Statistical Analysis for Uncertainty Quantification and Visualization of Scientific Data

Tushar M. Athawale

DOE CGF 2022

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

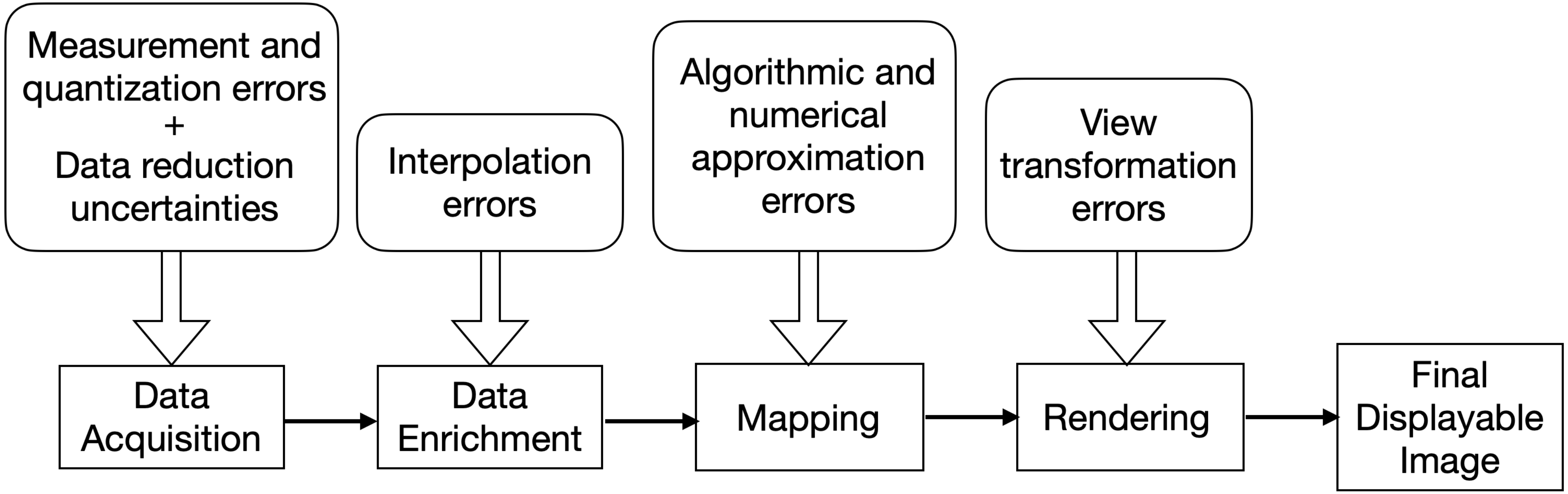
(Collaboration with the SCI Institute at the University of Utah)

www.sci.utah.edu

Uncertainty Visualization: Top Research Challenge

[A. T. Pang, C. M. Wittenbrink, and S. K. Lodha, "Approaches to Uncertainty Visualization", 1997]
[C. R. Johnson and A. R. Sanderson, "A Next Step: Visualizing Errors and Uncertainty", 2004]

Challenge: Not easy to quantify and convey uncertainties propagated through visualization algorithms!

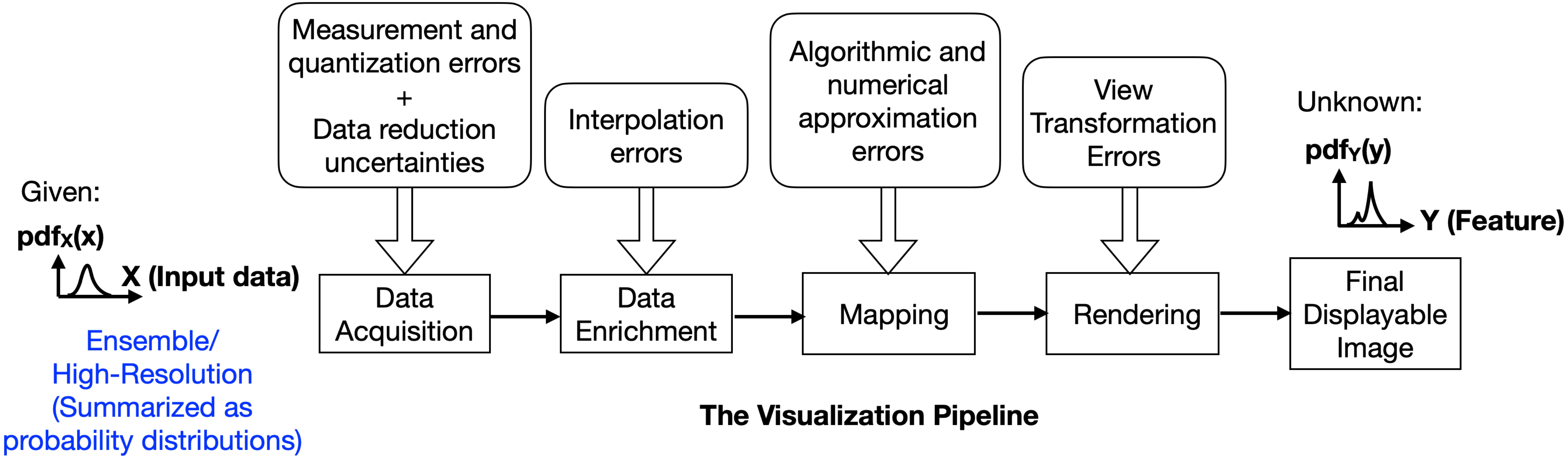


The Visualization Pipeline

[K. Brodlie, R. A. Osorio, and A. Lopes, "A Review of Uncertainty in Data Visualization", 2012]
[A. Kamal et al., "Recent Advances and Challenges in Uncertainty Visualization", 2021]

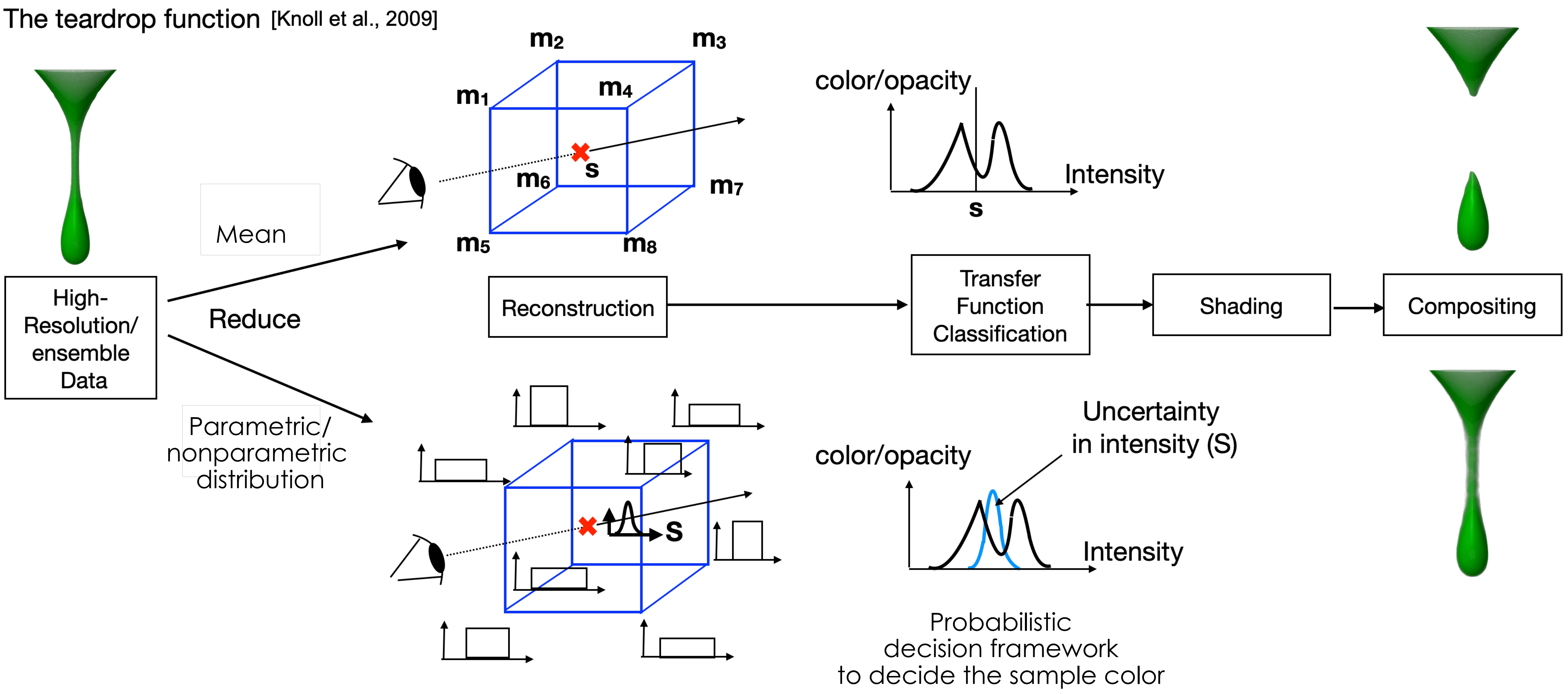
Our Approach to Uncertainty Quantification

Monte Carlo (easy but expensive) vs. Analytical (difficult but fast)



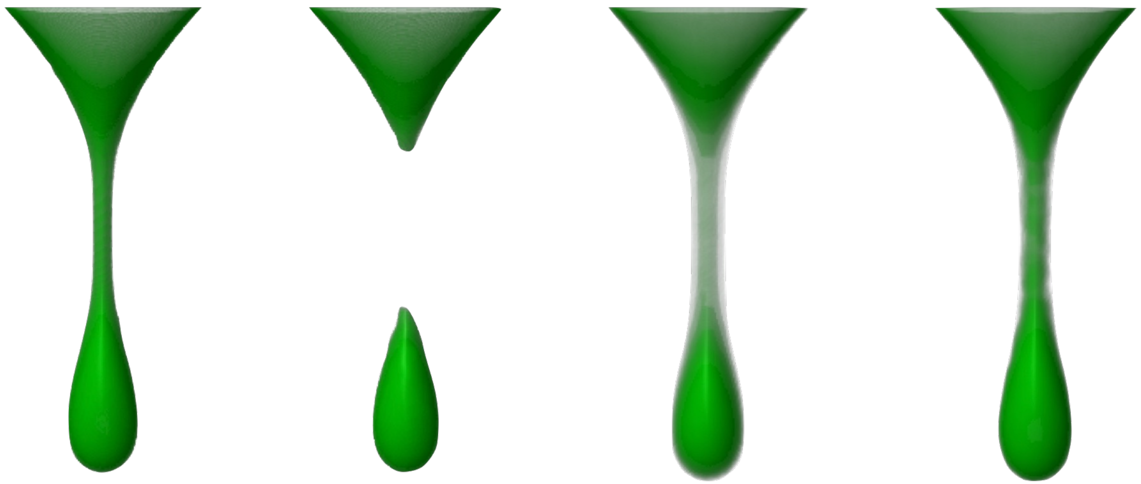
The Visualization Pipeline

Uncertainty-Aware Direct Volume Rendering



OAK RIDGE National Laboratory

Uncertainty-Aware Direct Volume Rendering



(a) Ground truth

(b) Mean

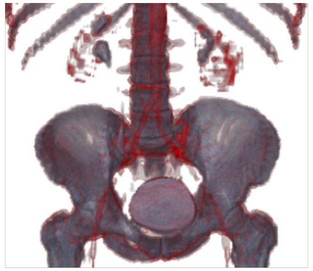
(c) Parametric [Sakhaee and Entezari, 2017]

(d) Nonparametric (Our contribution) [Athawale et al., 2020]

The teardrop function (ensemble dataset)

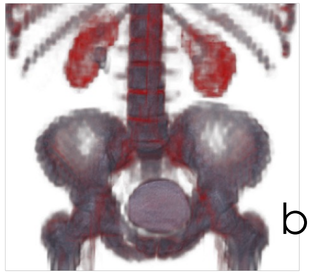
Visualization software: Voreen (http://voreen.uni-muenster.de)

(a) Ground truth (512x512x1559)



(b) Mean (64x64x195)

(c) Parametric (64x64x195) x 2 [Sakhaee and Entezari, 2017]



(d) Nonparametric (64x64x195) x b (our contribution) [Athawale et al., 2020]

b = # histogram bins

Osirix OBELIX dataset (http://medvis.org/datasets/)

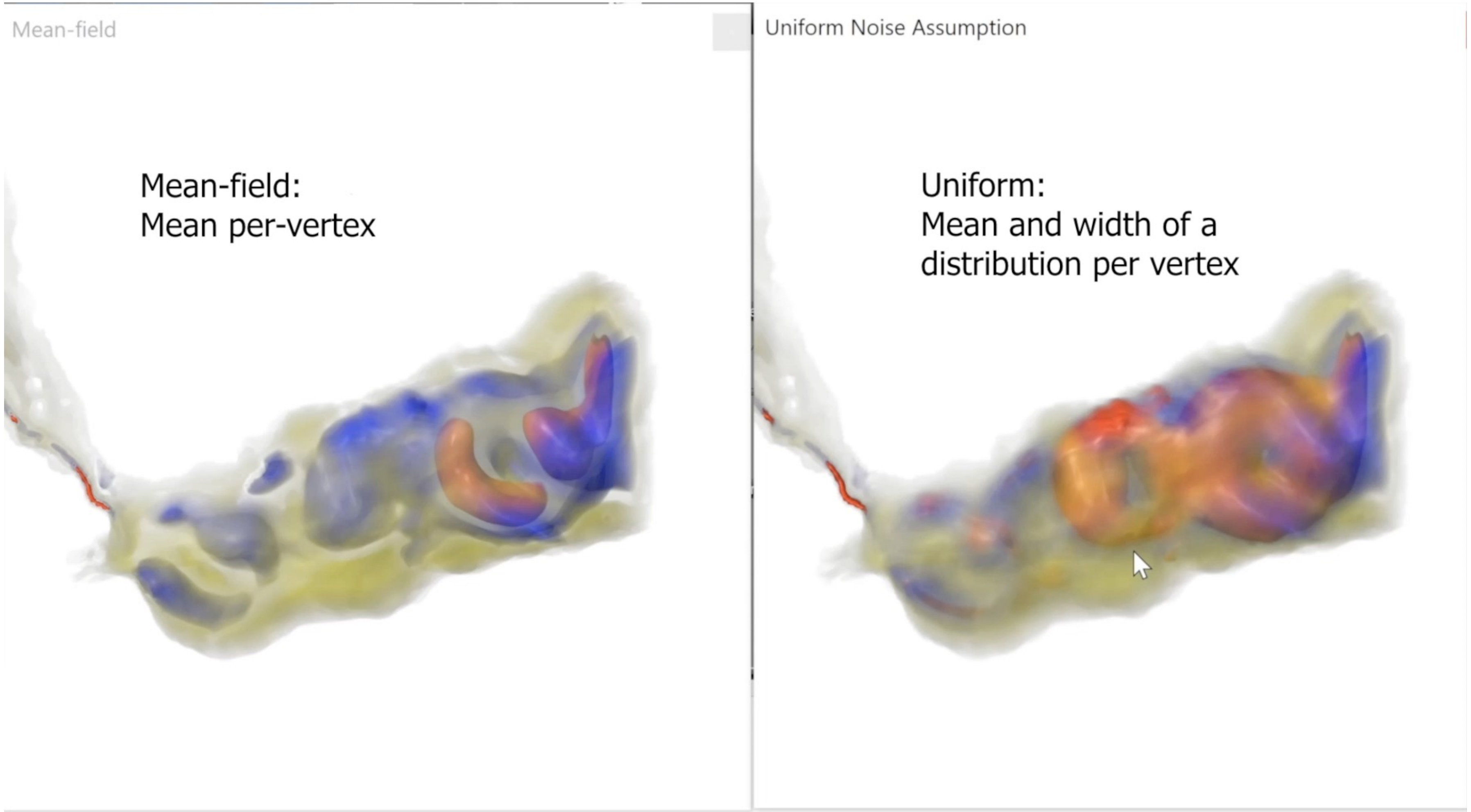
OAK RIDGE National Laboratory

The Red Sea Eddy Simulations

IEEE SciVis Contest 2020:

- Kaust Supercomputing Core Lab
- Large-scale eddy simulations (~1.5 TB)
- https://kaust-vislab.github.io/SciVis2020/

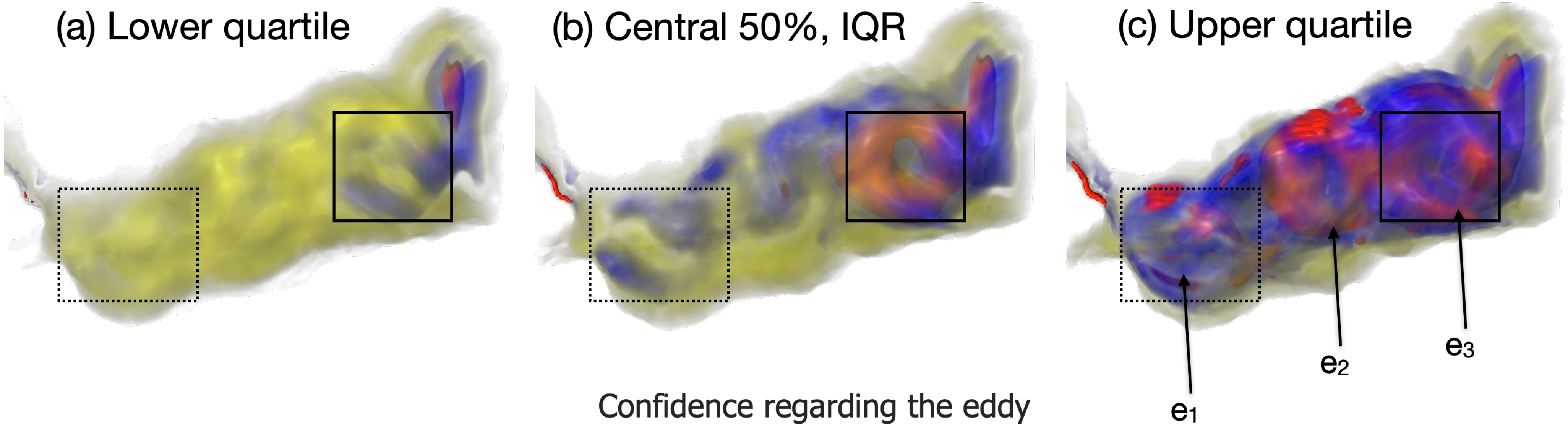
Uncertainty-Aware Volume Rendering: Interactive Exploration



OAK RIDGE National Laboratory

Quartile View: Uncertainty Visualization

Rendering uncertainty for 3D or high-dimensional data sets is an open research challenge.

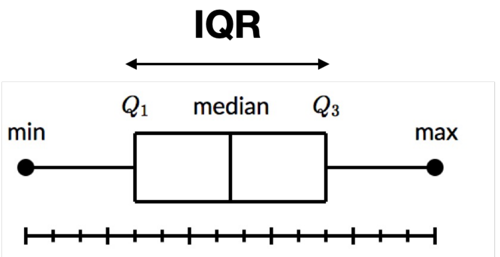


Confidence regarding the eddy presence/position:

e_1: Low

e_2: Moderate

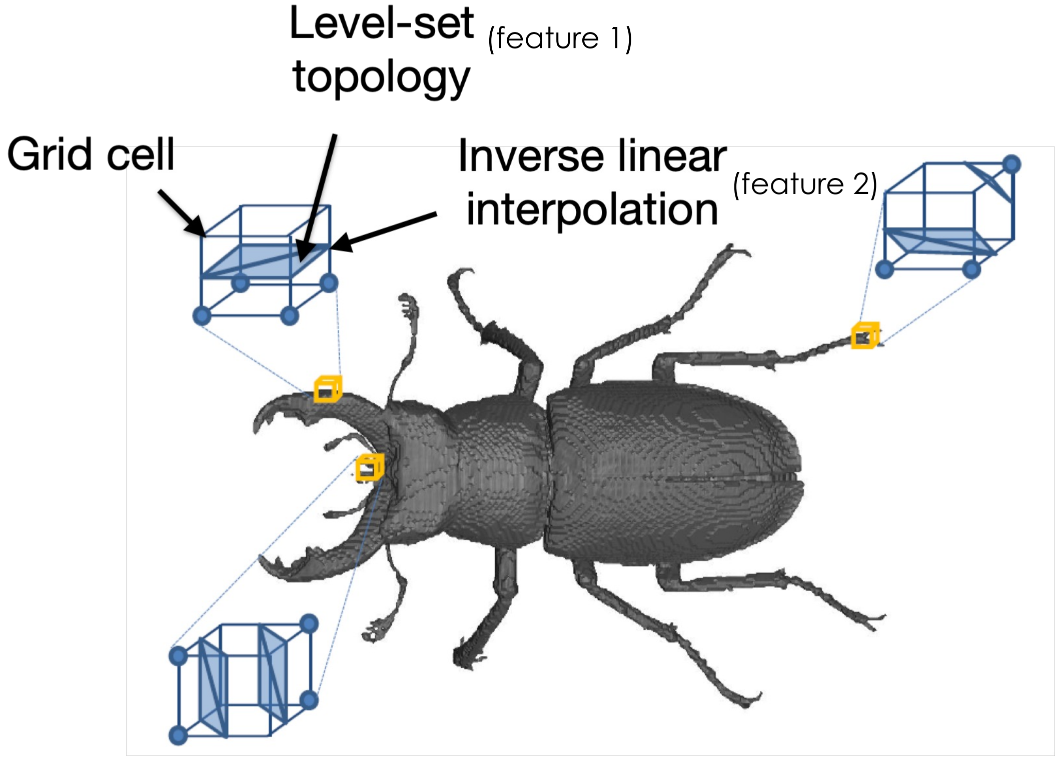
e_3: High



1D box plot

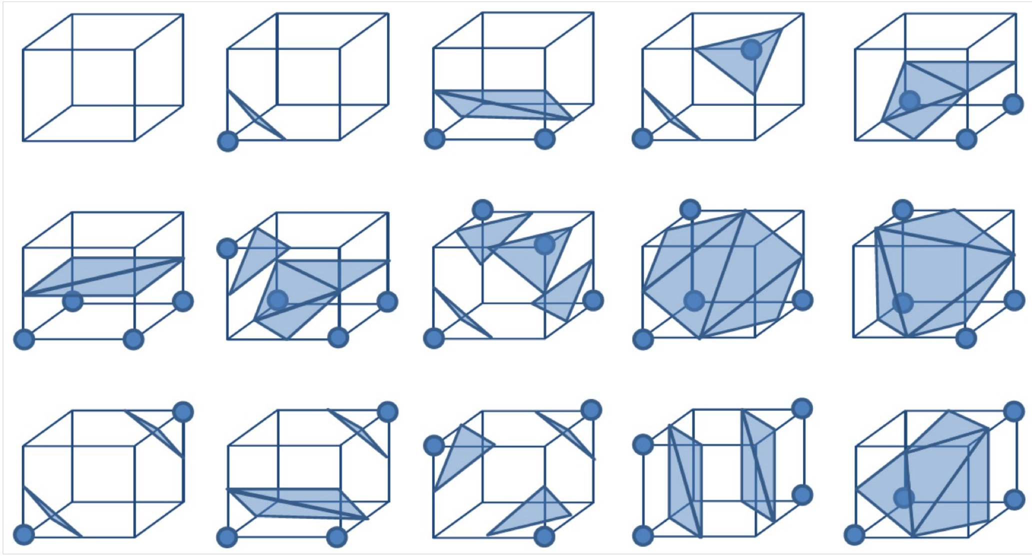
OAK RIDGE National Laboratory

Marching Cubes Algorithm [Lorensen and Cline, 1987]



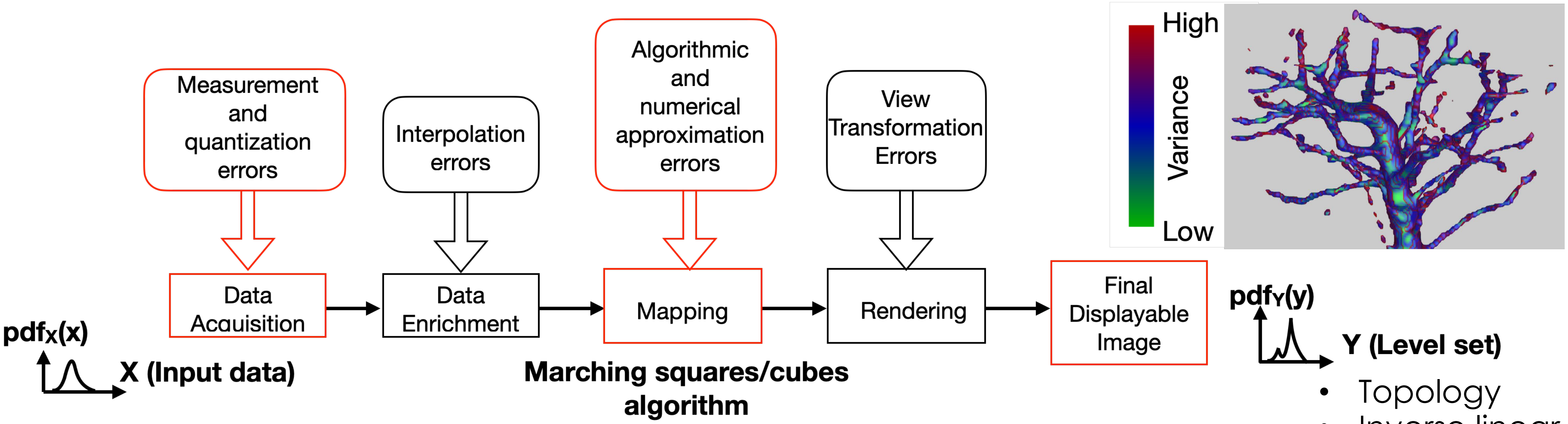
The Stag Beetle dataset is courtesy of Vienna University of Technology
https://www.cg.tuwien.ac.at/research/vis/datasets/

Topology cases



Level-Set Extraction from Uncertain Data (Analytical Approach)

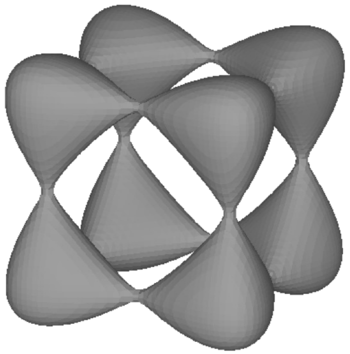
Marching cubes algorithm for certain [Lorensen and Cline, 1987] vs. uncertain data (our contribution!)



- Topology
- Inverse linear interpolation

Visualization of Inverse Linear Interpolation Uncertainty

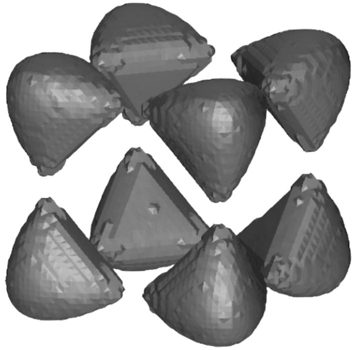
Ground truth



Memory consumption = 100*X

The tangle function [Knoll et al., 2009]

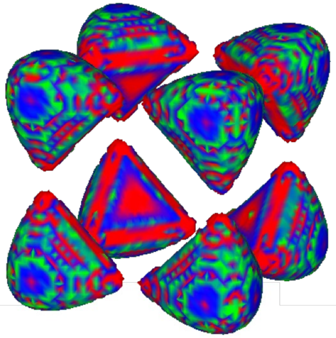
Mean



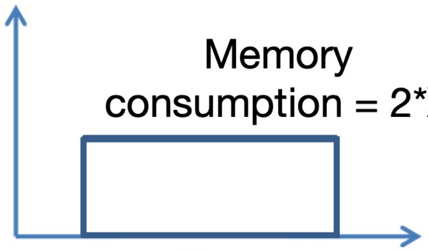
Memory consumption = X

Brick size = 100

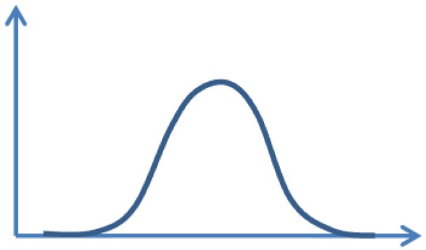
Parametric



Memory consumption = 2*X

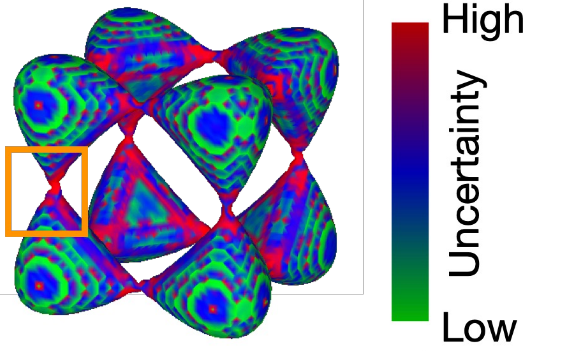


Uniform

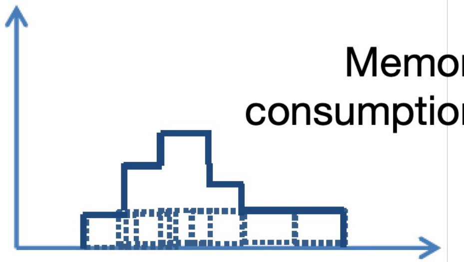


Gaussian

Nonparametric

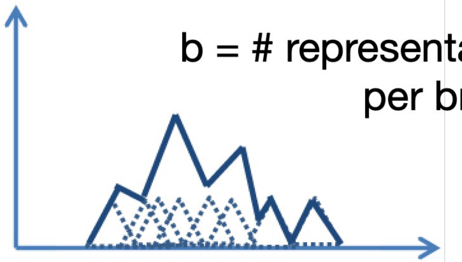


Memory consumption = b*X



Uniform kernel

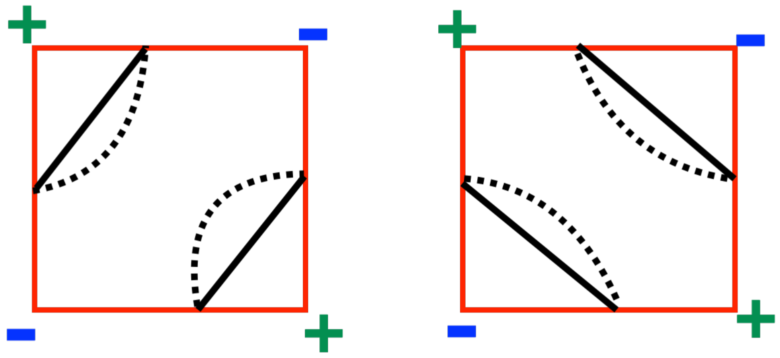
b = # representative samples per brick



Triangular kernel

[Athawale and Entezari, 2013; Athawale et al., 2016]

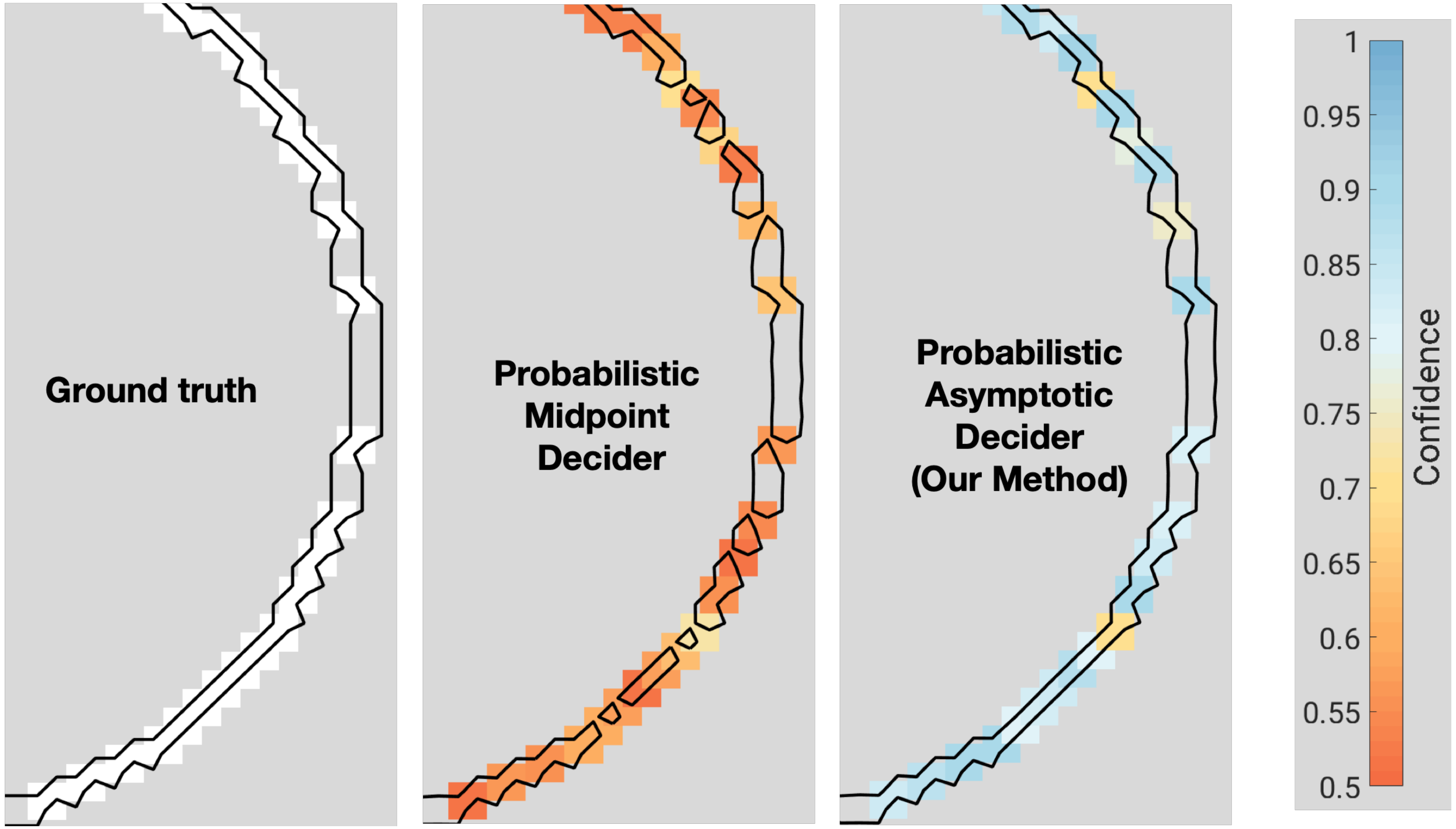
Uncertainty Visualization of Level-Sets (Topology)



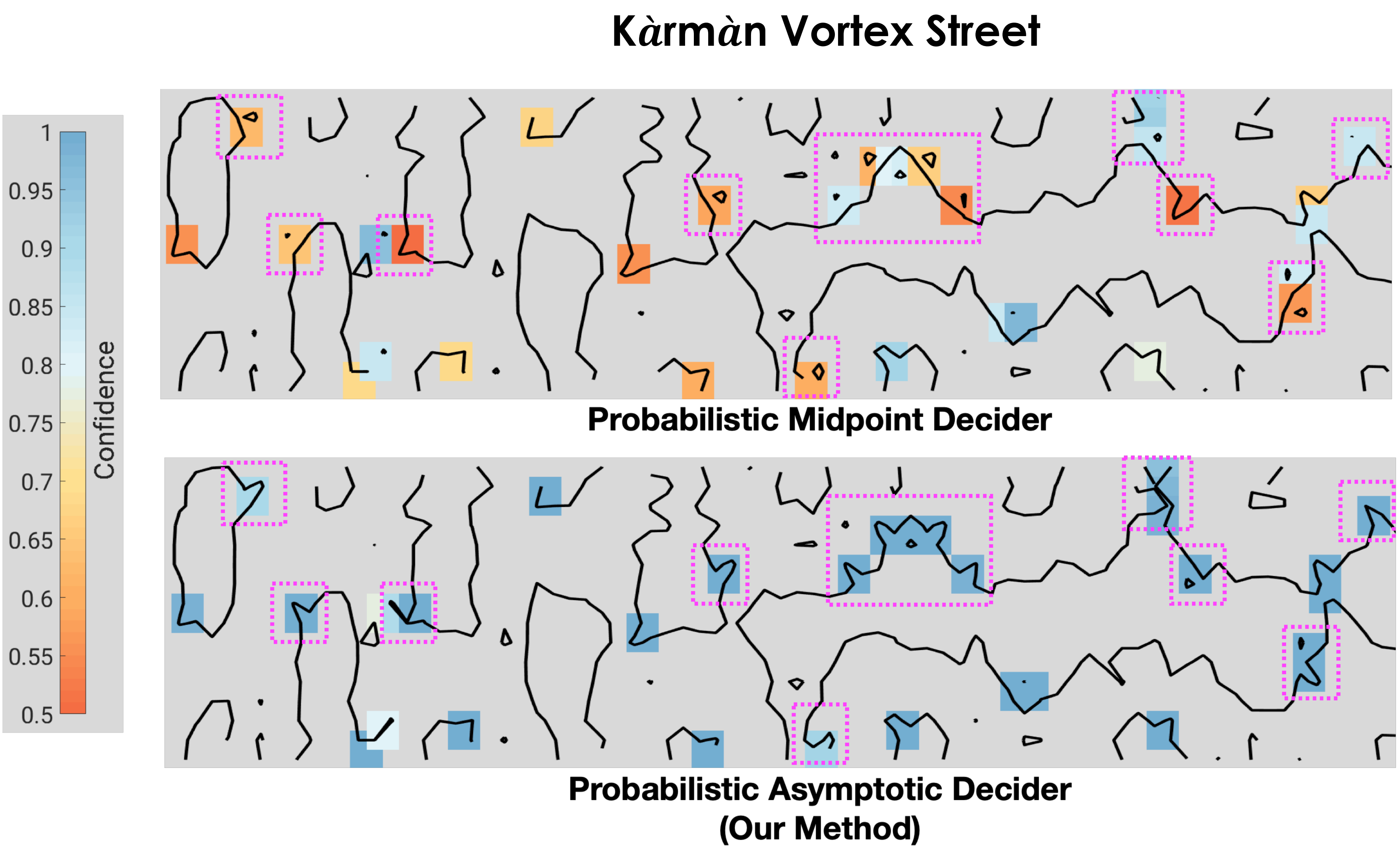
MSA ambiguous case

[Athawale and Johnson, 2018]

Concentric circles (synthetic data)

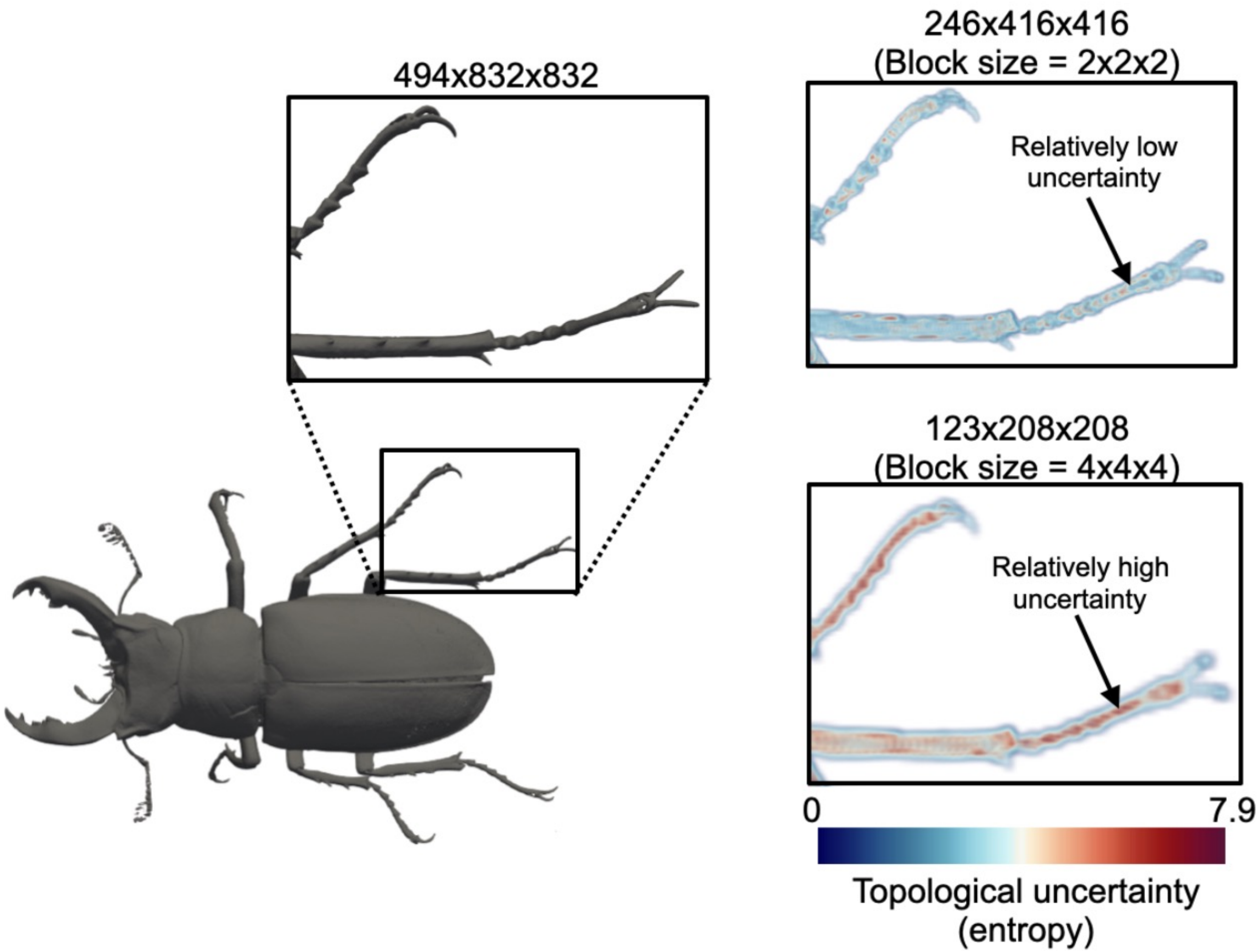


Uncertainty Visualization of Level-Sets (Topology)

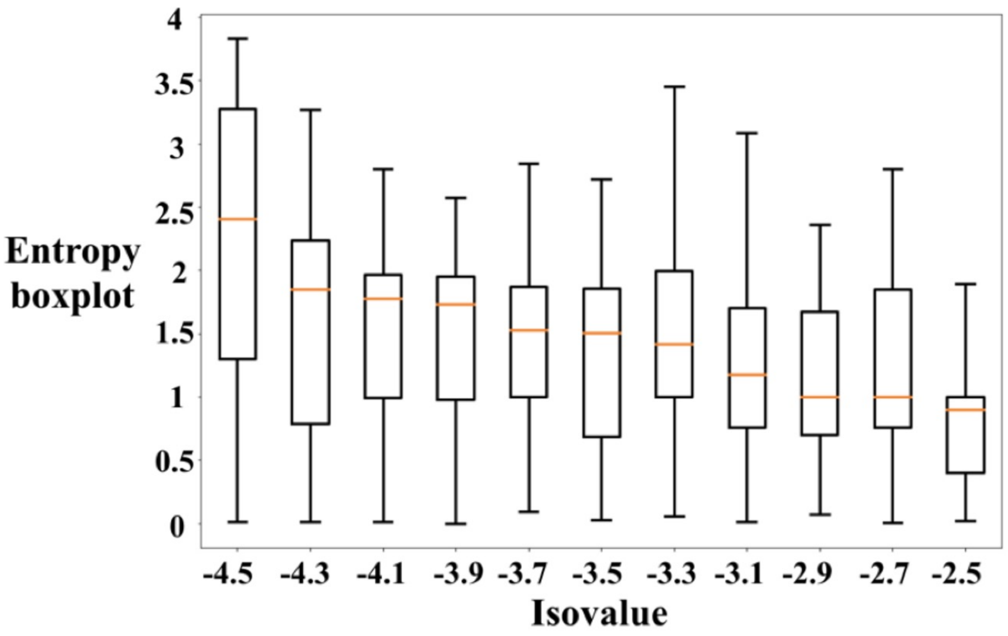


The flow simulation dataset is courtesy of the Gerris project [Popinet, 2003]

Uncertainty Visualization of Level-Sets (Topology)



Correlation between block size and uncertainty

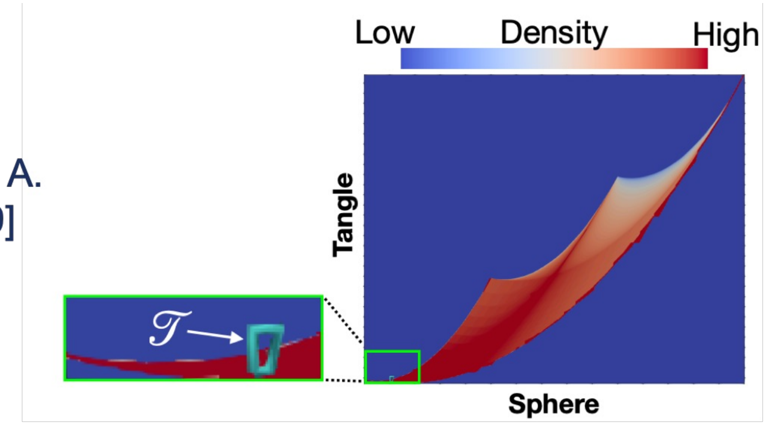


Correlation between isovalue and uncertainty

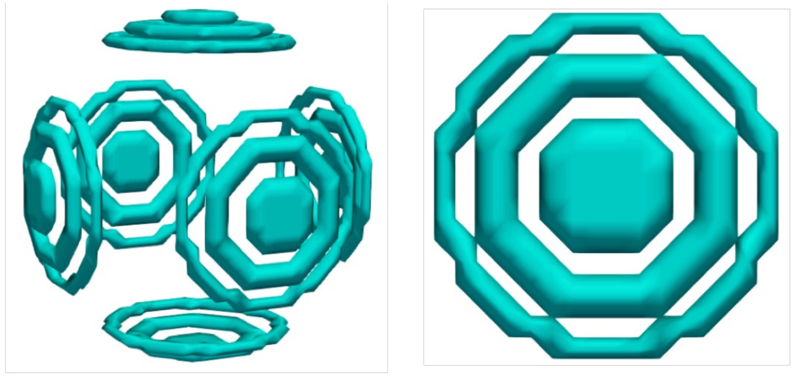
[T. Athawale, S. Sane, and C. R. Johnson, 2021]

Fiber Surface Extraction from Uncertain Data

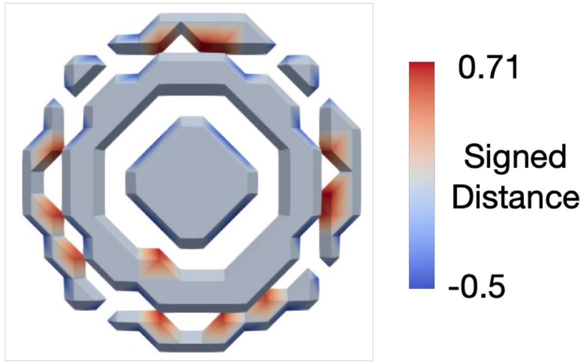
[Tangle function, A. Knoll et al., 2009]



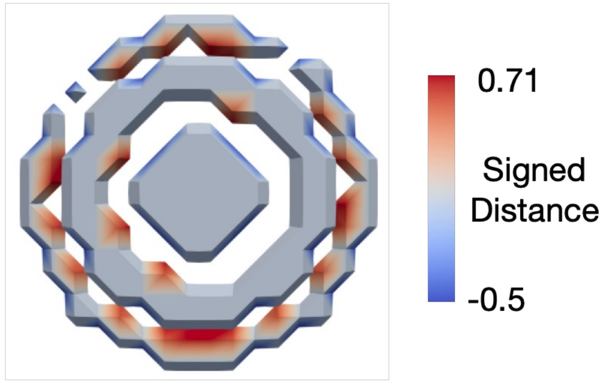
(a) Continuous scatterplot



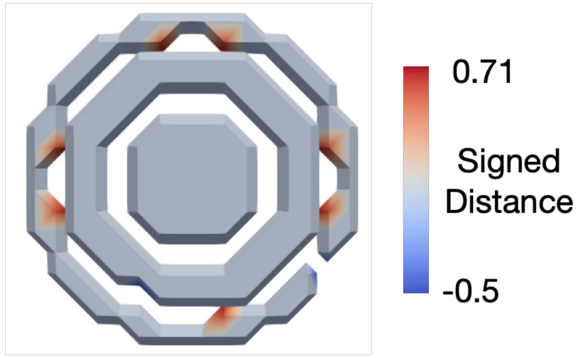
(b) Ground truth fiber surface



(c) Mean-field fiber surface



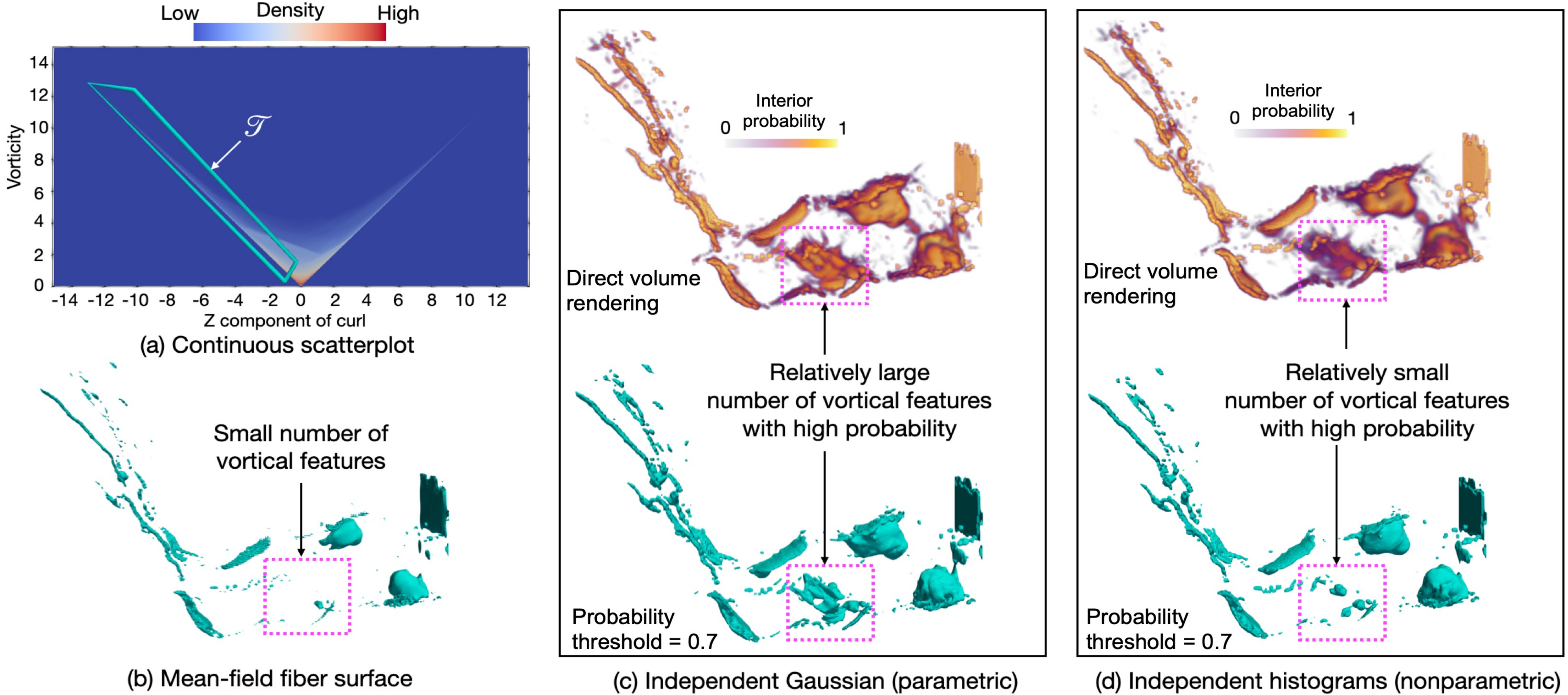
(d) Most probable fiber surface (parametric)



(e) Most probable fiber surface (nonparametric)

[T. M. Athawale, C. R. Johnson, S. Sane, and D. Pugmire, Accepted at IEEE VIS 2022]

Fiber Surface Extraction from Uncertain Data

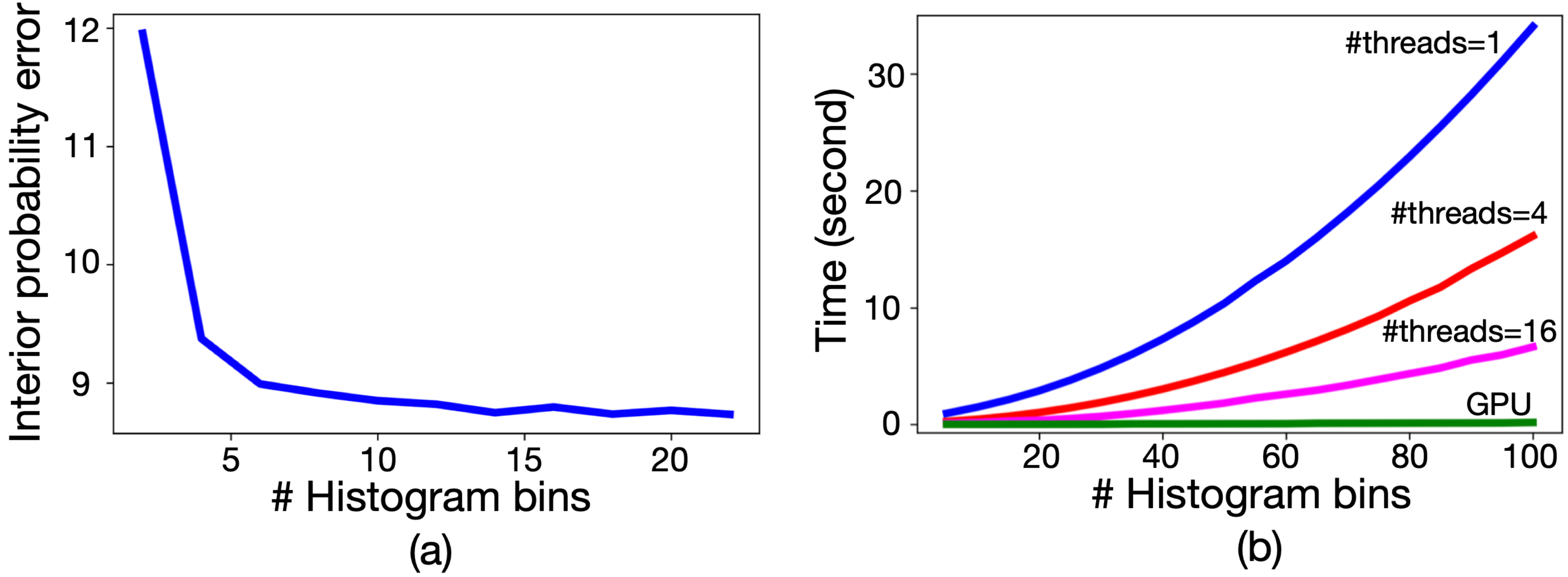


The dataset is downloaded from the IEEE 2020 SciVis Contest website (https://kaust-vislab.github.io/SciVis2020/)

[T. M. Athawale, C. R. Johnson, S. Sane, and D. Pugmire, Accepted at IEEE VIS 2022]

Parallel Implementation of Nonparametric Code

The GPU CUDA (NVIDIA V100 graphics card) and C++ openMP (Power9 CPU) implementations

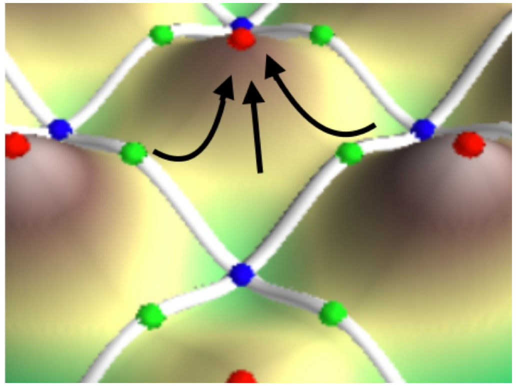


Accuracy (a) Vs. Timing (b) Curves

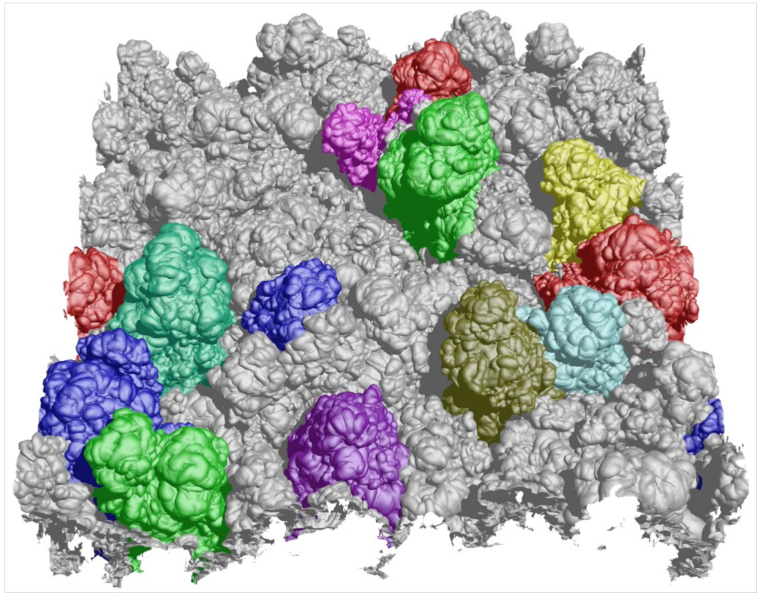
The computing resources are courtesy of the Summit Supercomputer at the Oak Ridge National Laboratory.

[T. M. Athawale, C. R. Johnson, S. Sane, and D. Pugmire, Accepted at IEEE VIS 2022]

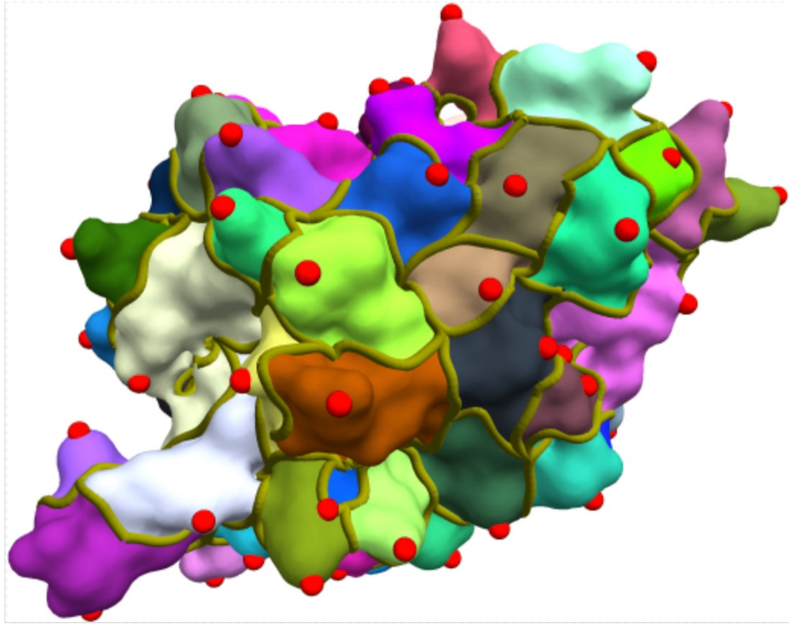
Morse Complex Visualizations



Topological descriptors of gradient flows of a scalar field

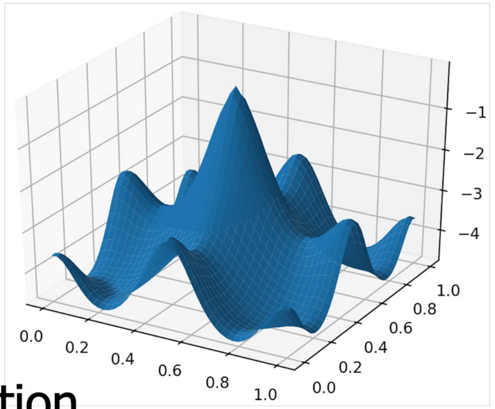


Understanding structure of turbulent mixing layers [Laney et al. 2006]



Segmenting molecular surfaces
[Natarajan et al., 2006]
[Shivashankar et al., 2012]

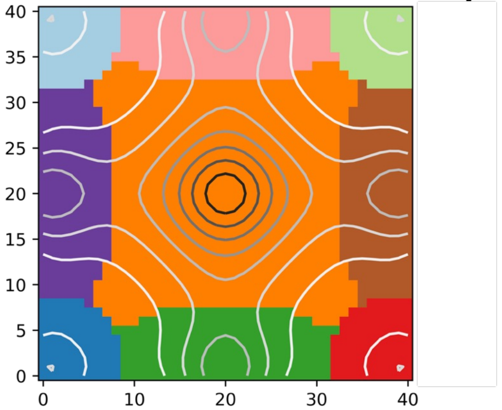
Effect of Noise on Morse Complexes



Ackley function
[Ackley, 1987]

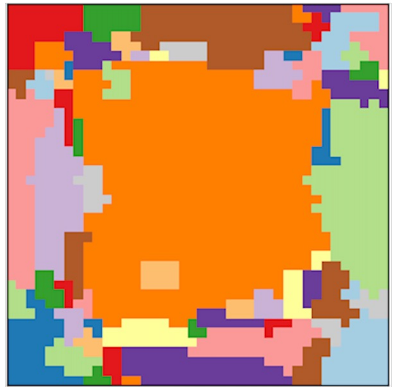
Morse complex extraction

Ground truth Morse complex

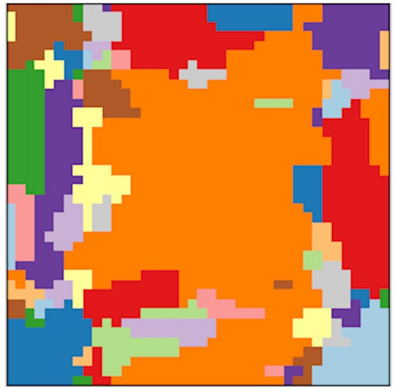


Mix noise and
extract Morse complex

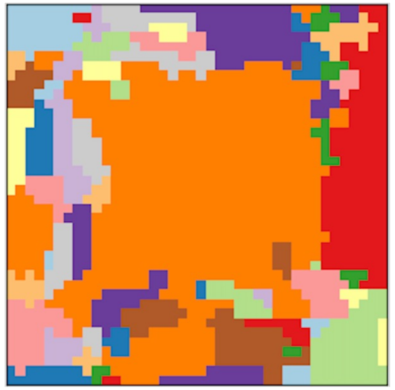
Ensemble member 1



Ensemble member 2

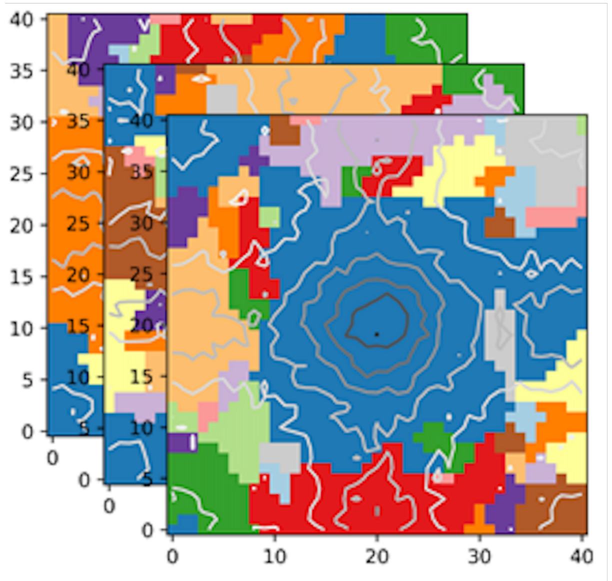


Ensemble member 3



OAK RIDGE
National Laboratory

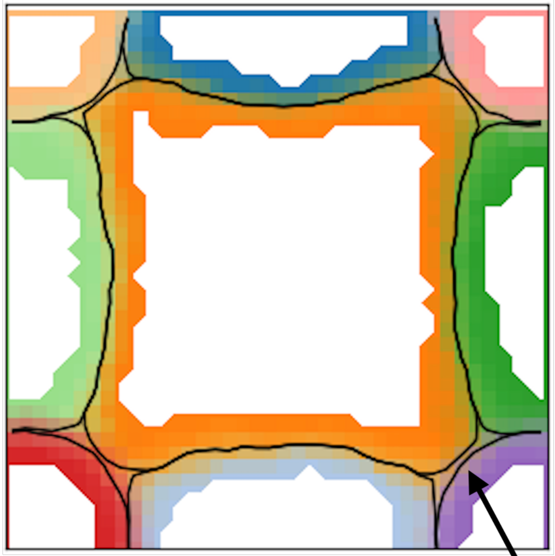
Morse Complex Uncertainty



Ensemble of Morse Complexes

Agreement Regions

+

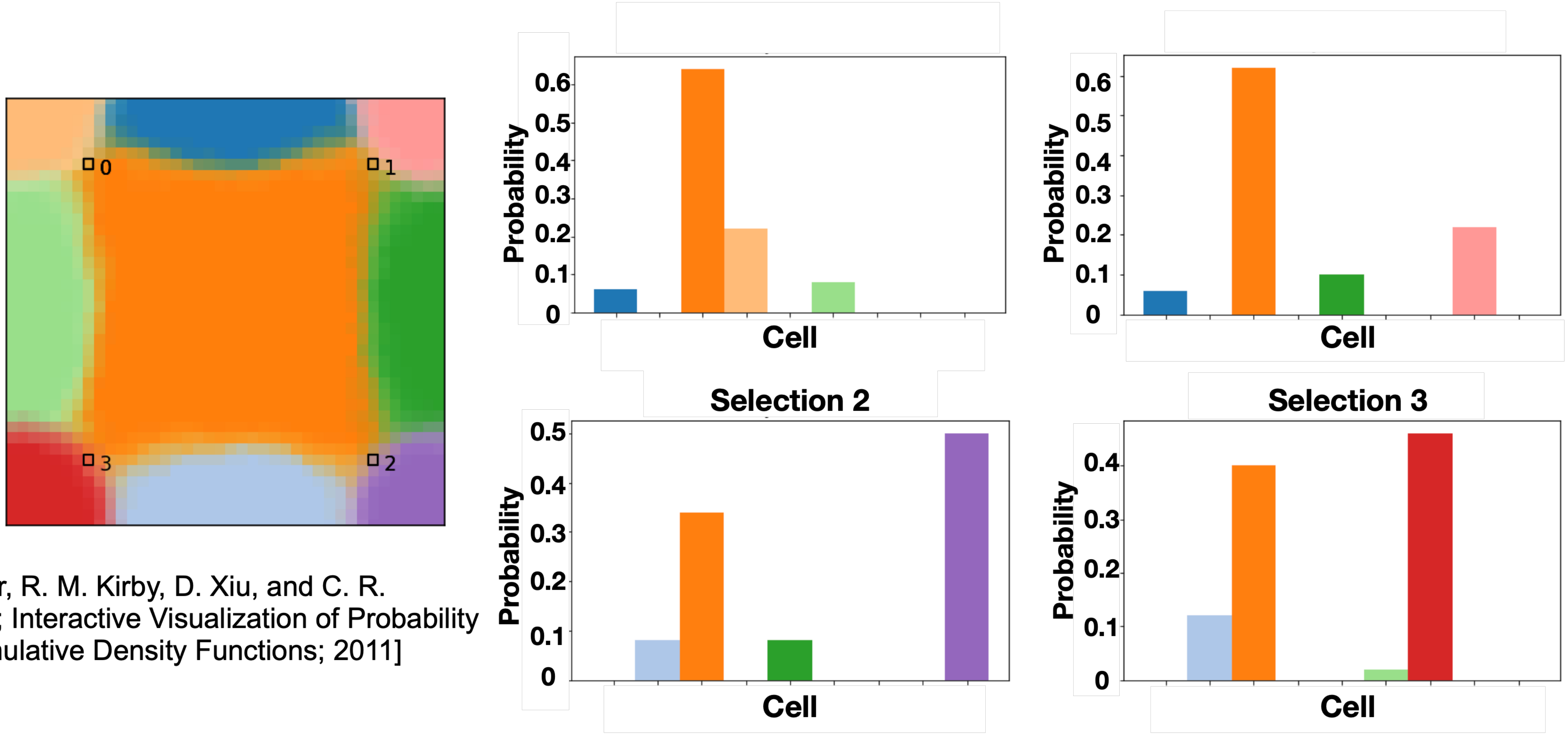


Uncertainty Regions

Expected boundaries

[T. M. Athawale, D. Maljovec, L. Yan, C. R. Johnson, V. Pascucci, and B. Wang, TVCG, 2022]

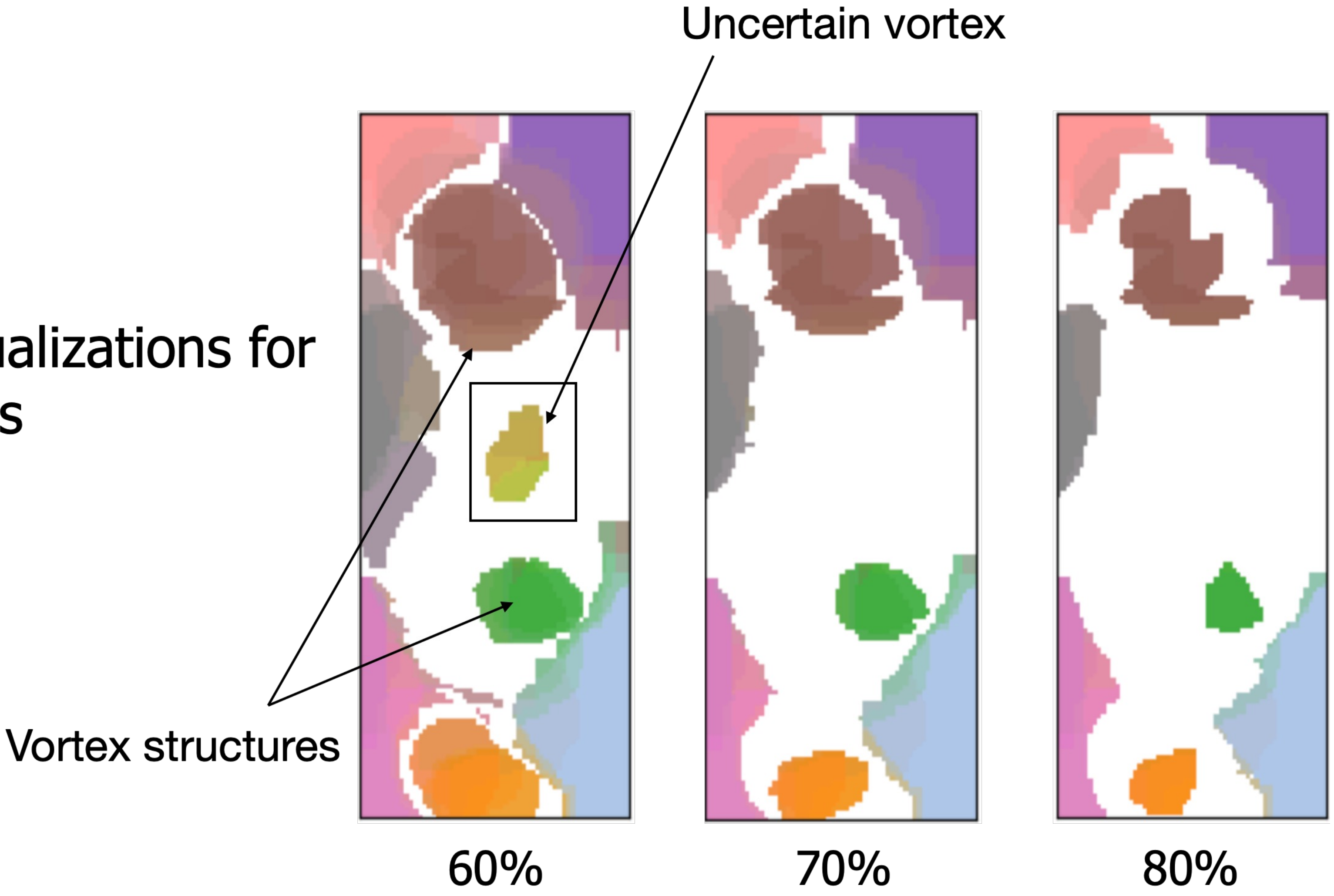
Interactive PDF Queries for Uncertain Regions



[K. Potter, R. M. Kirby, D. Xiu, and C. R. Johnson; Interactive Visualization of Probability and Cumulative Density Functions; 2011]

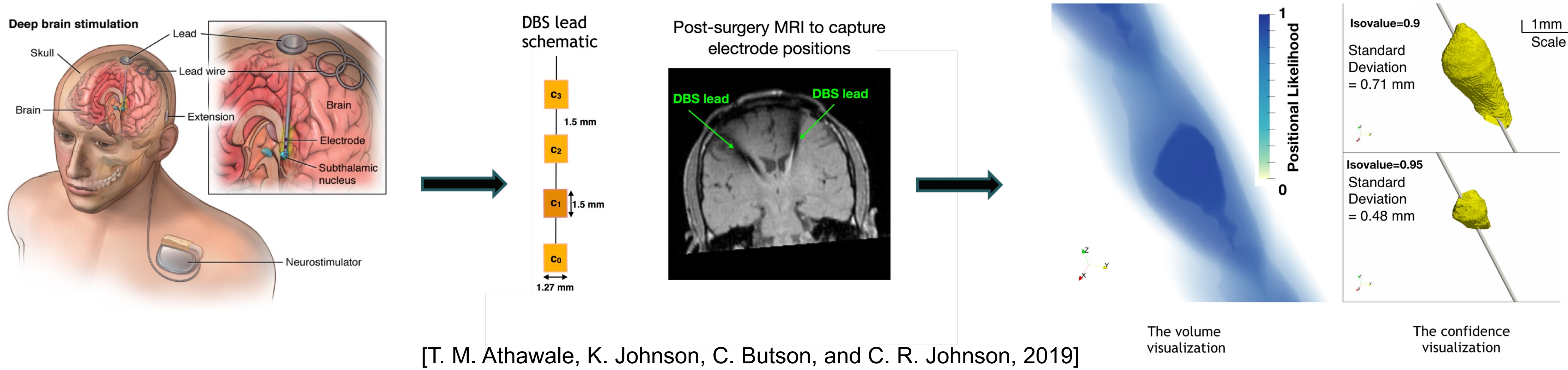
Agreement Exploration

Probabilistic visualizations for Morse complexes

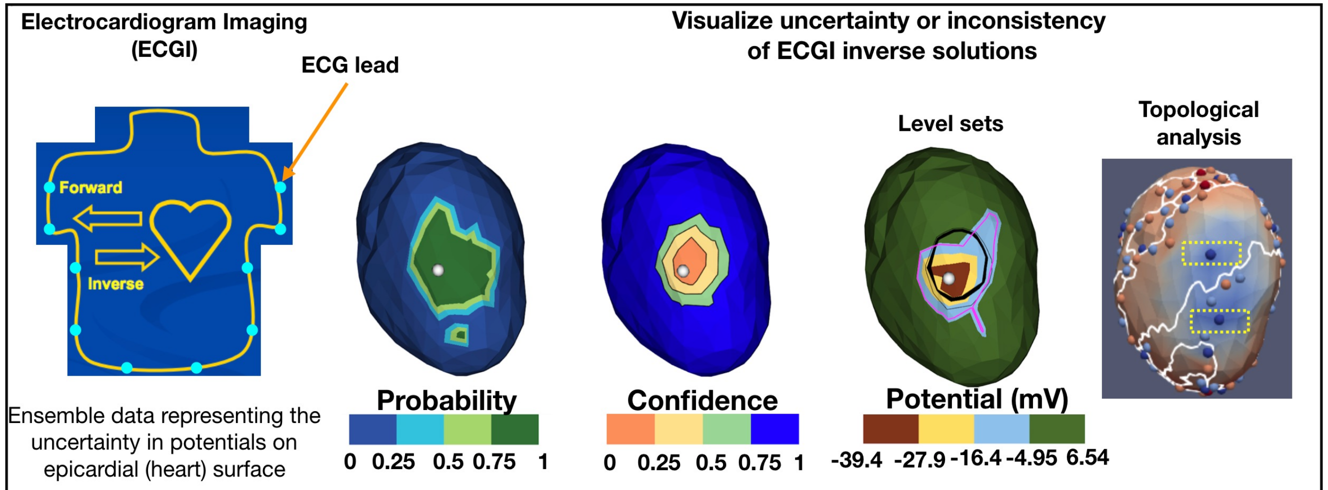


[T. M. Athawale, D. Maljovec, L. Yan, C. R. Johnson, V. Pascucci, and B. Wang, TVCG, 2022]

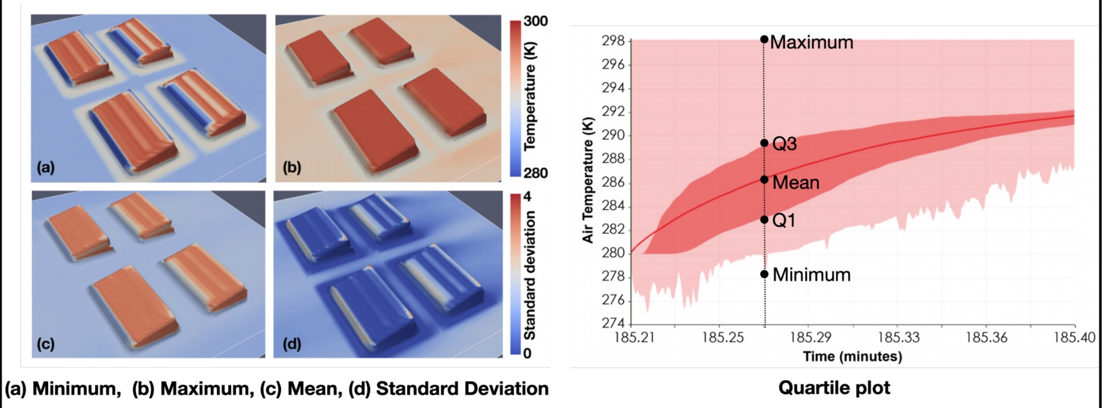
Uncertainty Visualization for Domain-Specific Data



[T. M. Athawale, K. Johnson, C. Butson, and C. R. Johnson, 2019]



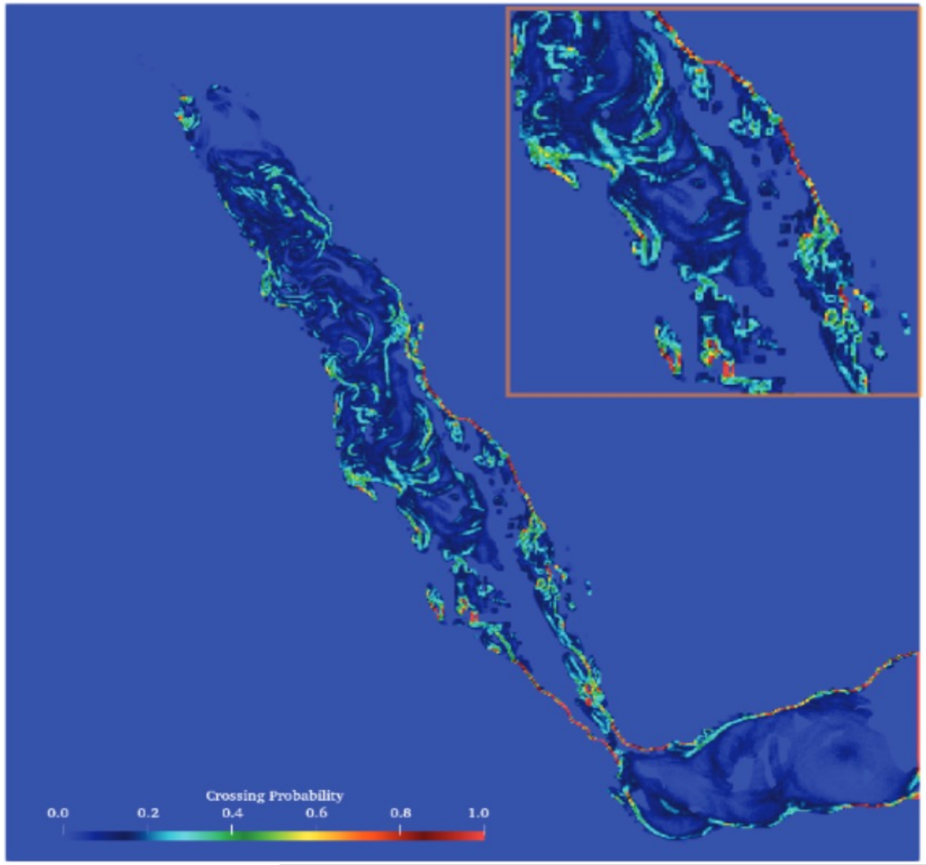
[D. Njeru, T. M. Athawale, J. France, and C. R. Johnson, 2022]



[T. M. Athawale, B. Stanislawski, S. Sane, and C. R. Johnson, 2021]

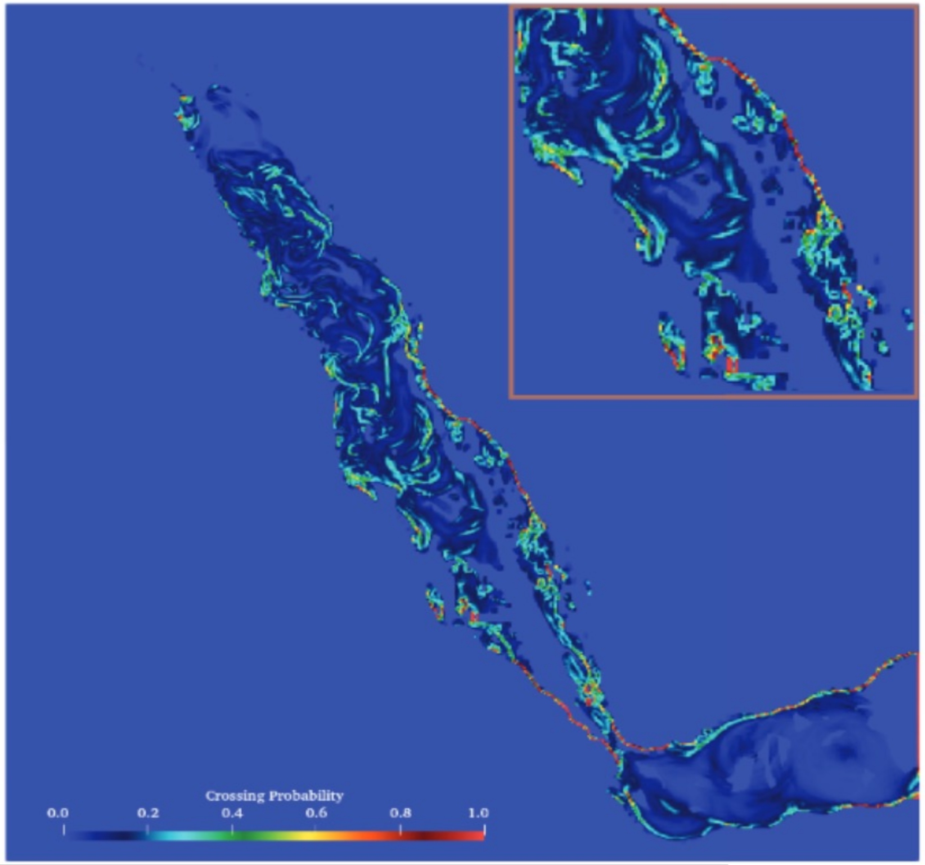
Future Work: Machine Learning for Uncertainty Visualization

Learn uncertainties pertinent to isosurfaces from a bunch of time steps and predict uncertainty for future time steps



Monte Carlo

[K. Pöthkow, B. Weber, and H.-C. Hege, "Probabilistic Marching Cubes", 2011]



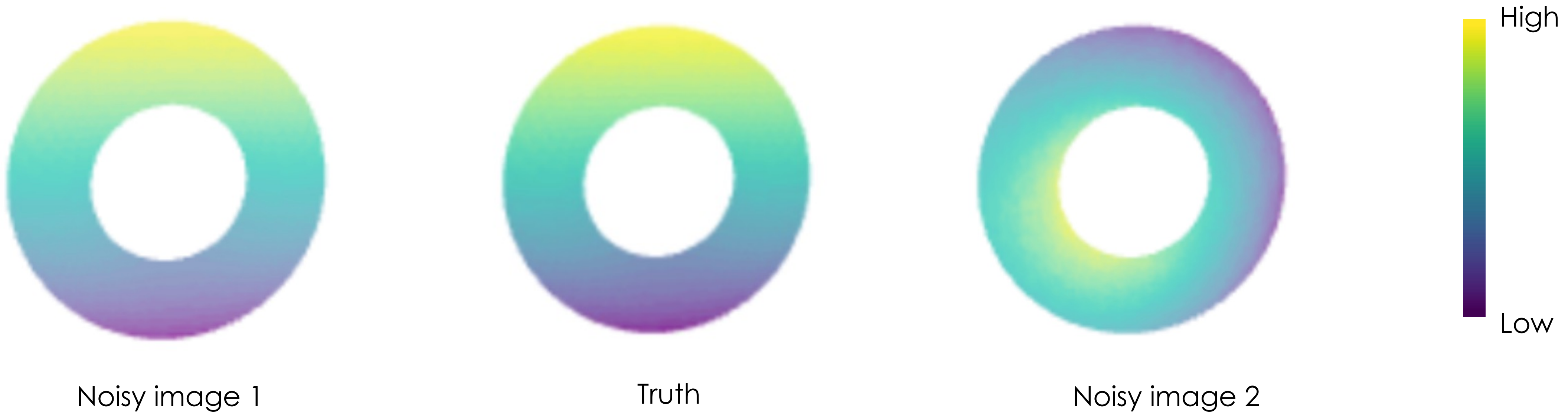
Machine Predicted
(170X faster)

[M. Han, T. M. Athawale, D. Pugmire, and C. R. Johnson, accepted at IEEE VIS 2022 short papers]

Future Work: Visualization and Decision-Making, A User Study

Under revision, [B. Triana, T. Kotha, T. M. Athawale, D. Pugmire, and P. Rosen]

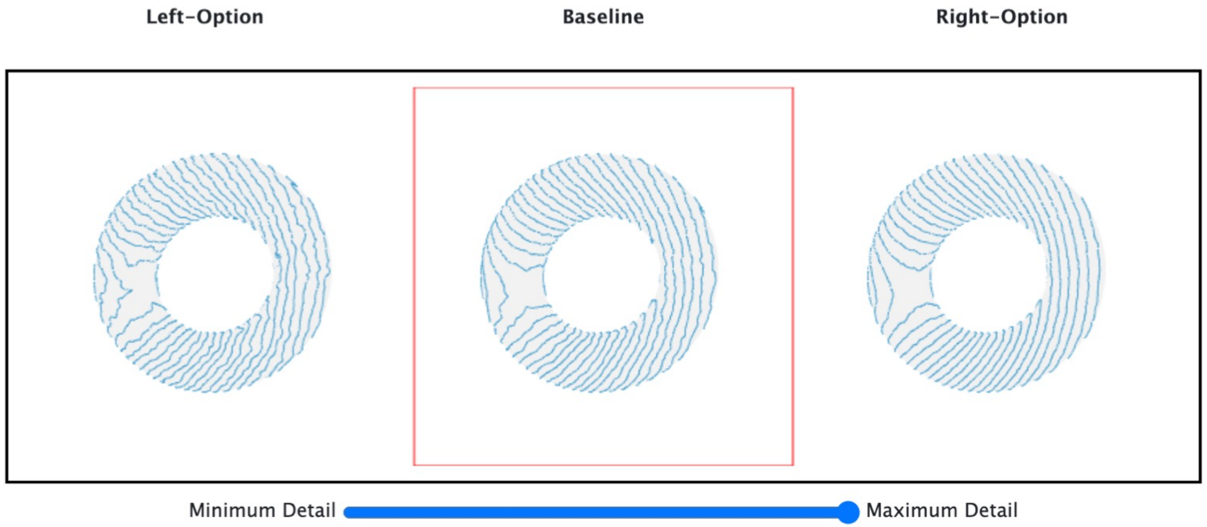
Which one of the following two noisy images is visually closer to the truth?



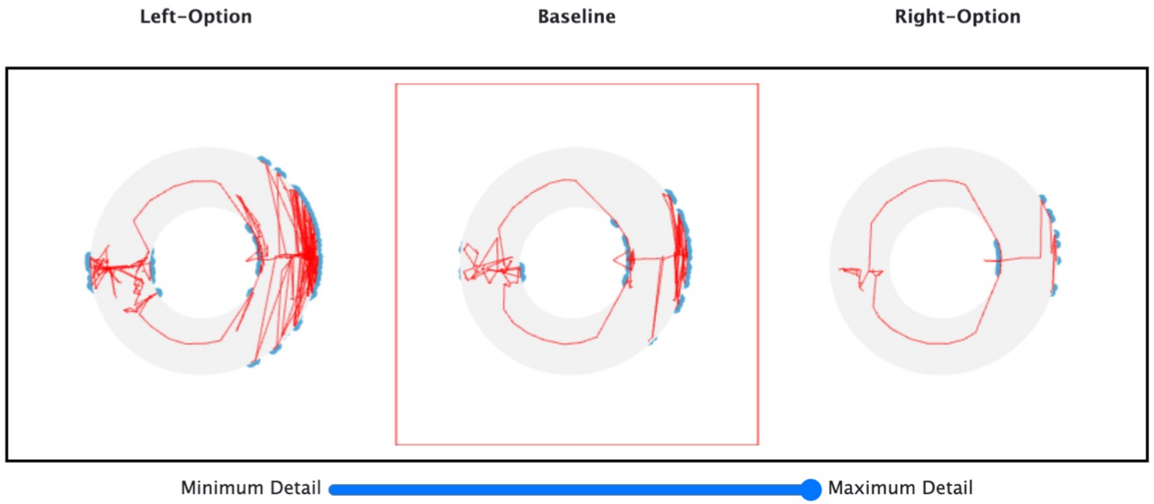
OAK RIDGE National Laboratory

Future Work: Visualization and Decision-Making, A User Study

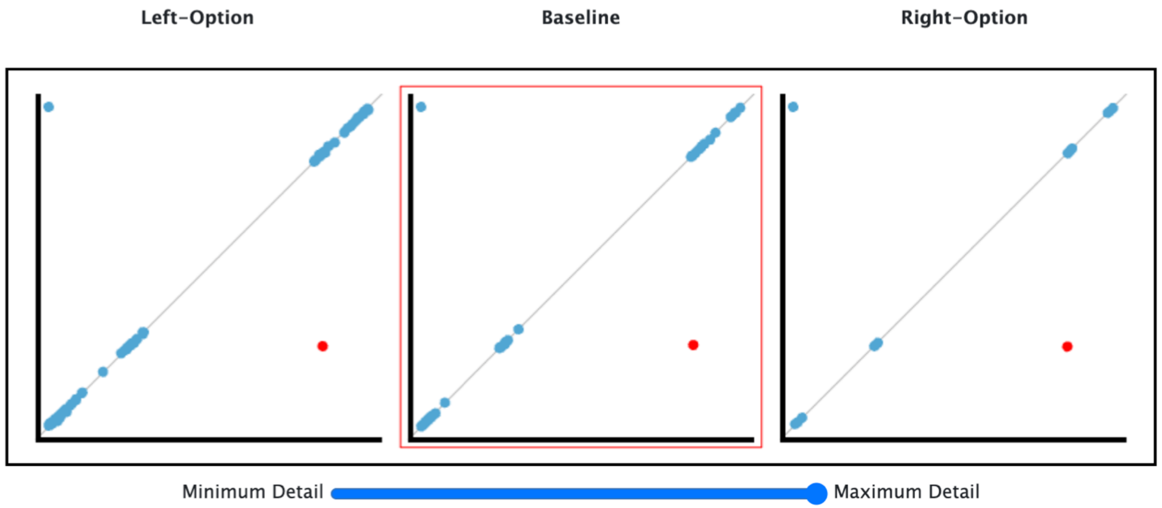
Understand the decision-making quality for topological visualizations under noise



Isocontours



Reeb graphs



Persistence diagrams

Oak Ridge National Laboratory

Open Research Challenges

- Uncertainty quantification for more visualization algorithms and high-dimensional data
- Devising uncertainty-aware decision frameworks to perform optimal algorithmic decisions, reduce uncertainty, and hence enhance quality of visualizations
- Improving interactivity of uncertainty visualization algorithms with machine learning models or GPU acceleration
- Understanding and managing tradeoffs between computational and memory requirements of uncertainty quantification techniques and timeliness of scientific applications
- Conducting community-wide surveys to assess cognition and decision-making quality

Thanks to Project Collaborators!

Dr. David Pugmire
(Group leader, Visualization group at ORNL)

Dr. Chris Johnson
(Postdoctoral advisor)

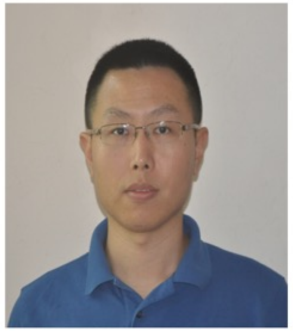
Dr. Alireza Entezari
(PhD advisor)

Dr. Chris Butson
(DBS project)

Dr. Bei Wang
(Morse complex project)

Dr. Valerio Pascucci
(Morse complex project)

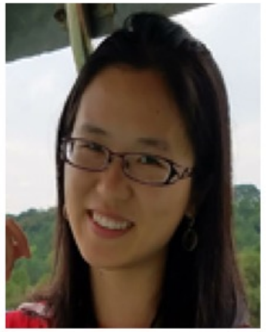
Dr. Paul Rosen
(Visualization sensitivity to noise project)



Dr. Bo Ma
(Direct volume rendering project)

Dr. Elham Sakhaee
(Direct volume rendering project)

Dr. Kara Johnson
(DBS project)



Dr. Lin Yan
(Morse complex project)

Dr. Dan Maljovec
(Morse complex project)

Dr. Sudhanshu Sane
(Multivariate uncertainty analysis project)

Dennis Njeru
(ECGI project)

Mengjiao Han
(Machine learning for uncertainty vis project)

Thank you!

This research is supported in part by by the the Scientific Discovery through Advanced Computing (SciDAC) program in the U.S. Department of Energy; NIH grants P41 GM103545-18 and R24 GM136986; the DOE grant DE-FE0031880; the Intel Graphics and Visualization Institutes of XeLLENCE; and the NSF grants IIS-1617101, IIS-1910733. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

For any questions, please contact me at:
Email: tushar.athawale@gmail.com
Personal website: http://tusharathawale.info