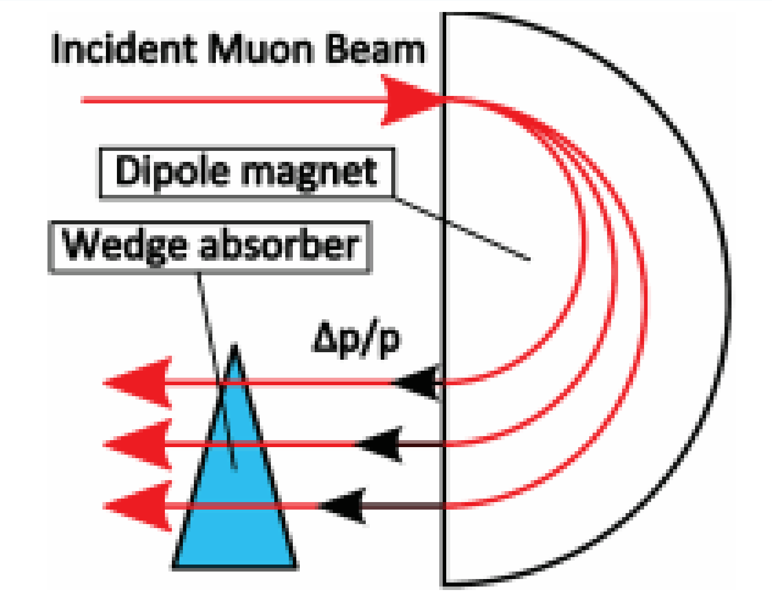
Emittance Exchange in MICE

Craig Brown

Brunel University

24 October 2019

Aims



- Demonstrate Emittance Exchange and Reverse Emittance Exchange in the Wedge using MICE data
- Emittance Exchange can be demonstrated by looking at the change in phase space density of the particle selection before and after having passed through a Wedge absorber
- Emittance Exchange is shown by a decreased transverse phase space density (x, px, y, py) and increased longitudinal phase space density (z, pz), (and vice versa for Reverse Emittance Exchange)
- Can use a number of techniques to calculate phase space density: KDE, KNN, Voronoi Tessellations, etc.
- MICE beam only has a small natural dispersion → Use beam reweighing techniques to select beams with desired dispersion

CM54

- Why do the phase space densities change depending on sample selection?
- Why do I get agreement with Francois for the upstream section but not the downstream section when comparing the 9th percentiles?
- Main question from CM54 to answer: Is KDE a poor estimator?
- Does KDE behave differently from KNN, or other density estimators?

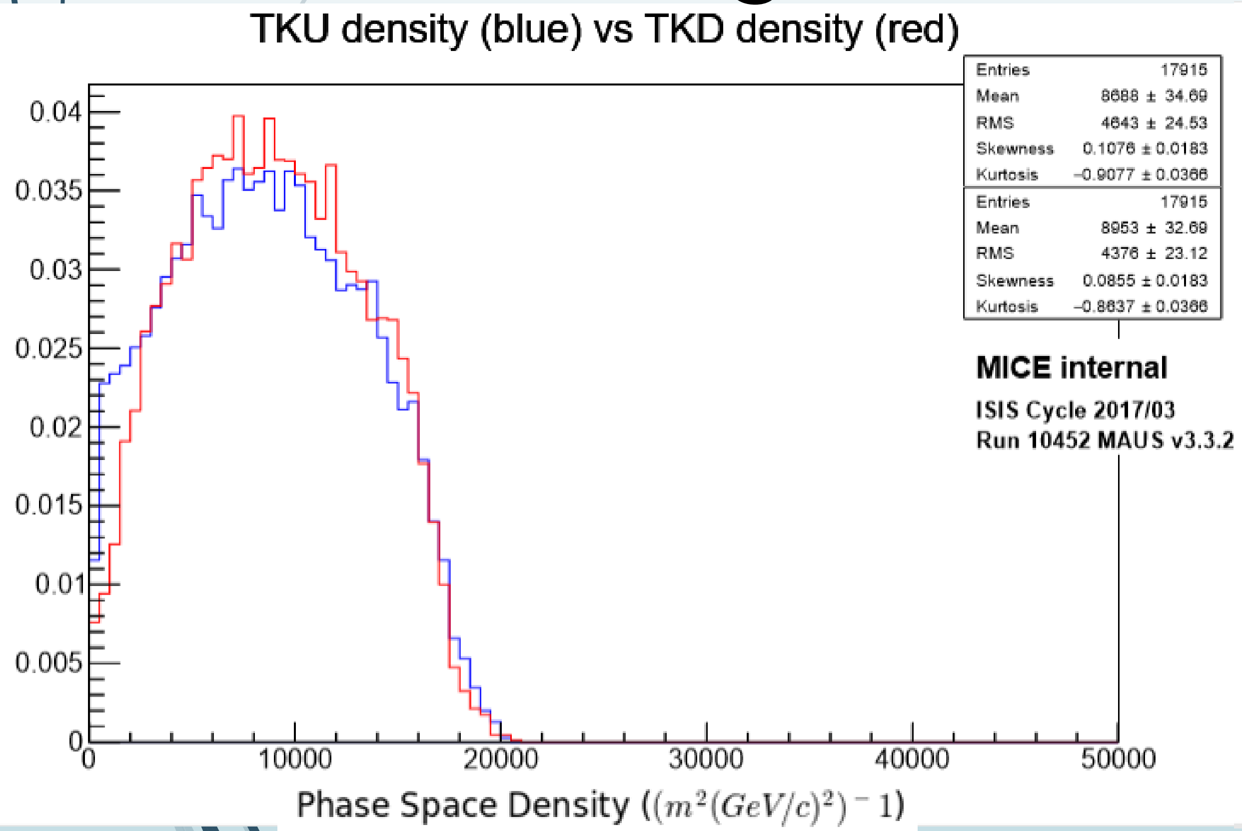
Previously:
Particle Selection – 4D transverse

- Will look at a number of selections for when the wedge is present/absent and see the advantages/disadvantages of selection cuts
- All will include:
  - TOF01 cut
  - Radius cut < 150 mm
  - Momentum cut 130 -150 MeV/c
  - Single track in the Upstream Tracker and a single track in the Downstream Tracker
- Will compare this cut with the selection for when there is an Upstream Track but no Downstream Track, to look at selection bias.

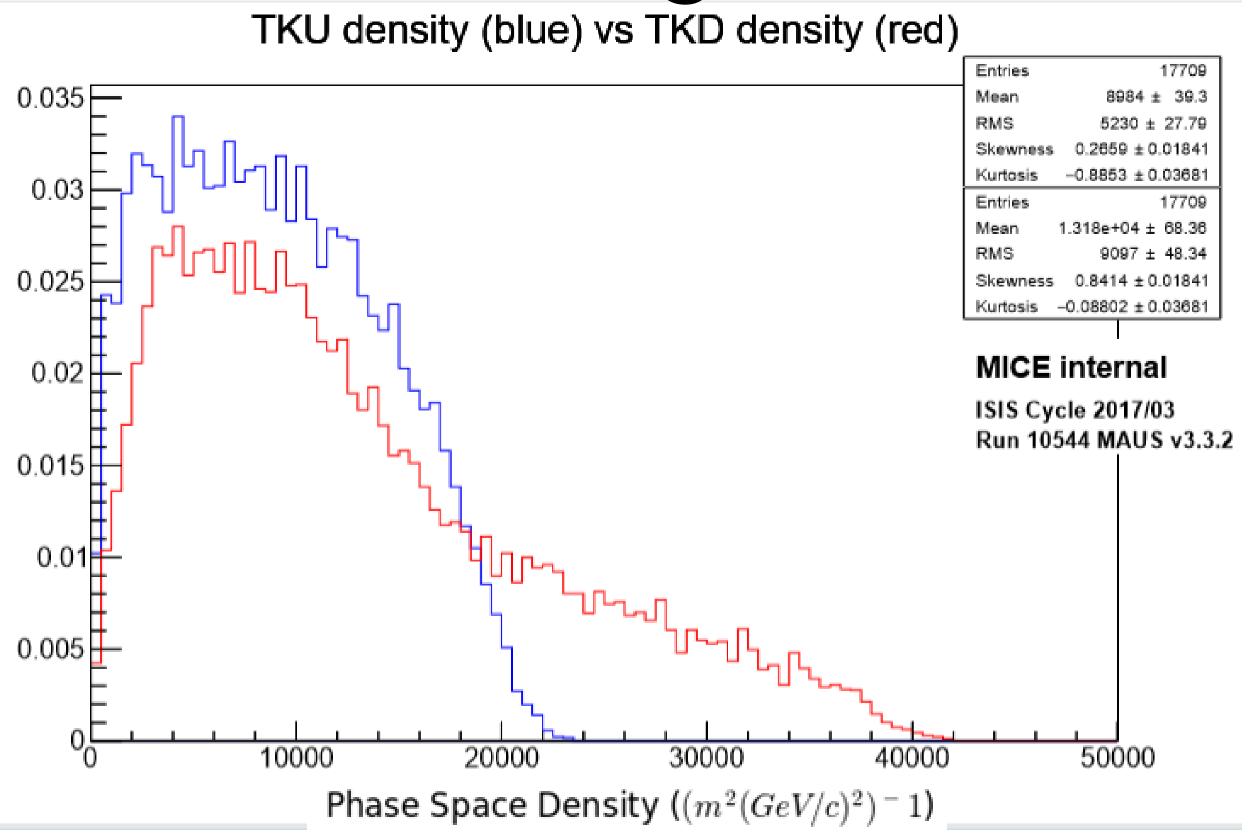
10-140 4D Transverse phase space density

- Single track that has gone both through Upstream and Downstream Tracker

No Wedge



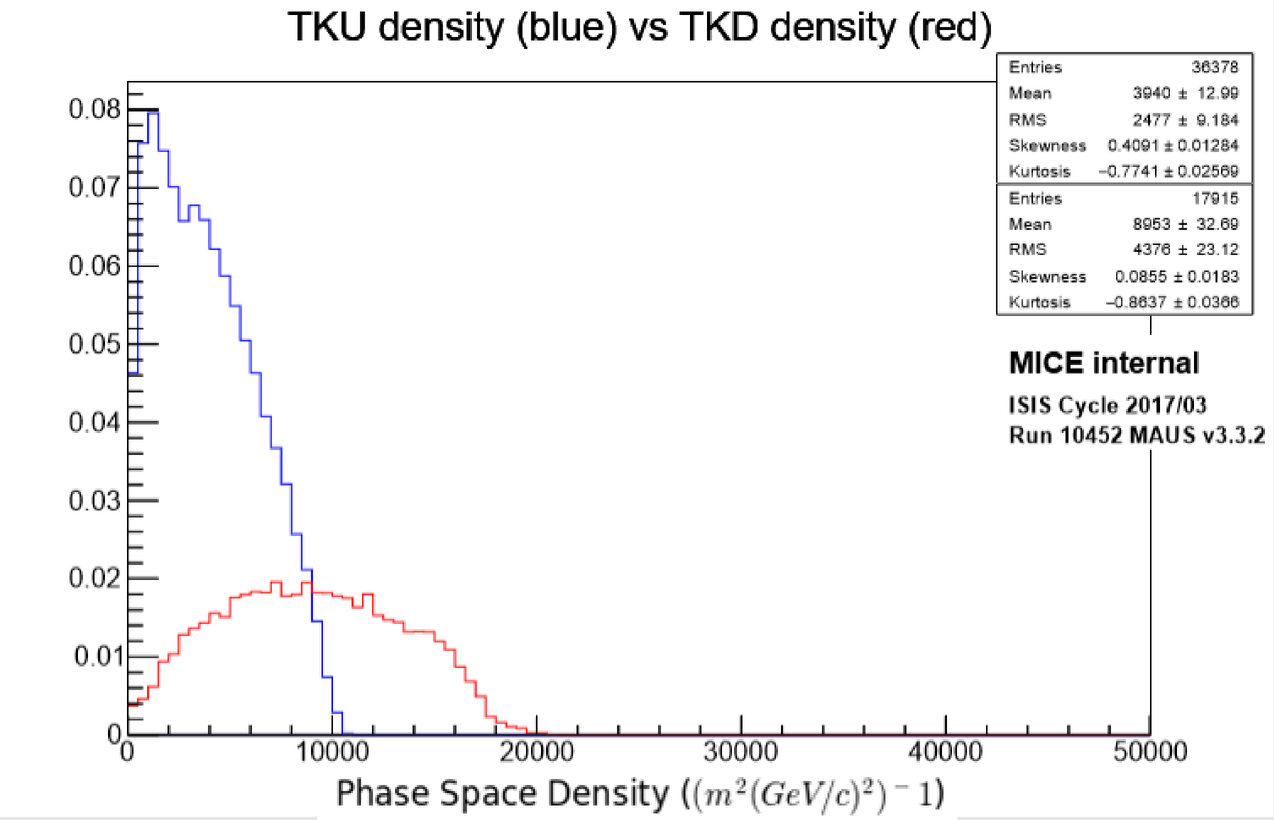
Wedge



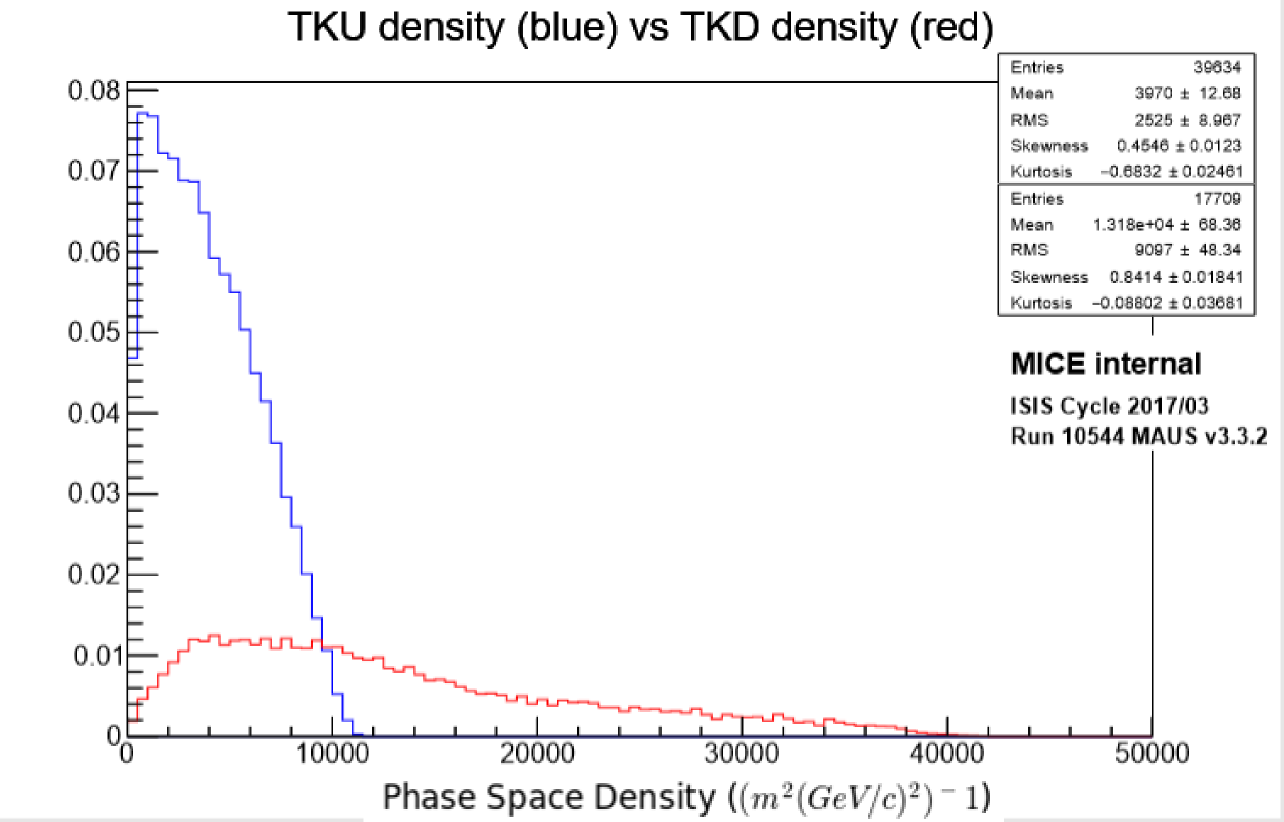
10-140 4D Transverse phase space density

- Single track that has gone both through Upstream and Downstream Tracker
- And single track that has only gone through Upstream Tracker

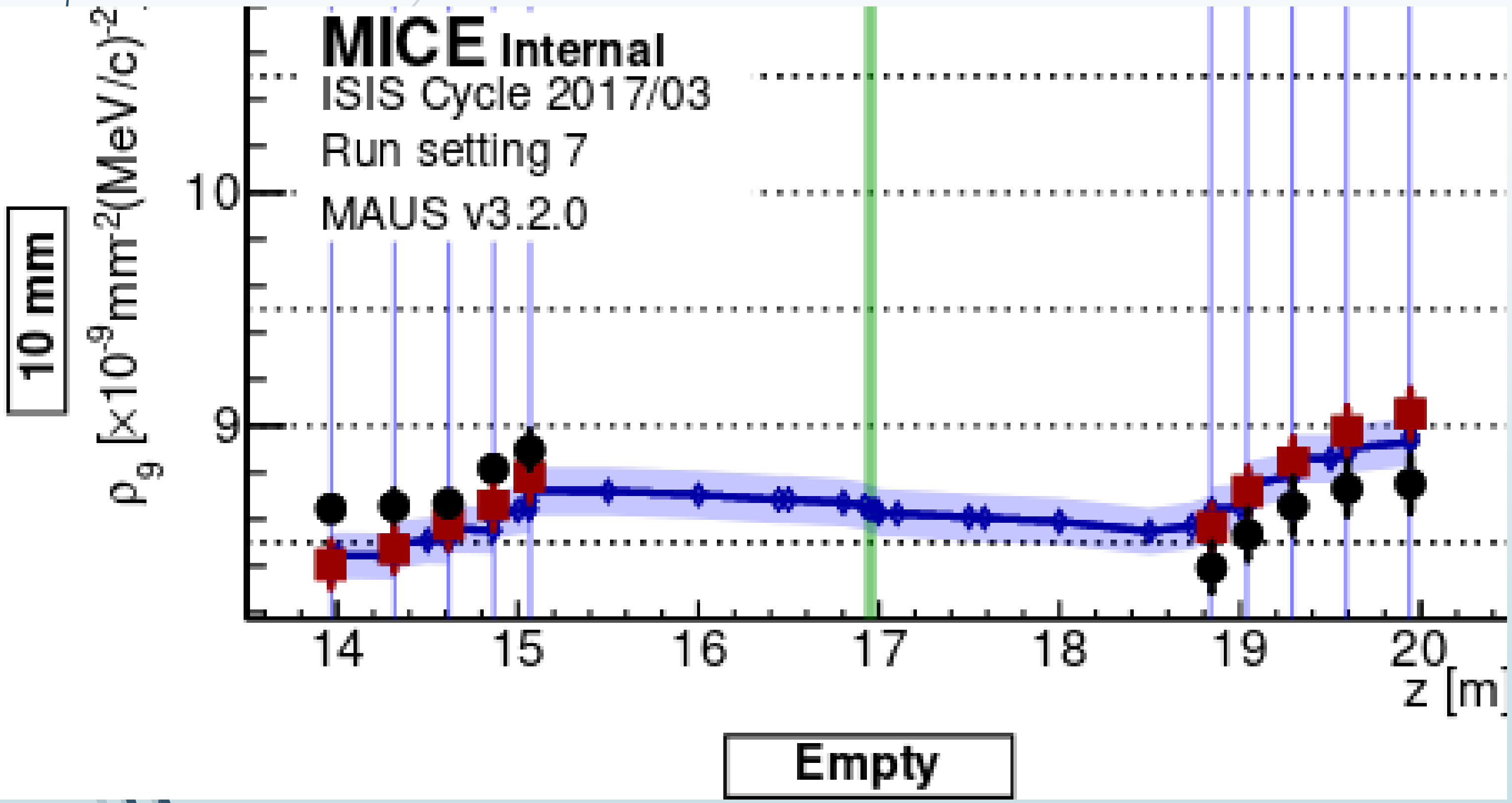
No Wedge



Wedge

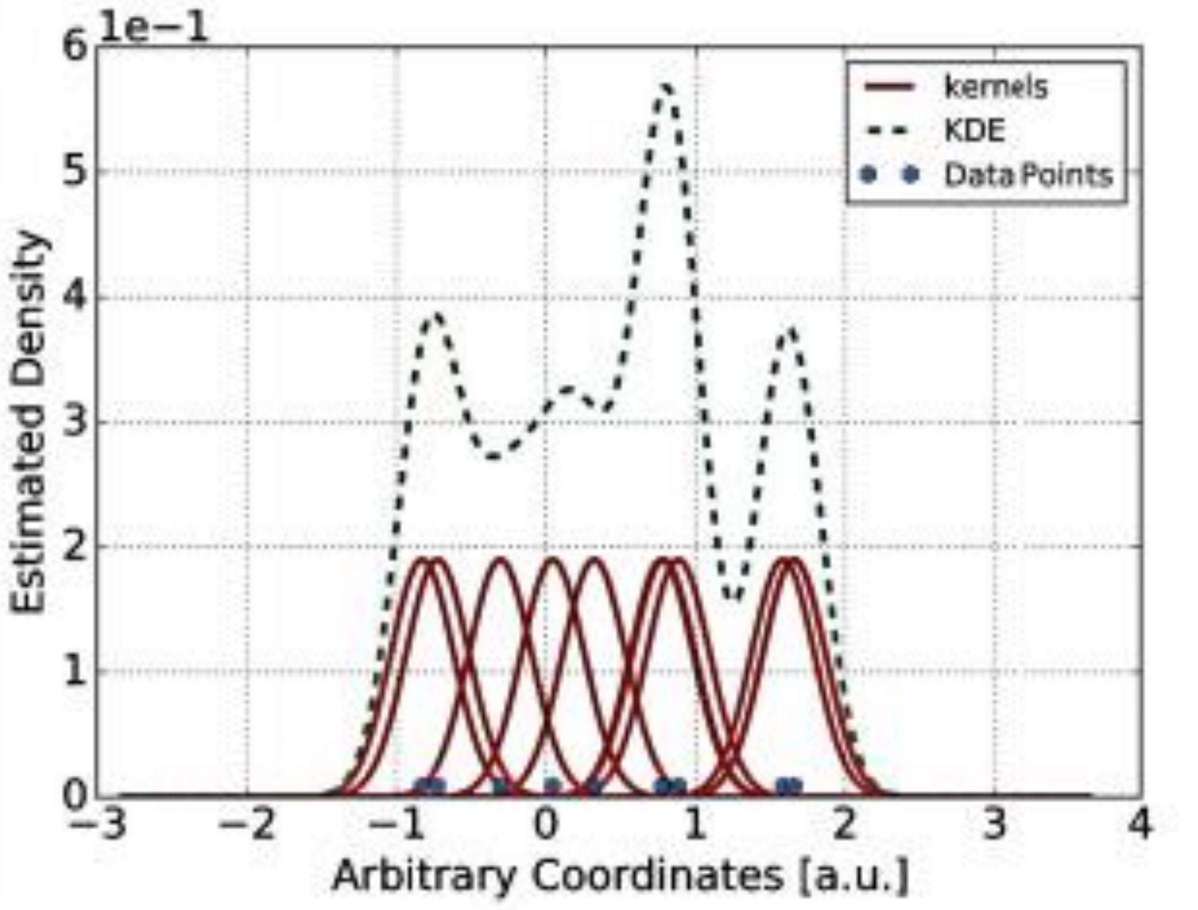


Is KDE a poor Estimator?



- Produces very different results depending on input selection
- It does show agreement with Francois' KNN estimate for the 9th percentile of the no absorber upstream sample where the data is comparable (bar for any small differences in magnetic fields)
- Why the different results?

Kernel Density Estimation



- Let each data point be represented by a kernel (e.g gaussian, uniform, epanechnikov)
- Sum of over all kernels to obtain the particle distribution

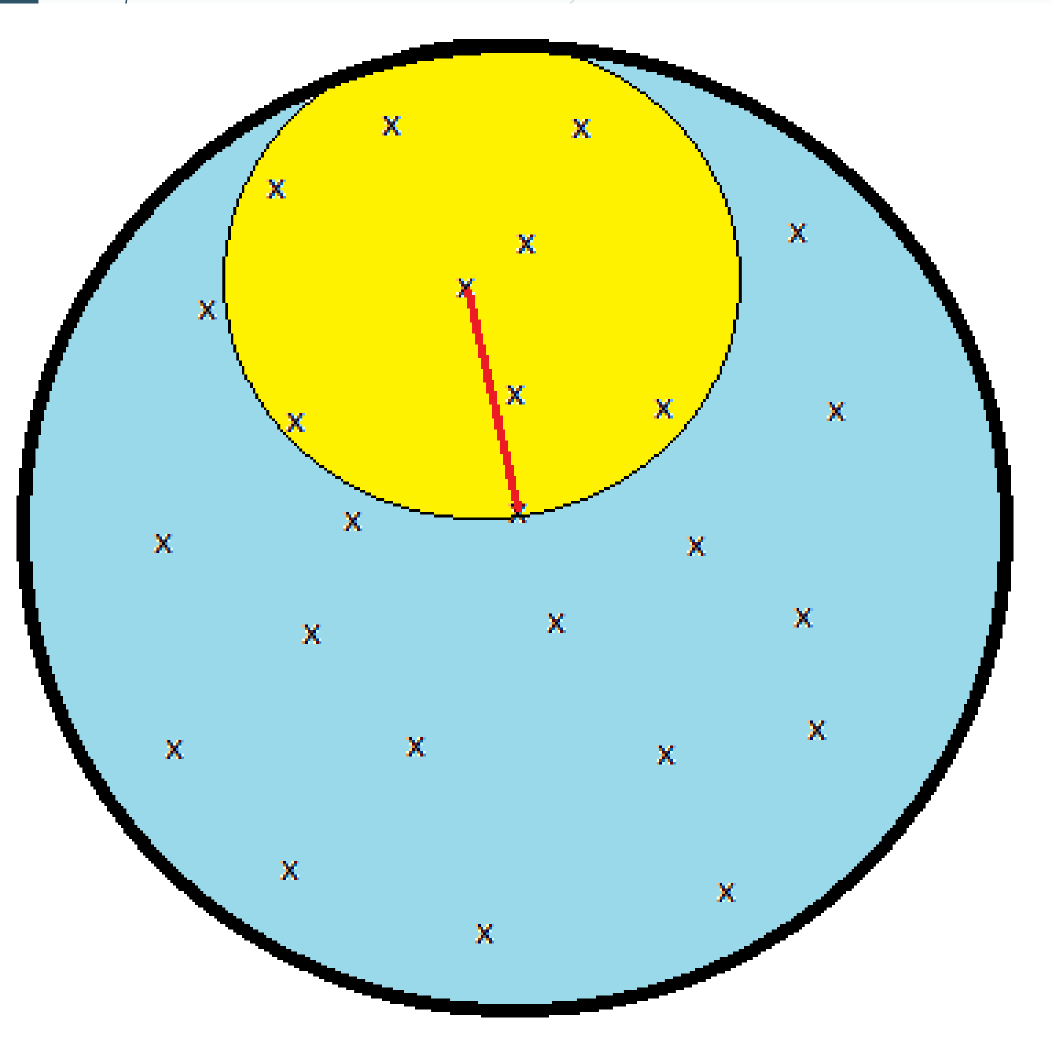
- The density estimate is obtained by applying:

$$\hat{\rho}(\vec{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\vec{x} - \vec{X}_i}{h}\right)$$

where K is the kernel choice, h is the bandwidth of the kernel, d is the dimension and n is the number of points.

For higher dimensions the covariance matrix is used to equally weight the bandwidth choice for each dimension

K-Nearest Neighbour



In KNN, for a particular choice of k, a bounding volume is placed around each particle. The bounding volume for each particle with that particle at its centre is the smallest volume to encompass k particles. The density for each particle can then be found by applying:

$$\vec{\rho}(x) = \frac{k}{n\kappa_d R_k^d} = \frac{k\Gamma(\frac{d}{2}+1)}{n\pi^{\frac{d}{2}} R_k^d}$$

where $d_k(t)$ is the Euclidean distance $R_i = ||\vec{x} - \vec{x_i}|| = \sqrt{(\vec{x} - \vec{x_i})^T(\vec{x} - \vec{x_i})}$, κ_d is the volume of a unit d-ball (in 1-D it is equal to two), $\Gamma(\frac{d}{2}+1)$ is Euler's gamma function.

For higher dimensions the covariance matrix is used to equally weight the distance for each dimension

Performance of density Estimators

- The Mean Integrated Squared Error (MISE) can describe the accuracy of $\vec{f}(t)$ as an estimator of f.
- $MISE(\vec{f}) = E\int\{\vec{f}(x) - f(x)\}^2 dx = \int\{E\vec{f}(x) - f(x)\}^2 dx + \int Var(\vec{f}(x))dx$
- $MISE(\vec{f}) = \int\left[Bias\left(\vec{f}(x)\right)^2 + Var(\vec{f}(x))\right]dx$
- As $\vec{f}(x) = \sum_{i=1}^{n}\vec{f}(x) \sim \frac{k}{r^d}$, the optimal choice of k is determined by a trade-off between the variance and the squared bias.

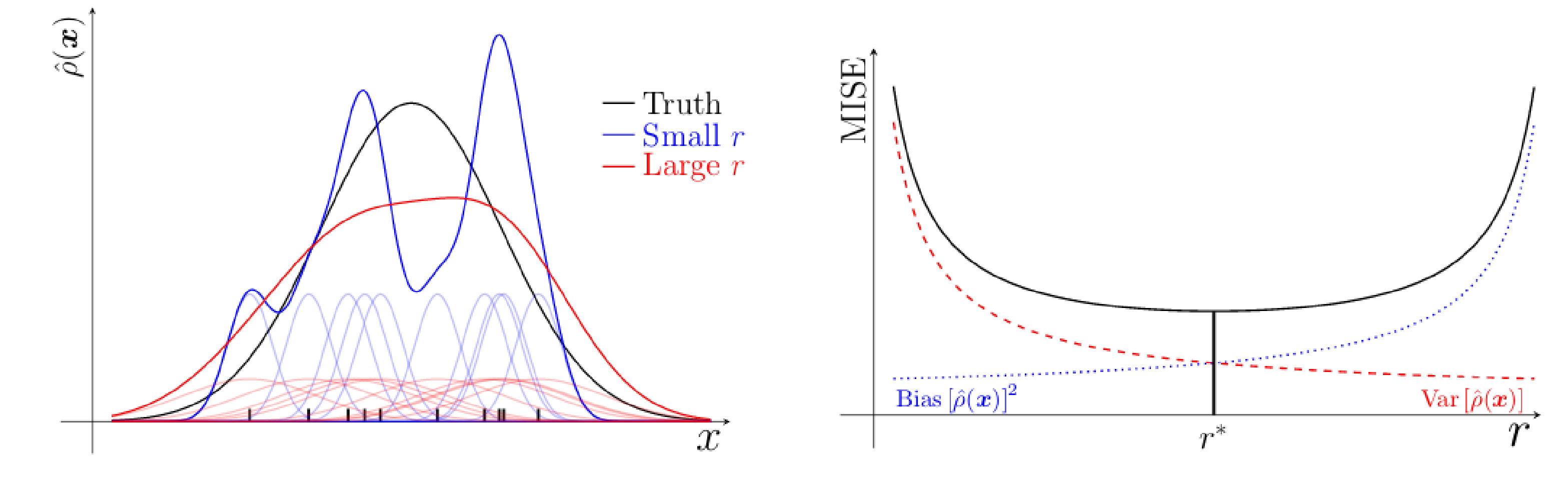


Figure 7.1: (Left) Illustration of the effect of the smoothing radius, r, on the behaviour of nonparametric density estimators. (Right) Schematic of the evolution of the bias, variance and MISE as a function of the smoothing radius, r.

For small r the estimate follows the data closely as its not biased but has a very large variance due to fluctuations.

For large r the estimate varies little as it is less sensitive to fluctuations, but becomes more biased.

Bias and Variance

- For KNN:

$$Bias\left[\vec{f}(x)\right] \cong \frac{\mu_2(w)\nabla^2 f(x)}{2(\kappa_d f(x))^{\frac{2}{d}}}\left(\frac{k}{n}\right)^2$$

$$Var\left[\vec{f}(x)\right] \cong \frac{f^2(x)}{k}$$

With $\mu_2(w)$ the second moment of the uniform kernel and $\nabla^2 f(x)$ the Laplacian of the density field. The MISE is of order:

$$MISE(k) = \mathcal{O}\left(\left(\frac{k}{n}\right)^{\frac{4}{d}} + \frac{1}{k}\right)$$

Which admits a minimum for a parameter k of order:

$$k \sim n^{-4/(4+d)}$$

The optimal rate of convergence for a KNN estimator is then (to the underlying density):

$$MISE(k) = \mathcal{O}\left(n^{-4/(4+d)}\right)$$

Bias and Variance

- For KDE (with second-order kernels):

$$Bias_h(\vec{x}) \approx \frac{1}{2}h^2\nabla^2 f(\vec{x})\int t_1^2 K(\vec{t})d\vec{t}$$

$$Var[\vec{f}(x)] \approx n^{-1}h^{-d}f(\vec{x})\int K(\vec{t})^2 d\vec{t}$$

The MISE is then approximated by

$$\frac{1}{4}h^4\left\{\int t_1^2 K(\vec{t})d\vec{t}\right\}^2 \int \{\nabla^2 f(\vec{x})\}^2 d\vec{x} + n^{-1}h^{-d}\int K(\vec{t})^2 d\vec{t}$$

The optimal window width to minimize MISE is given by

$$h_{opt}^{d+4} = d\int K(\vec{t})^2 d\vec{t}\left\{\int t_1^2 K(\vec{t})d\vec{t}\right\}^{-2}\left\{\int \{\nabla^2 f(\vec{x})\}^2 d\vec{x}\right\}^{-1} n^{-1}$$

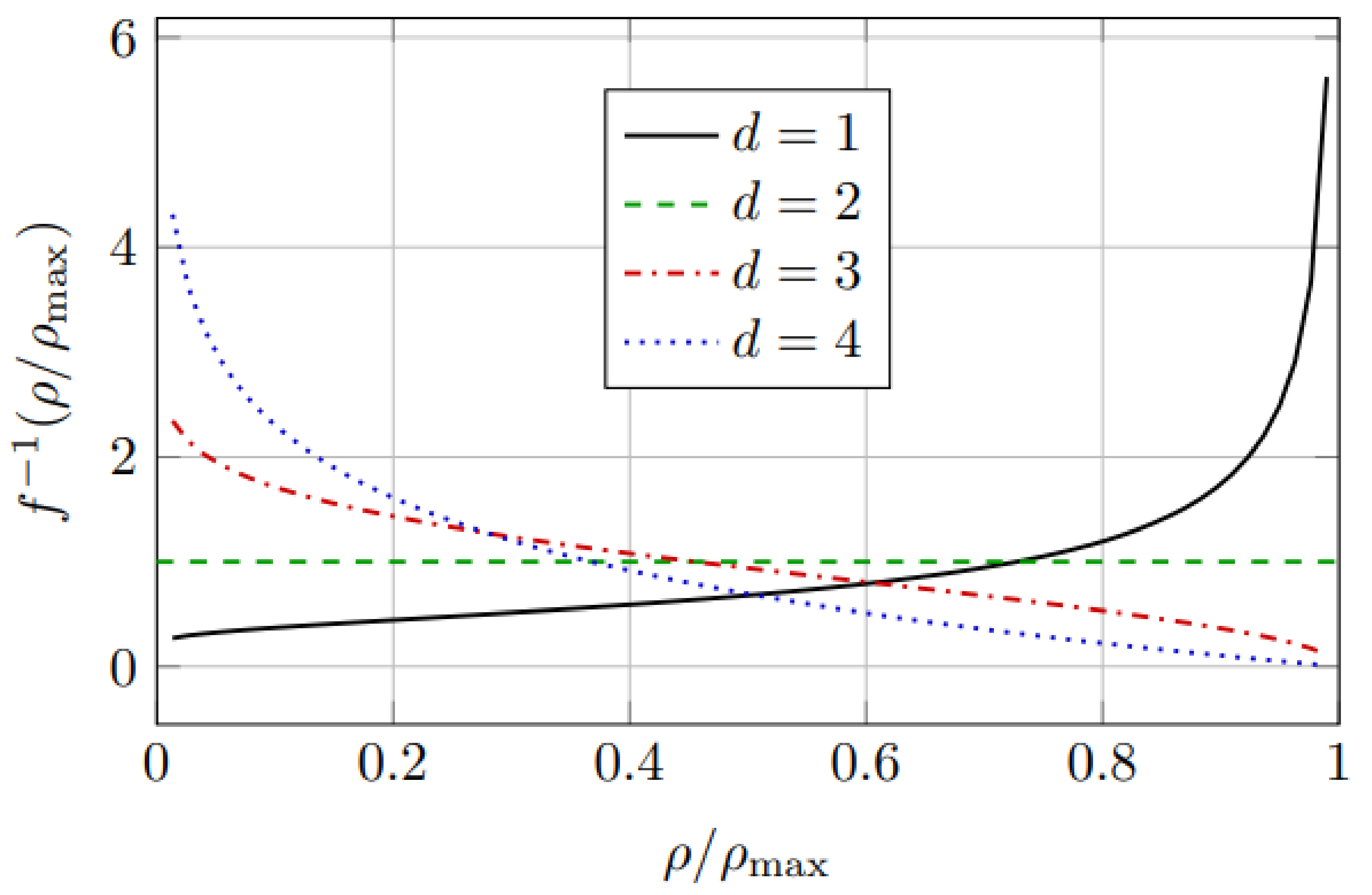
Therefore

$$MISE(h) = \mathcal{O}\left(n^{-4/(4+d)}\right)$$

This is same as for KNN, that is the rate of convergence to the density estimate is the same for KNN and KDE. The rate of convergence for the histogram is given by

$$MISE(\Delta) = \mathcal{O}\left(n^{-2/(2+d)}\right)$$

Phase Space Density to expect



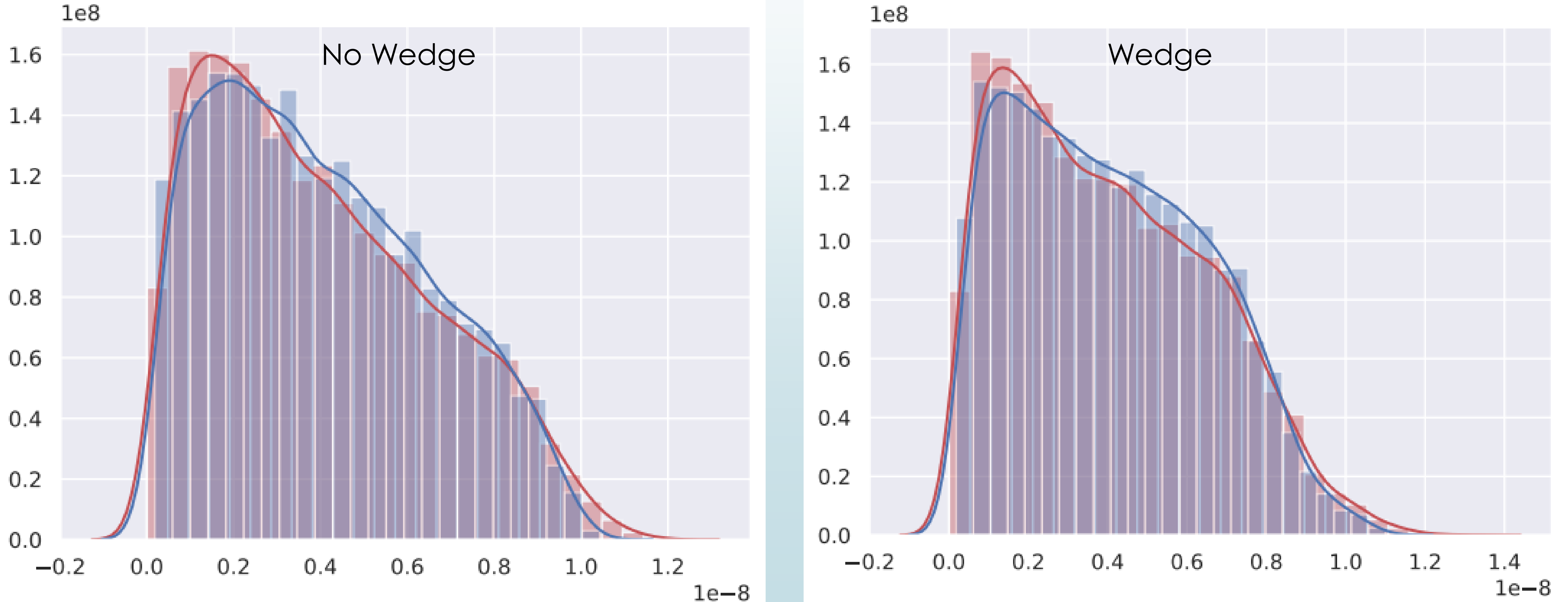
- For a Gaussian beam in one dimension, most particles will be found in the centre and are therefore at a high phase-space density
- As the dimension size increases, another variable is added to the phase-space density calculation
- It requires only one of the dimensions to be in a low phase space density to shift the global phase space density of the particle to a lower phase-space density
- As the dimension size increases most particles will be found at a lower phase space density

KDE vs KNN

Full Upstream Sample

Blue – KDE Red – KNN

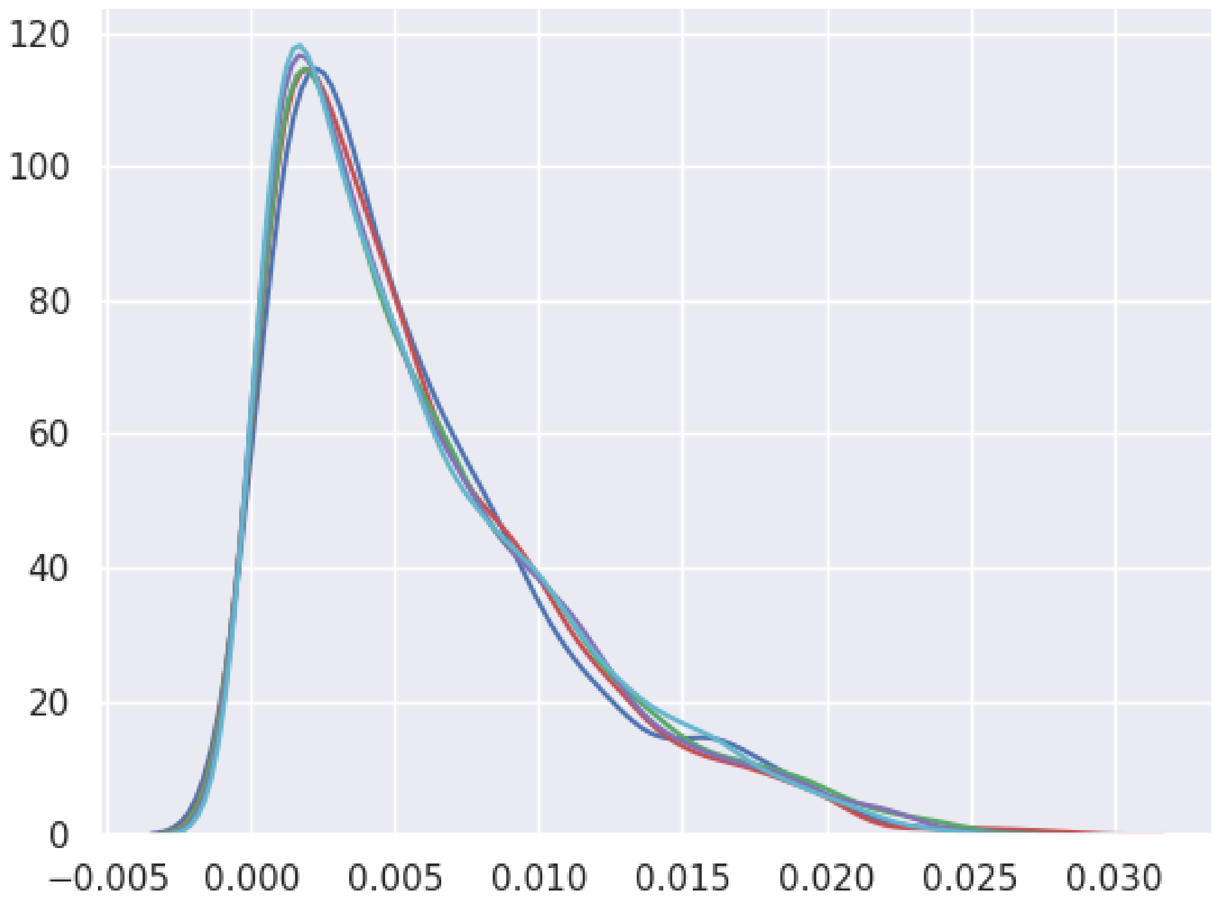
Slight differences due to KDE convolving the density with the kernel, while for KNN it has been smoothed to ensure area of graph is one



KNN and KDE behave the same

- Updated KNN routine to use KDTree algorithm approach to speed up calculation. Updated k-selection convention to match Francois' (i.e. whether to include test point, only minor difference)
- Checked changing the sample size(n), the k-value and changing one parameter while keeping the other constant. This resulted similar phase space densities, with differences due to increase in error due to choosing non-optimal parameters.
- Why do I get different phase-space densities?

Change in Sample Size – Toy Scenario



- See effect of change in sample size, as sample size increases, should approach underlying density of sample
- Random 4-D distribution with mean = 0, Standard Deviation = diag(1,1,1,1)

Blue: n =1000, k = 31

Red: n = 2000, k = 44

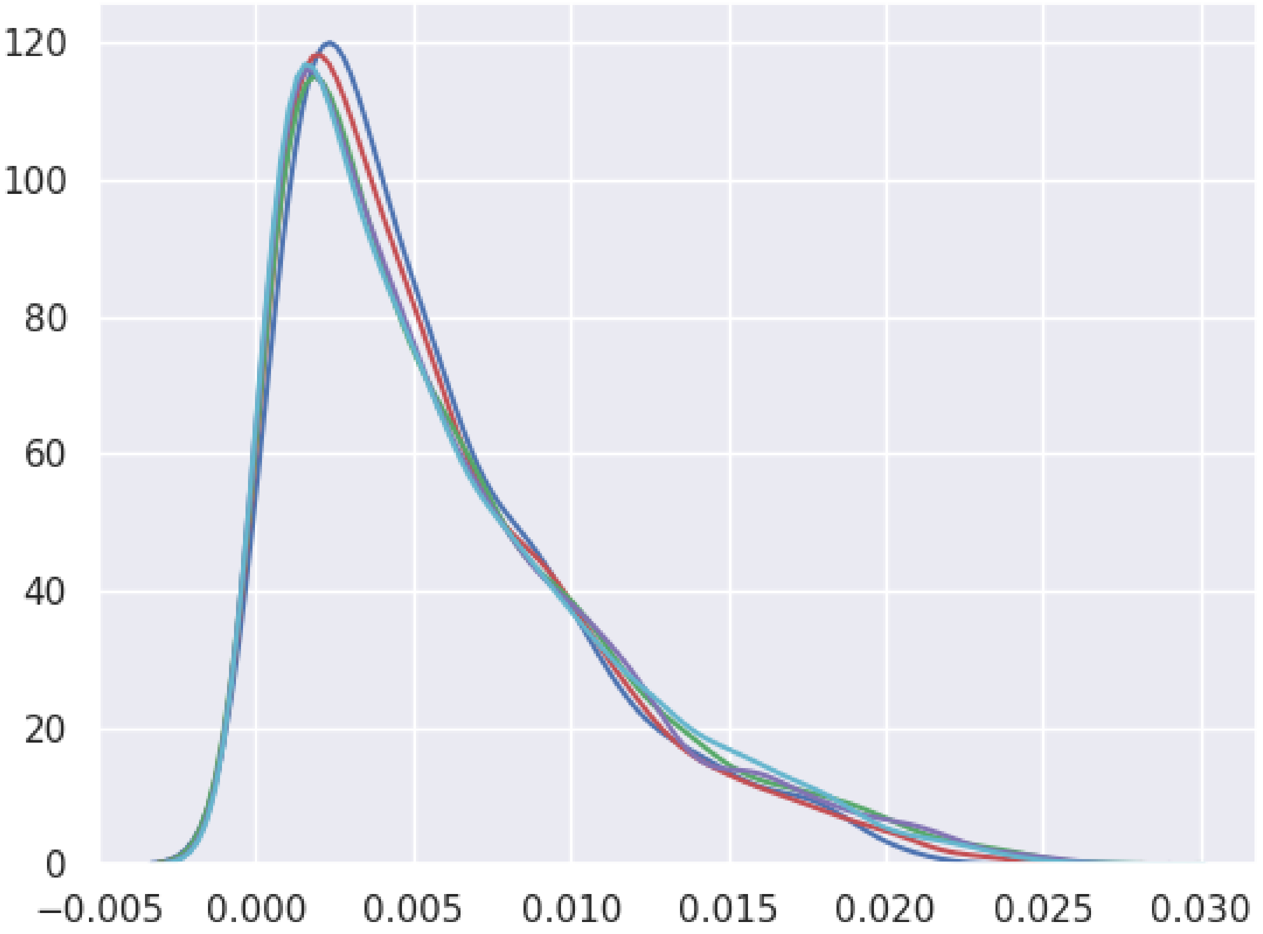
Green: n = 3000, k = 54

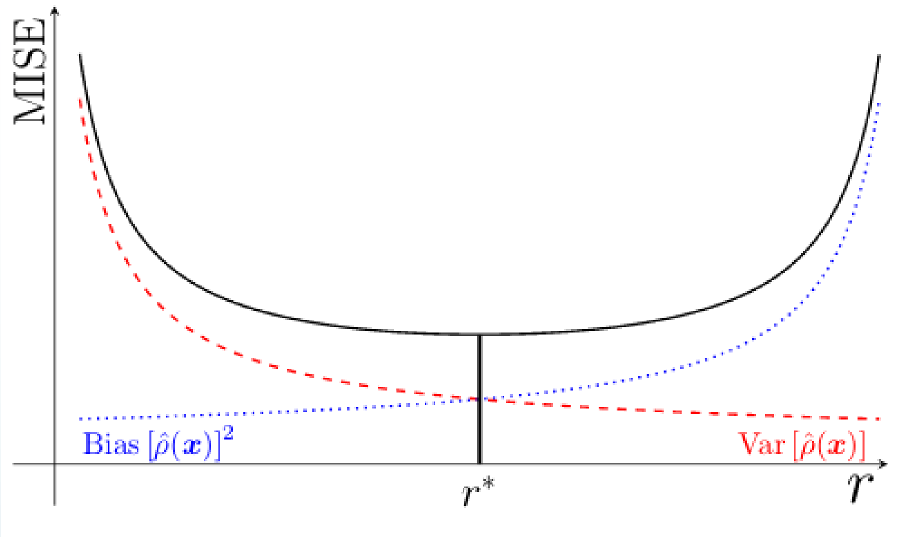
Magenta: n = 4000, k = 63

Cyan: n = 5000, k = 70

- Underlying sample density is approached as sample size increases, optimal k adjusts to reflect increase in sample size

Change in sample size, same k – Toy Scenario





- Changing sample size but keeping k constant increases MISE, as a suboptimal k is chosen
- D-dimensional radius for a test point increases/decreases as the test point needs to find more/less neighbours. This can give an apparent decrease/increase in the phase space density. As the sample size is increased, the phase space density becomes less susceptible to small changes in optimal k

Blue: n =1000, k = 54

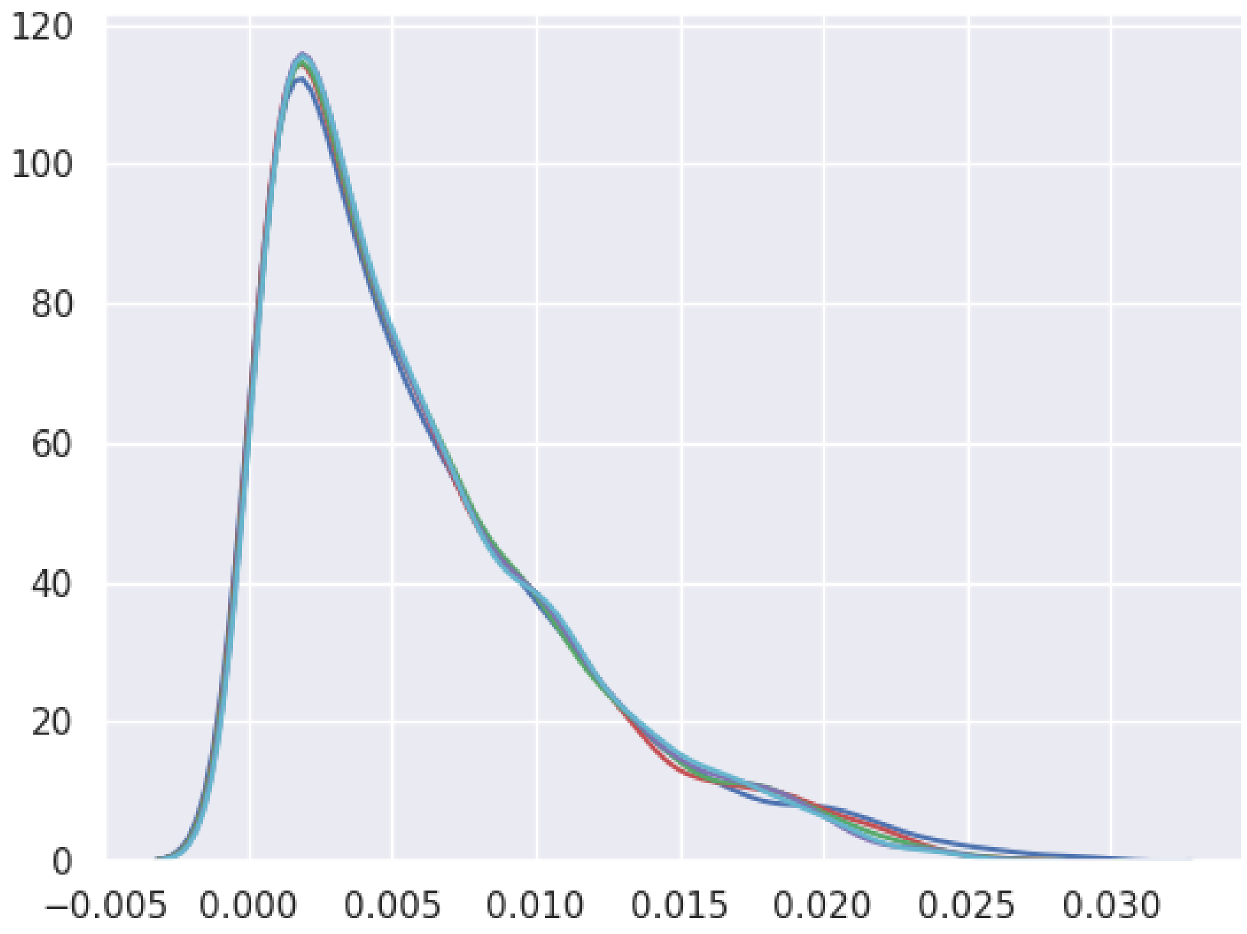
Red: n = 2000, k = 54

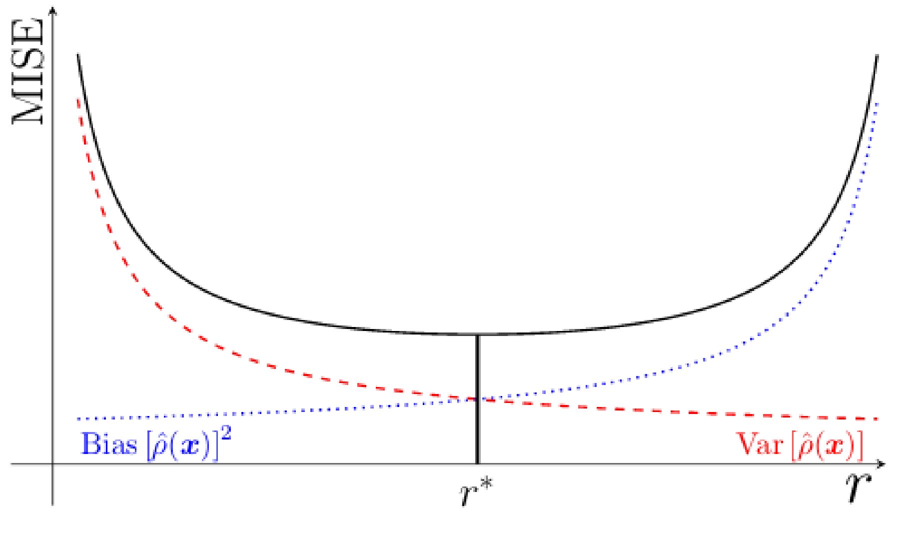
Green: n = 3000, k = 54

Magenta: n = 4000, k = 54

Cyan: n = 5000, k = 54

Change in k, same sample size – Toy scenario





- Choosing a suboptimal k leads to an increase in MISE
- When comparing data samples, one needs to use the same conditions for the sample i.e. use the same k to n relation e.g. $k \sim n^{-4/(4+d)}$
- A MISE that may not have been minimized may be desirable in areas that have been over or under smoothed

Blue: n =3000, k = 31

Red: n = 3000, k = 44

Green: n = 3000, k = 54

Magenta: n = 3000, k = 63

Cyan: n = 3000, k = 70

KNN and KDE behave the same

- Updated KNN routine to use KDTree algorithm approach to speed up calculation. Updated k-selection convention to match Francois' (i.e. whether to include test point, only minor difference)
- Checked changing the sample size(n), the k-value and changing one parameter while keeping the other constant. This resulted similar phase space densities, with differences due to increase in error due to choosing non-optimal parameters.
- Why do I get different phase-space densities?

Emittance in Experiments

- Emittance measurements can be biased
- The scraping of the beam on the aperture can give a false cooling effect
- Non-linearities can give rise to a false heating effect. The emittance of the beam has increased due to the non-linearities but the phase space volume hasn't changed size
- To see cooling, one can look at the change in phase-space volume or the change in density of that volume before and after it has gone through some material

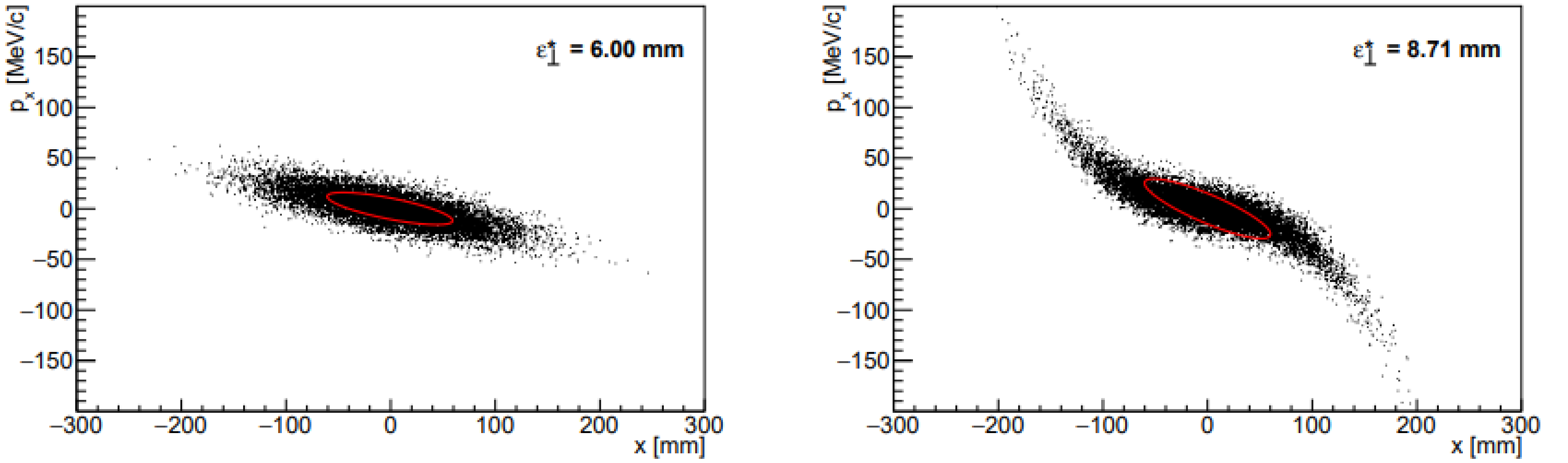
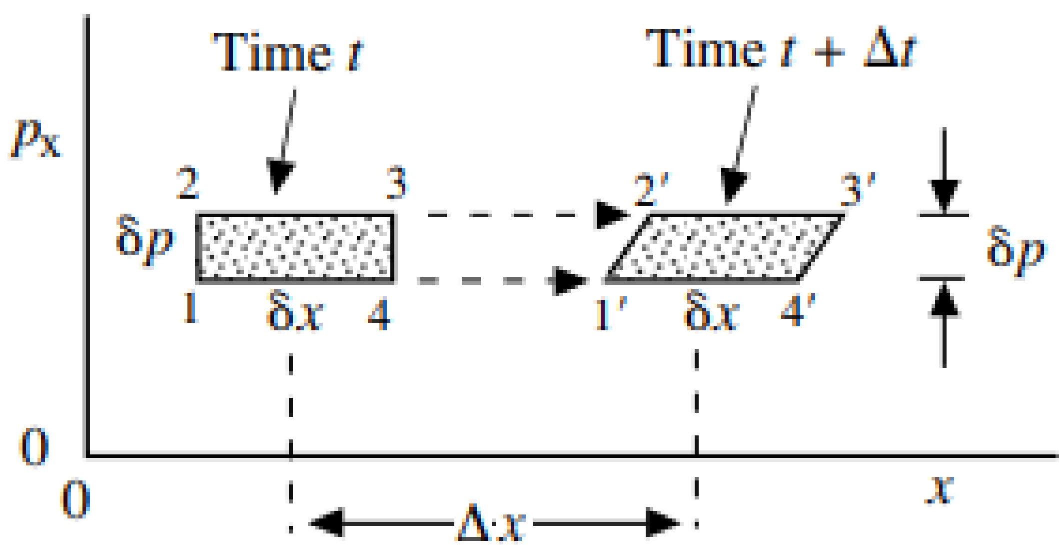
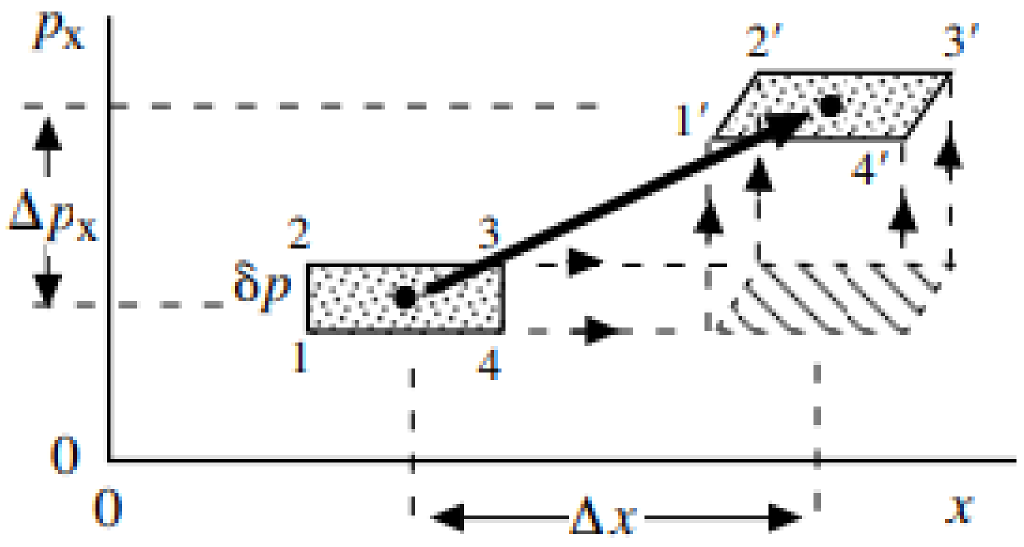


Figure 6.6: Scatter plot of a beam ($\epsilon_i = 6\,\text{mm}$, $\langle p_z \rangle = 140\,\text{MeV}/c$ and $\beta_\perp = 800\,\text{mm}$) after transport through a linear focusing lens of $f = 5\,\text{mm}^{-1}$ (left) and a similar nonlinear lens with $C_\alpha = 10^{-4}\,\text{mm}^{-2}$ (right). The red curve is the RMS ellipse.

Phase Space Volume and Density

- Take an arbitrary phase space volume upstream of the absorber and count the number of particles in that volume. Take the same volume downstream and count the number of particles in that volume. If it has changed then heating or cooling has taken place
- The problem is what does that phase space volume actually look like downstream as it has changed in shape due to differing momenta of particles in the beam and the magnetic forces of the cooling channel
- Transmission losses also need to be accounted for in an unbiased way





Liouville's theorem

- A particle beam can be described by the distribution of the particles in the beam also known as the phase space density $\rho(x, y, z, p_x, p_y, p_z)$.
- Liouville's theorem states that the density of particles in phase space is a constant i.e. ${}^{d\rho}/_{dt} = 0$ (providing there are no dissipative forces)
- The number of particles in a phase-space volume is then given by:

$$N = \int \rho(x, y, z, p_x, p_y, p_z) dx dy dz dp_x dp_y dp_z = \int \rho dV$$

- The phase-space density is directly related to the phase space volume
- The phase-space density can be calculated in a number of ways using density estimation techniques such as Kernel Density Estimation (KDE), the k-Nearest Neighbour Approach (KNN) plus many more
- Phase Space Density Estimation is a non-parametric technique to estimate the underlying probability density, the probability that a particle will be realized at a particular phase space density

Missing Data Problem

- No apparent problem with density estimation techniques, the correct density is calculated for that selection

- We know not all upstream particles make it downstream
- Transmission may introduce bias into calculations

- Missing Data can be classified in three different ways
  - - Missing Completely At Random (MCAR)
  - - Missing At Random (MAR)
  - - Missing Not At Random (MNAR)

Missing Completely At Random (MCAR)

- Good type of missing data
- Results not biased by missing data
- Think of 100 coin flips, and remove two results from the data
- The probability of getting heads remains the same bar for an increase in error based on those results

Missing At Random (MAR)

- Results biased by missing data
- But can use imputation techniques, interpolation or partial deletion
- Not perfect, but can give a confidence interval to the results based on the imputation method applied

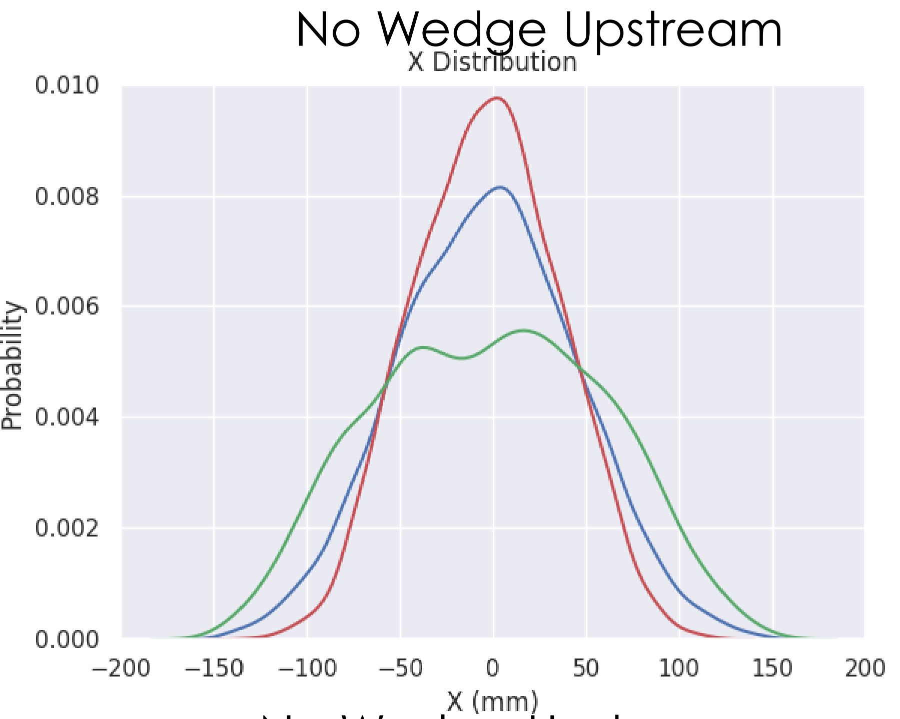
- Imagine a sensor with a particular operating temperature range from 100 K to 500 K
- The sensor is supposed to give a reading of the temperature every hour, but can only do so when in that operating range
- There are some occasions when no temperature is recorded.
- If one looks either side of the occasions when no temperature is recorded one can see it is close to 100 K
- One could assume that temperature went below 100 K on those occasions. Depending on the experiment it would be very unlikely to have gone above 500 K
- If the temperature changes behave in a smooth fashion one may even impute an actual temperature based on the temperature gradients before and after the missing data points instead of simply saying it is below 100 K
- Imputation techniques allows one to recover lost data points, but there is a natural level of uncertainty associated with it

Missing Not At Random (MNAR)

- Can lead to very biased results
- The missing data alters the interpretation of the results
- Significant problem in medical studies

- For example, a study in to the survival rate for a particular type of cancer.
- This study is followed up at regular time intervals into the future, but each time there are fewer and fewer responses.
- Could ignore the missing data and work on the completed data sets
- However, non-response may be directly correlated with non-survival
- The results are biased without accounting for this effect.

- MICE data is MNAR. Its missing due to scraping and magnetic field mismatches



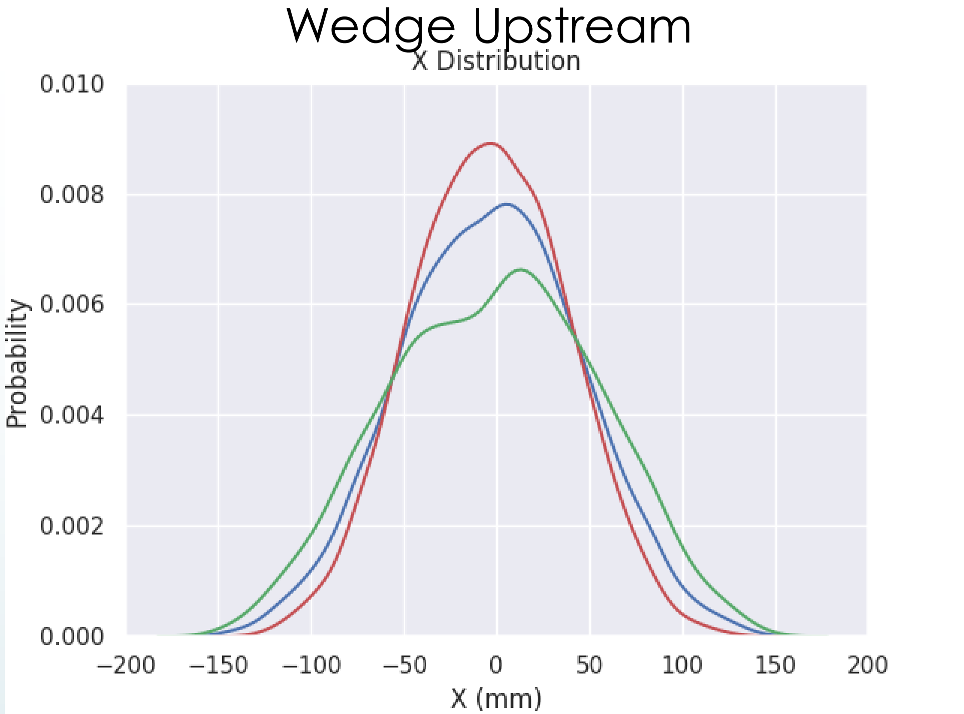
No Wedge (left) and Wedge (right)

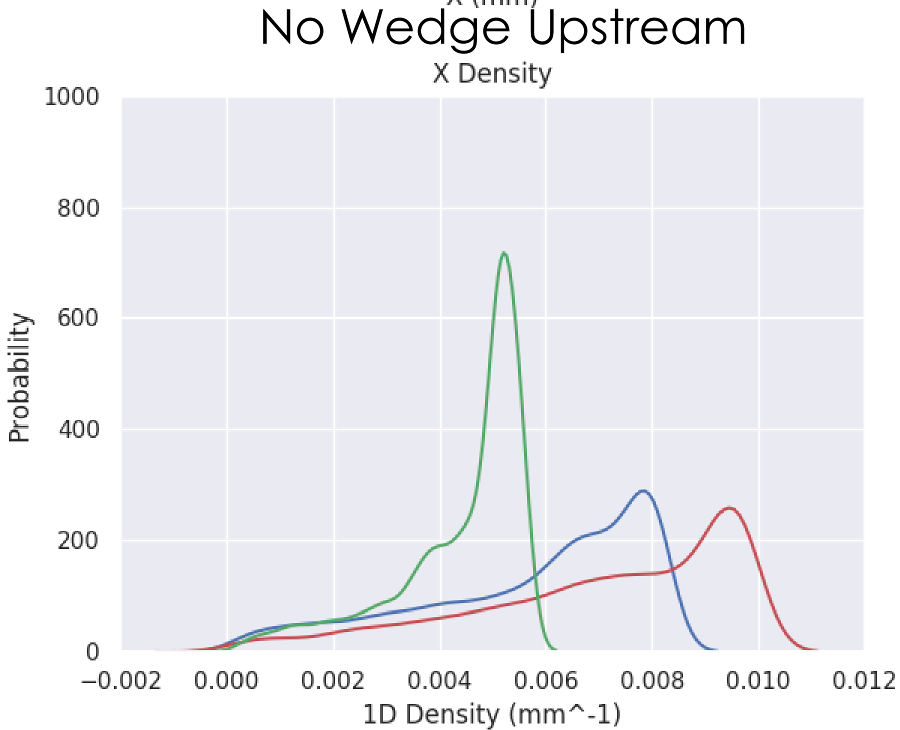
X Distribution (Top) and Density (Bottom)

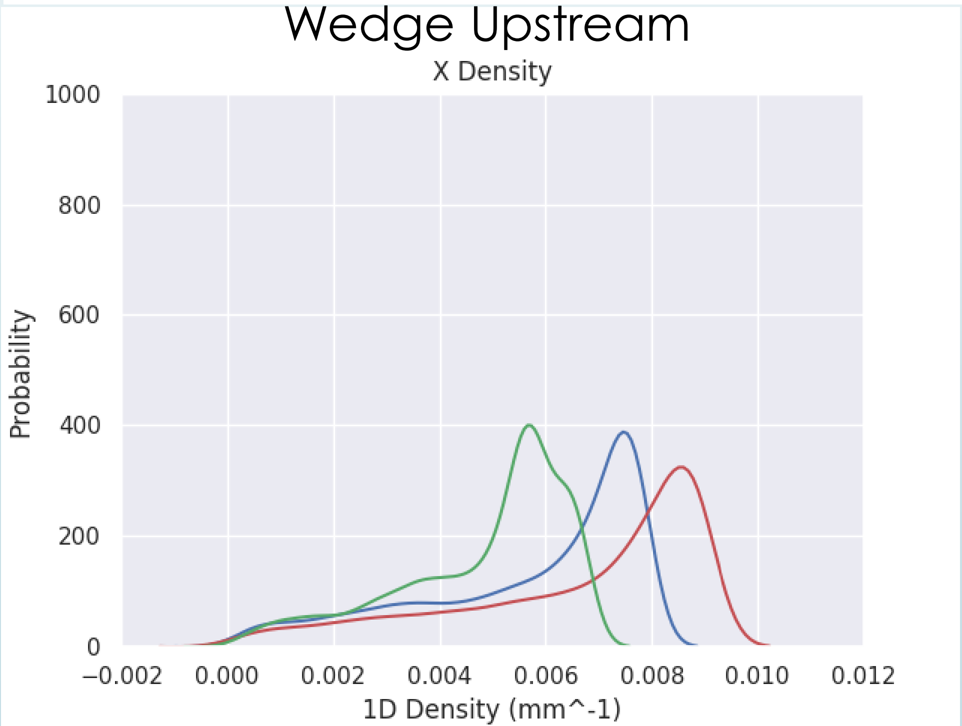
Blue – Full Upstream Sample
Red – Upstream Sample which makes it Downstream
Green – Upstream Sample which does not make it Downstream

Small preference for larger magnitude x not to make it downstream

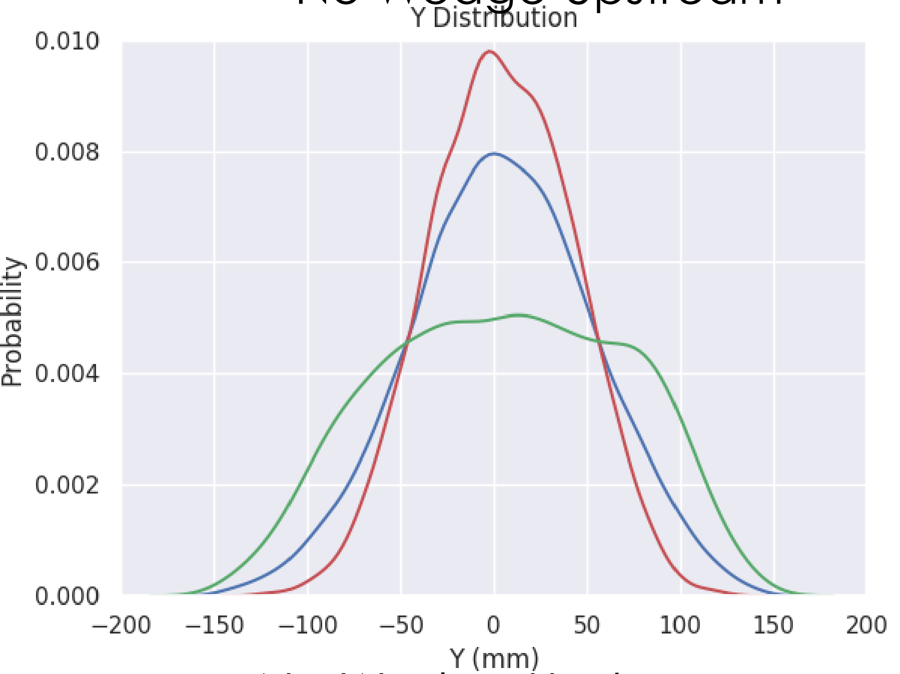
Wedge case shows slight directional bias as well. The Wedge does not transmit up to 15% of particles that would have made it downstream otherwise.



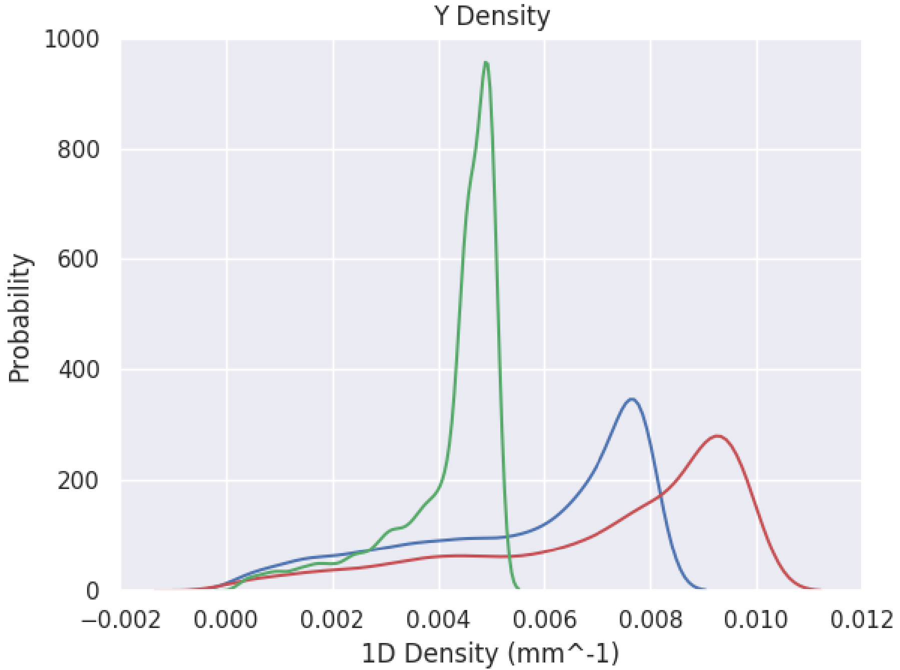




No Wedge Upstream



No Wedge Upstream



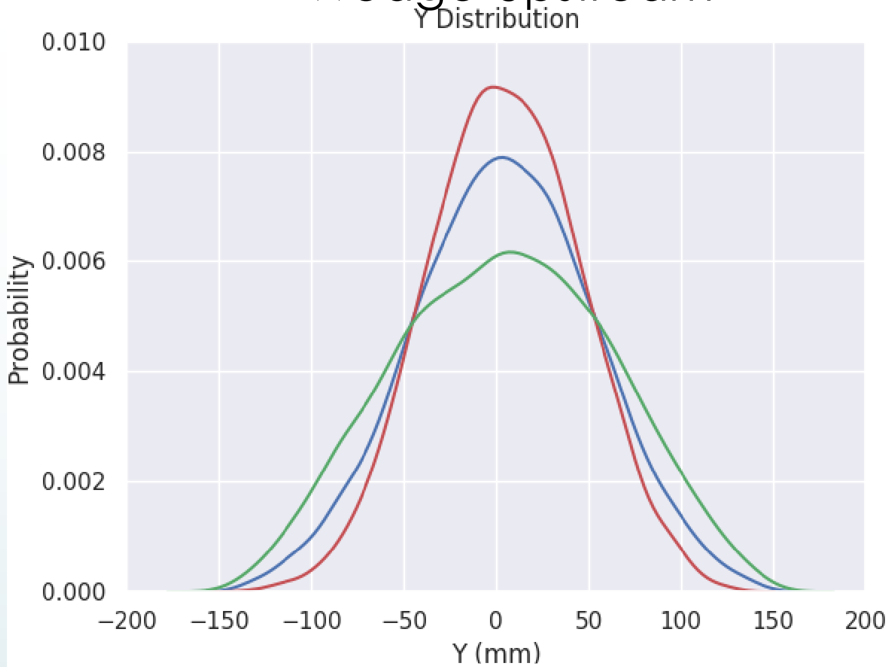
No Wedge (left) and Wedge (right)

Y Distribution (Top) and Density (Bottom)

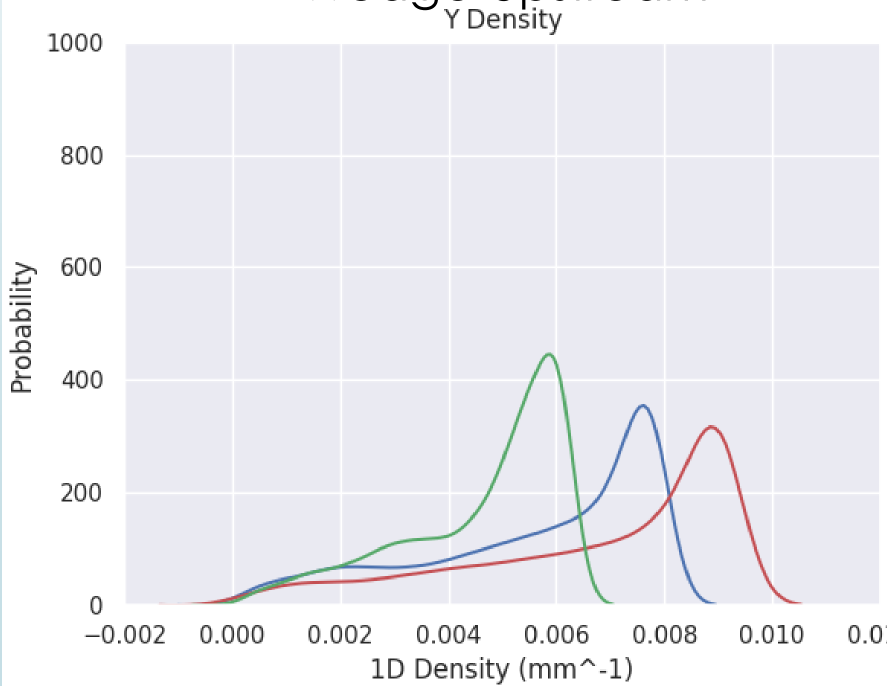
Blue – Full Upstream Sample
Red – Upstream Sample which makes it Downstream
Green – Upstream Sample which does not make it Downstream

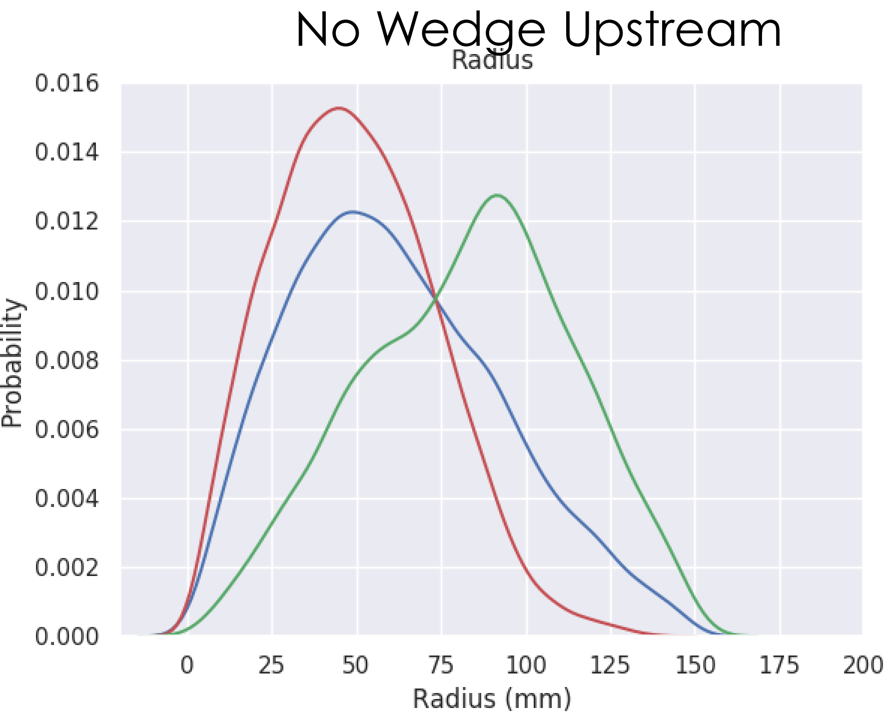
The Wedge counteracts some of the aperture cut effects, so that both low and high density particles do not make it downstream. This results in more similar distributions, however it is direction dependent.

Wedge Upstream



Wedge Upstream

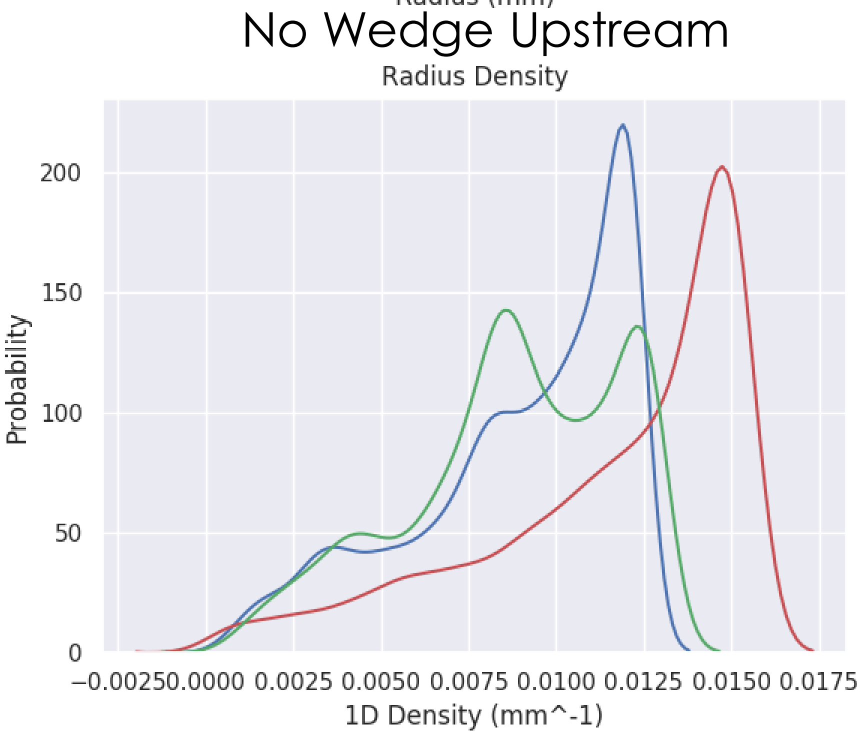




No Wedge (left) and Wedge (right)

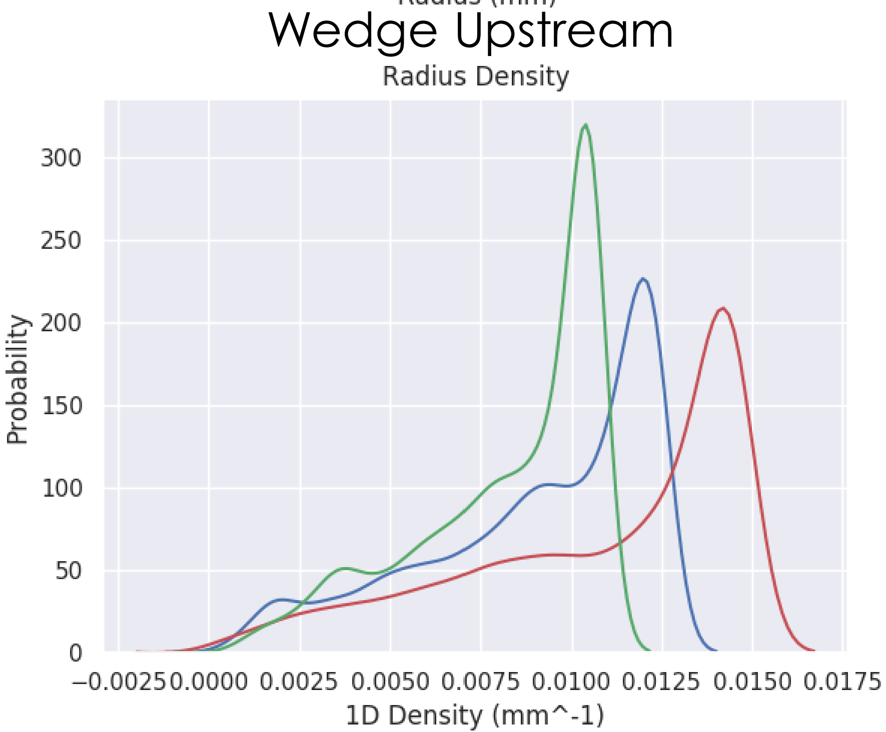
Radius Distribution (Top) and Density (Bottom)

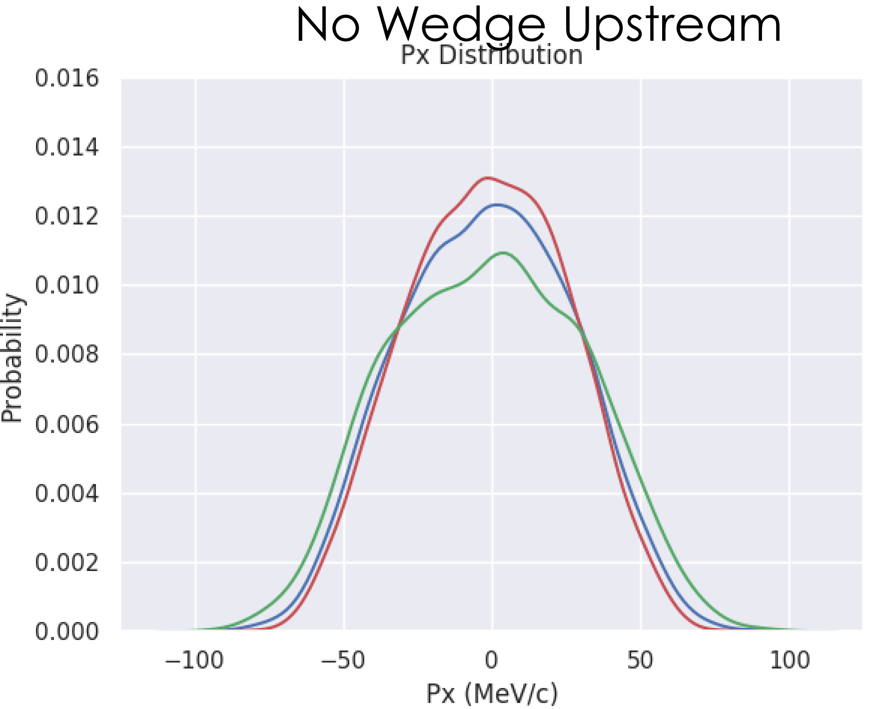
Blue – Full Upstream Sample
Red – Upstream Sample which makes it Downstream
Green – Upstream Sample which does not make it Downstream

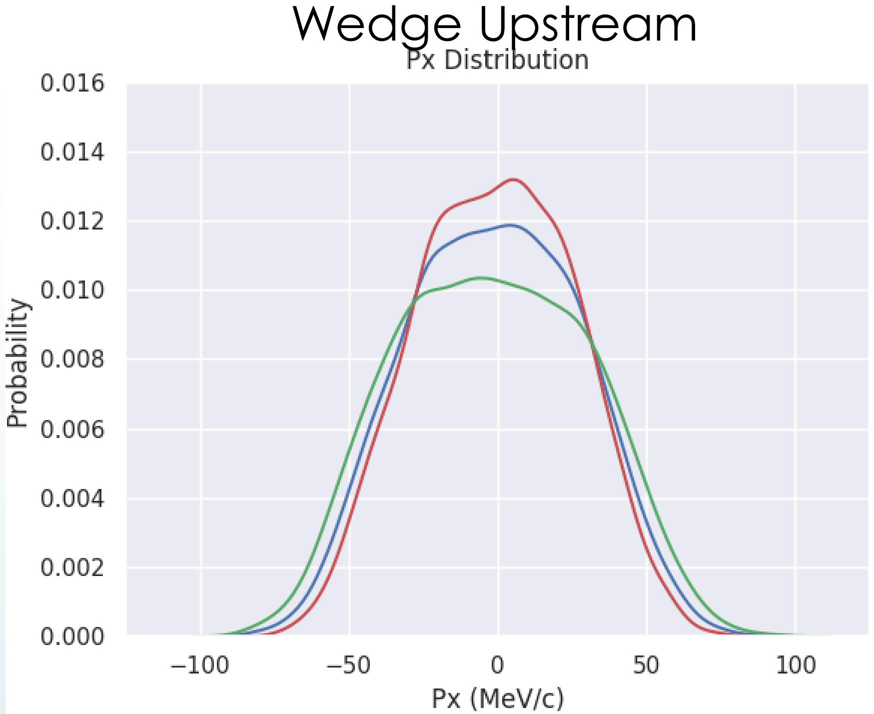


Not only high radius particles are eliminated. It is more likely for low to mid radius particles to be eliminated as there are simply more of them.

The double peak is due to the triangular shape of the distribution.







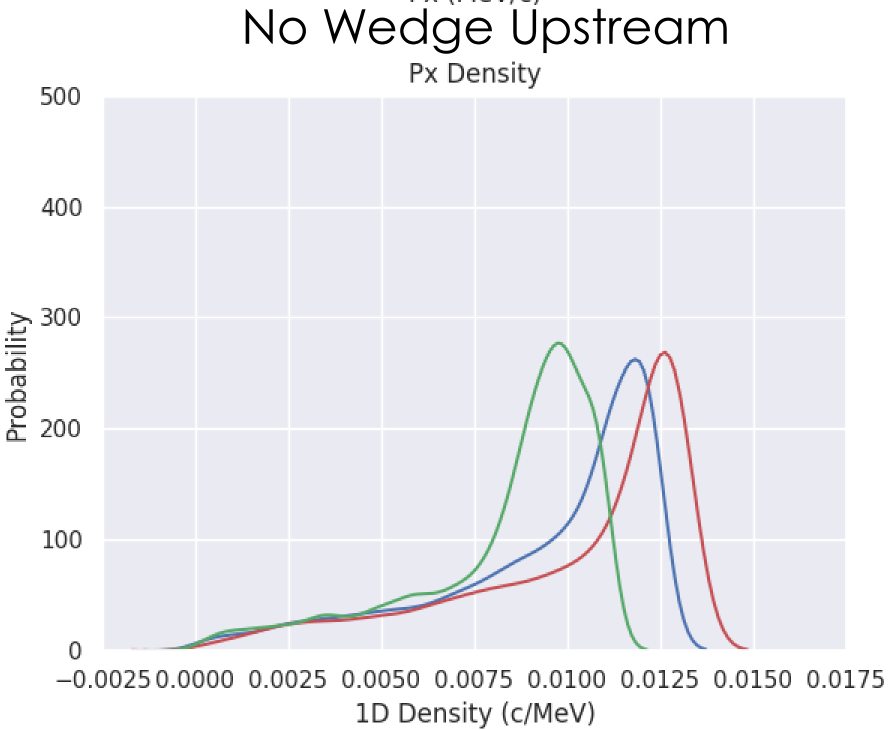
No Wedge (left) and Wedge (right)

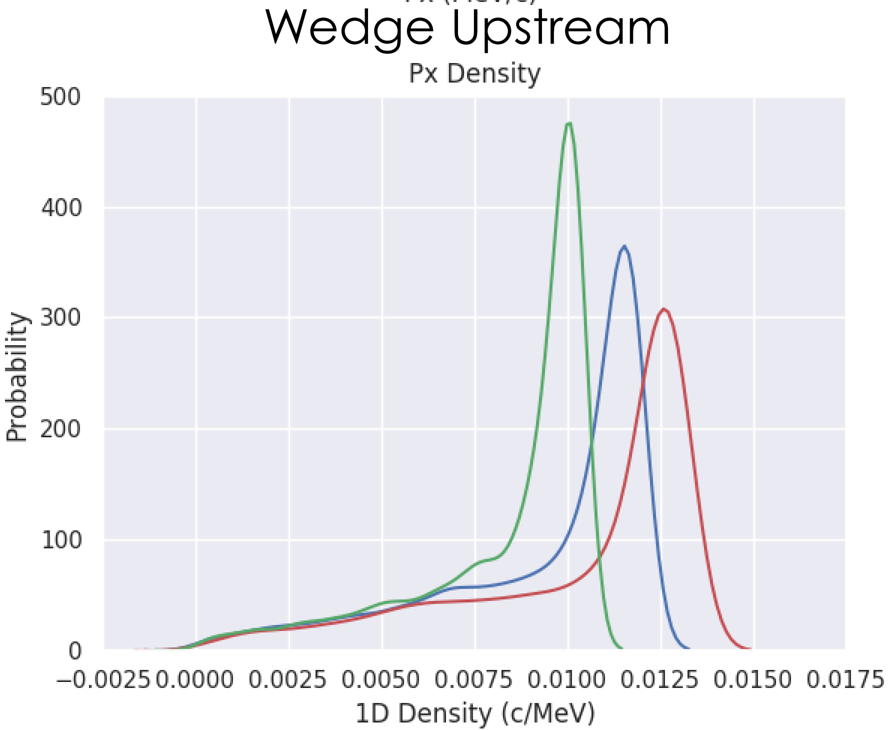
Px Distribution (Top) and Density (Bottom)

Blue – Full Upstream Sample
Red – Upstream Sample which makes it Downstream
Green – Upstream Sample which does not make it Downstream

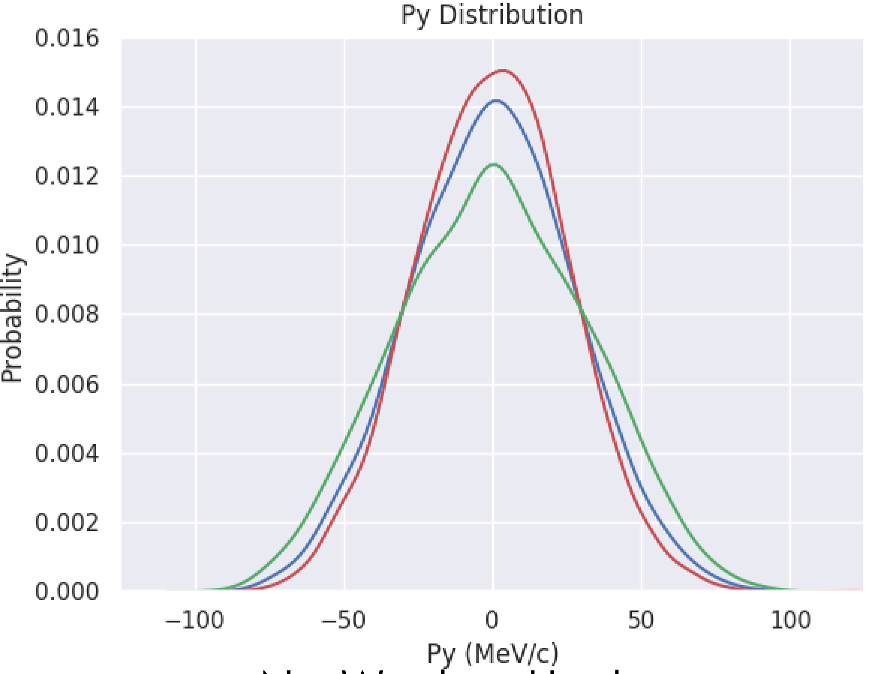
The Px and Py data are less affected by the aperture cut than the radius.

Px of higher density are more likely to be affected by the wedge than in the no wedge case

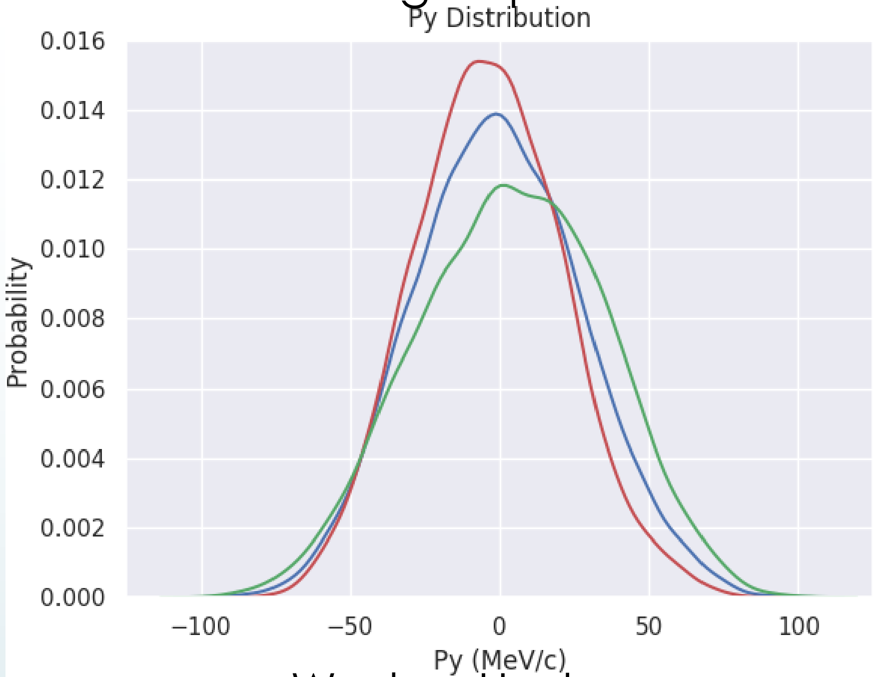




No Wedge Upstream



Wedge Upstream

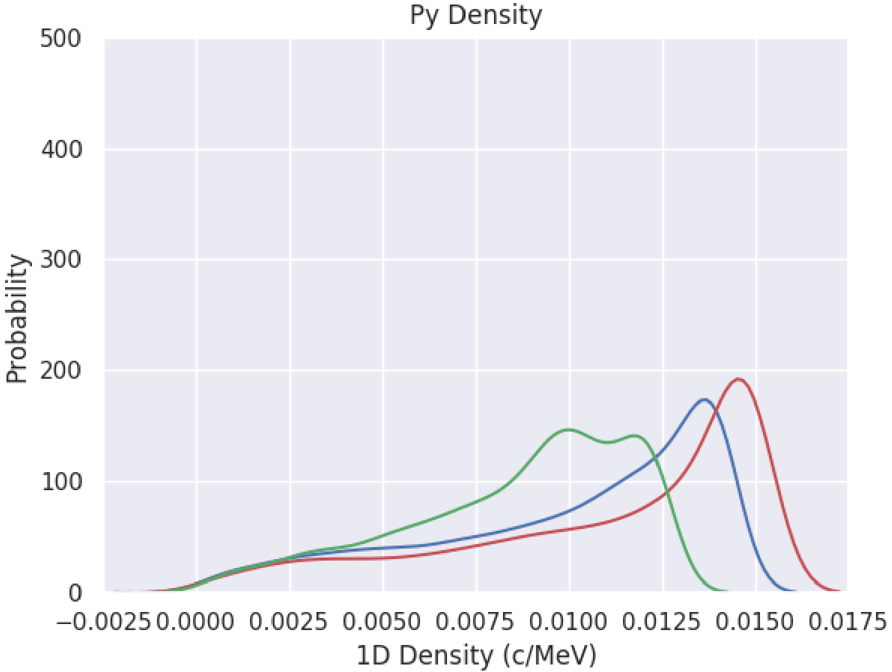


No Wedge (left) and Wedge (right)

Py Distribution (Top) and Density (Bottom)

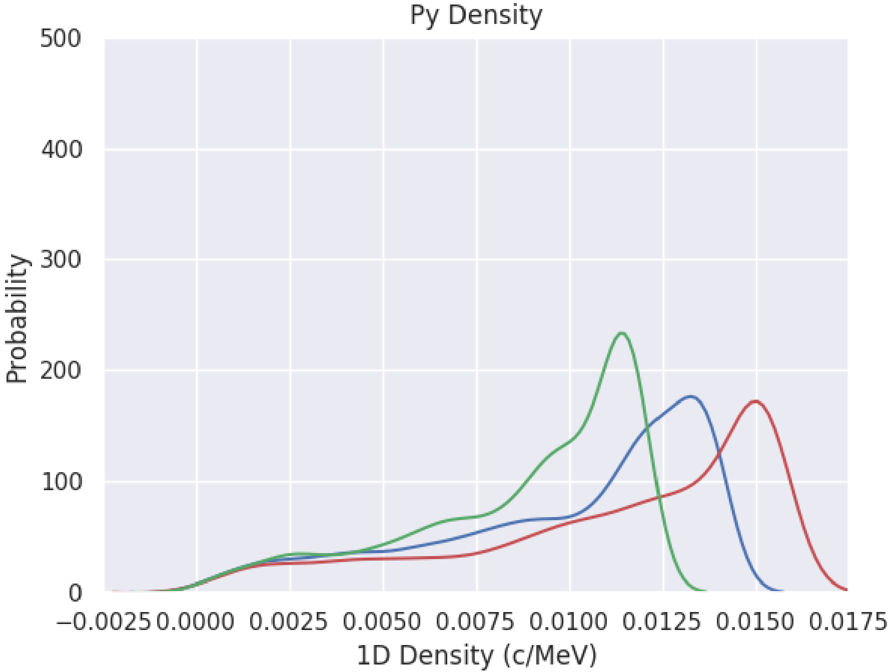
Blue – Full Upstream Sample
Red – Upstream Sample which makes it Downstream
Green – Upstream Sample which does not make it Downstream

No Wedge Upstream



The Py distribution shows a directional preference for particles that don't make it downstream. This is due to the x-py correlation

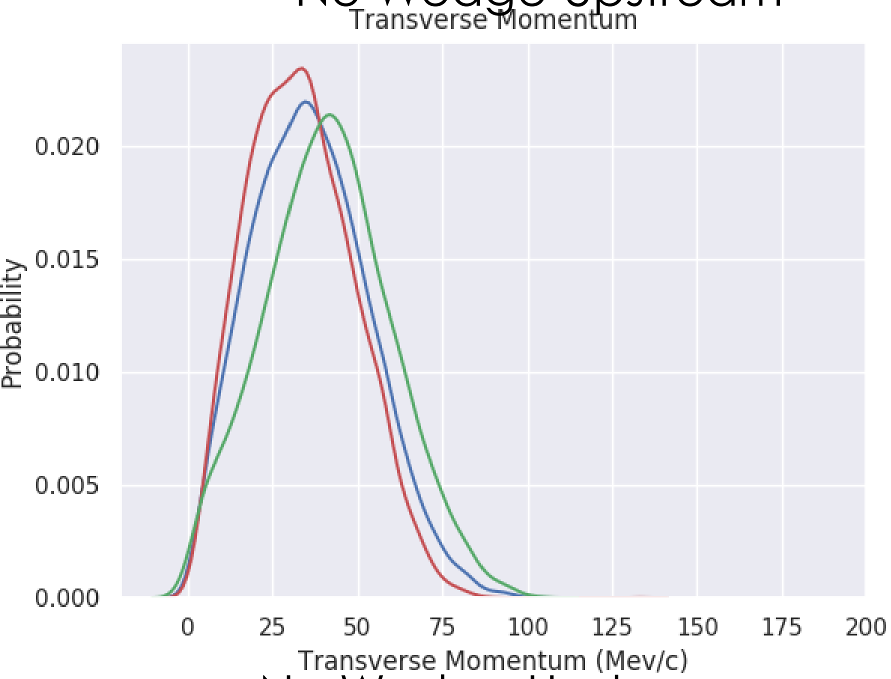
Wedge Upstream



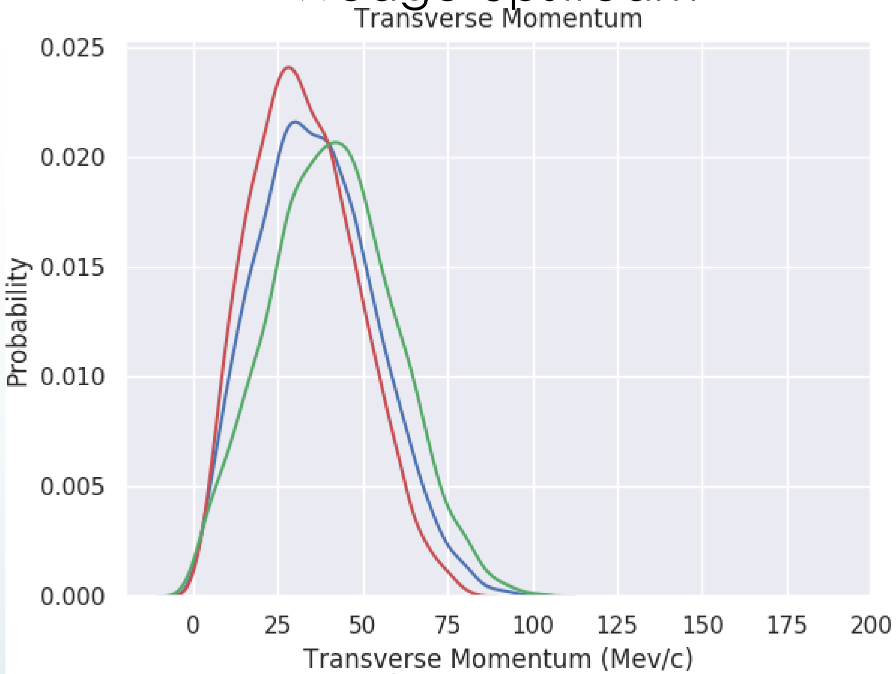
No Wedge (left) and Wedge (right)

Pt Distribution (Top) and Density (Bottom)

No Wedge Upstream



Wedge Upstream

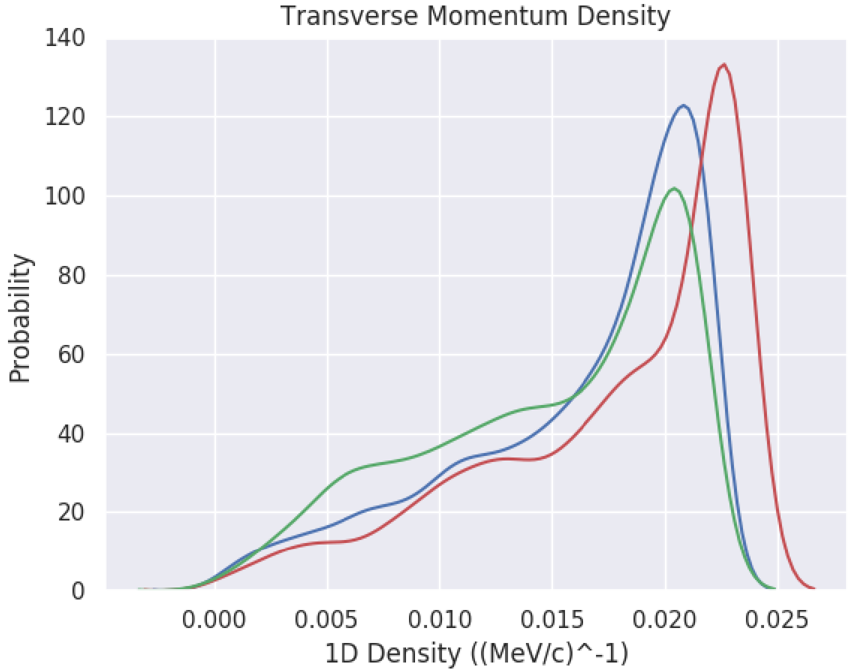


Blue – Full Upstream Sample

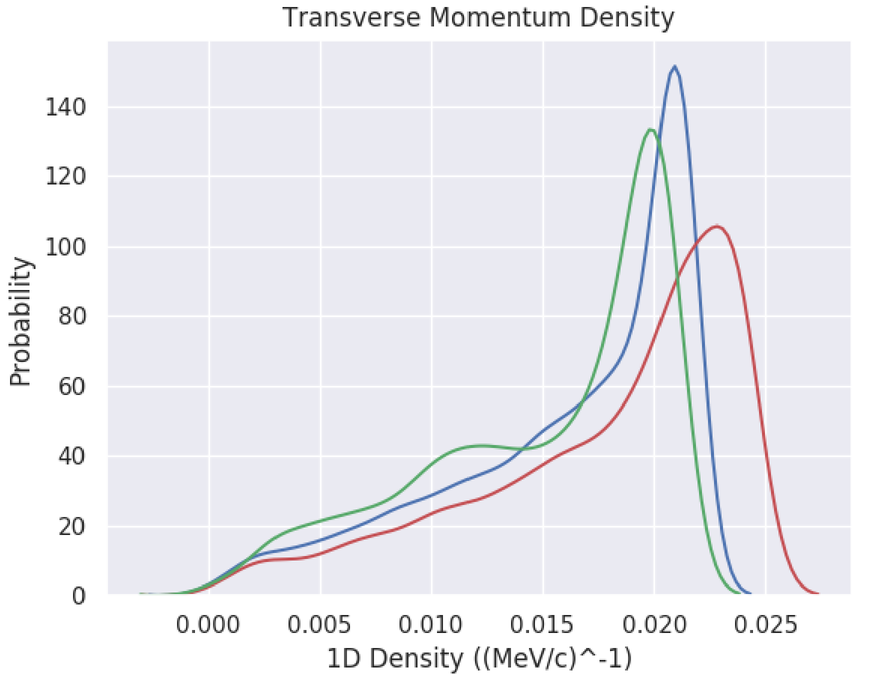
Red – Upstream Sample which makes it Downstream

Green – Upstream Sample which does not make it Downstream

No Wedge Upstream



Wedge Upstream



Higher Transverse momenta are less likely to make it downstream, but do not show the same distribution shape as for radius

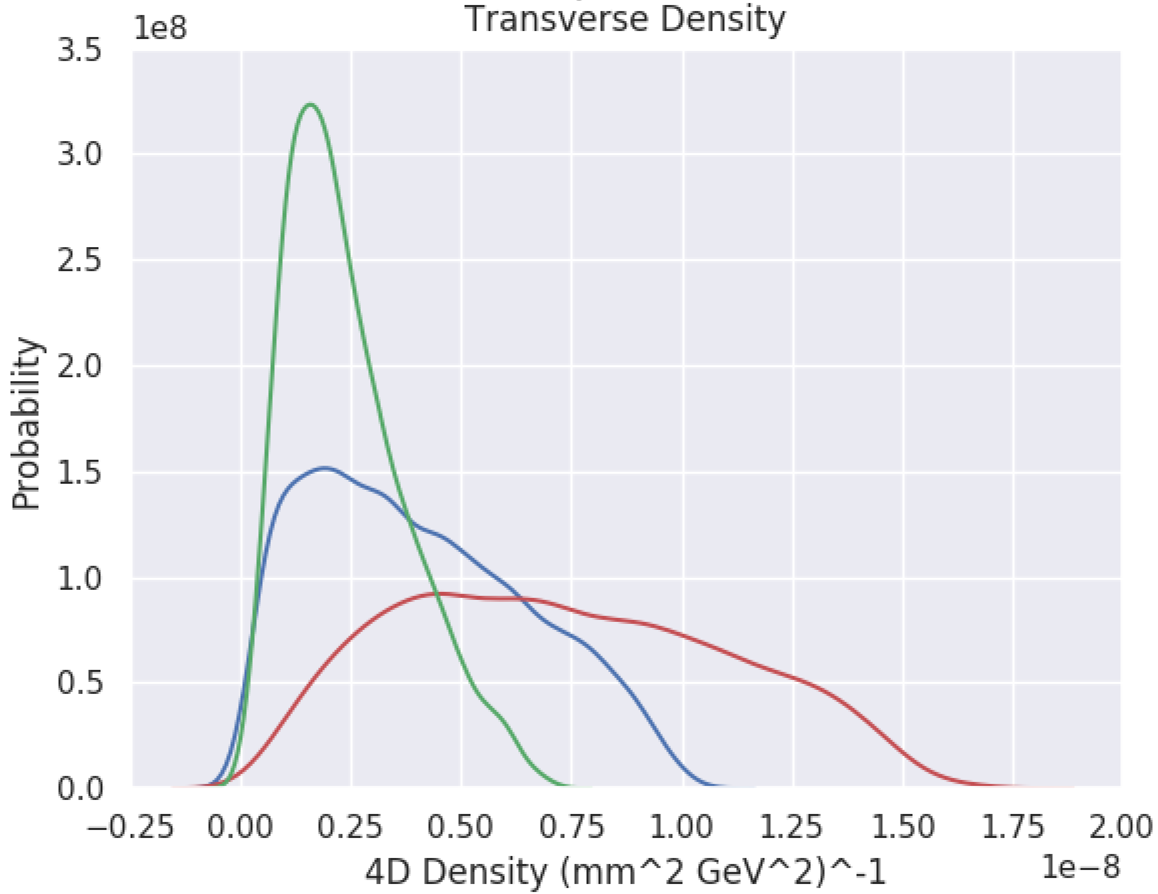
This results in the upstream and downstream samples being affected more in two of the four dimensions.

4D Transverse density

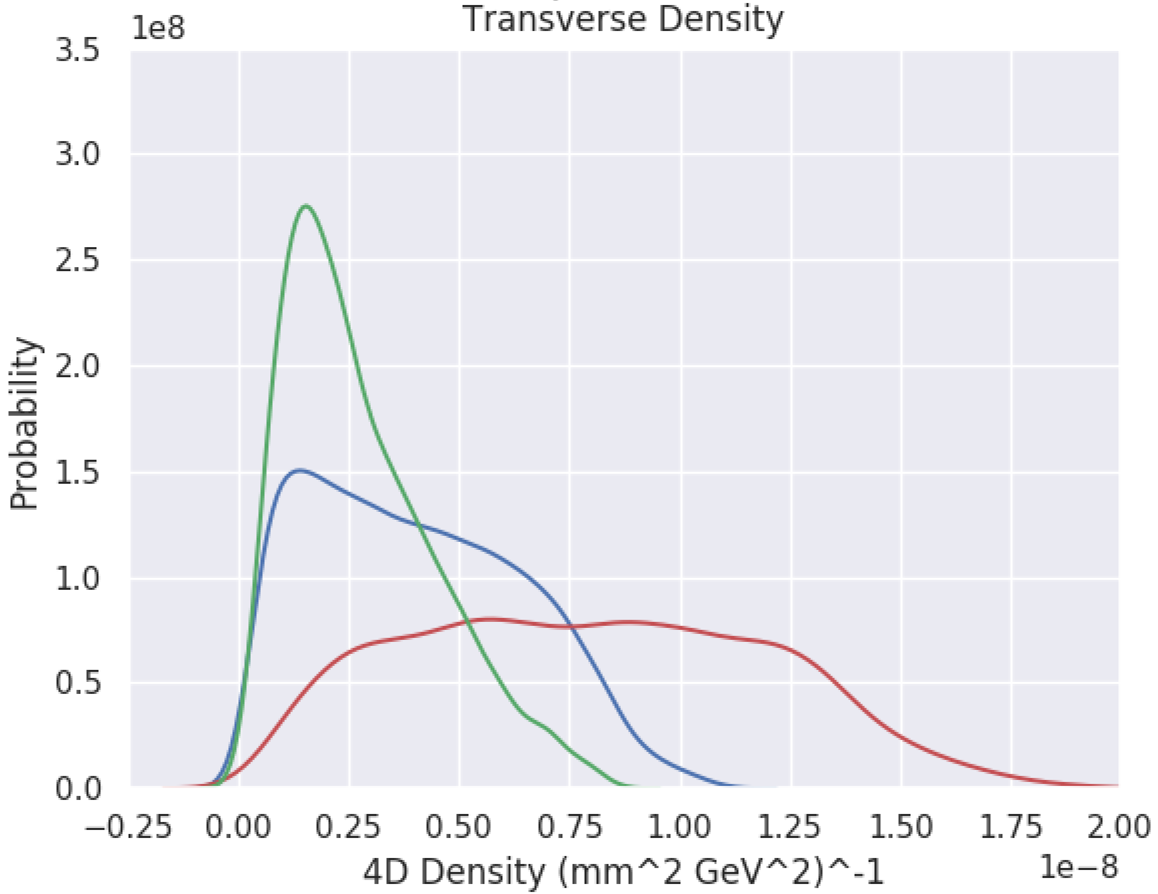
Blue – Full Upstream Sample
Red – Upstream Sample which makes it Downstream
Green – Upstream Sample which does not make it Downstream

Blue distributions are fairly similar, however the green distribution has become broader as some lower radius particles have been eliminated by the wedge

No Wedge Upstream



Wedge Upstream



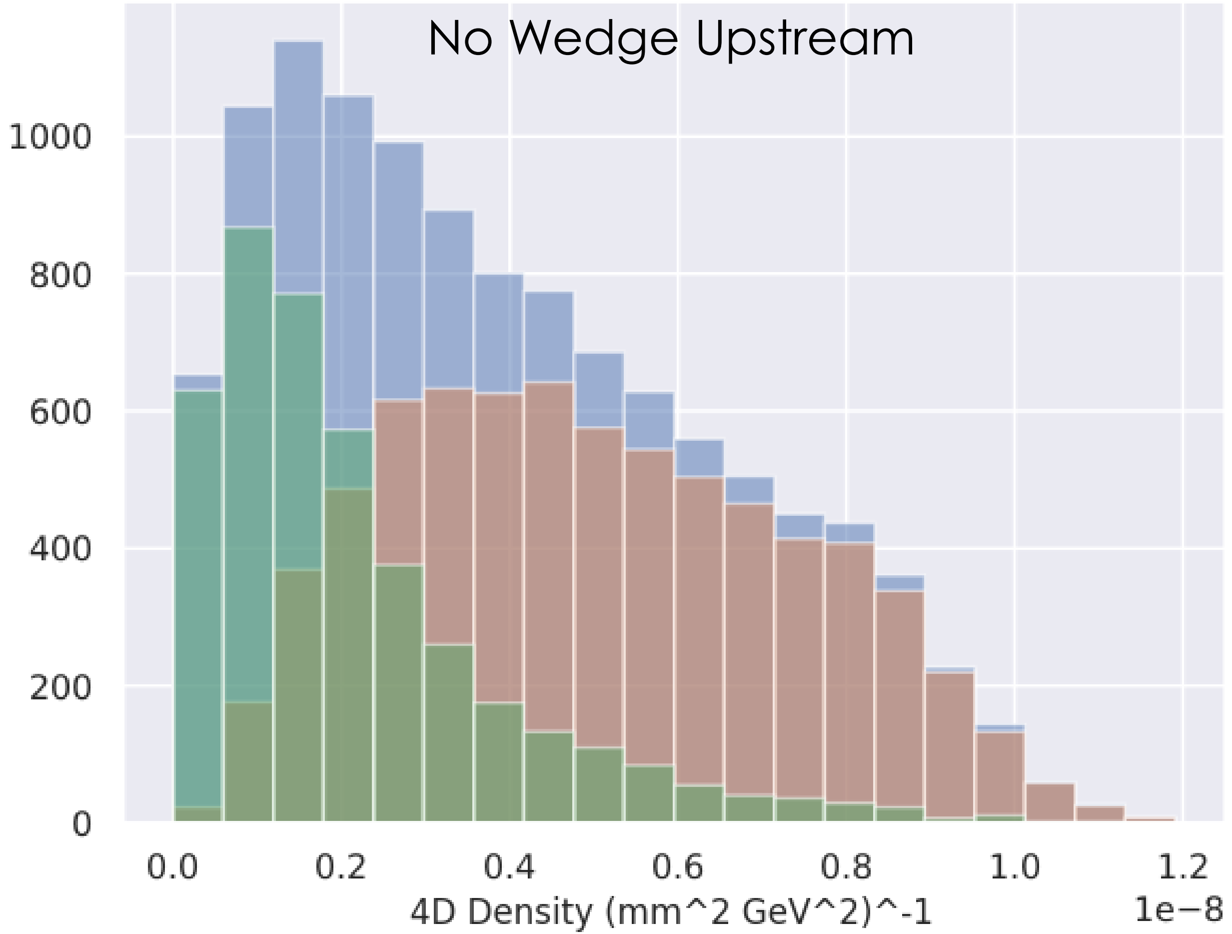
4D Transverse density

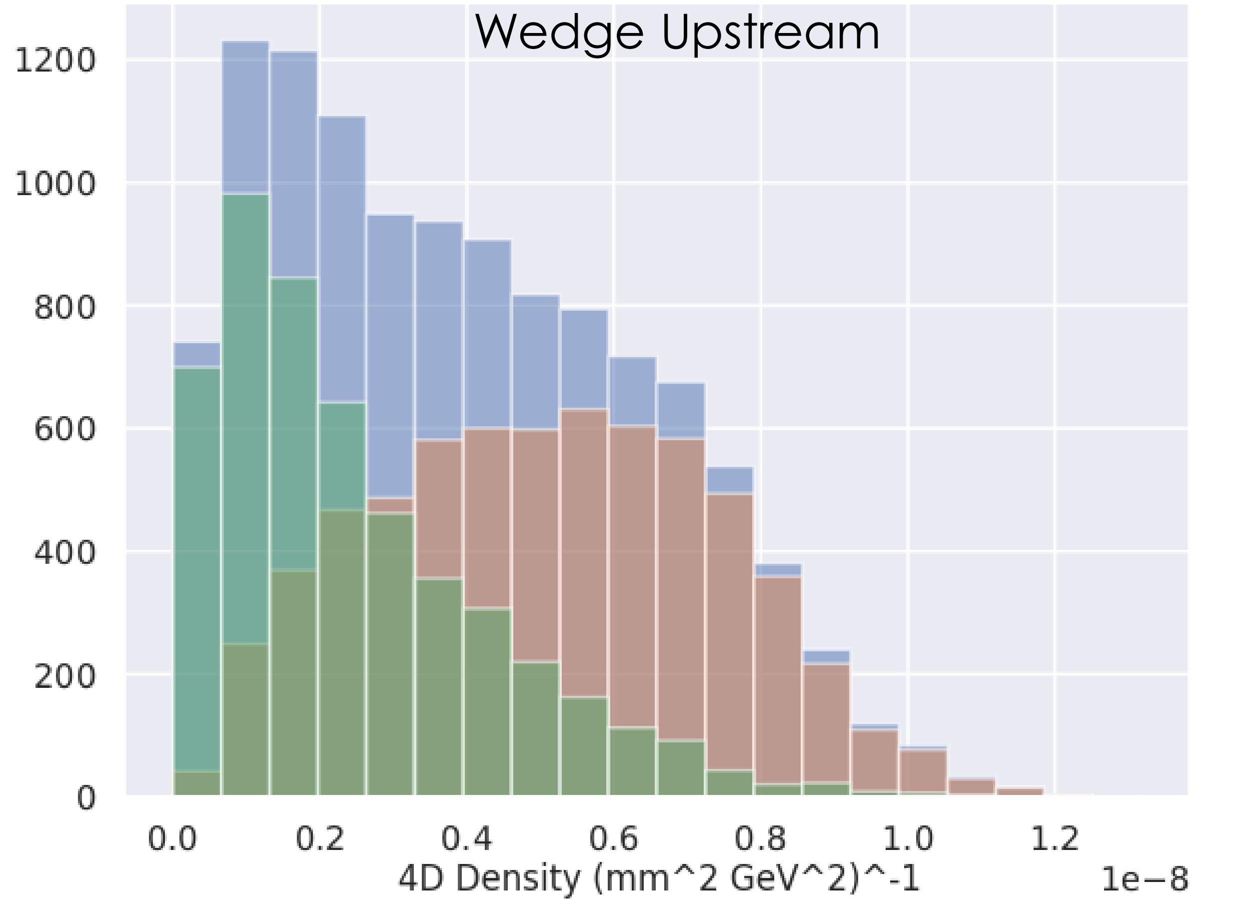
- The full upstream sample and the upstream sample which makes it downstream have different phase-space distributions and occupy different phase-space volumes
- This causes different phase-space densities to be calculated, however the correct one for that phase-space volume
- A solution to compare these two distributions is to express the upstream distribution which makes it downstream as part of the overall phase-space volume of the full upstream sample
- In the upstream sample this is quite simple as one has the full upstream sample. Downstream it is far more difficult as we only have the particles which have made it downstream

Not only low density particles are eliminated

Blue – Full Upstream Sample
Orange – Upstream Sample which makes it Downstream
Green – Upstream Sample which doesn't make it Downstream

The full upstream distribution (blue) can be divided into the upstream distribution which makes it downstream (orange) and upstream distribution which doesn't make it downstream (green) calculated over the full Upstream distribution volume.





Phase Space Density

- Phase Space Density tells you the distribution of particles for a particular sample over a given 6D volume.
- This volume is bounded by the particles position and momentum in 6D phase space. This volume remains conserved over time providing it doesn't encounter any dissipative forces (Liouville).
- The missing data from the downstream sample tends to have a lower phase space density. As these particles aren't used for the density calculation downstream, the density is calculated over a smaller volume.
- This gives an artificial rise in the phase-space density. It is the correct phase-space density, but has been calculated over a different phase-space volume then the upstream sample.
- Hence, normalization factors need to be applied to compare the upstream and downstream samples
- For the case of no absorber, the upstream and downstream densities should then match (when there is no scraping or dissipative forces)

Current Normalization

- Currently the Downstream density is normalized to the Upstream density by the ratio of the sample sizes.
- The density is calculated from:

$$\rho_{transverse} = norm\left(\frac{2k}{n\pi^2\sqrt{|\Sigma|}}\frac{1}{\left((X_i-\bar{X})\Sigma^{-1}(X_i-\bar{X})^T\right)^2}\right)$$

- Where $k = (n/norm)^{4/4+dim}$ and norm is the transmission.
- The Upstream density is then calculated as (norm = 1):

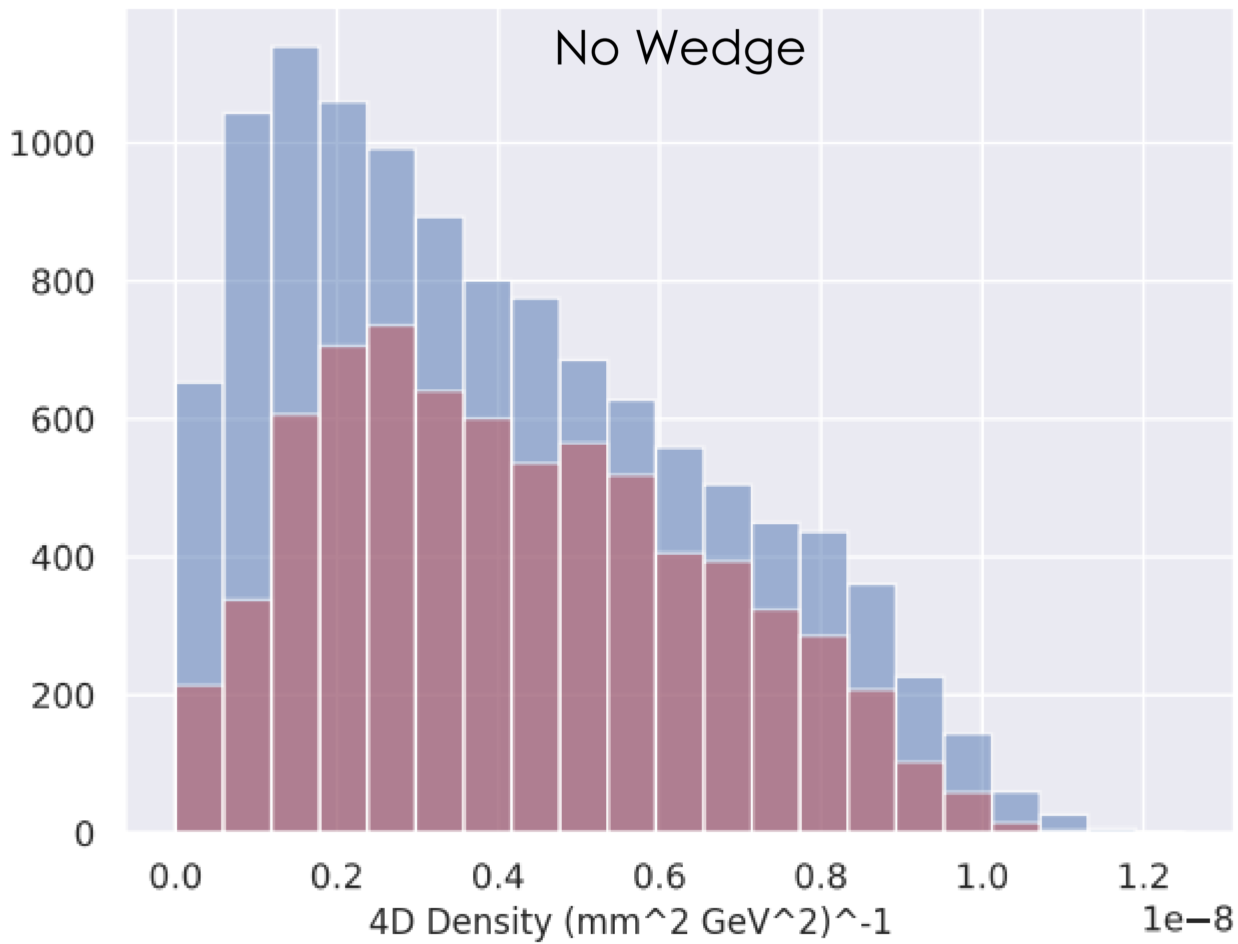
$$\rho_{transverse} = \left(\frac{2}{\sqrt{n_{up}}\pi^2\sqrt{|\Sigma_{up}|}}\frac{1}{\left((X_i-\overline{X_{up}})\Sigma_{up}{}^{-1}(X_i-\overline{X_{up}})^T\right)^2}\right)_{Up}$$

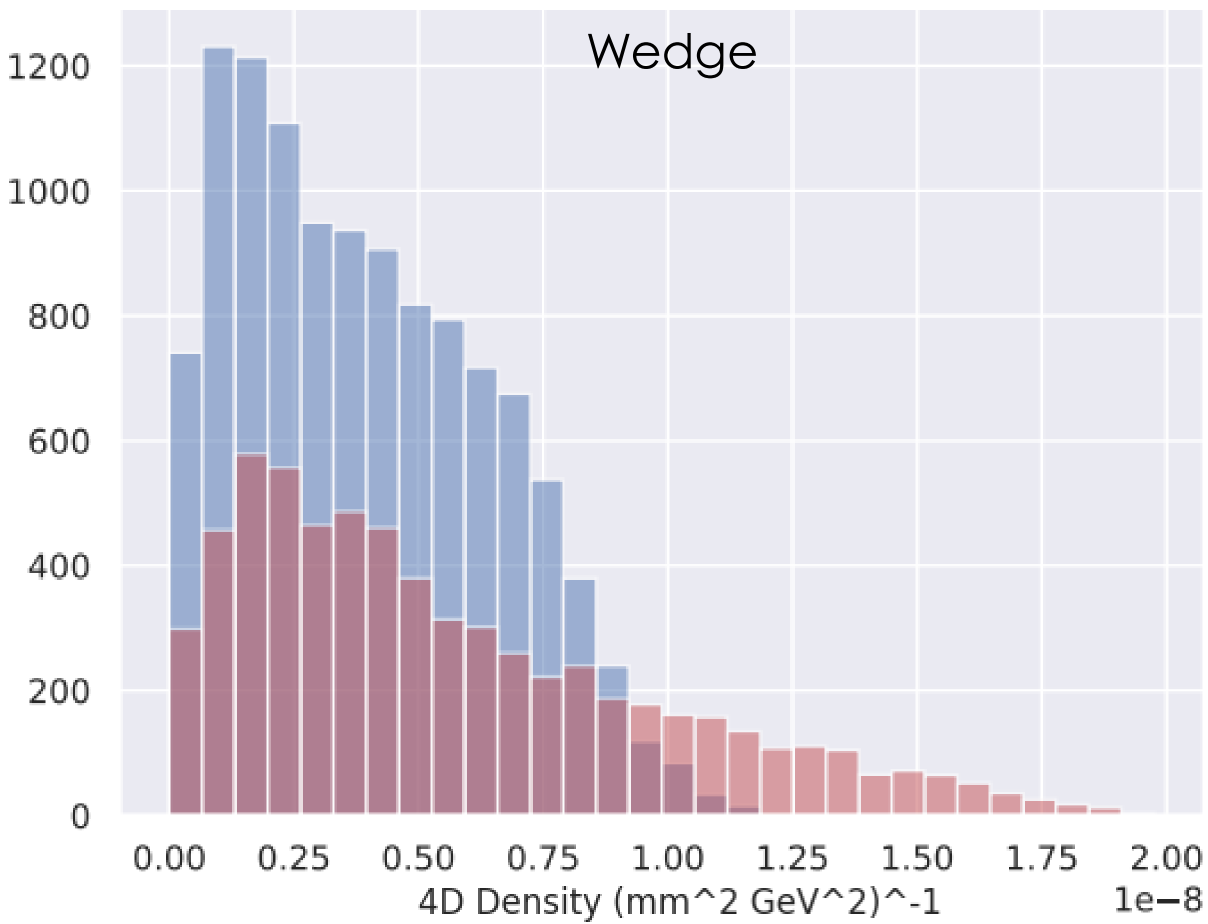
- The Downstream Normalized density is then calculated as $(norm = n_{down}/n_{up})$:

$$\rho_{transverse} = \left(\frac{2}{\sqrt{n_{up}}\pi^2\sqrt{|\Sigma_{down}|}}\frac{1}{\left((X_i-\overline{X_{down}})\Sigma_{down}{}^{-1}(X_i-\overline{X_{down}})^T\right)^2}\right)_{Down}$$

Upstream vs Downstream density using current normalization

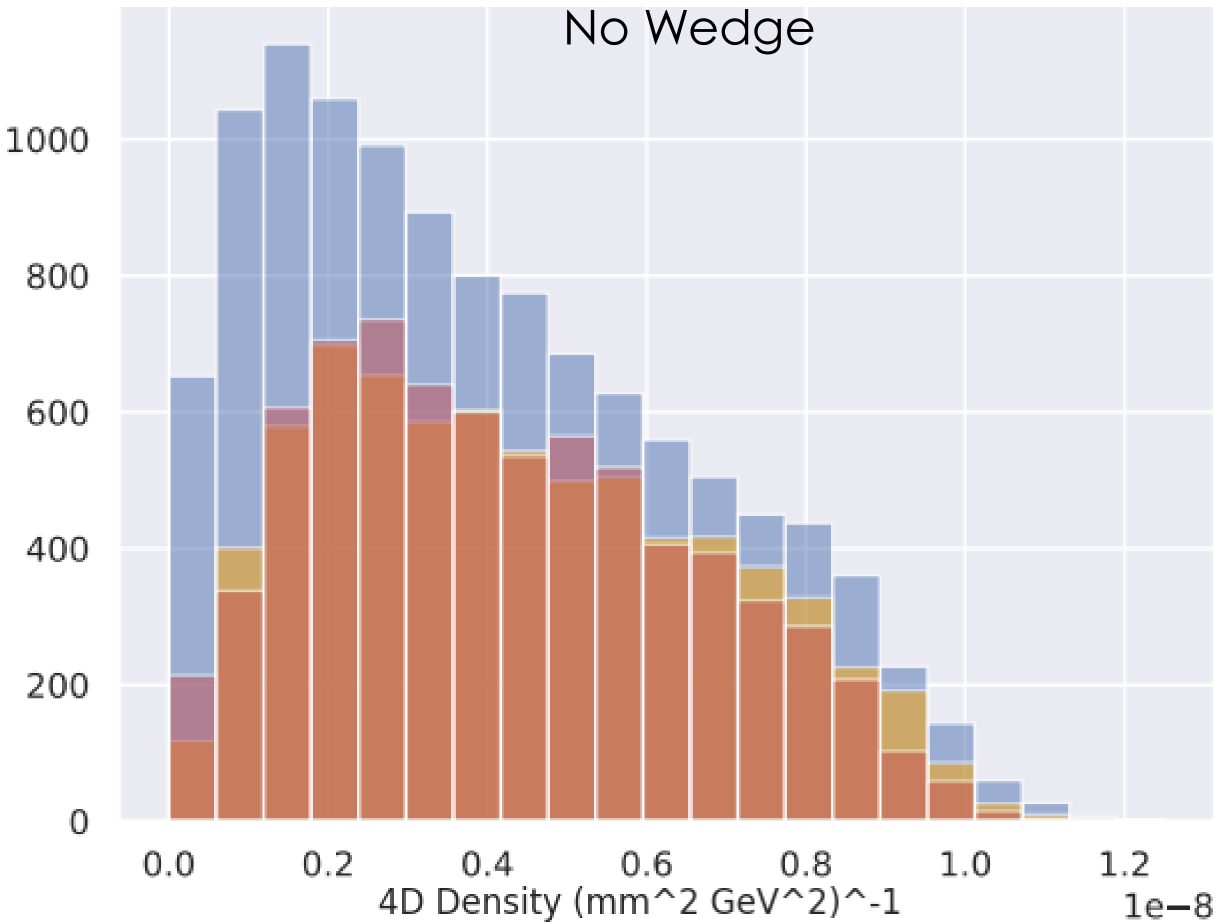
Blue – Full Upstream Sample
Red – Full Downstream Sample

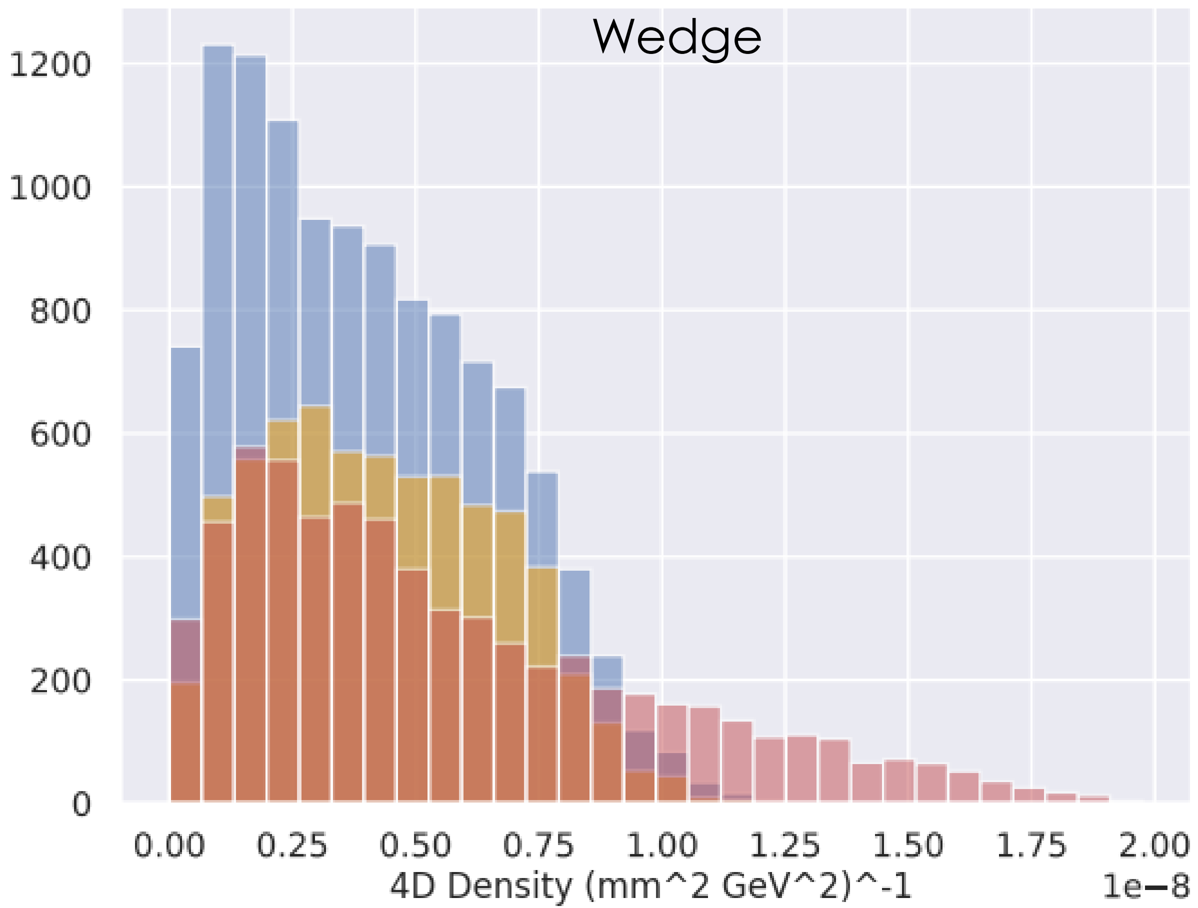




The same normalization can be used for the upstream sample which makes it downstream

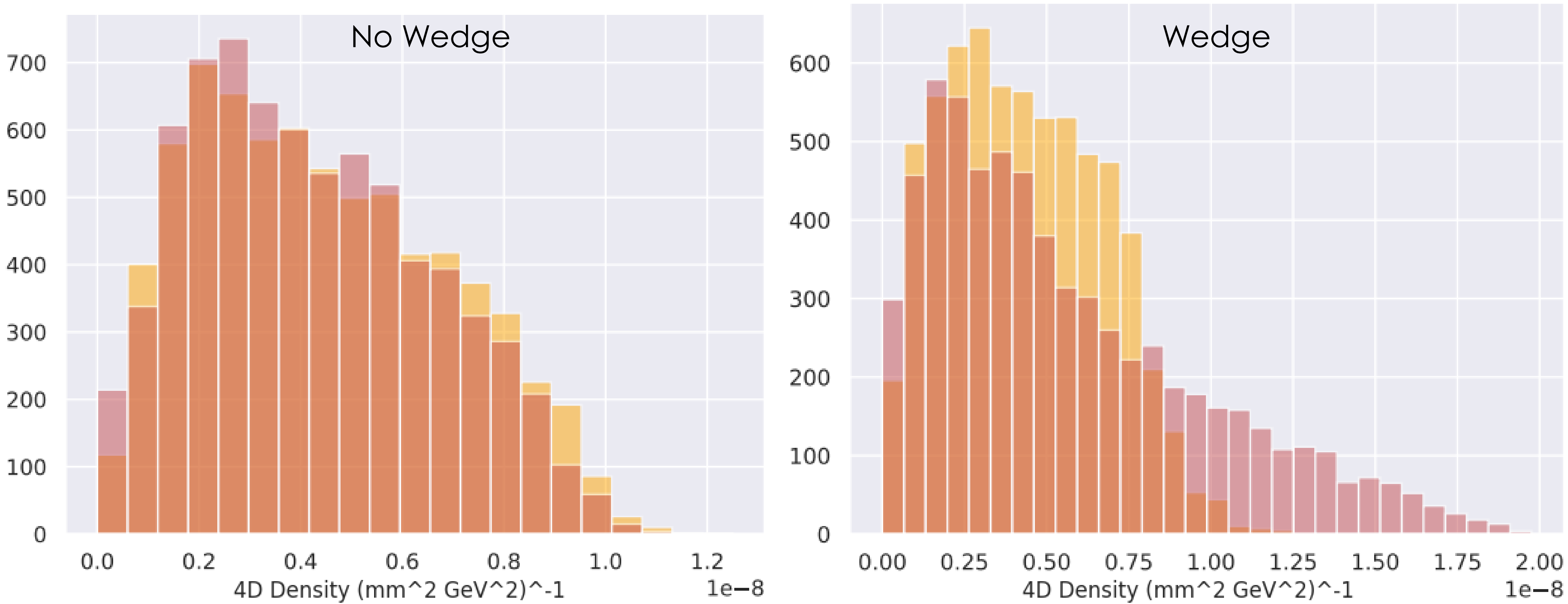
Blue – Full Upstream Sample
Red – Full Downstream Sample
Orange – Upstream Sample which made it Downstream





In the no absorber case, comparing the upstream sample which makes it downstream with the downstream sample, shows only very small heating, likely due to scraping and the tracker station window itself

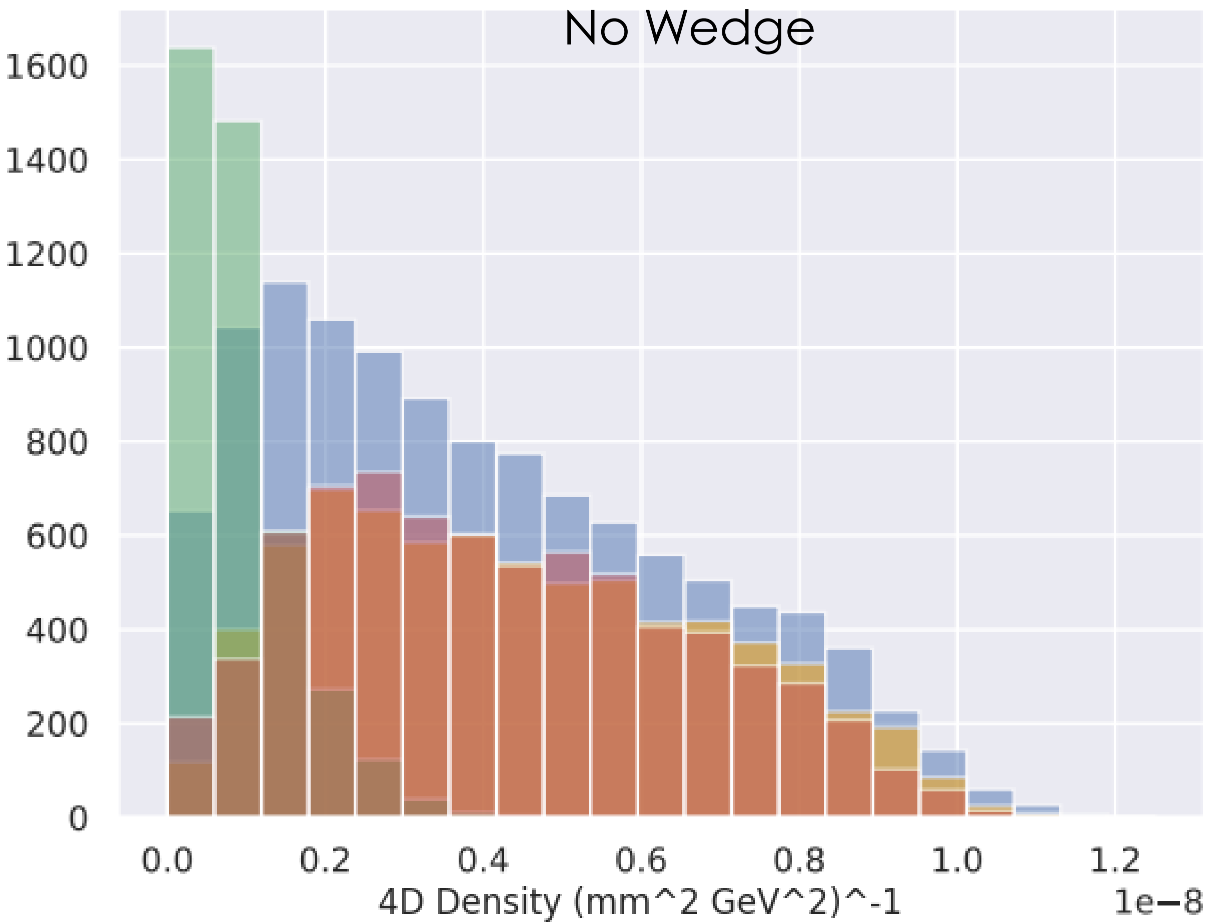
Red – Full Downstream Sample
Orange – Upstream Sample which made it Downstream

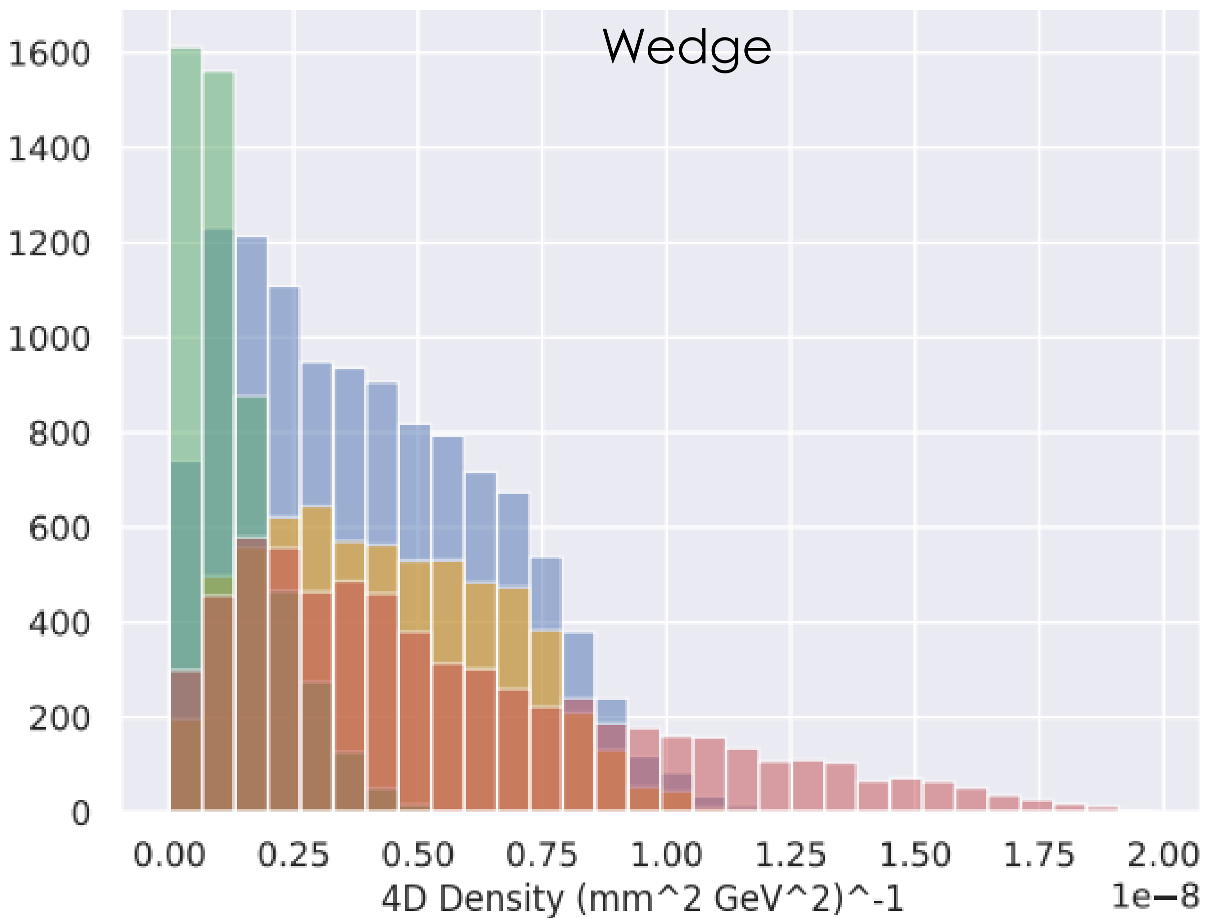


The normalization procedure can be used for the missing data as well

The upstream sample which makes it downstream and the upstream sample which doesn't make it downstream do not add up to the full upstream sample. This may indicate that the normalization is not correct

Blue – Full Upstream Sample
Red – Full Downstream Sample
Orange – Upstream Sample which made it Downstream
Green – Upstream Sample which doesn't make it downstream





Recall how the density was calculated

- The Upstream density is then calculated as (norm = 1):

$$\rho_{transverse} = \left(\frac{2}{\sqrt{n_{up}}\pi^2 \sqrt{|\Sigma_{up}|}} \frac{1}{\left((X_i - \overline{X_{up}}) \Sigma_{up}{}^{-1} (X_i - \overline{X_{up}})^T \right)^2} \right)_{Up}$$

- The Downstream Normalized density is then calculated as $(norm = n_{down}/n_{up})$:

$$\rho_{transverse} = \left(\frac{2}{\sqrt{n_{up}}\pi^2 \sqrt{|\Sigma_{down}|}} \frac{1}{\left((X_i - \overline{X_{down}}) \Sigma_{down}{}^{-1} (X_i - \overline{X_{down}})^T \right)^2} \right)_{Down}$$

- In the no absorber case the Upstream and Downstream densities will then match if:

$$\left(\sqrt{|\Sigma_{up}|} \left((X_i - \overline{X_{up}}) \Sigma_{up}{}^{-1} (X_i - \overline{X_{up}})^T \right)^2 \right)_{Up} = \left(\sqrt{|\Sigma_{down}|} \left((X_i - \overline{X_{down}}) \Sigma_{down}{}^{-1} (X_i - \overline{X_{down}})^T \right)^2 \right)_{Down}$$

Separating the Upstream Covariance matrix

Let the full upstream sample be denoted by:

$$\Sigma_1 = \sum_{1}^{N_1} (P_i - \bar{P}_1)(P_i - \bar{P}_1)/N_1$$

where P_i is the Phase Space vector (x, y, z, px, py, pz) and N_1 is the sample size. The Upstream sample which makes it downstream and the sample which doesn't make it downstream are respectively denoted by:

$$\Sigma_2 = \sum_{1}^{N_2} (P_i - \bar{P}_2)(P_i - \bar{P}_2)/N_2\,, \qquad \Sigma_3 = \sum_{1}^{N_3} (P_i - \bar{P}_3)(P_i - \bar{P}_3)/N_3$$

$$with \qquad N_1 = N_2 + N_3 \qquad \text{and} \qquad \bar{P}_1 = \frac{N_2\bar{P}_2 + N_3\bar{P}_3}{N_2 + N_3}$$

Then

$$\Sigma_1 = \sum_{1}^{N_1} \left(P_i - \frac{N_2\bar{P}_2 + N_3\bar{P}_3}{N_2 + N_3}\right)\left(P_i - \frac{N_2\bar{P}_2 + N_3\bar{P}_3}{N_2 + N_3}\right)/(N_2 + N_3)$$

Separating the Upstream Covariance matrix

$$\Sigma_1 = \sum_1^{N_1} \left(P_i - \frac{N_2\bar{P}_2 + N_3\bar{P}_3}{N_2 + N_3} \right) \left(P_i - \frac{N_2\bar{P}_2 + N_3\bar{P}_3}{N_2 + N_3} \right) / (N_2 + N_3)$$

Multiply above and below by $(N_2 + N_3)^2$

$$= \sum_1^{N_1} ((N_2 + N_3)P_i - N_2\bar{P}_2 + N_3\bar{P}_3)((N_2 + N_3)P_i - N_2\bar{P}_2 + N_3\bar{P}_3)/(N_2 + N_3)^3$$

$$= \sum_1^{N_1} \Big(N_2^2(P_i - \bar{P}_2)(P_i - \bar{P}_2) + N_2N_3(P_i - \bar{P}_2)(P_i - \bar{P}_3) + N_2N_3(P_i - \bar{P}_3)(P_i - \bar{P}_2) + N_3^2(P_i - \bar{P}_3)(P_i - \bar{P}_3) \Big) / (N_2 + N_3)^3$$

Multiply out and reorder

$$= \sum_1^{N_2} \Big(N_2^2(P_i - \bar{P}_2)(P_i - \bar{P}_2) + N_2N_3(P_i - \bar{P}_2)(P_i - \bar{P}_3) + N_2N_3(P_i - \bar{P}_3)(P_i - \bar{P}_2) + N_3^2(P_i - \bar{P}_3)(P_i - \bar{P}_3) \Big) / (N_2 + N_3)^3$$
$$+ \sum_1^{N_3} \Big(N_2^2(P_i - \bar{P}_2)(P_i - \bar{P}_2) + N_2N_3(P_i - \bar{P}_2)(P_i - \bar{P}_3) + N_2N_3(P_i - \bar{P}_3)(P_i - \bar{P}_2) + N_3^2(P_i - \bar{P}_3)(P_i - \bar{P}_3) \Big) / (N_2 + N_3)^3$$

Separate the sum into sums over N_2 and N_3

Separating the Upstream Covariance matrix

This is the general form partially separating out the covariance matrices

$$\Sigma_1 = \frac{N_2^3}{(N_2+N_3)^{\,3}}\Sigma_2 + \frac{N_3^3}{(N_2+N_3)^{\,3}}\Sigma_3 + \sum_1^{N_2}\Big(N_2N_3(P_i-\bar{P}_2)(P_i-\bar{P}_3) + N_2N_3(P_i-\bar{P}_3)(P_i-\bar{P}_2) + N_3^2(P_i-\bar{P}_3)(P_i-\bar{P}_3)\Big)/(N_2+N_3)^3 + \sum_1^{N_3}\Big(N_2N_3(P_i-\bar{P}_2)(P_i-\bar{P}_3) + N_2N_3(P_i-\bar{P}_3)(P_i-\bar{P}_2) + N_2^2(P_i-\bar{P}_2)(P_i-\bar{P}_2)\Big)/(N_2+N_3)^3$$

For a symmetric absorber and cooling channel (LiH, lH2, No absorber): $\bar{P}_2 \approx \bar{P}_3$, then

$$\Sigma_1 = \frac{N_2^3}{(N_2+N_3)^{\,3}}\Sigma_2 + \frac{N_3^3}{(N_2+N_3)^{\,3}}\Sigma_3 + \sum_1^{N_2}\Big(2N_2N_3(P_i-\bar{P}_2)(P_i-\bar{P}_2) + N_3^2(P_i-\bar{P}_2)(P_i-\bar{P}_2)\Big)/(N_2+N_3)^3 + \sum_1^{N_3}\Big(2N_2N_3(P_i-\bar{P}_3)(P_i-\bar{P}_3) + N_2^2(P_i-\bar{P}_3)(P_i-\bar{P}_3)\Big)/(N_2+N_3)^3$$

Substituting for $\bar{P}_2$ and $\bar{P}_3$ in their sums

$$= \Sigma_2\left(\frac{N_2^3 + 2N_2^2N_3 + N_2N_3^2}{(N_2+N_3)^3}\right) + \Sigma_3\left(\frac{N_3^3 + N_2^2N_3 + 2N_2N_3^2}{(N_2+N_3)^3}\right)$$

Separating the Upstream Covariance matrix

$$\Sigma_1 = \Sigma_2\left(\frac{N_2^3 + 2N_2^2N_3 + N_2N_3^2}{(N_2+N_3)^3}\right) + \Sigma_3\left(\frac{N_3^3 + N_2^2N_3 + 2N_2N_3^2}{(N_2+N_3)^3}\right)$$
$$= \Sigma_2\left(\frac{N_2}{N_2+N_3}\right) + \Sigma_3\left(\frac{N_3}{N_2+N_3}\right)$$

Therefore

$$N_1\Sigma_1 = N_2\Sigma_2 + N_3\Sigma_3$$

For a radially symmetric absorber (ignoring dissipative forces) the upstream distribution can be separated into the covariance matrix of the sample which makes it downstream and missing sample weighted by their sample sizes.

No absorber Covariance matrices

Full Upstream Sample (Σ_1)			
2320.82	182.66	56.49	-786.46
182.66	2427.89	675.30	163.87
56.49	675.30	829.47	-60.89
-786.46	163.87	-60.89	773.75

Upstream Sample that made it downstream(Σ_2)			
1518.43	85.98	90.92	-604.66
85.98	1477.41	518.44	73.55
90.92	518.44	711.03	-59.26
-604.66	73.55	-59.26	639.62

Full Downstream Sample (Σ_3)			
3712.64	349.76	-3.34	-1102.37
349.76	4077.46	947.56	320.47
-3.34	947.56	1035.08	-63.73
-1102.37	320.47	-63.73	1006.53

Recombined Upstream Sample ($\Sigma_2\left(\frac{N_2}{N_1}\right)+\Sigma_3\left(\frac{N_3}{N_1}\right)$)			
2320.42	182.39	56.46	-786.58
182.39	2427.71	675.28	163.80
56.46	675.28	829.47	-60.90
-786.58	163.80	-60.90	773.72

The determinant of a matrix

- The determinant of a matrix can be separated into parts using:

$$|\Sigma_1| = \sum_{i=0}^{n} \Gamma_n^i \left| {}^{\Sigma_2}\!/_{\Sigma_3}{}^{i} \right| = |\Sigma_2| + |\Sigma_3| + \sum_{i=1}^{n-1} \Gamma_n^i \left| {}^{\Sigma_2}\!/_{\Sigma_3}{}^{i} \right|$$

Where Γ_n^i represents substituting all combinations of i^{th} lines from Σ_2 by the same lines in Σ_3 and taking the subsequent determinant of the new matrix

- For the symmetric case (LiH, LH2 and no absorber) the previous and above substitutions could be made to compare the upstream and downstream densities. Due to the asymmetry this cannot be done for the wedge and requires further derivation for the asymmetric case.

Potential next step

- The missing data downstream is inaccessible, however the upstream sample which makes it downstream can be compared to the downstream sample
- The transport, M, of a covariance matrix from upstream to downstream can be given by:

$$\Sigma_{down} = \langle X_{down}\tilde{X}_{down}\rangle = \langle MX_{up}\tilde{M}\tilde{X}_{up}\rangle = M\langle X_{up}\tilde{X}_{up}\rangle\tilde{M} = M\Sigma_{up}\tilde{M}$$

- The determinant is given by:

$$|\Sigma_{down}| = \left|M\Sigma_{up}\tilde{M}\right| = |M|^2\left|\Sigma_{up}\right| = \left|\Sigma_{up}\right|$$

- The transfer matrix M has been previously investigated by Sophie Middleton and Chris Rogers
- A potential investigation would be to investigate the change in R for different fraction sizes of the beam. If stable it could be used to investigate the missing data downstream to see if it is due to scraping and magnet misalignment affects and nothing else

R Matrix

- The transfer Matrix components are then used to get the downstream distribution through:

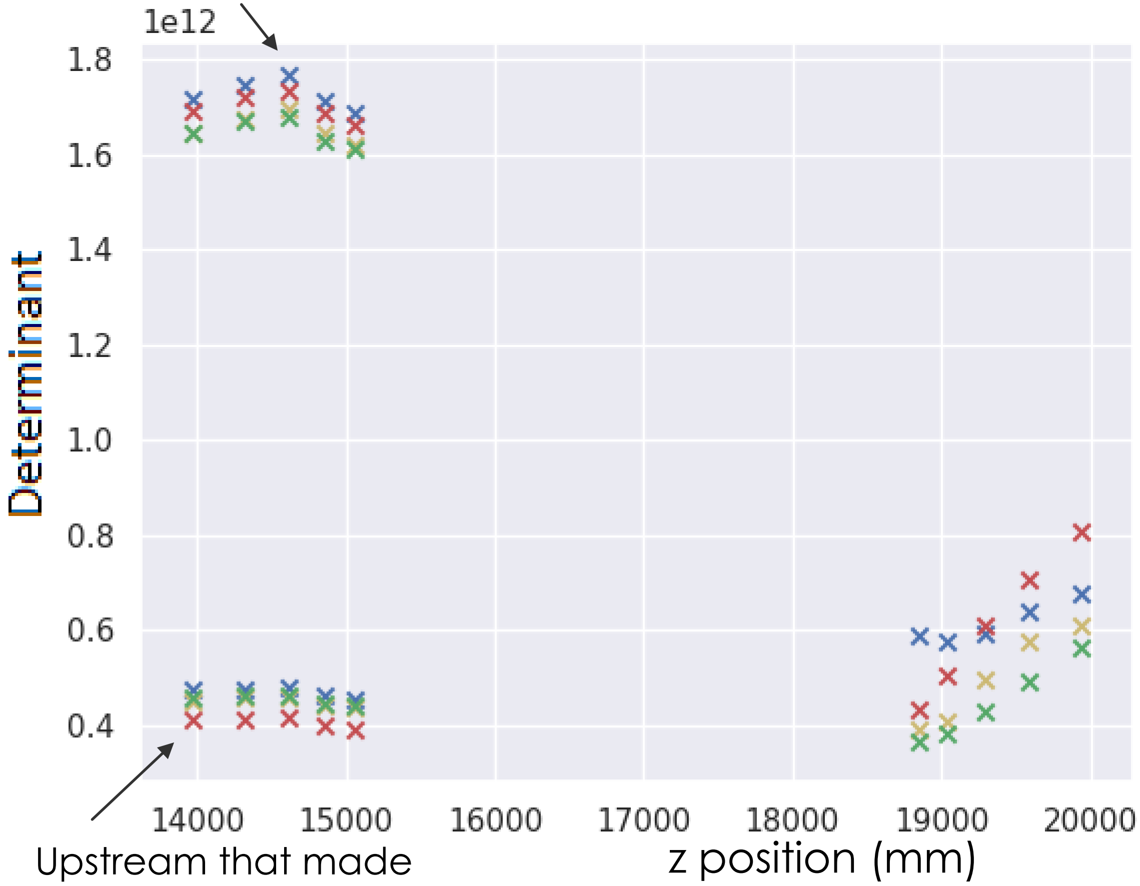
$$
\begin{aligned}
x_d &= M_{00} + M_{01}x_u + M_{02}x_u' + M_{03}y_u + M_{04}y_u' \\
x_d' &= M_{10} + M_{11}x_u + M_{12}x_u' + M_{13}y_u + M_{14}y_u' \\
y_d &= M_{20} + M_{21}x_u + M_{22}x_u' + M_{23}y_u + M_{24}y_u' \\
y_d' &= M_{30} + M_{31}x_u + M_{32}x_u' + M_{33}y_u + M_{34}y_u'
\end{aligned}
$$

- Note this assumes the tracker stations are parallel with no rotations. In this case many of the Matrix components are simply zero. Corrections can be made for these if that is not the case.
- Corrections may also need to be added for any deviations in the magnetic field in x and y. Further corrections could also be added to include higher order effects. This also assumes the transverse only description is valid, but it may need to be extended to a full 6D description. The main problem may then be large errors, where it may become difficult to discern anything.
- If the error for the transfer of individual particles proves too large, another way too look at it would be through the transfer of moments

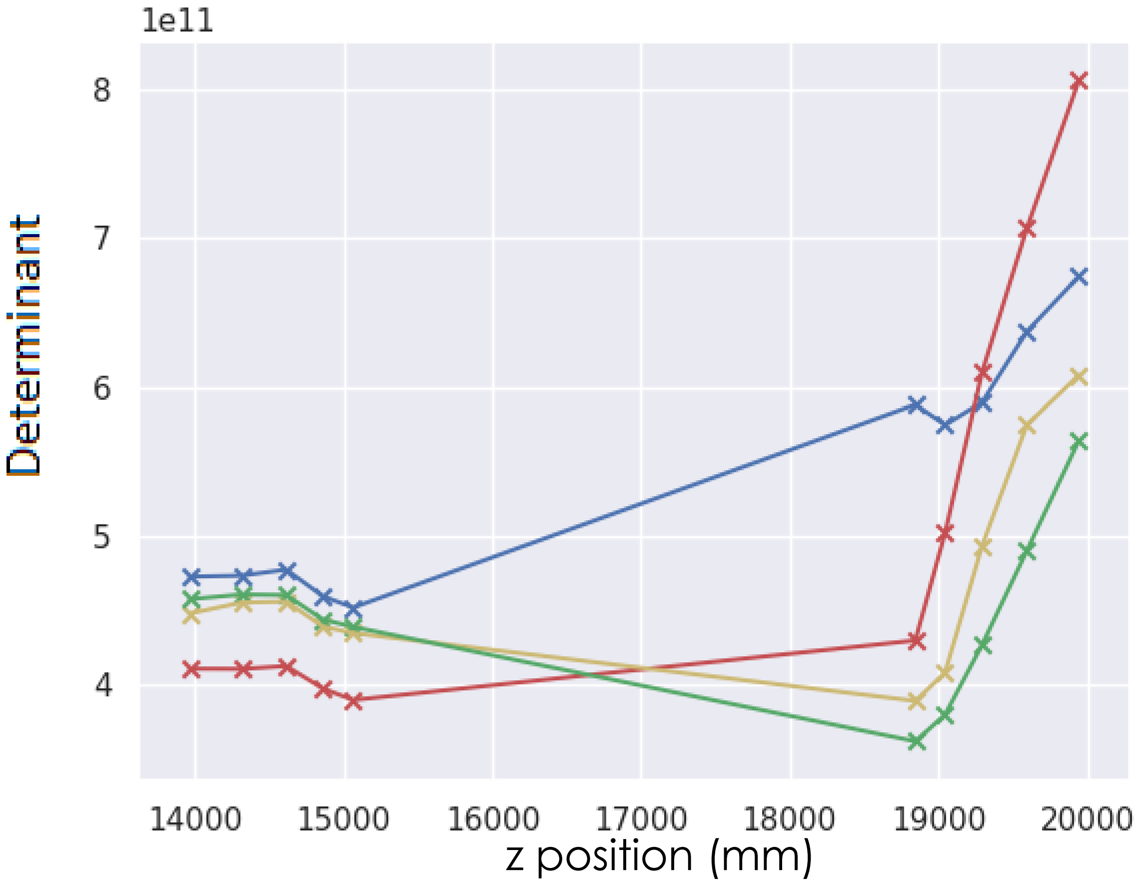
Caution: Change in the Determinant of the Covariance Matrix through the Cooling Channel

Blue – No Absorber
Red – Wedge
Yellow – LiH
Green – LH2

Full Upstream



Upstream that made it Downstream



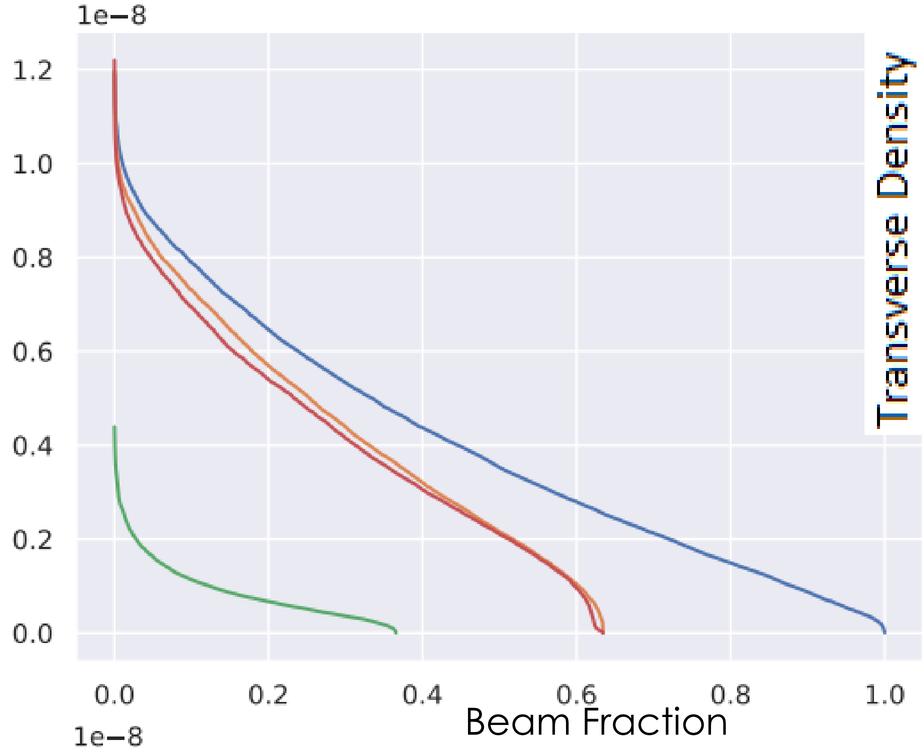
Change in Determinant

- For the absorber case it increases by a significant fraction
- None symplectic transverse transport may be due to inclusion of scattered particles, can try to use chi-squared cut to try to exclude them
- Determinant increases through the Downstream trackers (last two stations), likely due to M2 coil, may be getting non-transverse magnetic field effects. Can’t separate transverse from longitudinal anymore

- For other three absorbers the determinant begins to significantly increase through the downstream trackers (spin?)
- The LH2 Determinant is slightly lower than the LiH determinant (better cooling perhaps)
- Caution: Biased sample. Biased by downstream selection

Biased Cooling

- Cooling can be shown in a number of different ways
- One is to through cumulative plots showing what fraction of the beam is above a certain density
- If the downstream line is above the upstream line it shows cooling as the phase space density has increased.
- The opposite is the case for heating
- It is highly affected by transmission losses



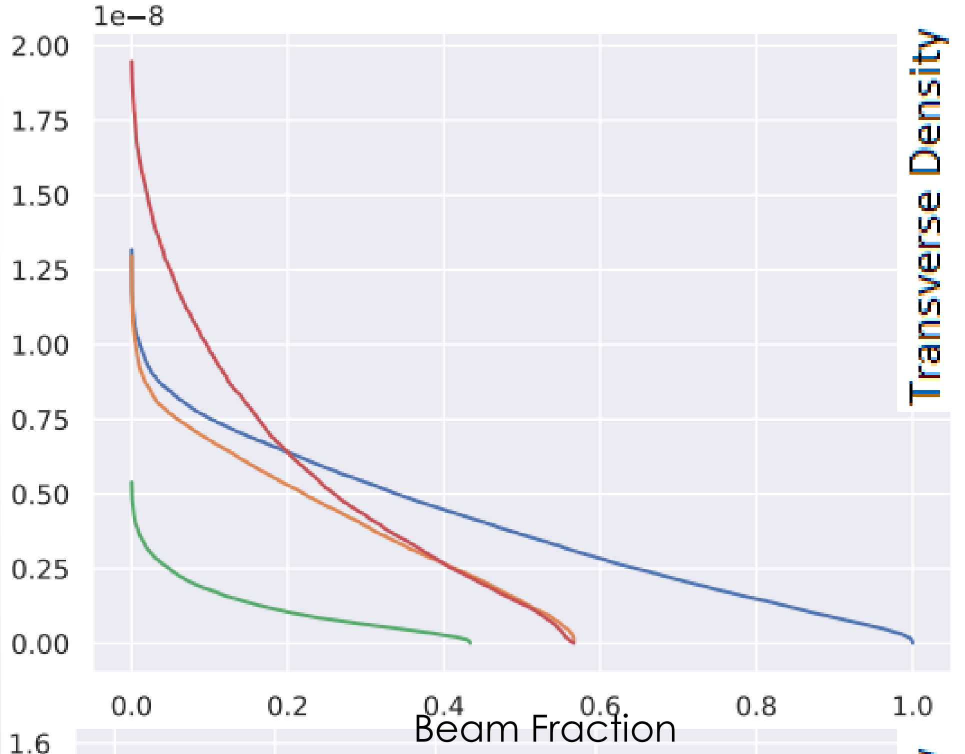
Fraction of beam above certain density

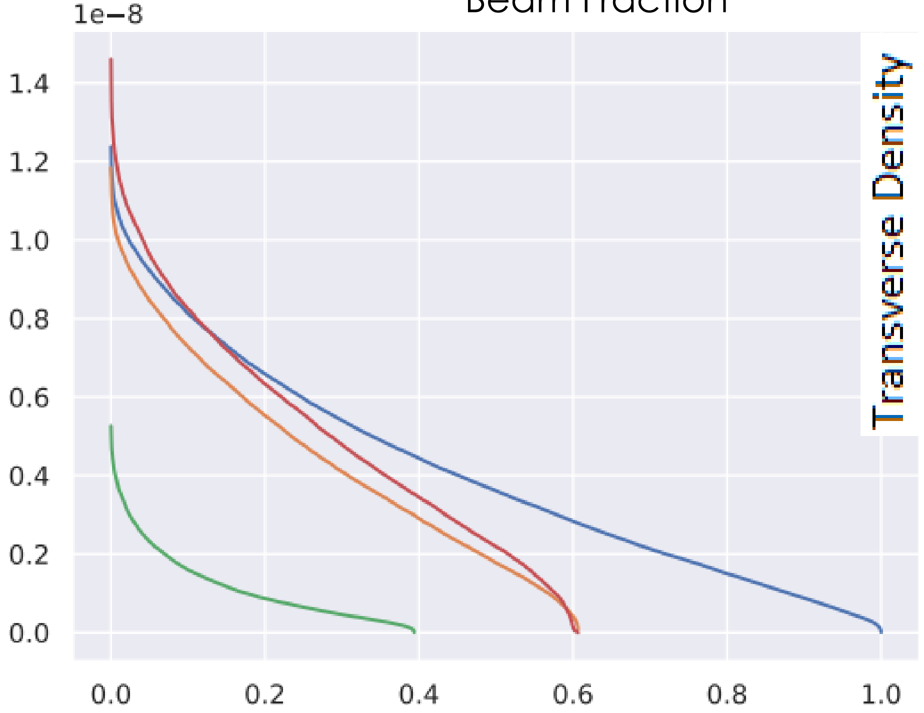
Top Left: No absorber

Top Right: Wedge

Bottom Left: LiH

Bottom Right: LH2



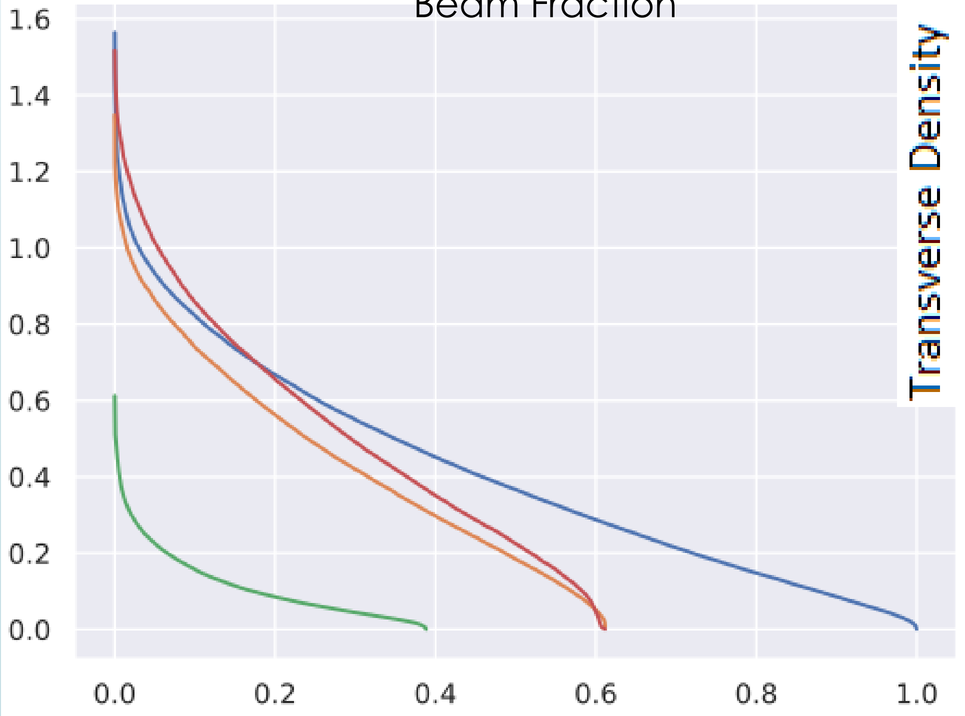


Blue – Full Upstream Sample

Red – Full Downstream Sample

Orange – Upstream Sample which made it Downstream

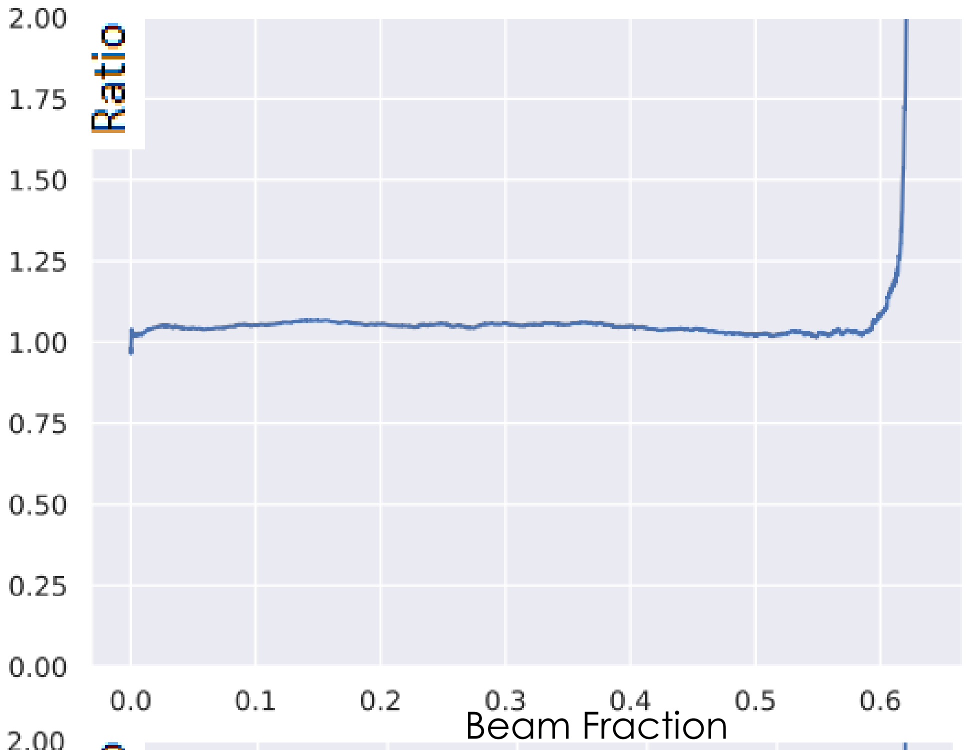
Green – Upstream Sample which doesn't make it downstream



Biased Cooling

- Another way to show cooling is to take the ratio of the fraction of the upstream sample which makes it downstream to the downstream sample. This should remain constant across the whole fraction of the beam for the symmetric absorber cases. For the wedge the ratio will be proportional to the thickness traversed
- It is biased however as it does not contain the full beam
- Particles are eliminated by the absorber, and some particles make it downstream because of the absorber

Ratio of the Downstream density to the Upstream density which makes it downstream

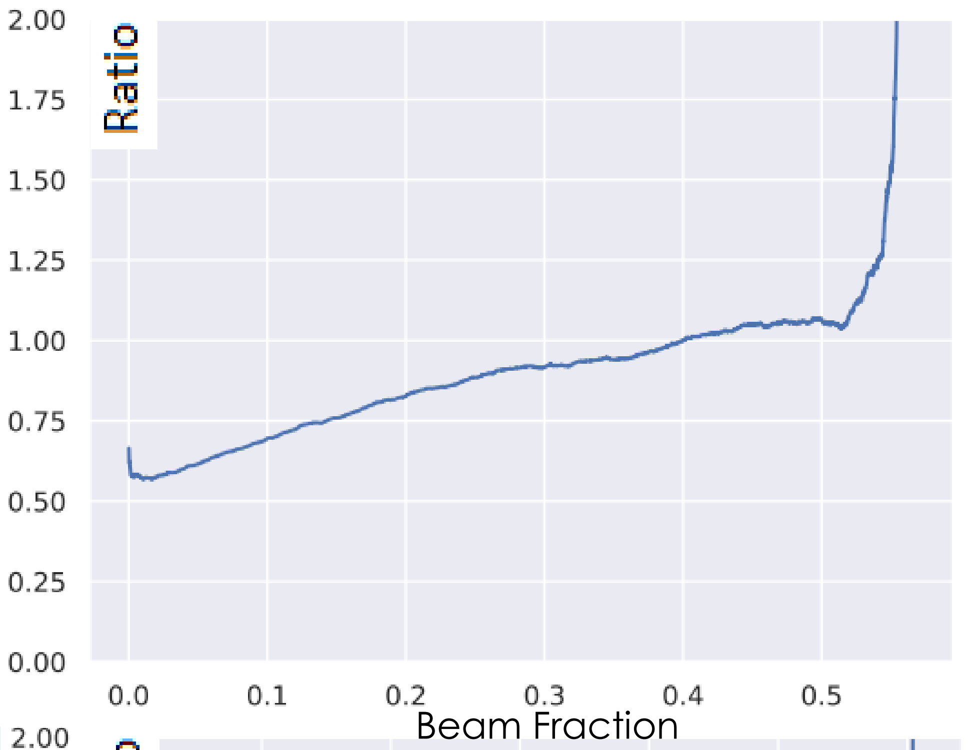


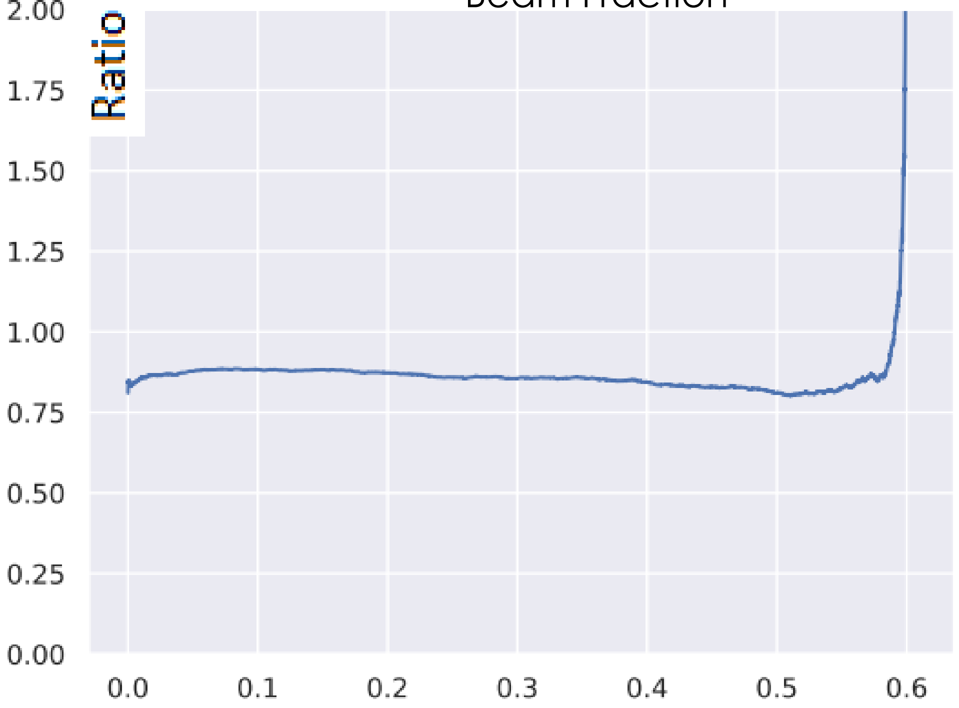
Top Left: No absorber

Top Right: Wedge

Bottom Left: LiH

Bottom Right: LH2

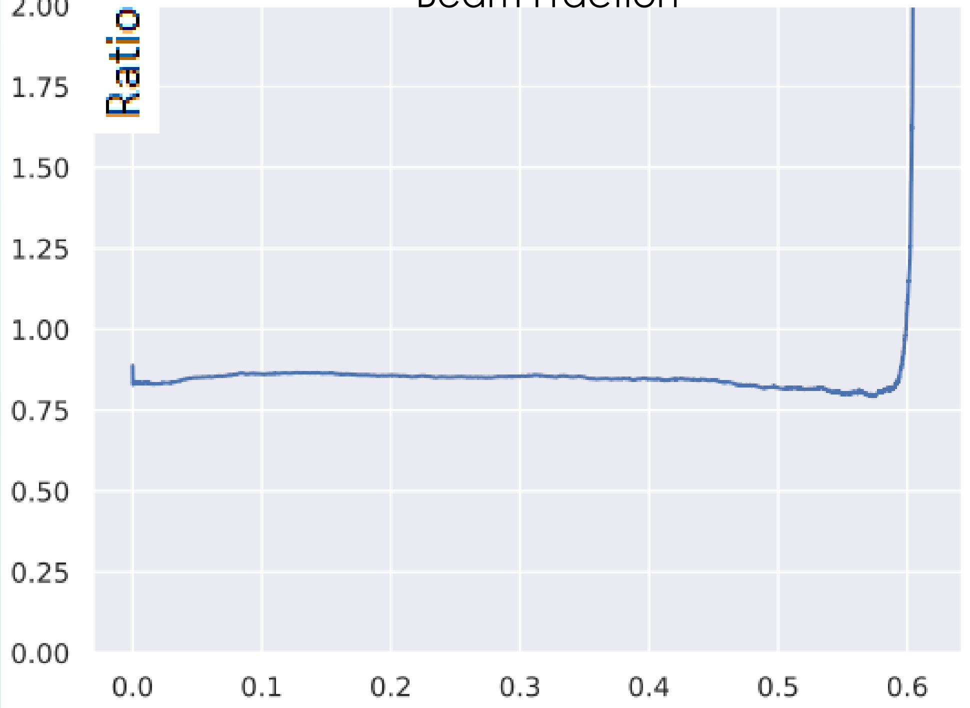




Ratio above one indicates heating while a ratio below one indicates cooling.

Transmission limits the beam to approximately 60% of the full upstream sample.

The min and max are limited by low sample size and scraping respectively



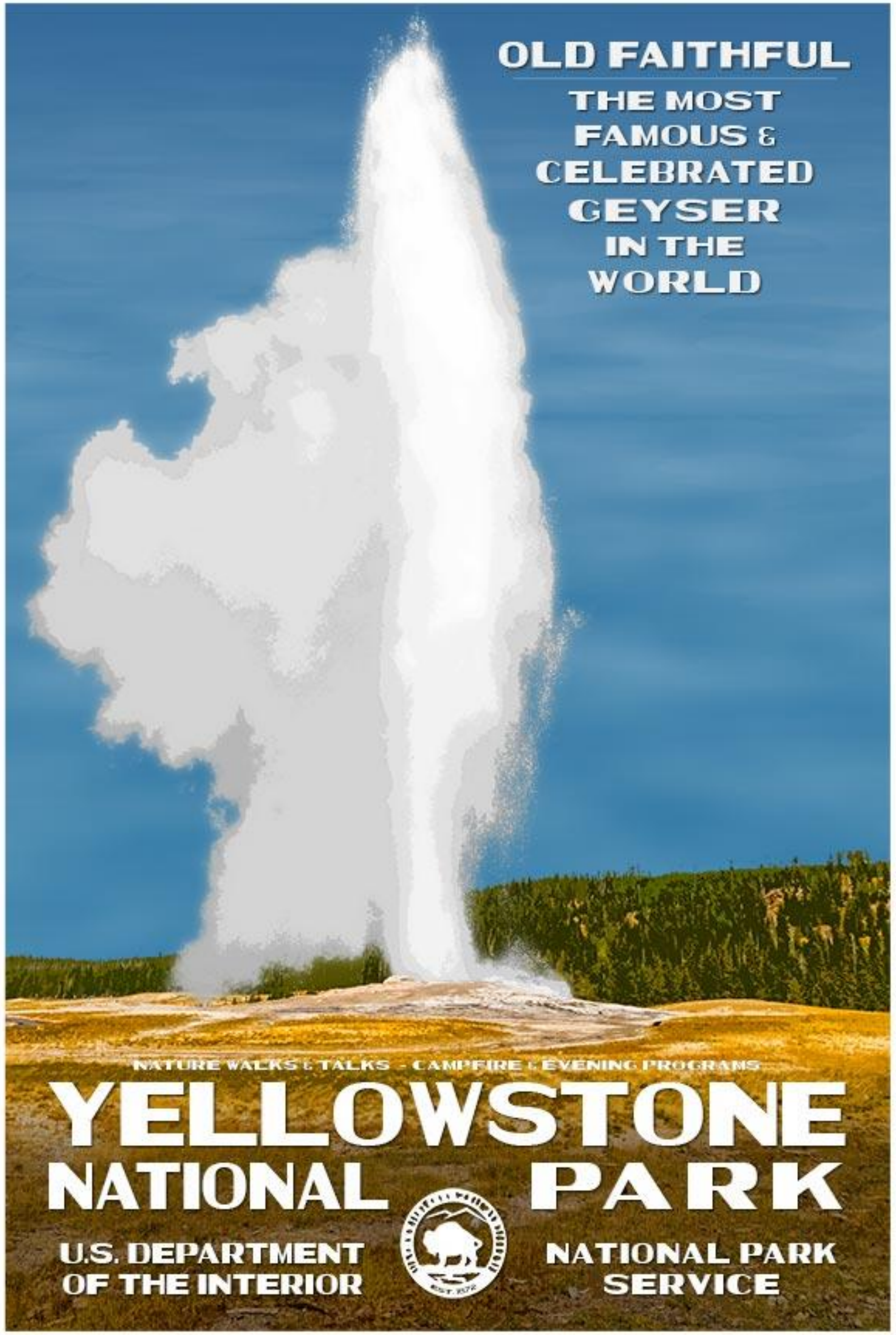
Conclusion

- KDE and KNN behave the same way
- The phase space density is calculated for a given volume
- Discrepancies seen before were due to changes in phase space volume from upstream to downstream by transmission losses
- The transmission losses are non-random
- This causes the normalization procedure for the downstream sample to underestimate the cooling performance

- The determinant of the transverse covariance matrix changes through the cooling channel (need to check it is only a scraping effect)
- Biased selection of only full tracks shows constant cooling across the beam fraction for LiH and LH2.
- The Wedge shows sloped cooling as the beam fraction increases. Need to check if asymmetries cause bias. Can then move onto longitudinal and 6D.
- Find a way to express downstream density in terms of upstream phase space volume in an unbiased way

THE END

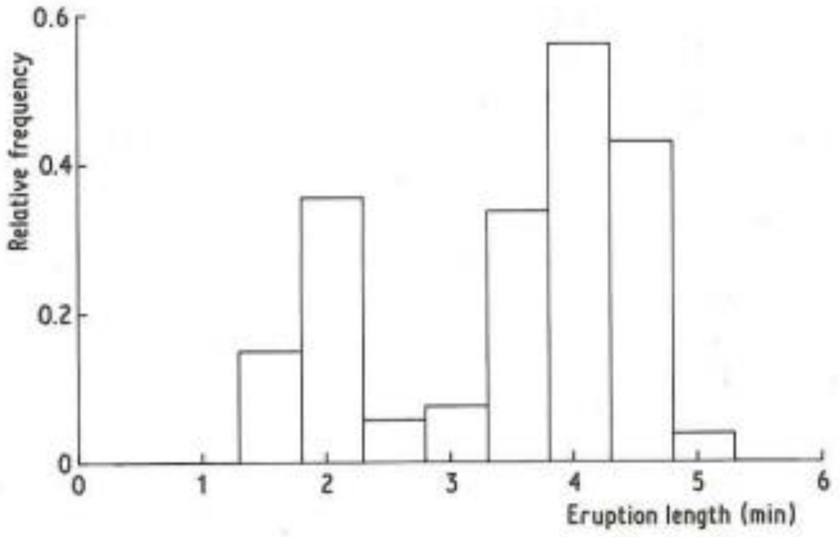
Extra Slides

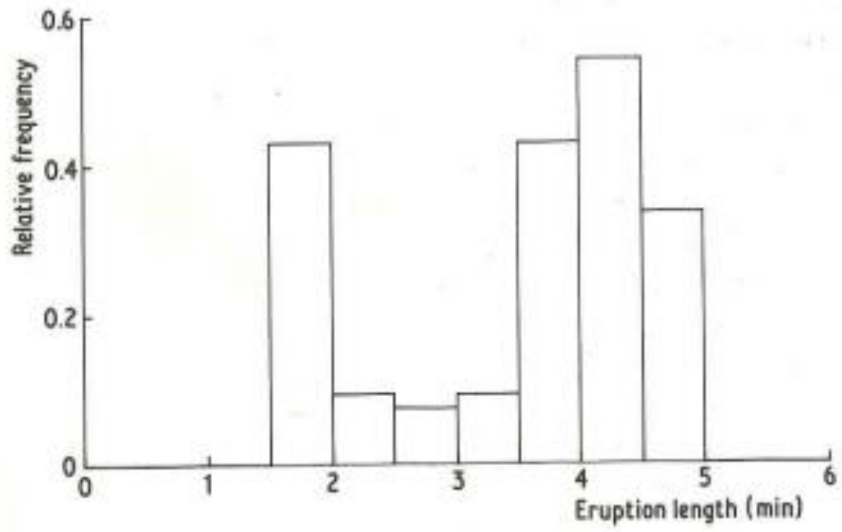


Histograms, KDE and KNN
Old Faithful Geyser Eruptions

- Highly predictable geothermal feature
- Spews boiling hot water 100 – 180 feet into the air
- Erupts 20 times a day. Eruptions can be predicted to within a 90% accuracy in a 10 minute interval
- Eruptions typically last 1.5 to 5 minutes
- Shows distinct bimodal feature
- The following will look at a sample of the data which should follow the parent distribution i.e. all eruptions in time
- This will be basis to determine if a density estimate follows the true underlying density

Histograms, KDE and KNN – Basics (from Silverman)





- Probability density function gives the probability a quantity is found in the interval:

$$P(a < X < b) = \int_a^b f(x)dx \qquad for\ all\ a < b$$

- The m^{th} histogram interval for origin x_0 and bin width h is given by:

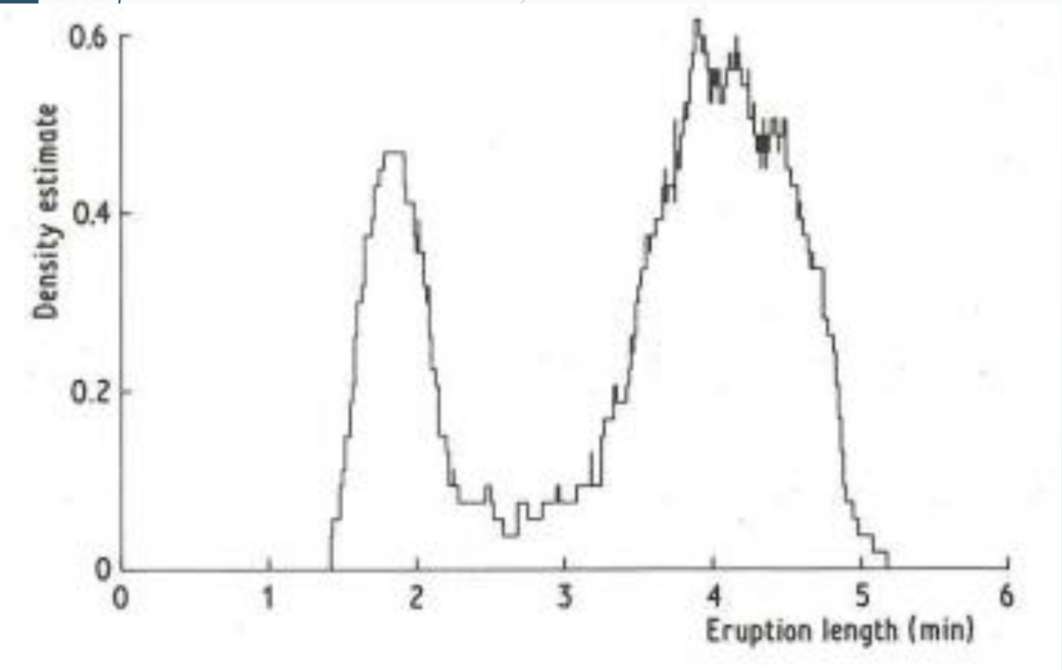
$$[x_0 + mh, x_0(m+1)h)$$

- The histogram is then defined by:

$$\hat{f}(x) = \frac{1}{nh}(no.of\ X_i\ in\ the\ same\ bin\ as\ x)$$

- Choice of origin and bin width can give "apparent structure effects" that are due to random error
- Discontinuity of histograms can cause difficulty if derivatives of the estimate are required

Naive Estimator



- For a random variable X with density f, then:

$$f(x) = \lim_{h \to 0} \frac{1}{2h} P(x - h < X < x + h)$$

- Then

$$\hat{f}(x) = \frac{1}{2hn}[no. of\ X_i, \dots, X_n\ falling\ in\ (x-h, x+h)]$$

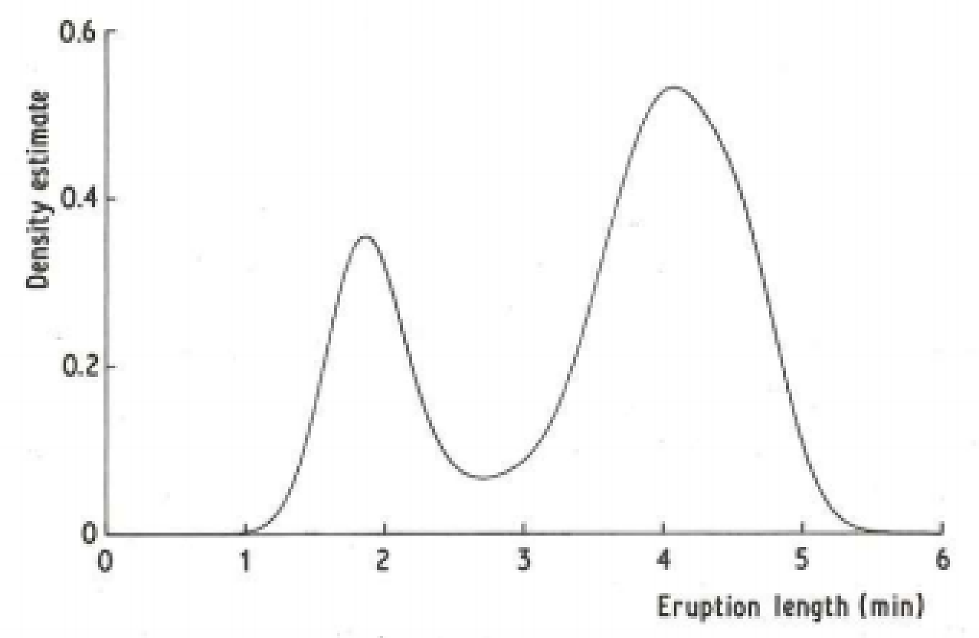
- Define weight function w by

$$w(x) = \begin{cases} 1/2 & if\ |x| < 1 \\ 0 & otherwise \end{cases}$$

- Then

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h} w\left(\frac{x - X_i}{h}\right)$$

Kernel Estimator



- The kernel estimator is obtained by replacing the weight function of the naive estimator by a kernel function K satisfying:

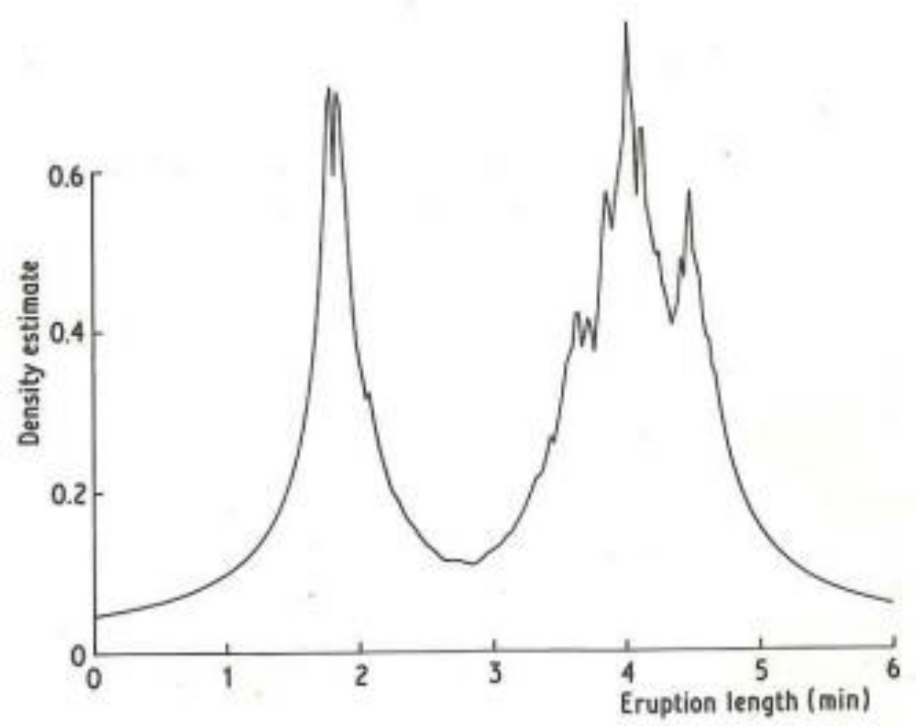
$$\int_{-\infty}^{\infty} K(x)dx = 1$$

- The kernel estimator of bandwidth h is then defined by

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

- Varying the bandwidth h determines the level of smoothing, as h tends to zero, the smoothing becomes a sum of Dirac delta spikes, but if h becomes large, all detail is obscured.
- If K is non-negative everywhere, then $\hat{f}$ itself will be a probability density. The probability density function of the sample has been convolved with the kernel
- This can lead to non-negative tails to naturally positive data, especially when the data distribution is long-tailed. Parameter choice can be used to minimize the undesired effects

K-nearest neighbour



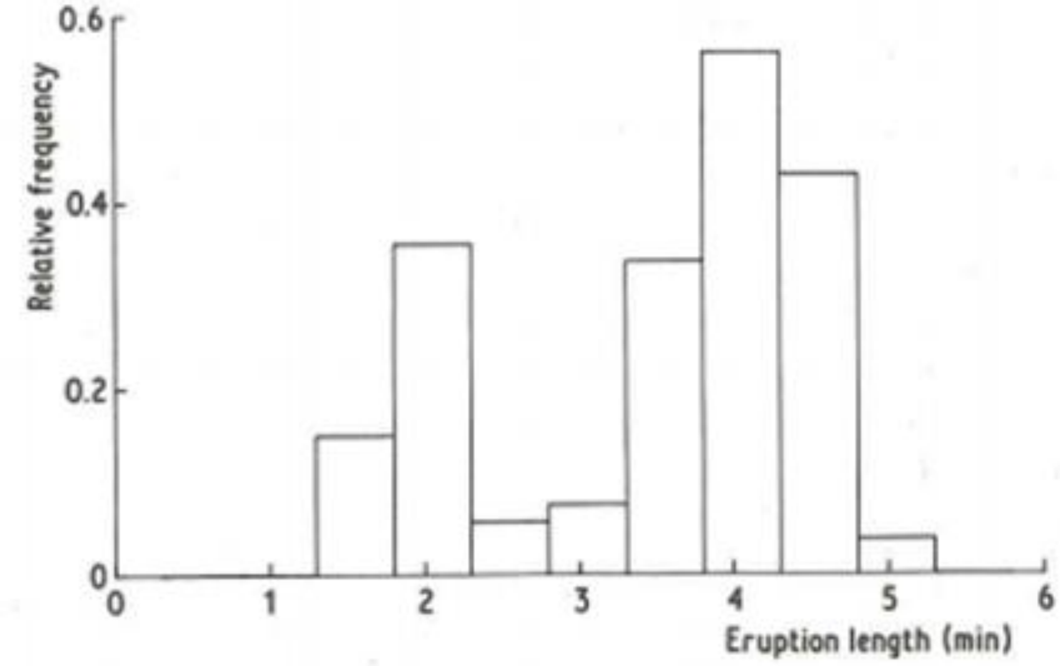
- Define the distance $d(x, y)$ between two points on the line to be $|x - y|$ and for each t define $d_1(t) \leq d_2(t) \leq \cdots \leq d_n(t)$ to be the distances arranged in ascending order.
- The k^{th} nearest neighbour density estimate is then given by
$$\hat{f}(t) = \frac{k-1}{2nd_k(t)}$$
- i.e. $(k-1)$ observations fall in the interval $[t - d_k(t), t + d_k(t)]$
- The nearest neighbour estimate is inversely proportional to the size of box needed to contain it -> undersmoothing in tails should be reduced
- $\hat{f}(t)$ is positive and continuous everywhere, but its derivative will be discontinuous at all the same points as d_k
- The nearest neighbour estimate will not be a probability density (but only an approximation) as it does not integrate to unity
- For t less than the smallest data point, $d_k(t) = X_k - t$ and for $t > X_n$: $d_k(t) = t - X_{(n-k+1)}$, thus $\int_{-\infty}^{\infty} \hat{f}\,(t)dt$ is infinite and the tails of $\hat{f}$ die away slowly

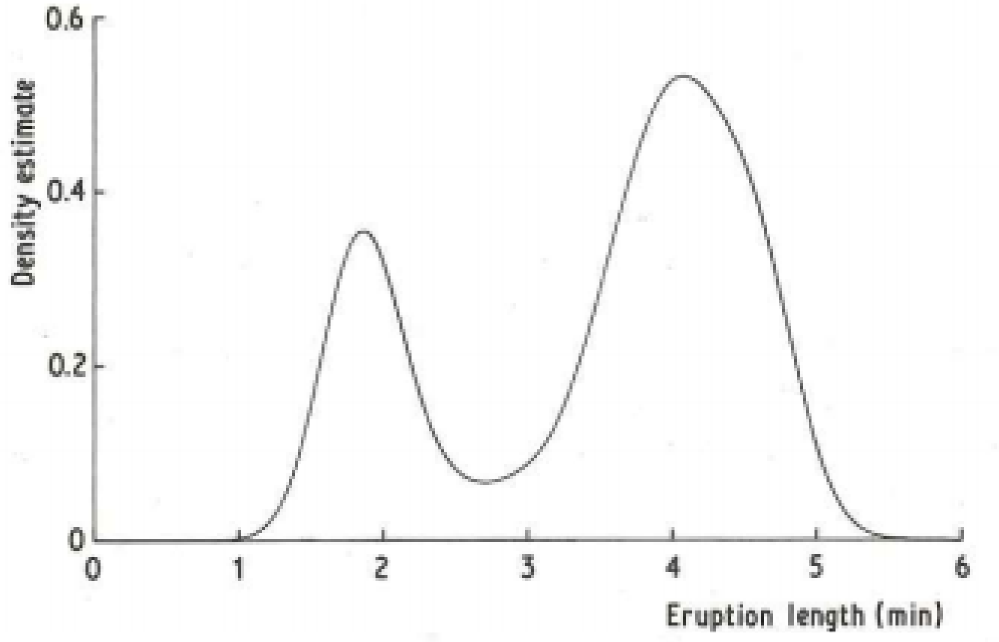
KNN relation to KDE

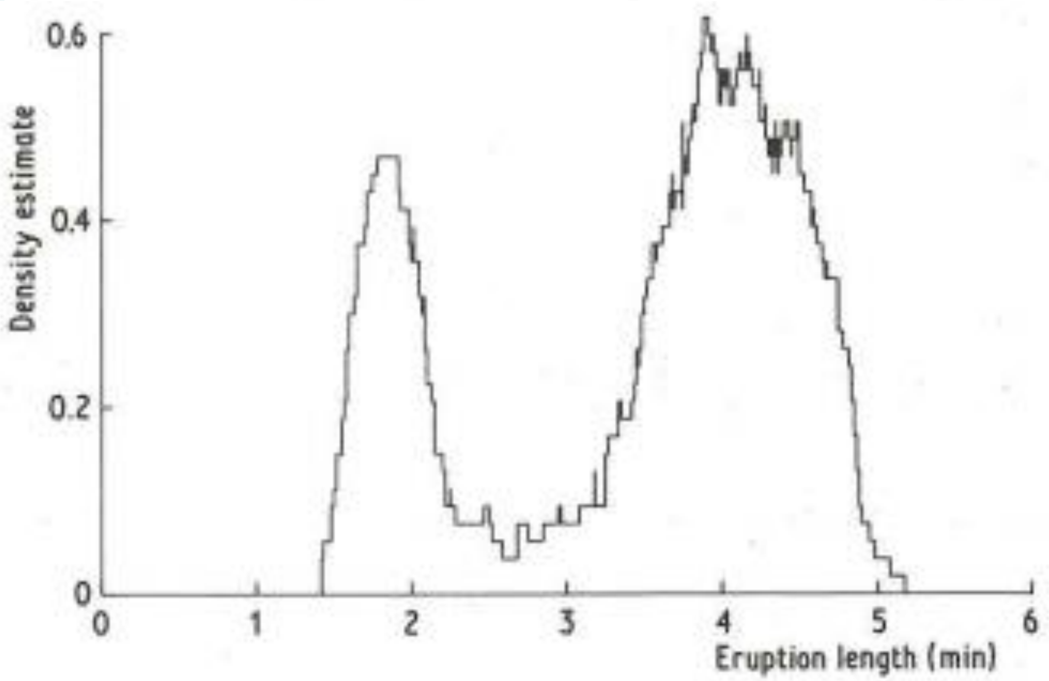
- Let $K(x)$ be a kernel function integrating to one
- The k^{th} nearest neighbour estimate is given by

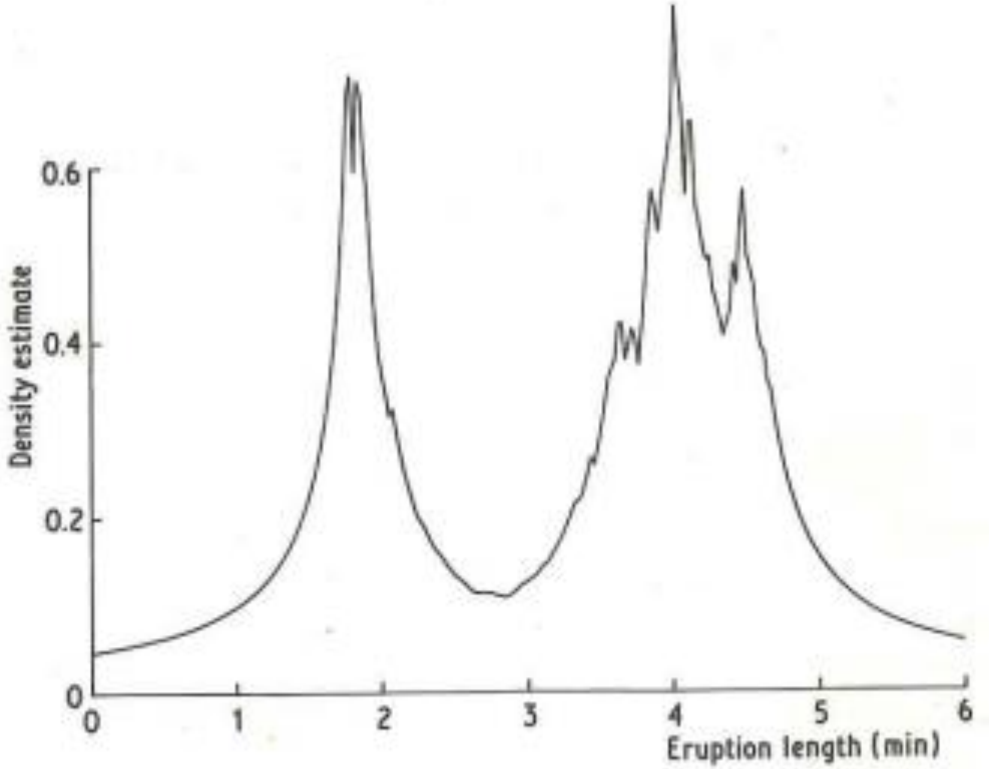
$$\hat{f}(t) = \frac{1}{nd_k(t)} \sum_{i=1}^{n} K\left(\frac{t - X_i}{d_k(t)}\right)$$

- $\hat{f}(t)$ is the kernel estimate evaluated at t with window width $d_k(t)$ where the choice of k governs the smoothing.


Relative frequency
0.6
0.4
0.2
0
0 1 2 3 4 5 6
Eruption length (min)


Density estimate
0.6
0.4
0.2
0
0 1 2 3 4 5 6
Eruption length (min)


Density estimate
0.6
0.4
0.2
0
0 1 2 3 4 5 6
Eruption length (min)


Density estimate
0.6
0.4
0.2
0
0 1 2 3 4 5 6
Eruption length (min)

1-D estimate to n-dimensional estimate

- KDE in 1-D

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

Becomes in n-D:

$$\hat{f}(\vec{x}) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left\{\frac{1}{h}(\vec{x} - \vec{X_i})\right\}$$

where $\int_{R^d} K(\vec{x})dx = 1$ for a d-dimensional space and h^d is the smoothing parameter for each particular dimension. h^d can also be given by a smoothing matrix e.g. the covariance matrix if it is representative of the underlying distribution.

- The choice of kernel only has a minor effect (slightly different efficiencies), and thus a gaussian kernel (most common) will be used to retain the differentiability of $\hat{f}(\vec{x})$. The gaussian kernel is given by:

$$K(\vec{x}) = (2\pi)^{-d/2}\exp(-\frac{1}{2}\vec{x}^T\vec{x})$$

1-D estimate to n-dimensional estimate

- KNN in 1-D

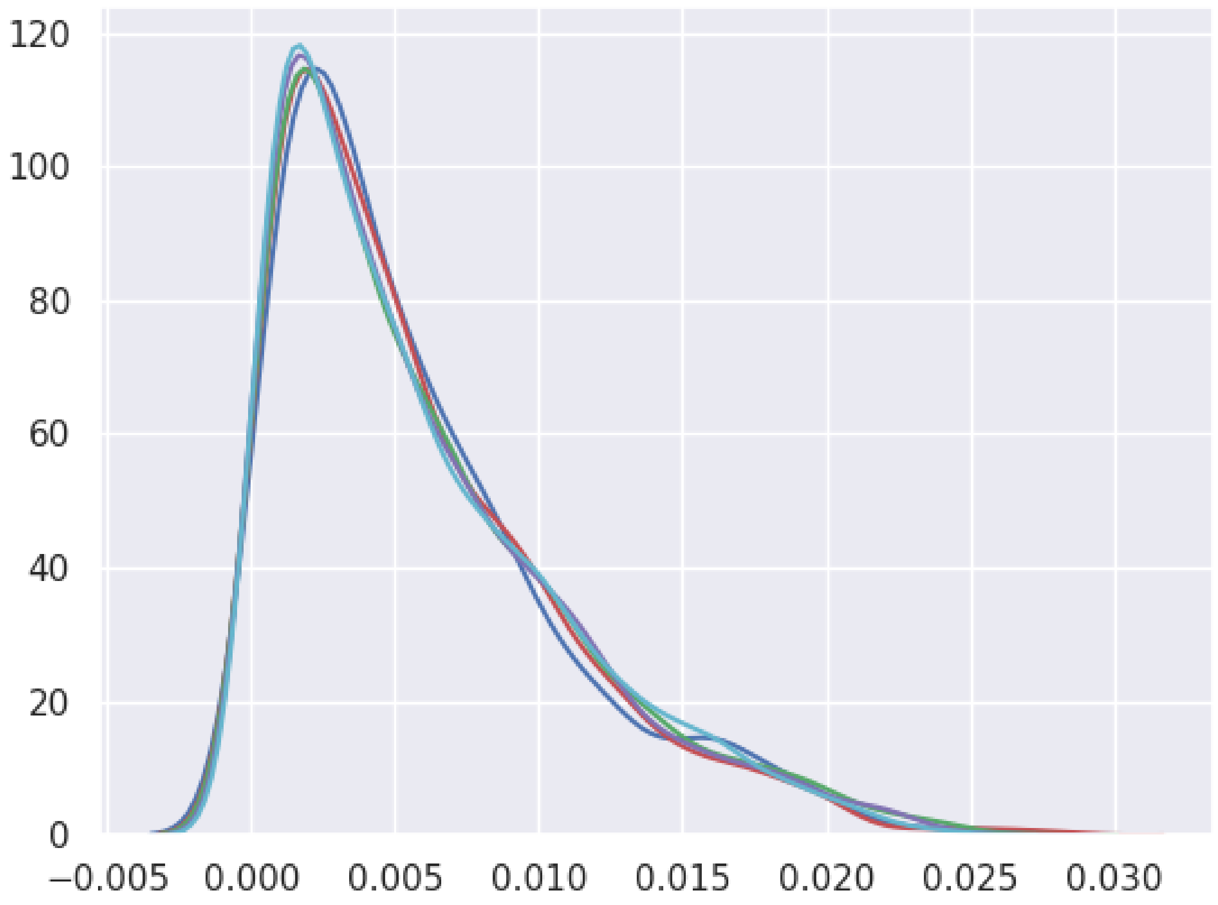
$$\hat{f}(t) = \frac{k-1}{2nd_k(t)}$$

Becomes in n-D (from Francois):

$$\vec{f}(x) = \frac{k}{n\kappa_d R_k^d} = \frac{k\Gamma(\frac{d}{2}+1)}{n\pi^{\frac{d}{2}} R_k^d}$$

where $d_k(t)$ is now the Euclidean distance $R_i = \left|\left|\vec{x} - \vec{x_i}\right|\right| = \sqrt{(\vec{x} - \vec{x_i})^T(\vec{x} - \vec{x_i})}$, κ_d is the volume of a unit d-ball (in 1-D it is equal to two), $\Gamma(\frac{d}{2}+1)$ is Euler's gamma function, while k and (k-1) differ due to counting conventions of whether the test point is included.

Change in Sample Size – Toy Scenario



- See effect of change in sample size, as sample size increases, should approach underlying density of sample
- Random 4-D distribution with mean = 0, Standard Deviation = diag(1,1,1,1)

Blue: n =1000, k = 31

Red: n = 2000, k = 44

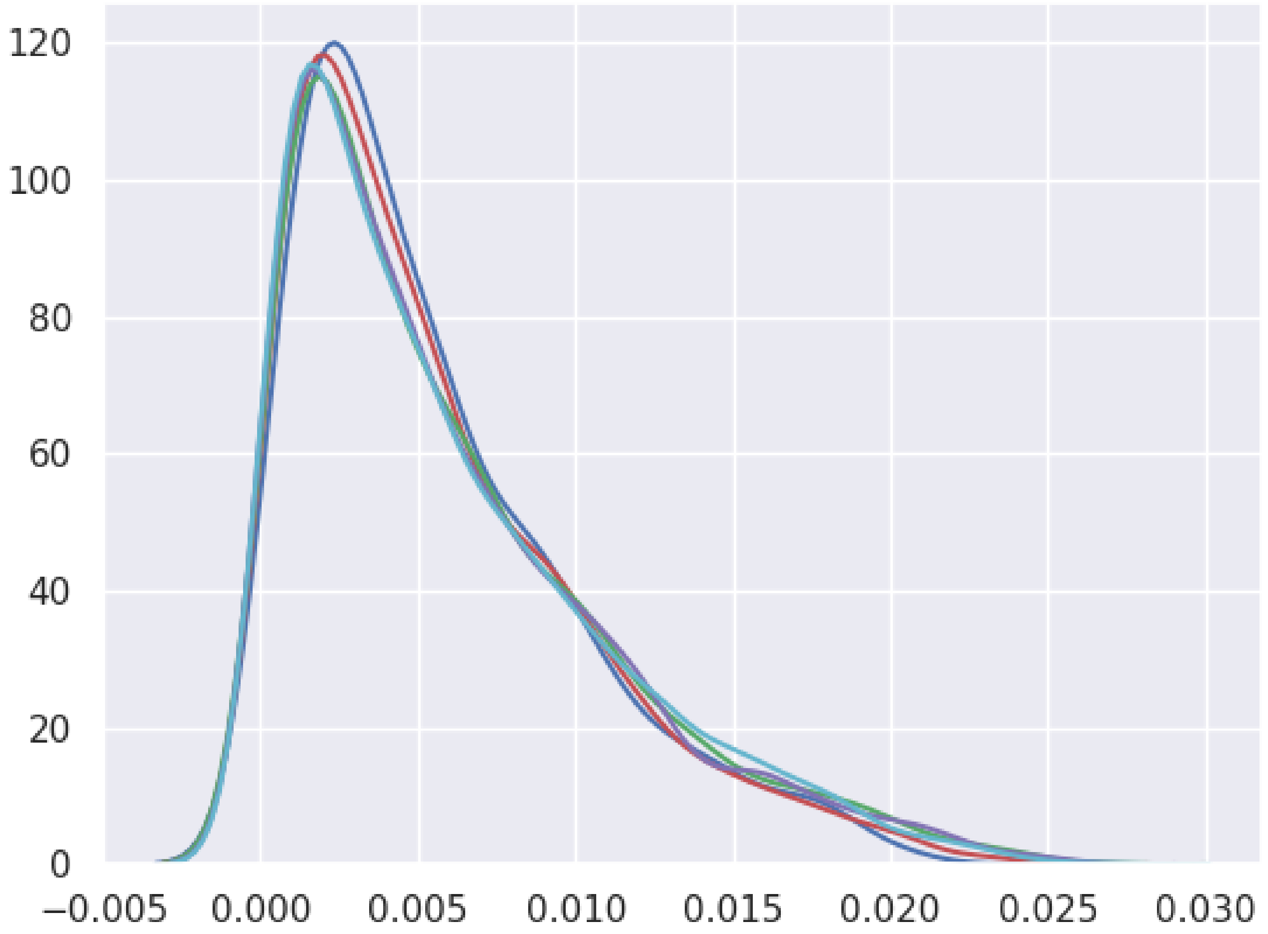
Green: n = 3000, k = 54

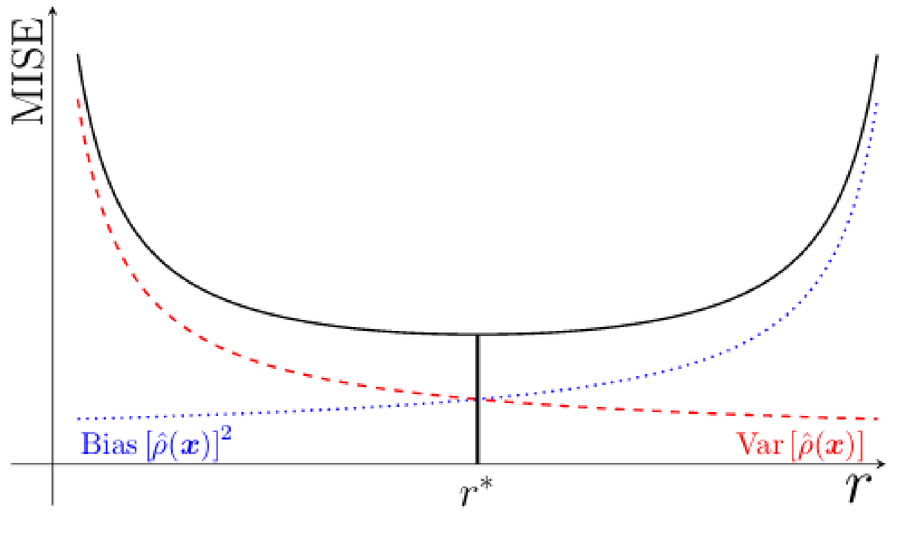
Magenta: n = 4000, k = 63

Cyan: n = 5000, k = 70

- Underlying sample density is approached as sample size increases, optimal k adjusts to reflect increase in sample size

Change in sample size, same k – Toy Scenario





- Changing sample size but keeping k constant increases MISE, as a suboptimal k is chosen
- D-dimensional radius for a test point increases/decreases as the test point needs to find more/less neighbours. This can give an apparent decrease/increase in the phase space density. As the sample size is increased, the phase space density becomes less susceptible to small changes in optimal k

Blue: n =1000, k = 54

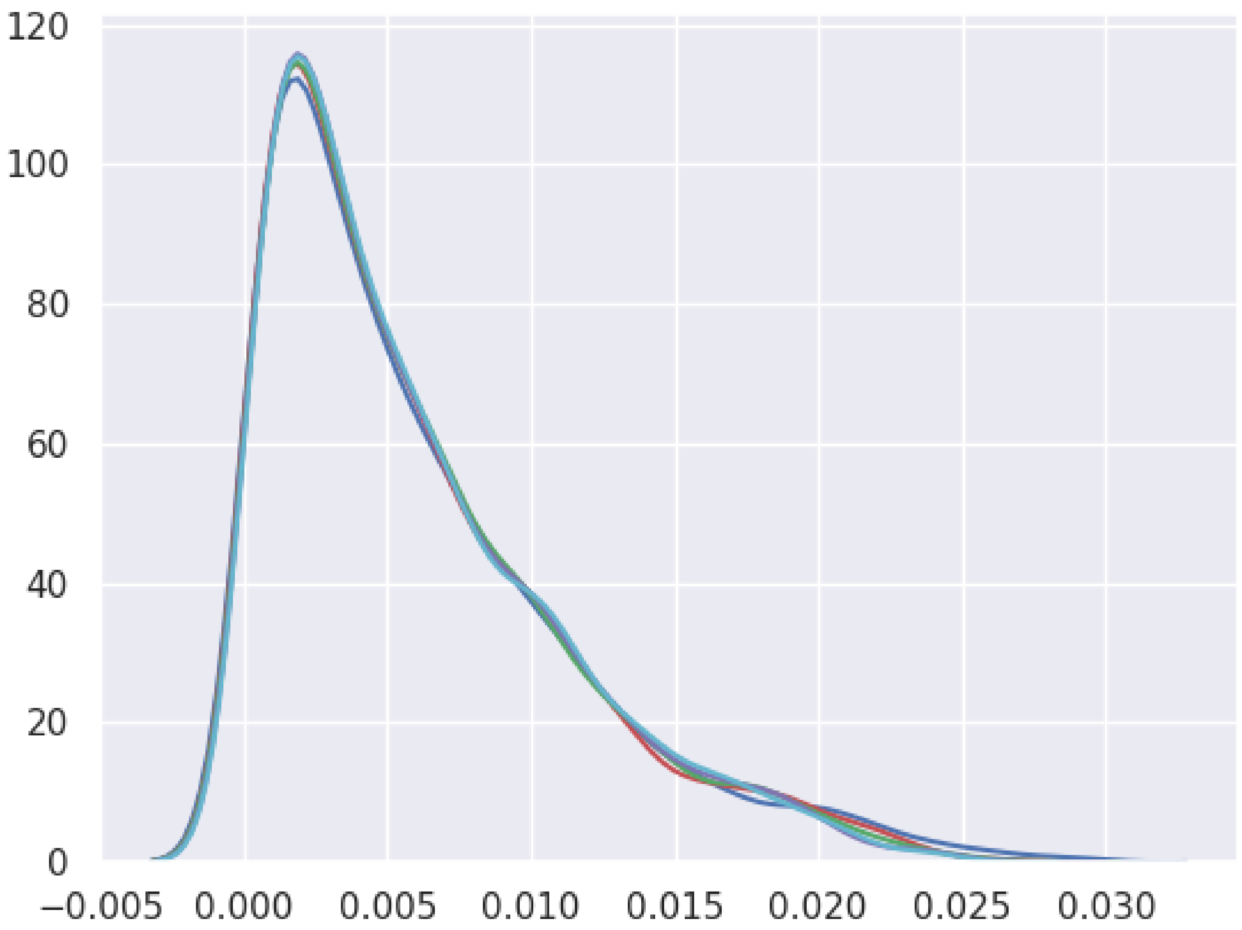
Red: n = 2000, k = 54

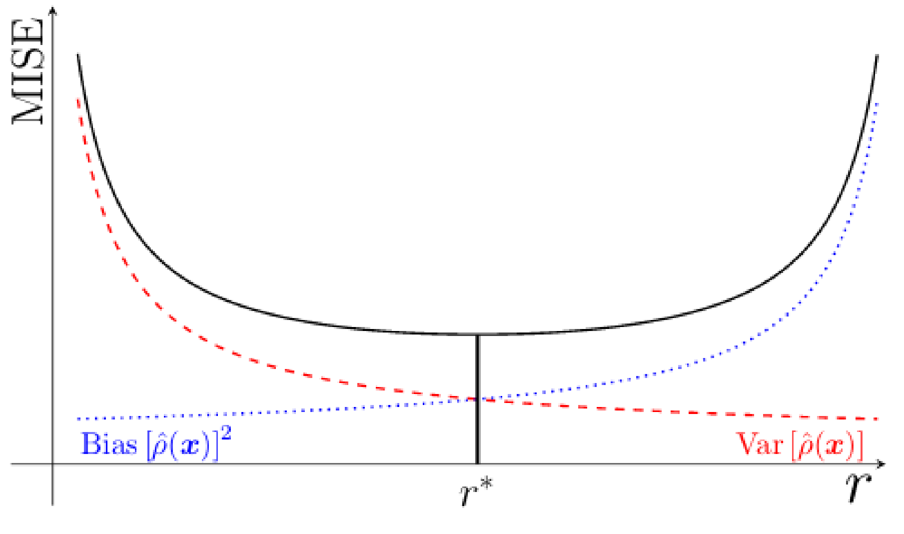
Green: n = 3000, k = 54

Magenta: n = 4000, k = 54

Cyan: n = 5000, k = 54

Change in k, same sample size – Toy scenario





- Choosing a suboptimal k leads to an increase in MISE
- When comparing data samples, one needs to use the same conditions for the sample i.e. use the same k to n relation e.g. $k \sim n^{-4/(4+d)}$
- A MISE that may not have been minimized may be desirable in areas that have been over or under smoothed

Blue: n =3000, k = 31

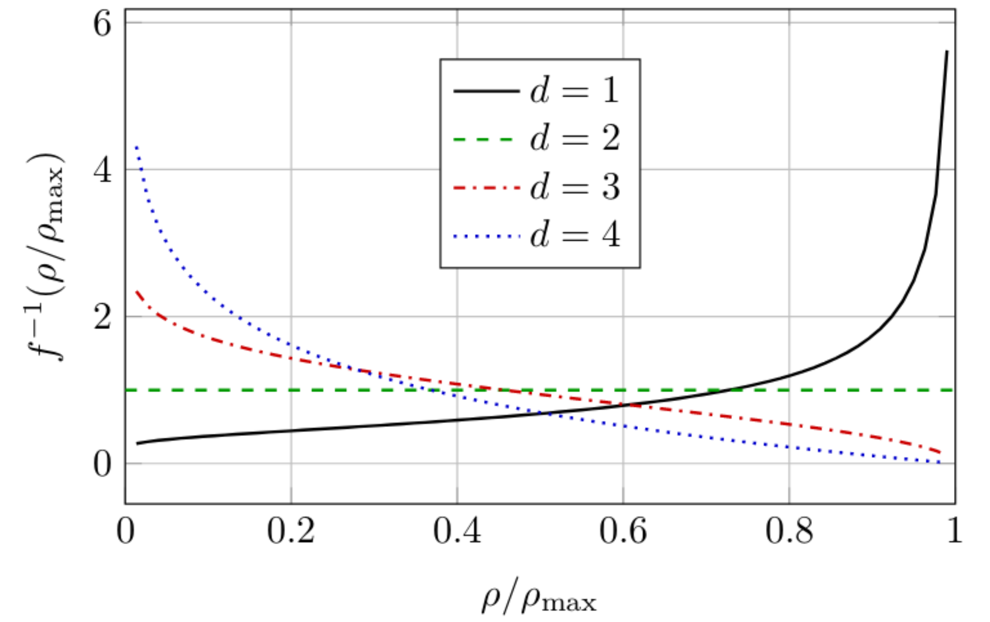
Red: n = 3000, k = 44

Green: n = 3000, k = 54

Magenta: n = 3000, k = 63

Cyan: n = 3000, k = 70

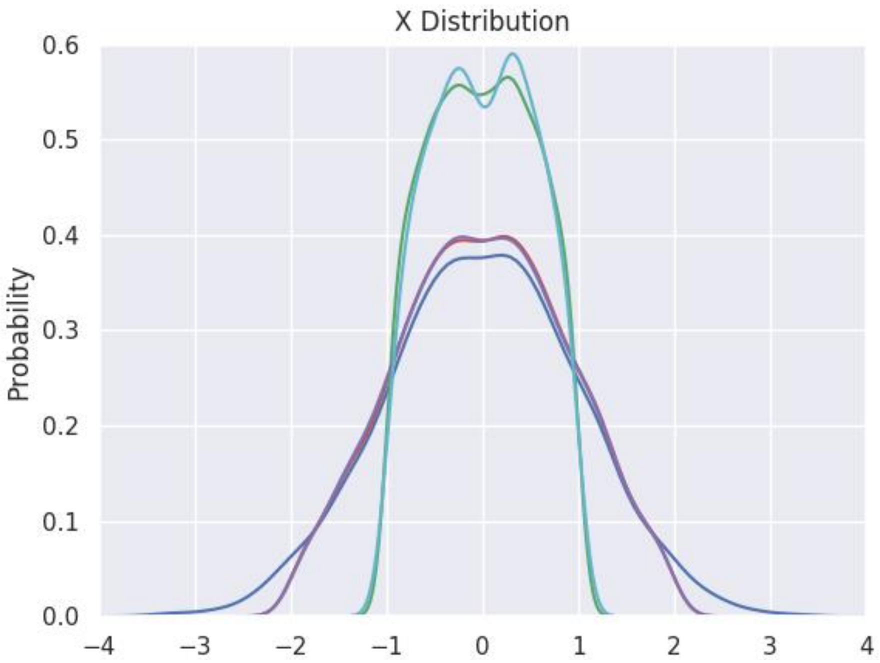
Missing Data - Toy Scenario
Scraping and Transmission Losses

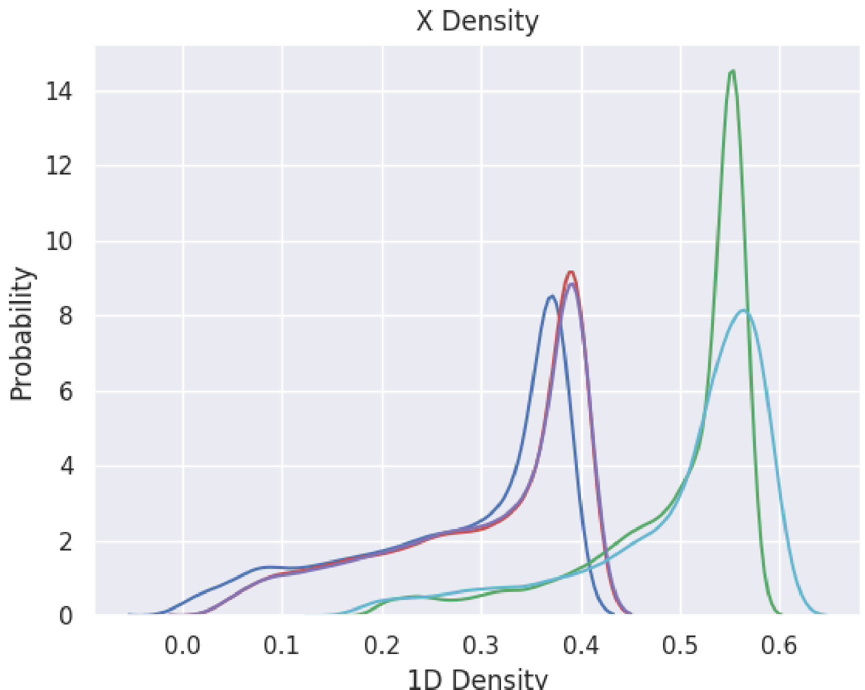


- Left – Expected Density for a Gaussian sample in each dimension normalized to the maximum density. As the dimension increases, particles more likely be found at a low phase space density

Toy example (next slides):

- 4D Gaussian sample – Mean = 0, Standard Deviation = diag(1,1,1,1)
- Full sample – No cuts – Blue
- Cut at +/- 2 sigma in one dimension called 'X' – red
- Cut at +/- 1 sigma in one dimension called 'X' – green
- Cut at +/- 2 sigma in each dimension – magenta
- Cut at +/- 1 sigma in each dimension – cyan





Full sample – No cuts – Blue

Cut at +/- 2 sigma in one dimension called 'X' – red

Cut at +/- 1 sigma in one dimension called 'X' – green

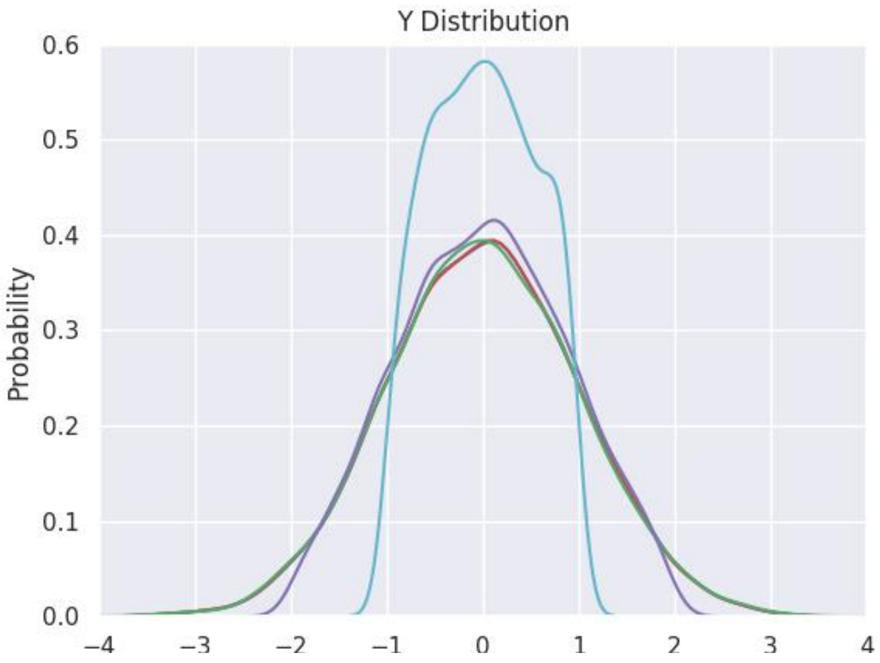
Cut at +/- 2 sigma in each dimension – magenta

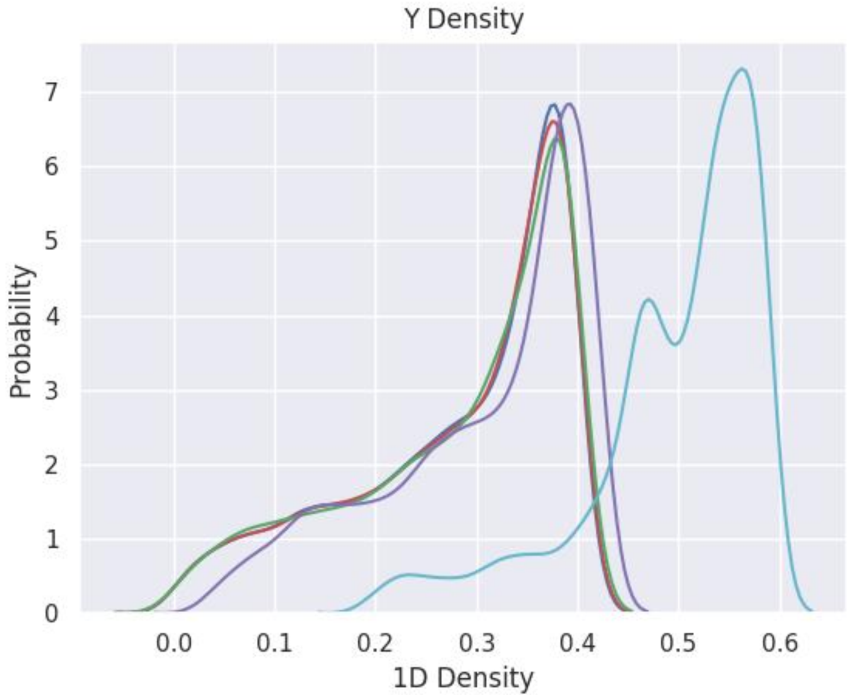
Cut at +/- 1 sigma in each dimension – cyan

1 sigma cut causes ~5% cut in 1D and ~17% in 4D which alters the density and distribution only slightly

2 sigma cut causes far greater change (~32% cut in 1D, ~79% cut in 4D)

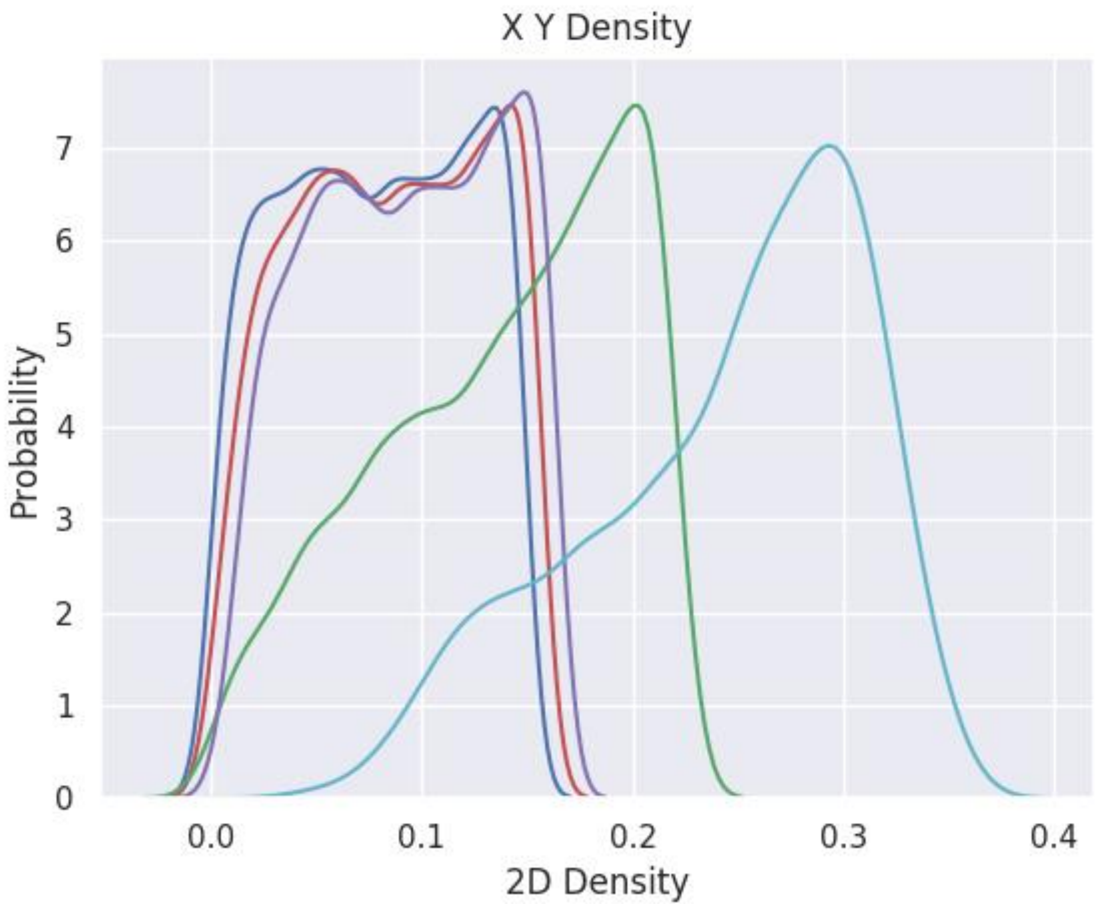
The k value is related to n, if the distribution is denser, than the calculated density will also be denser

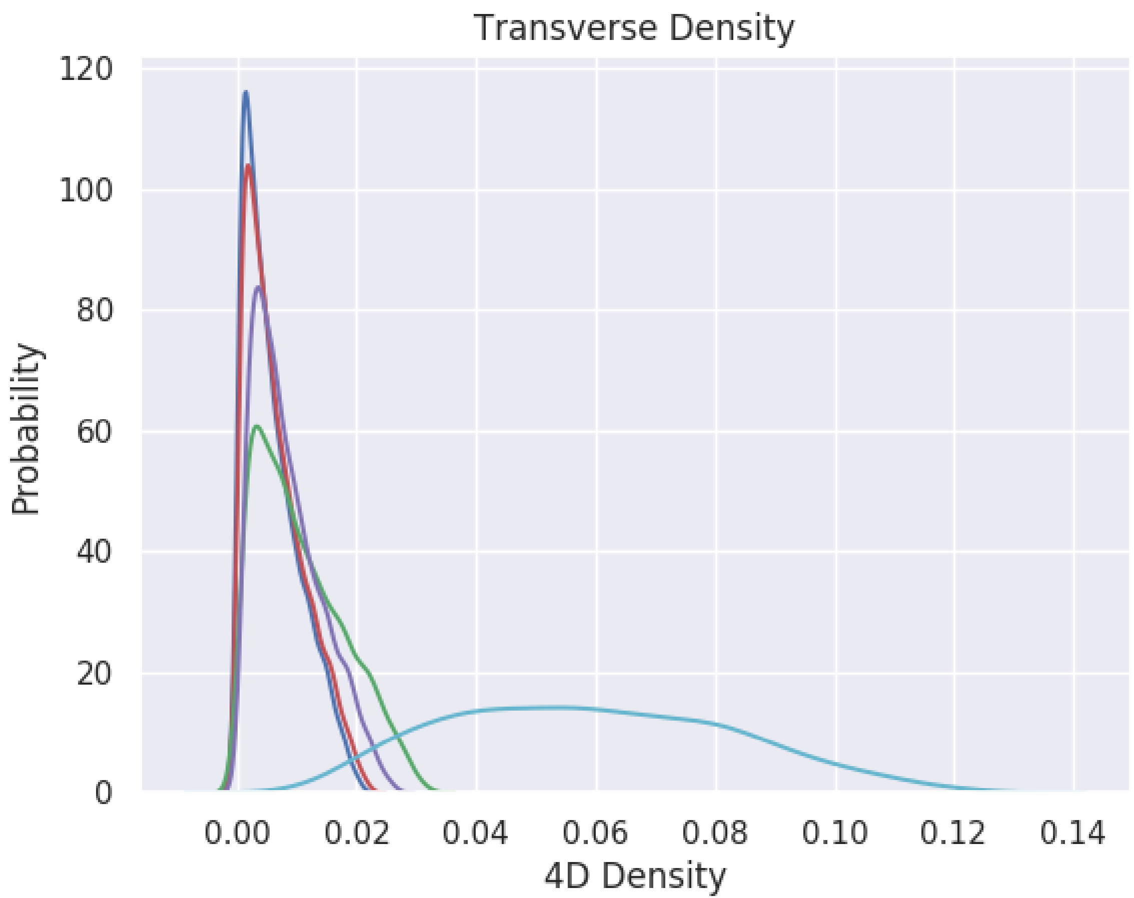




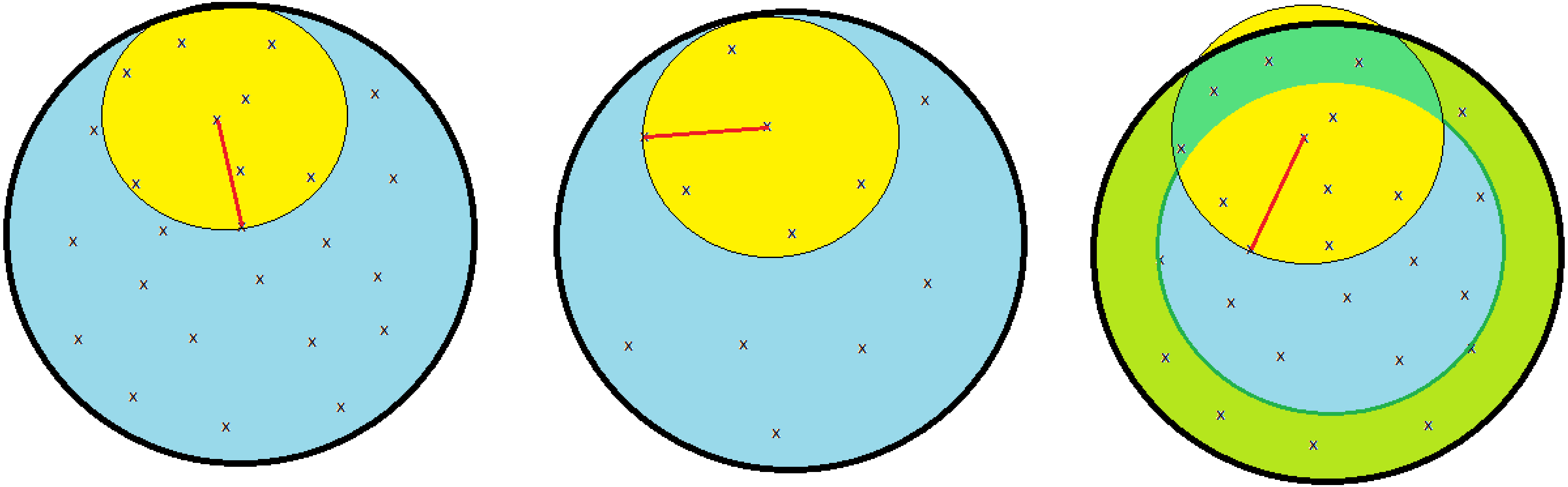
Missing Data Toy Scenario

Full sample – No cuts – Blue
Cut at +/- 2 sigma in one dimension called 'X' – red
Cut at +/- 1 sigma in one dimension called 'X' – green
Cut at +/- 2 sigma in each dimension – magenta
Cut at +/- 1 sigma in each dimension – cyan





- Left - original sample – red line shows k-nearest neighbour for point at centre of yellow circle
- Middle - subsample from original – red line distance only has slight change, k adjusts according to n. As n becomes small the error increases
- Right – Aperture cut by the green sub-circle – points at large radius are removed. While n has reduced, the k is now ideal for the subsample distribution.
- For points with a bounding circle affected by the aperture cut, the k-nearest neighbour may be further away, while for points at the centre of the sample the nearest neighbour is closer as the k is reduced, but no close points are removed



Sophie Middleton Transfer Matrix for Focus coil for some Pz momenta

$$31 < TOF_{12} < 31.14$$

$$\begin{pmatrix} 2.96 \pm 8.67 & -0.51 \pm 0.11 & 3.36 \pm 0.86 & 0.137 \pm 0.172 & -1.615 \pm 2.017 \\ -0.709 \pm 1.05 & -0.163 \pm 0.019 & -0.578 \pm 0.249 & 0.018 \pm 0.0129 & -0.249 \pm 0.151 \\ -2.97 \pm 10.28 & 0.059 \pm 0.207 & 1.12 \pm 2.476 & -0.721 \pm 0.234 & 4.58 \pm 2.16 \\ -0.432 \pm 043 & -0.0032 \pm 0.0115 & 0.157 \pm 0.154 & -0.154 \pm 0.125 & -0.882 \pm 0.117 \end{pmatrix}$$

$$31.14 < TOF_{12} < 31.28$$

$$\begin{pmatrix} 1.369 \pm 8.90 & -0.525 \pm 0.124 & 2.08 \pm 2.53 & 0.035 \pm 0.213 & -1.056 \pm 0.986 \\ 0.403 \pm 1.79 & -0.1456 \pm 0.0097 & -0.860 \pm 0.309 & -0.002 \pm 0.004 & -0.156 \pm 0.238 \\ -1.83 \pm 5.37 & -0.04 \pm 0.21 & 0.364 \pm 0.166 & 0.363 \pm 1.075 & 4.42 \pm 1.28 \\ 0.374 \pm 1.40 & 0.005 \pm 0.036 & -0.115 \pm 0.300 & -0.154 \pm 0.022 & -0.793 \pm 0.168 \end{pmatrix}$$

$$31.28 < TOF_{12} < 31.42$$

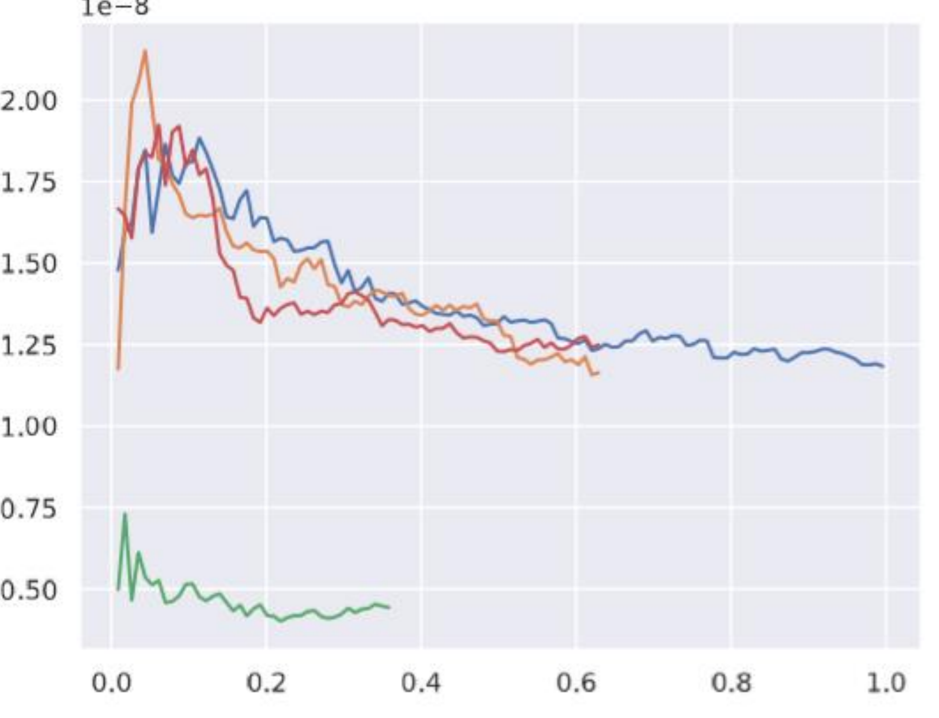
$$\begin{pmatrix} -5.34 \pm 4.45 & -0.644 \pm 0.08 & 2.66 \pm 0.60 & 0.1901 \pm 0.0526 & -0.91 \pm 0.71 \\ -0.74 \pm 0.98 & -0.153 \pm 0.019 & -0.689 \pm 0.151 & 0.0295 \pm 0.0029 & -0.278 \pm 0.061 \\ 0.351 \pm 5.55 & 0.0189 \pm 0.1121 & -0.67 \pm 0.47 & -0.516 \pm 0.035 & 2.42 \pm 1.05 \\ 1.213 \pm 1.227 & 0.0282 \pm 0.0301 & -0.315 \pm 0.088 & -0.144 \pm 0.010 & -0.688 \pm 0.119 \end{pmatrix}$$

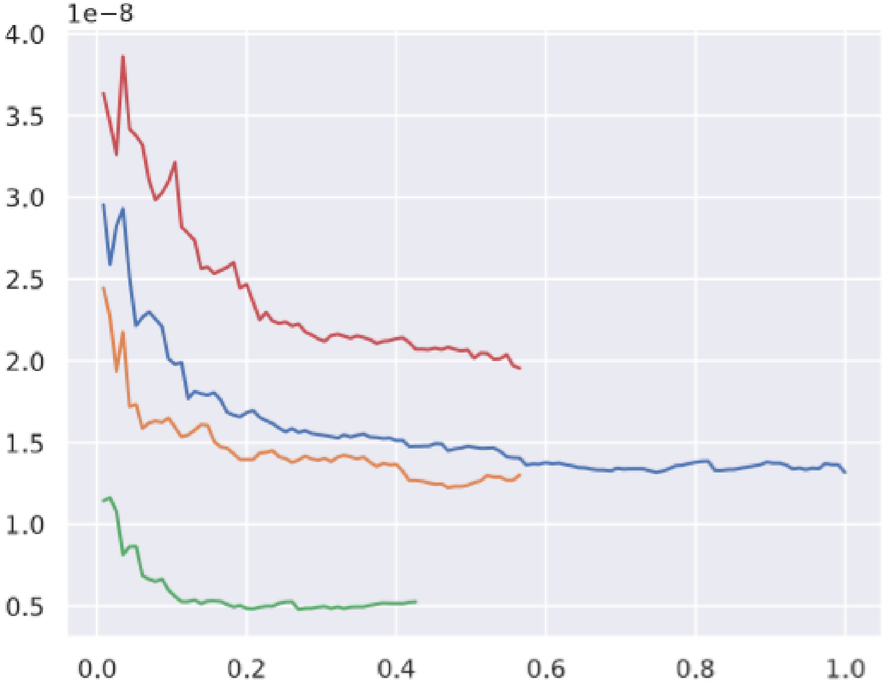
$$31.42 < TOF_{12} < 31.56$$

$$\begin{pmatrix} 9.29 \pm 10.05 & -0.62 \pm 0.10 & 1.331 \pm 1.071 & 0.112 \pm 0.134 & -2.29 \pm 2.11 \\ 1.05 \pm 1.42 & -0.148 \pm 0.019 & -0.941 \pm 0.133 & 0.0055 \pm 0.0100 & -0.142 \pm 0.194 \\ 13.4 \pm 9.50 & 0.209 \pm 0.252 & -2.16 \pm 1.46 & -0.484 \pm 0.064 & -0.34 \pm 0.92 \\ 1.92 \pm 1.53 & 0.0346 \pm 0.0332 & 0.166 \pm 0.260 & -0.152 \pm 0.009 & -1.135 \pm 0.088 \end{pmatrix}$$

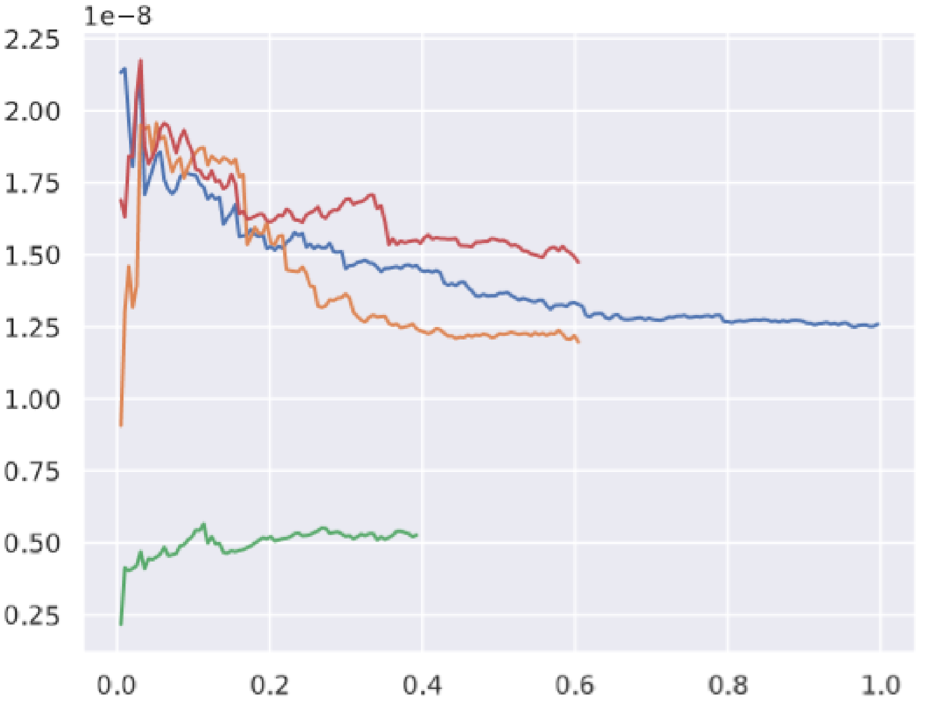
- Had very low statistics, can reduce error
- Shows pz dependence, may need to expand to 7 by 6 matrix

Change in Peak density vs beam fraction

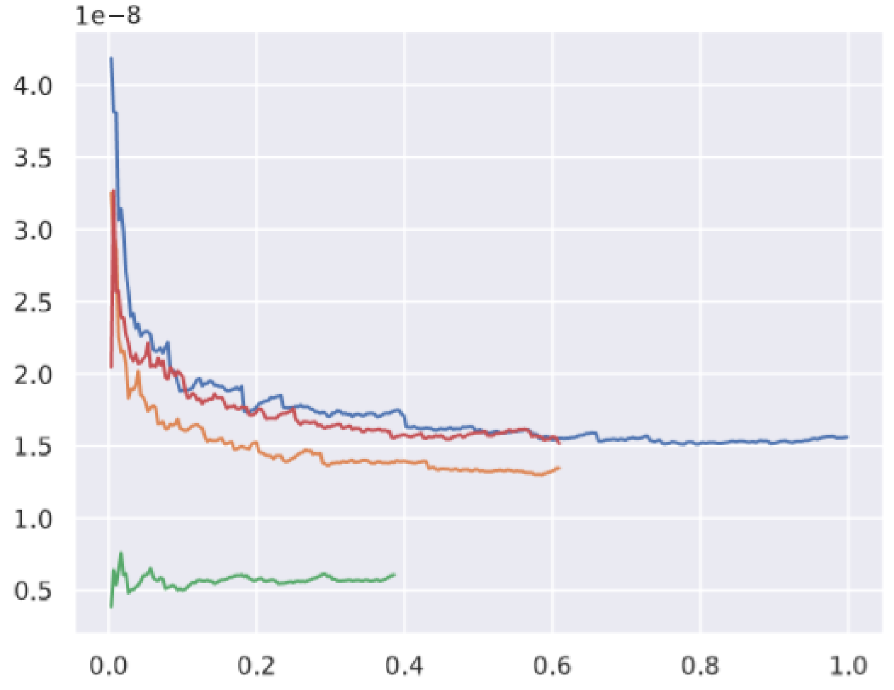




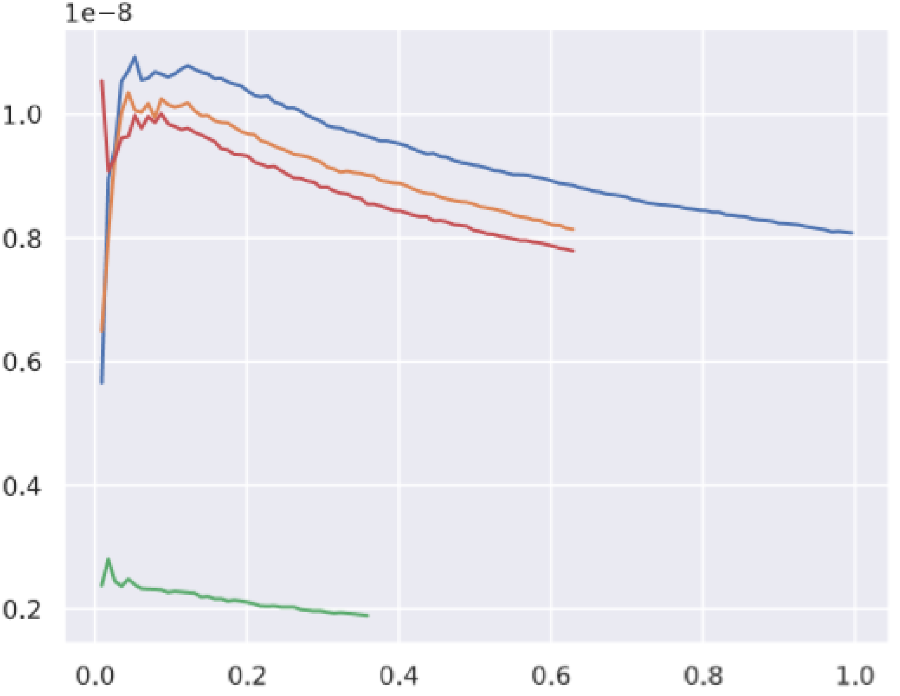
Top Left: No absorber
Top Right: Wedge
Bottom Left: LiH
Bottom Right: LH2

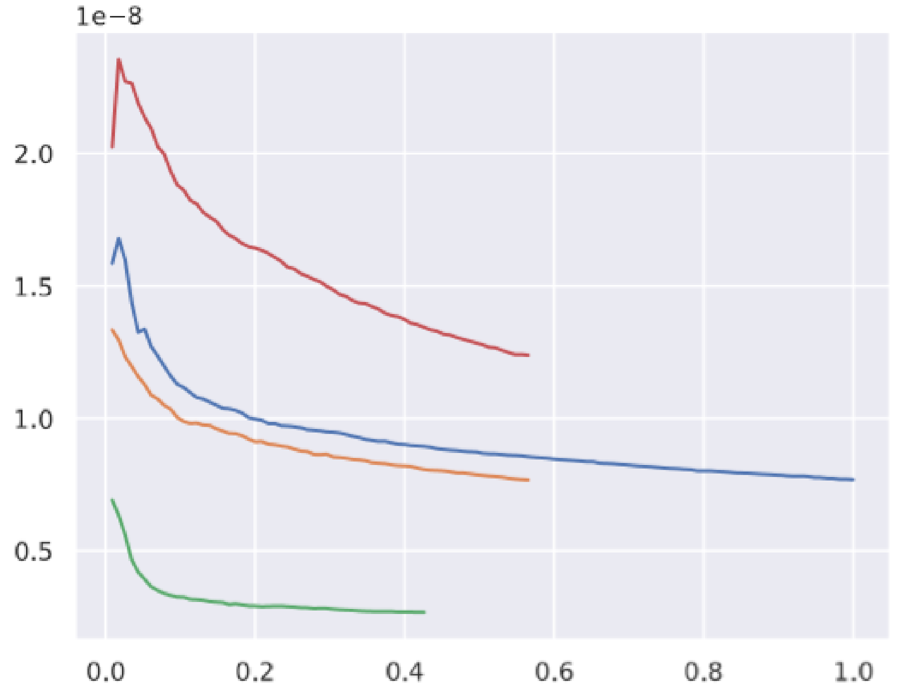


Blue – Full Upstream Sample
Red – Full Downstream Sample
Orange – Upstream Sample which made it Downstream
Green – Upstream Sample which doesn't make it downstream

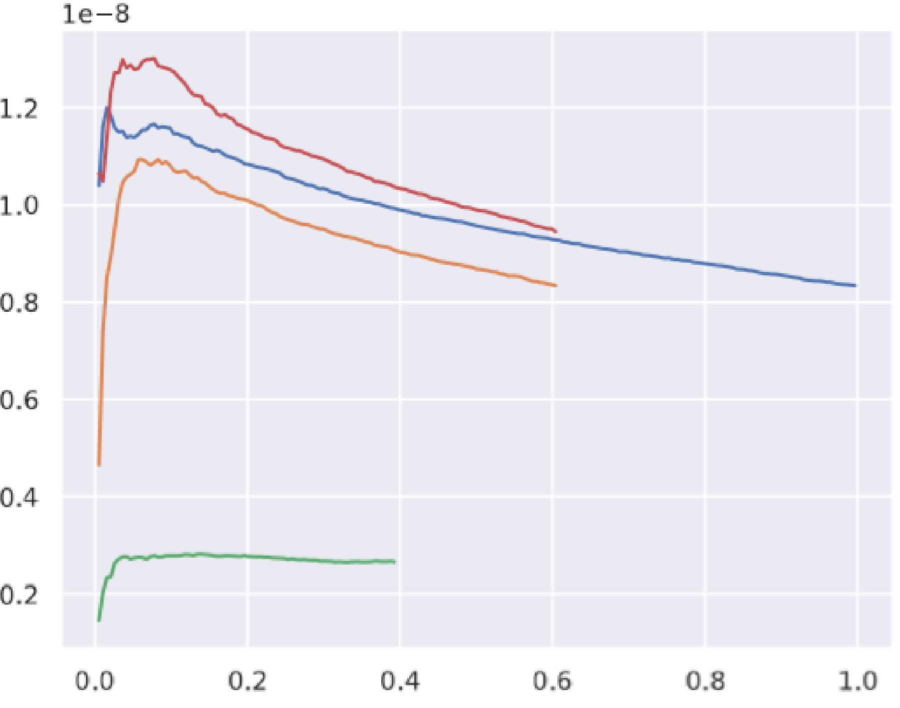


Change in 9th percentile density vs beam fraction

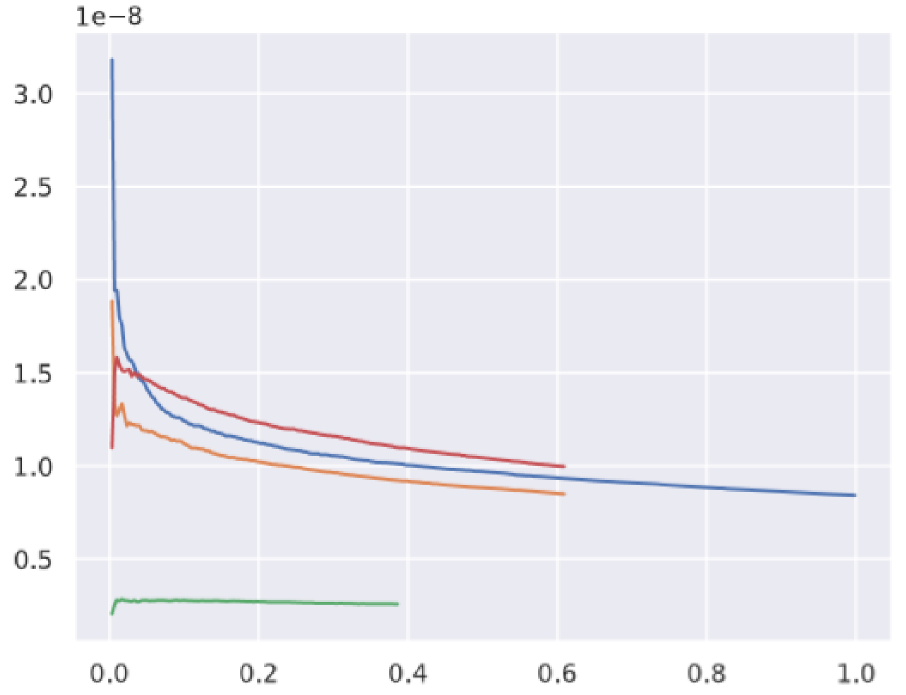


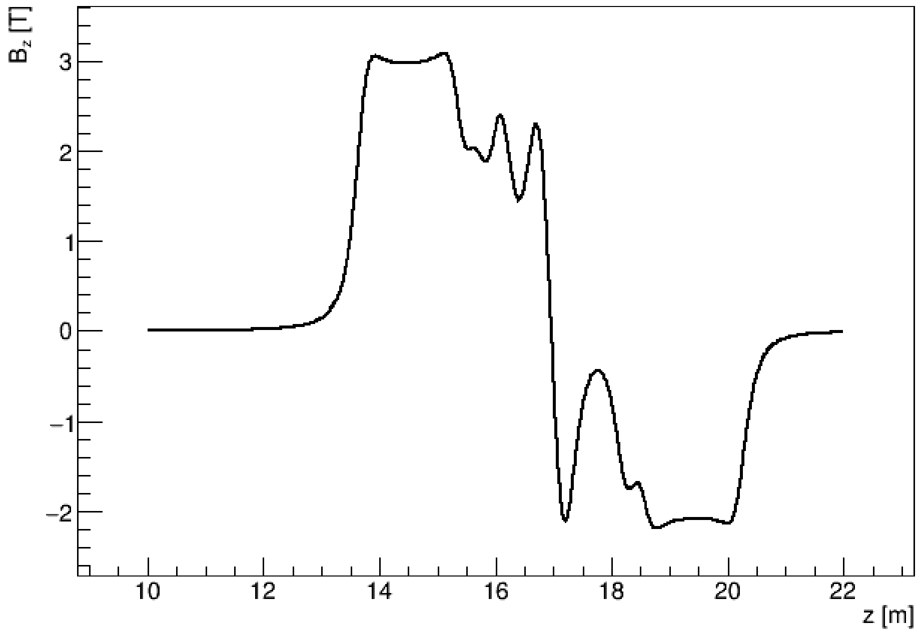


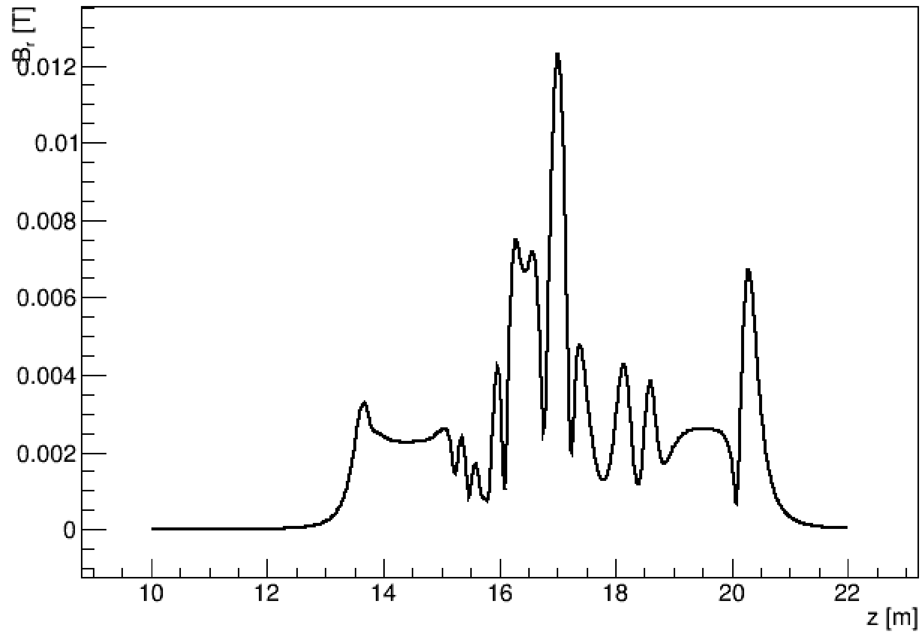
Top Left: No absorber
Top Right: Wedge
Bottom Left: LiH
Bottom Right: LH2



Blue – Full Upstream Sample
Red – Full Downstream Sample
Orange – Upstream Sample which made it Downstream
Green – Upstream Sample which doesn't make it downstream

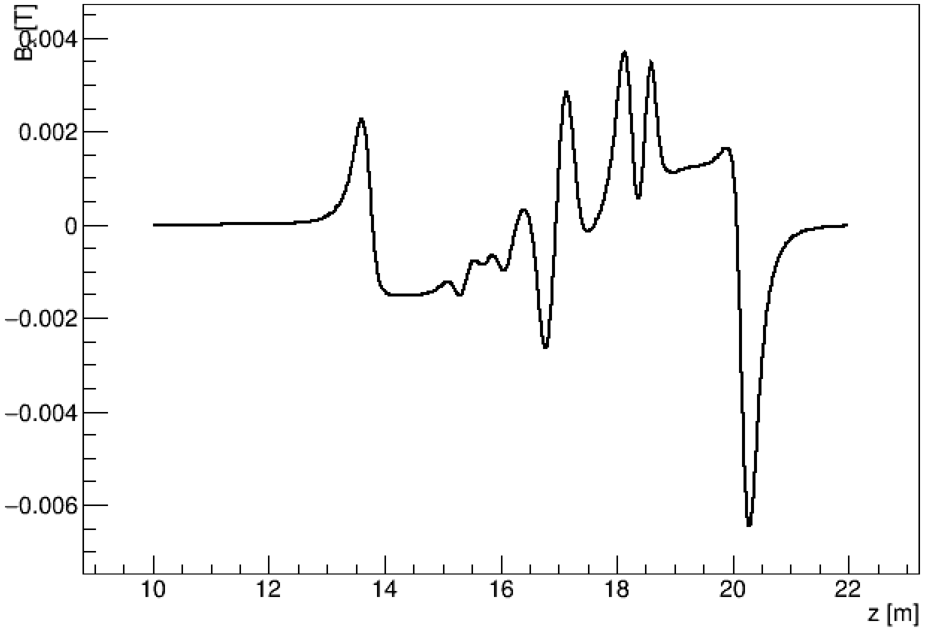


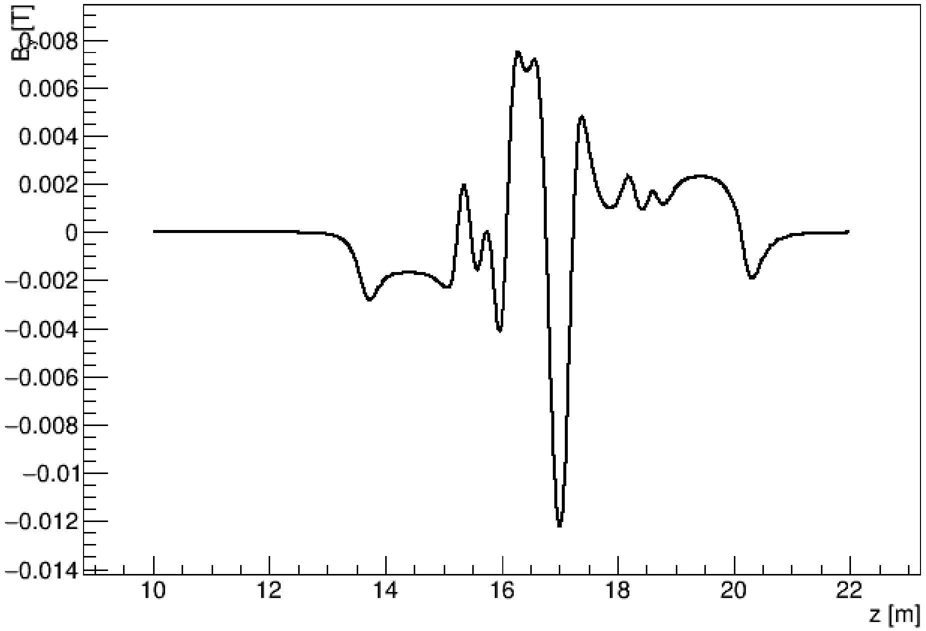


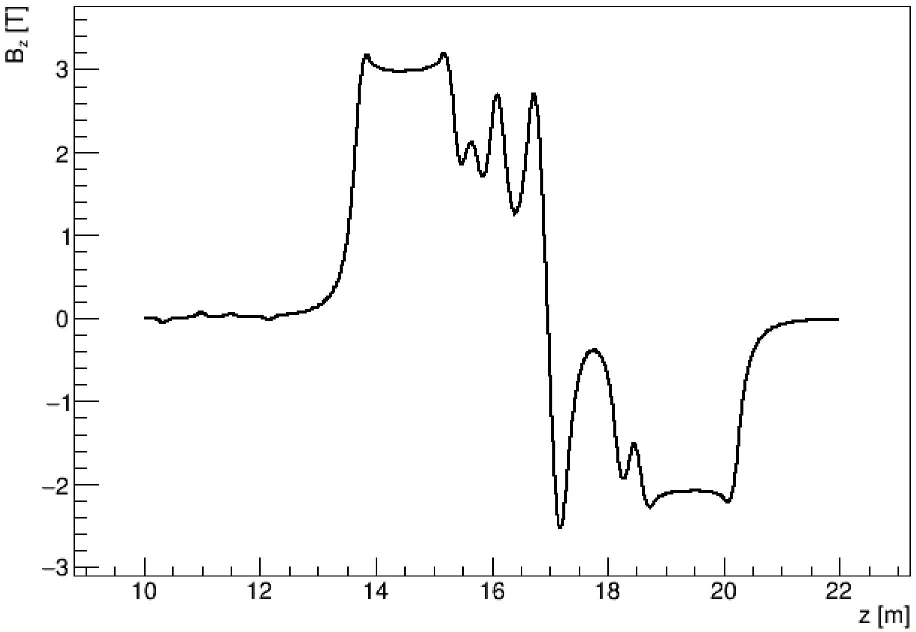


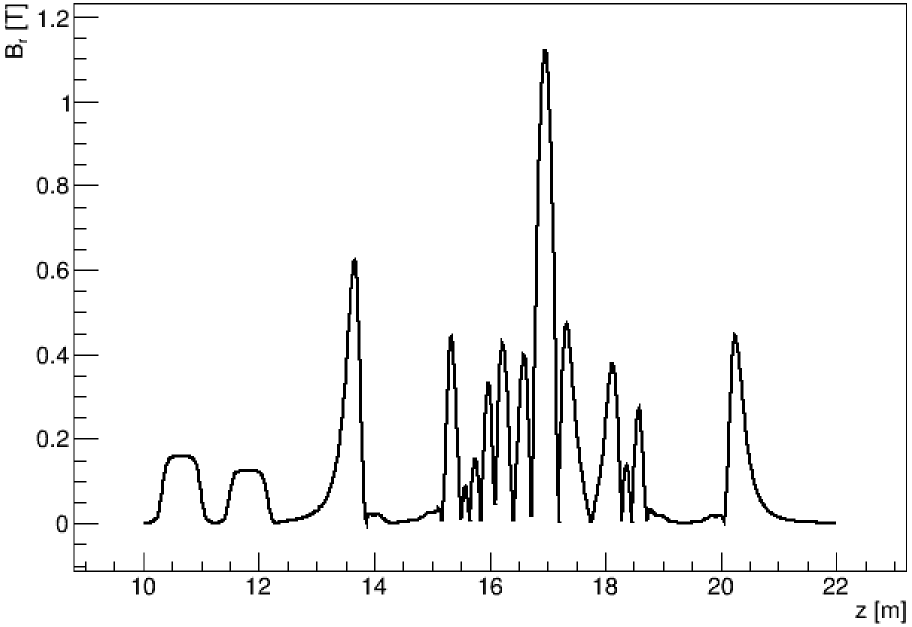
Makefieldmap.py at a radius of 0 mm

Top Left: Bz
Top Right: Br
Bottom Left: Bx
Bottom Right: By









Makefieldmap.py at a radius of 100 mm

Top Left: Bz
Top Right: Br
Bottom Left: Bx
Bottom Right: By