

# MC tuning in 2023

## More random thoughts on tools and techniques

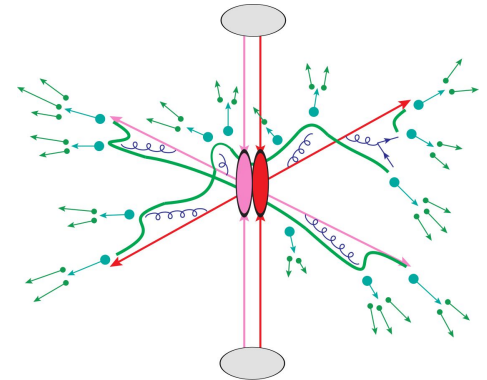
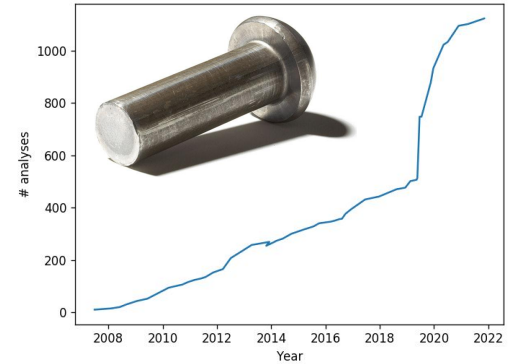
Andy Buckley,  
University of Glasgow

HSF MC Tuning Workshop  
28 June 2023



# What happened since 2015?

- ❖ Tool-wise, not an awful lot! Tuning of MPI in particular was crucial in ~2009-12, and was “good enough” by the time of LHC Run 2... people moved on
- ❖ Professor itself got used in neutrino physics, EFT studies, (and PDFs), but 2014-15 was the tuning high-water point.
- ❖ Times change: lots more data, observables, and a physics case from high-precision Run 3. Time to return... carefully
- ❖ Personnel changes & moves  $\Rightarrow$  development of Prof3 became Apprentice. Then main physics-dev moved on
- ❖ I still have a half-finished paper with Holger on making tune errors robust... but it needs a use-case to be worth the completion!
- ❖ Personal opinion: attempts at further tunes often got bogged down in death-by-committee. This is an area where
  - a) know specifically what problem you want to solve
  - b) get your hands dirty, and iterate; don't overplan



# Professor or Apprentice?

- ❖ You tell me! I updated Professor for the first time in ~8 years for an MSci project this year; another this summer will move forward more. It's not dead, but...



...sometimes they  
*don't* come!

# Professor or Apprentice?

- ❖ You tell me! I updated Professor for the first time in ~8 years for an MSci project this year; another this summer will move forward more. It's not dead, but...
  - Professor fell in the gap between old and new ways, particularly the rise of Numpy, etc.: v1 was "old Numpy", v2 was C++ core! Starting now, I'd write in "new Numpy"... which is pretty much Apprentice. Or hack from scratch
  - And rational interpolation is a clear solution to the problem of normalised polynomials, which are not themselves polynomials.
  - Is it supported? Is any?! How much hacking to expect?
  - We can do the core Professor numerical method in a few lines of Python now. [Good-enough Pade in Scipy](#). CPU not a bottleneck. Lessons learned about how to organise tunes and data-flow: maybe don't expect "frameworks"?

# More random thoughts

## ❖ Any room left in surrogate models?

- Sure... neural regression in place of polynomials or rationals is under-explored.
- And hypersphere sampling, quasirandom sampling, ... easy
- Even with polynomials, you can get custom: e.g., you don't *need* all the polynomial cross-terms: restricting correlations to lower powers can tame runaway scaling. Too custom for a UI...?

## ❖ Or something else?

- Serial or semi-serial methods like GAMPI, Bayesian optimisation, Dctr. Autotune as an attempt to systematise weight-setting. Papers, but no active use? Scaling past e+e- or single jet?
- Intellectually fun/impressive, but you need a killer app to beat trivially parallelisable surrogate models.

## ❖ Weights & well-defined uncertainties!

- More important! And how those uncertainties interplay with “theory” ambiguities such as scale choices. Need physics & holistic view to fairly squeeze down systematics.

