

# Interpretability and reliance on simulation

Benjamin Nachman

*Lawrence Berkeley National Laboratory*

[cern.ch/bnachman](https://cern.ch/bnachman)

[bpnachman@lbl.gov](mailto:bpnachman@lbl.gov)



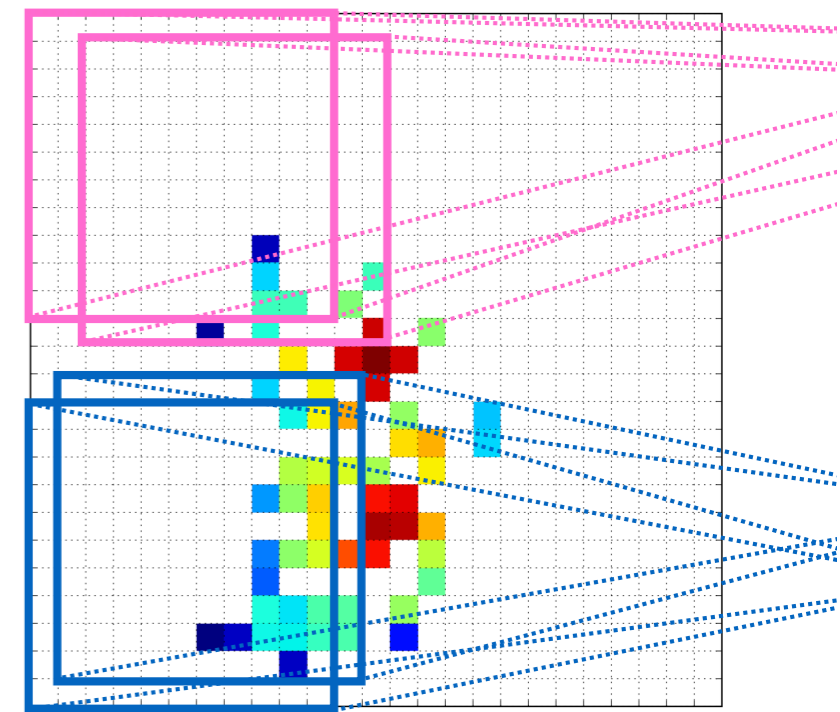
@bpnachman



bnachman



**BERKELEY  
EXPERIMENTAL  
PARTICLE  
PHYSICS**



**LHCP 2020**  
Virtual  
May 27, 2020

*I won't say much  
explicitly about this  
during the talk*

# Interpretability and reliance on simulation

Benjamin Nachman

*Lawrence Berkeley National Laboratory*

[cern.ch/bnachman](https://cern.ch/bnachman)

[bpnachman@lbl.gov](mailto:bpnachman@lbl.gov)



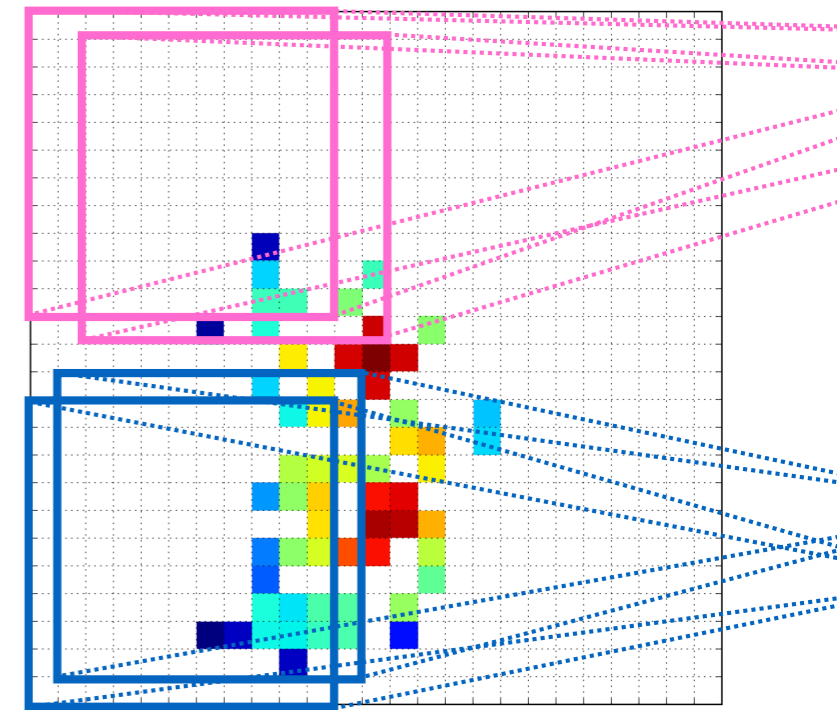
@bpnachman



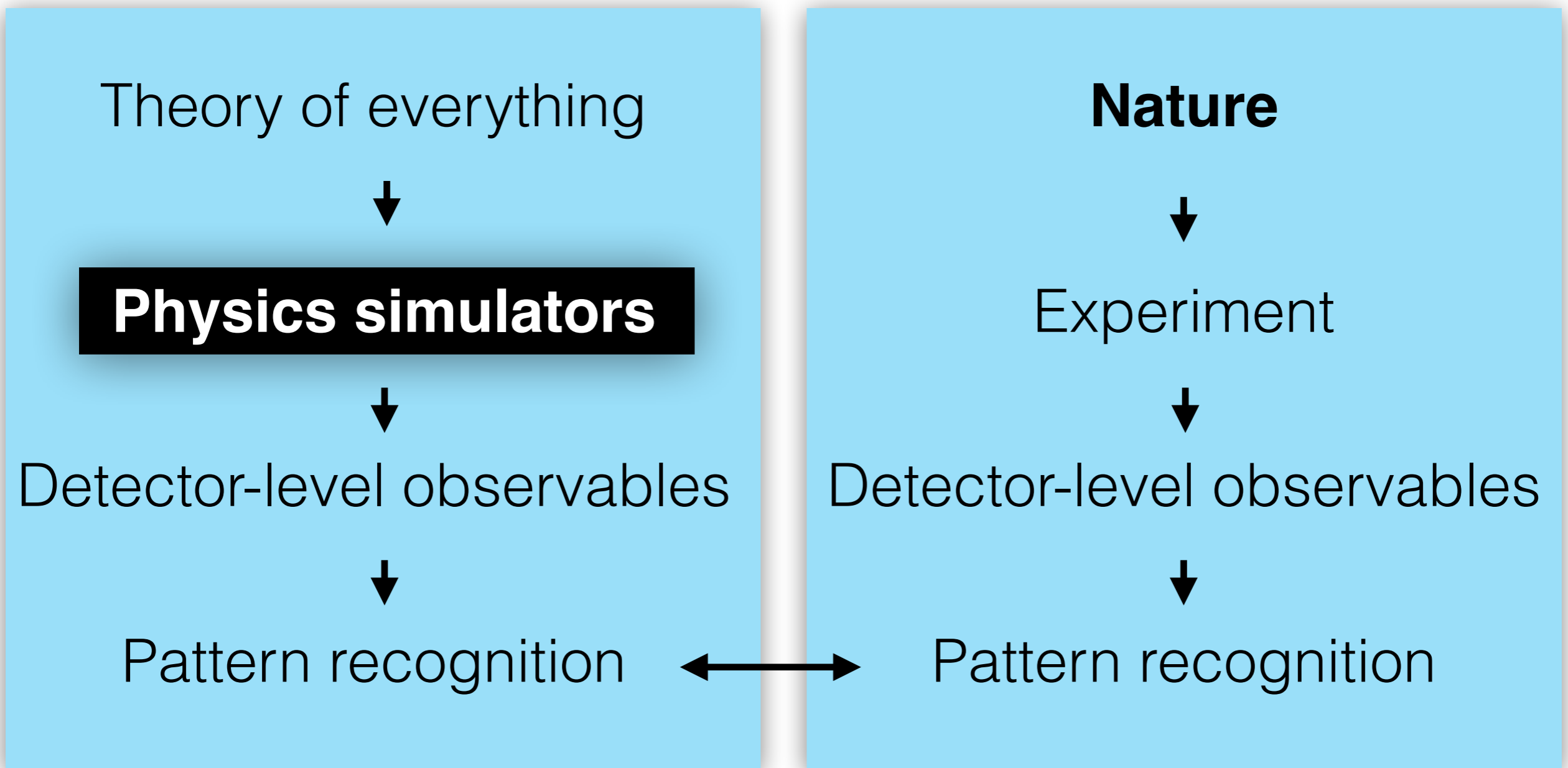
bnachman



**BERKELEY  
EXPERIMENTAL  
PARTICLE  
PHYSICS**

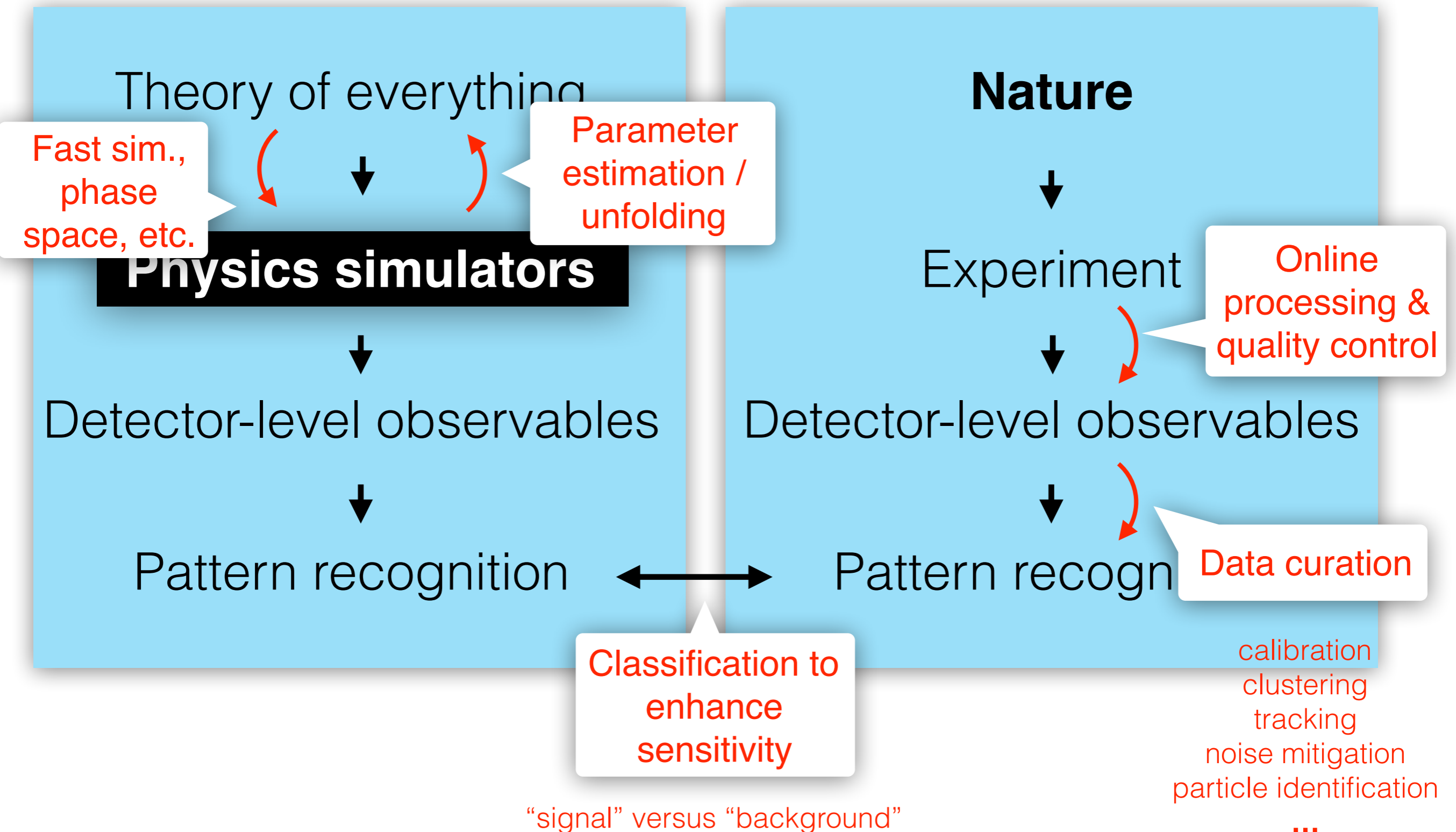


**LHCP 2020**  
Virtual  
May 27, 2020



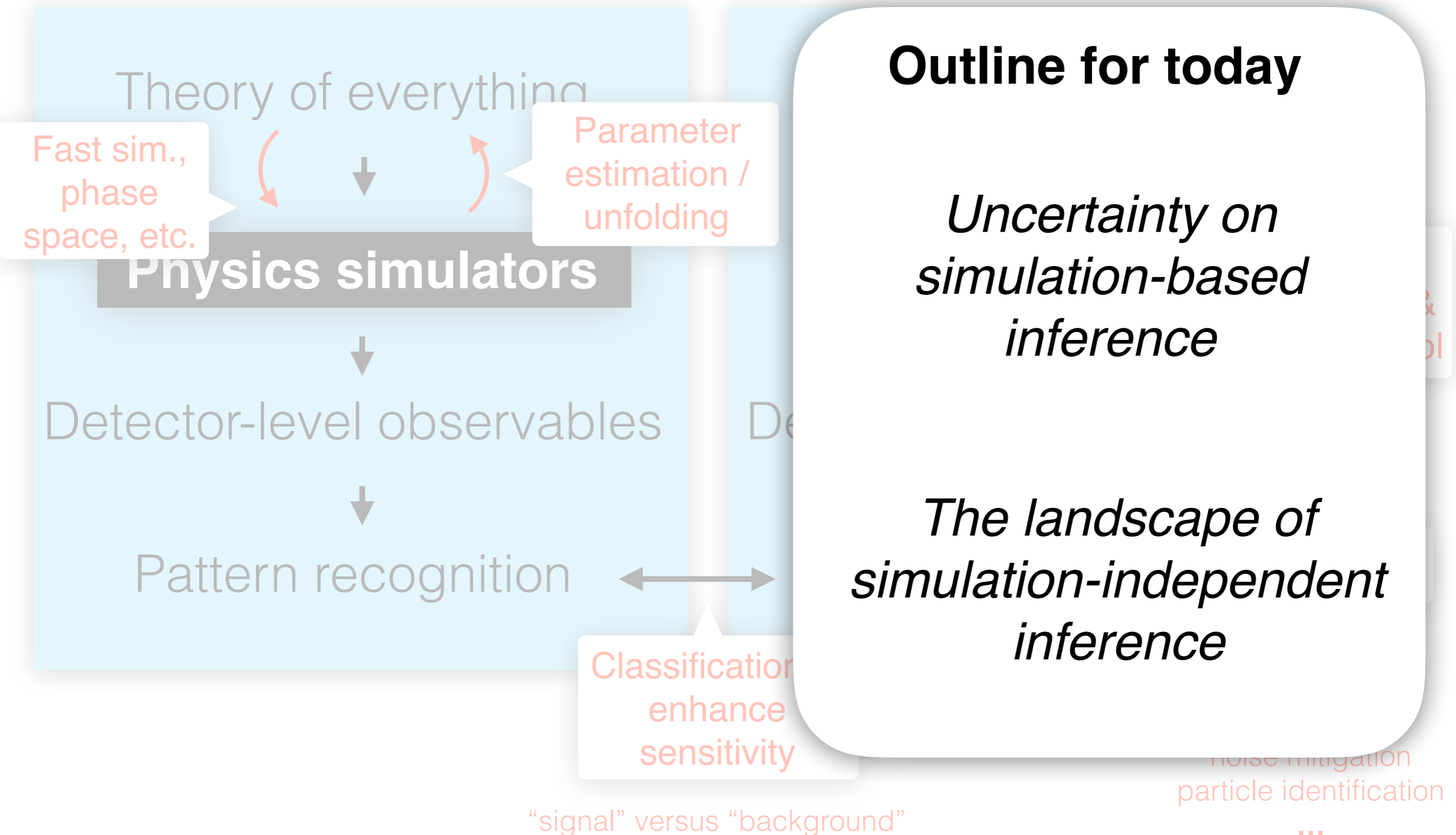
# Data analysis in HEP + Deep Learning

4



# Data analysis in HEP + Deep Learning

5



“But what are the uncertainties on the NN”?

- question asked by every reviewer



“But what are the uncertainties on the NN”?

- question asked by every reviewer

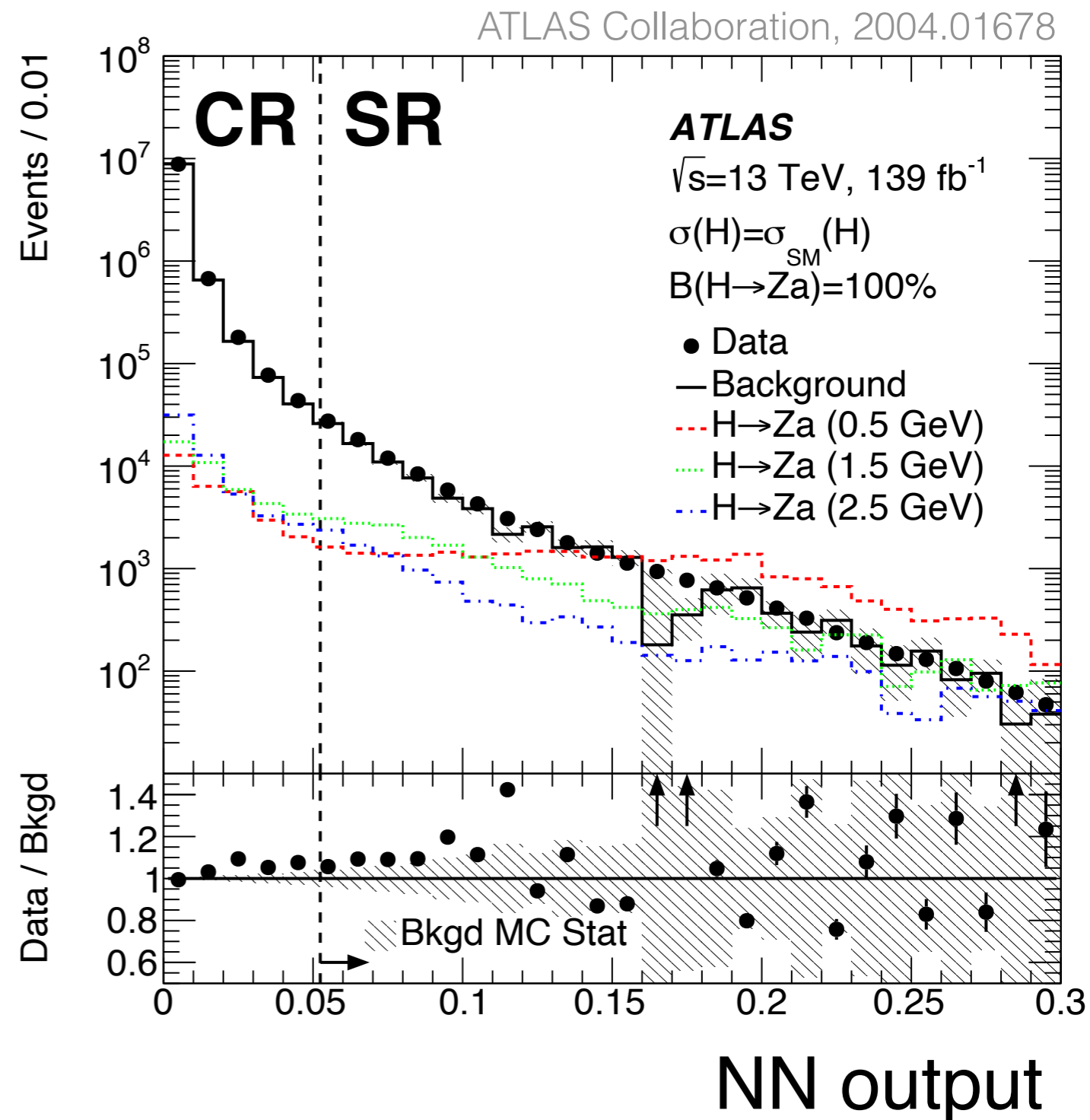
**Let's consider this question in the context of a search for new particles in collision events.**

# Setup



1. Train a classifier (in sim.) for signal vs. background.
2. Define a control region (**CR**) and a signal region (**SR**) using (1).
3. Check / modify simulation in CR.
4. Compare data and simulation in SR.

Significantly different? go to Stockholm : publish limits.



# Uncertainties for a NN-based analysis



Precision / Optimality

*Bad use of our data, time, money, etc. but **not wrong**.*

Accuracy / Bias

# Uncertainties for a NN-based analysis

10

Precision / Optimality:  $NN(x) \neq \frac{p_{\text{true}}(x|S+B)}{p_{\text{true}}(x|B)}$

↑  
Optimal by Neyman-Pearson  
(no nuisance parameters)

Accuracy / Bias

*Note that this is not  $p(x|S) / p(x|B)$ , however the two are monotonically related to each other.*

# Uncertainties for a NN-based analysis

11

Precision / Optimality:  $NN(x) \neq \frac{p_{\text{true}}(x|S+B)}{p_{\text{true}}(x|B)}$

Accuracy / Bias:  $p_{\text{prediction}}(NN) \neq p_{\text{true}}(NN)$

*The distribution of the (corrected) sim. is not correct.*

# Uncertainties for a NN-based analysis

12

Precision / Optimality:  $\text{NN}(\mathbf{x}) \neq \frac{p_{\text{true}}(\mathbf{x}|\mathbf{S}+\mathbf{B})}{p_{\text{true}}(\mathbf{x}|\mathbf{B})}$

*limited training statistics*

$p_{\text{train}}(\mathbf{x}) \neq p_{\text{true}}(\mathbf{x})$

*inaccurate training data*

$\text{NN}(\mathbf{x})|_{p_{\text{true}}=p_{\text{train}}} \neq \frac{p_{\text{true}}(\mathbf{x}|\mathbf{S}+\mathbf{B})}{p_{\text{true}}(\mathbf{x}|\mathbf{B})}$

*model/optimization flexibility*

**Statistical uncertainty**

**Systematic uncertainty**

*limited prediction statistics*

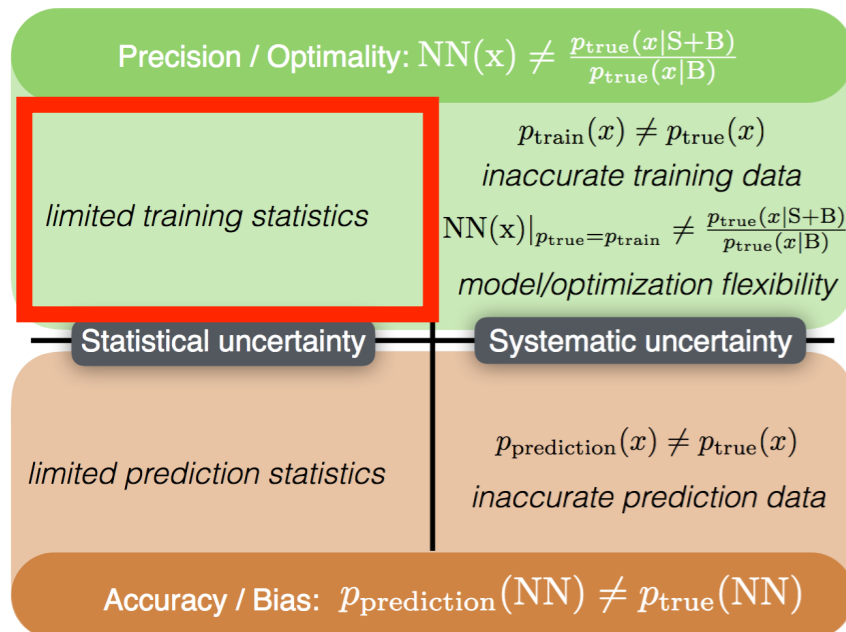
$p_{\text{prediction}}(\mathbf{x}) \neq p_{\text{true}}(\mathbf{x})$

*inaccurate prediction data*

Accuracy / Bias:  $p_{\text{prediction}}(\text{NN}) \neq p_{\text{true}}(\text{NN})$

# How to estimate precision stat. uncerts.

13



You can always accomplish this by bootstrapping: making pseudo-datasets from resampling and then retraining.

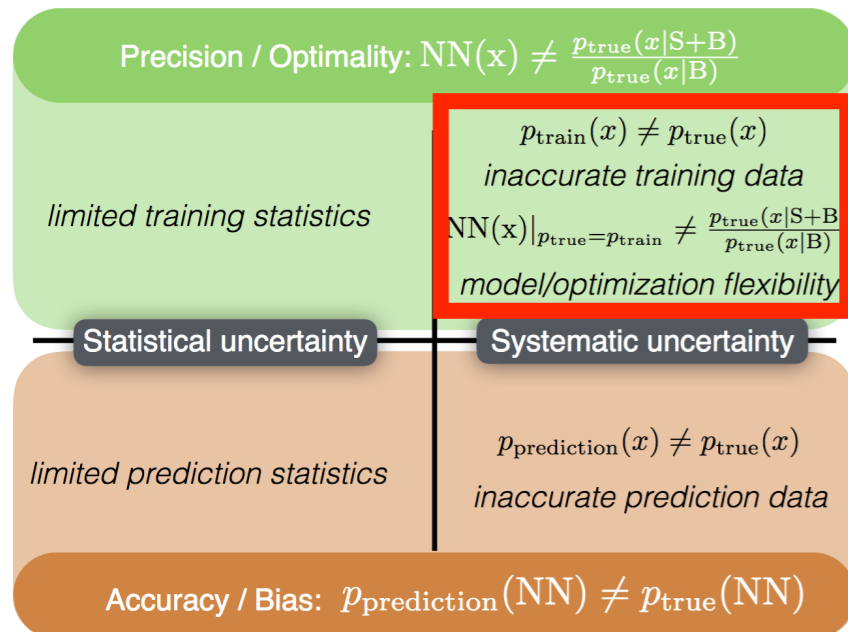
It is important to fix the NN initialization so that you are not also testing your sensitivity to that.

This can be painful because it requires retraining many NNs.

Maybe can accomplish with one Bayesian NN? See e.g. S. Bollweg, et al., SciPost Phys. 8, 006 (2020), 1904.10004 for a particle physics example.

# How to estimate precision syst. uncerts.

14



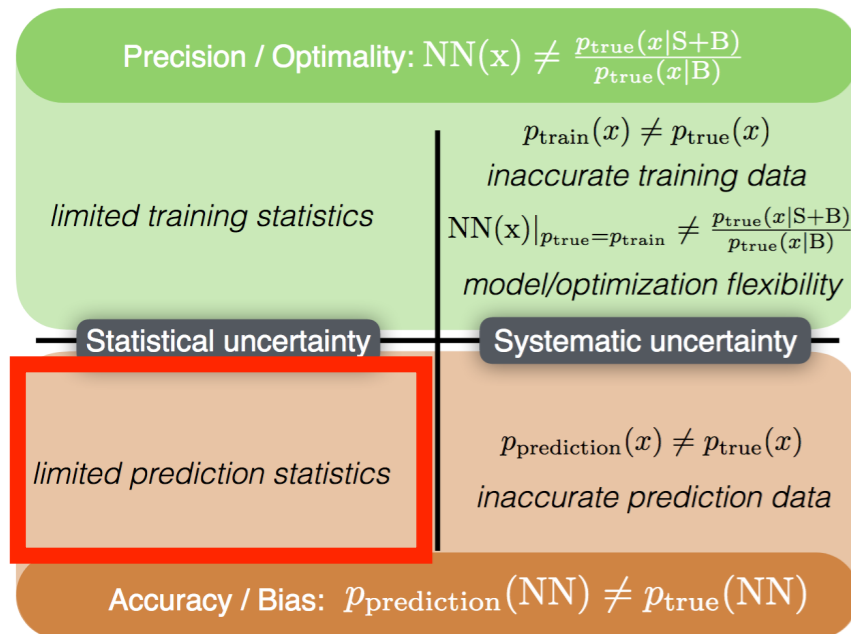
As with all systematic uncertainties,  
this is hard to quantify.

One component is due to the  
modeling of  $p(x)$  - more on this later.

Testing the flexibility of the network requires  
checking the sensitivity to the architecture  
(#layers, nodes/layer, etc.), the initialization, the  
training procedure (#epochs, learning rate, etc.)

# How to estimate bias stat. uncerts.

15

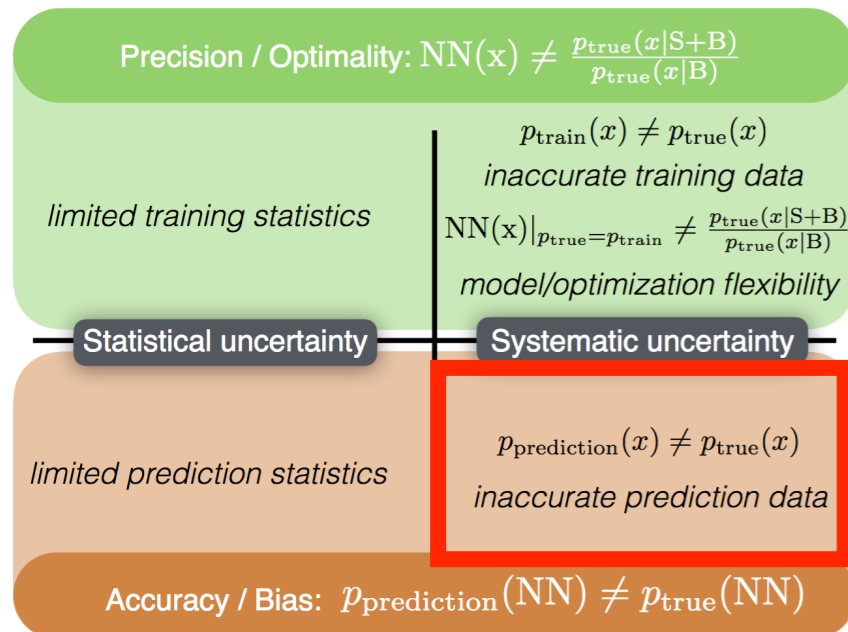


Can be estimated via bootstrapping. Less painful here because the NN's are fixed.

N.B. it may be possible to design a network that is designed to minimize uncertainty at inference. This does not work in all cases, but early studies in particle physics seem promising: S. Wunsch et al., 2003.07186, P. da Castro et al., CPC 244 (2019) 170, 1806.04743

# How to estimate bias syst. uncerts.

16



**This is the trickiest one...**

*...because we need the uncertainty on the modeling of  $x$  and  $x$  can be high-dimensional!*

In many cases, the uncertainties factorize, e.g. the uncertainty on two photon energies can be decomposed into the uncertainty on each photon.

However, in many cases, we simply do not know the full uncertainty model (= nuisance parameters and their distribution)

# High-dimensional Bias Uncertainties

17

One word of caution: current paradigm for uncertainties may be too naive for high-dimensional analysis!

(truly end-to-end)

e.g. for some uncertainties, we often compare two different models - one nuisance parameter.

How can we even see how sensitive we are to high-dimensional effects?

# High-dimensional Bias Uncertainties

18

One word of caution: current paradigm for uncertainties may be too naive for high-dimensional analysis!

(truly end-to-end)

e.g. for some uncertainties, we often compare two different models - one nuisance parameter.

How can we even see how sensitive we are to high-dimensional effects?

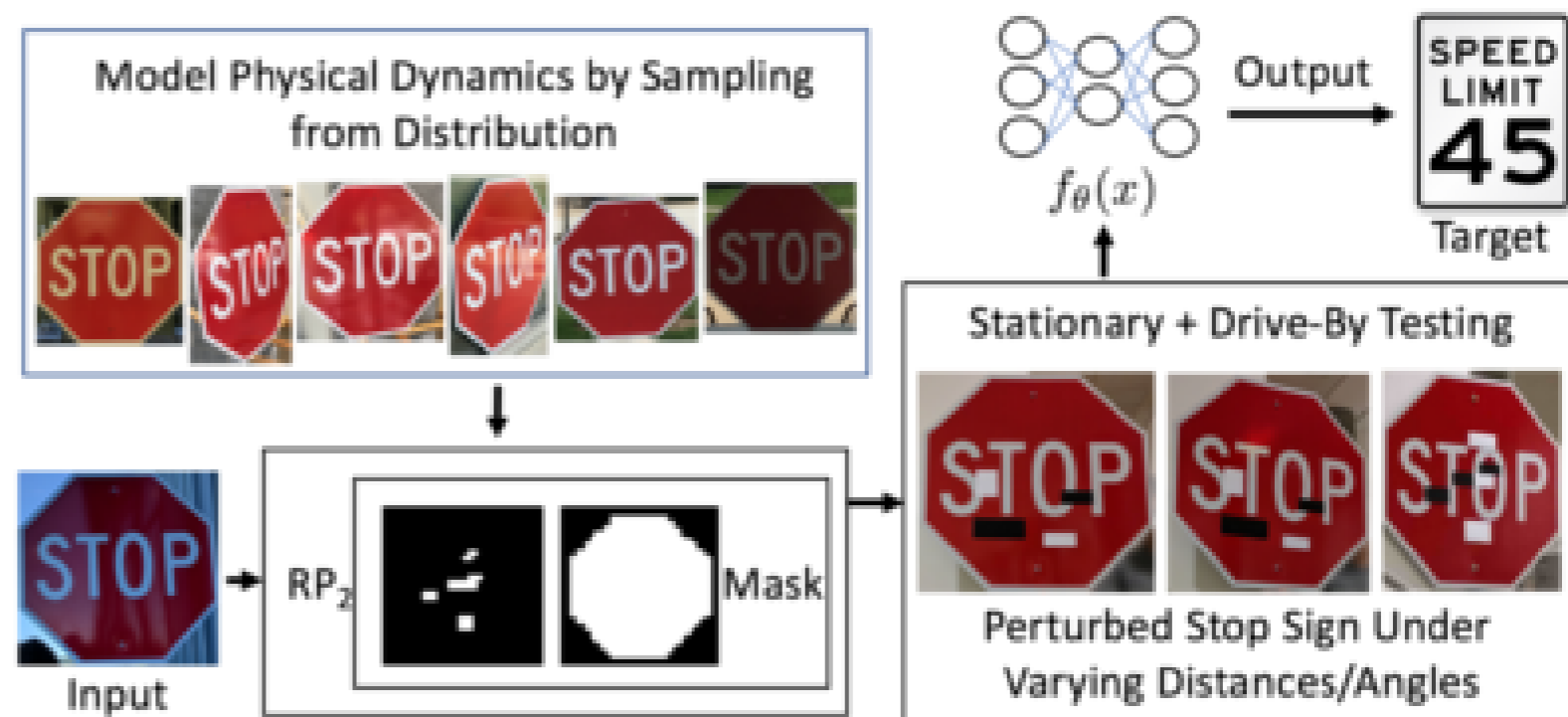
**Answer: borrow tools from AI Safety**



There is a vast literature on how easy it is to “attack” a NN.

*They want to know: how subtle can an attack be and still significantly impact the output.*

We know (hope?!) that nature is not evil, but these tools can help us probe the high-dimensional sensitivity of our NNs.



# Bounding high-dim. uncerts: strategy

20

$\mathbf{J}$  = collision event (in all of its high-dimensional glory)

$\mathbf{f}$  = fixed classifier for signal vs. background

Loss

$$\mathcal{L}_{\text{sig}} = \log(1 - f(g(\mathbf{J}))),$$

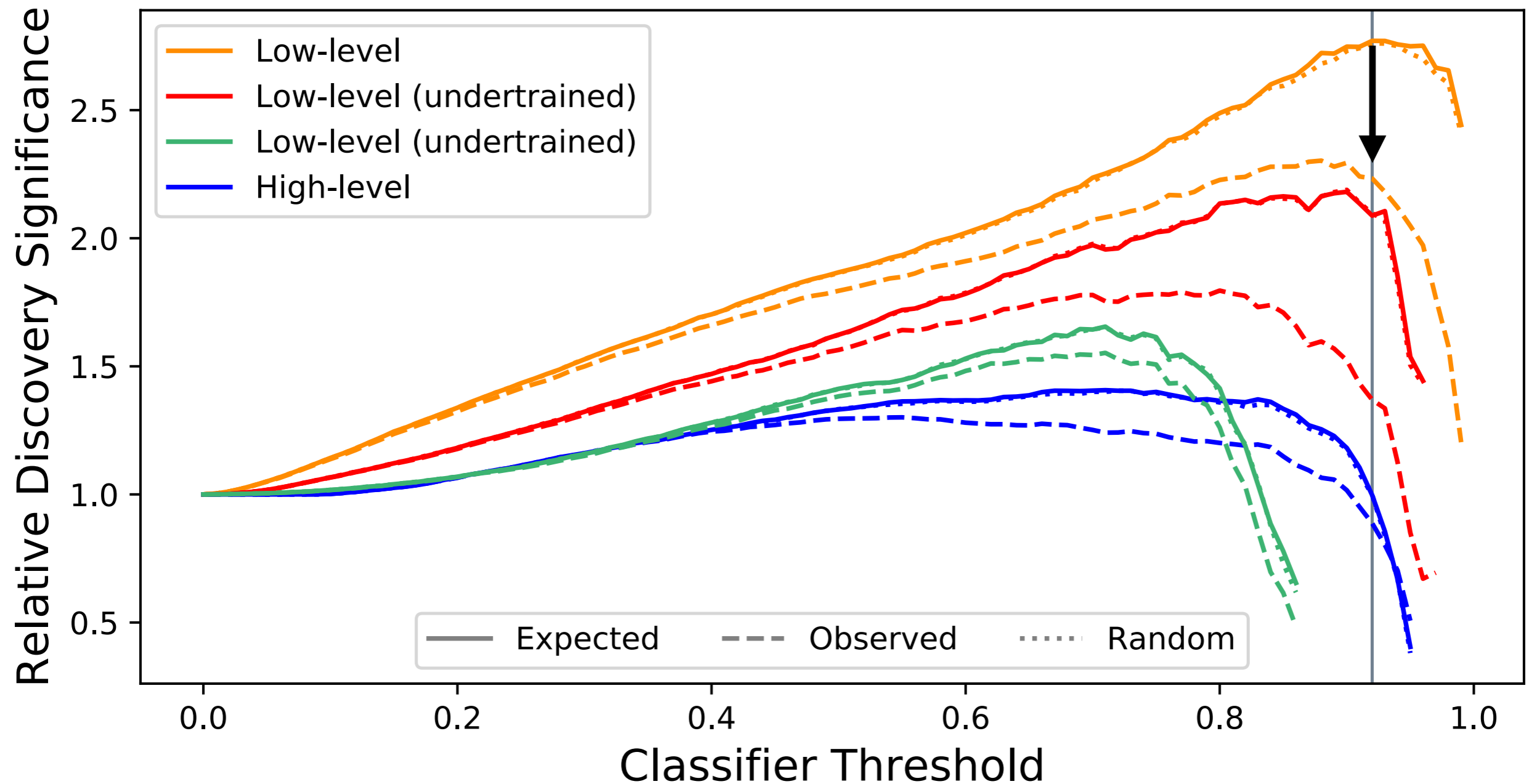
$$\begin{aligned} \mathcal{L}_{\text{bg}} = & \lambda_{\text{cls}} (f(\mathbf{J}) - f(g(\mathbf{J})))^2 \\ & + \sum_i \lambda_{\text{obs}}^{(i)} (\mathcal{O}^{(i)}(\mathbf{J}) - \mathcal{O}^{(i)}(g(\mathbf{J})))^2 \end{aligned}$$

$\mathbf{g}$  is a learned NN that maps  $\mathbf{J}$  to  $\mathbf{J} + \delta\mathbf{J}$ .

$\mathbf{O}(\mathbf{J})$  are observables that will be validated in the CR.

# High-dimensional Uncertainty

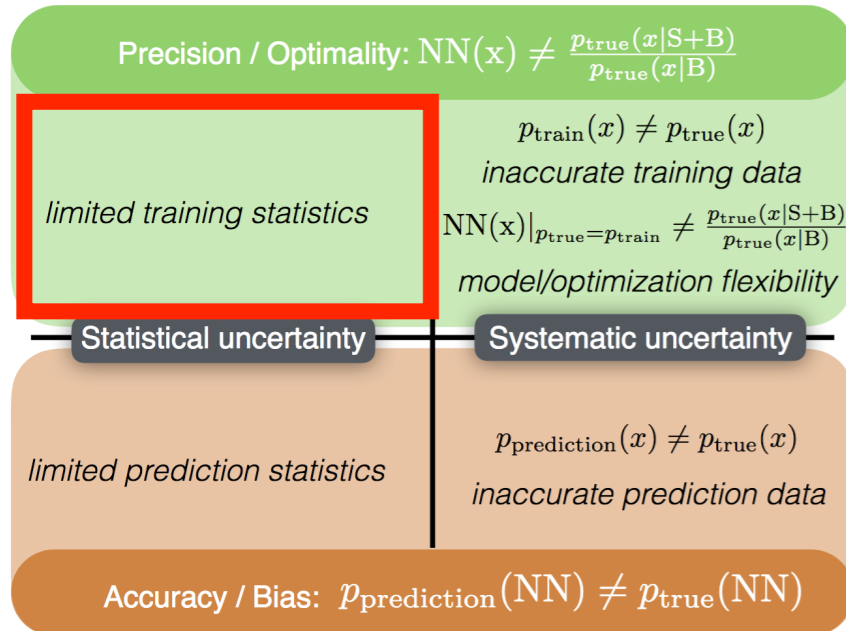
21



“worst-case uncertainty”

# How to reduce precision stat. uncerts.

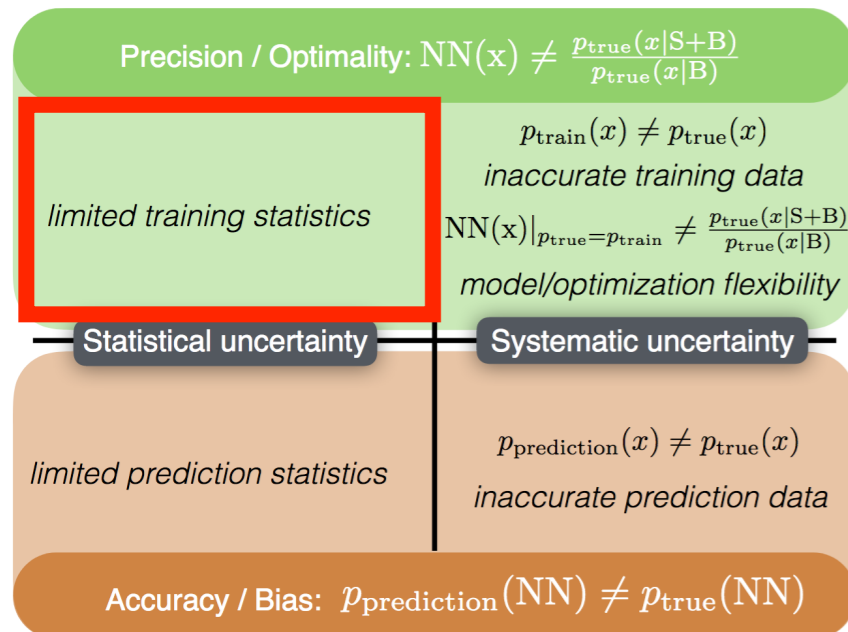
22



Train with more events!

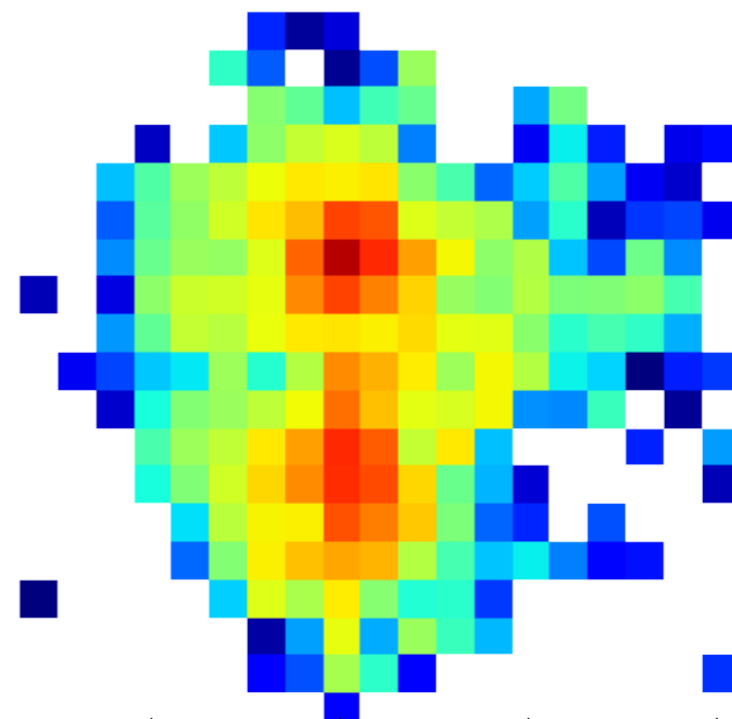
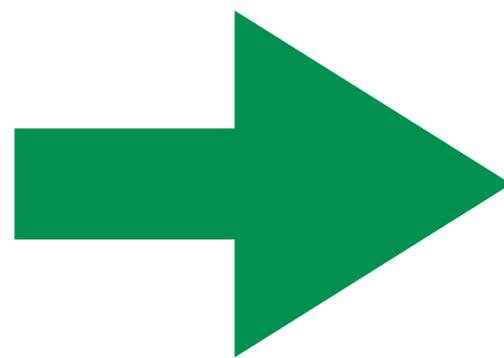
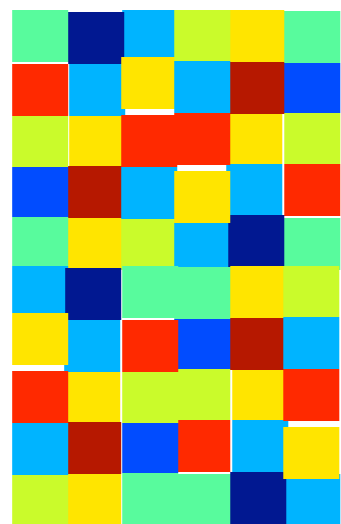
# How to reduce precision stat. uncerts.

23



Train with more events!

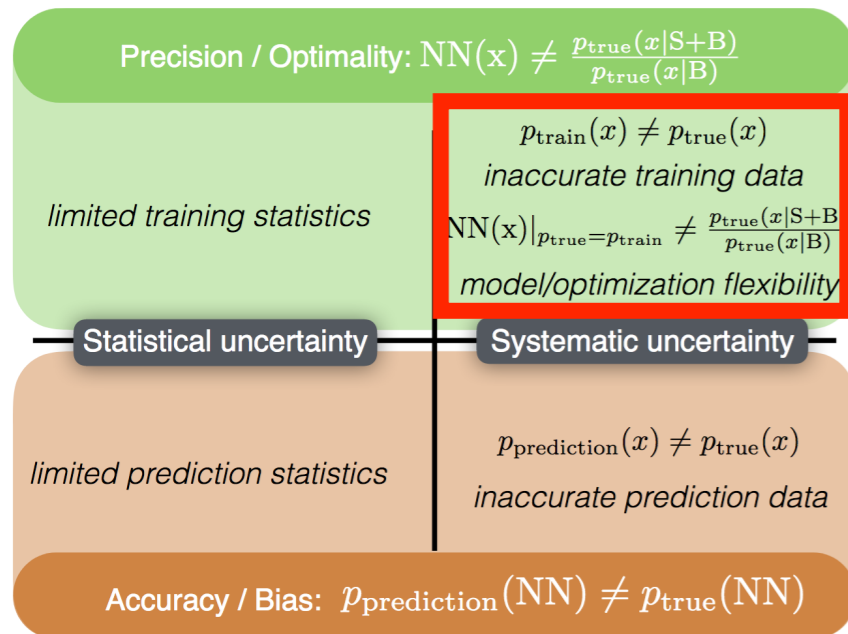
...maybe use NN's to help with that



M. Paganini, L. de Oliveira, BPN, PRL 120 (2018) 042003, 1705.02355 in particle physics and many more studies that have followed.

# How to reduce precision syst. uncerts.

24



Might be possible to reduce uncertainties or at least alleviate analysis complexity by making your NN independent of known nuisance parameters\*.

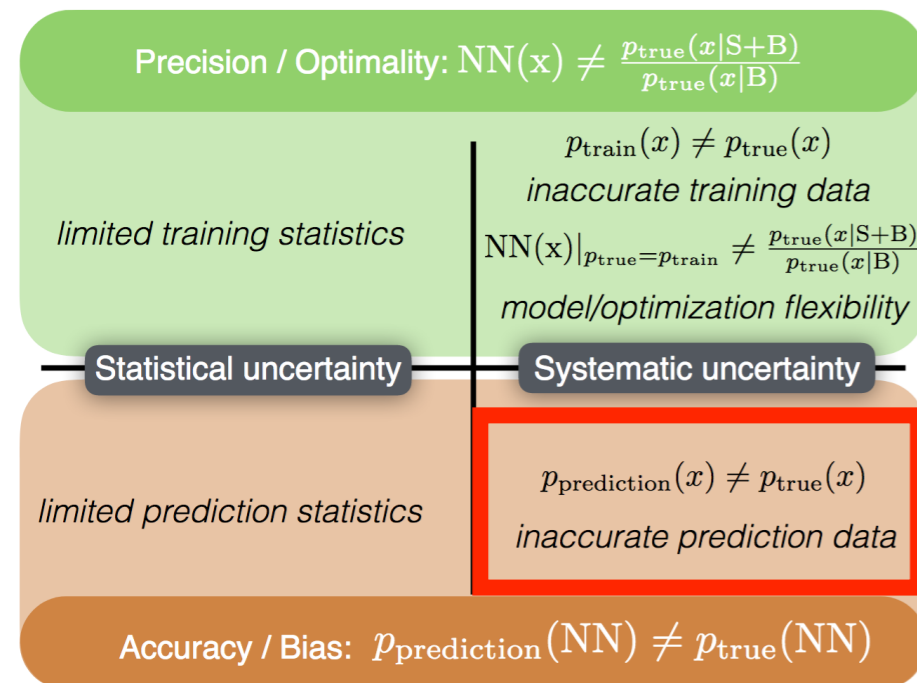
...might also be better to explicitly depend on the nuisance parameters and profile them in data.

\*see G. Louppe, et al., NIPS 2017, 1611.01046 for particle physics and many papers since.

# How to get around high-D bias uncersts?

25

Work hard to understand the true nuisance parameters in the hypervariate parameter space.



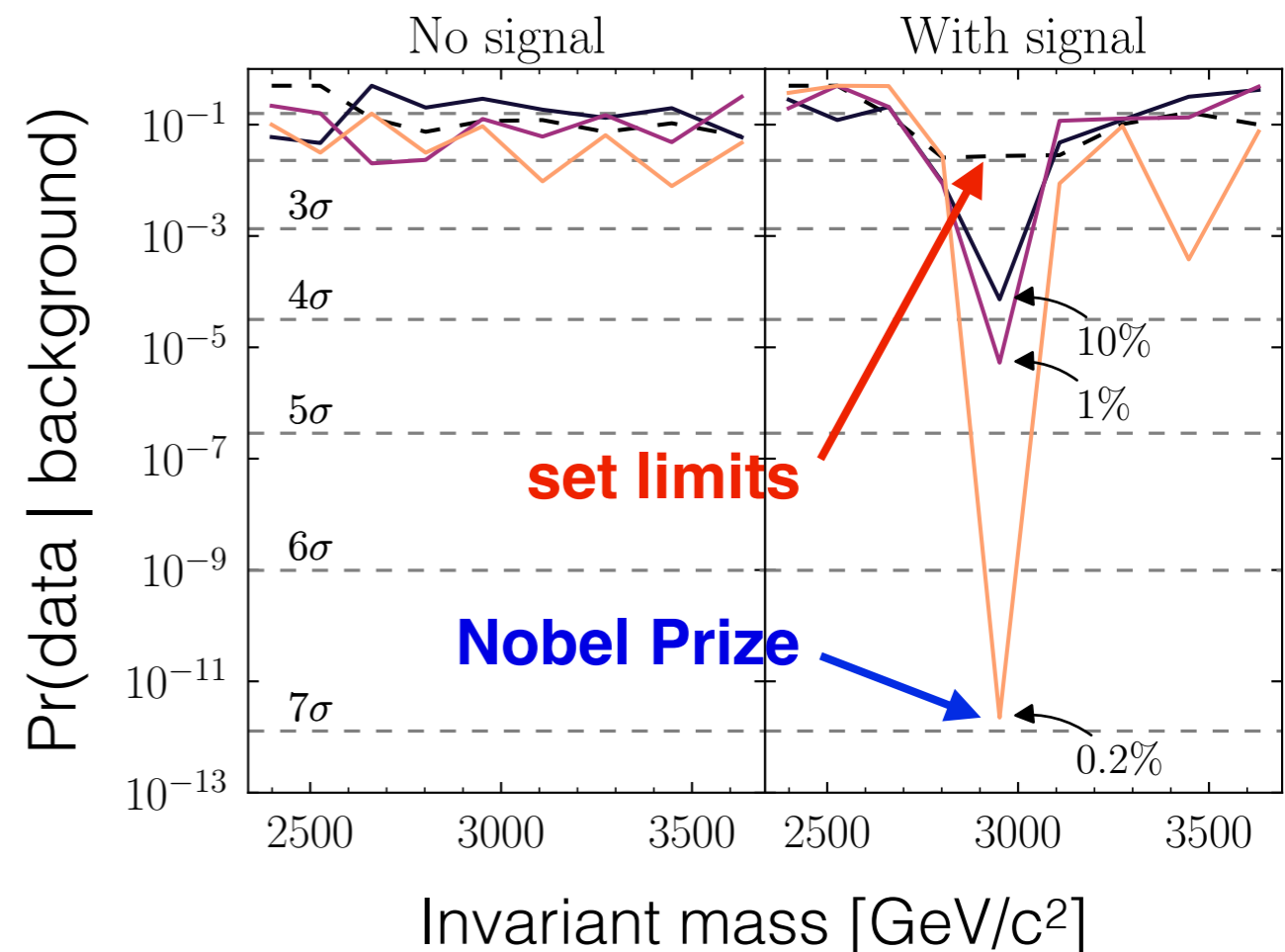
In my opinion, this is **THE** biggest challenge with deploying NN-based analyses ... solving it will require hard physics work.

# How to get around high-D bias uncerts?

26

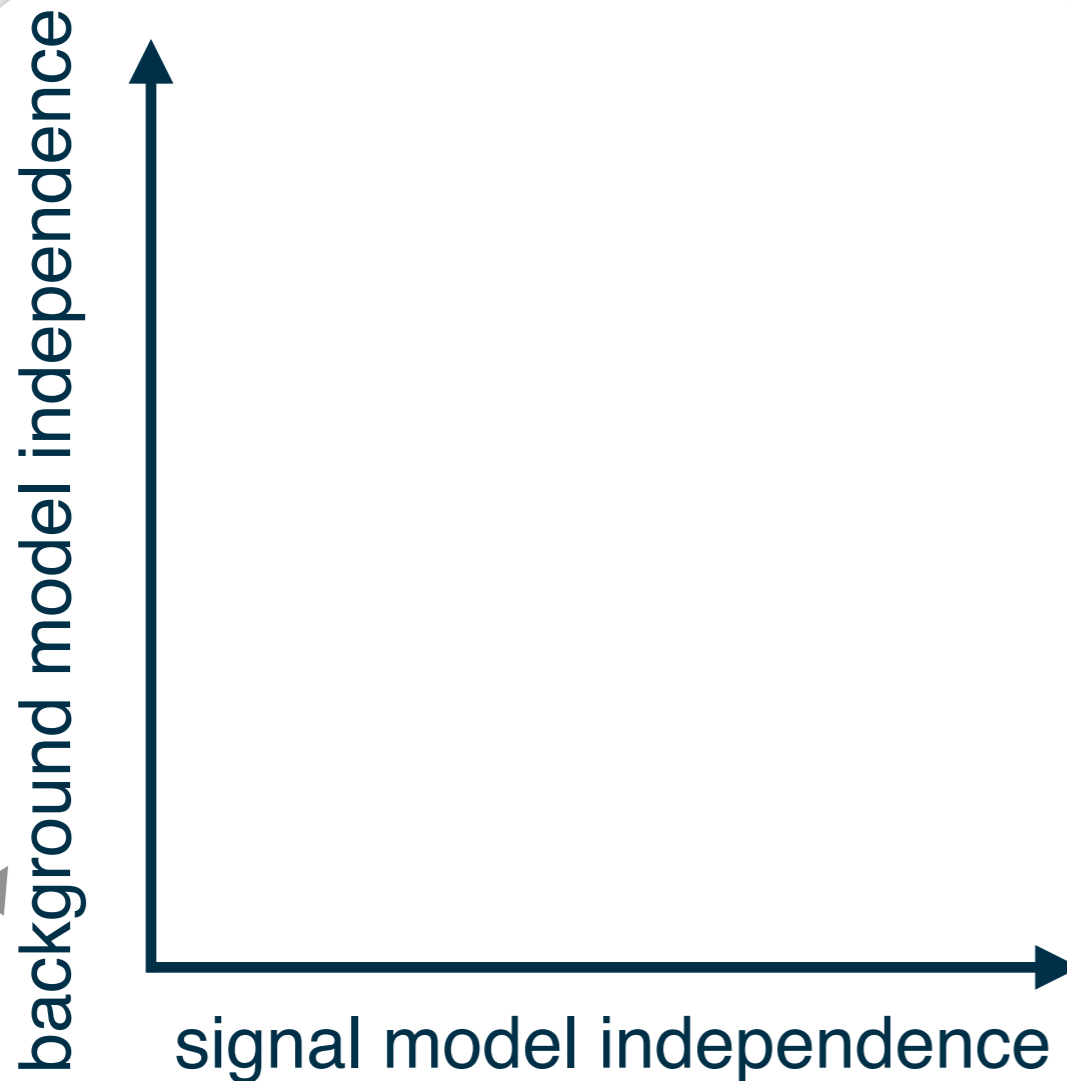
Work hard to understand the true nuisance parameters in the hypervariate parameter space.

**Don't use simulation!**  
(not always possible and of course, still has assumptions...)



# The landscape of model dependence

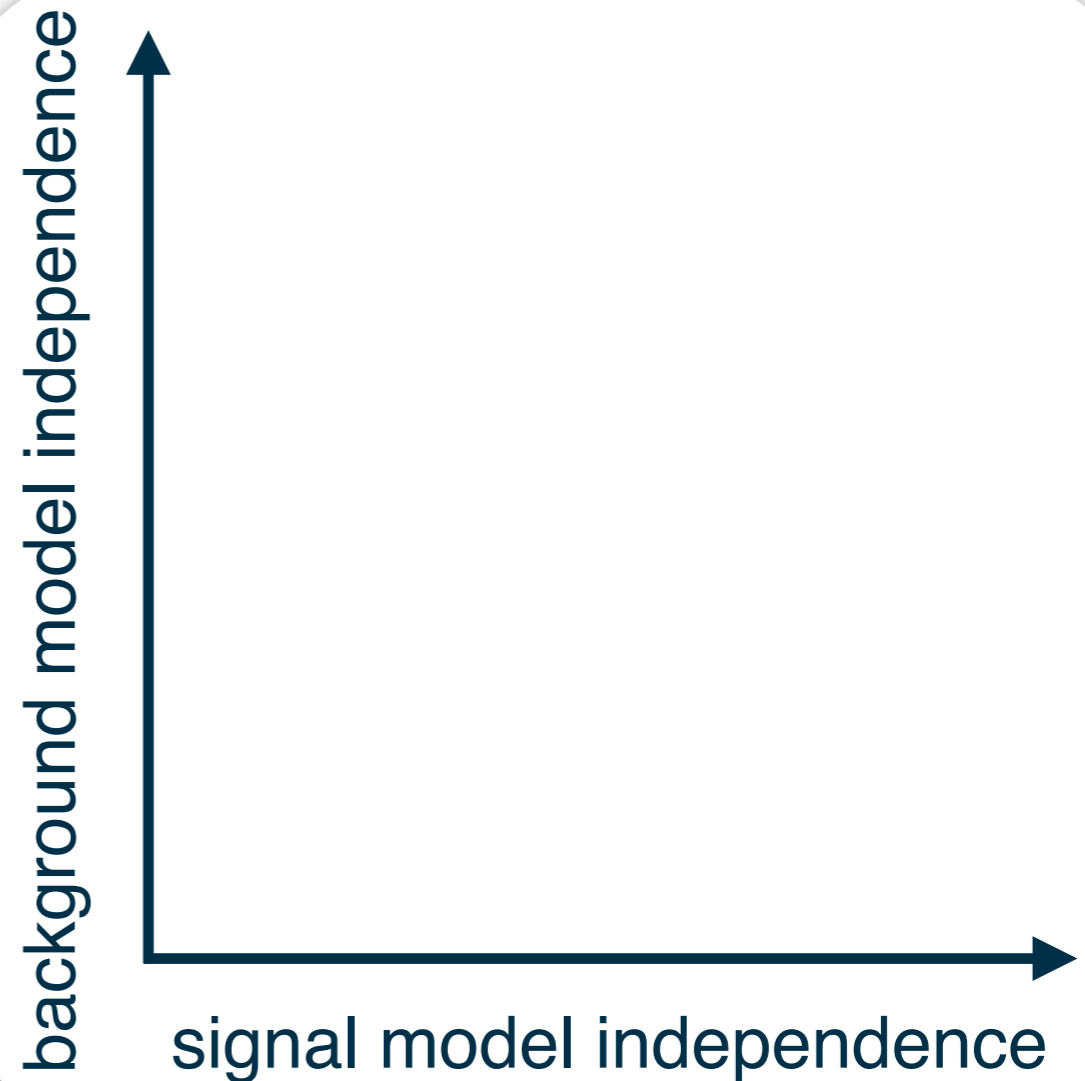
27



background model independence

signal model independence

Signal sensitivity



background model independence

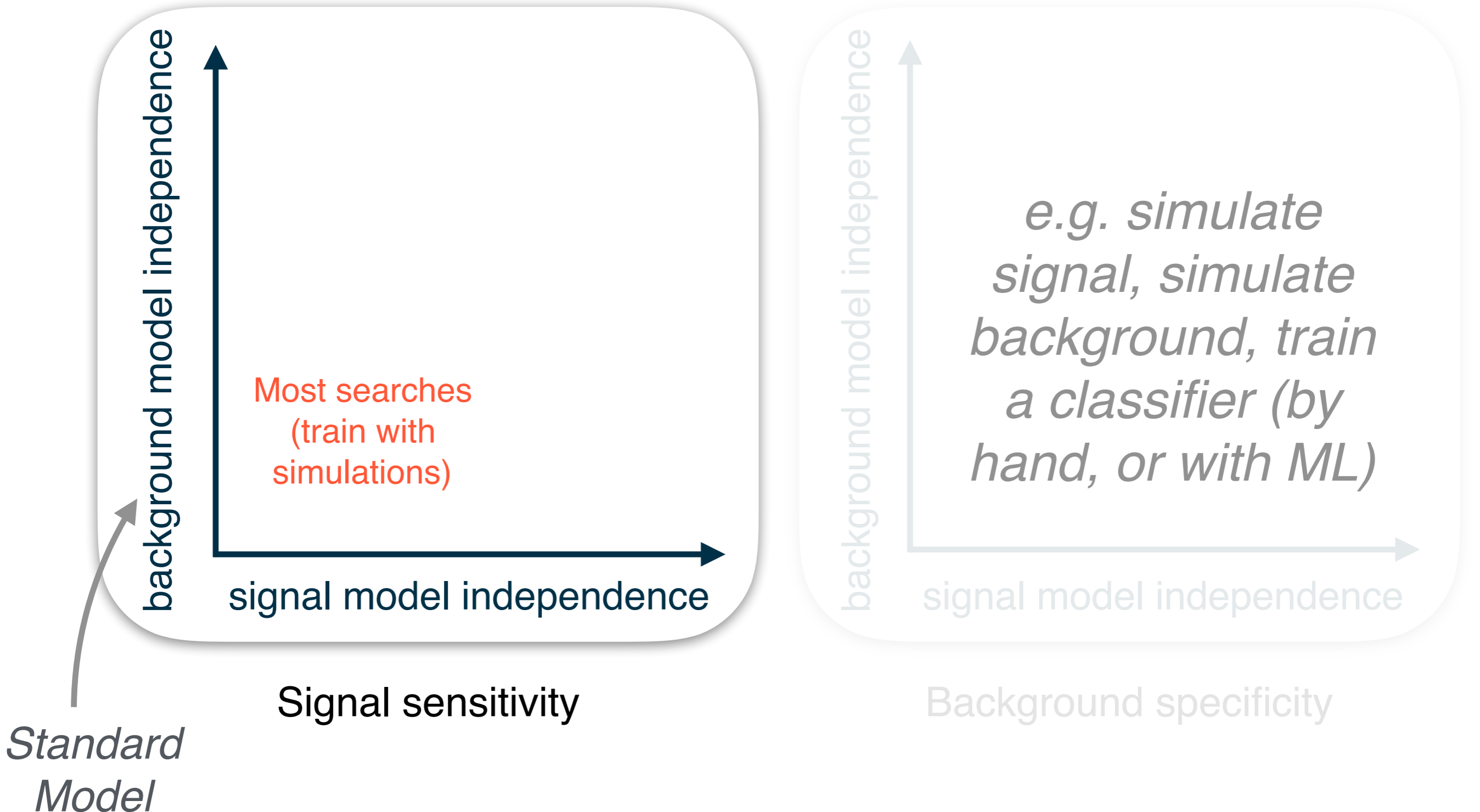
signal model independence

Background specificity

*Standard  
Model*

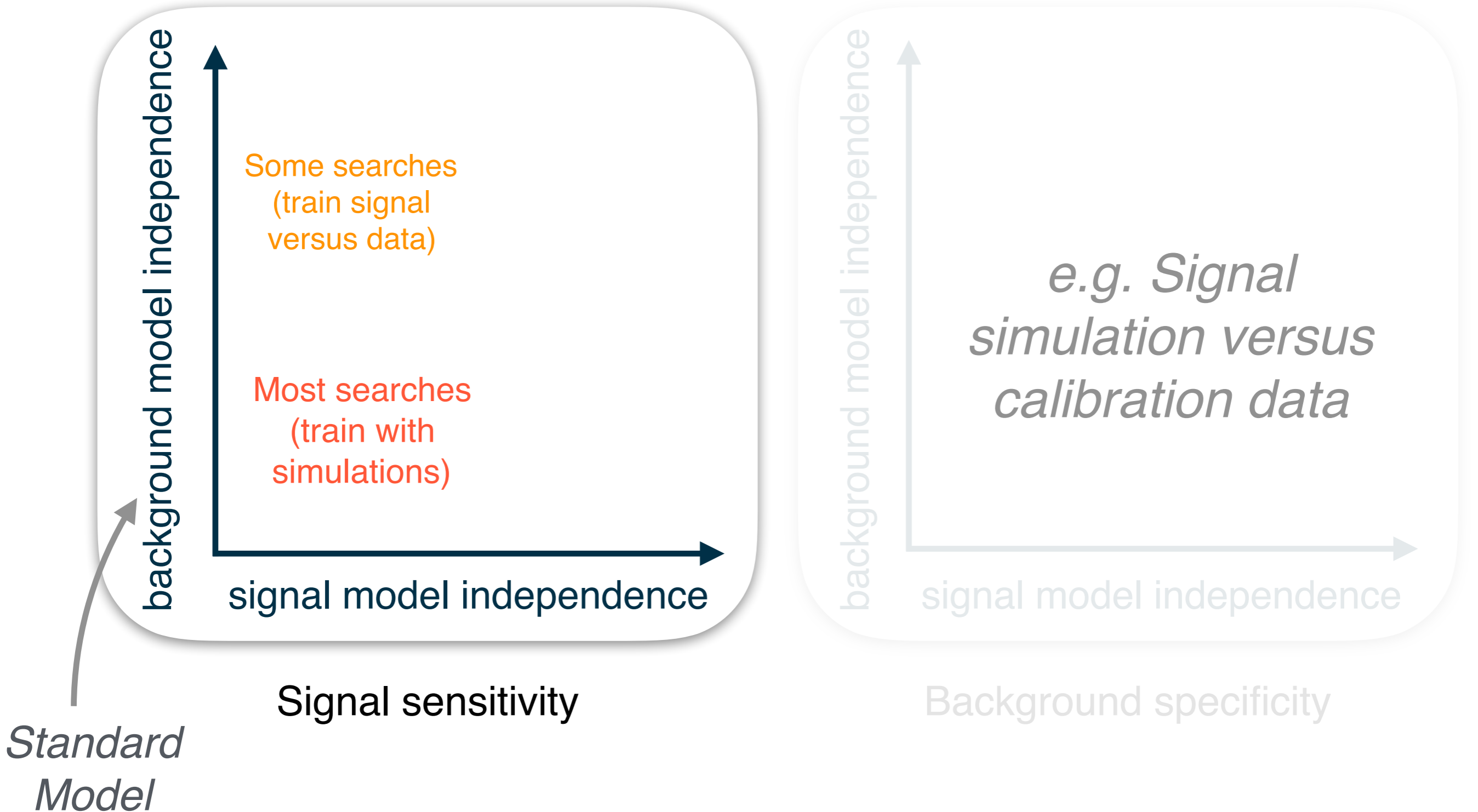
# The landscape of model dependence

28



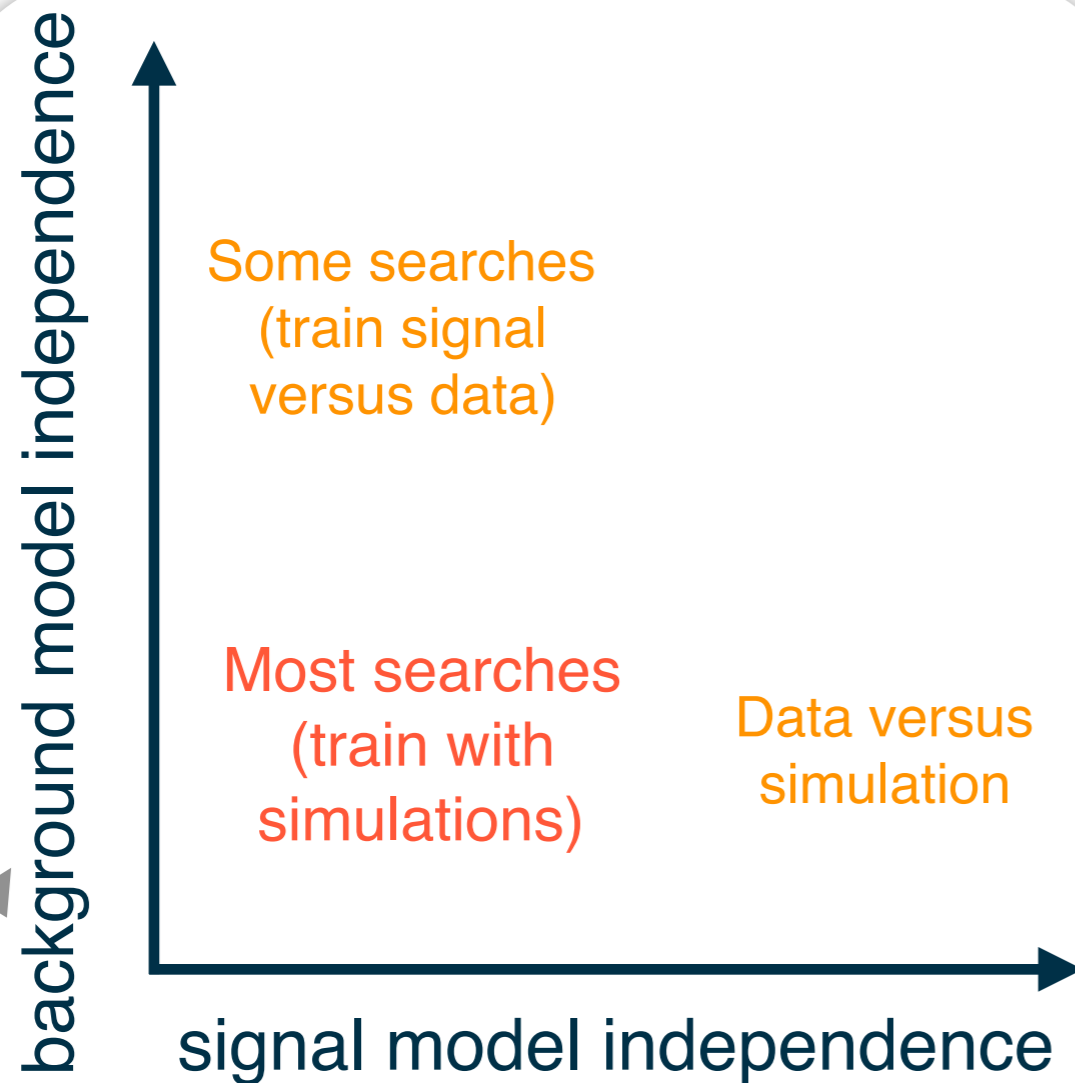
# The landscape of model dependence

29



# The landscape of model dependence

30



Signal sensitivity

*Standard  
Model*

This has a long history in the “non-ML” case, with the latest result from CMS earlier this week.

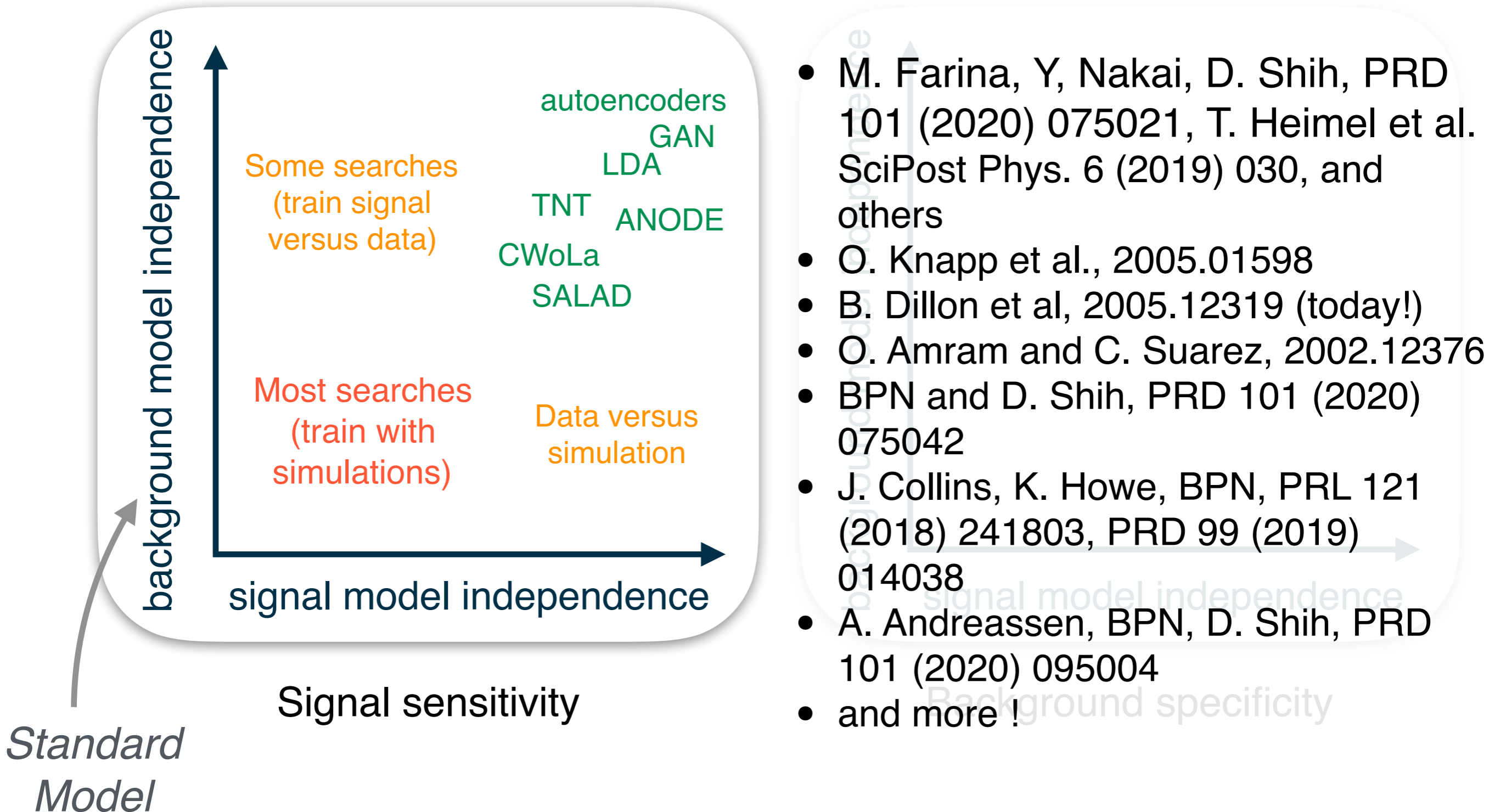
see B. Knuteson et al., Aleph, D0, H1, CDF, CMS (“MUSiC”), ATLAS (“General Search”)

This can be super-charged with machine learning, see e.g. R. T. D’Agnolo and A. Wulzer, PRD 99 (2019) 015014, and R. T. D’Agnolo et al. 1912.12155

Background specificity

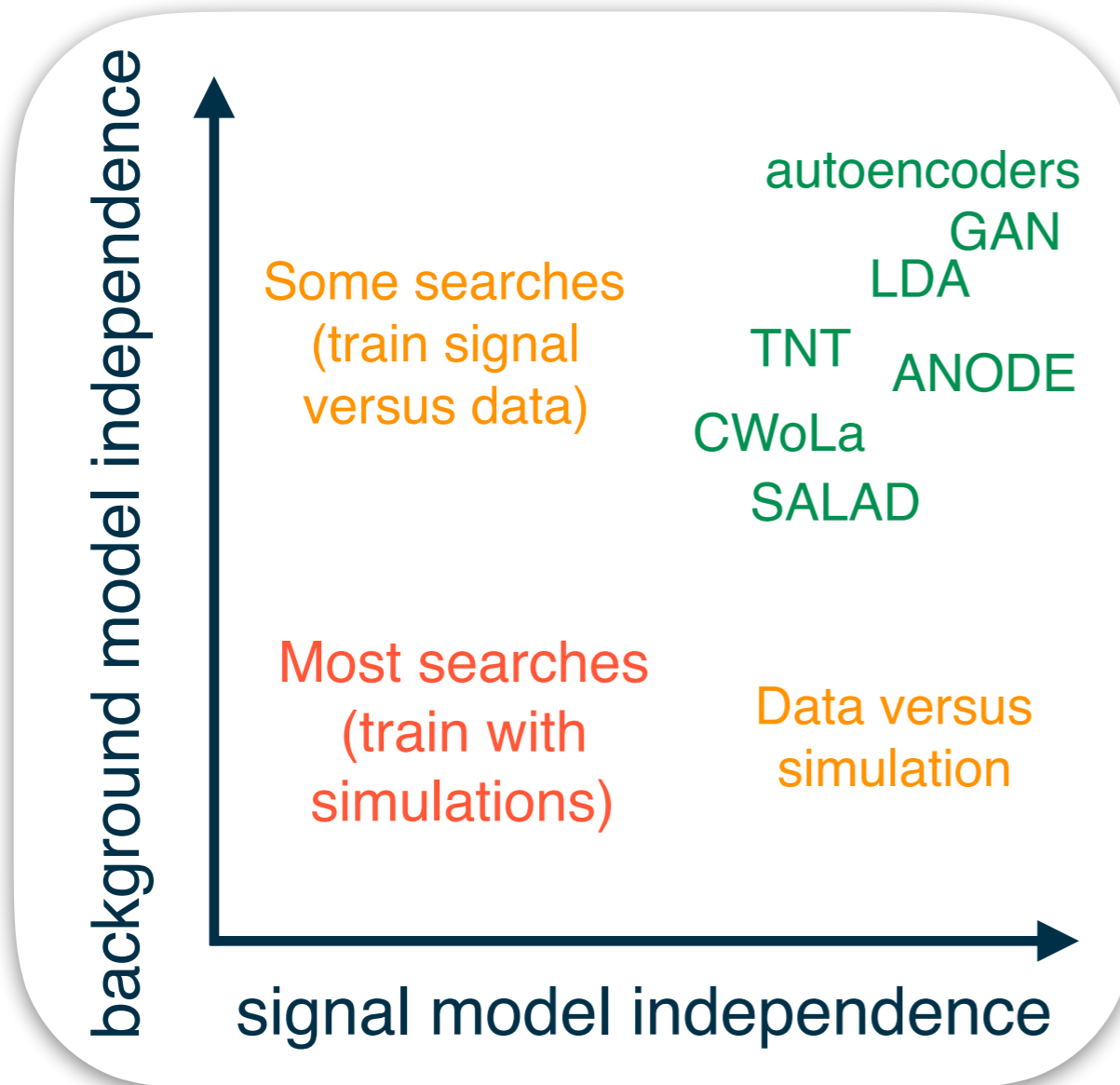
# The landscape of model dependence

31

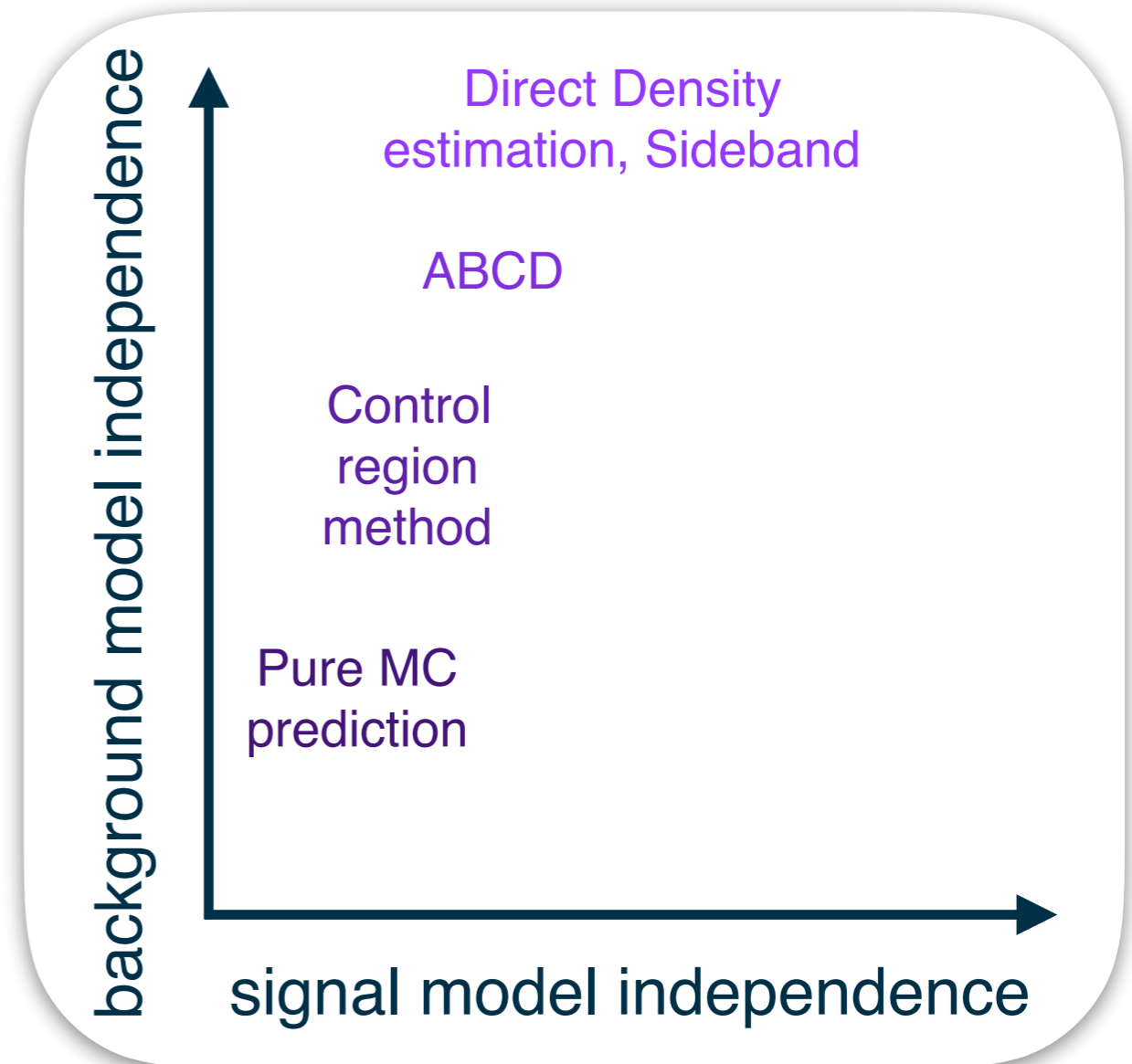


# The landscape of model dependence

32



Signal sensitivity



Background specificity

It is not enough to be sensitive to signal, need to also calibrate background ! Can mix and match some methods - some pairings are more natural than others.

# Anomaly detection future

33

*Rapidly developing  
area - **LHC Olympics**  
2020 to help facilitate!*



Summer Olympics  
will be virtual:  
[https://indico.desy.de/  
indico/event/25341/](https://indico.desy.de/indico/event/25341/)

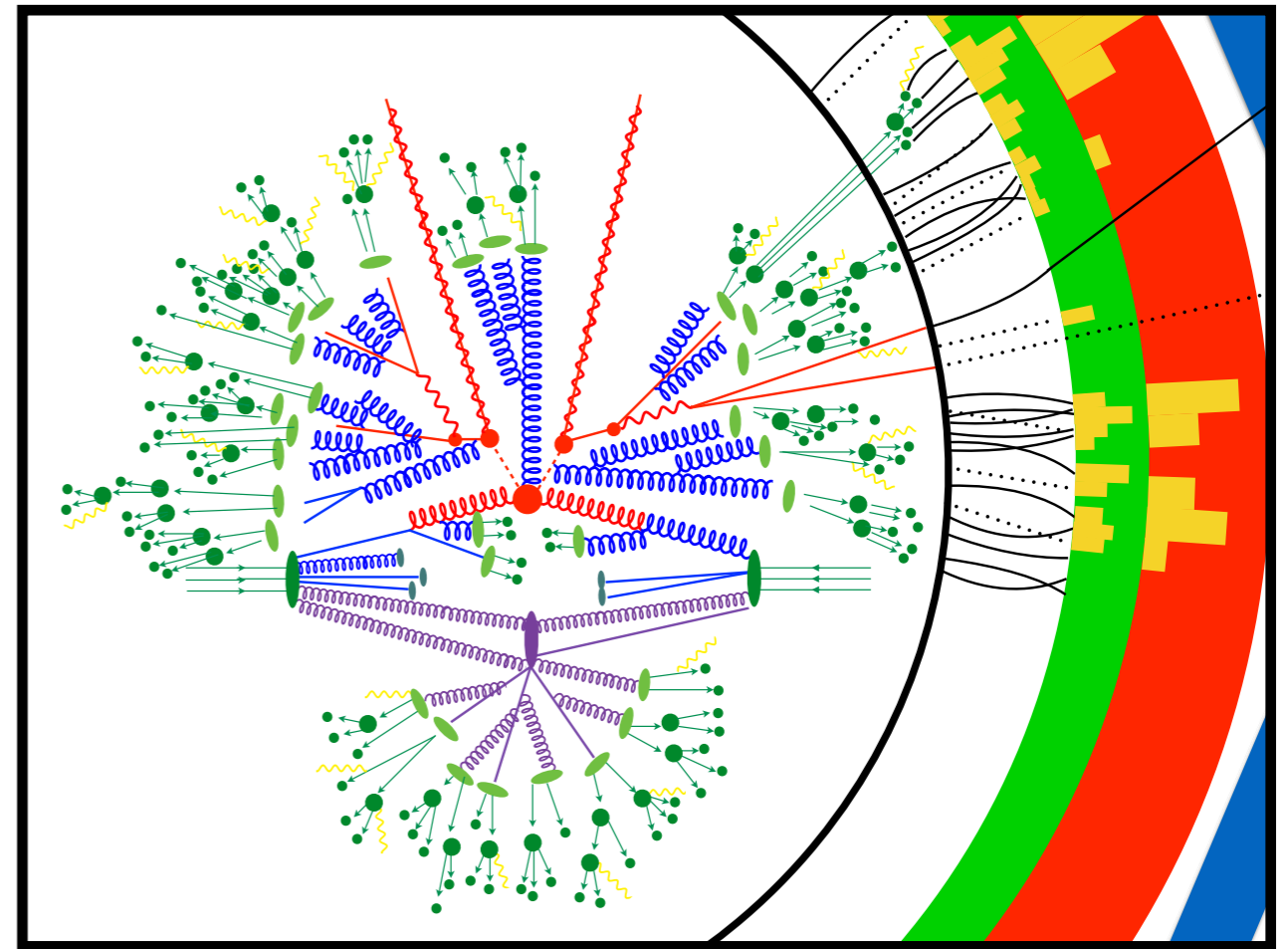
*G. Kasieczka. BPN, D. Shih  
<https://lhco2020.github.io/homepage/>*

# Conclusions and outlook

34

Deep learning has a great potential to **enhance**, **accelerate**, and **empower** HEP analyses

*I did not do justice to these topics, but many of them have been covered in other talks in this track!*



The **full phase space** of our experiments is now explorable, but we need to be cautious about new challenges from **uncertainty quantification in high dimensions**