Non-linear and Non-stationary Noise Removal from GW Data

Gabriele Vajente (vajente@caltech.edu)
LIGO Laboratory – Caltech

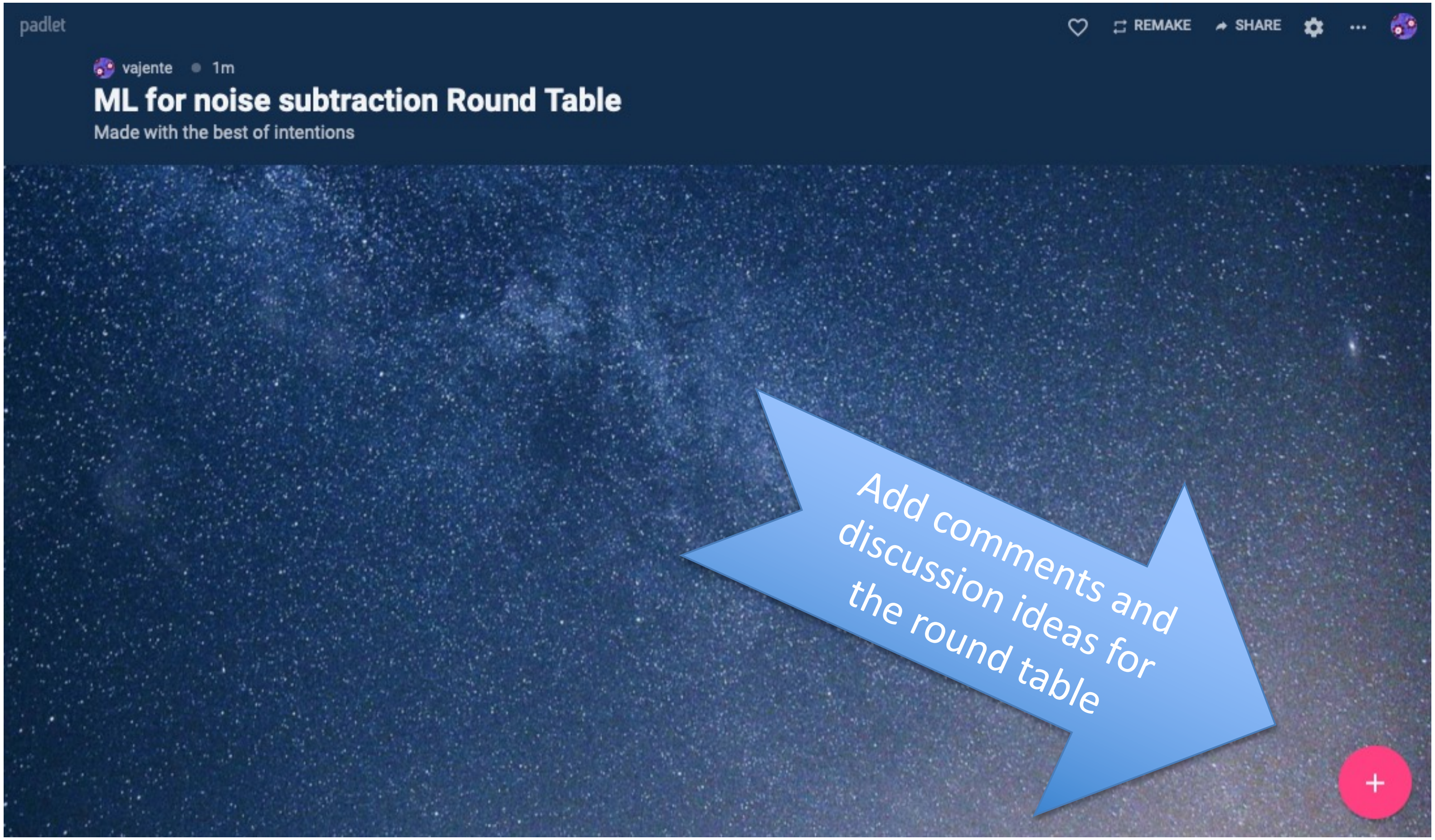
g2net WG3 workshop on Machine Learning for Advanced Control Techniques

March 22nd 2021

COST ACTION CA17137
A NETWORK FOR GRAVITATIONAL WAVES, GEOPHYSICS AND MACHINE LEARNING

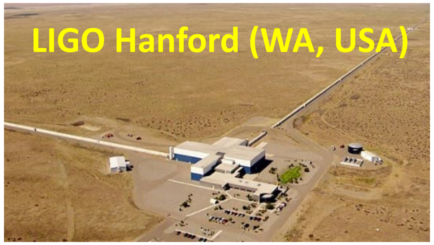
LIGO-G2100643-v1

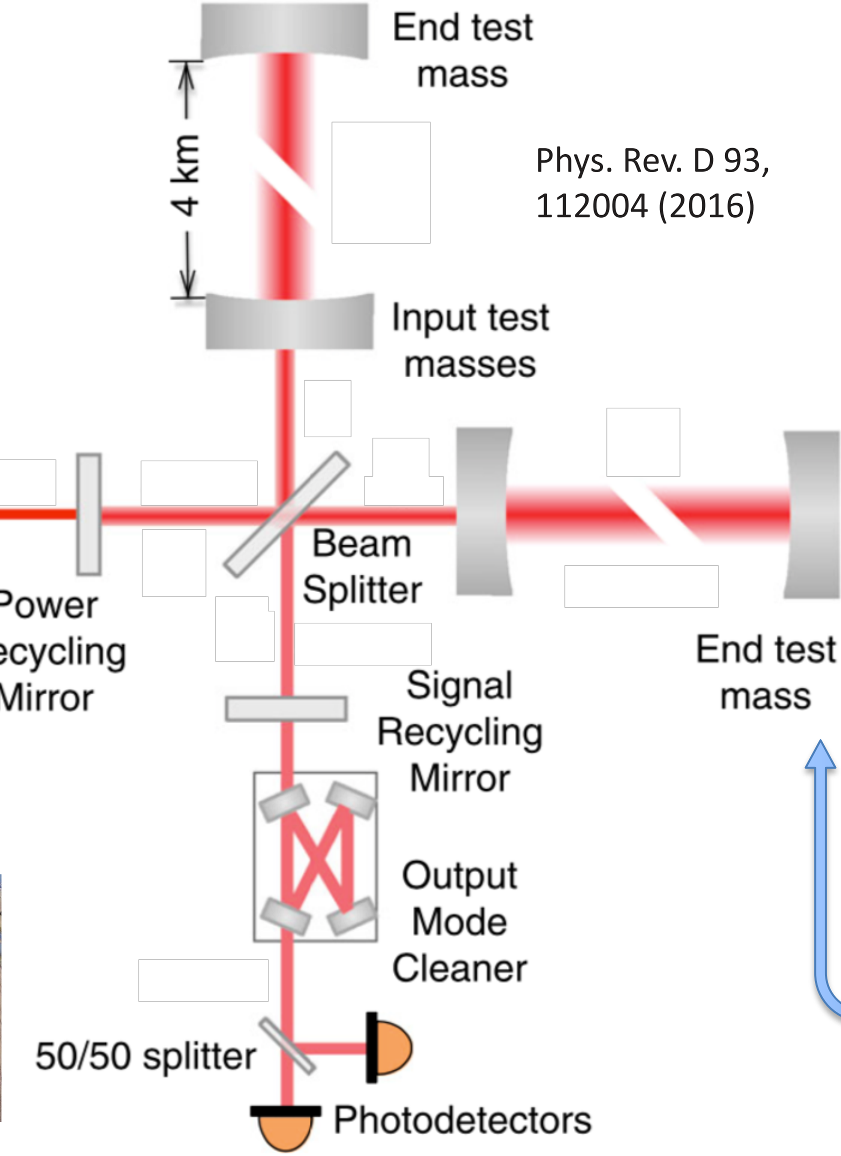
Round table: ML for noise subtraction



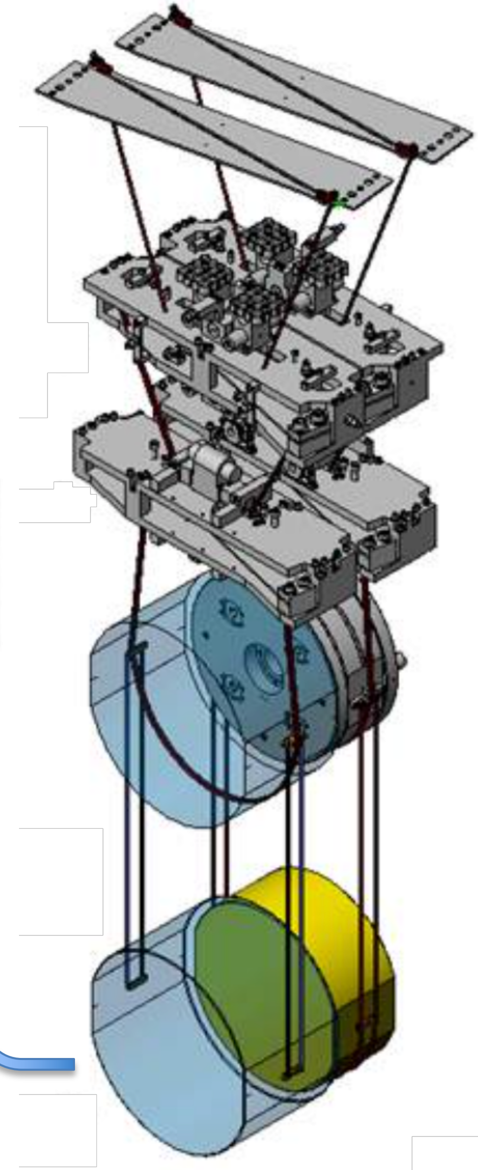
https://padlet.com/vajente/whthqq0lxa7nfy52

Advanced GW Detectors

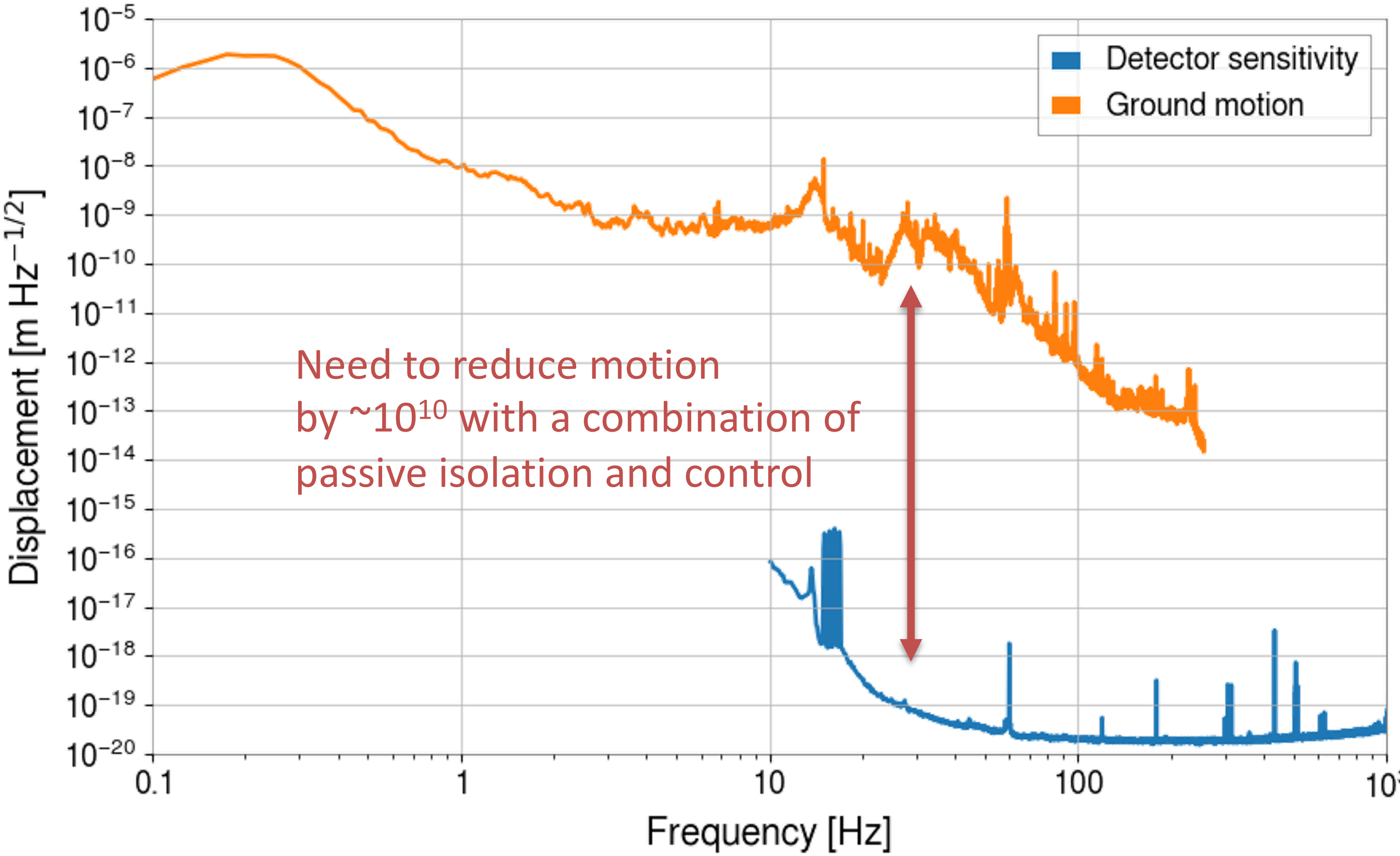




Phys. Rev. D 93, 112004 (2016)



High Sensitivity Operation

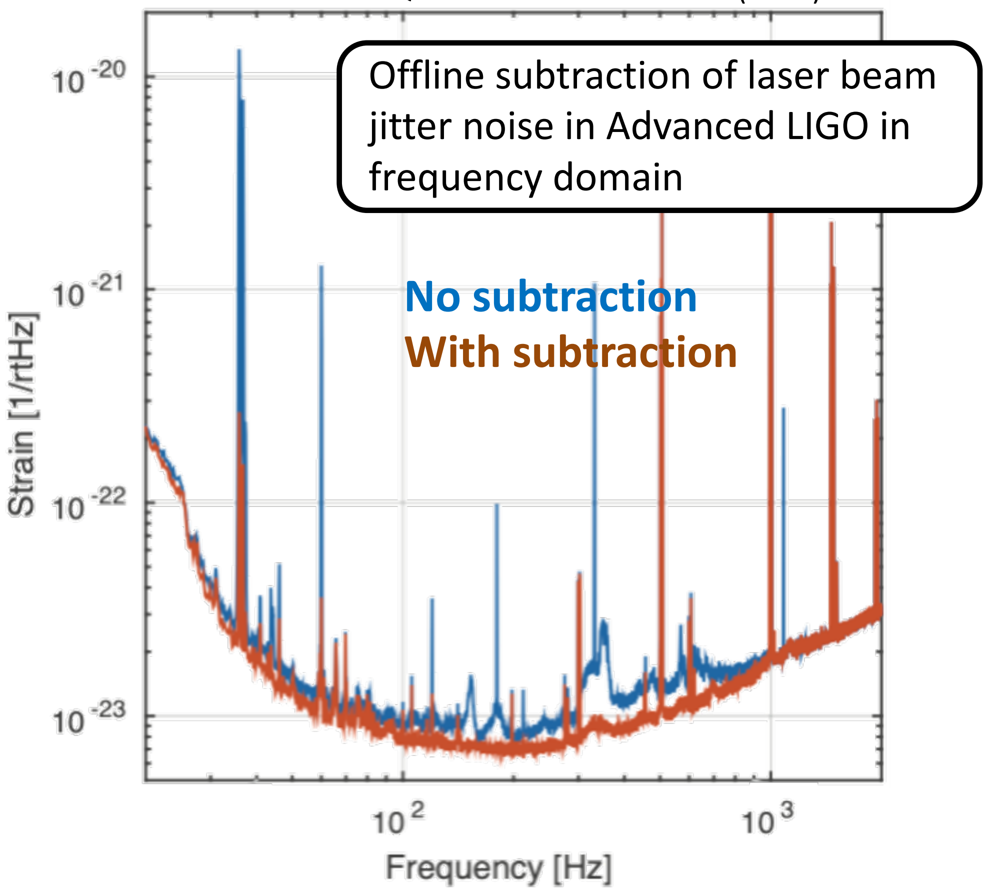


Compensate for **1-10 μm of motion below ~1 Hz**, without introducing spurious displacement larger than ~10^{-19} m/√Hz above ~10 Hz

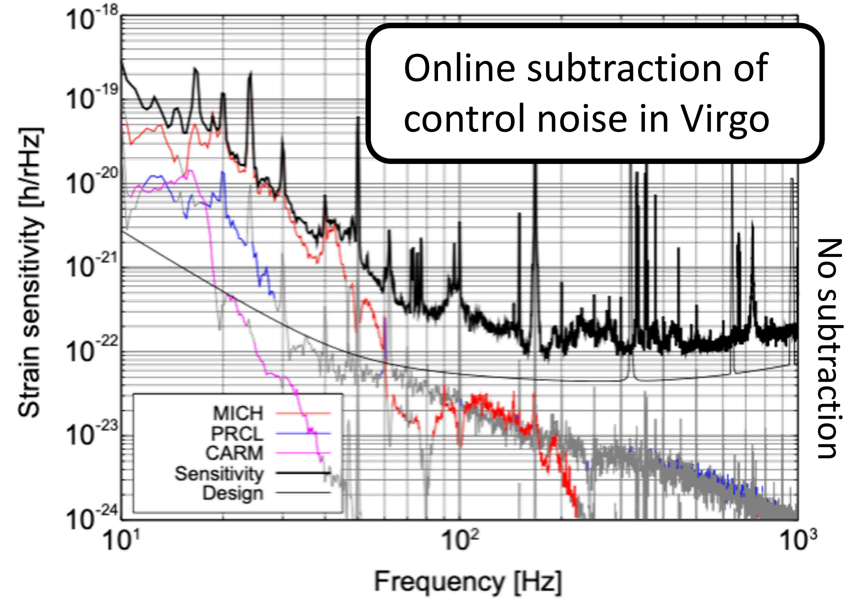
Feedback controls to **linearize** the system response and maintain the operating point close to optimal

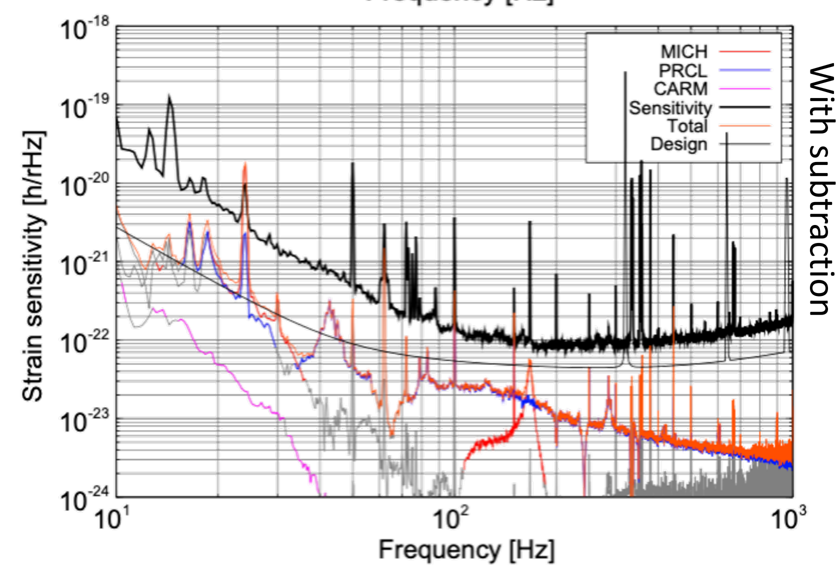
Linear Noise Subtraction

Phys. Rev. D 99, 042001 (2019)
Class. Quantum Grav. 36 055011 (2019)



Astroparticle Physics 33 (2010) 75





If the coupling is **linear** and **stationary**:

- Coherence / Transfer Function analysis, Wiener filter, online or offline

Linear vs Non-Linear

Linear

There's only one functional form

Simple parametrization

$$h(t) = \int_0^{+\infty} \alpha(\tau) x(t-\tau)\, d\tau$$

$$\tilde{h}(\omega) = \tilde{\alpha}(\omega)\tilde{x}(\omega)$$

$$\tilde{\alpha}(\omega) = k \frac{\prod_i (i\omega - z_i)}{\prod_i (i\omega - p_i)}$$

Coherence analysis is a powerful discovery tool

LIGO-G1500230

Non-Linear

Only constraint from causality

$$h(t) = \mathcal{F}[x(\tau < t)]$$

Volterra Series:

$$h(t) = \sum_{n=0}^{+\infty} \int_0^{+\infty} \dots \int_0^{+\infty} \alpha_n(\tau_1, \dots \tau_n) \prod_{i=1}^{n} x(t-\tau_i)\, d\tau_i$$

No easy parametrization

No discovery tool (higher order coherences are ill-defined and numerically unstable)

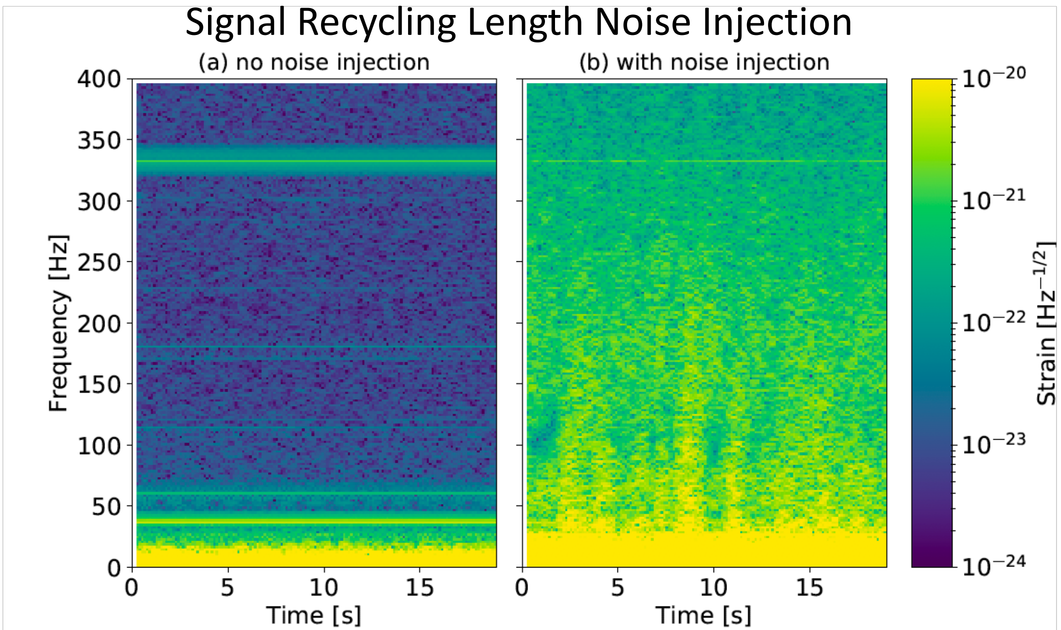
Non-Linear and Non-Stationary

- Noise in auxiliary channels can couple in a **non-linear** way

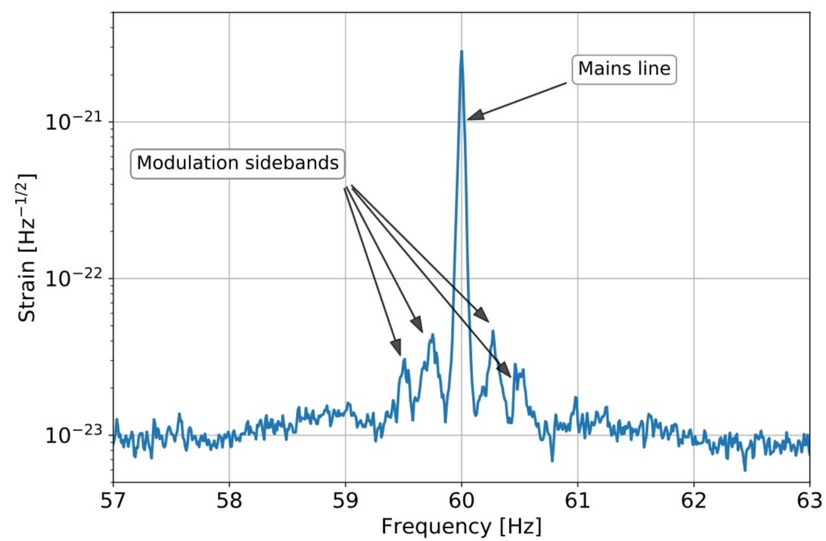
$$h(t) = \varepsilon_{\mathrm{B}}(t) + \mathcal{F}\left[w_1(\tau < t), \ldots, w_N(\tau < t)\right]$$

Some other background noise

- However, (in most cases*) we expect small deviations from linearity: **quadratic coupling** or **non-stationary coupling**



Phys. Rev. D 101, 042003 (2020)



*scattered light is an important exception: arXiv:2007.14876 (2020), Opt. Expr. 2 10546 (2013)

How does non-linear subtraction work?

- One source of noise with power in the frequency band of interest (e.g. 10-200 Hz)
- Coupling is instantaneously linear but modulated at low frequencies (e.g. <1 Hz)

G. Vajente, et al. Phys. Rev. D **101**, 042003 (2020)

Stationary linear coupling

Trained parameters

Fast noise witness

$$h(t) = H[s(t)] + \sum_{i=1}^{N} \alpha_i [x_i(t) s(t)]$$

Target: strain

Trained parameters

Non-stationary couplings

Modulation witness signals

- Parametrize the couplings and use ML-inspired cost optimization algorithms to minimize the noise in the strain signal

Parametrized approach

G. Vajente, et al. Phys. Rev. D **101**, 042003 (2020)

- Describe each transfer function using a parametrized form

- Laplace domain: $$\alpha_h(s) = \frac{\sum_{i=0}^{N_N} b_i s^i}{\sum_{j=0}^{N_D} a_j s^j}$$ Or a more efficient parametrization!

$$\tilde{r}(\omega) = \tilde{h}(\omega) - \tilde{y}(\omega) = \tilde{h}(\omega) - \sum_{i=0}^{N} \tilde{\alpha}_i(\omega)\tilde{s}_i(\omega)$$

- Define a weighted frequency-integrated cost function

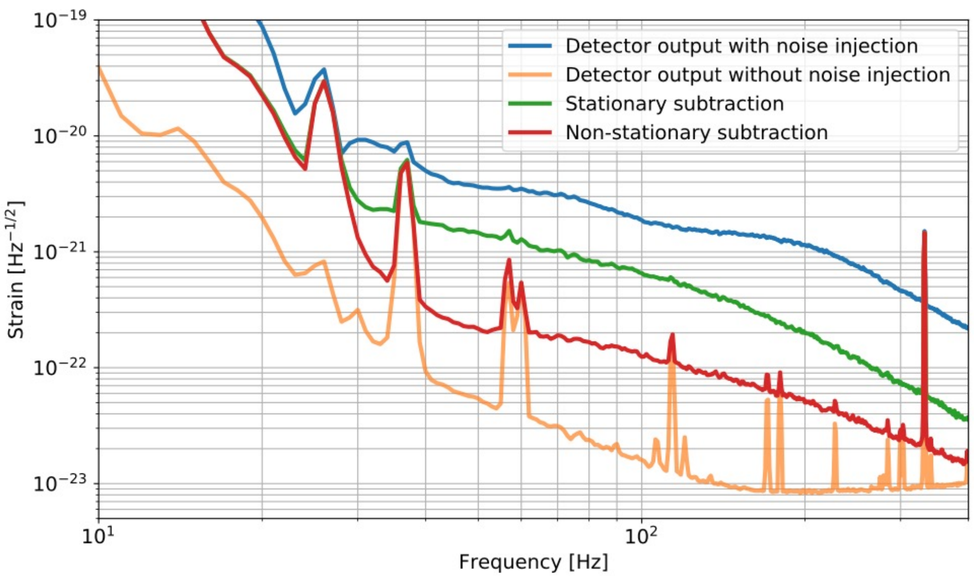
$$\mathcal{C}(\theta_{h,i}) = \frac{1}{\omega_2 - \omega_1} \int_{\omega_1}^{\omega_2} W(\omega) S_{rr}(\omega) d\omega \qquad W(\omega) = \frac{1}{S_{hh}(\omega)}$$

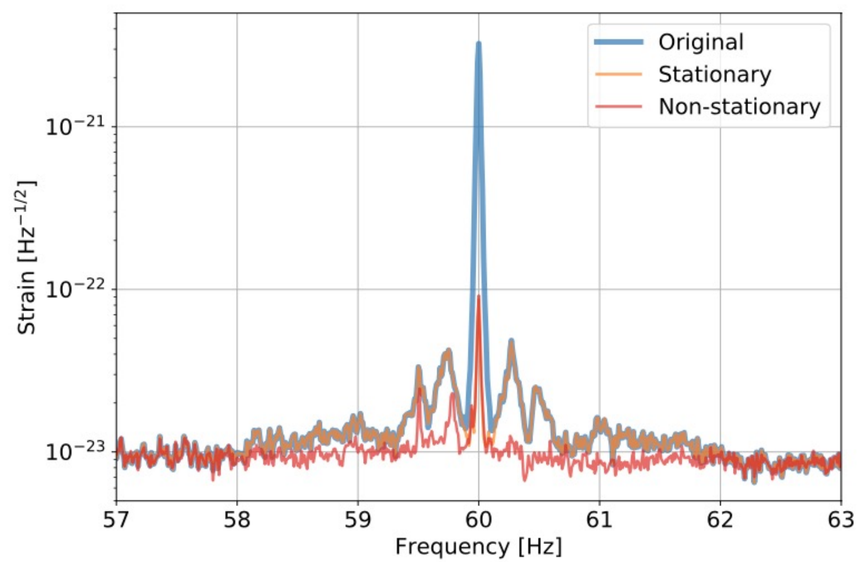
- Use gradient-based optimization algorithms to minimize residual (ADAM)

$$\frac{\partial \mathcal{C}}{\partial \theta_{h,n}} = \frac{1}{\omega_2 - \omega_1} \int_{\omega_1}^{\omega_2} W(\omega) \left[\frac{\partial S_{rr}}{\partial \alpha_h} \cdot \frac{\partial \alpha_h}{\partial \theta_{h,n}} + \frac{\partial S_{rr}}{\partial {\alpha_h}^*} \cdot \frac{\partial {\alpha_h}^*}{\partial \theta_{h,n}} \right] d\omega$$

Physics-informed Machine Learning Models

G. Vajente, et al. Phys. Rev. D **101**, 042003 (2020)

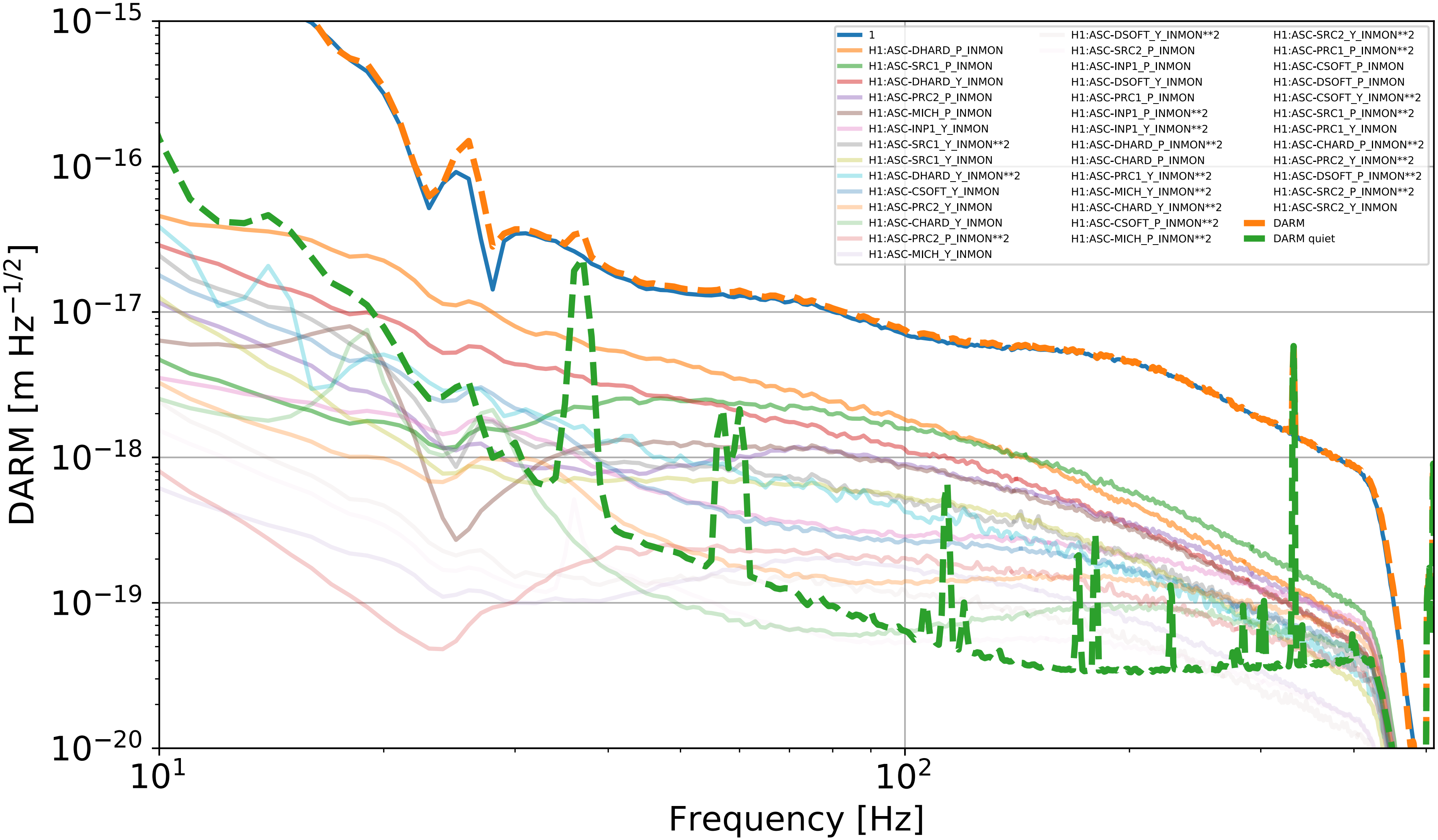




- Non-stationary noise modelling and subtraction **outperforms** the linear subtraction
- Not the most general non-linear model, but "**good enough**" to improve Advanced LIGO sensitivity

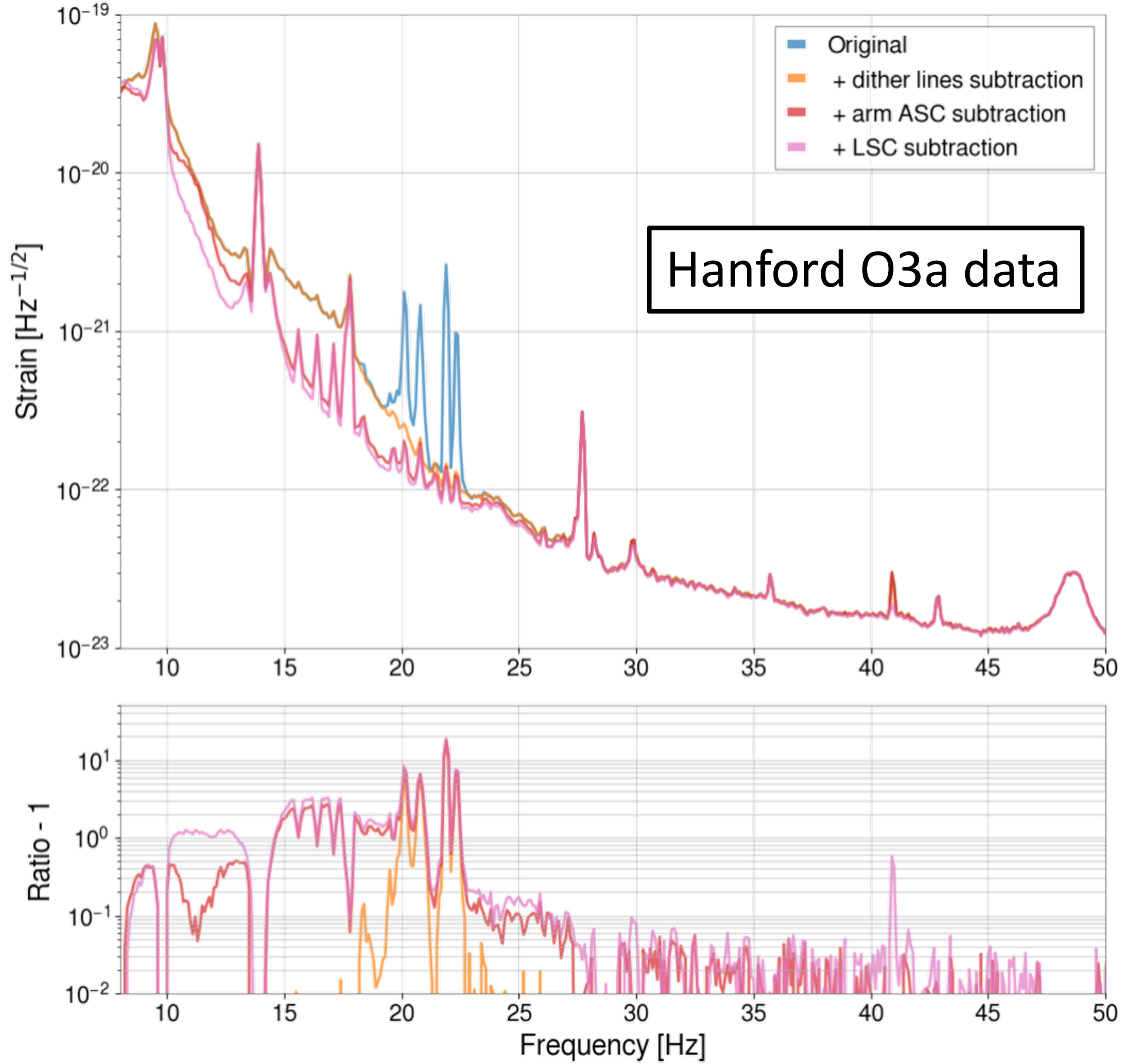
- **Interpretable results**: we can learn what degree of freedom modulates the coupling
- **Time domain implementation**

Coupling modulations are explicit



Low frequency control noise at LHO

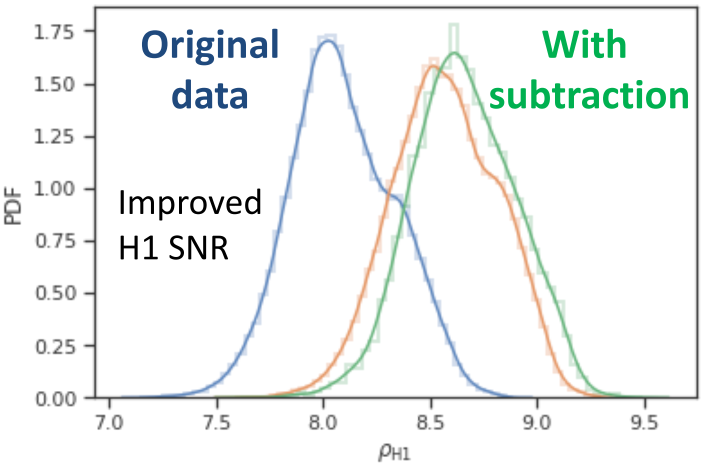
https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=57503
https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=56960
https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=57423

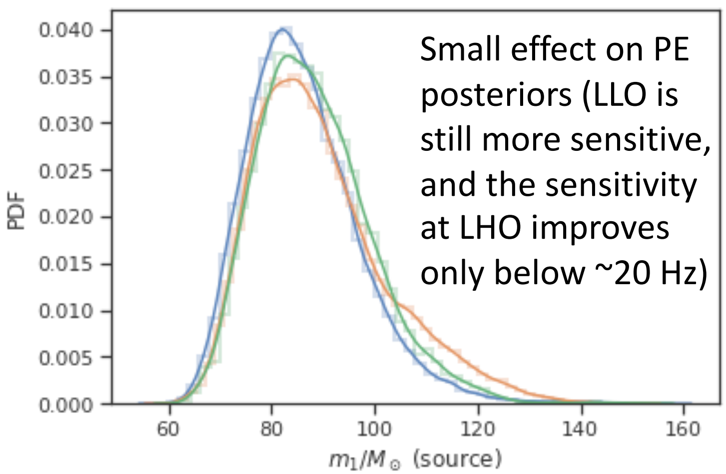


Potential astrophysical impact

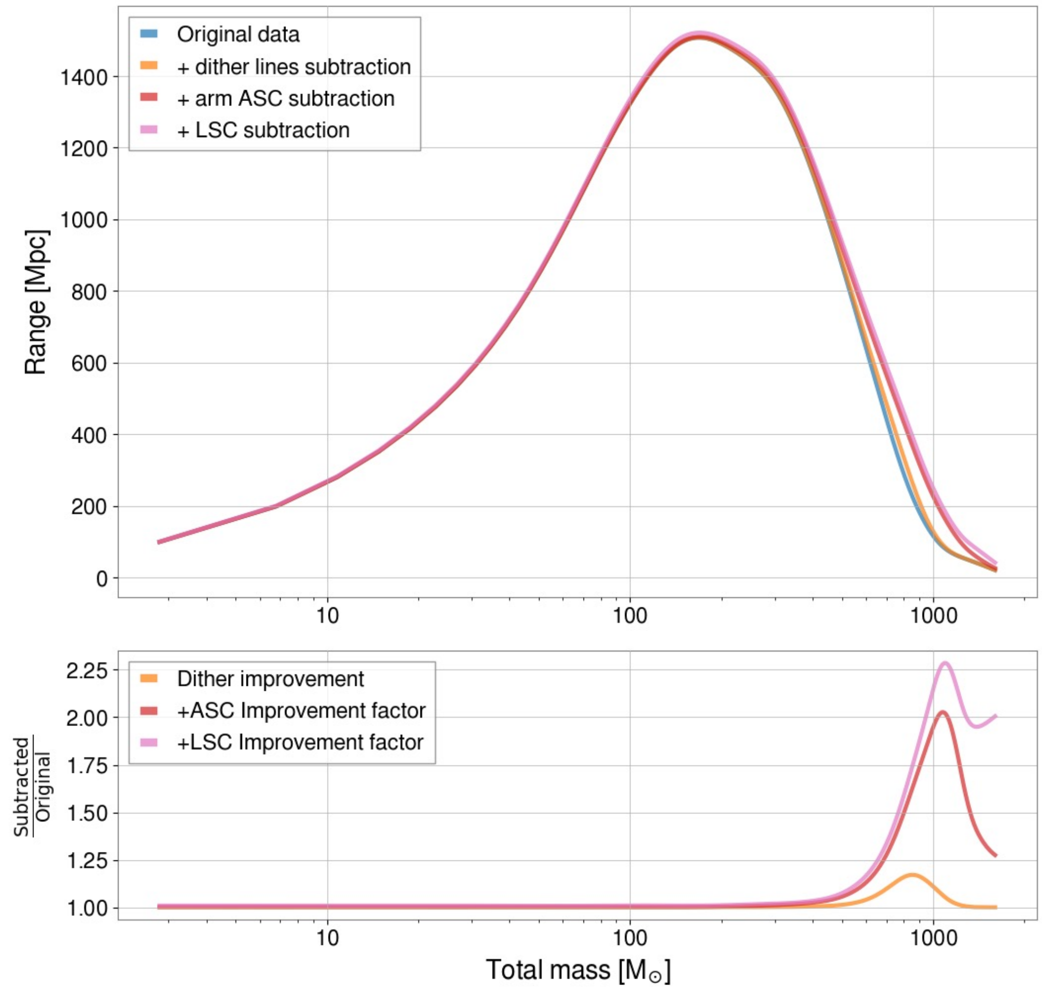
Parameter Estimation by M. Isi

High mass event GW190521



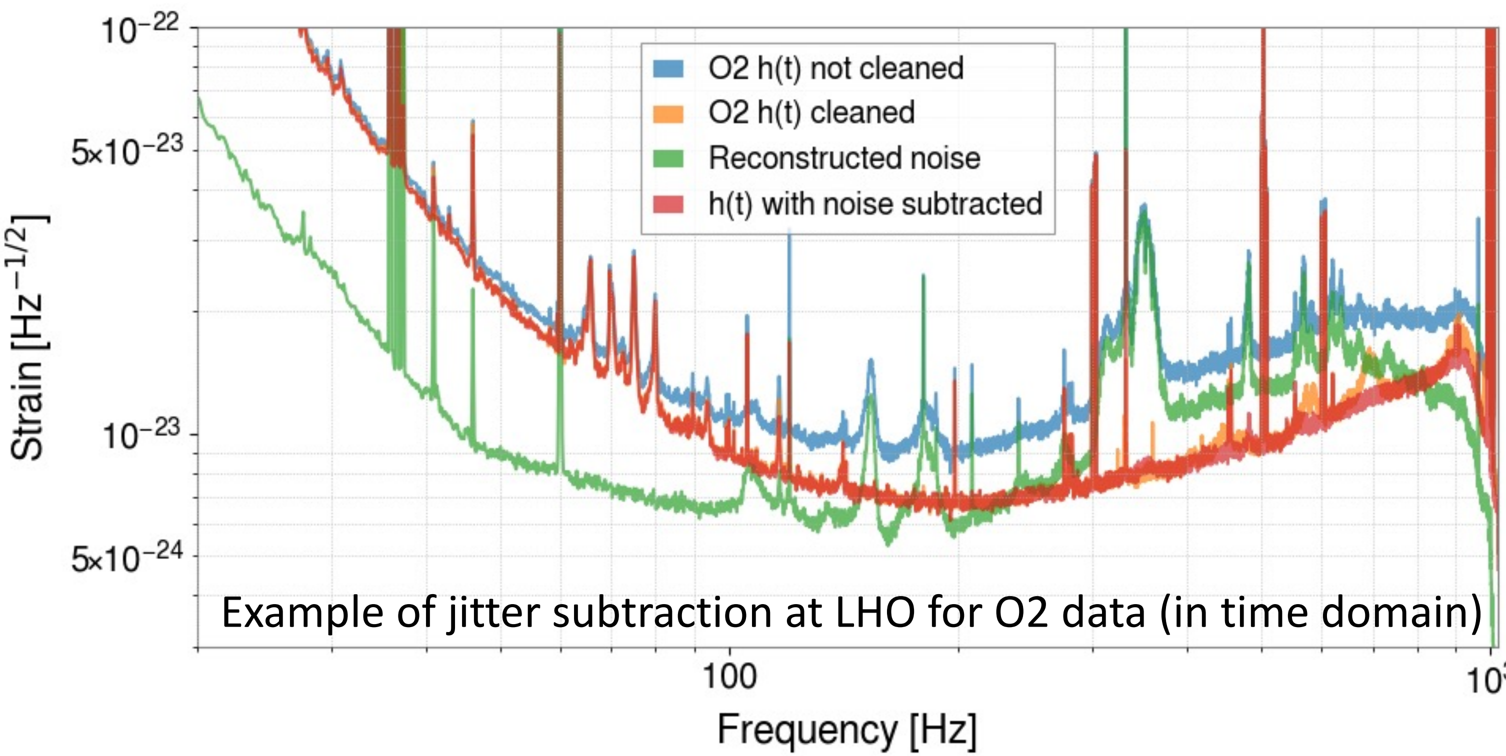


Effect on range for high masses



Going into O4 and beyond

- This is now a developed tool that can be used for **offline** noise subtraction (stationary and non-stationary)
- It could be adapted to run **online** by implementing IIR filters in the real time frontends or in low latency pipelines
- It could be made **adaptive** (with some more work and tests)



Time domain IIR implementation

20x Modulation witnesses

ASC-INP1_P_INMON
ASC-INP1_Y_INMON
ASC-MICH_P_INMON
ASC-MICH_Y_INMON
ASC-PRC1_P_INMON
ASC-PRC1_Y_INMON
ASC-PRC2_P_INMON
ASC-PRC2_Y_INMON
ASC-SRC1_P_INMON
ASC-SRC1_Y_INMON
ASC-SRC2_P_INMON
ASC-SRC2_Y_INMON
ASC-DHARD_P_INMON
ASC-DHARD_Y_INMON
ASC-CHARD_P_INMON
ASC-CHARD_Y_INMON
ASC-DSOFT_P_INMON
ASC-DSOFT_Y_INMON
ASC-CSOFT_P_INMON
ASC-CSOFT_Y_INMON

x

ASC-CHARD_P_OUT_DQ
ASC-CHARD_Y_OUT_DQ
ASC-DHARD_P_OUT_DQ
ASC-DHARD_Y_OUT_DQ
ASC-CSOFT_P_OUT_DQ
ASC-DSOFT_P_OUT_DQ

6x Noise witnesses

→

6x Stationary signals + 120x Modulated signals

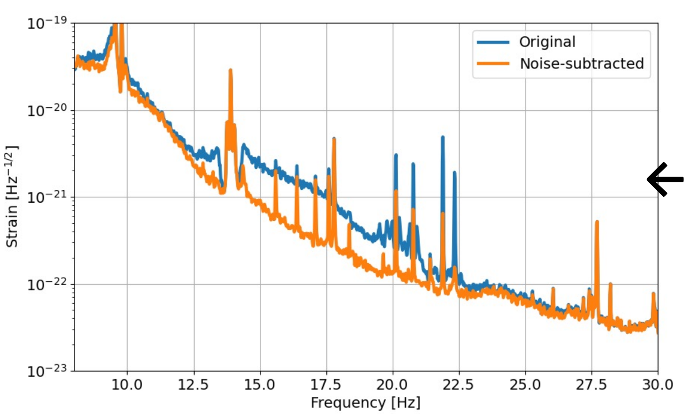
ASC-CHARD_P_OUT_DQ
ASC-CHARD_Y_OUT_DQ
ASC-DHARD_P_OUT_DQ
…
ASC-CHARD_P_OUT_DQ * ASC-INP1_P_INMON
ASC-CHARD_P_OUT_DQ * ASC-INP1_Y_INMON
ASC-CHARD_P_OUT_DQ * ASC-MICH_P_INMON
…
ASC-CHARD_Y_OUT_DQ * ASC-INP1_P_INMON
ASC-CHARD_Y_OUT_DQ * ASC-INP1_Y_INMON
ASC-CHARD_Y_OUT_DQ * ASC-MICH_P_INMON
…

↓

α_1 [ASC-CHARD_P_OUT_DQ]
α_2 [ASC-CHARD_Y_OUT_DQ]
α_3 [ASC-DHARD_P_OUT_DQ]
…
α_7 [ASC-CHARD_P_OUT_DQ * ASC-INP1_P_INMON]
α_8 [ASC-CHARD_P_OUT_DQ * ASC-INP1_Y_INMON]
α_9 [ASC-CHARD_P_OUT_DQ * ASC-MICH_P_INMON]
…
α_{27} [ASC-CHARD_Y_OUT_DQ * ASC-INP1_P_INMON]
α_{28} [ASC-CHARD_Y_OUT_DQ * ASC-INP1_Y_INMON]
α_{29} [ASC-CHARD_Y_OUT_DQ * ASC-MICH_P_INMON]
…

126x IIR filters (each is 12^{th} order)

← **Sum and subtract from h(t)** ←



How does non-linear subtraction work?

- Implemented as a python code, with mature user interface [NonSENS: Non-Stationary Estimation of Noise Subtraction]

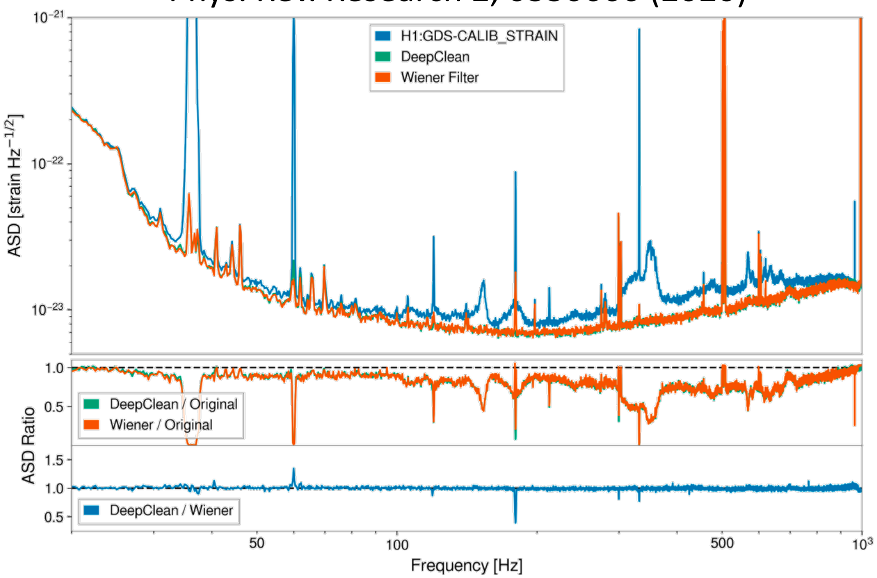
  https://git.ligo.org/gabriele-vajente/nonsens
- See in the `examples` folder for instructions and examples on how to use the package
- The algorithm compute coefficients of filters in Laplace or z-domain: direct implementation of **time-domain subtraction**
  - It can be done online and even adaptively
- The source of coupling modulation is explicitly exposed: allows understanding of the IFO dynamics

Why not a Neural Network?

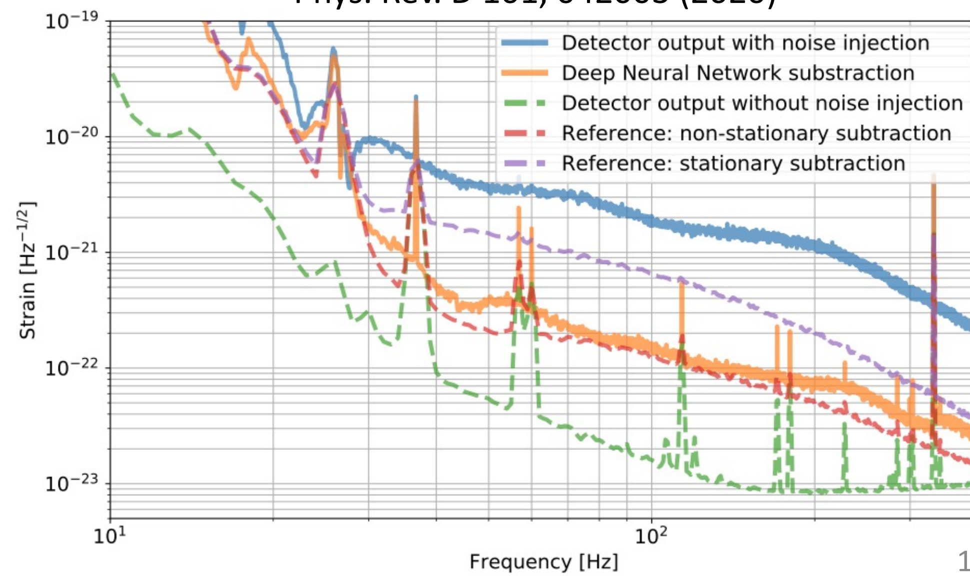
- In theory, a **Deep Neural Network** can express arbitrarily complicated non-linear functions
- Need to include time-dependency and history, therefore Convolutional NN or Recurrent NN
- DNN shown to reproduce linear subtraction performance
- Although a RNN is in principle capable of expressing a slowly modulated coupling, it's hard to train, at the end of training it underperforms, and it is not easily interpretable

Sampling problem: need fast sampling to capture noise and long times to capture variations.
Training problem: CNN and RNN do not scale well with very long time series

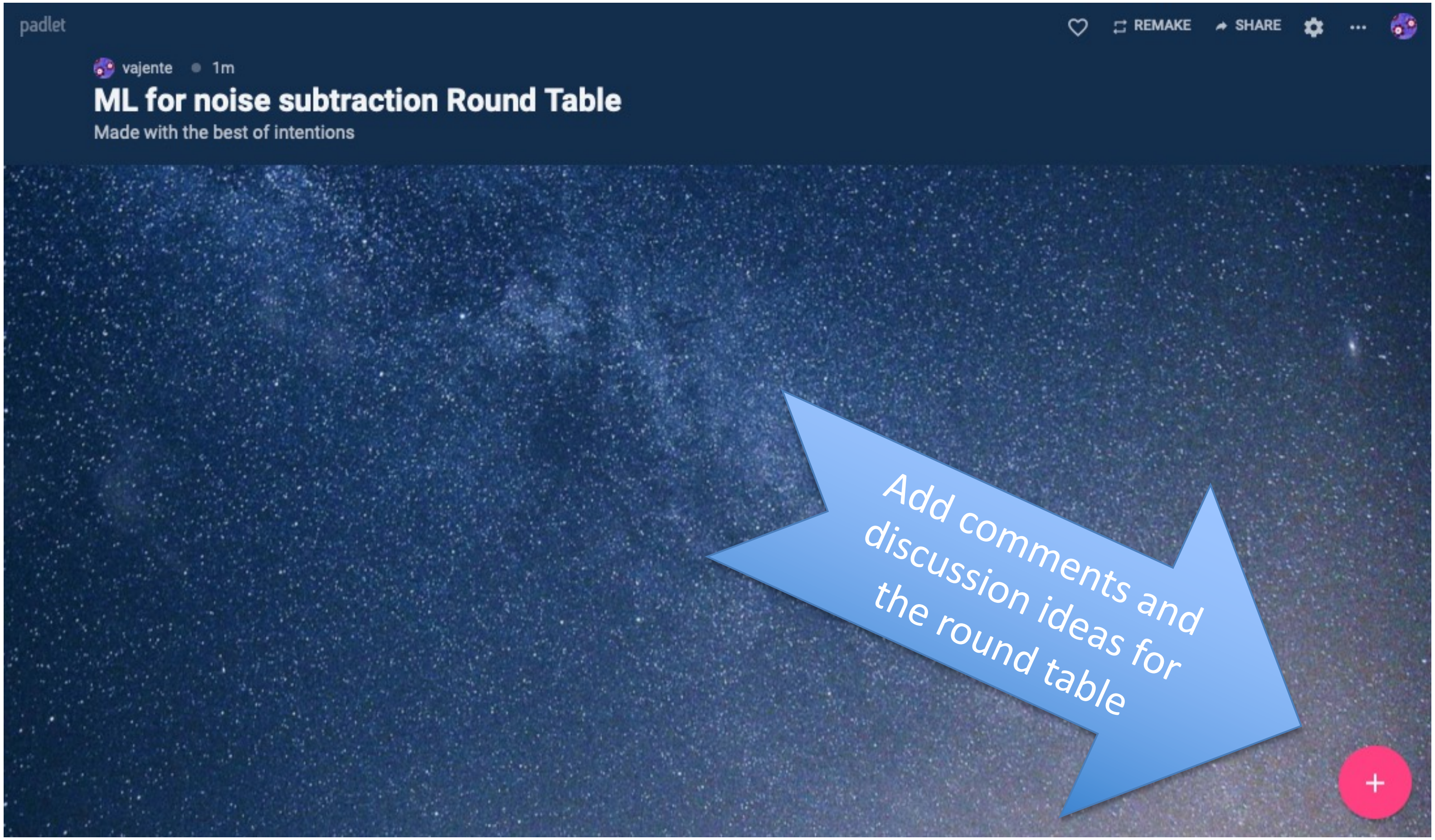
Phys. Rev. Research 2, 0330666 (2020)



Phys. Rev. D 101, 042003 (2020)



Round table: ML for noise subtraction



https://padlet.com/vajente/whthqq0lxa7nfy52