Interpretable Machine Learning for probing Kinematic Shapes

Ayan Paul

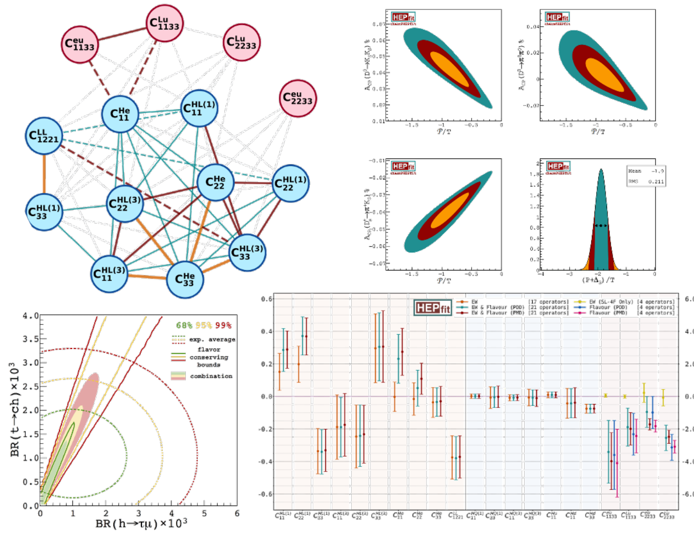
DESY, Hamburg & Humboldt Universität zu Berlin

A Sunday Afternoon on the Island of La Grande Jatte
-- Georges Seurat
On display at the Chicago Art Institute

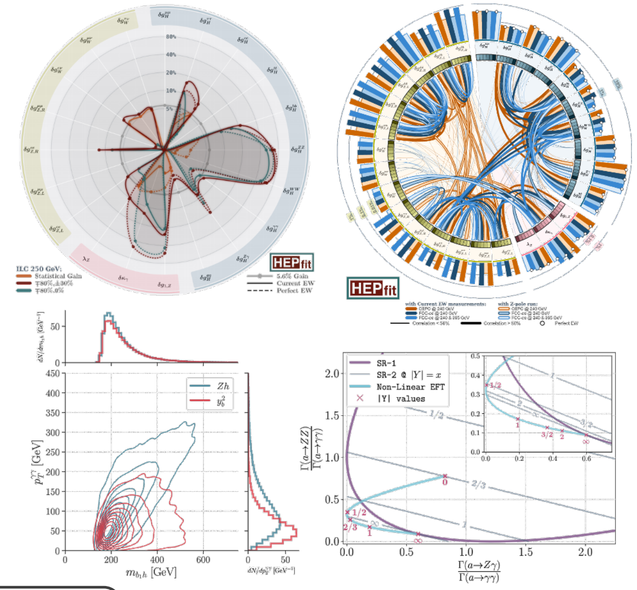
I am Ayan

- PhD, Theoretical Particle Physics, 2012
  - University of Notre Dame du Lac, USA
- Postdoctoral Fellow, 2012 – 2017
  - INFN, Sezione di Roma, Italia
- Fellow, 2017 – Present
  - Deutsches Elektronen-Synchrotron, Hamburg
- Senior Scientist, 2017 – Present
  - Humboldt Universität zu Berlin
- Chief Scientific Officer, 2020 – Present
  - Covis Inc, USA & Germany

PI – DESY Strategy Fund 2020
PI – VolkswagenStiftung Corona Response 2020



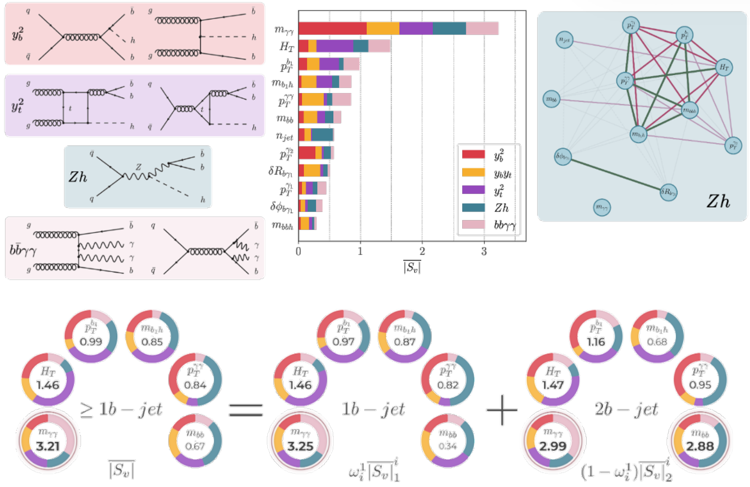
Flavour Physics



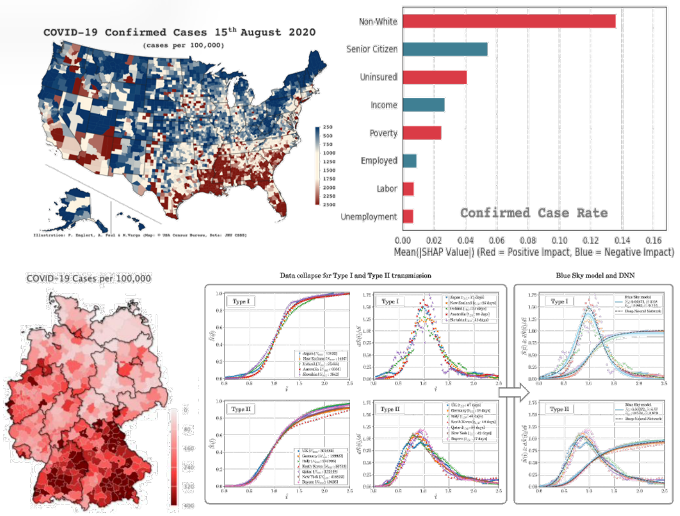
Higgs Physics

Understanding Nature

Interpretable Machine Learning



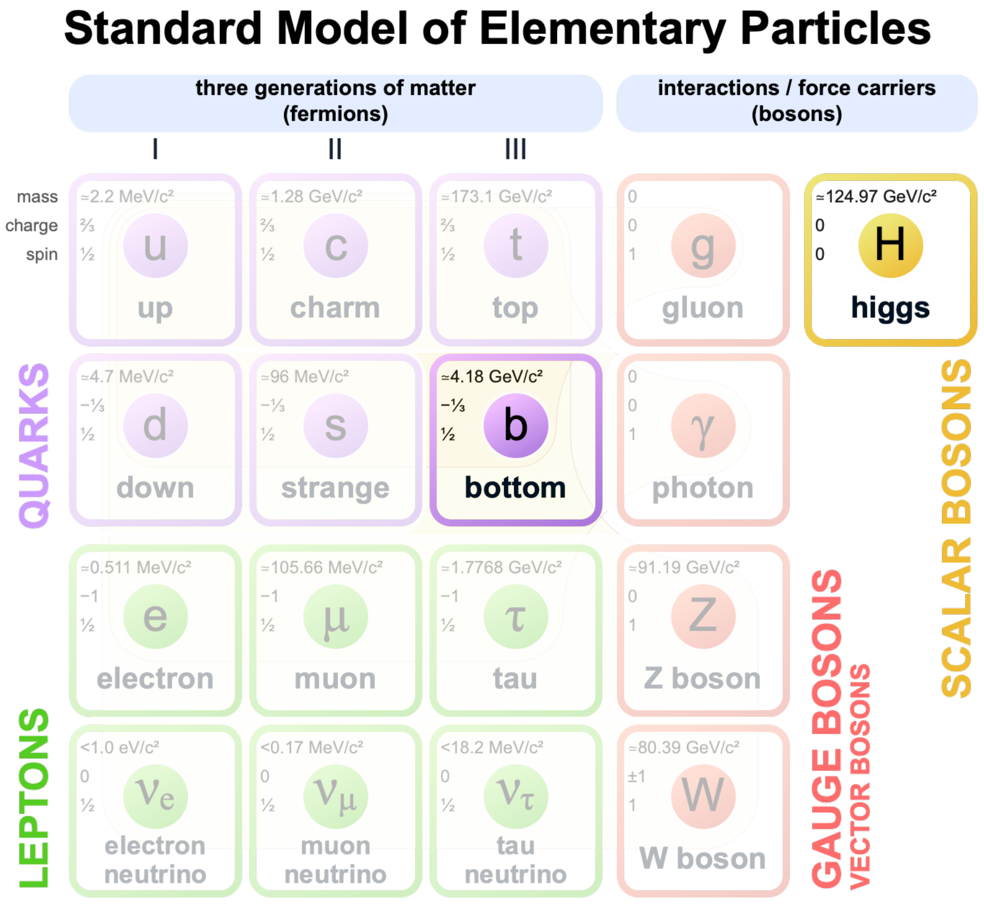
COVID-19



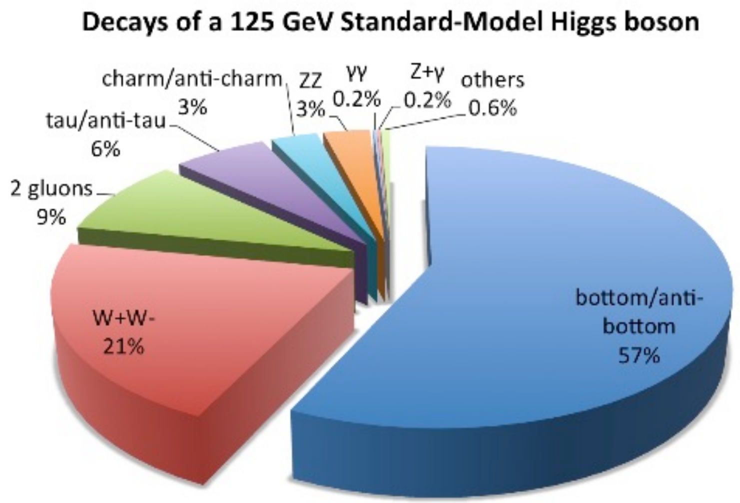
The Physics Case

Finding needles in a haystack:
when Higgs couplings become
hard to probe.

Higgs couplings and $b\bar{b}h$

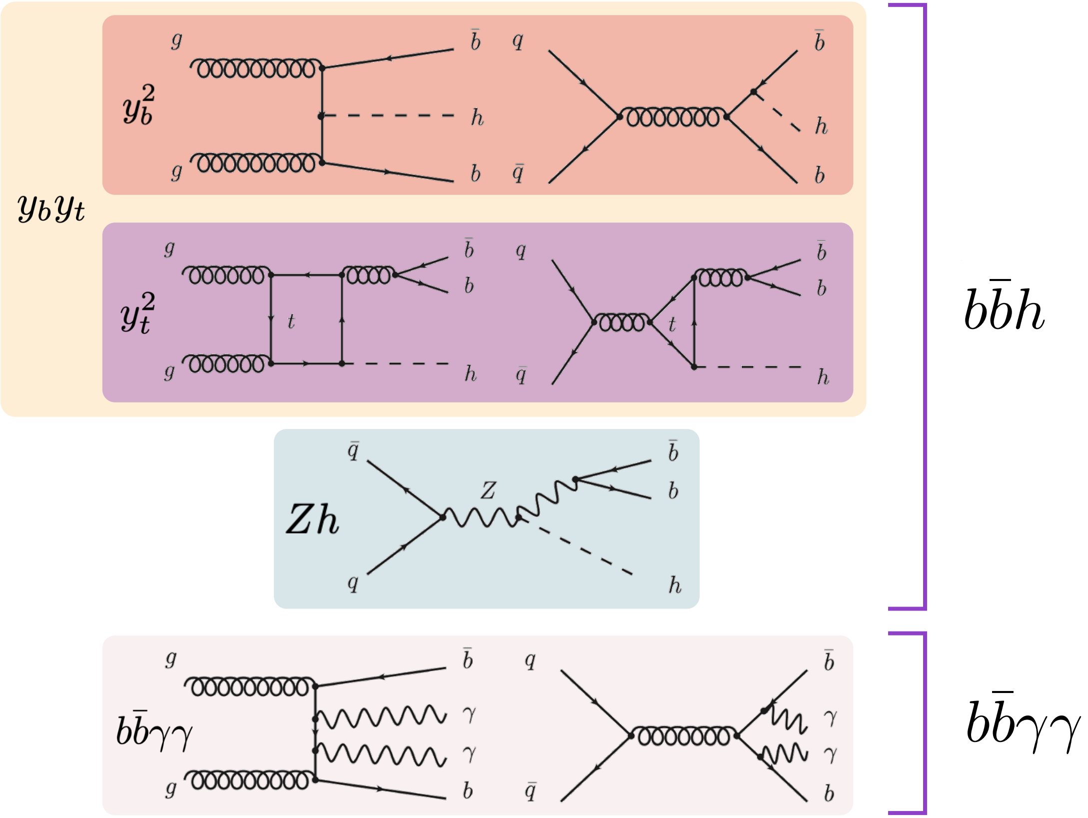


- Bottom Yukawa measurement is a recent achievement: $Vh, h \to b\bar{b}$
- The sign (or phase) of the Yukawa couplings have not been well measured
- There are possible interplays between Yukawa phases of various couplings in EDM measurements and collider physics



tiny signals embedded in multiple large backgrounds

Goal: measure bottom-Yukawa couplings



Traditional cut-based analysis cannot separate the different $b\bar{b}h$ contributions – no y_b sensitivity at HL-LHC

signals

Channel	LO σ (fb)	NLO-k-fact	6 ab^{-1}[#evt]	$2b$-jets[%]
y_b^2	0.0648	1.5	583	7.7%
$y_b y_t$	-0.00829	1.9	-95	4.0%
y_t^2	0.123	2.5	1,840	12%
Zh	0.0827	1.3	645	21%
$\sum b\bar{b}h$	0.262	-	2,970	-
$b\bar{b}\gamma\gamma$	12.9	1.5	116,000	14%

$b\bar{b}h$ background

QCD-QED background

C. Grojean, A. Paul and Zhuoni Qian, Resurrecting $b\bar{b}h$ with kinematic shapes.
JHEP 04 (2021) 139 [arXiv: 2011.13945]

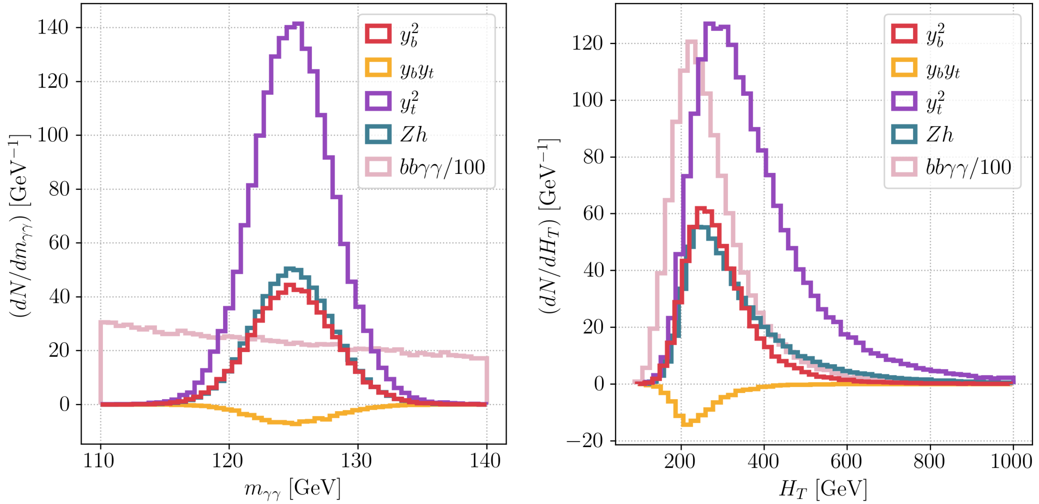
building an interpretable framework

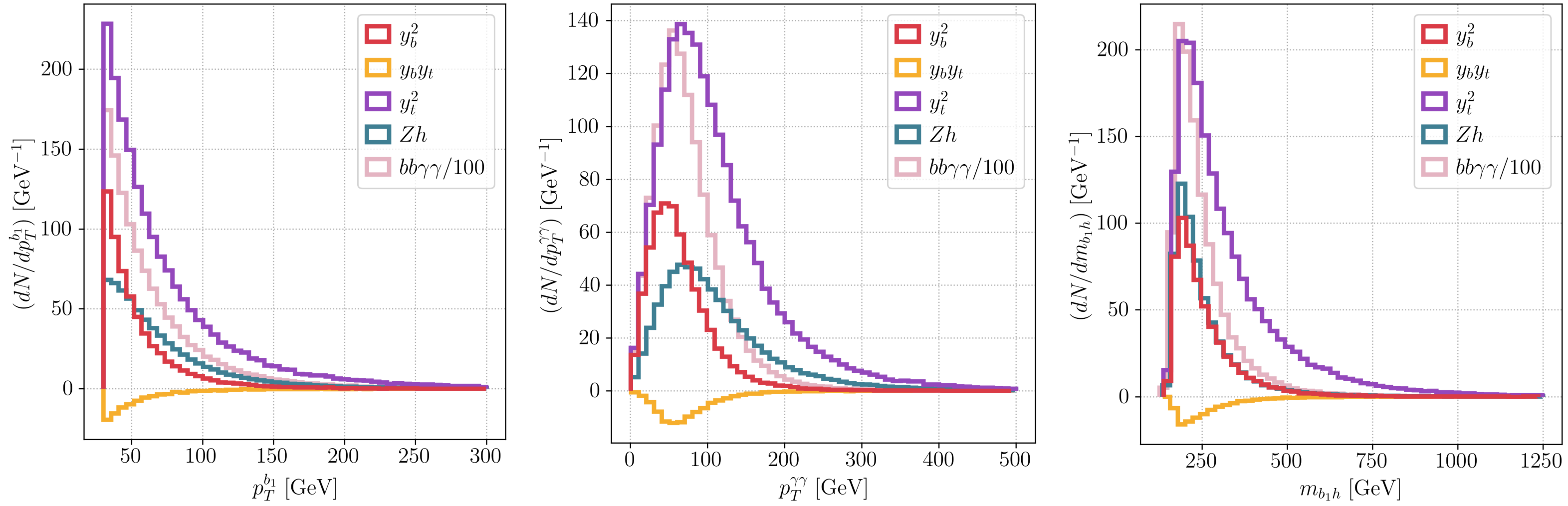
Understanding differences in shapes

- $p_T^{b_1}$, $p_T^{b_2}$, $p_T^{\gamma_1}$, $p_T^{\gamma\gamma}$,
- $\eta_{b_{j1}}$, $\eta_{b_{j2}}$, η_{γ_1}, $\eta_{\gamma\gamma}$,
- n_{bjet}, n_{jet}, $\Delta R_{\min}^{b\gamma}$, $\Delta\phi_{\min}^{bb}$,
- $m_{\gamma\gamma}$, m_{bb}, $m_{b_1 h}$, $m_{b\bar{b}h}$, H_T.

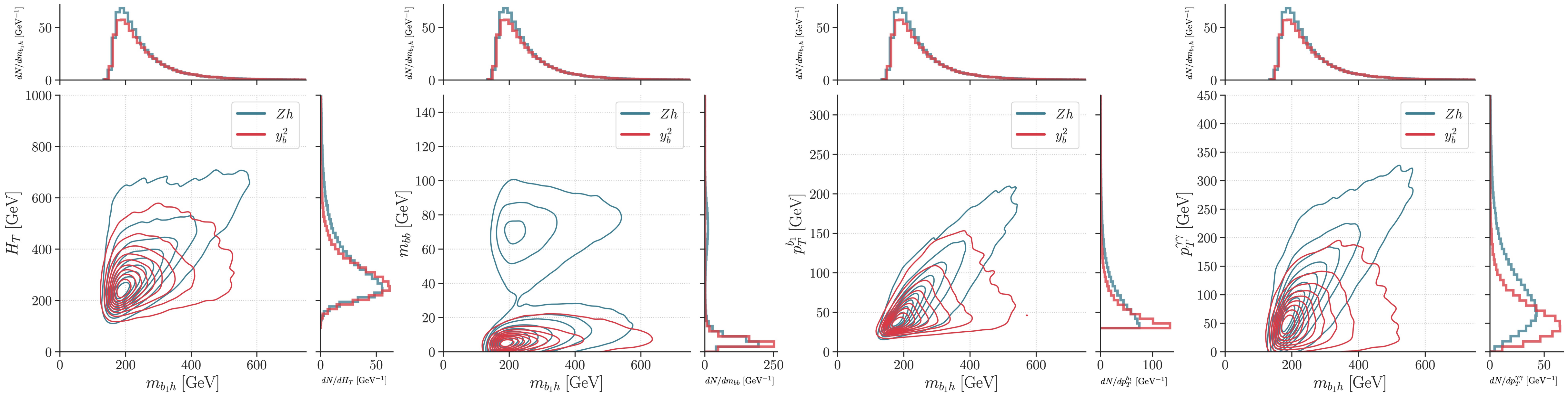
The choice of variables is important:

- Momenta four vectors are not easily interpretable
- Kinematic variables are interpretable but there is no clear "complete set"





the devil is in the correlation



- Cut-based analyses start to falter with multivariate correlations – difficult to visualize and interpret
- Machine learning algorithms excel at multivariate analyses

- Machine learning algorithms are essentially black-boxes – not good for understanding the underlying dynamics

References

Measurement of $H \to b\bar{b}$

- J. M. Butterworth, A. R. Davison, M. Rubin and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. 100 (2008) 242001, [arXiv:0802.2470].
- ATLAS collaboration, M. Aaboud et al., *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B 786 (2018) 59–86, [arXiv:1808.08238].
- CMS collaboration, A. M. Sirunyan et al., *Observation of Higgs boson decay to bottom quarks*, Phys. Rev. Lett. 121 (2018) 121801, [arXiv:1808.08242].

Recent theory papers on $b\bar{b}h$

- N. Deutschmann, F. Maltoni, M. Wiesemann and M. Zaro, *Top-Yukawa contributions to $b\bar{b}h$ production at the LHC*, JHEP 07 (2019) 054, [arXiv:1808.01660].
- D. Pagani, H.-S. Shao and M. Zaro, *RIP $Hb\bar{b}$: How other Higgs production modes conspire to kill a rare signal at the LHC*, JHEP 11 (2020) 036 [arXiv:2005.10277].

Papers on Higgs couplings fits

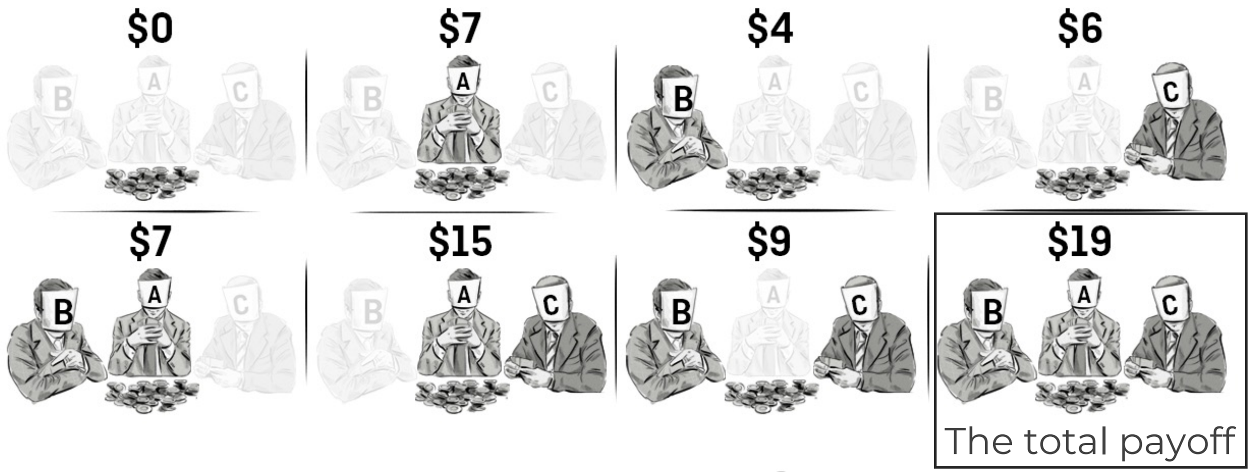
- M. Cepeda et al., *Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC*, vol. 7, pp. 221–584. 12, 2019. arXiv:1902.00134.
- J. de Blas et al., *Higgs Boson Studies at Future Particle Colliders*, JHEP 01 (2020) 139, [arXiv:1905.03764].

Cooperative Game Theory

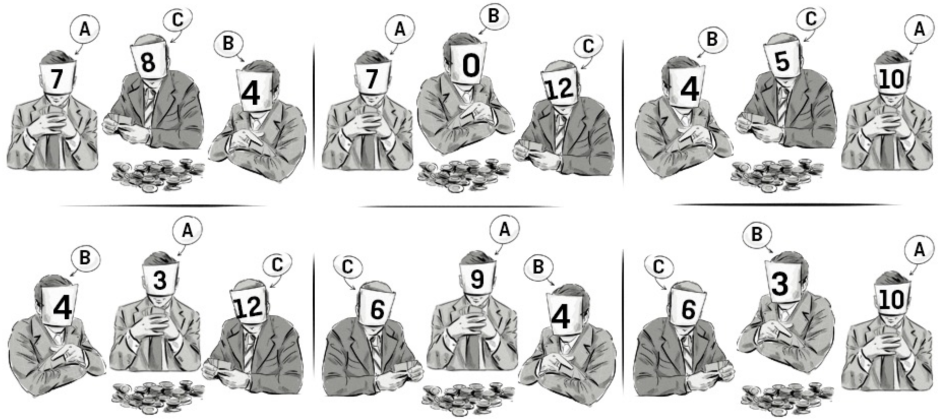
Because correlations are important, and a game needs to be played

a cooperative game

The value of each player and each combination of players



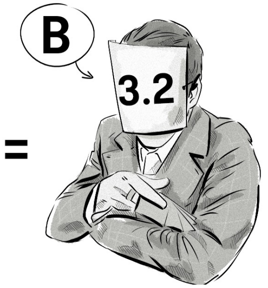
The value of the player in each game



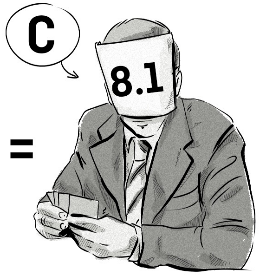
Marginalized values

(7+7+10+3+9+10) / 6 =

(4+0+4+4+4+3) / 6 =



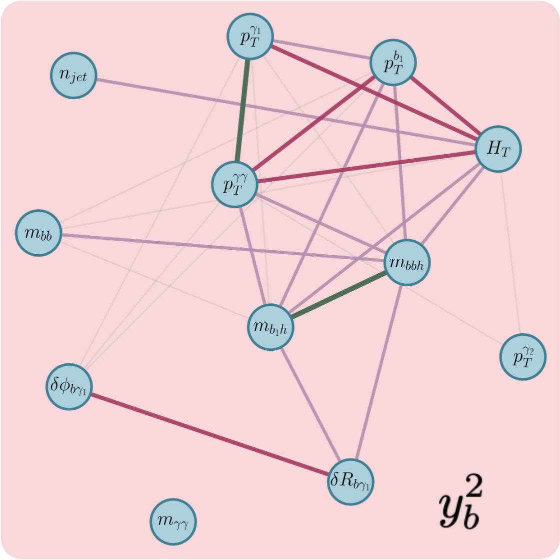
(8+12+5+12+6+6) / 6 =

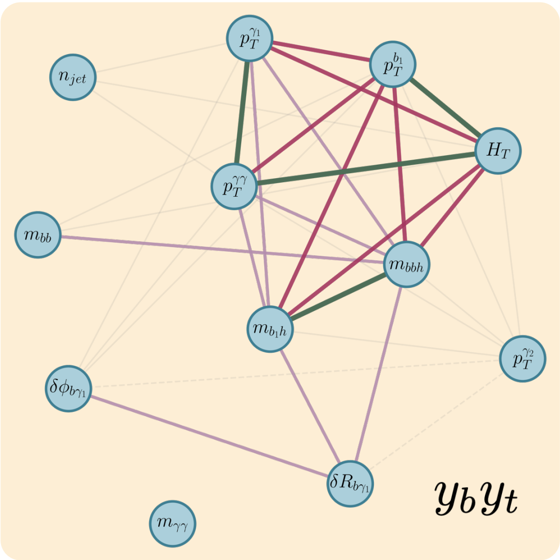


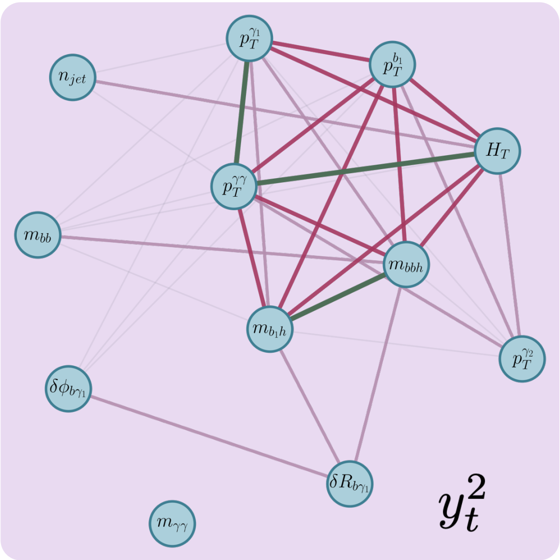
The most important player

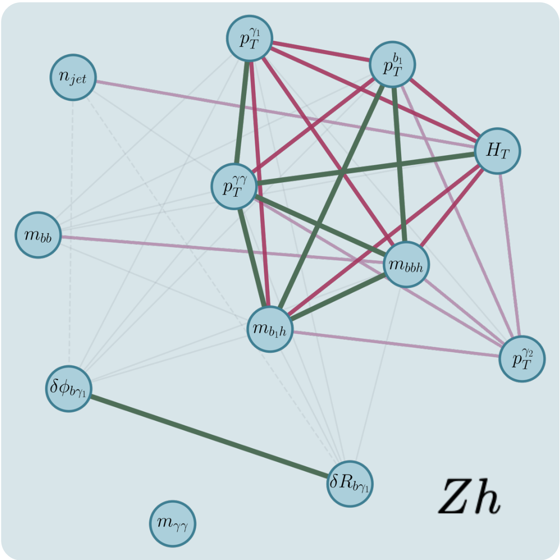
L. S. Shapley, Notes on the n-Person Game-II: The Value of an n-Person Game (1951).

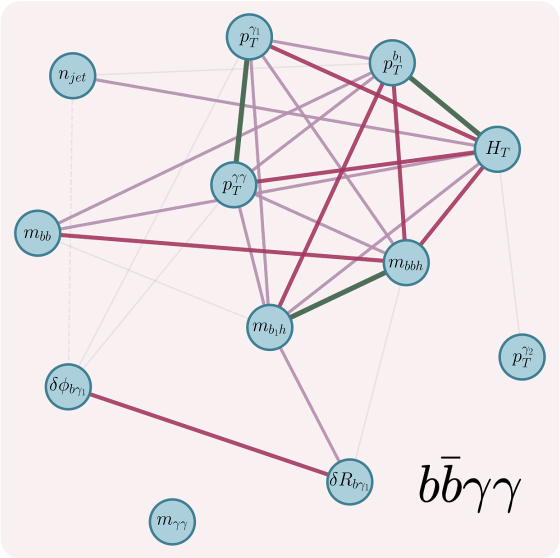
cooperation in Physics











$0.2 \leq |\rho| < 0.4$ $0.4 \leq |\rho| < 0.6$ $0.6 \leq |\rho| < 0.8$ $0.8 \leq |\rho| < 1.0$

$-\rho$ $+\rho$

multivariate inherits correlations!

- Variables "cooperate" to bring the outcome
- Outcome can be a measurable quantity or a probability of being of a certain kind
- This covers both regression and classification

some useful properties of Shapley values

For a game $\mathcal{G} = (\mathcal{K}, \nu)$ with a set $\mathcal{K}$ of players and a payoff ν:

Dummy Player: A player that doesn't contribute to any subset of players must receive zero attribution

$$\phi_k(\nu) = 0.$$

Efficiency: Attributions must add to the total gain

$$\sum_{k \in \mathcal{K}} \phi_k(\nu) = \nu\left(\mathcal{K}\right).$$

Symmetry: Symmetric players must receive equal attribution

$$\nu(\mathcal{A} \cup \{k\}) = \nu(\mathcal{A} \cup \{i\}) \implies \phi_k(\nu) = \phi_i(\nu).$$

Linearity: Attribution for the (weighted) sum of two games must be the same as the (weighted) sum of the attributions for each of the games

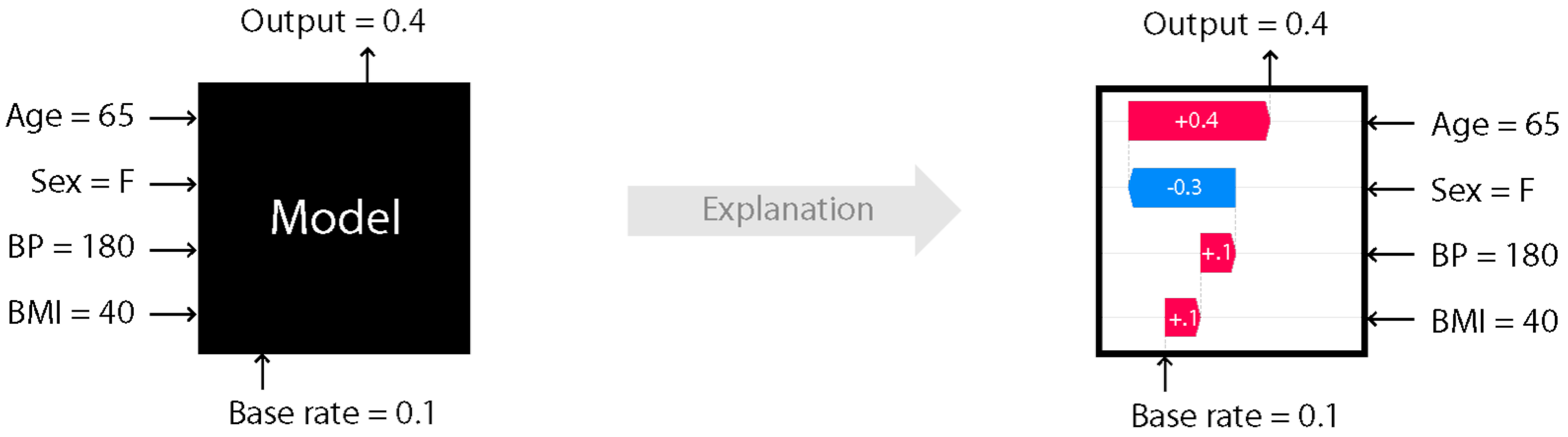
$$\phi_k\left(\nu + \omega\right) = \phi_k(\nu) + \phi_k(\omega) \qquad \forall\, k \in \mathcal{K}.$$

Shapley Additive exPlainers

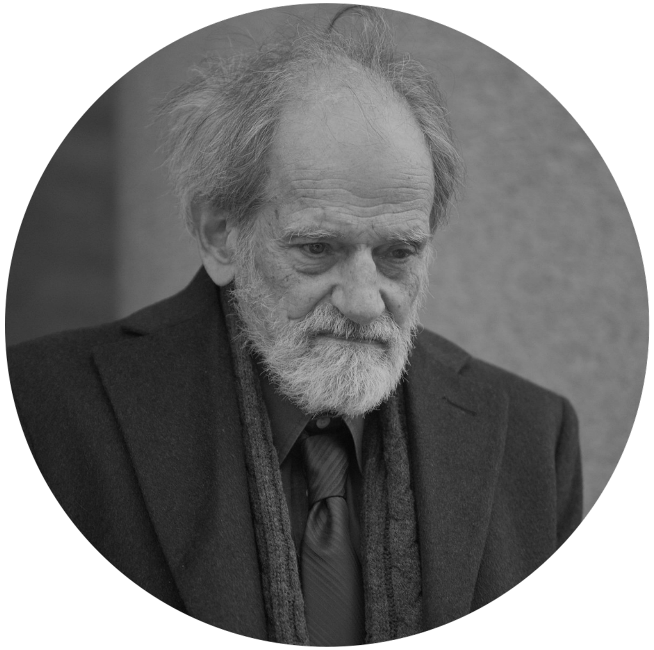
SHAP

Black-box model

Interpretable model



Local interpretation: event by event



Lloyd S. Shapley
Nobel Laureate 2012

S. M. Lundberg et al., *From local explanations to global understanding with explainable AI for trees.* [Nature Machine Intelligence 2, 56–67](#) (2020)

References

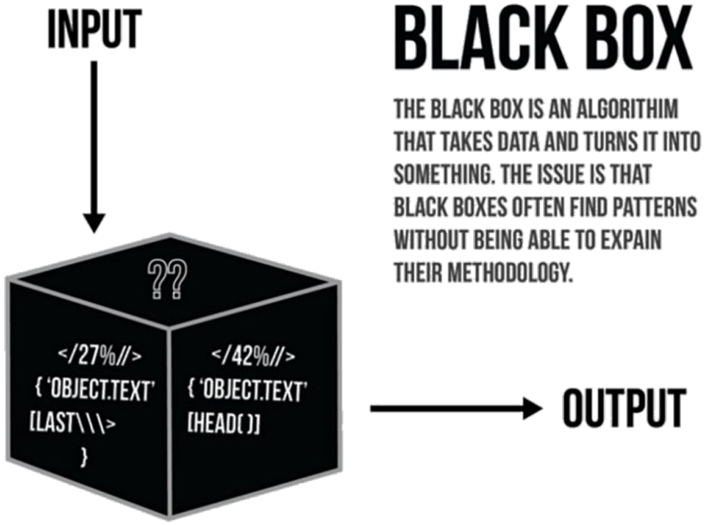
Shapley values and interpretable machine learning

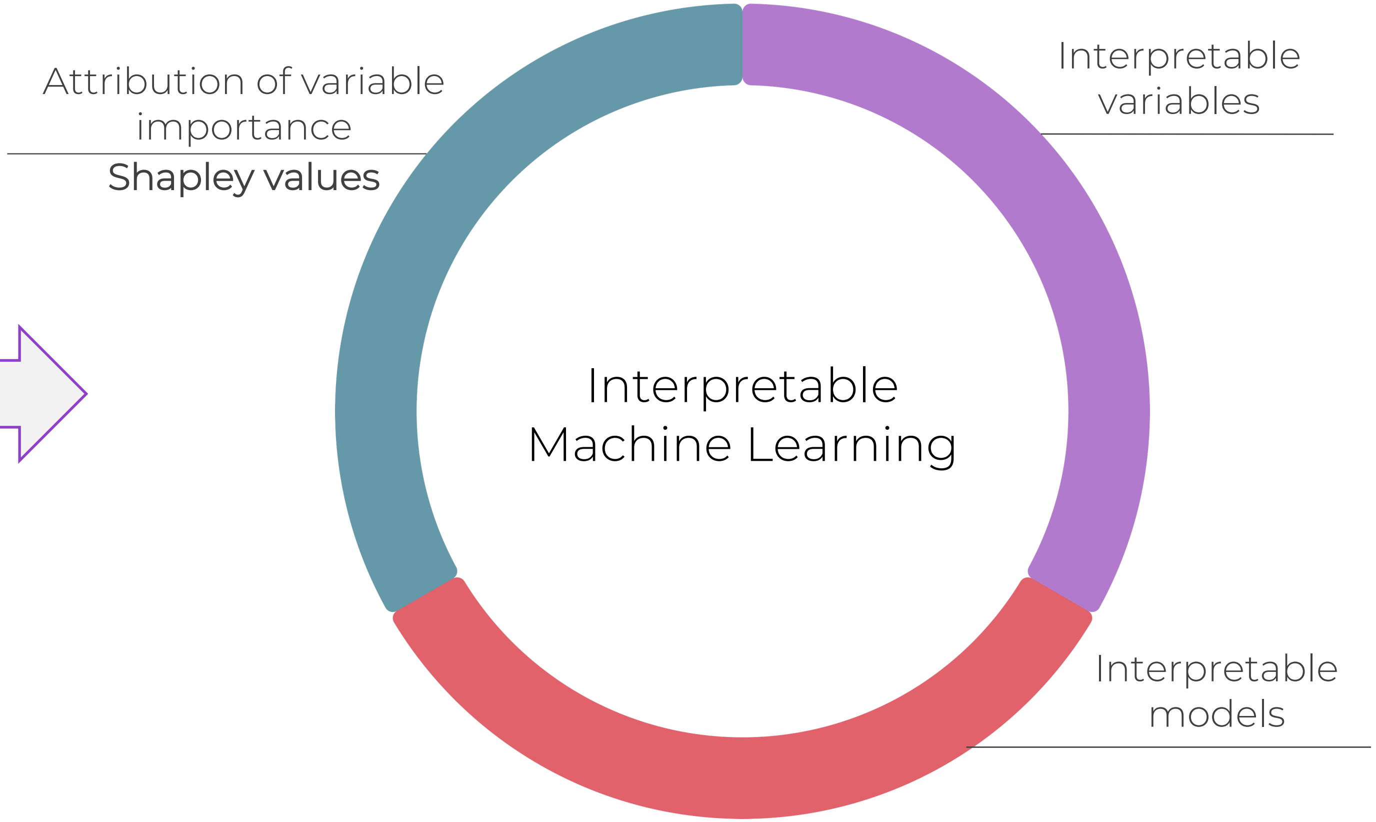
- L. S. Shapley, *Notes on the n-Person Game-II: The Value of an n-Person Game*, Rand Corporation (1951).
- L. S. Shapley, *A Value for n-person Games*. Contributions to the Theory of Games 2.28 (1953): 307-317.
- C. Molnar, *Interpretable Machine Learning*. Lulu, 2020. [Link]
- S. M. Lundberg and S.-I. Lee, *A unified approach to interpreting model predictions*, in Advances in Neural Information Processing Systems (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et al., eds.), vol. 30, pp. 4765–4774, Curran Associates, Inc., 2017. arXiv:1705.07874.
- S. M. Lundberg, G. G. Erion and S.-I. Lee, Consistent Individualized Feature Attribution for Tree Ensembles, arXiv e-prints (Feb. 2018) , [arXiv:1802.03888].
- S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair et al., *From local explanations to global understanding with explainable AI for trees*, Nature Machine Intelligence 2 (2020) 56–67.

Interpretable Machine Learning

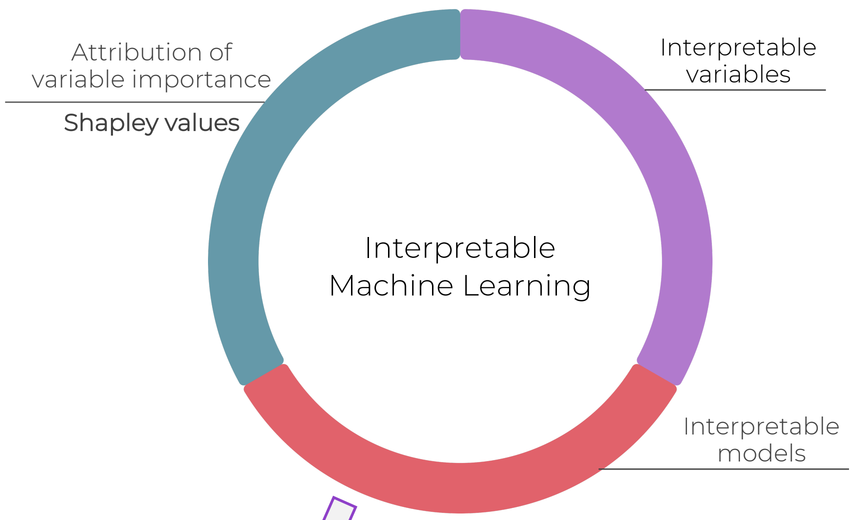
From Machine Learning to Human Understanding: turning black-boxes into Physics

the transition to interpretability

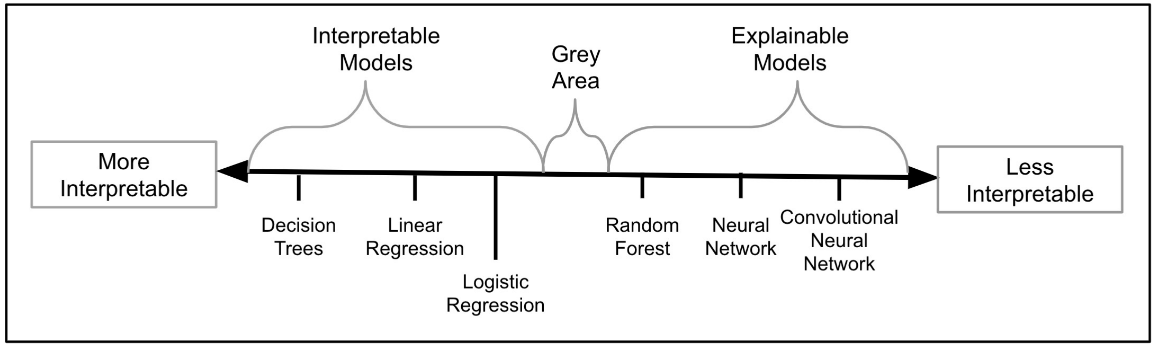




why explainable ML is not enough



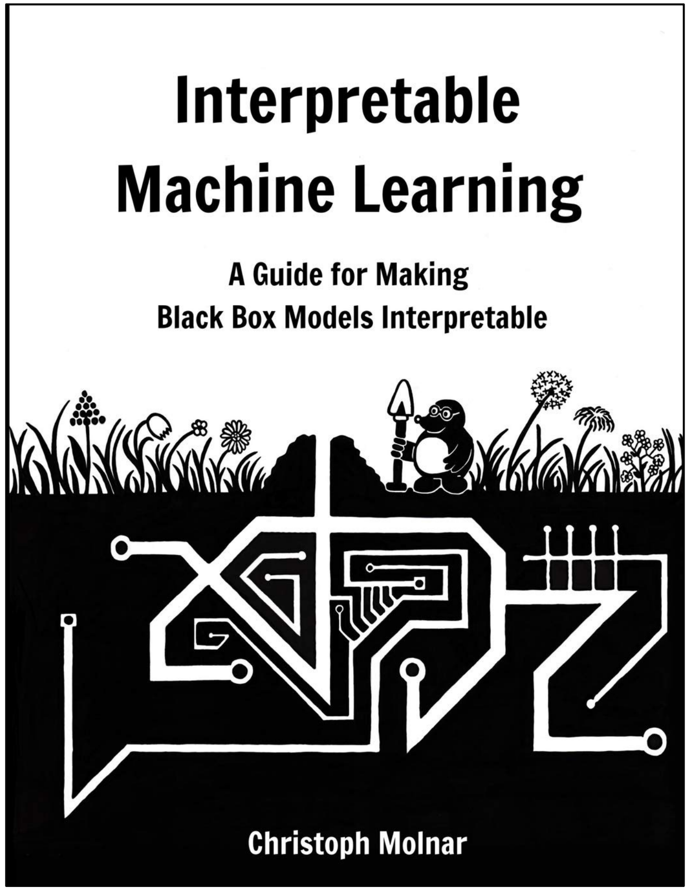
- Explainable models are not fully interpretable – proliferation of parameter can be a problem
- An interpretable model should be able to understandably map the input to the output
- Interpretability is important since an ML model should make the right decision for the right reasons



C. Rudin, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead.* Nature Machine Intelligence 1, 206–215 (2019)

how to make machine learning interpretable?

1. Read this book



2. Practice building interpretable models

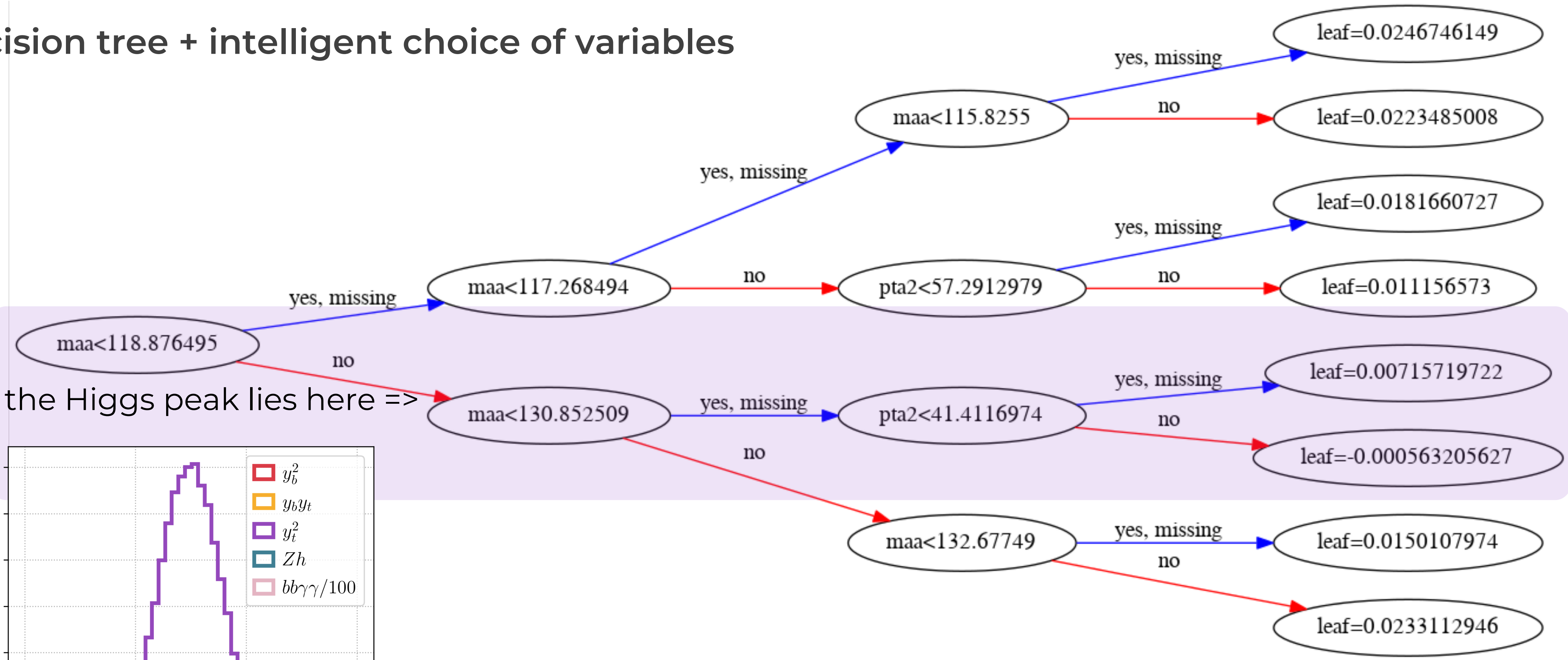
choose a variable set that is more interpretable

build a model that is interpretable

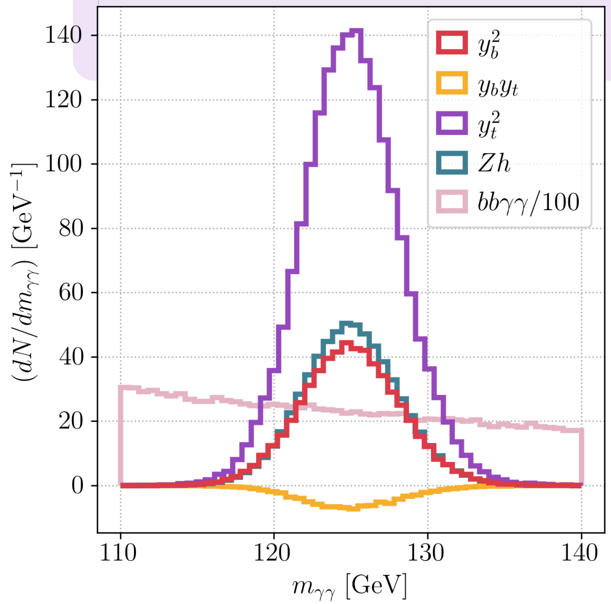
interpret how your variables are connected with the outcome

an interpretable model

Decision tree + intelligent choice of variables

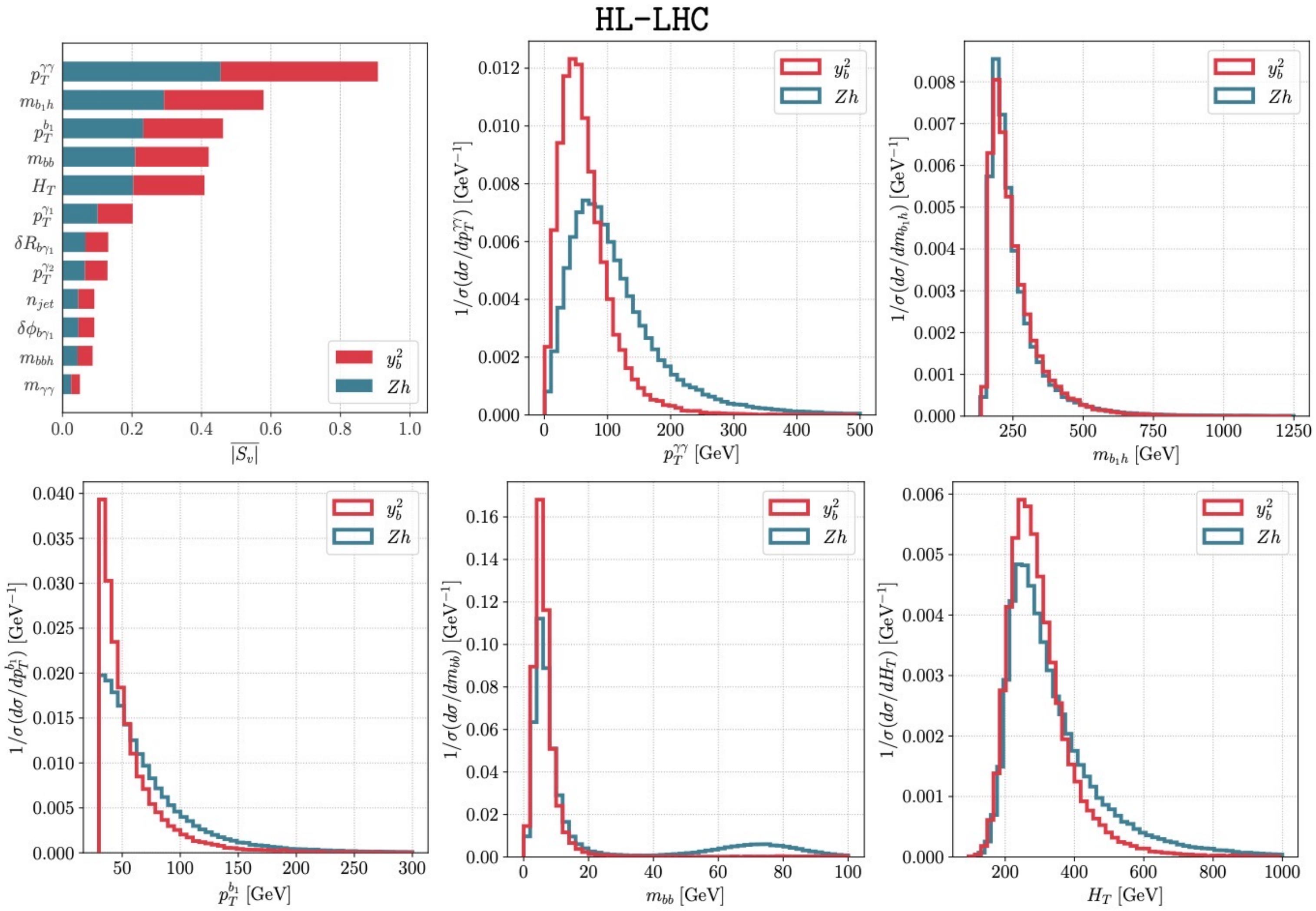


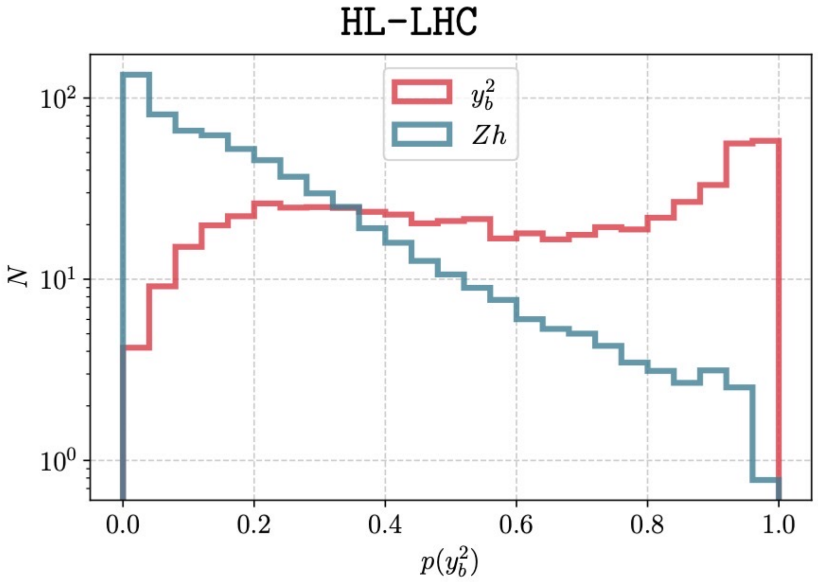
the Higgs peak lies here =>



NOTE: A different choice of variable basis might give higher accuracy, but you lose interpretability. (e.g., the four-momenta vectors of the final states)

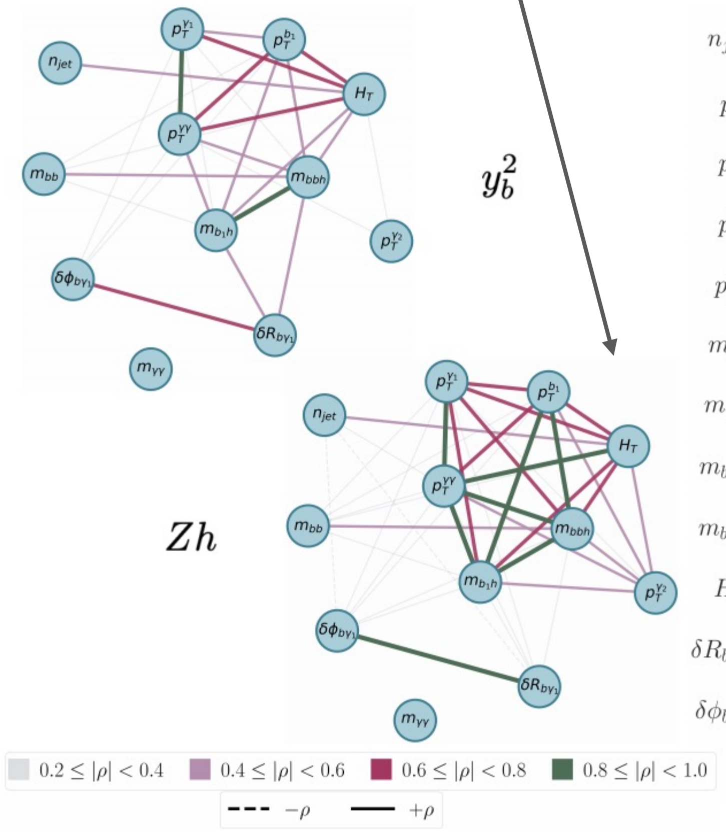
Zh $vs.$ y_b^2 discrimination

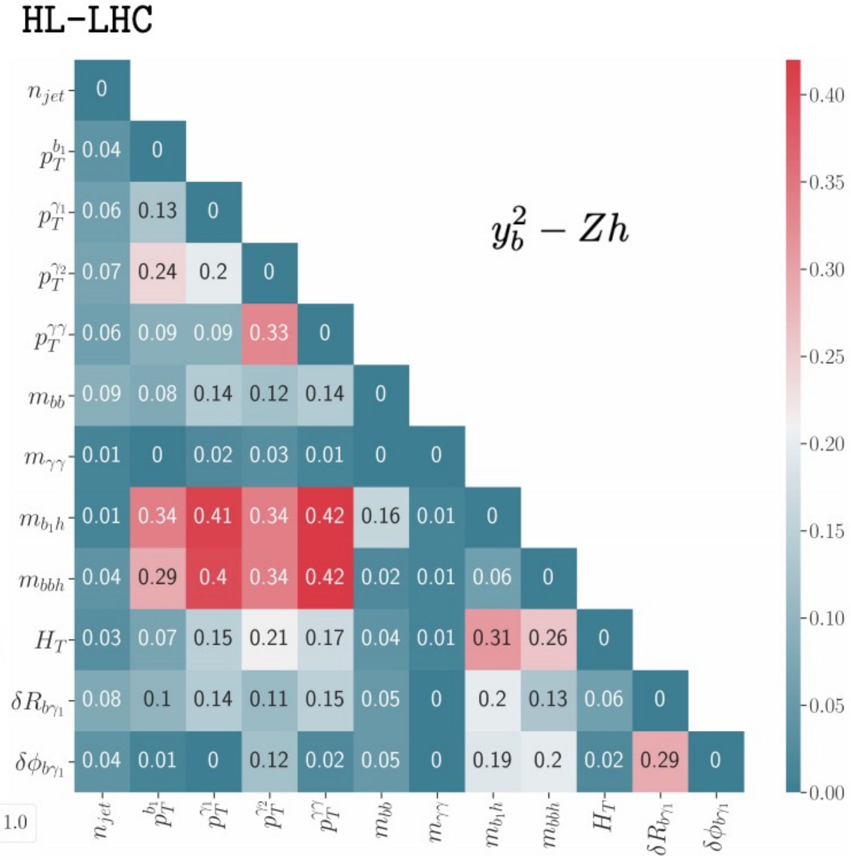


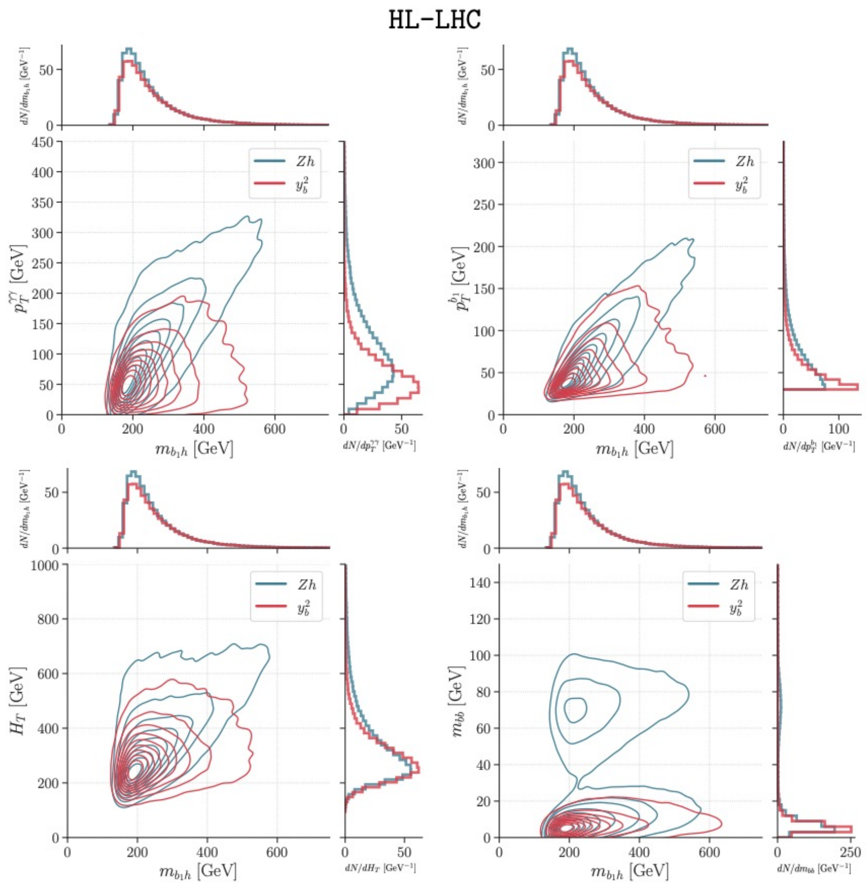


Zh $vs.$ y_b^2 discrimination

What really matters is the differences in the correlation between the two channels and this makes m_{b_1h} important



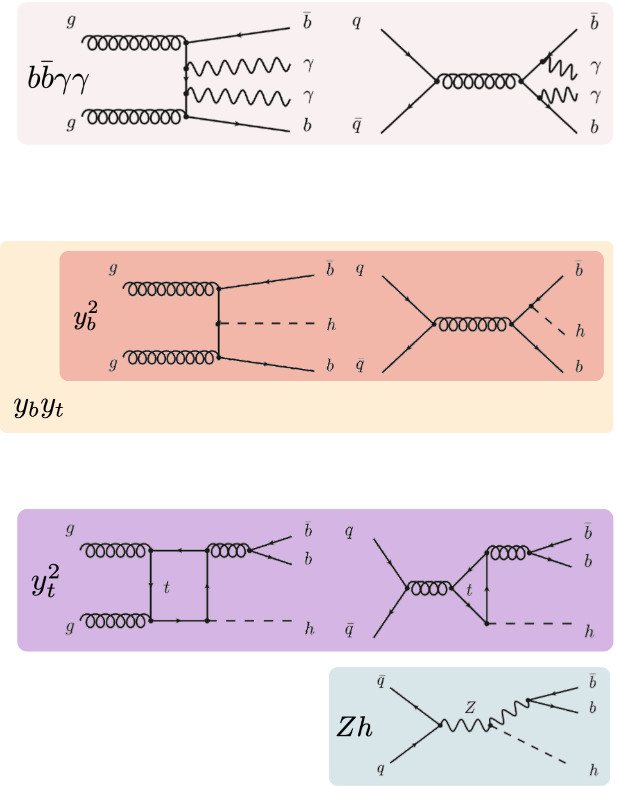




m_{b_1h} is highly correlated with some other kinematic variables that have variations in shape between the channels and hence it is a "team-player"

Interpretable ML for particle physics

Interpretable variables



QCD-QED Backgrounds

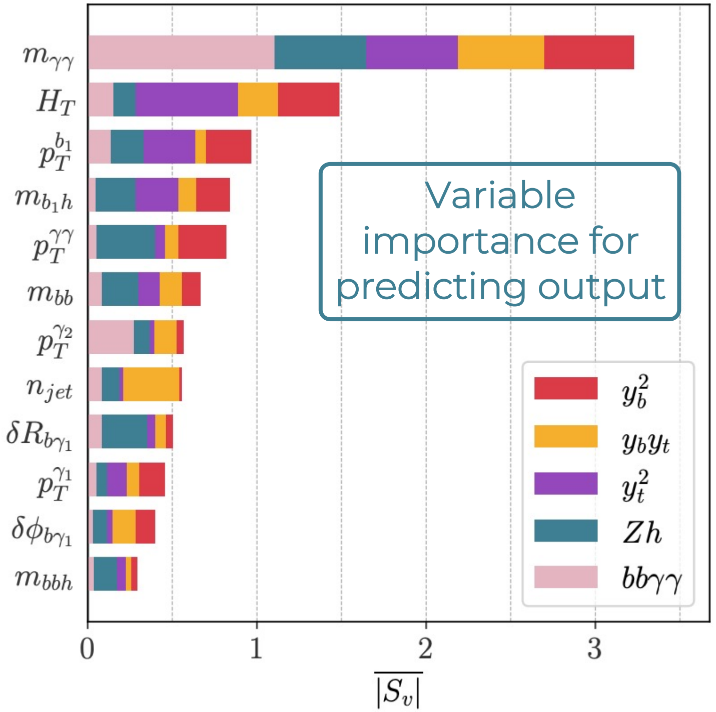
Signal

$b\bar{b}h$ backgrounds

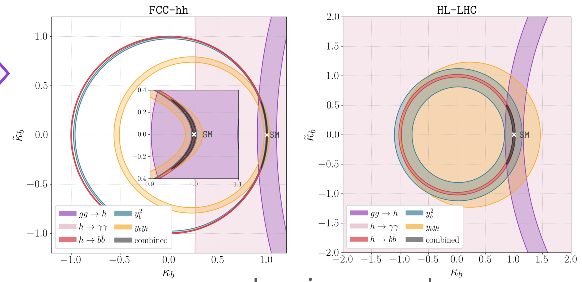
Shapley values + Physics Insights

Boosted Decision Trees + Signal Classification

Interpretable model



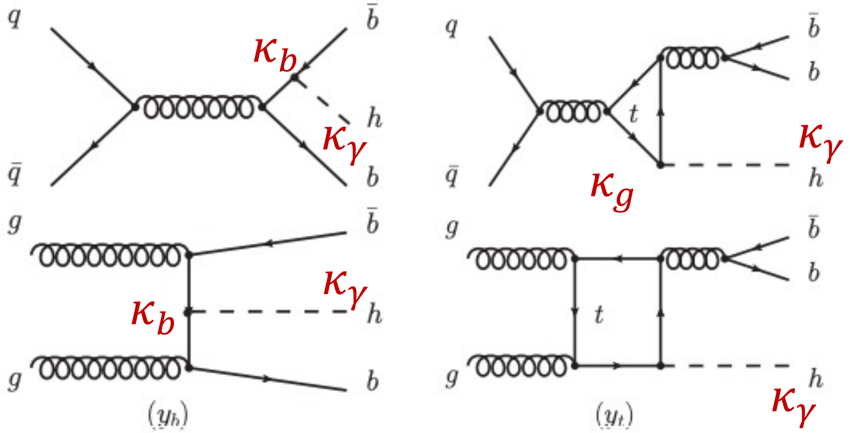
Variable importance for predicting output



new physics reach

- Gain of a factor of ~2 over traditional cut-based analysis
- Shapley values make the ML interpretable and allow for understanding the variable importance
- Separation of channels allow for simple rescaling to set bounds on anomalous couplings

results of the fit



Luminosities:

HL-LHC: 6 ab^{-1}

FCC-hh: 30 ab^{-1}

$$\mathcal{L} \supset -\frac{m_b}{v}(\kappa_b \bar{b} b + i\tilde{\kappa}_b \bar{b}\gamma_5 b)h$$

CP conserving — CP non-conserving

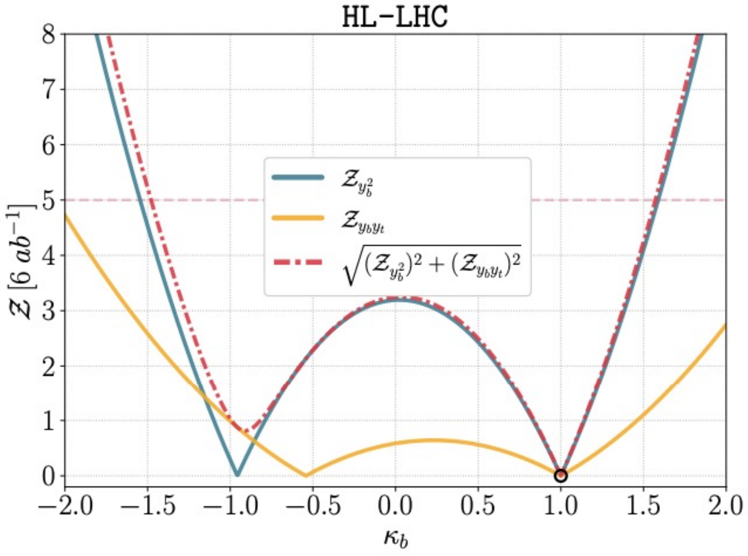
Direct bounds:

HL-LHC: 2.2%

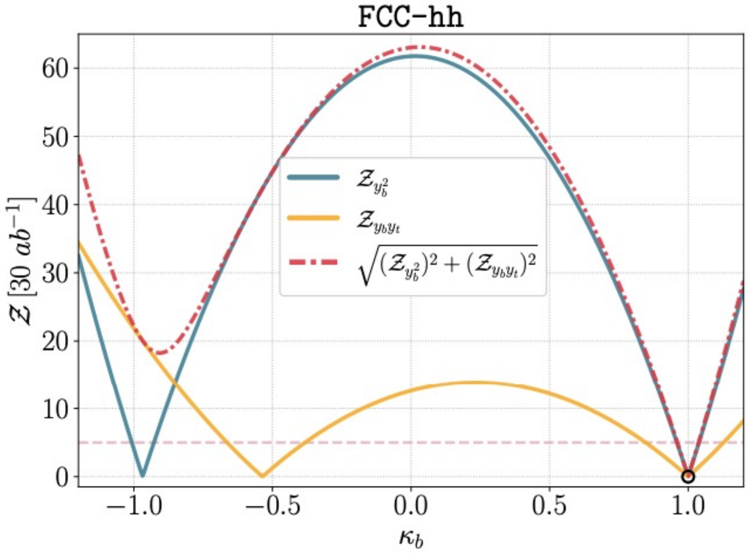
FCC-hh: 0.49%

κ_b generates κ_g and κ_γ through bottom-quark loops

$$\mathcal{Z}_j = \frac{|N_{jj}|}{\sqrt{\sum_i N_{ij}}} \quad [signal/\sqrt{signal + background}]$$



[-0.99,-0.82] ∪ [0.84,1.14]



[0.99,1.01]

Current measurement: ATLAS and CMS

$$\mu^{ATLAS}_{h\to b\bar{b}} = 1.01 \pm 0.12^{+0.16}_{-0.15},$$
$$\mu^{CMS}_{h\to b\bar{b}} = 1.04 \pm 0.14 \pm 0.14,$$

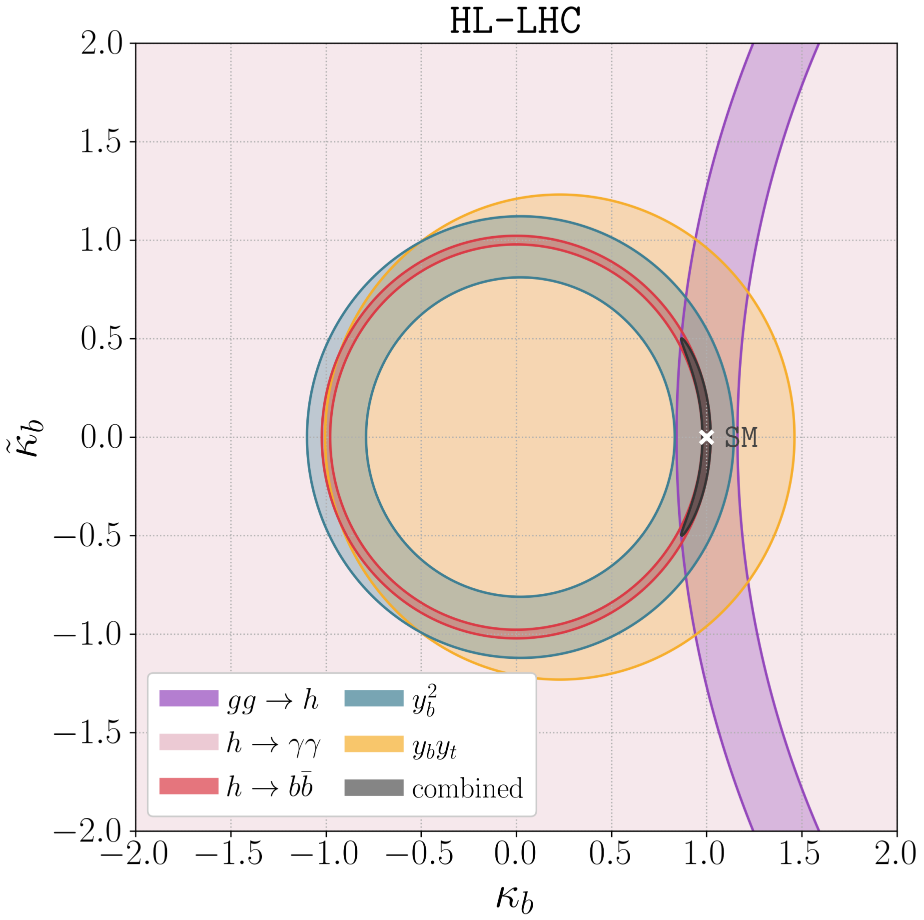
optimizing the analysis with machine learning

Predicted no. of events at HL-LHC (columns) / **Actual no. of events** (rows)

Channel	y_b^2	$y_b y_t$	y_t^2	Zh	$bb\gamma\gamma$	total
y_b^2	170	54	51	122	189	586
$y_b y_t$	-7	-24	-4	-20	-40	-95
y_t^2	238	112	452	546	487	1,835
Zh	22	28	21	416	161	648
$bb\gamma\gamma$	2,183	2,450	151	8,045	101,591	115,779
$\mathcal{Z}_j$	3.33	0.47	10.	4.36	317	

$$\mathcal{Z}_j = \frac{|N_{jj}|}{\sqrt{\sum_i N_{ij}}}$$

About ~60% gain in significance over traditional cut-based analyses (2σ).



$$\mathcal{L} \supset -\frac{m_b}{v}(\kappa_b \bar{b}b + i\tilde{\kappa}_b \bar{b}\gamma_5 b)h$$

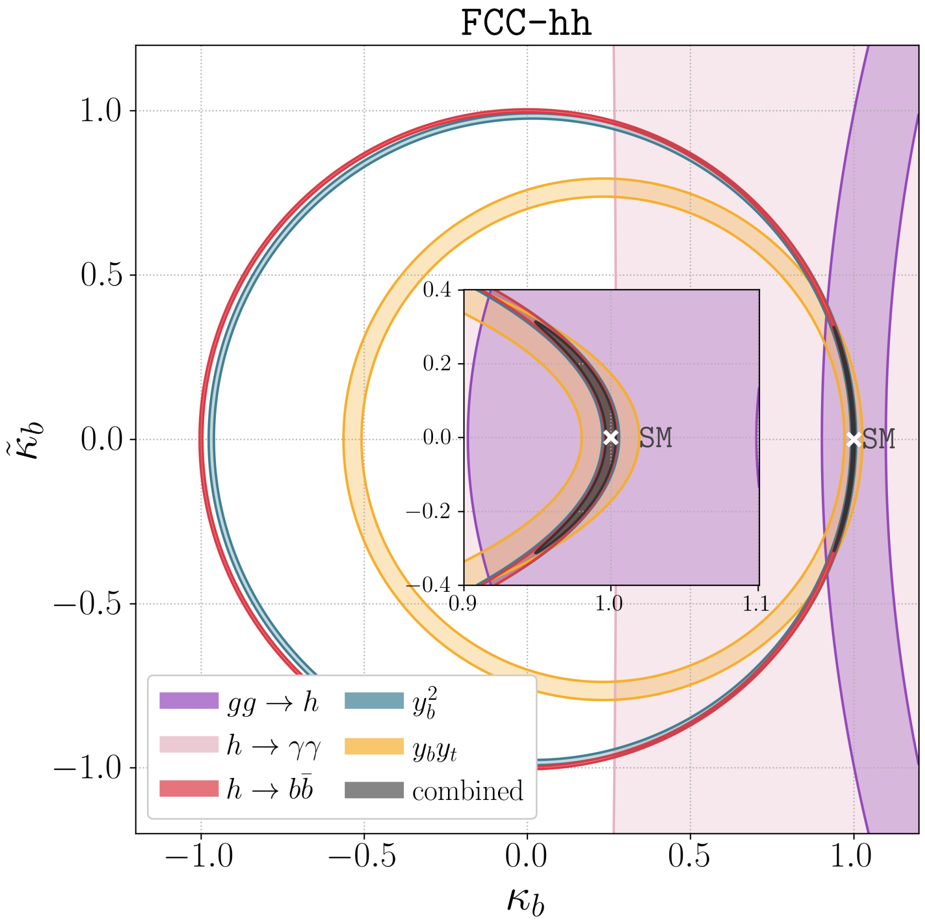
optimizing the analysis with machine learning

Predicted no. of events at FCC-hh

Actual no. of events	y_b^2	$y_b y_t$	y_t^2	Zh	$bb\gamma\gamma$	total
y_b^2	32,074	15,112	10,966	6,579	8,959	73,690
$y_b y_t$	-964	-6,815	-907	-583	-1,820	-11,089
y_t^2	48,772	45,751	148,669	39,598	26,484	309,274
Zh	1,860	4,498	2,280	12,661	2,282	23,581
$bb\gamma\gamma$	172,088	373,436	106,335	126,429	7,952,834	8,731,122
$\mathcal{Z}_j$	63.7	10.4	288	29.4	2,813	

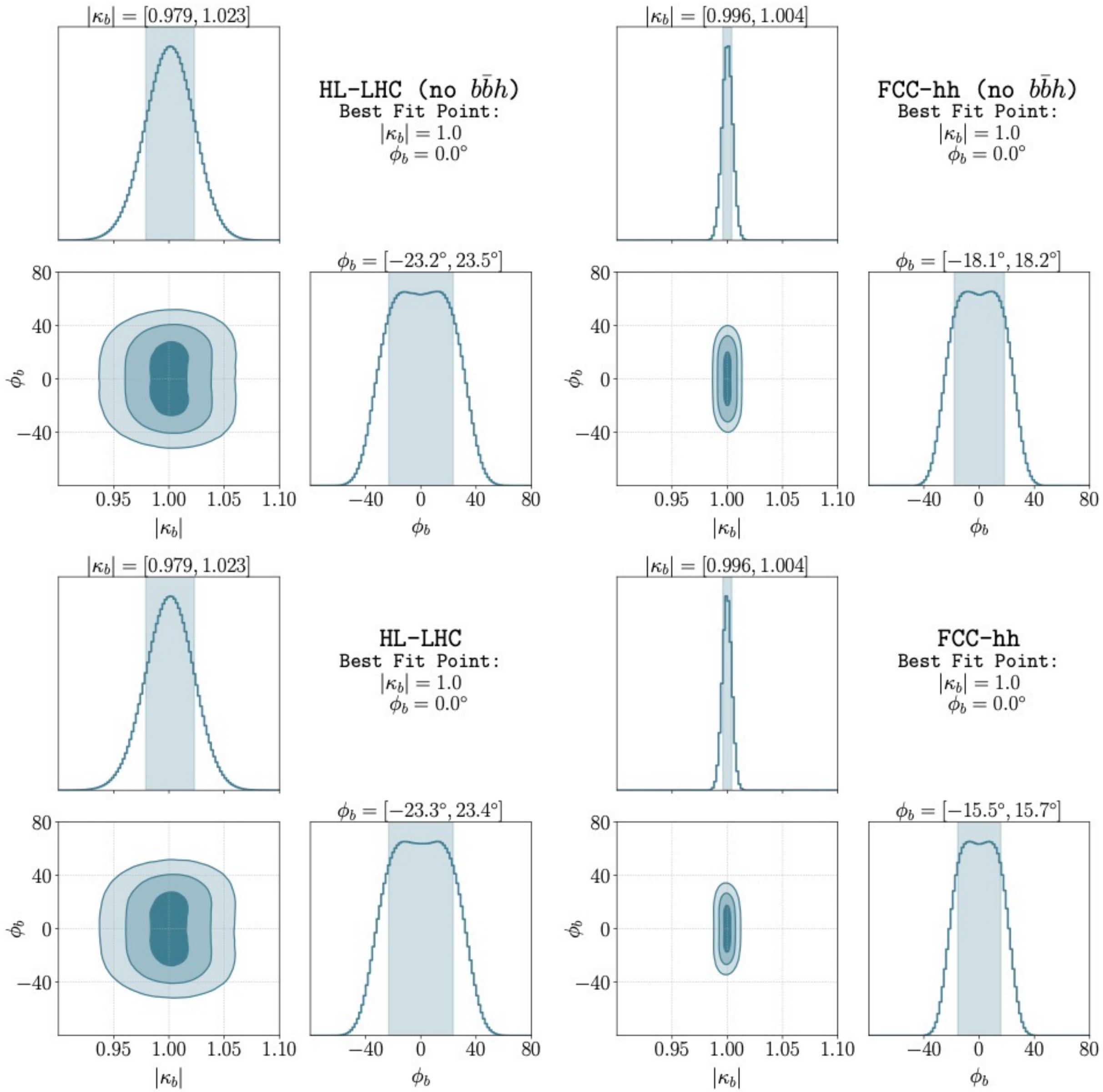
$$\mathcal{Z}_j = \frac{|N_{jj}|}{\sqrt{\sum_i N_{ij}}}$$

About ~60% gain in significance over traditional cut-based analyses.



$$\mathcal{L} \supset -\frac{m_b}{v}(\kappa_b \bar{b} b + i\tilde{\kappa}_b \bar{b}\gamma_5 b)h$$

results of fits



HL-LHC: $\phi_b = [-23.2°, 23.5°] \Rightarrow \tilde{\kappa}_b \lesssim 0.4$

Negligible improvement from including $b\bar{b}h$

FCC-hh: $\phi_b = [-15.5°, 15.7°] \Rightarrow \tilde{\kappa}_b \lesssim 0.3$

~15% improvement from including $b\bar{b}h$

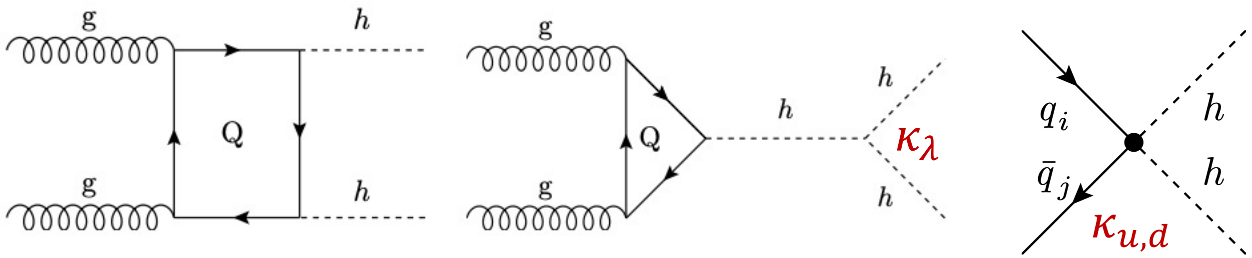
Comparison to:
Hadronic EDM (free of electron EDM assumption):

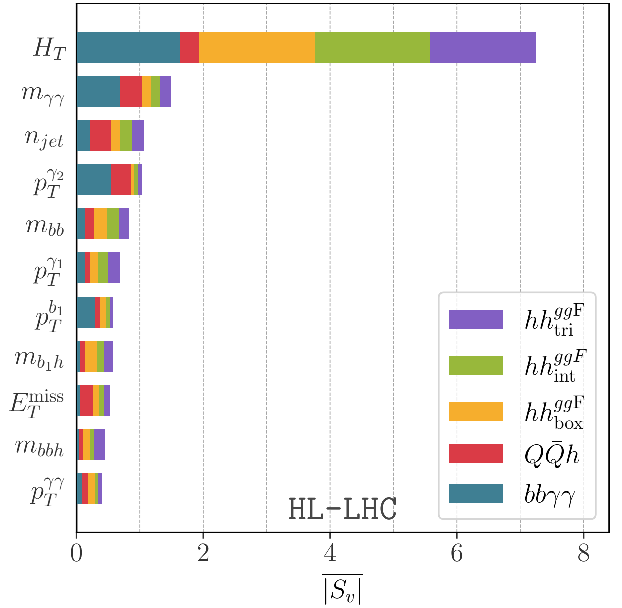
$$nEDM: \sum A\kappa_q\tilde{\kappa}_b + B\kappa_b\tilde{\kappa}_b \Rightarrow \tilde{\kappa}_b \lesssim 5$$

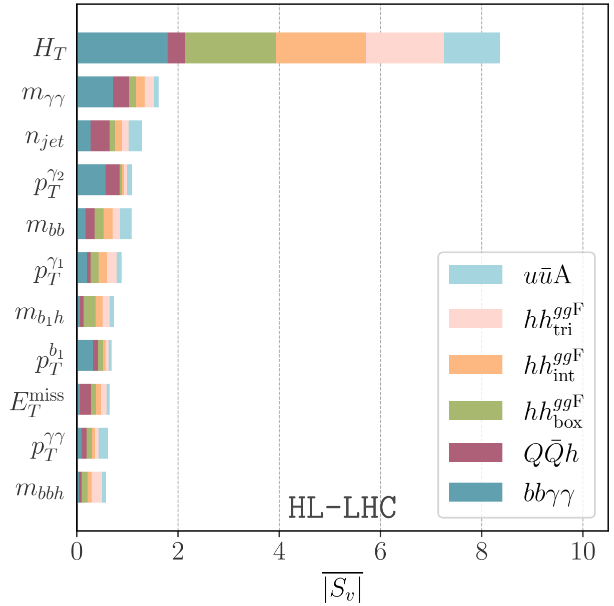
Electron EDM:

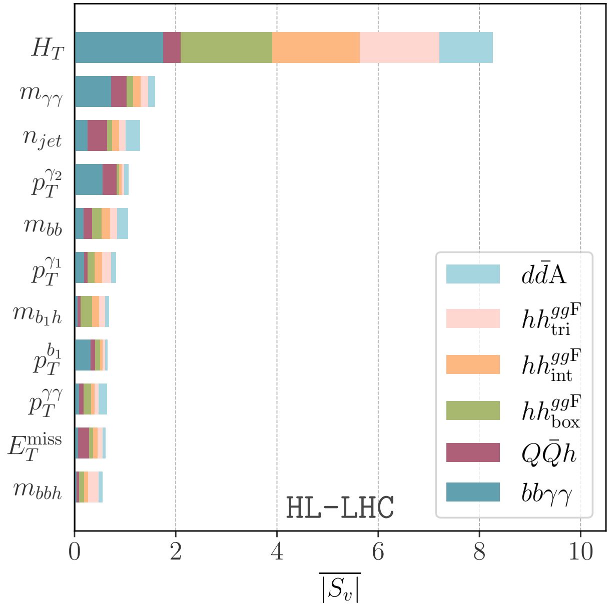
$$eEDM: \sum A\kappa_e\tilde{\kappa}_b + B\kappa_b\tilde{\kappa}_e \Rightarrow \tilde{\kappa}_b \lesssim 0.5$$

di-Higgs production at HL-LHC and FCC-hh









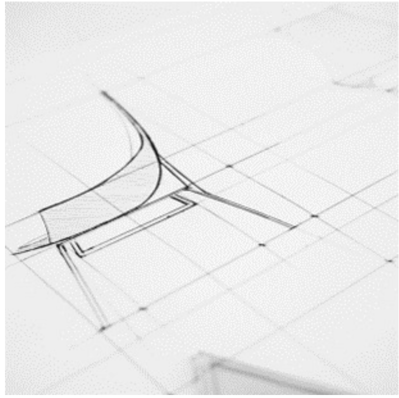
- Measurement of the trilinear can be made at HL-LHC possibly with more than 5σ significance
- Interestingly, light Yukawa measurements can also be made both from the production and decay rate modification
- The questions are:
  - How will the constraints change if the light Yukawa and trilinear couplings are measured simultaneously
  - What theoretical models can produce large light Yukawa modifications *and* Higgs trilinear modifications

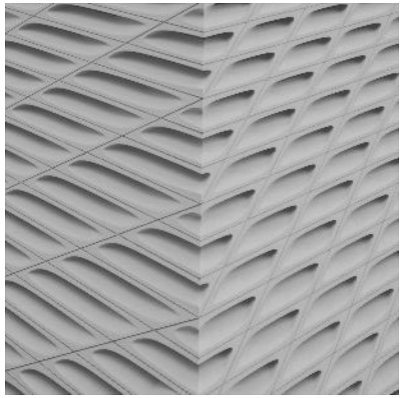
If you can't explain it simply, you don't understand it well enough.

- Albert Einstein

to a future of interpretability…

- Complex statistical machinery inherently reduce interpretability in terms of understandable physics
- Interpretability can be regained by intelligent modeling and using cooperative Game Theory
- We show the efficacy of this method by improving signal to background selection in $b\bar{b}h$
- Shapley values open a lot of possibilities for its application in Physics





Summary

the interpreters

Christophe Grojean

Th. Group Leader @ DESY, Hamburg
Professor @ Humboldt Universität zu Berlin

Ayan Paul

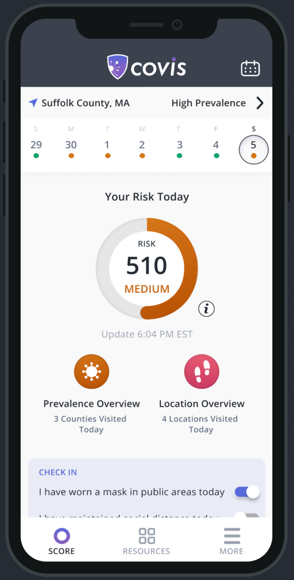
Fellow @ DESY, Hamburg &
Humboldt Universität zu Berlin

Zhuoni Qian

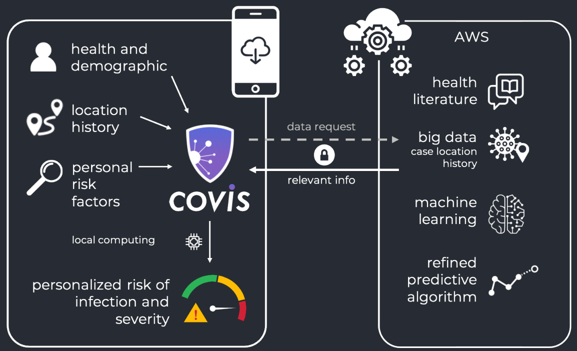
Fellow @ DESY, Hamburg &
Shandong University

Diversity@DESY-Theory

- **Diversity@DESY-Theory** was started in June 2020 by Postdocs and PhD students as consolidation of efforts for BLM throughout the international academic community.
- **Philosophy:** Diversity covers all kinds of variation in the academic world (gender, religion, etc.) and they should be accommodated for.
- A monthly remote meeting is held where different issues related to diversity are discussed which includes published articles and opinions.
- The Diversity Office of the Universe Cluster has been included with Eileen Schwanold providing expert advice on topics and actions.
- The plan is to form a core group of postdocs and students that can be approached by other members of DESY in case they want an unbiased discussion about any issues on diversity they might be facing.
- Topics related to diversity will be raised in workshops and conferences as a way of making people more aware of the core issues.
- Possible external outreach to other academic institutions to consolidate efforts of increasing inclusion in academia.
- Contact: Ayan Paul (apaul2@alumni.nd.edu)



COVID-19 risk assessment using public data, medical knowledge and AI algorithms



COVIS

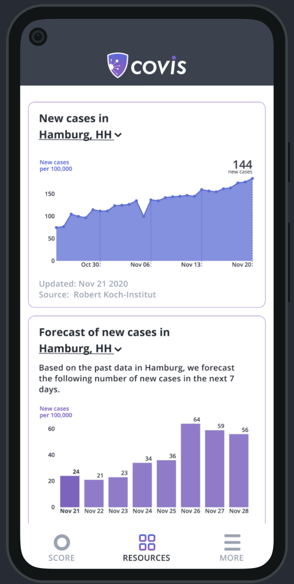
covishealth.com

Launched: August 2021
USA, Germany, Austria & Belgium
iOS & Android

Twice MIT COVID-19 hackathon winning solution implemented by a multidisciplinary team

User privacy and data security built into the architecture using federated computations

100,000€ in academic grants raised in the first 3 months



app features

app features

What CoVis brings to you:

- Real-time risk assessment as you go about your day.
- Next day prediction of COVID-19 cases globally.
- User score history and risk breakdown to understand what risks you have been facing.
- Global Information resources and international travel advisory at your fingertips.

What's your score?

Additional features include:

- Full score and history details beyond the past 14 days.
- 7 days of COVID-19 case rate prediction for all of Germany, USA, Austria and Belgium.
- User account to store score data (no personal data stored) for cross-device portability.
- Account and data management.

Developed With
Sidebench

CoVis - the future of disease management

Spin-off by DESY.

Thank You

Ayan Paul

DESY, Hamburg &

Humboldt Universität zu Berlin

www.desy.de/~apaul

apaul2@alumni.nd.edu

Back-up Slides

For the particularly curious

To my Mother and Father, who showed me what I could do,

and to Ikaros, who showed me what I could not.

"To know what no one else does, what a pleasure it can be!"

– adopted from the words of

Eugene Wigner.