

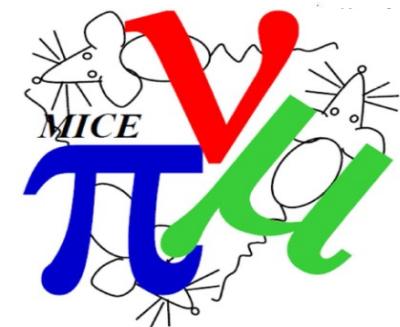
Novel Application of Density Estimation in MICE



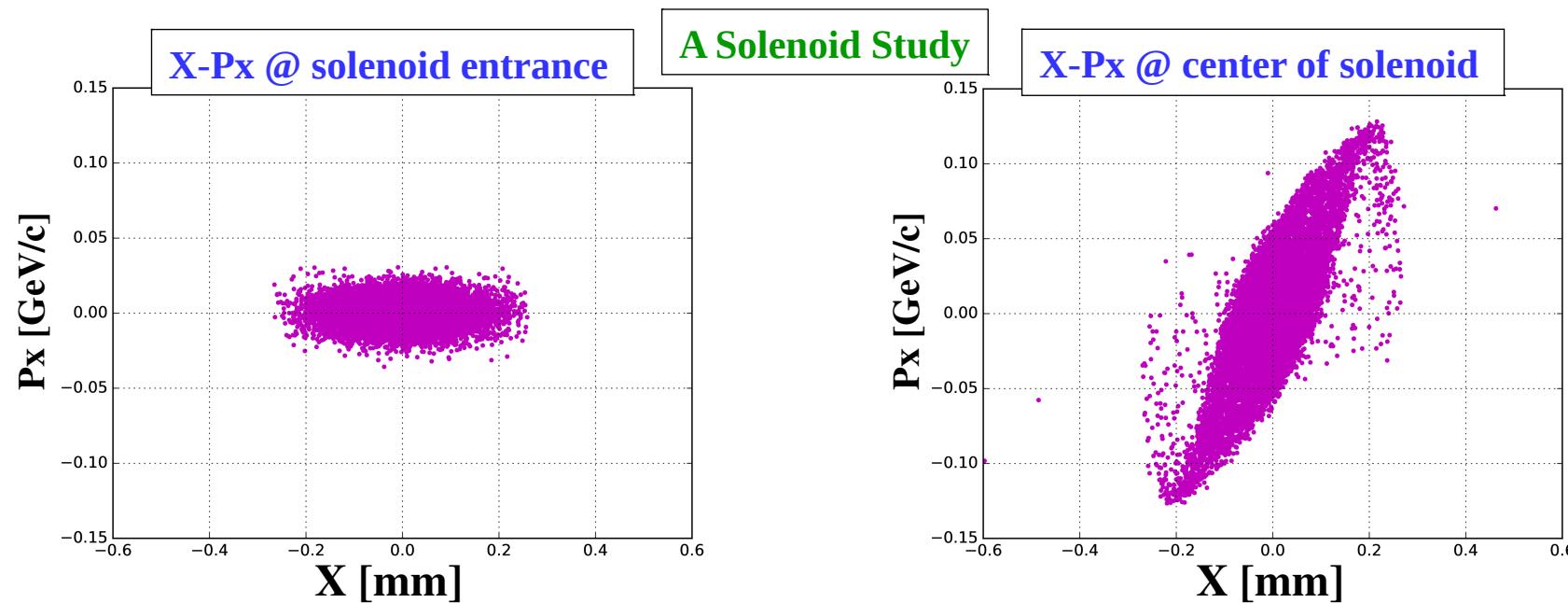
Tanaz A. Mohayai

CM47, RAL

February 13, 2017



Motivation

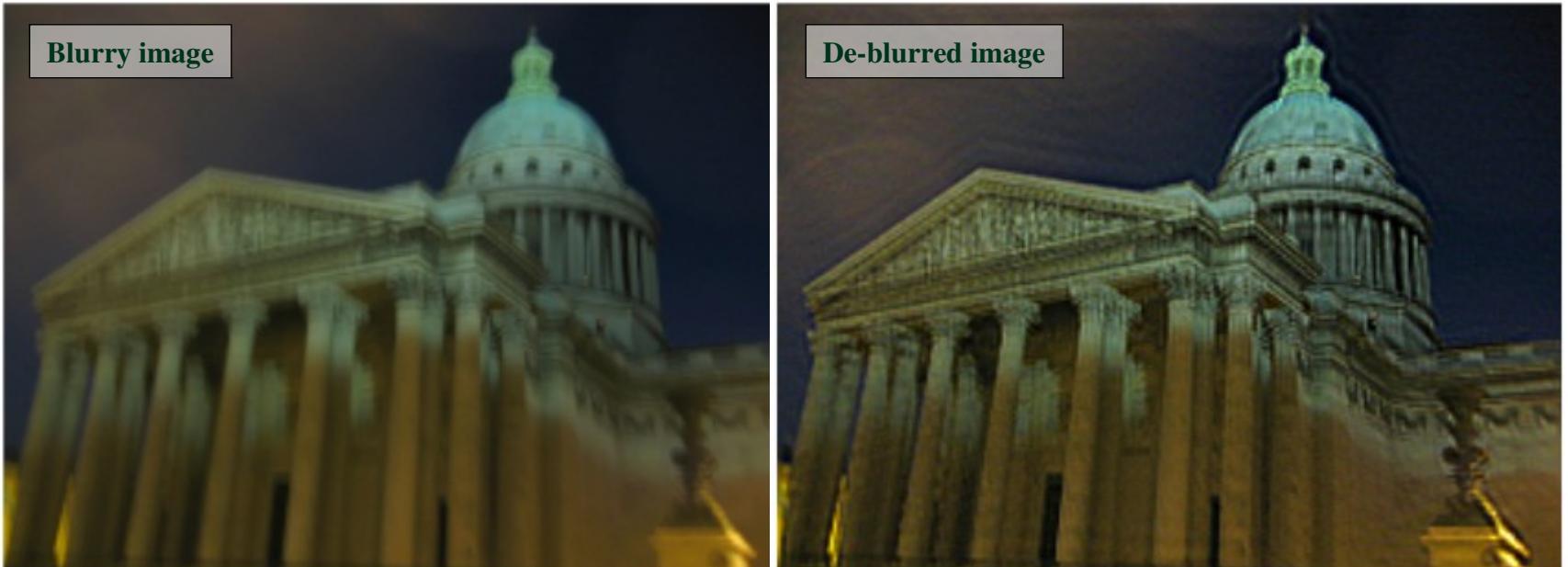


- **Problem:**
 - ★ RMS emittance assumes a Gaussian beam.
 - ★ Real-life beam is non-Gaussian (chromatic and non-linear effects).
- **Solution is Density Estimation:**
 - ★ Estimate PDF or density (normalized density) with few assumptions about the underlying distribution.
 - ★ Gives detailed single-particle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

Kernel Density Estimation (KDE)

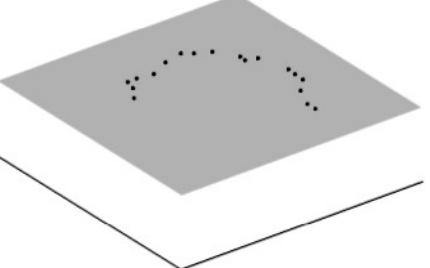
- Estimates the unknown probability density function (PDF) or density (normalized density) using kernels (smooth weight functions of certain widths).

Image processing with KDE

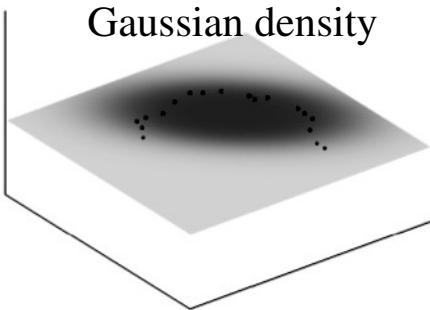


D. Krishnan et al., "Blind Deconvolution Using a Normalized Sparsity Measure", DOI: 10.1109/CVPR.2011.5995521

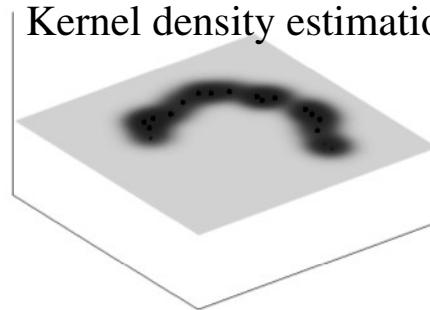
Actual distribution



Gaussian density



Kernel density estimation



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

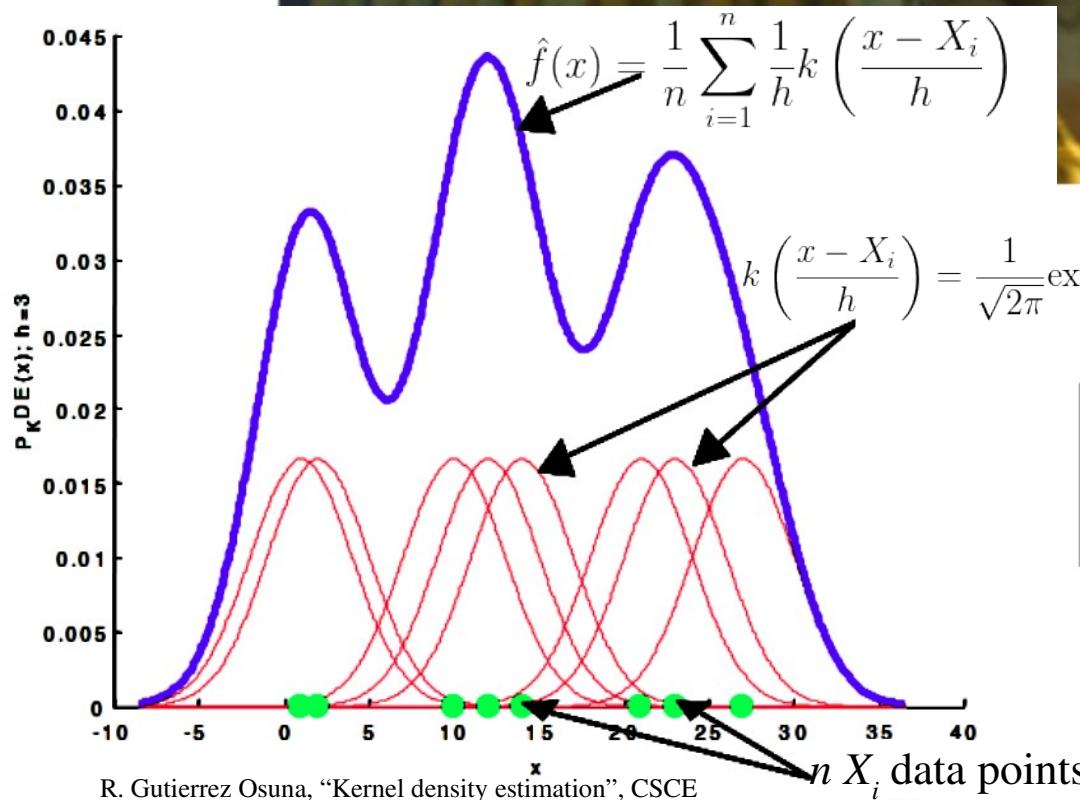
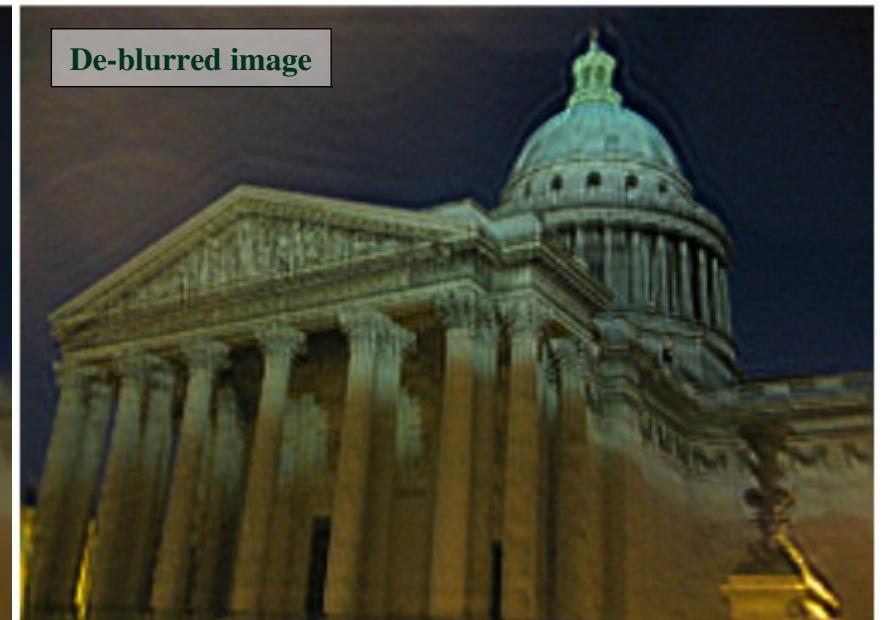
M. Rousson et al., "Efficient Kernel Density Estimation of Shape and Intensity Priors for Level Set Segmentation", DOI:10.1007/978-0-387-68343-0_13

Power of KDE

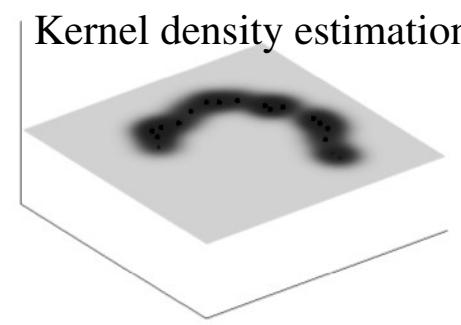
Kernel Density Estimation (KDE)

- 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),

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

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

Smoothness, bias noise, variance

- 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.

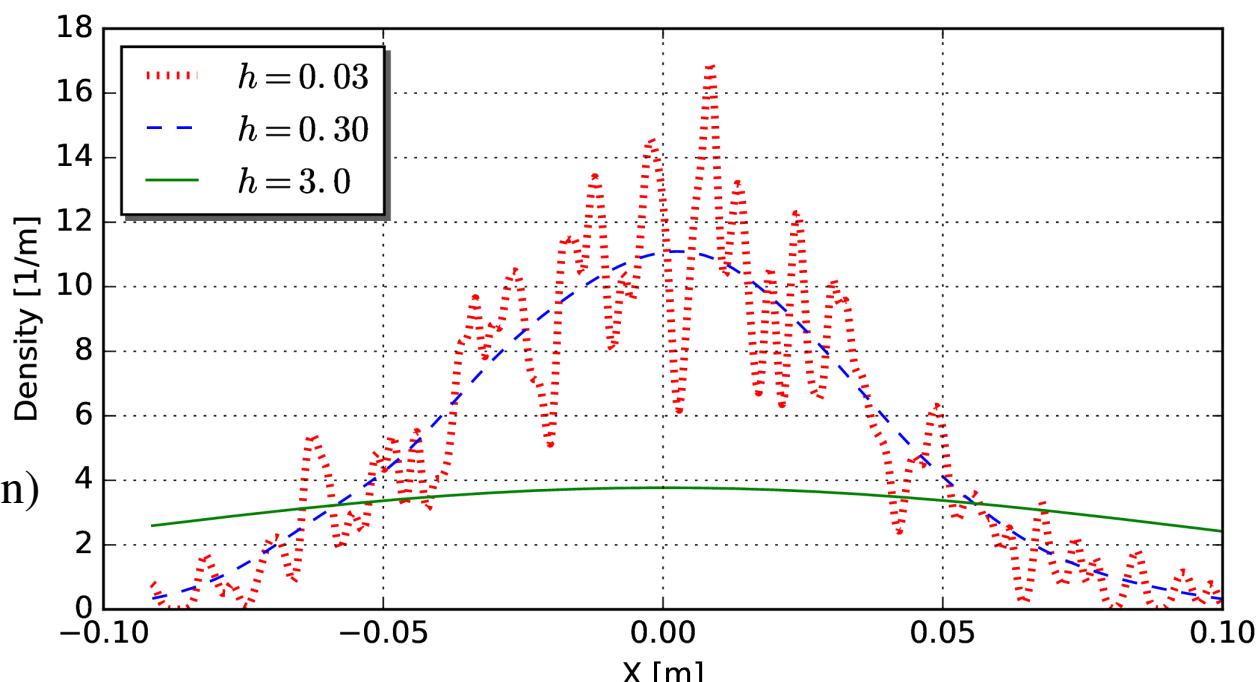
★ $h = 0.3$ (optimal h determined from minimizing MISE term above) reveals a Gaussian,

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

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

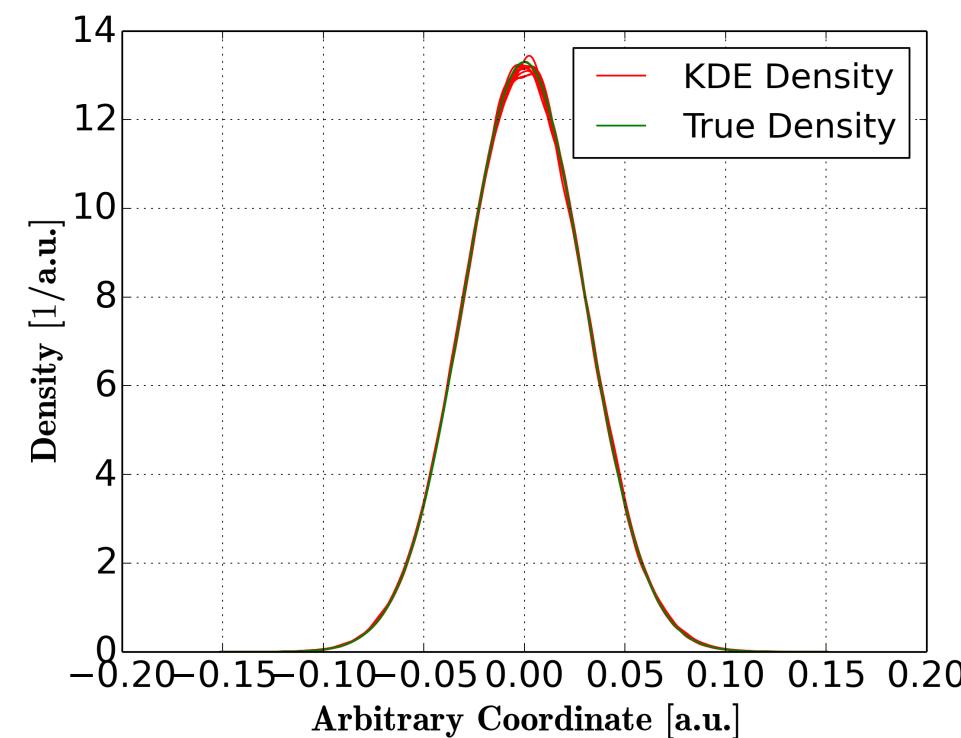
★ $h = 3.0$ over-smoothes the PDF.

★ $h = 0.03$ reveals a noisier PDF.

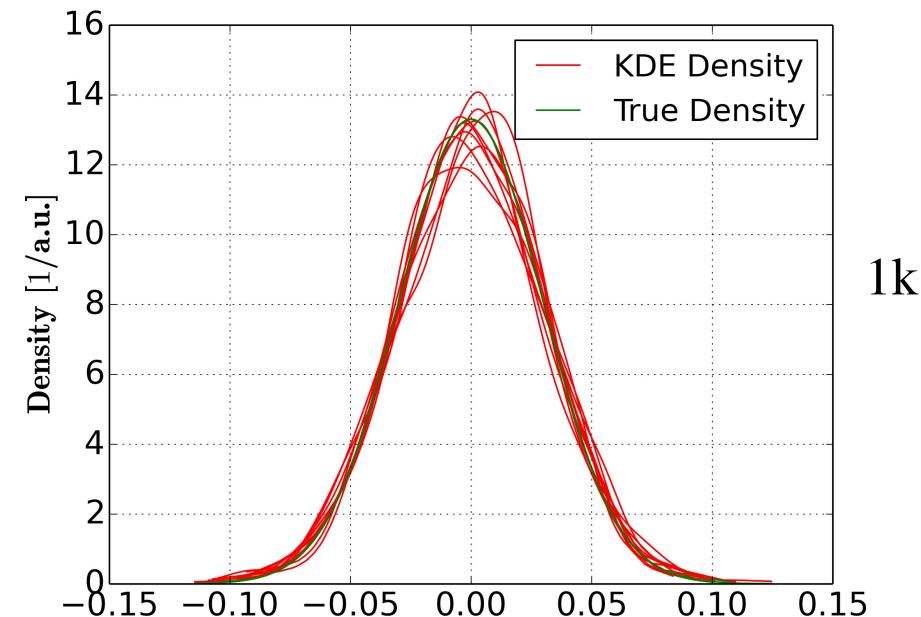


KDE Validation in 1D – sample size study, True vs. KDE

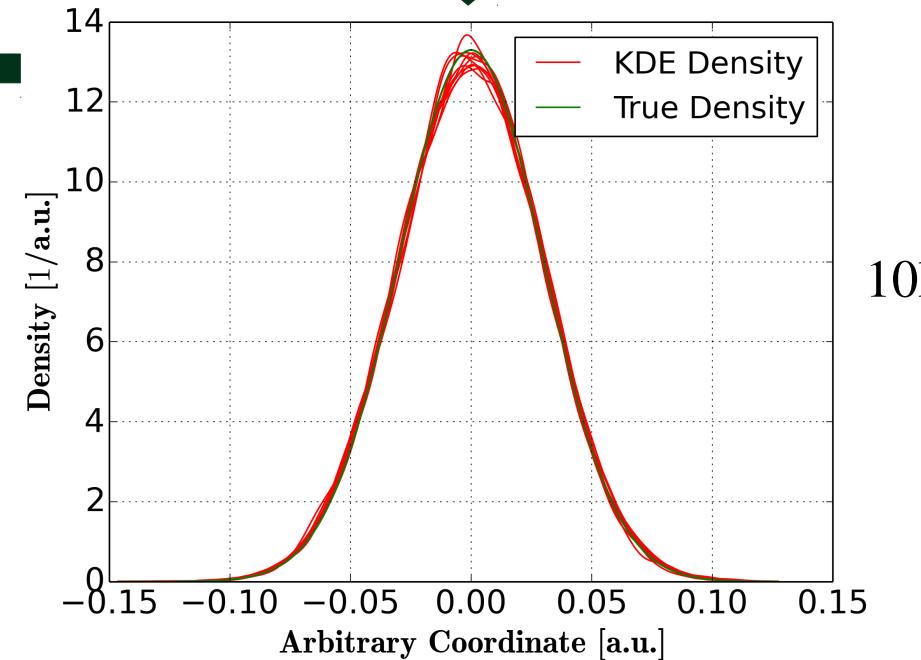
- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u), each with 1k, 10k, and 100k data points.
- Compared their KDE and true densities.
- The bandwidth is optimal (minimized MISE).



100k



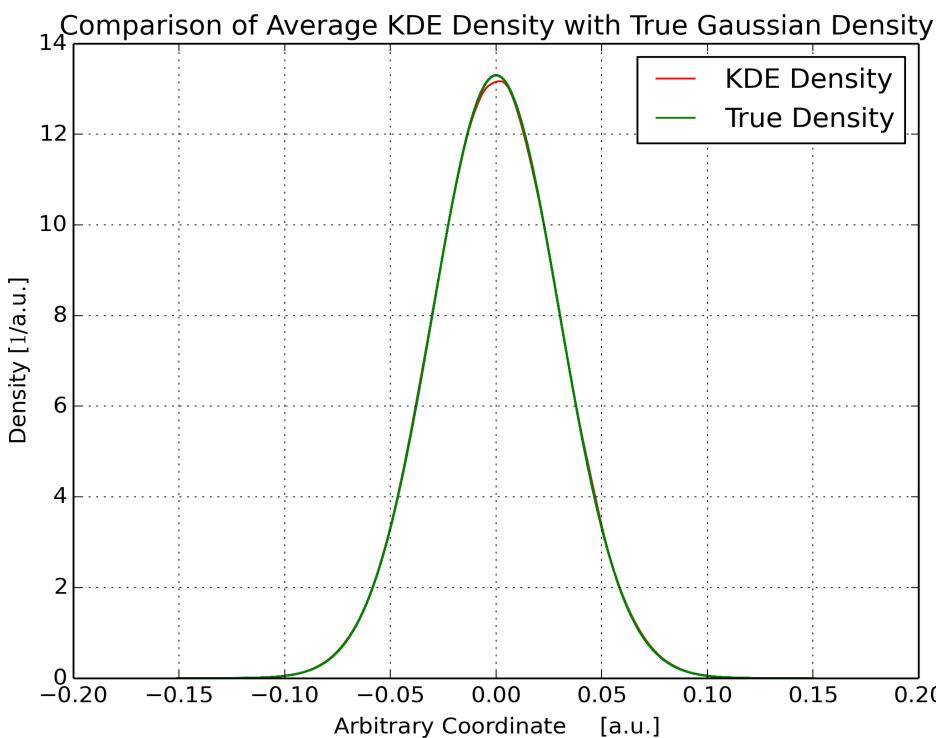
1k



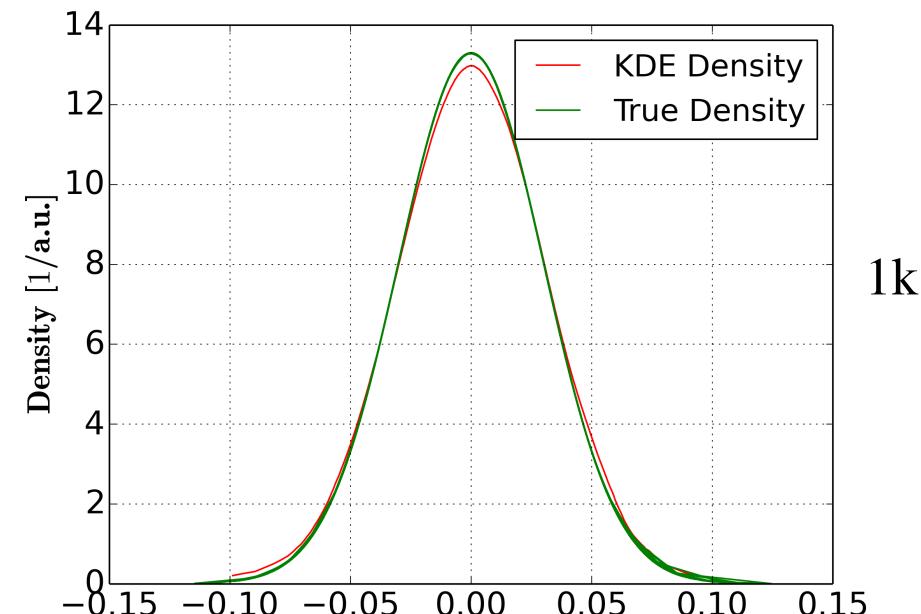
10k

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

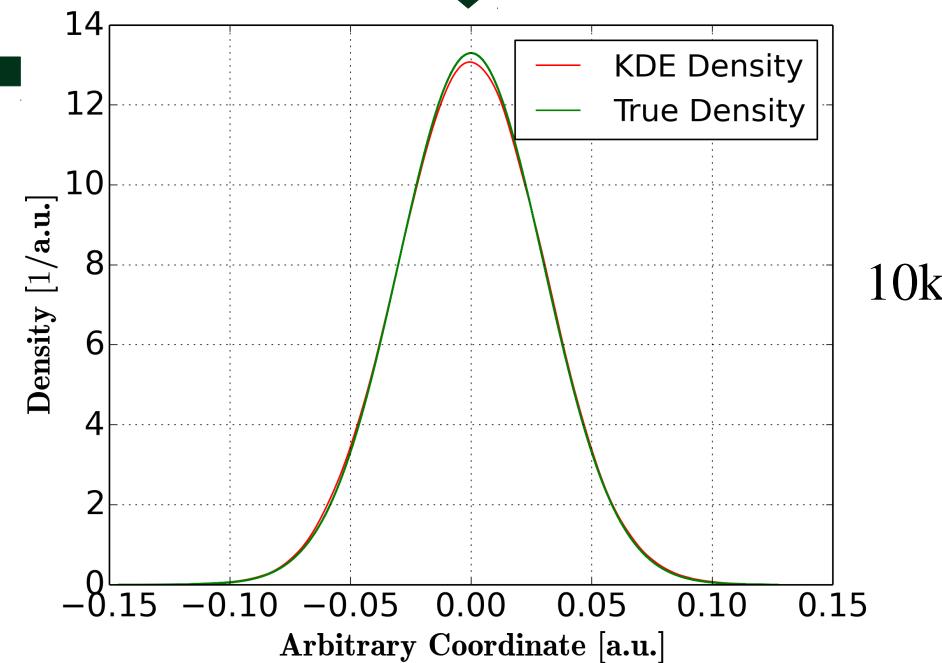
- ★ Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u), each with 1k, 10k, and 100k data points.
- ★ Compared the **averages** of their KDE and true densities.
- ★ The bandwidth is optimal (minimized MISE).



100k



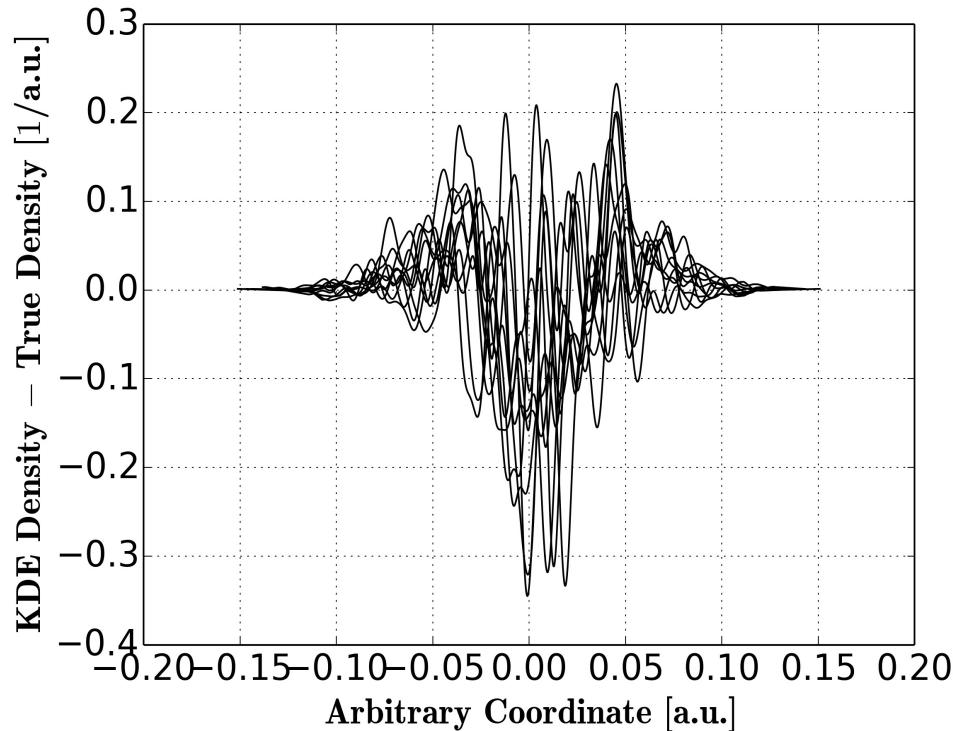
1k



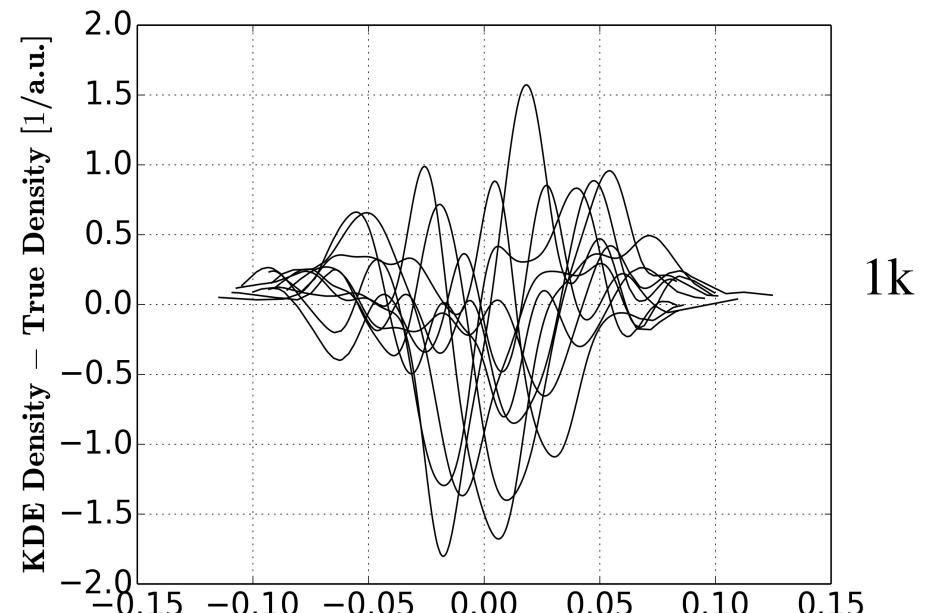
10k

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

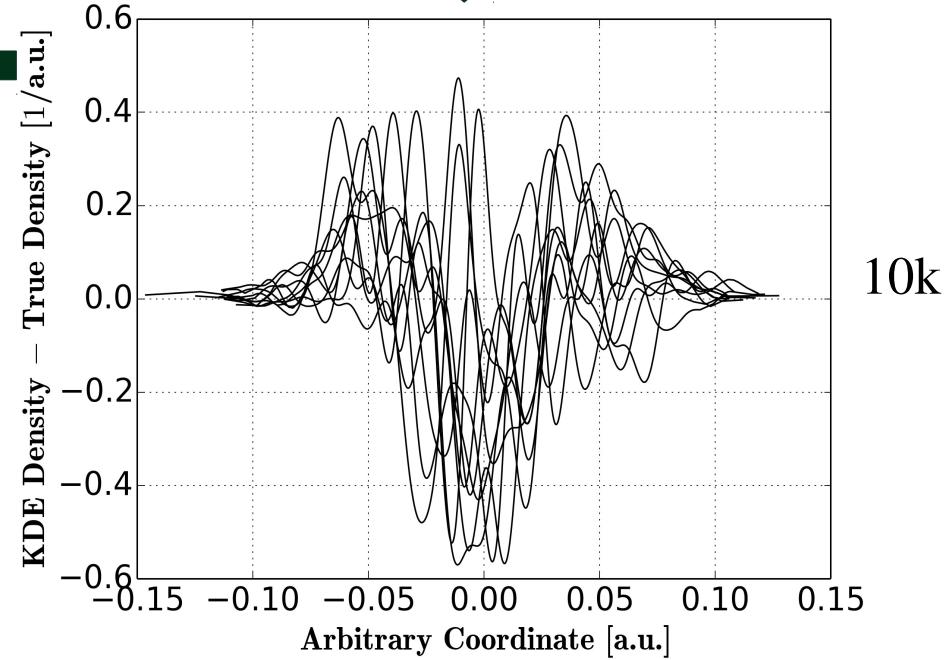
- Generated 10 toy Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points.
- Compared their **KDE errors** (differences between KDE density and true densities).
- The bandwidth is optimal (minimized MISE).



100k



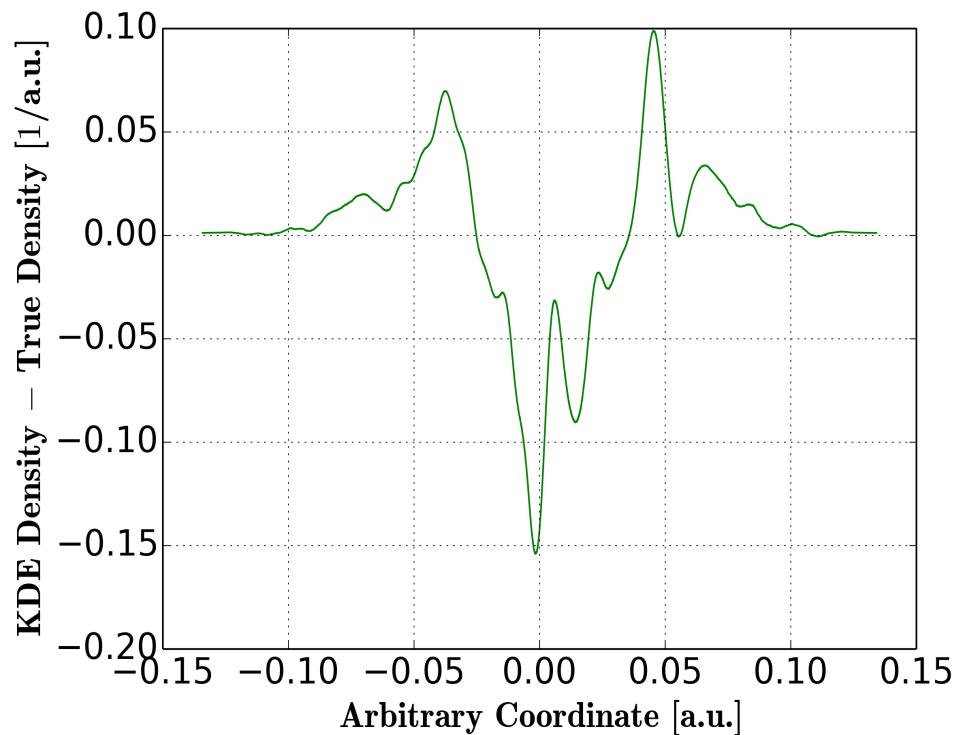
1k



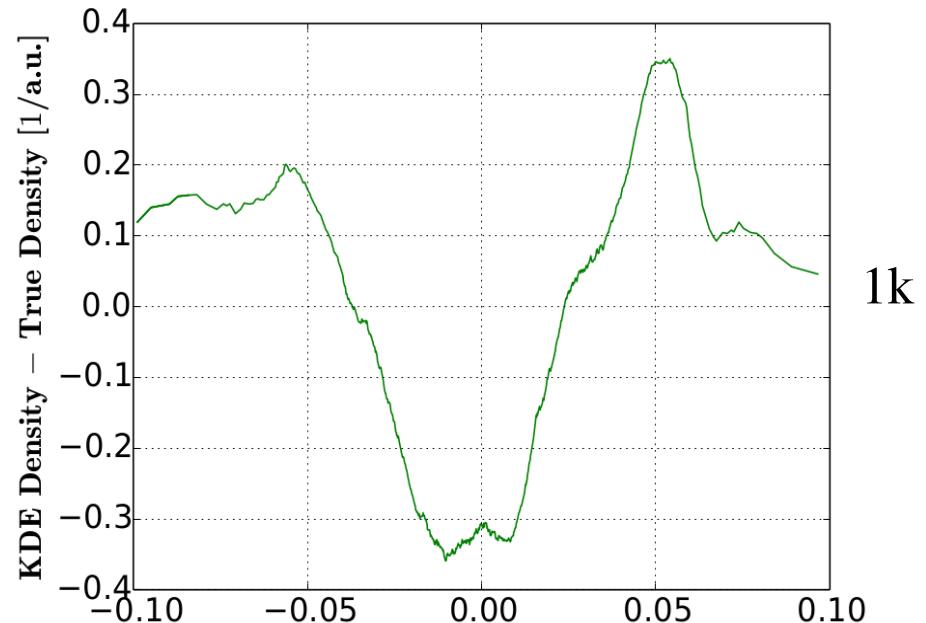
10k

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

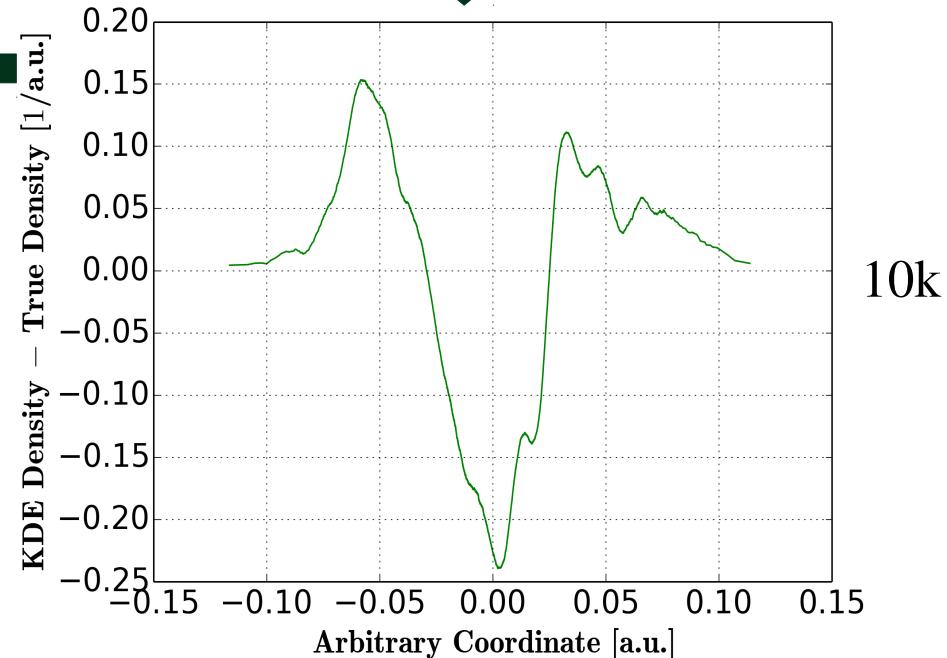
- Generated 10 toy Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points.
- Compared their **KDE error averages**.
- The bandwidth is optimal (minimized MISE).



100k



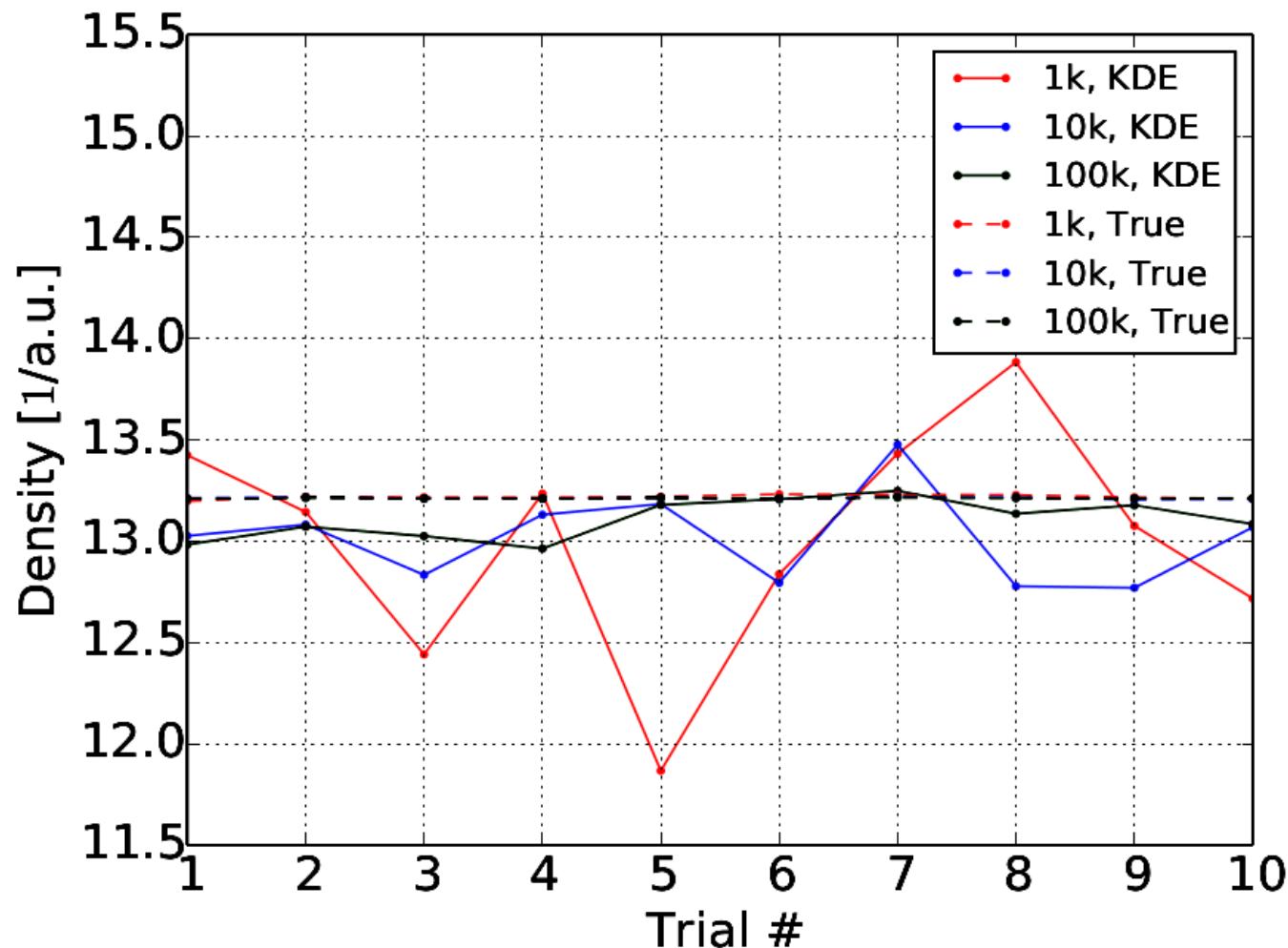
1k



10k

KDE Validation in 1D cont. – summary

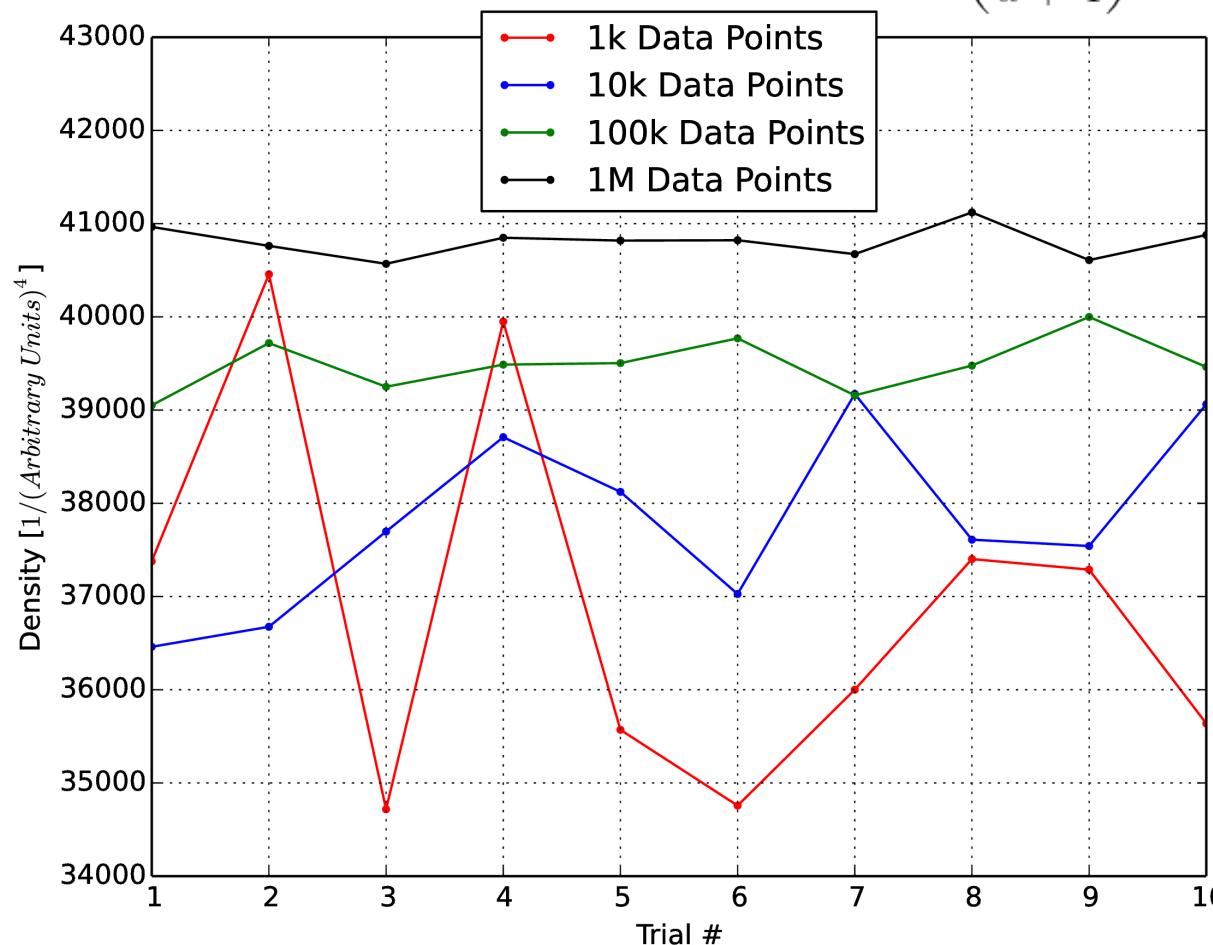
- To summarize, isolated each curve's peak density (density curve enclosing 9% of the sample size):
 - ★ KDE density stabilizes and approaches true density curve as sample size grows.
 - ★ Slight increase in mean density with growing sample size (caused by the optimal bandwidth's dependence on sample 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.
- Larger increase in mean density with sample size (compared with 1D) – optimal bandwidth's depends on sample size **and dimension variable, d**:

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

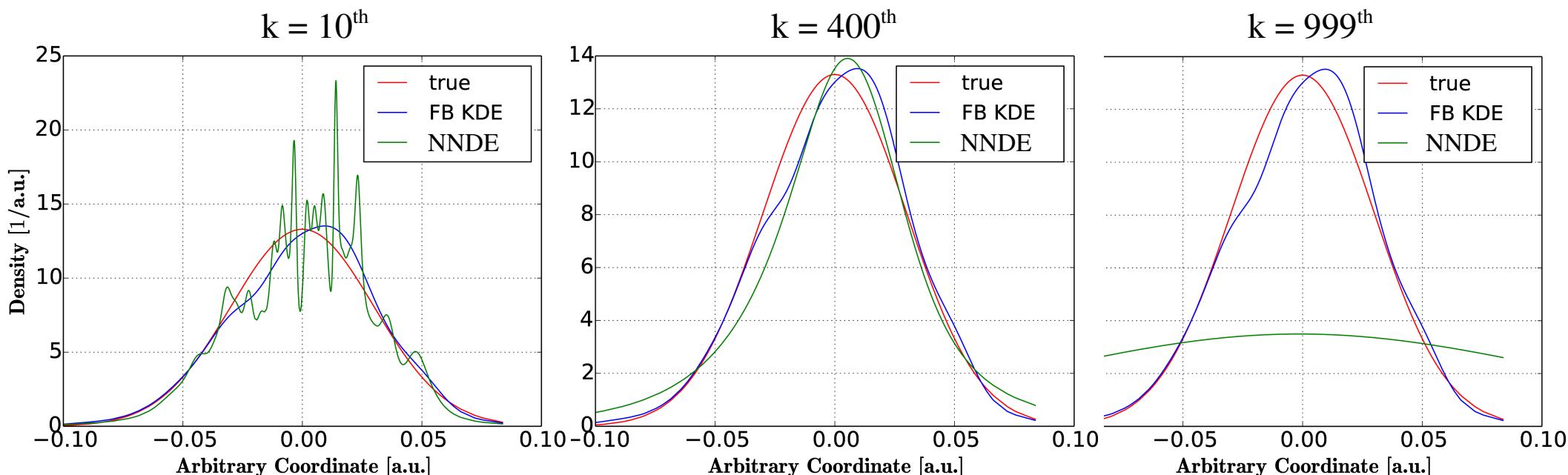


Nearest Neighbor Density Estimation (NNDE)

- KDE: equal contribution from each data point. Some sensitivity to distribution tails.
- 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^{th} nearest neighbor, sample size, n

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{d_i} k\left(\frac{|x - X_i|}{2d_i^2}\right) \quad d_i: \text{the euclidean distance between } i^{\text{th}} \text{ data point to } k^{\text{th}} \text{ nearest neighbor data point.}$$

- 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.

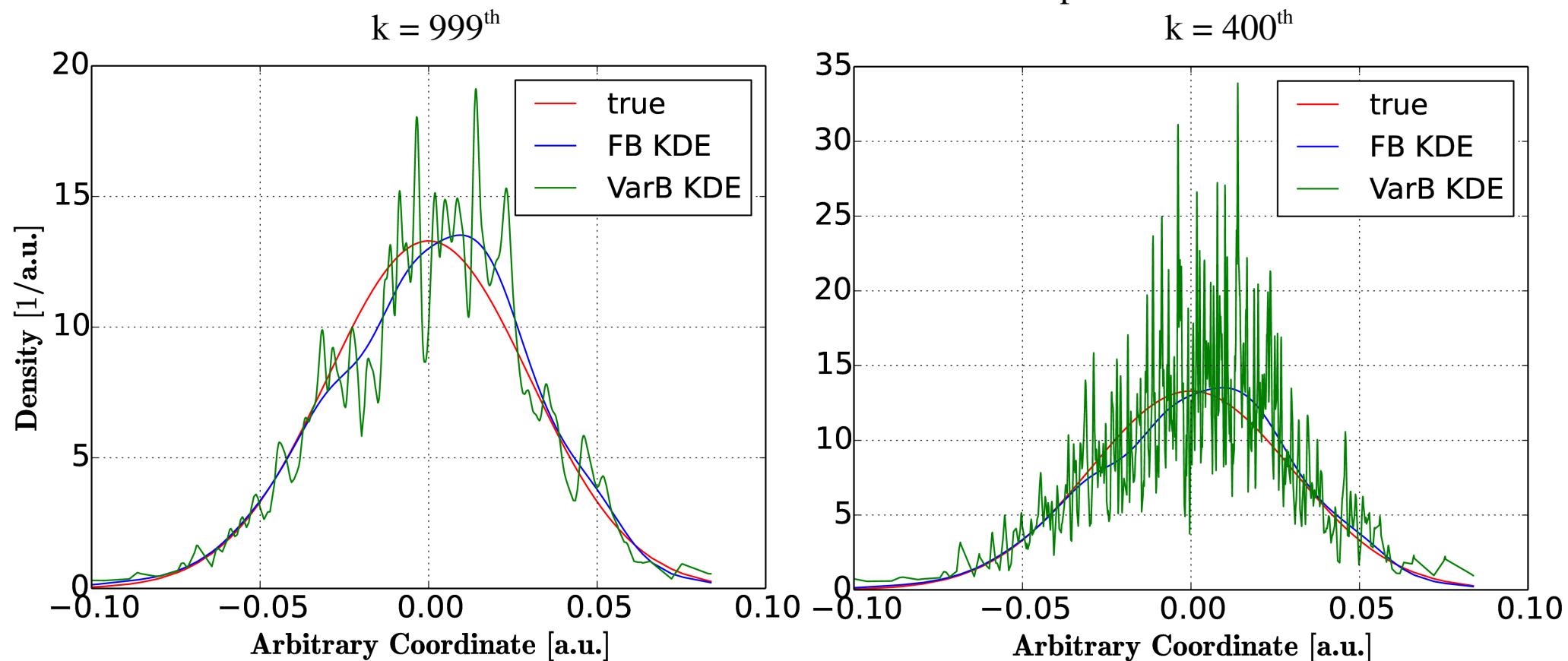


Combining KDE and NNDE – Variable KDE

- ★ Hybrid of KDE and NNDE:

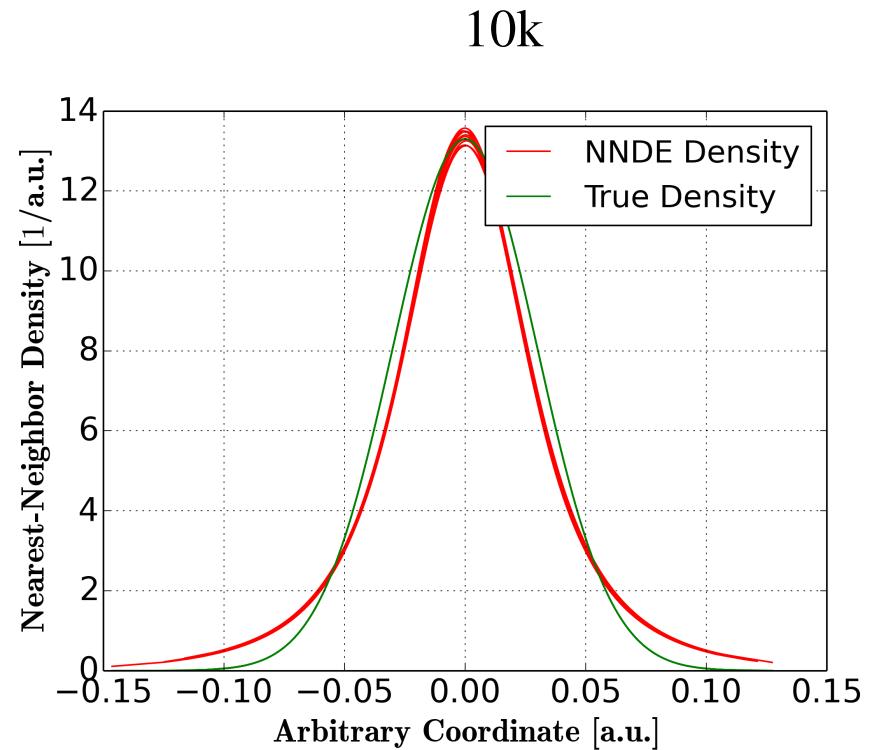
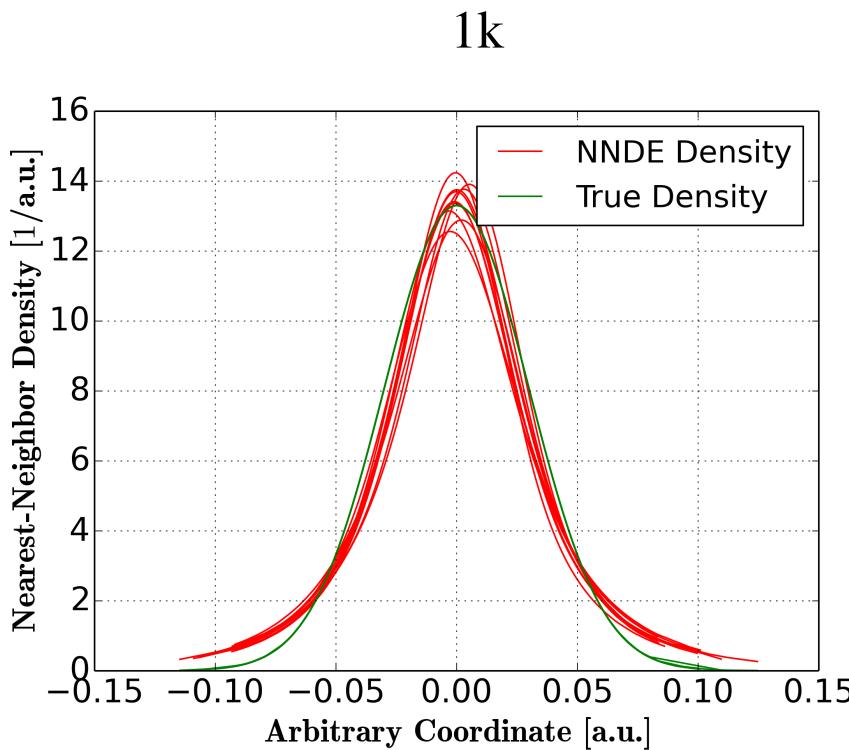
$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{hd_i} k\left(\frac{|x - X_i|}{2h^2 d_i^2}\right)$$

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



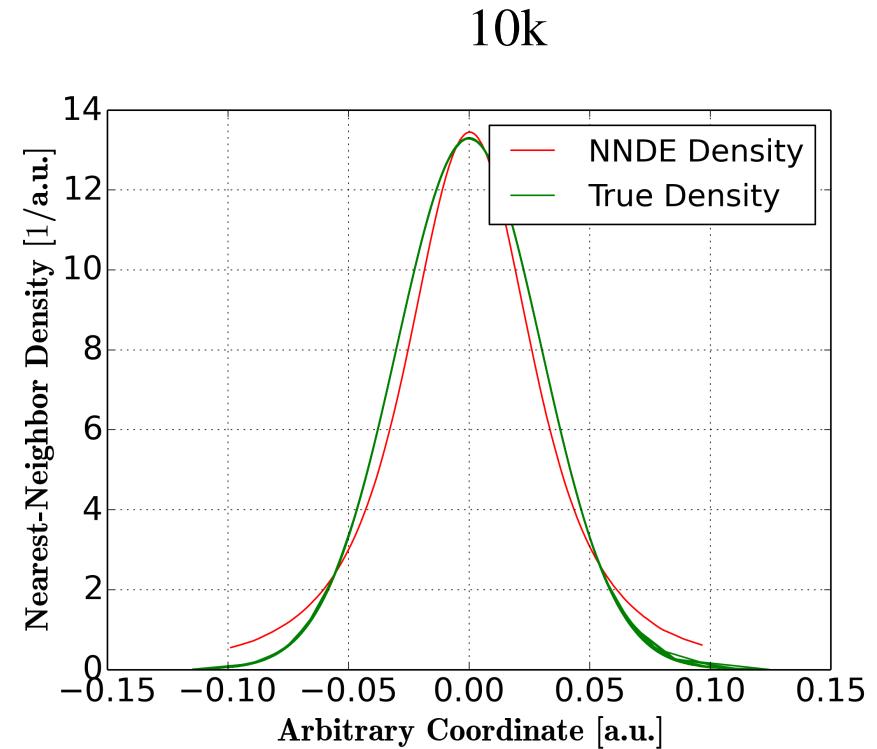
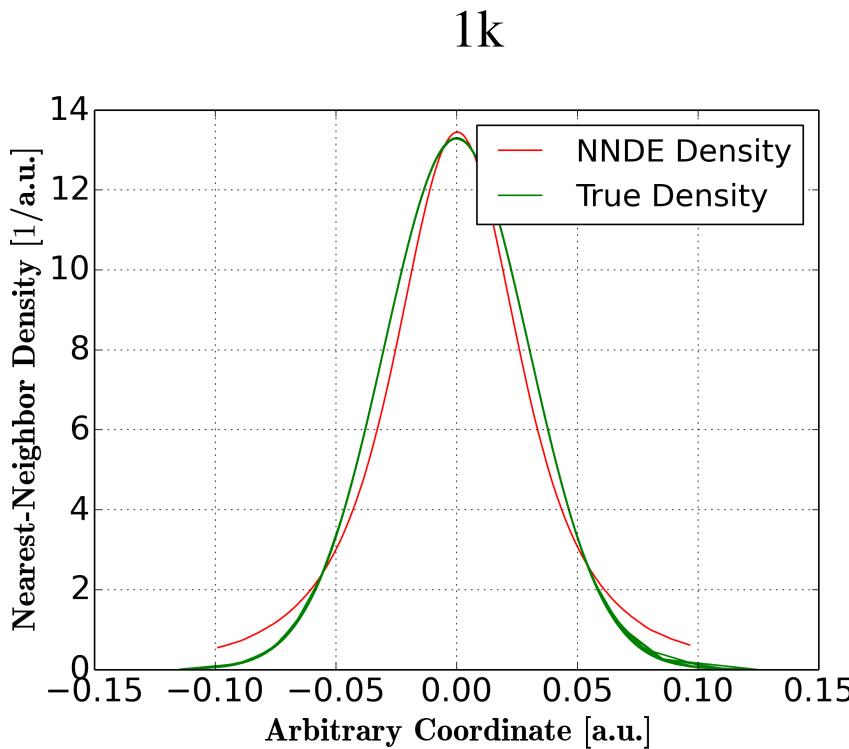
NNDE Validation in 1D – sample size study, True vs. NNDE

- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points, compared their NNDE and true densities.



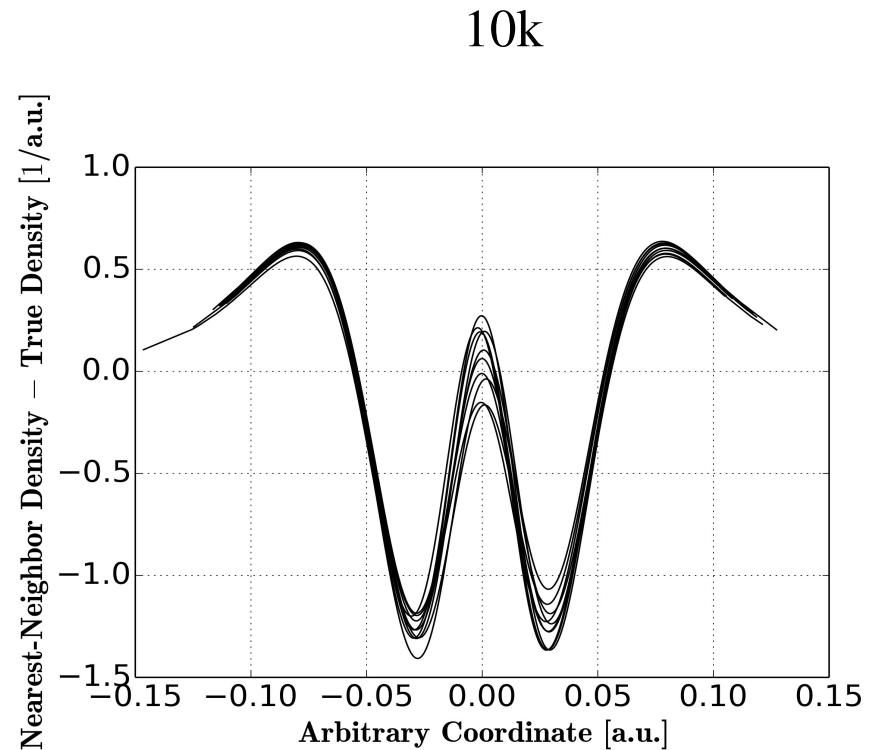
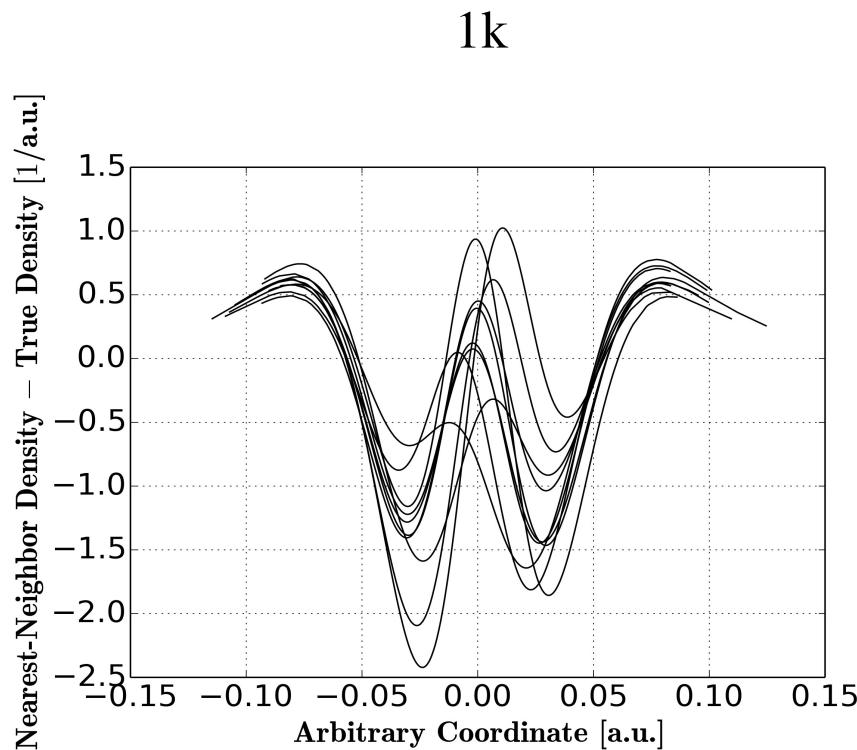
NNDE Validation in 1D – sample size study, True vs. NNDE

- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points, compared the averages of their NNDE and true densities.



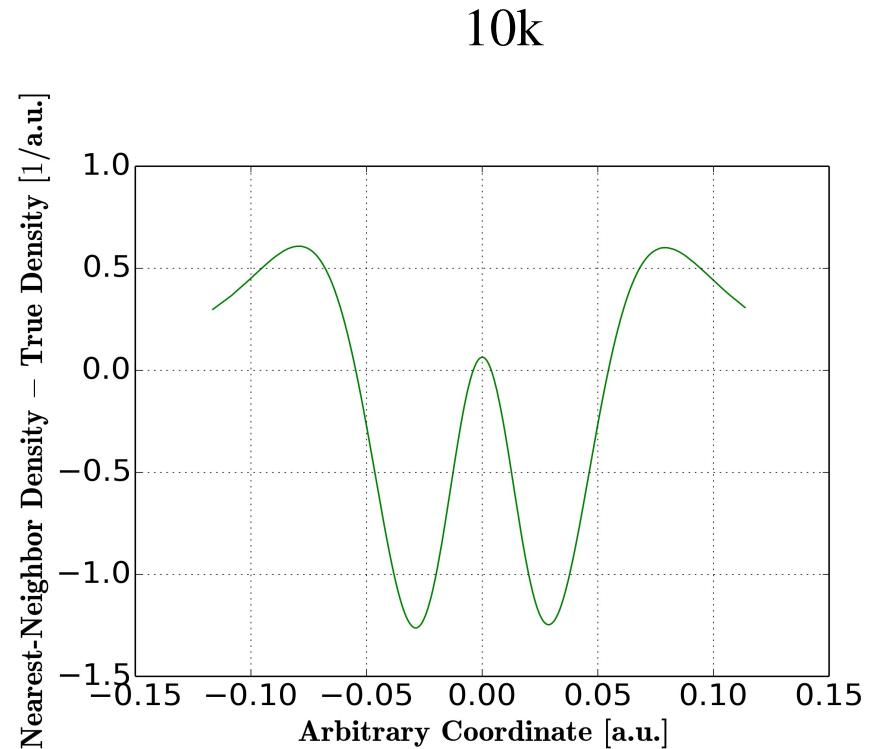
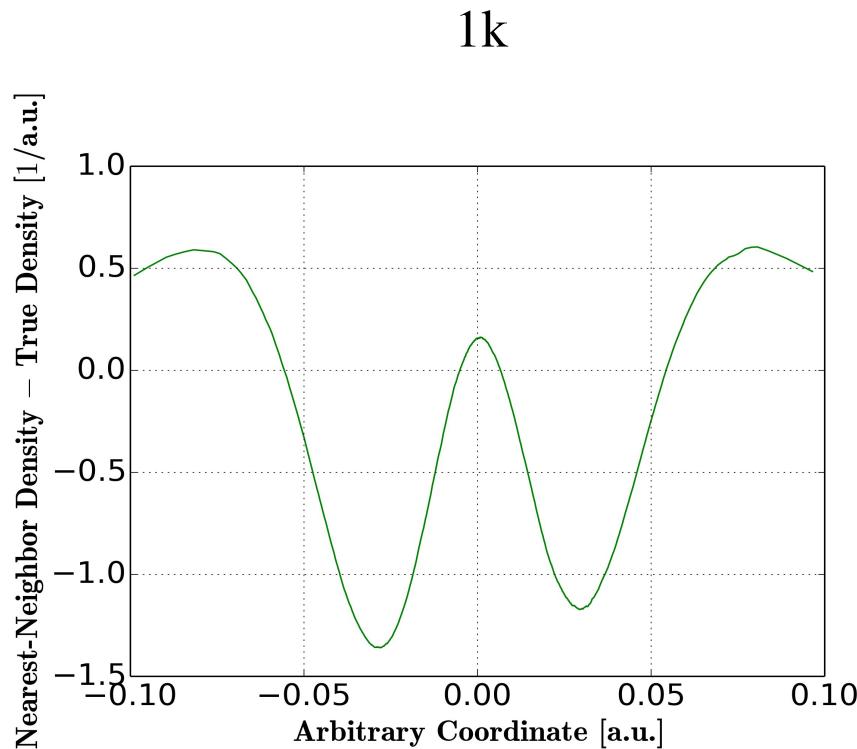
NNDE Validation in 1D – sample size study, True vs. NNDE

- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points, compared their NNDE errors (difference between NNDE and true density).



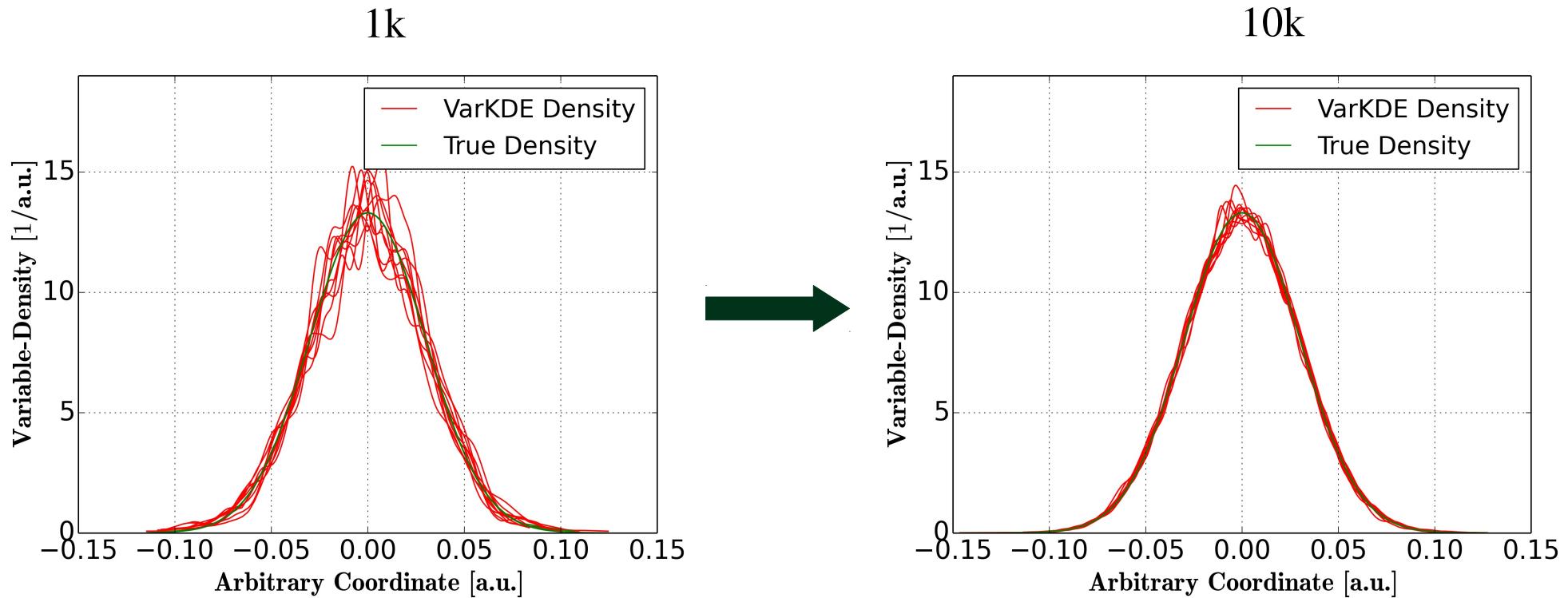
NNDE Validation in 1D – sample size study, True vs. NNDE

- ★ Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points, compared their averaged NNDE errors (difference between NNDE and true density).



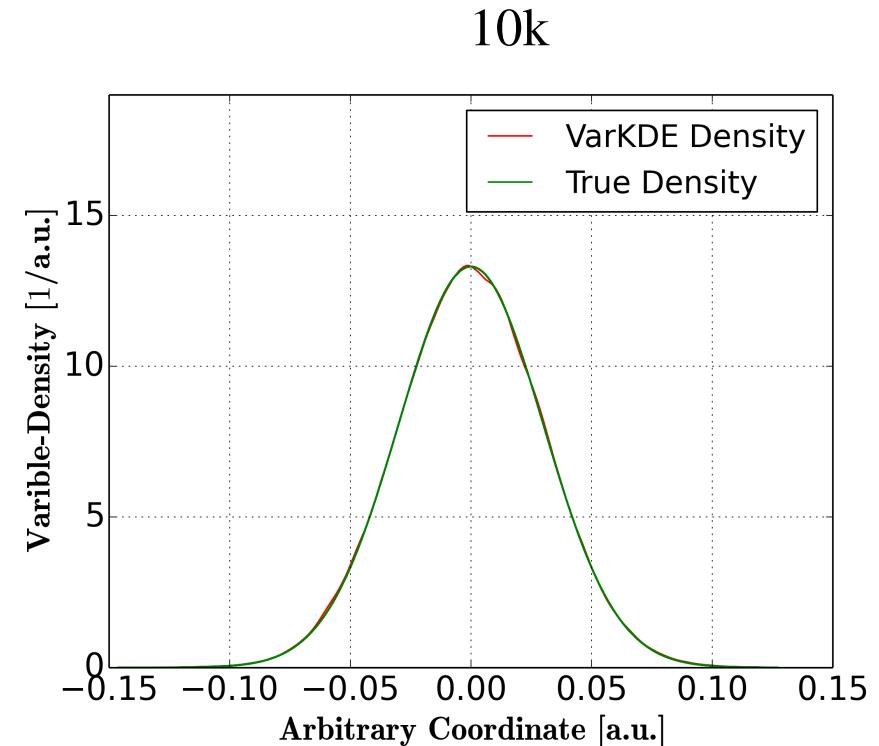
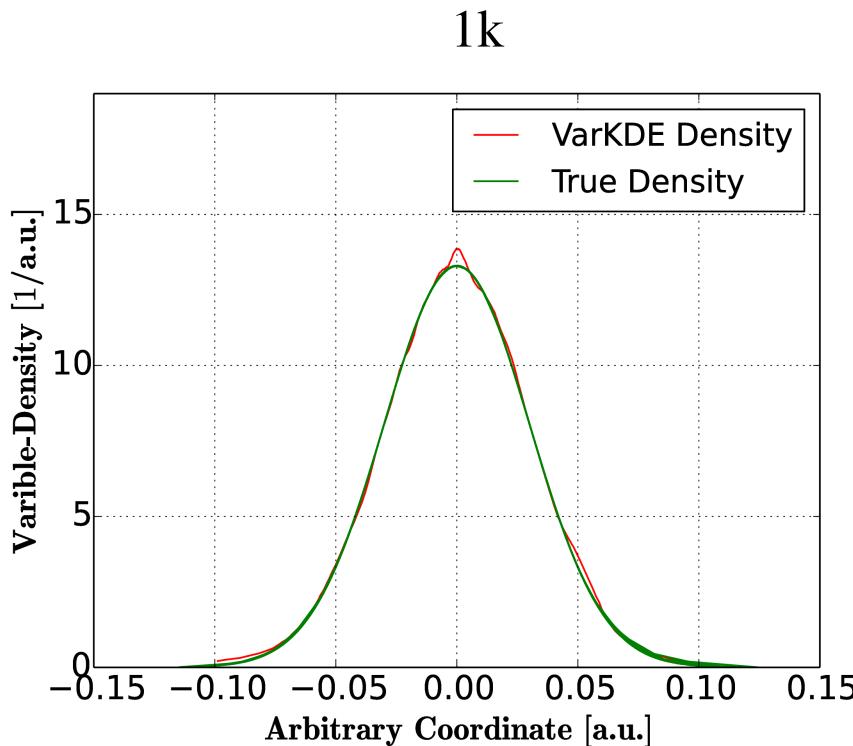
VarKDE Validation in 1D – sample size study, True vs. VarKDE

- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u), each with 1k, 10k, and 100k data points, compared their VarKDE and true densities.



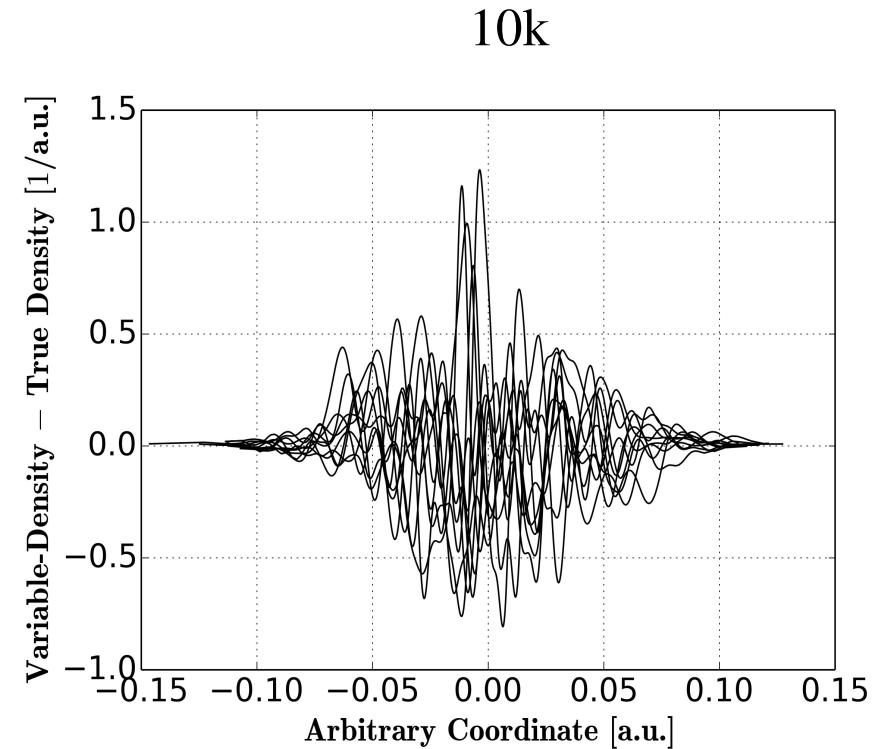
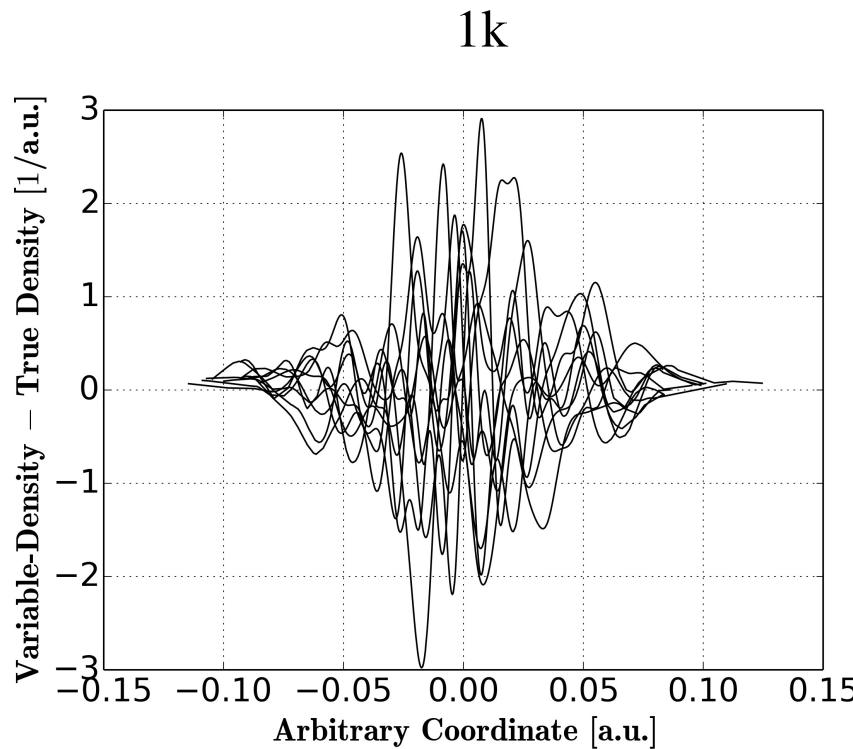
VarKDE Validation in 1D – sample size study, True vs. VarKDE

- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u), each with 1k, 10k, and 100k data points, compared the averages of their VarKDE and true densities.



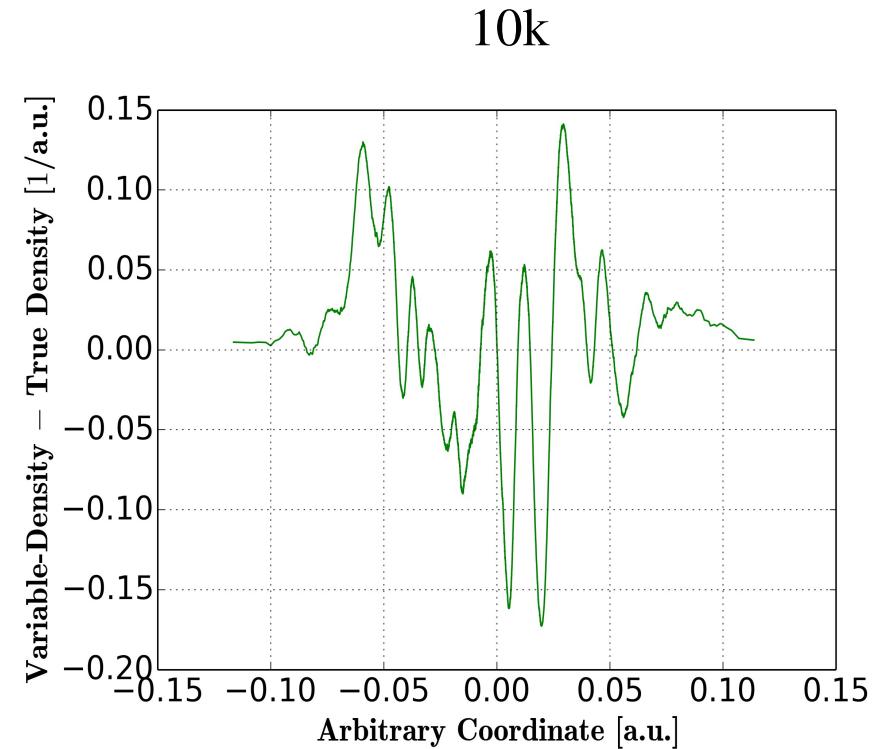
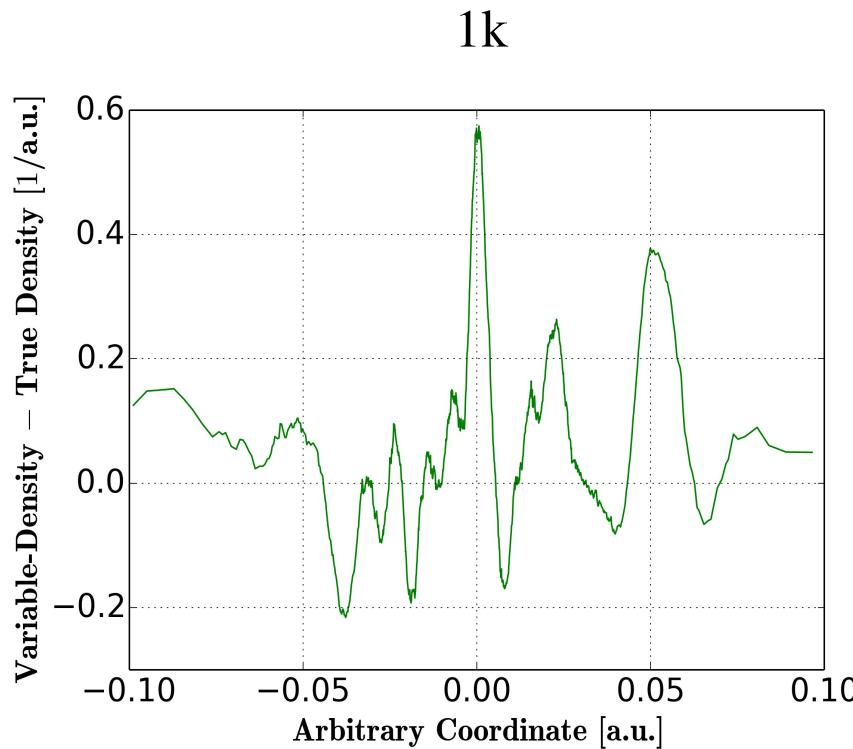
VarKDE Validation in 1D – sample size study, True vs. VarKDE

- ★ Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u.), each with 1k, 10k, and 100k data points, compared their VarKDE errors (difference between VarKDE and true density).



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

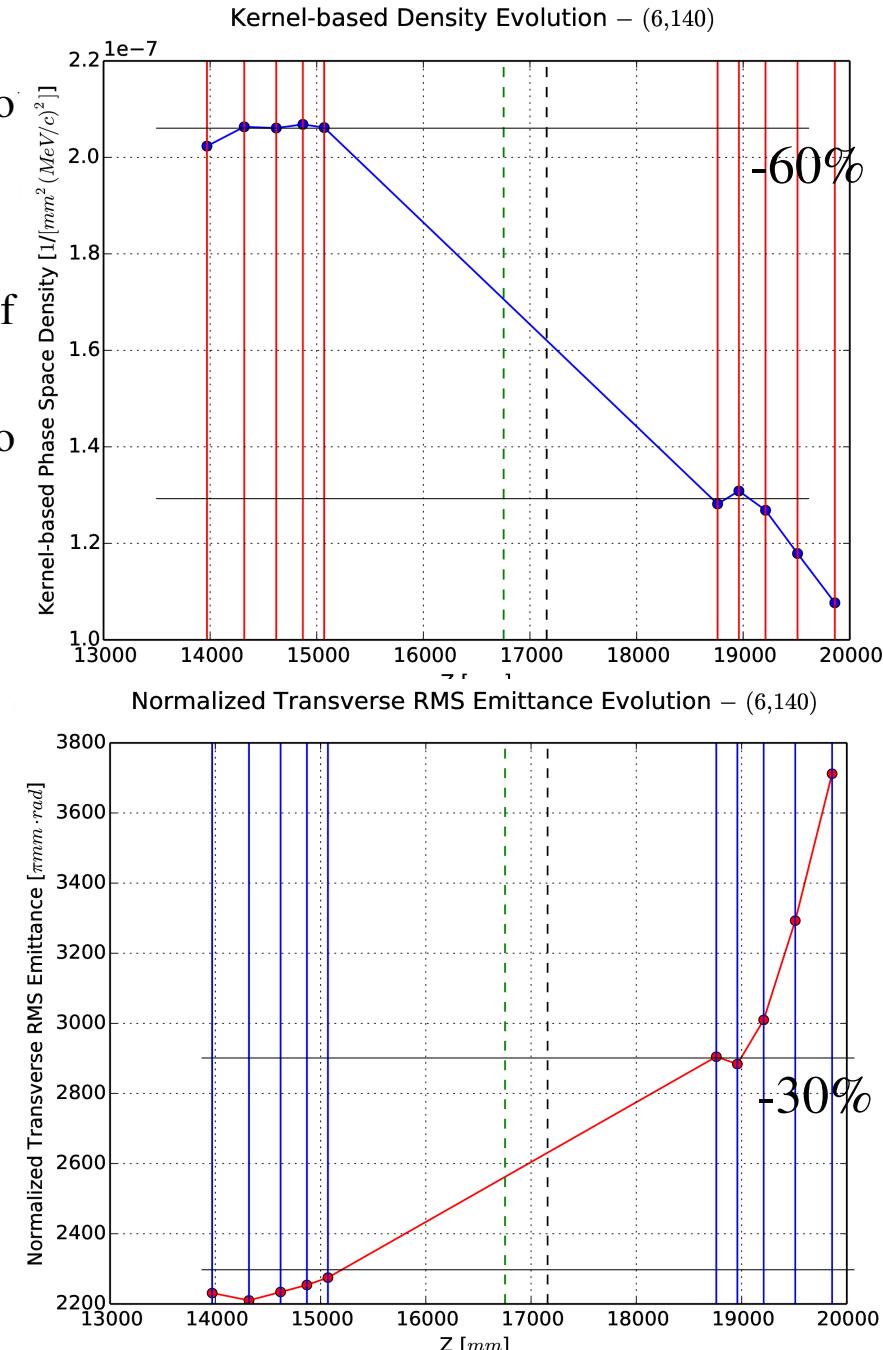
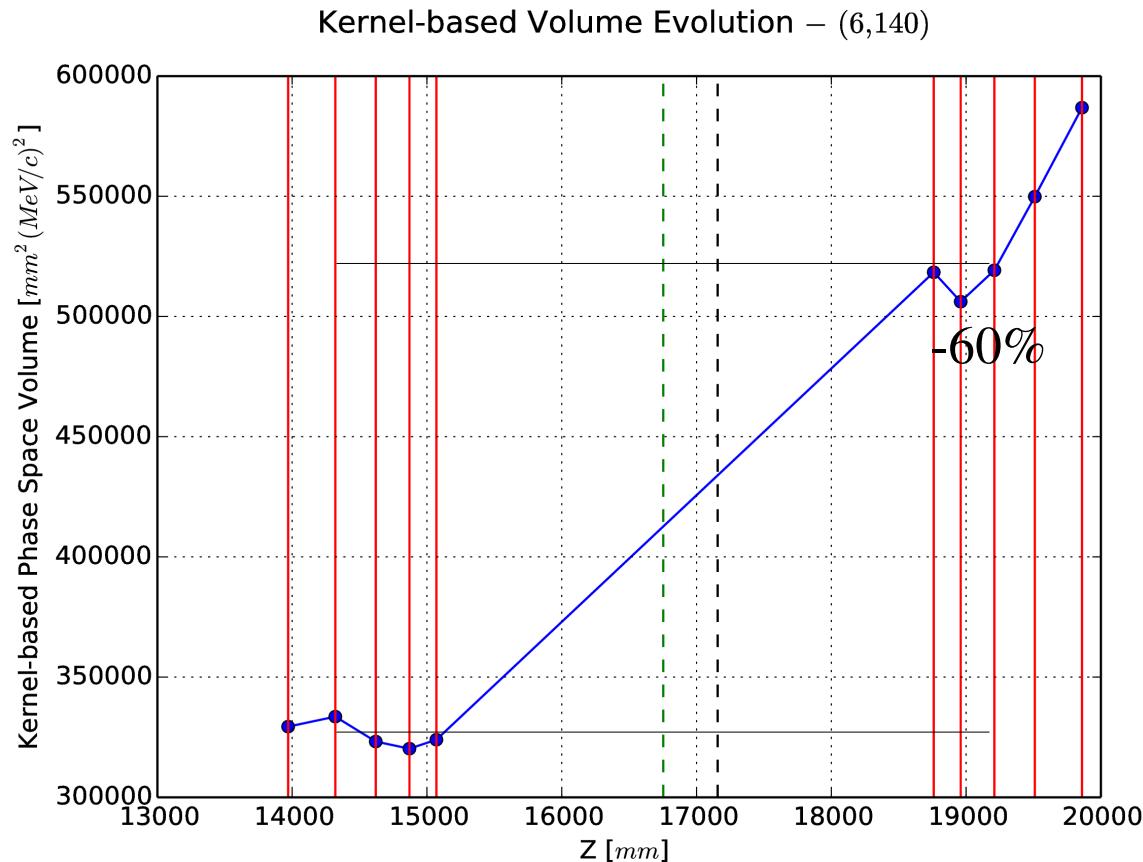
- Generated 10 Gaussian distributions ($\sigma = 0.03$ a.u), each with 1k, 10k, and 100k data points, compared their averaged VarKDE errors (difference between VarKDE and true density).



Preliminary KDE on recent MICE Data

Run 8681:

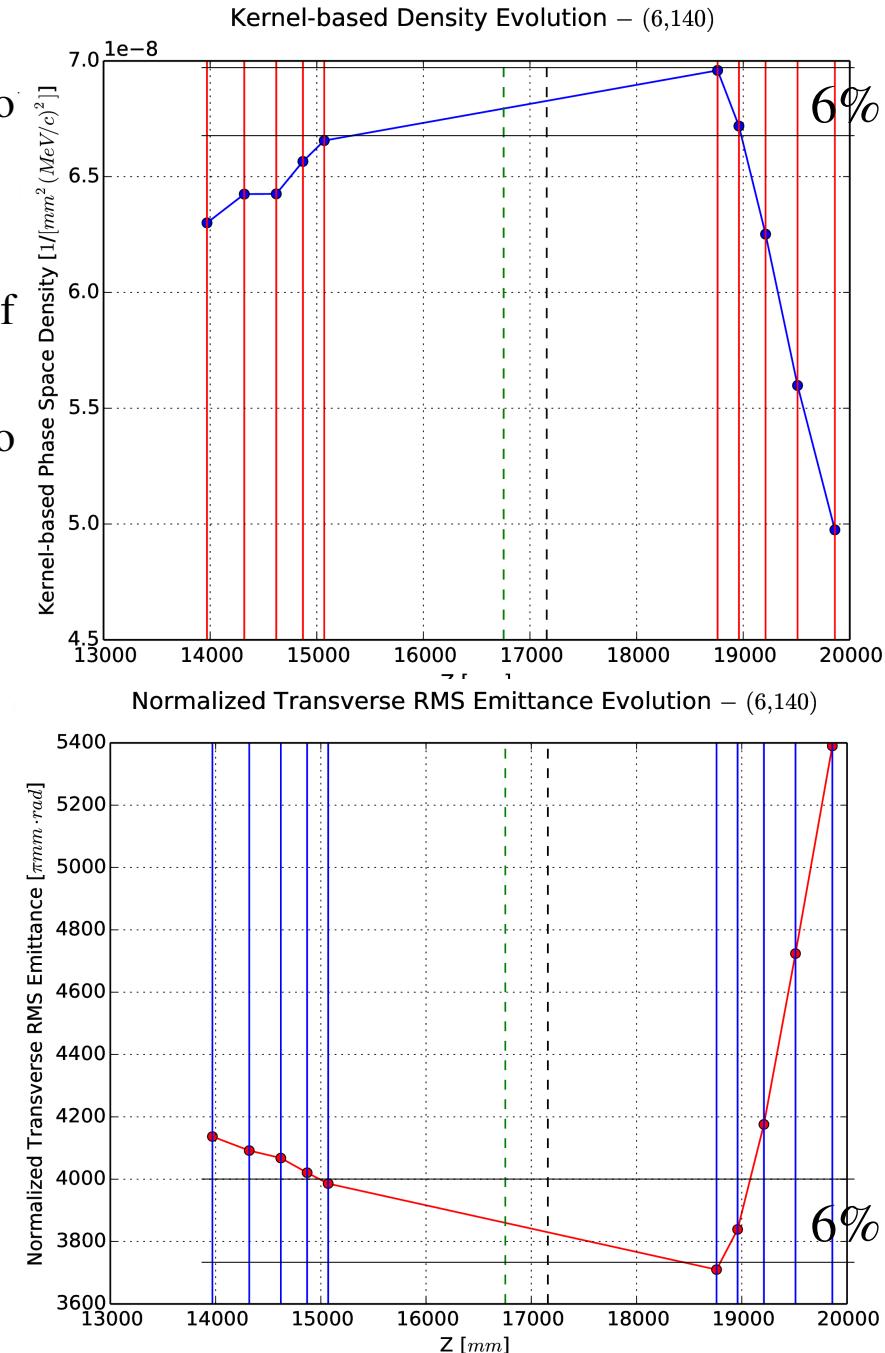
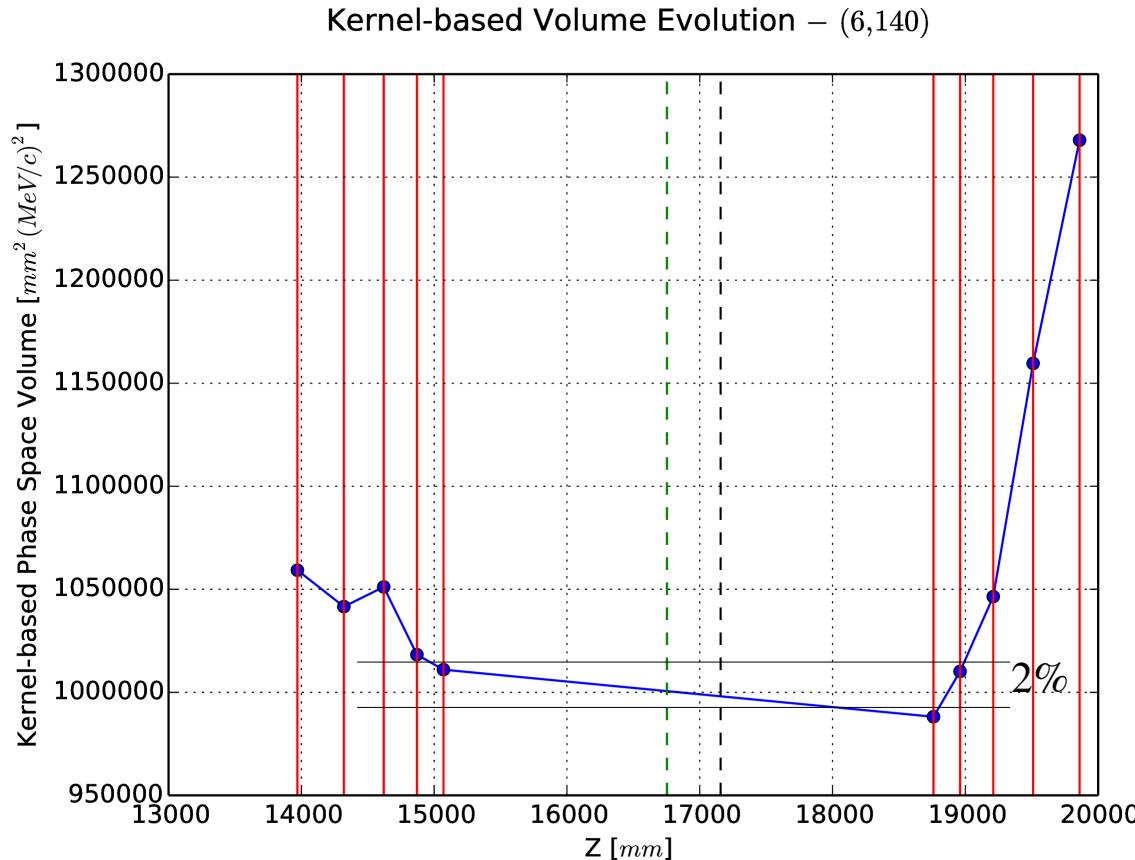
- ▶ 3-140 beam setting with LiH absorber in the channel and no currents in M1D & M2D coils.
- ▶ Beam heating according to changes in all 3 quantities.
- ▶ Dashed vertical lines: FCU & FCD. Solid lines: locations of the tracker stations.
- ▶ Good muon cut to discard muons which do not make it to TKD. $32 < \text{TOF12} < 39 \text{ ns}$, $100 < p < 220 \text{ MeV}/c$ cuts.



Preliminary KDE on recent MICE Data

■ Run 8699:

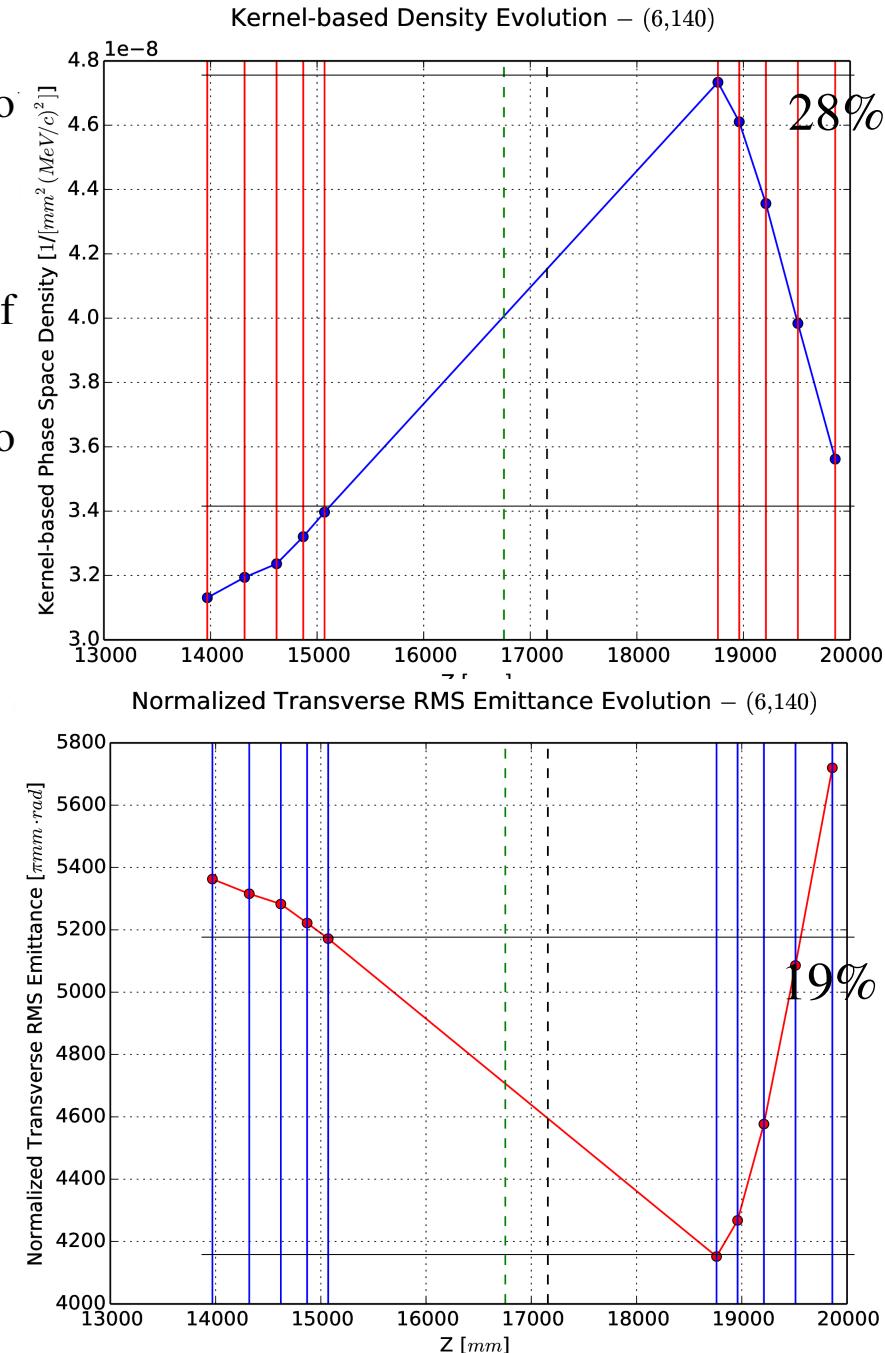
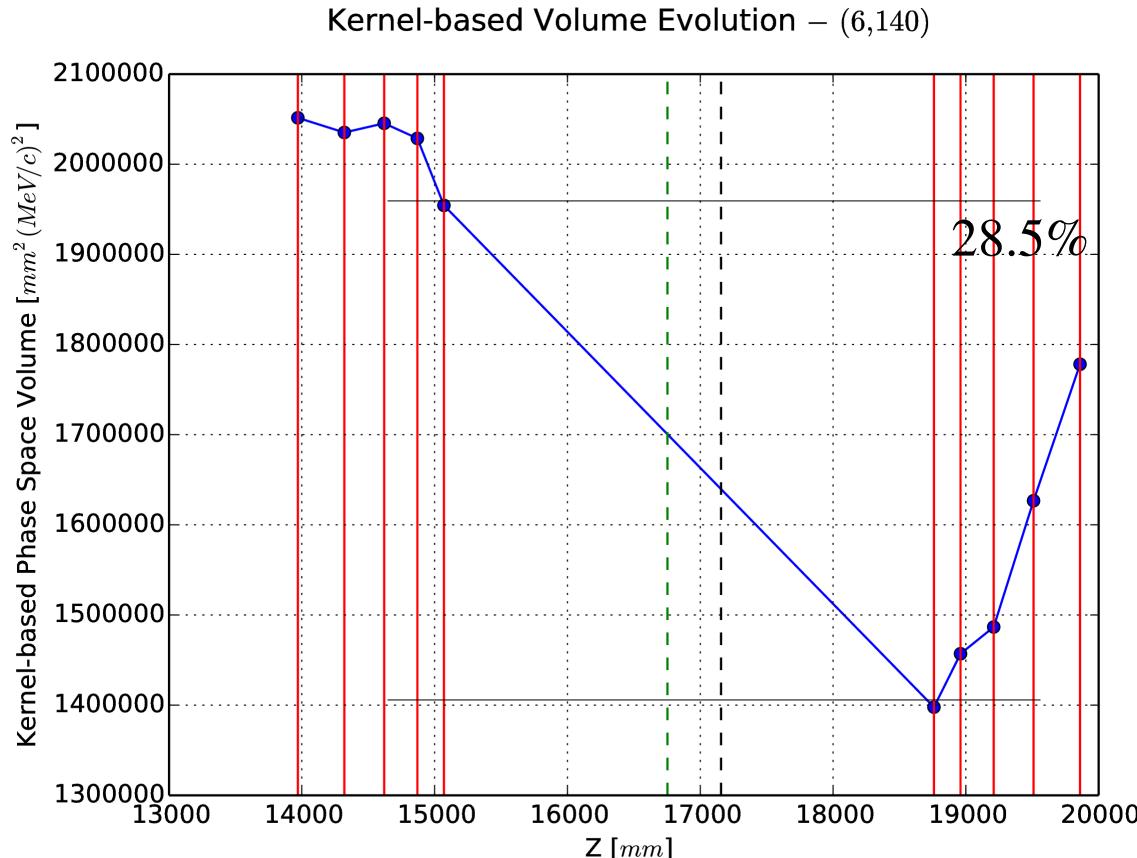
- ▶ 6-140 beam setting with LiH absorber in the channel and no currents in M1D & M2D coils.
- ▶ Beam cooling according to changes in all 3 quantities.
- ▶ Dashed vertical lines: FCU & FCD. Solid lines: locations of the tracker stations.
- ▶ Good muon cut to discard muons which do not make it to TKD. $32 < \text{TOF12} < 39 \text{ ns}$, $100 < p < 220 \text{ MeV}/c$ cuts.



Preliminary KDE on recent MICE Data

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 in 10-140 beam compared with the RMS emittance.
- VarKDE and NNDE perform well with a Gaussian distribution:
 - ★ Further application of the techniques to simulation, data, and long-tailed distributions (log Gaussian) in progress.
- 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!