

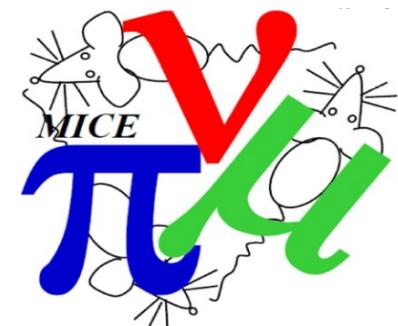
# Novel Application of Density Estimation in MICE



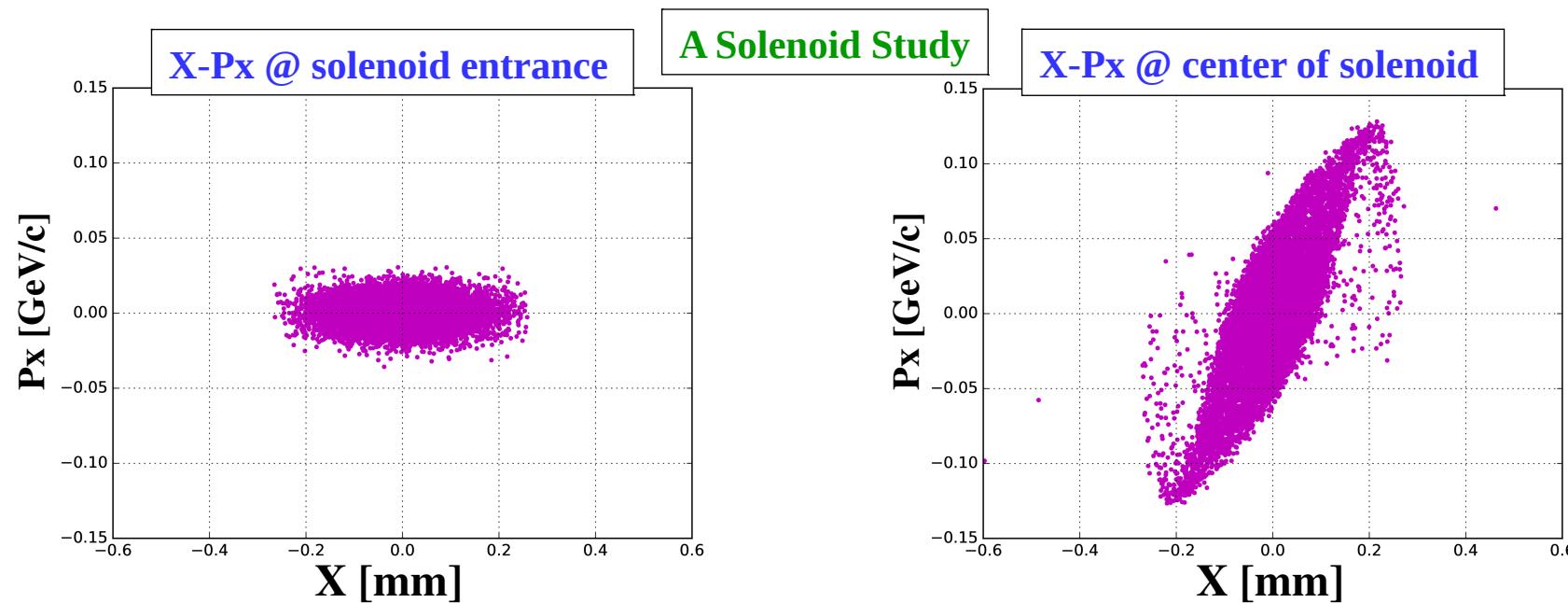
Tanaz Angelina Mohayai

CM47, RAL

February 14, 2017



# Motivation

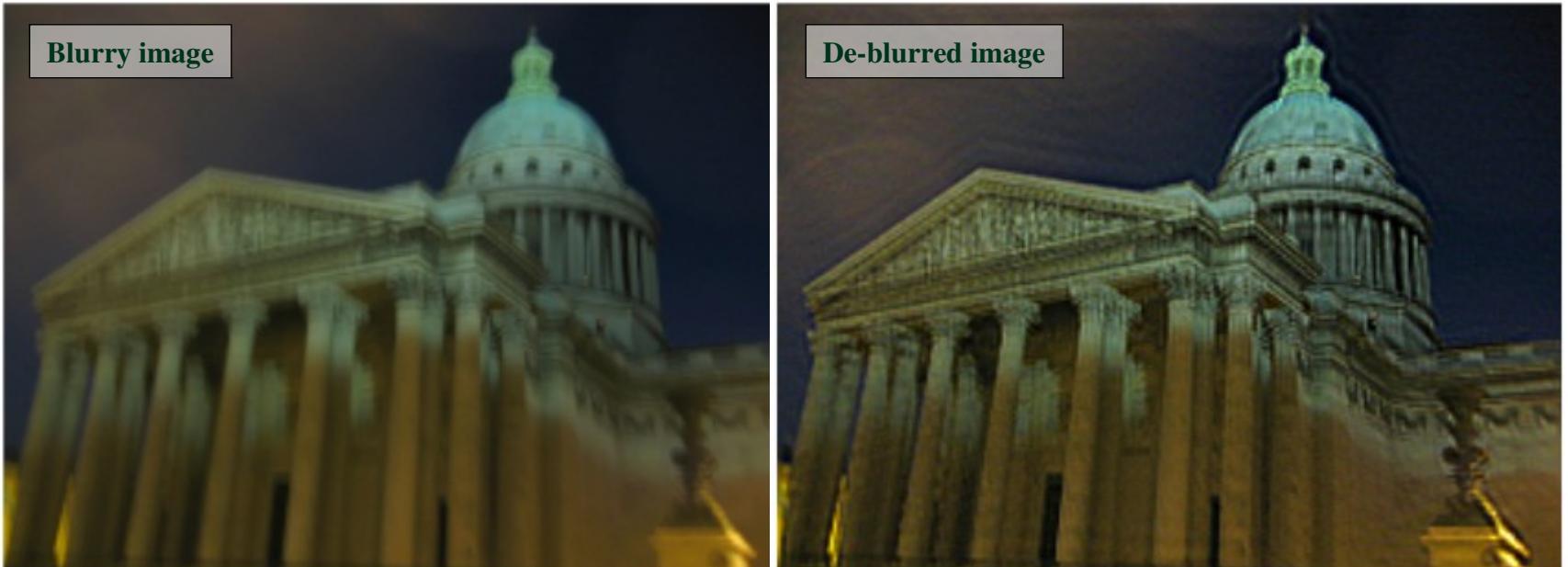


- **Problem:**
  - ★ RMS emittance assumes a Gaussian beam.
  - ★ Real-life beam is non-Gaussian (chromatic and non-linear effects).
- **Solution is Density Estimation:**
  - ★ Estimate PDF or density (normalized density) with few assumptions about the underlying distribution.
  - ★ Gives detailed single-particle diagnostics of the beam in a cooling channel.

# Kernel Density Estimation (KDE)

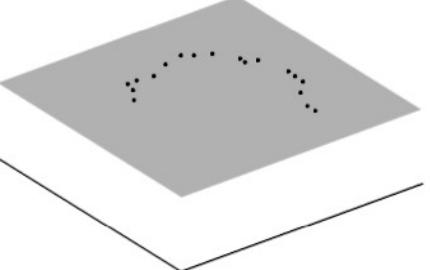
- Estimates the unknown probability density function (PDF) or density (normalized density) using kernels (smooth weight functions of certain widths).

**Image processing with KDE**

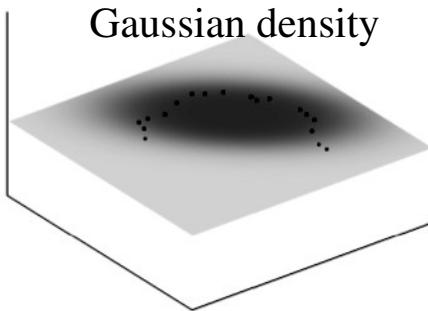


D. Krishnan et al., "Blind Deconvolution Using a Normalized Sparsity Measure", DOI: 10.1109/CVPR.2011.5995521

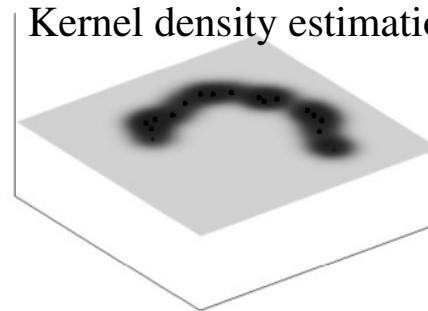
Actual distribution



Gaussian density



Kernel density estimation



- Kernel functions at each data point.
- Powerful single muon measurement tool for MICE.

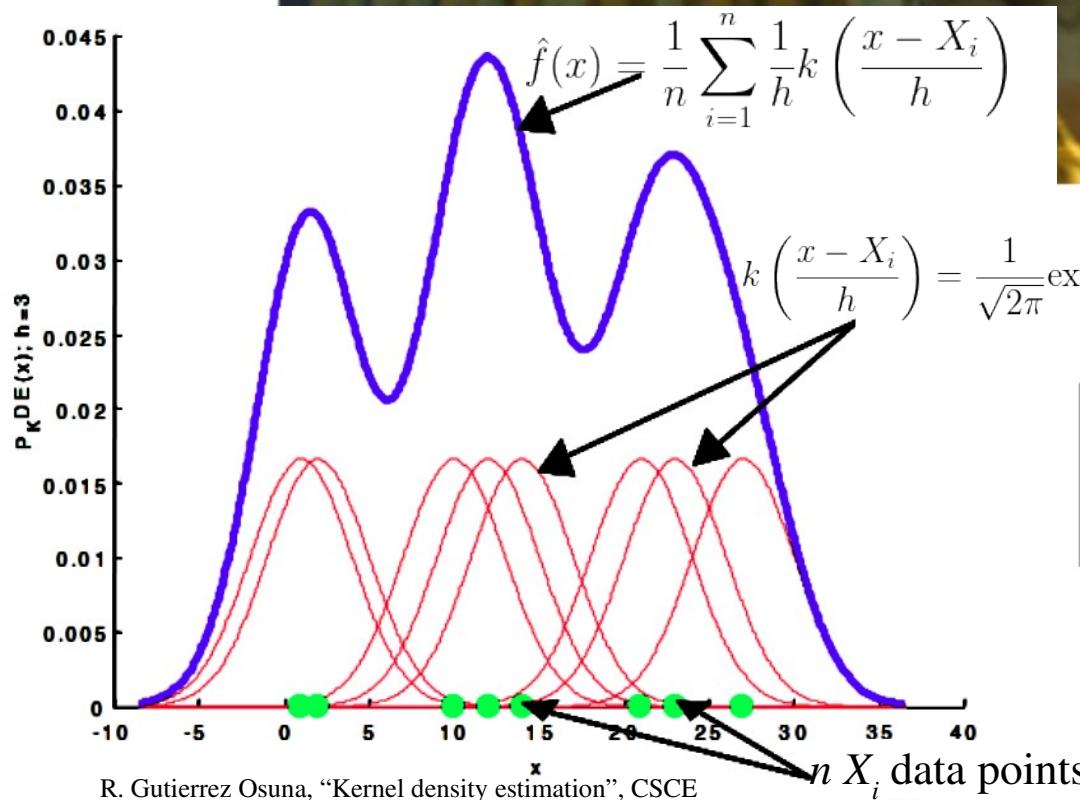
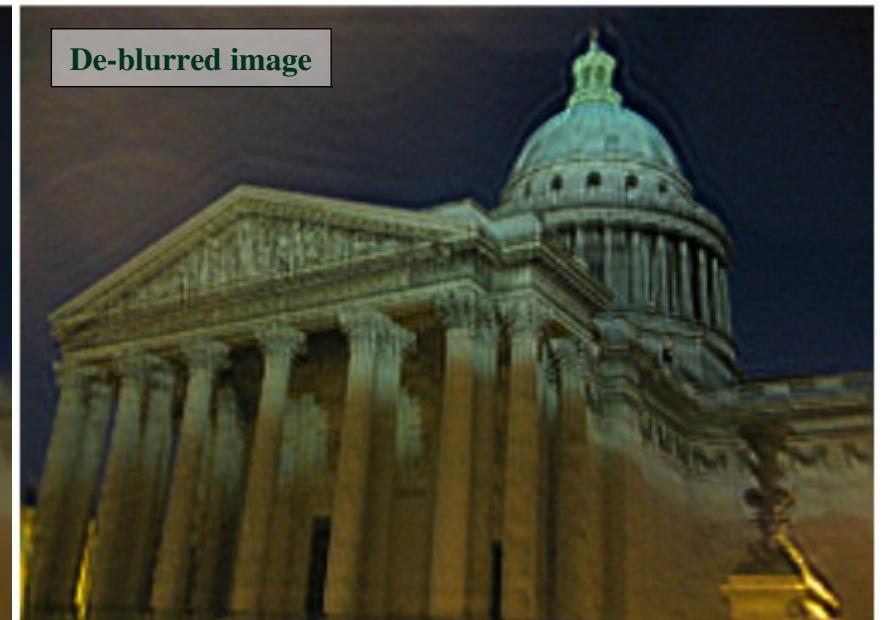
M. Rousson et al., "Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation", DOI:10.1007/978-0-387-68343-0\_13

**Power of KDE**

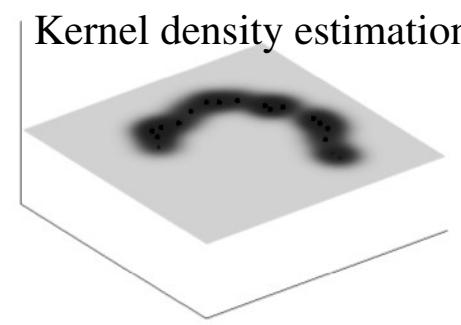
# Kernel Density Estimation (KDE)

- Estimates the unknown probability density function (PDF) or density (normalized density) using kernels (smooth weight functions of certain widths).

Image processing with KDE



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



Power of KDE

- Kernel functions at each data point.
- Powerful single muon measurement tool for MICE.

# Kernel Density Estimation (KDE) Validation

- Parameters affecting KDE density: **bandwidth parameter**,  $h$ , sample size,  $n$ , dimensionality,  $d$ .
  - KDE error: difference between true density  $f(x)$  and estimated density.
  - Common measure of error: mean integrated square error (MISE),

$$\text{Mean Integrated Square Error (MISE)},$$

$$\text{MISE}(\hat{f}(x)) = \mathbb{E} \int \left[ \hat{f}(x) - f(x) \right]^2 dx$$

- Bias-variance trade-off: reducing the bias (related to  $h^2$ ) leads to an increase in variance (inversely proportional to  $h$ ):

- ★  $x$ -coordinates of 500 muons at the entrance to TKU.

☆ Unknown true PDF.

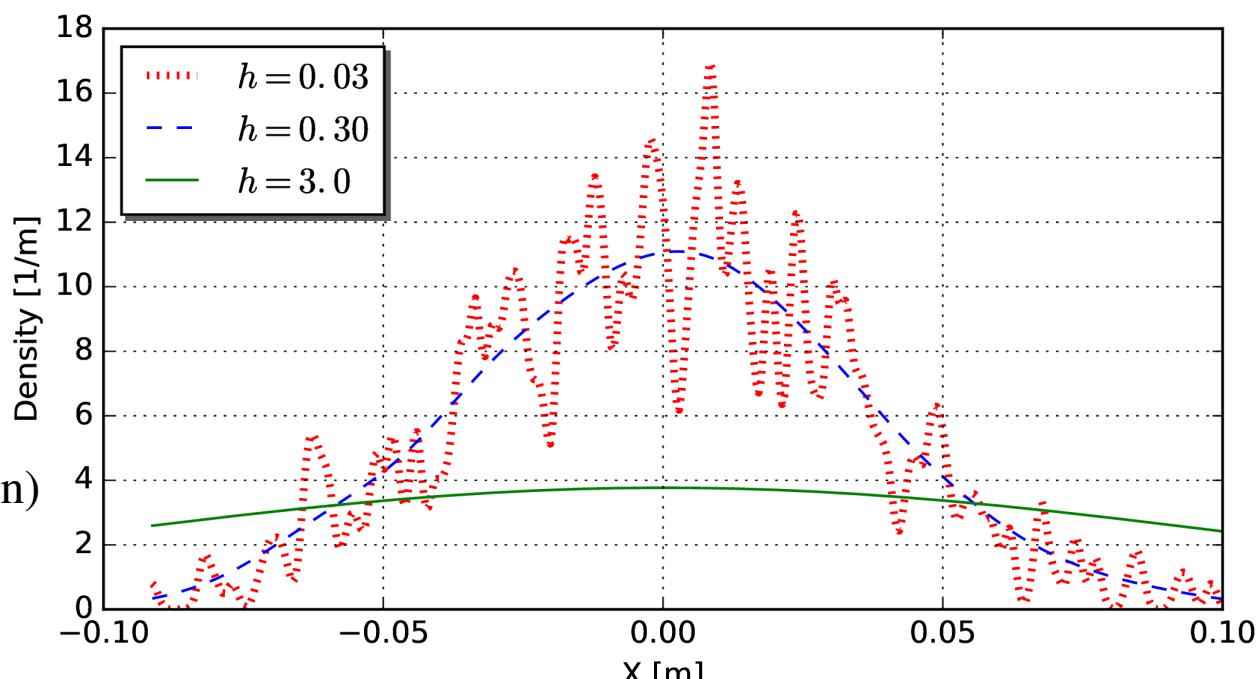
- ★  $h = 0.3$  (optimal  $h$  determined from minimizing MISE term above) reveals a Gaussian,

$$h = 1.06 n^{-1/5} \sigma$$

$$h = h_{\text{factor}} \sigma \quad (\sigma: \text{standard deviation})$$

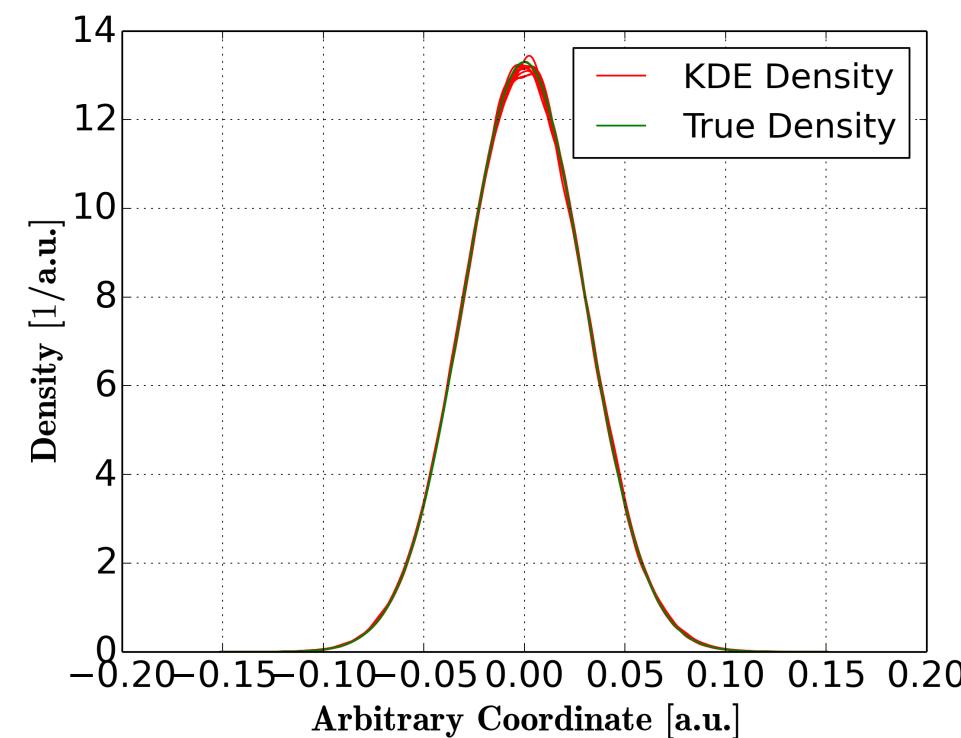
★  $h = 3.0$  over-smoothes the PDF.

\*  $h \equiv 0.03$  reveals a noisier PDE.

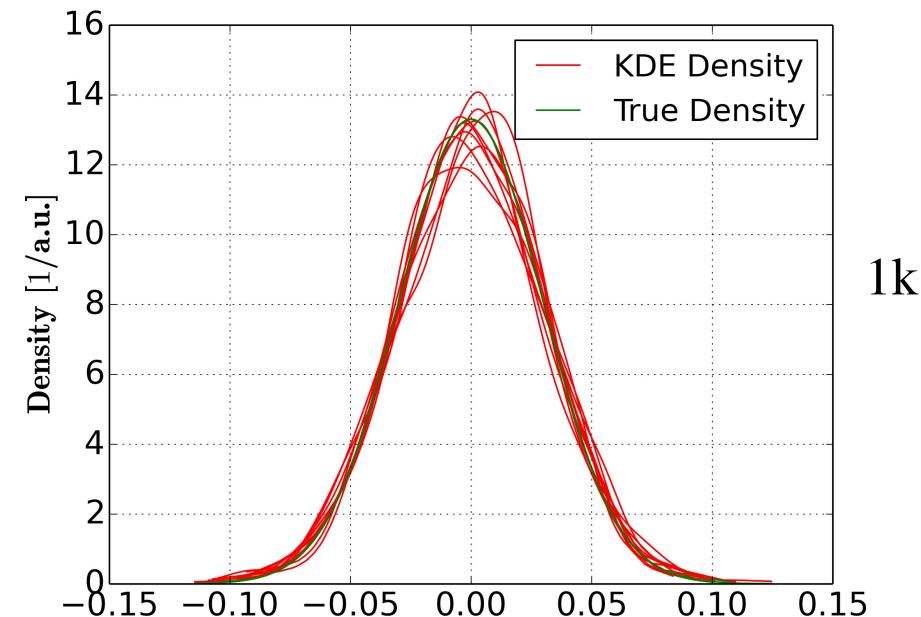


# KDE Validation in 1D – sample size study, True vs. KDE

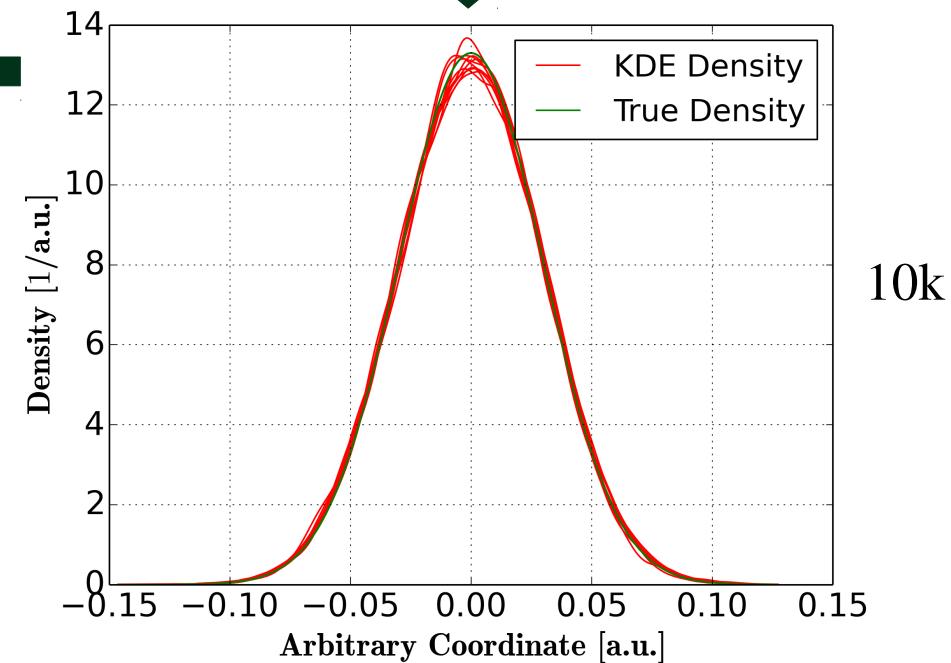
- Generated 10 Gaussian distributions ( $\sigma = 0.03$  a.u), each with 1k, 10k, and 100k data points.
- Compared their KDE and true densities.
- The bandwidth is optimal (minimized MISE).



100k



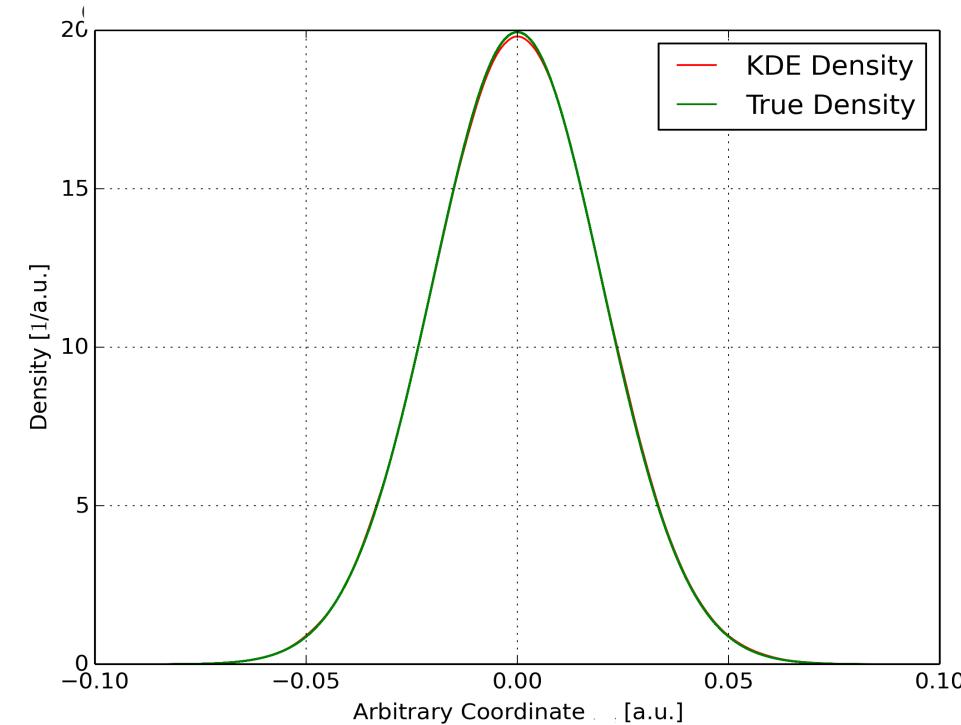
1k



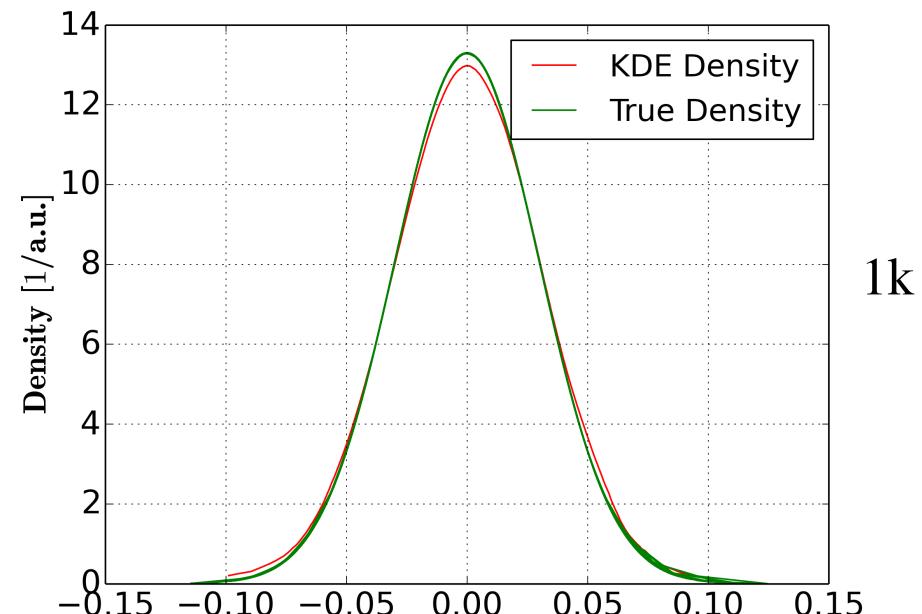
10k

# KDE Validation in 1D cont. – sample size study, True vs. KDE

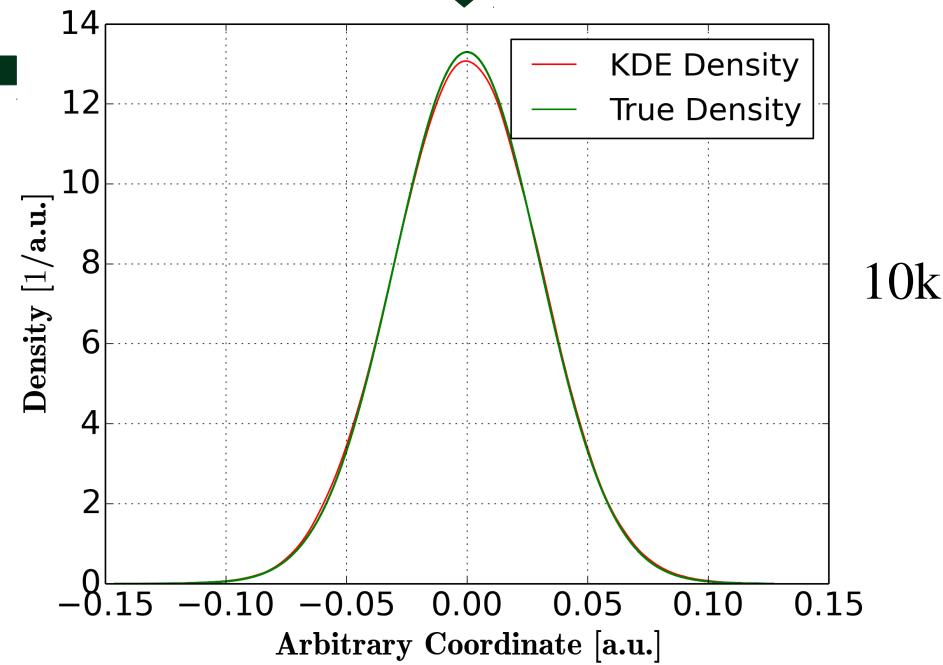
- Generated 10 Gaussian distributions ( $\sigma = 0.03$  a.u.), each with 1k, 10k, and 100k data points.
- Compared the **averages** of their KDE and true densities.
- The bandwidth is optimal (minimized MISE).



100k  
←



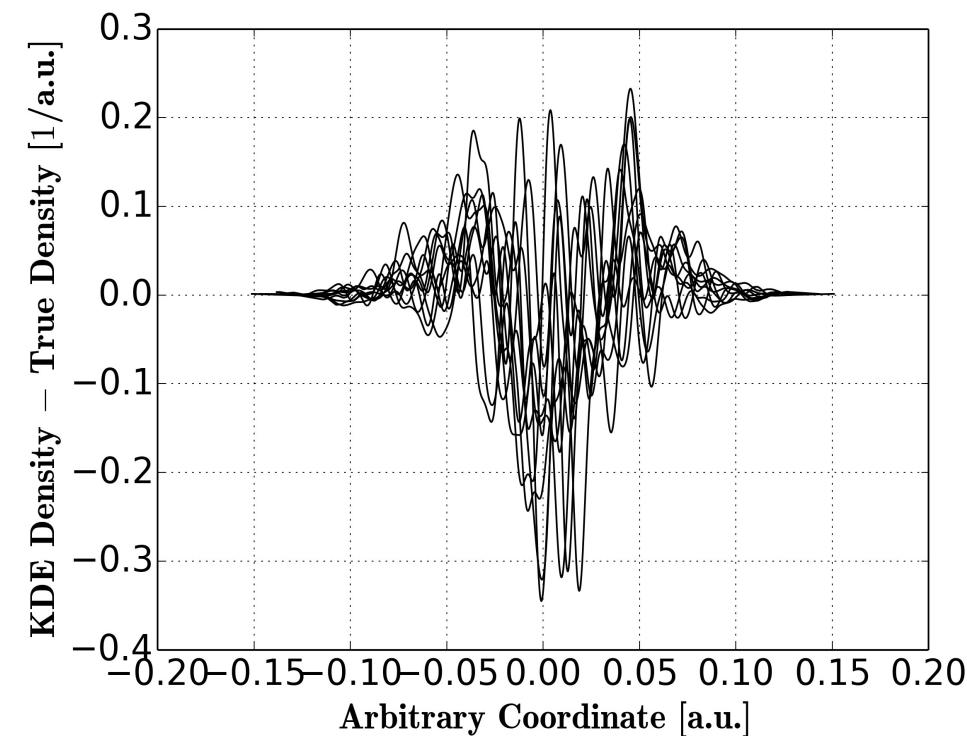
1k



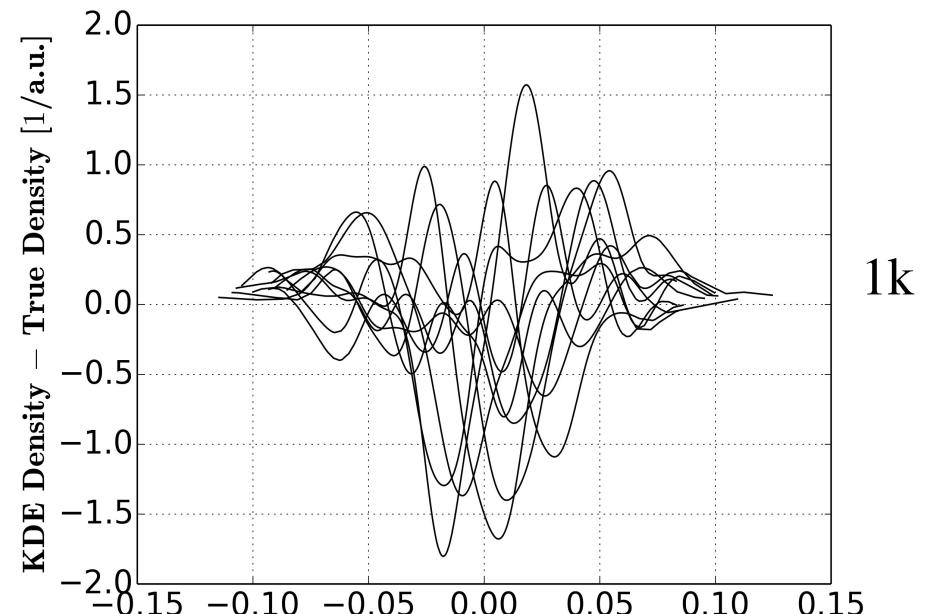
10k

# KDE Validation in 1D cont. – sample size study, True vs. KDE

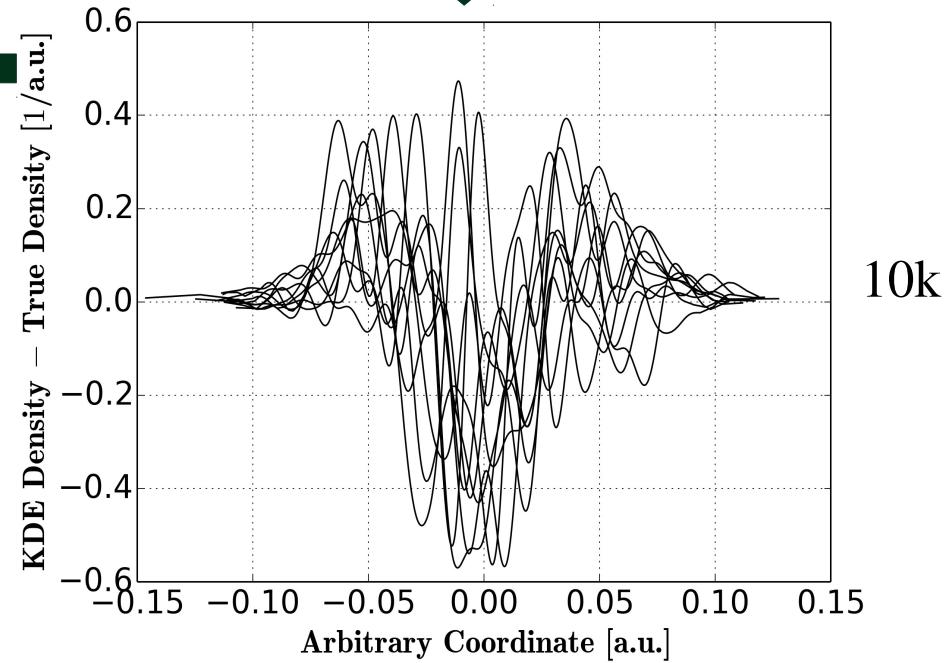
- Generated 10 toy Gaussian distributions ( $\sigma = 0.03$  a.u.), each with 1k, 10k, and 100k data points.
- Compared their **KDE errors** (differences between KDE density and true densities).
- The bandwidth is optimal (minimized MISE).



100k



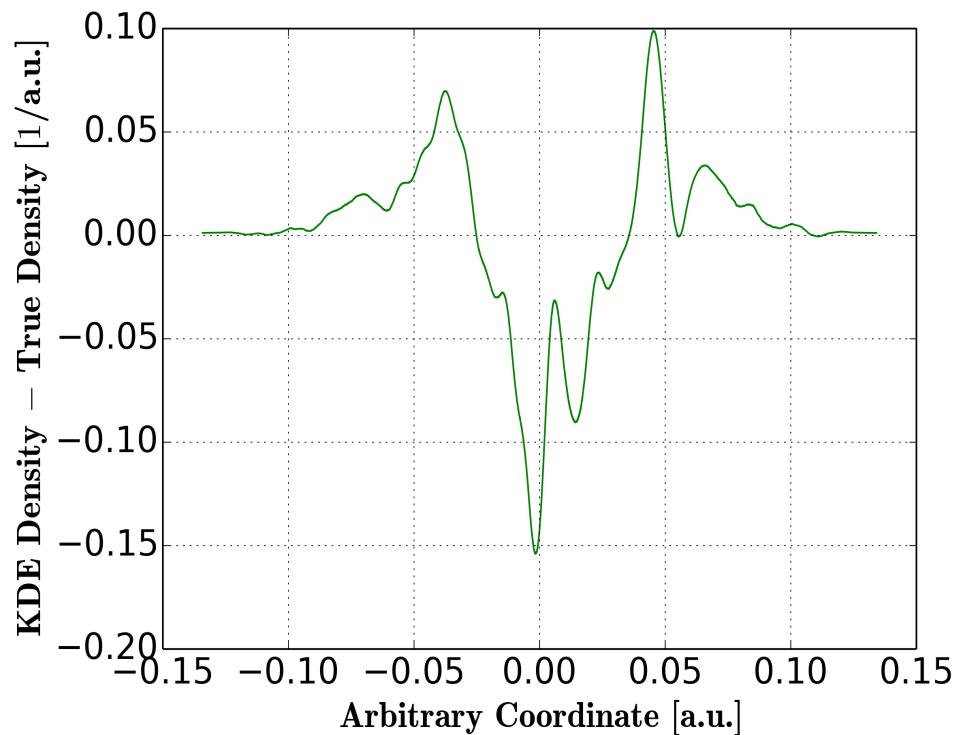
1k



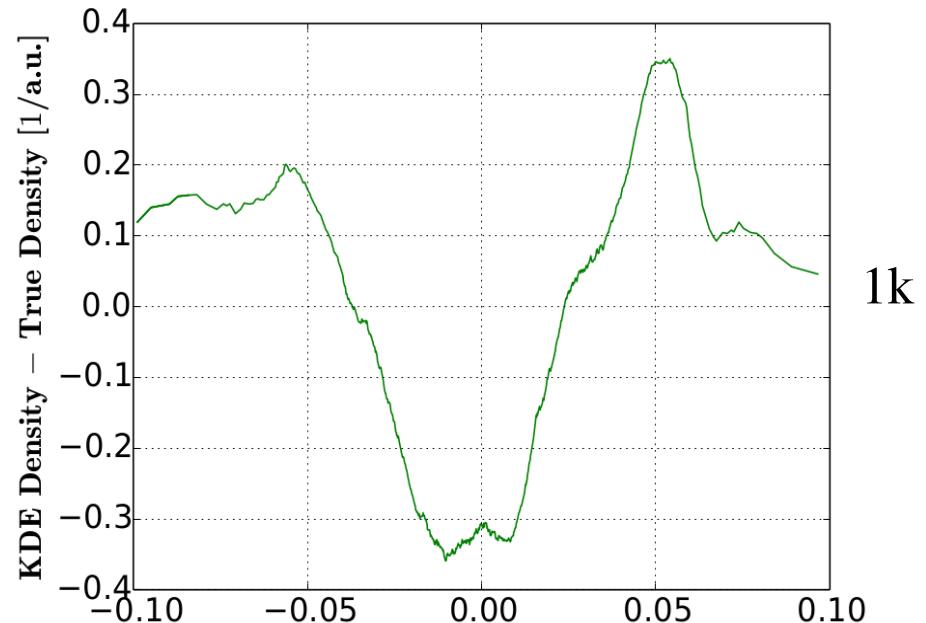
10k

# KDE Validation in 1D cont. – sample size study, True vs. KDE

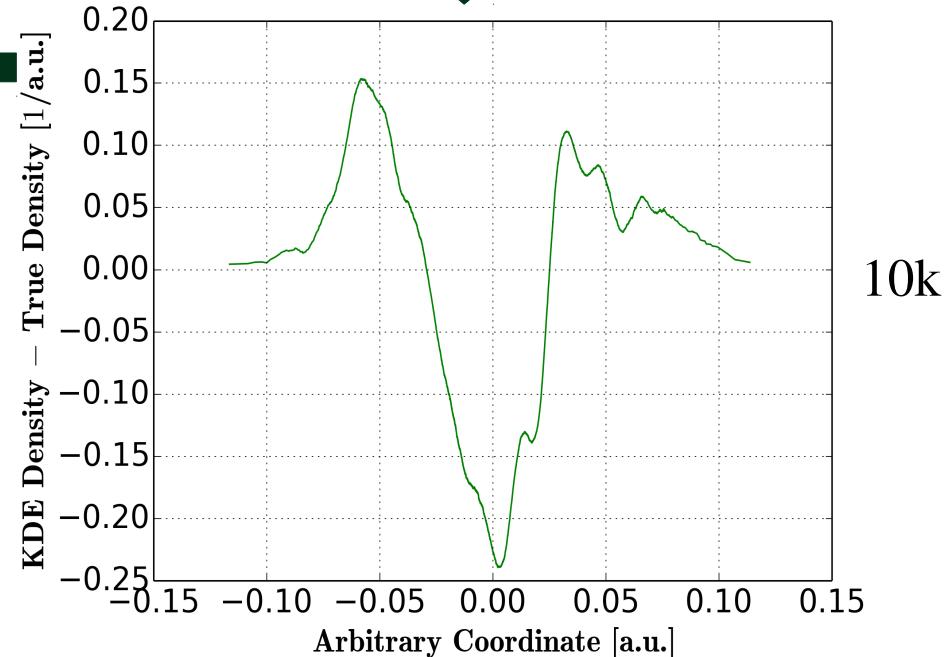
- Generated 10 toy Gaussian distributions ( $\sigma = 0.03$  a.u.), each with 1k, 10k, and 100k data points.
- Compared their **KDE error averages**.
- The bandwidth is optimal (minimized MISE).



100k



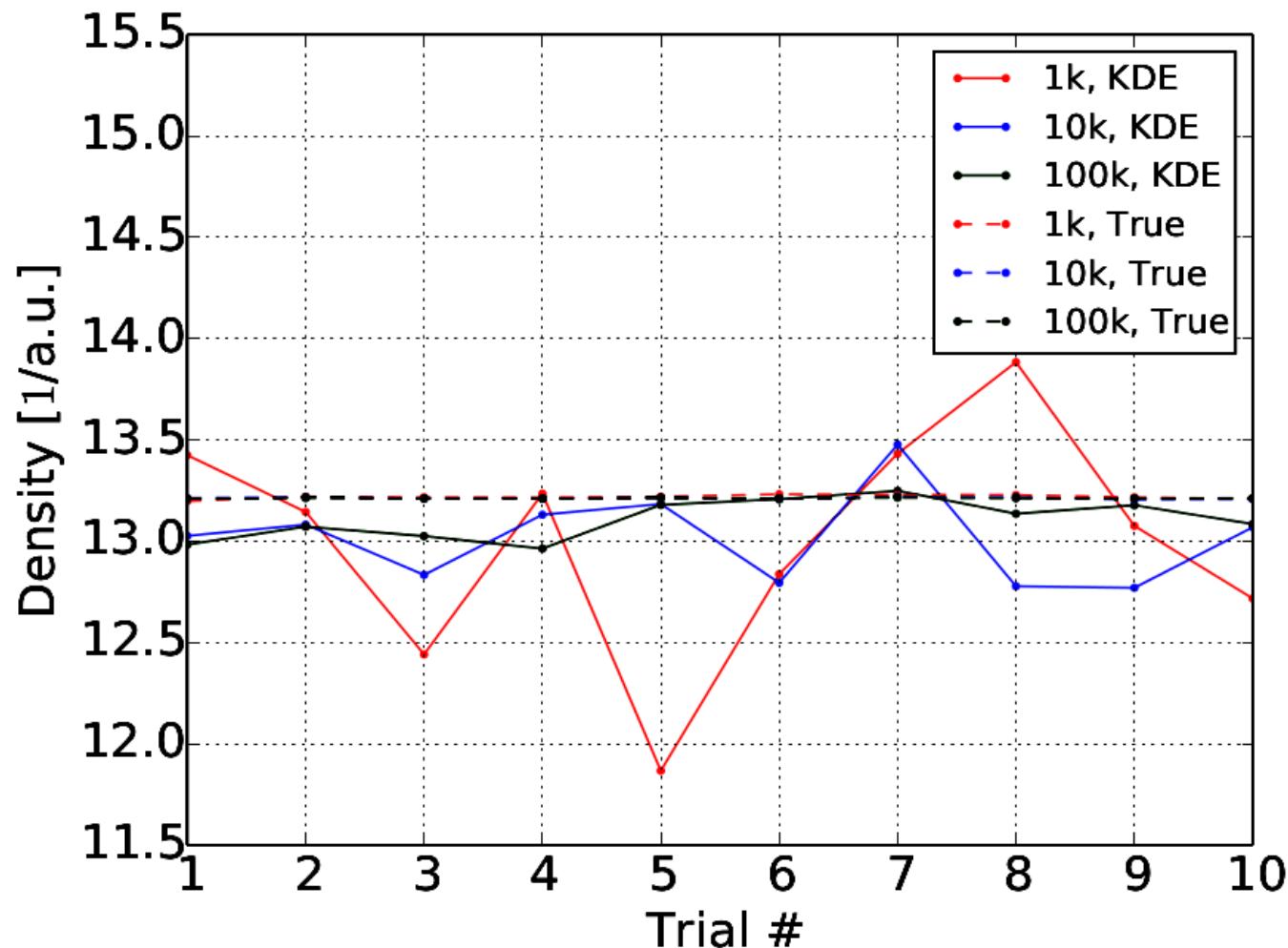
1k



10k

# KDE Validation in 1D cont. – summary

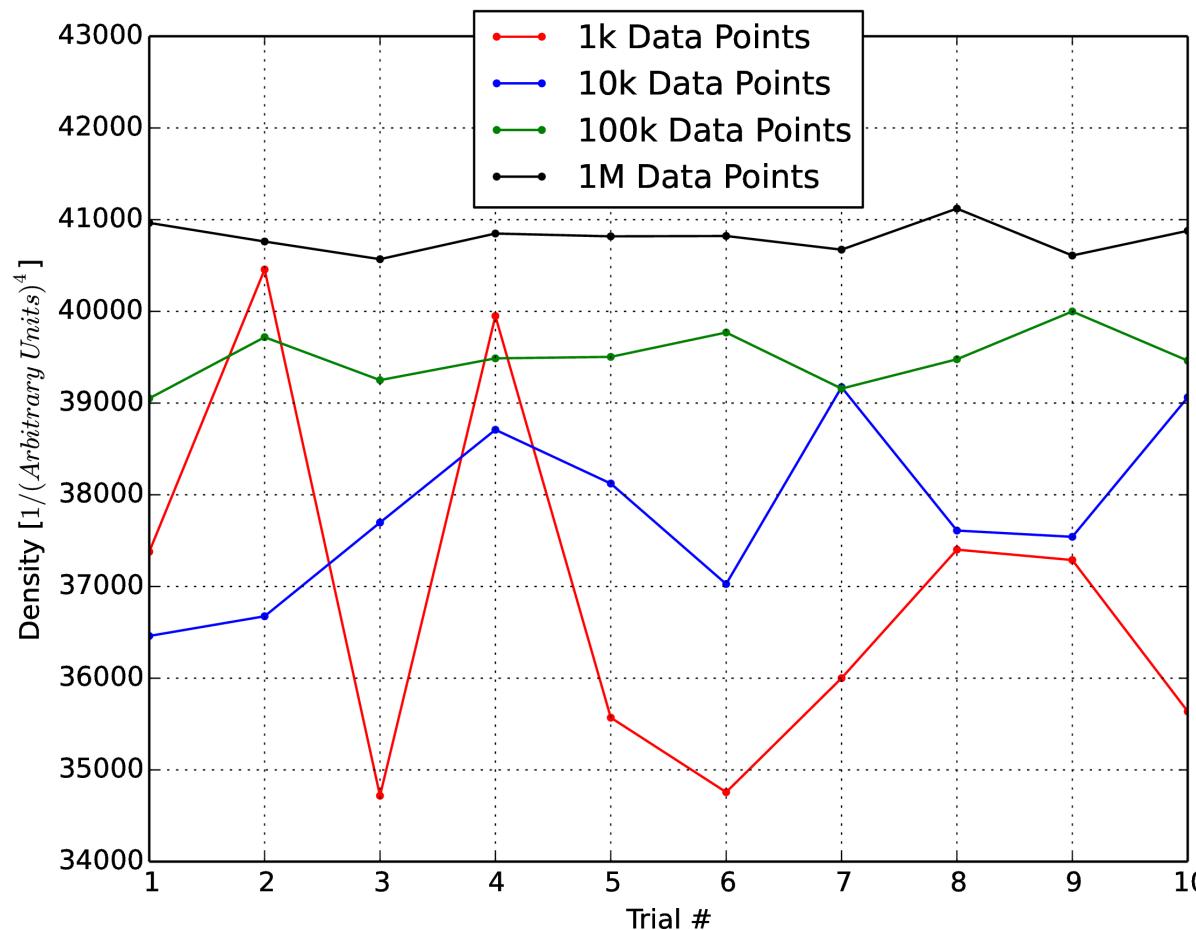
- To summarize, isolated each curve's peak density (density curve enclosing 9% of the sample size):
  - ★ KDE density stabilizes and approaches the true density curve as sample size grows.
  - ★ Slight increase in mean density with growing sample size (caused by the optimal bandwidth's dependence on sample size).



# KDE Validation in 4D

- Used the same 4-dimensional 10 distributions as previous slides; KDE applied to all four coordinates ( $\sigma_1 = \sigma_2 = 0.03$  a.u and  $\sigma_3 = \sigma_4 = 0.02$  a.u).
- KDE density stabilizes as sample size grows.
- Increase in mean density with sample size (compared with 1D) – optimal bandwidth's depends on sample size **and dimension variable, d**:

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



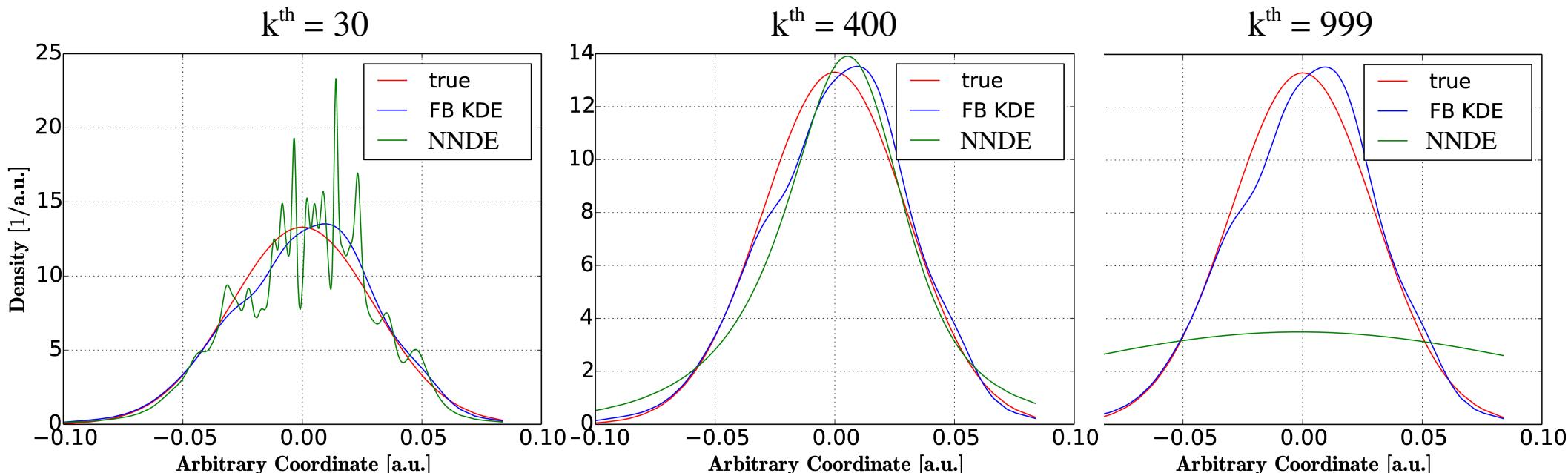
# Nearest Neighbor Density Estimation (NNDE)

- KDE: equal contribution from each data point:
  - ★ Some sensitivity to distribution tails.
  - ★ Kernel width fixed everywhere.
- NNDE: each data point's contribution depends on their distances to their neighboring points.
- Useful for long-tailed distributions (e.g. MICE beam in regions of turned off downstream match coils).
- Parameters affecting NNDE density:  $k^{\text{th}}$  nearest neighbor, sample size,  $n$

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{d_i} k\left(\frac{|x - X_i|}{2d_i^2}\right)$$

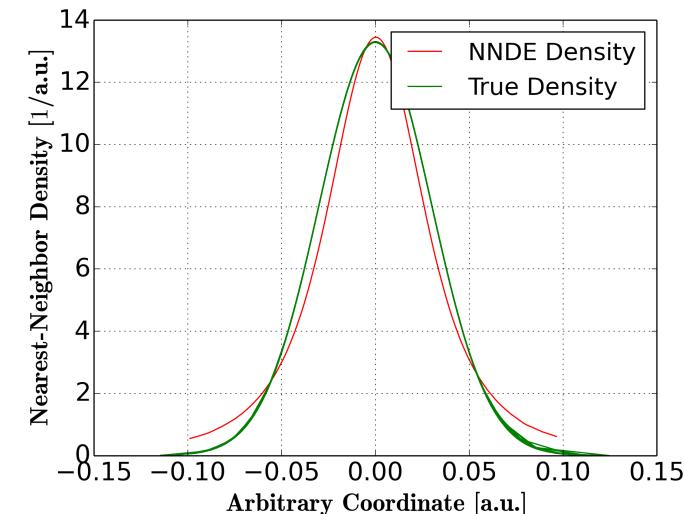
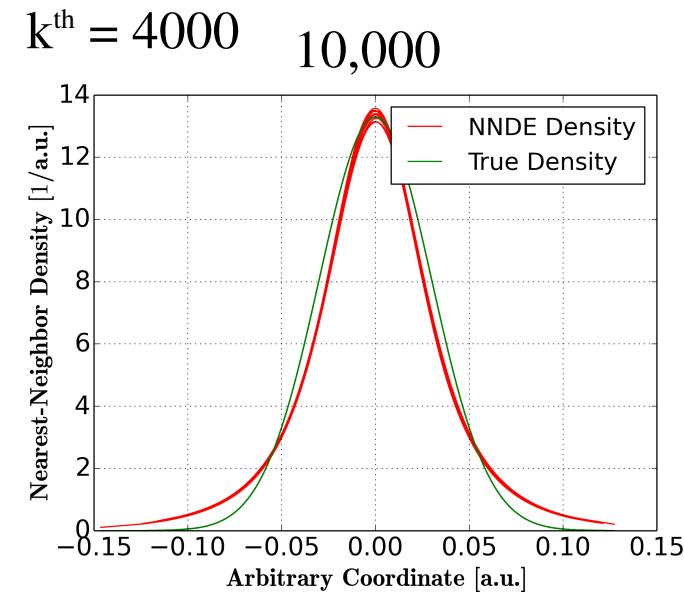
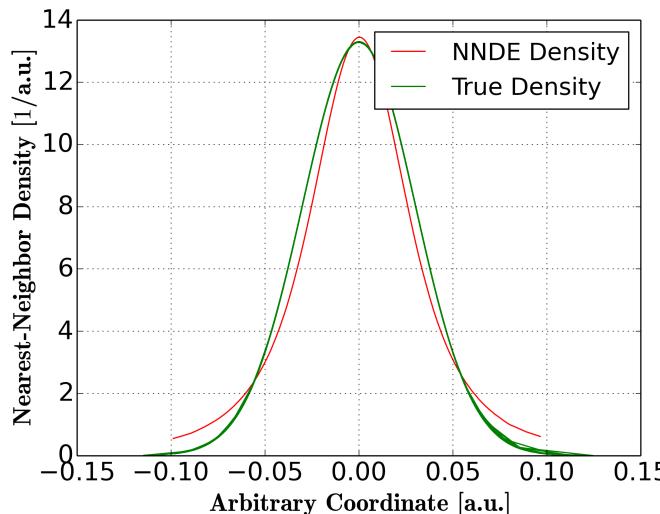
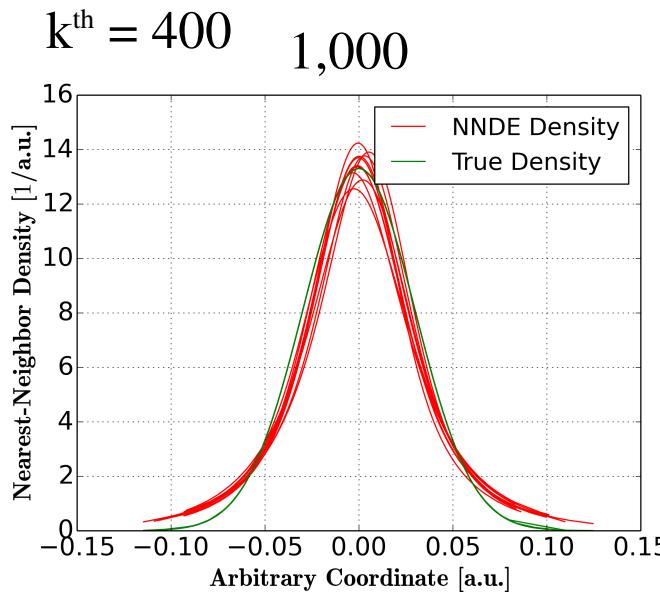
$d_i$ : the euclidean distance between  $i^{\text{th}}$  data point to  $k^{\text{th}}$  nearest neighbor data point.

- Generated Gaussian distribution with 1k data points ( $\sigma = 0.03$  a.u,  $n = 1000$ ).
- Compared NNDE with KDE (Fixed Bandwidth, FB KDE) and true density.



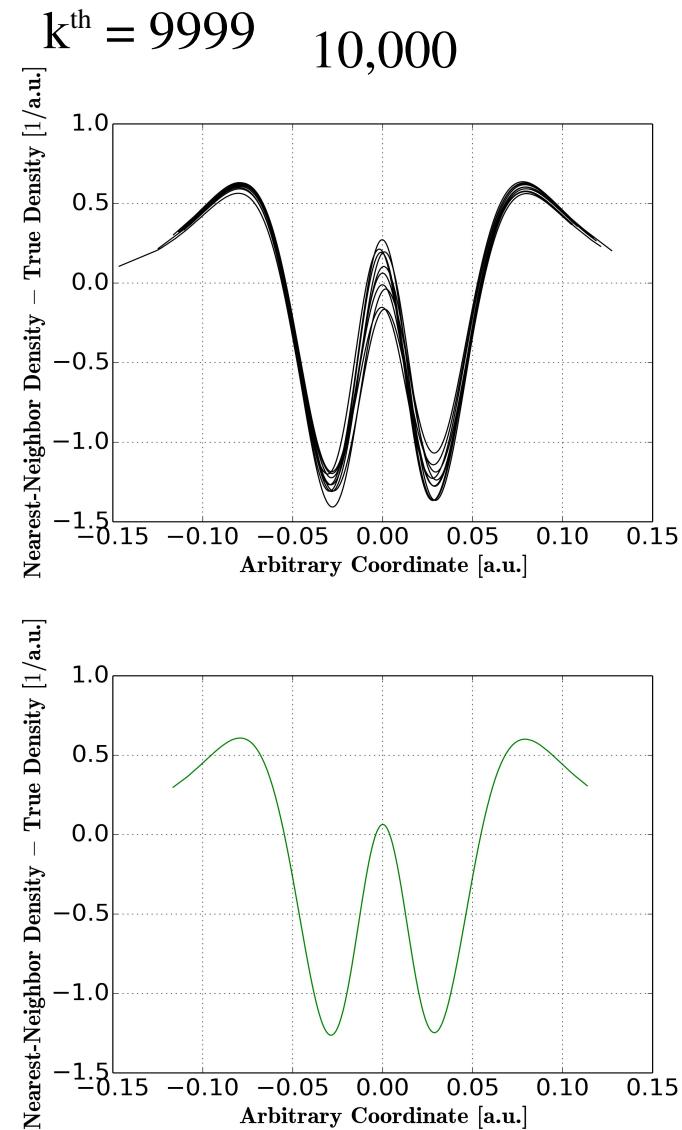
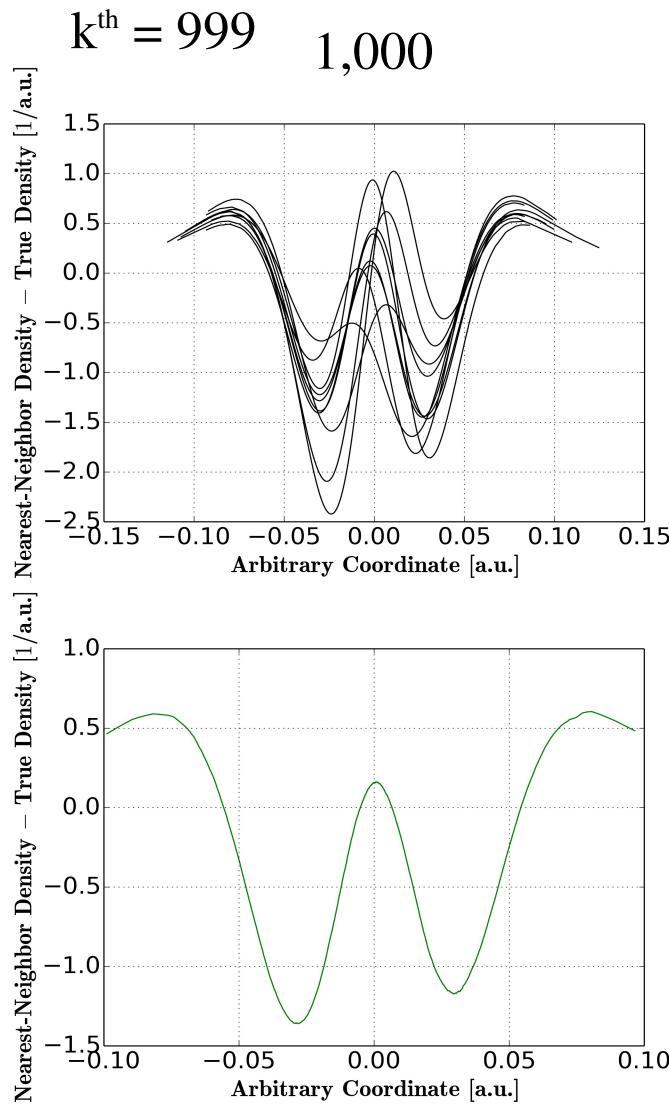
# NNDE Validation in 1D – sample size study, True vs. NNDE

- Generated 10 Gaussian distributions ( $\sigma = 0.03$  a.u), each with 1k, 10k data points, compared their NNDE and true densities.



# NNDE Validation in 1D – sample size study, True vs. NNDE

- ★ Generated 10 Gaussian distributions ( $\sigma = 0.03$  a.u), each with 1k, 10k data points, compared their errors and the averages of their errors.

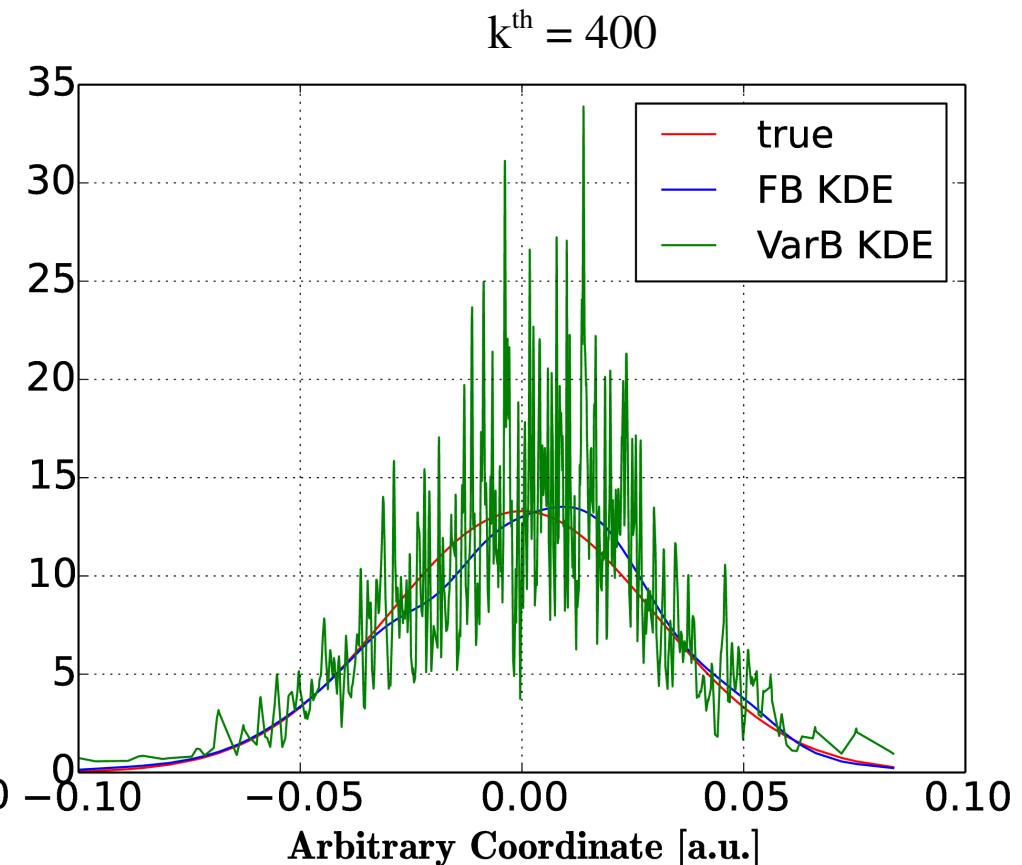
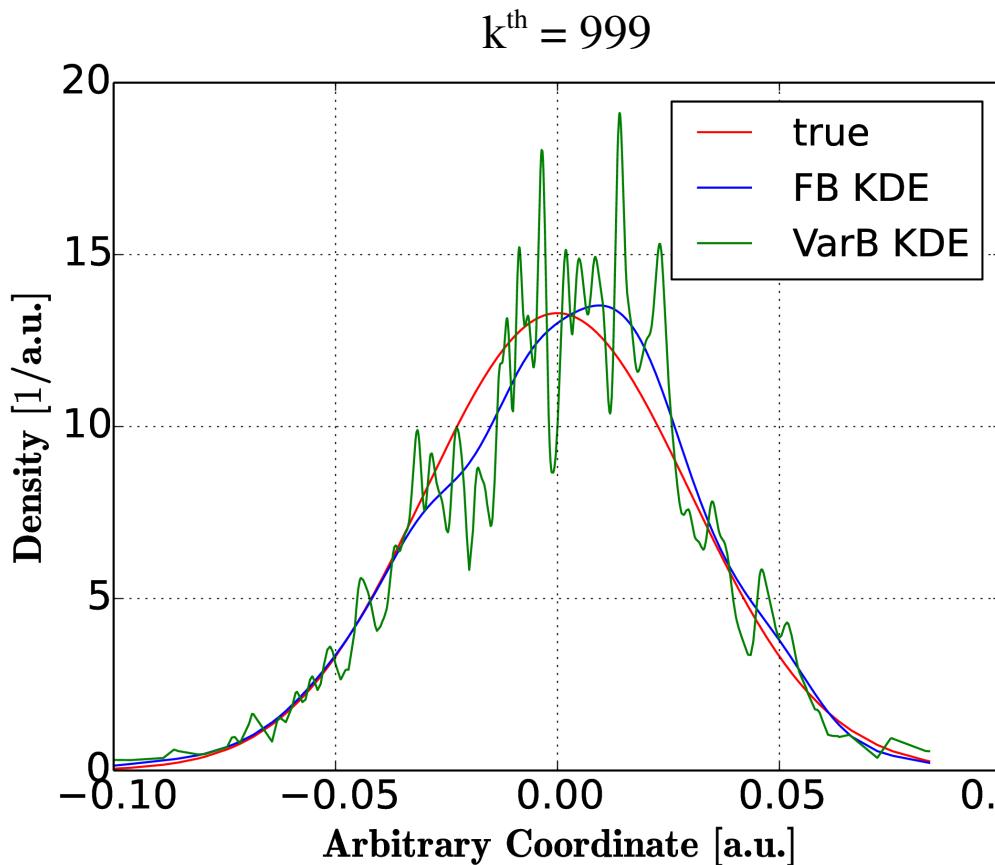


# Combining KDE and NNDE – Variable KDE

- ★ Hybrid of KDE and NNDE:

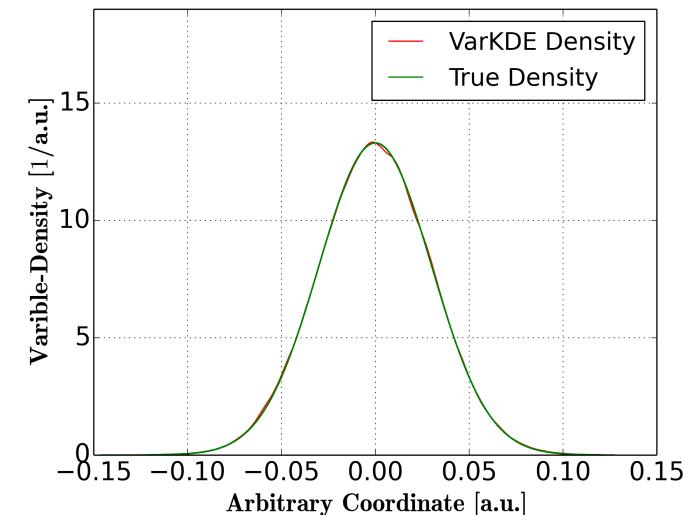
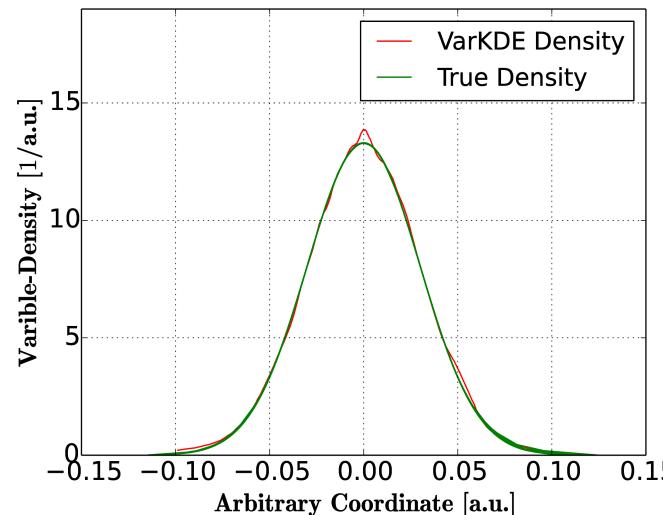
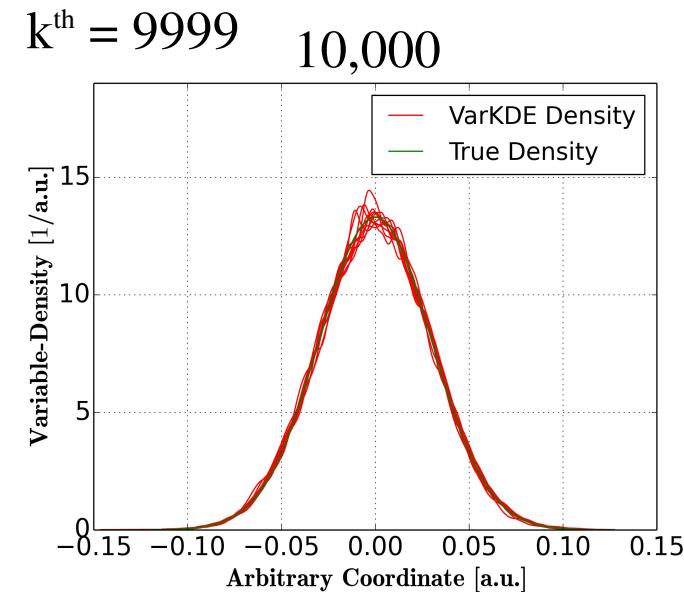
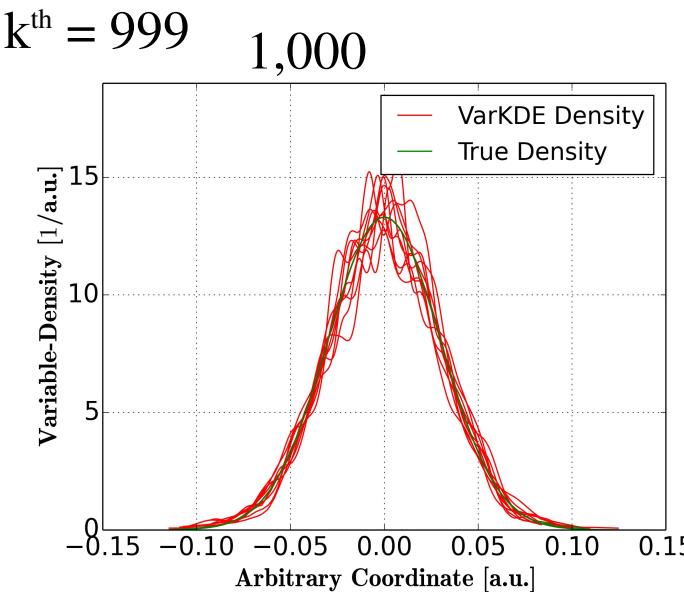
$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{hd_i} k\left(\frac{|x - X_i|}{2h^2 d_i^2}\right)$$

$d_i$ : the euclidean distance between  $i^{\text{th}}$  data point to  $k^{\text{th}}$  nearest neighbor data point.



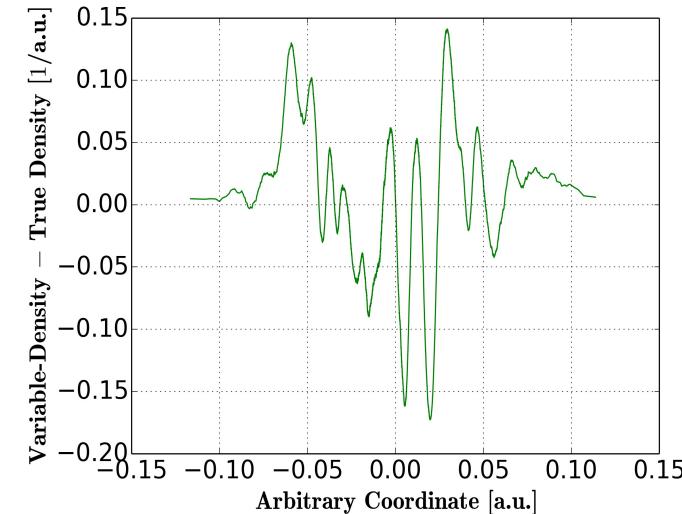
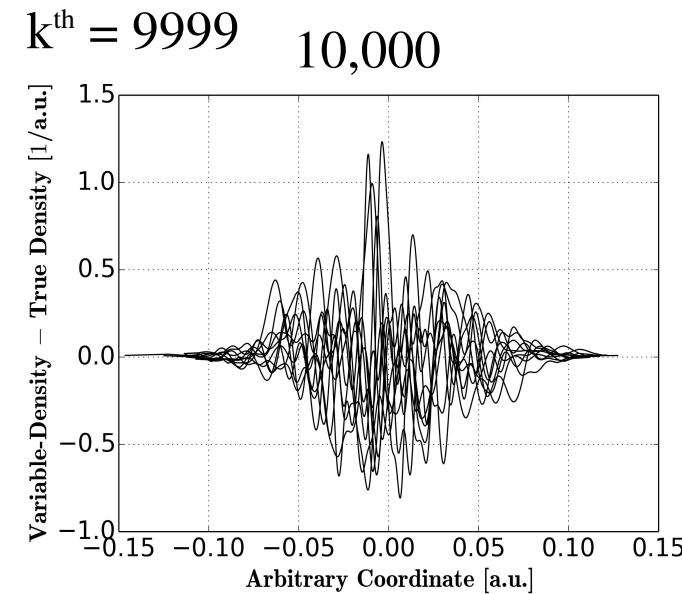
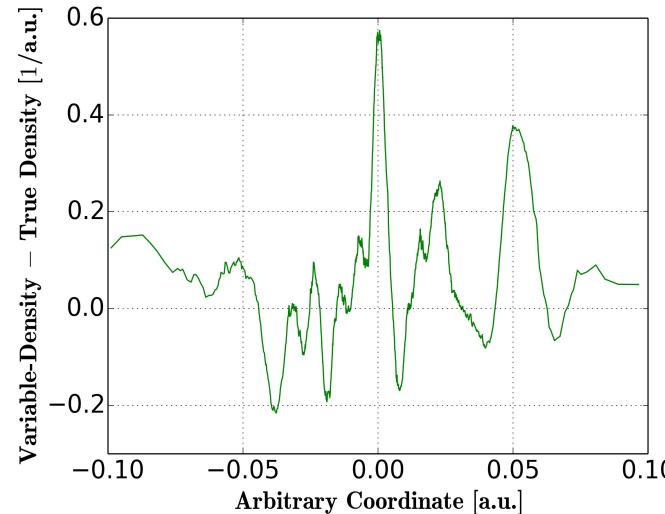
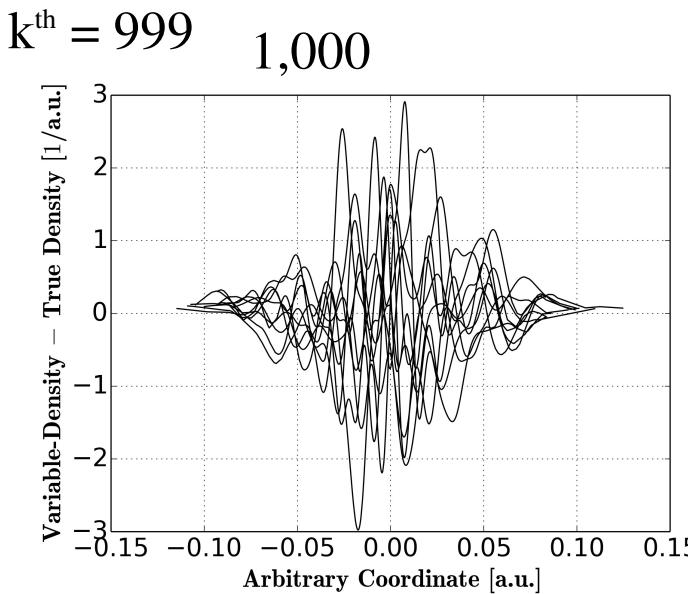
# VarKDE Validation in 1D – sample size study, True vs. VarKDE

- ★ Generated 10 Gaussian distributions ( $\sigma = 0.03$  a.u.), each with 1k, 10k data points, compared their VarKDE and true densities. Note: inserted a factor of 4 in VarKDE to make the curves smoother.



# VarKDE Validation in 1D – sample size study, True vs. VarKDE

- Generated 10 Gaussian distributions ( $\sigma = 0.03$  a.u), each with 1k, 10k data points, compared their errors. Note: inserted a factor of 4 in VarKDE to make the curves smoother.

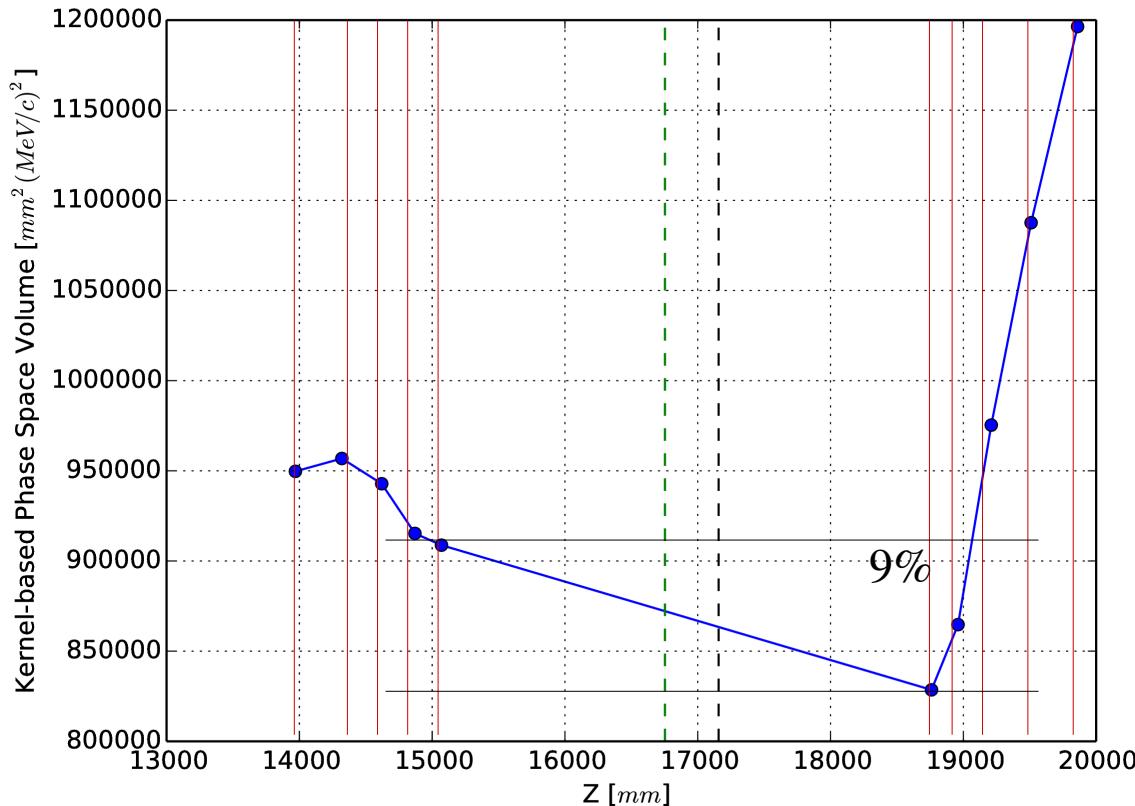


# Preliminary Application of KDE to MICE Data

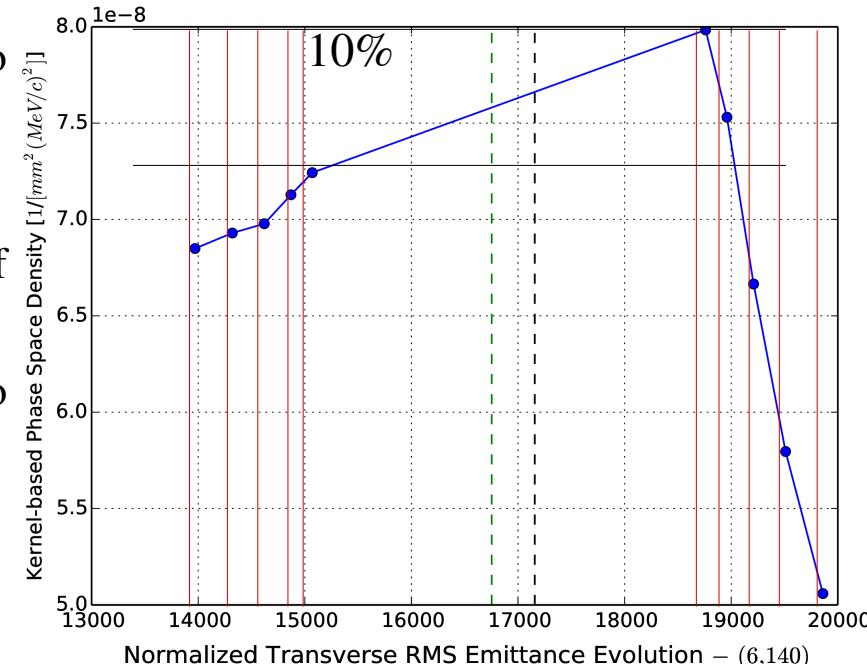
Run 8744:

- ▶ 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.  $32 < \text{TOF12} < 39 \text{ ns}$ ,  $100 < p < 220 \text{ MeV}/c$  cuts.

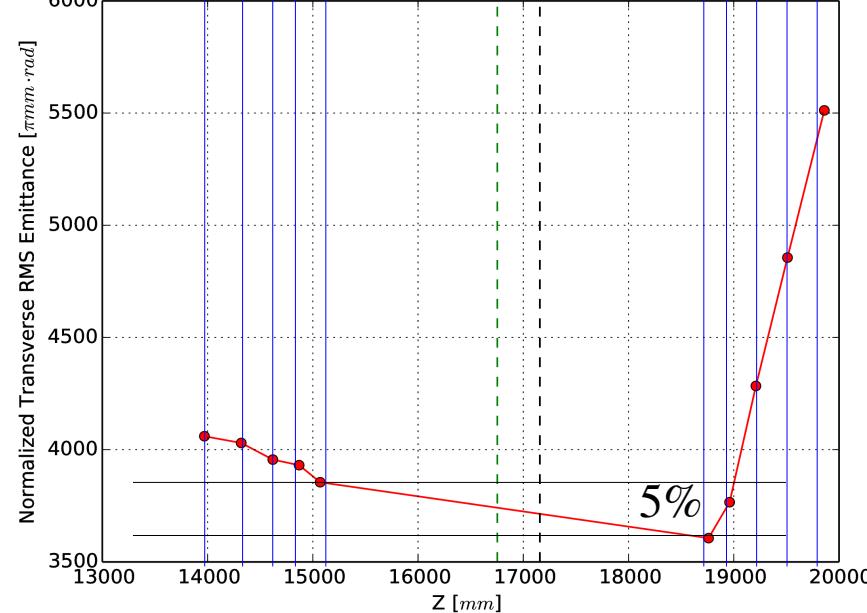
Kernel-based Volume Evolution – (6,140)



Kernel-based Density Evolution – (6,140)

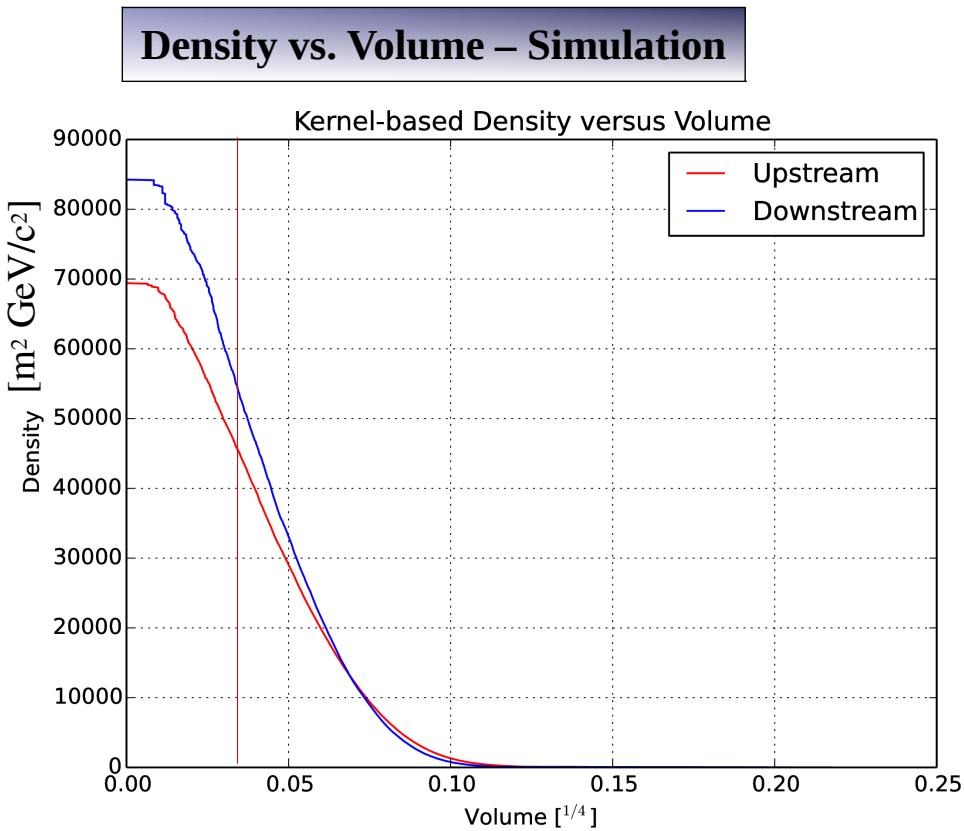
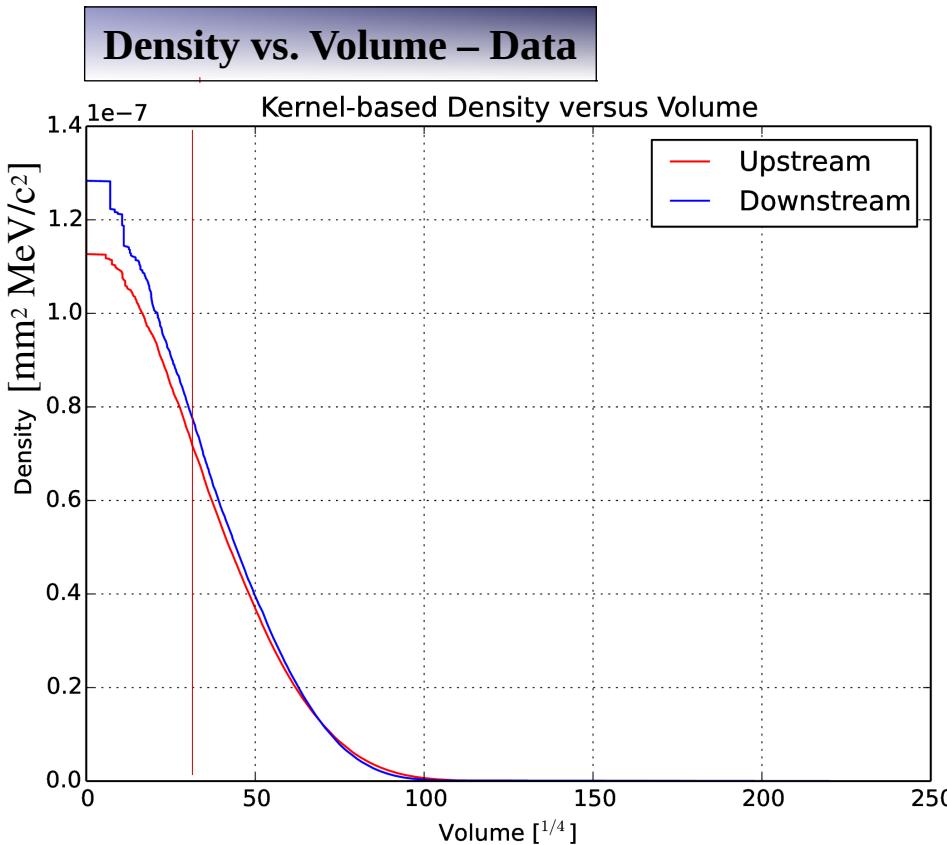


Normalized Transverse RMS Emittance Evolution – (6,140)



# Data vs. Simulation

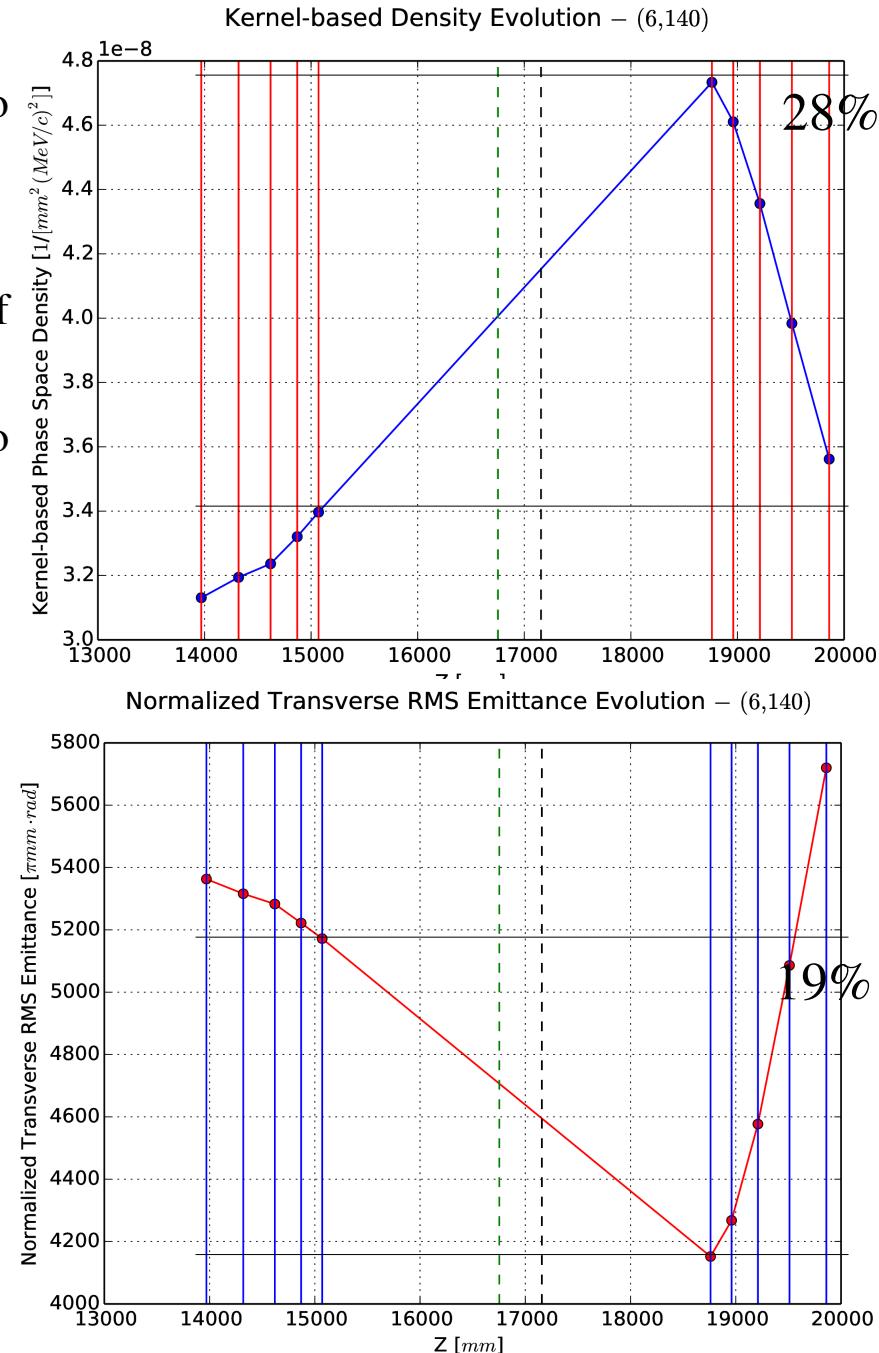
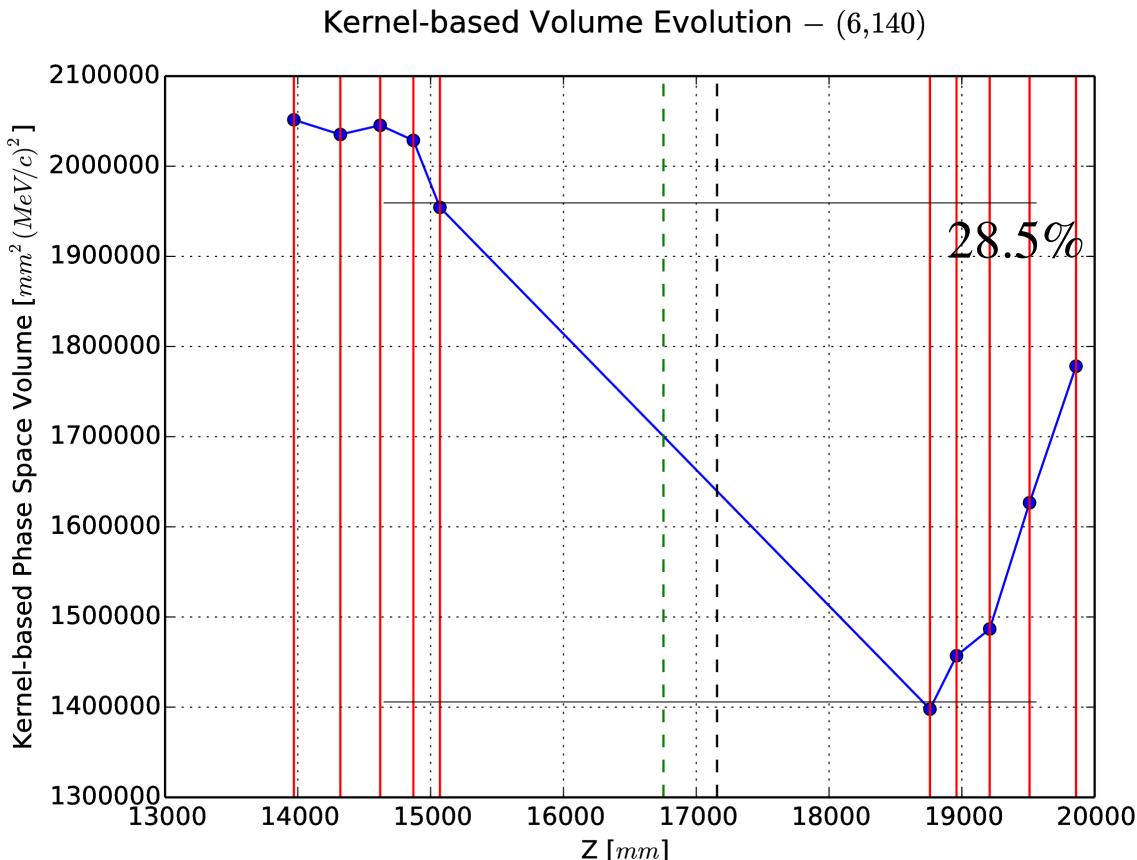
- Other phase-space contours:
  - ★ Volume $^{1/4}$ : the mean radius of hyper-ellipsoid.
  - ★ Beam center at volume $^{1/4} = 0$ , beam periphery at large volume $^{1/4}$  values.
  - ★ Increase in core density while decrease in density in the periphery.
  - ★ **Red vertical lines**: the locations of the one sigma contours (contour containing 9% of  $n$ ).



# Preliminary Application of KDE to MICE Data cont.

## Run 8685:

- ▶ 10-140 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.  $32 < \text{TOF12} < 39 \text{ ns}$ ,  $100 < p < 220 \text{ MeV}/c$  cuts.



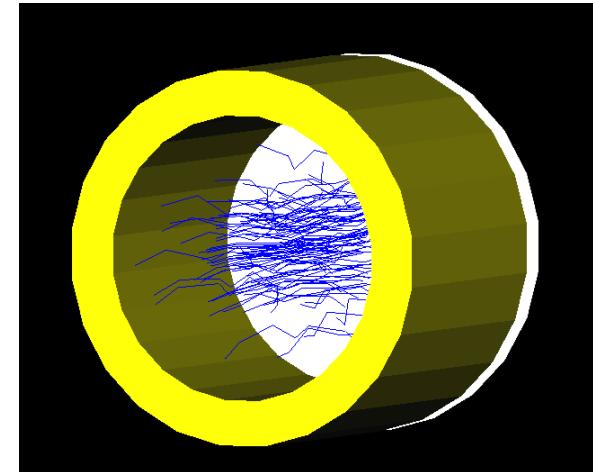
# Conclusion and Future Prospects

- Better beam cooling with KDE compared with the RMS emittance.
- VarKDE and NNDE perform well with a Gaussian distribution:
  - ★ Application to simulation, data, and long-tailed distributions (e.g. log Gaussian) in progress.
- NNDE method possibly useful for low statistics data.
- Improvements to the bandwidth parameter routine in KDE (e.g. cross validation) in progress.
- KDE-based beam weighting/sampling in progress (in addition to KDE-based beam cooling measurements).
- Stay tuned!

# Additional Slides

# Bandwidth's Effect on Density

- Motivation:
  - ✓ Fix the sample size (10,000) and study the effects of varying the bandwidth factor on density.
  - ✓ Study the sensitivity of KDE to long tails.
- Approach:
  - ✓ Generated a G4beamline Gaussian muon beam and passed it through a solenoid.
  - ✓ The regions in phase space to which KDE is applied: the entrance (upstream) and the center (downstream) of the solenoid.

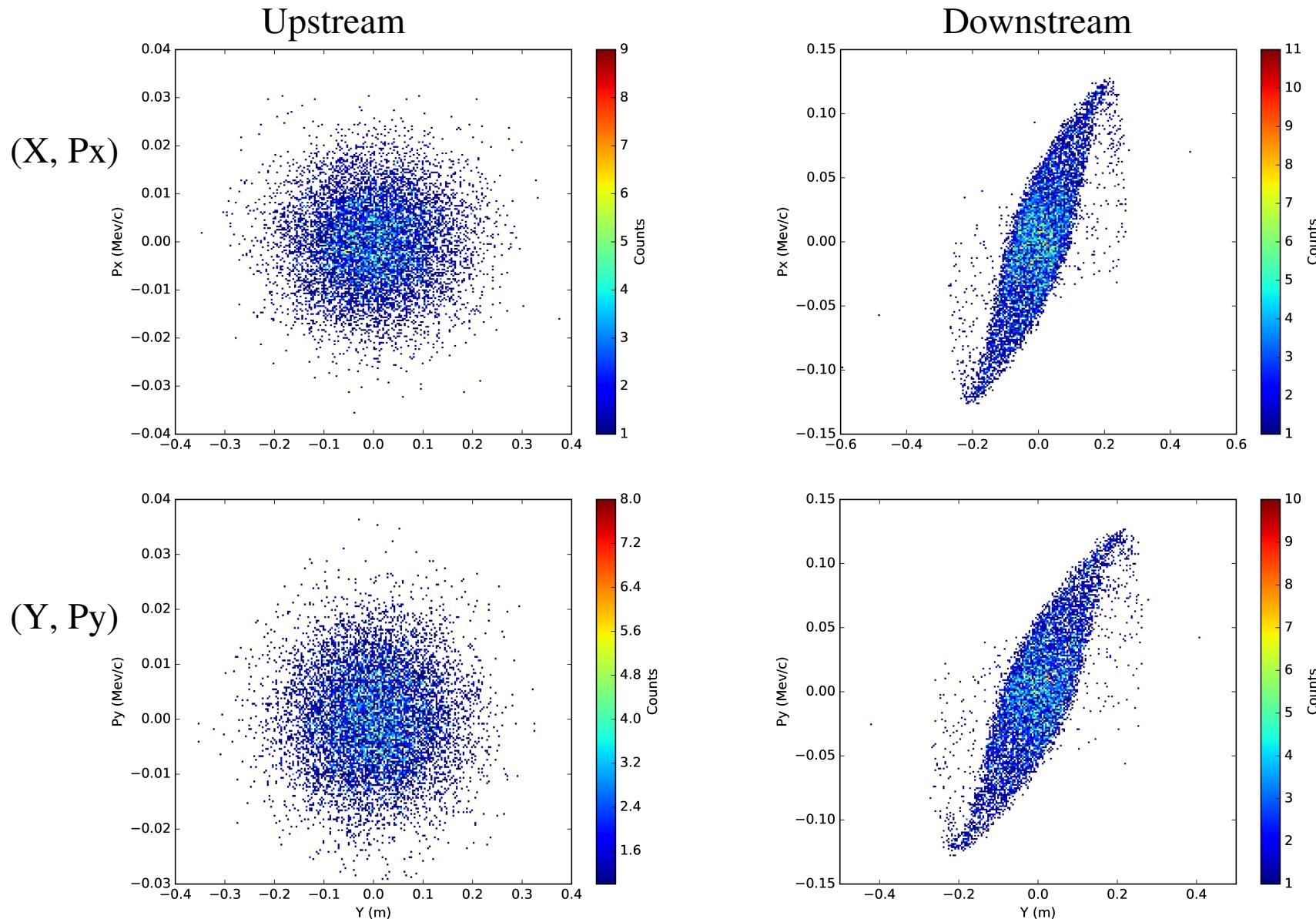


Simulation Parameters	Values
Number of events	10,000
Coil current	227.9 A/mm <sup>2</sup>
Solenoid Length	500 mm
$\sigma_x$	90 mm
$\sigma_y$	90 mm
$\sigma_{px}$	0.045 MeV/c
$\sigma_{py}$	0.045 MeV/c
Reference Momentum	200 MeV/c

Side Note: **Intentionally** produced the non-linear tails via an increase in beam size to study KDE of long-tailed distributions and the effects of bandwidth on the density in extreme conditions.

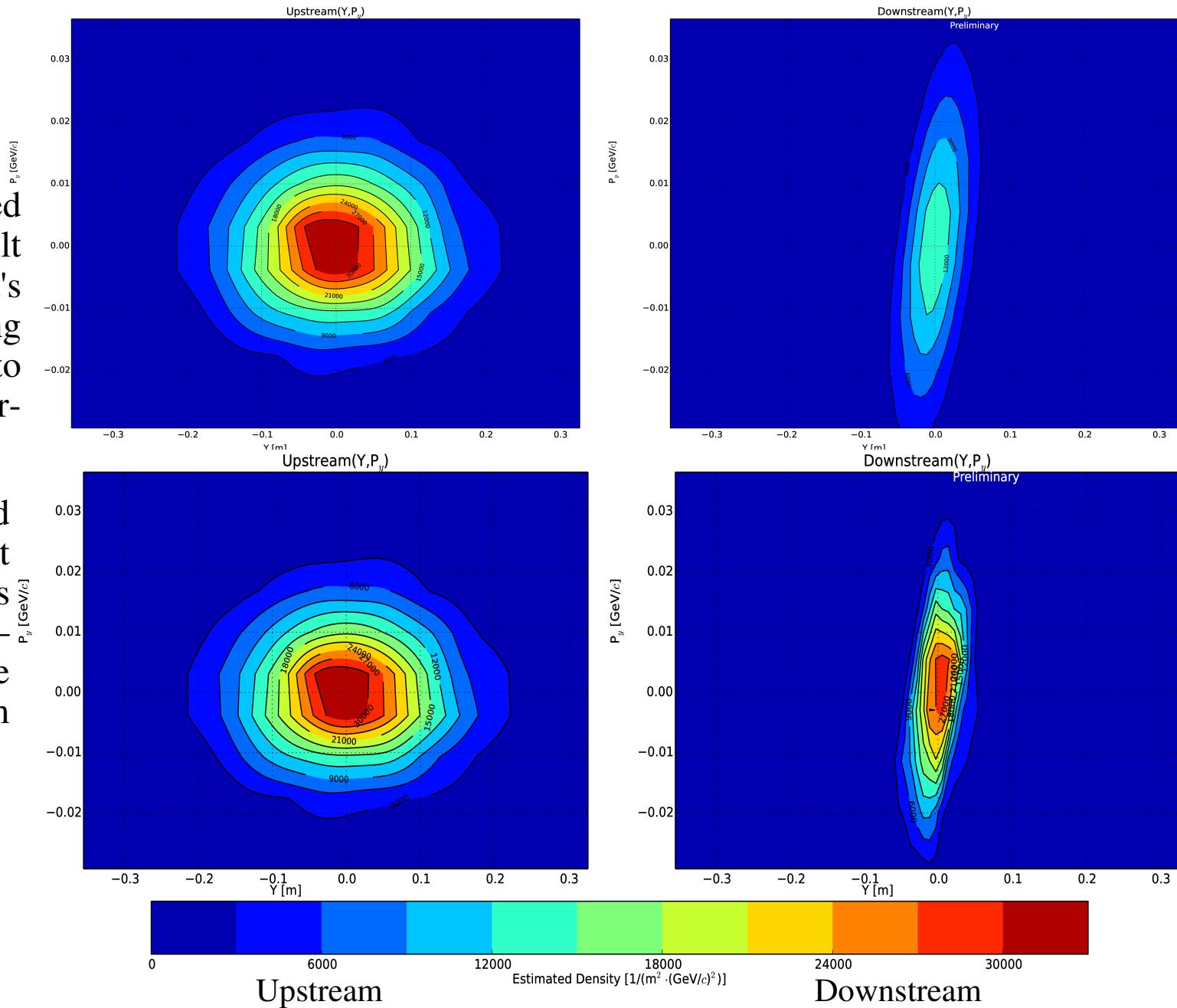
# Bandwidth's Effect on Density Cont. – Phase Space

- Clear display of the long-tails and the non-linear effects in the distribution, downstream.



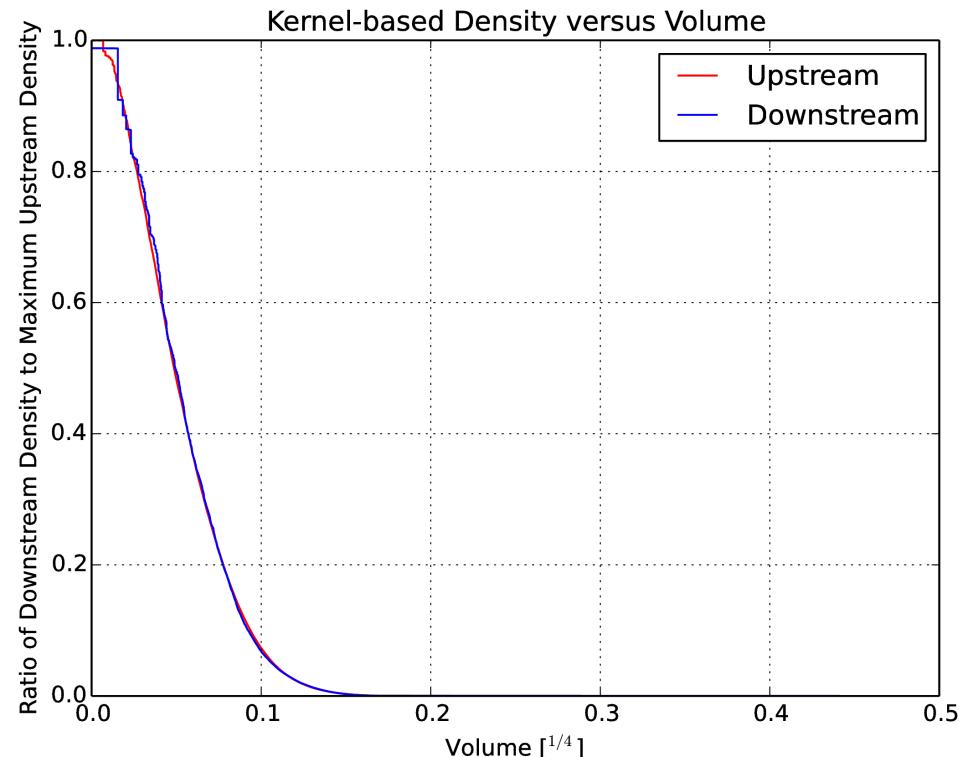
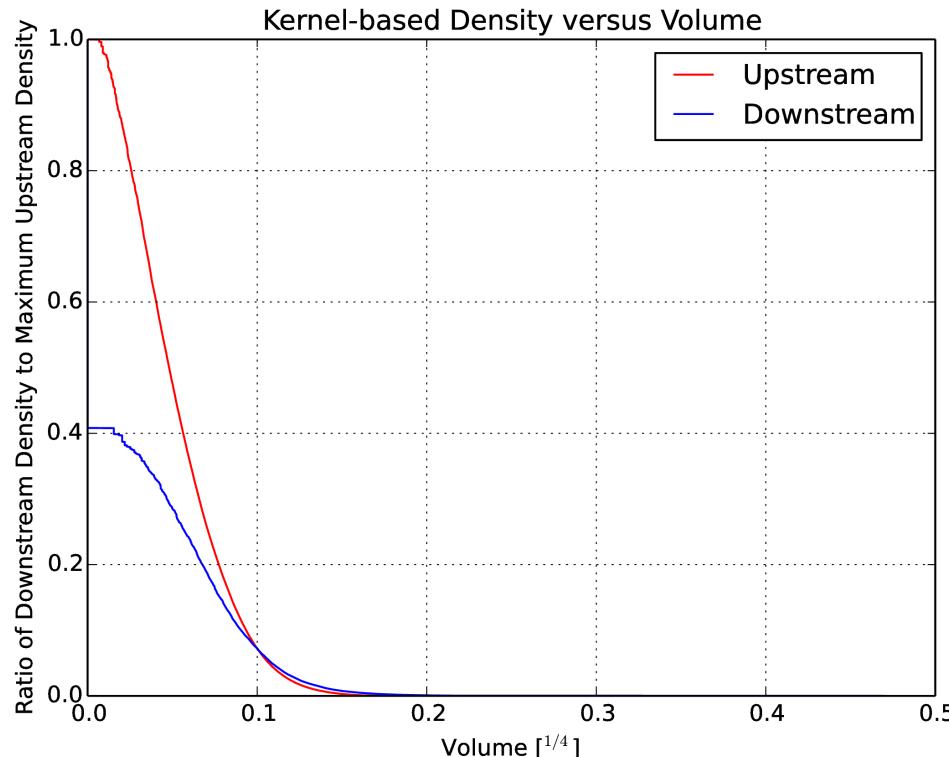
# Bandwidth's Effect on Density – Contours

- Top: density obtained with the default bandwidth factor, Scott's – demonstrates heating from upstream to downstream and is over-smoothed.
- Bottom: density obtained with the default bandwidth factor, Scott's multiplied by 0.42 – better represents the tailed phase space from last slide.



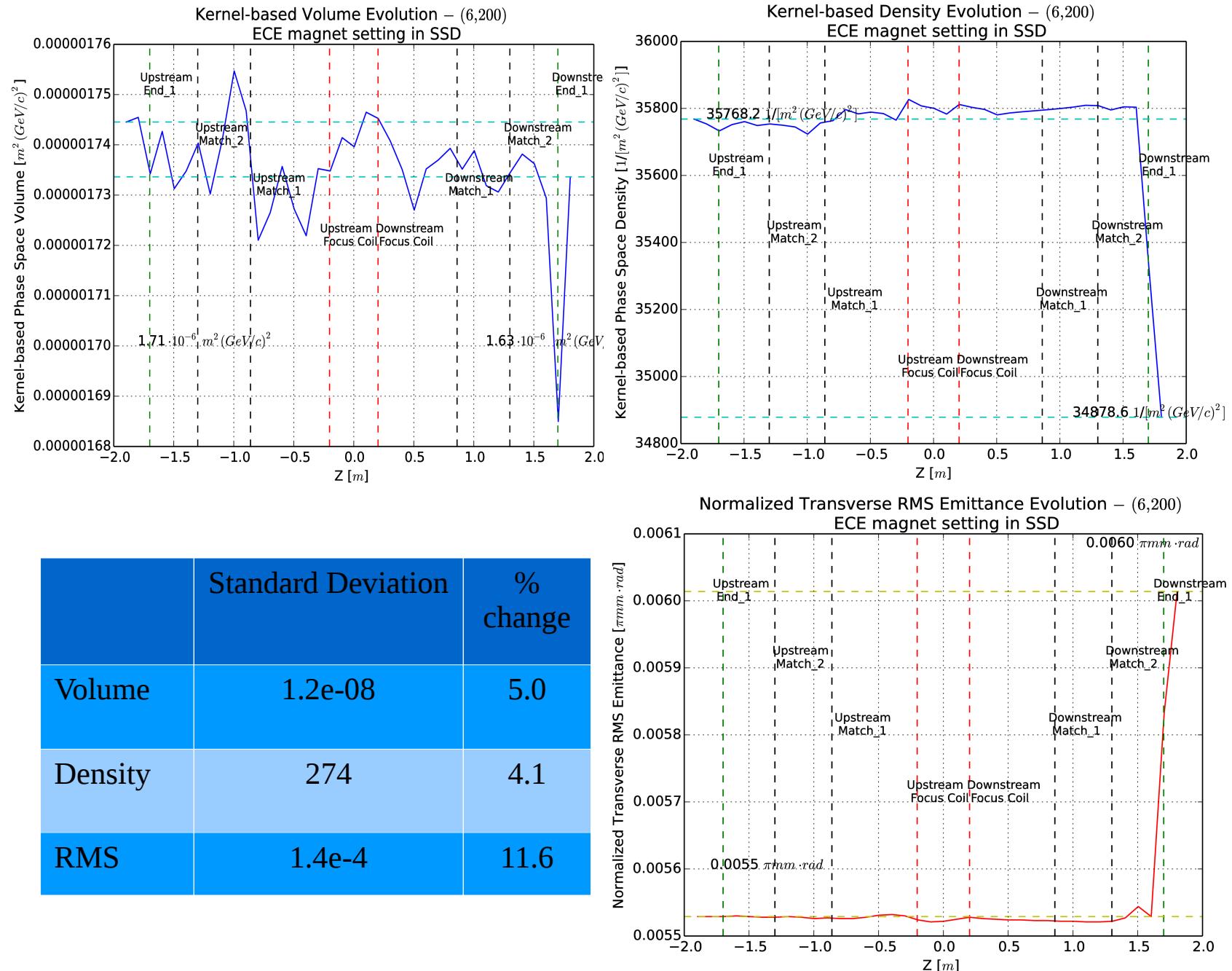
# Bandwidth's Effect on Density Cont. – Quantitative Comparison

- Left: ratio of downstream density to maximum upstream density vs. radius of four dimensional phase space with the larger bandwidth factor, Scott's factor.
- Right: ratio of downstream density to maximum upstream density vs. radius of four dimensional phase space with 0.423 times the Scott's factor. This bandwidth factor is a better representation of the conservation of density inside the solenoid.
- Results:
  - ✓ Choice of bandwidth factor has strong effects on the data interpretation.
  - ✓ KDE is sensitive to tails.



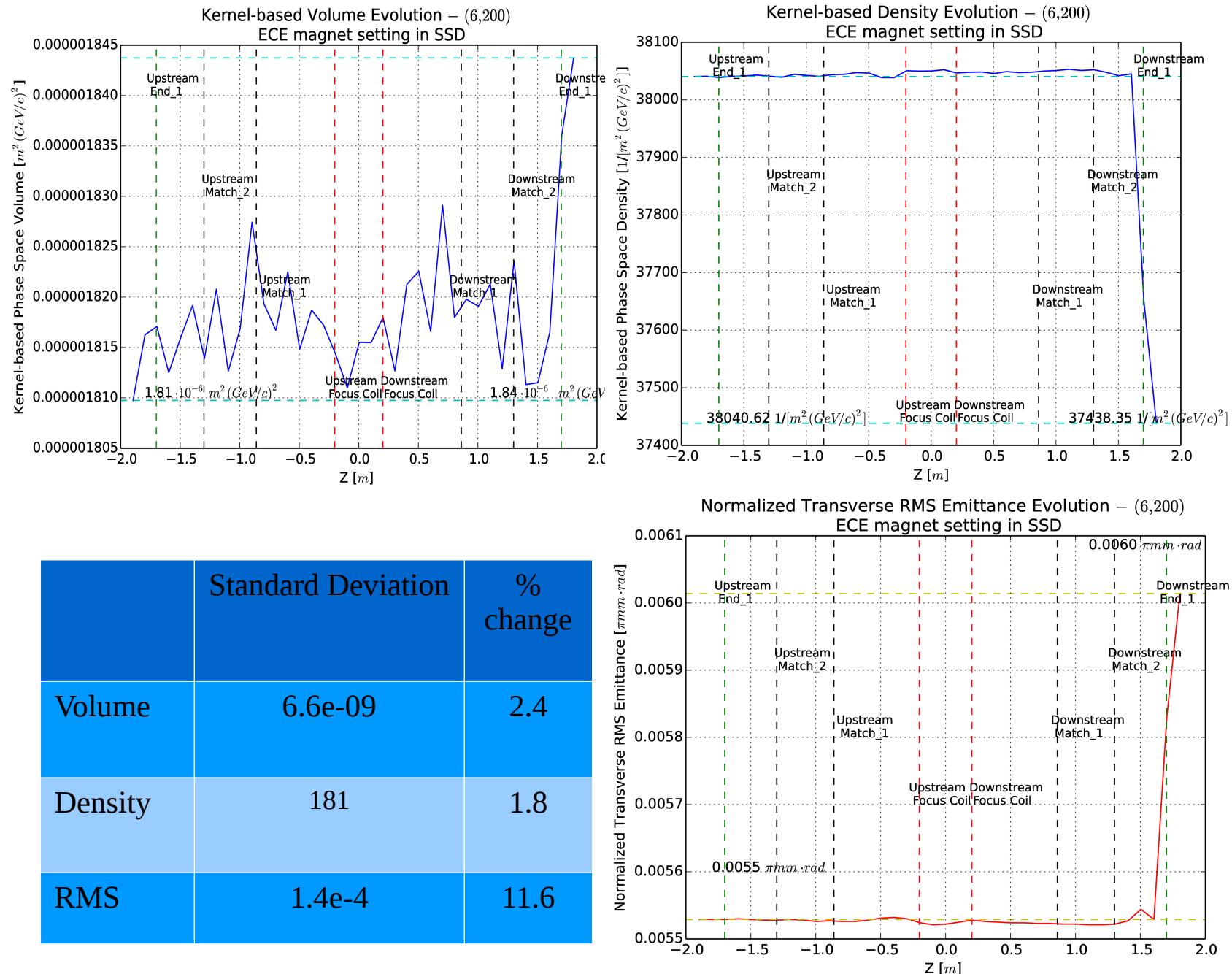
# MICE Optics with KDE

- Evolutions of emittance, beam-core volume and density along an empty-absorber channel:
  - ✓ Downstream match coils turned off.
  - ✓ Initial beam optics: (6, 200).
  - ✓ Trackers removed to ensure minimized energy loss effects.
  - ✓ 10,000 input with transmission of ~95%.
  - Result: spikes start to show up at  $z = 1.5$  m (close to the location of the downstream match coils).



# MICE Optics with KDE Cont. – Larger Sample Size

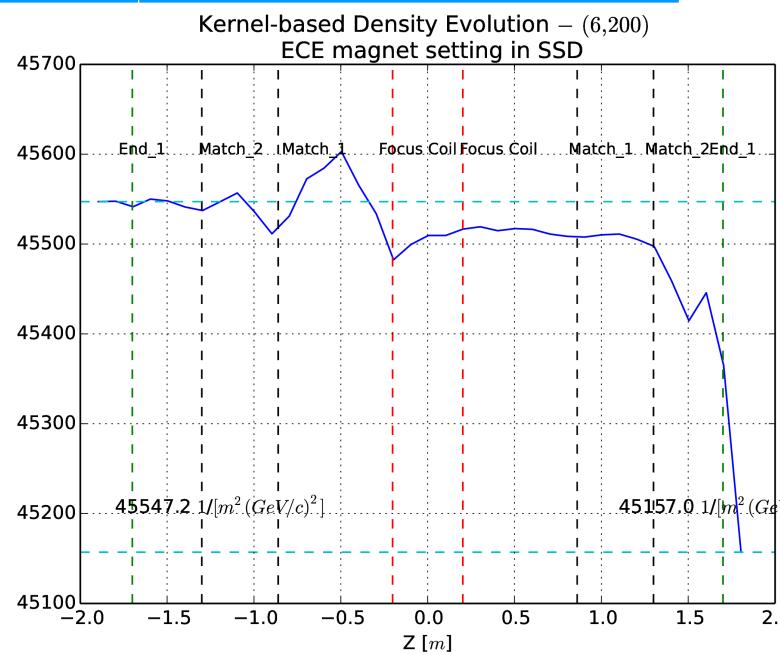
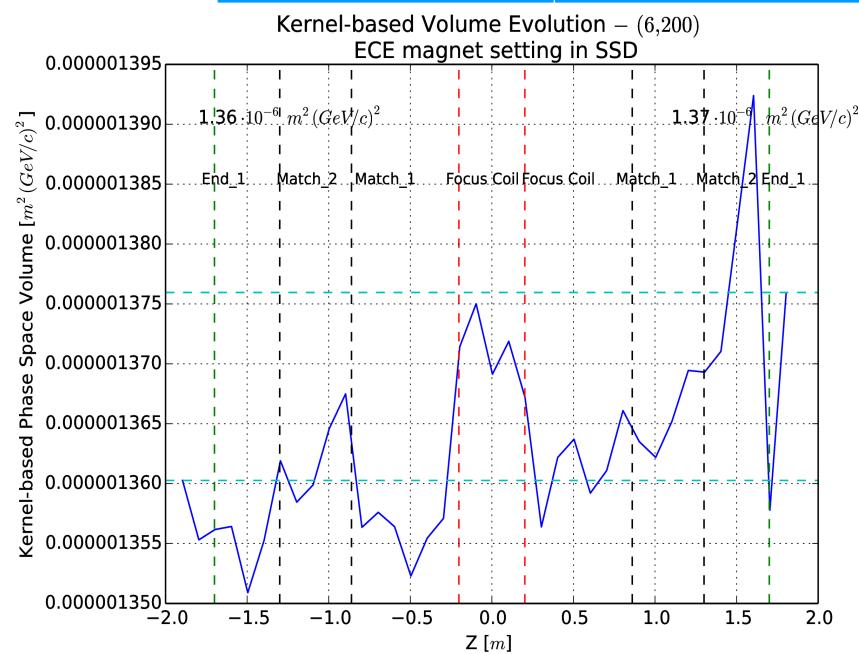
- Evolutions of emittance, beam-core volume and density along an empty-absorber channel:
  - ✓ Downstream match coils turned off.
  - ✓ Initial beam optics: (6, 200).
  - ✓ Trackers removed to ensure minimized energy loss effects.
  - ✓ 100,000 input with transmission of ~95%.
  - Results: shown on the table.



# MICE Optics with KDE Cont. – $5\sigma$ (High Amplitude) Cut

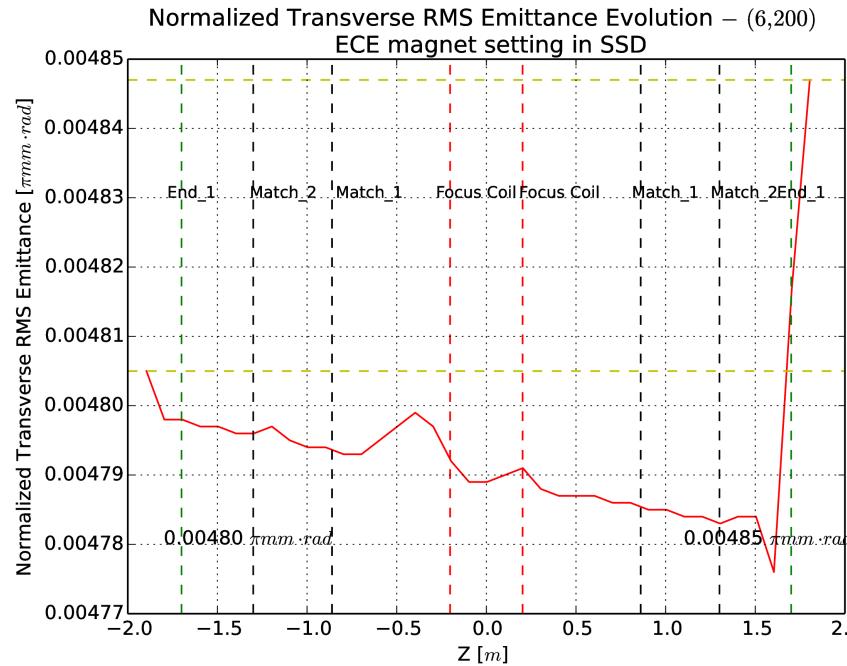
- Used KDE to apply a high amplitude cut,
- $5\sigma$  of the beam roughly contain 85% of the total muons in four dimensions (S.Y. Lee) → apply KDE on the distribution, do a binary search to isolate contour with 85% of all muons, and record  $(x, p_x, y, p_y)$  of the contained muons.
- After the cut, redo the “old routine” (described in the previous slide) on this subset of muons. Emittance same as slide 2 (did not apply this to emittance). Results are below,

	Standard Deviation	% change
Volume	1.0e-08	5.9
Density	108.6	1.3
RMS	0.00014	11.56

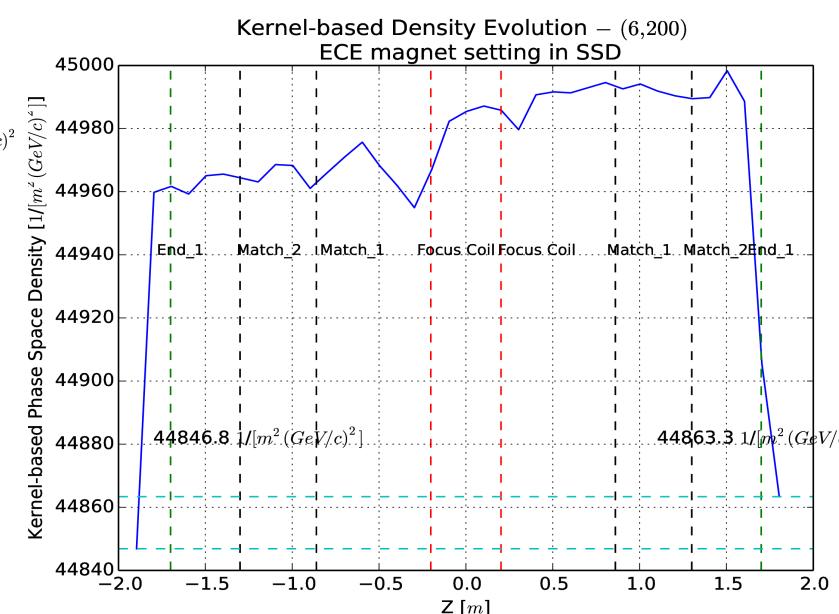
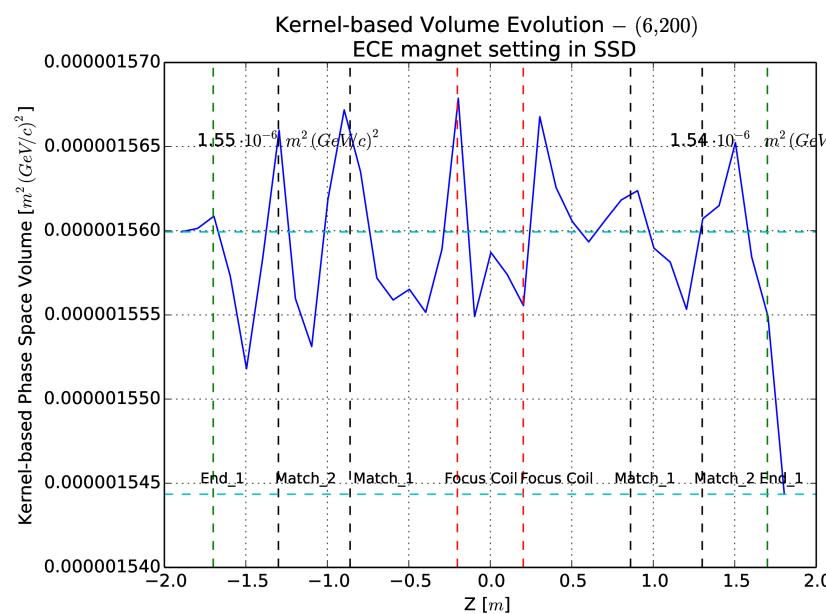


- Evolutions of emittance, beam-core volume and density along an empty-absorber channel:
- Downstream match coils turned off.
- Initial beam optics: (6, 200). **But with a  $5\sigma$  cut.**
- Trackers removed to ensure minimized energy loss effects.
- 10,000 input with transmission of ~95%.
- Results: shown on the table.

# MICE Optics with KDE Cont. – Cooling Channel Cont.



	Standard Deviation	% change
Volume	7.0e-09	2.8
Density	31.6	0.33
RMS	1.8e-05	2.28

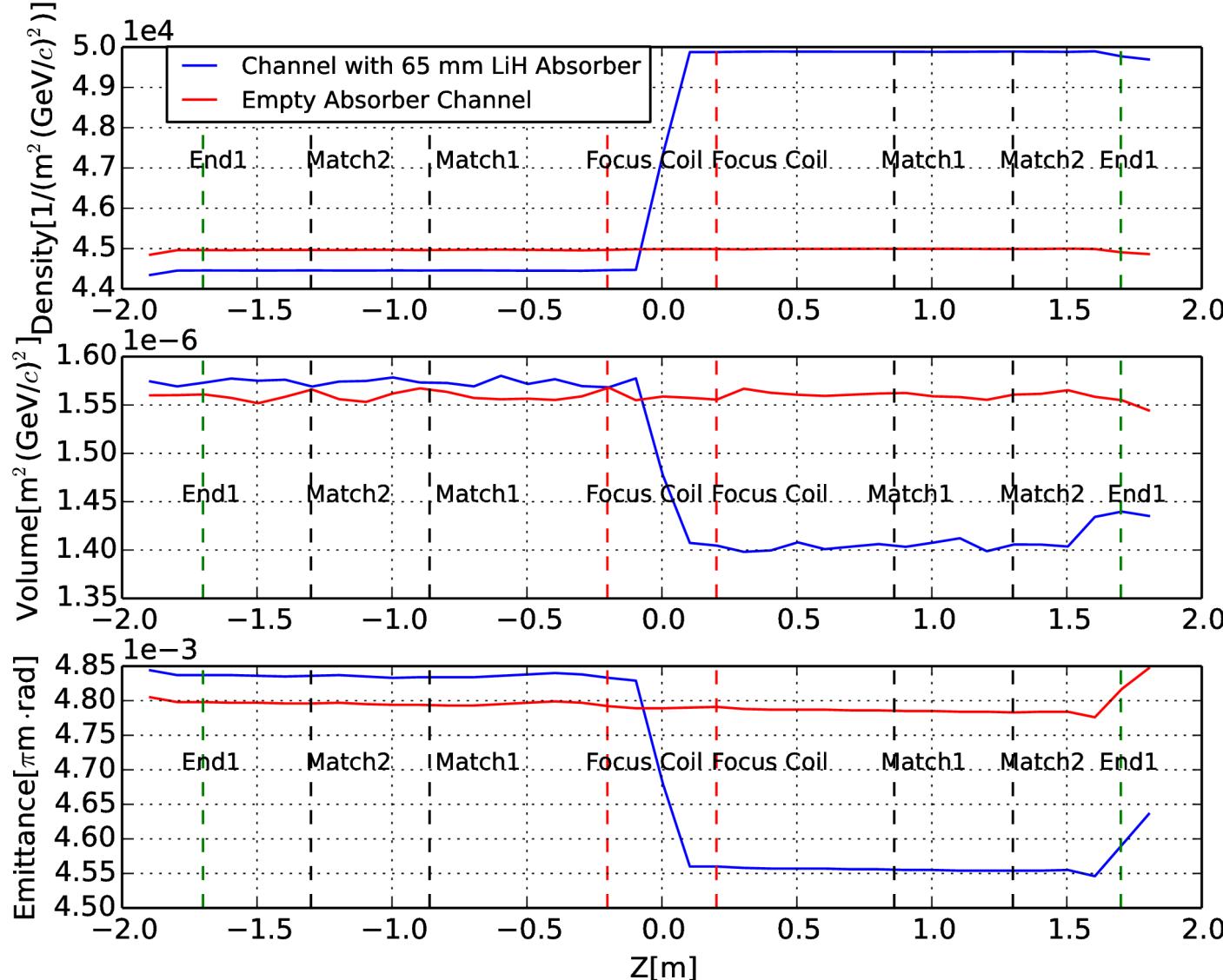


The evolutions of emittance along with beam-core volume and density in the empty-absorber lattice, zoomed in.

Results: shown on the table.

# MICE Optics with KDE Cont. – Cooling Channel

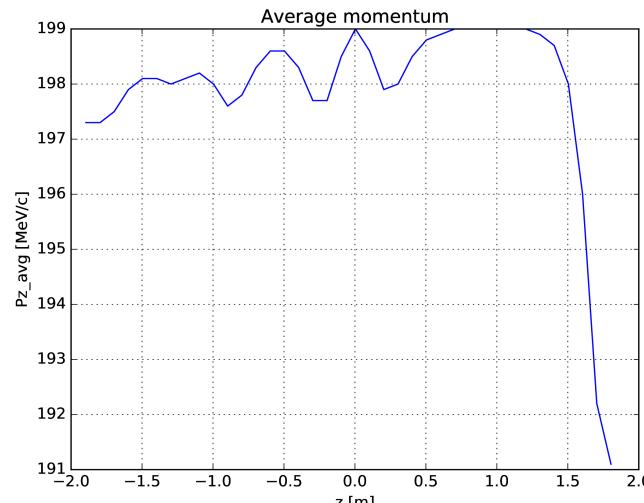
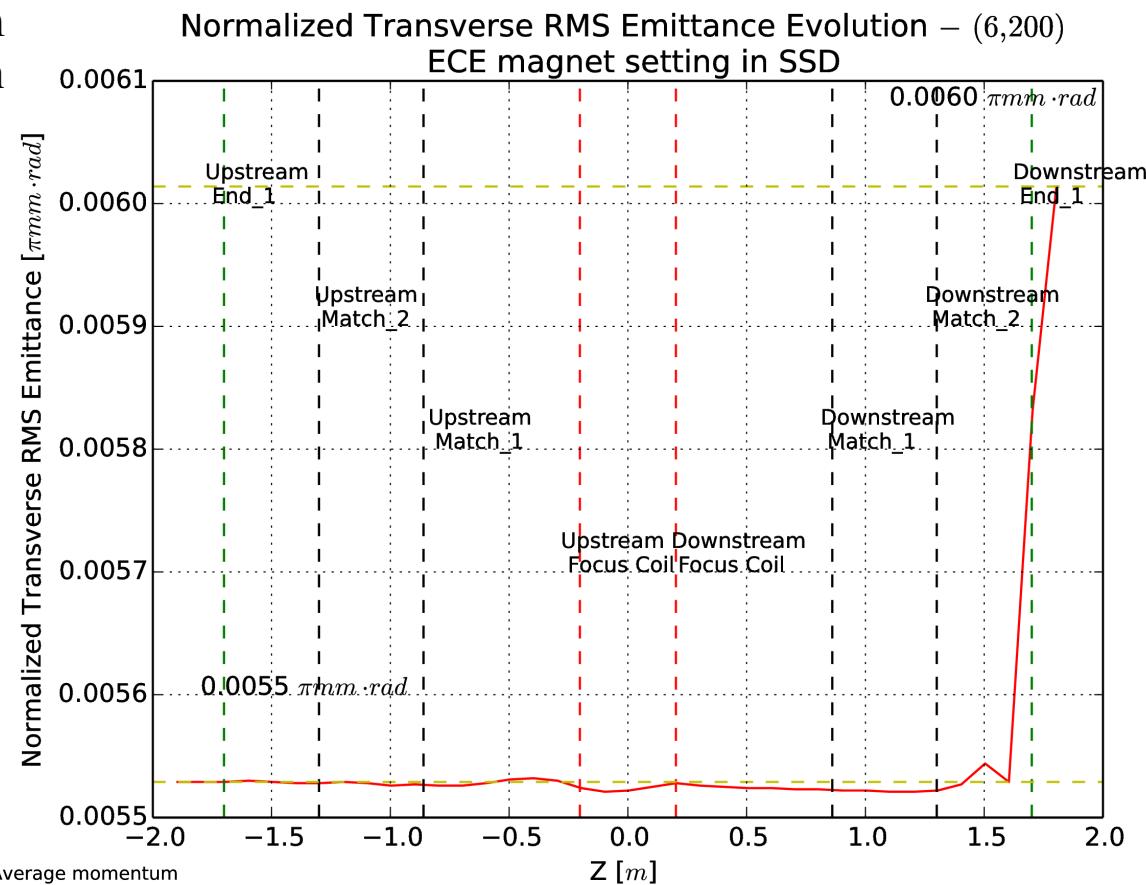
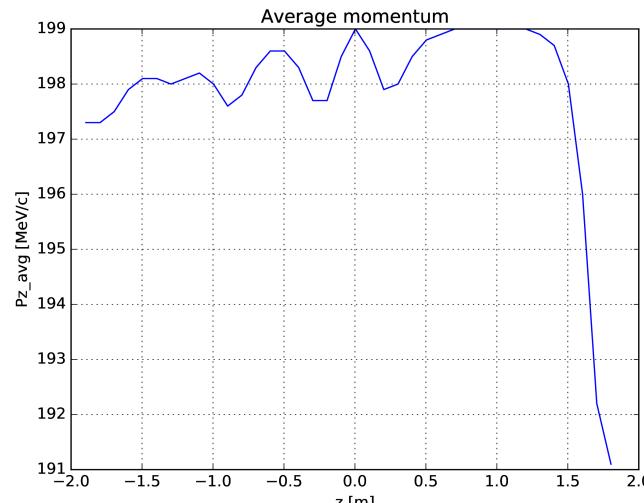
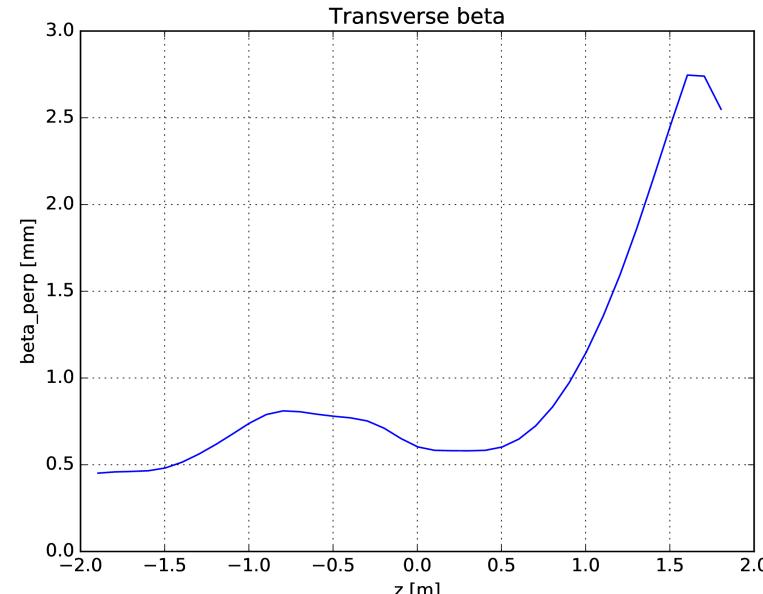
- Spikes still present in empty absorber channel but in the scale of cooling effect, they have smaller percentages.



- Evolutions of emittance, beam-core volume and density along the empty-absorber **and the cooling channels**:
  - Downstream match coils turned off.
  - Initial beam optics: (6, 200).
  - Trackers **placed back in the geometry** to ensure realistic energy loss processes.
  - 100,000** input muons with a transmission of  $\sim 85\%$ .
- Results: 4% and 9% reductions in emittance and volume with 12% increase in density.

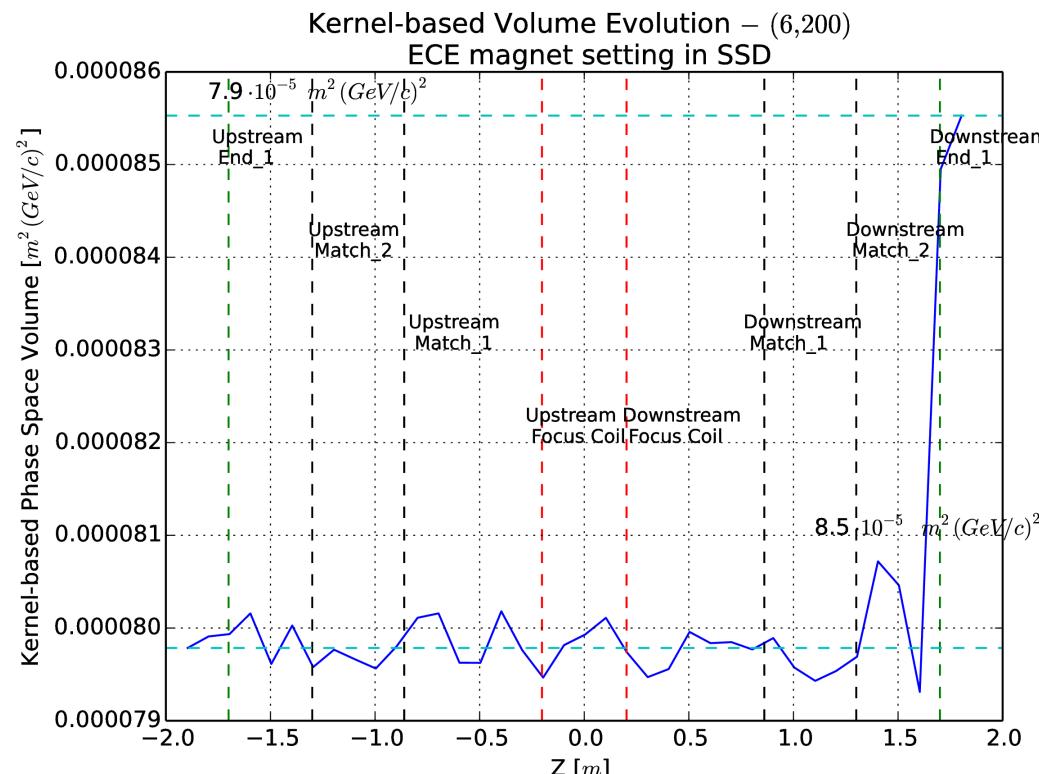
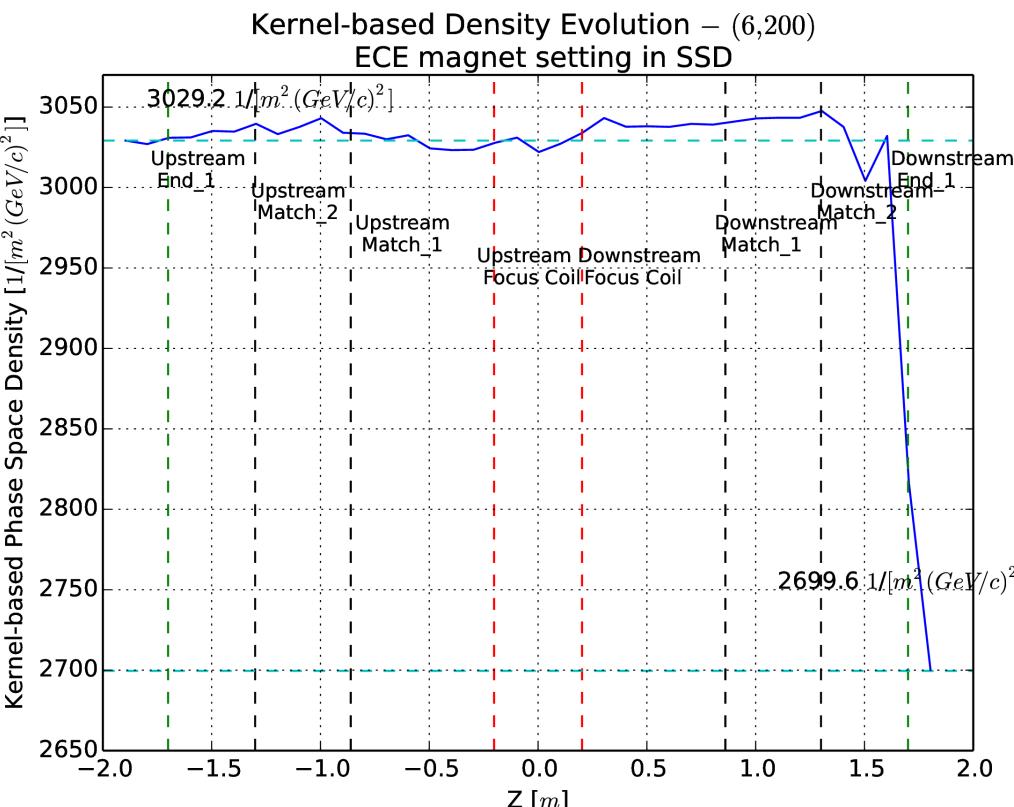
# Twiss Parameters + Emittance Evolution

- Substantial beam size growth after beam passes through the disabled match coils in SSD.
- Emittance growth of ~9% from US to DS.

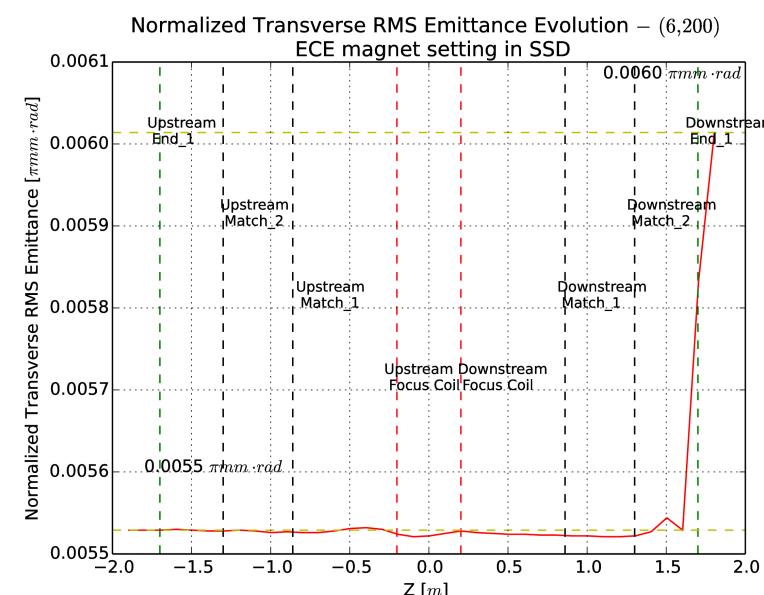


- Applied good muon cut (~500 muons do not make it to DS TRPs).
- Reminder: Only SSD in ECE setting and all solenoidal fields scaled to 3T.

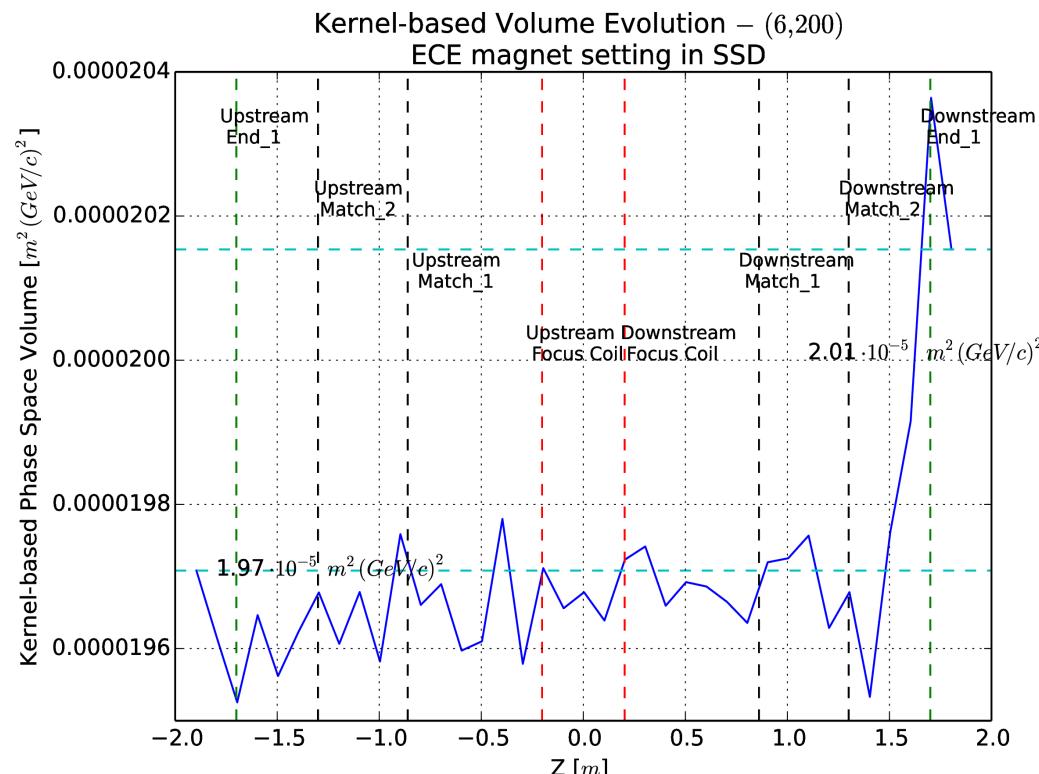
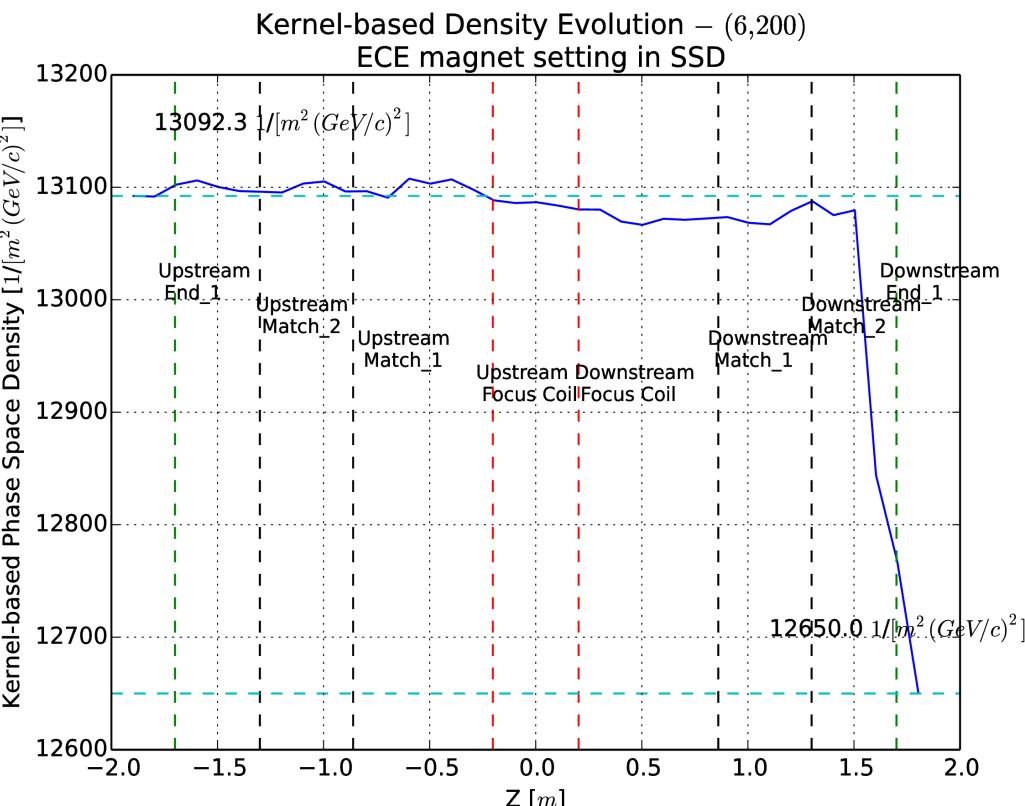
# Kernel-based Density and Volume - 90% Contour Evolution



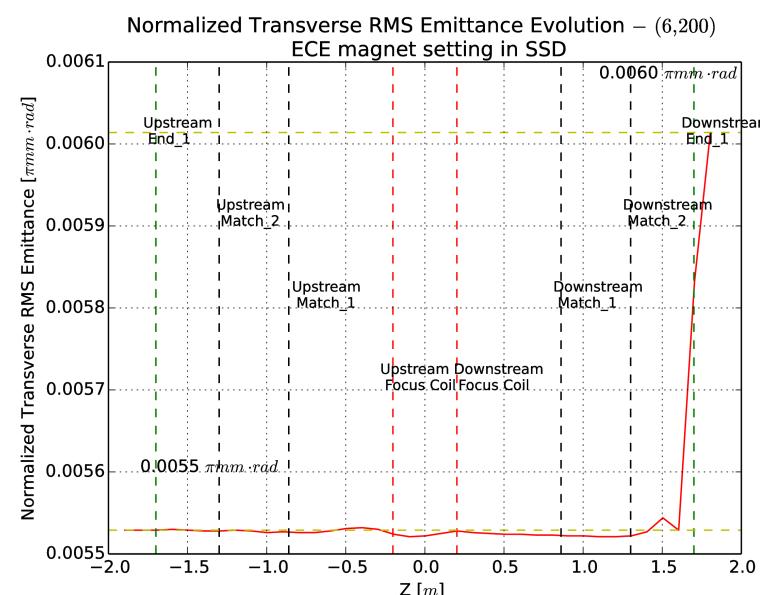
- Computed Kernel-based density and volume for the contour enclosing 90% of total muon.
- Density drops by ~12% and volume grows by ~7.5%.
- Measurement fluctuations are present. Expected for volume as its measurement relies on MC approach.
- Standard deviations in density and volume: 89 and 1.4e-06.



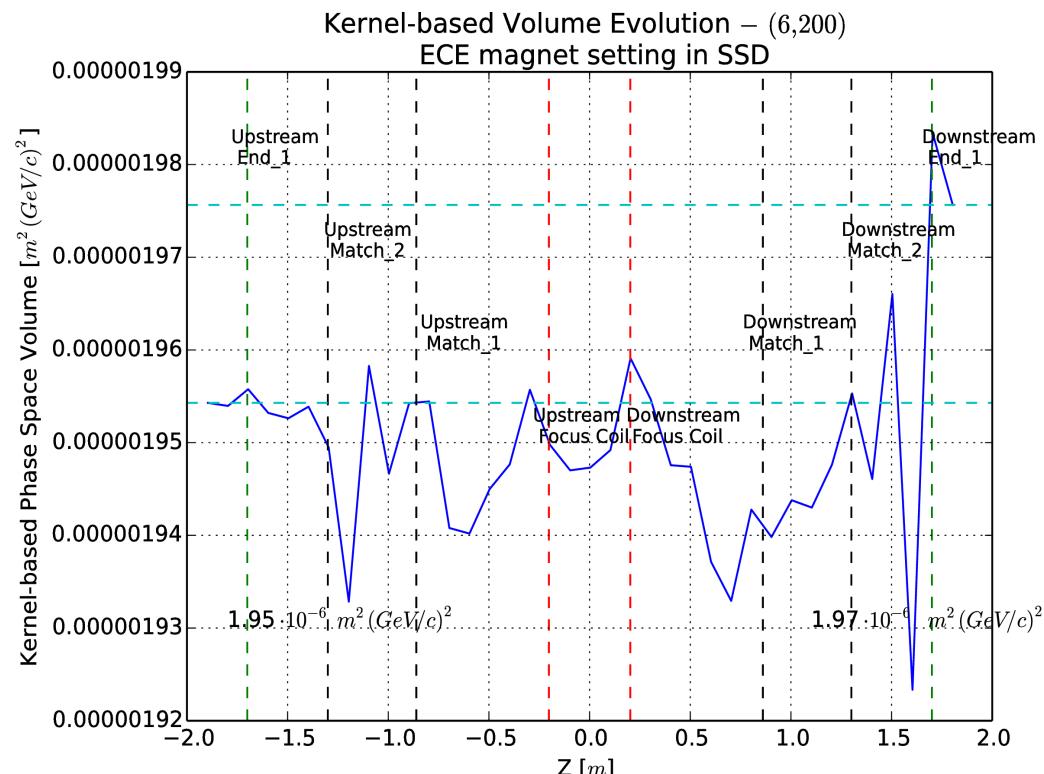
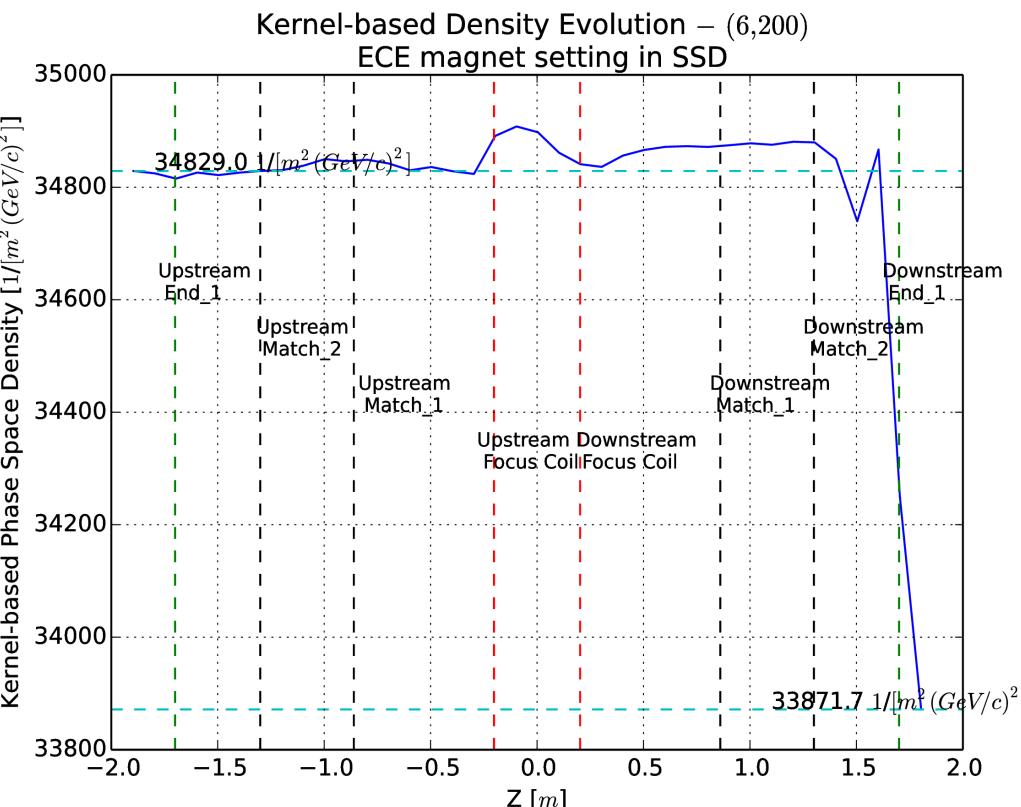
# Kernel-based Density and Volume - 50% Contour Evolution



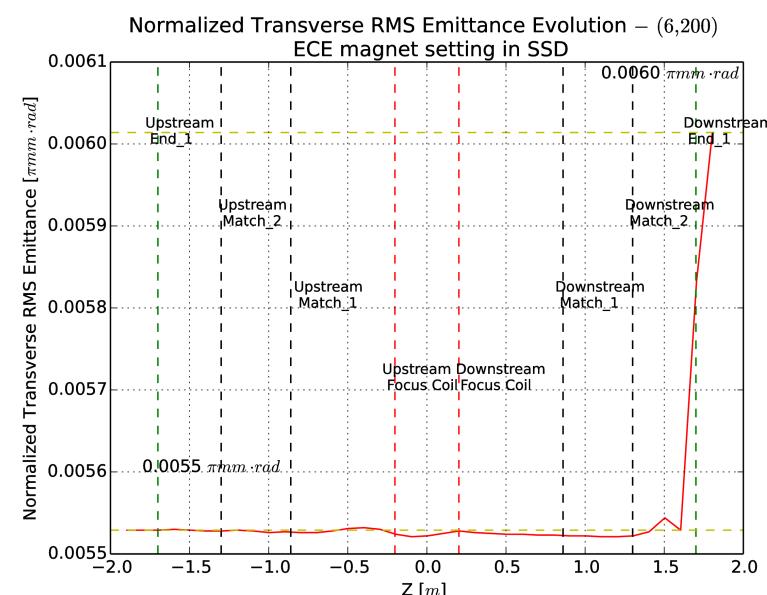
- Computed Kernel-based density and volume for the contour enclosing 50% of total muon.
- Density drops by ~3% and volume grows by ~2.0%.
- Measurement fluctuations are present. Expected for volume as its measurement relies on MC approach.
- Standard deviations in density and volume: 93.6 and 1.5e-07.



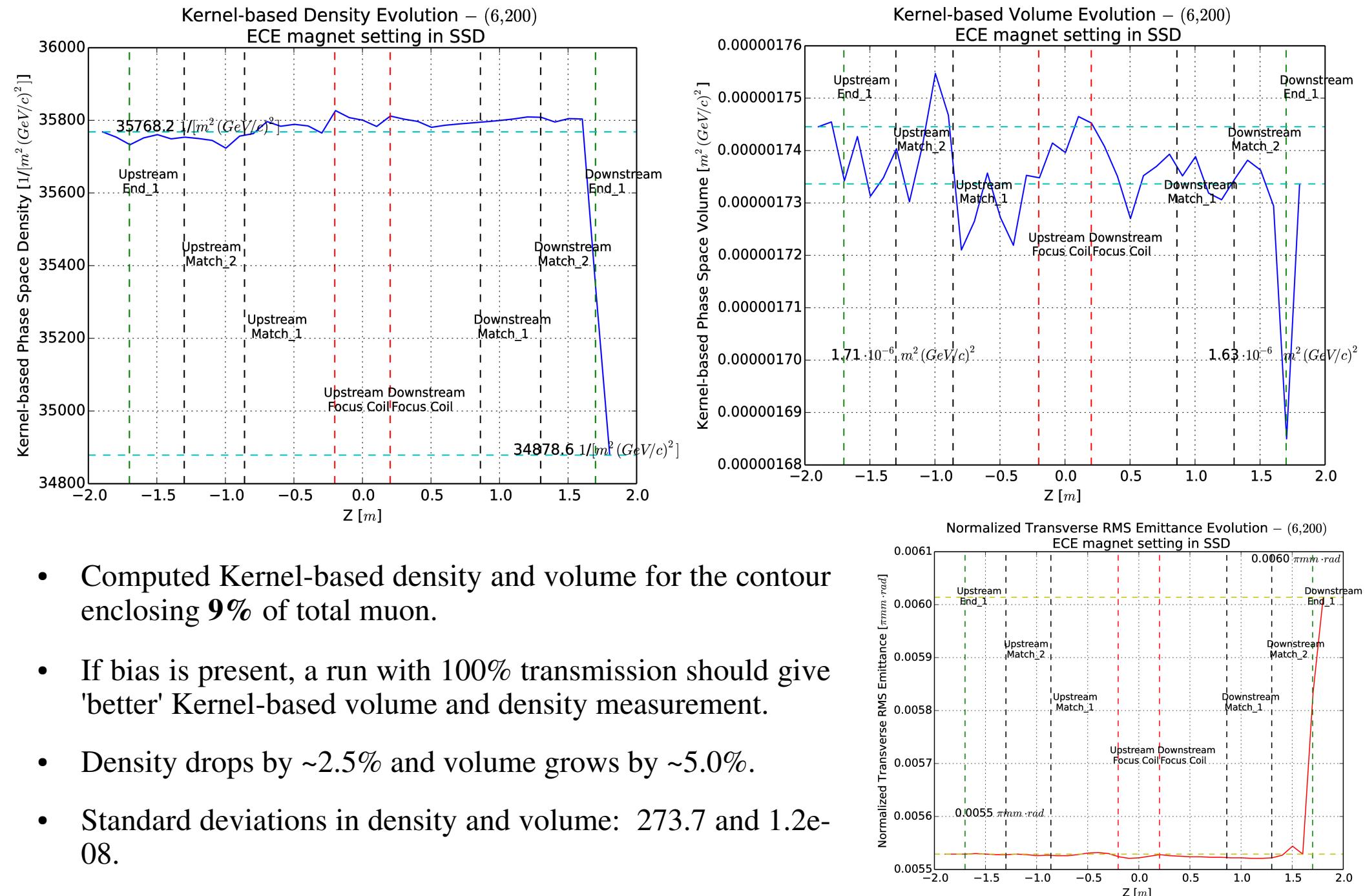
# Kernel-based Density and Volume - 10% Contour Evolution



- Computed Kernel-based density and volume for the contour enclosing 10% of total muon.
- Density drops by ~2.7% and volume grows by ~1.0%.
- Measurement fluctuations are present. Expected for volume as its measurement relies on MC approach.
- Standard deviations in density and volume: 295.2 and 1.31e-08.

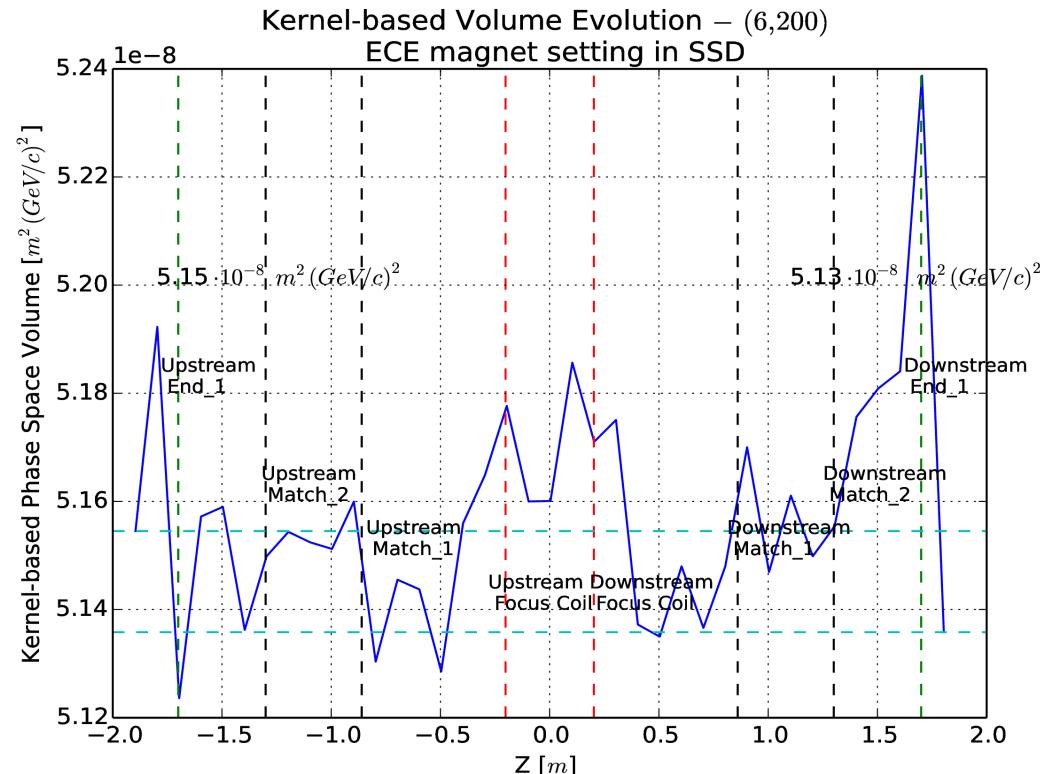
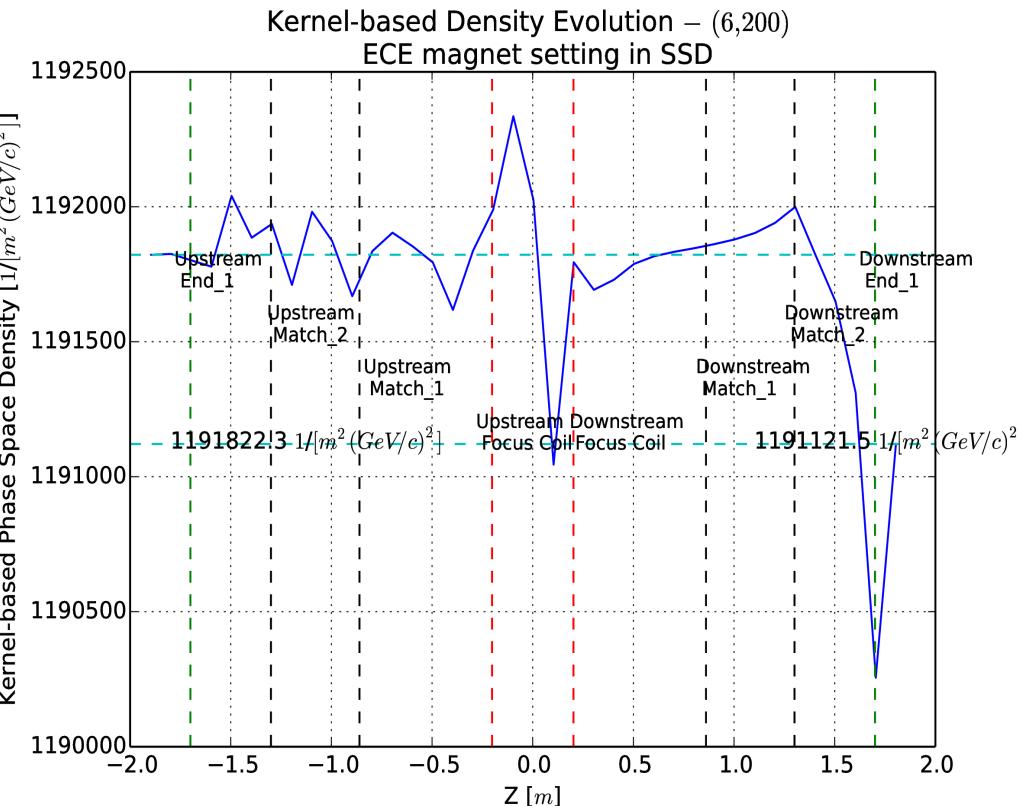


# Before - (6, 200) with SSD in ECE – ~95% Transmission

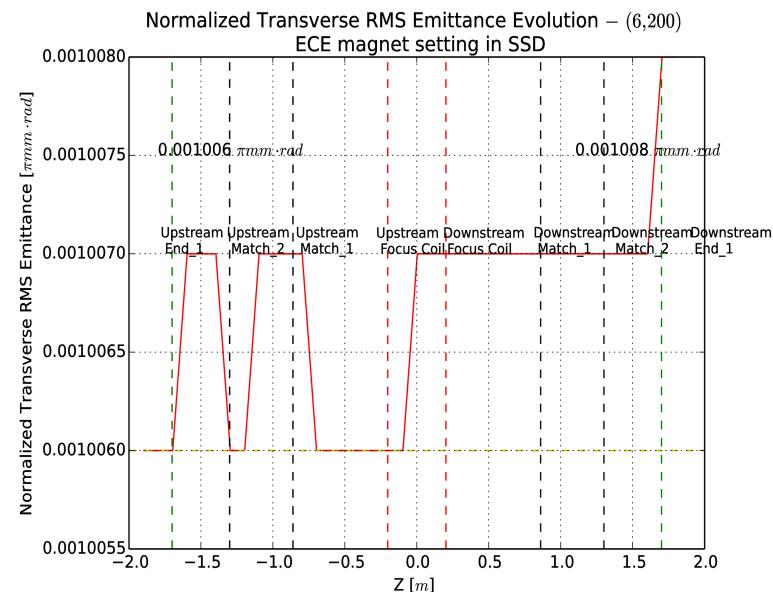


- Computed Kernel-based density and volume for the contour enclosing **9%** of total muon.
- If bias is present, a run with 100% transmission should give 'better' Kernel-based volume and density measurement.
- Density drops by  $\sim 2.5\%$  and volume grows by  $\sim 5.0\%$ .
- Standard deviations in density and volume: 273.7 and  $1.2 \times 10^{-8}$ .

# After - (1, 200) with SSD in ECE – 100% Transmission



- Computed Kernel-based density and volume for the contour enclosing 9% of total muon.
- Density drops by ~0.06% and volume grows by ~1.0%. Emittance growth is now improved as well ~0.1%
- Standard deviations in density and volume: 395.1 and  $2.4 \times 10^{-10}$ . 100% transmission run gives better KDE measurements. But fluctuations in measurement are worst.



# KDE Validation in 2D – Variance Test

- Generated 2D Gaussian distribution ( $\sigma_1 = 0.03$  a.u and  $\sigma_2 = 0.02$ ) of 1k, 10k, 100k sample sizes.
- Fixed the bandwidth factor ( $h_{\text{factor}} = 0.2$ ).
- Distribution gets less noisy (low variance) as sample size grows.

