Improving Vector Boson Fusion (VBF) LHC Higgs Analyses with Fox-Wolfram Moments

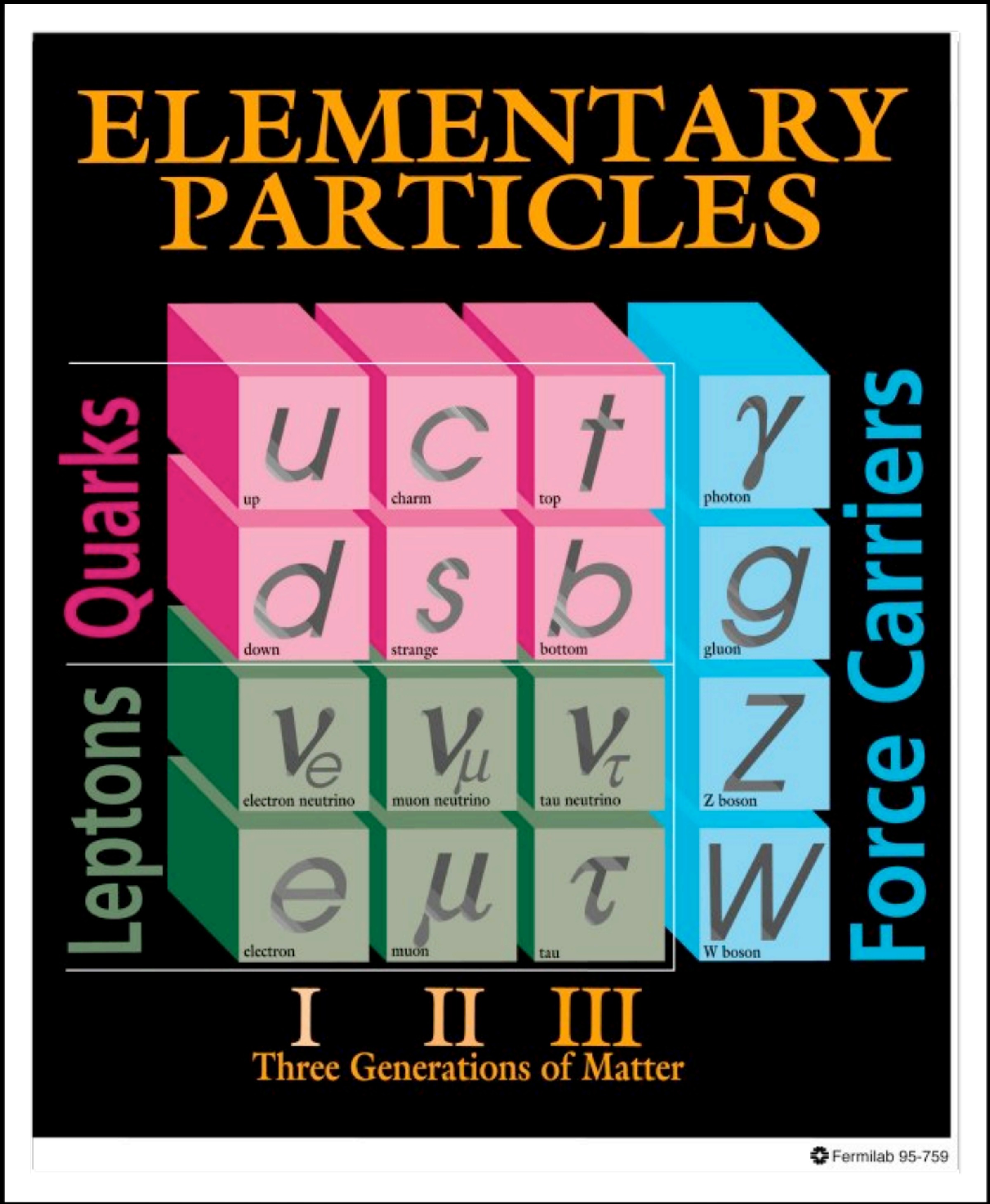
Dr. Catherine Bernaciak - ITP Universität Heidelberg

High Energy Physics Seminar, Department of Physics
University of Virginia, October 2, 2013

Outline

- Review of Standard Model Higgs Mechanism (4)
- Phenomenology of Standard Model Higgs (5)
- Higgs-like Boson Measurement at the LHC (2)
- Fox - Wolfram Moments (8)
- Results of Cut-Based and Boosted Decision Tree Analyses (16)

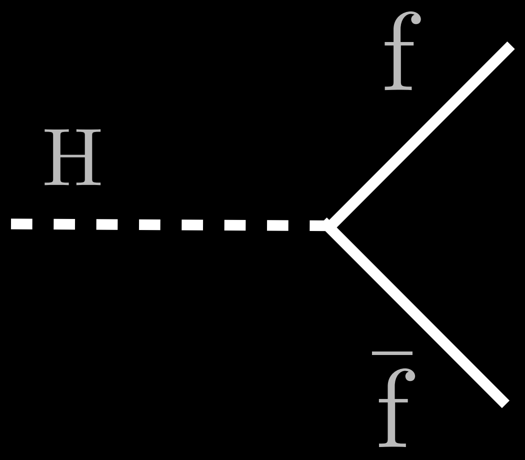
The Standard Model of Particle Physics

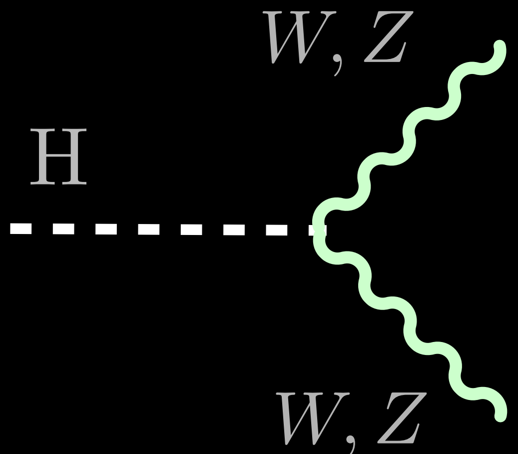


fermions = 6 quarks + 6 leptons

bosons = W, Z, photon, gluon, Higgs

imparts mass to W, Z, quarks, leptons





Electroweak Symmetry Breaking in the Standard Model - QED as a toy model

QED :

$$L = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu)$$

U(1) gauge transformation :

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \eta(x)$$

adding mass to photon field :

$$L = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) + \frac{1}{2}m^2 A_\mu A^\mu$$

violates gauge invariance

a simple, realistic solution:

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + |D_\mu \phi|^2 - V(\phi)$$

a new complex, scalar field

Electroweak Symmetry Breaking in the Standard Model - QED as a toy model

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + |D_\mu\phi|^2 - V(\phi)$$

covariant derivative: $D_\mu = \partial_\mu - ieA_\mu$

simplest, renormalizable,
U(1) invariant potential

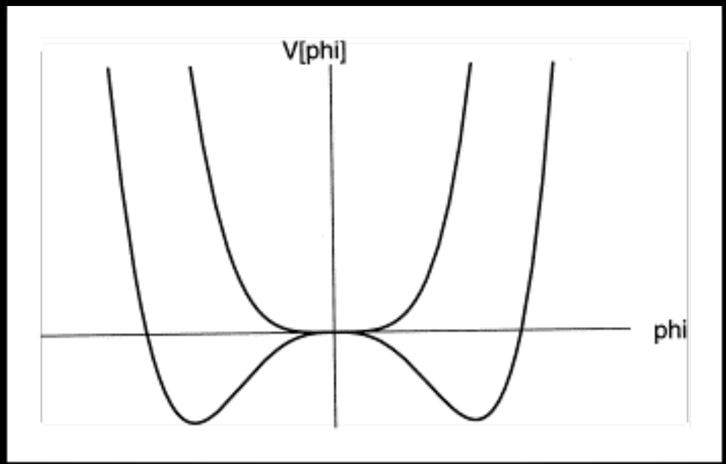
$$V(\phi) = \mu^2|\phi|^2 + \lambda(|\phi|^2)^2$$

$$\phi \to e^{-ie\eta(x)}\phi(x)$$

$\mu^2 < 0$

$\mu^2 > 0$

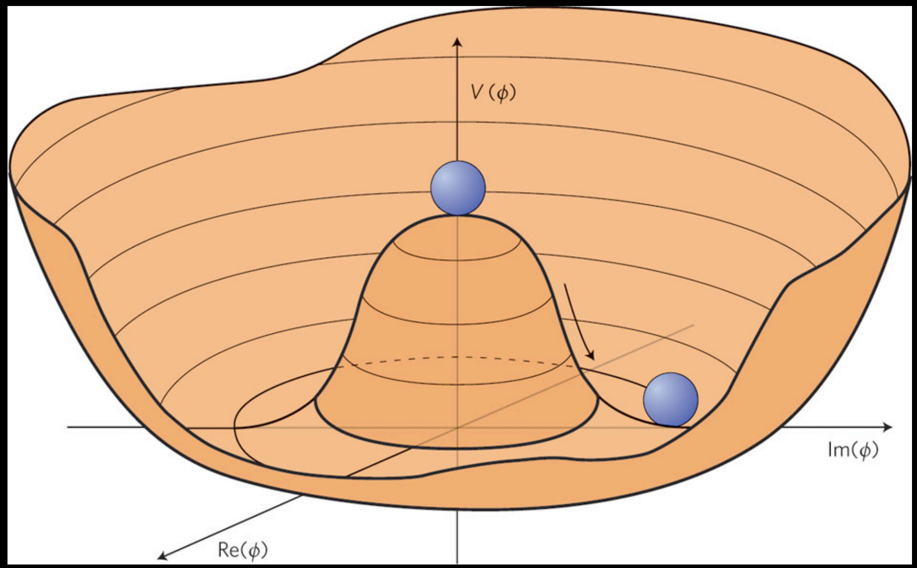
$$\langle\phi\rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}$$



$$\langle\phi\rangle = 0$$

$$\langle\phi\rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}$$

U(1) symmetry is broken with nonzero vev



expand ito of non-vev fields

$$\phi \equiv \frac{1}{\sqrt{2}} e^{i\frac{\chi}{v}}(v+h)$$

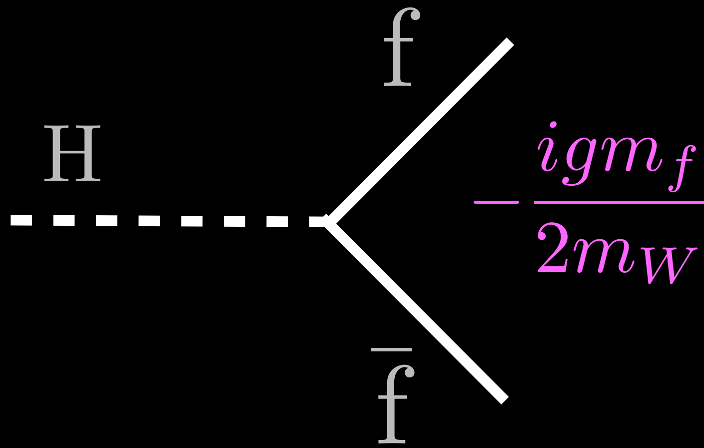
and mass is acquired:

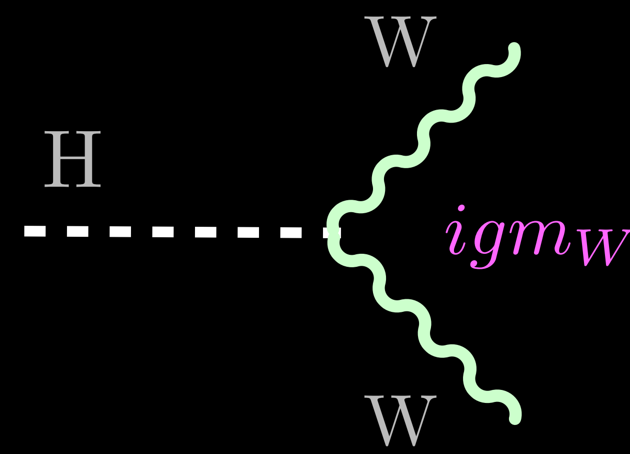
$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - evA_\mu\partial^\mu\chi + \frac{e^2v^2}{2}A_\mu A^\mu + \cdots$$

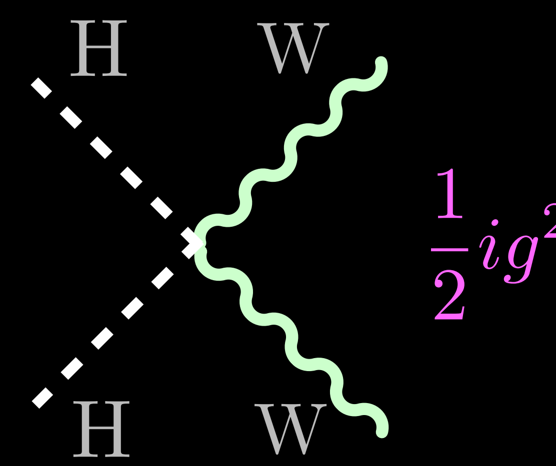
photon field with $M_A = ev$

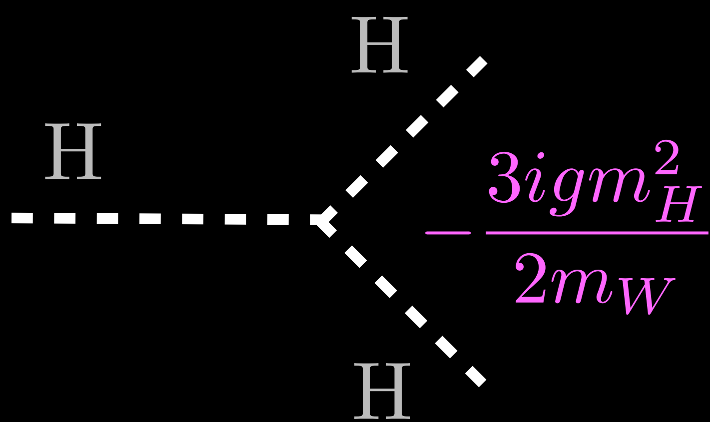
same principle, when applied to electroweak theory causes weak bosons to acquire mass - Higgs field emerges as physical particle....

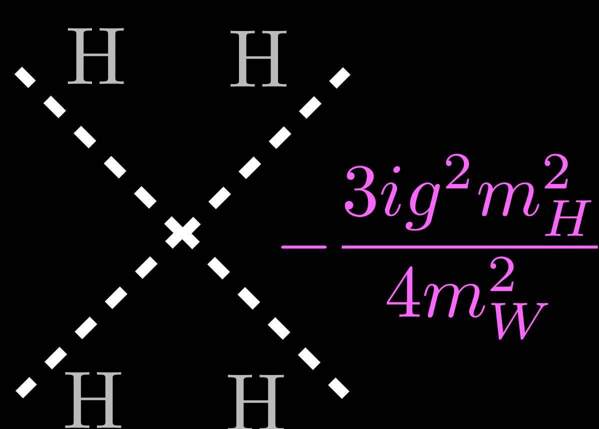
SM Higgs Interactions

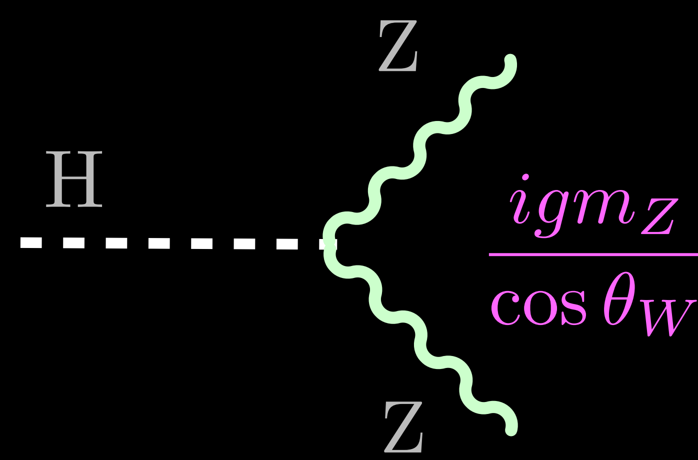


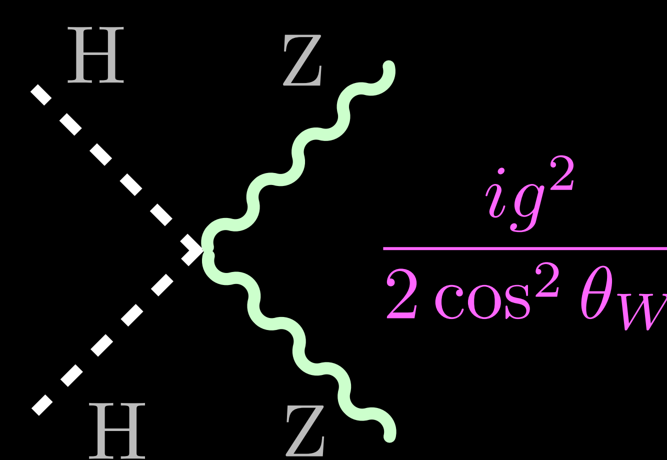










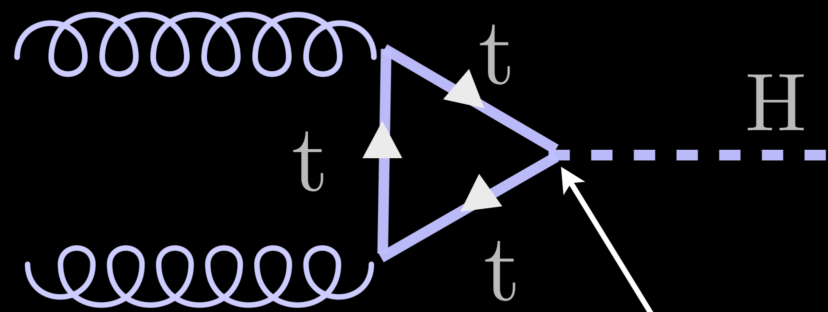


coupling strengths proportional to masses

- **once m_H is known, couplings can be measured and compared to SM prediction**

SM Higgs Production at the LHC

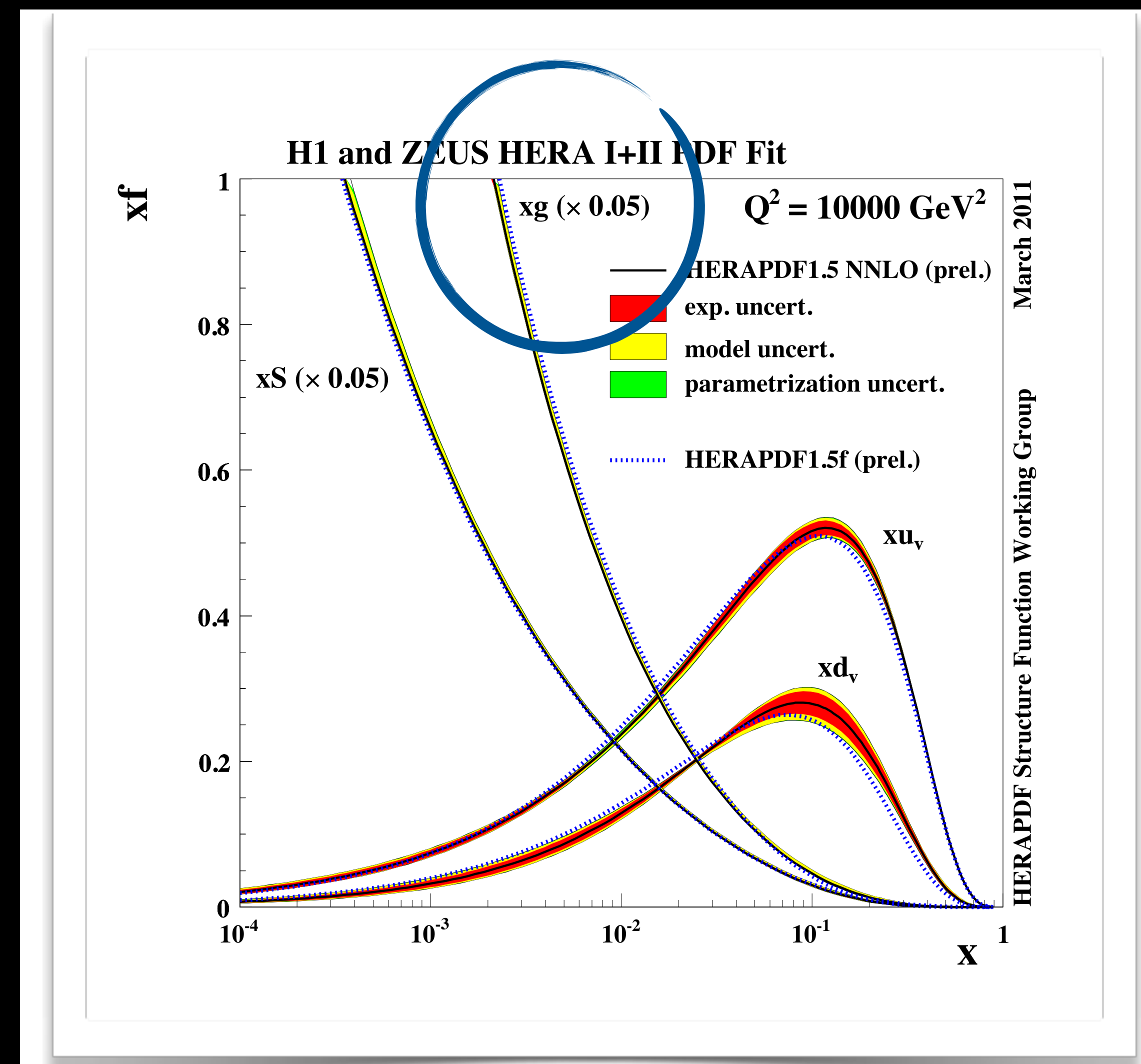
the gluon fusion channel - main LHC production mechanism



“gluon fusion” ggf

bottom loop suppressed by ~ 0.1% - lighter quark loops even less likely

why?! more likely to find a gluon in the proton

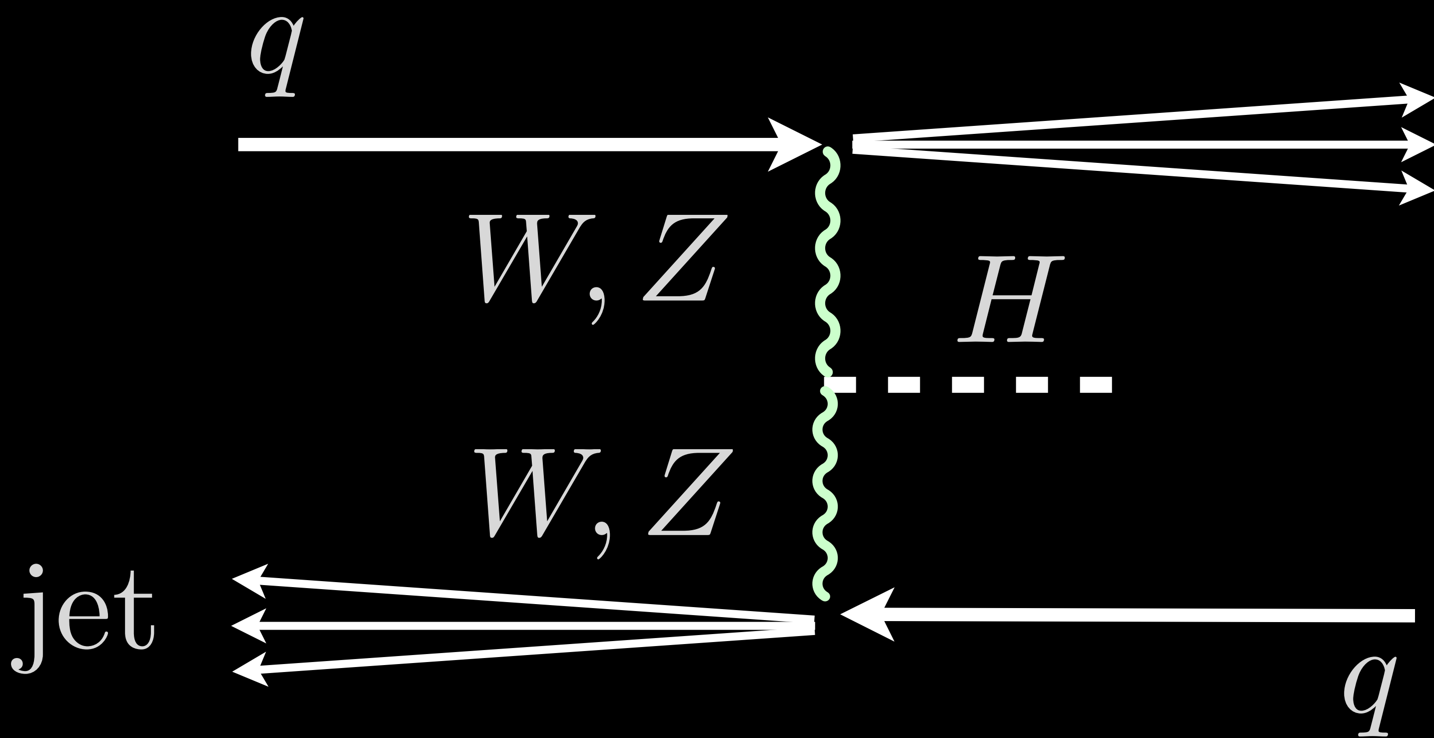


$$\sigma(gg \to H) \approx 15 \text{ pb at } 7 \text{ TeV}$$

$$\sigma(gg \to H) \approx 50 \text{ pb at } 14 \text{ TeV}$$

for $M_H = 125 \text{ GeV}$

SM Higgs Production at the LHC



Vector Boson Fusion

essential probe of EW higgs couplings - deviations from predicted rates could indicate BSM higgs physics

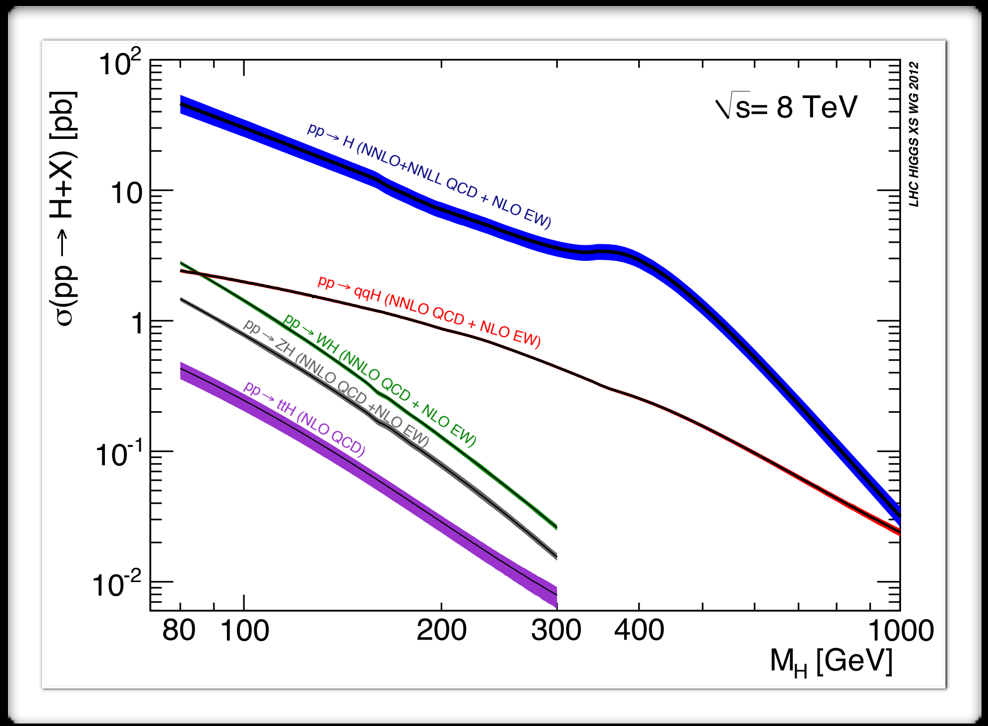
$$\sigma(qqH) \approx 1.3 \text{ pb at } 7 \text{ TeV}$$

$$\sigma(qqH) \approx 4 \text{ pb at } 14 \text{ TeV}$$

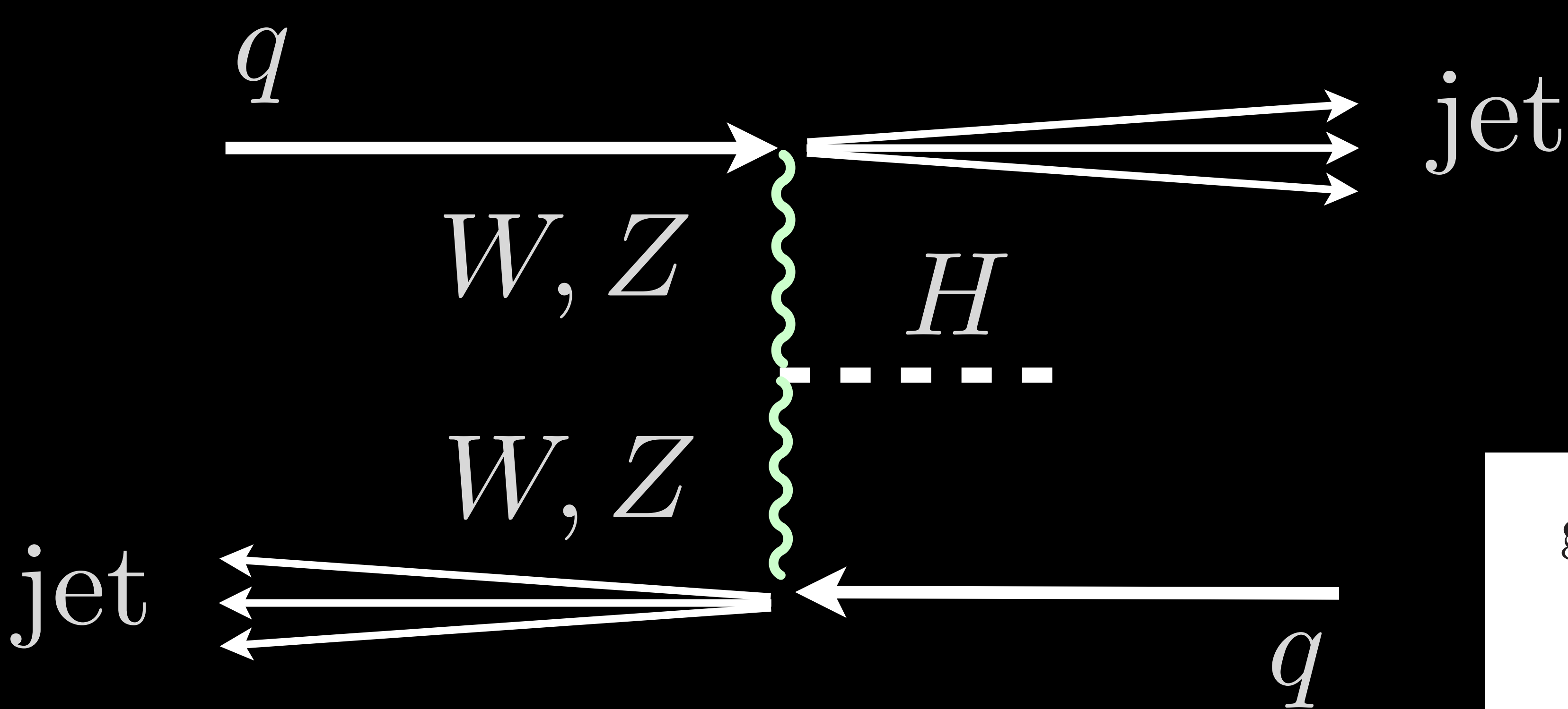
for $M_H = 125 \text{ GeV}$

distinctive "forward - backward" jet topology unlike any background processes

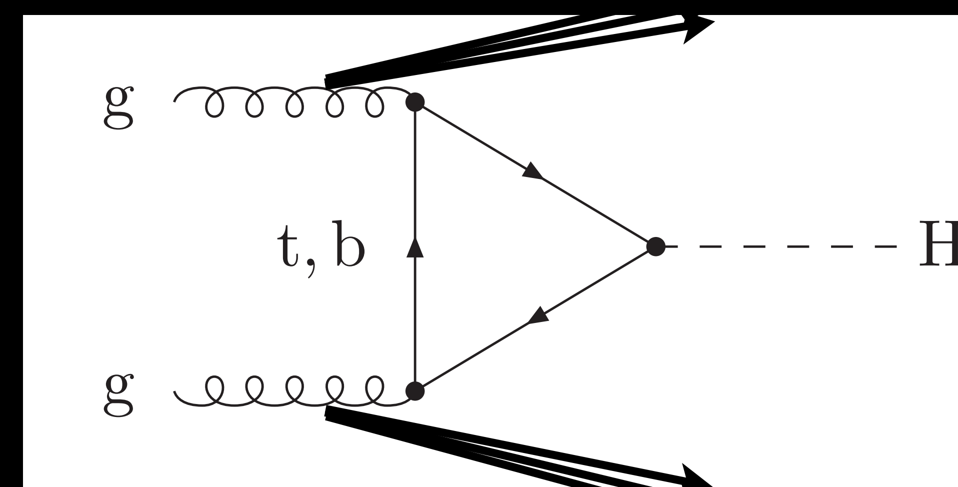
lack of central jet activity - handle for discerning from backgrounds



SM Higgs Production at the LHC



however, ggH + 2jet production could mimic VBF production



solution:
apply acceptance criteria on events to disfavor ggH + 2jet kinematics

$$p_{T j_1 j_2} > 20 \text{ GeV}$$

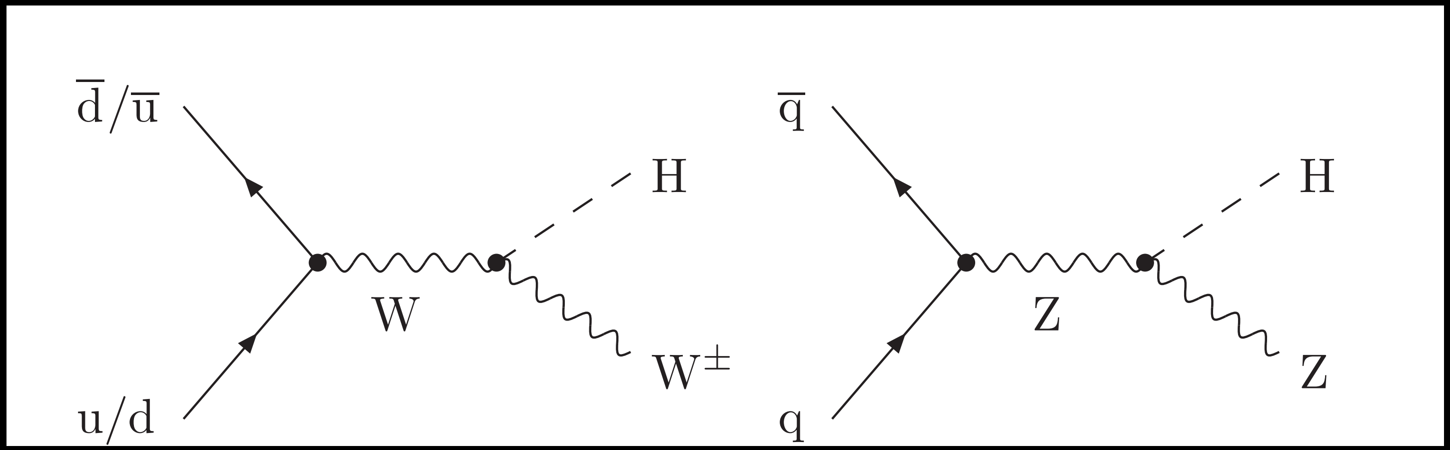
$$\eta_{j_1} \cdot \eta_{j_2} < 0$$

$$\Delta\eta_{j_1,j_2} > 4$$

for
after applying VBF selection, ggH events contribute only 4% - 5% to Higgs production

SM Higgs Production at the LHC

the **Higgs-strahlung** channel

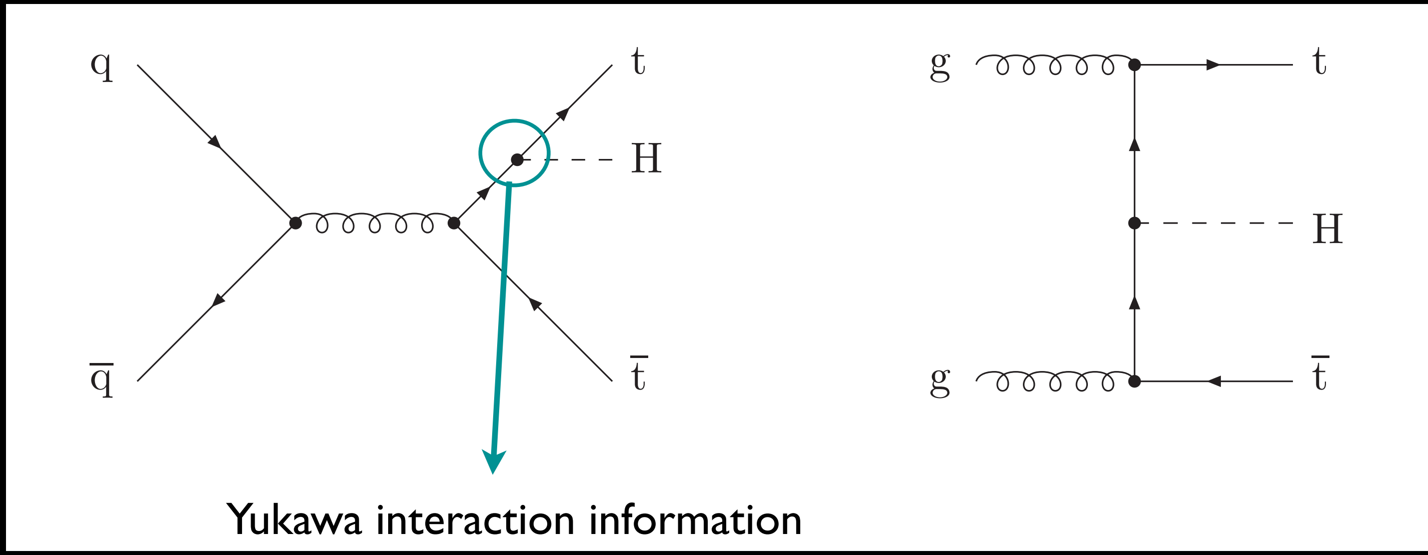


$$\sigma(W, ZH) \approx 0.6 \text{ pb at 7 TeV}$$

$$\sigma(W, ZH) \approx 1.5 \text{ pb at 14 TeV}$$

for $M_H = 125$ GeV

the **ttH** channel - Higgs in association with a top quark



$$\sigma(t\bar{t}H) \approx 88 \text{ fb at 7 TeV}$$

$$\sigma(t\bar{t}H) \approx 611 \text{ fb at 14 TeV}$$

for $M_H = 125$ GeV

SM Higgs Decay at the LHC

$H \to b\bar{b}$

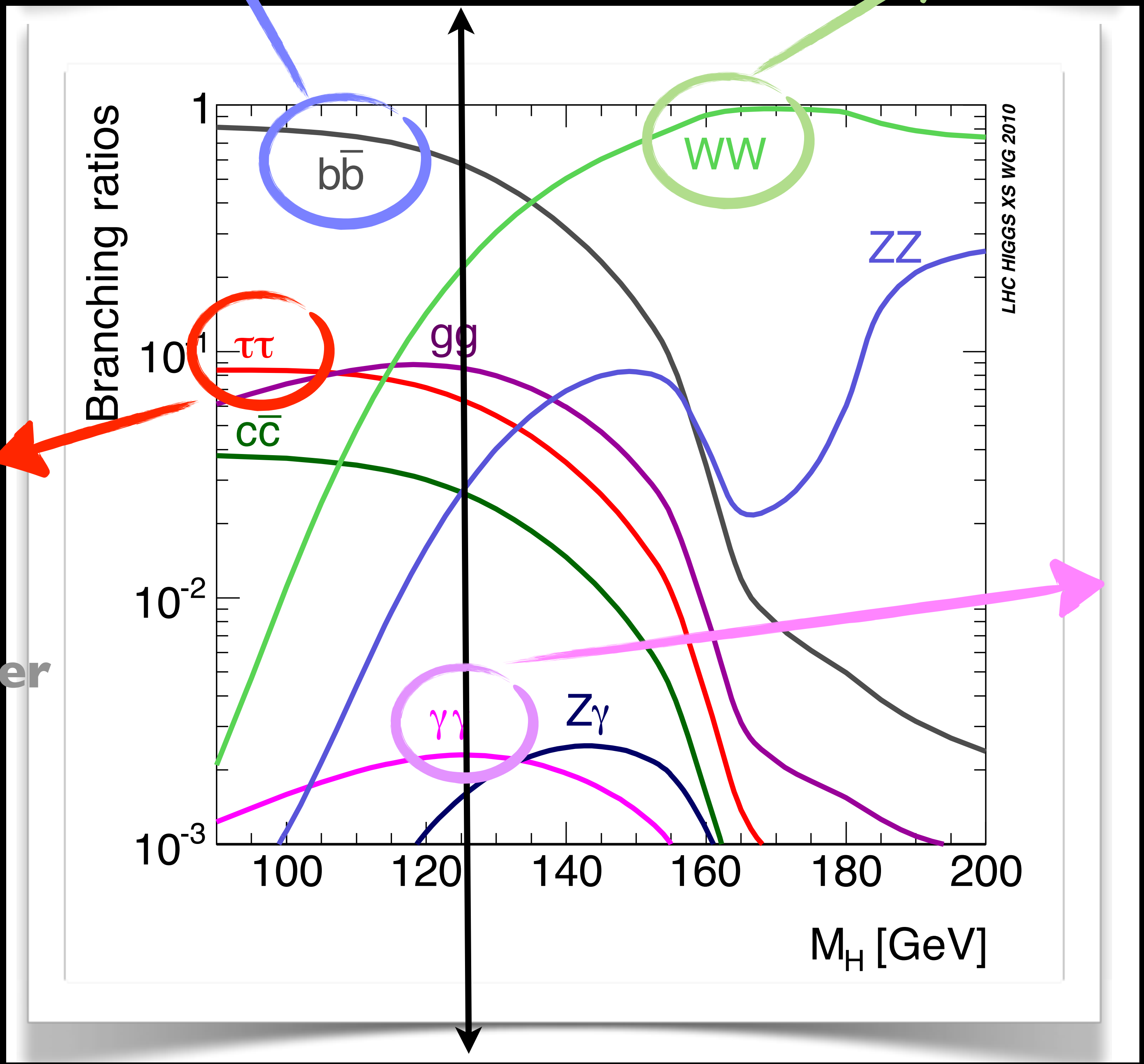
high decay rate, but b tagging efficiency at 60%

$WW \to l\nu l\nu$

clean and efficient, background under control

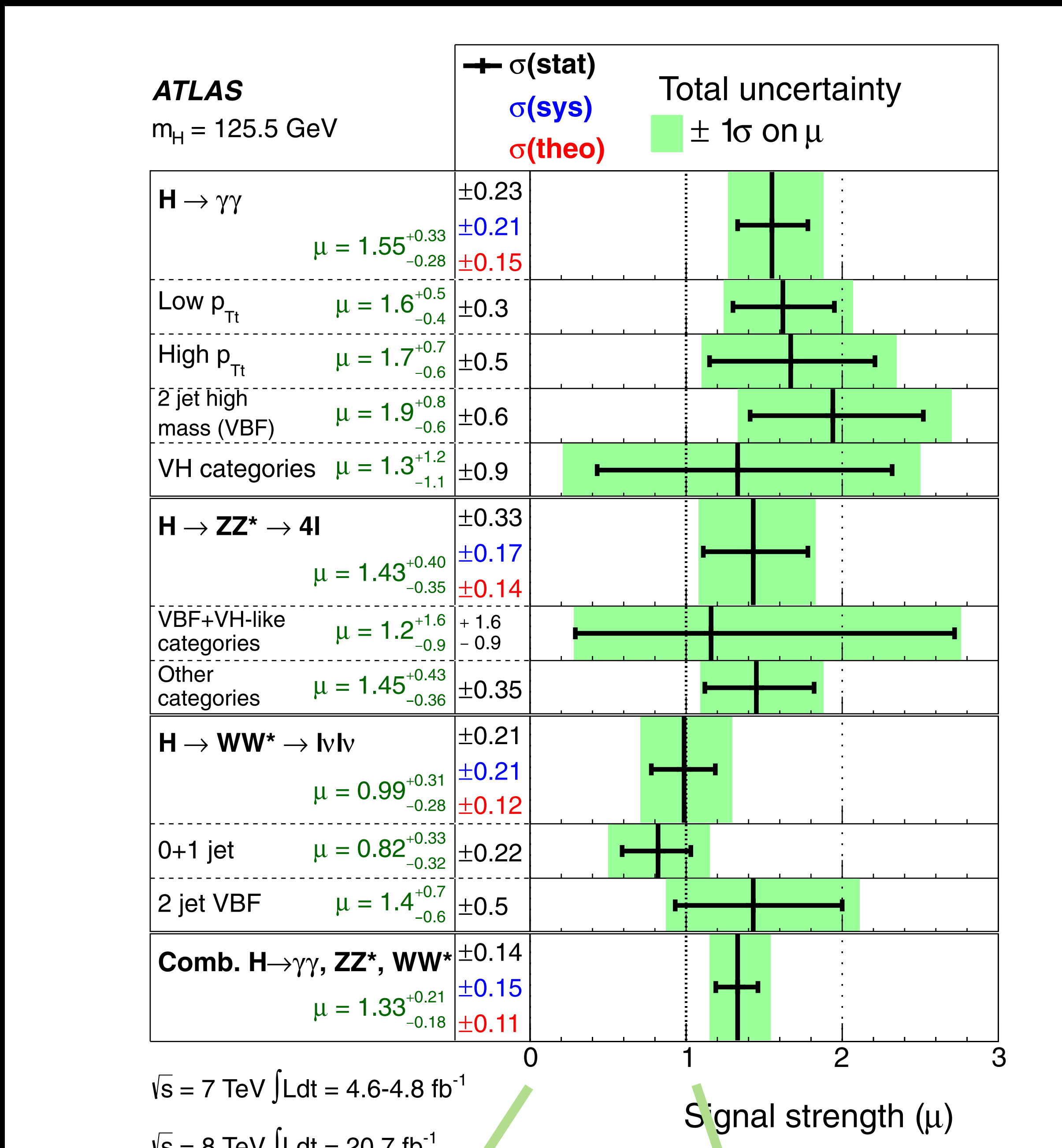
$\tau \to l\nu\nu$

clean and efficient, background under control



comparatively low rate, but very clean and efficient

LHC Higgs-like Boson Discovery



$$\mu = 0$$

background only hypothesis

$$\mu = 1$$

SM Higgs boson hypothesis

combined mass measurement:

ATLAS:

$$m_H = 125.5 \pm 0.2(\text{stat})^{+0.5}_{-0.6}(\text{sys})\ \text{GeV}$$

CMS:

$$m_H = 125.3 \pm 0.4(\text{stat}) \pm 0.5(\text{sys})\text{GeV}$$

combined signal strength measurement:

ATLAS:

$$\mu = 1.33 \pm 0.14(\text{stat}) \pm 0.15(\text{sys})$$

CMS:

$$\hat{\mu} = \frac{\sigma}{\sigma_{\text{SM}}} = 0.87 \pm 0.23$$

for M_H = 125 GeV

consistent with SM Higgs hypothesis

Theoretical Uncertainties in Higgs Measurement

large systematic uncertainty from **higher order QCD calculations matched to parton shower** - common to both ATLAS and CMS

ATLAS[1]

Source (theory)	Uncertainty (%)
QCD scale	± 8 (ggF), ± 1 (VBF, VH), $^{+4}_{-9}$ (ttH)
PDFs + α_s	± 8 (ggF, ttH), ± 4 (VBF, VH)

CMS[2]

Source	Range (%)
Integrated luminosity	2.2-4.4
Lepton identification and trigger efficiency (per lepton)	3
$Z(\nu\nu)H$ triggers	2
Jet energy scale	2-3
Jet energy resolution	3-6
Missing transverse energy	3
b-tagging efficiency	3-15
Signal cross section (scale and PDF)	4
Signal cross section (p_T boost, EWK/QCD)	5-10/10
Statistical precision of signal simulation	1-5
Backgrounds estimated from data	10
Backgrounds estimated from simulation	30

include more EW and QCD higher-order corrections, **resum EW Sudakov logs in VHbb ...**

better match parton shower to existing NLO and NNLO and implement in simulation tool **SHERPA, MC@NLO, MADGRAPH ...**

1 PLB 726 (2013) 88-119, 2 JHEP 06 (2013) 081

The Fox-Wolfram Moments[1]

a rotationally invariant set of observables constructed from Legendre polynomials

$$H_\ell = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{s} P_\ell(\cos\Omega_{ij})$$

- $\sum_{i,j}$ → correlations between hadrons, jets, calorimeter entries...
- $\frac{|\vec{p}_i||\vec{p}_j|}{s}$ → weight factor
- Ω_{ij} → total angle between objects

$$\cos\Omega_{ij} = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \cos(\phi_i - \phi_j)$$

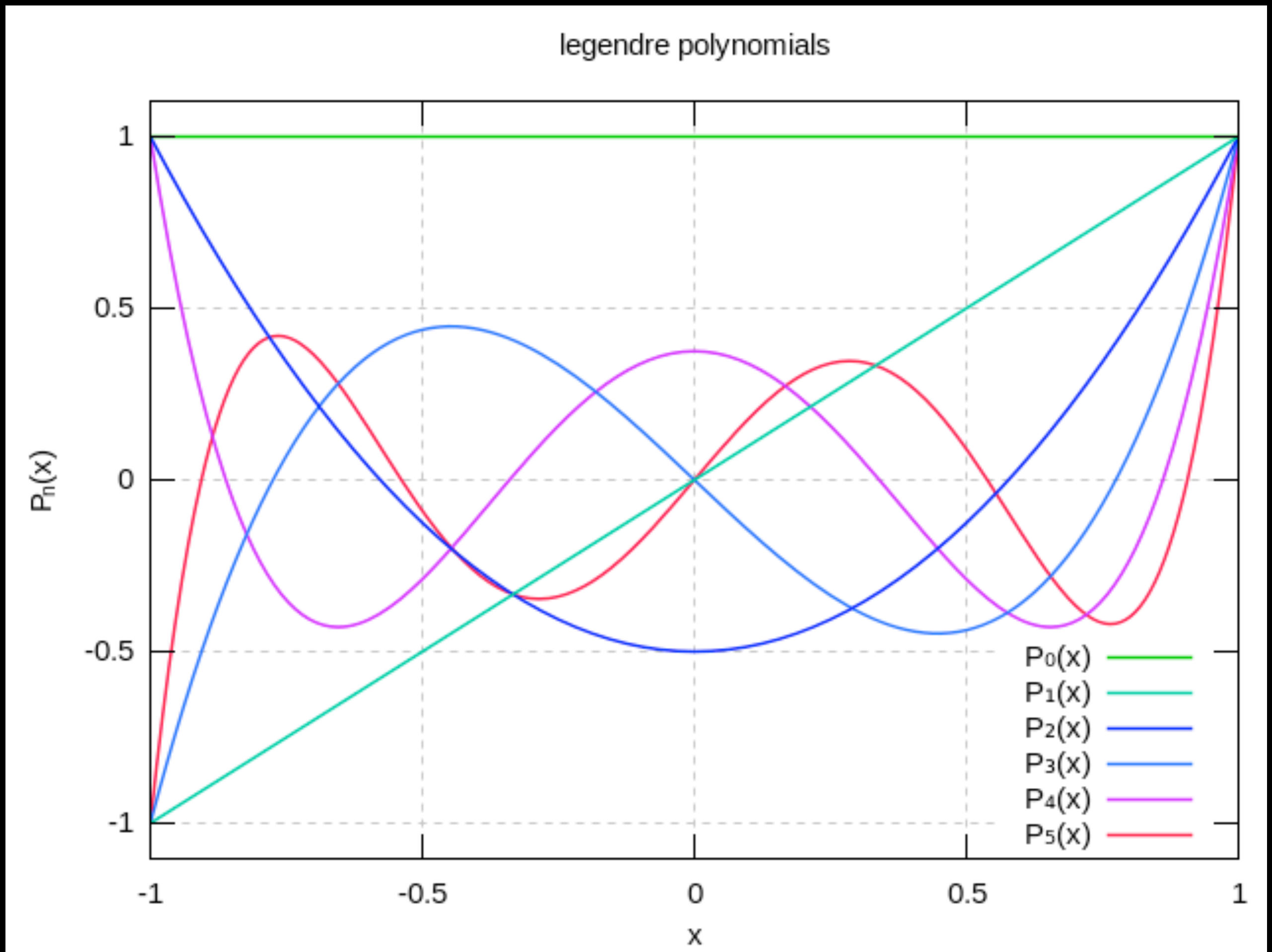
[1] Fox, Wolfram, PRL 1978

Legendre Polynomials

occur as series solution to Laplace's equation in spherical coordinates

$$\frac{d}{dx}\left[(1-x^2)\frac{d}{dx}P_n(x)\right] + n(n+1)P_n(x) = 0$$

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[(x^2-1)^n\right]$$



$$P_0(x) = 1\,,\; P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2-1)$$

$$P_3(x) = \frac{1}{2}(5x^3-3x)$$

$$\vdots$$

$$P_7(x) = \frac{1}{16}(429x^7 - 693x^5 + 315x^3 - 35x)$$

The Fox-Wolfram Moments

an event shape observable describing correlations between four-momentum objects

- **e+ e- to jets**

Fox, Wolfram Nucl. Phys. B 149 (1979) 413-496

- **Top Quark signal at Tevatron**

Field, Kanev, Tayebnejad PRD 55, 9 (1997)

- **B meson decays at Belle:**

Toru Iijima, hep-ex 0105005 (2001)

- **Higgs physics at the LHC:** **VBF H tautau vs Z+2j and Top Pair**

C.B., Buschmann, Butter, Plehn PRD 87, 073014 (2013)

- **A Multivariate study of Fox-Wolfram Moments for Higgs Analyses at the LHC**

C.B., Mellado, Plehn, Ruan, Schichtel, in preparation

The Fox-Wolfram Moments

$$H_\ell = \sum_{i,j} \frac{|\vec{p}_i||\vec{p}_j|}{s} P_\ell(\cos\Omega_{ij})$$

"weight factor"

$$0 \leq H_\ell \leq 1$$

$$W_{ij}^T = \frac{p_{Ti}p_{Tj}}{p_{T,\text{tot}}^2} \qquad W_{ij}^U = 1 \qquad W_{ij}^p = \frac{|\vec{p}_i||\vec{p}_j|}{|\vec{p}|_{\text{tot}}^2}$$

transverse momentum weight

unit weight

magnitude momentum weight

Fox-Wolfram Moments - 2 jet properties

$$H_\ell = \sum_{i,j=1}^{2} \frac{W_i W_j}{W_{\text{tot}}^2} P_\ell(\cos\Omega_{ij})$$

$$= \frac{1}{(W_1 + W_2)^2} \left[W_1^2 P_\ell(\cos 0) + W_2^2 P_\ell(\cos 0) + W_1 W_2 P_\ell(\cos\Omega_{12}) \right]$$

$$= 1 + \frac{2W_1 W_2}{(W_1 + W_2)^2 P_\ell(\cos\Omega_{12})}$$

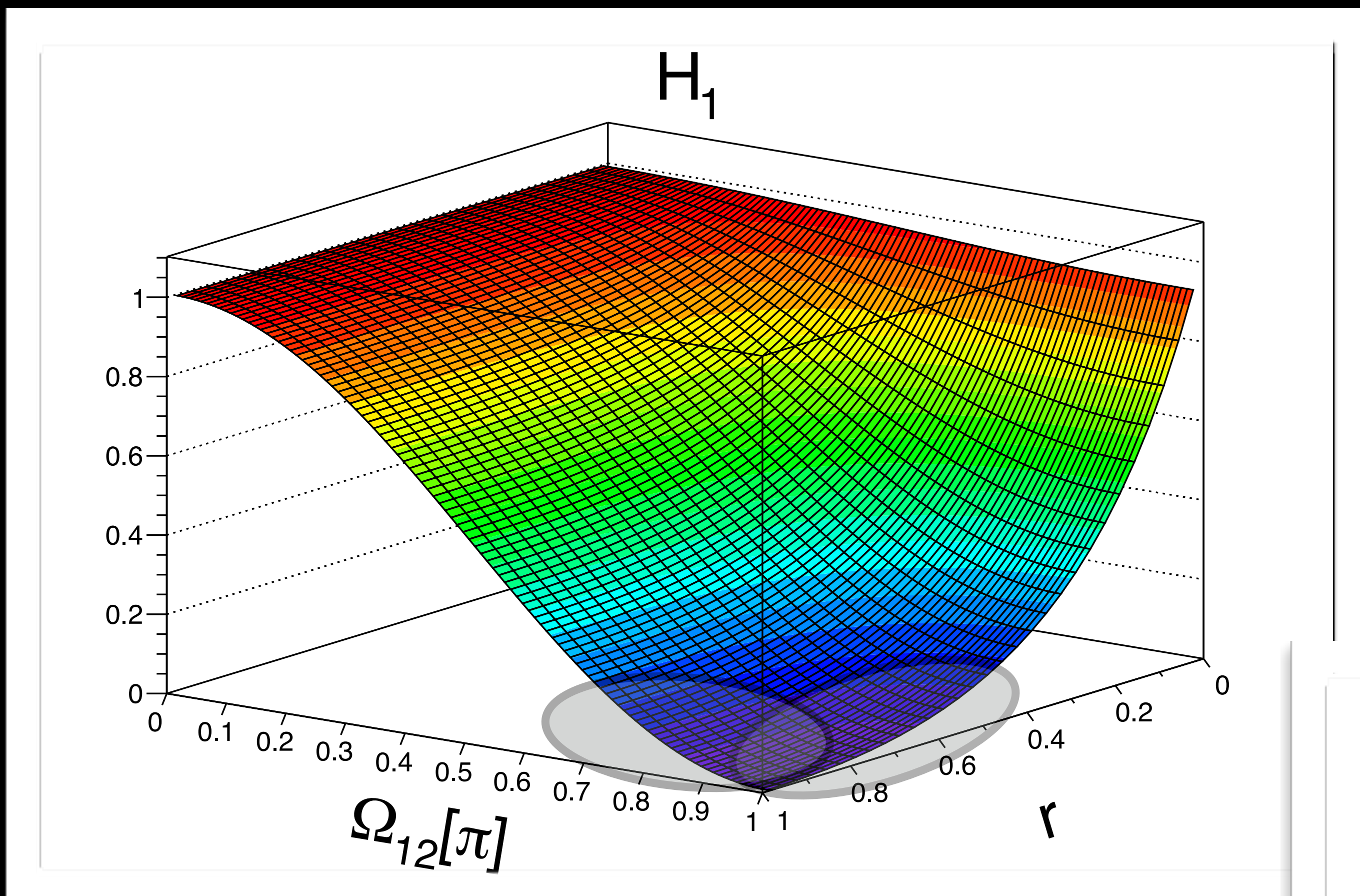
$$= \frac{1 + 2r P_\ell(\cos\Omega_{12})}{1 + 2r + r^2}$$

$$r = \frac{W_2}{W_1}$$

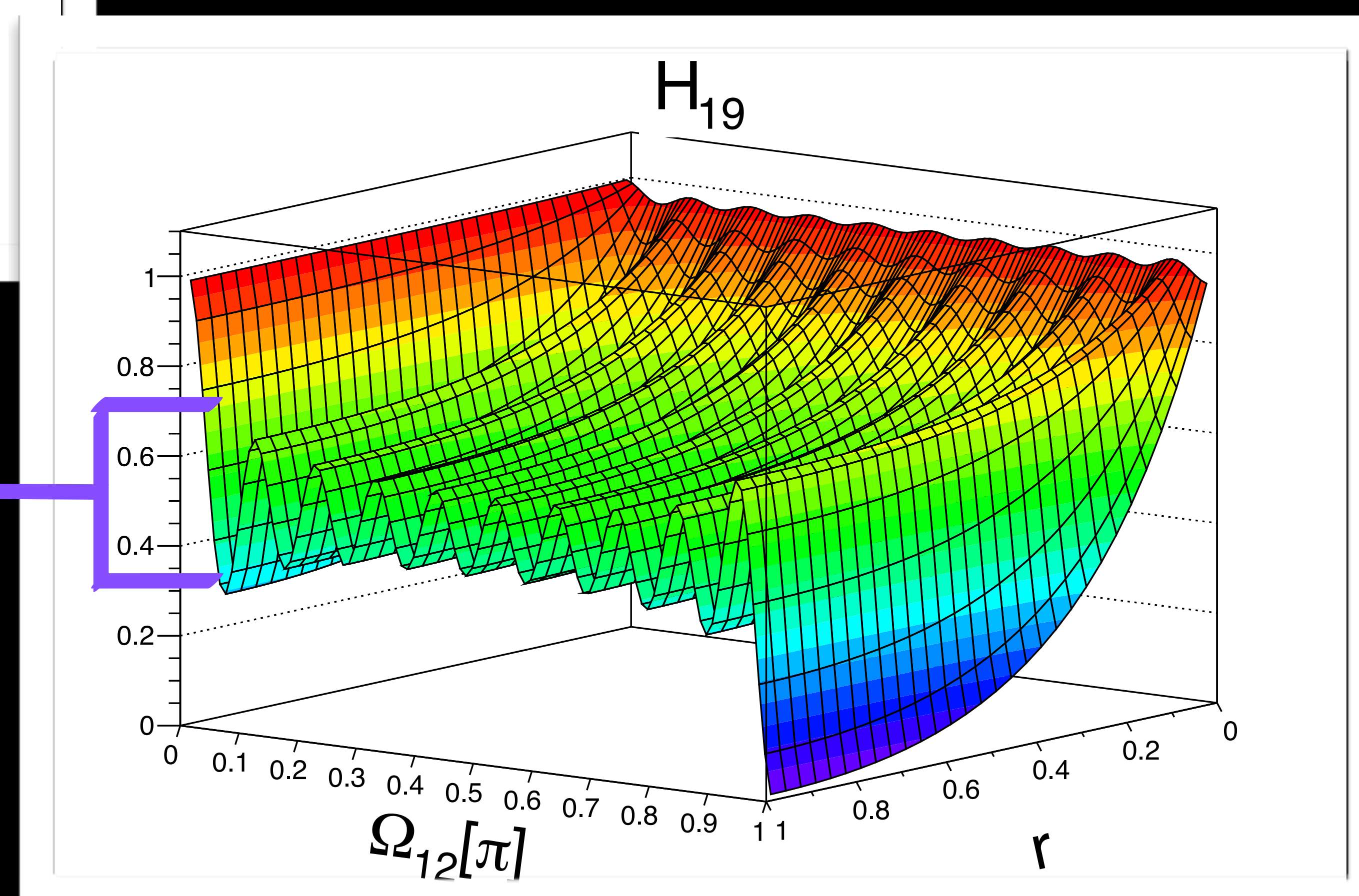
$$0 \le r \le 1$$

Fox-Wolfram moments - 2 jet properties

odd moments - best for discriminating back-to-back jets, higher moments resolve larger angular j_1 j_2 separation



$$r = \frac{W_2}{W_1}$$



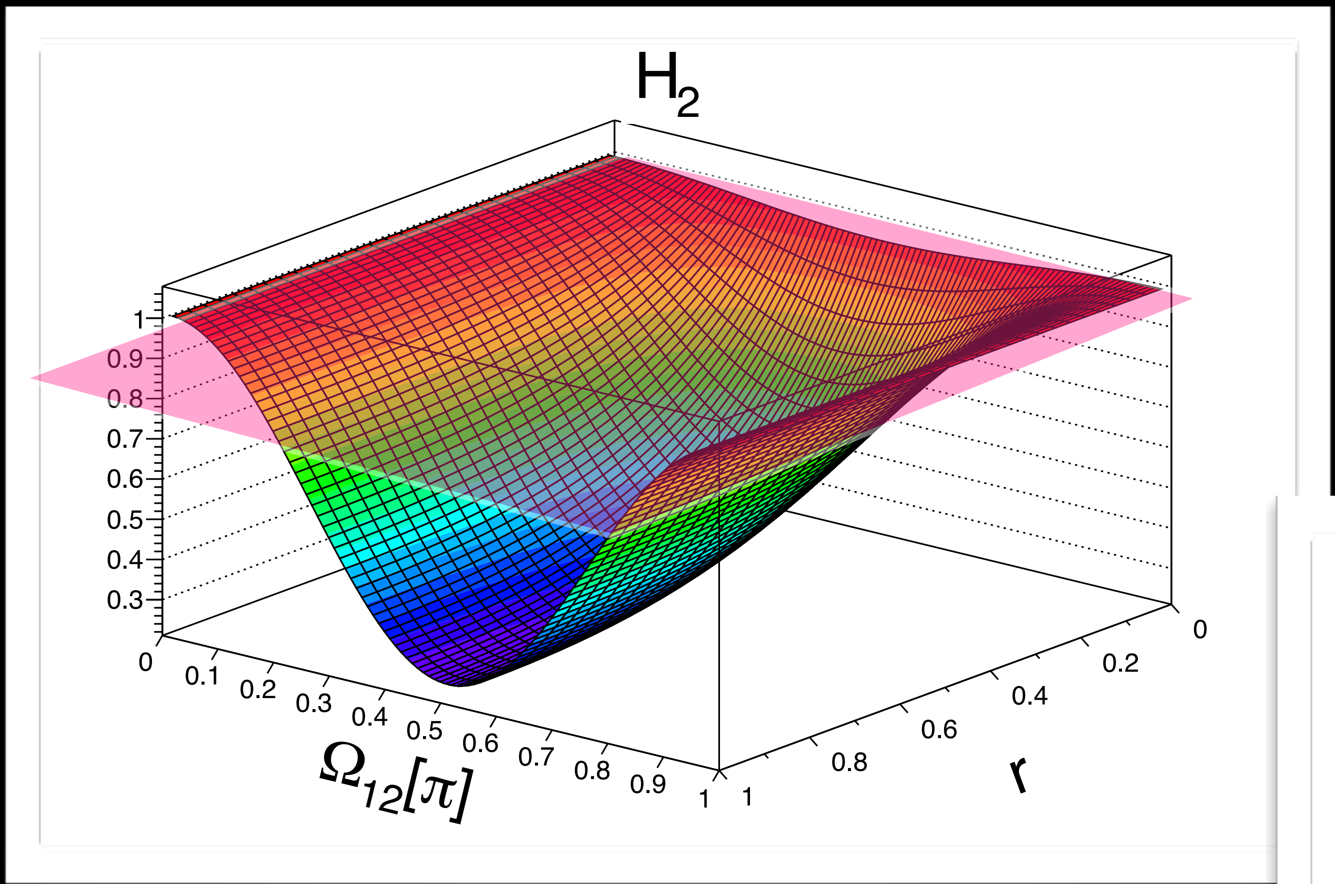
multivalued function, no resolution to intermediate values of Ω_{12}

$$H_\ell \to 0 \quad \text{for} \quad \Omega_{12} \to \pi$$

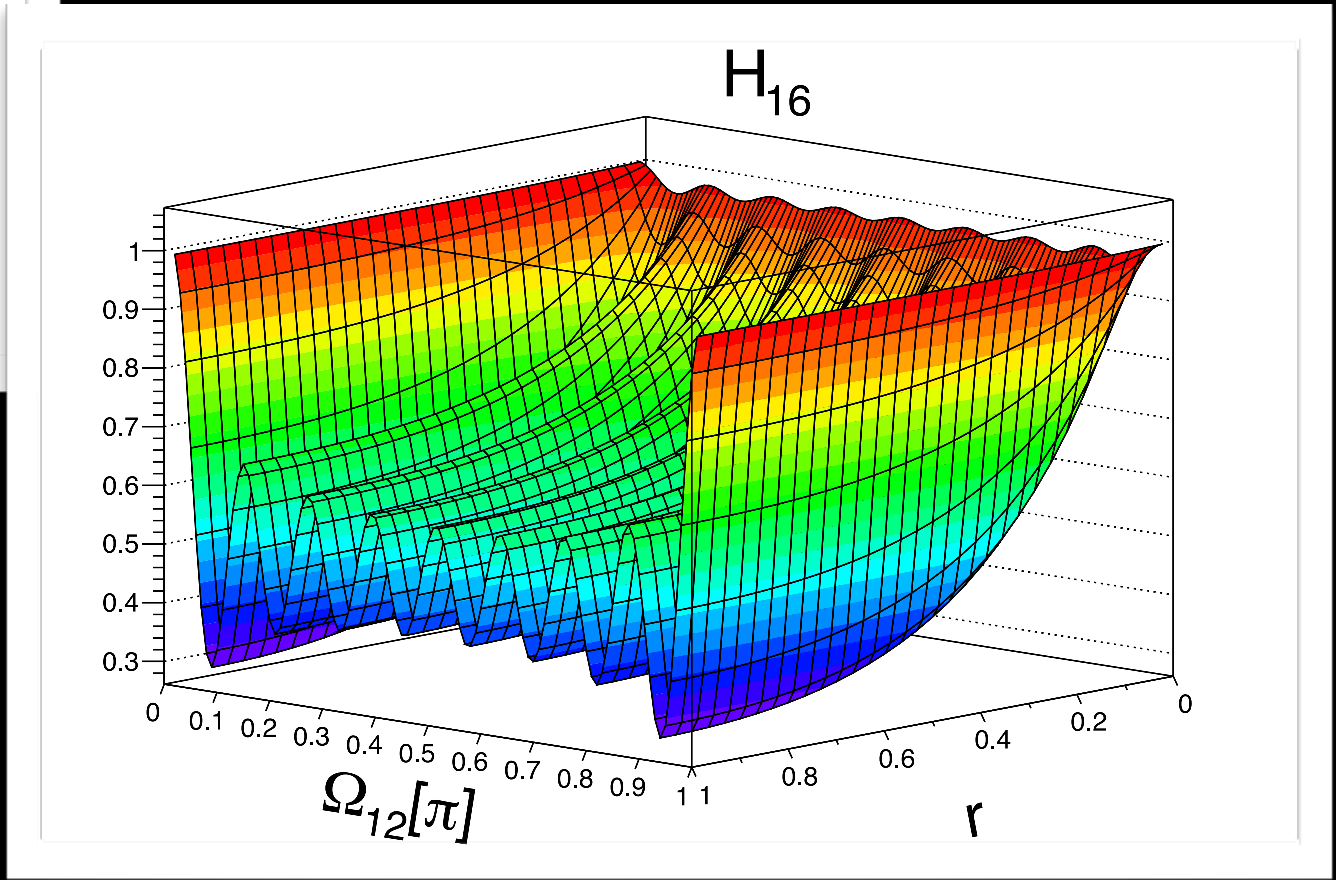
Fox-Wolfram moments - 2 jet properties

even moments - symmetry of even function reduces discriminatory power

$$H_\ell \to 1 \quad \text{for} \quad \Omega_{12} \to 0 \quad \text{AND} \quad \Omega_{12} \to \pi$$



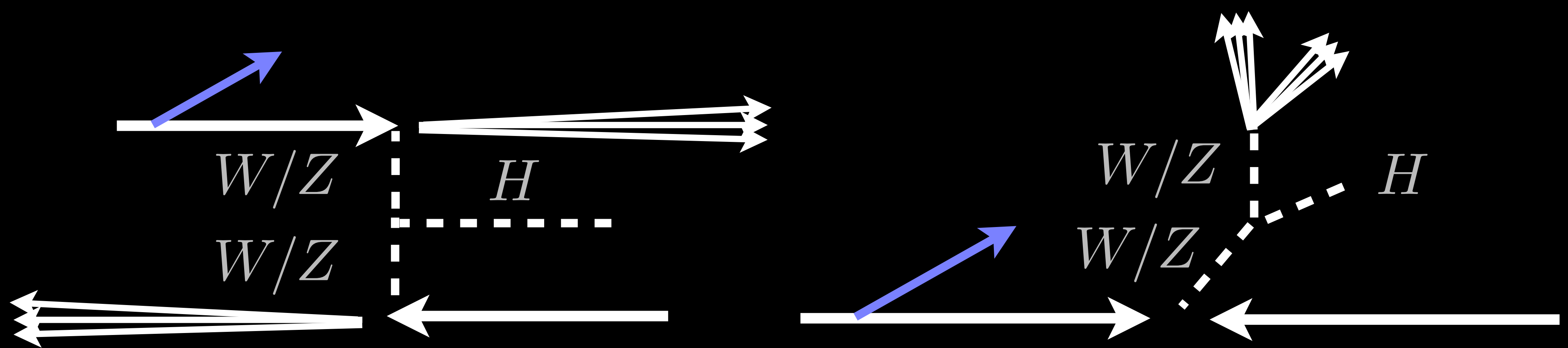
low, even moments may discern non forward-backward jets



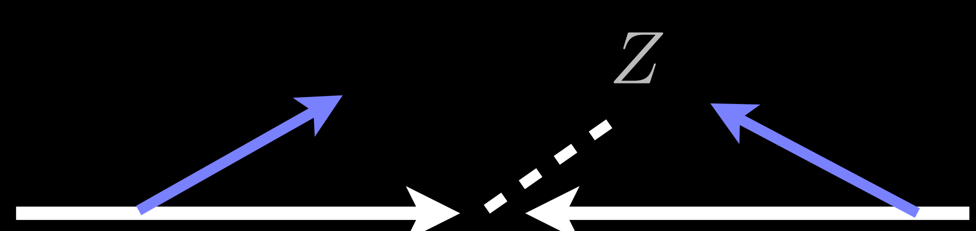
Analysis for H --> tau tau

(process + hard jet) x PS with CKKW using SHERPA

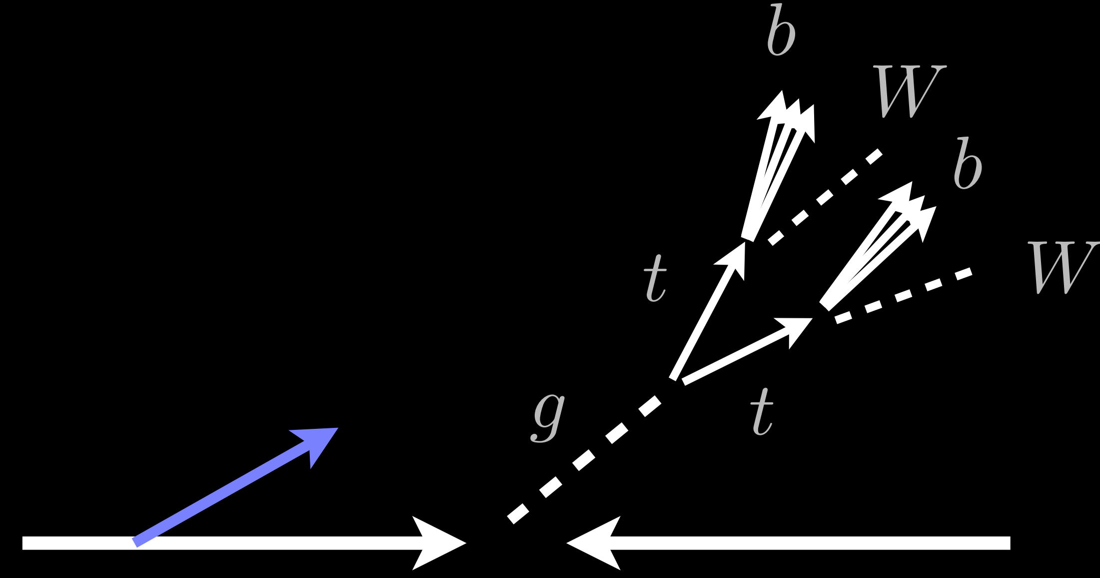
signal WBF



background QCD ZJJ



background Top Pair



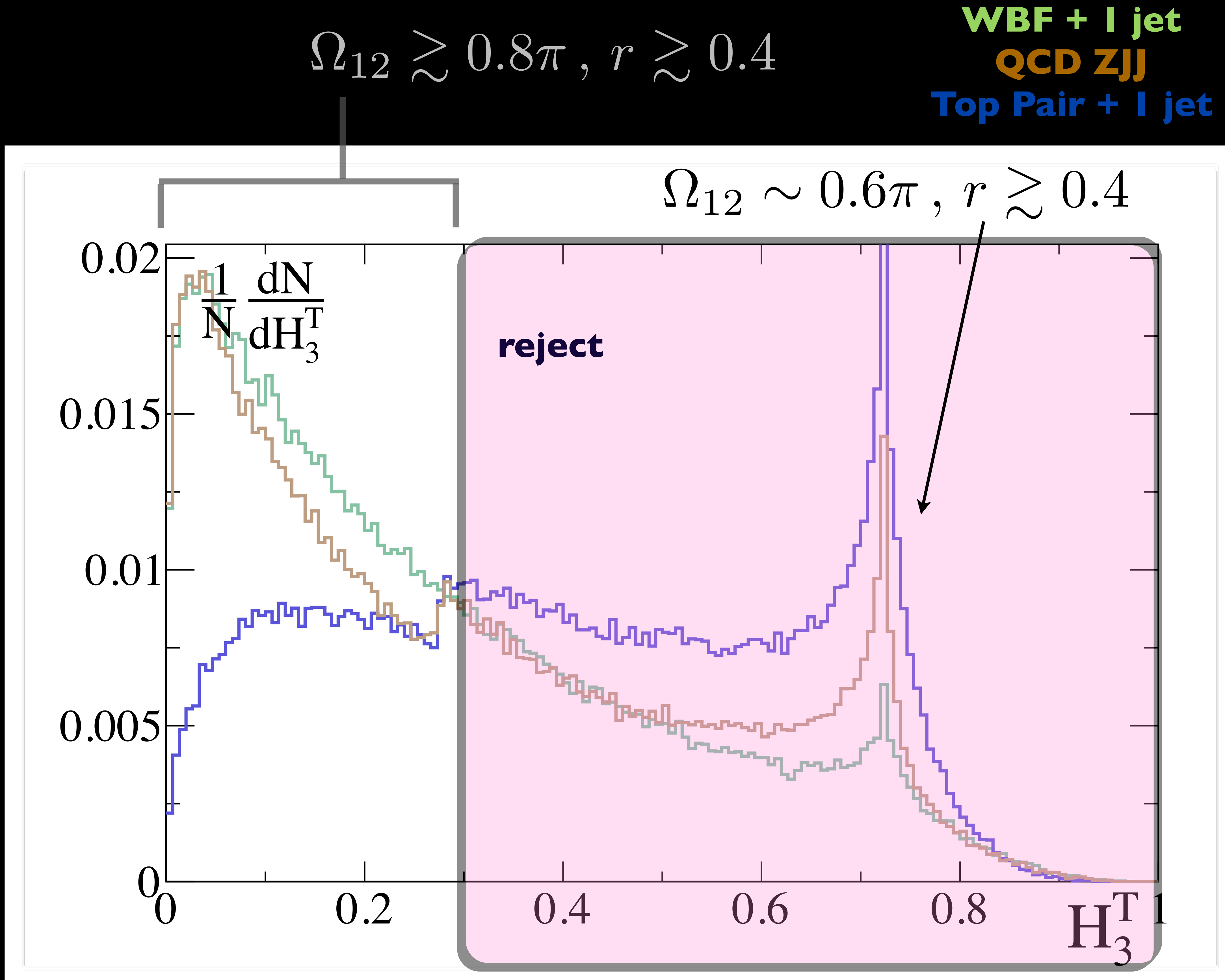
Fastjet anti-kT algorithm with R = 0.4, 8TeV

Cutflow Analysis

acceptance	WBF + 1 jet % fail	WBF + 1 jet XS (fb)	QCD ZJJ % fail	QCD ZJJ XS (fb)	Top Pair % fail	Top Pair XS (fb)	S / B
		18.7		115000		17200	1/7070
$p_{T j_1,j_2} > 20$ GeV	29.4	13.2	93.2	7820	9.63	15500	1/1767
$\|y_{j_1,j_2}\| < 5.0$	1.49	13.0	0.97	7740	0.182	15500	1/1788
$\Delta R_{j_1 j_2} > 0.7$	2.73	12.6	3.84	7440	2.32	15100	1/1789
$m_{j_1 j_2} > 600$ GeV	68.9	3.92	96.6	253	95.8	634	1/226
$b - veto$	NA	3.92	NA	253	54.0	292	1/139
$y_1 \cdot y_2 < 0$	1.41	3.86	9.17	230	13.8	252	1/125
$\|y_{j_1} - y_{j_2}\| > 4.4$	13.9	3.32	31.8	157	66.1	85.4	1/73

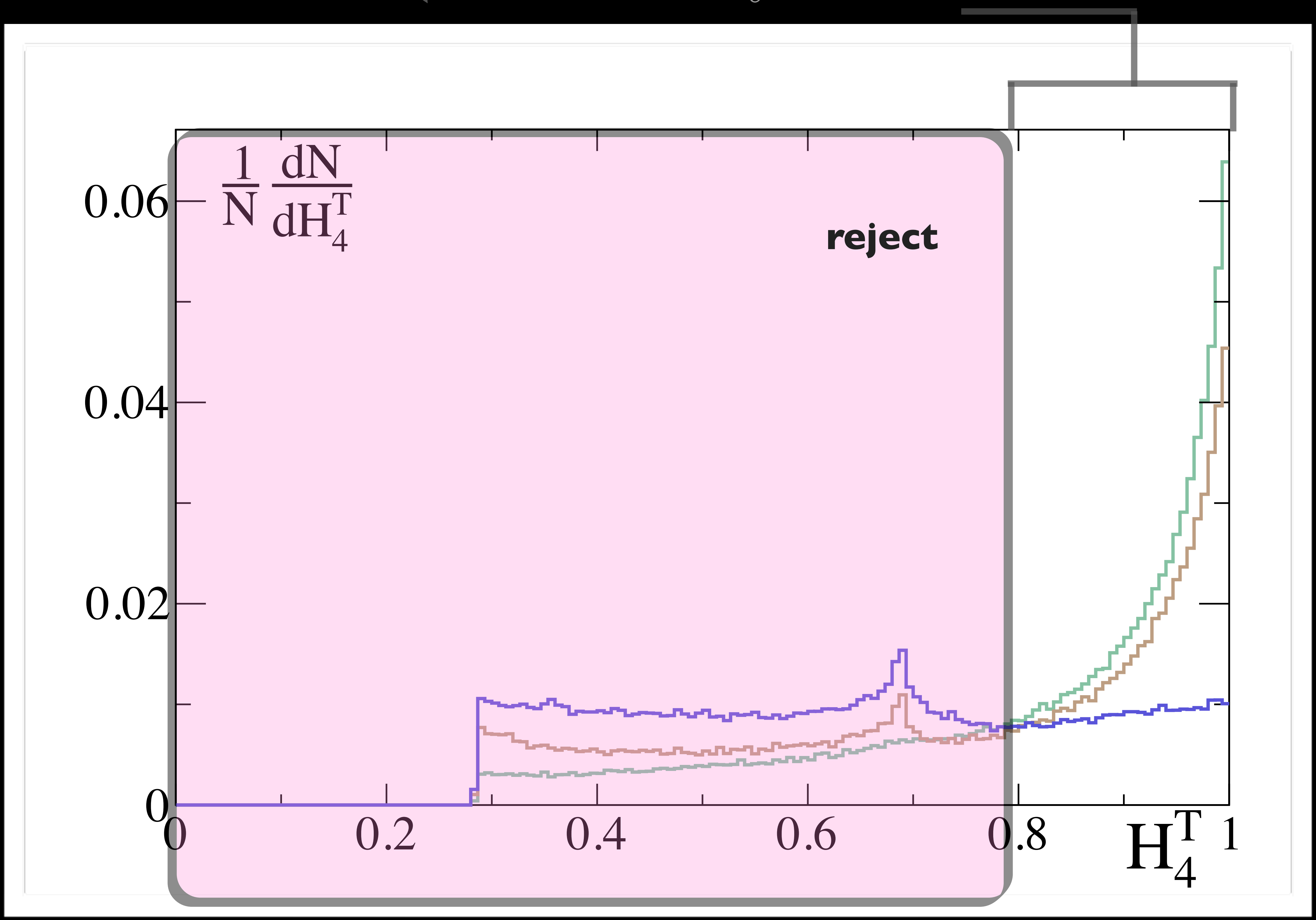
can cuts on FWM replace or be added to current cuts used for VBF event selection?

Cuts on FWM Distributions



Cuts on FWM Distributions

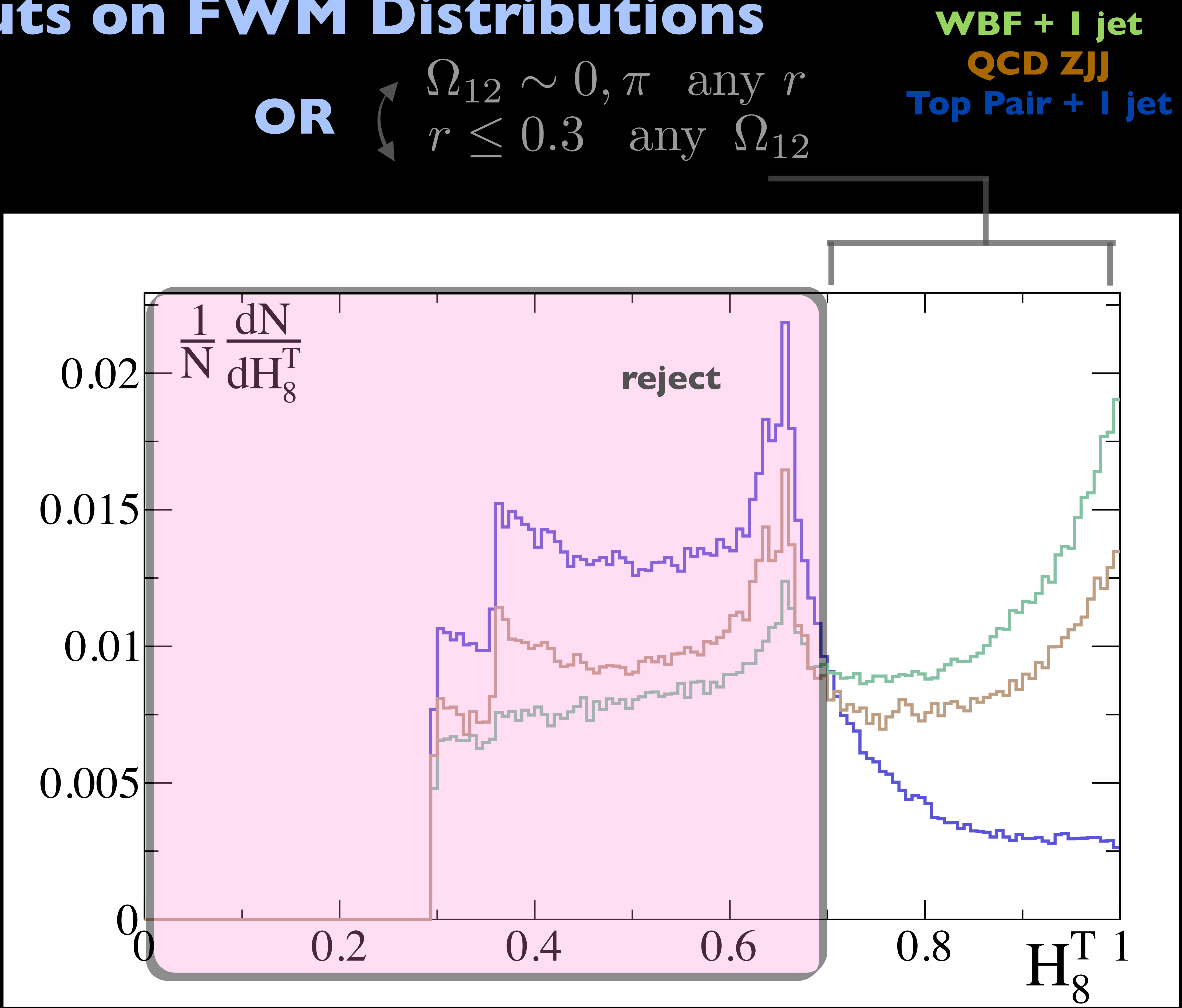
OR $\Omega_{12} \sim 0, \pi$ any r
$r \leq 0.3$ any Ω_{12}



Cuts on FWM Distributions

OR

$$\Omega_{12} \sim 0, \pi \quad \text{any } r$$
$$r \leq 0.3 \quad \text{any } \Omega_{12}$$



Cuts on FWM Distributions[1]

acceptance	WBF + 1 jet		QCD ZJJ		Top Pair		S / B
	% fail	XS (fb)	% fail	XS (fb)	% fail	XS (fb)	
min cuts + b-veto		3.92		253		292	1/139
$H_3^T < 0.3$	38.4	2.41	44.4	141	64.6	103	1/101
$H_4^T > 0.8$	35.8	2.52	48.1	131	73.3	78.0	1/83
$H_8^T > 0.8$	50.1	1.96	60.5	100	81.6	53.7	1/78
$H_{12}^T > 0.7$	64.5	1.39	73.0	68.3	88.0	35.0	1/74
rapidity gap	13.9	3.32	31.8	157	66.1	85.4	1/73

[1]C.B. et.al, PRD 87, 073014 (2013)

Analysis - Cutting on FWM

after typical WBF cuts are exhausted, can the moments help?

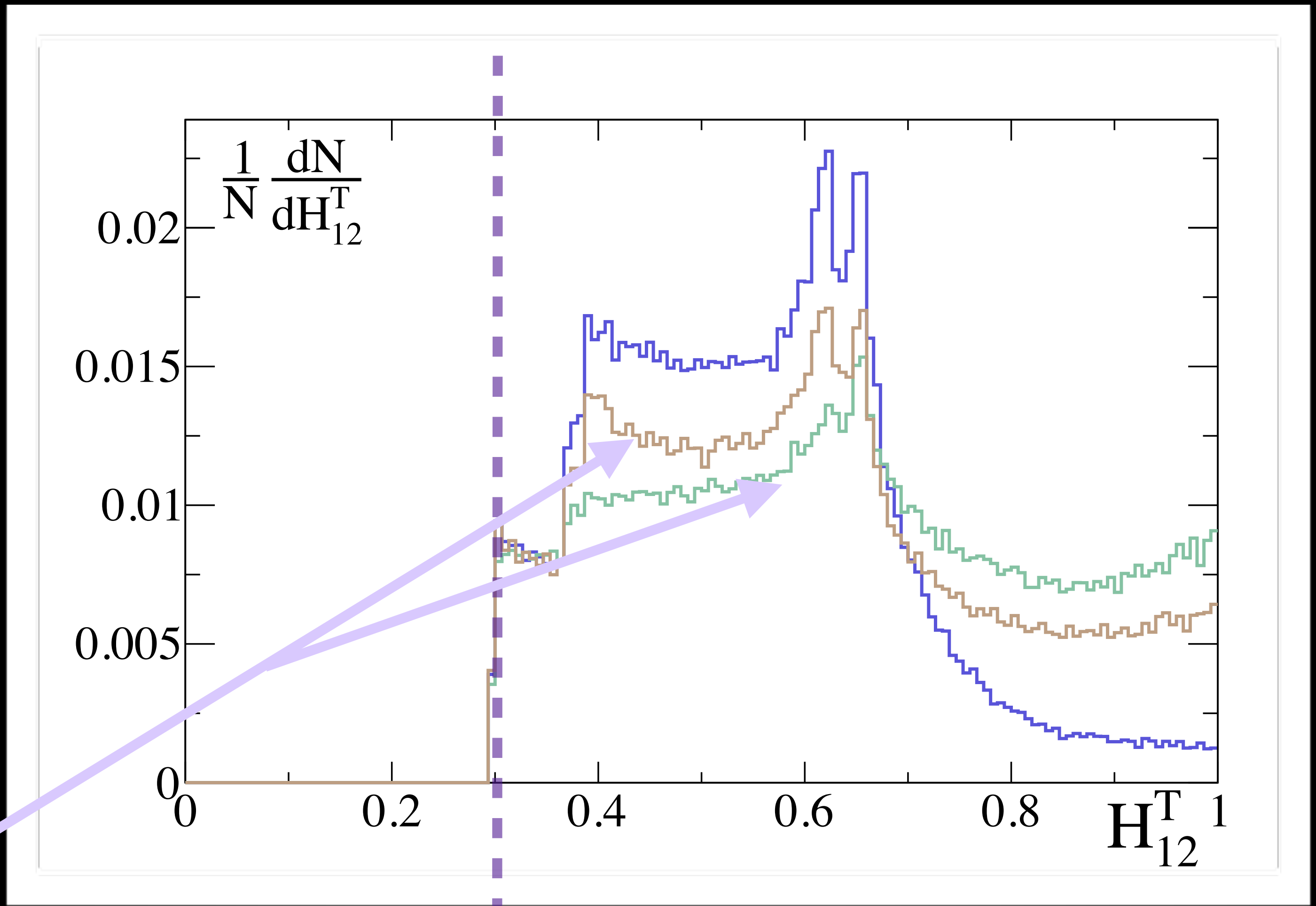
acceptance	WBF + 1 jet		QCD ZJJ		Top Pair		S / B
	% fail	XS (fb)	% fail	XS (fb)	% fail	XS (fb)	
		18.7		115000		17200	1/7070
minimal cuts + b veto	NA	3.92	NA	253	54.0	292	1/139
central jet cuts	13.9	3.32	31.8	157	66.1	85.4	1/73

$H_{12}^{T} > 0.7$ 1/57

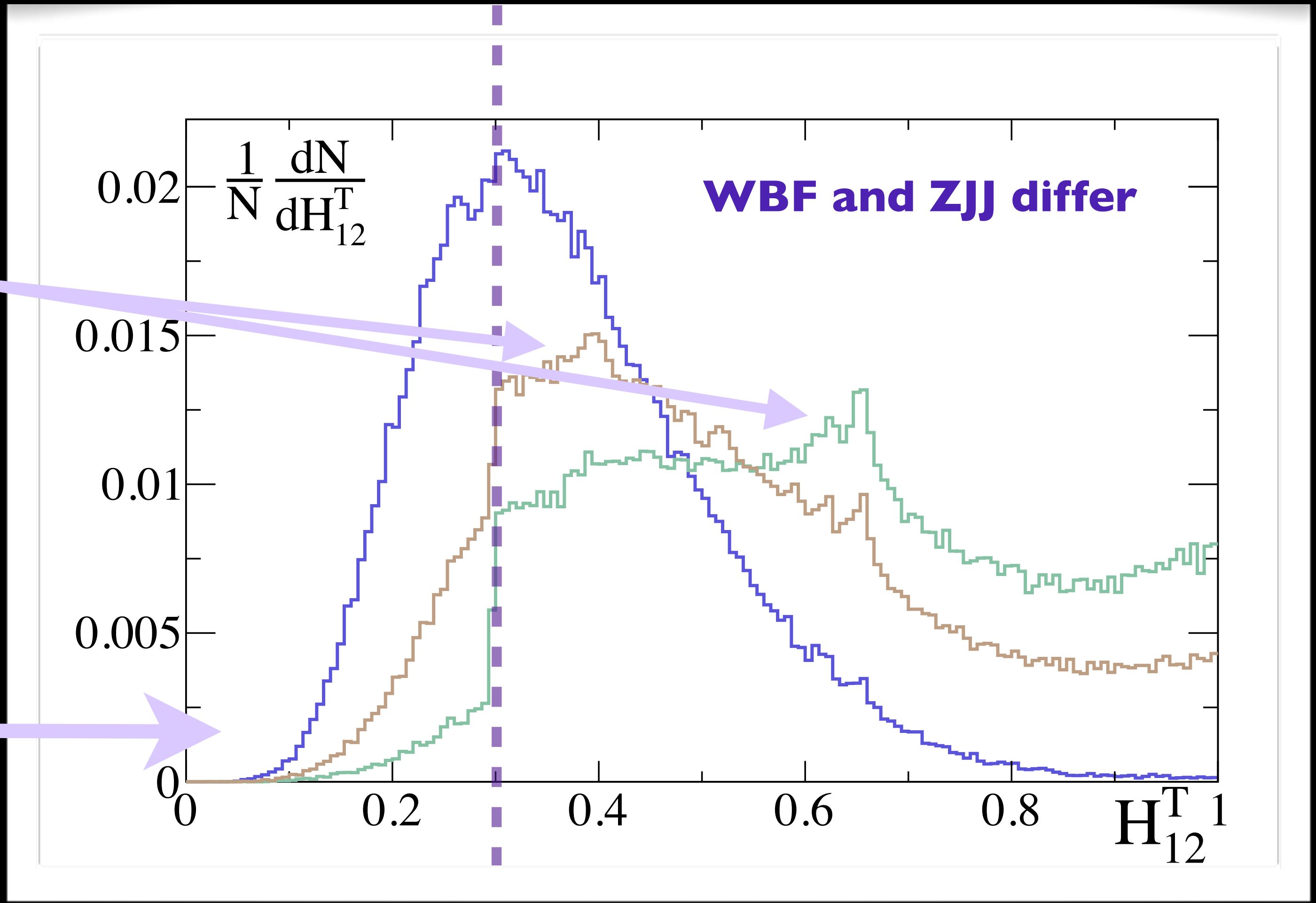
top pair background can be further supressed based on tagging jet correlations rephrased ito FWM

Inclusive FWM:

require *at least* 2 tagging jets satisfying minimal cuts



more power to discern WBF from ZJJ (3rd and higher jets have more drastically differing weights)



$H_\ell < 0.3$ **region populated**

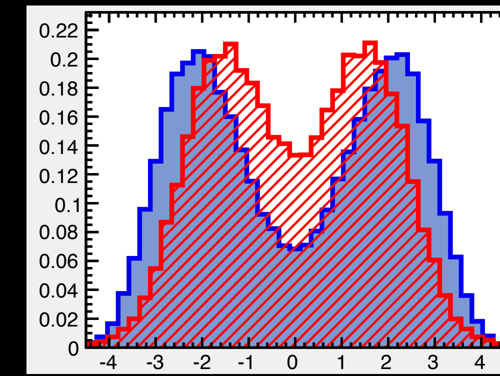
Classification Rule

A "classifier" is a rule for determining which class an instance of a set belongs to

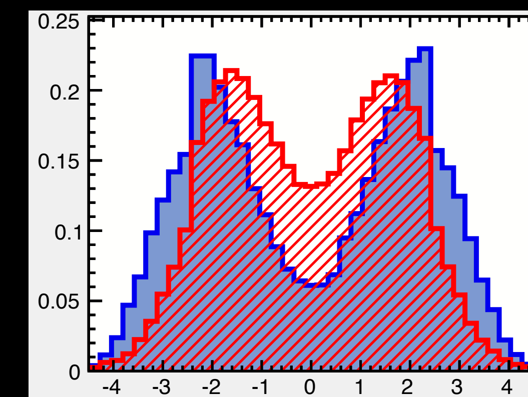
- class → sig/bkg
- instance → event
- set → data or MC sample

instance	y_{j_1}	y_{j_2}	Δy_{12}	m_{12}	class
event 1	2.79854	-1.33015	4.12869	264.056 GeV	S
event 2	1.5059	-1.09764	2.60354	156.285 GeV	B
⋮	⋮	⋮	⋮	⋮	⋮
event n	-1.10029	1.83929	2.93958	209.104 GeV	S

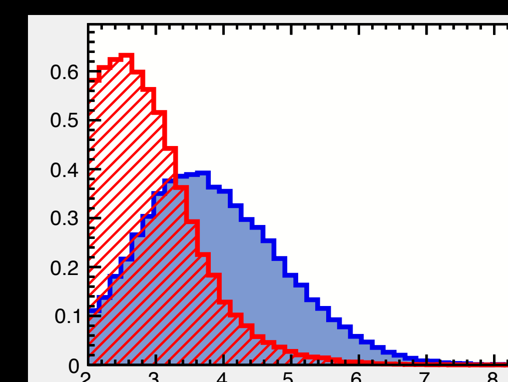
$y(\vec{x}) \in \mathbb{R}$ map all information of an event onto a real number - the "scalar output" of the classifier



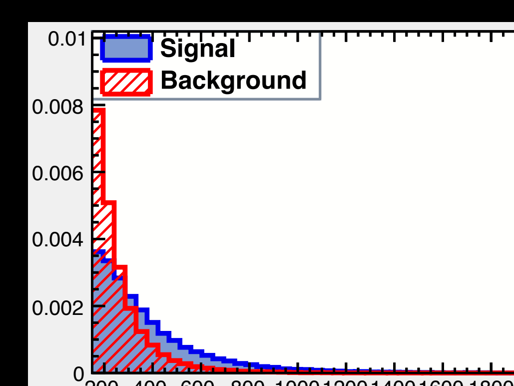
y_{j_1}



y_{j_2}



Δy_{12}



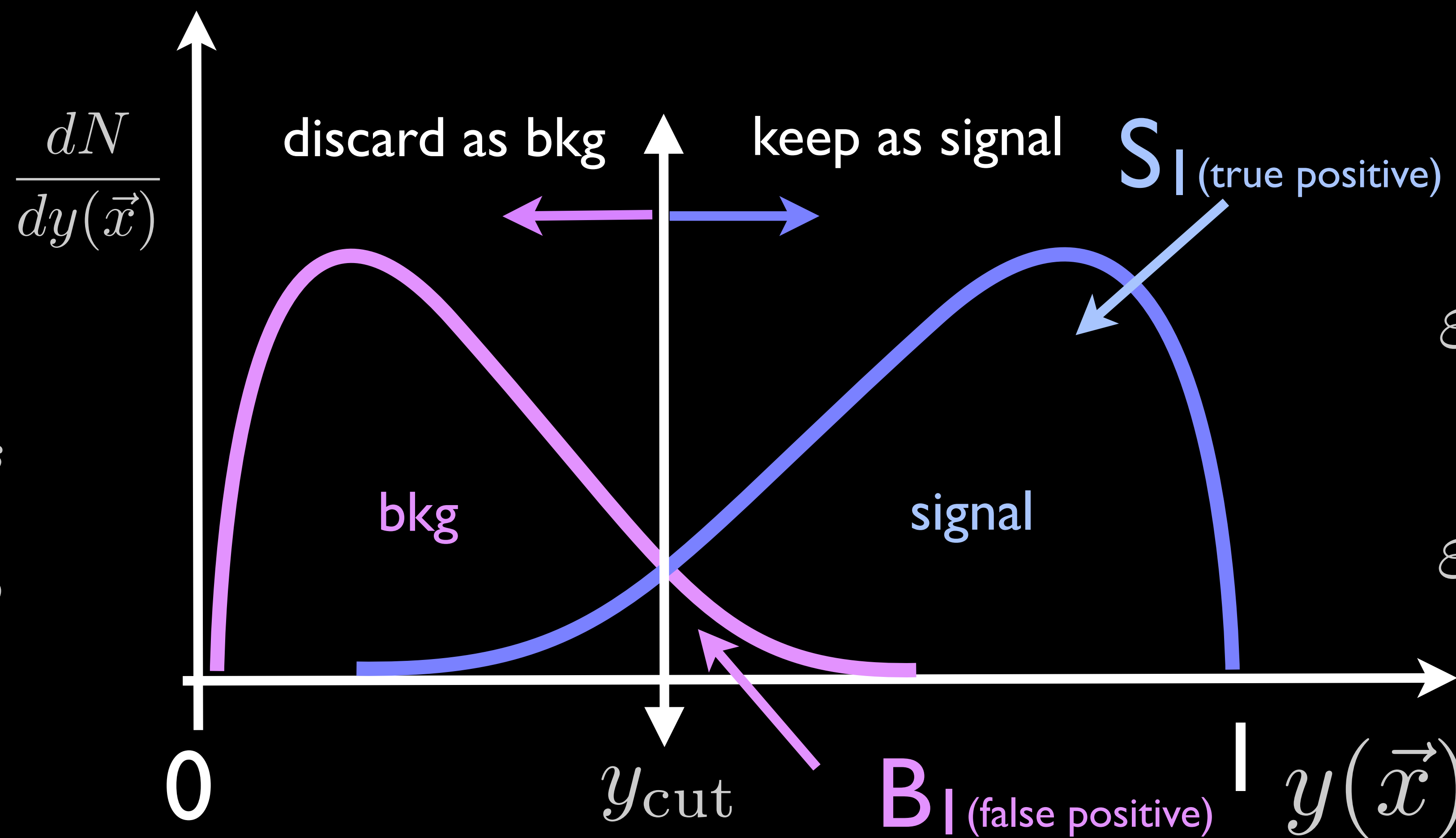
m_{12}

Classification Response and ROC Curves

rejection:
discard events

$$r_s = 1 - \varepsilon_s$$

$$r_b = 1 - \varepsilon_b$$



efficiency:
keep events

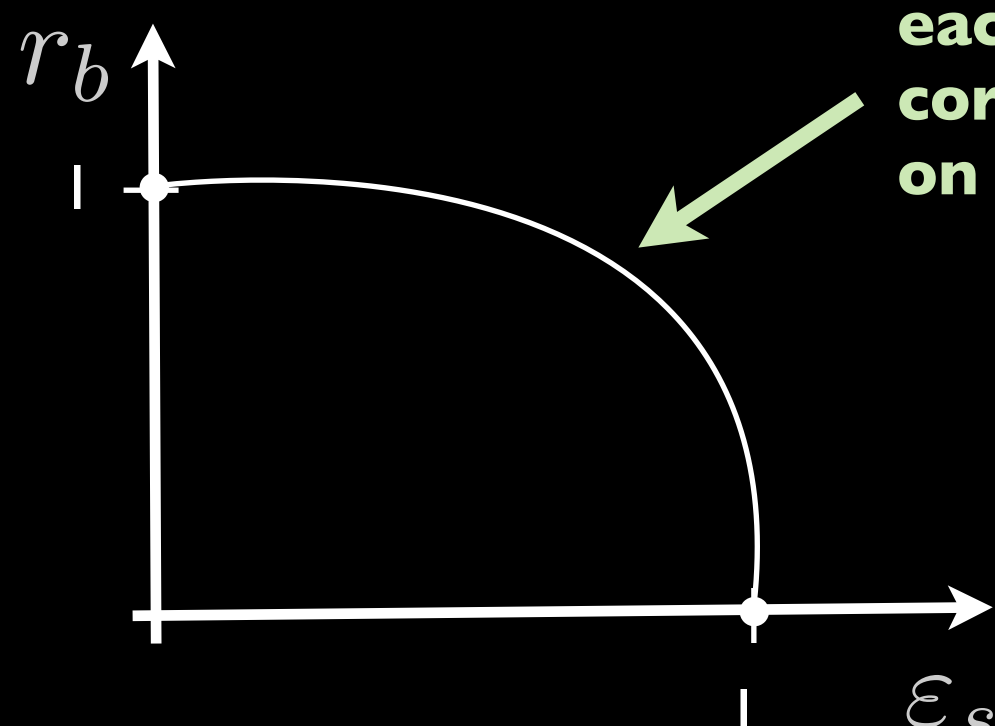
$$\varepsilon_s = \frac{S_1}{S_{\rm tot}}$$

$$\varepsilon_b = \frac{B_1}{B_{\rm tot}}$$

if $y_{\rm cut} = 1$

$$(\varepsilon_s, r_b) = (0, 1)$$

ROC curve
(Receiver Operating
Characteristic)



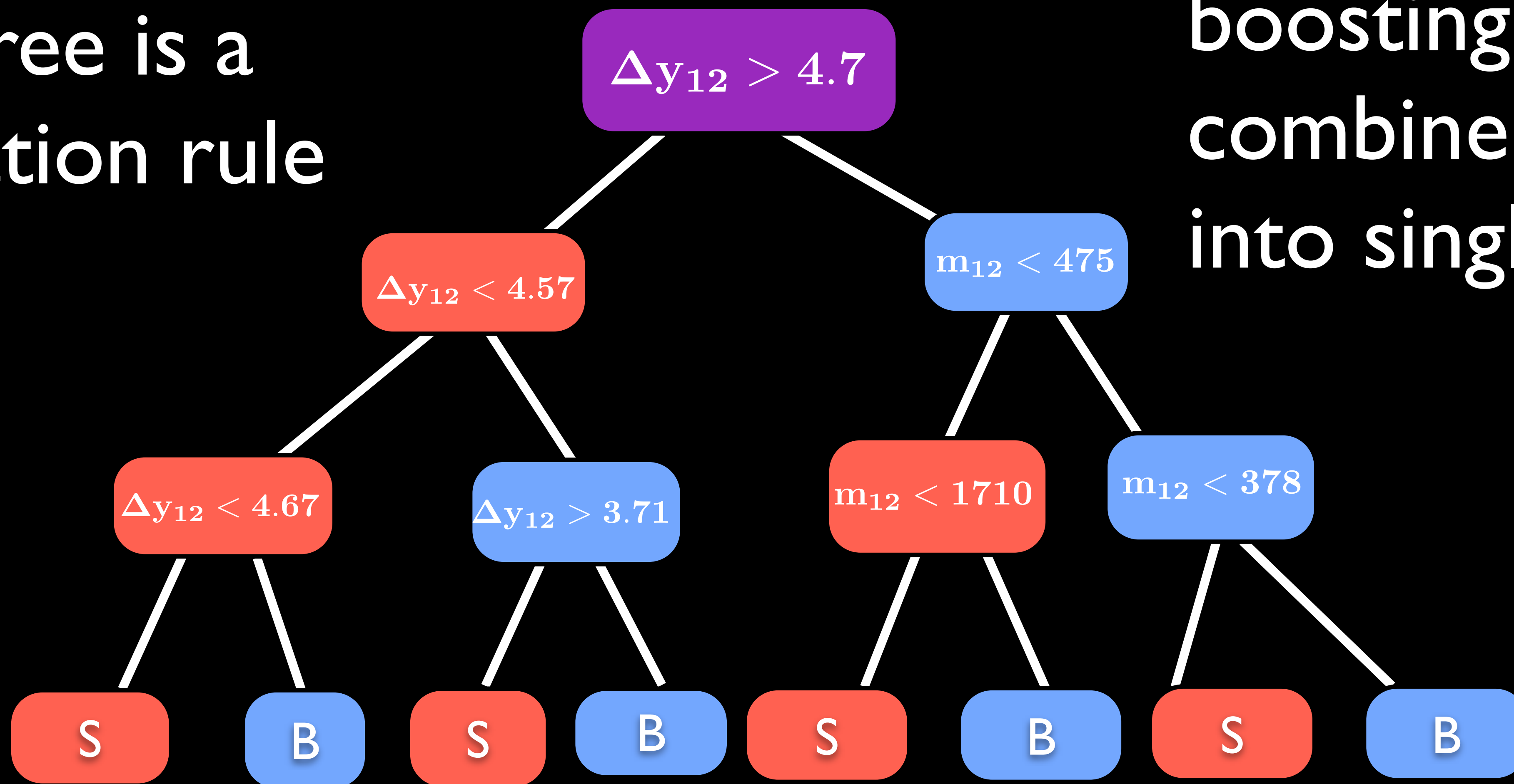
each point on curve corresponds to a cut on classifier response

if $y_{\rm cut} = 0$

$$(\varepsilon_s, r_b) = (1, 0)$$

Boosted Decision Trees

each tree is a classification rule



boosting: combine trees into single rule

Adaptive Boost Algorithm:

$$y(\vec{x}) = \frac{1}{N_{\mathrm{boost}}} \sum_{i}^{N_{\mathrm{boost}}} \ln(\alpha_i) h_i(\vec{x})$$

events misclassified are reweighted, another tree is built, misclassification rate is updated, event is reweighted, etc...

$$h_i(\vec{x}) = +1 \text{ (sig)}, \; -1 \text{ (bkg)}$$

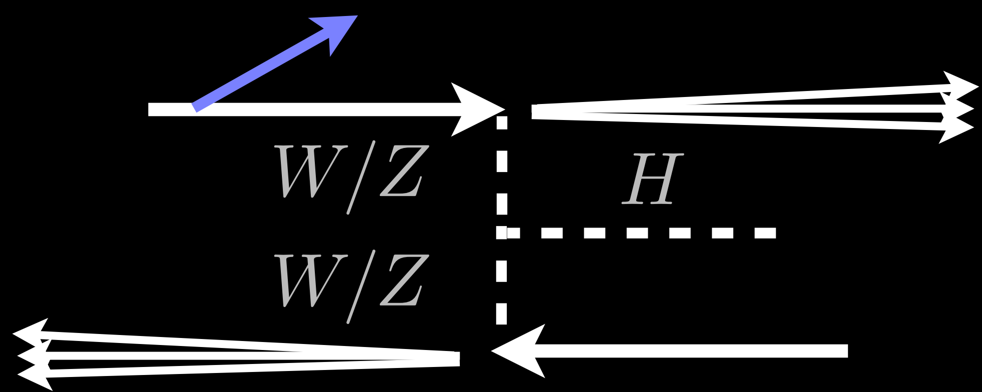
$$\alpha_i = \frac{1 - err_i}{err_i}$$

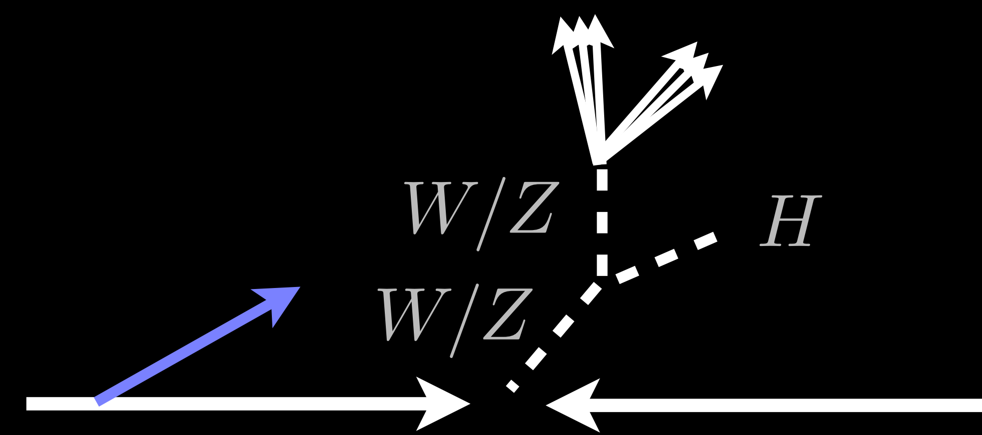
$$err_i = \text{ misclassification rate}$$

Analysis for H --> diphoton

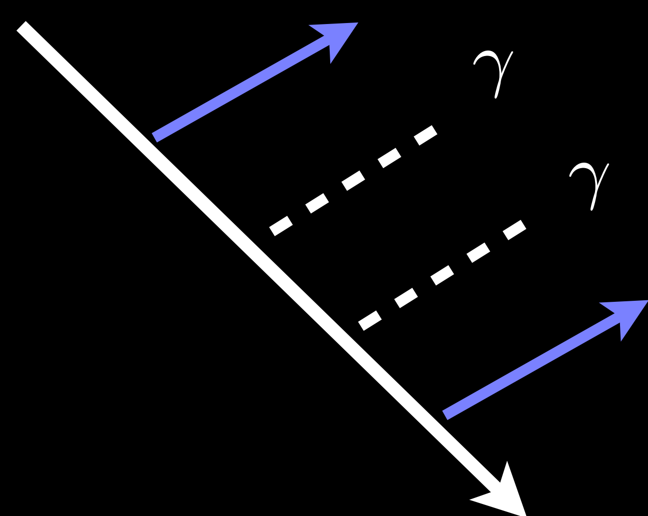
(process + hard jet) x PS with CKKW using SHERPA

signal VBF + 1 matrix element level jet





background diphoton + 2 matrix element level jets



Fastjet anti-kT algorithm with R = 0.4, 8TeV

BDT Analysis with only Tagging Jet Correlations

use FWM after applying acceptance criteria for jets:

$$p_{Tj} > 25 \text{ GeV} \quad \text{for} \quad |y_j| < 2.4$$

$$p_{Tj} > 30 \text{ GeV} \quad \text{for} \quad 2.4 \leq |y_j| < 4.5$$

$$|\Delta y_{j_1 j_2}| \geq 2 \quad \text{and} \quad m_{j_1 j_2} > 150 \text{ GeV}$$

compare FWM with tagging jet correlations used by ATLAS

$$\{m_{j_1 j_2} \,,\, y_{j_1} \,,\, y_{j_2} \,,\, \Delta y_{j_1 j_2}\}$$

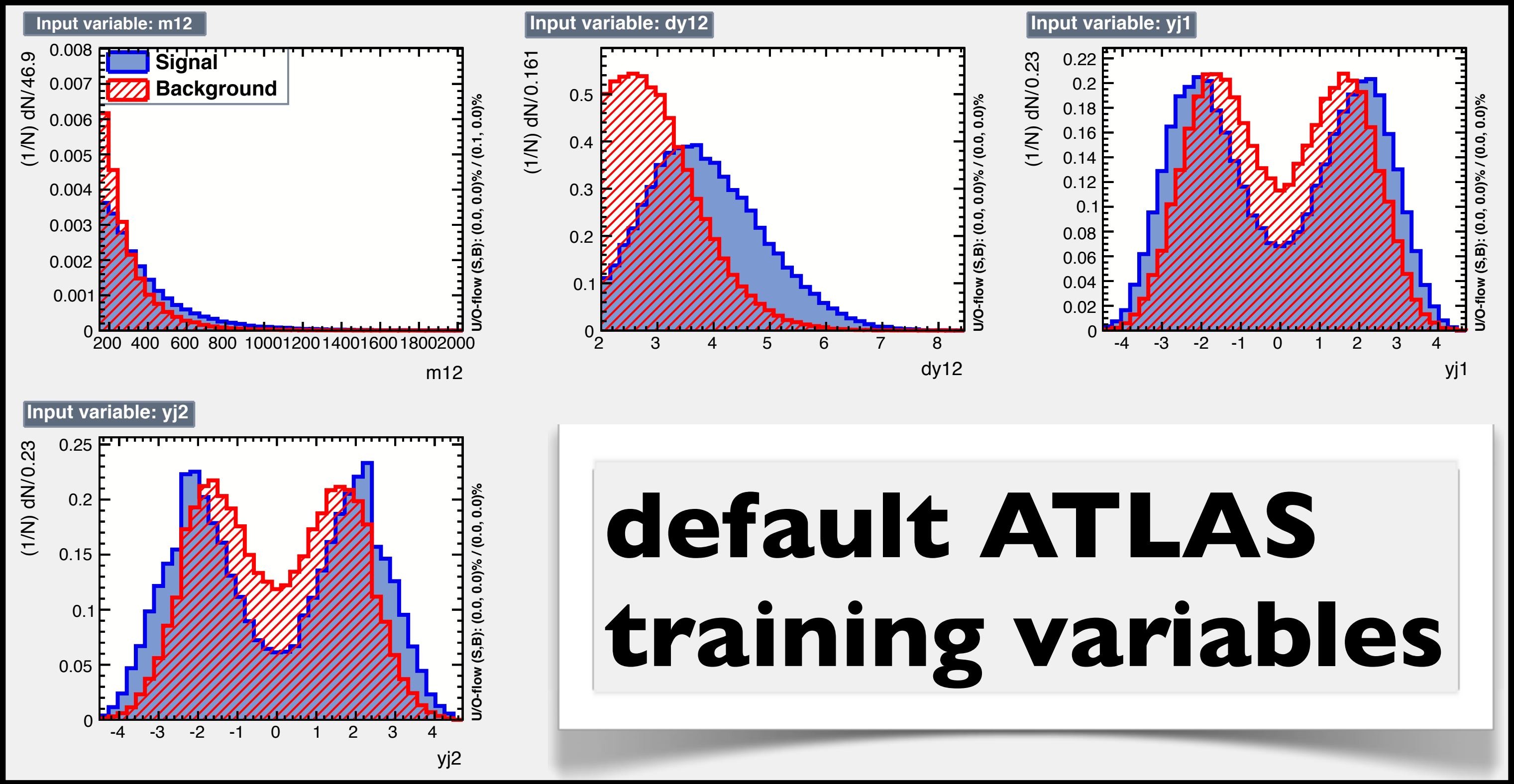
Decision Tree Settings[1]:

N_{train}, N_{test} = 100K, 50K

N_{trees}, N_{layers} = 400, 3

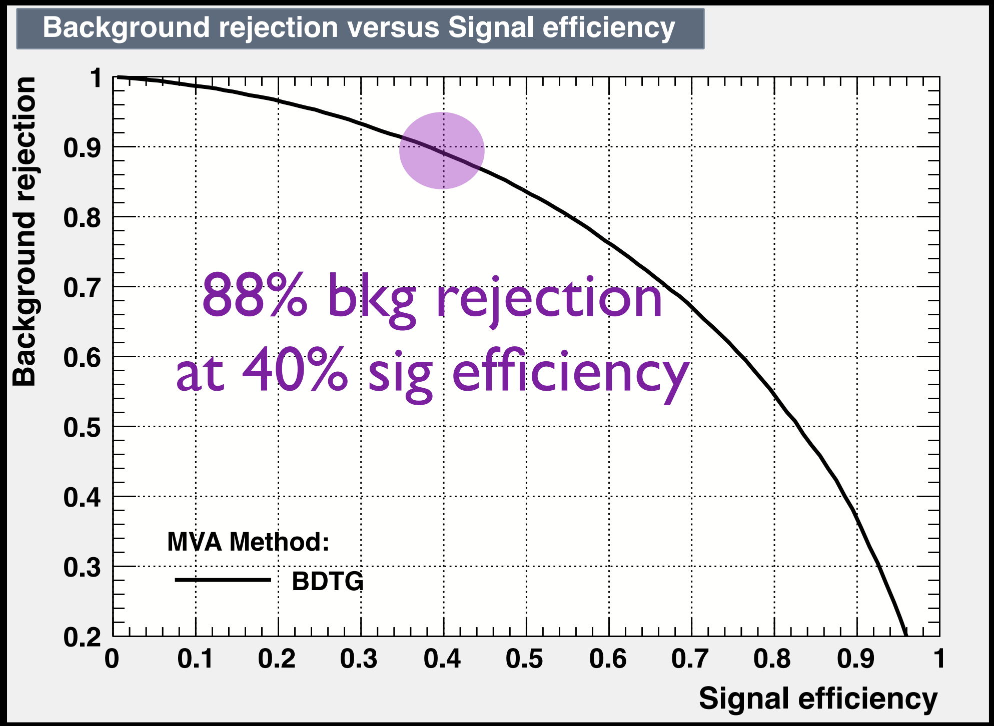
[1] Hoecker et.al. , Toolkit for Multivariate Analysis , http://tmva.sourceforge.net

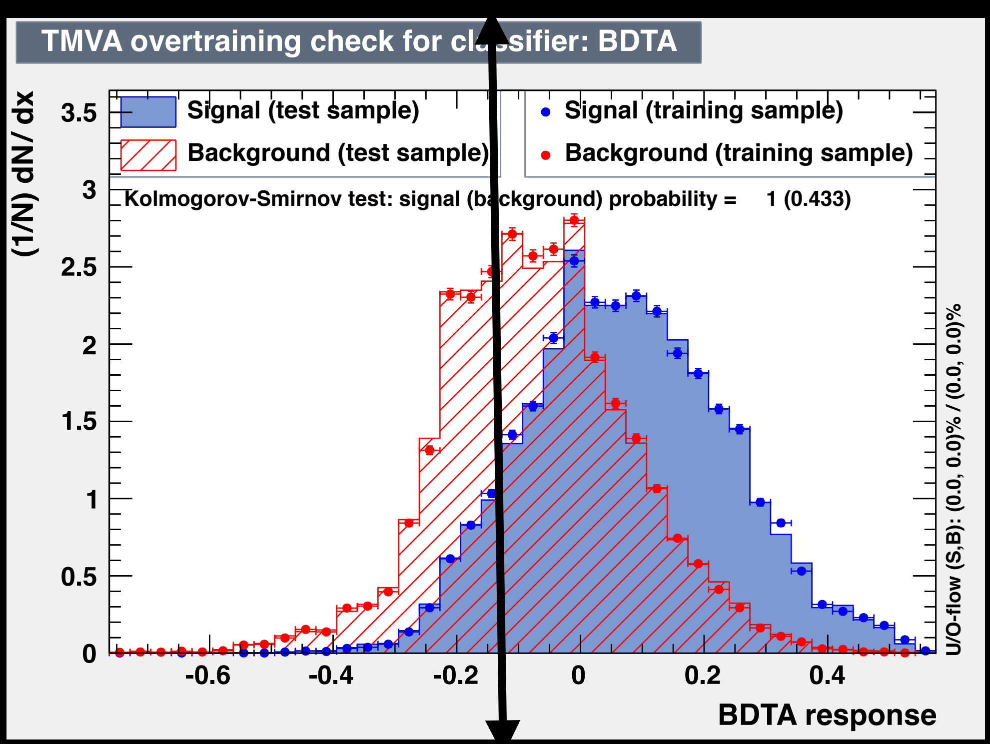
Results of BDT Analysis Including FWM



$$\frac{\mathrm{S}}{\sqrt{\mathrm{S}+\mathrm{B}}} = 198.7$$

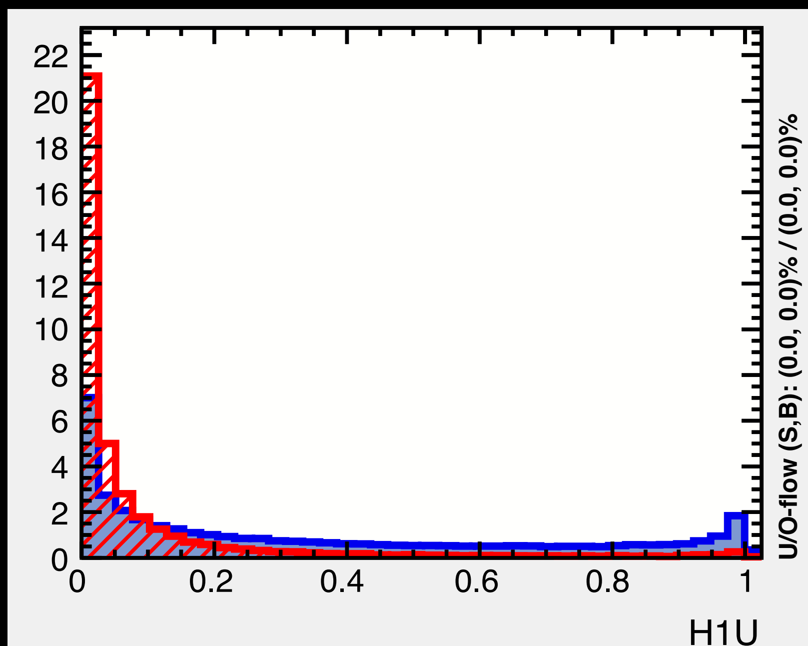
for cut at $y = -0.14$

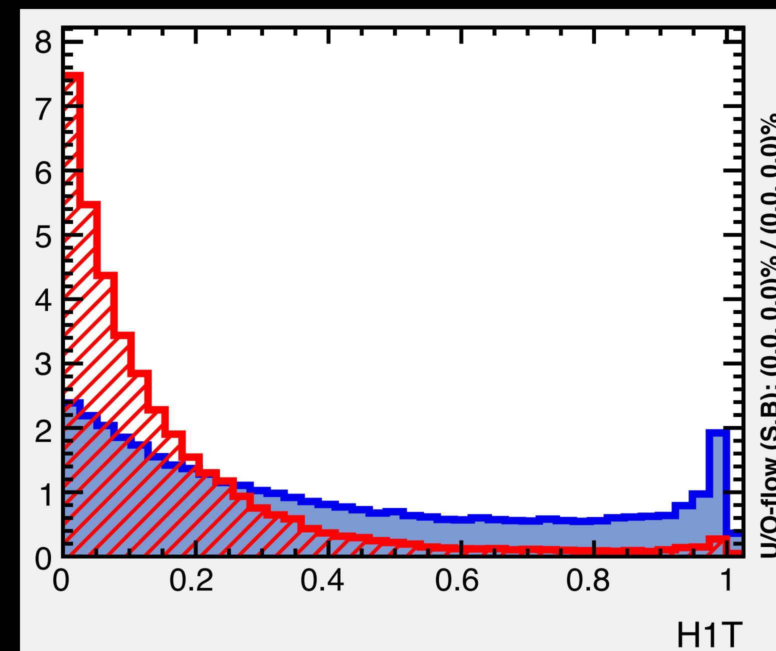




Results of BDT Analysis Including FWM[1]

in addition to default, train with: $H_\ell^{x,\phi} = \sum_{i,j=1}^{N} W_{ij}^x \; P_\ell(\cos \Delta\phi_{ij})$





$$H_1^{U,\phi} = \frac{1}{2} + \frac{1}{4} \cos \Delta\phi_{12}$$

$$H_1^{T,\phi} = \frac{p_{T1}^2 + p_{T2}^2}{p_{T\mathrm{tot}}^2} + \frac{p_{T1} p_{T2}}{p_{T\mathrm{tot}}^2} \cos \Delta\phi_{12}$$

	B rejection	$\frac{S}{\sqrt{S+B}}$	improvement
ATLAS default	88.7%	198.7 (-0.14)	
$H_1^{T,\phi}$, $H_1^{U,\phi}$	95.2%	209.166 (-0.07)	5.3%
$H_1^{T,\phi}$	94.9%	206.703 (-0.08)	4.0%
$H_1^{U,\phi}$	95.2%	208.821 (-0.08)	5.1%
$\cos \Delta\phi_{12}$	95.2%	208.821 (-0.08)	5.1%

[1] Bernaciak, Mellado, Plehn, Ruan, Schichtel, in preparation

Results of BDT Analysis Including FWM[1]

improvement with redefinition of FWM:

$$H_\ell^{x,\phi} = \sum_{i,j=1}^{N} W_{ij}^{x} \, P_\ell(\cos\Delta\phi_{ij})$$

	B rejection	$\frac{S}{\sqrt{S+B}}$	improvement
ATLAS default	88.7%	198.7 (-0.14)	
$H_1^{T,\phi} \to H_{20}^{T,\phi},\ H_1^{U,\phi} \to H_{20}^{U,\phi}$	95.0%	208.901 (-0.07)	5.1%
$H_1^{T,\phi},\ H_3^{T,\phi},\ H_1^{U,\phi},\ H_3^{U,\phi}$	95.3%	209.115 (-0.08)	5.3%
$H_1^{T,\phi},\ H_2^{T,\phi},\ H_2^{U,\phi},\ H_2^{U,\phi}$	95.2%	209.132 (-0.08)	5.3%
$H_1^{T,\phi},\ H_1^{U,\phi}$	95.2%	209.166 (-0.07)	5.3%
$H_1^{T,\phi}$	94.9%	206.703 (-0.08)	4.0%
$H_1^{U,\phi}$	95.2%	208.821 (-0.08)	5.1%
$\cos\Delta\phi_{12},\ W_{12}^{T}$	95.3%	209.299 (-0.08)	5.3%
$\cos\Delta\phi_{12}$	95.2%	208.821 (-0.08)	5.1%

redefinition of FWM offer modest improvement over ATLAS default variables

[1] Bernaciak, Mellado, Plehn, Ruan, Schichtel, in preparation

Conclusions - Future Work

✦FWM suitable for both cut-based and decision tree analysis - offer consistent 5% improvement for azimuthal angle definition

✦combinations of U and T weighted moments are better than T alone, U may be sufficient alone

✦total angle moments - offer 1% improvement - need to understand why

✦the FWM are an interesting addition to the variables currently used in Higgs analyses

Work Underway

✦compare with Neural Network MVA

✦incorporate 3rd jet and its scale uncertainty into this analysis

✦can moments be used as a modified jet veto?

BACKUP SLIDES

Gauge Invariance, Briefly

a key aspect of any realistic field theory description of matter is *gauge invariance*

↓

Lagrangian unchanged under a local change of coordinate system

↓ constructed from *gauge fields*

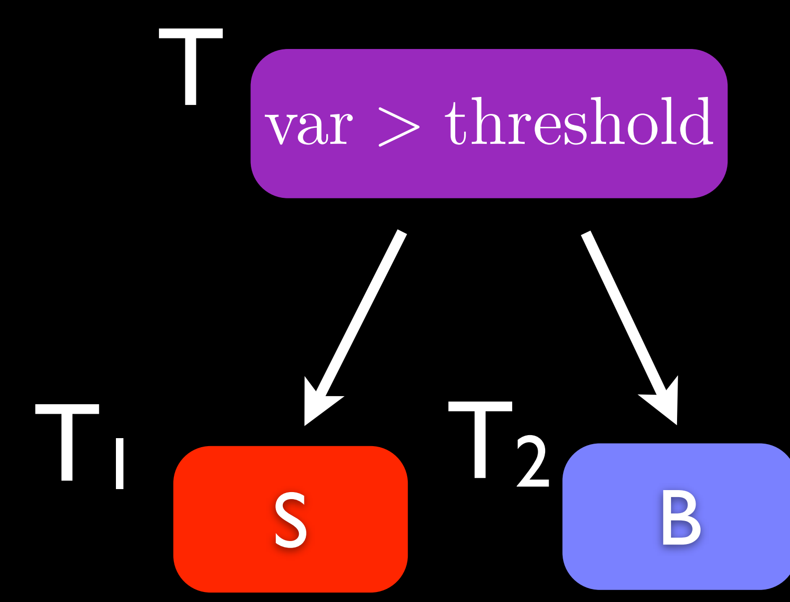
↓ spatial translation or rotation, internal field transformation

$$\psi \rightarrow e^{i\theta}\psi \qquad \psi \rightarrow e^{i\theta(x^\mu)}\psi$$

$$\psi(x^\mu) \rightarrow \psi(\Lambda^{\mu\nu}x_\nu + a^\mu)$$

Deciding Splitting Variables

avg. info needed to identify a class in T if it's partitioned into 2 subsets:



$$\mathrm{info}_X(\mathrm{T}) = \frac{|T_1|}{|T|}\mathrm{info}(\mathrm{T}_1) + \frac{|T_2|}{|T|}\mathrm{info}(\mathrm{T}_2)$$

information gain obtained by a particular test :

$$\mathrm{gain}(\mathrm{X}) = \mathrm{info}(\mathrm{T}) - \mathrm{info}_X(\mathrm{T})$$

variable 1

v_1

v_2

v_3

v_4

.

.

v_m

test 1: split variable 1 → $> v_2$ / $<= v_2$

test 2: split variable 1 → $> v_3$ / $<= v_3$

.

.

.

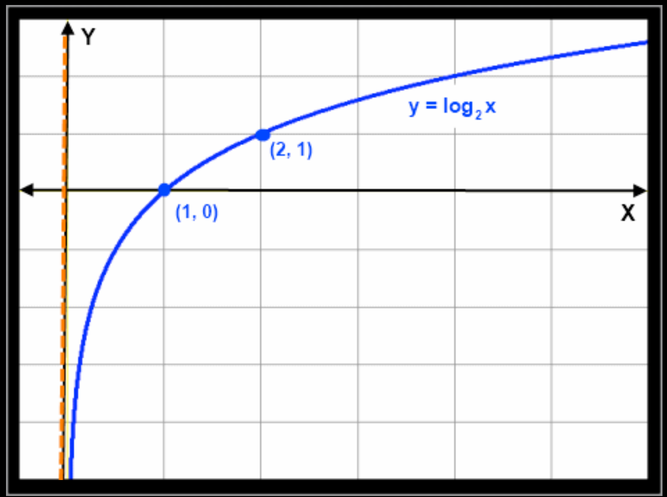
test m-1: split variable 1

repeat for all variables : test with largest gain ratio becomes root node ... repeat for subsequent nodes

Information Entropy

instance	y_{j_1}	y_{j_2}	Δy_{12}	m_{12}	class
event 1	2.79854	-1.33015	4.12869	264.056 GeV	S
event 2	1.5059	-1.09764	2.60354	156.285 GeV	B
⋮	⋮	⋮	⋮	⋮	⋮
event n	-1.10029	1.83929	2.93958	209.104 GeV	S

the set "T"



probability of finding an event belonging to S or B in the entire set T

$$P_{S,B} = \frac{N_{S,B}}{N_{\text{tot}}}$$

information entropy (general)

$$I_E = \log_2(P_{S,B}) \text{ bits}$$

information entropy of T

$$\text{info(T)} = -P_S \log_2(P_S) - P_B \log_2(P_B)$$

"avg. amount of info needed to identify the class of an event in T"

The Fox-Wolfram moments - brief history

an event shape observable describing correlations between four-momentum objects

e+ e- to jets

Fox, Wolfram Nucl. Phys. B 149 (1979) 413-496

Top Quark signal at Tevatron

Field, Kanev, Tayebnejad PRD 55, 9 (1997)

B meson decays at Belle:

Toru Iijima, hep-ex 0105005 (2001)

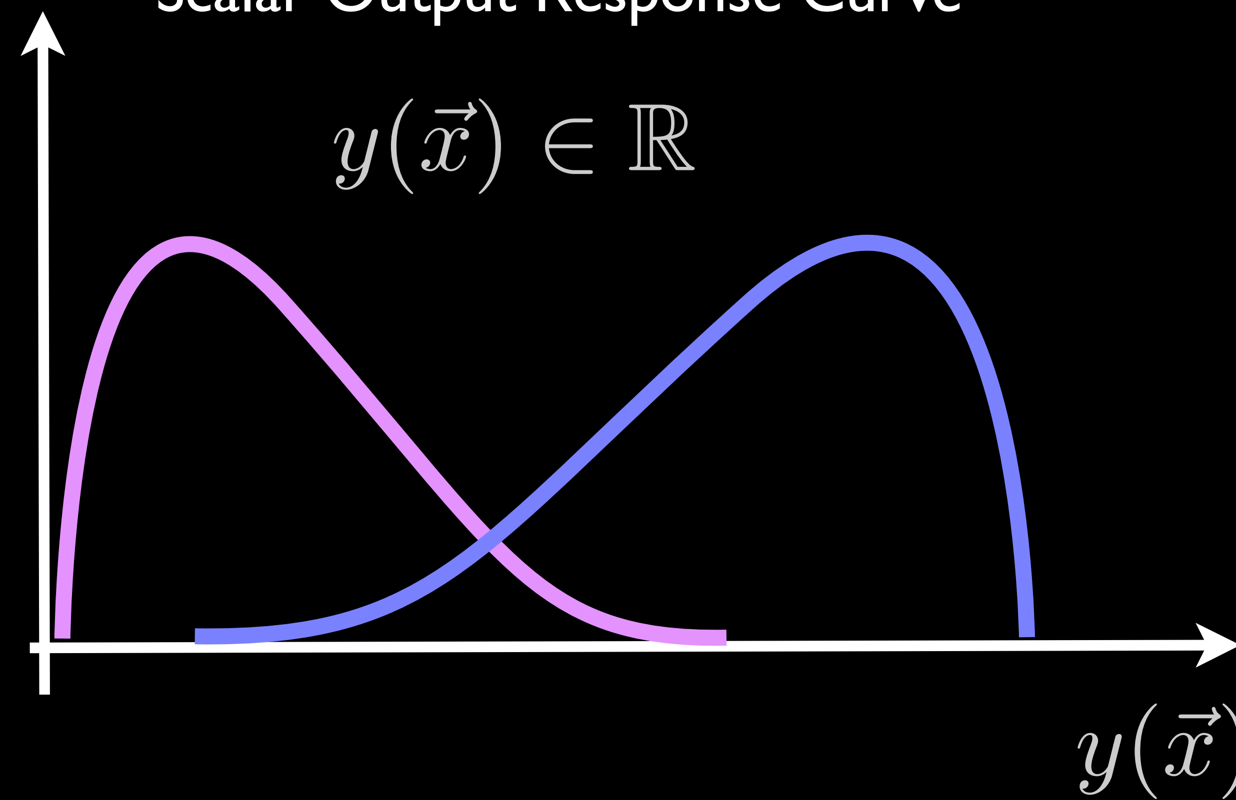
Higgs physics at the LHC: **WBF H tautau vs Z+2j and Top Pair**

C.B., Buschmann, Butter, Plehn PRD 87, 073014 (2013)

Classification Rules

y_{j_1}	y_{j_2}	Δy_{12}	m_{12}
2.79854	-1.33015	4.12869	264.056 GeV
1.5059	-1.09764	2.60354	156.285 GeV
⋮	⋮	⋮	⋮
-1.10029	1.83929	2.93958	209.104 GeV

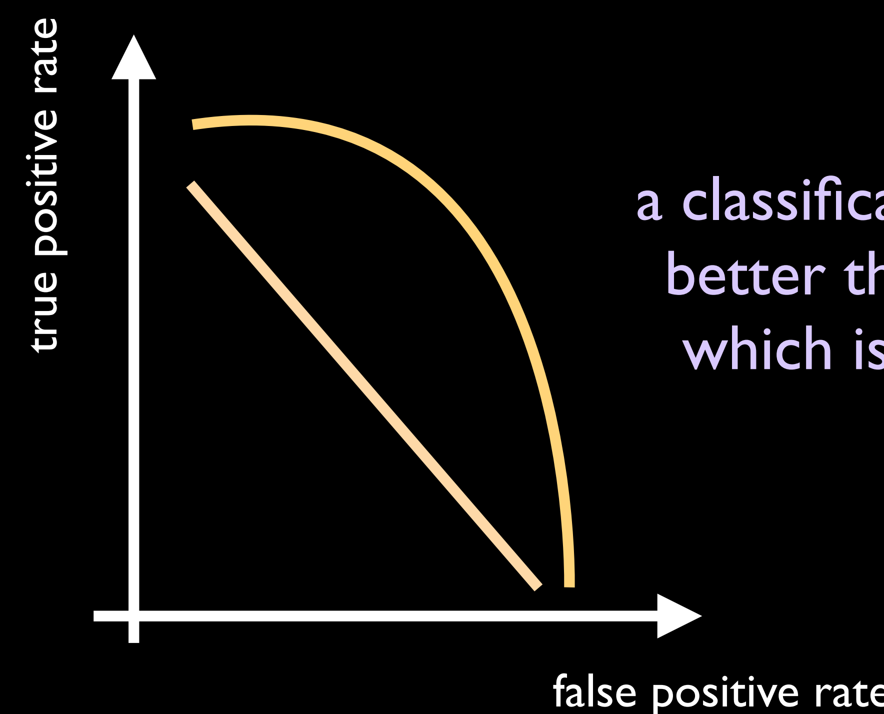
Scalar Output Response Curve



true class:

predicted class:	S	B
S	true positive	false positive
B	false negative	true negative

ROC curve (Receiver Operating Characteristic)



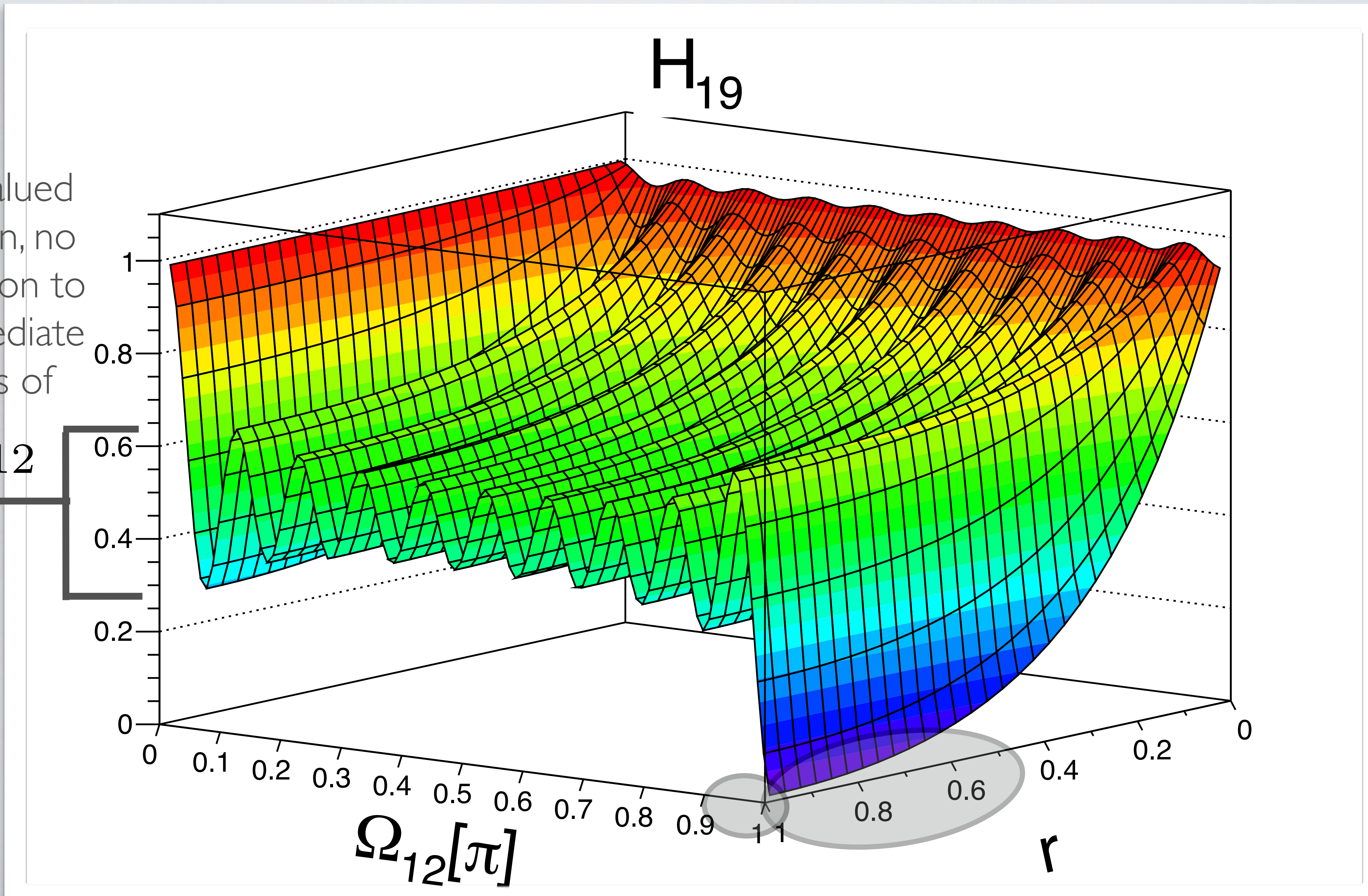
a classification rule seeks to do better than random guessing, which is correct 50% of the time

The Fox-Wolfram moments - 2 jet visualization

odd moments - best for discriminating back-to-back jets, higher moments resolve larger angular j_1 j_2 separation

$$H_\ell \to 0 \quad \text{for} \quad \Omega_{12} \to \pi$$

multivalued function, no resolution to intermediate values of Ω_{12}



The Fox-Wolfram moments - 2 jet visualization

even moments - symmetry of even function reduces discriminatory power

$$H_\ell \to 1 \quad \text{for} \quad \Omega_{12} \to 0 \quad \text{AND} \quad \Omega_{12} \to \pi$$