Ando Lab Seminar

March 9, 2018

Particle Swarm Optimization for Gravitational Wave Astronomy

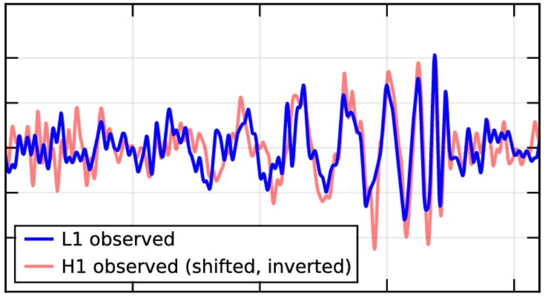
Yuta Michimura

Department of Physics, University of Tokyo

Contents

- Background
- Review of optimization methods
- Review of PSO application to GW-related research
- PSO for KAGRA design

Background



- Gravitational waves have been detected
- We have to focus more on how to extract physics from GWs, rather than on how to detect them
- The relationship between the detector sensitivity design and how much physics we can get is not always clear

- KAGRA and future detectors employ cryogenic cooling to reduce thermal noise
- Cryogenic cooling adds more complexity in sensitivity design compared with room temperature detectors because of the trade-off between mirror temperature and laser power

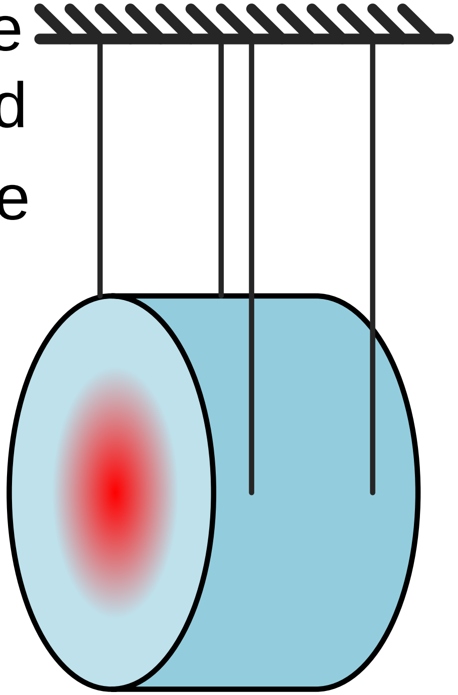
- More clever design of the sensitivity of GW detector?

Room Temperature Detector Design

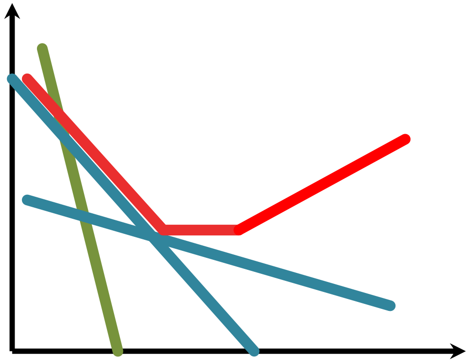
- **Seismic noise**: reduce as much as possible
  multi-stage vibration isolation, underground
- **Thermal noise**: reduce as much as possible
  larger mirror

  as thin as possible
  to support mirror mass

  thinner and longer suspensions

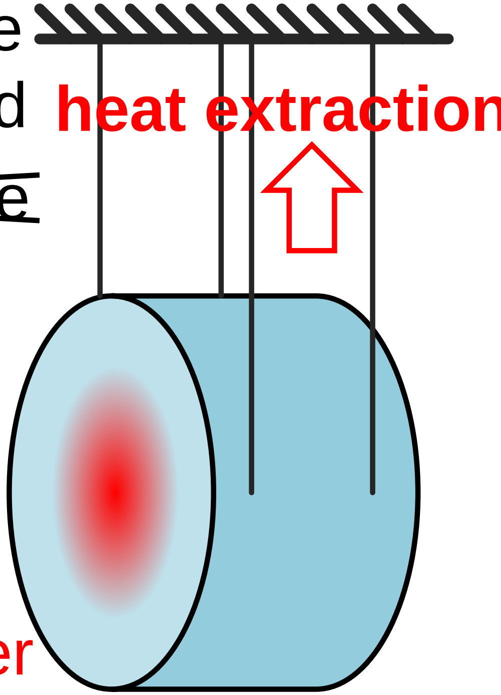


- **Quantum noise**: optimize the shape
  input laser power
  homodyne angle
  signal recycling mirror reflectivity
  detuning angle



Cryogenic Detector Design

- **Seismic noise**: reduce as much as possible
  multi-stage vibration isolation, underground



heat extraction

- **Thermal noise**: ~~reduce as much as possible~~
  larger mirror

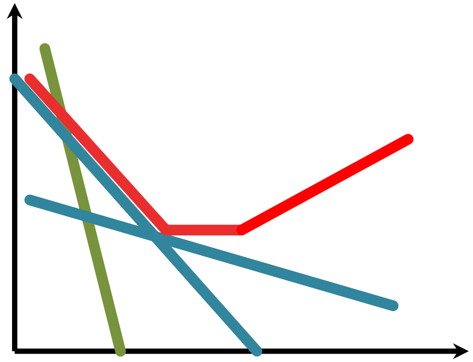
  as thin as possible
  to support mirror mass

  thinner and longer suspensions

  mirror cooling ⟷ worse cooling power
  mirror heating

  DILEMMA

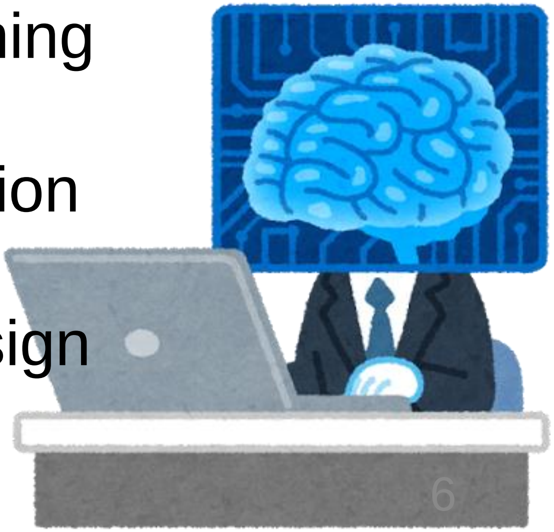
- **Quantum noise**: optimize the shape
  input laser power
  homodyne angle
  signal recycling mirror reflectivity
  detuning angle



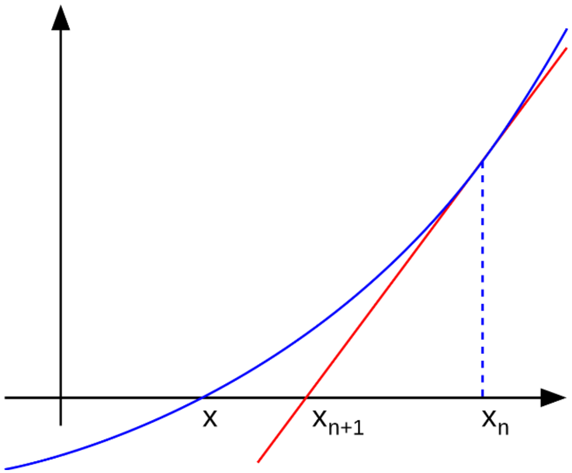
Optimization Problem

- Designing cryogenic GW detector is tough because thermal noise calculation and quantum noise optimization cannot be done independently
- Computers should do better than us

- Examples of computer-aided design / optimization
  MCMC for designing OPO
  N. Matsumoto, Master Thesis (2011)
  Machine learning for cavity mode-matching
  LIGO-G1700771
  Genetic algorithm for wave front correction
  JGW-G1706299
  Particle swarm optimization for filter design
  LIGO-G1700841 LIGO-T1700541



Optimization Algorithms



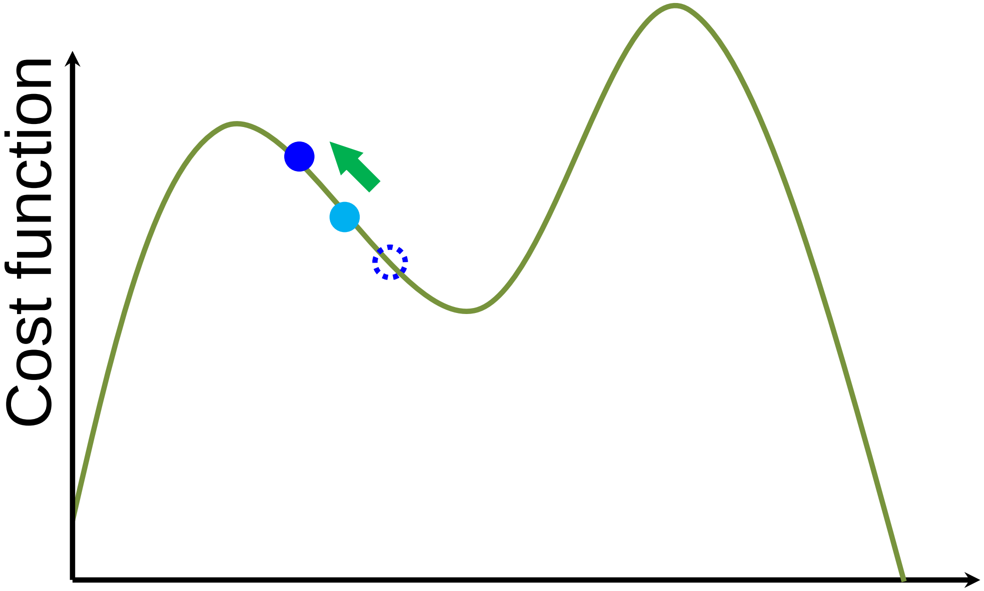
- Gradient methods
  - Gradient descent (最急降下法)
  - Newton’s method ……
- Derivative-free methods
  - Local search (局所探索法)
    - Hill climbing (山登り法)
    - Simulated annealing (焼きなまし法)
  - Evolutionally algorithms
    - Genetic algorithm
  - Swarm intelligence (群知能)
    - Ant colony optimization
    - Particle swarm optimization
- Markov chain Monte Carlo
- Machine learning (neural network, genetic programming…)

Metaheuristic

Stochastic optimization

Hill Climbing

- If neighboring solution is better, go that way

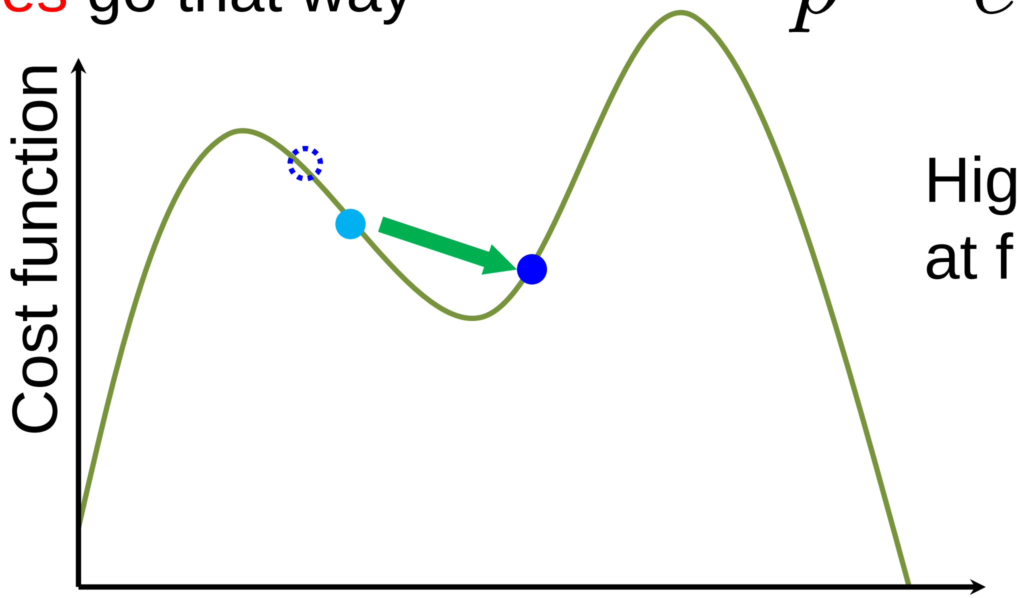


- Limitations
  - can only find local maximum/minimum

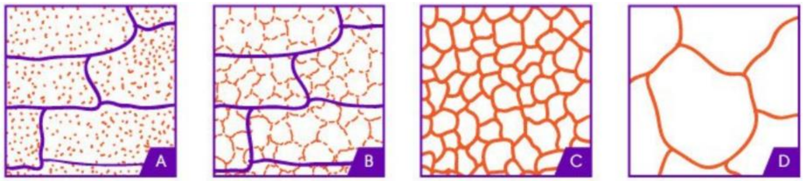
Simulated Annealing

- If neighboring solution is better, go that way
- Even if neighboring solution is worse, sometimes go that way

$$p = e^{\frac{f(x_i)-f(x_{i+1})}{T}}$$



Higher temperature at first, T=0 at last



Initial cold state

Heating; high stress areas dissipate

Recrystallization forms

Recrystallization forms

- Limitations
  - have to tune SA variables (especially cooling schedule) for different applications
  - takes time to find best solution

Particle Swarm Optimization

- Particles move based on own best position and entire swarm's best known position
- Position and velocity:

own best position so far

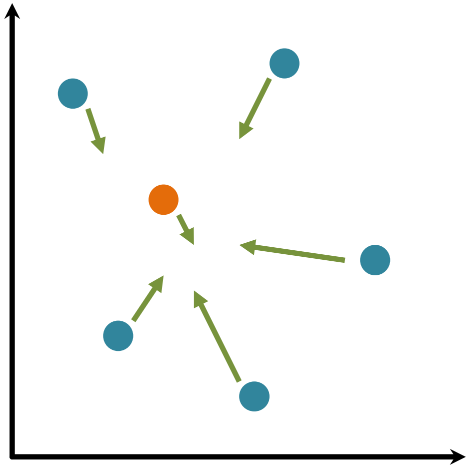
global best position so far

$$x_k(t+1) = x_k(t) + v_k(t)$$

$$v_k(t+1) = wv_k(t) + c_1 r_1(\hat{x}_k - x_k(t)) + c_2 r_2(\hat{x}_g - x_k(t))$$

inertia coefficient (~1)

coefficient c (~1)
random number r ∈[0,1]

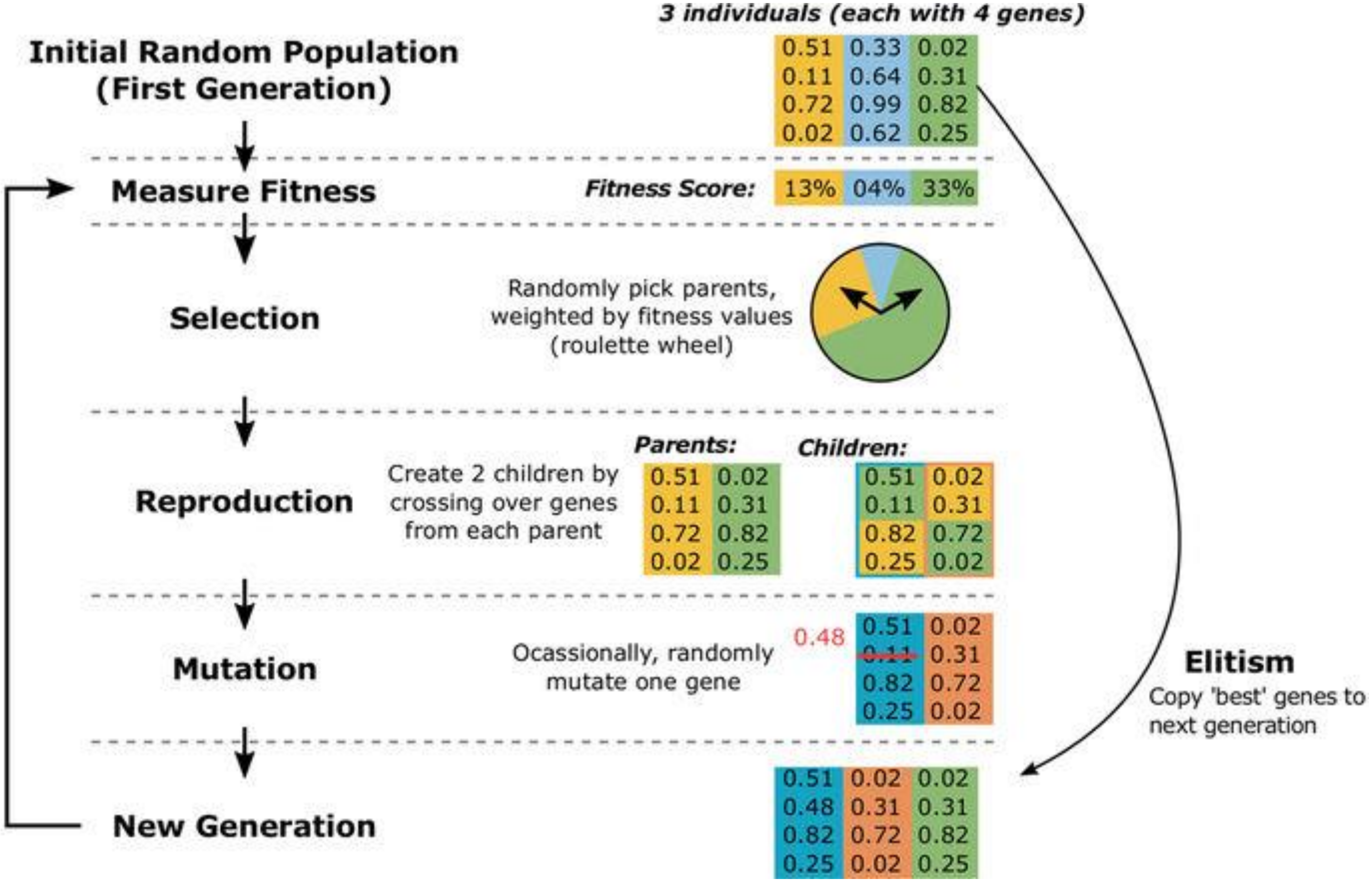


- Advantages
  - simple, fast (parallelized)
- Limitations
  - no guarantee for mathematically correct solution
  - tend to converge towards local maximum/minimum

Genetic Algorithm

- Individuals evolve based on
  - selection
  - crossover
  - mutation

Scientific Reports 6, 37616 (2016)



0.51	0.33	0.02
0.11	0.64	0.31
0.72	0.99	0.82
0.02	0.62	0.25

Fitness Score: 13% 04% 33%

Randomly pick parents, weighted by fitness values (roulette wheel)

Create 2 children by crossing over genes from each parent

Parents:

0.51	0.02
0.11	0.31
0.72	0.82
0.02	0.25

Children:

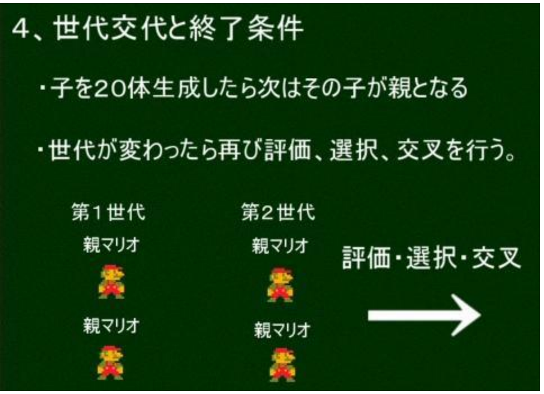
0.51	0.02
0.11	0.31
0.82	0.72
0.25	0.02

Ocassionally, randomly mutate one gene

0.51	0.02
~~0.11~~	0.31
0.82	0.72
0.25	0.02

Elitism
Copy 'best' genes to next generation

0.51	0.02	0.02
0.48	0.31	0.31
0.82	0.72	0.82
0.25	0.02	0.25



4、世代交代と終了条件

・子を20体生成したら次はその子が親となる

・世代が変わったら再び評価、選択、交叉を行う。

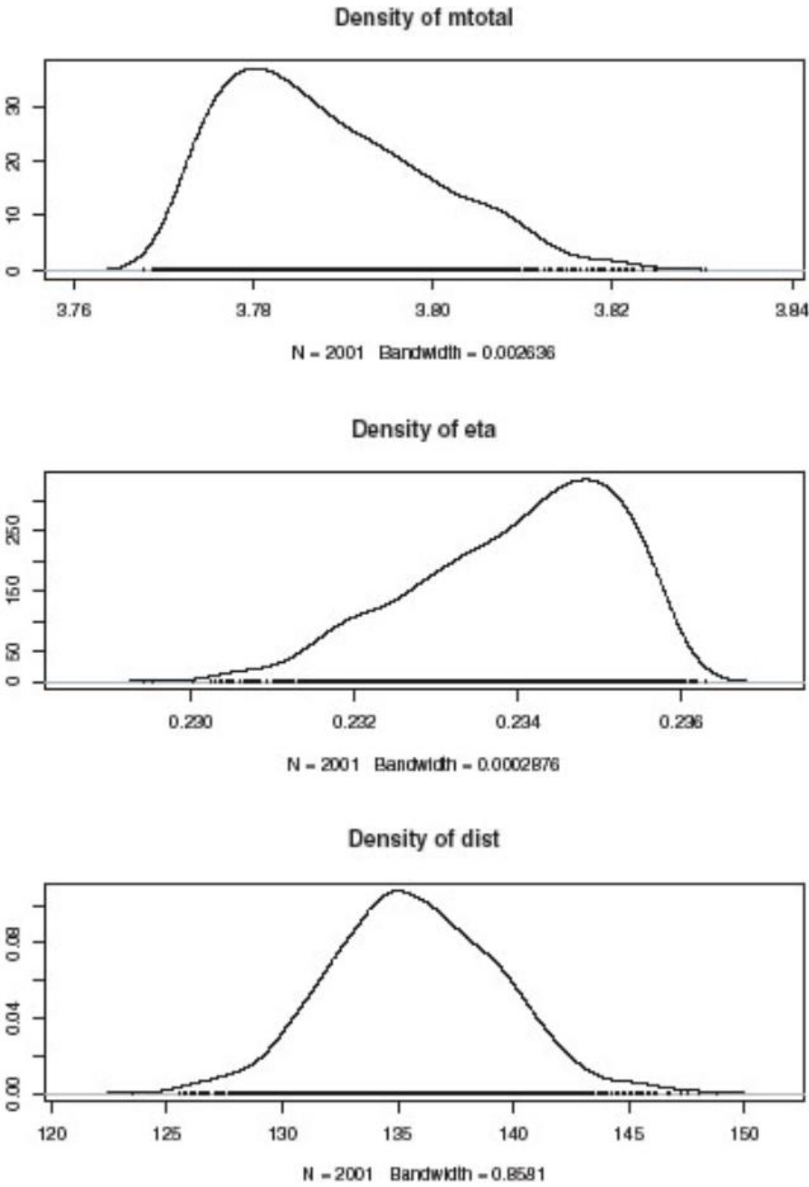
- Limitations
  - no guarantee for mathematically correct solution
  - solution could be local maximum/minimum
  - many variables for selection, crossover, mutation

Markov Chain Monte Carlo

- Not primarily for optimization
- Sample solutions with weighting (likelihood)
- Gives posterior probability density functions, and gives parameter estimation errors
- Also studied for use in GW parameter estimatio

Andrey Andreyevich Markov

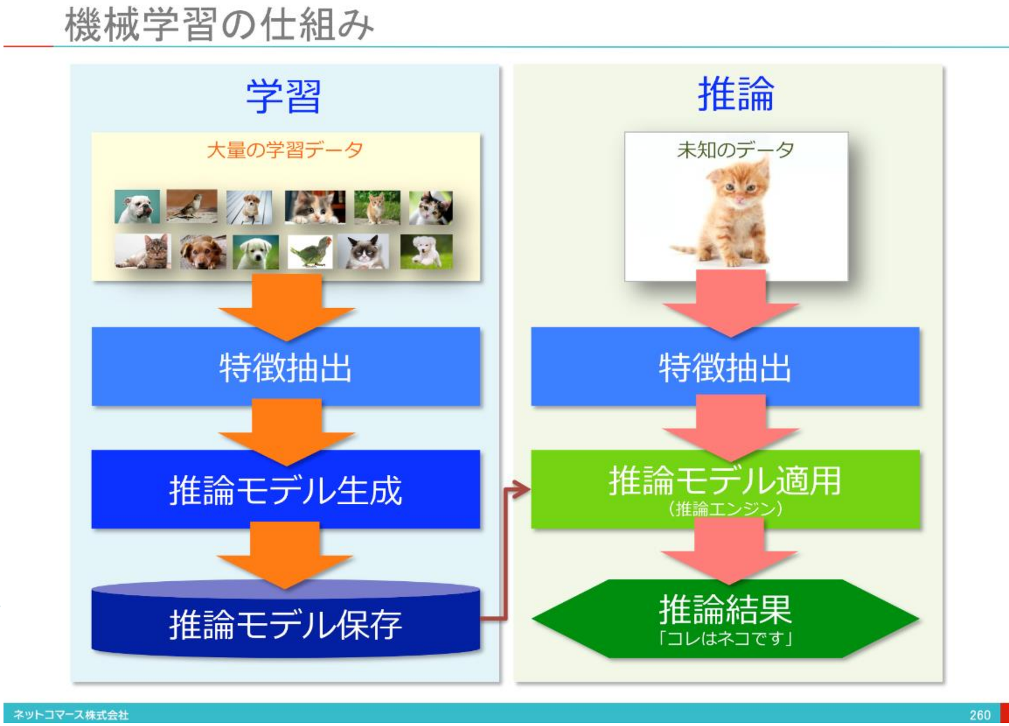
- Limitations
  - slow
  - needs prior information



CQG **21**, 317 (2004)

Machine Learning

- Not optimization algorithms
- Optimization algorithms are used for machine learning
- Prediction using statistics
  (by Jamie LIGO-G1700902)
- Limitations
  - needs big data for
    machine to learn



機械学習の仕組み

学習

大量の学習データ

特徴抽出

推論モデル生成

推論モデル保存

推論

未知のデータ

特徴抽出

推論モデル適用
（推論エンジン）

推論結果
「コレはネコです」

ネットコマース株式会社 260

http://blogs.itmedia.co.jp/itsolutionjuku/2015/07/post_106.html

- Machine learning for BEC production
  Scientific Reports 6, 25890 (2016)
- In my opinion, too much computation for optimization of function parameters

Why Particle Swarm Optimization?

- Looks simple!
- Python package Pyswarm available
  https://pythonhosted.org/pyswarm/
  https://github.com/tisimst/pyswarm/
- PSO can be done with only
  xopt, fopt = pso(func, lb, ub)

optimized parameter set

cost function to be minimized

lower / upper bounds

Additional parameters:
- swarm size
- minimum change of objective value before termination

- I'm not saying that PSO is the only best method for our use

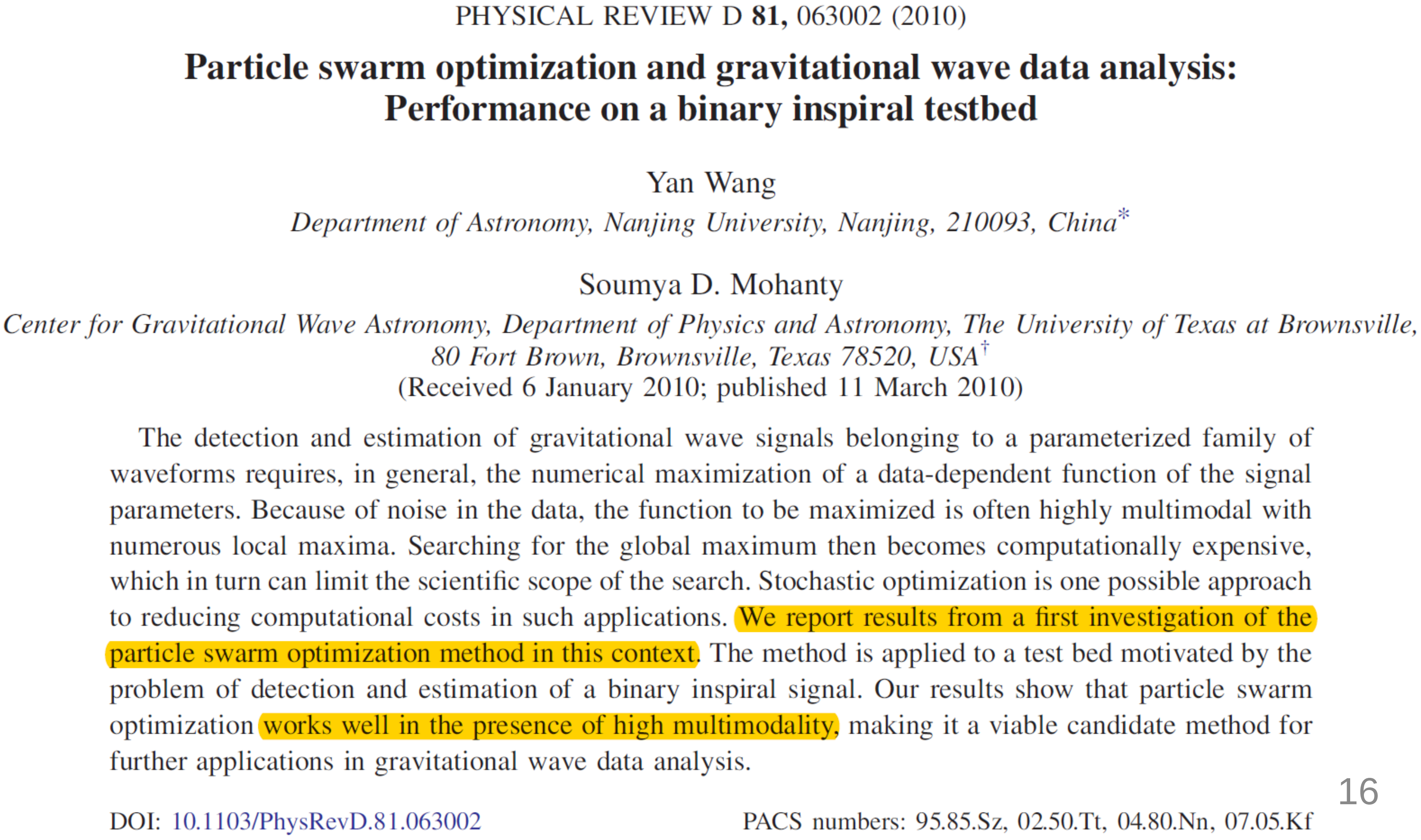
PSO for GW Related Research

- CBC search
  - Weerathunga & Mohanty, PRD 95, 124030 (2017)
  - Wang & Mohanty, PRD 81, 063002 (2010)
  - Bouffanais & Porter, PRD 93, 064020 (2016)
- CMBR analysis (WMAP data fit)
  - Prasad & Souradeep, PRD 85, 123008 (2012)
- Gravitational lensing
  - Rogers & Fiege, ApJ 727, 80 (2011)
- Continuous GW search using pulsar timing array
  - Wang, Mohanty & Jenet, ApJ 795, 96 (2014)

- Sensor correction filter design
  - Conor Mow-Lowry, LIGO-G1700841 LIGO-T1700541
- Voyager sensitivity design?

Wang & Mohanty (2010)

- Particle swarm optimization and gravitational wave data analysis: Performance on a binary inspiral testbed



PHYSICAL REVIEW D **81,** 063002 (2010)

Particle swarm optimization and gravitational wave data analysis: Performance on a binary inspiral testbed

Yan Wang

Department of Astronomy, Nanjing University, Nanjing, 210093, China[*]

Soumya D. Mohanty

Center for Gravitational Wave Astronomy, Department of Physics and Astronomy, The University of Texas at Brownsville, 80 Fort Brown, Brownsville, Texas 78520, USA[†]

(Received 6 January 2010; published 11 March 2010)

The detection and estimation of gravitational wave signals belonging to a parameterized family of waveforms requires, in general, the numerical maximization of a data-dependent function of the signal parameters. Because of noise in the data, the function to be maximized is often highly multimodal with numerous local maxima. Searching for the global maximum then becomes computationally expensive, which in turn can limit the scientific scope of the search. Stochastic optimization is one possible approach to reducing computational costs in such applications. We report results from a first investigation of the particle swarm optimization method in this context. The method is applied to a test bed motivated by the problem of detection and estimation of a binary inspiral signal. Our results show that particle swarm optimization works well in the presence of high multimodality, making it a viable candidate method for further applications in gravitational wave data analysis.

DOI: 10.1103/PhysRevD.81.063002 PACS numbers: 95.85.Sz, 02.50.Tt, 04.80.Nn, 07.05.Kf

Motivation for PSO

- Many local maxima in matched filtering
- Computationally expensive to search for global maxima
- Limiting search volume in parameter space, limiting the length of SNR integration affect the sensitivity of a search
- Computational efficiency is important

- Stochastic method (e.g. MCMC) may be sensitive to design variables and prior information
- Wide variety of stochastic method should be explored
- PSO has small number of design variables

- Note for stochastic method: additional computational cost of generating waveform on the fly

Setup

- Noise: iLIGO, single-detector
- Waveform: Upto 2PN,
  fmin= 40 Hz and fmax=700 Hz
  4 parameters (amplitude, time, phase, 2 chirp-time(←m1,m2))
- Tuned two PSO design variables (number of particles and change in intertia coefficient w) in a systematic (?) procedure based on computational cost and consistency of the result between individual PSO runs

$$w[k] = w_0 - m(k - k_0)/N_t,$$

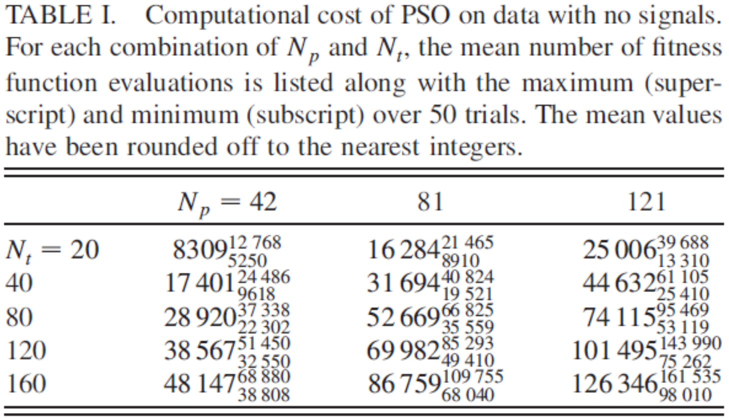


TABLE I. Computational cost of PSO on data with no signals. For each combination of N_p and N_t, the mean number of fitness function evaluations is listed along with the maximum (superscript) and minimum (subscript) over 50 trials. The mean values have been rounded off to the nearest integers.

	$N_p = 42$	81	121
$N_t = 20$	$8309^{12\,768}_{5250}$	$16\,284^{21\,465}_{8910}$	$25\,006^{39\,688}_{13\,310}$
40	$17\,401^{24\,486}_{9618}$	$31\,694^{40\,824}_{19\,521}$	$44\,632^{61\,105}_{25\,410}$
80	$28\,920^{37\,338}_{22\,302}$	$52\,669^{66\,825}_{35\,559}$	$74\,115^{95\,469}_{53\,119}$
120	$38\,567^{51\,450}_{32\,550}$	$69\,982^{85\,293}_{49\,410}$	$101\,495^{143\,990}_{75\,262}$
160	$48\,147^{68\,880}_{38\,808}$	$86\,759^{109\,755}_{68\,040}$	$126\,346^{161\,535}_{98\,010}$

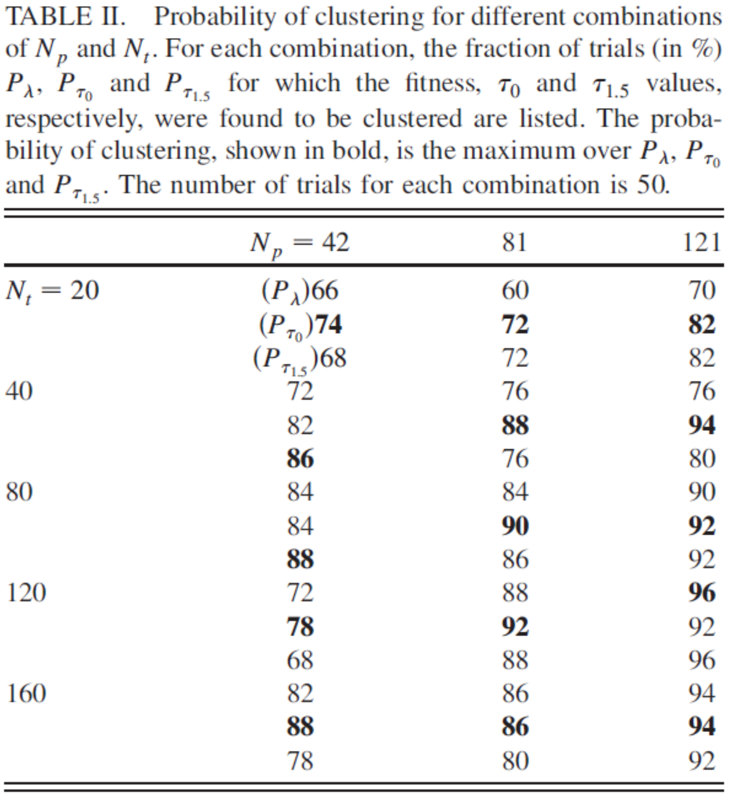


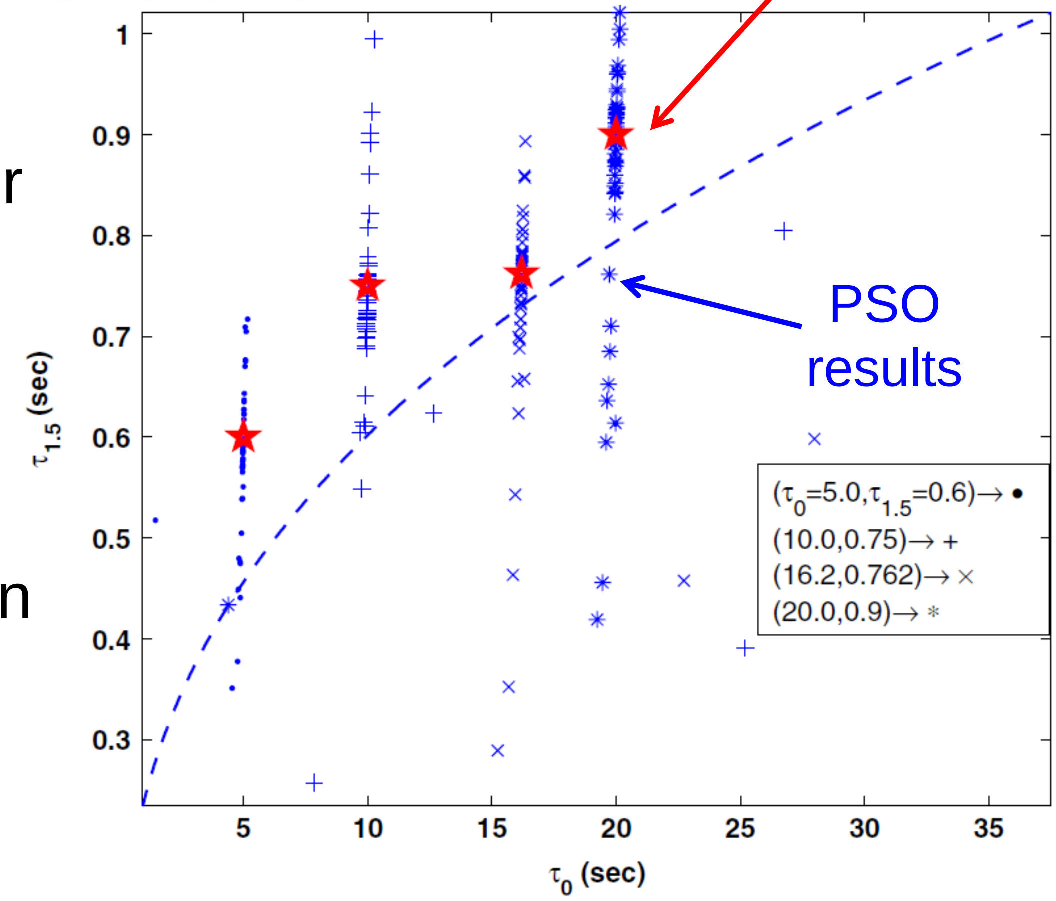
TABLE II. Probability of clustering for different combinations of N_p and N_t. For each combination, the fraction of trials (in %) P_λ, P_{τ_0} and $P_{\tau_{1.5}}$ for which the fitness, τ_0 and $\tau_{1.5}$ values, respectively, were found to be clustered are listed. The probability of clustering, shown in bold, is the maximum over P_λ, P_{τ_0} and $P_{\tau_{1.5}}$. The number of trials for each combination is 50.

	$N_p = 42$	81	121
$N_t = 20$	(P_λ)66	60	70
	(P_{τ_0})**74**	**72**	**82**
	$(P_{\tau_{1.5}})$68	72	82
40	72	76	76
	82	**88**	**94**
	86	76	80
80	84	84	90
	84	**90**	**92**
	88	86	92
120	72	88	**96**
	78	**92**	92
	68	88	96
160	82	86	94
	88	**86**	**94**
	78	80	92

Conclusion

- Looks OK
- Higher SNR gives better consistency in results, as expected

- Computational cost was ~7 times larger than grid-based search (because of low-dimensionality)
- With more dimensions, PSO should be cheaper



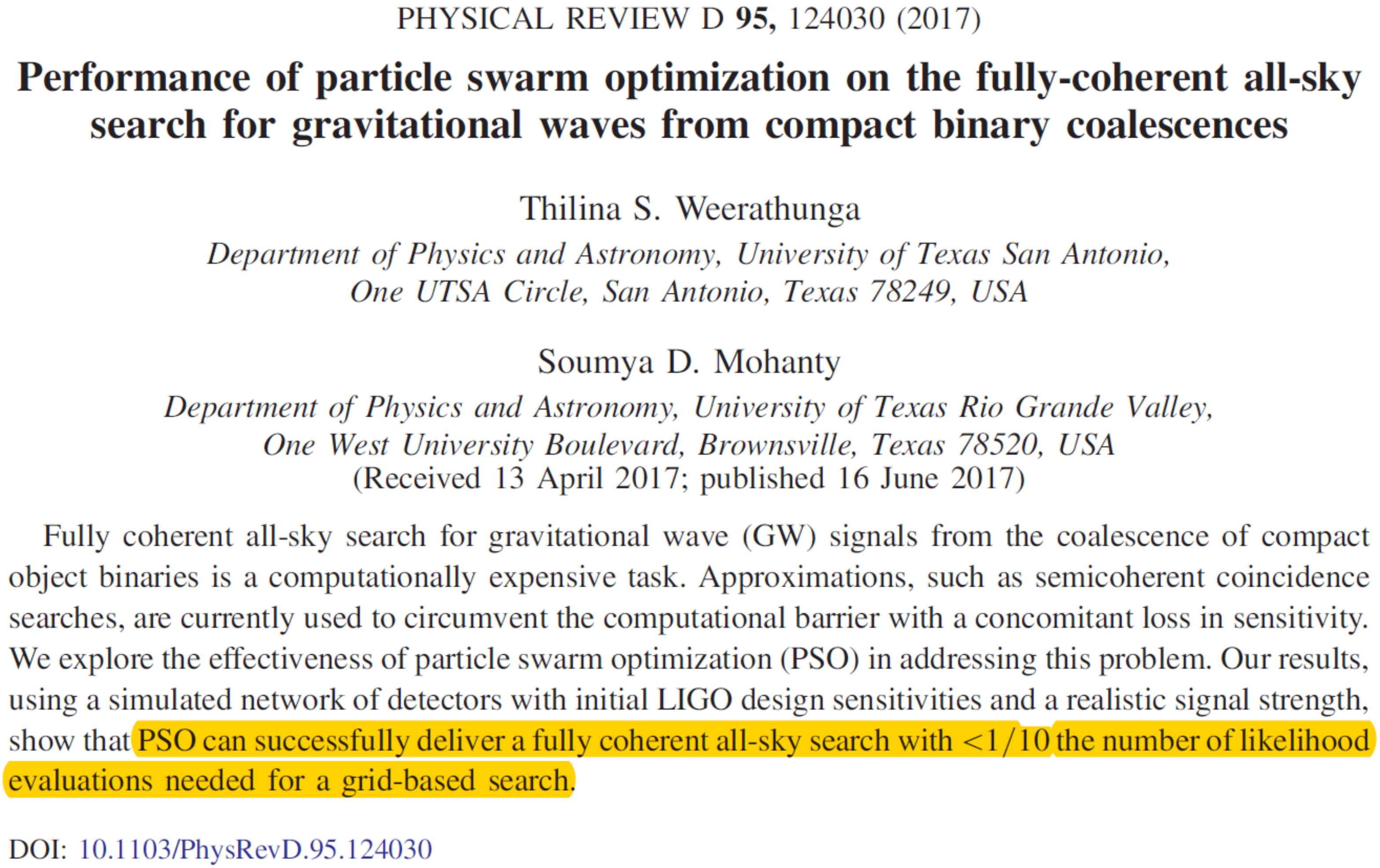
true value

PSO results

FIG. 5 (color online). Estimation of parameter values for a signal SNR of 8.0. The true locations of the signals are indicated by the ⋆ marker and each of the markers, ●, +, * and ×, indicates an estimated location corresponding to one of the true locations. The association between the markers and the true signal locations is indicated in the figure. For each true signal location, the simulation consisted of 50 trials.

Weerathunga & Mohanty (2017)

- Performance of particle swarm optimization on the fully-coherent all-sky search for gravitational waves from compact binary coalescences



PHYSICAL REVIEW D **95,** 124030 (2017)

Performance of particle swarm optimization on the fully-coherent all-sky search for gravitational waves from compact binary coalescences

Thilina S. Weerathunga

Department of Physics and Astronomy, University of Texas San Antonio,
One UTSA Circle, San Antonio, Texas 78249, USA

Soumya D. Mohanty

Department of Physics and Astronomy, University of Texas Rio Grande Valley,
One West University Boulevard, Brownsville, Texas 78520, USA

(Received 13 April 2017; published 16 June 2017)

Fully coherent all-sky search for gravitational wave (GW) signals from the coalescence of compact object binaries is a computationally expensive task. Approximations, such as semicoherent coincidence searches, are currently used to circumvent the computational barrier with a concomitant loss in sensitivity. We explore the effectiveness of particle swarm optimization (PSO) in addressing this problem. Our results, using a simulated network of detectors with initial LIGO design sensitivities and a realistic signal strength, show that PSO can successfully deliver a fully coherent all-sky search with <1/10 the number of likelihood evaluations needed for a grid-based search.

DOI: 10.1103/PhysRevD.95.124030

Setup

- HLVK network, with iLIGO noise
- Waveform: Upto 2PN,
  4 parameters (2 source locations, 2 chirp-time(←m1,m2))
- PSO design variables:
  Np=40 (swarm size)
  Niter=500 (number of iterations)

- For stochastic optimization methods, including PSO, convergence to the global maximum is not guaranteed
- Indirect check: check if fitness function is better than true signal parameters

Result: Detection Performance

- Fitness function is better in most cases

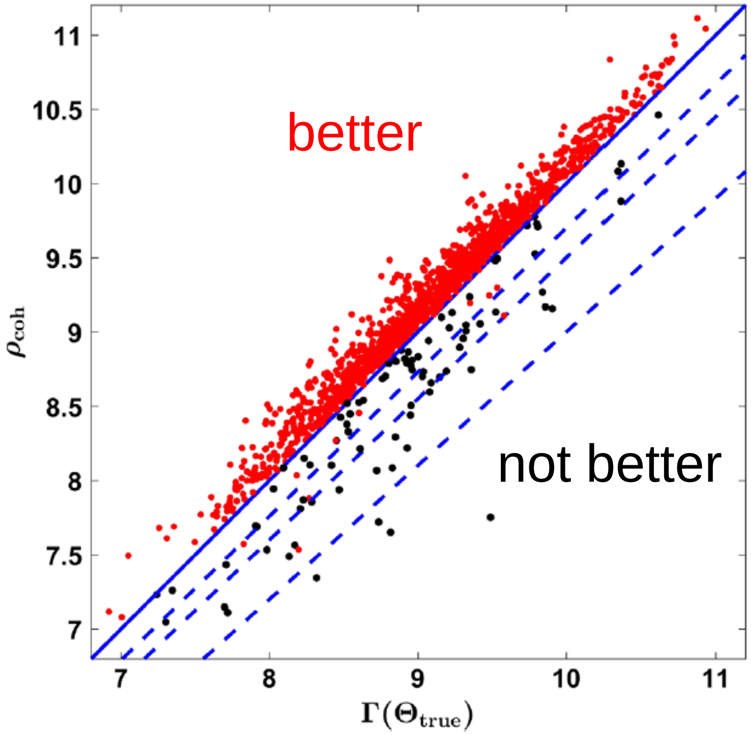


FIG. 3. Comparison of the coherent search statistic ρ_{coh} found by PSO with the coherent fitness value $\Gamma(\Theta_{\mathrm{true}})$ at the true signal parameters, Θ_{true}. Each dot corresponds to one data realization, from a total of 1440 realizations across all the source parameters used. Dashed lines show the 3%, 5%, and 10% drop from the coherent fitness value. Black dots indicate data realizations for which $\rho_{\mathrm{coh}} < \Gamma(\Theta_{\mathrm{true}})$ with $N_{\mathrm{runs}} = 12$ independent PSO runs, but recovered to $\rho_{\mathrm{coh}} \geq \Gamma(\Theta_{\mathrm{true}})$ when $N_{\mathrm{runs}} = 24$. The total number of points below the diagonal is 95.

Result: Source Location Estimate

- Estimation looks OK

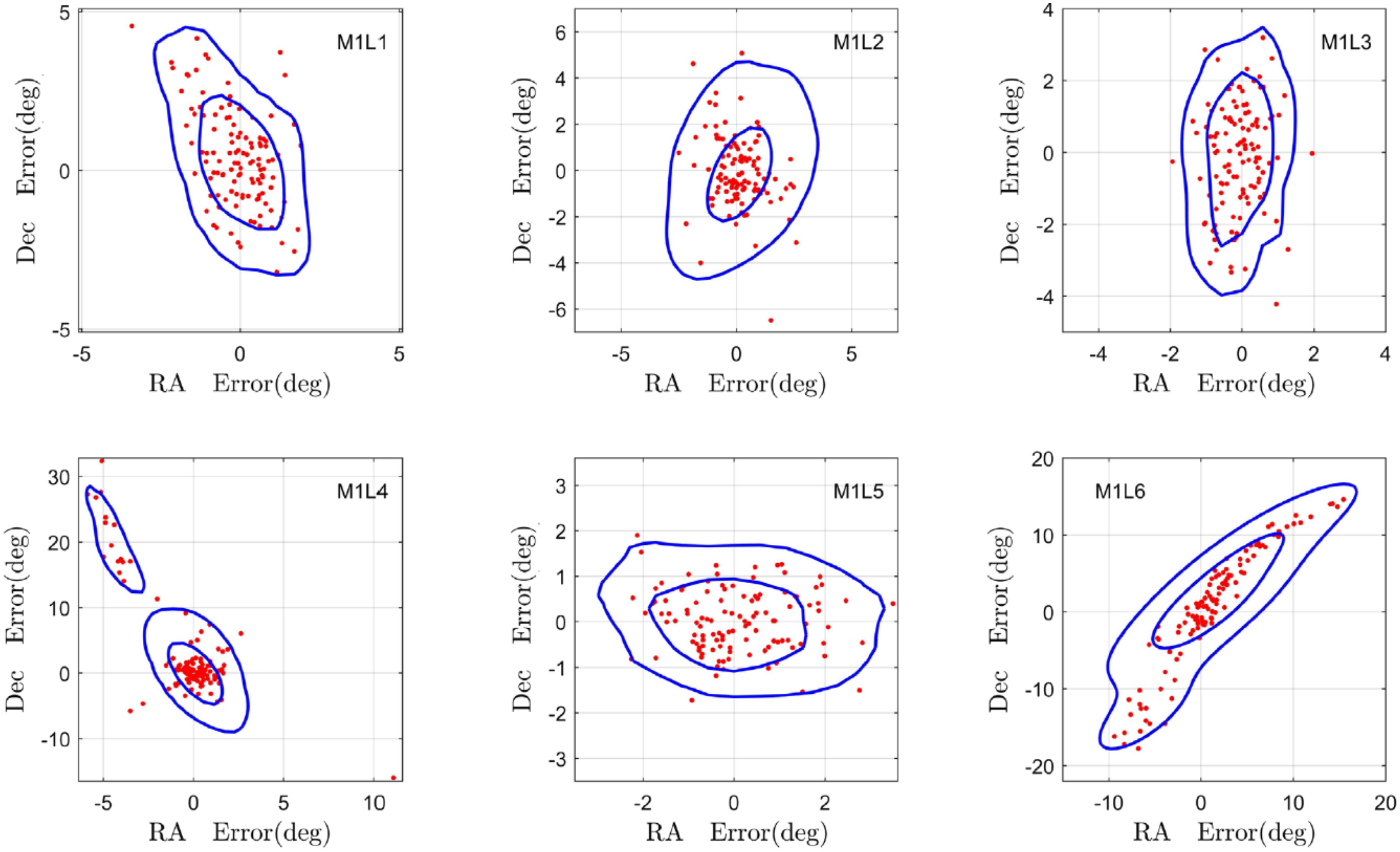


FIG. 5. Estimated sky locations (red dots) associated with the set M1 $(m_1 = 1.4\ M_\odot, m_2 = 1.4\ M_\odot)$ of sources. In each panel, the origin is centered at the true location of the source. The axes show the deviation of the estimated values of α and δ from their true values. Each panel also shows the contour levels of the bivariate probability density function, estimated using kernel density estimation (KDE) [48], that enclose 68% and 95% of the points. In these figures, the view has been zoomed in to show only the estimated locations that fall within or around the outer contour.

Result: Chirp Time Estimate

- Estimation looks OK

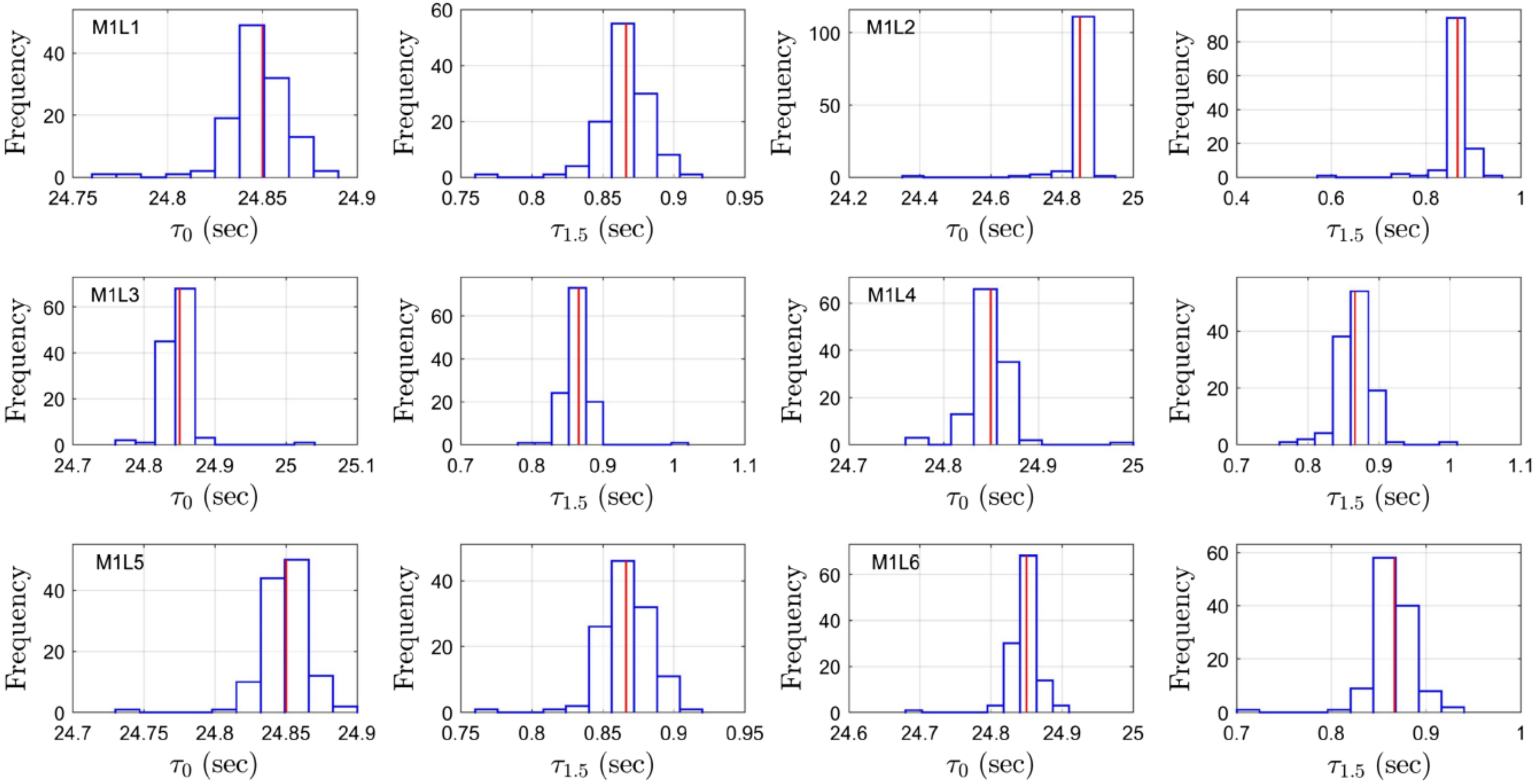


FIG. 7. Histograms of estimated chirp times, τ_0 and $\tau_{1.5}$, for all locations and mass set M1 (1.4 $M_\odot$ and 1.4 $M_\odot$). The true values of the chirp times are shown by the red line in each plot. For each source sky location, the τ_0 and $\tau_{1.5}$ distributions are adjacent and on the same row, with the $\tau_{1.5}$ distribution always to the right of the τ_0 one.

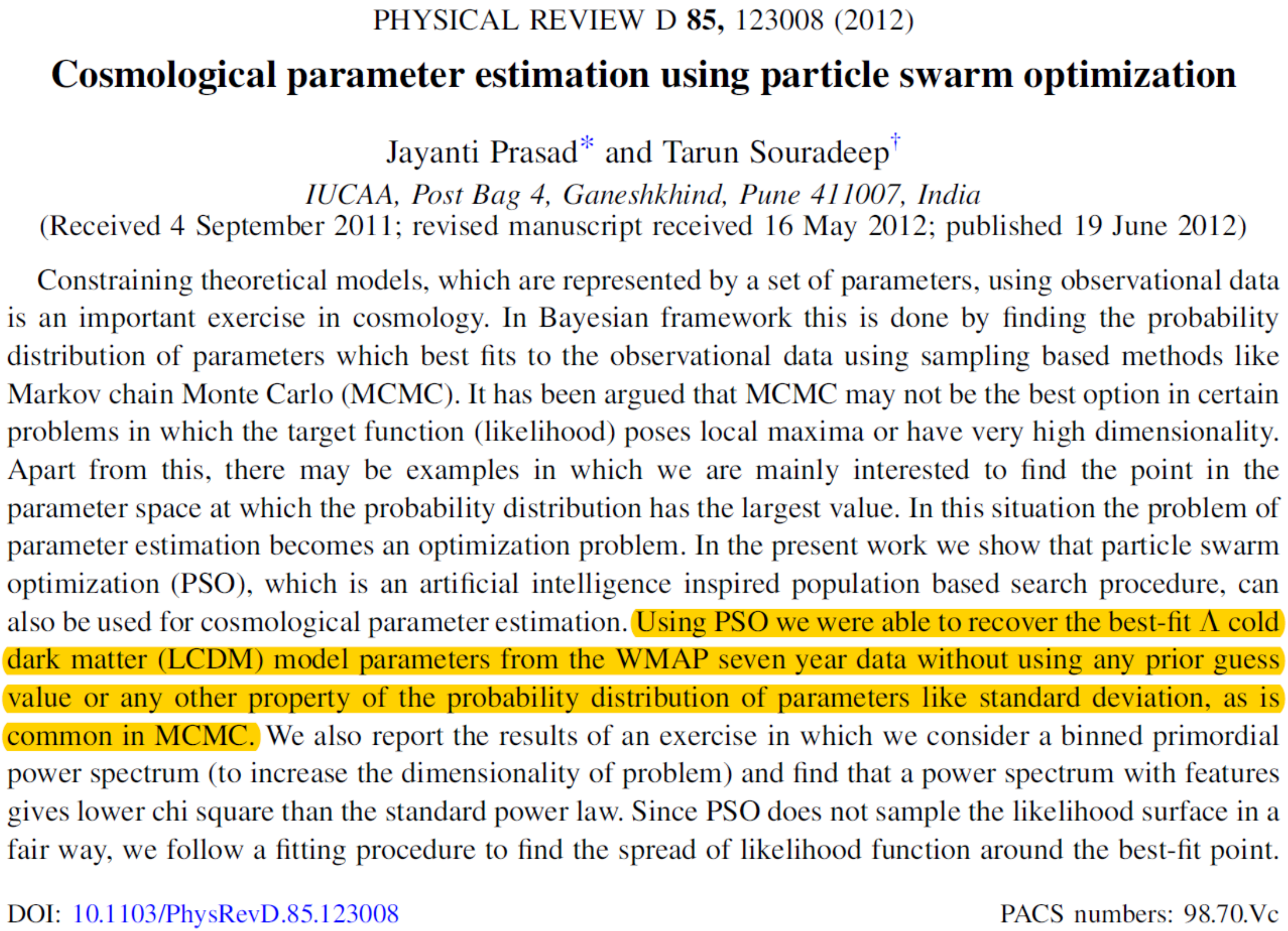
Conclusion

- Total number of fitness evaluations
  Np * Niter * Nrun = 40 * 500 * 12 = 2.4e5
- This is <1/10 of grid-based searches
- PSO can also be used for non-Gaussian noise

- Parameter estimation error comparison with Fisher information analysis is not meaningful (SNR is normalized to 9.0)
- Comparison with Bayesian approach is also difficult (error in Bayesian is different from frequentist one)

Prasad & Souradeep (2012)

- Cosmological parameter estimation using particle swarm optimization



PHYSICAL REVIEW D **85,** 123008 (2012)

Cosmological parameter estimation using particle swarm optimization

Jayanti Prasad[*] and Tarun Souradeep[†]

IUCAA, Post Bag 4, Ganeshkhind, Pune 411007, India

(Received 4 September 2011; revised manuscript received 16 May 2012; published 19 June 2012)

Constraining theoretical models, which are represented by a set of parameters, using observational data is an important exercise in cosmology. In Bayesian framework this is done by finding the probability distribution of parameters which best fits to the observational data using sampling based methods like Markov chain Monte Carlo (MCMC). It has been argued that MCMC may not be the best option in certain problems in which the target function (likelihood) poses local maxima or have very high dimensionality. Apart from this, there may be examples in which we are mainly interested to find the point in the parameter space at which the probability distribution has the largest value. In this situation the problem of parameter estimation becomes an optimization problem. In the present work we show that particle swarm optimization (PSO), which is an artificial intelligence inspired population based search procedure, can also be used for cosmological parameter estimation. Using PSO we were able to recover the best-fit Λ cold dark matter (LCDM) model parameters from the WMAP seven year data without using any prior guess value or any other property of the probability distribution of parameters like standard deviation, as is common in MCMC. We also report the results of an exercise in which we consider a binned primordial power spectrum (to increase the dimensionality of problem) and find that a power spectrum with features gives lower chi square than the standard power law. Since PSO does not sample the likelihood surface in a fair way, we follow a fitting procedure to find the spread of likelihood function around the best-fit point.

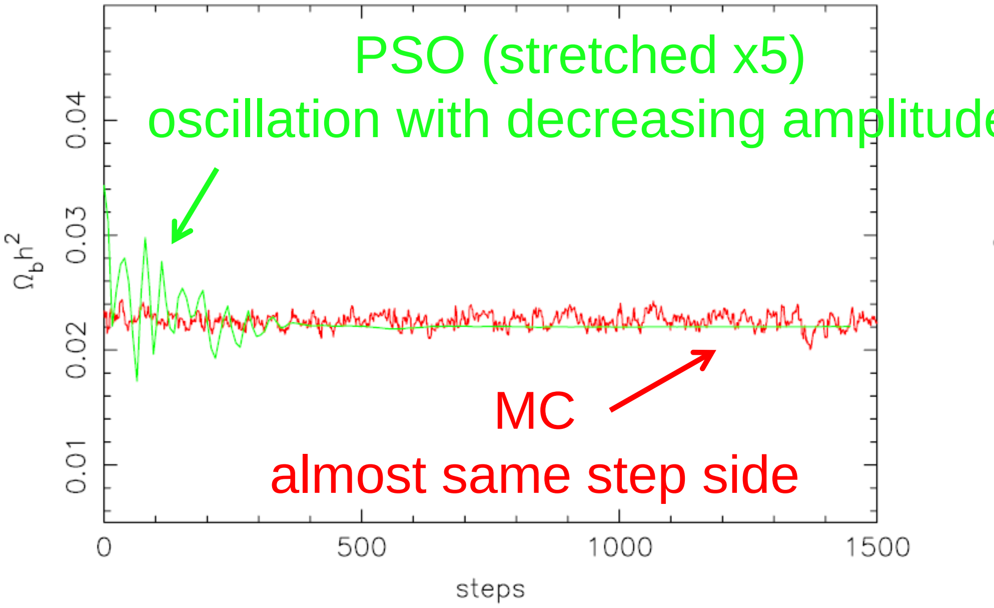
DOI: 10.1103/PhysRevD.85.123008 PACS numbers: 98.70.Vc

Motivation

- MCMC may not be the best option for problems which have local maxima or have very high dimensionality
- It has been recommended to use grid-based search first, and then MCMC
- PSO: computational cost does not grow exponentially with the dimensionality
- But, unlike MCMC, PSO does not give error bars (have to find some way to estimate)

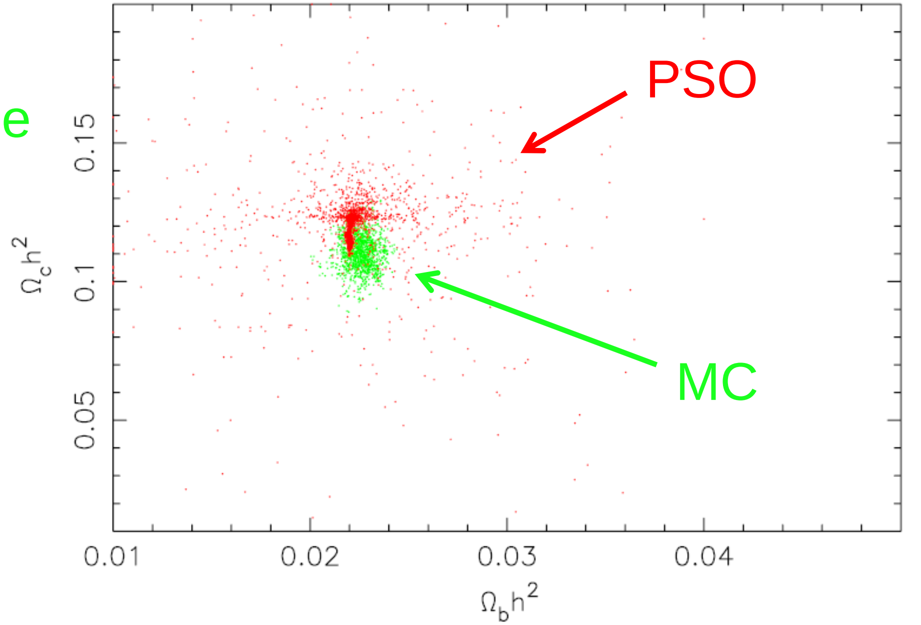
- ΛCDM model: six parameters
  cold dark matter density ($\Omega_c h^2$), baryon density ($\Omega_b h^2$), cosmological constant (Ω_Λ), primordial scalar power spectrum index (n_s), normalization (A_s), reionization optical depth (τ)

Comparison between MCMC



PSO (stretched x5)
oscillation with decreasing amplitude

MC
almost same step side



PSO

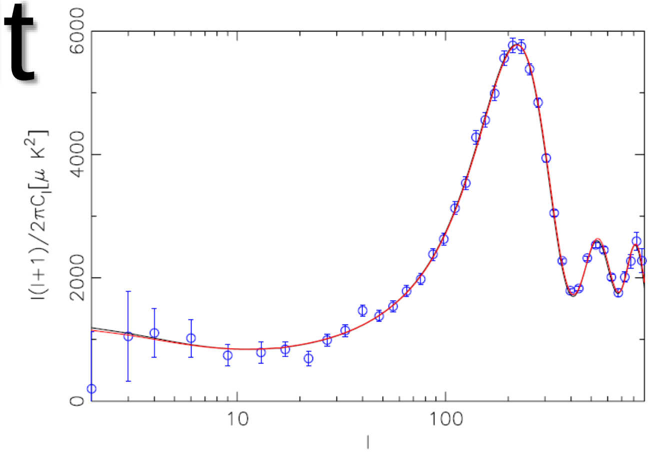
MC

FIG. 1 (color online). In this figure the red line shows a Markov chain which has been obtained from a typical run of COSMOMC and the green line shows the trajectory of a PSO particle, along the same dimension i.e. $\Omega_b h^2$. The Markov chain as well as a PSO trajectory can begin anywhere in the range and progressively move towards the best-fit location. However, in the case of PSO the particle approaches towards the best-fit location (Gbest) in an oscillatory manor with successively decreasing amplitude, which is not the case for a Markov chain since its step size does not vary much. Only after a sufficient number of PSO steps the particle positions and the Markov chain converge. Since there are more number of points for the Markov chain as compared to the PSO, we use x-scale such that we have five Markov points for every PSO point.

FIG. 2 (color online). In this figure the red and the green points show the distribution of the positions of PSO particles and samples from a Markov chain, respectively, in the same plane. From the figure it can be noticed that in the initial stage the scatter of PSO particles is very large (see Fig. 1 also), however, close to the convergence all particles get confined in a very compact region. The distribution of the sample points in the case of Markov chain is much more symmetric than in PSO. We suspect that this is due to the different role played by the stochastic variables (random numbers) in PSO as compared to that in the Markov chains. The nonsymmetric distribution makes PSO less favorable if we want to find the shape of the likelihood close to the best-fit values (in order to report errors) in comparison to the Markov chain.

Fitting Result

- Consistent with MCMC
- 50 times less fitness function call
- Only search range as an input



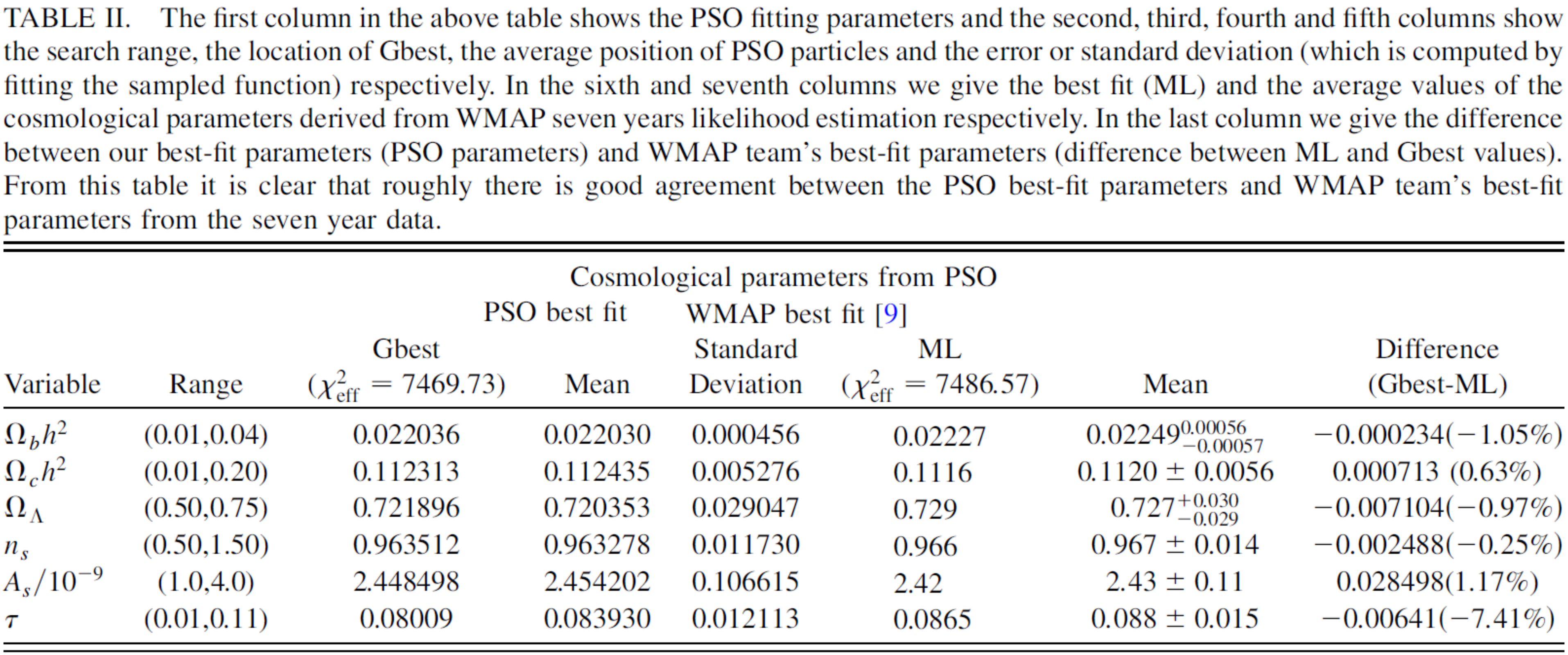


TABLE II. The first column in the above table shows the PSO fitting parameters and the second, third, fourth and fifth columns show the search range, the location of Gbest, the average position of PSO particles and the error or standard deviation (which is computed by fitting the sampled function) respectively. In the sixth and seventh columns we give the best fit (ML) and the average values of the cosmological parameters derived from WMAP seven years likelihood estimation respectively. In the last column we give the difference between our best-fit parameters (PSO parameters) and WMAP team's best-fit parameters (difference between ML and Gbest values). From this table it is clear that roughly there is good agreement between the PSO best-fit parameters and WMAP team's best-fit parameters from the seven year data.

Cosmological parameters from PSO							
		PSO best fit			WMAP best fit [9]		
		Gbest		Standard	ML		Difference
Variable	Range	($\chi^2_{\rm eff} = 7469.73$)	Mean	Deviation	($\chi^2_{\rm eff} = 7486.57$)	Mean	(Gbest-ML)
$\Omega_b h^2$	(0.01,0.04)	0.022036	0.022030	0.000456	0.02227	$0.02249^{0.00056}_{-0.00057}$	$-0.000234(-1.05\%)$
$\Omega_c h^2$	(0.01,0.20)	0.112313	0.112435	0.005276	0.1116	0.1120 ± 0.0056	0.000713 (0.63%)
Ω_Λ	(0.50,0.75)	0.721896	0.720353	0.029047	0.729	$0.727^{+0.030}_{-0.029}$	$-0.007104(-0.97\%)$
n_s	(0.50,1.50)	0.963512	0.963278	0.011730	0.966	0.967 ± 0.014	$-0.002488(-0.25\%)$
$A_s/10^{-9}$	(1.0,4.0)	2.448498	2.454202	0.106615	2.42	2.43 ± 0.11	0.028498(1.17%)
τ	(0.01,0.11)	0.08009	0.083930	0.012113	0.0865	0.088 ± 0.015	$-0.00641(-7.41\%)$

PPS with Features

- Primordial power spectrum is usually considered featureless
- PPS with power in bins (20 parameters in addition to $\Omega_c h^2$, $\Omega_b h^2$, Ω_Λ, τ)
- PSO fits better than MCMC
  χ_{eff}^2 is lower by 7

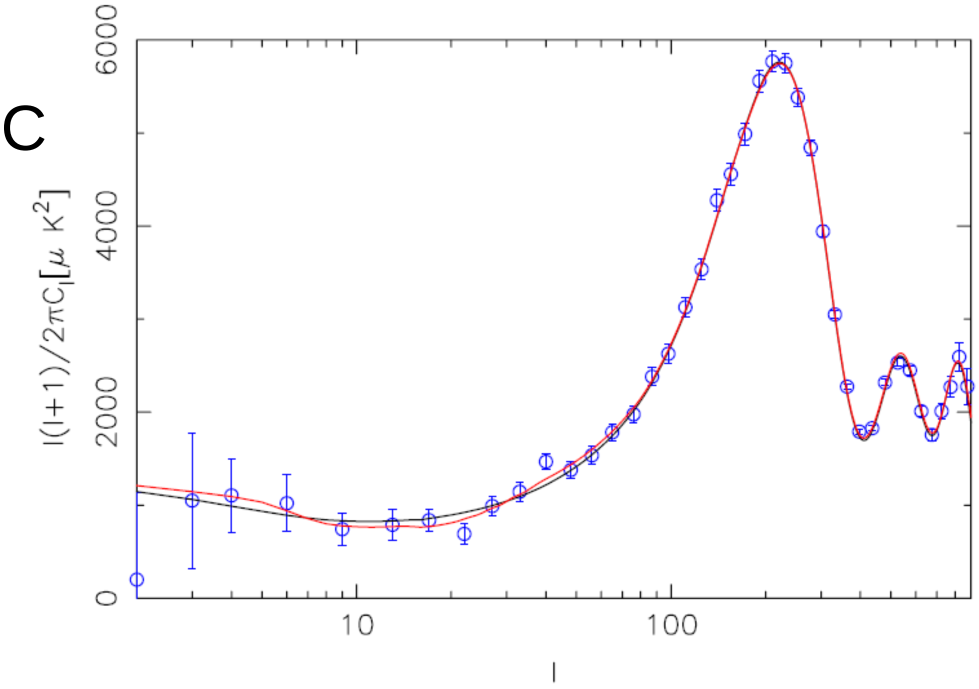


FIG. 9 (color online). The red, black and blue lines in the above figure represent the best-fit angular power spectrum recovered from PSO, standard LCDM power spectrum and the binned power spectrum of WMAP seven year data, respectively. Note that at low l the angular power spectrum with binned PPS fits better as compared to the standard power law PPS to the observed data (the improvement in $\Delta\chi^2_{\rm eff}$ is around 7).

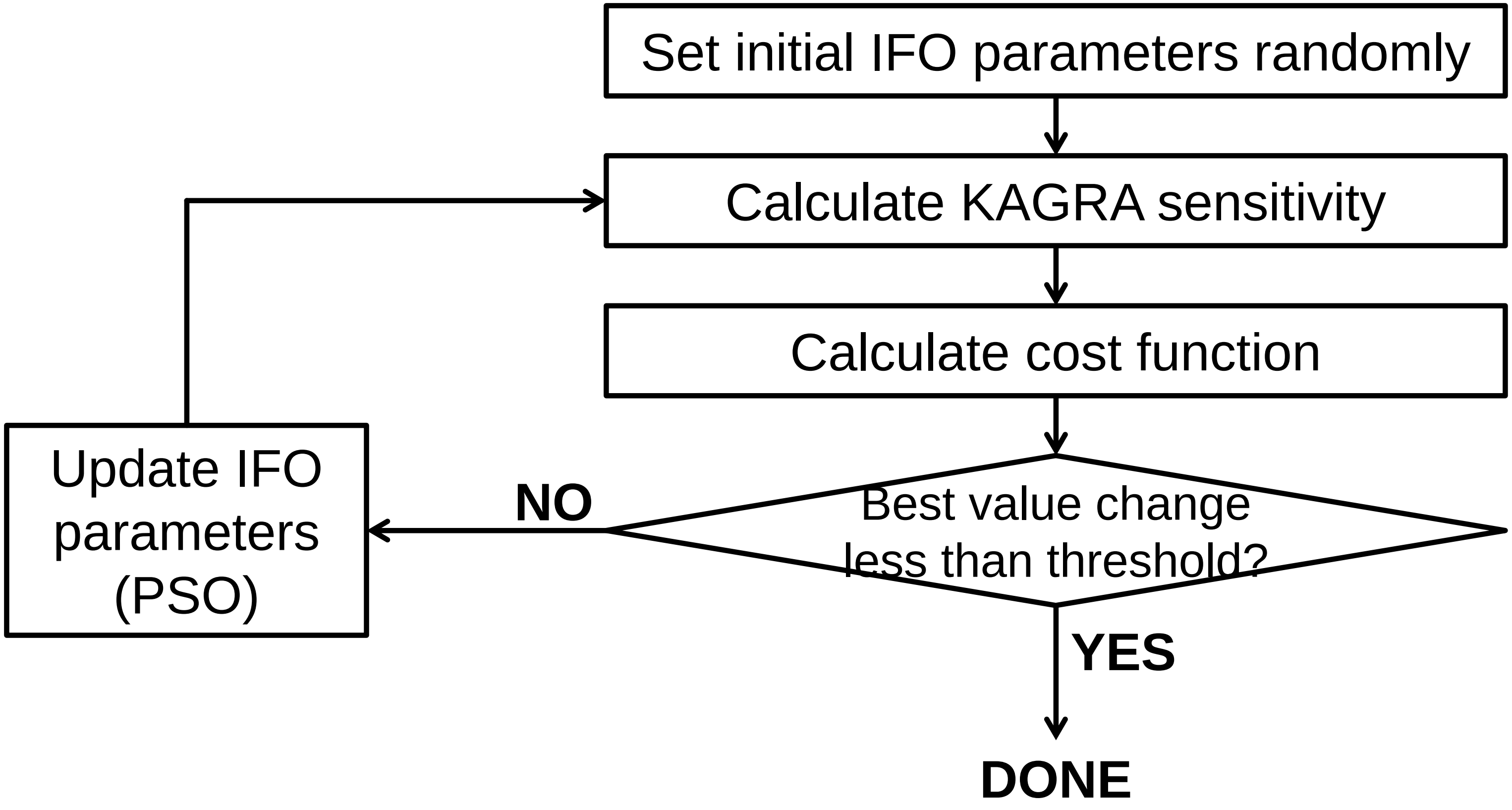
Summary on PSO

- Small number of design variables
- Almost no prior information necessary (only search range)
- Computationally cheaper for higher dimensionality
- No guarantee on convergence to the global optima
- Potential for further research

Particle Swarm Optimization for KAGRA Sensitivity Design

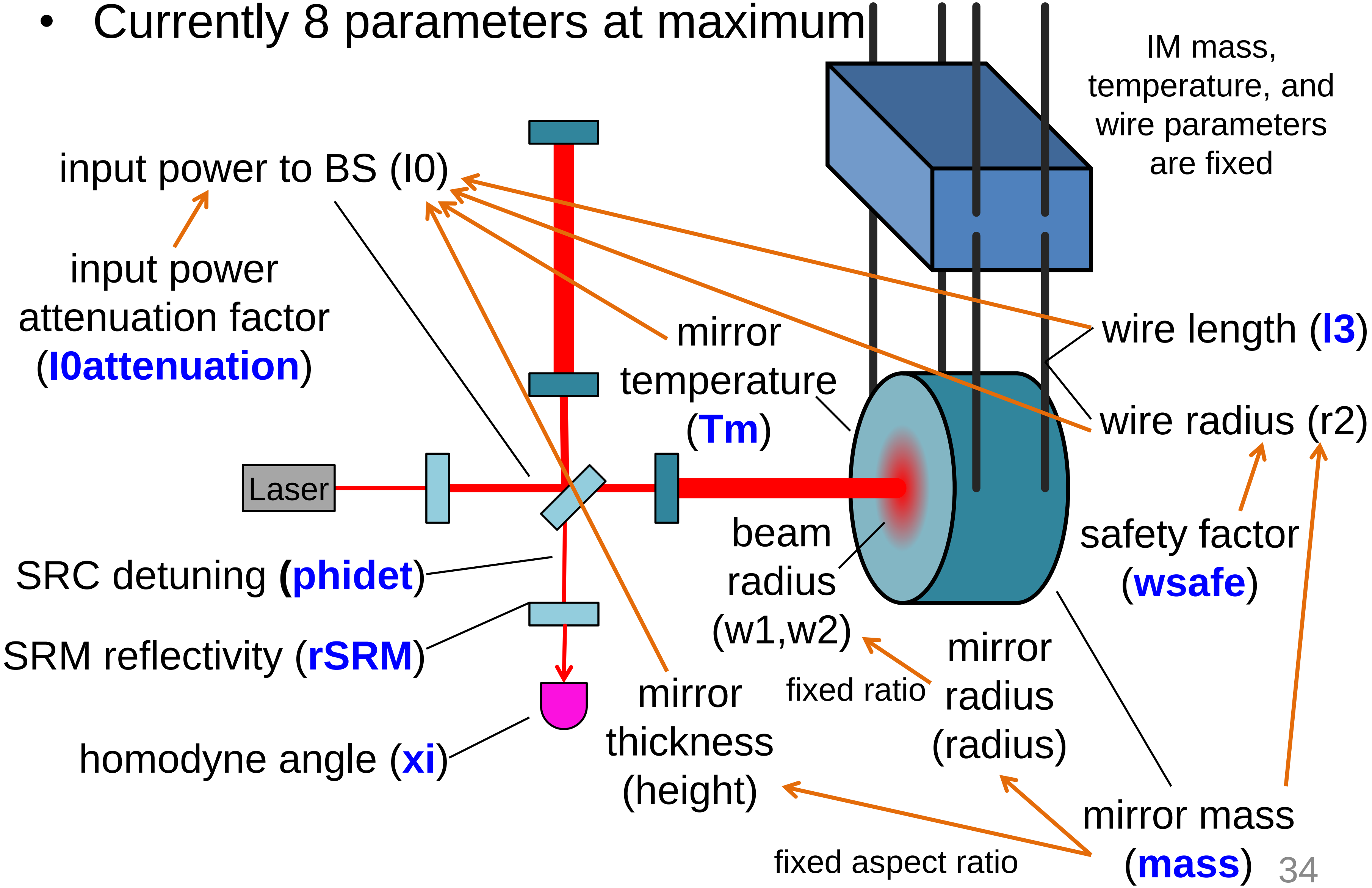
PSO for KAGRA Design

- Developed python codes to optimize KAGRA sensitivity using PSO (psokagra.py)
- Sensitivity calculation same as kagra_sensitivity.m by Komori *et al* (JGW-T1707038)


Set initial IFO parameters randomly
Calculate KAGRA sensitivity
Calculate cost function
Best value change less than threshold?
NO
Update IFO parameters (PSO)
YES
DONE

IFO Parameters to Optimize

- Currently 8 parameters at maximum



IM mass, temperature, and wire parameters are fixed

input power to BS (I0)

input power attenuation factor (**I0attenuation**)

mirror temperature (**Tm**)

wire length (**l3**)

wire radius (r2)

beam radius (w1,w2)

safety factor (**wsafe**)

SRC detuning **(phidet)**

SRM reflectivity (**rSRM**)

homodyne angle (**xi**)

mirror thickness (height)

fixed ratio

mirror radius (radius)

mirror mass (**mass**)

fixed aspect ratio

IFO Parameter Search Range

needs more money

	Lower bound	Upper bound	KAGRA Default
Detuning angle [deg]	86.5 (or 60) *	90	86.5
Homodyne angle [deg]	90	180	135.1
Mirror temperature [K]	20	30	22
Power attenuation	0.01	1	1
SRM reflectivity	0.6	1	0.92 (85%)
Wire length [cm]	20	100	35
Wire safety factor	3	100	12.57
Mirror mass [kg]	22.8	100	22.8

* Maximum detuning is 3.5 deg considering SRC nonlinear effect (Aso+ CQG 29, 124008 (2012))

- Boundary condition:
  if x>xmax, x=xmax; if x<xmin, x=xmin

Cost Functions

- Inspiral range for (equal mass) binary
  - calculation same as kagra_sensitivity.m by Komori *et al* (JGW-T1707038)
  - 10 Hz to f_ISCO, $f^{-7/3}/h^2$
  - might change to ir_ajith.m by M. Ando *et al* in the future (IMR waveform by Ajith+, PRL 106, 241101 (2011))
- Binary parameter estimation error for given source
  - calculation same as fisher analysis code by Nishizawa based on Khan+, PRD 93 044007 (2016) and Berti+, PRD 71, 084025 (2005)
  - 30 Hz to f_ISCO
  - only inspiral waveform for now
- SNR for given binary source
  - calculation same as fisher analysis code by Nishizawa
- Detection rate yet to be done (takes too much time)

Fisher Analysis

- Fisher information matrix

geometrical factor

waveform

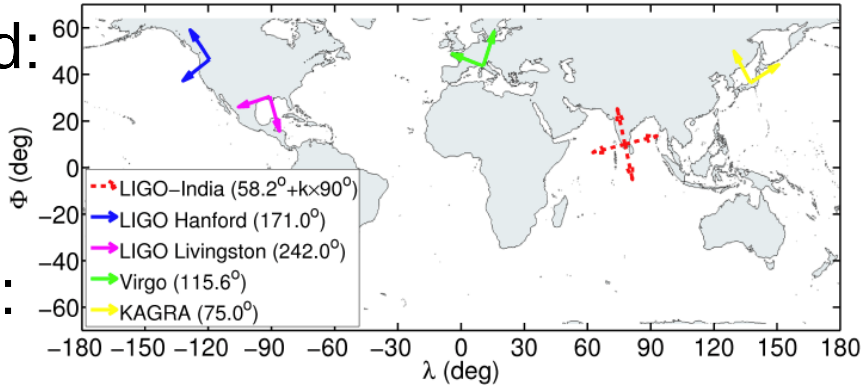
$$F_{ij} = 4\Re \int \sum_k \frac{|g_k|^2 \partial_i h^*(f) \partial_j h(f)}{S_k(f)} \mathrm{d}f$$

noise of k-th detector

$$\sigma_i = \sqrt{(F^{-1})^{ii}}$$

estimation error for i-th parameter

- GW detector network assumed: aLIGO H1, L1 and AdV with their designed sensitivities



Raffai+ CQG 30, 155004 (2013)

- Binary parameters considered:
  mc: chirp mass
  eta: symmetric mass ratio
  tc, phic: time and phase for coalescence
  dL: luminosity distance
  chis, chia: symmetric/asymmetric spin $\chi_{\mathrm{s/a}} = (\chi_1 \pm \chi_2)/2$
  thetas, phis: colatitude / longitude of source
  cthetai: inclination angle
  psip: polarization angle

PSO Design Variables

- Size of the swarm (swarmsize)
  - have to be tuned for each optimization (~100)
- Minimum change of swarm's best value before termination (minfunc)
  - precision you want to optimize the cost function (e.g. for inspiral range, 0.01 Mpc)
- That's it!

Optimization Speed

- O(1) minutes on my laptop, without multiprocessing

- Sensitivity calculation takes ~0.1 sec
- Inpiral range calculation takes ~0.00015 sec
- Fisher matrix calculation takes ~0.075 sec
  → sensitivity calculation limits the speed

- ~0.1 sec * ~100 particles * ~20 iterations = ~200 sec for optimization

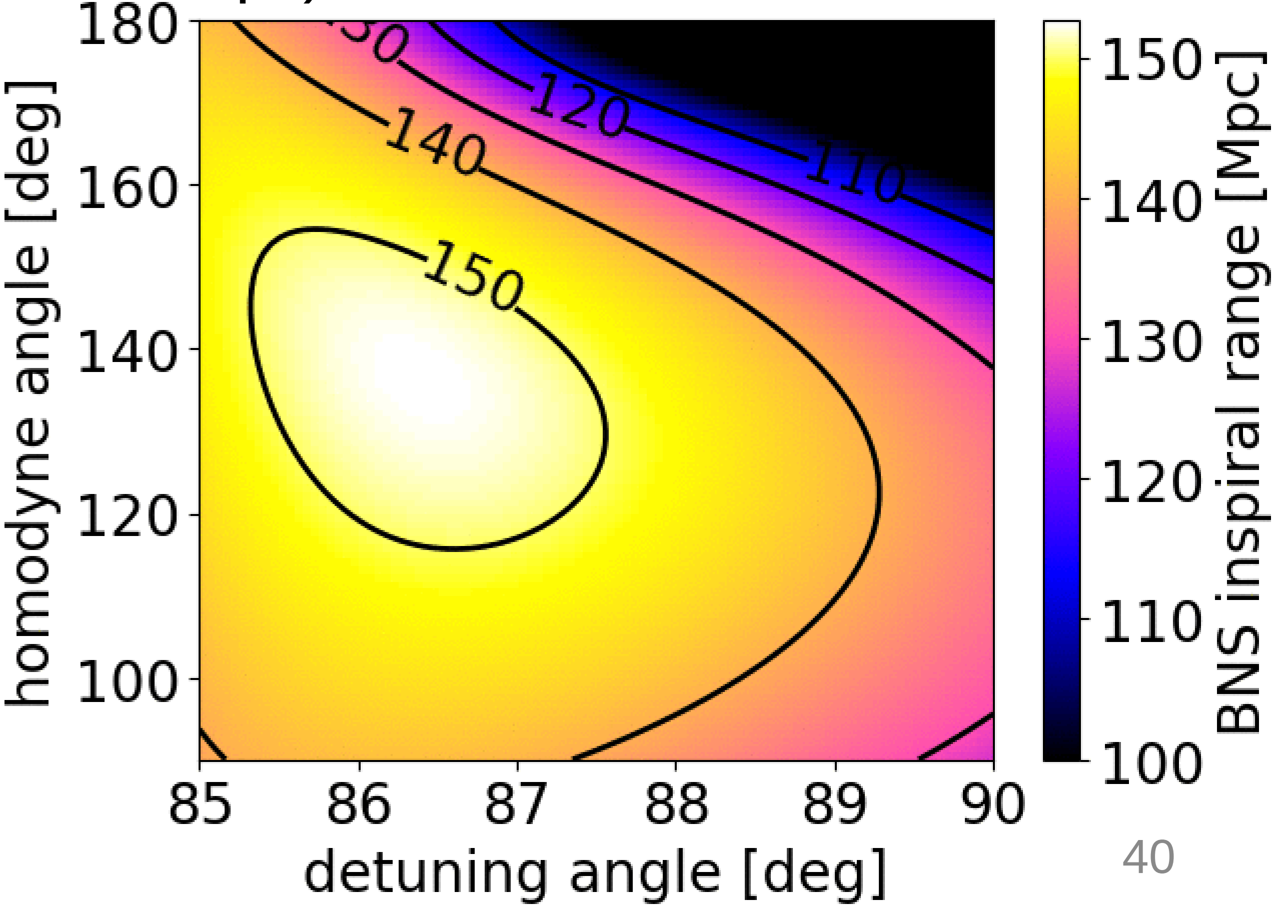
- Tolerable amount of time!

2 Params, for IR1.4

- phidet= 86.4 deg, xi= 136.7 deg
  gives IR1.4= 152.623 Mpc
  (Default: phidet= 86.5 deg, xi= 135.1 deg
  gives IR1.4= 152.598 Mpc)

Search range
for phidet = [85,90]

- Consistent with manual optimization

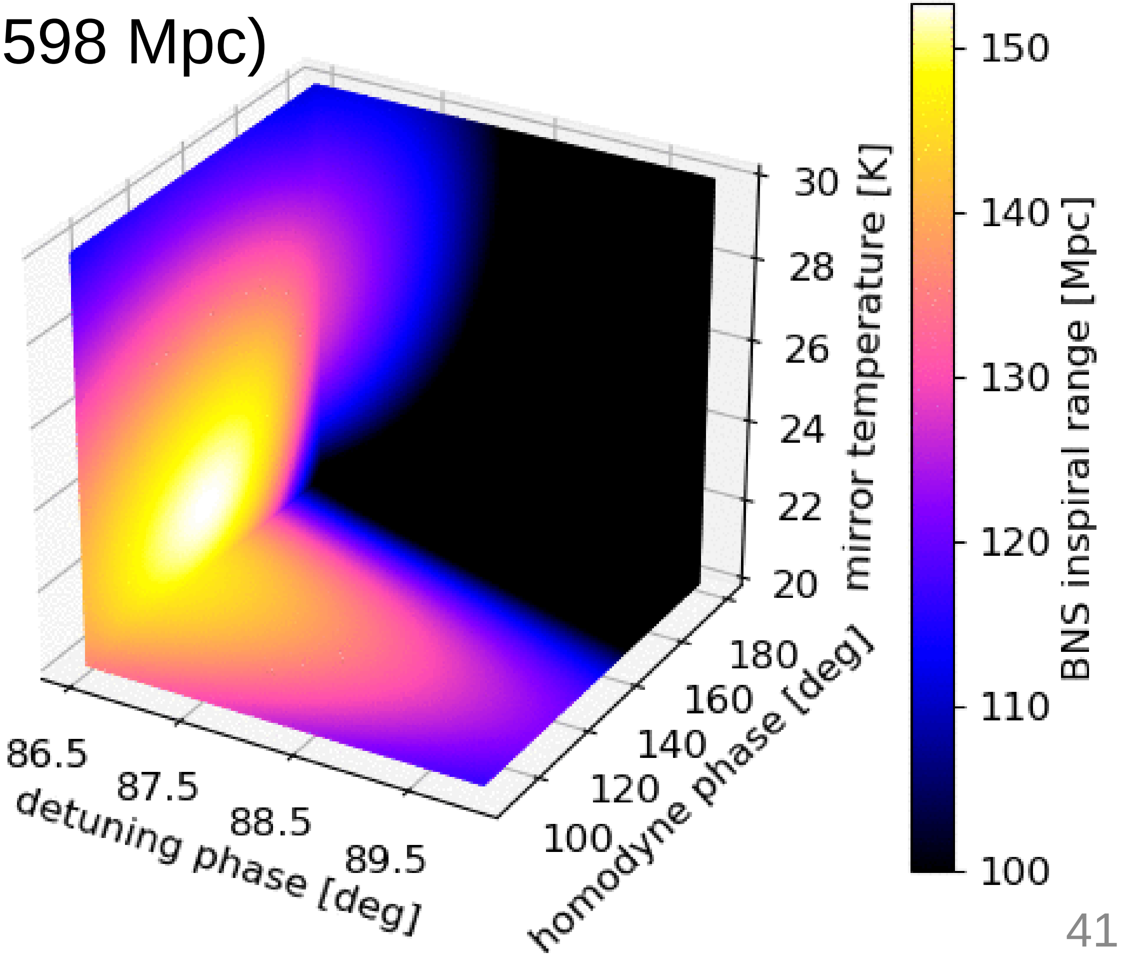


3 Params, for IR1.4

- phidet= 86.5 deg, xi= 134.3 deg, Tm= 21.6 K gives IR1.4= 152.794 Mpc (Default: phidet = 86.5 deg, xi = 135.1 deg, Tm =22 K gives IR1.4= 152.598 Mpc)

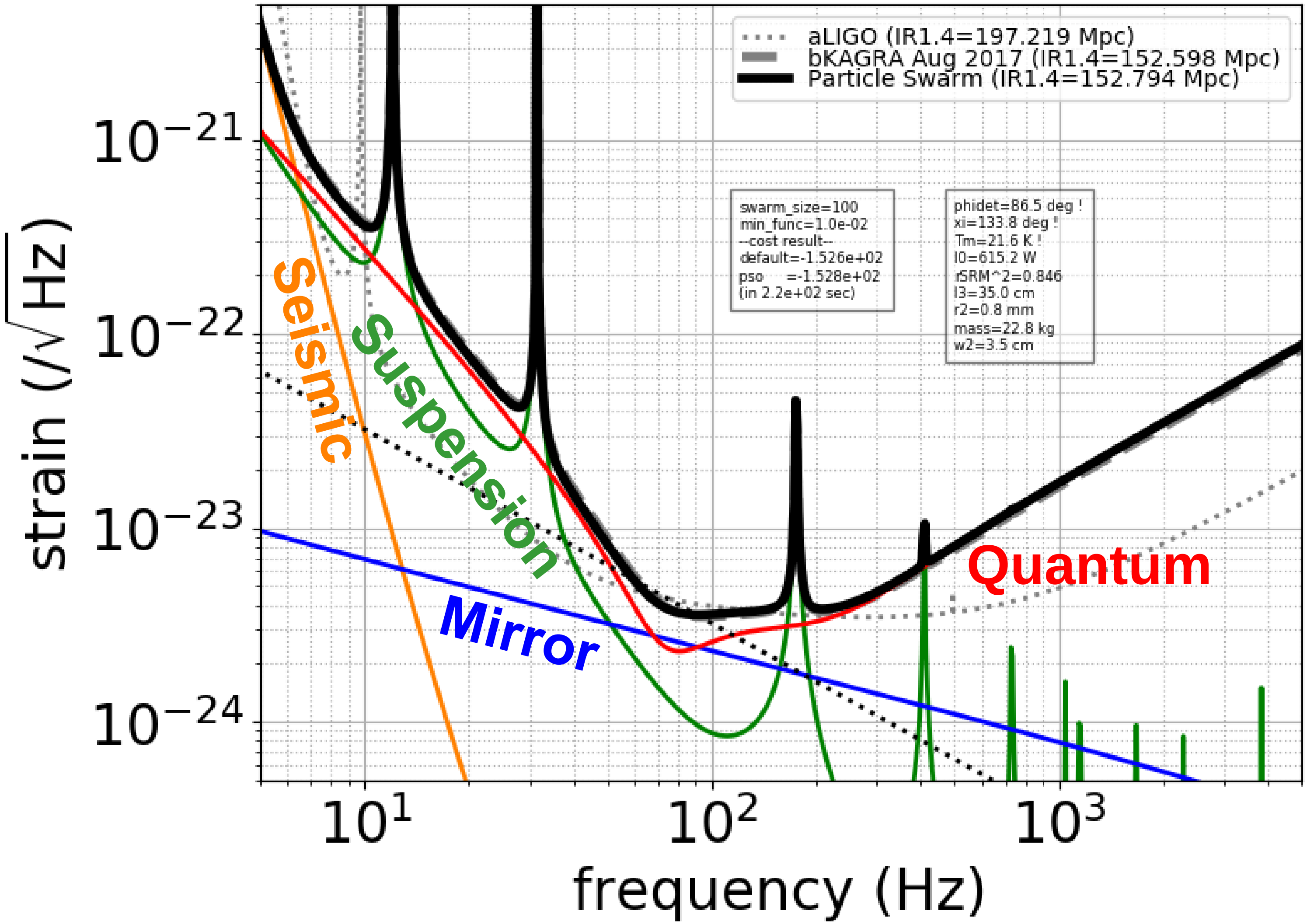
Resolution: 1 K

- Consistent with manual optimization



3 Params, for IR1.4

- Consistent with manual optimization

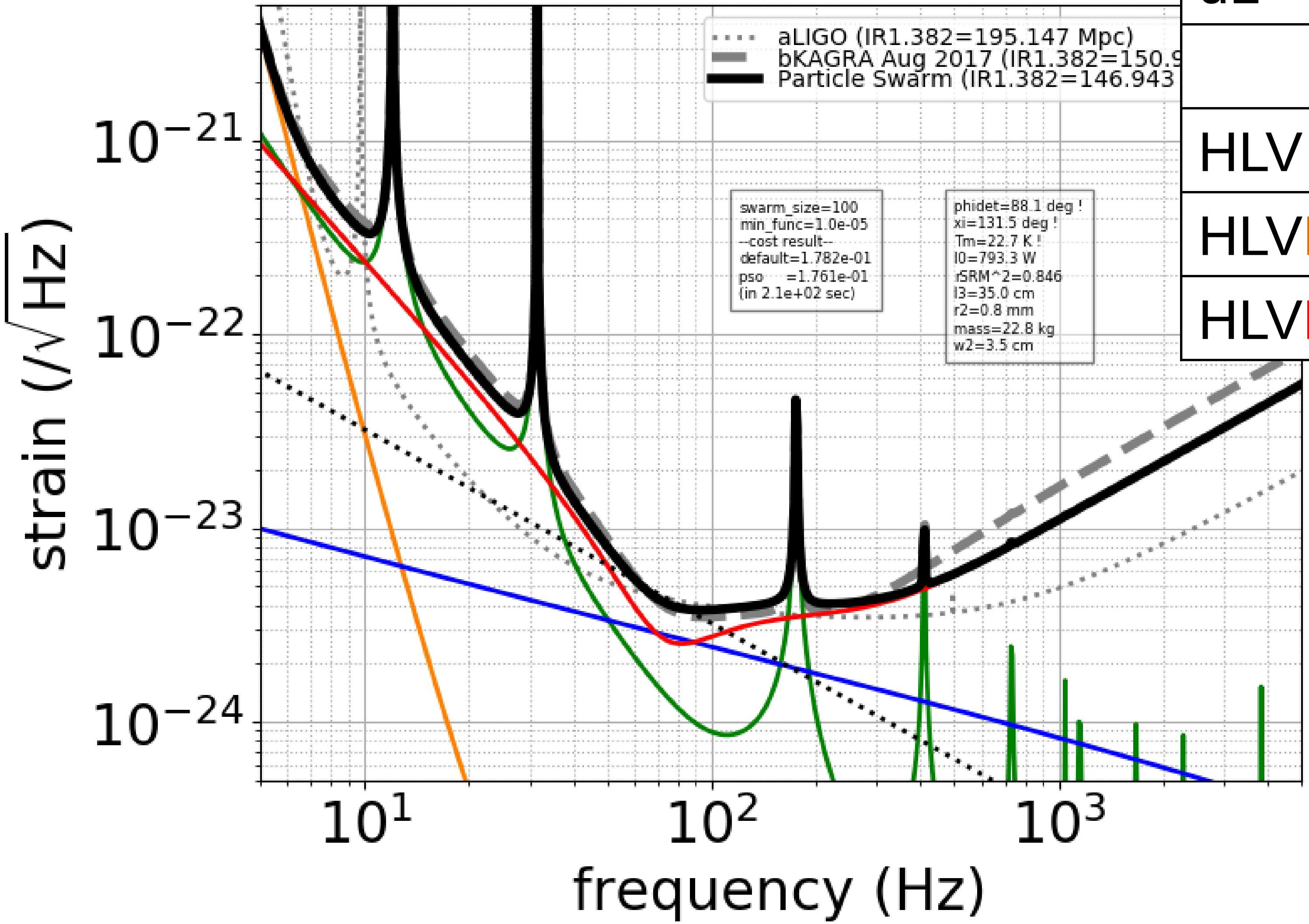


3 Params, GW170817 Distance

- Similar to IR1.4 optimization

  Better SNR gives smaller distance error

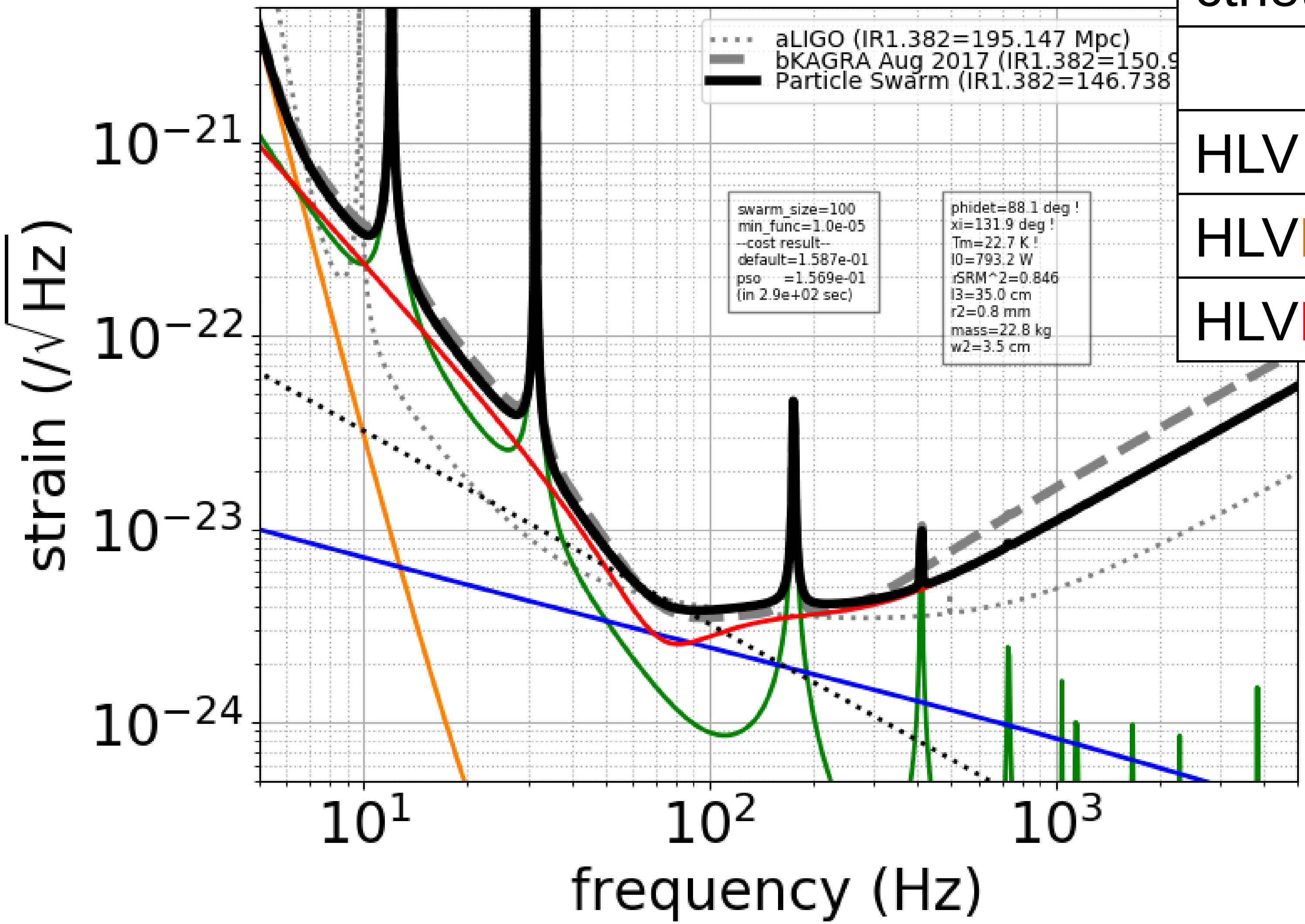
dL = 40 Mpc	
	log(dL)
HLV	26.1 %
HLVK	17.8 %
HLVK+	17.6 %



3 Params, GW170817 Inclination

- Almost same to distance optimization, as expected

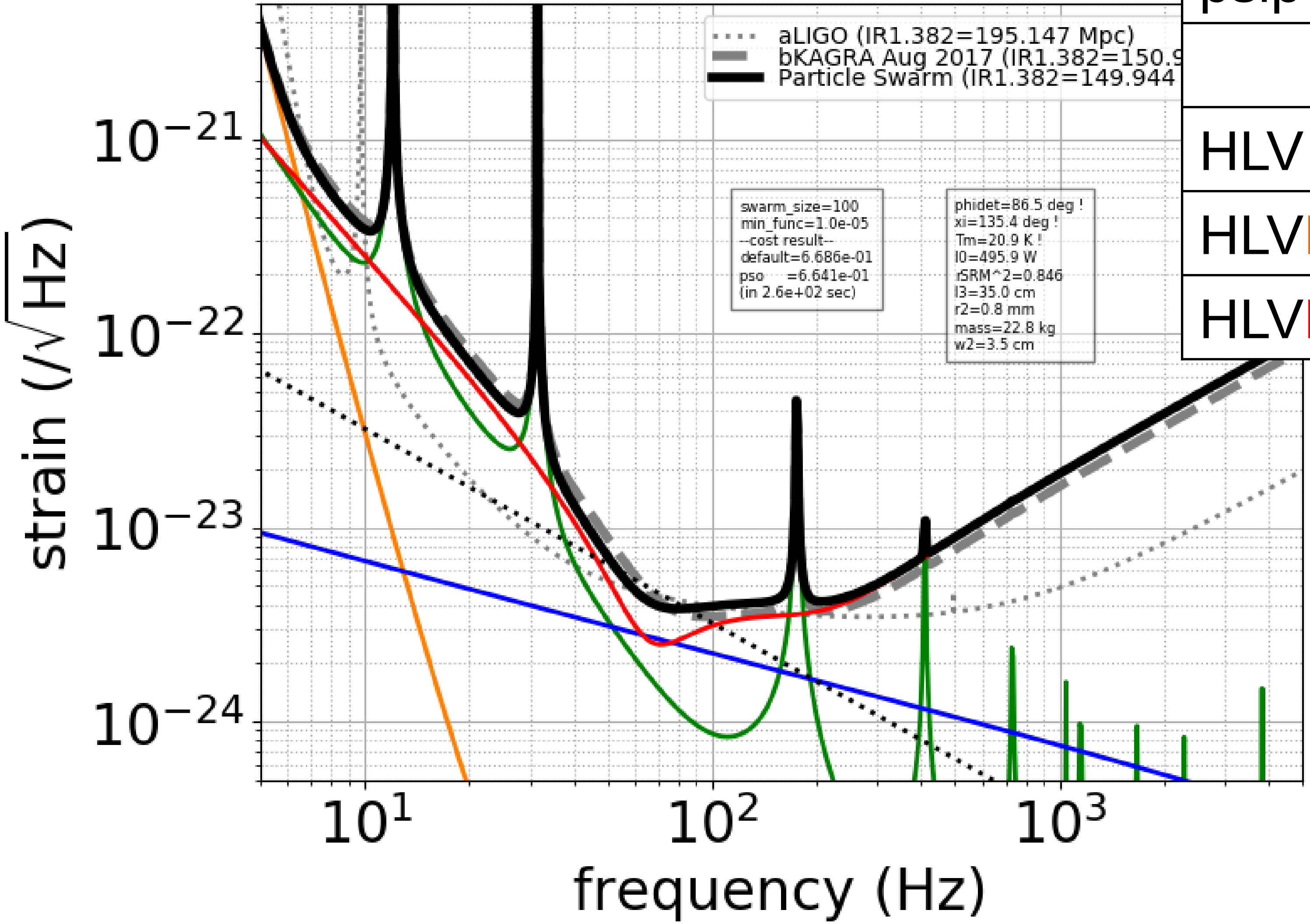
cthetai=cos(28°)	
	cthetai
HLV	0.231
HLVK	0.159
HLVK+	0.157



3 Params, GW170817 Polarization

- Similar to IR1.4 optimization

psip=0	
	psip [rad]
HLV	1.01
HLVK	0.668
HLVK+	0.664



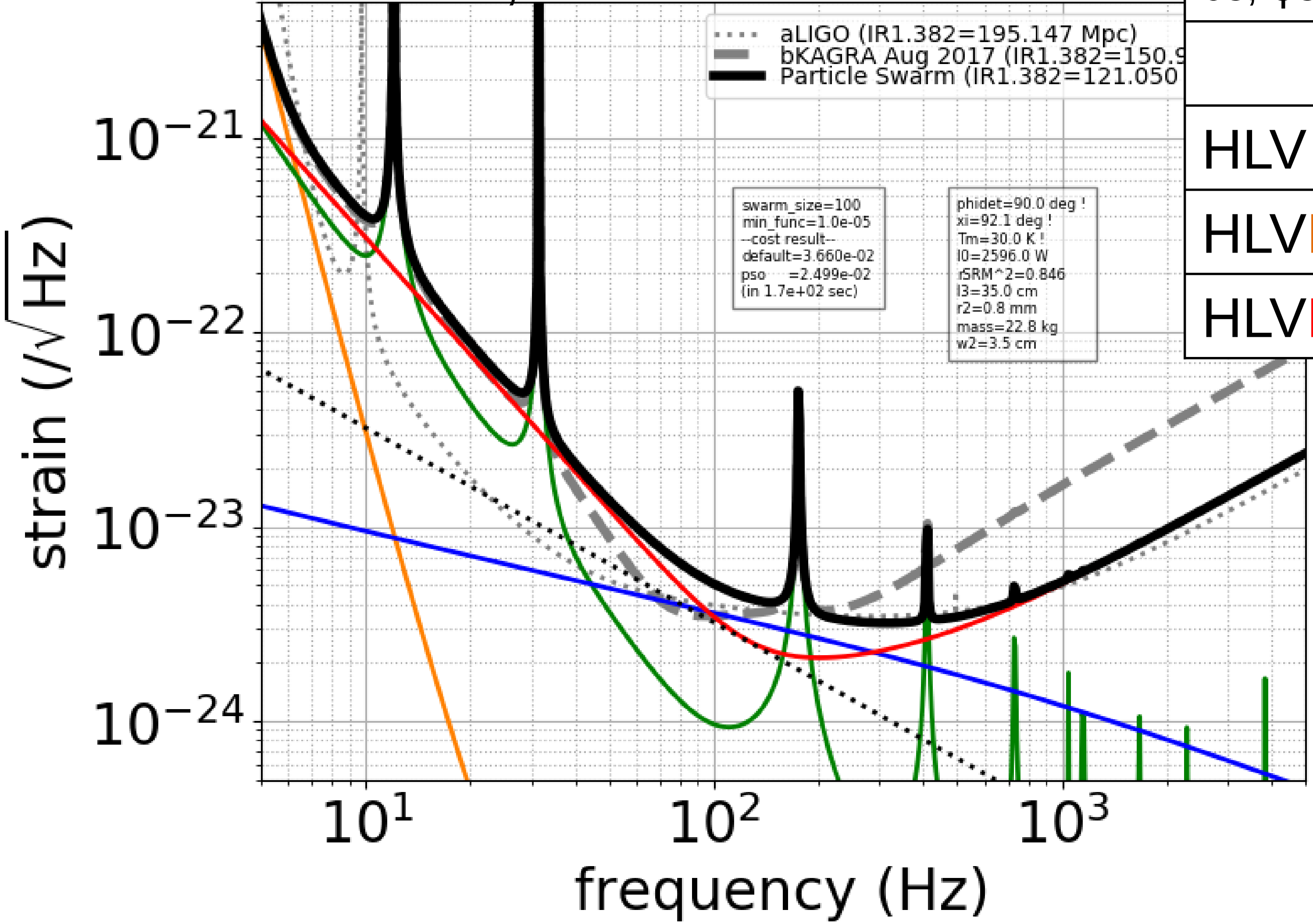
swarm_size=100
min_func=1.0e-05
--cost result--
default=6.686e-01
pso =6.641e-01
(in 2.6e+02 sec)

phidet=86.5 deg !
xi=135.4 deg !
Tm=20.9 K !
I0=495.9 W
rSRM^2=0.846
l3=35.0 cm
r2=0.8 mm
mass=22.8 kg
w2=3.5 cm

3 Params, GW170817 Localization

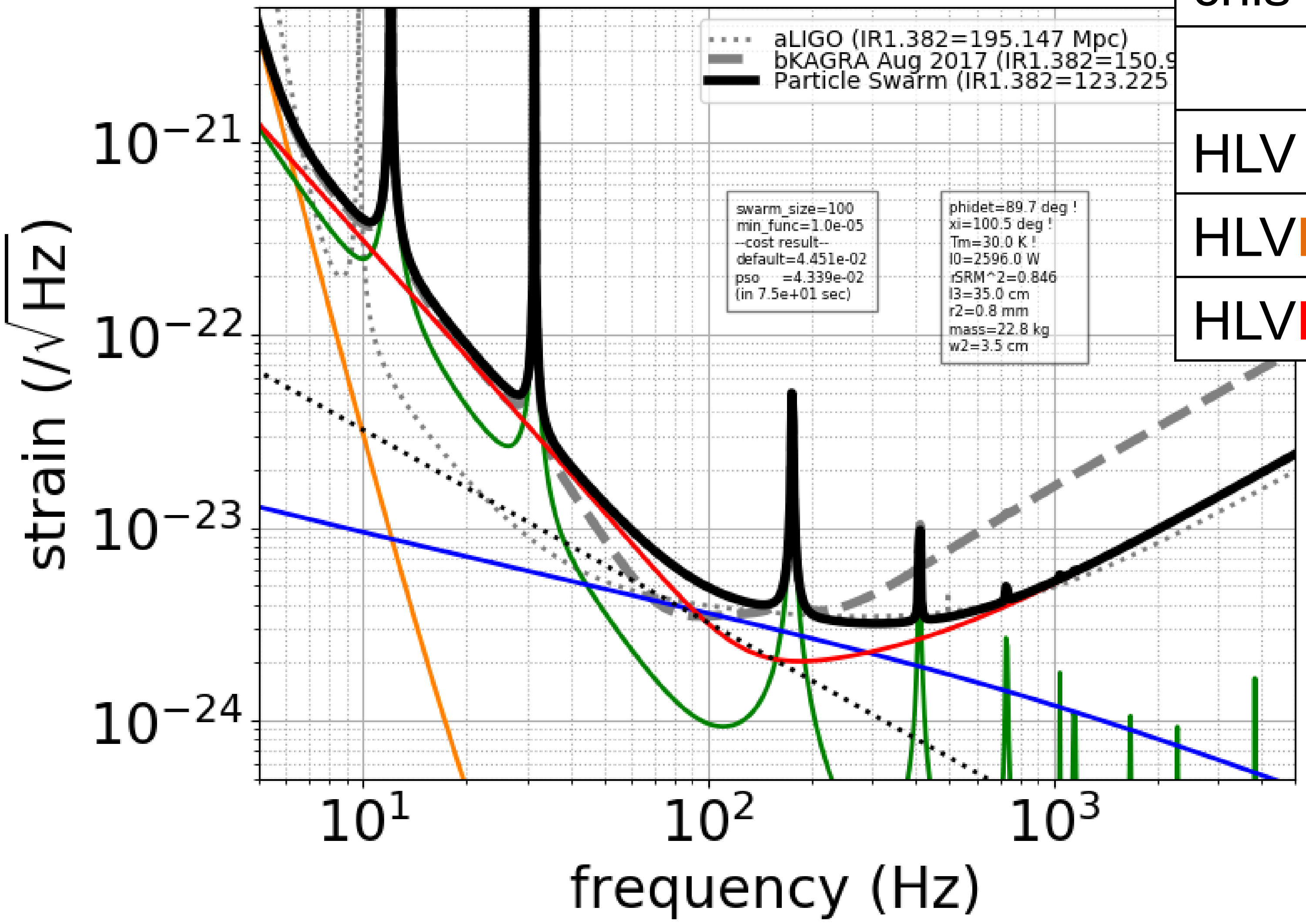
- ~120 Mpc, but better sensitivity at higher frequency
- half event rate, 1.5 times less error

θs, φs=(113.4° , 40°)	
	Ωs [deg]
HLV	0.0771
HLVK	0.0366
HLVK+	0.0250



3 Params, GW170817 SymSpin

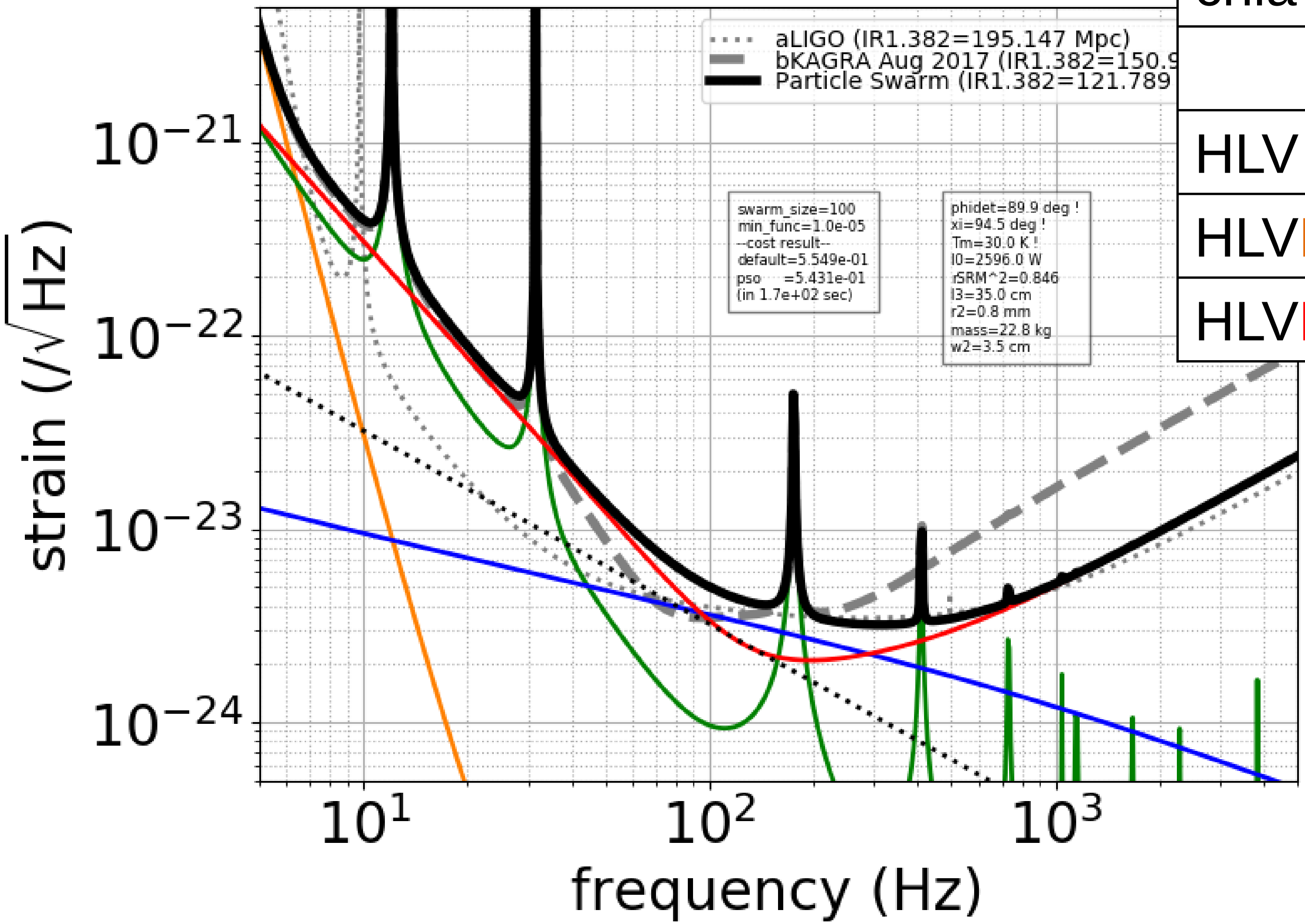
- Almost identical to localization optimization



chis=0	
	chis [rad]
HLV	0.0453
HLVK	0.0445
HLVK+	0.0434

3 Params, GW170817 AsymSpin

- Almost identical to localization optimization

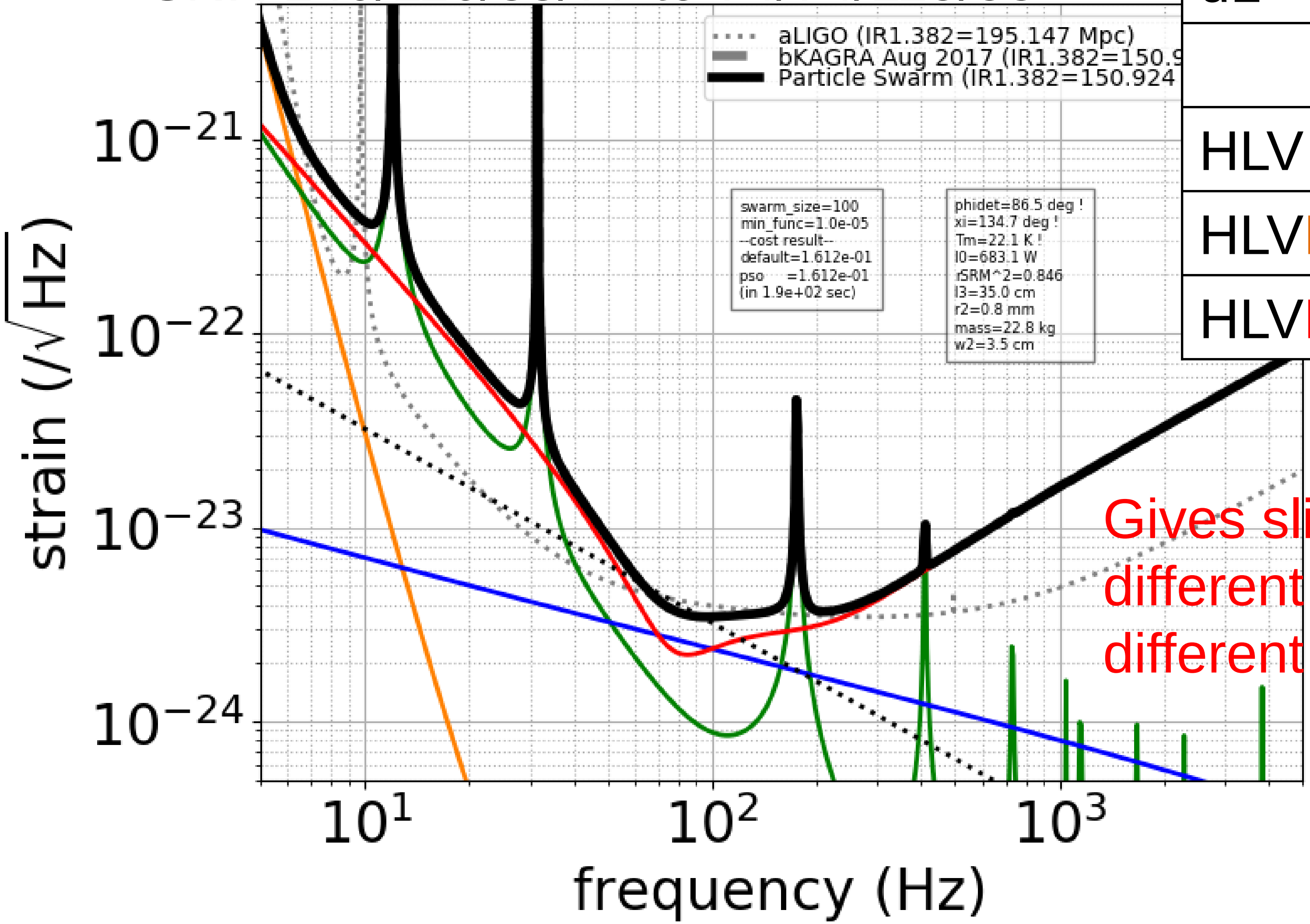


chia=0	
	chia [rad]
HLV	0.563
HLVK	0.555
HLVK+	0.543

3 Params, GW170817mod Distance

- Modified sky location from (113.4° ,40°) to (195° , 40°)
- SNR 140/116/38/77 to 74/74/115/58

dL = 40 Mpc	
	log(dL)
HLV	17.5 %
HLVK	16.1 %
HLVK+	16.1 %



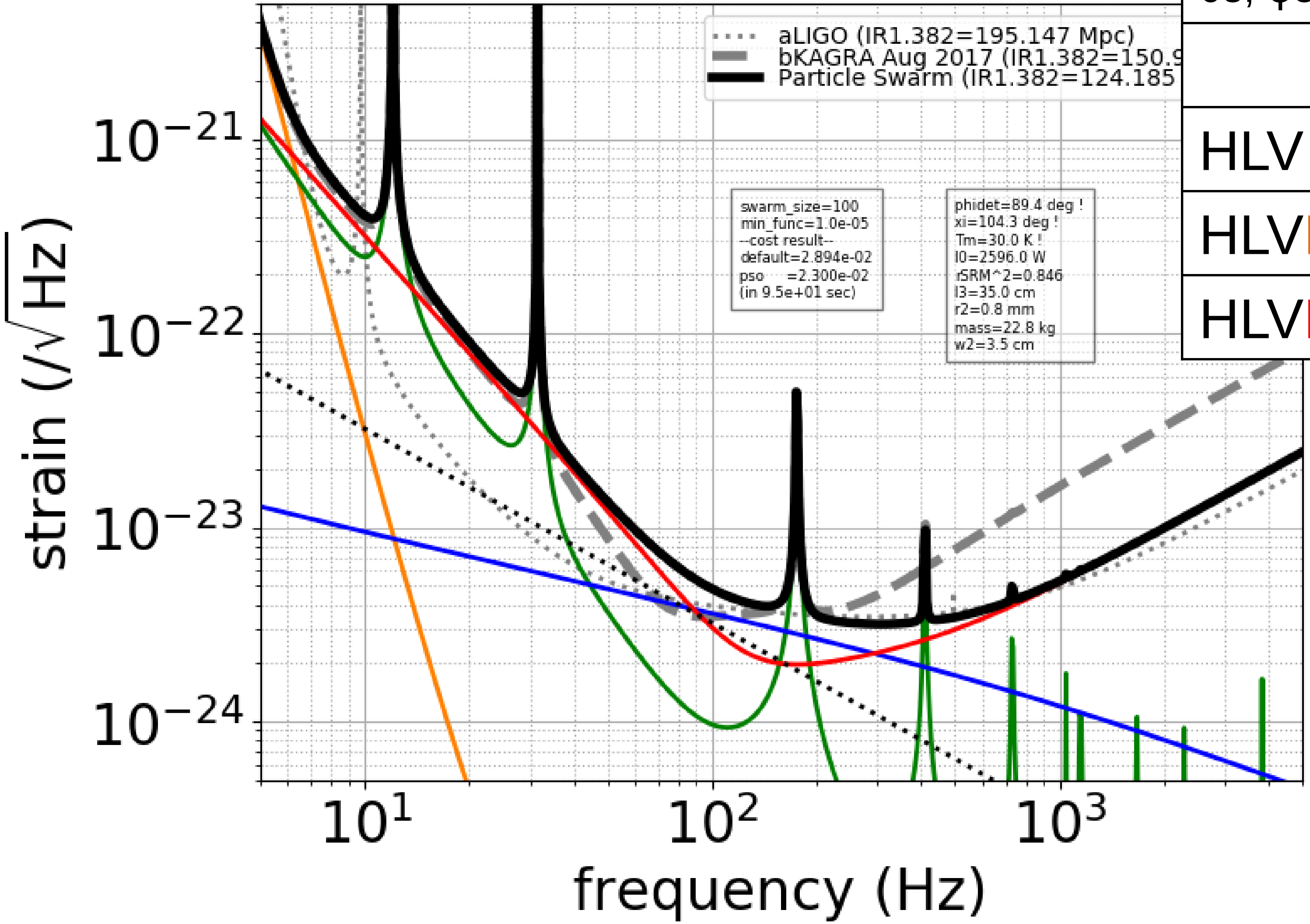
```
swarm_size=100
min_func=1.0e-05
--cost result--
default=1.612e-01
pso     =1.612e-01
(in 1.9e+02 sec)
```

```
phidet=86.5 deg !
xi=134.7 deg !
Tm=22.1 K !
I0=683.1 W
rSRM^2=0.846
l3=35.0 cm
r2=0.8 mm
mass=22.8 kg
w2=3.5 cm
```

Gives slightly different result for different location

3 Params, GW170817mod Localization

- Almost identical with original GW170817 localization even with different sky location

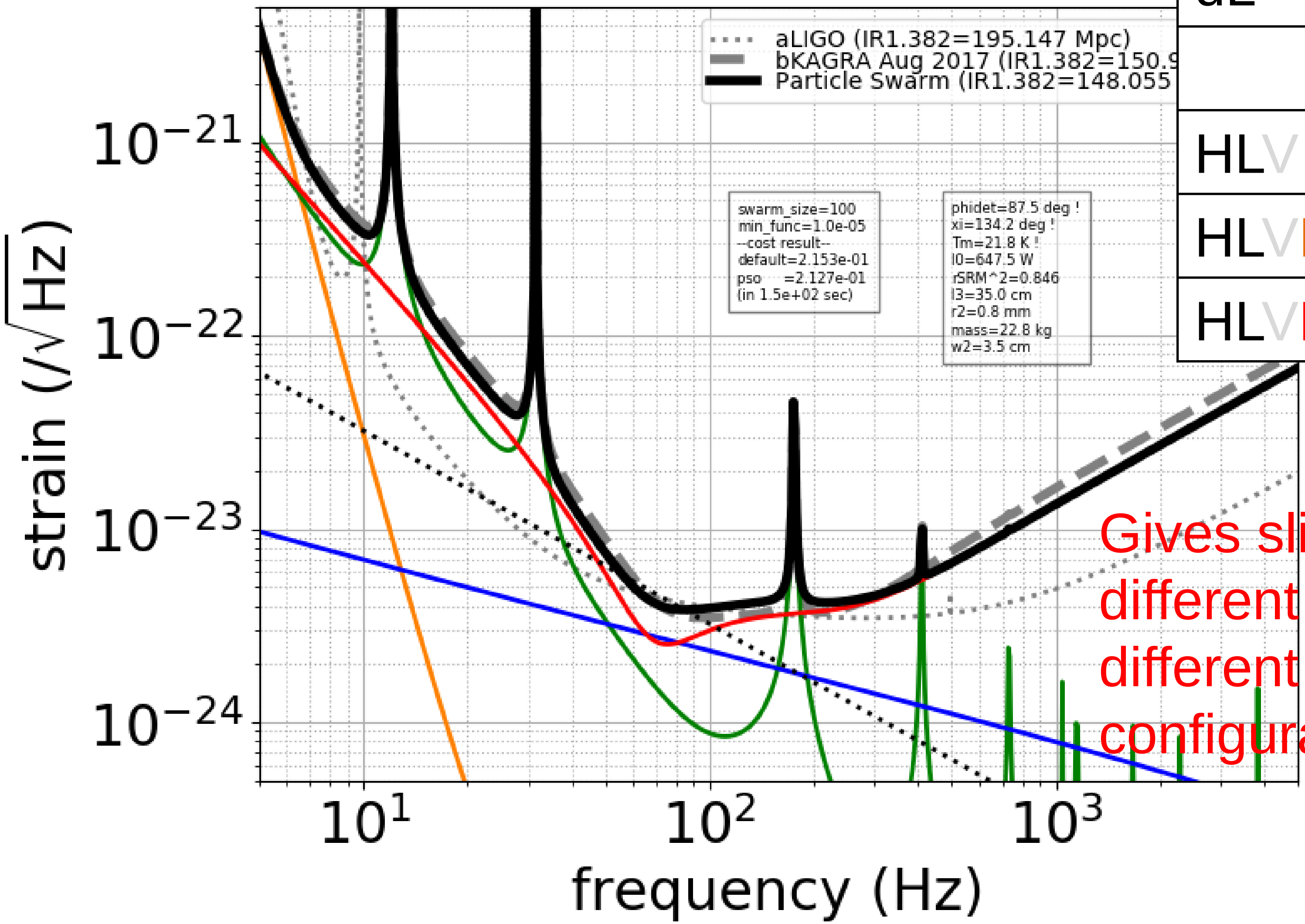


θs, φs=(195° , 40°)	
	Ωs [deg]
HLV	0.0534
HLVK	0.0289
HLVK+	0.0230

3 Params, GW170817 Distance

- No Virgo case

dL = 40 Mpc	
	log(dL)
HLV	193 %
HLVK	21.5 %
HLVK+	21.3 %

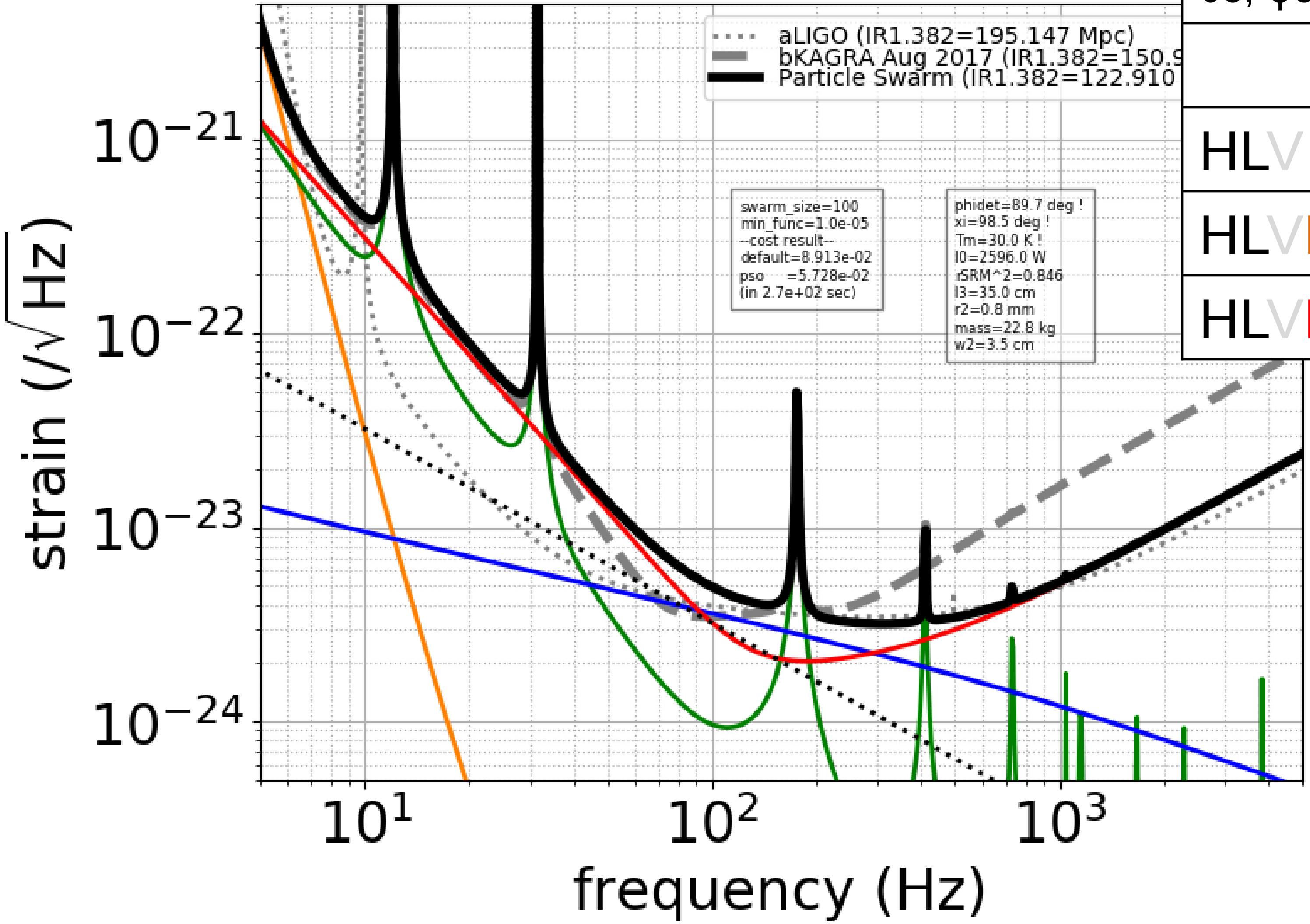


Gives slightly different result for different network configuration

3 Params, GW170817 Localization

- Almost identical with original GW170817 localization even with no Virgo case

θs, φs=(113.4° , 40°)	
	Ωs [deg]
HLV	3.34
HLVK	0.0891
HLVK+	0.0573



Thoughts on Parameter Estimation

- KAGRA sensitivity design can be done with PSO easily

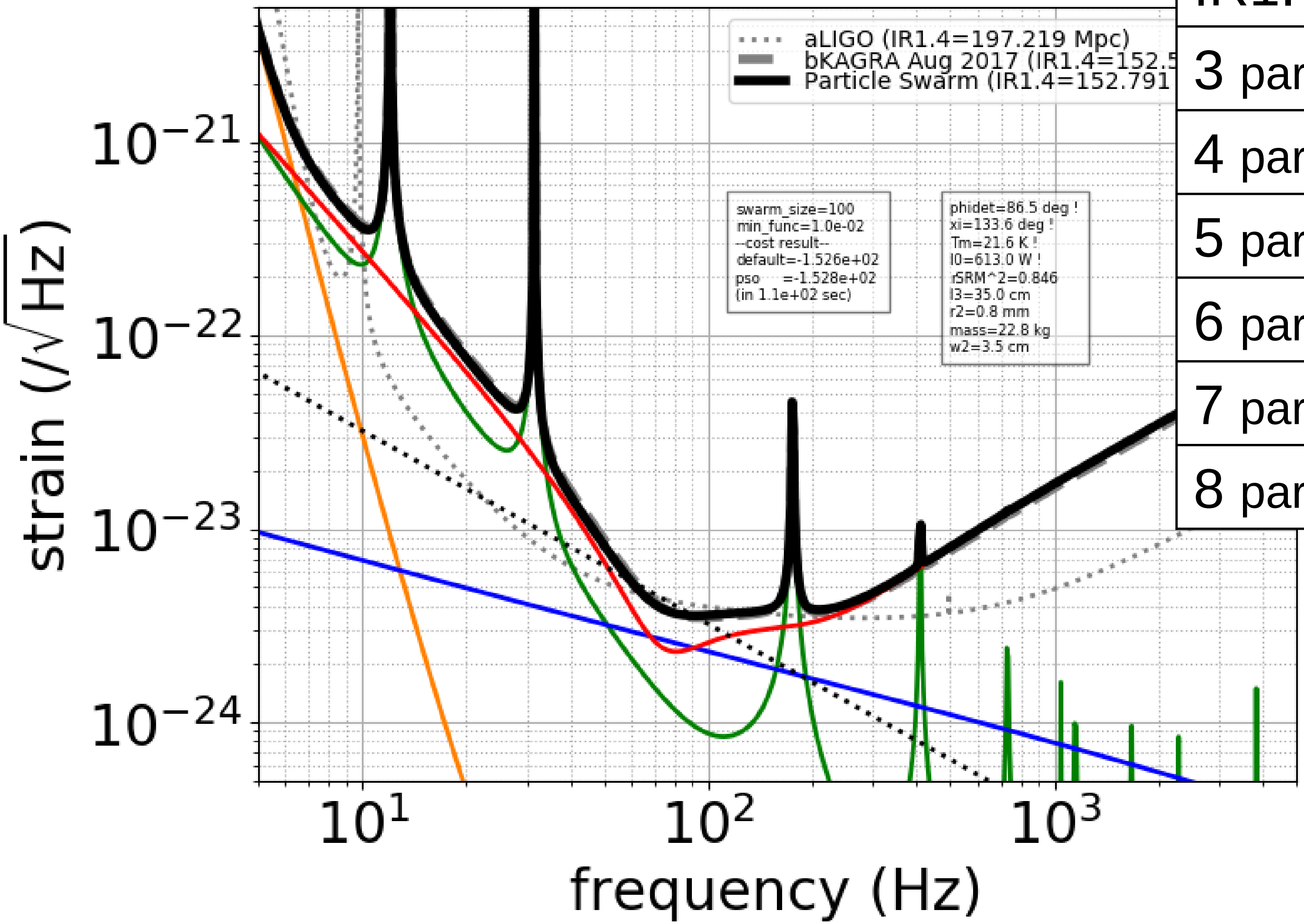
- IR optimization is basically optimum for distance, inclination, polarization estimations (depends on source location and detector network configuration)
- For sky localization and spin parameters, higher frequency sensitivity is important
- Even with higher frequency optimization at the cost of inspiral range degradation, improvement in binary parameter estimation is small

- **IR optimization (event rate optimization) seems like a reasonable choice**

4 Params, for IR1.4

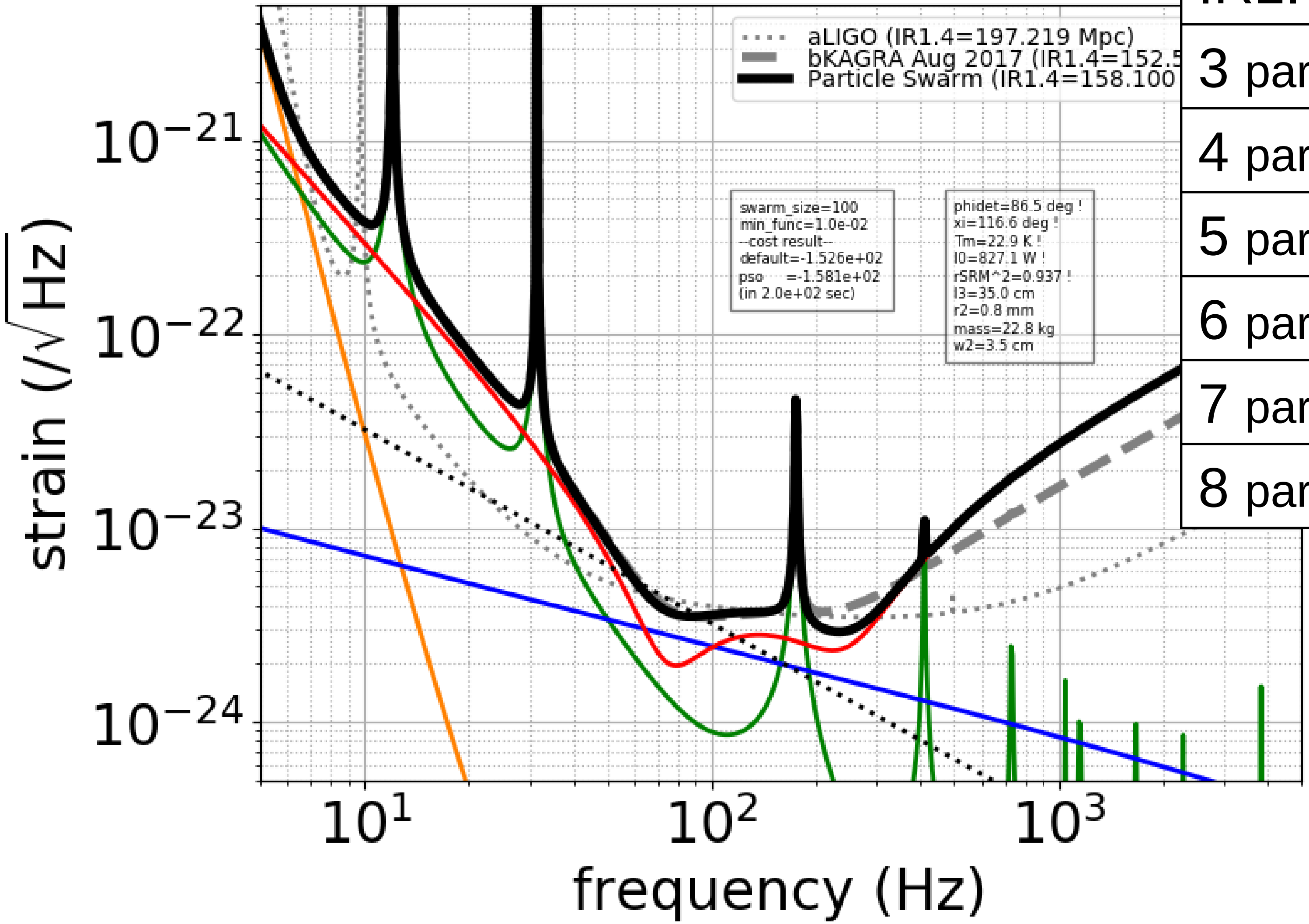
- Input power at maximum is good for IR1.4

IR1.4 [Mpc]	
3 params	153
4 params	153
5 params	
6 params	
7 params	
8 params	



5 Params, for IR1.4

- SRM reflectivity of 88% gives slightly better IR1.4



```
swarm_size=100
min_func=1.0e-02
--cost result--
default=-1.526e+02
pso     =-1.581e+02
(in 2.0e+02 sec)
```

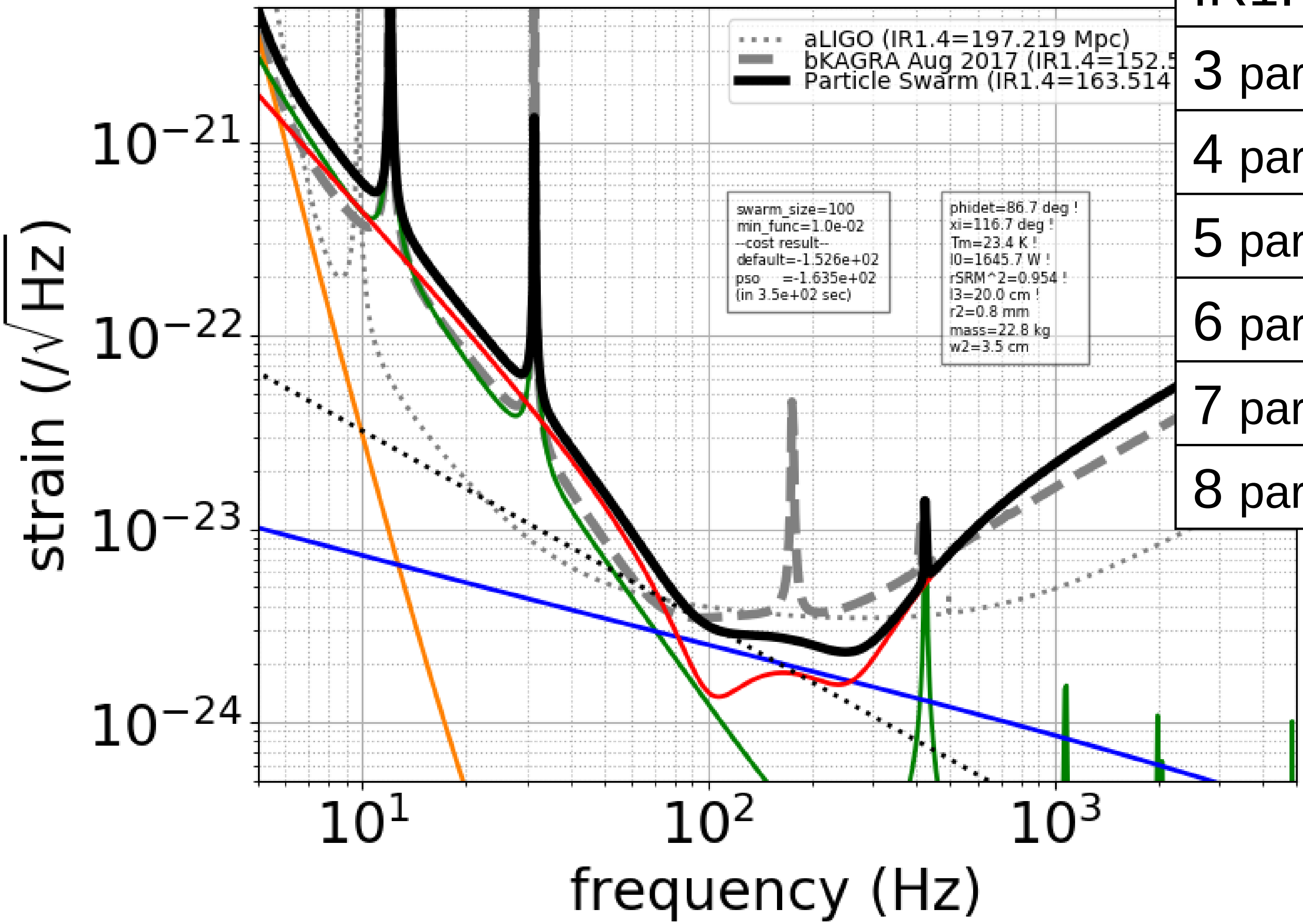
```
phidet=86.5 deg !
xi=116.6 deg !
Tm=22.9 K !
I0=827.1 W !
rSRM^2=0.937 !
l3=35.0 cm
r2=0.8 mm
mass=22.8 kg
w2=3.5 cm
```

IR1.4 [Mpc]	
3 params	153
4 params	153
5 params	158
6 params	
7 params	
8 params	

6 Params, for IR1.4

- Wire length is shorter the better for IR1.4

IR1.4 [Mpc]	
3 params	153
4 params	153
5 params	158
6 params	164
7 params	
8 params	

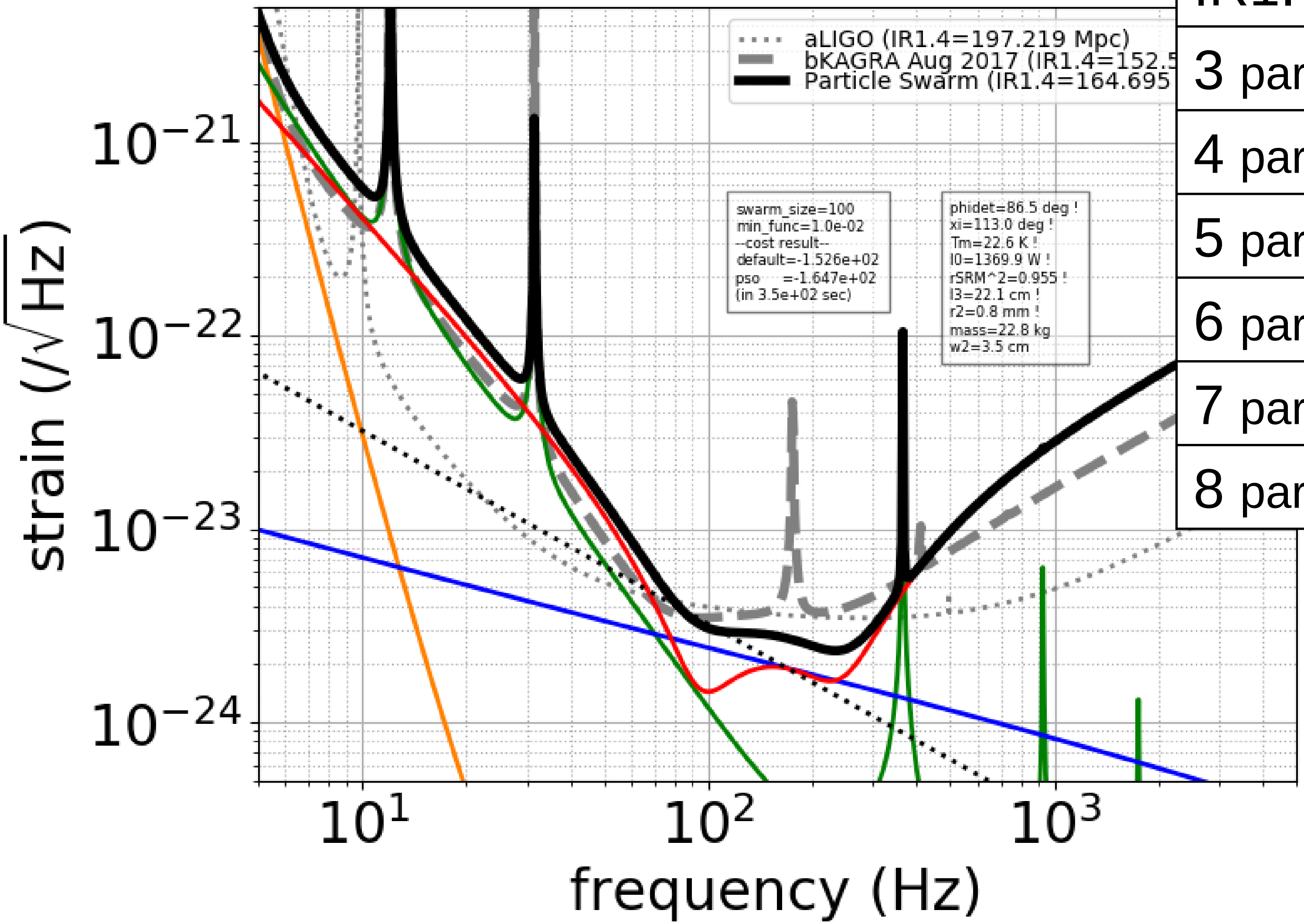


swarm_size=100
min_func=1.0e-02
--cost result--
default=-1.526e+02
pso =-1.635e+02
(in 3.5e+02 sec)

phidet=86.7 deg !
xi=116.7 deg !
Tm=23.4 K !
I0=1645.7 W !
rSRM^2=0.954 !
l3=20.0 cm !
r2=0.8 mm
mass=22.8 kg
w2=3.5 cm

7 Params, for IR1.4

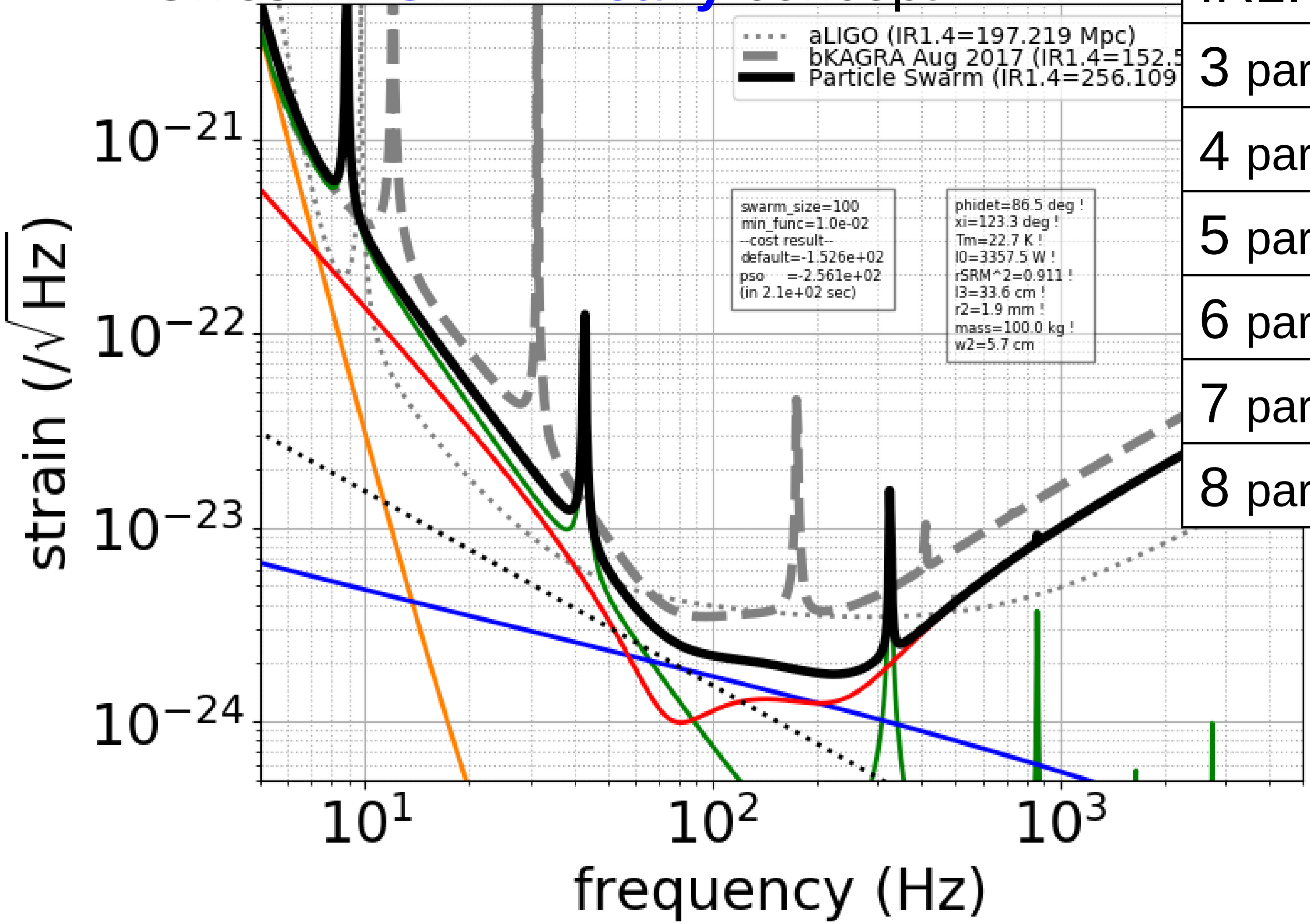
- Default wire radius is OK for IR1.4



IR1.4 [Mpc]	
3 params	153
4 params	153
5 params	158
6 params	164
7 params	165
8 params	

8 Params, for IR1.4

- Heavier mirror gives very good IR1.4
- Gives **KAGRA+ Heavy** concept



```
swarm_size=100
min_func=1.0e-02
--cost result--
default=-1.526e+02
pso     =-2.561e+02
(in 2.1e+02 sec)
```

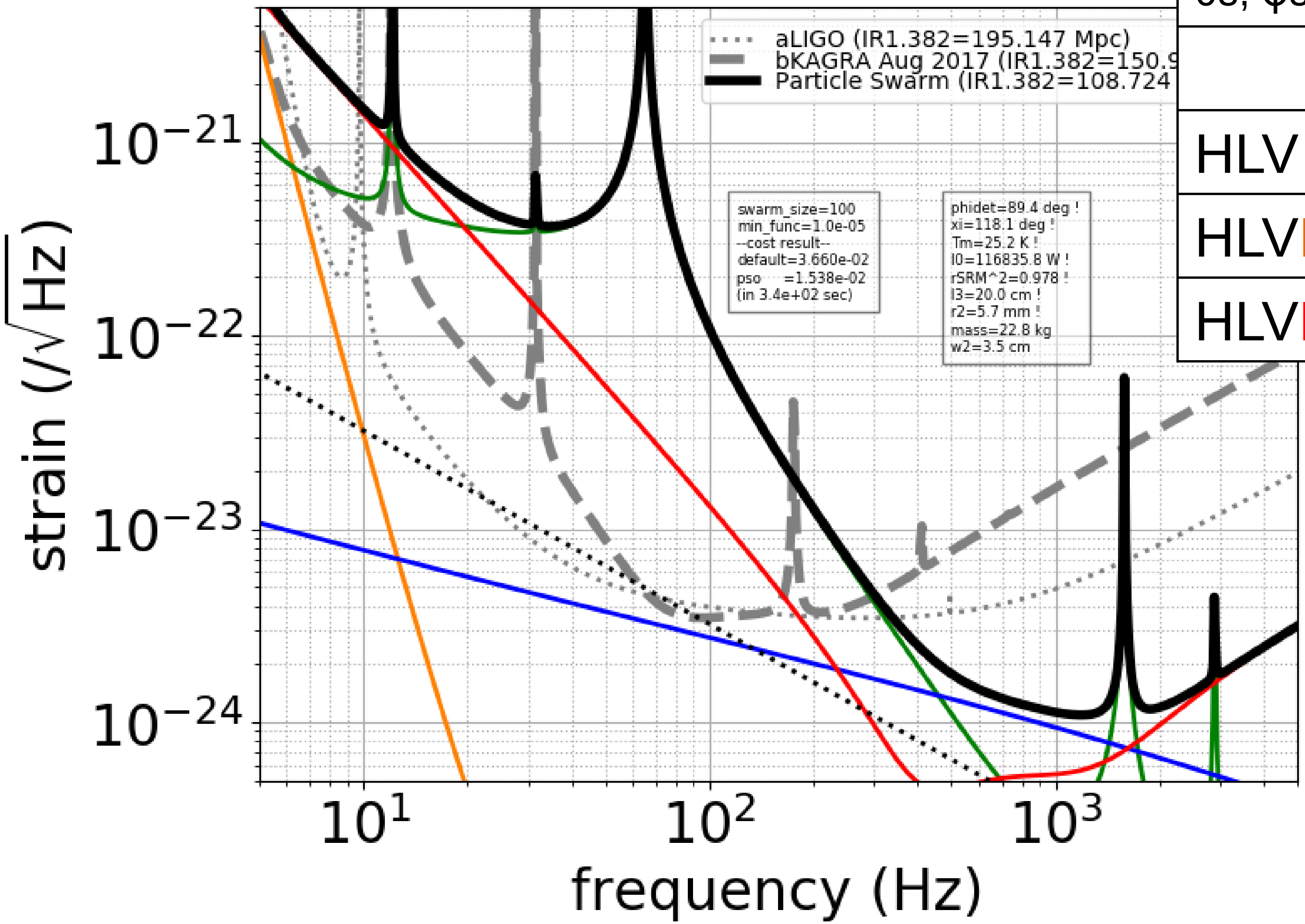
```
phidet=86.5 deg !
xi=123.3 deg !
Tm=22.7 K !
I0=3357.5 W !
rSRM^2=0.911 !
l3=33.6 cm !
r2=1.9 mm !
mass=100.0 kg !
w2=5.7 cm
```

IR1.4 [Mpc]	
3 params	153
4 params	153
5 params	158
6 params	164
7 params	165
8 params	256

7 Params, GW170817 Localization

- Gives **KAGRA+ HF** concept (without squeezing)

θs, φs=(113.4° , 40°)	
	Ωs [deg]
HLV	0.0771
HLVK	0.0366
HLVK+	0.0154



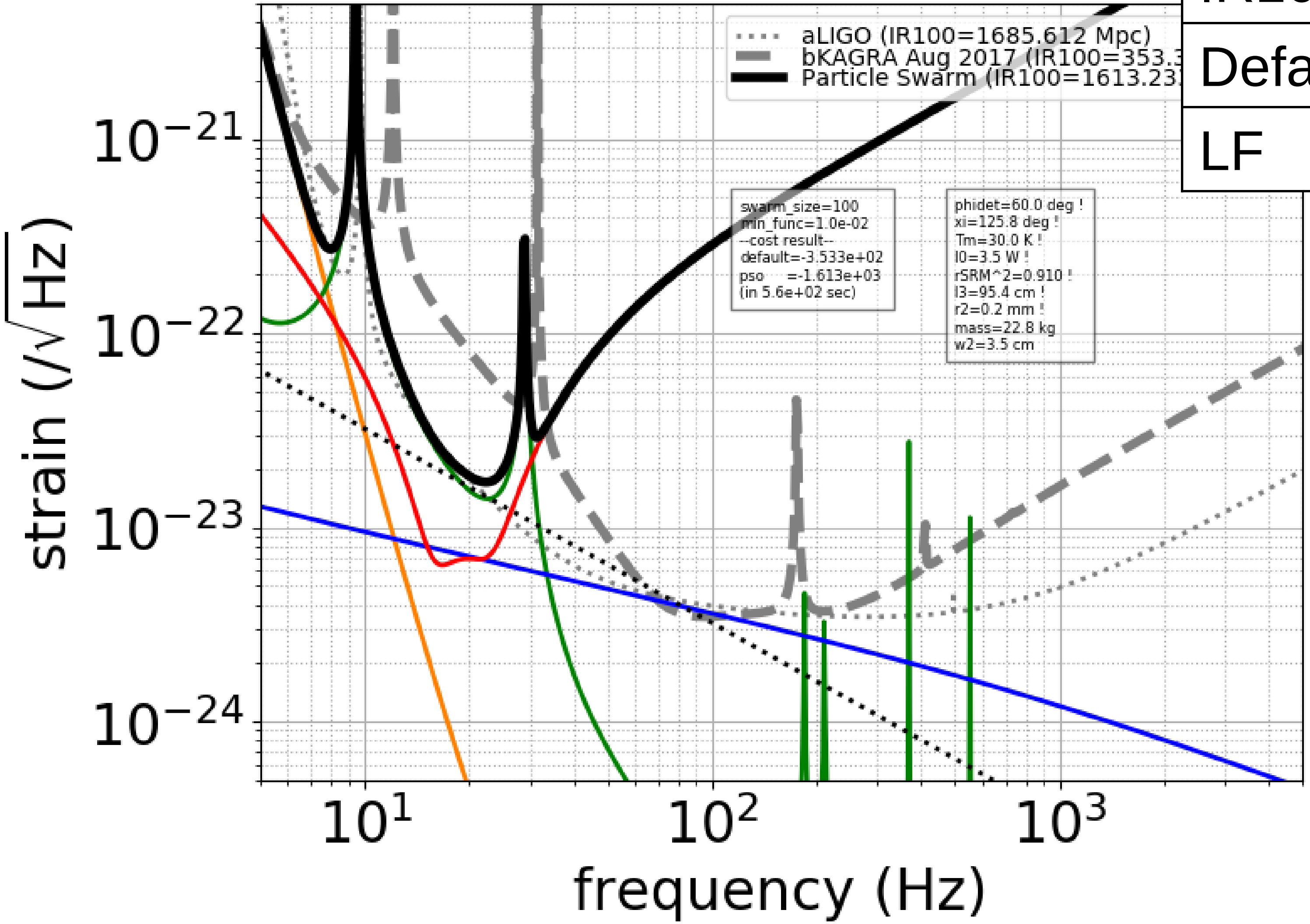
```
swarm_size=100
min_func=1.0e-05
--cost result--
default=3.660e-02
pso     =1.538e-02
(in 3.4e+02 sec)
```

```
phidet=89.4 deg !
xi=118.1 deg !
Tm=25.2 K !
I0=116835.8 W !
rSRM^2=0.978 !
l3=20.0 cm !
r2=5.7 mm !
mass=22.8 kg
w2=3.5 cm
```

7 Params/Large Detune, for IR100

- Gives **KAGRA+ LF** concept (without IM and ambient heat parameter tuning)

IR100 [Mpc]	
Default	353
LF	1613



Thoughts on KAGRA+

- Even with large degradation in inspiral range, parameter estimation improvement with **HF** is limited
- Design changes in cryogenics are necessary to realize **LF**
- **Heavy** mirror improves inspiral range a lot (which leads to reduction in distance, inclination, polarization angle error)

- Anyway, broader sensitivity improvement is good?
- I also want to see optimization for detection rate (and detection rate with PE error smaller than threshold)
- I also want to include IM mass and wire, squeezing parameters, ambient heat parameters for optimization

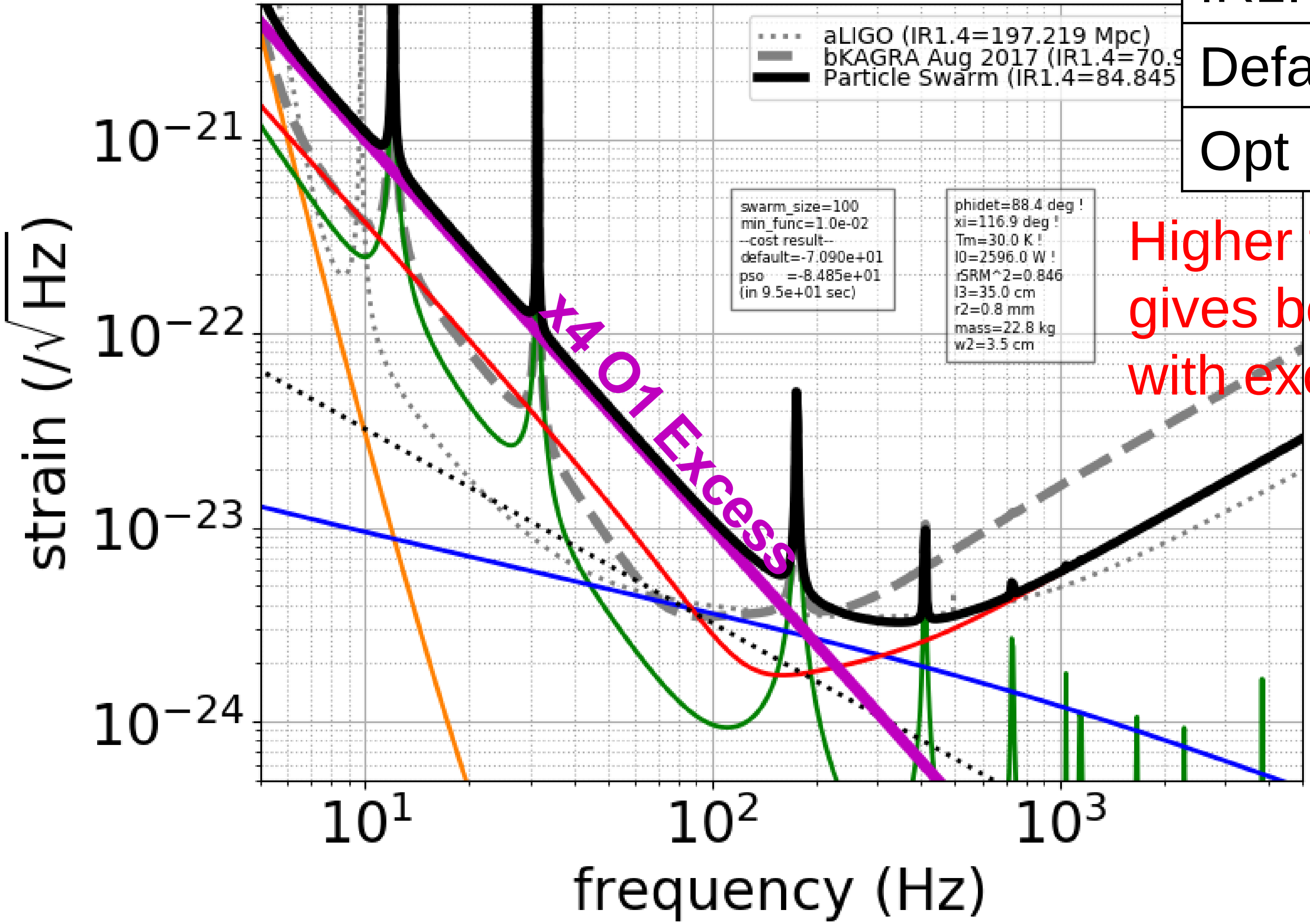
$$K_{\mathrm{abs}} = 2\beta_{\mathrm{sub}} t_m P_{\mathrm{mich}} + \gamma_{\mathrm{coa}} P_{\mathrm{circ}} + K_{\mathrm{rad}}$$

JGW-T1707038

4 Params, for IR1.4 with Excess

- PSO can be used to optimize configuration with excess noise during commissioning

IR1.4 [Mpc]	
Default	71
Opt	85



Higher temperature gives better IR1.4 with excess noise

Paper in Preparation

- First demonstration of PSO for GW detector design
- Maybe I don’t want to go into too much details of what is the best figure of merit
- How to validate PSO result, and show that PSO is useful?
- Focus on KAGRA upgrade and not PSO?

Summary

- GW astronomy started, and we need new figure of merits to design the sensitivity of GW detectors
- Cryogenics add more complexity in GW detector sensitivity design
- Developed a tool to optimize KAGRA sensitivity using particle swarm optimization
- PSO can be implemented easily, and it looks like it gives reasonable results with tolerable amount of time
- Cost functions available so far
  - inspiral range (SNR)
  - strain
  - binary parameter estimation error from Fisher analysis
- To be done:
  - optimization for detection rate
  - add more IFO parameters to be optimized (IM, squeezing, etc.)
  - faster calculation