Long-lived Particles Search with Deep Learning at Lepton Collider

Yulei Zhang[1], Cen Mo[1], Xiang Chen[1], Bingzhi Li[3], Hongyang Chen[3], Jifeng Hu[2], Liang Li[1]

[1] Shanghai Jiao Tong University

[2] South China Normal University

[3] Zhejiang Lab

7 April 2024
2024 European Edition of the International WorkShop on the CEPC
Marseille, France

Paper link [2401.05094]

New Physics Beyond SM- LLPs

Long-lived particles (LLPs) are important ways to new physics

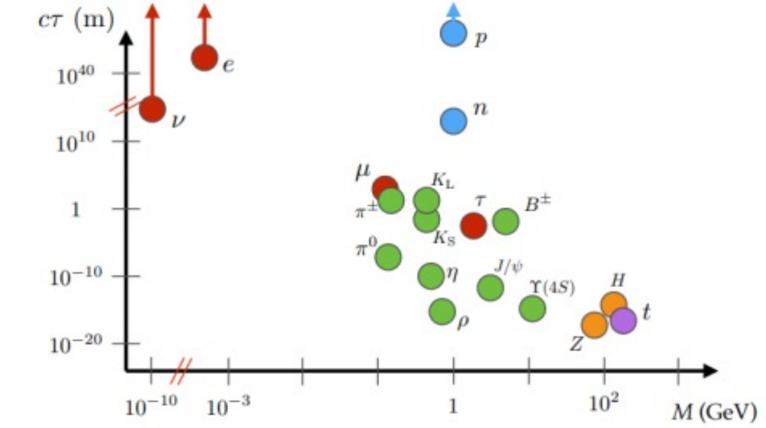
- Many particles in BSM models have a relatively long lifetime: weak coupling to SM particles, maybe new scalars, dark photons, ALP, SUSY….

LLP topology, a strong signature for detection:

- **Displaced vertex** with a long distance from the main vertex
- Different performance for neutral particles: a burst of energy appearing of nowhere and far away from the collision point

Potential on Lepton Collider

- The advantage of the lepton collider: clean environment
- Making use of **deep learning techniques**: Image recognition and pattern identification



JOURNAL OF PHYSICS G-NUCLEAR AND PARTICLE PHYSICS. 47(9) (2020)

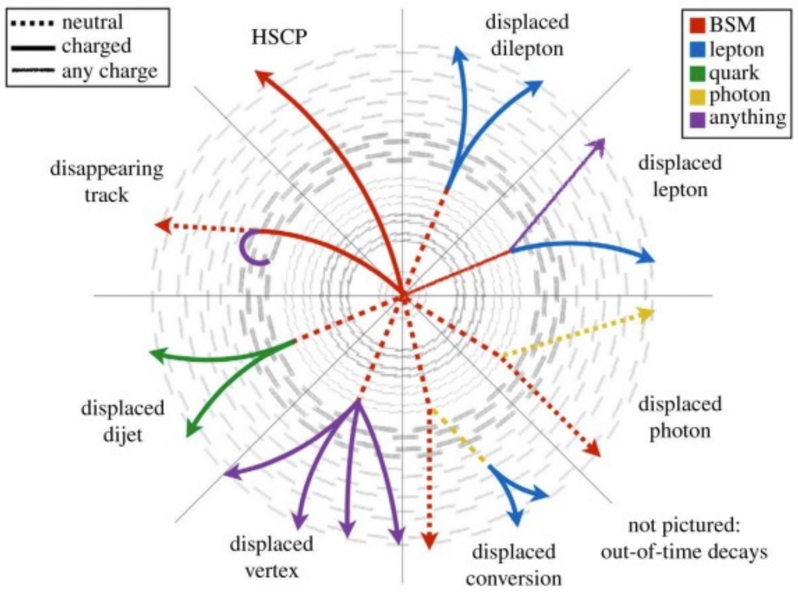
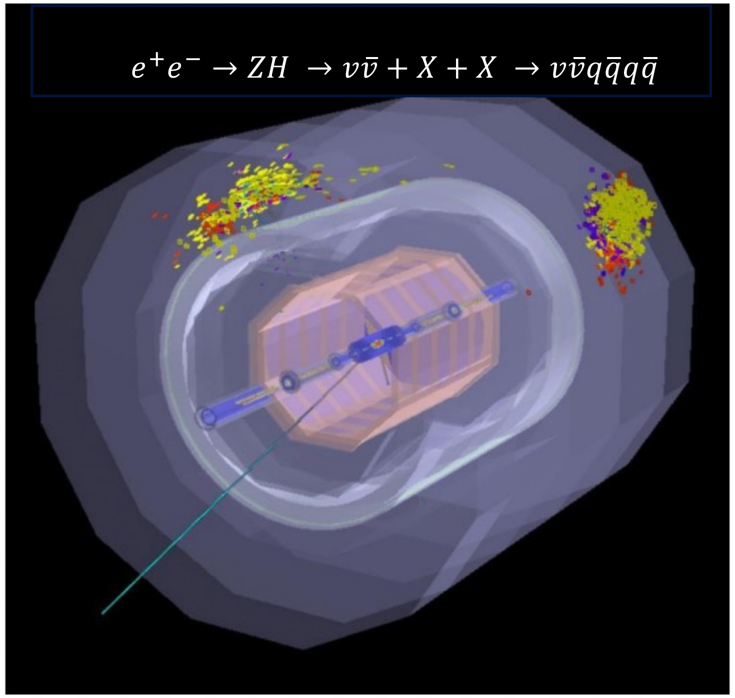


Figure from [A. De Roeck, Phil. Trans. Roy. Soc. Lond. A 377, 20190047 (2019)]

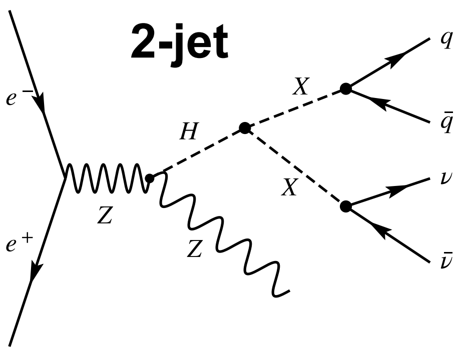
LLPs at CEPC

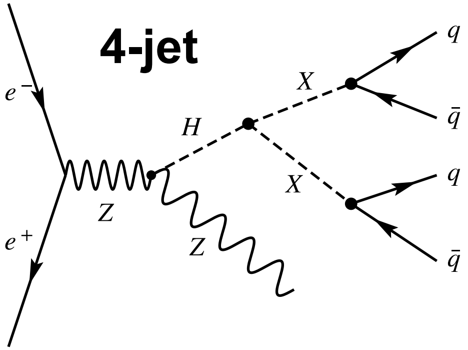
We consider two scenarios for LLP decay at CEPC: **2-jet final state** and **4-jet final state**

Acceptance (%)	Lifetime [ns]				
Mass [GeV]	0.001	0.1	1	10	100
1	100.00 ± 0.00	99.86 ± 0.01	48.76 ± 0.18	6.49 ± 0.09	0.67 ± 0.03
10	100.00 ± 0.00	100.00 ± 0.00	99.78 ± 0.01	46.80 ± 0.16	6.22 ± 0.08
50	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	99.31 ± 0.03	40.37 ± 0.16

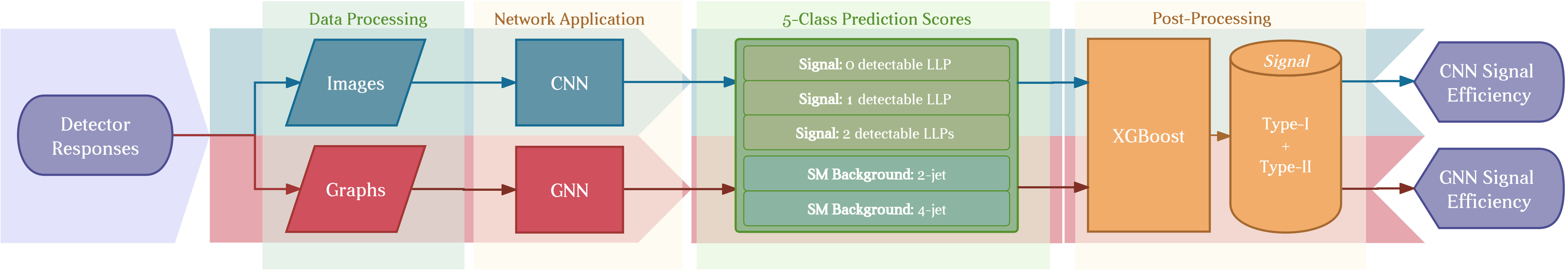


- We use the **full simulation** sample using CEPC official software (v4) to an integrated luminosity of 20 ab−1
- The LLP signal sample is generated by Madgraph5 and showered by Pythia8
- The decay vertex of LLPs: $0 \leq r_{\text{decay}} \leq 6\ [\text{m}]$





General Analysis Strategy



- In this study, we use **advanced neural networks trained with low-level detector information**:
  - No need for vertex reconstruction and object reconstruction
  - The input information from the detectors is all calibrated and considering detector resolution
  - Universal treatment for all decay channels

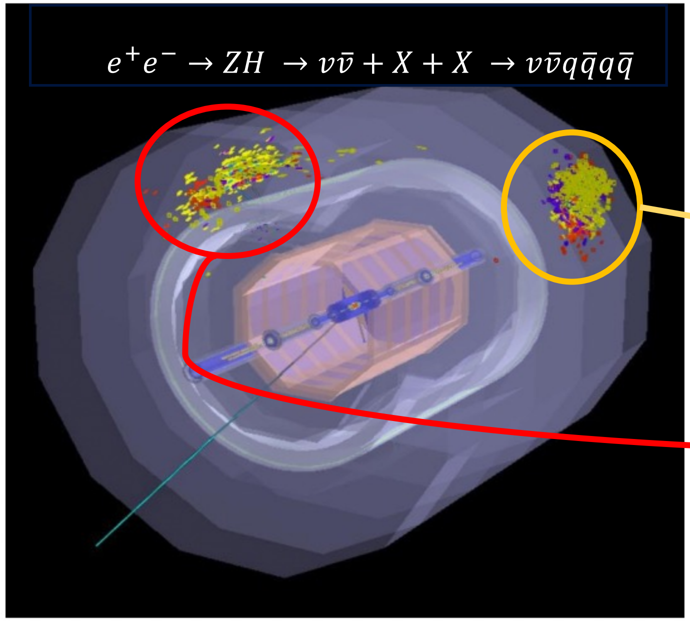
Process	LLP signal	SM ZH	$q\bar{q}$	ZZ	WW
σ[fb]	-	203.66	54106.86	1110.37	16721.77
Statistics	$\sim$ 15M	$\sim$ 1M	$\sim$ 10M	$\sim$ 1.1M	$\sim$ 8.9M

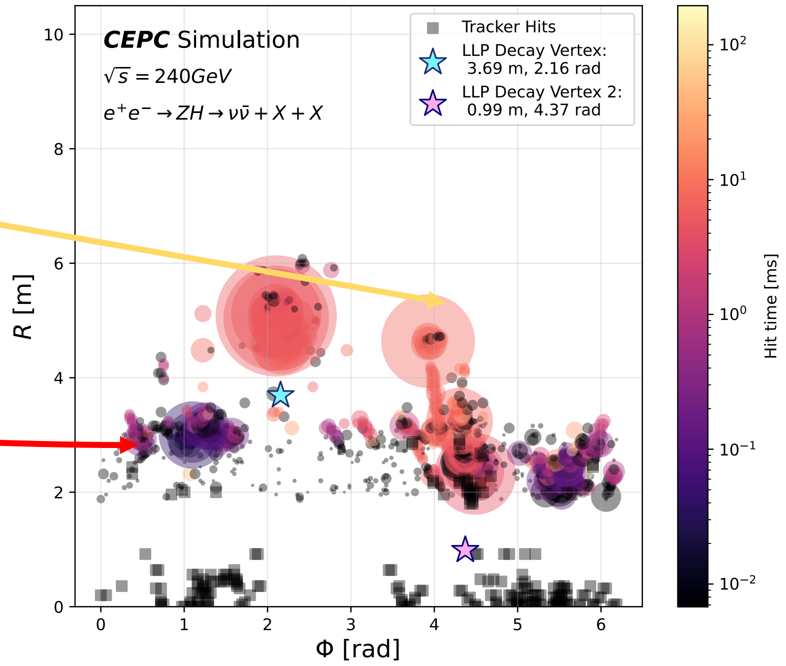
- LLPs are classified into **3 categories** based on the number of detectable LLPs
- Background samples: eeqq (2 fermions) and W/Z process (four fermions)

Image Conversion: CNN

Converting the detector information to 2D image

- The size is 200×200 according to R and phi in polar angle
- 2 channels: time and energy

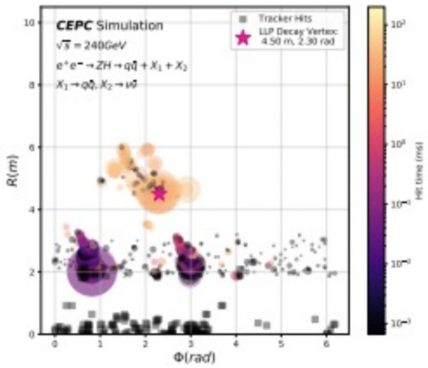


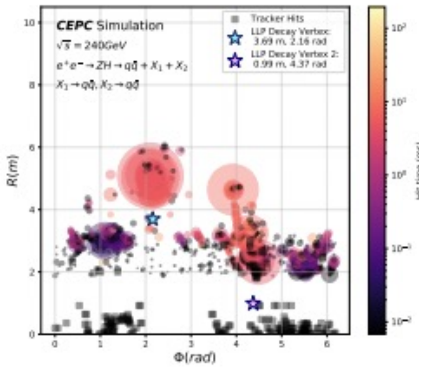


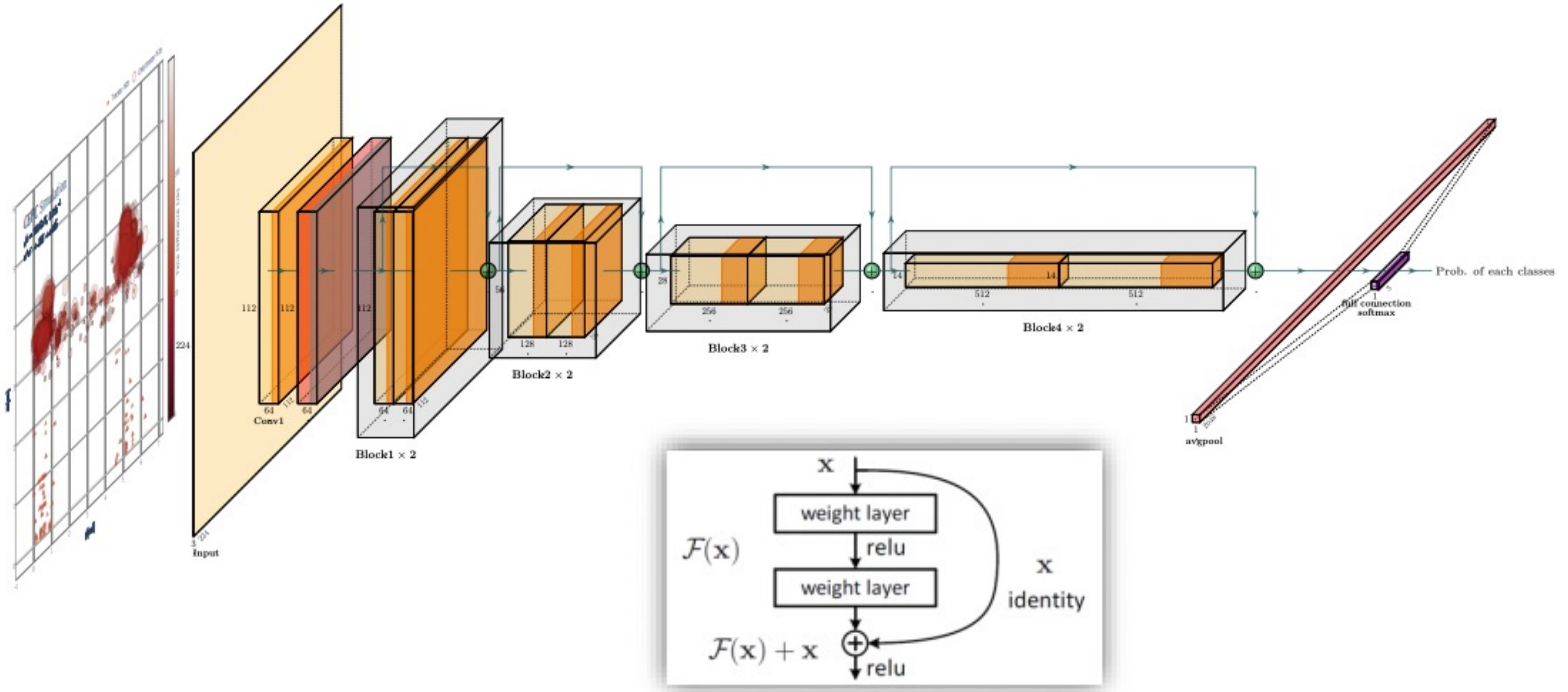
- Hits energy: bigger circle represents hits with larger energy
- Hits time: darker circles represent hits with smaller time differences $\Delta t = t_{\mathrm{hit,maxE}} - r_{\mathrm{hit,maxE}}/c$
- Include both calorimeter and tracker hits

CNN: Network Setup

Using ResNet18 model with the cross - entropy loss







- **Cross Entropy Loss:** $loss = -[\omega_0 * y_0 \log(x_0) + \omega_1 * y_1 \log(x_1) + \omega_2 * y_2 \log(x_2)]$

Class 0: 2-fermion bkg	Class 1: 4-fermion bkg	Class 2: LLP Signal
$\omega_0 = 0.5$	$\omega_1 = 0.25$	$\omega_2 = 0.25$

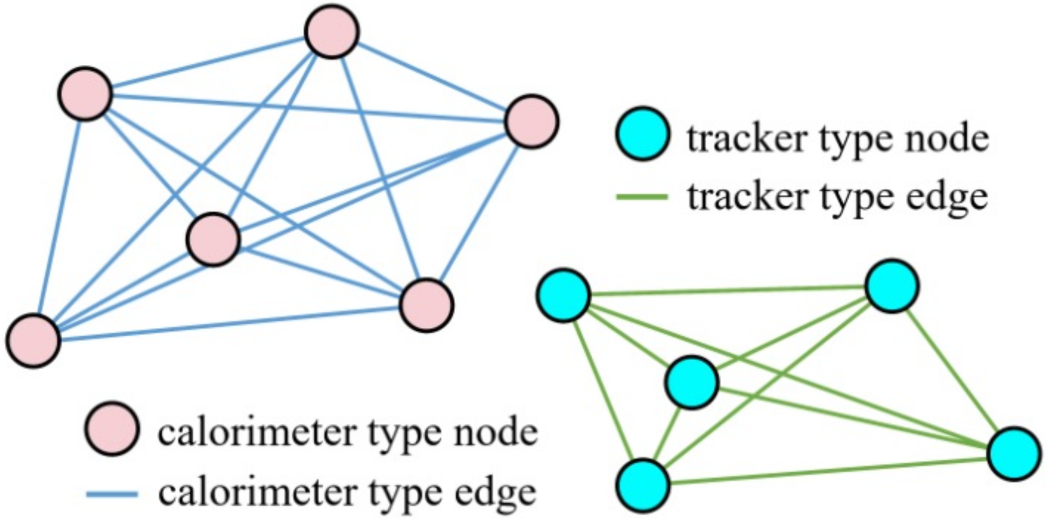
Graph Neutral Networks (GNN)

We change the representation of the information in the calorimeter and tracker to point-cloud dataset

- A clustering process is introduced to reduce graph complexity and extract the main information
- Nodes of the same detector type are interconnected comprehensively

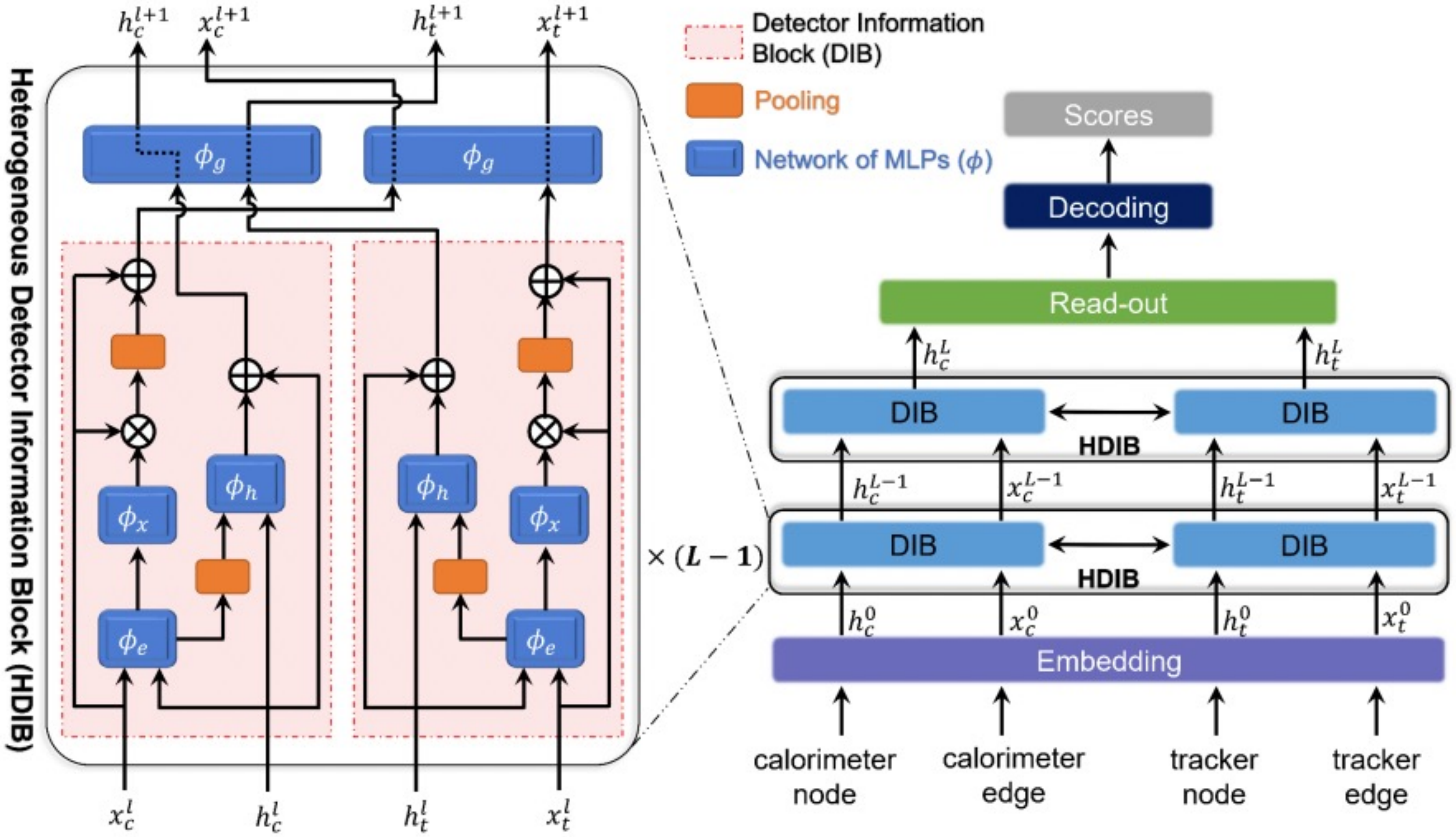
Features	Variable	Definition
calorimeter type node i	$\lvert x_i^\mu \rvert$	the space-time interval
	$\lvert p_i^\mu \rvert$	the invariant mass
	N_i	the number of hits
	η_i	$\frac{1}{2}\ln\frac{1+\frac{p_z}{p}}{1-\frac{p_z}{p}}$
	ϕ_i	$\arctan\frac{p_y}{p_x}$
	$\mathcal{R}_i$	$\sqrt{\eta^2+\phi^2}$
calorimeter type edge between node i and j	$x_i^\mu x_{j\mu},\ p_i^\mu p_{j\mu},\ x_i^\mu p_{j\mu},\ p_i^\mu x_{j\mu}$ $\lvert x_i^\mu - x_j^\mu \rvert,\ \lvert p_i^\mu - p_j^\mu \rvert,\ \eta_i-\eta_j,\ \phi_i-\phi_j,\ \mathcal{R}_i-\mathcal{R}_j$	
tracker type node i	$\lvert r \rvert$	euclidean distance
	N_i	the number of hits
	η_i	$\frac{1}{2}\ln\frac{1+\frac{z}{r}}{1-\frac{z}{r}}$
	ϕ_i	$\arctan\frac{y}{x}$
	$\mathcal{R}_i$	$\sqrt{\eta^2+\phi^2}$
tracker type edge between node i and j	$\lvert r_i - r_j \rvert,\ r_i r_j,\ \eta_i-\eta_j,\ \phi_i-\phi_j,\ \mathcal{R}_i-\mathcal{R}_j$	

- Features of nodes: calorimeter-type and tracker-type.
- Features of edges: interaction between neighbor nodes.



Heterogeneous GNN

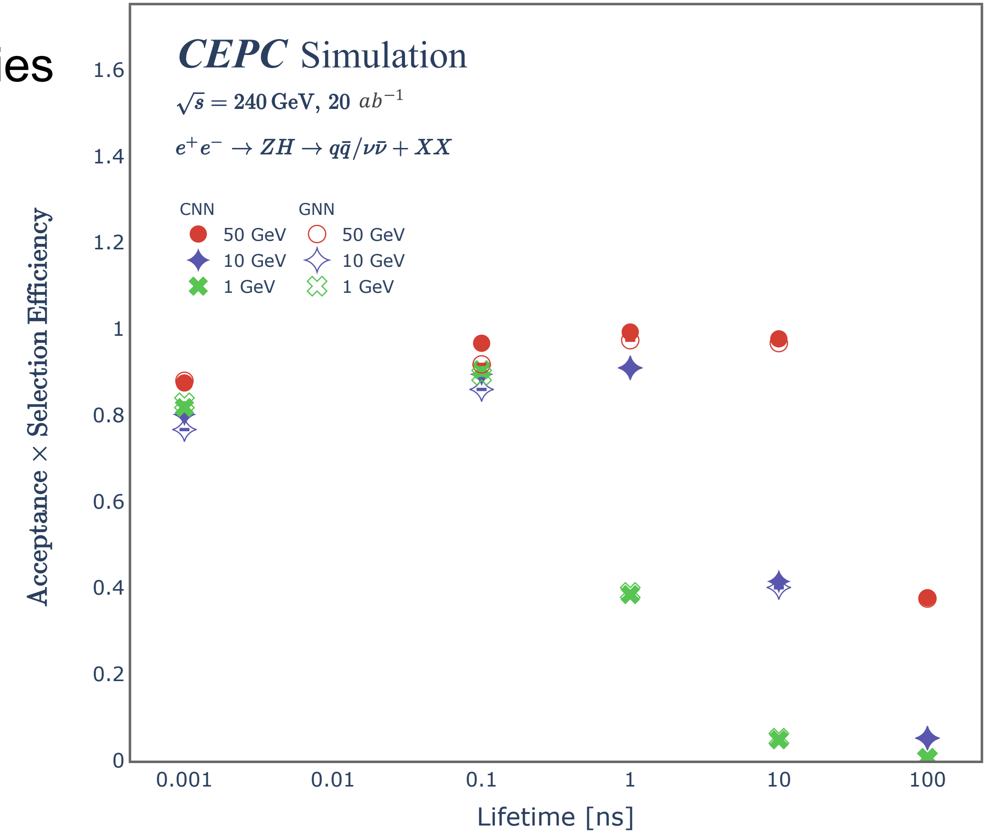
We use a GNN-based heterogeneous architecture for the graph inputs



ML- Bases Analysis Result

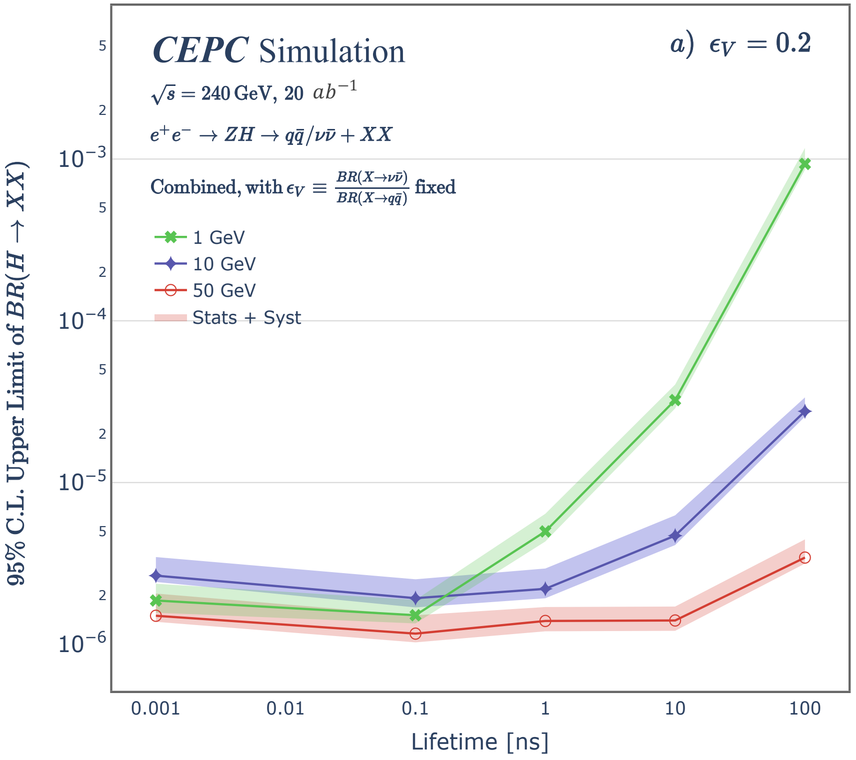
- Both CNN and GNN achieve high signal efficiencies with background-free
- The performance is consistent across different LLP mass and lifetime considerations.
- Systematics uncertainties of 1.7% / 2.3%:
  - Lumi uncertainties and neutral network variable uncertainties
  - Pile-up and cosmic rays is negligible

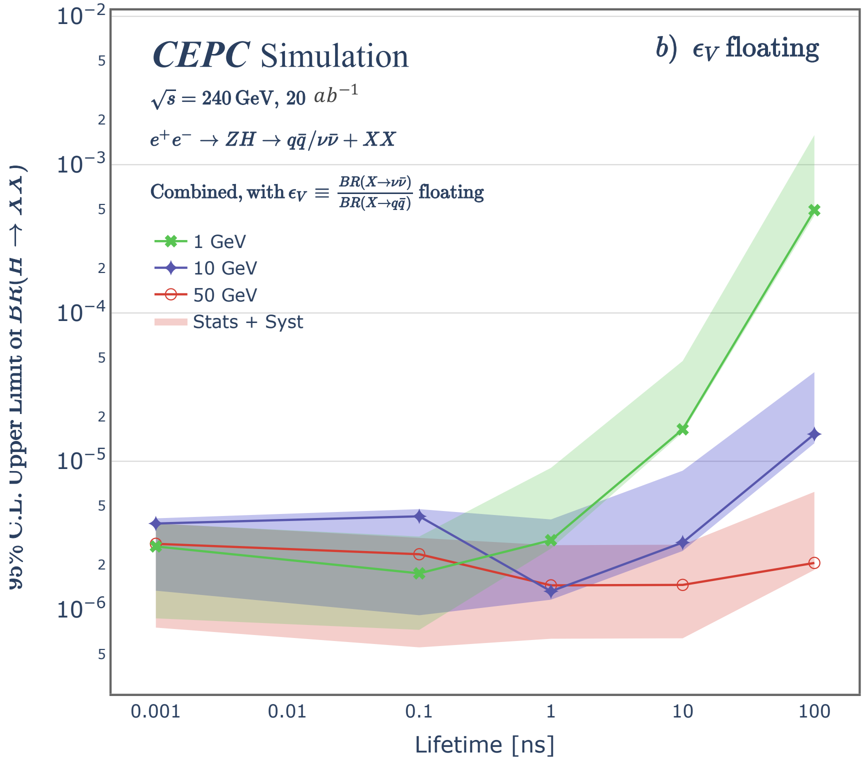
Approach	Efficiency (%)	Lifetime [ns]				
	Mass [GeV]	0.001	0.1	1	10	100
CNN	1	81.8 ± 0.1	90.7 ± 0.1	78.9 ± 0.2	74.4 ± 0.6	76.5 ± 1.9
	10	80.2 ± 0.1	89.5 ± 0.1	91.2 ± 0.1	88.7 ± 0.1	83.6 ± 0.5
	50	87.5 ± 0.1	96.7 ± 0.1	99.3 ± 0.0	98.4 ± 0.0	93.5 ± 0.1
GNN	1	82.9 ± 0.1	89.4 ± 0.1	79.9 ± 0.2	79.9 ± 0.6	80.2 ± 1.8
	10	76.7 ± 0.1	86.0 ± 0.1	91.2 ± 0.1	85.7 ± 0.2	83.7 ± 0.5
	50	88.0 ± 0.1	91.8 ± 0.1	97.4 ± 0.1	97.4 ± 0.1	93.0 ± 0.1

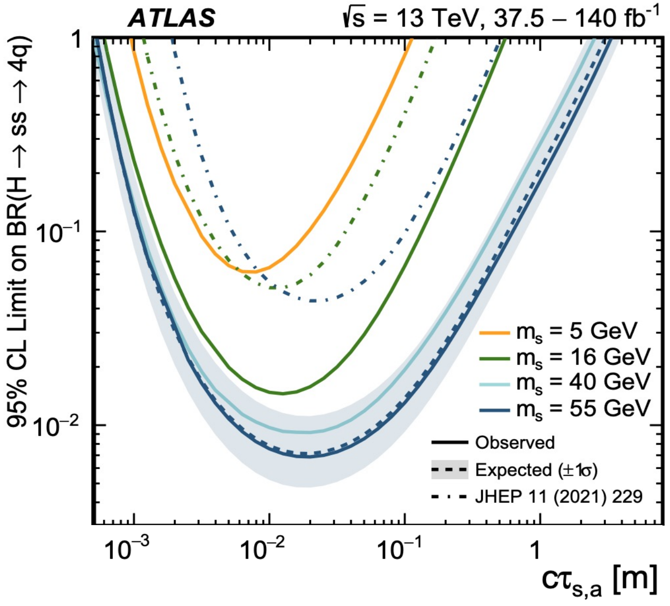


LLP Sensitivity

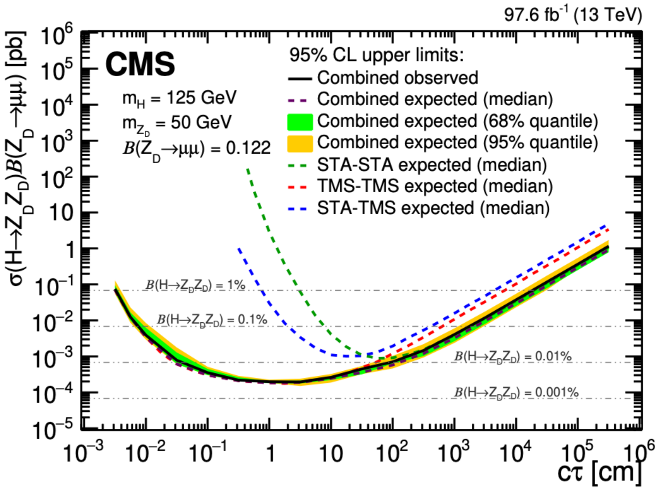
- The best expected limit of BR(H → XX) achieves 10e-6.
- Outperforming the current limit from ATLAS and CMS by two orders of magnitude.
- An order of magnitude better than the ILC's when the lifetime of LLP is over 1ns







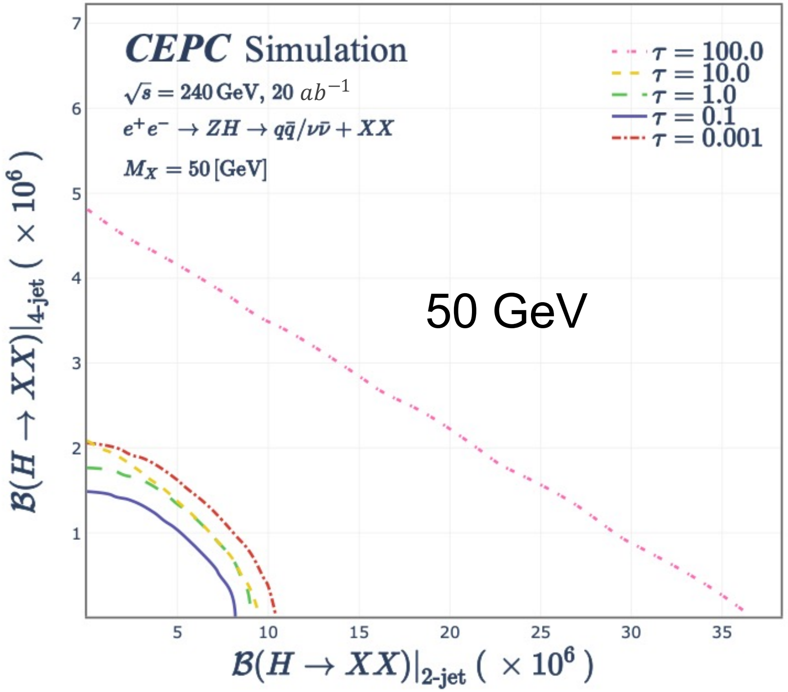
ATLAS [link]

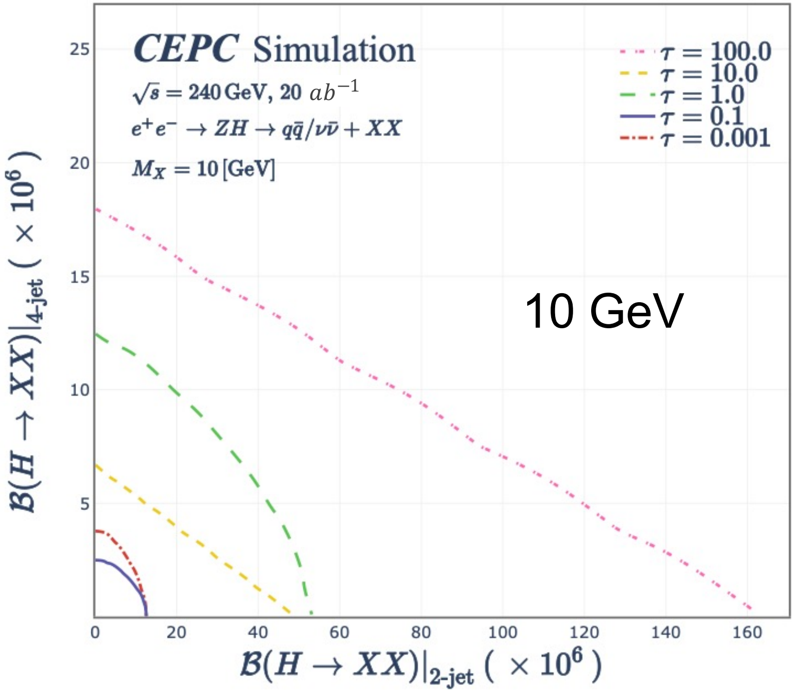


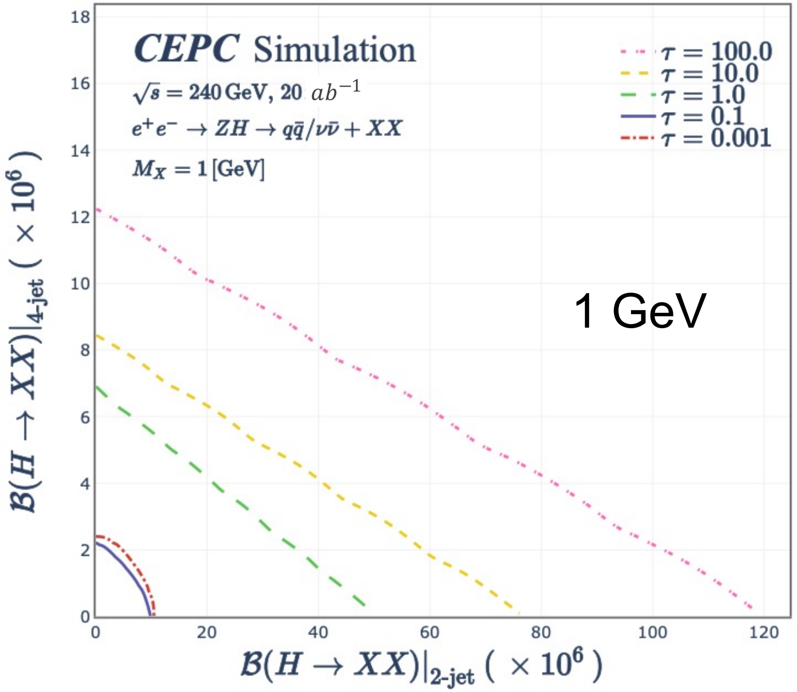
CMS [link]

LLP Sensitivity

- We also provide the 2D likelihood for 95% Confidence Level upper limit on BR(H → XX) with 2 jets and 4 jets final state
- Higher mass and shorter lifetime scenarios have better sensitivities







External Detector Design

Exploring the potential of enhancing the discovery sensitivity by placing an external detector far away from the baseline detector structure

- Multi layers using scintillators, extending from 6 to 106 meters
- The gaining factor is used to evaluate the performance: better improvements for long lifetime scenario

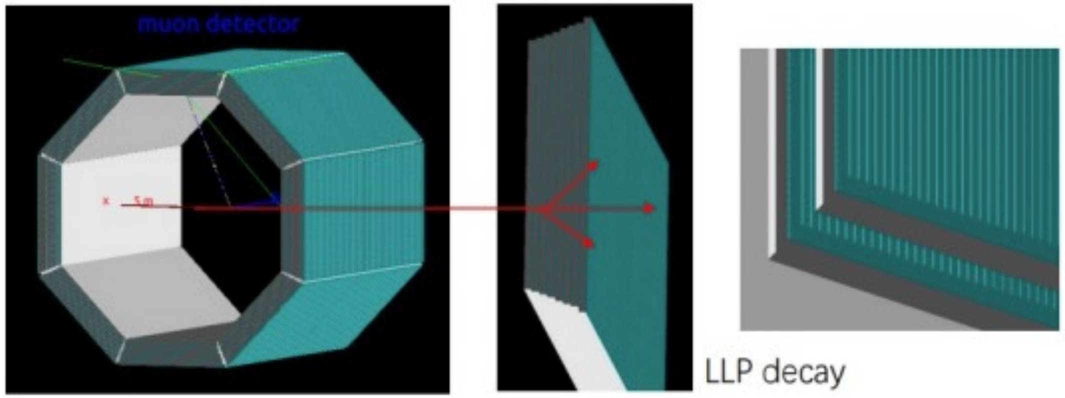
$$F_{\text{gain}} = \frac{N_{\text{obs}}}{N_{\text{gen}}} = \frac{\Delta\Omega}{4\pi}\frac{\Delta L}{d}e^{-\frac{L}{d}}$$

External detector length

Expected LLP decay length

-	Gain	Lifetime [ns]				
	Mass [GeV]	0.001	0.1	1	10	100
Ext. Detector	1	1	1	3.2	11.6	16.2
	10	1	1	1	3.3	11.8
	50	1	1	1	1.1	3.6

an external detector covering distance of 100 meters



Summary and Outlook

LLPs Search with Deep Learning at Lepton Collider

- **Clean environment with distinct detector signature**
- **Best exclusion limit on BR(H→LLPs) @ 20 ab-1: 1.2e-6**
    - 1D and 2D sensitivity results
- **Significant enhancement from deep learning techniques**
    - Simplified analysis strategy compared to the traditional method
        - Low-level detector information without full reconstruction
    - Signal efficiency as high as 99%
    - Short lifetime: biggest improvement
- **The external detector might help with long lifetime LLPs search**

Backups

Systematics Uncertainties

Several sources of systematic uncertainties have been carefully evaluated:

(i) **Uncertainty from Higgs Boson Count**: The uncertainty originating from the total number of Higgs bosons is estimated to be 1.0% [41]. This accounts for variations in the production rate and experimental measurement of Higgs boson events.

(ii) **Neural Network Training Variability**: To assess the robustness of our machine learning model, we trained 50 CNNs with different initial seeds. The training uncertainty for the neural networks is estimated as half of the difference between the maximum and minimum efficiencies observed, which amounts to approximately 1.7%.

A quadratic sum of the above two uncertainties yields a total systematic uncertainty of 2.0%.

LLP Limits

Table 5: The 95% C.L. exclusion limit on BR($h \to XX$) for all signal channels with both fixed and floating ϵ_V. The limits include $\pm 1\sigma$ uncertainties after taking into account both statistical and systematic contributions.

Scenario	$\mathcal{B}$ ($\times 10^{-6}$)	Lifetime [ns]				
	Mass [GeV]	0.001	0.1	1	10	100
Fixed	1	$1.9^{+0.5}_{-0.3}$	$1.5^{+0.4}_{-0.2}$	$5.0^{+1.4}_{-0.7}$	$32.4^{+8.0}_{-3.6}$	$933.1^{+237.3}_{-74.4}$
	10	$2.7^{+0.8}_{-0.2}$	$1.9^{+0.6}_{-0.2}$	$2.2^{+0.7}_{-0.3}$	$4.7^{+1.6}_{-0.6}$	$27.6^{+6.0}_{-2.0}$
	50	$1.5^{+0.6}_{-0.1}$	$1.2^{+0.4}_{-0.1}$	$1.4^{+0.3}_{-0.2}$	$1.4^{+0.3}_{-0.2}$	$3.4^{+1.0}_{-0.3}$
Floating	1	$2.7^{+1.1}_{-1.8}$	$1.8^{+1.4}_{-1.0}$	$2.9^{+6.1}_{-0.4}$	$16.4^{+31.0}_{-0.8}$	$493.4^{+1090.1}_{-39.7}$
	10	$3.8^{+0.3}_{-2.5}$	$4.3^{+0.5}_{-3.3}$	$1.3^{+2.7}_{-0.2}$	$2.8^{+5.8}_{-0.4}$	$15.3^{+24.6}_{-2.0}$
	50	$2.8^{+1.0}_{-2.0}$	$2.4^{+0.7}_{-1.8}$	$1.5^{+1.3}_{-0.8}$	$1.5^{+1.3}_{-0.8}$	$2.1^{+4.2}_{-0.2}$

Cut Bases Analysis Results

Table A1: Number of signal and background events after various selection criteria. MC simulation samples are produced corresponding to the integrated luminosity 5.6 ab^{-1}. Signal events have an LLP lifetime of 20 ns and a mass of 50 GeV.

Selections	LLPs Signal with $Z \to j\bar{j}$	$ee \to q\bar{q}$	$ee \to ZH$
generated	1.0×10^6	2.5×10^8	0.99×10^7
decay in muon detector	134559	6516657	796596
$\|m_{q\bar{q}} - m_Z\| < 15GeV$	113723	4013875	39631
$\|m_{q\bar{q}} - m_H\| < 15GeV$	104942	229703	26862
$0.23 < y_{12} < 0.72$	93,517	129,546	20,041
$E_{2jets} > 30\text{GeV}$	69,468	72	16
$min(\Delta T_{j_1}, \Delta T_{j_2}) > 3\text{ns}$	68,368	50	11
Efficiency	**50.80%**	7.7×10^{-6}	1.4×10^{-5}

Network Training Results