# ttH→γγ analysis: event selection and first fits

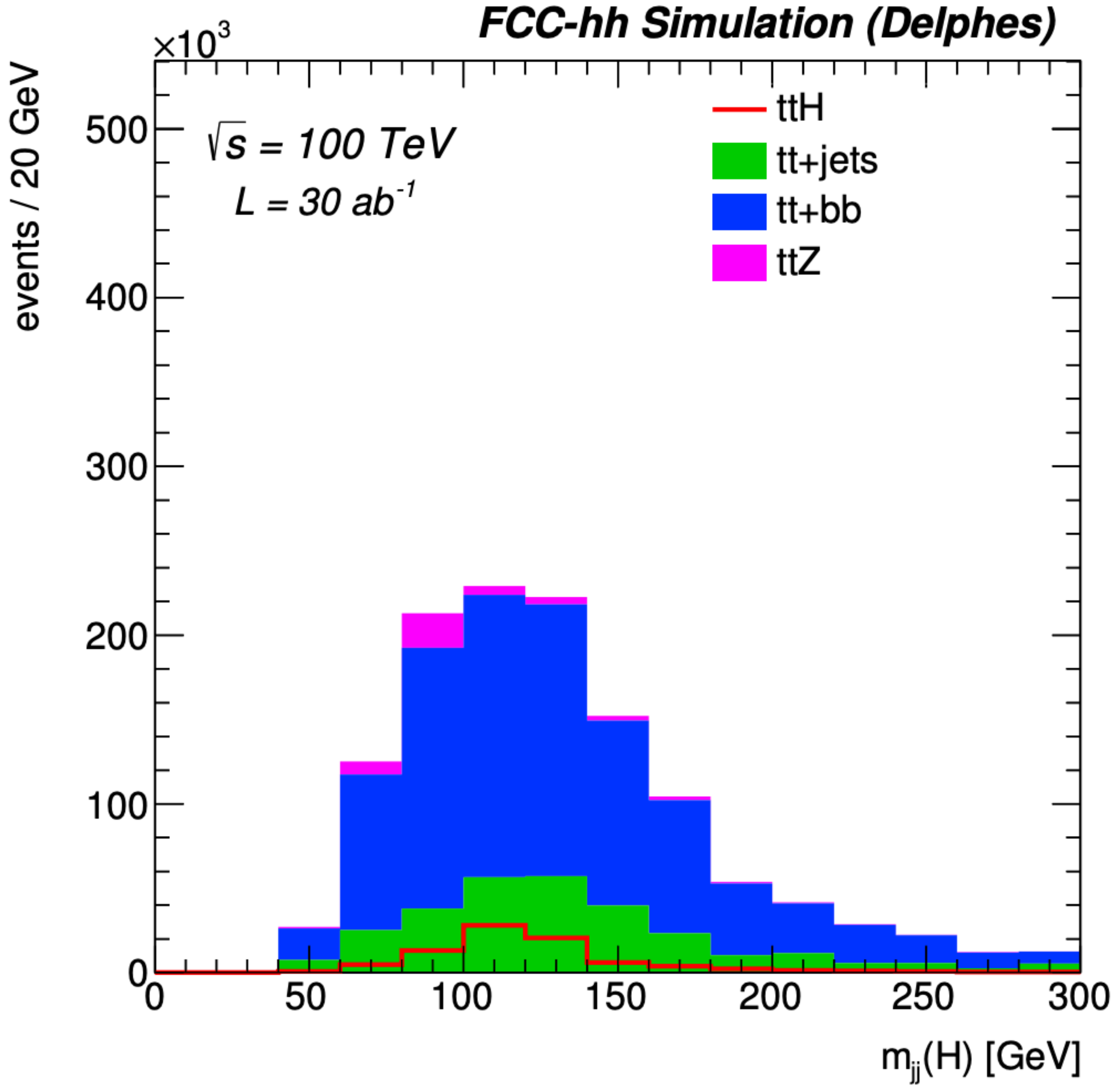
Elena Mazzeo, Stefano Manzoni, Birgit Sylvia Stapf

20 February 2025

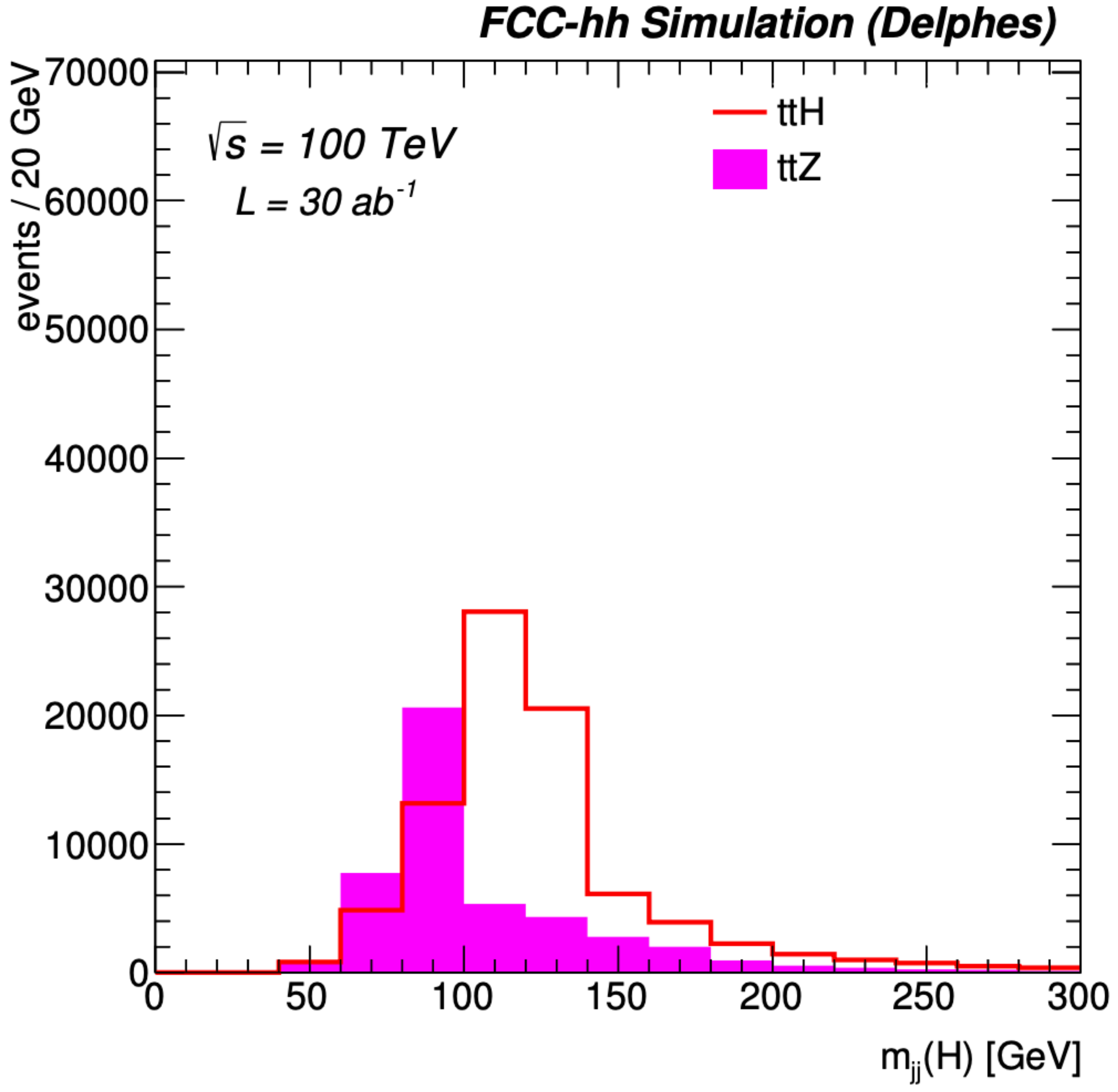
FCC-hh Physics & Performance meeting

# Outline

- The **ttH** production mode provides a **direct probe** to the **top quark Yukawa coupling** (= $y_t$).
  - **ttH→bb analysis** covered in [CERN-ACC-2018-0045] for the European strategy update of 2019.
    - **FCC-hh scenario** with $pp$ collisions at $\sqrt{s}$ = 100 TeV, $L$ = 30 ab$^{-1}$.




CERN-ACC-2018-0045




CERN-ACC-2018-0045

- Separating signal from backgrounds (= $t\bar{t}$+jets) using jet substructure observables.
- Suppress background from fake $b$-jets by requiring ≥ 4 $b$-tagged jets.
- $\mathbf{m_{jj}(H)}$ = final discriminant.
- Parameter of interest = **N(ttH) / N(ttZ)**.
- $\delta y_t / y_t \approx$ 1%.

- We would like to target the **ttH production mode @ FCC-hh** at $\sqrt{s}$ = 84 TeV, adding the **diphoton decay** of the Higgs boson (= **ttH→γγ**).
- We are considering **(semi-)leptonic top quark decays**: less statistics, but **cleaner signature** w.r.t. the hadronic channel (and much simpler handling of background samples).

# Signal and background samples

## Signal

- ttH→γγ (`mgp8_pp_tth01j_5f_haa`).
  - 306353 MC events before selections.
  - Inclusive in top quark decay.
  - Cross section(ttH) * BR(H→γγ) = 0.1018 pb, k-factor = 1.22, matching eff. = 0.613.

## Background

- ttγγ (`mgp8_pp_ttaa_semilep_5f_100TeV`).
  - 50000 MC events before selections.
  - Including semi-leptonic channel for top quark decays.
  - Cross section = 0.2852 pb, k-factor = 2, matching eff. = 1.

~100x larger w.r.t. LHC @ 13 TeV ( = 0.5071 pb * 0.00227 = 0.00115 pb).

## Some comments

- MC samples available for $\sqrt{s}$ = 100 TeV, new samples at $\sqrt{s}$ = 84 TeV will be produced.
- We do not have yet the MC samples for modeling the contribution **Vγγ + jets background**: we estimated that it amounts to **~30% of** the **ttγγ** contribution in our phase space, and, for now, we are covering this assigning an **additional k-factor** to the ttγγ sample.
  - Plan to have a **dedicated Vγγ + jets sample**!
- We could have an additional background from other processes, where (at least) one **photon** is a **fake**.
  - We estimated that it amounts to **~50% of** the **ttγγ** background, and we are accounting for it by applying an **additional 1.5 k-factor** to the ttγγ sample.
- Planning to check the background contributions from other H→γγ processes (e.g. VH).

# Event selection

- **Photon selection**
  - At least two photons with $\mathbf{p_T}$ **> 25 GeV** (see Delphes parametrisation of photon selection efficiency).
  - $\mathbf{p_T(\gamma_1)/m_{\gamma\gamma}} > 0.35$ and $\mathbf{p_T(\gamma_2)/m_{\gamma\gamma}} > 0.25$.
  - $105 < m_{\gamma\gamma} < 160$ GeV.

Dropped photon isolation requirement for now, since it caused an unexpectedly low photon efficiency.

→ To investigate.

- **b-jet selection**
  - At least two b-jets with $\mathbf{p_T}$ **> 25 GeV** passing **Medium b-tagging requirements** (see Delphes parametrisation of b-tagging efficiency).

- **Lepton selection**
  - At least one electron or muon with $\mathbf{p_T}$ **> 15 GeV** (see Delphes parametrisations of electron and muon efficiency).

- **$p_T$(H) binning**
  - We split the selected events into 5 bins in $\mathbf{p_T}(\gamma\gamma)$: [0, 60, 120, 200, 300, ∞] GeV. → Inspired by STXS 1.2 binning.

# Expected yields and efficiencies

## Expected yields

| Selection | $ttH \rightarrow \gamma\gamma$ | $tt\gamma\gamma$ |
|---|---|---|
| **All events** | 2.283668e+06 | 1.711200e+07 |
| **$\geq$ 2 photons** | 1.306010e+06 | 7.466238e+06 |
| **Rel. $p_T$ cuts** | 1.084243e+06 | 4.989717e+06 |
| **105 < $m_{\gamma\gamma}$ < 160 GeV** | 7.384012e+05 | 2.251650e+06 |
| **$\geq$ 2 b-jets** | 4.588242e+05 | 1.221150e+06 |
| **$\geq$ 1 lepton** | 1.259919e+05 | 6.430889e+05 |

## Efficiencies

| Selection | $ttH \rightarrow \gamma\gamma$ | $tt\gamma\gamma$ |
|---|---|---|
| **All events** | 1.000000 | 1.000000 |
| **$\geq$ 2 photons** | 0.571891 | 0.436316 |
| **Rel. $p_T$ cuts** | 0.474782 | 0.291592 |
| **105 < $m_{\gamma\gamma}$ < 160 GeV** | 0.323340 | 0.131583 |
| **$\geq$ 2 b-jets** | 0.200915 | 0.071362 |
| **$\geq$ 1 lepton** | 0.055171 | 0.037581 |

- As a first step, we have a look at the expected yields and the efficiency for each selection requirement.
- The single photon efficiency (~66% - 75%) is a bit lower w.r.t. the maximum photon efficiency from the Delphes card (~95%), but we also have a photon $\mathbf{p_T}$ cut.
- The efficiency for the full diphoton selection (~32%) is very similar to the efficiency obtained from the ATLAS ttH→γγ analysis!
- The single b-jet efficiency (~75%) seems consistent w.r.t. the medium b-tagging efficiency in the Delphes parametrisations.

# Expected yields and efficiencies

## Expected yields

| Selection | $ttH \to \gamma\gamma$ | $tt\gamma\gamma$ |
|---|---|---|
| **All events** | 2.283668e+06 | 1.711200e+07 |
| **Photon and $b$-jet sel., $\geq$ 1 lepton** | 1.259919e+05 | 6.430889e+05 |
| **0$\leq p_T(\gamma\gamma)$< 60 GeV** | 2.292593e+04 | 2.228106e+05 |
| **60$\leq p_T(\gamma\gamma)$< 120 GeV** | 3.508599e+04 | 2.388932e+05 |
| **120$\leq p_T(\gamma\gamma)$< 200 GeV** | 2.961359e+04 | 1.242280e+05 |
| **200$\leq p_T(\gamma\gamma)$< 300 GeV** | 4.837184e+04 | 1.629025e+05 |
| **$p_T(\gamma\gamma) \geq$ 300 GeV** | 1.960819e+04 | 1.848251e+04 |

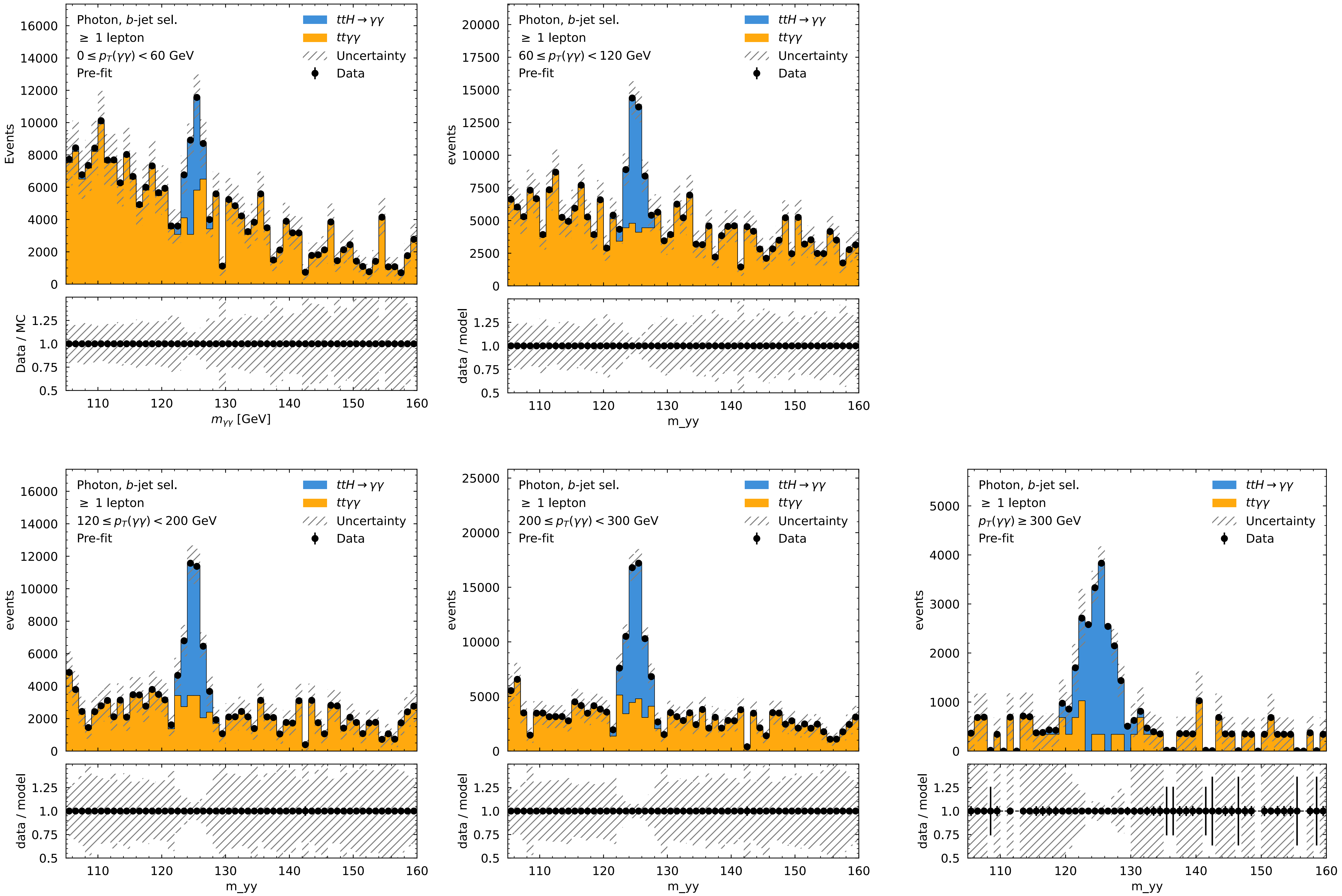
## Efficiencies

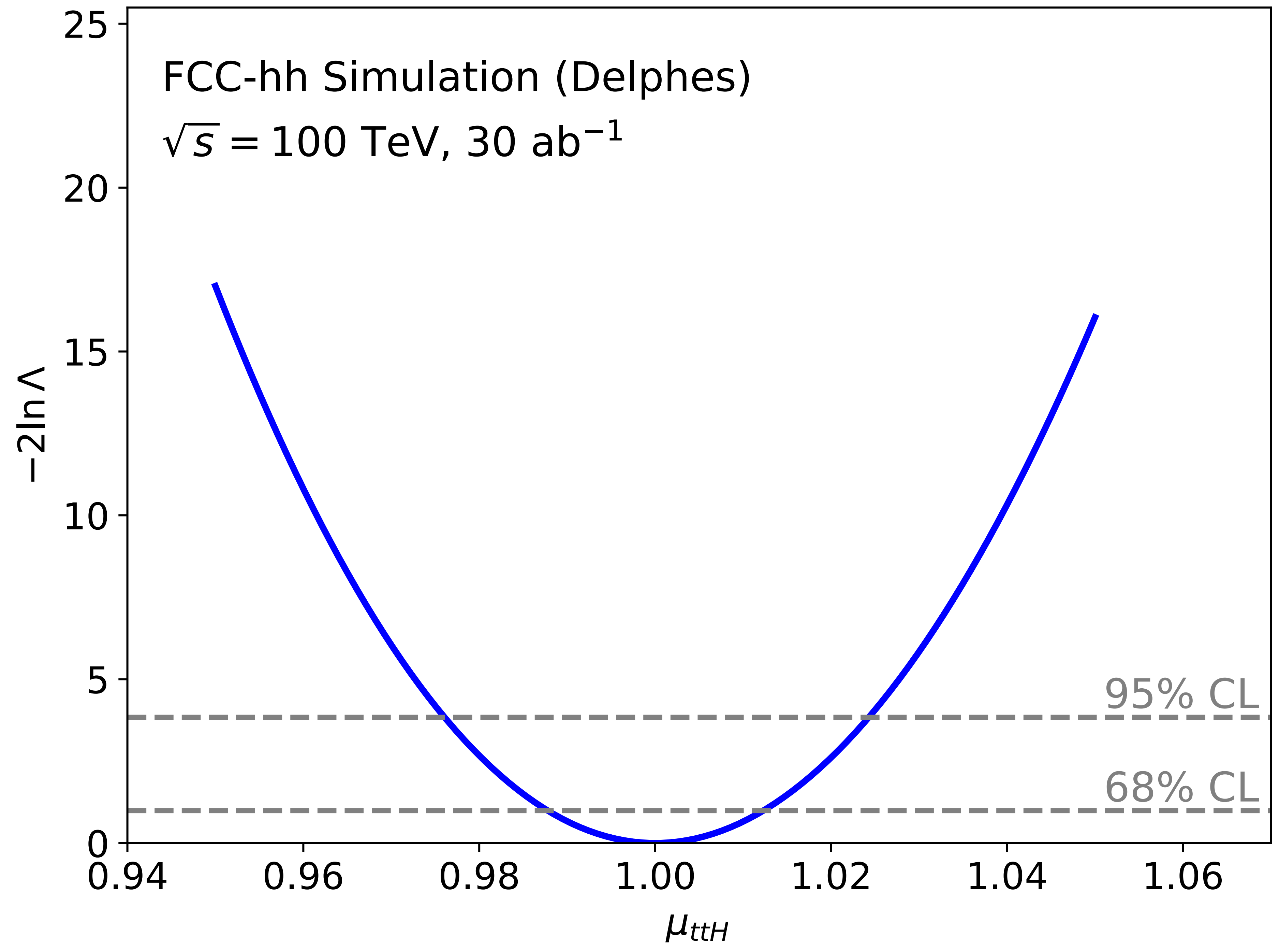
| Selection | $ttH \to \gamma\gamma$ | $tt\gamma\gamma$ |
|---|---|---|
| **All events** | 1.000000 | 1.000000 |
| **Photon and $b$-jet sel., $\geq$ 1 lepton** | 0.055171 | 0.037581 |
| **0$\leq p_T(\gamma\gamma)$< 60 GeV** | 0.010039 | 0.013021 |
| **60$\leq p_T(\gamma\gamma)$< 120 GeV** | 0.015364 | 0.013961 |
| **120$\leq p_T(\gamma\gamma)$< 200 GeV** | 0.012968 | 0.007260 |
| **200$\leq p_T(\gamma\gamma)$< 300 GeV** | 0.021182 | 0.009520 |
| **$p_T(\gamma\gamma) \geq$ 300 GeV** | 0.008586 | 0.001080 |

- The boosted $\mathbf{p_T}(\gamma\gamma)$ bins have much lower signal efficiency, but S / B ratio seems to improve (close to 1 for $\mathbf{p_T}(\gamma\gamma)$ > 300 GeV).

# Fit strategy

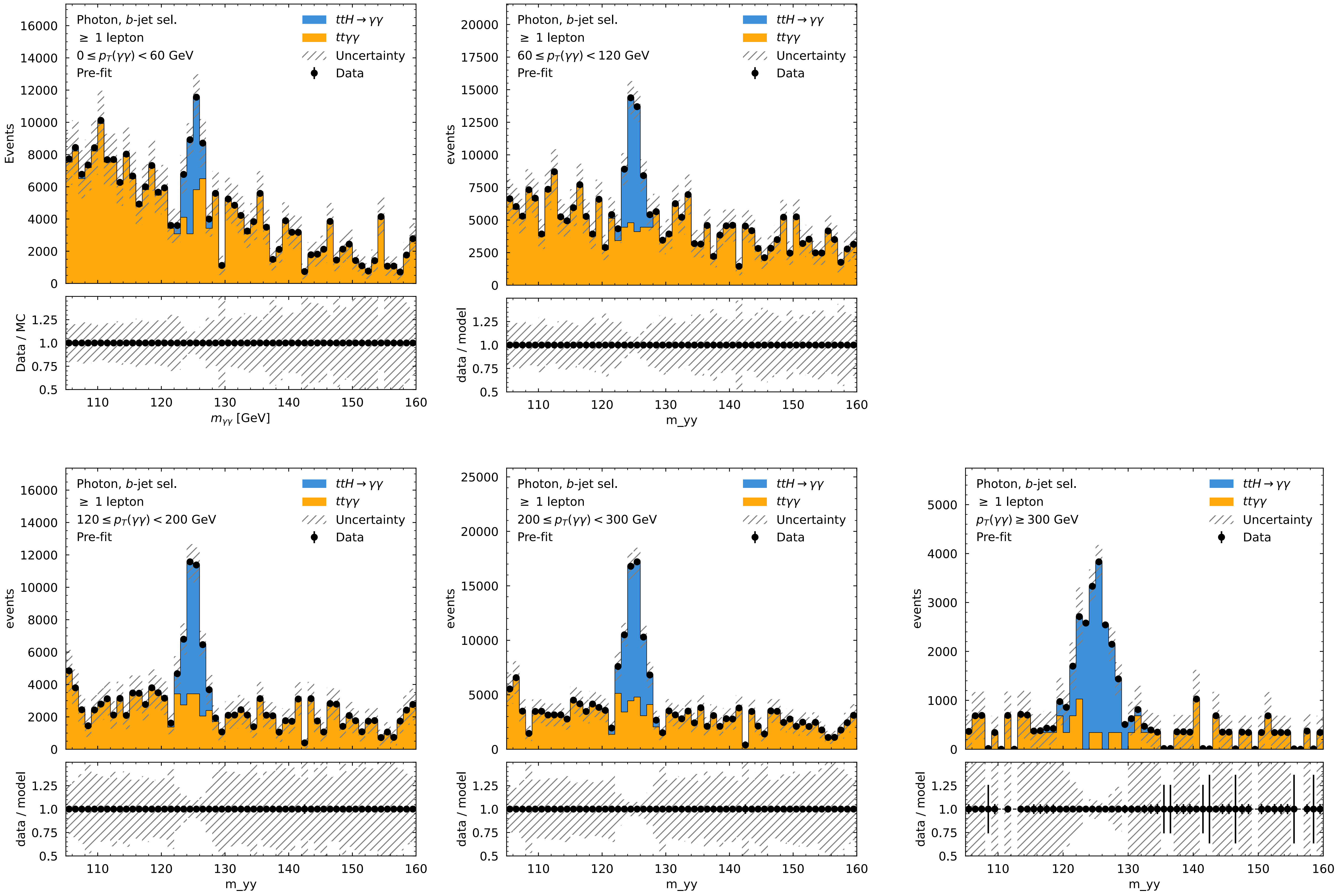
- The results are extracted via a (binned) maximum likelihood fit to the $m_{\gamma\gamma}$ distribution in the window 105 < $m_{\gamma\gamma}$ < 160 GeV, performed simultaneously across the $\mathbf{p_T}(\gamma\gamma)$ bins.
- We would like to measure the ttH signal strength (= $\mu_{\mathbf{ttH}}$) and quote an uncertainty.



# Fit strategy

- The results are extracted via a (binned) maximum likelihood fit to the $m_{\gamma\gamma}$ distribution in the window $105 < m_{\gamma\gamma} < 160$ GeV, performed simultaneously across the $\mathbf{p_T}(\gamma\gamma)$ bins.
- We would like to measure the ttH signal strength (= $\mu_{\mathbf{ttH}}$) and quote an uncertainty.




- $\mu(\mathbf{ttH})$ = 1.0000 +/- (-0.0123, +0.0124).
- Precision on $\mu(\mathbf{ttH})$ reaching almost 1.2%.

# Summary

- We would like to explore the **ttH→γγ analysis** @ **FCC-hh** at $\sqrt{s}$ = 84 TeV.

  ➡ Following the **ttH→bb analysis** in CERN-ACC-2018-0045.

- We are targeting the (semi-)leptonic channel for the top quark decay.

  ➡ Smaller signal efficiency (we loose ~50% of ttH→γγ events), but final state is cleaner and bkg. is easier to control.

- We have started to look at the available signal (**ttH→γγ)** and background (**ttγγ)** samples.

  ➡ We thought about additional sources of backgrounds (e.g. V**γγ+jets, fake photons contributions, single Higgs backgrounds**) and tried to make a plan on how to account for them.

- We have applied an event selection, targeting the **γγ+bb+leptons** final state.

- We split the selected events into mutually exclusive categories, defined using $p_T$(**γγ)**.

- We adopt $m_{\gamma\gamma}$ as the final observable, and extract the results via a binned maximum likelihood fit to the $m_{\gamma\gamma}$ distribution.

  ➡ $\mu(\mathbf{ttH})$ = 1.0000 +/- (-0.0123, +0.0124)

# Outlook

- **What's next?**
  - We would like to increase the statistics of our available signal and background samples.
    - After the event selection, the number of entries in our MC samples is quite low (~16k for ttH→γγ and ~2k for ttγγ).
    - Given that we are further splitting the selected events in categories, we would like to increase the stats. of a factor 5 for ttH→γγ (considering only the semi-leptonic top quark decays, i.e. 1.5 M events), and of a factor of 50 the ttγγ sample, reaching 2.5 M events.
    - For the background, we could increase the efficiency by reducing the width of the $m_{\gamma\gamma}$ window.
  - We would like to have a dedicated sample for the (subdominant) irreducible background from Vγγ + jets events.
  - We plan to check the contribution of backgrounds from other Higgs boson production modes.
  - See if there is room to still improve the sensitivity of this analysis (considering using a BDT to separate between signal and background).
  - Try to measure the ratio $\mu(\mathbf{ttH})/\mu(\mathbf{ttZ})$, from applying the same analysis strategy to the ttZ→ee process (on a longer timescale).

Thank you for your attention

# Backup

# MC statistics

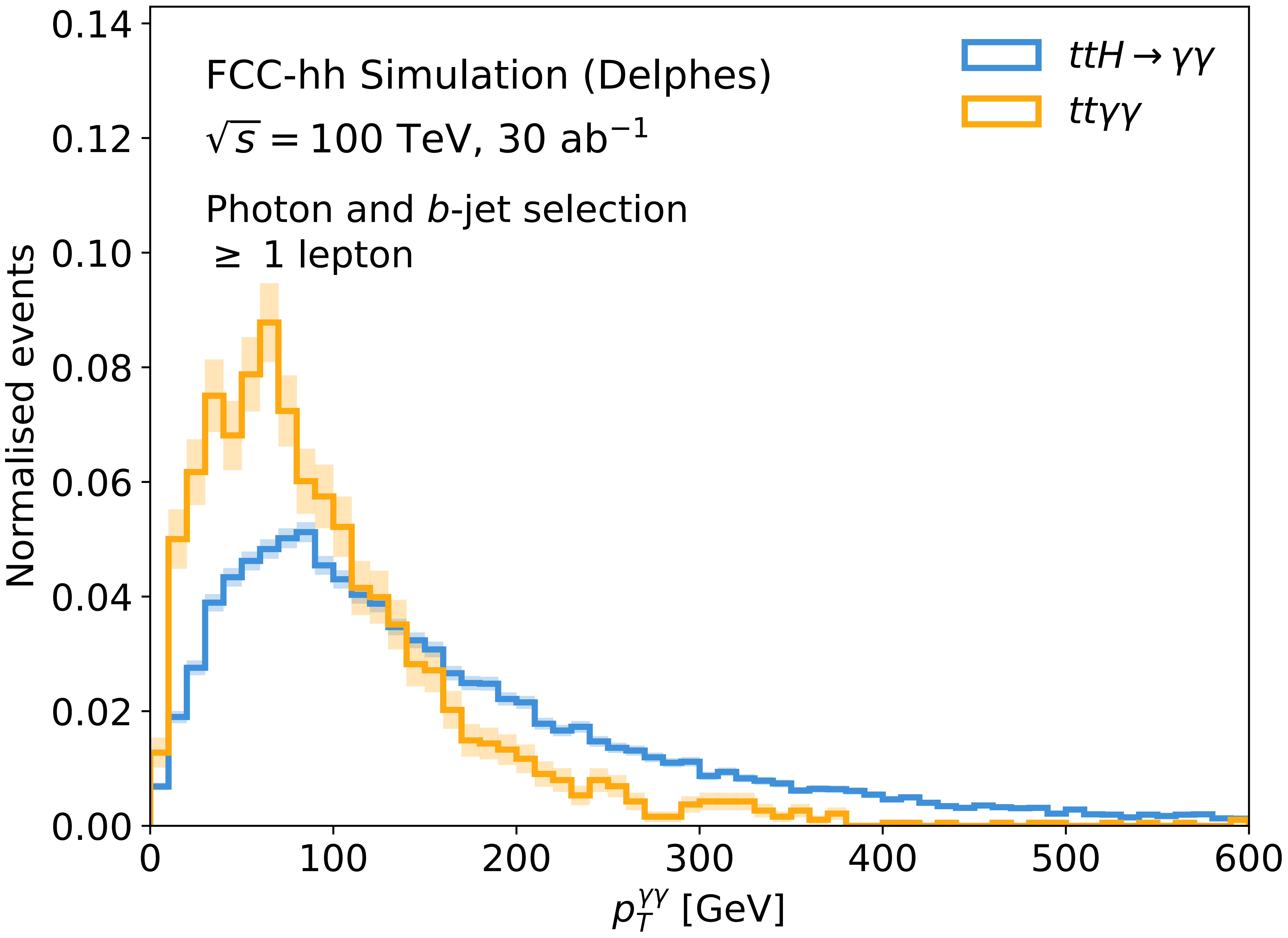
## Cutflow

| Selection | $ttH \to \gamma\gamma$ | $tt\gamma\gamma$ |
|---|---|---|
| All events | 306353 | 50000 |
| $\geq$ 2 photons | 175194 | 21815 |
| Rel. $p_T$ cuts | 145445 | 14579 |
| 105 < $m_{\gamma\gamma}$ < 160 GeV | 99050 | 6579 |
| $\geq$ 2 b-jets | 61545 | 3568 |
| $\geq$ 1 lepton | 16901 | 1879 |

## Analysis categories

| Selection | $ttH \to \gamma\gamma$ | $tt\gamma\gamma$ |
|---|---|---|
| All events | 306353 | 50000 |
| Photon and $b$-jet sel., $\geq$ 1 lepton | 16901 | 1879 |
| 0$\leq p_T(\gamma\gamma)$< 60 GeV | 3075 | 651 |
| 60$\leq p_T(\gamma\gamma)$< 120 GeV | 4706 | 698 |
| 120$\leq p_T(\gamma\gamma)$< 200 GeV | 3974 | 363 |
| 200$\leq p_T(\gamma\gamma)$< 300 GeV | 6490 | 476 |
| $p_T(\gamma\gamma) \geq$ 300 GeV | 2630 | 54 |

# $p_T(\gamma\gamma)$ distribution



# Inclusive analysis

- We tried to repeat the analysis, without splitting the selected events into $\mathbf{p_T}(\gamma\gamma)$ bins.

$\mu(\mathbf{ttH})$ = 1.0000 +/- (-0.0429, +0.0429)