Deep Loop Shaping: Reinforcement Learning Control for Gravitational-Wave Observatories

Tomislav Andric

4th EINSTEIN TELESCOPE ANNUAL MEETING
11-14/11/2025 OPATIJA

Overview: use neural networks for better feedback control

- In the early 2010s, discussions started at GWADW meetings about control noises as potential show-stoppers for future low-frequency sensitivity improvements
- Are there fundamental limits to the control performance?
- Linear control obeys fundamental constraints, e.g., derived from Kramers-Kronig relations. One can define optimal linear control.
- JH proposed reinforcement learning to evade the fundamental limitations.
- A first analysis was carried out as part of a PhD thesis submitted 2016 in Urbino.
- Shimmer Project started ~2020.
- Main aim was to have a stable HARD ASC loop with less noise injection > 10 Hz.
- Time domain simulation: Tomislav Andric, Jan Harms -> LightSaber.
- Linearized time-domain simulation: Chris Wipf - RT SimPlant
- Neural Network training: DeepMind / Caltech

Setup

- A GW detector is a system of optomechanical degrees of freedom (e.g., mechanical suspensions, beam phase/alignment/shape, laser amplitude and frequency)
- Something like 100 degrees of freedom need to be controlled, and there is an important coupling between most of them

Goals

- Laser interferometer must be operated as close as possible to its ideal state
- Low-frequency motion must be strongly reduced
- System must remain stable
- Noise injected by the controller must be minimized

Science
Current Issue First release papers Archive About

HOME > SCIENCE > VOL. 389, NO. 6764 > IMPROVING COSMOLOGICAL REACH OF A GRAVITATIONAL WAVE OBSERVATORY USING DEEP LOOP SHAPING

RESEARCH ARTICLE | ASTROPHYSICS

Improving cosmological reach of a gravitational wave observatory using Deep Loop Shaping

JONAS BUCHLI, BRENDAN TRACEY, TOMISLAV ANDRIC, CHRISTOPHER WIPF, YU HIM JUSTIN CHIU, MATTHIAS LOCHBRUNNER, CRAIG DONNER, RANA X. ADHIKARI, JAN HARMS, [...], AND THE LIGO INSTRUMENT TEAM +21 authors Authors Info & Affiliations

SCIENCE • 4 Sep 2025 • Vol 389, Issue 6764 • pp. 1012-1015 • DOI: 10.1126/science.adw1291

Editor's summary

Gravitational wave detectors have revolutionized astrophysics by detecting black holes and neutron stars. Most signals are captured in the 30- to 2000-Hz range, and the lower 10- to 30-Hz band remains largely unexplored because of persistent low-frequency control noise that limits sensitivity. Enhancing this sensitivity could increase cosmological reach. Using nonlinear optimal control through reinforcement learning with a frequency-domain reward, Buchli *et al*. developed a method that effectively reduces control noise in the low-frequency band. This method was successfully implemented at the Laser Interferometer Gravitational-Wave Observatory (LIGO) in Livingston and the Caltech 40 Meter Prototype, achieving control noise levels on LIGO's most demanding feedback control loop below the quantum noise, thus removing a critical obstacle to increased detector sensitivity. —Yury Suleymanov

Contributors


Jan Harms (GSSI)


Tomislav Andric (GSSI)


Christopher Wipf (Caltech)


Rana Adhikari (Caltech)


Jameson Rollins (Caltech)


Joseph Betzwieser (Caltech)


Abbas Abdolmaleki


Iain Barr


Charlie Beattie


Jonas Buchli


Serkan Cabi


Justin Chiu


Jonas Degrave


Yuzhu Dong


Craig Donner


Oliver Groth


Roland Hafner


Sandy Huang


Andrea Huber


Pushmeet Kohli


Matthias Lochbrunner


Hannah Openshaw


Martin Riedmiller


Gregory Thornton


Brendan Tracey


George van den Driessche


Markus Wulfmeier

Partner institutions

Caltech

GSSI

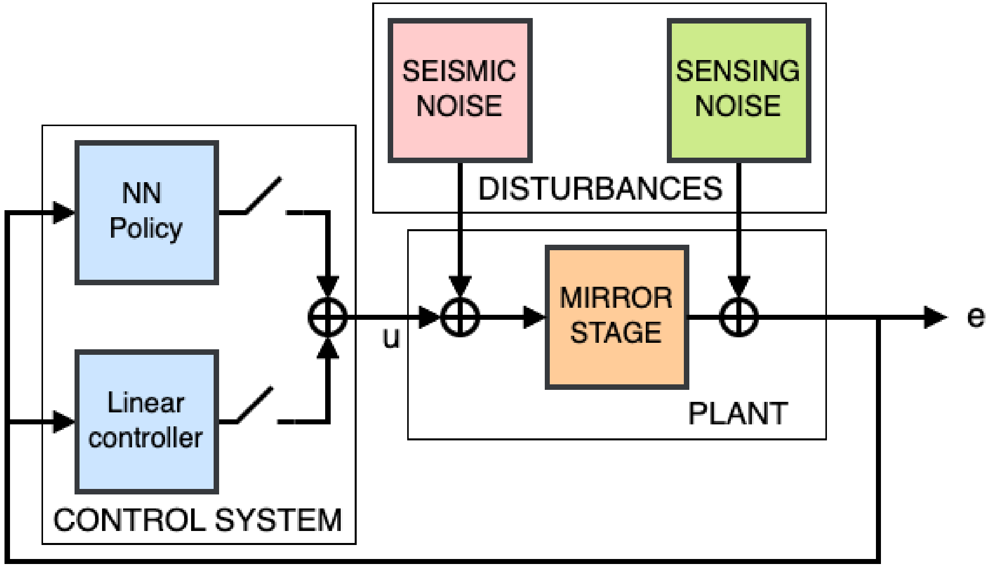
DeepMind

LIGO

+ The LIGO Instrument Team

Outline

1. Project Overview
    1. Reduce controls noise
        1. Comparison of LLO linear, and this work
    2. Working with Google DeepMind
        1. History / NDA
2. Main Results
    1. (40m IMC)
    2. LLO ASC
3. Summary
4. Future Work



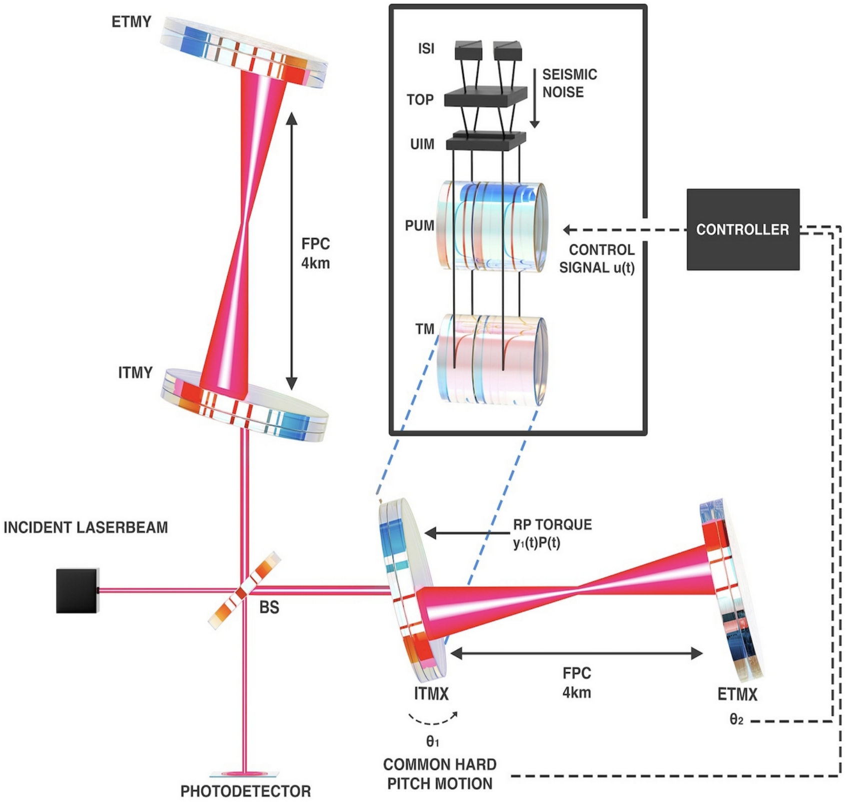
Mirror stabilization problem (ASC)

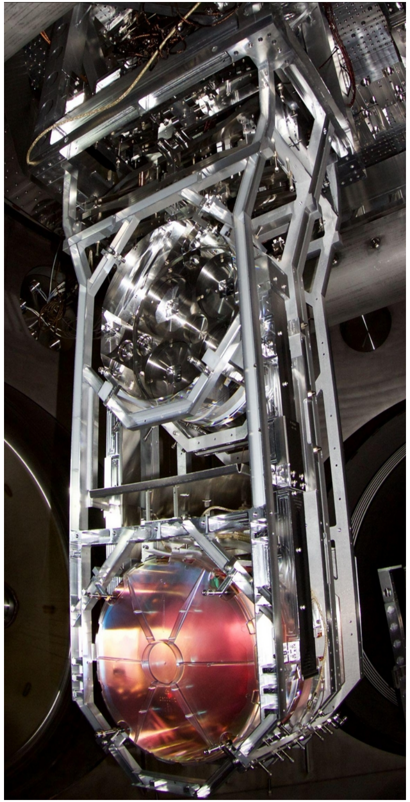
4-stage pendula, controlled by EM actuators

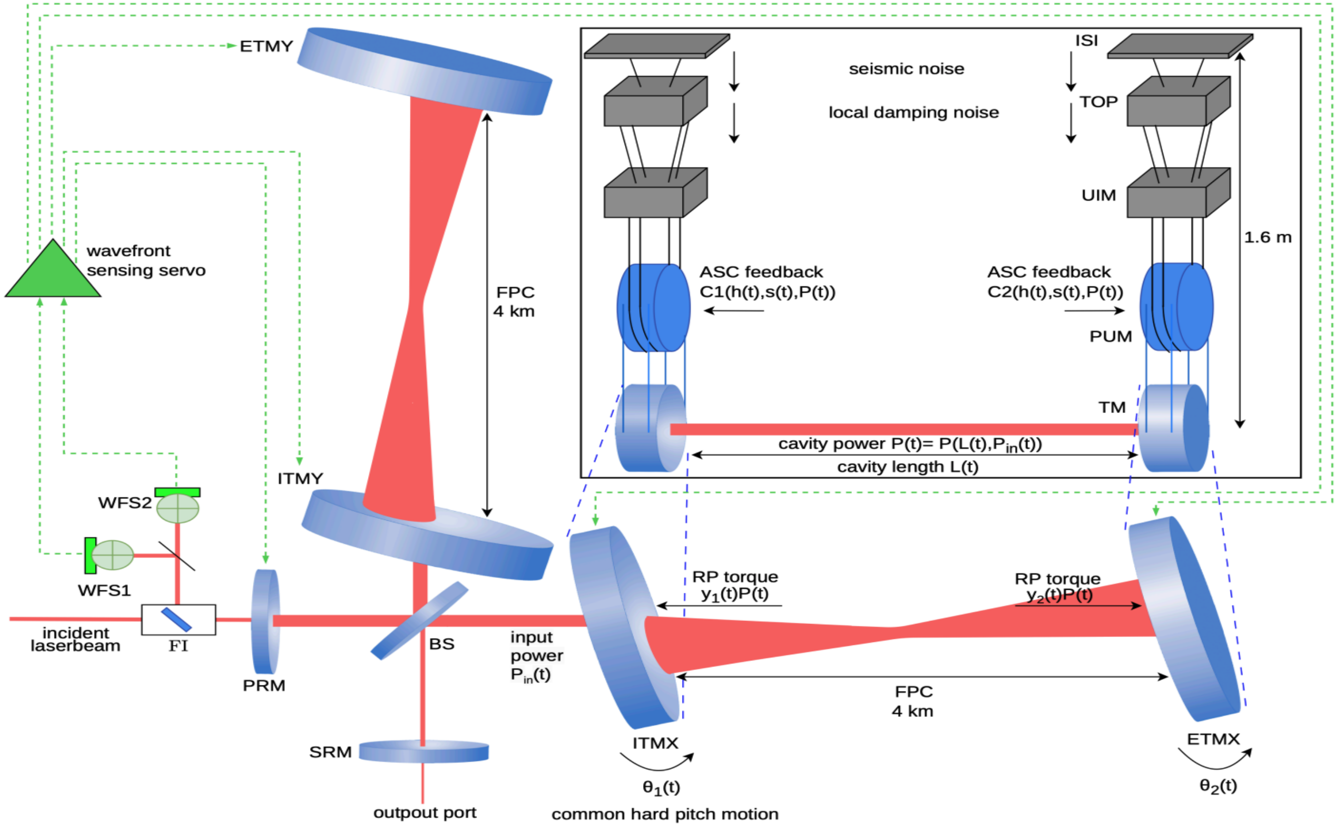
Main control challenges:
- Disturbance rejection: Seismic and other perturbations and noise coupling in through support
- Many eigenmodes ~0.5-2Hz

Control authority required up to ~3Hz

10-30Hz Observation band SNR requirement: angle error <<10^{-15}rad (femto rad)







ETMY
ISI
seismic noise
TOP
local damping noise
UIM
1.6 m
wavefront sensing servo
ASC feedback C1(h(t),s(t),P(t))
ASC feedback C2(h(t),s(t),P(t))
FPC 4 km
PUM
TM
cavity power P(t)= P(L(t),P_in(t))
cavity length L(t)
ITMY
WFS2
WFS1
RP torque y_1(t)P(t)
RP torque y_2(t)P(t)
incident laserbeam
FI
BS
input power P_in(t)
PRM
FPC 4 km
SRM
ITMX
ETMX
θ_1(t)
θ_2(t)
outpout port
common hard pitch motion

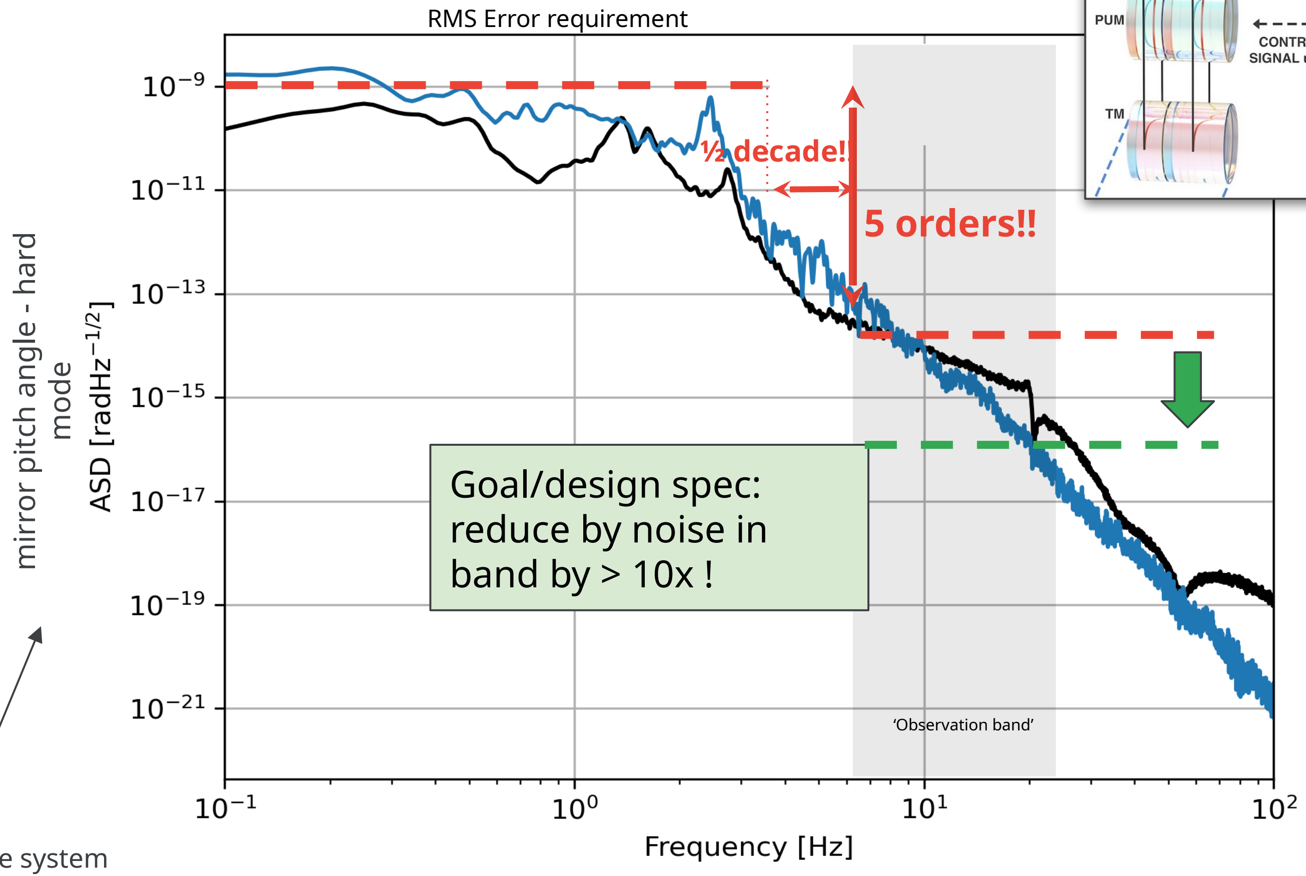
Mirror stabilization problem (ASC)

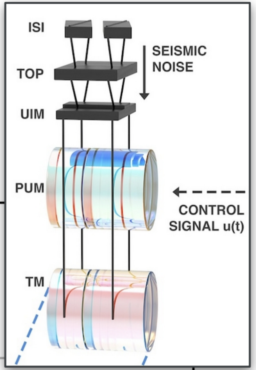
Main control challenges:

- Disturbance rejection: Seismic and other perturbations and noise coupling in through support
- Many eigenmodes around 0.5-2Hz

Control authority required up to ~3Hz

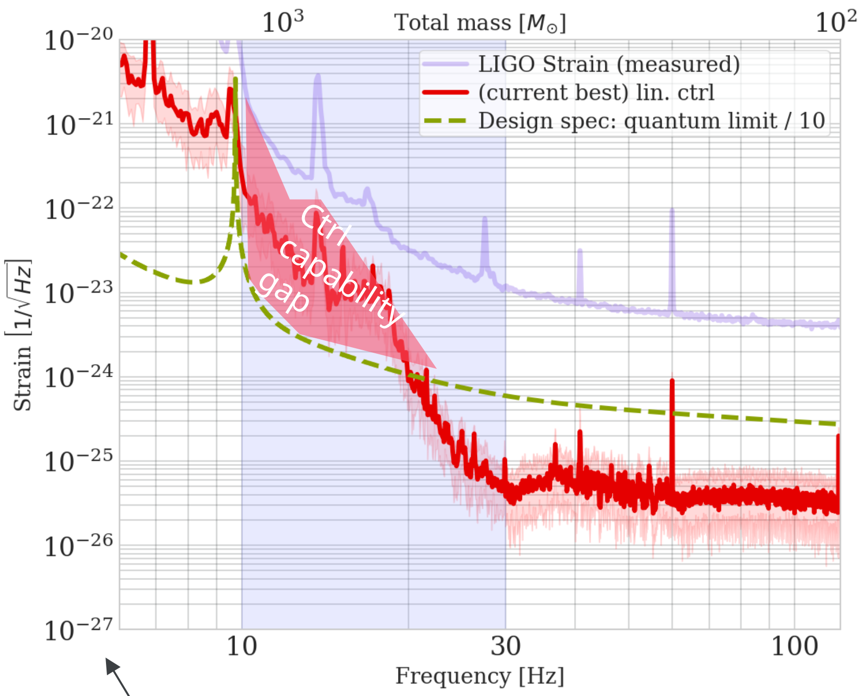
DARM SNR requirement
-> Angle error should roll off >200db/dec (>10 orders/dec) !!!





Different coordinate system

Current control is significant source of noise



Logarithmic scale!

The angular control noise is now one of main remaining blocker for increased low frequency sensitivity of LIGO!

Any reduction of the control noise will have a huge scientific impact.

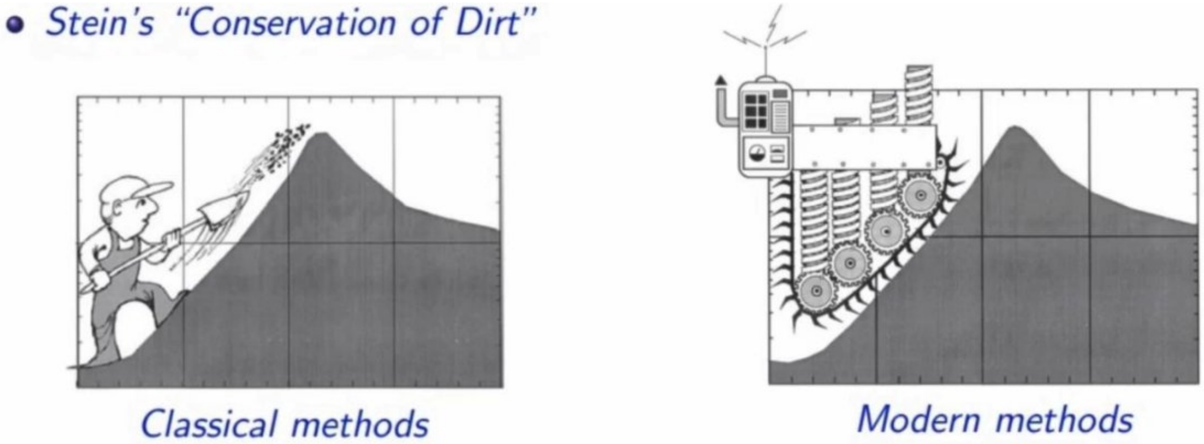
Big question: How can we improve the controller satisfying both stability and observation band performance?

Success criteria:

Improve the noise floor in the frequency range 10-30Hz for the LIGO system by at least one order of magnitude when compared to the currently used controllers.

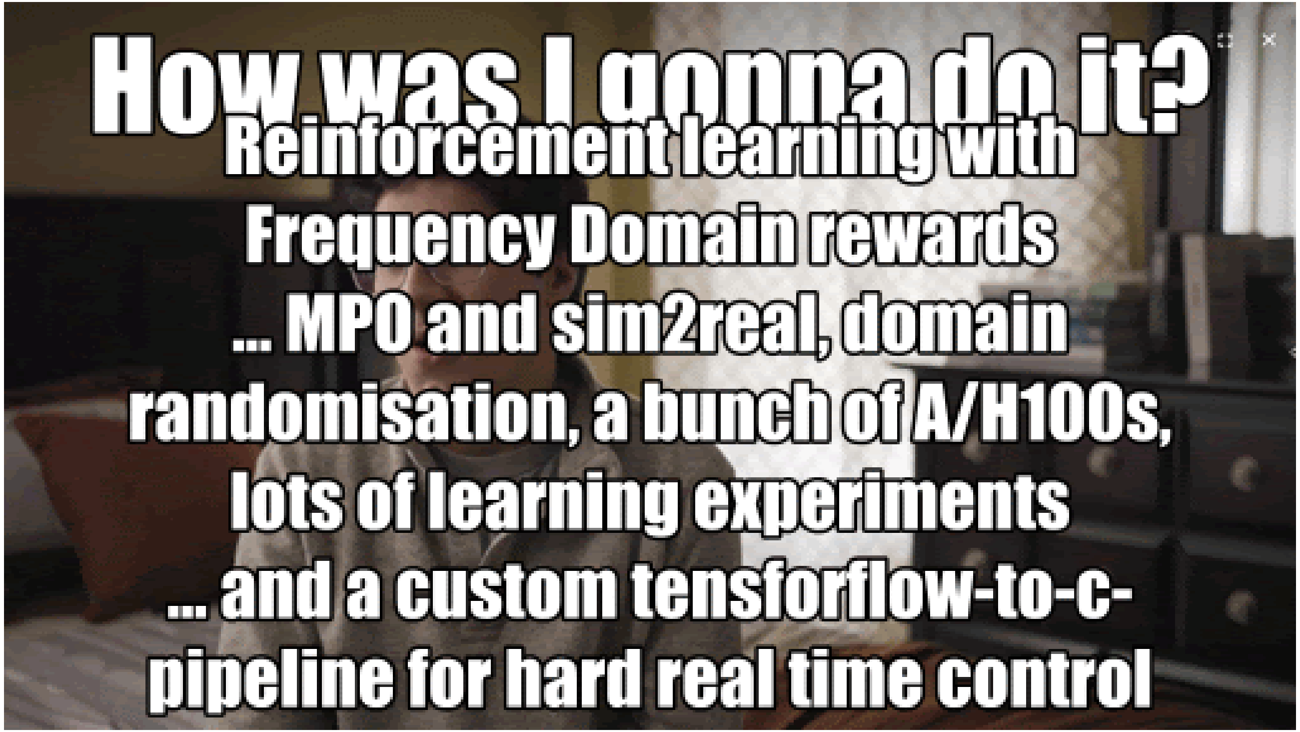
Bode’s sensitivity integral - the waterbed effect

$$\int_0^\infty \ln|S(j\omega)|d\omega = \int_0^\infty \ln\left|\frac{1}{1+L(j\omega)}\right|d\omega = \pi\sum Re(p_k) - \frac{\pi}{2}\lim_{s\to\infty} sL(s)$$



JSS, "Performance limitations in sensitivity reduction for nonlinear plants", *Systems & Control Letters*, 1991.
Stein, "Respect the unstable", *IEEE Control Systems Magazine*, 2003.
Chen, Fang, Ishii, "Fundamental limitations and intrinsic limits of feedback", *Annual Reviews in Control*, 2019.


How was I gonna do it?
Reinforcement Learning.

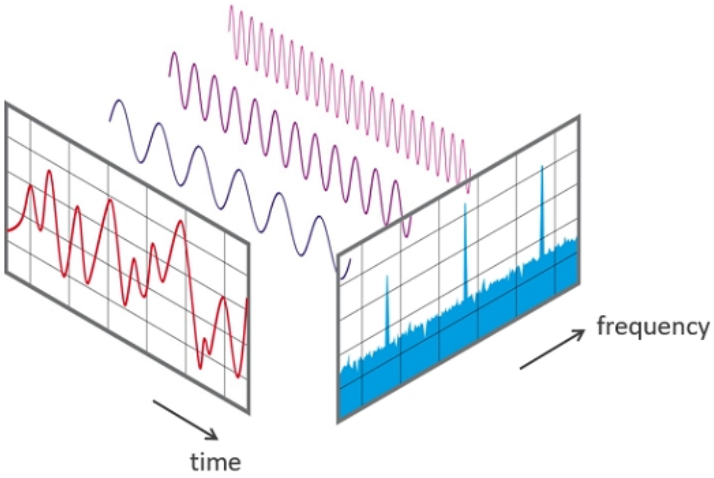

How was I gonna do it?
Reinforcement learning with
Frequency Domain rewards
... MPO and sim2real, domain
randomisation, a bunch of A/H100s,
lots of learning experiments
... and a custom tensorflow-to-c-
pipeline for hard real time control

Deep Loop Shaping

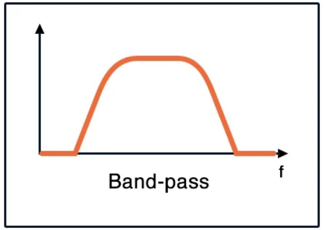
‘Reinforcement Learning with Frequency Domain Rewards’

Frequency domain rewards

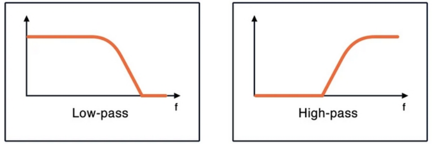
Basic idea: Formulate rewards in the frequency domain



Example: Use a bandpass pass filter to estimate the energy in a certain band and use this to compute a score



Filters have advantage that they can be run straightforwardly in 'online' scoring setups



Could also use (short-term) FFT or similar methods

Frequency domain rewards for LIGO ASC

Stability: Low frequency penalty

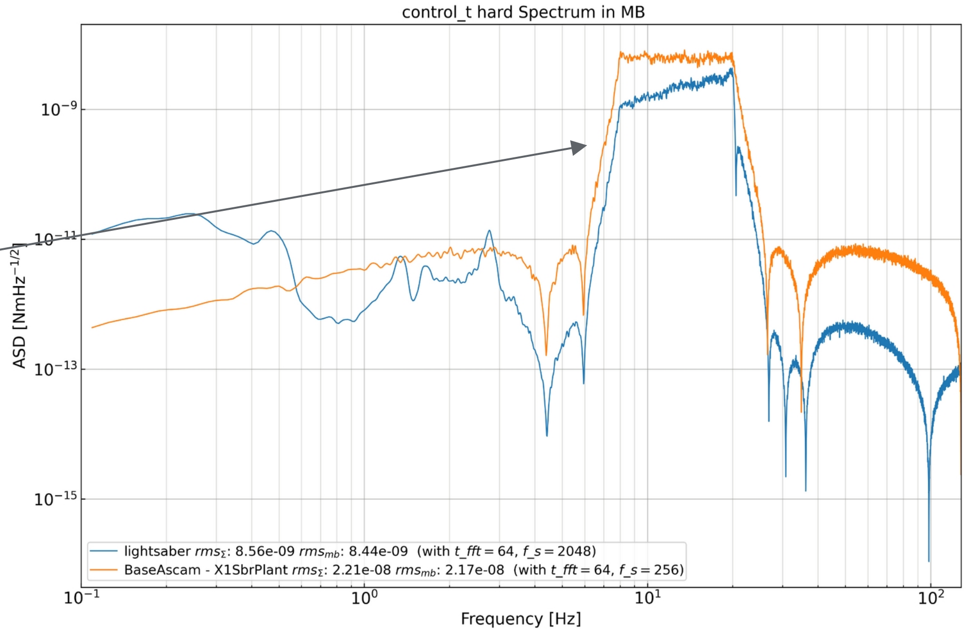
Reward aimed at reducing the overall magnitude (‘RMS’) of the error signal, encouraging stabilization. Computed using a lower pass filter with 3Hz corner frequency.

Sensitivity: Measurement band penalty

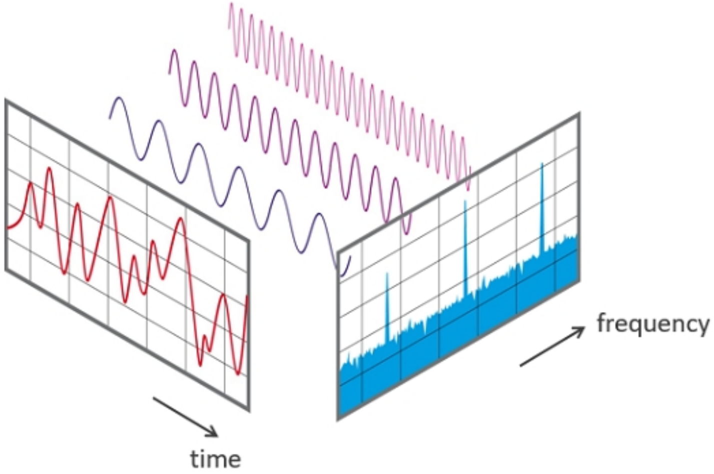
Reward aimed at reducing the control signal in the ‘measurement band’, i.e. the band of interest for GW signals. This is computed as a stepwise reward using an IIR.

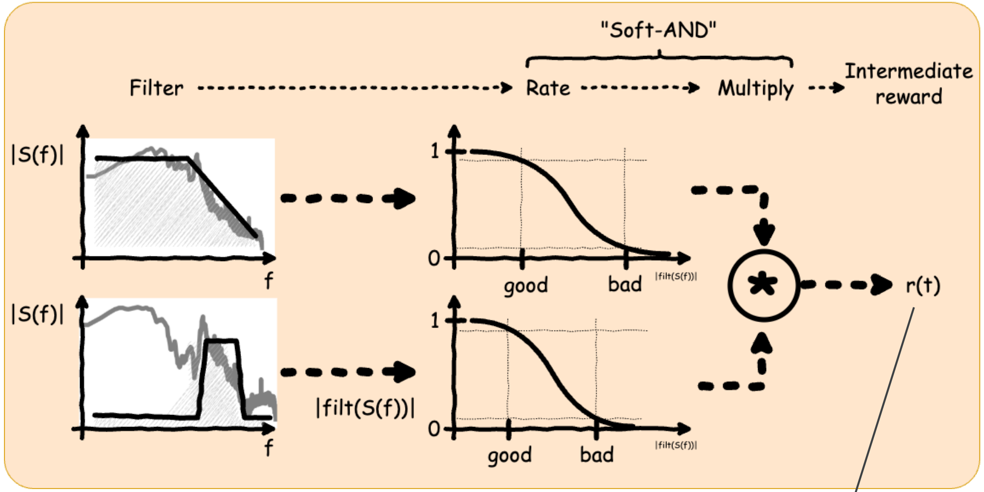
Regularization: High frequency penalty

Reward aimed at reducing the control signal at ‘high frequencies’ (i.e. above the measurement band). This is to keep the agent from introducing high frequency artifacts in that band.



Frequency domain rewards closed loop control design: Deep Loop Shaping

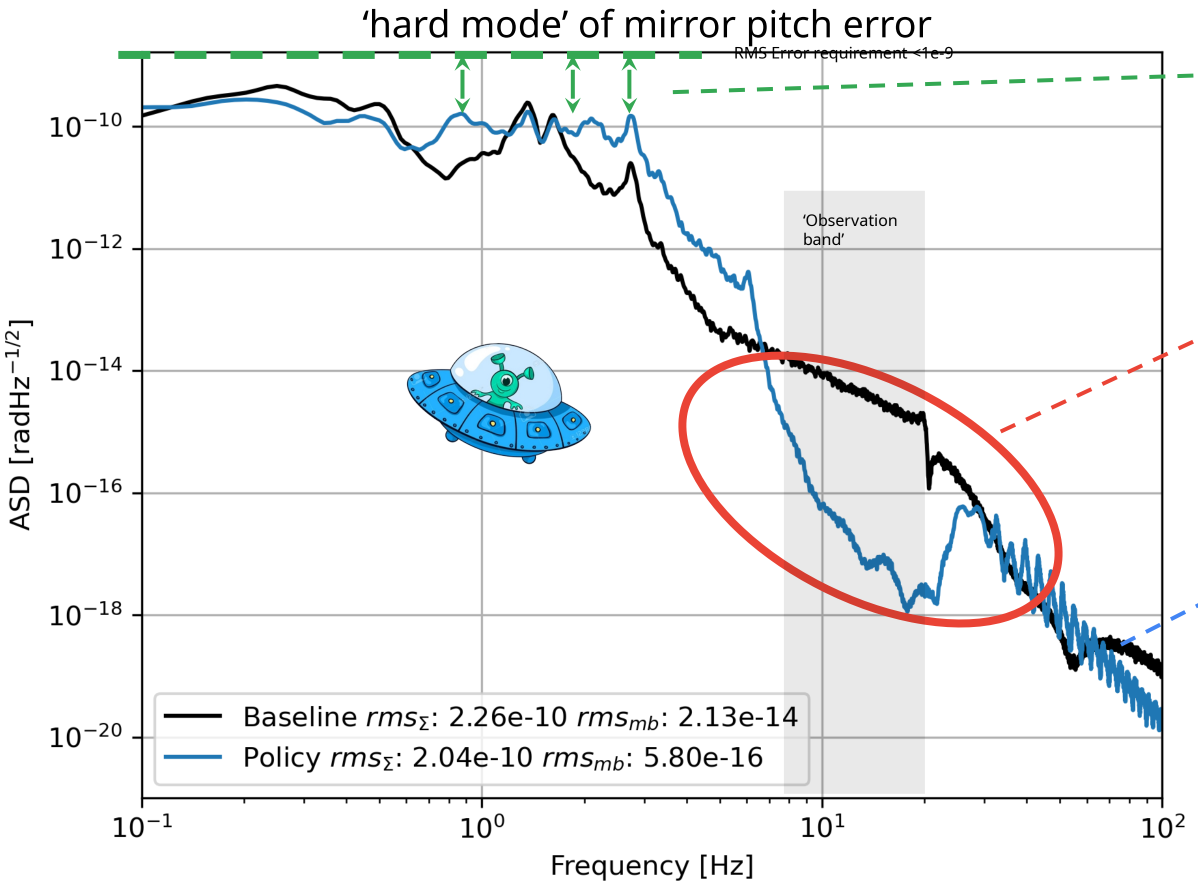




Deep Loop Shaping is a general RL/control design method!

$$\pi^* = \arg\max_{a \sim \pi} J = \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t, s_{t+1}) \right]$$

Frequency domain rewards FTW



SIMULATION RESULT - SEPTEMBER 2022

Stability: Low frequency penalty

Reward aimed at reducing the overall magnitude ('RMS') of the error signal, encouraging stabilization. Computed using a lower pass filter with 3Hz corner frequency.

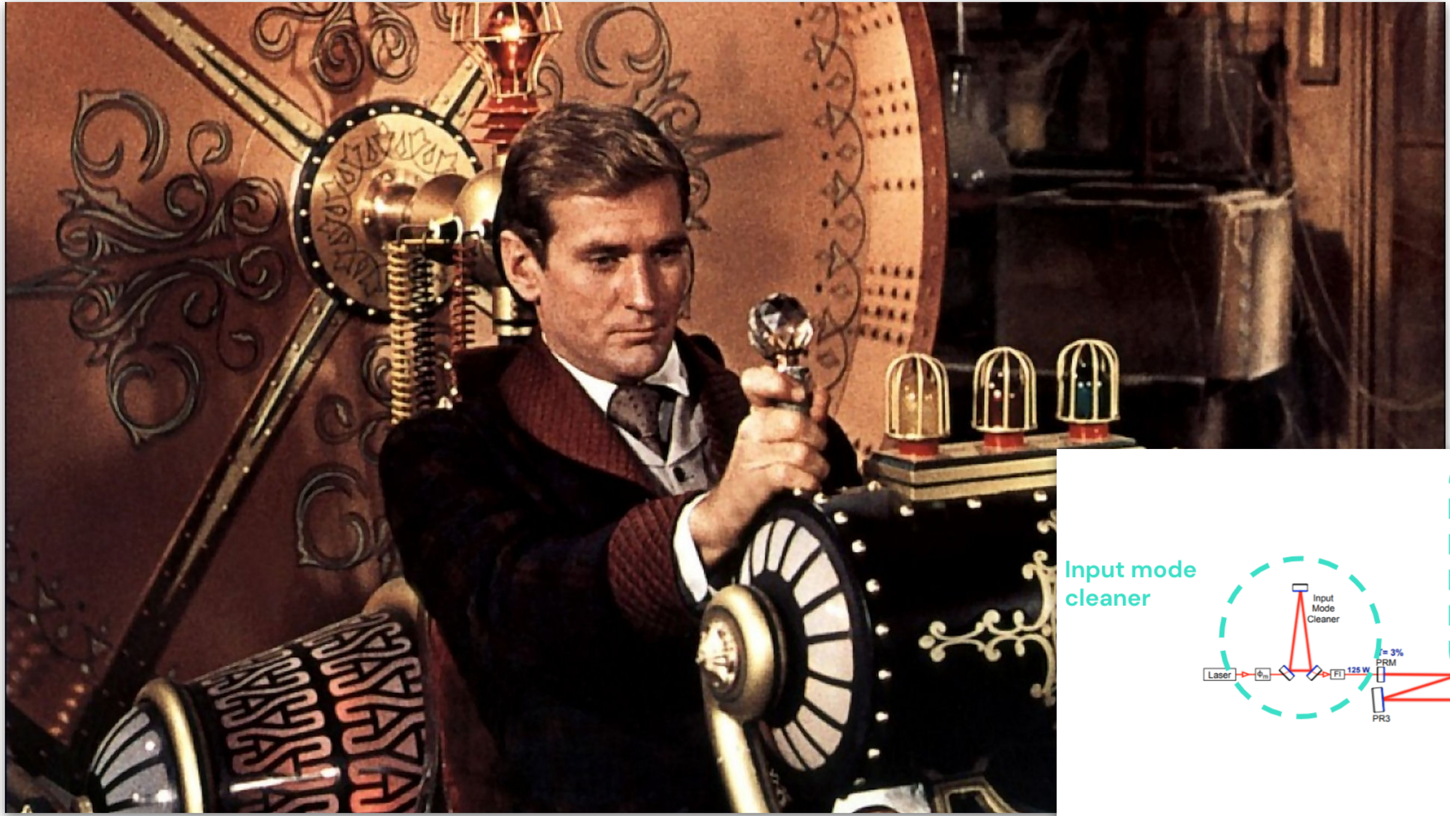
Sensitivity: Measurement band penalty

Reward aimed at reducing the control signal in the 'measurement band', i.e. the band of interest for GW signals. This is computed as a stepwise reward using an IIR.

Regularization: High frequency penalty

Reward aimed at reducing the control signal at 'high frequencies' (i.e. above the measurement band). This is to keep the agent from introducing high frequency artifacts in that band.

A ‘few experiments’ later… fast forward to 2023



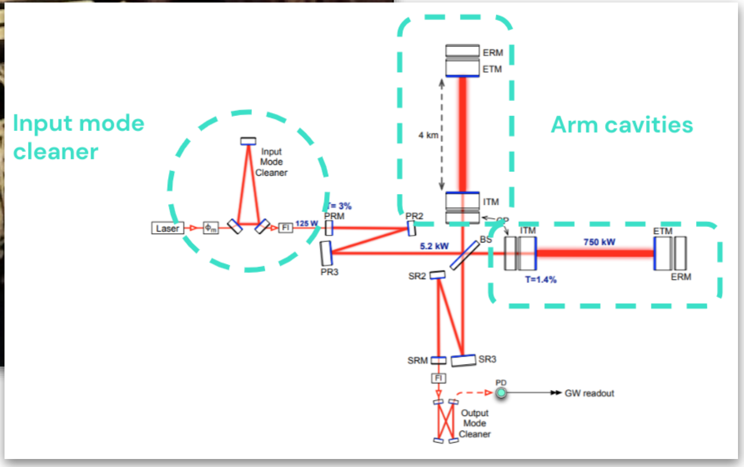
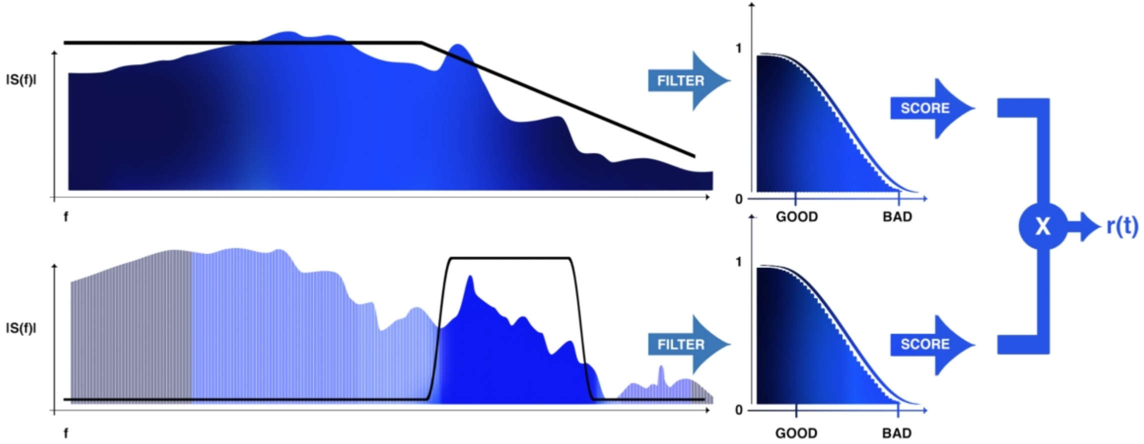


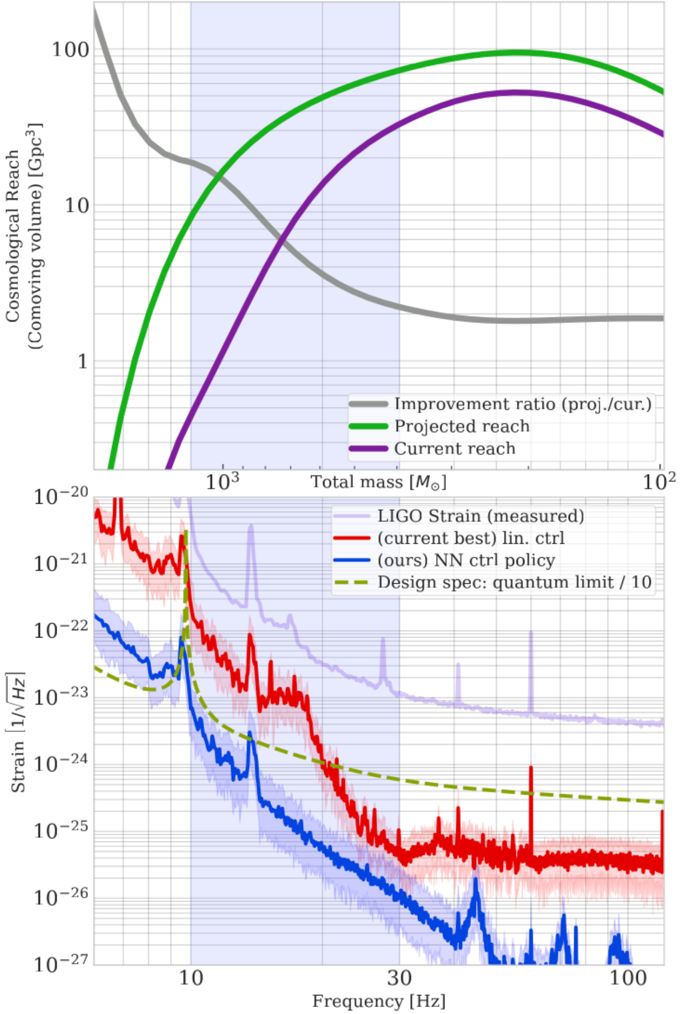
Illustration of method



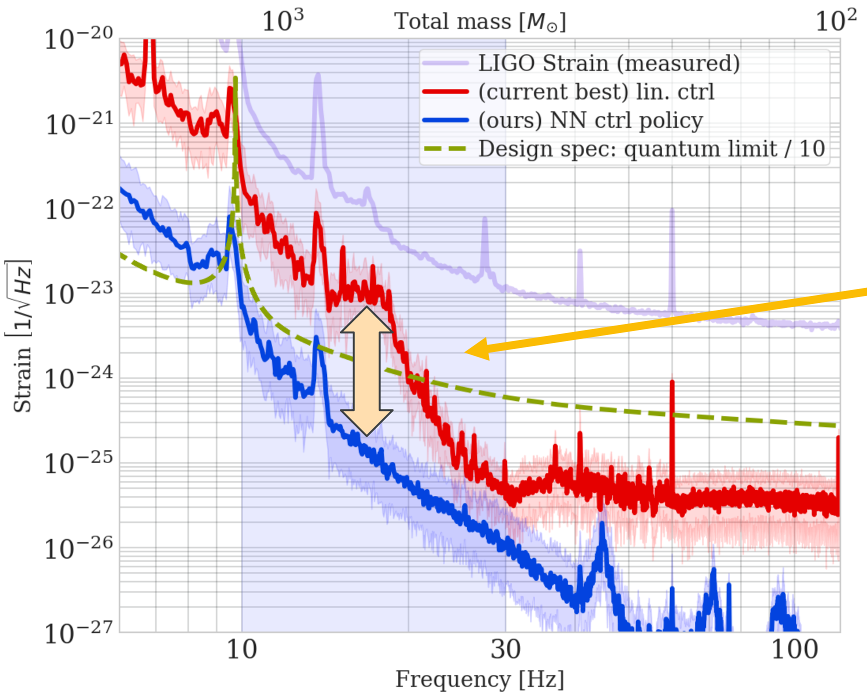
Key result: RL policy reduces noise up to 100x

Main Results:

1. Very good performance with linearized simulations
2. Good performance at LLO
3. Mostly short ~15 min tests.
4. Longer runs - no issues.



Key result: RL policy reduces noise up to 100x



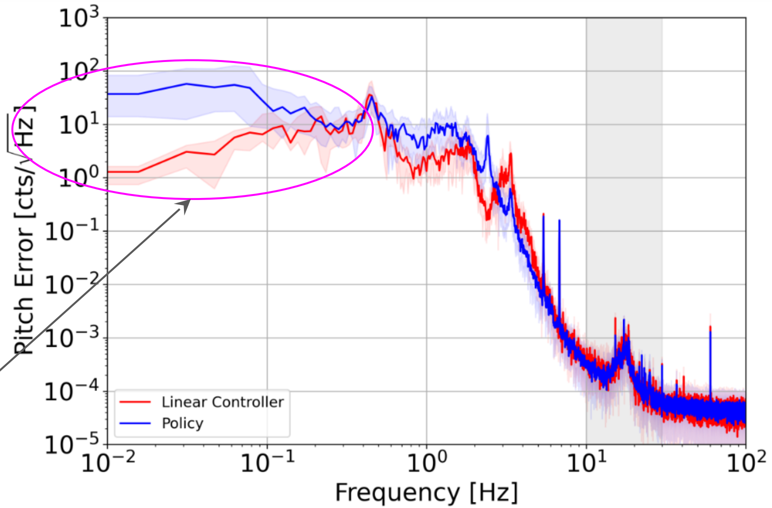
Success criteria:

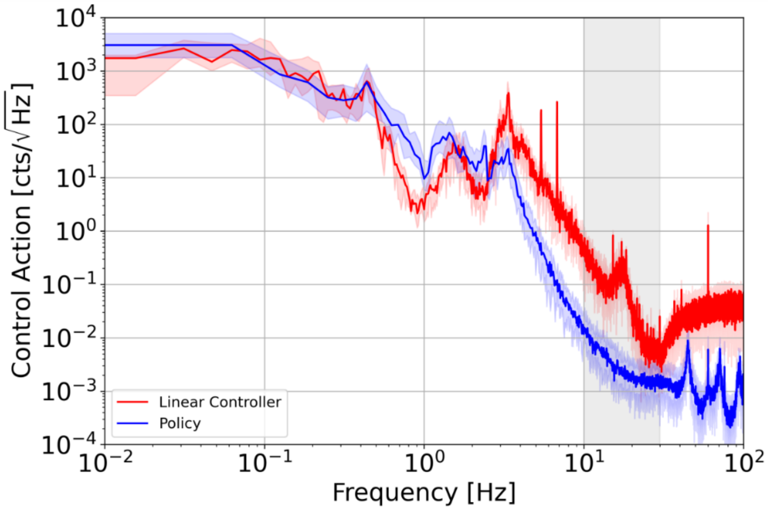
Improve the noise floor in the frequency range 10-30Hz for the LIGO system by at least one order of magnitude when compared to the currently used controllers.

We have removed the ASC control challenge as a blocker for increased low frequency sensitivity of LIGO 'in principle'

Fully implementing and operationalize this concept will have a huge scientific impact

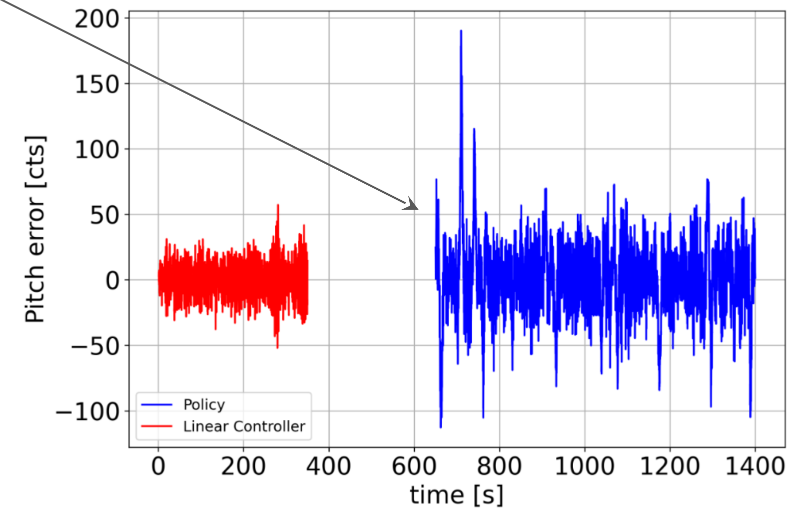
squidward - LLO 5.12.2024

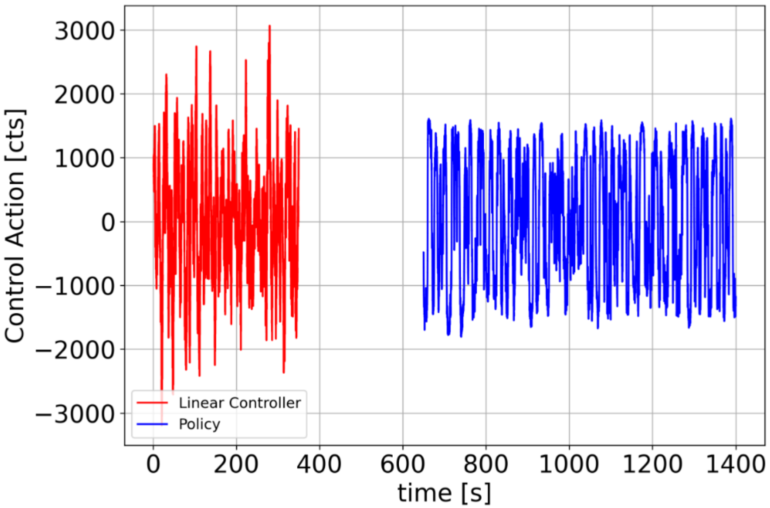




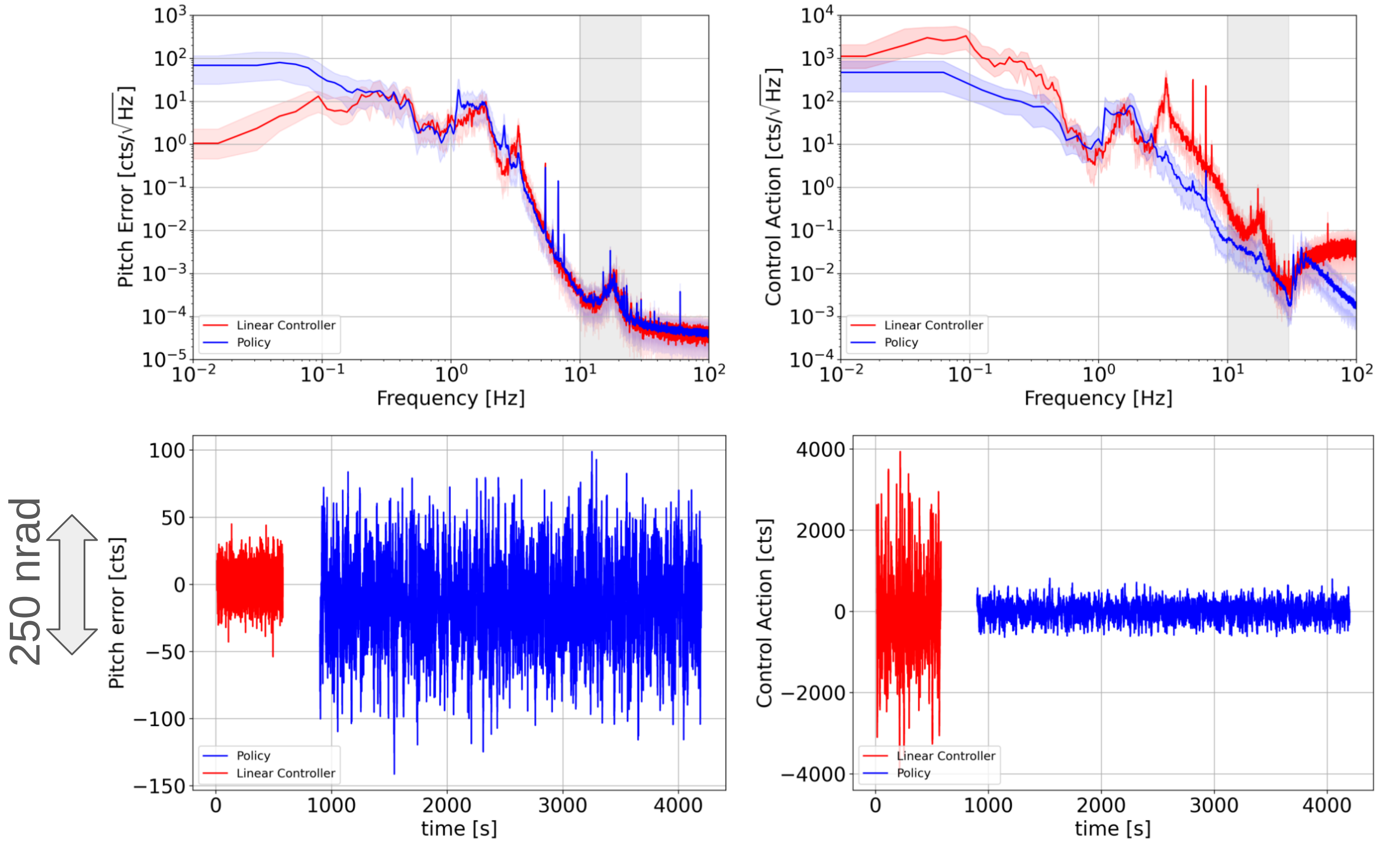
Low Freq Excess

- Needs better low freq SysID (L2/L1)

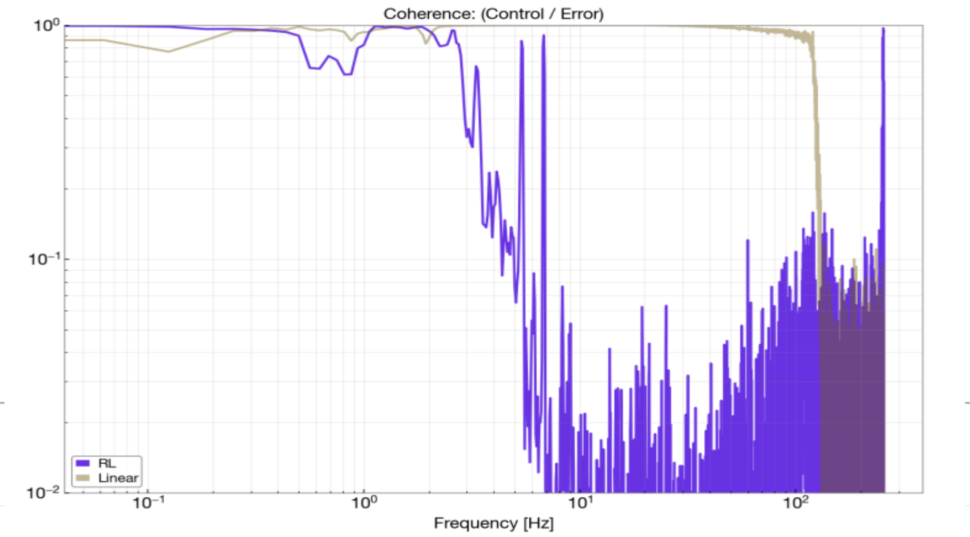


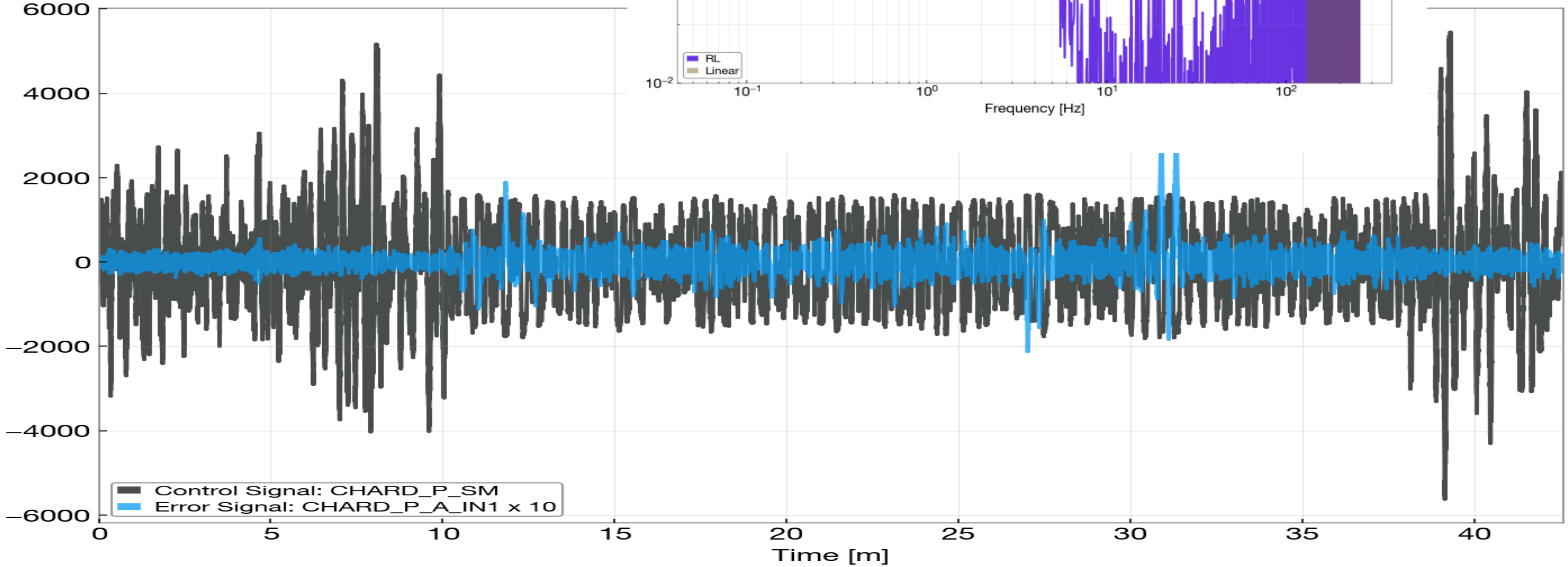


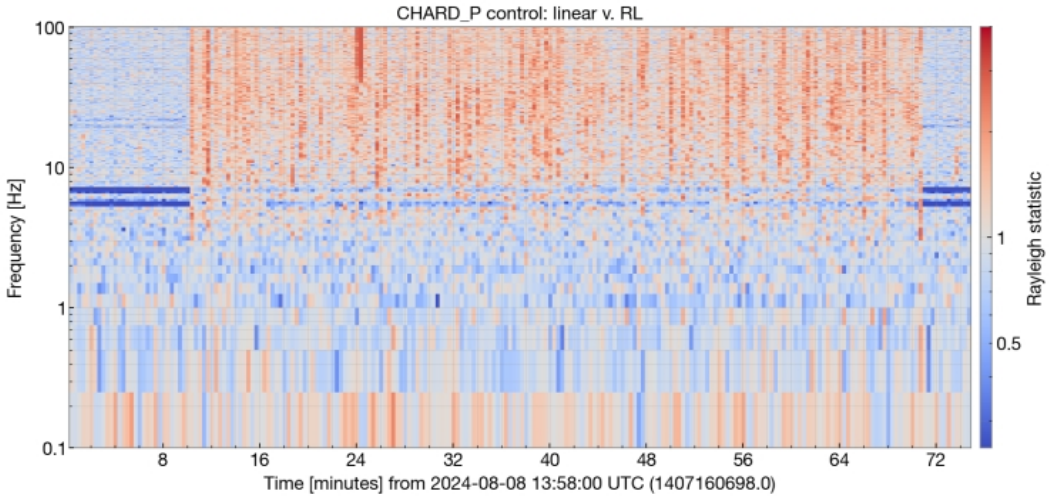
spongebob - 26.4.2024

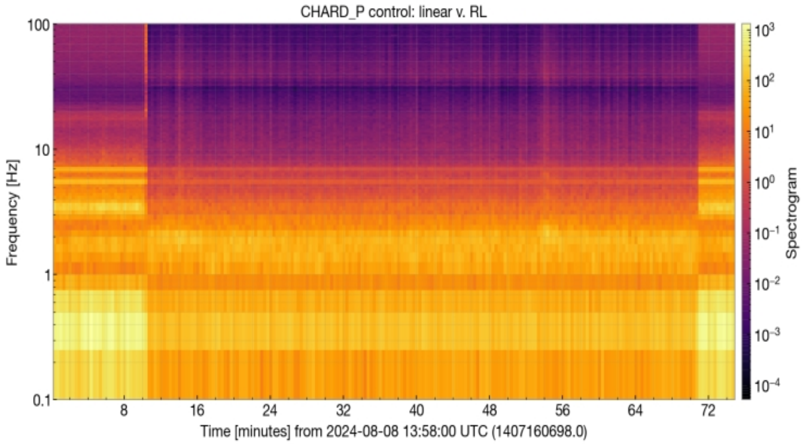


Coherence, Time series





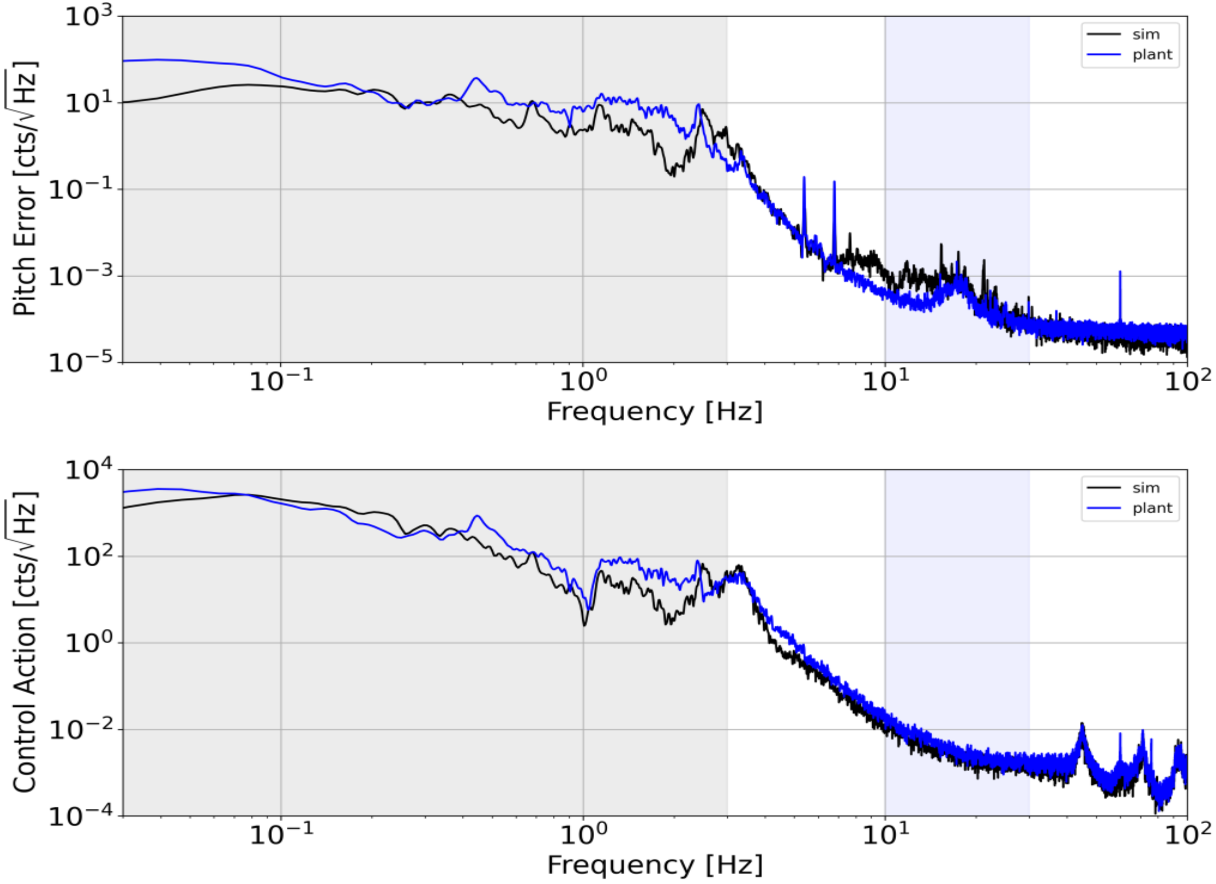




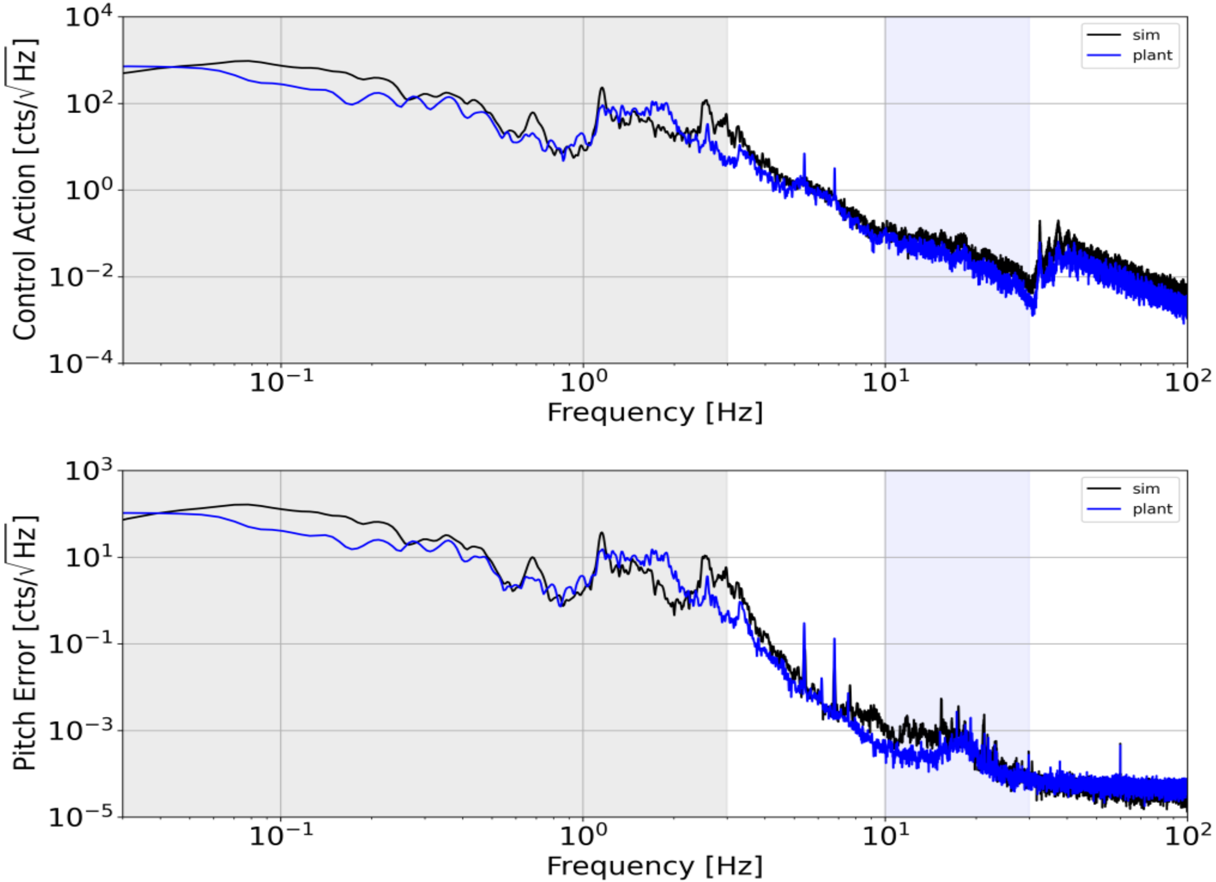
LIGO RayleighMonitor Algorithm

- Makes a set of short-time power spectra.
- Calculates the mean μ and the standard deviation σ of the power spectrum in each frequency bin.
- Ratio $R := \sigma/\mu$ is an interesting statistic:
  - » $R = 1$ is what you expect for Gaussian noise.
  - » $R < 1$ indicates coherent variation.
  - » $R > 1$ indicates glitchy/ratty data.
- RayleighMonitor plots scrolling spectrograms (μ) and “Rayleighgrams” (R) for visual inspection of data characteristics.

squidward - sim2real transfer is excellent above >0.1Hz



spongebob - sim2real transfer is excellent



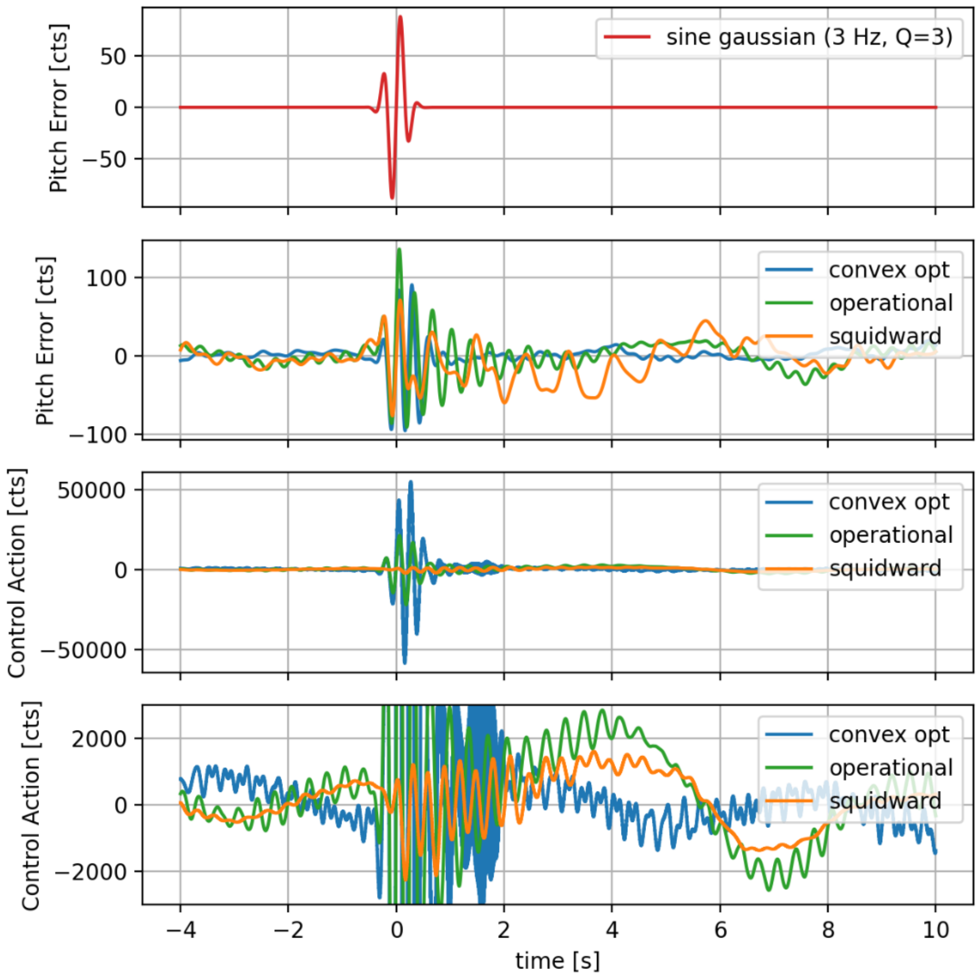
Robustness of controller to transients

Cvx opt controller reacts very aggressively, for not much difference in error

RL controller has longer settling time, low frequency oscillation after the event.

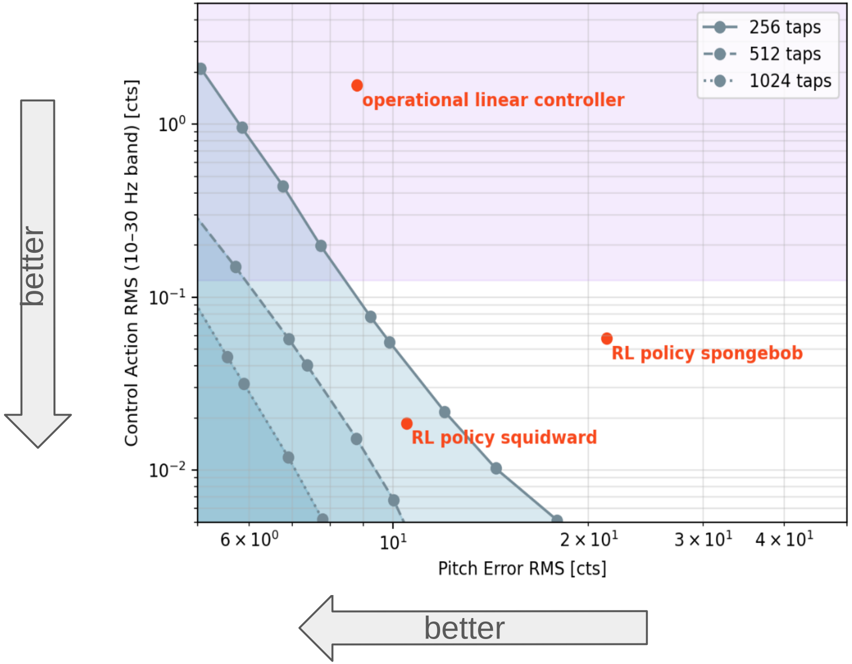
But that’s fine, as long as the mirror remains stable within bounds, what we mostly care about is the noise injected in the GW observation band (10-30Hz)

Note: RL controller is limited *by design* using a soft nonlinearity



Optimized Linear Controllers

- Since the 1980s, convex optimization was recognized as a powerful tool for optimizing linear control loops
- This method lets us map out a "Pareto frontier" of high-performance linear CHARD_P controllers. These can provide a baseline of comparison for the nonlinear policies
- Optimization performed over: FIR filters with varying tap length (i.e., history window size)
- RL policies operate without an auxiliary stabilizer, yet outperform convex-optimized controllers that require one
- Not yet fit for deployment (not robust under plant variation)



Summary

1. The LIGO Controls/ASC feedback noise is ~significant noise source at 10-30 Hz in DARM.
2. This is due to 2 effects:
   a. too much feedback noise at 10-20 Hz
   b. too much beam spot motion at 0.1-1 Hz
3. **Improving the low freq noise will improve several science targets:** IMBH, BNS early warning, BBH eccentricity, high Z sirens
4. We have tried filtering / loopology for years, with some success▼, but are still 10x above the fundamental limits: quantum/gravity
5. This technique (RL/ML/AI) can and should be implemented for MICH/SRCL/ISI/SUS (similar issue - want LF control and less HF noise)

1. Have been working with Deepmind (now Google Deepmind).
2. Collaboration with Caltech (Rana Adhikari)
3. **Great** simulated performance on CHARD using time domain simulation
4. Good real performance on 40m IMC ASC (6x6 MIMO system w/ WFS + QPDs)
5. Success in LLO CHARD_P tests
6. We will have a 'open house' zoom workshop so that people can get some hands on time with the tools.

▼ see LLO elog 64660 on RHP zeroes by Valera and Chris Wipf

Questions / Worries

1. Is it safe? Does it inject fake black holes?
2. Can it go crazy and damage the optics?
3. Long term stability?
4. Robustness to transients?
5. How long does it take to train?

1. We record the RL control signals in frames as usual for any controller.
2. Poorly trained controllers can be unstable and make oscillations, but we have limits on the controller's output as usual.
3. We have run it for some hours at the 40m and the performance is stationary (as per our rough eyeball estimates of Rayleigh-grams)
4. Is robust under these tests:
    a. turned off and on the sensors
    b. turned off and on one mirror actuator
    c. big step in actuator (reduced trans power by 2x)
    d. banging on chamber
    e. walking around chamber
    f. turned on linear controller in parallel (!!)
5. Now that the exploration space has been reduced, the training takes ~1-2 days on a good machine with a few GPUs. Can be done in AWS or Google Cloud. Has ~700 free parameters.

Future Work

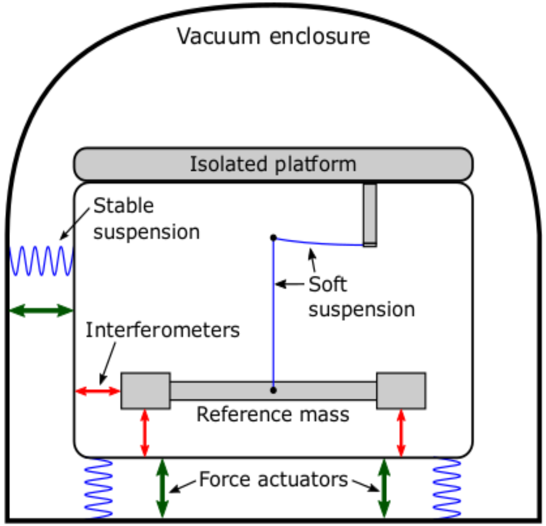
1. Implement on Virgo loops
2. Make the plant sysID more automated and robust
3. Adapt the Lightsaber model continuously to the live data to extract physically meaningful plant parameters.
4. Explore hybrid linear + nonlinear control
5. Make it run robustly during high-noise conditions
6. Reduce non-stationarity in controller output (c.f. Rayleigh grams)

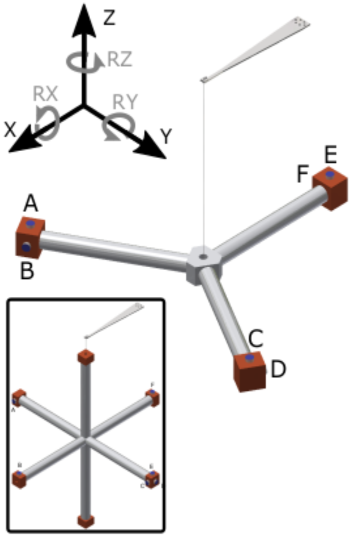
Gemini's idea of an improved gravitational wave detector

ASC for ET-LF

- ET – issue needs to be addressed already with its design
- RMS not filtered with mechanical suspension
- Sensors will not become better
- Natural resonant frequency ~0.05 Hz in pitch and ~0.2 Hz in yaw
- Resonant frequencies for ET-LF are:
  - soft mode pitch: 0.0218 Hz
  - hard mode pitch: 0.1413 Hz
  - soft mode yaw: 0.1949 Hz
  - hard mode yaw: 0.2397 Hz
- **Need for RL for improvements**
- **ET-LF-Lightsaber**

- Omnisens





C M Mow-Lowry and D Martynov 2019 Class. Quantum Grav. 36 245006

GS
SI

Thank you!

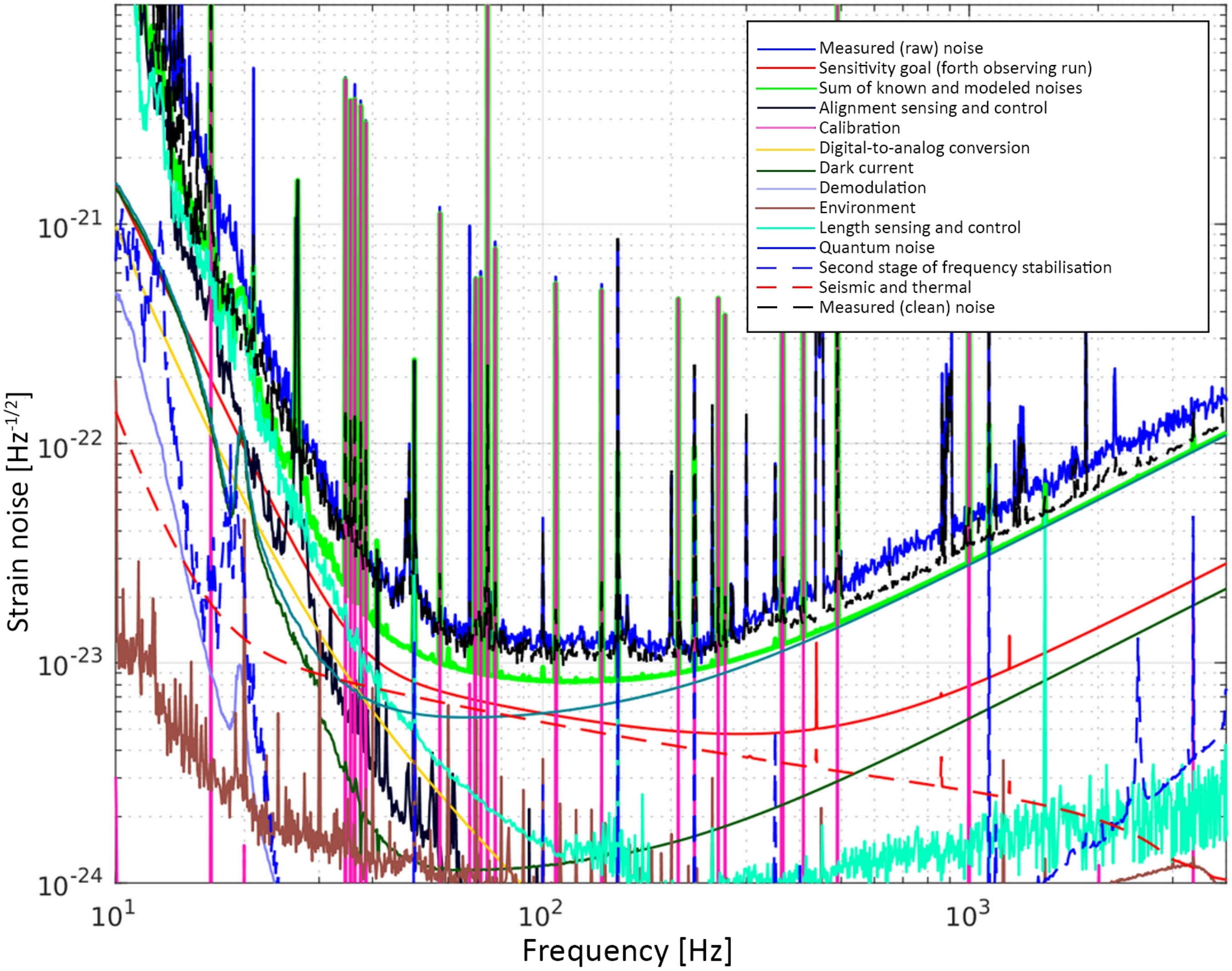
GSSI Google DeepMind Caltech LIGO

tomislav.andric@gssi.it

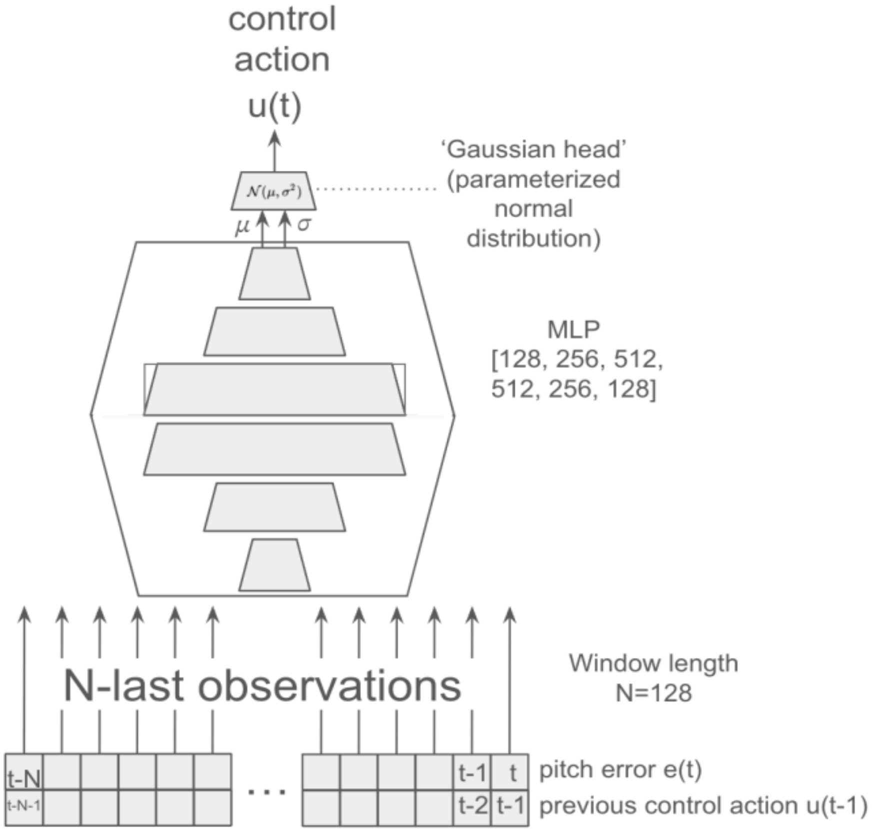
Backup

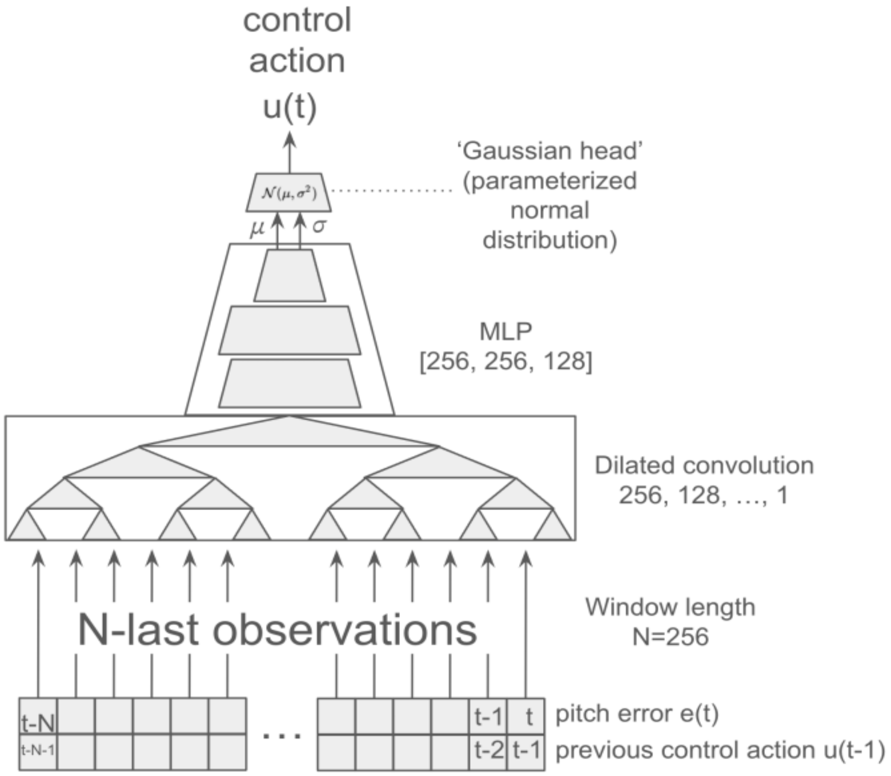
GSSI

INFN Istituto Nazionale di Fisica Nucleare

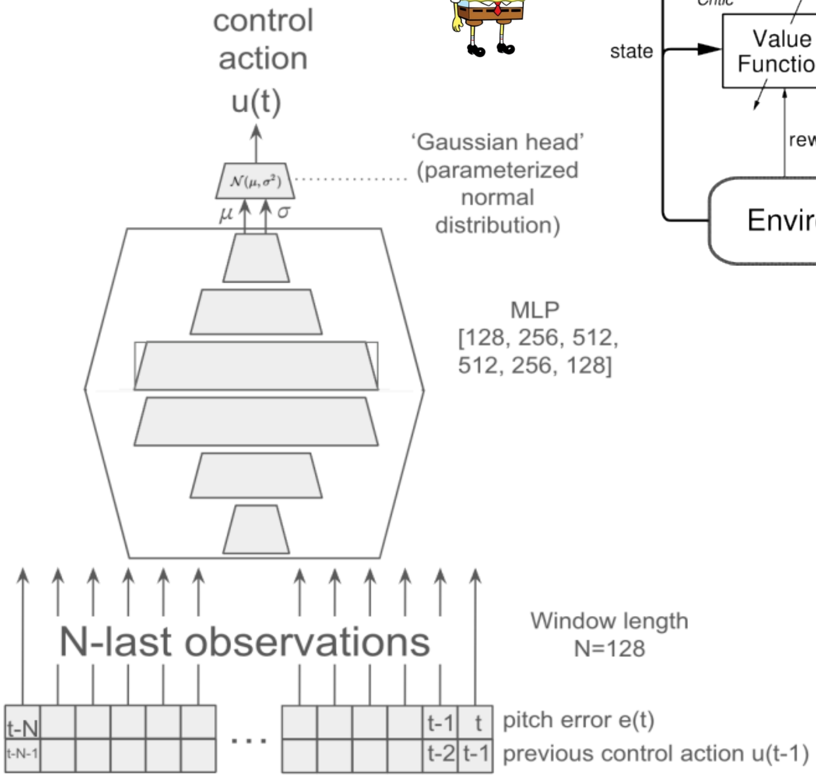


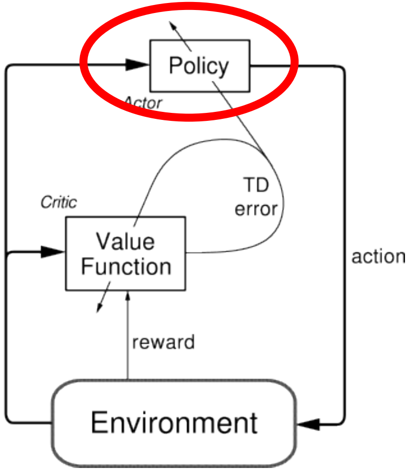
Architectures

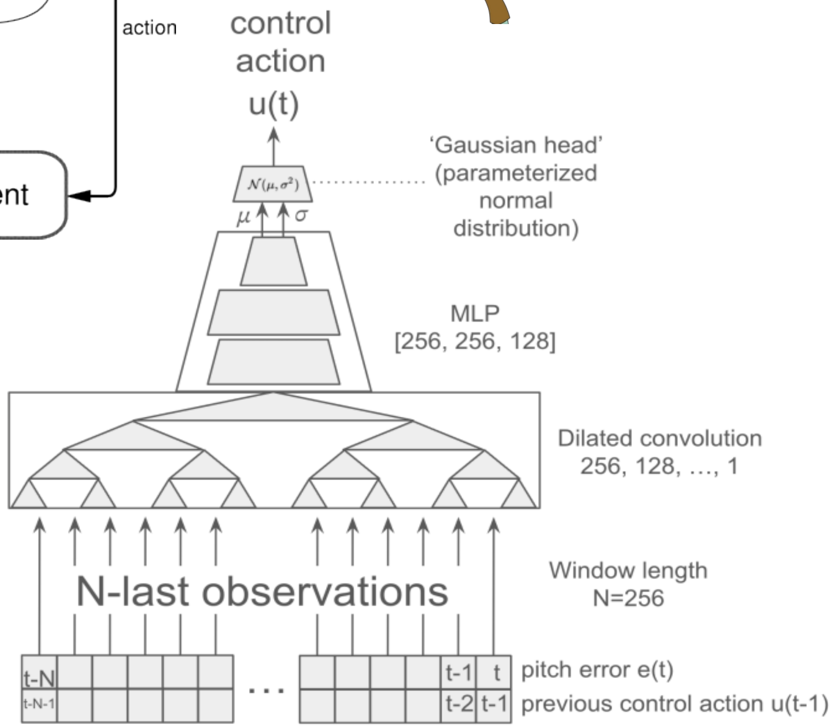




Policy Architectures



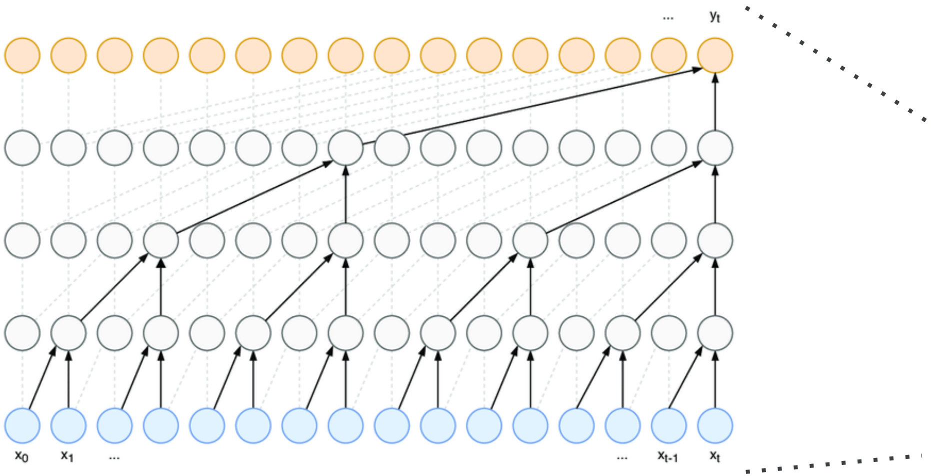


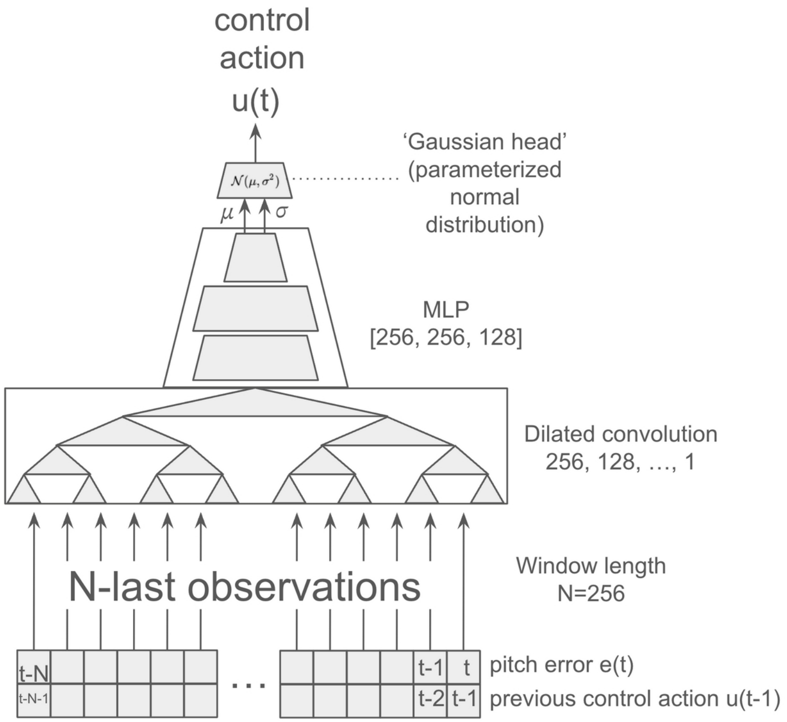


Policy network architecture

Network needs long history for signal processing

Dilated convolutions are an efficient way to enable long histories
...without overwhelming parameter count



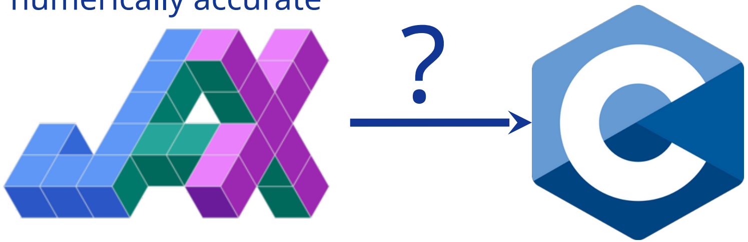


How not to crash a multi-100M-$ experiment

Simple ask: Could you please run your neural networks in the kernel space of our hard-real-time control system?

Requirements: hard-real time, kernel space

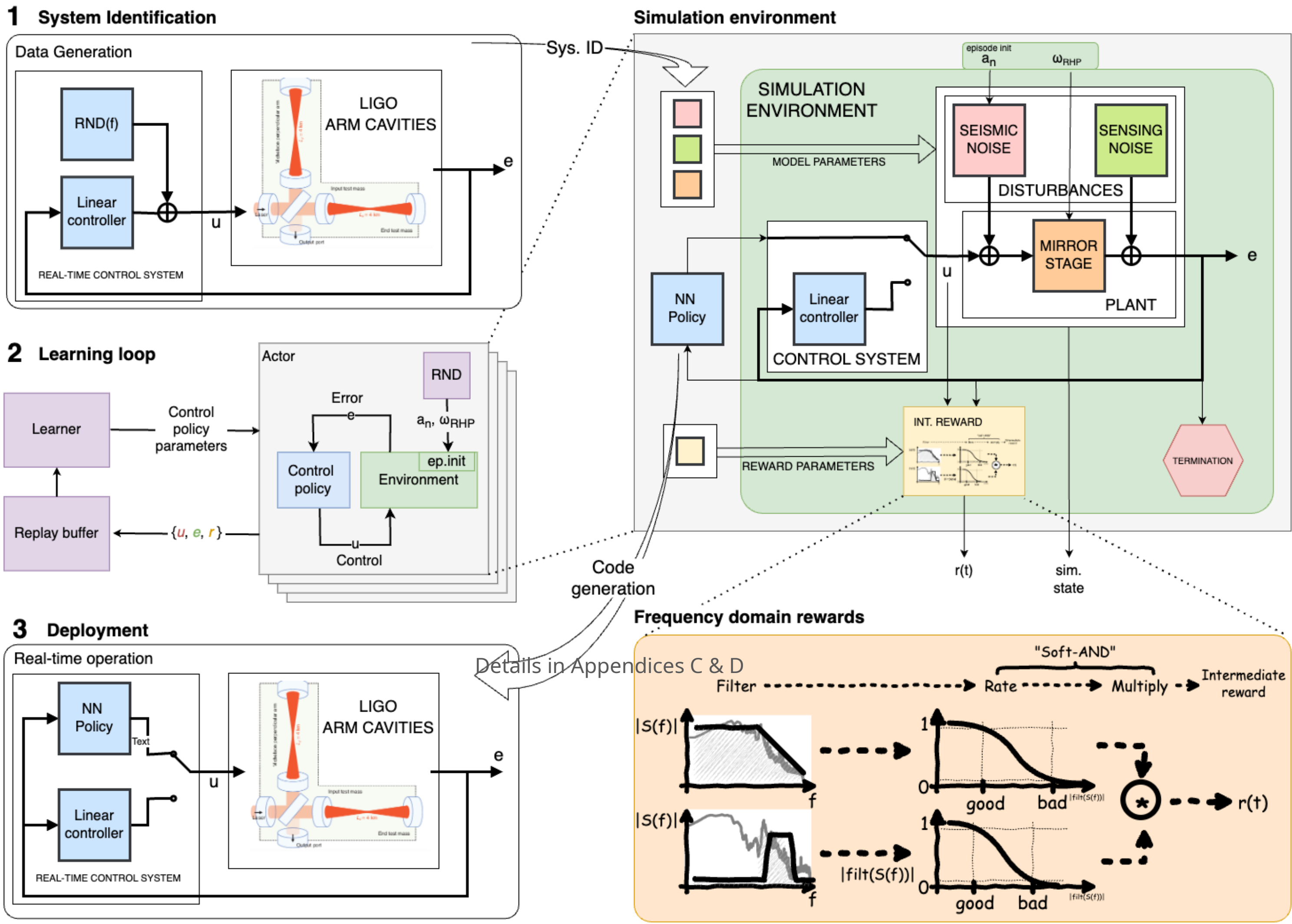
- Compile with LIGO toolchain / workflow as a **Linux kernel module**
- Zero dependencies, real time guarantees
- Run deterministically at 2kHz, small binary size
- No memory allocation, syscalls, libm (!), etc.
- Be absolutely stable and numerically accurate



Solution: custom code generation JAX/TF to C. Enabled running networks in hard real-time Kernel space … with zero crashes

From concept to closed loop operation on LIGO in ~6 weeks

Illustration of method



CDS Issues

- How does this work?
  - Runs in kernel as a C user function, or in userspace as a loadable library
  - In either case, it is hardcoded (changes require a recompile and restart)
- What’s the difference between kernel and userspace?
  - Kernel mode runs internal to the operating system. Userspace mode runs as an application program
  - Normally controllers have run in the kernel. Userspace controllers are still an experimental development
- How straightforward is it to propagate this elsewhere?
  - Userspace build uses standard tools. C code generator for kernel build remains proprietary for now
  - Build process requires some manual intervention (but less so in the latest RTS release)
- How much control bandwidth is possible?
  - Typical run time on present hardware is short enough for sample rates of 2048 Hz, BW ~10s of Hz
- How could we scale to run 10 or 100 of these in the future?
  - They are more hardware intensive than linear filtering, so more CPU cores would have to be provided
  - Expect that we could run the 4 HARD loops on a single core
- How do we train these networks in the future?
  - Involve more people
  - Need medium scale GPU resources
- What’s the turnaround time to get it running for a new loop?
  - Limited by SysID time: need faster, automated sysID

Future work - Gravitational Wave Observatories

- Address all 3 additional 'HARD' mode control loops and demonstrate sensitivity improvement of the instrument
- Improve low frequency control authority

Mid term

- Establish properties of the nonlinear control policy, ensure no interference with measurement
- Prove concept in daily operations

Longer term

- Remove dependence on system identification: Online/offline learning with LLO data?
- Solve the 'right' problem: Use 'all channels', do feedforward+feedback control (catch: big step up in engineering complexity)

Understand and apply Deep Loop Shaping as a generic RL/control method!

Gemini's idea of an improved gravitational wave detector.

Future work for control affine people

Understand and apply
Deep Loop Shaping as a
generic RL/control method!

- Variants of frequency estimates (FFTs, wavelets, filters…)
- Post episode vs online scoring
- Causality?
- Improve low frequency control authority
- What’s the role and importance of nonlinearities
- Use some of the ‘verifiable’ NN within this framework?
- Difficult control problems to try DLS on?
- ….

Plan / Impact

- A **first demonstration in Virgo** will take a few months of work involving a few key experts (real-time system, control, Project Shimmer)
- Afterwards, turnaround time to **apply to a new control loop** will be mostly set by system identification: need faster, automated sys ID directly interfacing with Lightsaber simulations
- **It might take years until Deep Loop Shaping is used for all important control loops**, but it provides a clear path to mitigate control noises enough for all future Virgo upgrades.
- Develop **adaptive Deep Loop Shaping to maximize the benefit** [requires GPU real-time infrastructure]
- **Deep Loop Shaping enables ET-LF**

Suggestions for Virgo

1. Auxiliary longitudinal DOFs, including the Power Recycling cavity, Signal Recycling cavity, and Michelson DOF, to minimize control noise injection by optimizing loop shapes and reducing required control bandwidth.

2. Angular DOFs, where RL could mitigate angular-to-longitudinal coupling by learning optimal actuation strategies tailored to real-time beam alignment, with the potential to outperform static dithering or feed-forward techniques.

3. Can also be used for Wavefront Sensing and Control as demonstrated in GEO600, but this is a qualitatively different controls problem also from RL perspective.

IMC

Need-to-Know