Offline Electron Seeding Validation - Update

Caleb Fangmeier
Ilya Kravchenko, Greg Snow

University of Nebraska - Lincoln

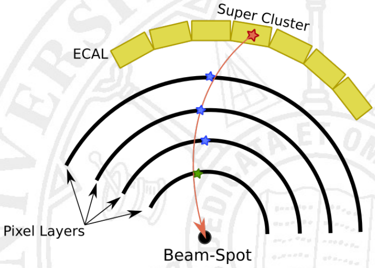
EGamma Reco/Comm/HLT Meeting — March 16, 2018

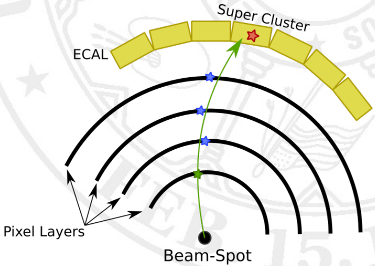
Introduction

- Our goal is to study **seeding** for the **offline** GSF tracking with the **new pixel detector**.
- Specifically, we want to optimize the new pixel-matching scheme from HLT for use in off-line reconstruction.
- This Talk:
  - Show corrected performance comparisons between old and new seeding
  - Show reduction in number of seeds not resulting in GSF tracks
- Additional plots are available here https://eg.fangmeier.tech/seeding_studies_2018_03_08_17/output/

N-Hit Electron Seeding

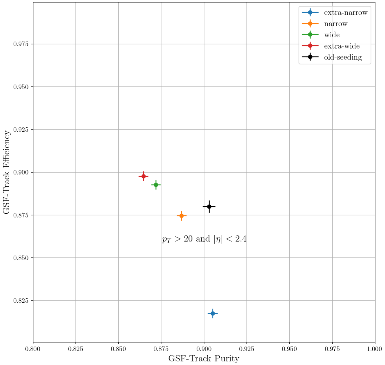
1. Using the beam spot, the SC position, and SC energy, propagate a path through the pixels.
2. Require the first hit to be within a $\delta\phi$ and δz window. ($\delta\phi$ and δR for FPIX)
3. δz window for first hit is huge as SC and beam spot positions give very little information about z.
4. Forget the SC position, and propagate a new track based on the vertex and first hit positions, and the SC energy.
5. Progress one-by-one through the remaining hits in the seed and require each one fit within a specified window around the track.
6. Quit when all hits are matched, or a hit falls outside the window. No skipping is allowed.





Previous status-quo

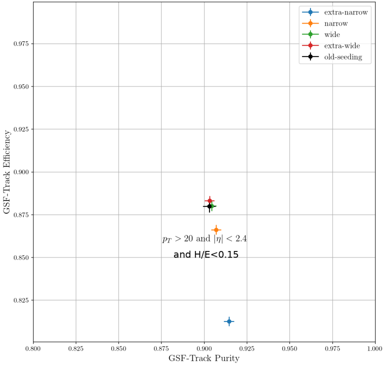
- In a previous presentation[1], I showed efficiency vs. purity for
  - Old pair-match seeding (`ElectronSeedProducer`)
  - New triplet+ seeding (`ElectronNHitSeedProducer`) for several choices of matching windows.
- Old seeding produced far fewer fake (non-truth matched) seeds at similar efficiency.
- Unclear why. Perhaps not optimal matching windows?



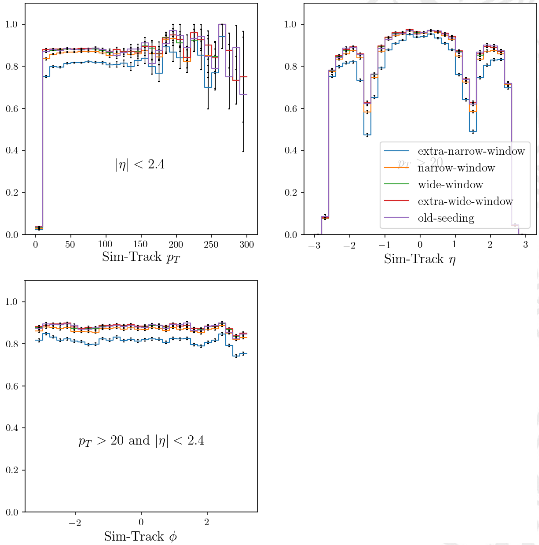
[1] https://indico.cern.ch/event/697074/contributions/2898322/attachments/1602057/2540261/main.pdf

H/E Requirement on Super-Clusters

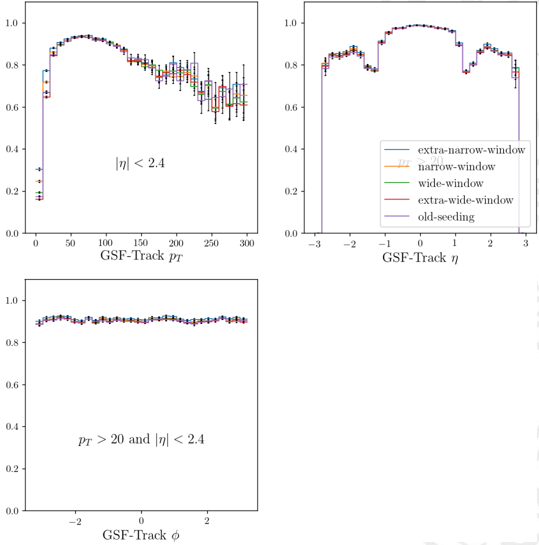
- Investigating the cause of this revealed that the old seeding had applied a $H/E < 0.15$ cut on super-clusters.
- After applying this same cut on the new seeds, the performance gap becomes negligible.
- The **narrow** working point of the new seeding uses HLT window sizes(see backup).
- Performance of the old seeding can be closely matched with the **wide** matching windows.



Kinematic Distributions - Efficiency



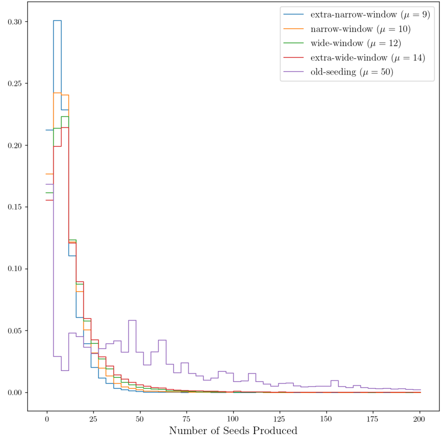
Kinematic Distributions - Purity



wide and **old-seeding** are comparable across all kinematic ranges

Seed Counts

- The amount of electron seeds is dramatically reduced by the new matching scheme.
- Part of the motivation for use in HLT.
- Factor of ≈ 4 reduction comparing to the `wide` working point



Conclusions & Outlook

- The HLT `NHit` settings are more restrictive than the current **offline**/**old-seeding** ones
- Opening them up to the **wide** settings matches current **offline** performance.
- `NHit` seeding can match performance while producing far fewer candidate seeds.

- Next Steps:
  - Verify performance is still acceptable in high fake environments ($t\bar{t}$ for example)
  - Settle on an "optimal enough" set of windows
  - Decide on which CMSSW release to target and begin preparing (private branch/merge request/etc.)
- Other Thoughts
  - What are expert's opinions on continued window optimization? (Likely some small gains still to be had)
  - Are there other ideas for cross-checks to be done before proceeding further?

BACKUP

Definitions

- **Sim-Track -** A track from a simulated electron originating from the luminous region of CMS (beam-spot +- 5σ)
- **ECAL-Driven Seed -** A seed created via a matching procedure between Super-Clusters and General Tracking Seeds (Either from `ElectronSeedProducer` or `ElectronNHitSeedProducer`)
- **GSF Track -** A track from GSF-Tracking resulting from an **ECAL-Driven Seed**
- **Seeding Efficiency -** The fraction of **Sim-Tracks** that have a matching **ECAL-Driven Seed** (based on simhit-rechit linkage)
- **GSF Tracking Efficiency -** The fraction of **Sim-Tracks** that have a matching **GSF Track** (again, based on simhit-rechit linkage)
- **ECAL-Driven Seed Purity -** The fraction of **ECAL-Driven Seeds** that have a matching **Sim-Track**
- **GSF Tracking Purity -** The fraction of **GSF Tracks** that have a matching **Sim-Track**

Matching Window Parameters

		extra-narrow	narrow(HLT)	wide	extra-wide
Hit 1	dPhiMaxHighEt	**0.025**	**0.05**	**0.1**	**0.15**
	dPhiMaxHighEtThres	20.0	20.0	20.0	20.0
	dPhiMaxLowEtGrad	-0.002	-0.002	-0.002	-0.002
	dRzMaxHighEt	9999.0	9999.0	9999.0	9999.0
	dRzMaxHighEtThres	0.0	0.0	0.0	0.0
	dRzMaxLowEtGrad	0.0	0.0	0.0	0.0
Hit 2	dPhiMaxHighEt	**0.0015**	**0.003**	**0.006**	**0.009**
	dPhiMaxHighEtThres	0.0	0.0	0.0	0.0
	dPhiMaxLowEtGrad	0.0	0.0	0.0	0.0
	dRzMaxHighEt	**0.025**	**0.05**	**0.1**	**0.15**
	dRzMaxHighEtThres	30.0	30.0	30.0	30.0
	dRzMaxLowEtGrad	-0.002	-0.002	-0.002	-0.002
Hit 3+	dPhiMaxHighEt	**0.0015**	**0.003**	**0.006**	**0.009**
	dPhiMaxHighEtThres	0.0	0.0	0.0	0.0
	dPhiMaxLowEtGrad	0.0	0.0	0.0	0.0
	dRzMaxHighEt	**0.025**	**0.05**	**0.1**	**0.15**
	dRzMaxHighEtThres	30.0	30.0	30.0	30.0
	dRzMaxLowEtGrad	-0.002	-0.002	-0.002	-0.002

`NHit` Seeding window parameters. Bold designates modified values.