Stony Brook University

Unearthing kinematic information in WH production

Based on `arXiv:1908.XXXX`

Samuel Homiller

C. N. Yang Institute for Theoretical Physics
& Brookhaven Natl. Lab

In collaboration with J. Brehmer, S. Dawson, F. Kling, and T. Plehn

APS DPF Meeting, August 1, 2019
Northeastern University

The Higgs Legacy of the LHC

With the Standard Model complete, the next steps are to measure the theory as precisely as possible

SMEFT: Parameterizes BSM effects in terms of higher dimensional operators

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_{d,k} \frac{C_k^d}{\Lambda^{d-4}} \mathcal{O}_k^d$$

In this work: truncate at d = 6, and consider effects only up to $\mathcal{O}(1/\Lambda^2)$

Goal: Understand what the legacy measurements of the LHC will tell us about BSM physics

The Higgs Legacy of the LHC

With the Standard Model complete, the next steps are to measure the theory as precisely as possible

SMEFT: Parameterizes BSM effects in terms of higher dimensional operators

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_{d,k} \frac{C_k^d}{\Lambda^{d-4}} \mathcal{O}_k^d$$

In this work: truncate at d = 6, and consider effects only up to $\mathcal{O}(1/\Lambda^2)$

Goal: Understand what the legacy measurements of the LHC will tell us about BSM physics

Where do global analyses of Higgs-Gauge Sector get their information?

A. Butter, O. Éboli, J. Gonzalez-Fraile, M. C. Gonzalez-Garcia, T. Plehn, M. Rauch, 1604.03105
J. Ellis, C. Murphy, V. Sanz, T. You, 1803.03252
A. Biekötter, D. Gonçalves, T. Plehn, M. Takeuchi, D. Zerwas, 1811.08401
A. Biekötter, T. Corbett, T. Plehn, 1812.07587

Here: consider example of WH production, in the $\ell\nu b\bar{b}$ channel

Compare traditional analysis methods with modern inference techniques

Use Information Geometry to make these questions quantitative

WH Production in the SMEFT

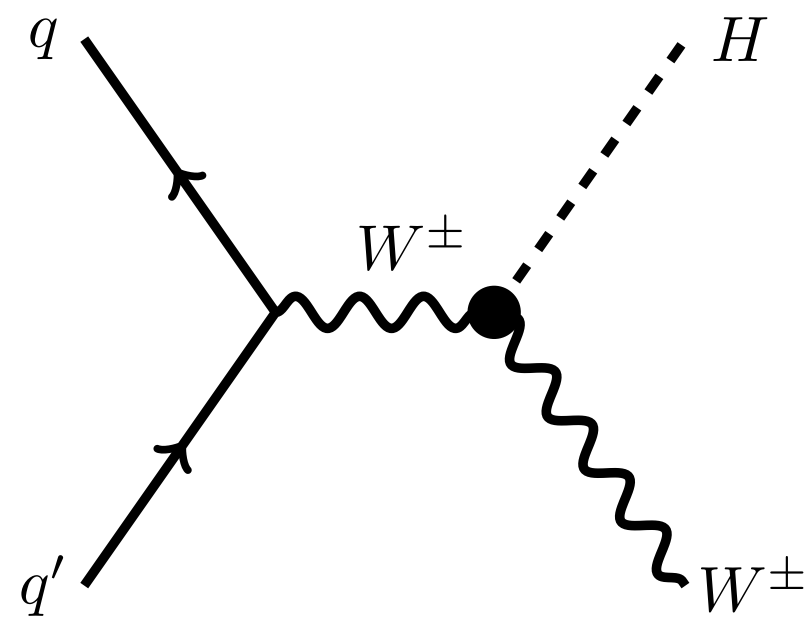
In the Warsaw Basis:

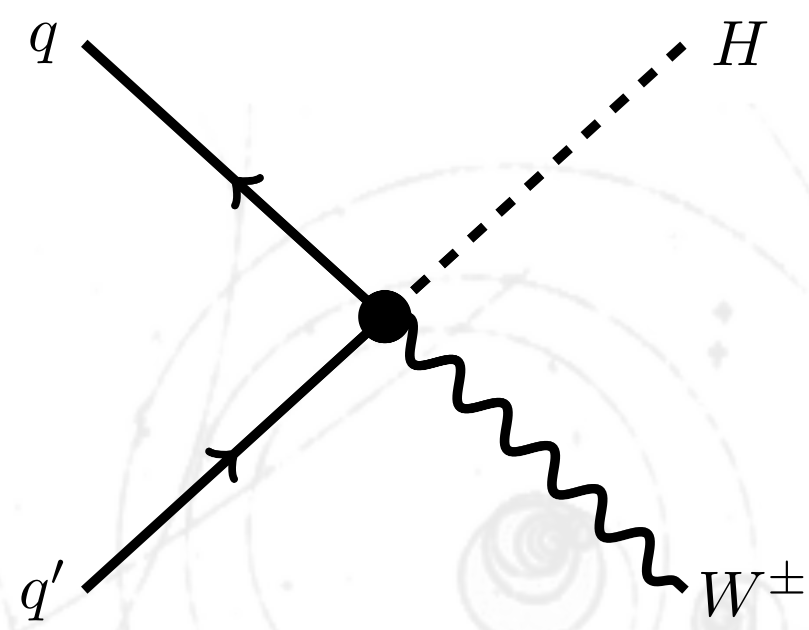
There are four relevant operators*

$$\mathcal{O}_{HD} = |H^\dagger D^\mu H|^2$$
$$\mathcal{O}_{H\Box} = (H^\dagger H)\Box(H^\dagger H)$$

} Finite Higgs wave-function renormalization

$$\mathcal{O}_{HW} = (H^\dagger H)\, W^a_{\mu\nu} W^{a\,\mu\nu}$$
$$\mathcal{O}^{(3)}_{Hq} = (H^\dagger i \overleftrightarrow{D}^a_\mu H)(Q_L \sigma^a \gamma^\mu Q_L)$$





$\mathcal{O}_{HD}$, $\mathcal{O}_{H\Box}$ always enter in the combination

$$\frac{\tilde{C}_{HD}}{\Lambda^2}\tilde{\mathcal{O}}_{HD} \equiv \frac{\tilde{C}_{HD}}{\Lambda^2}\left(\mathcal{O}_{H\Box} - \frac{1}{4}\mathcal{O}_{HD}\right)$$

So we're left with 3 theory parameters:

$$\left\{\tilde{C}_{HD},\, C_{HW},\, C^{(3)}_{Hq}\right\}$$

*neglecting operators modifying the Higgs decay, which are not well measured here

How do we estimate theory parameters?

The object linking observed events to theory parameters is the likelihood function:

$$p(x|\theta) = \int dz_d \int dz_s \int dz_p \, p(x|z_d) \underbrace{p(z_d|z_s)}_{\text{detector smearing}} \overbrace{p(z_s|z_p)}^{\text{shower-splitting}} \overbrace{p(z_p|\theta)}^{\text{parton-level}}$$

$$\left(\sim |\mathcal{M}(z_p|\theta)|^2\right)$$

Doing this integral explicitly is intractable

How do we estimate theory parameters?

The object linking observed events to theory parameters is the likelihood function:

$$p(x|\theta) = \int dz_d \int dz_s \int dz_p \, p(x|z_d) \underbrace{p(z_d|z_s)}_{\text{detector smearing}} \overbrace{p(z_s|z_p)}^{\text{shower-splitting}} \overbrace{p(z_p|\theta)}^{\text{parton-level}}$$

$$\left(\sim |\mathcal{M}(z_p|\theta)|^2\right)$$

Doing this integral explicitly is intractable

But we can make progress by choosing simple observables

Rate: $x = N_{\text{events}}$

Histogram: $x = \{N_{\text{bin}}\}$

(e.g., Simplified Template Cross Sections)

Traditional methods involve throwing out some kinematic information!

How do we estimate theory parameters?

The object linking observed events to theory parameters is the likelihood function:

$$p(x|\theta) = \int dz_d \int dz_s \int dz_p \, p(x|z_d) \underbrace{p(z_d|z_s)}_{\text{detector smearing}} \overbrace{p(z_s|z_p)}^{\text{shower-splitting}} \overbrace{p(z_p|\theta)}^{\text{parton-level}}$$

$$\left(\sim |\mathcal{M}(z_p|\theta)|^2\right)$$

Doing this integral explicitly is intractable

But we can make progress by choosing simple observables

Rate: $x = N_{\text{events}}$

Histogram: $x = \{N_{\text{bin}}\}$

(e.g., Simplified Template Cross Sections, STXS)

Traditional methods involve throwing out some kinematic information!

Close to the Standard Model, we can use *optimal observables*

$$\mathcal{O}_n(x) \sim \frac{\partial}{\partial \theta_n} \log p(x\,|\,\theta)\Big|_{\theta=0} \equiv t_n(x\,|\,\theta = 0)$$

D. Atwood, A. Soni PRD 45, 7 (1992)
M. Diehl, O. Nachtmann, Z. Phys. C 62 (1994)

Information Geometry

J. Brehmer, K. Cranmer, F. Kling, T. Plehn, 1612.05261
J. Brehmer, F. Kling, T. Plehn, T. Tait, 1712.02350

Expected limits can be nicely summarized by the *Fisher Information*:

$$I_{ij}(\theta_0) \equiv -\int dx\, p(x\,|\,\theta_0)\, \frac{\partial \log p(x\,|\,\theta)}{\partial \theta_i} \frac{\partial \log p(x\,|\,\theta)}{\partial \theta_j} = \langle t_i\, t_j\,|\,\theta_0 \rangle$$

Information Geometry

J. Brehmer, K. Cranmer, F. Kling, T. Plehn, 1612.05261
J. Brehmer, F. Kling, T. Plehn, T. Tait, 1712.02350

Expected limits can be nicely summarized by the *Fisher Information*:

$$I_{ij}(\theta_0) \equiv - \int dx\, p(x\,|\,\theta_0)\, \frac{\partial \log p(x\,|\,\theta)}{\partial \theta_i} \frac{\partial \log p(x\,|\,\theta)}{\partial \theta_j} = \left\langle t_i\, t_j \,|\, \theta_0 \right\rangle$$

which has a number of useful properties:

- Expected limits bounded by the Inverse Fisher Information (Cramèr-Rao)
- Transforms covariantly under parameter transformations (EFT bases)
- Additive for different processes & phase space regions
- Can include systematic uncertainties as nuisance parameters

Information Geometry

J. Brehmer, K. Cranmer, F. Kling, T. Plehn, 1612.05261
J. Brehmer, F. Kling, T. Plehn, T. Tait, 1712.02350

Expected limits can be nicely summarized by the *Fisher Information*:

$$I_{ij}(\theta_0) \equiv -\int dx\, p(x\,|\,\theta_0)\, \frac{\partial \log p(x\,|\,\theta)}{\partial \theta_i} \frac{\partial \log p(x\,|\,\theta)}{\partial \theta_j} = \langle t_i\, t_j \,|\, \theta_0 \rangle$$

which has a number of useful properties:

- Expected limits bounded by the Inverse Fisher Information (Cramèr-Rao)
- Transforms covariantly under parameter transformations (EFT bases)
- Additive for different processes & phase space regions
- Can include systematic uncertainties as nuisance parameters

$$\begin{matrix} C_{H\Box} & C_{HD} & C_{HW} & C_{Hq}^{(3)} \end{matrix}$$

$$I_{ij}^{\text{Rate}} = \begin{pmatrix} 28.0 & -7.0 & 101 & 272 \\ -7.0 & 1.75 & -25.2 & -68 \\ 101 & -25.2 & 360.1 & 979 \\ 272 & -68 & 979 & 2660 \end{pmatrix}$$

Degeneracy between rate only operators is apparent

Information Geometry

J. Brehmer, K. Cranmer, F. Kling, T. Plehn, 1612.05261
J. Brehmer, F. Kling, T. Plehn, T. Tait, 1712.02350

Expected limits can be nicely summarized by the *Fisher Information*:

$$I_{ij}(\theta_0) \equiv -\int dx\, p(x|\theta_0) \frac{\partial \log p(x|\theta)}{\partial \theta_i} \frac{\partial \log p(x|\theta)}{\partial \theta_j} = \langle t_i t_j | \theta_0 \rangle$$

which has a number of useful properties:

- Expected limits bounded by the Inverse Fisher Information (Cramèr-Rao)
- Transforms covariantly under parameter transformations (EFT bases)
- Additive for different processes & phase space regions
- Can include systematic uncertainties as nuisance parameters

$$\begin{matrix} & C_{H\Box} & C_{HD} & C_{HW} & C_{Hq}^{(3)} \end{matrix}$$

$$I_{ij}^{\mathrm{Rate}} = \begin{pmatrix} 28.0 & -7.0 & 101 & 272 \\ -7.0 & 1.75 & -25.2 & -68 \\ 101 & -25.2 & 360.1 & 979 \\ 272 & -68 & 979 & 2660 \end{pmatrix}$$

Degeneracy between rate only operators is apparent

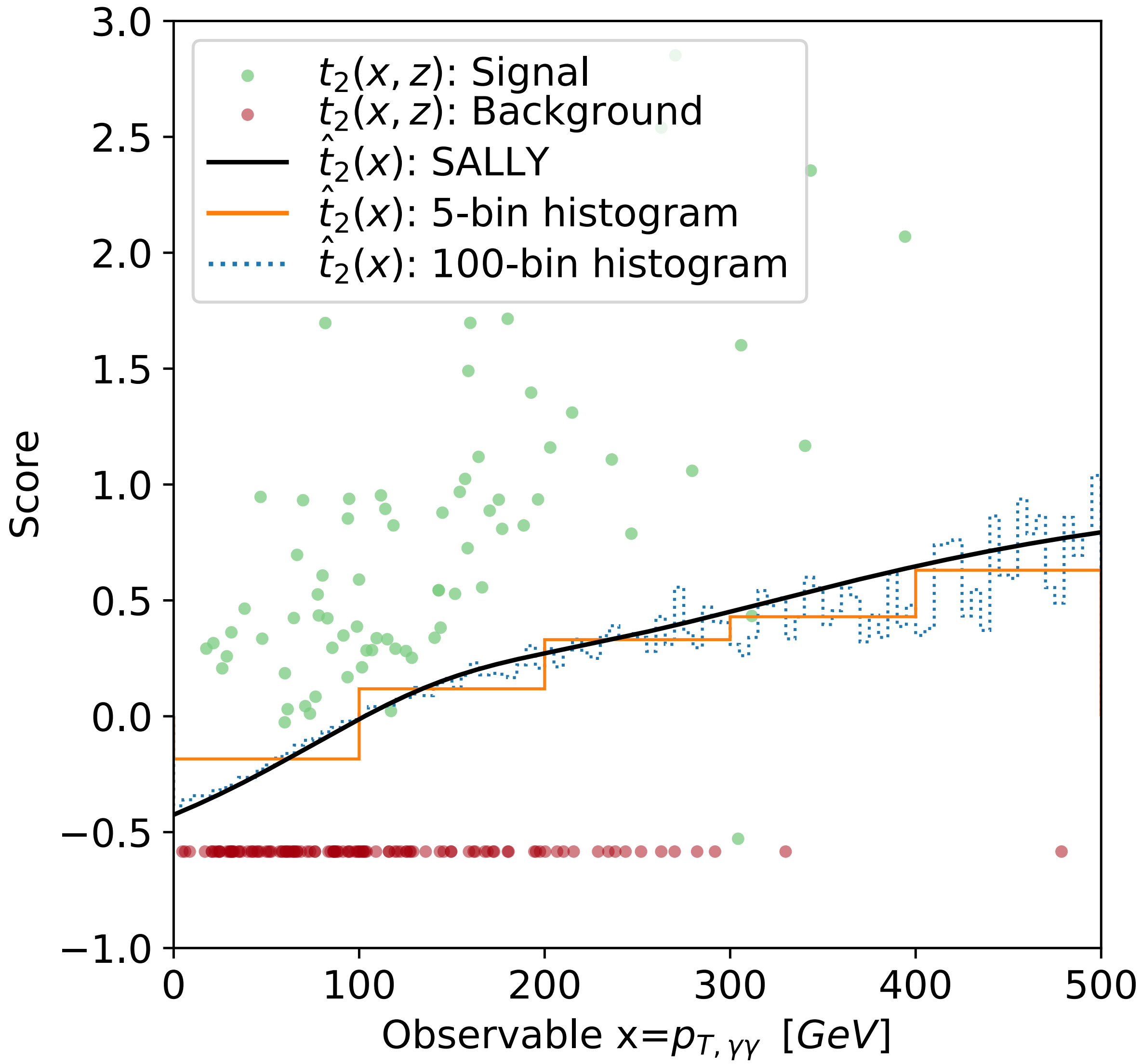
Moreover, the Fisher Information can be reliably estimated *in the presence of detector effects* using Machine Learning techniques

J. Brehmer, K. Cranmer, G. Louppe, J. Pavez 1805.00013, 1805.00020,1805.12244
J. Brehmer, F. Kling, I. Espejo, K. Cranmer 1907.10621

MadMiner

An Inference Toolkit for Particle Physics
J. Brehmer, F. Kling, I. Espejo, K. Cranmer, 1907.10621

- Combines the power of ML inference methods with intuition of the Matrix Element Method
- Automizes score & likelihood ratio estimation techniques, fully interfaced with simulation tools
- Out of the box: pheno-level analysis
  - MadGraph, Pythia, Delphes
  - Backgrounds
  - PDF/scale uncertainties
  - ML uncertainties
- Scalable to state-of-the-art experimental tools
- Python package
  - Modular interface
  - Extensive documentation
  - On GitHub: github.com/diana-hep/madminer
  - Easy to install: `pip install madminer`



MadMiner Workflow

Setup Morphing Basis
(choose theory parameters)

$$\left\{ \tilde{C}_{HD},\ C_{HW},\ C_{Hq}^{(3)} \right\}$$

MadMiner Workflow

Setup Morphing Basis

(choose theory parameters)

$$\left\{ \tilde{C}_{HD},\ C_{HW},\ C_{Hq}^{(3)} \right\}$$

→

Simulate Signal & Backgrounds

WH in SMEFT (Using Reweighting)

$Wb\bar{b}$, $t\bar{t}$, and tb backgrounds

MadMiner Workflow

Setup Morphing Basis
(choose theory parameters)
$$\left\{ \tilde{C}_{HD},\ C_{HW},\ C_{Hq}^{(3)} \right\}$$

→

Simulate Signal & Backgrounds

WH in SMEFT (Using Reweighting)

$Wb\bar{b}$, $t\bar{t}$, and tb backgrounds

→

Choose Observables:

- Full set of four momenta,
- Only physical momenta,
- a subset of interesting variables…

MadMiner Workflow

Setup Morphing Basis
(choose theory parameters)
$$\left\{ \tilde{C}_{HD},\ C_{HW},\ C_{Hq}^{(3)} \right\}$$

Simulate Signal & Backgrounds

WH in SMEFT (Using Reweighting)

$Wb\bar{b}$, $t\bar{t}$, and tb backgrounds

Choose Observables:

- Full set of four momenta,
- Only physical momenta,
- a subset of interesting variables…

Train SALLY Estimator
Score Approximates Likelihood LocallY

MadMiner Workflow

Setup Morphing Basis
(choose theory parameters)
$$\left\{ \tilde{C}_{HD},\ C_{HW},\ C_{Hq}^{(3)} \right\}$$

Simulate Signal & Backgrounds

WH in SMEFT (Using Reweighting)

$Wb\bar{b}$, $t\bar{t}$, and tb backgrounds

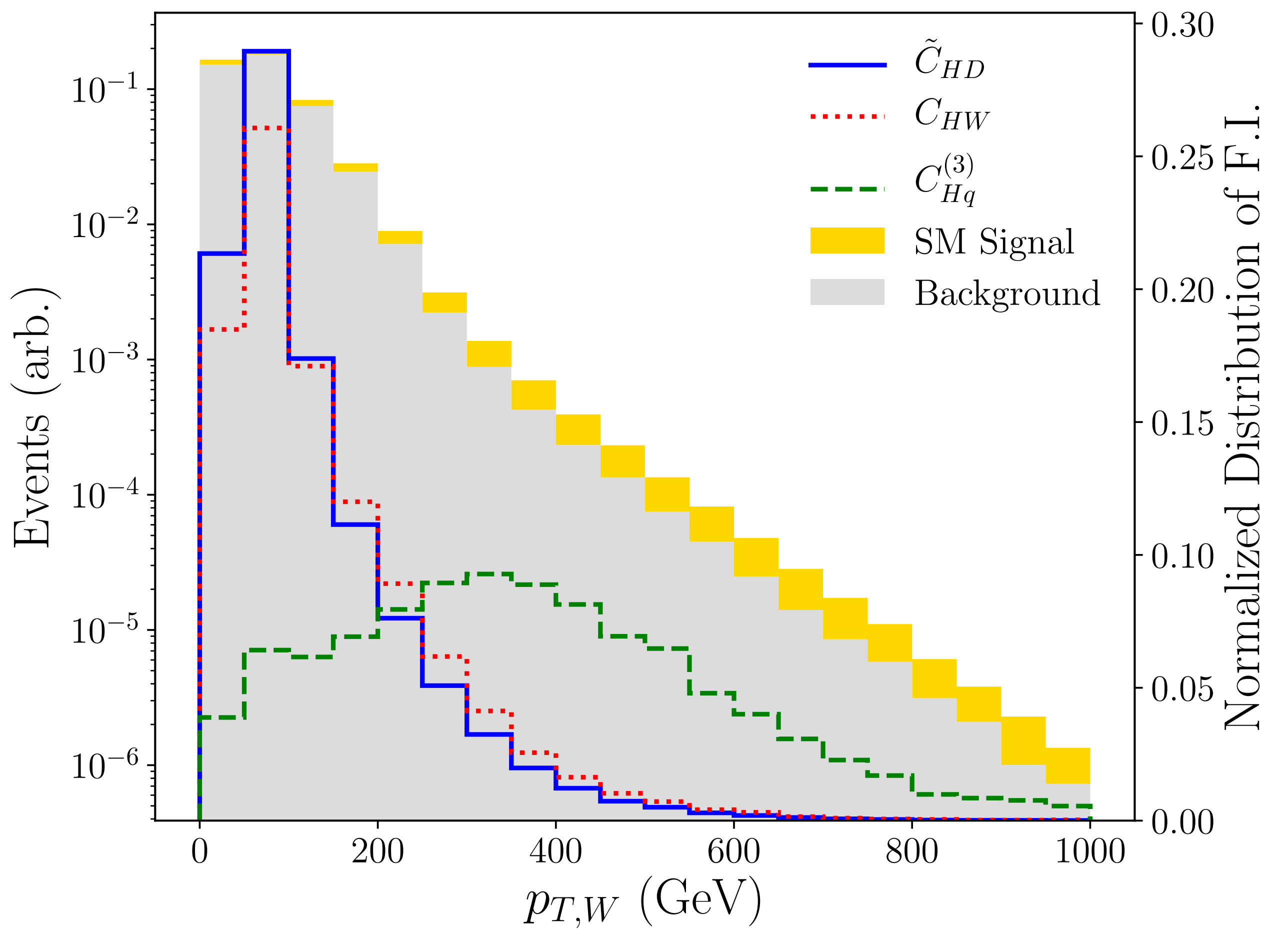
Choose Observables:

- Full set of four momenta,
- Only physical momenta,
- a subset of interesting variables…

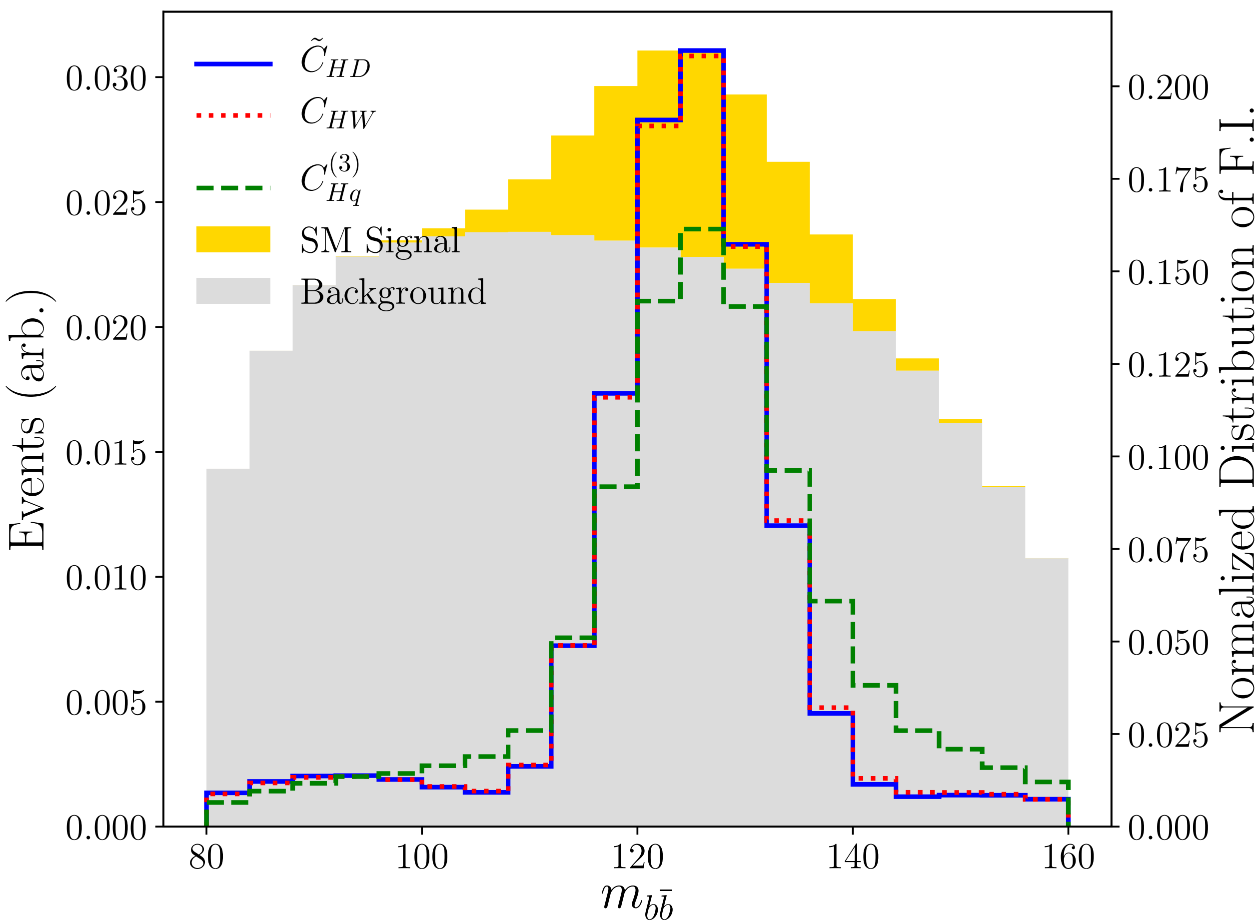
Train SALLY Estimator
Score Approximates Likelihood LocallY

Evaluate Fisher Information
Calculate Expected Limits

Where is the Information?



Momentum enhanced operators have Fisher Information peaked in high momentum bins



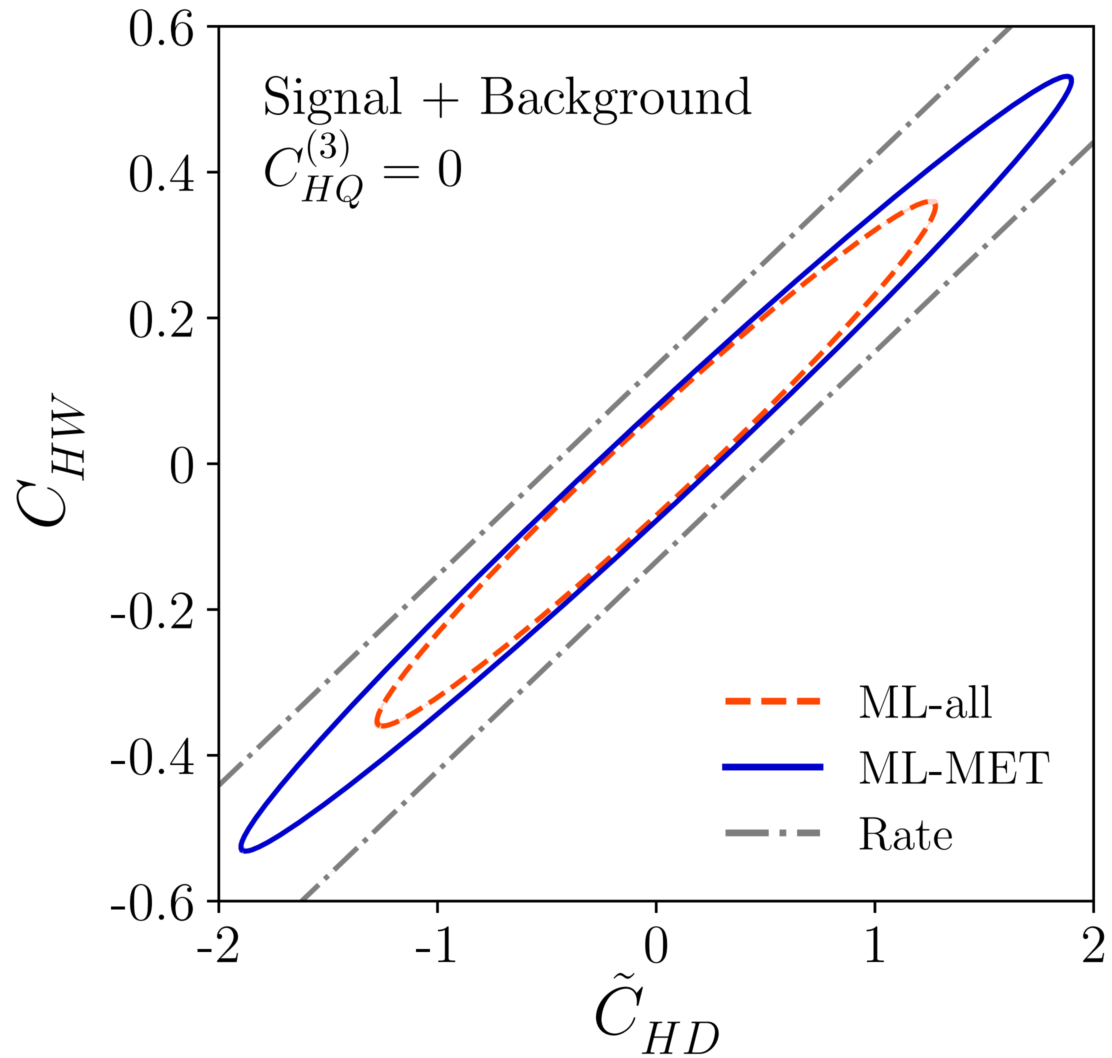
$$\mathcal{M}^0 \sim 1 + \frac{sv^2}{M_W^2 \Lambda^2} C_{Hq}^{(3)}$$

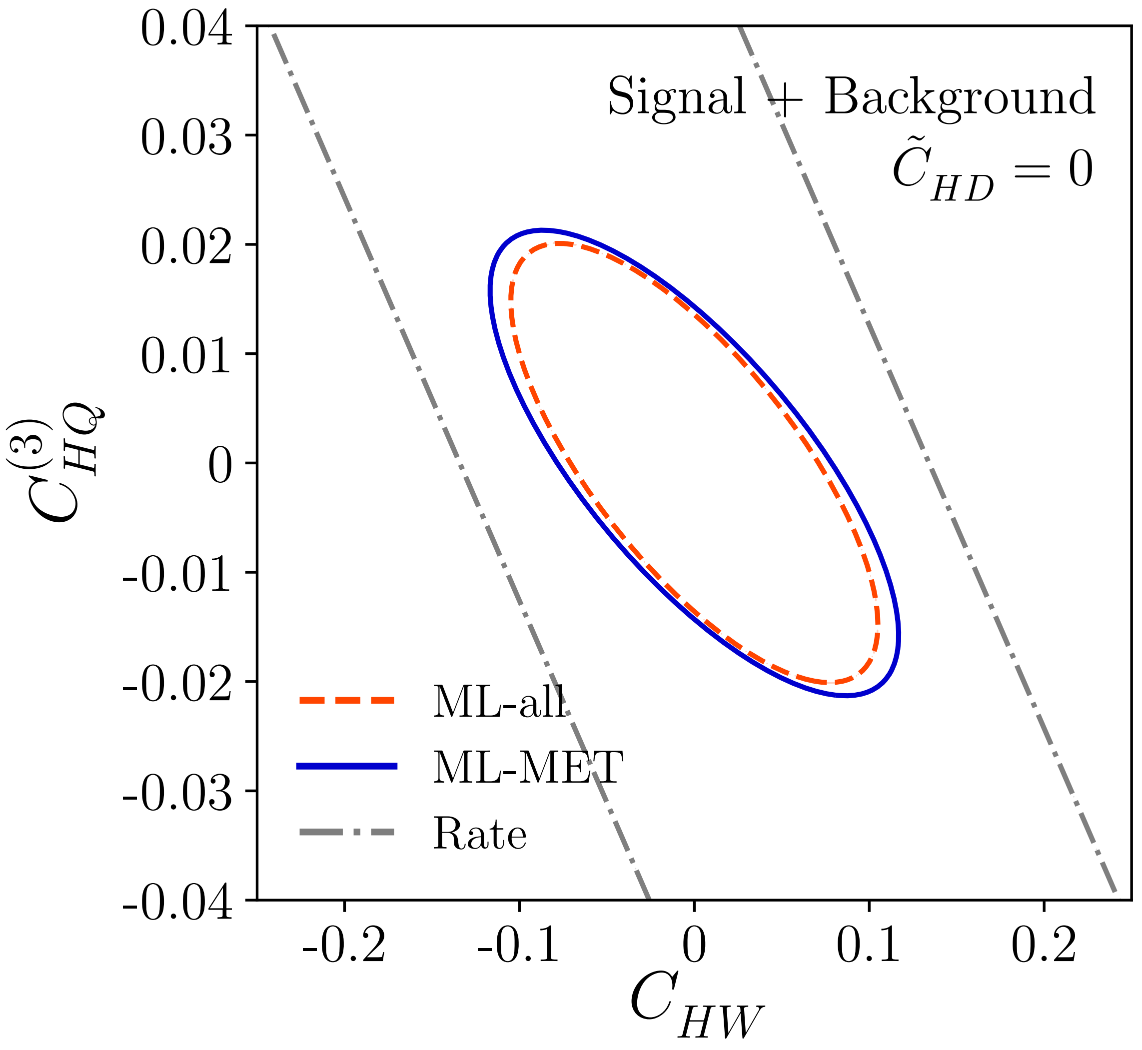
$$\mathcal{M}^{\pm} \sim \frac{M_W}{\sqrt{s}} \left(1 + \frac{sv^2}{M_W^2 \Lambda^2} \left(C_{HW} + C_{Hq}^{(3)} \right) \right)$$

Including backgrounds can change naive distribution of information

What is lost in the Neutrino Momentum?

We can train our estimator on different sets of observables:
with and without the neutrino energy & longitudinal momentum





This allows us to quantify how much information is lost in missing momentum!

How Many Bins are Necessary?

We can compute the Fisher Info for a histogram of $p_{T,W}$ starting with bins at 0, 150, and 250 GeV

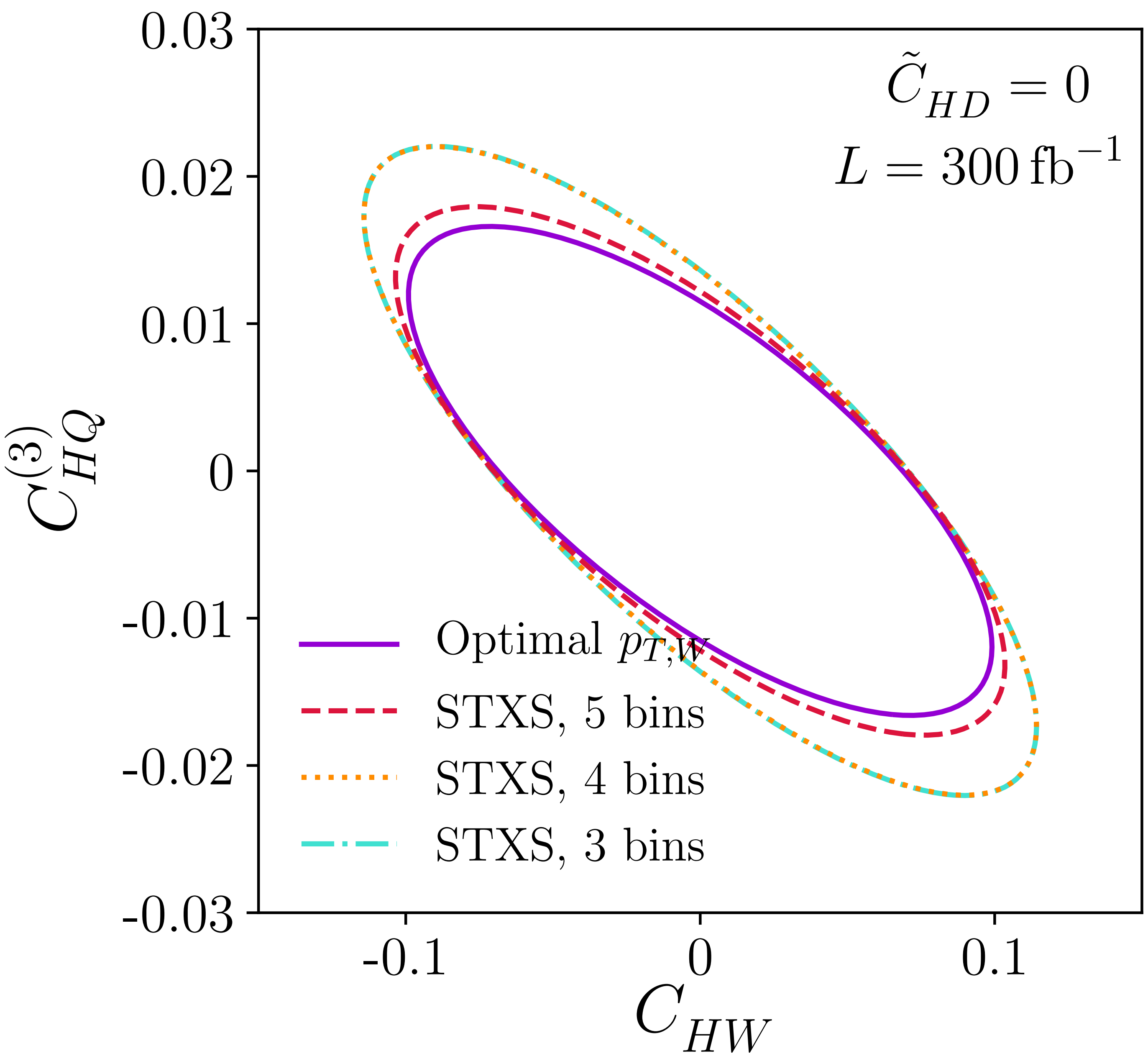
STXS Stage 1 [1610.07922]

Include the additional bins at 75 GeV (4 bins) and 400 GeV (5 bins)

STXS Stage 1.1 [1906.02754]

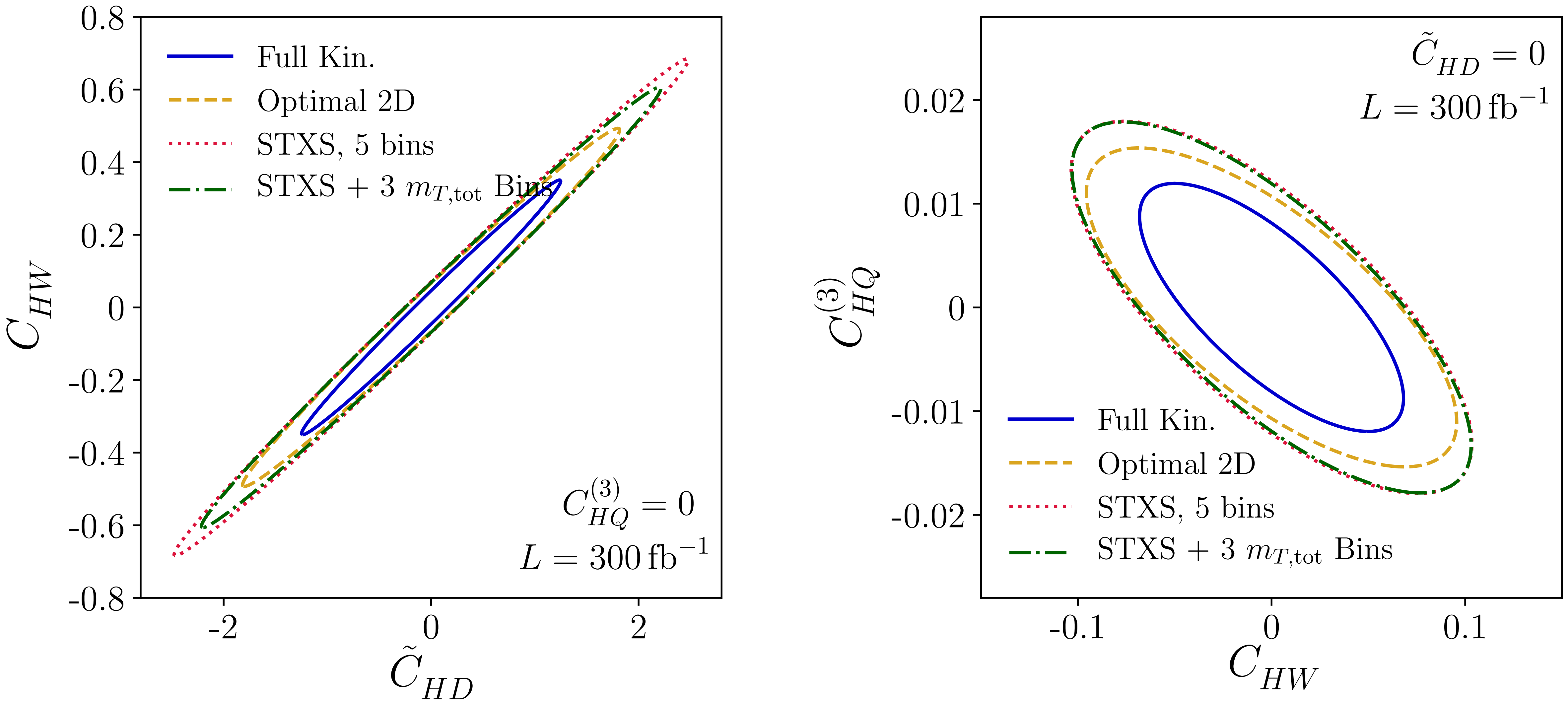
Compare to SALLY Estimator trained on only $p_{T,W}$

→ Additional high momentum bin essential for constraining $C^{(3)}_{Hq}$



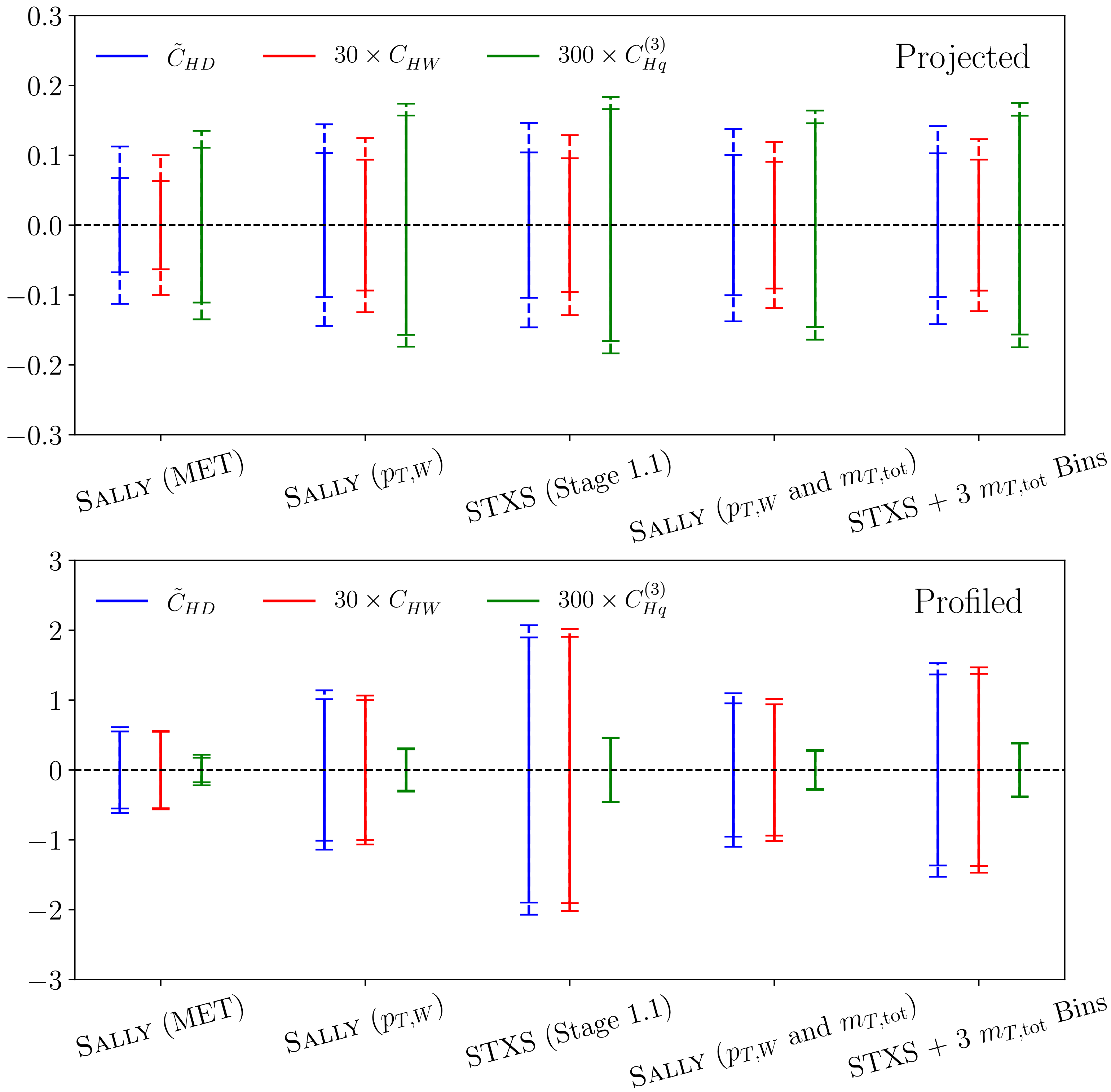
Is One Observable Enough?

Compare STXS to Estimator trained on 2 Observables & 2D Histogram



Additional observables help discriminate against background

Conclusions

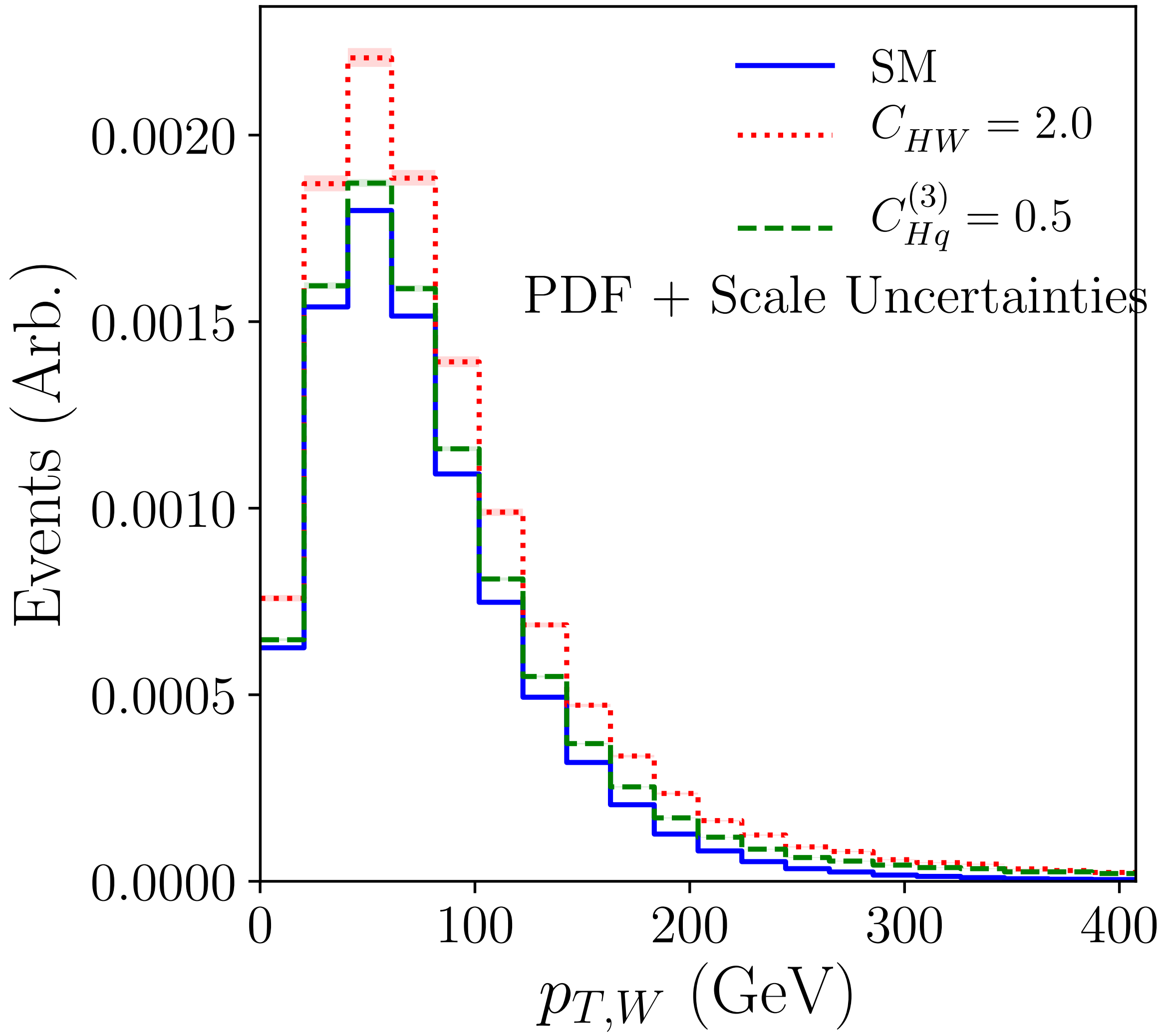


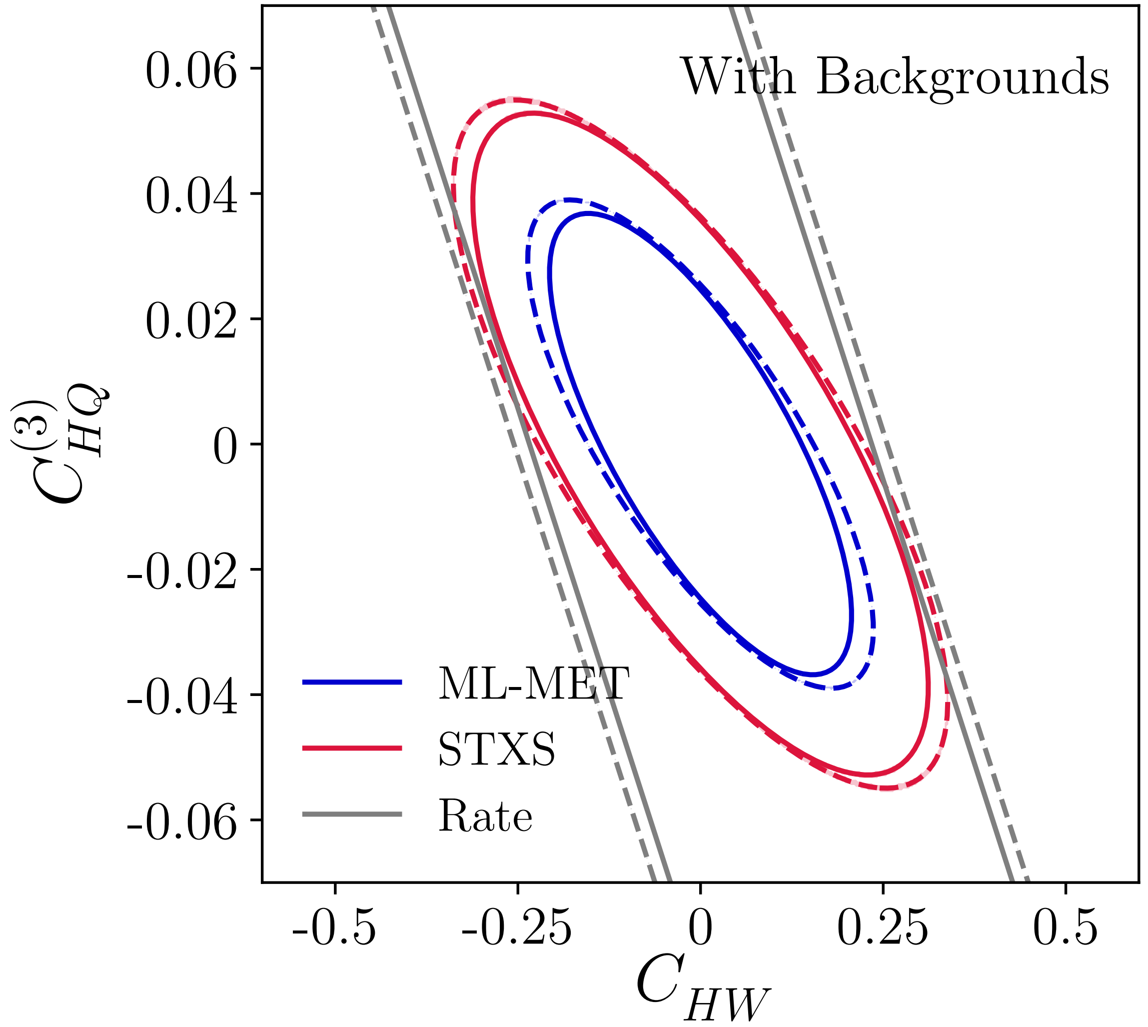
- Our analysis captures effects at order $1/\Lambda^2$
- Momentum-enhanced 4-point interactions require high momentum bins to constrain
- Can be further constrained with additional observables
- These effects are well understood in the context of the Fisher Information with the help of MadMiner

Backup

Including Systematics

Scale & PDF Uncertainties can be treated as additional (nuisance) parameters





We include these for the WH Signal and marginalize over them in the Fisher Information.