2021 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, OCT. 25–28 2021, GOLD COAST, AUSTRALIA

FAST AND ACCURATE GRAVITATIONAL-WAVE MODELLING WITH PRINCIPAL COMPONENT REGRESSION

Cyril Cano[1] *Éric Chassande-Mottin*[2] *Nicolas Le Bihan*[1]

[1] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-Lab, F-38000 Grenoble, France
[2] Université de Paris, CNRS, Astroparticule et Cosmologie, F-75006 Paris, France

ABSTRACT

Inference from gravitational-wave observations relies on the availability of accurate theoretical waveform models to compare with the data. This paper considers the rapid generation of surrogate time-domain waveforms consistent with the gravitational-wave signature of the merger of spin-aligned binary black holes. Building on previous works, a machine-learning model is proposed that allows for highly-accurate waveform regression from a set of examples. An improvement of about an order of magnitude in accuracy with respect to the state of the art is demonstrated, along with a significant speed up in computing time with respect to the reference generation software tools.

Index Terms— Principal component regression – Chirp signals – Astrophysics – Gravitational waves

1. MOTIVATIONS

The first observation of gravitational waves by the LIGO/Virgo collaboration in 2015 [1] marked the advent of gravitational-wave astronomy. Since this date about 50 events have been detected [2] associated to the merger of compact star binaries, primarily binary black holes, the focus of this work.

Theoretical waveform models (or templates) are used to detect the gravitational-wave signals using matched filtering [2], or to infer the astrophysical parameters of their source using Bayesian samplers [3]. For both those tasks, a large number of template waveforms (of the order of 10^5 to 10^6) are required to be computed to cover the relevant parameter space.

Waveform models are deduced from the resolution of the source dynamics which is a difficult relativistic problem (see [4] for a recent review). The evaluation of recent and accurate waveform models is computationally expensive. In fact, this computation dominates the computational budget for parameter inference.

In the next decade, LIGO and Virgo detectors are expected to conduct at least two major observing runs with improved sensitivity, leading to a large increase in the number of detected signals. The analysis of those future observations calls for numerically efficient, yet accurate waveform generators.

The waveform morphology being reasonably smooth and slowly varying with respect to its parameters (i.e., the binary component masses and their spins), it is possible to fit a signal model based on a generic amplitude and phase evolution from a set of example waveforms. This has been successfully realized using reduced-order modelling (based on tensor spline fitting) [5, 6] and more standard machine learning approach such as mixture-of-experts regression [7] and artificial neural networks [8, 9].

In this work, we build on the latter and propose a model based on principal component regression able to improve the overall regression accuracy by about an order of magnitude. Also, this model has a reduced complexity. It can be implemented using off-the-shelf algorithms from `Scikit-learn` software library [10], resulting in a more compact code, which is easier to maintain. These performances are notably achieved thanks to a different choice of features.

The outline of this article is as follows: in Section 2, we provide background information on the gravitational-wave signals and their characterization as amplitude and frequency modulated signals. In Section 3 we describe the regression model composed of a data reduction step using principal component analysis (PCA), followed by a regression step using linear regression of polynomial combinations of the original features. Section 4 details the application of the model and the optimization of its hyperparameters, and shows its performances in terms of accuracy and computation time.

2. GRAVITATIONAL-WAVE WAVEFORMS

Gravitational waves are oscillations of the space-time curvature propagating through space [1]. They are caused by the acceleration of asymmetrical mass distributions. There is a variety of potential astrophysical sources, however the only sources detected so far are coalescences of compact star binaries, composed of either black holes and/or neutron stars. A binary system of two black holes (BBH) orbiting around each other loses energy by emission of gravitational radiation inducing an orbital decay till the two objects eventually merge. The details of the trajectory depend on 8 astrophysical param-

978-1-7281-6662-9/20/$31.00

eters, namely m_i for $i = 1, 2$ the mass of each black hole and $S_i \in \mathbb{R}^3$ for $i = 1, 2$ their respective spins. Here we focus on "spin-aligned" BBH where the spins of both objects are aligned with the orbital angular momentum $\vec{L}$ i.e., the z axis normal to the orbital plane. In this case, the parameter space reduces to 4 parameters $\{m_1, m_2, S_{1z}, S_{2z}\}$ [11] as we have $S_{ix} = S_{iy} = 0$ for $i = 1, 2$. The gravitational-wave signal is also determined by extrinsic parameters linked to the position and orientation of the detector relative to the source. These include the luminosity distance r in megaparsecs (Mpc), i.e. the distance between the source and the observer, and the direction of the line of sight parametrized by the two angles (ι, φ), namely the inclination of the source and the initial phase. The parameters are depicted in Figure 1.

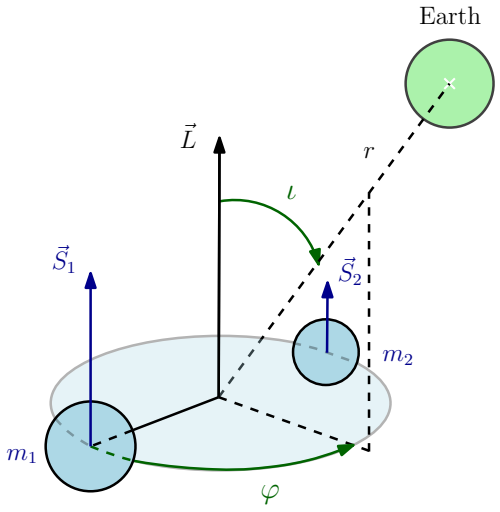

Fig. 1. Physical parameters of a compact star binary – Intrinsic parameters $\{m_1, m_2, S_1, S_2\}$ (component masses and spins) and extrinsic parameters such as the direction of the line of sight (ι, φ) and the distance r. For spin-aligned binaries component spins are aligned with the orbital angular momentum $\vec{J}$ so the intrinsic parameters reduce to $\{m_1, m_2, S_{1z}, S_{2z}\}$.

General relativity predicts the existence of two polarization modes for gravitational waves, the *plus* and *cross* polarizations, denoted $h_+(t)$ and $h_\times(t)$. It is customary to express the two polarizations as a single complex-valued time series $h(t) = h_+(t) - \boldsymbol{i}h_\times(t)$ referred to as gravitational-wave strain.

The computation of $h(t)$ needs to resolve the source dynamics. There are no exact close-form solution for this relativistic problem, however accurate approximations are available. In the spin-aligned case and for binaries with low mass ratio q (defined as $q = m_1/m_2$ and with $q \geq 1$) and low total mass $M = m_1 + m_2$, it is sufficient to retain the dominant $(\ell = 2, m = \pm 2)$ mode in the spin-weighted spherical harmonics expansion of $h(t)$ [4], thus leading to the following expression:

$$h(t) \propto \frac{M}{r} a(t) \Big[\frac{1 + \cos^2 \iota}{2} \cos(\Phi(t) - 2\varphi) - \boldsymbol{i} \cos \iota \sin(\Phi(t) - 2\varphi) \Big] \quad (1)$$

where $a(t)$ and $\Phi(t)$ stand for the amplitude and phase of the dominant mode.

General relativity's scale invariance implies $h(t; m_1, m_2) = h(\lambda t; \lambda m_1, \lambda m_2)$ [7]. The total mass can thus be fixed to a fiducial value $M^* = 20\, M_\odot$. For an arbitrary total mass M, $h(t; M)$ can be deduced from $h(t; M^*)$ by compressing/stretching the time axis by a factor of M/M^*. Consequently, for spin-aligned BBH the waveforms can be described using a reduced number of three intrinsic parameters, e.g., $\{q, S_{1z}, S_{2z}\}$. The extrinsic parameters correspond to simple scaling or phase factors that can be applied *a posteriori* in the waveform computation. We thus fix $r^* = 1\,\text{Mpc}$, and $(\iota^*, \varphi^*) = (0, 0)$ in the sequel, with no loss of generality. With such a setting, the GW signal $h(t) \propto a(t) e^{-\boldsymbol{i}\Phi(t)}$ appears as an amplitude and frequency modulated signal, often referred to as "chirp" for short.

A range of models allows the generation of approximated yet accurate waveforms for data analysis purposes [4]. Here, we use SEOBNRv4 [11] based on the effective-one-body formalism (EOB). Figure 2 shows an example of the expected gravitational-wave signal from a BBH merger.

The waveform amplitude and phase can be simply deduced from the polarizations using:

$$a(t) = |h(t)| \quad \text{and} \quad \Phi(t) = \arctan \frac{h_+(t)}{h_\times(t)} \quad (2)$$

Parameters $a(t)$ and $\phi(t)$ are preferred to $h_+(t)$ and $h_\times(t)$ in the generation procedure for their non-oscillatory and smooth behaviours (see Figure 2).

3. LAYOUT OF THE MACHINE LEARNING MODEL

Following [7] we now introduce a generative model able to regress waveforms given $\theta = \{m_1, m_2, S_{1z}, S_{2z}\}$ the set of features that collects the input astrophysical parameters.

3.1. Main steps

The regression targets are chirp waveforms with the specific amplitude and phase evolution shown in the previous section. We first map those complex oscillatory signals (not well suited for dimension reduction) into more regular amplitude and phase attributes that exhibit slow and smooth variations with respect to the features, that can thus be more accurately fitted. In a nutshell, the regressor is a principal component regressor [12], that essentially consists in regressing the PCA coefficients of the attributes from polynomial combining of

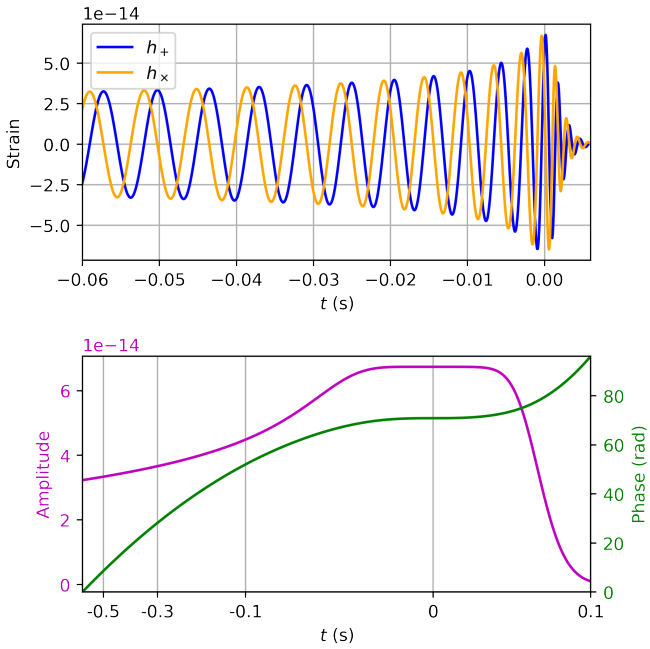

Fig. 2. Top: example of gravitational-wave waveform obtained with the SEOBNRv4 model for a spin-aligned BBH. Bottom: associated amplitude and phase waveform attributes obtained from the waveform following Section 3.1.1.

the features. The waveform generation essentially performs the reverse process: obtain estimates for the PCA coefficients from the regressor, that are inverted to compute attributes, which in turn are mapped to waveforms.

This section describes the main steps of the model in more details.

3.1.1. *Mapping to attributes*

Over its entire duration, the chirp signal $h(t)$ goes through a sequence of three phases associated to different dynamical regimes: *inspiral* when the two black holes are far apart, *merger* when they are close and "plunge" onto each other and *ringdown* when the merger remnant settles down to equilibrium. The amplitude and phase evolve over different timescales during these phases. To capture their variations with a uniform accuracy over the entire waveform duration, amplitude and phase are discretized in time with a varying sampling resolution, thus leading to the attributes represented in Figure 2.

Time grid goes from $t_{\text{start}} = -20\,\text{s}$ to $t_{\text{end}} = 0.006$ with $t = 0$ at maximum amplitude. Concretely, 4096 sample points are computed over the non-uniform time grid $\tilde{t} = \text{sign}\,(t)\,|t|^{\frac{1}{\alpha}}$ where $\alpha = 0.35$ [7]. The sampling rate is reduced in the inspiral phase where both the amplitude and phase evolutions are slow, while it is increased when getting closer to the merger where the amplitude reaches its maximum, and where the phase evolves rapidly. This resampling has two benefits: data reduction during inspiral [1], and high accuracy near the merger.

3.1.2. *Data reduction*

The training set for attributes is reduced through a principal component analysis (PCA) after an alignment to 0 at $t = t_{\text{start}}$ for PCA efficency purposes[2]. The first three principal components/vectors obtained for the amplitude and phase are displayed in Figure 3. The minimal number of principal components (PC) required to reach the maximal capabilities of the model is determined by cross-validation (see Section 4.2).

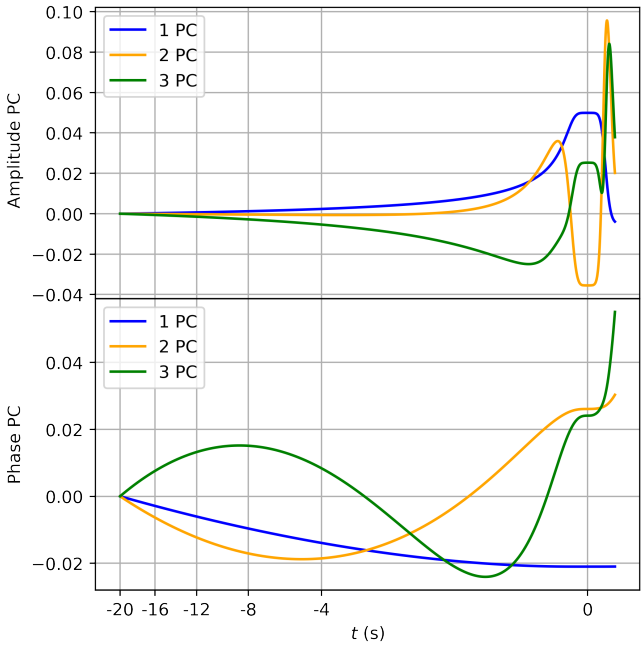

Fig. 3. First three principal components for the amplitude (upper panel) and phase (lower panel) attributes.

3.1.3. *Regression*

Schmidt et al. [7] use a mixture of experts (MoE) regressor to infer the PCA coefficients from the set of features $\{q, \chi_{1z}, \chi_{2z}\}$ where $\vec{\chi}_i = \vec{S}_i/m_i^2$ for $i = 1, 2$ are the dimensionless spins. MoE is an ensemble learning method, based on a weighted sum of linear regressors called "experts" [12].

[1] This allows to easily process waveforms sampled at 16384 Hz on a standard laptop.

[2] The amplitude and phase offsets subtracted by the alignment procedure can be fitted and added back at the generation stage to produce the full waveform. Though we don't detail this part here but this can be done with good accuracy with the same regressor.

In this work, a single linear regressor is preferred and applied to a different set of features, expanded with polynomial combinations up to a pre-determined order. This maximum order is computed to maximize a score, see Section 4.2.

Choosing the initial set of features has a significant impact on the final score and regression accuracy. To leading order, the amplitude and phase evolution are known [4] to depend on the chirp mass $\mathcal{M} = (m_1 m_2)^{3/5}/M^{1/5}$, the mass ratio q and the effective spin $\chi_{\text{eff}} = (q\chi_{1z} + \chi_{2z})/(1+q)$. It is natural to think that those physically motivated parameters are good candidates to fit the data.

Systematic tests with subsets made of different feature combinations were performed and the final model use the feature set leading to the best score (see Section 4.2.2).

3.1.4. Waveform generation

Once the model is trained, the usage goes as follows: the PC coefficients of the attributes are predicted for the parameters of the desired compact binary. The attribute time series are deduced from the predicted PC coefficients and the waveform amplitude and phase are interpolated from the attributes on the (uniform) time grid used for data analysis. The GW polarizations $h_+(t)$ and $h_\times(t)$ are finally computed by applying (1) with the requested distance, inclination and initial phase.

3.2. Implementation and scoring

The principal component regressor described in Section 3.1 relies on off-the-shelf building blocks (polynomial feature expansion, PCA, linear regression). The full pipeline is implemented using the `Scikit-learn` software library [10] resulting in a compact code, easy to maintain.

The regression accuracy is evaluated by a specific metric called mismatch or unfaithfulness [13]. The mismatch between two waveforms $h, g \in \mathbb{C}^N$ is defined as:

$$\text{mismatch}\,(h,g) = \min_{\tau\in\mathbb{R}} \left[1 - \frac{|\langle h_\tau, g\rangle|}{\|h_\tau\|\,\|g\|}\right] \qquad (3)$$

where $h_\tau(t) := h(t-\tau)$ and with the scalar product:

$$\langle f, g\rangle = \int \frac{h(f)g^*(f)}{S(f)}\,df$$

defined in the Fourier domain and with f being the frequency variable. This metric is a loss function (smaller is better) and it is phase-shift and time-shift invariant. The definition in (3) allows for a frequency-dependent weighing, usually fixed to the GW detector noise power spectrum density $S(f)$. Here, we assume a flat noise curve $S(f) = 1$, which leads to a conservative constraint as the regression should then be equally accurate at all frequencies.

Errors in the waveform approximation lead to systematic errors in the astrophysical parameters estimates obtained from the observations. Those systematic errors from mis-modeling should be smaller than the statistical errors (due to the presence of noise in the observations). This principle leads to the rule of thumb (see e.g., App. G of [13]) stating that the mismatch should be $< N/(2\,\text{SNR}^2)$ where N is the effective number of intrinsic parameters and SNR is the signal-to-noise ratio. Schmidt et al. [7] achieve a median mismatch value of 5×10^{-4}, with tails going to 10^{-1} in the worst case. In the spin aligned case with $N = 3$ effective parameters, this corresponds to an applicability range that goes up to $\text{SNR} = 55$ (3 in the worst case).

4. RESULTS

We now train and apply the regression model, tune the model hyperparameters to optimize the final score and evaluate its performances.

4.1. Training and testing datasets

A dataset of 4000 randomly distributed BBH waveforms was simulated with the SEOBNRv4 model available from the `LALSimulation` software library [14]. The mass ratio q was randomly drawn uniformly in the range $[1,\,20]$ and spins χ_{1z}, χ_{2z} uniformly over $[-0.8,\,0.95]$. Such dataset is comparable to the one used in [7].

For the subsequent analysis, the dataset is splitted into a training and a testing set that corresponds to $80\,\%$ and $20\,\%$ of the main dataset respectively.

4.2. Hyperparameter tuning

The model presented in Section 3 has a small number of hyperparameters that can be adjusted in order to optimize the mismatch score.

4.2.1. Number of PC after truncation

The overall modelling error is essentially due to the truncation of the PC expansion and the regression for each attribute. We first evaluate the former in Figure 4 (green circles). The phase approximation plays a dominant role. For instance, the truncation to only one PC leads to a mismatch of 5×10^{-5} for the amplitude and 7×10^{-1} for the phase.

Figure 4 shows that the overall error after regression (blue crosses) stabilizes at 6 PCs for the phase with a mismatch median score of about 10^{-5}. This is the retained number of PCs after truncation as this indicates the regressor fails to estimate higher order PC coefficients. For simplicity, the PCA expansion is truncated to the same number of PCs for both the amplitude and phase.

4.2.2. Feature selection

Different feature combinations taken were tested and compared through their mismatch median value and dispersion.

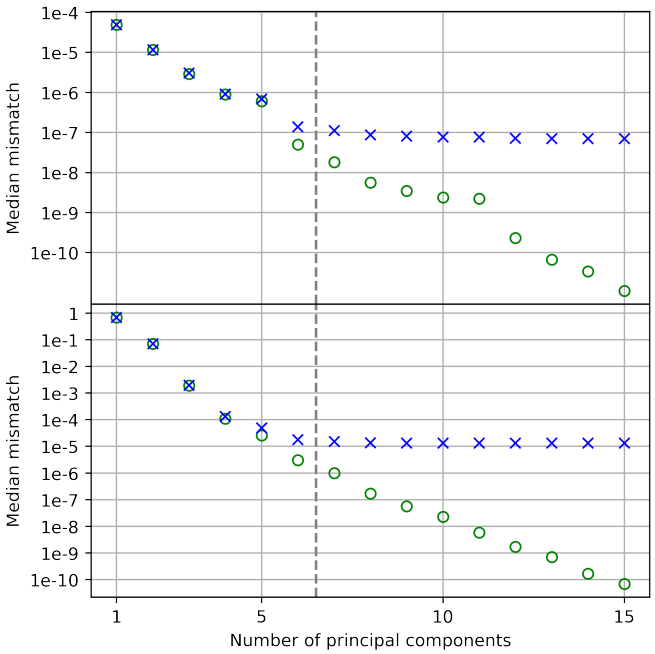

Fig. 4. Median mismatch *vs* number of principal components retain after PCA truncation. The mismatch is computed after fitting the amplitude only (phase is exact) in the upper panel, and the phase only in the lower. Circles represent mismatches solely due to the PCA truncation and crosses represent mismatches due to the overall model. The dashed line indicates the selected PC truncation level at 6 PC.

Promising sets of features were first pre-selected based on their r^2 score [10] obtained for the regression of the first PC of the phase, which essentially governs the overall performance (see Section 4.2.1). This procedure was performed for subsets that collect up to 6 parameters from the pool that includes q, $\mathcal{M}$, $\chi_{\rm eff}$ and χ_{iz}, m_i and $1/m_i$ for $i = 1, 2$.

The two following feature sets $\{q,\, m_2,\, \chi_{1z},\, \chi_{\rm eff}\}$ and $\{\chi_{2z},\, \chi_{\rm eff},\, \mathcal{M}\}$ were found to provide the best median mismatch of order 10^{-5}. The latter has the nice property to possess 3 features only, similarly to the number of intrinsic physical parameters, however the former results in a much smaller variance for the mismatch and is thus preferred. Interestingly feature sets such as $\{q,\, \chi_{1z},\, \chi_{2z}\}$ (used in [7]) or $\{q,\, \chi_{\rm eff},\, \mathcal{M}\}$ (a "natural" candidate composed of the physically motivated parameters) do not perform as well (by orders of magnitude).

4.2.3. Maximum order of polynomial features

The feature set $\{q,\, m_2,\, \chi_{1z},\, \chi_{\rm eff}\}$ is expanded by polynomial combinations of the initial features up to a maximum order [10]. The minimum order necessary to achieve best performances is determined by cross-validation. As shown in Figure 5 the mismatch is minimized with a $7^{\rm th}$ order polynomial feature expansion.

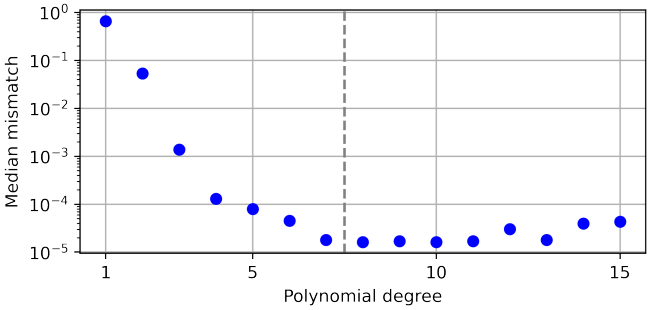

Fig. 5. Median mismatch *vs* order of the polynomial expansion of features $\{q,\, m_2,\, \chi_{1z},\, \chi_{\rm eff}\}$. The dashed line shows the selected degree of 7.

4.3. Accuracy and runtime benchmarking

Figure 6 shows the mismatch distribution obtained with the testing set using the best feature set. The median mismatch is 2×10^{-5} (average is 4×10^{-5}) and the 5% and 95% percentiles are 2×10^{-6} and 1.5×10^{-4}, resp. The worst case mismatch is $\sim 10^{-3}$. These results show an improvement of about an order of magnitude compared to [7]. This makes the approximation error much smaller than the numerical and modelling error intrinsic to the SEOBNRv4 model [11], uniformly over the considered parameter space. Following the rule given in Section 3.2, this corresponds to an applicability range that goes up to SNR ≈ 225 (18 in the worst case) which covers the loudest SNR expected during the up-coming LIGO and Virgo science runs.

We now evaluate the speed up factor defined as the ratio between the runtimes required to compute the original waveform model (using the `LALSimulation` software library [14]) to that of the proposed regression model. Figure 7 displays the speed up factor distribution for the SEOBNRv4 model[3]. To perform this measurement, a set of 500 waveforms was generated by drawing the total mass uniformly between 40 and 100 $M_\odot$ with the other properties as described in Section 4.1. The median speed up is measured to be $\sim 10^2$ (comparable to [7]). This shows that the proposed method can greatly accelerate parameter estimation with Bayesian samplers [3].

The runtime of the generation with the regression model is dominated by the interpolation step from the non-uniform time grid of attributes to the uniform time grid of the user. The speed up factor could thus be further improved by dropping the non-uniform time grid (and thus the need for interpolation). Preliminary tests indicate that the same accuracy level can be obtained without the non-uniform time grid.

[3] The initial frequency is set to $f_{\rm min} = 15$ Hz

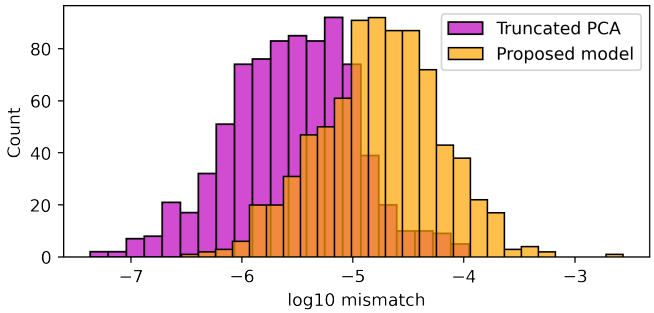

Fig. 6. Histograms of mismatch obtained with testing set. In violet the mismatch due to phase approximation with 6 principal components and in orange the mismatch of the model.

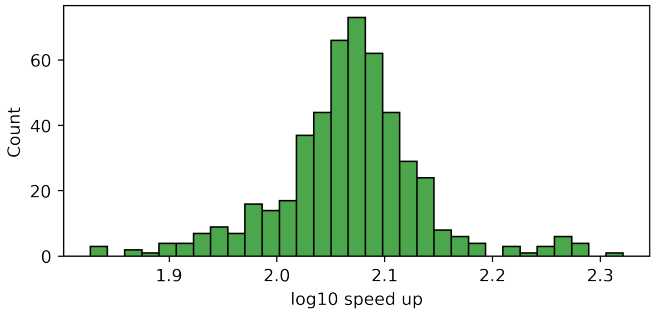

Fig. 7. Speed up factor in the waveform generation with the regression model compared to `LALSimulation` [14]. The waveforms generated for this benchmarking are associated to compact star binaries, with random total mass uniformly taken between 40 and 100 $M_\odot$.

5. PERSPECTIVES

The results presented here show that a judicious choice of features and regression technique is crucial to improve the accuracy of gravitational waveform generation techniques. This work paves the way to the rapid generation of time-domain polarized waveforms associated with precessing binaries (with arbitrary, misaligned spins). In this more complex case with more input parameters and a larger waveform variability high-accuracy modelling is anticipated to be more difficult to obtain. However, waveform generation acceleration will be important for the next science runs as current observations already indicate the presence of relativistic spin-induced precession [2]. Another straightforward but important extension is the inclusion of higher/subdominant modes other than $(\ell = 2, m = \pm 2)$ in the generative model in order to enlarge the scope of rapidly generated gravitational-wave signatures.

6. REFERENCES

[1] B. P. Abbott et al., "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, pp. 061102, 2016.

[2] R. Abbott et al., "GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the 3rd observing run," arXiv:2010.14527, 2020.

[3] E. Thrane and C. Talbot, "An introduction to bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models," *PASA*, vol. 36, pp. e010, 2019.

[4] P. Schmidt, "Gravitational waves from binary black hole mergers: Modelling and observations," *Front. Astron. Space Sci.*, vol. 7, no. 28, 2020.

[5] M. Pürrer, "Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins," *Phys. Rev. D*, vol. 93, no. 6, pp. 064041, 2016.

[6] B. D. Lackey et al., "Surrogate model for an aligned-spin effective one body waveform model of binary neutron star inspirals using Gaussian process regression," *Phys. Rev. D*, vol. 100, no. 2, pp. 024002, 2019.

[7] S. Schmidt et al., "Machine learning gravitational waves from binary black hole mergers," *Phys. Rev. D*, vol. 103, pp. 043020, 2021.

[8] S. Khan and R. Green, "Gravitational-wave surrogate models powered by artificial neural networks," *Phys. Rev. D*, vol. 103, pp. 064015, Mar 2021.

[9] A. J. K. Chua et al., "Reduced-order modeling with artificial neurons for gravitational-wave inference," *Phys. Rev. Lett.*, vol. 122, pp. 211101, May 2019.

[10] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[11] A. Bohé et al., "Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors," *Phys. Rev. D*, vol. 95, pp. 044028, 2017.

[12] T. Hastie et al., *The elements of statistical learning*, Springer, 2nd edition, 2017.

[13] K. Chatziioannou et al., "Constructing Gravitational Waves from Generic Spin-Precessing Compact Binary Inspirals," *Phys. Rev. D*, vol. 95, no. 10, pp. 104004, 2017.

[14] LIGO Scientific Collaboration, "LIGO Algorithm Library - LALSuite," 2018, DOI: 10.7935/GT1W-FZ16.