Probing the CP nature of the Higgs couplings in $b\bar{b}h$ and $t\bar{t}h$ at the LHC

Braga meeting – Universidade do Minho, Escola de Ciências

Rodrigo Capucha

Collaboration: A. Onofre, D. Azevedo, R. Santos

30-31 January 2020

Introduction

- In **2012**, a new scalar particle with a mass close to 125 GeV, later identified as **the Higgs boson, was discovered at the Large Hadron Collider (LHC)**. **Electroweak symmetry breaking** confirmed and SM completed. ATLAS, Phys. Lett. B716 (2012), and CMS, Phys. Lett. B 716 (2012).

- Tested in a wide range of precise experimental measurements at colliders. **No experimental result strongly deviates from its predictions** (**few exceptions:** B meson decays and muon's anomalous magnetic moment).

- But **it is incomplete**. Some unanswered question: **dark matter, matter-antimatter asymmetry, gravity**.

CP-Violation (CPV)

- **CP:** symmetry that changes particle into antiparticle and inverts the spatial coordinates of a particle.

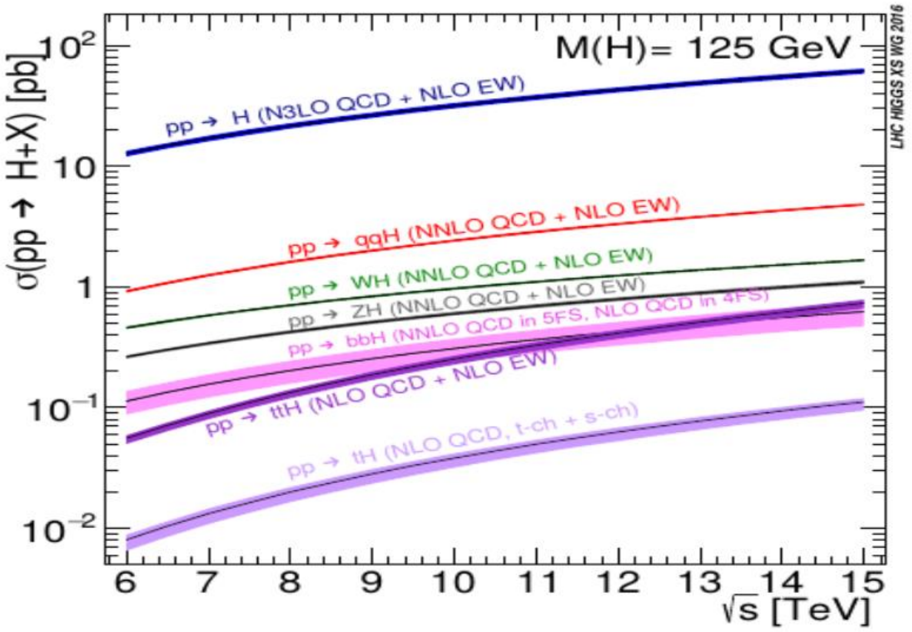
- Why does it matter? According to **Sakharov's conditions for baryogenesis**, matter-antimatter discrepancy may be explained if CP is violated.

- Is there **CPV in the SM?** Yes, in the Yukawa sector, but **not enough**. **New sources of CPV are needed**. A possibility is to add CPV in an extended scalar sector (like the C2HDM hep-ph/0211371).

- How to look for CPV at the LHC?

CP nature of the Higgs

The **SM Higgs** is **CP-even**. What about the **discovered scalar**? **Pure CP-odd Higgs already ruled out** at 99.98% in VH production Phys. Lett. B 759, 672 (2016). Is it CP-even? Not necessarily. It **may have a CP-even and a CP-odd component**. This would imply **CPV in the scalar sector.**



- The Higgs CP can be probed at the LHC by looking for CP-sensitive observables in Higgs decays into boson pairs, fermions, or in Higgs production channels. **hVV (V=Z,W) couplings: only** the **CP-even component** is **projected out**. **Fermions: both** components can **contribute equally**. Precise **measurement of** the **Yukawa couplings still lacking**.

- **Two processes** were considered: $b\bar{b}h$ **and** $t\bar{t}h$. **Complicated backgrounds** and relatively **low rates**, but we are **directly probing the vertices**. Difficult channel, but $t\bar{t}h$ **has already been observed.** CMS, Phys. Rev. Lett. 120 (2018) and ATLAS, Phys. Lett. B 784 (2018)

The Model

- **SM + generic CP-violating Yukawa coupling** for f = b, t.

$$\mathcal{L}_{hf\bar{f}} = -\sum_f g_{hf\bar{f}} \kappa_{hf\bar{f}} \bar{f} \left(\cos\alpha + i \sin\alpha \gamma_5\right) f h,$$

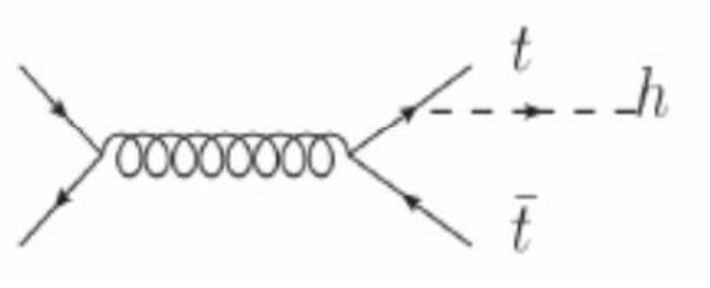
- $\boldsymbol{g_{hf\bar{f}}}$ are the SM couplings. We set $\boldsymbol{k_{hf\bar{f}}}$ **= 1**. **Two limits** considered: **CP-even** (α = 0, h = H) and **CP-odd** (α = π/2, h = A).

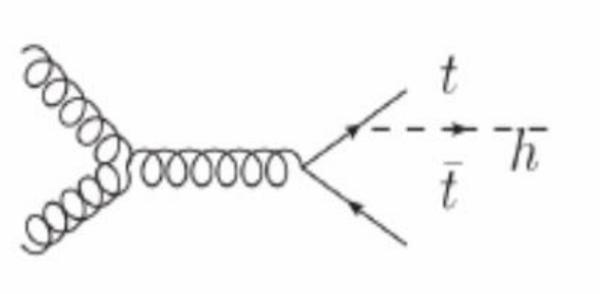
- For $\boldsymbol{b\bar{b}h}$**:** h → $\tau^+\tau^-$. m_h= 10 – 125 GeV.

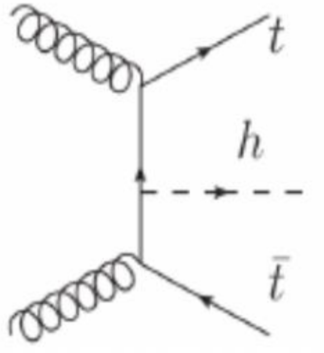
- For $\boldsymbol{t\bar{t}h}$**:** h → $b\bar{b}$, t ($\bar{t}$) → W^+b ($W^-\bar{b}$) and $W^+(W^-)$ → $l^+\upsilon_l(l^-\bar{\upsilon}_l)$ (**dileptonic state**), with $\boldsymbol{l} = \boldsymbol{e}, \boldsymbol{\mu}$. m_h= 40 – 500 GeV.

Generation of signal and background events

- Generation of **pp colisions at 13 TeV** done with the Monte Carlo event generator **MadGraph5_aMC@NLO.** JHEP **1407**, 079 (2014)
- h = H, A implemented with the **HC_NLO_X0 model**, with **NLO** corrections. JHEP 1311 (2013) Backgrounds simulated with the **SM implementation**.
- **MadSpin** used to decay heavy resonances (t, h, and V) preserving full spin correlations. JHEP **1303**, 015 (2013)
- **Pythia 6** used for showering and hadronization. JHEP **0605**, 026 (2006)
- **DELPHES 3** used for a fast detector simulation of a LHC-like experiment. JHEP **1402**, 057 (2014)
- Analysis done with **MadAnalysis 5**. EPJC **74, no 10, 3103 (2014)**







Topology	Order	Generated cross-section (pb)
$t\bar{t}H^{SM}$	NLO	0.025
$t\bar{t}b\bar{b}$	NLO	0.79
$t\bar{t}$+3 jets	LO	37.89
$t\bar{t}V$+ jets	LO	0.0618
Single top s-channel	LO	2.1916
Single top t-channel+ jets	LO	46.863
Single top Wt-channel	LO	15.1827
W+4 jets	LO	34500
$Wb\bar{b}$+2 jets	LO	289
Z+4 jets	LO	3120
$Zb\bar{b}$+2 jets	LO	123
WW+3 jets	LO	84.2
WZ+3 jets	LO	37.9
ZZ+3 jets	LO	11

- **Backgrounds (only for $t\bar{t}h$):** all processes that can lead to 4 jets and two leptons of opposite charge. Jets: gluons or non-b quarks.

Event Selection and Reconstruction (for $t\bar{t}h$ and backgrounds)

- Results for $\boldsymbol{b\bar{b}h}$ **only at parton level**.

- **Selection cuts:**

➢ $N_{jets} \geq 4$ and $N_{lep} \geq 2$ (one-to-one correspondence)

➢ p_T> 20 GeV and |η| < 2.5 (experimental limitations)

➢ $|m_{l^+l^-} - m_Z|$ > 10 GeV and at least 3 b-tagged jets (after reconstruction).

- **Reconstruction:** matching and actual reconstruction done separately.

➢ Boosted decision tree (BDT) algorithm to assign jets to the b-quarks of the top-quarks and Higgs. Constraints: only 6 jets with highest p_T. $m_{l^+b_t}$< 150, $m_{l^-\bar{b}_{\bar{t}}}$< 150 and 20 < $m_{b_h\bar{b}_h}$< 300 GeV.

➢ After matching, 6 objects are identified. These are used to find the momentum of the remaining particles.

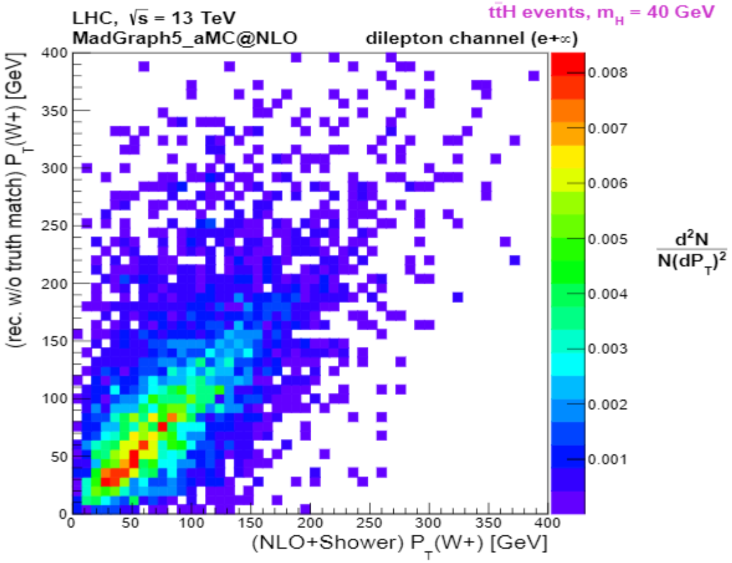
Event Selection and Reconstruction (for $t\bar{t}h$)

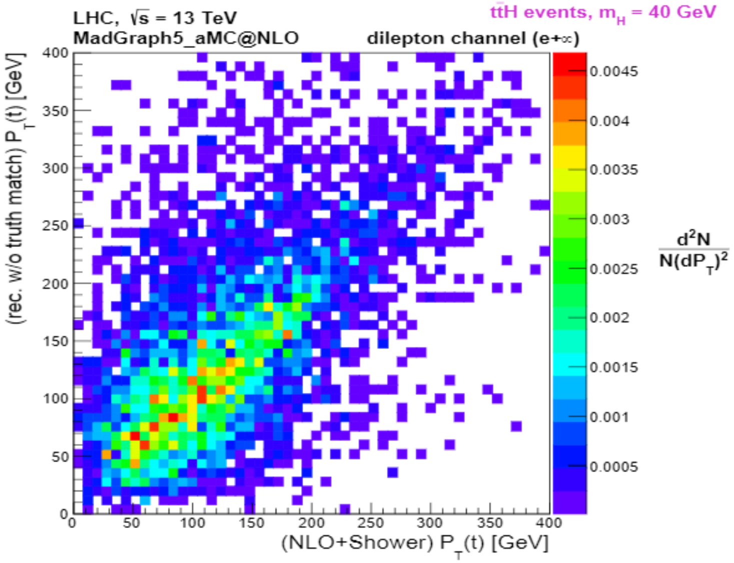
- All **missing energy due to neutrinos**. With m_{υ}= 0, we are left with six unknowns:

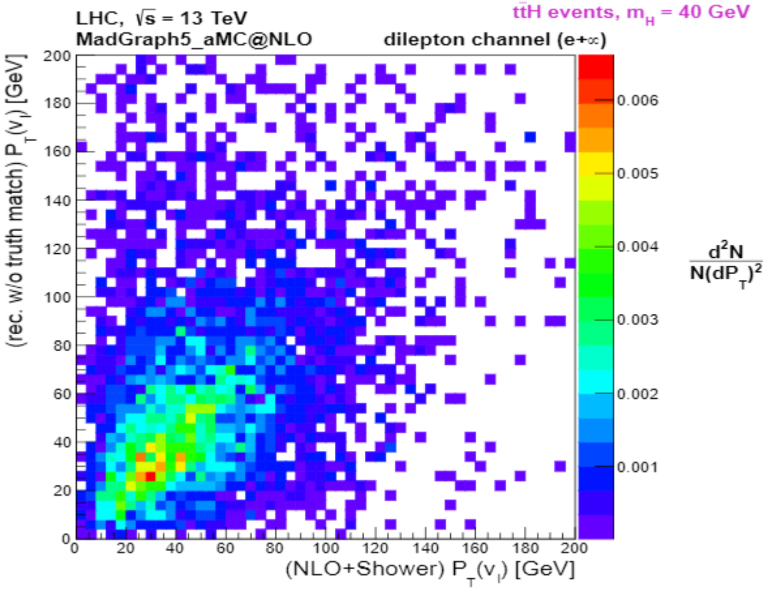
$$
\begin{aligned}
(p_{l^+} + p_\nu)^2 &= m_{W^+}^2 \\
(p_{l^-} + p_{\bar{\nu}})^2 &= m_{W^-}^2 \\
(p_{W^+} + p_{b_t})^2 &= m_t^2 \\
(p_{W^-} + p_{\bar{b}_{\bar{t}}})^2 &= m_{\bar{t}}^2 \\
p_\nu^x + p_{\bar{\nu}}^x &= \not{E}^x \\
p_\nu^y + p_{\bar{\nu}}^y &= \not{E}^y .
\end{aligned}
$$

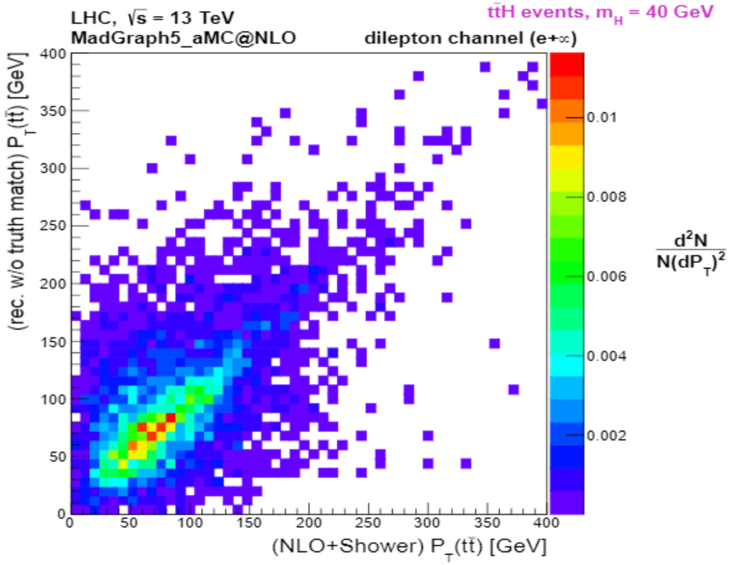
- $m_{W^\pm}$ and $m_t(m_{\bar{t}})$ randomly generated from 2D p.d.f.s. If no solution is found up to 500 trials, event is discarded. If several solutions are found, a **likelihood function is built using parton level information** and the **solution with the highest likelihood is picked**.

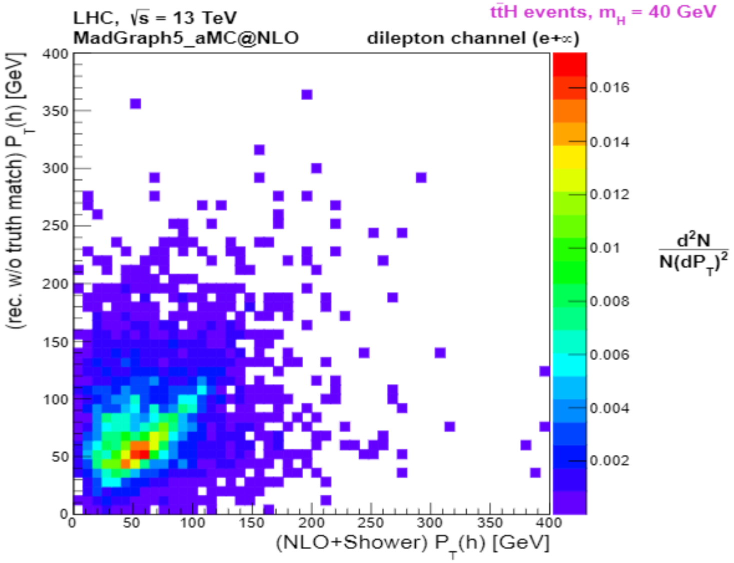
$$
L_{t\bar{t}h} \propto \frac{1}{p_{T_\nu} p_{T_{\bar{\nu}}}} P(p_{T_\nu}) P(p_{T_{\bar{\nu}}}) P(p_{T_t}) P(p_{T_{\bar{t}}}) P(p_{T_{t\bar{t}}}) P(m_t, m_{\bar{t}}) P(m_h),
$$

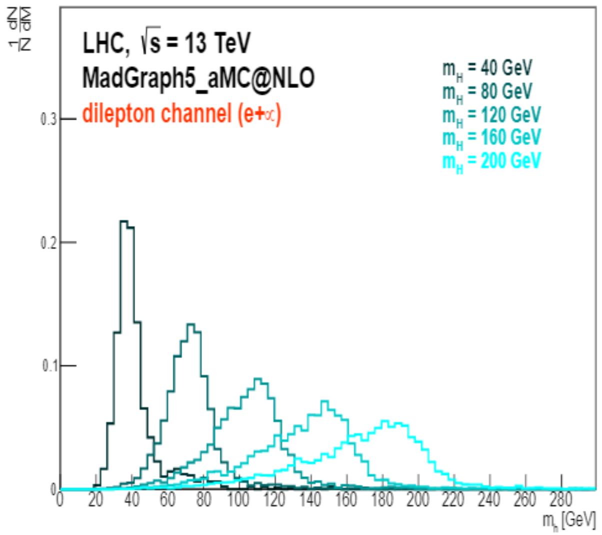

LHC, √s = 13 TeV
MadGraph5_aMC@NLO
tt̄H events, m_H = 40 GeV
dilepton channel (e+∝)
(rec. w/o truth match) P_T(W+) [GeV]
(NLO+Shower) P_T(W+) [GeV]
d²N / N(dP_T)²


LHC, √s = 13 TeV
MadGraph5_aMC@NLO
tt̄H events, m_H = 40 GeV
dilepton channel (e+∝)
(rec. w/o truth match) P_T(t) [GeV]
(NLO+Shower) P_T(t) [GeV]
d²N / N(dP_T)²

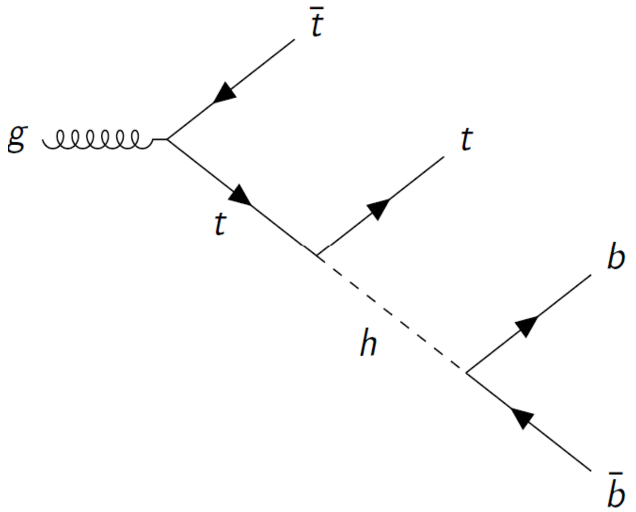

LHC, √s = 13 TeV
MadGraph5_aMC@NLO
tt̄H events, m_H = 40 GeV
dilepton channel (e+∝)
(rec. w/o truth match) P_T(ν_l) [GeV]
(NLO+Shower) P_T(ν_l) [GeV]
d²N / N(dP_T)²

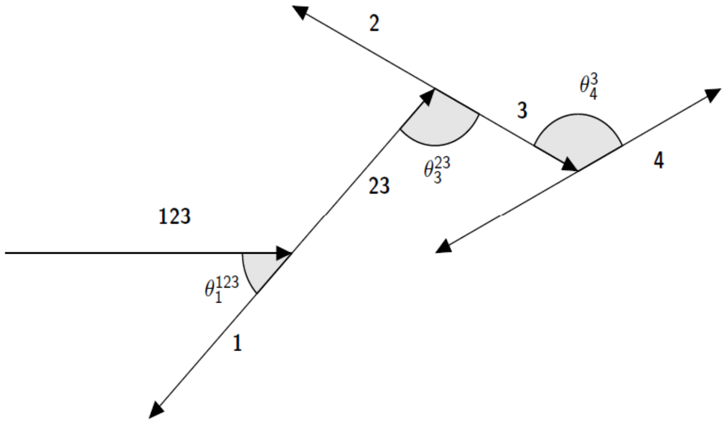

LHC, √s = 13 TeV
MadGraph5_aMC@NLO
tt̄H events, m_H = 40 GeV
dilepton channel (e+∝)
(rec. w/o truth match) P_T(tt̄) [GeV]
(NLO+Shower) P_T(tt̄) [GeV]
d²N / N(dP_T)²


LHC, √s = 13 TeV
MadGraph5_aMC@NLO
tt̄H events, m_H = 40 GeV
dilepton channel (e+∝)
(rec. w/o truth match) P_T(h) [GeV]
(NLO+Shower) P_T(h) [GeV]
d²N / N(dP_T)²


LHC, √s = 13 TeV
MadGraph5_aMC@NLO
dilepton channel (e+∝)
m_H = 40 GeV
m_H = 80 GeV
m_H = 120 GeV
m_H = 160 GeV
m_H = 200 GeV
1/N dN/dM
m_h [GeV]

CP observables





- **Two types of observables**, based on previous works. Designed to increase the sensitivity in discriminating CP signals from irreducible backgrounds, but also to probe the CP nature of the Yukawa coupling.

- **Angular distributions θ_Y^X:** Phys. Rev. D 92, 034021 (2015); Phys. Rev. D 96 (2017)

➢ angle between the 4-momenta of the Y system, measured in the rest frame of X, with respect to the direction of X in the rest frame of its mother.

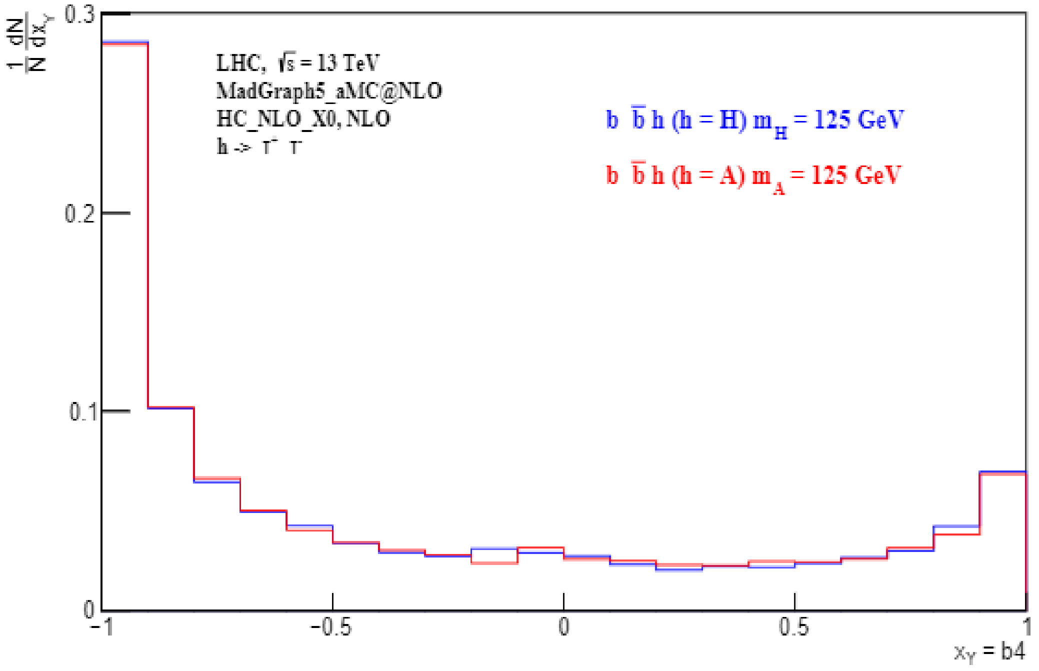
➢ Two ways to compute: direct or sequential boost.

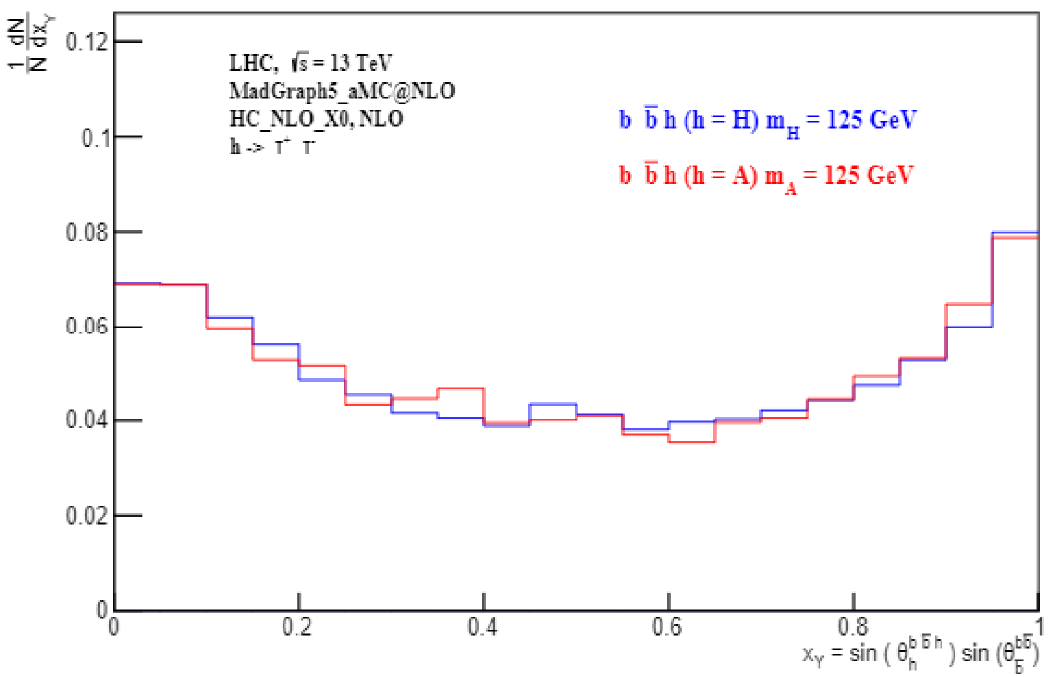
➢ Set of functions $f(\theta_Y^X)g(\theta'^X_Y)$ were considered, with f, g, simple trigonometric functions.

- **Gunion-He variables** (both in the LAB and $t\bar{t}h$ CM system): Phys. Rev. Lett. 76 (1996)

$$b_2 = (\vec{p}_t \times \hat{k}_z).(\vec{p}_{\bar{t}} \times \hat{k}_z)/|\vec{p}_t||\vec{p}_{\bar{t}}|, \qquad b_4 = (p_t^z.p_{\bar{t}}^z)/(|\vec{p}_t|.|\vec{p}_{\bar{t}}|)$$

Results for $b\bar{b}h$

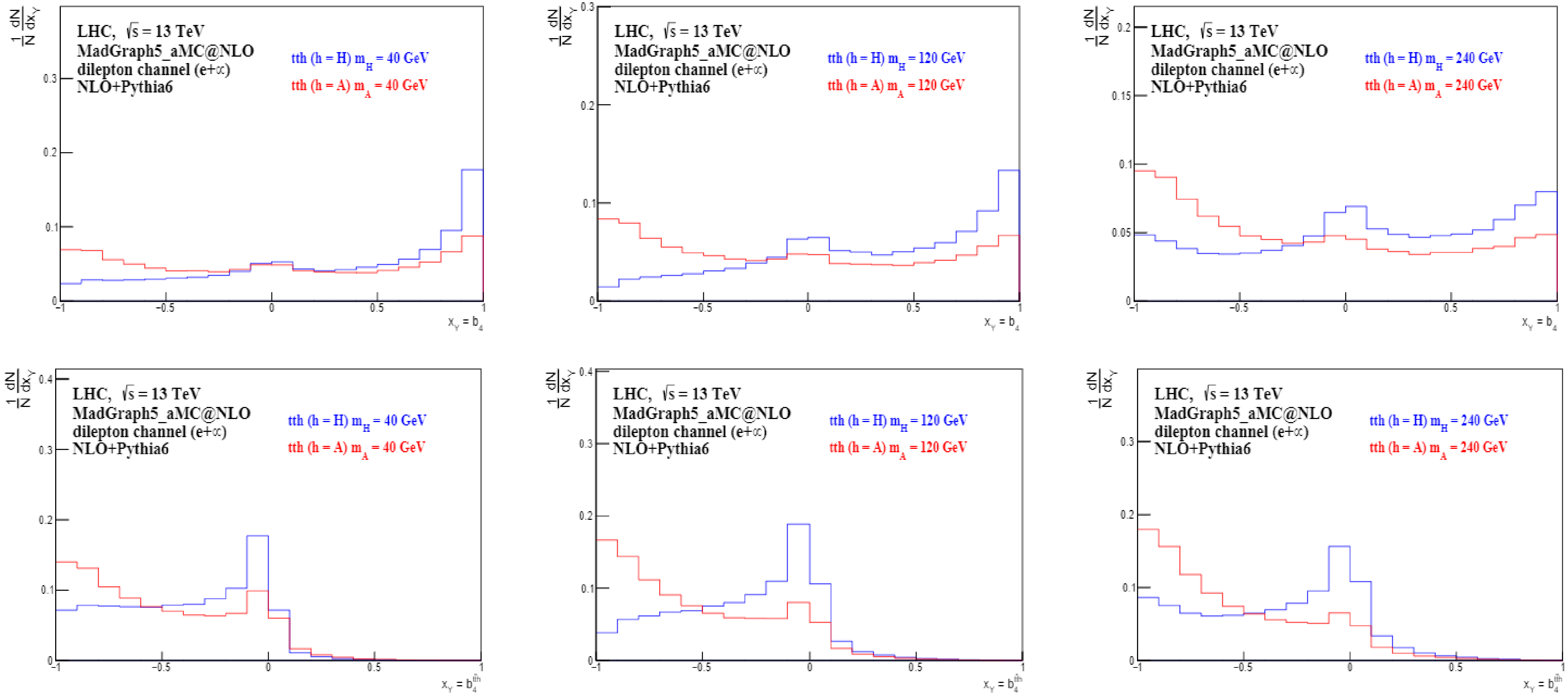




- **No visible asymmetry for $b\bar{b}h$.** Explicit evaluation of the $f\bar{f}h$ production shows that the **asymmetry term is proportional to m_f^2.** Only significant when m_f is big enough relative to m_h. Phys. Rev. Lett. 76 (1996)
- What if we lower m_h (down to 10 GeV)? Still nothing.
- What about single bottom? Also nothing.

Results for $t\bar{t}h$

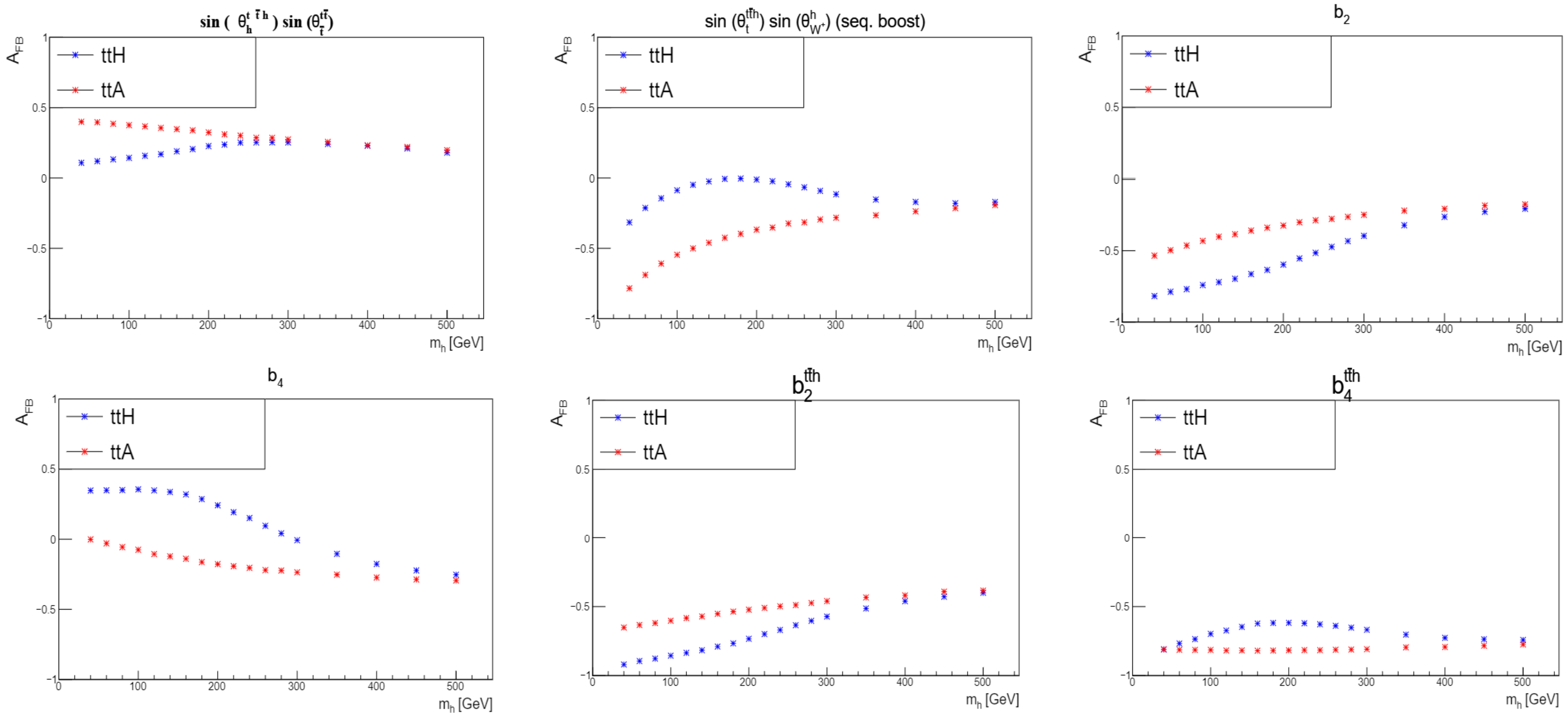
- In **previous works** m_h**= 125 GeV. Here** m_h**= 40-300 GeV.**



- **For the tops significant differences are found**.

Asymmetries

$$A_{FB}^{Y} = \frac{\sigma(x_Y > x'_Y) - \sigma_Y(x_Y < x'_Y)}{\sigma_Y(x_Y > x'_Y) + \sigma_Y(x_Y < x'_Y)},$$



- **Asymmetry vanishes for very large Higgs masses**. Exact value depends on the variable.

Expected confidence levels (CL) for $t\bar{t}h$

- Computed from **likelihood ratios** obtained from binned distributions of the observables, as a function of the integrated luminosity. Both **signal and backgrounds are** described by **Poisson distributions.**
- **Only statistical uncertainties** are considered. **Dileptonic channel alone.**
- Shown up to the **High Luminosity LHC (HL-LHC)**, maximum expected is 3000 fb^{-1}, for m_h= 40, 80, 120, 160 and 200 GeV.

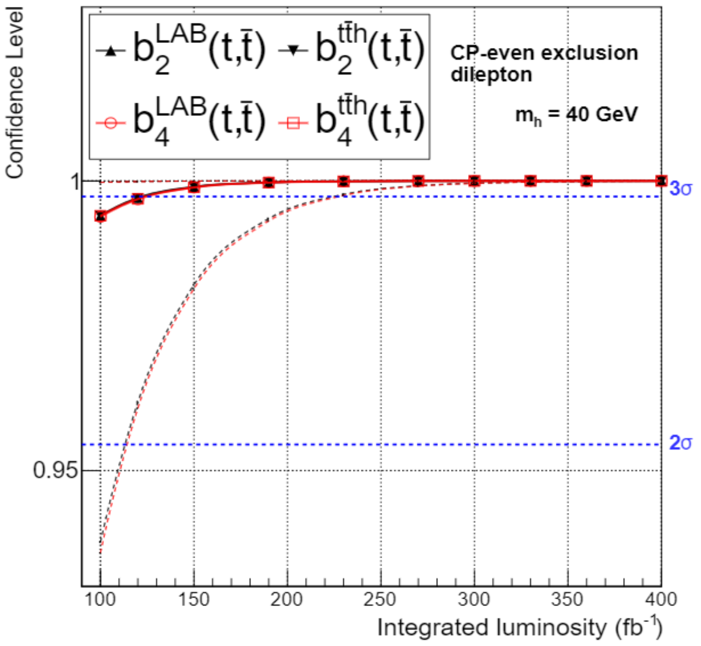
- **Four different scenarios:**

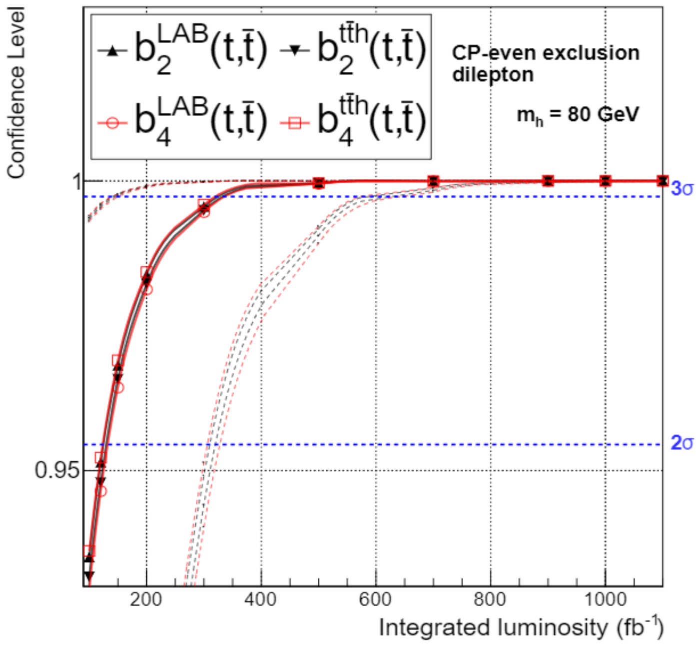
➢ 1: **CP-even exclusion**. H_0= SM and H_1= SM+CP-even signal.

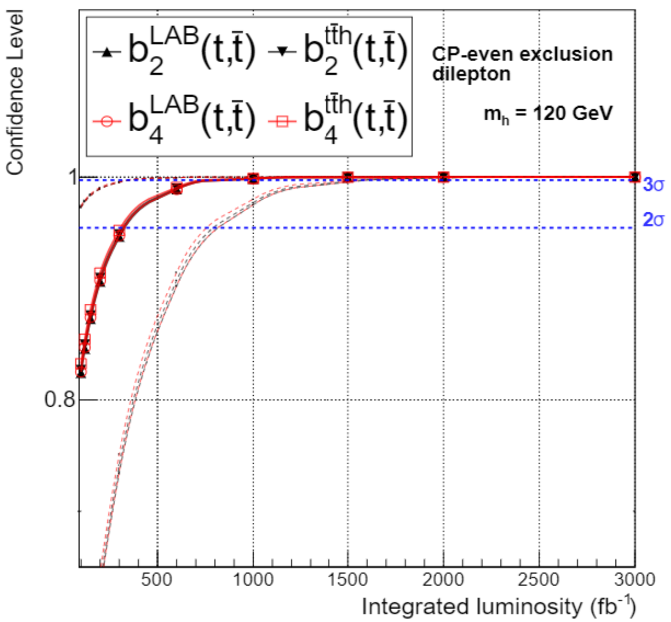
➢ 2: **CP-odd exclusion**. H_0= SM and H_1= SM+CP-odd signal.

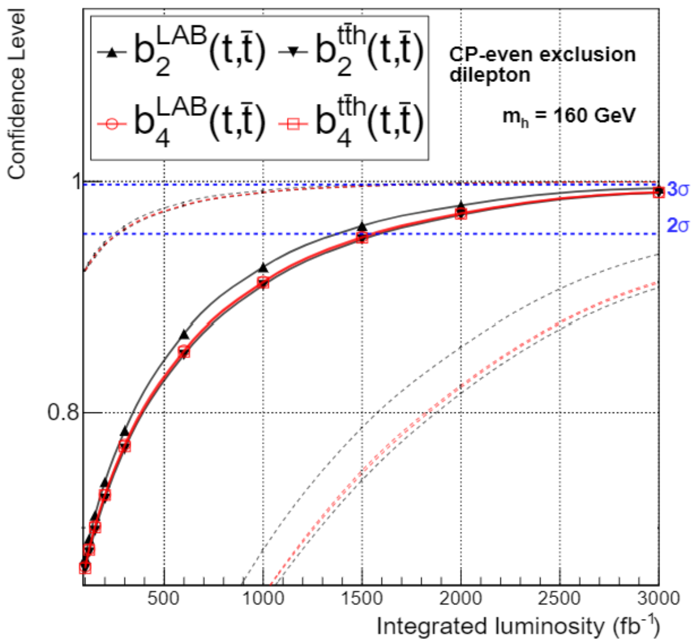
➢ 3: **CP-odd exclusion (vs CP-even)**. H_0= SM +CP-even signal and H_1= SM+CP-odd signal.

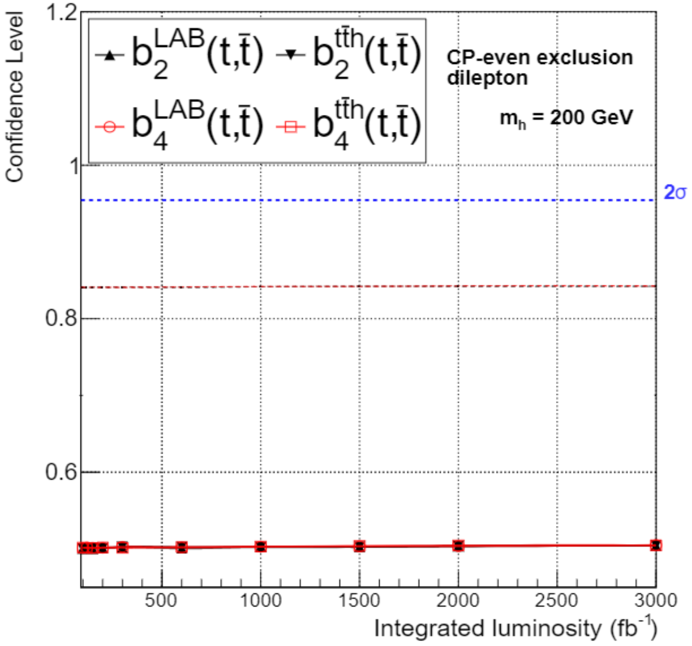
➢ 4: **SM exclusion (vs CP-even)**. H_0= SM+CP-even signal and H_1= SM.



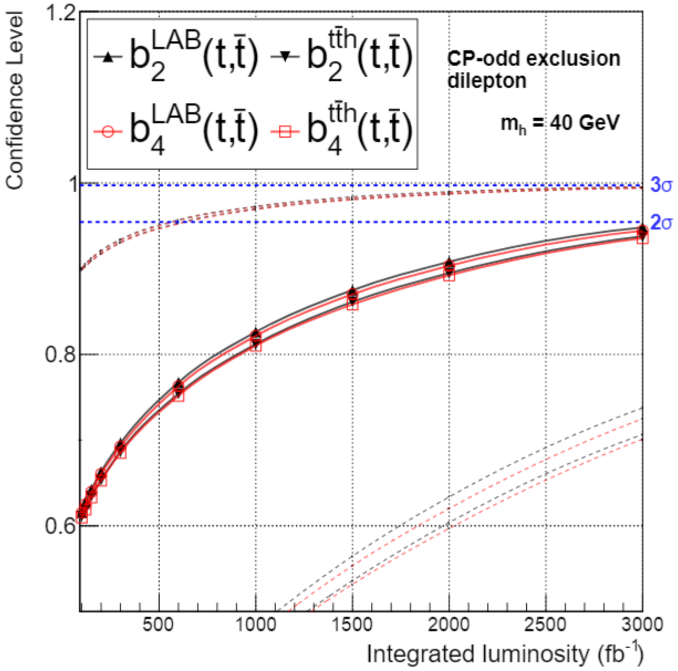


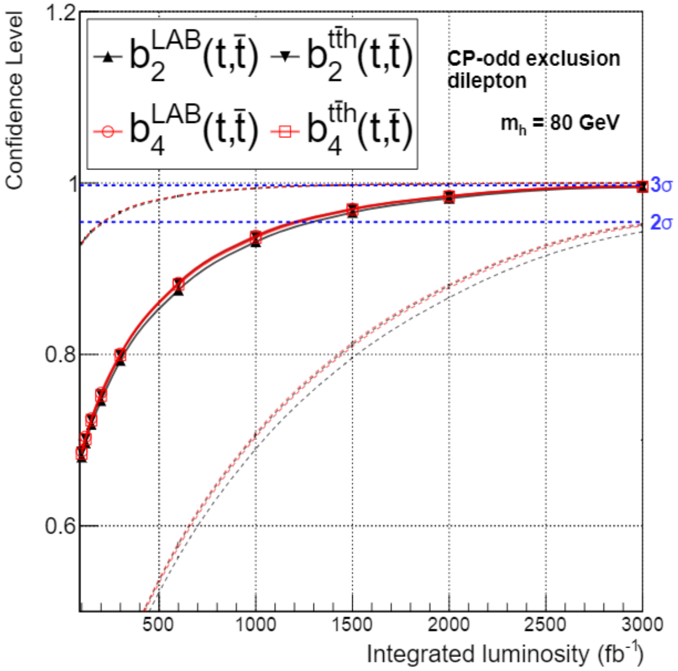


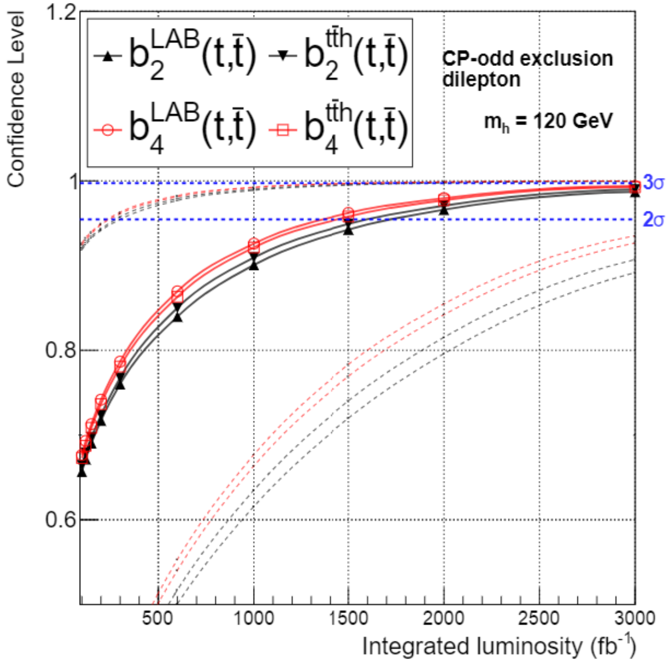


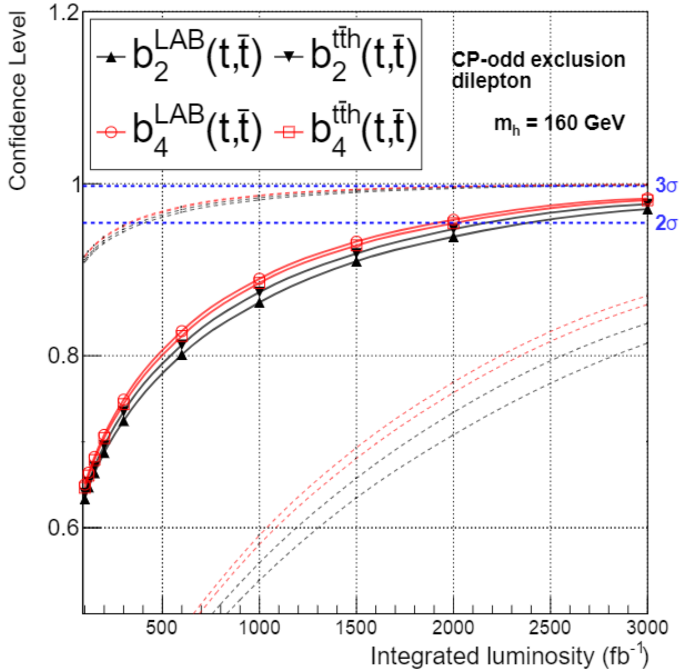


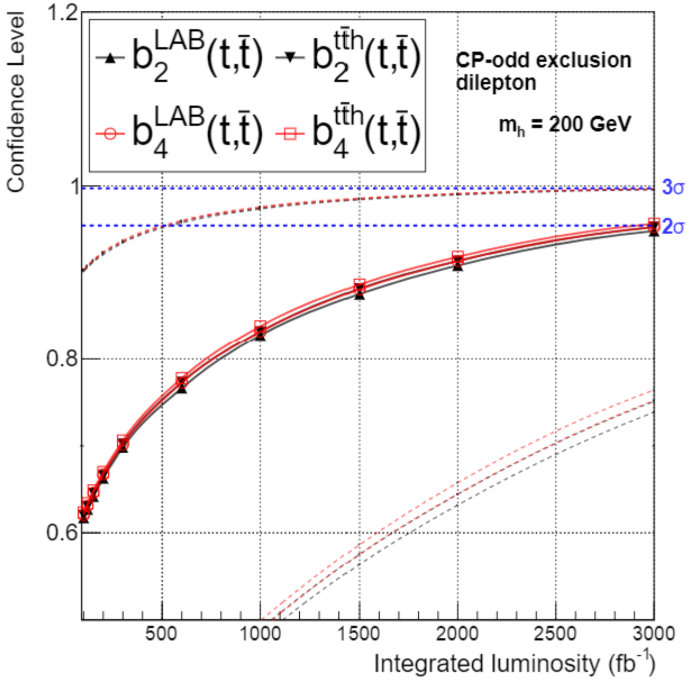
- **CP-even with m_h= 40, 80 GeV can already be excluded** at the LHC using the dilepton channel alone.
- **Heavier masses harder to exclude**, since $\sigma_{t\bar{t}h}$ decreases with m_h.
- Scenarios 1 and 4 give similar results.





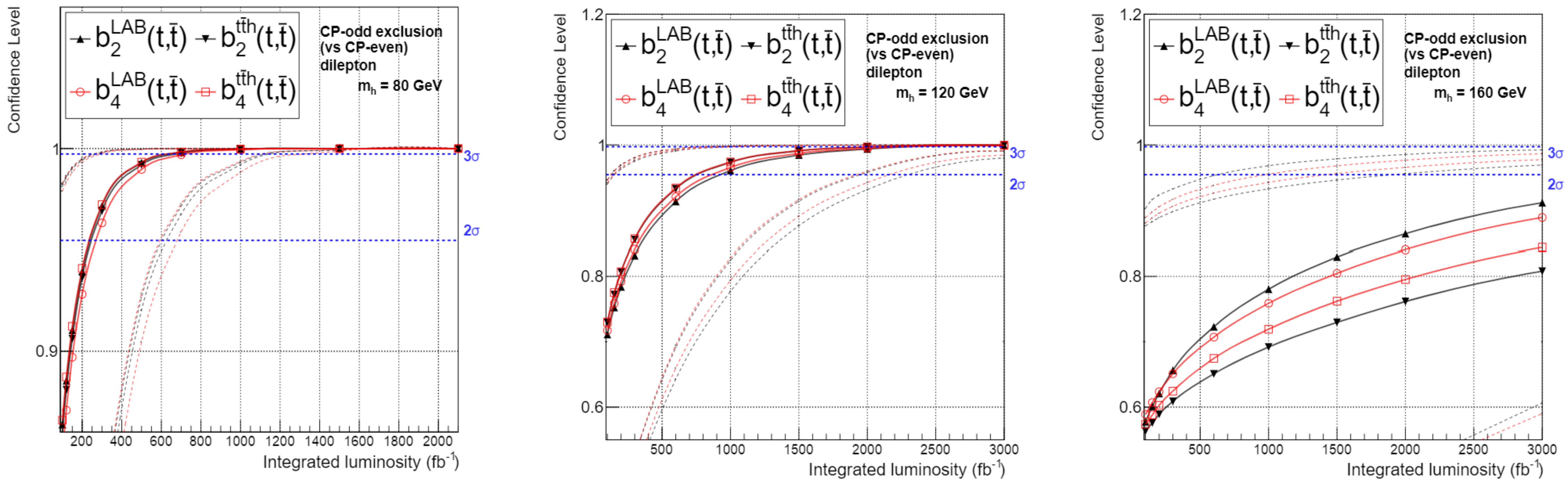






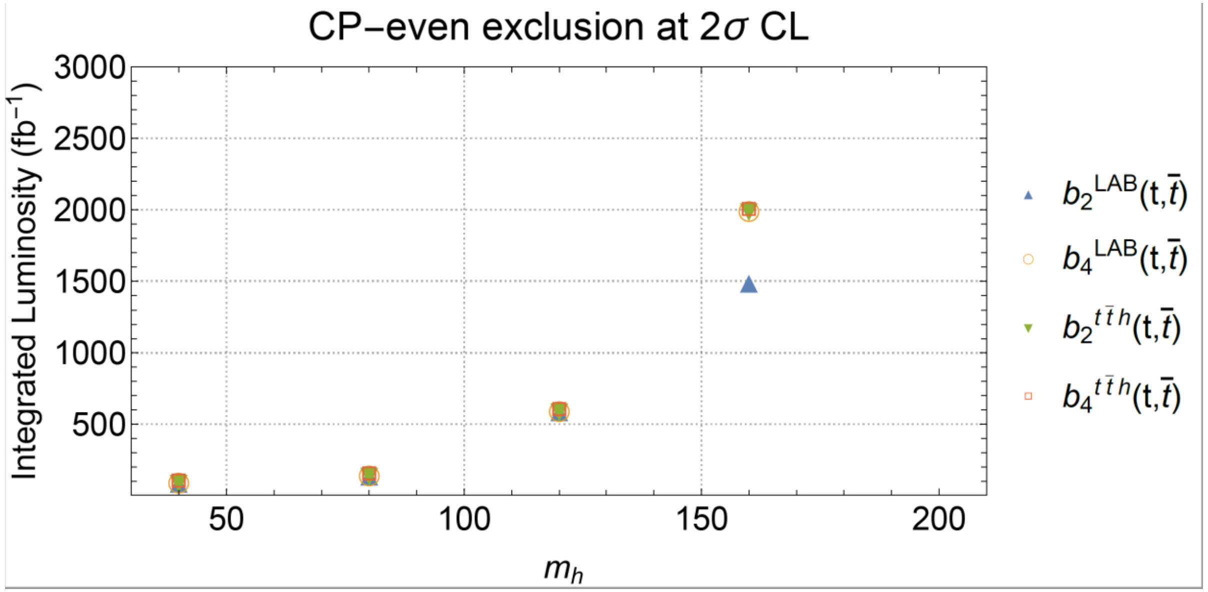
- **CP-odd harder to exclude for lighter masses, but better for heavier masses.**

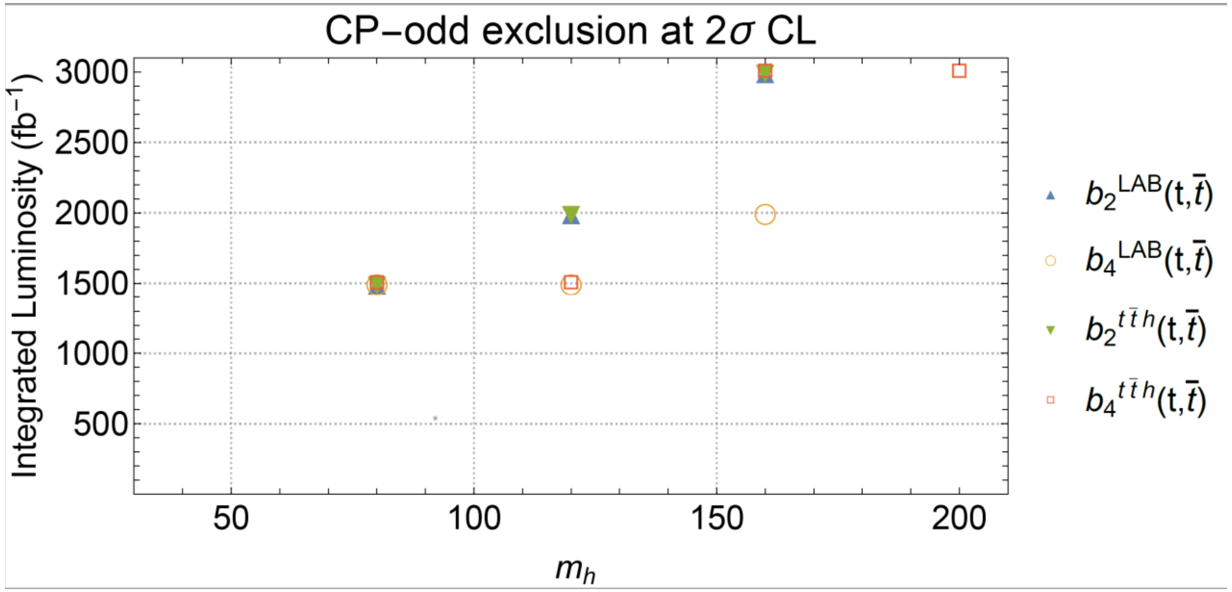
- $\sigma_{t\bar{t}A}$ starts lower than $\sigma_{t\bar{t}H}$. Decreases more smoothly. Around m_h= 160-180 GeV, it becomes larger.

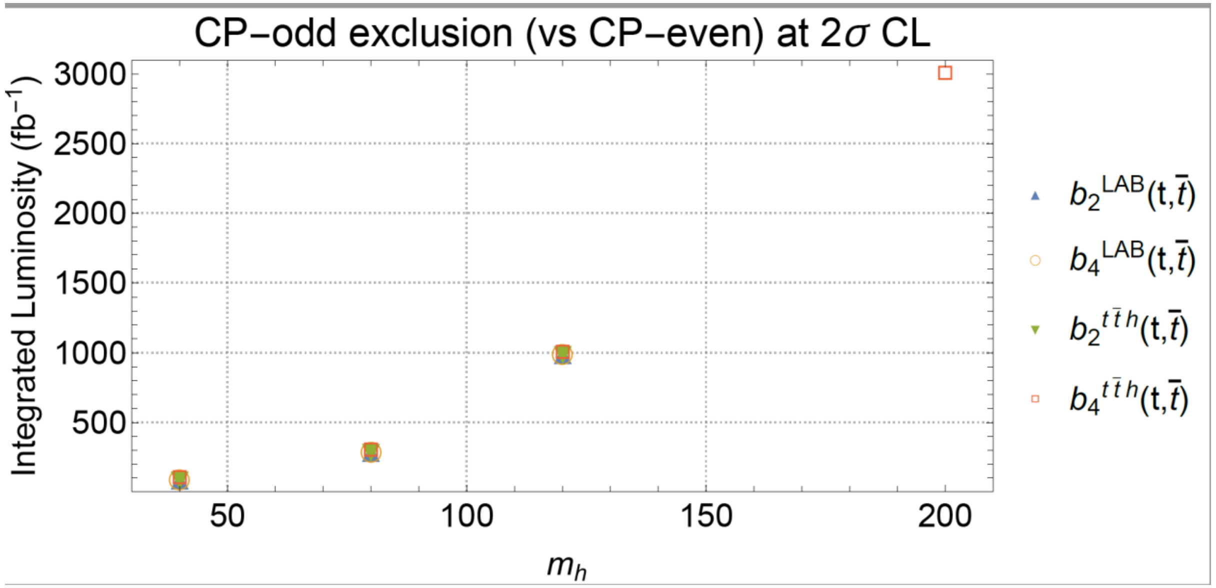


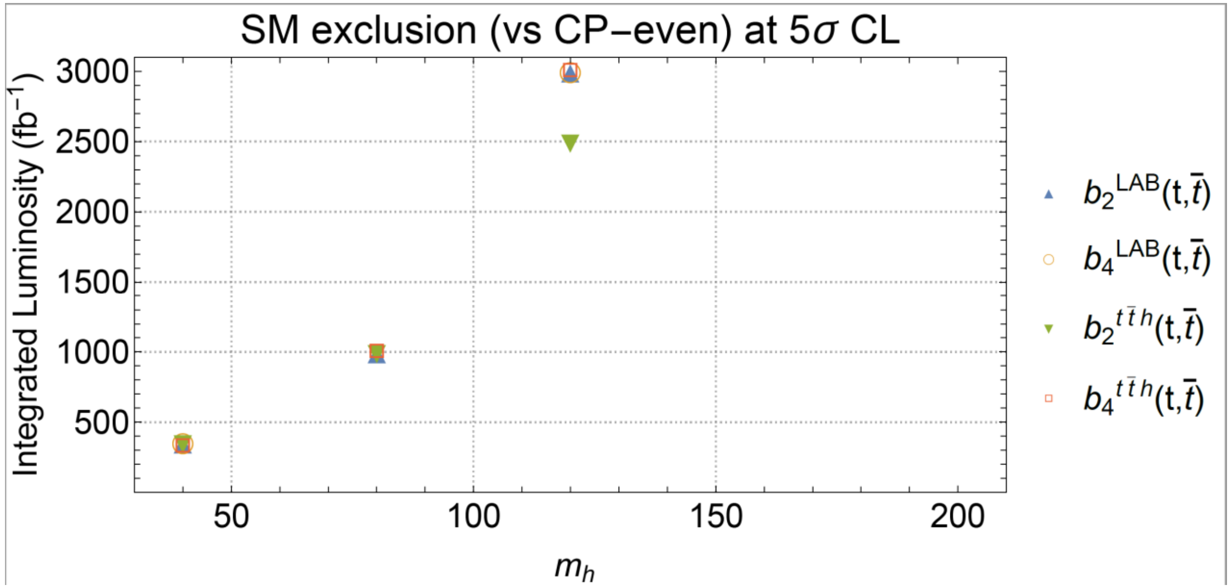
- m_h= 40 and m_h = 200 similar to CP-even and CP-odd exclusion, respectively.

- **If CP-even is found, a pure CP-odd can be excluded** from a pure CP-even scalar.


CP−even exclusion at 2σ CL
Integrated Luminosity (fb^{-1})
3000
2500
2000
1500
1000
500
50
100
150
200
m_h
$b_2^{\mathrm{LAB}}(t,\bar{t})$
$b_4^{\mathrm{LAB}}(t,\bar{t})$
$b_2^{t\bar{t}h}(t,\bar{t})$
$b_4^{t\bar{t}h}(t,\bar{t})$


CP−odd exclusion at 2σ CL
Integrated Luminosity (fb^{-1})
3000
2500
2000
1500
1000
500
50
100
150
200
m_h
$b_2^{\mathrm{LAB}}(t,\bar{t})$
$b_4^{\mathrm{LAB}}(t,\bar{t})$
$b_2^{t\bar{t}h}(t,\bar{t})$
$b_4^{t\bar{t}h}(t,\bar{t})$

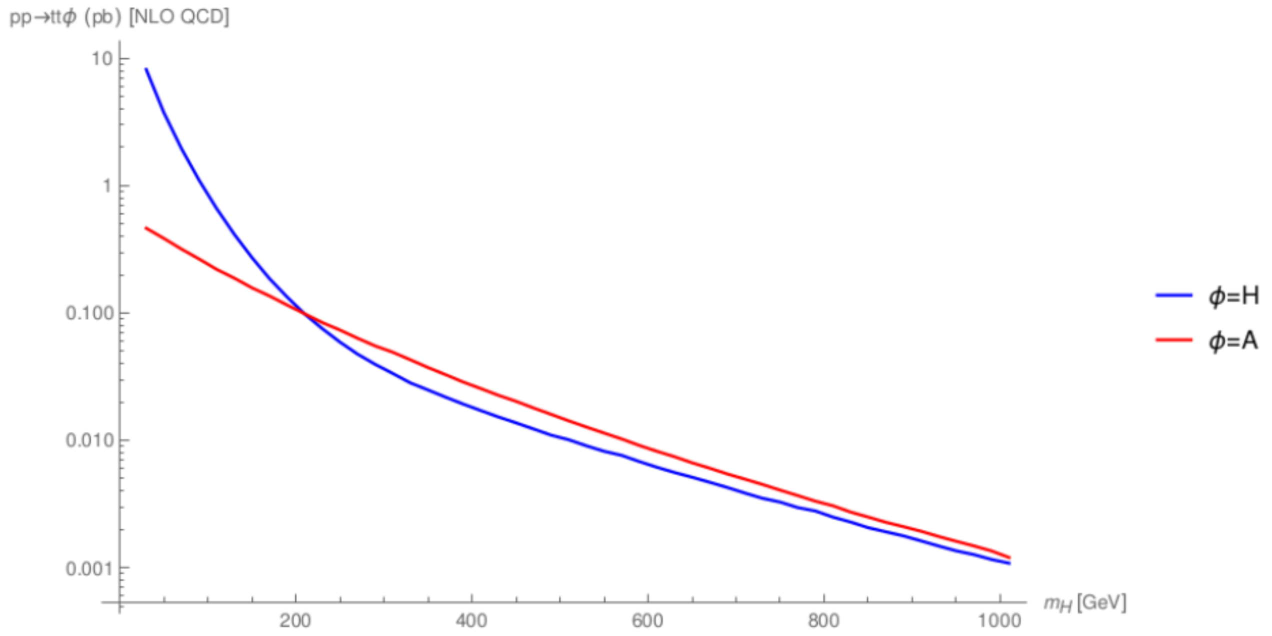

CP−odd exclusion (vs CP−even) at 2σ CL
Integrated Luminosity (fb^{-1})
3000
2500
2000
1500
1000
500
50
100
150
200
m_h
$b_2^{\mathrm{LAB}}(t,\bar{t})$
$b_4^{\mathrm{LAB}}(t,\bar{t})$
$b_2^{t\bar{t}h}(t,\bar{t})$
$b_4^{t\bar{t}h}(t,\bar{t})$


SM exclusion (vs CP−even) at 5σ CL
Integrated Luminosity (fb^{-1})
3000
2500
2000
1500
1000
500
50
100
150
200
m_h
$b_2^{\mathrm{LAB}}(t,\bar{t})$
$b_4^{\mathrm{LAB}}(t,\bar{t})$
$b_2^{t\bar{t}h}(t,\bar{t})$
$b_4^{t\bar{t}h}(t,\bar{t})$

Conclusions

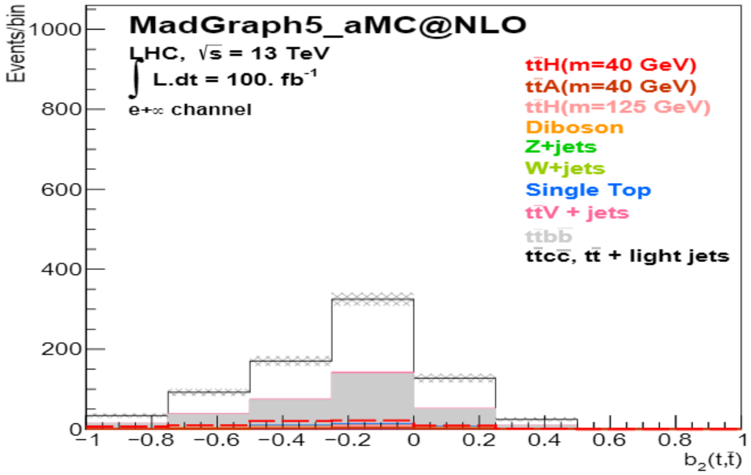
- **Asymmetry term proportional to** $\boldsymbol{m_f^2}$ for $f\bar{f}h$ production.
- Very **difficult to determine the nature of the bottom quark Yukawa couplings** at the LHC, even for very light Higgses.
- **For the tops, significant differences found** for several Higgs masses. **Those vanish for large scalar masses.**
- Expected confidence levels computed in four different scenarios. **Lighter CP-even can already be excluded at the LHC using the dilepton channel alone**. **Heavier CP-even scalars not excluded**. **CP-odd exclusion requires higher luminosity, but improves for heavier masses. If new CP-even Higgs is found, CP-odd exclusion is possible.**

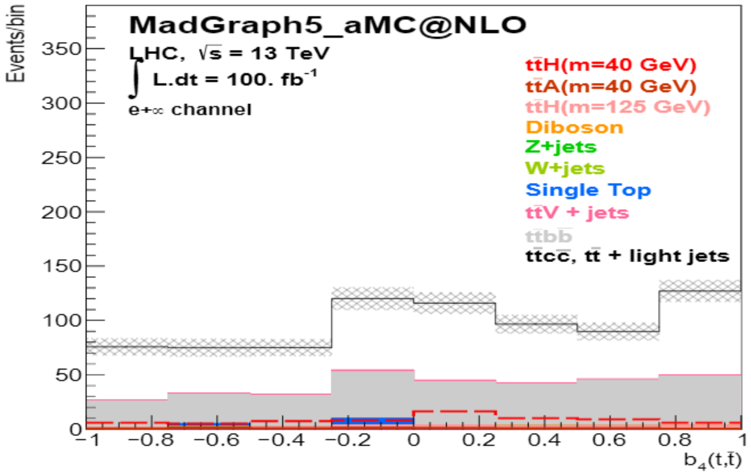
$\sigma_{t\bar{t}h}$ vs m_h

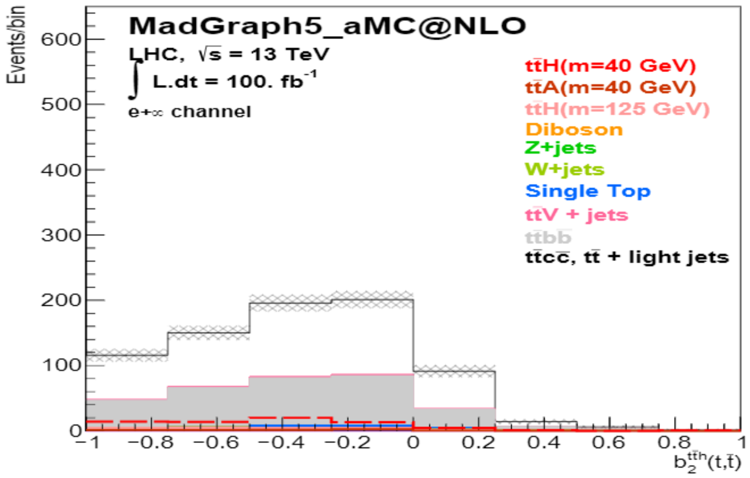


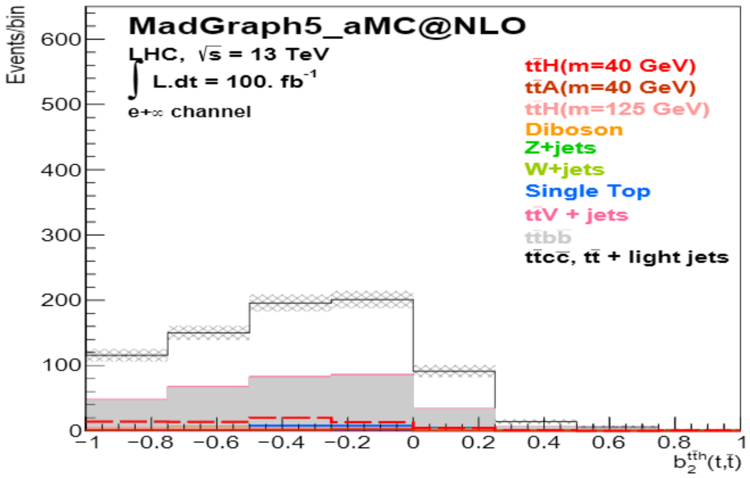
- Computed in MadGraph5_aMC@NLO at NLO. No decays.

Signal vs Background that go into CL









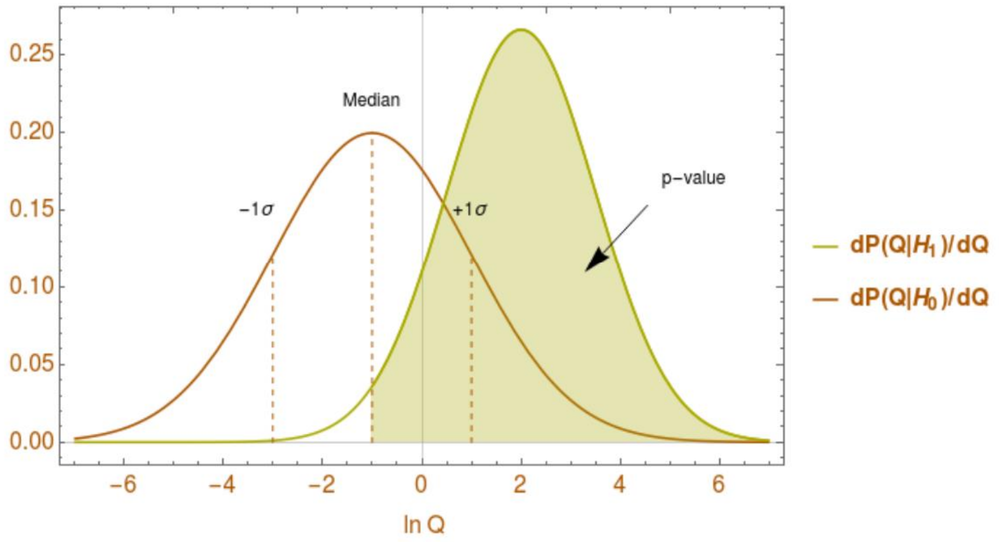
Confidence levels computation

$$\ln Q = -(\lambda_{1\mathrm{tot}} - \lambda_{0\mathrm{tot}}) + \sum_{i=1}^{N_{chan}} n_i \ln\left(\frac{\lambda_{1i}}{\lambda_{0i}}\right).$$

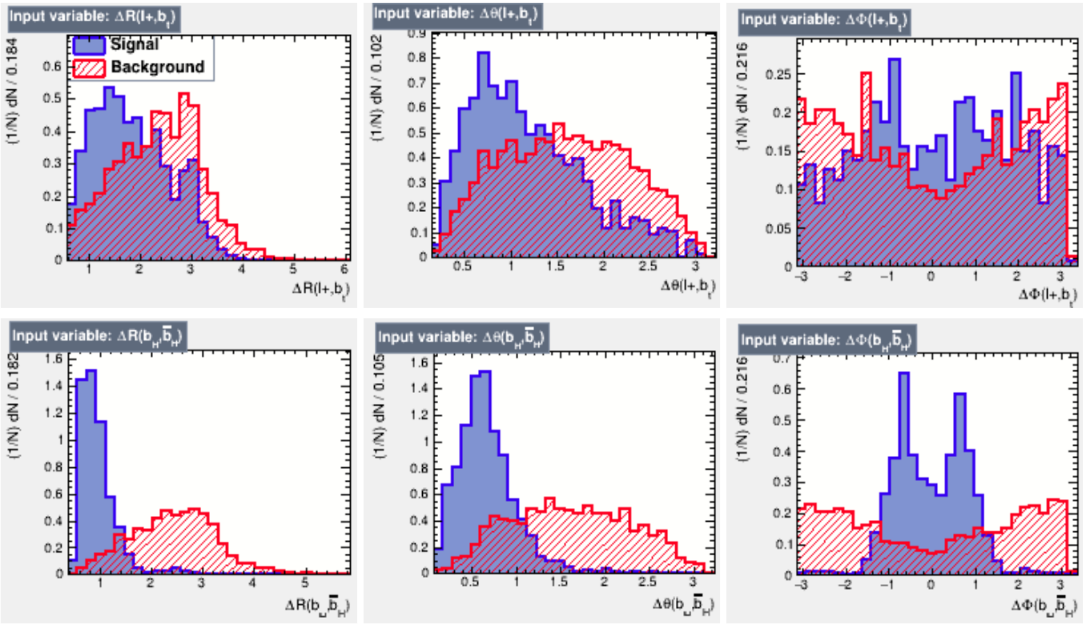
- n_i - number of observed events.
- λ_{1tot} and λ_{0tot} - total number of predicted events assuming H_1 or H_0.

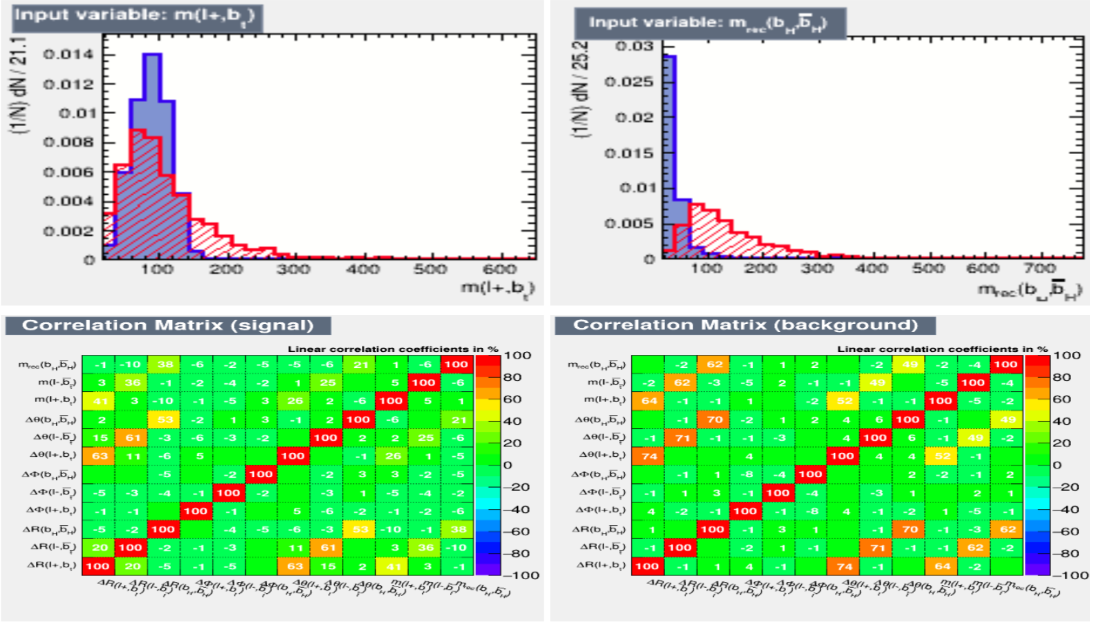
$$CL_s = P\left(Q \geq Q_{obs} \middle| H_1\right),$$

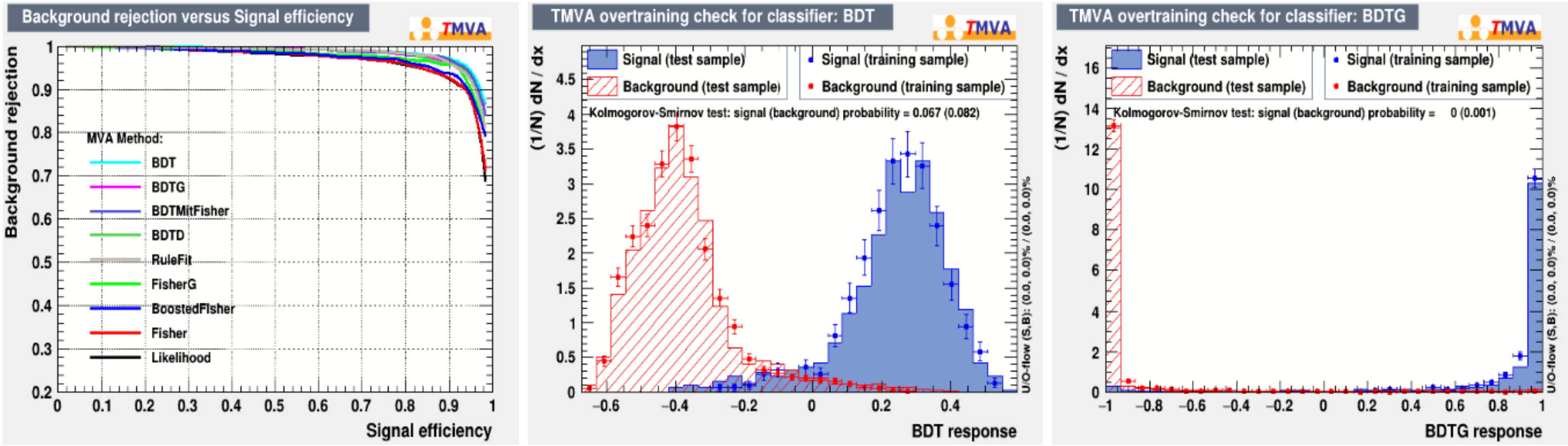
- 1 000 000 toy experiments
- n_i random, using λ_{1i} or λ_{0i} for mean value.
- ln Q in each experiment is computed, for H_1 and H_0.
- ln Qobs = Median of dP(Q|H_0)/dQ



Reconstruction plots







Reconstruction plots

	Efficiency (%)
$N_{jets} \geq 4$ & $N_{lep} \geq 2$	5-20
Matching with TM	18-61
Reconstruction with TM	66-73
Reconstruction without TM	49-63