

Double Higgs Production at HL-LHC



K.C. Kong
University of Kansas

Based on: [Kim, Kong, Matchev, Park, PRL 2019](#)
[Kim, Kim, Kong, Matchev, Park, JHEP 2019](#)
[Huang, Kang, Kim, Kong, Pi, JHEP 2022](#)

HET Seminar, June 16, 2022



Why double Higgs (hh) ?

- Measurement of the Higgs self-couplings provides crucial input to confirm SM prediction, and is essential to understand EWSB.
- Triple Higgs coupling can be probed via double Higgs production at the LHC.
- Resonant / non-resonant double Higgs production is interesting, phenomenologically and experimentally.
 - Triple Higgs coupling is easily modified in many extensions of SM.
 - Double Higgs production is a guaranteed physics at HL-LHC with high impact (a new collider is needed to probe quartic coupling).
 - Double Higgs production provides measurement of the first non-trivial term (cubic term) in the Higgs potential.
 - Destructive interference between box and triangle diagrams in SM makes it difficult to probe the triple Higgs coupling.
 - The cubic coupling is sensitive at lower-energy bins where the backgrounds are large.
 - It is challenging experimentally.
 - It brings many different final states.

$$V = \frac{m_h^2}{2}h^2 + \kappa_3\frac{m_h^2}{2v}h^3 + \kappa_4\frac{m_h^2}{8v^2}h^4 + \dots$$

Why double Higgs (hh) ?

$$\sigma(hh)_{SM}^{NNLO} \simeq 40.7 \text{ fb} \quad (14 \text{ TeV})$$

higher branching ratios

cleaner final state

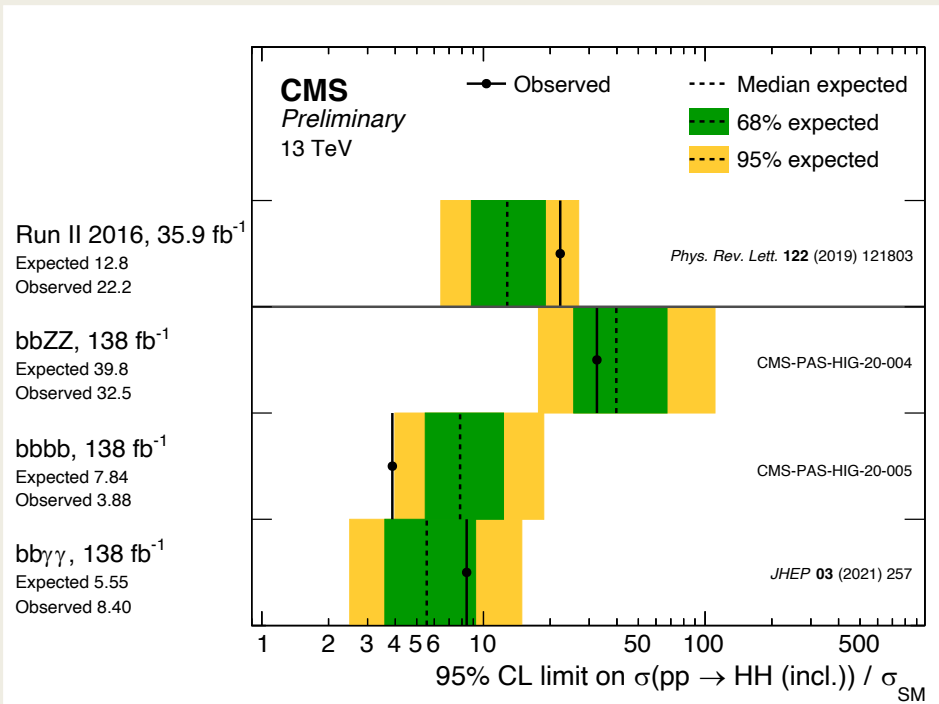
	bb	WW^*	$\tau\tau$	ZZ^*	$\gamma\gamma$
bb	33%				
WW^*	25%	4.6%			
$\tau\tau$	7.3%	2.7%	0.39%		
ZZ^*	3.1%	1.1%	0.33%	0.069%	
$\gamma\gamma$	0.26%	0.1%	0.028%	0.012%	0.0005%

probe the triple Higgs coupling.

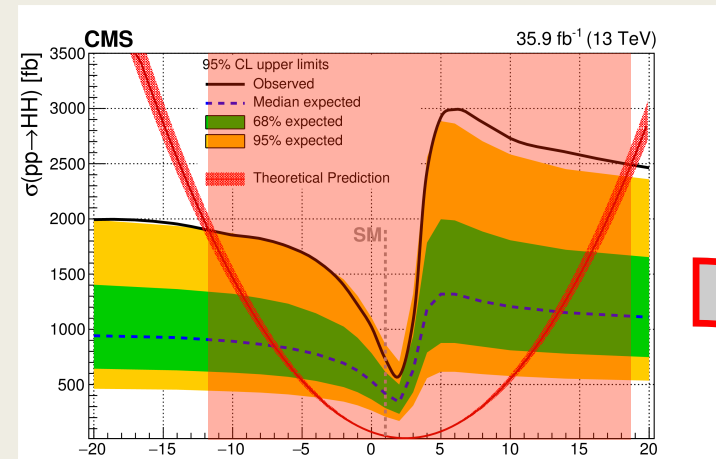
- The cubic coupling is sensitive at lower-energy bins where the backgrounds are large.
- It is challenging experimentally.
- It brings many different final states.

$$V = \frac{m_h^2}{2} h^2 + \kappa_3 \frac{m_h^2}{2v} h^3 + \kappa_4 \frac{m_h^2}{8v^2} h^4 + \dots$$

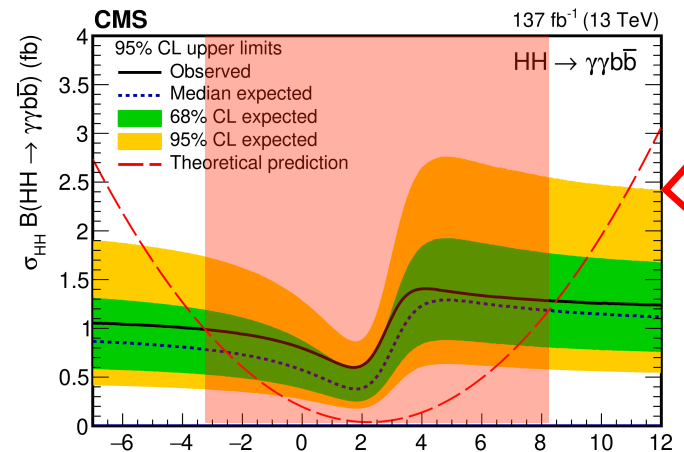
HH searches summary



Each of the expected limits of 4b and bbyγ with run 2 statistics are 2 times more stringent than the 2016 combination!



Allowed at 95% CL: $-11.8 < k_\lambda < 18.8$



Allowed at 95% CL: $-3.3 < k_\lambda < 8.5$

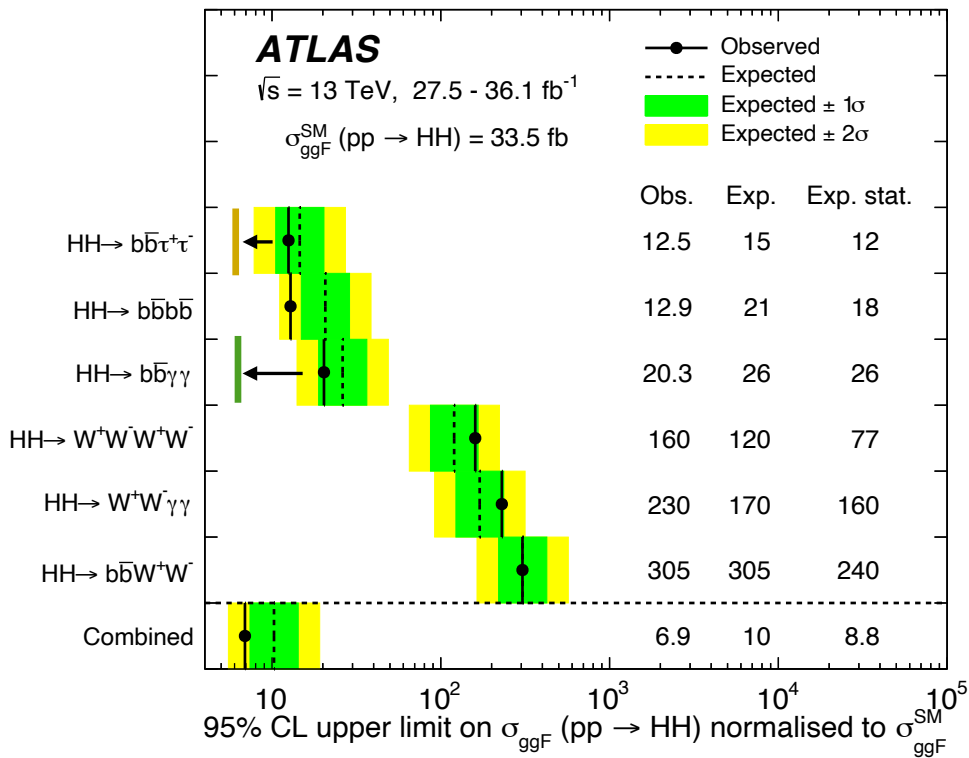
2016 comb vs bbyγ Full Run 2

Combination

$b\bar{b}l\nu l\nu$ final state : $\mathcal{L} = 139\text{fb}^{-1}$ [Phys. Lett. B 801 \(2020\) 135145](#)
 Combination: $\mathcal{L} = 36\text{fb}^{-1}$ [Phys. Lett. B 800 \(2020\) 135103](#)
 Combination: $\mathcal{L} = 139\text{fb}^{-1}$ [ATLAS-CONF-2021-052](#)



Combination done with most of the analyses with $\mathcal{L} = 36\text{fb}^{-1}$:



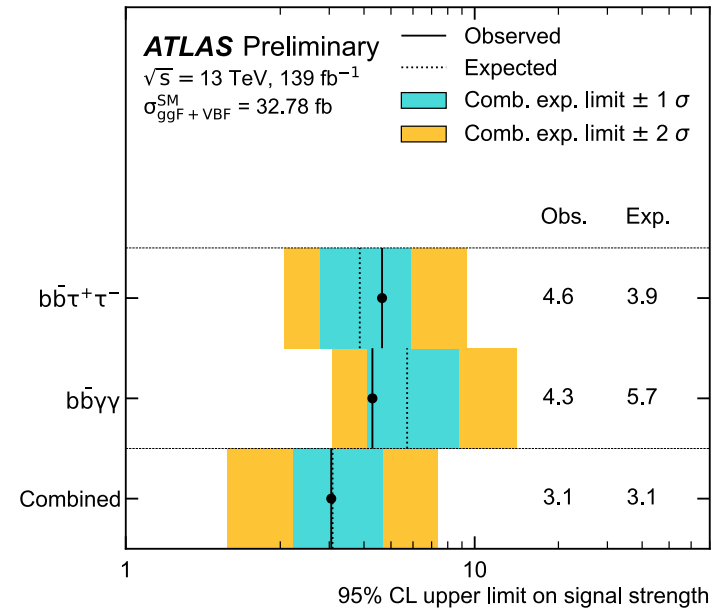
Additional results with $\mathcal{L} = 139\text{fb}^{-1}$:

$b\bar{b}\gamma\gamma$ and $b\bar{b}\tau\tau$ final states:

New full Run-2 combination with the two strongest channels.

$\mu_{HH}^{ggF+VBF}$ observed (expected) limit is 3.1 (3.1).

Best limit observed up to now!



15

Louis D'Eramo (NIU) - 19/05/2022 - Higgs self-coupling at ATLAS - LHCP2022

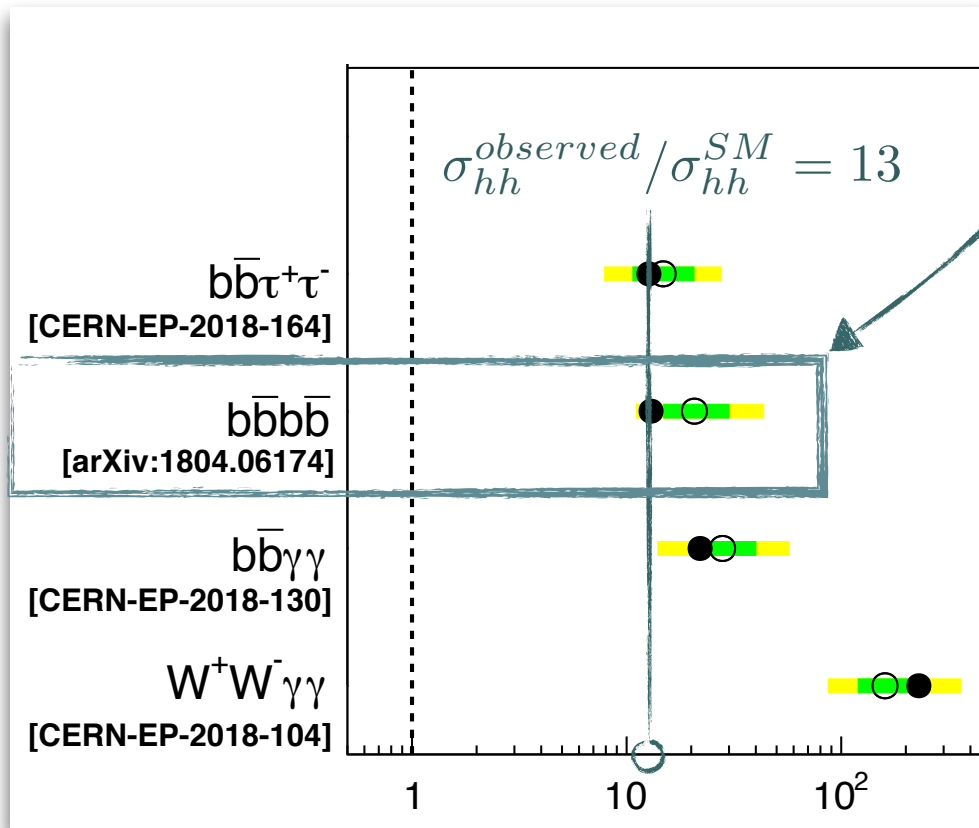
Talk by Louis D'Eramo at LHCP2022

Observed limit: $-1.0 \leq \kappa_3 \leq 6.6$

Expected limit: $-1.2 \leq \kappa_3 \leq 7.2$

Experimental status on c_3 @ LHC 13 TeV

27.5 – 36.1 fb⁻¹ (13 TeV)



- The $bbbb$ channel is significantly improved!

- Using an improved b -tagging algorithm ($MV2c10$)

$$\epsilon_b = 70 \%$$

$$\epsilon_{c \rightarrow b} = 8.3 \sim 14.1 \%$$

$$\epsilon_{j \rightarrow b} = 0.26 \sim 0.83 \%$$

95% CL on $\sigma_{hh} / \sigma_{hh}^{SM}$

ATLAS, PRL 117 (2016) 079901

ATLAS, PRL 117 (2016) 012001

Decays

$$\sigma(hh)_{SM}^{NNLO} \simeq 40.7 \text{ fb} \quad (14 \text{ TeV})$$

higher branching ratios

	<i>bb</i>	<i>WW*</i>	$\tau\tau$	<i>ZZ*</i>	$\gamma\gamma$
<i>bb</i>	33%				
<i>WW*</i>	25%	4.6%			
$\tau\tau$	7.3%	2.7%	0.39%		
<i>ZZ*</i>	3.1%	1.1%	0.33%	0.069%	
$\gamma\gamma$	0.26%	0.1%	0.028%	0.012%	0.0005%

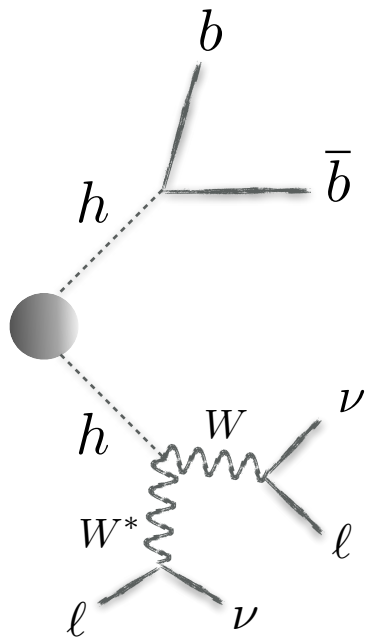
cleaner final state

1902.00134	Statistical-only		Statistical + Systematic	
	ATLAS	CMS	ATLAS	CMS
$HH \rightarrow bbbb$	1.4	1.2	0.61	0.95
$HH \rightarrow b\bar{b}\tau\tau$	2.5	1.6	2.1	1.4
$HH \rightarrow b\bar{b}\gamma\gamma$	2.1	1.8	2.0	1.8
$HH \rightarrow b\bar{b}VV (ll\nu\nu)$	-	0.59	-	0.56
$HH \rightarrow bbZZ (4l)$	-	0.37	-	0.37
combined	3.5	2.8	3.0	2.6
	Combined 4.5		Combined 4.0	

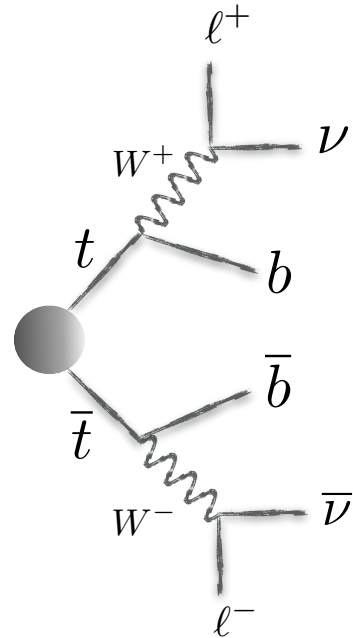
4 σ expected for ATLAS+CMS!

- Measurements of the triple Higgs coupling is challenging due to a small $\sigma(hh)$ and large backgrounds.
- No single channel is expected to reach 3 sigma at HL-LHC.
- The combination of various channels is crucial. **$bbWW$ dilepton channel has good potential for further improvement. (Focus of this talk)**

Previous (theory) study on $hh \rightarrow bbWW^*$



Signal



Major background

$$\sigma_{bknd} \sim 10^5 \sigma_{hh}$$

- $hh \rightarrow bbWW^*$ channel suffers from the large $t\bar{t}$ background.
- Most studies report that the sensitivity of signal in the di-leptonic channel is very poor.

CMS-FTR-15-002-PAS

Adhikary, Banerjee, Barman, Bhattacharjee, Niyogi 2017

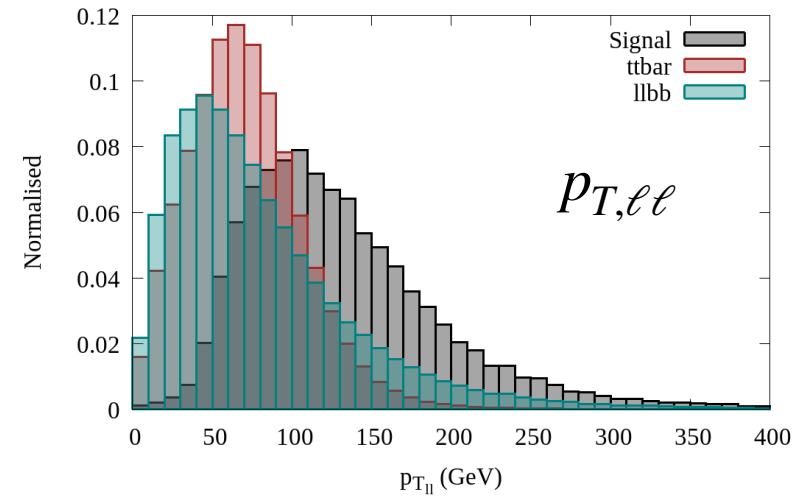
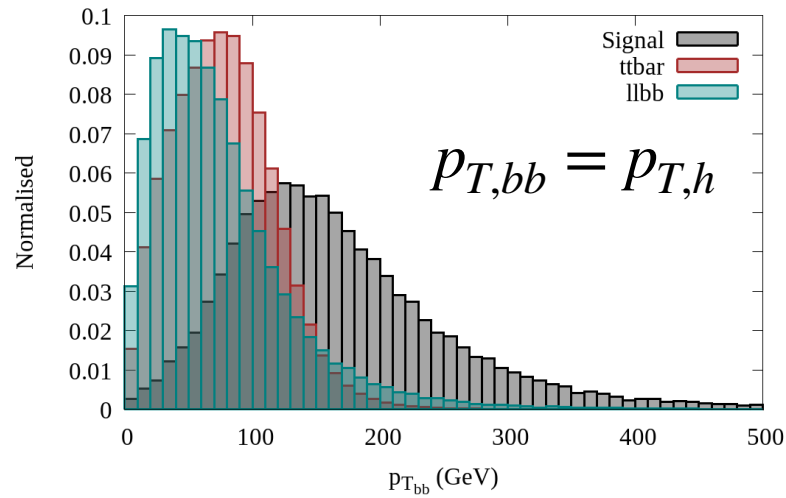
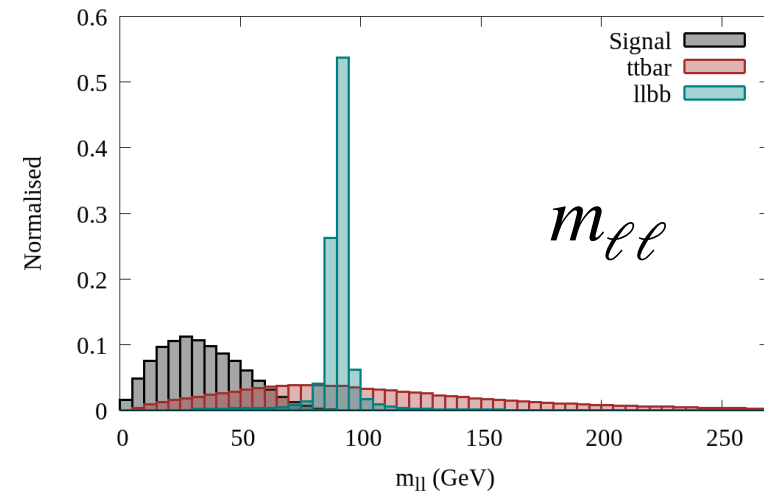
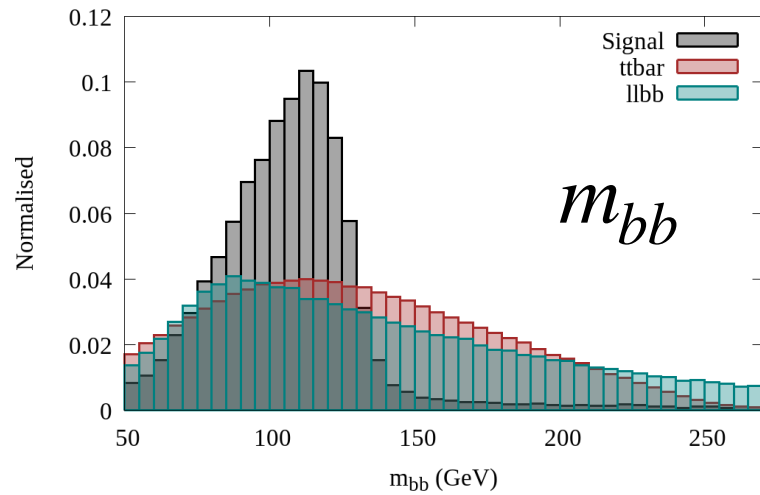
- The situation in the semi-leptonic mode is even worse.

Dolan, Englert, Spannowsky 2012

Adhikary, Banerjee, Barman, Bhattacharjee, Niyogi 2017

cf) Papaefstathiou, Yang, Zurita 2012

$hh \rightarrow bbWW^*$: dilepton channel



10 variables: $p_{T,\ell_{1/2}}$, \cancel{E}_T , m_{ll} , m_{bb} , ΔR_{ll} , ΔR_{bb} , $p_{T,bb}$, $p_{T,ll}$, $\Delta\phi_{bb\ell\ell}$

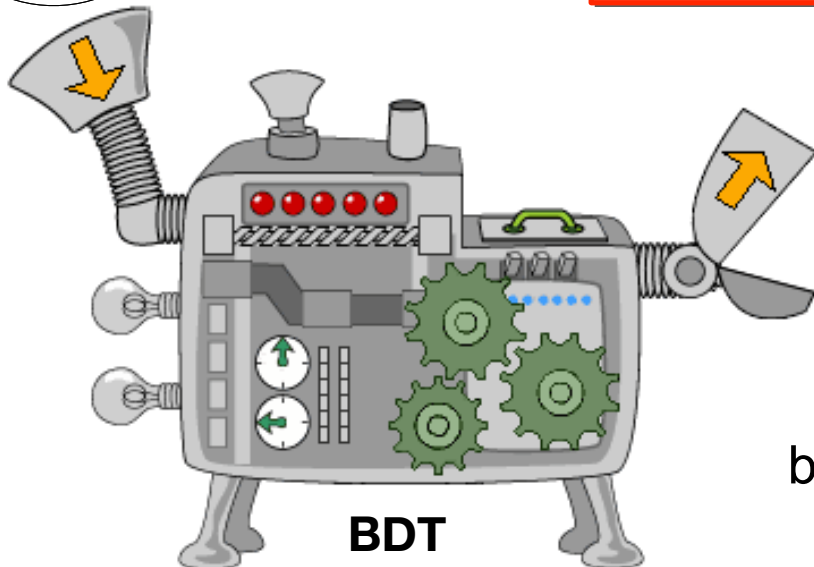
$hh \rightarrow bbWW^*$: dilepton channel

HL-LHC, 14 TeV, $L=3 \text{ ab}^{-1}$

Adhikary, Banerjee, Barman, Bhattacharjee, Niyogi JHEP 2017

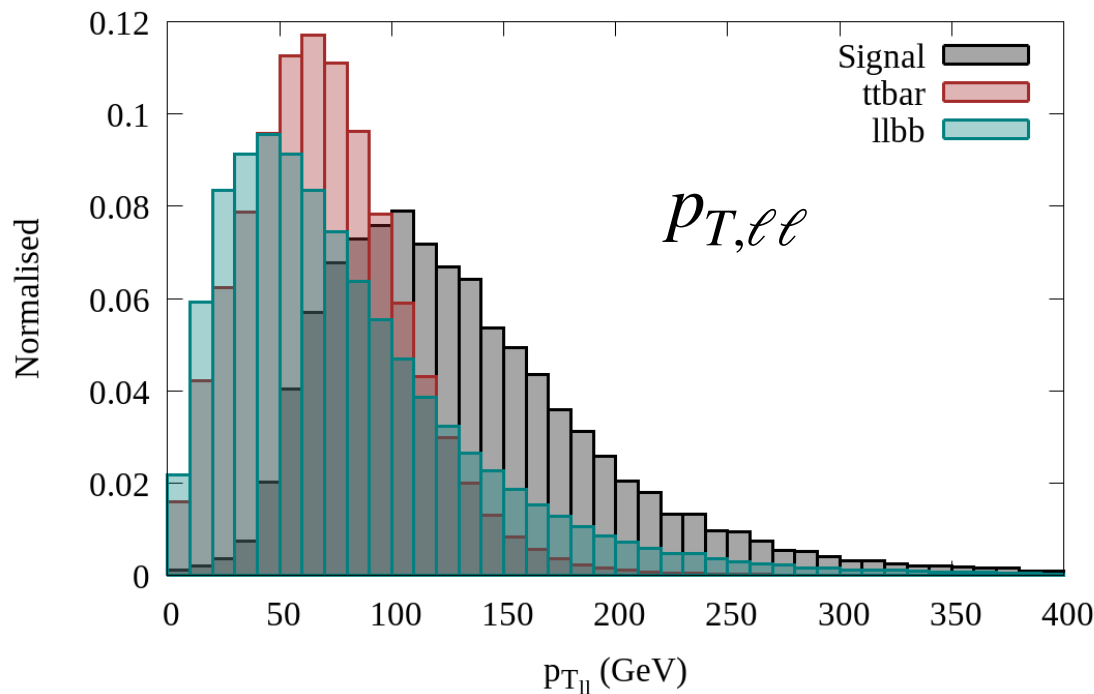
\cancel{E}_T
 $p_{T,\ell_{1/2}}$ $\Delta\phi_{bb\ell\ell}$
 $\Delta R_{\ell\ell}$ ΔR_{bb}
 m_{bb} $m_{\ell\ell}$
 $p_{T,bb}$ $p_{T,\ell\ell}$

Sl. No.	Process	Order	Events
	$t\bar{t} \text{ lep}$	NNLO [128]	2080.52
Background	$t\bar{t}h$	NLO [111]	131.66
	$t\bar{t}Z$	NLO [130]	106.31
	$t\bar{t}W$	NLO [129]	35.97
	$hb\bar{b}$	NNLO (5FS) + NLO (4FS) [111]	~ 0
	$lb\bar{b}$	LO	842.72
	Total		3197.18
	Signal ($hh \rightarrow b\bar{b}WW \rightarrow b\bar{b}\ell\ell + \cancel{E}_T$)	NNLO [70]	35.20
Significance (S/\sqrt{B})			0.62



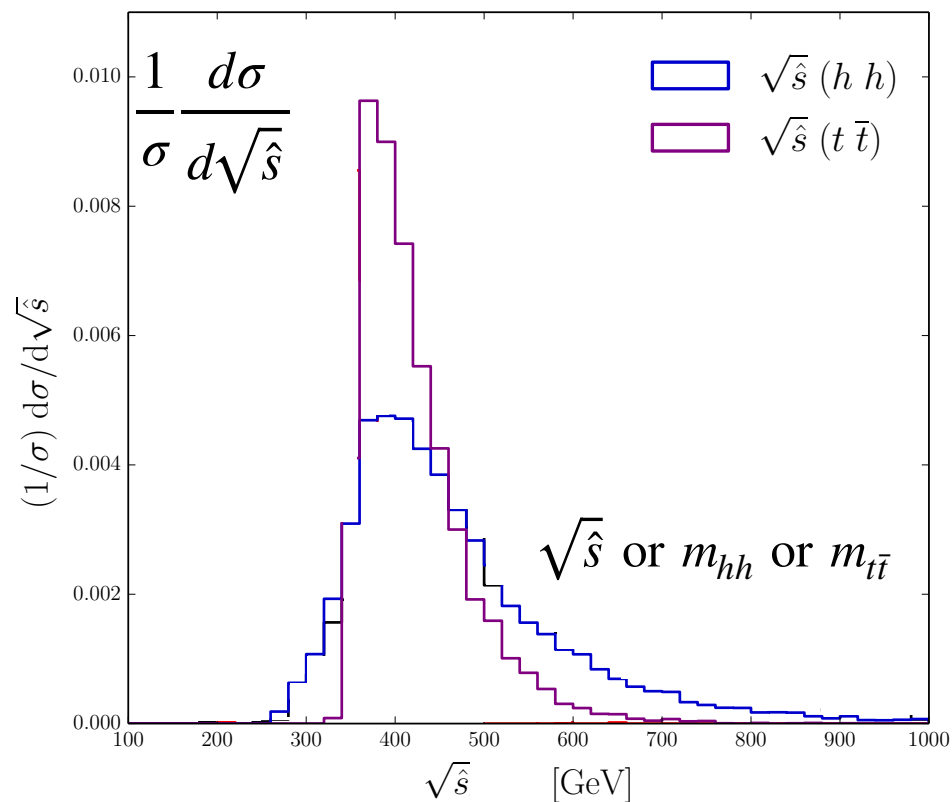
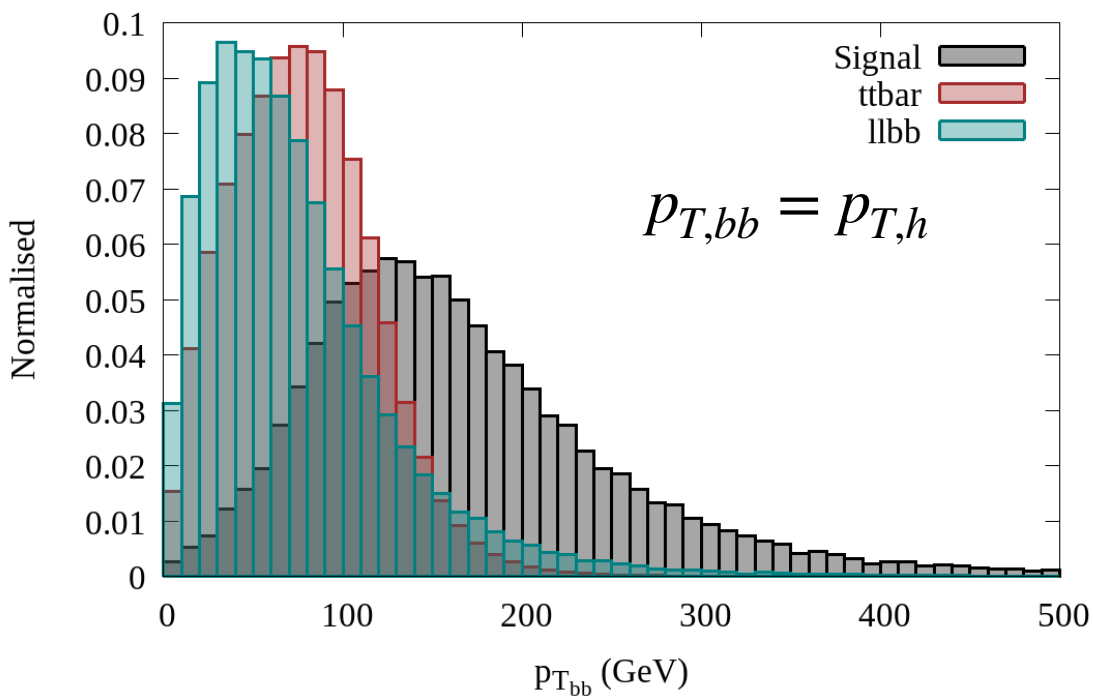
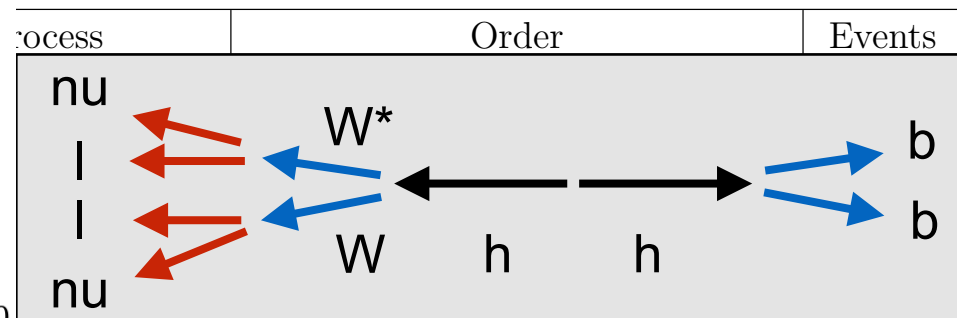
Note that tW background is not included here.

10 variables: $p_{T,\ell_{1/2}}$, \cancel{E}_T , $m_{\ell\ell}$, m_{bb} , $\Delta R_{\ell\ell}$, ΔR_{bb} , $p_{T,bb}$, $p_{T,\ell\ell}$, $\Delta\phi_{bb\ell\ell}$

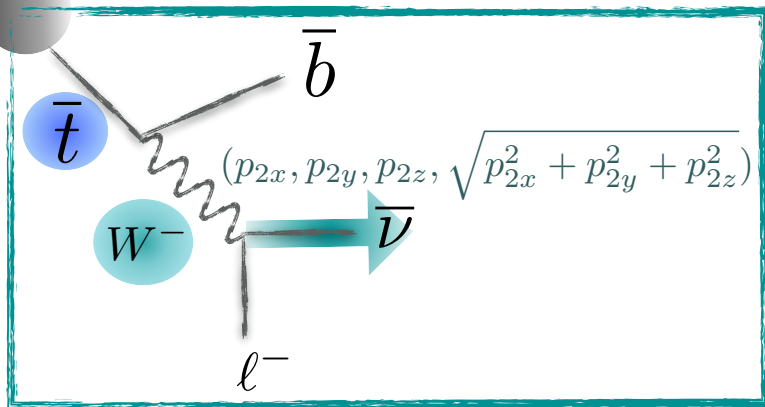
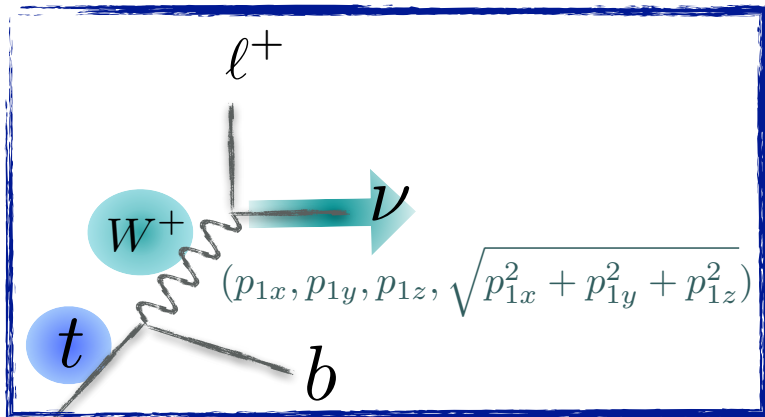


dilepton channel

Chikary, Banerjee, Barman, Bhattacharjee, Niyogi JHEP 2017



How to reduce $t\bar{t}$ background: Topness (T)



$$\chi_{ij}^2 \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} \left[\frac{(m_{b_i l^+ \nu}^2 - m_t^2)^2}{\sigma_t^4} + \frac{(m_{l^+ \nu}^2 - m_W^2)^2}{\sigma_W^4} \right. \\ \left. + \frac{(m_{b_j l^- \bar{\nu}}^2 - m_t^2)^2}{\sigma_t^4} + \frac{(m_{l^- \bar{\nu}}^2 - m_W^2)^2}{\sigma_W^4} \right]$$

$$T \equiv \min(\chi_{12}^2, \chi_{21}^2)$$

two possible ways of pairing b and l

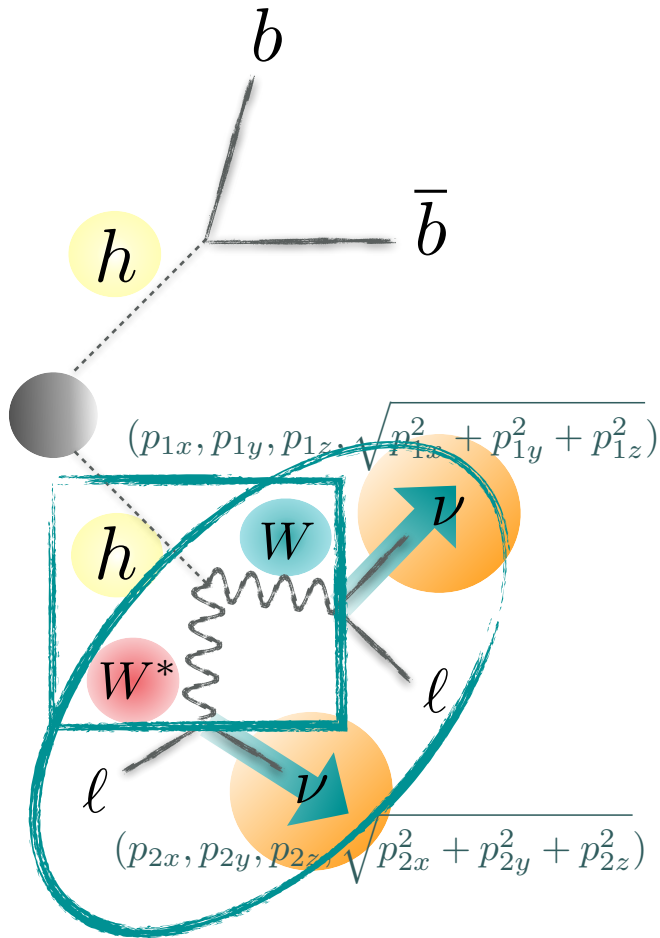
- Topness provides a degree of consistency to dileptonic $t\bar{t}$ production.
- It scans over 6 unknowns of neutrino momenta with four on-shell masses and missing E_T constraints.
- And find the minimum of the likelihood function.
- $t\bar{t}$ events will give a smaller value of Topness than hh events.

Grasser, Shelton, Park, PRL 2013

Kim, Kong, Matchev, Park, PRL 2019

Kim, Kim, Kong, Matchev, Park, JHEP 2019

How to reduce $t\bar{t}$ background: Higgsness (H)



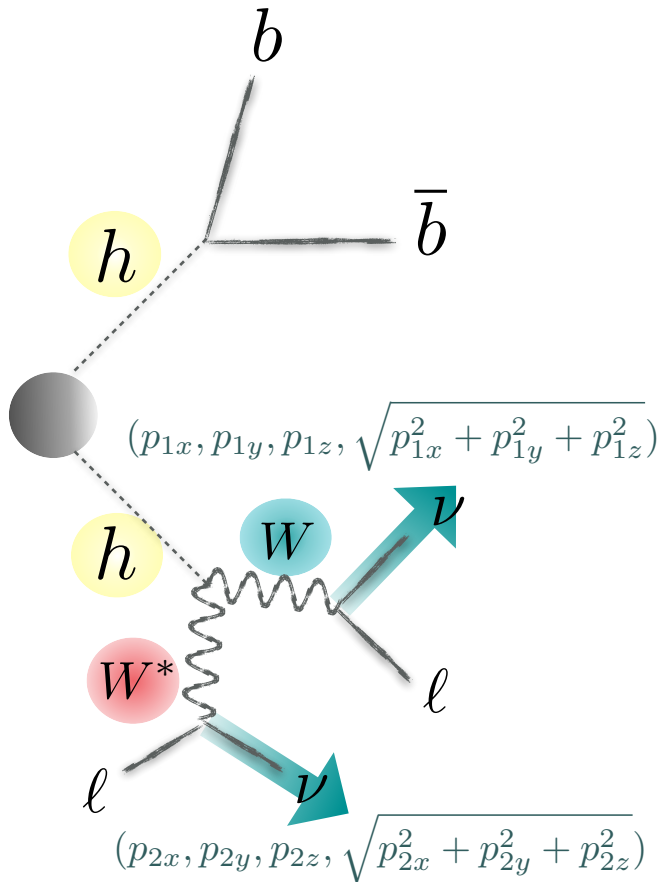
$$H \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} \left[\frac{(m_{\ell^+ \ell^- \nu \bar{\nu}}^2 - m_h^2)^2}{\sigma_{h\ell}^4} + \frac{(m_{\nu \bar{\nu}}^2 - m_{\nu \bar{\nu}, peak}^2)^2}{\sigma_{\nu}^4} \right. \\ \left. + \min \left(\frac{(m_{\ell^+ \nu}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^- \bar{\nu}}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4}, \right. \right. \\ \left. \left. \frac{(m_{\ell^- \bar{\nu}}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^+ \nu}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4} \right) \right],$$

two possible ways of pairing ν and ℓ

$\sim m_h - m_W$
off-shell

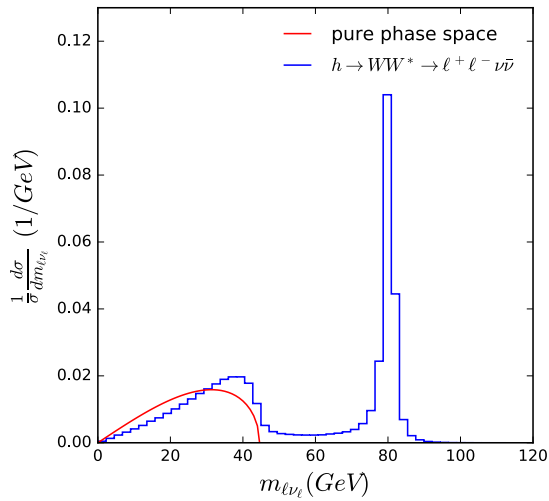
- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

Higgsness (H)



$$\begin{aligned}
 H \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} & \left[\frac{(m_{\ell^+ \ell^- \nu \bar{\nu}}^2 - m_h^2)^2}{\sigma_{h\ell}^4} + \frac{(m_{\nu \bar{\nu}}^2 - m_{\nu \bar{\nu}, peak}^2)^2}{\sigma_{\nu}^4} \right. \\
 & + \min \left(\frac{(m_{\ell^+ \nu}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^- \bar{\nu}}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4}, \right. \\
 & \left. \left. \frac{(m_{\ell^- \bar{\nu}}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^+ \nu}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4} \right) \right], \\
 & \sim m_h - m_W \text{ off-shell}
 \end{aligned}$$

- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

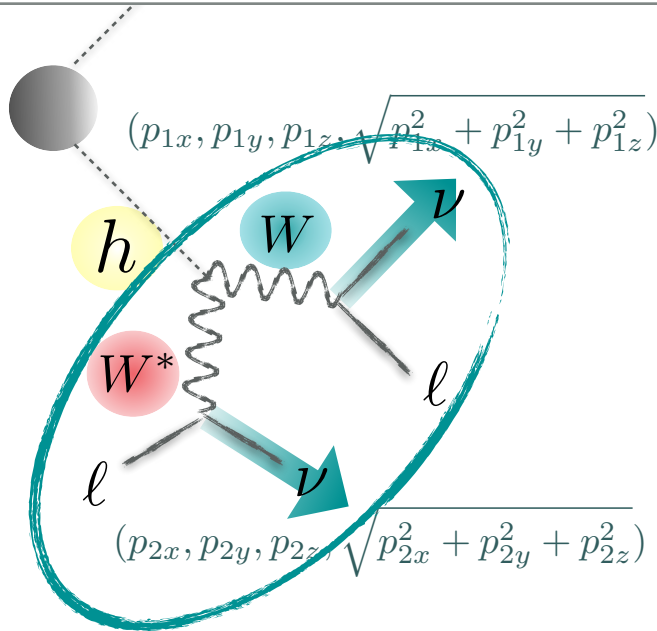


$$m_{W^*}^{peak} = \frac{1}{\sqrt{3}} \sqrt{2(m_h^2 + m_W^2) - \sqrt{m_h^4 + 14m_h^2 m_W^2 + m_W^4}}$$

$$E = \sqrt{m_W m_{W^*}} e^\eta,$$

$$\cosh \eta = \left(\frac{m_h^2 - m_W^2 - m_{W^*}^2}{2m_W m_{W^*}} \right)$$

$$+ \frac{\left(m_{\nu\bar{\nu}}^2 - m_{\nu\bar{\nu},peak}^2 \right)^2}{\sigma_\nu^4}$$

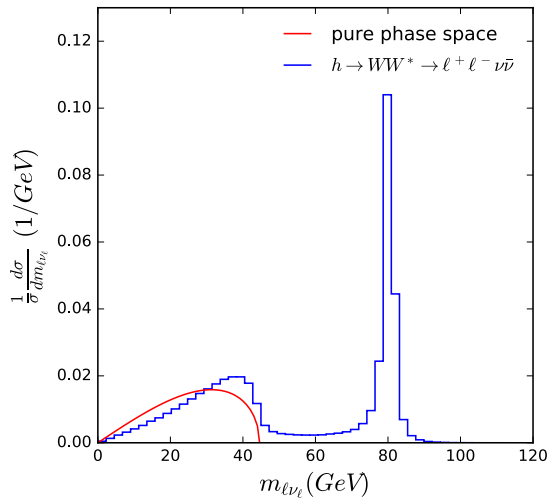


$$+ \min \left(\frac{\left(m_{\ell+\nu}^2 - m_W^2 \right)^2}{\sigma_W^4} + \frac{\left(m_{\ell-\bar{\nu}}^2 - m_{W^*,peak}^2 \right)^2}{\sigma_{W^*}^4} \right),$$

$$\left. \frac{\left(m_{\ell-\bar{\nu}}^2 - m_W^2 \right)^2}{\sigma_W^4} + \frac{\left(m_{\ell+\nu}^2 - m_{W^*,peak}^2 \right)^2}{\sigma_{W^*}^4} \right],$$

$\sim m_h - m_W$
off-shell

- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

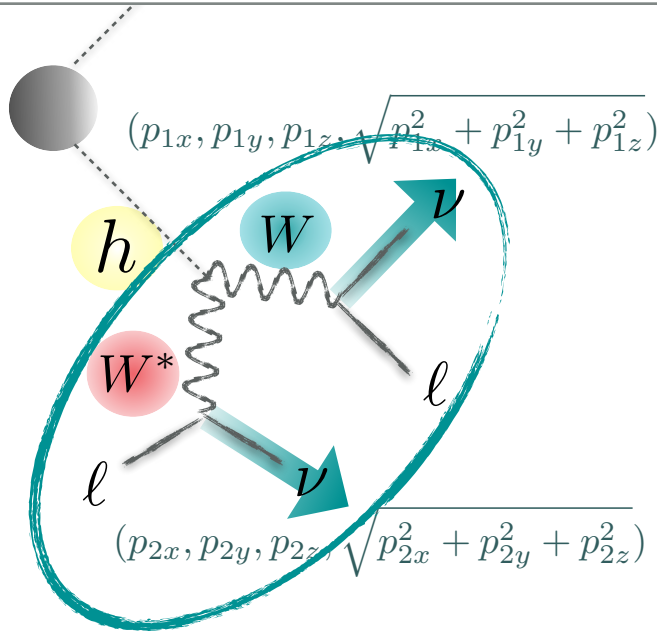


$$m_{W^*}^{peak} = \frac{1}{\sqrt{3}} \sqrt{2(m_h^2 + m_W^2) - \sqrt{m_h^4 + 14m_h^2 m_W^2 + m_W^4}}$$

$$E = \sqrt{m_W m_{W^*}} e^\eta,$$

$$\cosh \eta = \left(\frac{m_h^2 - m_W^2 - m_{W^*}^2}{2m_W m_{W^*}} \right)$$

$$+ \frac{\left(m_{\nu\bar{\nu}}^2 - m_{\nu\bar{\nu},peak}^2 \right)^2}{\sigma_\nu^4}$$

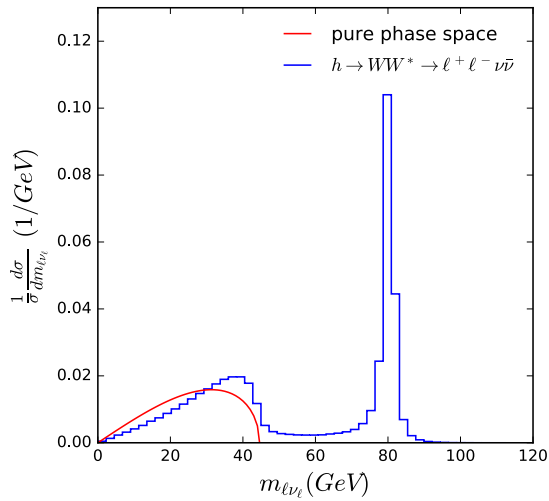


$$+ \min \left(\frac{\left(m_{\ell+\nu}^2 - m_W^2 \right)^2}{\sigma_W^4} + \frac{\left(m_{\ell-\bar{\nu}}^2 - m_{W^*,peak}^2 \right)^2}{\sigma_{W^*}^4}, \right.$$

$$\left. \frac{\left(m_{\ell-\bar{\nu}}^2 - m_W^2 \right)^2}{\sigma_W^4} + \frac{\left(m_{\ell+\nu}^2 - m_{W^*,peak}^2 \right)^2}{\sigma_{W^*}^4} \right),$$

$\sim m_h - m_W$
off-shell

- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

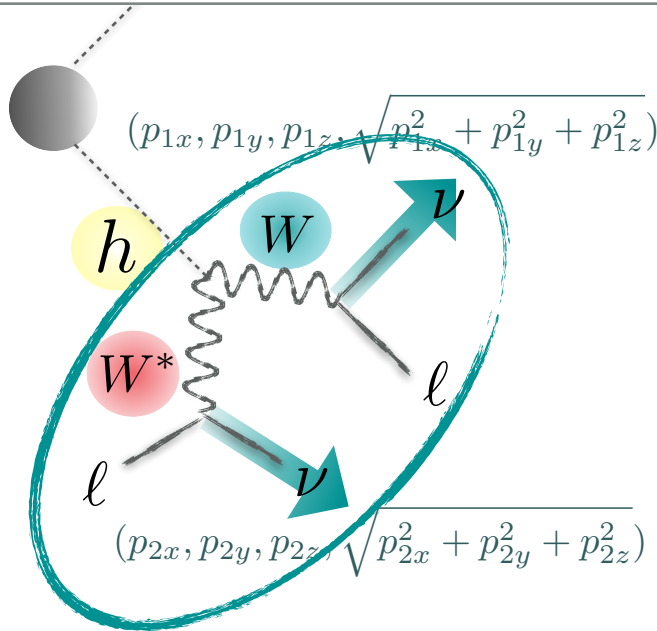


$$m_{W^*}^{peak} = \frac{1}{\sqrt{3}} \sqrt{2(m_h^2 + m_W^2) - \sqrt{m_h^4 + 14m_h^2 m_W^2 + m_W^4}}$$

$$E = \sqrt{m_W m_{W^*}} e^\eta,$$

$$\cosh \eta = \left(\frac{m_h^2 - m_W^2 - m_{W^*}^2}{2m_W m_{W^*}} \right)$$

$$+ \frac{\left(m_{\nu\bar{\nu}}^2 - m_{\nu\bar{\nu},peak}^2 \right)^2}{\sigma_\nu^4}$$



$$+ \min \left(\frac{\left(m_{\ell+\nu}^2 - m_W^2 \right)^2}{\sigma_W^4} + \frac{\left(m_{\ell-\bar{\nu}}^2 - m_{W^*,peak}^2 \right)^2}{\sigma_{W^*}^4}, \right.$$

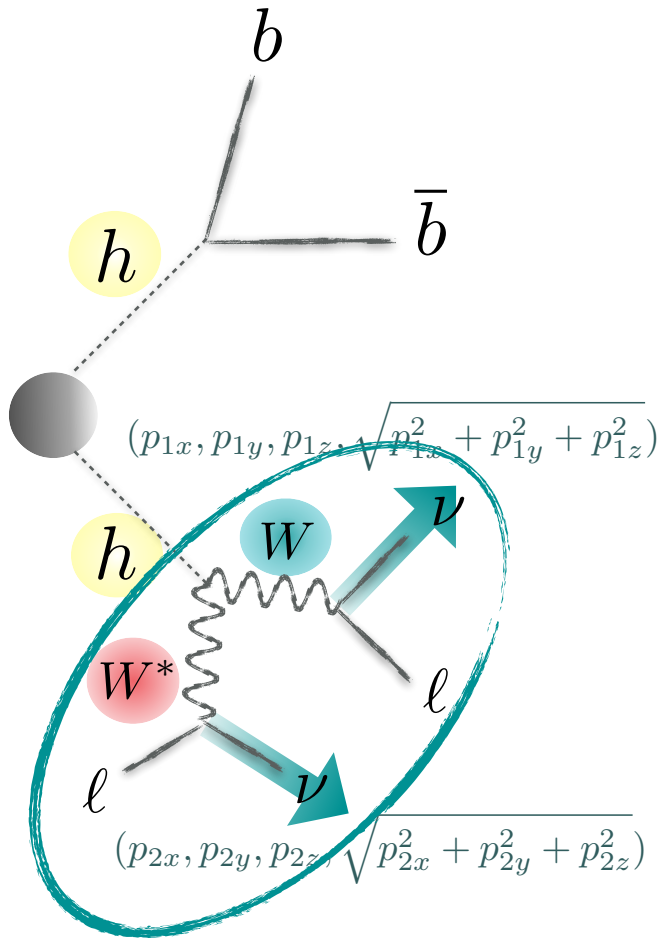
$$\left. \frac{\left(m_{\ell-\bar{\nu}}^2 - m_W^2 \right)^2}{\sigma_W^4} + \frac{\left(m_{\ell+\nu}^2 - m_{W^*,peak}^2 \right)^2}{\sigma_{W^*}^4} \right),$$

two possible ways of paring ν and ℓ

$\sim m_h - m_W$
off-shell

- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

Higgsness (H)



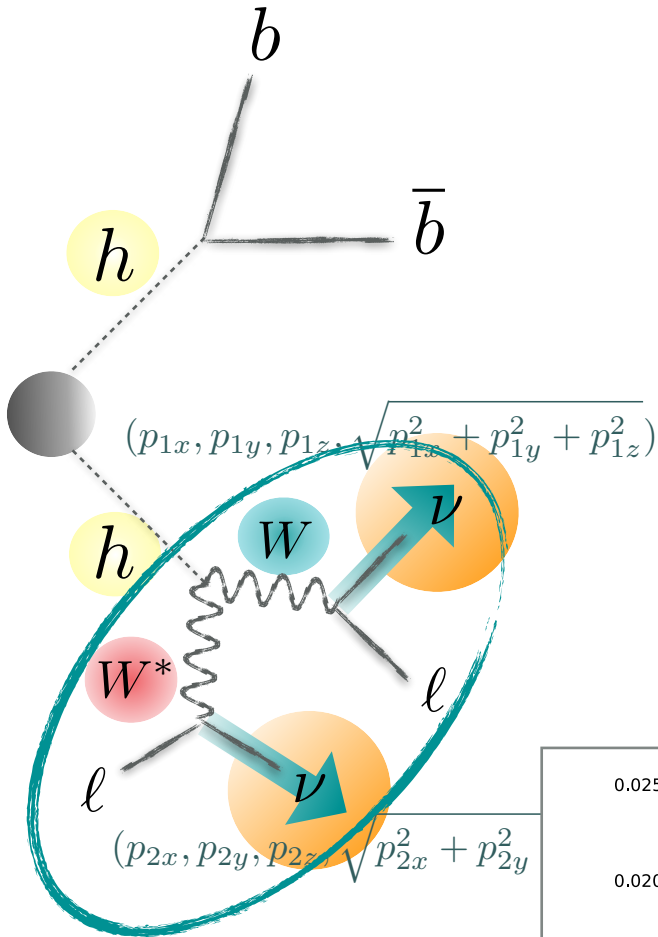
$$H \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} \left[\frac{(m_{\ell^+ \ell^- \nu \bar{\nu}}^2 - m_h^2)^2}{\sigma_{h\ell}^4} + \frac{(m_{\nu \bar{\nu}}^2 - m_{\nu \bar{\nu}, peak}^2)^2}{\sigma_{\nu}^4} \right. \\ \left. + \min \left(\frac{(m_{\ell^+ \nu}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^- \bar{\nu}}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4}, \right. \right. \\ \left. \left. \frac{(m_{\ell^- \bar{\nu}}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^+ \nu}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4} \right) \right],$$

two possible ways of pairing ν and ℓ

$\sim m_h - m_W$
off-shell

- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

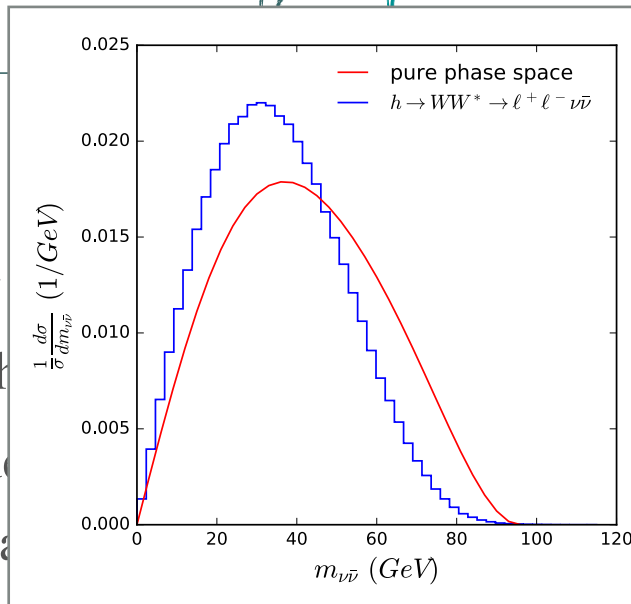
Higgsness (H)



$$H \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} \left[\frac{(m_{\ell^+ \ell^- \nu \bar{\nu}}^2 - m_h^2)^2}{\sigma_{h\ell}^4} + \frac{(m_{\nu \bar{\nu}}^2 - m_{\nu \bar{\nu}, peak}^2)^2}{\sigma_{\nu}^4} \right]$$

$$+ \min \left(\frac{(m_{\ell^+ \nu}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^- \bar{\nu}}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4}, \right.$$

$$\left. \frac{(m_{\ell^- \bar{\nu}}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^+ \nu}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4} \right],$$



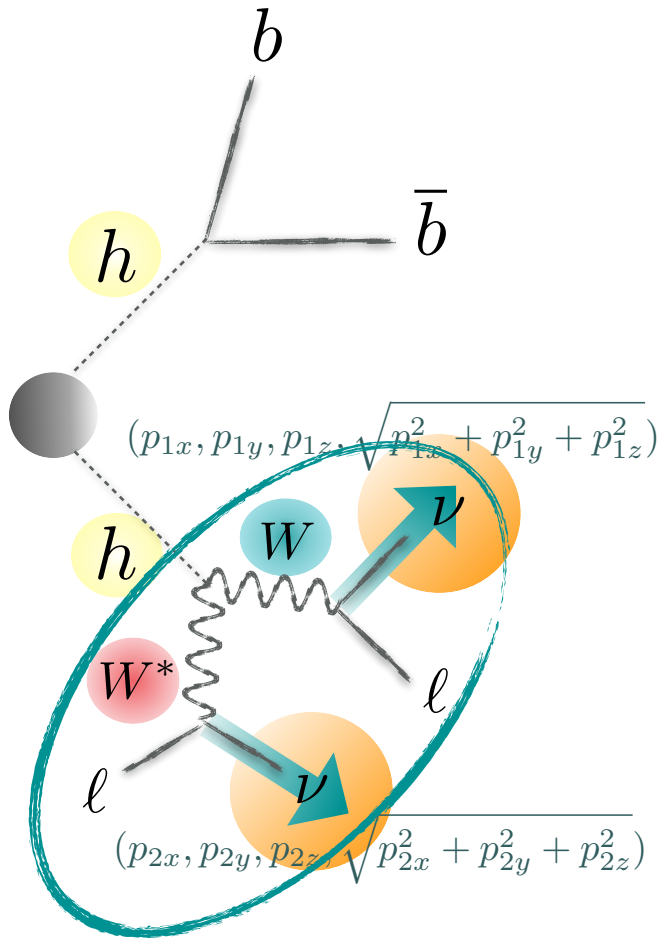
$$\frac{d\sigma}{dm_{\nu\bar{\nu}}} \propto \int dm_{W^*}^2 \lambda^{1/2}(m_h^2, m_W^2, m_{W^*}^2) f(m_{\nu\bar{\nu}})$$

$$f(m) \sim \begin{cases} \eta m, & 0 \leq m \leq e^{-\eta} E, \\ m \ln(E/m), & e^{-\eta} E \leq m \leq E, \end{cases}$$

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$$

- Higgsness provides a
- The off-shell W also h
- Its distribution is wide
- $t\bar{t}$ events will give a l

Higgsness (H)



$$H \equiv \min_{\vec{p}_T = \vec{p}_{\nu T} + \vec{p}_{\bar{\nu} T}} \left[\frac{(m_{\ell^+ \ell^- \nu \bar{\nu}}^2 - m_h^2)^2}{\sigma_{h\ell}^4} + \frac{(m_{\nu \bar{\nu}}^2 - m_{\nu \bar{\nu}, peak}^2)^2}{\sigma_{\nu}^4} \right]$$

$$+ \min \left(\frac{(m_{\ell^+ \nu}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^- \bar{\nu}}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4}, \right.$$

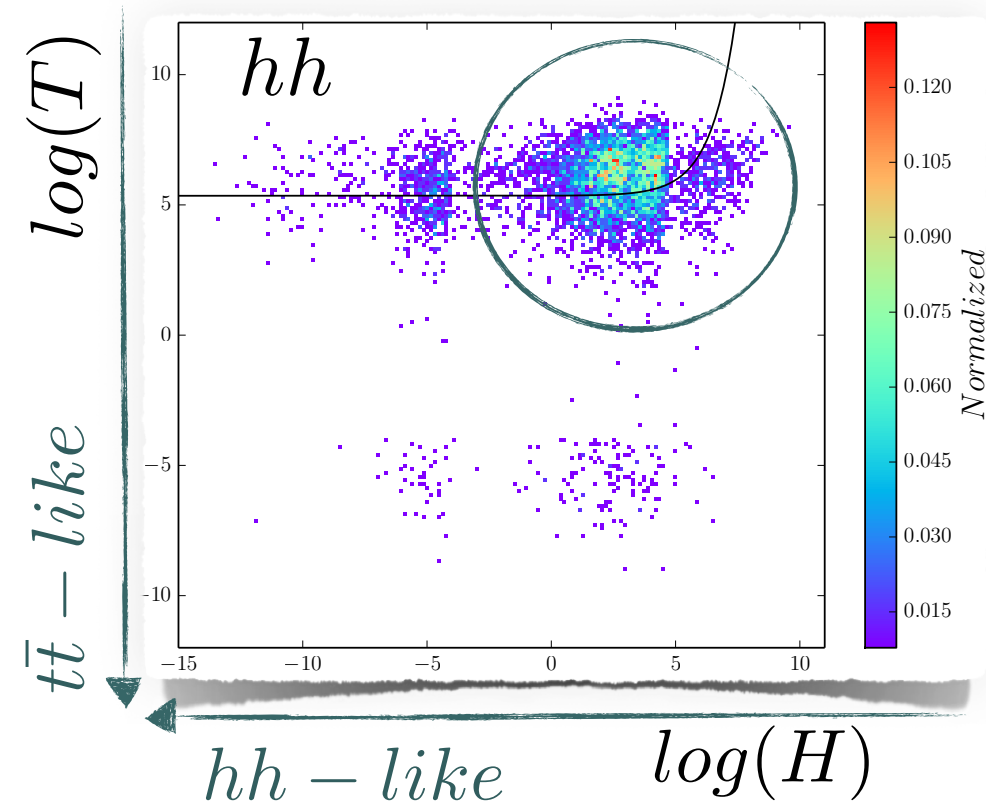
$$\left. \frac{(m_{\ell^- \bar{\nu}}^2 - m_W^2)^2}{\sigma_W^4} + \frac{(m_{\ell^+ \nu}^2 - m_{W^*, peak}^2)^2}{\sigma_{W^*}^4} \right],$$

two possible ways of pairing ν and ℓ

$\sim m_h - m_W$
off-shell

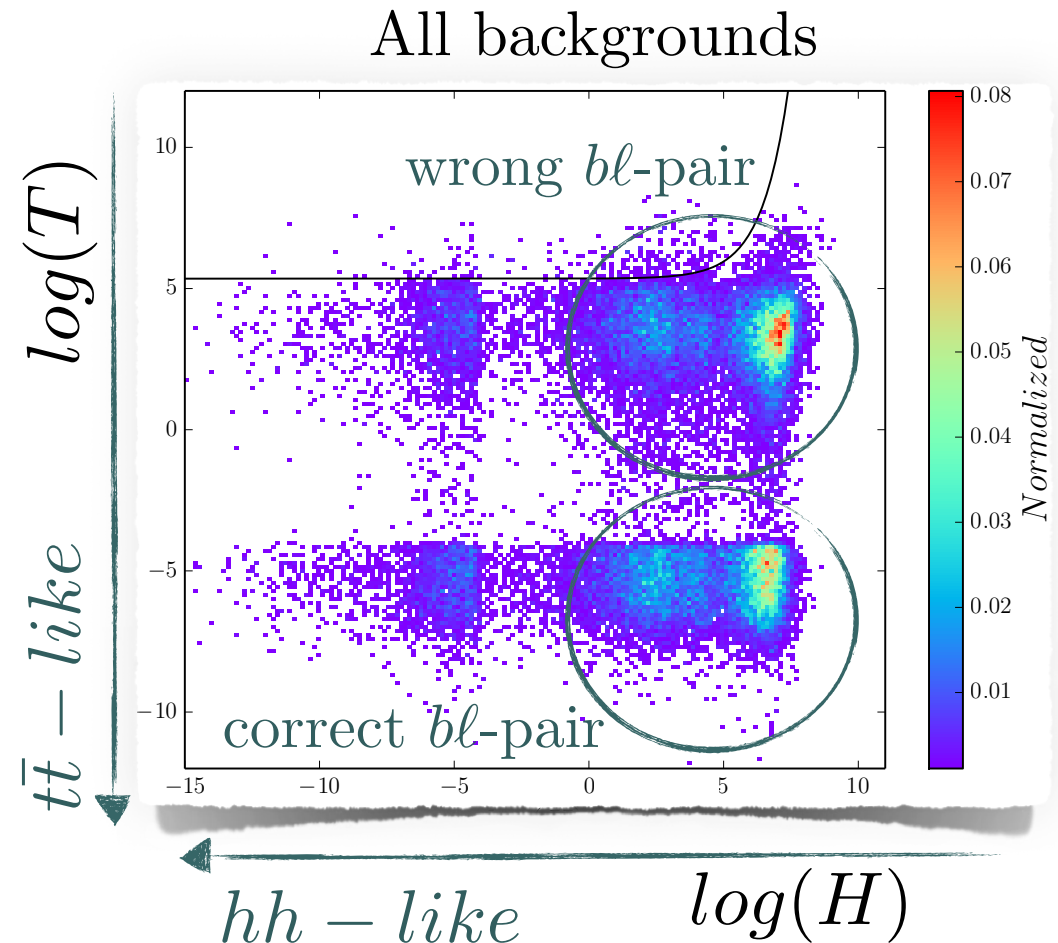
- Higgsness provides a degree of consistency to dileptonic $h \rightarrow WW^*$ system.
- The off-shell W also has an end-point near $m_h - m_W$.
- Its distribution is wide, but there is a peak, which can constrain hh system further.
- $t\bar{t}$ events will give a larger value of Higgsness than hh events.

Distributions of $(\log H, \log T)$ after baseline selection cuts

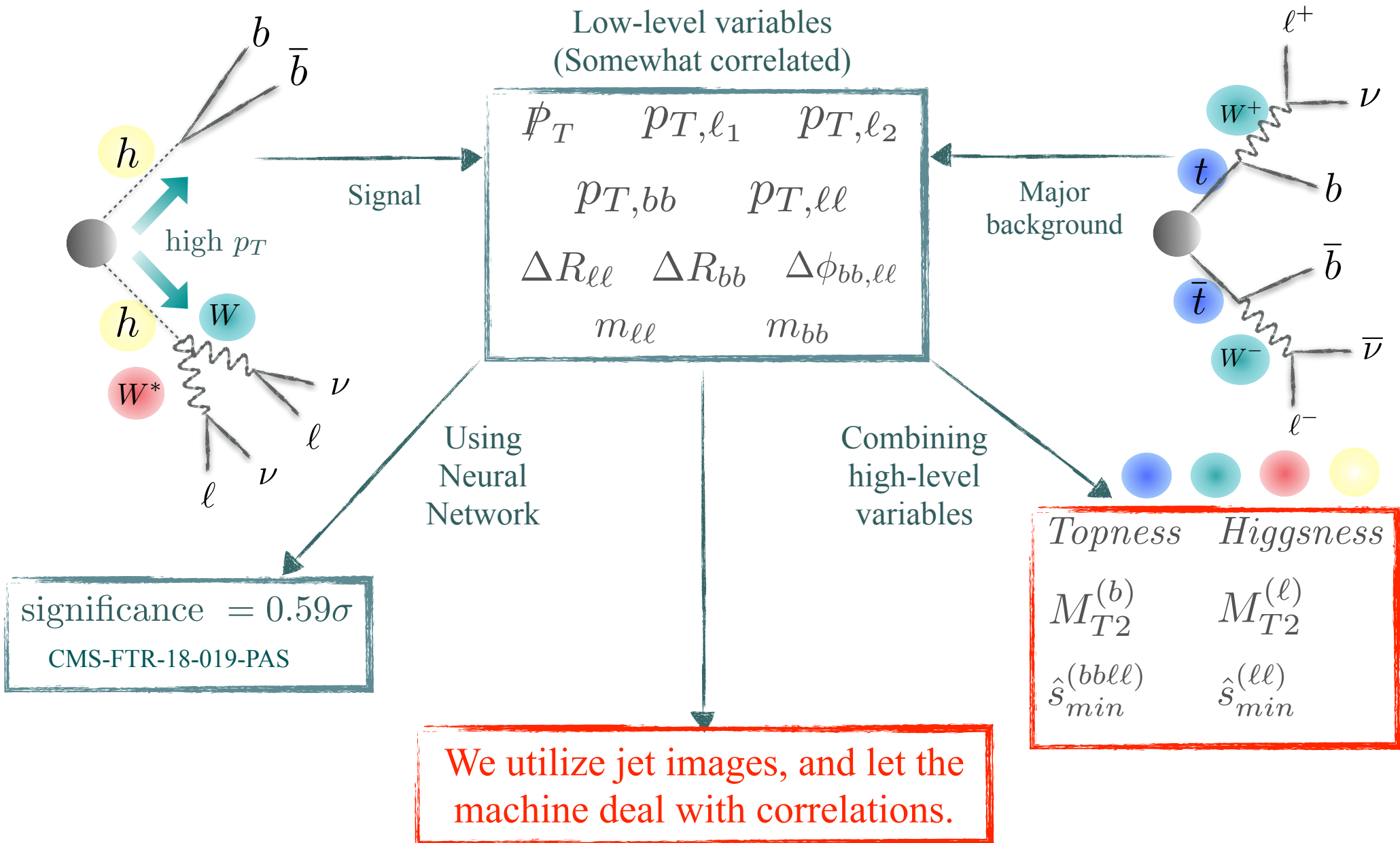


- A clear separation between *hh* and backgrounds (*t \bar{t}* is dominant)

- Since there is a two-fold ambiguity in *bl*-paring, Topness displays the island-nature.



How to reduce backgrounds further



Different Color flows

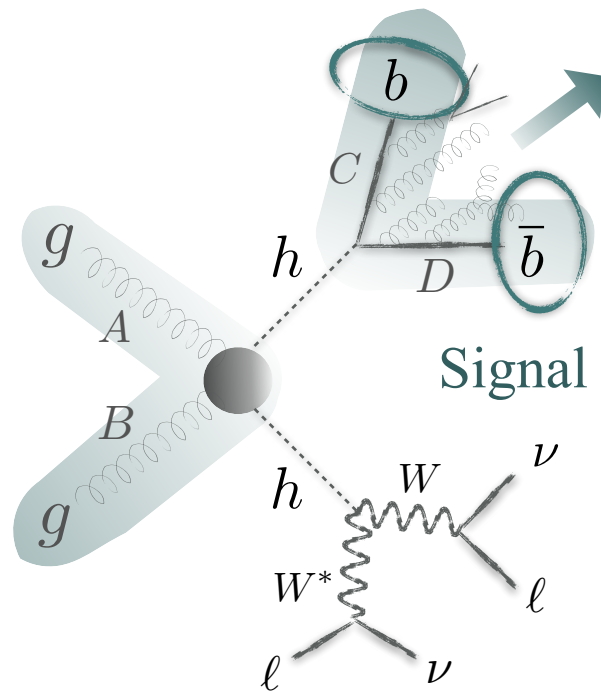
J. Gallicchio and M. D. Schwartz [2010]

L. Oliveira, M. Kagan, L. Mackey, B. Nachman, A. Schwarzman [2017]

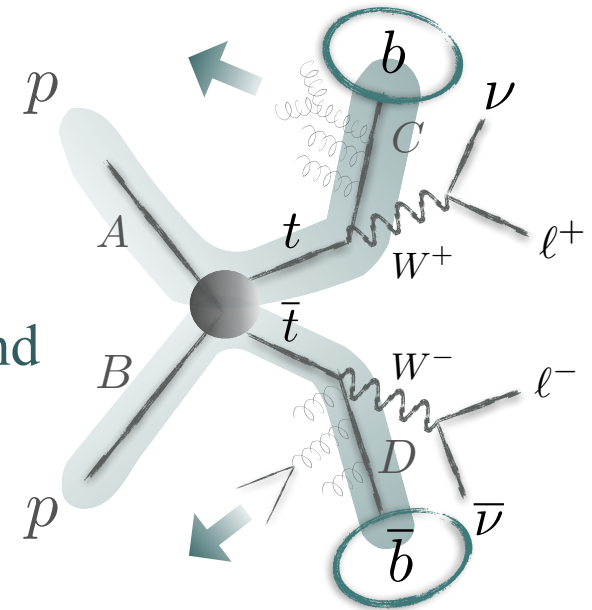
J. Lin, M. Freytsis, I. Moutl, B. Nachman [2018]

See also P. T. Komiske, E. M. Metodiev, J. Thaler [2019]

J. H. Kim, M. Kim, K. C. Kong, K. T. Matchev, M. Park [2019]



Major background



- The advantage of using jet images is that we can better capture color-flow effects.
- Since the Higgs is a color-singlet, two b jets are color-connected with each other.
- However, two b jets from $t\bar{t}$ are color-connected with initial states.
- Parton showering dominantly occurs in the direction of color string.

Different Color flows

J. Gallicchio and M. D. Schwartz [2010]

L. Oliveira, M. Kagan, L. Mackey, B. Nachman, A. Schwarzman [2017]

J. Lin, M. Freytsis, I. Moutl, B. Nachman [2018]

See also P. T. Komiske, E. M. Metodiev, J. Thaler [2019]

J. H. Kim, M. Kim, K. C. Kong, K. T. Matchev, M. Park [2019]

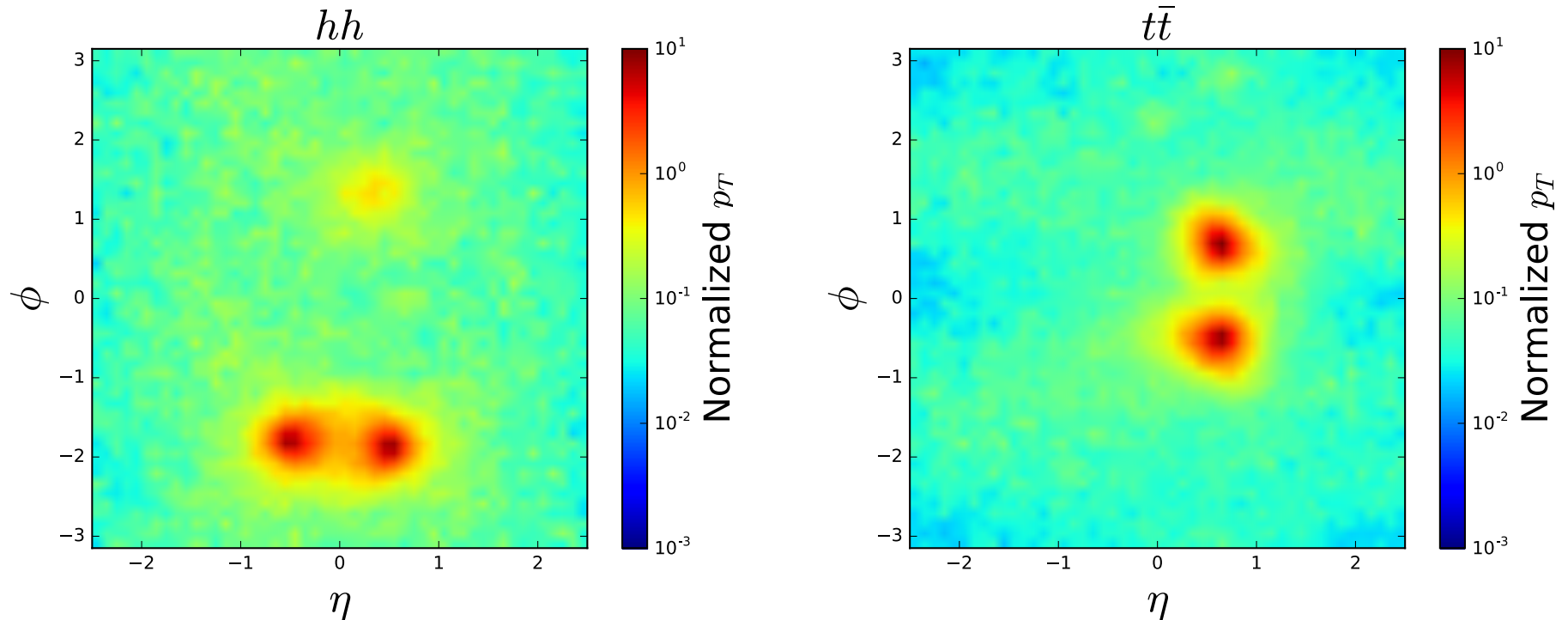
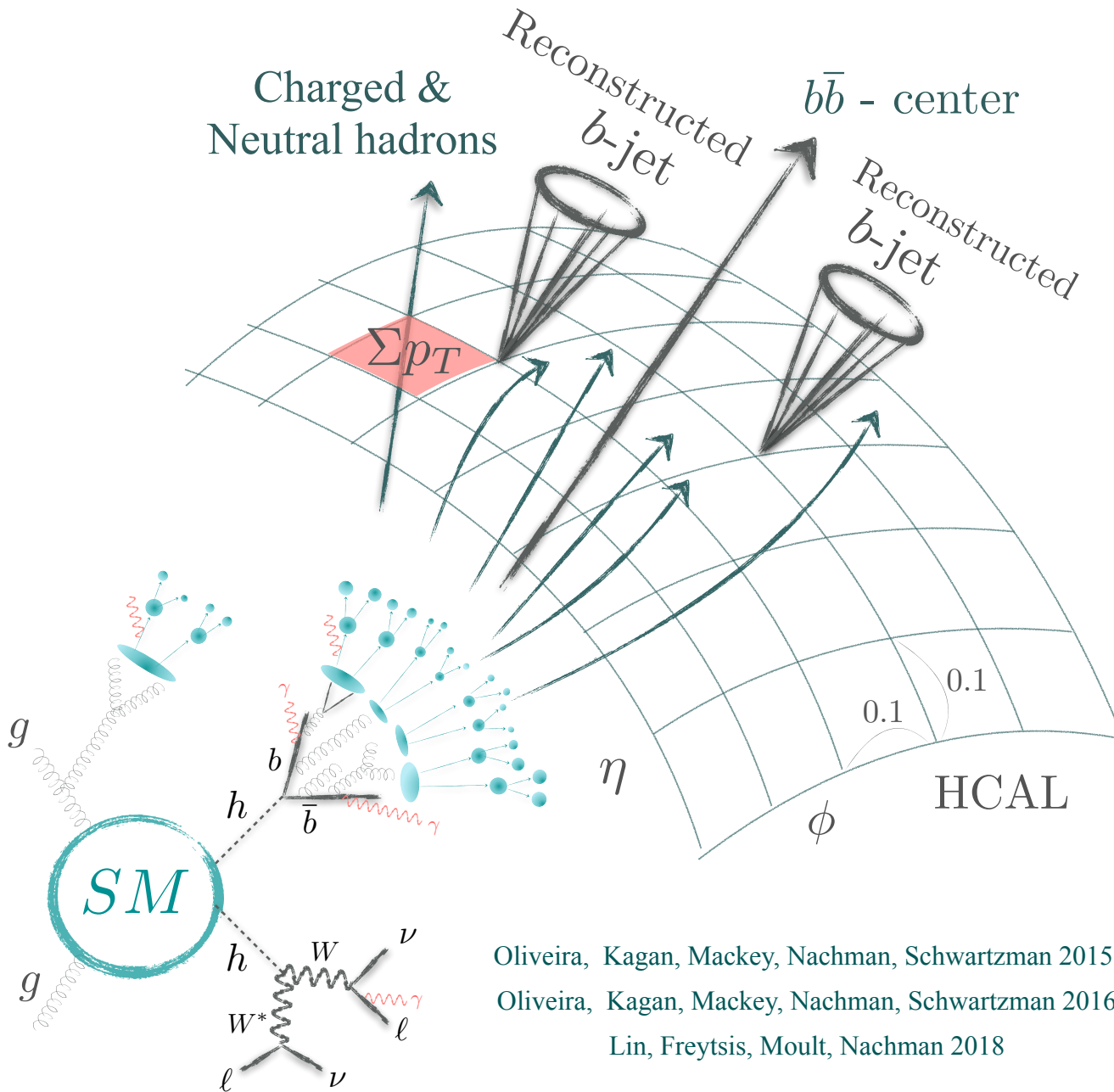


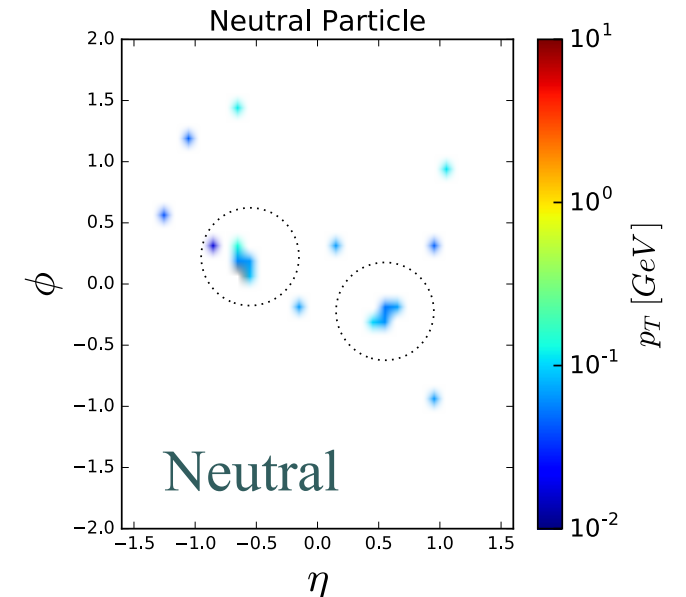
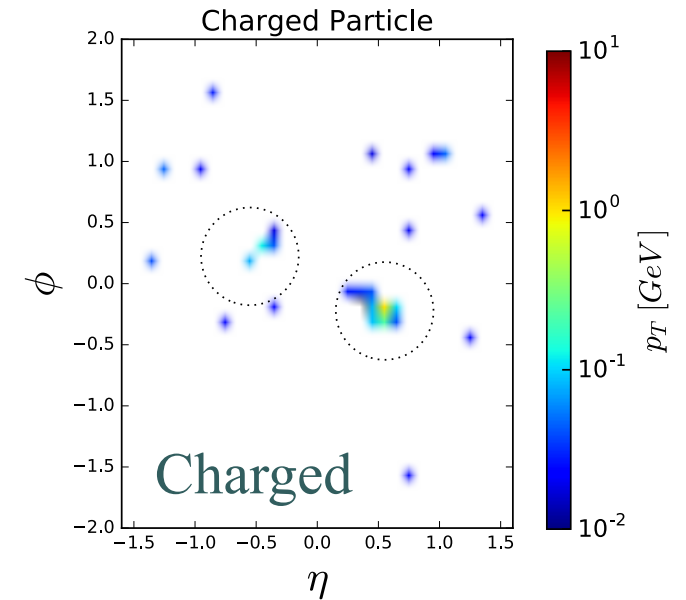
Figure 5. Cumulative p_T distributions resulting from showering 10,000 times a single partonic event for the signal (left) and $t\bar{t}$ production (right). The two b quarks from $h \rightarrow b\bar{b}$ are color-connected to each other and the soft radiation tends to fill in the region between them (left panel), while the two b quarks from $t\bar{t}$ production are not color-connected and the two clusters from their hadronization tend to be more isolated (right panel).

Processing Hadron Images (hh)



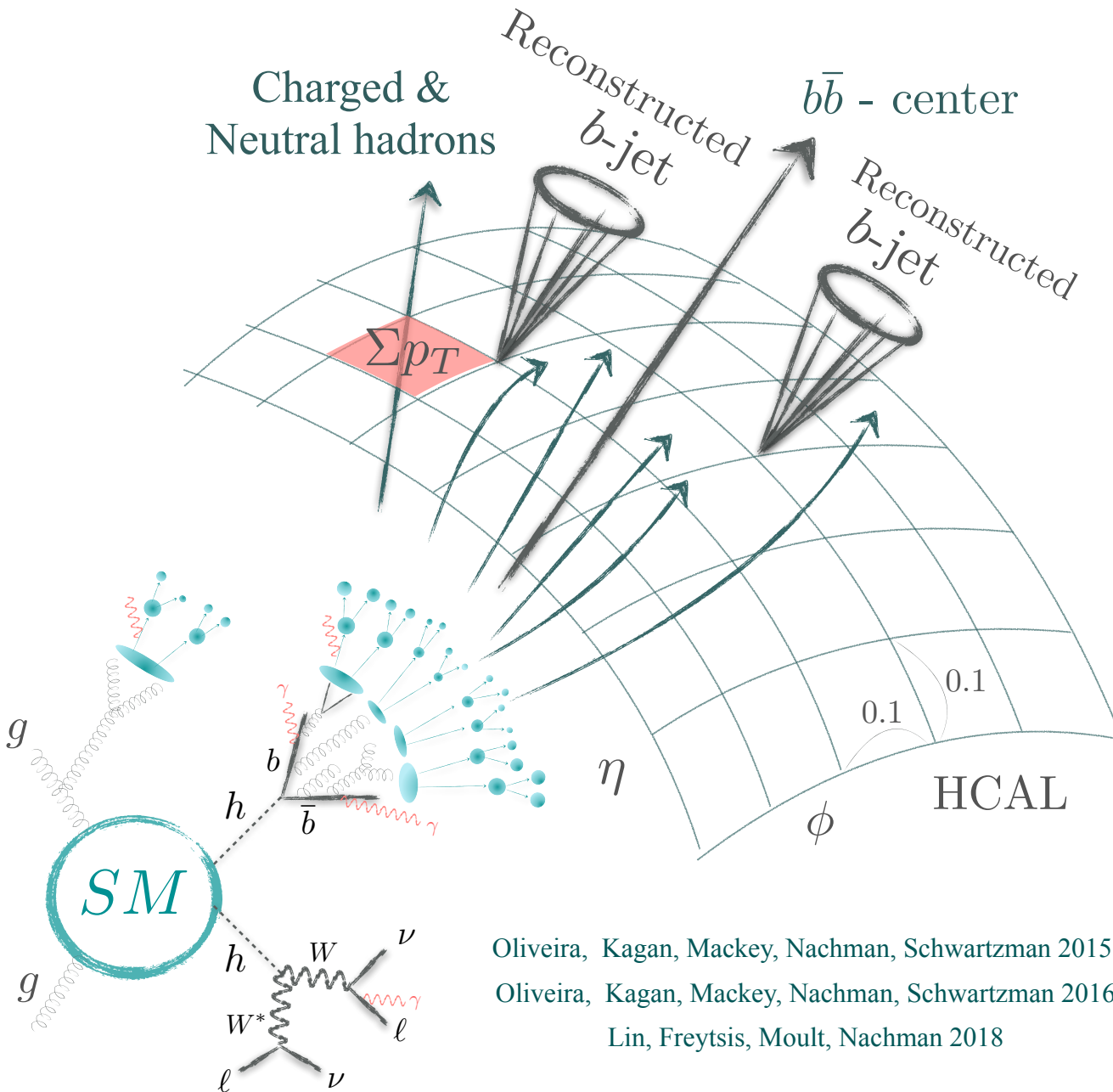
Oliveira, Kagan, Mackey, Nachman, Schwartzman 2015,
 Oliveira, Kagan, Mackey, Nachman, Schwartzman 2016
 Lin, Freytsis, Moul, Nachman 2018

Each event

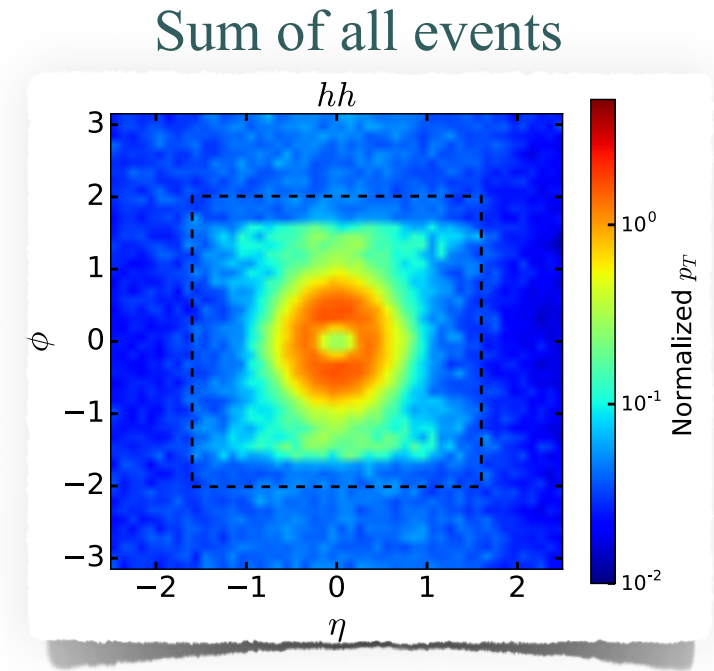


Kim, Kong, Matchev, Park JHEP 2019

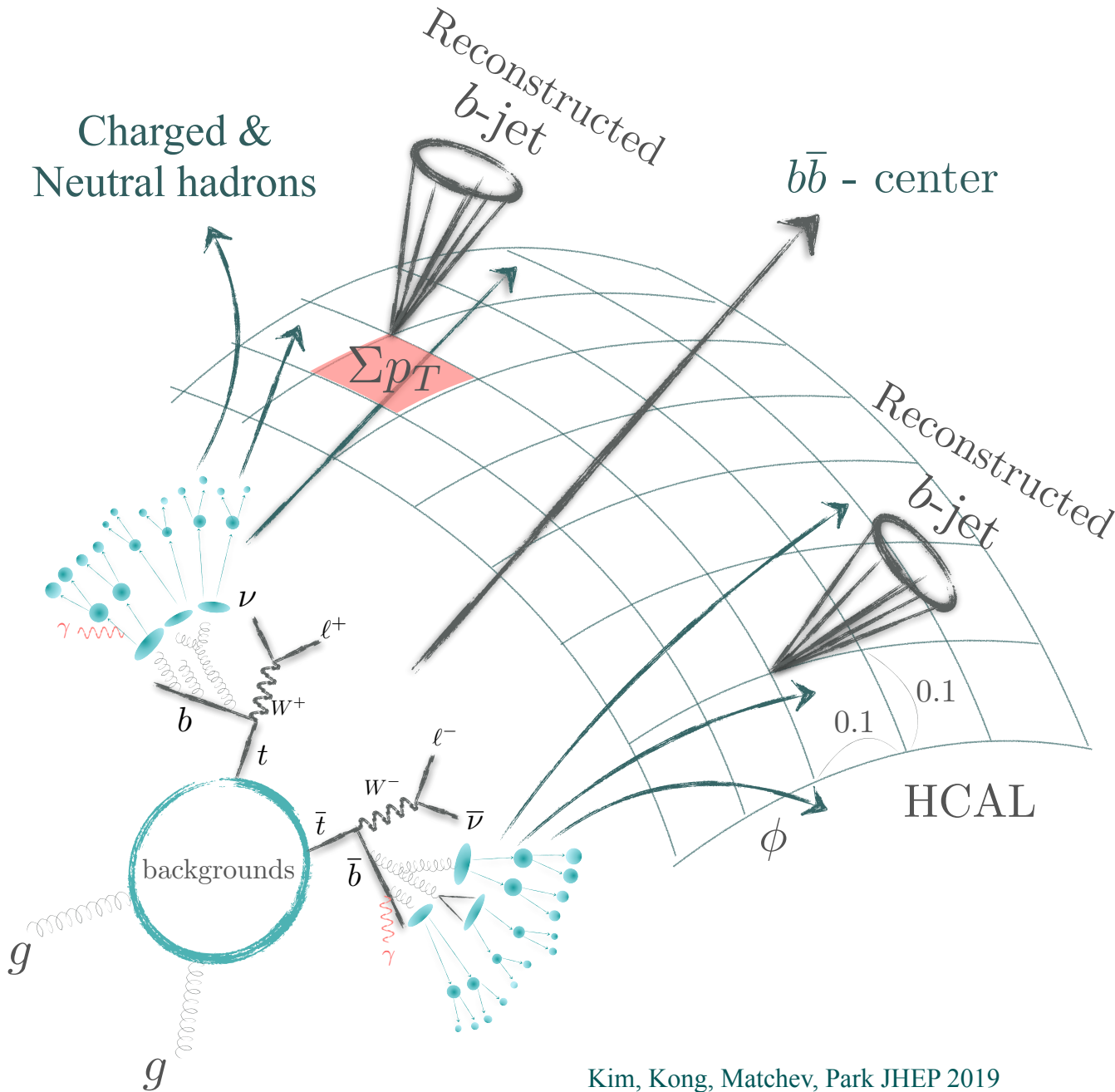
Processing Hadron Images (hh)



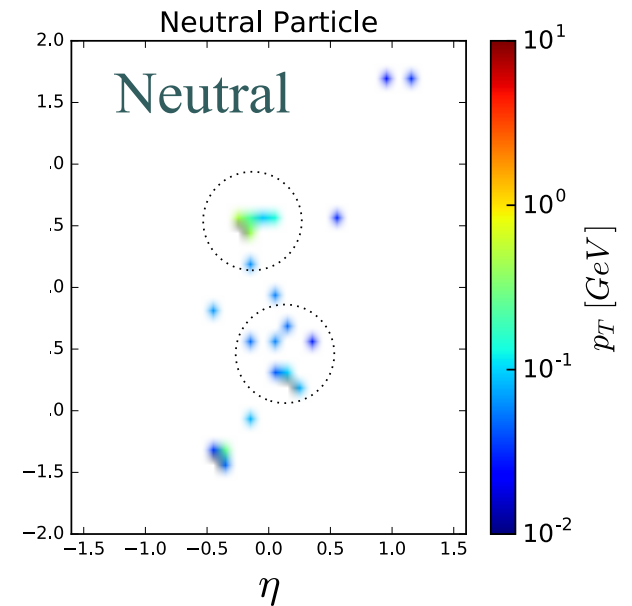
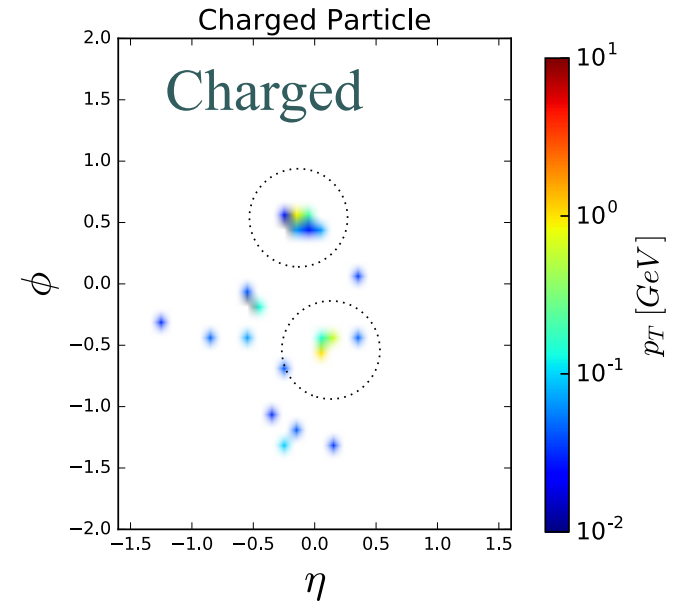
Oliveira, Kagan, Mackey, Nachman, Schwartzman 2015,
 Oliveira, Kagan, Mackey, Nachman, Schwartzman 2016
 Lin, Freytsis, Moutl, Nachman 2018



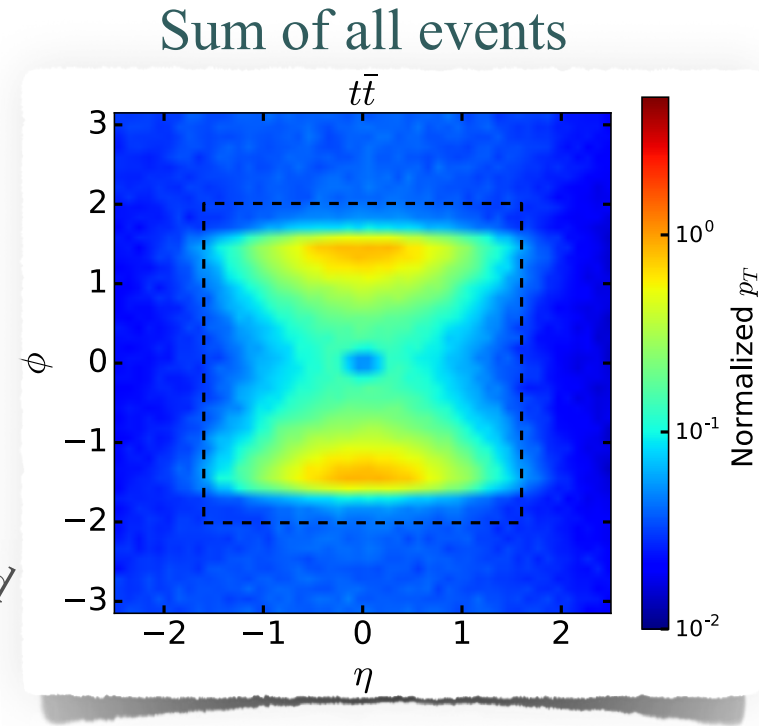
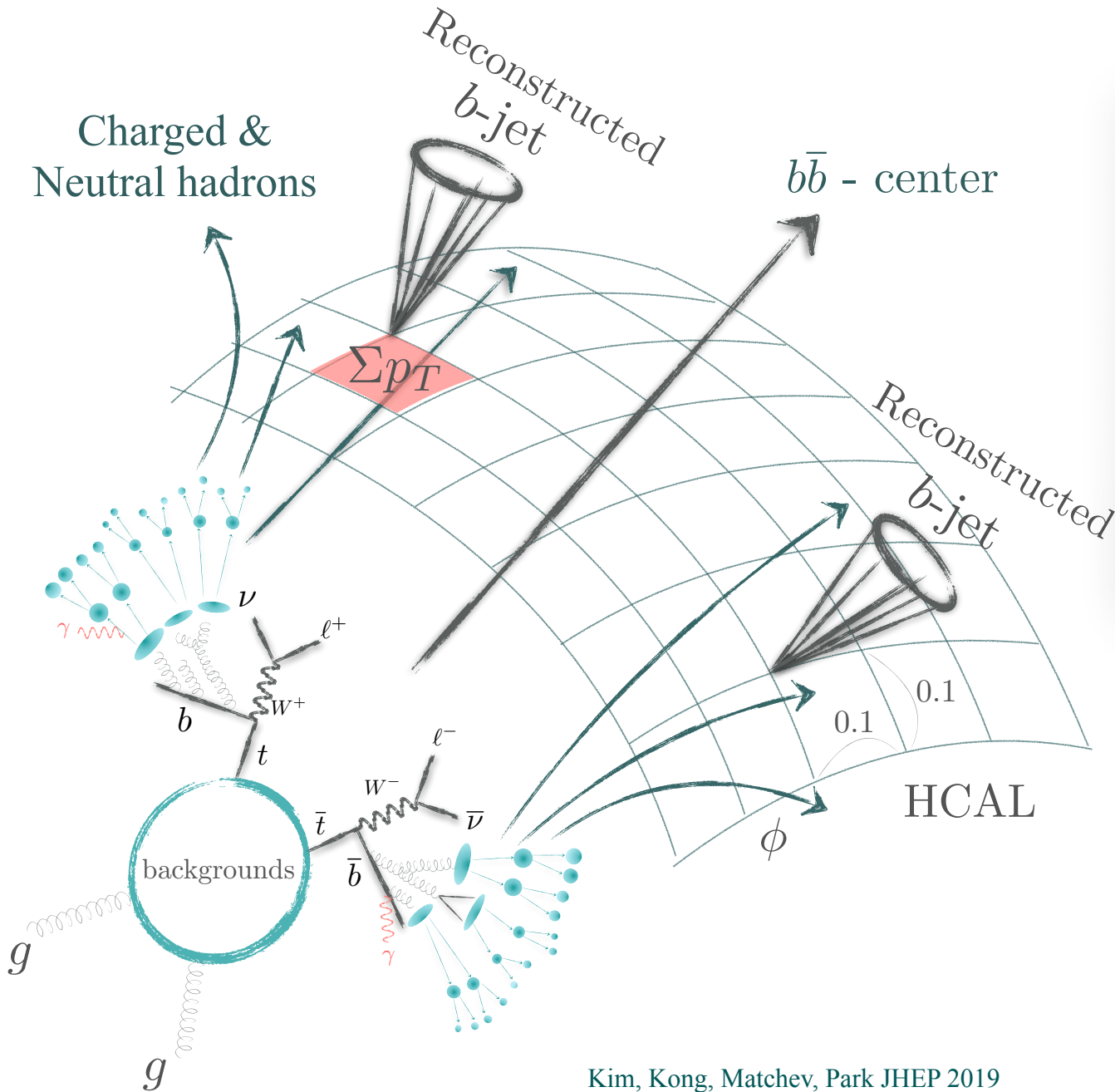
Processing Hadron Images (tt)



Each $t\bar{t}$ event

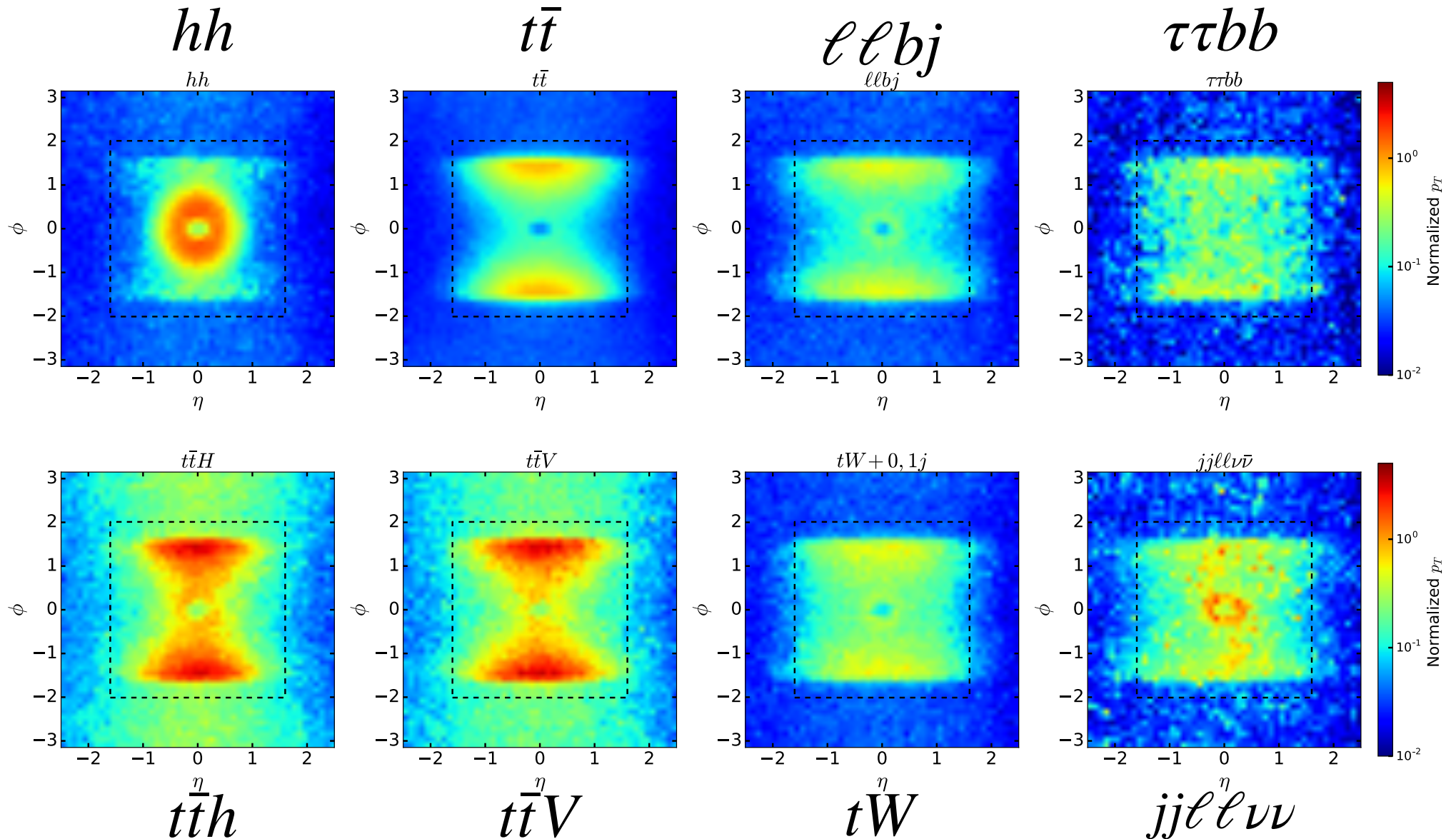


Processing Hadron Images (tt)



Jet images before baseline cuts

Kim, Kong, Matchev, Park JHEP 2019



Baseline cuts: $\cancel{p}_T > 20$ GeV,

$p_{T,l} > 20$ GeV, $\Delta R_{\ell\ell} < 1.0$,

$p_{T,b} > 30$ GeV, $\Delta R_{bb} < 1.3$,

$m_{\ell\ell} < 65$ GeV, $95 < m_{bb} < 140$ GeV

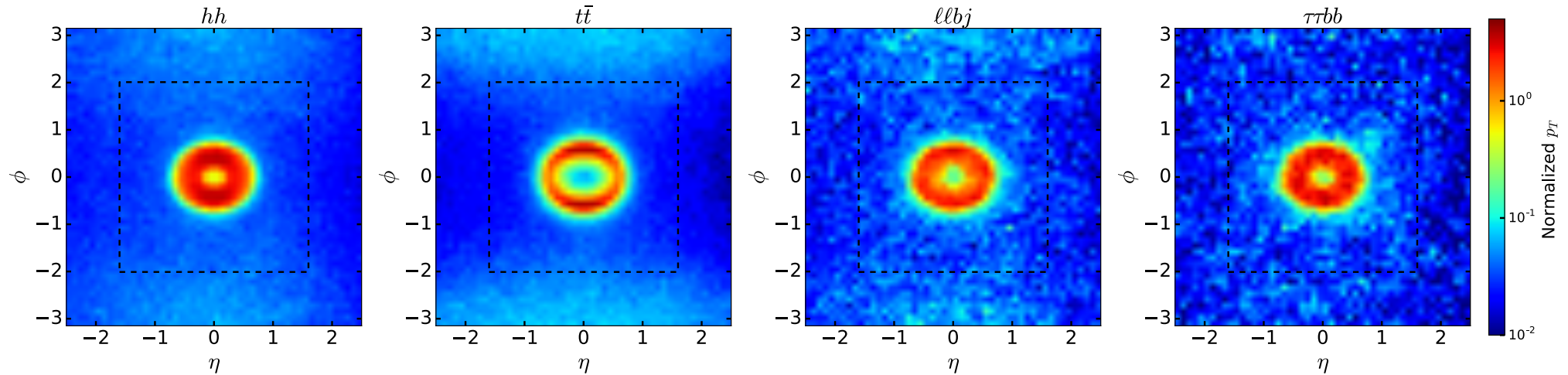
Jet images after baseline cuts

hh

$t\bar{t}$

$\ell\ell bj$

$\tau\tau bb$

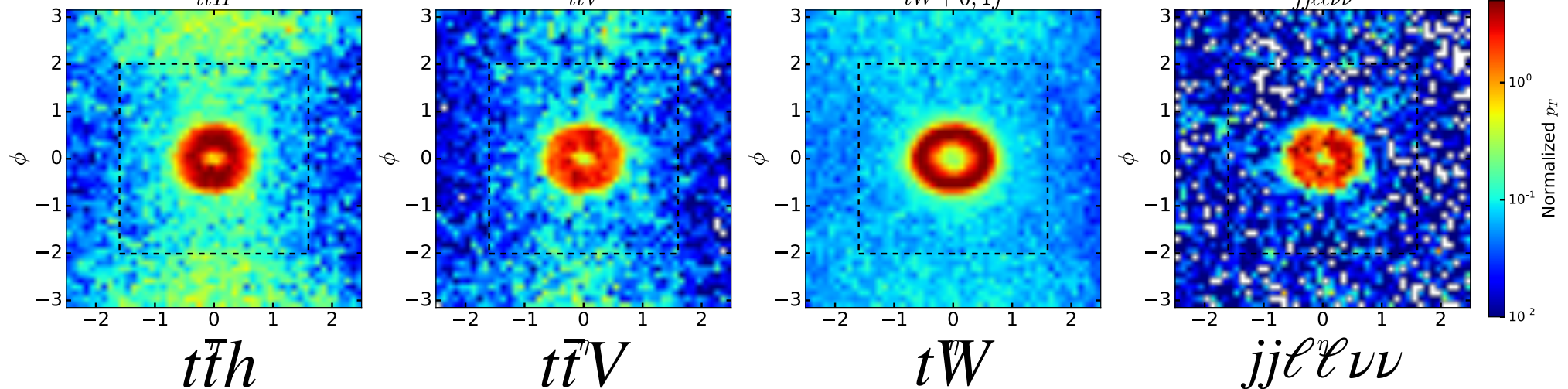


$t\bar{t}H$

$t\bar{t}V$

$tW+0,1j$

$jj\ell\ell\nu\nu$



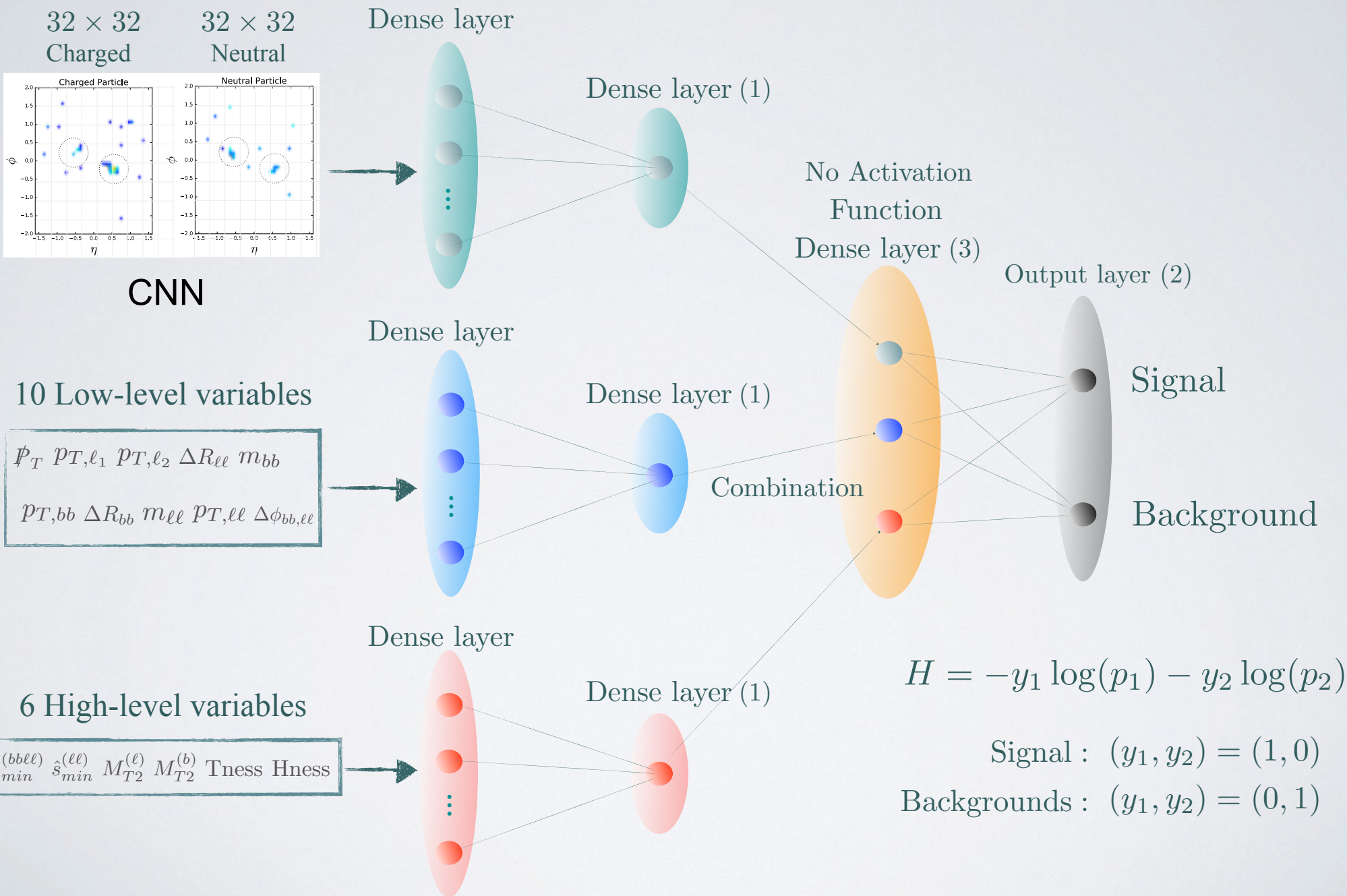
$t\bar{t}h$

$t\bar{t}V$

tW

$jj\ell\ell\nu\nu$

Combining dense neural networks



$hh \rightarrow bbWW^*$ discovery significance

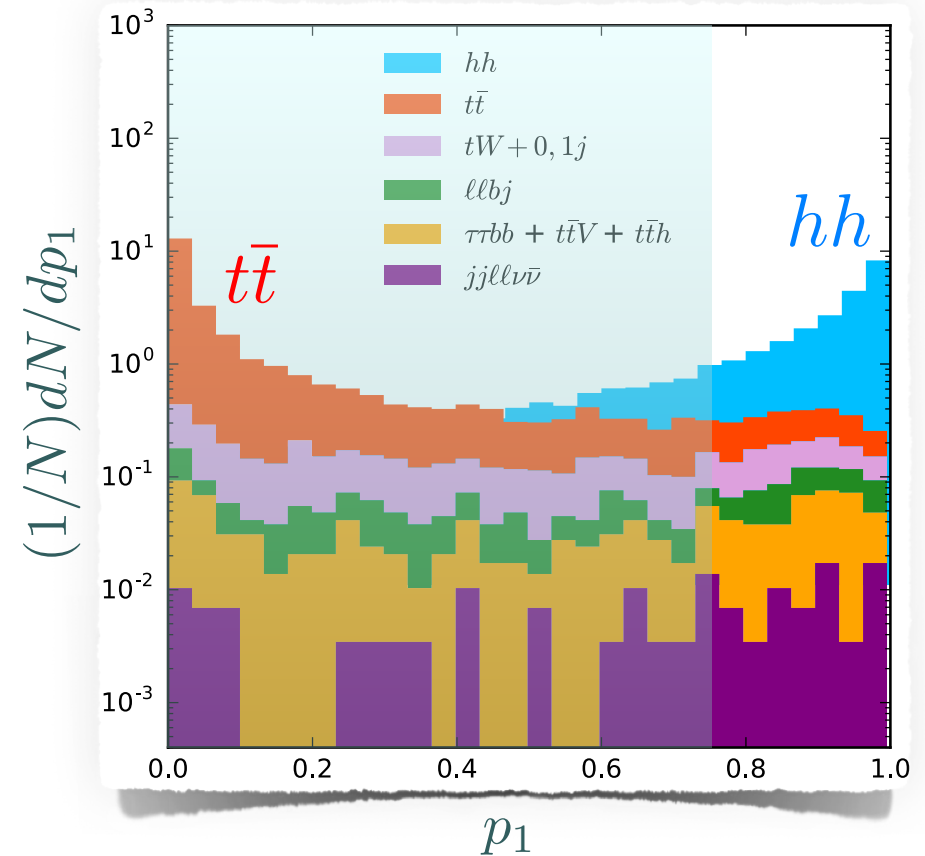
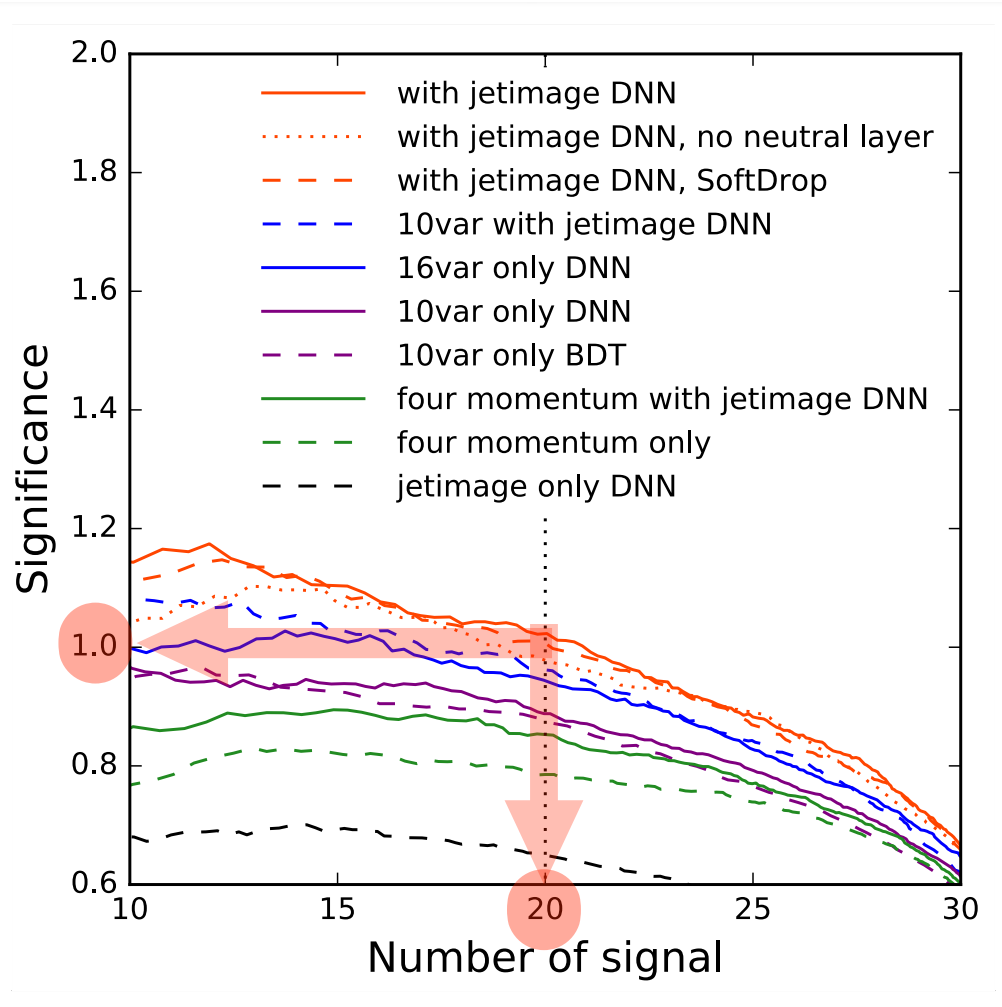
Using Delphes

3 ab^{-1} (14 TeV)

$c_3 = 1$

significance = 0.59σ

CMS-FTR-18-019-PAS

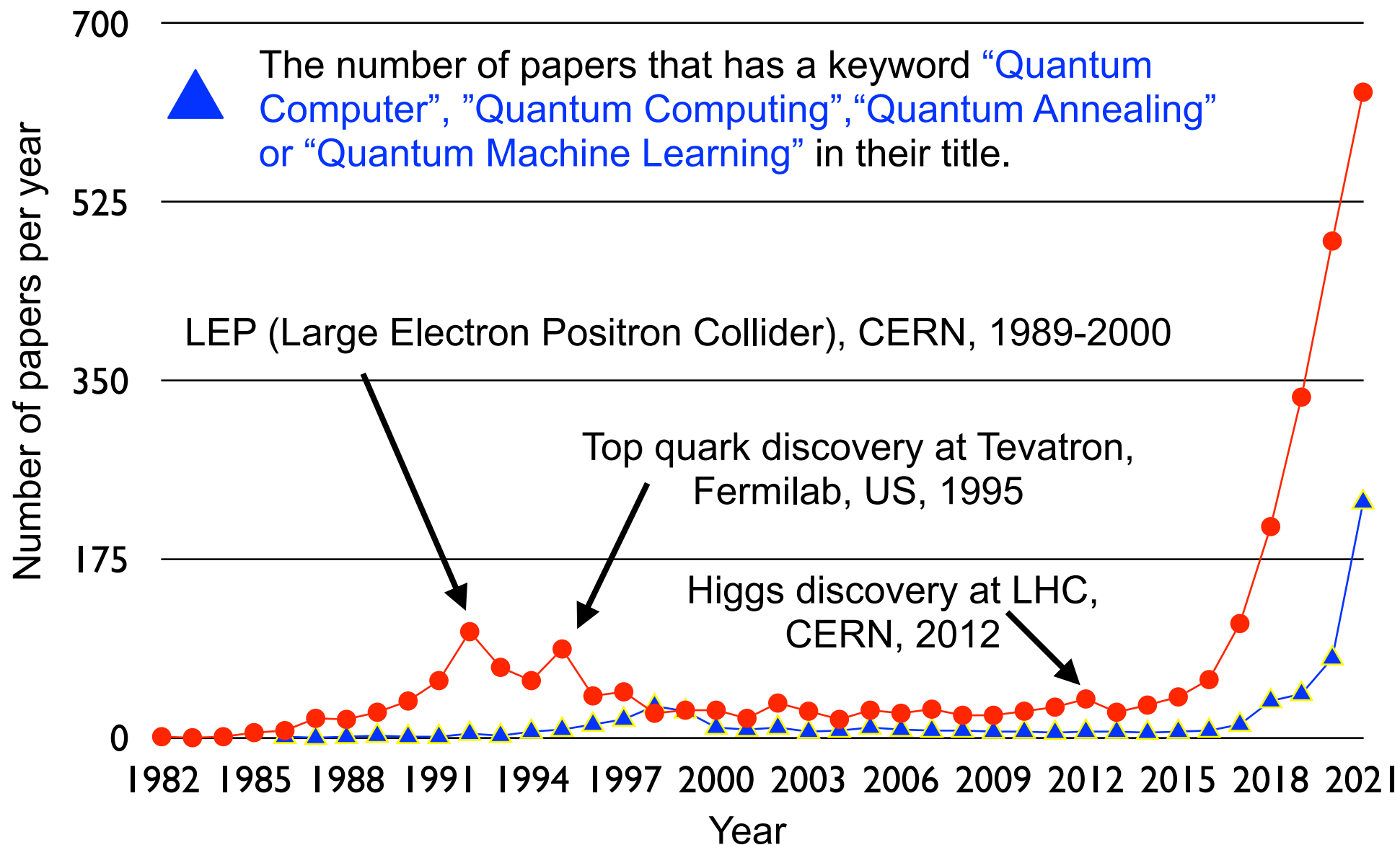


- We cut on NN score and count the number of signal and background events. For $NS=20$, we obtain the significance of ~ 1 .
- The DNN with jet images and high-level variables improves the final significance.

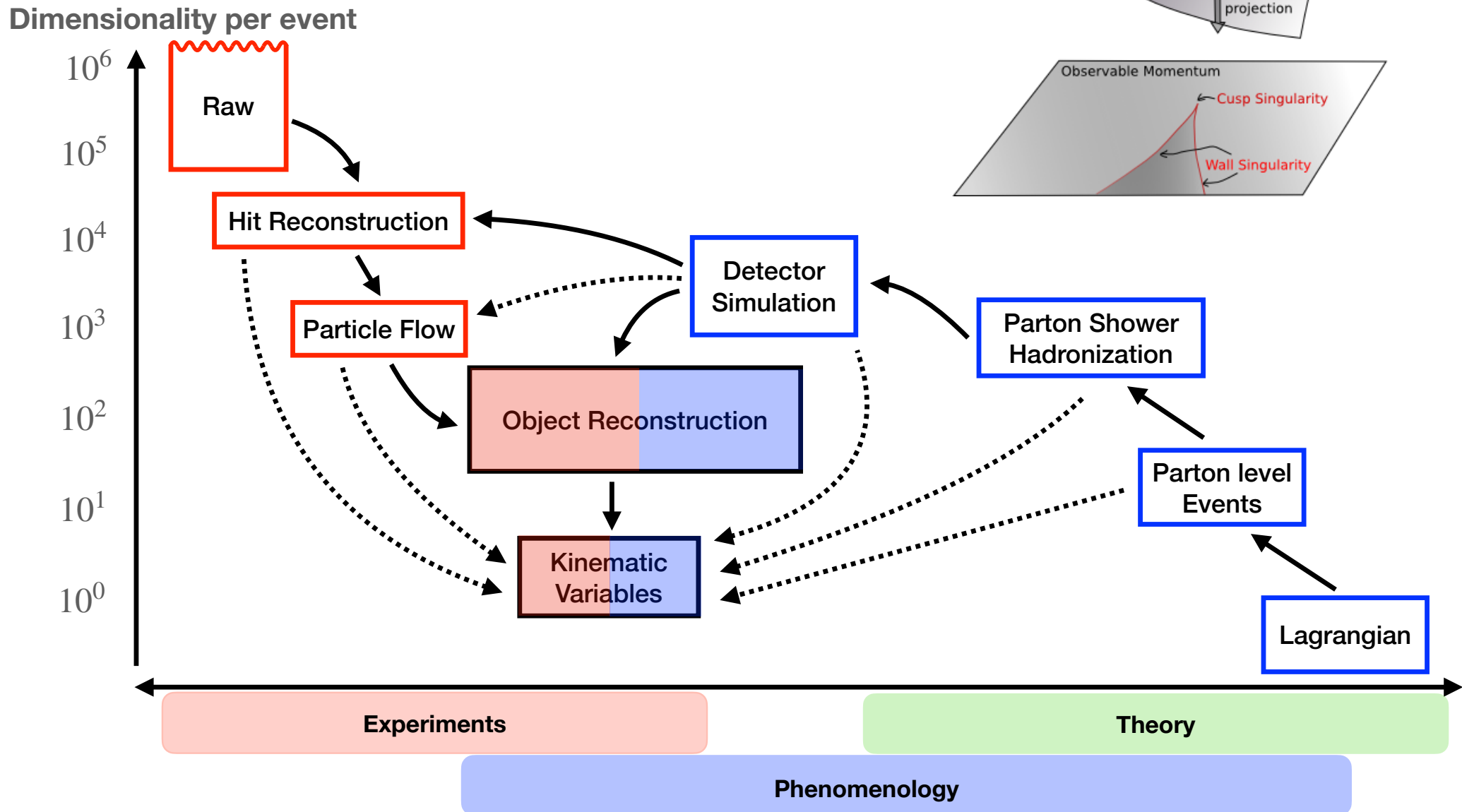
Data is obtained from **InspireHEP**

● The number of papers (in high energy physics) that has a keyword “Machine Learning”, “Deep Learning”, “Artificial Intelligence” or “Neural Networks” in their title.

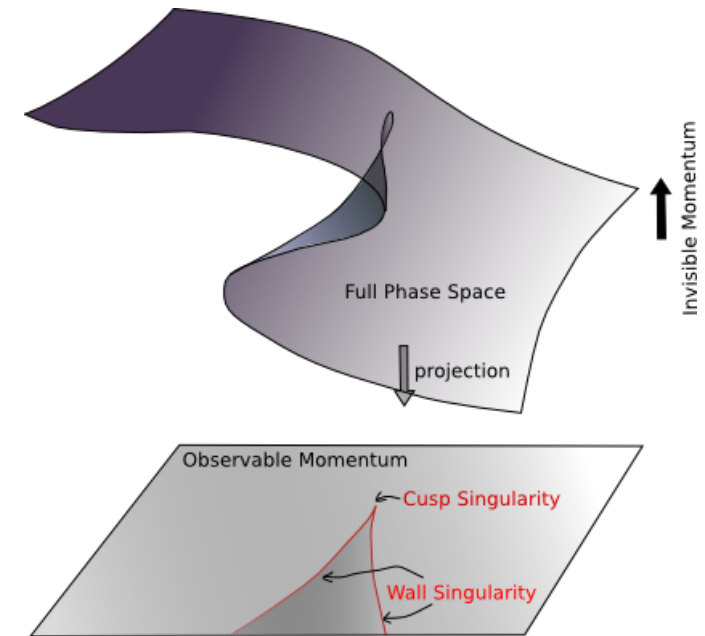
▲ The number of papers that has a keyword “Quantum Computer”, “Quantum Computing”, “Quantum Annealing” or “Quantum Machine Learning” in their title.



Dimensional Reduction in Collider Experiments and Phenomenological Studies



Dimensional Reduction in Collider Experiments and Phenomenological Studies



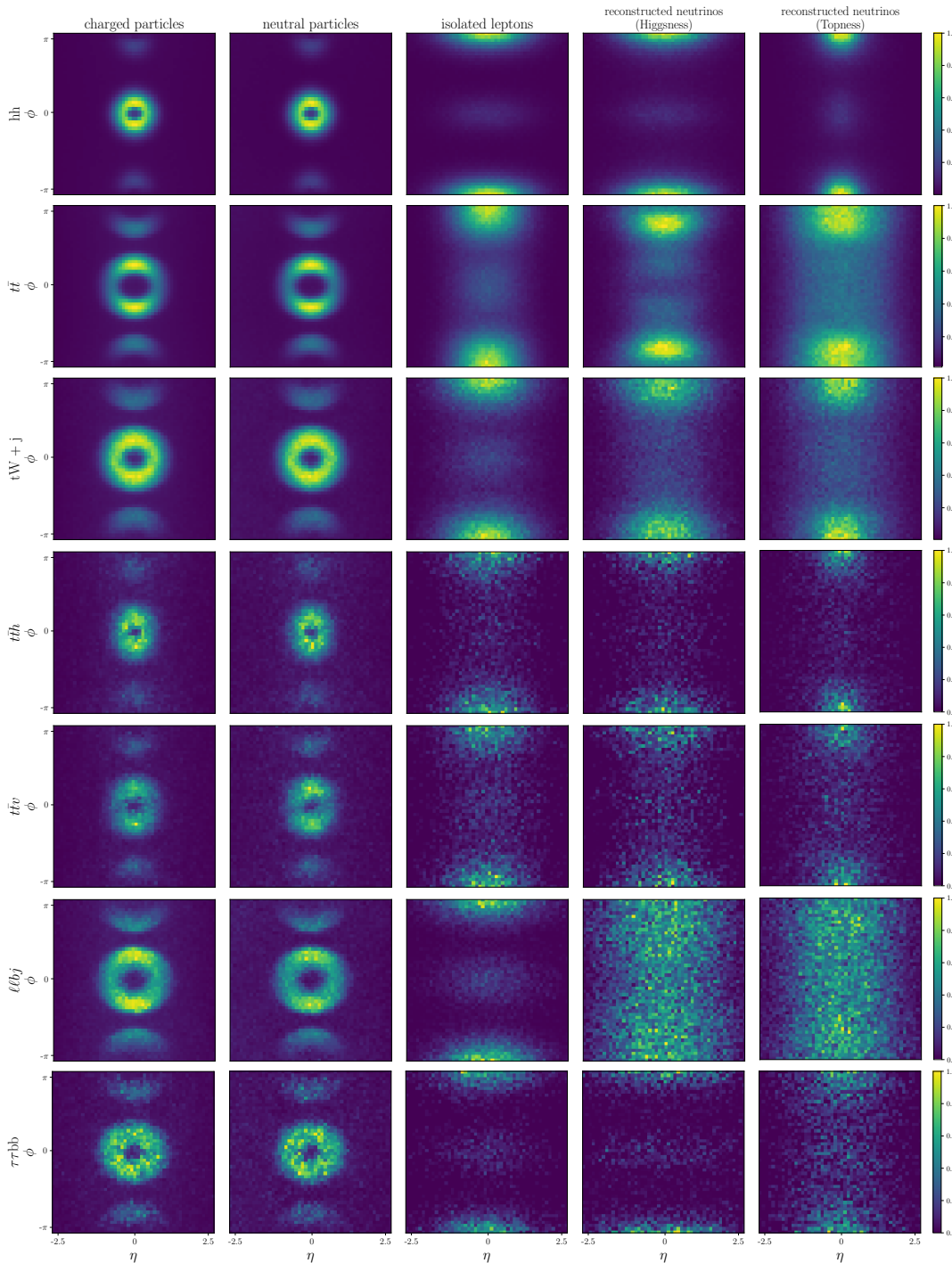
Low-Level Inputs → Kinematic Variables → Physics Task (a)

Low-Level Inputs → Machines → Physics Task (b)

Kinematic Variables → Machines → Physics Task
Low-Level Inputs → Kinematic Variables
Low-Level Inputs → Machines (dashed arrow)

Low-Level Inputs → Machines → Deep Learned Kinematic Variables → Physics Task (d)

Charged Hadrons Neutral Hadrons Isolated Leptons Neutrinos Higgsness Neutrinos Topness



The Di-Higgs Photography

hh

$t\bar{t}$

tW

$t\bar{t}h$

$t\bar{t}V$

$\ell\ell bj$

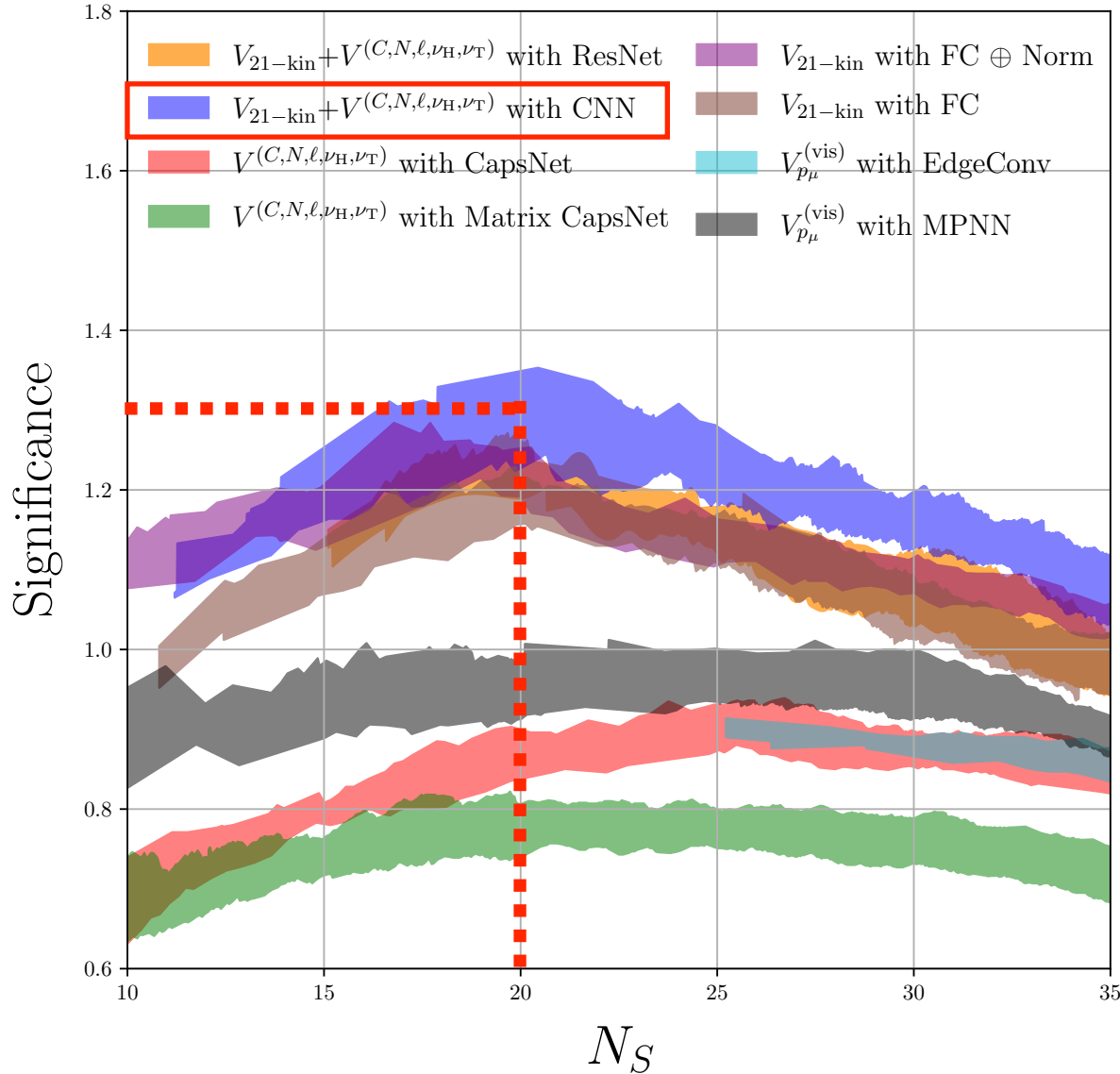
$\tau\tau bb$

- Topness and Higgsness provide **approximate neutrino momenta**, which allow a complete reconstruction of the final state.
- From the W decay, we know that neutrino and lepton distributions must be similar.
- Use the additional **lepton and neutrino images** in NNs. Need enough training data and deep networks to catch correlation of all images.

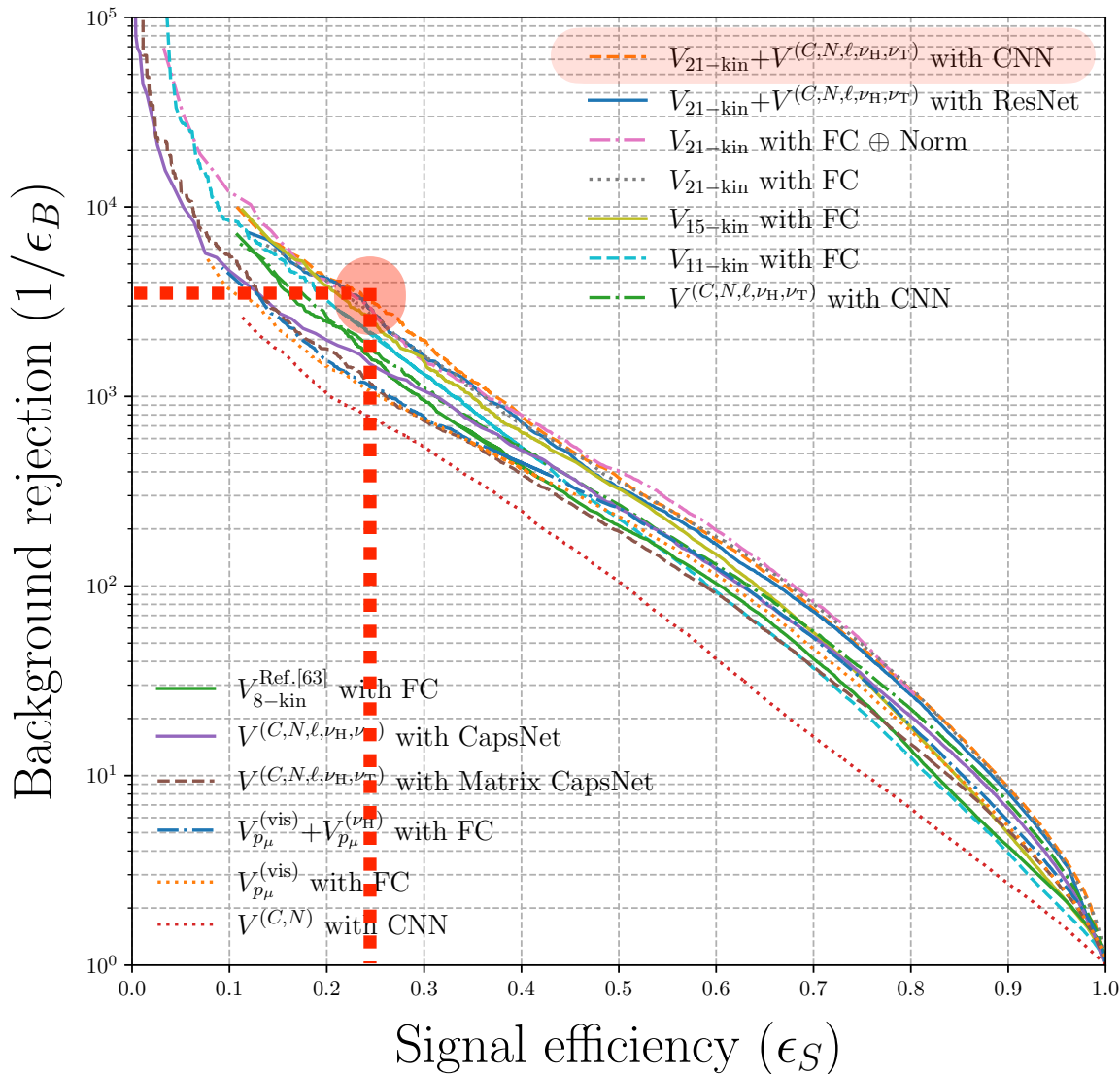
$$V_{\text{image}}^{(C,N,\ell,\nu_H,\nu_T)} = (5 \times 50 \times 50)$$

$$V_{21\text{-kin}} = \{ p_T(\ell_1), p_T(\ell_2), p_{Tbb}, p_{T\ell\ell}, \cancel{p}_T, \Delta R_{bb}, \Delta R_{\ell\ell}, \Delta\phi_{bb,\ell\ell}, m_{\ell\ell}, m_{bb}, \\ \min[\Delta R_{b\ell}], \Delta R_{\nu\nu}^H, m_{\nu\nu}^H, \Delta R_{\nu\nu}^T, m_{\nu\nu}^T, \sqrt{\hat{s}_{\min}^{(bb\ell\ell)}}, \sqrt{\hat{s}_{\min}^{(\ell\ell)}}, M_{T2}^{(b)}, M_{T2}^{(\ell)}, H, T \}$$

$$V_{\text{image}}^{(C,N,\ell,\nu_H,\nu_T)} = (5 \times 50 \times 50)$$



- We have tried various NN architectures (DNN, CNN, ResNet, CapsNet, MPNN etc) with various combinations of input features (four momenta, kinematic variables, images).
- CNN with 21 kinematic variables + 5 images gives the best significance of ~ 1.3 for $NS=20$.
- We have repeated the same runs with 10 different random initializations, which give similar results.
- Images with leptons and neutrinos improve the results slightly.



Cross section after baseline cuts (Before cutting on NN score)

	Cross sections [fb]
hh ($\kappa_3 = 1$)	2.81×10^{-2}
$t\bar{t}$	2.52×10^2
$tW + j$	5.73
$t\bar{t}h$	2.53×10^{-1}
$t\bar{t}V$	3.18×10^{-1}
$\ell\ell bj$	1.61
$\tau\tau bb$	1.49×10^{-2}

$\sim 97\%$
 $\sim 2\%$

$$\sigma_{bknd}/\sigma_{hh} \approx 9250$$

After cutting on NN score
for CNN with 21 variables + 5 images

NS=20 and NB=220

$$\sigma_{bknd}/\sigma_{hh} \approx 11$$

background rejection = $1/\epsilon_B = 3,500$

signal efficiency = $\epsilon_S = 0.23$

$t\bar{t}$: 44 %

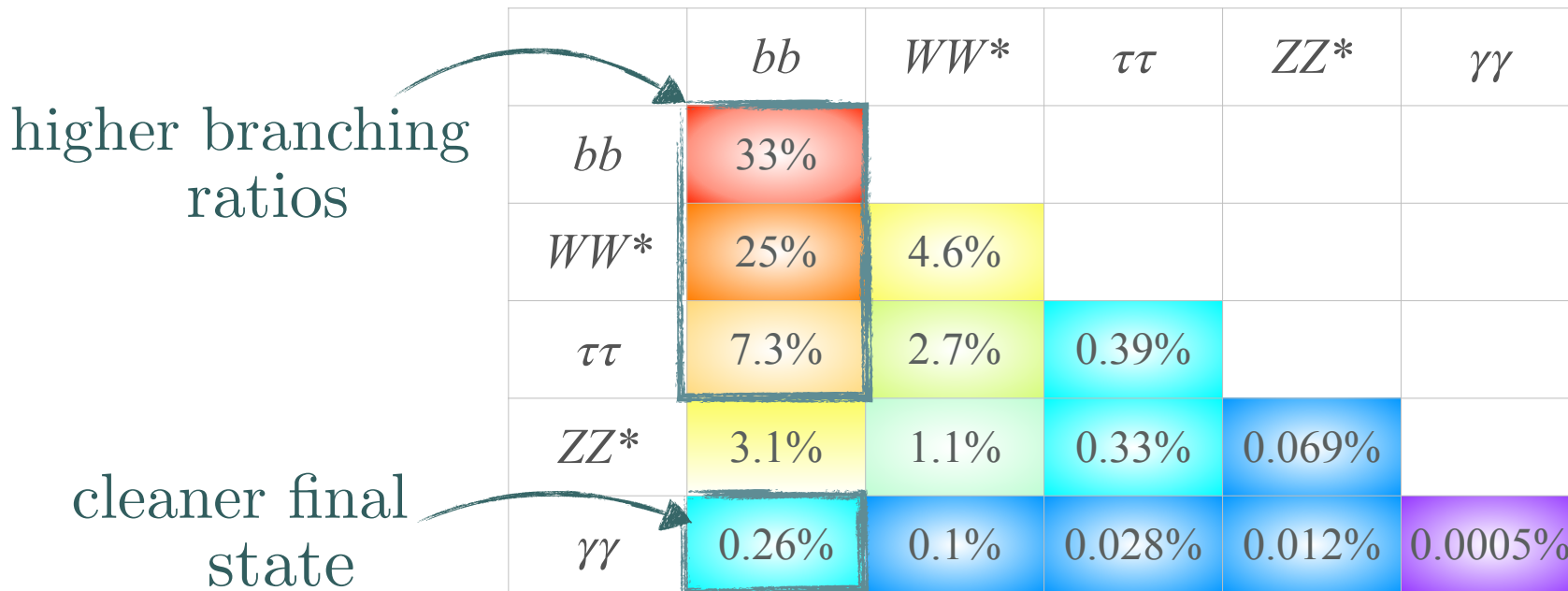
$t\bar{t}h + t\bar{t}V$: 22 %

tW : 24 %

$\ell\ell bj + \tau\tau bb$: 10 %

Combination of various channels

$$\sigma(hh)_{SM}^{NNLO} \simeq 40.7 \text{ fb} \quad (14 \text{ TeV})$$



1902.00134	Statistical-only		Statistical + Systematic	
	ATLAS	CMS	ATLAS	CMS
$HH \rightarrow b\bar{b}b\bar{b}$	1.4	1.2	0.61	0.95
$HH \rightarrow b\bar{b}\tau\tau$	2.5	1.6	2.1	1.4
$HH \rightarrow b\bar{b}\gamma\gamma$	2.1	1.8	2.0	1.8
$HH \rightarrow b\bar{b}VV (ll\nu\nu)$	-	0.59	-	0.56
$HH \rightarrow b\bar{b}ZZ (4l)$	-	0.37	-	0.37
combined	3.5	2.8	3.0	2.6
	Combined 4.5		Combined 4.0	

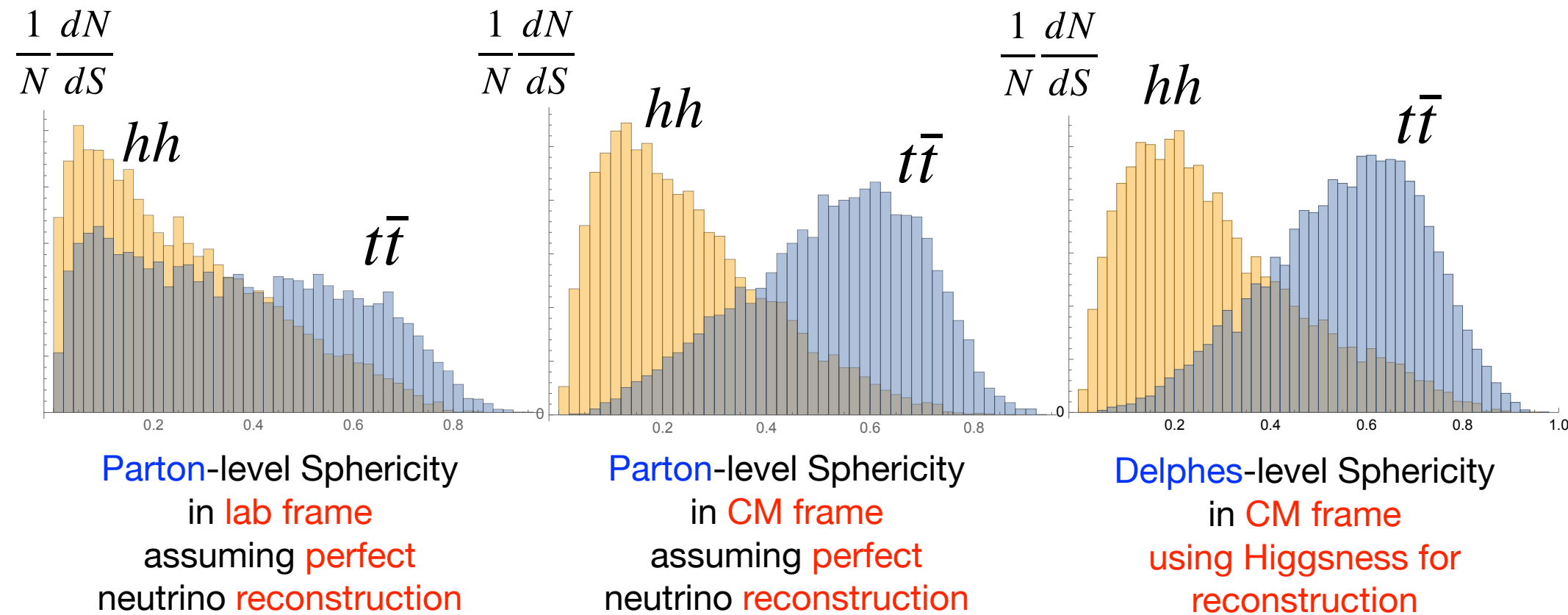
4 σ expected for ATLAS+CMS!

- These measurements are challenged by a low $\sigma(hh)$ and small branching ratios (BR).
- No single channel is expected to reach 3 sigma at HL-LHC.
- The combination of different channels is crucial. $bbWW$ has good potential for further improvement.

Channel	Statistical only		Statistical + Systematic	
	ATLAS	CMS	ATLAS	CMS
$hh \rightarrow b\bar{b}b\bar{b}$	1.4	1.2	0.61	0.95
$hh \rightarrow b\bar{b}\tau^+\tau^-$	2.5	1.6	2.1	1.4
$hh \rightarrow b\bar{b}\gamma\gamma$	2.1	1.8	2.0	1.8
$hh \rightarrow b\bar{b}VV(\ell\ell\nu\nu)$	-	0.59	-	0.56
$hh \rightarrow b\bar{b}ZZ(4\ell)$	-	0.37	-	0.37
combined	3.5	2.8	3.0	2.6
	combined 4.5		combined 4.0	
combined with the new results on $hh \rightarrow b\bar{b}VV(\ell\ell\nu\nu)$ in this study	3.8	3.0	3.2	2.8
	combined 4.8		combined 4.2	

- We roughly reproduce all significances in the other channels following 1902.00134, and combine the new/updated result from $bb\ell\ell$ channel.
- The significances are added in quadrature, and the channels are treated as uncorrelated, assuming that the systematic uncertainties such as the theory uncertainties and the luminosity uncertainty, have little impact on the individual results.
- We assume 10% reduction in the signal significance, to take into account the systematics in the $bb\ell\ell$ channel,

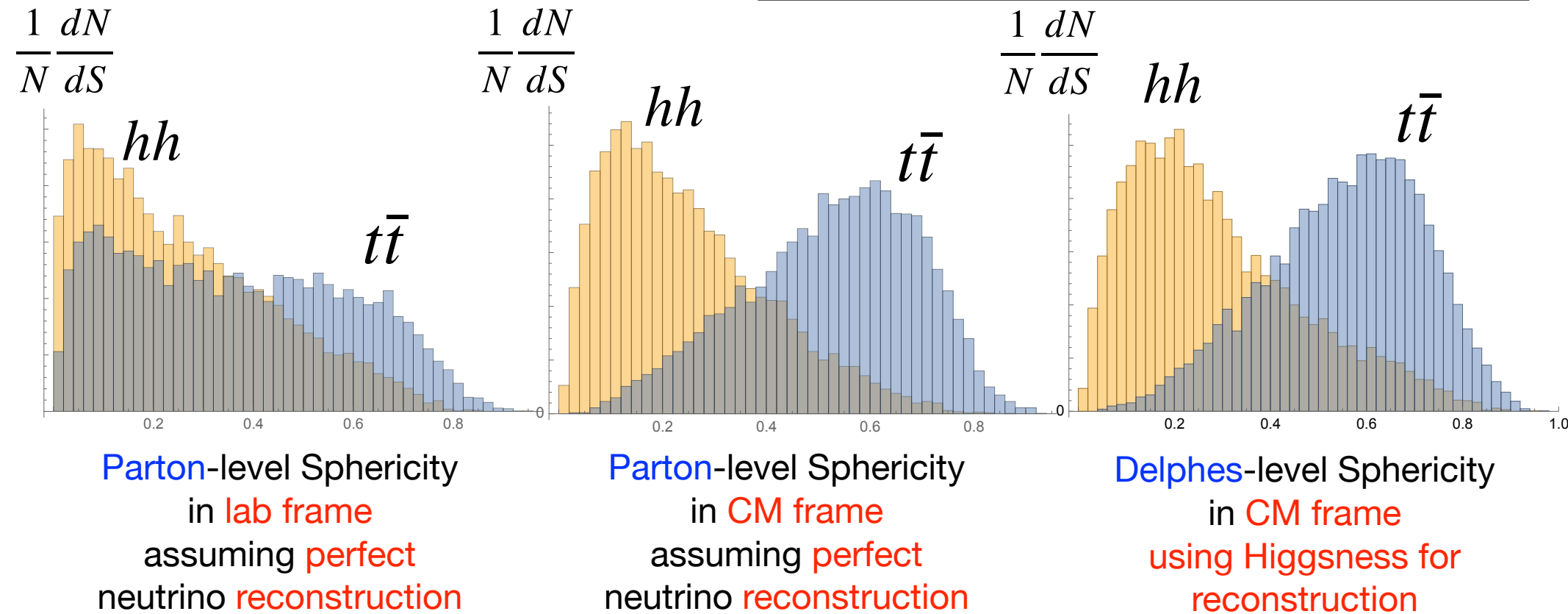
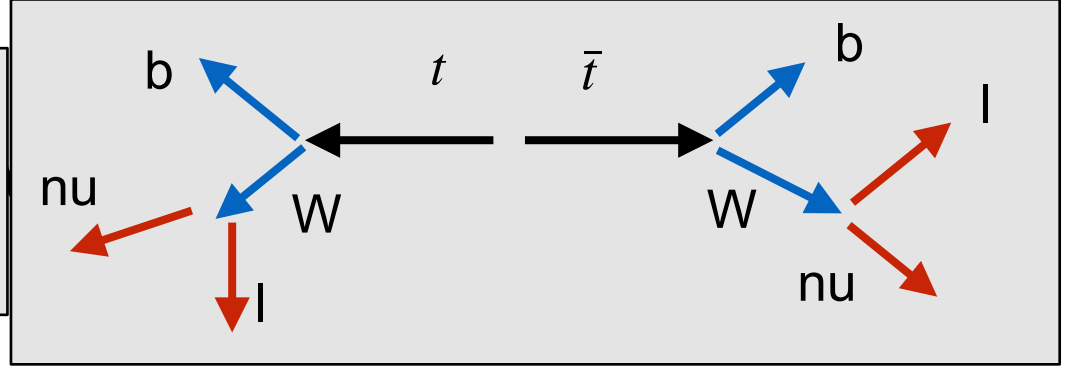
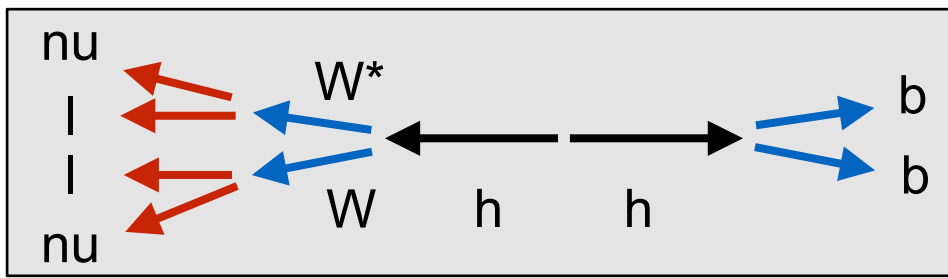
Shape variables



$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2}, \quad \lambda_1 + \lambda_2 + \lambda_3 = 1$$

$$S = \frac{3}{2}(\lambda_2 + \lambda_3), \quad \lambda_1 \geq \lambda_2 \geq \lambda_3$$

- $S \rightarrow 0$: pencil-like event
- $S \rightarrow 1$: isotropic event
- Results using Topness are similar.



$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2}, \quad \lambda_1 + \lambda_2 + \lambda_3 = 1$$

$$\lambda_1 \geq \lambda_2 \geq \lambda_3$$

$$S = \frac{3}{2}(\lambda_2 + \lambda_3)$$

- $S \rightarrow 0$: pencil-like event
- $S \rightarrow 1$: isotropic event
- Results using Topness are similar.

HH: semi-leptonic channel

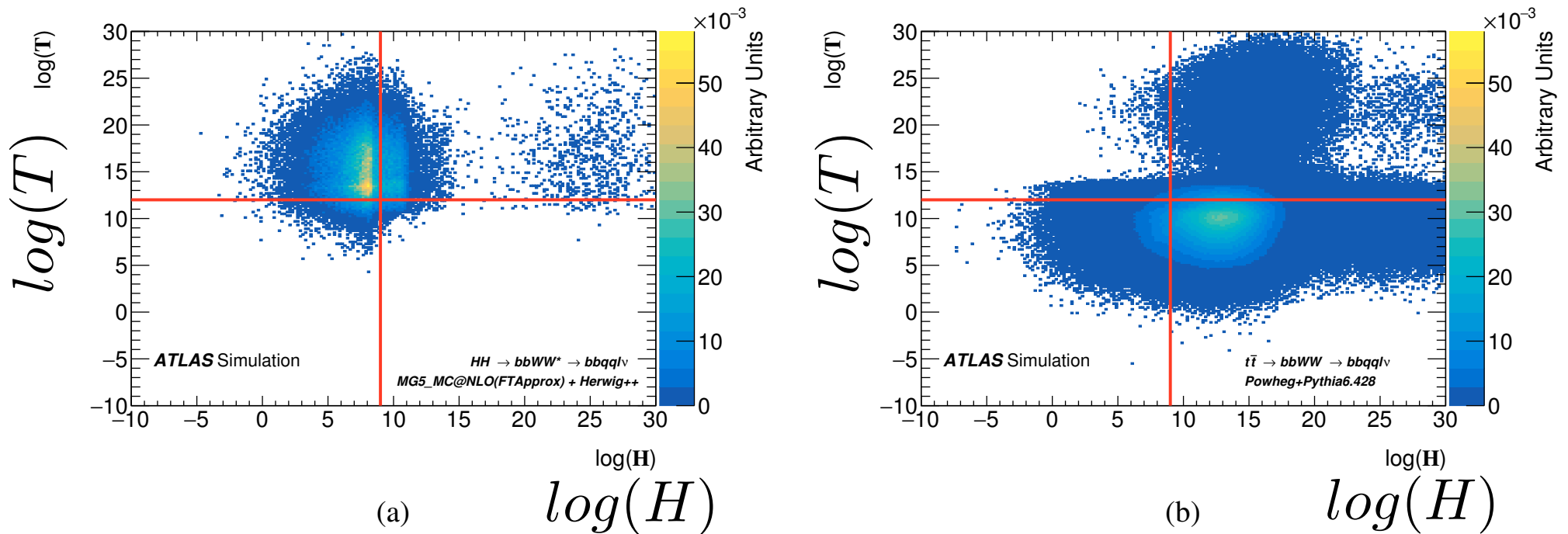
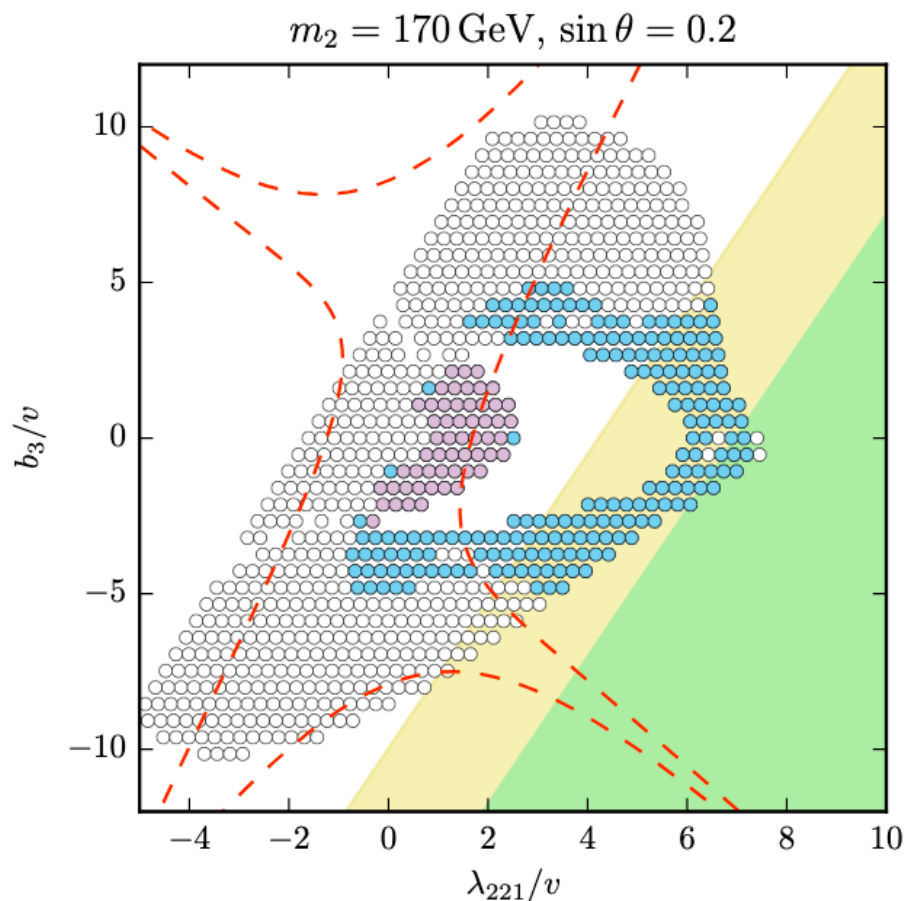


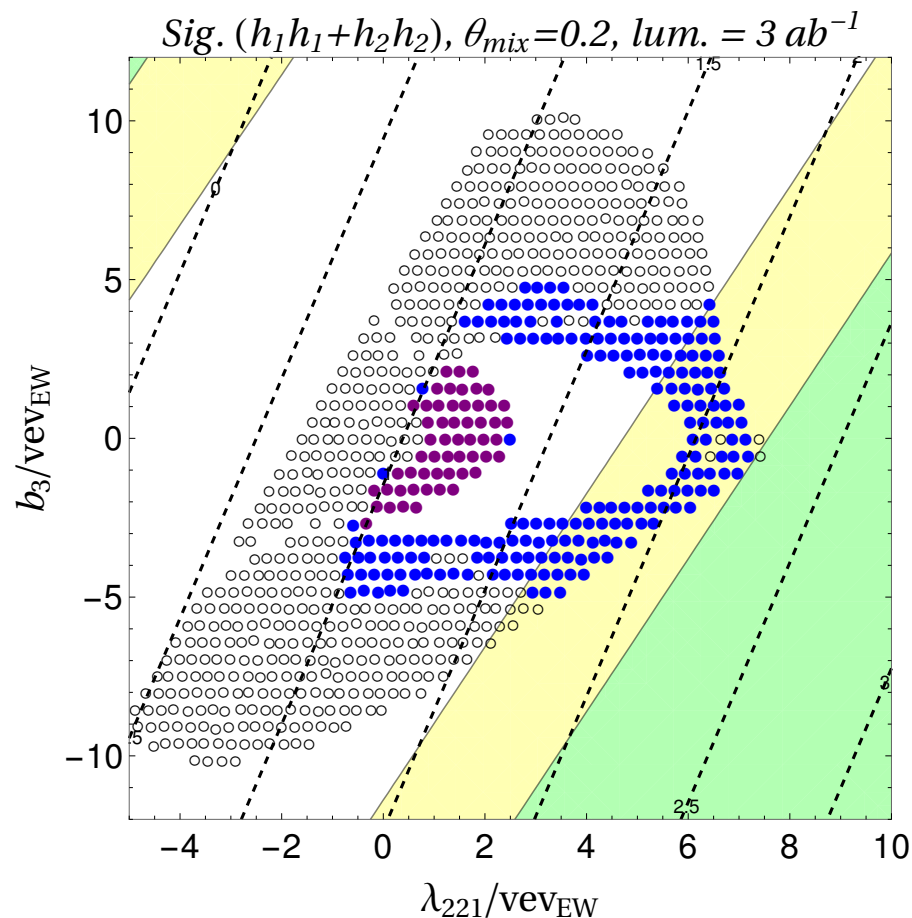
Figure 2: Distribution of Higgsness and Topness in a two-dimensional plane ($\log(H)$, $\log(T)$) for simulated signal $HH \rightarrow bbWW^* \rightarrow bbqql\nu$ (a) and background $t\bar{t} \rightarrow bbWW \rightarrow bbqql\nu$ (b) events without selection requirements. The signal sample is generated with MG5_MC@NLO(FTApprox) + Herwig++, while the background sample is generated with Powheg + Pythia6.428. The distributions are normalised to unit area. Red lines are drawn to give a visible reference for a possible separation between signal and background.

ATL-PHYS-PUB-2019-040

Exclusion (yellow) and discovery reach (green) in $h_2 h_2 \rightarrow 2j 3l + \text{met}$ channel at the HL-LHC.

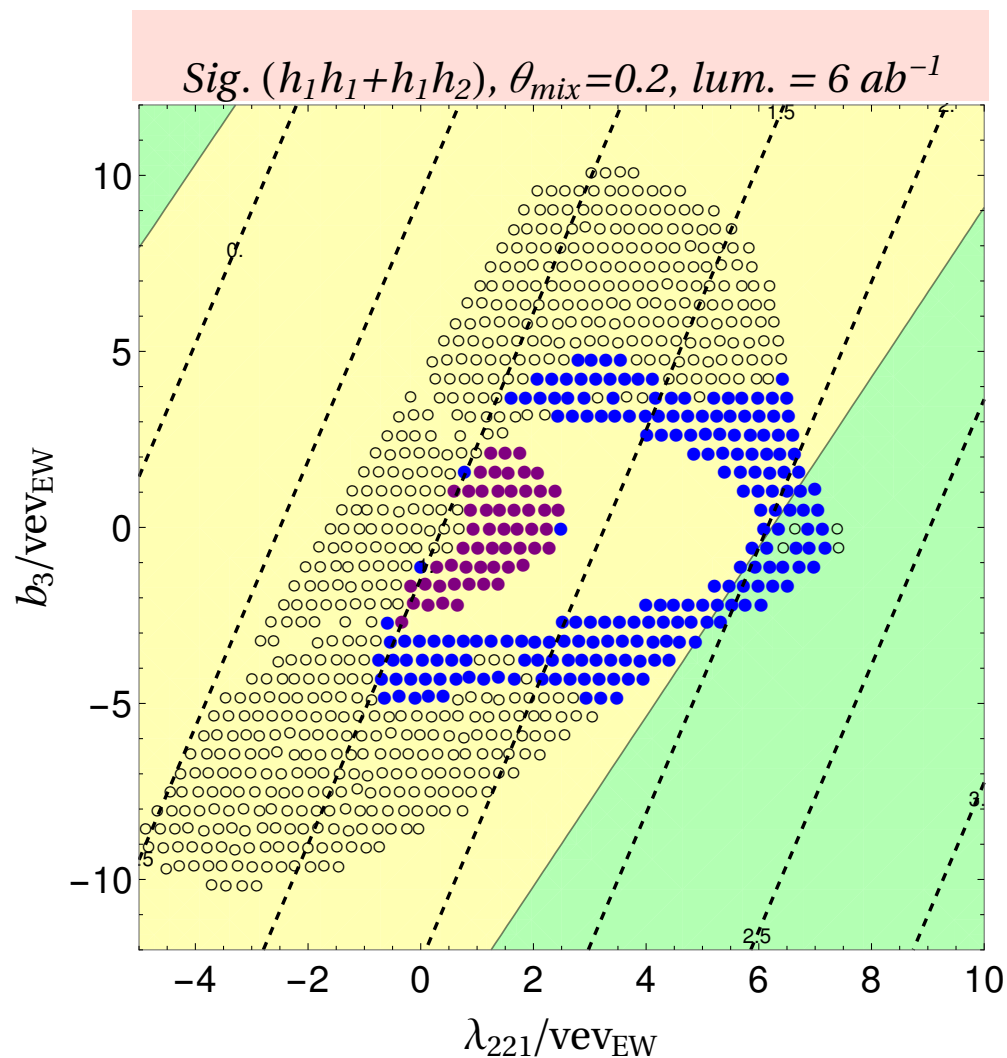
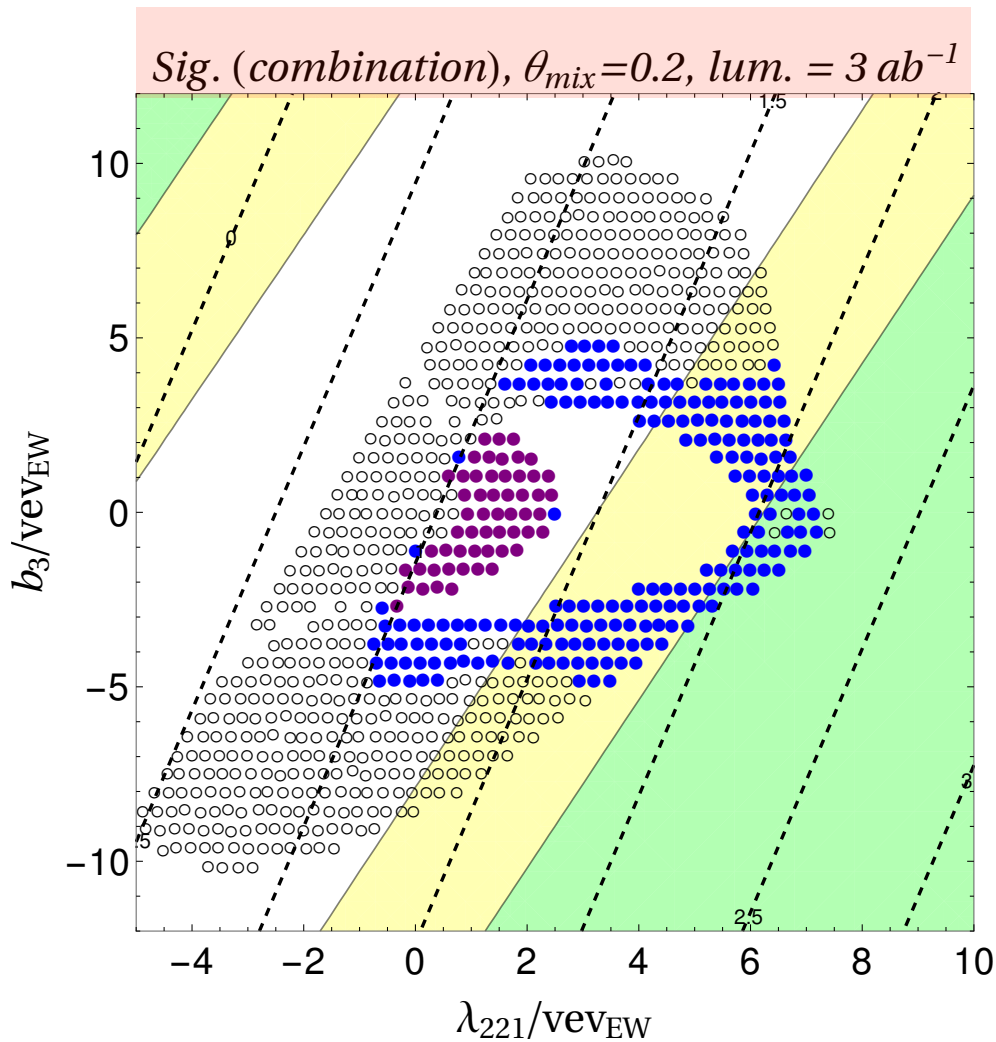


Exclusion (yellow) and discovery reach (green) in $h_1 h_2 + h_1 h_1 \rightarrow bb 2l + \text{met}$ channel at the HL-LHC.



Blue points feature an EWPT with $\phi_h(T_c)/T_c \geq 1$ for some value of $b_4 > 0.01$ utilizing the one-loop daisy-resummed thermal effective potential. **Purple points additionally feature a strong first-order electroweak phase transition** as predicted by the gauge-invariant high-T approximation (which drops the Coleman-Weinberg potential and is thus only applied to regions with tree-level vacuum stability). Strong electroweak phase transitions are typically correlated with sizable values of λ_{221} .

Exclusion (yellow) and discovery reach (green) for combining $h_2 h_2 \rightarrow 2j 3l + \text{met}$ and $h_1 h_2 + h_1 h_1 \rightarrow bb 2l + \text{met}$ channel at the HL-LHC.
 (Mixing angle=0.2 and $m_H=170$ GeV)



Further improvement possible with image inputs

Alhazmi, Kim, Kong, Lewis, preliminary

Summary

- Higgs self couplings are important to understand the nature of electroweak symmetry breaking. The HL-LHC will have a sensitivity to the measurement of the triple Higgs coupling via double Higgs production.
- Double Higgs production is challenging due to small signal cross section / large SM backgrounds, which requires combination of multiple channels for discovery.
- $bbWW$ dilepton channel is one of difficult channels due to strong correlation among many kinematic variables.
- Multivariate analysis could benefit from deep neural networks using jet images and novel kinematic variables such as Topness / Higgsness with mass information.
- Topness/Higgsness provide approximate momenta of the missing neutrinos, which allow to study the event shape variables.
- $bbWW$ channel could make a significant contribution in the combination of multiple channels for the triple Higgs coupling measurement.
- Application in the semi-leptonic channel, and non-resonant HH production in the singlet extension of SM.
- Investigating Spiking Neural Network with double Higgs production.

Additional slides

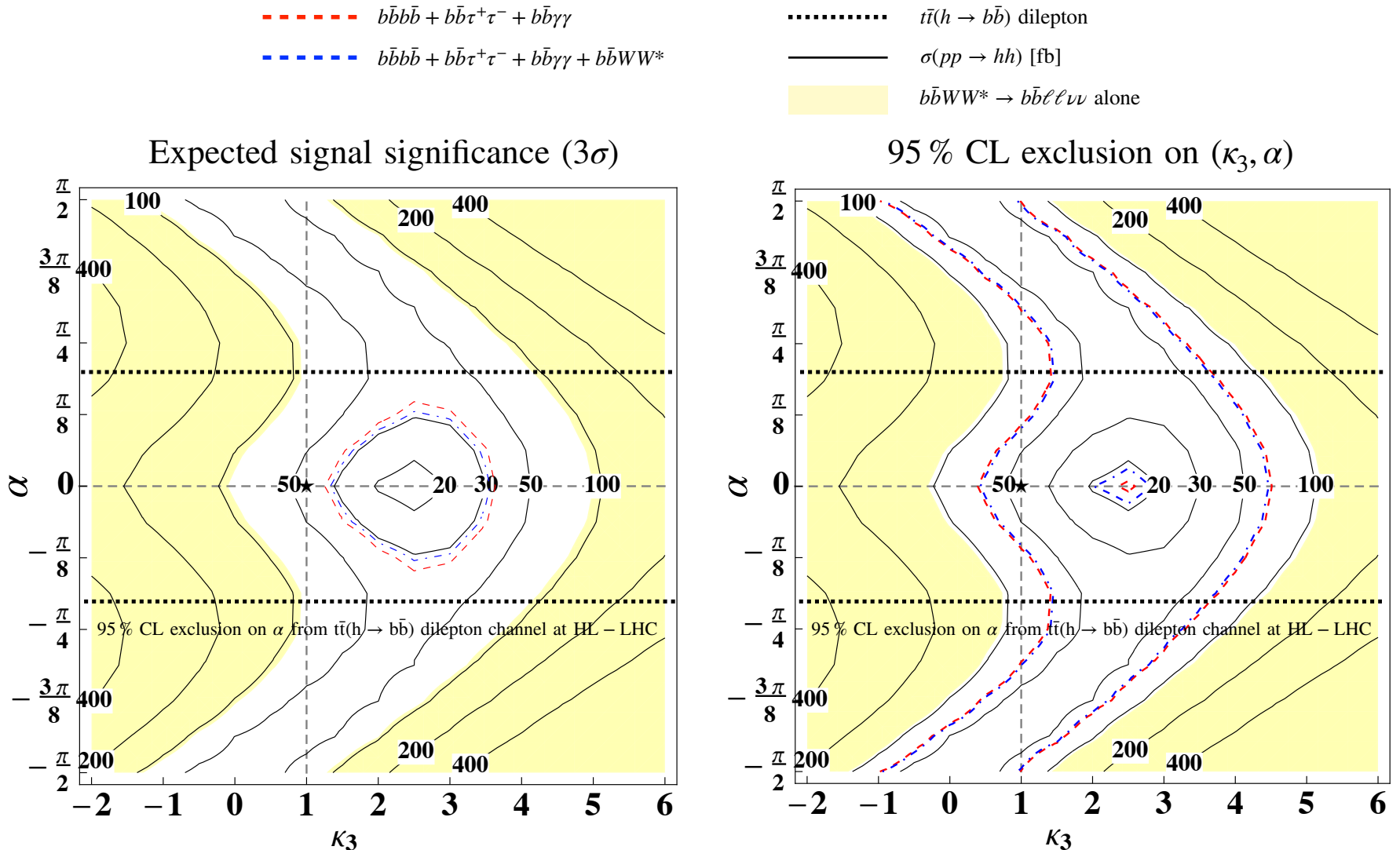


Figure 15. Expected 3σ significance of observing Higgs boson pair production (left) and 95% C.L. exclusion (right) in the (κ_3, α) plane at the HL-LHC with 3 ab^{-1} . We used the binned log-likelihood analysis with statistical uncertainties only, assuming the same efficiencies for all (κ_3, α) values as one for $(\kappa_3, \alpha) = (1, 0)$ (SM point denoted by \star). Contours of the double Higgs production cross section (in fb) are shown in black-solid curves. The yellow shaded region is obtained using results in this study for the dilepton channel ($hh \rightarrow b\bar{b}WW^* \rightarrow b\bar{b}\ell\ell\nu\bar{\nu}$). The red dashed curve is obtained combining three channels, $b\bar{b}b\bar{b} + b\bar{b}\tau^+\tau^- + b\bar{b}\gamma\gamma$ following Ref. [45], while the blue dashed curve includes all four channels. The horizontal-black dotted line represents a sample 95% exclusion on the CP angle from the dilepton channel of $t\bar{t}h$ production with $h \rightarrow b\bar{b}$ [103], $|\alpha| \lesssim 35^\circ$.

Pile-up

- Soft Drop method: a powerful pile-up mitigation technique
- Tried without neutral hadrons (for which pile up effect would be worse)
- We adopt the definition for a missing transverse momentum from ATLAS, which excludes contributions from soft neutral particles

$$\vec{\cancel{P}}_T = - \left(\sum \vec{p}_{T\ell} + \sum \vec{p}_{T\gamma} + \sum \vec{p}_{Tj} + \sum \vec{p}_{T(\text{track})} \right)$$

Here the last term is added to consider unused soft tracks. These tracks are required to have $p_T > 0.4$ GeV, $|\eta| < 2.5$ and transverse (longitudinal) impact parameter $|d_0| < 1.5$ mm ($|z_0 \sin \theta| < 1.5$ mm). To reduce effects from pile-up, we only use particles which have track information.

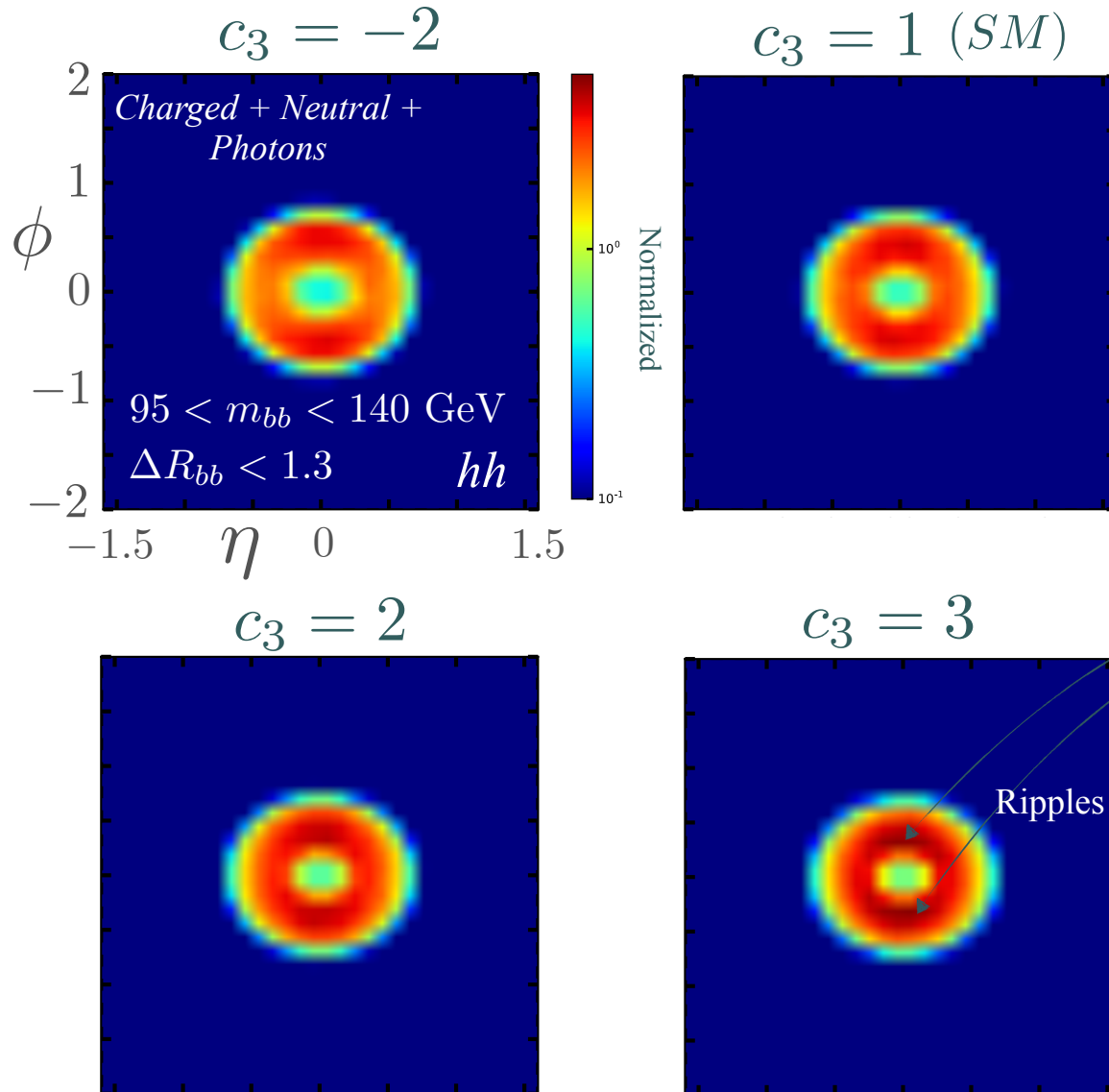
Pile-up

$$\text{Soft Drop Condition: } \frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^\beta$$

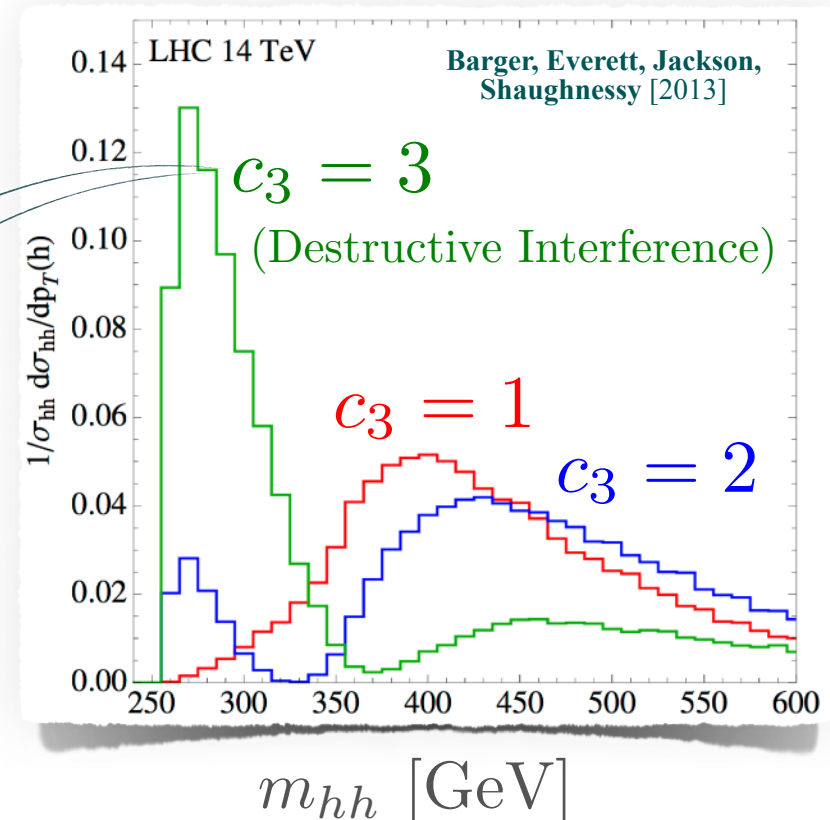
In order to examine the effects of pile-up, we use several methods as follows. In the first method, we use the Soft Drop algorithm [58] to remove soft jet activity which is exacerbated by pile-up. We set $\beta = 0$ and $z_{\text{cut}} = 0.1$ with $R = 1.2$ anti- k_T clustered fatjets. Then we select the closest fatjet to the $b\bar{b}$ momentum in the η - ϕ plane and replace the particle flow data with the charged and neutral jet constituents of the selected fatjet. Soft Drop does not affect the jet images and retains the same shapes as in Fig. 8. In second method, we remove the neutral jet image layer in the analysis. Unlike charged particles, which can be cleaned up from pile-up relatively easily by checking the longitudinal vertex information [106], neutral particles cannot be treated the same way and suffer from non-removable pile-up effects. The corresponding results with these two pile-up mitigation methods are also shown in Fig. 11 with the red dotted line labelled “16var with jetimage DNN, SoftDrop” and the red, dashed line labelled “16var with jetimage DNN, no neutral layer”, respectively.

Shifting the Higgs triple coupling c_3

- How does the hh image vary by shifting the Higgs triple coupling c_3 ? $\mathcal{L} \supset -c_3 \frac{m_h^2}{2v} h^3$



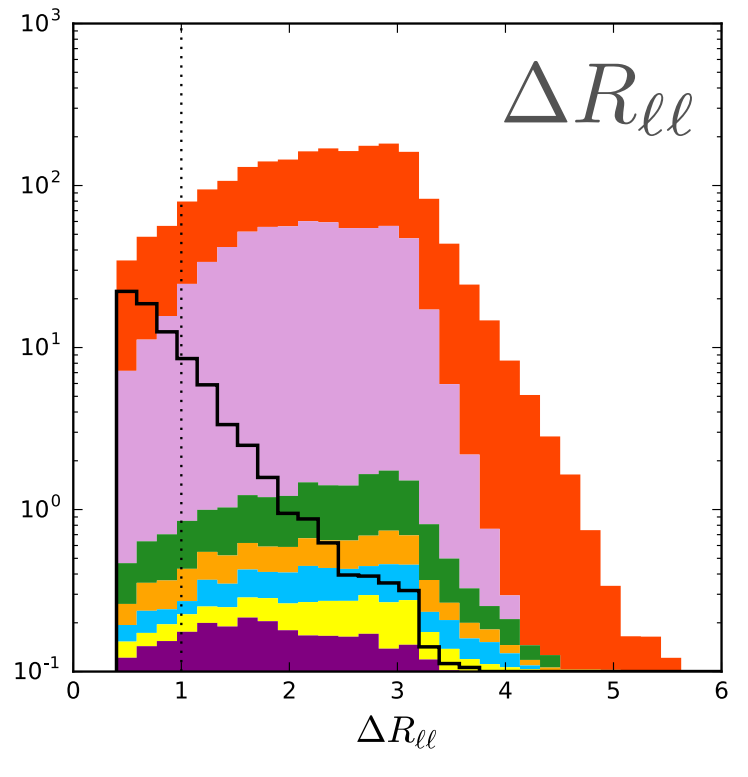
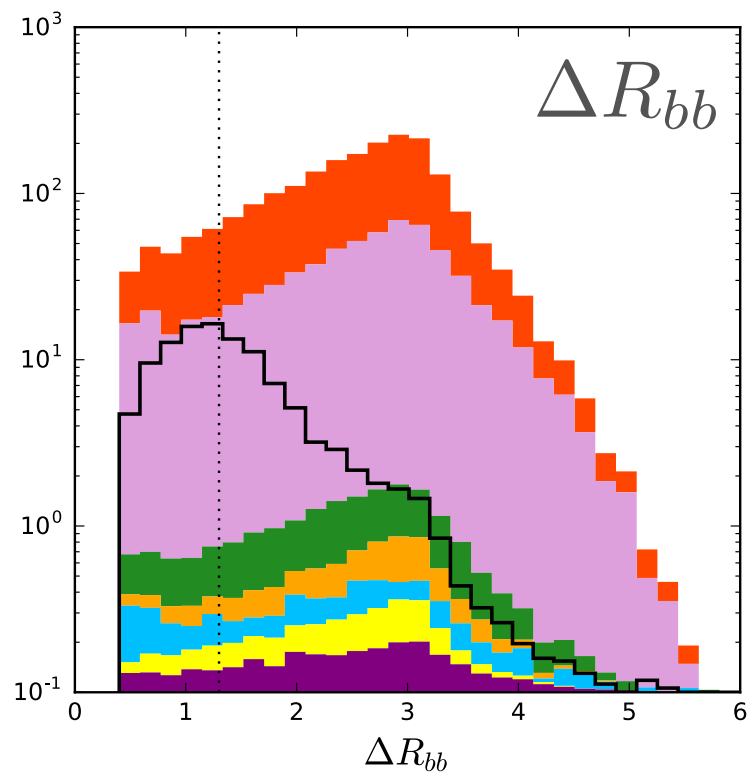
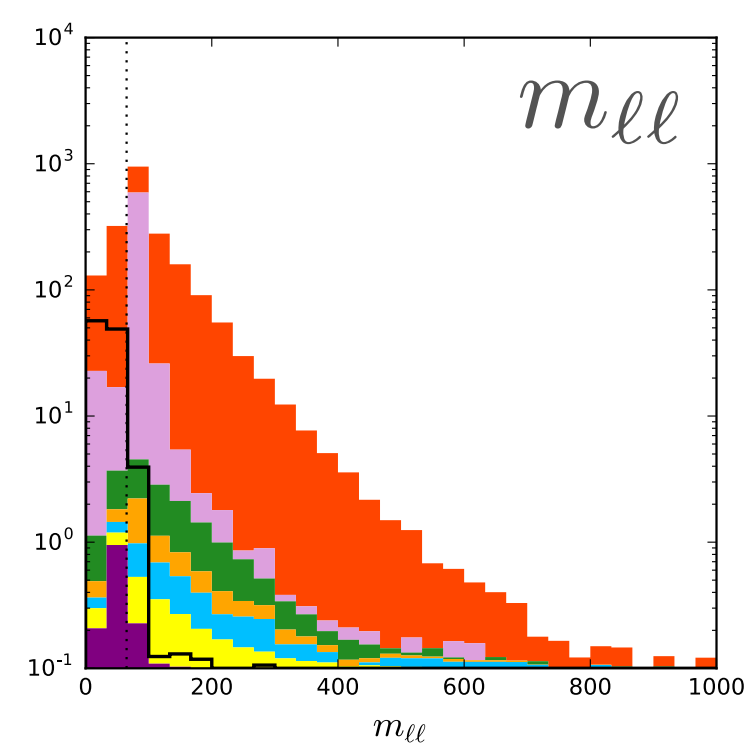
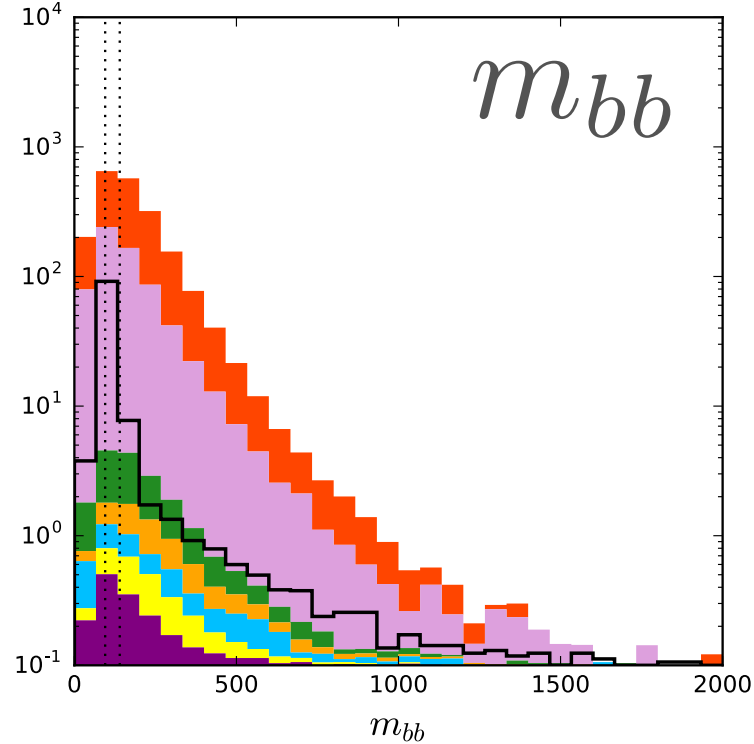
- Interestingly, the gradient of images change as moving into the region of a destructive interference.
- We are working on how much neural network can be sensitive to this change (we don't know yet, sorry).

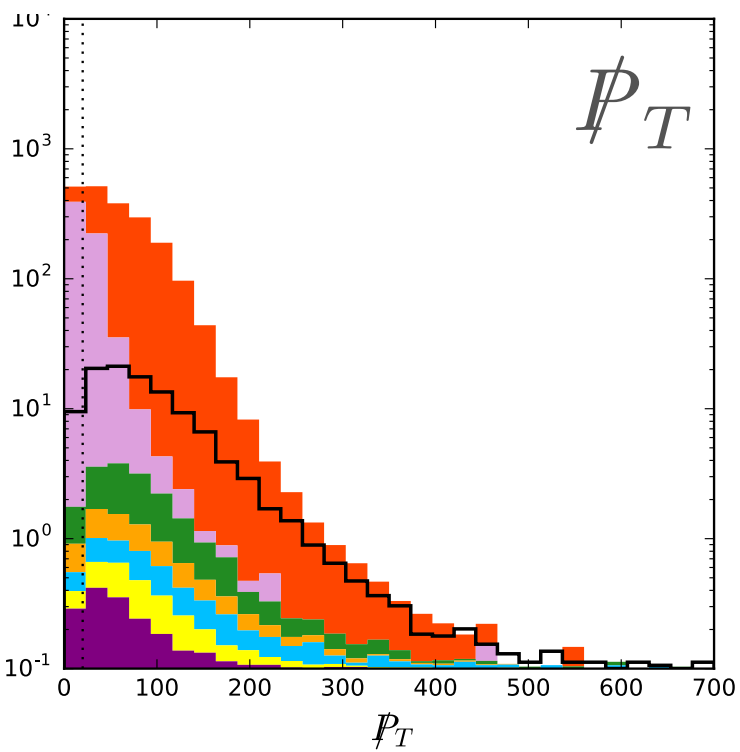
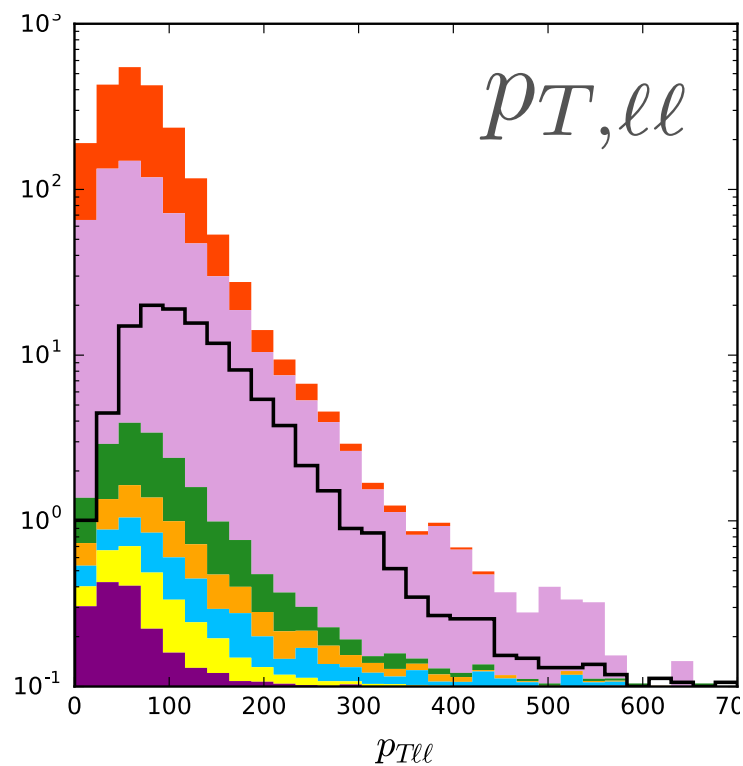
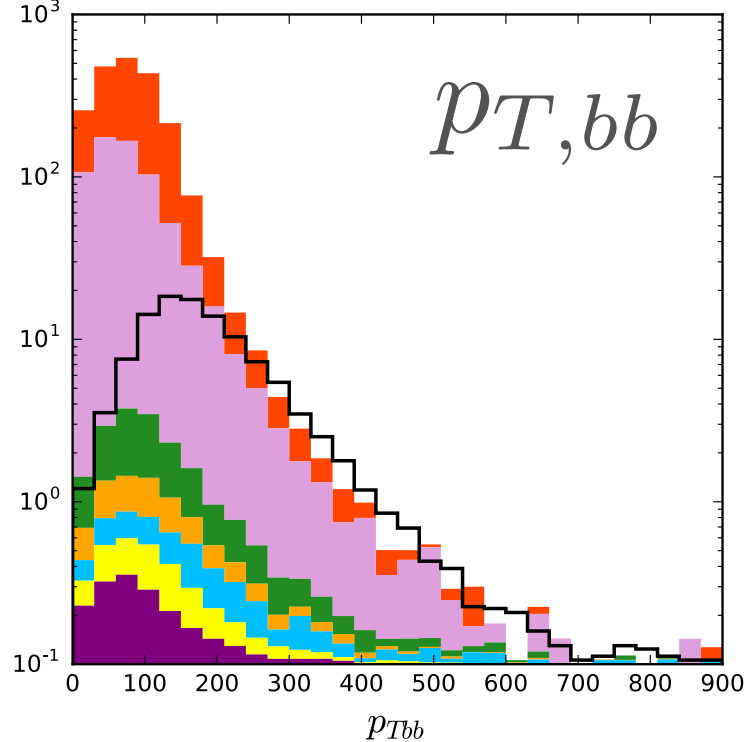
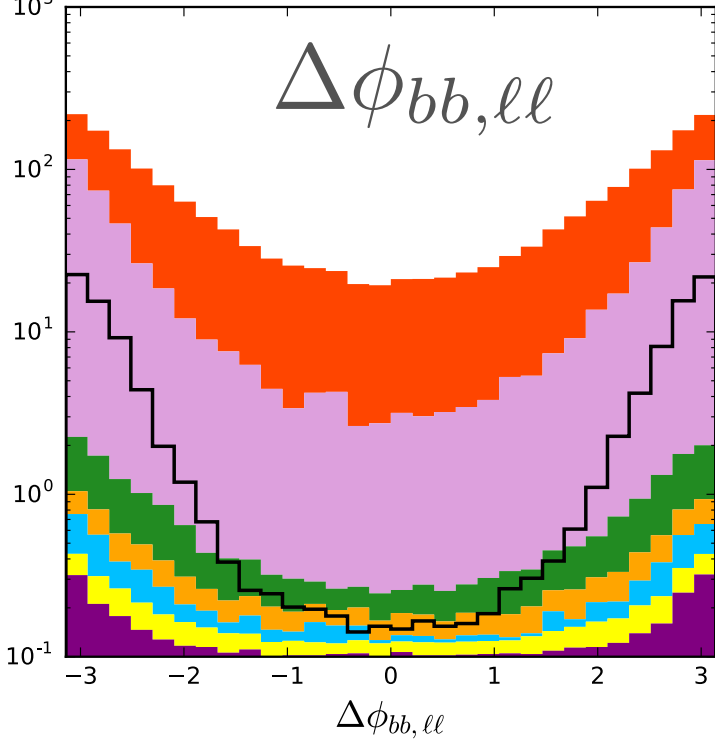


Some modification

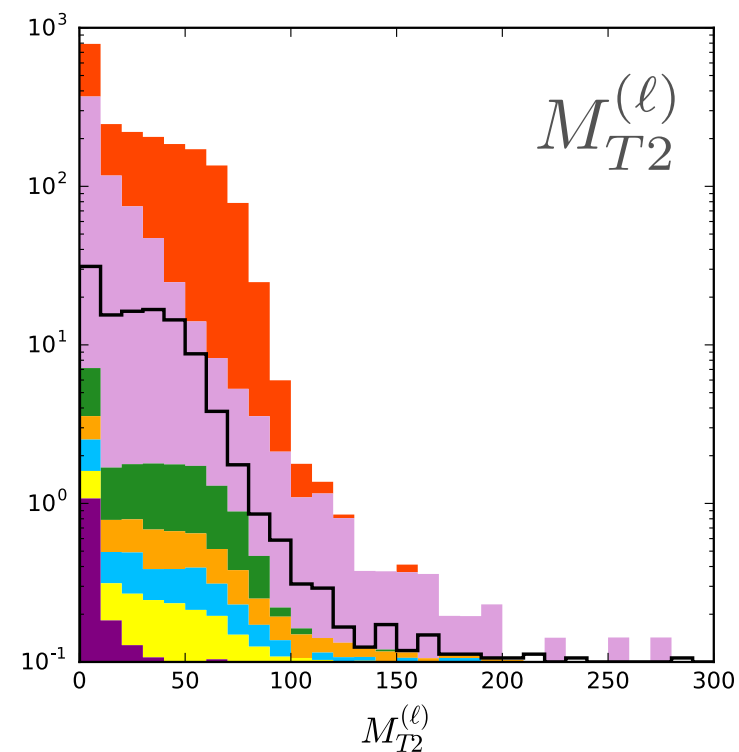
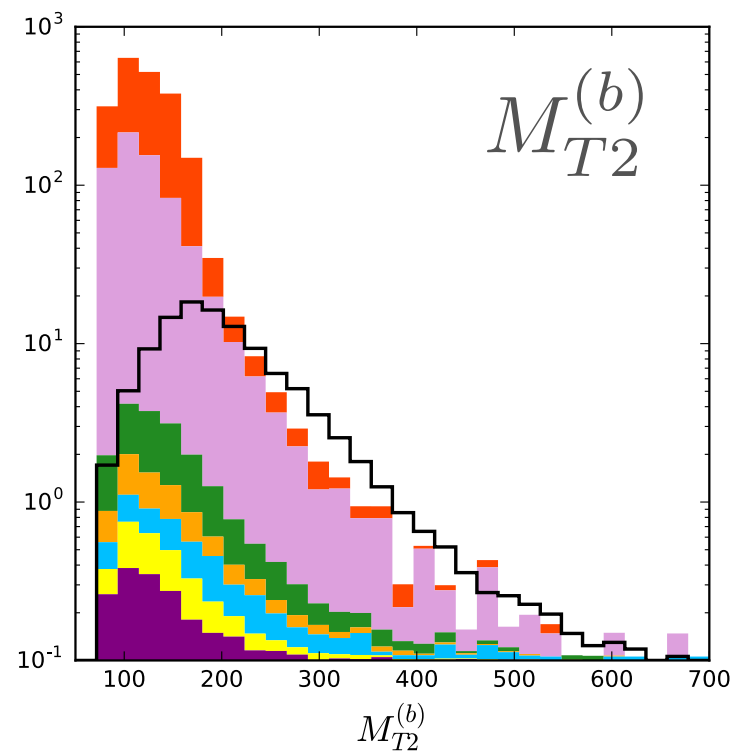
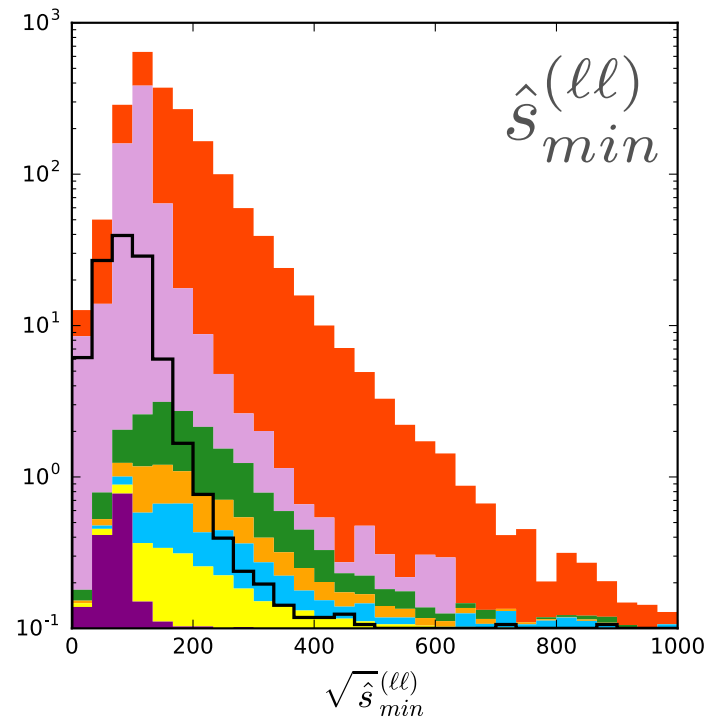
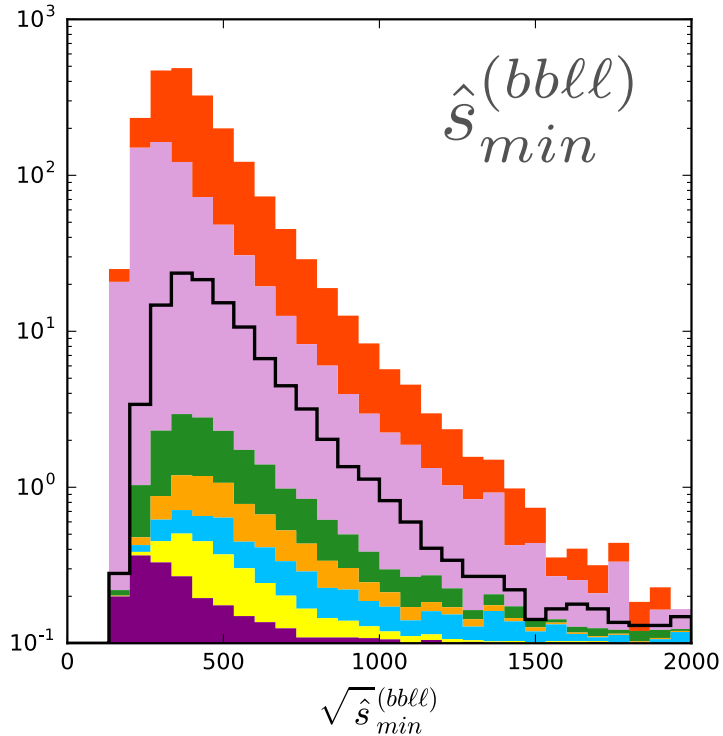
- Jets are clustered with the anti- k_T algorithm [65] with cone-size $\Delta R = 0.4$, where ΔR is the distance (2.1) in the (ϕ, η) space. Jets are also required to have $p_T > 30$ GeV and $|\eta| < 2.5$.
- For lepton isolation, we require $\frac{p_{T\ell}}{p_{T\ell} + \sum_i p_{Ti}} > 0.7$, where the sum is taken over the transverse momenta p_{Ti} of all final states particles i , $i \neq \ell$, with $|\eta_i| < 2.5$, $p_{Ti} > 0.5$ GeV and within $\Delta R_{i\ell} < 0.3$ of the lepton candidate ℓ .
- For photon isolation, we analogously require $\frac{\sum_i p_{Ti}}{p_{T\gamma}} < 0.12$ for particles within $\Delta R_{i\gamma} < 0.3$ of the photon candidate γ .
- The missing transverse momentum \vec{P}_T is defined as the negative vector sum of the transverse momenta of the reconstructed jets, leptons and photons.
- We use the a flat b -tagging efficiency, $\epsilon_{b \rightarrow b} = 0.75$, and flat mis-tagging rates for non- b jets of $\epsilon_{c \rightarrow b} = 0.1$ and $\epsilon_{j \rightarrow b} = 0.01$ [66].

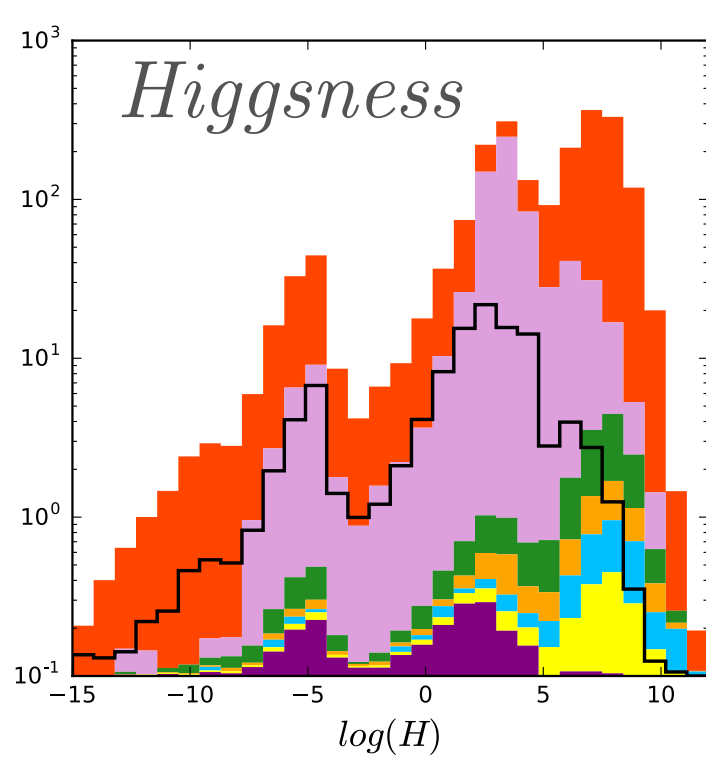
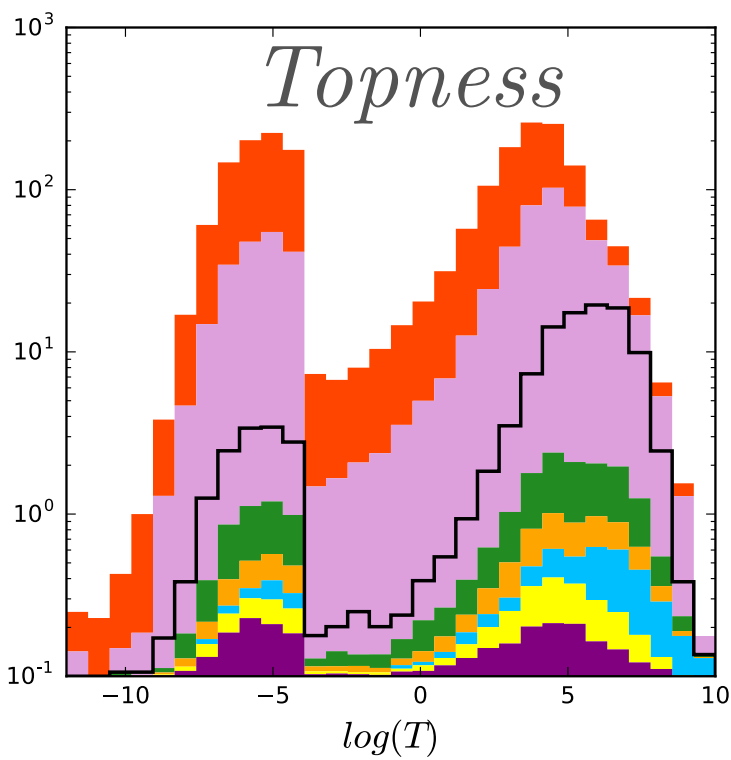
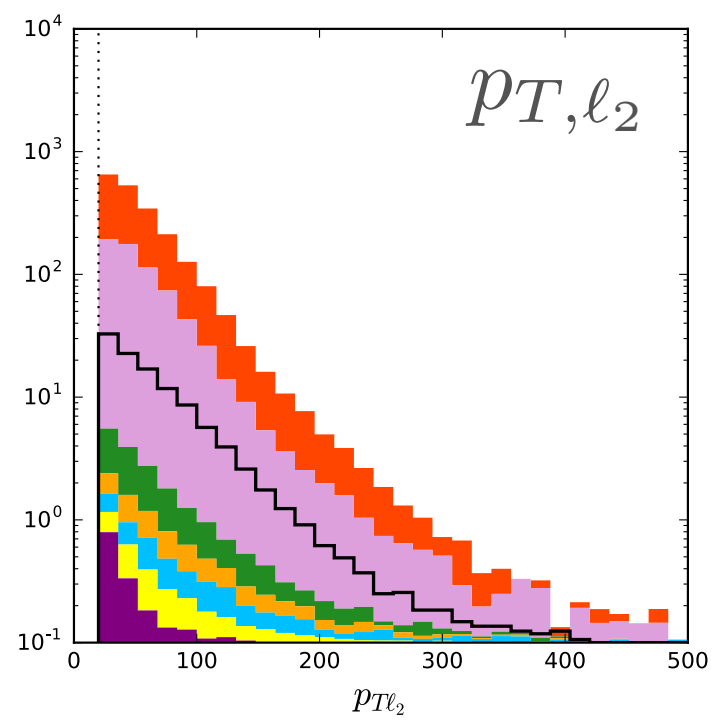
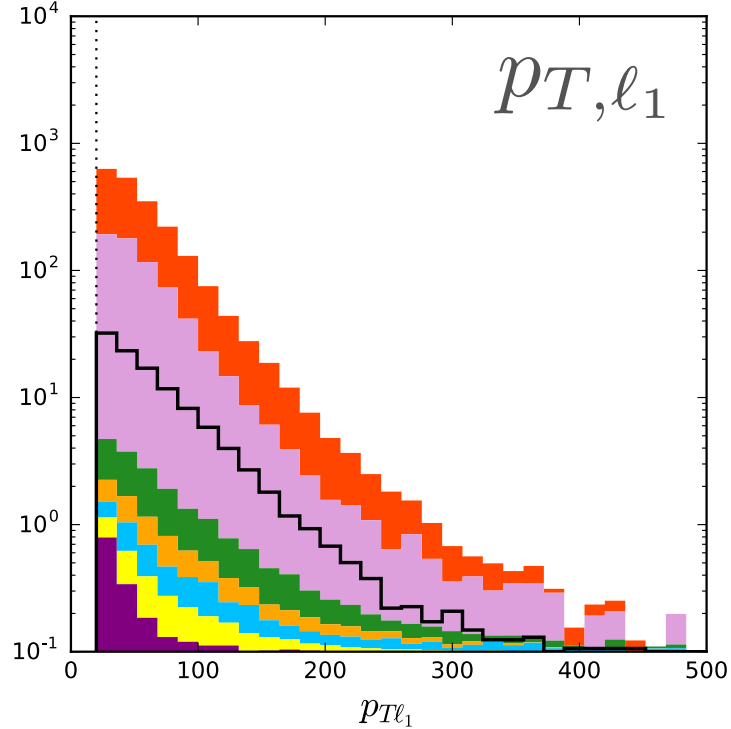
Dotted lines are baseline cuts.

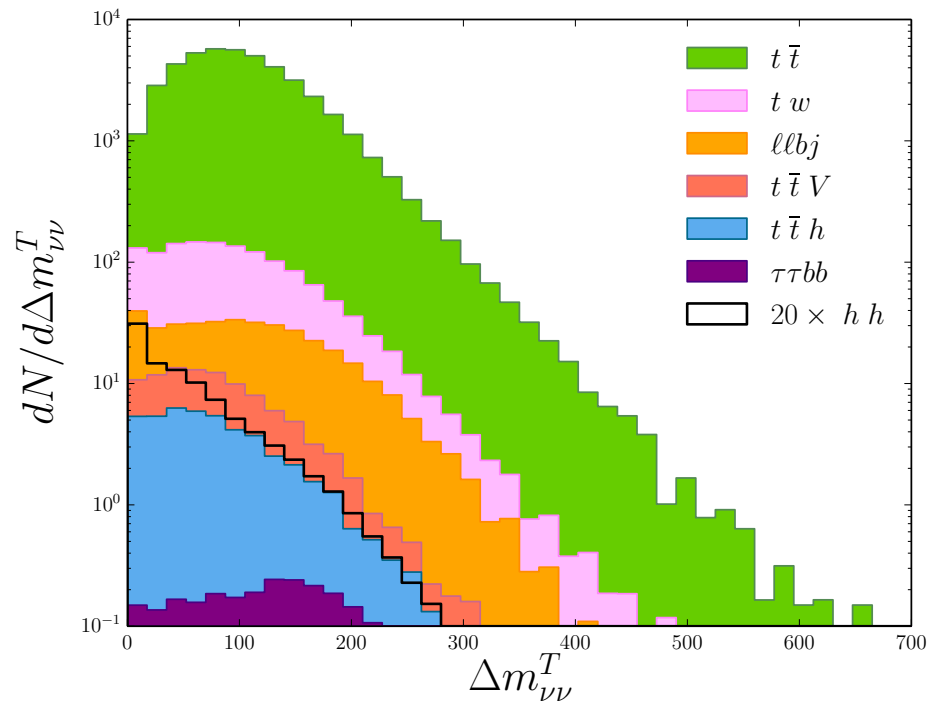
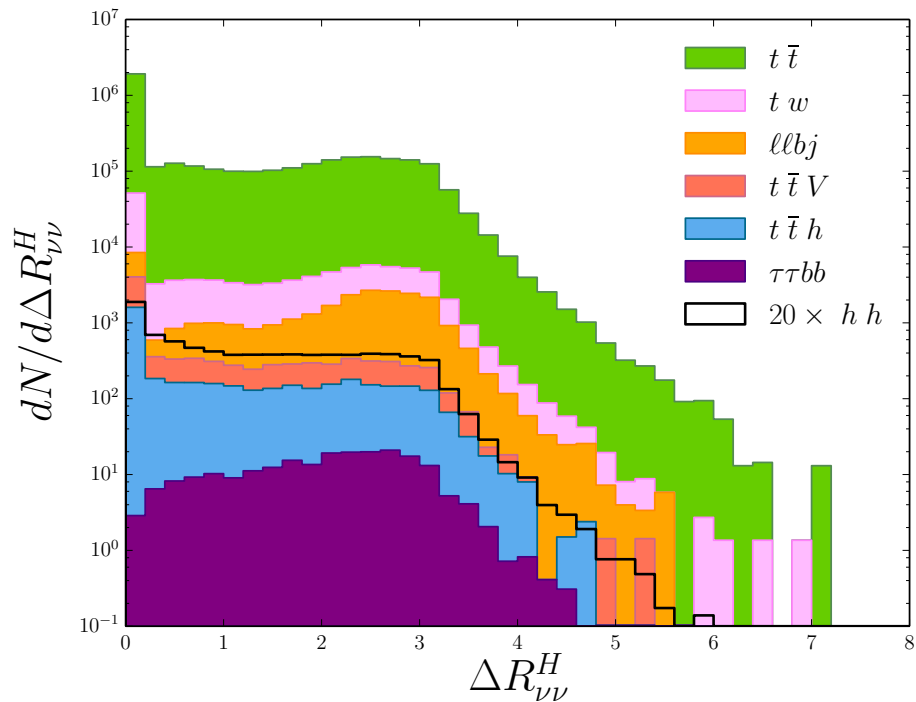
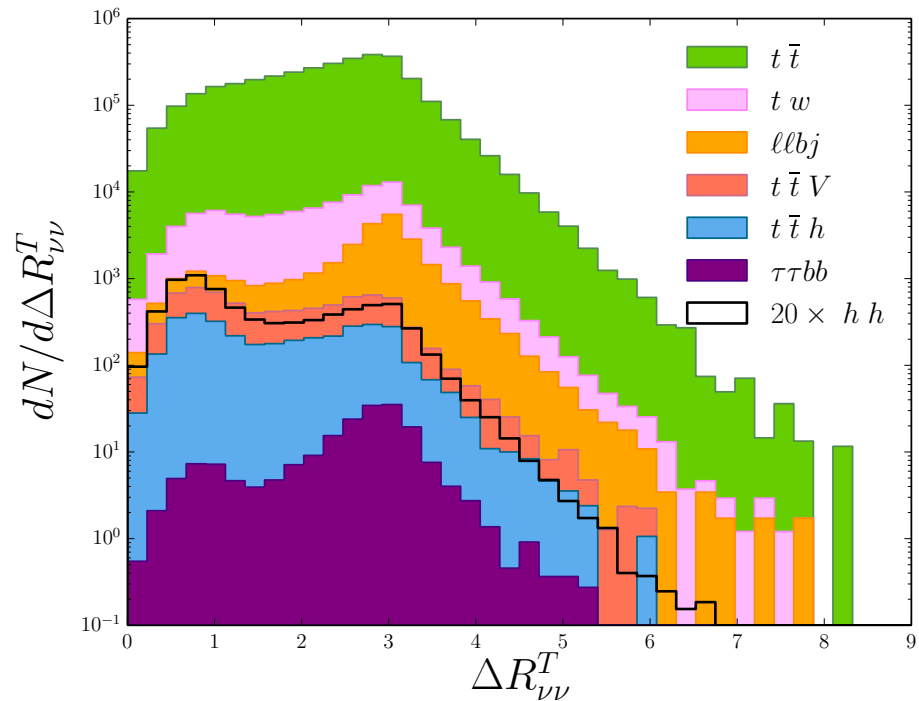
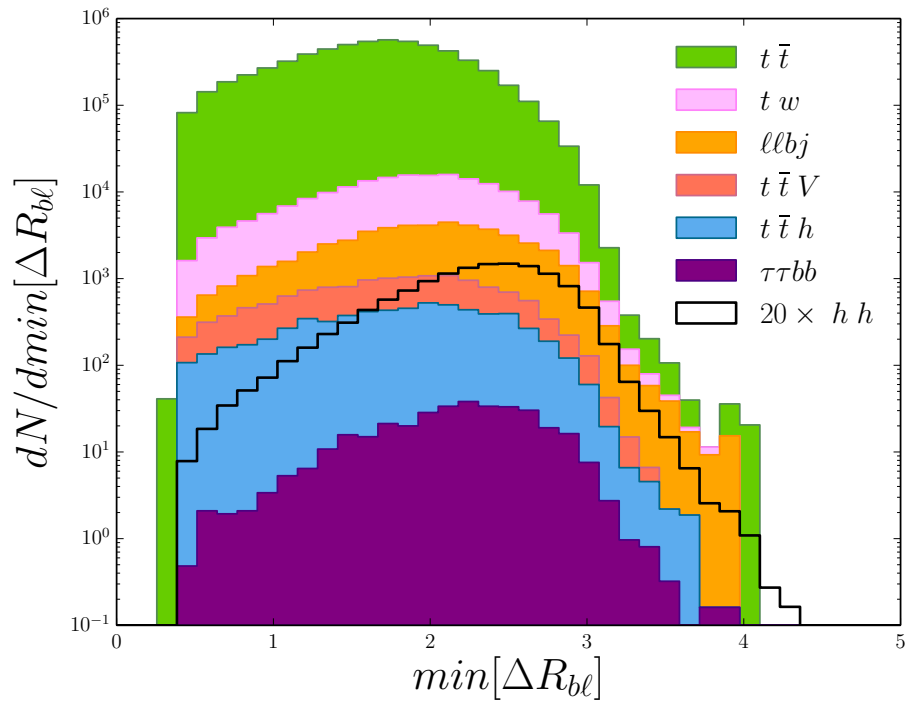


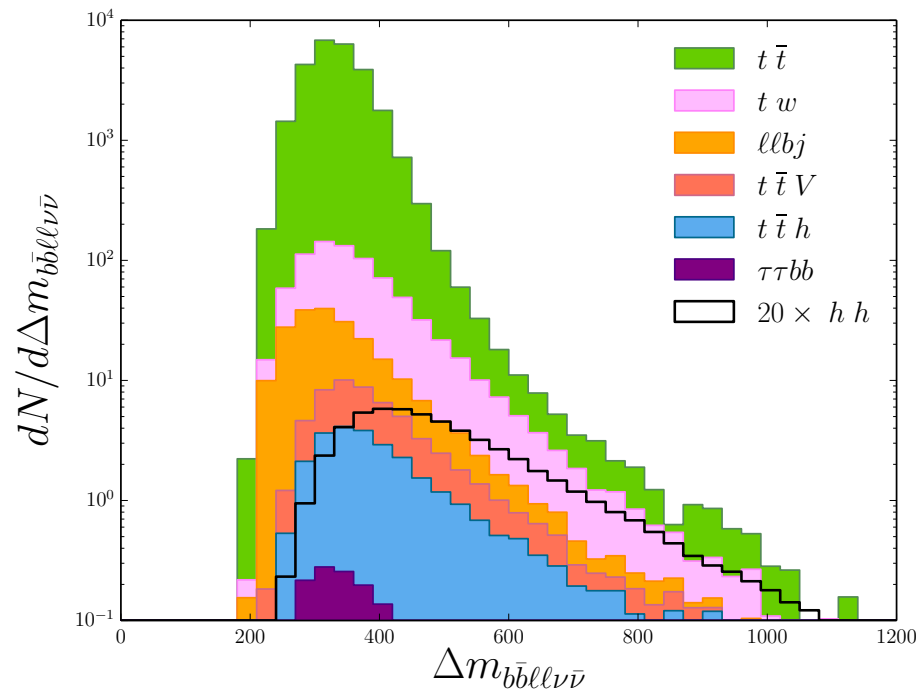
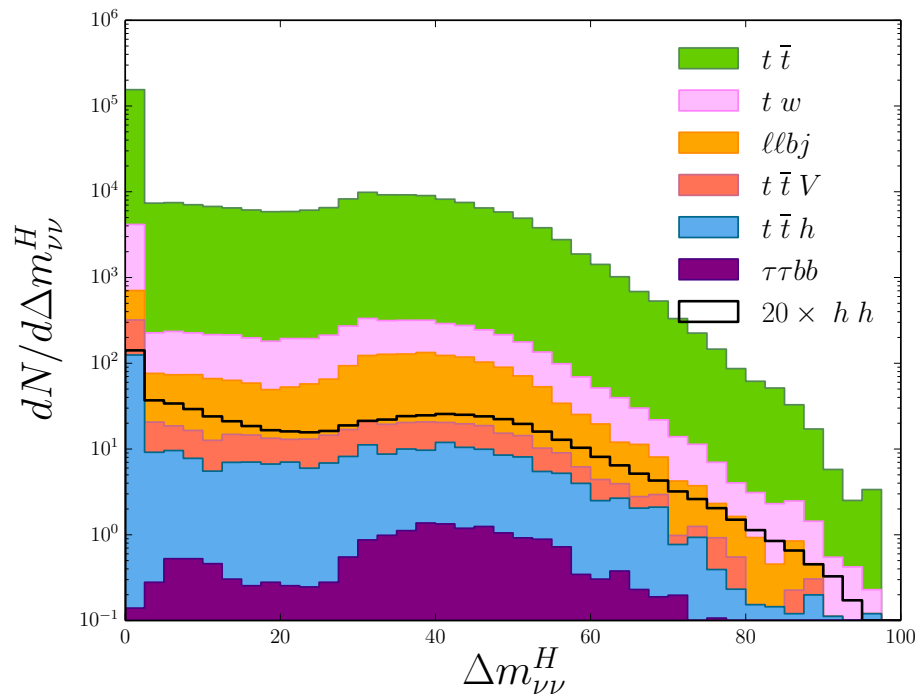


$$\hat{s}_{min}^{(v)} = m_{\tilde{\nu}}^2 + 2 \left(\sqrt{|\vec{P}_T^{\tilde{\nu}}|^2 + m_{\tilde{\nu}}^2} |\vec{P}_T| - \vec{P}_T^{\tilde{\nu}} \cdot \vec{P}_T \right)$$

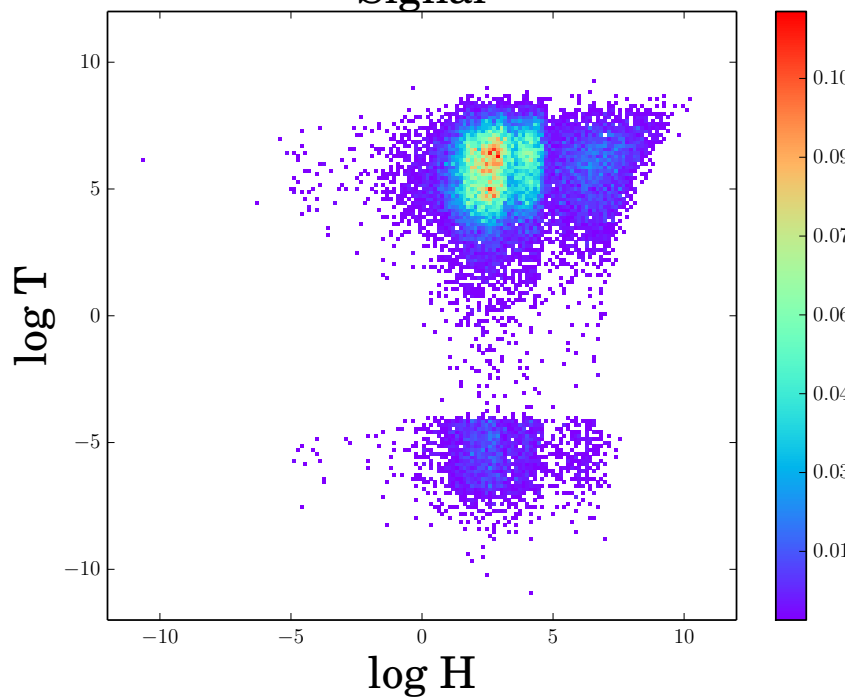




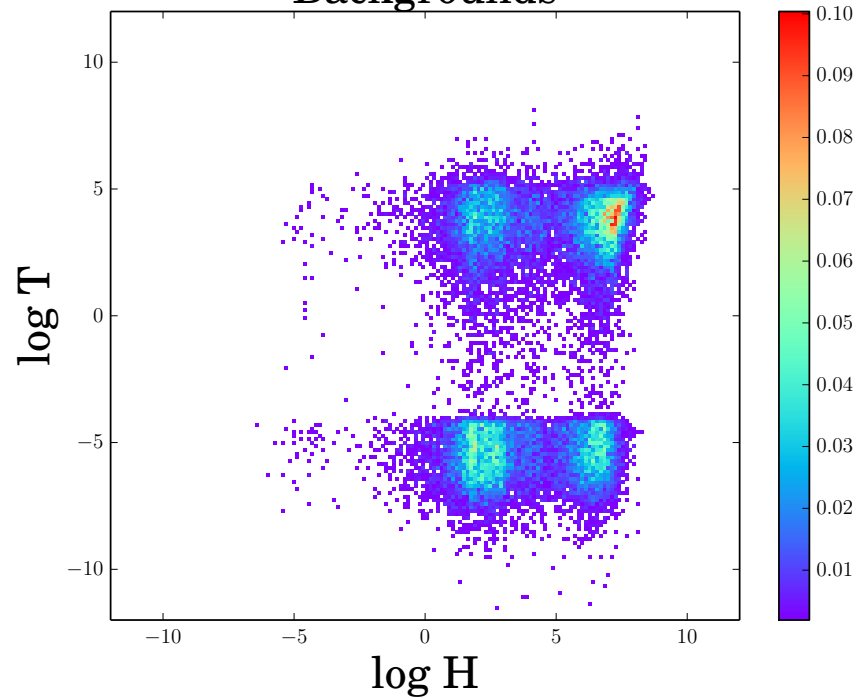




Signal



Backgrounds



Cuts at event generation

$$p_{Tj} > 20 \text{ GeV}, p_{Tb} > 20 \text{ GeV},$$

$$p_{T\gamma} > 10 \text{ GeV}, p_{T\ell} > 10 \text{ GeV}, \eta_j < 5, \eta_b < 5, \eta_\gamma < 2.5,$$

$$\eta_\ell < 2.5, \Delta R_{bb} < 1.8, \Delta R_{\ell\ell} < 1.3,$$

$$70 \text{ GeV} < m_{jj}, m_{bb} < 160 \text{ GeV} \text{ and } m_{\ell\ell} < 75 \text{ GeV}$$

$$5 \text{ GeV} < m_{\ell\ell} < 75 \text{ GeV} \quad \text{For } jj\ell\ell\nu\bar{\nu}, \ell\ell bj \text{ and } tW + j \text{ backgrounds,}$$

Cross sections @ 14 TeV LHC

$$\sigma_{hh} = 40.7 \text{ fb (NNLO)}$$

Kim, Kong, Matchev, Park, PRL 2019

Kim, Kim, Kong, Matchev, Park, JHEP 2019

$$\sigma_{hh} \cdot 2 \cdot \text{BR}(h \rightarrow b\bar{b}) \cdot \text{BR}(h \rightarrow WW^* \rightarrow \ell^+ \ell^- \nu \bar{\nu}) = 0.648 \text{ fb}$$

ℓ denotes an electron or a muon, including leptons from tau decays.

- tt: 953.6 pb (NNLO)
- tth: 611.3 fb (NLO)
- ttV (V=W, Z): 1.71 pb (NLO)
- DY: $k_{QCD \otimes QED}^{NNLO, DY} \approx 1$
- Irreducible jjllnunu: $k_{NLO} = 2$
- tWj: 0.51 pb (after cuts, including all relevant branching fractions)

$$\sigma_{bknud} \sim 10^5 \sigma_{hh}$$

	Signal	$t\bar{t}$	$t\bar{t}h$	$t\bar{t}V$	$llbj$	$\tau\tau bb$	$tw + j$	$jjll\nu\nu$	σ	S/B
Baseline cuts: $p_T > 20 \text{ GeV}$, $p_{T,\ell} > 20 \text{ GeV}$, $\Delta R_{\ell\ell} < 1.0$, $p_{T,b} > 30 \text{ GeV}$, $\Delta R_{bb} < 1.3$, $m_{\ell\ell} < 65 \text{ GeV}$, $95 < m_{bb} < 140 \text{ GeV}$	0.01046	1.8855	0.0269	0.0179	0.0697	0.0250	0.2209	0.0113	0.38	0.0046

cross section in fb

tt: 84% tW: 9.8% DY+jets: 3.1% tth: 1.2% tautau + bb: 1.1% ttV: 0.8%