Particle Physics session of the First MODE workshop

“Differentiable Programming for Experimental Design”

Pietro Vischia[1]

[1] F.N.R.S. and CP3 — IRMP, Université catholique de Louvain

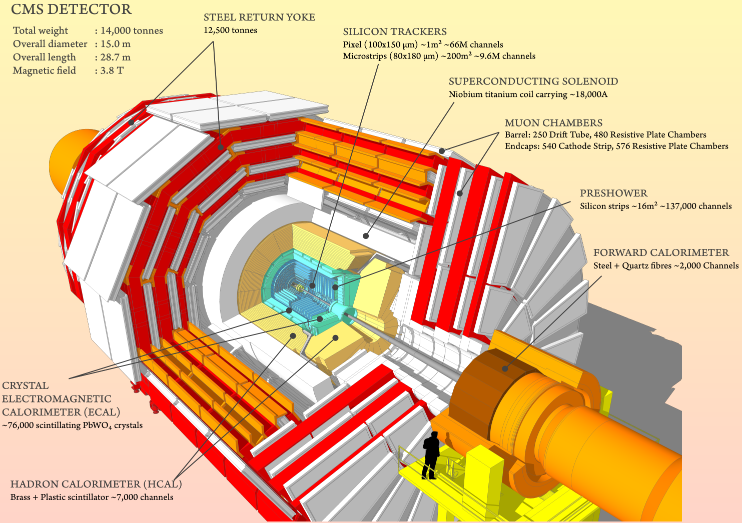
First MODE Workshop

What is MODE about

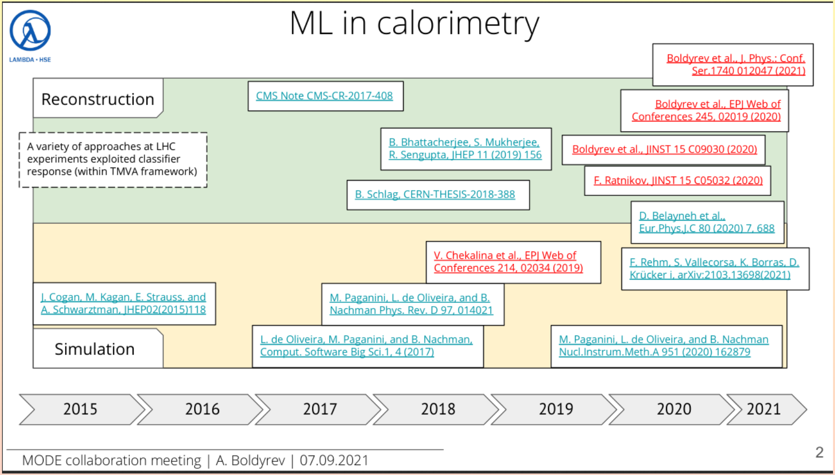
- MODE is mostly about these people being right

$$argmax_{x,y}(\mathcal{L}(x,y)) \neq \left[argmax_x\left(\int \mathcal{L}(x,y)dy\right), argmax_y\left(\int \mathcal{L}(x,y)dx\right)\right]$$

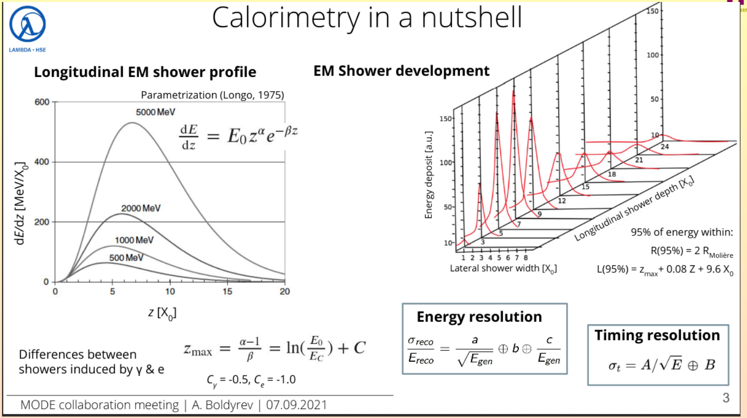
A typical HEP detector is obscenely complex

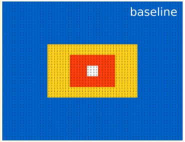


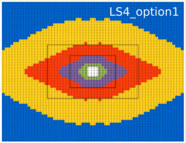
Calorimetry as a first obvious application



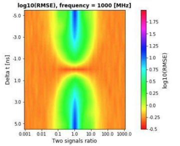
Geometry+timing=challenge (A. Boldyrev)

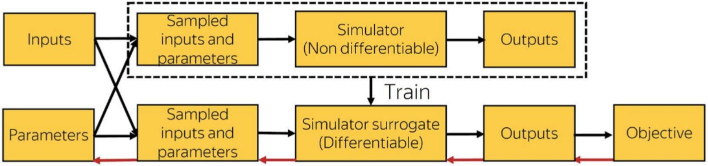






Two signals discriminating

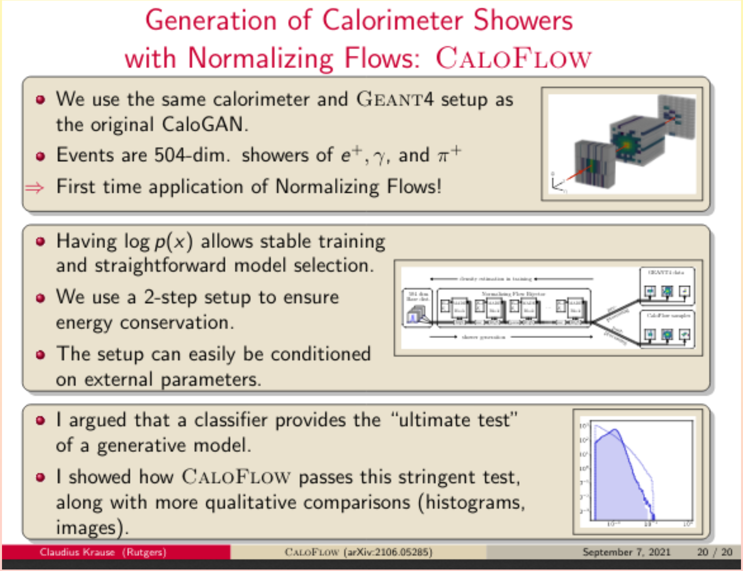




Shirobokov S., Belavin V., Kagan M, AU, Baydin A., NeurIPS'20 paper, arXiv:2002.04632 [cs.LG]

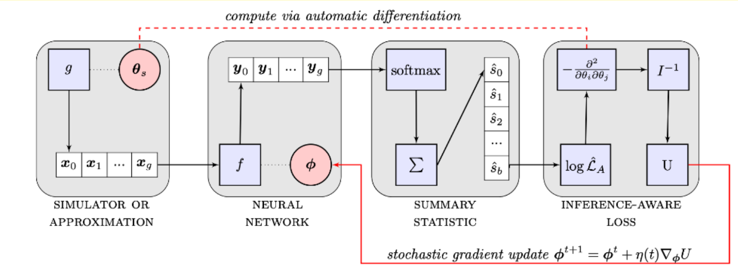
How to generate the inputs to these optimizations? (C. Krause)

- Fast generation of calorimeter showers by generative models



Make your inference aware of uncertainties (L. Layer)

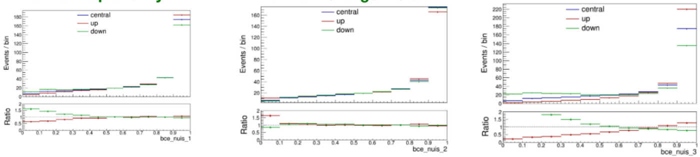
- Extend prototype cases to multiple variables, multiple types of uncertainties



Study II: artificial shift in multiple variables

- Introduce **linear shift** in **up to 3 variables** and train one INFERNO model for each case
- Predict the same **up/down shift** for signal template with BCE model and INFERNO

Binary Cross Entropy



INFERNO

Shift in aplanarity Shift in charge x eta Shift in MET

MC-ensemble-based methods need robust error analysis too (A. Ramos)

UCLouvain
Institut de recherche
en mathématique et physique

Pieces of LQCD computation: From simulations to the proton mass

What are the challenges?

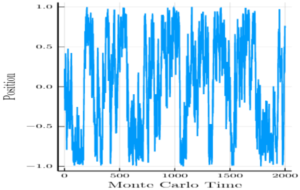
- Observables measured on the **same** configurations are correlated
- Measurements show autocorrelations

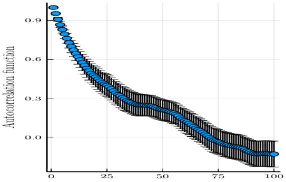
Error analysis in lattice QCD code should be aware of:

- Derived observables: **Complicated functions** of Monte Carlo measurements
- Autocorrelation of observables (affect error estimates)
- Correlation of observables measured on the same ensemble

Handling Monte Carlo data

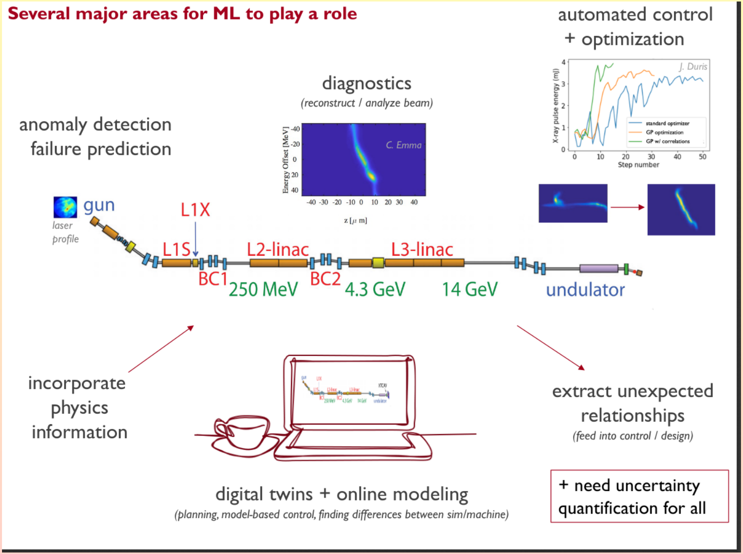
- Generate a random walk in the interval $[-1, 1]$





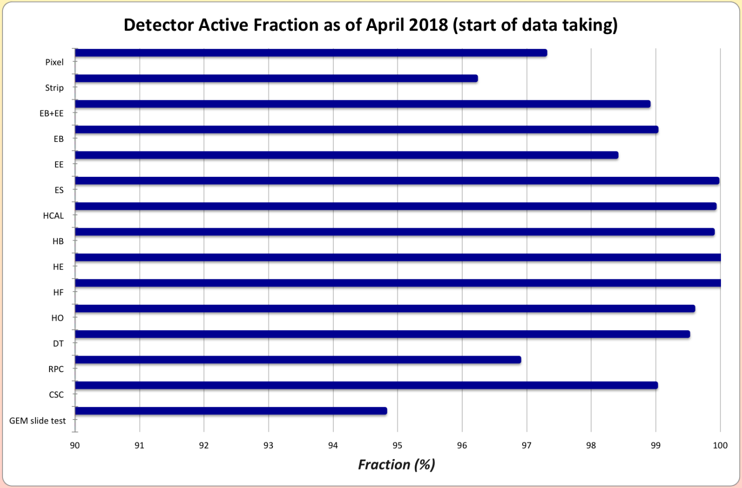
- Input a Monte Carlo history as uwreal. Correlations handled automatically

What about colliders? (A. Edelen)



A personal note on detectors: can we push this further?

- Not all detector channels are active at any given time. Can we account for this?
- Worst-case scenarios, or identification of areas at risk as a function of DC fraction
- Realtime reconfiguration of the detector (shutting down channels?)



Maybe premature, but if interested then let's talk!

Some sociology (or “how to make the community digest this”)

- Hardcore detector developers may resist the introduction of these methods
  - Can all the elements necessary for the AI training be quantified in a way applicable to the AI?
  - There may be the perception that we aim at substituting “many very experienced people working many years” in a naïve and simplistic way
  - There may be feelings that factorization is still the best approach, even if this is easily mathematically disproven
- We must convey our objectives with clarity
- We must make clear that domain knowledge will still be crucial in setting up these pipelines
- We must explain the way we model domain knowledge into our pipelines in a way that is digestible to resisting people

Summary

- HEP detectors are ultracomplex
- Calorimetry is an obvious first candidate to MODE methods
  - Some work already done by MODE members
  - Interested people face similar optimization problems, e.g. in future muon colliders (see L. Sestini slides after this summary)
  - Synergies and collaborations can and should be built
- We are physicists, uncertainties are our daily bread (an uncertain bread, until tenure ☺)
  - Make machine learning aware of systematic uncertainties
  - Deal with tricky correlations and autocorrelations
- Push this further! Can we optimize a detector geometry in real time?
  - Maybe premature, but if interested then let's talk!
- We must not forget sociology
  - The way we present our results may be as crucial as the methods themselves if we want to pioneer these concepts and methods