Offline Electron Seeding Validation - Update

Caleb Fangmeier
Ilya Kravchenko, Greg Snow

University of Nebraska - Lincoln

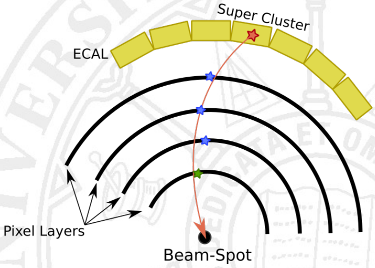
EGM general meeting **CMS week** — April 18, 2018

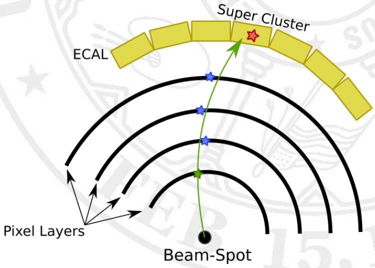
Introduction

- 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 performance comparison between new and old seeding in fake-rich environment
  - Show alternative efficiency/purity measurements using ΔR matching

N-Hit Electron Seeding

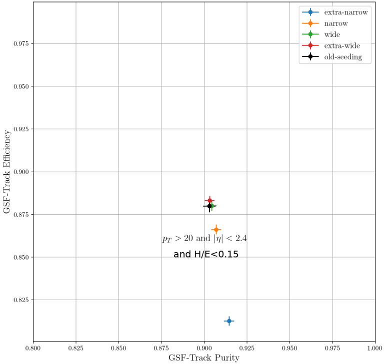
1. Using the beam spot, the SC position, and SC energy, propagate a path through the pixels.
2. Require the first hit to be within a $\delta\phi$ and δz window. ($\delta\phi$ and δR for FPIX)
3. δz window for first hit is huge as SC and beam spot positions give very little information about z.
4. Forget 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

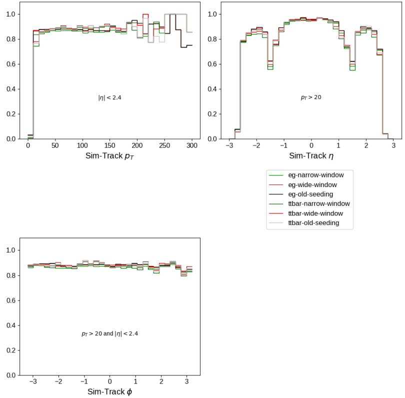
- In a previous presentation[1], I showed efficiency vs. purity for
  - Old pair-match seeding (`ElectronSeedProducer`)
  - New triplet seeding (`ElectronNHitSeedProducer`) for several choices of matching windows.
- Performance of new seeding at the `wide` working point was comparable to old seeding in low-fake ($Z \to e^+e^-$) environment
- Needed to validate performance in a high fake environment.



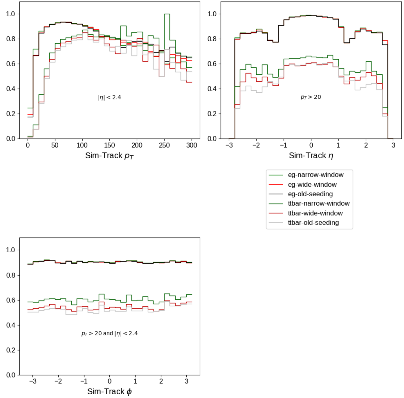
[1] https://indico.cern.ch/event/697077/contributions/2936039/attachments/1618649/2573874/main.pdf

Relative Performance

GSF Tracking Efficiency



GSF Tracking Purity



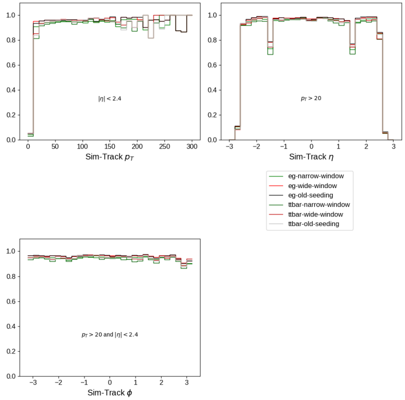
Samples

/ZToEE_NNPDF30_13TeV-powheg_M_120_200/RunIISummer17DRStdmix-NZSFlatPU28to62_92X_upgrade2017_realistic_v10-v1

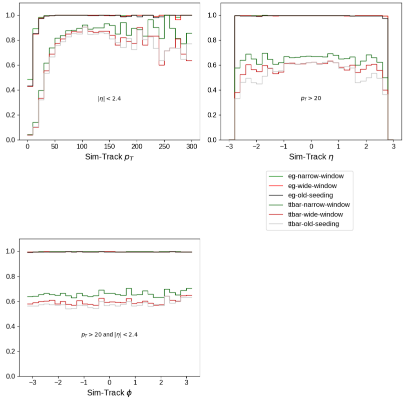
/TT_TuneCUETP8M2T4_13TeV-powheg-pythia8/RunIISummer17DRStdmix-NZSFlatPU28to62_92X_upgrade2017_realistic_v10-v2

ΔR Matching

GSF Tracking Efficiency (ΔR Matched)



GSF Tracking Purity (ΔR Matched)



- Previous efficiency/purity defintions based on shared tracker hits between `SimTracks` and `GSFTracks`.
- An alternative is to use simple ΔR matching.
- Above figures use $\Delta R < 0.2$ for matching criteria.
- Overall numbers improve and show fewer detector effects.

Overall Performance

Integrating over all tracks with $p_T > 20\text{GeV}$ and $\eta < 2.4$ yields the performance numbers below.

Sample	Algo	Efficiency (Hit Matched)	Purity (Hit Matched)
$Z \rightarrow ee$	`old-seeding`	$88.05 \pm 0.28\%$	$90.30 \pm 0.29\%$
	`narrow`	$86.63 \pm 0.28\%$	$90.69 \pm 0.29\%$
	`wide`	$88.01 \pm 0.28\%$	$90.43 \pm 0.29\%$
$t\bar{t}$	`old-seeding`	$88.06 \pm 0.77\%$	$52.35 \pm 0.60\%$
	`narrow`	$86.89 \pm 0.79\%$	$60.56 \pm 0.67\%$
	`wide`	$88.30 \pm 0.77\%$	$54.38 \pm 0.61\%$
Sample	**Algo**	**Efficiency (ΔR Matched)**	**Purity (ΔR Matched)**
$Z \rightarrow ee$	`old-seeding`	$96.08 \pm 0.28\%$	$99.54 \pm 0.29\%$
	`narrow`	$94.49 \pm 0.28\%$	$99.72 \pm 0.29\%$
	`wide`	$96.00 \pm 0.28\%$	$99.60 \pm 0.29\%$
$t\bar{t}$	`old-seeding`	$94.84 \pm 0.77\%$	$57.49 \pm 0.60\%$
	`narrow`	$93.54 \pm 0.79\%$	$65.84 \pm 0.67\%$
	`wide`	$95.06 \pm 0.77\%$	$59.52 \pm 0.61\%$

Note that the `wide` working point of the new seeding matches the `old-seeding` within errors except for purity is $\approx 2\%$ better in the $t\bar{t}$ sample.

Conclusions & Outlook

- The new seeding algorithm at the `wide` working point has been verified to perform as well as, and in some cases better, than the current pair seeding based on MC studies in both low and high purity environments.
- Relative performance is not an artifact of Hit Matching, but can be reproduced with simple ΔR matching.
- Unless there are objections, propose to move forward with implementing the new algorithm as the default in the next available SW release.

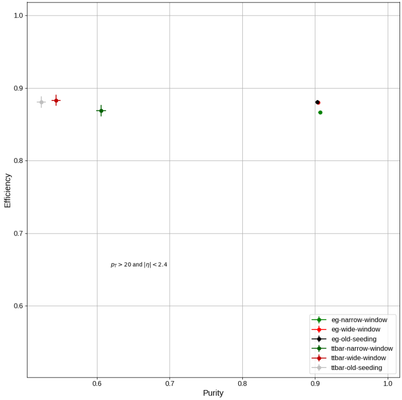
BACKUP

Definitions

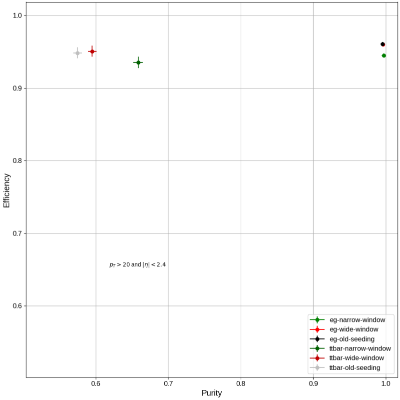
- **Sim-Track -** A track from a simulated electron originating from the luminous region of CMS (beam-spot +- 5σ)
- **ECAL-Driven Seed -** A seed created via a matching procedure between Super-Clusters and General Tracking Seeds (Either from `ElectronSeedProducer` or `ElectronNHitSeedProducer`)
- **GSF 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**

Overall Performance

GSF Tracking Performance (Hit Matched)



GSF Tracking Performance (ΔR Matched)



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.