Analytic Optimal Observables from Symbolic Regression

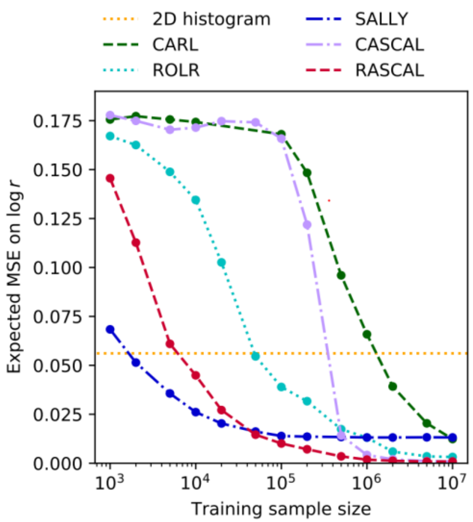
Nathalie Soybelman
ITP, Universität Heidelberg

Content

- What is an optimal observable?
- Associated Higgs production
- CP violation in weak boson fusion

J. Brehmer, A. Butter, K. Cranmer, T. Plehn, N. Soybelman

What is an optimal observable?



arXiv: 1805.00020

Best observable for measuring a physical parameter

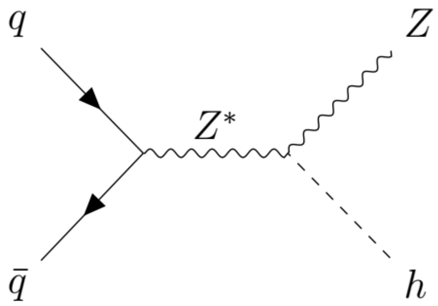
Likelihood $$p(z|\theta) = \frac{1}{\sigma(\theta)} \frac{\mathrm{d}\sigma(z|\theta)}{\mathrm{d}z}$$

Score $$t(z|\theta) = \nabla_\theta \log p(z|\theta) = \frac{\nabla_\theta \mathrm{d}\sigma(z|\theta)}{\mathrm{d}\sigma(z|\theta)} - \frac{\nabla_\theta \sigma(\theta)}{\sigma(\theta)} \sim \frac{\nabla_\theta |\mathcal{M}|^2}{|\mathcal{M}|^2}$$

z contains latent variables (e.g. initial quark flavor)

Marginalize over latent variables → Optimal Observable

Associated Higgs production



Parameter $\theta = f_B$

Operator $\mathcal{O}_B = \frac{ig'}{2}(D^\mu\phi)^\dagger D^\nu \phi B_{\mu\nu}$

Feynman rules $i\frac{2m_Z^2}{v}g_{\mu\nu} + i\frac{f_B}{\Lambda^2}\frac{g'^2 v}{4}\left(p_\nu^H p_\mu^{Z_1} + p_\mu^H p_\nu^{Z_2} - p^H \cdot p^{Z_1} g_{\mu\nu} - p^H \cdot p^{Z_2} g_{\mu\nu}\right)$

Matrix element $|\mathcal{M}|^2 = p_0 + a\theta + b\theta^2$

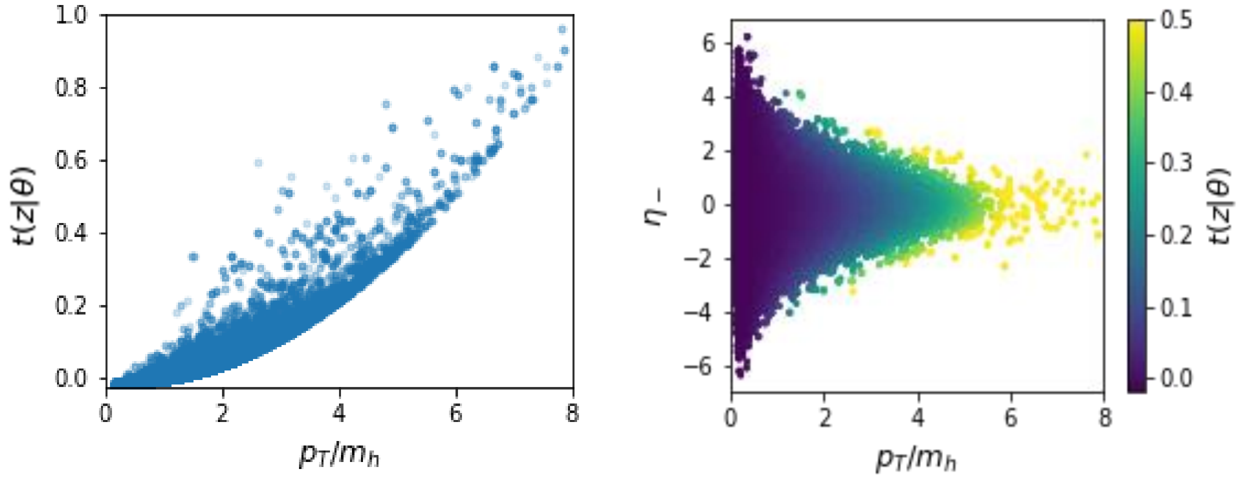
Score $t(z|\theta) \sim \frac{\nabla_\theta |\mathcal{M}|^2}{|\mathcal{M}|^2} = \frac{a + 2b\theta}{p_0 + a\theta + b\theta^2} \sim \frac{a}{p_0} + \frac{1}{p_0}\left(2b - \frac{a^2}{p_0}\right)\theta$

Choose one quark type → no unobservable variables

Optimal Observable for $f_B = 0$

$$\frac{\nabla_\theta |\mathcal{M}|^2}{|\mathcal{M}|^2} = \frac{a}{p_0}$$

Use MadMiner to compute score

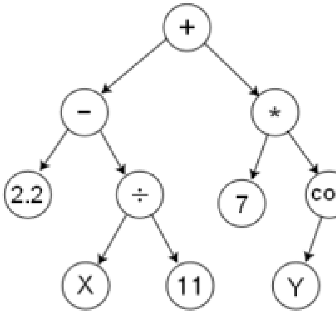


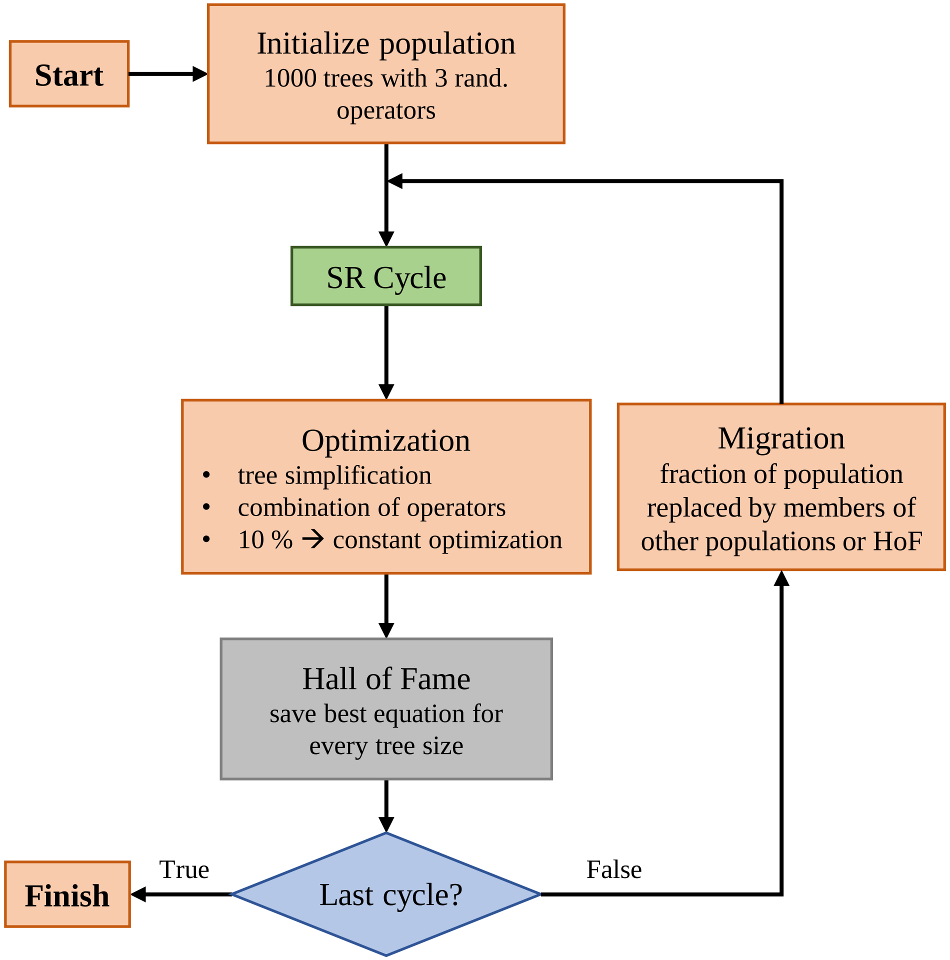
$$\eta_- = \eta_Z - \eta_H$$

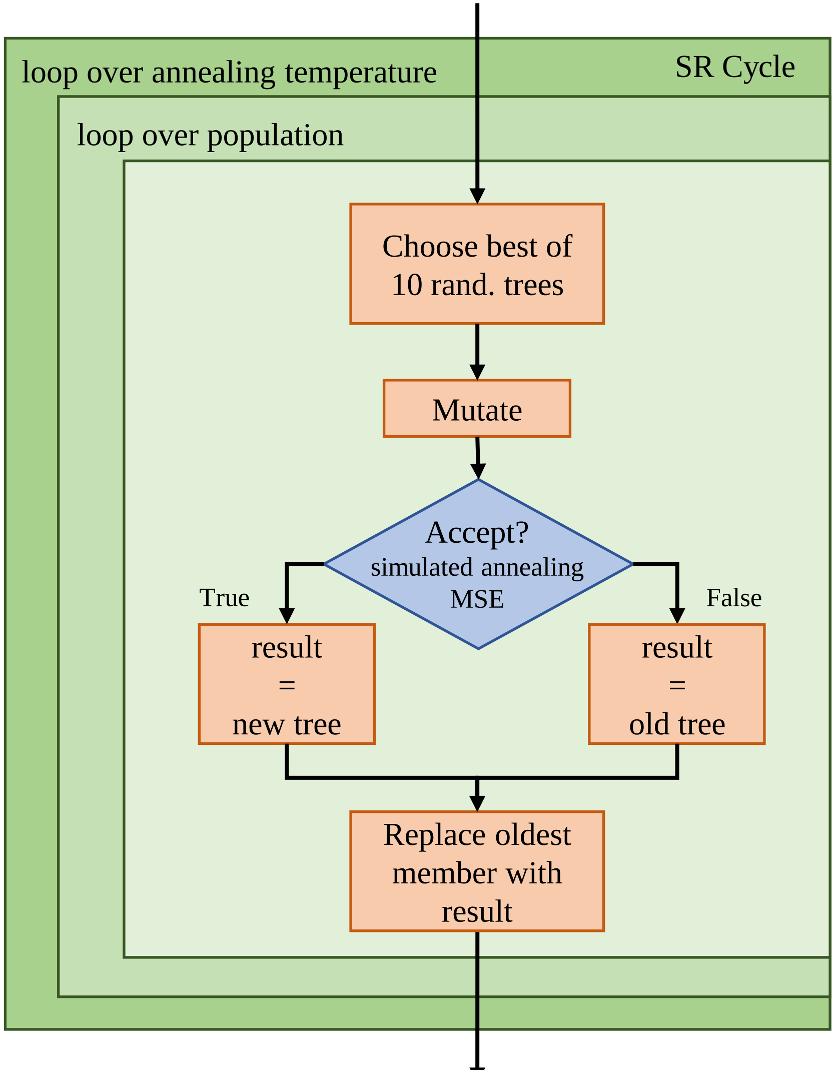
Score depends fully on two observables

Symbolic Regression

PySR: Python + Julia

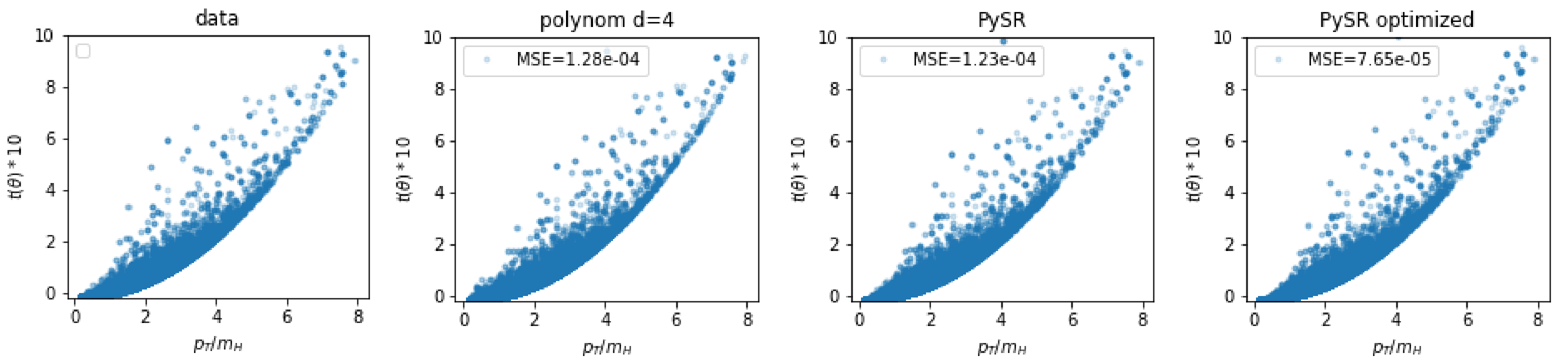






Nathalie Soybelman, Universität Heidelberg

Results $f_B = 0$



- Polynomial d = 4 agrees with data
- PySR recquires small data set → select data points by binning
- Postprocessing: optimizer → fitting parameters on full data set

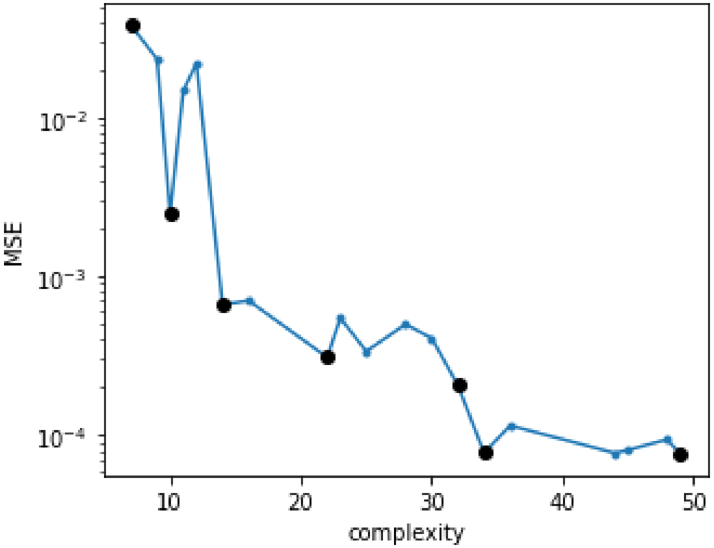
PySR + optimizer → best agreement with data

Hall of Fame

Best equation for every complexity

$x \to p_T$

$y \to |\eta_-|$



c.	p.	function	MSE
7	1	$ax(x+y)$	$3.81 \cdot 10^{-2}$
10	3	$ax^2(b+y)-c$	$2.49 \cdot 10^{-3}$
14	3	$ax^2+bx^2y^2-c$	$6.64 \cdot 10^{-4}$
22	4	$ax^2+bx^2y^2-cxy-d$	$3.09 \cdot 10^{-4}$
32	6	$a(x^2+y)+bx^2y-(cx-d)^2+ex^2y^3-f$	$2.06 \cdot 10^{-4}$
34	7	$a(x^2+y)+bx^2y-(cx-d)^2+ey^3(x-f)^2-g$	$7.77 \cdot 10^{-5}$
49	9	$ax^2+bx^2y-cy(x-d)+ey^3(x-f)^2+gx^2y^2-hx-i$	$7.65 \cdot 10^{-5}$

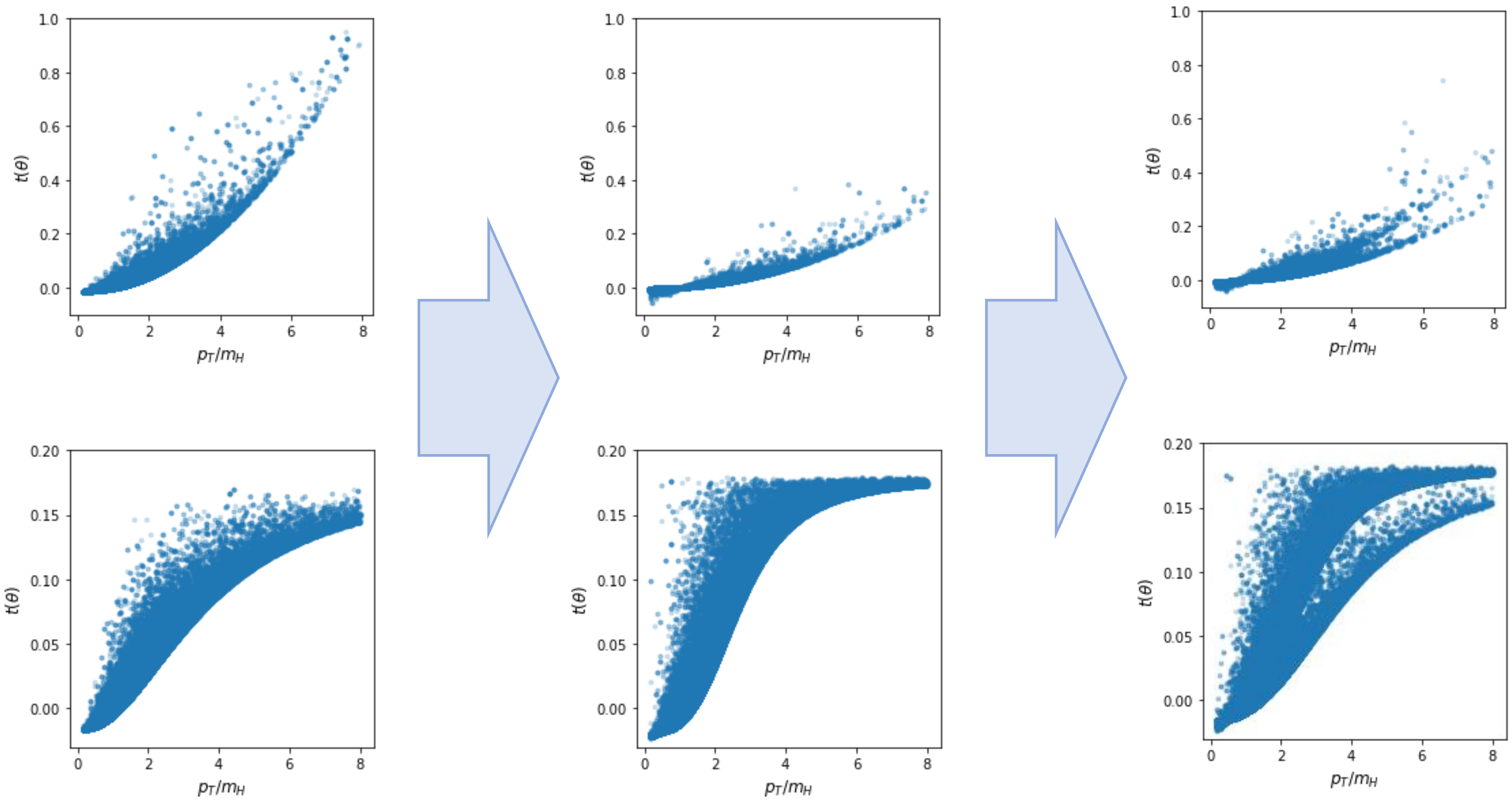
Choose “best“ function according to complexity and MSE

Generalize Process

only Z-boson | with photon | all quarks

$f_B = 0$

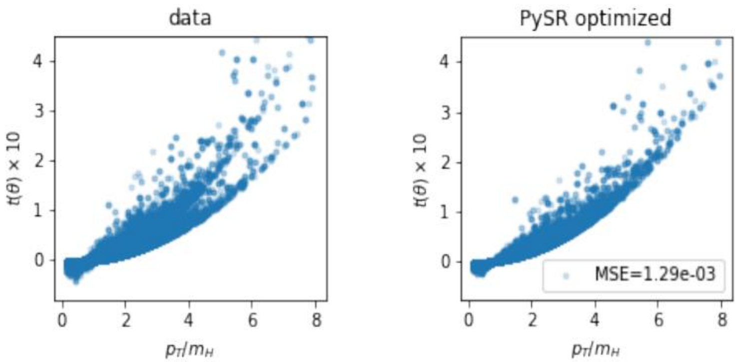
$f_B = 10$



Double branch structure

Process with all quarks

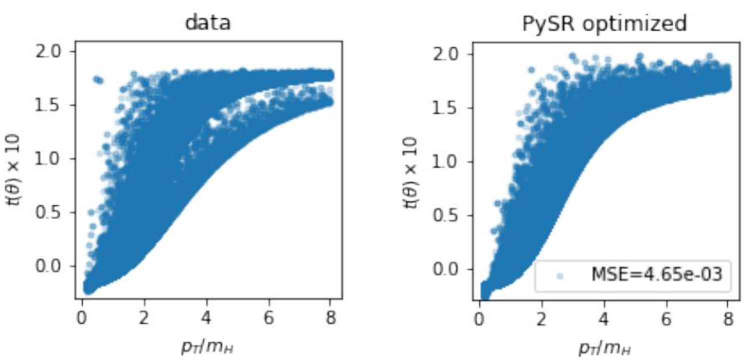
$f_B = 0$



- Quark flavor → Latent variable
- Marginalize over latent variable

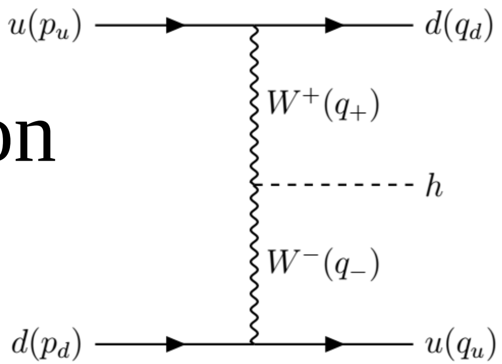
$$\hat{t} = ap_T^2 + b\left(|\eta_-|^2 + c\right)\left[|\eta_-|(dp_T - e)(p_T - f) + p_T + g\right] - h$$

$f_B = 10$



Good result for complex structure

$$\hat{t} = ap_T + b|\eta_-| - c(p_T - d)^3 + e - \frac{f}{gp_T^3|\eta_-|^3 - |\eta_-|(hp_T + i) + j(p_T + k)^6 + 1}$$



CP Violation in Weak Boson Fusion

Parameter $\theta = f_{W\widetilde{W}}$

Operator $\mathcal{O}_{W\widetilde{W}} = -(\phi^\dagger \phi)\widetilde{W}^k_{\mu\nu} W^{\mu\nu k}$

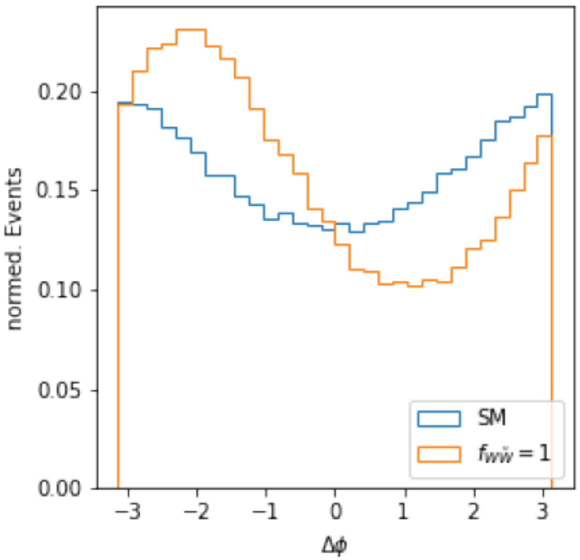
Feynman rules $2i\frac{m_W^2}{v} g_{\mu\nu} + 4i\frac{f_{W\widetilde{W}}v}{\Lambda^2}\epsilon_{\mu\nu\rho\sigma} q_+^\sigma q_-^\rho$

CP observable $\epsilon_{\mu\nu\rho\sigma} p_u^\mu p_d^\nu q_u^\sigma q_d^\rho \sim p_{T1} p_{T2} \sin \Delta\phi$

CPV observable is part of matrix element

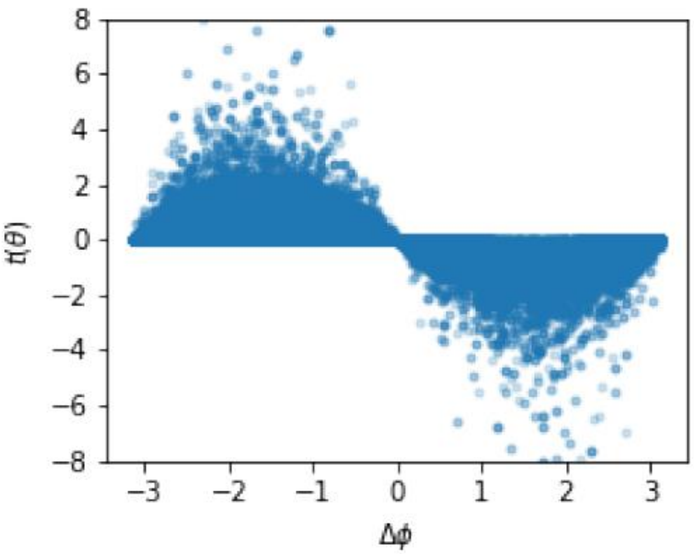
Optimal Observable for $f_{W\widetilde{W}} = 0$

Event distribution



Optimal Observable for $f_{W\widetilde{W}} = 0$

Score for $f_{W\widetilde{W}} = 0$

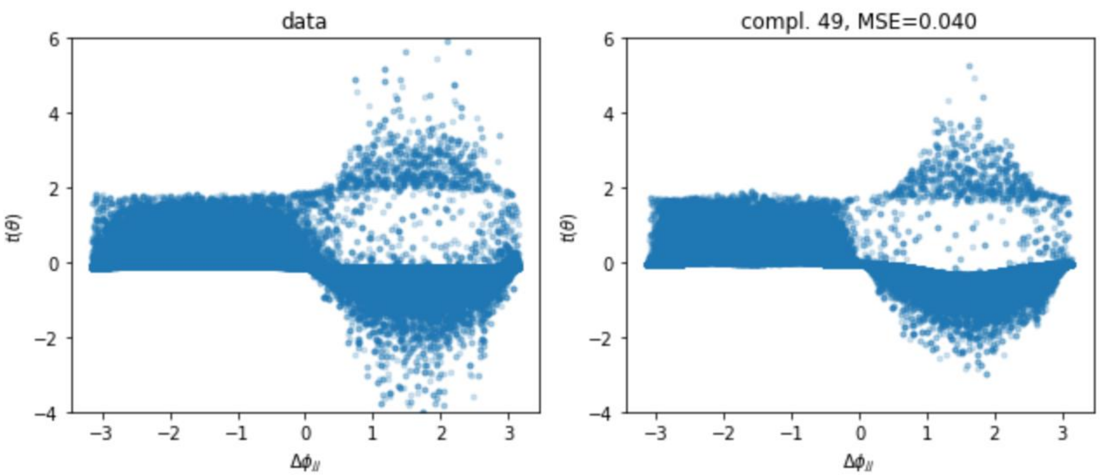


Process more complex → 2 variables not enough

Best fit using
$\Delta\phi,\ p_{Tj1},\ p_{Tj2},\ \Delta\eta$

c.	function	MSE
4	$a\sin(\Delta\phi)$	$1.03 \cdot 10^{-1}$
8	$a\sin(\Delta\phi)p_{Tj1}p_{Tj2}$	$1.49 \cdot 10^{-2}$
28	$(p_{Tj2}+a)(bp_{Tj1}(c-\Delta\phi)-p_{Tj1}(d\Delta\eta+ep_{Tj2}+f)\sin(\Delta\phi+g))$	$8.18 \cdot 10^{-3}$

Optimal Observable for $f_{W\widetilde{W}} = 1$



Input variables
$\sin\Delta\phi,\ p_+,\ p_\times,\ \Delta\eta$

$$p_+ = p_{Tj1} + p_{Tj2}$$

$$p_\times = p_{Tj1} p_{Tj2}$$

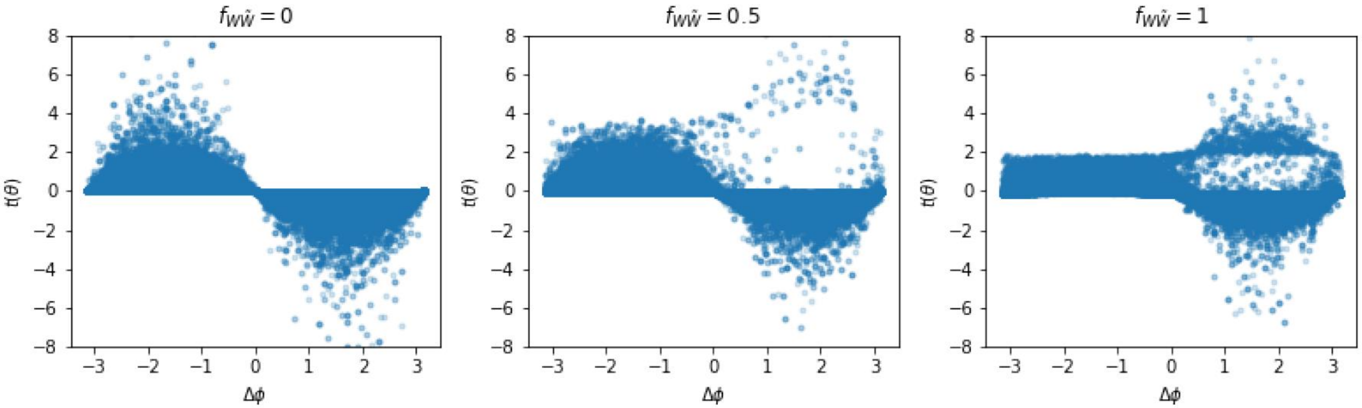
$$\hat{t} = a s_\phi^3 + \frac{b(s_\phi^3 + p_+)}{\dfrac{c\Delta\eta^2}{s_\phi^2\left(-\Delta\eta p_\times + p_\times^3 - d p_\times\left(\dfrac{\Delta\eta s_\phi}{s_\phi^2 + e} + s_\phi - p_+\right)\right)} - f s_\phi + p_\times} + g$$

Summary & Outlook

- Optimal Observable – best observable for parameter measurement
- Found analytic expression for Optimal Observable with SR
- Hall of Fame → find best equation with minimal complexity
- SR fits even complex structures (e.g. double branches)

More complex process – SR will choose the relevant observables

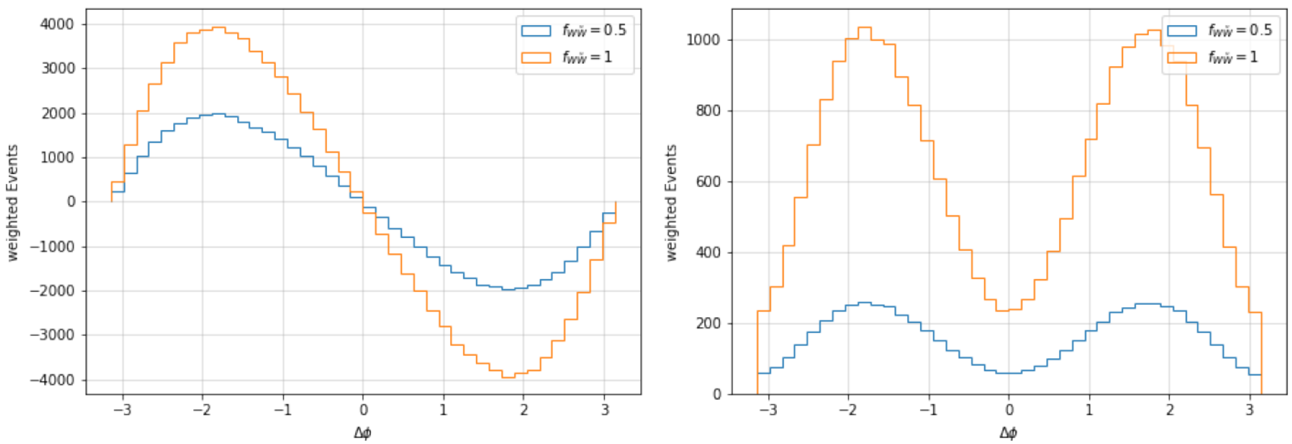
Back up



Quadratic term dominates → $$\frac{\nabla_\theta |\mathcal{M}|^2}{|\mathcal{M}|^2} = \frac{a + 2b\theta}{p_0 + a\theta + b\theta^2} \sim \frac{2}{\theta}$$

Interference term

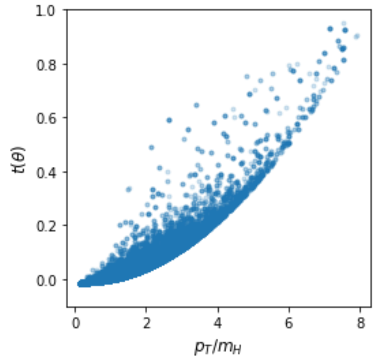
Quadratic term

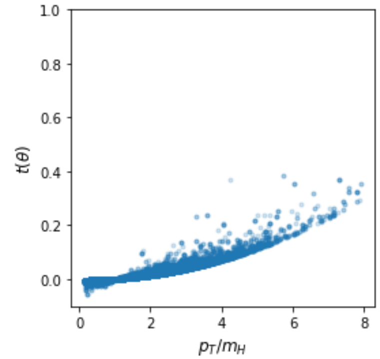


Back up

only Z-boson | with photon

$f_B = 0$

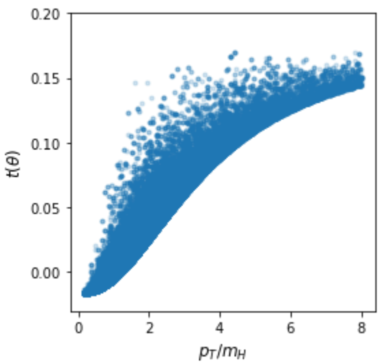


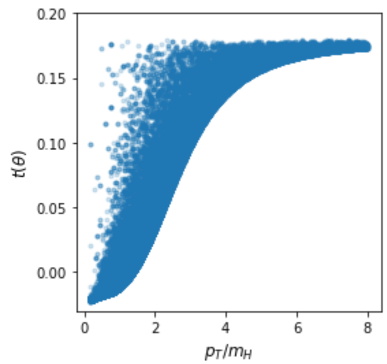


Relative minus sign between photon- and Z-vertex

→ Interference SM/NP smaller

$f_B = 10$





→ Quadratic term dominates

→ Score approaches constant

$$\frac{\nabla_\theta |\mathcal{M}|^2}{|\mathcal{M}|^2} = \frac{a + 2b\theta}{p_0 + a\theta + b\theta^2} \sim \frac{2}{\theta}$$