Equiprobability Anomaly Detector Using Angular Distributions

[2610.23667]



Alex Whitten, John Cicero, Tae Min Hong*

PIKIMO 2026

indico.global/event/19193/contributions/172574

Outline

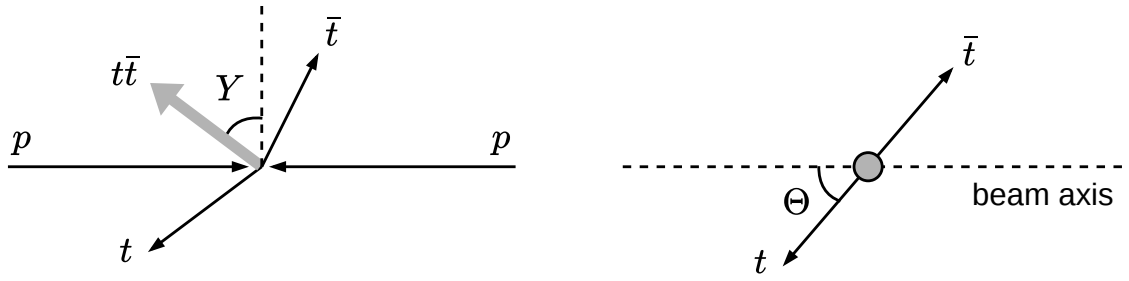
T. M. Hong
Pittsburgh

1	**10 angles for $t\bar{t}$ decays**	Wigner D functions expand the $t\bar{t}$ density
	■ **Orthogonal series density estimation**	Technique to estimate the coefficients
2	**Equiprobable surfaces**	Bins of equal $\lvert M \rvert^2$
	■ **Toy: 1d, 2d**	Distribution of probability density
	■ **Discussion**	Link to information theory
3	**Likelihood ratio**	Parton level against a normalizing flow
	■ **Limitations**	W-mass constraint & jet assignment

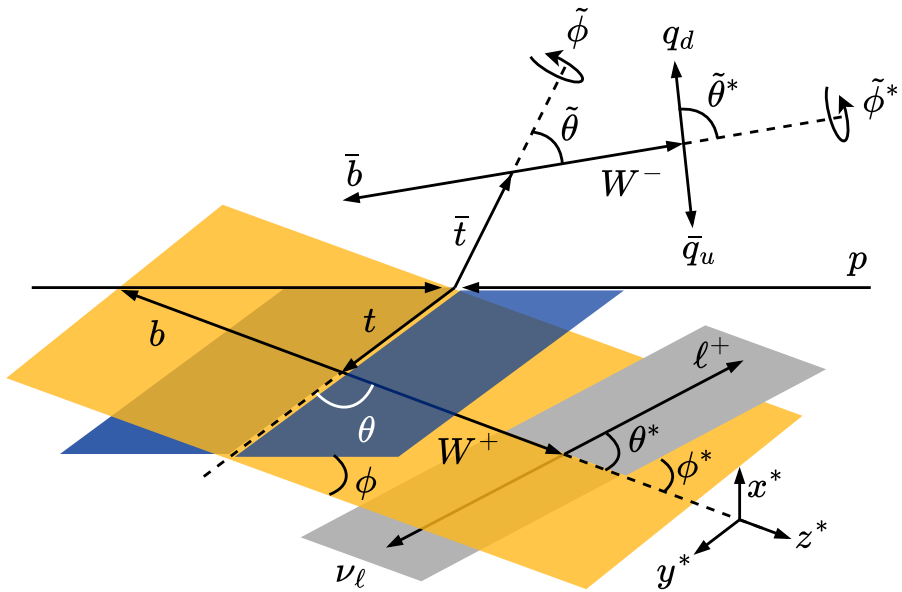
Distribution of probability density: $t\bar{t}$ decays

T. M. Hong
Pittsburgh

(a) Rapidity Y of $t\bar{t}$ in the lab frame (b) Polar angle Θ in the $t\bar{t}$ COM frame



(c) Angles defined in the mother's COM with respect to her direction



Ten angles per decay

- **Orthonormal basis**
  - Wigner D for each decay

Angular momentum conservation leaves 24 terms

$$\underbrace{\frac{1}{\Gamma}\frac{\partial\Gamma}{\partial\Omega\,\partial\Omega'}(\boldsymbol{\theta})}_{f_t(\boldsymbol{\theta})} = \sum_{j_1 j_2 m' m}^{j_1\le 1,\, j_2\le 2} c_{m'm}^{j_1 j_2} \cdot M_{m'm}^{j_1 j_2}(\boldsymbol{\theta})$$

  - Legendre in $\cos\Theta$
  - Hermite in Y

$$f_{t\bar{t}}(\mathbf{x}) = \sum_{j\tilde{\jmath} ln} c_{j\tilde{\jmath} ln} \cdot \underbrace{M_j(\boldsymbol{\theta})\, M_{\tilde{\jmath}}(\tilde{\boldsymbol{\theta}})\, P_l(\cos\Theta)\, H_n(Y)}_{B_{j\tilde{\jmath} ln}(\mathbf{x})},$$

Number of terms

- $24^2 \cdot 5^2 = 14400$ coefficients

Extract coefficients?
Use **orthogonal series density estimation**

T. M. Hong
Pittsburgh

1 Expand orthonormal basis

$$f(x) = \sum_n c_n \varphi_n(x)$$

2 Extract coefficients

$$c_n = \int f(x)\, \varphi_n(x)\, \mathrm{d}x \qquad \textit{exact}$$

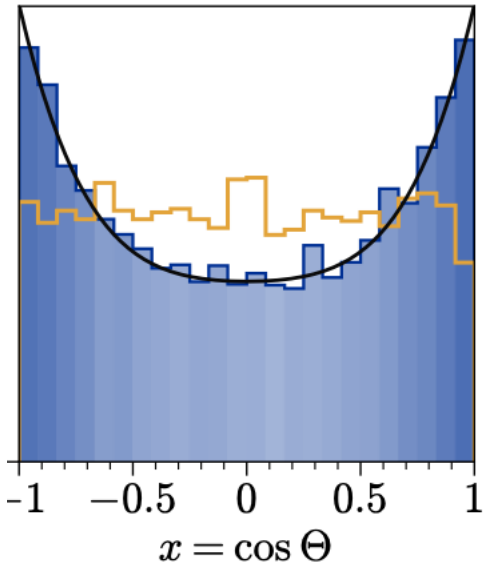
OSDE

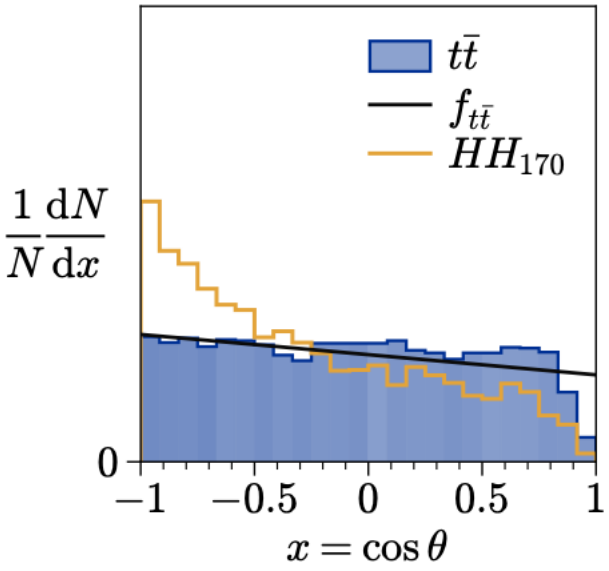
$$\hat{c}_n = \frac{1}{N} \sum_{i=1}^{N} \varphi_n(x_i) \qquad \textit{sample mean}$$

3 Keep significant terms

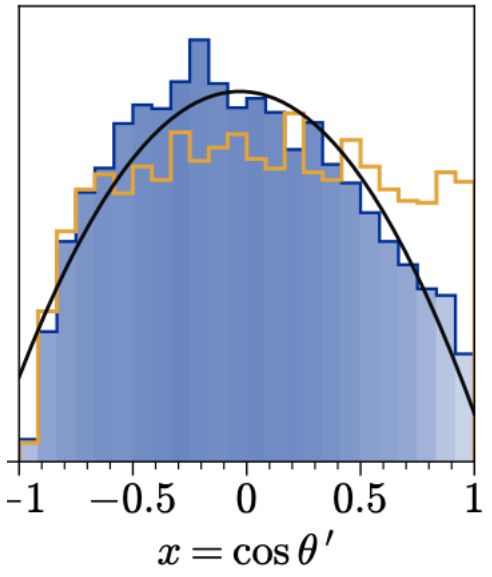
$$\left|\hat{c}_n\right| > 2.5\sigma_n, \quad \sigma_n = \frac{\mathrm{std}(\varphi_n)}{\sqrt{N}} \qquad \textit{noise cut}$$

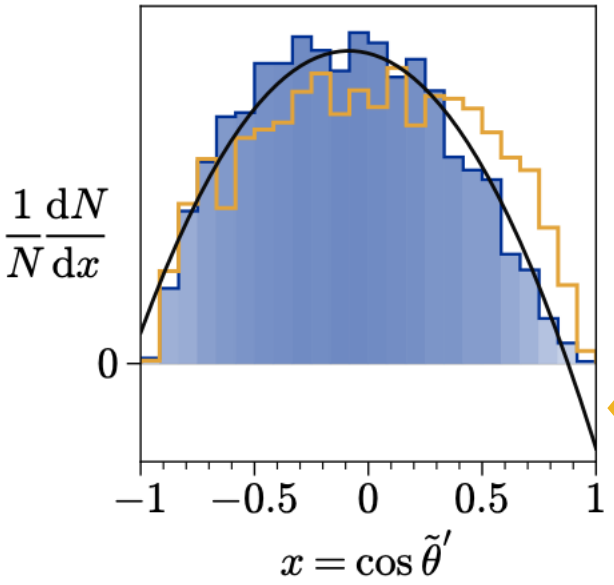
- 14400 coefficients in total
- ~300 kept via cross-validation





fit





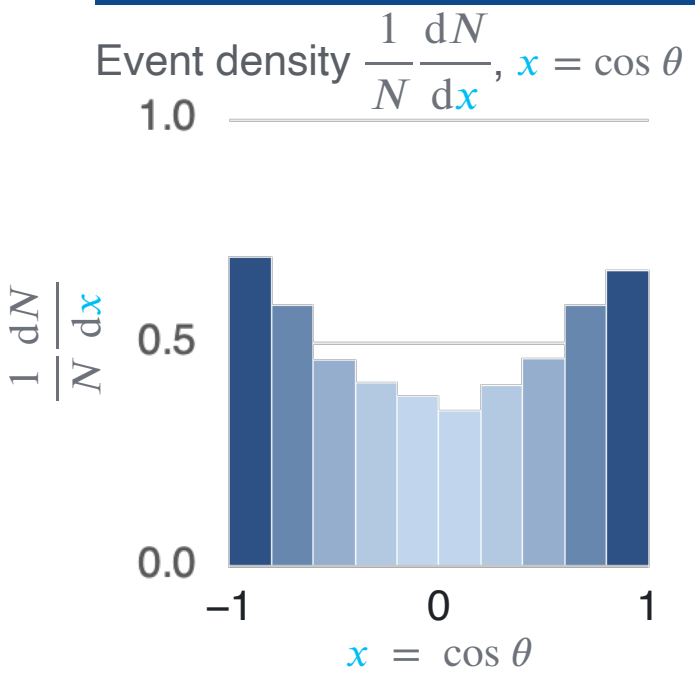
negative probability

Gibbs overshoot

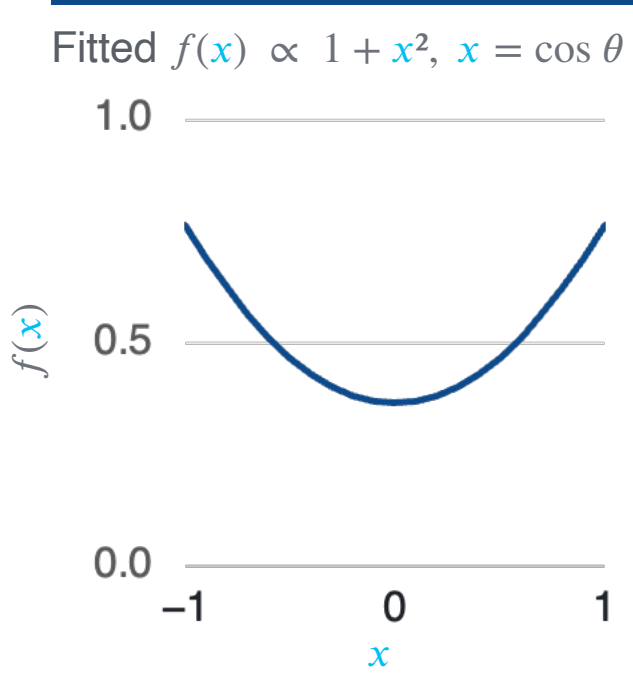
Distribution of probability density: **1d toy**

T. M. Hong
Pittsburgh

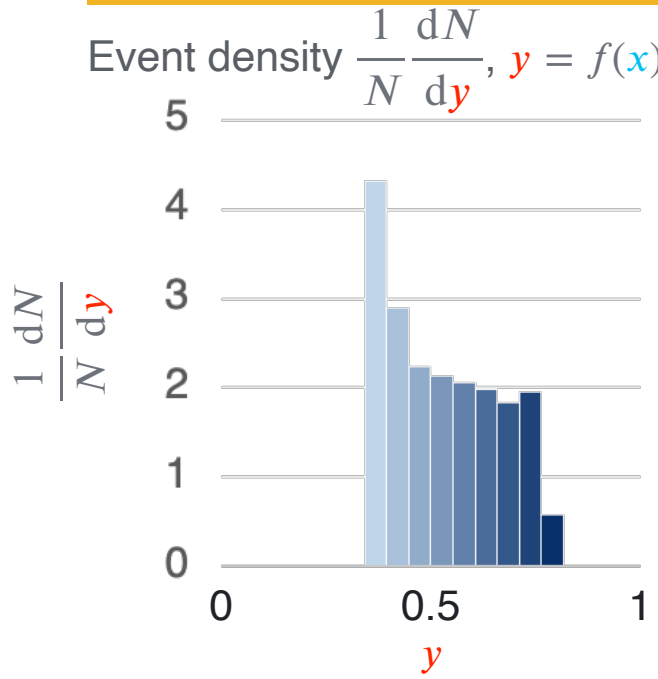
1. Distribution of x



2. Fit the pdf



3. Distribution of y



- Definition: x are angles, y the probability density
- $p = \frac{1}{N}\frac{dN}{dy}$ is the **ordered** probability density of the probability density in x
- The value $y = f(x)$ can be computed per event characterized by x

Distribution of probability density: **2d toy**

T. M. Hong
Pittsburgh

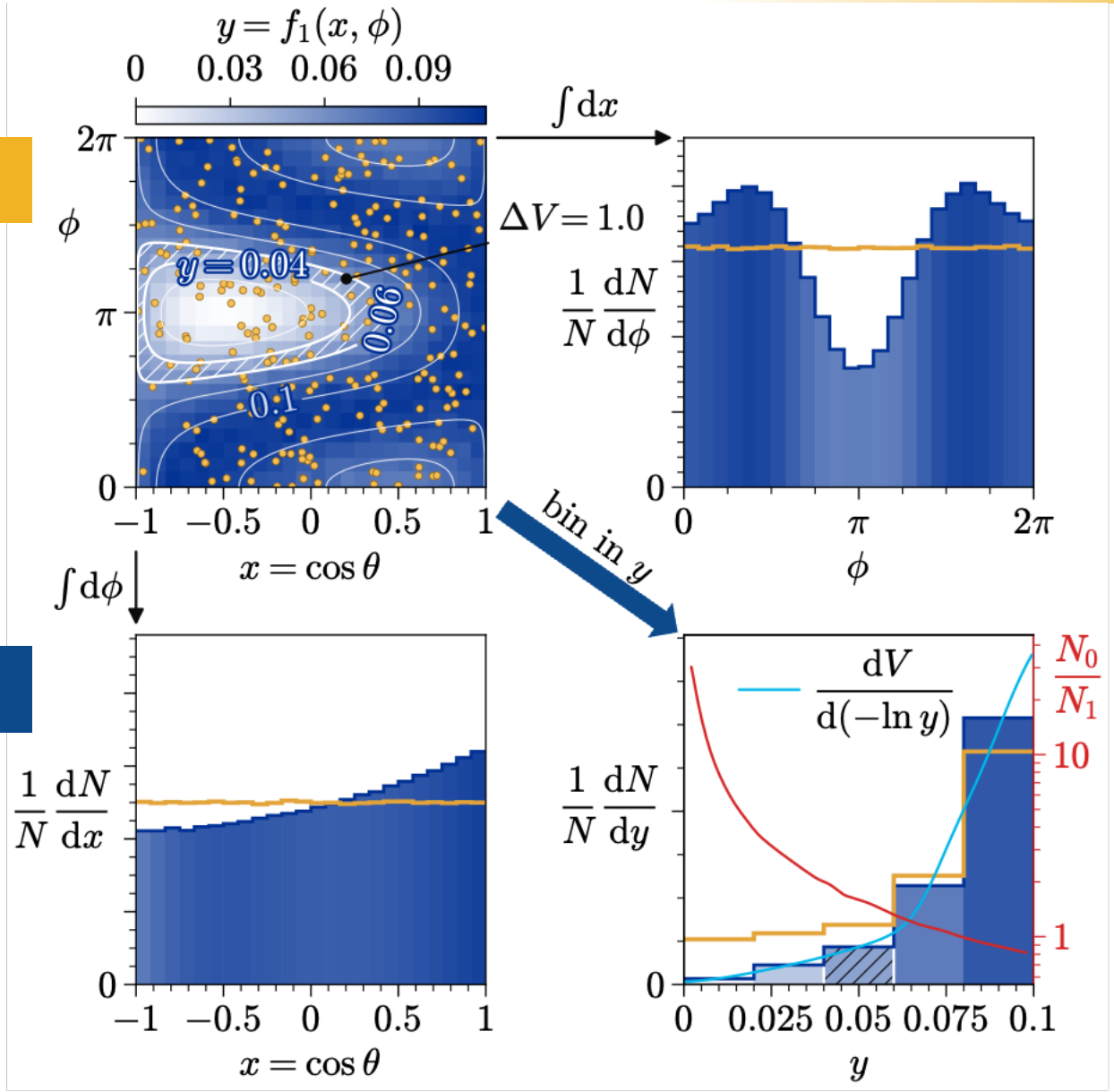
1. Distribution of x

Event density $\frac{1}{N}\frac{\mathrm{d}2N}{\mathrm{d}x\,\mathrm{d}\varphi}$

- **Blue** is like W decay and $f_1(x)$ describes it

2. Projection

- **Yellow** dots represent uniform decay of spin-0



2. Projection

- **Yellow** dots represent uniform decay of spin-0

3. Distribution of y

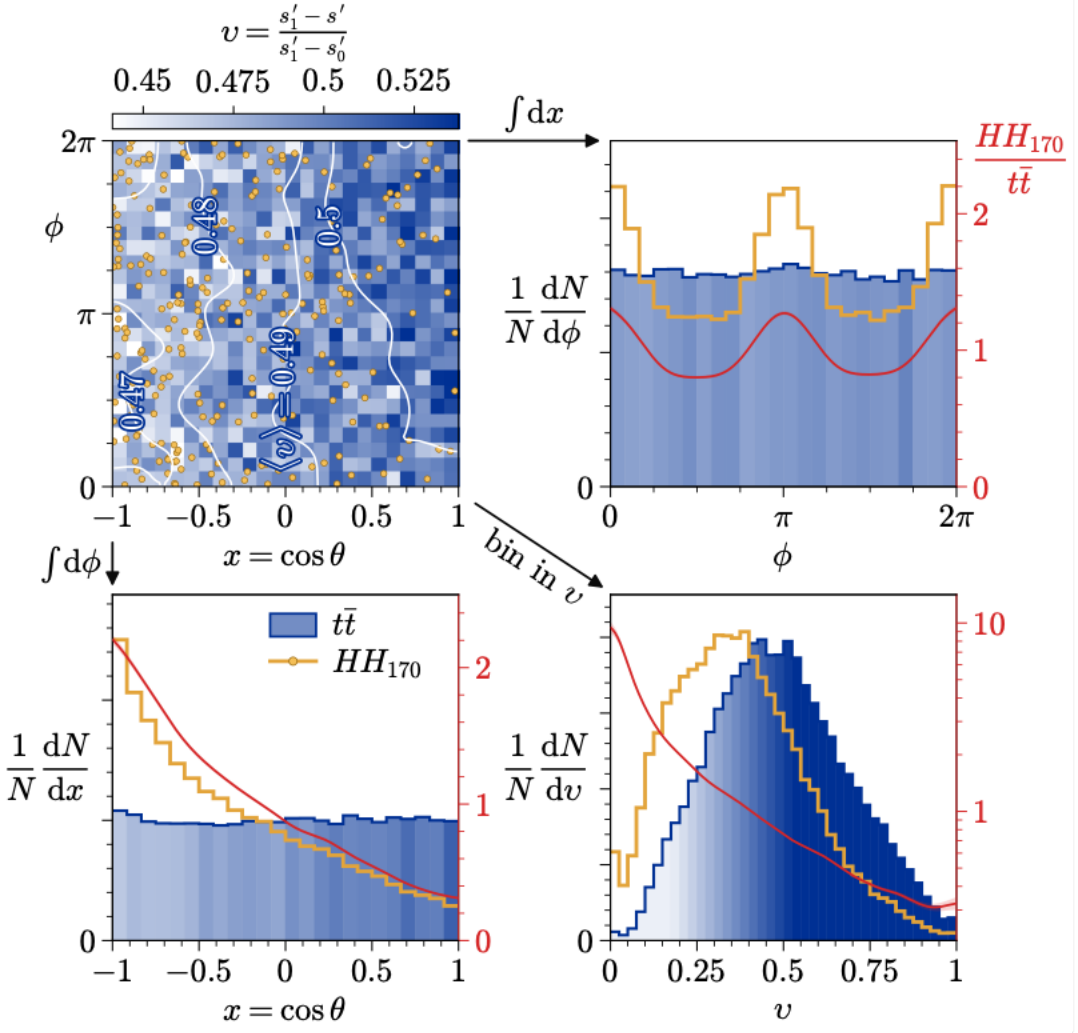
$p(y) = \frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}y}$ where $y = f(x)$

- **Red** curve = $\frac{\textbf{Yellow}}{\textbf{Blue}}$ increases as $y \to 0$

Distribution of probability density: **10d of $t\bar{t}$**

T. M. Hong
Pittsburgh

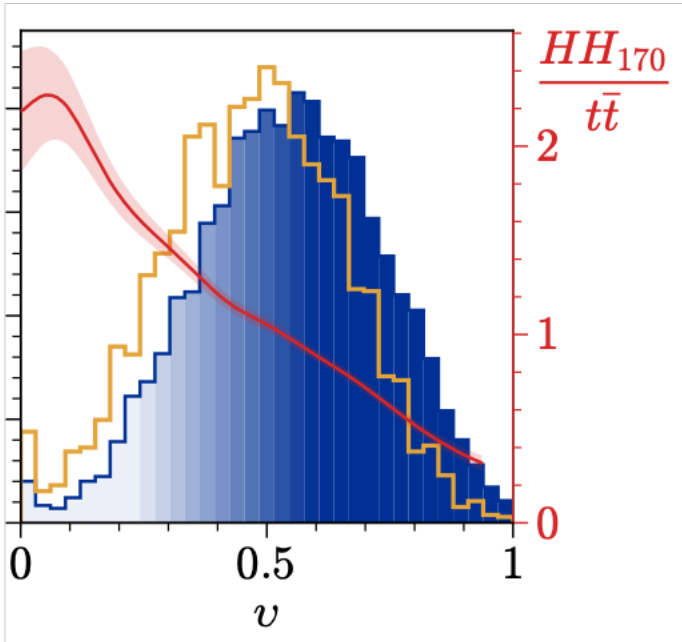
Parton level

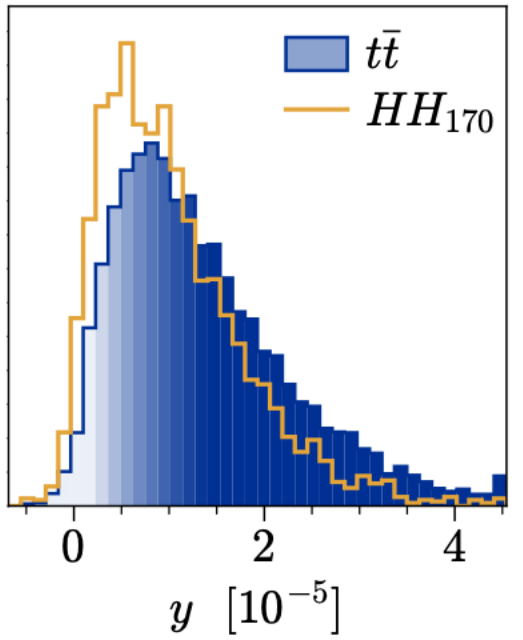


Reco matched to truth

Results

- Fit is for $t\bar{t}$ angles
- Fit does not know about **HH**, an ‘anomaly’





Discussion: **Equiprobable surfaces**

T. M. Hong
Pittsburgh

1 Physics

$$y \propto |M|^2$$

- Score is the matrix element method
- Contours are equal-$|M|^2$ surfaces

2 Geometric connection

$$p(y) = \frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}y} = \frac{\mathrm{d}V}{\mathrm{d}(-\ln y)}$$

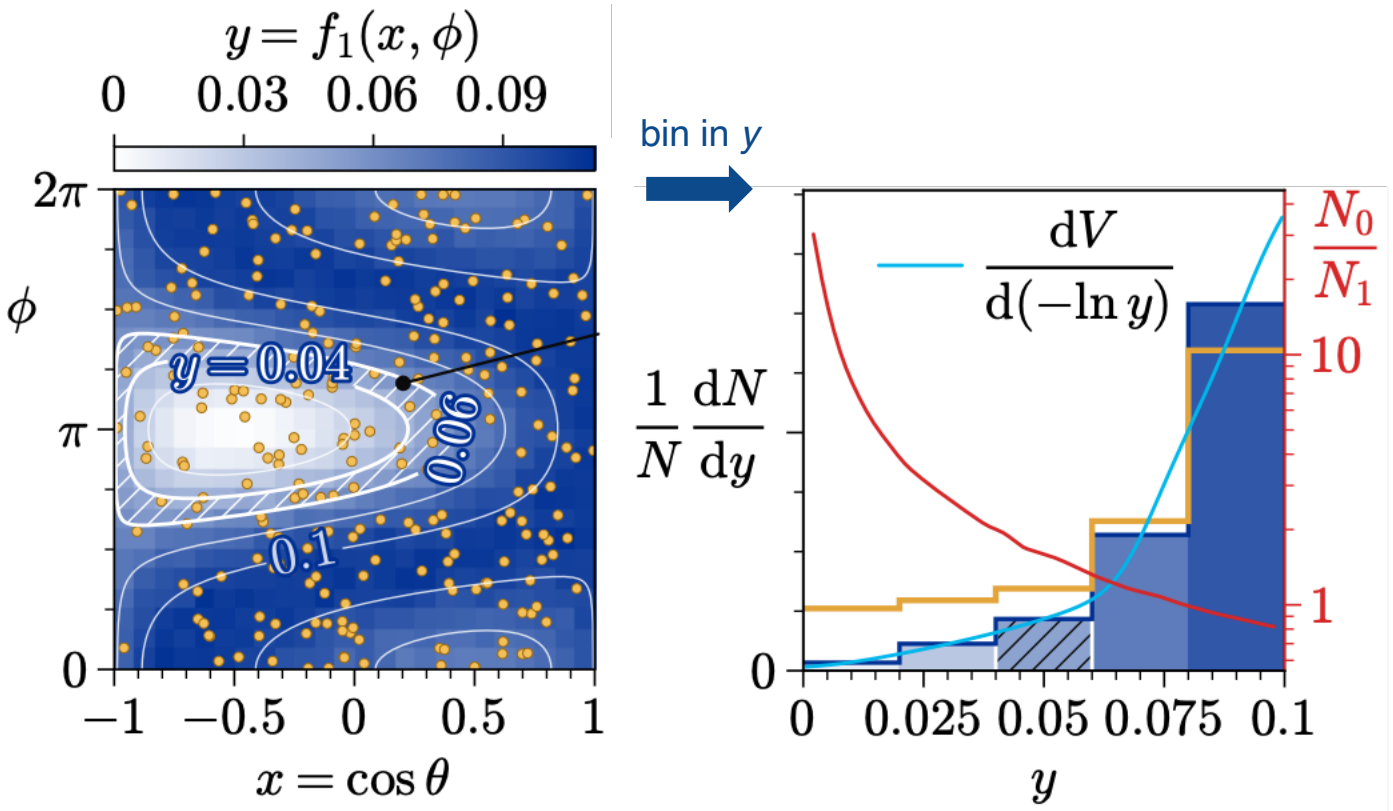
- $V(y)$ is the phase-space volume $f > y$
- Boundary is the equiprobable surface

3 Information connection

$$D_{\mathrm{KL}} = \langle s_i \rangle_j - \langle s_j \rangle_j \geq 0$$

- Surprisal $s = -\ln y$, with $\langle s_i \rangle_i$ as entropy
- D_{KL} is the KL divergence in info theory

2d toy to illustrate



Likelihood ratio of an anomaly

T. M. Hong
Pittsburgh

Of an anomaly that we don't know of?

- We extracted coefficients of $t\bar{t}$ and treated others as an anomaly. But given a particular anomaly, like HH, a likelihood ratio can be formed:

$$L_{HH}(y) = \frac{p_{HH}(y)}{p_{t\bar{t}}(y)}$$

unknown

known

Likelihood ratio

1 L_{HH} is the optimal discriminant

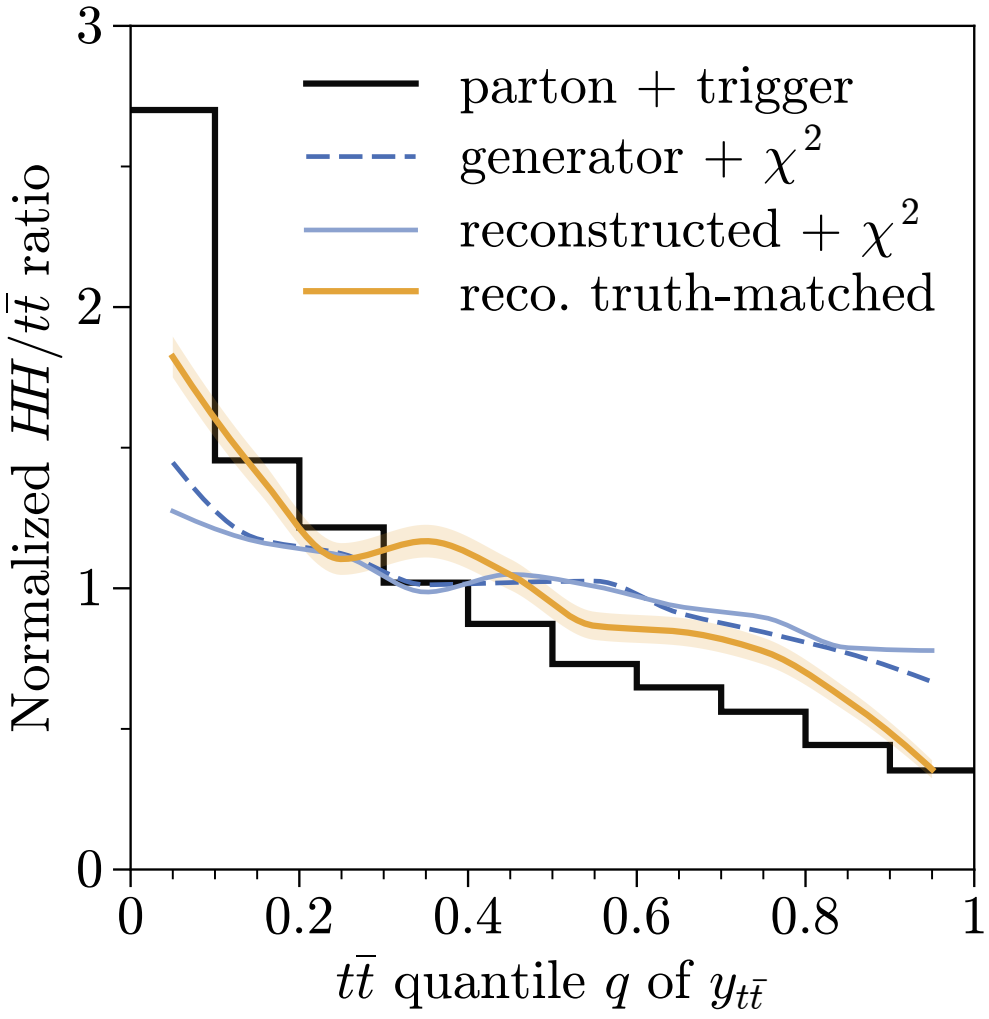
- via Neyman-Pearson lemma

2 Each y bin is an equiprobable surface

- L_{HH} is constant on it

3 $D_{KL} = \langle \ln L_{HH} \rangle$ is the information

- *HH*-weighted average of the log of the plot



Limitations

T. M. Hong
Pittsburgh

AUC comparison vs other ML

Method	Parton	Reco χ^2
Unsupervised · trained on $t\bar{t}$ only		
Equiprobability (this work)	0.67	0.54
Normalizing flow	0.66	—
Autoencoder	0.57	—
Supervised · trained with HH labels		
BDT		
MLP	0.92	0.69
FT-Transformer		

Reco-level issues

1 hh_{125} - W-mass constraint

- vs. hh_{125} fitting off-shell W* forces it to look like on-shell W
- Need a different neutrino determination

2 HH_{170} - Jet assignment

- vs. HH_{170} no longer has the above W* issue, but need to match jets
- But χ^2 based on jet combination forces H_{170} to look like top
- Need a different jet assignment metric

Conclusions

T. M. Hong
Pittsburgh

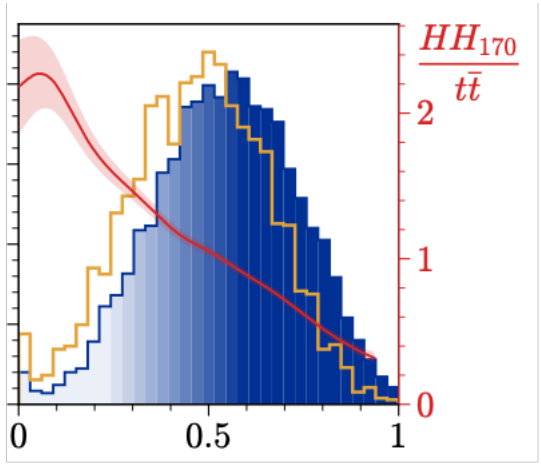
1 Fit the full 10d $t\bar{t}$ density

$$\sum_{j_1 j_2 m' m}^{j_1 \le 1,\, j_2 \le 2} c_{m'm}^{j_1 j_2} \cdot M_{m'm}^{j_1 j_2}(\boldsymbol{\theta})$$

- Wigner D basis: 14400 coefficients to describe all ten angles
- Extract via orthogonal series density estimation

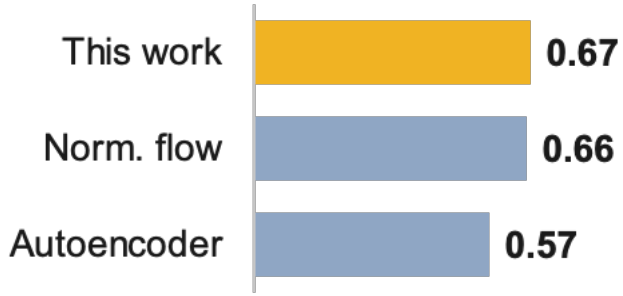
2 Score events on equiprobable surfaces

- Fitted density **y = f(x)** ranks every event
- Its bins are surfaces of equal $|M|^2$, linked to KL divergence



3 Matches ML without labels

- Trained on $t\bar{t}$ only: AUC 0.67 at parton level vs. other ML
- Reco level (0.54) needs better neutrino & jet assignment



Preprint arXiv:2610.23667

Code github.com/tmhong/equiprobable

Backup: samples and simulation

T. M. Hong
Pittsburgh

Samples

$t\bar{t}$ · background

- SM default, semileptonic $b\bar{b}lvq\bar{q}'$
- 10^5 events to fit, 10^4 independent test

hh_{125} · SM anomaly

- hh → $b\bar{b}WW^*$, $2 \cdot 10^4$ events

HH_{170} · BSM anomaly

- HH → $b\bar{b}WW$, m_H = 170 GeV, 10^5 events
- hh and HH from the HEFT package

Generation

- MadGraph5_aMC@NLO 3.6.7, LO
- pp at $\sqrt{s}$ = 13 TeV, NNPDF2.3 LO

Simulation & reconstruction

Detector

- Pythia 8.3 shower; Delphes 3 ATLAS card, no pileup
- Anti-k_t R = 0.6 jets, $p_T > 25$ GeV, $|\eta| < 2.5$

Selection

- 1 lepton ($p_T > 25$ GeV), ≥ 4 jets, ≥ 1 b jet
- Passes **24%** of $t\bar{t}$, **20%** of HH

Reconstruction

- v p_z from the m_W constraint
- χ^2 jet assignment over six leading jets: correct for **19%** of $t\bar{t}$, **7%** of HH
- Truth match $\Delta R < 0.4$: **7.1%** of $t\bar{t}$, **3.5%** of HH