technische universität dortmund

Fakultät Physik

ML & experimental particle physics

Johannes Albrecht (TU Dortmund & Lamarr)

20.03.2024

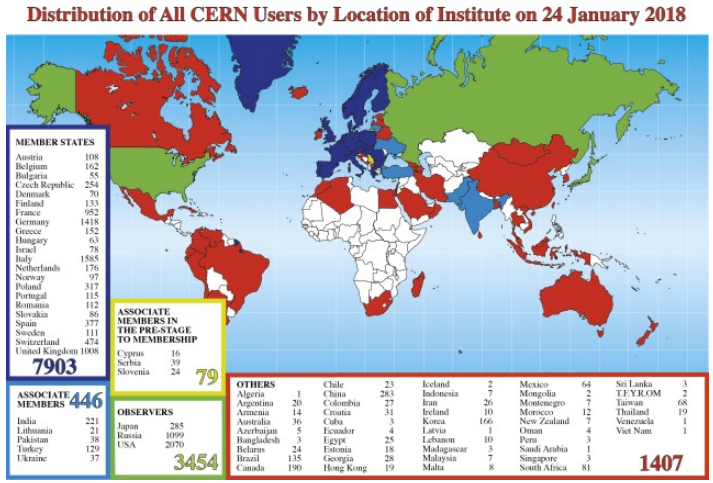
LAMARR INSTITUTE FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

B3

Many references to B. Mitreska and Q. Führing, Lamarr meeting 8.11.23

Content

- Particle Physics: Experimental situation very interesting
  - Entire community joined forces at one accelerator center at a time
    → extremely focussed field
  - We have 10 years of LHC data and will collect another 20
    → huge dataset to exploit optimally

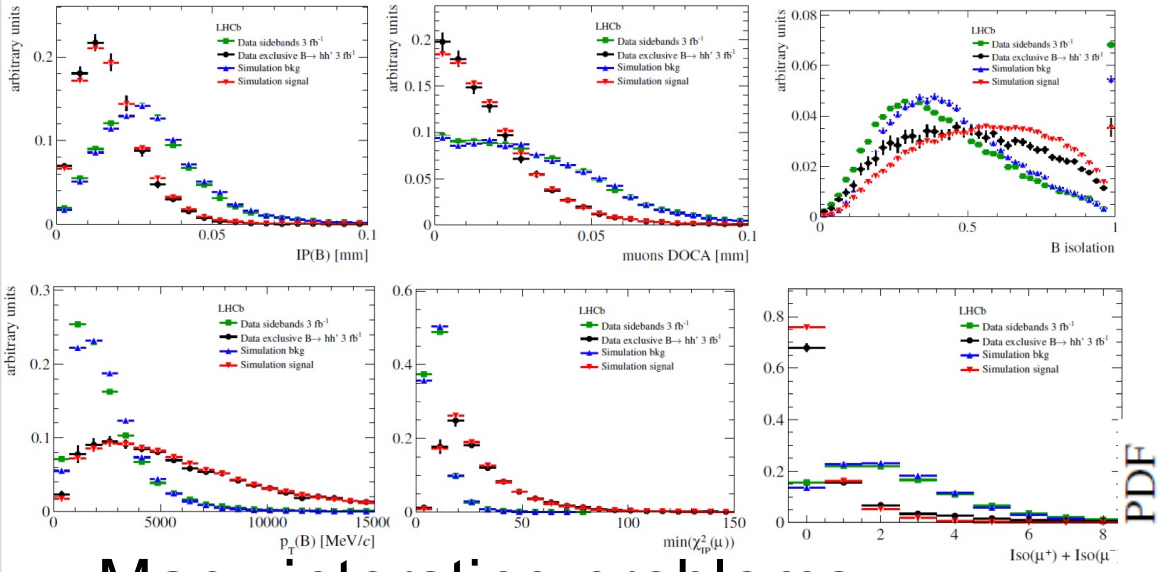


- Modern Particle Physics is (in a part) data science
  - Process 4TB/s of data in „real time“
  - Many hypothesis based approaches, often optimal, sometimes not?
  - Better access to data in any way has large potential

- Particle physicsts are not (entirely) naive
  - No quick solutions, joint projects needed and started within Lamarr, more collaborations welcome

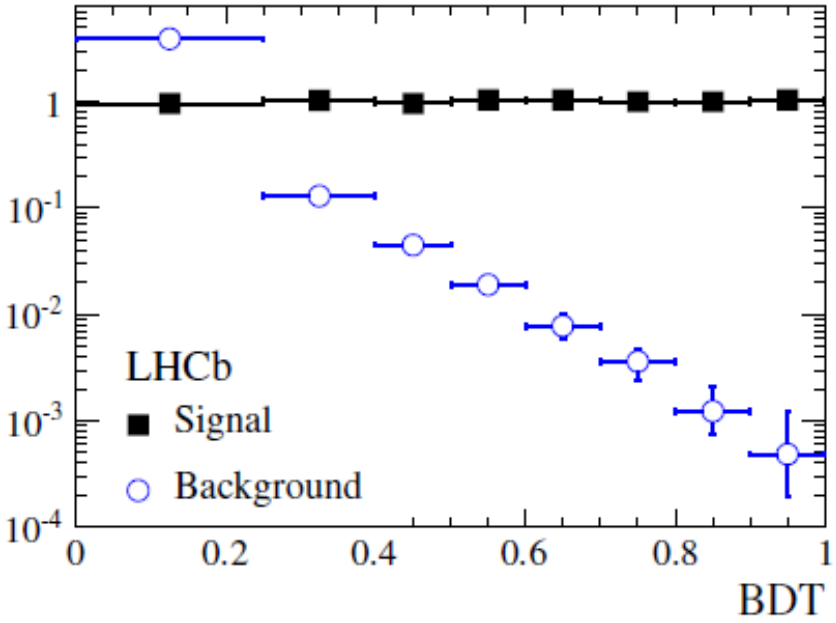
Analysis: signal separation

- Typical: many variables, all with some separation power
  - Total of 12 variables, including:

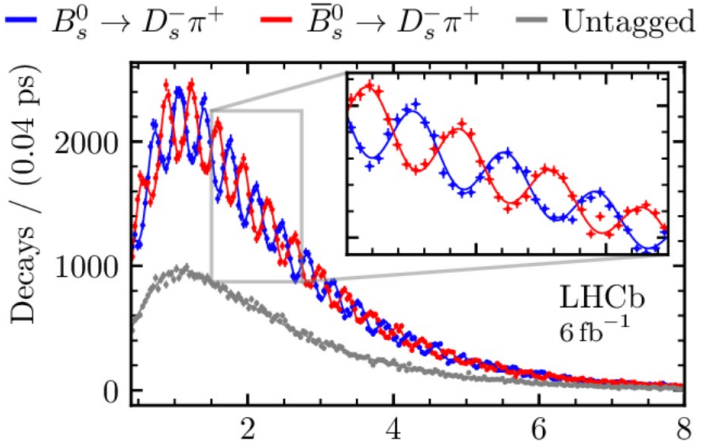


- Many intersting problems:
  - Simulation: enourmous ressources, event generation (learning possible?), detector response, etc
  - Smart use of entire event?
  - Hypothesis based analysis
    → learned Standard Model ?

- Combine variables in clasifier (BDT, NN, ..)
- Problem: simulation & signal proxy
- Put some thought in Labelling ?

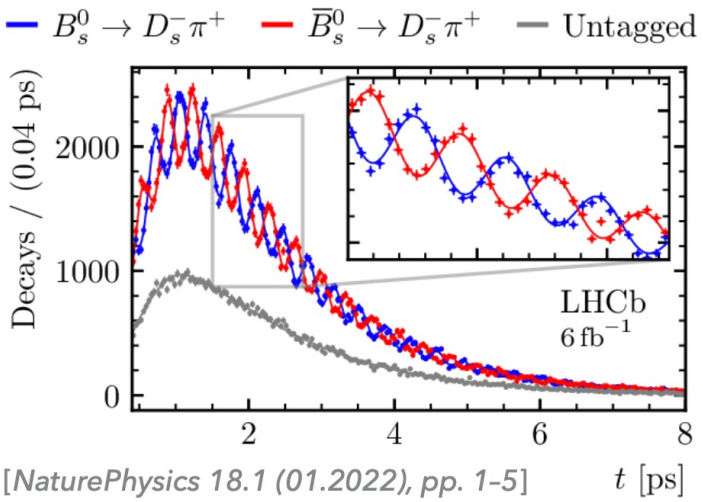


Flavor Tagging: a whole playground

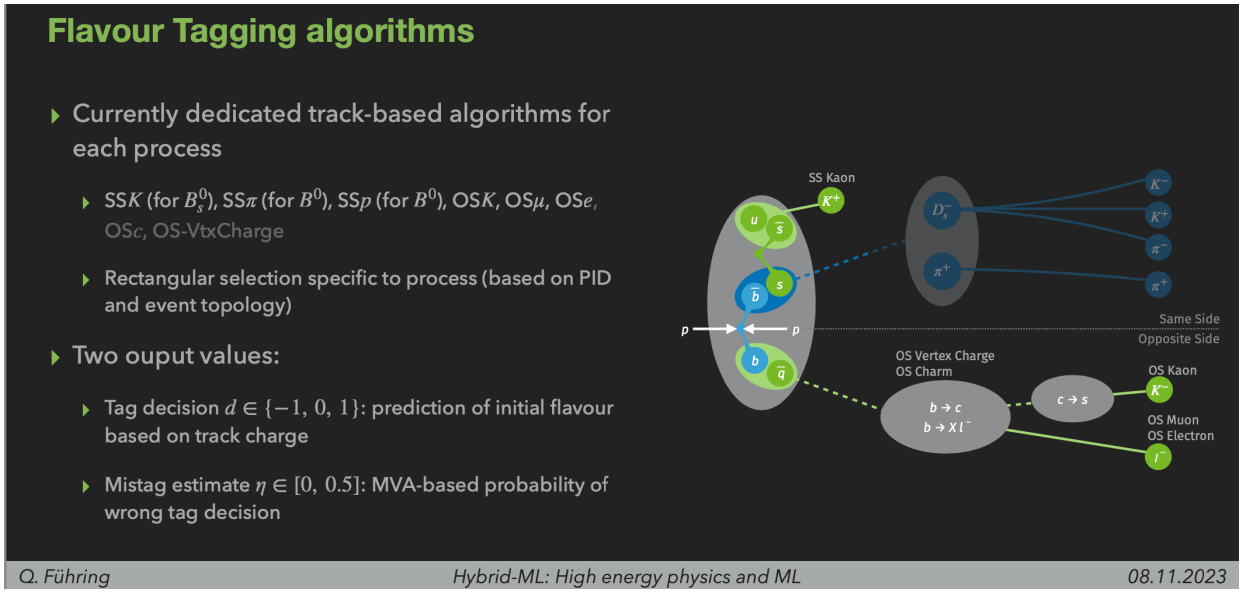


[*NaturePhysics 18.1 (01.2022), pp. 1–5*]

Flavor Tagging: a whole playground

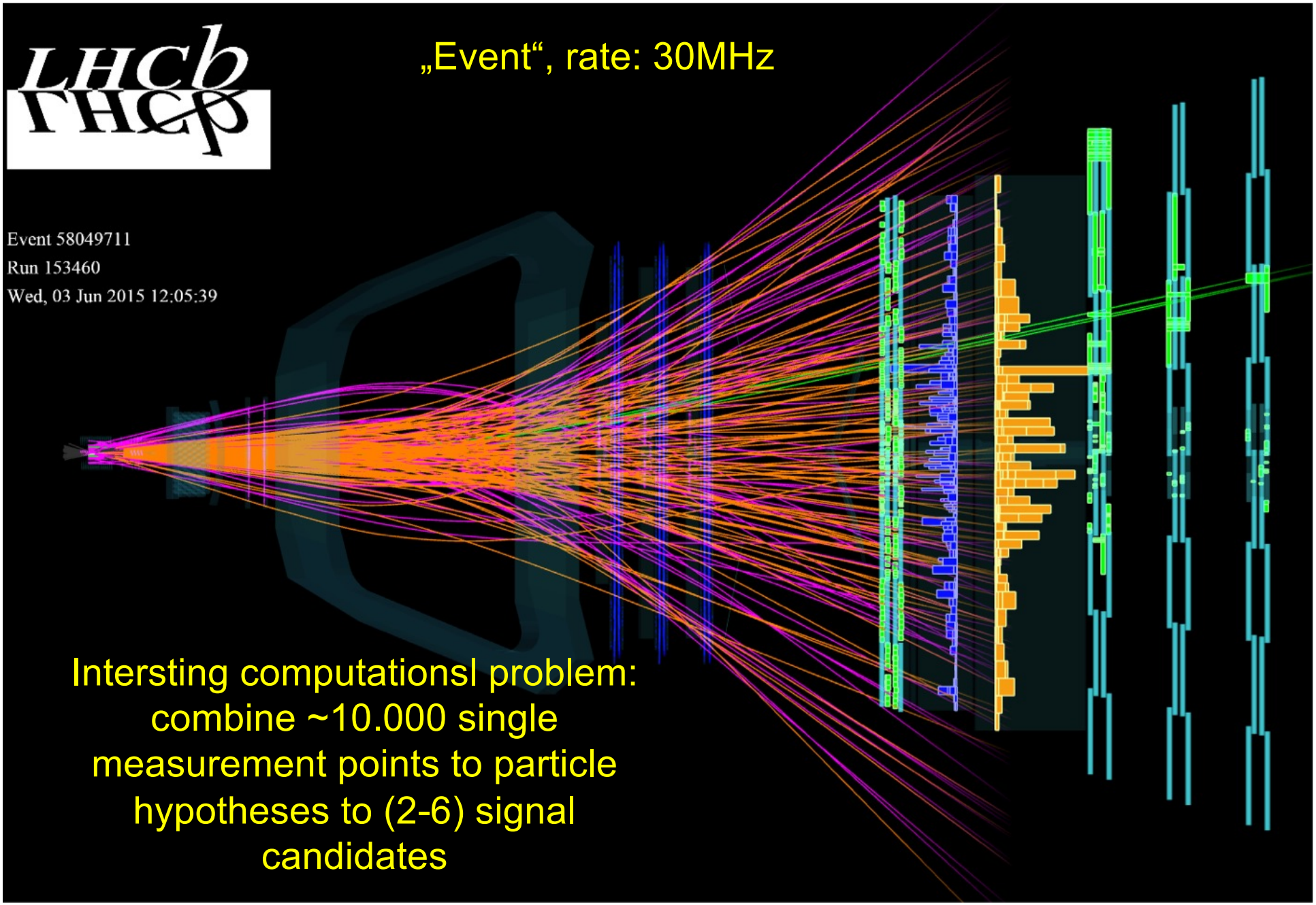


[NaturePhysics 18.1 (01.2022), pp. 1–5]



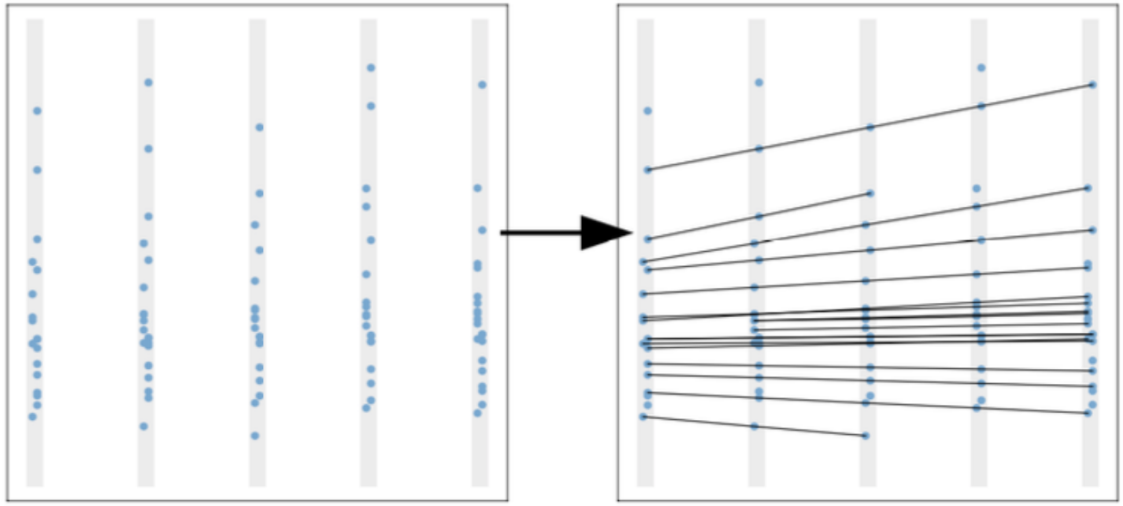
- Limits performance of analysis, currently, we loose factor 20
- Complex, hypothesis based analysis → can we do better?
- Inclusive tagger in development (since many years …)
- Differences: simulation-data known, but complex
  - Often recovered by multivariate reweighting
  - Alternative approaches to be studied (e.g. Domain Adaptation)
- Work in progress (Führing, Bunse, Popp)

Detector images

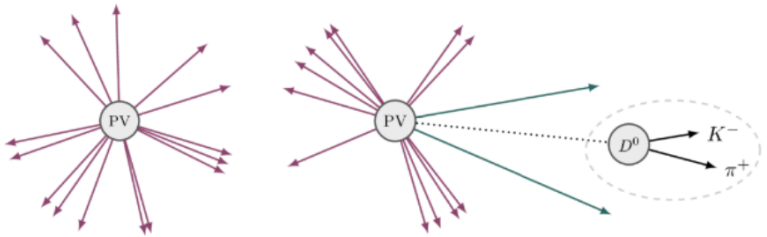


Track and vertex reconstruction

- Correctly match the hits in detector planes
- p-p colisions are busy with producing many tracks



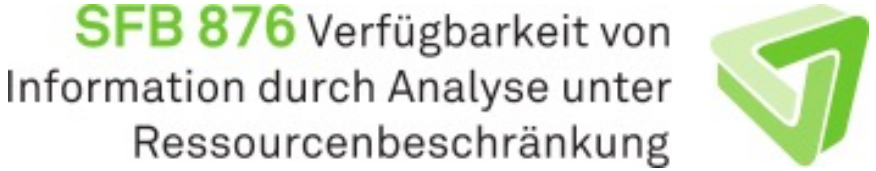
- We need reconstructed tracks to form particles



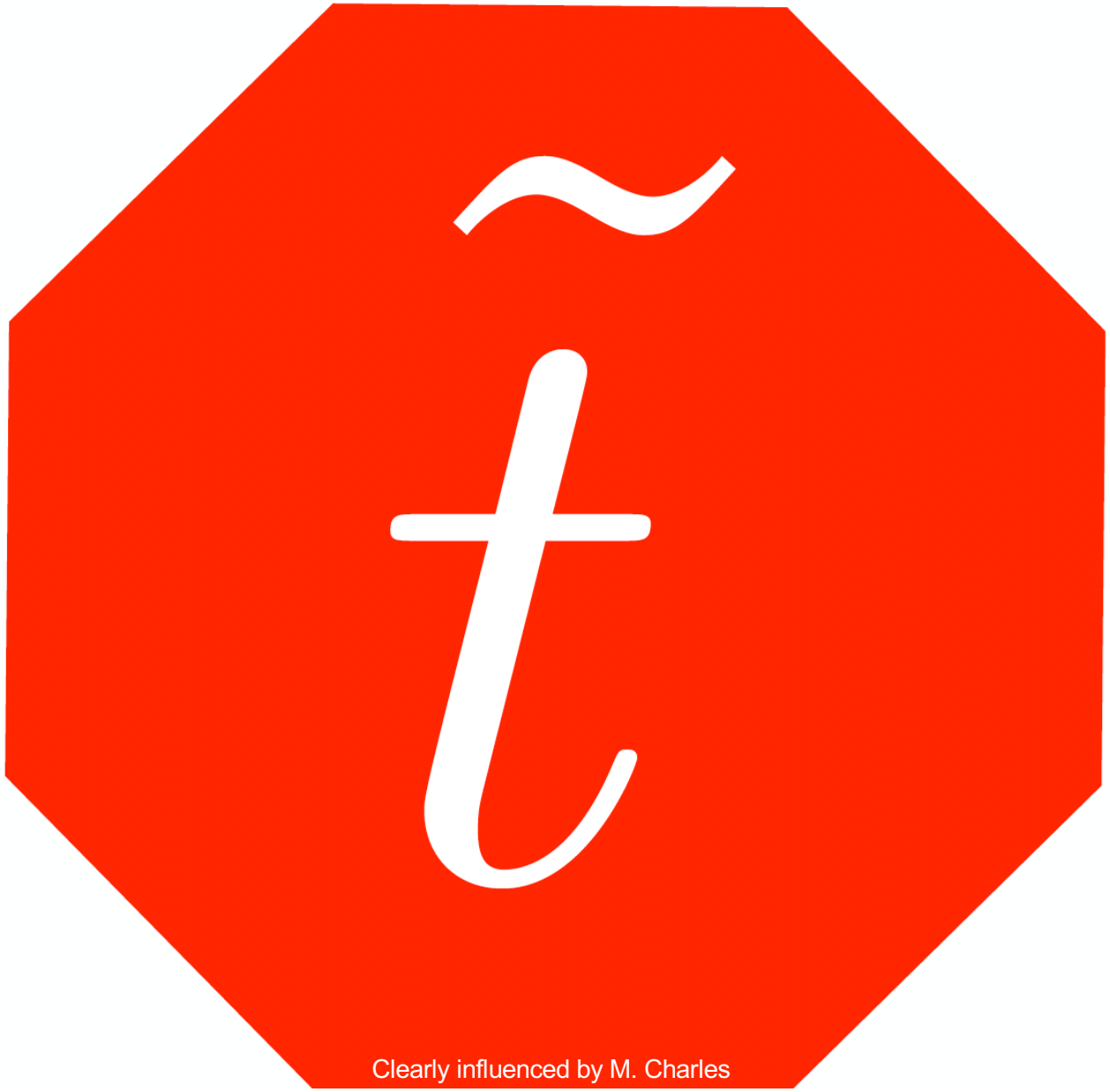
Many problems where expert knowledge might be able to help:

- Pattern recognition
- Parameter estimate
- Ring finding
- Anomaly testeing
- System architecture, hybrid optimization
  (Volume: 400 nvidia A5000 GPUs + O(70k) Intel E5-2630 equivalent physical cores, 30PB disc, 1-2M network cost)
- Fast and efficient candidate selection ~signal fraction up to $1{:}10^{10}$
- …

Stamp Collection / Danksagung and Fördergeber



There is still some space left…


Clearly influenced by M. Charles