Observables for EFT Measurements

Andrei Gritsan
Johns Hopkins University

April 25, 2024

LPC EFT Workshop at Notre Dame

University of Notre Dame, Indiana

Observables in the Context of EFT fits

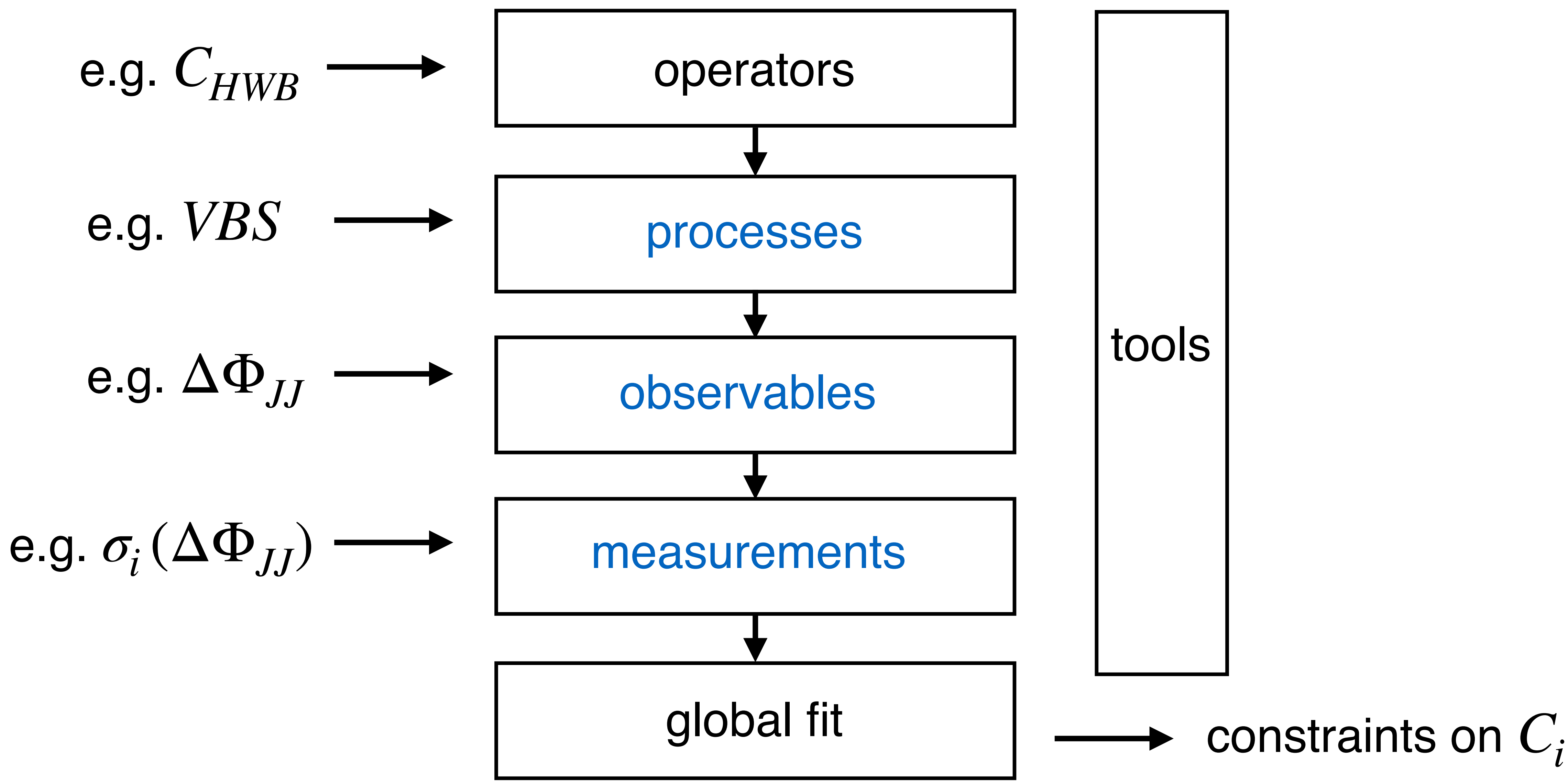
- LHC EFT WG effort (Area 3): [arXiv:2211.08353](https://arxiv.org/abs/2211.08353)

CERN-LHCEFTWG-2022-001
CERN-LPCC-2022-05

November 15, 2022

LHC EFT WG Report: Experimental Measurements and Observables

Nuno Castro[1], Kyle Cranmer[2], Andrei V. Gritsan[3], James Howarth[4], Giacomo Magni[5,6], Ken Mimasu[7], Juan Rojo[5,6], Jeffrey Roskes[3], Eleni Vryonidou[8], Tevong You[9,10,11]



Observables

measurements ← observables ← processes

$$\mathscr{P}(\vec{x}_{\rm reco}\,|\,\vec{\theta}) = \int d\vec{x}_{\rm part} \quad p(\vec{x}_{\rm reco}\,|\,\vec{x}_{\rm part}) \quad \mathscr{P}(\vec{x}_{\rm part}\,|\,\vec{\theta})$$

reco observables

$\vec{x}_{\rm reco}$

measurement

parton shower

detector effects

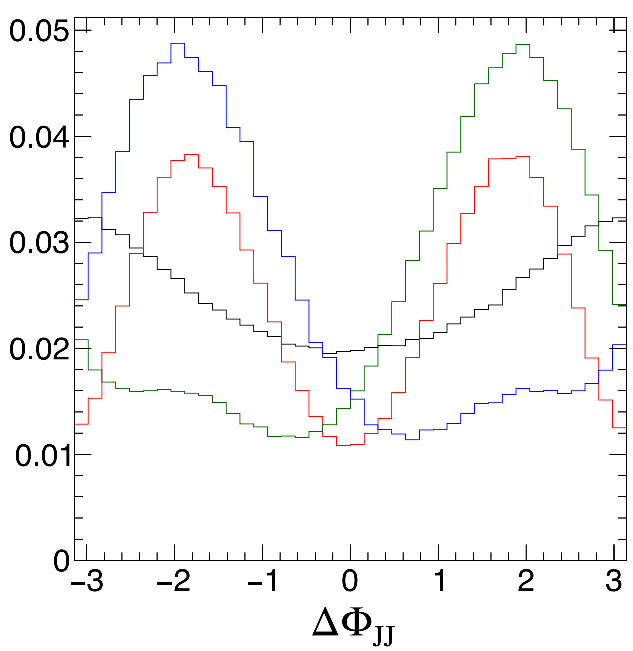
reconstruction

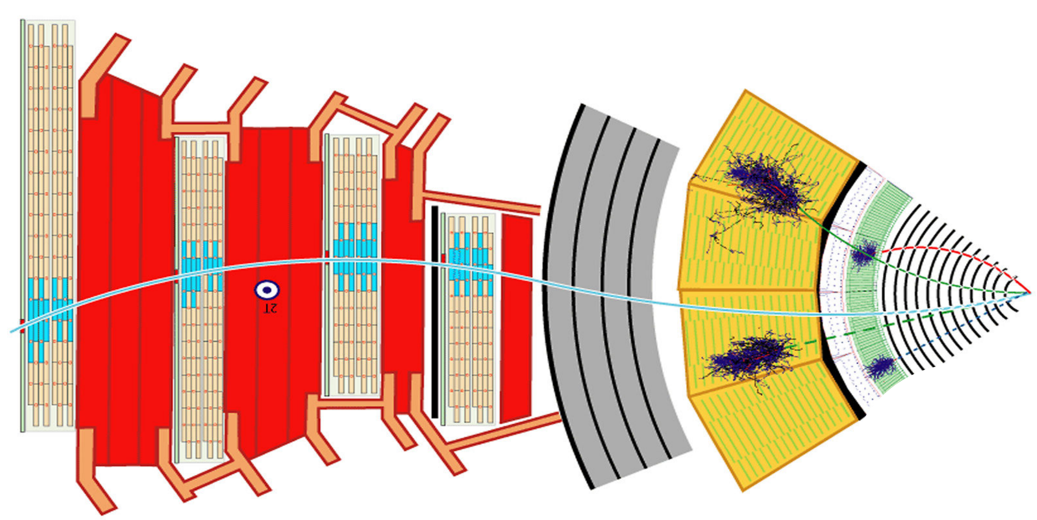
hard process

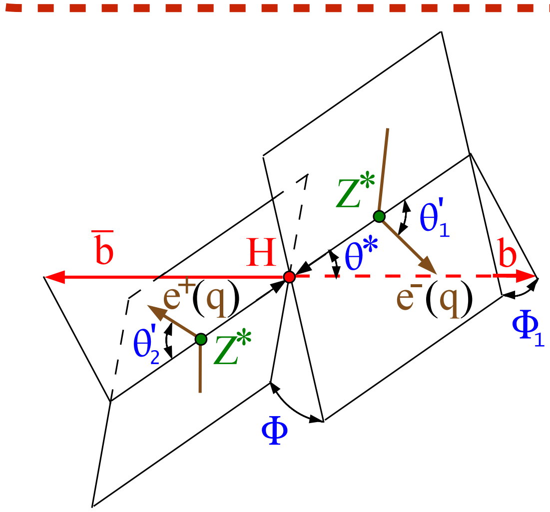
matrix elements

EFT params $\vec{\theta}$

parton mom $\vec{x}_{\rm part}$







Observables

$$\mathscr{P}(\vec{x}_{\rm reco}\,|\,\vec{\theta}) = \int d\vec{x}_{\rm part}\;\; p(\vec{x}_{\rm reco}\,|\,\vec{x}_{\rm part})\;\; \mathscr{P}(\vec{x}_{\rm part}\,|\,\vec{\theta})$$

reco observables

- typical SM observables (to suppress background)
- EFT-sensitive observables (e.g. angular, q^2, etc)
- optimized observables (matrix element, machine learning)
- full accessible information $\vec{x}^{\rm full}_{\rm reco}$ (e.g. all four-vectors)

Example: VBF $\Delta\Phi_{JJ}$ (EFT-sensitive)

EFT:

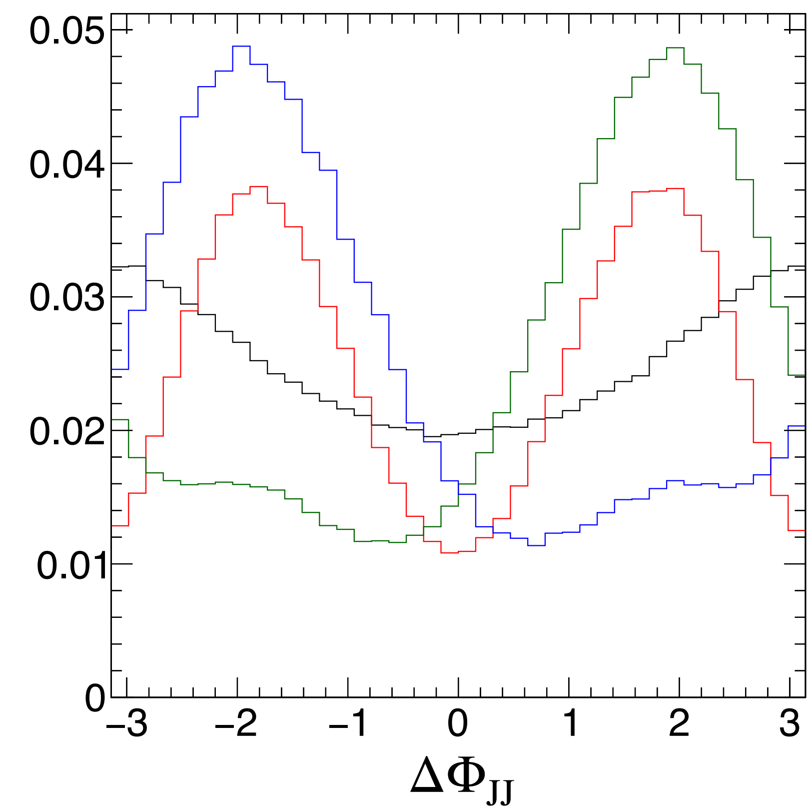
- new tensor structures
- higher q dimensions

SM — $(\theta_0, 0)$

CP-odd — $(0, \theta_1)$

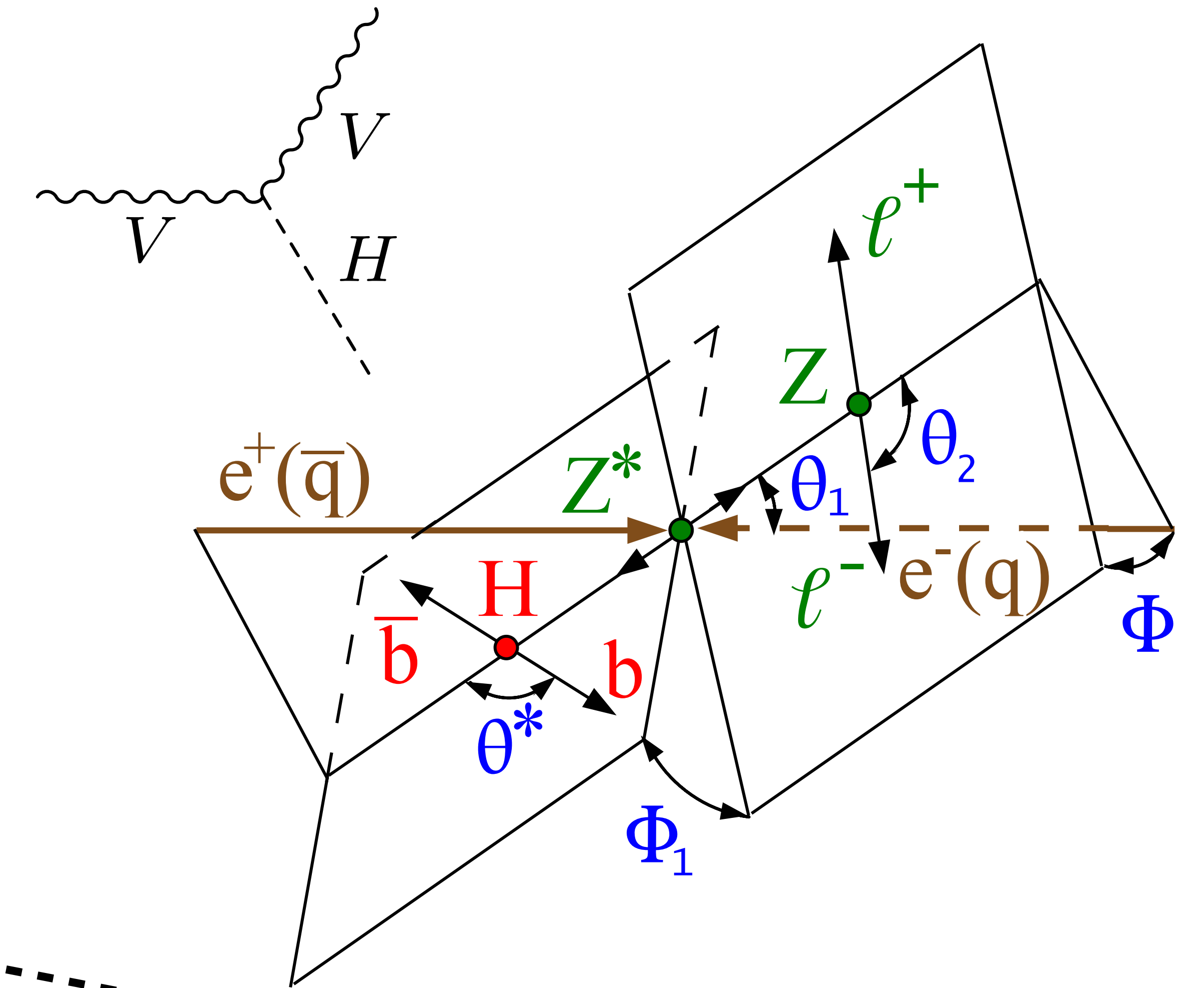
+mix — $(\theta_0, +\theta_1)$

- mix — $(\theta_0, -\theta_1)$



Observables

Example: VH process on LHC



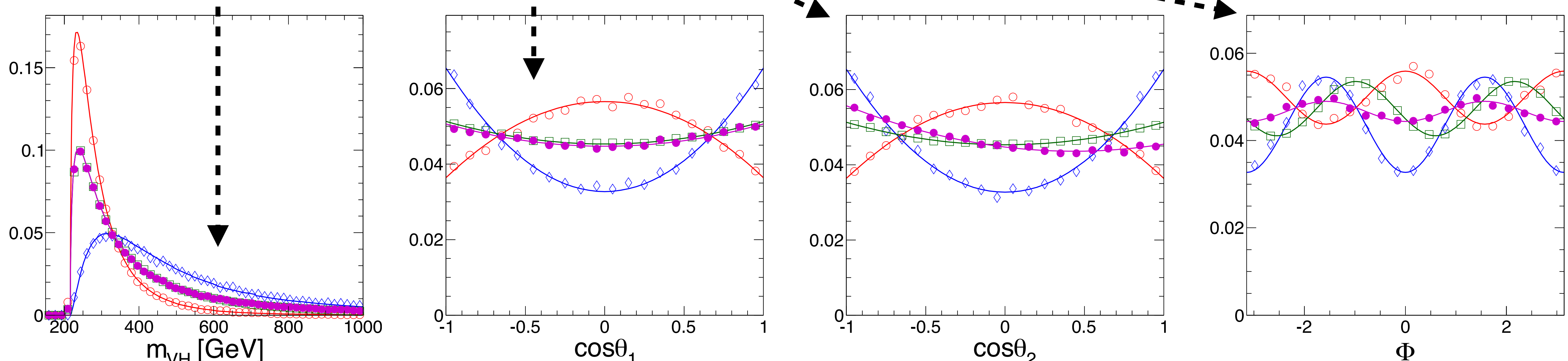
— full information $\vec{x}^{\,\text{full}}_{\text{reco}}$: take all below

— SM observables: e.g. m_{bb}, $m_{\ell\ell}$

(useless for EFT)

— EFT-sensitive observables

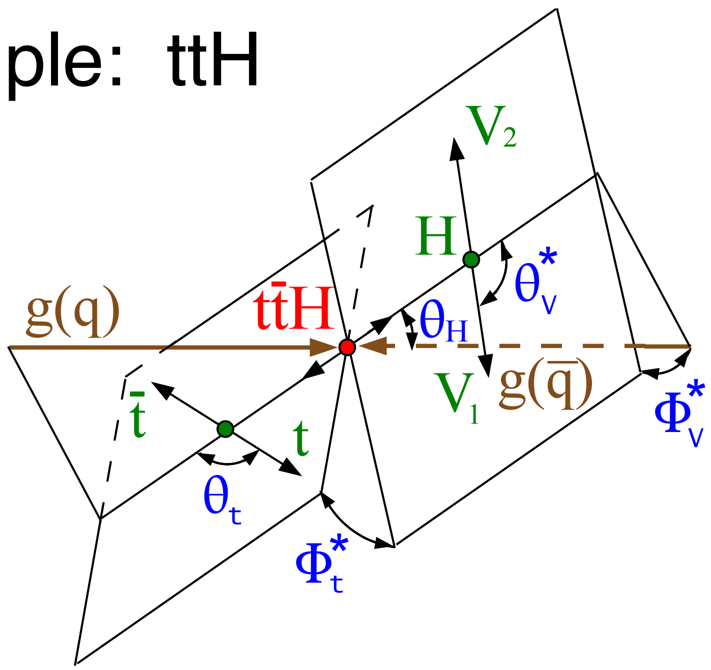
q^2- enhancement   angular

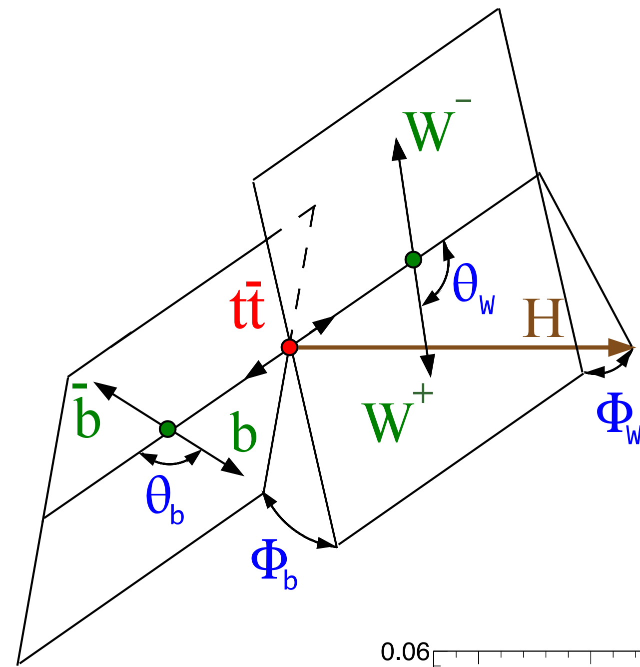


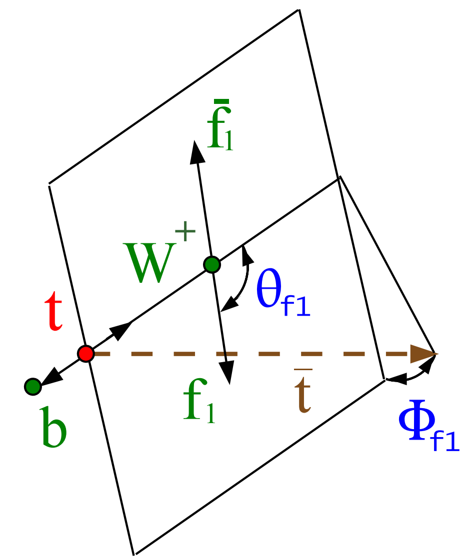
arXiv:1309.4819

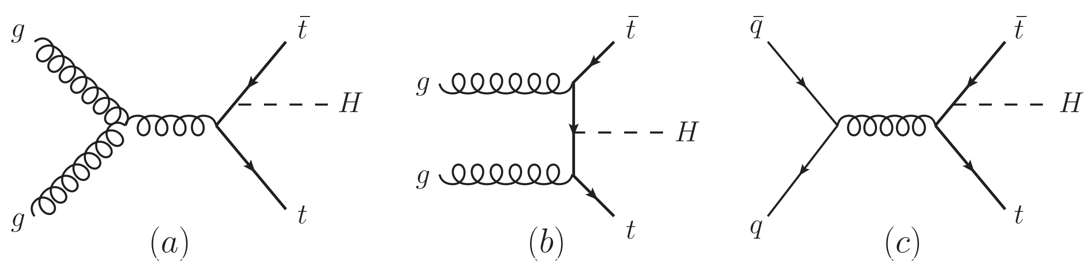
Observables

Example: ttH

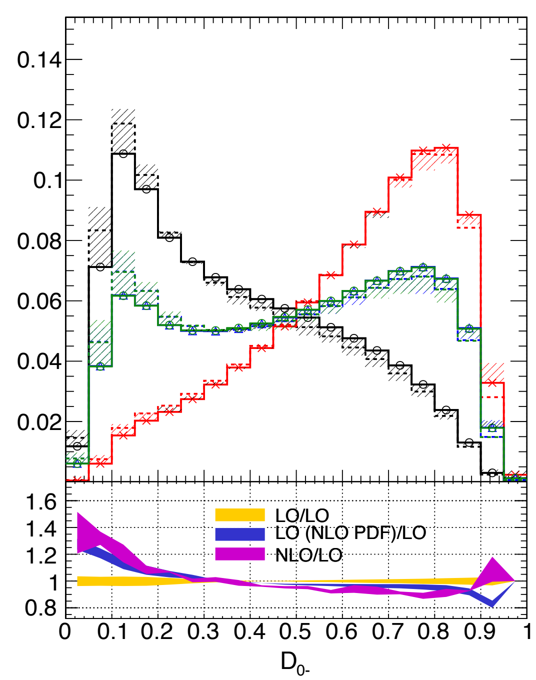


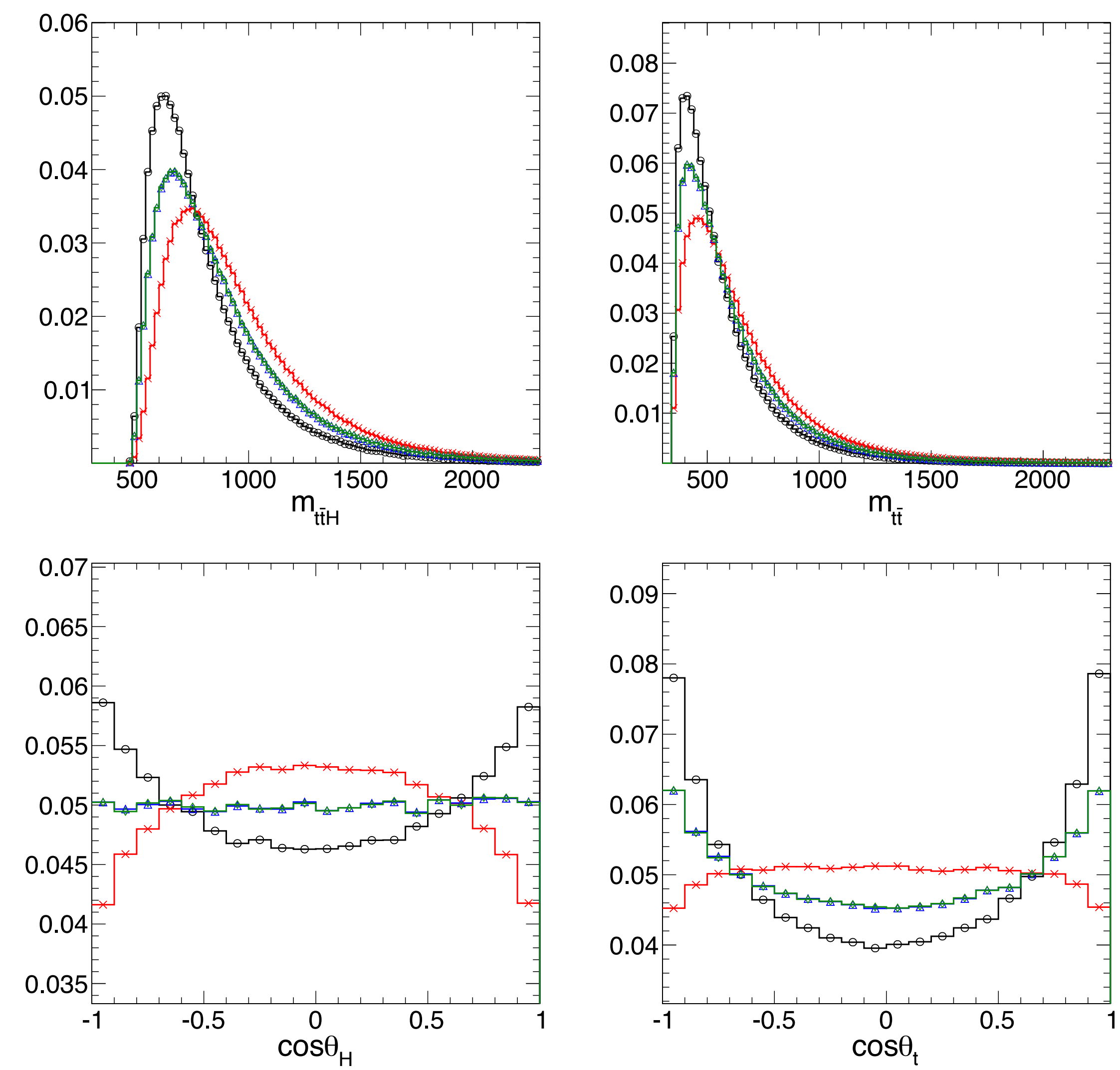






- ~20 observables
- ~1 optimal observable if consider 2 Operators (κ_t vs $\tilde{\kappa}_t$)



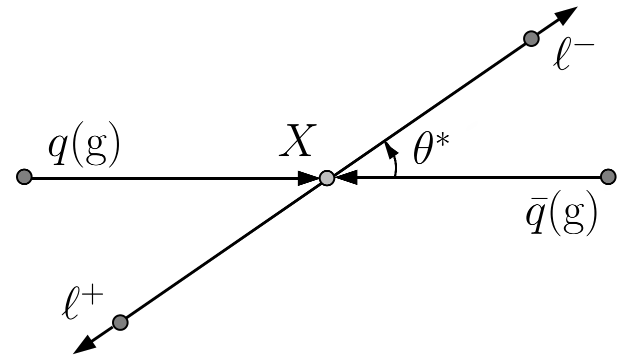


arXiv:1606.03107

Measurement: "Matrix Element Method"

- "Best" measurement: — full accessible information $\vec{x}^{\,\rm full}_{\rm reco}$ (e.g. 4-vectors)

- Example: 1st $\sin^2\theta_W$ on LHC (arXiv:1110.2682)



(θ^*, Y, s)

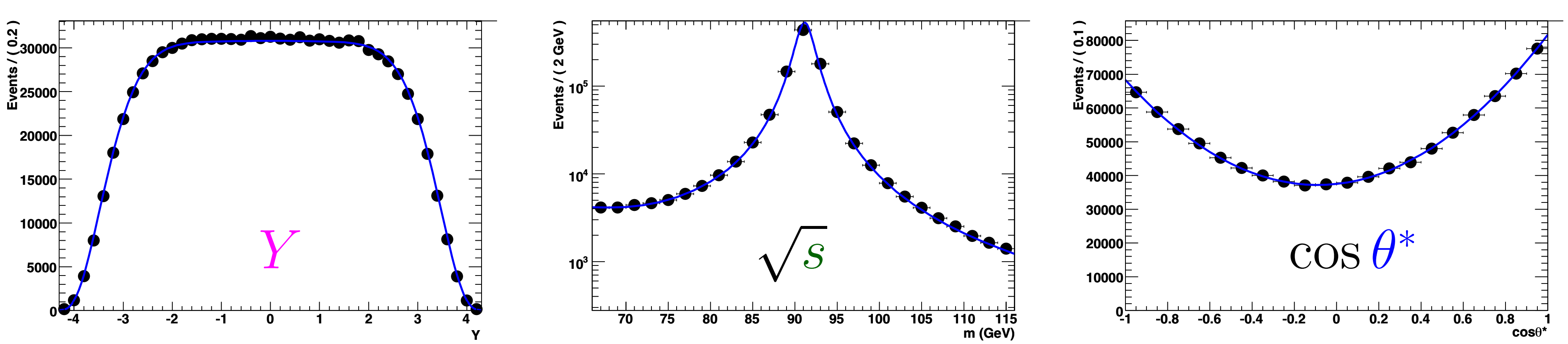
— hard process "matrix elements" (MEM)

$$\frac{d\sigma_{pp}(Y, s, \theta^*)}{dY\, ds\, d\cos\theta^*} \propto \frac{1}{s_{pp}} \sum_{q=udscb} \hat{\sigma}_{q\bar{q}}(s, \theta^*)\, \tilde{f}_q\left(e^{Y}\sqrt{s/s_{pp}}, s\right)\, \tilde{f}_{\bar{q}}\left(e^{-Y}\sqrt{s/s_{pp}}, s\right)$$

depends on $\sin^2\theta_W$ & C_i

— detector effects, reconstruction

$$\mathcal{P}_{\rm detect}(\theta^*, Y, s) = \mathcal{G}(\theta^*, Y, s) \times \int_{-\infty}^{+\infty} dx\, \mathcal{R}(x)\, \mathcal{P}_{\rm observe}(Y, s-x, \theta^*)$$



thesis of N.Tran

Measure: MEM and Simulation-Based Inference

- not the same as optimized observables (though can be used to compute):
  — ME or ML observables can be used in any approach (e.g. differential)

- Matrix Element Method (MEM) — compute the likelihood from first principles

$$\mathscr{P}(\vec{x}_{\rm reco}\,|\,\vec{\theta}) = \int {\rm d}\vec{x}_{\rm part}\, p(\vec{x}_{\rm reco}\,|\,\vec{x}_{\rm part})\,\mathscr{P}(\vec{x}_{\rm part}\,|\,\vec{\theta})$$

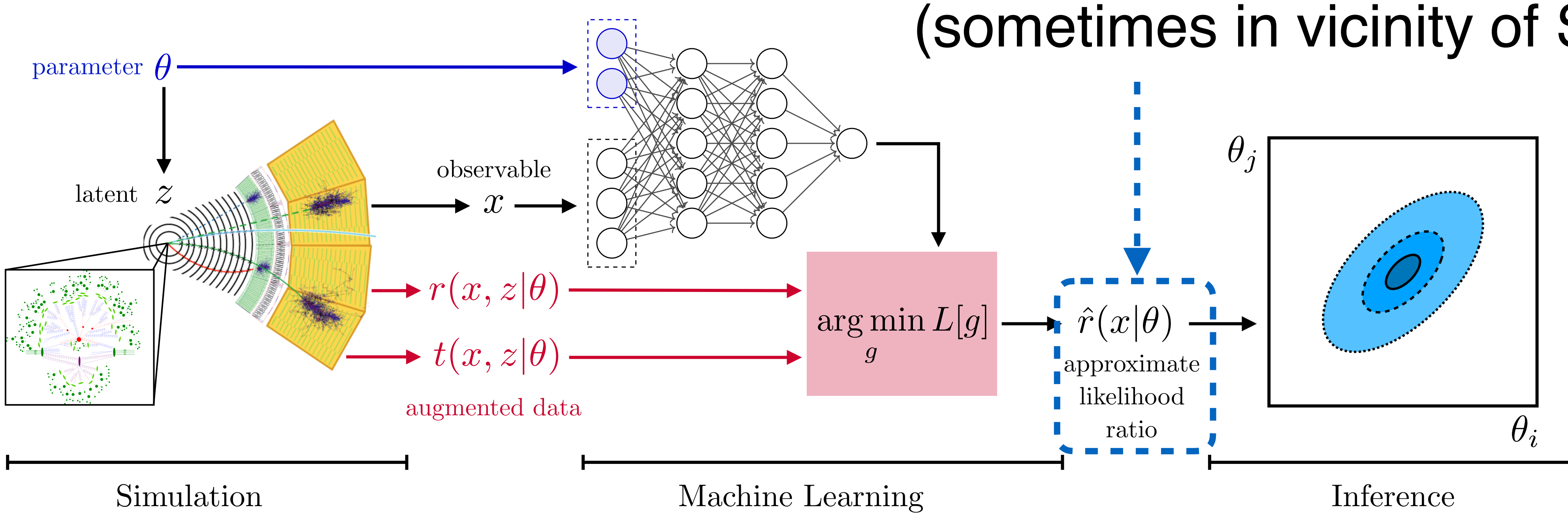
full info

ideal for EFT, but: hard to model transfer function p, ME not available for all processes… (e.g. backgrounds)

popular in Flavor, few examples in EW, top, Higgs

- Simulation-based (ML) inference — learn the full likelihood ratio (sometimes in vicinity of SM)

see talk by Harrison Prosper



Measurements: Template fit

MEM and ML aside:

- most measurements are based on

  templates of Observables $\mathscr{P}_k(\vec{x}_{\rm reco})$

$$\mathscr{P}(\vec{x}_{\rm reco}\,|\,\vec{\theta}) \propto \mathscr{P}_0(\vec{x}_{\rm reco}) + \sum_k \left(\frac{2\theta_k}{\theta_0}\right) \mathscr{P}_k(\vec{x}_{\rm reco}) + \dots$$

- Single-step (folded)

— can be optimal and unbiased

— most difficult and no re-interpretation

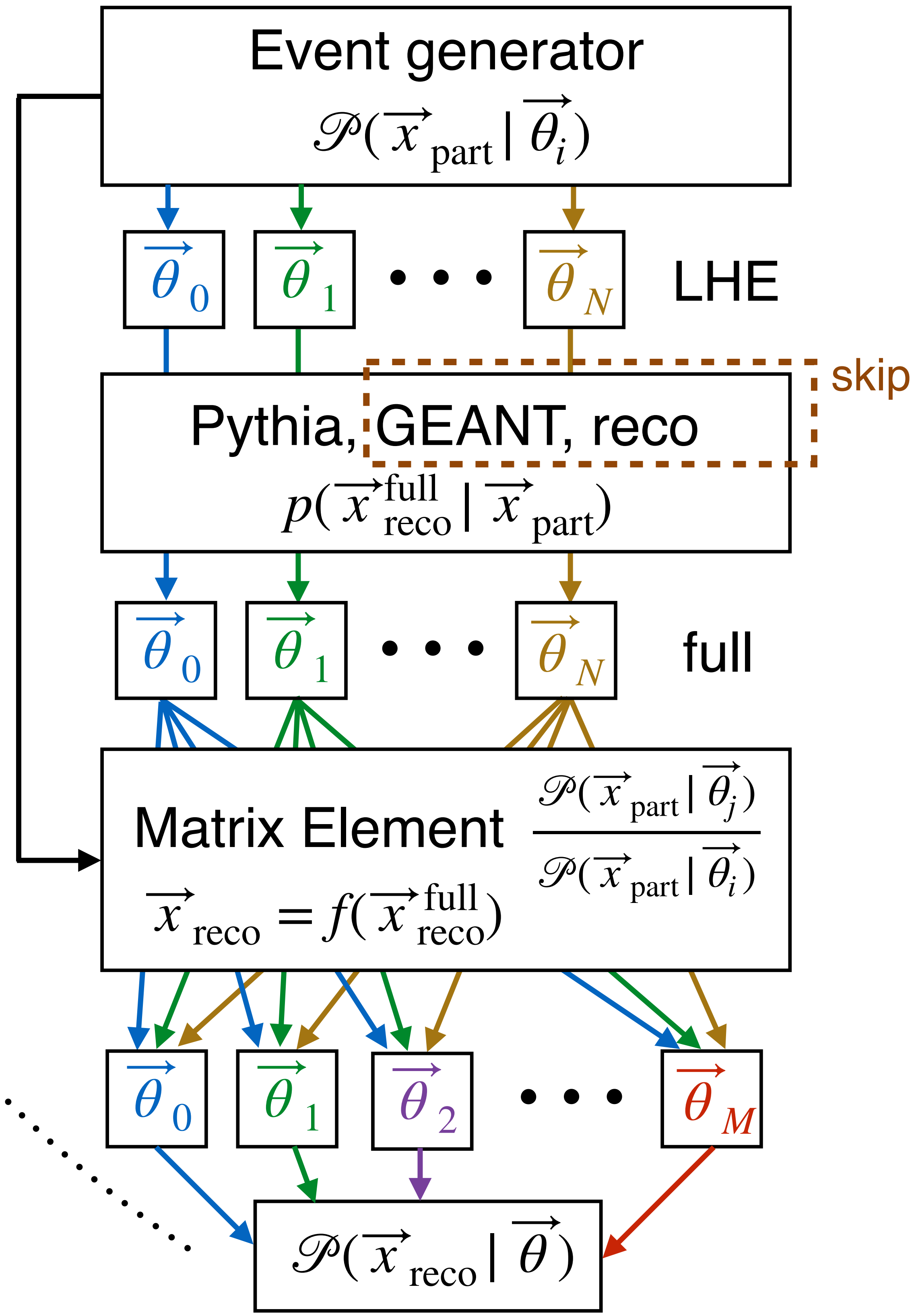
example: direct fullsim fit for θ_i

- Two-step (unfolded)

— easier and open for re-interpretation

— not full information, SM assumption

example: differential, STXS



Observables for Unfolded Measurements

- Differential cross sections — detector corrected measurement

  — historically tools for theorists to test calculations and MC tuning

  — more recently EFT applications — as step-1 in interpretation (shape dependence of Observables)

- Considerations for Observables:

— best with diagonal response matrix in the unfolding procedure

— best with flat acceptance effects: biased to SM in unfolding — often small

step-1: “unfold” to parton-level distribution

$\vec{x}'_{\rm part} \subset \vec{x}_{\rm part}$

$$\mathcal{P}(\vec{x}_{\rm reco}|\vec{\theta}) = \int {\rm d}\vec{x}'_{\rm part}\; p'(\vec{x}_{\rm reco}|\vec{x}'_{\rm part};\vec{\theta})\,\mathcal{P}(\vec{x}'_{\rm part}|\vec{\theta})$$

usually assume SM θ_0

— often exclude optimized Observables, but some examples exist

— EFT effect in background — best with high S/B

Which Observables to use?

- — typical SM observables (to suppress background)
- — EFT-sensitive observables (e.g. angular, q^2, etc)
- — optimized observables (matrix element, machine learning)
- — full accessible information $\vec{x}^{\,\rm full}_{\rm reco}$ (e.g. all four-vectors)

— full information is the best, input to MEM, ML
but hard to deal with $ND, N \gg 1$, e.g. in templates

— optimized Observables: pack full information in $1D$ optimal for one target
works if the number of targets is small (e.g. most sensitive C_i)

— SM or EFT observables: most often used in the unfolded measurements
as input to step-2

— observable choice does not limit its usage
(e.g. differential measurement of an optimized Observable is an option)

Illustration of Observables

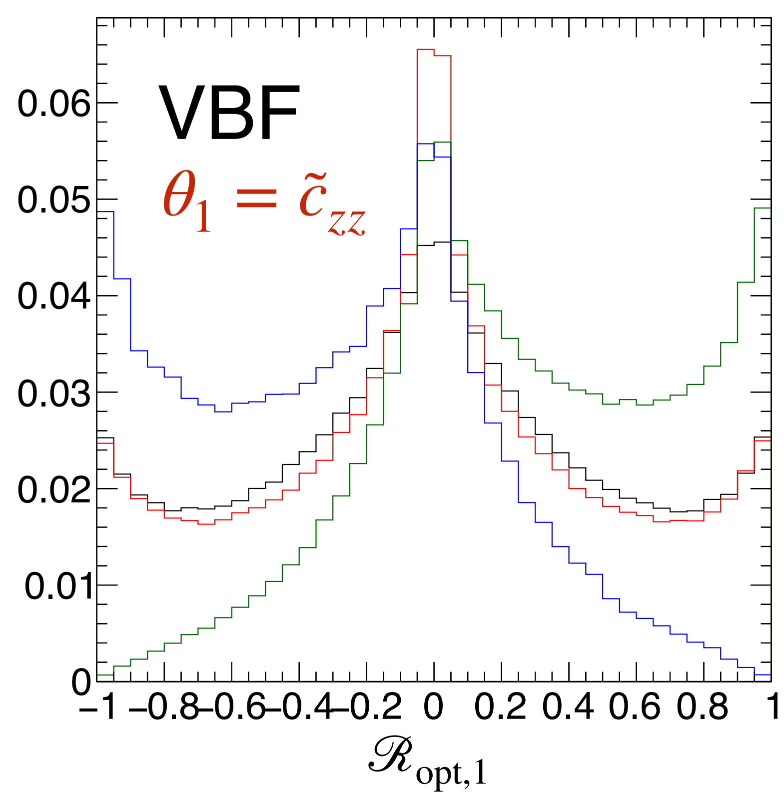
- — typical SM observables (to suppress background)
- — EFT-sensitive observables (e.g. angular, q^2, etc)
- — optimized observables (matrix element, machine learning)
- — full accessible information $\vec{x}^{\,\rm full}_{\rm reco}$ (e.g. all four-vectors)

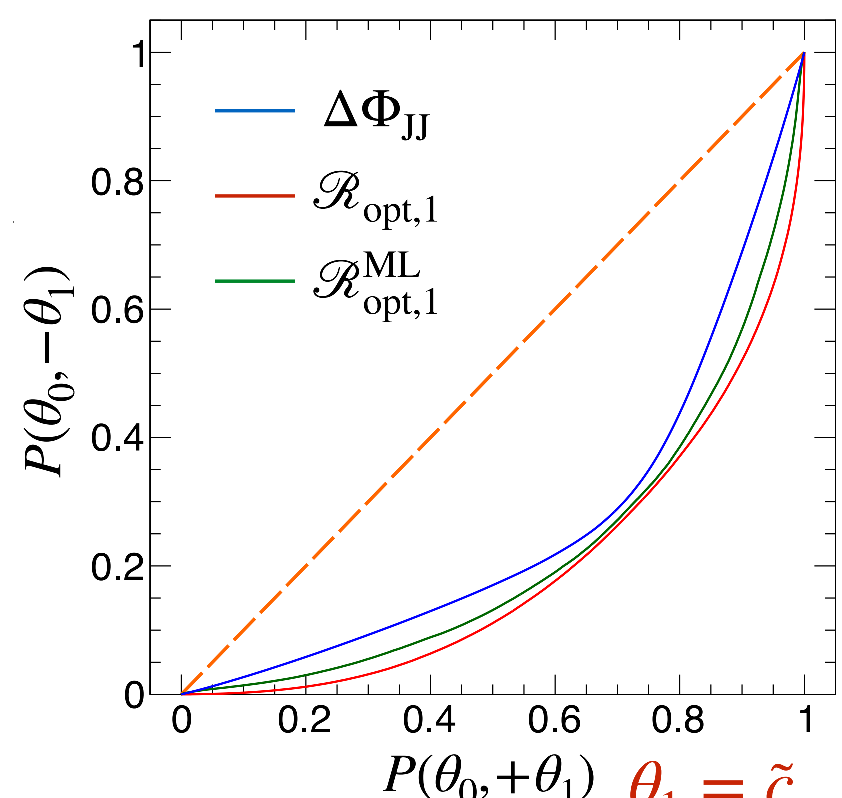
Performance

optimal Observable

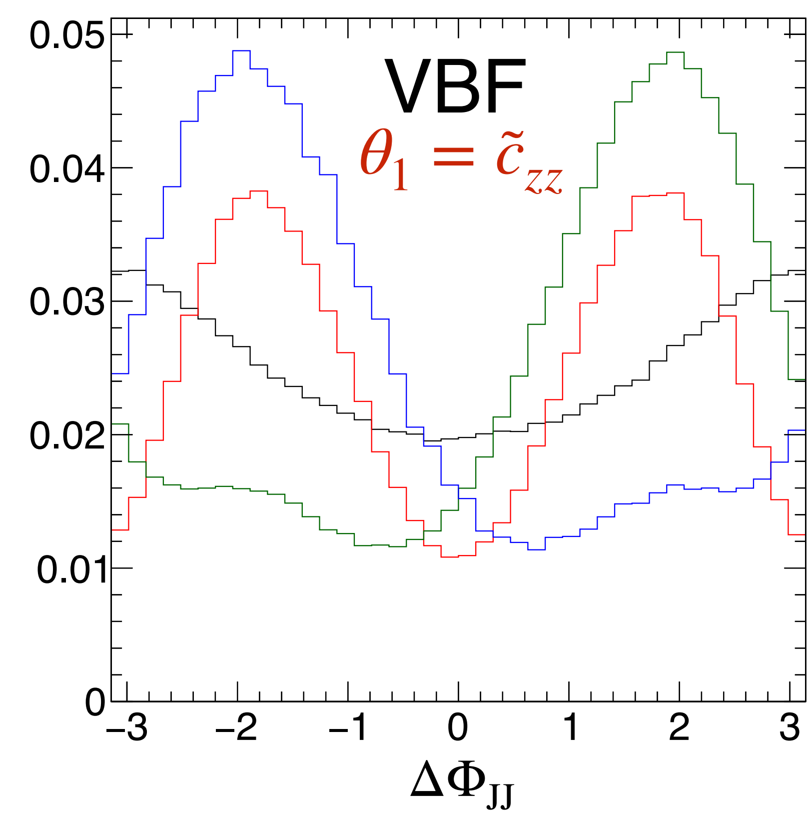
application to $\theta_1 = \tilde{c}_{zz}$

EFT observable





$\theta_1 = \tilde{c}_{zz}$



Optimized Observables: type 2

$$\mathscr{P}(\vec{x}_{\rm reco}\,|\,\vec{\theta}) \propto \mathscr{P}_0(\vec{x}_{\rm reco}) + \sum_k \left(\frac{2\theta_k}{\theta_0}\right)\mathscr{P}_{0k}(\vec{x}_{\rm reco}) + \sum_k \left(\frac{\theta_k}{\theta_0}\right)^2 \mathscr{P}_k(\vec{x}_{\rm reco}) + \sum_{i<j}\left(\frac{2\theta_i\theta_j}{\theta_0^2}\right)\mathscr{P}_{ij}(\vec{x}_{\rm reco})$$

$$\mathcal{R}_{\rm opt,2} = \frac{\mathcal{P}_1(\vec{x}^{\rm\,full}_{\rm reco})}{\mathcal{P}_0(\vec{x}^{\rm\,full}_{\rm reco}) + c\cdot\mathcal{P}_1(\vec{x}^{\rm\,full}_{\rm reco})}$$

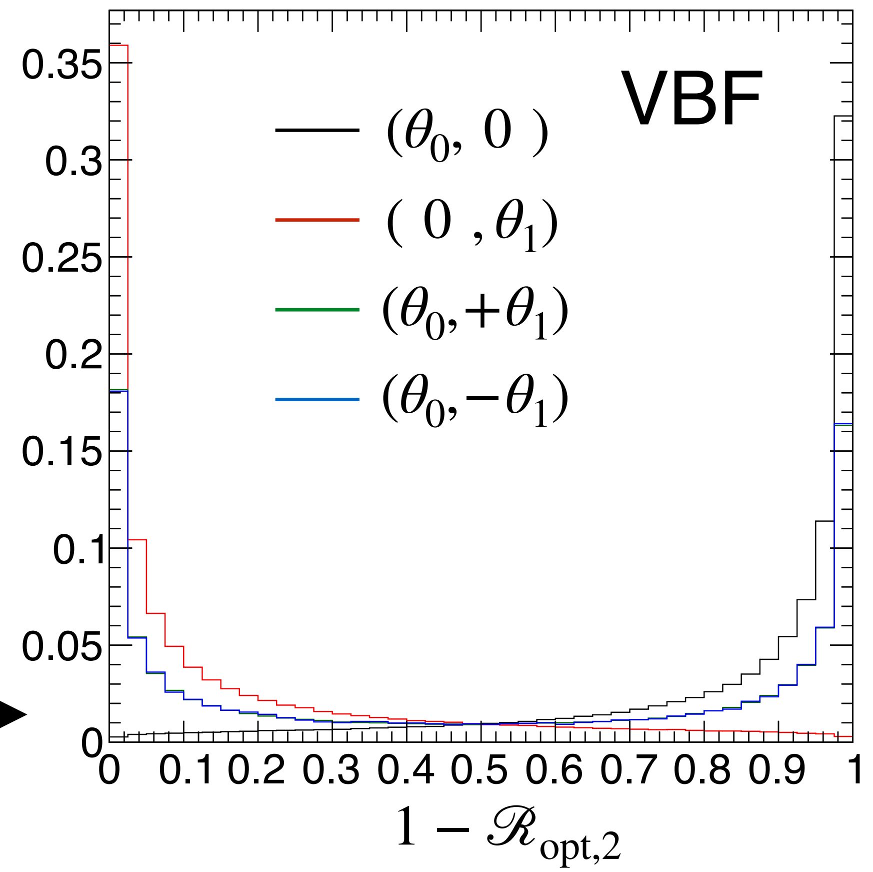
$\mathscr{P}_0, \mathscr{P}_1$
"matrix elements"
for models θ_0, θ_1

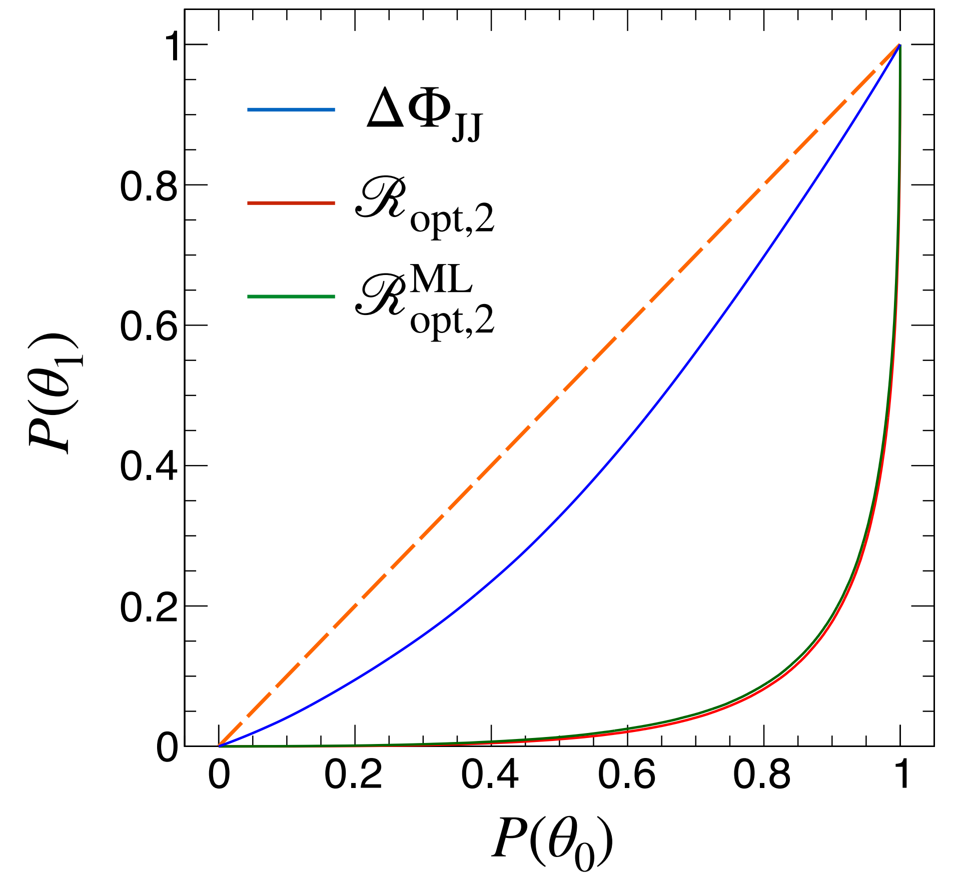
- Type-2 Optional Observable with Matrix Elements:
  - — classical signal-to-background
  - — in EFT: SM term vs. quadratic term

- ML equivalent: parton shower, detector effects

train against 2 samples

e.g. in VBF: $\theta_1 = \tilde{c}_{zz}$





Optimized Observables: type 1

$$\mathscr{P}(\vec{x}_{\rm reco}\,|\,\vec{\theta}) \propto \mathscr{P}_0(\vec{x}_{\rm reco}) + \sum_k \left(\frac{2\theta_k}{\theta_0}\right)\mathscr{P}_{0k}(\vec{x}_{\rm reco}) + \sum_k \left(\frac{\theta_k}{\theta_0}\right)^2 \mathscr{P}_k(\vec{x}_{\rm reco}) + \sum_{i<j}\left(\frac{2\theta_i\theta_j}{\theta_0^2}\right)\mathscr{P}_{ij}(\vec{x}_{\rm reco})$$

$$\mathcal{R}_{\rm opt,1} = \frac{2\mathcal{P}_{01}(\vec{x}^{\rm \,full}_{\rm reco})}{\mathcal{P}_0(\vec{x}^{\rm \,full}_{\rm reco}) + c\cdot\mathcal{P}_1(\vec{x}^{\rm \,full}_{\rm reco})}$$

$\mathscr{P}_{01}$ matrix element for interference θ_0, θ_1

$c = 0$ or $\mathscr{R}_{\rm opt,1}$ & $\mathscr{R}_{\rm opt,2}$

- Type-1 Optional Observable with Matrix Elements:
  - — no “classical” analogy
  - — in EFT: SM term vs. interference (linear) term
- ML equivalent:

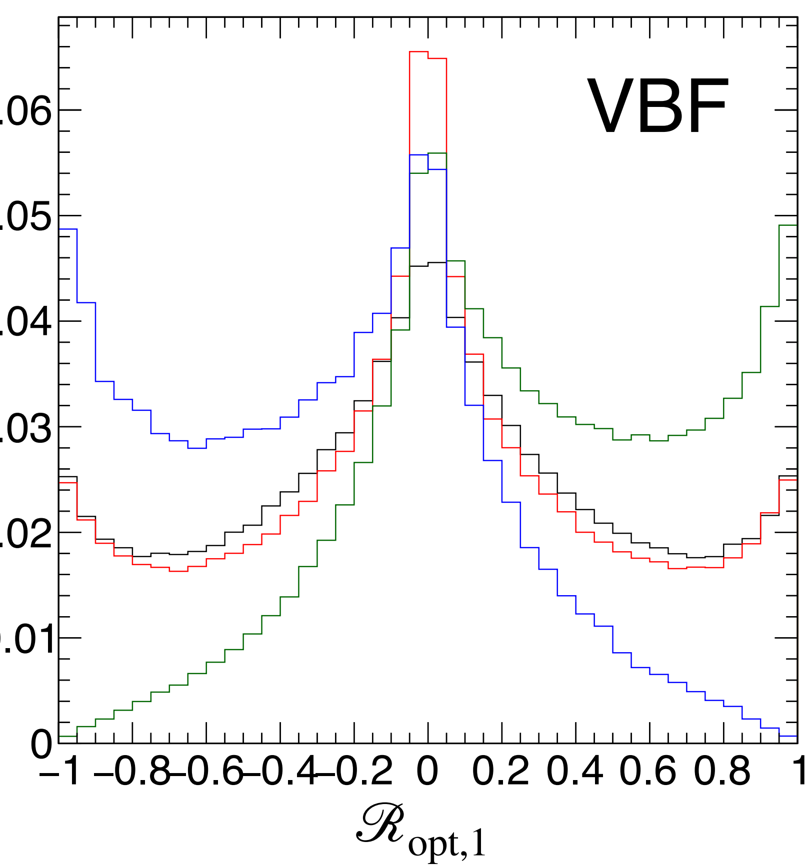
train against 2 samples

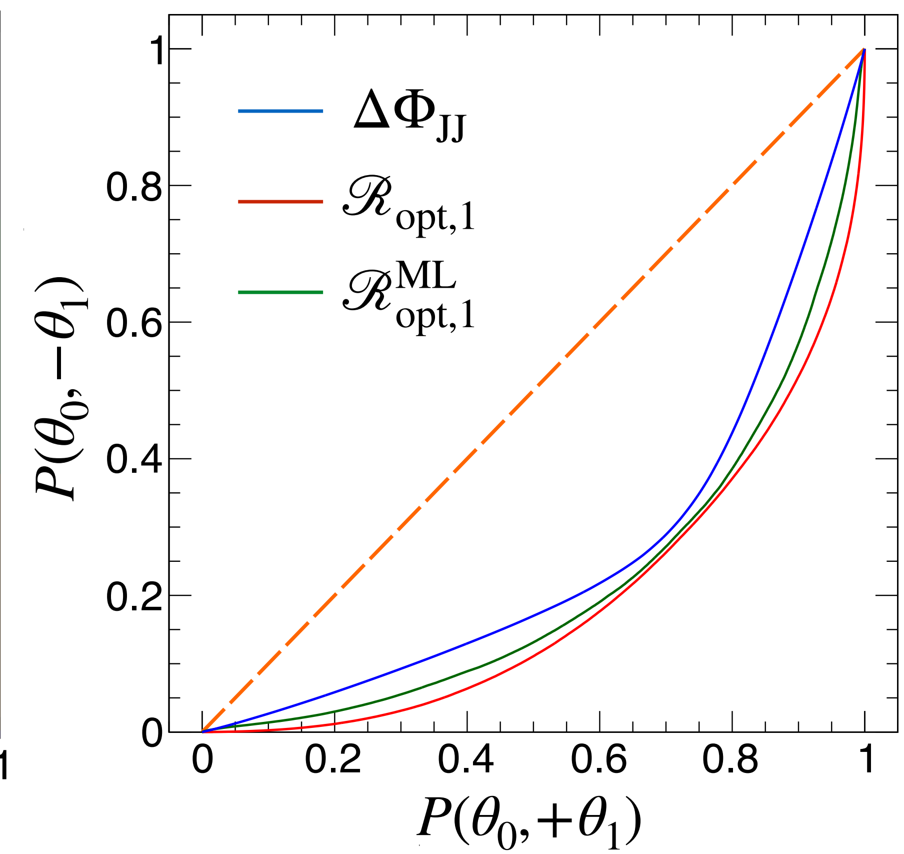
+mix vs -mix

(SM ± interference θ_0, θ_1)

(no quadratic term for $c = 0$)

e.g. in VBF: $\theta_1 = \tilde{c}_{zz}$





Question from Organizers: ML or ME?

- When to use OO with ML:

  account for parton shower, strong reconstruction effects, missing particles

  account for permutations of particles (combinatorics)

  when ME not available (readily)

- When to use OO with ME: when all of the above is not a problem

- Optimal Observable with ML:

  — trained on simulation based on ME

  — guided by Matrix Element Approach

  (a) use full information as input ! ($\vec{x}^{\text{full}}_{\text{reco}}$)

  (b) follow ME prescription in training

pair $(\mathscr{R}_{\text{opt},1}, \mathscr{R}_{\text{opt},2})$ guaranteed to be optimal for any size of 1 coupling

$$\begin{cases} \mathcal{R}_{\text{opt},2} = \dfrac{\mathcal{P}_1(\vec{x}^{\text{full}}_{\text{reco}})}{\mathcal{P}_0(\vec{x}^{\text{full}}_{\text{reco}}) + c \cdot \mathcal{P}_1(\vec{x}^{\text{full}}_{\text{reco}})} \\[2ex] \mathcal{R}_{\text{opt},1} = \dfrac{2\mathcal{P}_{01}(\vec{x}^{\text{full}}_{\text{reco}})}{\mathcal{P}_0(\vec{x}^{\text{full}}_{\text{reco}}) + c \cdot \mathcal{P}_1(\vec{x}^{\text{full}}_{\text{reco}})} \end{cases}$$

$(0 \leq c \leq 1)$

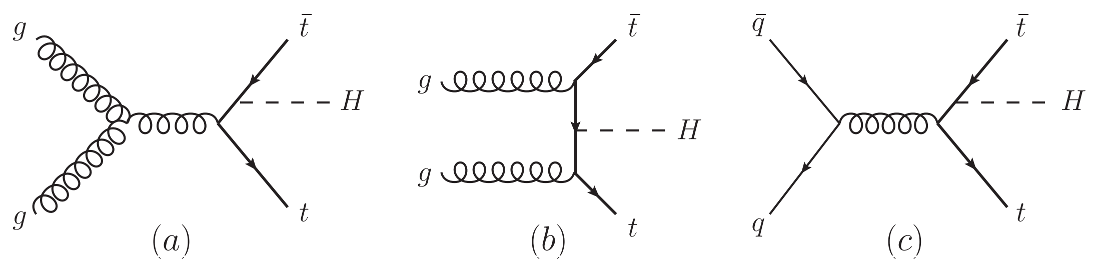
$\mathscr{R}^{\text{ML}}_{\text{opt},2}$ → train $100\,\%$ state $\mathscr{O}$ against SM

$\mathscr{R}^{\text{ML}}_{\text{opt},1}$ → train $50\,\%$ state $\mathscr{O}$/SM against $-50\,\%$

($c = 1$, remove quadratic term for $c = 0$)

Using Optimized Observables with ML

- When to use OO with ML:



- first CP analysis in $t\bar{t}H$ process

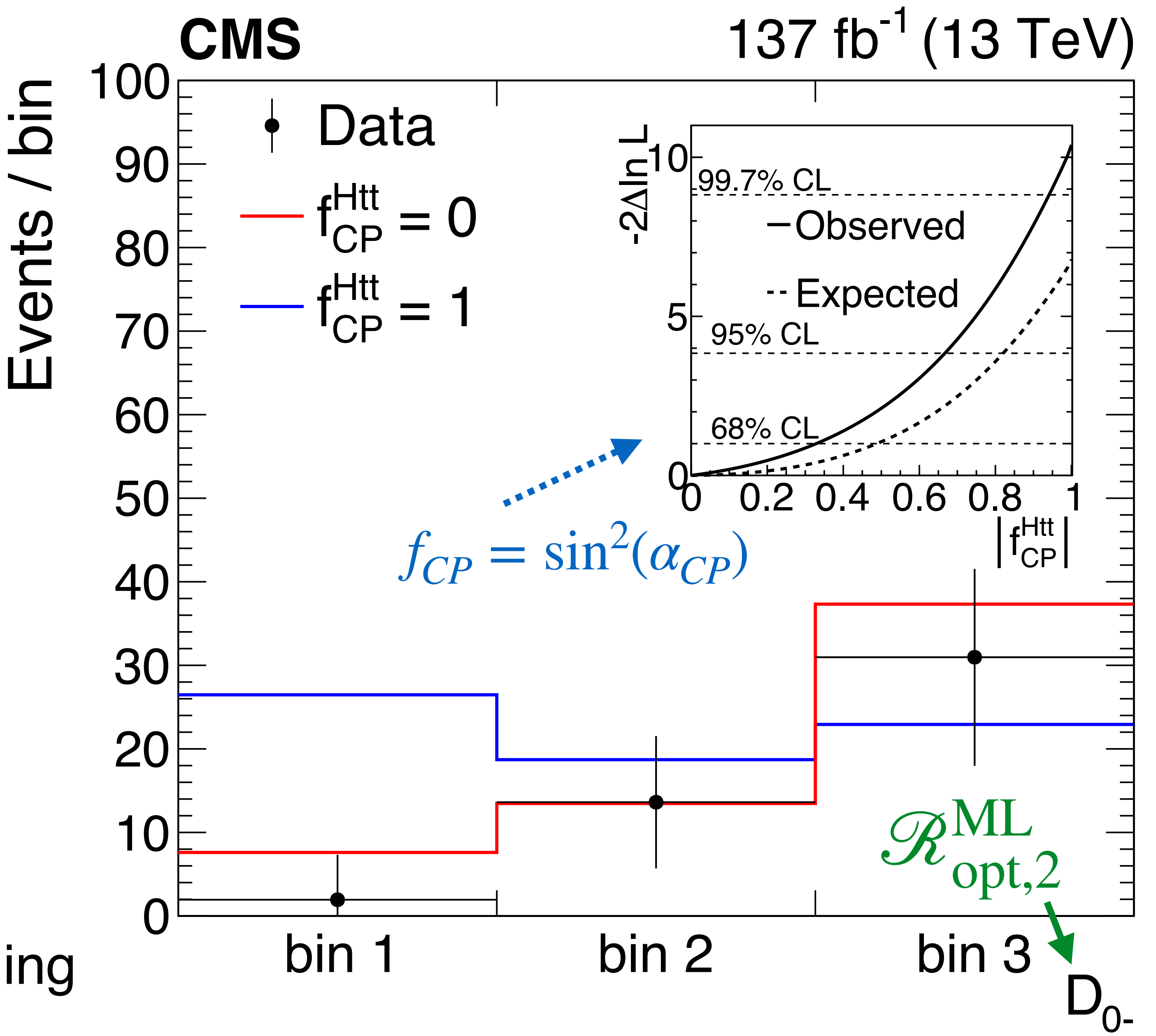
even though feasibility was done with MELA
arXiv:1606.03107

— permutation of particles
— lost particles
motivated ML

(a) used full information as input !

(b) followed ME prescription in training

CMS arXiv:2003.10866



$f_{CP} = \sin^2(\alpha_{CP})$

$\mathscr{R}^{ML}_{opt,2}$ → D_{0-}

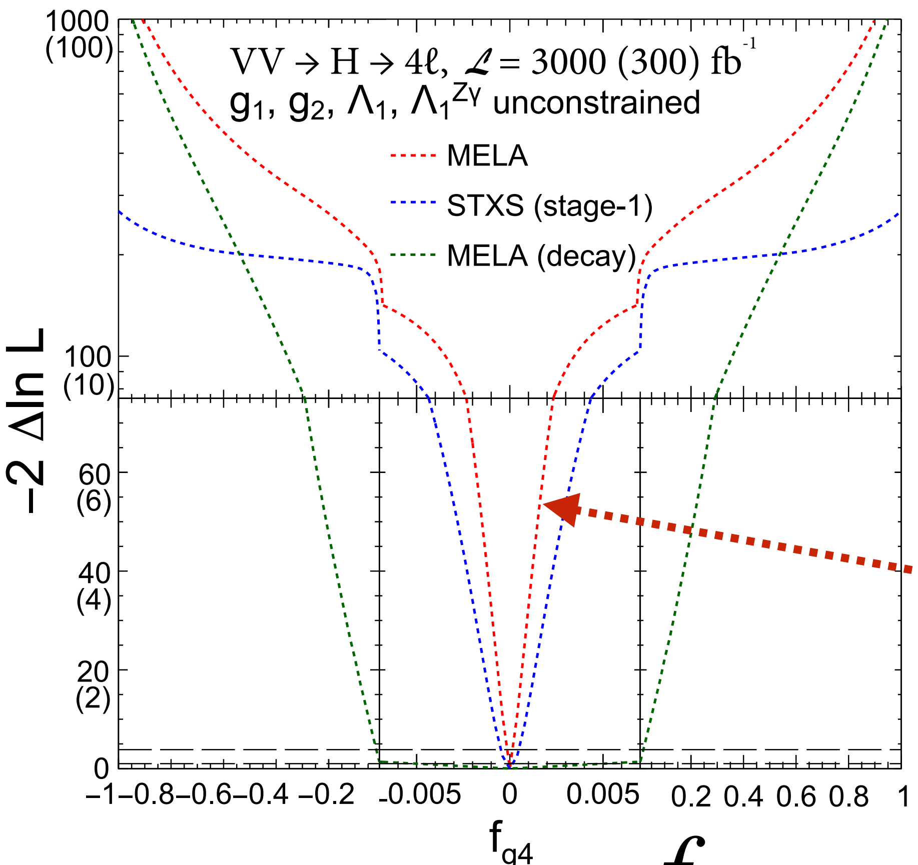
Using Optimized Observables with MELA

- MELA (Matrix Element Likelihood Approach) — tutorial yesterday by Mohit Srivastav
- Example process $VV \to H \to 4\ell$

— all production mechanisms: ggH, VBF, VH, $t\bar{t}H$, tH, $b\bar{b}H$

— three approaches to binning: dedicated, decay-only, STXS v1.1

— 5 HVV, 2 Hgg, 2 Htt couplings free ⇒ Optimal Observables targeting each

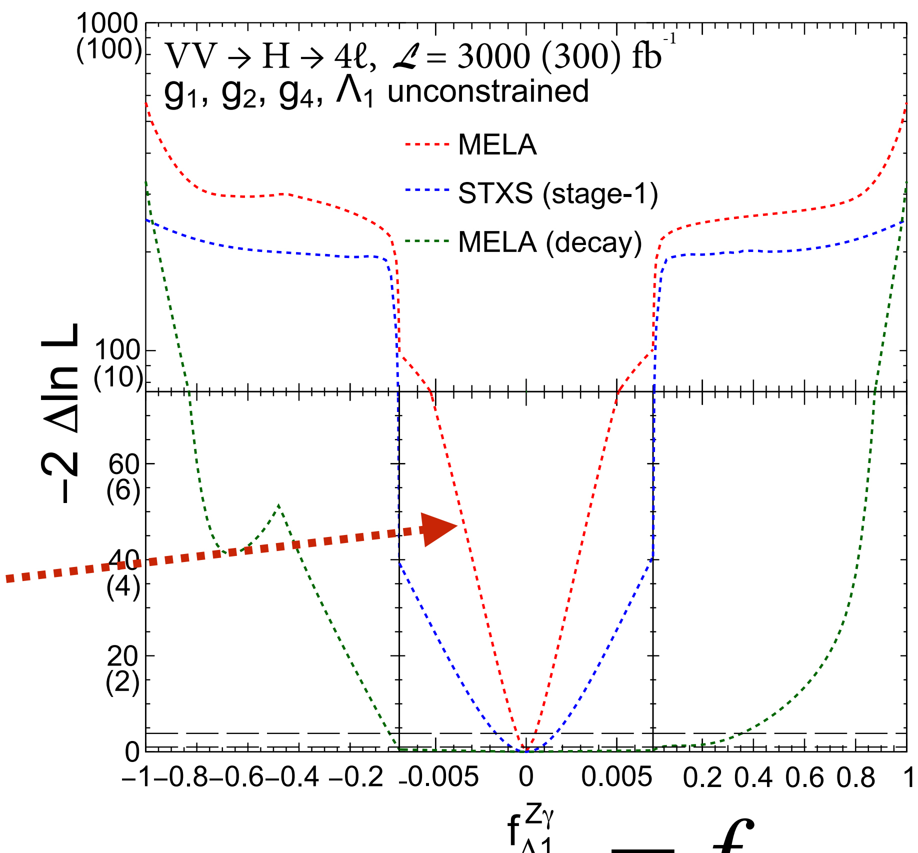


$= f_{\tilde{c}_{zz}}$

$\times 2-3$ tighter constraints with Optimal MELA vs Differential STXS

production information $VV \to H$

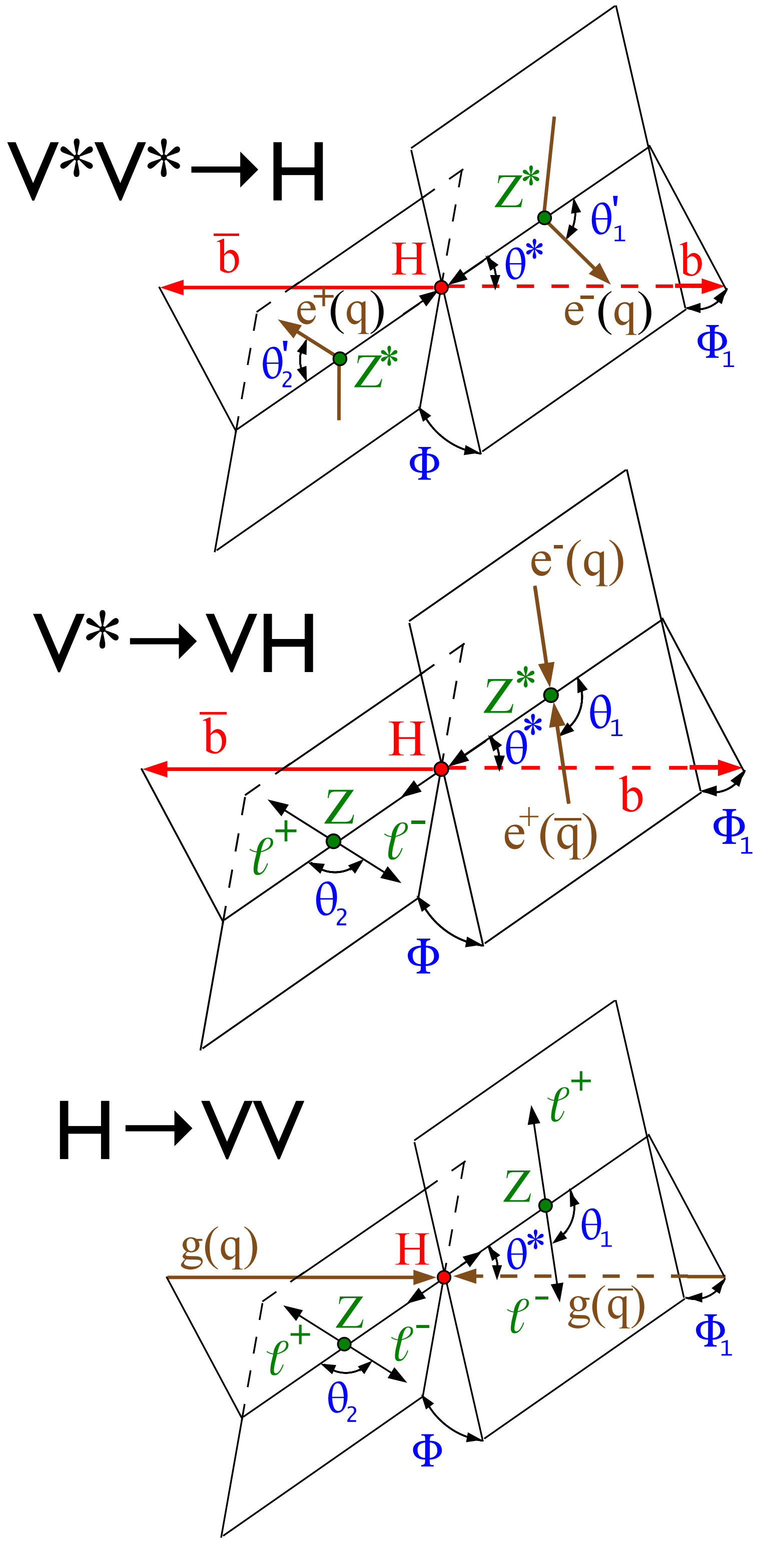
arXiv:2002.09888

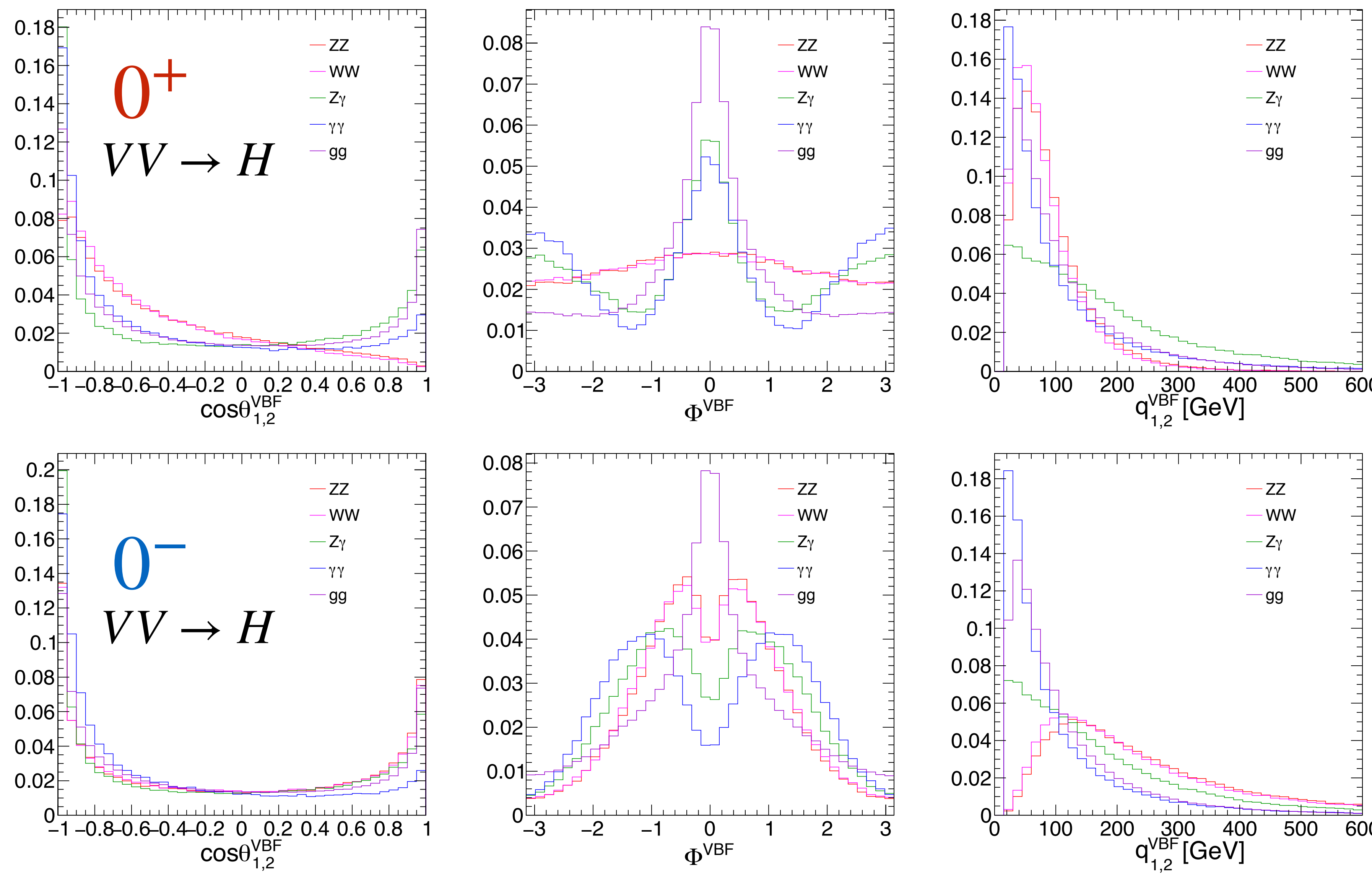


$= f_{c_{z\Box}}$

Using Optimized Observables with MELA

- Take VBF topology ⇒ $HWW, HZZ, HZ\gamma, H\gamma\gamma, Hgg$ couplings
  - — unique multi-D. kinematics in each case:
    (want to use it all to isolate each operator)





hep-ph arXiv:2002.09888

use JHUGen/MELA in this study, but ideas are generic

Question from Organizers:

- How >1 observables can be selected and designed to better constrain coefficients?
- Ideally 1 or 2 observables per target operator, if distinct and important enough
  - — up to 7 observables use in [CMS-HIG-19-009]

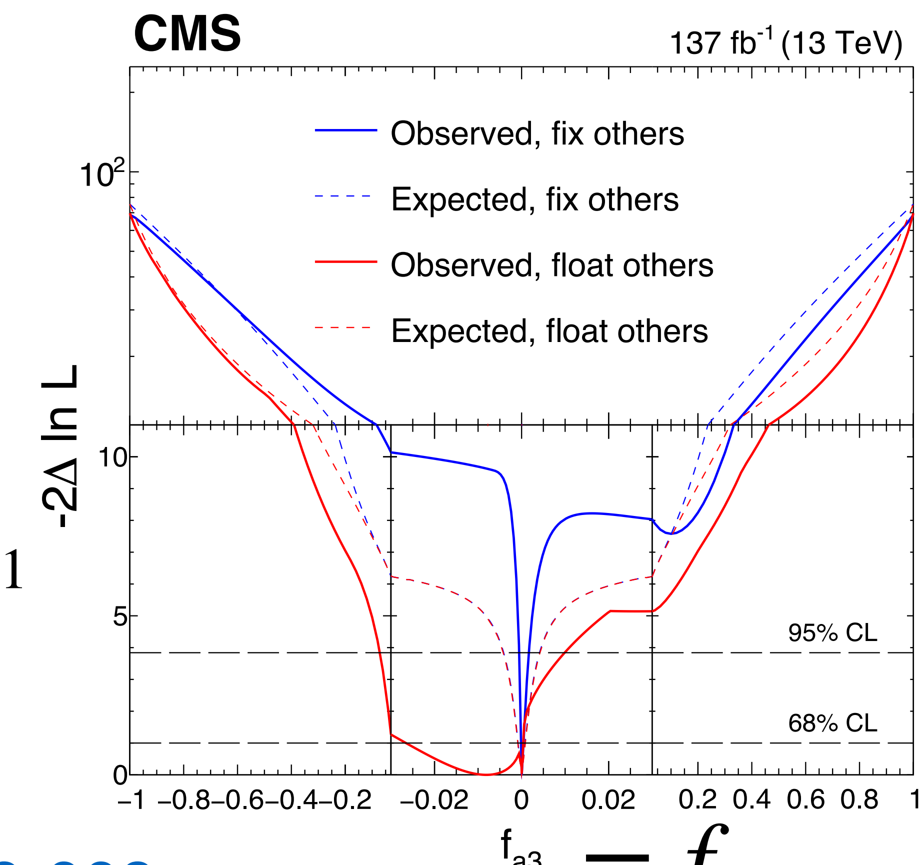
    depending on category (targeting a process):

    1 to suppress background

    2 pairs $(\mathscr{R}_{\mathrm{opt},1}, \mathscr{R}_{\mathrm{opt},2})$ for $c_{zz}, \tilde{c}_{zz}$

    2 of $\mathscr{R}_{\mathrm{opt},2}$ for $\delta c_z, c_{z\Box}$ due to correlation to $\mathscr{R}_{\mathrm{opt},1}$

    used OO of both types due to limited sensitivity



$= f_{\tilde{c}_{zz}}$

[CMS-HIG-19-009]

Boosted	$p_T^{4\ell} > 120\,\mathrm{GeV}$	$\mathcal{D}_{\mathrm{bkg}}, p_T^{4\ell}$
VBF-1jet	$\mathcal{D}_{\mathrm{1jet}}^{\mathrm{VBF}} > 0.7$	$\mathcal{D}_{\mathrm{bkg}}, p_T^{4\ell}$
VBF-2jet	$\mathcal{D}_{\mathrm{2jet}}^{\mathrm{VBF}} > 0.5$	$\mathcal{D}_{\mathrm{bkg}}^{\mathrm{EW}}, \mathcal{D}_{0h+}^{\mathrm{VBF+dec}}, \mathcal{D}_{0-}^{\mathrm{VBF+dec}}, \mathcal{D}_{\Lambda 1}^{\mathrm{VBF+dec}}, \mathcal{D}_{\Lambda 1}^{Z\gamma,\mathrm{VBF+dec}}, \mathcal{D}_{\mathrm{int}}^{\mathrm{VBF}}, \mathcal{D}_{CP}^{\mathrm{VBF}}$
VH-hadronic	$\mathcal{D}_{\mathrm{2jet}}^{\mathrm{VH}} > 0.5$	$\mathcal{D}_{\mathrm{bkg}}^{\mathrm{EW}}, \mathcal{D}_{0h+}^{\mathrm{VH+dec}}, \mathcal{D}_{0-}^{\mathrm{VH+dec}}, \mathcal{D}_{\Lambda 1}^{\mathrm{VH+dec}}, \mathcal{D}_{\Lambda 1}^{Z\gamma,\mathrm{VH+dec}}, \mathcal{D}_{\mathrm{int}}^{\mathrm{VH}}, \mathcal{D}_{CP}^{\mathrm{VH}}$
VH-leptonic	see Section 3	$\mathcal{D}_{\mathrm{bkg}}, p_T^{4\ell}$
Untagged	none of the above	$\mathcal{D}_{\mathrm{bkg}}, \mathcal{D}_{0h+}^{\mathrm{dec}}, \mathcal{D}_{0-}^{\mathrm{dec}}, \mathcal{D}_{\Lambda 1}^{\mathrm{dec}}, \mathcal{D}_{\Lambda 1}^{Z\gamma,\mathrm{dec}}, \mathcal{D}_{\mathrm{int}}^{\mathrm{dec}}, \mathcal{D}_{CP}^{\mathrm{dec}}$

Considerations for Optimized Observables

- We should be able to explore unique kinematics in our detectors

  — using dedicated tools (ME, MVA, $\Delta\Phi_{JJ}$,…) (optimal)

  — using full detector simulation of EFT effects (correct)

(1) Challenges in number of terms ⇒ bookkeeping, re-weighting (MELA)

(2) Challenges in number of observables ⇒ optimize bins

— often keep quadratic terms: to have positive probability

— may keep optimal Observable type-2: when sensitivity still limited

$$\mathcal{P}\left(\vec{x};\vec{f}\right) = \sum_{k\le l\le m\le n=1}^{K} \mathcal{P}_{klmn}\left(\vec{x}\right) \sqrt{|f_k \cdot f_l \cdot f_m \cdot f_n|}\, \mathrm{sign}(f_k \cdot f_l \cdot f_m \cdot f_n)$$

K - number of couplings, N - number of products (4 in VV→H→VV, 2 in gg→H)

total # terms = $\dfrac{(N+K-1)!}{N!(K-1)!}$ =

3 for N=2, K=2 in gg→H, ttH

15 for N=2, K=5 in decay H→VV→ 4ℓ

70 for N=4, K=5 in VV→H→VV

495 for N=4, K=9 in VV→H→VV (offshell+bkg)

linear terms = K

Considerations for Optimized Observables

- With the large number of Operators, e.g. in VBS / VBF / VH / H→VV:

$$C^{\varphi W}, C^{\varphi B}, C^{\varphi WB} \leftrightarrow c_{zz}, c_{z\gamma}, c_{\gamma\gamma}$$

$$C^{\varphi \tilde{W}}, C^{\varphi \tilde{B}}, C^{\varphi \tilde{W}B} \leftrightarrow \tilde{c}_{zz}, \tilde{c}_{z\gamma}, \tilde{c}_{\gamma\gamma}$$

$$C^{\varphi D}, C^{\varphi \Box}, \delta v \leftrightarrow \delta c_z, \delta c_w, c_{z\Box}$$

$$C^{\varphi G}, C^{\varphi \tilde{G}} = c_{gg}, \tilde{c}_{gg}$$

$$\kappa_t, \tilde{\kappa}_t, \kappa_b, \tilde{\kappa}_b$$

$$C_L^{Ztt}, C_R^{Ztt}$$

$$C^W, C^{\tilde{W}}$$

- Define the target set of EFT operators θ_i is important in advance:
  - — rotate operators to remove flat directions
  - — determine sensitive θ_i e.g. $C^{\varphi W}, C^{\varphi B}, C^{\varphi WB} \leftrightarrow c_{zz}, c_{z\gamma}, c_{\gamma\gamma}$

- Cannot keep ~20 optimal Observables
  - — define optimal Observables for a limited set of operators for a given process
  - — sometimes using direct angular and q^2 information (~full) is optimal (recall VH)

Summary

- General considerations: for observables
  - variety of approaches, some historical
  - no unique recommendation
  - awareness of pros / cons, tools
  - practical choices

- Observables for EFT
  - — from “simple” to optimized observables
  - — clear prescription if optimization is desired
    needed: clear target, choice of operators

- Observables for EFT
  - — full information — use in MEM, ML
  - — optimized Observables — in dedicated fits
  - — EFT observables — in differential / dedicated fits
  - — SM observables — in differential