Non-Sampling Parameter Estimation for Gravitational Wave Physics

Kyungseop Yoon, Christina Reissel, and Philip Harris

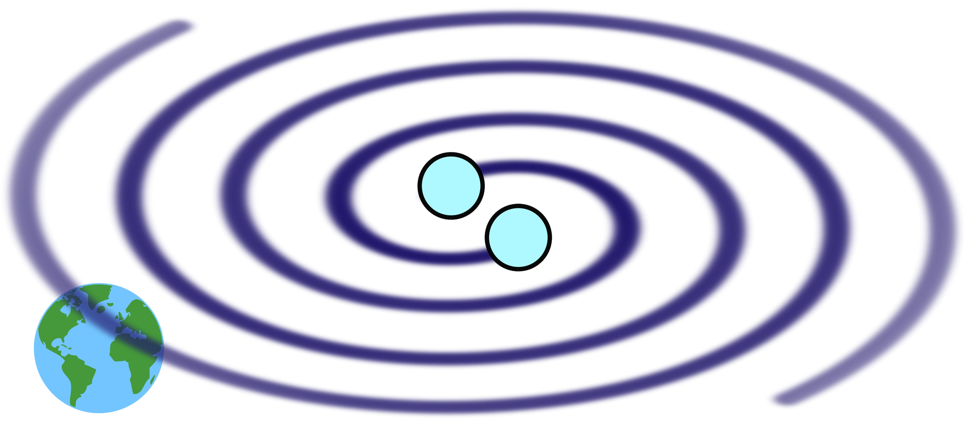
Department of Physics, Massachusetts Institute of Technology

Introduction

Gravitational waves (GW) are the first messengers of merger events, arriving before electromagnetic or neutrino signals. **Fast and accurate parameter estimation** help to prioritize observations, efficiently allocate telescope time, and capture critical early-time data. **Current approaches**, such as Markov-Chain Monte Carlo (MCMC) and Likelihood-Free Inference (LFI), **rely on sampling techniques**.

We propose an **end-to-end pipeline** based on **State Space Models (SSM)**, which

1. **bypasses the need for sampling** techniques altogether, and
2. **regresses** not only the **mean predicted values** but also their **uncertainties**



State Space Model

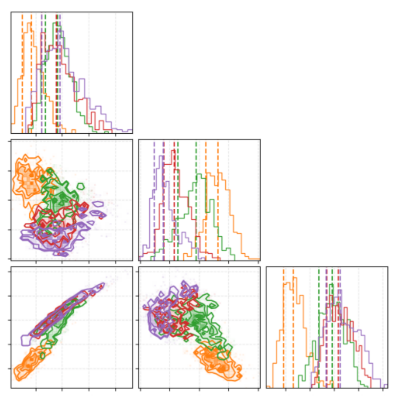
SSM maps 1-D inputs $x(t)$ to outputs $y(t)$ via N-D hidden variables through the first-order linear differential equations,

$$h'(t) = \mathbf{A}h(t) + \mathbf{B}x(t)$$
$$y(t) = \mathbf{C}h(t) + \mathbf{D}x(t).$$

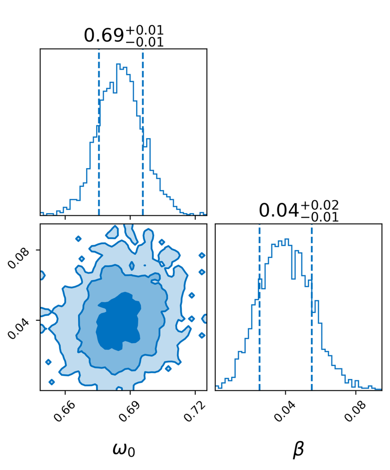
In this work, we use the **Diagonlized Structured State Space sequence model (S4D)** [1], which is built on the SSM concept to **handle long sequences efficiently**.

Benchmarks

1. Likelihood-Free Inference (LFI)
This method [2] obtains the posterior distribution from training a normalizing flow on an embedding network, which has been trained to marginalize over the time dimension.

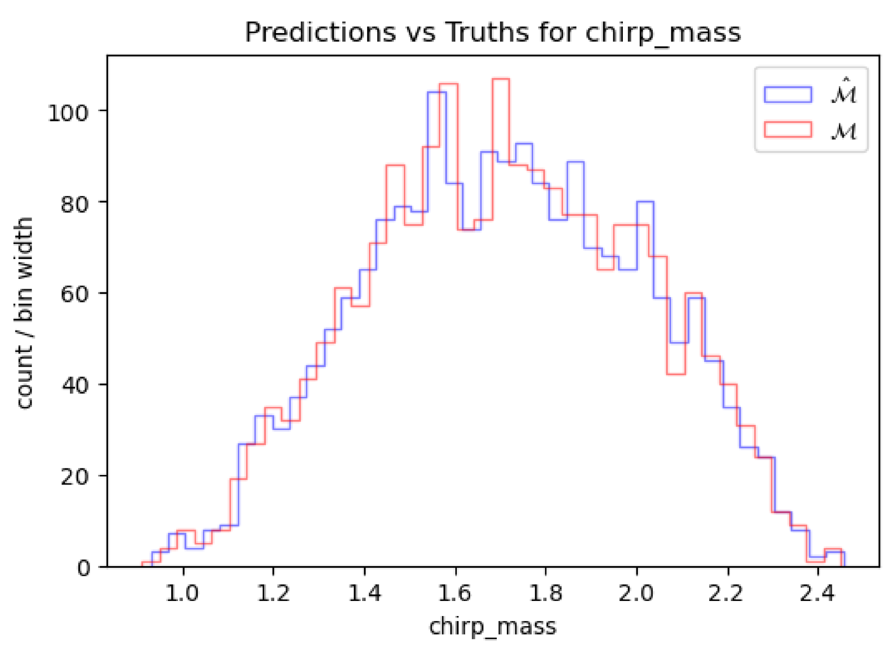


2. Bilby MCMC
Bilby is a popular software in astrophysics that provides various stochastic sampling algorithms [3]. We use the implementation of Markov-Chain Monte Carlo (MCMC) in Bilby [4].



BNS Mergers & Outlook

Binary neutron star (BNS) merger signals are notable for EM counterparts, long mergers, and weak amplitudes. We aim to regress the parameters and their uncertainties, comparing with the benchmarks. We are interested in **mass parameters** ($\mathcal{M}, M, q$, etc.) and **sky localization** (z, α, δ, etc.). Preliminary result shows chirp mass $\mathcal{M}$ regression on a signal-only dataset.

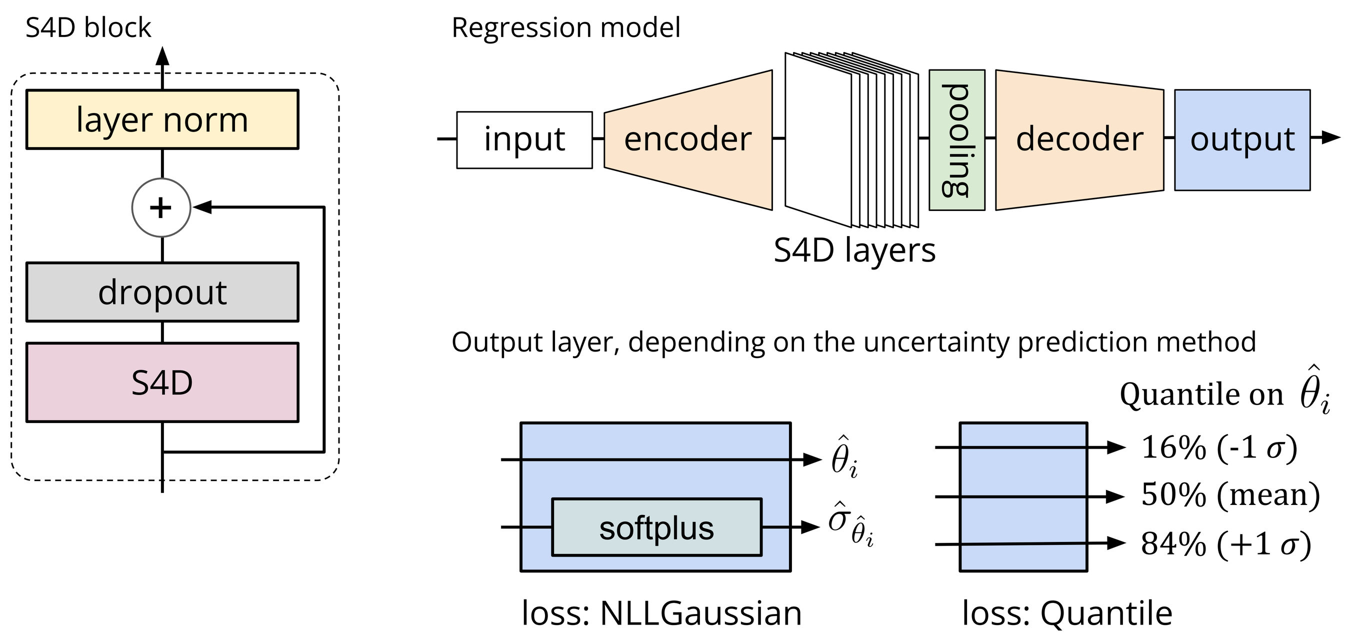


Outlook
We are currently preparing realistic realistic LIGO O3 waveforms including background noise.

Setup

Regression model
Given unknown parameters $\theta_i \in \Theta$ of our signal, our model regresses the mean predicted values $\hat{\theta}_i$ and their uncertainties $\hat{\sigma}_{\hat{\theta}_i}$.

Uncertainty prediction
For each output $\hat{\theta}_i$, we assign additional output(s) for the uncertainty $\hat{\sigma}_{\hat{\theta}_i}$. This is done by *two* choices of loss functions:



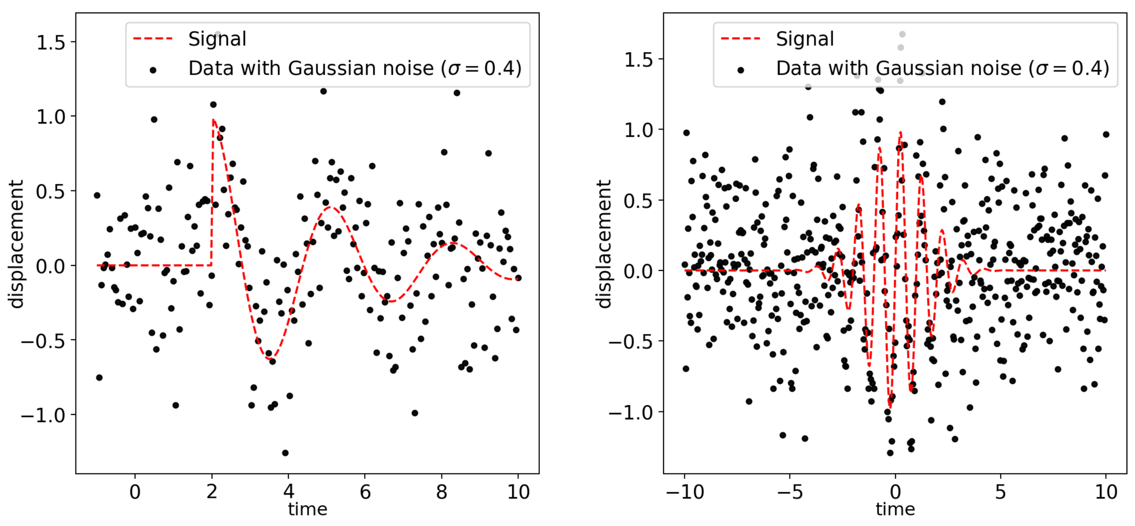
1. **NLL Gaussian loss**: $l \propto \sum_i (\hat{\theta}_i - \theta_i)^2 / \hat{\sigma}^2_{\hat{\theta}_i}$
2. **Quantile loss**: uncertainty is half the difference between the 16^{th} and 84^{th} quantile predictions.

Interpretation of result
We expect $z_i = (\hat{\theta}_i - \theta_i)/\hat{\sigma}_{\hat{\theta}_i}$ to follow a standard normal distribution, $\mathcal{N}(0, 1)$, based on the assumption that $\hat{\sigma}_{\hat{\theta}_i} \sim (\hat{\theta}_i - \theta_i)$ is Gaussian-shaped.

Toy Experiments

Follwing the setup in [2], we validate the regression model on each toy model:

- **Damped Harmonic Oscillator (SHO)**
  Θ = {natural freq ω_0, damping coeff β}
- **Sine-Gaussian (SG)**
  Θ = {central freq f_0, pulse width τ}

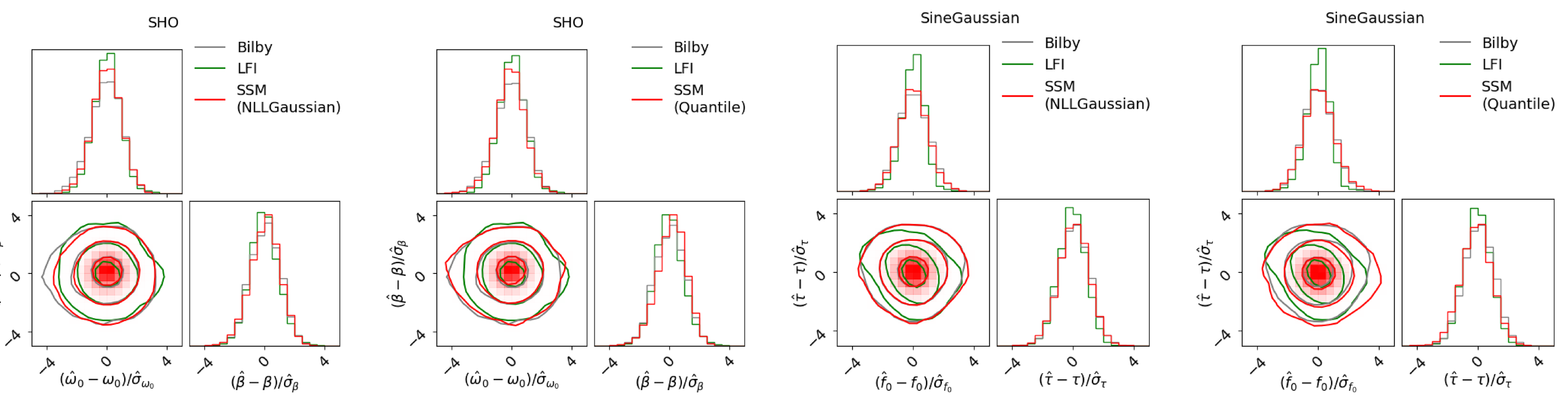


Parameters and start time t_0 are sampled from uniform priors; Gaussian noise ($\sigma = 0.4$) is added per data point. 100,000 simulations are split 80-10-10 into training, validation, and testing sets.

Comparison of the regression to the benchmarks: LFI and Bilby MCMC

While the biases are broadly comparable across methods, SSM achieves lower uncertainties than LFI. The Bilby MCMC baseline yields uncertainties on par with those from SSM.

Dataset	SHO		SG	
Method	$\hat{\omega}_0 - \omega_0$	$\hat{\beta} - \beta$	$\hat{f}_0 - f_0$	$\hat{\tau} - \tau$
SSM (NLLGaussian loss)	$0.000^{+0.039}_{-0.037}$	$0.003^{+0.044}_{-0.041}$	$0.000^{+0.014}_{-0.013}$	$0.016^{+0.249}_{-0.230}$
SSM (Quantile loss)	$-0.003^{+0.040}_{-0.045}$	$0.008^{+0.052}_{-0.038}$	$0.001^{+0.027}_{-0.026}$	$-0.016^{+0.198}_{-0.215}$
LFI	$0.002^{+0.134}_{-0.141}$	$-0.002^{+0.058}_{-0.066}$	$0.002^{+0.040}_{-0.041}$	$-0.020^{+0.279}_{-0.354}$
Bilby	$-0.001^{+0.030}_{-0.044}$	$0.002^{+0.040}_{-0.037}$	$-0.000^{+0.009}_{-0.008}$	$0.014^{+0.195}_{-0.163}$



References

[1] Gu A, Goel K and Ré C 2021 *CoRR* **abs/2111.00396** (*Preprint* 2111.00396) https://arxiv.org/abs/2111.00396
[2] Chatterjee D, Harris P C, Goel M, Desai M *et al.* 2023 Optimizing likelihood-free inference using self-supervised neural symmetry embeddings (*Preprint* 2312.07615) https://arxiv.org/abs/2312.07615
[3] Ashton G *et al.* 2019 *Astrophys. J. Suppl.* **241** 27 (*Preprint* 1811.02042)
[4] Ashton G and Talbot C 2021 *Monthly Notices of the Royal Astronomical Society* **507** 2037–2051 http://dx.doi.org/10.1093/mnras/stab2236