Better Higgs Measurements through Information Geometry

work with Johann Brehmer, Kyle Cranmer, Tilman Plehn and Tim Tait

arXiv:1612.05261, 1712.02350

Felix Kling

UCIrvine

April 19th 2018, HEFT 2018

Introduction

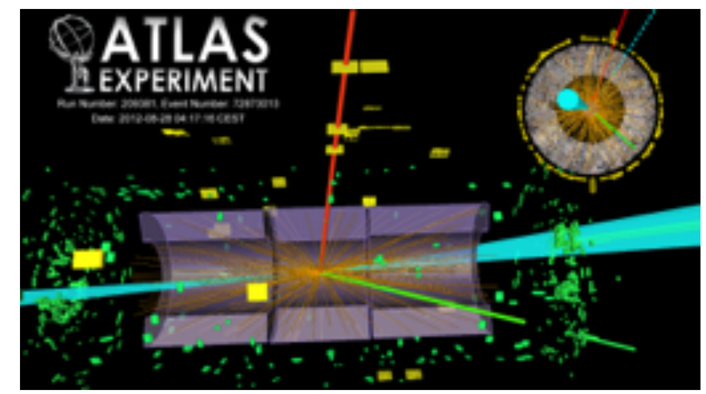
Motivation

- Higgs discovery: Standard Model complete
- there is probably* new physics in the Higgs sector:
  hierarchy problem, dark matter, CP-violation, …
- measurement of Higgs properties most exciting mission in the future

* no warranty expressed or implied

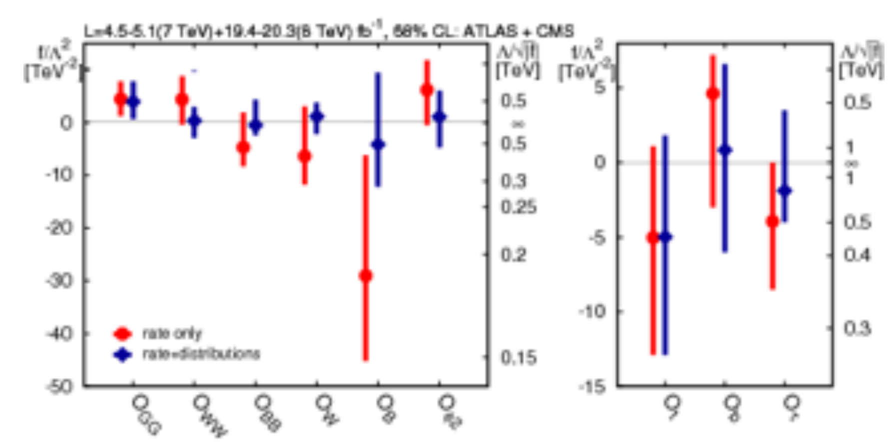
until the LHC find something really cool

Era of Data:



- large statistics at LHC, HL-LHC, HE-LHC
- complex data, contains lots of information
- modern multivariate analysis techniques [T. Martini, P. Uwer 1506.08798]
- correlations between measurements

Theory:

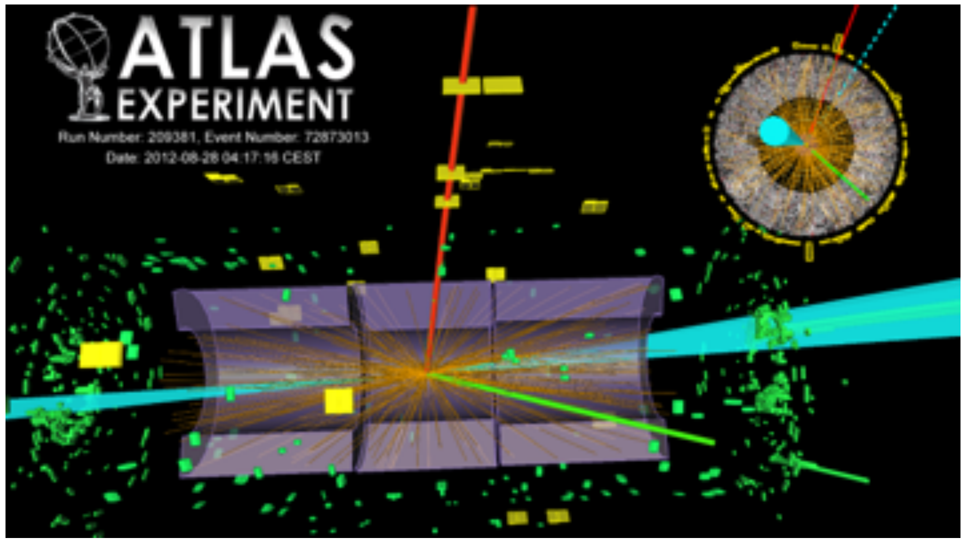


- theory description more and more complex
  coupling modifiers $\kappa \longrightarrow$ EFT
- predicts lots of features:
  rate, kinematic distribution, asymmetries

How to do Theory in an Era of Data?

Introduction

complex data: x



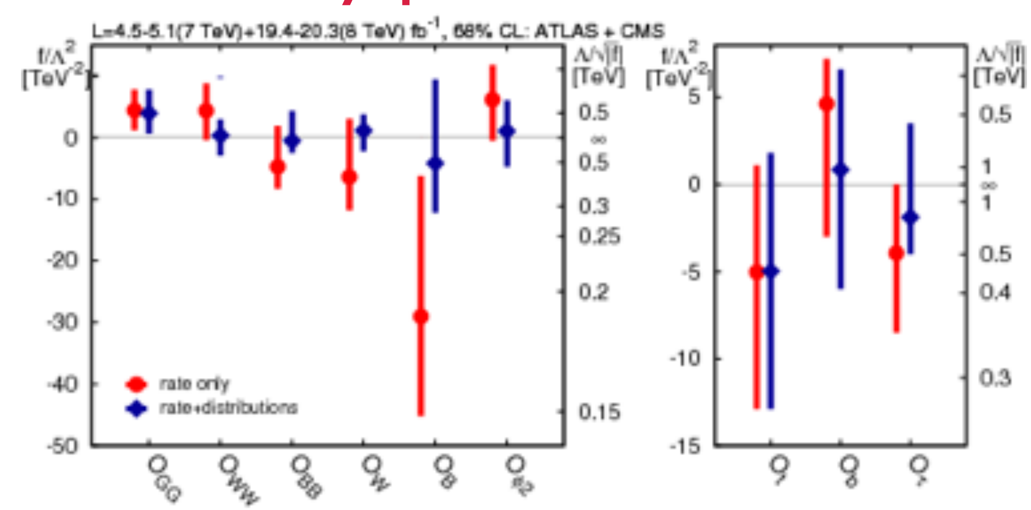
Conventional Analysis:

- rate or histogram based
- use standard kinematic observables
  → reproducible and **transparent**
- throw away lots of information
  → limited performance
- we already did that in the 80th …

Multivariate Methods:

- matrix-element-based, machine learning
- many recent developments
- use all phase-space information
  → **optimized sensitivity**
- black boxes
  → unsatisfying for theorists

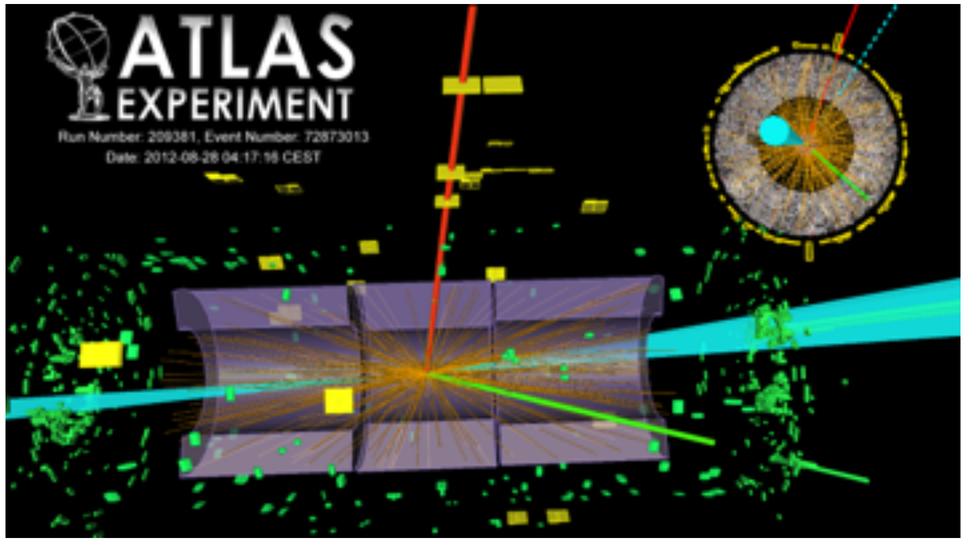
theory parameters: θ



[T. Corbett et al 1505.05516]

Introduction

complex data: x

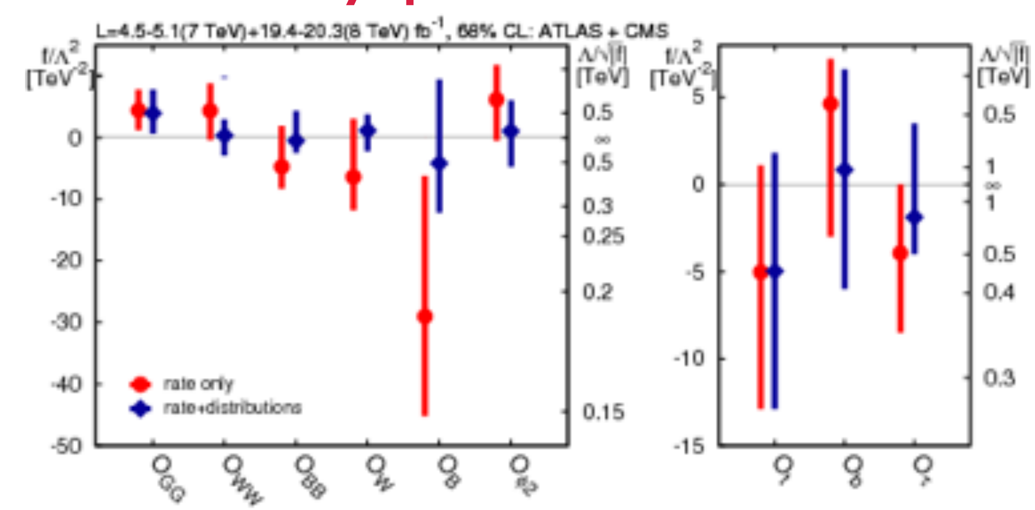


Conventional Analysis:
- rate or histogram based
- use standard kinematic observables
  → reproducible and **transparent**
- throw away lots of information
  → limited performance
- we already did that in the 80th …

Can we compute the **maximum sensitivity** of LHC data to theory in a **transparent** way?
→ Information Geometry

Multivariate Methods:
- matrix-element-based, machine learning
- many recent developments
- use all phase-space information
  → **optimized sensitivity**
- black boxes
  → unsatisfying for theorists

theory parameters: θ



[T. Corbett et al 1505.05516]

Introduction and Outline

Higgs CP - Which observables are sensitive to Higgs CP?
- What assumptions link those observables to CP?

Information Geometry - What is information?

Probing Higgs CP with Information Geometry - How well can we quantitatively test CP in the Higgs-gauge sector?

Total Information - What is the maximum precision to measure theory parameters?

Differential Information - Where in phase space is the information?

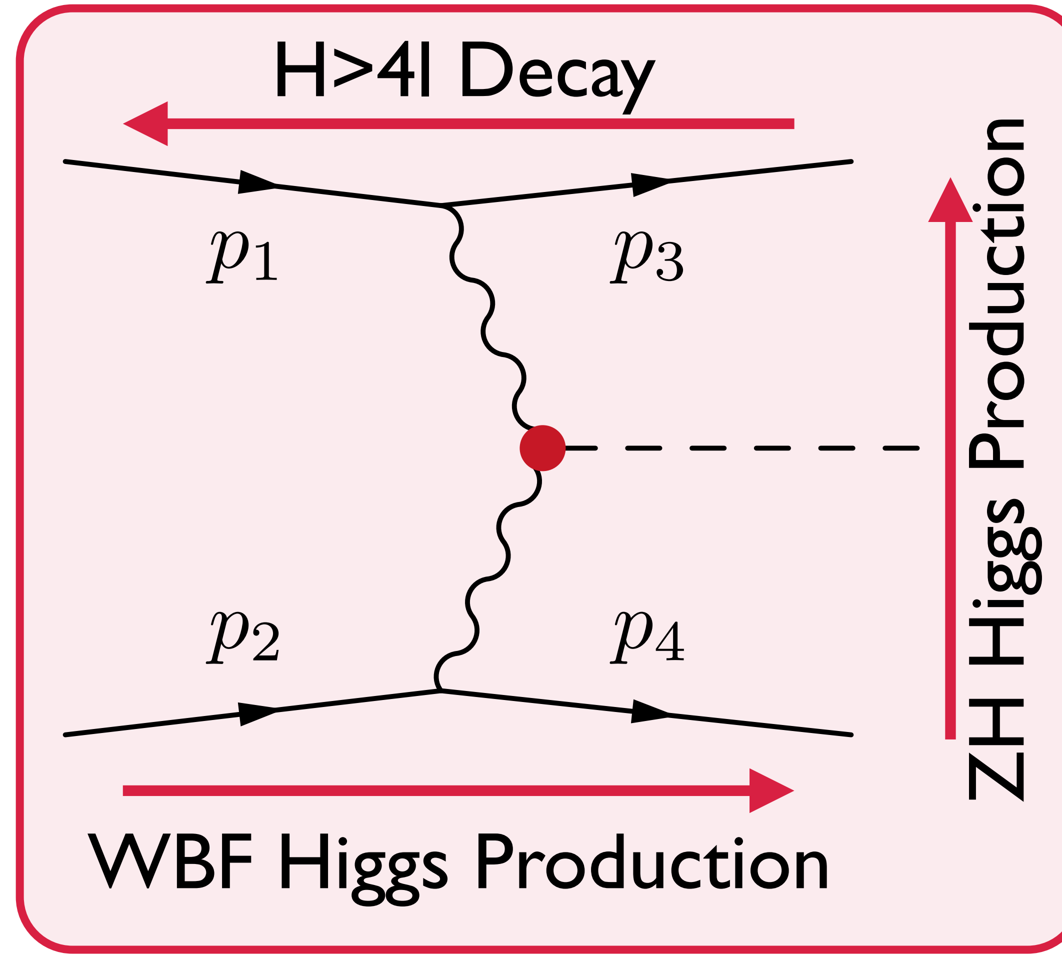
Information in Distributions - What are the most powerful observables?

Information in Analyses - How do histogram-based and multivariate analyses compare?

Summary and Outlook

Which observables are sensitive Higgs CP?

Higgs-Gauge Coupling
- WBF and ZH production, H>4l decay
- same hard process
- different final state (charge measurement)



Theory Language:
- dim-6-operators of SMEFT: $\mathcal{L} \supset \sum_i \frac{f_i}{\Lambda^2}\mathcal{O}_i$
- operators such as (be carefull, see Ben Gripaios)

  CP-even: $\mathcal{O}_{WW} \sim (\phi^\dagger \phi)\, W_{\mu\nu}\, W^{\mu\nu}$

  CP-odd: $\mathcal{O}_{W\widetilde{W}} \sim (\phi^\dagger \phi)\, W_{\mu\nu}\, \widetilde{W}^{\mu\nu}$
- goal: measure Wilson coefficients: f_i

Observables: 4 independent 4-momenta p_i

4 C-even and P-even scalar products

2 C-odd and P-even scalar products:

1 C-even and P-odd $\epsilon_{\alpha\beta\gamma\delta}\, p_1^\alpha p_2^\beta p_3^\gamma p_4^\delta$:

up to 3 CP sensitive observables

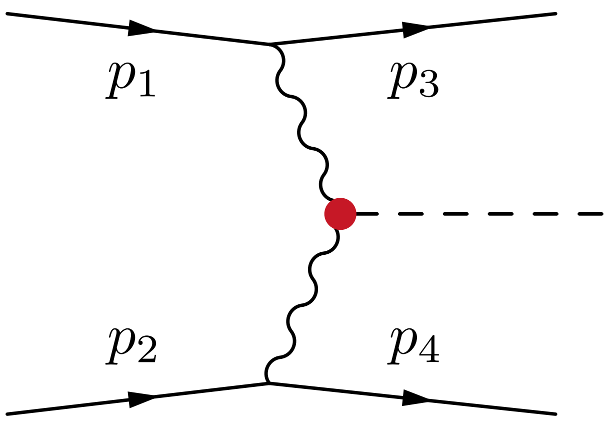
	WBF	ZH	H>4l
		$\Delta p_{T,ll}, \Delta E_{ll}$	$\theta_{1,2}$
	$\Delta\phi^s_{jj}$	$\Delta\phi^s_{ll}$	Φ

[WBF: Hankele, Klamke, Zeppenfeld hep-ph/0609075,
ZH: Christensen, Han, Li 1005.5393,
H>4l: Bolognesi et al. 1208.4018]

What assumptions link those observables to CP?

Why is WBF Higgs production sensitive to CP?

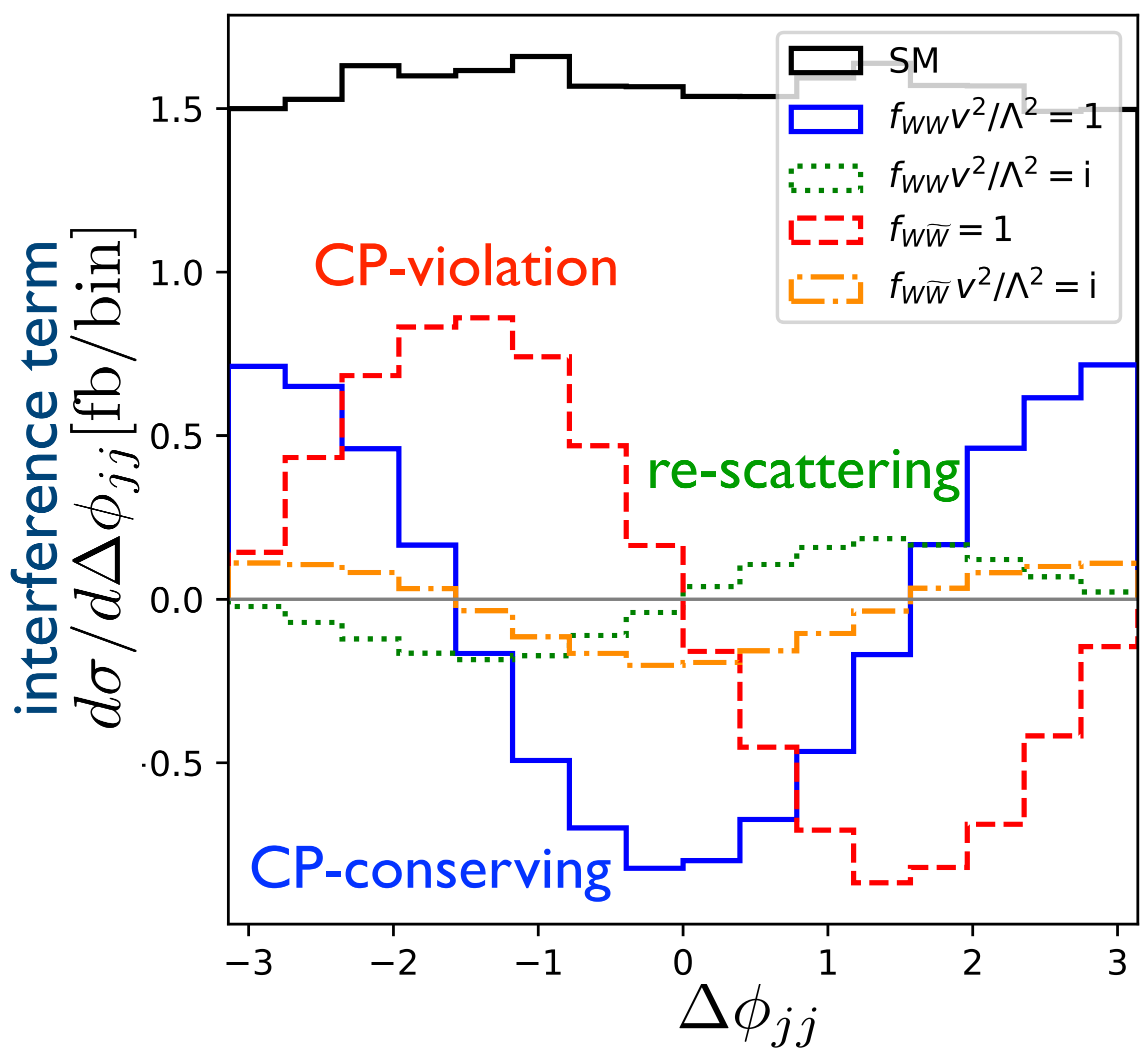
- naive time reversal $\hat{T}: \ |\vec{p},\vec{s}\rangle \to |-\vec{p},-\vec{s}\rangle$
- $\hat{T}$-symmetric initial state at pp-collider
- $\hat{T}$-invariant squared matrix element in absence of CP-violation and re-scattering



$$\langle f|\mathcal{T}|i\rangle \underset{\text{CPT-theorem}}{\overset{CP\text{-invariant}}{=}} \langle i_T|\mathcal{T}|f_T\rangle \underset{\text{optical theorem}}{\overset{\text{no re-scattering}}{=}} \langle f_T|\mathcal{T}|i_T\rangle^* \quad \Rightarrow \quad |\langle f|\mathcal{T}|i\rangle|^2 = |\langle f_T|\mathcal{T}|i_T\rangle|^2$$

- genuine $\hat{T}$-odd observable $\epsilon_{\alpha\beta\gamma\delta}\ p_1^\alpha p_2^\beta p_3^\gamma p_4^\delta$
  → signed angle $\Delta\phi_{jj}^s$

$\Delta\phi_{jj}^s$ is sensitive to CP-violation
if re-scattering effects are known to be small



Information Geometry

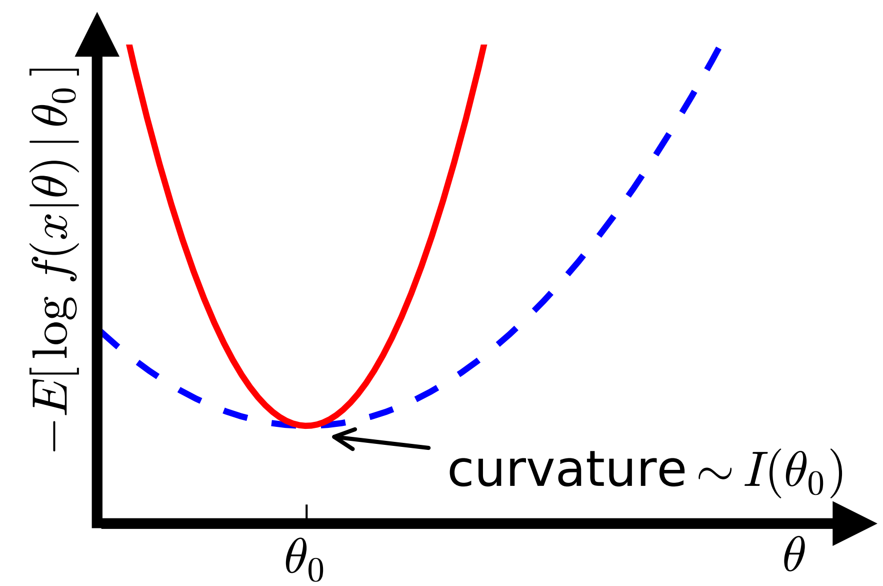
Measurement process
(we saw that before, see David Straub)

$$\theta_0 \xrightarrow{f(x|\theta_0)} x \xrightarrow{\hat{\theta}(x)} \hat{\theta}$$

true parameters — observables — estimator

Cramer-Rao Bound [C. R. Rao 1945; H. Cramér 1946]

$$\mathrm{cov}[\hat{\theta}|\theta_0] \leq I_{ij}^{-1}(\theta_0)$$



[Figure by J. Brehmer]

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

- encodes the maximum sensitivity of observables to model parameters

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

- **calculable** using Monte Carlo: $I_{ij}(\theta) = L \sum_{\text{events}} \frac{1}{\Delta\sigma(\theta)} \frac{\partial\Delta\sigma(\theta)}{\partial\theta_i} \frac{\partial\Delta\sigma(\theta)}{\partial\theta_j}$

[Plehn, Schichtel, Wiegand 1311.2591; Brehmer, Cranmer, FK, Plehn 1612.05261]

- **Additive** between experiments / phase-space regions
- **Independent** of parametrization of x
- **Covariant** under θ→ θ\`

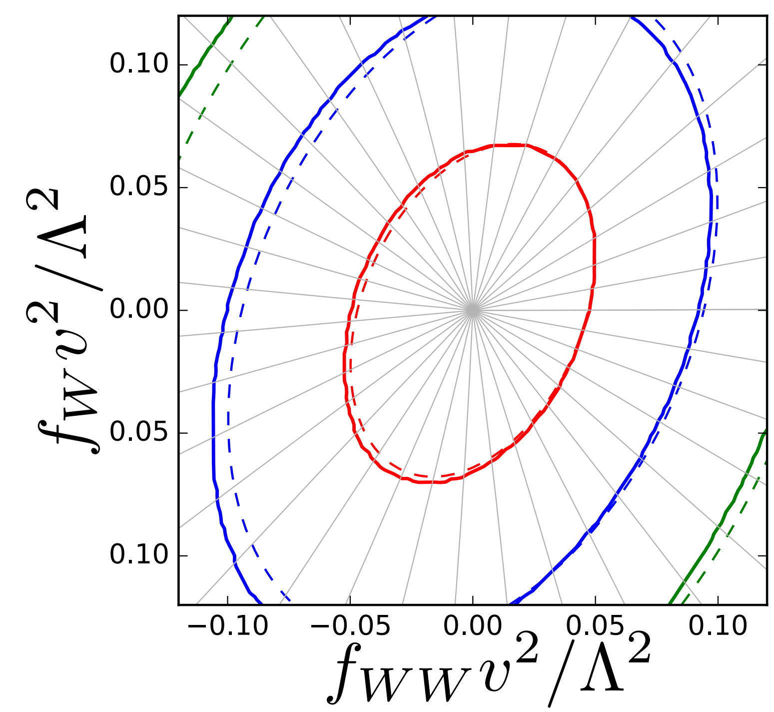
Total Information

What is the maximum precision to measure theory parameters?

- encoded in Fisher Information $I = \sum_{all\ events} I_{event}$

Example: WBF Higgs Production with $H \to \tau\tau$

$$
\begin{array}{cc}
 & \begin{array}{cccc} f_W & f_{WW} & f_{W\tilde{W}} & \mathrm{Im} f_{WW} \end{array} \\
I_{ij}(\mathbf{0}) = & \left(\begin{array}{cccc} 715 & -191 & 1 & 0 \\ -191 & 321 & -1 & 0 \\ 1 & -1 & 359 & -81 \\ 0 & 1 & -81 & 23 \end{array}\right) \begin{array}{c} f_W \\ f_{WW} \\ f_{W\tilde{W}} \\ \mathrm{Im} f_{WW} \end{array}
\end{array}
$$



- sensitivity to CP-violating operator
- large mixing between CP-conserving operators
- no mixing between CP-conserving and CP-violating operators
- re-scattering can mimic CP-violation
- Minimal Errors: $\Delta\theta > 1\sqrt{I}$

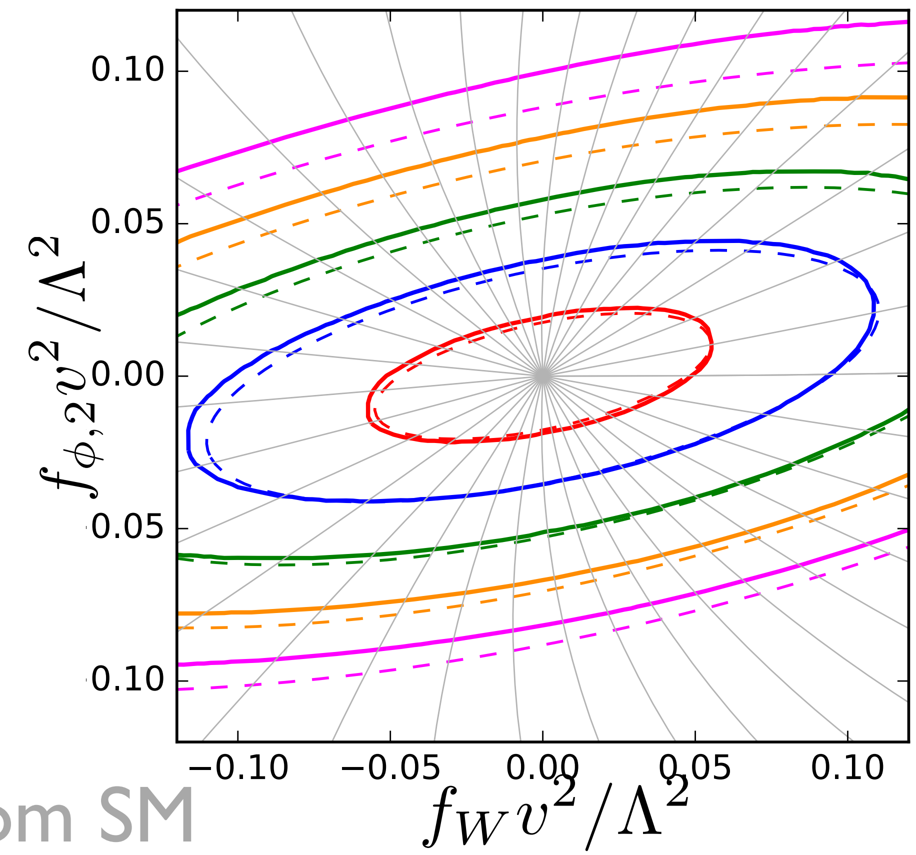
→ calculate the maximum sensitivity of any LHC process

we assume13TeV LHC, L=100 fb^{-1}, take into account ggF and Z+jets BG,
for more analysis details see 1612.05261, 1712.02350

Total Information

Geometric Interpretation of Fisher Information

- Distance Measure ~ unlikeliness to measure θ if θ_0 is true 'in sigmas'
- local distance: $d^2 = I_{ij}(\theta_0)(\theta^i - \theta_0^i)(\theta^j - \theta_0^j)$
  (dashed)
- global distance: $d = \min_{\theta(s)} \int_{s_a}^{s_b} ds \sqrt{I_{ij} \frac{d\theta_i}{ds} \frac{d\theta_j}{ds}}$
  (solid)



Contours of distance d=1,2,3,4,5 from SM

- $I_{ij}(\mathbf{0})$ only sensitive to linear effects: $\Delta\sigma \sim \theta_i \Delta\sigma_i$
- Information geometry for dim-6 operators $\theta_i = f_i^{d=6} v^2/\Lambda^2$

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

Difference between local/global distance ⟷ size of $\mathcal{O}(\Lambda^{-4})$ effects

Differential Information

Where in phase space is the information?

- binned kinematic distribution of information

$$I_{bin} = \sum_{events \in bin} I_{event}$$

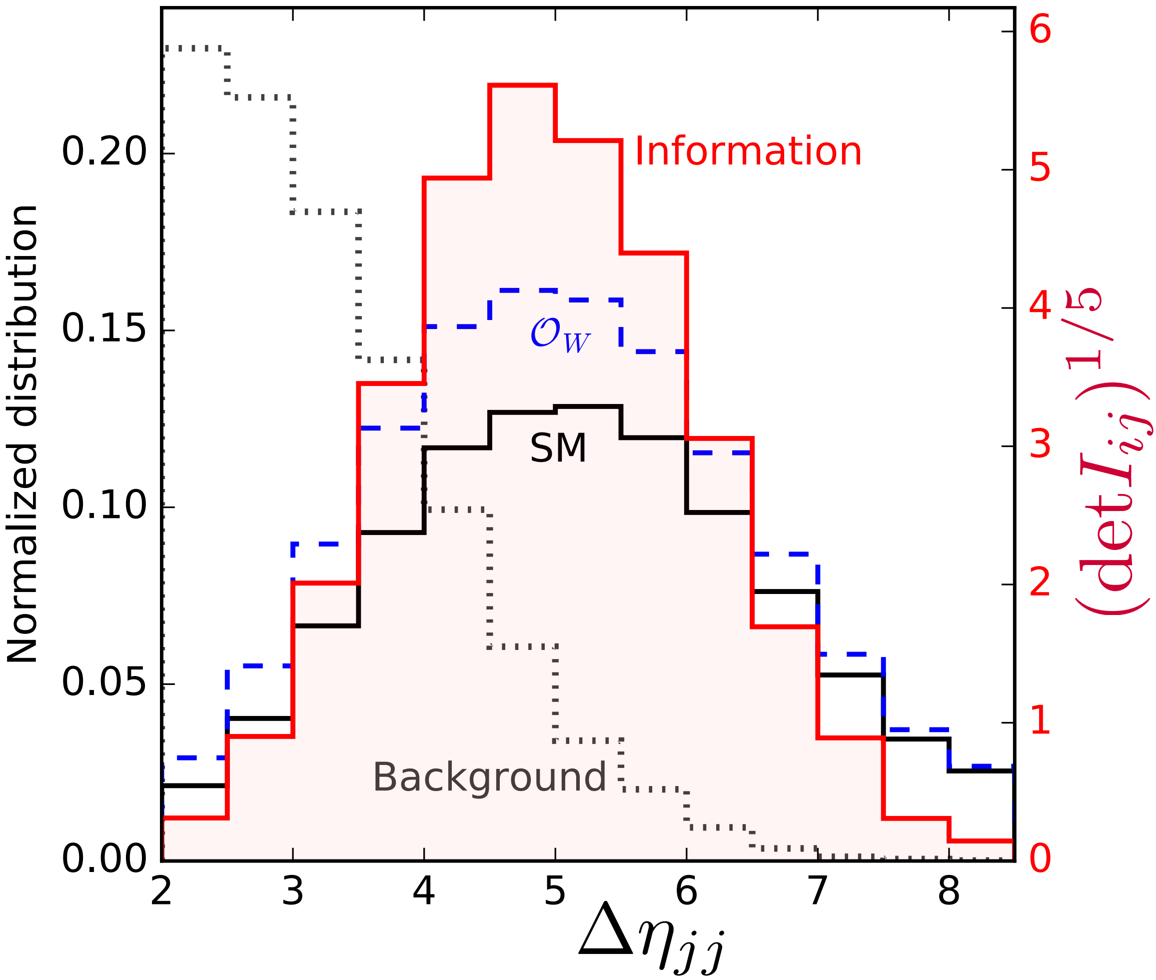
Example: Jet Rapidity Difference in WBF

- smaller background at large $\Delta\eta_{jj}$
- momentum dependent operator

  → largest effect at medium $\Delta\eta_{jj}$
- strong WBF cuts ($\Delta\eta_{jj}$> 4.2):

  → lose information of dim-6 operators

→ identify relevant phase-space regions



Information in Distributions

What are the most powerful observables?

- information of binned kinematic distribution

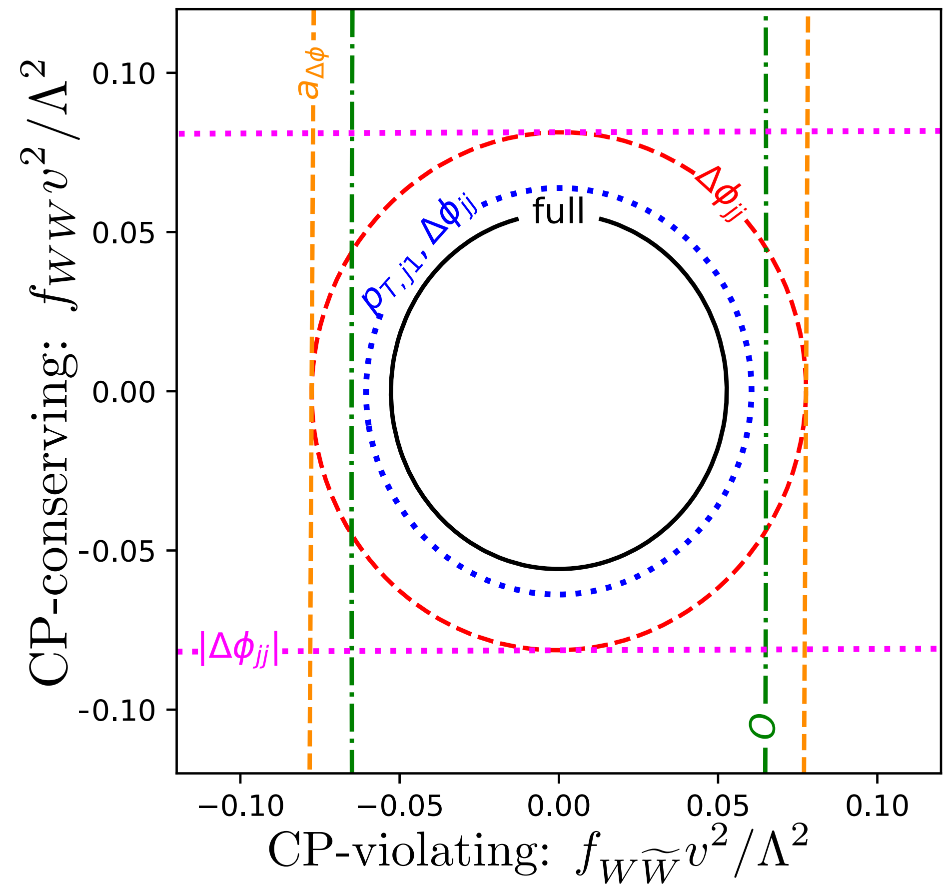
$$I = \sum_{bins} I_{bin}$$

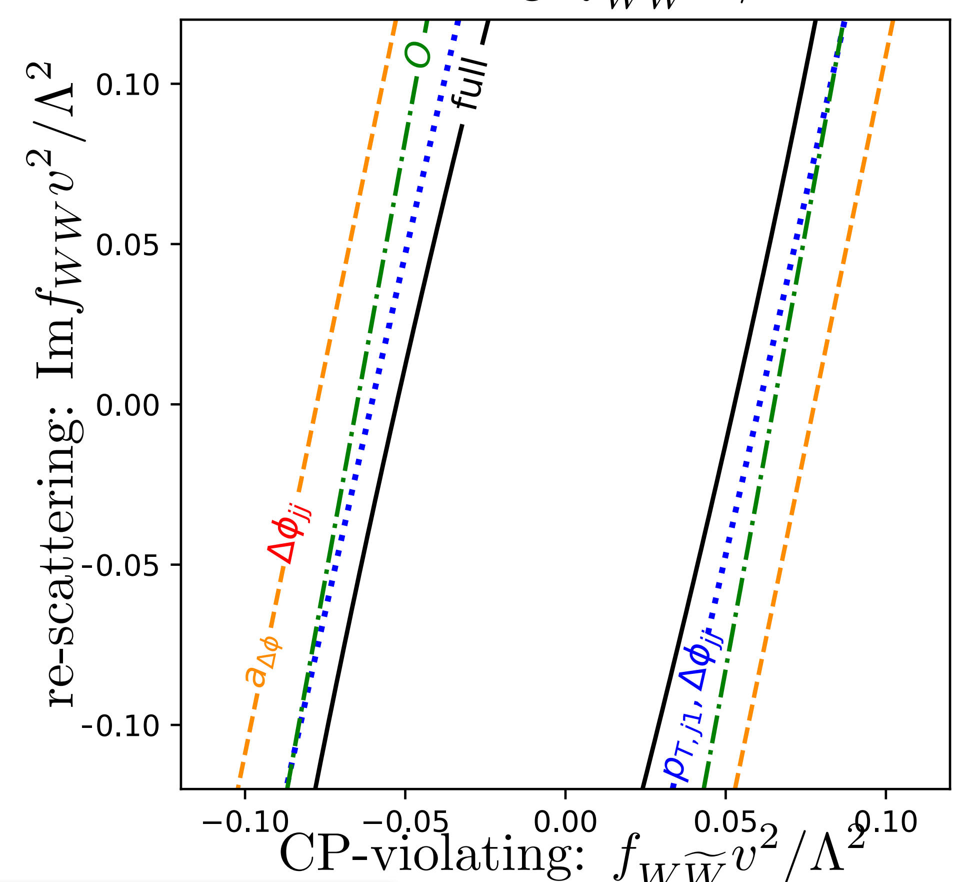
- minimum measurement error $\Delta f \geq 1/\sqrt{I}$

Example: Higgs coupling measurement in WBF

- $|\Delta\phi_{jj}|$ sensitive to CP-conserving physics only
- asymmetry sensitive to CP-violating physics only
- signed $\Delta\phi_{jj}$ probes both
- 2D histogram better, but still not close to **full** information
- re-scattering effects can mimic CP-violation
- asymmetry in $\Delta\phi_{jj}$ implies CP violation (in the absence of re-scattering)
- re-scattering small in SM

→ identify most powerful observables

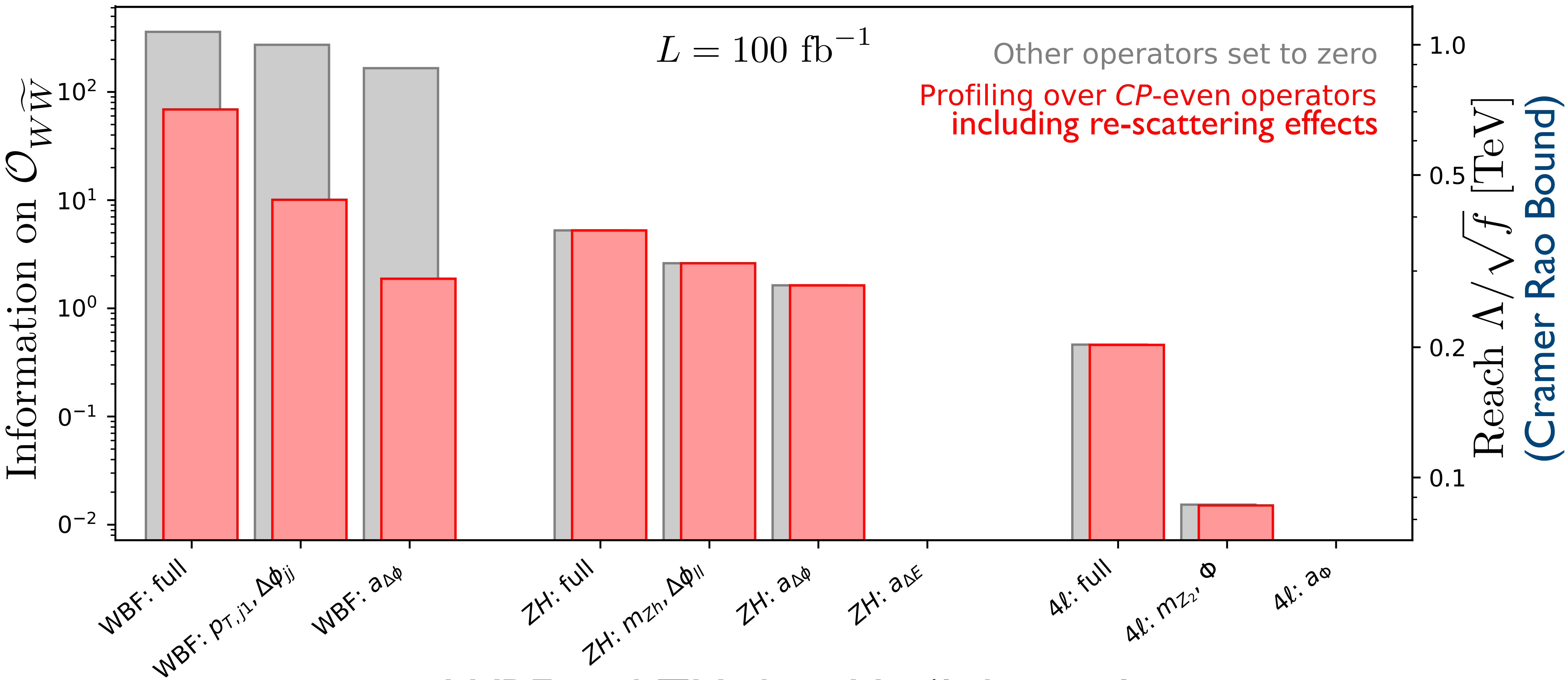




Information in Analyses

How do histogram-based and multivariate analyses compare?

Example: Information on CP-violating Higgs couplings



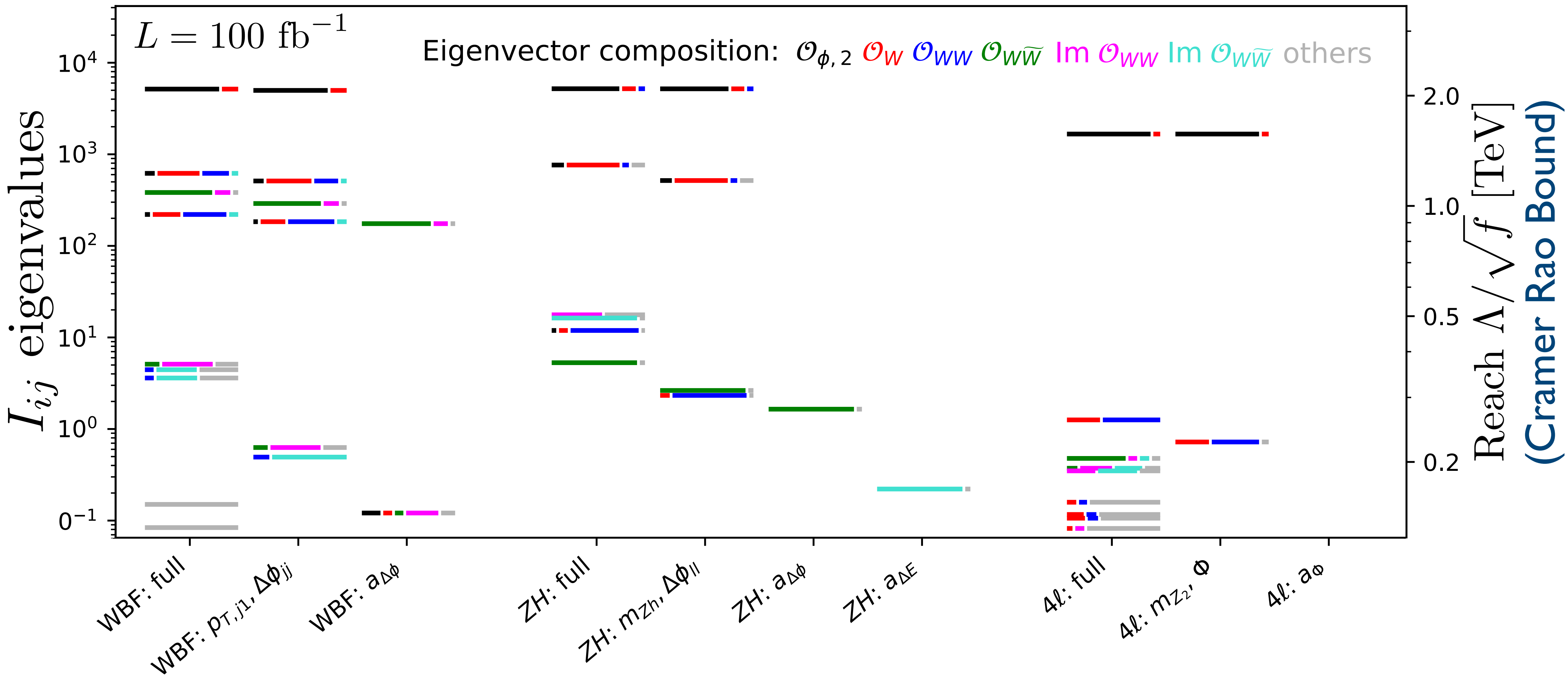
- more sensitivity in WBF and ZH than H>4l due to larger momentum transfer
- WBF requires additional theory assumption on re-scattering
- CP-information mostly captured in asymmetry of $\epsilon_{\alpha\beta\gamma\delta}\, p_1^\alpha p_2^\beta p_3^\gamma p_4^\delta \sim \Delta\phi$
- adding momentum transfer measures/multivariate analysis increase sensitivity

⟶ quantitatively compare histogram-based vs. multivariate analyses

Information in Analyses

How do histogram-based and multivariate analyses compare?

Example: Information and correlation of all Higgs couplings



- $a_{\Delta\varphi}$ not sensitive to real CP-even Wilson coefficients
- WBF requires additional theory assumption on re-scattering

Conclusion

Theory in an Era of Data

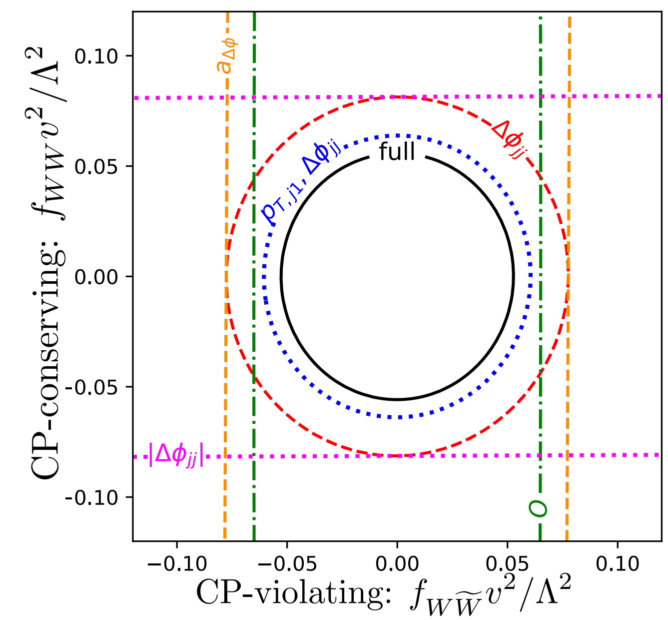
- lots of data, powerful multivariate tools
- constrain high-dimension theory space

$$
\begin{array}{c c l}
 & \begin{array}{cccc} f_W & f_{WW} & f_{W\tilde{W}} & \mathrm{Im} f_{WW} \end{array} & \\
I_{ij}(\mathbf{0}) = & \begin{pmatrix} 715 & -191 & 1 & 0 \\ -191 & 321 & -1 & 0 \\ 1 & -1 & 359 & -81 \\ 0 & 1 & -81 & 23 \end{pmatrix} & \begin{array}{l} f_W \\ f_{WW} \\ f_{W\tilde{W}} \\ \mathrm{Im} f_{WW} \end{array}
\end{array}
$$

Information Geometry

- fisher information encodes the maximum sensitivity of observables to model parameters

- calculate maximum sensitivity
- identify important phase space regions
- identify most powerful observables
- quantitatively compare analyses



- powerful and transparent analysis tool
- particularly easy to apply to EFT

Outlook:

- include systematics,
- detector effects, missing information