Seeding in the Tracker Subsystem: Part 2

LUXE Simulation and Analysis Meeting

Arka Santra, December 15, 2020
Weizmann Institute of Science,
Rehovot, Israel

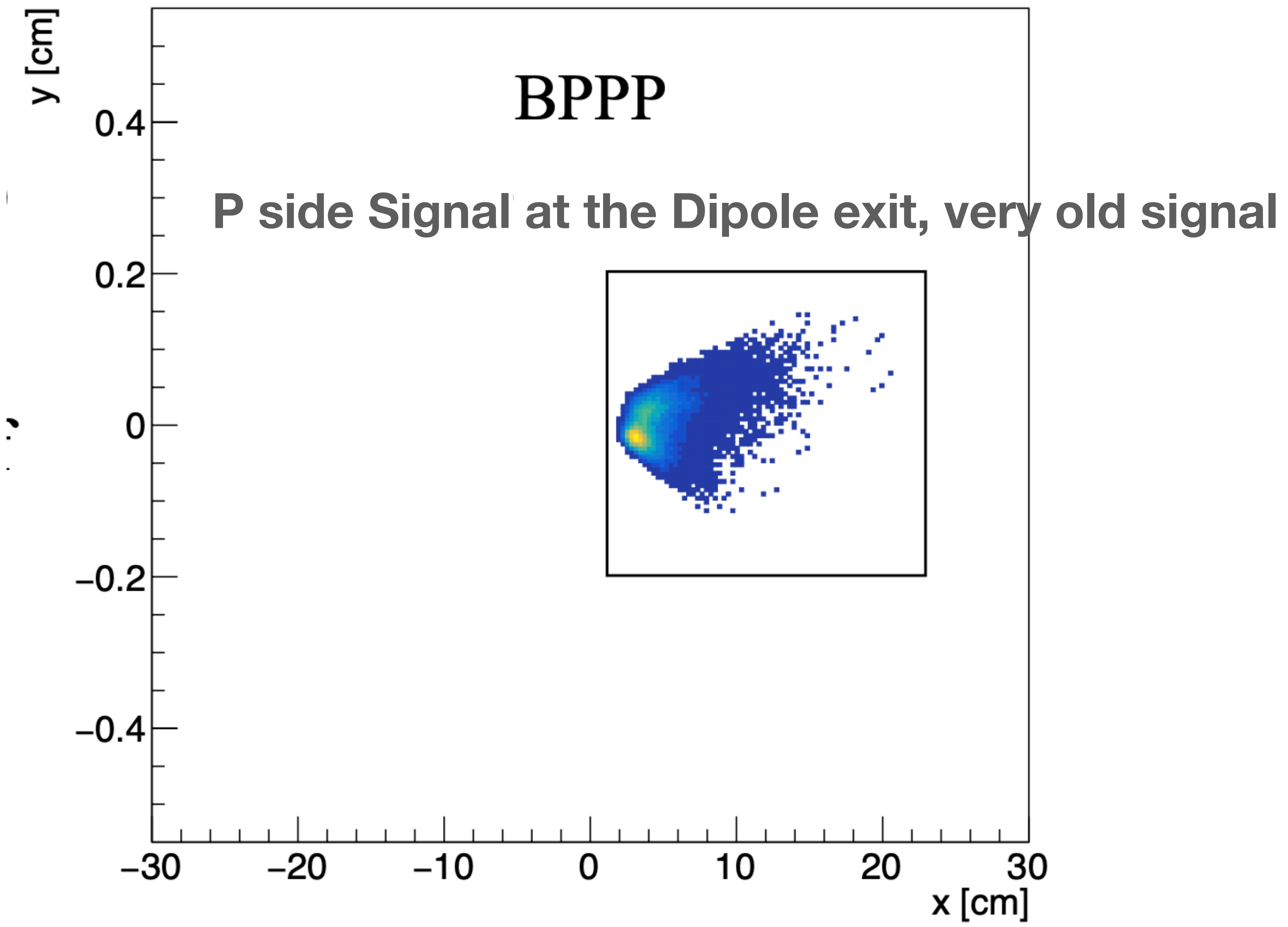
Signal and Background samples

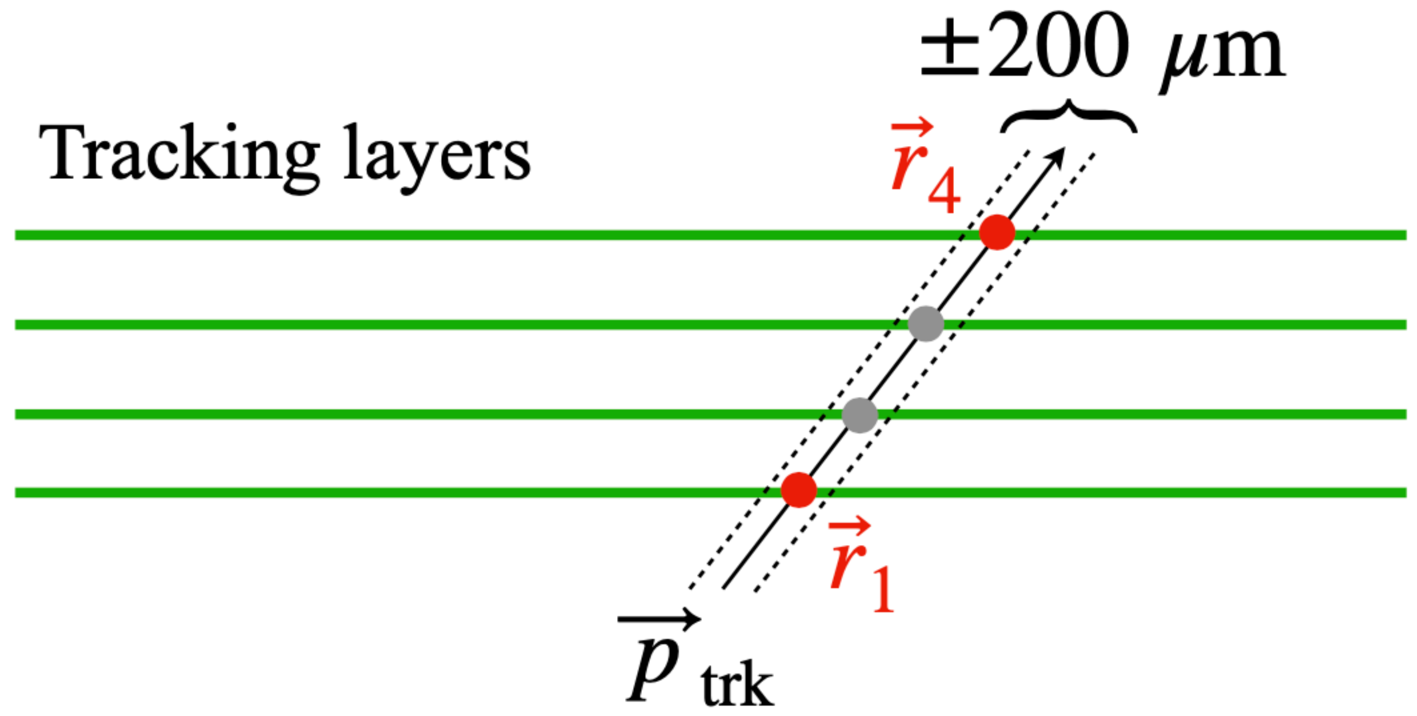
Recipe to merge signal and background

- Taking background tracks from the electron beam only setup.
  - At present, we have ~4 BX of electron beam only background.
  - Applied the factor 0.826 so that only those electrons not interacting with laser are taken.
  - 1 BX is 1.5e9*0.826 electrons produced at the simulation.
    - Make bunches of electron beam only simulated files such that each bunch now has 1.5e9*0.826 electrons.
- Signal tracks are from the e+laser JETI40 setup.
  - Used w0 3000 nm files (Old, low number of signal tracks, ~1/BX)
  - Used w0 5000 nm files(Updated, high number of signal tracks, ~500/BX)
- Mixing the signal and background per BX
  - For each BX of background, add 1 BX of signal (take signal BXs having highest number of signal tracks)

The seeding algorithm (Old)

- Keep unique set of tracks from first layer and last layer of tracker.
  - Overlap region is removed (by cutting on x value of the tracks) from the outer stave of first layer (innermost) and from the inner stave of the last layer (outermost).
- Loop over all pairs in layers 4 and 1, now only positron side (x>0).
- Reject pair of clusters if
  - |x1| > |x4| or they have different sign
  - |z1| == |z4|
  - |y_exit| > 20/2 mm
  - |x_exit| < 5 mm and |x_exit| > xDipoleWidth (330 mm) / 2
  - If not one cluster in the road of 200 um connecting vector r1 and r4 in both layer 2 and layer 3.
  - The seed energy is less than 1 GeV or 16.5 GeV.
    - The seed energy is calculated from the track.





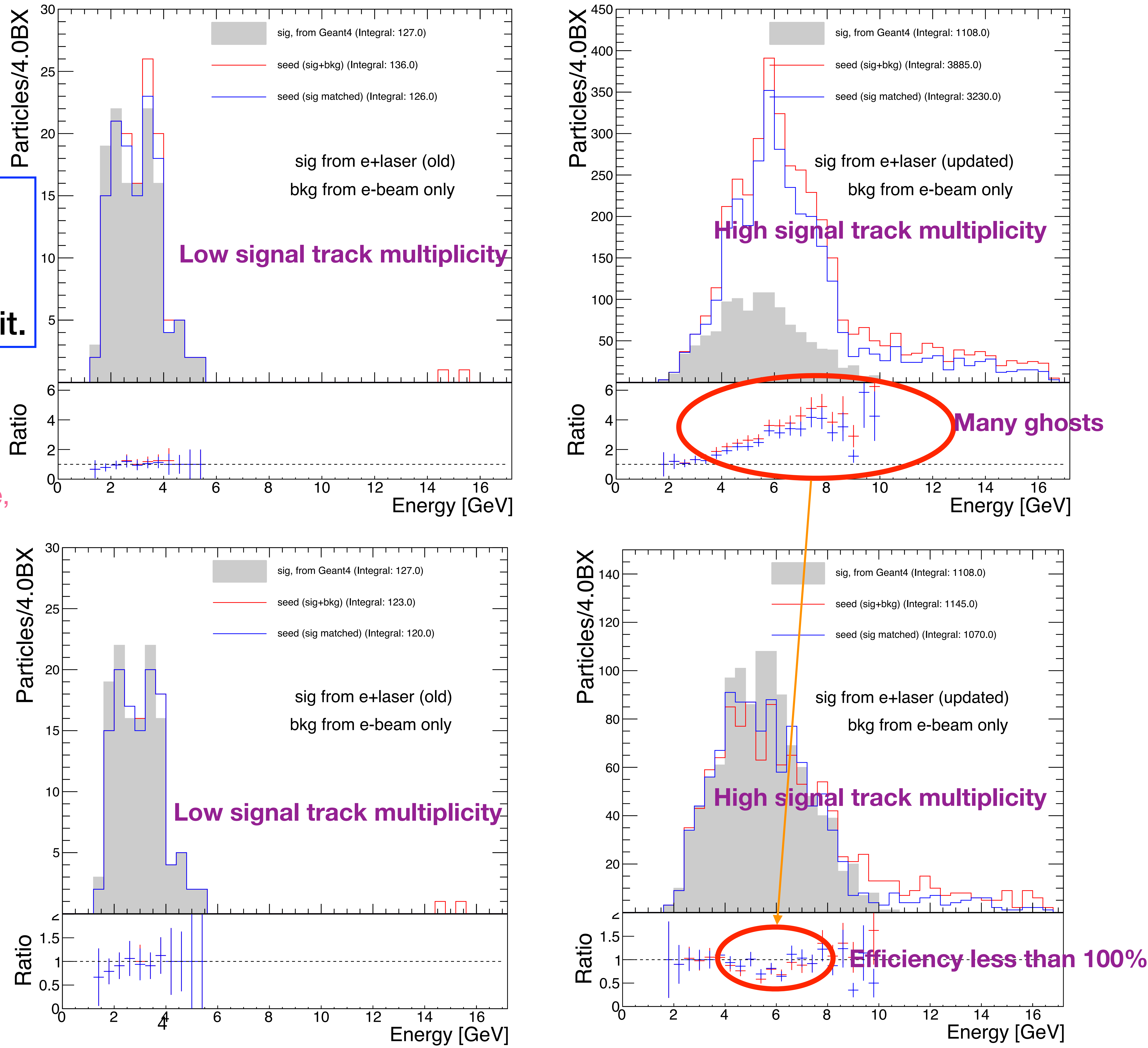
Seed energy plots

- 4 BXs are used to improve the statistics, seeding was done per BX
- Grey fill is the true energy from Geant4
- Red line is the energy from seed tracks, both signal and background combined.
- Blue is the energy from seed tracks, only signal tracks
  - Matching of signals cannot be done track per track wise, as all signal tracks have the same track id (==1).

keep layer 1 hit even if one seed was created from it.

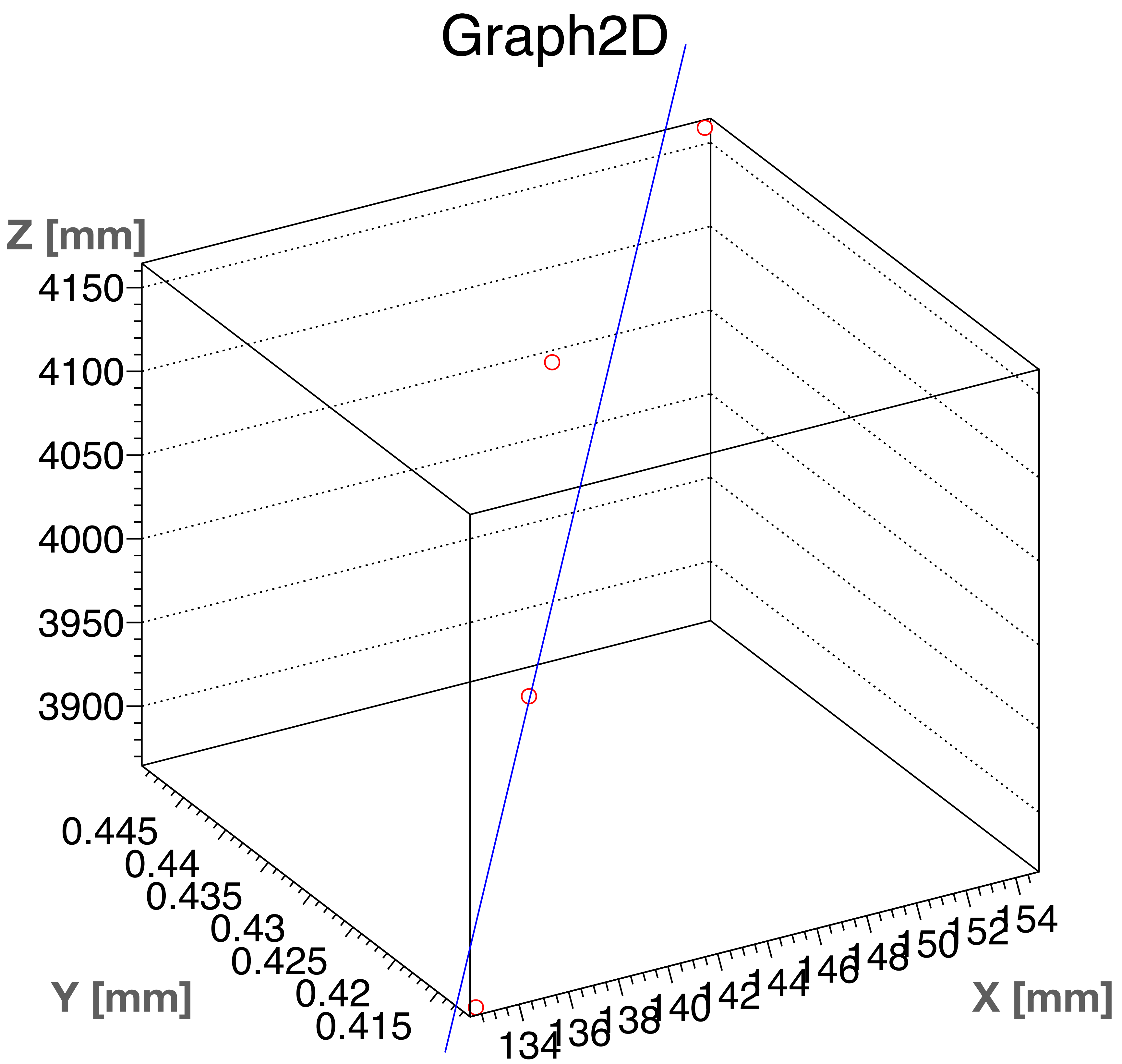
- High signal tracks create many more ghost tracks.
  - We can reduce the ghosts if we remove layer 1 "hit" which is already in a seed.
  - At this point, not too worried about having too many seeds. Rather we need to see if we have enough seeds not to have inefficiency.
- **We need to improve the scenario.**

remove layer 1 hit if one seed was created from it.

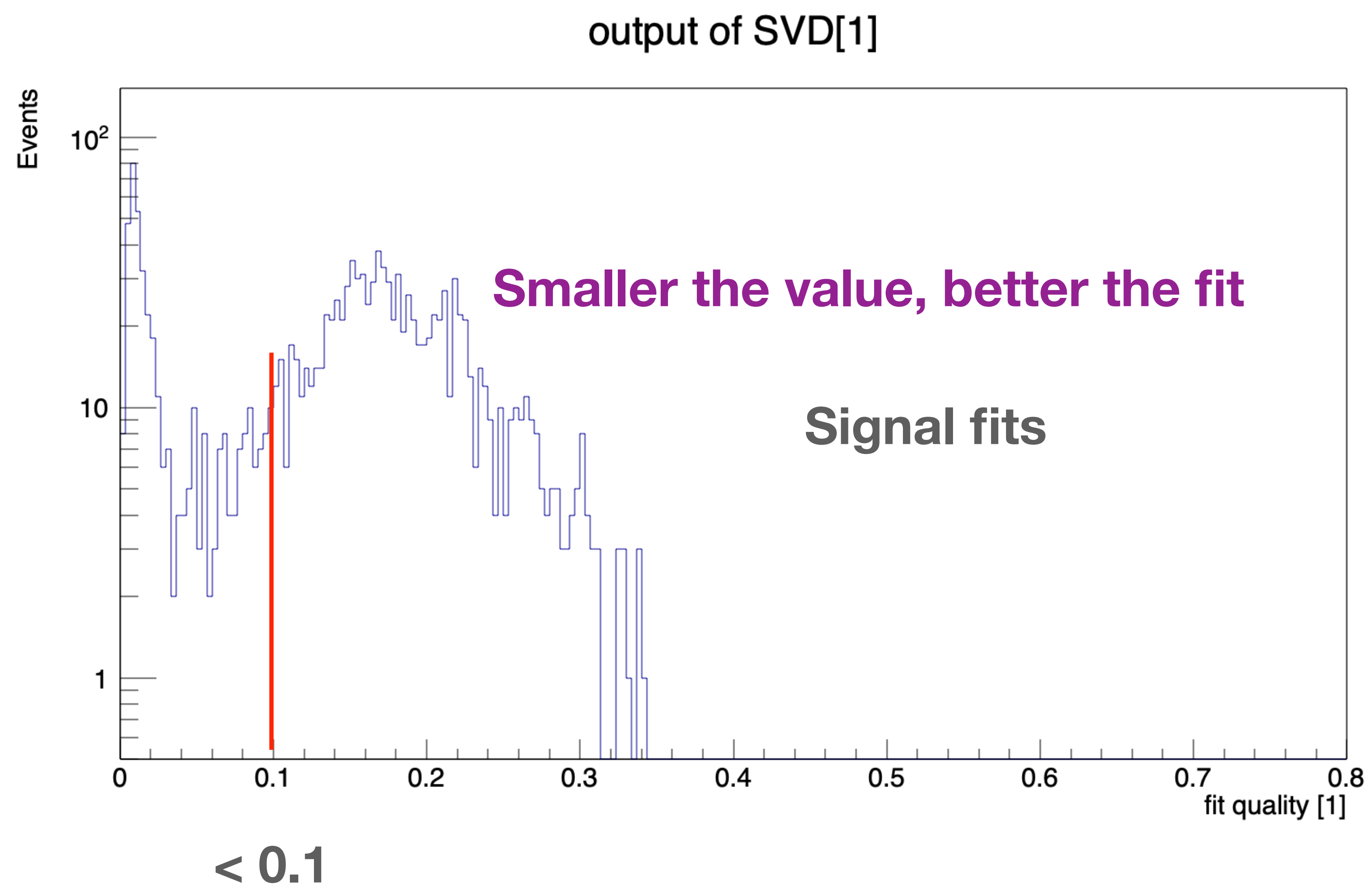


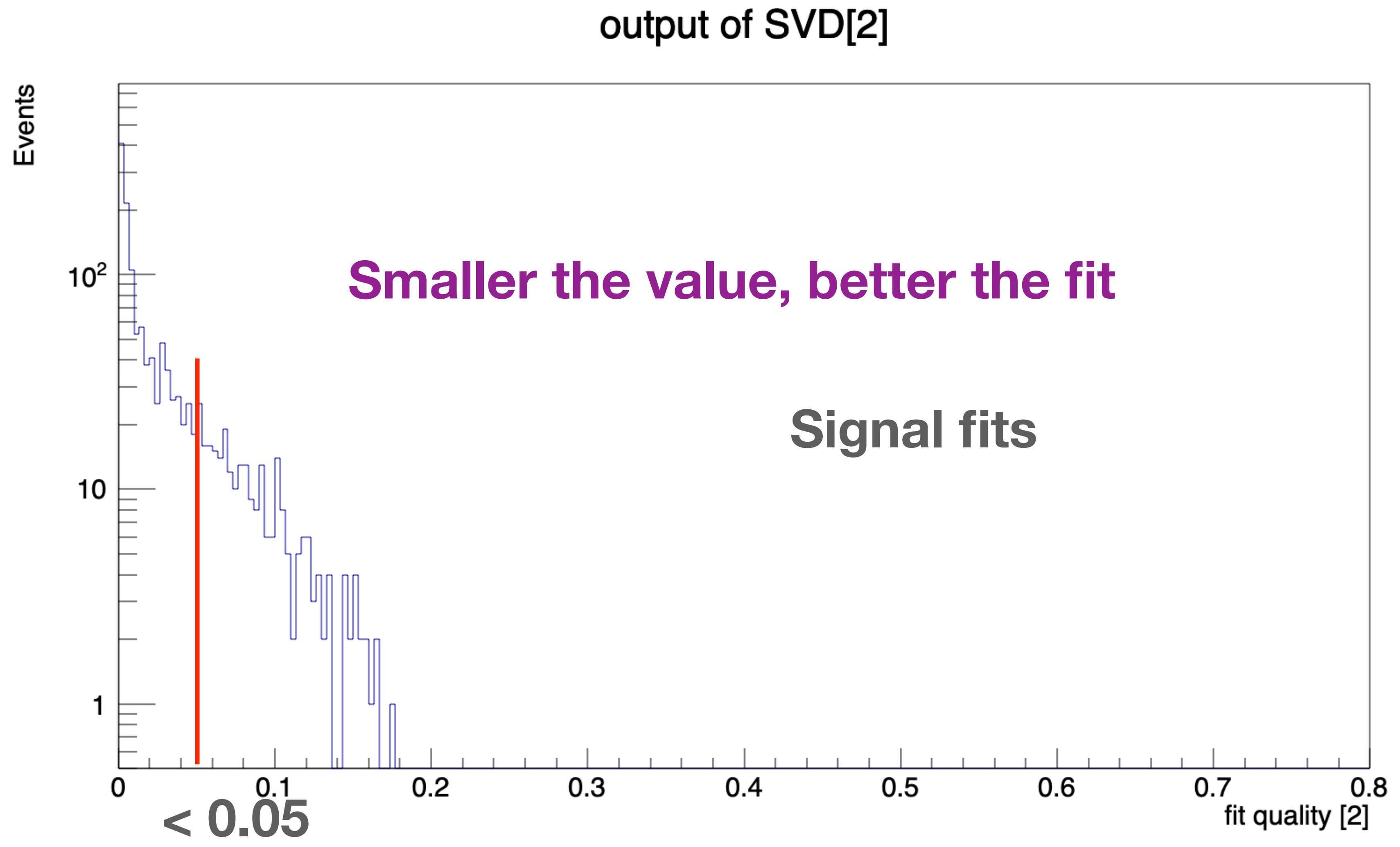
Track fitting: SVD algorithm

- Another way to reduce the ghost is to make some track fit.
  - We use Singular Value Decomposition (SVD) algorithm to fit, and then take the best track.



- The smallest singular value corresponds to the solution of linear least square fitting problem.
- We take fit quality [1] < 0.1 and fit quality [2] < 0.05 as the optimum cuts

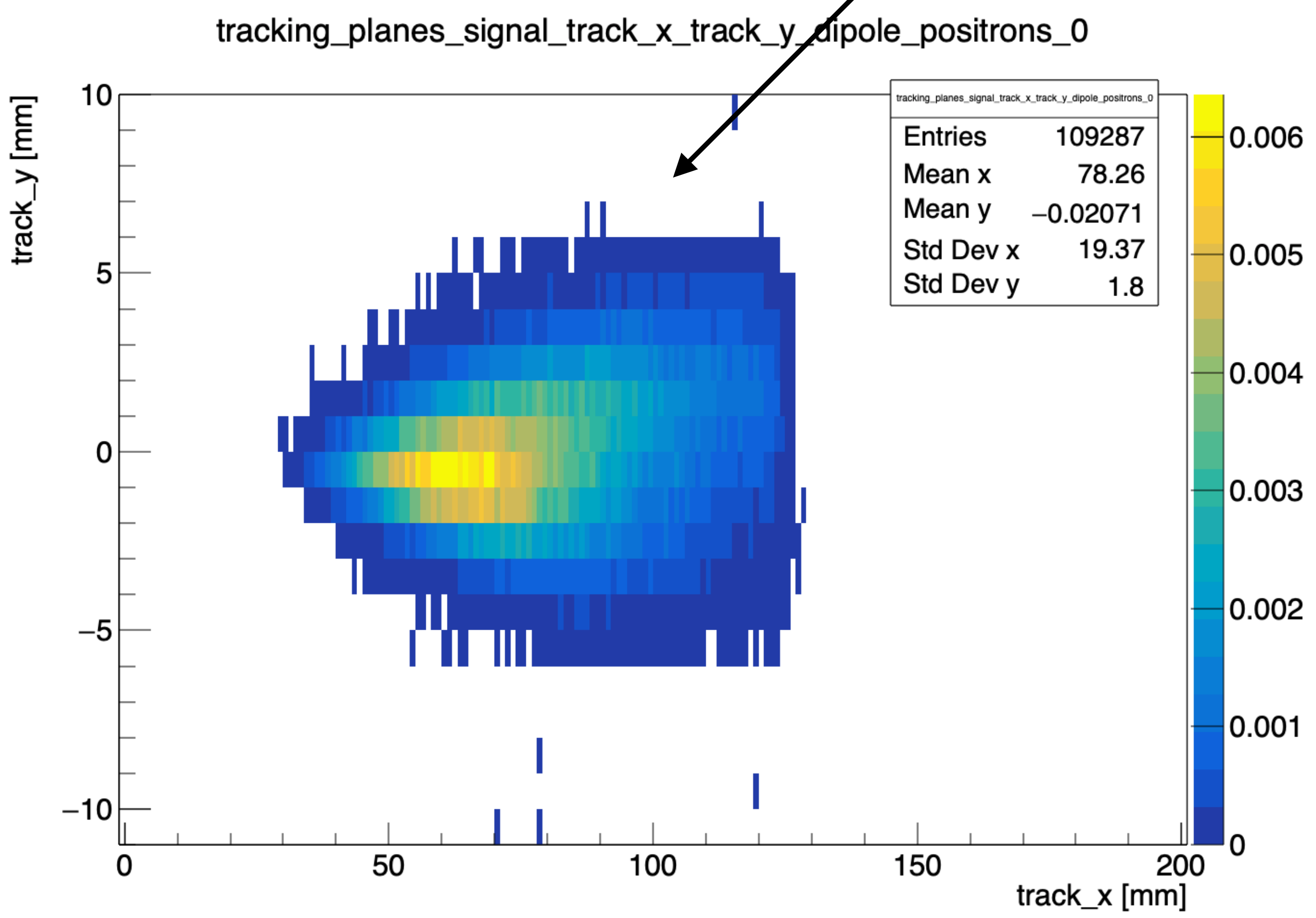




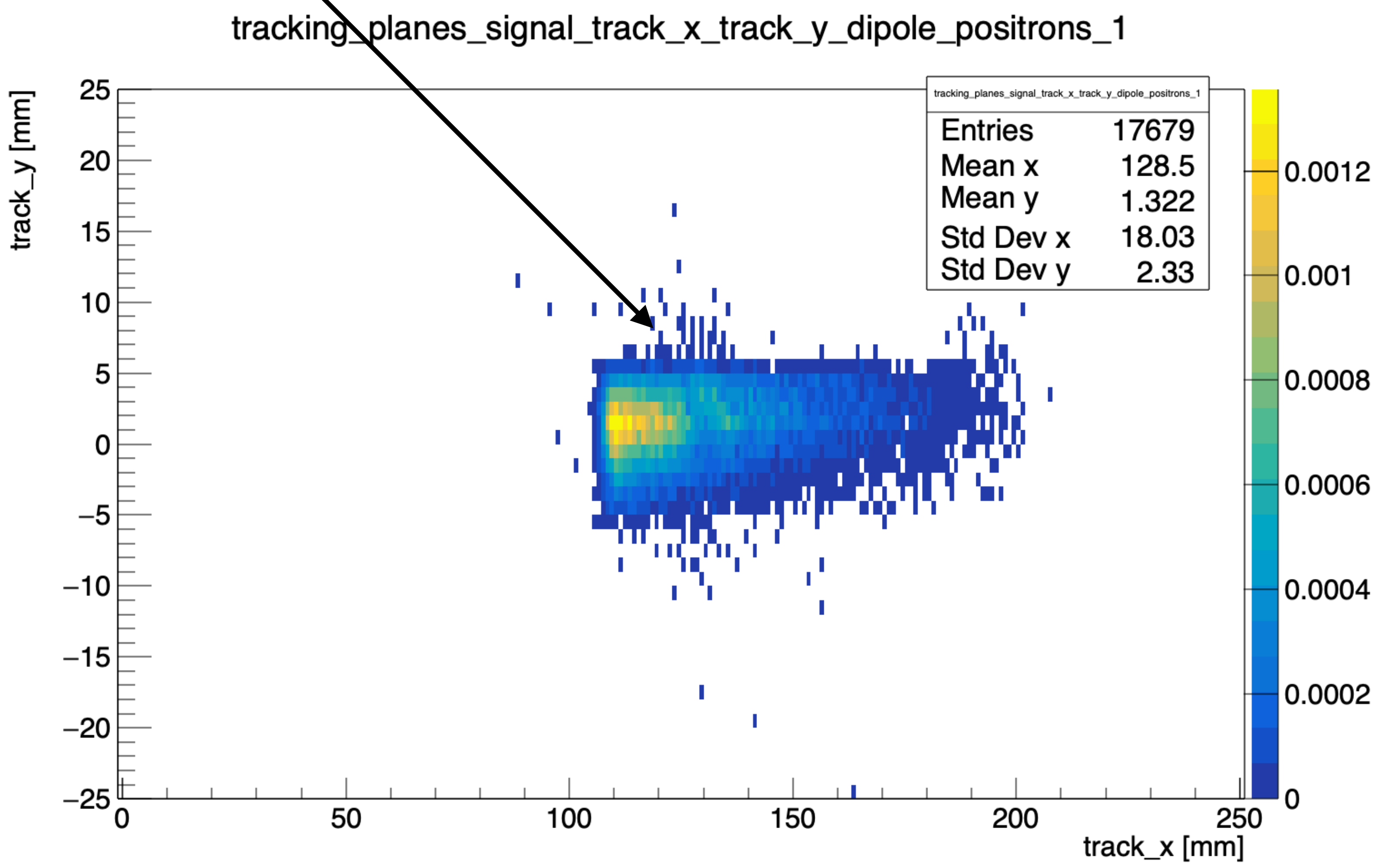
The seeding algorithm (Updated, with loose cuts)

- Keep unique set of tracks from first layer and last layer of tracker.
    - Overlap region is removed (by cutting on x value of the tracks) from
      the outer stave of first layer (innermost) and from the inner stave of
      the last layer (outermost).
- Loop over all pairs in layers 4 and 1, now only positron side (x>0).
- Reject pair of clusters if
    - |x1| > |x4| or they have different sign
    - |z1| == |z4|
    - |y_exit| > 10 mm
    - |x_exit| < 5 mm and |x_exit| > 200 mm
    - If not one cluster in the road of 130 um connecting vector r1 and r4 in both layer 2 and layer 3.
    - The seed energy is greater than 20.0 GeV.
        - The seed energy is calculated from the track.
    - Apply the SVD fit parameter cut

P side Signal at the Dipole exit (New signal)

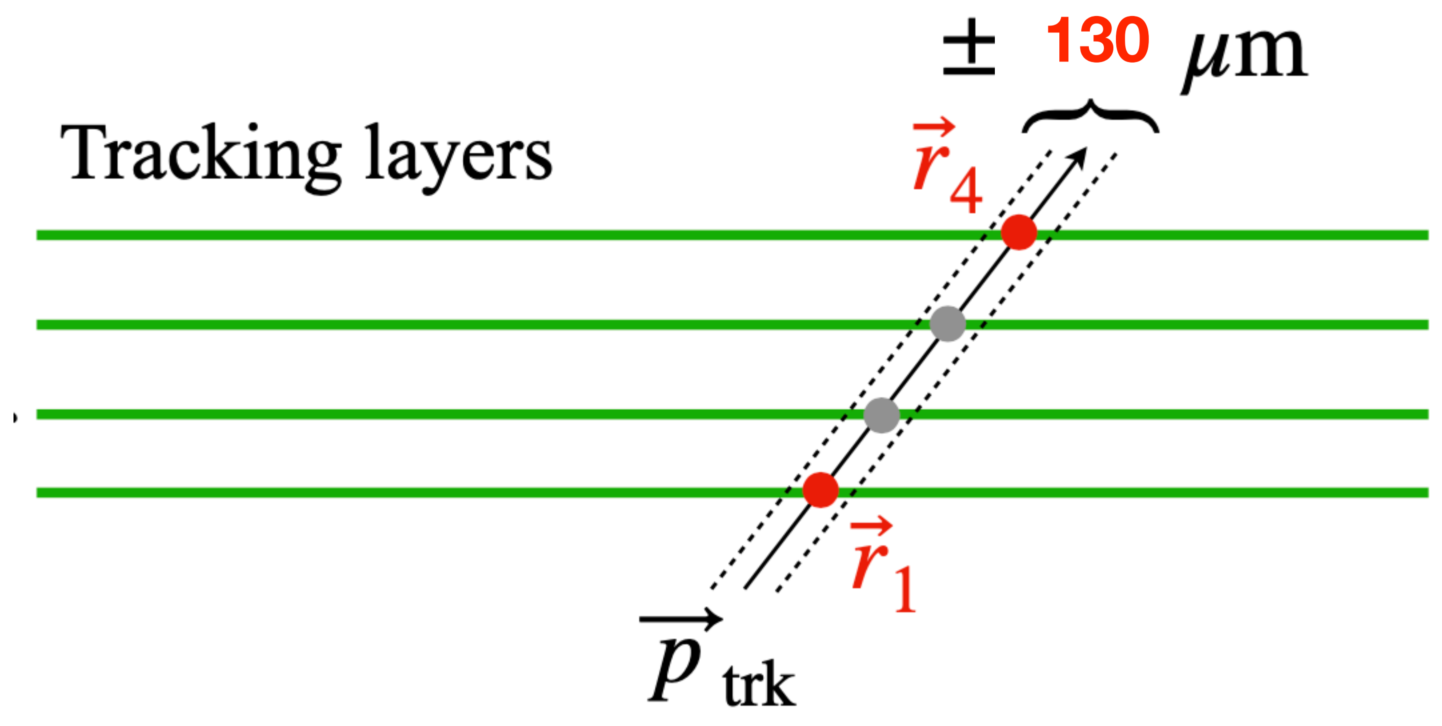


Stave 0



Stave 1

Changes are in red



Seed energy plots (Updated)

- Improvement over previous algorithm for high signal track scenario.

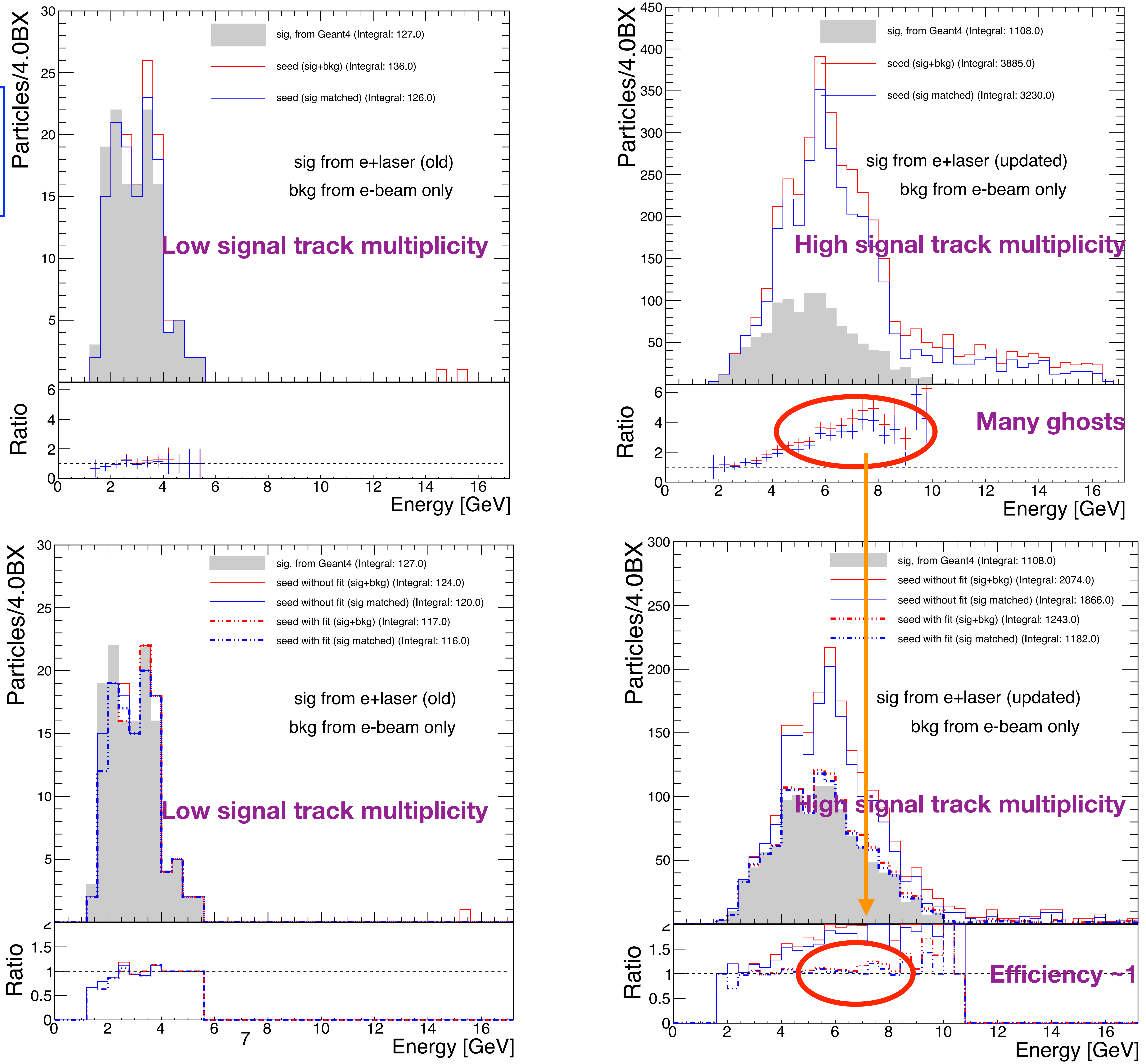
- Two more histograms added (dashed lines):
  - After applying the SVD fit cut

- While loosening of the seed selection window increases the seed tracks, application of SVD fit makes things better.

- There are much room to improve the algorithm, one thing is to apply the Kalman filter in place of SVD fit.

Old seeding algorithm (same as slide 4)

New seeding algorithm + SVD fit



Summary and to do:

		Actual signal tracks	Require hits in 3 layers	Ratio over actual signal tracks, 3 layer hits	sig+bkg tracks, Old algorithm, 4 layer hits	Ratio over actual signal tracks, 4 layer hits	sig+bkg tracks, new algorithm, 4 layer hits	Ratio over actual signal tracks, 4 layer hits	sig+bkg tracks, new algorithm +SVD fit, 4 layer hits	Ratio over actual signal tracks, 4 layer hits
	Low signal multiplicity	127	352	277.17%	136	107.09%	124	97.64%	117	92.13%
Present scenario	High signal multiplicity	1108	12109	1092.87%	3885	350.63%	2074	187.18% %	1243	112.18%

- Looked at the signal and background mixture per BX for the first time.
- Played with the seeding algorithm: Added an SVD track fit to improve the scenario.
- For high signal track multiplicity, the algorithm+fit efficiency is good.
- This will be repeated for other spot sizes once the fixed signal samples arrive.
- Clustering from hits will be added for the seeding, which will reduce the background.
- Lots of room for improvement
    - Kalman filter.
    - Remove ambiguity from the signal tracks.

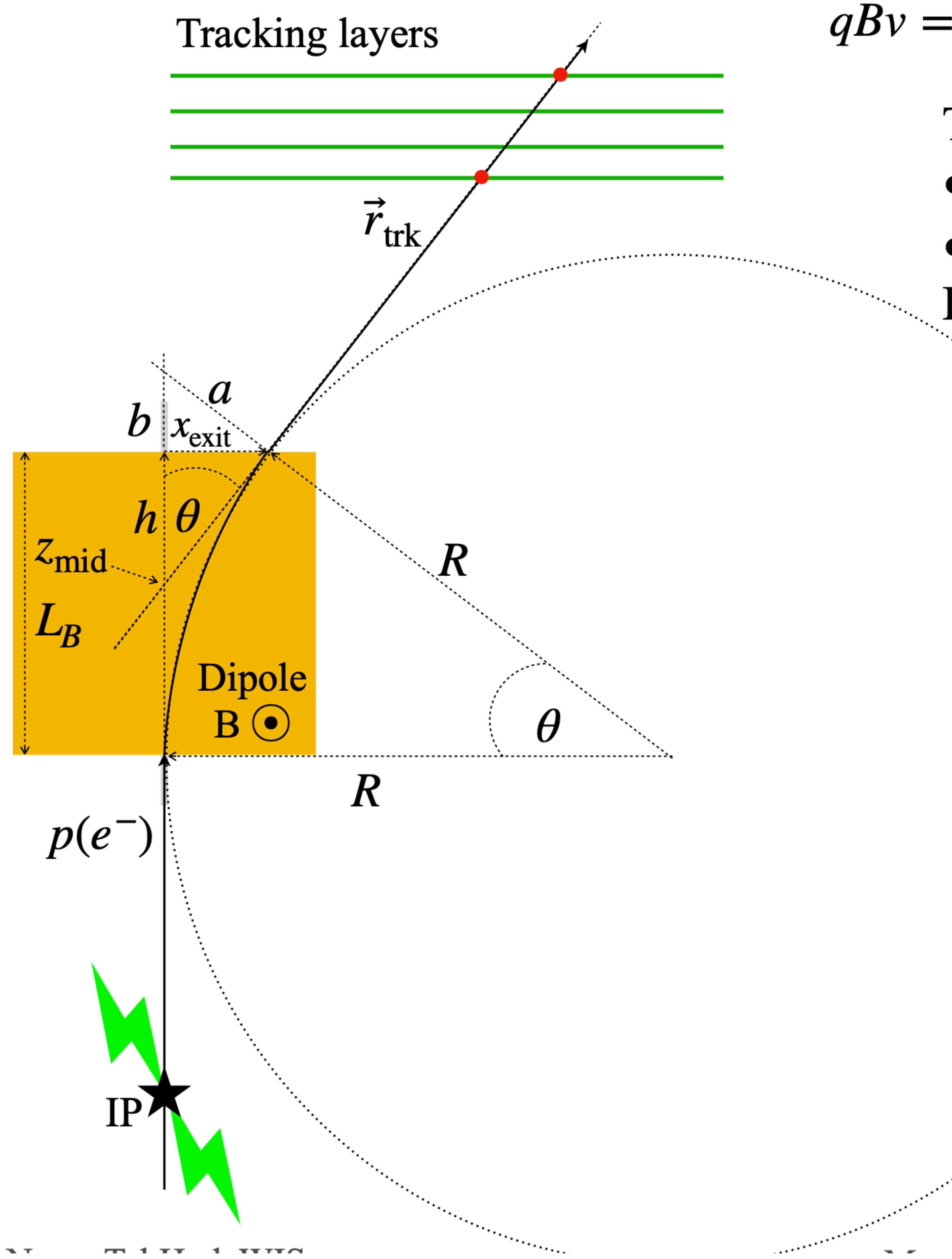
Bonus slides

SVD track fitting

- A common application of the singular value decomposition (SVD) is in fitting solutions to linear equations.
- Suppose we collect (x_i, y_i) data, which can be fit to some linear homogeneous equation $ax_i + by_i - c = 0$. If we have our data in a matrix A, and the coefficients in some vector v, we can write this problem as $Av = 0$, where we'd like to figure out what v is given A
- We usually have more data than coefficients, in which case we can't solve this
- However, we can fit v to A in order to minimise the value of Av via SVD
- Suppose we take the singular value decomposition of $A \rightarrow A = U\Sigma V^{\mathrm{T}}$
- If one of the singular values is 0 then we have an exact solution
- If none of the singular values are zero, we have no exact solution
- However, it can be shown that the smallest singular value corresponds to the solution of the linear least squares fitting problem
- Namely, if we want to find the least squares fit to the data, we need to look at the smallest (usually the last) singular value and read the respective column of V to get the best fit coefficients (a, b, c)

From Noam

Electron in a B-field of a dipole



$$qBv = \frac{mv^2}{R} \longrightarrow p = qBR \longrightarrow p[\text{GeV}] = 0.3 \cdot B[\text{T}] \cdot R[m]$$

To get R, need to extrapolate the track backwards to

- the dipole exit plane and obtain the $x_{\rm exit}$ coordinate
- then intercept with the $x = 0$ line to find $z_{\rm mid} \to h$

Extrapolation is done using 2 points at layer 4 and 1

$$\tan\theta = \frac{x_{\rm exit}}{h} = \frac{L_B + b}{R} \longrightarrow R = h\frac{L_B + b}{x_{\rm exit}}$$

$$\tan\theta = \frac{b}{x_{\rm exit}} \longrightarrow R = h\frac{L_B + x_{\rm exit}\tan\theta}{x_{\rm exit}}$$

$$R = h\frac{L_B + x_{\rm exit}^2/h}{x_{\rm exit}} = \frac{hL_B}{x_{\rm exit}} + x_{\rm exit}$$

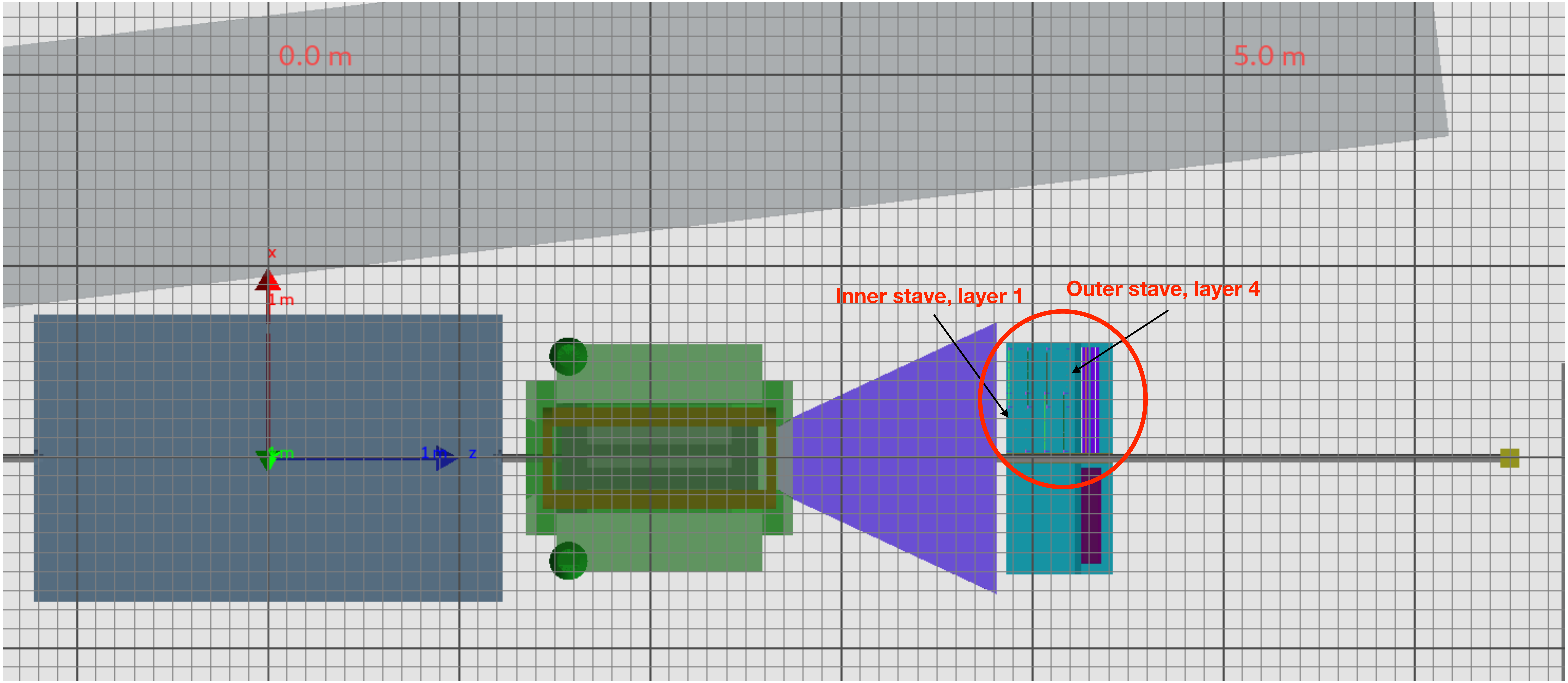
$$p_{\rm seed} = 0.3B \cdot \left(\frac{hL_B}{x_{\rm exit}} + x_{\rm exit}\right)$$

$$\vec{p}_{\rm seed} = \left(\sim 0,\ p_{\rm seed}\frac{y_{\rm trk}}{r_{\rm trk}},\ p_{\rm seed}\frac{z_{\rm trk}}{r_{\rm trk}}\right)$$

Random number drawn from the truth p_x distribution at the vertex

From Noam

The subsystems near the IP



Plots from Sasha