Di-Higgs Production Associated with Dark Matter at the LHC: A Machine-Learning Analysis

Ernesto Arganda

Departamento de Física Teórica &
Instituto de Física Teórica UAM-CSIC
Universidad Autónoma de Madrid

ernesto.arganda@uam.es | @ernesto_arganda

May, 30th 2024

References

- EA, M. Epele, N. I. Mileo, R. A. Morales, [arXiv:2401.03178 [hep-ph]], under review in *EPJ Plus*.

Outline

Introduction

- Machine learning (ML) basic tool for exp and pheno HEP studies:
  - ML crucial to take full advantage of LHC data to probe SM and BSM.
  - Could ML replace traditional cut-and-count methods?
- Extended Higgs sector searches intensive experimental program by ATLAS and CMS.
  - Additional Higgs bosons portals to dark sectors.
  - Multi-Higgs final states + E_T^{miss} coming from dark matter (DM).
  - Many ATLAS and CMS searches for di-Higgs production + E_T^{miss}.
- Case study: di-Higgs + E_T^{miss} LHC signatures in general frameworks of extended Higgs sectors with `XGBoost` and deep neural networks (DNN).

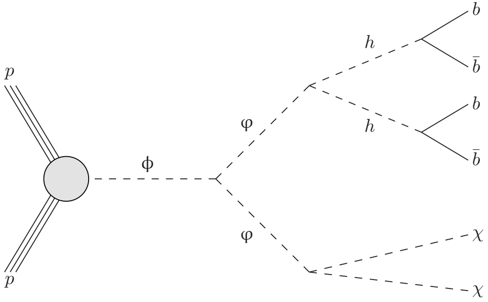
Main goal: improvement capability of modern ML tools over cut-based analyses applied to LHC Higgs-pair production associated with DM.

Phenomenological Framework

Collider analysis within general class of simplified models [Blanke *et al.*, 1901.07558] with an extended scalar sector of 3 real scalar particles:

- The heaviest scalar ϕ produced via gluon fusion at LHC and predominantly decays to a pair of
- intermediate scalars φ that interacts with visible sector only through its coupling with SM Higgs h.
- The lightest scalar χ is the DM candidate.

Resonant topology: $pp \to \phi \to \varphi\varphi \to (hh)(\chi\chi)$ [Arganda *et al.*, 1710.07254]



Signal and Background Simulation

- MC events generated with MadGraph + Pythia + Delphes.
- $\sqrt{s} = 14$ TeV and $\mathcal{L} = 1$ ab^{-1}.
- ATLAS card provided by Delphes 3.3.3:
  - b-tagging efficiency: $\epsilon_b = \frac{24 \tanh(0.003\, p_T)}{1+0.086\, p_T}$ (maximum b-tagging efficiency of $\sim 73\%$ for $p_T \sim 120$ GeV).
  - Mistag rates: $\epsilon_j = 0.002 + 7.3 \times 10^{-6}\, p_T$, $\epsilon_c = \frac{0.2 \tanh(0.02\, p_T)}{1+0.0034\, p_T}$.
- 4-flavor scheme and no jet matching.

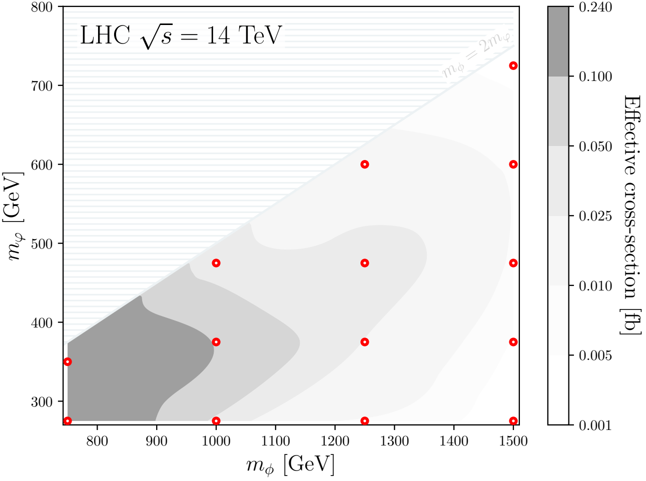
Signal Region (SR)

$$N_b = 4,\ N_{lep} = 0,\ p_T^{j,b} > 20 \text{ GeV},\ E_T^{\text{miss}} > 200 \text{ GeV}.$$

Signal Effective Cross-Sections (XS)

$$\sigma_{\text{eff}} = \sigma(pp \to b\bar{b}b\bar{b}\chi\chi) \cdot \varepsilon_{\text{SR}}$$

where ε_{SR} is fraction of simulated signal events which satisfies SR.



Background Simulation

- Irreducible
  - $Z + b\bar{b}b\bar{b}$ with $Z \to \nu\bar{\nu}$: $Z|_{\rm inv}\, b\bar{b}b\bar{b}$.
  - $t\bar{t} + b\bar{b}$ with $t(\to bjj)\bar{t}(\to \bar{b}\ell^-\bar{\nu})$ and $t(\to b\ell^+\nu)\bar{t}(\to \bar{b}jj)$: $t\bar{t}|_{\rm semilep}\, b\bar{b}$ and $t\bar{t}|_{\rm semitau}\, b\bar{b}$.
- Reducible
  - $t\bar{t}$+jets.
  - V+jets ($Z + b\bar{b} + jj$, $Z + jjjj$, $W^\pm + b\bar{b} + jj$, and $W^\pm + jjjj$).
  - QCD multijet: negligible since no genuine source of $E_T^{\rm miss}$.

Background Effective XS in SR

Process	$\sigma_{\rm eff}$ [fb]
$Z\|_{\rm inv}\, b\bar{b}jj$	0.937
$t\bar{t}\|_{\rm semilep}\, b\bar{b}$	0.680
$t\bar{t}\|_{\rm semitau}\, b\bar{b}$	0.580
$W^\pm\|_{\rm semilep}\, b\bar{b}jj$	0.242
$Z\|_{\rm inv}\, b\bar{b}b\bar{b}$	0.112

Low-Level and High-Level Features for ML Algorithms

Low-level feature	Description
N_j	Number of light-jets
p_T^i	Transverse momentum of the four leading b-jets ($i = bst, bnd, brd, bth$)
η^i	Pseudorapidity of the four leading b-jets ($i = bst, bnd, brd, bth$)
ϕ^i	Azimuthal angle of the four leading b-jets ($i = bst, bnd, brd, bth$)
E_T^{miss}	Missing transverse momentum
ϕ^{miss}	Azimuthal angle of the missing transverse momentum

High-level feature	Description
χ_{hh}	$\sqrt{\left(\frac{m_{2b}^{\text{lead}}-m_h}{0.1m_{2b}^{\text{lead}}}\right)^2+\left(\frac{m_{2b}^{\text{subl}}-m_h}{0.1m_{2b}^{\text{subl}}}\right)^2}$
p_T^i	Transverse momentum of the two reco Higgs bosons ($i = Hst, Hnd$)
η^i	Pseudorapidity of the two reco Higgs bosons ($i = Hst, Hnd$)
ϕ^i	Azimuthal angle of the two reco Higgs bosons ($i = Hst, Hnd$)
m_{hh}	Invariant mass of the reco Higgs boson pair
$\Delta\eta_{hh}$	Difference $\eta^{Hst} - \eta^{Hnd}$ of the two reco Higgs bosons
$\Delta\phi_{hh}$	Difference $\phi^{Hst} - \phi^{Hnd}$ of the two reco Higgs bosons
ΔR_{hh}	Distance $\sqrt{\Delta\eta_{hh}^2 + \Delta\phi_{hh}^2}$ of the two reco Higgs bosons
$\Delta\phi^i_{MET}$	Differences $\phi^{\text{miss}} - \phi^i$ for $i = bst, bnd, brd, bth, Hst, Hnd$
E_T^{miss} significance	Computed as $E_T^{\text{miss}}/\sqrt{p_T^{bst}+p_T^{bnd}+p_T^{brd}+p_T^{bth}}$

Cut-Based Analysis

Signal significance with $\Delta = 15\%$ systematic uncertainty:

$$\mathcal{S}_{\text{sys}} = \sqrt{2\left((B+S)\log\left(\frac{(\Delta^2B^2+B)(B+S)}{\Delta^2B^2(B+S)+B^2}\right) - \frac{1}{\Delta^2}\log\left(\frac{\Delta^2B^2S}{B(\Delta^2B^2+B)}+1\right)\right)}$$

Only 2 Benchmark Points (BP) exceed evidence (3σ) significance level:

BP 750_350_25

- Cuts: SR + $N_j \leq 4$, $p_T^{bst} < 160$ GeV, $p_T^{bnd} < 110$ GeV, $p_T^{brd} < 80$ GeV, $\chi_{hh} < 3.5$, $m_{hh} < 380$ GeV, $|\Delta\phi_{MET}^{bst}| < 1.8$.
- $\mathcal{S}_{\text{sys}} = 3.77\sigma$.

BP 1000_275_25

- Cuts: SR + $N_j \leq 2$, $\chi_{hh} < 3$, $m_{hh} < 300$ GeV, $\Delta R_{hh} < 1$.
- $\mathcal{S}_{\text{sys}} = 3.48\sigma$.

Machine-Learning Algorithms for Collider Analyses

To tackle binary classification problem we consider two ML algorithms:

- `XGBoost`, an ensemble method based on gradient-boosted trees.
- Deep Neural Networks (DNN), neural networks that use multiple layers to progressively extract higher-level features from the raw input.

XGBoost: Overview and Architecture

- binary:logistic as objective function.
- Receiver Operating Characteristic Area under the Curve (AUC) as the evaluation metric.
- Hyper-parameters optimization carried out via GridSearchCV.
- Number of boost rounds controlled by using early stopping.

Parameter	Description
learning_rate	Step size shrinkage of weights used at each boosting step
max_depth	Maximum depth of a tree
min_child_weight	Minimum sum of instance weight required in a child
subsample	Fraction of training instances to be random samples for each tree
colsample_bytree	Subsample ratio of columns when constructing each tree
gamma	Specifies the minimum loss reduction required to split a node
reg_alpha	L1 regularization term on weights
reg_lambda	L2 regularization term on weights

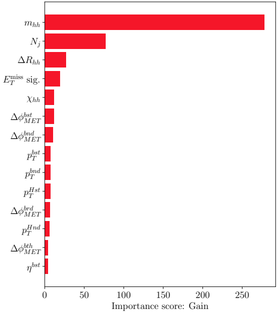
XGBoost: Feature Importance

Simulated sample:

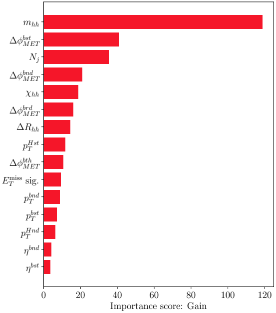
- Split in training, validation, and test: $\sim$ 65%, 15%, and 20% of total number of events.
- 1:1 composition of signal and background events.

Best classification power obtained with set of 14 features:

$$p_T^{bst}, \eta^{bst}, p_T^{bnd}, \eta^{bnd}, p_T^{Hst}, p_T^{Hnd}, E_T^{\text{miss}} \text{ sig}, \chi_{hh}, m_{hh}, \Delta R_{hh}, \Delta\phi_{MET}^{bst},$$
$$\Delta\phi_{MET}^{bnd}, \Delta\phi_{MET}^{brd}, \text{and } \Delta\phi_{MET}^{bth}.$$



BP 750_275_25



BP 750_350_25

DNN: Overview and Architecture

- Sigmoid as activation function.
- 500, 500, 250, 100, and 50 neurons in 5 consecutive layers.
- Dropout-rate of 21% to reduce overfitting.
- Binary cross-entropy as training loss function, using Adam algorithm with initial learning rate η of 0.001.
- AUC as evaluation metric.
- Early stopping approach to avoid overtraining.

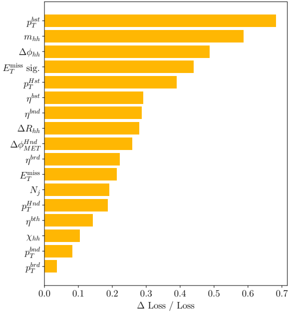
DNN: Feature Importance

Simulated sample:

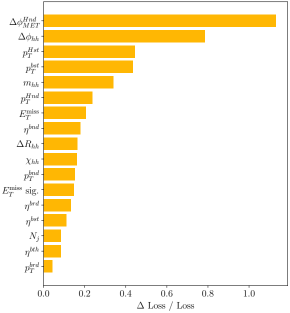
- Split in training, validation, and test: $\sim$ 65%, 15%, and 20% of total number of events.
- 1:1 composition of signal and background events.

Best classification power obtained with set of 17 features:

N_j , p_T^{bst} , η^{bst} , p_T^{bnd} , η^{bnd} , p_T^{brd} , η^{brd} , η^{bth} , p_T^{Hst} , p_T^{Hnd} , E_T^{miss} , E_T^{miss} sign. , χ_{hh} , m_{hh} , ΔR_{hh} , $\Delta\phi_{hh}$, and $\Delta\phi_{MET}^{Hnd}$.

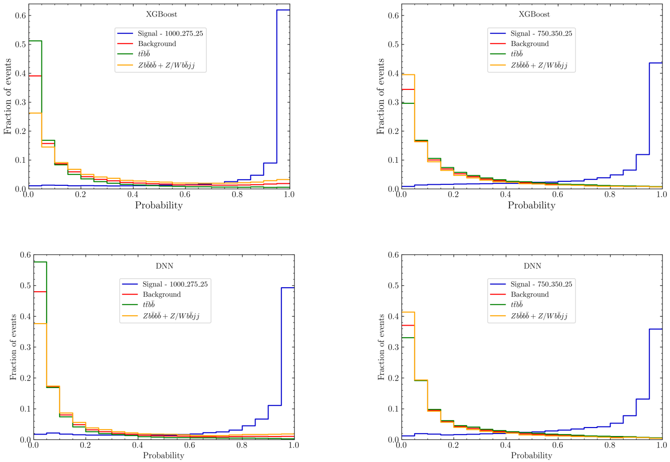


BP 750_275_25



BP 750_350_25

Results: XGBoost and DNN Probability Distributions



- Both ML algorithms lead to efficient classification: all AUC > 0.9.
- In general, both classifiers better at identifying $t\bar{t}b\bar{b}$ than other backgrounds.
- `XGBoost` seems to label signal better, DNN classifies background more efficiently.
- Signal acceptance and background rejection depend on probability threshold.

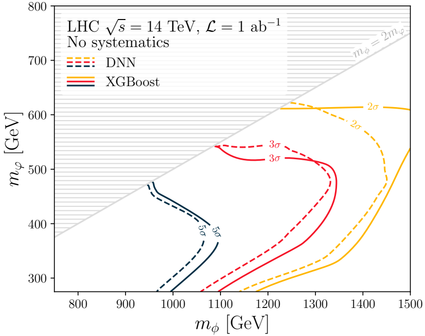
Choice of Probability Threshold

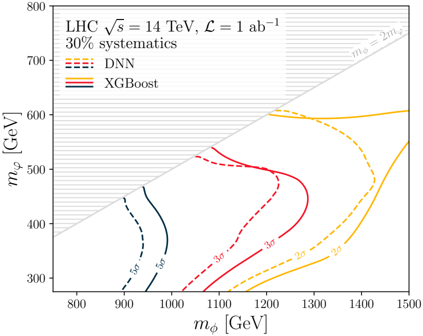
For each BP, probability threshold is chosen to maximize the statistical significance:

$$\mathcal{S} = \sqrt{-2\left((S+B)\ln\left(\frac{B}{S+B}\right)+S\right)},$$

- Systematic uncertainties neglected.
- S and B are signal and background rates after applying SR cuts and additional cut on classifier output.

Significance level curves in $[m_\phi, m_\varphi]$ plane

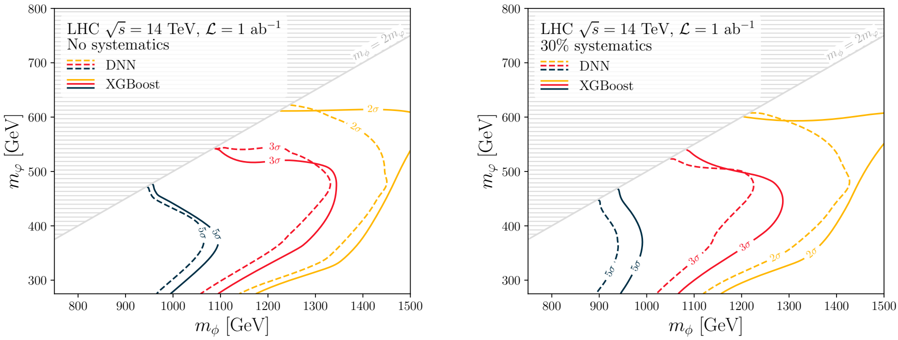




No systematics:

- For $m_\phi \lesssim 950$ GeV, 5σ reached for all m_φ.
- For $m_\varphi \sim 380$ GeV, discovery region extends in m_ϕ-range up to ~ 1090 GeV with `XGBoost` and ~ 1060 GeV with DNNs.
- 3σ obtained for $m_\phi \sim 1340$ GeV providing $m_\phi \sim 480$ GeV.
- Exclusion-limit (2σ) sensitivity in most of parameter space.
- In general, `XGBoost` results tend to slightly improve DNN ones.

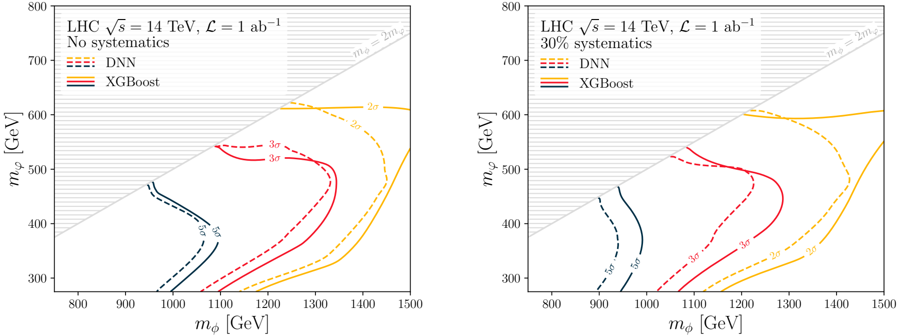
Significance level curves in $[m_\phi, m_\varphi]$ plane



30% systematics:

- Thresholds optimizing $\mathcal{S}_{\text{sys}}$ shift to increase B rejection at the cost of decreasing S acceptance.
- 5σ obtained for all range of m_φ for $m_\phi < 900$ GeV. Discovery region reaches $m_\phi \sim 990$ GeV for $m_\varphi \sim 370$ GeV.
- 3σ reached now for m_ϕ up to ~ 1290 GeV when $m_\varphi \sim 450$ GeV.
- Again, `XGBoost` classifiers appear to provide better prospects.

Significance level curves in $[m_\phi, m_\varphi]$ plane



In summary:

- Significant improvement with respect to cut-based analysis.
- 15% systematics: $\mathcal{S} > 3\sigma$ for 7 (XGBoost) and 6 (DNN) of 14 BP, while for cut-based analysis only 2 BP.
- 30% systematics: cut-based analysis $\mathcal{S}$ drop below 3σ. Not the case when the ML classifiers are used.

Conclusions

- Performance study of `XGBoost` and DNN classifiers on LHC signature consisting of 4 b-jets+$E_T^{\rm miss}$ (scalars+DM simplified model), comparing against cut-based analyses.
- Dominant irreducible backgrounds are $Z + b\bar{b}b\bar{b}$ and $t\bar{t} + b\bar{b}$; main reducible backgrounds are V+jets ($Z/W + b\bar{b} + jj$); QCD multijet safely ignored since does not provide true source of $E_T^{\rm miss}$.
- Scan of $m_\phi \in [750, 1500]$ GeV and $m_\varphi \in [275, m_\phi/2]$ GeV. 15 low-level and 18 high-level kinematic features to feed ML algorithms.
- ML performance based on maximum significance reached for 14 TeV and 1 ab^{-1}. Both algorithms present very similar performances and a significant improvement with respect to cut-based analysis.

Main take-home message: our pheno analysis shows that proposed LHC signature deserves dedicated searches by exp collaborations, for which modern ML algorithms would play a crucial role.

BACKUP

Simplified model

Inspired by [Blanke *et al.*, 1901.07558], the interaction Lagrangian is given by

$$\mathcal{L} = \frac{C_{\phi gg}}{\Lambda}\phi G_{\mu\nu}G^{\mu\nu} + \frac{m_{\phi\varphi\varphi}}{2}\phi\varphi\varphi + \frac{m_{\varphi hh}}{2}\varphi hh + \frac{m_{\varphi\chi\chi}}{2}\varphi\chi\chi\,, \tag{1}$$

and the resulting cross-section for the resonant topology is factorized as

$$\begin{aligned}\sigma(pp \to b\bar{b}b\bar{b}\chi\chi) \quad &= \quad \sigma(pp \to \phi)\cdot \mathrm{BR}(\phi \to \varphi\varphi)\cdot 2 \cdot \mathrm{BR}(\varphi \to hh) \\ &\quad \cdot\mathrm{BR}(\varphi \to \chi\chi)\cdot \left[\mathrm{BR}(h \to b\bar{b})\right]^2 \,. \end{aligned} \tag{2}$$

Effective coupling $C_{\phi gg}$ collects the effect of heavy quarks at scale $\Lambda = 1$ TeV and production cross-section $\sigma(pp \to \phi)$ is:

$$\sigma(pp \to \phi) = \left(\frac{v}{1\,\mathrm{TeV}}\right)^2 \sigma(pp \to S)\,, \tag{3}$$

while the two decay channels for the heavy scalar are

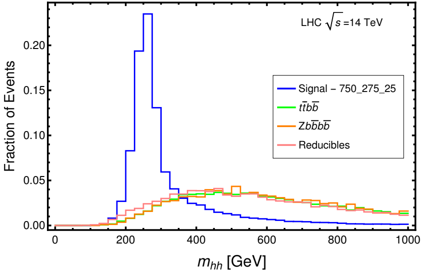
$$\Gamma(\phi \to \varphi\varphi) = \frac{m^2_{\phi\varphi\varphi}}{32\pi m_\phi}\sqrt{1 - \frac{4m^2_\varphi}{m^2_\phi}} \quad \text{and} \quad \Gamma(\phi \to gg) = \frac{C^2_{\phi gg}}{1\,\mathrm{TeV}^2}\frac{2m^3_\phi}{\pi}\,. \tag{4}$$

With $m_{\phi\varphi\varphi}$ equal to EW scale and far to $\varphi\varphi$ threshold, $\mathrm{BR}(\phi \to \varphi\varphi) \sim 1$. Couplings $m_{\varphi hh}$ and $m_{\varphi\chi\chi}$ are such that branching ratios of φ decaying to hh and $\chi\chi$ maximize their product: $\mathrm{BR}(\varphi \to hh) = \mathrm{BR}(\varphi \to \chi\chi) = 0.5$. $\mathrm{BR}(h \to b\bar{b}) = 0.58$ is also considered.

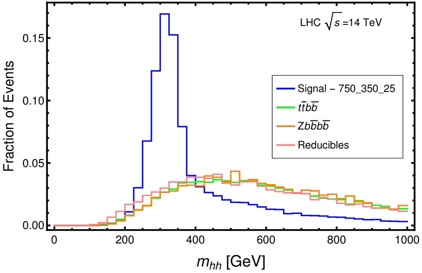
Specific Effective XS for Benchmarks

m_Φ	m_φ	$\sigma_{\rm eff}$ [fb]
750	275	0.21
750	350	0.19
1000	275	0.044
1000	375	0.098
1000	475	0.036
1250	275	0.0094
1250	375	0.024
1250	475	0.030
1250	600	0.013
1500	275	0.0013
1500	375	0.0046
1500	475	0.0085
1500	600	0.0087
1500	725	0.0037

Relevant Kinematic Distributions

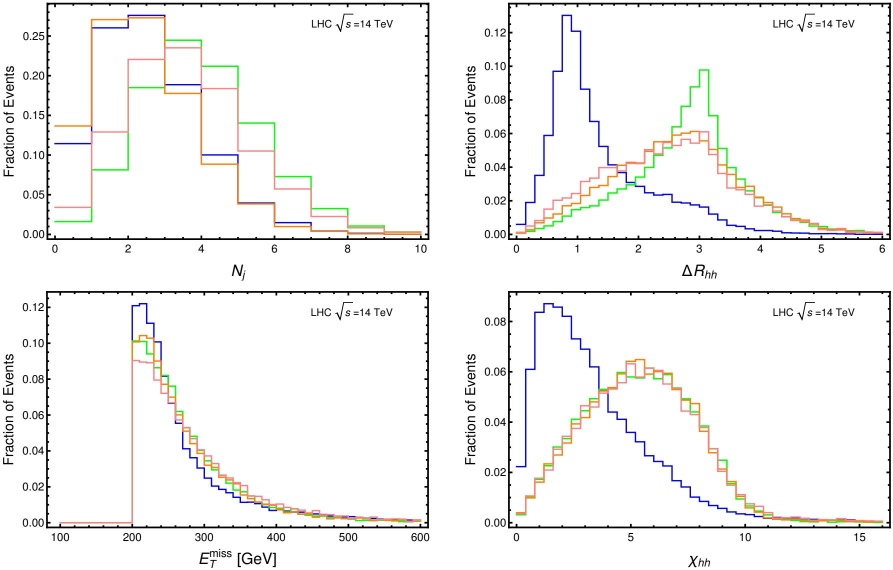


BP 750_275_25

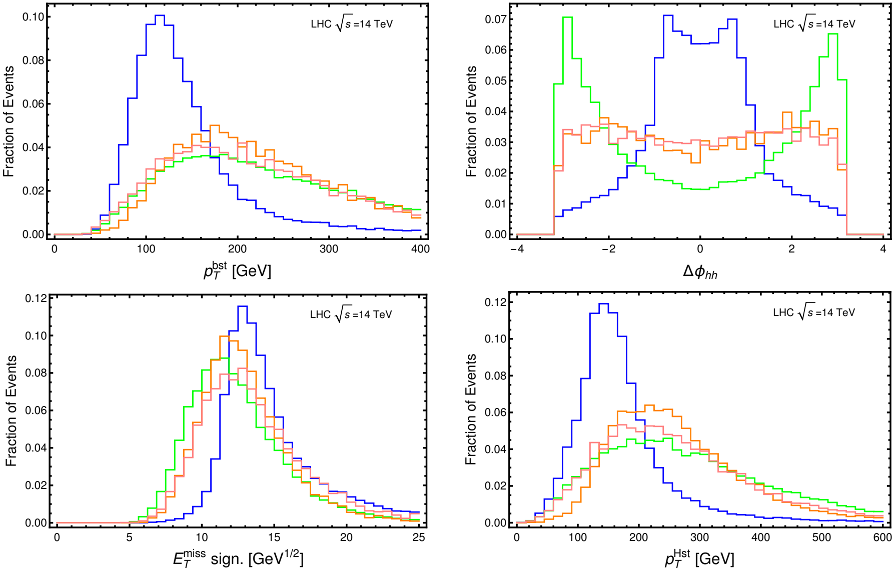


BP 750_350_25

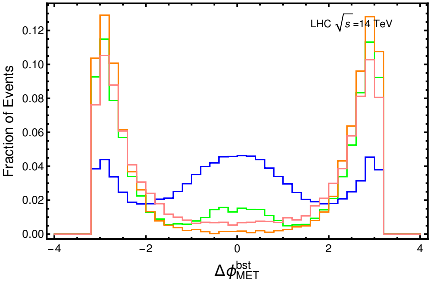
Relevant Kinematic Distributions: BP 750_275_25

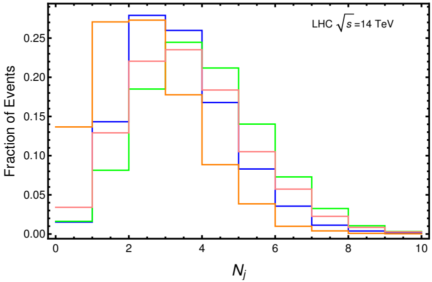


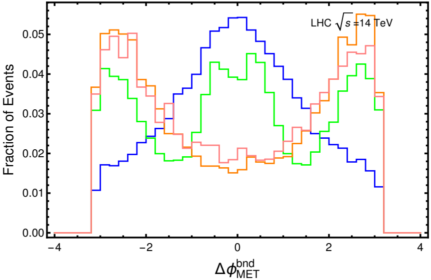
Relevant Kinematic Distributions: BP 750_275_25

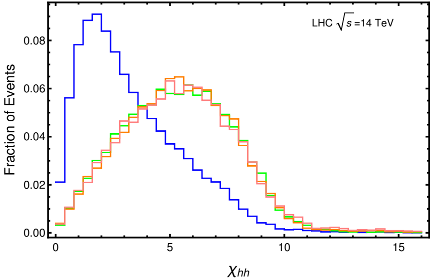


Relevant Kinematic Distributions: BP 750_350_25

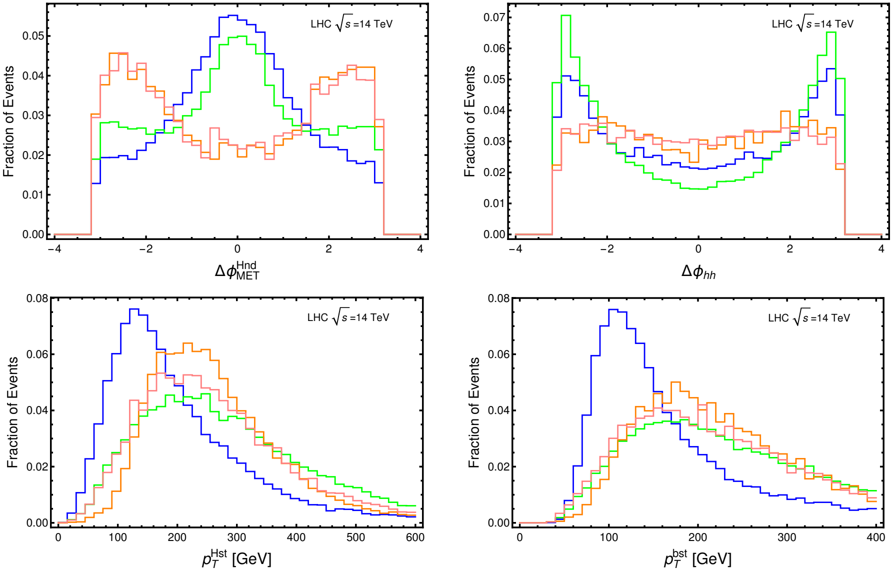








Relevant Kinematic Distributions: BP 750_350_25



Tables of Acceptance

m_ϕ	m_φ	Signal	$Z\|_{\rm inv}\, b\bar{b}jj$	$W^\pm\|_{\rm semilep}\, b\bar{b}jj$	$t\bar{t}\|_{\rm semilep}\, b\bar{b}$	$t\bar{t}\|_{\rm semitau}\, b\bar{b}$	$Z\|_{\rm inv}\, b\bar{b}b\bar{b}$
750	275	0.4047	0.0139	0.0135	0.0047	0.0053	0.0162
750	350	0.4172	0.0066	0.0075	0.0088	0.0078	0.0010
1000	275	0.3136	0.0054	0.0054	0.0005	0.0008	0.0039
1000	375	0.3551	0.0126	0.0129	0.0045	0.0054	0.0159
1000	475	0.2485	0.0007	0.0010	0.0014	0.0014	0.0001
1250	275	0.2141	0.0015	0.0012	0.0001	0.0002	0.0008
1250	375	0.3032	0.0033	0.0030	0.0004	0.0006	0.0021
1250	475	0.2384	0.0019	0.0020	0.0004	0.0003	0.0018
1250	600	0.2485	0.0006	0.0009	0.0005	0.0004	0.0001
1500	275	0.2537	0.0019	0.0021	0.0001	0.0002	0.0011
1500	375	0.2786	0.0013	0.0013	0.0001	0.0002	0.0007
1500	475	0.2458	0.0009	0.0009	0.0001	0.0001	0.0004
1500	600	0.2450	0.0009	0.0010	0.0001	0.0001	0.0008
1500	725	0.2560	0.0003	0.0005	0.0004	0.0002	0.0001

Table: Acceptances for DNN classifiers after applying a probability threshold that maximizes the significance without including systematic uncertainties.

Tables of Acceptance

m_ϕ	m_φ	Signal	$Z\|_{\rm inv}\, b\bar{b}jj$	$W^\pm\|_{\rm semilep}\, b\bar{b}jj$	$t\bar{t}\|_{\rm semilep}\, b\bar{b}$	$t\bar{t}\|_{\rm semitau}\, b\bar{b}$	$Z\|_{\rm inv}\, b\bar{b}b\bar{b}$
750	275	0.3805	0.0093	0.0084	0.0029	0.0034	0.0126
750	350	0.4147	0.0065	0.0071	0.0067	0.0059	0.0007
1000	275	0.3265	0.0036	0.0038	0.0004	0.0003	0.0017
1000	375	0.3134	0.0077	0.0076	0.0018	0.0032	0.0086
1000	475	0.1061	8.0902	$< 10^{-4}$	0.0001	0.0001	$< 10^{-4}$
1250	275	0.1951	0.0008	0.0010	0.0001	0.0001	0.0007
1250	375	0.2561	0.0015	0.0017	$< 10^{-4}$	0.0001	0.0012
1250	475	0.2735	0.0024	0.0029	0.0003	0.0004	0.0035
1250	600	0.2995	0.0009	0.0017	0.0009	0.0009	0.0002
1500	275	0.1914	0.0009	0.0010	$< 10^{-4}$	$< 10^{-4}$	0.0010
1500	375	0.2627	0.0009	0.0009	0.0001	$< 10^{-4}$	0.0007
1500	475	0.2324	0.0007	0.0009	$< 10^{-4}$	$< 10^{-4}$	0.0002
1500	600	0.0855	$< 10^{-4}$	0.0001	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
1500	725	0.2297	0.0002	0.0006	0.0001	0.0002	0.0002

Table: Acceptances for `XGBoost` classifiers after applying a probability threshold that maximizes the significance without including systematic uncertainties.

Tables of Acceptance

m_ϕ	m_φ	Signal	$Z\vert_{\rm inv}\, b\bar{b}jj$	$W^{\pm}\vert_{\rm semilep}\, b\bar{b}jj$	$t\bar{t}\vert_{\rm semilep}\, b\bar{b}$	$t\bar{t}\vert_{\rm semitau}\, b\bar{b}$	$Z\vert_{\rm inv}\, b\bar{b}b\bar{b}$
750	275	0.1736	0.0024	0.0021	0.0009	0.0007	0.0031
750	350	0.1528	0.0007	0.0006	0.0008	0.0008	$< 10^{-4}$
1000	275	0.1999	0.0021	0.0021	0.0001	0.0003	0.0012
1000	375	0.1411	0.0021	0.0021	0.0007	0.0008	0.0022
1000	475	0.1706	0.0003	0.0004	0.0006	0.0007	$< 10^{-4}$
1250	275	0.1632	0.0008	0.0006	$< 10^{-4}$	0.0001	0.0004
1250	375	0.1990	0.0013	0.0013	0.0002	0.0003	0.0008
1250	475	0.1678	0.0008	0.0008	0.0003	0.0001	0.0007
1250	600	0.2063	0.0004	0.0006	0.0004	0.0003	0.0001
1500	275	0.2150	0.0014	0.0014	0.0001	0.0001	0.0007
1500	375	0.2410	0.0010	0.0010	0.0001	0.0001	0.0004
1500	475	0.2243	0.0007	0.0007	0.0001	0.0001	0.0003
1500	600	0.2134	0.0006	0.0008	0.0001	0.0001	0.0005
1500	725	0.2296	0.0002	0.0004	0.0003	0.0002	$< 10^{-4}$

Table: Acceptances for DNN classifiers after applying a probability threshold that maximizes the significance including 30% of systematic uncertainties in the total background yield.

Tables of Acceptance

m_ϕ	m_φ	Signal	$Z\|_{\rm inv}\, b\bar{b}jj$	$W^\pm\|_{\rm semilep}\, b\bar{b}jj$	$t\bar{t}\|_{\rm semilep}\, b\bar{b}$	$t\bar{t}\|_{\rm semitau}\, b\bar{b}$	$Z\|_{\rm inv}\, b\bar{b}b\bar{b}$
750	275	0.1722	0.0015	0.0009	0.0004	0.0003	0.0015
750	350	0.1650	0.0006	0.0005	0.0005	0.0003	$< 10^{-4}$
1000	275	0.2371	0.0017	0.0016	0.0002	0.0002	0.0012
1000	375	0.1391	0.0014	0.0014	0.0002	0.0005	0.0020
1000	475	0.1061	8.0902	$< 10^{-4}$	0.0001	0.0001	$< 10^{-4}$
1250	275	0.0749	0.0001	0.0001	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
1250	375	0.2561	0.0015	0.0017	$< 10^{-4}$	0.0001	0.0012
1250	475	0.1180	0.0003	0.0006	$< 10^{-4}$	$< 10^{-4}$	0.0005
1250	600	0.1983	0.0004	0.0007	0.0002	0.0004	0.0002
1500	275	0.1914	0.0009	0.0010	$< 10^{-4}$	$< 10^{-4}$	0.0010
1500	375	0.1259	0.0009	0.0002	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
1500	475	0.2324	0.0007	0.0009	$< 10^{-4}$	$< 10^{-4}$	0.0002
1500	600	0.0855	$< 10^{-4}$	0.0001	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
1500	725	0.2298	0.0002	0.0006	0.0001	0.0002	0.0002

Table: Acceptances for `XGBoost` classifiers after applying a probability threshold that maximizes the significance including 30% of systematic uncertainties in the total background yield.