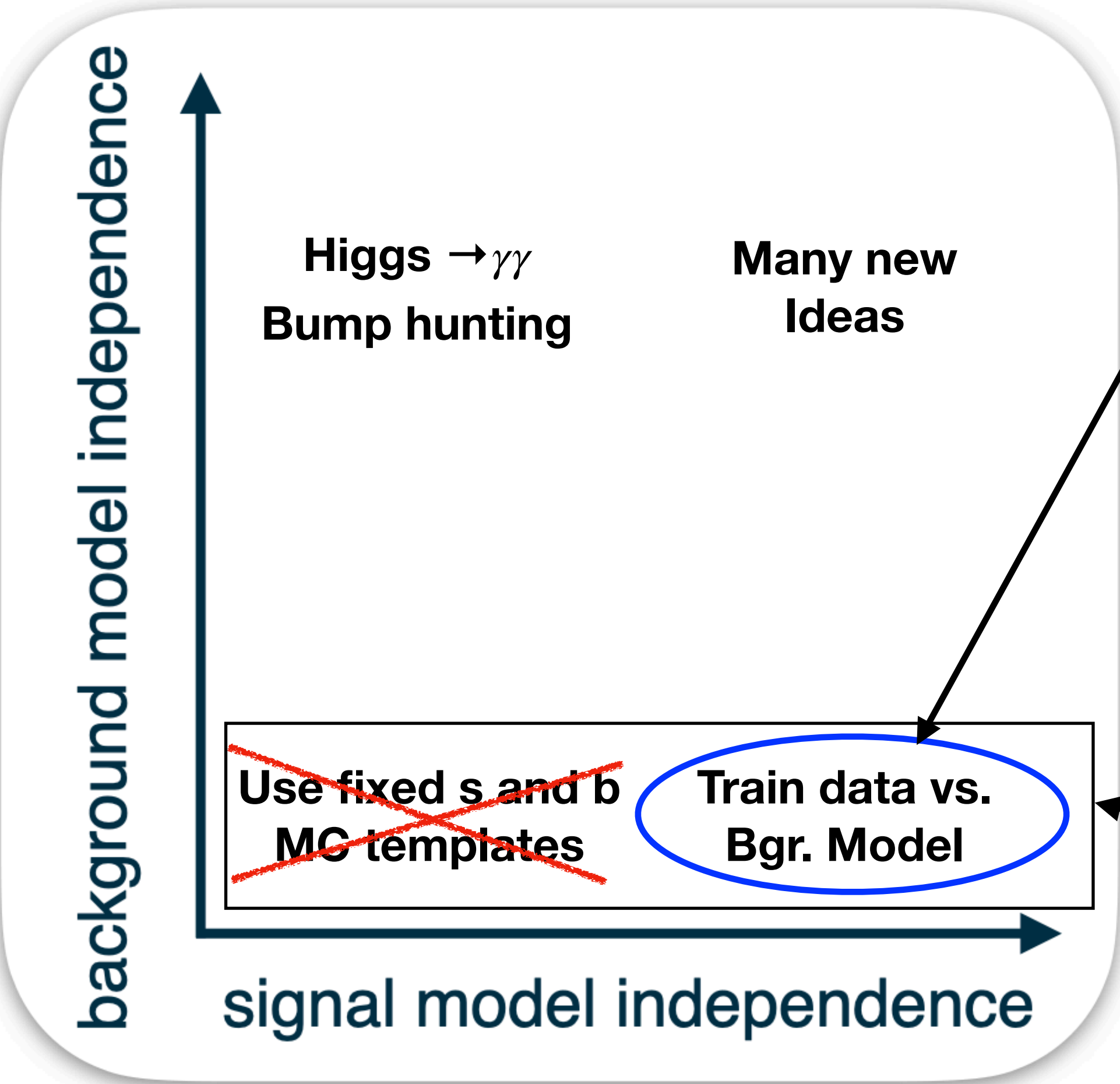


Workshop motto: Model independent searches

Ben Nachman’s landscape (slightly adapted)



A. Wulzer: Use Machine Learning

$n(x|\mathbf{w}) = n(x|R) e^{f(x;\mathbf{w})}$
Alternative in parametrised form.
 $f(x;\mathbf{w})$ is flexible function approximant
If $f(x;\mathbf{w})$ is a **neural network**

→ Is there some equivalence here?
ML can handle better bumpy signals?
ML: universal approximation with
superposition of sigmoidal functions
(G. Cybenko 1989)

$p(y) = b(y) e^{\sum_j \theta_j \phi_j(y) - Z}$
GOF-tests L. Wasserman

Neyman Smooth Tests

See also S. Algeri
arXiv:2009.00503

- Neyman Smooth Tests, Level Set Test,
- Bickel-Breiman Nearest Neighbour Test,
- LRT, Score Test,
- Bump test, Two Sample Test, RKHS,
- MMD, Energy Classifier tests

Louis' topic list:

Issues

Are model-independent and model-dependent distinct; or opposite ends of a spectrum?

Options for methods:

Believe background MC? Or choose control region depleted in signal

Choice of data

Input variables: (p_x , p_y , p_z , m) or higher level variables

Binned or unbinned data

Data statistic

Loss function

Form of anomaly: 1-D v multi-D; Peak or enhancement or any?

Can method 'target' types of anomalies? (cf. Kolmogorov-Smirnov or Anderson-Darling)

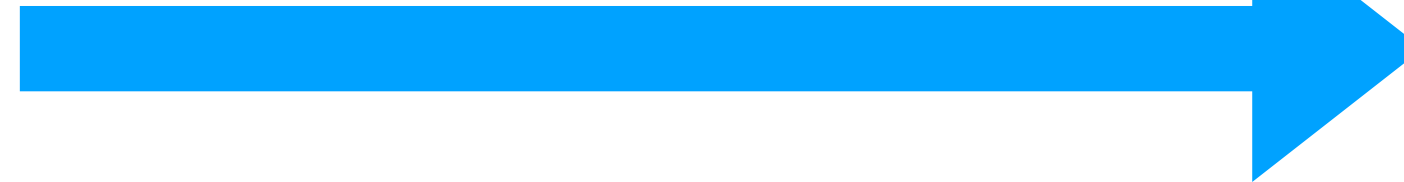
Provide info on where anomaly is?

Systematics from ML procedure; and other systematic issues

Publish discrepancies straight away or first confirm via model dependent approach?

What to publish if nothing found?

Our ~7 hours workshop



How to compare different anomaly detectors? Benchmarks for checking power.

LEE: Local or global p-value?

Where is 'Elsewhere'?

PHYSTAT-Anomalies: Setting the Scene

→ **Organise follow-up PHYSTAT events, formats/details to be discussed!**

For getting feedback on how you liked the workshop we will create a survey in indico and send you a link

Thanks to

Speakers **M. Kuusela, L. Lyons, B. Nachman, A. Wulzer, L. Wasserman, G. Kasieczka, I. Ochoa, S. Caron**

**Discussion
Leaders** **R. Cousins Jr., R. Lockhart**

**Session
Chairs** **M. Kuusela, T. Junk, A. Lee, D. Van Dyk**



and all participants for joining and contributing to lively discussions!

On behalf of the Orga team:



Olaf Behnke
Louis Lyons
Ben Nachman
Gregor Kasieczka
Mikael Kuusela



The PHYSTAT saga carries on

Next seminars:

15th June 2022, Tommaso Dorigo (Padova), "Sticking to the roots of machine learning: old-school approaches to physics analysis"

20th July 2022, Mikael Kuusela (CMU), "Gaussian Processes for Particle Physicists"

Next workshop:

27-30 September 2022: PHYSTAT-Gamma on high Energy Gamma Ray Astronomy in a Multi-Wavelength Context

~~Details/Links at home page:~~ **Details/Links at home page:** <https://espace.cern.ch/phystat/>

Please send suggestions for further topics/speakers to olaf.behnke@desy.de, louis.lyons@physics.ox.ac.uk



The PHYSTAT saga carries on

Next seminars:

15th June 2022, Tommaso Dorigo (Padova), "Sticking to the roots of machine learning: old-school approaches to physics analysis"

20th July 2022, Mikael Kuusela (CMU), "Gaussian Processes for Particle Physicists"

Next workshop:

Good bye!

27-30 September 2022: PHYSTAT-Gamma on high Energy Gamma Ray Astronomy in a Multi-Wavelength Context

~~Details/Links at home page:~~ <https://espace.cern.ch/phystat/>

Please send suggestions for further topics/speakers to olaf.behnke@desy.de, louis.lyons@physics.ox.ac.uk