Prospects to measure the di-Higgs production at the HL-LHC

Sebastian Olivares on behalf of the ATLAS di-Higgs prospects analysis group

Pontificia Universidad Católica de Chile

Thursday August 30th, 2018

Introduction

- Recent efforts from ATLAS and CMS collaborations to measure the di-Higgs coupling with 13 TeV data independently, reached new limits of almost 10 times the SM expectation

- A proper measurement of the Higgs self-interaction term might unveil characteristics of the Higgs potential that would verify if the symmetry breaking is due to a SM-like Higgs sector

- The following slides presents an update of the prospects for the measurement of the di-Higgs production coupling with the ATLAS detector at the HL-LHC, using three different final states:

  - $\mathrm{hh} \rightarrow b\bar{b}b\bar{b}$
  - $\mathrm{hh} \rightarrow b\bar{b}\gamma\gamma$
  - $\mathrm{hh} \rightarrow b\bar{b}\tau^{+}\tau^{-}$

➢ A first attempt to combine these channels is given at the end

❖ **Please take into consideration that the results showed here are not approved by ATLAS yet**

$hh \rightarrow b\bar{b}\gamma\gamma$

- **Analysis used 14 TeV samples**, with the upgraded detector geometry and performance functions, for the HL-LHC at $<\mu> = 200$

- **Event selection:**
  - At least two isolated photons ($p_T > 43\ or\ 30\ \text{GeV}, |\eta| < 2.37\ or\ 1.52 < |\eta| < 2.37$)
  - At least two b-tagged $R = 0.4$ jets ($p_T > 35\ \text{GeV}, |\eta| < 2.5$)
  - Demand less than six jets with $p_T > 35\ \text{GeV}, |\eta| < 2.5$
  - No isolated electrons with $p_T > 30\ \text{GeV}, |\eta| < 2.37\ or\ 1.52 < |\eta| < 2.37$
  - No isolated muons with $p_T > 25\ \text{GeV}, 0.1 < |\eta| < 2.5$

- **Background estimation:**
  - Main background arise from a continuum of multiple jets and photons that mimic the $b\bar{b}\gamma\gamma$ final state
  - A second smaller component come from single Higgs boson production
  - Overlaps between samples are taken into account (e.g. $b\bar{b}\gamma\gamma$ vetoed by $b\bar{b}j\gamma$)

- **Fit $m_{\gamma\gamma}$ spectrum**

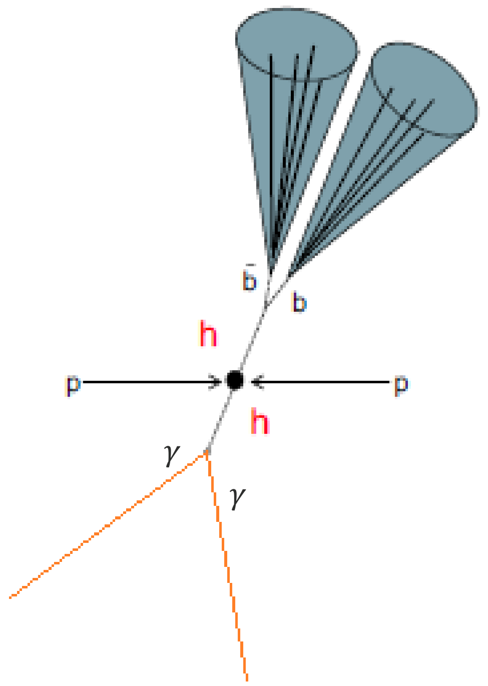

Fig.6: di-Higgs $b\bar{b}\gamma\gamma$ candidate.

$hh \rightarrow b\bar{b}\gamma\gamma$

- **Projections for HL-LHC**
  - In order to increase the available statistics in some samples, **weighted events** from all possible final state combinations are considered as last year's PUB note
  - **b-tagging algorithm** (mv2c10) updated to the most recent ITk layout with improved c-jet rejection is considered
  - The **photon algorithm** used has an improved resolution, with an efficiency of 60% for photons of 50 GeV and 85% above 150 GeV

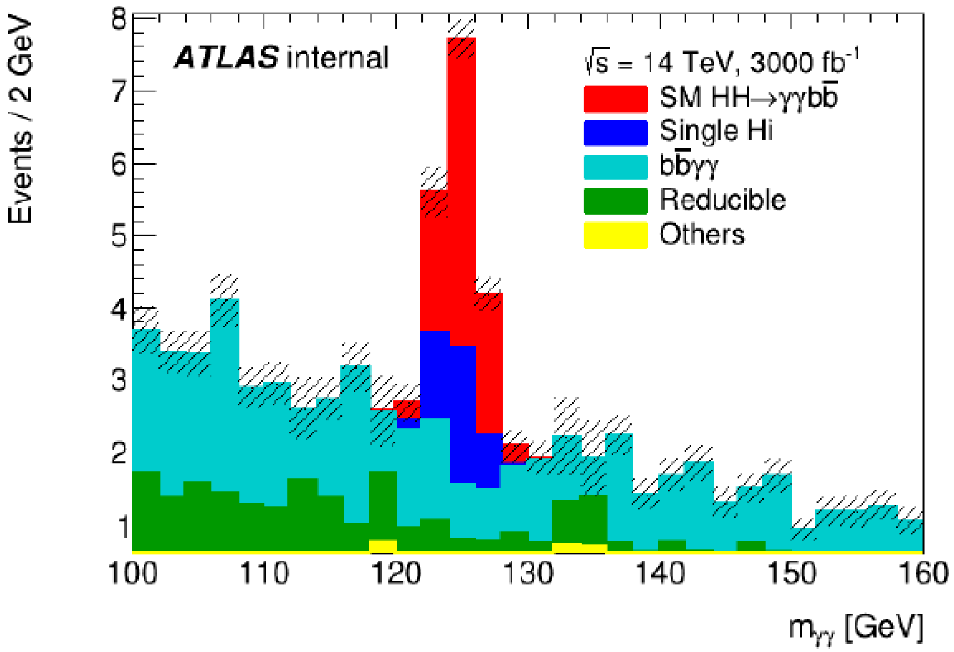

Fig.7: $m_{\gamma\gamma}$ distribution for signal and background samples in the fit region.

- **Multivariate discriminant BDT selection**
  - **BDT (from TMVA) discriminant** is used to improve the separation of signal from background, using 21 variables that provided good discrimination with minimal correlation
  - BDT training done between the range 120 < $m_{\gamma\gamma}$ < 130 GeV to optimize sensitivity in signal region

$hh \rightarrow b\bar{b}\gamma\gamma$

- **Systematic uncertainties**
  - **Theory uncertainties taken from YR4**, except for **ggF QCD** scale (which is assumed with a **100% uncertainty**), conservative PDF scale factors used
  - Experimental uncertainties taken from Run-2 analysis applying scale factors to HL-LHC
  - **Few percentage impact on final results**

- **Limits**

	significance
w/o syst.	2.34σ
with syst	2.29σ

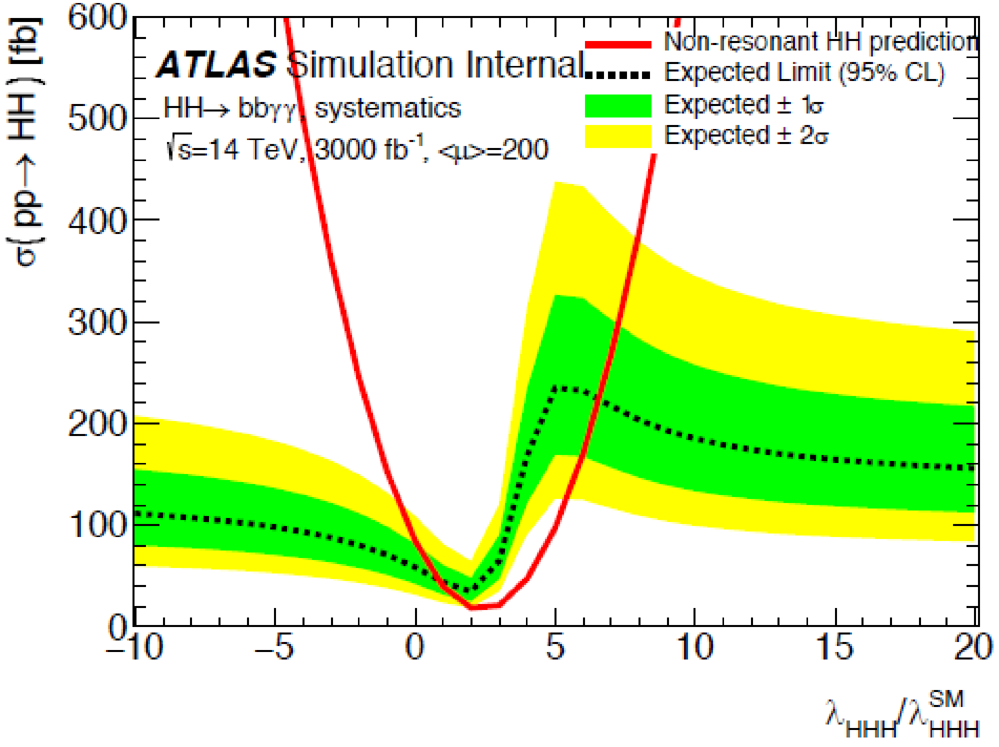

Fig.8: 95% CL upper limit on the cross-section as a function of the Higgs boson self-coupling ratio $\lambda_{hhh}/\lambda_{SM}$ with systematics.

$hh \rightarrow b\bar{b}\gamma\gamma$

- **Likelihood scan**
  - **BSM effects on the Higgs trilinear coupling** might affects single-Higgs rates at higher order corrections
  - Is not enough to look at inclusive rates if we want to check the effects on single-Higgs observables
  - To look for this effect, **we can examine the log-likelihood ratio of different values of λ_{hhh}**

$$-\ln\left[\frac{\mathcal{L}(\lambda_{hhh})}{\mathcal{L}(\lambda_{hhh}^{SM})}\right]$$

  - Likelihood (including NP) are fit to pseudo-data independently
  - **Data has SM signal injected**
  - $-1.03 < {}^{\lambda_{hhh}}/_{\lambda_{SM}} < 7.94$ (w/o sys)
  - $-1.23 < {}^{\lambda_{hhh}}/_{\lambda_{SM}} < 8.08$ (with sys)

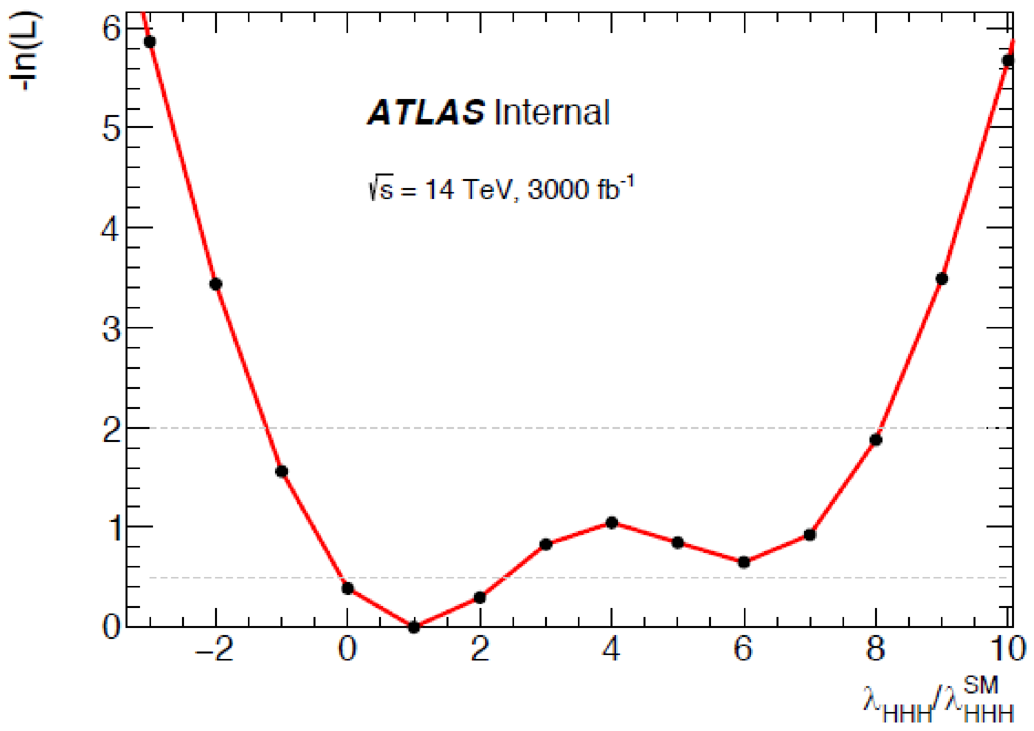

Fig.9: 95% CL upper limit on the log-likelihood ratio as a function of the Higgs boson self-coupling ratio ${}^{\lambda_{hhh}}/_{\lambda_{SM}}$ with systematics.

$hh \rightarrow b\bar{b}\tau^{+}\tau^{-}$

- Analysis is **an extrapolation of the Run-2 publication**, searching for **non-resonant Higgs boson** pair production

- **Event selection:**

  ❖ p_T cut depends on the trigger used

  - Two b-tagged jets associated with two taus or a tau + lepton (opposite electric charge)
  - Hadronic tau jets are distinguish from quark and gluon jets by using a BDT discriminant
  - Electron candidates are identified using a likelihood technique
  - Only isolated muons and electrons are considered

- **Background estimation:**

  - Dominant background processes are $t\bar{t}$, QCD multi-jet and Z boson produced in association with heavy-flavour-jets
  - SM Higgs boson production in association with a Z boson that subsequently decays into a $b\bar{b}$ final state, is an irreducible background

- **Fit BDT discriminants**

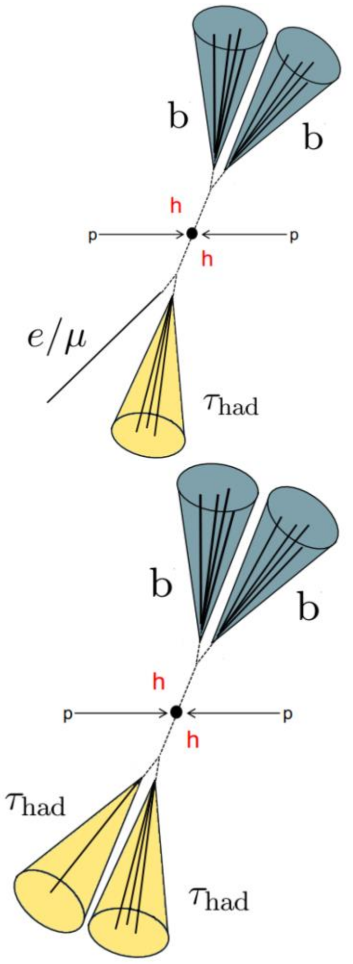

Fig.10: di-Higgs $b\bar{b}\tau^{+}\tau^{-}$ candidate in the semi-leptonic and hadronic mode

$hh \rightarrow b\bar{b}\tau^{+}\tau^{-}$

- **Projections for HL-LHC**
  - Signal and background distributions **scaled** by the integrated luminosity
  - All distributions **scaled** to 14 TeV cross sections
  - Assuming the **same** detector performance, the same trigger threshold, same identification efficiency with same rejection
  - **No degradation** from pile-up
  - **b-tagging algorithm** (mv2c10) updated to the most recent ITk layout with improved c-jet rejection is considered
  - **Finer binning** with respect to Run-2 for the BDT discriminant distribution

- **Multivariate discriminant BDT selection**
  - BDT discriminant is used to improve the separation of signal from background, using 6 variables that provided good discrimination with minimal correlation
  - BDT distributions from the three signal regions, along with control region yields, form the inputs to the final fit

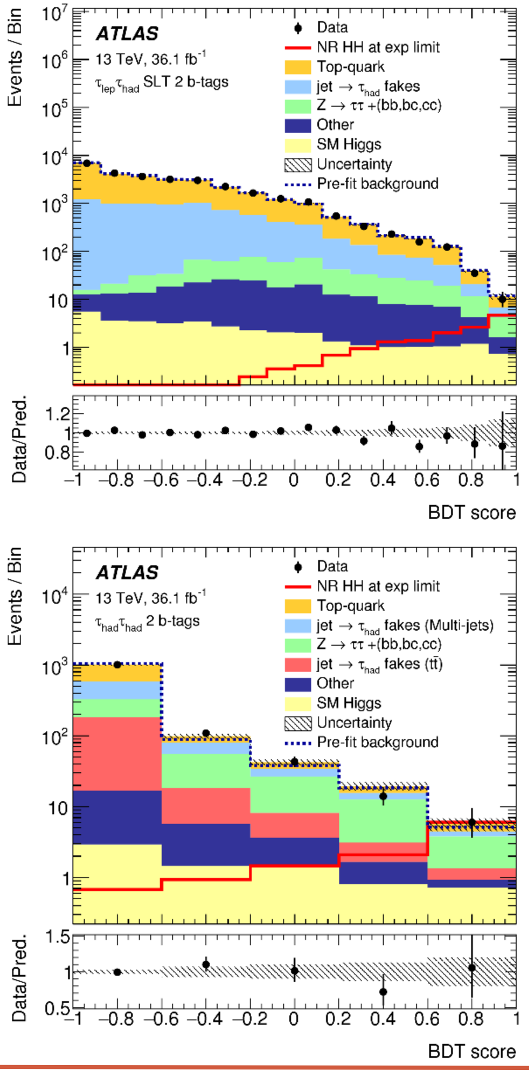

Fig.11: BDT discriminant distribution for signal and background samples, for the $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$

$hh \rightarrow b\bar{b}\tau^+\tau^-$

- **Systematic and statistical uncertainties**
  - **Normalisation fixed to the best Run-2 fit** values with corresponding uncertainties:
    - Z+Heavy-flavour, scaled up by 1.34 with normalization uncertainty of 12%
    - $t\bar{t}$ no change in normalization
    - $\tau_{had}\tau_{had}$ 37% up, 34% down; $\tau_{had}\tau_{had}$ 13%
- **Results scenarios**
  - Results considers four scenarios:
    - Full systematics
    - Full systematics, MC statistical uncertainty neglected
    - Stat-only
    - **Baseline scenario**: Z+Heavy-flavour and $t\bar{t}$ scale down by a factor of 10 (negligible effect), and $t\bar{t}H$ and VH scaled down to 10% (significant effect) and not considering MC stats. unc.

Significance:

channel	stat-only	systs	no MC stats	Baseline
$\tau_{lep}\tau_{had}$	1.11	0.19	0.65	0.67
$\tau_{had}\tau_{had}$	2.47	0.68	1.81	2.13
$b\bar{b}\tau^+\tau^-$	2.71	0.71	1.92	2.27

Table 1: Significance for each scenario.

Data statistics	0.08
Full systematic unc	0.92
all normalizations	0.07
all but normalizations	0.86
b-tagging	0.01
jets and MET	0.01
τ_{had}	0.02
fake τ_{had} estimation	0.005
Top Modeling	0.01
$Z \rightarrow \tau\tau$	0.01
SM Higgs Modeling	0.05
MC stat	0.84
Signal	0.01

Table 2: Fractional impact of nuisance parameter sets, quadratically subtracted from total uncertainty.

$hh \rightarrow b\bar{b}\tau^{+}\tau^{-}$

- **Limits**
  - Signal and background distributions are modeled by WSMaker with its corresponding upper and lower limits calculation
- **Likelihood scan**
  - **Data has SM signal injected**

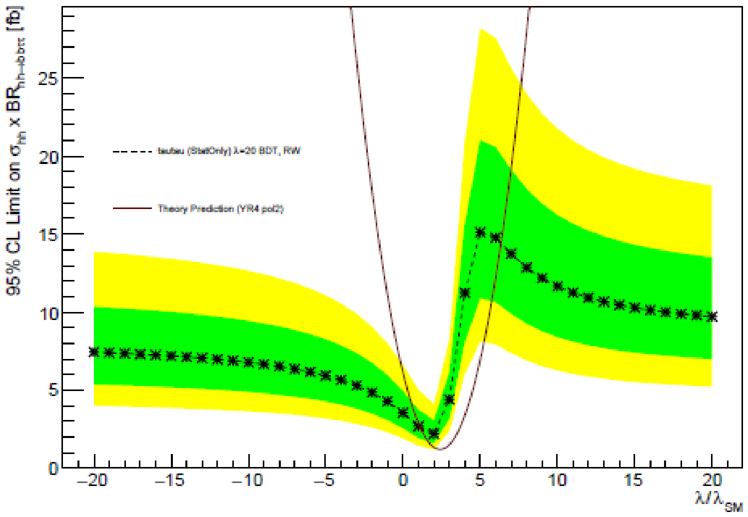

Fig.12: 95% CL upper limit on the cross-section x BR as a function of the Higgs boson self-coupling ratio ${}^{\lambda_{hhh}}/_{\lambda_{SM}}$ without systematics.

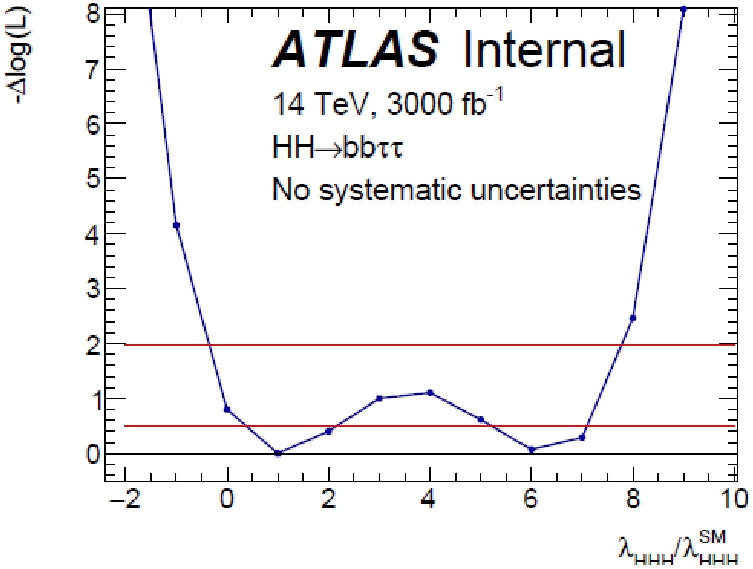

Fig.13: 95% CL upper limit on the log-likelihood ratio as a function of the Higgs boson self-coupling ratio ${}^{\lambda_{hhh}}/_{\lambda_{SM}}$ without systematics.

	significance	${}^{\lambda_{hhh}}/_{\lambda_{SM}}$
w/o syst.	2.71σ	[-0.3,7.7]
baseline	2.27σ	[-3.9,12.5]

$hh \to b\bar{b}b\bar{b}$

- Analysis is an extrapolation of the Run-2 publication, searching for non-resonant Higgs boson pair production using the "resolved" analysis method (not boosted, reconstruction of four b-jets), **same as previous PUB Note**

- **Event selection:**
  - Select four b-tagged $R = 0.4$ jets ($p_T > 40$ GeV, $|\eta| < 2.5$)
  - Pair into two Higgs boson candidates
  - Demand Higgs boson candidates satisfy requirements on p_T and $\Delta R(h, h)$
  - Veto events with hadronic top candidates
  - Insist masses of both Higgs boson candidates are consistent with m_h

- **Background estimation:**
  - Background dominated by multi-jet and $t\bar{t}$
  - Difficult to model multi-jet background using MC simulation
  - Multi-jet background modeled using data-driven methods

- **Fit m_{4j} spectrum**

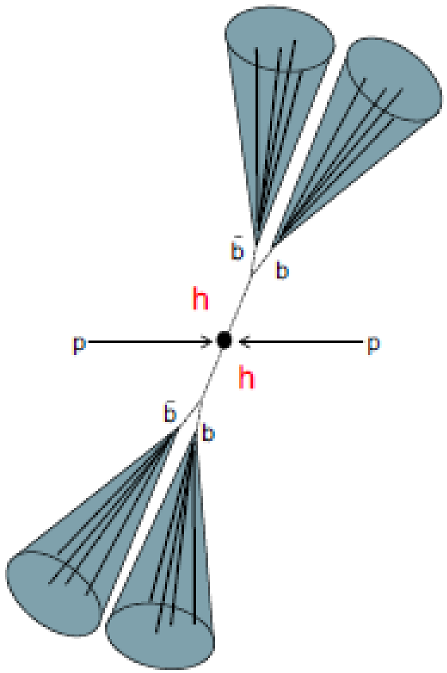

Fig.1: di-Higgs $b\bar{b}b\bar{b}$ candidate using resolved method.

$hh \to b\bar{b}b\bar{b}$

- **Projections for HL-LHC**
  - **Extrapolation approach described in last year's PUB note**
  - Forced to do extrapolation because of lack of MC-based background samples, but projection has full systematic uncertainty treatment and was validated with real data
  - Current Run-2 analysis performance will be maintained (ignores improvements from upgraded detector, **except ITk**, but neglects pile-up degradation)

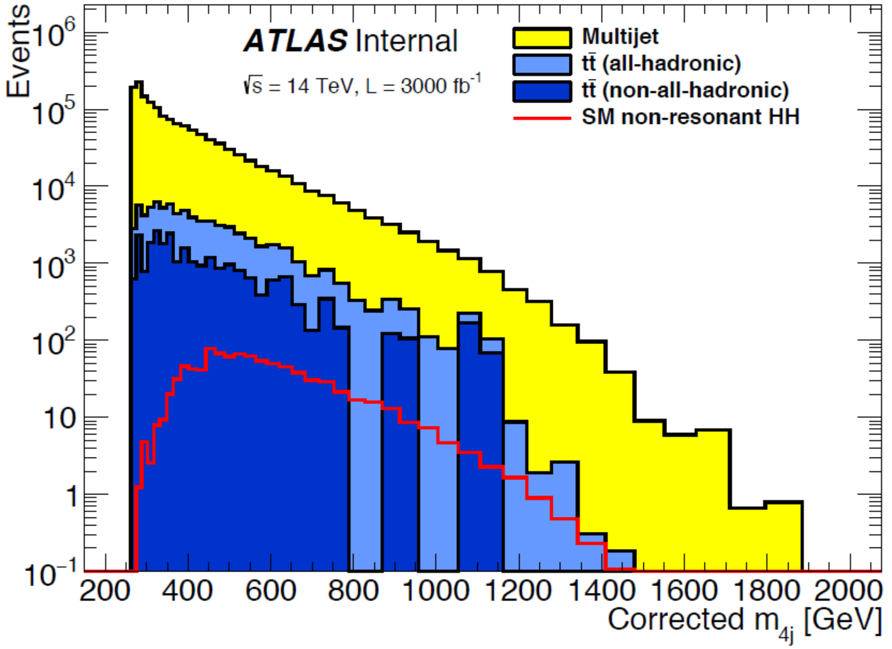

Fig.2: HL-LHC 3000 fb^{-1} extrapolated samples distribution; normalising $\sqrt{s} = 13 \to 14$ TeV and scaling by the expected ITk efficiency.

- **Statistical uncertainties**
  - Signal and $t\bar{t}$ are taken from MC simulation (independent of $\int \mathcal{L}dt$), **but multi-jet is data-driven (scales with $\int \mathcal{L}dt$)**

- **Systematic uncertainties**
  - JER, b-tagging and luminosity are not statistically limited (independent of $\int \mathcal{L}dt$)

$hh \rightarrow b\bar{b}b\bar{b}$

- **Systematic uncertainties**
  - Multi-jet and $t\bar{t}$ modelling uncertainties decrease $\propto \int \mathcal{L}dt$
  - **Systematic uncertainties lead to a worsening of sensitivity by a factor of 1.1 – 2.5**
  - Fit can constrained largest impact nuisance parameters are by ~40%
  - Detector systematic have zero impact

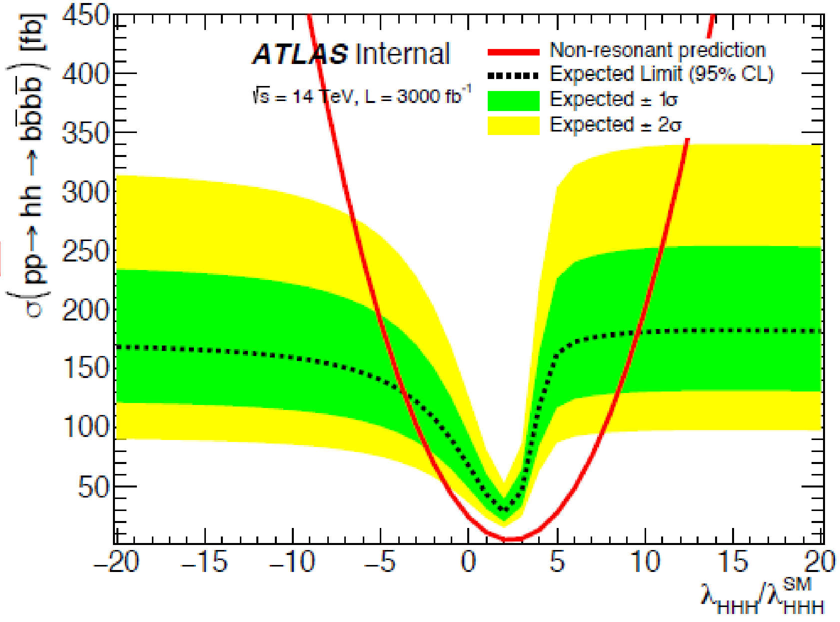

Fig.4: 95% CL upper limit on the cross-section as a function of the Higgs boson self-coupling ratio ${}^{\lambda_{hhh}}/_{\lambda_{SM}}$ including systematics.

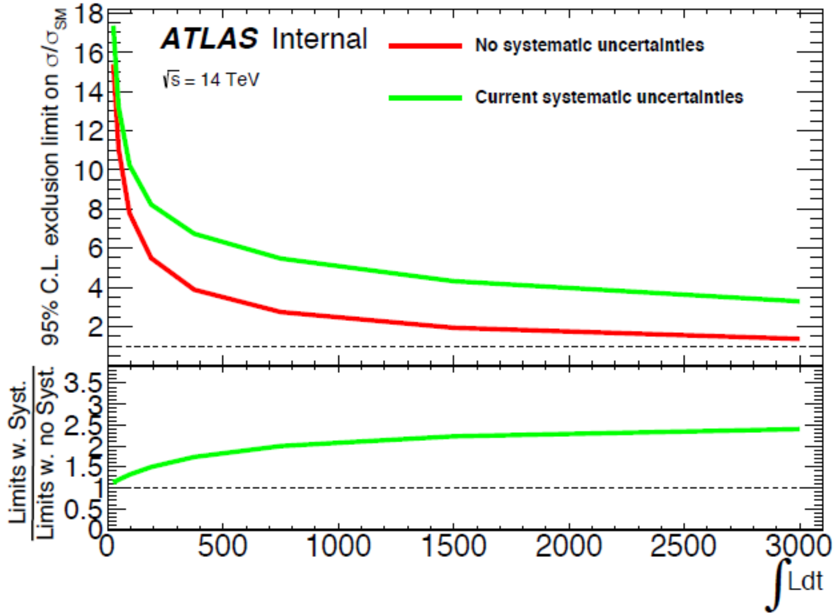

Fig.3: 95% CL upper limit on the cross-section as a function of the luminosity. Lower panel shows the ratio between the limit calculated without and with systematics.

	significance	${}^{\lambda_{hhh}}/_{\lambda_{SM}}$
w/o syst.	X	[-1.2, 8]
with syst	X	[-4.1, 8.7]

$hh \rightarrow b\bar{b}b\bar{b}$

- **HL-LHC trigger minimum jet p_T threshold**
    - High pile-up events cause difficulties in maintaining high acceptance when triggering on multi-jets
    - ATLAS HL-LHC trigger menu requires all four jets to have offline reconstruction **$p_T > 75$ GeV**
    - This **degrades the sensitivity by 30%** relative to the current analysis threshold of $p_T > 30$ GeV; equivalent to reduce the integrated luminosity of the final dataset by 1500 fb^{-1}
    - Systematic uncertainties increase by a factor of 2.5
    - $-2.6 < {}^{\lambda_{hhh}}/_{\lambda_{SM}} < 8.4$ (w/o sys)
    - $-4.8 < {}^{\lambda_{hhh}}/_{\lambda_{SM}} < 8.6$ (with sys)

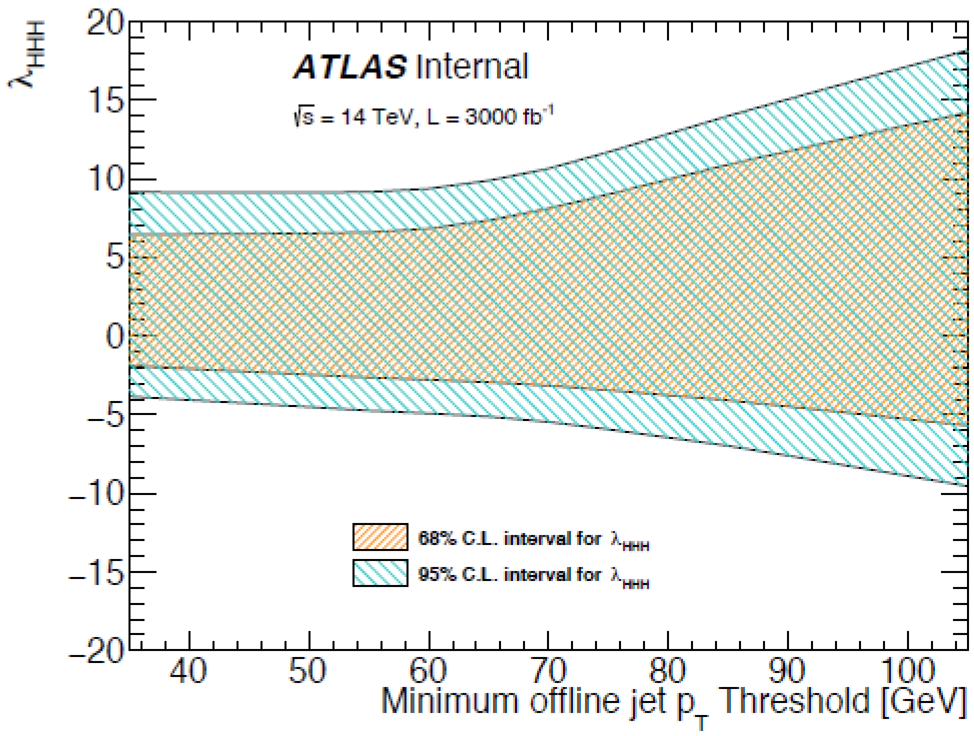

Fig.5: 95% CL upper limit on the Higgs boson self-coupling ratio as a function of the minimum offline jet p_T threshold including systematics.

- **Impact of ITk**
    - b-tagging efficiency improved by 1.375
    - Limit with systematics uncertainty improves from 3.7 → 3.3

Combination

- **Independent statistical frameworks**
  - **Different statistical frameworks** have been used for each channel, which complicates matters:
    - $hh \to b\bar{b}b\bar{b}$: HistFactory
    - $hh \to b\bar{b}\gamma\gamma$: HistFactory and Hfitter
    - $hh \to b\bar{b}\tau^+\tau^-$: WSMaker
- **Combination tool**
  - CombinationTool can handle workspaces created with different frameworks
  - There are inconsistencies in the treatment of NP (under study)
  - **Results, for the moment, without systematics**

	-2σ	-1σ	Exp.	+1σ	+2σ
bbbb	1.71	2.29	3.18	4.42	5.93
bbττ	0.553	0.743	1.03	1.52	2.08
bbγγ	0.492	0.661	0.918	1.35	1.98
bbbb+bbττ+bbγγ	0.495	0.665	0.923	1.28	1.72

Table 3: Limits on $\mu = {}^{\sigma}/_{\sigma_{SM}}$ for the combination of $hh \to b\bar{b}b\bar{b}$, $hh \to b\bar{b}\gamma\gamma$ and $hh \to b\bar{b}\tau^+\tau^-$ channels.

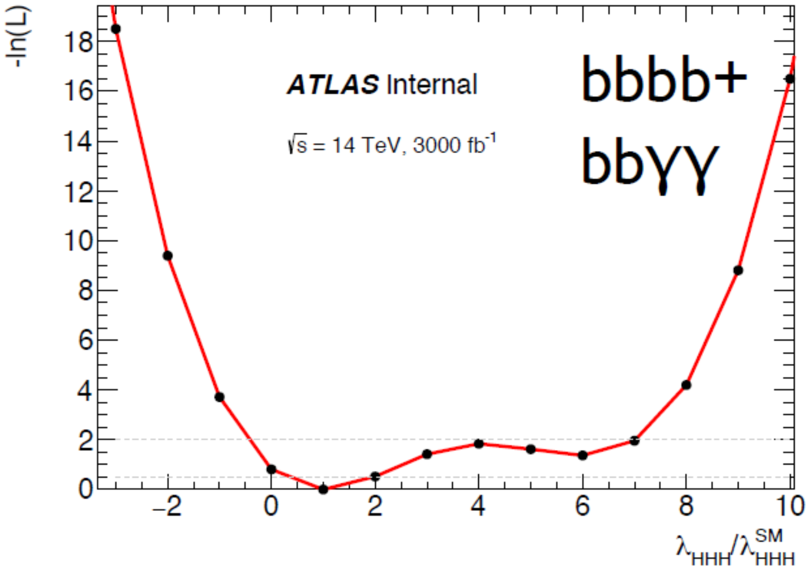

Fig.14: 95% CL upper limit on the log-likelihood ratio as a function of the Higgs boson self-coupling ratio ${}^{\lambda_{hhh}}/_{\lambda_{SM}}$ without systematics for the combination of $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}\gamma\gamma$ channels.

Summary

- An update of the study for the measurement of the non-resonant Higgs boson pair production, using the combination of the $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ channels, with the ATLAS detector at the HL-LHC was presented

- Systematics uncertainties made a considerable impact on the sensitivity obtained by the $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ analyses, while the consequence on the $b\bar{b}\gamma\gamma$ is almost negligible

- ATLAS di-Higgs prospects group is almost ready to finish the combination. If numbers agree with CMS group, then we could do a simple combination using $2 \times$ CMS or ATLAS luminosity

- A naïve combination of $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ gives a significance > 3σ with systematics

w/o systematics

channel	significance	$\lambda_{hhh}/\lambda_{SM}$
$b\bar{b}b\bar{b}$	X	[-1.2, 8.0]
$b\bar{b}\gamma\gamma$	2.34σ	[-1.03, 7.94]
$b\bar{b}\tau^+\tau^-$	2.71σ	[-0.3,7.7]
combined	X	X

with systematics

channel	significance	$\lambda_{hhh}/\lambda_{SM}$
$b\bar{b}b\bar{b}$	X	[-4.1, 8.7]
$b\bar{b}\gamma\gamma$	2.29σ	[-1.23, 8.08]
$b\bar{b}\tau^+\tau^-$	0.71σ	[-3.9,12.5]
combined	X	X

BACKUP

SM di-Higgs cross section

- Elisabeth noticed discrepancies between the cross-sections we are using for the di-Higgs production:

  - The $hh \rightarrow b\bar{b}\gamma\gamma$ channel has been using a total cross-section of 39.71 fb, which is an old value from a previous PUB Note
  - The twiki page of the LHC XS WG for the YR4 recommended a SM di-Higgs cross-section for ggF of 36.69 fb
  - The draft of the current YR recommended a total cross-section of 32.88 fb

  - Which numbers should we use?
  - Shall we use the ggF cross-section only (thinking that all analyses are using just this production mode), or the total cross-section?
  - How to parametrize the λ dependence? In the YR draft there are given cross-sections for different λ values, does it match with the morphing?

  ✓ We are going to use the YR4 cross-sections since they are NNLO