Search for additional Higgs bosons at the FCC-ee

A.-M. Magnan, T. Curtis, Imperial College London
N. Nasser, CERN summer student
T. Robens, IRB Zagreb

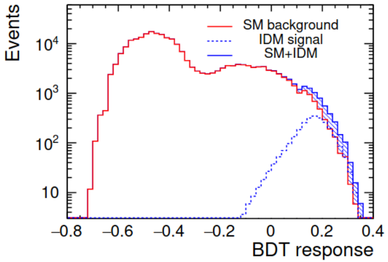
21/10/2024, FCC Physics and Performance meeting

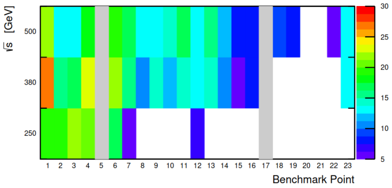
The Inert Doublet model (IDM)

- 2HDM: 5 scalars, h, H, A, H+, H-.
- h: SM Higgs with constraints from SM measurements.
- Add Z2 symmetry: dark scalars don't couple to fermions.
- Dark Matter candidate(s): lightest, choose H.
- Five free parameters: m_H, m_A, $m_{H^\pm}$, λ_{345}, λ_2.
- Can ignore λ_{345}, λ_2 for this analysis: related to quartic iDM-H couplings and coupling to SM H, subdominant processes - Plan to add study with max-allowed λ_2 for ee$\rightarrow Zh, h \rightarrow HH$ contribution.
- Similarly, can set $m_{H^\pm} = m_A + 50\,\text{GeV}$ as same-flavour dilepton channel less sensitive to $m_{H^\pm}$ production $\Rightarrow$ Down to 2 free parameters! m_H, m_A
- Plan to add study with max-allowed $M_{H^\pm}$ for small mass splitting $M_A - M_H$.
- Constraints from all experimental results: JHEP 1812 (2018) 081, set of 20 benchmark points relevant for FCC-ee $\Rightarrow$ here will use (M_H,M_A) parameter space more uniformly.

Previous studies

- IDM model studied with ILC/CLIC setup at different $\sqrt{s}$ scenarios, with HH$\mu\mu(+\nu\nu)$ final state, and also HHe$\mu\nu\nu$ and semi-leptonic final states.
- T. Robens et al, JHEP 07 (2019) 053 (@$\sqrt{s} = 380$ GeV)
- Snowmass report: arXiv:2002.11716 (@$\sqrt{s} = 250$ GeV up to 3 TeV)
- Main backgrounds: inclusive $ee \to \ell\ell$, WW, ZZ, ZH [and ttbar when kinematically available].
- Strategy based on BDT, extraction of maximum significance S/$\sqrt{S+B}$.
- 5-σ discovery possible up to m$_A + m_H = 220$ GeV with 1 ab^{-1} at $\sqrt{s}$ =250 GeV.

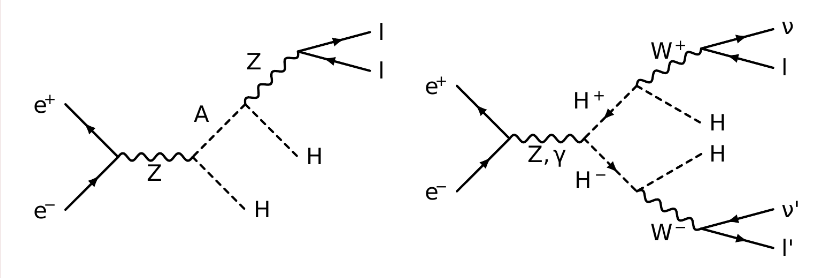




Significance for points $>$ 5-σ threshold.

Outline of this analysis

Final state considered: 2ℓ(=e or μ) + HH, mainly produced through AH and H^+H^-



1. Define parameter space for the signal MC production - reduced to 2 free parameters, m_H and m_A.
2. Set of preselection cuts to reject main background, keeping signal for most of the parameter space under study.
3. Multivariate analysis with set of input features: parametric Neural Network.
4. Fit pNN output and extract 95% CL upper limit on signal XS using CMS Combine package, and also 5-σ discovery contours.
5. Use $\mathcal{L} = 10.8$ (3) ab^{-1} of total integrated luminosity for $\sqrt{s} = 240$ (365) GeV.

Setup

Signal generation

- Generate Madgraph5_aMC@NLO cards for all benchmark points, see https://github.com/amagnan/FCC
- Generated 500k events per points, for FCC-ee @ $\sqrt{s} = 240$ and 365 GeV.
- Directly by final state: ee$\rightarrow$llHH or ll$\nu\nu$HH, with l=e, μ, τ.
- Use Pythia8 for hadronisation and τ decays.
- Finally, run through Delphes using FCC IDEA cards from official FCC-ee repository, Winter2023 production.

FCC software

- Produce events in the EDM4Hep format, root-based.
- Analyse with FCCAnalyses package, forked here: https://github.com/amagnan/FCCAnalyses
- Python-based config files to create branches with analysis variables in 2 stages,
- then python-based config files to define selection in 2 steps also: output separate histograms for each sample and selection, then plot together signal and backgrounds.

Definition of Delphes objects

Electrons and photons

- Delphes electrons, p$>$ 5 GeV.
- Delphes photons, p$>$ 5 GeV.

Muons

- Delphes muons, p$>$ 5 GeV.

Jets and MET

- Reclustered jets removing selected electrons and muons.
- Durham algo, exclusive clustering N=2, E-scheme: JetClustering::clustering_ee_kt(2, 2, 1, 0)(pseudo_jets)
- MissingET collection from Delphes.

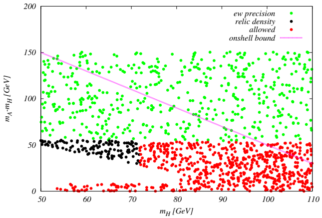
Algorithm	Objects	Selection requirements	efficiency
Tracking	e, μ, charged hadrons	p$_T$ $>$ 0.1 GeV, $\|\eta\| < 2.56$	1
Identification	γ, e, μ	E$>$ 2 GeV, $\|\eta\| < 3$	0.99

Lepton pair and recoil

- Z candidates: ReconstructedParticle::resonanceBuilder(91)(selected_leptons)
- ReconstructedParticle::recoilBuilder(240)(Zcandidates)

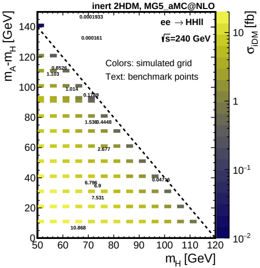
Simulated parameters

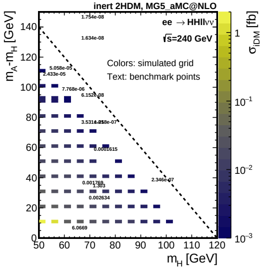
- $M_{H\pm} = M_A + 50\,\text{GeV}$ constraints also $M_A - M_H < 50\,\text{GeV} \Rightarrow$ This is not wanted, will be fixed using appropriate value for $M_{H\pm}$.



- HHll dominated by AH production: $M_{H\pm}$ choice irrelevant, for now can still use sensitivity obtained for $M_A - M_H > 50\,\text{GeV}$.

- HHll$\nu\nu$ depends on $M_{H\pm}$ choice! But stays subdominant...





Background samples

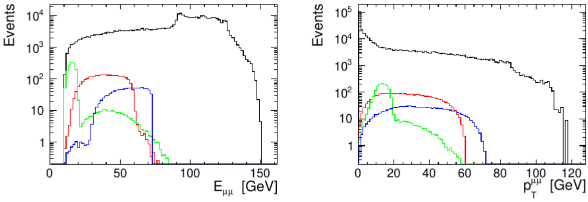
- Using FCC officially generated samples.
- Winter 2023 production.
- ee collisions at $\sqrt{s} = 240$ and 365 GeV.
- Inclusive W, Z and Higgs decays.
- ee to ee,$\mu\mu$,$\tau\tau$ production via t- and s-channel.
- Generators: Pythia8 and Whizard+Pythia6.

Ecm	Process	N Gen	xs (pb)	Eq. L (ab^{-1})
240	ZZ	56162093	1.359	41
	WW	373375386	16.4385	23
	eeH	1200000	0.00716	168
	mumuH	1200000	0.00676	178
	tautauH	1200000	0.00675	178
	nunuH	3500000	0.0462	76
	ee M30-150	85400000	8.305	10
	mumu	53400000	5.288	10
	tautau	52400000	4.668	11
365	ZZ	11470944	0.6428	18
	WW	11754213	10.7165	1.1
	eeH	1000000	0.00739	135
	mumuH	1200000	0.004185	287
	tautauH	1100000	0.004172	264
	nunuH	2200000	0.05394	41
	ee M30-150	3000000	1.527	2.0
	mumu	6600000	2.2858	2.9
	tautau	12800000	2.01656	6.3
	ttbar	2700000	0.8	3.4

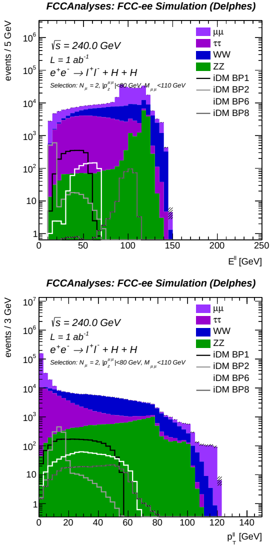
PS: contribution from Noor as CERN Summer Student: addition of the $\sqrt{s}$ =365 GeV samples.

Cross-check against CLIC Setup

$\sqrt{s} = 250$ GeV



T. Robens et al, JHEP 07 (2019) 053 (@$\sqrt{s} = 380$ GeV)
Snowmass report: arXiv:2002.11716 (@$\sqrt{s} = 250$ GeV)



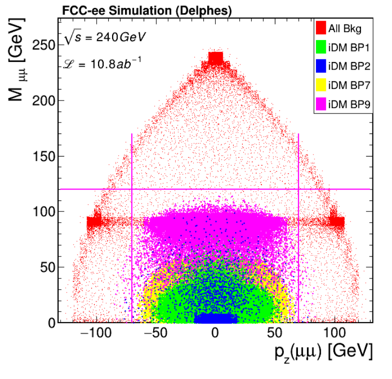
Selection

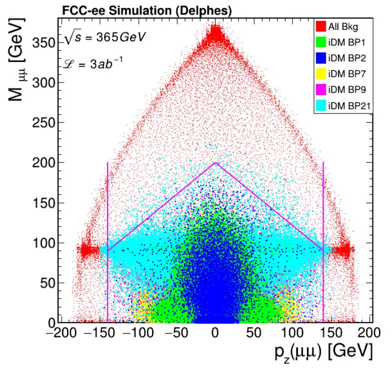
- Largely inspired from arXiv:2002.11716 with adaptation to new $\sqrt{s}$ scenarios, and addition of ee channel.
- Final state: exactly 2e (2μ) with $p > 5$ GeV [including $\tau \to e, \mu$], no other μ (e) or γ, jets.
- Some transverse missing energy to further reject inclusive ee ($\mu\mu$) production.
- Note: for ee production, missing $M_{ee} < 30$ GeV contribution ! Combining channels only for mass splitting $m_A - m_H >$30 GeV.

Step	Selection at $\sqrt{s}$ =240 GeV	Selection at $\sqrt{s}$ =365 GeV	target background
Preselection	$\lvert p_z(\ell\ell)\rvert < 70$ GeV	$\lvert p_z(\ell\ell)\rvert < 140$ GeV	ZZ, ee→ $ee, \mu\mu$
	$M_{\ell\ell} < 120$ GeV	$M_{\ell\ell} < (-9.0/14.0 \times \lvert p_z(\ell\ell)\rvert + 200)$	WW, ee→ $ee, \mu\mu$
	$E_T^{miss} > 5$ GeV		ZZ, ee→ $ee, \mu\mu$
Object veto	3rd lepton E> 5 GeV, jet, photon E> 5 GeV		WW, ZZ, ee→ $\ell\ell$
Leptons p_T	$p_T < 80, 60$ GeV	$p_T < 140, 80$ GeV	WW, ZZ, ee→ $\tau\tau$
E/p	$p_{ll}/E_{ll} > 0.1$		ee→ $\ell\ell$

Main variables for background rejection

Powerful 1st level rejection using 2-D cut

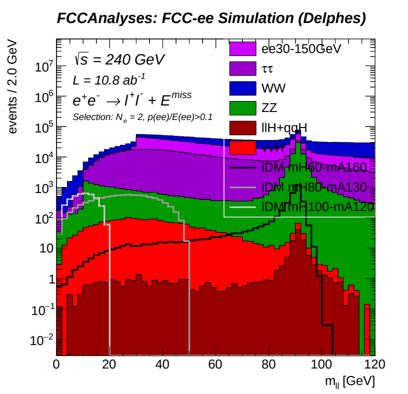




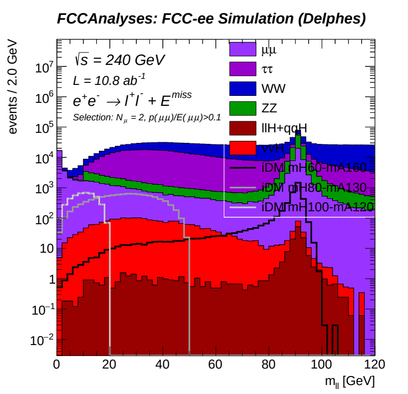
- Reproducing previous results from [arXiv:2002.11716].

Dilepton mass after selection

M_{ee} after ee selection



$M_{\mu\mu}$ after $\mu\mu$ selection



- Next step: enhance signal with machine learning.

Parametric Neural Network

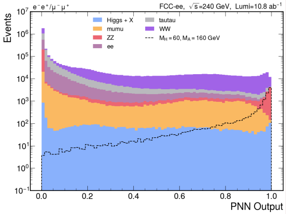
Input features

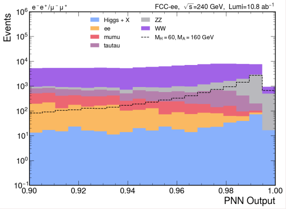
- the dilepton pair $E_{\ell\ell}$,
- the dilepton pair $p_T^{\ell\ell}$,
- the dilepton invariant mass $M_{\ell\ell}$,
- the dilepton recoil mass calculated assuming the nominal $\sqrt{s}$,
- the dilepton $p_z^{\ell\ell}$,
- the dilepton Lorentz boost $p_{\ell\ell}/E_{\ell\ell}$,
- the polar angle of the dilepton pair $\cos\theta$,
- the leptons p_T,
- the leptons $\cos(\Delta\phi)$,
- ℓ^- production angle with respect to the beam direction calculated in the dilepton centre-of-mass frame $\cos(\theta*)$,
- ℓ^- production angle with respect to the dilepton pair boost direction, calculated in the dilepton centre-of-mass frame $\cos(\theta R)$

- Single network for all parameter space
- Interpolation between points simulated to get smooth limits everywhere.
- Implemented in PyTorch.
- MC split into 3 datasets: training, validation and test with fractions 50%, 20% and 30%.
- ee and $\mu\mu$ summed together.
- Cross-checked against a simple BDT with same input features on theory benchmark point. Similar sensitivity as in arXiv:2002.11716.

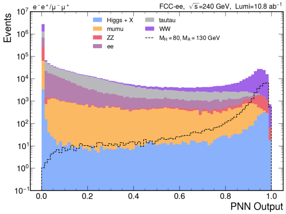
Parametric Neural Network output

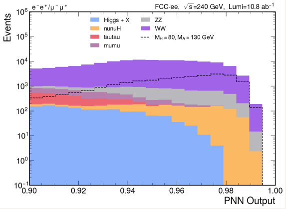
$M_H = 60$ GeV, $M_A - M_H = 100$ GeV



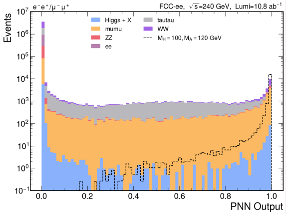


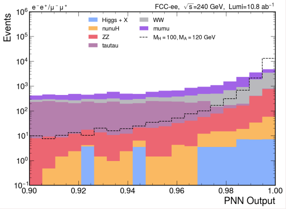
$M_H = 80$ GeV, $M_A - M_H = 50$ GeV





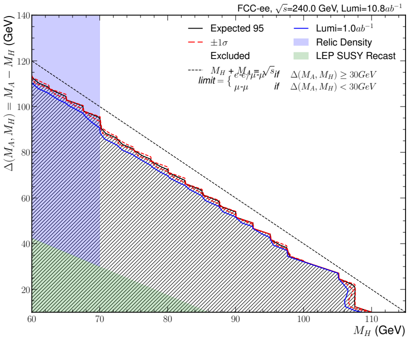
$M_H = 100$ GeV, $M_A - M_H = 20$ GeV

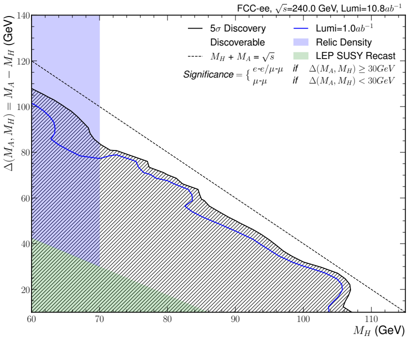




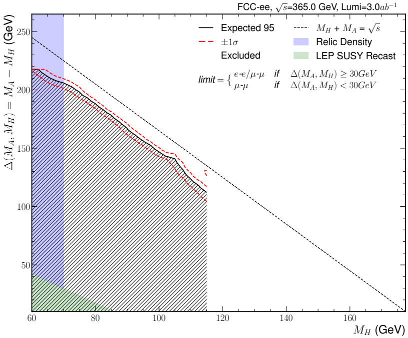
Extraction of the results

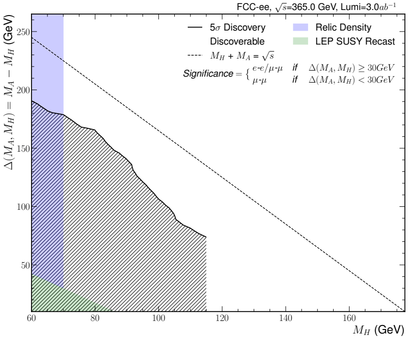
- Perform maximum likelihood fit of the pNN output above > 0.9, using for now just MC stat uncertainties as nuisance, within CMS Combine [arXiv:2404.06614].
- extract both 95% CL exclusion region, and 5-σ discovery contours, for several luminosity scenarios: nominal, half, 1/5 and 1/10.





Results at $\sqrt{s}$ =365 GeV

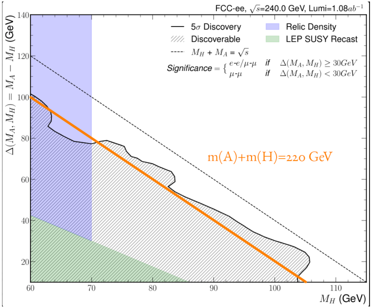


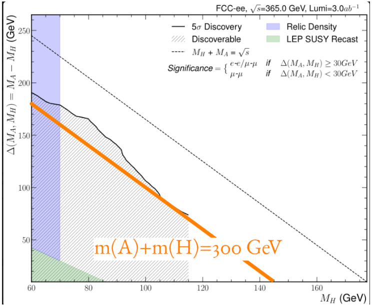


- Work-in-progress: missing simulation with higher m_H.

Conclusion

- Explored the IDM model with FCC-ee at $\sqrt{s}$ =240 (365) GeV.
- Reproduced CLIC/ILC setup results, extending a little the reach with parametric Neural Network approach with smooth limit/significance extraction.
- Will fix the model parameters to "allowed" choices but not expecting large impact on the sensitivity.
- Add a study of additionnal sensitivity from λ_2- and $M_{H\pm}$-driven diagrams.
- Next: ready to incorporate realistic systematic uncertainty scenarios !





Backups

BACKUPS

The iDM lagrangian

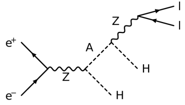
$$V = -\frac{1}{2}\left[m_{11}^2(\phi_S^\dagger\phi_S)+m_{22}^2(\phi_D^\dagger\phi_D)\right] + \frac{\lambda_1}{2}(\phi_S^\dagger\phi_S)^2+\frac{\lambda_2}{2}(\phi_D^\dagger\phi_D)^2$$
$$+\lambda_3(\phi_S^\dagger\phi_S)(\phi_D^\dagger\phi_D)+\lambda_4(\phi_S^\dagger\phi_D)(\phi_D^\dagger\phi_S) + \frac{\lambda_5}{2}\left[(\phi_S^\dagger\phi_D)^2+(\phi_D^\dagger\phi_S)^2\right]$$

Benchmark points

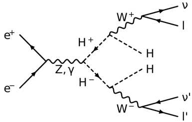
- Benchmark points: JHEP 1812 (2018) 081

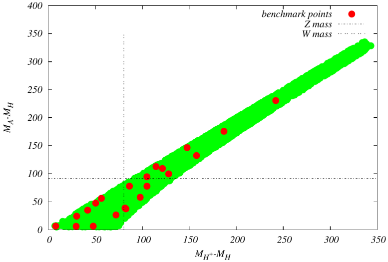
No.	M_H	M_A	$M_{H^\pm}$	$\sigma(250)$	$\sigma(380)$	$\sigma(500)$
BP1	72.77	107.803	114.639	77.2	65.9	45.7
BP2	65	71.525	112.85	155	85.1	53.4
BP3	67.07	73.222	96.73	149	83.5	52.8
BP4	73.68	100.112	145.728	89.2	69.1	46.9
BP6	72.14	109.548	154.761	75.1	65.4	45.4
BP7	76.55	134.563	174.367	31.2	52.3	40.1
BP8	70.91	148.664	175.89	20	47.5	38.1
BP9	56.78	166.22	178.24	14.1	43	36
BP10	76.69	154.579	163.045	9.44	43	36.2
BP11	98.88	155.037	155.438	-	35.6	33.2
BP12	58.31	171.148	172.96	9.01	40.4	34.8
BP13	99.65	138.484	181.321	5.17	42.5	36.2
BP14	71.03	165.604	175.971	5.13	39.6	34.7
BP15	71.03	217.656	218.738	-	18.2	24.2
BP16	71.33	203.796	229.092	-	23.3	26.9
BP18	147	194.647	197.403	-	6.14	18.7
BP19	165.8	190.082	195.999	-	3.02	16.6
BP20	191.8	198.376	199.721	-	-	11.3
BP21	57.475	288.031	299.536	-	2.66	12.6
BP22	71.42	247.224	258.382	-	8.94	18.6
BP23	62.69	162.397	190.822	13.2	43.3	36.2

ee→llHH



ee→ll$\nu\nu$HH





Note: my numbering goes from 1 to 20, corresponding to BP 1-4,6-10,12-16,18-23.