PSO Applications

Gravitational Wave Data Analysis School in China

Soumya D. Mohanty

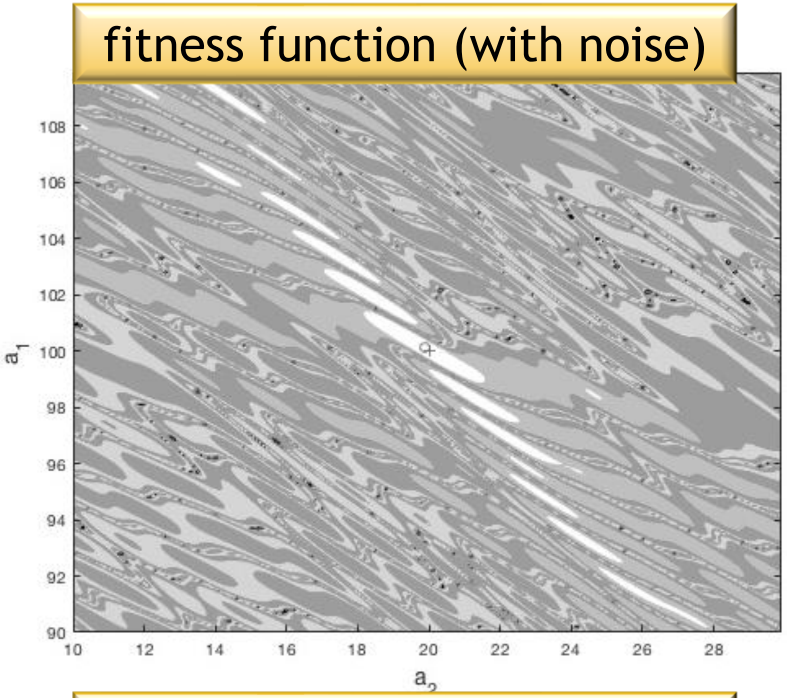
UTRGV

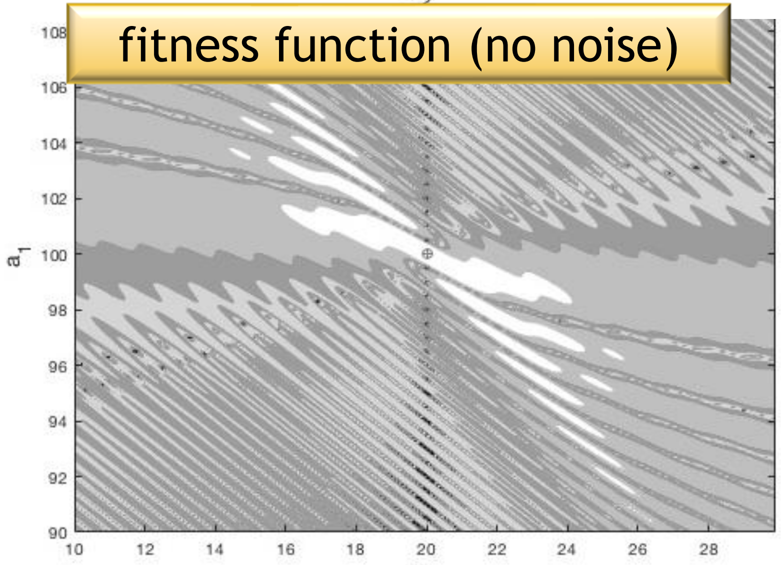
Toy application

- Use PSO to find the GLRT and MLE for data containing the quadratic chirp signal added to colored Gaussian noise

$$L_G = \max_{\Theta} \langle \bar{y}, \bar{q}(\Theta) \rangle^2$$

  - The fitness function to be minimized is $-\langle \bar{y}, \bar{q}(\Theta) \rangle^2$
  - $\Theta = (a_1, a_2, a_3)$
- Results for WGN can be found in the textbook (Chapter 5)





Fitness function cross-section

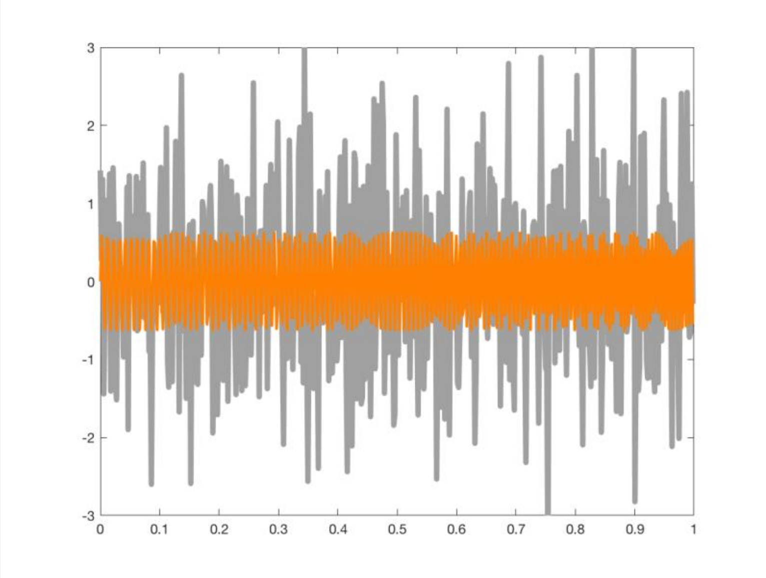
- Quadratic chirp: Multiple local minima in fitness function even in the absence of noise
- PSO has to search for the global minimum while avoiding local minima

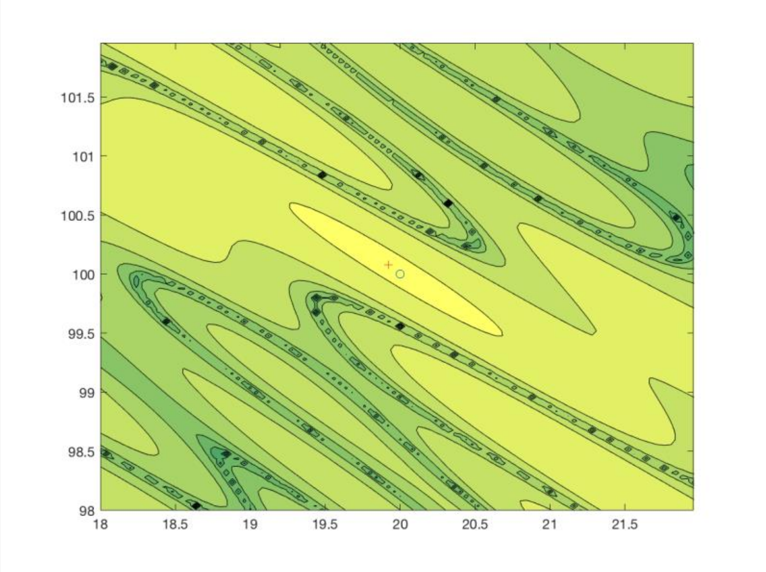
Tuning PSO

Simulate data realizations based on assumed noise model

Each data realization leads to a different fitness function $\Rightarrow$ Variation in PSO performance

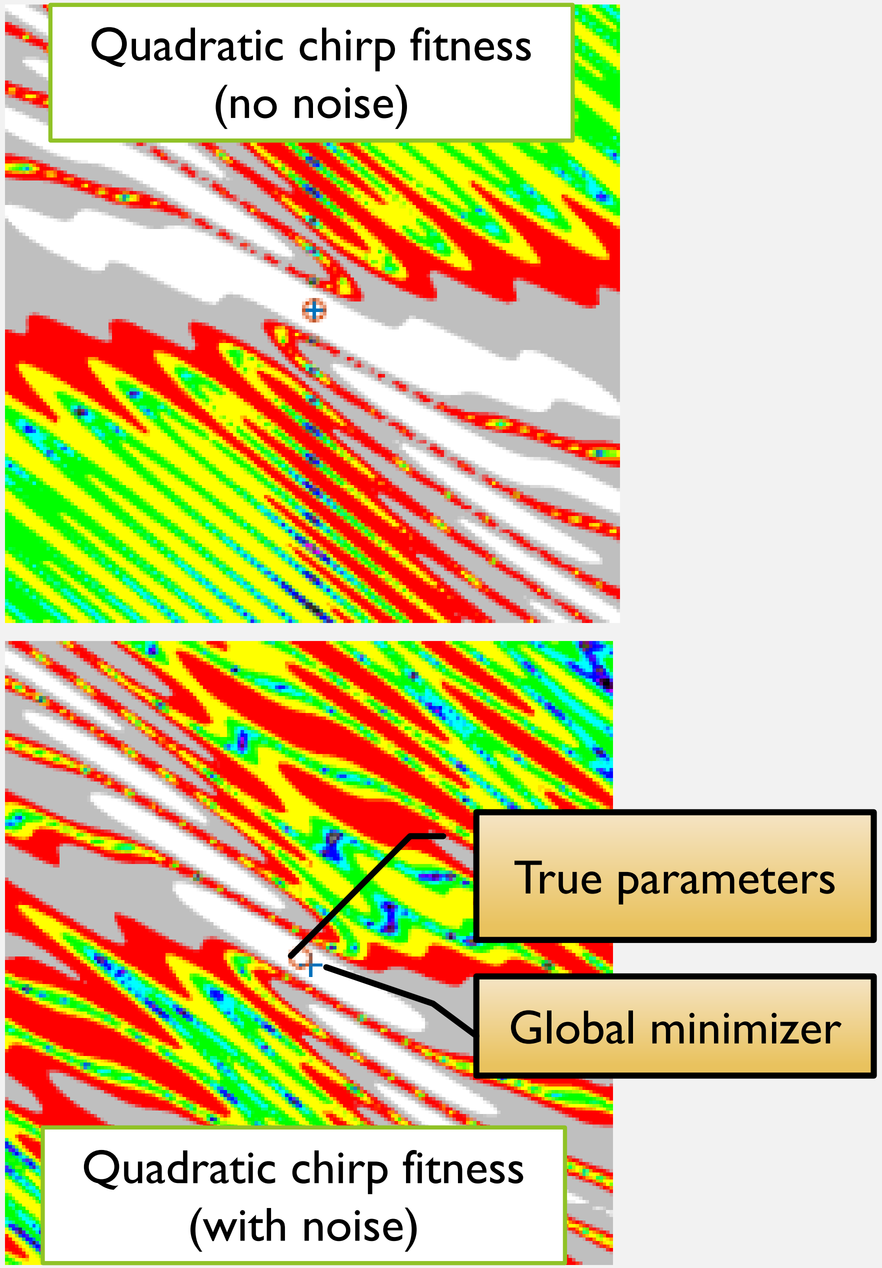
Use statistical approach to tune PSO





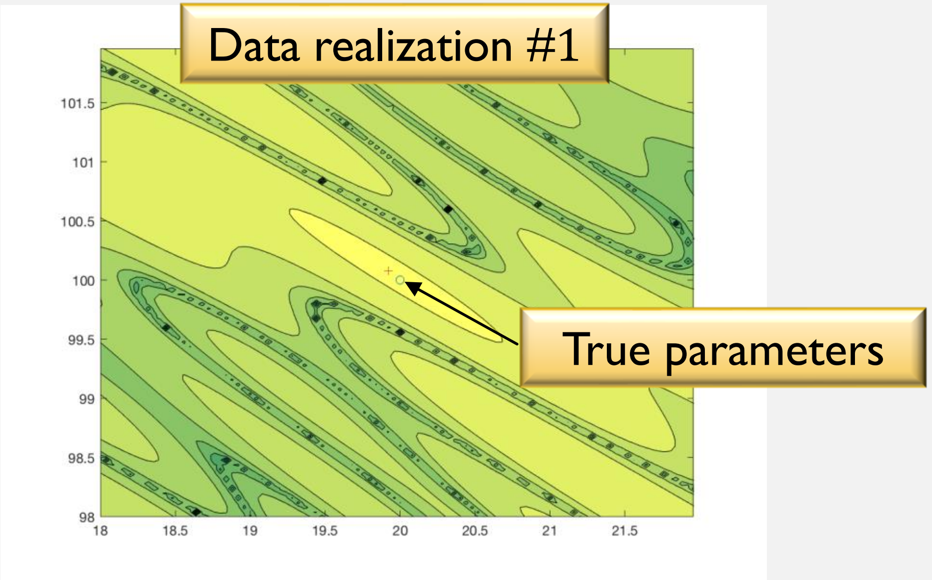
DATA SIMULATION

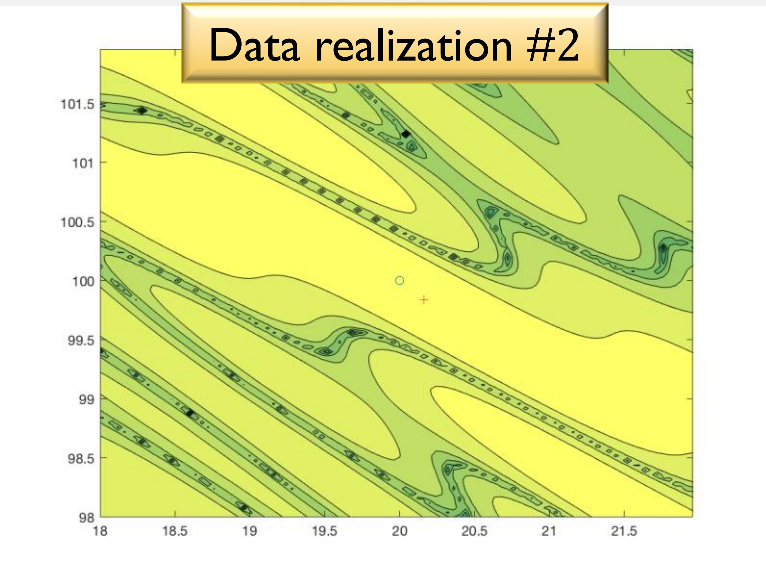
- Keep the parameters of the true signal (e.g. quadratic chirp) fixed
- Add different noise realizations
- Each data realization ⇒ one fitness function realization



Statistical tuning approach

- For the same true signal parameters, the global minimizer will be different for different data realizations
- The best fitness value will always occur away from the true parameters
  - This is why we get error in parameter estimates in the presence of noise
- This fact can be used to develop a tuning procedure



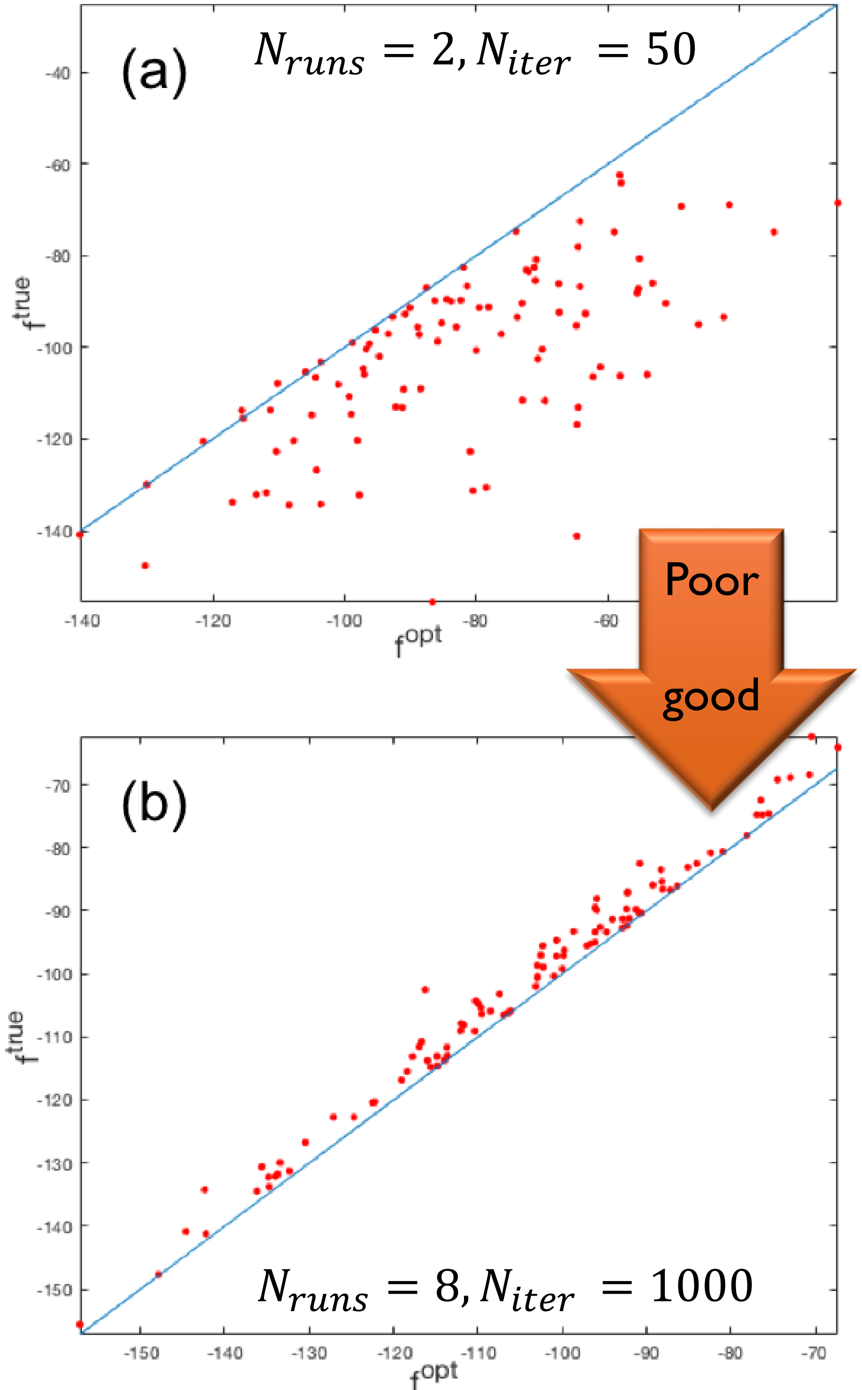


PSO Tuning for regression problems

- Key idea: The global minimum must be lower than the fitness at the true parameters

$$f^{opt} < f^{true}$$

- PSO is working well if this condition is satisfied for a sufficiently high fraction of data realizations
- Proposed in:
  - Wang, Mohanty, Physical Review D, 2010
  - Normandin, Mohanty, Weerathunga, Physical Review D, 2018



Parametric regression

- The true parameters are known for simulated data
- $\Rightarrow$ Possible to check $f^{opt} < f^{true}$ for each data realization
- Set up a grid of values in
  - N_{iter}: Number of iterations
  - N_{runs}: Number of runs in BMR strategy
- For each combo (N_{iter}, N_{runs}): Get fraction X of N data realizations where this condition is satisfied
- Get all (N_{iter}, N_{runs}) for which X is below some preset value
- Pick the combo in this set with the lowest computational cost

Results

Chapter 5 of "Swarm intelligence methods for statistical regression"

Parametric regression

- Quadratic chirp:

$$f\left(x;\bar{\theta}\right) = A\sin(2\pi\Phi(x))\,;\ \bar{\theta} = (A, a_1, a_2, a_3)$$
$$\Phi(x) = a_1 x + a_2 x^2 + a_3 x^3$$

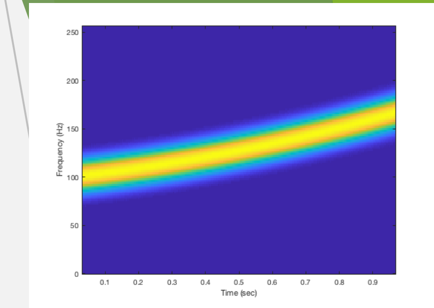
- True parameters

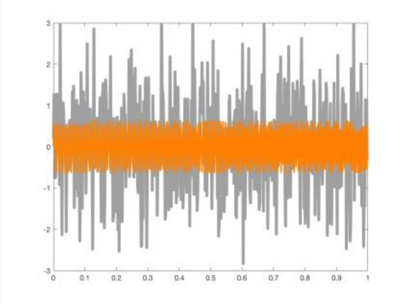
$$A = 0.625, a_1 = 100, a_2 = 20, a_3 = 10$$

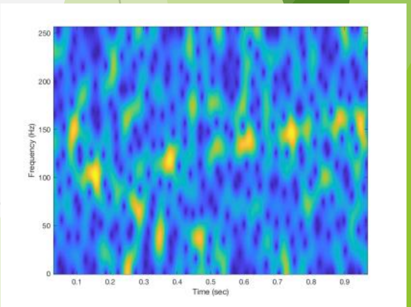
- White Gaussian Noise (WGN): iid Normal with mean =0 and variance =1
- 100 data realizations
- PSO Search space:

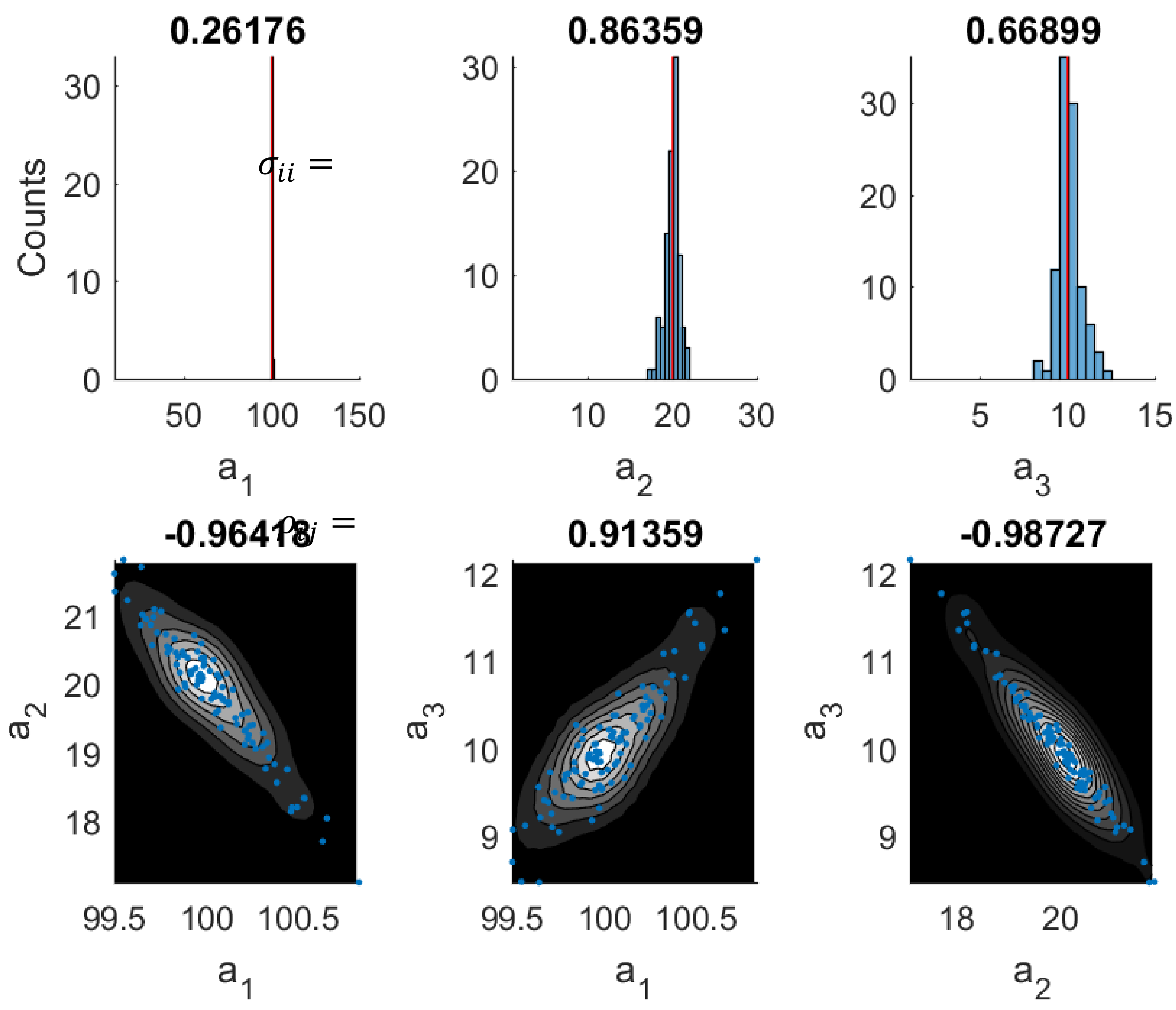
$$a_1 \in [10{,}150], a_2 \in [1{,}30], a_3 \in [1{,}15]$$

*True parameters not centered in search space









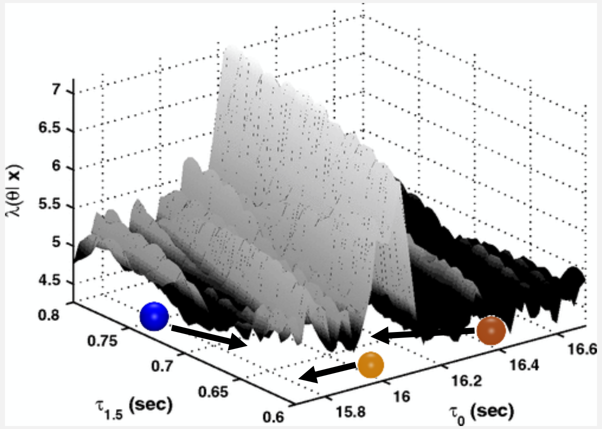
$\sigma_{ii} =$
$\rho_{ij} =$
0.26176
0.86359
0.66899
-0.96418
0.91359
-0.98727
Counts
a1
a2
a3
a1
a1
a2
a2
a3
a3

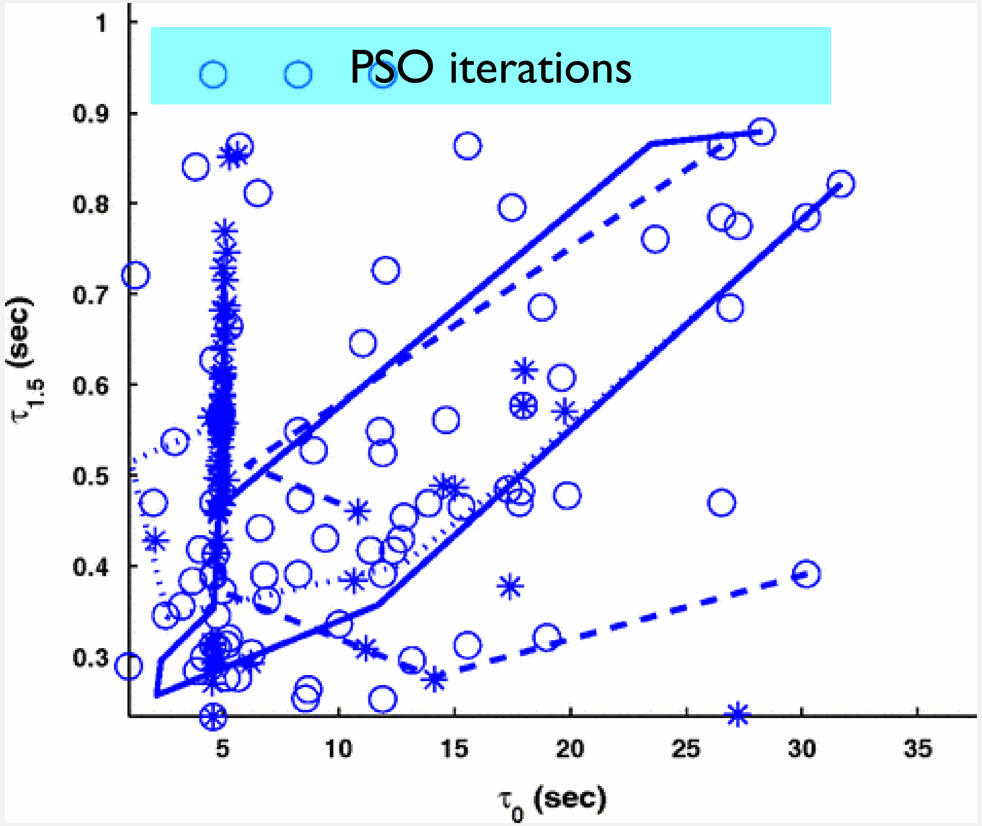
PSO SUCCESS STORIES

Applications in gravitational wave astronomy

NSF NASA TACC TEXAS ADVANCED COMPUTING CENTER

Binary inspiral



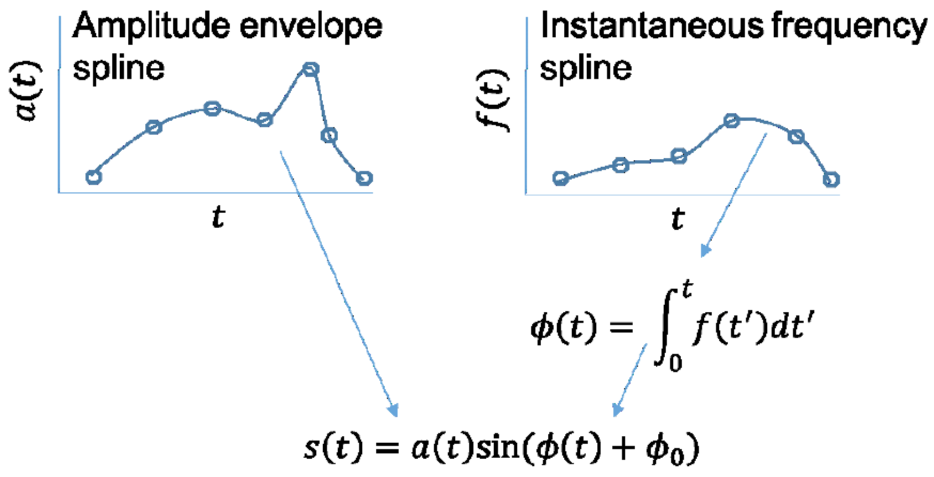


PSO-BASED BINARY INSPIRAL SEARCH

- First use of PSO in GW data analysis:
  - Wang, Mohanty, Physical Review D, 2010
- PSO: factor of $\approx$ 10 fewer evaluations
  - Weerathunga, Mohanty, 2017
- On the threshold of a real-time optimal search:
  - Normandin, Mohanty, Weerathunga, 2018
  - Srivastava, Nayak, Bose, 2018

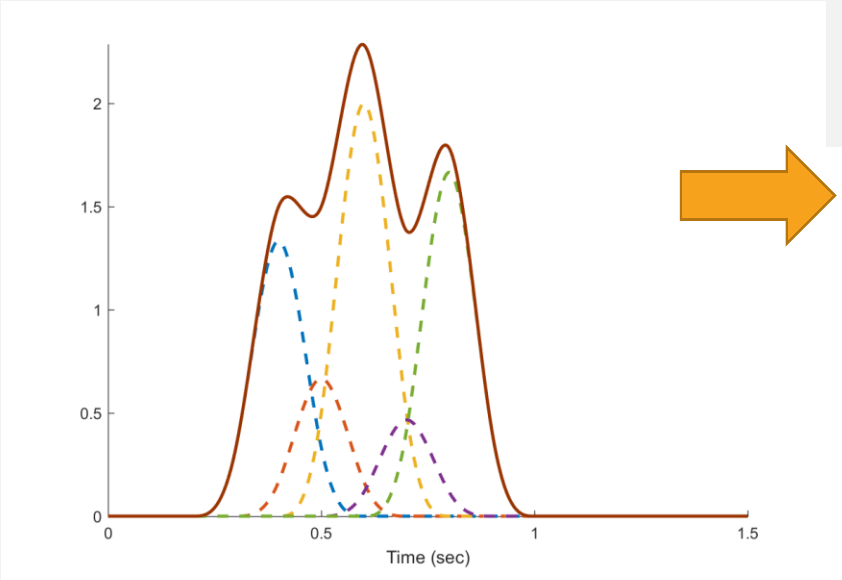
SEARCH FOR UNMODELED CHIRPS

- New approach: model the unknown functions with splines and optimize over their breakpoints
  - Soumya D. Mohanty, Physical Review D (2017).
- SEECR: Spline-Enabled Effective-Chirp Regression



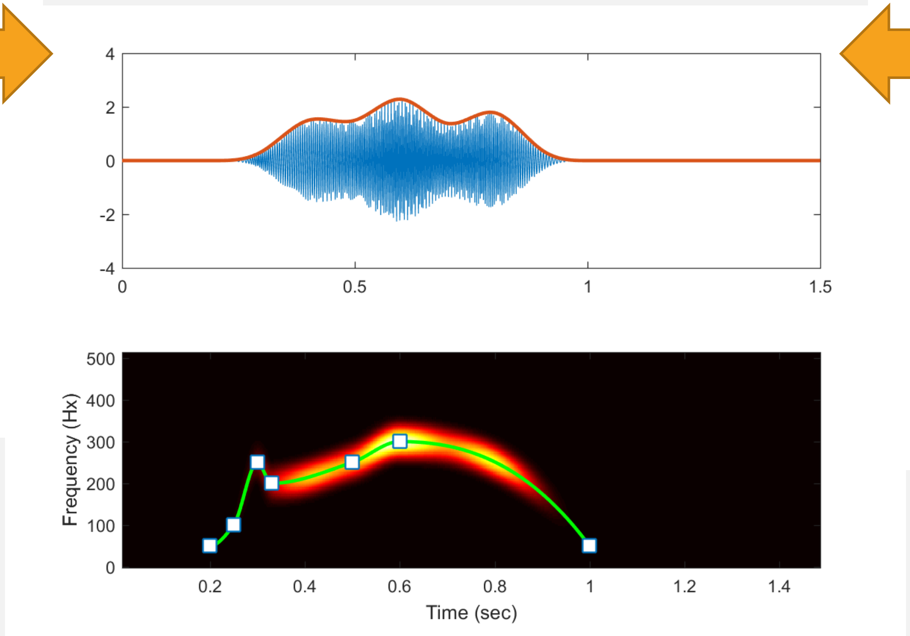
Unmodeled chirps

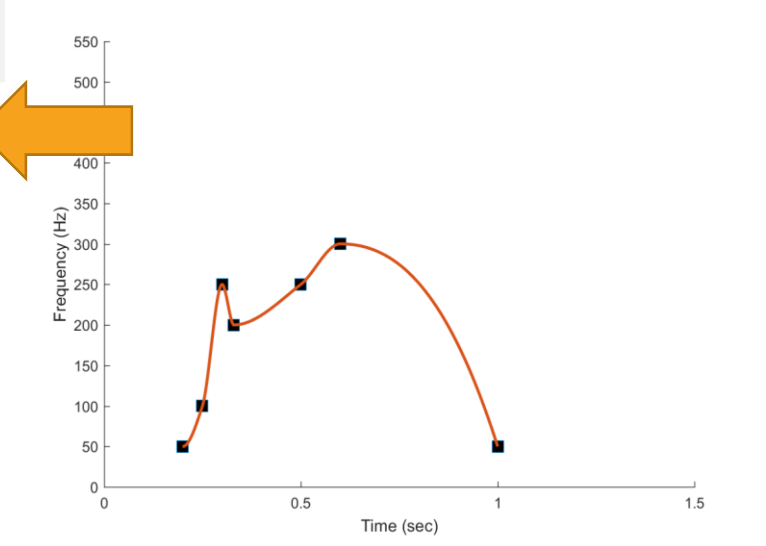
SEECR: SIGNAL MODEL



$$a(t) = \sum_{i=0}^{M-1} \alpha_i B_{i,k}(t)$$

$$s(t) = a(t)\sin(\phi(t) + \phi_0)$$



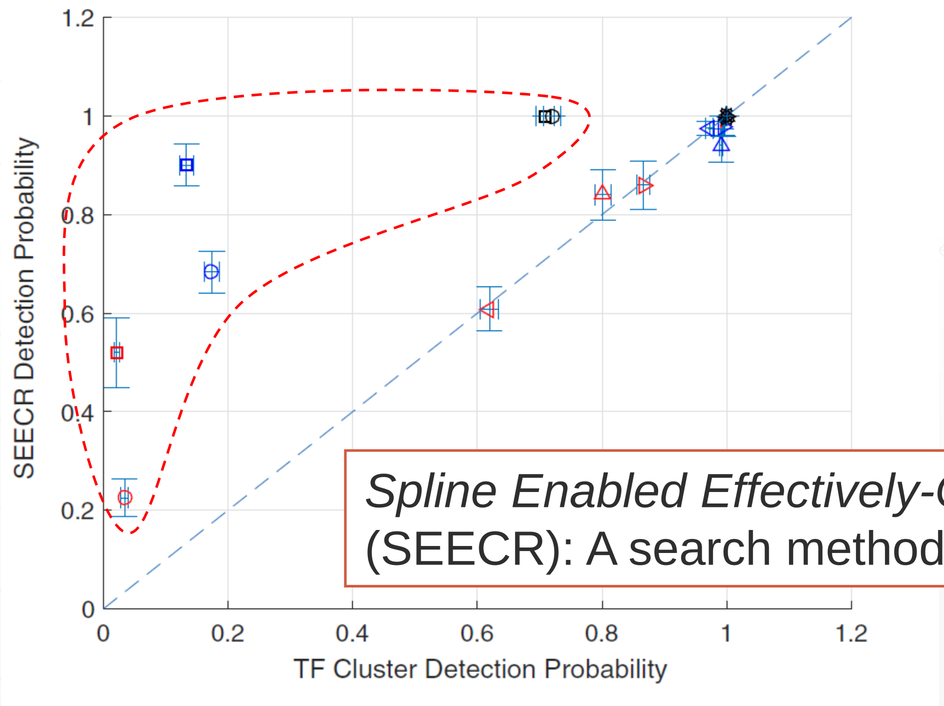


$$\phi(t) = 2\pi \int_0^t f(t')dt'$$

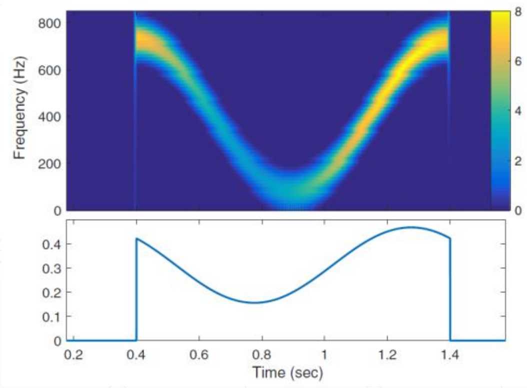
TF Clustering and Chirps

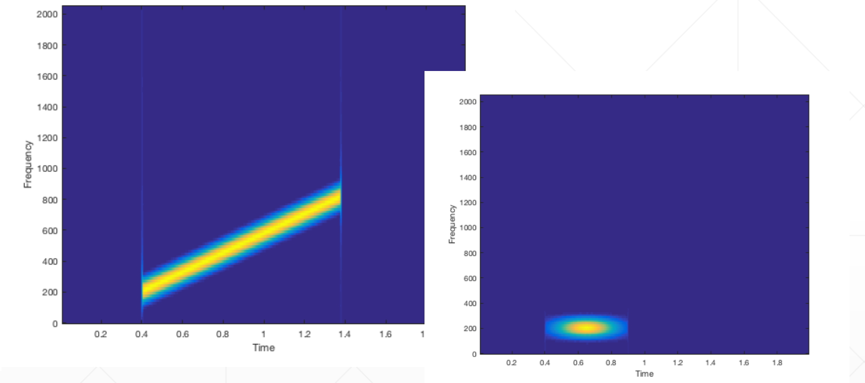
- Chirp signal: distributes signal power along **well-defined** tracks in the TF plane
- TF clustering performs surprisingly poorly on some simple chirp signals
  - Example: Linear chirp and cosine-chirp @SNR= 10, 12, 15

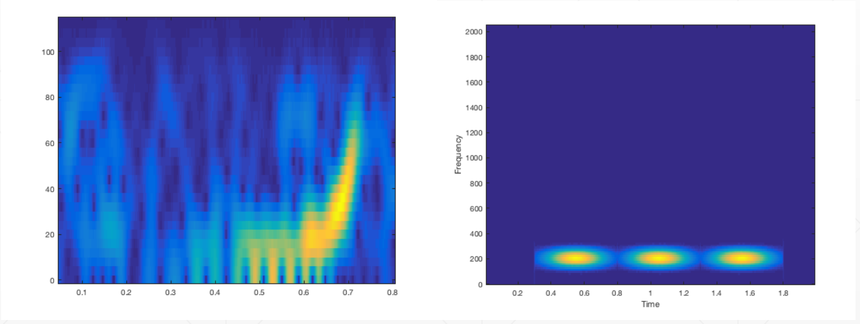
SOUMYA D. MOHANTY PHYSICAL REVIEW D **96,** 102008 (2017)



Spline Enabled Effectively-Chirp Regression
(SEECR): A search method for unmodeled chirps



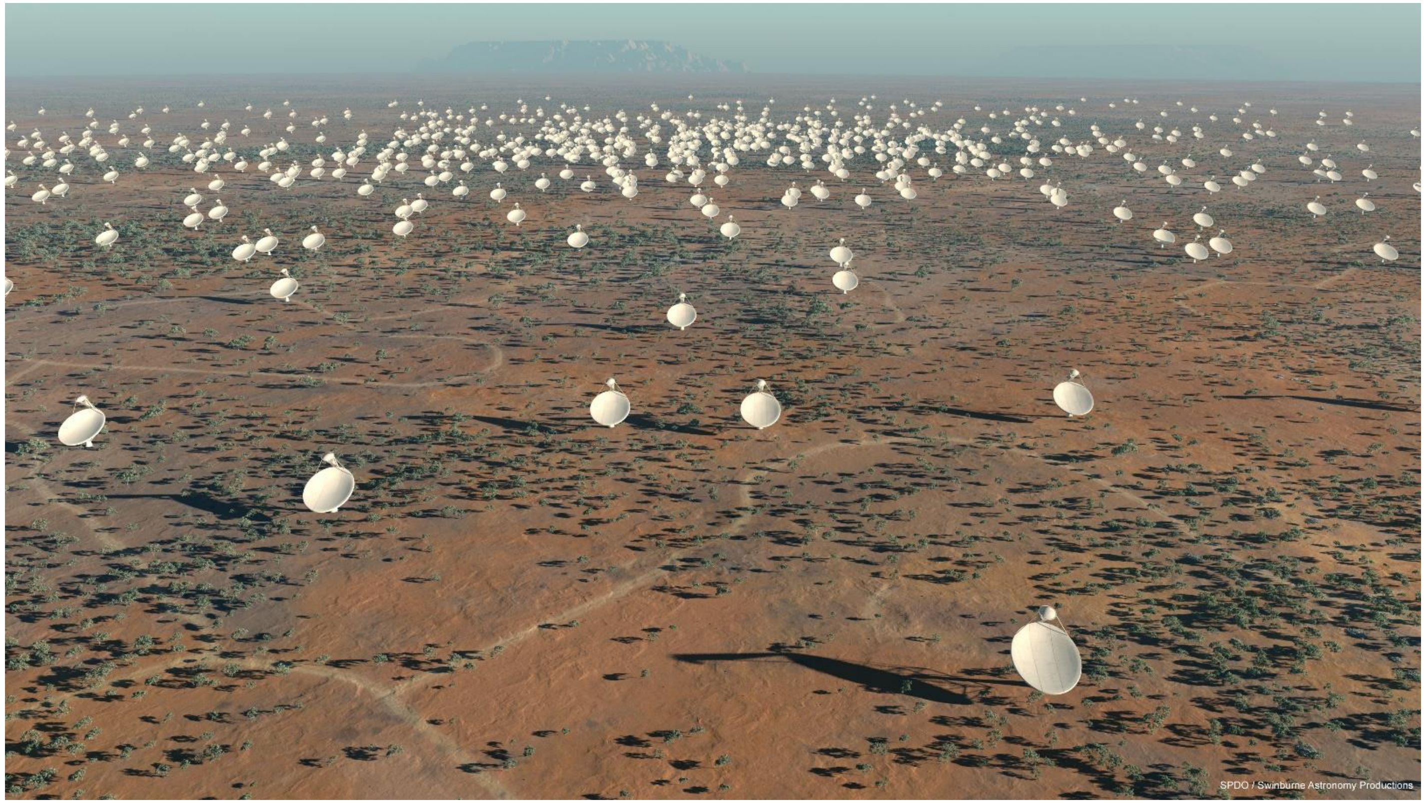




Large-scale PTA

Next Gen Instruments: Square Kilometer Array

- The **Square Kilometre Array** (**SKA**) is a large multi radio telescope project aimed to be built in Australia and South Africa.
- If built, it would have a total collecting area of approximately one square kilometre.
- 50 times more sensitive than any other radio instrument
- Construction of the SKA is scheduled to begin in 2018 for initial observations by 2020, but the construction budget is not secured at this stage.
- The SKA would be built in two phases, with Phase 1 (2018-2023) representing about 10% of the capability of the whole telescope.



Artist's impression of the 5km diameter central core of SKA antennas

SKA era PTA

Current	SKA era PTA
IPTA: About 30 pulsars	Anticipated: 6000 millisecond pulsars R. Smits et al, A & A, (2009)
A few pulsars with timing residual noise level ~ 100 ns	Several hundreds timed to better than 100 ns accuracy
1 pulsar already timed to ~ 80 ns accuracy (Arzoumanian et al, ApJ, 2016)	100 ns is conservative since SKA will have much higher sensitivity

Question: What can a SKA era PTA with 1000 pulsars achieve in terms of GW astronomy?

Answer: A realistic assessment requires overcoming a **data analysis challenge**: “Pulsar phase parameters”

SMBHB GW signal

- Data and signal model for single GW source:

$$\underbrace{\begin{pmatrix} d_1(t) \\ d_2(t) \\ \vdots \\ d_N(t) \end{pmatrix}}_{\text{Timing Residuals from } N \text{ Pulsars}} = \left[\mathbf{1} - \underbrace{\begin{pmatrix} T[\tau_1] & 0 & \cdots & 0 \\ 0 & T[\tau_2] & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & T[\tau_N] \end{pmatrix}}_{\text{Time Delay Ops.}} \right] \overbrace{\underbrace{\begin{pmatrix} F_{+,1} & F_{\times,1} \\ F_{+,2} & F_{\times,2} \\ \vdots & \vdots \\ F_{+,N} & F_{\times,N} \end{pmatrix}}_{\text{Antenna Patterns}}}^{\mathbf{A}} \overbrace{\begin{pmatrix} h_+(t) \\ h_\times(t) \end{pmatrix}}^{\mathbf{h}(t)} + \underbrace{\begin{pmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_3(t) \end{pmatrix}}_{\text{Noise}}$$

- Here τ_i is a time delay that depends on the
  - Earth-pulsar distance: **not known accurately**
  - and Earth-Pulsar-SMBHB lines-of sight geometry
- Unknown signal parameters: Amplitude, sky location, frequency, Observer-Binary orbit geometry
- Time delay ⇒ Phase shift (Pulsar phase parameter)⇒ **Additional unknown parameter** (1 per pulsar)

Signal detection and estimation

Global Minimization over all the signal parameters:

$$MLE\ or\ GLRT \Rightarrow \min_{(\text{parameters})}(.) \rightarrow \min_{(\text{intrinsic})}\left(\min_{(\text{extrinsic})}(.)\right)$$

Carrying out the inner maximization analytically/semi-analytically reduces the computational cost

Intrinsic parameters: Pulsar phases, sky location ("F-statistic" approach)	
Markov Chain Monte Carlo (MCMC)	Particle Swarm Optimization (PSO)
• Corbin & Cornish 2010 • Taylor et al, 2014 • Others…	• Wang, Mohanty, Jenet, ApJ, 2014 • Zhu et al, MNRAS, 2016

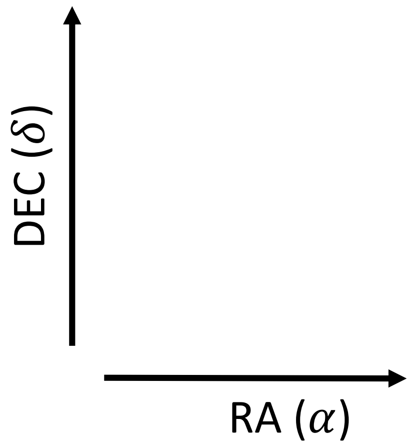
MaxPhase Algorithm

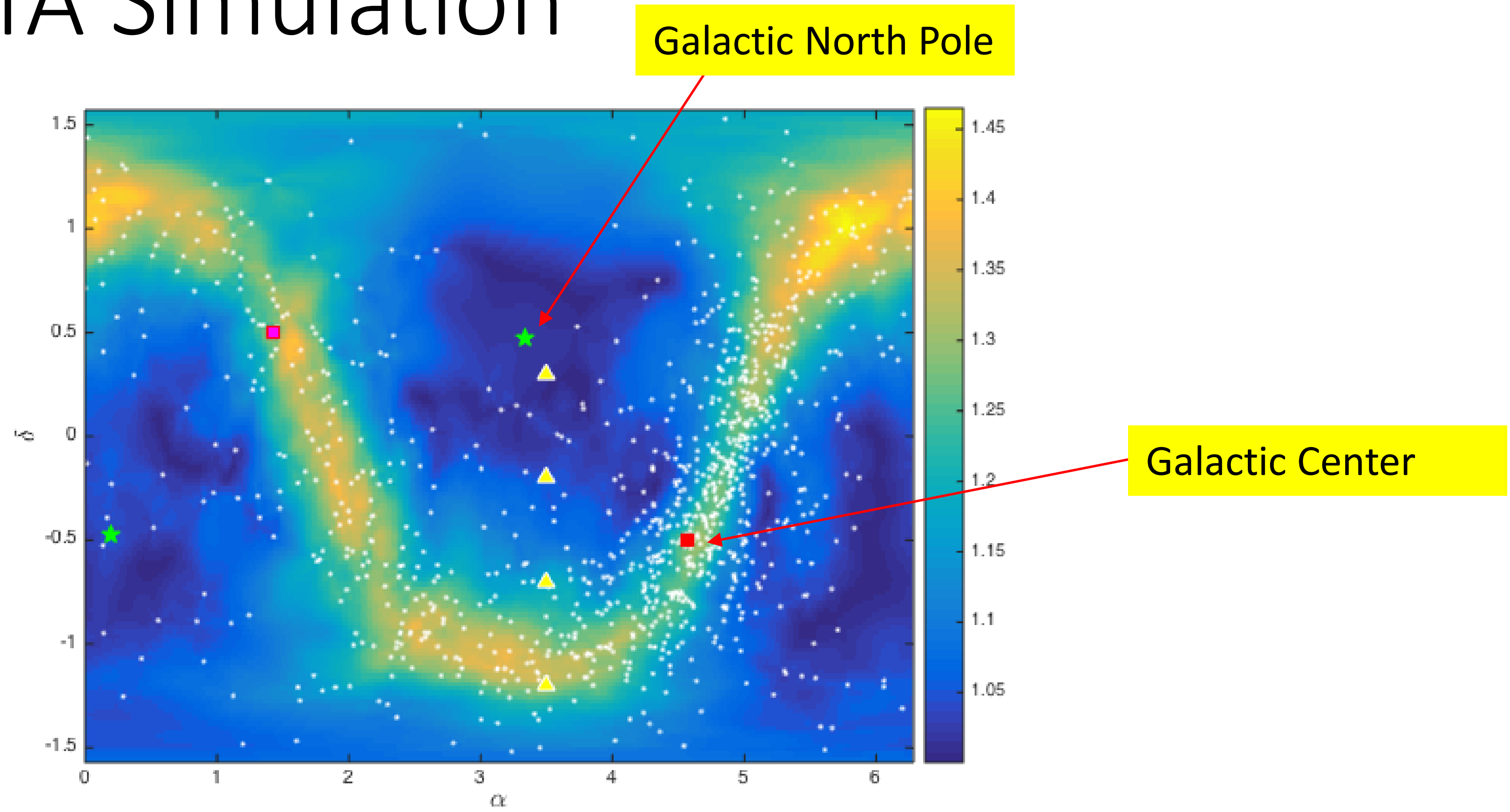
A SKA era PTA will contain **hundreds** of pulsar phase parameters!
Analysis algorithm must be scalable if all the pulsars are included in the analysis.

Minimize: $f(x, \phi_1, \phi_2) = f_1^2(x, \phi_1) + f_2^2(x, \phi_2)$; 10 grid points along each parameter		
ϕ_1, ϕ_2: intrinsic	Solution $\hat{x}$ for inner minimization depends on $\phi_1, \phi_2 \Rightarrow$ $\mathrm{Minimize:}\ f_1^2(\hat{x}(\phi_1,\phi_2),\phi_1) + f_2^2(\hat{x}(\phi_1,\phi_2),\phi_2) = g(\phi_1,\phi_2)$	Cost: $10 \times 10 \times 10$ Not scalable
x: intrinsic	Inner minimization: $\min_{\phi_1} f_1^2(x,\phi_1) + \min_{\phi_2} f_2^2(x,\phi_2)$ Minimize: $f_1^2\left(x, \hat{\phi}_1(x)\right) + f_2^2\left(x, \hat{\phi}_2(x)\right) = k(x)$	Cost: $(10 + 10) \times 10$ Scalable

- MaxPhase: Choose pulsar phases as extrinsic parameters
  - Semi-analytical minimization by solving quartic equations
  - Number of intrinsic parameters is fixed at 7 irrespective of the number of pulsars
  - PSO used for the outer minimization

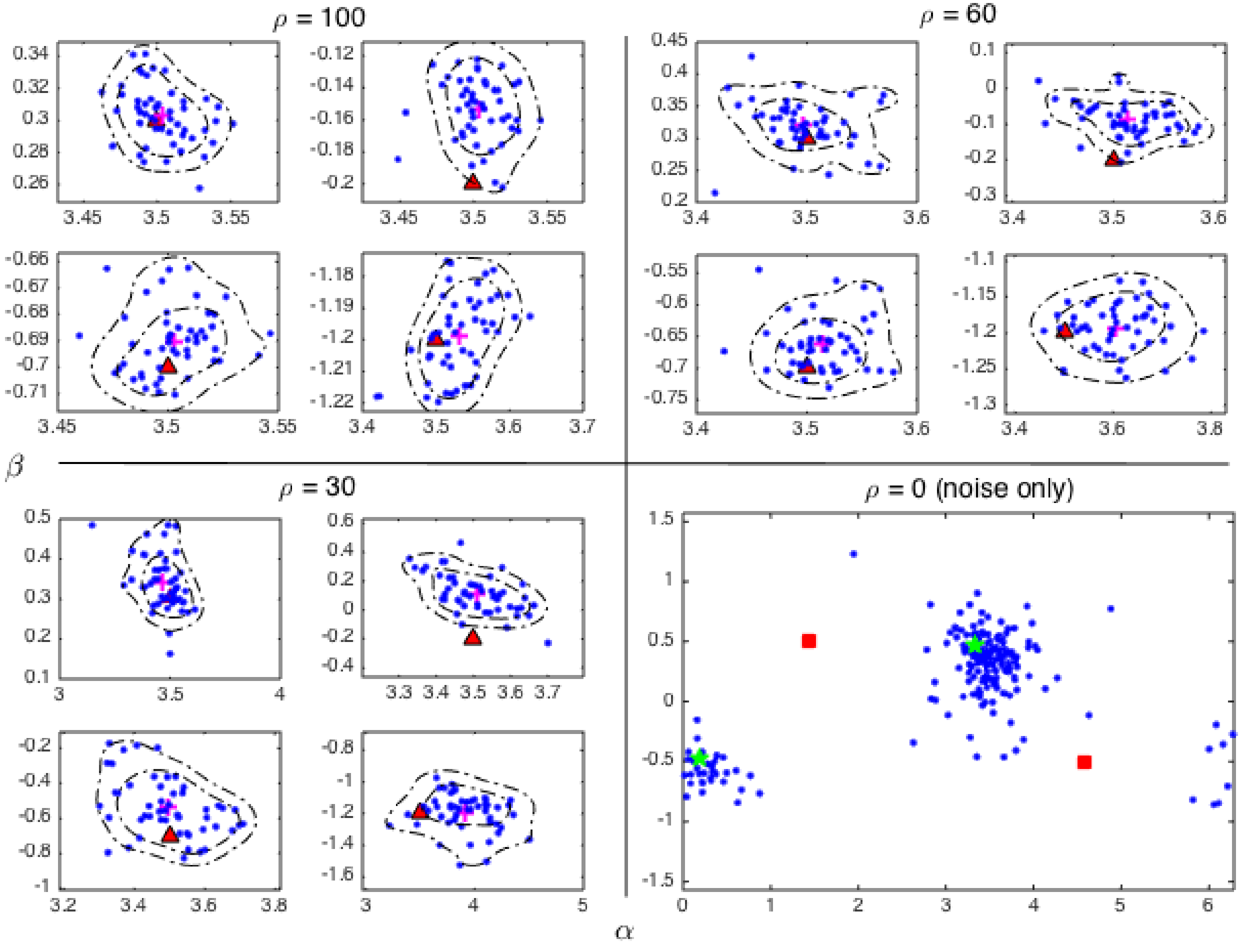
SKA era PTA Simulation





- 1000 pulsars; timing residual noise rms = 100 ns (White, Gaussian)
- 4 SMBHB locations
- Observation period: 5 years; Cadence: one sample per two weeks;
- Wang, Mohanty, Physical Review Letters, 2017

Direction Estimation



- Good condition number spots attract noise only estimates
- Location B and D show significant bias (towards nearest good condition number spot)

- Conservative error area: $2\sigma_\alpha \times 2\sigma_\delta \times \cos\delta$
- Localization to within $\sim$70 to $\sim$180 deg^2 at $\rho = 30$.
- Search for PSO J334 (Liu et al, ApJL, 2015): 80 deg^2 field from Pan-STARRS1 Medium survey
- Optical counterpart searches possible for even the most distant sources (SKA + LSST)