

Machine learning for quantum-enhanced gravitational-wave observatories

Chris Whittle^{1,2}, Ge Yang², Matthew Evans¹ and Lisa Barsotti^{1,2}

¹*LIGO, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

²*Institute of Artificial Intelligence and Fundamental Interaction, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*



(Received 8 June 2023; accepted 8 August 2023; published 31 August 2023)

Machine learning has become an effective tool for processing the extensive datasets produced by large physics experiments. Gravitational-wave detectors are now listening to the universe with quantum-enhanced sensitivity, accomplished with the injection of squeezed vacuum states. Squeezed state preparation and injection is operationally complicated, as well as highly sensitive to environmental fluctuations and variations in the interferometer state. Achieving and maintaining optimal squeezing levels is a challenging problem and will require development of new techniques to reach the lofty targets set by design goals for future observing runs and next-generation detectors. We use machine learning techniques to predict the squeezing level during the third observing run of the Laser Interferometer Gravitational-Wave Observatory (LIGO) based on auxiliary data streams, and offer interpretations of our models to identify and quantify salient sources of squeezing degradation. The development of these techniques lays the groundwork for future efforts to optimize squeezed state injection in gravitational-wave detectors, with the goal of enabling closed-loop control of the squeezer subsystem by an agent based on machine learning.

DOI: [10.1103/PhysRevD.108.043034](https://doi.org/10.1103/PhysRevD.108.043034)

I. INTRODUCTION

Machine learning has seen an abundance of applications in experimental physics: active control of high energy particle experiments [1], quantum state tomography [2,3], Bose-Einstein condensate preparation [4] and adaptive feedback for phase estimation [5], to name some recent uses. Laser interferometric gravitational-wave detectors like LIGO [6], Virgo [7], GEO600 [8], and KAGRA [9] are complex instruments that require hundreds of control loops to operate simultaneously. The incorporation of machine learning into the operation of these detectors is an active area of research, following the successful application of machine learning in gravitational-wave data characterization and analysis, including glitch classification [10–13], noise subtraction [14–16], waveform modeling [17] and signal searches [18]. A first demonstration of neural network-based alignment of the GEO600 optical cavities was recently performed [19].

Here we investigate the application of machine learning to minimize quantum noise in gravitational-wave detectors by optimizing the injection of squeezed states.

The third observing run (O3) of the gravitational-wave detector network saw the first demonstration of quantum-enhanced gravitational-wave detection [20–22]. At high ($\gtrsim 50$ Hz) frequencies, interferometers are limited by quantum shot noise due to the random arrival times of the uncorrelated photons that make up the electromagnetic field striking the readout photodiodes. Squeezed states can

circumvent this limit by inducing correlations between photons, thereby reducing uncertainty in the readout quadrature [23]. Squeezed vacuum states enabled a reduction of quantum noise by up to 3.2 dB in LIGO [20] and Virgo [21], and up to 6 dB in GEO600 [22].

Although already a triumph in quantum metrology and astrophysics, future gravitational-wave detectors with even greater sensitivity require further improvements. Inspection of the detector data recorded over the course of O3 reveals that it is difficult to maintain optimal squeezing performance throughout the 1-year run, with an average observed squeezing of 2.23 dB, nearly 1 dB less than the maximum observed. Moreover, there is significant variability in the observed squeezing level. For example, the histogram of squeezing levels observed in the Livingston detector [Fig. 1(c)] shows a 0.36 dB standard deviation in the distribution of squeezing levels measured throughout the run. Similar performance is reported by the Virgo (Fig. 3 in Ref. [21]) and GEO600 (Fig. 2 in Ref. [22]) detectors. Indeed, squeezed vacuum states are highly sensitive to a number of parameters of the squeezer subsystem and the interferometer: most prominently the temperature of, and pump power circulating within, the optical parametric oscillator (OPO), the thermal state of the interferometer, and the alignment of the optics directing the squeezed state into the interferometer. Beyond these direct dependencies, other fluctuations can affect one of many control systems—power in the locking field or magnitude of the local ground

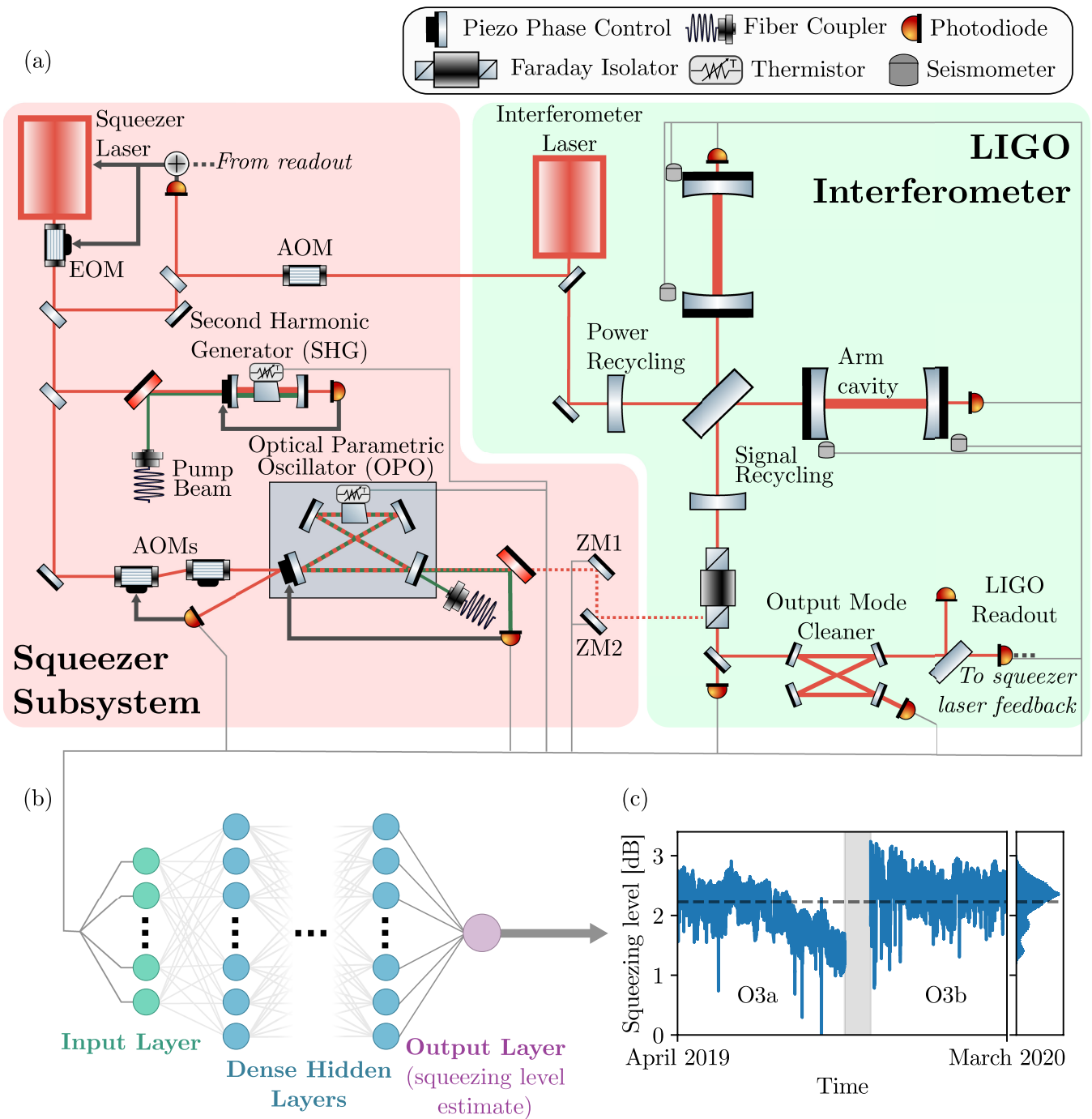


FIG. 1. (a) High-level schematic of the Advanced LIGO squeezer subsystem injection into the main interferometer. Control sidebands are imprinted onto the squeezer laser using acousto-/electro-optic modulators (A/EOMs). These sidebands are used to attain phase stability of the second harmonic generator (SHG)—used to generate the green pump field—and the optical parametric oscillator (OPO), which produces the squeezed vacuum state. Thermoelectric coolers keep each crystal at a steady temperature. Steering mirrors (ZM1 and ZM2) direct the squeezed vacuum into the interferometer. Gray lines trace paths from readbacks from the squeezer feedback loops and representative witness sensors to the input of the machine learning model. Included among these auxiliary channels are data produced by photodiodes (both power measurements and alignment error signals), seismometers and thermistors. (b) The network is comprised of a series of hidden dense layers with which it calculates an estimate of the squeezing level observed in the interferometer. The full list of channels used as input is given in Table I. (c) A time series of the squeezing level observed throughout O3, measured using the cross-correlation of the interferometer readout photodiodes, also plotted as a histogram. The dashed line denotes the mean squeezing level of 2.23 dB throughout the run.

motion for instance—and, in turn, the squeezing level realized in the interferometer. Additionally, optical filter cavities [9,24,25] have recently been installed in existing facilities to provide a reduction in quantum noise across the entire frequency band, further increasing the complexity of the squeezed state manipulation and injection. On the more distant horizon, next-generation gravitational-wave detectors demand squeezing levels up to 10 dB [26,27]. Extrapolating the effect of current degradation mechanisms to the high squeezing levels of future detectors: the fluctuation in optical loss required to explain the O3 squeezing level variation [for the middle 90% of squeezing levels; Fig. 1(c)] would cause a squeezing source capable of 10 dB of squeezing to deteriorate to below 6 dB. Consistently operating at the desired amount of squeezing in the face of increasingly complex detectors will require new methods for optimizing and tuning the injection of squeezed states throughout observing runs.

In the work described in this paper we trained neural network models to predict the squeezing level in the interferometer using various squeezer and witness channels—which monitor the physical environment of the interferometer—as input. The models were trained on historical data recorded at the LIGO Livingston Observatory (LLO) during O3. The exact inputs to the models were selected using a genetic algorithm, which additionally allowed us to identify and rank the statistical correlations present between each of the monitor channels and the squeezing level. We performed a sensitivity analysis on the resulting model to further study correlations with witness channels and identify avenues through which the squeezer might be optimized. This work is intended to lay the foundation for future closed-loop control of squeezer systems using machine learning by demonstrating the sensing side of the machine learning feedback loop. Future work will close the loop by actuating on the squeezer subsystem and interferometer to implement the optimizations we identify here.

In Sec. II we describe the experimental setup of the LIGO interferometer and the squeezed state generation. We lay out and motivate the particular model architecture used in Sec. III, and outline the specific data channels we use as inputs to our model (also in the Appendix). Section IV describes our analysis of the generated models and presents some findings on the behavior of the LIGO squeezer. A discussion of a roadmap for the future is offered in Sec. V.

II. EXPERIMENTAL SETUP

The LIGO detector is a Michelson-based laser interferometer with 4 km long arms which transduces space-time perturbations produced by the passage of gravitational waves into electrical signals at the interferometer readout. The mirrors reside inside an ultra-high-vacuum system, and are isolated from the ground by multi-stage pendulums

mounted on a seismically isolated platform. The relative distance and alignment of the interferometer’s mirrors are sensed by photodetectors and wavefront sensors respectively, and actively controlled by digital servo systems that feed control signals back to actuators which steer the mirrors’ relative longitudinal positions and orientations.

Photon shot noise is a fundamental limit of the interferometer that arises from statistical fluctuations in the photon arrival time at the interferometer’s readout port, and it scales inversely with the square root of the circulating power. Nearly 250 kW of laser light was circulating in the LIGO arm cavities during O3, requiring a thermal compensation system to compensate for the thermal lensing produced by heating from the laser beam. In addition to maximizing the amount of circulating laser light in the interferometer arms, the injection of squeezed vacuum states is a complementary method to further reduce shot noise. A simplified scheme of the LIGO squeezed vacuum system is shown in Fig. 1(a), with a full description available in [20]. It is built around a bow-tie OPO placed on a seismically isolated platform within the ultrahigh-vacuum envelope. A 1064 nm wavelength laser (“squeezer laser” in Fig. 1) is phase-locked to the interferometer laser operating at the same wavelength. From this laser, a pump 532 nm field is produced by a second harmonic generator (SHG) and injected via optical fiber into the OPO cavity. A 1064 nm squeezed field is produced by the second-order nonlinearity in the periodically poled potassium titanyl phosphate (PPKTP) crystal contained within the OPO. The squeezing level of the vacuum state produced by the OPO is determined by the amount of pump light injected and by the temperature of the nonlinear crystal, both of which are under feedback control.

The resulting squeezed vacuum field is steered via two suspended mirrors (“ZMs”) [30] to the output Faraday isolator which delivers the field *into* the readout port of the interferometer. An auxiliary control field consisting of a single radio-frequency sideband (at +3.1 MHz relative to the main laser frequency, generated by two acousto-optic modulators in series) is injected via a second optical fiber into the OPO as a proxy for sensing the phase of the squeezed vacuum field, and is additionally used to control this phase [31,32].

In full operation, the control field co-propagates with the squeezed field, and its beat note with the interferometer field is sensed at the readout photodetectors in transmission of the output mode cleaner (OMC) [33]. This signal is then fed back to the squeezer laser to lock the squeezed quadrature angle. The relative angular alignment between the squeezed and interferometer fields is sensed by wavefront sensors placed at a pickoff before the OMC [34], with alignment signals generated from the beat note between the control field and a 45 MHz sideband of the interferometer field, and fed back to the steering optics (ZM1 and ZM2 in Fig. 1).

With the squeezed field phase-stabilized and well-aligned to the interferometer, the noise at the interferometer readout will decrease at frequencies where photon shot noise dominates. The extent to which shot noise gets suppressed is dictated by the decoherence and degradation of the squeezed state along its optical path to the readout, in addition to the amount of squeezing generated within the OPO. Incoherent losses act to mix the squeezed state with the unsqueezed vacuum state. For instance, imperfect mode-matching between the optical mode produced by the OPO and the interferometer mode is a significant source of loss and is addressed by alignment of the ZMs. Moreover, residual fluctuations in the quadrature angle (“phase noise”) of the squeezed state contaminate the squeezing with the noise-amplifying antisqueezed quadrature, also leading to a reduction in the squeezing level observed.

We estimate the effective level of quantum noise reduction from the two photodetector readout signals after the OMC: the difference between these channels yields the shot noise alone, the classical noises having been removed by the common mode rejection. While the difference is unchanged by the injection of squeezing, the sum will show the reduction in noise, and we therefore build an estimate of the squeezing level using the readout sum relative to the difference. We can improve this slightly by subtracting an estimate of the classical noise sources from the readout sum, which itself is calculated by cross-correlating the readout channels at a reference time without any squeezing. Finally, the squeezing level is calculated within a frequency band 1.1–1.9 kHz where quantum noise is dominant and classical noises are minimal [20].

III. MACHINE LEARNING MODEL TO PREDICT SQUEEZING LEVELS

The primary goal of this work was to build a machine-learning-based model to accurately predict the measured squeezing level using information from auxiliary data streams (“channels”). A successful model must capture the squeezing degradation mechanisms described in Sec. II—loss, mode-matching, and phase noise—and should illuminate potentially actionable correlations with the auxiliary channels.

A. Inputs

As input to the model, we selected channels that should capture the state of both the squeezer subsystem and the general alignment of the interferometer. (See Table I in the Appendix for the complete list of channels used.)

As part of the former category, we include the pump power incident on the OPO crystal (using the power reflected and transmitted from the OPO cavity as proxies), and the temperature of the OPO crystal (as controlled by the thermoelectric cooler), both of which have a direct impact

on the amount of squeezing generated. We additionally include the temperature of the SHG crystal, used to generate the OPO pump field.

We include a number of channels to characterize the interaction between the generated squeezed state and the interferometer. The pitch and yaw of the two suspended squeezed state relay optics are the most direct signals that track this alignment. Other signals used for general alignment sensing and control (ASC) of the full interferometer are also included. We also use channels that track the amount of thermal compensation applied to the test masses as a proxy of the thermal state of the interferometer. Finally, data from seismometers tracking microseismic motion at various frequency bands from 0.1 to 1 Hz are also included in our analysis.

We used data from all of these channels over the course of O3 to train and validate our model. As we are focused here on long-term drifts in performance of the squeezer, we take the mean value of each channel over a minute timescale. For some alignment channels, we also include the root mean square (RMS) values: the average will show long-term alignment drifts while the RMS shows alignment noise. Times that coincide with large transients in the interferometer are cut from our pool of data, identified by a short-term elevation of noise or inconsistency in squeezing estimated from different frequency bands.

B. Architecture

Related previous works have used linear techniques, such as coherence calculations or lasso regression [35], to identify correlations between auxiliary channels and the astrophysical detector range. While the efficacy of these models has been demonstrated, a more effective model of the squeezer must be capable of fitting the nonlinearities associated with squeezed state generation and alignment-dependent mode-matching losses [36,37]. Compared to simply fitting to an analytical model, machine learning techniques also allow us to identify and fit to new, unexpected dependencies. To this end, we choose to use neural networks comprised of a series of hidden internal layers with nonlinear activation functions.

After the initial input layer, we additionally incorporate a random Fourier features (RFF) layer, mapping input features to random Fourier modes along random lines in parameter space. Such RFF layers allow a network to uniformly sample the Fourier space of a function, thereby allowing one to circumvent the spectral bias of standard neural networks toward low frequency features [38,39]. This technique is commonly used in spatial machine learning applications like image regression and three-dimensional object regression, and in our system it ensures we capture both fine and coarse dependencies of our squeezing level on other parameters. For example, we aim to capture both small misalignments about the optimum as well as how the squeezing level changes when

optics are highly misaligned, the former of which would be typically underfitted by standard neural networks, but is a more common scenario than the latter. Further, use of such a layer confers a practical benefit by generally offering much faster convergence when training the model.

A series of nonlinear dense layers follow this mapping, before the network outputs the estimated squeezing level. We conducted a brute-force search of the hyperparameter space to empirically identify a successful architecture by first selecting several representative time intervals from each half of O3 (i.e., O3a and O3b). A model was trained with the given hyperparameters and evaluated using separate validation data. Our search space spanned different activation functions, various counts of sequential dense layers, different dimensions of said dense layers, and the dimensionality (and presence) of the RFF layer.

We additionally tested different durations of training data; longer training intervals increase the amount of training data available but sacrifice accuracy as the interferometer undergoes long-term drifts that may not be captured by the auxiliary witness channels. By using a training data duration that is shorter than the characteristic timescale of these long-term drifts, we can still achieve a model that can accurately predict the squeezing level in the short term. We ultimately found that one month of training struck the balance between sufficient training data and model specificity. The optimal architecture was found to use two 64-dimensional dense layers with sigmoid activation functions fitting to 4096 Fourier features output from the RFF layer.

The model performance is demonstrated in Fig. 2, where it is shown predicting the squeezing level over approximately one week of O3b immediately following one month of training. The model is compared against a simple linear model, as well as a model that additionally allows for cross-terms between different witness channels. While some segments of O3 show similar performance for each of these models, there is a significant fraction of times that do benefit from capturing the nonlinear behavior to achieve better predictions, as shown in Fig. 2.

The periods of significant discrepancy between the observed squeezing and the squeezing level predicted in Fig. 2 (for instance, around 0.5–1 days) suggest that our model does not capture all dependencies of the squeezing level in the interferometer. Although more work is required to determine the exact cause, this could be due to some combination of an insufficient set of input channels, a poor choice of model architecture, not enough training data or too slow a sampling rate used. However, the sensitivity analysis performed in Sec. IV is still effective at identifying correlations and ranking channels within the set used for this model (with some higher-order error in the rankings due to these effects).

During our hyperparameter optimization, we additionally tested architectures that are well-suited to time series

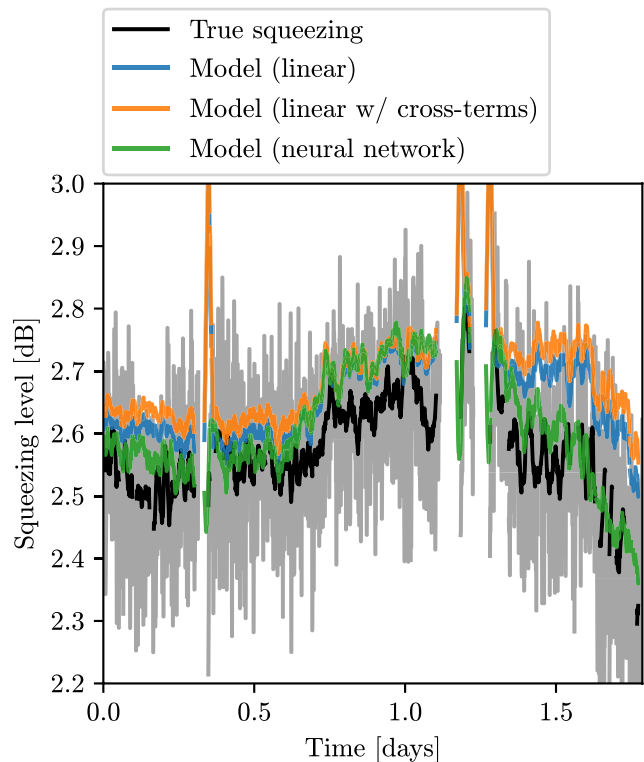


FIG. 2. Measured squeezing level calculated from cross-correlation at the interferometer readout compared with model estimates based on squeezer, alignment and environment auxiliary channels. The true and estimated squeezing levels are plotted for a validation dataset with a duration of just under two days. Although each datapoint is a 60 s average, we smooth the data with a rolling window of 20 minutes to aid readability and emphasize longer-term trends (the presmoothed true squeezing level is plotted in gray). The models are all trained on the 30 days of data preceding this validation dataset. The full nonlinear neural network model (green) is compared with two more simple models: a linear combination of witness channels (blue) and a linear model that also allows for cross-terms (orange). The gaps in data correspond with periods for which the interferometer was either unlocked or not squeezing, or other excess noise resulted in a veto from our data analysis pipeline.

data, including long short-term memory (LSTM) networks and one-dimensional convolutional neural networks (1D CNN). We ultimately discovered that these alternative architectures granted marginal improvements in accuracy ($\lesssim 5\%$ lower average loss) at the cost of a loss of interpretability. Future work could further investigate time series models, for example through use of those that specialize in interpretability of dependencies over variable timescales [40].

C. Feature selection

One of the primary goals of this study was to identify interpretable and actionable dependencies of the squeezing level on other parameters of the interferometer. To aid with

this interpretability, as well as assuage some of the other issues with high-dimensional regression like over-fitting or lengthy training processes, we performed feature selection to identify the most predictive channels among our chosen inputs, discarding the rest. This was accomplished with a genetic algorithm, used to select an optimal subset of channels as inputs for our model [41]. This approach benefits from being scalable and system-agnostic, but can be computationally intensive.

We seeded our initial populations with random subsets of the channels listed in Table I. During each iteration, models were trained on data from their respective channels over a selection of time intervals throughout O3. The average error of each model was computed from subsequent validation intervals, which was then used to weight the probability of that subset propagating to the following generation. The results of the genetic algorithm are shown in Fig. 3: the top panel shows the average loss (blue trace) of models in each generation stochastically decreasing as more accurate models are preferentially selected. We also kept track of the lowest loss observed at each iteration (orange trace)

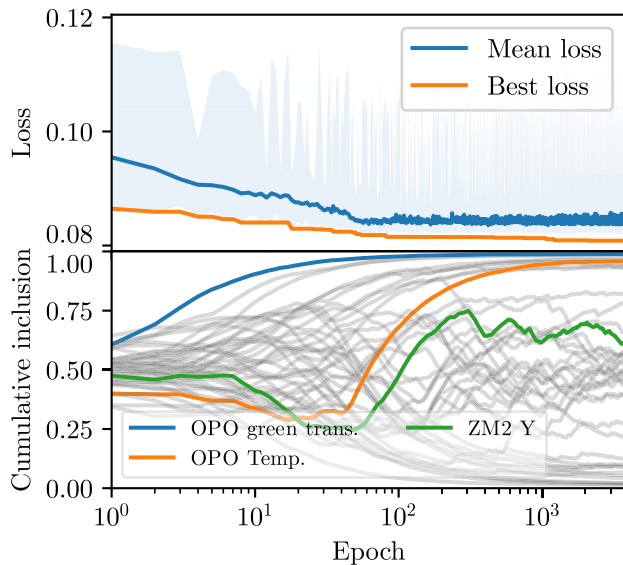


FIG. 3. Results from feature selection of the auxiliary interferometer channels using a genetic algorithm. The average loss (model error, blue line) within each generation is plotted as a function of epoch number, along with the lowest loss (orange line) model trained so far (top). To compare the relative importance placed on each channel by the algorithm, we also plot the fraction of models in which each channel is included, cumulatively averaged for each generation over its ancestors (bottom). We highlight properties on which we expect the squeezing level to depend: the laser power and temperature within the optical parametric oscillator, as well as the alignment between the squeezed field and the interferometer (for instance, the yaw of the second steering mirror, ZM2).

and ultimately used the corresponding channel subset for the subsequent analyses (marked in Table I).

IV. SENSITIVITY ANALYSIS

After building a model architecture and determining a predictive set of input channels, we looked for interpretable correlations with the detector squeezing level. The field of sensitivity analysis deals with studying how the output of a model varies with each of its inputs, as well as the relative importance of each of its features. Here, we applied sensitivity analysis techniques to identify specific channels to which the squeezing level is most sensitive, and quantify the extent of this dependence on alignment and other interferometer channels.

A. Global sensitivity analysis

Global sensitivity analyses explore the behavior of the model throughout the full extent of the input feature space, providing a holistic view of its dependencies.

A byproduct of our feature selection is that we have already computed a quantitative measure of the importance of each channel for squeezing level prediction. The average fraction of models in which each channel is included gives an indication of how that feature is being weighted. In the lower panel of Fig. 3, we plot the cumulative average inclusion of each channel as the genetic algorithm evolves. As a sanity check, we inspected channels for which we know the squeezing level has a strong dependence (see the highlighted curves): laser power and crystal temperature both appear directly in the analytical form for amount of squeezing generated by an OPO [36,42]. We likewise know that a shift in alignment into the interferometer (from e.g. motion in one of the steering mirrors ZM) should result in a change in squeezing observed. However, because the interferometer alignment can also drift, looking at the change of a single optic in isolation has a less clear-cut consequence for squeezing. The inclusion fractions for all channels are given in Table I. In addition to the channels already highlighted, we saw that some seismometer channels show strong correlation with the squeezing level, as well as a handful of alignment channels for other cavities in the interferometer.

A well-established technique in machine learning literature for studying global sensitivities is by computing Sobol indices [43]. This technique requires decomposing a predictive model $f(\mathbf{x})$, which takes n -dimensional inputs \mathbf{x} , into summands of different dimensions,

$$f(\mathbf{x}) = f_0 + \sum_i f_i(x_i) + \sum_{i<j} f_{ij}(x_i, x_j) + \dots + f_{12\dots n}(x_1, x_2, \dots, x_n). \quad (1)$$

Sobol indices are computed as variances of these constituent functions, normalized to the variance of the total model

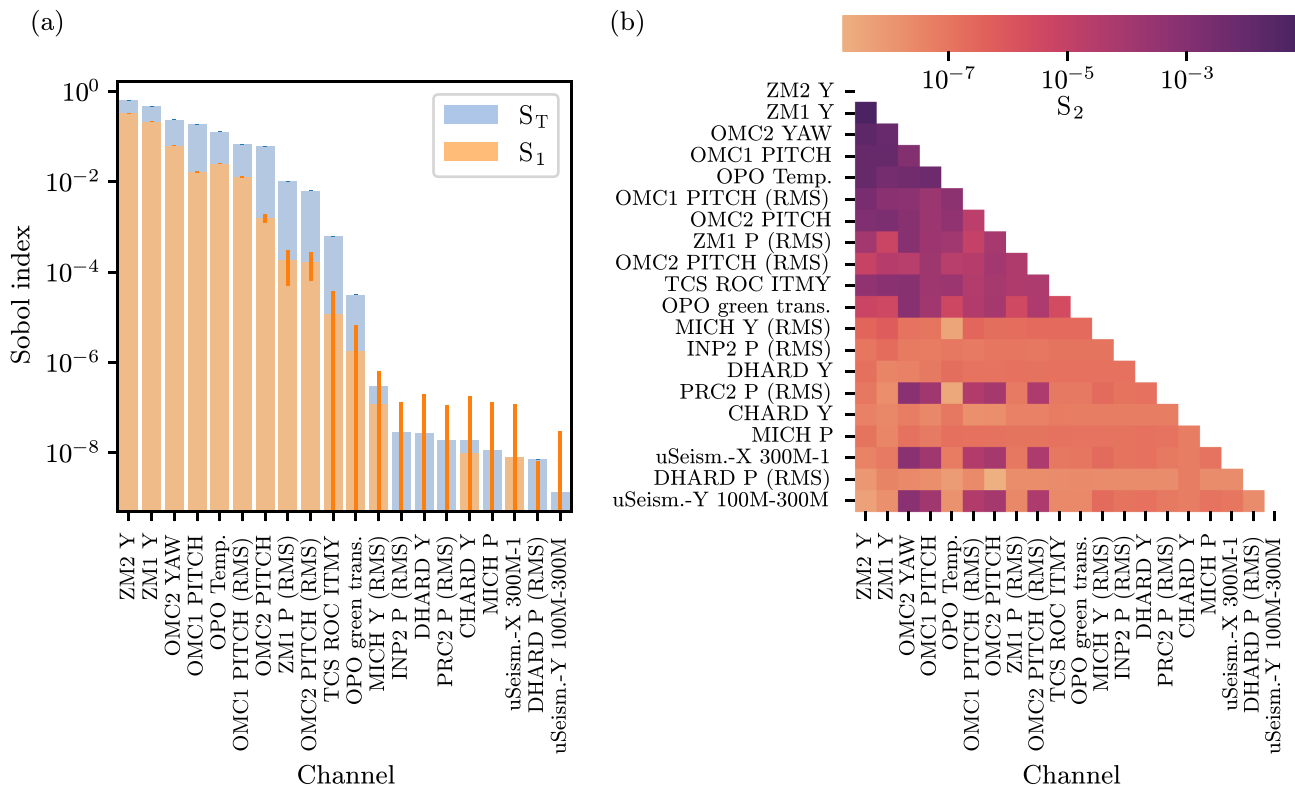


FIG. 4. (a) First-order and total Sobol indices, showing the global dependence of the squeezing model on each channel when varied independently (first-order) and jointly with other channels (total) respectively, computed by Monte Carlo sampling. The orange error bars show the 95% confidence interval for the first-order indices. (b) Second-order Sobol indices, showing the global dependence of the squeezing model when two channels are varied simultaneously.

output [44]. In this way, each Sobol index represents the contribution of a given channel x_i —or the contribution from its interaction with one or more other channels—to the output. The first-order indices S_1^i encapsulate variations due to fluctuations in just that channel x_i , while the second-order indices $S_2^{i,j}$ capture variability due to the joint variation of two channels. These effects, in addition to higher-dimension components, sum together to yield the total Sobol indices S_T^i .

We trained models on four sets of month-long training data, then used Monte Carlo sampling over each of their respective input spaces to compute the variances of the decomposed model outputs. We subsequently evaluated the Sobol indices associated with each input channel and average over each model to produce a plot of first-, second- and total-order indices, shown in Fig. 4. The computational cost associated with sampling this high-dimensional input space demands the dimensionality reduction we performed through feature selection in Sec. III C—more than a handful of additional channels would render this approach infeasible.

The first- and total-order sensitivities highlight a particular sensitivity of the model to drifts in the yaw of the steering mirrors (ZM1 and ZM2). Fluctuations in pitch, which manifest through the RMS channels, of the first

steering mirror and the output mode cleaner (OMC) optics are also shown to be significant, suggesting that improvements in the OMC pitch alignment loops could lead to more stable squeezing levels. Further investigation is required to determine the reason for the sensitivity to these particular degrees of freedom. The second-order indices seem to mostly track the sensitivities depicted by the first- and total-order indices.

B. Local sensitivity analysis

We now present a local sensitivity analysis of our model, in which we study how the model dependencies vary throughout the input space. For instance, we might expect the model sensitivities to change at times of low versus high seismic activity or in different alignments of the interferometer. To identify these different operational regimes occupied throughout O3 in our high-dimensional input data (e.g. “high seismic activity”), we performed a k -means clustering over all of the input channels (Fig. 5). In this process, we compute five [45] points in parameter space that act as “cluster centroids.” Each O3 data point (time sample) can then be assigned to the cluster associated with the nearest cluster centroid. The positions of these centroids are set to minimize the total sum of distances between data points and their assigned

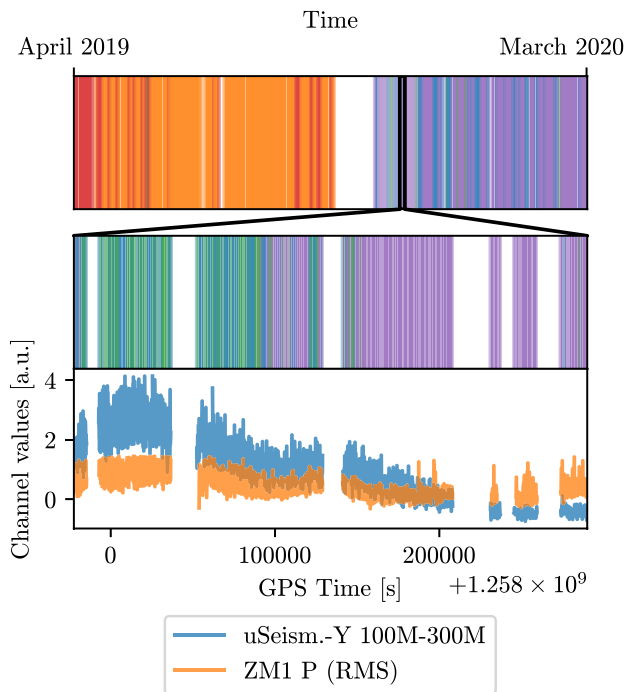


FIG. 5. The segmentation of O3 data into k -means clusters, where each of the five colors represent a quantitatively distinct operating point. Times with similar interferometer behavior (according to the k -means clustering) are grouped together by color. The zoomed-in middle panel shows the cluster partitioning over a period of ~ 3 days. The blue and green clusters are characterized by heightened levels of seismic noise in comparison with nominal operating conditions (the purple cluster). The timeseries in the bottom panel highlights this by showing how the power in the 0.1–0.3 Hz frequencies in the Y-arm seismometer and the fluctuations in the ZM1 pitch vary over these ~ 3 days (from relatively high to low).

centroids. In this way, we partition our O3 time series data into time periods of similar behavior.

The partial derivative of the network output can be computed with respect to each of the input factors and evaluated at particular fixed points. We trained models over various time intervals, as in Sec. IV A, and computed partial derivatives at each of the cluster centroids calculated from k -means, using them as points representative of the interferometer behavior in each of the operating regimes. We then compared the magnitudes of the partial derivatives, after normalizing by the standard deviation of each input channel, to evaluate sensitivity to each of the input parameters. Figure 6 shows this procedure for a month of training data in O3b, evaluated at a typical O3b interferometer configuration [Fig. 6(a)] and in a regime with significant micro-seismic activity and alignment noise [Fig. 6(b)]. We see that the relative effects of fluctuations in various alignment signals change during times with substantial ground motion. We also observe that some channels to which the squeezing level is globally sensitive

(shown in Fig. 4) are locally insensitive at particular points in parameter space.

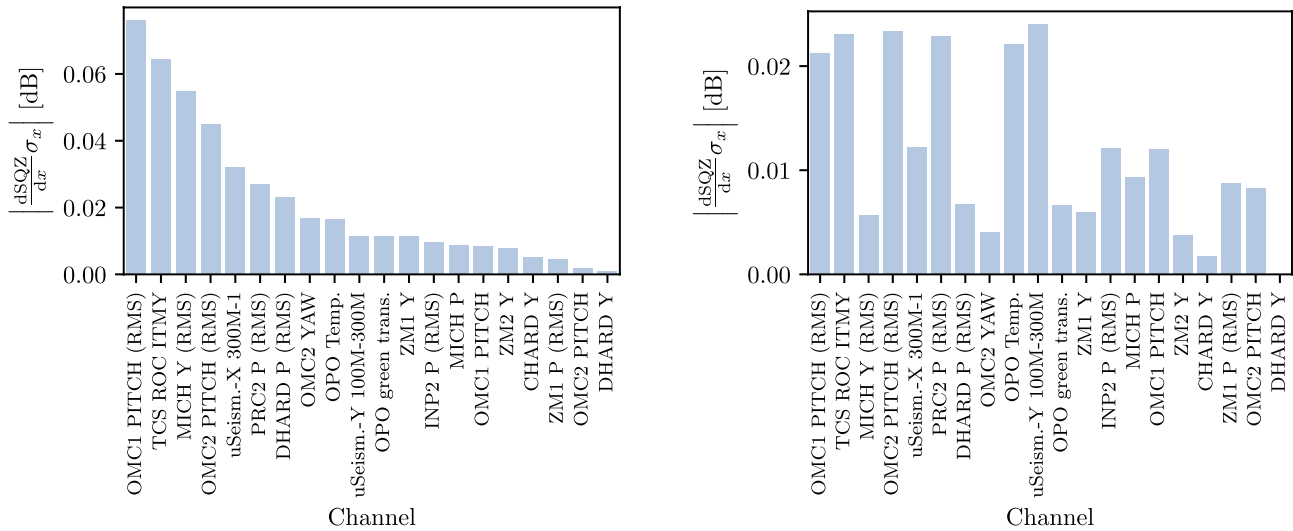
In principle, the derivatives computed can be used to determine actions that can be taken to optimize squeezing. For instance, a nonzero derivative in the OPO temperature suggests that it is suboptimal, and the sign of the derivative will indicate the direction in which it should be adjusted. Note, however, that this analysis ignores interactions between channels, which can somewhat muddy direct interpretation of these local sensitivity indices (and, in particular, their signs).

V. ROADMAP FOR THE FUTURE

The work described here demonstrated that machine learning models are capable of sensing the state of the squeezer and interferometer (Sec. III), and can produce actionable diagnostic information about sources of squeezing degradation (Sec. IV). A fruitful direction for future work is toward continuous optimization of the squeezer based on these real-time machine-learning-based prognoses. An artificial intelligent (AI) agent that replaces *ad hoc* and infrequent manual tuning of the squeezing system with a controller that is *always on* may be a critical step toward achieving the 10 dB suppression of quantum noise required by the next-generation gravitational wave observatories [26,27]. This offers a fertile ground for importing ideas from other fields, and the development of new control methods.

Capturing all of the complexity of gravitational-wave interferometers in a model can be challenging, which makes applying analytical control methods such as model-predictive control difficult. An alternative is to adopt learning-based approaches, such as model-free reinforcement learning (RL), that are able to produce control behaviors similar to those from the human operators by learning via trial-and-error in a physics simulator. “Learning-to-control” schemes as such have so far achieved impressive results on complex games such as the ATARI arcade learning environments (ALE, see Bellemare *et al.* [46]), StarCraft, and Dota, without making explicit assumptions on the underlying decision processes [47–49]. This also comes with the added benefit that the AI agent can continuously improve in performance during deployment, given additional supervision in the form of reinforcement.

Past learning-based control methods struggle with learning and operating under confounding signals. The feature analysis in Sec. III C helps by removing confounding signals through an information-theoretic approach that has been shown to improve learning on complex learning tasks [50]. Recent results in sim-to-real transfer approaches, which show promise in robotics for implicitly identifying system parameters under partial observability [51,52], may also be applicable to gravitational-wave detectors.



(a) The model gradients computed at the cluster centroid representative of typical operation during O3b (shown in Fig. 5 in purple).

(b) The model gradients computed at the cluster centroid characterized by higher-than-usual micro-seismic activity and alignment noise (shown in Fig. 5 in blue).

FIG. 6. Partial derivatives shown at different points in parameter space for a neural network trained on ~ 1 month of O3b data. The derivative of the predicted squeezing level with respect to each auxiliary channel $\partial\text{SQZ}/\partial x$ is shown. Each derivative is multiplied by the auxiliary channel standard deviation σ_x , giving the squeezing level variations that result from typical fluctuations in the given channel. The channels are listed in order of decreasing sensitivity, according to the partial derivatives, during nominal O3b operation. This analysis hints at the variables on which one should actuate to optimize the squeezing level in the interferometer in a particular regime.

Any move toward machine-learning techniques and learning-to-control will also necessitate next-generation infrastructure in the data acquisition, storage, analysis, and control pipeline [53]. Figure 7 illustrates a simplified

example architecture for a computational and inference infrastructure capable of running a learned controller.

VI. CONCLUSIONS

The application of machine learning in experimental physics is growing, and it is providing new methods to optimize performance. Here we describe an exploration of trained neural network models that estimate the quantum-enhanced performance of the LIGO Livingston detector via squeezing injection, using historical data from the past observing run O3. We have demonstrated the sensing, processing and interpretation steps of our training models, including various sensitivity analysis techniques to identify correlations between auxiliary channels and the squeezing observed in the interferometer. Specific alignment signals for the squeezed state steering mirrors and other cavities within the interferometer have been found to be the most correlated. The most computationally expensive processes in our pipeline are the hyperparameter optimization, feature selection steps and, to a lesser extent, computation of Sobol indices. Once a network architecture has been selected, training, evaluation and local sensitivity analysis is relatively lightweight, making the approaches demonstrated here inexpensive.

Our analysis shows that machine learning could be used to continuously monitor the performance of the squeezer system and the interferometer in a high-dimensional parameter space to identify causes of squeezing degradation.

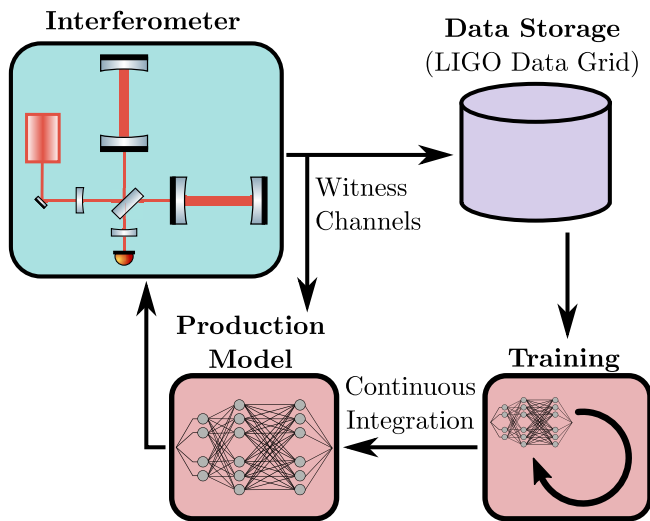


FIG. 7. Proposed workflow for optimization of the squeezer (or other) subsystem with machine learning. In this scenario, the latest trained model continuously runs on the interferometer data as it is recorded, identifying possible optimizations and feeding back to the LIGO real-time control system. Meanwhile, additional hardware runs training on the latest data, with the updated model weights continuously deployed to the production model.

Such continuous monitoring of squeezing degradation sources is presently not possible in existing gravitational-wave detectors, and would be a significant step toward the realization of persistently high levels of squeezing. The next step of this work would be to use this newly acquired information to actuate on the system and maximize squeezing levels. This approach could be extended to other quantum precision experiments in which maintaining optimal performance for long periods of time is important.

Future gravitational-wave detectors like Cosmic Explorer [26] will require high levels of squeezing (up to 10 dB) to reach their sensitivity targets. In the next decade AI might become not just a useful tool, but an indispensable one to guarantee optimal performance over long periods of time. Moreover, similar techniques can evolve to address many other interferometer control problems, with the potential to revolutionize the way current gravitational-wave detectors are controlled, and the way future observatories are designed.

ACKNOWLEDGMENTS

LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation, and operates under Cooperative Agreement No. PHY-1764464. Advanced LIGO was built under Grant No. PHY-0823459.

This work is also supported by the National Science Foundation under Cooperative Agreement No. PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions). The authors acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this paper. We thank Maggie Tse for the initial intuition that squeezing optimization could benefit from an AI agent, for brainstorming on a possible approach for this study, and for providing comments on the manuscript. We thank Lee McCuller, Dhruva Ganapathy, Wenxuan Jia, and Kevin Kuns for brainstorming in the initial phase of the project on causes of squeezing degradation. We thank Pulkit Agrawal for discussions on reinforcement learning, and Erik Katsavounidis, Philip Harris and Alec Gunny for discussions on deployment of machine learning-based algorithms in LIGO. Finally, we thank Nikhil Mukund for carefully reading the manuscript and providing comments, and Hang Yu for his comments as part of the LIGO review process. This paper has LIGO Document Number LIGO-P2300114.

APPENDIX: CHANNEL LIST

The channels used as input for this analysis are listed in Table I.

TABLE I. List of channels used as input for the squeezing level estimator, listed both by their standardized channel name as well as more readable, abbreviated versions. The mean and rms suffixes refer to the statistic computed over 60 s intervals. The average inclusion fraction of each channel during the feature selection process (Fig. 3) is also shown: a higher number indicates a greater importance of the channel for squeezing level estimation. Channels marked with a circle (•) were included in the optimal subset found during feature selection and used for subsequent sensitivity analysis (Sec. IV).

LIGO channel name	Description	Average inclusion fraction
Interferometer alignment signals (see Ref. [28])		
ASC-CHARD_P_OUTPUT.mean,m-trend	CHARD	0.40
ASC-CHARD_P_OUTPUT.rms,m-trend	CHARD P (RMS)	0.07
• ASC-CHARD_Y_OUTPUT.mean,m-trend	CHARD Y	0.60
ASC-CHARD_Y_OUTPUT.rms,m-trend	CHARD Y (RMS)	0.10
ASC-DHARD_P_OUTPUT.mean,m-trend	DHARD P	0.36
• ASC-DHARD_P_OUTPUT.rms,m-trend	DHARD P (RMS)	0.40
• ASC-DHARD_Y_OUTPUT.mean,m-trend	DHARD Y	0.40
ASC-DHARD_Y_OUTPUT.rms,m-trend	DHARD Y (RMS)	0.06
ASC-INP2_P_OUTPUT.mean,m-trend	INP2 P	0.03
• ASC-INP2_P_OUTPUT.rms,m-trend	INP2 P (RMS)	0.99
ASC-INP2_Y_OUTPUT.mean,m-trend	INP2 Y	0.02
ASC-INP2_Y_OUTPUT.rms,m-trend	INP2 Y (RMS)	0.02
• ASC-MICH_P_OUTPUT.mean,m-trend	MICH P	0.93
ASC-MICH_P_OUTPUT.rms,m-trend	MICH P (RMS)	0.93
ASC-MICH_Y_OUTPUT.mean,m-trend	MICH Y	0.44
• ASC-MICH_Y_OUTPUT.rms,m-trend	MICH Y (RMS)	0.95
• OMC-ASC_P1_I_OUTPUT.mean,m-trend	OMC1 PITCH	0.82
• OMC-ASC_P1_I_OUTPUT.rms,m-trend	OMC1 PITCH (RMS)	0.98
• OMC-ASC_P2_I_OUTPUT.mean,m-trend	OMC2 PITCH	0.46

(Table continued)

TABLE I. (Continued)

LIGO channel name	Description	Average inclusion fraction
Interferometer alignment signals (see Ref. [28])		
• OMC-ASC_P2_I_OUTPUT.rms,m-trend	OMC2 PITCH (RMS)	0.28
OMC-ASC_Y1_I_OUTPUT.mean,m-trend	OMC1 YAW	0.65
OMC-ASC_Y1_I_OUTPUT.rms,m-trend	OMC1 YAW (RMS)	0.05
• OMC-ASC_Y2_I_OUTPUT.mean,m-trend	OMC2 YAW	0.69
OMC-ASC_Y2_I_OUTPUT.rms,m-trend	OMC2 YAW (RMS)	0.08
ASC-PRC2_P_OUTPUT.mean,m-trend	PRC2 P	0.50
• ASC-PRC2_P_OUTPUT.rms,m-trend	PRC2 P (RMS)	0.62
ASC-PRC2_Y_OUTPUT.mean,m-trend	PRC2 Y	0.43
ASC-PRC2_Y_OUTPUT.rms,m-trend	PRC2 Y (RMS)	0.27
Ground motion signals (in the 100–300 mHz band)		
ISI-GND_STS_ETMX_Y_BLRMS_100M_300M.mean,m-trend	uSeism.-X 100M-300M	0.14
• ISI-GND_STS_ETMX_Y_BLRMS_300M_1.mean,m-trend	uSeism.-X 300M-1	0.96
• ISI-GND_STS_ETMY_Y_BLRMS_100M_300M.mean,m-trend	uSeism.-Y 100M-300M	0.95
ISI-GND_STS_ETMY_Y_BLRMS_300M_1.mean,m-trend	uSeism.-Y 300M-1	0.09
Squeezer OPO and SHG trends (see Ref. [20])		
• SQZ-OPO_IR_LF_OUTPUT.mean,m-trend	OPO green transmission	0.99
SQZ-OPO_REFL_DC_POWER.mean,m-trend	OPO green reflection	0.02
• SQZ-OPO_TEC_THERMISTOR_TEMPERATURE.mean,m-trend	OPO temperature	0.95
SQZ-SHG_TEC_THERMISTOR_TEMPERATURE.mean,m-trend	SHG temperature	0.02
Squeezed beam steering mirror signals (see Ref. [20])		
SUS-ZM1_M1_DAMP_P_OUTPUT.mean,m-trend	ZM1 P	0.57
• SUS-ZM1_M1_DAMP_P_OUTPUT.rms,m-trend	ZM1 P (RMS)	0.73
• SUS-ZM1_M1_DAMP_Y_OUTPUT.mean,m-trend	ZM1 Y	0.55
SUS-ZM1_M1_DAMP_Y_OUTPUT.rms,m-trend	ZM1 Y (RMS)	0.09
SUS-ZM2_M1_DAMP_P_OUTPUT.mean,m-trend	ZM2 P	0.53
SUS-ZM2_M1_DAMP_P_OUTPUT.rms,m-trend	ZM2 P (RMS)	0.47
• SUS-ZM2_M1_DAMP_Y_OUTPUT.mean,m-trend	ZM2 Y	0.69
SUS-ZM2_M1_DAMP_Y_OUTPUT.rms,m-trend	ZM2 Y (RMS)	0.44
Interferometer thermal state signal (see Ref. [29])		
• TCS-SIM_ITMY_SURF_ROC_FULL_OUTPUT.mean,m-trend	TCS ROC ITMY	0.97

- [1] V. Kain, S. Hirlander, B. Goddard, F. M. Velotti, G. Z. Della Porta, N. Bruchon, and G. Valentino, *Phys. Rev. Accel. Beams* **23**, 124801 (2020).
- [2] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Nat. Phys.* **14**, 447 (2018).
- [3] H.-Y. Hsieh, Y.-R. Chen, H.-C. Wu, H. L. Chen, J. Ning, Y.-C. Huang, C.-M. Wu, and R.-K. Lee, *Phys. Rev. Lett.* **128**, 073604 (2022).
- [4] Z. Vendeiro, J. Ramette, A. Rudelis, M. Chong, J. Sinclair, L. Stewart, A. Urvoy, and V. Vuletić, *Phys. Rev. Res.* **4**, 043216 (2022).
- [5] A. Hentschel and B. C. Sanders, *Phys. Rev. Lett.* **104**, 063603 (2010).
- [6] The LIGO Scientific Collaboration, *Classical Quantum Gravity* **32**, 074001 (2015).
- [7] F. Acernese *et al.*, *Classical Quantum Gravity* **32**, 024001 (2014).
- [8] K. L. Dooley *et al.*, *Classical Quantum Gravity* **33**, 075009 (2016).
- [9] T. Akutsu, M. Ando, K. Arai, and Y. Arai (KAGRA Collaboration), *Nat. Astron.* **3**, 35 (2019).
- [10] R. Biswas, L. Blackburn, J. Cao, R. Essick, K. A. Hodge, E. Katsavounidis, K. Kim, Y.-M. Kim, E.-O. Le Bigot, C.-H. Lee, J. J. Oh, S. H. Oh, E. J. Son, Y. Tao, R. Vaulin, and X. Wang, *Phys. Rev. D* **88**, 062003 (2013).
- [11] J. Powell, D. Trifirò, E. Cuoco, I. S. Heng, and M. Cavaglià, *Classical Quantum Gravity* **32**, 215012 (2015).
- [12] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. K. Katsaggelos,

- S. L. Larson, T. K. Lee, C. Lintott, T. B. Littenberg, A. Lundgren, C. Østerlund, J. R. Smith, L. Trouille, and V. Kalogera, *Classical Quantum Gravity* **34**, 064003 (2017).
- [13] N. Mukund, S. Abraham, S. Kandhasamy, S. Mitra, and N. S. Philip, *Phys. Rev. D* **95**, 104059 (2017).
- [14] R. Ormiston, T. Nguyen, M. Coughlin, R. X. Adhikari, and E. Katsavounidis, *Phys. Rev. Res.* **2**, 033066 (2020).
- [15] G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepańczyk, and S. Vitale, *Phys. Rev. D* **101**, 042003 (2020).
- [16] H. Yu and R. X. Adhikari, *Front. Artif. Intell.* **5**, 811563 (2022).
- [17] Z. Doctor, B. Farr, D. E. Holz, and M. Pürrer, *Phys. Rev. D* **96**, 123011 (2017).
- [18] P. T. Baker, S. Caudill, K. A. Hodge, D. Talukder, C. Capano, and N. J. Cornish, *Phys. Rev. D* **91**, 062004 (2015).
- [19] N. Mukund, J. Lough, A. Bisht, H. Wittel, S. L. Nadji, C. Affeldt, F. Bergamin, M. Brinkmann, V. Kringel, H. Lück, M. Weinert, and K. Danzmann, [arXiv:2301.06221](https://arxiv.org/abs/2301.06221).
- [20] M. Tse *et al.*, *Phys. Rev. Lett.* **123**, 231107 (2019).
- [21] F. Acernese *et al.* (Virgo Collaboration), *Phys. Rev. Lett.* **123**, 231108 (2019).
- [22] J. Lough, E. Schreiber, F. Bergamin, H. Grote, M. Mehmet, H. Vahlbruch, C. Affeldt, M. Brinkmann, A. Bisht, V. Kringel *et al.*, *Phys. Rev. Lett.* **126**, 041102 (2021).
- [23] L. Barsotti, J. Harms, and R. Schnabel, *Rep. Prog. Phys.* **82**, 016905 (2018).
- [24] L. McCuller, C. Whittle, D. Ganapathy, K. Komori, M. Tse, A. Fernandez-Galiana, L. Barsotti, P. Fritschel, M. MacInnis, F. Matichard, K. Mason, N. Mavalvala, R. Mittleman, H. Yu, M. E. Zucker, and M. Evans, *Phys. Rev. Lett.* **124**, 171102 (2020).
- [25] Y. Zhao *et al.*, *Phys. Rev. Lett.* **124**, 171101 (2020).
- [26] M. Evans *et al.*, [arXiv:2109.09882](https://arxiv.org/abs/2109.09882).
- [27] P. Jones, T. Zhang, H. Miao, and A. Freise, *Phys. Rev. D* **101**, 082002 (2020).
- [28] L. Barsotti, M. Evans, and P. Fritschel, *Classical Quantum Gravity* **27**, 084026 (2010).
- [29] A. F. Brooks *et al.*, *Appl. Opt.* **55**, 8256 (2016).
- [30] B. J. J. Slagmolen, A. J. Mullavey, J. Miller, D. E. McClelland, and P. Fritschel, *Rev. Sci. Instrum.* **82**, 125108 (2011).
- [31] H. Vahlbruch, S. Chelkowski, B. Hage, A. Franzen, K. Danzmann, and R. Schnabel, *Phys. Rev. Lett.* **97**, 011101 (2006).
- [32] S. Chelkowski, H. Vahlbruch, K. Danzmann, and R. Schnabel, *Phys. Rev. A* **75**, 043814 (2007).
- [33] H. Grote, K. Danzmann, K. L. Dooley, R. Schnabel, J. Slutsky, and H. Vahlbruch, *Phys. Rev. Lett.* **110**, 181101 (2013).
- [34] E. Schreiber, K. L. Dooley, H. Vahlbruch, C. Affeldt, A. Bisht, J. R. Leong, J. Lough, M. Prijatelj, J. Slutsky, M. Was, H. Wittel, K. Danzmann, and H. Grote, *Opt. Express* **24**, 146 (2016).
- [35] M. Walker, A. F. Agnew, J. Bidler, A. Lundgren, A. Macedo, D. Macleod, T. Massinger, O. Patane, and J. R. Smith, *Classical Quantum Gravity* **35**, 225002 (2018).
- [36] M. J. Collett and C. W. Gardiner, *Phys. Rev. A* **30**, 1386 (1984).
- [37] L. McCuller, S. E. Dwyer, A. C. Green, H. Yu, K. Kuns, L. Barsotti *et al.*, *Phys. Rev. D* **104**, 062006 (2021).
- [38] A. Rahimi and B. Recht, *Advances in Neural Information Processing Systems* (Curran Associates, Red Hook, New York, 2007), Vol. 20.
- [39] M. Tancik, P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. Barron, and R. Ng, *Advances in Neural Information Processing Systems* (2020), Vol. 33, p. 7537.
- [40] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, *Int. J. Forecast.* **37**, 1748 (2021).
- [41] B. Xue, M. Zhang, W. N. Browne, and X. Yao, *IEEE Trans. Evol. Comput.* **20**, 606 (2016).
- [42] K. Goda, K. McKenzie, E. E. Mikhailov, P. K. Lam, D. E. McClelland, and N. Mavalvala, *Phys. Rev. A* **72**, 043819 (2005).
- [43] I. M. Sobol, *Math. Comput. Simul.* **55**, 271 (2001).
- [44] In particular, the global sensitivity (Sobol) indices are computed as

$$S_{i_1 \dots i_s} = \int f_{i_1 \dots i_s}^2 dx_{i_1} \dots dx_{i_s} / \left(\int f^2 dx - f_0^2 \right).$$

See Ref. [43] for more details on this definition.

- [45] We empirically determined that five cluster centroids was sufficiently numerous to capture different regimes that may be interesting, and few enough that we can ascribe some intuitive physical meaning to each.
- [46] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling, *J. Artif. Intell. Res.* **47**, 253 (2013).
- [47] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu *et al.*, *Nature (London)* **518**, 529 (2015).
- [48] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu *et al.*, *Nature (London)* **575**, 350 (2019).
- [49] C. Berner, G. Brockman, B. Chan, V. Cheung *et al.*, [arXiv:1912.06680](https://arxiv.org/abs/1912.06680).
- [50] X. Fu, G. Yang, P. Agrawal, and T. Jaakkola, in *Proceedings of the 38th International Conference on Machine Learning*, Proceedings of Machine Learning Research Vol. 139, edited by M. Meila and T. Zhang (PMLR, 2021), pp. 3480–3491.
- [51] A. Kumar, Z. Fu, D. Pathak, and J. Malik, in *Robotics Science and Systems* (2021).
- [52] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, *Sci. Rob.* **5**, eabc5986 (2020).
- [53] A. Gunny, D. Rankin, J. Krupa, M. Saleem, T. Nguyen, M. Coughlin, P. Harris, E. Katsavounidis, S. Timm, and B. Holzman, *Nat. Astron.* **33**, 075009 (2022).