Referee report for *On modeling for Kerr black holes: Basis learning, QNM frequencies, and spherical-spheroidal mixing coefficients* by Lionel London and Edward Fauchon-Jones,

This manuscript summarizes two methods for fitting of high-dimensional data using non-standard basis functions. In particular, two greedy algorithm variants, GMVR and GMVP regression, are introduced. These methods are then used to build models of quasi-normal mode (QNM) frequencies of a Kerr black hole as well as spherical-spheroidal mixing coefficients. Together, these two models allow one to evaluate efficient and accurate gravitational ringdown waveforms, which is of interest to LIGO.

The work nicely introduces, summarizes, and explores the application of GMV{R,P} and provides an illustrative toy problem. The paper describes the algorithms implemented in a publicly available code, and I suspect anyone seeking to use this code will find the current manuscript to be helpful.

This paper certainly warrants publication, and I have no hesitation recommending it to be published once the following points are addressed.

My main concern is that the new physics results of this paper are not well differentiated (yet) from existing work. To summarize, it seems like the new physics result is the accurate fitting formula Eq 22 - 30 (QNMs) and the mixing coefficients from appendix A. However, at least for someone who is not an expert in modeling QNMs, the paper failed to make it clear in what ways these formula are improvements over existing formula in the literature. Again, as a non-expert I'm not sure what the current state-of-the-art is (I'm mostly aware of resources like `https://pages.jh.edu/~eberti2/ringdown/`). In my opinion, the paper should state what the current state-of-the-art fits are, and then provide a comparison between those fits and the fits provided in this manuscript. Note that there is a qualitative comparison/discussion of the improvements, but there really should be accompanying quantitative comparisons too.

Additional comments:

1. Page 1: While it became clear later on, please state after equation 1 the difference between barred and un-barred harmonic indicies.

2. Page 1: Why is the initial mass, M_i, used? Even brining up BBH systems and relating the correct ringdown parameters to the initial mass seems to be opening up an unnecessary can of worms.

3. Page 1: Missing "(" on the -2 spherical harmonic

4. Page 1: I don't understand the phrase "correlate with the morphology". Can this either be rephrased or the idea expanded upon?

5. Page 2: While perhaps obvious, please state that the pseudo-inverse isn't guaranteed to always exist.

6. Page 3: there's a lot of new notation in the algorithms that is not defined anywhere in the paper. Please define all notation

7. Page 3: Alg 1, line 8: How is this error computed? Is it relative or absolute error, L2, max, something else?

8. Page 3: Alg 1, line 14: is this really supposed to be the error between the two successive iterations? Shouldn't it be $|\epsilon_{min}| < tol$?

9. Page 3: Please provide any additional references related to algorithm's 1 and 2, since they seem somewhat similar to others that have appeared in the literature (e.g. the broad category of stepwise regression). Please provide any important theoretical properties, like convergence rates, if they are known.

10. Page 3: Is an "n-fold cartesian inner-product" the same as a tensor-product?

11. Page 4: Is degree tempering new? If no, please cite a reference. If yes, it would be worthwhile to point out a new aspect of the base methodology.

12. Page 6: Figure 2: are the errors relative or absolute? Errors should also be reported by comparing against validation data that was not used for training. I would also recommend removing the top right figure as it seems to add no value (the bottom right one contains the same information but displayed in a better way).

13. Page 6: Its not clear how the points are found? Is this a 2D tensor product grid of 25^2 points in total?

14. Page 8: Please add a reference to the statement of the "well known nonlinear oscillations". Perhaps this is well known to those working with QNMs on a regular basis, but I don't think its well known (it was previously unknown to me).

15. Throughout the paper it is stated that there are N variables and K terms. Since the summations start at 0, the equations suggest there are K+1 terms and N+1 variables.

16. While the paper is clear to follow (and was enjoyable to read), I found numerous grammatical errors (missing articles, subject-verb disagreement, etc) that should be corrected.