FJPPN 2024 project HEP_17

"A path toward the discovery of the Higgs-pair production in ATLAS (ATLAS HH)"

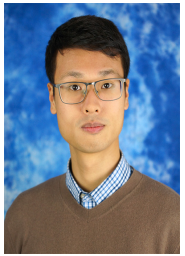
French activities

(a very brief summary)

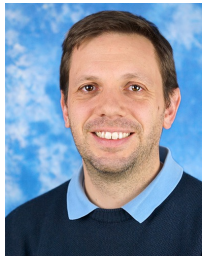
Marco Delmastro

24/2/2025

French contributors


Zhibo Wu


Oleksii Kurdysh


Nicolas Berger


Marco Delmastro

- **LAPP Annecy**
  - ✓ **Marco** Delmastro
  - ✓ **Nicolas** Berger
  - ✓ **Zhibo** Wu (post-doc) → b-tagging for HH→yybb
  - ✓ **Oleksii** (post-doc) → EFT interpretation of HH→yybb
  - ✓ *Marta Cinci (Master student, 2023-2024, left physics)*
  - ✓ *1 future PhD student → ML photon ID for HH→yybb, full Run 3 HH→yybb*
    - *from October 2025, M2 internship March-July, PhD 2025-2028, already financed by ENIGMASS+*
  - ✓ *1 future postdoc → EFT, HH combination?*
    - *from May 2025, financed by ANR EFT@LHC*
  - ✓ *Possibly a CRCN permanent hire in 2025 → if so, possibly open another HH channel, e.g. HH→bbττ*
- **LPCA Clermont** (cannot join today)
  - ✓ Djamel Boumediene → ttHH
  - ✓ Louis D'Eramo (Higgs "Photons" group co-conveners) → HH→yybb
  - ✓ Arthur Lafarge (PhD student, 2024-2027) → HH→yybb

From the FJPPN 2024 proposal, to be discussed…

The aim of this project is to build up a new France-Japan collaboration toward the HH discovery and the measurement of the Higgs self-coupling, through a series of coherent ATLAS analyses with bbγγ and bbττ final states with 400 fb^{-1} data in Run 2 and Run 3. The milestones are as follows:

- 2024: ramp-up and consolidate individual analyses; prepare harmonized analysis framework, study trigger usage, and, establish reconstructions of boosted or soft objects using novel machine learning;
- 2025: establish the methodology to measure the self-coupling constant, and make public results with a partial dataset;
- 2026: conclude the publications of individual analyses with the full dataset;
- 2027: combine the results of the ATLAS bbγγ and bbττ analyses with the bbbb one, and initiate the combination with corresponding results in CMS, to conclude on the Higgs self-coupling constraint at LHC Run 3.

Ongoing!

HH combination is moving quite slowly (not to speak of the ATLAS+CMS combination): any idea how to speed this up for full Run 3!

A support from FJPPL (TYL) in 2024 is critical for timely ramp-up of two harmonized analyses with EM and Hadronic final states, discussing together in-person to settle analysis details. Marco and Louis will visit KEK to work with Yu, Tatsuya and several early-career colleagues in Japan, and Tatsuya and Yu will visit LAPP and LPC to work with Marco, Nicolas, Djamel, Louis and other experts.

We have money for a visit to Japan, to be organized…

Today!

Toward Full Run 3 HH→yybb analysis: new Photon ID

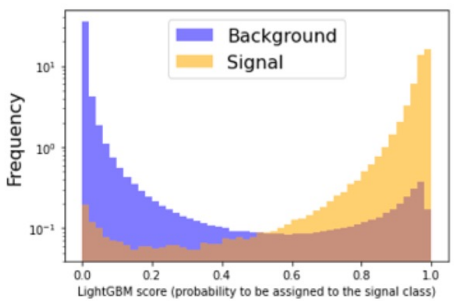
Objective: to investigate the possibility of improving the analysis sensitivity by introducing a **Boosted Decision Tree** (BDT) based Photon Identification algorithm.

Tested on the $HH \rightarrow \gamma\gamma\bar{b}b$ channel

- The BDT (from LightGBM package) was trained on **Pythia (direct) γj** and **Pythia (JFXX) jj MC samples**.

→ Used **mc16 Run2 Rel21** tuples provided by the Photon ID group for the BDT training

1. **signal** → the leading reco photon matched to *true* γ in Pythia γ-jet;
2. **background** → the leading reco photon matched to *true background* in Pythia JF.

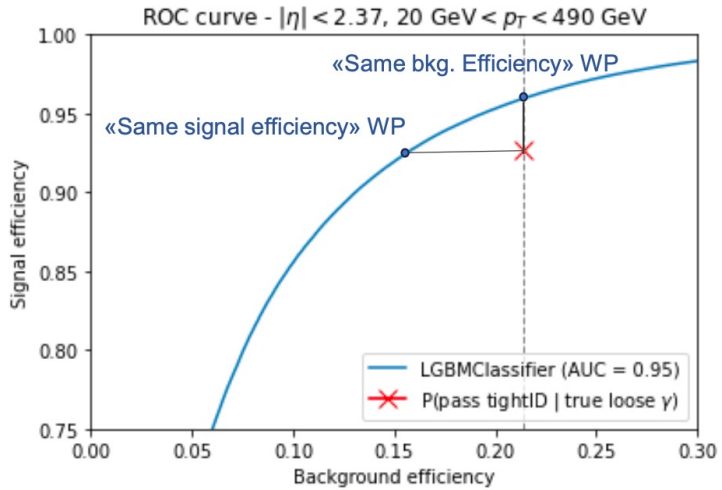


Training variables:

- **Photon kinematics:** η, pT ;
- **Shower shapes:** Rη , Rφ , ωη2, fracs1, ωη1, ωtots1 , Rhad , Rhad1 , Eratio , ΔE ;
- **Photon conversion:** convRadius, convType.

- **Reweighting:** reweighting the η and p_T distribution of the Py γ -jet signal onto the Py JF background

→ BDT does not have to learn from signal and background distributions of kinematic variables.



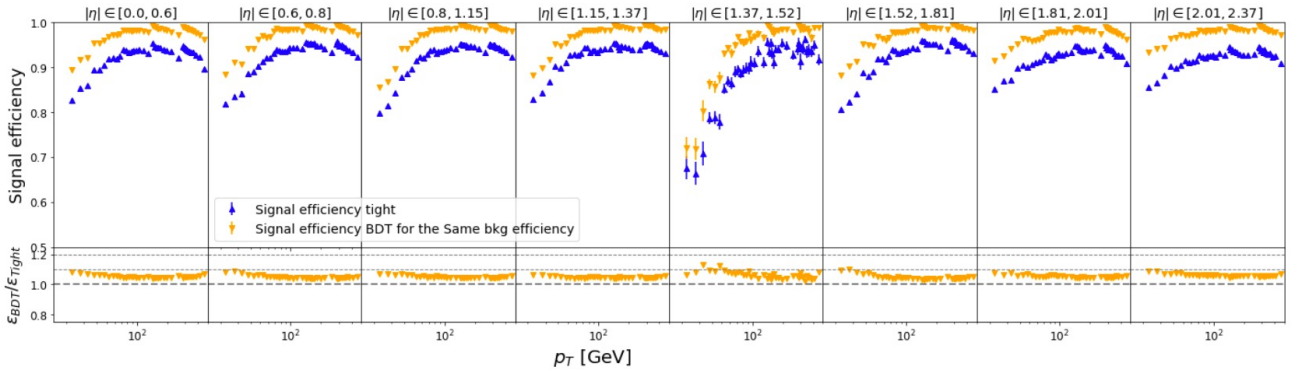
Marta Cinci

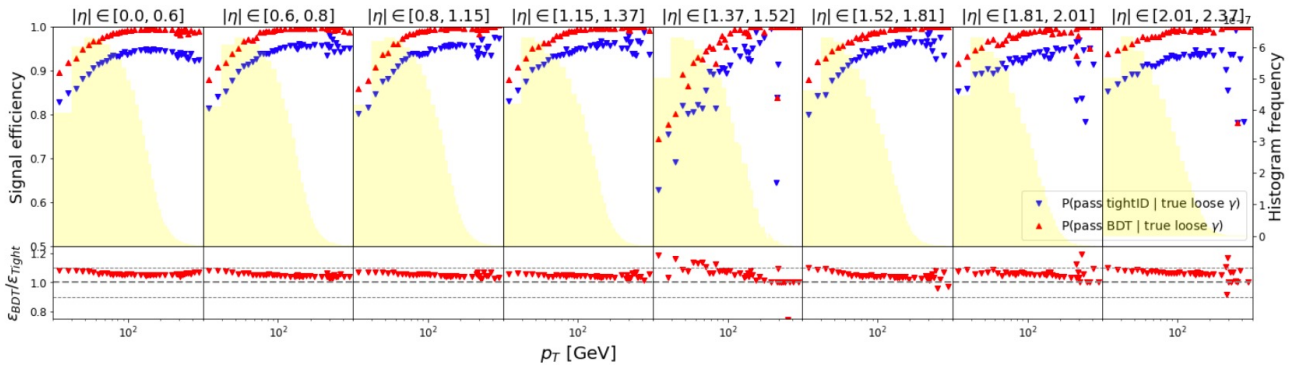
BDT-based Photon ID: same *tight* rejection, higher efficiency

WP: Cut on BDT is selected in order to obtain a **background efficiency** equal to the one of the tight selection in each (η, p_T) bin.

ϵ_{sig} = Prob(pass ID | loose ID γ matched to true γ)

Pythia γ-jet MC samples
(test events only)





ggF HH SM sample

A **5%-10% improvement** in the signal efficiency w.r.t. tight ID is observed in both samples with this new "Same bkg. efficiency" WP.

Marta Cinci

Impact of BDT-based Photon ID on HH→yybb MC samples

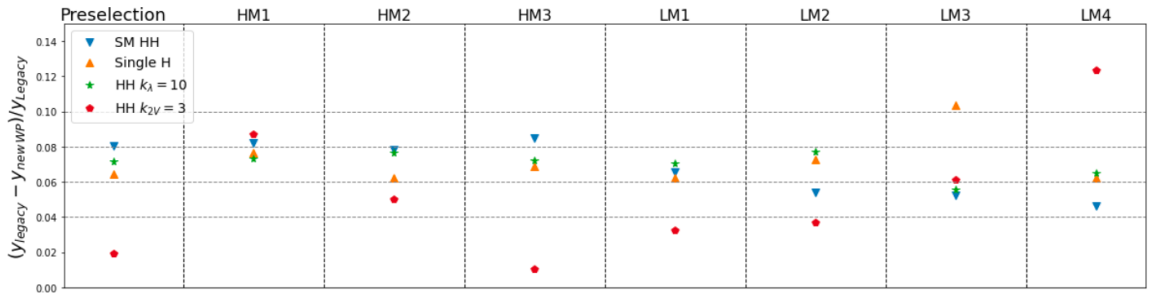
By modifying the event preselection (**BDT-based photon ID** instead of **tight ID**)

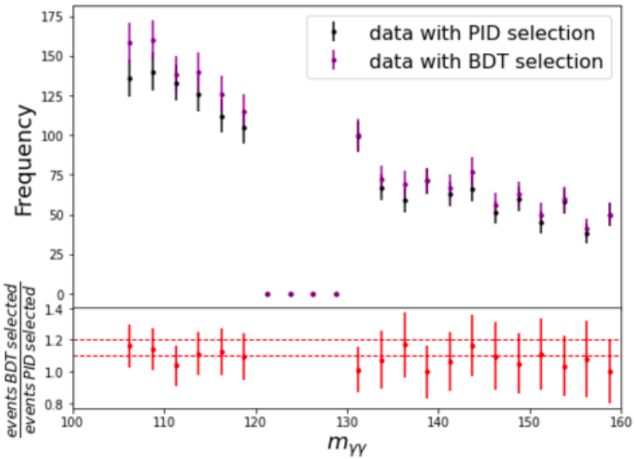
→ The **expected improvement** in SM signal yield with the new "Same bkg efficiency" w.r.t. the tight ID WP is **~ 6 – 8%**

We expect an increased signal yield, and increased bkg. yields

- $H \to \gamma\gamma$
- $\gamma\gamma$ and γj production

An increase of data events in the sidebands considering the purities in data (majority of true γ in the sidebands).





Marta Cinci

Impact of BDT-based Photon ID on HH→yybb analysis

1. We have re-extracted via the fit on $m_{\gamma\gamma}$ the expected **stat. only** upper limits **at 95% CL** on the **signal strength μ_{HH}**;

2. Via a profile log-likelihood (2Δln(L)) scan we have also derived **expected 68%** and **95% CLs** for k_λ and k_{2V} constraints

3. **Work in progress:** maximal exploitation of the BDT output:

	BDT-based photonID analysis	Nominal Legacy analysis
μ_{HH}	4.54	4.86

→ improvement of **~ 7% w.r.t. Legacy stat. only** analysis!

→ ~ 1.2% improvement on k_λ, no improvement found on k_{2V}.

- Add the photonID scores of both photons in the training
  → move away from the fixed WP and let the analysis BDT make good use of the ID scores.
- Re-optimize the categories on the new BDT to maximise the expected di-Higgs significance.

Marta Cinci