

# 重力波の初検出と情報処理技術

— LIGO と KAGRA で活用されている情報処理技術 —

Kipp Cannon<sup>\*1</sup> 端山和大<sup>\*1</sup> 伊藤洋介<sup>\*1</sup> 高橋弘毅<sup>\*2</sup>

<sup>\*1</sup> 東京大学

<sup>\*2</sup> 長岡技術科学大学

## LIGO による重力波の初検出

日本時間 2015 年 9 月 14 日 18 時 50 分 45 秒～46 秒、米国 Hanford と Livingston にそれぞれ 1 台ずつ、計 2 台のレーザ干渉計からなる advanced Laser Interferometer Gravitational wave Observatory (aLIGO) がブラックホール (Black Hole, BH) 連星合体现象からの重力波信号 (GW150914 と呼称) を記録した<sup>☆1</sup>。高速解析パイプラインである coherent Wave Burst (cWB) が 3 分以内に信号を抽出、Wisconsin-Milwaukee 大学にサーバがあるデータベース Gravitational Wave Candidate Event Database (GraceDb)<sup>☆2</sup> に登録した。GraceDb は LIGO の研究者 (LIGO Scientific Collaboration : LSC) に信号検出速報を出し、その日のうちに彼らは、Albert Einstein がその存在を予言してから 100 年目にして、人類が初めて重力波信号を直接検出したことを知ることになった<sup>☆3</sup>。

このとき、速報システムの存在は本質的であった。実は公式の観測は 2015 年 9 月 19 日 0 時 (日本時間)

開始であって、信号検出時、検出器はエンジニアリングランと呼ばれる試験観測を行っていたのだ<sup>☆4</sup>。速報がなければ研究者たちは最後の調整に勤しみ、検出器の構成や設定を変えてしまっていた可能性がある。実験装置の状況が変わってしまえば、ノイズ特性の研究を行えず、たとえば信号の偽検出率<sup>☆5</sup>を求めることはできなかった。速報システムが有効に働いたため、研究者たちは検出時点での装置の設定・構成を詳細に記述、設定パネル・コネクタなどの写真約 500 枚を撮って、その後約 1 カ月間ほど同一構成・設定のまま、検出器を運用することができた。

cWB は本質的に時間・周波数領域に局在する飛び値を検出するため、連星合体、超新星爆発や、未知の重力波源を探索できるなど、汎用性が高い<sup>☆6</sup>。一方で、検出した信号の解釈にはなんらかのモデルが必要である。また検出効率の点でも、連星合体に特化すれば、信号対雑音比を最大化するという意味で最適な線形フィルタであるマッチドフィルタを使うことができる。LSC の gstlal<sup>☆7</sup>、pycbc<sup>☆8</sup> といった解析パイプラインは、モデルを連星合体现象に限定してマッチドフィルタを使うことによって、BH の質量、自転、到来方向などの情報をさ

☆1 電荷を持った物体が加速度運動すると電磁波を放射するように、質量を持った物体が加速度運動すると、これに伴う重力場の時間変動は重力波として光速以下で伝わる。ただし、重力相互作用はとても弱いので、検出できるほどの重力波が放射されるのは、天体が光速に近い速度で運動をしているようなときである。1974 年の連星パルサー PSR 1913+16 の発見は、重力波の存在の間接的な証明につながり、発見者の Hulse と Taylor にはノーベル物理学賞 (1993 年) が授与された。電波・X 線天文学など観測手段の多様化が 20 世紀の天文学の爆発的な発展を引き起こしたことを思い出せば、今回の発見によって我々が宇宙の新しい観測手段を手に入れた興奮を共有してもらえらるだろう。

☆2 MySQL を利用し、Web インターフェースを備える。一般公開はされていない。

☆3 物理的意義については日本物理学会誌 2016 年 4 月号に掲載される中村氏と田越氏の解説や、LIGO Open Science Center (<https://losc.ligo.org/about/>) などを参照。ただしあまりにもきれいな信号であったため、筆者の 1 人である Cannon は最初、人為的にテストのために入れた偽の信号だと疑っていた。なぜ疑うのか、Harry Collins の "Gravity's Ghost and Big Dog" を一読されたい。

☆4 試験観測とはいえ、サイエンスに足る質のデータである。ただし、エンジニアリングランであったことから、GraceDb から電磁波観測網への通報は当日マニュアルで行われた。詳細は省くが、電磁波との同時観測は重力波源の物理を探る上できわめて重要である。今回も GW150914 発生直後に Fermi 衛星がガンマ線を検出しており、BH 連星合体が発見以来 40 年以上正体不明であるショートガンマ線バーストを引き起こしている可能性が議論されている。

☆5 ノイズを誤って信号として検出してしまう確率のこと。

☆6 wavelet の一種である Wilson-Daubechies-Meyler 変換を使い、wavelet の係数を調整することによって BH 連星合体にも対応できるようにしている。一般相対論を仮定した精密な理論波形を利用するわけではないので重力波源の方向決定精度など、パラメータ決定精度は低い。

☆7 <https://wiki.ligo.org/DASWG/GstLAL>

☆8 <http://ligo-cbc.github.io/pycbc/latest/html/>

らに正確に決定できる。信号到来時はBH探査を行っていなかったが、これらのパイプラインは2日後にGW150914がBH連星合体からの信号であることを確認した。特に、GStreamer<sup>☆9</sup>を利用したgstlalは高速解析パイプラインであり、信号到達から50秒以内にGraceDbに速報を出すことができる。

## LIGOが活用する情報処理技術

マッチドフィルタ法では、質量などのいくつかのパラメータのさまざまな値について、フィルタを生成しデータに適用する。gstlalではそのフィルタ数は約25万にもなる。オンラインでこの膨大なフィルタを生成、データを解析し、電磁波観測網に速報を出すために、gstlalは、特異値分解、マルチレートフィルタリングなどの手法を活用している。なおgstlalを含むLSCのコードの多くはC/C++、Python、Matlab、ROOT<sup>☆10</sup>等で書かれ、GNU Public Licenseなどのもと、一般に公開されている<sup>☆11</sup>。

必要となる計算機資源もそれなりにあり、gstlalはカリフォルニア工科大学(Caltech)の計算機クラスターのうち1,000コアを利用している。実際、重力波天文学は強力な計算機資源が必要な分野である<sup>☆12</sup>。LSCはHanford, Livingstonの検出器サイト、Wisconsin-Milwaukee大学、Caltech、ドイツ・HannoverのAlbert Einstein Institute(AEI)<sup>☆13</sup>などに計28,000コアほどの計算機資源を擁する<sup>☆14</sup>。ジョブ管理システムとしては、HT-Condor/DAGMan<sup>☆15</sup>が多くの場合採用されている。ほかに、分散コンピューティングプラットフォームであるBerkeley Open Infrastructure for Network Computing<sup>☆16</sup>を利

用したEinstein@Home<sup>☆17</sup>がLSCのプロジェクトであることをご存知の方もいるだろう。

重力波データといくつかの重要なデータは低遅延解析を行う世界各地に高速に転送される一方、20万を超える検出器診断チャンネルのデータはCaltechにいったん集められ(データ量は10MB/s超)、5PBのテープストレージに保存される。メタデータとよく使うデータは旧Sun Microsystemsが開発したZFSを利用したHDDアレイにキャッシュされる。Caltechからは、LSCの研究者、またLSCと共同研究を行っている欧州のVirgo重力波検出器<sup>☆18</sup>の研究者など、世界各地の共同研究者へ転送される。データを誤りや見落としなく各地へ配布するために、LSCではgridミドルウェアを利用している<sup>☆19</sup>。またデータ転送前後でファイルおよび時系列データのチェックサムを確認・記録し、さらに誰がどのファイルを転送し、受領したかを記録している。

以上駆け足だったが、LSCが、重力波の検出と、KAGRAやVirgoなどによる国際重力波観測網の充実によって今後の展開が期待される重力波天文学創成のために、情報処理技術を、ソフトウェア・ハードウェアにおいて重視・活用していたことを感じ取ってもらえれば幸いである。さて、ひるがえって日本の状況はどうなっているのだろうか? 次章では、日本の重力波検出プロジェクトKAGRAの解説と、情報処理技術に関連した取り組みを解説する。

## 大型低温重力波望遠鏡KAGRAとその情報処理技術

大型低温重力波望遠鏡KAGRA(愛称を公募し選考委員(小川洋子選考委員長)により愛称KAGRAに決定)は、腕の長さが3kmのレーザ干渉計型検

☆9 <https://gstreamer.freedesktop.org> GStreamerはオープンソースのオーディオ・ビデオストリーミングライブラリ。

☆10 <https://root.cern.ch>

☆11 <https://www.lsc-group.phys.uwm.edu/daswg/>

☆12 計算機資源が必要な対象としては、低質量X線連星という天体がある。2008年には、その検出には当時のLSCの計算機資源の少なくとも50倍の能力が必要と報告された。

☆13 [http://www.aei.mpg.de/96456/20\\_Computer\\_Clusters](http://www.aei.mpg.de/96456/20_Computer_Clusters) AEIのAtlasクラスターは約3,000のCPU(計約15,000コア)、約1,000のGPGPU、5PBのHDDと4.5PBのテープストレージを備える。

☆14 小規模なクラスターを除く。

☆15 <https://research.cs.wisc.edu/htcondor/>

☆16 BOINCはSETI@Homeを支えるプラットフォームとして開発が始まった。URLは<https://boinc.berkeley.edu/>

☆17 <http://www.einsteinathome.org>

☆18 <https://www.ego-gw.it/index.aspx>

☆19 ただし、データがどこにあるのか、計算機資源がどこにあるのかかわらないというデータグリッド、コンピューティンググリッドは利用していない。あくまで、gridはバルクデータの転送に利用し、解析はローカルにあるデータをローカルにある計算機資源で行う。またある割合でファイルが壊れていることがあり、宇宙線によるものと考えられている。



図-1 大型低温重力波望遠鏡 KAGRA (図:東京大学宇宙線研究所)

出器であり、ニュートリノ観測装置スーパーカムイオカンデと同じ岐阜県飛騨市神岡の鉱山内に建設を進めている(図-1)。KAGRAは坑内(地下)に建設することで地面振動の影響を抑え、また、熱雑音の影響を避けるため鏡とそれを懸架する振り子の低温化を行うという、他国の検出器にはない特徴を持つ。まず常温で基本的な光学的構成を持つ干渉計を完成させ、2016年3月中にテスト観測を開始する予定である(この原稿は3月上旬に書いている)。この観測では、後に述べる一連のデータ転送作業や解析パイプラインなどが問題なく動作するかなどを検証する。その後、さらにアップグレード作業を行い、鏡を低温(20K)に冷やし、2017年度中に運転を開始することを目指している。

KAGRAにおけるデータの流れの概要を図-2に示す。坑内に存在する検出器データや検出器の状態を把握するための各種センサなどからの環境データを常時取得し、そのデータ(raw data)はファイルに記録され、図-2のDetectorのシステム内のデータ保管場所にファイルが一定間隔で生成される。データ転送・保管システムが、そのファイルの生成を検知し、検知したファイルの転送を図-2の流れに沿うかたちで順次開始する。KAGRAでは検出器(Detector)で取得したデータ(raw data)を神岡坑外解析棟に設置され

ているシステムに転送し、神岡坑外のシステムはraw dataを解析用のデータ(processed data. Proc. dataとも略される)に変換する。raw dataとproc. dataは、神岡坑外のシステムから東京大学宇宙線研究所(柏キャンパス)に設置されているシステムに転送され長期間保管される。これらのデータ転送については、Socketを用いて行い、データ転送の要求速度は20MB/sである。また、5年間の稼働におけるデータ容量の想定値は3PiBである。さらに、神岡坑外から大阪市立大学などの解析を行う各機関のシステムへの転送も行う。データの損失や大きな遅延がないように処理・転送する必要があり、データ処理・転送の稼働状態を把握するためのログを記録・管理することも重要である。これらのデータ転送・保管システムはC言語やLinux inotify, syslog-ngd, SharedMemoryなどを駆使して開発が進められている。また、海外のグループとのデータのやりとりやデータの冗長性を保つためのmirrorサイトへのデータ転送は、LSCと同じくgridミドルウェアを用いる可能性を模索しており、gridミドルウェアであるGlobus Toolkit<sup>☆20</sup>を用いて実際にgrid環境を構築しそのテストも進めている。

ここで、KAGRAの現在の主な計算機資源<sup>☆21</sup>について触れておく(図-3)。まず、神岡坑内に20TiBのデータspool用ストレージが2台、神岡坑外解析棟に200TiBのストレージ、データ転送のためのサーバ、および、データ処理用の計算サーバ計64コアが設置されている。また、東京大学宇宙線研究所(柏キャンパス)に100TiBのストレージが設置されている。LSCのgstlalのような即時検出・速報を目的とした高速データ解析のために、大阪市立大学に、288TiBのストレージ、および、データ処理用の計算サーバ計392コアを保有している。これら神岡-柏-大阪市立大学間の各システムは仮想プライベートネットワーク(VPN)によって接続されている。なお、これらの計算機資源は、当然十分なものではなく順次追加され増強されていく予定である。このほかにも東京大学宇宙線研究所の共同利用計算機の利用も念頭においている。

☆20 <http://toolkit.globus.org/toolkit/>

☆21 データ転送・保管システム、データ解析に使用できるもの。

データ解析ソフトウェアについては、LIGO の話でも紹介しているが、LSC が開発した非常に大きなソフトウェア群がある。分かりやすいエラーハンドリングやバグトラッキングシステム、巧妙なメモリマネジメントが実装されている一方で、非常に複雑な構造をしており、ユーザに徹するなら問題はないが、新しい解析方法の開発とその実装を行うために、新しい関数やルーチンを追加するのは我々には容易ではない。そこで、LSC が開発したソフトウェア群を参考にすることが依存しない、KAGRA 独自の解析ソフトウェアライブラリ KAGRA Algorithmic Library (KAGALI) を開発している。また、KAGALI だけでなく必要であれば LSC が開発したソフトウェアなどを利用し、それだけで動く重力波データ解析アプリケーション・ソフトウェア KAGALI-Apps も開発している。主に C 言語を用いて開発を進めているが、Python やライブラリ開発ツールでは Perl

も用いている (Haskell を用いての解析ソフトウェアライブラリ開発も進めており、それについては後に紹介する) KAGALI は、さまざまなシステムでインストール・動作できることを基本とし、現在のところ Red Hat Enterprise Linux V6.x (Scientific Linux, CentOS を含む)、および、Mac OS X において、GNU C コンパイラ (gcc)、または、Intel C コンパイラ (icc) で動作確認をしている。

開発するにあたり、バグが入り込む可能性を減らす、デバッグを容易にする、共有財産を作ることでデータ解析チームのコード作成の効率化を図る、仕様を統一することで再利用しやすくする、などのた

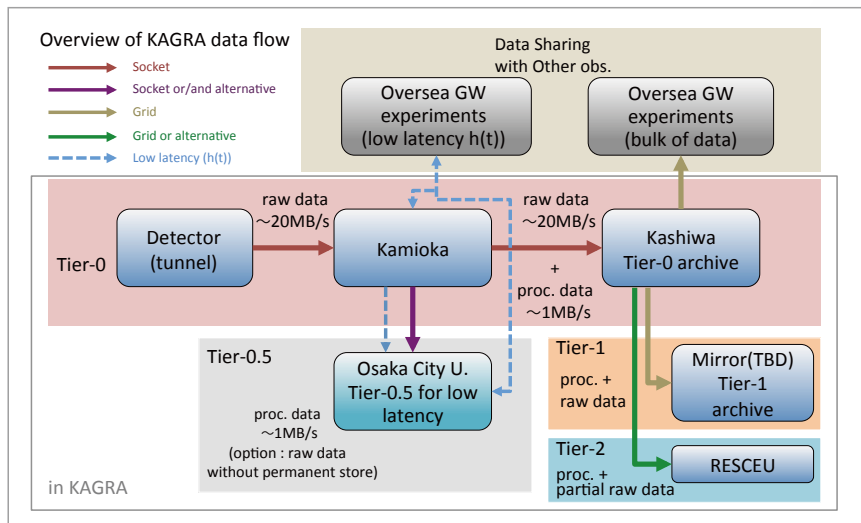


図-2 KAGRA におけるデータの流れ



図-3 KAGRA の現在の主な計算機資源 (データ転送・保管システム, データ解析関連)

めに、コーディングスタイルガイドラインを作成した。バージョン管理には git を利用し、GNU Auto-tools (autoconf/automake/libtools) を用いている。ソースがサーバに push されたら毎晩自動的にチェックするなどソースコードのエラーハンドリングやバグトラッキング、メモリリーク発見などのためのさまざまな仕組みが導入されている<sup>☆22</sup>。開発が現在進行中 (pre-alpha バージョン) であり、2016 年

<sup>☆22</sup> Git や jenkins のセットアップには ALminium (<https://github.com/aluminium/aluminium>) を利用させていただいている。  
 Jenkins : <https://jenkins-ci.org>  
 cppcheck : <http://cppcheck.sourceforge.net/>  
 Valgrind : <http://valgrind.org>

3月に始まるテスト観測のデータ解析の一部で利用予定である。また、2017年度のKAGRAの運転までに最初のバージョンをリリースすることも目標に開発を進めていく。

## 検出器特性診断

地上のレーザ干渉計型重力波望遠鏡は、光を2つに分けるビームスプリッターとその光を打ち返す鏡で干渉させる。両腕の光路長が重力波により変わる。したがって光路長を変えるありとあらゆるものは雑音源となる。代表的なものは地面振動である。数10Hz以下という低周波で突発的に混入する地面振動は、超新星爆発からの重力波のように波形の不定性が高い重力波信号や、今回のブラックホール連星合体のような特徴的な継続時間が短い重力波信号との区別が非常に難しい。また地面振動が問題となる低周波は望遠鏡の制御帯域とも重なっているため、制御エラーによる非定常雑音や、望遠鏡の安定運転にも影響を与える。さらに、光学台の共振周波数を励起し散乱光雑音として数100Hzという望遠鏡感度が最も良い周波数帯域にアップコンバージョンされて混入するなど、地面振動はさまざまに変化して重力波検出を妨げる。気象も観測に大きな影響を及ぼす。LIGO Hanfordは砂漠にあるために、気圧変動による低周波地面振動が安定運転を困難にした。また、雷の放電などによる電離層の震動は、7.8Hzおよびその倍波に緩やかなピークを持つSchumann共振として観測される。そのコヒーレンス長は数1000kmにも及ぶため、複数の望遠鏡に同じような影響を与え、ある種の重力波の検出を困難にする。KAGRAは池ノ山と呼ばれる、水を多く含む山の地下に建設されているために、春季の雪解け水による重力勾配雑音も懸念されている。また望遠鏡の構造に起因した雑音も問題となる。たとえばミラーを吊るすワイヤの共振周波数は鋭い狭帯域信号となり、周辺の感度を大きく悪化させる。鏡の位置を制御するために鏡にコイルマグネットアクチュエータを取り付けるが、そのマグネットが環境磁場

の影響を受け不連続に磁化が進み、バルクハウゼンノイズを生み出す。その他電源雑音、突発的に現れる回路雑音、飛行機やヘリコプターなどの音響雑音も観測に影響を及ぼす。最近のaLIGOの観測では複数の電圧制御発振器や電波発信機がビートして観測帯域内にダウンコンバージョンして混入してくる、数kHzの周波数変動をする「口笛雑音」もそのカップリング経路が不明で対処が問題となった。

こうした雑音の影響を特定して、データから排除していくことで重力波の信頼性の高い検出が可能になる。aLIGOでは環境雑音を把握し、その重力波データへの混入経路を把握するため、20万もの検出器診断チャンネルを駆使している。KAGRAでも環境雑音を把握する同様なチャンネルが用意される予定である。我々はこの多チャンネルデータを詳細に解析する情報処理技術が必要になる。多くの環境外乱は、重力波チャンネルを含め、複数のチャンネルで検出される。そこで多変量解析が非常に強い武器になる。また、チャンネル間の相関を調べるために膨大な組合せのコヒーレンスや線形相関、非線形相関を調べることも連続して行う。データのスペクトルの変動や、突発的な現象を調べるためにはデータを時間周波数空間上に展開し、そのチェンジポイントを検出するモニタ、検出器の不具合によって現れる発振を捉えるための、狭帯域信号モニタ、データのパワーを帯域ごとに調べるモニタ、データのガウス性を調べるモニタなど多くのモニタを膨大なチャンネルデータに対して24時間走らせる必要がある。

## Haskell

さまざまな統計的性質を24時間安定してモニタできる多くのソフトウェアの開発を、小人数で短期間で行わなければならない。それを実現するために、少々逆説的ではあるが、我々は膨大な過去の遺産を持つ動的なプログラミング言語をメインに使うことを諦めた。その代わりに、強い静的型付け言語であるHaskellを導入した。重力波の業界ではC、C++、Pythonなどが中心に使われており、Haskellをブ

プロジェクトで使うものは今までいなかったため、相当の覚悟が必要であったし、それは今でも続いている。しかし、プロジェクトが走り始めて2年、良い兆候がいくつも見られている。まずプロジェクトに参加する学生にとって珍しいこともあり、強い学習意欲が見られ、そのため週1回の勉強会を経て2カ月もすれば拙いながらもさまざまな有用な関数を書けるようになった。この学習意欲の向上が最も良かった点だと考えている。また、多くの者にとって初めてのものなので、プログラムのコーディング作法をそるえることは比較的容易であった。コンパイル時にエラーを捉え、可能な限りランタイムエラーを避けることが可能となり、Haskellは開発時間の短縮に寄与している。関数をベースとすることで、副作用をあまり気にせずに安全にコードを再利用できるようになった。統計解析やデータ処理のコードは多くの部分が類似しているために、このことは特に有効だった。また関数の使い方が型を見ることによって容易に推定できることはプロジェクトの推進力になった。全体的にコードが短くなったことで開発者がプロジェクト全体を把握しやすくなった。速度が重要な場面、メモリを多く使う部分や、重力波業界特有のライブラリを使わざるを得ない部分ももちろんあり、その場合は、Foreign Function Interfaceを用いることで、非常に容易にCライブラリを呼び出せることも便利である。導入して良い点はあるが、もちろん（潜在的な）問題も含んでいるので、これからも油断はできない。たとえば、Haskellはコンパイラ、モジュールの開発速度が比較的早いために、ふと気づいたときには今まで使っていた関数が別のモジュールに移っていたり、依存関係が壊れることがある。こうした問題はstackと呼ばれる新しいビルドツールである程度緩和されてきてはいるが、十分に気を使わなければならない。また、遅延評価のためにメモリの使用量が大きい。最近のコンパイラでは正格評価をデフォルトにできるプラグマも導入されてきてはいるがやはり注意すべきだろう。低温干渉計である、地下に設置されているという、重力波業界に今までなかった特徴を持った重力波望

遠鏡の診断ツールライブラリを、この業界で今まで使われてこなかったHaskellを用いて実装した。ライブラリ名はHasKALである。KAGRA観測が始まろうとしているこれからHasKAL/Haskellが活躍する。

## 重力波天文学の幕開け

重力波源の位置は複数の検出器で重力波を同時観測し、その到着時間の差より導きだされる。GW150914においては、aLIGOの2台の検出器が検出した時間差より、その位置の決定精度は600平方度程度の南天ということしか分からなかった。重力波源の方向をより正確に特定するには少なくとも3台、できれば4台以上の検出器が必要である。そのため、KAGRAが世界の検出器と重力波観測ネットワークを形成し、重力波観測を行うことは重力波天文学の創成に向け非常に意義が大きい。

我々は、ちょうど今、重力波天文学が始まるスタートラインに立ったといえる。さあ、重力波天文学の幕開けです。

(2016年3月10日受付)

Kipp Cannon kipp@resceu.s.u-tokyo.ac.jp

2016年2月より東京大学ビッグバン宇宙国際研究センターに准教授（KAGRA重力波データ解析国際協力部門）として着任。LIGO Scientific CollaborationおよびKAGRA Collaborationのメンバで、今回の発見で大きな役割を果たしたgstlalの作者の1人。

端山和大 hayama@icrr.u-tokyo.ac.jp

東京大学宇宙線研究所特任助教。神岡で研究を行う。かつてLIGOに所属し検出器特性の研究や、バースト性重力波の探索を行っていた。現在はKAGRA検出器特性解析チームのトップでありまたKAGRAデータ解析チームに所属し、バースト性重力波探索、マルチメッセンジャー解析、検出器特性の研究と診断方法についての研究を行う。KAGRAにおけるHaskell伝導者。

伊藤洋介 yosuke\_ito@resceu.s.u-tokyo.ac.jp

東京大学ビッグバン宇宙国際研究センター重力波データ解析国際協力部門特任助教。KAGRAデータ解析チームに所属し、回転する中性子星からの重力波検出を目論む。Einstein@Home立ち上げ時LIGOに所属し、Einstein@Homeのポスト解析に携わった。

高橋弘毅（正会員）hirotaka@kjs.nagaokaut.ac.jp

長岡技術科学大学大学院工学研究科情報・経営システム工学専攻准教授。かつてLSCにおいて連星合体からの重力波を探索するチームに所属し、特に日本の前世代の重力波検出器TAMA300とinitial LIGOとの共同観測にかかわった。現在はKAGRAデータ解析チームに所属し、データ転送・保管システムの開発やデータ解析手法の開発に携わっている。