Rapid generation of fully relativistic EMRI waveforms for data analysis

Alvin Chua (+ Michael Katz, Niels Warburton, Scott Hughes)
JPL-Caltech

23rd Capra Meeting (virtual)
22 June 2020

Jet Propulsion Laboratory
California Institute of Technology

The EMRI problem

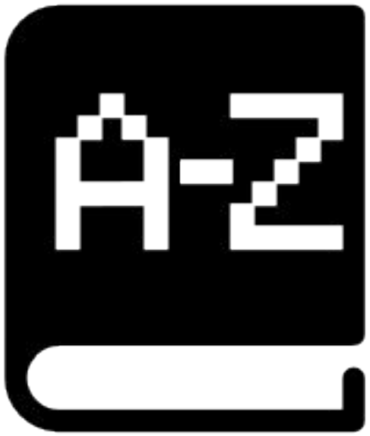
- Waveforms are the forward models of GW science: Parameters $\mapsto$ observables
- Data analysis uses waveforms to find inverse solutions: Data $\mapsto$ parameters
- Both are harder for EMRIs (+ astrophysics, but let's not go into that here)

The EMRI problem

- Waveforms are the forward models of GW science: Parameters $\mapsto$ observables
- Data analysis uses waveforms to find inverse solutions: Data $\mapsto$ parameters
- Both are harder for EMRIs (+ astrophysics, but let's not go into that here)
- Difficulty 1: Accuracy
  - EMRIs are strong-field, high-SNR sources that require accurate modeling to find & characterize
  - Phasing accurate to post-1-adiabatic order should be enough, but we are not there yet
- Difficulty 2: Efficiency
  - EMRI signals are long-lived with rich harmonic content; they are costly to model & analyze
  - Stochastic algorithms in data analysis require bulk generation of waveforms (at least billions)
- Difficulty 3: Extensiveness
  - Even the "leading-order" space of EMRI orbits is gargantuan in terms of information volume
  - Waveforms are only half the battle: detector response encodes important extrinsic effects

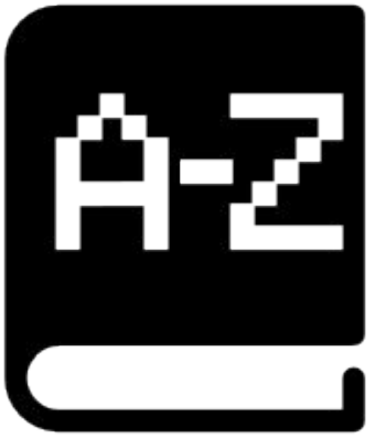
A brief definition of terms

- Waveform model:
  - Accurate (fully relativistic)
  - Not necessarily efficient & not necessarily extensive
- Template model:
  - Not necessarily accurate
  - Efficient & extensive
- Kludge models (AK, NK, AAK) are template models:
  - Not accurate (semi-relativistic)
  - Efficient & extensive



A brief definition of terms

- Waveform model:
  - Accurate (fully relativistic)
  - Not necessarily efficient & not necessarily extensive
- Template model:
  - Not necessarily accurate
  - Efficient & extensive
- Kludge models (AK, NK, AAK) are template models:
  - Not accurate (semi-relativistic)
  - Efficient & extensive
- The model we are introducing is accurate & efficient, but not yet extensive
- What should we call accurate, efficient & extensive models?
  - I proposed taking back the term “surrogate” last year, but it is a bit too loaded now
  - More on this later



A waveform model for LISA data analysis

- Standard modular description
  - Angular & frequency-based decomposition
  - Osculating geodesics
- Generic Kerr orbits
  - Need schemes to evolve through resonances
  - Need secondary spin, mass/spin evolution, etc.
- Angular dependence
  - Spheroidal harmonics with spin weight -2
- Inspiral trajectory (+ mode phasing)
  - Post-1-adiabatic order
- Mode amplitudes
  - Adiabatic order

$$h(t) = \frac{1}{r} \sum_{lmkn} A_{lmkn}(t) e^{-i\Phi_{mkn}(t)} V_{lmkn}(\theta, \phi)$$

$$G(t) \equiv (p(t), e(t), \iota(t))$$

$$V_{lmkn}(\theta, \phi) = {}_{-2}S_{lmkn}(\theta) e^{im\phi}$$

$$\Phi_{mkn}(t) = \text{init.} + \int_{t_0}^{t} dt' \, \omega_{mkn}(G(t')) + \text{osc.}$$

$$A_{lmkn}(t) = -\frac{2Z^{\infty}_{lmkn}(G(t))}{\omega^2_{mkn}(G(t))}$$

A waveform model for LISA data analysis (so far)

- Standard modular description
  - Angular & frequency-based decomposition
  - Osculating geodesics
- Eccentric Schwarzschild orbits
  - Neglect resonances
  - Neglect secondary spin, mass/spin evolution, etc.
- Angular dependence
  - Spherical harmonics with spin weight -2
- Inspiral trajectory (+ mode phasing)
  - Adiabatic order
- Mode amplitudes
  - Adiabatic order

$$h(t) = \frac{1}{r}\sum_{lmn} A_{lmn}(t) e^{-i\Phi_{mn}(t)} V_{lm}(\theta,\phi)$$

$$G(t) \equiv (p(t), e(t))$$

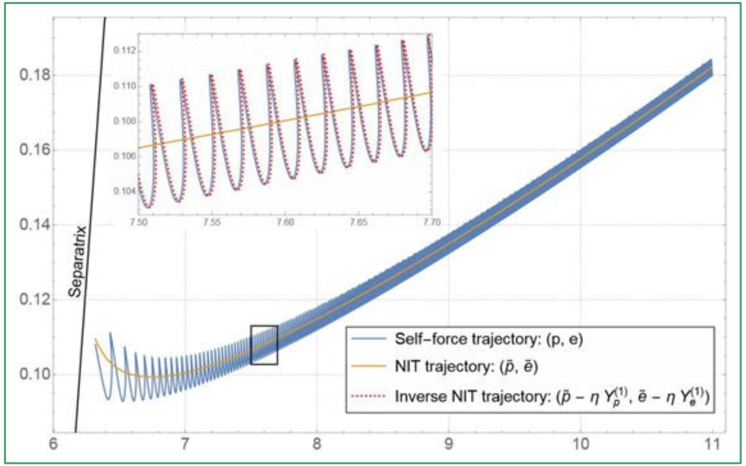
$$V_{lm}(\theta,\phi) = {}_{-2}Y_{lm}(\theta) e^{im\phi}$$

$$\Phi_{mn}(t) = \text{init.} + \int_{t_0}^{t} dt'\, \omega_{mn}(G(t'))$$

$$A_{lmn}(t) = -\frac{2Z^{\infty}_{lmn}(G(t))}{\omega^2_{mn}(G(t))}$$

Inspiral trajectory

- Options for trajectory model
  - PN flux-based (less likely for actual analysis)
  - Teukolsky flux-based (OK for detection)
  - Two-timescale framework
  - NIT (van de Meent & Warburton, 2018)



van de Meent & Warburton (2018)

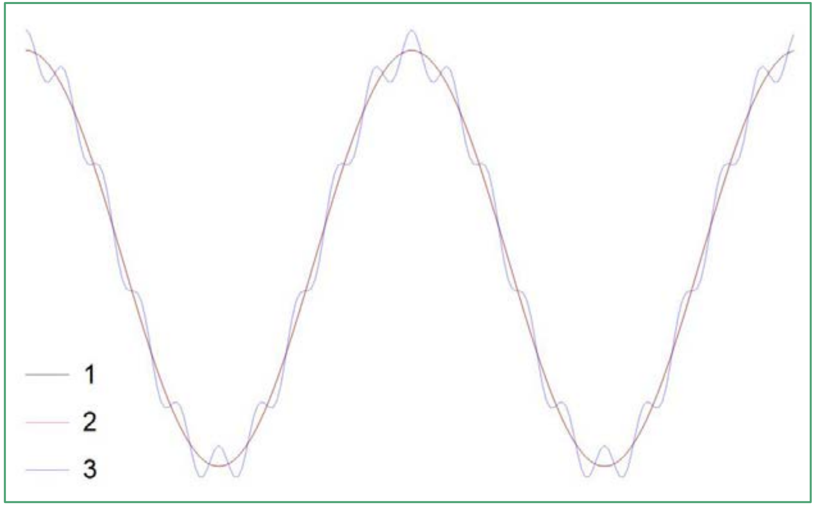
Inspiral trajectory

- Options for trajectory model
  - PN flux-based (less likely for actual analysis)
  - Teukolsky flux-based (OK for detection)
  - Two-timescale framework
  - NIT (van de Meent & Warburton, 2018)
- Current model: Teukolsky flux-based
  - Numerical data: 10^{-12} fractional error on amplitudes, 1640 points in geodesic space
  - 8th-order Runge-Kutta method
  - Cubic-spline interpolation

$$(p_0, e_0, \eta) \mapsto (G(t), \Phi_{mn}(t))$$

Mode amplitudes

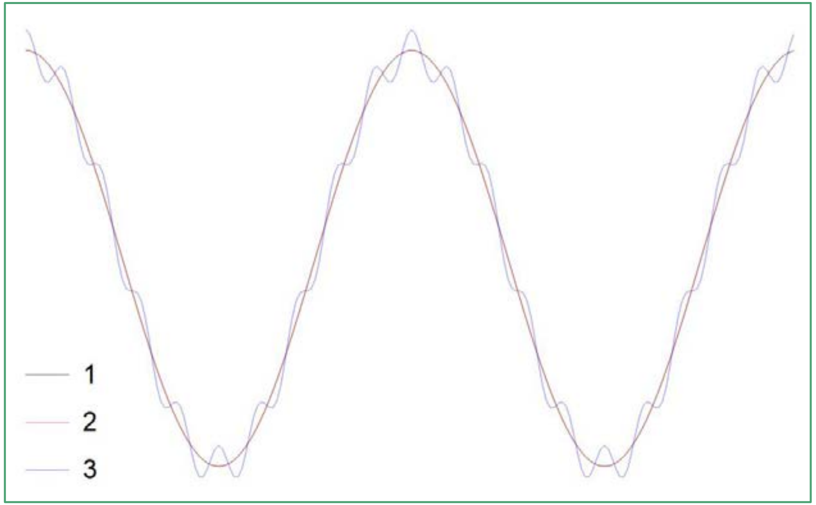
- Options for amplitude model
  - PN amplitudes (partial coverage)
  - Interpolate Teukolsky amplitudes (local)
  - Fit Teukolsky amplitudes (global)
- Are adiabatic amplitudes OK?



$$1 : \cos(\omega t),\ T = 10^3\,\text{cycles}$$
$$2 : \cos((1 + 10^{-4})\omega t)$$
$$3 : \cos(\omega t) + 10^{-1}\cos(100\omega t)$$

Mode amplitudes

- Options for amplitude model
  - PN amplitudes (partial coverage)
  - Interpolate Teukolsky amplitudes (local)
  - Fit Teukolsky amplitudes (global)
- Are adiabatic amplitudes OK?
  - Yes, probably



$1 : \cos(\omega t),\ T = 10^3$ cycles

$2 : \cos((1 + 10^{-4})\omega t)$ $\quad$ $\mathrm{over}(1, 2) = 0.936$

$3 : \cos(\omega t) + 10^{-1}\cos(100\omega t)$ $\quad$ $\mathrm{over}(1, 3) = 0.995$

Mode amplitudes

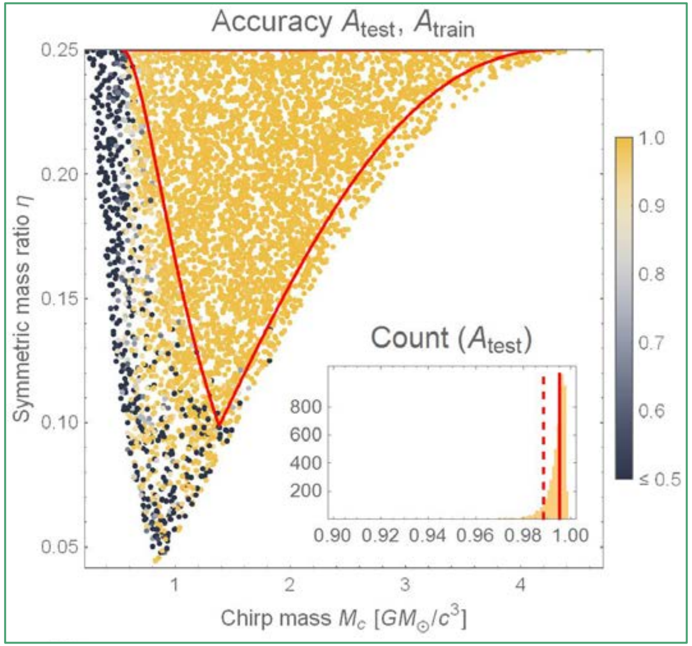
- Options for amplitude model
  - PN amplitudes (partial coverage)
  - Interpolate Teukolsky amplitudes (local)
  - Fit Teukolsky amplitudes (global)
- Are adiabatic amplitudes OK?
  - Yes, probably
- Interpolation/fitting is difficult though
  - High-dimensional mode space
  - Kerr: > 2-dimensional geodesic space
  - Many evaluations along trajectory

$$2 \leq l \leq 10,\ m \leq l,\ |n| \leq 30$$

$$(p, e) \mapsto \mathrm{vec}(A_{lmn}) \in \mathbb{C}^{3843} \cong \mathbb{R}^{7686}$$

Mode amplitudes

- Current model: Neural-network fit
  - ROMAN (Chua, Galley & Vallisneri, 2019)
  - Combination of ROM & deep learning
  - Alternative to surrogate + ROQ framework



Chua, Galley & Vallisneri (2019)

Mode amplitudes

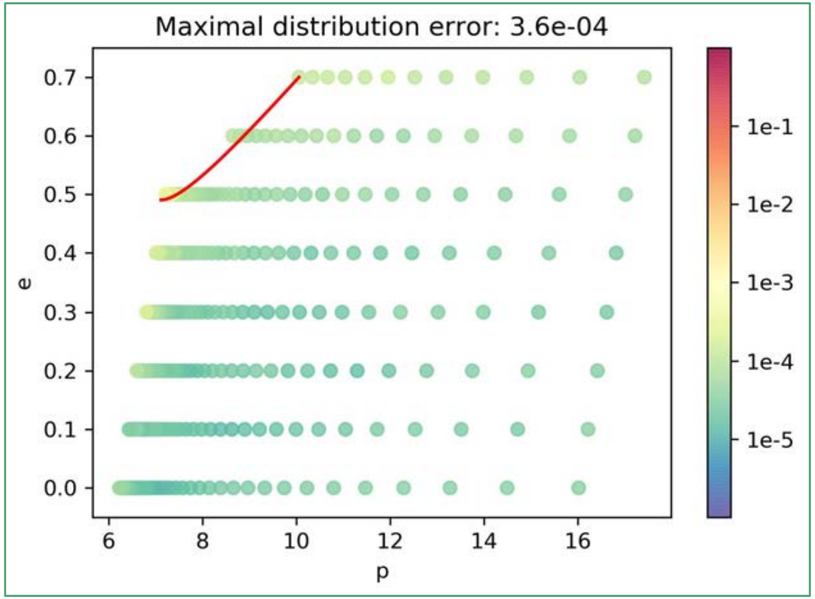
- Current model: Neural-network fit
  - ROMAN (Chua, Galley & Vallisneri, 2019)
  - Combination of ROM & deep learning
  - Alternative to surrogate + ROQ framework
- Construct reduced basis for mode set
  - Numerical data: Same as trajectory model
  - Order reduction: 7686 to 198

$$\mathrm{vec}(A_{lmn})(p,e) = \sum_i \alpha_i(p,e)\mathbf{e}_i \equiv \alpha(p,e)$$

$$(p,e) \mapsto \alpha \in \mathbb{C}^{99} \cong \mathbb{R}^{198}$$

Mode amplitudes

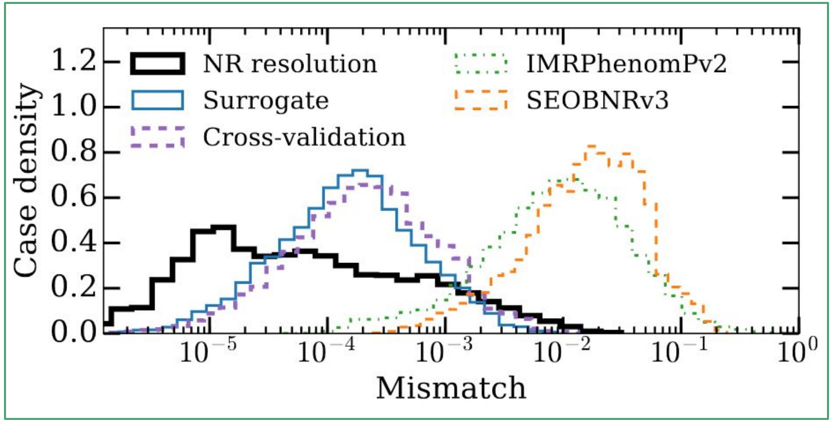
- Current model: Neural-network fit
  - ROMAN (Chua, Galley & Vallisneri, 2019)
  - Combination of ROM & deep learning
  - Alternative to surrogate + ROQ framework
- Construct reduced basis for mode set
  - Numerical data: Same as trajectory model
  - Order reduction: 7686 to 198
- Train neural network on reduced map
  - Network: Multilayer perceptron, 20 layers
  - Test domain: Initial eccentricities up to 0.7, plunge eccentricities up to 0.5, separations from LSO + 0.2M to LSO + 10M
  - Distribution error: 1 - cos(angle)



$$\text{error} = 1 - \frac{\Re(\alpha^\dagger \alpha_{\text{num}})}{\sqrt{(\alpha^\dagger \alpha)(\alpha^\dagger_{\text{num}} \alpha_{\text{num}})}}$$

Mode amplitudes

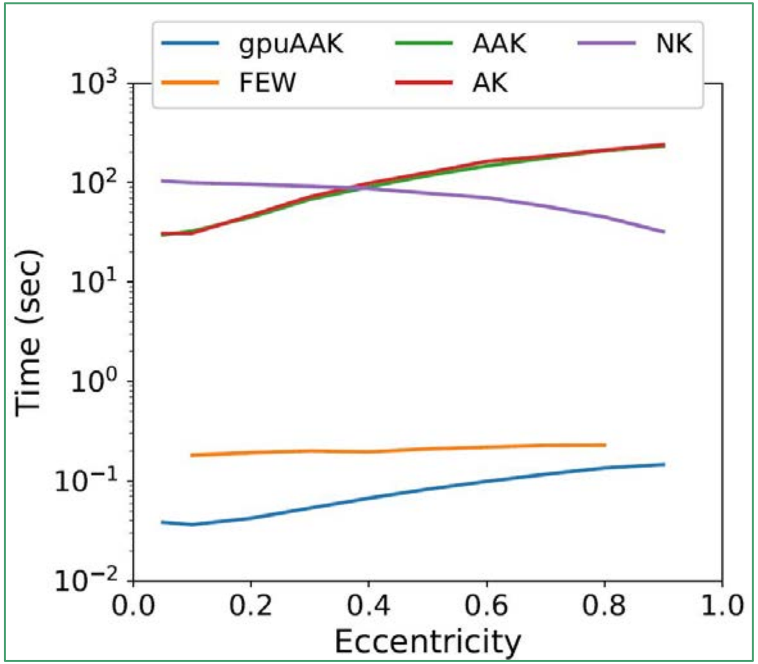
- Why not apply ROM directly to waveforms?
  - Circular Schwarzschild IMRI: 1 parameter, < 200 cycles, 22 modes (Rifat et al. 2020)
  - A usable EMRI surrogate is unlikely to cover more than a miniscule patch of parameter space
  - Accuracy is also an issue: Even best NR surrogates have maximal mismatches > 10^{-3}



Blackman et al. (2017)

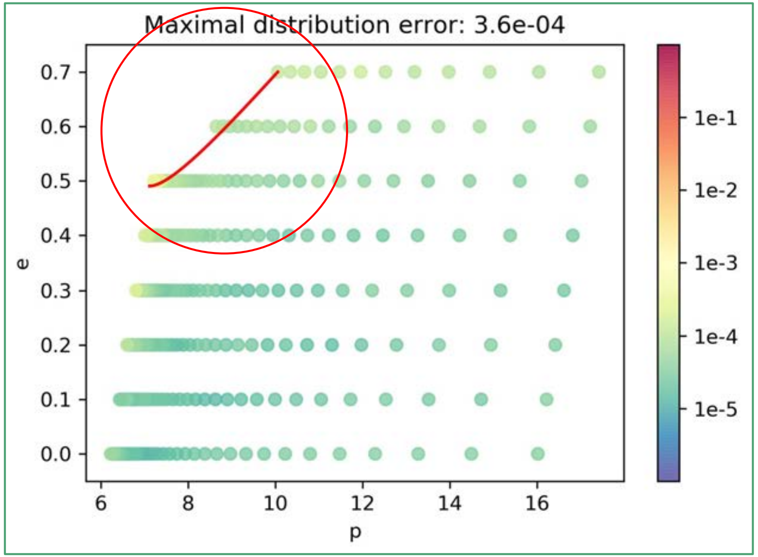
GPU implementation

- GPU acceleration in GW modeling
  - MBH waveforms (Katz et al. 2020)
  - Kludge waveforms (EMRI Kludge Suite)
- Neural networks are highly parallelizable
- Speed: Milliseconds to < 1 second
  - Analysis-length waveforms (1 year at 0.1 Hz)
  - Full harmonic content (all relevant modes)



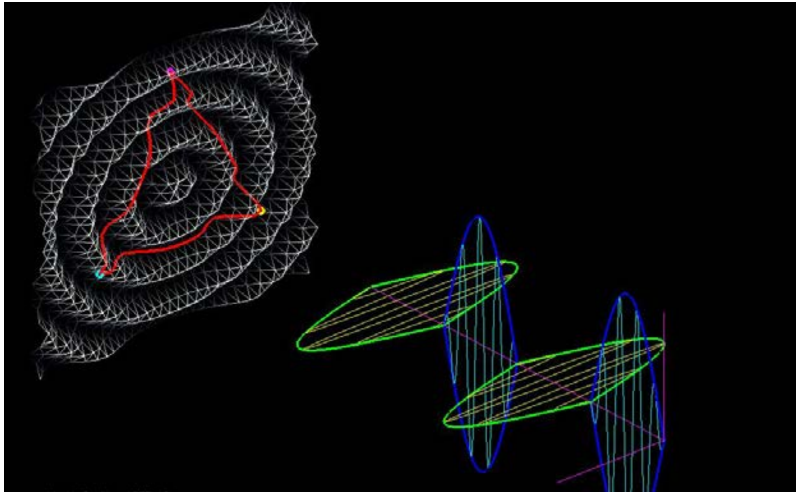
GPU implementation

- GPU acceleration in GW modeling
  - MBH waveforms (Katz et al. 2020)
  - Kludge waveforms (EMRI Kludge Suite)
- Neural networks are highly parallelizable
- Speed: Milliseconds to < 1 second
  - Analysis-length waveforms (1 year at 0.1 Hz)
  - Full harmonic content (all relevant modes)
- Accuracy: Maximal mismatch $< 4 \times 10^{-4}$
  - Compared to slow fiducial waveform with dense interpolation (about 1 hour on 10 cores)
  - Enough to estimate parameters without bias for most signals (up to around SNR 100)



Strategies for future work

- Improve extensiveness (source-side)
  - Eccentric equatorial Kerr: Add spin
  - Partial post-adiabatic trajectory: 1st-order SF
  - Incorporate resonance schemes
- Other representations
  - Frequency domain: Higher-order SPA
  - Time-frequency domain: STFT
- Improve extensiveness (detector-side)
  - Integrate with LISA response models



N. Douillet

Summary & references

- EMRI waveforms for LISA data analysis must be accurate, efficient & extensive
- We introduce the first fast & fully relativistic model for eccentric Schwarzschild
- Our framework is designed to scale well to the full EMRI analysis problem

---

- A. J. K. Chua, M. L. Katz, N. Warburton & S. A. Hughes, Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for LISA data analysis, in prep.
- M. L. Katz, S. Marsat, A. J. K. Chua, S. Babak & S. L. Larson, GPU-accelerated massive black hole binary parameter estimation with LISA, in rev., arXiv:2005.01827
- A. J. K. Chua, C. R. Galley & M. Vallisneri, Reduced-order modeling with artificial neurons for gravitational-wave inference, Phys. Rev. Lett. 122, 211101 (2019).
- M. van de Meent & N. Warburton, Fast self-forced inspirals, Class. Quantum. Grav. 35, 144003 (2018).