Variational autoencoders for noise reduction in industrial LLRF systems

Jonathan Edelen, Joshua Einstein-Curtis, Christopher Hall, Morgan Henderson (RadiaSoft)

Auralee Edelen, Jorge Diaz Cruz (SLAC)

October 24th 2023

LLRF 2023: Gyeongju, Republic of Korea

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Accelerator R&D and Production award number DE-SC0023641

Industrial Accelerators

Security and Defense
- Directed energy
- Single effects testing

Imaging
- X-ray sources
- Gamma-ray sources
- electron microscopy

Manufacturing
- Polymer treatment
- Industrial Curing
- Welding
- Ion Implantation

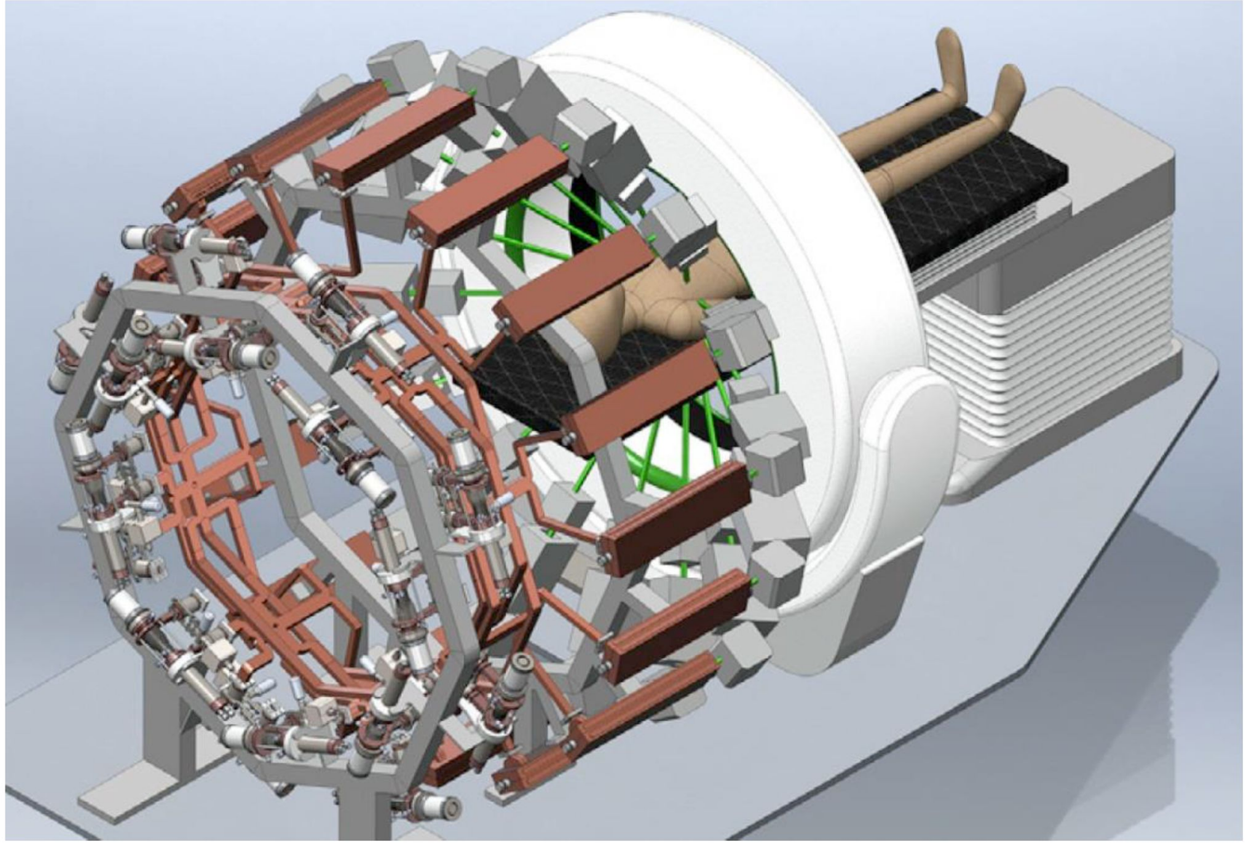
Medical
- Proton Therapy
- E-beam therapy
- X-ray therapy

Sterilization
- Medical device sterilization
- Food irradiation
- Wastewater treatment

- Legacy systems lack complexity, automation is straightforward
  - Single RF structure controlled with a PLL or similar
  - loose beam tolerances
- Next generation of industrial systems are increasing in complexity
  - Synchronization of multiple structures for higher energy applications
  - Tighter tolerances on output beams for emerging applications

Case study: FLASH Radiotherapy

- Delivery of doses at significantly higher rates
  - result in less damage to normal tissue
  - reduce reliance on mechanical moving parts, such as multi-leaf collimators
- SLAC led developments
  - PHASER (pluridirectional high-energy agile scanning electronic radiotherapy)
  - VHE (very high energy)
- The RF system is much more complex compared to single cavity legacy systems
  - synchronization between structures
  - noise / variation in individual cavity signals
- Beam control via steering magnets



Opportunities for Industrial Accelerators

- Focus areas for improving controls
  - Improvement of feedback systems for beam stabilization
  - Automation of startup routines (calibrations and synchronization)
  - Improvement of signal quality for RF systems

Opportunities for Industrial Accelerators

- Focus areas for improving controls
  - Improvement of feedback systems for beam stabilization
  - Automation of startup routines (calibrations and synchronization)
  - **Improvement of signal quality for RF systems**

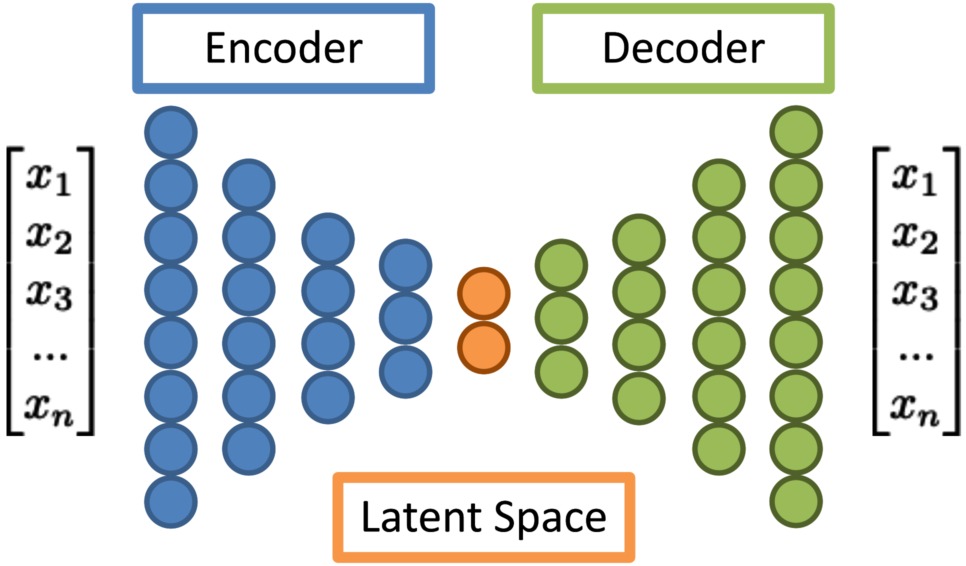
Opportunities for Industrial Accelerators

- Focus areas for improving controls
  - Improvement of feedback systems for beam stabilization
  - Automation of startup routines (calibrations and synchronization)
  - **Improvement of signal quality for RF systems**

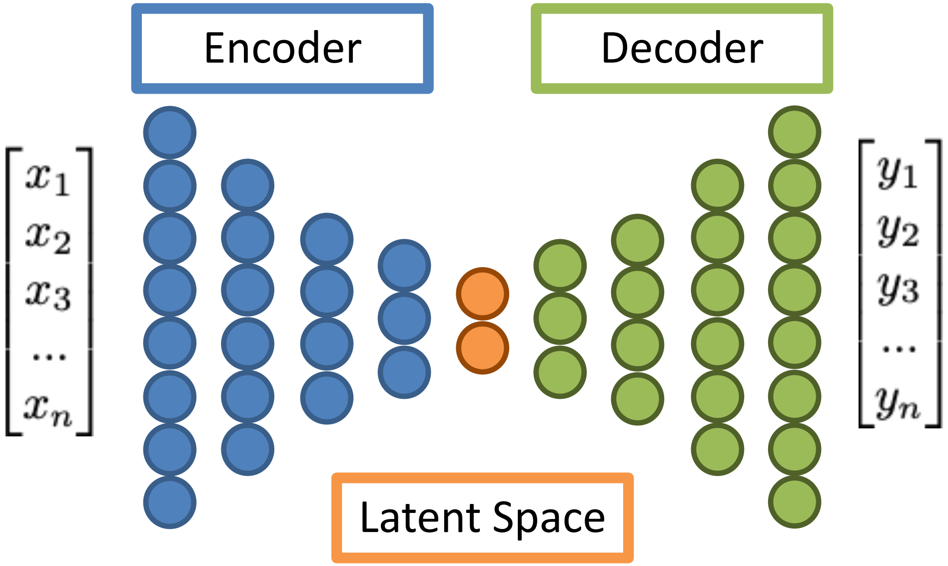
Autoencoders for Noise Reduction in RF Signals

What is an Autoencoder?

- Autoencoder
  - Type of neural network
  - Transforms data into a latent space and performs a reconstruction
  - Inputs and Outputs are the same: i.e. it is an identity transformation for a given dataset

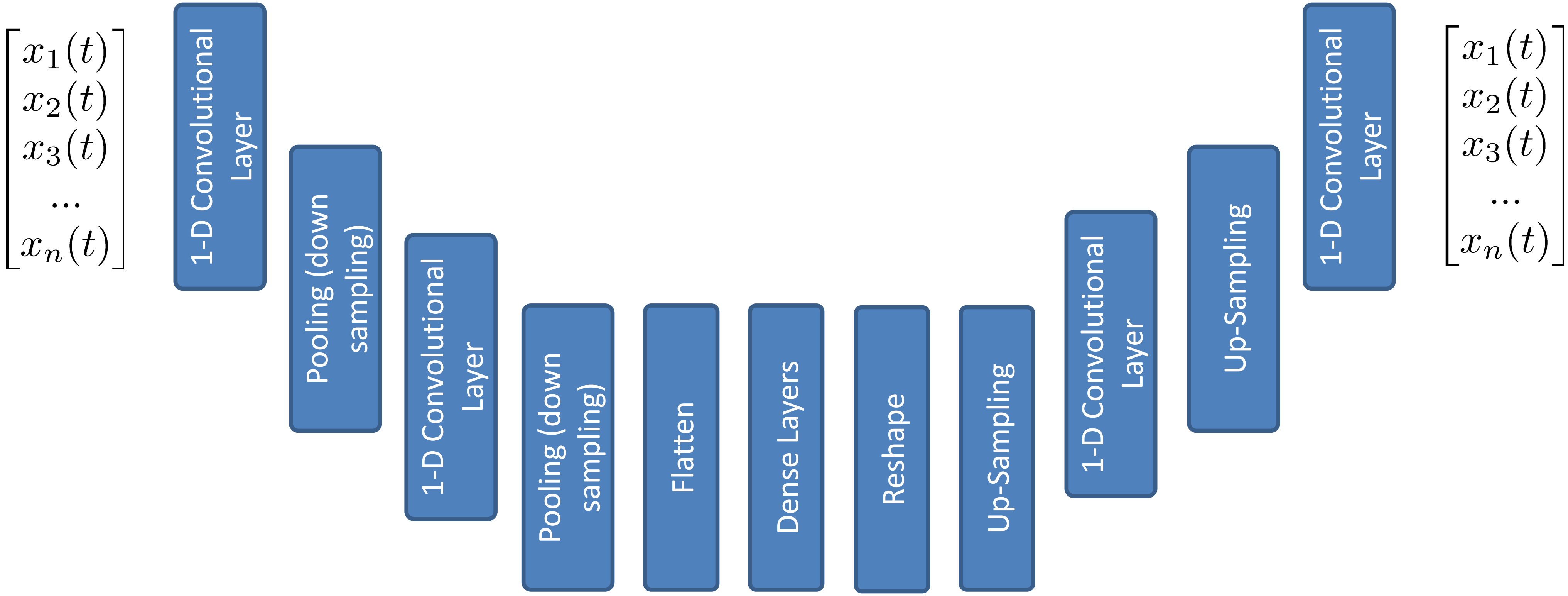


- Encoder-Decoder network
  - Transforms data into a latent space which is mapped to an output space



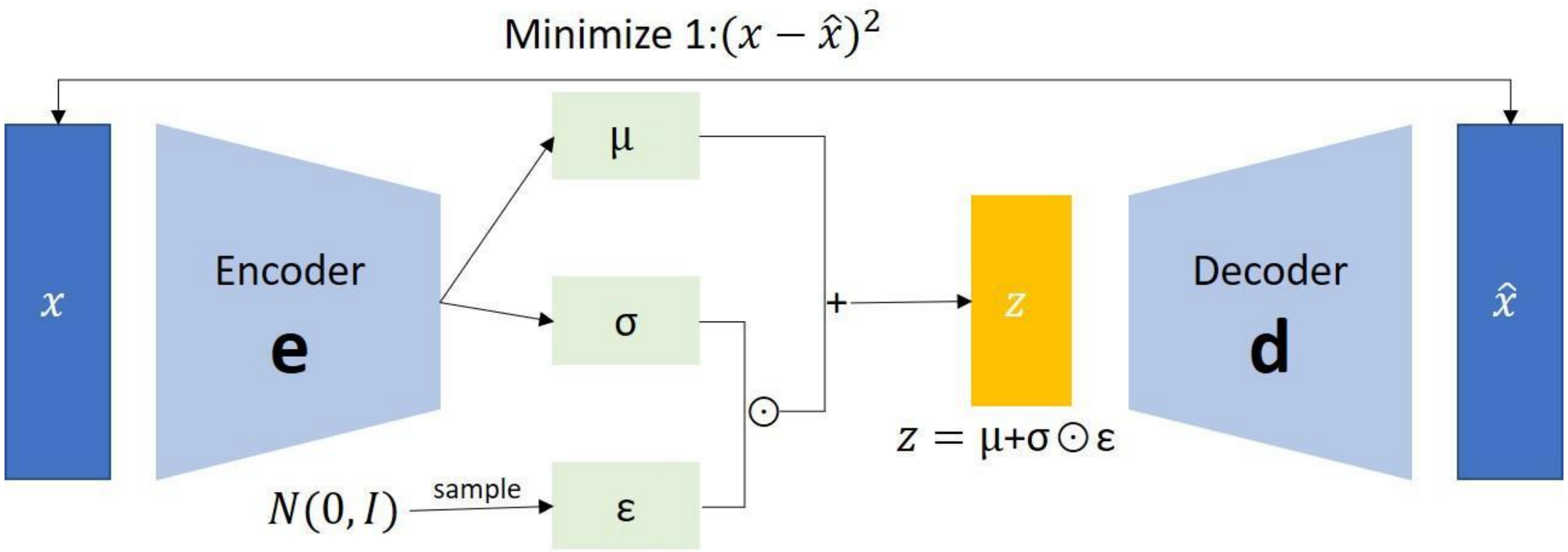
Convolutional Autoencoders

- Neural network that converts 1-D sequence into a latent-space
  - Filters learn translation invariant features much like an image based U-net
  - Pooling layers for downsampling
  - Signal is upsampled and filtered to reconstruct the original signal
  - Deconvolutional layers can also be used



Variational Autoencoders

- Variational autoencoders enforce smoothness condition in the latent space
- Dimensionality reduction removes complexity of noise
- Tests done using simulated BPM data
- Logically extended to RF data
- Could implement the autoencoder on a FPGA for near-real-time noise reduction



Cavity simulator

- Based on an equivalent RLC circuit model
  - Transmitted voltage differential equation:

$$\frac{d}{dt}\begin{bmatrix}\mathrm{Re}(V_t)\\ \mathrm{Im}(V_t)\end{bmatrix} = \begin{bmatrix}-\omega_{1/2} & -\Delta\omega\\ \Delta\omega & -\omega_{1/2}\end{bmatrix}\begin{bmatrix}\mathrm{Re}(V_t)\\ \mathrm{Im}(V_t)\end{bmatrix} + \frac{R_L\omega_{1/2}}{m}\begin{bmatrix}\mathrm{Re}(I_{fwd})\\ \mathrm{Im}(I_{fwd})\end{bmatrix}$$

  - Reflected voltage computed from transmitted:

$$V_r = \frac{1}{m}V_t - \frac{Z_0}{2}I_{fwd}$$

V_t: transmitted voltage   *$\omega_{1/2}$: half band-width*   *R_L: loaded "shunt" resistance*   *I_{fwd}: forward current*

V_r: reflected voltage   *$\Delta\omega$: frequency detuning*   *m: cavity/waveguide coupling ratio*   *Z_0: reference impedance*

Cavity Simulator

```yaml
---

# Configure pulse settings
Pulse :
  frequency : 2855.95e6 # Driving frequency, Hz
  time_step : 1.e-7 # Time resolution of simulator
  rate : .5 # Pulse rate, Hz
  duration : 20.e-6 # Pulse duration, s

# Configure LLRF elements
Elements :

  # Primary signal generator (klystrino)
  Gen :
    type : Generator
    ports : [GenTee]
    signal_type : tophat
    signal_params :
      amp : 50. # A
      start : 1.e-6 # s
      duration : 5.e-6 # s
      rho : 2.5e8 # s^-1
    noise:
      Ig : 1. # A
      phase : .15 # rad
```

```yaml
  # First RF cavity
  Cav1 :
    type : Cavity
    ports : [TeeCav1, Beam]
    f0 : 2856e6 # Hz
    Q : 2.e4
    R : 3.e3 # Ohm
    beta : 1.
    noise :
      Pr : 1.e3 # W
      Pt : 1.e3 # W
      rphase : .15 # rad
      tphase : .15 # rad

  # Line from magic tee to phase shifter
  TeeShift :
    type : Line
    ports : [Tee, Shift]
    alpha : 0. # m^-1
    beta : 0. # m^-1
    length : 1. # m
```

```
LLRFSim Command Line Interface

--------------------------------------------------------------------------------
EPICS PVA Server: ON
PV Prefix: ACCEL

Simulator Commands:

        1) Start or stop RF pulses
        2) List all LLRF components
        3) Modify an LLRF component
        4) Start or stop an EPICS PVA server
        5) Print EPICS PVA server configuration
        6) Write an LLRFSim config file
        7) Load an LLRFSim config file

        0) Exit

Enter a # for one of the commands listed above:
->
```

LLRF Elements

Name	Type	Port Connections	Free Parameters
Gen	Generator	1: GenTee	phiG: 0.000 signal_type: tophat signal_params: {'amp': 50.0, 'start': 1e-06, 'duration': 5e-06, 'rho': 250000000.0}
Beam	Beam	1: Cav1 2: Cav2	phiB: 0.000 signal_type: step signal_params: {'amp': 0.0, 'start': 0.0, 'duration': 0.0}
GenTee	Line	1: Gen 2: Tee	alpha: 0.000 beta: 0.000 length: 1.000
Tee	MagicTee	1: GenTee 2: TeeLoad 3: TeeCav1 4: TeeShift	
TeeLoad	Line	1: Tee 2: Load	alpha: 0.000 beta: 0.000 length: 1.000
Load	Load	1: TeeLoad	
TeeCav1	Line	1: Tee 2: Cav1	alpha: 0.000 beta: 0.000 length: 1.000
Cav1	Cavity	1: TeeCav1 2: Beam	f0: 2.856e+09 Q: 2.000e+04 R: 3.000e+03 beta: 1.000 Z0: 50.000
TeeShift	Line	1: Tee 2: Shift	alpha: 0.000 beta: 0.000 length: 1.000
Shift	Shifter	1: TeeShift 2: ShiftCav2	shift: 1.571
ShiftCav2	Line	1: Shift 2: Cav2	alpha: 0.000 beta: 0.000 length: 1.000
Cav2	Cavity	1: ShiftCav2 2: Beam	f0: 2.856e+09 Q: 2.000e+04 R: 3.000e+03 beta: 1.000 Z0: 50.000

Kalman Filtering

- Continuous time model
  - Prediction of the reflected and transmitted from the drive signal
  - Linear cavity model including detuning
  - Noise included in the model

$$\boldsymbol{x} = \begin{bmatrix} \mathrm{Re}(V_t) \\ \mathrm{Im}(V_t) \end{bmatrix}, \qquad \boldsymbol{u} = \begin{bmatrix} \mathrm{Re}(I_{fwd}) \\ \mathrm{Im}(I_{fwd}) \end{bmatrix}, \qquad \boldsymbol{y} = \begin{bmatrix} \mathrm{Re}(V_t) \\ \mathrm{Im}(V_t) \\ \mathrm{Re}(V_r) \\ \mathrm{Im}(V_r) \end{bmatrix}$$

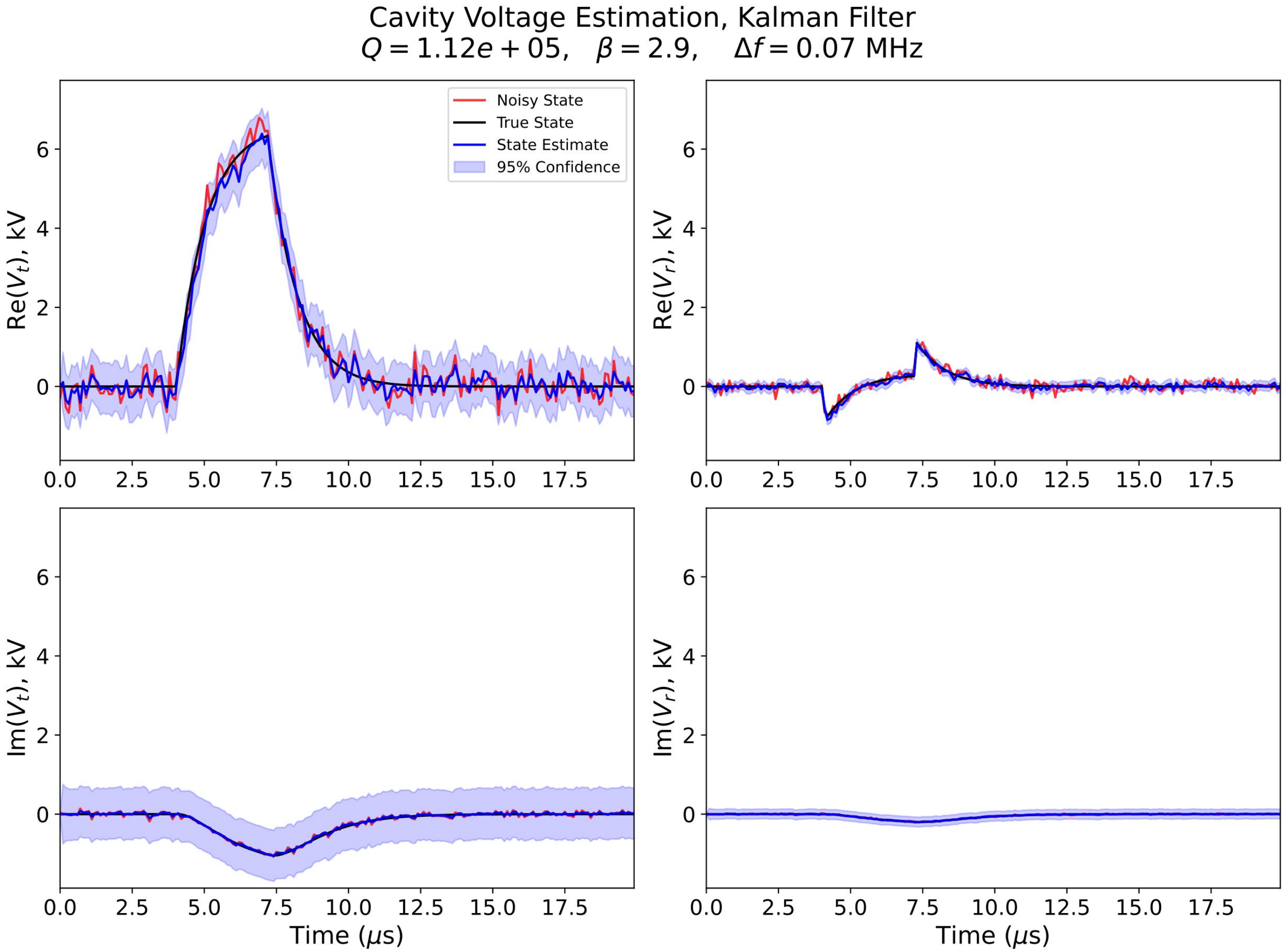
$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{u} + \Gamma\tilde{\boldsymbol{w}}, \qquad \boldsymbol{y} = C\boldsymbol{x} + D\boldsymbol{u} + \tilde{\boldsymbol{v}}$$

$$A = \begin{bmatrix} -\omega_{1/2} & -\Delta\omega \\ \Delta\omega & -\omega_{1/2} \end{bmatrix}, \qquad B = \frac{R_L \omega_{1/2}}{m} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/m & 0 \\ 0 & 1/m \end{bmatrix}, \qquad D = -\frac{Z_0}{2} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\tilde{\boldsymbol{w}} \sim \mathcal{N}(\boldsymbol{0}, W), \qquad \tilde{\boldsymbol{v}} \sim \mathcal{N}(\boldsymbol{0}, R)$$

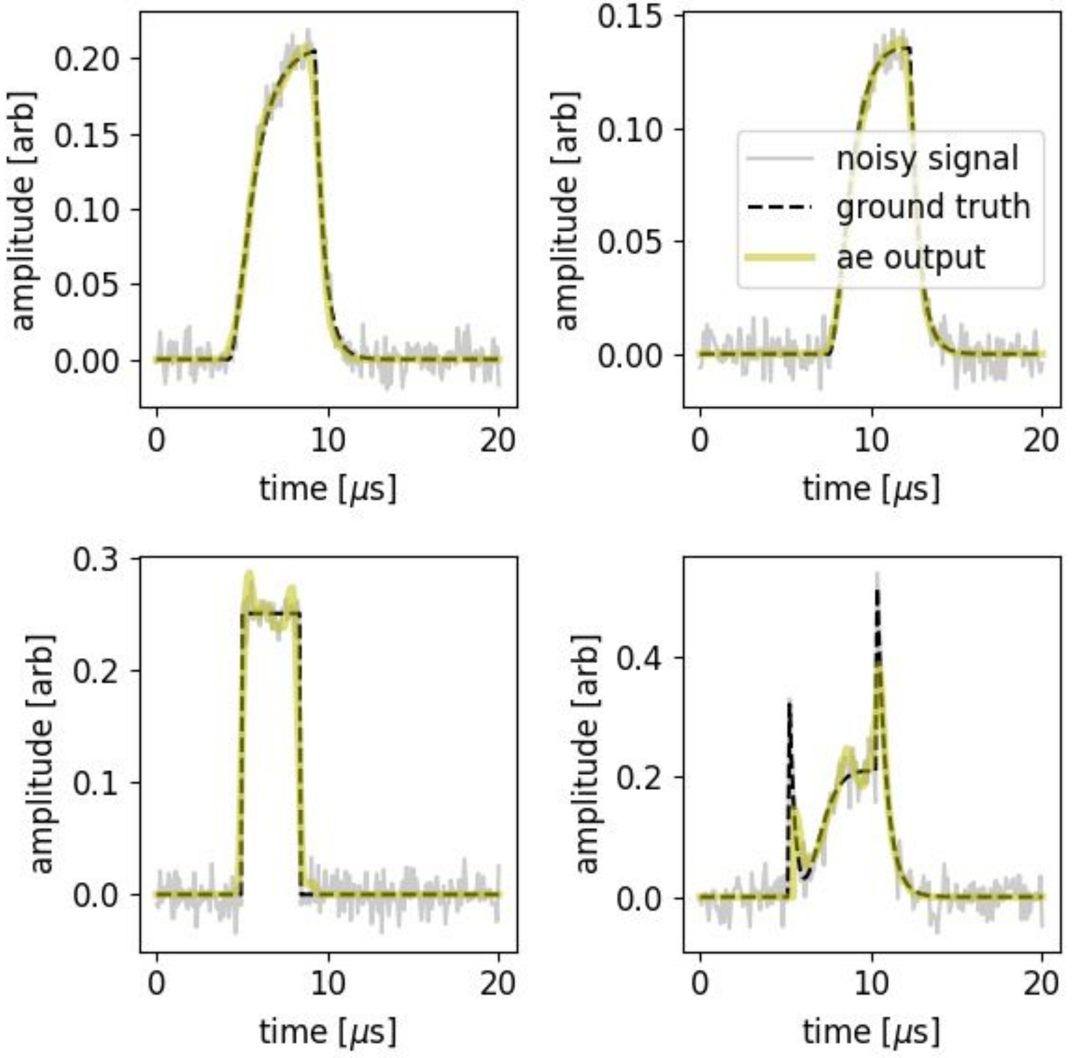
Kalman Filtering

- Prediction of transmitted and reflected signal given a drive signal
  - State estimation include noise and variance



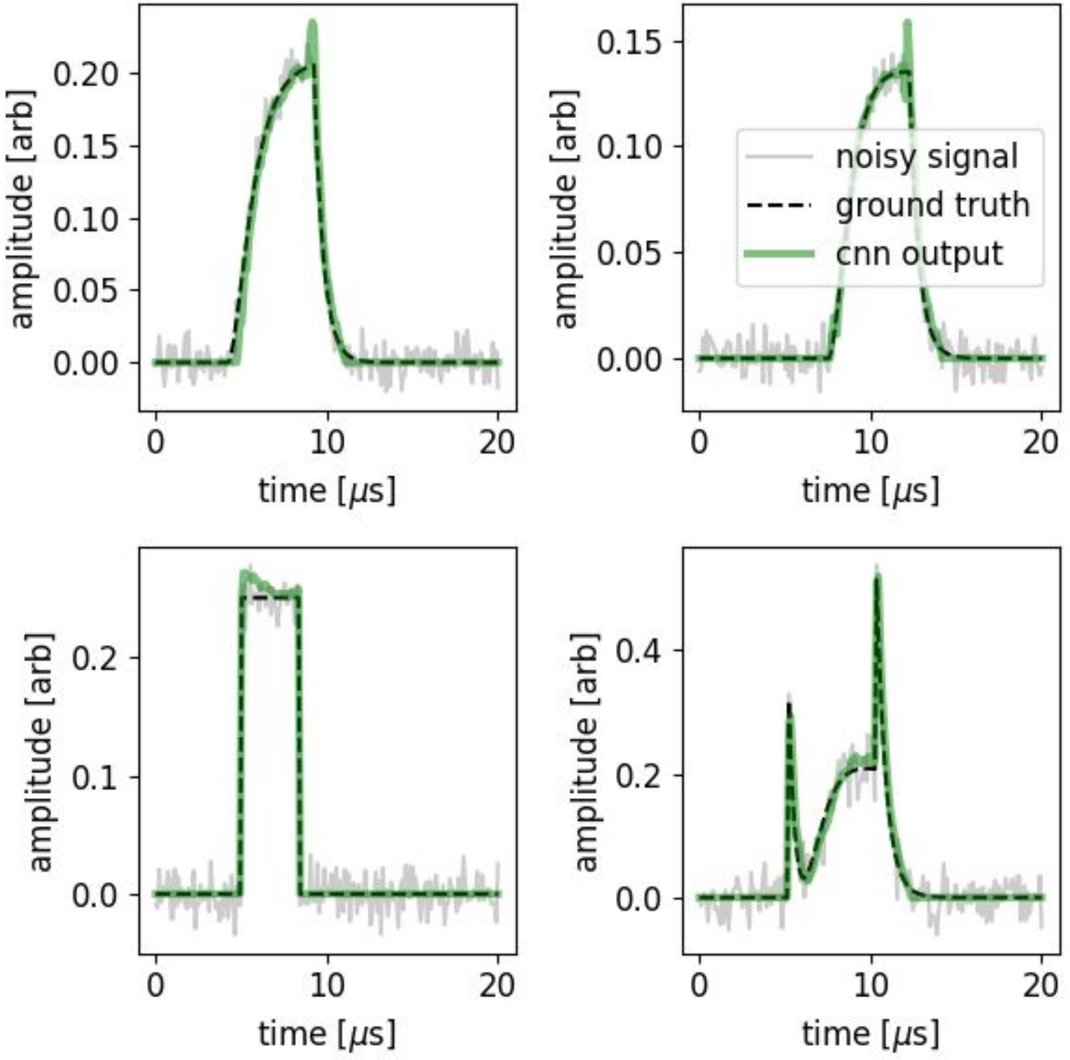
Feedforward Autoencoder

- Each time point in the sequence is considered a feature
- Each type of signal (forward, reflected, probe) are used to train a separate model
- Model parameters
  - latent space dimension 32
  - single encoder and decoder layer
  - trained for 100 epochs
  - batch size of 100



Convolutional Autoencoder

- Trained on RF simulator data with noise
  - Each signal used as a separate input for one model
- Model architecture
  - 4 encoding layers with filters and pooling
  - flatten and then dense layers reducing down to a latent dimension of 10
  - 4 decoding layers with filters and upsampling
- Batch size varied between 20 and 1000
- Trained for 1000 - 2000 epochs

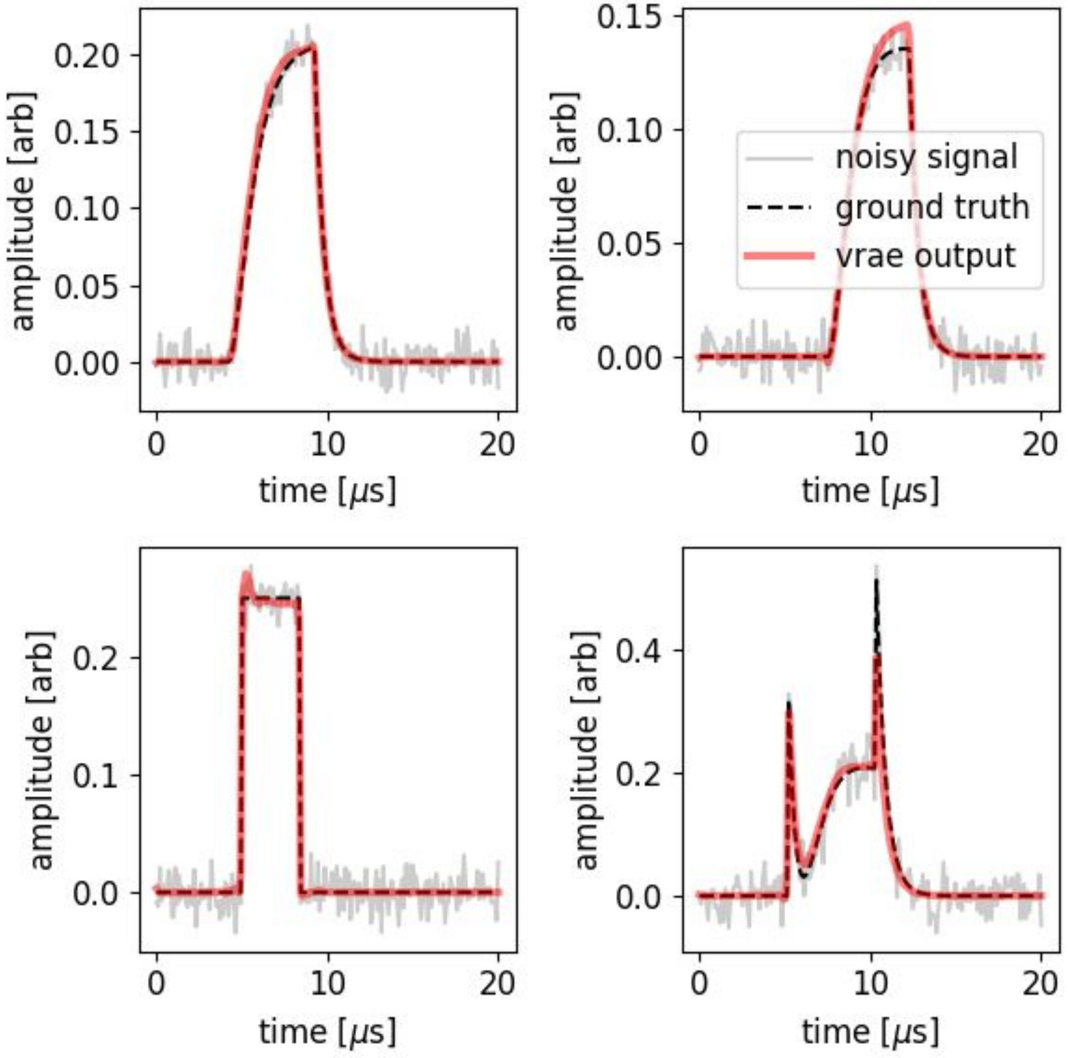


Variational Recurrent Autoencoders

- Forward, reflected, and probe are all considered features to the model

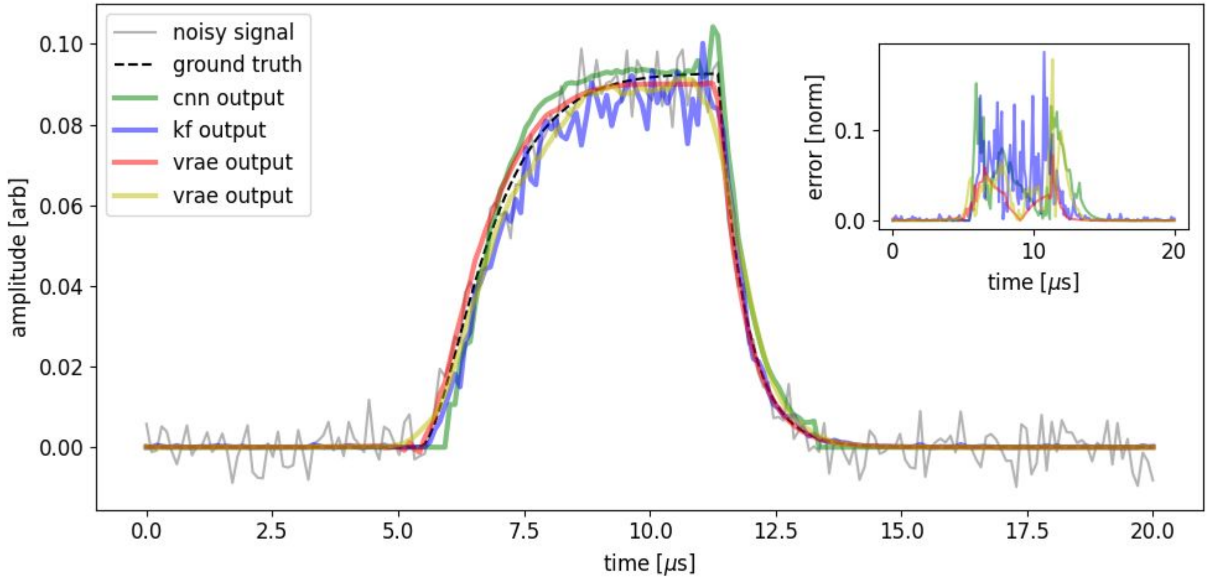
- Model architecture
  - 3 hidden layers
  - latent space of 12
  - recurrent layers are used for the encoder and the decoder

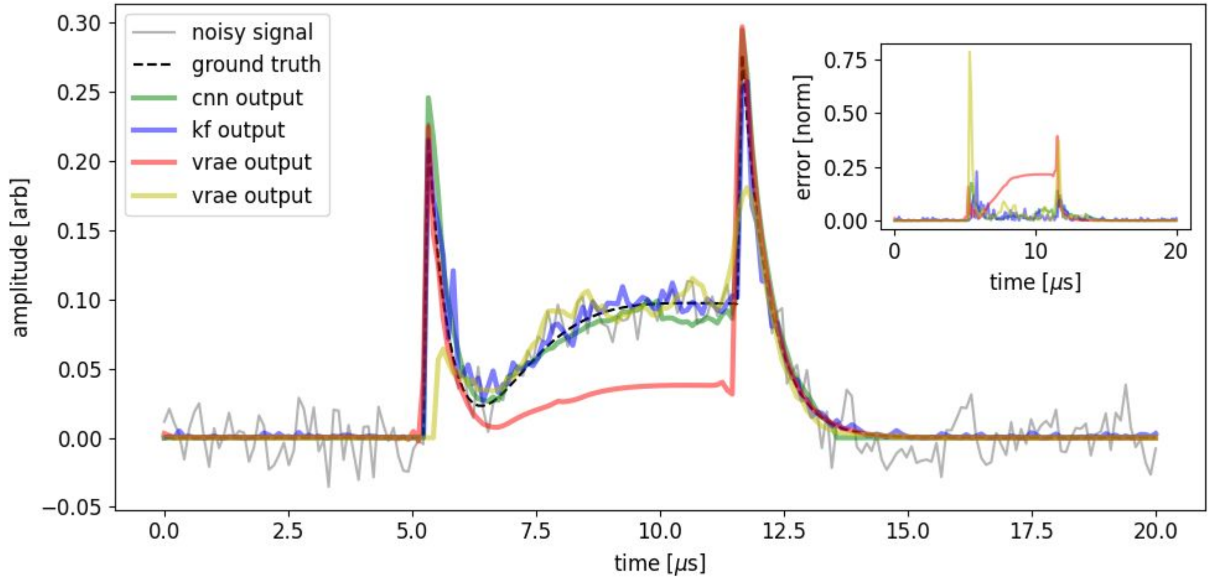
- Batch size of 16
- Train for 840 epochs
- Gaussian reconstruction loss



Comparison of Noise Reduction Methods

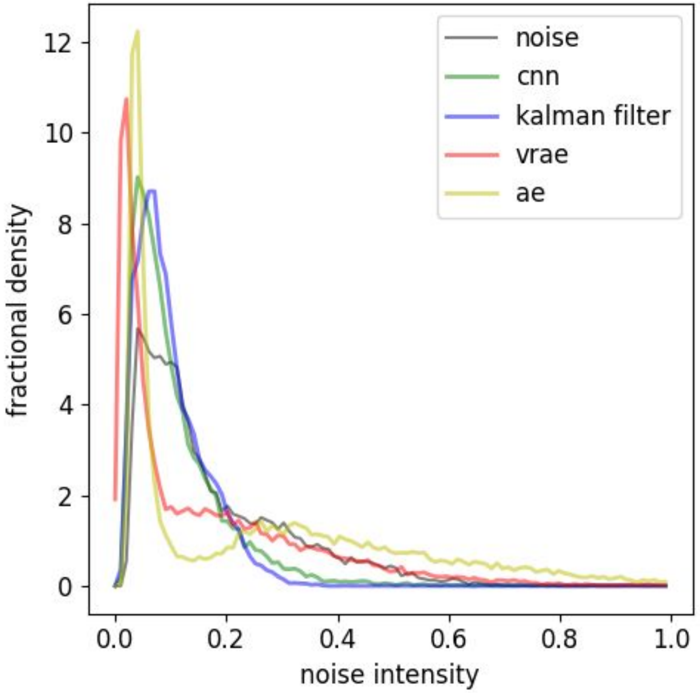
- Two test waveforms
  - Probe signal (top)
  - Reflected (bottom)
- Inset plots show the error between the ground truth and the reconstructed signal

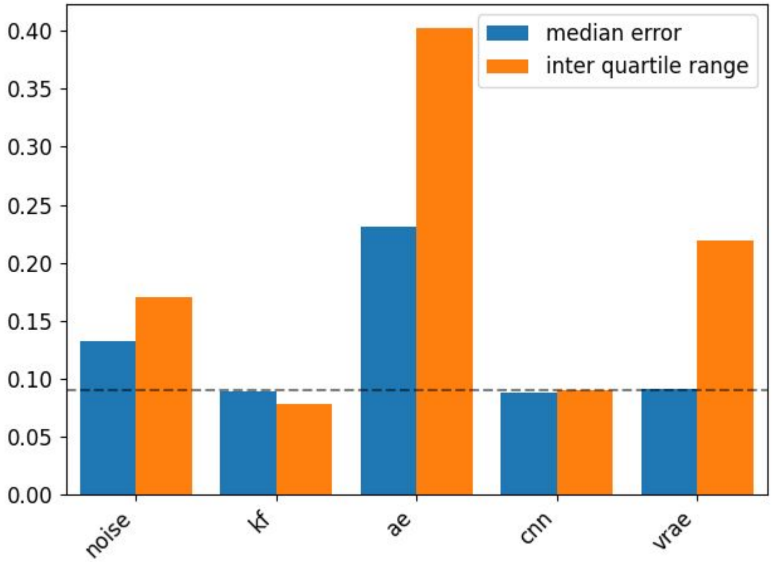


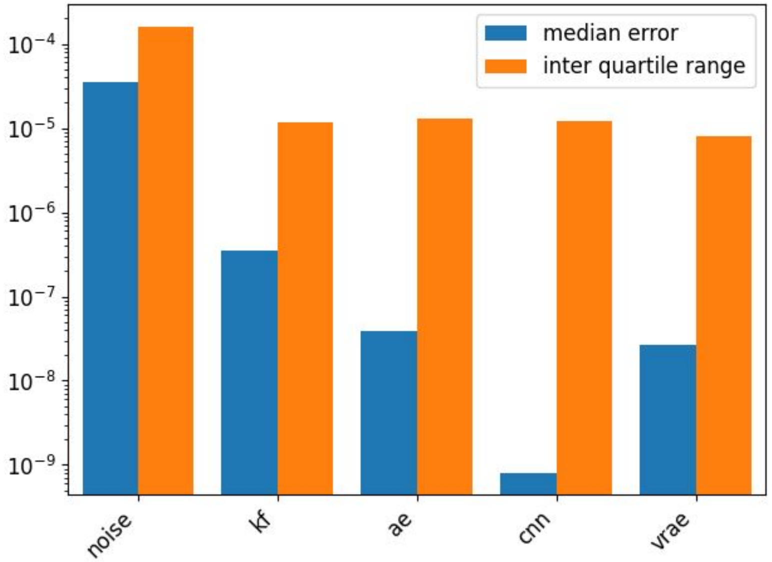


Comparison of Noise Reduction Methods

- Compute the root-sum-squared error for each example (across the whole waveform)
  - Histogram of the squared error on the test data comparing the autoencoder, convolutional autoencoder, variational autoencoder, and the kalman filter (left)
  - Error statistics showing median and interquartile range for each noise reduction method (middle)
- Compute median and interquartile range for the squared error across the whole dataset (right)

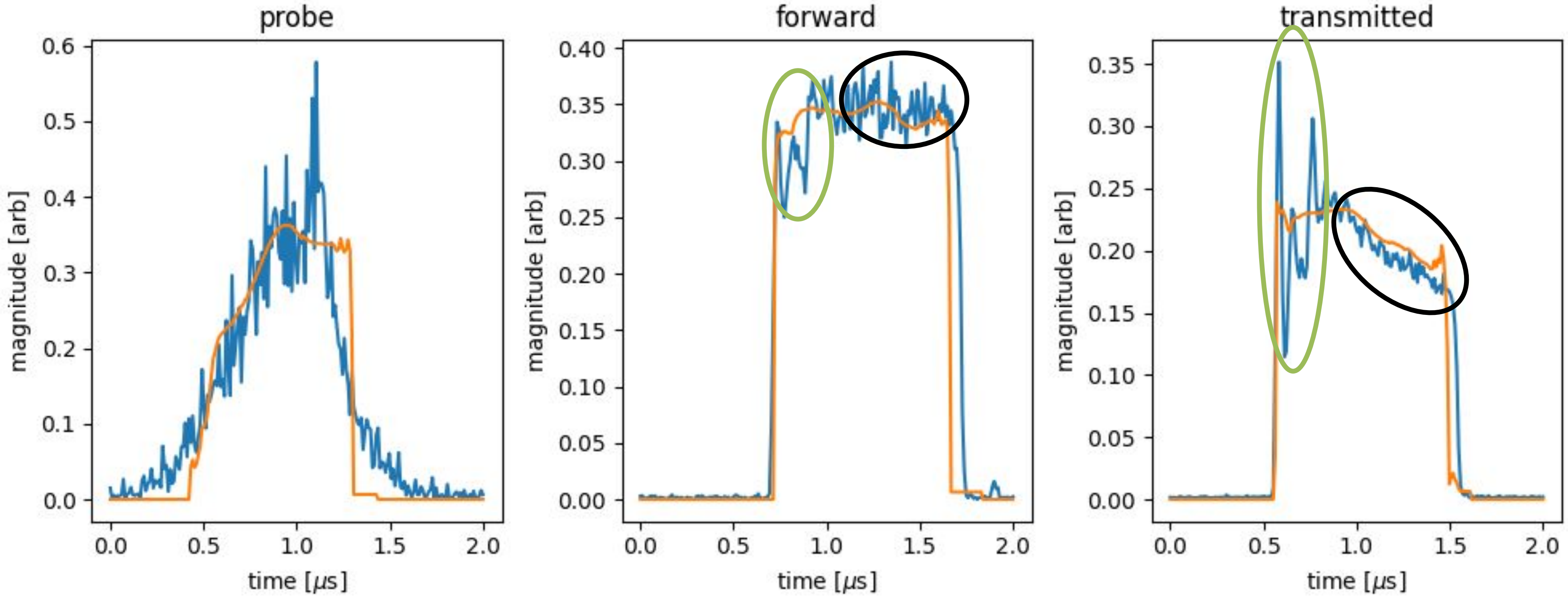






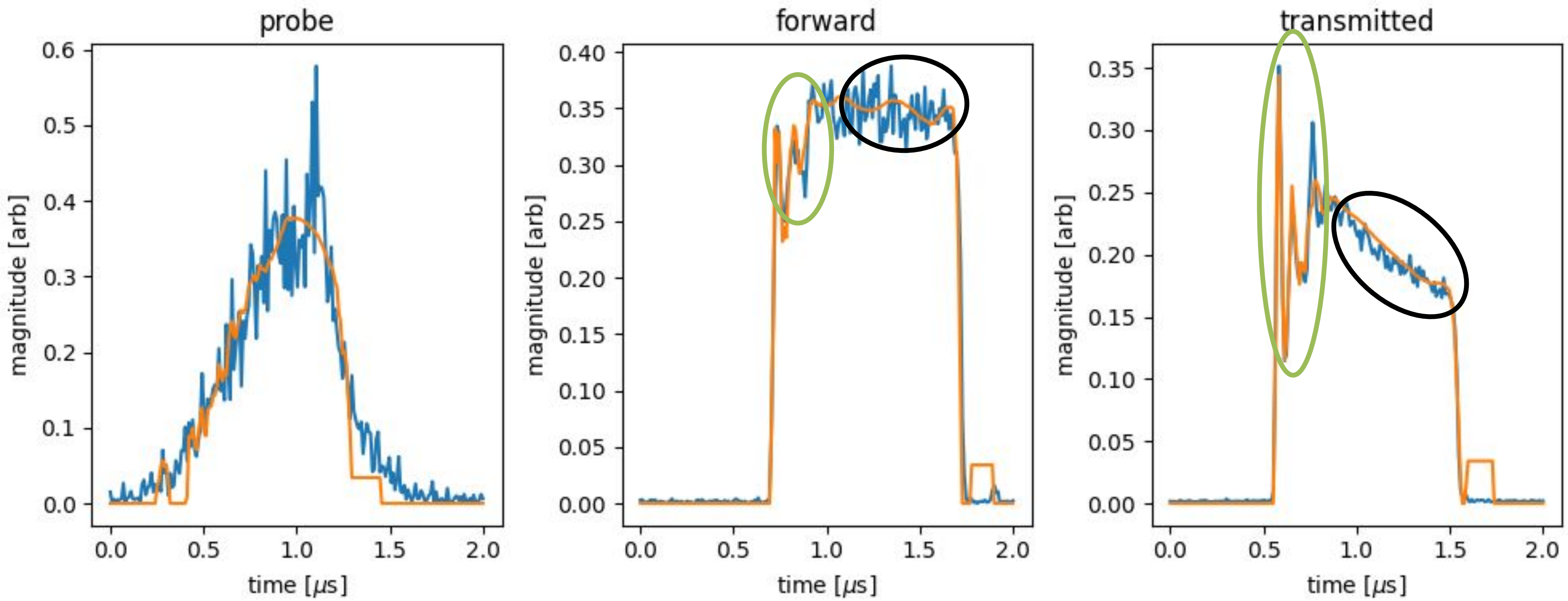
CNN Autoencoder Performance on Measurements

- Data collected from industrial system under development at RadiaBeam
  - Low signal levels with a high degree of klystron noise
  - Novel structure has unique RF characteris
  - Black circles are noise
  - Green circles are not noise



CNN Autoencoder Retraining

- Data collected from industrial system under development at RadiaBeam
  - Low signal levels with a high degree of klystron noise
  - Novel structure has unique RF characteris
  - Black circles are noise
  - Green circles are not noise



Conclusions

- Industrial accelerators have a large landscape of applications
  - growing demand for industrial systems
  - complexity of industrial accelerators is increasing
  - automation is critical when operating outside the laboratory environment

- Developing ML tools for automation
  - Initial studies focused on noise reduction
  - Various ML methods show promise for this application
  - Model development using simulator
  - Testing shows promise

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.