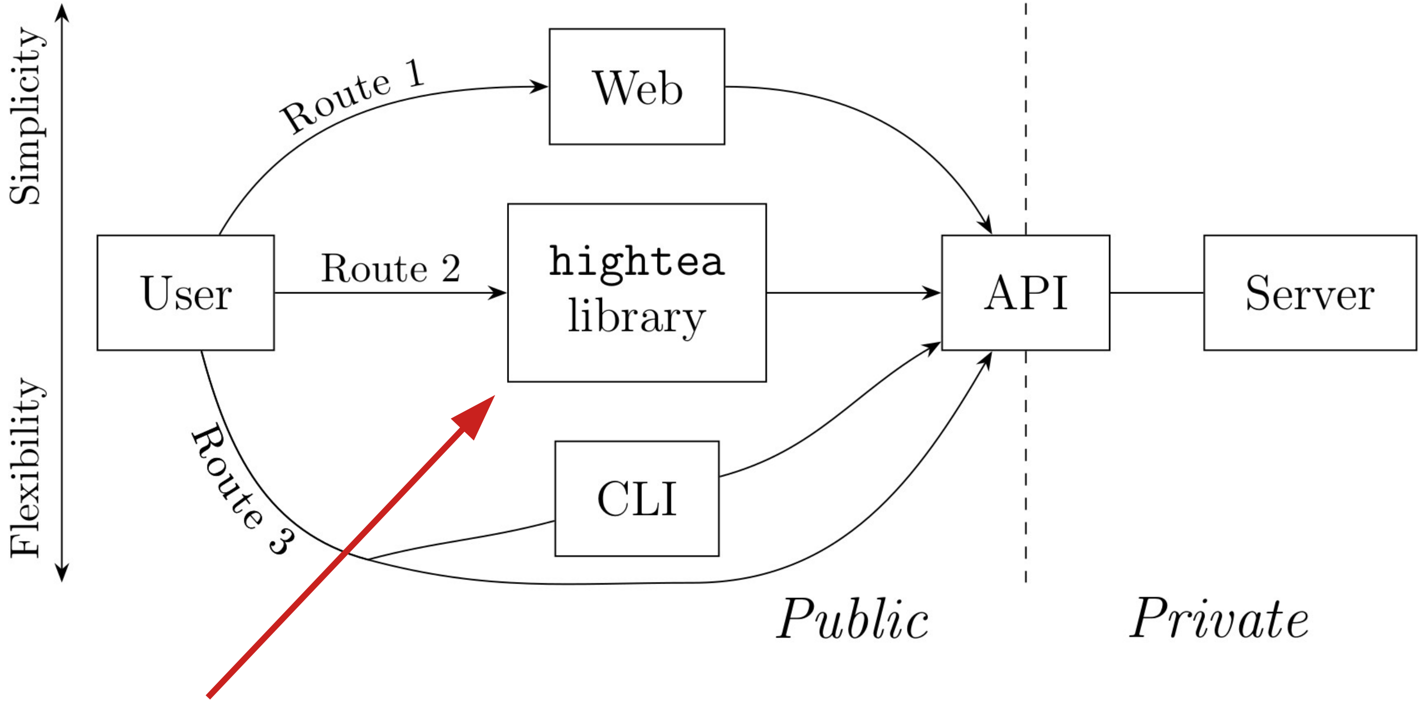
Central Homepage:

https://www.precision.hep.phy.cam.ac.uk/hightea/

- Overview
- General information
- Access point:
  - Running HighTEA via Google Colab
  - The Github page: https://github.com/HighteaCollaboration

Disclaimer: We are in the polishing phase, so small changes possible

The setup from the user perspective



Demonstration Google Colab

Current capabilities of the API

- Selecting a process: currently WW (+decays), ttbar, diphoton
- Asking for histograms of observables:
  - Some observables are pre-implemented
  - Own observables from some basic 4-momenta
  - Free specification of bins
- Renormalization/Factorization Scale variation:
  - Change of pre-factors and functional form
- PDF variation (and alphaS)
- Specify phase space cuts
- If jets are present -> jet radius can be changed too

HighTEA-client example: top-quark pair production

1) Give it a name

2) Select a process

3) Customise your request:

- Observables
- Contributions
- Scale choice
- PDF choice
- Binning
- Variations

```
job = hightea('Example-ttbar-simple-job',directory=USERDIR)      # define new job
job.process('pp_tt_13000_172.5',verbose=False)                   # specify process for job
job.define_new_variable('mtt',                                   # specify a new variable
                        'sqrt((p_t_0+p_tbar_0)**2-(p_t_1+p_tbar_1)**2-(p_t_2+p_tbar_2)**2-(p_t_3+p_tbar_3)**2)')
```

```
job.contribution('NNLO')                                           # specify contribution
job.scales('mtt','mtt*2')                                          # choose renormalization and factorization scale
job.pdf('CT14nnlo')                                                # choose pdf
job.observable('mtt',[345, 400, 450, 500, 550, 600, 650, 700])    # specify binning: variable and bin edges
job.scale_variation('3-point')                                     # add scale variation
```

4) Submit and have a tea:

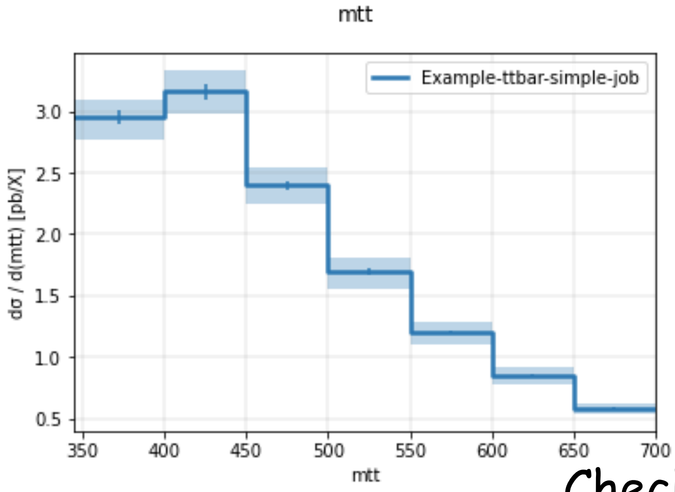
```
job.request()
```

```
request submitted :  2022-12-12 20:10:06.097307
request finished  :  2022-12-12 20:31:30.017514
```

HighTEA example: top-quark pair production

5) Look at the results
- Plots
- Numbers

```
1  plot(job.result());
```




```
1  job.show_result()
```

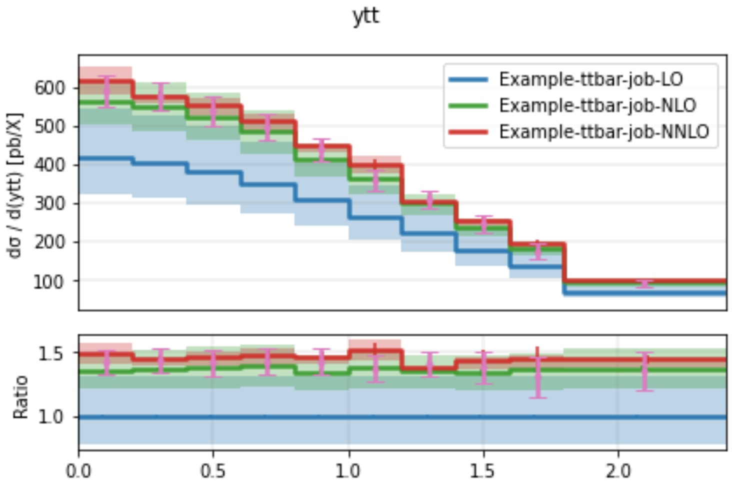
```
Name                        :  Example-ttbar-simple-job
Contributions               :  ['NNLO']
muR                         :  mtt
muF                         :  mtt*2
pdf                         :  CT14nnlo , 0
fiducial xsection           :       741.88
fiducial xsection error     :       5.4971
systematic unc. [%]         : scale (3)
                            :  +      6.4/ -      7.2
Histogram     : mtt
 bin1 low   | bin1 high | sigma [pb]  | mc-err [pb] |        scale (3) [%] |
       345  |       400 |      162.43 |      3.4369 | +      4.6/ -      6.1|
       400  |       450 |      158.09 |      3.0948 | +      5.3/ -      5.9|
       450  |       500 |      119.76 |      1.8652 | +      6.1/ -      6.5|
       500  |       550 |      84.537 |      1.3853 | +      7.3/ -        8|
       550  |       600 |      60.065 |     0.96371 | +      6.9/ -      8.3|
       600  |       650 |      42.524 |     0.67758 | +      7.5/ -      8.1|
       650  |       700 |      29.135 |     0.43426 | +      6.9/ -      7.7|
```

Check out the examples:
https://github.com/HighteaCollaboration/hightea-examples

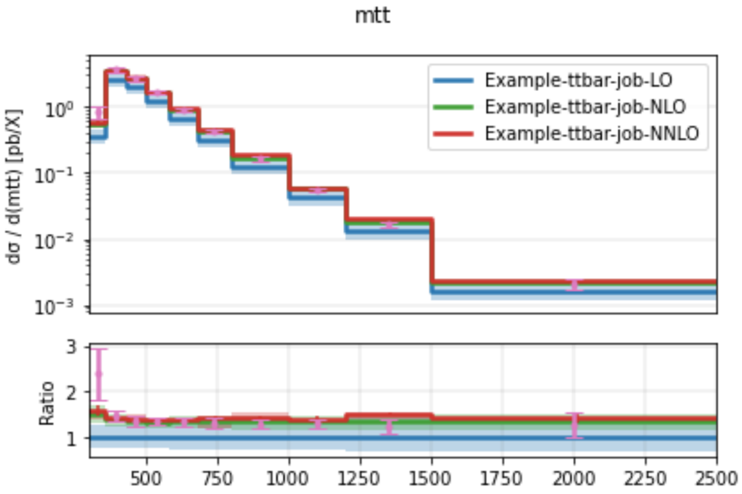
HighTEA example: CMS ttbar

CMS 1803.08856

Rapidity of top-quark pair:



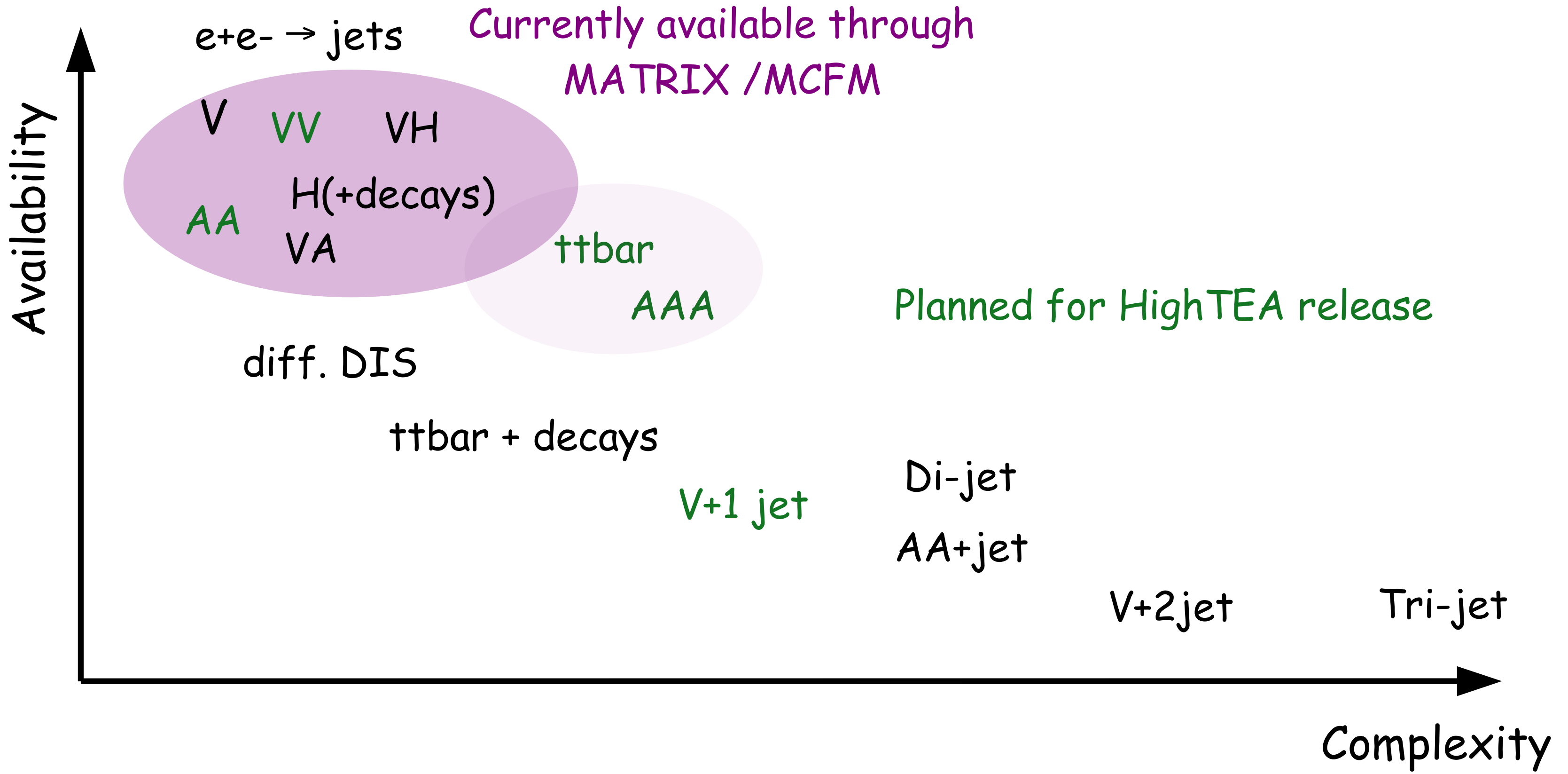
Invariant mass of top-quark pair:



Outlook & Summary

Outlook – We will add more content!

Processes currently implemented in our STRIPPER framework through NNLO QCD



* V processes include leptonic decay mode(s)

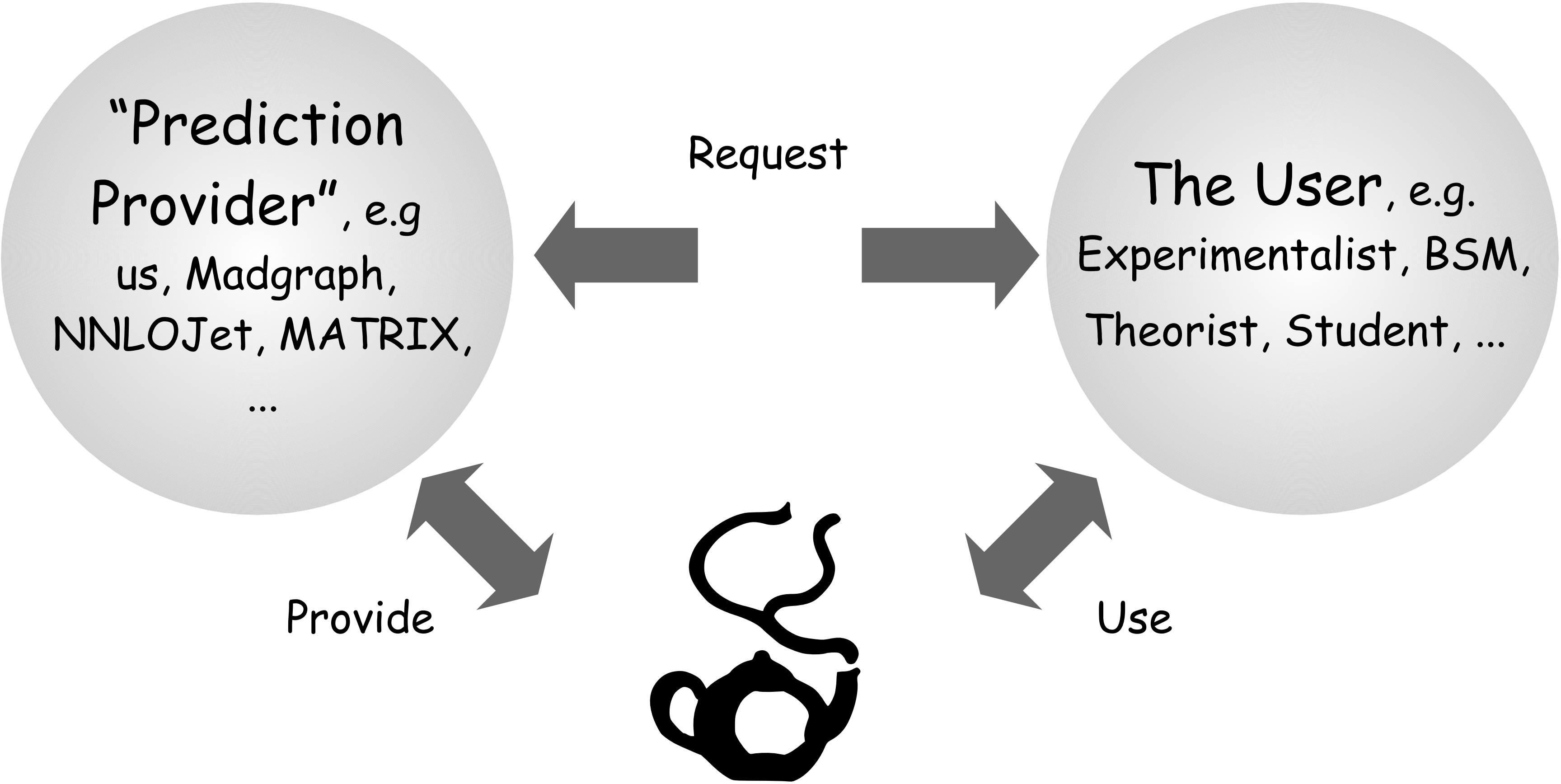
Outlook - More functionality/applications

- Functionality:
  - SMPDFs → More efficient PDF error estimations
  - More control over initial state, selection of specific parton fluxes
  - Incorporation of HEPMC/LHE files. (Basically putting Madgraph into HighTEA)
- Applications:
  - FastNLO/PineAPPL grids could be generated from our database
  - PDF fits using directly NNLO QCD predictions (instead of K-factors)
    - Might be interesting in case of new channels at NNLO, for example: pp → AA (gg → AA)
  - EFT operators:

$$\mathrm{d}\sigma = \mathrm{d}\sigma^{SM} + \sum \frac{c_i}{\Lambda} \underbrace{\mathrm{d}\sigma^{\mathcal{O}_i}} + \sum \frac{c_i c_j}{\Lambda^2} \underbrace{\mathrm{d}\sigma^{\mathcal{O}_i \mathcal{O}_j}} + \ldots$$

Individual "datasets"

The Vision


"Prediction Provider", e.g
us, Madgraph, NNLOJet, MATRIX, ...
Request
The User, e.g.
Experimentalist, BSM, Theorist, Student, ...
Provide
Use

Summary

- HighTEA: High-energy Theory Event Analysis
  - A tool to make state-of-the-art phenomenology: available, accessible and sustainable!
- Main functionalities to get NNLO QCD implemented, including:
  - "Arbitrary" observables and binnings
  - PDF/scale variations
  - Phase space restrictions
- Plan for release:
  Examples and Tutorials,
  Providing datasets for NNLO QCD in AA, VV, ttbar, V+jet(not yet publicly available)

high tea

for your freshly brewed analysis

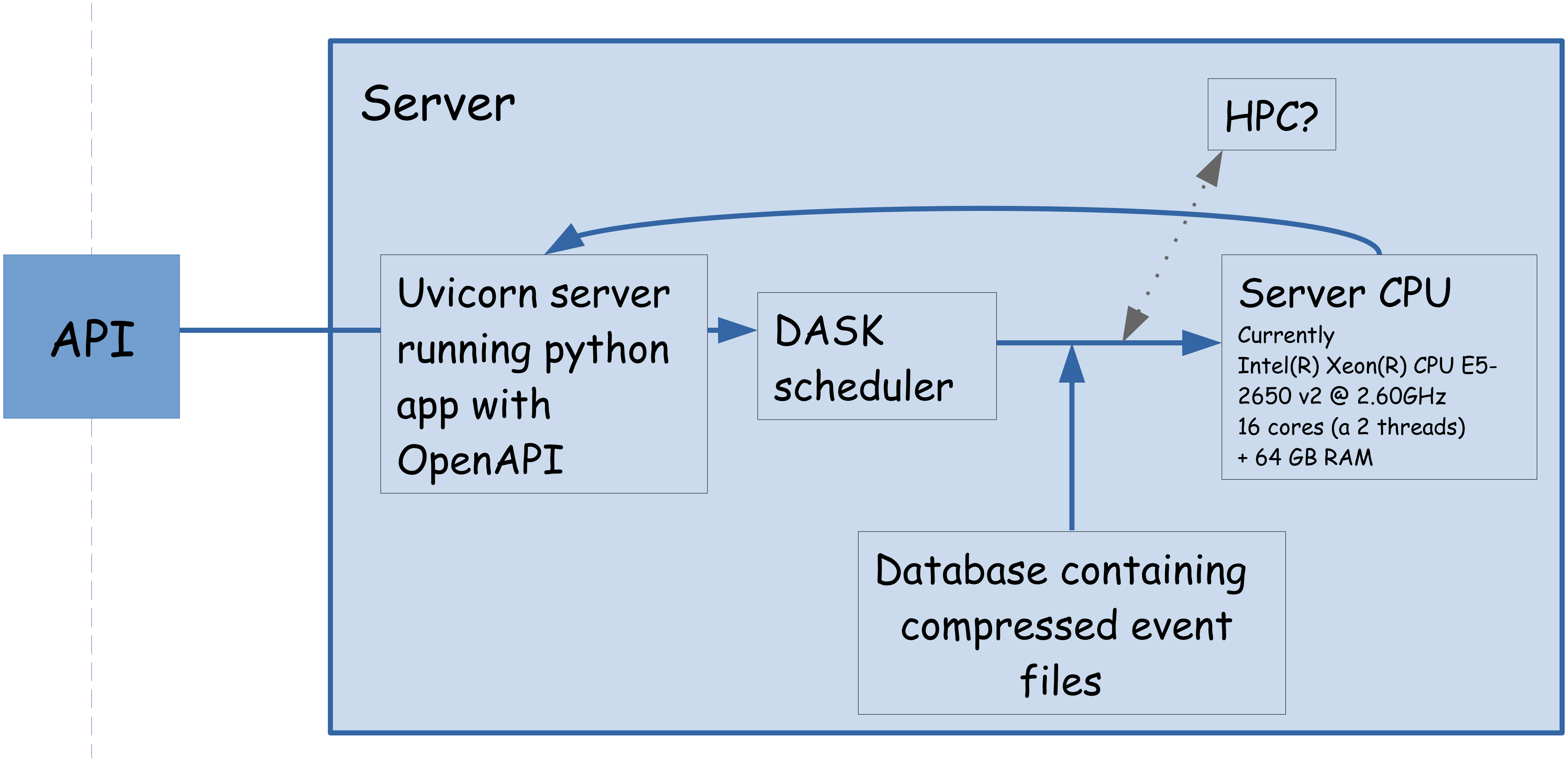
Thank you and stay tuned!

The HighTEA collaboration

Michal Czakon, Zahari Kassabov, Alexander Mitov, **Rene Poncelet** and Andrei Popescu

Backup

The server


API
Server
HPC?
Uvicorn server running python app with OpenAPI
DASK scheduler
Server CPU
Currently Intel(R) Xeon(R) CPU E5-2650 v2 @ 2.60GHz
16 cores (a 2 threads)
+ 64 GB RAM
Database containing compressed event files

Partially unweighted events

The hadronic cross section in collinear factorization:

$$d\sigma(P_1, P_2) = \sum_{ab} \iint_0^1 dx_1 dx_2 f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) d\hat{\sigma}_{ab}(x_1 P_1, x_2 P_2)$$

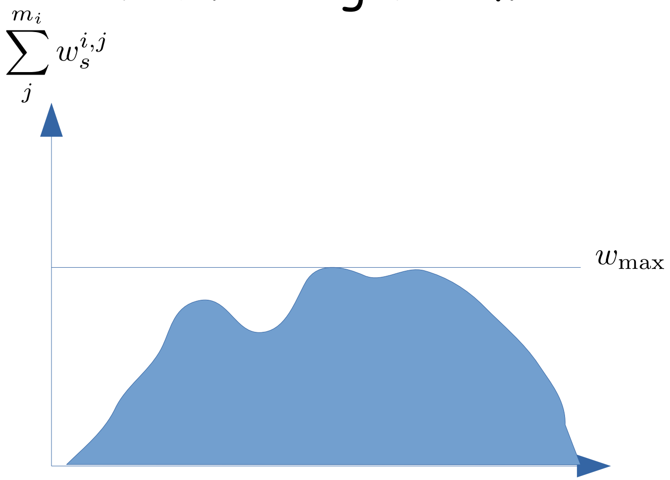
The partonic cross section can be expanded in aS: $\hat{\sigma}_{ab\to X} = \hat{\sigma}^{(0)}_{ab\to X} + \hat{\sigma}^{(1)}_{ab\to X} + \hat{\sigma}^{(2)}_{ab\to X} + \mathcal{O}(\alpha_s^3)$

Using MC method for integration:

$$\sigma_{\text{tot}} = \frac{1}{n}\sum_i^n \left(\sum_j^{m_i} w_s^{i,j} \right)$$

Beyond LO events might correspond to more than one kinematic: Subtraction events!

Hit-And-Miss Algorithm:



Store each sub-event with weight: $w_s^{i,j}/w_{\max}$

Reweighting

Factorizing renormalization and factorization scale dependence:

$$w_s^{i,j} = w_{\rm PDF}(\mu_F, x_1, x_2) w_{\alpha_s}(\mu_R) \left(\sum_{i,j} c_{i,j} \ln\big(\mu_R^2\big)^i \ln\big(\mu_F^2\big)^j \right)$$

PDF dependence:

$$w_{\rm PDF}(\mu, x_1, x_2) = \sum_{ab \in \text{channel}} f_a(x_1, \mu) f_b(x_2, \mu)$$

AlphaS dependence:

$$w_{\alpha_s}(\mu) = (\alpha_s(\mu))^m$$

Allows full control over scales and PDF