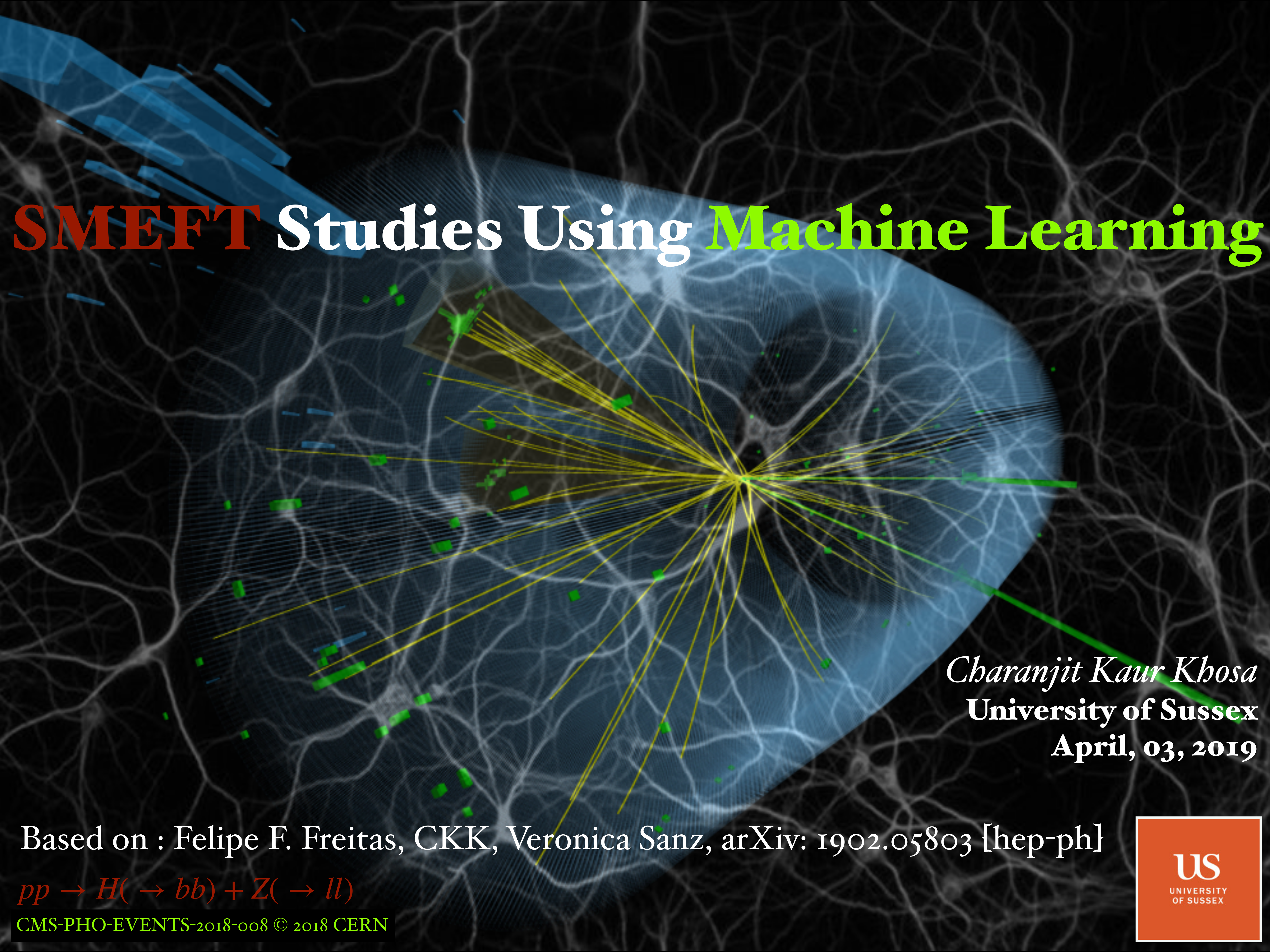


SMEFT Studies Using Machine Learning

Charanjit Kaur Khosa
University of Sussex
April, 03, 2019

Based on : Felipe F. Freitas, CKK, Veronica Sanz, arXiv: 1902.05803 [hep-ph]

$pp \to H(\to bb) + Z(\to ll)$

US UNIVERSITY OF SUSSEX

SM Effective Field Theory(SMEFT)

$$\mathscr{L}^{d=6} = \mathscr{L}_{SM} + \sum_i \frac{C_i}{\Lambda^2}\mathscr{O}_i$$

$$\begin{aligned}
\mathcal{L}^{\rm SILH}_{\rm SMEFT} \supset\ & \frac{\bar c_W}{m_W^2}\frac{ig}{2}\left(H^\dagger\sigma^a \overleftrightarrow{D}^\mu H\right) D^\nu W^a_{\mu\nu} + \frac{\bar c_B}{m_W^2}\frac{ig'}{2}\left(H^\dagger \overleftrightarrow{D}^\mu H\right)\partial^\nu B_{\mu\nu} + \frac{\bar c_T}{v^2}\frac{1}{2}\left(H^\dagger \overleftrightarrow{D}_\mu H\right)^2 \\
& + \frac{\bar c_{ll}}{v^2}(\bar L\gamma_\mu L)(\bar L\gamma^\mu L) + \frac{\bar c_{He}}{v^2}(iH^\dagger \overleftrightarrow{D}_\mu H)(\bar e_R\gamma^\mu e_R) + \frac{\bar c_{Hu}}{v^2}(iH^\dagger \overleftrightarrow{D}_\mu H)(\bar u_R\gamma^\mu u_R) \\
& + \frac{\bar c_{Hd}}{v^2}(iH^\dagger \overleftrightarrow{D}_\mu H)(\bar d_R\gamma^\mu d_R) + \frac{\bar c'_{Hq}}{v^2}(iH^\dagger\sigma^a \overleftrightarrow{D}_\mu H)(\bar Q_L\sigma^a\gamma^\mu Q_L) \\
& + \frac{\bar c_{Hq}}{v^2}(iH^\dagger \overleftrightarrow{D}_\mu H)(\bar Q_L\gamma^\mu Q_L) + \frac{\bar c_{HW}}{m_W^2} ig (D^\mu H)^\dagger\sigma^a(D^\nu H)W^a_{\mu\nu} + \frac{\bar c_{HB}}{m_W^2} ig'(D^\mu H)^\dagger(D^\nu H)B_{\mu\nu} \\
& + \frac{\bar c_{3W}}{m_W^2} g^3\epsilon_{abc}W^{a\,\nu}_{\mu} W^b_{\nu\rho}W^{c\,\rho\mu} + \frac{\bar c_g}{m_W^2} g_s^2|H|^2 G^A_{\mu\nu}G^{A\mu\nu} + \frac{\bar c_\gamma}{m_W^2} g'^2|H|^2 B_{\mu\nu}B^{\mu\nu} \\
& + \frac{\bar c_H}{v^2}\frac{1}{2}(\partial^\mu|H|^2)^2 + \sum_{f=e,u,d}\frac{\bar c_f}{v^2} y_f|H|^2\bar F_L H^{(c)} f_R \\
& + \frac{\bar c_{3G}}{m_W^2} g_s^3 f_{ABC}G^{A\nu}_{\mu} G^{B\rho}_{\nu} G^{C\mu}_{\rho} + \frac{\bar c_{uG}}{m_W^2} g_s y_u \bar Q_L H^{(c)}\sigma^{\mu\nu}\lambda_A u_R G^A_{\mu\nu}
\end{aligned}$$

SMEFT : Global Analysis

- Precision electroweak data, LHC Run 1 & 2 data (Higgs production, pair of gauge bosons)

Observable	Measurement	Ref.	SM Prediction	Ref.
Γ_Z [GeV]	2.4952 ± 0.0023	[41]	2.4943 ± 0.0005	[40]
σ^0_{had} [nb]	41.540 ± 0.037	[41]	41.488 ± 0.006	[40]
R^0_ℓ	20.767 ± 0.025	[41]	20.752 ± 0.005	[40]
$A^{0,\ell}_{\text{FB}}$	0.0171 ± 0.0010	[41]	0.01622 ± 0.00009	[120]
$\mathcal{A}_\ell\,(P_\tau)$	0.1465 ± 0.0033	[41]	0.1470 ± 0.0004	[120]
$\mathcal{A}_\ell$ (SLD)	0.1513 ± 0.0021	[41]	0.1470 ± 0.0004	[120]
R^0_b	0.021629 ± 0.00066	[41]	0.2158 ± 0.00015	[40]
R^0_c	0.1721 ± 0.0030	[41]	0.17223 ± 0.00005	[40]
$A^{0,b}_{\text{FB}}$	0.0992 ± 0.0016	[41]	0.1031 ± 0.0003	[120]
$A^{0,c}_{\text{FB}}$	0.0707 ± 0.0035	[41]	0.0736 ± 0.0002	[120]
$\mathcal{A}_b$	0.923 ± 0.020	[41]	0.9347	[120]
$\mathcal{A}_c$	0.670 ± 0.027	[41]	0.6678 ± 0.0002	[120]
M_W [GeV]	80.387 ± 0.016	[42]	80.361 ± 0.006	[120]
M_W [GeV]	80.370 ± 0.019	[100]	80.361 ± 0.006	[120]

LEP data

+

WW LEP2 data

+

M_W Tevatron

J.Ellis, C.W.Murphy, V.Sanz and T.You JHEP06(2018)146

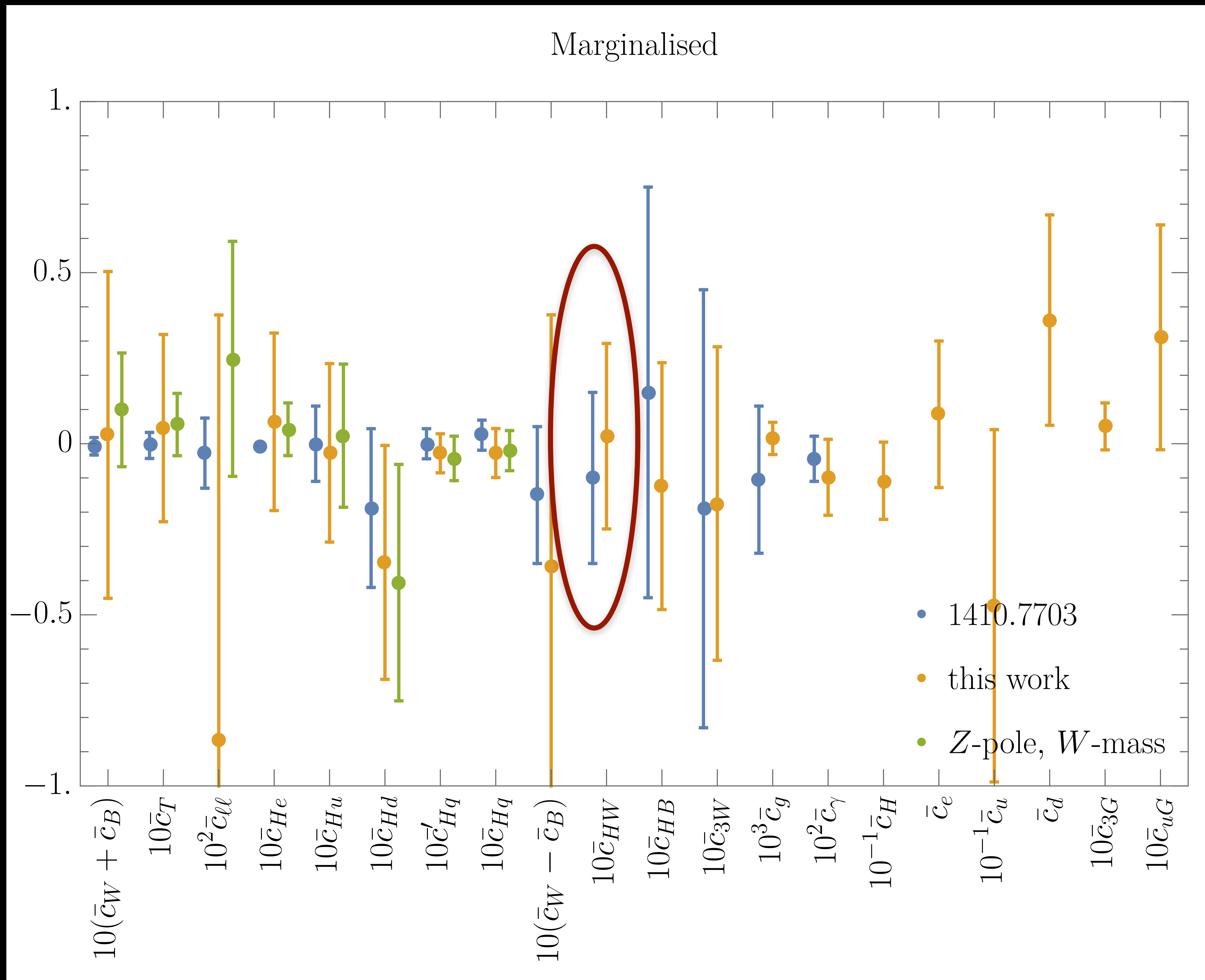
Run 1 data

Production	Decay	Signal Strength	Production	Decay	Signal Strength
ggF	$\gamma\gamma$	$1.10^{+0.23}_{-0.22}$	Wh	$\tau\tau$	-1.4 ± 1.4
ggF	ZZ	$1.13^{+0.34}_{-0.31}$	Wh	bb	1.0 ± 0.5
ggF	WW	0.84 ± 0.17	Zh	$\gamma\gamma$	$0.5^{+3.0}_{-2.5}$
ggF	$\tau\tau$	1.0 ± 0.6	Zh	WW	$5.9^{+2.6}_{-2.2}$
VBF	$\gamma\gamma$	1.3 ± 0.5	Zh	$\tau\tau$	$2.2^{+2.2}_{-1.8}$
VBF	ZZ	$0.1^{+1.1}_{-0.6}$	Zh	bb	0.4 ± 0.4
VBF	WW	1.2 ± 0.4	tth	$\gamma\gamma$	$2.2^{+1.6}_{-1.3}$
VBF	$\tau\tau$	1.3 ± 0.4	tth	WW	$5.0^{+1.8}_{-1.7}$
Wh	$\gamma\gamma$	$0.5^{+1.3}_{-1.2}$	tth	$\tau\tau$	$-1.9^{+3.7}_{-3.3}$
Wh	WW	$1.6^{+1.2}_{-1.0}$	tth	bb	1.1 ± 1.0
pp	$Z\gamma$	$2.7^{+4.6}_{-4.5}$	pp	$\mu\mu$	0.1 ± 2.5

(Early) Run 2 data
+
STXS

	Production	Decay	Sig. Stren.		Production	Decay	Sig. Stren.
[102]	1-jet, $p_T > 450$	$b\bar{b}$	$2.3^{+1.8}_{-1.6}$	[110]	pp	$\mu\mu$	-0.1 ± 1.5
[103]	Zh	$b\bar{b}$	0.9 ± 0.5	[111]	Zh	$b\bar{b}$	$1.12^{+0.50}_{-0.45}$
[103]	Wh	$b\bar{b}$	1.7 ± 0.7	[111]	Wh	$b\bar{b}$	$1.35^{+0.68}_{-0.59}$
[104]	$t\bar{t}h, \geq 1\ell$	$b\bar{b}$	0.72 ± 0.45	[112]	$t\bar{t}h$	$b\bar{b}$	$0.84^{+0.64}_{-0.61}$
[105]	$t\bar{t}h$	$1\ell + 2\tau_h$	$-1.52^{+1.76}_{-1.72}$	[113]	$t\bar{t}h$	$2\ell os + 1\tau_h$	$1.7^{+2.1}_{-1.9}$
[105]	$t\bar{t}h$	$2\ell ss + 1\tau_h$	$0.94^{+0.80}_{-0.67}$	[113]	$t\bar{t}h$	$1\ell + 2\tau_h$	$-0.6^{+1.6}_{-1.5}$
[105]	$t\bar{t}h$	$3\ell + 1\tau_h$	$1.34^{+1.42}_{-1.07}$	[113]	$t\bar{t}h$	$3\ell + 1\tau_h$	$1.6^{+1.8}_{-1.3}$
[105]	$t\bar{t}h$	$2\ell ss$	$1.61^{+0.58}_{-0.51}$	[113]	$t\bar{t}h$	$2\ell ss + 1\tau_h$	$3.5^{+1.7}_{-1.3}$
[105]	$t\bar{t}h$	3ℓ	$0.82^{+0.77}_{-0.71}$	[113]	$t\bar{t}h$	3ℓ	$1.8^{+0.9}_{-0.7}$
[105]	$t\bar{t}h$	4ℓ	$0.9^{+2.3}_{-1.6}$	[113]	$t\bar{t}h$	$2\ell ss$	$1.5^{+0.7}_{-0.6}$
[106]	0-jet DF	WW	$1.30^{+0.24}_{-0.23}$	[114]	ggF	WW	$1.21^{+0.22}_{-0.21}$
[106]	1-jet DF	WW	$1.29^{+0.32}_{-0.27}$	[114]	VBF	WW	$0.62^{+0.37}_{-0.36}$
[106]	2-jet DF	WW	$0.82^{+0.54}_{-0.50}$	[115]	$B(h \to \gamma\gamma)/B(h \to 4\ell)$		$0.69^{+0.15}_{-0.13}$
[106]	VBF 2-jet	WW	$0.72^{+0.44}_{-0.41}$	[115]	0-jet	4ℓ	$1.07^{+0.27}_{-0.25}$
[106]	Vh 2-jet	WW	$3.92^{+1.32}_{-1.17}$	[115]	1-jet, $p_T < 60$	4ℓ	$0.67^{+0.72}_{-0.68}$
[106]	Wh 3-lep	WW	$2.23^{+1.76}_{-1.53}$	[115]	1-jet, $p_T \in (60, 120)$	4ℓ	$1.00^{+0.63}_{-0.55}$
[107]	ggF	$\gamma\gamma$	$1.10^{+0.20}_{-0.18}$	[115]	1-jet, $p_T \in (120, 200)$	4ℓ	$2.1^{+1.5}_{-1.3}$
[107]	VBF	$\gamma\gamma$	$0.8^{+0.6}_{-0.5}$	[115]	2-jet	4ℓ	$2.2^{+1.1}_{-1.0}$
[107]	$t\bar{t}h$	$\gamma\gamma$	$2.2^{+0.9}_{-0.8}$	[115]	"BSM-like"	4ℓ	$2.3^{+1.2}_{-1.0}$
[107]	Vh	$\gamma\gamma$	$2.4^{+1.1}_{-1.0}$	[115]	VBF, $p_T < 200$	4ℓ	$2.14^{+0.94}_{-0.77}$
[108]	ggF	4ℓ	$1.20^{+0.22}_{-0.21}$	[115]	Vh lep	4ℓ	$0.3^{+1.3}_{-1.2}$
[109]	0-jet	$\tau\tau$	0.84 ± 0.89	[115]	$t\bar{t}h$	4ℓ	$0.51^{+0.86}_{-0.70}$
[109]	boosted	$\tau\tau$	$1.17^{+0.47}_{-0.40}$	[116]	Wh	WW	$3.2^{+4.4}_{-4.2}$
[109]	VBF	$\tau\tau$	$1.11^{+0.34}_{-0.35}$				
[106]	Zh 4-lep	WW	$0.77^{+1.49}_{-1.20}$				

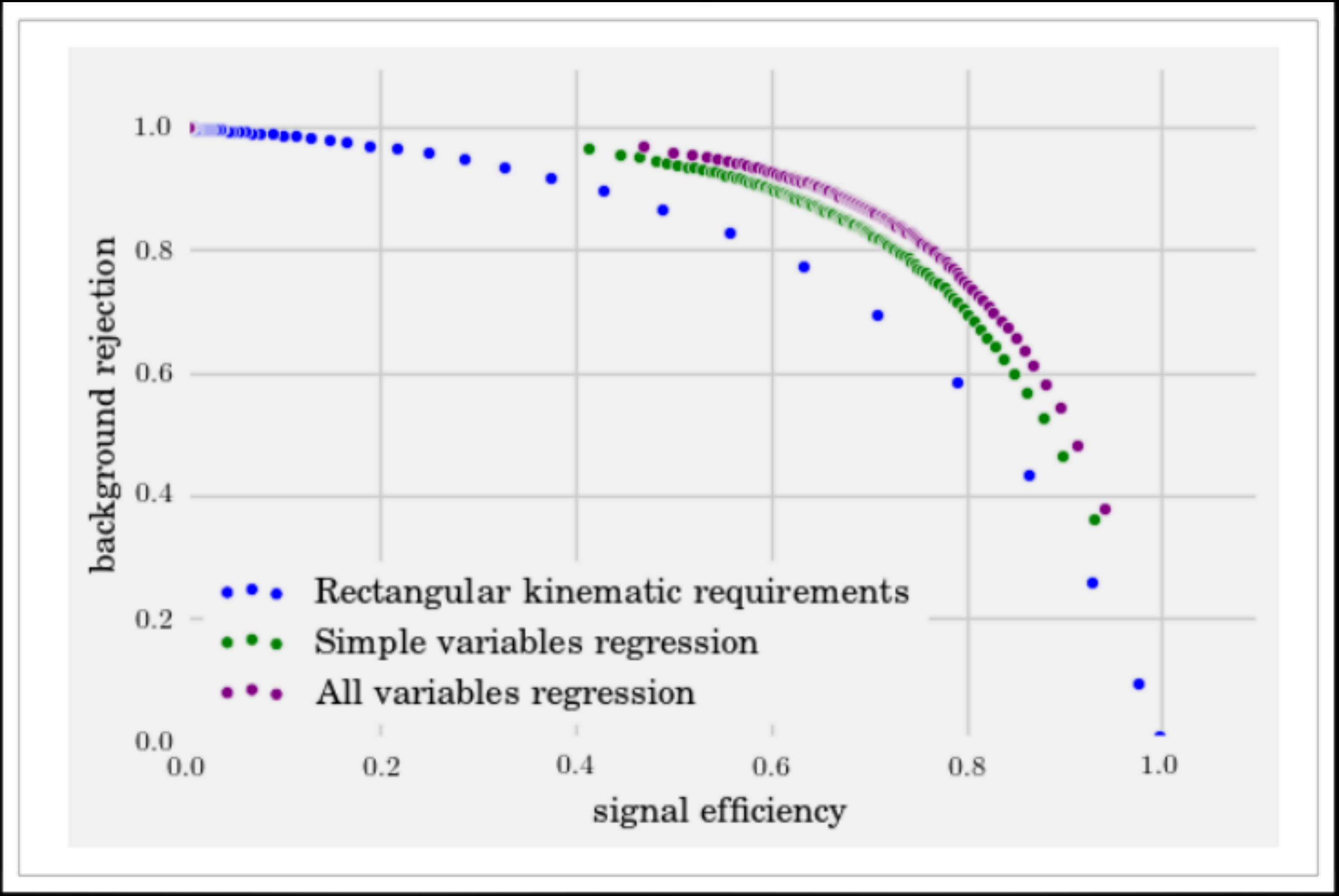
J.Ellis, C.W.Murphy, V.Sanz and T.You JHEP06(2018)146



More recent analysis : A.Biekötter, T.Corbett and T.Plehn, arXiv:1812.07587 [hep-ph]

Why Machine Learning is required for HEP analyses?

Classification between signal and background process



Need to quickly identify subtle effects in multidimensional distributions

P. Mehta et al., arXiv:1803.08823[physics.comp-ph]

SMEFT analyses using ML

VBF channel (not a reality yet)

J. Brehmer, K. Cranmer, G. Louppe and J Pavez,
Phys. Rev. Lett. 121 (2018) 111801, Phys. Rev. D 98 (2018) 052004

VH channel (data available)

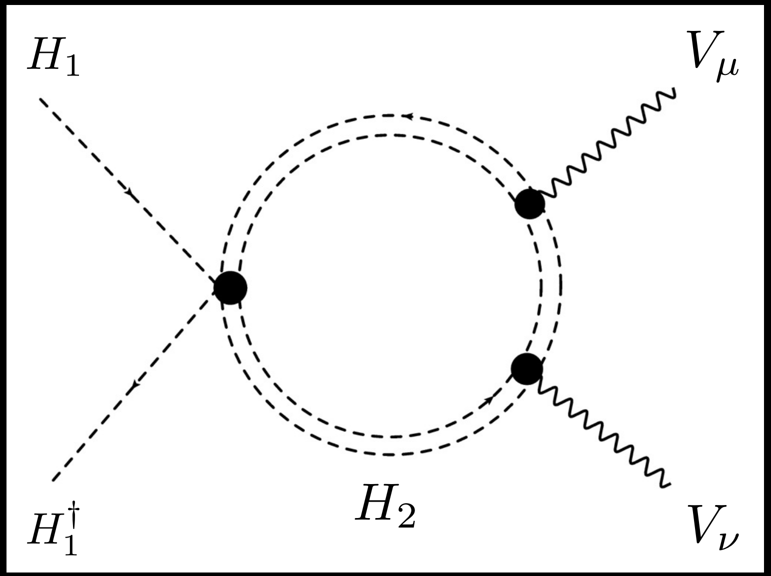
Felipe F. Freitas, CKK, Veronica Sanz, arXiv: 1902.05803 [hep-ph]

Our approach is to use a better algorithm to analyse the data which is available now

More subtle kinematic signatures
Higgs EFT via VH channel

Generic EFT effect

$$\eta_{\mu\nu}\, g m_V \Rightarrow \eta_{\mu\nu}\, g m_V - \frac{2\, g\, c_{HW}}{m_W}\, p_\mu^V\, p_\nu^V + \dots$$



$$\mathscr{L}_{SM} \Rightarrow \mathscr{L}_{SM} + \frac{2igc_{HW}}{m_W^2}\left[D^\mu H^\dagger T_{2k} D^\nu H\right] W_{\mu\nu}^k + \dots$$

Specific operator which produces it

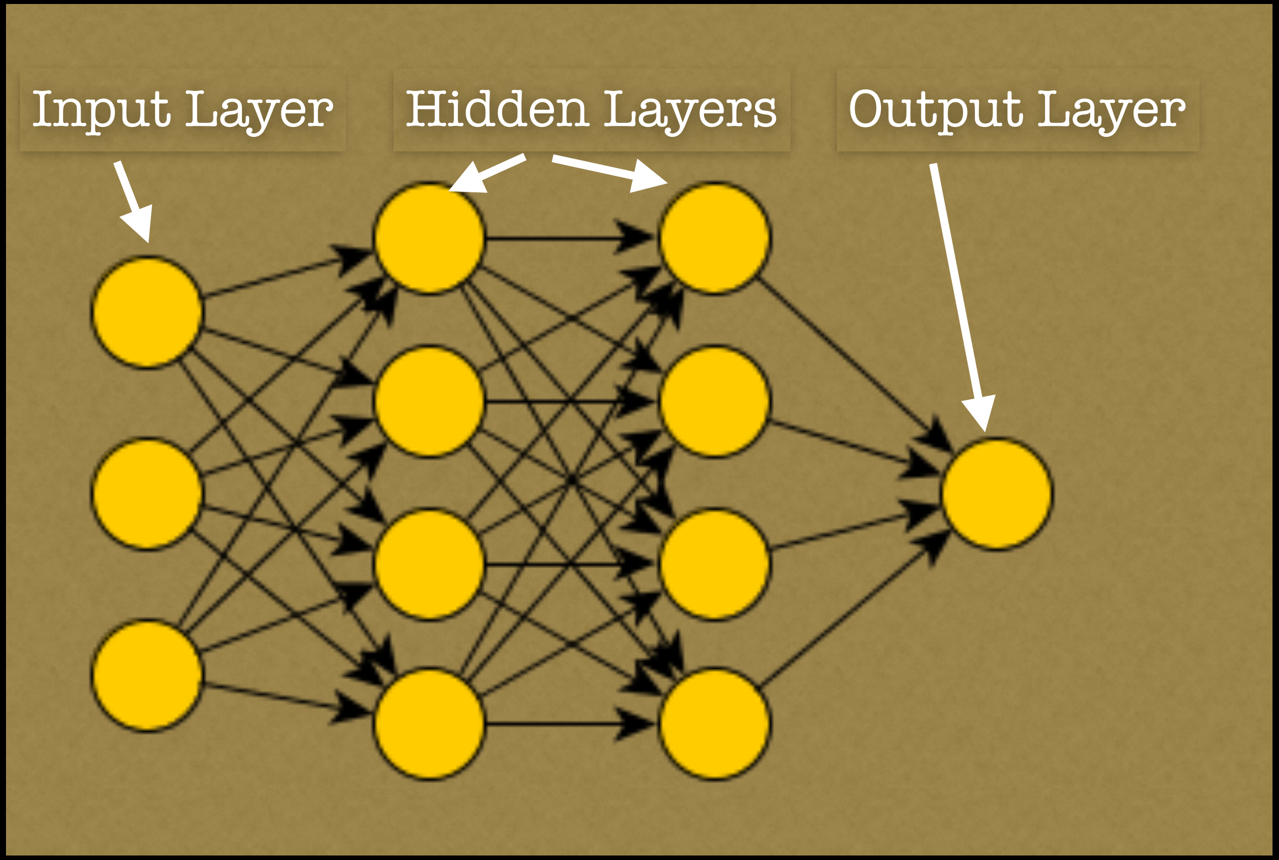
Felipe F. Freitas, CKK, Veronica Sanz, arXiv: 1902.05803 [hep-ph]

Broad Categories of ML

- Supervised Learning (labelled data)
- Unsupervised Learning (finding patterns and structure)
- Reinforcement Learning (interaction with an environment)

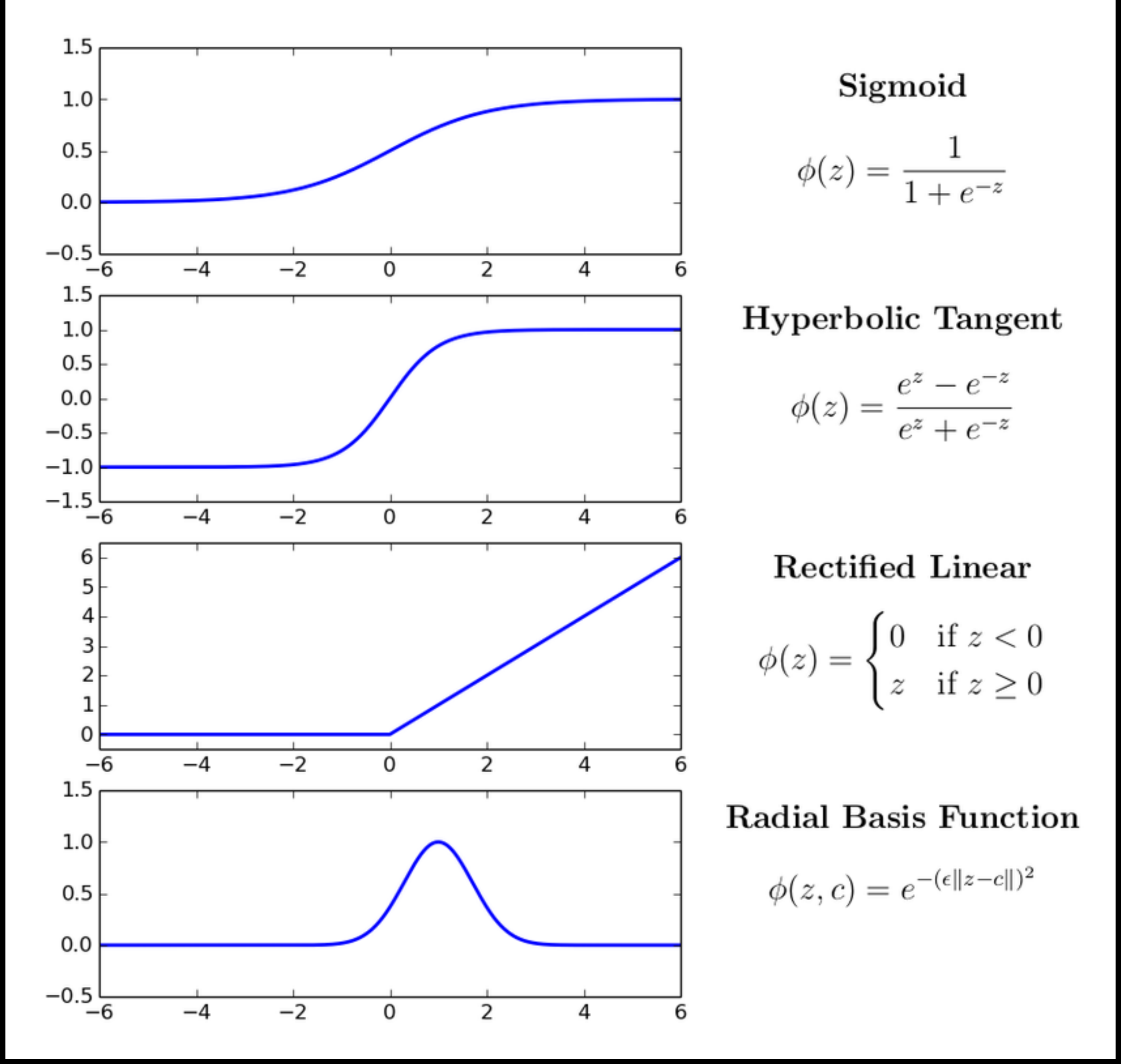
Neural Network : basic structure

- Input layer nodes : set of observables(kinematical features)
- Number of hidden layers (shallow or deep NN)
- Output layer : binary/multi-class classification

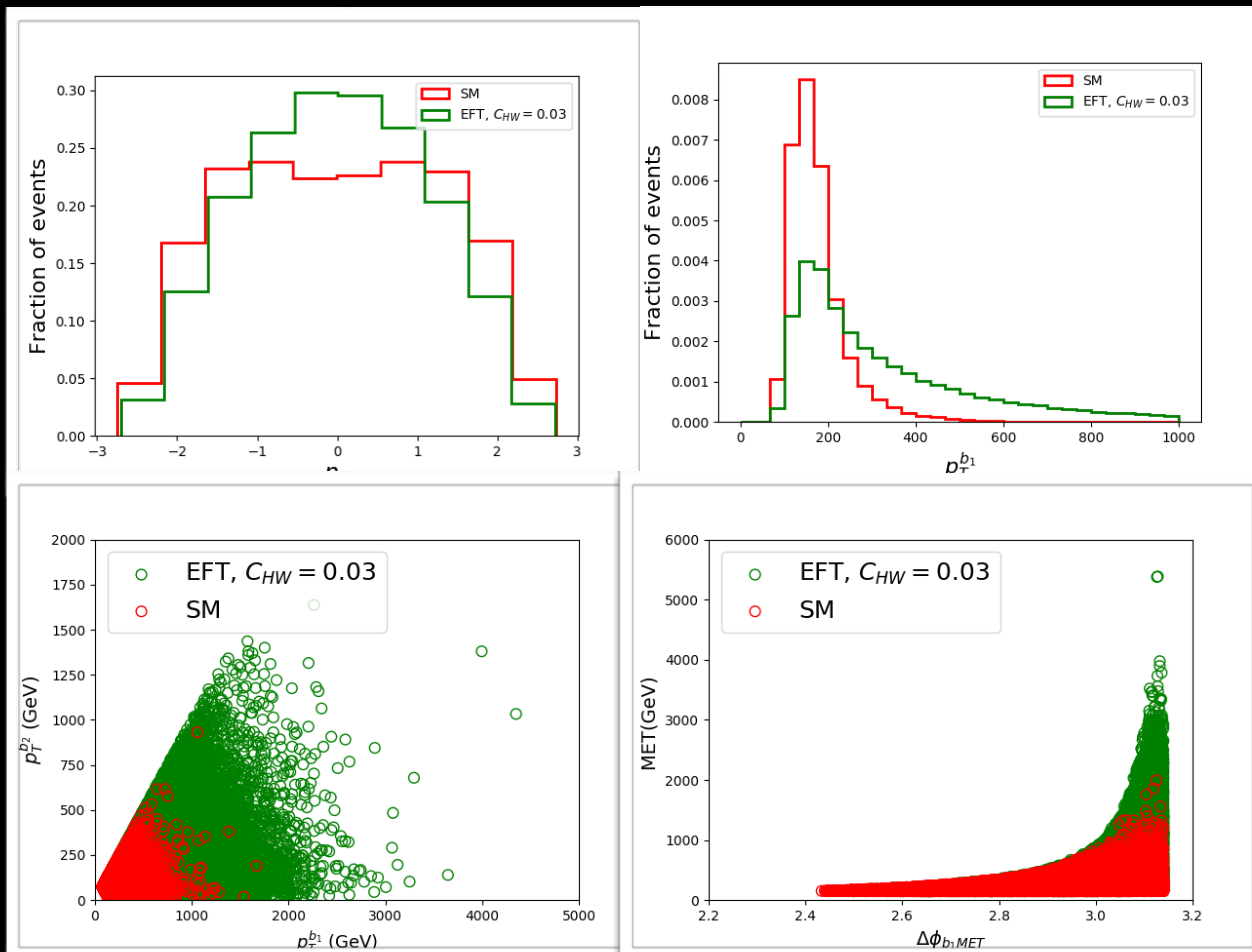


Key Insights

- Loss function
- Dropouts
- Activation function
- Backpropagation
- Epochs
- Receiver operative characteristic (ROC)



1D and 2D features



Appropriate Performance Measure for HEP Analysis

Discovery Significance $\frac{S}{\sqrt{B}}$

Asimov Significance

$$Z_A = \left[2\left((s+b)\ln\left[\frac{(s+b)(b+\sigma_b^2)}{b^2+(s+b)\sigma_b^2}\right] - \frac{b^2}{\sigma_b^2}\ln\left[1+\frac{\sigma_b^2 s}{b(b+\sigma_b^2)}\right]\right)\right]^{1/2}$$

Specific (Asimov) Loss function

$$\ell_{s/\sqrt{s+b}} = (s+b)/s^2$$

Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, arXiv:1007.1727[physics.data-an]

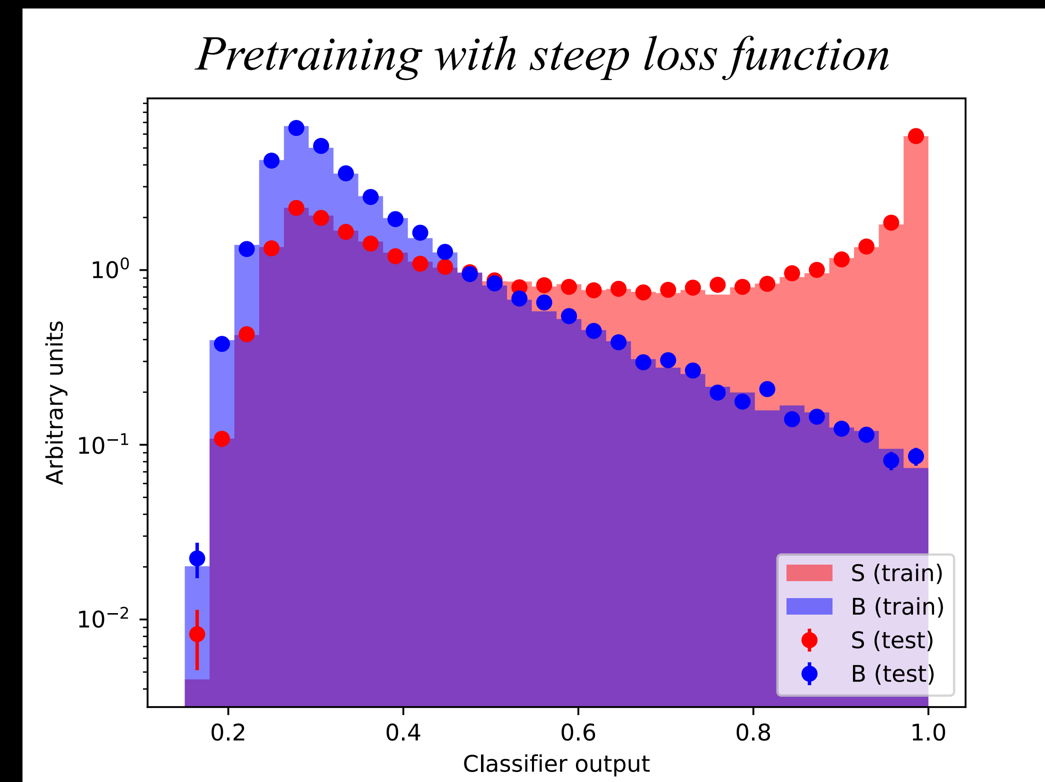
Adam Elwood and Dirk Krücker, arXiv: 1806.00322[hep-ex]

Neural Network Architecture

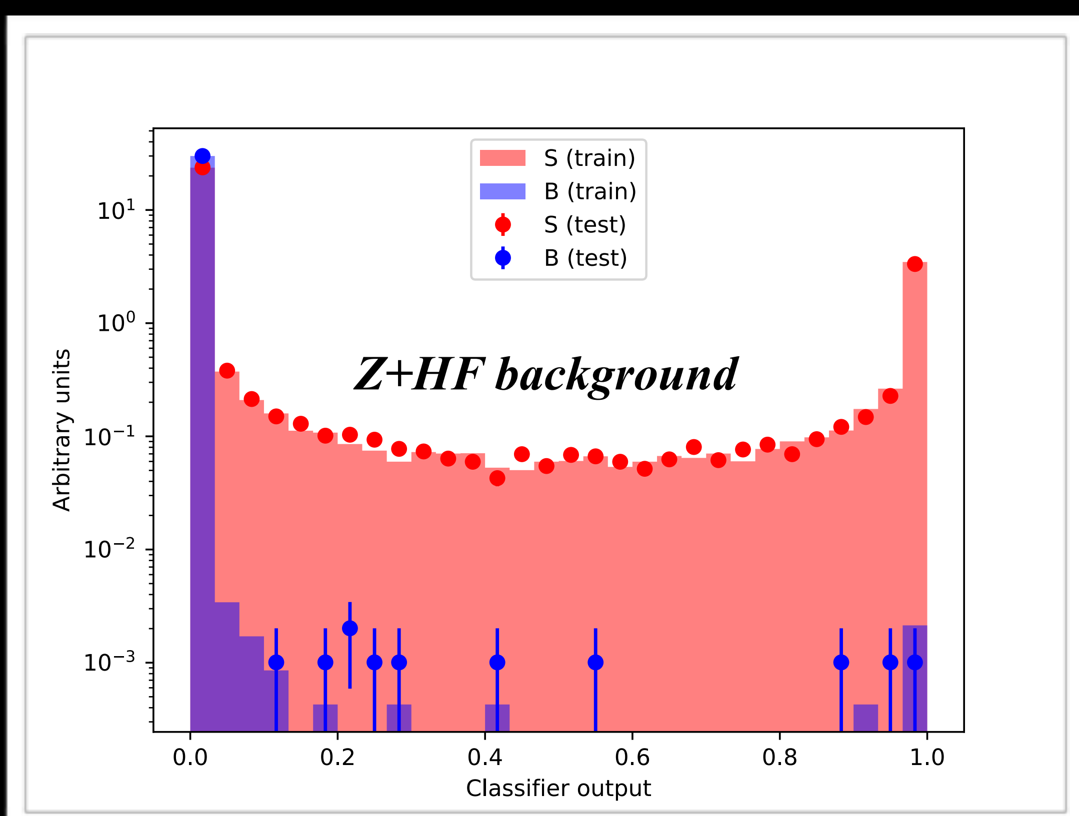
- Hidden layer : 1 (optimised)
- Dropouts : 0.2
- Activation function : ReLu
- Training set : Test set = 70 % : 30% (data scaling)
- Loss function : Asimov loss function (pre-training with MSE)

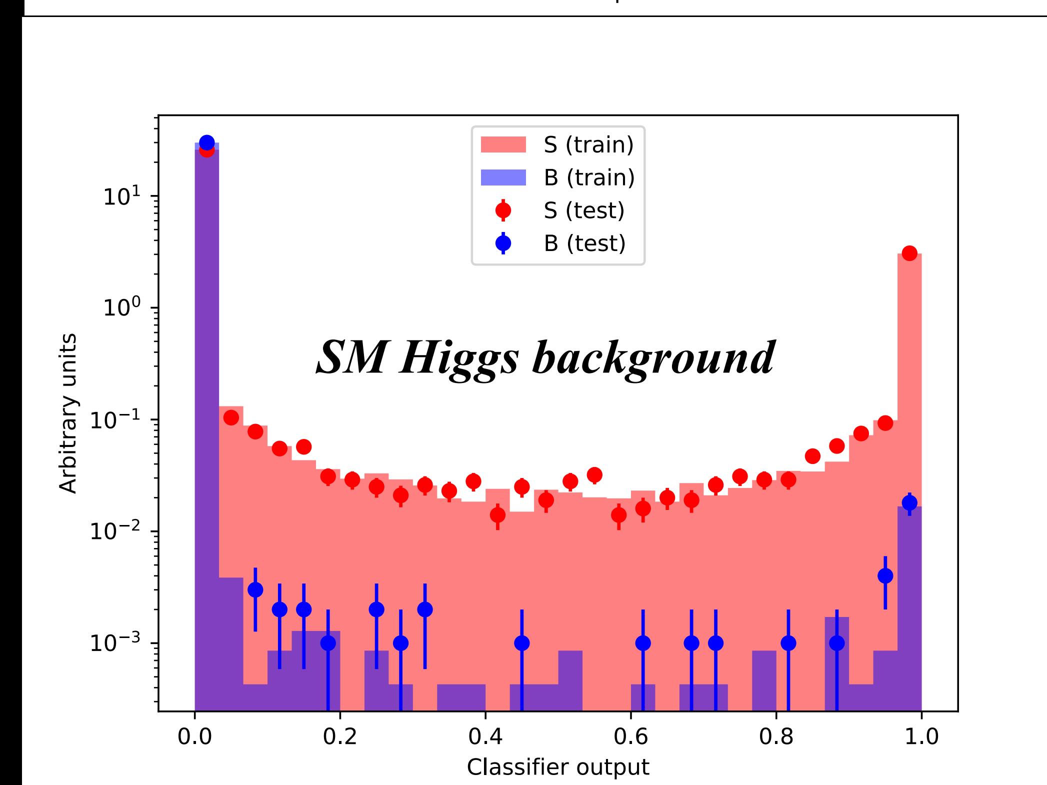
Classifier Output

Pre-training with usual loss function and 5 epochs

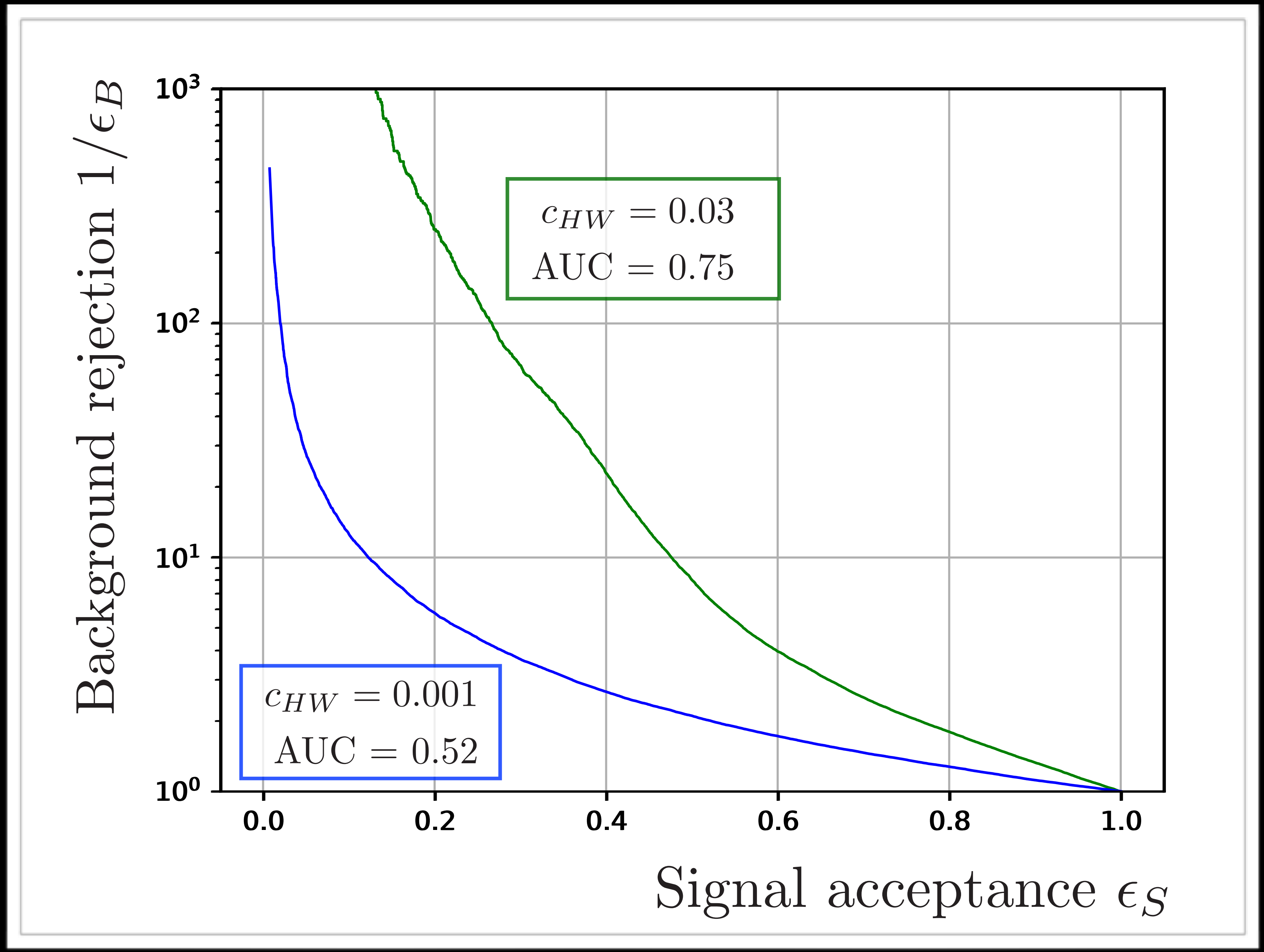


Final distributions with Asimov loss function

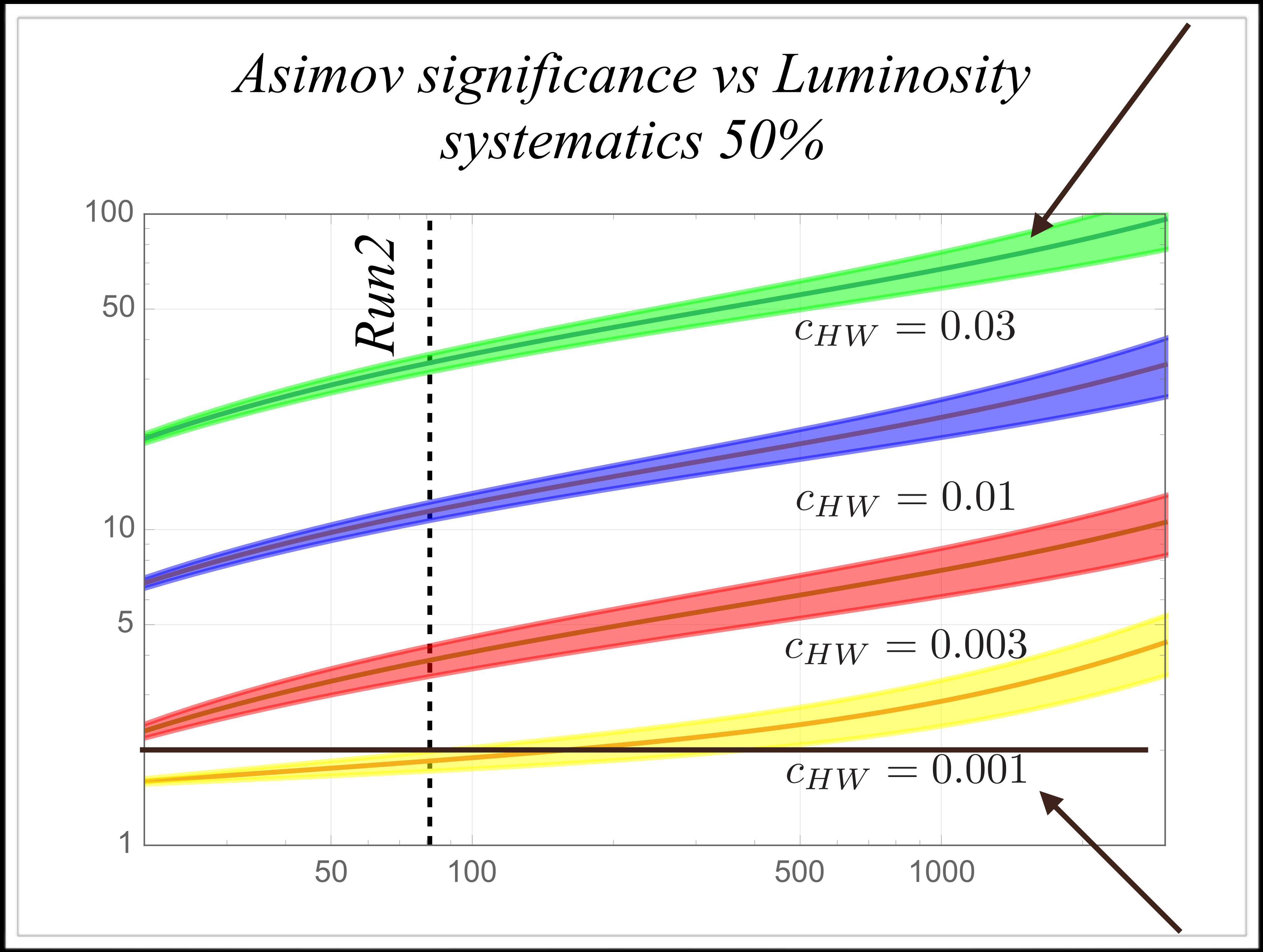




ROC (SMEFT vs SM Higgs background)

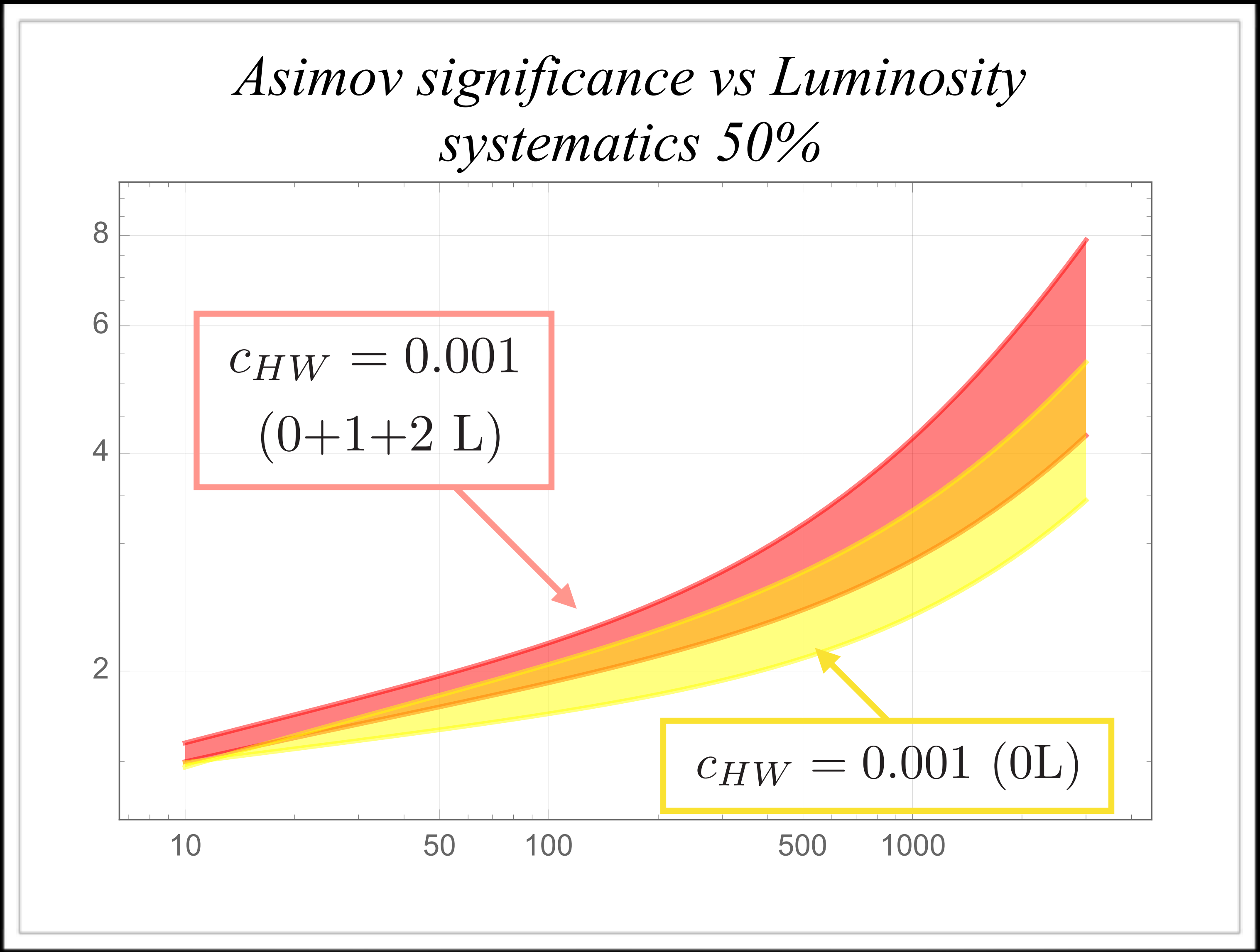


Current limit at 95% CL



Possible improvement

Combining 0L+1L+2L for a limiting case



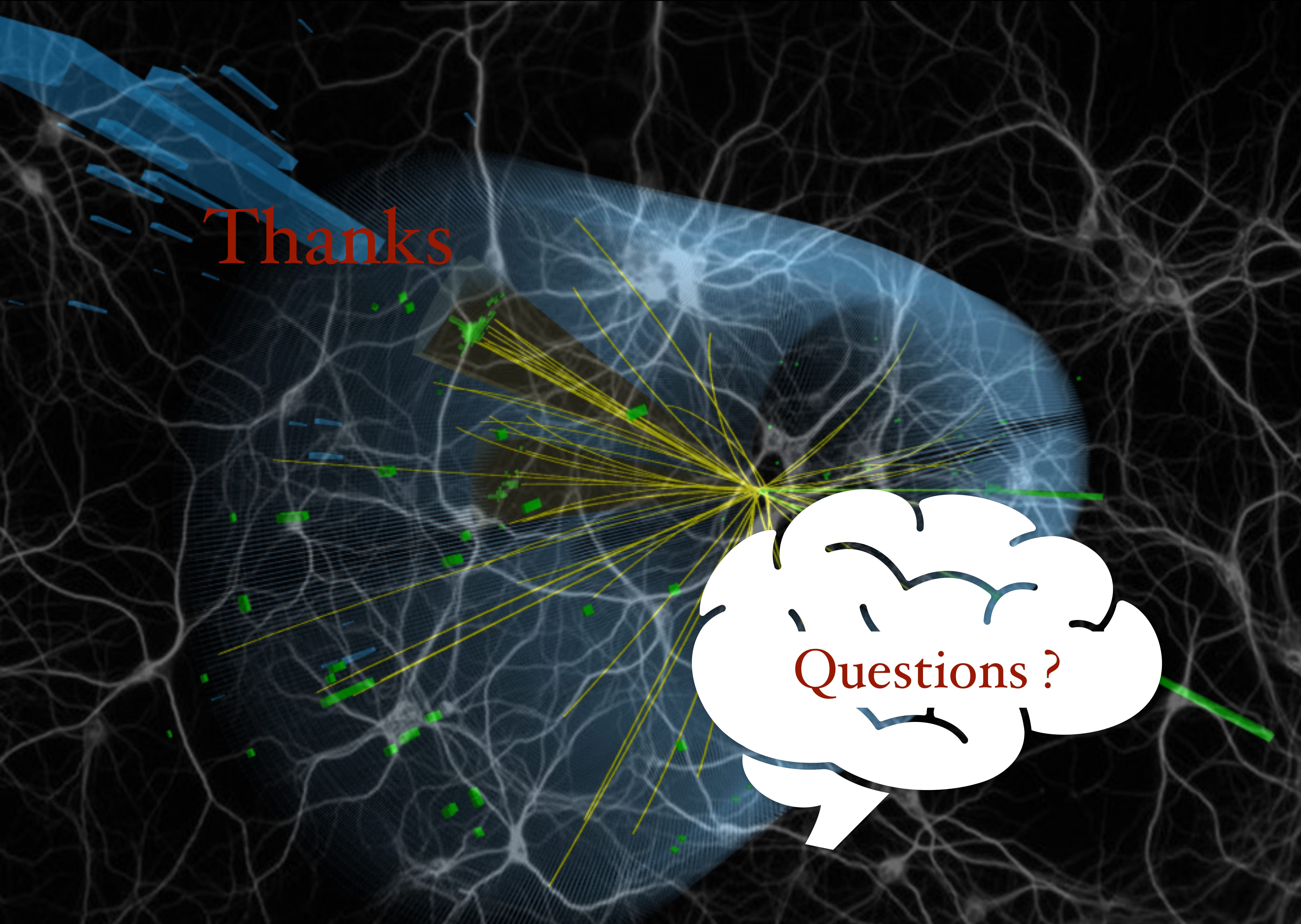
But for a realistic analysis combining different channels may help!!

Further …

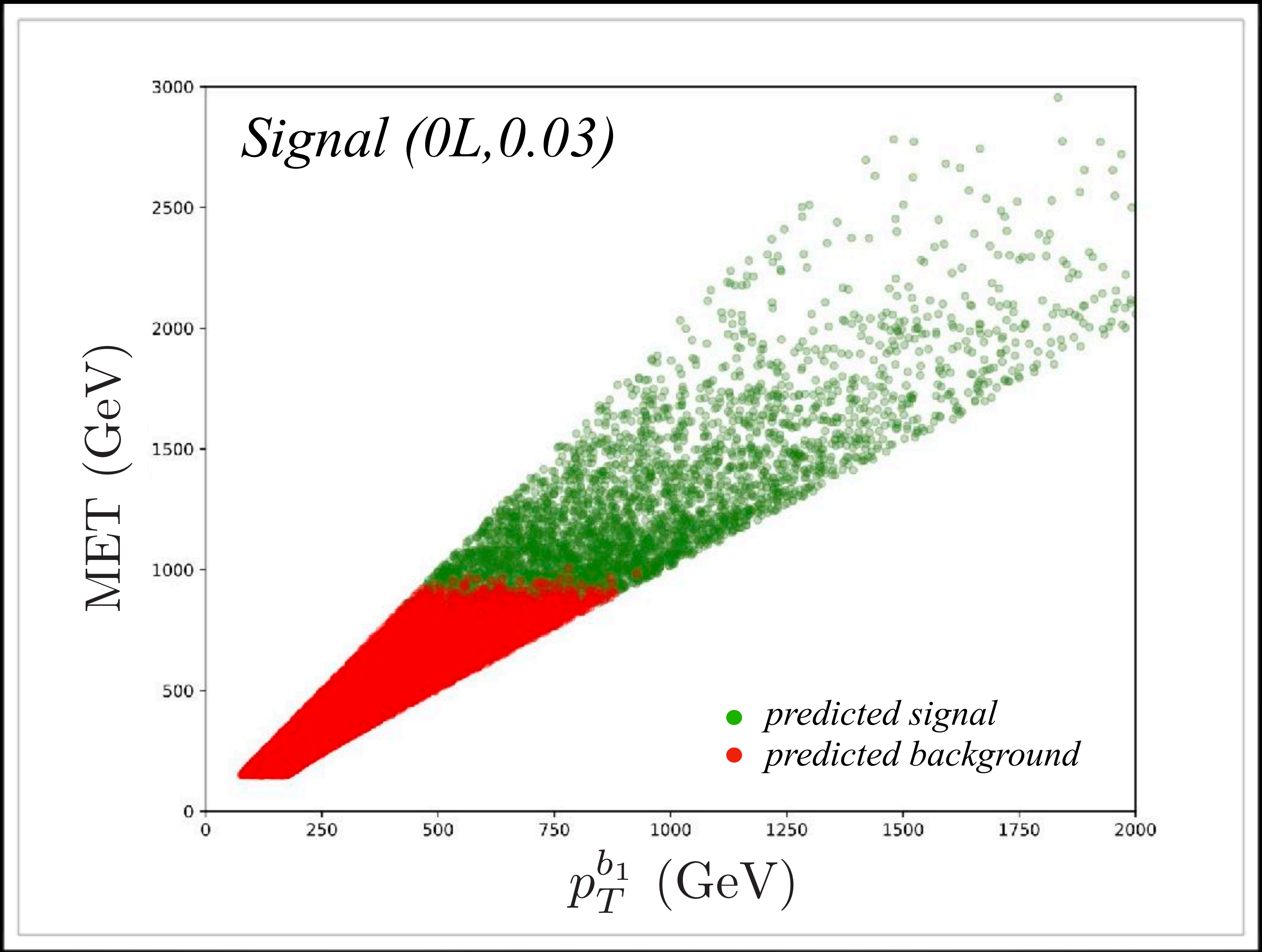
- This analysis is a "proof-of-concept" analysis
- Scaling for more operators and coefficients
- Hadronization, showering effects, detector simulation
- Global analysis using this algorithm (is in progress)
- We are discussing with ATLAS (VH) subgroup
- Different algorithms for same analysis set-up

Summary and Outlook

- Use of ML techniques to exploit kinematic information in SMEFT framework
- VH channel analysis (proof-of-concept) : one operator
- This analysis may provide significant stronger bounds on EFT coefficient
- Scalability for more operators
- Use/test this approach for actual LHC data



Reference: Felipe F. Freitas, CKK, Veronica Sanz, arXiv: 1902.05803 [hep-ph]

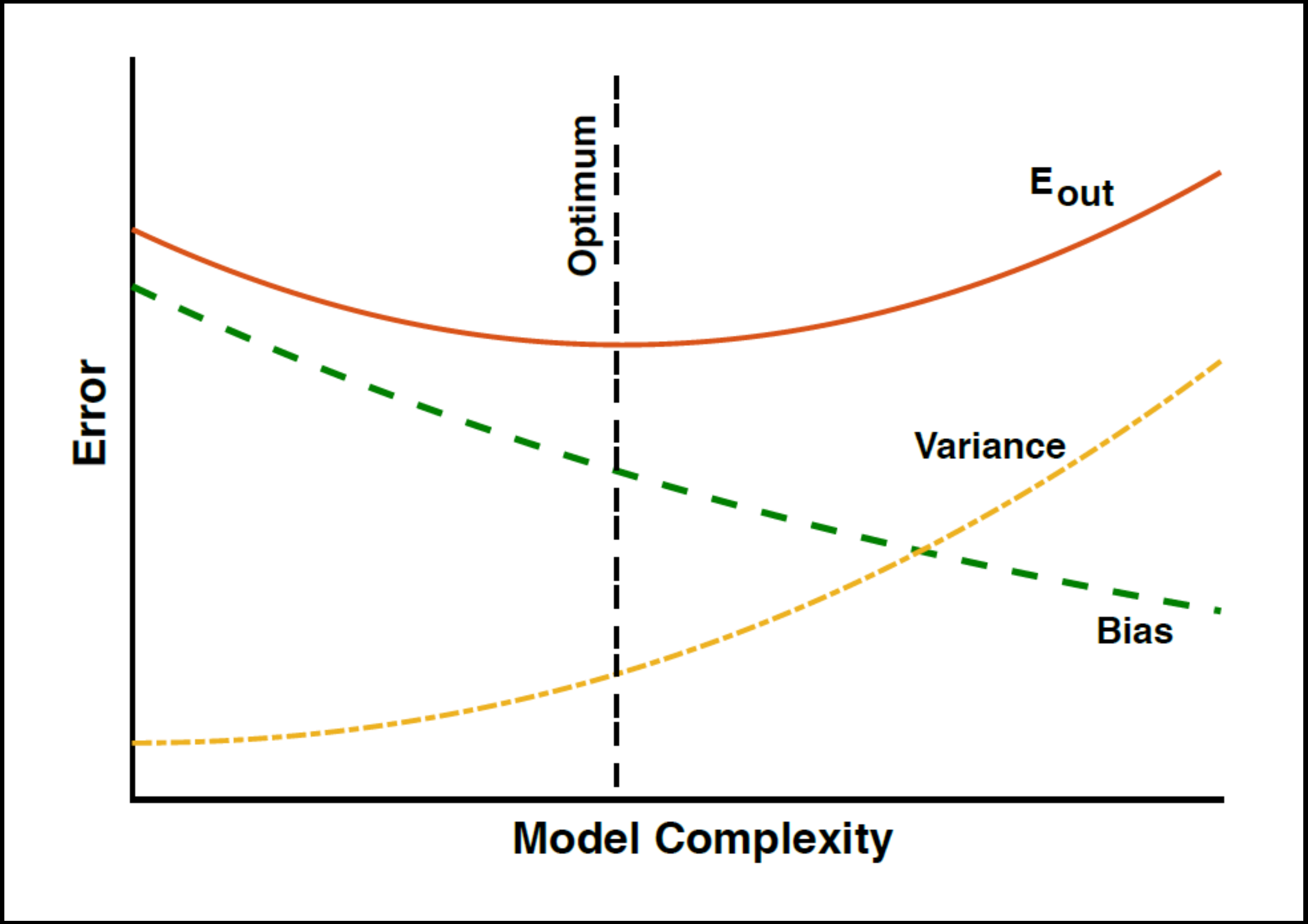


Signal (0L,0.03)
MET (GeV)
$p_T^{b_1}$ (GeV)
0
500
1000
1500
2000
2500
3000
0
250
500
750
1000
1250
1500
1750
2000
predicted signal
predicted background

Bias-Variance Trade-off

Bias : Deviation from the true value

Variance : fluctuations due to sample size



P. Mehta et al., arXiv:1803.08823[physics.comp-ph]