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 5th 2018, HE/HL LHC Meeting

EFT at HL/HE-LHC

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

arXiv:1612.05261, 1712.02350

Felix Kling

UCIrvine

April 5th 2018, HE/HL LHC Meeting

Introduction

Motivation

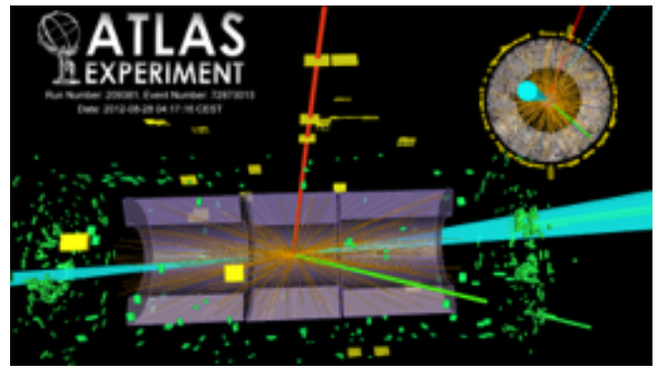
- Higgs discovery: Standard Model complete
- there is probably* new physics in the Higgs sector:
  hierarchy problem, dark matter, CP-violation, …
- measurement of Higgs properties most exciting mission in the future

* no warranty expressed or implied

until the LHC find something really cool

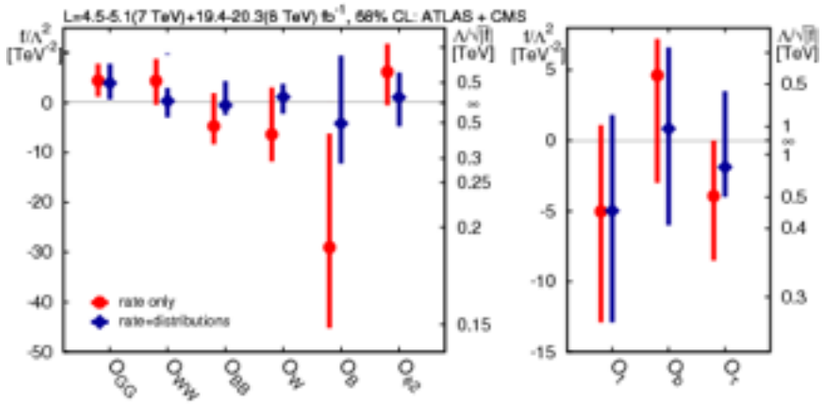
Era of Data:

- large statistics at LHC, HL-LHC, HE-LHC
- complex data, contains lots of information
- modern multivariate analysis techniques
- correlations between measurements



Theory:

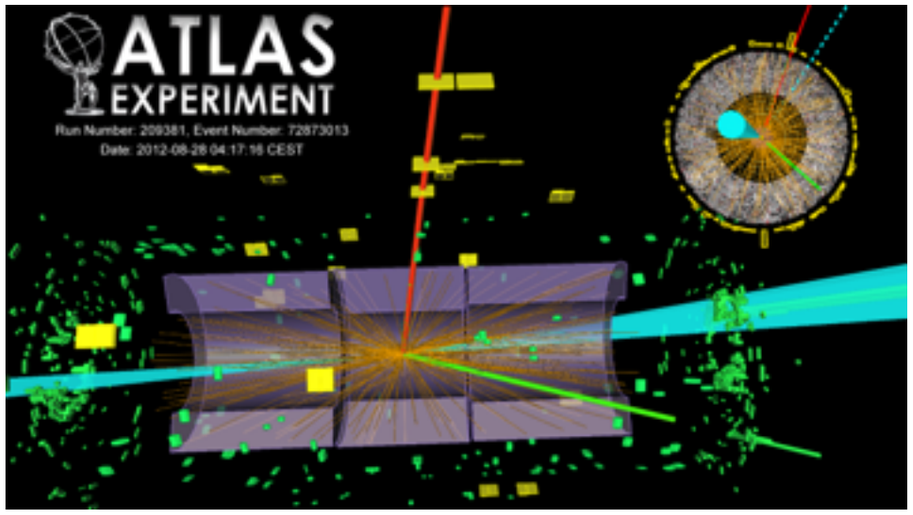
- theory description more and more complex
  coupling modifiers κ ⟶ EFT
- predicts lots of features:
  rate, kinematic distribution, asymmetries



How to do Theory in an Era of Data?

Introduction

complex data: x



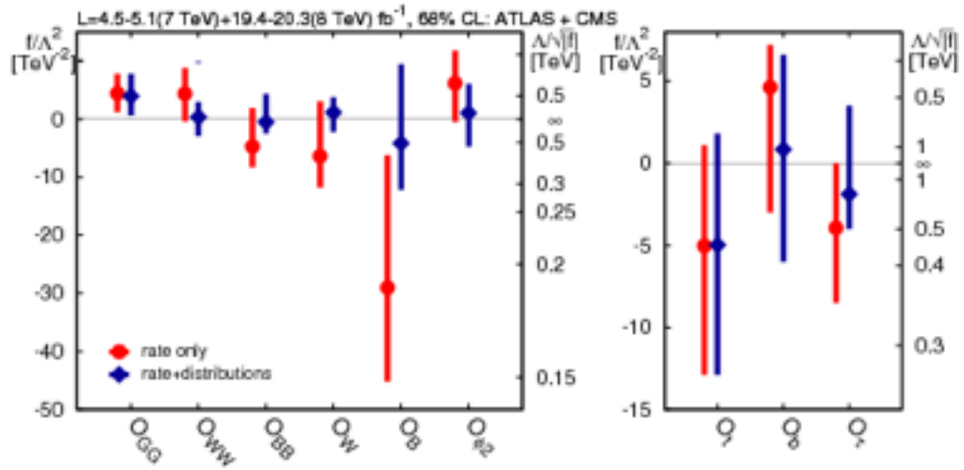
Conventional Analysis:

- 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

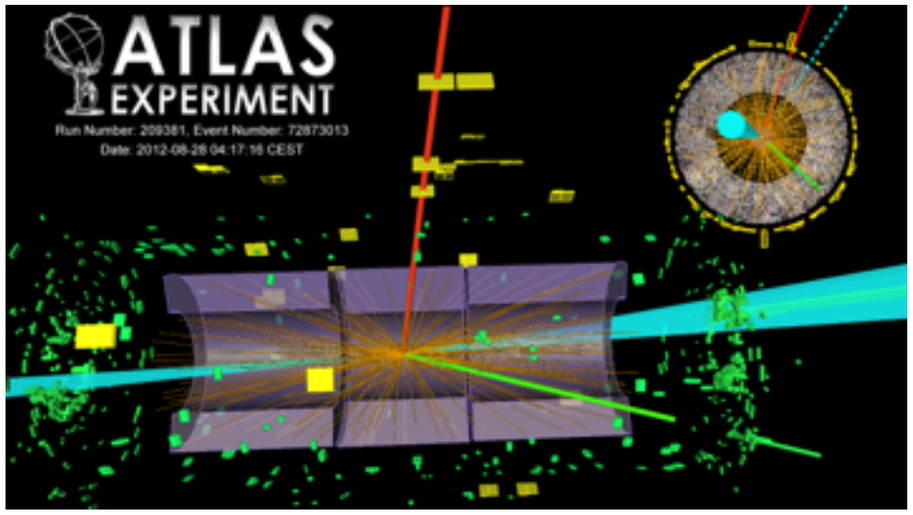
theory parameters: θ



[T. Corbett et al 1505.05516]

Introduction

complex data: x



Conventional Analysis:

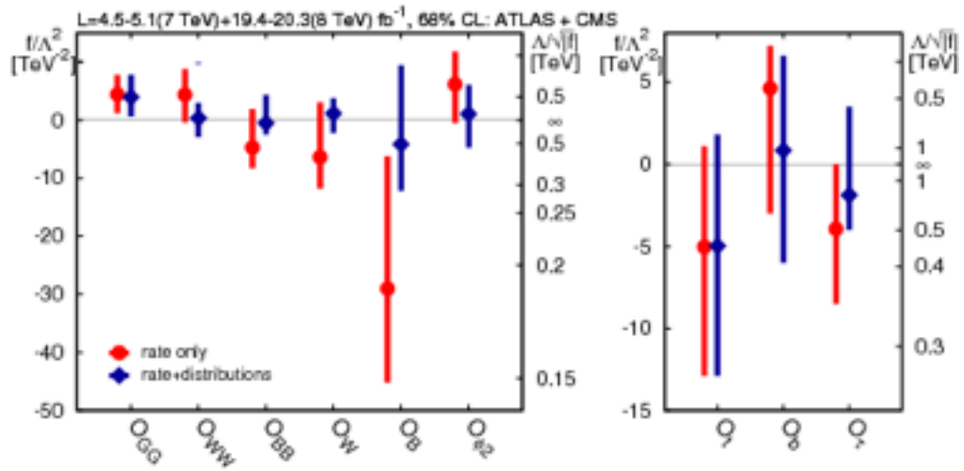
- rate or histogram based
- use standard kinematic observables
  → reproducible and **transparent**
- throw away lots of information
  → limited performance
- we already did that in the 80th …

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

Information Geometry - What is information?

Physics Example - 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

Information Geometry

Measurement process

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

θ_0: true parameters — x: observables — $\hat{\theta}$: estimator

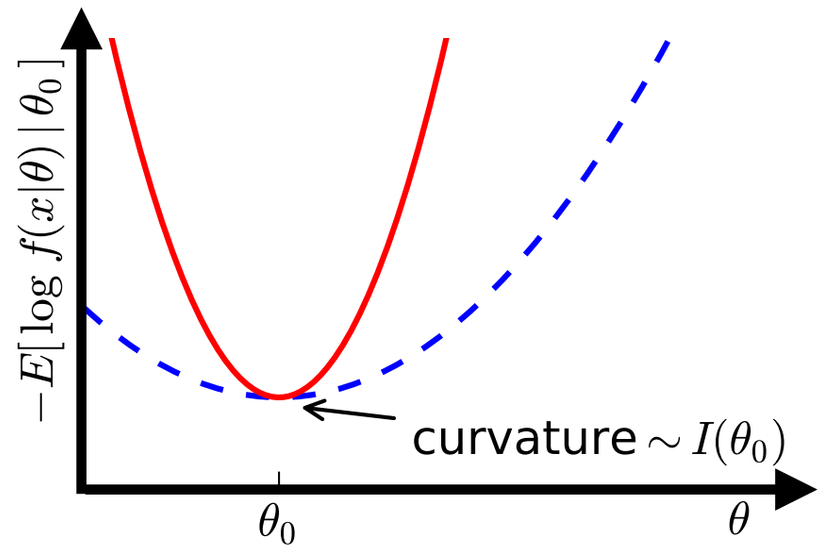
Cramer-Rao Bound [C. R. Rao 1945; H. Cramér 1946]

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

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

- encodes the maximum sensitivity of observables to model parameters

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



[Figure by J. Brehmer]

- 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`

CP in the Higgs-gauge sector

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

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

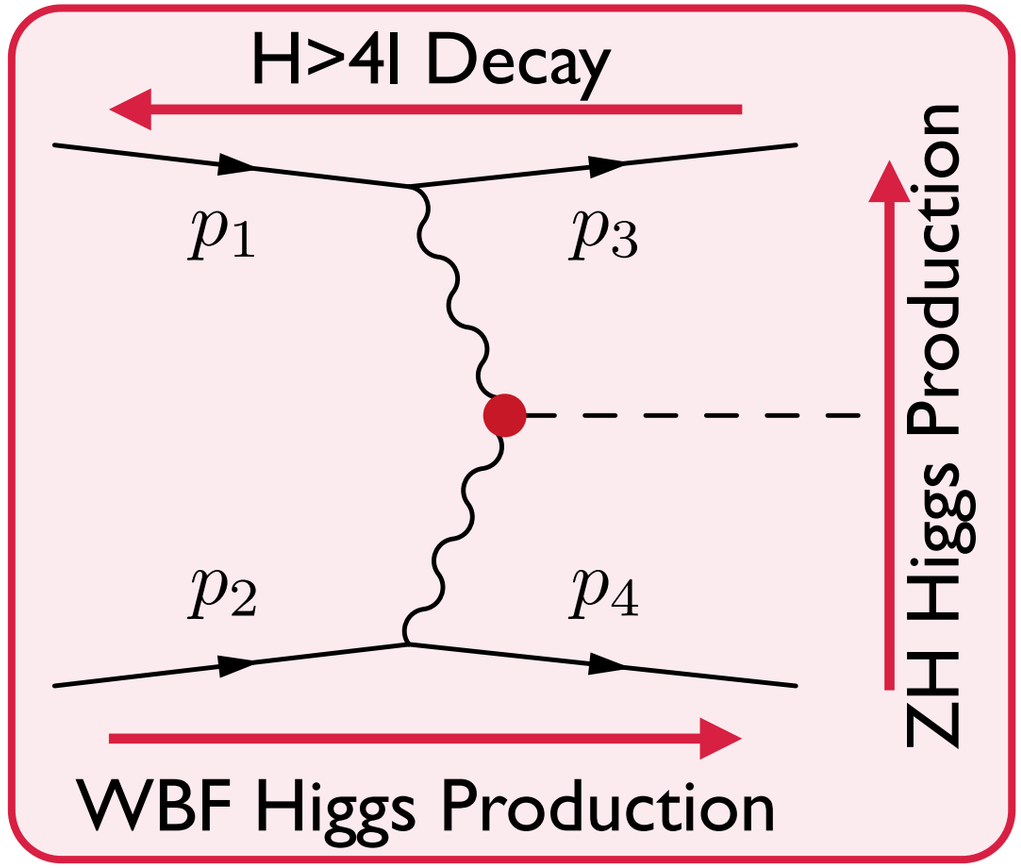
  CP-odd: $\mathcal{O}_{W\widetilde{W}} \sim (\phi^\dagger \phi)\, W_{\mu\nu}\, \widetilde{W}^{\mu\nu}$
- goal: measure Wilson coefficients: f_i

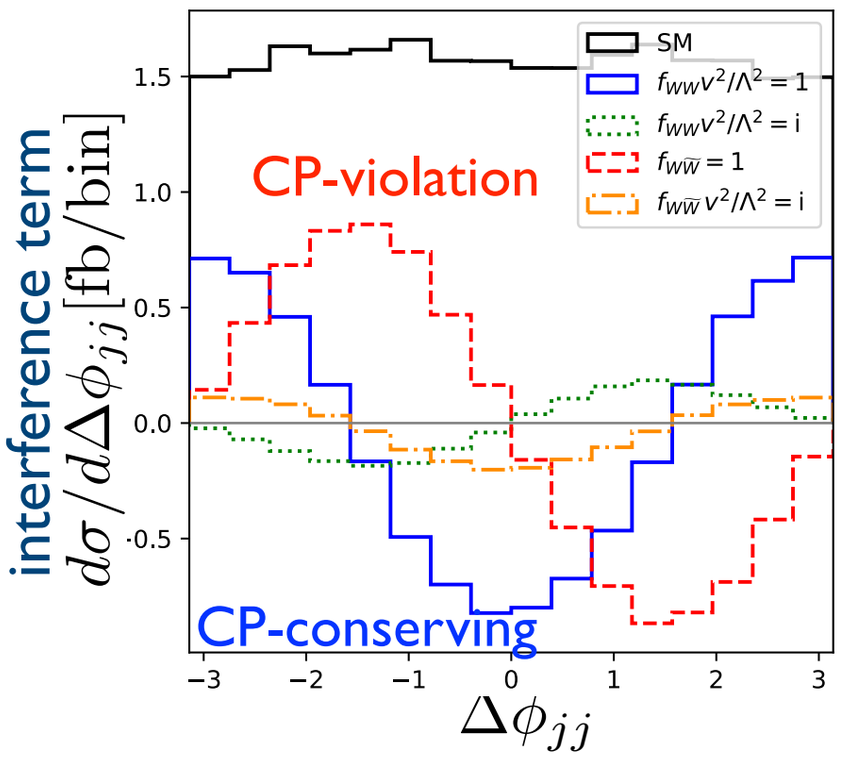
Observables: 4 independent 4-momenta p_i
- CP-sensitiv observable: $\epsilon_{\alpha\beta\gamma\delta}\, p_1^\alpha p_2^\beta p_3^\gamma p_4^\delta$

  WBF: $\Delta\phi_{jj}^s$ Hankele, Klamke, Zeppenfeld 0609075

  ZH: $\Delta\phi_{ll}^s$ Christensen, Han, Li 1005.5393,

  H>4l: Φ Bolognesi et al. 1208.4018





Total Information

What is the maximum precision to measure theory parameters?

- encoded in Fisher Information

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

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

- sensitivity to CP-violating operator
- large mixing between CP-conserving operators
- no mixing between CP-conserving and CP-violating operators

→ 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

Differential Information

Where in phase space is the information?

- binned kinematic distribution of information

Example: Jet Rapidity Difference in WBF

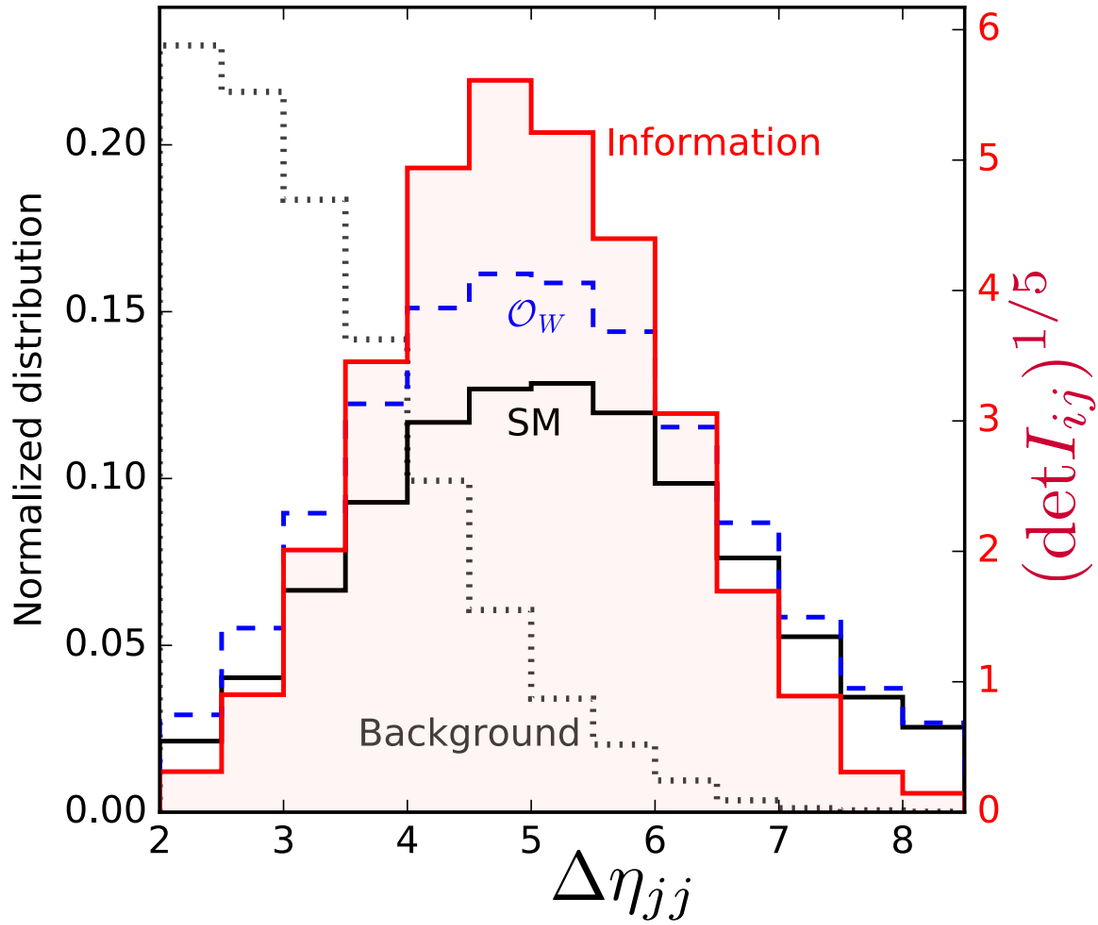
- smaller background at large $\Delta\eta_{jj}$
- momentum dependent operator

→ 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

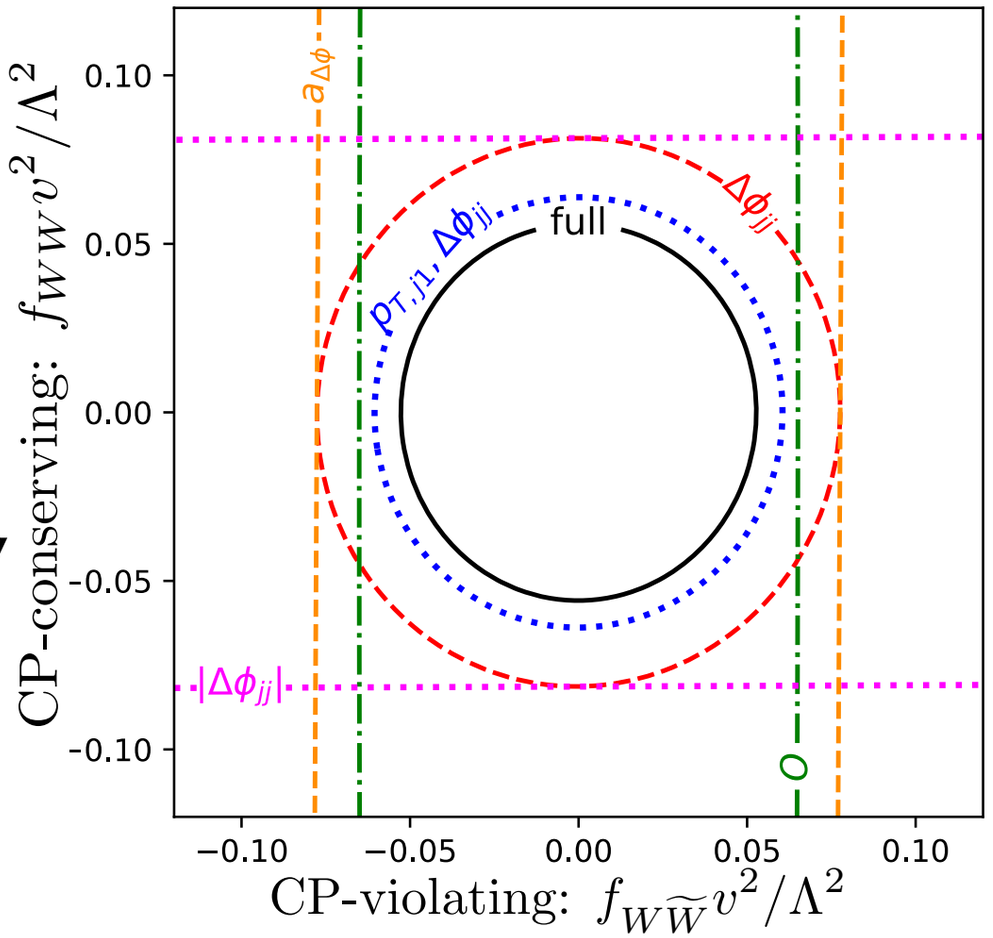
What are the most powerful observables?

- information of binned kinematic distribution
- 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

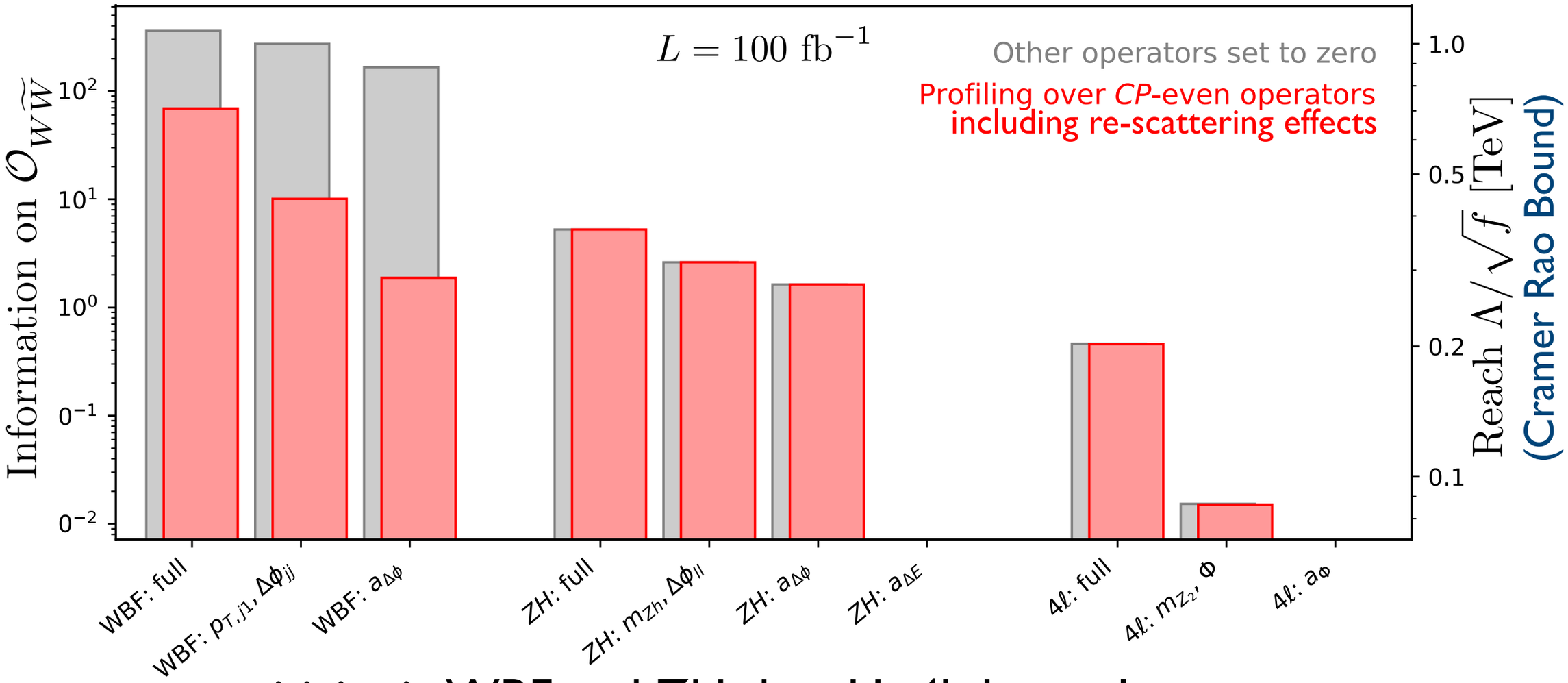
→ identify most powerful observables



Information in Analyses

How do histogram-based and multivariate analyses compare?

Example: Information on CP-violating Higgs couplings

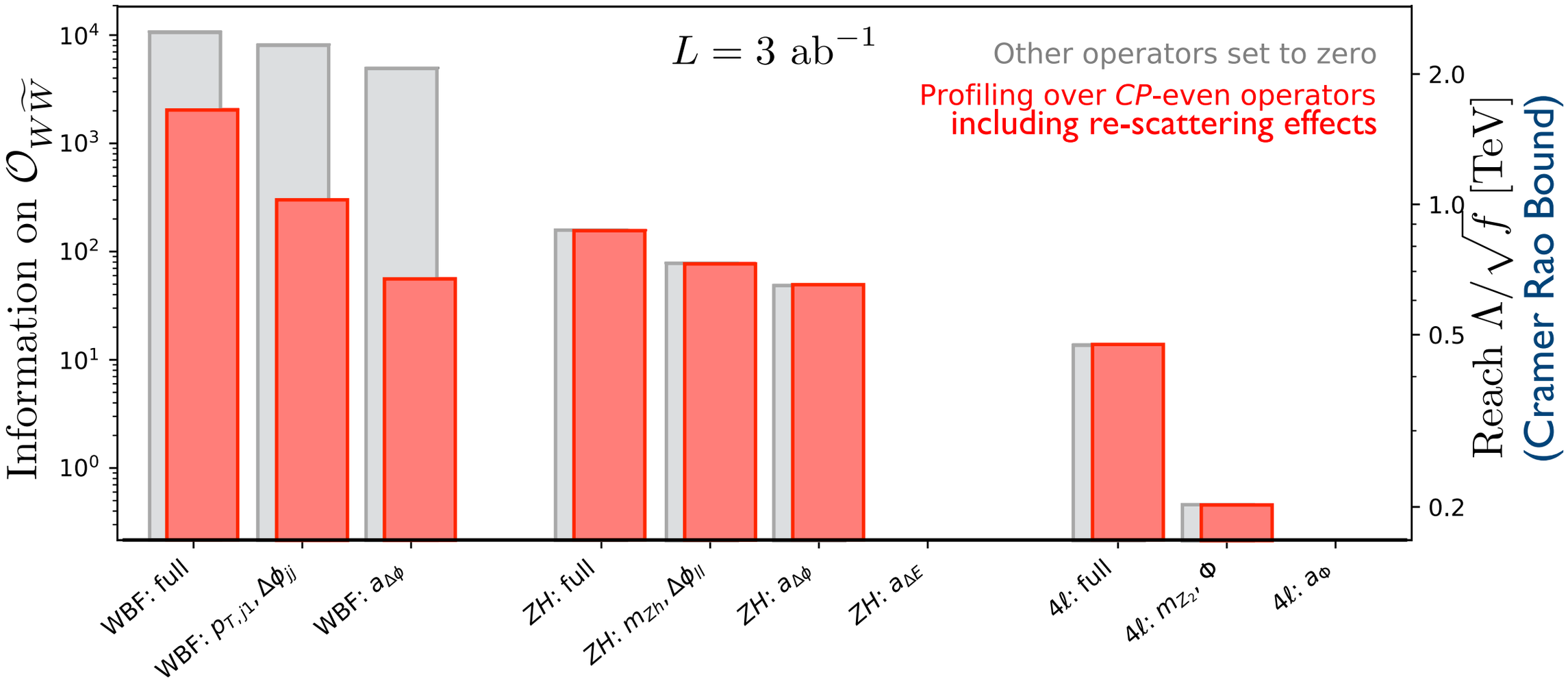


- more sensitivity in WBF and ZH than H>4l due to larger momentum transfer
- WBF requires additional theory assumption on re-scattering
- CP-information mostly captured in asymmetry of $\epsilon_{\alpha\beta\gamma\delta}\ p_1^\alpha p_2^\beta p_3^\gamma p_4^\delta \sim \Delta\phi$
- adding momentum transfer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

→ quantitatively compare histogram-based vs. multivariate analyses

Information in Analyses

- much better precision possible at HL-LHC



Conclusion

Theory in an Era of Data

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

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

Information Geometry

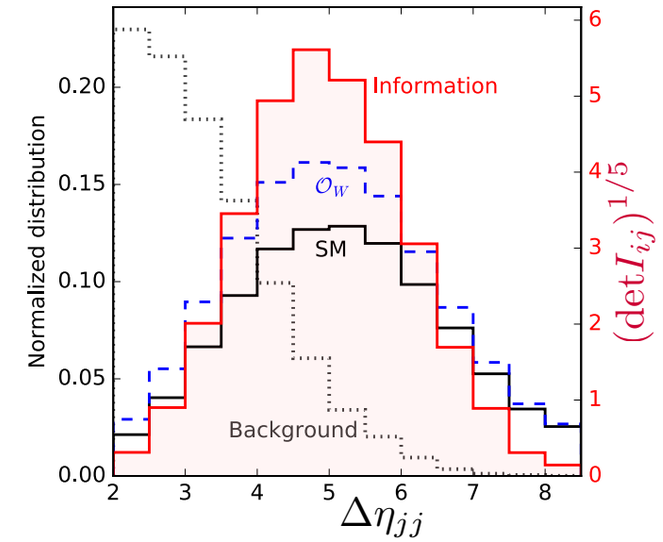
- fisher information encodes the maximum sensitivity of observables to model parameters

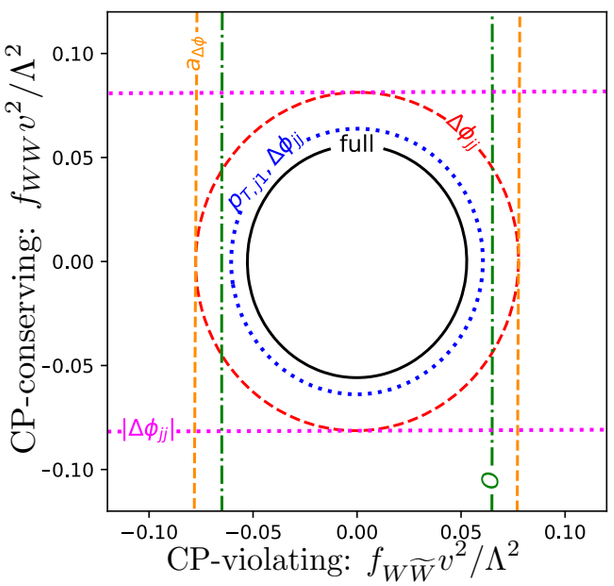
- calculate maximum sensitivity
- identify important phase space regions
- identify most powerful observables
- quantitatively compare analyses

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

Outlook:

- include systematics,
- detector effects, missing information

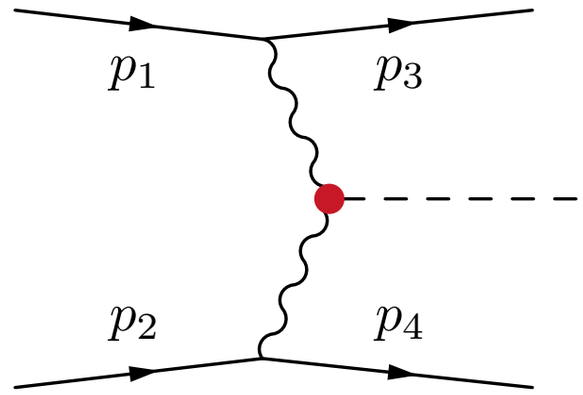




What assumptions link those observables to CP?

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



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

- 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