Astrophysics Challenge Problems

Gianni Mossi (KBR, Inc.)

Erik Gustafson (KBR, Inc.)

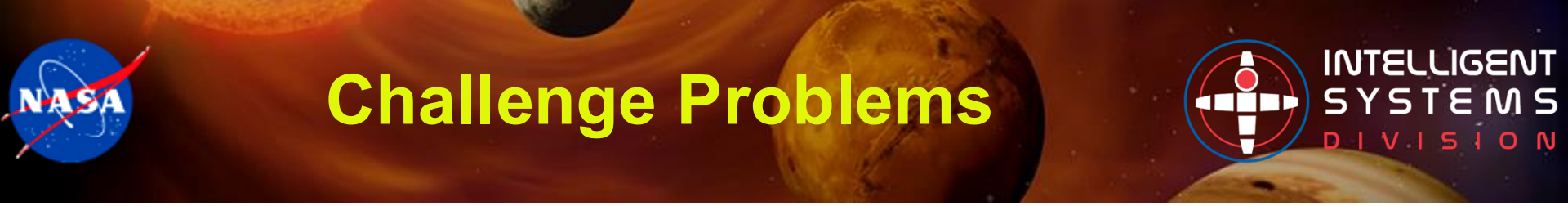
Super Intelligence (SI) Usage Disclosure

This document was created with assistance from NASA's Office 365 Copilot for background research, problem specification and identification. The content has been reviewed and edited by all authors. More information on the extent and nature of SI usage is included in the Acknowledgements sections of the document





INTELLIGENT SYSTEMS DIVISION



Challenge Problems

1. Numerical Relativity Simulations of Binary Black hole mergers in computationally difficult regimes

2. Global Fit: Laser Interferometer Space Antenna

3. N-body cosmology



Binary Black Hole Merger Simulations

Binary Black Hole Mergers

INTELLIGENT SYSTEMS DIVISION

Application Details

Goal: construct gravitational waveforms for most difficult parameter regime of binary black hole (BBH) merges

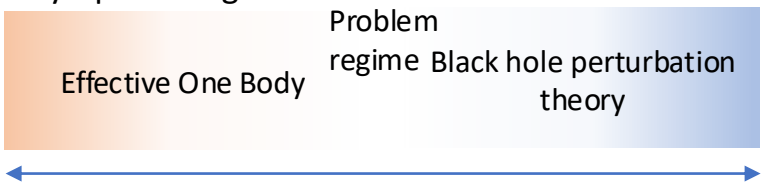
- Perform theory calculations that are of interest and support LISA experiments
- Simulations of black hole mergers for mass ratios of $1{:}10^3$ to $1{:}10^6$ high eccentricity, and off-axis spins are difficult

Problem Details

- Numerically solving Einsteins equations
- Solving complicated partial differential equations of fields
- Inputs: Mass ratio, spins, gridding methods,
- Output: noiseless Gravitational waveforms for merger that can be used to disambiguate signals.

Current Approaches

- Einstein Toolkit or SpEC for numerical relativity
- For asymptotic limits use effective theories:
  - Black hole perturbation theory (large mass ratio)
  - Effective one body theories (equal mass ratio)
- Numerical relativity can solve all regimes but is computationally expensive; effective theories are good for asymptotic regime but bad for intermediate



Impact of New/Better Solution

- Want ~5% or greater accuracy in waveforms and/or order of magnitude improvement in computational speed.
- Improved accuracy will improve analysis of LISA data; improved speed enables reduced costs and greater number of mergers available.
- Improvements to simulating PDE's will enable improvement across a broad scope of other fields including fluid dynamics.

Application Details

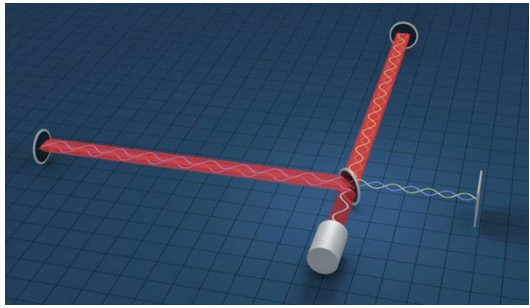
Goal: construct gravitational waveforms for most difficult parameter regime of binary black hole merges (**and possibly other processes of interest**)

- Perform theory calculations that are of interest and support LISA experiments
- Simulations of black hole mergers for mass ratios of $1{:}10^3$ to $1{:}10^6$ high eccentricity, and off-axis spins are difficult
- **Background:** current Earth-based observatories (LIGO, Virgo, etc) are limited by seismic noise that prevents them from studying Gravitational Wave (GW) signals in the low-frequency band f ≤10 Hz. Several astronomical events/processes produce GW signals supported in this band, for example:
  - The early stage of merges that LIGO will eventually detect. This gives LISA an advantage of years over Earth-based observations (*cf.* Multi-Messenger astronomy).
  - Supermassive black-hole binaries (10^{-4}-10^{-1} Hz signals during inspirals)
  - Extreme-Mass-Ratio Inspirals (never reach high-frequency band before plunge)
  - Galactic Binary White Dwarfs (~ 10^{-3} Hz)
  - Cosmic Strings? Phase transitions? Stochastic backgrounds?

LISA is expected to be able to detect signals in the 10^{-4} Hz ≤f ≤ 10^{-1} Hz range.

Application Details

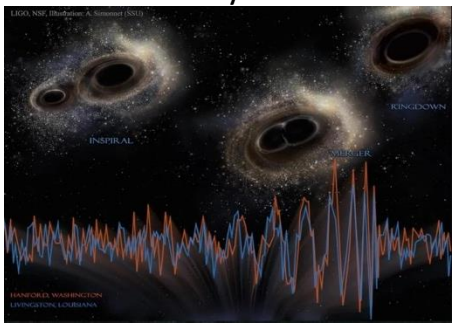
Gravitational waves (GW) are distortions in the spacetime curvature generated by accelerating massive objects. GWs can be measured using a Michelson interferometer by observing a change in the interference pattern produced by the relative stretching and squeezing (gravitational strain) of its two arms.



Michelson interferometer setup. T. Pyle Courtesy Caltech/MIT/LIGO Laboratory.

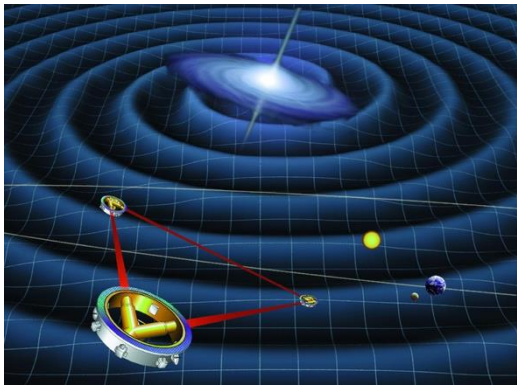
The **strain signal** (timeseries) carries information about the process that originally generated the gravitational waves. Thus, GWs are a new probing instrument that we have at our disposal for studying the universe. **We need to know what (ideal) strain signals look like in order to be able to identify them in the empirically-observed, noisy data recorded by the observatory.**

The process where two black holes merge is one of the most common source of GW that we can currently detect.



Above: the three stages of the merger.
Below: the strain timeseries
Image Credit: LIGO , NSF , Aurore Simonnet (Sonoma State U.)





Left: Virgo interferometer. Image: The Virgo collaboration/CCO 1.0 Right: NASA concept of the LISA satellite observatory (three arms) (NASA illustration of LISA)

Current Approaches

Simulations of BBH mergers can be solved through a variety of techniques.

Equal Mass Ratio

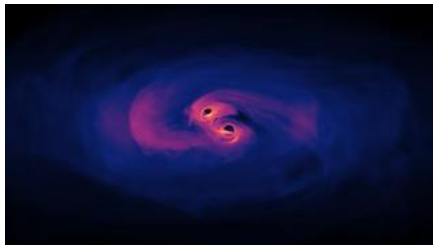
- **Effective one body** (EOB) techniques allow for analytic descriptions of equal mass ratio binaries through the entire lifecycle.
- Some surrogate techniques for numerical relativity are trained on analytic methods such as **Black hole perturbation theory**.

Intermediate Mass Ratio

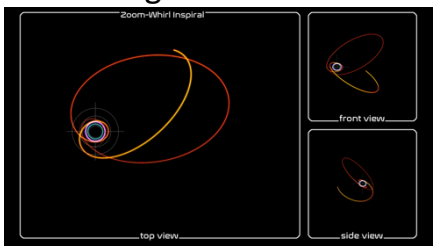
- Typically solved using **numerical relativity.**
- **Numerical relativity** uses software packages such as Einstein ToolKit or internal packages such as at CalTech.
- This is the computationally difficult regime we want to solve.

Extreme Mass Ratio

- **Black Hole perturbation** theory is a numerical / analytic technique that treats the small black hole as a point particle. See https://arxiv.org/pdf/gr-qc/9609012
- Can also be simulated slightly more cheaply using **numerical relativity** as only one ”blackhole” is modelled
- Some work has been done to extend range by using ML methods or surrogates.



NASA's Goddard Space Flight Center/Scott Noble; simulation data, d'Ascoli et al. 2018



NASA Goddard Space Flight Center

Numerical Relativity

Problem area

Equal Mass Ratio

Extreme Mass Ratio

State of the Art Methods



Black hole perturbation theory

This method treats the one of the blackholes as stationary Schwarzschild object of mass M and the other black hole as a moving point particle in the metric of the larger object. This reduces computational complexity drastically [Lou1998].

Software: https://bhptoolkit.org

Post-Newtonian Methods

These methods utilize a perturbative expansion in the reduced velocity $\beta = v/c$ and using a power series in the metric tensor to allow for analytic solutions of general relativity in a "weakly" general relativistic limit.

Numerical Relativity

The standard numerical approach for solving the Einstein equations for black hole binary mergers is the Baumgarte-Shapiro-Shibata-Nakamura (BSSN) reformulation of the equations [Baum1998, Shi1995]. This method forms the basis for many modern-day stable solutions to the Einstein equations for studying the merger of black hole binaries. And solve the equations numerically. These use widely available software toolkits such as:

- Einstein Toolkit
  - https://einsteintoolkit.org
- Others:

Many methods use a moving puncture method which "pierces" the metric off a grid point on the surface and moves that pierce around in a space-time snapshot.

Other methods include pseudo-spectral ones developed by Caltech/Cornell but are closed source.

Alternative Computing Methods

There has been rapid development of surrogate methods that attempt to stitch together numerical relativity results with the analytic / perturbative techniques to help better generate waveforms in the less well understood intermediate mass ratio regime. These methods effectively work like a machine learning algorithm.

It is possible that quantum computing can also be leveraged for solving the partial differential equations, but so far this has not been attempted.

Currently Available Data

Numerical Relativity

Catalog	Link	$q = \frac{m_1}{m_2}$ range	Sim. Method	L,m waveforms	Instances	Catalog Gen time
RIT	Link	q < 128	Finite Diff.	<4,4	1881	-
MAYA	Link	q < 15	Finite Diff.	<4,4	1087	-
BAM	Link	q < 8	Finite Diff.	<4,4	80	155M Core Hours
SXS	Link	q < 20	Spectral	<4,4	3756	480M Core Hours
GR-Athena	Link	q < 4	Finite Diff	<4,4	4(?)	26M Core Hours

Simulations can take O(10^5) CPU hours depending on optimizations to grid specifications for finite difference method.
Simulations with SpEC can often run along the O(10^5) Core hours per simulation

Acronym definitions
RIT (Rensselaer Institute of Technology), MAYA (), BAM (bifunctional adaptive mesh), SXS (Simulating eXtreme Spacetimes), GR-Athena (General Relativity-Athena),

Problem Details

- Specification: solving less well understood BBH mergers. Simulation of PDEs corresponding to numerical relativity.
- Core enabling computation: Solving PDEs
- Inputs:
  - Problem: starting conditions, internal spin for black holes, velocities, mass ratios,
  - Solver: gridding specifications, lattice spacings, simulation method
- Outputs: noiseless gravitational waveforms from mergers for binary black hole systems.

Solve or approximate partial differential equations:

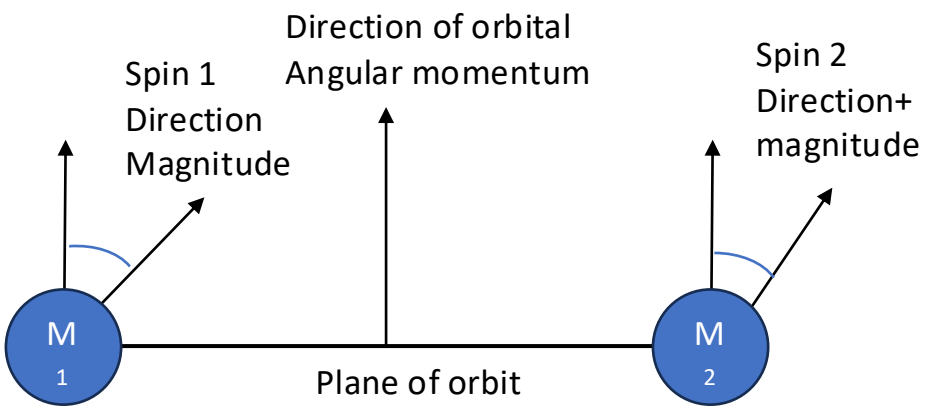
$$\frac{d}{dt}\tilde{\gamma}_{ij} = -2\alpha\,\tilde{A}_{ij} + \tilde{\gamma}_{il}\frac{\partial\beta^l}{\partial x^j} + \tilde{\gamma}_{jl}\frac{\partial\beta^l}{\partial x^i} - \frac{2}{3}\tilde{\gamma}_{ij}\frac{\partial\beta^l}{\partial x^l}$$

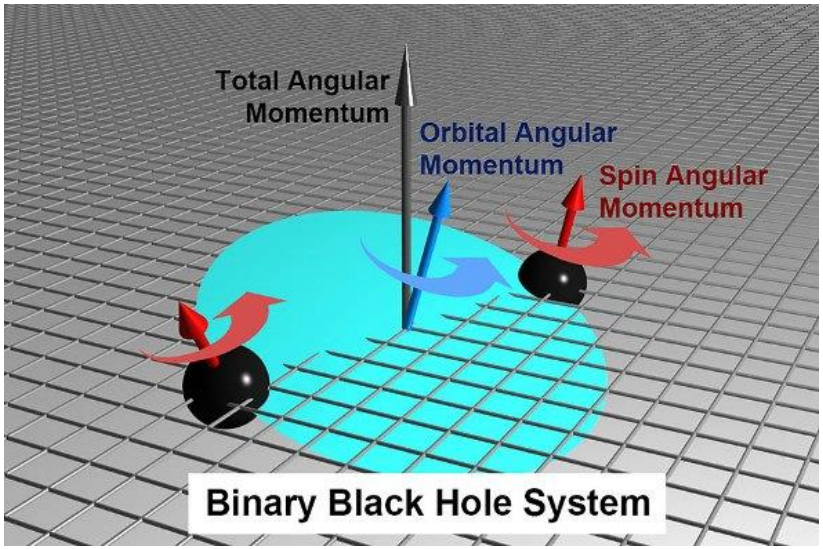
$$\frac{d}{dt}\tilde{A}_{ij} = e^{-4\phi}\left[\alpha\left(R_{ij} - \frac{1}{3}\gamma_{ij}R\right) - \left(D_iD_j\alpha - \frac{1}{3}\gamma_{ij}D^kD_k\alpha\right)\right] + \alpha\left(K\,\tilde{A}_{ij} - 2\,\tilde{A}_{il}\,\tilde{A}^l_j\right) + \frac{\partial\beta^m}{\partial x^i}\,\tilde{A}_{mj} + \frac{\partial\beta^m}{\partial x^j}\,\tilde{A}_{mi} + \frac{2}{3}\frac{\partial\beta^m}{\partial x^m}\,\tilde{A}_{ij}$$

$$\frac{d}{dt}\phi = \frac{1}{6}\left(-\alpha K + \frac{\partial\beta^i}{\partial x^i}\right)$$

$$\frac{d}{dt}K = \alpha\left(\tilde{A}_{ij}\,\tilde{A}^{ij} + \frac{1}{3}K^2\right) - D^iD_i\alpha + \alpha(\rho + S)$$

$$\frac{d}{dt} = \frac{\partial}{\partial t} - \beta^i\frac{\partial}{\partial x^i}$$





This Photo is by Midori Kitigawa and is licensed under CC BY

Problem Details

Solve or approximate partial differential equations:

$$\frac{d}{dt}\tilde{\gamma}_{ij} = -2\alpha\,\tilde{A}_{ij} + \tilde{\gamma}_{il}\frac{\partial\beta^l}{\partial x^j} + \tilde{\gamma}_{jl}\frac{\partial\beta^l}{\partial x^i} - \frac{2}{3}\tilde{\gamma}_{ij}\frac{\partial\beta^l}{\partial x^l}$$

$$\frac{d}{dt}\tilde{A}_{ij} = e^{-4\phi}\left[\alpha\left(R_{ij} - \frac{1}{3}\gamma_{ij}R\right) - \left(D_iD_j\alpha - \frac{1}{3}\gamma_{ij}D^kD_k\alpha\right)\right] + \alpha\left(K\,\tilde{A}_{ij} - 2\,\tilde{A}_{il}\,\tilde{A}^l_j\right) + \frac{\partial\beta^m}{\partial x^i}\,\tilde{A}_{mj} + \frac{\partial\beta^m}{\partial x^j}\,\tilde{A}_{mi} + \frac{2}{3}\frac{\partial\beta^m}{\partial x^m}\,\tilde{A}_{ij}$$

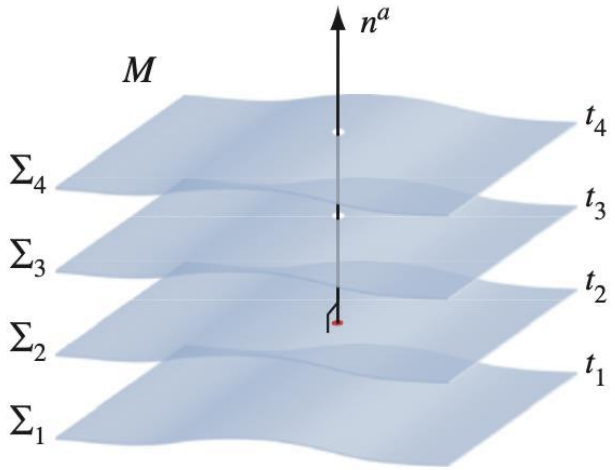
$$\frac{d}{dt}\phi = \frac{1}{6}\left(-\alpha K + \frac{\partial\beta^i}{\partial x^i}\right)$$

$$\frac{d}{dt}K = \alpha\left(\tilde{A}_{ij}\,\tilde{A}^{ij} + \frac{1}{3}K^2\right) - D^iD_i\alpha + \alpha(\rho + S)$$

$$\frac{d}{dt} = \frac{\partial}{\partial t} - \beta^i\frac{\partial}{\partial x^i}$$

- Specification
  - γ_{ij} is the special three metric with $\tilde{\gamma}_{ij} = e^{-4\phi}\gamma_{ij}$
  - $\phi = \frac{\ln(\det(\gamma_{ij}))}{12}$
  - D_i is the covariant derivative
  - β^l is the shift vector
  - α is the lapse function
  - R_{ij} is the three dimensional Ricci tensor and $R = \gamma^{ij}R_{ij}$
  - K_{ij} is the extrinsic curvature with $K = \gamma^{ij}K_{ij}$
  - tilde factors indicate a multiplication by $e^{-4\phi}$
  - x_i is the spacial components
  - See M. Shibata and T. Nakamura, Phys. Rev. D 52, 5428 ~1995 for more details.

Mathematical Description



This image is by Faisal Alotaibi from fig 1 from their masters thesis used unchanged under cc-by-4.0

The manifold can be reconstructed completely by starting with the **local data on a given "initial time slice"** and then **propagating it to different neighboring time slices**

Given a space time $(\mathcal{M}, g_{\mu\nu})$, define a 3+1 foliation in spacelike hypersurfaces (3d time slices)

$$M = \bigcup_i \Sigma_i$$

The metric $g_{\mu\nu}$ induces a

- Spatial metric γ_{ij}
- Extrinsic curvature K_{ij}

On each time slice Σ_i

Einstein's Field Equations

$$G_{\mu\nu} = 8\pi T_{\mu\nu}$$

Projected onto a time slice generate:

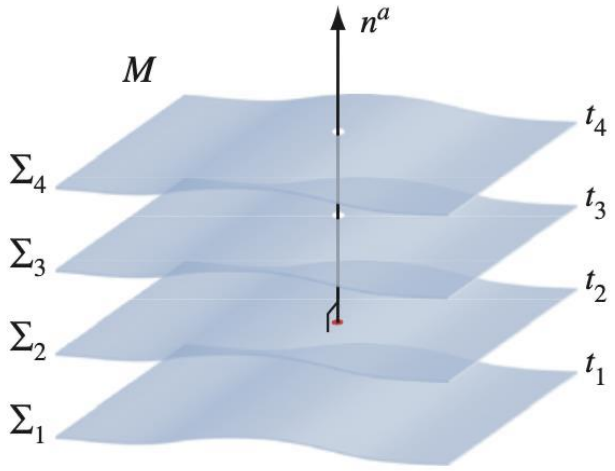
- A Hamiltonian constraint

$R + K^2 - K_{ij}K^{ij} = 16\pi\rho.$ $\rho = n_i n_j T^{ij}$

- And a Momentum Constraint

$D_j K_i^j - D_i K = 8\pi S_i$ $S = -\gamma_i^j n^k T_{jk}$

Mathematical Description



This image is by Faisal Alotaibi from fig 1 from their masters thesis used unchanged under cc-by-4.0

Initial Data Problem

We do not know how to specify the geometric data on the initial time slice for a binary Black Hole merger in a vacuum

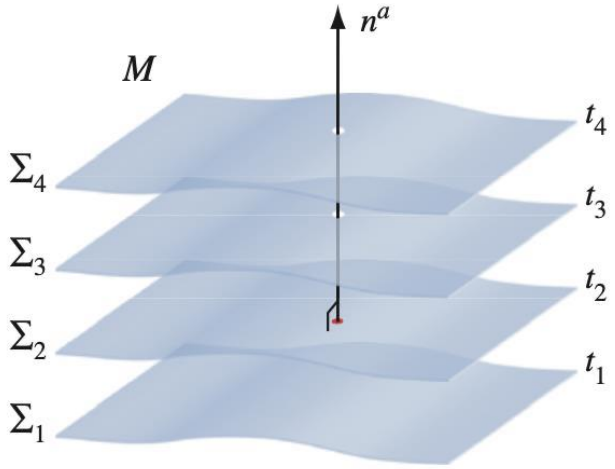
- **One black hole:** *e.g.* Schwarzschild metric
- **Two (or more) black holes:** no analytically known results

The problem is nonlinear: the BHs interact gravitationally so we cannot easily build the two-BH solution from two copies of the one BH solution

Would likes for a binary black hole merger's initial time slice

1. We must have two apparent horizons (*i.e.* we are looking at two black holes)
2. The spin of each black hole (calculated for the apparent horizon) must have the correct value
3. The Christodoulou mass of each black hole (calculated for the apparent horizon) must have the correct value
4. Separation and eccentricity.

Mathematical Description



This image is by Faisal Alotaibi from fig 1 from their masters thesis used unchanged under cc-by-4.0

Finding the Apparent Horizon

I.e. find the outermost *marginally outer trapped surface (MOTS) S* and compute its area from the induced metric

Calculating the spin of the black hole

$$J = \frac{1}{8\pi}\int_S K_{ij}\phi^i s^j dA$$

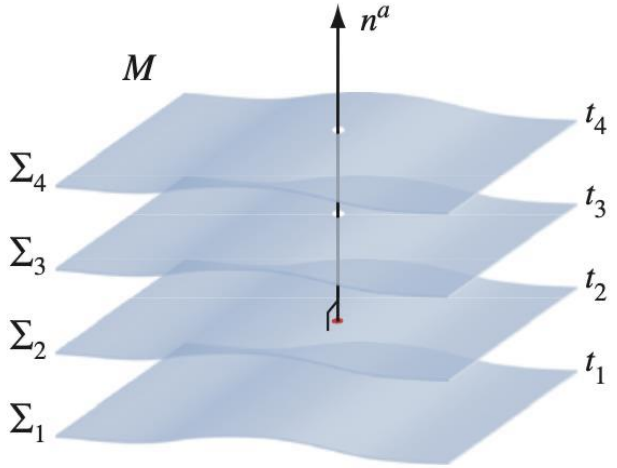
Where s^j is outward-pointing normal unit vector to the horizon, ϕ^i is an approximate rotational Killing vector on the horizon.

Calculating the masses

1. **Irreducible mass** $M_{irr} = \sqrt{A/16\pi}$ *where A is the area of the apparent horizon.*

2. **Christodoulou mass** $M_{Chr} = \sqrt{M_{irr}^2 + J^2/\left(4M_{irr}^2\right)}$

Mathematical Description



This image is by Faisal Alotaibi from fig 1 from their masters thesis used unchanged under cc-by-4.0

Bowen-York Recipe

We can use the Bowen-York recipe to define the initial conditions for the black holes. The ingredients are:

- Conformal transverse-traceless decomposition
- Trace/Traceless decomposition $K_{ij} = A_{ij} + \frac{1}{3}\gamma_{ij}K$
- Maximal Slicing $K = 0$
- Conformal decomposition $\gamma_{ij} = \psi^A \ \bar{\gamma}_{ij}$
- Asymptotically flat metric $\bar{\gamma}_{ij} = n_{ij}$
- Helmholtz theorem $\bar{A}_{ij} = \bar{A}^L_{ij} + \bar{A}^{TT}_{ij}$ w.r.t. vector potential W^i

How to set spin and momentum

Momentum constraint becomes linear

$$\partial^j \partial_j W^i + \frac{1}{3}\partial^i \partial_j W^j = 0$$

then the solution $W^i = \frac{\bar{\epsilon}^{ijk} l_j J_k}{s^2}$ sets the spin

the solution: $W^i = -\frac{1}{4s}\left(7P^i + l^i l_k P^k\right)$ sets the momentum

How to set the mass

One chooses puncture masses in the conformal factor

$$\psi = 1 + \frac{\mathcal{M}_1}{2s_1} + \frac{\mathcal{M}_2}{2s_2} + u$$

Where *u* is a solution of an auxiliary equation depending on other data

$$\bar{D}^2 u = -\beta(\alpha + \alpha u + 1)^{-7}$$

Caveat
The Bowen-York parameters are not identical to the physical parameters they represent

Impacts of a New / Better Solution

- Reduce operational runtime or improve use efficiency of high-performance computing clusters.
- Enable a more robust catalog of sources and parameters and waveforms for LISA data analysis.
- Enable faster solution of problem with less effort / time. Allows for building waveform banks faster and explore regions of parameter space that have not been mapped out entirely.
- Grand challenge problems:
  - Developing advanced techniques for General relativistic magneto hydrodynamics (GRMHD) to study neutron stars and their behavior

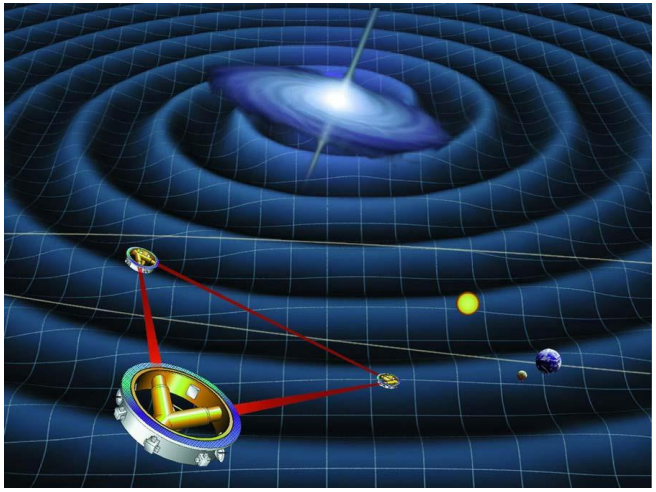


Image Credit: NASA

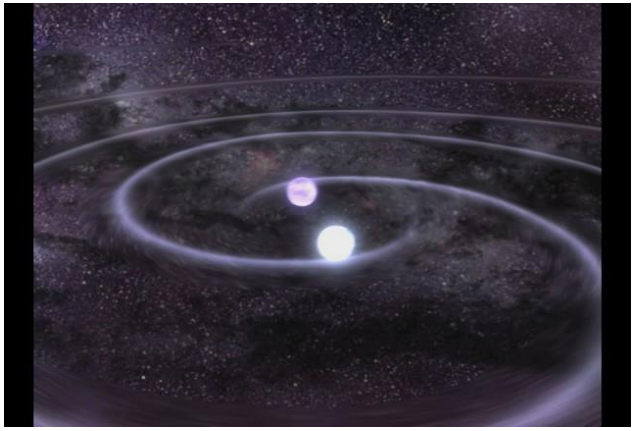
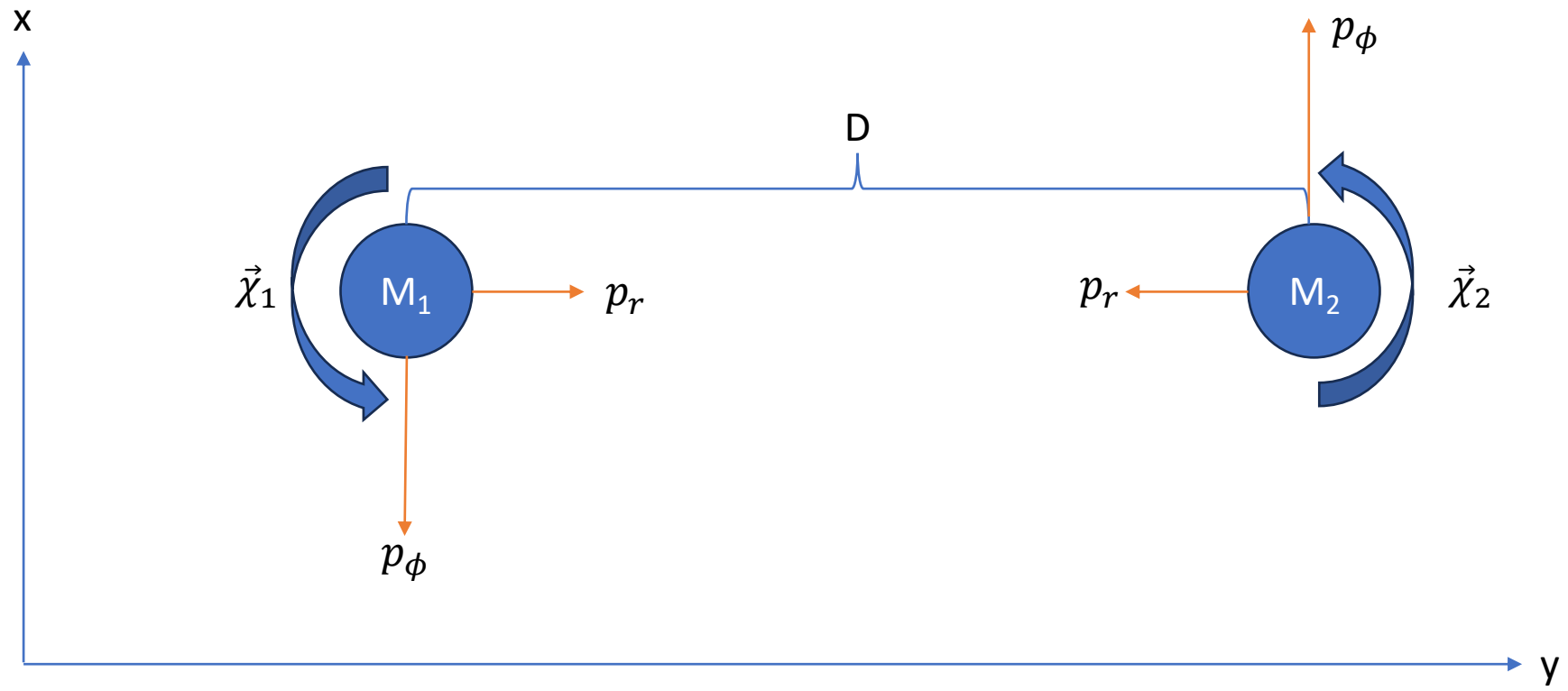


Image Credit: NASA/Goddard Space Flight Center

What specifies and instance

Instances are specified by the:

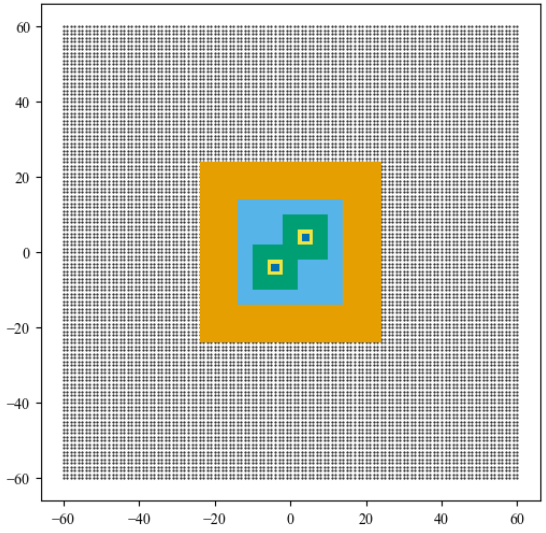
physical parameters (black hole features)



What Specifies an instance

Instances are specified by the:

Numerical parameters (grid specifications and solver specifications)



Grid discretization around the black holes. The black grid is the coarsest while the ones nearer to the black hole become substantially more refined.

Origin is the center-of-mass. We want the values of the spherical harmonic components of the gravitational Weyl scalar $\Psi_4(t,\theta,\phi,R)$ for a fixed "observation radius" R , for times $t_k = k\Delta t$ with $k = 0,1,\ldots,k_{max}$ at points on the spherical surface described by angles (θ,ϕ) for a discretization $\theta_k = k\pi/n_\theta$ and $\phi_j = 2\pi j/n_\phi$ for $n_\theta = 120$, $n_\phi = 240$. Typical values for the spherical harmonics measured are:

$$(\ell,m) = (2,\pm2),(3,\pm2),(4,\pm2),$$
$$(4,\pm3),(3,\pm3),(4,\pm4)$$

Observation radius is usually $R\sim100m$ and measured at multiple values of R.

The values for Δt are dependent on the chosen method of simulation.

Extracting Strain Data

Then the strain h can be obtained from Ψ_4 by

$$h = h_+ - ih_\times = -\int_{-\infty}^{t} dt' \int_{-\infty}^{t'} \Psi_4(t'')dt''$$

The details of how this integral is performed depend on the actual simulation technique especially when looking at the early time limits.

This can be mapped to spherical harmonics as follows

$$h(r,t,\theta,\phi) = \frac{1}{r}\sum_{l,m} h^Y_{lm}(t)\,{}_{-2}Y_{lm}(\theta)e^{im\phi}$$

With:

$${}_{-2}Y_{lm}(\theta) = \left[\frac{(l+s)!}{(l-s)!}\right]\bar{\partial}^{-2}\left(Y_{lm}(\theta)e^{im\phi}\right)$$

Where

$$\bar{\partial} = -\left(\partial_\theta + \frac{i}{\sin(\theta)}\partial_\phi - s\cot(\theta)\right)$$

Löffler *et al*. The Einstein Toolkit: A Community Computational Infrastructure for Relativistic Astrophysics. arXiv:1111.3344v1

Baker, Campanelli, Lousto. "The Lazarus project: a pragmatic approach to binary black hole evolution". *Phys Rev D* **65** 044001 (2002)

Important Caveats

Choice of Units

Typically work in units where:

$$c = G = M_{total} = 1$$

Dimensionful quantities are restored by using the physical values for the various quantities

Consistency Checks

We need to do consistency checks, e.g. energy conservation that full process yields a physical accurate result. (Explain further)

Matching for initial conditions

Various definitions for masses, relative distance, spin, and linear momentum may not agree as input parameter for various methods: Post-Newtonian, Spectral, and Numerical Relativity

That values we provided in the following problem cases are estimates for quasi-circular orbits in Numerical Relativity. However, some amount of nonperturbative tuning will be required to ensure minimal eccentricities and spins.

Important Caveats

Choice of Units

Typically work in units where:

$$c = G = M_{total} = 1$$

Dimensionful quantities are restored by using the physical values for the various quantities

Consistency Checks

We need to do consistency checks, e.g. energy conservation that full process yields a physical accurate result.

Estimating Total simulation time

Estimating a prior how long a simulation should run is not typically easy. The leading methods involve utilizing black-hole perturbation theory to roughly estimate time to merger and then some amount of additional buffer time to account for discrepancies.

Toy Problem

Numerical parameters (We are following the ET GW150914 par file parametrization)

User-defined parameters

Param	Abbrev.	Value
Separation	D	10
Mass ratio	$q = \frac{m_1}{m_2}$	36/29
Total mass	M	1.0
Dimensionless spin BH1	$\vec{\chi}_1$	[0, 0, 0.31]
Dimensionless spin BH2	$\vec{\chi}_2$	[0, 0, -0.46]
Radial Momentum	p_r	-0.0000845
Azimuthal Momentum	p_ϕ	0.09530152

Derived parameters

Param	Abbrev.	Value
Linear momentum of BH1	$\vec{P}_1 \in \boldsymbol{R}^3$	$[P_r, P_\phi, 0]$
Linear momentum of BH2	$\vec{P}_2 \in \boldsymbol{R}^3$	$[-P_r, -P_\phi, 0]$
Mass BH1	m_1	$M \frac{q}{q+1}$
Mass BH2	m_2	$\frac{M}{q+1}$
Position BH1	$\vec{X}_1 \in \boldsymbol{R}^3$	$[m_1\ D, 0, 0]$
Position BH2	$\vec{X}_2 \in \boldsymbol{R}^3$	$[m_2\ D, 0, 0]$
Dimensionful Spin BH1	$\vec{S}_1 \in \boldsymbol{R}^3$	$m_1^2\ \vec{\chi}_1$
Dimensionful Spin BH2	$\vec{S}_2 \in \boldsymbol{R}^3$	$m_2^2\ \vec{\chi}_2$

Units: G=c=1 and we measure "masses" in multiples of the total mass $M = m_1 + m_2$. With these units we have that $M = 1$ by definition.

Stepping Stone Instance (1/3)

INTELLIGENT SYSTEMS DIVISION

Physical parameters (derived from RIT data set here)

User-defined parameters

Param	Abbrev.	Value
Separation	D	7
Mass ratio	$q = \frac{m_1}{m_2}$	64/1
Total mass	M	1.0
Dimensionless spin BH1	$\vec{\chi}_1$	[0, 0, 0]
Dimensionless spin BH2	$\vec{\chi}_2$	[0, 0, 0]
Radial Momentum	p_r	0.0
Azimuthal Momentum	p_ϕ	0.00772
Simulation Time	T	8000

Derived parameters

Param	Abbrev.	Value
Linear momentum of BH1	$\vec{P}_1 \in \boldsymbol{R}^3$	$[P_r, P_\phi, 0]$
Linear momentum of BH2	$\vec{P}_2 \in \boldsymbol{R}^3$	$[-P_r, -P_\phi, 0]$
Mass BH1	m_1	$M \frac{q}{q+1}$
Mass BH2	m_2	$\frac{M}{q+1}$
Position BH1	$\vec{X}_1 \in \boldsymbol{R}^3$	$[m_1\ D, 0, 0]$
Position BH2	$\vec{X}_2 \in \boldsymbol{R}^3$	$[m_2\ D, 0, 0]$
Dimensionful Spin BH1	$\vec{S}_1 \in \boldsymbol{R}^3$	$m_1^2\ \vec{\chi}_1$
Dimensionful Spin BH2	$\vec{S}_2 \in \boldsymbol{R}^3$	$m_2^2\ \vec{\chi}_2$

Goals: order of magnitude improvement in runtime over simulation of parameters using Einstein toolkit or related software package

Stepping Stone Instance (2/3)

INTELLIGENT SYSTEMS DIVISION

Physical parameters (derived from RIT data set here)

User-defined parameters

Param	Abbrev.	Value
Separation	D	7
Mass ratio	$q = \frac{m_1}{m_2}$	64/1
Total mass	M	1.0
Dimensionless spin BH1	$\vec{\chi}_1$	[0, 0, 0]
Dimensionless spin BH2	$\vec{\chi}_2$	[0, 0, 0]
Radial Momentum	p_r	0.0
Azimuthal Momentum	p_ϕ	0.00772
Simulation Time	T	8000

Derived parameters

Param	Abbrev.	Value
Linear momentum of BH1	$\vec{P}_1 \in \boldsymbol{R}^3$	$[P_r, P_\phi, 0]$
Linear momentum of BH2	$\vec{P}_2 \in \boldsymbol{R}^3$	$[-P_r, -P_\phi, 0]$
Mass BH1	m_1	$M \frac{q}{q+1}$
Mass BH2	m_2	$\frac{M}{q+1}$
Position BH1	$\vec{X}_1 \in \boldsymbol{R}^3$	$[m_1\ D, 0, 0]$
Position BH2	$\vec{X}_2 \in \boldsymbol{R}^3$	$[m_2\ D, 0, 0]$
Dimensionful Spin BH1	$\vec{S}_1 \in \boldsymbol{R}^3$	$m_1^2\ \vec{\chi}_1$
Dimensionful Spin BH2	$\vec{S}_2 \in \boldsymbol{R}^3$	$m_2^2\ \vec{\chi}_2$

Goals: order of magnitude improvement in runtime over simulation of parameters using Einstein toolkit or related software package

Stepping Stone Instance (3/3)

Physical parameters (derived from RIT data set here)

User-defined parameters

Param	Abbrev.	Value
Separation	D	7
Mass ratio	$q = \frac{m_1}{m_2}$	128/1
Total mass	M	1.0
Dimensionless spin BH1	$\vec{\chi}_1$	[0, 0, 0]
Dimensionless spin BH2	$\vec{\chi}_2$	[0, 0, 0]
Radial Momentum	p_r	0.0
Azimuthal Momentum	p_ϕ	0.003922
Time to merger	T	8000

Derived parameters

Param	Abbrev.	Value
Linear momentum of BH1	$\vec{P}_1 \in \boldsymbol{R}^3$	$[P_r, P_\phi, 0]$
Linear momentum of BH2	$\vec{P}_2 \in \boldsymbol{R}^3$	$[-P_r, -P_\phi, 0]$
Mass BH1	m_1	$M \frac{q}{q+1}$
Mass BH2	m_2	$\frac{M}{q+1}$
Position BH1	$\vec{X}_1 \in \boldsymbol{R}^3$	$[m_1\ D, 0, 0]$
Position BH2	$\vec{X}_2 \in \boldsymbol{R}^3$	$[m_2\ D, 0, 0]$
Dimensionful Spin BH1	$\vec{S}_1 \in \boldsymbol{R}^3$	$m_1^2\ \vec{\chi}_1$
Dimensionful Spin BH2	$\vec{S}_2 \in \boldsymbol{R}^3$	$m_2^2\ \vec{\chi}_2$

Goals: order of magnitude improvement in runtime over simulation of parameters using Einstein toolkit or related software package

Challenge Problem

Physical parameters

User-defined parameters

Param	Abbrev.	Value
Separation	D	7
Mass ratio	$q = \frac{m_1}{m_2}$	1000/1
Total mass	M	1.0
Dimensionless spin BH1	$\vec{\chi}_1$	[0, 0, 0]
Dimensionless spin BH2	$\vec{\chi}_2$	[0, 0, 0]
Radial Momentum	p_r	0.0
Azimuthal Momentum	p_ϕ	0.003922
Sim Time	T	48000

Derived parameters

Param	Abbrev.	Value
Linear momentum of BH1	$\vec{P}_1 \in \boldsymbol{R}^3$	$[P_r, P_\phi, 0]$
Linear momentum of BH2	$\vec{P}_2 \in \boldsymbol{R}^3$	$[-P_r, -P_\phi, 0]$
Mass BH1	m_1	$M \frac{q}{q+1}$
Mass BH2	m_2	$\frac{M}{q+1}$
Position BH1	$\vec{X}_1 \in \boldsymbol{R}^3$	$[m_1\ D, 0, 0]$
Position BH2	$\vec{X}_2 \in \boldsymbol{R}^3$	$[m_2\ D, 0, 0]$
Dimensionful Spin BH1	$\vec{S}_1 \in \boldsymbol{R}^3$	$m_1^2\ \vec{\chi}_1$
Dimensionful Spin BH2	$\vec{S}_2 \in \boldsymbol{R}^3$	$m_2^2\ \vec{\chi}_2$

Goals: Implementation of problem that is can be cross-validated with a numerical relativity simulation. And is faster than existing studies.
This could be done in principle with sufficient super computing capabilities but is likely too expensive and a more energy/cost-efficient solution is desired.

Other Impacts / Useful References

Other NASA

Computational improvements for binary mergers might lend themselves to studying accretion disk physics in general relativistic magneto-hydrodynamics.

Other Agencies

Studies of Gravitational waves and binary mergers coupled to actual LISA experiments may enable discoveries of new physics such as what is dark matter and dark energy enabling discoveries in particle physics.

Financial Benefits

Reduced or more efficient usage of HPC resources would enable s

Black Hole Perturbation Theory:

[Dav1971] M. Davis, R. Ruffini, W. H. Press, and R. H. Price, Phys. Rev. Lett. 27, 1466 (1971).

[Dav1972] M. Davis, R. Ruffini, and J. Tiomno, Phys. Rev. D 5, 2932 (1972).

[Lou1995] C. O. Lousto and R. H. Price, Phys.Rev. D55, 2124 (1997), arXiv:gr-qc/9609012 [gr-qc].

[Lou1995a] C. O. Lousto and R. H. Price, Phys. Rev. D56, 6439 (1997), arXiv:gr-qc/9705071.

[Lou1998] C. O. Lousto and R. H. Price, Phys. Rev. D 57, 1073 (1998), arXiv:gr-qc/9708022.

Numerical Relativity BSSN

[Shi1995] M. Shibata and T. Nakamura, Phys. Rev. D 52, 5428 1995.

[Baum1998] Baumgarte, T. W., Hughes, S. A., & Shapiro, S. L., 1999, Phys. Rev. D 60, 087501

Post Newtonian Methods

[Bla2024] Blanchet, Luc. "Post-Newtonian theory for gravitational waves." *Living Reviews in Relativity* 27.1 (2024): 4.

Useful Data sets / software

arXiv:2407.18319 has spin aligned with rotation. $1 \leq q \leq 1000, \quad -0.8 \leq \chi_1 \leq 0.8$ Black hole perturbation theory

arXiv:2203.10109 has spin aligned with rotation. $1 \leq q \leq 30, -0.5 \leq \chi_1 \leq 0.5$ uses Numerical relativity.

arXiv:2006.04818 has no spin. $q = 15, 32, 64, 128$. No eccentricities.

arXiv:2202.00018 $0 \leq e \leq 1, 1 \leq q \leq 128,$

arXiv:2309.00262 $q < 15, \ |\chi| \leq 1$

arXiv:2505.13378 $q \leq 8, |\vec{\chi}| \leq 0.8, e \sim 1$ (total catalog time 480,000,000 core hours)

https://github.com/gmorras/pyEFPE/tree/main License: MIT

https://github.com/nrpy/nrpypn License: BSD-2

Discussions

NASA Domain Experts we communicated with

- Alexander Ruestle (GSFC-Adnet Systems Inc.)
  - Google Scholar Profile
- Bernard J. Kelly (GSFC-University of Maryland Baltimore County)
  - Google Scholar Profile
- Jeremy Schnittman



LISA Global Fit

LISA Global Fit

Application Details

Goal: accurately identify gravitational wave signatures for mergers from noise **strain measurement data streams**

- Disambiguate noisy data from strain measurements on LISA. Background noise and various other simultaneous sources can conflate measurements contributing to backgrounds and conflicting signals.
- Perform global fit across the frequency spectrum of strain measurements to adequately disambiguate signals.

Problem is repeated analysis of ongoing signal acquisitions from LISA.

Problem Details

- Accurately discerning gravitational wave signals from an unknown number of sources and corresponding types of gravitational waves.
- Inputs: Fit models, strain data time series calibrated and noise corrected Time-Delay Interferometry (TDI) data streams [ESA2024].
- Output: Probability density functions for identified sources. Sky locations for possible mergers and events.
- Repeatedly analyzing data.

Current Approaches

- Markov Chain Monte Carlo is frequently used to select combinations and parameters for various models due to its ability to find accurate solutions in high dimensional parameter spaces, such as the one LISA global fit falls under (~10^5 parameters)

Impact of New/Better Solution

- Faster fitting and identification of gravitational wave sources would allow for possible observations.
- Accurate identification of GW sources would enable concrete categorization of source parameters.

Problem Details

Problem description

LISA output strain data stream is given by:

$$s_\alpha(t) = h_\alpha\left(t, \vec{\lambda}\right) + n_\alpha(t) = \sum_{i=1}^{N} h_\alpha^{(i)}\left(t, \vec{\lambda}_i\right) + n_\alpha(t)$$

Detector Response In channel

Source parameters

Instrument noise In channel

We want to find a solution which fits to N' "visible" sources and can adequately describe a good fit to the data.

Specification for Problem:

- Core enabling computation: Solving a high dimensional version of the "Cocktail party problem" of identifying various sources under noise of other sources
- Inputs: Catalogs (possible models) to fit against, noise processed strain data
- Outputs: Probability density functions for identified sources. Sky locations for possible mergers and events.

Complications to the global fit:

- Individual source models typically require between 7-17 parameters depending on the complexity of source.
- Number of sources resolvable above noise is very large, possibly ~10^4 sources and 10^5 parameters to fit with different model catalogs.
- Number of resolvable sources is unknown so a convergence criterion is required to avoid overfitting.
- Signal-to-noise ratios for background or distant sources may contaminate readings.
- Issues with identifying models via label-switching ambiguities and the indistinguishability of sources [arXiv:2502.14818]

Selection of models M_i

- Probability $p(M_i|d) = \frac{p(d|M_i)\pi(M_i)}{p(d)}$ (Bayes Theorem)
- $p(d|M_i) = \int p\left(\vec{\theta}\middle|d, M_i\right)\pi\left(\vec{\theta}\middle|M_i\right) d\,\vec{\theta}$
- Use Bayes factor to select updates

$$B_{M_i/M_j} = \frac{p(d|M_i)}{p\left(d\middle|M_j\right)}$$

- Find model with greatest likelihood.

State of the Art

LISA Global Fit

State of the Art methods for solving the global fit problem include using **Markov Chain Monte Carlo** methods to search through both model and parameter space to identify the best solution [Cor2005] this forms the basis for most existing methods for solving the global fit problem.

Most existing algorithms such as Erebor [Kat2024] or others [Den2025]. These pipelines are split into a **detection** and **optimization** step.

- **Detection:** this includes identification of possible sources and adds them to the underlying model to be fit (Massive Binary Black Holes, White Dwarf Binary, Extreme Mass Ratio Inspirals, stellar-origin black hole binaries, unmodelled GW sources, backgrounds)
- **Optimization:** this takes the model and optimizes the parameters to understand what the underlying gravitational wave sources are (e.g., explicit mass ratio, intrinsic spin).

Computational concerns:
Given the nature of the search methods there are strong needs for parallelism.

Formal complexity results:
The **Markov Chain Monte Carlo** methods for LISA global fits typically linearly with the number of parameters. Depending on underlying factors and implementations on measuring the statistical significance of certain signals as the number of parameters grows.

Existing Datasets Instance Specifications:

- LISA Data set Challenge 2: M. Le Jeune and S. Babak, 10.5281/zenodo.7132178 (2022)
- LISA Data Challenge Spritz (LDC2b) (Noisy Data)

Any solution needs to be able to accurately reproduce data in the LISA Data set challenge or other corresponding data set.

[Cor2025] Cornish and Chowder “LISA Data Analysis using MCMC methods” https://arxiv.org/pdf/gr-qc/0506059
[Kat2024] Katz, Michael L., et al. "Efficient GPU-accelerated multisource global fit pipeline for LISA data analysis." *Physical Review D*111.2 (2025): 024060.
[Den2025] Deng, Senwen, et al. "Modular global-fit pipeline for LISA data analysis." *Physical Review D* 111.10 (2025): 103014.

Understanding Noise

Types of Noise

- Easy to handle background type noise:
  - Gaussian $n(t) = \mathcal{N}(0, \sigma)$
  - Circulant (infinite duration or periodic)
  - Toeplitz (infinite duration or periodic)
- Hard to handle
  - Data gaps or instrumental glitches
  - Ones that interrupt data sequences
  - Or fall outside the gaussian and stationary structure

Computational Complexity

- Gaussian/Toeplitz $\mathcal{O}(N \log_2 N)$ when computations are done in Frequency space
- Data gaps can have scaling between $\mathcal{O}(N^3)$ and $\mathcal{O}(N^2 \log_2 N)$ depending on type of data gaps

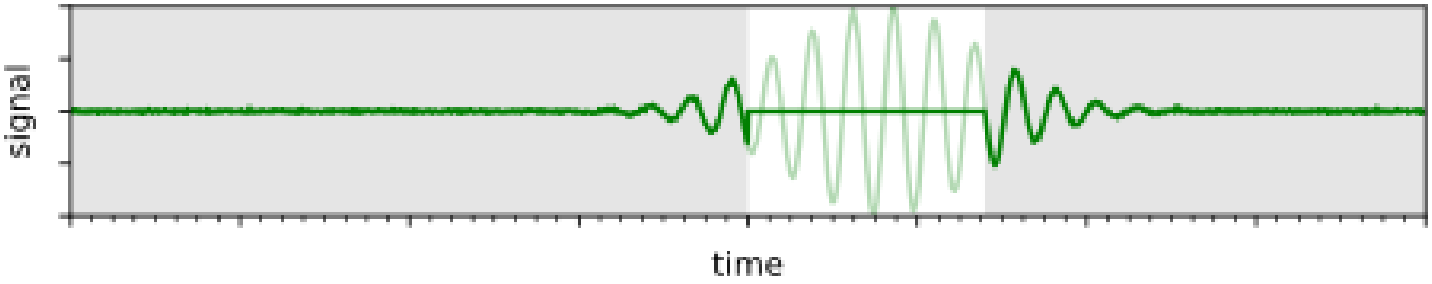
Methods for tackling data gaps

- Data augmentation [Bag2019]
  - This method utilizes Bayesian analysis to gap fill missing regions of data by estimating the type of noise in the background
- Windowing [Zac2021]
  - This method works around the gap by ignoring the missing data and only treating the information present using a ”window” to block the data gap from the signals.

Windowing Side effects

- loss of correlation between data segments
- Short data segments can introduce edge effect errors from the Fourier transform
- Loss of resolvable frequencies if the continuous data segments are too short in time

Figure (right): cartoon example of a data gap in the gravitational wave time delay interferometry signal. Unshaded region indicates the gap in the signal



[Bag2019] Q. Baghi, I. Thorpe, J. Slutsky, J. Baker, T. Dal Canton, N. Korsakova, and N. Karnesis, Phys. Rev. D 100, 022003 (2019),
[Zac2021] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, Phys. Rev. D 104, 063034 (2021)

Impact of New/Better Solution

LISA Global Fit

- Improve overall utilization of the low-bandwidth data from the LISA strain data stream.
- Better solutions may enable an improved categorization and detection of GW sources of scientific interest to the LISA experiment.
- Improvement sought: an improvement that provides improved confidence intervals or substantially improved runtime (order of magnitude faster).
- Faster solutions may help improve catalogs and detections.
- Grand Challenge extensions
  - Are there ways to better understand crowded signals and disambiguation of them in general for ”Cocktail Party Problem”
  - Can we couple GW studies with other multi-messenger studies to better understand astrophysical observables.
  - Possible violations of the strong equivalence principle heralding new physics.

Challenges to Global Fit

- Missing data issues. The Time delayed interferometry precludes many fast-to-evaluate models for the parameter inference with respect to the underlying noise [Bur2025].
- Data glitches are often of the form of:
  - non uniform or white noise.
  - Gaps and missing pieces of data
- All of these types of data essentially make Fourier sampling of the data substantially more challenging for signal extraction.

[Bur2025] Burke, Ollie, et al. "Mind the gap: addressing data gaps and assessing noise mismodeling in LISA." *arXiv preprint arXiv:2502.17426* (2025).

Definitions: Accuracy

A common metric to assign accuracy or "goodness-of-fit" is the signal overlap:

$$\mathcal{O}\big(s(\theta_{rec}), s(\theta_{act})\big) = \frac{\langle s(\theta_{rec})|s(\theta_{act})\rangle}{\sqrt{\langle s(\theta_{rec})|s(\theta_{rec})\rangle\langle s(\theta_{act})|s(\theta_{act})\rangle}}$$

Where,

$$\langle s(\theta_{rec})|s(\theta_{inj})\rangle = 4\,Re\int_{f_{low}}^{f_{high}} df\ \frac{\tilde{s}^*(f;\theta_{rec})s(f;\theta_{act})}{\mathcal{S}_n(f)}$$

- With $\mathcal{S}_n(f)$ being the single sided power spectral density calculated using the `AnalyticNoise` which is available at https://pypi.org/project/lisa-data-challenge/,
- $s(\theta)$ is the time domain signal,
- θ_{rec} is the value determined by the fitting procedure
- θ_{act} are the ones underlying the signal in the synthetic data sets

An additional metric frequently used is the signal mismatch denoted as

$$\mathcal{M} = 1 - \mathcal{O}\big(s(\theta_{rec}), s(\theta_{act})\big)$$

Underlying data sets and previous work

The toy, stepping stone, and challenge problem instances that we consider in this work recommend utilizing the LISA data challenge pipeline software which contains many of the necessary configurations to develop the test case instances https://lisa.pages.in2p3.fr/LDC/data_generation/running.html

There is some set of heavy lifting for the challenge problems that will require introducing new types of data glitches beyond what is present in the various LISA data challenge problem instances. This will require modifications to the pipeline to enable more complex data glitch artifacts.

Additionally beyond the scope of the problem instances provided we suggest comparing against recent LISA data challenge problems for the early steps for which existing solutions already exist (a non exhaustive list is provided below):

Toy problem instance

arXiv:2511.07385 , arXiv:2510.20891 , arXiv:2505.09600 , arXiv:2502.14818, arXiv:2501.10277 , arXiv:2405.04690 , arXiv:2403.15318

Stepping stone problem instance 1

https://arxiv.org/abs/2502.17426

Toy Problem Instance (1/2)

Problem: using the Radler1a challenge, https://sbgvm-151-90.in2p3.fr/challenge1a, to accurately identify all the single gravitational wave source in each of the six challenge instances.

Target Accuracy

The mismatch between the Sangria signal and the derived signal should be:

$$\mathcal{M} = 1 - \mathcal{O} \leq 10^{-3}$$

Lower values of $\mathcal{M}$ are preferred

Time-to-solution improvement

At least as fast as $\mathcal{O}(N^2)$ and hopefully up to as fast as $\mathcal{O}(Nlog(N))$ instance cases for best known solutions

Key features to solve for

The key goal is to solve for the signal $h(t, \vec{\theta})$ buried under signal $s(t, \vec{\theta})$ that includes noise. For each of the 6 challenge problems.

Availability

The software for the Radler data instance is available at the following repository https://sbgvm-151-90.in2p3.fr/challenge1a

Toy Problem Instance (2/2)

Problem: using the Sangria synthetic data sets available at https://lisa.pages.in2p3.fr/LDC/data_generation/running.html and https://zenodo.org/records/7332221 accurately identify all mergers and corresponding system parameters

Target Accuracy

The mismatch between the Sangria signal and the derived signal should be:

$$\mathcal{M} = 1 - \mathcal{O} \leq 10^{-3}$$

Lower values of $\mathcal{M}$ are preferred

Time-to-solution improvement

At least as fast as $\mathcal{O}(N^2)$ and hopefully up to as fast as $\mathcal{O}(Nlog(N))$ instance cases for best known solutions

Key features to solve for

The key goal is to solve for the parameters defining the waveform signals for the 15 super massive black hole mergers embedded in a galactic background defined by white dwarf binaries.

Correctly identify the type and parameters of the supermassive binary black hole mergers within 5%. And have an overall signal mismatch, $\mathcal{M} < 10^{-3}$.

Availability

The software for the Sangria data instance is available at the following repository https://zenodo.org/records/7332221

Stepping-Stone Instance (1/3)

INTELLIGENT SYSTEMS DIVISION

Problem: using the Spritz synthetic data sets available at https://sbgvm-151-90.in2p3.fr/challenge2b accurately identify the two mergers with existing glitches present in the data.

Target Accuracy

The mismatch between the Sangria signal and the derived signal should be:

$$\mathcal{M} = 1 - \mathcal{O} \leq 10^{-3}$$

Lower values of $\mathcal{M}$ are preferred

Time-to-solution improvement

Faster than $O(N^2 log(N))$ or compelling case that prefactors for higher scaling, e.g. $O(N^3)$, are sufficiently small that that in meaningful study values of N.

Key features to solve for

The key goal is to solve for the parameters defining the waveform signals for the gravitational wave source in both instance cases.

Correctly identify the type and parameters of the gravitational wave source mergers to within 5%. And have an overall signal mismatch, $\mathcal{M} < 10^{-3}$.

Availability

The software for the Spritz data instance is available at the following repository https://sbgvm-151-90.in2p3.fr/challenge2b

Stepping Stone Instance (2/3)

INTELLIGENT SYSTEMS DIVISION

Problem: Using the Sangria synthetic data sets available at accurately identify all 3 mergers and corresponding system parameters under the following systematic noise with a missing signal of (3hrs every 14 days) modelling antenna pointing.

Target Accuracy

The mismatch between the Sangria signal and the derived signal should be:

$$\mathcal{M} = 1 - \mathcal{O} \leq 10^{-3}$$

Lower values of $\mathcal{M}$ are preferred

Time-to-solution improvement

Faster than $O\left(N^2 log(N)\right)$ or compelling case that prefactors for higher scaling, e.g. $O(N^3)$, are sufficiently small that that in meaningful study values of N.

Key features to solve for

The key goal is to solve for the parameters defining the waveform signals for the gravitational wave source in both instance cases.

Correctly identify the type and parameters of the gravitational wave source mergers to within 5%. And have an overall signal mismatch, $\mathcal{M} < 10^{-3}$.

Availability

The software for the Sangria data instance is available at the following repository https://sbgvm-151-90.in2p3.fr/challenge2a but will need to be modified to include the 3hrs gaps present in the system.

Stepping Stone Instance (3/3)

INTELLIGENT SYSTEMS DIVISION

Problem: Using the Sangria synthetic data sets available at accurately identify all 3 mergers and corresponding system parameters under the following systematic noise with a missing signal of 100 seconds 3x / day.

Target Accuracy

The mismatch between the Sangria signal and the derived signal should be:

$$\mathcal{M} = 1 - \mathcal{O} \leq 10^{-3}$$

Lower values of $\mathcal{M}$ are preferred

Time-to-solution improvement

Faster than $O(N^2 log(N))$ or compelling case that prefactors for higher scaling, e.g. $O(N^3)$, are sufficiently small that that in meaningful study values of N.

Key features to solve for

The key goal is to solve for the parameters defining the waveform signals for the gravitational wave source in both instance cases.

Correctly identify the type and parameters of the gravitational wave source mergers to within 5%. And have an overall signal mismatch, $\mathcal{M} < 10^{-3}$.

Availability

The software for the Sangria data instance is available at the following repository https://sbgvm-151-90.in2p3.fr/challenge2a but will need to be modified to include the 100 second gaps present in the system.

Challenge Problem

- **Problem**: Using the Sangria synthetic data sets available at accurately identify all 3 mergers and corresponding system parameters under the following systematic noise with the following modifications:
  1. (3 hrs every 14 days) modelling antenna pointing
  2. Point ahead adjustment 100 secs 3x / day
  3. Tilt-to-length adjustment 2 days 4x / year
  4. Unplanned outages such as micrometeoroid impacts (1 day each at 30x / year)
  5. Minor instrument outages 10 seconds / year
- **Target Accuracy:** mismatch between fit and signal $\mathcal{M} = 1 - \mathcal{O} \leq 10^{-3}$
- **Time-to-Solution Improvement:** Faster than $O\left(N^2 log(N)\right)$ or compelling case that prefactors for higher scaling, e.g. $O(N^3)$, are sufficiently small that that in meaningful study values of N.

Other Impacts

Non-SMD and External Impacts

The signal disambiguation problem sometimes referred to as the cocktail party problem. This has a related form in biological systems such as auditory processing of signals, i.e. picking up the signal you want to find buried in a background of related sounds.

Solutions to this set of problems could enable a variety of developments in biological / neurological / machine learning systems or other problems where one is trying to extract a signal from overlapping background systems.

It is possible there could be benefits from the solution to this problem for various other detector technologies such as for CERN/SLAC/DUNE. Especially w/r/t disambiguating signals. Such as calorimetry

There are applications to solving eigenmode decompositions for quantum simulation tasks which may prove useful for Early Fault-tolerant quantum simulations.

Wave and turbulence data for oceanographic time series data.

Discussions

NASA Domain Experts we communicated with

- Alexander Ruestle (GSFC-Adnet Systems Inc.)
  - Google Scholar Profile
- Bernard J. Kelly (GSFC-University of Maryland Baltimore County)
  - Google Scholar Profile
- Jeremy Schnittman



N-Body Cosmology

Astrophysical N Body Problems

Application Details

- **Scientific Goal:** probing the properties of the early universe (~10^7 years before present)
- **Technical Goal:** interpreting the observations of the Hubble telescope and the Roman Space Telescope (RST)
- **Numerical Goal:** developing more efficient N-body simulation techniques that allows one to perform statistical analysis (inverse problem) on the observed data and infer Bayesian posteriors of physical quantities of interest about the early universe

Problem Details

- **Core computation:** integrating the extremely long time (Newtonian or Post-Newtonian) dynamics for a system of a large number of objects, in regimes where interactions cannot be neglected
- **Input:** initial conditions
- **Output:** desired observables in the time-propagated configuration(s), at given times over the full time-interval targeted by the simulation

Current Approaches

- **State-of-the-art methods**
- NBODY6++ variants use a Hermite interpolation scheme + regularizations
- PeTaR uses a tree-like representation for the particles that separates long-range from short range interactions, which are then treated independently (hybrid time integration).

- Simulations of ~10^6 bodies for O(10) Gigayears (Gyr) are barely doable with very large effort, but are far too costly for solving the inverse problem (*i.e.* Bayesian inference)

Impact of New or Better Solution

- Current N-body simulations are runtime-limited: specific galaxy formation processes simply cannot be studied as collisional systems of 10^7 stars evolving for ~10^{10} years would require years-long simulations on currently-available supercomputers
- Significantly faster computations would allow one to infer properties of the the early universe by solving the inverse problem for such star systems (currently impossible to even tackle), and more generally study the processes of formation of galaxies.

Astrophysical N Body Problems



Image by NASA's Goddard Space Flight Center.
https://svs.gsfc.nasa.gov/15069/#section_credits

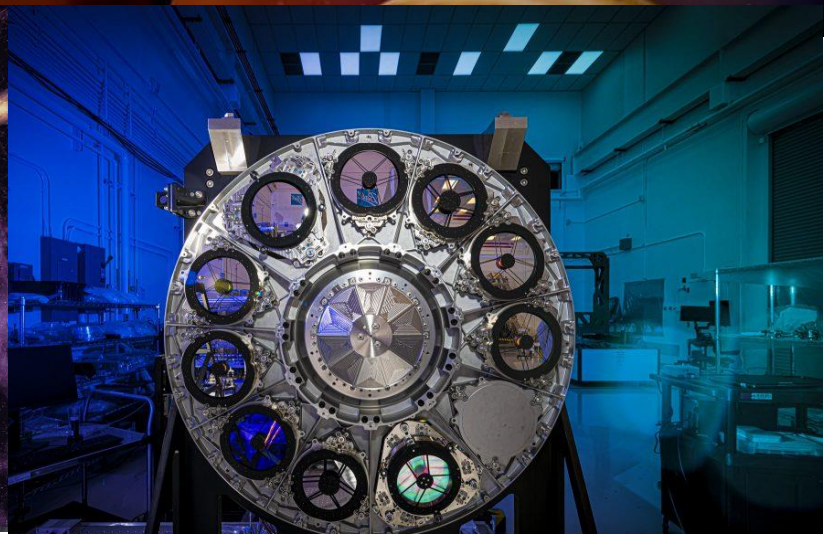
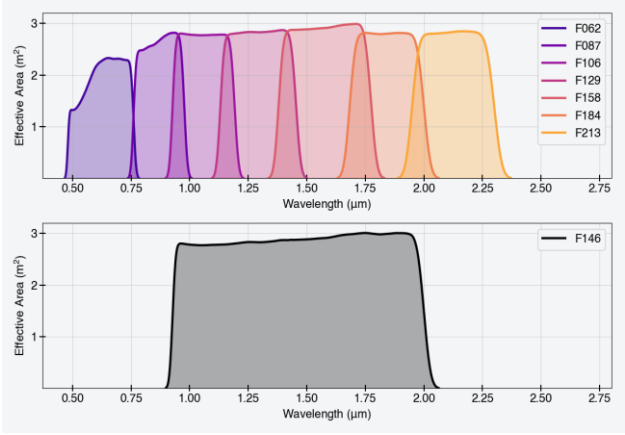


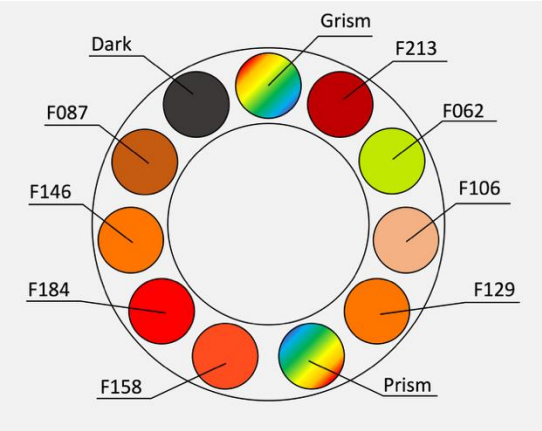
Image by NASA's Goddard Space Flight Center.
https://etd.gsfc.nasa.gov/our-work/roman/

The Roman Space Telescope

Launched on August 30, 2026

- **Specs:** Wide-Field Instrument (WFI) and Chronograph Instrument (CGI)
- **Orbit:** L2 Lagrange point
- **Objective:** imaging with Hubble-like resolution over a much wider areas (~100x) compared to the Hubble Space or James Webb Telescopes.
- **Main requirement:** satellites need guidance on what information to collect and when to do it.





Above images under public domain from the STScI Roman Telescope Repository at this link.

Astrophysical N Body Problems



Image by NASA's Goddard Space Flight Center.
https://svs.gsfc.nasa.gov/15069/#section_credits

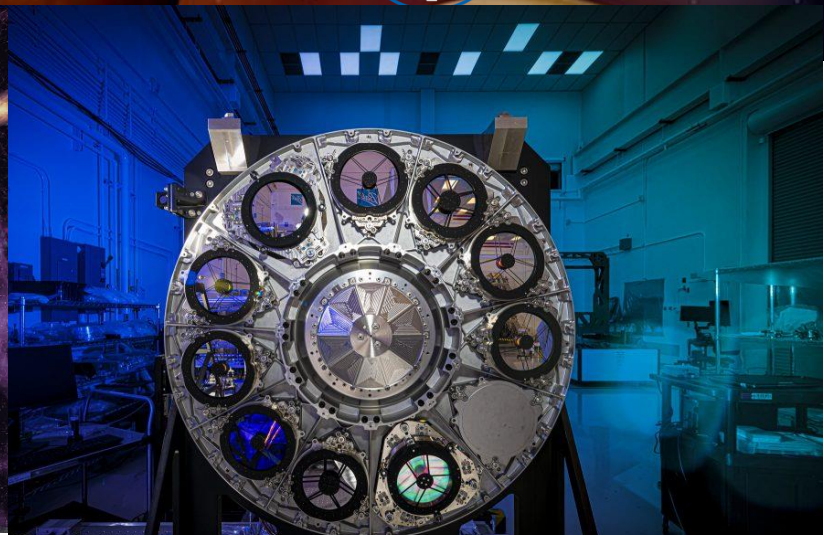
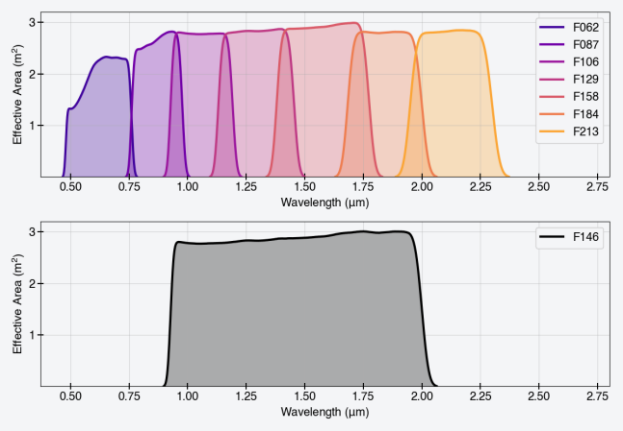
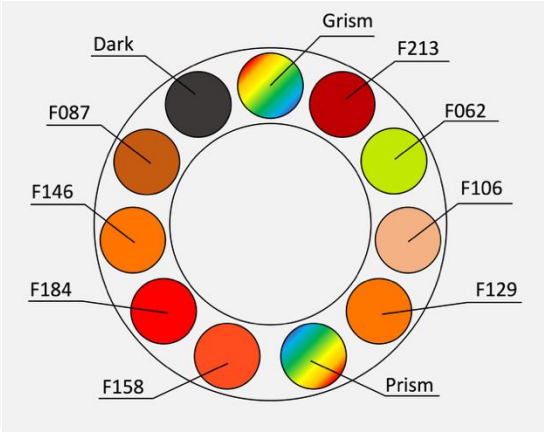


Image by NASA's Goddard Space Flight Center.
https://etd.gsfc.nasa.gov/our-work/roman/

The Roman Space Telescope

- **Goal: Infer dynamical history from a present-day snapshot.** Roman observes a cluster at one stage of its multi-Gyr evolution. N-body calculations provide the time dimension: they let us ask which initial masses, sizes, concentrations, binary fractions, etc. evolve into the structural and kinematic state Roman sees today. In this sense the simulations are the forward component of an eventual inverse problem.





Astrophysical N Body Problems

Globular Clusters

- These are clusters of gravitationally-bound 10^4-10^6 stars
- Their tight structure can make them extremely long-lived (possibly 10^9 years)
- There are many in globular clusters in a galaxy. In the Milky Way they are in the bulge and the halo.
- Their dynamics are collisional: stars "bump" into one another (needs direct-summation)

Open Clusters

- Smaller and less densely populated clusters of 10^1-10^2 stars
- Not as stable as globular clusters, can disperse after ~10^6 years
- In the Milky Way they are in and between the spiral arms
- Their dynamics are collisionless

Nuclear Star Clusters

- These are very dense (a few parsecs wide) cluster of 10^6-10^8 stars
- There's typically only one per galaxy, located in the center of the galaxy
- The dynamics are collisional

Scientific Topics and Questions about the Early Universe

- Globular-cluster internal dynamical evolution, relaxation, and core evolution
- Mass segregation and evolution of the stellar mass function in globular clusters
- Compact-remnant retention and dynamical evolution in globular clusters
- Tidal stripping, stellar escape, tidal tails, and dissolution of star clusters

Ideally, one would need to propagate the current state of the universe backwards in time (for billions of years). However, observed data is partial, noisy so the goal is Bayesian inference for target parameters (compute the posterior distribution for the parameter values given the observed data)

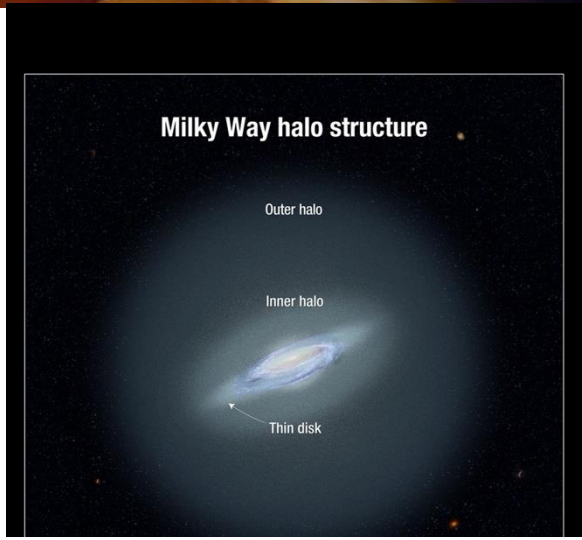


Image credit NASA, ESA, A. Feild (STScI); https://science.nasa.gov/missions/hubble/stellar-archaeology-traces-milky-ways-history/



The globular cluster, NGC 2298, sparkles in this new NASA Hubble Space Telescope image.
NASA, ESA, G. Piotto (Universita degli Studi di Padova), and A. Sarajedini (Florida Atlantic University); Processing: Gladys Kober (NASA/Catholic University of America)

Astrophysical N Body Problems

Utility and Benefits

Goal

- Faster N-body simulations, without decrease in accuracy

Utility Argument

- Today, N-body simulations of dense star clusters require massive computational resources. This limits scientists to studying a small number of cases through "heroic" computational efforts
- Significantly faster algorithms will be needed to turn these rare efforts into practical research tools
- Inverse problems require integrating the dynamics of the system from many initial conditions (forward model) in order to compute likelihoods of physical parameters of interest (typically describing long-past), given the data observed today.

Candidate Applications for Faster N-body Simulations

1) *The origin of merging black holes*
   Black-hole mergers are the flagship of Gravitational Wave research, but did black-hole binaries evolve in isolation or were they assembled in dense star clusters? Inverse problem can infer distributions of physical parameters that one can match with GW catalogues that we have observed.
2) *Can globular clusters generate intermediate-mass black holes?*
   Intermediate-mass black holes lie in between stellar black holes (formed by the gravitational collapse of a star) and the supermassive black holes at galactic centers. One proposed formation mechanism for intermediate-mass black holes is a runaway sequence of stellar collisions in a young, exceptionally dense cluster. N-body simulations could tell us which cluster conditions produce these black holes, whether they remain in the cluster, and what gravitational-wave signatures would reveal them.
3) *Can we use stellar streams to infer the presence of dark matter?*
   As globular clusters orbit a galaxy, they can lose stars that form into long and narrow "streams". Gravitational interactions with dark matter can create distortions in these stellar streams. However, the streams are also distorted through ordinary cluster evolution, tidal stripping, etc. These more typical effects must be "factored out" before one can conclude the residual effect is evidence of dark matter. The Roman telescope is expected to resolve many of these factors
4) *How did nuclear star cluster originate?*
   Many galaxies have an extremely dense star cluster at their center, but how did it form? One possibility is that several globular clusters gradually spiraled inward and eventually merged together. Another is that the dense cluster formed from gas already present near the center of the galaxy. N-body simulations can test the likelihood of the infall scenario, but the statistical analysis would require immense computational effort
5) *"Archaeology" of star clusters*
   What did today's globular clusters look like in the early universe? How did they originate, how many of them were destroyed and how? Uncovering the state of the early universe requires connecting hydrodynamical and cluster-evolution simulations, and current computational costs restrict systematic explorations.

Astrophysical N Body Problems

Current State-of-the-Art Algorithms

- **NBODY6++ and GPU versions**

- **Hermite interpolation:** the NBODY6 integrator uses an interpolation scheme to improve time discretization error $O(\Delta t^4)$ from the standard leap-frog scheme error $O(\Delta t^2)$.
- **Timestep Choice:** the timestep is adaptively chosen to empirically minimize time-discretization error and limit computational effort
- **Regularization:** when two bodies are too close together (a "close encounter" happens), their timesteps become small and truncation effects get big. In this case the 2-body system's dynamics is excised from the algorithm's usual behavior and is resolved through *ad hoc* methods ("regularizations") that propagate the 2-body state forward in time until their relative distance is again large enough that the usual simulation can resume.

- **PeTar**

- **Hybrid force calculation (P3T):** PeTar splits gravitational interactions into short-range and long-range parts. Nearby particles are integrated accurately with direct particle-particle methods, while distant forces are approximated with a Barnes–Hut tree, reducing the long-range computational cost to roughly $O(N \ln N)$.
- **Hybrid time integration:** long-range forces are evolved with a second-order leapfrog/kick-drift-kick scheme, while short-range interactions of single stars and subsystem centers of mass are integrated with a fourth-order Hermite method.
- **Close encounters and binaries (SDAR):** compact binaries, triples, and other strongly interacting few-body subsystems are identified and treated with slow-down algorithmic regularization (SDAR). This regularizes close encounters and avoids resolving every orbit of a weakly perturbed hard binary, greatly reducing the cost associated with very short orbital timescales.
- **Parallelism:** the tree/direct decomposition is designed for large-scale parallel execution using MPI, OpenMP, SIMD, and GPUs, allowing PeTar to handle much larger NN and much higher binary fractions efficiently.

Astrophysical N Body Problems

INTELLIGENT SYSTEMS DIVISION

Stepping-Stone Instances

- **Few-body Simulations: binary-single interactions**
  - **Informal Goal:** reproduce the single-binary scattering process from figure 4 of paper [1]: two bound stars and one unbound star, all of equalmass, binary is initially circular. Produce dimensionless cross-section as a function of the velocity at infinity normalized by the critical velocity. Separate by outcome (shown: exchange, ionization)

Pipeline for datapoints production

1) Choose value of control variable x
2) Sample the ensemble and produce initial conditions
3) Evolve the initial condition to the final state
4) Based on outcome X ∈ {exchange, ionization}, update observed frequencies N_X/N_{trials}
5) Repeat steps 2-4 to compute cross-section integrals
6) Plot two datapoints (for A=exchange, or A=ionization) for given x value

Units	$G = m = a = 1$
Force law: Newtonian	$a_i = G\sum_{j\neq i} m_j \frac{r_j - r_i}{\lvert r_j - r_i\rvert^3}$
Critical speed	$v_c = \sqrt{3Gm/(2a)}$
Particles involved	2 in binary (A, B), 1 intruder (C)
Control variable	$x = v_\infty / v_c$
Starting distance	$R_0/a = (4/\delta)^{1/3} = 73.68$
Ensemble parameters	$b, \Omega_{\rm bin}, \eta$

$$\sigma(x) = \int db\, d\Omega_{\rm bin}\, d\eta\, p(b, \Omega_{\rm bin}, \eta)\mathbf{1}[\text{outcome} = A]$$

impact parameter $p(b) = \frac{2b}{b_{\max}^2}, \quad 0 \le b \le b_{\max} = 5a\sqrt{1 + \frac{6Gm}{5av_\infty^2}}$

$$d\Omega_{\rm bin} = d(\cos\theta)\, d\phi$$

inclination $\cos(\theta) \sim \text{Unif}(-1, 1)$

azimuthal orientation $\phi \sim \text{Unif}(0, 2\pi)$

orbital phase $\eta \sim \text{Unif}(0, 2\pi)$

[1] Fregau *et al.* "stellar collisions during binary-binary and binary-single star interactions" Monthly Notices of the Royal Astronomical Society, 2004

Astrophysical N Body Problems

Intermediate Challenge

- **Informal Goal 1: Globular Cluster Simulations.** Simulate the relaxation of a globular star cluster (of $N=10^6$ stars with equal mass) of total mass $M = 1.2 \times 10^6$ solar masses.

Pipeline:

1) **Generate the initial conditions according to the King model**
2) Evolve the globular cluster forward in time for 12 billion years
3) Report the values of the target quantities at the required times

[1] I. R. King, "The Structure of Star Clusters. III. Some Simple Dynamical Models," Astronomical Journal 71, 64–75 (1966).
[2] M. Gieles and A. Zocchi, "A family of lowered isothermal models," MNRAS 454, 576–592 (2015)
[3] A. H. W. Küpper, T. Maschberger, P. Kroupa, and H. Baumgardt, "Mass segregation and fractal substructure in young massive clusters: I. The McLuster code and method calibration," MNRAS 417, 2300–2317 (2011), Appendix A

Globular Cluster Model

The stars in the globular cluster are generated by sampling the King model (see *e.g.* [1],[2],[3] and the box in the next slide for the definition of the relevant quantities).

After defining the mass $m_i = M/N$, for each star $i=1,\ldots,N$:

Generate the position vector

Sample the radial coordinate r of the star from the density, *e.g.*

1) Sample $u \simeq \text{Unif}(0,1)$
2) Solve for x the equation $I(x) = u\, I_t$
3) Obtain the value of the radial coordinate by $r = r_0 x$

Sample the angular components of the position vector uniformly at random (spherical symmetry)

Generate the velocity vector

Sample the velocity modulus

$$A(y,W) \equiv \exp\left(W(1-y^2)\right)/\exp(W)$$

1) Sample $u,v \simeq \text{Unif}(0,1)$ independently
2) Define $y \equiv u^{1/3}$
3) Compute $A(y, W)$
4) Accept if $v < A(y,W)$ and reject and repeat 4)-7) until a candidate is accepted
5) Obtain the radial component of the velocity $v \equiv \sigma_K \sqrt{2Wy}$

Sample the angular components of the velocity vector uniformly at random (spherical symmetry)

Remove finite-sample bulk motion:
("internal centering")

$$\mathbf{r}_i \to \mathbf{r}_i - \mathbf{R}_{CM} \qquad \mathbf{v}_i \to \mathbf{v}_i - \mathbf{V}_{CM}$$

$$\mathrm{R}_{CM} \equiv \frac{1}{M}\sum_{i=1}^{N} m_i \mathbf{r}_i \qquad \mathrm{V}_{CM} \equiv \frac{1}{M}\sum_{i=1}^{N} m_i \mathbf{v}_i$$

Astrophysical N Body Problems

Intermediate Challenge

Finite-N sampling of the King model

Starting from the parameters (for concreteness, choose $W_0 = 7$, M = 1.2 x 10^6 solar masses, $r_h = 2$ pc)

$$W_0 = \text{dimensionless shape}, \qquad M = \text{mass scale}, \qquad r_h = \text{length scale}$$

We define the function

$$g(w) \equiv e^w \mathrm{erf}(\sqrt{w}) - 2\sqrt{\frac{w}{\pi}}\left(1 + \frac{2}{3}w\right), \qquad \text{for } w \geq 0$$

1) Solve the King equations

For unknown functions

$W(x)$ and $I(x) = \int_0^x y^2 \frac{g(W(y))}{g_0} dy$

solve the system

$$\frac{dW}{dx} = -\frac{9I}{x^2} \qquad \frac{dI}{dx} = x^2 \frac{g(W)}{g_0}$$

with boundary conditions

$$W(0) = W_0, \quad I(0) = 0$$

2) Find the parameters

From the solution of 1) above, find the smallest zero of the W function

x_t where $W(x_t) = 0$

and then define

I_t such that $I_t = I(x_t)$

x_h where $I(x_h) = \frac{1}{2} I_t$

3) Define parameters

$$r_0 \equiv r_h / x_h$$
$$r \equiv r_0 \, x$$
$$r_t \equiv r_0 \, x_t$$

4) Obtain density

Define the density

$$\rho(r) = \rho_0 \frac{g(W(r/r_0))}{g(W_0)}$$

where $\rho_0 \equiv \frac{M}{4\pi r_0^3 I_t}$

and the velocity scale

$$\sigma_K^2 \equiv \frac{GM}{9 r_0 I_t}$$

Astrophysical N Body Problems

Intermediate Challenge

- **Informal Goal 1: Globular Cluster Simulations.** Simulate the relaxation of a globular star cluster (of $N=10^6$ stars with equal mass) of total mass $M = 1.2 \times 10^6$ solar masses.

Pipeline:

1) Generate the initial conditions according to the King model
2) **Evolve the globular cluster forward in time for 12 billion years**
3) **Report the values of the target quantities at the required times**

[1] D. Guszejnov et al., "Effects of the environment on the multiplicity properties of stars in the STARFORGE simulations," MNRAS 518, 4693–4712 (2023), §2.2.
[2] Baumgardt, Hut & Heggie (2002), "Long-term evolution of isolated N-body systems," MNRAS 336, 1069–1081
[3] Casertano & Hut (1985), "Core radius and density measurements in N-body experiments: Connections with theoretical and observational definitions," ApJ 298, 80–94

Evolution

Evolve the initial data for the cluster forward in time for a dimensionless time $\tau = 20$ where

$$\tau \equiv t/t_{rh,0}$$

is defined by the initial half-mass relaxation time $t_{rh,0} \equiv \dfrac{0.138 N^{1/2} r_{h,0}^{3/2}}{\sqrt{Gm}\ln(0.4N)}$

Extract Target Quantities

For times $t_n = n\,\Delta t$ for $\Delta t \equiv 0.05$ compute the Lagrangian radii

- **Relaxation:** calculate the Lagrangian radii (see e.g. [1],[2],[3]) for fractions f = 1/100, 5/100, 10/100, 50/100, 90/100.
  - Compute the *COM-reduced system-particle set* and the bound cluster. Call M_b its total mass.
  - Calculate the density center of the cluster
  - For each bound system particle, calculate its distance from the density center
  - Sort the particles so that $R_1 \le R_2 \le \ldots \le R_N$.
  - For a given mass fraction $0 < f < 1$, define the radius $r_f \equiv \inf\left\{r \ :\ M_b(r) \ge f\, M_b\right\}$ where $M_b(r)$ is the mass included in a radius r around the density center, and M_b is the total bound mass.

Astrophysical N Body Problems

Intermediate Challenge

- **Informal Goal 1: Globular Cluster Simulations.** Simulate the relaxation of a globular star cluster (of $N=10^6$ stars with equal mass) of total mass $M = 1.2 \times 10^6$ solar masses.

Pipeline:

1) Generate the initial conditions according to the King model
2) Evolve the globular cluster forward in time for 12 billion years
3) **Report the values of the target quantities at the required times**

Benchmark against: PeTar extrapolated performance.

From [1] we can take PeTar performance for similar simulations and extrapolate. The published figures use the Hénon time (standard N-body time unit) and study PeTar wall time which runs on 640 CPU cores. The $N=10^6$ particle simulations require approximately 63 seconds for the model without primal binaries. By converting into 20 half-mass relaxation time units (for the formula used, see e.g. [1])

$$t_{rh,0} = \frac{0.138N}{\ln(0.4N)} \left(\frac{r_h}{r_v}\right)^{3/2} t_{NB}$$

For our $W_0 = 7$ King model and $N=10^6$, we get $r_h/r_v \approx 0.8113$ and

$$t_{rh,0} \approx 7.82 \times 10^3 t_{NB}, \qquad 20 t_{rh,0} \approx 1.56 \times 10^5 t_{NB}$$

$$T_{\text{walltime}} = 1.56 \times 10^5 \times 63\, s \approx 2736\, \text{hours}$$

and a wall time

which means

$$T_{\text{normalized}} = 2736 \times 640 \approx 1.75 \times 10^6\, \text{core hours}$$

Astrophysical N Body Problems

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) **Generate the initial conditions for the host galaxy according to the host-galaxy model**
2) Generate the initial conditions for the nuclear star cluster
3) Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years
4) Report the values of the target quantities at the required times

Host Galaxy Model

We choose a tapered Dehnen model for the generation of the **positions** of particles that represent collisionless entities (stellar host and dark matter) that make up the host galaxy where the nuclear star cluster is located. For total mass *M*, radial scale *a*, and inner density slope γ

$$\rho_{M,a,\gamma}(r) = \frac{(3-\gamma)M}{4\pi}\frac{a}{r^{\gamma}(r+a)^{4-\gamma}}$$

The profile is truncated by the function

$$T_{r_1,r_2}(r) = \begin{cases} 1 & \text{if } r \le r_1 \\ 1 - 10s^3 + 15s^4 - 6s^5 & \text{if } r_1 \le r < r_2 \\ 0 & \text{if } r_2 \le r \end{cases} \qquad s \equiv \frac{r - r_1}{r_2 - r_1}$$

So that the empirical density of the model is given by

$$\rho_{\text{eff}}(r) \propto \rho_{M,a,\gamma}(r) T_{r_1,r_2}(r)$$

Where the proportionality constant is fixed by the condition

$$4\pi \int_0^{\infty} \rho_{\text{eff}}(r)\, r^2 dr = M$$

Astrophysical N Body Problems

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) **Generate the initial conditions for the host galaxy according to the host-galaxy model**
2) Generate the initial conditions for the nuclear star cluster
3) Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years
4) Report the values of the target quantities at the required times

Host Galaxy Model

For each of the two components of the host galaxy (*i.e.* for the host stars and the dark matter) we fix a truncated Dehnen profile given by parameters

	Total mass M	Radial scale a	r_1	r_2	softening
Host stars	5×10^8	500	8000	10000	0.5
Dark matter	10^{10}	3000	30000	40000	2.0

and we **generate 2^{22} particles** each, where the position of each particle is sampled independently at random, with spherical coordinates given by

$$r \sim \frac{4\pi r^2}{M}\rho_{\text{eff}}(r), \qquad \cos(\theta) \sim \text{Unif}(-1,1), \qquad \phi \sim \text{Unif}(0,2\pi)$$

(masses are given in units of solar mass, distances in parsecs)

Astrophysical N Body Problems

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) **Generate the initial conditions for the host galaxy according to the host-galaxy model**
2) Generate the initial conditions for the nuclear star cluster
3) Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years
4) Report the values of the target quantities at the required times

[1] Eddington, A. S., The distribution of stars in globular clusters, MNRAS, 76, 572 (1916)
[2] Binney, J. & Tremaine, S. (2008), Galactic Dynamics, 2nd ed.

Host Galaxy Model

For the initial **velocities** of the host galaxy particles, we use the standard Eddington formula construction (see e.g. [1][2])

$$f_\alpha(E) = \frac{1}{\sqrt{8}\pi^2}\left[\int_0^E \frac{d^2\rho_\alpha}{d\Psi^2}\frac{1}{\sqrt{E-\Psi}}d\Psi + \frac{1}{\sqrt{E}}\frac{d\rho_\alpha}{d\Psi}\Big|_{\Psi=0}\right]$$

with respect to the total potential of the host (which includes contributions from both the host stars and the dark matter "particles"). One arrives at the distribution for the velocity modulo $v = ||\mathbf{v}||$ conditioned on the radial distance

$$p_\alpha(v\,|\,r) = 4\pi v^2 \frac{f_\alpha\left(\Psi(r) - v^2/2\right)}{\rho_\alpha(r)}$$

The other components of the velocity vector are generated isotropically.

Astrophysical N Body Problems

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) Generate the initial conditions for the host galaxy according to the host-galaxy model
2) **Generate the initial conditions for the nuclear star cluster**
3) Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years
4) Report the values of the target quantities at the required times

Star Cluster Model

The star cluster is composed of 12 globular clusters, each generated according to the dynamical King model (see "Intermediate Challenge" section).

For all star clusters we set $W_0 = 7$ and $r_h = 2$ pc, and for index $k=1,\ldots,12$ a number of one-solar-mass stars given in the table to the right.

Each cluster is placed in an orbit inside of the host. We specify the orbit of the center of mass by computing the position X_k and velocity V_k.

$$\mathbf{X}_k \equiv r_{a,k}\,\mathbf{e}_{r,k}$$

where the directions $\mathbf{e}_{r,k}$ lie on a Fibonacci sphere (see next slide) and $r_{a,k}$ is the apocenter

$$\mathbf{V}_k \equiv v_{t,k}\,\mathbf{e}_{t,k}$$

where the tangential directions $\mathbf{e}_{t,k}$ are chosen to be orthogonal to the Fibonacci directions $\mathbf{e}_{r,k}$ (see next slide) and the moduli are defined by

$$v_{t,k} \equiv L_k / r_{a,k} \qquad L_k^2 = 2\frac{\Phi_0(r_{a,k}) - \Phi_0(r_{p,k})}{r_{p,k}^{-2} - r_{a,k}^{-2}}$$

$$\Phi_0 = -4\pi G\left[\frac{1}{r}\int_0^r \rho_{\rm tot}(s)s^2 ds + \int_r^{r_{\rm max}} \rho_{\rm tot}(s) s ds\right]$$

where $\rho_{\rm tot}$ is the total density of the host.

Index	Num. of stars	Apocenter r_a	Pericenter r_p
1	1,200,000	400	120
2	1,000,000	500	250
3	800,000	350	175
4	650,000	600	300
5	550,000	300	90
6	450,000	450	225
7	350,000	250	125
8	300,000	550	110
9	250,000	200	100
10	220,000	700	350
11	180,000	150	75
12	150,000	800	400

Astrophysical N Body Problems

INTELLIGENT SYSTEMS DIVISION

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) Generate the initial conditions for the host galaxy according to the host-galaxy model
2) **Generate the initial conditions for the nuclear star cluster**
3) Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years
4) Report the values of the target quantities at the required times

Star Cluster Model

Fibonacci Sphere

$$\mathbf{e}_{r,k}, \qquad k = 0, 1, \ldots 11$$

$$\varphi_g \equiv (1+\sqrt{5})/2$$
$$\mathrm{frac}(x) \equiv x - \lfloor x \rfloor$$

For each value of k, define

$$\phi_k = 2\pi\,\mathrm{frac}(k/\varphi_g)$$

$$z_k \equiv 1 - 2\frac{k+1/2}{12}$$

$$\mathbf{e}_{r,k} = \begin{bmatrix} \sqrt{1-z_k^2}\cos\phi_k \\ \sqrt{1-z_k^2}\sin\phi_k \\ z_k \end{bmatrix}$$

Tangential vectors

For each value of k, define

$$\mathbf{a} = \begin{cases} \mathbf{a}_z \equiv (0,0,1) \in \mathbb{R}^3 & \text{if } |\mathbf{a}_z \cdot \mathbf{e}_{r,k}| \le 9/10 \\ \mathbf{a}_y \equiv (0,1,0) \in \mathbb{R}^3 & \text{otherwise} \end{cases}$$

$$\mathbf{e}_{1,k} \equiv \frac{\mathbf{a} \times \mathbf{e}_{r,k}}{\|\mathbf{a} \times \mathbf{e}_{r,k}\|}$$

$$\mathbf{e}_{2,k} \equiv \mathbf{e}_{r,k} \times \mathbf{e}_{1,k}$$

$$\psi_k \equiv 2\pi\,\mathrm{frac}(k\sqrt{2})$$

$$\mathbf{e}_{t,k} = \cos(\psi_k)\mathbf{e}_{1,k} + \sin(\psi_k)\mathbf{e}_{2,k}$$

Cluster insertion

For each cluster *k*, and each particle *i* in the cluster, with *internally centered* position and velocity

$$\mathbf{r}_i^{\mathrm{int}}, \quad \mathbf{v}_i^{\mathrm{int}}$$

we apply a translation

$$\mathbf{r}_i \equiv \mathbf{r}_i^{\mathrm{int}} + \mathbf{X}_k$$

and a boost

$$\mathbf{v}_i \equiv \mathbf{v}_i^{\mathrm{int}} + \mathbf{V}_k$$

This completes the specification of the initial configuration for the nuclear star cluster

Astrophysical N Body Problems

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) Generate the initial conditions for the host galaxy according to the host-galaxy model
2) Generate the initial conditions for the nuclear star cluster
3) **Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years**
4) **Report the values of the target quantities at the required times**

Evolution

Evolve the initial data for the cluster forward in time for 12 Gyr.

Compute

Every 20 Myr, compute and report the following quantities.
For the orbital-decay history:

- The galaxy center $R_{center}(t)$ using the shrinking-sphere algorithm applied to the system of host-particles only
- For each star cluster k=1…,12, the center of the bound remnant of the cluster $R_k(t)$ through an iterative gravitational ubinding procedure (*e.g.* the one indicated in [1]). Report the galactocentric distance of the bound remnant

$$r_k(t) \equiv \left| \mathrm{R}_k(t) - \mathrm{R}_{\mathrm{center}}(t) \right|$$

The (spatially-discretized) radial density profile $\rho_{\mathrm{NSC}}(r)$ of the nuclear star cluster:

- For binning radii $r_j^{(b)} = 10^{-1+j/10}$ pc (logarithmic binning) with j = 0, 1, …, 31 define the j-th spherical shell around $R_{center}(t)$ as the points at distance $r_j^{(b)} \le r \le r_{j+1}^{(b)}$ from $R_{center}(t)$. The volume of the shell is given by

$$V_j = \frac{4}{3}\pi\left({r_j^{(b)}}^3 - {r_{j+1}^{(b)}}^3 \right)$$

- and the total mass of the stars (of all the 12 clusters) included in the j-th shell is indicated with M_j. Then

$$\rho_{\mathrm{NSC}}(r_j) \equiv M_j / V_j$$

[1] Han, J., Jing, Y.P., Wang, H., Wang, W. Resolving subhaloes' lives with the hierarchical bound-tracing algorithm, MNRAS 427, 2437-2449 (Appendix A)

Astrophysical N Body Problems

Grand Challenge

- **Informal Goal 2: Nuclear Star Cluster Simulations.** Simulate a 12 Gyr evolution of a nuclear star cluster (10^7 particles) made of ~10 clusters of 10^6 compact objects immersed in a background of dark matter and lumps of collisionless particles ("softened interactions")

Pipeline:

1) Generate the initial conditions for the host galaxy according to the host-galaxy model
2) Generate the initial conditions for the nuclear star cluster
3) Evolve the host galaxy + nuclear star cluster forward in time for 12 billion years
4) **Report the values of the target quantities at the required times**

Benchmark against: PeTar extrapolated performance (order of magnitude).

For the grand challenge the system contains 12 collisional globular clusters embedded in a live host, with a total particle number of order $N \sim 10^7$. PeTar benchmarks for $N=10^7$ require approximately 1100–1350 seconds per standard N-body time unit on 960 CPU cores. To convert the required 12 Gyr evolution time we use the internal time unit of a representative cluster. Taking

$$M_{\rm GC} = 1.2 \times 10^6\, M_\odot, \qquad r_h = 3\,{\rm pc}, \qquad W_0 = 7,$$

And using $r_h/r_v \approx 0.8113$ as before, we obtain

$$r_v = \frac{r_h}{0.8113} \approx 3.70\,{\rm pc}.$$

The corresponding physical N-body time-unit is

$$t_{\rm NB} = \left(\frac{r_v^3}{GM_{\rm GC}}\right)^{1/2} \approx 0.097\,{\rm Myr}.$$

Consequently we arrive at an estimated

$$T_{\rm walltime} \approx 1.24 \times 10^5 \times 1344\,{\rm s} \approx 4.6 \times 10^4\,{\rm hours}.$$

run in parallel on 960 cores. Normalizing for this gives

$$T_{\rm normalized} \approx 4.6 \times 10^4\,{\rm hours} \times 960 \approx 4.4 \times 10^7\,{\rm core\ hours}.$$

Allowing an order-unity penalty for the 12 distinct dense clusters, the live host, and the associated neglected complexity gives the order-of-magnitude estimate

$$\boxed{T_{\rm NSC} \approx 10^8\,{\rm core\ hours}}$$



Astrophysical N Body Problems



INTELLIGENT SYSTEMS DIVISION

Whom did we talk to?

- John Fregeau (IDA Center for Computing Sciences)
- Daniel Atallah (UNC Chapel Hill & Northwestern University)



Acknowledgements



INTELLIGENT SYSTEMS DIVISION

This work was supported by NASA's Office of the Chief Science Data Officer.

Gianni Mossi and Erik Gustafson are KBR, Inc. employees working under the Prime Contract No. 80ARC020D0010 with the NASA Ames Research Center.

The authors acknowledge the use of NASA LLM tools for help with background research, problem identification and specification.