Better Higgs measurements through information geometry

Johann Brehmer

Heidelberg University

Based on 1612.05261
with Kyle Cranmer, Felix Kling, and Tilman Plehn

Joint particle seminar, UC Irvine
3 May 2017

To new physics through Higgs measurements

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

- There's probobably[1] new physics in the Higgs sector
    - Hierarchy problem
    - Fermion masses
    - DM
    - Baryon asymmetry
    - …

with a bit of creativity

- Measurement of Higgs properties most exciting mission for Run 2

until the LHC finds something really cool

- Need model-independent parametrisation of Higgs properties

[1] No warranty, expressed or implied

SM effective field theory

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

[W. Buchmuller, D. Wyler 85; K. Hagiwara, S. Ishihara, S. R. Szalapski, D. Zeppenfeld 93;
B. Grzadkowski, M. Iskrzynski, M. Misiak, J. Rosiek 1008.4884; …]

- New physics at $\Lambda \gg E_{\text{LHC}} \sim m_h$?

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \underbrace{\sum_i^{59} \frac{f_i^{d=6}}{\Lambda^2} \mathcal{O}_i^{d=6}} + \sum_k \frac{f_k^{d=8}}{\Lambda^4} \mathcal{O}_k^{d=8} + \dots$$

e. g. $\mathcal{O}_W = (D^\mu \phi)^\dagger \sigma^k (D^\nu \phi) W_{\mu\nu}^k \;\dots$

SM effective field theory

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

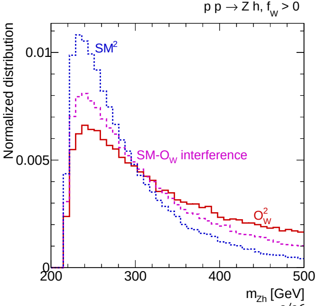
[W. Buchmuller, D. Wyler 85; K. Hagiwara, S. Ishihara, S. R. Szalapski, D. Zeppenfeld 93; B. Grzadkowski, M. Iskrzynski, M. Misiak, J. Rosiek 1008.4884; …]

- New physics at $\Lambda \gg E_{\text{LHC}} \sim m_h$?

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \underbrace{\sum_i^{59} \frac{f_i^{d=6}}{\Lambda^2} \mathcal{O}_i^{d=6}} + \sum_k \frac{f_k^{d=8}}{\Lambda^4} \mathcal{O}_k^{d=8} + \dots$$

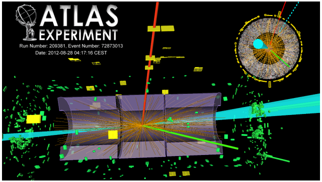
$$\text{e.g. } \mathcal{O}_W = (D^\mu \phi)^\dagger \sigma^k (D^\nu \phi) W^k_{\mu\nu} \dots$$

- Dimension-6 operators: perfect language for new physics signatures in Higgs sector?
  - Model independence?
  - Correlations between Higgs, LHC TGC, LEP, …
  - Total rates + distributions



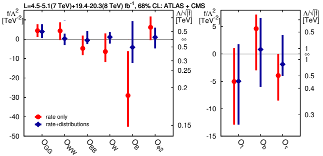
Infering Higgs properties

Complex data $\mathbf{x}$



[ATLAS 1501.04943]

Many parameters $\boldsymbol{\theta}$

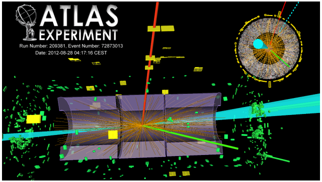


[T. Corbett et al 1505.05516]

Infering Higgs properties

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Complex data $\mathbf{x}$

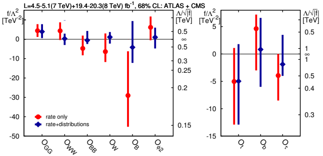


[ATLAS 1501.04943]

- Conventional analyses:
  - standard kinematic observables

  $\Rightarrow$ reproducible and transparent;
  don't scale well with complexity

Many parameters $\boldsymbol{\theta}$

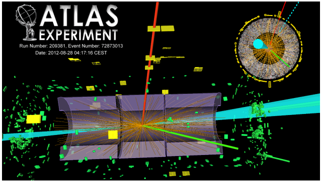


[T. Corbett et al 1505.05516]

Infering Higgs properties

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Complex data $\mathbf{x}$



[ATLAS 1501.04943]

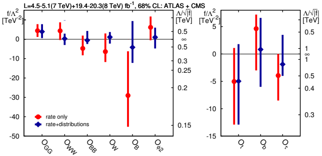
- Conventional analyses:
  - standard kinematic observables

  $\Rightarrow$ reproducible and transparent; don't scale well with complexity

- Multivariate methods:
  - matrix-element-based
  - likelihood-free inference (machine learning)

  $\Rightarrow$ powerful black boxes

Many parameters $\boldsymbol{\theta}$



[T. Corbett et al 1505.05516]

Efficient measurements need guidelines

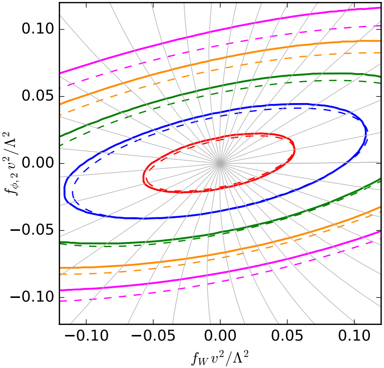
UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

1. What is the maximum sensitivity of a measurement?
2. Does the EFT series converge?
3. Where in phase space is the information?
4. What are the most powerful observables?

Efficient measurements need guidelines

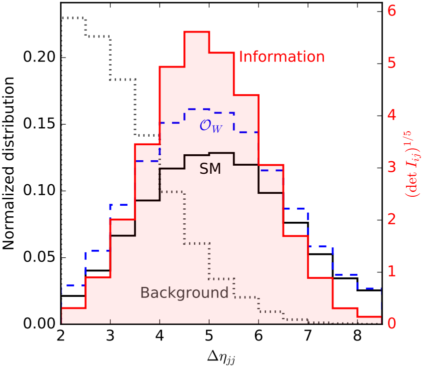
1. What
2. Does
3. Whe
4. What

Sneak preview:



Efficient measurements need guidelines

Sneak preview:



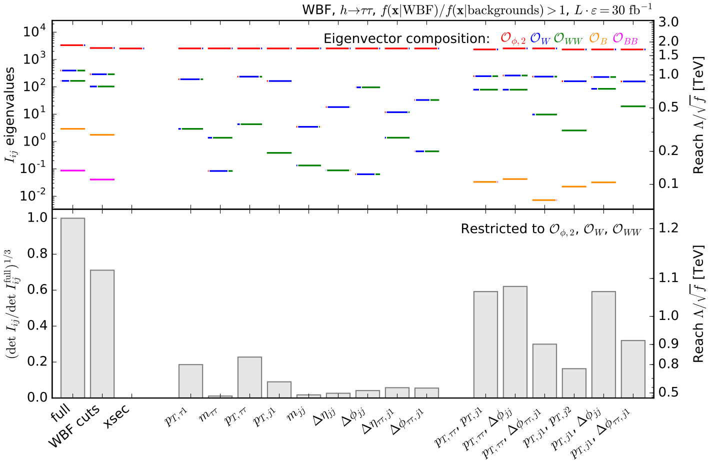
1. What
2. Does
3. Whe
4. What

Efficient measurements need guidelines

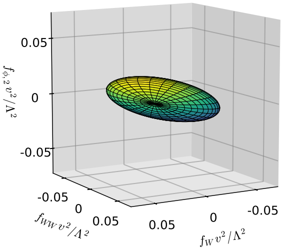
UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

1. What
2. Does
3. Whe
4. What

Sneak preview:



1. What is the maximum sensitivity of a measurement?



Cramér-Rao bound

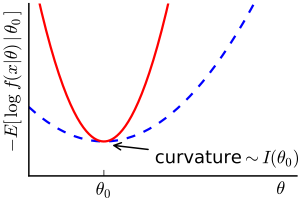
UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

- Starting point: probability distribution $f\big(\underbrace{\mathbf{x}}_{\text{observables}} \,\big|\, \underbrace{\boldsymbol{\theta}}_{\text{theory parameters}}\big)$

- Fisher information

$$I_{ij}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \log f(\mathbf{x}|\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\middle|\boldsymbol{\theta}\right]$$

[F. Edgeworth 1908; R. Fisher 1925; ...]



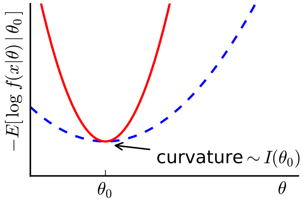
Cramér-Rao bound

- Starting point: probability distribution $f\Big(\underbrace{\mathbf{x}}_{\text{observables}}\;\Big|\;\underbrace{\boldsymbol{\theta}}_{\text{theory parameters}}\Big)$

- Fisher information

$$I_{ij}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \log f(\mathbf{x}|\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\middle|\boldsymbol{\theta}\right]$$

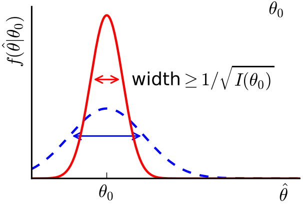
[F. Edgeworth 1908; R. Fisher 1925; …]



- Cramér-Rao bound: all unbiased estimators $\hat{\boldsymbol{\theta}}(\mathbf{x})$ satisfy

$$\operatorname{cov}\left[\hat{\boldsymbol{\theta}}\middle|\boldsymbol{\theta}_0\right]_{ij} \geq I^{-1}_{ij}(\boldsymbol{\theta}_0)$$

[C. R. Rao 1945; H. Cramér 1946]



$\Rightarrow$ I_{ij} ~ maximal knowledge on $\boldsymbol{\theta}$ we can derive from an observation

The Fisher information matrix

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

- Properties:
  - Describes all directions in theory space
  - Additive between experiments / phase-space regions
  - Independent of parametrization of $\mathbf{x}$
  - Covariant under $\boldsymbol{\theta} \to \boldsymbol{\Theta}(\boldsymbol{\theta})$:

$$I_{ab}(\boldsymbol{\Theta}) = \frac{\partial \theta_i}{\partial \Theta_a} I_{ij}(\boldsymbol{\theta}) \frac{\partial \theta_j}{\partial \Theta_b}$$

The Fisher information matrix

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

- Properties:
  - Describes all directions in theory space
  - Additive between experiments / phase-space regions
  - Independent of parametrization of $\mathbf{x}$
  - Covariant under $\boldsymbol{\theta} \to \boldsymbol{\Theta}(\boldsymbol{\theta})$:

$$I_{ab}(\boldsymbol{\Theta}) = \frac{\partial \theta_i}{\partial \Theta_a} I_{ij}(\boldsymbol{\theta}) \frac{\partial \theta_j}{\partial \Theta_b}$$

- Frequently used to optimize experiments
  - E. g. cosmology

    [P. Jaranowski, A. Krolak 1994; L. Wolz, M. Kilbinger, J. Wellter, T. Giannantonio 1205.3984; T. Edwards, C. Weniger 1704.05458; . . .]

  - Rarely in particle physics

    [F. Ferreira, S. Fichet, V. Sanz 1702.05106; CMS 1704.06142]

Fisher information in LHC processes

- Extended likelihood ansatz:

$$f(\mathbf{x}|\boldsymbol{\theta}) = \mathrm{Pois}(n|\sigma L) \prod_{i=1}^{n} f^{(1)}(x_i|\boldsymbol{\theta})$$

Fisher information in LHC processes

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

- Extended likelihood ansatz:

$$f(\mathbf{x}|\boldsymbol{\theta}) = \mathrm{Pois}(n|\sigma L) \prod_{i=1}^{n} f^{(1)}(x_i|\boldsymbol{\theta})$$

- Fisher information:

$$I_{ij}(\boldsymbol{\theta}) = L \frac{\partial\sigma}{\partial\theta_i} \frac{1}{\sigma} \frac{\partial\sigma}{\partial\theta_j} + L\,\sigma\,E\left[\frac{\partial \log f^{(1)}(x|\boldsymbol{\theta})}{\partial\theta_i} \frac{\partial \log f^{(1)}(x|\boldsymbol{\theta})}{\partial\theta_j}\middle|\boldsymbol{\theta}\right]$$

MC integration, $\int \mathrm{d}x\ f^{(1)}(x) \to \sum_{\text{events}} \Delta\sigma/\sigma$:

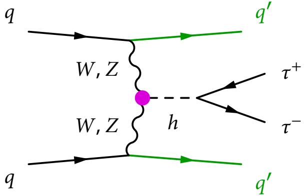
$$I_{ij}(\boldsymbol{\theta}) = L \sum_{\text{events } k} \frac{\partial\Delta\sigma_k}{\partial\theta_i} \frac{1}{\Delta\sigma_k} \frac{\partial\Delta\sigma_k}{\partial\theta_j}$$

$\Rightarrow$ Can calculate all $I_{ij}(\boldsymbol{\theta})$ from a single MC run

Weak boson fusion (WBF), $h \to \tau\tau$

- Kinematics of tagging jets sensitive to Higgs-gauge interaction

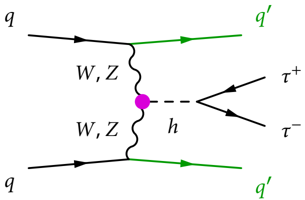
  [D. Rainwater, D. Zeppenfeld, K. Hagiwara hep-ph/9808468;
  T. Plehn, D. Rainwater, D. Zeppenfeld hep-ph/0105325;
  C. Englert, D. Gonçalves-Netto, K. Mawatari, T. Plehn 1212.0843; ...]



Weak boson fusion (WBF), $h \to \tau\tau$

- Kinematics of tagging jets sensitive to Higgs-gauge interaction

  [D. Rainwater, D. Zeppenfeld, K. Hagiwara hep-ph/9808468; T. Plehn, D. Rainwater, D. Zeppenfeld hep-ph/0105325; C. Englert, D. Gonçalves-Netto, K. Mawatari, T. Plehn 1212.0843; …]

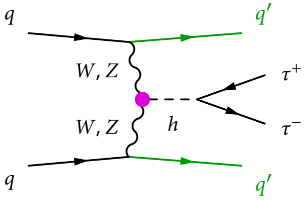


- Model: dimension-6 Higgs-gauge operators

$$\boldsymbol{\theta} = \frac{v^2}{\Lambda^2} \begin{pmatrix} f_{\phi,2} \\ f_W \\ f_{WW} \\ f_B \\ f_{BB} \end{pmatrix} \longrightarrow \mathcal{O}_{\phi,2} = \frac{1}{2}\,\partial^\mu(\phi^\dagger\phi)\,\partial_\mu(\phi^\dagger\phi)$$

rescales all h couplings

Weak boson fusion (WBF), $h \to \tau\tau$



- Kinematics of tagging jets sensitive to Higgs-gauge interaction

  [D. Rainwater, D. Zeppenfeld, K. Hagiwara hep-ph/9808468;
  T. Plehn, D. Rainwater, D. Zeppenfeld hep-ph/0105325;
  C. Englert, D. Gonçalves-Netto, K. Mawatari, T. Plehn 1212.0843; ...]

- Model: dimension-6 Higgs-gauge operators

$$\boldsymbol{\theta} = \frac{v^2}{\Lambda^2} \begin{pmatrix} f_{\phi,2} \\ f_W \\ f_{WW} \\ f_B \\ f_{BB} \end{pmatrix}$$

$f_{\phi,2} \longrightarrow \mathcal{O}_{\phi,2} = \frac{1}{2} \partial^\mu (\phi^\dagger \phi) \partial_\mu (\phi^\dagger \phi)$ — rescales all h couplings

$f_W \longrightarrow \mathcal{O}_W = \mathrm{i} \frac{g}{2} (D^\mu \phi)^\dagger \sigma^k (D^\nu \phi) W^k_{\mu\nu}$

$f_{WW} \longrightarrow \mathcal{O}_{WW} = -\frac{g^2}{4} (\phi^\dagger \phi) W^k_{\mu\nu} W^{\mu\nu\, k}$

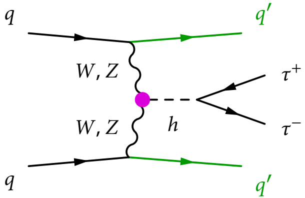
hWW, hZZ kinematics

Weak boson fusion (WBF), $h \to \tau\tau$

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

- Kinematics of tagging jets sensitive to Higgs-gauge interaction

  [D. Rainwater, D. Zeppenfeld, K. Hagiwara hep-ph/9808468; T. Plehn, D. Rainwater, D. Zeppenfeld hep-ph/0105325; C. Englert, D. Gonçalves-Netto, K. Mawatari, T. Plehn 1212.0843; ...]



- Model: dimension-6 Higgs-gauge operators

$$\boldsymbol{\theta} = \frac{v^2}{\Lambda^2} \begin{pmatrix} f_{\phi,2} \\ f_W \\ f_{WW} \\ f_B \\ f_{BB} \end{pmatrix}$$

$f_{\phi,2} \longrightarrow \mathcal{O}_{\phi,2} = \frac{1}{2}\,\partial^\mu(\phi^\dagger\phi)\,\partial_\mu(\phi^\dagger\phi)$ — rescales all h couplings

$f_W \longrightarrow \mathcal{O}_W = \mathrm{i}\frac{g}{2}\,(D^\mu\phi)^\dagger\sigma^k(D^\nu\phi)\,W^k_{\mu\nu}$

$f_{WW} \longrightarrow \mathcal{O}_{WW} = -\frac{g^2}{4}\,(\phi^\dagger\phi)\,W^k_{\mu\nu}\,W^{\mu\nu\,k}$

hWW, hZZ kinematics

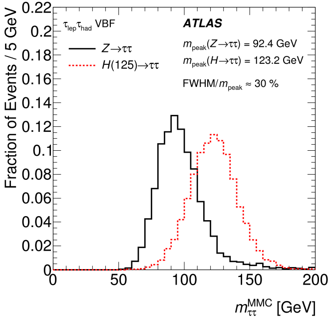
$f_B \longrightarrow \mathcal{O}_B = \mathrm{i}\frac{g}{2}\,(D^\mu\phi^\dagger)(D^\nu\phi)\,B_{\mu\nu}$

$f_{BB} \longrightarrow \mathcal{O}_{BB} = -\frac{g'^2}{4}\,(\phi^\dagger\phi)\,B_{\mu\nu}\,B^{\mu\nu}$

hZZ kinematics

Setup

- Tools: MadGraph 5, MadMax
  [J. Alwall et al 1405.0301;
  K. Cranmer, T. Plehn hep-ph/0605268;
  T. Plehn, P. Schichtel, D. Wiegand 1311.2591;
  F. Kling, T. Plehn, P. Schichtel 1607.07441]
- Backgrounds:
  - QCD and electroweak $Z \to \tau\tau$
  - Gluon-fusion Higgs production



- Approximations:
  - τ decays not simulated — BR for semileptonic $\tau\tau$ mode
  - Parton level — CJV survival probabilities from literature [D. Rainwater, D. Zeppenfeld, K. Hagiwara hep-ph/9808468]
  - No detector simulation — $m_{\tau\tau}$ smeared by single / double Gaussian fitted to ATLAS results [ATLAS 1501.04943, see above]
  - No systematic or theory uncertainties
- $\sqrt{s} = 13$ TeV, $L \cdot \varepsilon = 30$ fb^{-1}
- Cuts: $p_{T,j} > 20$ GeV, $|\eta_j| < 5.0$, $\Delta\eta_{jj} > 2.0$, $\Delta R_{jj} > 0.4$

Maximum sensitivity in WBF, $h \to \tau\tau$

- Fisher information at the SM:

$$
I_{ij}(\mathbf{0}) = \begin{array}{c} \\ \left(\begin{array}{rrrrr} 3202 & -625 & -7 & -35 & 0 \\ -625 & 451 & -110 & 23 & -2 \\ -7 & -110 & 244 & -6 & 3 \\ -35 & 23 & -6 & 4 & 0 \\ 0 & -2 & 3 & 0 & 0 \end{array}\right) \end{array}
\begin{array}{l} \\ \mathcal{O}_{\phi,2} \\ \mathcal{O}_{W} \\ \mathcal{O}_{WW} \\ \mathcal{O}_{B} \\ \mathcal{O}_{BB} \end{array}
$$

with column labels $\mathcal{O}_{\phi,2}$, $\mathcal{O}_{W}$, $\mathcal{O}_{WW}$, $\mathcal{O}_{B}$, $\mathcal{O}_{BB}$

Maximum sensitivity in WBF, $h \to \tau\tau$

- Fisher information at the SM:

$$
I_{ij}(\mathbf{0}) = \begin{array}{c c}
\begin{array}{ccccc} \mathcal{O}_{\phi,2} & \mathcal{O}_W & \mathcal{O}_{WW} & \mathcal{O}_B & \mathcal{O}_{BB} \end{array} & \\
\begin{pmatrix}
3202 & -625 & -7 & -35 & 0 \\
-625 & 451 & -110 & 23 & -2 \\
-7 & -110 & 244 & -6 & 3 \\
-35 & 23 & -6 & 4 & 0 \\
0 & -2 & 3 & 0 & 0
\end{pmatrix} & \begin{array}{c} \mathcal{O}_{\phi,2} \\ \mathcal{O}_W \\ \mathcal{O}_{WW} \\ \mathcal{O}_B \\ \mathcal{O}_{BB} \end{array}
\end{array}
$$

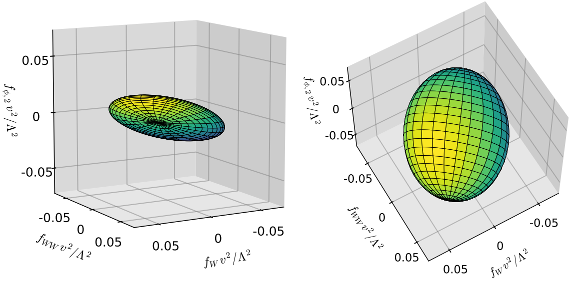
- Minimal errors $\Delta\theta \geq 1/\sqrt{I}$:
  - Largest eigenvalue along $\mathcal{O}_{\phi,2}$:

    $$\Delta(f/\Lambda^2) \gtrsim 0.3\ \text{TeV}^{-2}$$

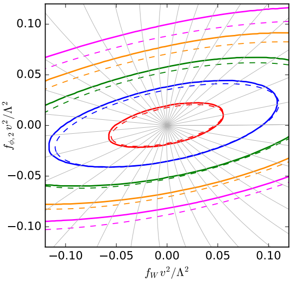
  - $\mathcal{O}_W$-$\mathcal{O}_{WW}$ plane:

    $$\Delta(f/\Lambda^2) \gtrsim 1.0\ \text{TeV}^{-2}$$

  - Large mixing



2. Does the EFT series converge?



A hierarchy of scales?

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

- EFT approach based on $E^2/\Lambda^2 \ll 1$
- Test this scale hierarchy!
  - Limit momentum flow with cut

    $$E \sim p_{T,j} < p_{T,j}^{\text{max}}$$

  - Precision

    $$\Delta(f\, v^2/\Lambda^2) = 1/\sqrt{I}$$

    corresponds to new physics reach

    $$\Lambda = \sqrt{f}\, v\, I^{1/4}$$

A hierarchy of scales?

- EFT approach based on $E^2/\Lambda^2 \ll 1$

- Test this scale hierarchy!
  - Limit momentum flow with cut

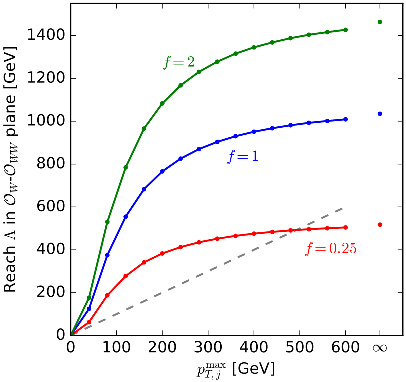
$$E \sim p_{T,j} < p_{T,j}^{\text{max}}$$

  - Precision

$$\Delta(f\, v^2/\Lambda^2) = 1/\sqrt{I}$$

corresponds to new physics reach

$$\Lambda = \sqrt{f}\; v\; I^{1/4}$$



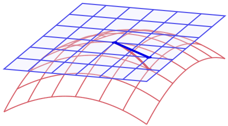
Information geometry

- Geometric interpretation: [C. R. Rao 1945, S. Amari 1968; ...]
  - Parameter space of theory $\rightsquigarrow$ manifold
  - Parametrization θ_i $\rightsquigarrow$ map (coordinates)
  - Fisher information I_{ij} $\rightsquigarrow$ Riemannian metric

- Distance measures:
  - Local / tangent space at $\boldsymbol{\theta}_0$:

$$d_{\text{local}}(\boldsymbol{\theta};\boldsymbol{\theta}_0) = \sqrt{I_{ij}(\boldsymbol{\theta}_0)\,(\theta^i - \theta_0^i)\,(\theta^j - \theta_0^j)}$$

$\sim$ unlikeliness to measure $\boldsymbol{\theta}$ if $\boldsymbol{\theta}_0$ is true, 'in sigmas'



  - Global along geodesics:

$$d(\boldsymbol{\theta}_a,\boldsymbol{\theta}_b) = \min_{\boldsymbol{\theta}(s)} \int_{s_a}^{s_b} \mathrm{d}s\,\sqrt{I_{ij}(\boldsymbol{\theta})\,\frac{\mathrm{d}\theta_i(s)}{\mathrm{d}s}\,\frac{\mathrm{d}\theta_j(s)}{\mathrm{d}s}}$$

Geometry of effective field theories

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

- Remember

$$I_{ij}(\boldsymbol{\theta}) = L \sum_{\text{events}} \frac{\partial \Delta\sigma(\boldsymbol{\theta})}{\partial \theta_i} \frac{1}{\Delta\sigma(\boldsymbol{\theta})} \frac{\partial \Delta\sigma(\boldsymbol{\theta})}{\partial \theta_j}$$

$\Rightarrow$ $I_{ij}(\mathbf{0})$ only sensitive to linear effects $\Delta\sigma \sim \theta_i \Delta\sigma_i$

Geometry of effective field theories

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

- Remember

$$I_{ij}(\boldsymbol{\theta}) = L \sum_{\text{events}} \frac{\partial \Delta\sigma(\boldsymbol{\theta})}{\partial \theta_i} \frac{1}{\Delta\sigma(\boldsymbol{\theta})} \frac{\partial \Delta\sigma(\boldsymbol{\theta})}{\partial \theta_j}$$

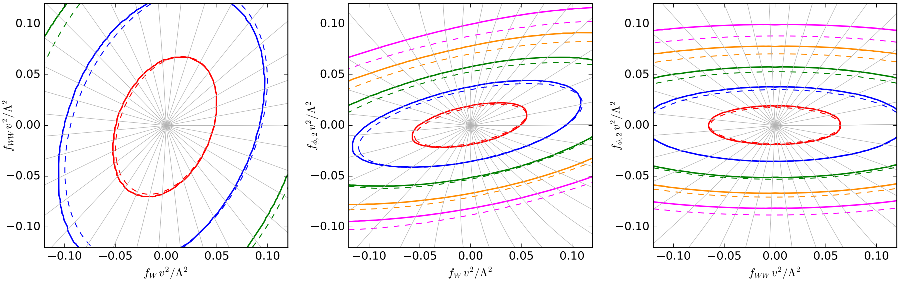
$\Rightarrow$ $I_{ij}(\mathbf{0})$ only sensitive to linear effects $\Delta\sigma \sim \theta_i \Delta\sigma_i$

- Information geometry for dimension-6 operators, $\theta_i = f_i^{d=6} v^2/\Lambda^2$:

$$\Delta\sigma = \underbrace{\underbrace{\Delta\sigma_{SM} + \sum_i \frac{f_i^{d=6}}{\Lambda^2} \Delta\sigma_i}_{I_{ij}(\mathbf{0}),\ \text{local distances at SM}} + \sum_{i,j} \frac{f_i^{d=6} f_j^{d=6}}{\Lambda^4} \Delta\sigma_{ij}}_{I_{ij}(\boldsymbol{\theta} \neq 0),\ \text{global distances}} + \underbrace{\sum_k \frac{f_k^{d=8}}{\Lambda^4} \Delta\sigma_k}_{\text{always missing}} + \mathcal{O}\left(1/\Lambda^6\right)$$

$\Rightarrow$ Difference between local and global distances $\leftrightarrow$ size of $\mathcal{O}\left(1/\Lambda^4\right)$ effects

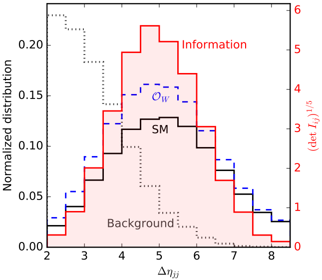
Global vs local distances for WBF



Contours of local (dashed) and global (solid) distances $d = 1, 2, 3, \ldots$ from SM

Other parameters set to zero

3. Where in phase space is the information?



The differential information

- Differential information with respect to any observable:

$$
\begin{aligned}
I_{ij}(\boldsymbol{\theta}) &= \sum_{\text{events}} \frac{\partial \Delta\sigma}{\partial \theta_i} \frac{1}{\Delta\sigma} \frac{\partial \Delta\sigma}{\partial \theta_j} \\
&= \sum_{\text{bins } b} \underbrace{L \sum_{\text{events in } b} \frac{\partial \Delta\sigma}{\partial \theta_i} \frac{1}{\Delta\sigma} \frac{\partial \Delta\sigma}{\partial \theta_j}}_{\equiv I_{ij}^b(\boldsymbol{\theta})}
\end{aligned}
$$

- Defines most important phase-space regions

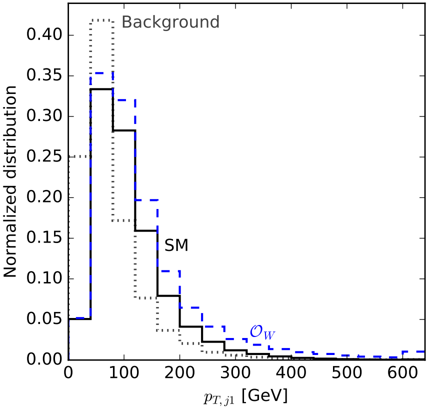
Differential information over $p_{T,j}$

Strongly correlated with momentum transfer E through production vertex: measures $\mathcal{O} \sim \partial^2/\Lambda^2 \sim E^2/\Lambda^2$

SM

$f_W/\Lambda^2\, v^2 = 0.5$

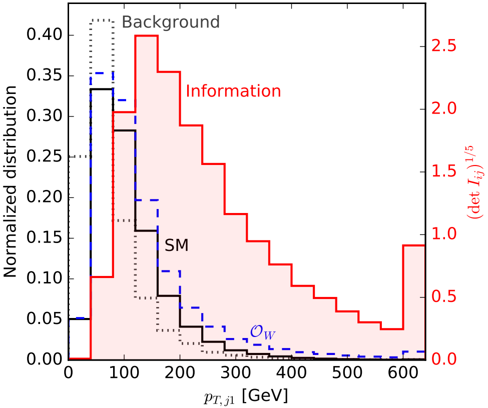
QCD $Z \to \tau\tau$



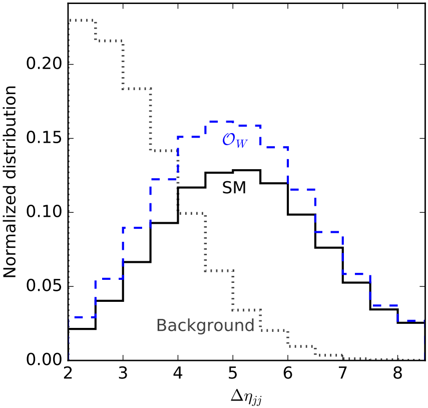
Differential information over $p_{T,j}$

Strongly correlated with momentum transfer E through production vertex:
measures $\mathcal{O} \sim \partial^2/\Lambda^2 \sim E^2/\Lambda^2$

SM
$f_W/\Lambda^2\, v^2 = 0.5$
QCD $Z \to \tau\tau$
$\det I_{ij}(\mathbf{0})$



Differential information over $\Delta\eta_{jj}$



Trade-off:

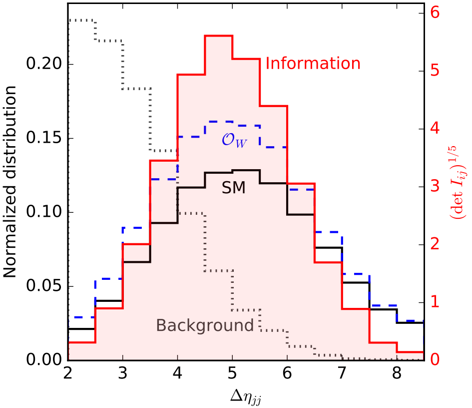
- Background suppression better at large $\Delta\eta_{jj}$
- Momentum-dependent operators have largest effects at medium $\Delta\eta_{jj}$

SM

$f_W/\Lambda^2\, v^2 = 0.5$

QCD $Z \to \tau\tau$

Differential information over $\Delta\eta_{jj}$



Trade-off:

- Background suppression better at large $\Delta\eta_{jj}$
- Momentum-dependent operators have largest effects at medium $\Delta\eta_{jj}$

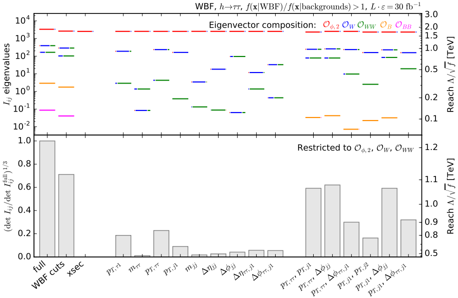
SM

$f_W/\Lambda^2\, v^2 = 0.5$

QCD $Z \to \tau\tau$

$\det I_{ij}(\mathbf{0})$

4. What are the most powerful observables?



Information in individual distributions

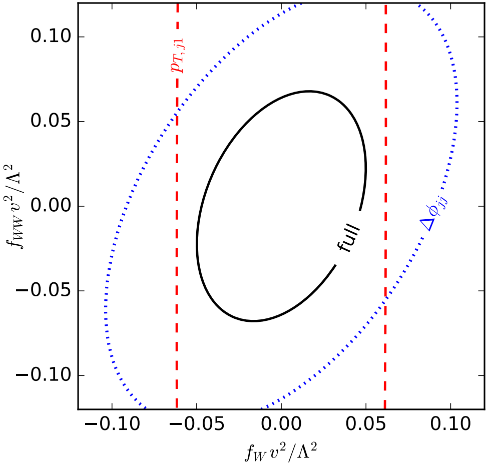
- Reduced information in histogram (rather than in full kinematics):

$$I_{ij}^{\text{distribution}}(\boldsymbol{\theta}) = \sum_{\text{bins } b} L \, \frac{\partial \sigma_b(\boldsymbol{\theta})}{\partial \theta_i} \, \frac{1}{\sigma_b(\boldsymbol{\theta})} \, \frac{\partial \sigma_b(\boldsymbol{\theta})}{\partial \theta_j}$$

- Defines most powerful observables
- Allows to compare histogram-based and multivariate analyses

WBF observables (1)

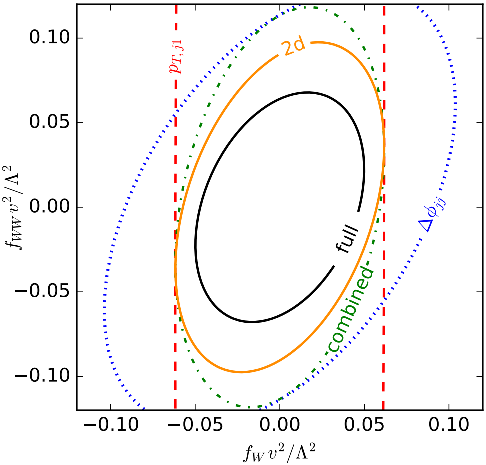
- Virtuality measure probes only $\mathcal{O}_W$
- Angular correlation between tagging jets sensitive to $\mathcal{O}_{WW}$



Other parameters set to zero, likelihood-based event selection

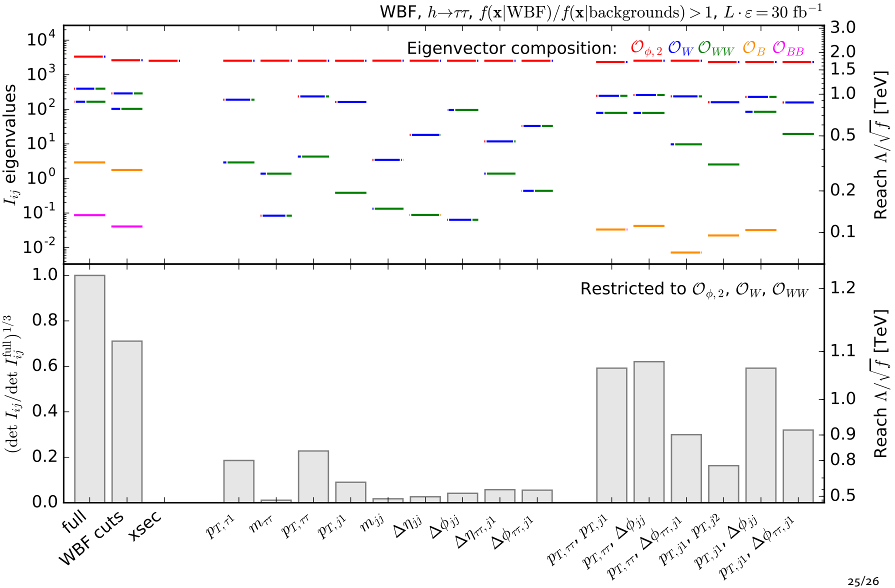
WBF observables (1)

- Virtuality measure probes only $\mathcal{O}_W$
- Angular correlation between tagging jets sensitive to $\mathcal{O}_{WW}$
- Two-dimensional histo necessary for stringent constraints, but still not close to full information



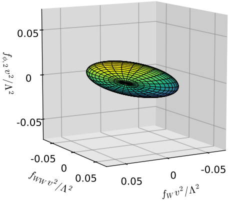
Other parameters set to zero, likelihood-based event selection

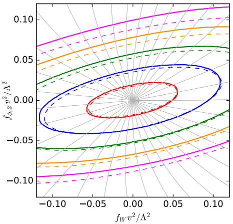
WBF observables (2)

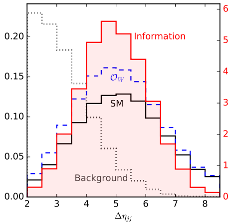


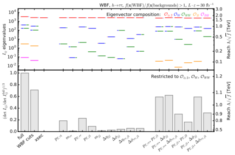
Conclusions

Information geometry lets us…







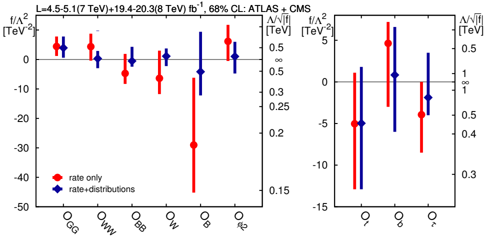


- calculate the maximum sensitivity of any LHC process
- discuss the EFT convergence
- select the important phase-space regions
- define the most powerful observables, compare them to multivariate methods

Bonus material

Sensitivity vs validity

- Run I fit: [T. Corbett, O. Eboli, D. Goncalves, J. Gonzalez-Fraile, T. Plehn, M. Rauch 1505.05516]

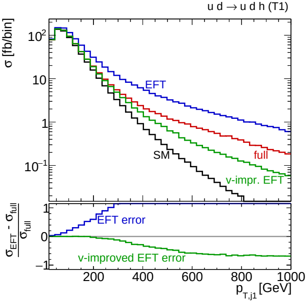


- Kinematic information up to $E \lesssim 400$ GeV crucial
- Sensitive to NP scales $\Lambda \sim \sqrt{f} \cdot 400$ GeV

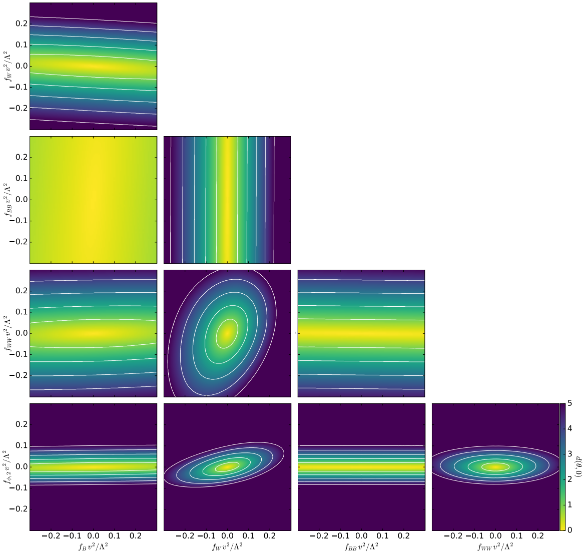
- Is the dimension-six model still useful?
  - Strongly coupled NP: works fine
  - Weakly coupled NP: no guarantee, but works in many scenarios (with v-improved matching)

    [JB, A. Freitas, D. Lopez-Val, T. Plehn 1510.03443]
  - EFT less reliable in high-energy tails

[See YR4 and references therein ;), note similar questions for DM EFTs]



WBF distances

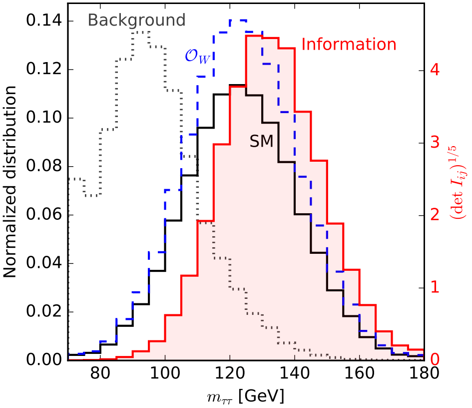


Distances from SM $d(\boldsymbol{\theta}, \mathbf{0})$

Optimal precision ($d = 1$):

$$\Delta(f_{\phi,2}\, v^2/\Lambda^2) \approx 0.02$$
$$\Delta(f_W\, v^2/\Lambda^2) \approx 0.05$$
$$\Delta(f_{WW}\, v^2/\Lambda^2) \approx 0.05$$

Differential information over $m_{\tau\tau}$



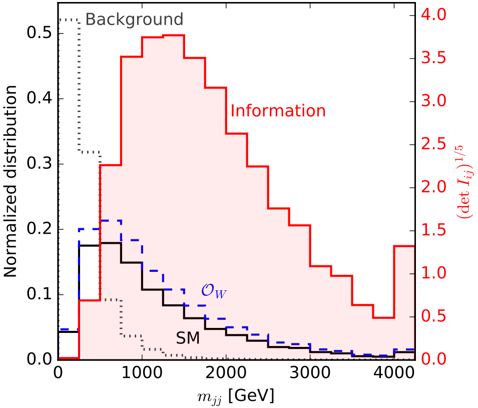
SM

$f_W/\Lambda^2\, v^2 = 0.5$

QCD $Z \to \tau\tau$

$\det I_{ij}(\mathbf{0})$

Differential information over m_{jj}



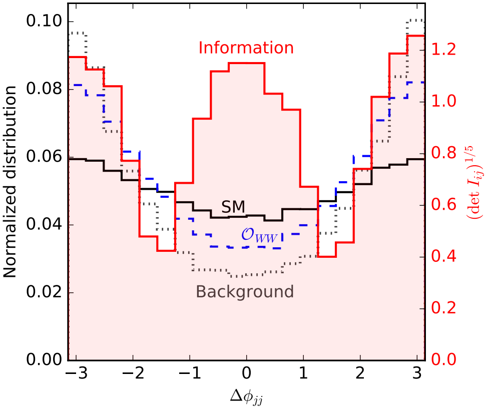
SM

$f_W/\Lambda^2\, v^2 = 0.5$

QCD $Z \to \tau\tau$

$\det I_{ij}(\mathbf{0})$

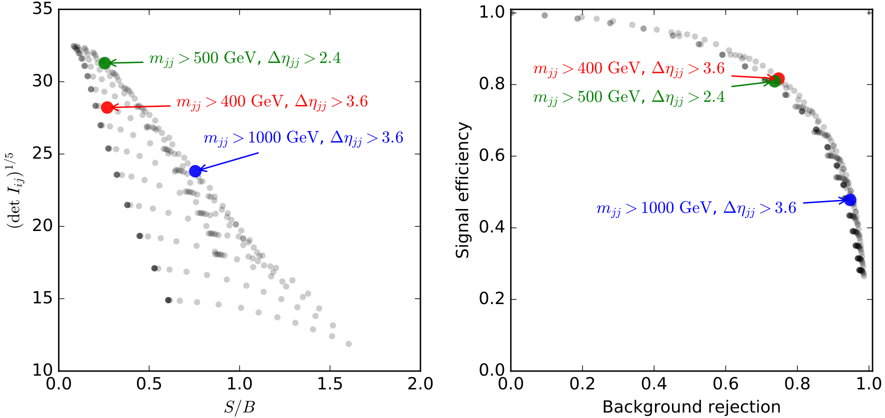
Differential information over $\Delta\phi_{jj}$



SM
$f_{WW}/\Lambda^2\,v^2 = 0.5$
QCD $Z \to \tau\tau$
$\det I_{ij}(\mathbf{0})$

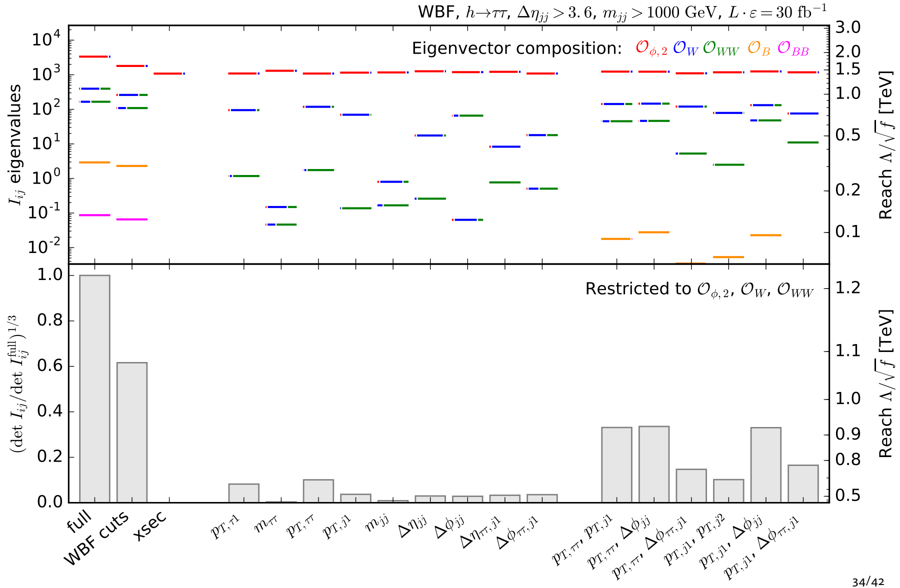
Optimizing cuts

- Scan over m_{jj} and $\Delta\eta_{jj}$ cuts $\rightsquigarrow$ signal and background rate, $I_{ij}(\mathbf{0})$
- Trade-off between information and purity (left)
- Standard ROC curves (right) can be misleading



Common cuts: 105 GeV $< m_{\tau\tau} <$ 165 GeV, $p_{T,j1} >$ 50 GeV

WBF observables after conventional cuts



Adding systematic uncertainties

Procedure:

- Add nuisance parameter to Fisher information:

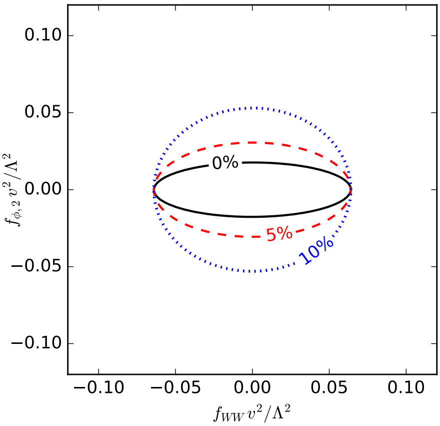
$$I_{ij} = \begin{pmatrix} I_t & I_m^T \\ I_m & I_n \end{pmatrix}$$

- Profiled information:

$$I_{\text{profiled}} = I_t - I_m^T \, I_n^{-1} \, I_m$$

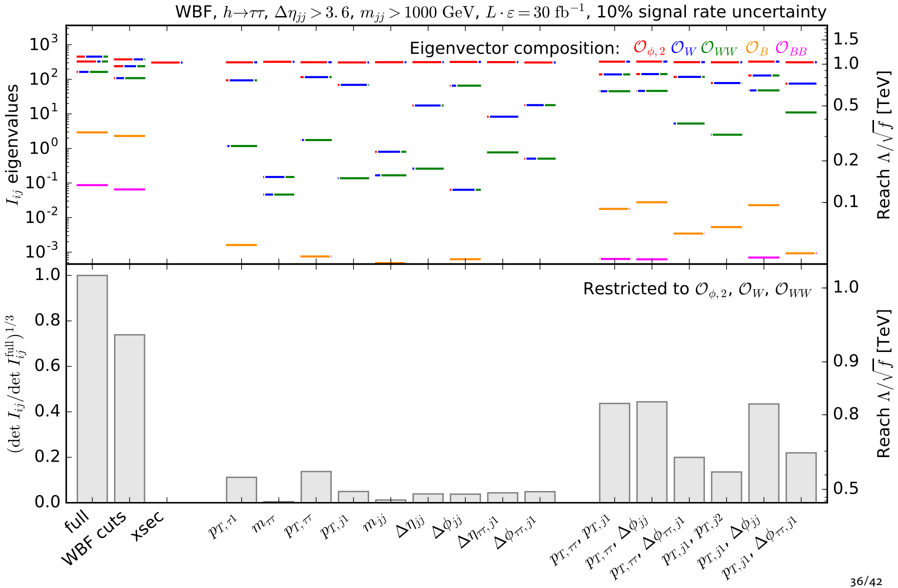
[T. Edwards, C. Weniger 1704.05458]

Local distances from SM, profiled over Gaussian uncertainties of 5% or 10% on signal rate:



Other parameters set to zero

WBF observables with systematics

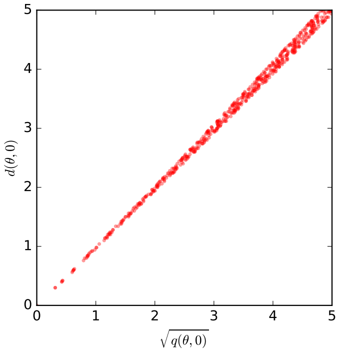


Fisher information vs likelihood ratio

- Confidence intervals based on hypothesis tests with likelihood ratio: are the Fisher information results relevant?

- Check!
  - Sample points $\boldsymbol{\theta}$ in $\mathcal{O}_W$-$\mathcal{O}_{WW}$ plane
  - Compare information distance $d(\boldsymbol{\theta}, \mathbf{0})$ to expected log likelihood ratio

$$q(\boldsymbol{\theta}|\mathbf{0}) = E\left[-2\,\log\frac{f(\mathbf{x}|\boldsymbol{\theta})}{f(\mathbf{x}|\mathbf{0})}\middle|\mathbf{0}\right]$$



$\Rightarrow$ Conclusions from information approach should also apply to limit setting

Ambient Fisher distance

- Distance measure in space of all distributions:

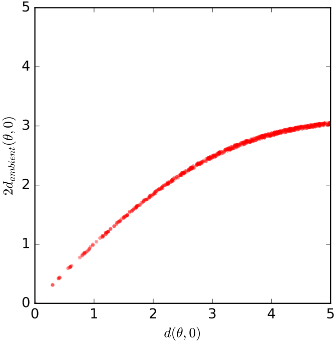
[J. Streets; ...]

$$d_{\text{ambient}}(f_a, f_b) = \arccos \int \mathrm{d}\mathbf{x} \sqrt{f_a(\mathbf{x}) f_b(\mathbf{x})}$$

- Comparison with 'our' information distances (as in last slide):

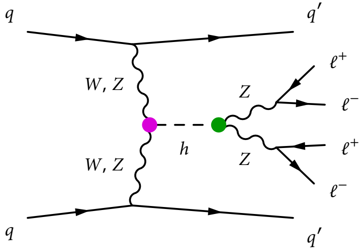
$$2\, d_{\text{ambient}}(f(\mathbf{x}|\boldsymbol{\theta}_a), f(\mathbf{x}|\boldsymbol{\theta}_b)) \sim d(\boldsymbol{\theta}_a, \boldsymbol{\theta}_b)$$

at small distances,
in agreement with literature
[K. Carter, R. Raich, W. Finn, A. Hero 0802.2050]



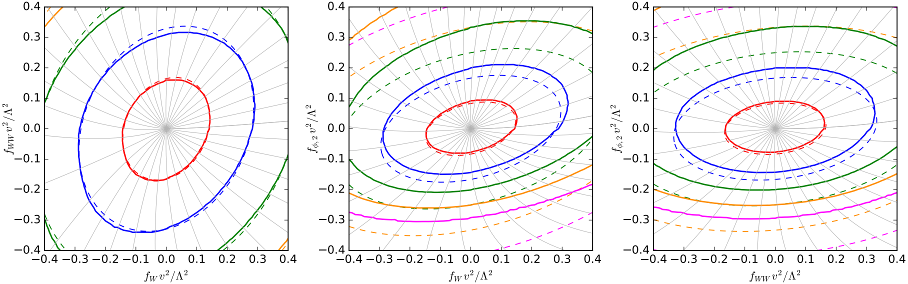
Weak boson fusion, $h \to 4\ell$

- Production vs decay



- Setup as for $h \to \tau\tau$, except:
  - No backgrounds, no smearing
  - $L \cdot \varepsilon = 100\ \text{fb}^{-1}$
  - Cuts: $p_{T,j} > 20$ GeV, $|\eta_j| < 5.0$, $p_{T,\ell} > 10$ GeV, $|\eta_\ell| < 2.5$

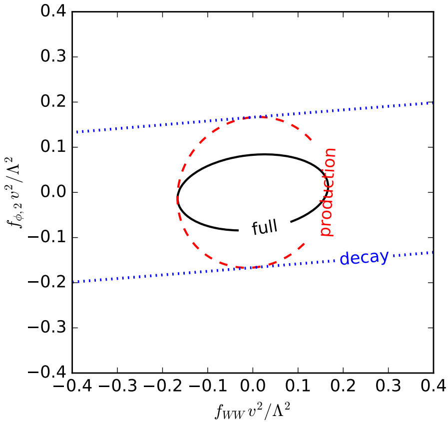
Information in WBF, $h \to 4\ell$



Contours of local (dashed) and global (solid) distances $d = 1, 2, 3, \ldots$ from SM

Other parameters set to zero

Production vs decay



Operator effects limited to production / decay vertex

Observables in WBF, $h \to 4\ell$