

# Computationally Efficient and accurate ML algorithms for PV finding

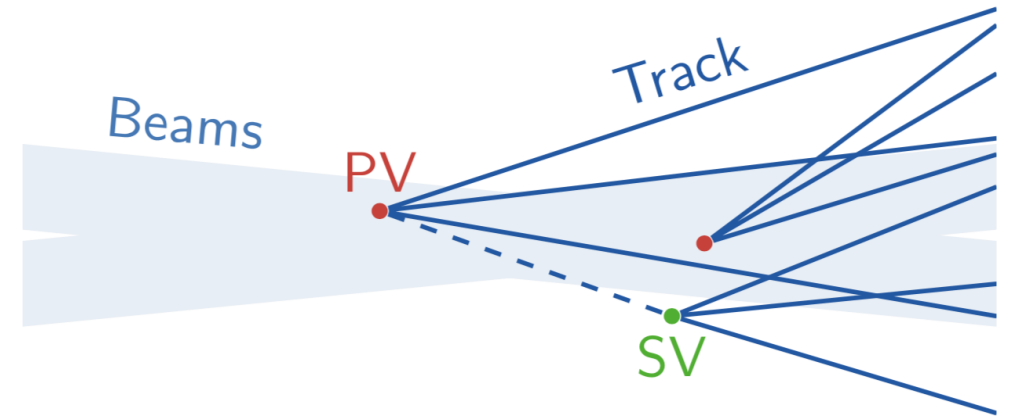
Gowtham Atluri

Mike Sokoloff

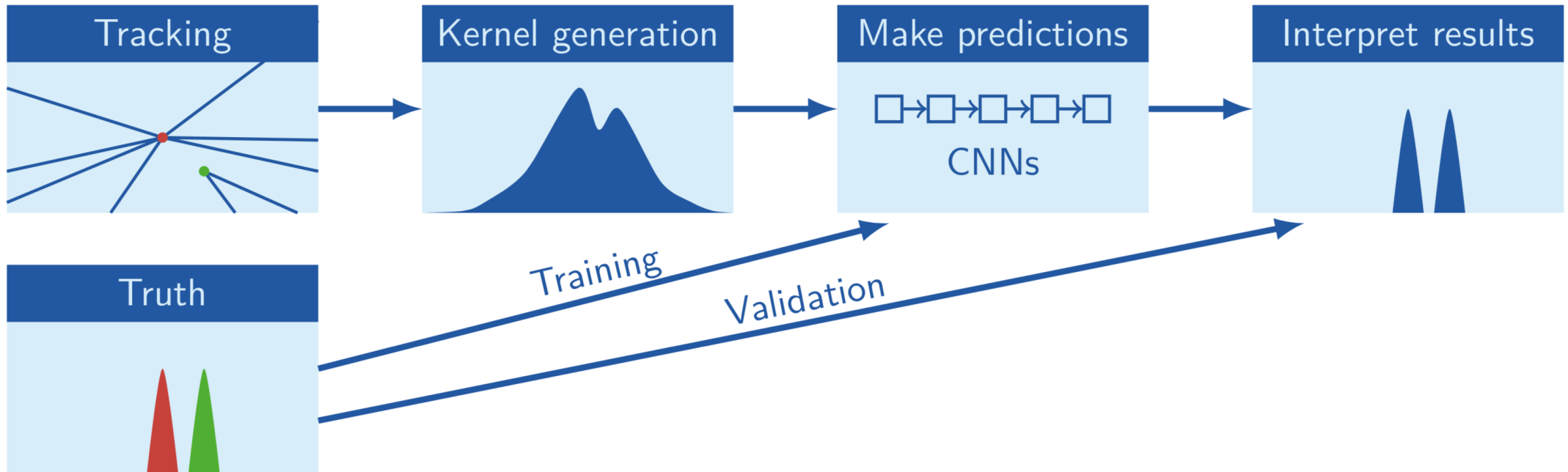


# PV finder

- LHCb upgrade
  - Much higher pileup
  - Need for faster algorithms
- Goal: Discovering PVs and SVs
- Use of ML approaches
  - Transform 3D data to 1D
  - 1D Convolutional neural nets
  - Highly parallelizable, GPU friendly
  - Room for interpretability

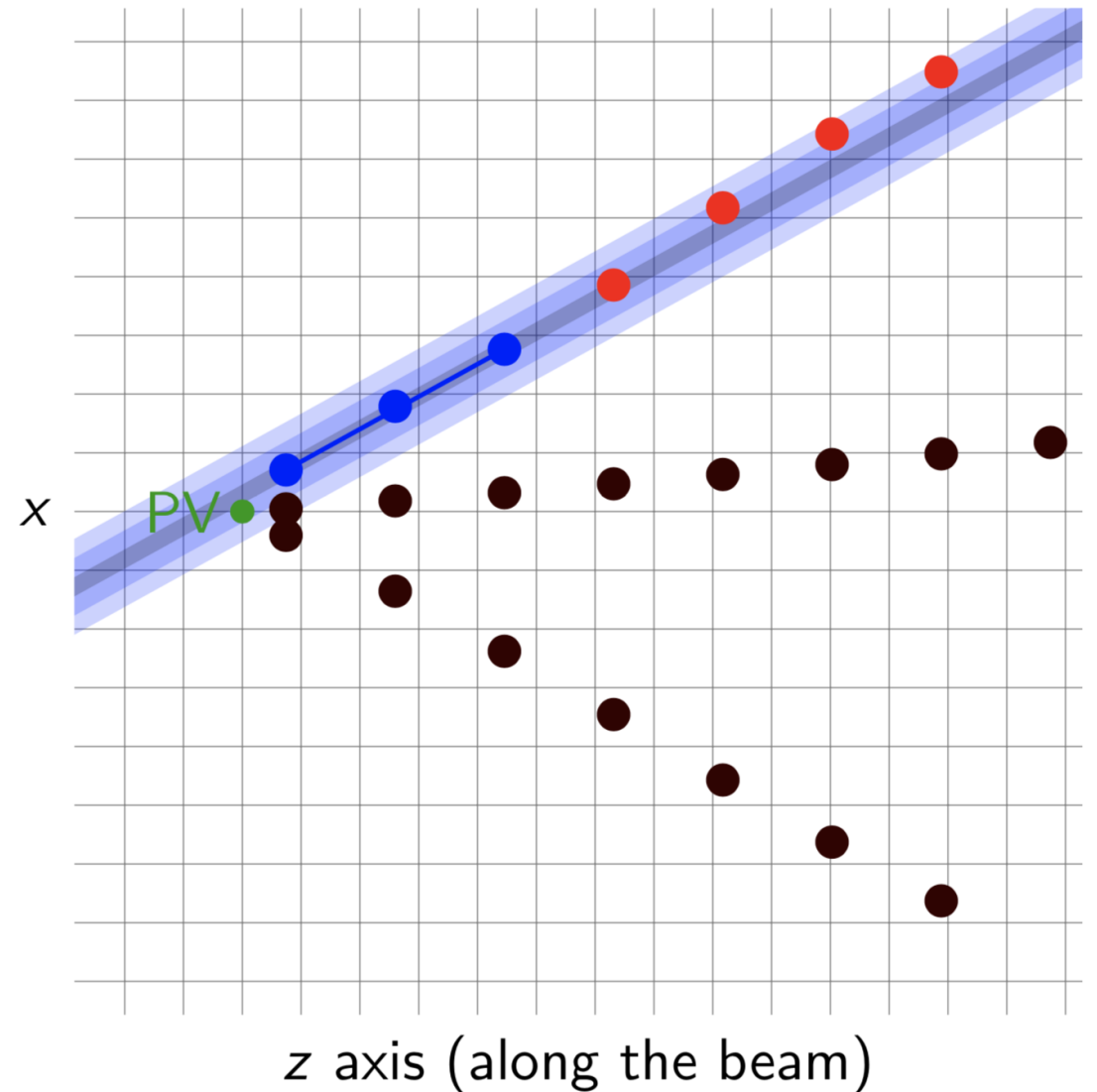


# Hybrid ML approach

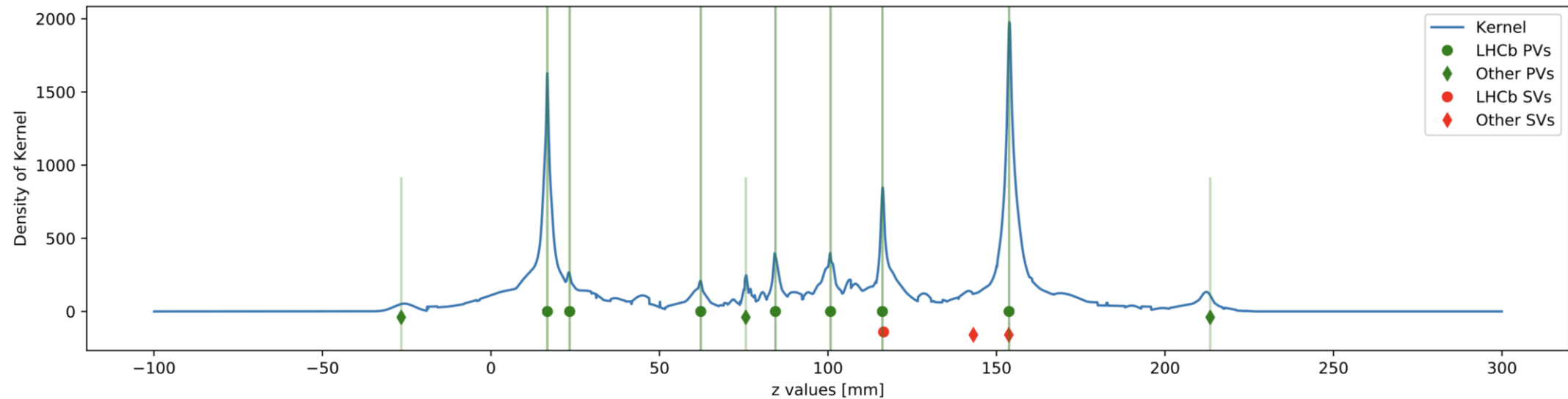


# Generating KDE

- Partial or full Tracking
- Fill each voxel center with a Gaussian PDF
- PDF for each track is combined
- Fill z histogram with maximum KDE in xy.



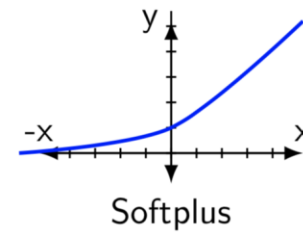
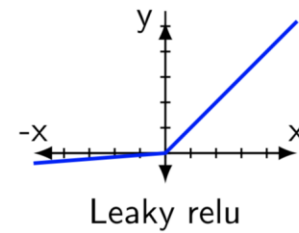
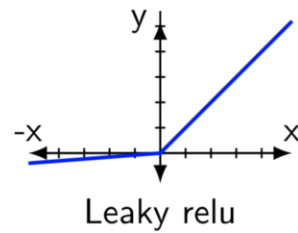
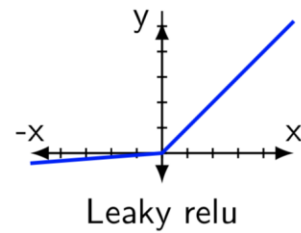
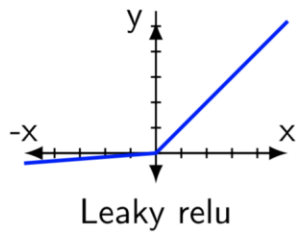
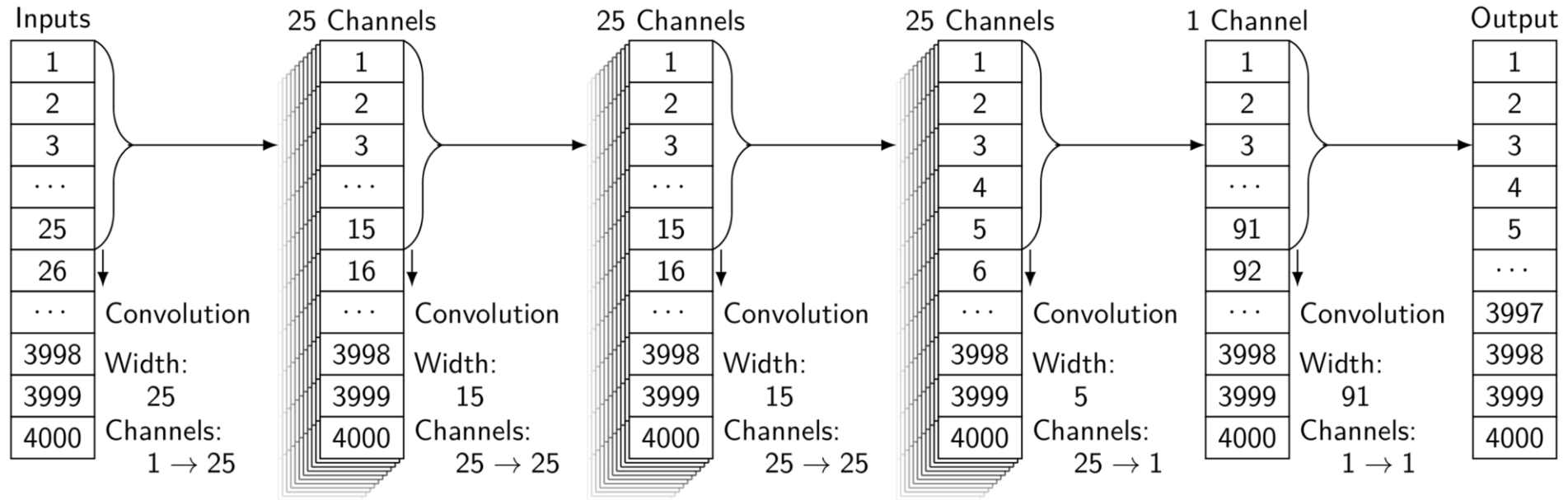
# Sample z KDE Histogram



## Challenges:

- True PVs may be offset from peak
- Peaks in close proximity interact

# CNNs developed by Cincinnati LHCb group



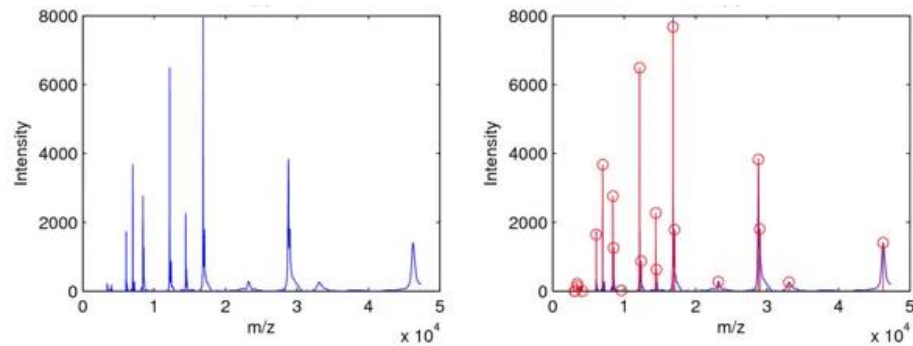
# Opportunities for CS contributions

- Characterize PVs where CNN is accurate
  - Identify the filters that capture PVs
  - Refine the CNN design for computational efficiency
- Characterize PVs where CNN is **not** accurate
  - Insufficient training examples?
  - Need for multiple models?
  - Ensemble models?

# Opportunities for CS contributions

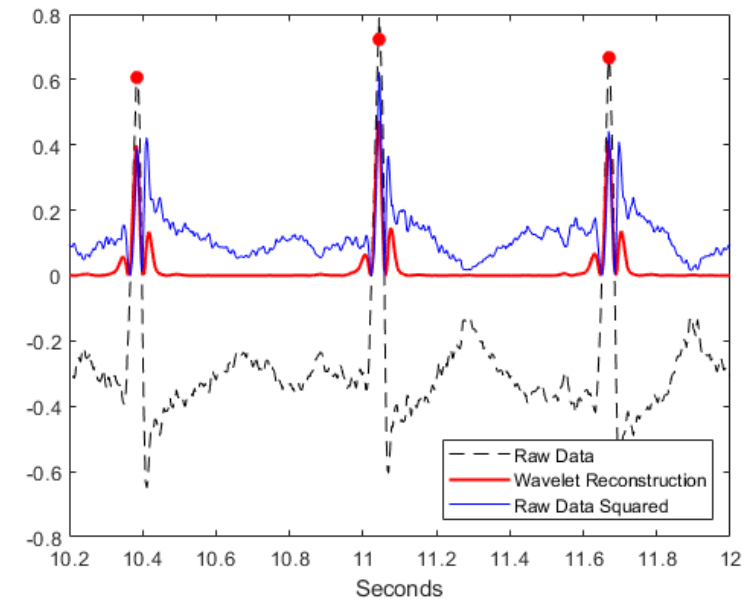
- Crosspollination of ideas with wider ML community
  - Heuristic one-pass algorithms
  - Hidden Markov Models, etc.

Mass Spectrometry [Yang et al. 2009]



Also in smart watches and mobile devices for step counting.

ECG Data



# Software deliverables

- Tools to visualize CNN filters
  - Visualizing the filters that are learned
  - Visualizing which filters contribute to correct predictions
  - Visualizing which filters contribute to incorrect predictions
- Tools to investigate CNN learning
  - Visualizing filters as a function of epoch#

# Other Outcomes

- Evaluation of alternative design choices for CNNs
  - One full 'z' histogram vs. separate 'chunks'
  - Impact of #layers, size of filters, etc.
  - Use of Recurrent NNs? When PVs are in close proximity.
- Evaluation of existing ML/heuristic approaches for peak detection
  - Heuristic computational approaches
  - HMMs
  - CNNs

Questions?