# Machine Learning and Controls in GW detectors

**Gabriele Vajente**
LIGO Laboratory Caltech

**GWADW 18**

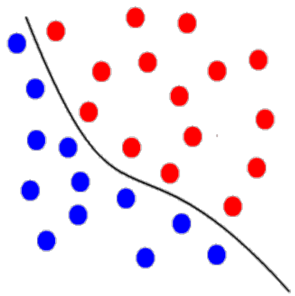
13th May 2018
Girdwood, Alaska

LIGO-G1800874-v2

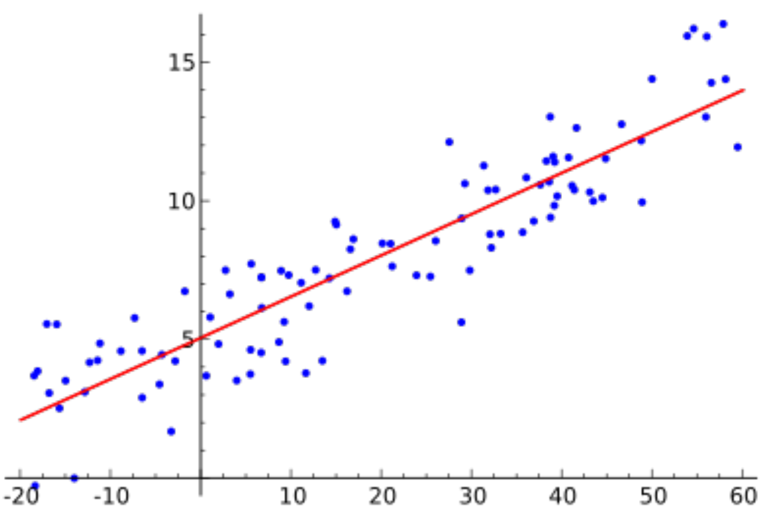
Image credit: www.1and1.com

# What is machine learning?

In computer science **Artificial Intelligence** research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of success to achieve some goal.
[Poole, Mackworth & Goebel 1998]



**Machine learning** is a field of computer science that gives computers the ability to learn without being explicitly programmed.
[supposedly Arthur Samuel, A.I. pioneer 1959]




***Linear regression*** *is the simplest example of machine learning: the parameters of a model can be obtained based on data, without explicitly setting their value*

Images credit: wikipedia

# Machine learning applications

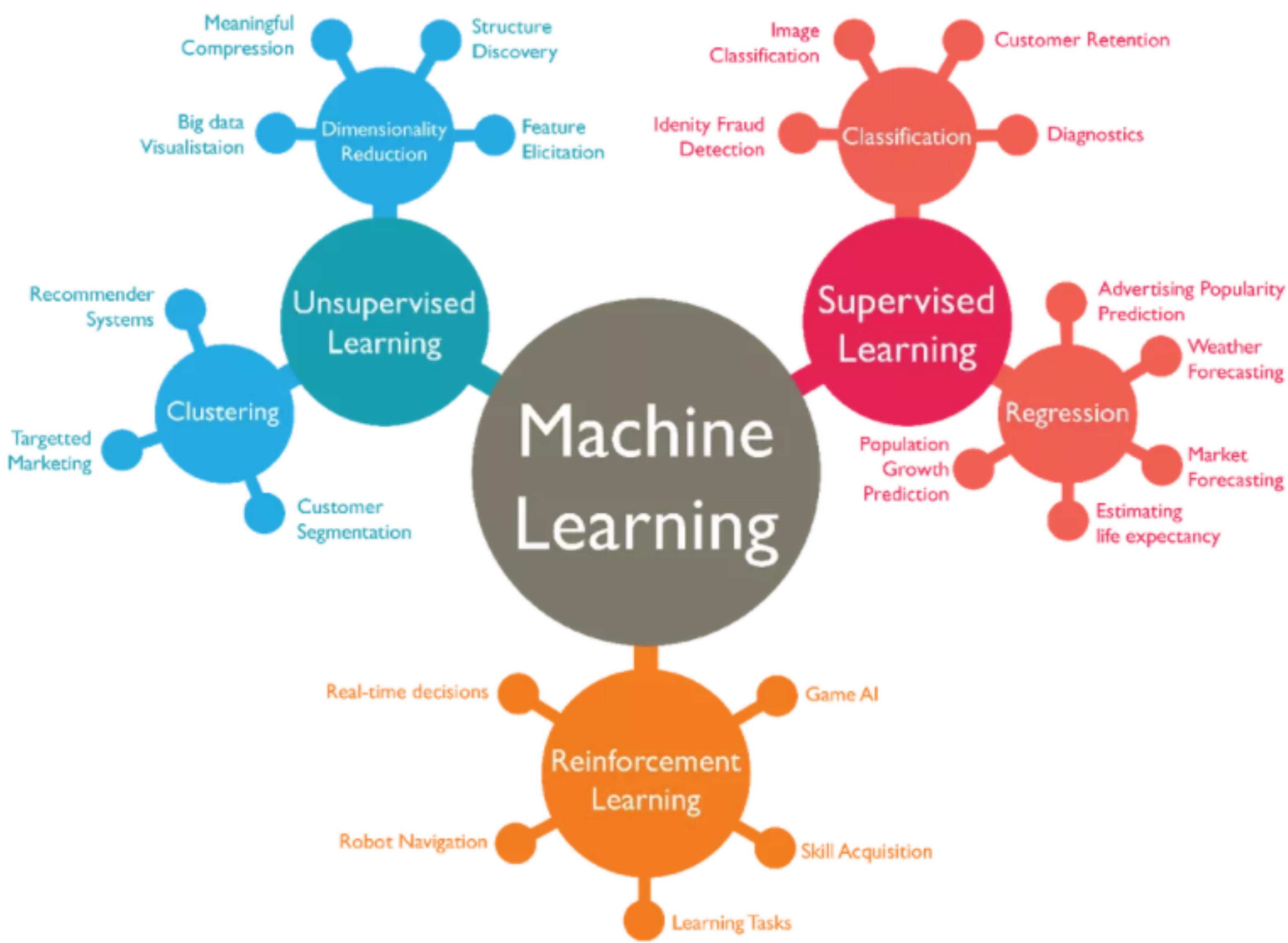



Image credits:www.datasciencecentral.com

# Machine Learning in LIGO

- Many successful applications to data analysis and detector characterization

**Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science**

M Zevin[1], S Coughlin[1], S Bahaadini[2], E Besler[2], N Rohani[2], S Allen[3], M Cabero[4], K Crowston[5], A K Katsaggelos[2], S L Larson[1,3] +Show full author list

Published 28 February 2017 • © 2017 IOP Publishing Ltd

Classical and Quantum Gravity, Volume 34, Number 6

ELSEVIER

Physics Letters B

Volume 778, 10 March 2018, Pages 64-70

open access

**Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data**

Daniel George a, b, E.A. Huerta b

Show more

https://doi.org/10.1016/j.physletb.2017.12.053

Get rights and content

**Machine Learning in Characterization and Commissioning @ GW Detectors**

LIGO-G1701113

Nikhil Mukund Menon

IUCAA Pune

**Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy**

Hunter Gabbard, Michael Williams, Fergus Hayes, and Chris Messenger

Phys. Rev. Lett. **120**, 141103 – Published 6 April 2018

https://wiki.ligo.org/MLA/ML_at_LIGO_and_VIRGO

# How about control problems?

System parameter identification and modeling

Optimal control
Kalman filter

Adaptive control

Genetic programming

Reinforcement learning

Time series forecasting

Anomaly detection

Robotics

Self driving vehicles

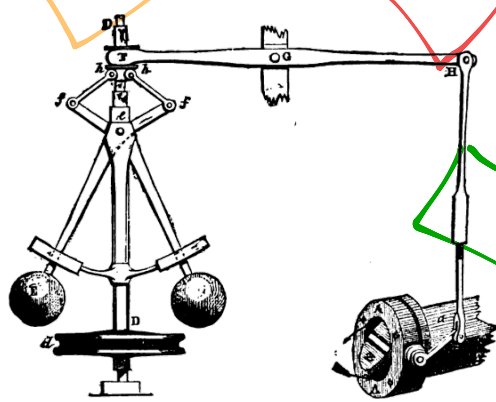


FIG. 4.—*Governor and Throttle-Valve.*

Video games

Dynamical systems numerical integration

By R. Routledge - Image from "Discoveries & Inventions of the Nineteenth Century" by R. Routledge, 13th edition, published 1900., Public Domain, https://commons.wikimedia.org/w/index.php?curid=231047

# Deep Learning

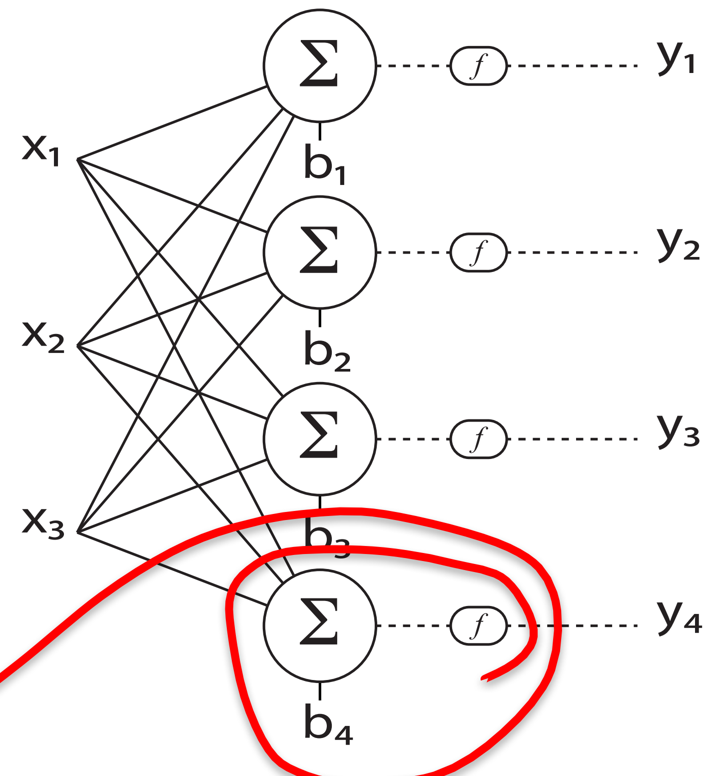
Deep Learning $\equiv$ ( Linear regression $+$ A pinch of non-linearity )$^{N}$

*and an awful lot of data and computational power…*

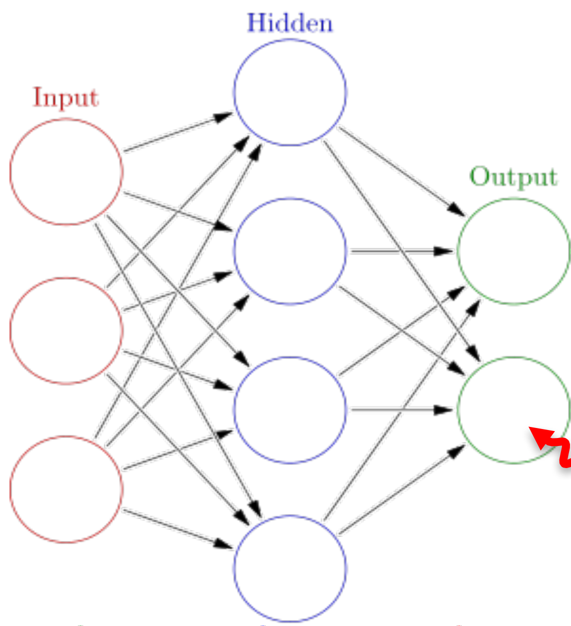
# Artificial neural networks

- Neurons can be organized into layers
- There are always an input and an output layer
- Intermediate layers are called **hidden**
- “Many” hidden layers = **deep neural network (DNN)**

$$y_i = f\left(\sum_j W_{ij} x_j + b_i\right)$$




https://en.wikipedia.org/wiki/Artificial_neural_network




Here each circle represents a unit, including non-linear activation function

$$Y = f\left(W^{(O)} f\left(W^{(H)} f\left(W^{(I)} X + b^{(I)}\right) + b^{(H)}\right) + b^{(O)}\right)$$

# Deep Learning in Controls

**Reinforcement learning**: not ready yet to get out of research and into real world applications


Google DeepMind

# Applications of Deep Learning in GW detectors

- Non-linear estimator of longitudinal degrees of freedom to aid lock acquisition
- Camera image processing to extract beam spot position
- Discovery of non-linear noise couplings and subtraction

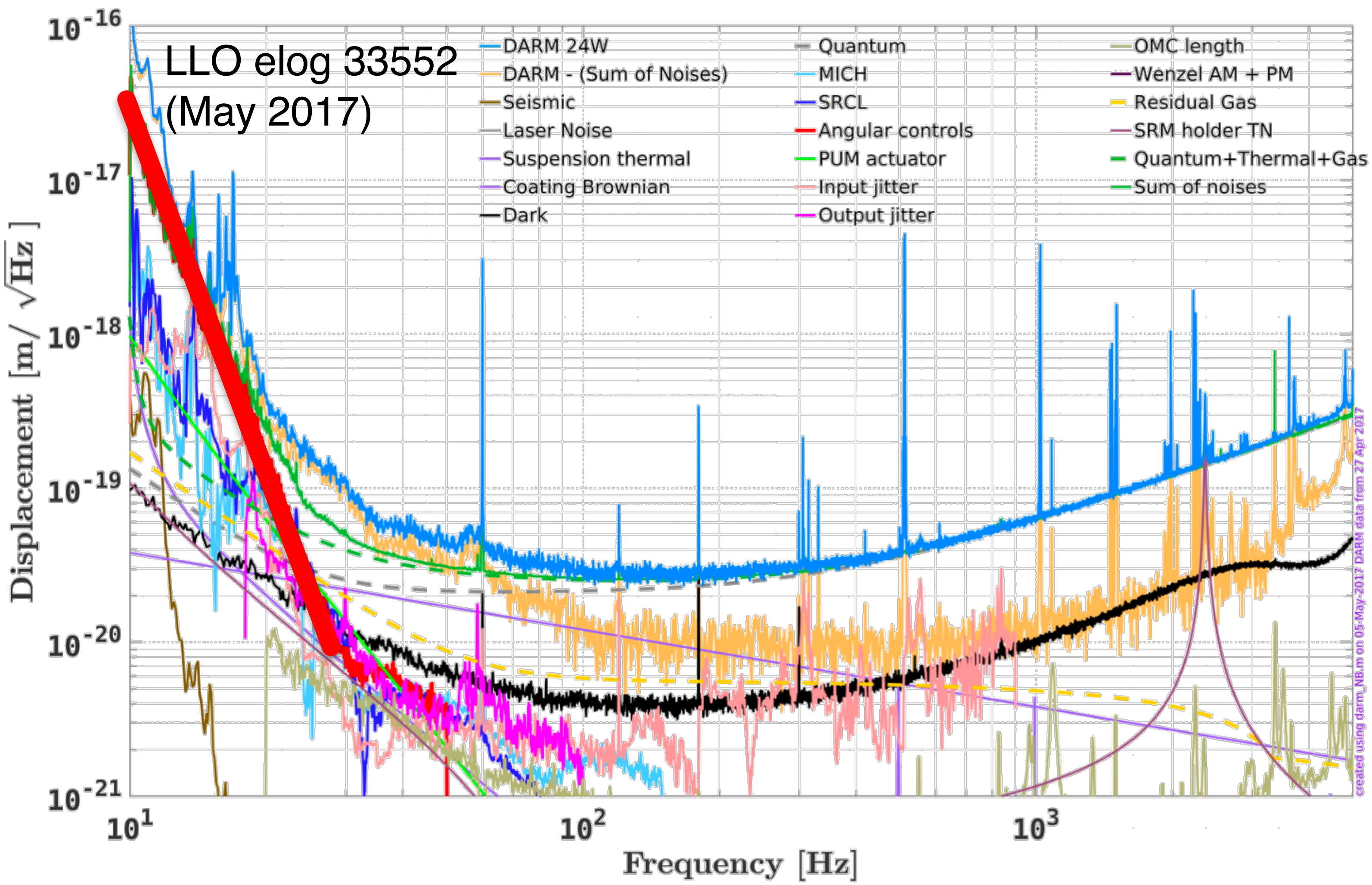

Huffington Post

https://techbuf.com/human-brain-neural-network/

# IMAGE PROCESSING FOR BEAM CENTERING

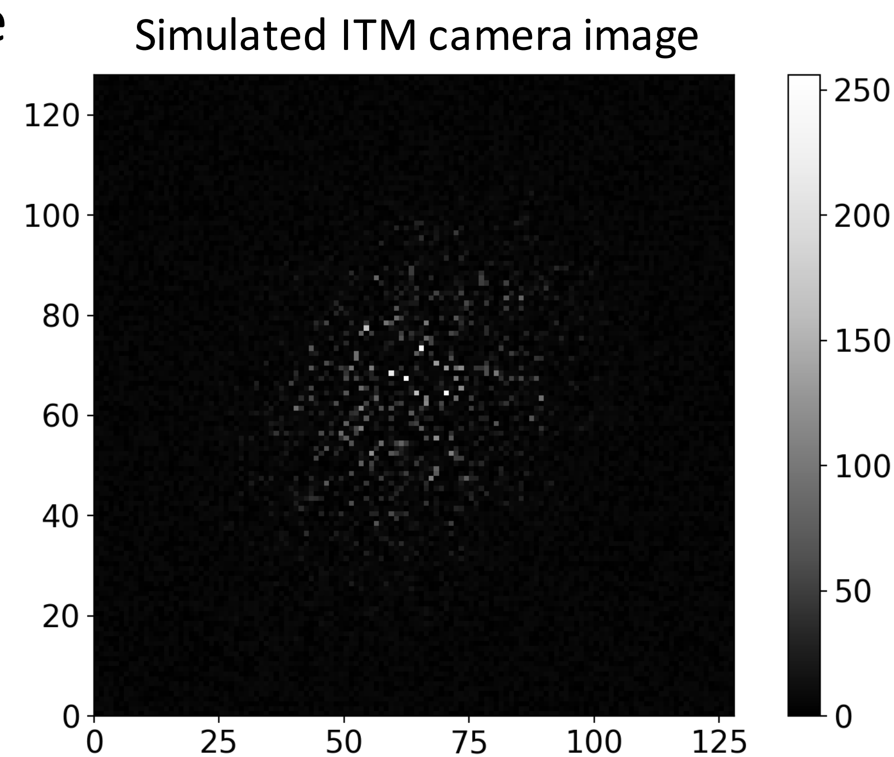
# Beam spot centering

- If the beam is miscentered on the test masses, angular noise couples to h(t): bad!
- Beam centering / actuator balance / angle to length tuning: all different ways to look at the same problem

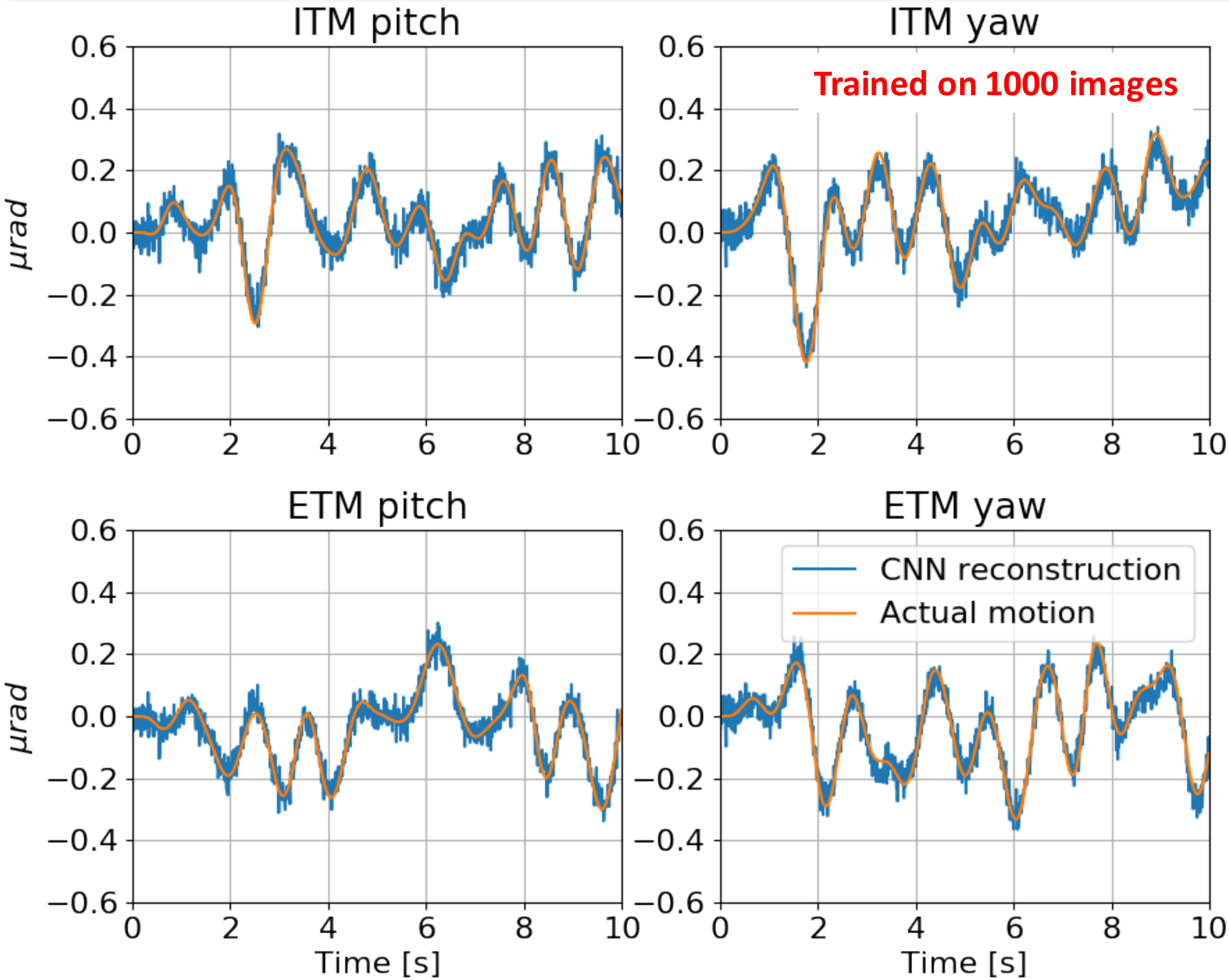


# Simulation

- Random seismic angular motion of both test masses
- Compute the cavity axis and the beam spot position
- Simulate a scatter points on the mirrors: uniform spatial distribution, Gaussian distribution of scatter intensity
- Shine the beam on the mirror and simulate a camera image with some angle of view and some background noise
- Input: a series of image pairs (ITM and ETM) for each time, with the beam moving as dictated by the mirror angular motions

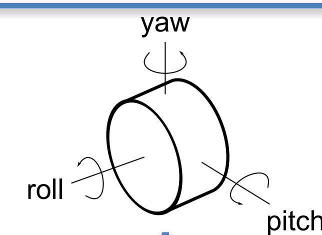


# Simulation results

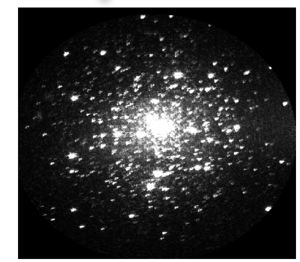


# But how to train it for the real system?

- Image depends on (unknown!) scatter distribution and camera angle / magnification
- We need to train on real images
- Create a training set
- We can modulate pitch and yaw of all test masses in a known way
- Use high frequency high amplitude angular dither lines, demodulate to measure beam spot position (similar to current A2L procedure)
- Collect some tens of minutes of data at ~10-30 fps
- Train on images and measured beam spot
- The trained network can continuously reconstruct the beam spot positions (mirror angles) without need for angular lines




In-lock A2L continuous measurement



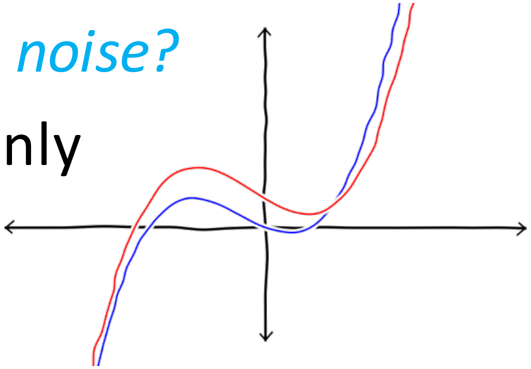
Train CNN offline

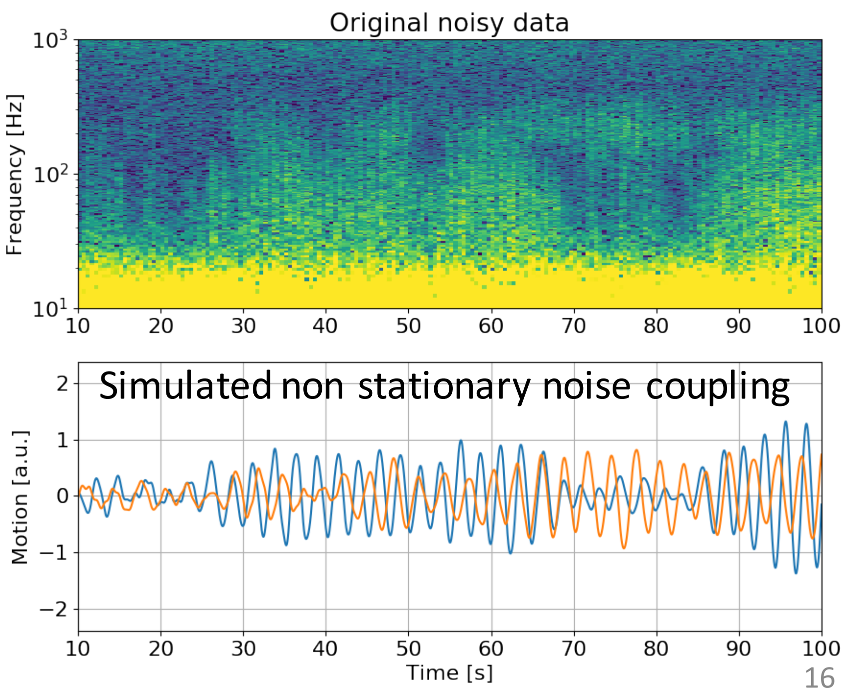
Michael Coughlin, Rich Ormistom, Gabriele Vajente, Rana Adhikari

# NON-LINEAR NOISE COUPLINGS

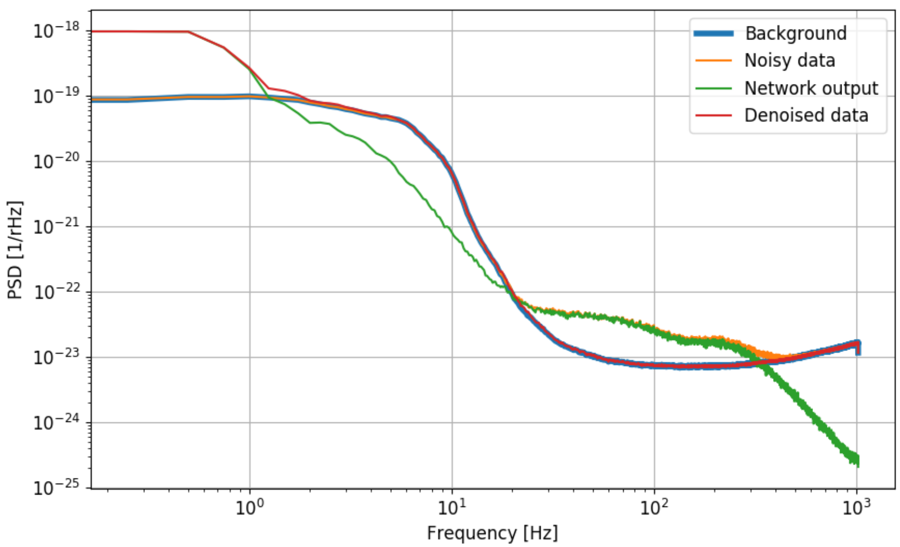
# Neural networks as function approximators

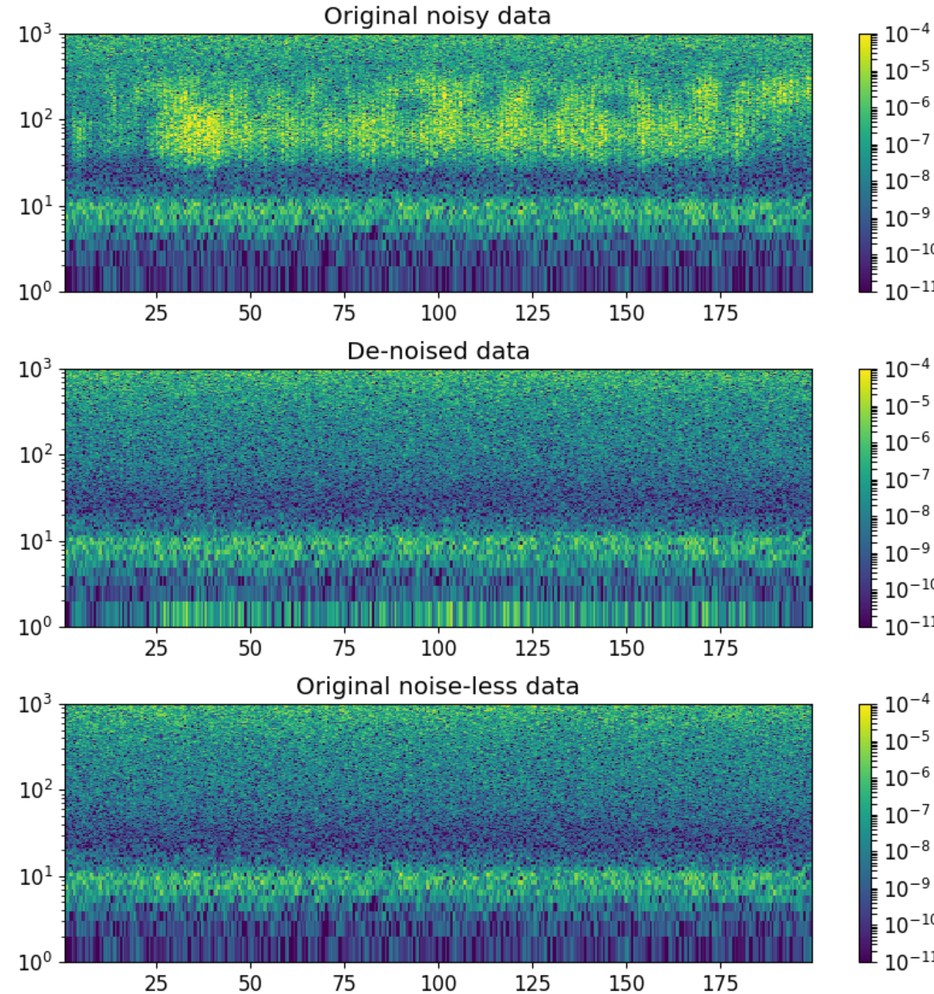
- *Can a neural network learn non-linear and non-stationary couplings in Advanced LIGO noise?*
- Start simple by using simulated data with only one noise coupling at a time
  - Modulated jitter noise
  - SRCL-like sensing noise with double modulation
- Rest of the talk:
  - Examples of generated data
  - Network architecture
  - Results on **simulated data**
  - **Some trial with real data**





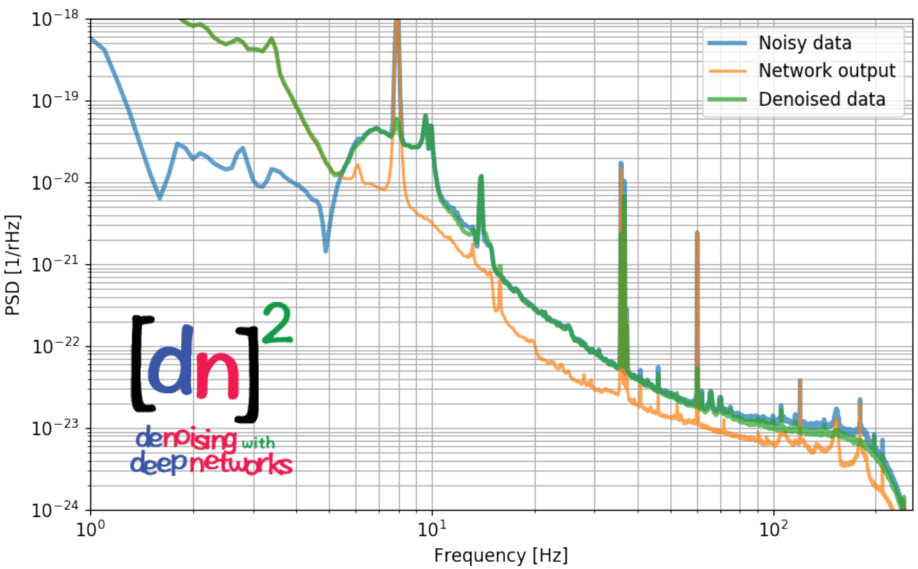
# Simulated data

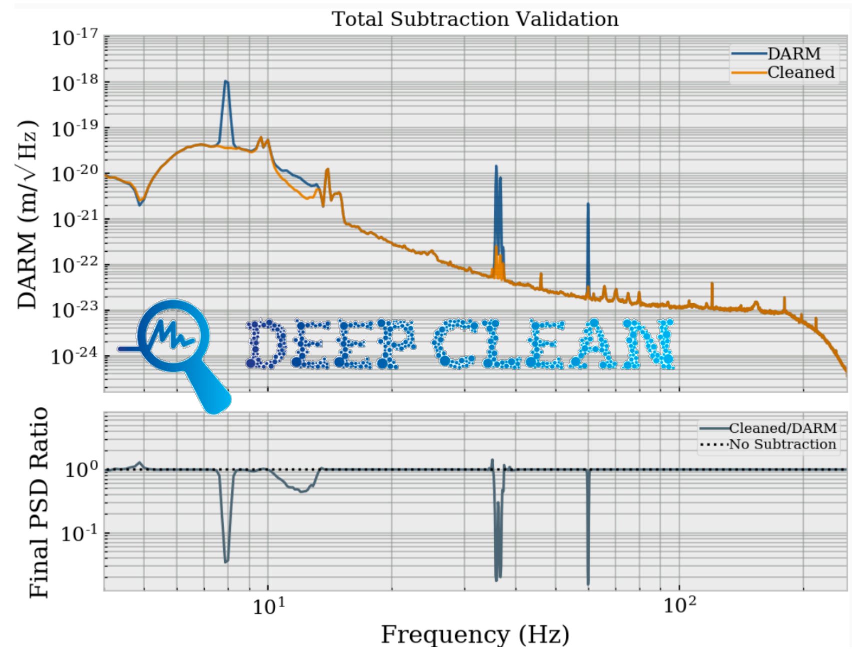



- Two noise coupling paths, modulated by different seismic motion
- Witness: sensing noise, seismic motions
- The network discovers a non trivial transfer function modulation, training on a few tens of minutes of data

# Real Data – LHO during O2



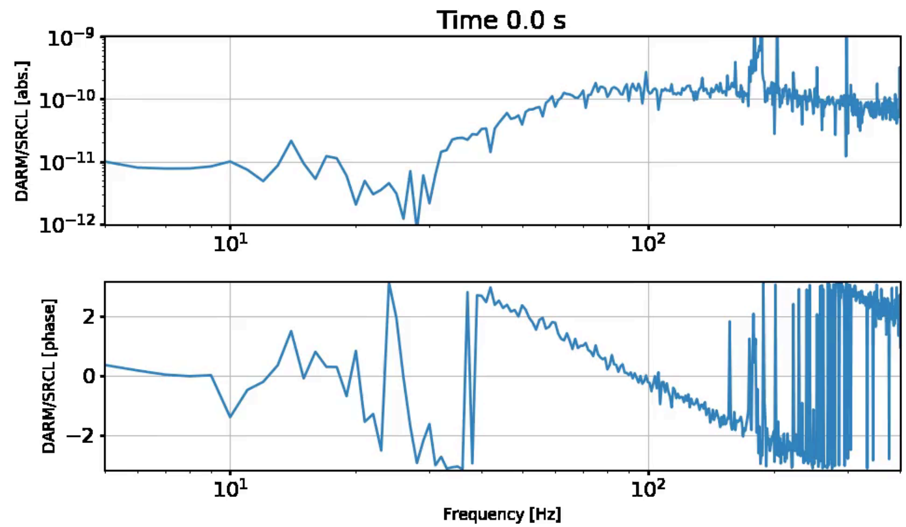



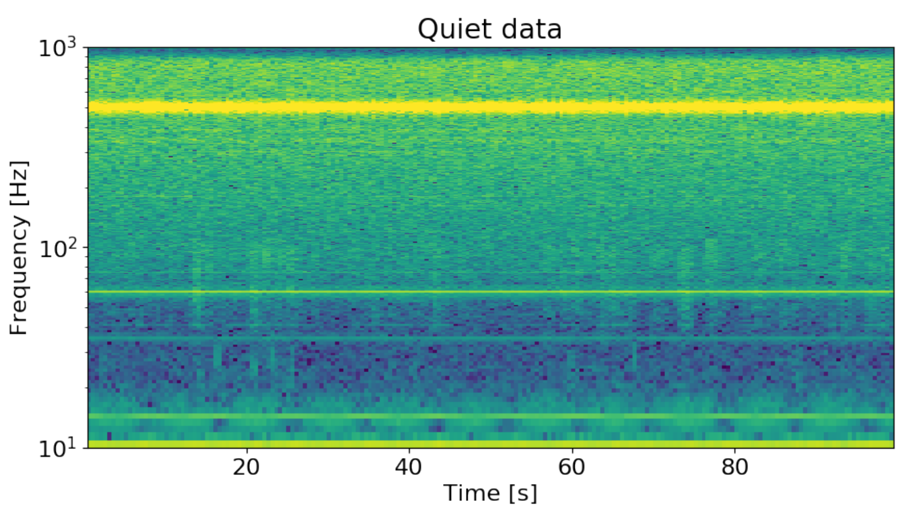

H1:GDS-CALIB_STRAIN
H1:PSL-DIAG_BULLSEYE_PIT_OUT_DQ
H1:PSL-DIAG_BULLSEYE_YAW_OUT_DQ
H1:PSL-DIAG_BULLSEYE_WID_OUT_DQ
H1:IMC-WFS_A_DC_PIT_OUT_DQ
H1:IMC-WFS_B_DC_PIT_OUT_DQ
H1:IMC-WFS_A_DC_YAW_OUT_DQ
H1:IMC-WFS_B_DC_YAW_OUT_DQ
H1:ASC-DHARD_P_OUT_DQ
H1:ASC-DHARD_Y_OUT_DQ
H1:ASC-CHARD_P_OUT_DQ
H1:ASC-CHARD_Y_OUT_DQ
H1:LSC-CAL_LINE_SUM_DQ
H1:LSC-SRCL_IN1_DQ
H1:LSC-MICH_IN1_DQ
H1:LSC-PRCL_IN1_DQ
H1:PEM-EY_MAINSMON_EBAY_1_DQ
H1:PEM-EY_MAINSMON_EBAY_2_DQ
H1:PEM-EY_MAINSMON_EBAY_3_DQ
H1:CAL-CS_LINE_SUM_DQ
H1:CAL-PCALY_TX_PD_OUT_DQ
H1:CAL-PCALY_EXC_SUM_DQ
H1:SUS-ETMY_L3_CAL_LINE_OUT_DQ

- The network is able to discover the beam jitter coupling and subtract
- Some improvement at low frequency too
- We did not expect much non-linearity or non-stationarity in LHO/O2

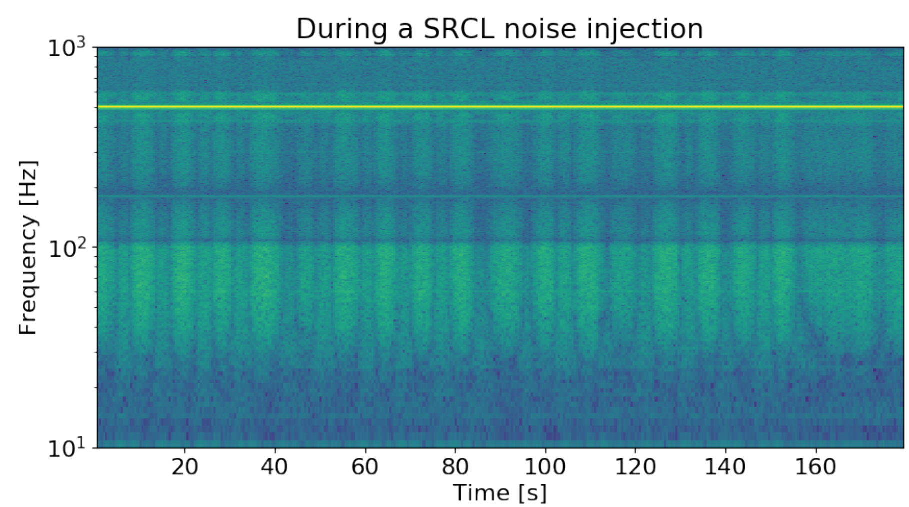
# Real data with known non-stationarity

- Back to (the future) 2015, non-stationary coupling of SRCL to DARM elogs 17912, 18026
- Dataset available for experiments (contact me)



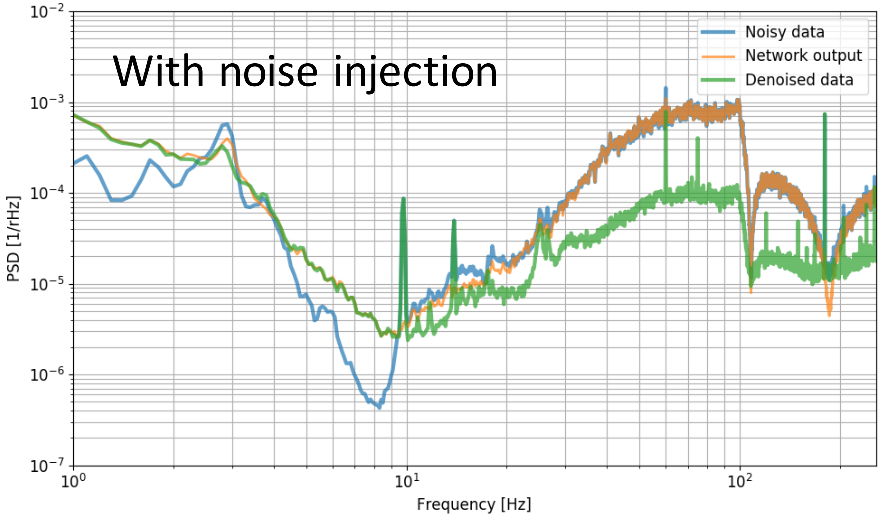

DARM during a quiet period: non stationary noise at low frequencies

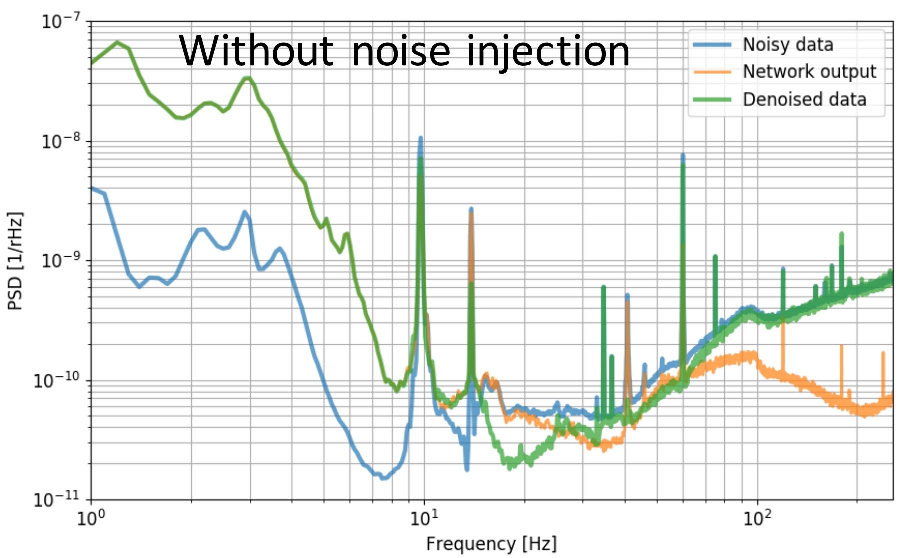


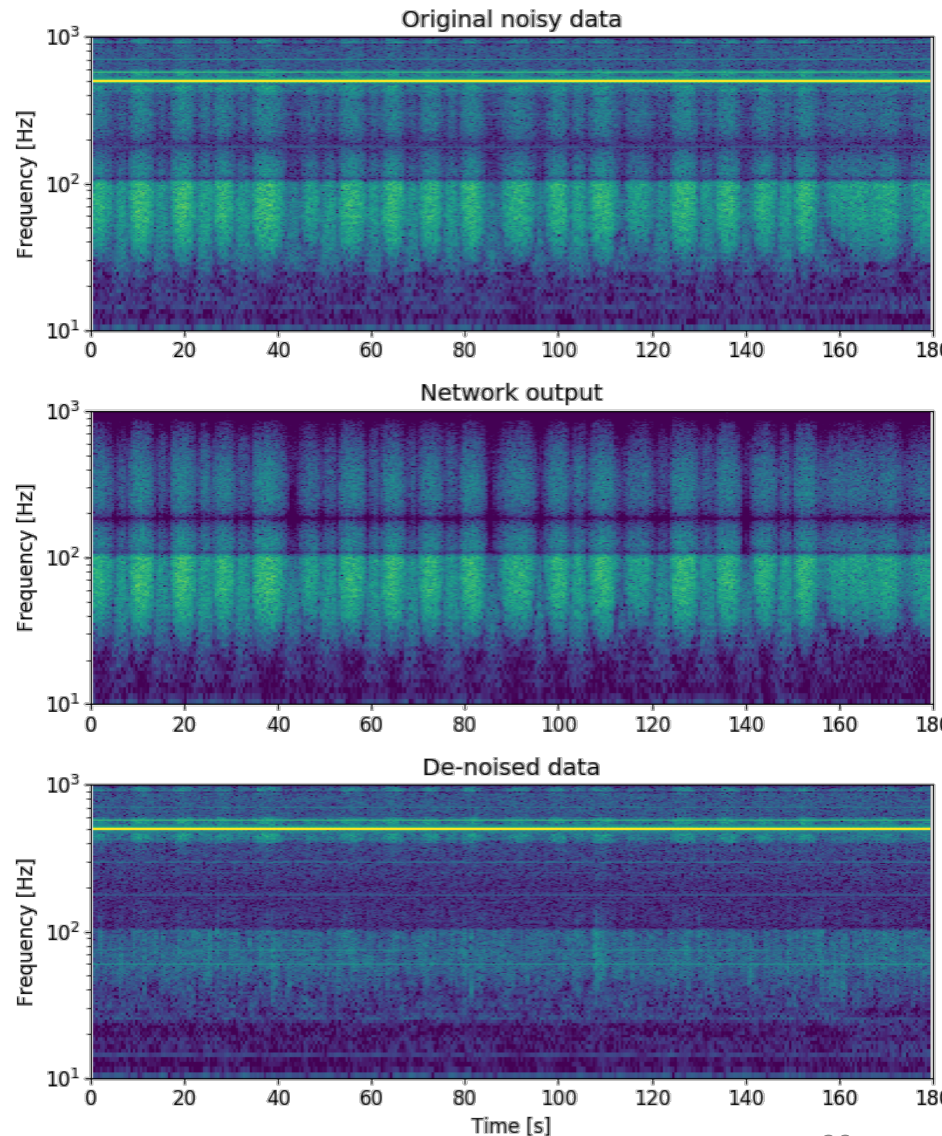

DARM during a SRCL (stationary) noise injection: coupling was modulated by SRC alignment fluctuations

# Results

## DARM signal [uncalibrated]



Jamie Rollins, Gabriele Vajente, Gautam Venugopalan

# STATE ESTIMATOR

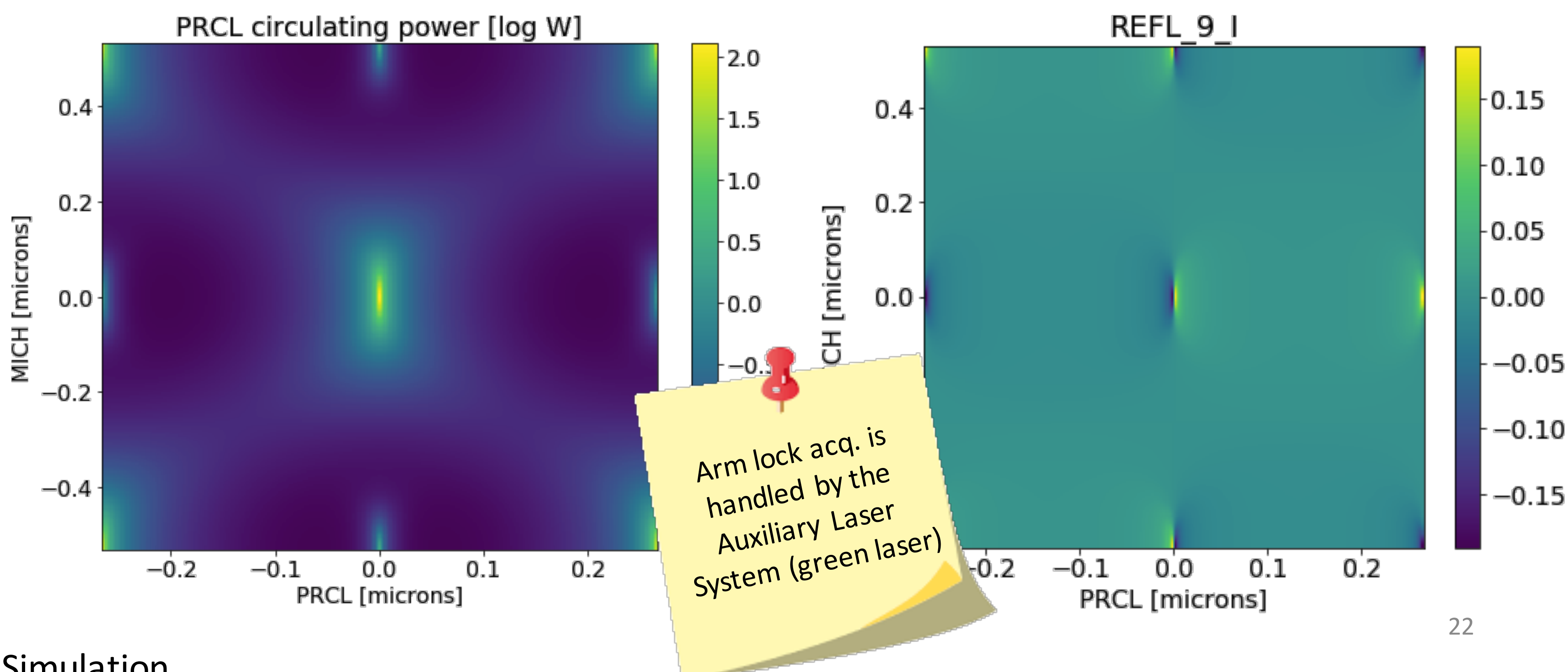
# Interferometers are highly non-linear*

*when they're not locked

- Highly non linear problem
- Alternative view: good linearized error signals exist in a small fraction of phase space
- Let's focus on PRMI for this talk (easy to represent in 2D plots, enough complexity to make it interesting)

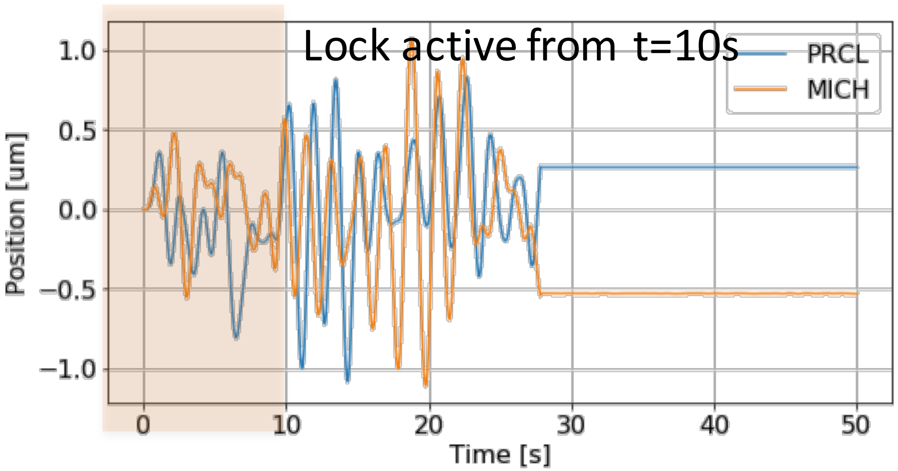
**DRMI:** **Dual Recycled Michelson Interferometer**

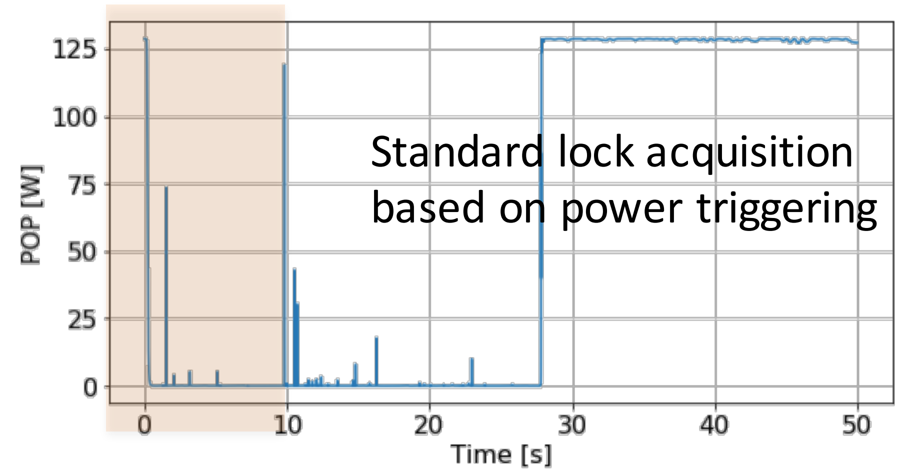
**PRMI:** **Power Recycled Michelson Interferometer**



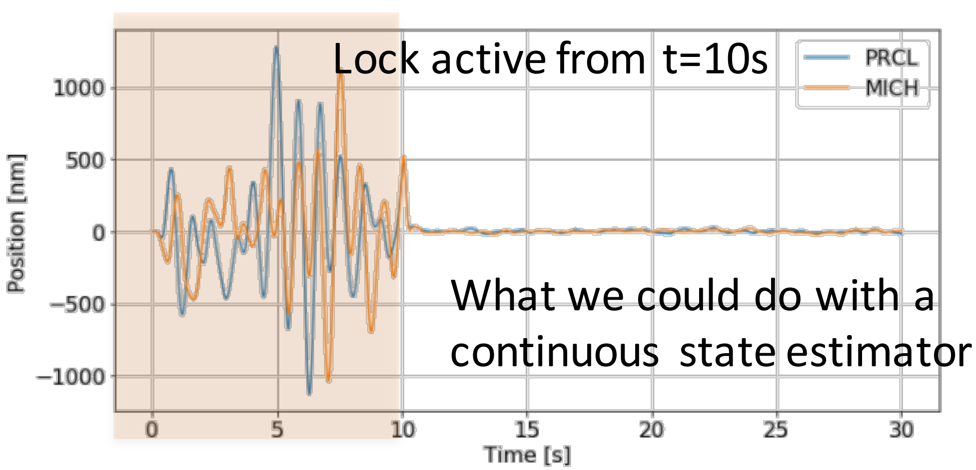

Simulation

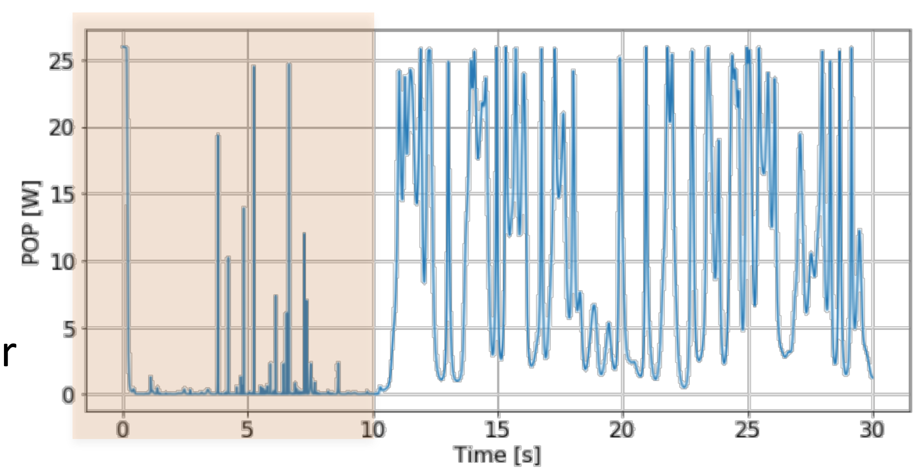
# If we knew the mirror position at all times…




*versus*

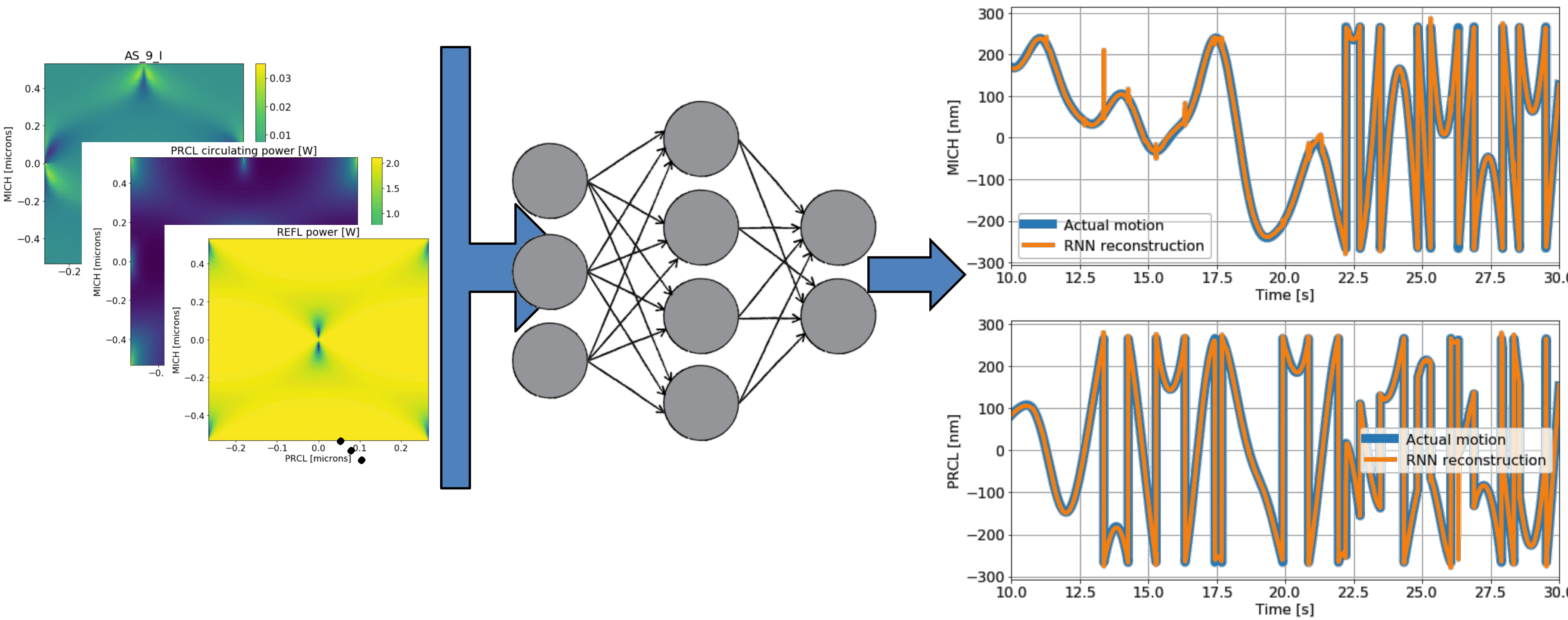



All plots are SIMULATIONS

# Use a Deep Recurrent Neural Network

- **Inputs**: optical error signals (POP_DC, REFL_DC, POP_1F_I/Q, etc...)
- **Outputs**: MICH, PRCL, etc.. positions

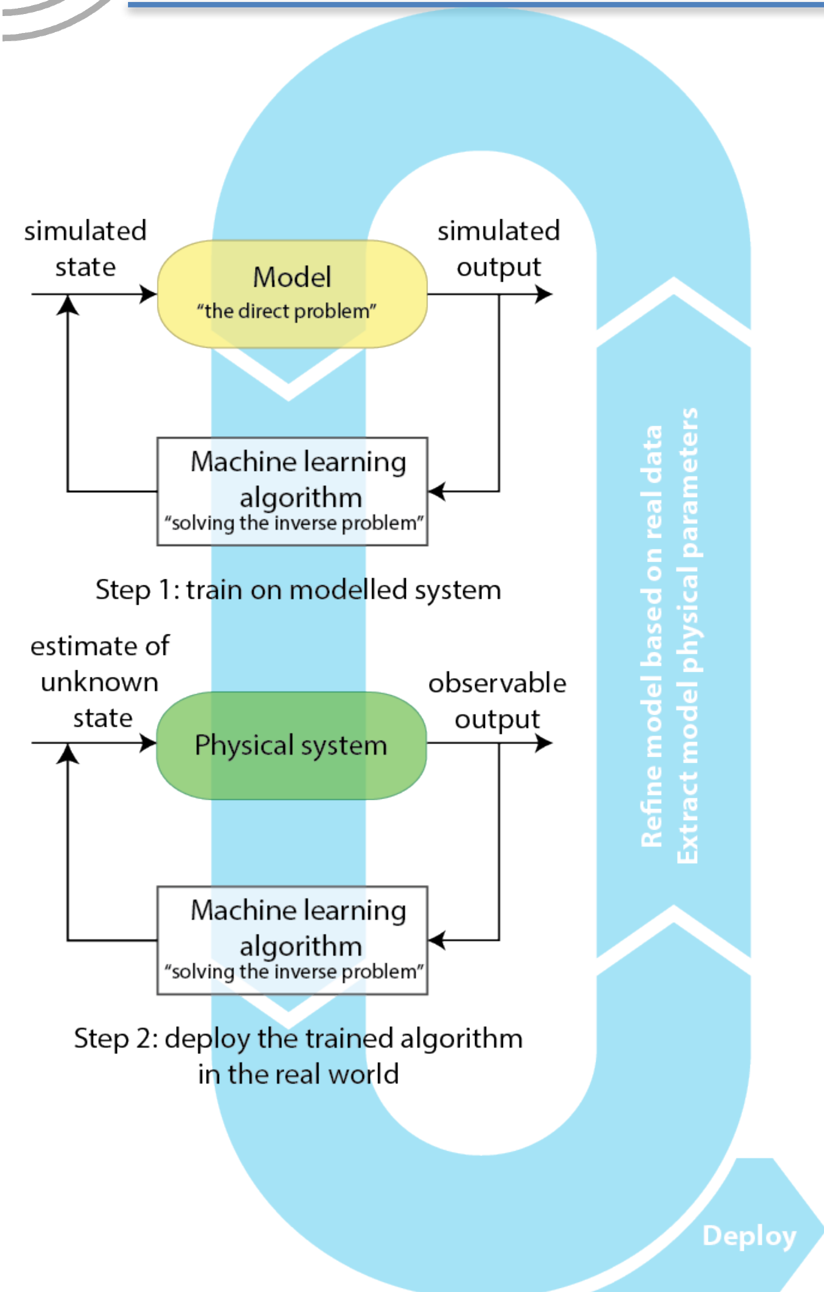



We are dealing with time series of signals: the instantaneous values are not enough to predict the MICH/PRCL positions.

We need to feed the network some past history of optical error signals:

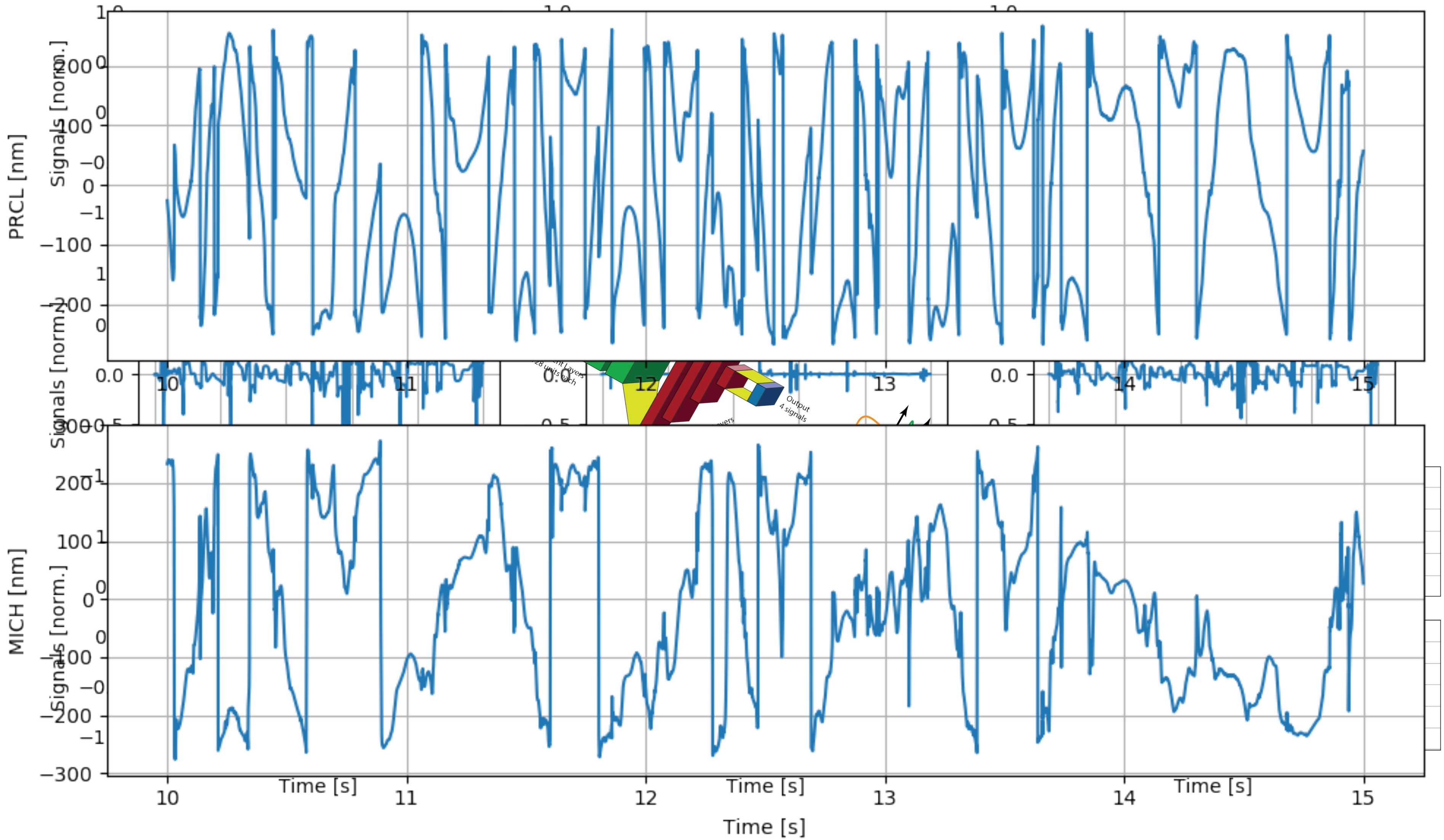
**Recurrent Neural Networks (RNN) maintain an internal memory**

# Train in simulation, deploy to real world



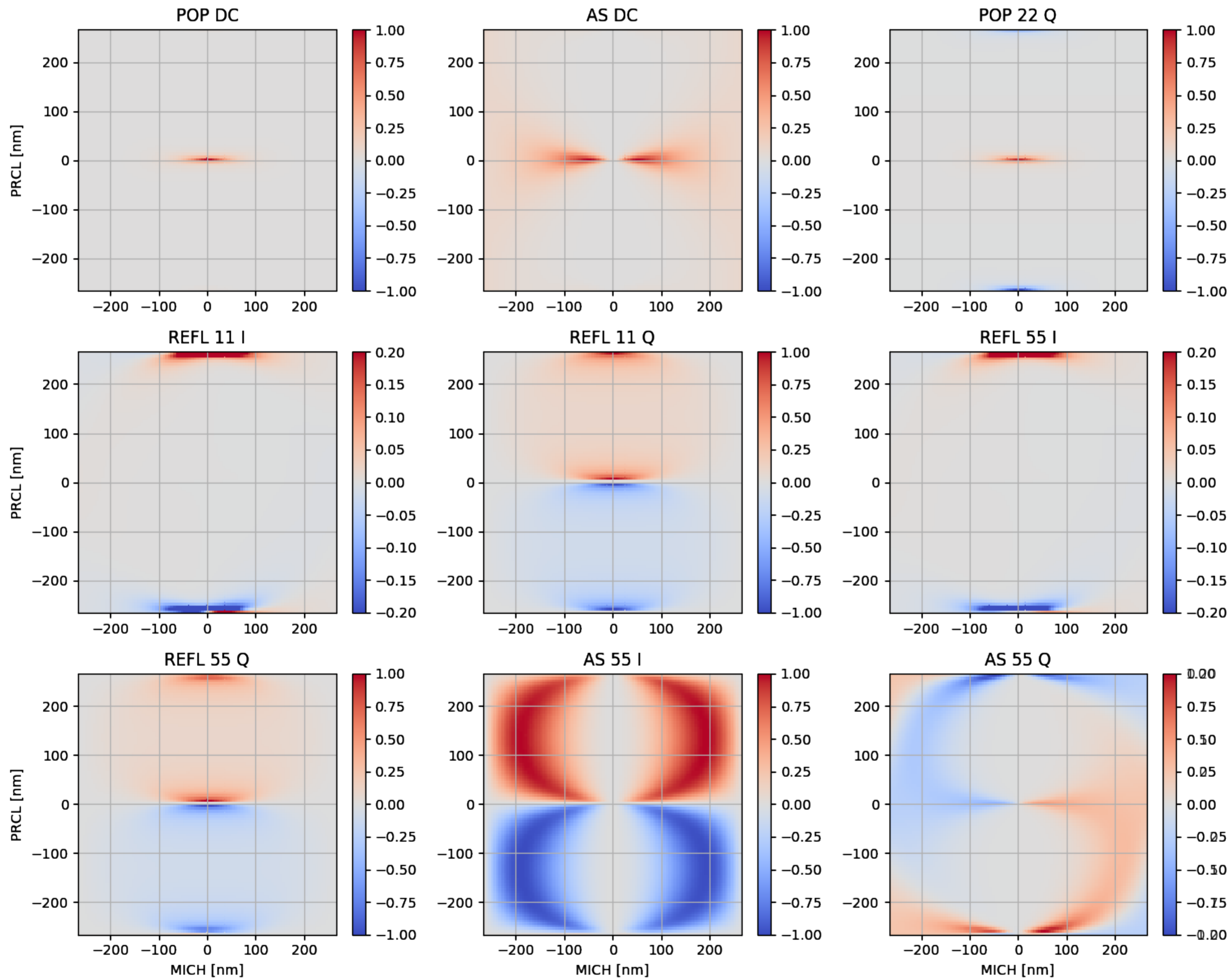

- Deep Neural Network learning needs a lot of training examples (**$10^5$ - $10^6$**)
- Not practical to do it online (and **we don't have the targets** in the real system!!)
- Use a **simulation** of the system as accurate as possible (including uncertainties)
- **Train on the simulated data**
- Deploy on the real system and test the performance
- Fine tune if needed

# Results



# Real data / Simulation

SIMULATION



# Conclusions

- Machine learning techniques especially Deep Learning, look very promising
- Our problems and applications are quite different from main stream Deep Learning
- Nowadays it's easy to implement them (lots of ready to use libraries
- **But Deep Learning or Machine Learning are not always the best tool for the job**


www.mrmoneymustache.com

# References

## Deep Learning introductions:

- www.deeplearningbook.org
- www.deeplearning.ai
- course.fast.ai
- MIT course on AI 6.034 (online)
- Stanford Machine Learning Course CS229 (online)
- A. Geron 'Hands-on Machine Learning with Scikit-Learn and TensorFlow" O'Reilly 2017
- TensorFlow: www.tensorflow.org
- PyTorch: www.pytorch.org

## Deep Learning for Lock Acquisition:

- G1701455 Talk at CSWG call 08/02/17
- G1701589 Talk at LVC meeting 08/28/17
- G1702072 Talk at CSWG call 10/19/17
- G1702213 Talk at MLA call 11/08/2017
- T1700466 Technical note "Deep Learning for Lock Acquisition"
- https://git.ligo.org/gabriele-vajente/machine-learning-lock-acquisition the actual code

## Noise subtraction:

- G1800334 talk at LVC meeting
- G1800589 talk at LVC meeting
- https://git.ligo.org/gabriele-vajente/dn2 [dn]$^2$ code
- https://git.ligo.org/rich.ormiston/DeepClean DeepClean code

## Beam spot position:

- G1800359 Talk at LVC meeting
- https://git.ligo.org/gabriele-vajente/beam-spot-centering The actual simulation and network code