Cambridge 4 Feb 2022

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

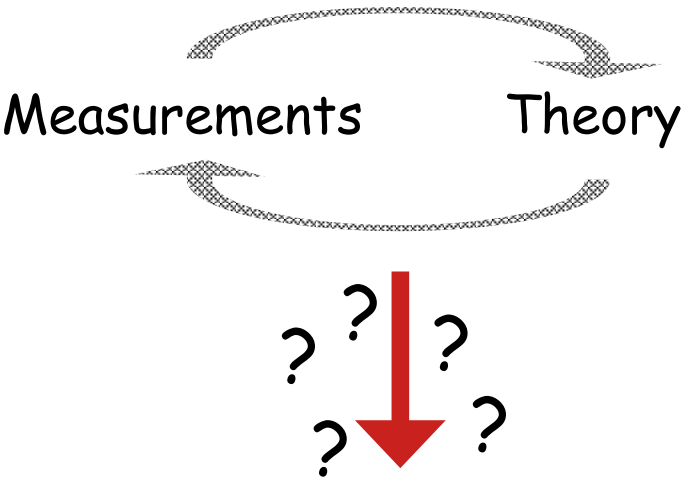
*who has at least a smartphone

Why is this needed?

Fact of Life:

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

Precise & accurate:



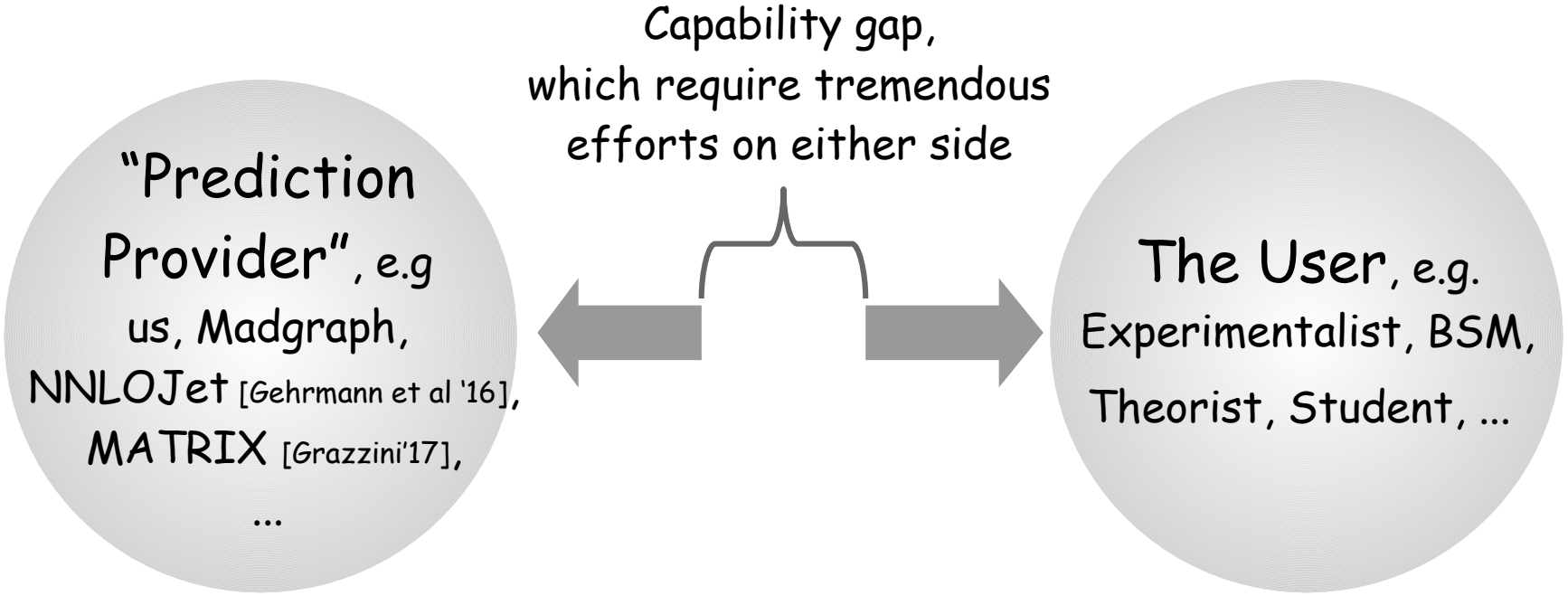
Where do those who do the comparisons get hold of the "best" predictions?

The present situation

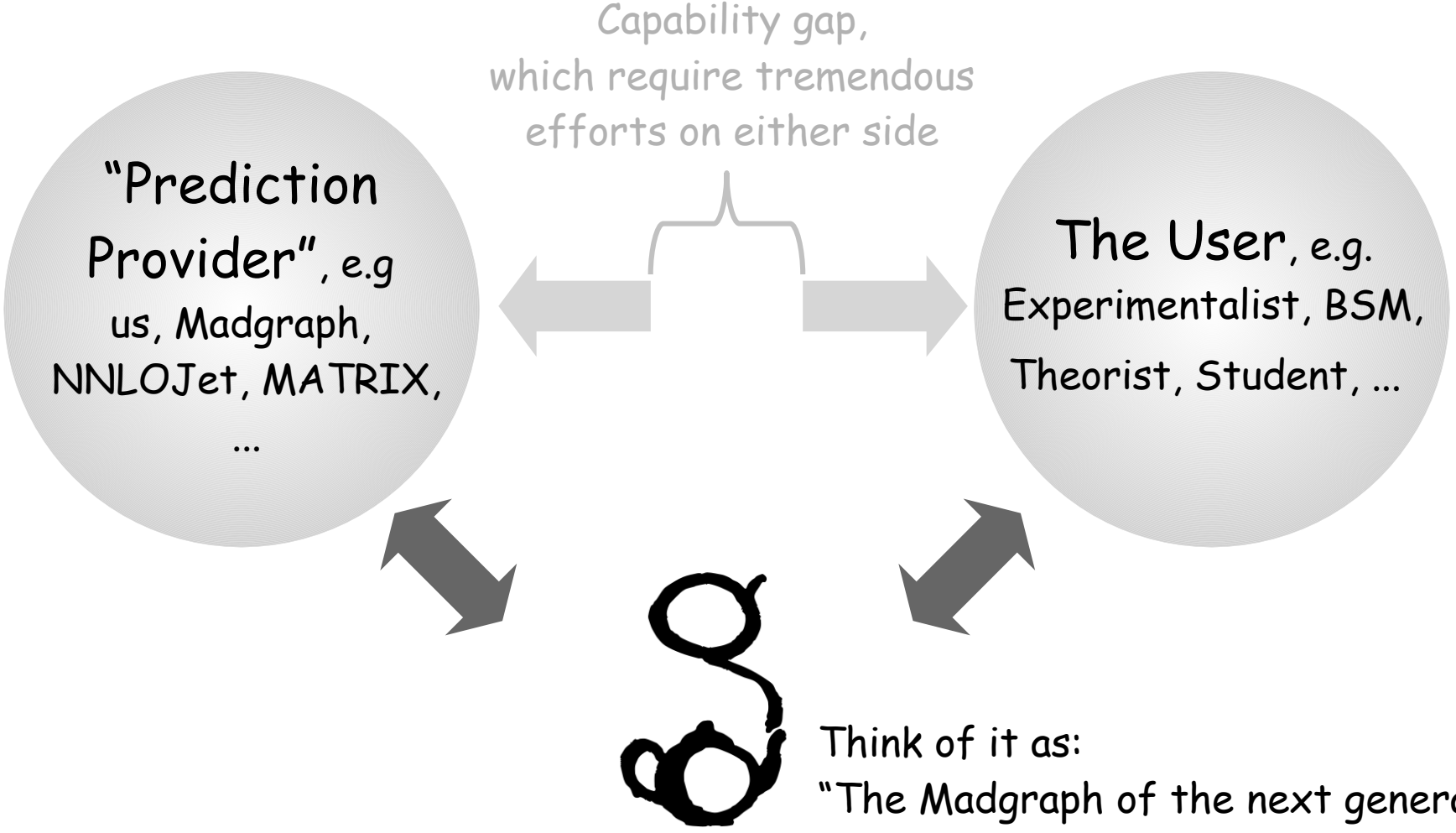
"I'm interested in observable X for process Y"

- Run a public general purpose code (like MadGraph[Alwall'14], Sherpa[Gleisberg'08],...)
  - Implementation of X
  - Expertise to get 'sensible' results
  - 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

The role of HighTEA



The role of HighTEA



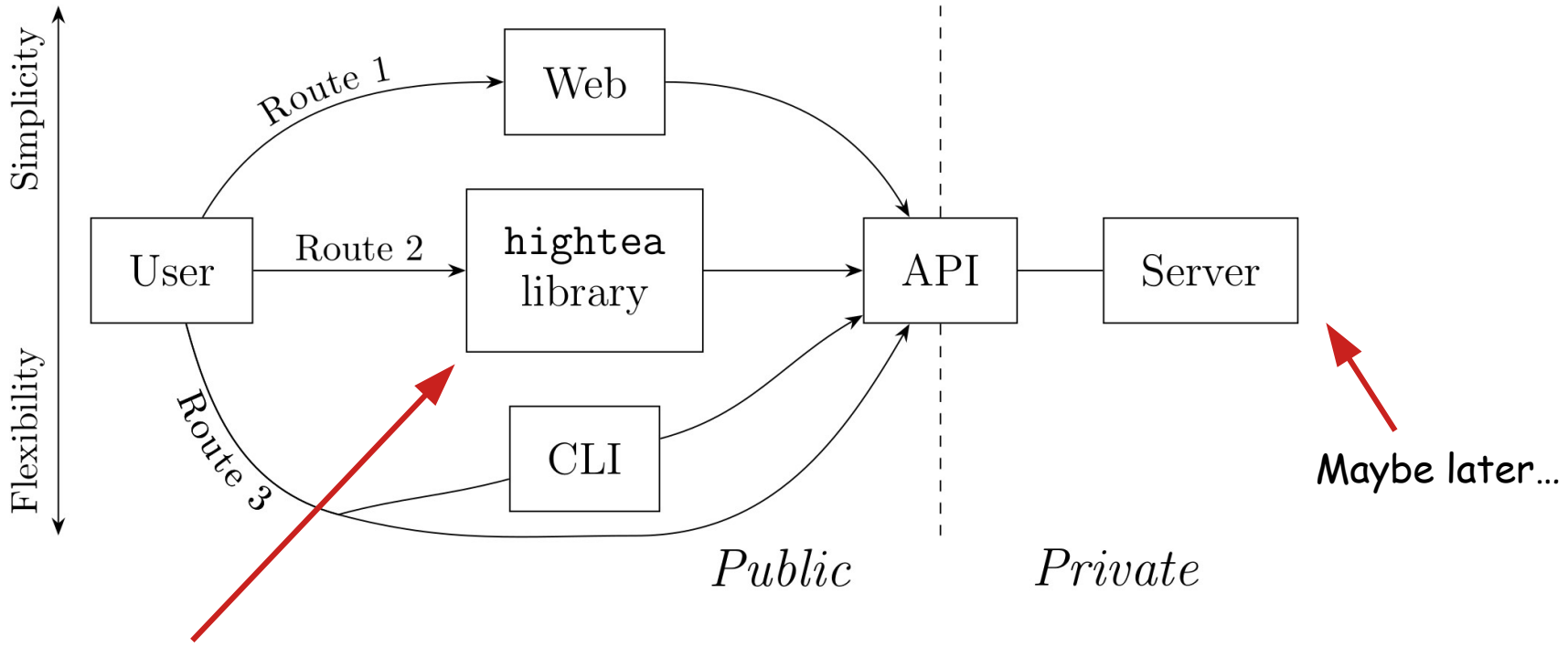
How ?

- Development of a database of precomputed "Theory Events" (Particle level / MC Truth)
  - Currently this means partonic fixed order events
  - Extensions to included showered/hadronized events is feasible
  - Equivalent to a full fledged computation

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

The setup from the user perspective



Demonstration (~20-30 mins)

Current capabilities of the user interface

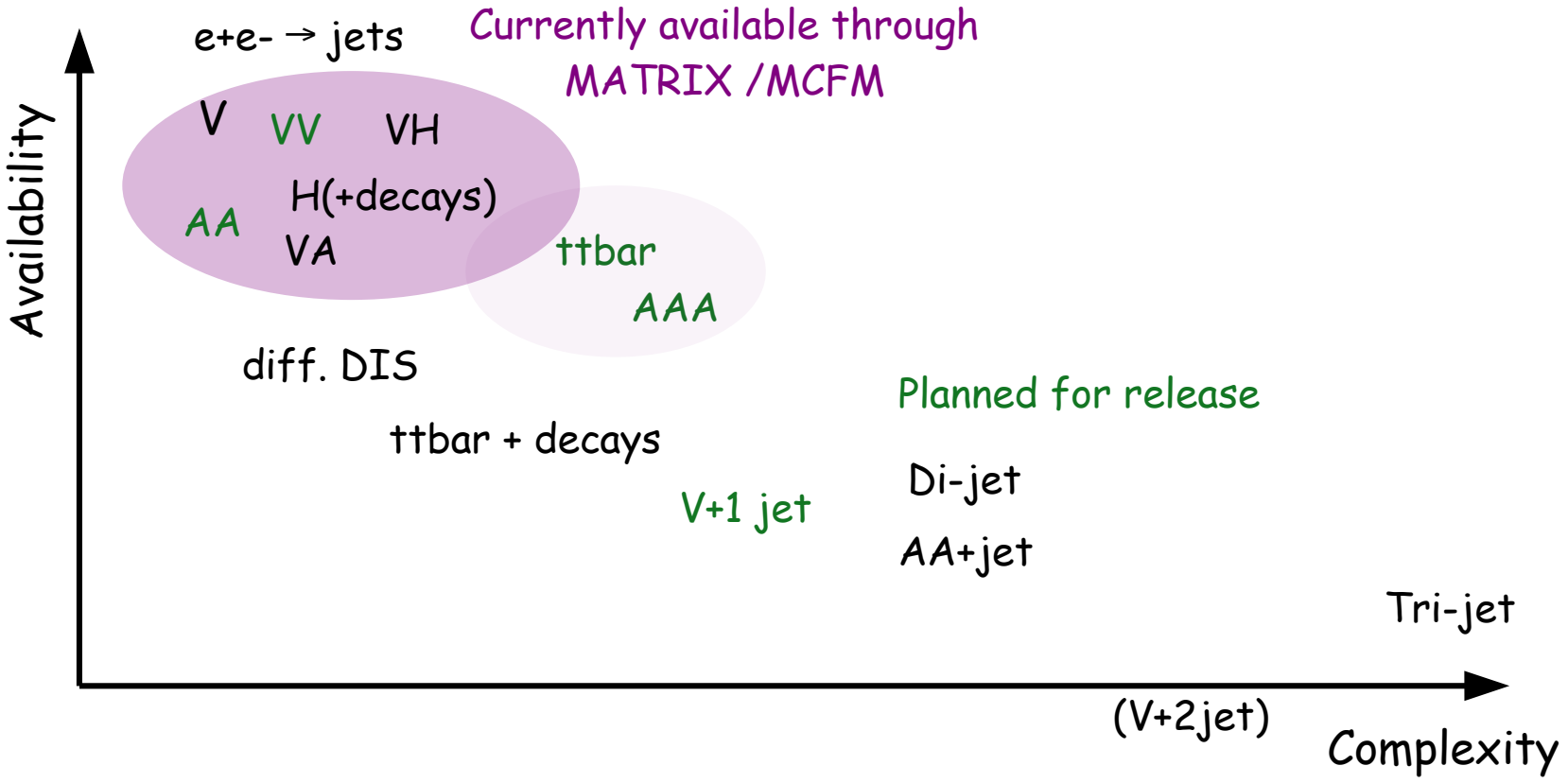
- Selecting a process
- Asking for histograms for observables of the available processes
  - Some observables are pre-implemented
  - Own observables from some basic 4-momenta
  - Free specification of bins including +-Infinity
- Renormalization/Factorization Scale variation:
  - Change of pre-factors and functional form (any observable)
- PDF (member) variation
- Specify phase space cuts
- If jets are present -> jet radius can be changed too.

Web Tour / Interactive Part

- Website & Demo page
- HighTEA-examples: Start.ipynb
  - Tutorial
  - ttbar example
  - The Github page & ReadTheDocs

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
  - Convergence improvement techniques?! → Improving statistical uncertainties
  - Incorporation of HEPMC/LHE files. (Basically putting Madgraph into HighTEA
- Applications:
  - FastNLO/PineApple 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!

Effective Field Theory

Basic idea: Encode potential heavy new physics in terms of effective operators and Wilson coefficients

Data + SM + EFT → constraints of coefficients and new physics

Full potential requires the 'best' theory for the SM, to minimize the effect of fitting 'higher order SM effects'

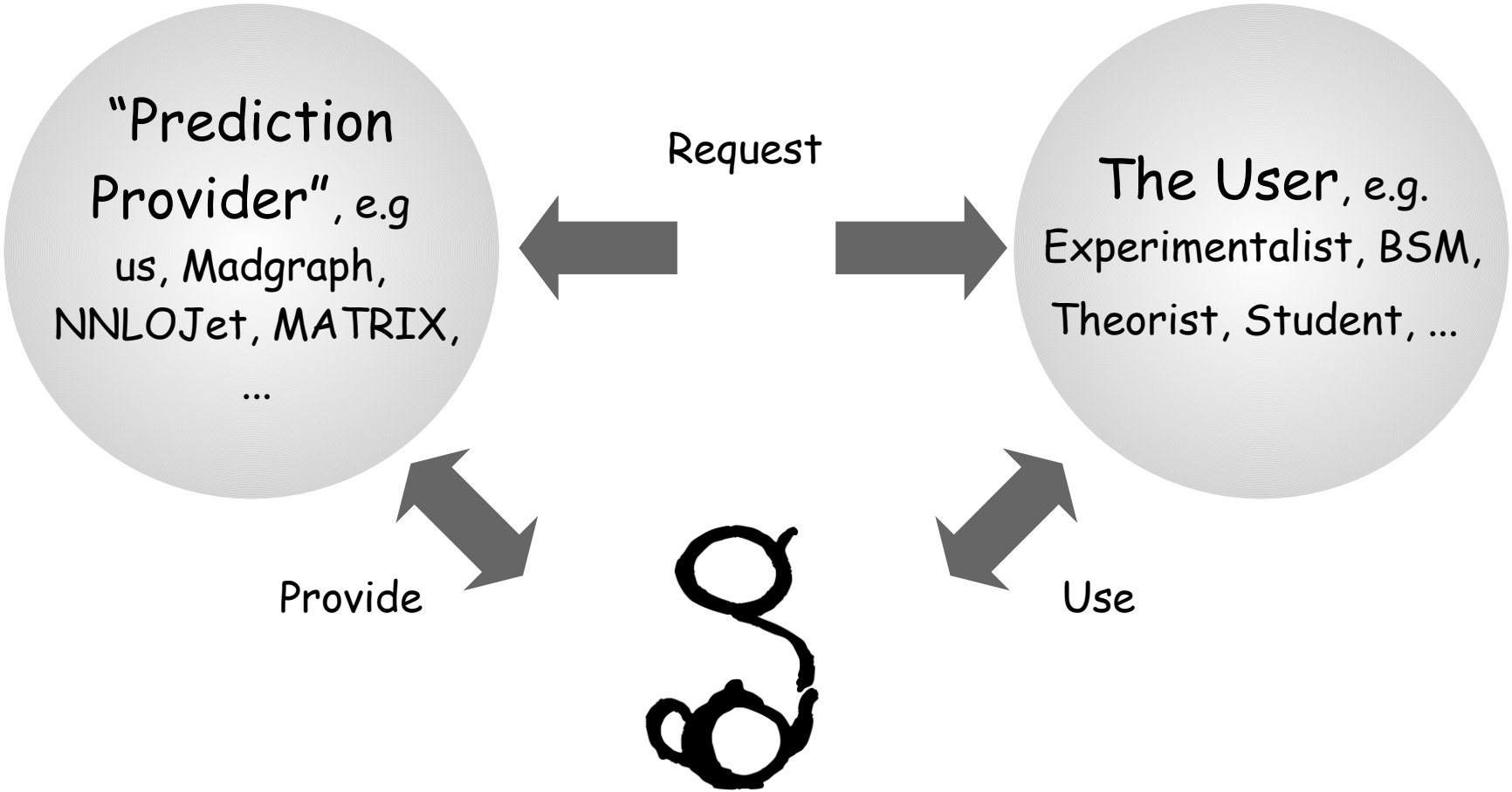
Where to get 'best' theory? → The ideal use case for HighTEA

A possible extension: using Madgraph+"EFT provider" to generate data samples for operators/SM interferences!

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

Individual "datasets"

The Vision


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

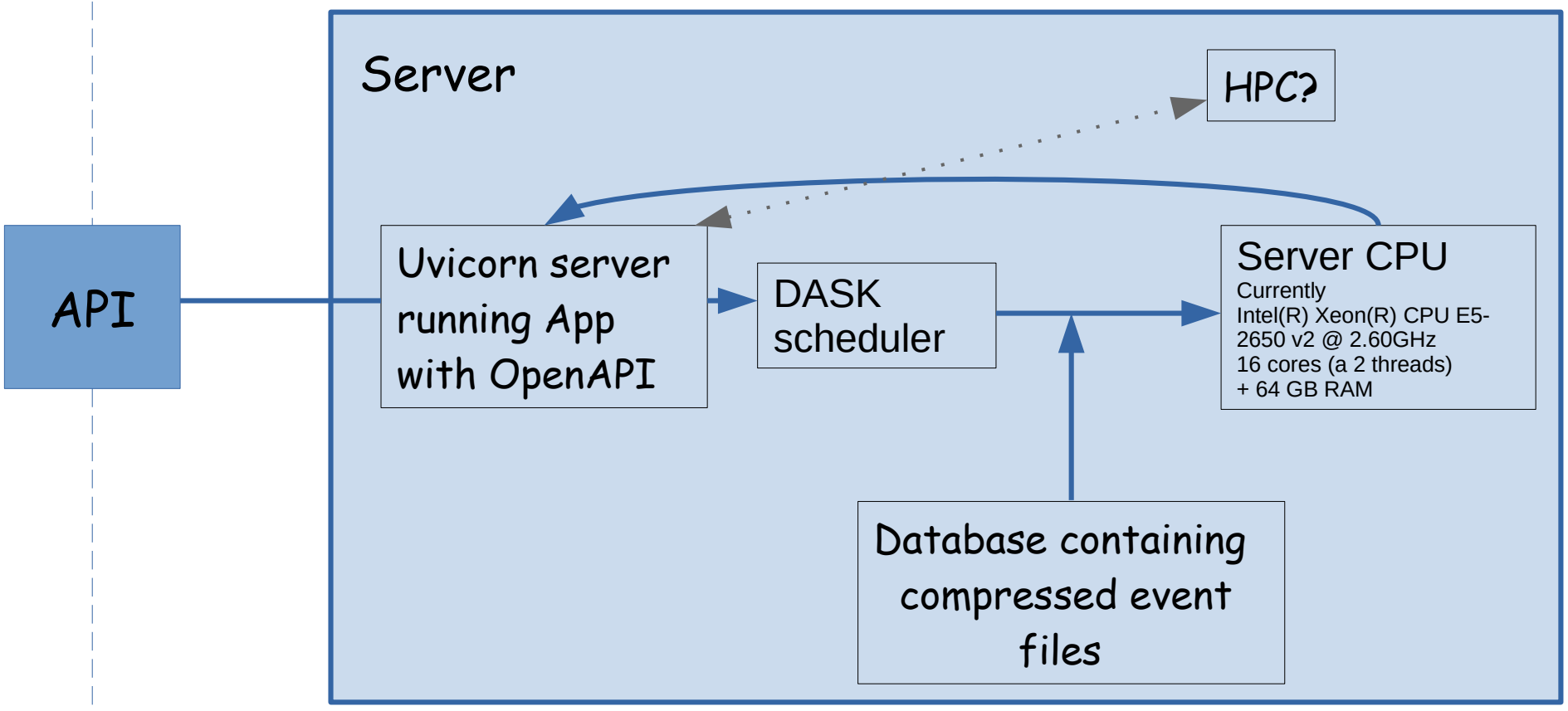
Summary

- 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)

- Outlook: Many, many ideas to implement ;)

Backup

The server


Server
HPC?
API
Uvicorn server running 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:

$$\mathrm{d}\sigma(P_1,P_2) = \sum_{ab} \int\int_0^1 \mathrm{d}x_1 \mathrm{d}x_2 f_a(x_1,\mu_F^2) f_b(x_2,\mu_F^2) \mathrm{d}\hat{\sigma}_{ab}(x_1 P_1, x_2 P_2)$$

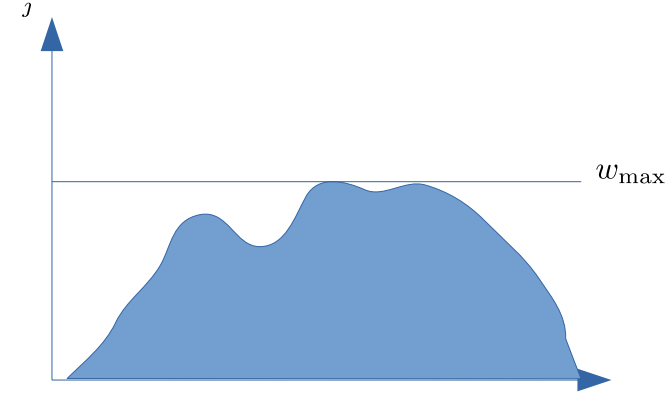
The partonic cross section can be expanded in aS: $\hat{\sigma}_{ab\to X} = \hat{\sigma}^{(0)}_{ab\to X} + \hat{\sigma}^{(1)}_{ab\to X} + \hat{\sigma}^{(2)}_{ab\to X} + \mathcal{O}(\alpha_s^3)$

Using MC method for integration:

Hit-And-Miss Algorithm:

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

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



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

Reweighting

Factorizing renormalization and factorization scale dependence:

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

PDF dependence:

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

AlphaS dependence:

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

Allows full control over scales and PDF