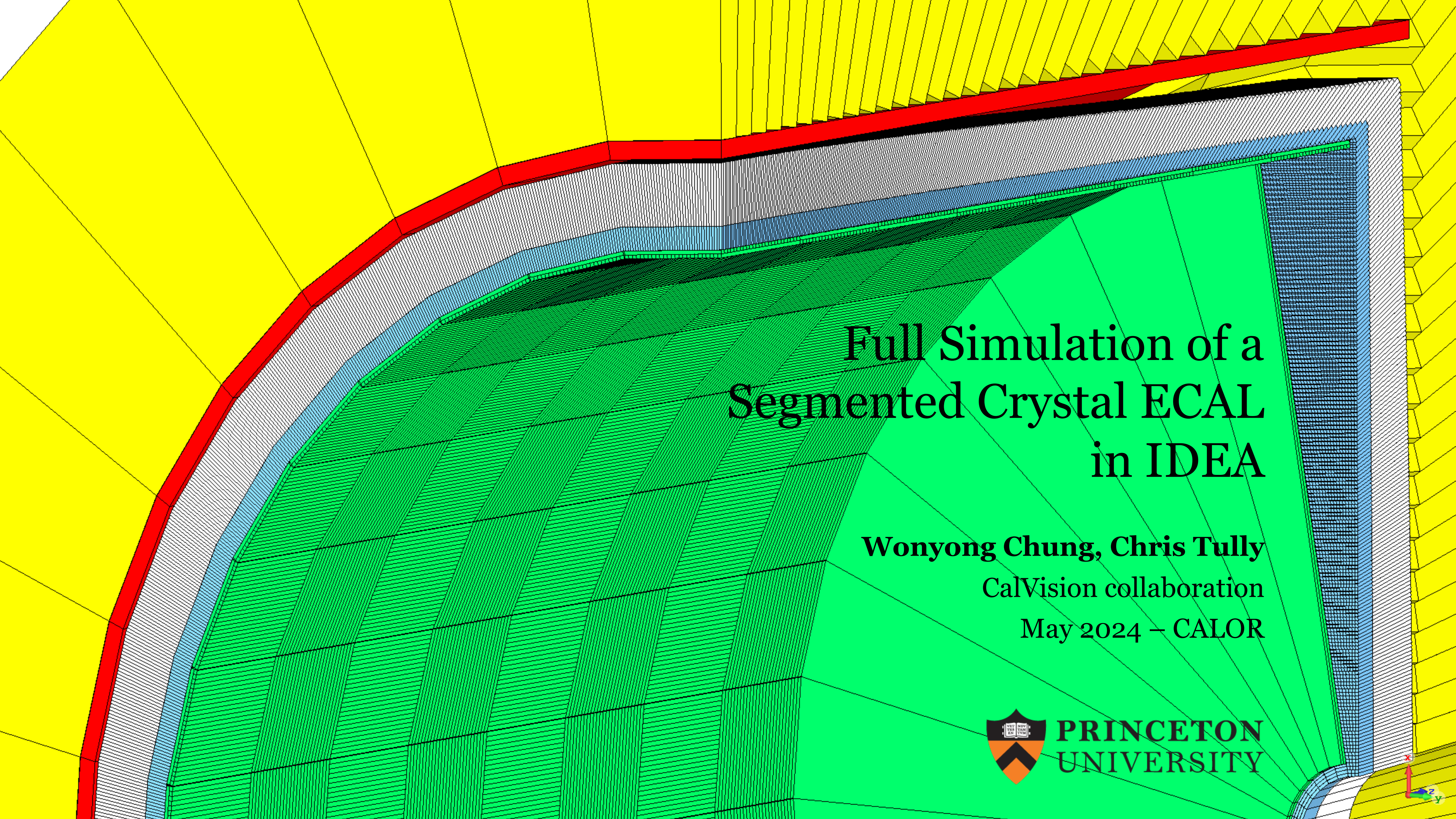


Full Simulation of a Segmented Crystal ECAL in IDEA

Wonyong Chung, Chris Tully

CalVision collaboration

May 2024 – CALOR

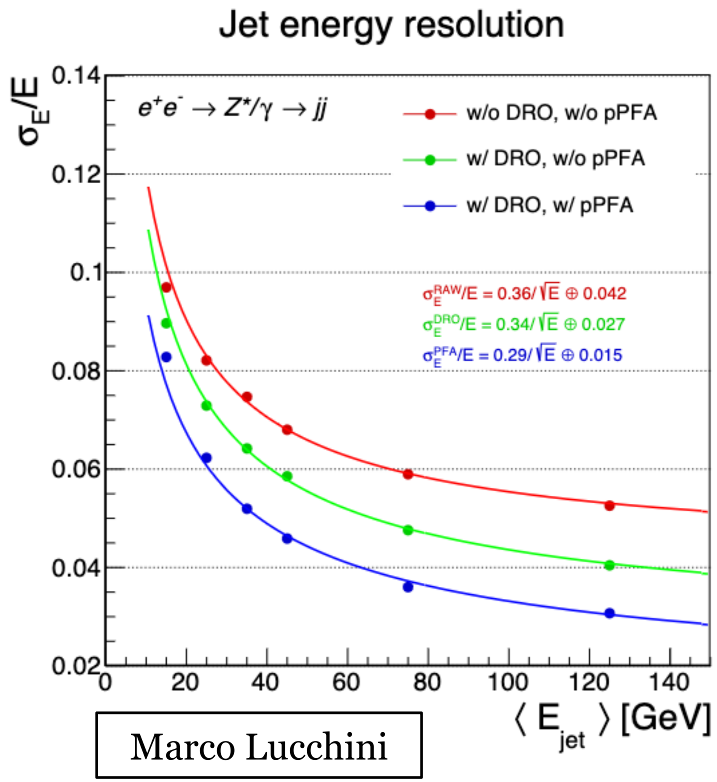
PRINCETON UNIVERSITY

Segmented Crystal EM Precision Calorimeter (SCEPCal)

New perspectives on segmented crystal calorimeters for future colliders
JINST 15 (2020) P11005 [2008.00338]

Particle Flow with a Hybrid Segmented Crystal and Fiber Dual-Readout Calorimeter
JINST 17 (2022) P06008 [2202.01474]

Basic idea covered in several talks
References above remain sound
"proto-PFA" studies done in pure Geant4



Marco Lucchini

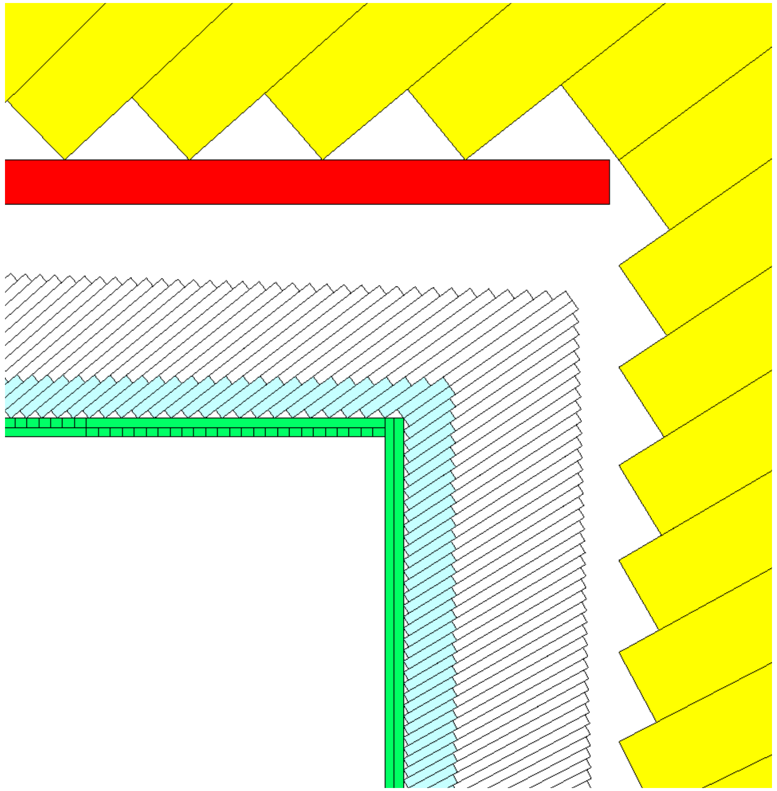
This talk:

Implementation in dd4hep from scratch

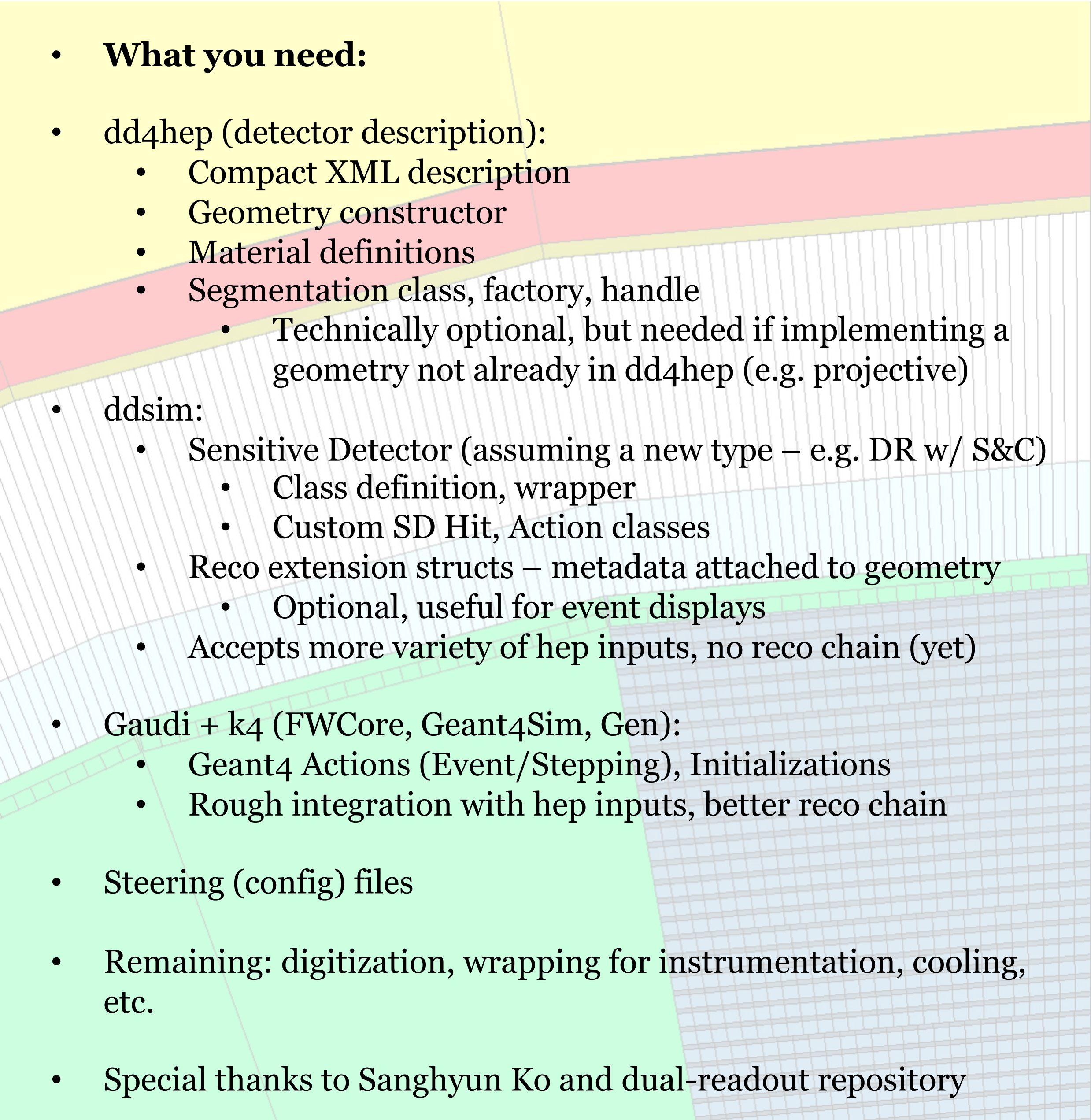
Development environment

Detector description

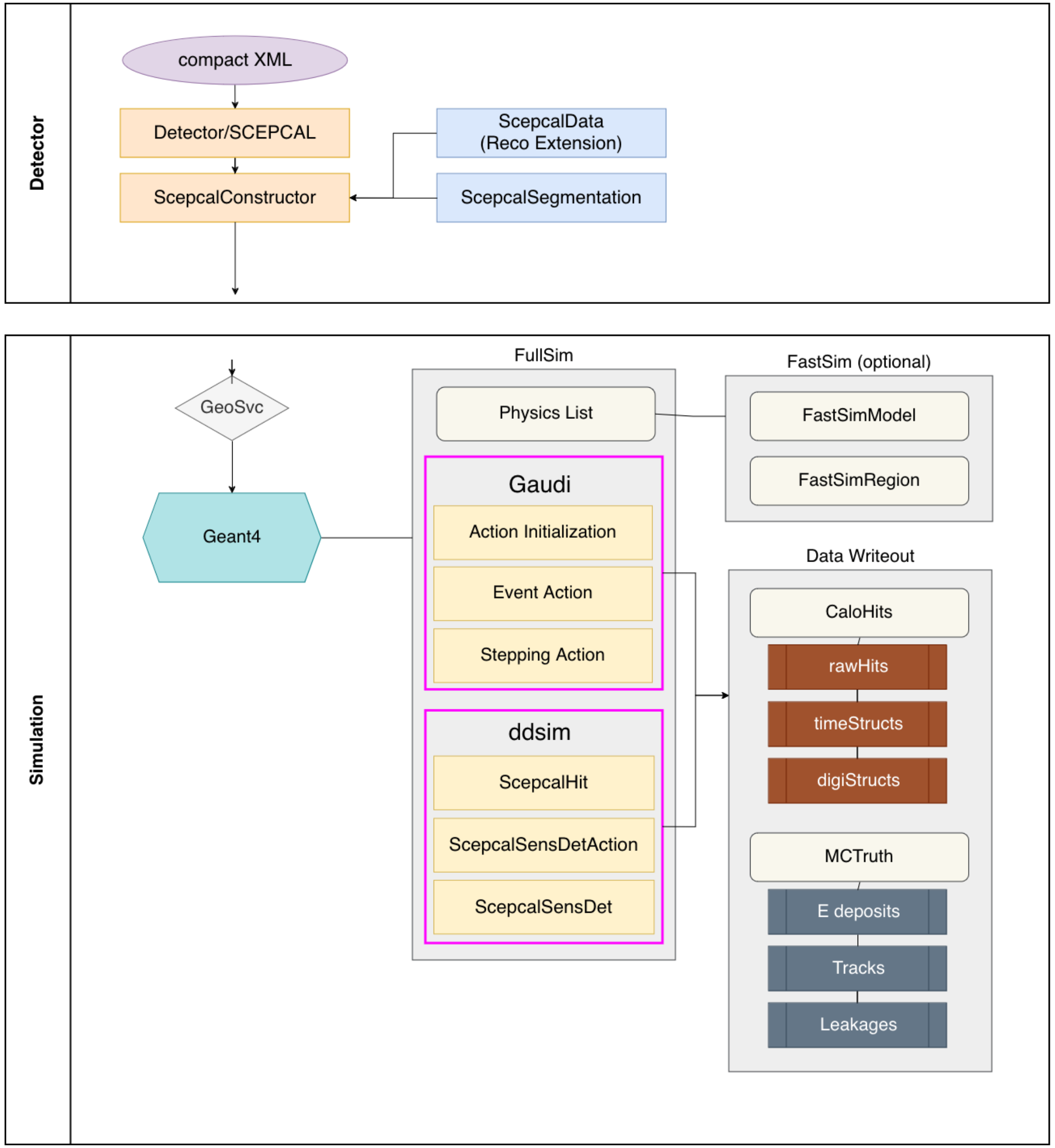
Reconstruction strategies and effective use of neural nets



Implementing a new detector from scratch in FCC-SW



- **What you need:**
- dd4hep (detector description):
  - Compact XML description
  - Geometry constructor
  - Material definitions
  - Segmentation class, factory, handle
    - Technically optional, but needed if implementing a geometry not already in dd4hep (e.g. projective)
- ddsim:
  - Sensitive Detector (assuming a new type – e.g. DR w/ S&C)
    - Class definition, wrapper
    - Custom SD Hit, Action classes
  - Reco extension structs – metadata attached to geometry
    - Optional, useful for event displays
  - Accepts more variety of hep inputs, no reco chain (yet)
- Gaudi + k4 (FWCore, Geant4Sim, Gen):
  - Geant4 Actions (Event/Stepping), Initializations
  - Rough integration with hep inputs, better reco chain
- Steering (config) files
- Remaining: digitization, wrapping for instrumentation, cooling, etc.
- Special thanks to Sanghyun Ko and dual-readout repository



Development environment

New detector development necessitates local visualization

- Remote options exist (XForwarding, mount /cvmfs, JSROOT)
- **High-granularity geometries, >100ms latency becomes painful**

Recommendation: AlmaLinux9 + key4hep-spack

- Previously used: Debian, Fedora, nixOS
- Some local path mods needed in key4hep-spack
- Some deep dependencies (e.g. Cython) do not compile on arm64 – so no macOS
- **Forces you to learn software stack – always good**

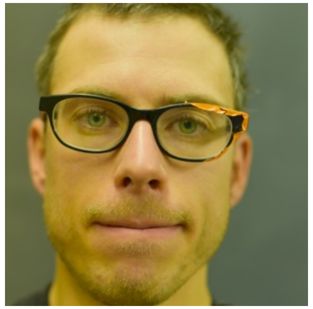
Takeaway:

- Sometimes it’s just hard to know in advance

AlmaLinux

key4hep and spack have matured in 2023

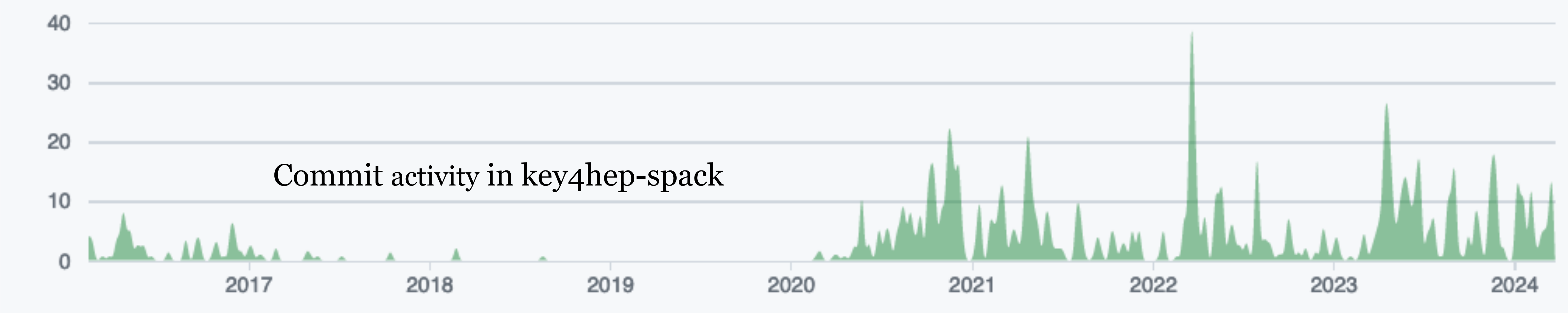
- But someone still had to add the hashes manually



Valentin Volkl

Thomas Madlener

Juan Miguel
Carceller



Commit activity in key4hep-spack

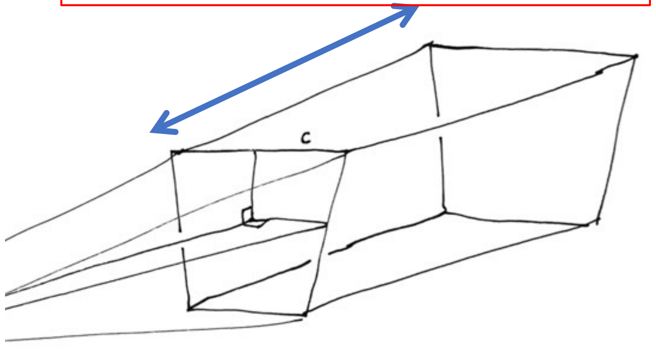
Detector description

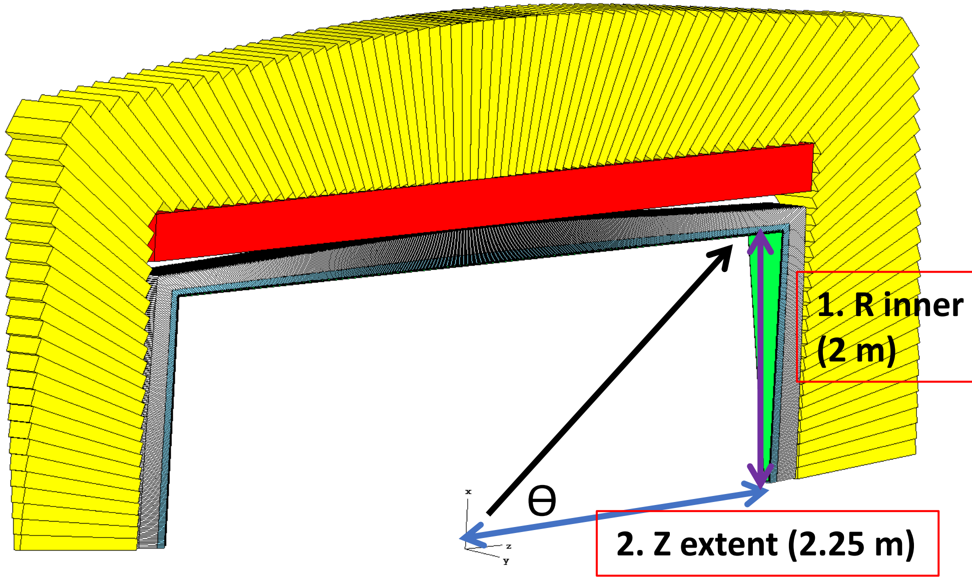
- **Fully parametrized construction**

- Only 7 parameters needed:
    - Inner radius
    - Z extent of barrel
    - Crystal face width (nominal)
    - Front crystal length
    - Rear crystal length
    - Timing crystal thickness (nominal)
    - Number of phi segments
        - Ensures hermeticity
        - Enables timing layer
        - Takes care of projective gaps

- Geometry optimizations
    - Intermediate envelope volumes, <1000 volumes each
    - Orange slices (barrel)
    - Rings (endcap)

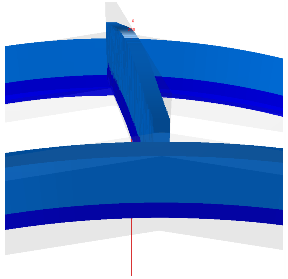
    - ~10x speed/memory improvement vs. monolithic single container

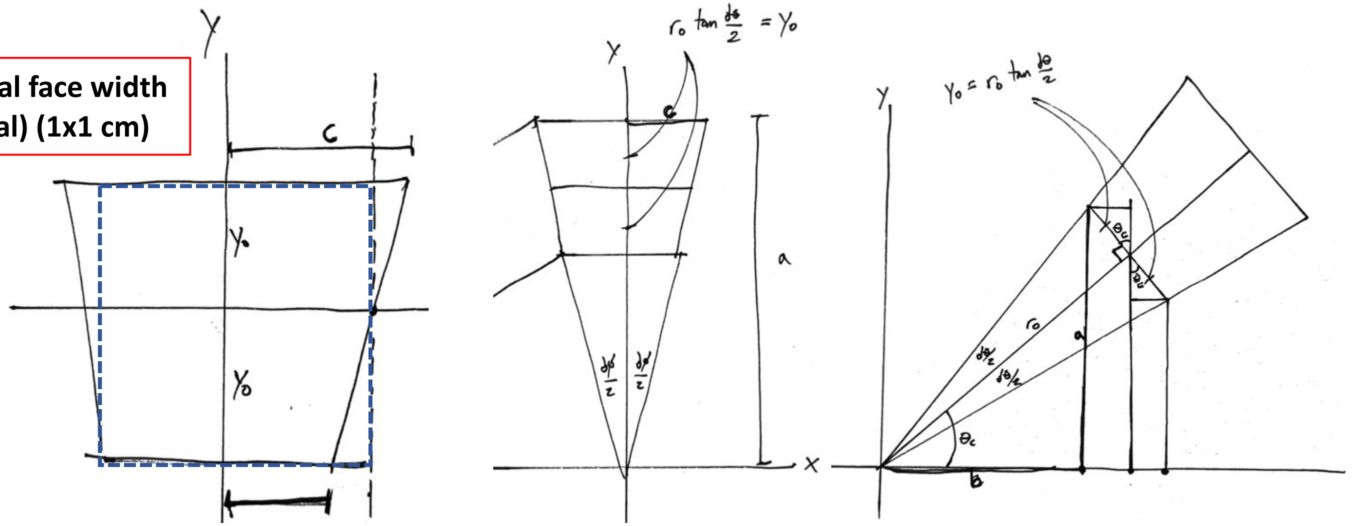

4/5. Front/Rear crystal lengths (5, 15 cm)


7. Number of global phi segments (32)

1. R inner (2 m)

2. Z extent (2.25 m)




3. Crystal face width (nominal) (1x1 cm)

Simulation validation

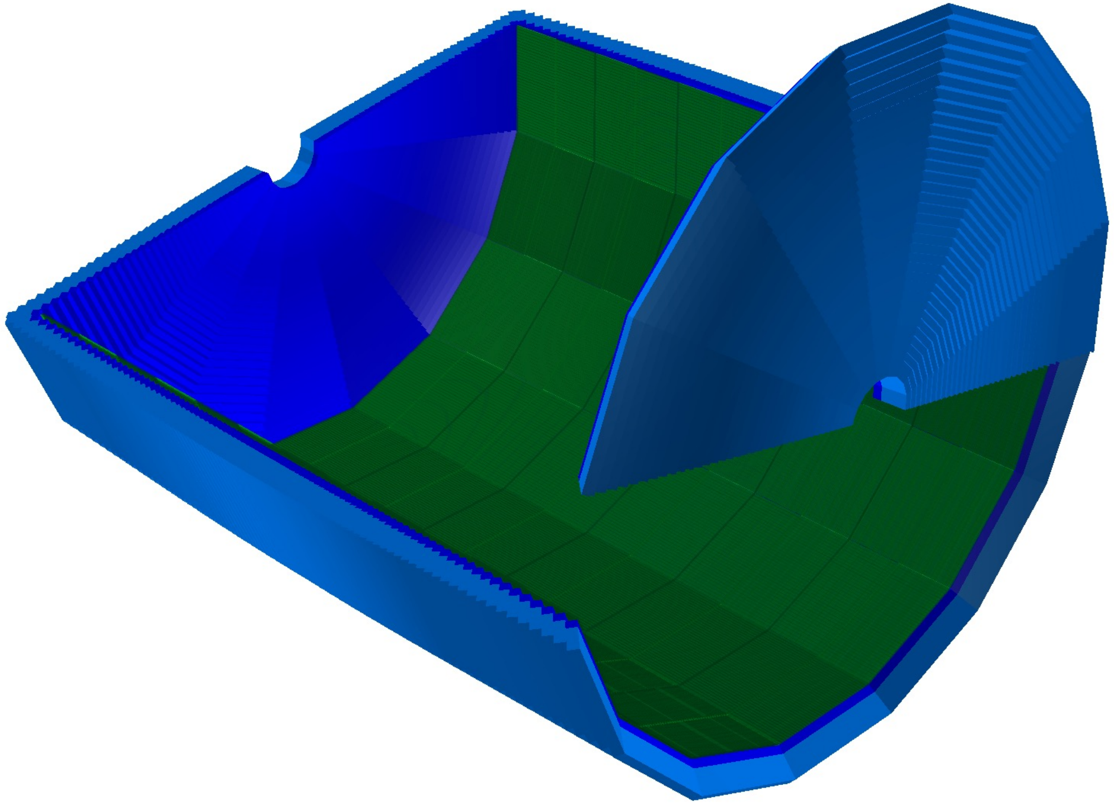
- Runs FCC-ee events and outputs Calo hits

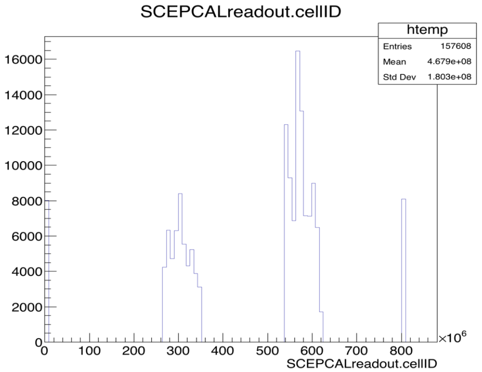
- Baseline w/o instrumentation, wrapping, cooling, etc.
    - Add these in parallel with reco perf tests, understand/isolate systematics
- Plots - ***wzp6_ee_ZZ_test_ecm240*** (10 events)

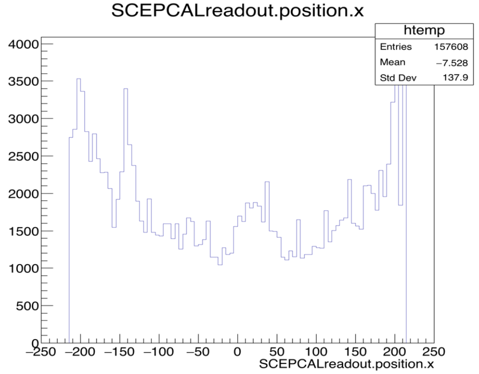
- No optical physics, simply count & terminate S/C
    - GPU efforts an active area (Celeritas, AdePT, CaTS)

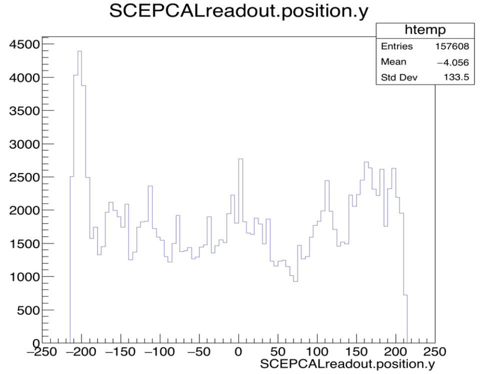
- PbWO4 + LYSO timing layer

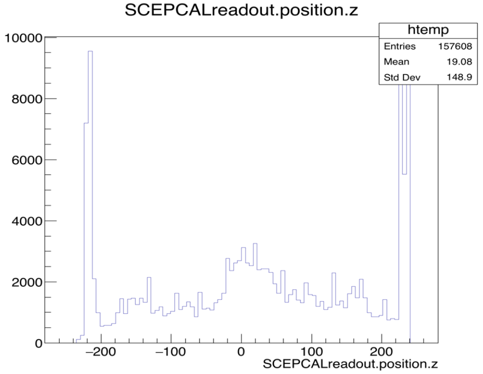
- **With 1x1cm crystal faces/thickness:**
    - **~1.12 million** barrel crystals
    - **~400,000** endcap crystals
    - **~30,000** timing crystals











Reconstruction strategy

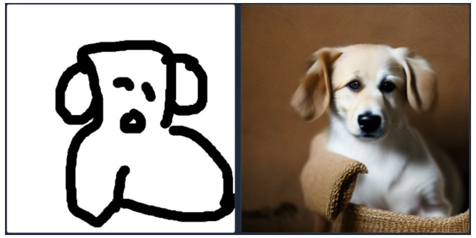
GravNet [1902.07987] (GNN, 2019, distance-weighted, HGCAL)
MLPF [2101.08578] (GNN, 2021, jets/pileup, CMS, comparable to PFA)
ParticleNet [2309.13231] (GNN, 2019, jet tagging unordered particle clouds, CEPC)
Particle Transformer [2202.03772] (Transformers, 2022, kinematics + pairwise interactions, JetClass 100M open data set)
… and many more

- Trend towards unordered particle clouds
  - **Clustering becomes an inner detail of overall reconstruction**
- **Join the fray?**
  - Could start throwing calo hits at the latest NNs…
  - Ongoing work to converge into full detector sim (tracker + ECAL + HCAL)

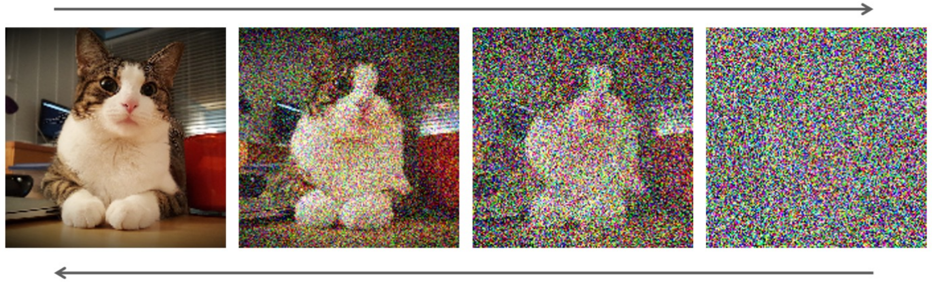
- Meanwhile, need to understand the different neural networks
  - **Unfortunate reality of AI/ML: physics follows computer science**
  - Different types of NNs were originally built, optimized, tuned for a specific task
    - Image recognition (CNNs)
    - Image generation (GANs, Diffusion)
    - Etc.
- **Study how to reframe the physics question into these contexts as much as possible**
  - Understand and integrate the strengths of rule-based PFAs into NNs

Pick the right neural network for the detector

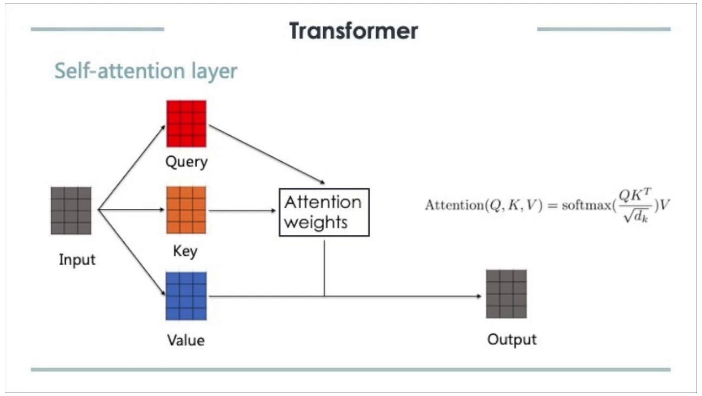
StableDiffusion + ControlNet



- **Tile-based sampling calorimeters:**
  - Diffusion models – **inherently designed to reconstruct from noise**
  - ControlNet/Inpainting – can use **tracks as a guide**



  - Reconstructing from noise – potential to integrate systematics into NNs



- **Multi-segmented/homogeneous calorimeters:**
  - Transformers – **designed to maintain long range context in sequential data, i.e. NLP**
  - Use **tracks for attention, segmentation and timing provide sequence**

`The quick brown fox jumped over the lazy dog`

  - GNNs are generally short range and generally use distance, arguably same kind of assumptions as rule-based PFA

- **Mixed/"Semi-Homogeneous"**:
  - Generative Adversarial Networks – Diffusion model in sampling portion, transformer/GNN in homogeneous, directly leverage the detector design

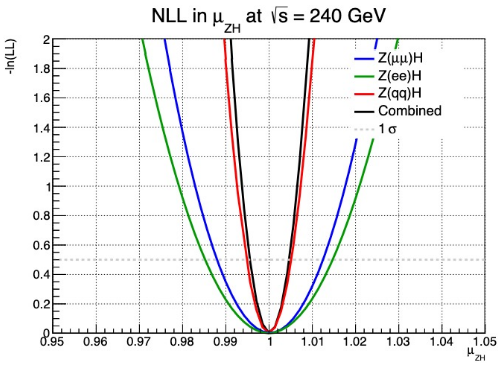
Train according to the physics case

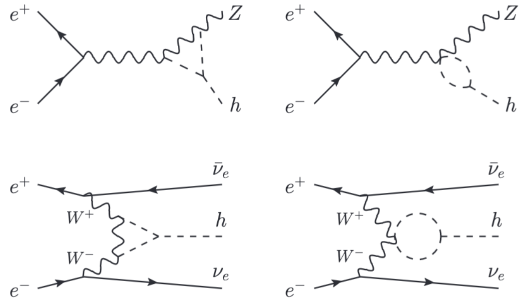
- **Physics case is essential to training approach**

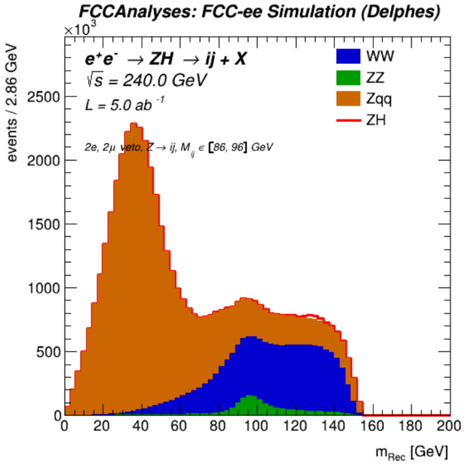
- **For an e+e- collider:**
    - HHH from loop correction to the HZ cross section
    - Z(qq)H apparently dominates
    - Hinges on ability of detector to reduce ZZ background

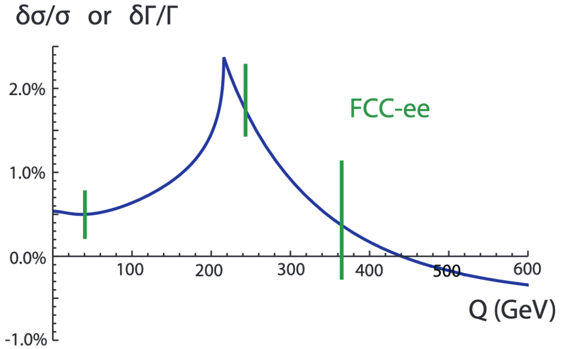
- Foundational approach? – orders of magnitude more data, unexpected performance gains seen in LLMs (circa ~2016-17)
    - Feed all processes/jets at once instead of training on specific subsets, see what happens

- Unclear how to combine separately trained networks on a single readout chain









A special Higgs challenge
https://arxiv.org/pdf/2106.15438

Probing the Higgs sector at the FCC-ee
https://cds.cern.ch/record/2835483

	All classes		$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to 4q$	$H \to \ell\nu qq'$	$t \to bqq'$	$t \to b\ell\nu$	$W \to qq'$	$Z \to q\bar{q}$
	Accuracy	AUC	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{99\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{99.5\%}$	$\text{Rej}_{50\%}$	$\text{Rej}_{50\%}$
PFN	0.772	0.9714	2924	841	75	198	265	797	721	189	159
P-CNN	0.809	0.9789	4890	1276	88	474	947	2907	2304	241	204
ParticleNet	0.844	0.9849	7634	2475	104	954	3339	10526	11173	347	283
ParT	**0.861**	**0.9877**	**10638**	**4149**	**123**	**1864**	**5479**	**32787**	**15873**	**543**	**402**
ParT (plain)	0.849	0.9859	9569	2911	112	1185	3868	17699	12987	384	311

Baselines, new technologies, new benchmarks

- Fully parametrized, hot-swappable sub-detector designs in new frameworks
  - Fast iteration and optimization possible

- Top-down view of detector geometry and data readout
  - "Triggerless" DAQ, flexible online software (see LHCb)
  - Next-gen ASICs *should* be fast enough to run real-time inference on front-end electronics

- Opportunity for unprecedented vertical integration

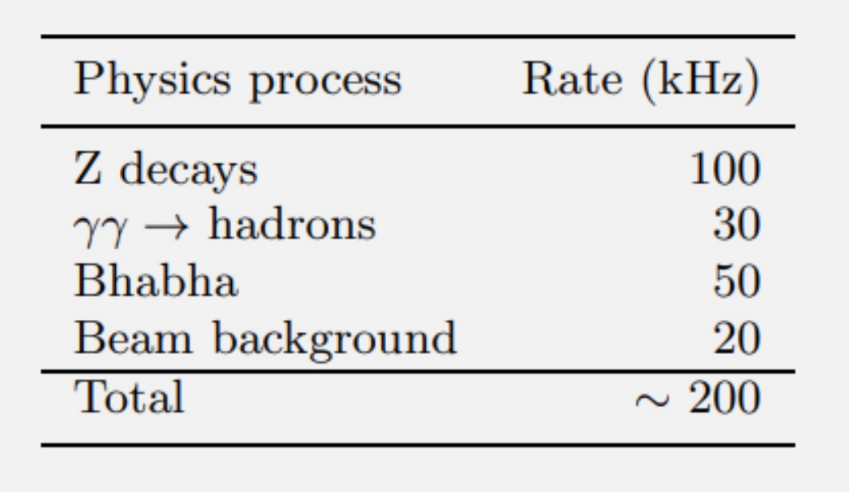
- Current designs are arguably the 'final forms' of CMS, ATLAS, etc.
  - We need more new technologies
    - Lattice-oriented crystals
    - Chromatic calorimetry
    - Quantum sensors for HEP [2311.01930]

- Growing understanding of NN maths as relevant to physics
  - "A mathematical perspective on Transformers" [2312.10794]

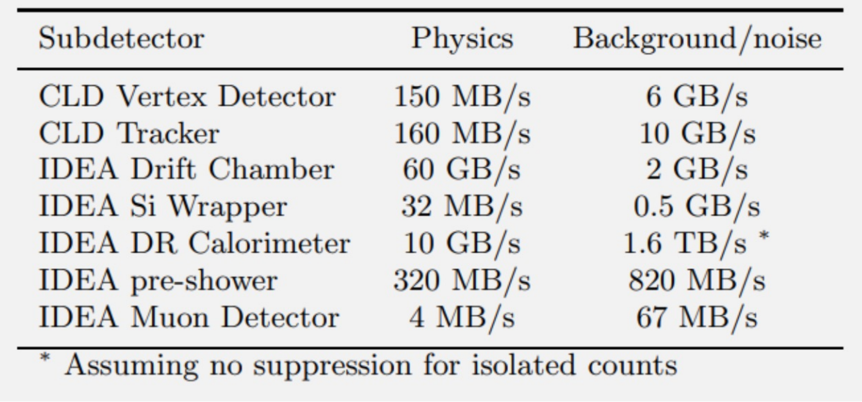
> We develop a mathematical framework for analyzing Transformers based on their interpretation as interacting particle systems, which reveals that clusters emerge in long time.

- **Thank you for your attention!**

Constraints and limitations for FCC detectors



Physics process	Rate (kHz)
Z decays	100
$\gamma\gamma \to$ hadrons	30
Bhabha	50
Beam background	20
Total	~ 200



Subdetector	Physics	Background/noise
CLD Vertex Detector	150 MB/s	6 GB/s
CLD Tracker	160 MB/s	10 GB/s
IDEA Drift Chamber	60 GB/s	2 GB/s
IDEA Si Wrapper	32 MB/s	0.5 GB/s
IDEA DR Calorimeter	10 GB/s	1.6 TB/s *
IDEA pre-shower	320 MB/s	820 MB/s
IDEA Muon Detector	4 MB/s	67 MB/s

* Assuming no suppression for isolated counts