Novel Approaches to Measuring Beam Cooling

Tanaz A. Mohayai

MICE-US Analysis Meeting

Dec 16, 2016

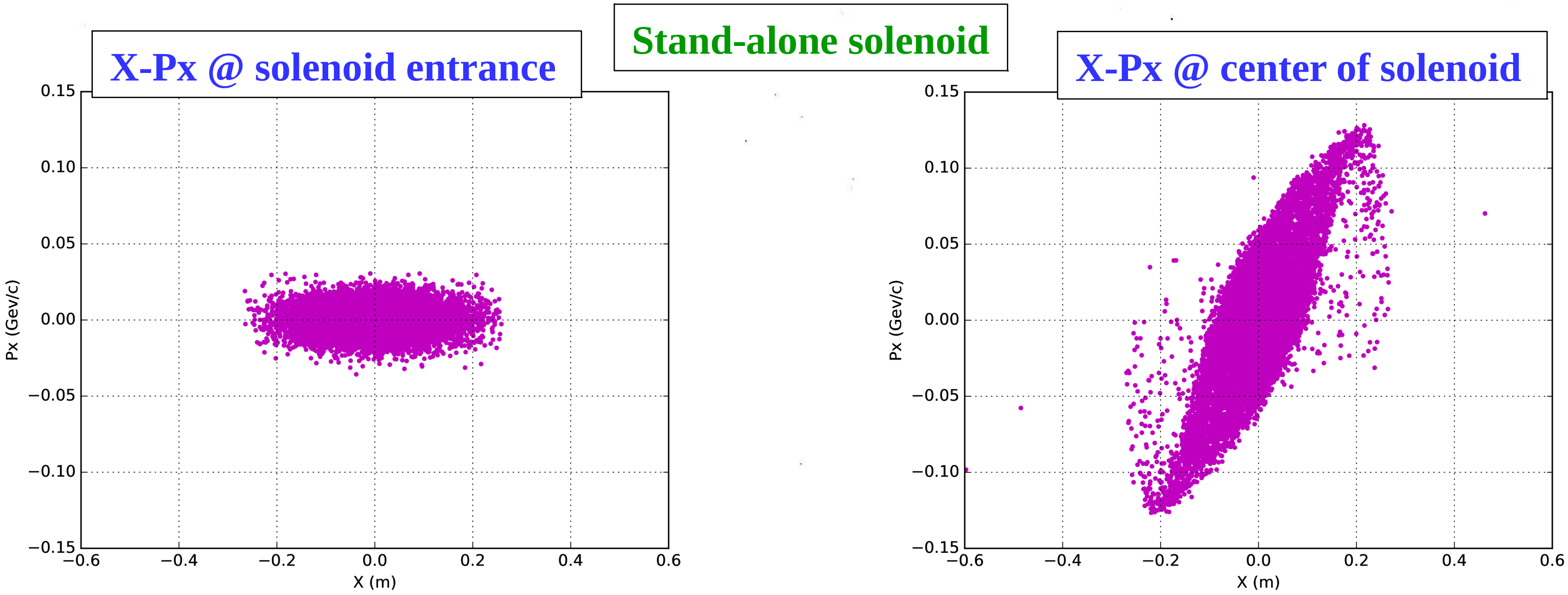
ILLINOIS INSTITUTE OF TECHNOLOGY

Fermilab

MICE

Motivation

- Solenoid beam optics prone to **filamentation** & **non-linear effects**.



- **Problem**:
  - "Apparent" RMS emittance growth due to non-linear effects causing distortions in beam optics.
  - Defining a model for such distributions becomes a challenge.
- **Solution**:
  - Estimate distribution using density estimation (DE) techniques.
  - DE captures non-linear effects (e.g. Kernel Density Estimation).

Overview – Density Estimation (DE) Techniques

- Some basics:
  - Density estimation: widely-used statistics technique for estimating the **probability density function (PDF)** of a **random variable**.
  - Note: density & **PDF** terms are used interchangeably in my analysis.
- Finding **PDF** using DE:
  - Choose a DE technique for which the error (deviations of estimated **PDF** from the true **PDF**) is minimized.
  - Two types:
    1) Parametric: make an assumption about underlying **PDF** to estimate distribution parameters.
    2) Non-parametric: make fewer assumptions about the underlying **PDF** to estimate the distribution using some smoothing technique. **DE techniques used in MICE are non-parametric.**
- Oldest and simplest example:
  - Histograms:
    - Estimate density by counting data points which fall within bins of certain widths.
    - Features:
      1) Estimated density changes depending on choice of origin.
      2) As bin width, h changes, the level of smoothing changes.
      3) Distribution not generally smooth affecting data interpretation.

Overview cont. – Kernel Density Estimation (KDE) technique

- Make use of a **kernel function** *k(x)* (instead of a bin) to estimate density at an arbitrary point, *x*. Acts as a weight function: points closer to *x* get higher weights. Satisfies following conditions:

$$\int_{-\infty}^{\infty} k(x)dx = 1. \qquad \int_{-\infty}^{\infty} xk(x)dx = 0. \qquad k_2(k) = \int_{-\infty}^{\infty} x^2 k(x)dx > 0.$$

- The estimator using the kernel function is called **kernel density estimation (KDE),**

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

n: sample size

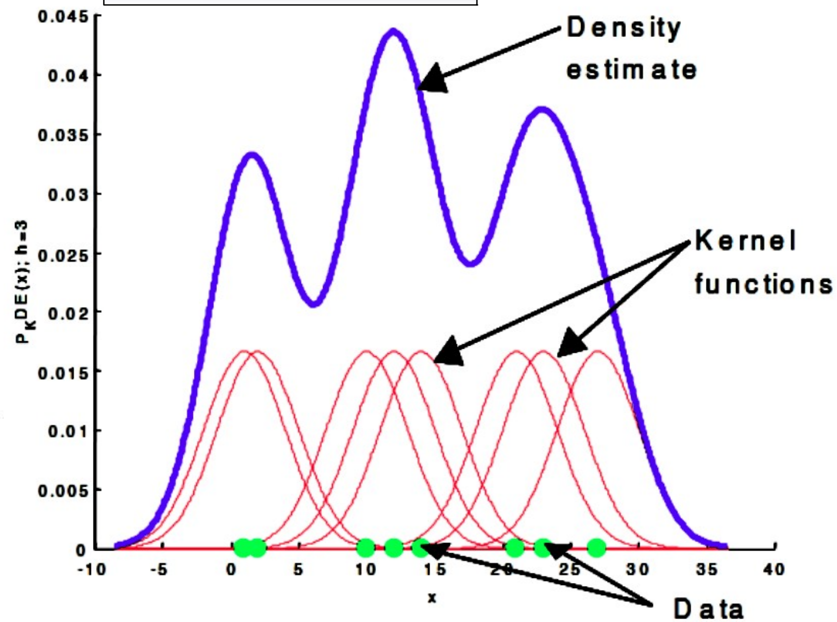
X_i: i^{th} data point

h: the width of **kernel function** known as the **bandwidth parameter**

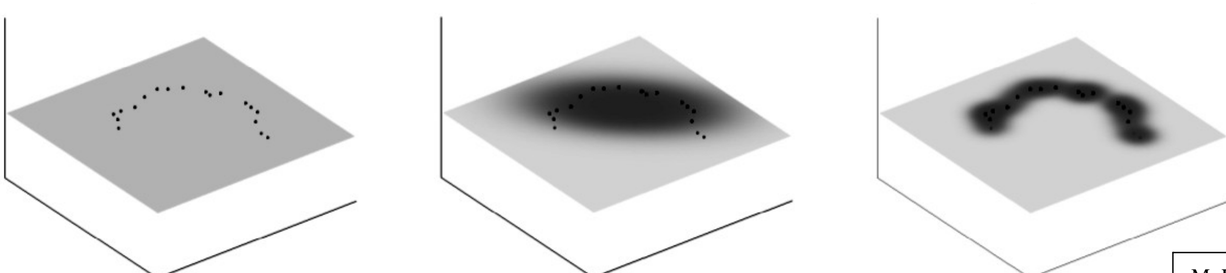
- **Kernel function** can be a symmetric distribution. e.g. Gaussian kernel function:

$$K\left(\frac{\vec{x} - \vec{x_i}}{h}\right) = \frac{1}{(2\pi)^{\frac{d}{2}}}\exp\left(-\frac{|\vec{x} - \vec{x_i}|^2}{2h^2}\right).$$

1-D KDE illustration



R. Gutierrez Osuna, "Kernel density estimation", CSCE 666 Pattern Analysis, Texas AM University.



KDE comparison with Gaussian distribution

M. Rousson, et. al., "Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation", (MICCAI) (2005)

KDE in One Dimension

- In DE, it is important to minimize errors.
- Measures of error – error being the discrepancy between estimated **PDF** ($\hat{f}(x)$) & true **PDF** ($f(x)$):

Mean square error (MSE):

- Average of the squares of the errors at a single point x: $\mathrm{MSE}(\hat{f}(x)) = E\left\{\hat{f}(x) - f(x)\right\}^2$

Mean integrated square error (MISE):

- Average of the squares of the errors at all values of x. MSE and MISE can be written in terms of **bias** & **variance**,

$$\mathrm{MISE}(\hat{f}(x)) = \int_{-\infty}^{\infty} \mathrm{bias}(\hat{f}(x))^2 dx + \int_{-\infty}^{\infty} \mathrm{variance}(\hat{f}(x)) dx =$$
$$\int_{-\infty}^{\infty} \left[E\hat{f}(x) - f(x)\right]^2 dx + \int_{-\infty}^{\infty} \mathrm{variance}(\hat{f}(x)) dx.$$

- **Systematic** error or **bias** of the estimated density – larger bias leads to over-smoothed distribution.
- **Random** error or **variance** of the estimated density – larger variance leads to noisier distribution.

Asymptotic MISE or AMISE:

- MISE can be derived using asymptotic approximations - the limit in which estimated **PDF** approaches true **PDF**.

KDE in One Dimension cont.

- Use properties below to obtain AMISE:

✯ Reminder: **kernel function** should satisfy:

$$\int_{-\infty}^{\infty} k(x)dx = 1. \qquad \int_{-\infty}^{\infty} xk(x)dx = 0. \qquad k_2(k) = \int_{-\infty}^{\infty} x^2 k(x)dx > 0.$$

✯ **Bandwidth parameter**, h should be small and n should be large (limit as estimated **PDF** $\rightarrow$ true **PDF**).

- AMISE,

$$\text{AMISE}\left(\hat{f}(x)\right) \approx \left(\overbrace{\frac{1}{4}h^4 k_2^2(k) \int_{-\infty}^{\infty} f''(x)^2 dx}^{\textbf{Integrated bias}} + \overbrace{\frac{1}{n}\frac{1}{h}\int_{-\infty}^{\infty} k(u)^2 du}^{\textbf{Integrated variance}}\right).$$

where $u = \dfrac{x-z}{h}$

- Minimize AMISE with respect to h & solve for optimal **bandwidth parameter**

,

$$h_{\text{optimal}} = k_2(k)^{-\frac{2}{5}} \left(\int_{-\infty}^{\infty} k(u)^2 du\right)^{\frac{1}{5}} \left(\int_{-\infty}^{\infty} f''(x)^2 dx\right)^{\frac{-1}{5}} n^{-\frac{1}{5}}.$$

- Optimal **bandwidth parameter** depends on the second derivate of the true **PDF**. Assume true **PDF** is the standard Gaussian distribution, and **kernel function** the Gaussian kernel, then h is

$$h_{\text{optimal}} = 1.06\sigma n^{-1/5}. \quad \text{or} \quad h_{\text{optimal}} = h_{\text{factor}}\sigma$$

σ*:* the variance of the standard Gaussian density.

KDE in One Dimension cont.

- **Reminder:** AMISE in terms of the asymptotic **bias** & **variance** terms,

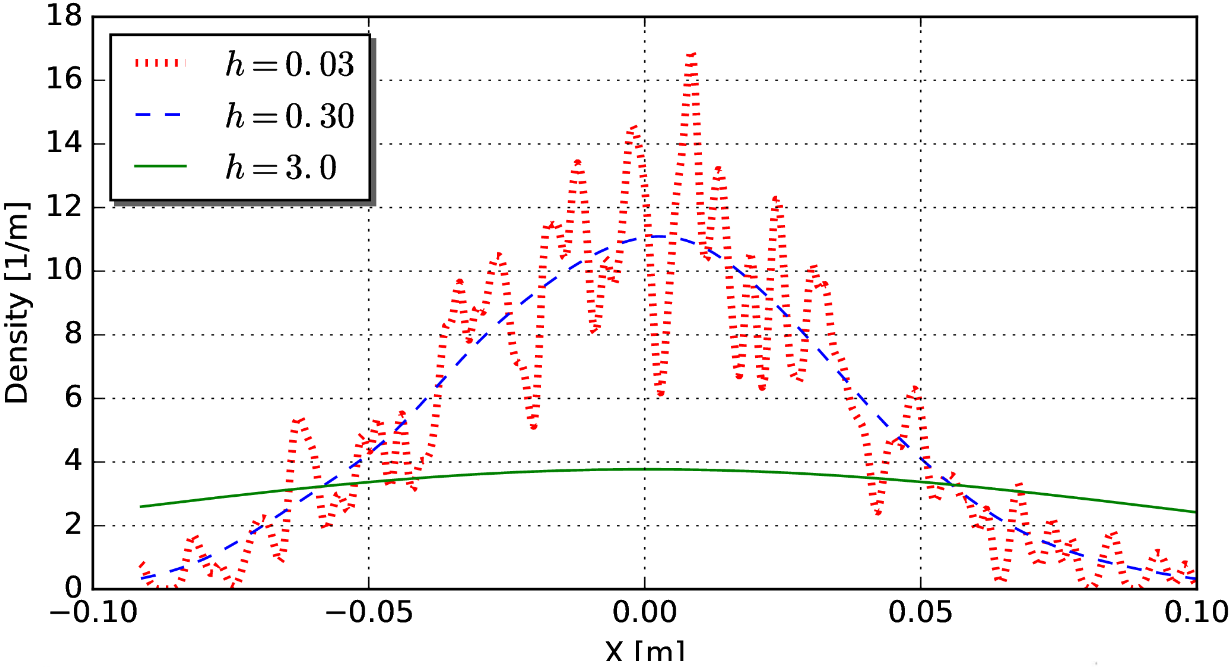
$$\text{AMISE}(\hat{f}(x)) = \frac{1}{4}h^4 k_2^2(k)\int_{-\infty}^{\infty} f''(x)^2 dx + \frac{1}{n}\frac{1}{h}\int_{-\infty}^{\infty} k(u)^2 du.$$

- Note the **bias**-**variance** trade-off feature in AMISE:

- To get small **bias**, h should be small, BUT small h leads to large **variance** & vice versa.
- Effect of changing h on estimated PDF: large h over-smooths density, while small h under-smooths it.

- x coordinates of a subsample of *500* muons at the entrance of TKU.
- True PDF approximately Gaussian.
- $h = 0.3$ reveals a Gaussian.
- $h = 3.0$ over-smooths the PDF.
- $h = 0.03$ reveals a noisier PDF.

Effects of varying bandwidth on estimated density



KDE in Multiple Dimensions

- The AMISE in multiple dimensions,

$$\text{AMISE}(\hat{f}(\vec{x})) \approx \underbrace{\frac{1}{4}h^4 \int_{-\infty}^{\infty} u^2 k(u) du \int_{-\infty}^{\infty} \left(\nabla^2 f(\vec{x})\right)^2 dx}_{\text{Integrated bias}} + \underbrace{n^{-1} h^{-d} \int_{-\infty}^{\infty} k(u)^2 du}_{\text{Integrated variance}}.$$

- The multi-dimensional optimal bandwidth,

$$h_{\text{optimal}} = \left(d \int_{-\infty}^{\infty} k(u)^2 du \left(\int_{-\infty}^{\infty} u^2 k(u) du \right)^{-2} \right)^{\frac{1}{d+4}} \left(\int_{-\infty}^{\infty} \left(\nabla^2 f(\vec{x})\right)^2 dx \right)^{\frac{-1}{d+4}} n^{\frac{-1}{d+4}}.$$

$$h_{\text{optimal}} = \left(\frac{4}{d+4}\right)^{\frac{1}{d+4}} \Sigma n^{\frac{-1}{d+4}} \quad \text{or} \quad h_{\text{optimal}} = h_{\text{factor}} \Sigma$$

- Multivariate **KDE**,

$$\hat{f}(\vec{x}) = \frac{|\Sigma|^{-1/2}}{n h_{\text{factor}}^d \sqrt{(2\pi)^d}} \sum_{i=1}^{n} \exp\left[-\frac{(\vec{x} - \vec{X}_i)^T \Sigma^{-1} (\vec{x} - \vec{X}_i)}{2 h_{\text{factor}}^2} \right]$$

where Σ is the covariance matrix of the data set, representing the variances of each of the phase-space coordinate in d-dimensions.

- **Note**: KDE is applied to the 4-dimensional MICE phase space.

DE and KDE in MICE

- Why use DE and KDE in MICE:

1) New figures of merit for beam cooling:
   - Phase-space density
   - Phase-space volume
   - Count of muons
2) Beam sampling to demonstrate stronger cooling effect:
   - More accurate classifications of the beam tails and cores (k-nearest neighbor classifier/density estimation).
   - Cutting high amplitude muons to reduce tail effects.
   - Re-weighting of the beam to obtain an ideal Gaussian beam.

DE and KDE in MICE cont. – Approach

- Phase-space density measurement using KDE:

1) Measure the bandwidth parameter, h (slide 8).
2) Estimate density for each muon of coordinates (x,p_x,y,p_y) with contributions from all the other muons of coordinates (x_i,p_{xi},y_i,p_{yi}).
3) Produce phase-space density evolution plot: this means drawing contour lines of constant density to track the density of a specific contour as it evolves in the cooling channel. I track the one sigma contour (in 4 dimensions, contains 9% of muons).

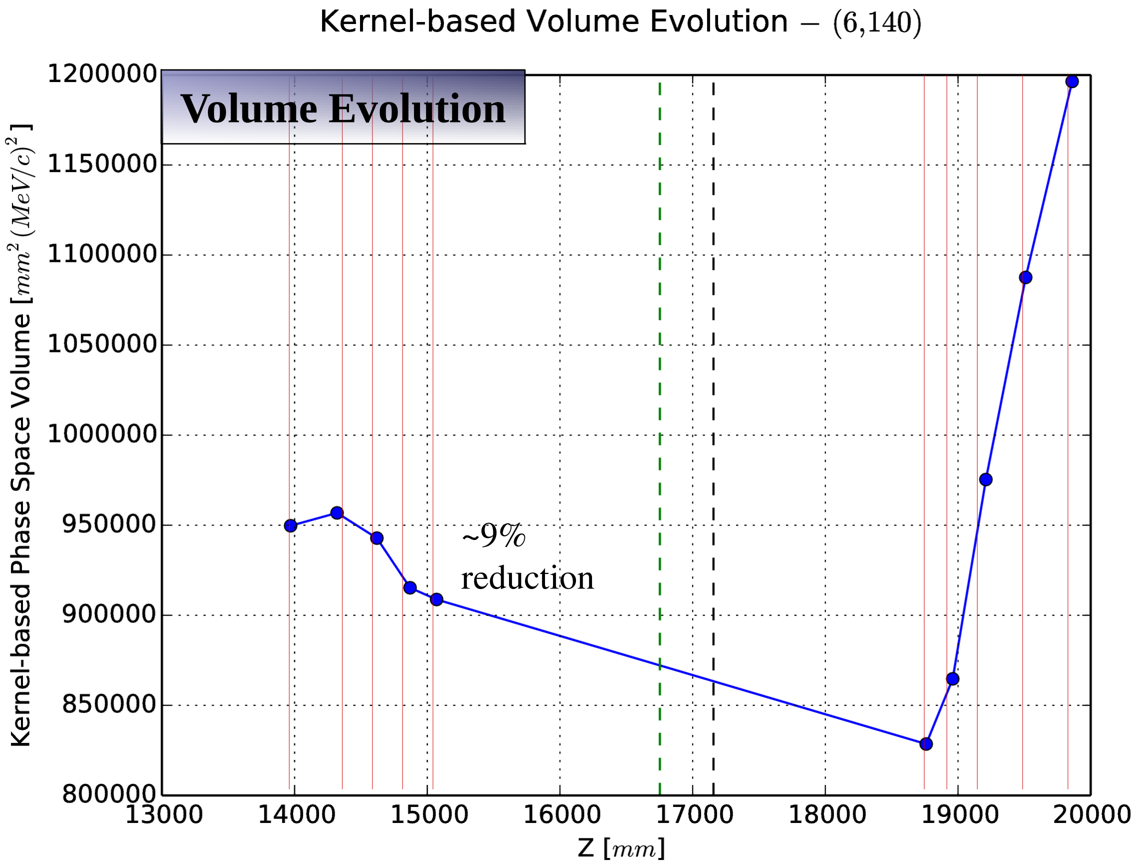
- Phase-space volume measurement using KDE:

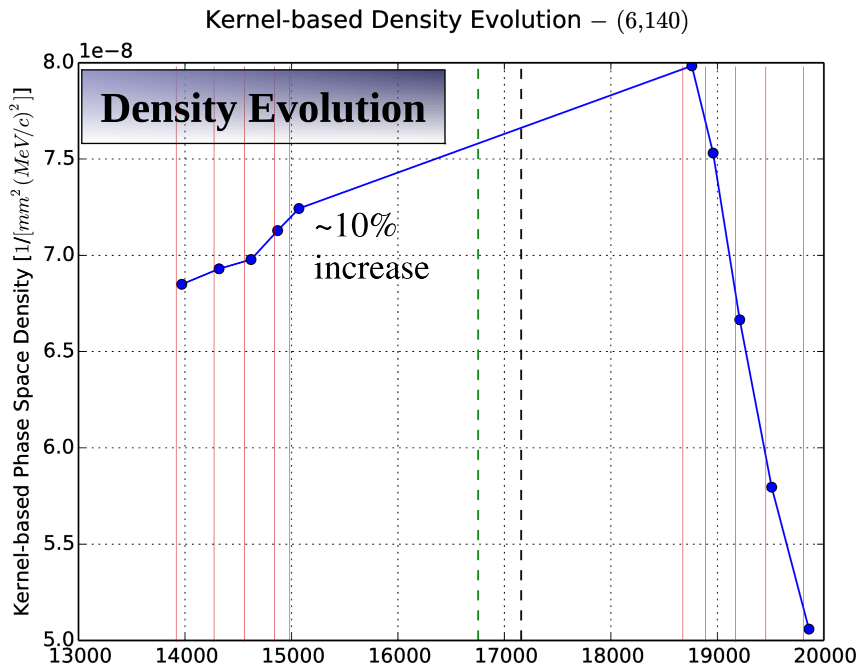
1) Draw a bounding hyper-rectangle around the contour under study.
2) Throw random Monte Carlo (MC) points at the hyper-rectangle. A fraction of MC points end up inside the contour under study.
3) Volume of the contour is the volume of the hyper-rectangle multiplied by the fraction of the MC points inside the contour.

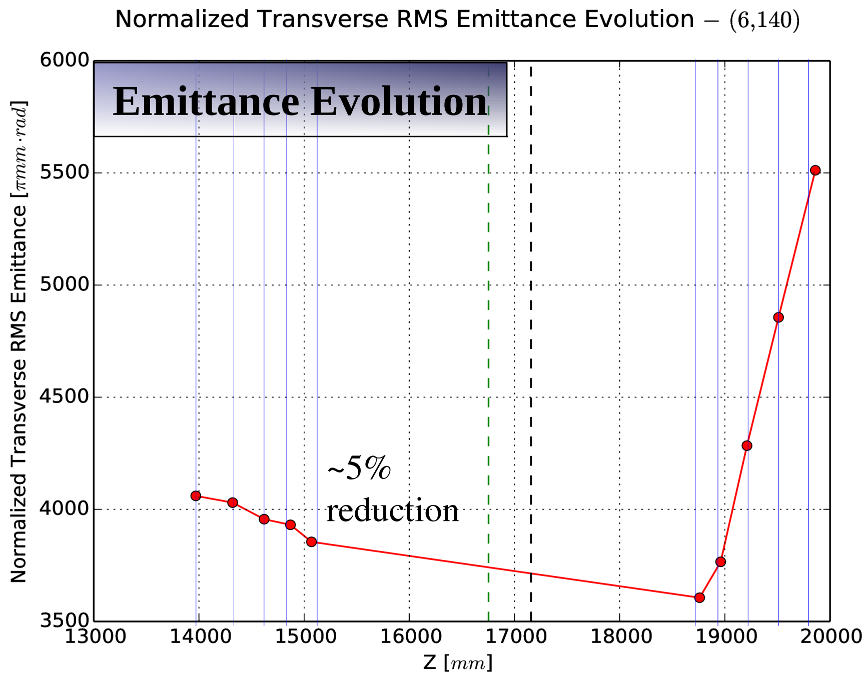
- Investigate the contours of different densities by plotting density versus volume.

Preliminary Application of KDE to MICE Data

- Run 8744:
  - 6-140 beam setting with LiH absorber in the channel and no currents in M1D & M2D coils.
  - Beam cooling according to changes in all 3 quantities.
  - Dashed vertical lines: FCU & FCD. Solid lines: locations of the tracker stations.
  - Good muon cut to discard muons which do not make it to TKD.

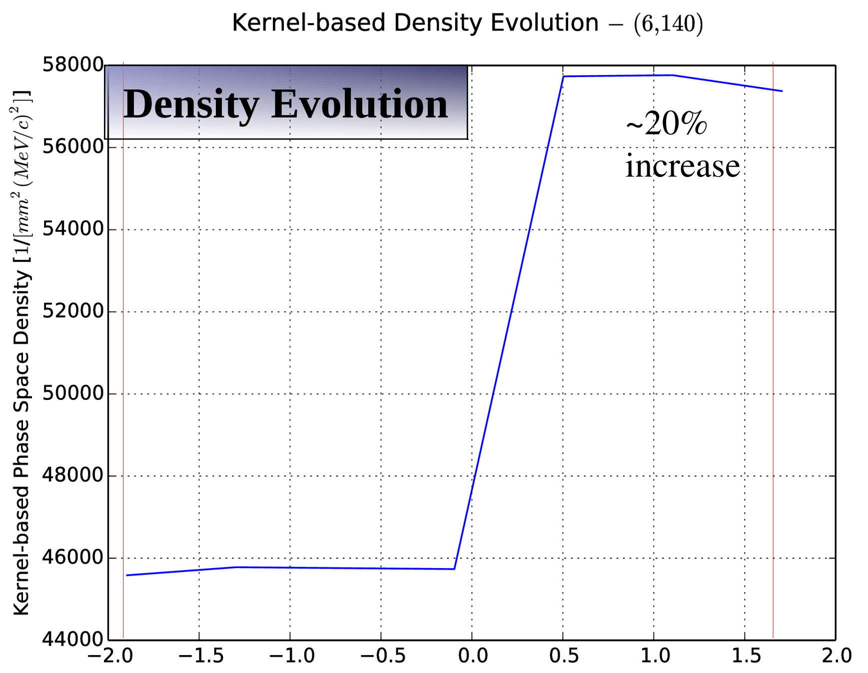


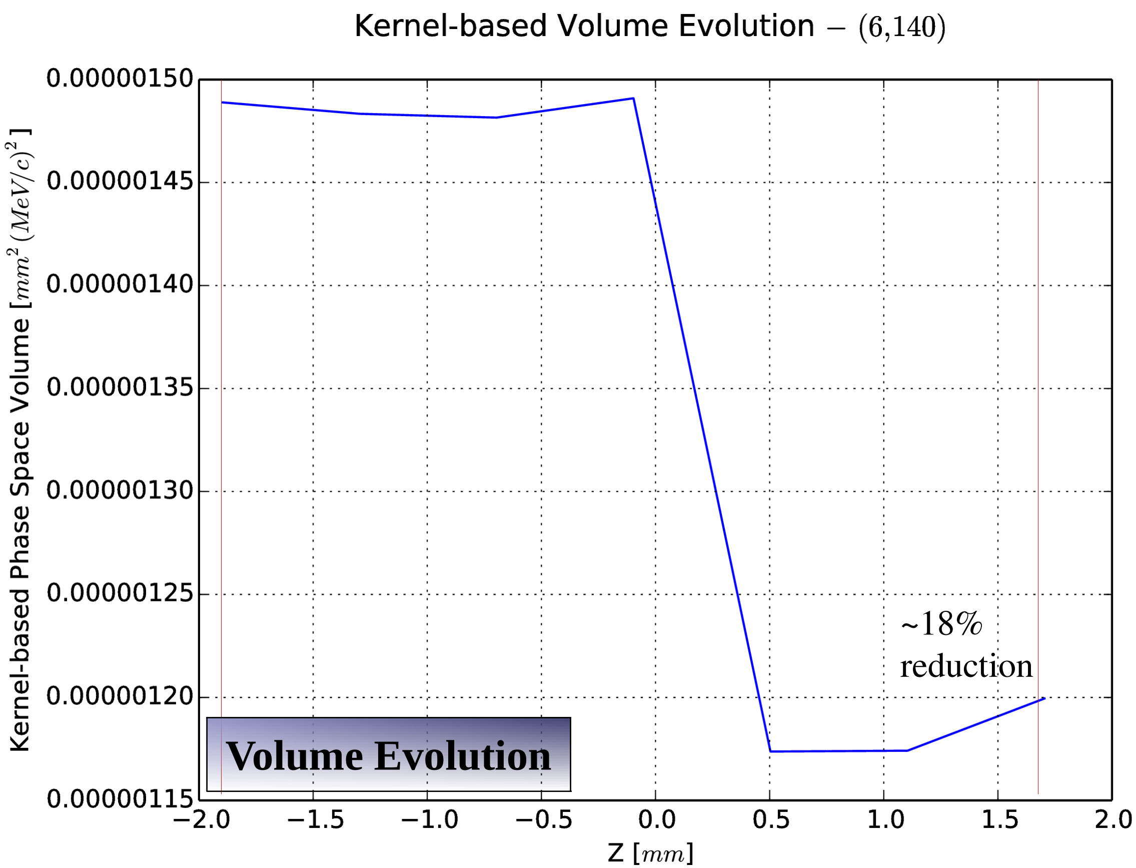


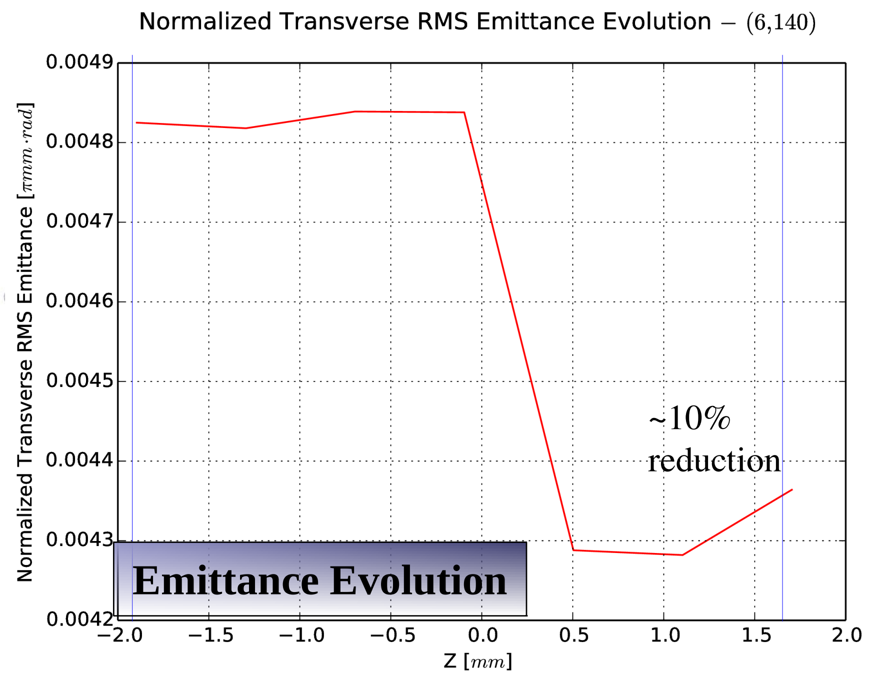


Comparison with Simulation

- 6-140 beam setting with LiH absorber in the channel and no currents in M1D & M2D coils.
- Muons only tracked between TKU to TKD reference planes.
- Beam cooling according to changes in all 3 quantities.
- Good muon cut to discard muons which do not make it to TKD.



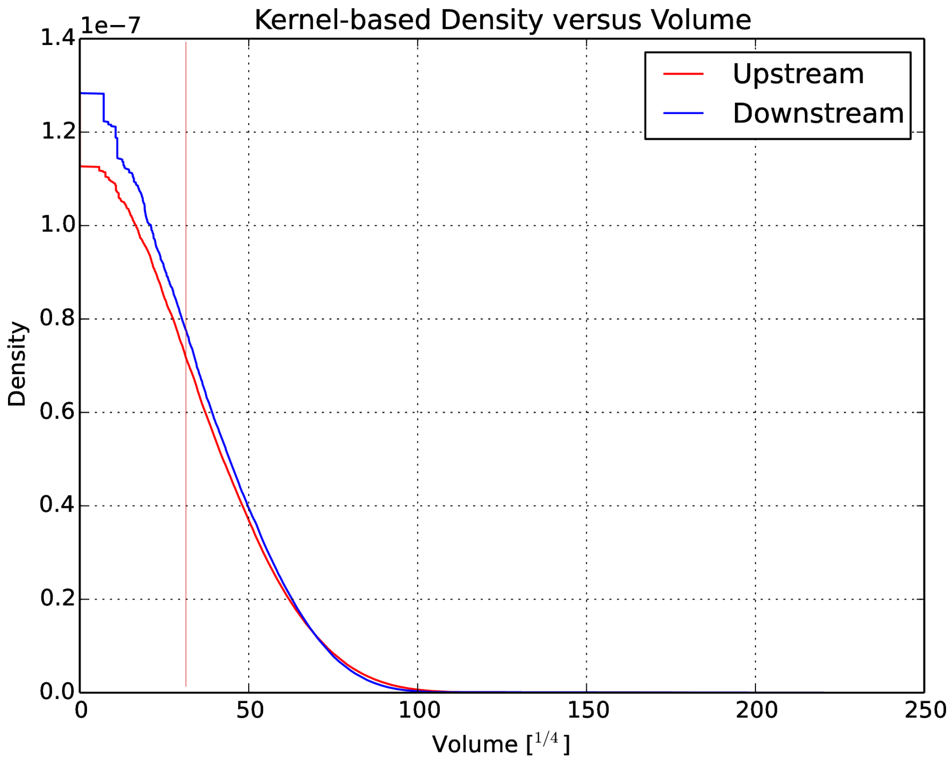




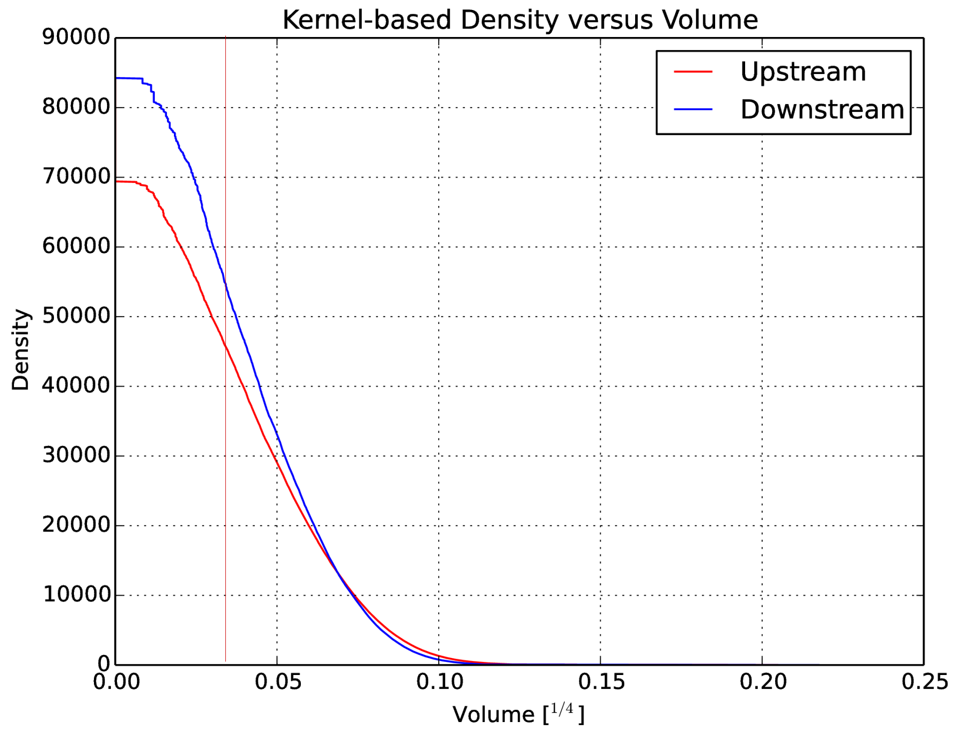
Density vs. Volume – Data vs. Simulation

- Other phase-space contours:
  - Volume*1/4*: the mean radius of hyper-ellipsoid.
  - Beam center at volume*1/4* = *0* & beam periphery at large volume*1/4* values.
  - Increase in core density while decrease in density in beam periphery.
  - **Red vertical lines**: the locations of the one sigma contour.

Density vs. Volume – Data



Density vs. Volume – Simulation



Conclusion and Future Prospects

- Density and volume behavior consistent with emittance. If the beam is distorted (e.g. at the locations of the disabled M1D), KDE needs to be extended from its standard form. Standard form: optimal bandwidth parameter is the same at the beam core and tail. One can adapt the bandwidth parameter to take the distances between muons into account. For this, I am applying the nearest-neighbor & the variable KDE methods.
- Necessary to make further improvements to the simulation routine.
- In 8744, KDE shows stronger cooling effect compared with emittance.
- Re-weight MICE beam to get to the ideal Gaussian distribution using KDE.
- We have plenty of excellent data! Will be looking at more data with different settings.