

A decorative horizontal bar with a teal segment on the left and an orange segment on the right.

MadMiner on REANA

Sinclert Pérez,
Johann Brehmer,
Irina Espejo,
Ben Galewsky,
Felix Kling,
Kyle Cranmer



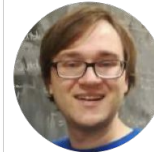
Ben Galewsky



Irina Espejo



Johann Brehmer



Felix Kling



Sinclert Pérez



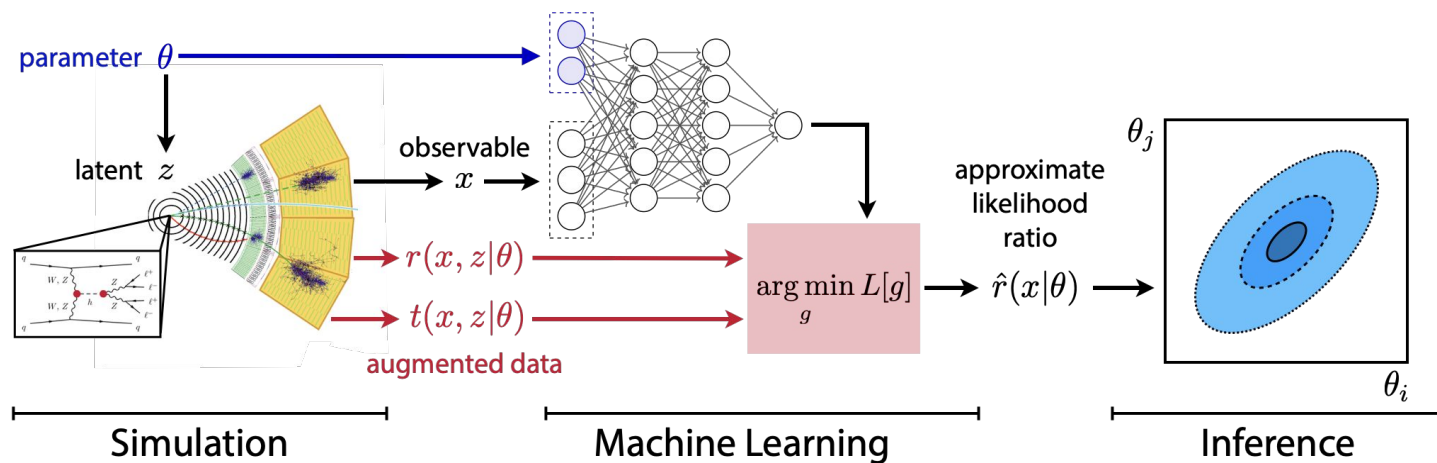
Kyle Cranmer



Madminer: simulation based inference

Simulation-based inference: we want to infer theoretical parameters using a simulator to describe predictions.

MadMiner: a tool that implements various simulation-based inference strategies for particle physics.





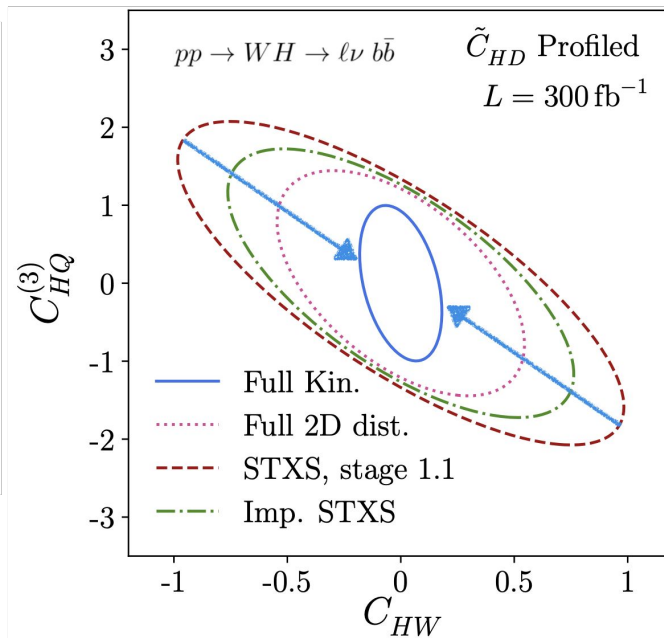
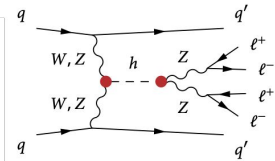
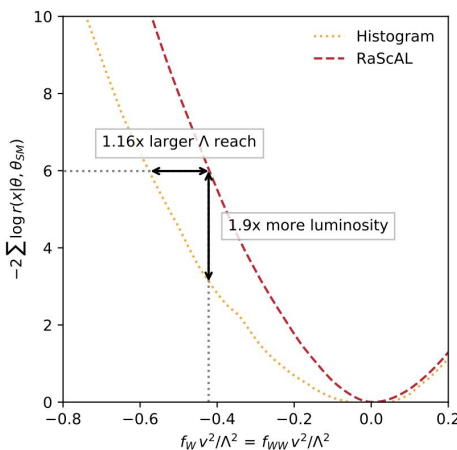
Madminer: impact on physics

Disclaimer: I am not a physicist

Claim: These inference strategies exploit more information from the LHC data and lead to more sensitive measurements.

For more information:

- [Constraining Effective Field Theories with Machine Learning.](#)
- [MadMiner: Machine learning-based inference for particle physics.](#)





Madminer-workflow: containers

MadMiner can be containerized depending on the functionality it provides:

Physics simulation:

Makes use of the following stack:

- **MadGraph**: for describing theoretical models, generating Monte Carlo, “*gold mining*”.
- **Numpy f2py**: to translate Fortran into Python.
- **Pythia8**: for events simulation.
- **Delphes**: for detector simulation & reconstruction.



[madminertool/madminer-workflow-ph](https://github.com/madminertool/madminer-workflow-ph)

ML inference:

Makes use of the following stack:

- **Numpy**: to work with math functions.
- **Pytorch**: to train the evaluate the ML model.
- **Matplotlib**: to show how theoretical values can be approximated.



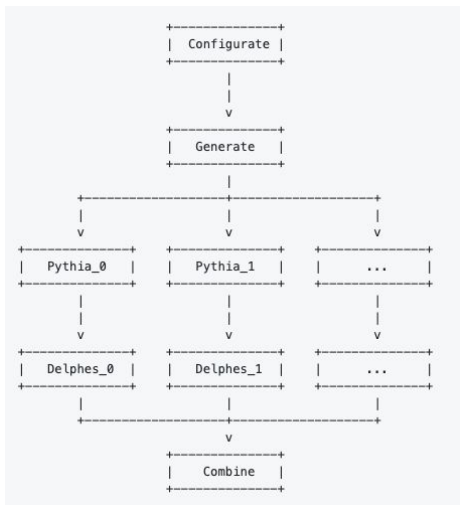
[madminertool/madminer-workflow-ml](https://github.com/madminertool/madminer-workflow-ml)



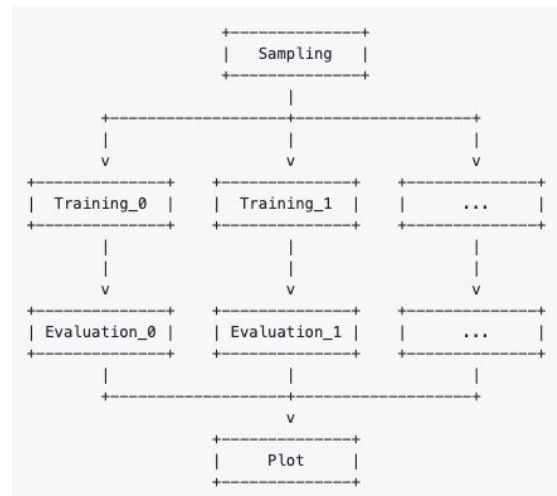
Madminer-workflow: REANA

MadMiner images can be used to coordinate Yadage *sub-workflows*:

Physics simulation:



ML inference:



Madminer-workflow: REANA

Both *sub-workflows* can be combined.

Making the resulting *REANA workflow* a direct **translation** of the original paper MadMiner diagram.

Check out [Scailfin/madminer-workflow](https://github.com/Scailfin/madminer-workflow) for **good documentation** on how to have it deployed.

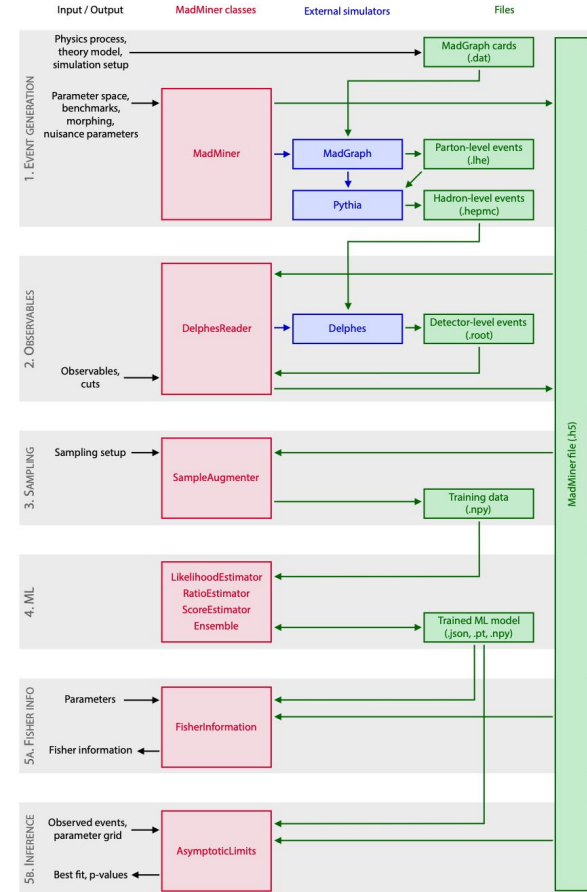
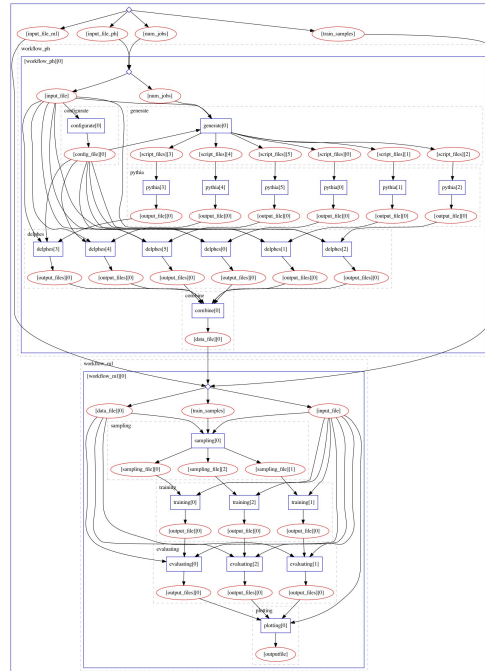


Figure 1. Example workflow, with classes in red, external simulations in blue, and files in green.



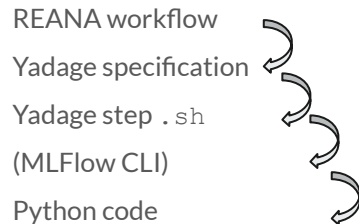
Madminer-workflow: MLFlow

Thanks to its REANA workflow format, *Madminer-workflow* can be run **multiple times**, with **different hyper-parameter configurations**.

How to keep track of each *run* information?



How to propagate hyper-parameters *downfall*?



Search Runs: State: Active Search Clear

Showing 3 matching runs Compare Delete Download CSV

| | Start Time | Run Name | User | Source | Version | Parameters > | | | Metrics > | | | Tags | |
|--------------------------|---------------------|----------|----------|-------------|---------|---------------|-----------------|-----------------|-------------|---------------|-------------|----------|--------|
| <input type="checkbox"/> | | | | | | data_file | eval_folder | inputs_file | theta 0 LLR | theta 0 score | theta 1 LLR | context | method |
| <input type="checkbox"/> | 2020-10-23 15:54:07 | - | sinclert | madminer:ev | - | /madminer/... | /Users/sincl... | /Users/sincl... | 0.083 | - | 0.089 | workflow | alices |
| <input type="checkbox"/> | 2020-10-23 15:54:07 | - | sinclert | madminer:ev | - | /madminer/... | /Users/sincl... | /Users/sincl... | - | -0.014 | - | workflow | sally |
| <input type="checkbox"/> | 2020-10-23 15:54:07 | - | sinclert | madminer:ev | - | /madminer/... | /Users/sincl... | /Users/sincl... | -0.041 | - | -0.048 | workflow | alice |



Madminer: goals and scalability

The overall **goal** is to reproduce the [MadMiner paper](#) results with the *MadMiner-workflow* (REANA).

Until now, papers using *Madminer* were manually managing the complete workflow, taking **weeks** to publish quality results from a Monte Carlo generation.

For the **scalability** of MadMiner:

- Join the ongoing discussions within:
 - SSL (*Kubernetes*).
 - Notre Dame (*VC3/HPC/Bluewaters*).
 - BNL (*Slurm*).
- See Kenyi's talk on **Tuesday** around SCALFIN efforts.





Madminer: References and Use

Review of simulation-based inference

K. Cranmer, **J. Brehmer**, G. Louppe:
“The frontier of simulation-based inference”
PNAS, 1911.01429

Machine learning–based inference methods

J. Brehmer, G. Louppe, J. Pavez, K. Cranmer:
“Mining gold from implicit models to improve likelihood-free inference”
PNAS, 1805.12244

M. Stoye, **J. Brehmer**, K. Cranmer, G. Louppe, J. Pavez:
“Likelihood-free inference with an improved cross-entropy estimator”
NeurIPS workshop, 1808.00973

Probabilistic models

J. Brehmer and K. Cranmer:
“Flows for simultaneous manifold learning and density estimation”
ICML workshop, 2003.13913

Particle physics

J. Brehmer, K. Cranmer, G. Louppe, J. Pavez:
“Constraining Effective Field Theories with machine learning”
PRL, 1805.00013

J. Brehmer, K. Cranmer, G. Louppe, J. Pavez:
“A guide to constraining Effective Field Theories with machine learning”
PRD, 1805.00020

J. Brehmer, F. Kling, I. Espejo, K. Cranmer:
“MadMiner: Machine learning—based inference for particle physics”
CSBS, 1907.10621, <https://github.com/diana-hep/madminer>

J. Brehmer, K. Cranmer, F. Kling, and T. Plehn:
“Better Higgs Measurements Through Information Geometry”
PRD, 1612.05261

Astrophysics

J. Brehmer, S. Mishra-Sharma, J. Hermans, G. Louppe, K. Cranmer
“Mining for Dark Matter Substructure: Inferring subhalo population properties from strong lenses with machine learning”
ApJ, 1909.02005