Higgs physics with information geometry

Johann Brehmer

Heidelberg University

Work in progress with Kyle Cranmer, Felix Kling, and Tilman Plehn

Higgs Couplings 2016

1. Information what?

Statistics for the Higgs

	Higgs physics, circa 2012
Physics question	Higgs or no Higgs (also, its mass)
Statistics question	Hypothesis test (for each m_h)
Performance bound	Likelihood ratio (Neyman-Pearson)

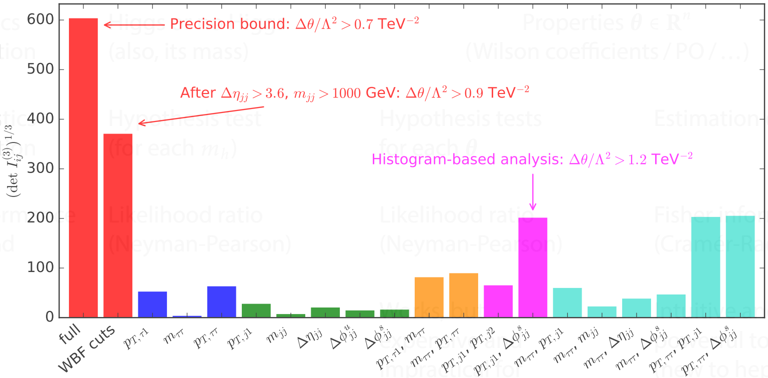
Statistics for the Higgs

	Higgs physics, circa 2012	Higgs physics, circa 2016	
Physics question	Higgs or no Higgs (also, its mass)	Properties $\boldsymbol{\theta} \in \mathbf{R}^n$ (Wilson coefficients / PO / …)	
Statistics question	Hypothesis test (for each m_h)	Hypothesis tests for each $\boldsymbol{\theta}$	Estimation of $\boldsymbol{\theta}$
Performance bound	Likelihood ratio (Neyman-Pearson)	Likelihood ratio (Neyman-Pearson)	Fisher information (Cramér-Rao)
		Works, but expensive and impractical for planning measurements	Intuitive and powerful tools (new to hep)

Statistics for the Higgs

Higgs physics, circa 2012 | Higgs physics, circa 2016

Sneak preview:



planning measurements

What can we learn from an experiment?

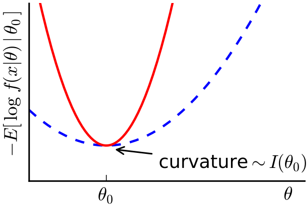
- Starting point: probability distribution $f\Big(\underbrace{\mathbf{x}}_{\text{observables}} \Big| \underbrace{\boldsymbol{\theta}}_{\text{theory parameters}}\Big)$

"All models are wrong, but some are useful." — H. G. Wells

- Fisher information

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

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



What can we learn from an experiment?

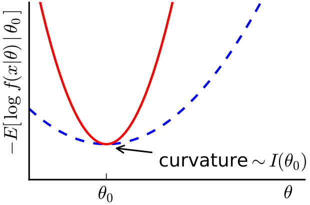
- Starting point: probability distribution $f(\underbrace{\mathbf{x}}_{\text{observables}} \mid \underbrace{\boldsymbol{\theta}}_{\text{theory parameters}})$

"All models are wrong, but some are useful." — H. G. Wells

- Fisher information

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

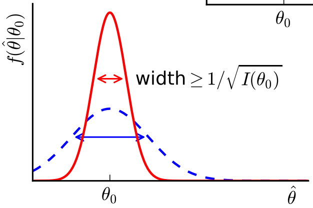
[F. Edgeworth 1908; R. Fisher 1925; …]



- 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

More on the Fisher information

- Properties of the Fisher information $I_{ij}(\boldsymbol{\theta}) = -E\left[\partial_i \partial_j \log f(\mathbf{x}|\boldsymbol{\theta}) \middle| \boldsymbol{\theta}\right]$:
  - Additive between experiments / phase-space regions
  - Invariant under reparametrizations of observables $\mathbf{x}$
  - Covariant under reparametrizations of theory parameters $\boldsymbol{\theta}$
- Can be calculated from MC samples

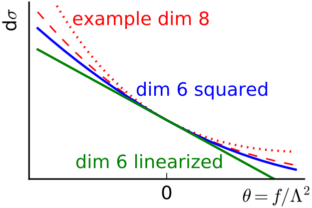
More on the Fisher information

- Properties of the Fisher information $I_{ij}(\boldsymbol{\theta}) = -E\left[\partial_i \partial_j \log f(\mathbf{x}|\boldsymbol{\theta}) \middle| \boldsymbol{\theta}\right]$:
  - Additive between experiments / phase-space regions
  - Invariant under reparametrizations of observables $\mathbf{x}$
  - Covariant under reparametrizations of theory parameters $\boldsymbol{\theta}$
- Can be calculated from MC samples
- Information geometry: [C. R. Rao 1945, S. Amari 1968; …]
  - Parameter space of theory $\rightsquigarrow$ manifold
  - Parametrization θ_i $\rightsquigarrow$ map (coordinates)
  - Fisher information $\rightsquigarrow$ Riemannian metric
- Distance measures:
  - Local / tangent space at $\boldsymbol{\theta}_0$: $\sqrt{I_{ij}(\boldsymbol{\theta}_0)(\theta^i - \theta_0^i)(\theta^j - \theta_0^j)}$
    - Gaussian limit: corresponds to maximal expected significance (in sigmas) with which $\boldsymbol{\theta}_0$ can be excluded if $\boldsymbol{\theta}$ is true
  - Global along geodesics: $d(\boldsymbol{\theta}_a, \boldsymbol{\theta}_b) = \min_{\boldsymbol{\theta}(s)} \int_{s_a}^{s_b} \mathrm{d}s \sqrt{I_{ij} \frac{\mathrm{d}\theta^i(s)}{\mathrm{d}s} \frac{\mathrm{d}\theta^j(s)}{\mathrm{d}s}}$

Effective field theory

- New physics at $\Lambda \gg E$?

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



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

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

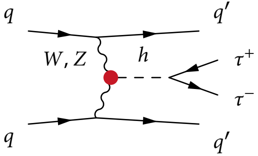
  - Curvature $\leftrightarrow$ size of $\mathcal{O}\left(1/\Lambda^4\right)$ effects
  - Depending on UV physics, this can give hints about EFT validity

2. Application to weak boson fusion

Weak boson fusion, $h \to \tau\tau$

- Well-known probe of Higgs-gauge structure
  - Interesting kinematics of tagging jets

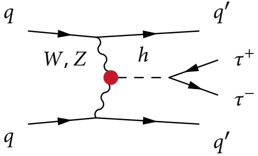
[D. Rainwater, D. Zeppenfeld, K. Hagiwara hep-ph/9808468;
T. Plehn, D. Rainwater, D. Zeppenfeld hep-ph/0105325;
C. Englert, D. Gonçalves-Netto, K. Mawatari, T. Plehn 1212.0843; …]



Weak boson fusion, $h \to \tau\tau$

- Well-known probe of Higgs-gauge structure
  - Interesting kinematics of tagging jets

[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; …]



- Theory language: dimension-6 operators of SM EFT, $\mathcal{L} \supset \sum_i \frac{f_i}{\Lambda^2}\mathcal{O}_i$

[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; …]

  - Total rate: $$\mathcal{O}_{\phi,2} = \frac{1}{2}\,\partial^\mu(\phi^\dagger\phi)\,\partial_\mu(\phi^\dagger\phi)$$
  - New kinematic structures:

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

$$\mathcal{O}_{BB} = -\frac{g'^2}{4}\,(\phi^\dagger\phi)\,B_{\mu\nu}\,B^{\mu\nu} \qquad \mathcal{O}_{WW} = -\frac{g^2}{4}\,(\phi^\dagger\phi)\,W^k_{\mu\nu}\,W^{\mu\nu\,k}$$

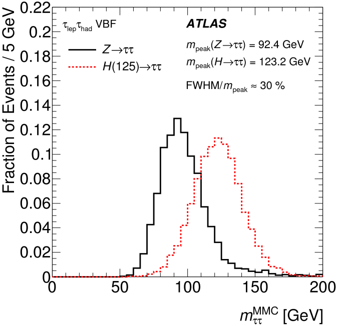
  - CP violation: $$\mathcal{O}_{W\widetilde{W}} = -\frac{g^2}{4}\,(\phi^\dagger\phi)\,W^k_{\mu\nu}\,\widetilde{W}^{\mu\nu\,k}$$
  - Others strongly constrained by EWPD or redundant

Setup

- MC: `MadMax` / `MadGraph 5`
  - Just one run necessary for exact full geometry

[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]
- $\sqrt{s} = 13$ TeV, $L \cdot \varepsilon = 30\ \text{fb}^{-1}$
- Cuts: $p_{T,j} > 20$ GeV, $|\eta_j| < 5.0$, $\Delta\eta_{jj} > 2.0$, $\Delta R_{jj} > 0.4$

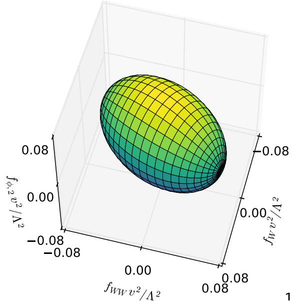
Fisher information at SM

- Simple matrix that describes sensitivity to theory directions:

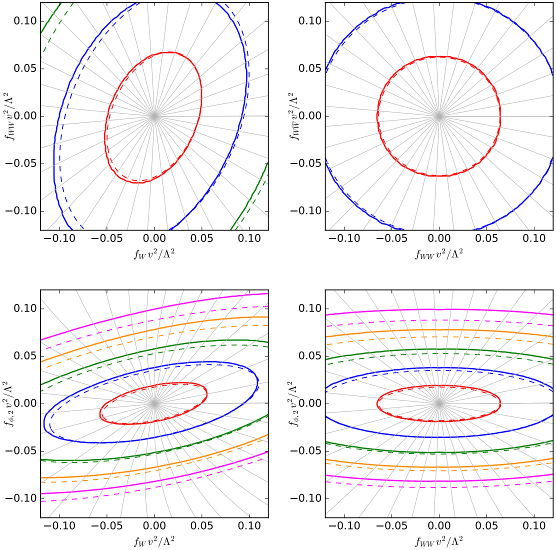
$$I_{ij}(\mathbf{0}) = \begin{array}{c} \begin{array}{cccccc} \mathcal{O}_B & \mathcal{O}_W & \mathcal{O}_{BB} & \mathcal{O}_{WW} & \mathcal{O}_{\phi,2} & \mathcal{O}_{W\widetilde{W}} \end{array} \\ \begin{pmatrix} 4 & 23 & 0 & -6 & -35 & 0 \\ 23 & \color{blue}{451} & -2 & \color{green}{-109} & \color{green}{-625} & -1 \\ 0 & -2 & 0 & 3 & 0 & 0 \\ -6 & \color{green}{-109} & 3 & \color{blue}{244} & -9 & -0 \\ -35 & \color{green}{-625} & 0 & -9 & \color{red}{3213} & 2 \\ 0 & -1 & 0 & 0 & 2 & \color{blue}{257} \end{pmatrix} \end{array} \begin{array}{c} \\ \mathcal{O}_B \\ \mathcal{O}_W \\ \mathcal{O}_{BB} \\ \mathcal{O}_{WW} \\ \mathcal{O}_{\phi,2} \\ \mathcal{O}_{W\widetilde{W}} \end{array}$$

  - $\mathcal{O}_{\phi,2}$ direction can be measured best, followed by $\mathcal{O}_W$, $\mathcal{O}_{WW}$, and $\mathcal{O}_{W\widetilde{W}}$
  - Large mixing between operators
- Visualize as linearized distance contours

$$I_{ij}(\boldsymbol{\theta}_0)(\theta^i - \theta_0^i)(\theta^j - \theta_0^j) = d^2$$



Information geometry



Contours of distances $d = 1, 2, 3, \ldots$ from SM (solid: full geometry, dashed: linearized)

Geodesics

Geometry is approximately flat: small effects from $\mathcal{O}\left(1/\Lambda^4\right)$ terms

Other parameters set to zero

Where information comes from: $p_{T,j}$

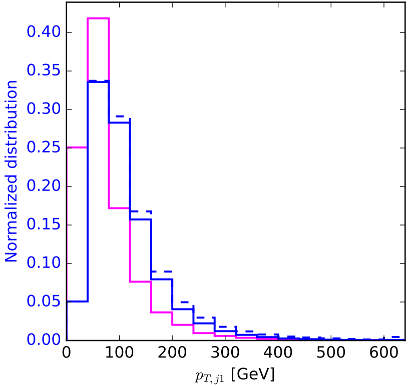
UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

Strongly correlated with virtuality of initial W / momentum transfer q through production vertex: measures $\mathcal{O} \sim \partial^2/\Lambda^2 \sim q^2/\Lambda^2$

Signal distribution

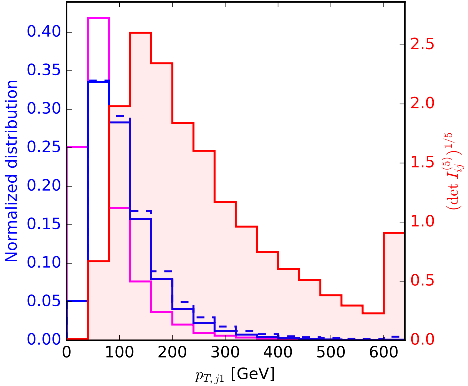
(dotted: $f_{WW}/\Lambda^2 \, v^2 = 0.5$)

QCD $Z \to \tau\tau$ distribution



Where information comes from: $p_{T,j}$

Strongly correlated with virtuality of initial W / momentum transfer q through production vertex: measures $\mathcal{O} \sim \partial^2/\Lambda^2 \sim q^2/\Lambda^2$



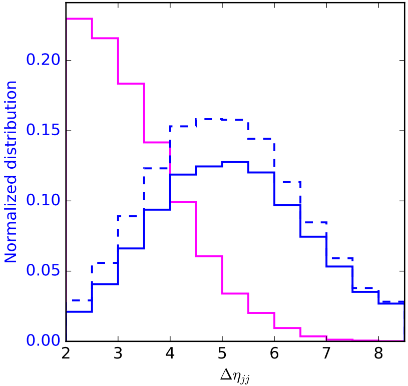
Signal distribution

(dotted: $f_{WW}/\Lambda^2\, v^2 = 0.5$)

QCD $Z \to \tau\tau$ distribution

$\det I_{ij}(\mathbf{0})$ over CP-even operators

Where information comes from: $\Delta\eta_{jj}$



Trade-off:

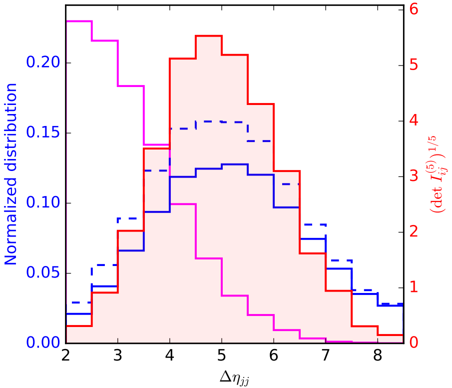
- Background suppression better at large $\Delta\eta_{jj}$
- Momentum-dependent operators have largest effects at medium $\Delta\eta_{jj}$

Signal distribution

(dotted: $f_W/\Lambda^2\, v^2 = 0.5$)

QCD $Z \to \tau\tau$ distribution

Where information comes from: $\Delta\eta_{jj}$



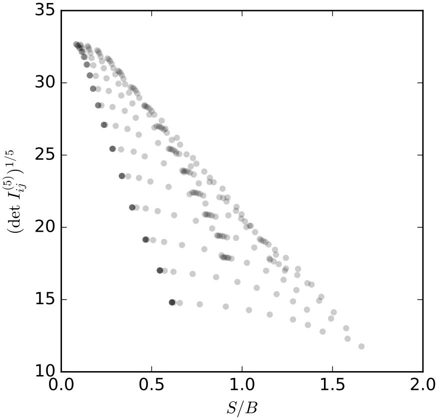
Trade-off:

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

Signal distribution
(dotted: $f_W/\Lambda^2 v^2 = 0.5$)
QCD $Z \to \tau\tau$ distribution
$\det I_{ij}(\mathbf{0})$ over CP-even operators

Optimizing cuts

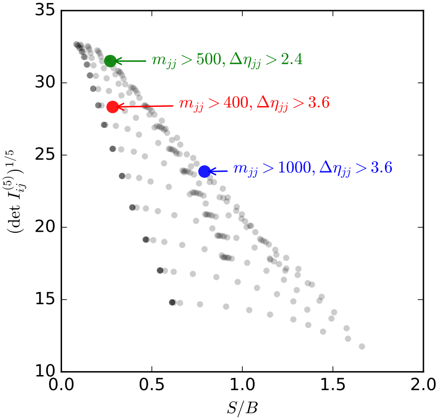
- Scan over m_{jj} and $\Delta\eta_{jj}$ cuts $\rightsquigarrow$ signal and background rate, $I_{ij}(\mathbf{0})$
- Trade-off between information and purity (left)

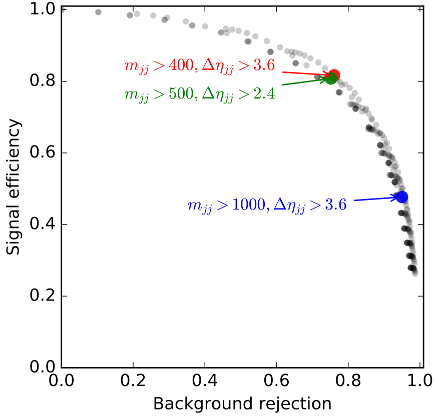


Common cuts: $105\ \text{GeV} < m_{\tau\tau} < 165\ \text{GeV}$, $p_{T,j1} > 50\ \text{GeV}$

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

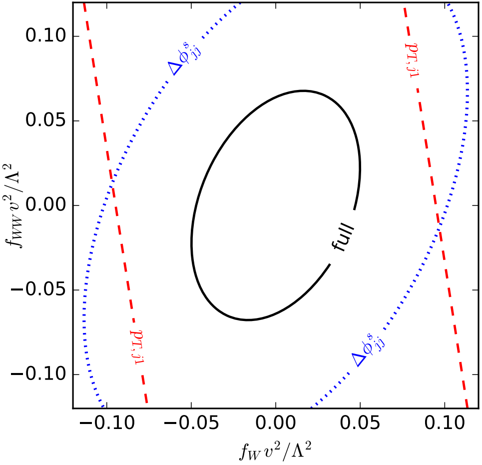




Common cuts: 105 GeV < $m_{\tau\tau}$ < 165 GeV, $p_{T,j1}$ > 50 GeV

Information in individual distributions

- How much information is preserved when projecting the full kinematics into histograms?
- Virtuality measure and angular correlations sensitive to different directions

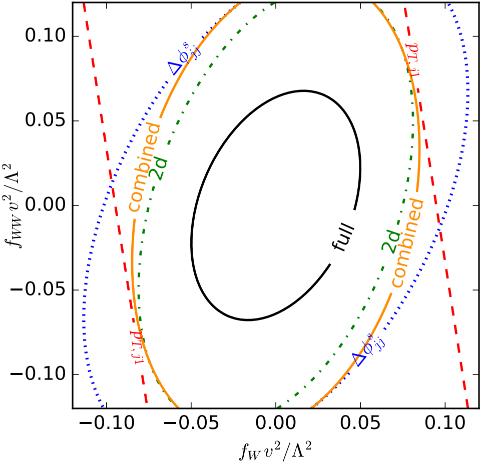


Other parameters set to zero

Information in individual distributions

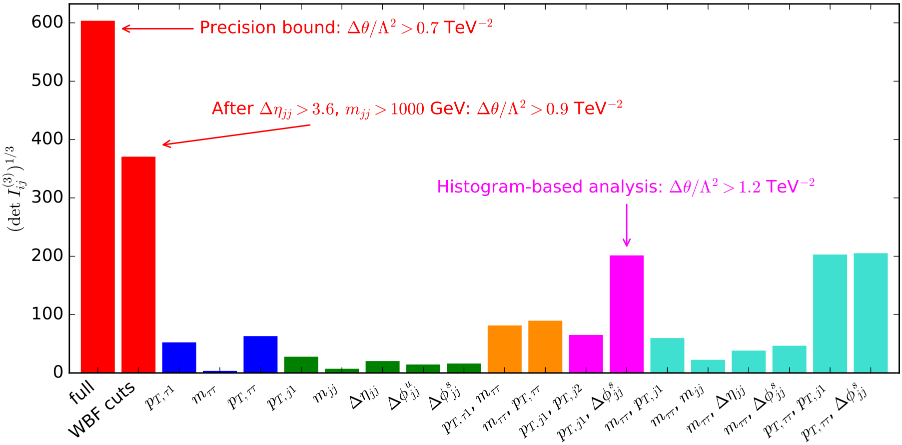
UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

- How much information is preserved when projecting the full kinematics into histograms?
- Virtuality measure and angular correlations sensitive to different directions
- Two-dimensional histos carry more information, but still not close to maximum



Other parameters set to zero

Comparing distributions



Determinant over $\mathcal{O}_W, \mathcal{O}_{WW}, \mathcal{O}_{\phi,2}$

Conclusions: statistics

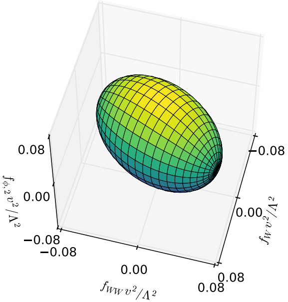
- The Fisher information $I_{ij}(\boldsymbol{\theta}) = -E\left[\partial_i \partial_j \log f(\mathbf{x}|\boldsymbol{\theta}) \middle| \boldsymbol{\theta}\right]$ is…
  - important: limit of knowledge we can derive from measurement
  - compact: simple matrix for each $\boldsymbol{\theta}$
  - well-behaved: additive between measurements, phase-space regions; invariant / covariant under reparametrization
  - intuitive: defines distances in theory space
  - easy to calculate: from MC samples

Conclusions: statistics

- The Fisher information $I_{ij}(\boldsymbol{\theta}) = -E\left[\partial_i \partial_j \log f(\mathbf{x}|\boldsymbol{\theta}) \middle| \boldsymbol{\theta}\right]$ is…
  - important: limit of knowledge we can derive from measurement
  - compact: simple matrix for each $\boldsymbol{\theta}$
  - well-behaved: additive between measurements, phase-space regions; invariant / covariant under reparametrization
  - intuitive: defines distances in theory space
  - easy to calculate: from MC samples
- With it we can…
  - analyse the distribution of information over phase space
    $\rightsquigarrow$ define fiducial regions
    $\rightsquigarrow$ focus attention for calibration and tuning
  - compare information in different distributions
    $\rightsquigarrow$ pick which differential xsec to measure
    $\rightsquigarrow$ design statistical analysis
  - check $\mathcal{O}\left(1/\Lambda^4\right)$ terms in EFT $\rightsquigarrow$ judge EFT validity

Conclusions: application to weak boson fusion

- Calculated Fisher information in $\tau\tau$ and 4ℓ channels
- Distribution of information:
  - Orthogonal information in rate and shapes
  - Kinematic information in high-energy tails
  - Careful with $\Delta\eta_{jj}$ cuts
- Compression into histograms:
  - Best: momentum transfer proxies, angular correlations between jets
  - Two-dimensional histograms more informative than one-dimensional
  - Histogram-based analysis only close to Cramér-Rao bound for clean 4ℓ channel
- WBF production carries much more information than on-shell decay
  - Momentum transfer not limited by Higgs mass



3. Backup

Cramér-Rao bound

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

- Any estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$ with expectation value $\bar{\boldsymbol{\theta}}(\boldsymbol{\theta})$ has

$$\operatorname{cov}\left[\hat{\boldsymbol{\theta}}\middle|\boldsymbol{\theta}\right]_{ab} \equiv E\left[(\hat{\theta}_a - \bar{\theta}_a)(\hat{\theta}_b - \bar{\theta}_b)\middle|\boldsymbol{\theta}\right] \geq \frac{\partial\bar{\theta}_a}{\partial\theta_i}\, I^{-1}_{ij}(\boldsymbol{\theta})\, \frac{\partial\bar{\theta}_b}{\partial\theta_j}$$

  - Unbiased estimators: $\bar{\boldsymbol{\theta}} = \boldsymbol{\theta}$, so

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

  - Unbiased estimator for one parameter:

$$\operatorname{var}\left[\hat{\theta}\middle|\theta\right] \geq \frac{1}{I(\theta)}$$

- Maximum-likelihood estimators saturate this bound asymptotically

Calculating the Fisher information

- Event count in one channel, $f(n|\nu) = \mathrm{Pois}(n|\nu)$:

$$I^{\text{xsec}}(\nu) = \frac{1}{\nu} \quad \Rightarrow \quad I^{\text{xsec}}(\theta) = \frac{\partial \nu}{\partial \theta} \frac{1}{\nu} \frac{\partial \nu}{\partial \theta}$$

- Histogram, bins with expectation ν_a:

$$I^{\text{histo}}_{ab}(\boldsymbol{\nu}) = \frac{\delta_{ab}}{\nu_a} \quad \Rightarrow \quad I^{\text{histo}}_{ij}(\boldsymbol{\theta}) = \sum_{\text{bins } a} \frac{\partial \nu_a}{\partial \theta_i} \frac{1}{\nu_a} \frac{\partial \nu_a}{\partial \theta_j}$$

- Extended likelihood, $f(\mathbf{x}|\boldsymbol{\theta}) = \mathrm{Pois}(n|\sigma(\boldsymbol{\theta})L) \; \prod_{i=1}^n f_1(x_i|\boldsymbol{\theta})$:

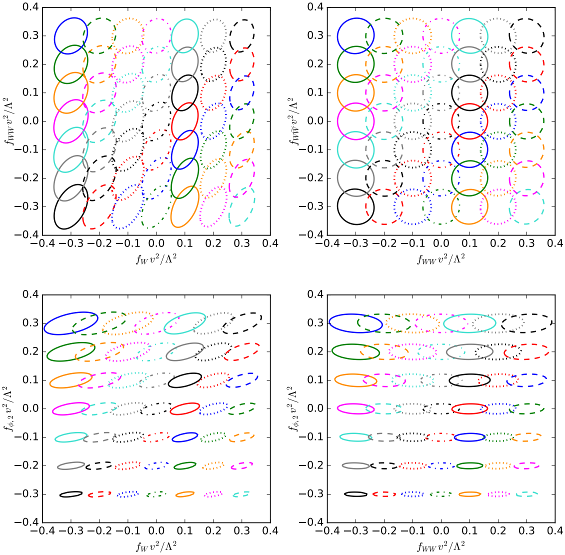
$$I_{ij}(\boldsymbol{\theta}) = L \frac{\partial \sigma(\boldsymbol{\theta})}{\partial \theta^i} \frac{1}{\sigma(\boldsymbol{\theta})} \frac{\partial \sigma(\boldsymbol{\theta})}{\partial \theta^j} + L\, \sigma(\boldsymbol{\theta})\, E\left[\frac{\partial \log f_1(x|\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \log f_1(x|\boldsymbol{\theta})}{\partial \theta_j} \middle| \boldsymbol{\theta} \right]$$

MC integration, $\int \mathrm{d}x \; f_1(x) \to \sum_{\text{events}} \frac{\Delta\sigma}{\sigma}$:

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

$\Rightarrow$ We can calculate the Fisher information from standard MC samples

Fisher information over theory space

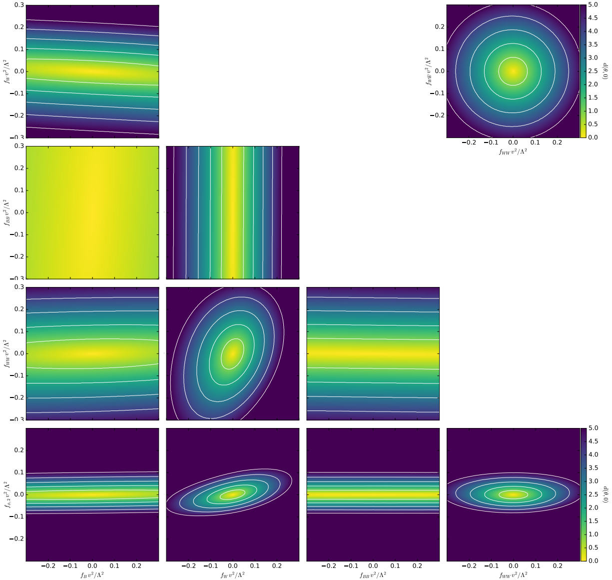


Linearized distance contours

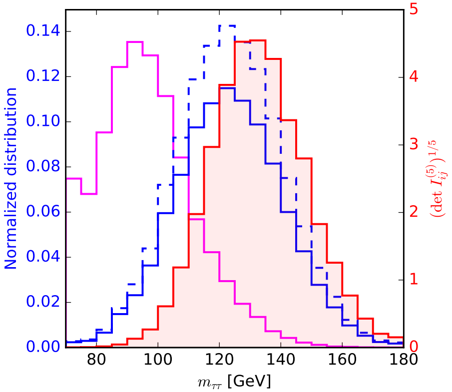
$$I_{ij}(\boldsymbol{\theta}_0)(\theta^i - \theta^i_0)(\theta^j - \theta^j_0) = 1$$

Other parameters set to zero

Distances from SM



Where information comes from: $m_{\tau\tau}$



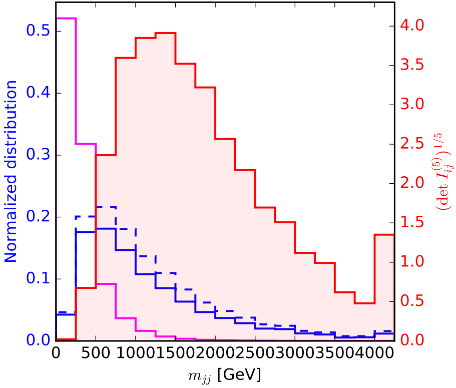
Signal distribution

(dotted: $f_W/\Lambda^2\, v^2 = 0.5$)

QCD $Z \to \tau\tau$ distribution

$\det I_{ij}(\mathbf{0})$ over CP-even operators

Where information comes from: m_{jj}



Signal distribution

(dotted: $f_W/\Lambda^2\, v^2 = 0.5$)

QCD $Z \to \tau\tau$ distribution

$\det I_{ij}(\mathbf{0})$ over CP-even operators

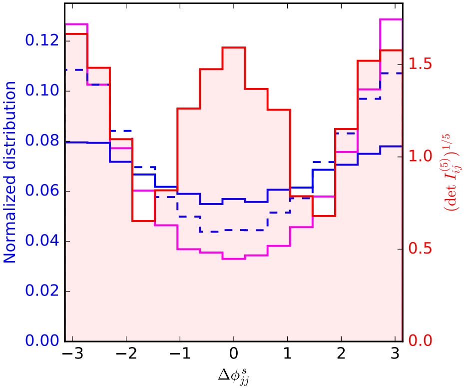
Where information comes from: $\Delta\phi_{jj}$ (1)

Signed version:

$$\Delta\phi^s_{jj} = \phi_{j(\eta<0)} - \phi_{j(\eta>0)}$$

CP-even operators: shifts to parallel / back-to-back geometry

Signal distribution
(dotted: $f_{WW}/\Lambda^2\, v^2 = 0.5$)
QCD $Z \to \tau\tau$ distribution
$\det I_{ij}(\mathbf{0})$ over CP-even operators

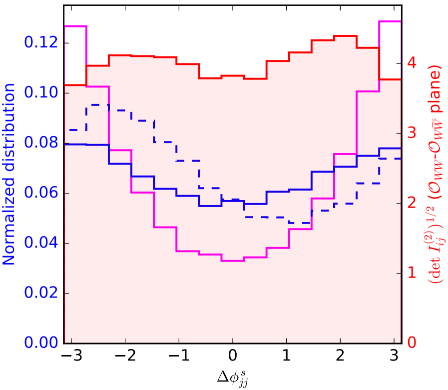


Where information comes from: $\Delta\phi_{jj}$ (2)

Signed version:

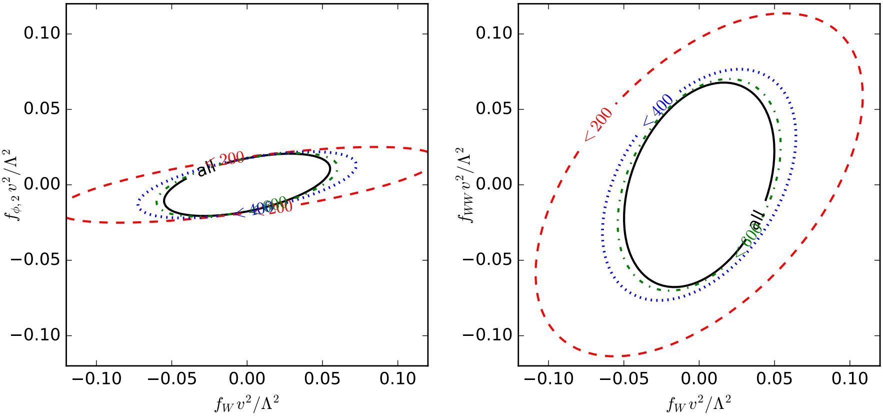
$$\Delta\phi^s_{jj} = \phi_{j^{\eta<0}} - \phi_{j^{\eta>0}}$$

CP-odd operators: sign important



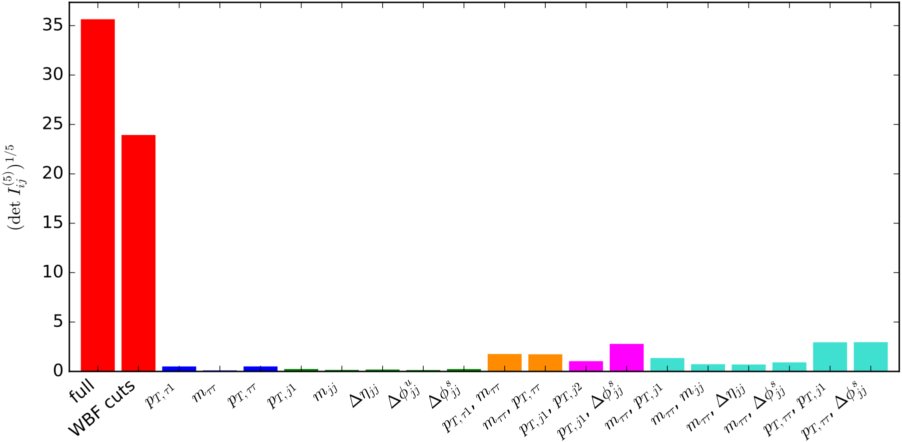
Signal distribution
(dotted: $f_{W\widetilde{W}}/\Lambda^2\, v^2 = 0.5$)
QCD $Z \to \tau\tau$ distribution
$\det I_{ij}(\mathbf{0})$ over $\mathcal{O}_{WW}$-$\mathcal{O}_{W\widetilde{W}}$ space

Limiting the jet p_T

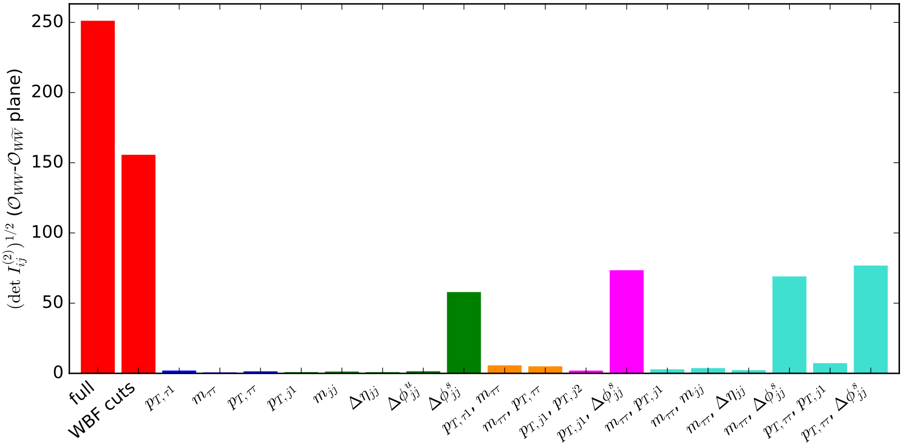


Upper limit on $p_{T,j}$ given along curves in GeV

Information in distributions: 5d

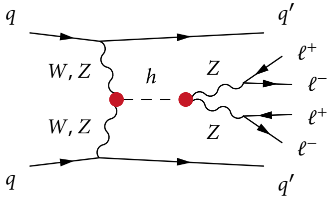


Information in distributions: CP odd



4. Physics in production and decay: $h \to 4\ell$

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



- Production vs decay
  - hZZ decay vertex: many angular structures
  - Very clean
- Same operators as before:

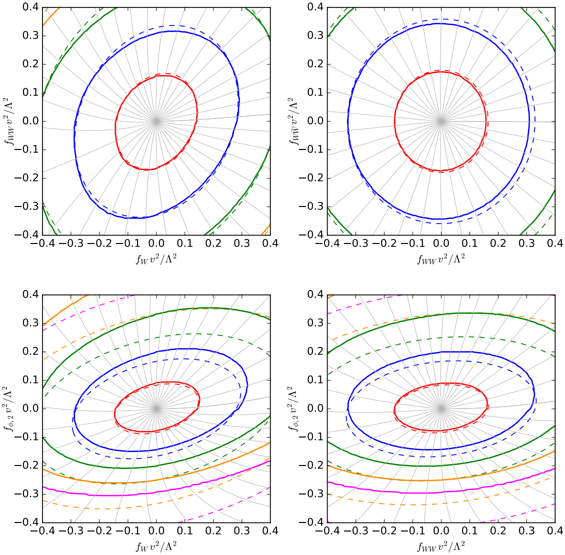
$$\mathcal{O}_B = \mathrm{i}\frac{g}{2}\,(D^\mu\phi^\dagger)(D^\nu\phi)\,B_{\mu\nu} \qquad \mathcal{O}_W = \mathrm{i}\frac{g}{2}\,(D^\mu\phi)^\dagger\sigma^k(D^\nu\phi)\,W^k_{\mu\nu}$$

$$\mathcal{O}_{BB} = -\frac{g'^2}{4}\,(\phi^\dagger\phi)\,B_{\mu\nu}\,B^{\mu\nu} \qquad \mathcal{O}_{WW} = -\frac{g^2}{4}\,(\phi^\dagger\phi)\,W^k_{\mu\nu}\,W^{\mu\nu\,k}$$

$$\mathcal{O}_{\phi,2} = \frac{1}{2}\,\partial^\mu(\phi^\dagger\phi)\,\partial_\mu(\phi^\dagger\phi) \qquad \mathcal{O}_{W\widetilde{W}} = -\frac{g^2}{4}\,(\phi^\dagger\phi)\,W^k_{\mu\nu}\,\widetilde{W}^{\mu\nu\,k}$$

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

Information geometry



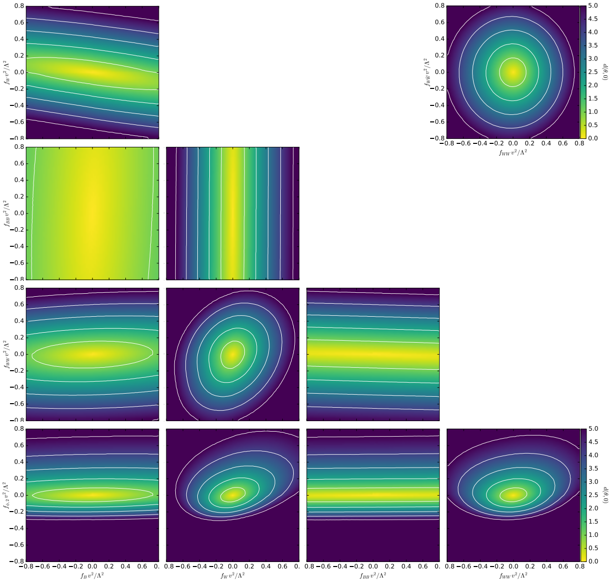
Contours of distances $d = 1, 2, 3, \ldots$ from SM (solid: full geometry, dashed: linearized)

Geodesics

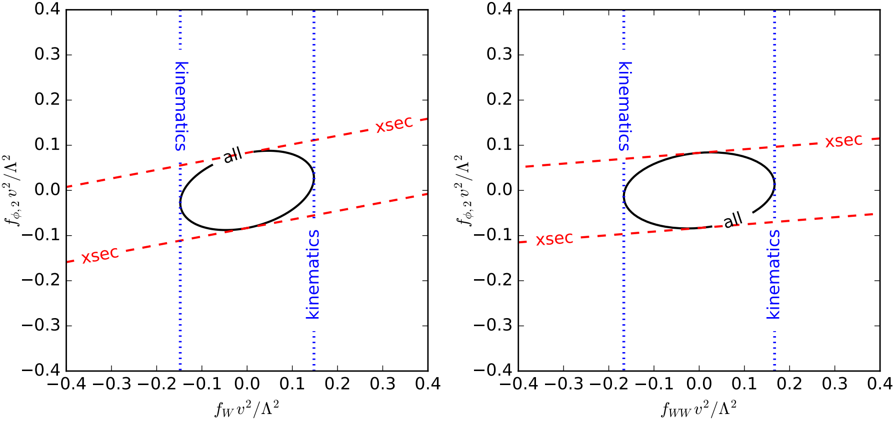
Larger effects from $\mathcal{O}(1/\Lambda^4)$ terms

Other parameters set to zero

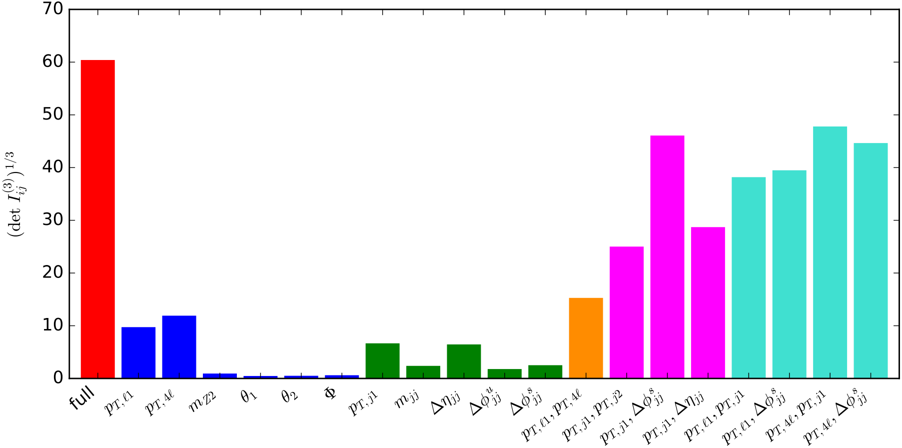
Distances from SM



Rate vs kinematics information

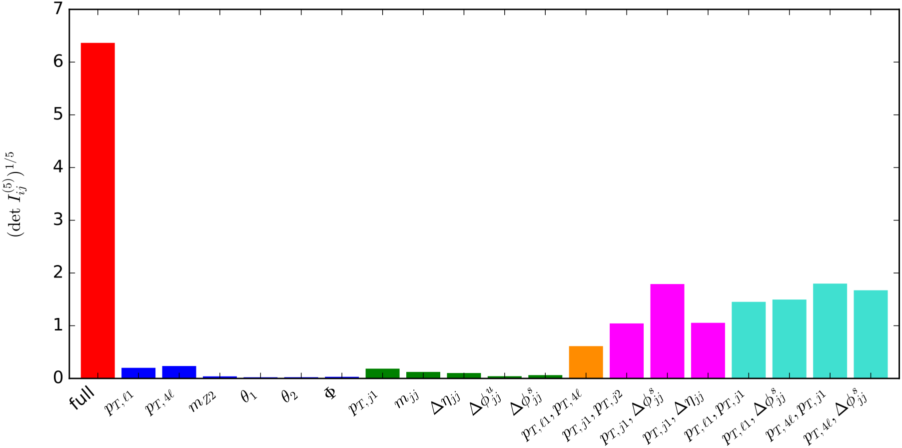


Information in distributions: 3d

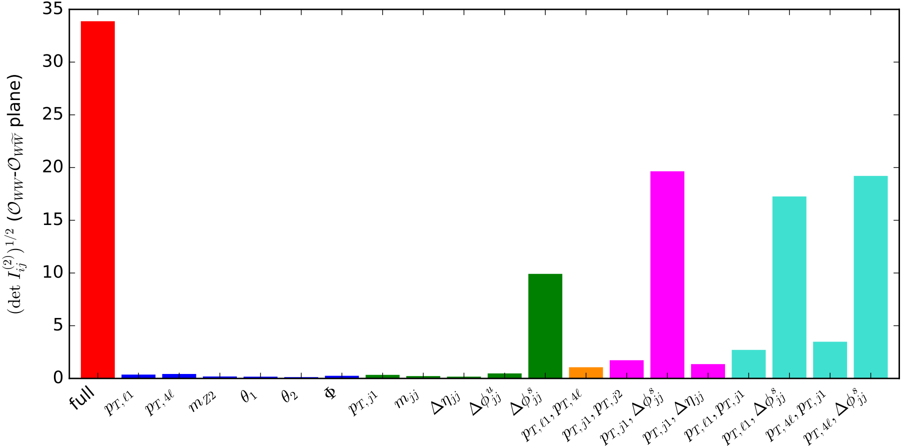


Determinant over $\mathcal{O}_W, \mathcal{O}_{WW}, \mathcal{O}_{\phi,2}$

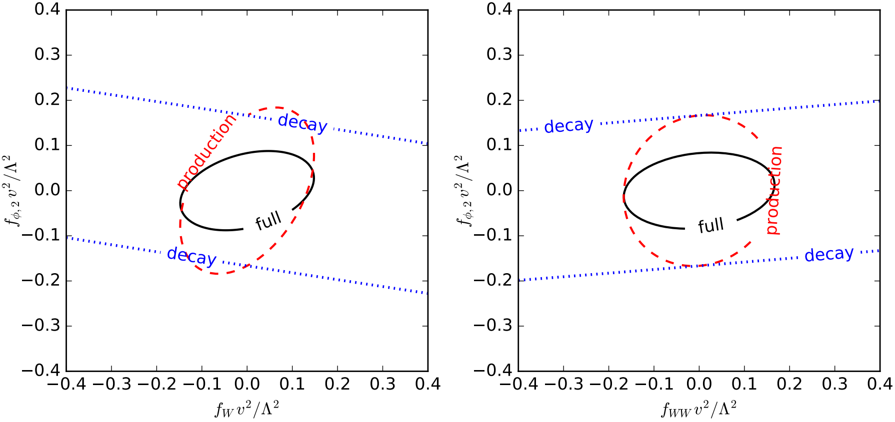
Information in distributions: 5d



Information in distributions: CP odd



Production vs decay



Operator effects limited to production / decay vertex

Comparison with likelihood ratio

Information distance

$$d(\boldsymbol{\theta}_1, \boldsymbol{\theta}_0) = \min_{\boldsymbol{\theta}(s)} \int_{s_0}^{s_1} \mathrm{d}s \sqrt{I_{ij} \frac{\mathrm{d}\theta^i(s)}{\mathrm{d}s} \frac{\mathrm{d}\theta^j(s)}{\mathrm{d}s}}$$

is strongly correlated with expected log likelihood ratio

$$q(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_0) = E\left[-2 \log \frac{f(\mathbf{x}|\boldsymbol{\theta}_1)}{f(\mathbf{x}|\boldsymbol{\theta}_0)} \middle| \boldsymbol{\theta}_0 \right]$$