WHICH CONTINENT IS THIS MOUNTAIN IN?

Kudos to N. Wardle

SEANPIPPDA: ONE PHYSICIST'S VIEW

A. David (CERN and IST-Lisboa)

SEANPIPPDA: ONE PHYSICIST'S VIEW

A. David (CERN and IST-Lisboa)

MY VIEW IS LIMITED

I cannot claim to have understood everything.

- There's a reason Mikael has 2× more time.
  - Remember what physicists think of factors of 2.

This view is **incomplete** and **has systematics.**

- Both are my fault, not the speakers' or discussion partakers'.

My goal: share what I learned and what I think the future can bring.

SEANPIPPDA: ONE PHYSICIST'S VIEW

A. David (CERN and IST-Lisboa)

SEANPIPPDA

PHYSTAT

SEANPIPPDA

PHYSTAT 2021 ± 2

A BEPHYTTING DIMIDATION?

PHYTISTICIAN

A BIASED SAMPLE FROM OUR CORRESPONDING DISTRIBUTIONS' TAILS

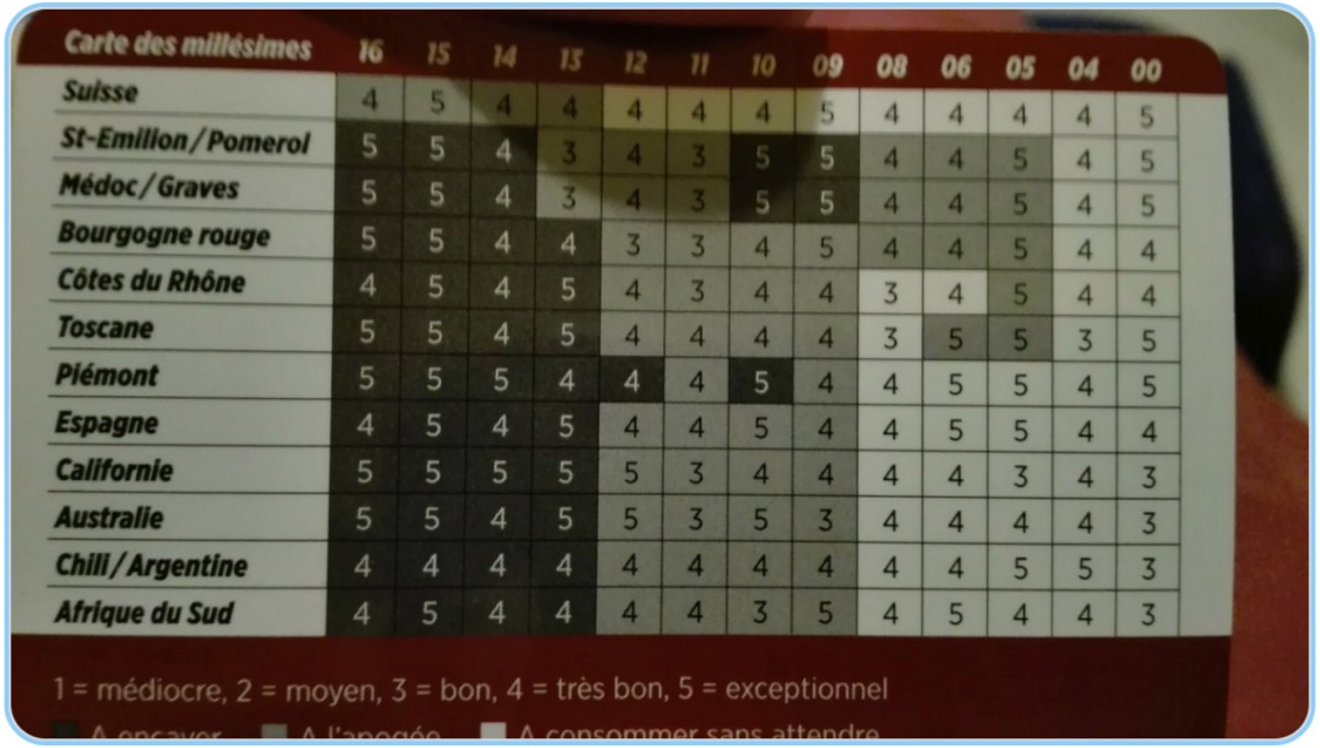
Statistically-concerned physicists who can do some math.

Physically-intrigued statisticians who are willing to help.

Thank you for being interested across the disciplines!

Patrick Koppenburg 🐧
@PKoppenburg

If wines can do it why not research?



Carte des millésimes	16	15	14	13	12	11	10	09	08	06	05	04	00
Suisse	4	5	4	4	4	4	4	5	4	4	4	4	5
St-Emilion/Pomerol	5	5	4	3	4	3	5	5	4	4	5	4	5
Médoc/Graves	5	5	4	3	4	3	5	5	4	4	5	4	5
Bourgogne rouge	5	5	4	4	3	3	4	5	4	4	5	4	4
Côtes du Rhône	4	5	4	5	4	3	4	4	3	4	5	4	4
Toscane	5	5	4	5	4	4	4	4	3	5	5	3	5
Piémont	5	5	5	4	4	4	5	4	4	5	5	4	5
Espagne	4	5	4	5	4	4	5	4	4	5	5	4	4
Californie	5	5	5	5	5	3	4	4	4	4	3	4	3
Australie	5	5	4	5	5	3	5	3	4	4	4	4	3
Chili/Argentine	4	4	4	4	4	4	4	4	4	4	5	5	3
Afrique du Sud	4	5	4	4	4	4	3	5	4	5	4	4	3

1 = médiocre, 2 = moyen, 3 = bon, 4 = très bon, 5 = exceptionnel

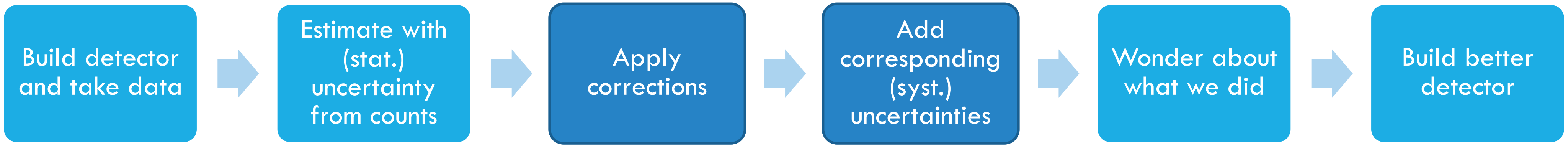
À encaver · À l'apogée · À consommer sans attendre

11:57 PM · Jan 30, 2023 · **443** Views

OUR SYSTEMATIC PLIGHT WITH BIAS

Build detector and take data → Estimate with (stat.) uncertainty from counts → Apply corrections → Add corresponding (syst.) uncertainties → Wonder about what we did → Build better detector

OUR SYSTEMATIC PLIGHT WITH BIAS



OUR SYSTEMATIC PLIGHT WITH BIAS

Build detector and take data → Estimate with (stat.) uncertainty from counts → Apply corrections → Add corresponding (syst.) uncertainties → Wonder about what we did → Build better detector

USUALLY

BANFF
CENTRE
FOR ARTS AND
CREATIVITY

90 YEARS
1933–2023

Banff International Research Station
for Mathematical Innovation and Discovery

THIS WEEK

BANFF

90 YEARS 1933–2023

CENTRE FOR PHYSICAL DARK ARTS AND CREATIVE SOLUTIONS

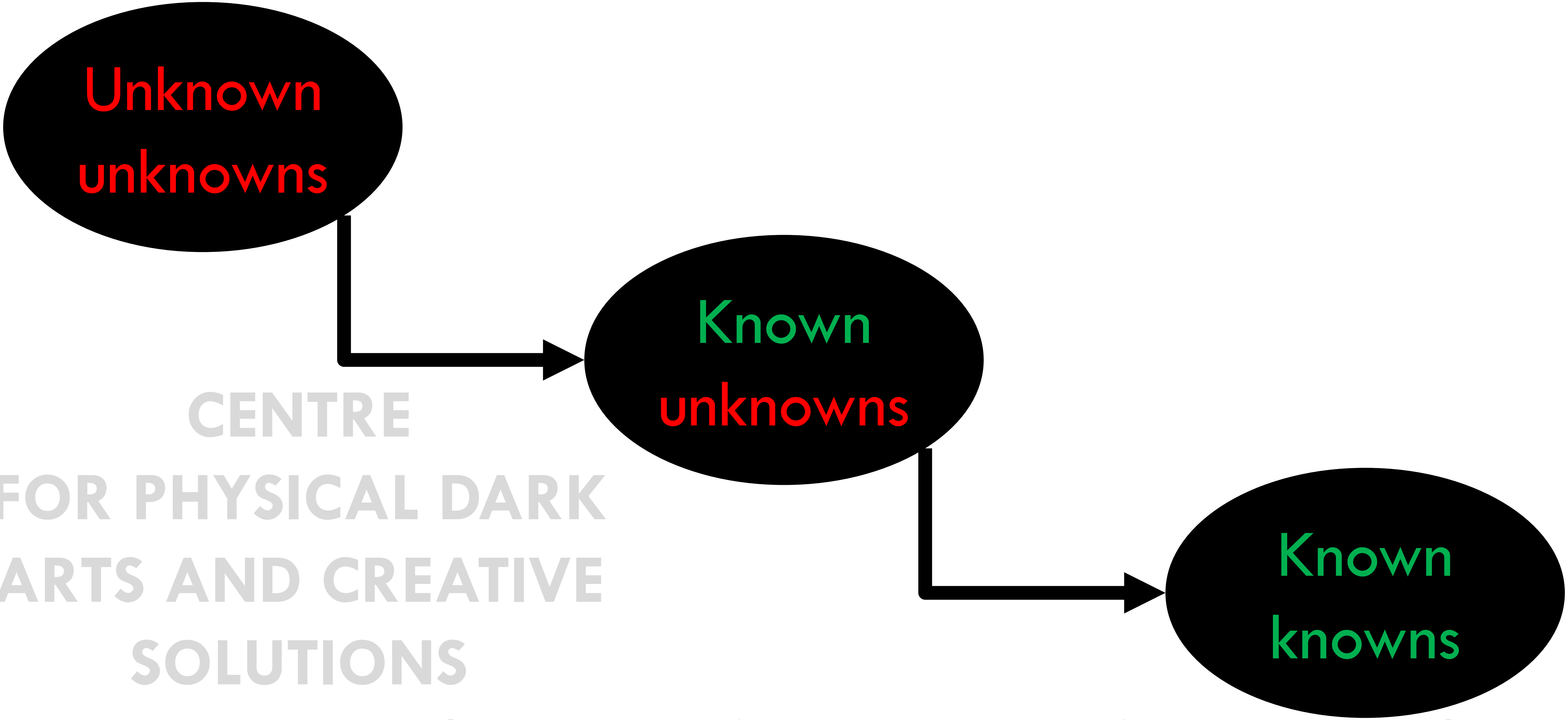
Banff International Research Station for Mathematical Innovation and Discovery

THIS WEEK

BANFF 90 YEARS 1933–2023

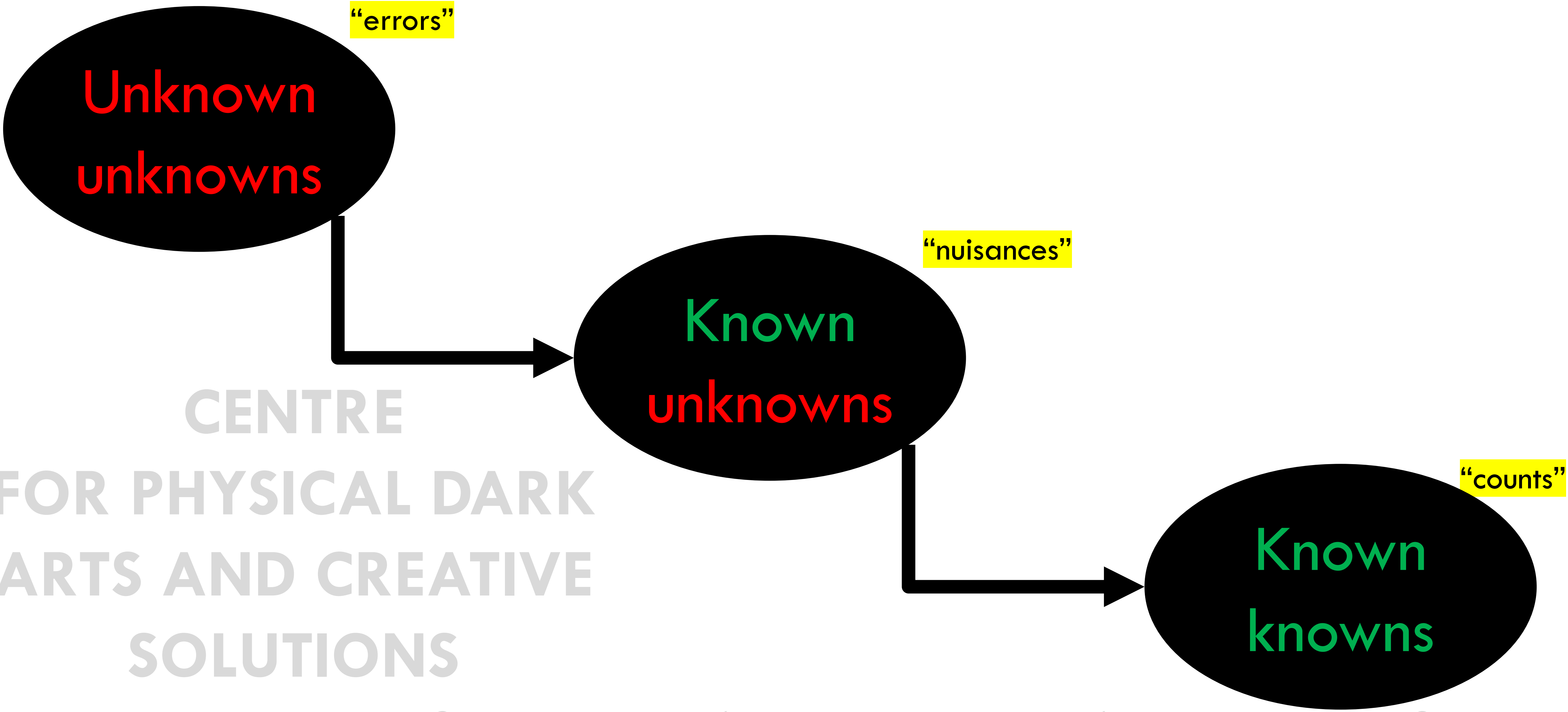
CENTRE
FOR PHYSICAL DARK
ARTS AND CREATIVE
SOLUTIONS

Banff International Research Station
for Statistical Innovation and Discovery of
Particle Physics Analysis ingredients



CENTRE
FOR PHYSICAL DARK
ARTS AND CREATIVE
SOLUTIONS

for Statistical Innovation and Discovery of
Particle Physics Analysis ingredients



CENTRE
FOR PHYSICAL DARK
ARTS AND CREATIVE
SOLUTIONS

for Statistical Innovation and Discovery of
Particle Physics Analysis ingredients

Unknown unknowns

"errors"

Typically Physicists' hunting grounds

Known unknowns

"nuisances"

Typ. Statisticians' hunting grounds

Known knowns

"counts"

CENTRE FOR PHYSICAL DARK ARTS AND CREATIVE SOLUTIONS

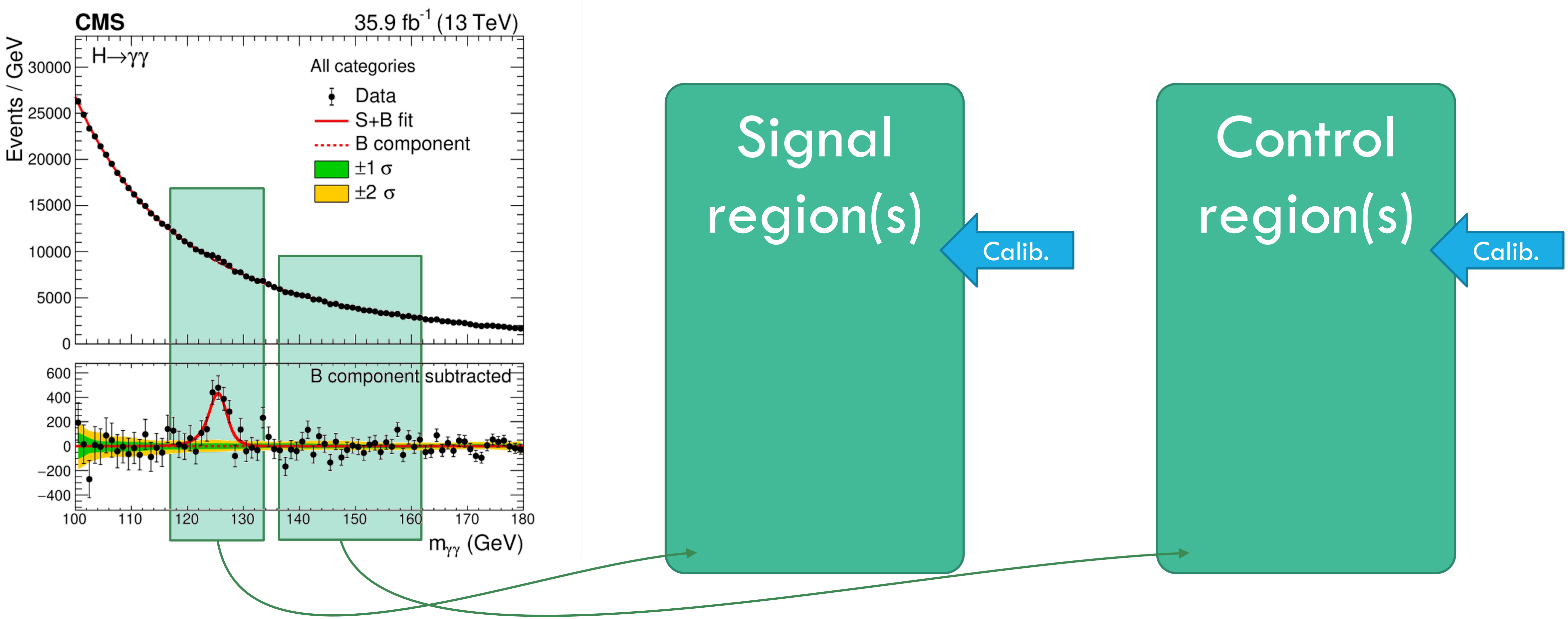
for Statistical Innovation and Discovery of Particle Physics Analysis ingredients

Area closed east of Banff to give cougars room to hunt

Dave Dormer
CTVNewsCalgary.ca Digital Producer
@dave_dormer | Contact

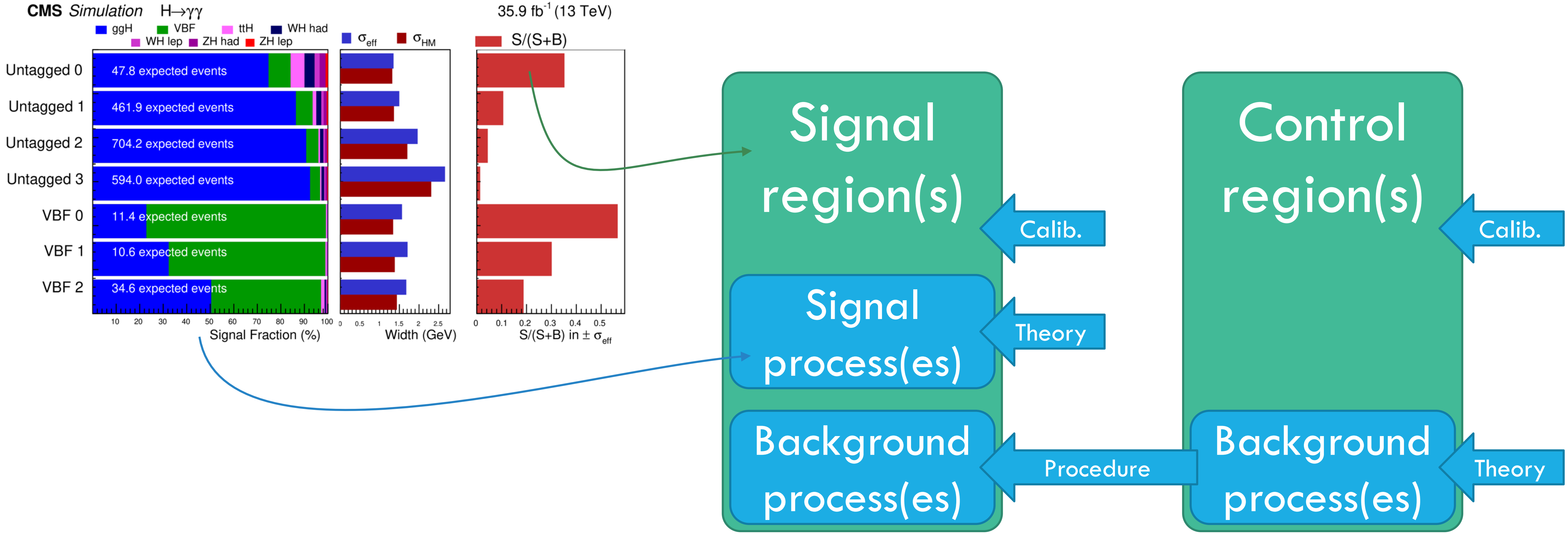
Published Monday, January 25, 2021 2:34PM MST

ERRORS — NUISANCES — COUNTS



arXiv:2002.06398

ERRORS – NUISANCES – COUNTS



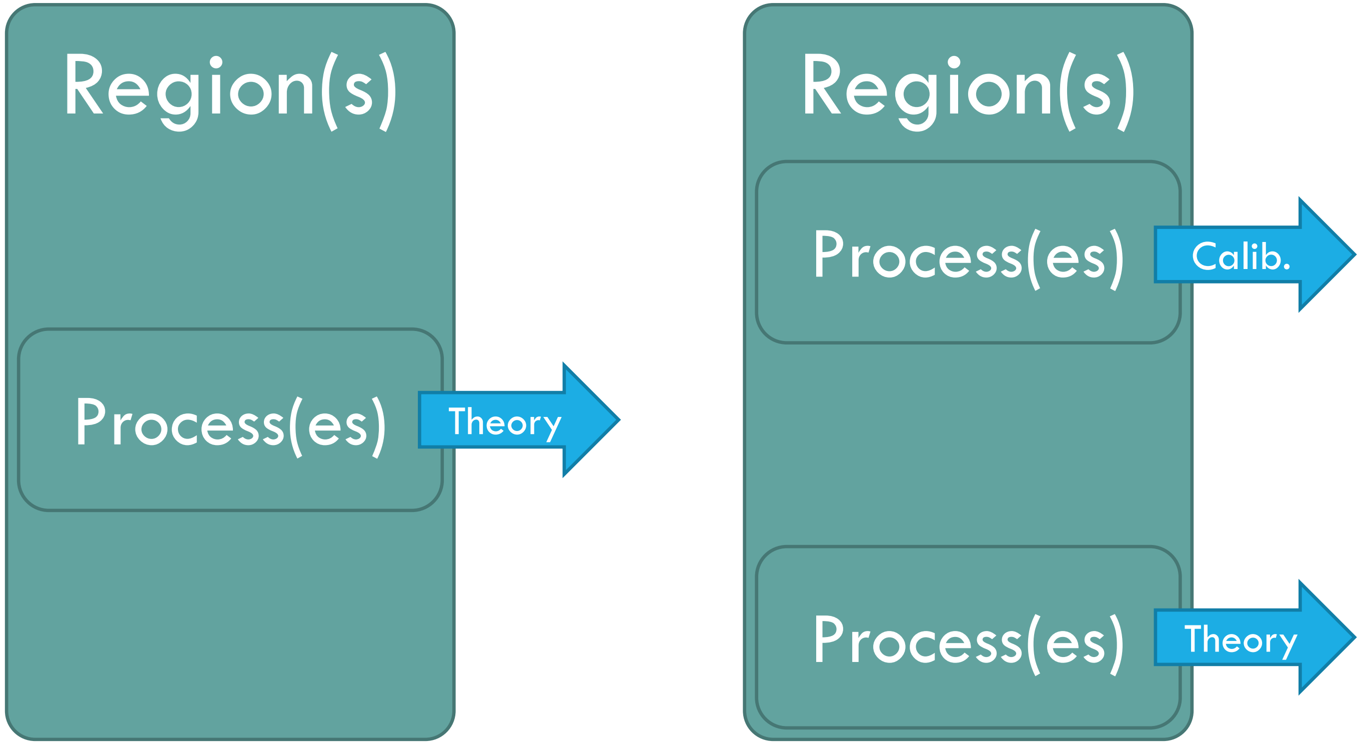
A REVERSIBLE PROCESS

Data also used to:

- Calibrate.
- Constrain theory parameters.
- Constrain non-perturbative inputs.
  - Perennial concern that parton distribution function fits may subsume BSM physics effects.

Same events ≠
(Double-counting = Double-dipping)

- Avoiding circularity always in the back of our minds.



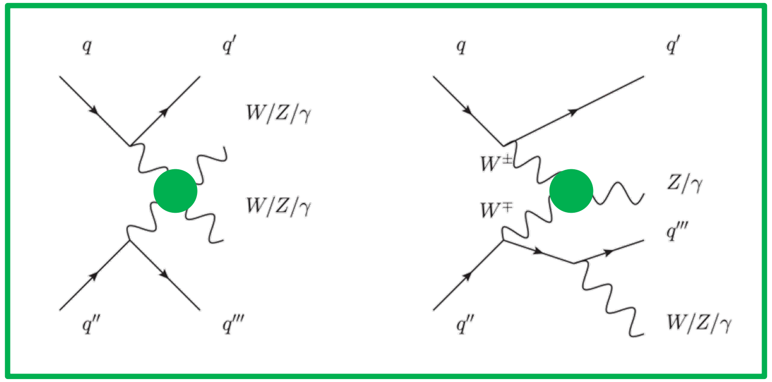
BEYOND S AND B – PROCESSES

Quantum-mechanically indistinguishable example.

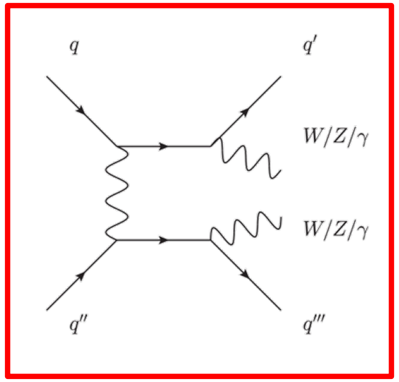
- Use interference as systematic?
- Avoid interfering phase space?
- Estimate effects on the total?
- …

Generally-speaking there are:

- Processes sensitive to the inference you want to make.
- Processes that are not.
  - Some you can estimate from MC.
  - Others may be better estimated from data.
  - Many have an impact on the power of your inference.
- Detector limitations (like noise).



More interesting



Less interesting

or $\mu^+\mu^-$, a photon, and two jets are selected. The electroweak component is measured with observed and expected significances of 4.1 standard deviations. The fiducial cross-section for electroweak production is measured to be $\sigma_{Z\gamma jj-\mathrm{EW}} = 7.8 \pm 2.0$ fb, in good agreement with the Standard Model prediction.

MadGraph5_aMC@NLO 2.3.3 MC cross-section prediction in the fiducial region ($\sigma^{\mathrm{fid., MC}}_{Z\gamma jj-\mathrm{EW}}$). Because the effect of interference between the $Z\gamma jj$–QCD and the $Z\gamma jj$–EW processes is not accounted for in the $Z\gamma jj$–QCD contribution, the observed cross-section $\sigma^{\mathrm{fid.}}_{Z\gamma jj-\mathrm{EW}}$ formally corresponds to electroweak production plus the interference effects.

ALTERNATIVES AND MORPHING

Some alternatives are **physical deformations** with meaning.

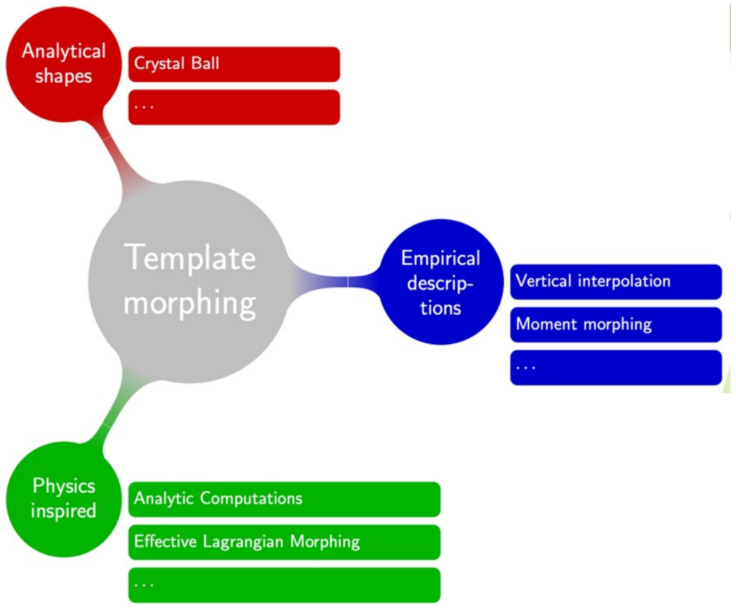
- “Average”/morphing makes sense.

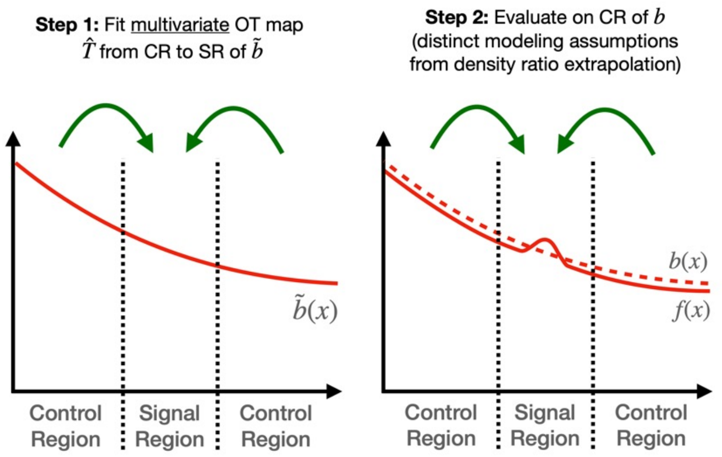
Some alternatives are really **just alternatives**.

- And if they end up mattering we’ll likely throw one out as unphysical. (Cousins)

Perturbative theory uncertainties are a whole different beast altogether.

- Limited but non-zero knowledge on the next term.





ALTERNATIVES AND MORPHING

Some alternatives are physical deformations with meaning.

- "Average"/morphing makes sense.

Some alternatives are really just alternatives.

- And if they end up mattering we'll likely throw one out as unphysical. (Cousins)

Perturbative theory uncertainties are a whole different beast altogether.

- Limited but non-zero knowledge on the next term.

Parametrize and estimate the actual source of the uncertainty: $f''(0)$

$$f(x) = f(0) + f'(0)\,x \underbrace{+ f''(0)\,\frac{x^2}{2} + \mathcal{O}(x^3)}_{\text{source of the theory uncertainty}}$$

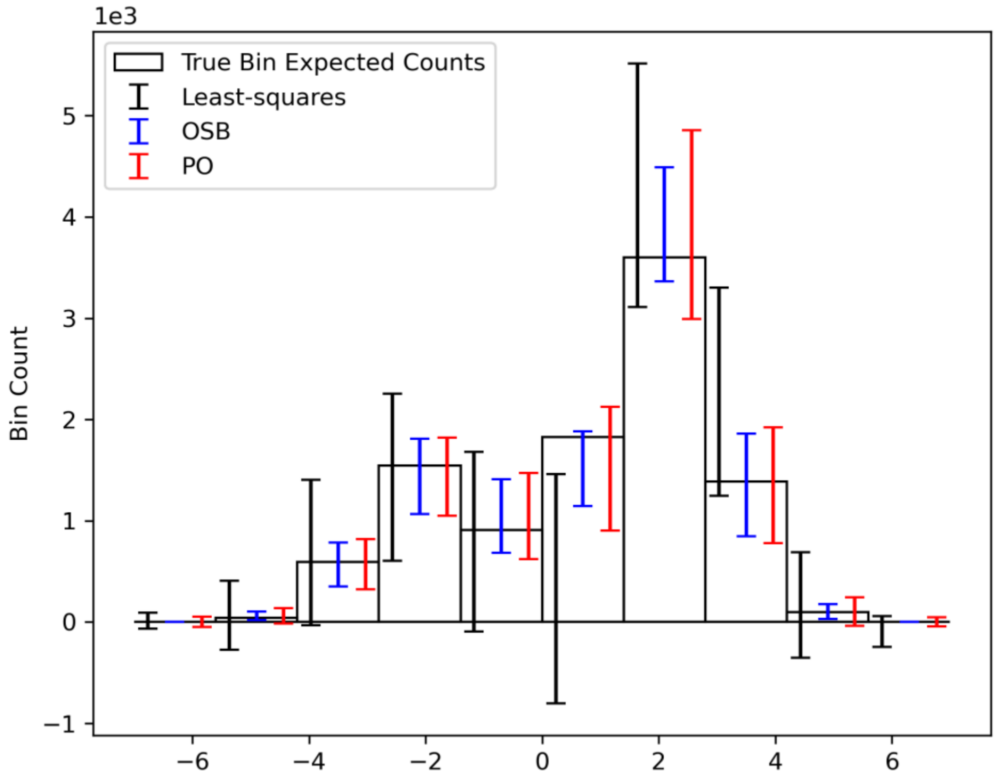
- We typically know a lot about the general structure of $f''(0)$ even without explicitly calculating it

PROGRESS IN UNFOLDING

Important physics tool for theory-experiment communication.

- Avoids theorists having to turn their calculations into full-fledged simulations.

Exciting progress with many open questions for future work.



ASTRONOMERS AND CALIBRATORS

All-in-one in HEPP but not universal.

- Also makes HEPP papers have very long, uninformative, author lists.

Cases in LHC where “interpreters” are “calibrators”.

Cases where “interpretation” is blunted to not step beyond “calibration” stated ability.

My rule of thumb: if an analysis constrains a calibration-provided nuisance parameter, stop and think.

- And then possibly take action.

THE MAGIC OF DISCRETE PROFILING

Perhaps there is **hope to understand discrete profiling** in the model selection context.

- How do you feel about model averaging being the (weighted) average of estimates across different models?

Unavoidable comparison with spurious signal; **both are prescriptions using statistical uncertainty under the signal as the gauge**

- Discrete profiling functions are chosen to have bias smaller than O(10%) stat. unc.
- Spurious signal is chosen on similar basis and added to signal model.

Model	MLE of θ	Log likelihood at MLE	AIC	Akaike Weight
Linear	0.13	-174.79	355.58	≈ 0
Quadratic	0.84	-155.85	319.70	0.29
Cubic	1.14	-154.81	319.62	0.30
Quartic	1.21	-154.02	320.05	0.24
Quintic	1.18	-153.84	321.67	0.11
Sextic	1.24	-153.83	323.65	0.040
Septic	1.25	-153.34	324.69	0.024

A WELCOME SYSTEMATISATION

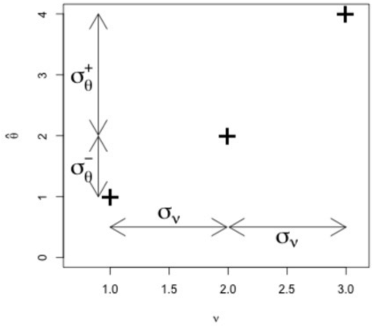
OPAT vs APAST

Combination of measurements vs combined measurement.

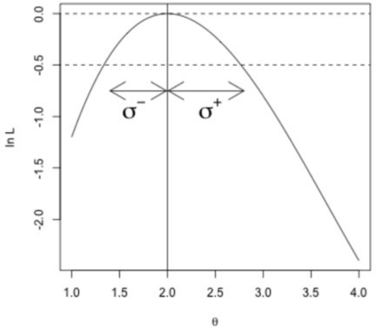
Discussion on **simplified likelihoods**:

- Taylor expansion seems to be founded.
- For PDFs a whole different story: cumulants, saddle point approximation, etc.

"Systematic" OPAT systematics evaluation

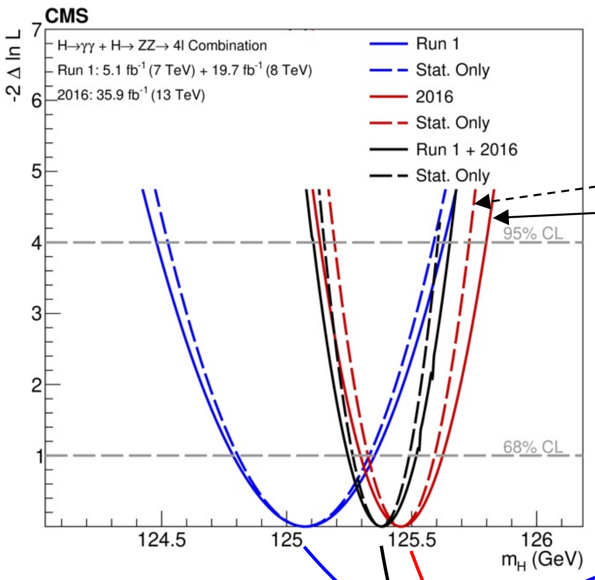


"Statistical" From ML estimation



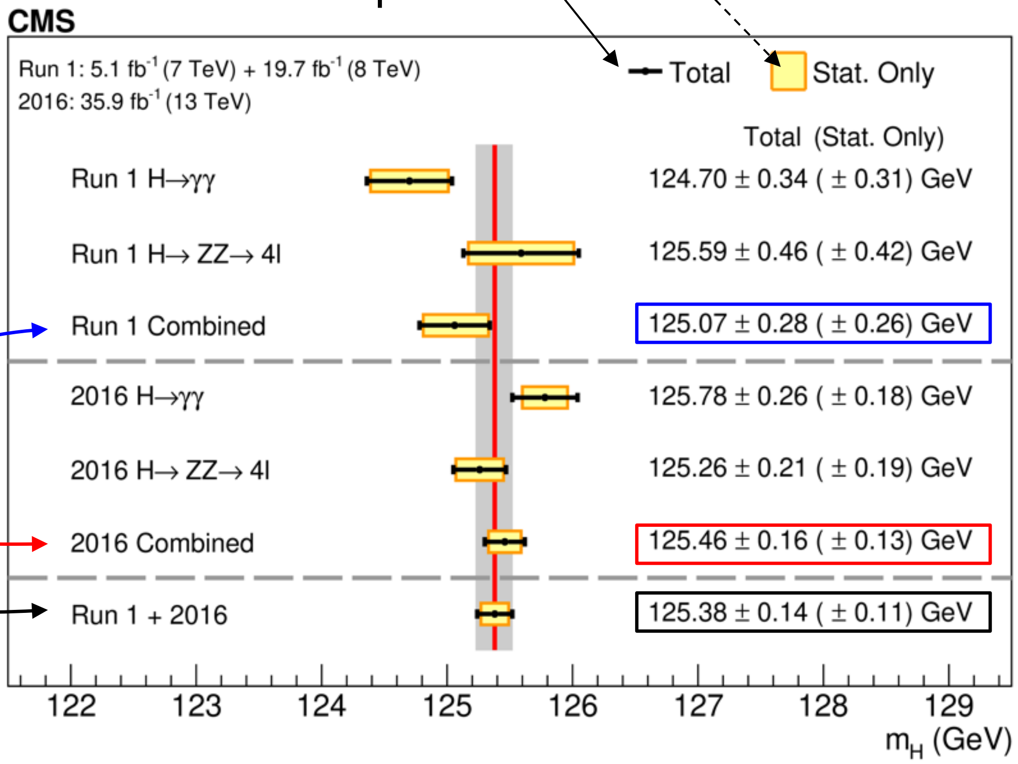
From PDFs	From Likelihoods
Error is variance of result	Error is 68% central CL
You are probably Combining Errors, in quadrature + skew	You are probably Combining Results
Goodness of fit is irrelevant	Compatibility vital & straightforward
You are probably not combining results (but you can if you work at it)	You are probably not combining errors (you can if you work at it, but not in quadrature)
"Systematic" Asymmetric Error formulæ	"Statistical" Asymmetric Error formulæ

APAT – ALL PARAMETERS AT A TIME



All fixed to MLE

All profiled



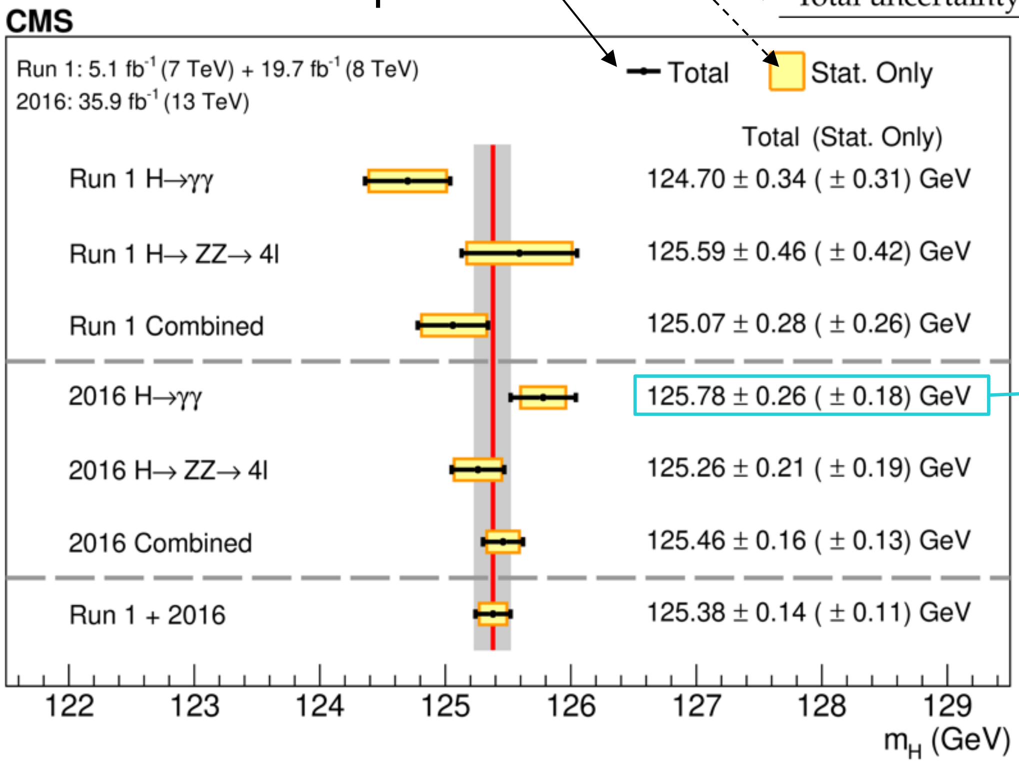
APAT – ALL PARAMETERS AT A TIME

Subsets profiled

All fixed to MLE

All profiled

Source	Contribution (GeV)
Electron energy scale and resolution corrections	0.10
Residual p_T dependence of the photon energy scale	0.11
Modelling of the material budget	0.03
Nonuniformity of the light collection	0.11
Total systematic uncertainty = Total unc. $\ominus$ Stat. unc.	0.18
Statistical uncertainty	0.18
Total uncertainty	0.26



BETTER ASYMPTOTICS

Sine qua non for ”errors on errors” that can benefit all.

Correction can also be used as coverage diagnostic tool.

I wonder what happens in asymmetric cases…

$$r_\mu^* = r_\mu + \frac{1}{r_\mu}\log\frac{q_\mu}{r_\mu} \quad = \frac{r_\mu - \mathrm{E}[r_\mu]}{\mathrm{V}[r_\mu]^{1/2}} + \mathcal{O}(n^{-3/2})$$

$$r_\mu^* \sim \mathcal{N}(0,1) + \mathcal{O}(n^{-3/2})$$

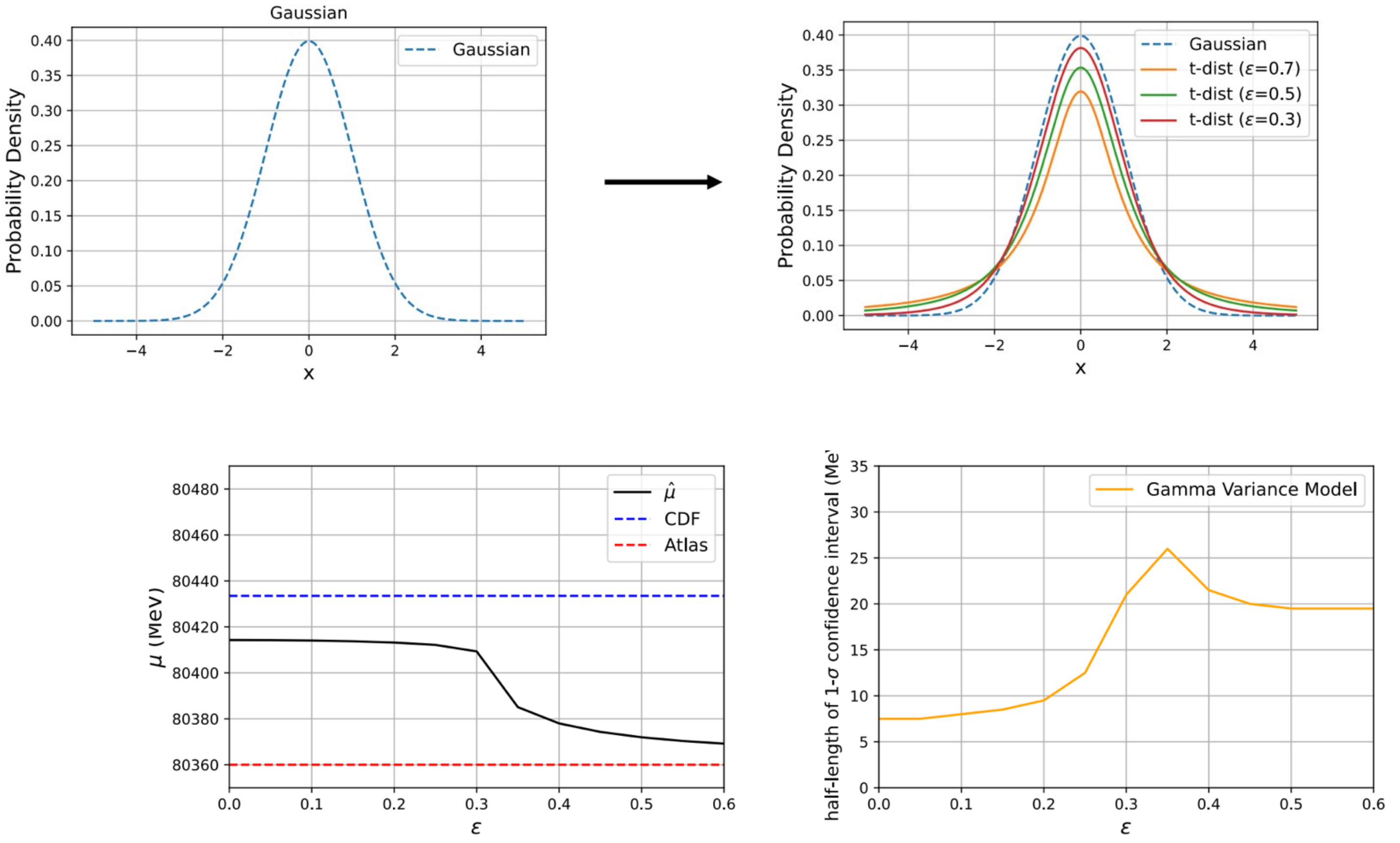
Added after the presentation.

HOW CERTAIN IS THAT UNCERTAINTY?

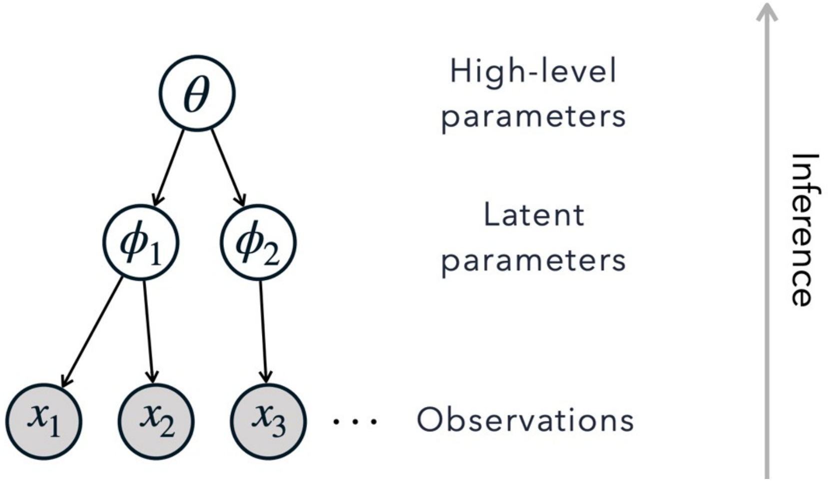
"Unleash the tails !"

Discussion focused on applying these foremost to theory inputs.

- For exp. uncs. I wonder what the evaluation experiment_k by physicist_i would be.
  - Especially when k = i.
- Lots of interesting ideas to pursue to understand how it deals with outliers.
- I know at least one theorist seriously studying the method.



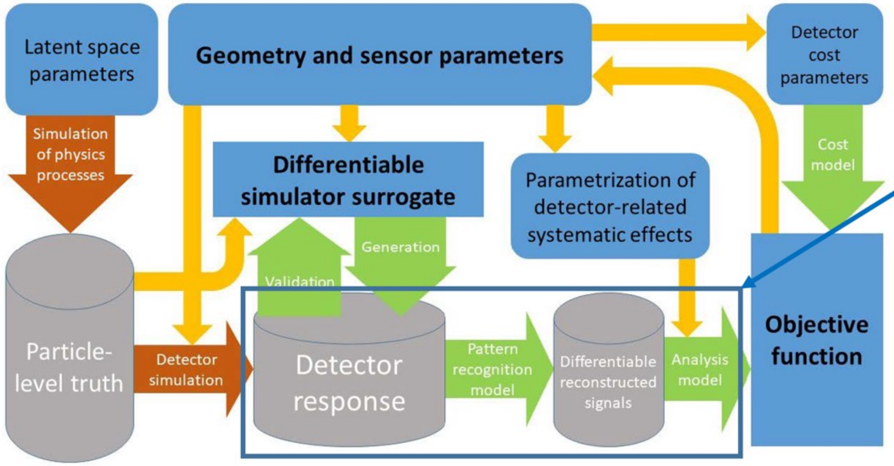
HIERARCHIES TO DIVIDE AND CONQUER



Specifying intermediate "quantities of interest" or "observables".

Not new: we calibrate energies of individual hits and reconstruct momenta of individual tracks.

Not a conclusion, just a feeling; a theme.



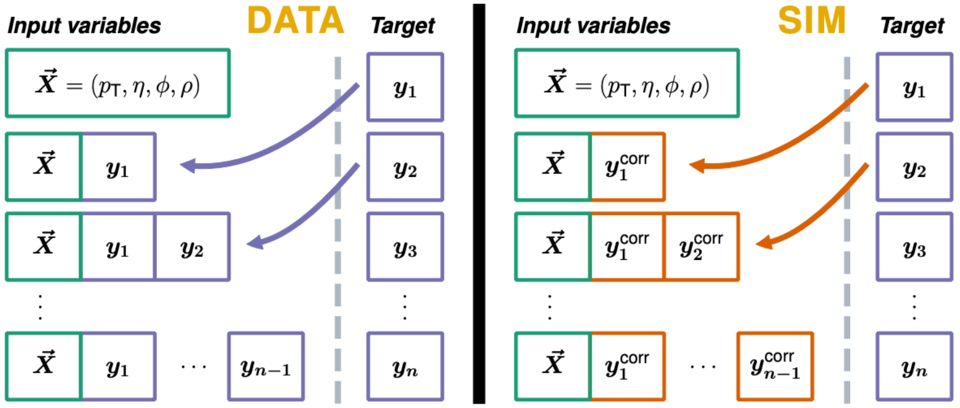
The model must include a **model** of the absolute state-of-the-art (or even extrapolated future performance!) of reconstruction and inference to avoid any **misalignment**

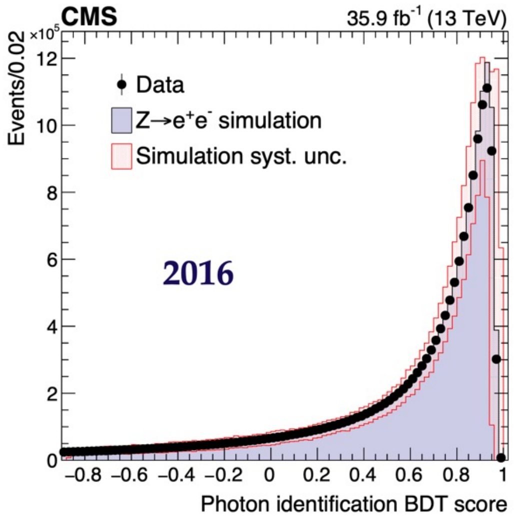
"LAST MILE" CORRECTIONS

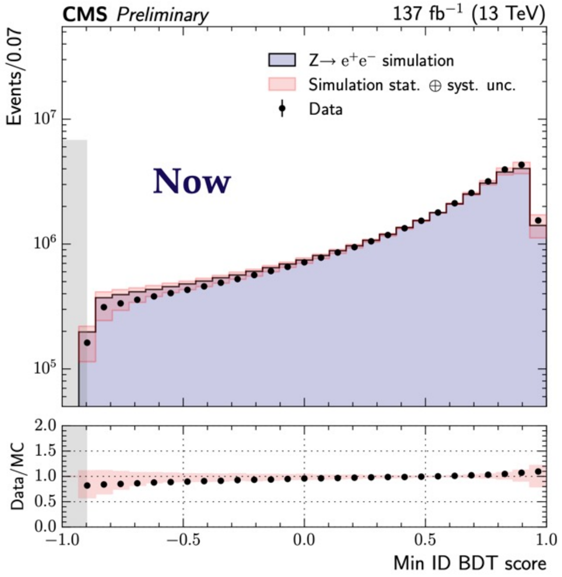
Simulation imperfections can have substantial impact on inference.

Example evolution with time:

- "Multiply and smear".
- 1-D quantile regression.
- Chained quantile regression.



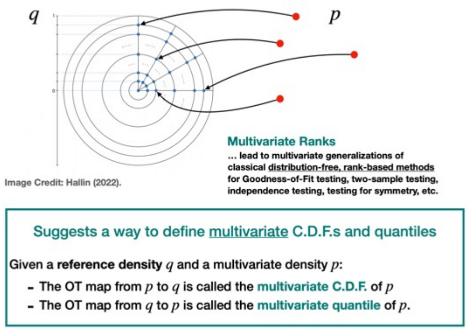




Is this the best that can be done?

I heard this week about:

- Multi-dim. quantile regression.
- Multi-dim. CDF.
- Optimal Transport maps.



THE BULK AND THE TAILS

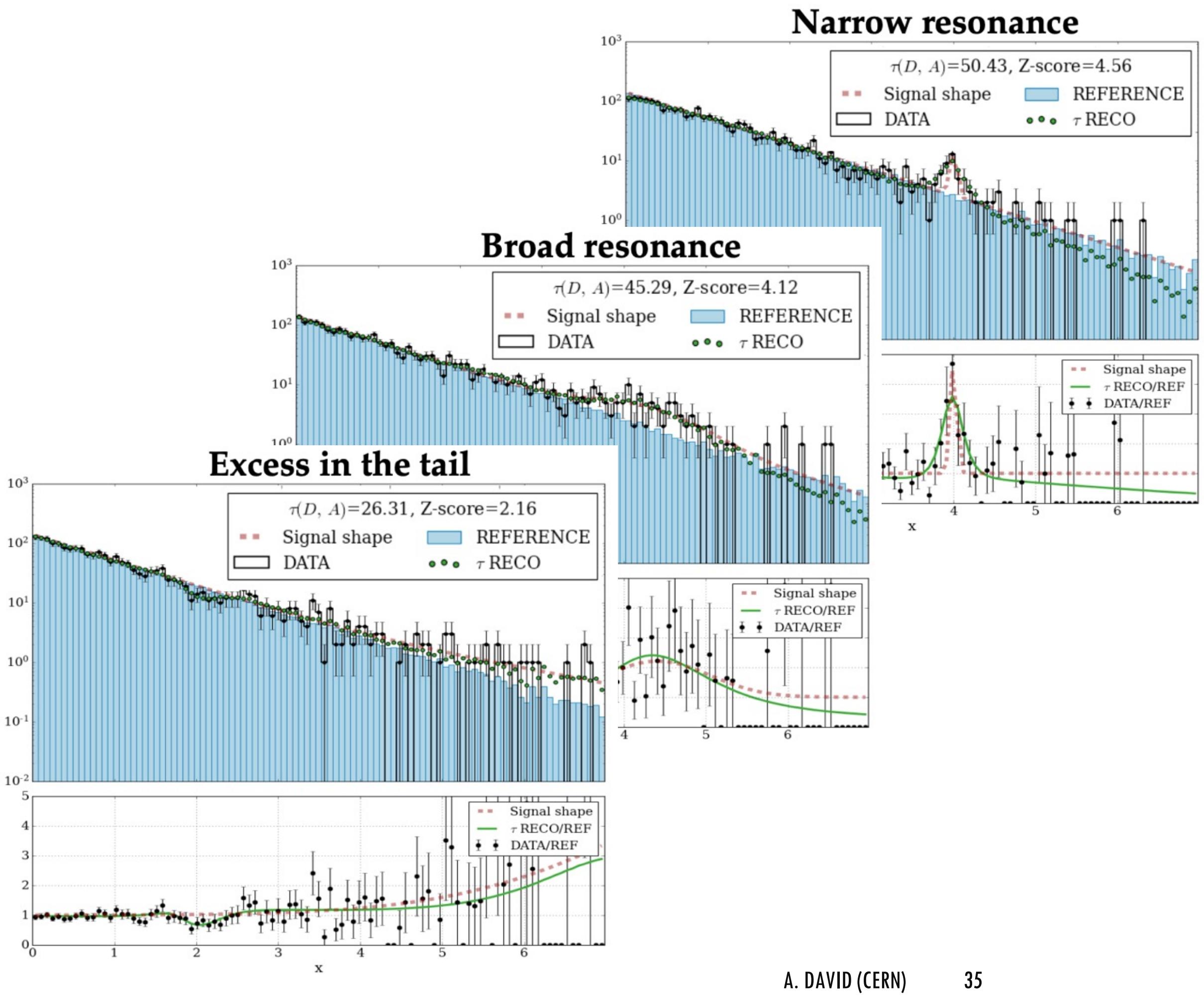
BSM physics unlikely to be the obvious stuff already looked for in the last 40 years.

- Must be within reach and be very subtle (bulk), or
- Out of reach and very energetic (tails).

Requiring same support as the SM simulation **does not cover second case above**.

- I.e. events beyond SM sim. support that could still be SM.
- Connected also to amount of SM sim. that can be afforded.

Can outlier estimation come to the rescue?



BUMP CLIMBERS

ML FOR AI – I.E. FOR ACTUAL INTELLIGENCE*

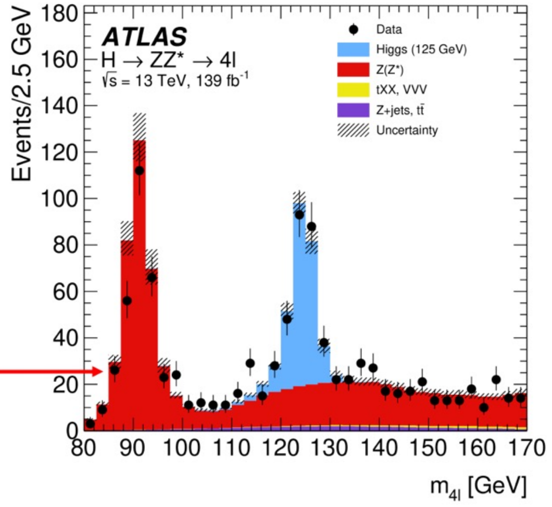
Progress: agreement that optimality and correctness are not the same.

- **90% of cases.**
  - **Can live with consequences.**
- 10% of cases.
  - Can have dire consequences.

Question of optimality:

- Did ML get best reconstruction or event selection?
- Effects definition of discriminating variables, but doesn't affect compatibility with data

Things that affect $p(\cdot|\,\lambda(\theta))$



Questions of correctness:

- Did ML learn an accurate fast simulation?
- Did ML learn a good background estimate?
- Effects statistical model & compatibility with data!

EPJC 80 (2020) 942

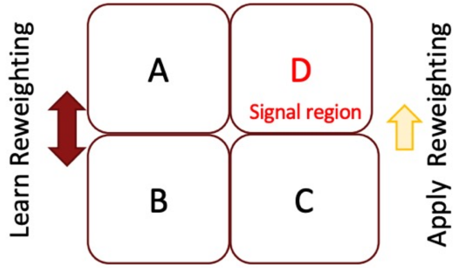
ML FOR AI – I.E. FOR ACTUAL INTELLIGENCE

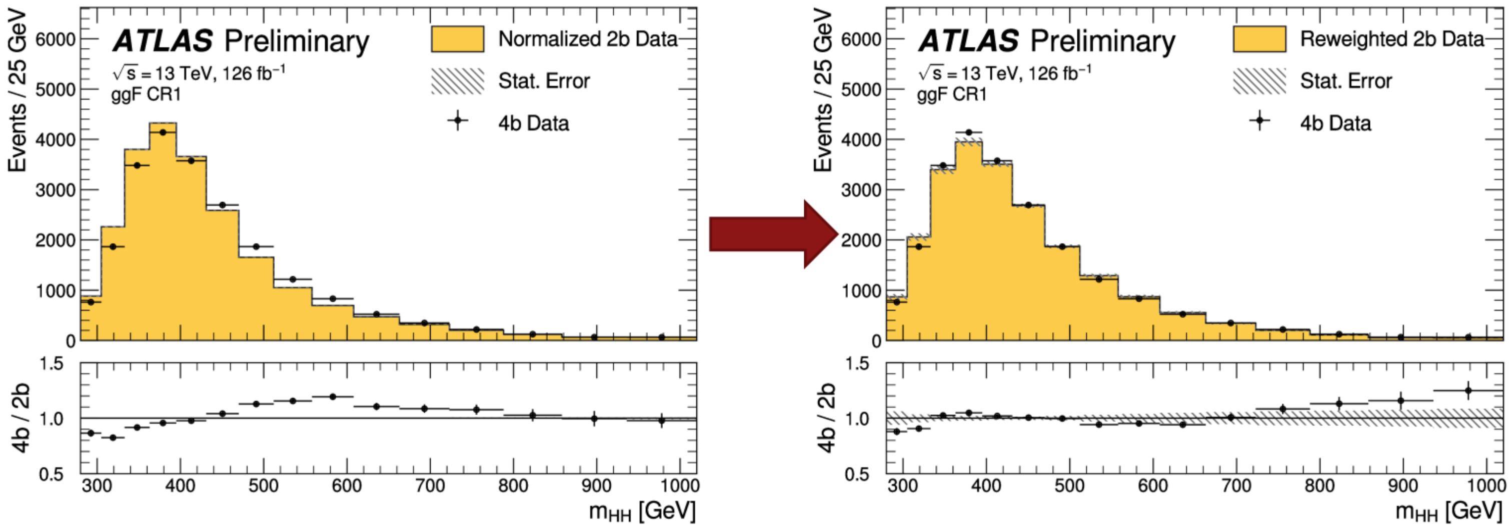
Progress: agreement that optimality and correctness are not the same.

- 90% of cases.
  - Can live with consequences.
- **10% of cases.**
  - **Can have dire consequences.**

ABCD excellent playground to test and learn.

- Also on density learning vs OT mapping.



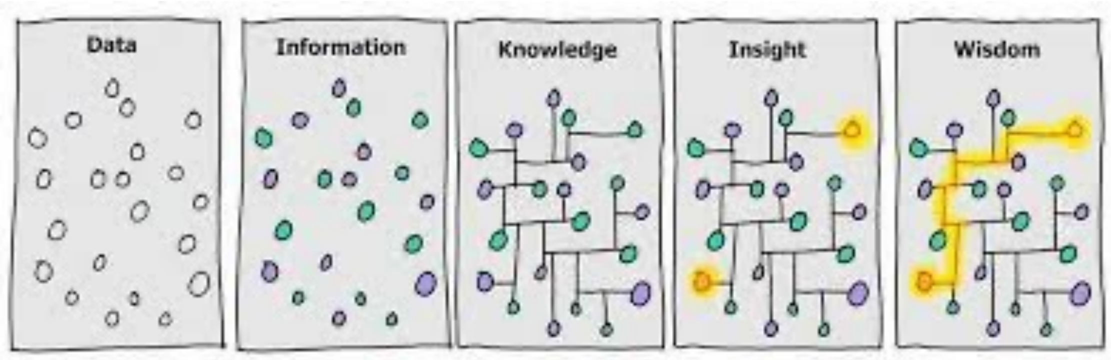


ML FOR AI – I.E. FOR ACTUAL INTELLIGENCE

Large potential and broad applicability

- Detector operation.
- Construct observables.
- Detector designs.
- Model-independent methods vs SM sim. statistics.
- Skirt systematically-affected phase spaces.
- …

My take: algorithms can more easily explore outside the box **iff** we manage to write loss functions that can do that.
Also, ML is not yet wise.



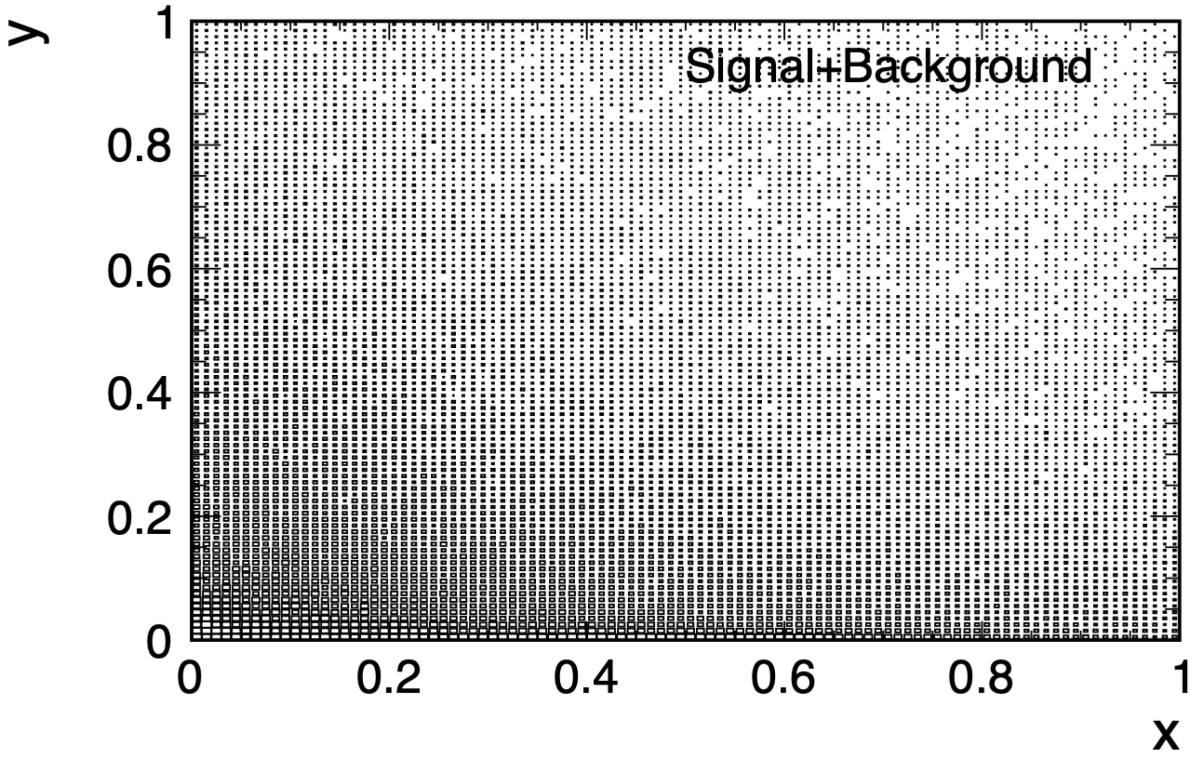
THE IMPORTANCE OF THE MWE (FOR ACTUAL INTELLIGENCE)

Physicists using Open Data took one year to reproduce an analysis.

Crucial to have **minimum working examples and/or challenges** that flesh out the essential of the problems being faced.

Fundamental to continued flow of knowledge.

- I dare suggest a "hackathon" with talks before over Zoom and work together on-premise for a week.



LHC'S BUT ONE CORNER OF PARTICLE PHYSICS

Specific issues that deserve just as much attention from statisticians.

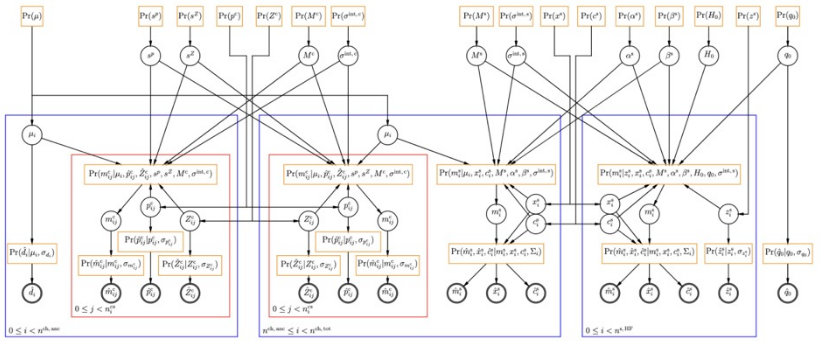
- Fertile (safe?, welcoming?) ground for Bayesian methods.

LHCb and Belle II: measurement style

For measuring $\mathscr{B}$ of rare B-decays **LHCb uses mostly relative** and **Belle II absolute approach**:

- $\mathscr{B}(B^+ \to \mu^+\mu^-\mu^+\nu) = \mathscr{B}(B^+ \to J/\psi(\to \mu^+\mu^-)K^+) \times \dfrac{\epsilon(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)}{\epsilon(B^+ \to \mu^+\mu^-\mu^+\nu)} \times \dfrac{N(B^+ \to \mu^+\mu^-\mu^+\nu)}{N(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)}$
- $\mathscr{B}(B^+ \to K^+\nu\bar{\nu}) = \dfrac{N(B^+ \to K^+\nu\bar{\nu})}{\epsilon(B^+ \to K^+\nu\bar{\nu})}$

$$N^i_{pred} = \int_{E_{\min}}^{E_{\max}} \overbrace{P(\nu_\alpha \to \nu_\beta)}^{\text{signal}} \times \underbrace{\overbrace{\Phi(E_\nu)}^{\text{beam}} \times \overbrace{\sigma(E_\nu),\, \vec{x})}^{\text{cross section}} \times \overbrace{\epsilon(\vec{x})}^{\text{detector}}}_{\text{nuisance parameters}} \, dE_\nu$$



[Feeney+ MNRAS 2018]

SEANPIPPDA WAS VERY EXCITING

UNTIL NEXT TIME

Statistics

- A continued treasure trove of techniques that shape our practices.
  - Ever more pervasive and more advanced.

Particle physics

- Particularly good conditions for application of statistical methods.
  - Some may call our null not challenging enough.
  - Some may find Poisson counts too boring.
- Variety of problems in multiple aspects of statistical practice.
  - Likely applicable to other (physical) sciences.

The workshop does not end today.

- Can keep Slack alive or move it to Mattermost?
- Can we consider a hackathon-like format in the future?

"THERE IS NO SPOON"