GPU use cases in Gravitational Wave Research

Chun-Yu Lin

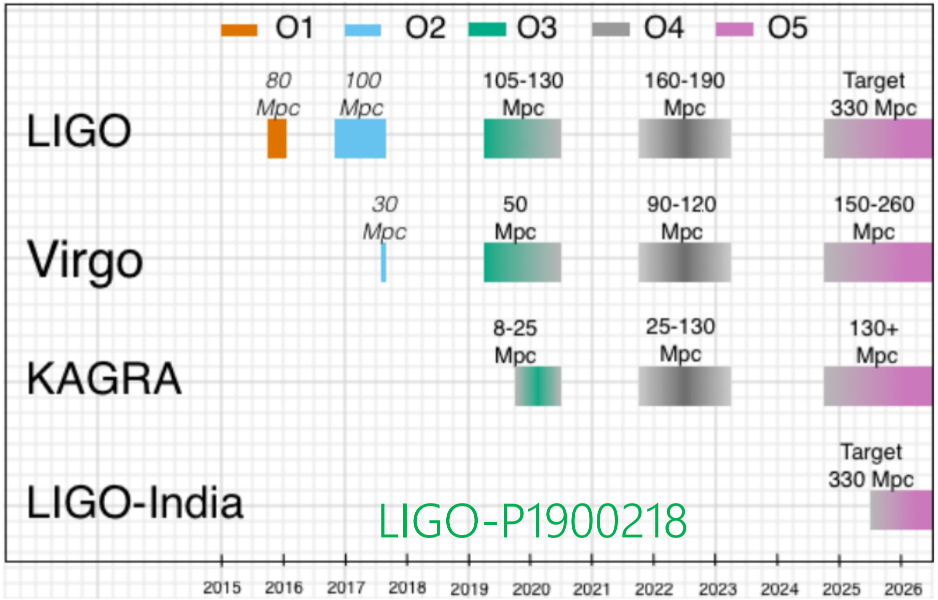
Natl Center for High-performance Computing (NCHC)

Natl Applied Research Labs, Taiwan

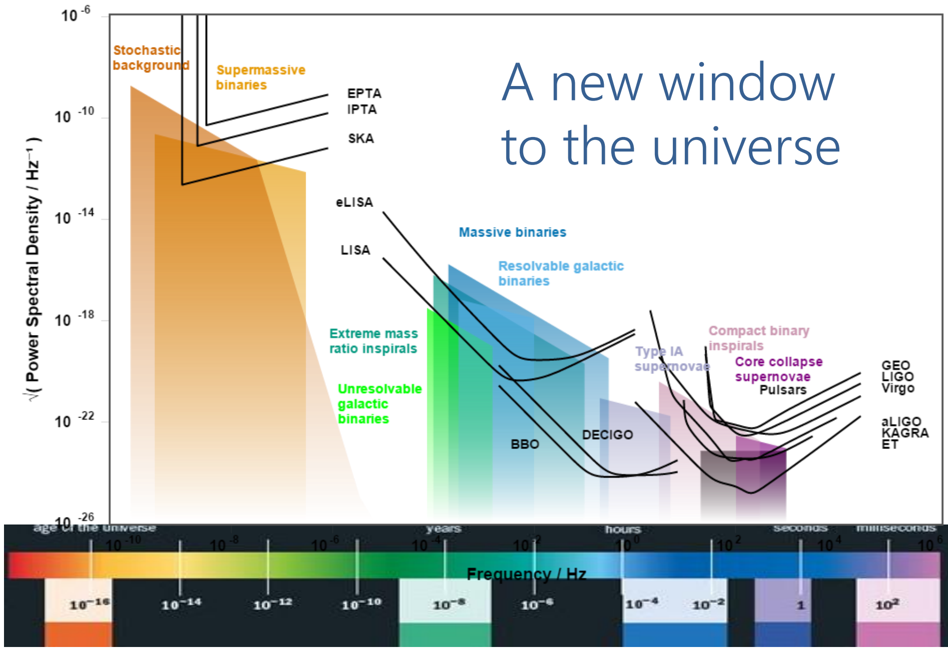
OpenACC annual meeting, Sep 3, 2019

Background

- Global GW detection network LIGO-Virgo made the first detection of GW from the black hole binary (GW150914) and neutron star binary (GW170817), predicted by Einstein 100yr ago.
- They've entered the third observation phase (O3) on April 2019.
- KAGRA project in Japan is also ready to join O3 by 2020.



LIGO-P1900218

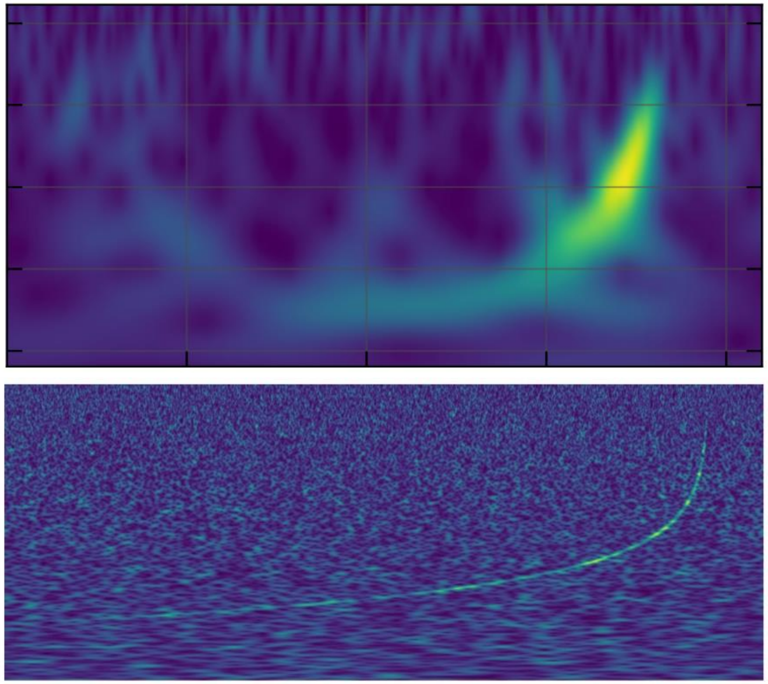


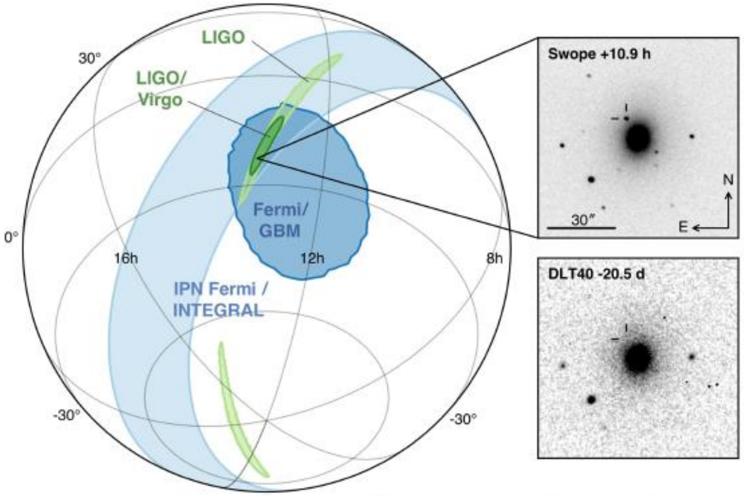
A new window
to the universe

http://rhcole.com/apps/GWplotter/

A new window to the universe

- 1st binary black hole (BBH) event (GW150914): collide into ~60 solar mass BH at ~1 B yr ago, refining astrophysical models, consistent with General Relativity
- 1st binary neutron star (BNS) event (GW170817): Mark the dawn of multi-messenger GW astronomy era, its optical/infrared emission (due to the r-process) explain the major origin of heavy elements like Au, Pb

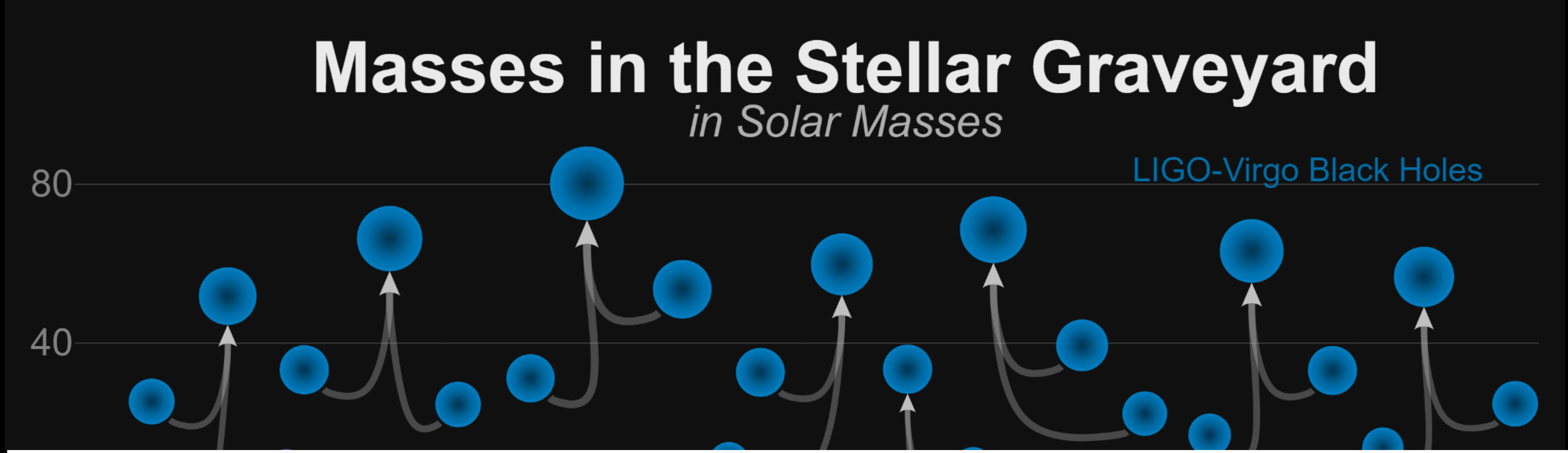




40 Mpc in NGC 4993 (GW170817)

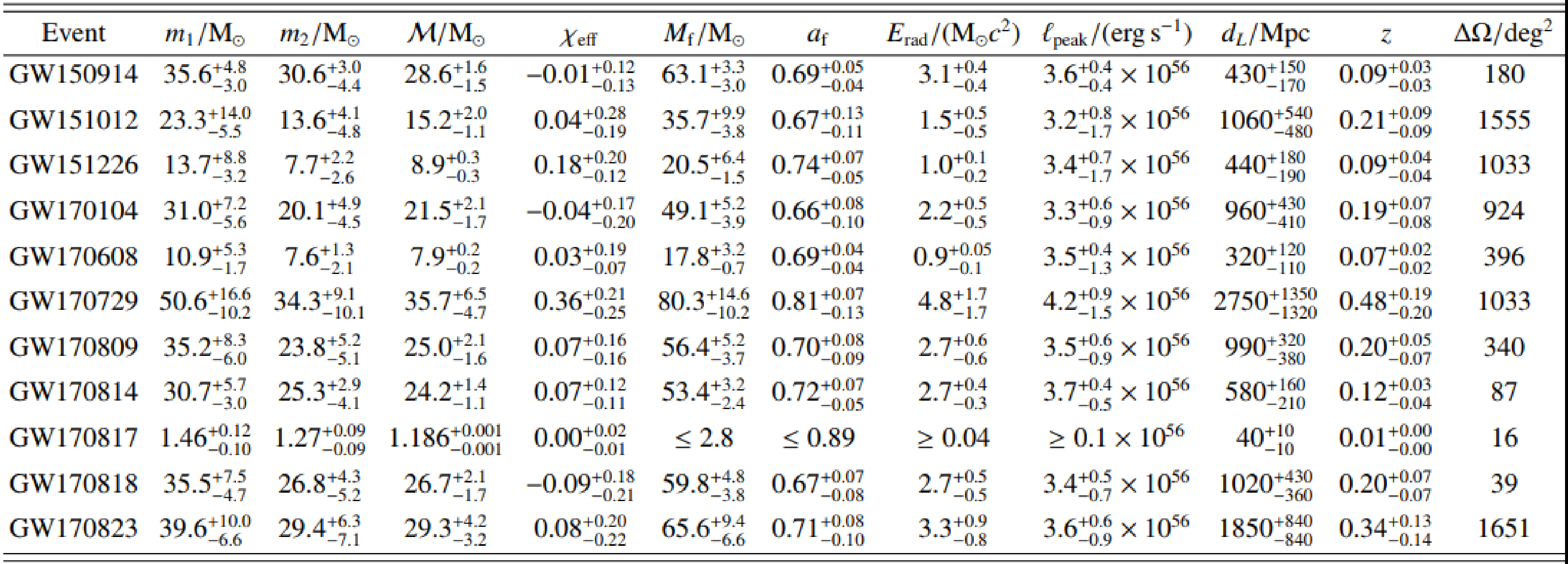
Masses in the Stellar Graveyard

in Solar Masses



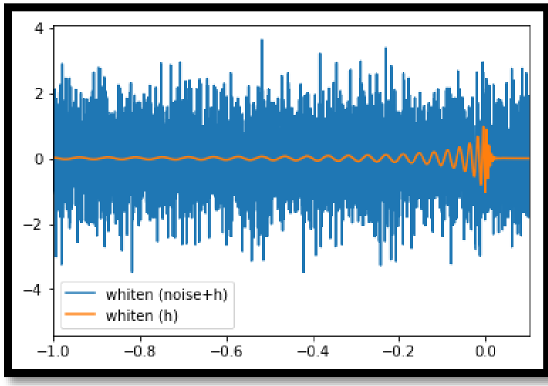
11 events announced (O1/O2)

arXiv:1811.12907



Event	$m_1/\mathrm{M}_\odot$	$m_2/\mathrm{M}_\odot$	$\mathcal{M}/\mathrm{M}_\odot$	χ_eff	$M_\mathrm{f}/\mathrm{M}_\odot$	a_f	$E_\mathrm{rad}/(\mathrm{M}_\odot c^2)$	$\ell_\mathrm{peak}/(\mathrm{erg\,s^{-1}})$	d_L/Mpc	z	$\Delta\Omega/\mathrm{deg}^2$
GW150914	$35.6^{+4.8}_{-3.0}$	$30.6^{+3.0}_{-4.4}$	$28.6^{+1.6}_{-1.5}$	$-0.01^{+0.12}_{-0.13}$	$63.1^{+3.3}_{-3.0}$	$0.69^{+0.05}_{-0.04}$	$3.1^{+0.4}_{-0.4}$	$3.6^{+0.4}_{-0.4}\times10^{56}$	430^{+150}_{-170}	$0.09^{+0.03}_{-0.03}$	180
GW151012	$23.3^{+14.0}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$15.2^{+2.0}_{-1.1}$	$0.04^{+0.28}_{-0.19}$	$35.7^{+9.9}_{-3.8}$	$0.67^{+0.13}_{-0.11}$	$1.5^{+0.5}_{-0.5}$	$3.2^{+0.8}_{-1.7}\times10^{56}$	1060^{+540}_{-480}	$0.21^{+0.09}_{-0.09}$	1555
GW151226	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.6}$	$8.9^{+0.3}_{-0.3}$	$0.18^{+0.20}_{-0.12}$	$20.5^{+6.4}_{-1.5}$	$0.74^{+0.07}_{-0.05}$	$1.0^{+0.1}_{-0.2}$	$3.4^{+0.7}_{-1.7}\times10^{56}$	440^{+180}_{-190}	$0.09^{+0.04}_{-0.04}$	1033
GW170104	$31.0^{+7.2}_{-5.6}$	$20.1^{+4.9}_{-4.5}$	$21.5^{+2.1}_{-1.7}$	$-0.04^{+0.17}_{-0.20}$	$49.1^{+5.2}_{-3.9}$	$0.66^{+0.08}_{-0.10}$	$2.2^{+0.5}_{-0.5}$	$3.3^{+0.6}_{-0.9}\times10^{56}$	960^{+430}_{-410}	$0.19^{+0.07}_{-0.08}$	924
GW170608	$10.9^{+5.3}_{-1.7}$	$7.6^{+1.3}_{-2.1}$	$7.9^{+0.2}_{-0.2}$	$0.03^{+0.19}_{-0.07}$	$17.8^{+3.2}_{-0.7}$	$0.69^{+0.04}_{-0.04}$	$0.9^{+0.05}_{-0.1}$	$3.5^{+0.4}_{-1.3}\times10^{56}$	320^{+120}_{-110}	$0.07^{+0.02}_{-0.02}$	396
GW170729	$50.6^{+16.6}_{-10.2}$	$34.3^{+9.1}_{-10.1}$	$35.7^{+6.5}_{-4.7}$	$0.36^{+0.21}_{-0.25}$	$80.3^{+14.6}_{-10.2}$	$0.81^{+0.07}_{-0.13}$	$4.8^{+1.7}_{-1.7}$	$4.2^{+0.9}_{-1.5}\times10^{56}$	2750^{+1350}_{-1320}	$0.48^{+0.19}_{-0.20}$	1033
GW170809	$35.2^{+8.3}_{-6.0}$	$23.8^{+5.2}_{-5.1}$	$25.0^{+2.1}_{-1.6}$	$0.07^{+0.16}_{-0.16}$	$56.4^{+5.2}_{-3.7}$	$0.70^{+0.08}_{-0.09}$	$2.7^{+0.6}_{-0.6}$	$3.5^{+0.6}_{-0.9}\times10^{56}$	990^{+320}_{-380}	$0.20^{+0.05}_{-0.07}$	340
GW170814	$30.7^{+5.7}_{-3.0}$	$25.3^{+2.9}_{-4.1}$	$24.2^{+1.4}_{-1.1}$	$0.07^{+0.12}_{-0.11}$	$53.4^{+3.2}_{-2.4}$	$0.72^{+0.07}_{-0.05}$	$2.7^{+0.4}_{-0.3}$	$3.7^{+0.4}_{-0.5}\times10^{56}$	580^{+160}_{-210}	$0.12^{+0.03}_{-0.04}$	87
GW170817	$1.46^{+0.12}_{-0.10}$	$1.27^{+0.09}_{-0.09}$	$1.186^{+0.001}_{-0.001}$	$0.00^{+0.02}_{-0.01}$	≤ 2.8	≤ 0.89	≥ 0.04	$\geq 0.1\times10^{56}$	40^{+10}_{-10}	$0.01^{+0.00}_{-0.00}$	16
GW170818	$35.5^{+7.5}_{-4.7}$	$26.8^{+4.3}_{-5.2}$	$26.7^{+2.1}_{-1.7}$	$-0.09^{+0.18}_{-0.21}$	$59.8^{+4.8}_{-3.8}$	$0.67^{+0.07}_{-0.08}$	$2.7^{+0.5}_{-0.5}$	$3.4^{+0.5}_{-0.7}\times10^{56}$	1020^{+430}_{-360}	$0.20^{+0.07}_{-0.07}$	39
GW170823	$39.6^{+10.0}_{-6.6}$	$29.4^{+6.3}_{-7.1}$	$29.3^{+4.2}_{-3.2}$	$0.08^{+0.20}_{-0.22}$	$65.6^{+9.4}_{-6.6}$	$0.71^{+0.08}_{-0.10}$	$3.3^{+0.9}_{-0.8}$	$3.6^{+0.6}_{-0.9}\times10^{56}$	1850^{+840}_{-840}	$0.34^{+0.13}_{-0.14}$	1651

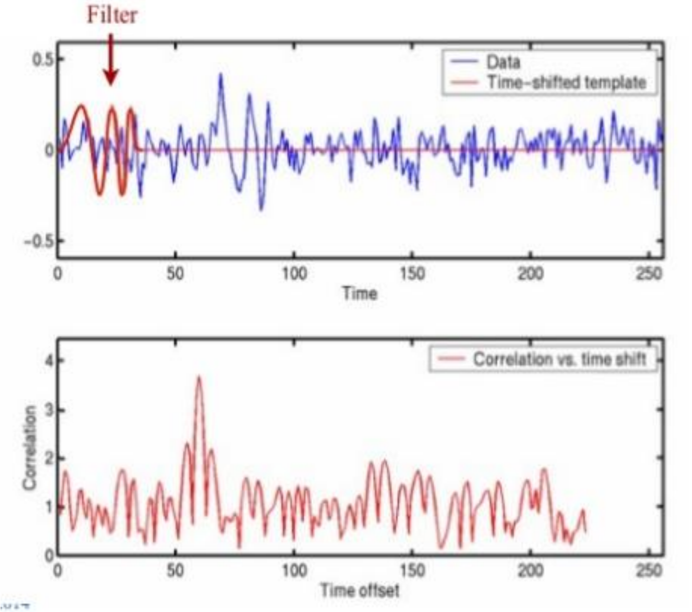
Matched Filtering



- Gravitational wave is weak, burying in noise.
  To identify event from noise channel:

$$s(t) = h(t, \theta^{\mu}) + n(t)$$

- Correlation-based approach by maximizing



$$\mathrm{C(t)} = \int s(t')h(t - t', \theta^{\mu})dt'$$

$$= \int \frac{\tilde{s}(f)\widetilde{h}^{*}(f, \theta^{\mu})}{S_n(f)} e^{2\pi i f t} df$$

with waveform templates

$S_n(f)$: power spectrum density

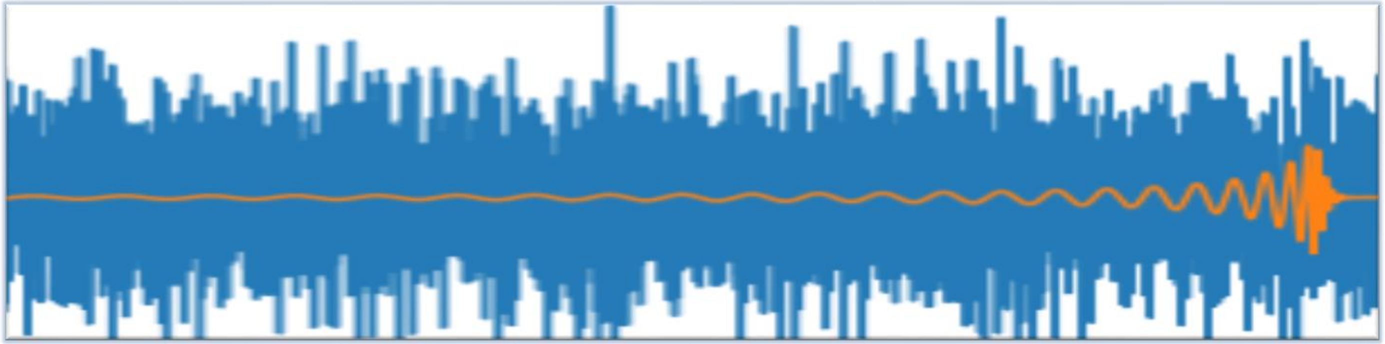
Ilya Mandel (2014)

- Need accurate template -- the first detection use ~250k template !

Computing challenges in GW detection

- Waveform modeling
  - Numerical relativity: Simulate Einstein equation for accurate waveform – Multiphysics PDE with gravity/ matter/ microphysics
  - (Semi-)Analytical waveforms

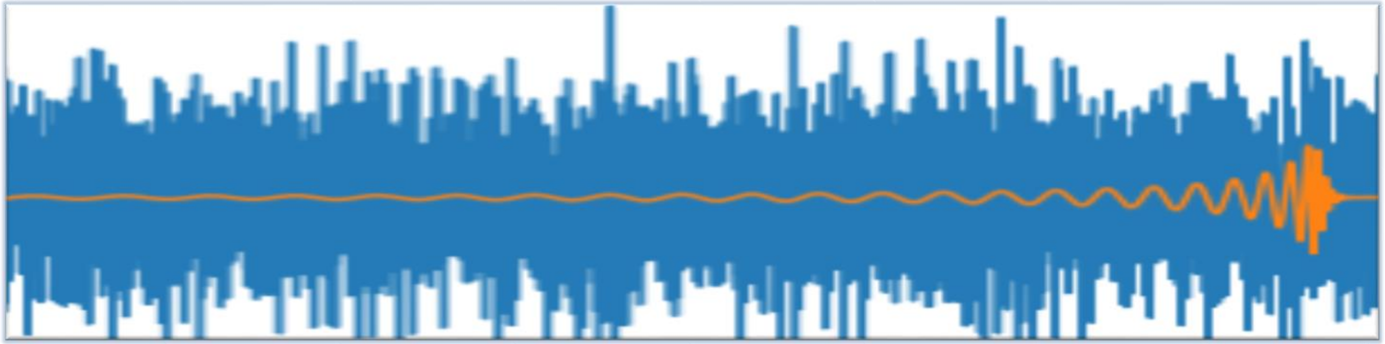
- Data analysis of GW
  - Modeled and un-modeled search
  - Parameter estimation of GW



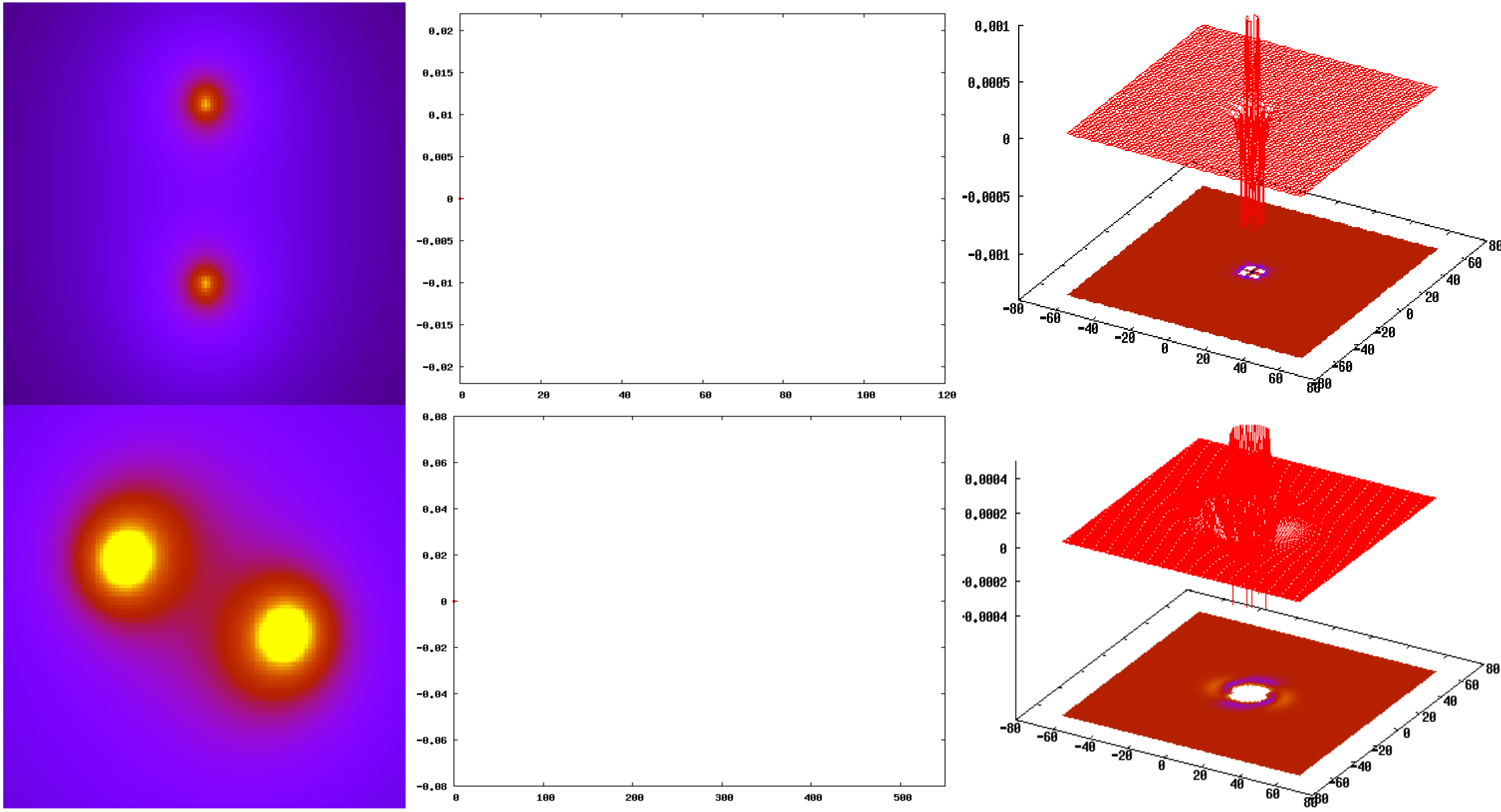
Computing challenges in GW detection

- Waveform modeling
  - Numerical relativity: Simulate Einstein equation for accurate waveform – Multiphysics PDE with gravity/ matter/ microphysics
  - (Semi-)Analytical waveforms

- Data analysis of GW
  - Modeled and un-modeled search
  - Parameter estimation of GW



Binary black hole evolution

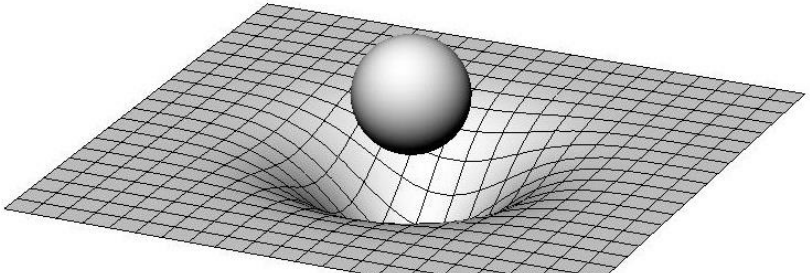


~ O(100 CPU-core) and O(10) days in 2009

Credit: Zhoujian Cao (BNS)

Einstein’s eq of General Relativity

Geometry $G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mathrm{u}\nu}$ Matters



(Hyperbolic-type) Evolution equation :

$$\begin{aligned} \partial_t \gamma_{ij} &= -2\alpha K_{ij} + \pounds_\beta \gamma_{ij} \\ \partial_t K_{ij} &= \alpha \left\{ R_{ij} - {}^{(4)}R_{ij} - 2K_{ik}K^k_j + KK_{ij} \right\} - D_i D_j \alpha + \pounds_\beta K_{ij} \end{aligned}$$

(Elliptic-type) Constraints :

$$R + K^2 - K_{ij}K^{ij} = 16\pi E$$ Hamiltonain constraint

$$D_j(K^{ij} - \gamma^{ij}K) = 8\pi p^i$$ Momentum constraint

Baumgarte-Shapiro-Shibata-Nakamura (BSSN) Formulation

24+ variables

$$
\begin{aligned}
\phi &= \frac{1}{12}\ln\gamma \\
\tilde{\gamma}_{ij} &= e^{-4\phi}\gamma_{ij} \\
K &= \gamma^{ij}K_{ij} \\
\tilde{A}_{ij} &= e^{-4\phi}K_{ij}^{TF} \\
\tilde{\Gamma}^i &= \tilde{\gamma}^{ij}\tilde{\Gamma}^i_{jk}
\end{aligned}
$$

Evolution eqs

$$
\begin{aligned}
\frac{d}{dt}\phi &= -\frac{1}{6}\alpha K \quad , \qquad \frac{d}{dt} = \frac{\partial}{\partial t} + \mathcal{L}_\beta \\
\frac{d}{dt}\tilde{\gamma}_{ij} &= -2\alpha\tilde{A}_{ij} \\
\frac{d}{dt}K &= \alpha\left(\tilde{A}^{ij}\tilde{A}_{ij} + \frac{1}{3}K^2\right) - D^2\alpha \\
\frac{d}{dt}\tilde{A}_{ij} &= \alpha\left(K\tilde{A}_{ij} - 2\tilde{A}_{ik}\tilde{A}^k_j\right) + e^{-4\phi}\left(\alpha R_{ij}^{TF} - D_iD_j\alpha^{TF}\right) \\
\frac{\partial}{\partial t}\tilde{\Gamma}^i &= 2\alpha\left(\tilde{\Gamma}^i_{jk}\tilde{A}^{jk} - \frac{2}{3}\tilde{\gamma}^{ij}\partial_jK + 6\tilde{A}^{ij}\partial_j\phi\right) - 2\tilde{A}^{ij}\partial_j\alpha \\
&\quad + \tilde{\gamma}^{jk}\partial_j\partial_k\beta^i + \frac{1}{3}\tilde{\gamma}^{jk}\partial_j\partial_k\beta^k + \beta^j\partial_j\tilde{\Gamma}^i - \tilde{\Gamma}^j\partial_j\beta^i + \frac{2}{3}\tilde{\Gamma}^i\partial_j\beta^j
\end{aligned}
$$

Gauge sector

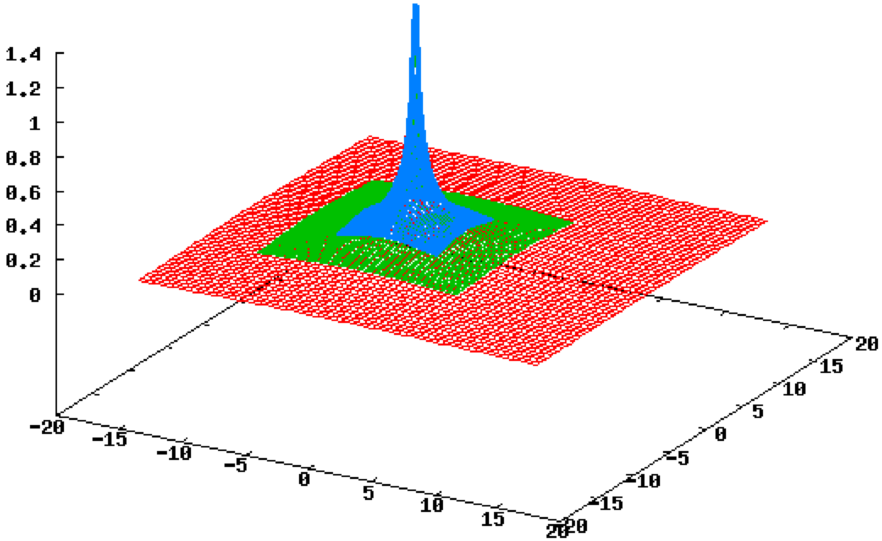
$$
\begin{aligned}
\partial_t\alpha &= D_i\beta^i - \alpha K \\
\partial_t\beta^i &= \frac{3}{4}B^i, \\
\partial_t B^i &= \alpha^2\partial_t\tilde{\Gamma}^i - \eta B^i
\end{aligned}
$$

- BSSN proposed in 1995-1999
- 1st-order in time, 2-nd order in space
- First long-term stable BBH simulation in 2005 (GW150914)
- RHS at each iteration >> O(1,000) ops.

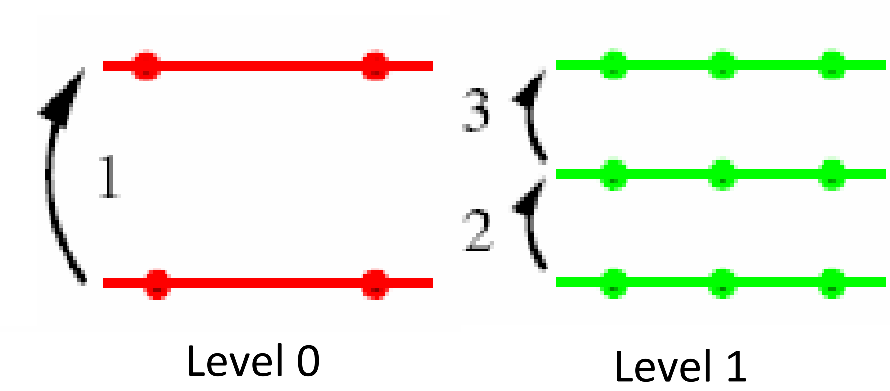
Solve BSSN with FMR with moving box

- Simulate **binary black hole evolution** with simple RK4, 4th order FD, and Berger-Oliger-type fixed mesh refinement (~Chombo AMR)

Level-wise domain decomposition restricts efficient parallelism !



Z Cao+ (BNS)
arXiv:0812.0641

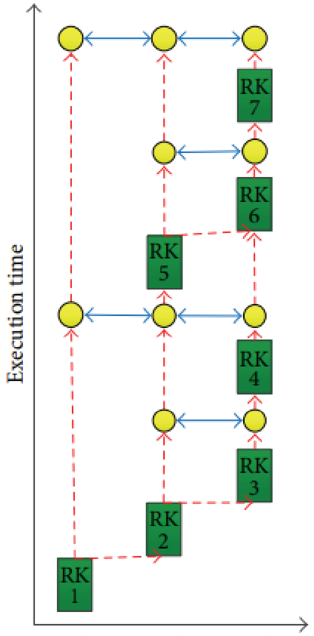


F. Loeffler+ CQG **29**, 115001 (2012)

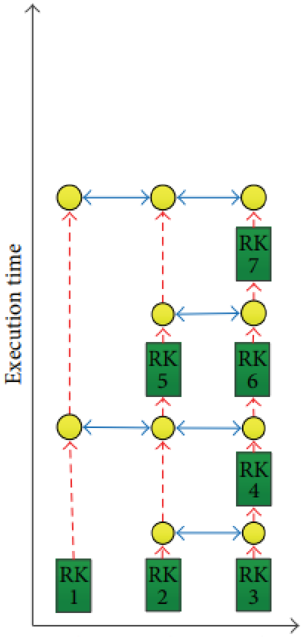
GPU port strategy

AMS-NCKU GPU code
Zhihui Du, Zhoujian Cao+
Sci. Prog. **7** (2016)

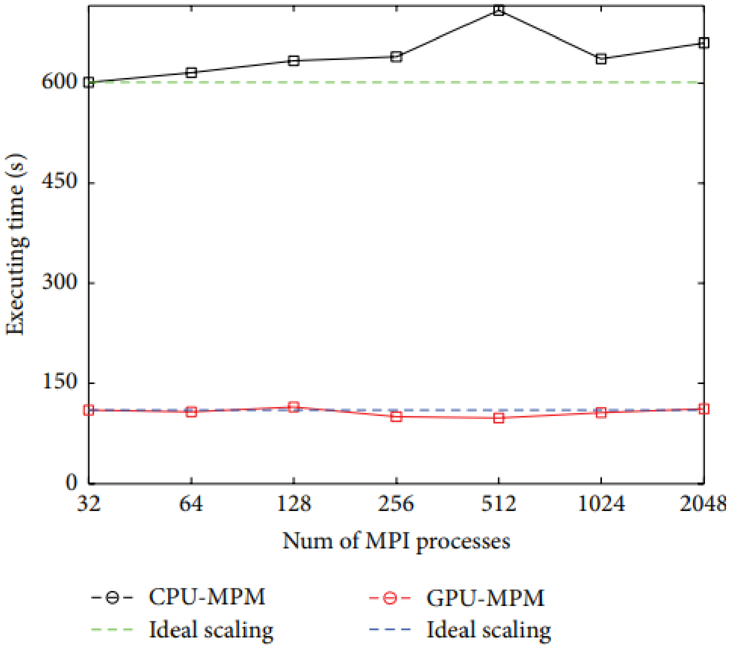
- MPI + OpenMP + CUDA
- Improved mesh-refinement concurrency (~2x)
- Put RHS (> 80%, ~O(5000) ops per grid) in GPU (~5x)
- Typically, overall ~10x with good scaling upto 2k MPI ranks.



(a) SPM executes different mesh levels one by one



(b) MPM executes different mesh levels in parallel

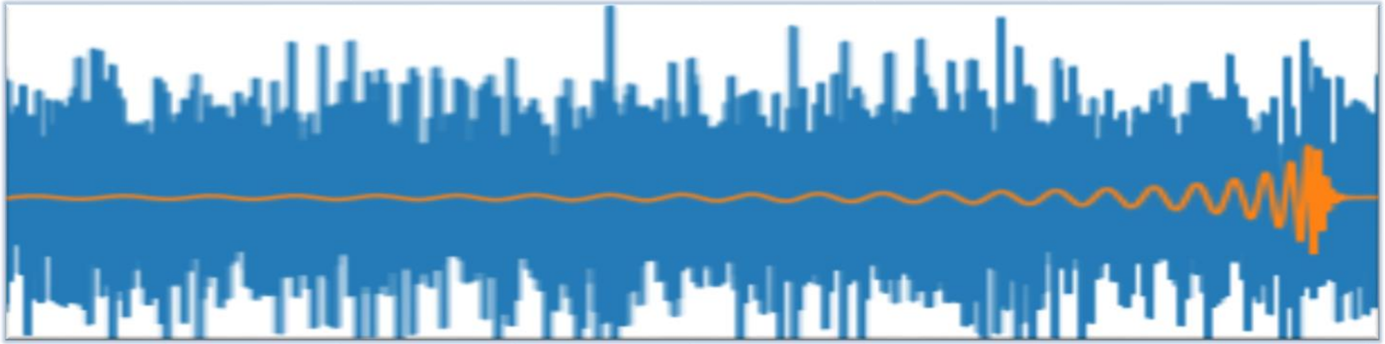


(b) MPM

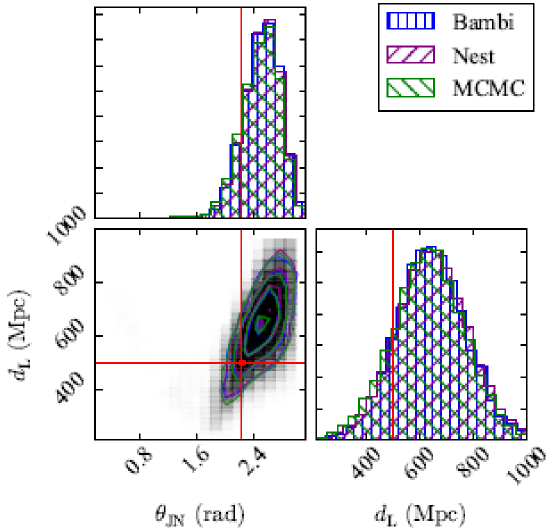
Computing challenges in GW detection

- Waveform modeling
  - Numerical relativity: Simulation of Einstein equation for accurate waveform – Multiphysics PDE with gravity/ matter/ microphysics
  - (Semi-)Analytical waveforms

- Data analysis of GW
  - Modeled and un-modeled search
  - Parameter estimation of GW



Parameter estimation of GW



- To find distribution of source parameter θ_i for detection *d* in Bayesian framework:

Likelihood Prior

posterior

$$P(\theta_i|d) = \frac{P(d|\theta_i) \quad P(\theta_i)}{P(d)}$$

Evidence

V: Prior: astrophysics models / assumptions

V: Likelihood: statistics of observation $P(d|\theta) \propto exp\left[-\frac{(d-h(\theta_i)|d-h(\theta_i))}{2}\right]$

X: Evidence: not easy to get $P(d) = \int P(d|\theta_i)P(\theta_i)d\theta_i$

- How to sample Posterior ∝ Likelihood x Prior ?

Markov chain Monte Carlo (MCMC)

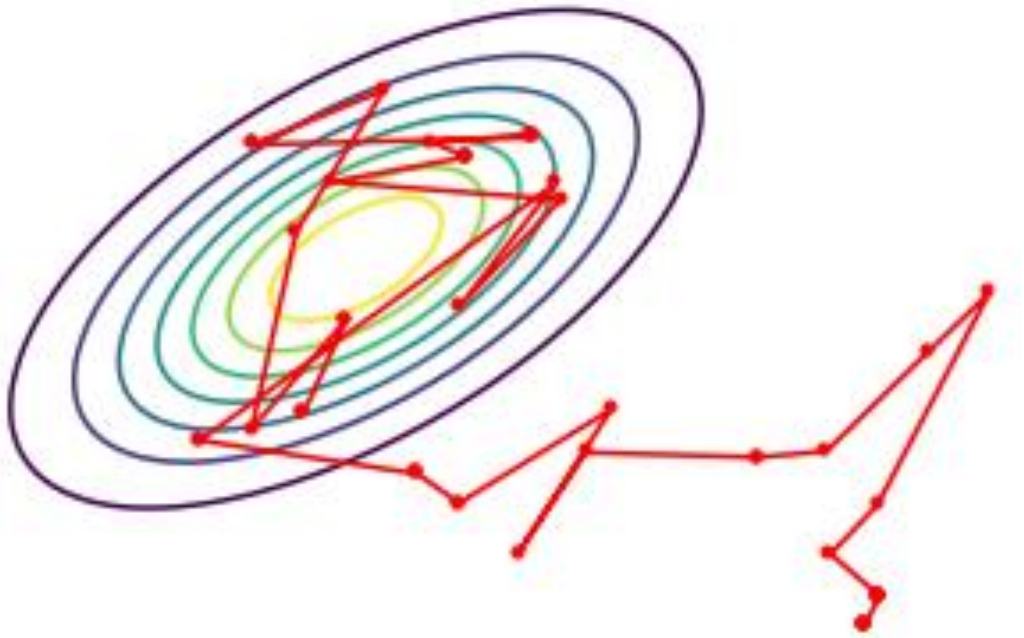
To sample from unknown $P(x)$ by Markov chain with an initial $P^0(x)$ and transition rate $T(x'|x)$.

If the chain exists →

1. It satisfies detailed balance (equilibrium)

$$T(x'|x)P(x) = T(x|x')P(x')$$

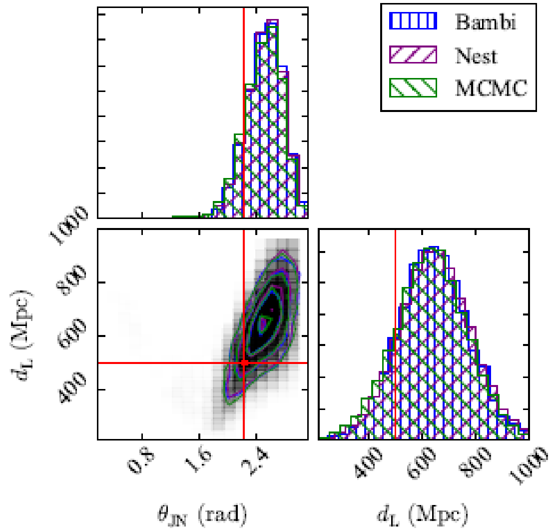
2. It converges to the $P(x)$ starting from any distribution (ergodicity)



MCMC – Metropolis-Hastings (1953/1970)

- Start from a sample in $P(x)$, and draw a next sample in $P(x')$ accepted only by the rate:

$$a = min\left(1, \frac{P(x')}{P(x)}\frac{T(x|x')}{T(x'|x)}\right)$$



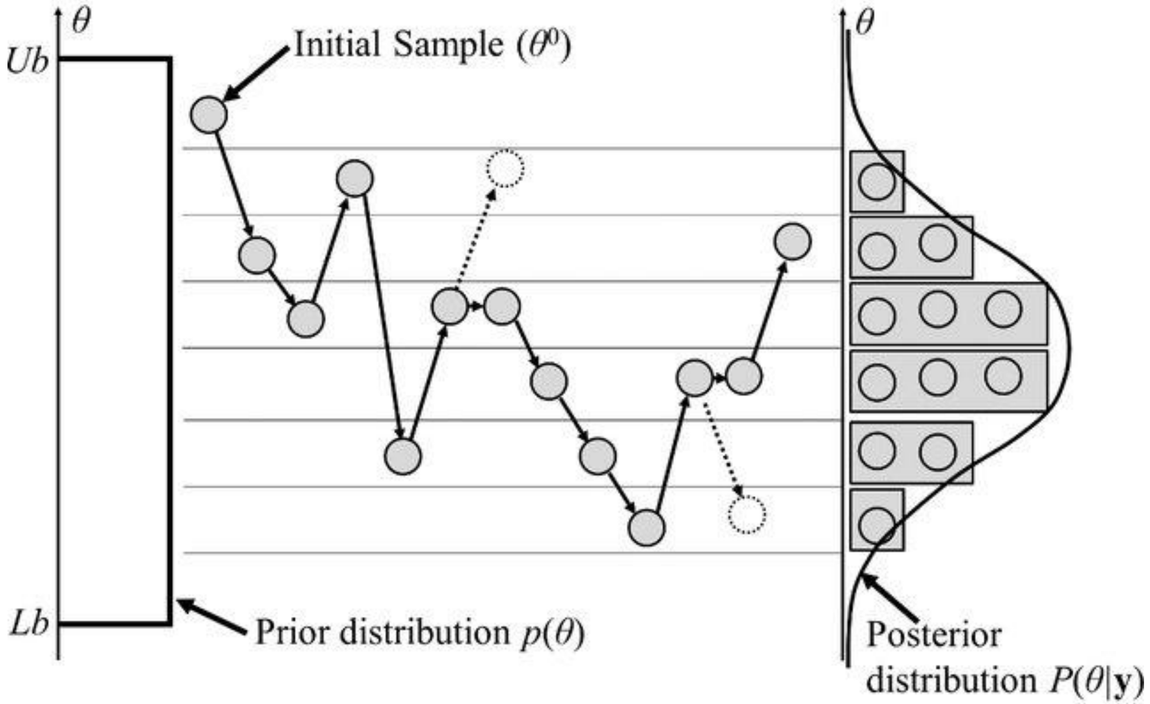


Image credit: Lee+, Energies **8** (2015) 5538

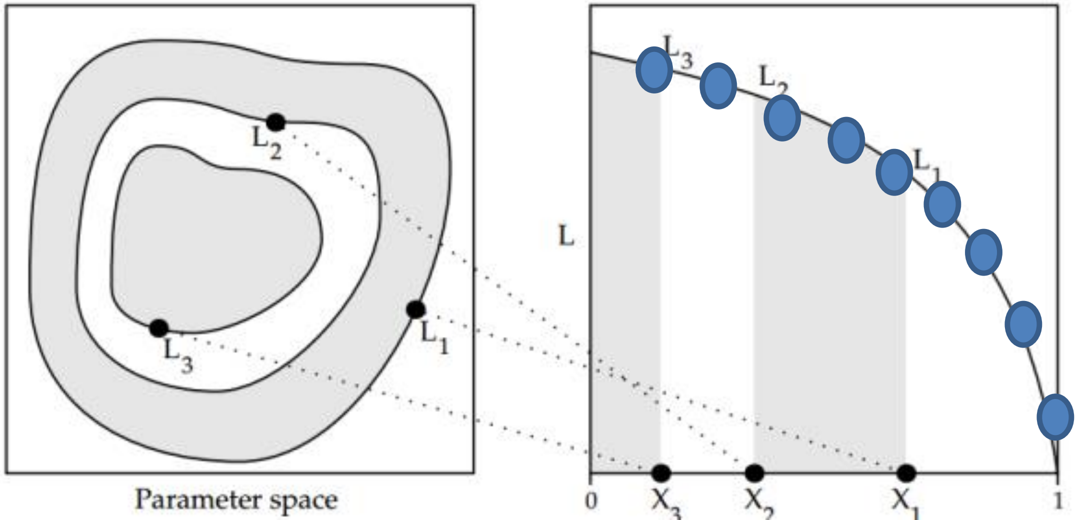
Nested Sampling (2004)

- LIGO's community CPU code : LALSuite
- Computing the **evidence** at the same time, it maps multi-D integral into 1-D integral:

$$Z \equiv \int L(\theta_i)\ P(\theta_i) d\theta_i = \int L\ dX,$$

- Prior mass X: fraction of parameter with $L(\theta) > \lambda$

$$X(\lambda) = \int_{L(\theta)>\lambda} P(\theta_i) d\theta_i$$



Sampling with increasing L

The key is :
$X_i = \exp\left(-\frac{i}{N}\right)$
statistically

John Skilling, 2004

Computing challenges for NS

- Likelihood involves heavy waveform calculation

$$P(d|\theta) \propto exp\left[-\frac{(d-h(\theta_i)|d-h(\theta_i))}{2}\right]$$

Element-wise op.

- **>15 parameters** for waveforms (of circular BH binaries) :
  - mass, spin, location, inclination, polarization, …….

- Typically ~1 day / PE run on single core
  - 10 k Nested sampling iterations / 1 PE run
  - O(1000) waveform / iteration
  - 4 kHz * 8 sec * 4 detectors ~ 128 k ops / waveform
- Cost of (analytical) waveform is model dependant

Core of GPU-accelerated PE code

Credit: Sadakazu Haino (AS)

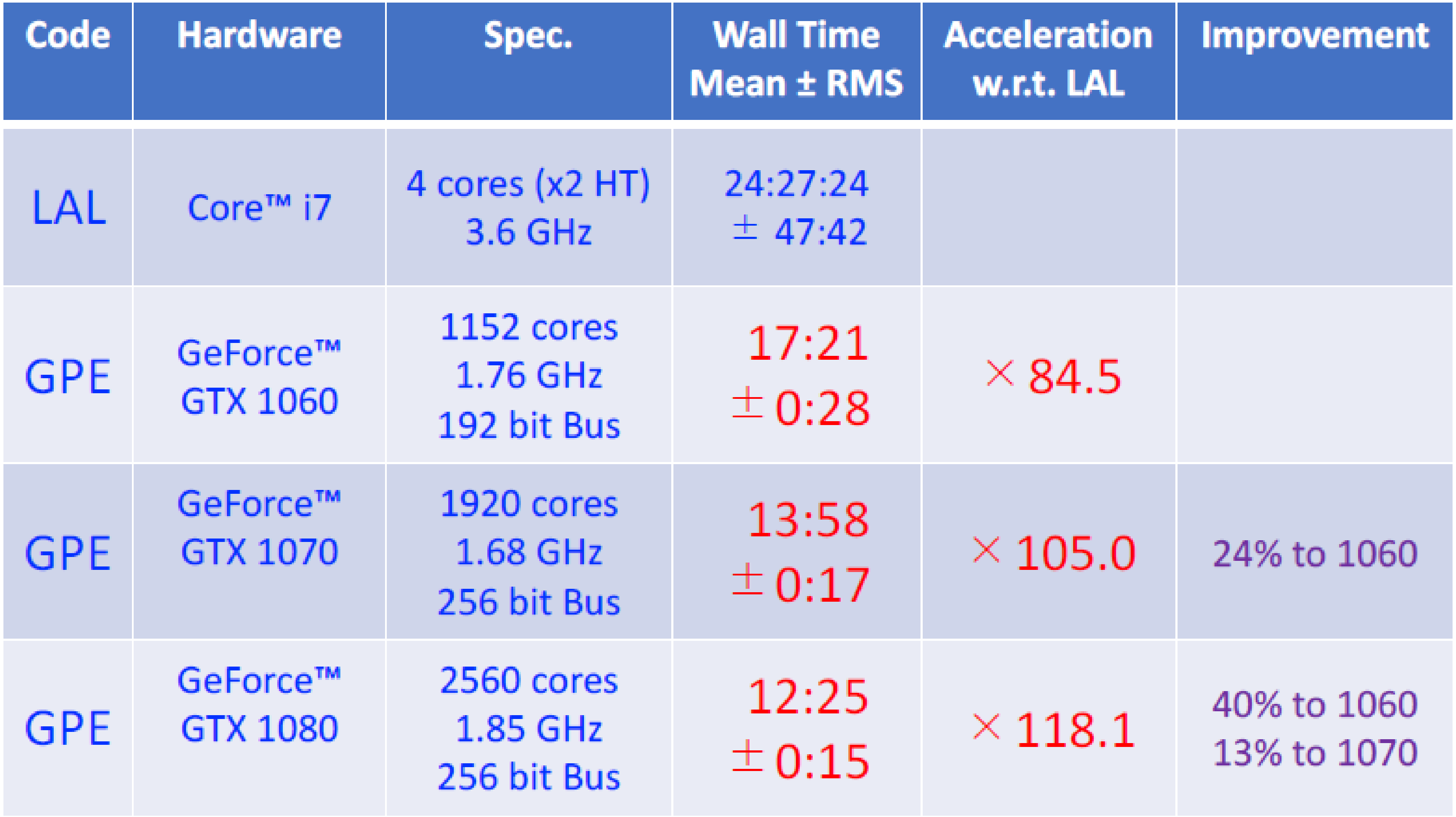
```
__global__ void _geval(. . .)
{
  int i = threadIdx.x;
  int j = blockIdx.x+blockIdx.y*gridDim.x;
  int k = blockIdx.y;
  int x = blockIdx.x*blockDim.x+i;

  __shared__ real cs[BD], ss[BD], dd[BD], rr[BD], ri[BD];
                  cs[i] = ss[i] = dd[i] = rr[i] = ri[i] = 0;

  real   fr = df*x, ph = 0, hpr = 0, hpi = 0, hcr = 0, hci = 0;
  real  czp = *(ppr++), szp = *(ppr++), tc = *(ppr++);

  if (Model == IMRP) cimr(ph, hpr, hpi, hcr, hci, fr);
  else               ctyf(ph, hpr, hpi, hcr, hci, fr);

  int l = k*Nfr+x;
  core(fr, dr[l], di[l], ps[l], hpr, hpi, hcr, hci, czp, szp, tc,
       cer, cei, fp, fc, dt, Slen, cs[i], ss[i], dd[i], rr[i], ri[i]);
  __syncthreads();

  ...
}
```

Waveform $h(\theta_i)$ at each f

Element-wise Likelihood calculation

GPE (GPU-acc. PE) vs. LALSuite

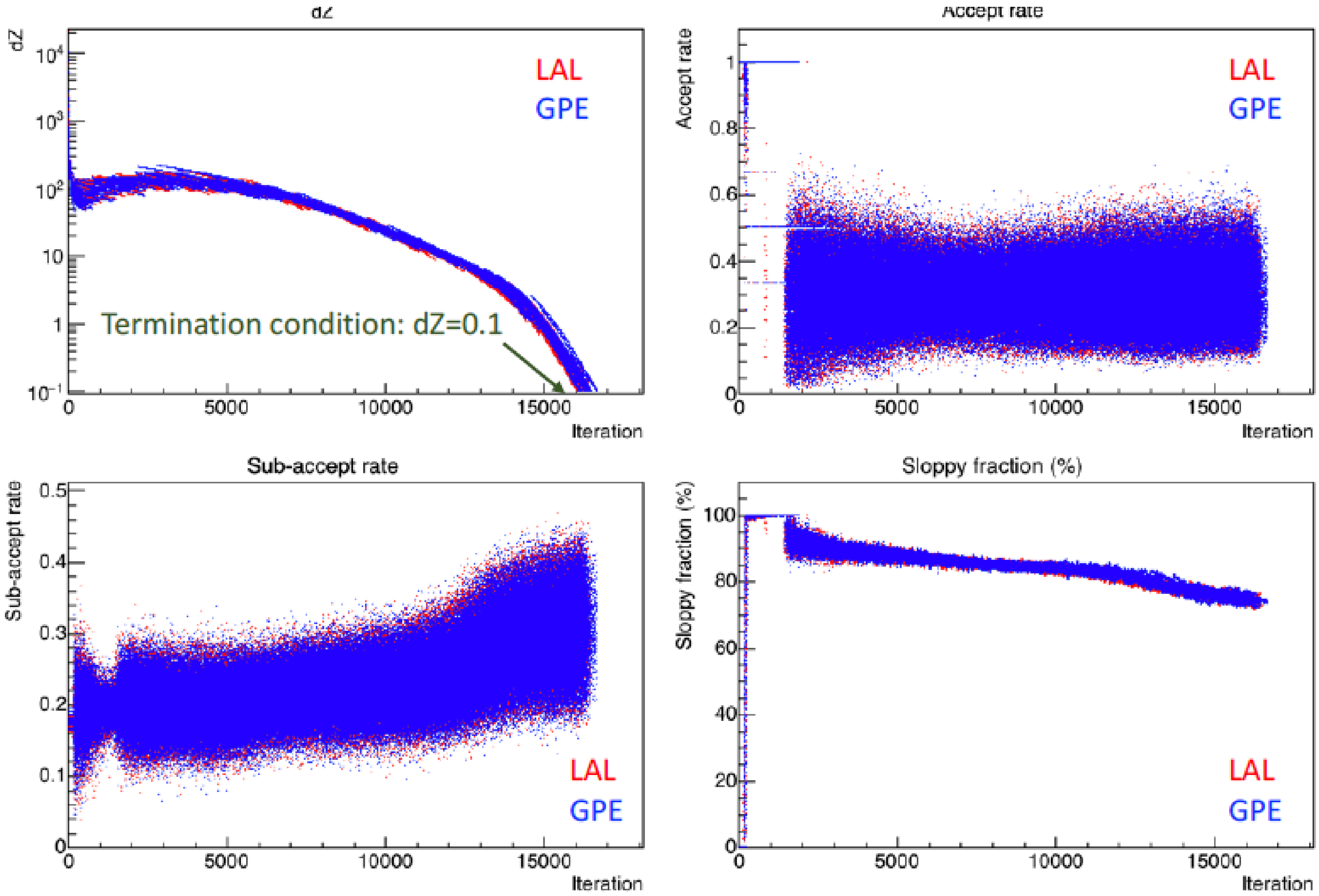
Credit: Sadakazu Haino (AS)



Code	Hardware	Spec.	Wall Time Mean ± RMS	Acceleration w.r.t. LAL	Improvement
LAL	Core™ i7	4 cores (x2 HT) 3.6 GHz	24:27:24 ± 47:42		
GPE	GeForce™ GTX 1060	1152 cores 1.76 GHz 192 bit Bus	17:21 ± 0:28	× 84.5	
GPE	GeForce™ GTX 1070	1920 cores 1.68 GHz 256 bit Bus	13:58 ± 0:17	× 105.0	24% to 1060
GPE	GeForce™ GTX 1080	2560 cores 1.85 GHz 256 bit Bus	12:25 ± 0:15	× 118.1	40% to 1060 13% to 1070

Validation of GPE with LAL

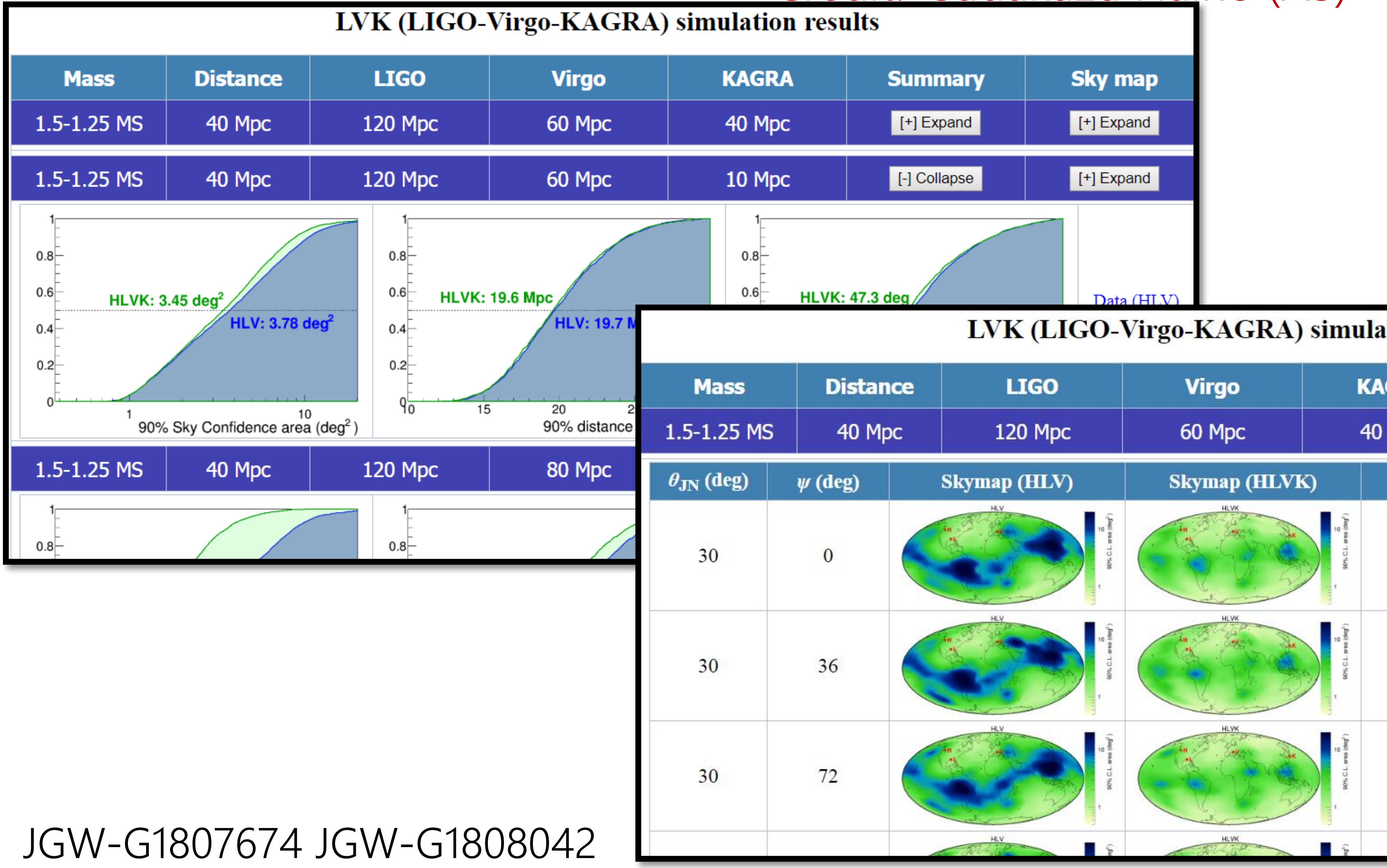
Credit: Sadakazu Haino (AS)



Termination condition: dZ=0.1

For testing detector efficiency

Credit: Sadakazu Haino (AS)



LVK (LIGO-Virgo-KAGRA) simulation results

Mass	Distance	LIGO	Virgo	KAGRA	Summary	Sky map
1.5-1.25 MS	40 Mpc	120 Mpc	60 Mpc	40 Mpc	[+] Expand	[+] Expand
1.5-1.25 MS	40 Mpc	120 Mpc	60 Mpc	10 Mpc	[-] Collapse	[+] Expand
1.5-1.25 MS	40 Mpc	120 Mpc	80 Mpc			

LVK (LIGO-Virgo-KAGRA) simula

Mass	Distance	LIGO	Virgo	KA
1.5-1.25 MS	40 Mpc	120 Mpc	60 Mpc	40

θ_{JN} (deg)	ψ (deg)	Skymap (HLV)	Skymap (HLVK)
30	0		
30	36		
30	72		

JGW-G1807674 JGW-G1808042

Summary

- Gravitational wave events will be common in the coming observation phase, which need intense study on waveform modeling and parameter estimation.

- GPU has been evaluated in GW research community:
  - ▶ Numerical relativity: AMS-NCKU code ~ 10x (Z Du/ Z Cao)
  - ▶ Parameter estimation: GPE code ~100x (Sadakazu Haino)
  - ▶ .....

- Our GPU strategy is quite straightforward

  Directive-based programming model would be appealing.

Thank you

Chun-Yu Lin **林俊鈺**
lincy@nchc.org.tw