The Architecture of the LISA Science Analysis

Final report for a Keck Institute for Space Studies workshop, January 16–19, 2018

- Rachel Akeson - California Institute of Technology
- Anne Archibald - Universiteit van Amsterdam
- Stanislav Babak - Université Paris Diderot
- Katelyn Breivik - Northwestern University
- C. Titus Brown - University of California, Davis
- Alvin Chua - Jet Propulsion Laboratory
- Neil Cornish - Montana State University
- Curt Cutler - Jet Propulsion Laboratory
- Scott Davidoff - Jet Propulsion Laboratory
- Francois Foucart - University of New Hampshire
- Chad Galley - Jet Propulsion Laboratory
- Lawrence Kidder - Cornell University
- Prayush Kumar - Cornell University
- Astrid Lamberts - California Institute of Technology
- Geoffrey Lovelace - California State University, Fullerton
- Ashish Mahabal - California Institute of Technology
- Christine Corbett Moran - California Institute of Technology
- Laura Nuttall - Cardiff University
- Maria Okounkova - California Institute of Technology
- Travis Robson - Montana State University
- Mark Scheel - California Institute of Technology
- Deirdre Shoemaker - Georgia Institute of Technology
- Stephen Taylor - California Institute of Technology
- Saul Teukolsky (co-organizer) - California Institute of Technology
- Massimo Tinto - University of California, San Diego
- Michele Vallisneri (co-organizer) - Jet Propulsion Laboratory
- Vijay Varma - California Institute of Technology
- Alberto Vecchio - University of Birmingham

Overview

Our goals

The space-based gravitational-wave (*henceforth, GW!*) observatory **LISA** will offer **unparalleled science returns**, including a view of massive black-hole mergers to high redshifts, precision tests of general relativity and black-hole structure, a census of thousands of compact binaries in the Galaxy, and the possibility of detecting stochastic signals from the early Universe.

While the Mock LISA Data Challenges (2006–2011) gave us confidence that LISA will be able to fulfill its scientific potential, we still have a rather incomplete idea of what the **end-to-end LISA science analysis** should look like. The task at hand is substantial.

- Our data reduction needs to ensure the phase coherence of GW measurements across data gaps and instrument glitches over multiple years.
- Our waveform models need to reach part-in-10^5 accuracy (to achieve full performance when testing general relativity), with sufficient computational efficiency to sample parameter space broadly.
- Our algorithms need to resolve thousands of individual sources of different types and strengths, all of them superimposed in the same multi-year dataset, and simultaneously characterize the underlying noise-like stochastic background.
- Our catalogs need to represent the complex and high-dimensional joint distributions of estimated source parameters for all sources.

It is tempting to assume that current algorithms and prototype codes will scale up to this challenge, thanks to the greatly increased computational power that will become available by LISA's launch in the early 2030s. In reality, harnessing that power will require very different **methods, adapted to future high-performance computational architectures** that we can only glimpse now. Thus, we need to begin our exploration at this time, seeking inspiration from other disciplines (e.g., big-data processing, computational biology, the most advanced applications in astroinformatics), and learning to pose the same physical questions in different, future-proof ways—or even daring to imagine questions that will be tractable only with future machines.

The broad objective of this study program was to imagine **how evolved or rethought data-analysis algorithms and source-modeling codes will solve the LISA science analysis on the computers of the future**, with the hope of guiding LISA science and data-analysis research and development in the years to come.

Our study

The study began with a short course open to the Caltech community, with video available online,[1] with sections on LISA science (by Curt Cutler), the numerical modeling of LISA sources (by Deirdre Shoemaker), and LISA data analysis (by Stas Babak).

[1] "The Architecture of Lisa Science Analysis Short Courses (Video)," 2018, `https://kiss.caltech.edu/short_courses/LISA.html`.

The study participants then convened for three sessions consisting of tutorials, discussions, and breakout groups:

- The session on **data-analysis problems and approaches** included talks on stochastic techniques, machine/deep learning, the LISA Data Processing Center being prototyped in Paris, modern astronomical archives, as well as a panel on social aspects of writing scientific software. Breakout groups convened to discuss the global-fit problem, data products, and the detection of unexpected signals.
- The session on **numerical modeling** included talks on new computational paradigms, parameter estimation, hydrodynamics in numerical relativity, and reduced-order-modeling waveforms, as well as a panel on EMRI waveforms, which are especially challenging analytically and numerically. Breakout groups convened to discuss new ideas for waveform-accuracy estimates, searches and parameter estimation, and the astronomy-community interface.
- The session on **computation** included talks on data-intensive architectures, the scientific use of GPUs, "live" data products, and advanced visualization techniques.

The study participants then discussed follow up work and the organization of report-writing. Slides from all presentations are available on the KISS website.[2]

[2] "The Architecture of Lisa Science Analysis Presentations (Pdf)," 2018, `http://kiss.caltech.edu/workshops/LISA/LISA.html#Presentations`.

This report

This report does not aim at providing an exhaustive retelling of the many interesting discussions that took place at the study; rather, it focuses on the *stories and questions* that we consider likely to generate the most interesting and urgent directions of research.

The report includes three main sections. The **complexity and richness of LISA science** covers the requirements of efficiency and accuracy imposed on waveforms by LISA's remarkable sensitivity, as well as the many use astronomical cases for the LISA data, which will need to be fulfilled with *ad hoc* data products. **Hard problems** focuses on the difficulty of detecting and characterizing many thousands of superimposed sources; of compiling and serving a source catalog that can fully represent the LISA findings and enable insightful science; and of building confidence in the detection of previously unknown sources that inhabit LISA's proverbial *discovery space*. **New tools** is concerned with the mostly untapped power of deep learning for GW analysis applications, and with emerging tools (such as computational notebooks) that promise agile, efficient, and reproducible access to astronomical archives and to all computations that can be performed with them.

An executive summary, such as would be customary for a report such as this, could not do justice to the explorational nature of this study, and to the many strands that we touched briefly and that yet may figure in the ultimate solution to the LISA science analysis. We offer instead an **unordered list of takeaways** that one of us (M.V.) compiled shortly after the study. As you will see, some of these points are represented more generously than others in the rest of this report, so if you do read through, you will come out with your own different list. Nevertheless:

- There are many very deep research questions that stand between us and the solution to the LISA science analysis. It will take time to answer then, and we will need to think carefully and purposefully about them. However, since we need to get going practically, we can begin by **providing infrastructure for experimentation that will later become infrastructure for implementation**. We should be especially careful to remain flexible while we get oriented.

- As for numerical modelers, they need **practical performance and accuracy benchmarks** as soon as possible, so that they can strategize on making LISA-grade waveforms across all source classes. As Saul Teukolsky points out, "This is a billion dollar experiment, and we do not want to sacrifice even the last vestiges of science that can be extracted from it."

- The awesome demonstrations of **deep learning** (DL!) in many domains (and the first tentative applications in GW data analysis) show the potential of this technique, which will only become more important because it is already driving processor development; so our wish to continue riding Moore's law will only be fulfilled if we embrace DL as a data-analysis workhorse, or at least if we recast standard calculations in DL's mathematical language.

- Speaking of Moore's law, much of its ascent in recent times has come in the shape of ever more powerful graphical processing units, with larger and larger collections of elementary compute cores. **GPU programming** used to be a perilous and abstruse land, but thankfully now the right tools (especially compilers, mathematical libraries) are now available to wield all the Gigaflops we are promised.

- We need to **look outside the GW community** and provide diverse venues for engagement of scientists of different kinds who will make the most of the LISA data, and of untapped talent more generally. We can take advantage of "citizen" programmers, but we need virtuous policies and modern practices in organizing scientific software development and in distributing credit fairly.

- The high-level LISA data products will need to serve a diverse community ranging from cosmographers to population modelers, from seekers of counterparts to experts in alternative theories of gravity. We can take inspiration from recent astronomical archives (Fermi, Kepler, SDSS) in **tailoring and diversifying our data products**. However, we should remember that the best internal data representation (for source characterization and data cleaning) may not be same as needed for archiving and for external users. We need help from database professionals.

- There are now powerful **collaborative tools that are fulfilling a vision of open science** and of computations brought to otherwise unwieldy datasets *in the cloud*. In addition, bespoke visualizations can be extremely insightful. While LISA will fly with onboard technology that is necessarily frozen at the beginning of its arc, the same need not be true for its science segment, if we remain flexible and observant.

The complexity and richness of LISA science

LISA needs waveforms!

In discussing **waveform accuracy**, it is useful to distinguish between the intrinsic error in the underlying *models* of physics[3,4] (e.g., a numerical relativity simulation, or a post-Newtonian expansion of Einstein's field equations) and the error incurred in making these models available as *templates*: *hybridization* error when marrying analytical and numerical waveforms; *approximation* error when fitting phenomenological models to simulations; *representation* error when casting waveforms on reduced-order-modeling bases. (One speaks generally of *surrogates* to encompass waveforms that can be evaluated efficiently and that approximate more careful and expensive computations.)

Physical-model errors are addressed by theoretical work; template errors by more careful and potentially costly computation. **Waveform**

[3] Alessandra Buonanno and B. S. Sathyaprakash, "Sources of Gravitational Waves: Theory and Observations," in *General Relativity and Gravitation: A Centennial Perspective*, ed. Abhay Ashtekar et al. (Cambridge University Press, 2015), 287–346, `https://doi.org/10.1017/CBO9781139583961.009`.

[4] Matthew W. Choptuik, Luis Lehner, and Frans Pretorius, "Probing Strong-Field Gravity Through Numerical Simulations," in *General Relativity and Gravitation: A Centennial Perspective*, ed. Abhay Ashtekar et al. (Cambridge University Press, 2015), 361–411, `https://doi.org/10.1017/CBO9781139583961.011`.

accuracy requirements affect both the theoretical inputs and their data-analytic representations, and they will depend strongly on the data-analysis *application* under consideration. A roughly accuracy-ranked list of applications (with the most demanding on top) would be

1. testing general relativity;
2. astrophysical parameter estimation;
3. minimizing residuals after the hierarchical subtraction of loud signals;
4. detection (which will be certainly satisfied when 2 and 3 are)

We should target the application with the most stringent needs. However, how do we define **waveform accuracy requirements** formally for LISA? It is unclear that we have all the mathematical tools and criteria to pose this question profitably (although we have strong starting points[5]). In space-project jargon, **what is the modeling and templating error budget for LISA?** Once we know, we can ask questions such as:

- Are post-Newtonian (PN), effective-one-body (EOB), and numerical-relativistic (NR) binary models sufficiently accurate to extract all science from loud massive black-hole binary inspirals? E.g, are numerical-relativity simulations sufficient to capture the fine details of mergers? How *long* do they need so that the "handoff" from PN/EOB to numerical is sufficiently accurate?
- What about the model representations as phenomenological (Phenom) or reduced-order-modeling (ROM) templates? Can these be improved at the cost of more computation, whether *online* during parameter estimation or in a preliminary stage?

The other important consideration is that we should concentrate on the specific **region of source parameter space to which LISA will be sensitive**: for instance, significant binary eccentricity and precession are challenging to numerical simulations, and they will be more relevant to LISA than they are to ground-based detections.

In terms of work that is happening within the LISA Consortium as we write, it is relevant to ask if the **LISA Data Challenges** (LDC)[6] could be used to answer these questions empirically. (However, the template families used currently in the LDC need to be updated toward the state of the art.) Specifically, we would address the level of waveform error that is tolerable for different parameter-estimation tasks. To do so, we need to designate a waveform family *as truth*, then perform parameter estimation with a different family.

For instance, if the reference waveforms are provided by long(er) numerical-relativity simulations and parameter estimation is per-

[5] Lee Lindblom, Benjamin J. Owen, and Duncan A. Brown, "Model waveform accuracy standards for gravitational wave data analysis" 78 (December 2008): 124020, `https://doi.org/10.1103/PhysRevD.78.124020`.

[6] "LISA Data Challenges," 2018, `https://lisa-ldc.lal.in2p3.fr/ldc`.

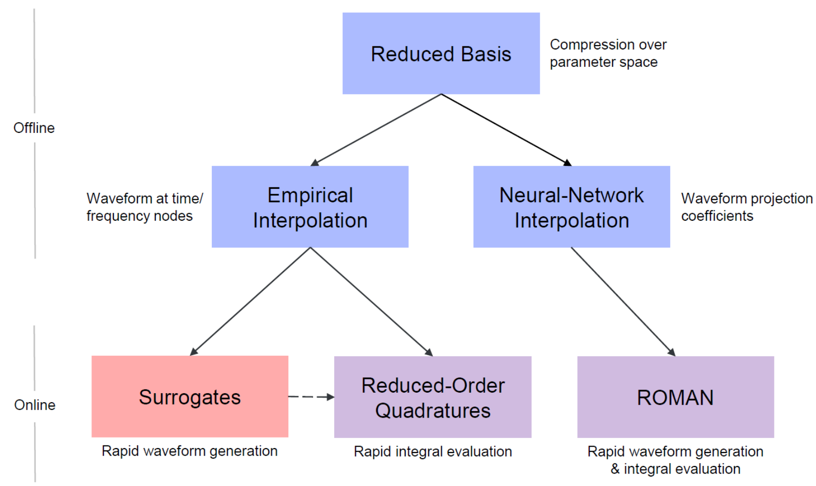


Figure 1: The reduced-order-modeling approach involves the offline construction and interpolation of a reduced-basis for the signal space, in order to greatly accelerate the online use of templates for detection and parameter estimation. Credit: C. Galley & A. Chua

formed with Phenom surrogates, we would be testing the accuracy of the Phenom template implementation (which can also be done directly[7]). Conversely, we can use a surrogate family as reference and perform parameter estimation with a modified surrogate that includes *plausible* phasing/amplitude errors across parameter space (e.g., as obtained by comparing numerical-grid refinements): doing so would exemplify the parameter-estimation consequences of mismodeling error and potentially lead to quantitative accuracy criteria for models. In doing such tests, we would:

- where possible, seed numerical investigations with analytical results, such as Cutler and Vallisneri's analysis of parameter-estimation bias,[8] formulated for PN waveforms;
- make sure the analysis is sufficiently general—we should be able to apply the same methods to all waveform families using compatible definitions of intrinsic errors;
- perform the study up to the highest expected signal-to-noise ratios (for binary black-hole coalescences, 10^4).

Once accuracy requirements are understood, the modeling community can set out to fulfill them. In the case of **numerical relativity**, it is unclear that current paradigms (e.g., domain-based parallelism) will suffice, even when coupled with the continued growth of computing power. Indeed, the movement toward sharing power among many compute cores raises many problems that limit performance: communication bandwidth and latency, memory contention, synchronization bottlenecks.

Recent efforts to reformulate the numerical solution of Einstein's equation in terms of **task-based parallelism**[9] aim to address these

[7] Prayush Kumar et al., "Accuracy of binary black hole waveform models for aligned-spin binaries" 93 (May 2016): 104050, `https://doi.org/10.1103/PhysRevD.93.104050`.

[8] C. Cutler and M. Vallisneri, "LISA detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms" 76, no. 10 (November 2007): 104018, `https://doi.org/10.1103/PhysRevD.76.104018`.

[9] Lawrence E. Kidder et al., "SpECTRE: A task-based discontinuous Galerkin code for relativistic astrophysics," *Journal of Computational Physics* 335 (April 2017): 84–114, `https://doi.org/10.1016/j.jcp.2016.12.059`.

limits, dividing computation into tasks within a dependency tree. The tasks correspond (roughly) to evaluating the functions required to advance the evolution equations; the parameters needed by the functions define their mutual dependencies. A smart scheduler assigns tasks that are ready to compute cores, managing asynchronous information transfer, and (crucially) handling load balancing and *task stealing*.

It may be possible to improve the *status quo* for other aspects of NR modeling; for instance, by applying machine learning to automation in the "hand holding" of simulation runs, and possibly replacing certain computational bottlenecks with appropriately trained neural networks (an approach used in particle physics for expensive Monte Carlo likelihoods).

We would hope that the eventual modeling and representation errors would be so small that they can simply be ignored in data analysis. If this is not the case, we must formulate a robust method to propagate a characterization of the errors into data-analysis pipelines, where the corresponding degrees of freedom can be treated as nuisance parameters and marginalized over.[10]

[10] C. J. Moore et al., "Improving gravitational-wave parameter estimation using Gaussian process regression" 93, no. 6 (March 2016): 064001, https://doi.org/10.1103/PhysRevD.93.064001.

Beyond the general schemes outlined above, there are specific considerations to be made for each class of LISA sources. For **massive and stellar-mass black-hole binaries with comparable mass ratios**, we already have banks of numerical-relativity simulations that can provide fiducial injections, although we need more simulations of eccentric binaries; PN waveforms also need to include eccentricity organically. To orient effort in both domains, we need to understand astrophysical expectations for eccentric systems. NR simulations are currently able to comfortably handle eccentricity and spin precession, for mass ratios of up to around 1/20 and spin magnitudes of up to around 0.8. An extension to higher spins would require improved initial data, while going to more pronounced mass ratios will likely be better achieved through black-hole perturbation theory.

Intermediate–mass-ratio black-hole binaries (IMRIs) are very problematic already at the modeling stage. They would require pushing perturbative analytical expansions to third order in the mass ratio – very difficult – although Le Tiec's "q to ν" trick may help.[11] At worst, we may need to investigate new perturbative schemes for the selfforce. Looking at currently available models and simulations, we could test EOB waveforms at high mass ratio with numerical-relativity runs, which however would be very expensive.

[11] Alexandre Le Tiec et al., "Periastron Advance in Spinning Black Hole Binaries: Gravitational Self-Force from Numerical Relativity," *Phys. Rev. D* 88, no. 12 (December 2013): 124027, https://doi.org/10.1103/PhysRevD.88.124027.

The modeling of **extreme–mass-ratio inspirals** (EMRIs) for parameter estimation requires up to second-order self-force corrections to a Kerr geodesic orbit, or specifically their average dissipative effects. (The complete second-order contribution would be useful for IMRIs.)

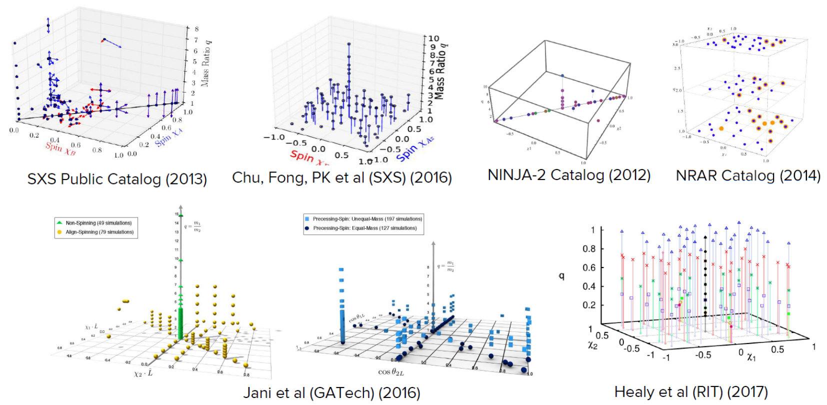
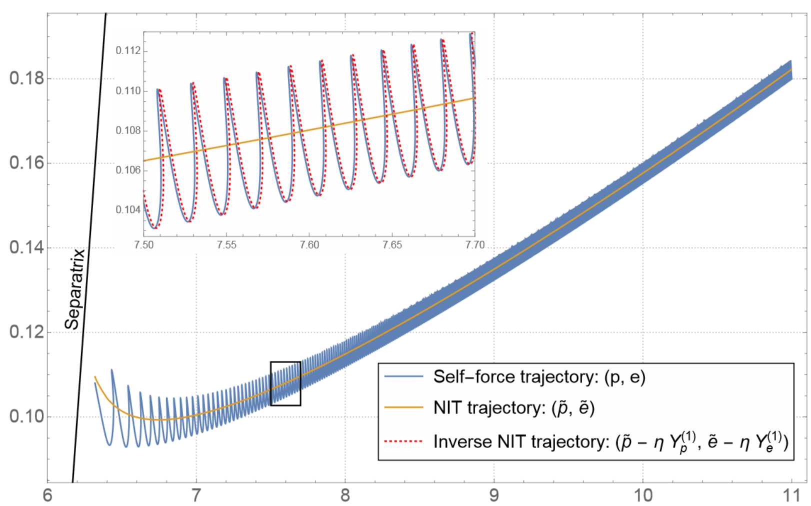


Figure 2: Parameter distributions of current NR-simulation catalogs. Credit. P. Kumar

Waveforms produced directly from self-force calculations would be too computationally expensive for use in parameter estimation algorithms, and must be approximated by a surrogate with very minimal loss of accuracy. The fast near-identity-transform method of van de Meent and Warburton[12] is promising for this purpose, although it depends on the construction of an accurate interpolant for the self-force when the calculations become available. For EMRI detection, fast **kludge waveforms**[13,14,15] that utilize various approximations have been shown to be adequate,[16] but the presently available ones should still be improved further in terms of both their speed and accuracy.



[12] Maarten van de Meent and Niels Warburton, "Fast Self-Forced Inspirals," *Classical and Quantum Gravity* 35, no. 14 (June 2018): 144003, https://doi.org/10.1088/1361-6382/aac8ce.

[13] Leor Barack and Curt Cutler, "LISA Capture Sources: Approximate Waveforms, Signal-to-Noise Ratios, and Parameter Estimation Accuracy," *Phys. Rev. D* 69, no. 8 (April 2004): 082005, https://doi.org/10.1103/PhysRevD.69.082005.

[14] Stanislav Babak et al., "'Kludge' Gravitational Waveforms for a Test-Body Orbiting a Kerr Black Hole," *Phys. Rev. D* 75, no. 2 (January 2007): 024005, https://doi.org/10.1103/PhysRevD.75.024005.

[15] Alvin J K Chua and Jonathan R Gair, "Improved Analytic Extreme-Mass-Ratio Inspiral Model for Scoping Out eLISA Data Analysis," *Classical and Quantum Gravity* 32, no. 23 (November 2015): 232002, https://doi.org/10.1088/0264-9381/32/23/232002.

[16] Alvin J. K. Chua, Christopher J. Moore, and Jonathan R. Gair, "Augmented Kludge Waveforms for Detecting Extreme-Mass-Ratio Inspirals," *Phys. Rev. D* 96, no. 4 (August 2017): 044005, https://doi.org/10.1103/PhysRevD.96.044005.

LISA makes catalogs!

The LISA observations will seed many studies, investigations, and followups across the astronomical community. Because the LISA archive will be the first of its kind (notwithstanding many similarities with source catalogs for mature ground-based GW detectors) it is important to **reach out to astronomers** to teach them the peculiarities of LISA data, and to learn about their requirements and use cases. KISS study participants outlined several avenues to do so, including holding regular workshops at professional conferences; providing worked-out data-analysis examples (e.g., as computational notebooks, see below); structuring the LISA archive to accept contributions from outside the project. This outreach could extend to the public, through a carefully designed *citizen science project* similar to Gravity Spy.[17]

[17] "Gravity Spy," 2018, https://www.zooniverse.org/projects/zooniverse/gravity-spy/about/results.

Thus, the LISA data archive will need to fulfill **many different use cases** in the astronomical community. The following are five examples (certainly not exhaustive) which we chose to probe the range of possible data products, and their connections to other astronomical catalogs.

- **A census of massive black holes throughout the Universe.** LISA will observe massive black-hole mergers (MBHs) throughout the entire history of the Universe, and EMRI events out to $z \sim 1$. With several tens of events per year and a mission length in the range 5–10 years, LISA will build up a catalog of potentially hundreds of these events that can be used to form a census of MBH properties over cosmic time. In particular, LISA will be sensitive to MBH mergers with constituent BHs in the range 10^4–$10^7 M_\odot$, which will still harbor a signature of seed formation at high redshifts. This catalog will also shed light on the impact of metallicity on MBH formation, MBH accretion efficiencies, and MBH accretion geometries. Models of hierarchical structure formation that have been constructed for consistency with observed AGN properties can agree with a wide variety of BH seed conditions,[18,19],[20]. Hence, LISA offers an unparalleled opportunity to peer into seed formation and accretion models of BHs across throughout the history of the Universe. **Required data products**: Catalog of events, with independent posterior samples for each. The recovered event parameters should include luminosity distance, component spins, and component masses.

- **Mapping Galactic structure and formation using white-dwarf binaries.** The population of white dwarfs (WDs) residing in the Milky Way is an important tracer of the Galactic gravitational potential and formation history.[21] While the most interesting probes into Galactic structure and evolution history reside with the distant populations in the Milky Way halo, Gaia will observe the WD population with 100% completeness out to only 100 pc.[22] This is largely due to the combined restrictions leading from Gaia's limit-

[18] J. R. Gair, C. Tang, and M. Volonteri, "LISA extreme-mass-ratio inspiral events as probes of the black hole mass function" 81, no. 10 (May 2010): 104014, https://doi.org/10.1103/PhysRevD.81.104014.

[19] J. R. Gair et al., "Constraining properties of the black hole population using LISA," *Classical and Quantum Gravity* 28, no. 9 (May 2011): 094018, https://doi.org/10.1088/0264-9381/28/9/094018.

[20] A. Sesana et al., "Reconstructing the massive black hole cosmic history through gravitational waves" 83, no. 4 (February 2011): 044036, https://doi.org/10.1103/PhysRevD.83.044036.

[21] J. S. Kalirai, "The age of the Milky Way inner halo" 486 (June 2012): 90–92, https://doi.org/10.1038/nature11062.

[22] J. M. Carrasco et al., "Gaia photometry for white dwarfs" 565 (May 2014): A11, https://doi.org/10.1051/0004-6361/201220596.

ing magnitude (20) and the inherent dim nature of WDs. Studies have shown, however, that LISA will observe gravitational waves from hundreds to thousands of resolved double WDs (DWDs) in the outer regions of the Galaxy with high signal to noise. Included in this large population is the subset of several dozens of detached and accreting DWDs observable by *both* LISA and Gaia,[23,24]. The combined observations of WDs from Gaia in the regions out to 100 pc and LISA in the regions beyond could provide an unprecedented view into the structure and formation of the Milky Way. **Required data products**: Catalog of parameters of resolved WDs, including three dimensional positions and masses.

- **Understanding unknown GW signals.** Although there are a number of sources LISA will definitely observe (such as massive black hole mergers and white dwarf binaries), the mission may also discover completely unexpected signals. (See also the section dedicated to such detections in *Hard problems* below.) In order to rule out instrumental origins there will have to be a standard (and if possible) automated checklist of investigations to conduct, such as checking the time delay between the excitations felt at each spacecraft, as well as auxiliary sensors data. Once an astrophysical origin can be established, as much information as possible concerning the unknown signal will need to be appear in the LISA catalog, for follow-up study and comparisons with electromagnetic catalogs. This is a exciting opportunity for scientists across many disciplines to work together to understand a new source of GWs. Lessons on how to investigate this form of signal can certainly be learnt from other areas of astrophysics, such as fast radio bursts. **Required data products**: Strain data containing the signal; as much information about the signal as possible, such as sky localization, length of the signal, peak strain amplitude, etc.

- **Episodic accretion in white dwarf binary populations.** Cataclysmic variables (CVs) and AM Canum Venaticorum stars (AM CVn) will comprise a large subset of the resolved compact binaries observed by LISA in the Milky Way[25,26]. Based on EM observations, several of these systems (*verification binaries*) are known to emit GWs detectable by LISA.[27] Observing AM CVns in outburst leads to a wide range of opportunities to witness astrophysical processes of great interest, including accretion disk spiral density waves, quasi-period oscillations, and the launching of jets[28,29,30,31]. AM CVn outbursts have been shown to have orbital-period–dependent brightness variations, recurrence times, and durations, with shorter orbital periods leading to smaller brightness variations and shorter durations and recurrence times.[32] While

[23] V. Korol et al., "Prospects for detection of detached double white dwarf binaries with Gaia, LSST and LISA" 470 (September 2017): 1894–1910, `https://doi.org/10.1093/mnras/stx1285`.

[24] K. Kremer et al., "Accreting Double White Dwarf Binaries: Implications for LISA" 846 (September 2017): 95, `https://doi.org/10.3847/1538-4357/aa8557`.

[25] G. Nelemans, "AM CVn stars," in *The Astrophysics of Cataclysmic Variables and Related Objects*, ed. J.-M. Hameury and J.-P. Lasota, vol. 330, Astronomical Society of the Pacific Conference Series, 2005, 27..330. . . 27N.

[26] J.-E. Solheim, "Update on AM CVn stars," in *American Institute of Physics Conference Series*, ed. K. Werner and T. Rauch, vol. 1273, American Institute of Physics Conference Series, 2010, 299–304, `https://doi.org/10.1063/1.3527826`.

[27] A. Stroeer and G. Nelemans, "The influence of short-term variations in AM CVn systems on LISA measurements" 400 (November 2009): L24–L28, `https://doi.org/10.1111/j.1745-3933.2009.00754.x`.

[28] D. Steeghs, E. T. Harlaftis, and K. Horne, "Spiral structure in the accretion disc of the binary IP Pegasi" 290 (September 1997): L28–L32, `https://doi.org/10.1093/mnras/290.2.L28`.

duration times are 10–40 days over orbital periods of 22–36 mins, the recurrence times are 30–500 days over the same orbital period range, highlighting the necessity for rapid release of GW data from accreting WD binaries. **Required data products**: Catalog of parameters for accreting WDs, especially mass, chirps, and positions, made available *on timescales of days to weeks*.

- **Multi-messenger observations of massive black hole mergers.** Combined detections of the GW and EM signals powered by merging MBH binaries can significantly enhance the scientific return of LISA. EM signals provide improved localization and distance estimates, as well as information about the environment of the merging MBHs, thus opening new ways to use LISA detections in cosmology, astrophysics, and for tests of general relativity. Given the large uncertainties in the sky localization of LISA events, this will only be possible if the EM signals are sufficiently well understood to limit the rate of false positives in the LISA detection volume. The most promising EM counterparts to merging MBHs are powered by the accretion of gas onto the MBHs, both before and after merger. Our theoretical understanding of that process remains limited, in part due to the difficulty involved in simulating the many different time scales relevant to this problem, and in part due to large uncertainties regarding the feeding mechanism of MBH binaries and the large scale structure of the accretion flow. **Required data products**: Prompt alerts provided by LISA to EM observers, with tight sky localization.

The design of the LISA archive can benefit from the experience of releasing the data of ground-based detectors, and of survey space missions such as Kepler. The **LIGO–Virgo Collaboration** currently releases prompt event alerts, limited data stretches (4,096 s) around events published in refereed journals, and full strain datasets two years after the end of an observing run (although this delay may be shortened in the future). The **Gravitational Wave Open Science Center**[33] (formerly LIGO Open Science Center) hosts public datasets along with tutorials and tools to aid data use. It recently released the first *catalog of detected GW transients*[34,35], as well as *posterior samples* for selected events. Some lessons applicable to LISA are that:

- we should carefully consider the content and format of data products ahead of time (and possibly release mock examples), so that future users have time to prepare downstream tools;
- we should think carefully about giving guidance about the interpretation of the data. In the case of LIGO–Virgo, strain-data releases were never meant to allow the reproduction of GW searches,

[33] "Gravitational Wave Open Science Center," 2018, `https://gw-openscience.org`.

[34] The LIGO Scientific Collaboration and the Virgo Collaboration, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *arXiv E-Prints*, November 2018, arXiv:1811.12907, `http://arxiv.org/abs/1811.12907`.

[35] "GWOSC Catalogs (Webpage)," 2018, `https://www.gw-openscience.org/catalog`.

but were nevertheless used in that way, resulting in almost certainly spurious claims.[36]

[36] James Creswell et al., "On the time lags of the LIGO signals," *Journal of Cosmology and Astro-Particle Physics* 2017 (August 2017): 013, https://doi.org/10.1088/1475-7516/2017/08/013.

NASA's **Kepler**[37] employed the *transit method* to survey our Galactic neighborhood and discover hundreds of Earth-size and smaller planets in or near the habitable zone, characterizing the fraction of stars that might have such planets. The statistical nature of Kepler's goals, its increase in sensitivity with longer baselines, and its resurrection in a different operating mode after a hardware failure resulted in **multiple versions and multiple kinds of astronomical data catalogs** (see a full list[38]). Some lessons applicable to LISA are that:

[37] "Kepler and K2," 2018, https://www.nasa.gov/mission_pages/kepler.

[38] "Kepler and K2 Data Products," 2018, https://keplerscience.arc.nasa.gov/data-products.html.

- the astronomical community can deal with multiple versions of object catalogs and with objects changing status between catalogs (or even appearing and disappearing), but consistent versioning is paramount;
- the full range of products should be defined early to allow for good interfaces and dataflow in processing and archive;
- project definitions and thresholds should be enforced strictly to avoid confusion (e.g., we should avoid changes in SNR thresholds or object status definitions between groups in project).

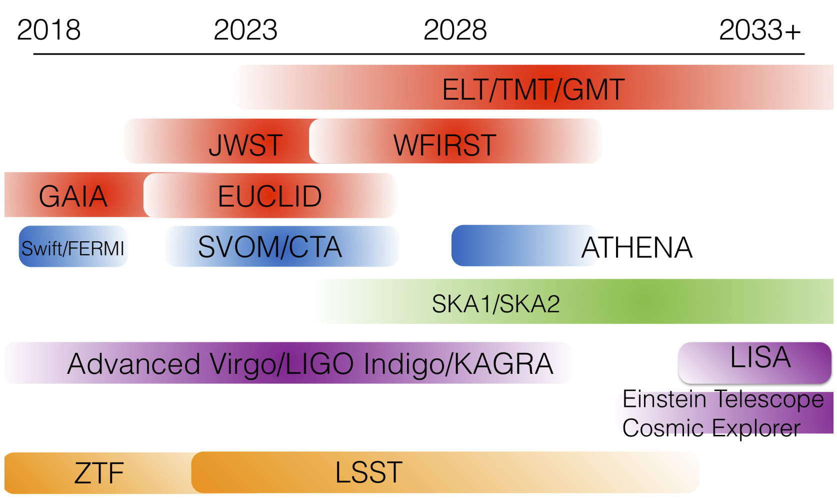


Figure 3: Astronomical observatories that will provide input to LISA science before LISA flies. Credit:

Hard problems

The global fit

It was realized early[39] in the development of the LISA mission that

[39] Tom Prince, Bonny Schumaker, and Michele Vallisneri, "Analysis Methods for Interferometric Gravitational-Wave Observations from Space (Amigos)," 2005, http://www.vallis.org/publications/amigos-2005.pdf.

the problem of combining the identification and characterization of thousands of individual sources would be a difficult one, especially in the presence of data gaps and artifacts, or of poorly measured, nonstationary/non-Gaussian noise.

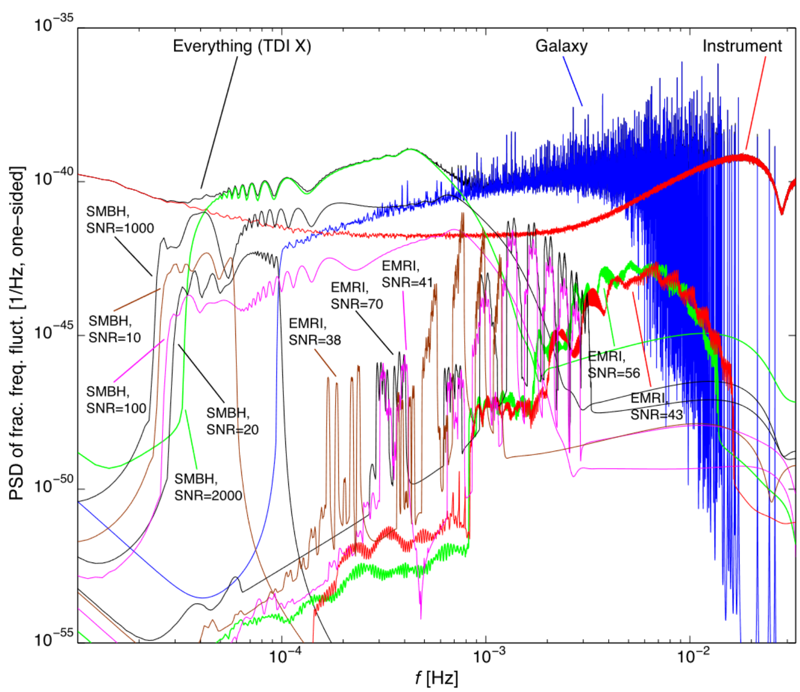
Many reasons for worry. The "confusion" problem in LISA data analysis lies in the fact that multiple signals (if *overlapping* geometrically in signal space) interact in the likelihood. Typical overlaps are small, but small groups of sources will be in close proximity, and be poorly resolvable. Furthermore, with many signals the confusion adds up and unresolved signals (including many millions of Galactic binaries) act as an effective non-Gaussian and nonstationary noise term,[40,41].

[40] C. Cutler and J. Harms, "Big Bang Observer and the neutron-star-binary subtraction problem" 73, no. 4 (February 2006): 042001, https://doi.org/10.1103/PhysRevD.73.042001.

[41] T. Robson and N. Cornish, "Impact of galactic foreground characterization on a global analysis for the LISA gravitational wave observatory," *Classical and Quantum Gravity* 34, no. 24 (December 2017): 244002, https://doi.org/10.1088/1361-6382/aa9601.

Furthermore, even the number of sources is unknown; this calls for *transdimensional* likelihood exploration algorithms, where goodness of fit (which will always increase with the number of model parameters) is tempered by a regularization term (which happens naturally in a Bayesian framework [laumann2018]), or ideally by a strong astrophysical prior.

The basic building block of GW analysis is the likelihood/sampling distribution of instrument noise, which is evaluated most efficiently in a diagonal basis. For stationary Gaussian noise, the frequency domain provides such a representation. However LISA instrument noise will be nonstationary because of instrument effects, and also because of the "seasonal" confusion background from the Galaxy.



Exploiting randomness. Stochastic methods such as Markov-

Chain Monte Carlo (MCMC)[42] have long been a workhorse in GW analysis. Applying them to specific families of GW signals is more in the realm of art (or at least craft) than science. The choice of efficient proposals that provide a reliable guide for stochastic exploration. In addition to the basic strategy of local posterior approximation with Fisher-matrix covariance, GW work has successfully employed *ad hoc* steps based on physical symmetries (e.g., jumps between multiple EMRI harmonics), precomputed (and necessarily approximated) global likelihood maps (e.g., an F-statistic map for global binaries), *differential evolution* (i.e., using past stochastic history to suggest jumps), and more. Approaches that improve on the basic Metropolis–Hastings exploration algorithm include *parallel tempering*,[43,44], Hamiltonian Monte Carlo,[45] several variants of nested sampling, and more. Other GW-specific tricks involve the analytical maximization or marginalization of the likelihood, whenever possible.

Glimpses of solutions. Cornish suggests that a wavelet representation similar to what used by LIGO's WaveBurst may allow the efficient representation of slowly varying nonstationary noise,[46,47], as well as an expedient treatment of gaps. Such a technique creates a need for efficient waveform generation in the wavelet domain.

In the LIGO context, *BayesWave*[48] represents a sophisticated and efficient approach to modeling signals, instrument glitches, and noise out of basic components (wavelets for signals and glitches) assembled in trans-dimensional fashion. Its approach to signal and noise priors is pragmatic rather than fundamental (what "true" prior probability can after all be assigned to a given wavelet?), and it is validated by the standard strategy of injecting and recovering simulated signals in actual and synthesized instrument noise.

To tackle the global-fit problem, Cornish suggests a strategy of "solving for anything", using a Bayesian model with dimension $\sim$ 500,000, possibly updating sources and parameters in turn. The problem can be parallelized by creating many equivalent chains simultaneously (although each will need to deal with all parameters), and by adopting parallel tempering.

After surveying this landscape, study participants identified a list of **key needs and challenges**:

- **Waveform overlaps** need to be quantified extensively, both between different waveform families, and between different parameter regions for the same type of waveforms (e.g., for EMRIs, which have very complex waveforms with multiple harmonics that may lead to partial "matches" between distant regions of parameter space).

[42] S. Sharma, "Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy" 55 (August 2017): 213–59, https://doi.org/10.1146/annurev-astro-082214-122339.

[43] R. H. Swendsen and J.-S. Wang, "Replica Monte Carlo simulation of spin glasses," *Physical Review Letters* 57 (November 1986): 2607–9, https://doi.org/10.1103/PhysRevLett.57.2607.

[44] W. D. Vousden, W. M. Farr, and I. Mandel, "Dynamic temperature selection for parallel tempering in Markov chain Monte Carlo simulations" 455 (January 2016): 1919–37, https://doi.org/10.1093/mnras/stv2422.

[45] Radford M Neal and others, "MCMC Using Hamiltonian Dynamics," *Handbook of Markov Chain Monte Carlo* 2, no. 11 (2011): 2.

[46] R. Dahlhaus, "Fitting Time Series Models to Nonstationary Processes," *Ann. Statist.* 25, no. 1 (February 1997): 1–37, https://doi.org/10.1214/aos/1034276620.

[47] Guy P Nason, Rainer Von Sachs, and Gerald Kroisandt, "Wavelet Processes and Adaptive Estimation of the Evolutionary Wavelet Spectrum," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 62, no. 2 (2000): 271–92, https://www.jstor.org/stable/3088859.

[48] N. J. Cornish and T. B. Littenberg, "Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches," *Classical and Quantum Gravity* 32, no. 13 (July 2015): 135012, https://doi.org/10.1088/0264-9381/32/13/135012.

- Priors for the **number of detectable sources** are poorly known; too strong assumptions may bias astrophysical inference from catalogs down the line.
- Extreme efficiency will be needed in **generating waveforms** *and* in producing the corresponding **LISA responses** (not a trivial operation for LISA).
- Bias from systematic/**unmodeled effects** (e.g., white-dwarf triples[49]) needs to be excluded or addressed. Likewise, instrument glitches will need to be recognized, then fit or excised.
- The **implementation** of the global fit is still defined very vaguely, and it is likely to be challenging technically. Issues to be clarified include:
  - Computational **parallelism**, which may be based on tasks, or populations, or source classes. Can we have multiple pipelines that fit different clusters of "similar" signals?
  - The logistics of **iterative exploration** (possibly proceeding from the strongest to the weakest sources). If iteration proceeds by subtracting progressively weaker signals, then the imperfect fit of the louder signals may leave significant residuals that confuse the fit of the weaker signals.
  - The **representation of very highly dimensional posteriors**.
- The global-fit solution will **evolve with time**: it will become better defined but more complex as we accrete more and more data, with subthreshold events transitioning into detectability.
- The quality of data and prevalence of noise will also vary naturally across time. Should our analysis proceed from the quietest times?
- How do we **evaluate and visualize** the **performance** (and correctness!) of likelihood exploration? (E.g., how do we know that we are exploring all relevant modes?)

[49] T. Robson et al., "Detecting hierarchical stellar systems with LISA" 98, no. 6 (September 2018): 064012, `https://doi.org/10.1103/PhysRevD.98.064012`.

Beyond these well-specified areas of investigations, study participants brainstormed several more open **ideas to explore**:

- Are there any **new methods** that could be applied to the problem of finding global maxima in high-dimensional search problems?
  - Is MCMC truly the only appropriate paradigm?
  - Are there alternatives to reversible-jump MCMC to handle transdimensionality?
  - Study participants mentioned the Hilbert–Huang transform, empirical mode decomposition, intrinsic time-scale decomposition.
- Could we use subsets of data to **soften likelihood surfaces** and subsets of parameters to reduce the dimensionality of the prob-

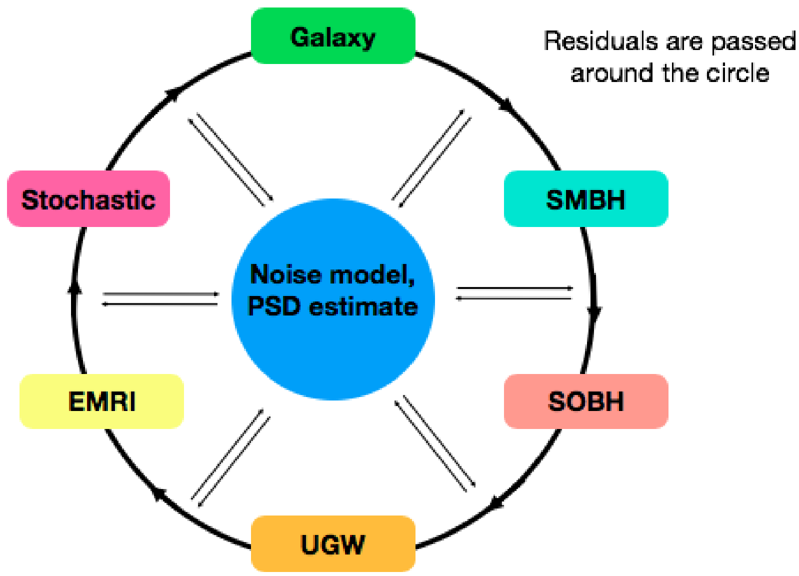


Figure 4: A schematic representation of the global-fit algorithm. Credit: N. Cornish & C. Cutler

lem? (E.g., with only a week of data all binaries would be well-described by two parameters.) These fits/posteriors can then be evolved incrementally into more complete and complex ones.

- It may be possible to take sub-threshold EMRI fits on short data spans and use **pattern recognition** to assemble them into a consistent EMRI solution.
- Can **machine learning** help with the parallel allocation of stochastic chains (possibly for subsets of sources)? Likewise, can we use machine learning to tune other search parameters?
- How can we **accelerate likelihood evaluations** (e.g., with GPUs, *tensor processing units*, precomputation/reuse of computation) and where can we usefully approximate them?
- Can we **fit confusion backgrounds of subthreshold sources**? (E.g., EMRIs, by somewhat keying on their distinctive features?)

Designing the LISA catalog

When Hogg and Lang state forcefully that "Telescopes don't make catalogs"[50] they mean that the direct output of telescopes are images (arrays of pixels), whereas the desired scientific outcome of observations and surveys is *statistical inference* about astrophysical properties or astronomical populations. Standard catalogs, which typically describe best-fit parameters and error bars for all detected sources, are neither; they represent a historically entrenched intermediate data product, which may nevertheless diminish science payoffs if informa-

[50] D. W. Hogg and D. Lang, "Telescopes don't make catalogues!" in *EAS Publications Series*, ed. C. Turon, F. Meynadier, and F. Arenou, vol. 45, EAS Publications Series, 2010, 351–58, `https://doi.org/10.1051/eas/1045059`.

tion is lost in compiling them.

There are several reasons to believe that a "traditional" LISA catalog may suffer mightily from this problem:

- **Source parameter correlations.** Most of LISA's sources are long-lived (radiating GWs in LISA's detection band for years or more), so tens of thousands of them will be present simultaneously in the LISA time series, and will need to be disentangled. If two detectable sources have strong overlap *in signal space*, then the parameter errors for one will be strongly correlated with the parameter errors of the other. Such correlations would need to be represented in the LISA catalog. As discussed above, even the *number* of sources may be uncertain: e.g., three WD–WD binaries with similar sky positions and frequencies may be fit reasonably well by a set of two binaries. It is not obvious how such alternative explanations of the data may even be represented in a catalog. That such things will occur is to some extent an unavoidable fact of this mission, and the catalog must be designed with this in mind. Understanding the extent of these sorts of "overlaps" will surely be important to designing a useful catalog. Hints of understanding appeared in the literature (e.g., Cutler and Racine[51] use large-deviation theory to show that it is extremely improbable that a large number of low-SNR WD binaries could mimic either an EMRI signal or MBHB signal), but much remains murky.
- **Quasi-detections.** The catalog should include a list of sub-threshold sources that come near the statistical threshold of confident detection, as well as preliminary estimates of their parameters (and possibly estimates of when the detector threshold would be reached, for the purpose of possible EM follow up). Such sources contain significant information for astrophysical rate and population studies; furthermore, they play an important role at the boundary between confirmed individual detection and diffuse signal components. Sub-threshold sources will suffer even more from parameter correlations.
- **Non-Gaussian posteriors.** Even for clearly detected sources, the inferred posterior probability distributions of parameters may have complicated, non-normal shapes, and possibly multiple local maxima with comparable probability weights. (The Mock LISA Data Challenges indicated that this is the case for binaries of rapidly spinning MBHs, as well as EMRIs.) The most benign catalog presentation for these sources would be a Gaussian mixture (a weighted sum of normal distributions) in parameter space, but a more expensive description as a Monte Carlo sampling of parameter values may be needed. Inter-source correlations then become

[51] É. Racine and C. Cutler, "Gaussianity of LISA's confusion backgrounds" 76, no. 12 (December 2007): 124033, https://doi.org/10.1103/PhysRevD.76.124033.

even more problematic: for each source, we would want distributions that are marginalized with respect to all other sources and systematics; but we would also need a way to represent the correlations.

A vision for the LISA catalog. Looking beyond problems to aspirations and solutions, we may adopt Hogg and Lang's a goal for a maximally performing catalog to allow statistical inference *just as if we were analyzing the raw data directly*. In mathematical terms, this means that we need access to the likelihood of the observed data as a function of a model (i.e., the parameters of all sources, plus parameters describing instrument noise and other effects). Hogg and Lang suggest three possible implementation steps, of increasing technical complexity and computational cost:

1. A **traditional catalog**, with sufficient information (parameter correlations, systematic effects) such that we are able reconstruct the likelihood of the data given variations with respect to nominal parameter estimates. The Gaia catalog[52] will come close to meeting these criteria. Because of the problems outlined above, it seems difficult for a LISA catalog to do the same, except perhaps for some special subsets of detected sources.
2. A **collection** (indeed, a sampling) **of traditional catalogs**, which collectively approximate the joint posterior distribution of all source parameters. A LISA data-analysis pipeline that culminates in a global fit would in principle be able to produce a sampling of catalogs, but we would need to ensure that the sampling is actually representative of the underlying joint posterior. Hybrid solutions may be expedient. For the purpose of astrophysical population inference, it would be especially useful to commission a service (an application or web server) that could return equal-weight (sub-)catalog samples for a user-selected category of sources.
3. A self-contained software application or web service that returns the **likelihood of the data given an input catalog**. Computing one such response would amount to computing waveforms for all postulated sources, and cross-correlating their sum with the LISA data—obviously a very expensive computational task, which may nevertheless be possible on large dedicated computing clusters, or by exploiting a reservoir of pre-computed waveforms, waveform basis decompositions, partially evaluated likelihoods, or other approximations.

[52] "Gaia Archive," 2018, `https://gea.esac.esa.int/archive`.

The analytical and computational feasibility of these steps needs to be scoped out. In addition, the LISA catalog should embody a set of mission-specific features, which again pose intriguing data-analytical or data-science challenges.

- **Updates.** The signal-to-noise ratios for most detected sources will grow continuously throughout the mission, so frequent catalog updates will be useful for the broader scientific community and, crucially, to organize EM follow ups and request *protected observations* at the expected time of binary mergers. Indeed, the updates would both refine parameter estimates and add new sources that have passed the detection threshold.
- **A living catalog.** The LISA catalog should be capable of archiving and presenting the results of EM observations, whether searches for EM transient counterparts for MBH mergers, or optical observations of Galactic binaries (as discussed in above under "Community use cases").
- **Powerful visualization.** Powerful visualization tools will be especially useful for many tasks. For instance, mission scientists will want to perform spot checks of the largely automated catalogs, by zooming in on sources that have recently transitioned from *subthreshold* to *detected*, and/or sources whose parameter estimates are changing rapidly. The relatively high dimensionality of posterior distributions makes their visualization challenging.

Preparing for the unexpected

In one of the breakout sessions, study participants were asked to discuss and document a chain of events that leads to detecting a GW signal in LISA data that cannot be mapped to any known astrophysical source. How would LISA scientist convince the world that the signal is real? Ground-based observatories may face the same problem for signals that are only detected in *unmodeled* search pipelines, but the LISA case is complicated by the presence of many concomitant signals, and by the relative paucity of auxiliary channels monitoring instrument performance.

The participants discussed several scenarios:

- signals with clearly attributable EM counterparts (the easiest case);
- signals with weird or mismatched EM counterparts;
- signals found by black-box machine-learning algorithms with strange parameter attributions;
- signals that match alternative-gravity dynamics better than general relativity;
- signals found in open data releases by nonstandard third-party pipelines;

- a set of similar low-significance "triggers" (according to standard pipelines) that were reanalyzed at the end of the mission;
- finally, the absence of a GW signal where one is expected (e.g., for a verification binary).

All of these have very interesting implications (both scientific and sociological), and may be instructive examples for roleplaying exercises in future mock data challenges. The participants converged on a strong imaginary candidate, **GW350914**, which they characterized as follows:

- very strong, with SNR = 200, and high Bayes factor (above 5σ);
- not periodic, lasting 1.2 days, with peak frequency of 3 mHz;
- localized to within 25 square degrees in the sky, but without apparent EM counterparts;
- with no associated features in auxiliary channels (i.e., "green" data quality);
- a very good way to celebrate the 20th anniversary of GW detection!

The participants further developed a *real-vs-not-real* **checklist**,[53] which begins by questioning the state of the instrument, then transitions to data analysis. Such a process is formalized well for LIGO.[54,55]

1. All spacecraft systems are nominal.
2. Nothing (too) suspicious is seen in the auxiliary channels.
3. The rate of test-mass decharging was not increased.
4. The spacecraft antennae were not moving or settling.
5. The spacecraft gas tanks were not experiencing any unusual phenomena.
6. Space weather was not significant.
7. Consistent time delays are observed between the excitations seen at the three LISA spacecraft. (Best assessed if all six LISA links are available.)
8. The signal cannot be explained as the residual from the imperfect subtraction of other sources.
9. Nothing is seen in the null/breathing mode channel (if the frequency is sufficiently low that the null channel is actually GW-insensitive).
10. Signal consistency tests pass [editor's note: though it is unclear what these would be for an unmodeled signal].
11. The signal can be consistently attributed to one region in the sky.
12. The distance estimate is sensible (i.e., not in our planetary backyard).
13. A reasonable combination of search pipelines detects the signal.

[53] A. Gawande, *The Checklist Manifesto: How to Get Things Right* (Henry Holt; Company, 2010), `https://books.google.com/books?id=x3IcNujwHxcC`.

[54] B. P. Abbott et al., "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914," *Classical and Quantum Gravity* 33, no. 13 (July 2016): 134001, `https://doi.org/10.1088/0264-9381/33/13/134001`.

[55] B. P. Abbott et al., "Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGOs first observing run," *Classical and Quantum Gravity* 35, no. 6 (March 2018): 065010, `https://doi.org/10.1088/1361-6382/aaaafa`.

14. The possibility of a hoax (through either hardware of software injection) can be investigated an excluded. This was done very carefully for the first LIGO detection.[56]

[56] Michael Landry and Brian O'Reilly, "LLO, LHO Site Status for Gw150914," 2015, `https://dcc.ligo.org/LIGO-L1500138/public`.

New tools

Machine learning

Modern astronomical surveys and time-domain–astronomy programs produce huge amounts of data, often characterized by non-ideal properties (gaps, irregular sampling, heteroskedastic errors) that are ignored by many standard statistical methods. To tackle this onslaught, astronomers have been increasingly turning to *machine learning*, which largely eschews the principled application of statistical theory in favor of a **training–testing–application** sequence. Machine-learning algorithms draw inferences *directly from the data*, either by comparing their evolving guesses with the "right" answers provided by the user (supervised learning), or by discovering hidden structure or groupings from first principles (unsupervised learning).

Traditional machine learning relies on designing and deriving relatively low-dimensional **feature vectors** from homogeneous data, and then feeding the features to classification (guessing the discrete class of an input) or regression (estimating continuous parameters) algorithms. For the same type of data, many different features can be defined (and many end up being very correlatd in practice), and the choice of features can drastically alter performance. Furthermore, "naturally" defined features can carry physical meaning that is lost when they are chosen from a more mathematically amorphous set, a in Principal Component Analysis.[57] When features or classes are obvious to human eyes but elusive for mathematical definition, citizen science[58] can provide training "labels" for supervised learning.

Deep learning[59] (which generally refers to stochastically trained *deep* neural networks, DNNs, with many layers) works directly from "raw" data, but its box is rather black. In its most impressive application so far, Google DeepMind's *AlphaZero* application learned to play chess, shogi, and Go at superhuman levels by unsupervised reinforcement learning—that is, by *playing itself* on the basis of the rules of the game alone.[60] Among other techniques, AlphaZero relies on **convolutional neural networks** (CNNs), which apply multiple linear convolutions to input vectors and arrays, and which are employed broadly in computer vision and other image processing. Trained convolutional networks are opaque in how they achieve their tasks, but the technique of *deconvnet*s can be used to project the "feature" activations, such as the notion of "cat" for an animal image classi-

[57] C. Donalek et al., "Feature Selection Strategies for Classifying High Dimensional Astronomical Data Sets," in *2013 Ieee International Conference on Big Data*, 2013, 35–41.

[58] "Zooniverse," 2018, `https://www.zooniverse.org`.

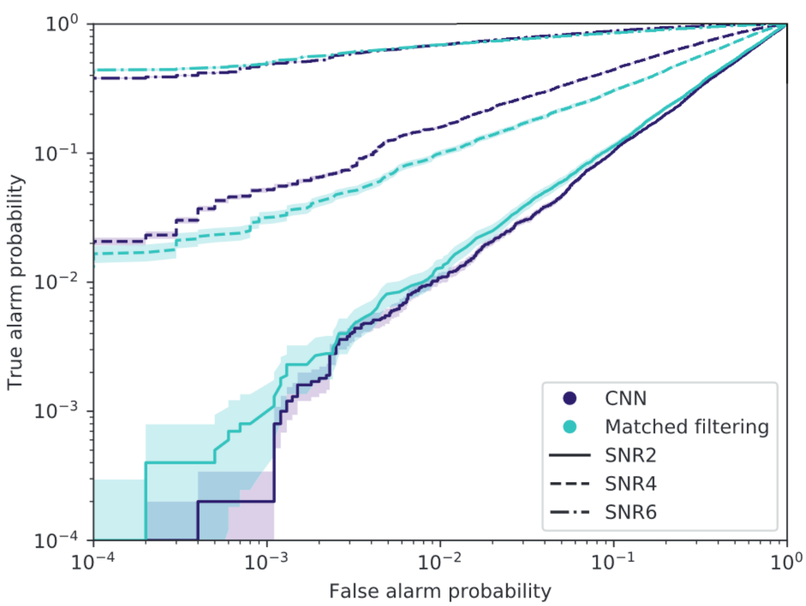
[59] Jürgen Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks* 61 (January 2015): 85–117, `https://doi.org/10.1016/j.neunet.2014.09.003`.

[60] David Silver et al., "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go Through Self-Play," *Science* 362, no. 6419 (2018): 1140–4, `https://doi.org/10.1126/science.aar6404`.

fier back to the input pixel space[61,62]. In fact, studying activation patterns can identify useful features, learned directly from the data.

Study participant Mahabal described the application of an encoder–decoder CNN to the task of image differencing in transient detection.[63] The approach "encapsulates all the steps of a traditional image-subtraction pipeline — image registration, background subtraction, noise removal, PSF matching and subtraction — in a single real-time convolutional network." Datasets characterized by an "arrow of time" can be tackled by a deep-learning technique known as *recurrent neural networks*[64] (RNNs) which allow for loops, or equivalently for the propagation of information across multiple iterations of the network. *Long Short Term Memory* networks[65] (LSTMs) are RNNs that are capable of learning long-term dependencies, and can be used for light curves and other time series.

Convolutional networks have been applied to the detection and localization of GW signals from LIGO black-hole mergers, in both the time domain[66] and the frequency domain[67,68]. The analysis spanned simulated LIGO-type signals[69,70,71] as well as real LIGO data.[72] CNNs were also tasked with the classification of transient noise events ("glitches") in LIGO data[73,74], and with the denoising of LIGO signals within *recurrent autoencoders*.;[75,76]



Beyond CNNs, LIGO glitches were tackled with *dictionary learning*,[77] and denoising with dictionary learning[78] and total-variation methods[79,80]. Finally, study participants Chua, Galley, and Vallisneri have demonstrated the use of DNNs as fast waveform interpolants in a ROM representation, providing benefits such as native parameter estimation and analytic waveform derivatives.[81] (**Obviously the ref-**

[61] Matthew D. Zeiler and Rob Fergus, "Visualizing and Understanding Convolutional Networks," in *Computer Vision – Eccv 2014*, ed. David Fleet et al. (Cham: Springer International Publishing, 2014), 818–33.

[62] Chris Olah et al., "The Building Blocks of Interpretability," *Distill*, 2018.

[63] N. Sedaghat and A. Mahabal, "Effective image differencing with convolutional neural networks for real-time transient hunting" 476 (June 2018): 5365–76, `https://doi.org/10.1093/mnras/sty613`.

[64] Andrej Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks (Blog Post)," 2015, `https://karpathy.github.io/2015/05/21/rnn-effectiveness`.

[65] Christopher Olah, "Understanding Lstm Networks (Blog Post)," 2015, `https://colah.github.io/posts/2015-08-Understanding-LSTMs`.

[66] Timothy Gebhard and Niki Kilbertus, "ConvWave: Searching for Gravitational Waves with Fully Convolutional Neural Nets," 2017, `https://github.com/nikikilbertus/convwave`.

[67] D. George and E. A. Huerta, "Deep neural networks to enable real-time multimessenger astrophysics" 97, no. 4 (February 2018): 044039, `https://doi.org/10.1103/PhysRevD.97.044039`.

[68] H. Gabbard et al., "Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy," *Physical Review Letters* 120, no. 14 (April 2018): 141103, `https://doi.org/10.1103/PhysRevLett.120.141103`.

[69] Fan XiLong et al., "Applying Deep Neural Networks to Detection and Space Parameters Estimation of Compact Binary Coalescence with a Network of Gravitational Wave Detectors," *SCIENCE CHINA Physics, Mechanics & Astronomy*, n.d.

[70] A. Rebei et al., "Fusing numerical relativity and deep learning to detect higher-order multipole waveforms from eccentric binary black hole mergers," *arXiv E-Prints*, July 2018, `http://arxiv.org/abs/1807.09787`.

[71] H. Nakano et al., "Comparison of various methods to extract ringdown frequency from gravitational wave data," *arXiv E-Prints*, November 2018, `http://arxiv.org/abs/1811.06443`.

[72] D. George and E. A. Huerta, "Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data," *Physics Letters B* 778 (March 2018): 64–70, `https://doi.org/10.1016/j.physletb.2017.12.053`.

[73] M. Razzano and E. Cuoco, "Image-based deep learning for classification of noise transients in gravitational wave detectors," *Classical and Quantum Gravity* 35, no. 9 (May 2018): 095016, `https://doi.org/10.1088/1361-6382/`

erences do not all fit in the margin; I will shorten them somehow.)

The KISS discussion of machine learning (and deep learning especially), raised a number of questions (Q) and (wishful?) suggestions (S). These are some of them.

- Q: can DNNs learn the **ideal proposal distribution for MCMC**? — The ideal Bayesian proposal is in fact the posterior itself; it may be possible, even if it seems very ambitious, to train DNNs to return approximations to Bayesian posteriors given the data, by exposing them to a large variety of noisy datasets. More in line with the spirit of the question, DNNs may learn to offer "smart" proposal moves within MCMC runs, since for each proposed move we know immediately whether it is accepted; however the Metropolis–Hastings rule requires that for each proposal $x \to y$ we know the probability of the inverse proposal $y \to x$. This may be difficult to arrange.

- Q: how are DNN architecture (e.g., depth, connectivity. . .) and hyperparameters (e.g., learning rate) decided? — It's a **black art**. Some groups use Bayesian methods.

- S: while deep learning is opaque (and by extension not statistically "pure"), it could provide a **rough map of where to look for signals** in GW searches. The parallel with stochastic methods is interesting: these are used to solve problems that we can specify analytically, but not solve easily; supervised DNNs excel where the desired outcome is clear, but we cannot even formulate a way to get the solution. However, we may still need some reframing of GW problems before we can make headway.

- S: training sets containing GW signals are generated quickly with simulation; *generative networks* may be more appropriate for glitch modeling, where we have limited data.

- S: instead of using machine learning to replace classical solutions to the big problems of GW data analysis (detection, parameter estimation, data quality), we should go over the pieces of pipelines, and figure out which could be made faster by replacing it with a fast (but expensive to train) black box. For instance, likelihoods can be *emulated*; [10.1111/j.1365-2966.2011.20288.x] or as suggested above, the control systems that govern the solution of NR could benefit from a steady automated hand.

- S: we should tap into the large and enthusiastic user/scholar community flocking to deep learning. If we describe problems well (e.g., on "competition" portals such as Kaggle[82]), and abstract

[82] "Kaggle: Your Home for Data Science," 2018, `https://www.kaggle.com`.

them as much as possible from GW jargon, talented people may be interesting in applying their own methods.

- S (based on presentation by Nvidia GPU software engineer): deep learning is increasingly driving the development and capabilities of new GPU generations; conversely, the continued increase in computing power in modern computers relies increasingly on the teraflops unleashed by the massively parallel arrays of GPU compute cores. It follows that if GW data analysis is to continue benefiting from Moore's law, it must reformulate its tasks as DL problems, or at least as algorithms that use the same mathematical primitives. (In fact, Google, Apple, and other large players are now building very fast and efficient *tensor processing units* that are strictly designed for DNNs.[83])

[83] Kaz Sato, "What Makes Tpus Fine-Tuned for Deep Learning? (Google Cloud Blog)," n.d., `https://cloud.google.com/blog/products/ai-machine-learning/what-makes-tpus-fine-tuned-for-deep-learning`.

Other talking points from workshop group sessions

Notes from breakout session in Oort Cloud - new ideas

- EM side: can we use archival data from the EM community to look in LISA for sources. And also vice versa.
- large scale data-mining
- deep learning correlations for EM-GW
- Denial of service attack on theorists - you take down a website by having many computers visit website. Same thing with data generation. E.G. we found these 10 signals in these 10 periods and some of them are null. See if the theorists pick out the fake parts
- Improving DQ
- Better input for MCMC, using neural nets for feature extraction. Something similar has been done for multi-nests. Have so far looked at toy problems.
- Generate enough waveforms (loads), would we be able to make progress on the inverse problem? Use ML to use the parameters to come out with the waveform.
- PE exploration. Whether cloud services can be used to distribute workload. Use ROMs and just grid based brute force exploration with loads of computers.
- Kaggle competition - online platform were you have competitions involving datasets with answers to them. Train ML on the dataset and see how you do. Submit answers/algorithms. Many different sites where we can host something like this.

- How we use astrophysical priors for NR and analysis - to find out where the theorists should spend most of their time

Useful to have the astrophysical modelling, and plug in to the inverse problem so you are agnostics. Go from LISA data to go back to formation channel.

Going to have 1000s of waveforms for LISA. Vector basis and randomly choose coefficients to find the best match. Need to solve the inverse problem to do the PE. Detection is probably 'easy' but won't have the parameters easily. Want to go from wiggle in the data to figuring out waveform rather than matching the waveform to the data.

*) Can create simulated LISA data based on EM catalogues

*) Create targeted GW search for something from Gaia etc archives

*) Find a way to map probabilities / joint probabilities - say found something in EM to 80% but GW to 20% - would you believe it?

Biblio

Abbott, B. P., R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, M. Adamo, et al. "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914." *Classical and Quantum Gravity* 33, no. 13 (July 2016): 134001. `https://doi.org/10.1088/0264-9381/33/13/134001`.

Abbott, B. P., R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, et al. "Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGOs first observing run." *Classical and Quantum Gravity* 35, no. 6 (March 2018): 065010. `https://doi.org/10.1088/1361-6382/aaaafa`.

Babak, Stanislav, Hua Fang, Jonathan R. Gair, Kostas Glampedakis, and Scott A. Hughes. "'Kludge' Gravitational Waveforms for a Test-Body Orbiting a Kerr Black Hole." *Phys. Rev. D* 75, no. 2 (January 2007): 024005. `https://doi.org/10.1103/PhysRevD.75.024005`.

Barack, Leor, and Curt Cutler. "LISA Capture Sources: Approximate Waveforms, Signal-to-Noise Ratios, and Parameter Estimation Accuracy." *Phys. Rev. D* 69, no. 8 (April 2004): 082005. `https://doi.org/10.1103/PhysRevD.69.082005`.

Bonnet-Bidaud, J. M., M. Mouchet, C. Busschaert, E. Falize, and C. Michaut. "Quasi-periodic oscillations in accreting magnetic white dwarfs. I. Observational constraints in X-ray and optical" 579 (July 2015): A24. `https://doi.org/10.1051/0004-6361/201425482`.

Buonanno, Alessandra, and B. S. Sathyaprakash. "Sources of Gravitational Waves: Theory and Observations." In *General Relativity and Gravitation: A Centennial Perspective*, edited by Abhay Ashtekar, Beverly K. Berger, James Isenberg, and MalcolmEditors MacCallum, 287–346. Cambridge University Press, 2015. `https:`

//doi.org/10.1017/CBO9781139583961.009.

Carrasco, J. M., S. Catalán, C. Jordi, P.-E. Tremblay, R. Napiwotzki, X. Luri, A. C. Robin, and P. M. Kowalski. "Gaia photometry for white dwarfs" 565 (May 2014): A11. https://doi.org/10.1051/0004-6361/201220596.

Choptuik, Matthew W., Luis Lehner, and Frans Pretorius. "Probing Strong-Field Gravity Through Numerical Simulations." In *General Relativity and Gravitation: A Centennial Perspective*, edited by Abhay Ashtekar, Beverly K. Berger, James Isenberg, and MalcolmEditors MacCallum, 361–411. Cambridge University Press, 2015. https://doi.org/10.1017/CBO9781139583961.011.

Chua, A. J. K., C. R. Galley, and M. Vallisneri. "ROMAN: Reduced-Order Modeling with Artificial Neurons." *arXiv E-Prints*, November 2018. http://arxiv.org/abs/1811.05491.

Chua, Alvin J K, and Jonathan R Gair. "Improved Analytic Extreme-Mass-Ratio Inspiral Model for Scoping Out eLISA Data Analysis." *Classical and Quantum Gravity* 32, no. 23 (November 2015): 232002. https://doi.org/10.1088/0264-9381/32/23/232002.

Chua, Alvin J. K., Christopher J. Moore, and Jonathan R. Gair. "Augmented Kludge Waveforms for Detecting Extreme-Mass-Ratio Inspirals." *Phys. Rev. D* 96, no. 4 (August 2017): 044005. https://doi.org/10.1103/PhysRevD.96.044005.

Coppejans, D. L., E. G. Körding, J. C. A. Miller-Jones, M. P. Rupen, G. R. Sivakoff, C. Knigge, P. J. Groot, P. A. Woudt, E. O. Waagen, and M. Templeton. "Dwarf nova-type cataclysmic variable stars are significant radio emitters" 463 (December 2016): 2229–41. https://doi.org/10.1093/mnras/stw2133.

Cornish, N. J., and T. B. Littenberg. "Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches." *Classical and Quantum Gravity* 32, no. 13 (July 2015): 135012. https://doi.org/10.1088/0264-9381/32/13/135012.

Creswell, James, Sebastian von Hausegger, Andrew D. Jackson, Hao Liu, and Pavel Naselsky. "On the time lags of the LIGO signals." *Journal of Cosmology and Astro-Particle Physics* 2017 (August 2017): 013. https://doi.org/10.1088/1475-7516/2017/08/013.

Cutler, C., and J. Harms. "Big Bang Observer and the neutron-star-binary subtraction problem" 73, no. 4 (February 2006): 042001. https://doi.org/10.1103/PhysRevD.73.042001.

Cutler, C., and M. Vallisneri. "LISA detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms" 76, no. 10 (November 2007): 104018. https://doi.org/10.1103/PhysRevD.76.104018.

Dahlhaus, R. "Fitting Time Series Models to Nonstationary Processes." *Ann. Statist.* 25, no. 1 (February 1997): 1–37. https://doi.

org/10.1214/aos/1034276620.

Donalek, C., S. G. Djorgovski, A. A. Mahabal, M. J. Graham, A. J. Drake, A. A. Kumar, N. S. Philip, et al. "Feature Selection Strategies for Classifying High Dimensional Astronomical Data Sets." In *2013 Ieee International Conference on Big Data*, 35–41, 2013.

Gabbard, H., M. Williams, F. Hayes, and C. Messenger. "Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy." *Physical Review Letters* 120, no. 14 (April 2018): 141103. https://doi.org/10.1103/PhysRevLett.120.141103.

"Gaia Archive," 2018. https://gea.esac.esa.int/archive.

Gair, J. R., A. Sesana, E. Berti, and M. Volonteri. "Constraining properties of the black hole population using LISA." *Classical and Quantum Gravity* 28, no. 9 (May 2011): 094018. https://doi.org/10.1088/0264-9381/28/9/094018.

Gair, J. R., C. Tang, and M. Volonteri. "LISA extreme-mass-ratio inspiral events as probes of the black hole mass function" 81, no. 10 (May 2010): 104014. https://doi.org/10.1103/PhysRevD.81.104014.

Gawande, A. *The Checklist Manifesto: How to Get Things Right*. Henry Holt; Company, 2010. https://books.google.com/books?id=x3IcNujwHxcC.

Gebhard, Timothy, and Niki Kilbertus. "ConvWave: Searching for Gravitational Waves with Fully Convolutional Neural Nets," 2017. https://github.com/nikikilbertus/convwave.

George, D., and E. A. Huerta. "Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data." *Physics Letters B* 778 (March 2018): 64–70. https://doi.org/10.1016/j.physletb.2017.12.053.

———. "Deep neural networks to enable real-time multimessenger astrophysics" 97, no. 4 (February 2018): 044039. https://doi.org/10.1103/PhysRevD.97.044039.

George, D., H. Shen, and E. A. Huerta. "Classification and unsupervised clustering of LIGO data with Deep Transfer Learning" 97, no. 10 (May 2018): 101501. https://doi.org/10.1103/PhysRevD.97.101501.

"Gravitational Wave Open Science Center," 2018. https://gw-openscience.org.

"Gravity Spy," 2018. https://www.zooniverse.org/projects/zooniverse/gravity-spy/about/results.

Groot, P. J. "Evolution of Spiral Shocks in U Geminorum during Outburst" 551 (April 2001): L89–L92. https://doi.org/10.1086/319826.

"GWOSC Catalogs (Webpage)," 2018. https://www.gw-openscience.org/catalog.

Hogg, D. W., and D. Lang. "Telescopes don't make catalogues!"

In *EAS Publications Series*, edited by C. Turon, F. Meynadier, and F. Arenou, 45:351–58. EAS Publications Series, 2010. https://doi.org/10.1051/eas/1045059.

"Kaggle: Your Home for Data Science," 2018. https://www.kaggle.com.

Kalirai, J. S. "The age of the Milky Way inner halo" 486 (June 2012): 90–92. https://doi.org/10.1038/nature11062.

Karpathy, Andrej. "The Unreasonable Effectiveness of Recurrent Neural Networks (Blog Post)," 2015. https://karpathy.github.io/2015/05/21/rnn-effectiveness.

"Kepler and K2," 2018. https://www.nasa.gov/mission_pages/kepler.

"Kepler and K2 Data Products," 2018. https://keplerscience.arc.nasa.gov/data-products.html.

Kidder, Lawrence E., Scott E. Field, Francois Foucart, Erik Schnetter, Saul A. Teukolsky, Andy Bohn, Nils Deppe, et al. "SpECTRE: A task-based discontinuous Galerkin code for relativistic astrophysics." *Journal of Computational Physics* 335 (April 2017): 84–114. https://doi.org/10.1016/j.jcp.2016.12.059.

Korol, V., E. M. Rossi, P. J. Groot, G. Nelemans, S. Toonen, and A. G. A. Brown. "Prospects for detection of detached double white dwarf binaries with Gaia, LSST and LISA" 470 (September 2017): 1894–1910. https://doi.org/10.1093/mnras/stx1285.

Kremer, K., K. Breivik, S. L. Larson, and V. Kalogera. "Accreting Double White Dwarf Binaries: Implications for LISA" 846 (September 2017): 95. https://doi.org/10.3847/1538-4357/aa8557.

Kumar, Prayush, Tony Chu, Heather Fong, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Mark A. Scheel, and Bela Szilagyi. "Accuracy of binary black hole waveform models for aligned-spin binaries" 93 (May 2016): 104050. https://doi.org/10.1103/PhysRevD.93.104050.

Landry, Michael, and Brian O'Reilly. "LLO, LHO Site Status for GW150914," 2015. https://dcc.ligo.org/LIGO-L1500138/public.

Le Tiec, Alexandre, Alessandra Buonanno, Abdul H. Mroué, Harald P. Pfeiffer, Daniel A. Hemberger, Geoffrey Lovelace, Lawrence E. Kidder, et al. "Periastron Advance in Spinning Black Hole Binaries: Gravitational Self-Force from Numerical Relativity." *Phys. Rev. D* 88, no. 12 (December 2013): 124027. https://doi.org/10.1103/PhysRevD.88.124027.

Levitan, D., P. J. Groot, T. A. Prince, S. R. Kulkarni, R. Laher, E. O. Ofek, B. Sesar, and J. Surace. "Long-term photometric behaviour of outbursting AM CVn systems" 446 (January 2015): 391–410. https://doi.org/10.1093/mnras/stu2105.

Lindblom, Lee, Benjamin J. Owen, and Duncan A. Brown. "Model waveform accuracy standards for gravitational wave data analysis" 78 (December 2008): 124020. https://doi.org/10.1103/PhysRevD.78.124020.

"LISA Data Challenges," 2018. https://lisa-ldc.lal.in2p3.fr/ldc.

Llorens-Monteagudo, M., A. Torres-Forné, J. A. Font, and A. Marquina. "Classification of gravitational-wave glitches via dictionary learning." *arXiv E-Prints*, November 2018. http://arxiv.org/abs/1811.03867.

Meent, Maarten van de, and Niels Warburton. "Fast Self-Forced Inspirals." *Classical and Quantum Gravity* 35, no. 14 (June 2018): 144003. https://doi.org/10.1088/1361-6382/aac8ce.

Moore, C. J., C. P. L. Berry, A. J. K. Chua, and J. R. Gair. "Improving gravitational-wave parameter estimation using Gaussian process regression" 93, no. 6 (March 2016): 064001. https://doi.org/10.1103/PhysRevD.93.064001.

Nakano, H., T. Narikawa, K.-i. Oohara, K. Sakai, H.-a. Shinkai, H. Takahashi, T. Tanaka, N. Uchikata, S. Yamamoto, and T. S. Yamamoto. "Comparison of various methods to extract ringdown frequency from gravitational wave data." *arXiv E-Prints*, November 2018. http://arxiv.org/abs/1811.06443.

Nason, Guy P, Rainer Von Sachs, and Gerald Kroisandt. "Wavelet Processes and Adaptive Estimation of the Evolutionary Wavelet Spectrum." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 62, no. 2 (2000): 271–92. https://www.jstor.org/stable/3088859.

Neal, Radford M, and others. "MCMC Using Hamiltonian Dynamics." *Handbook of Markov Chain Monte Carlo* 2, no. 11 (2011): 2.

Nelemans, G. "AM CVn stars." In *The Astrophysics of Cataclysmic Variables and Related Objects*, edited by J.-M. Hameury and J.-P. Lasota, 330:27. Astronomical Society of the Pacific Conference Series, 2005.

Olah, Chris, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev. "The Building Blocks of Interpretability." *Distill*, 2018.

Olah, Christopher. "Understanding Lstm Networks (Blog Post)," 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs.

Prince, Tom, Bonny Schumaker, and Michele Vallisneri. "Analysis Methods for Interferometric Gravitational-Wave Observations from Space (Amigos)," 2005. http://www.vallis.org/publications/amigos-2005.pdf.

Racine, É., and C. Cutler. "Gaussianity of LISA's confusion backgrounds" 76, no. 12 (December 2007): 124033. https://doi.org/10.1103/PhysRevD.76.124033.

Razzano, M., and E. Cuoco. "Image-based deep learning for classification of noise transients in gravitational wave detectors." *Classical and Quantum Gravity* 35, no. 9 (May 2018): 095016. https://doi.org/10.1088/1361-6382/aab793.

Rebei, A., E. A. Huerta, S. Wang, S. Habib, R. Haas, D. Johnson, and D. George. "Fusing numerical relativity and deep learning to detect higher-order multipole waveforms from eccentric binary black hole mergers." *arXiv E-Prints*, July 2018. http://arxiv.org/abs/1807.09787.

Robson, T., and N. Cornish. "Impact of galactic foreground characterization on a global analysis for the LISA gravitational wave observatory." *Classical and Quantum Gravity* 34, no. 24 (December 2017): 244002. https://doi.org/10.1088/1361-6382/aa9601.

Robson, T., N. J. Cornish, N. Tamanini, and S. Toonen. "Detecting hierarchical stellar systems with LISA" 98, no. 6 (September 2018): 064012. https://doi.org/10.1103/PhysRevD.98.064012.

Sato, Kaz. "What Makes Tpus Fine-Tuned for Deep Learning? (Google Cloud Blog)," n.d. https://cloud.google.com/blog/products/ai-machine-learning/what-makes-tpus-fine-tuned-for-deep-learning.

Schmidhuber, Jürgen. "Deep Learning in Neural Networks: An Overview." *Neural Networks* 61 (January 2015): 85–117. https://doi.org/10.1016/j.neunet.2014.09.003.

Sedaghat, N., and A. Mahabal. "Effective image differencing with convolutional neural networks for real-time transient hunting" 476 (June 2018): 5365–76. https://doi.org/10.1093/mnras/sty613.

Sesana, A., J. Gair, E. Berti, and M. Volonteri. "Reconstructing the massive black hole cosmic history through gravitational waves" 83, no. 4 (February 2011): 044036. https://doi.org/10.1103/PhysRevD.83.044036.

Sharma, S. "Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy" 55 (August 2017): 213–59. https://doi.org/10.1146/annurev-astro-082214-122339.

Shen, H., D. George, E. A. Huerta, and Z. Zhao. "Denoising Gravitational Waves using Deep Learning with Recurrent Denoising Autoencoders." *arXiv E-Prints*, November 2017. http://arxiv.org/abs/1711.09919.

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, et al. "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go Through Self-Play." *Science* 362, no. 6419 (2018): 1140–4. https://doi.org/10.1126/science.aar6404.

Solheim, J.-E. "Update on AM CVn stars." In *American Institute of Physics Conference Series*, edited by K. Werner and T. Rauch, 1273:299–304. American Institute of Physics Conference Series, 2010. https://doi.org/10.1063/1.3527826.

Steeghs, D., E. T. Harlaftis, and K. Horne. "Spiral structure in the accretion disc of the binary IP Pegasi" 290 (September 1997): L28–L32. https://doi.org/10.1093/mnras/290.2.L28.

Stroeer, A., and G. Nelemans. "The influence of short-term variations in AM CVn systems on LISA measurements" 400 (November 2009): L24–L28. https://doi.org/10.1111/j.1745-3933.2009.00754.x.

Swendsen, R. H., and J.-S. Wang. "Replica Monte Carlo simulation of spin glasses." *Physical Review Letters* 57 (November 1986): 2607–9. https://doi.org/10.1103/PhysRevLett.57.2607.

"The Architecture of Lisa Science Analysis Presentations (Pdf)," 2018. http://kiss.caltech.edu/workshops/LISA/LISA.html#Presentations.

"The Architecture of Lisa Science Analysis Short Courses (Video)," 2018. https://kiss.caltech.edu/short_courses/LISA.html.

The LIGO Scientific Collaboration, and the Virgo Collaboration. "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs." *arXiv E-Prints*, November 2018, arXiv:1811.12907. http://arxiv.org/abs/1811.12907.

Torres, A., A. Marquina, J. A. Font, and J. M. Ibáñez. "Total-variation-based methods for gravitational wave denoising" 90, no. 8 (October 2014): 084029. https://doi.org/10.1103/PhysRevD.90.084029.

Torres-Forné, A., E. Cuoco, A. Marquina, J. A. Font, and J. M. Ibáñez. "Total-variation methods for gravitational-wave denoising: Performance tests on Advanced LIGO data" 98, no. 8 (October 2018): 084013. https://doi.org/10.1103/PhysRevD.98.084013.

Torres-Forné, A., A. Marquina, J. A. Font, and J. M. Ibáñez. "Denoising of gravitational wave signals via dictionary learning algorithms" 94, no. 12 (December 2016): 124040. https://doi.org/10.1103/PhysRevD.94.124040.

Vousden, W. D., W. M. Farr, and I. Mandel. "Dynamic temperature selection for parallel tempering in Markov chain Monte Carlo simulations" 455 (January 2016): 1919–37. https://doi.org/10.1093/mnras/stv2422.

Wei, W., and E. A. Huerta. "Gravitational Wave Denoising of Binary Black Hole Mergers with Deep Learning." *arXiv E-Prints*, January 2019. http://arxiv.org/abs/1901.00869.

XiLong, Fan, Li Jin, Li Xin, Zhong Yuanhong, and Cao Junwei. "Applying Deep Neural Networks to Detection and Space Parameters Estimation of Compact Binary Coalescence with a Network of Gravitational Wave Detectors." *SCIENCE CHINA Physics, Mechanics & Astronomy*, n.d.

Zeiler, Matthew D., and Rob Fergus. "Visualizing and Understanding Convolutional Networks." In *Computer Vision – Eccv 2014*, edited by David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, 818–33. Cham: Springer International Publishing, 2014.

"Zooniverse," 2018. https://www.zooniverse.org.