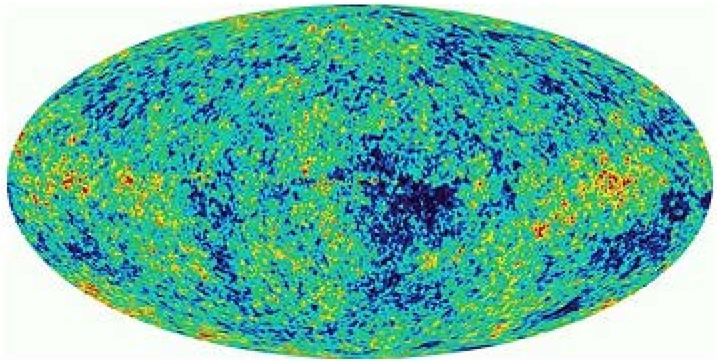
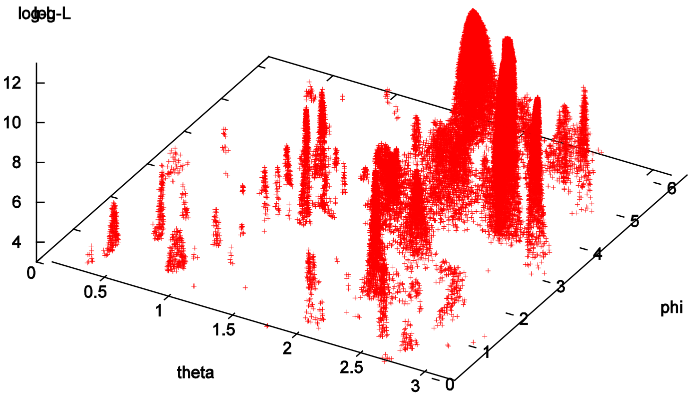
Probabilistic Source Detection



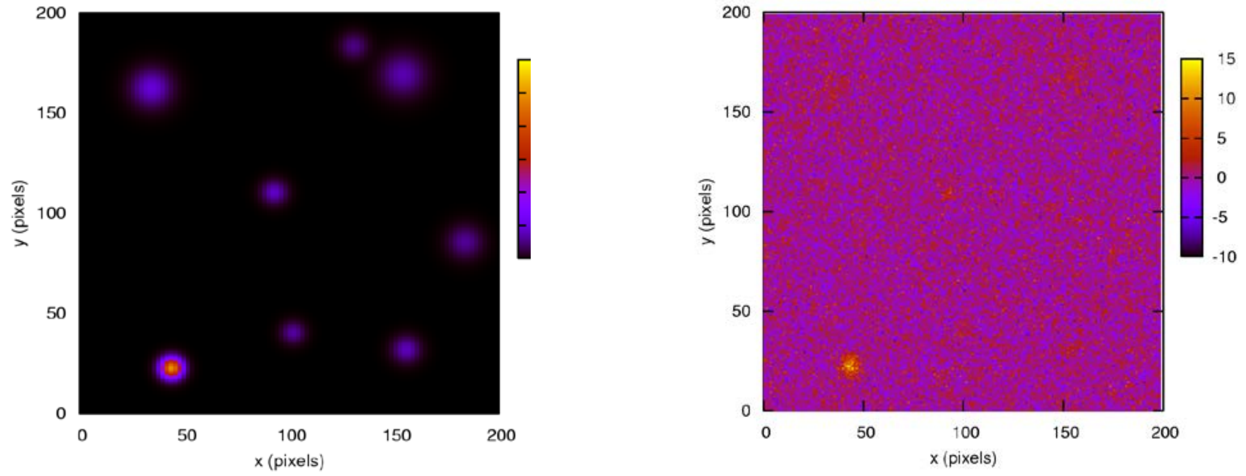


Farhan Feroz
f.feroz@mrao.cam.ac.uk

Cavendish Astrophysics

UNIVERSITY OF CAMBRIDGE

Source/Object Detection: Problem Definition

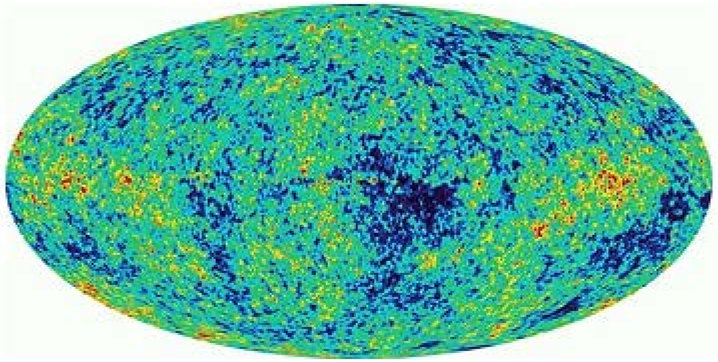


- Aim: Detect and characterize discrete sources in a background
- Each source described by a template $f(p)$ with parameters p.
- With source and noise contributions being additive and N_s source

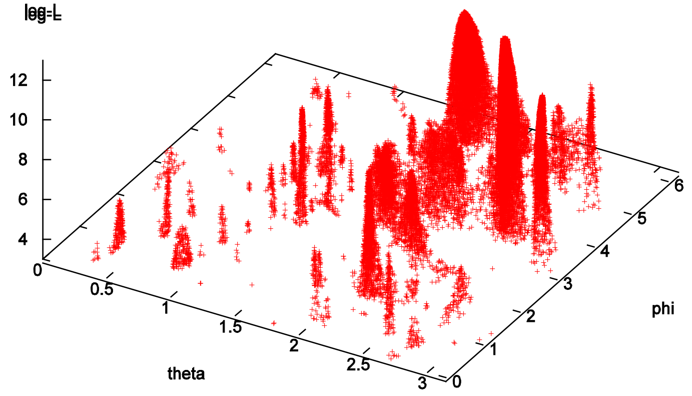
$$d = b(q) + n(r) + \sum_{k=1}^{N_s} f(p_k)$$

- Inference goal: Use data d to constrain source parameters N_s, p_k ($k = 1, 2, \ldots, N_s$). Margnialize over background and noise parameters (q and r).

Probabilistic Source/Object Detection



- Problems in Object Detection
  - Identification
  - Quantifying Detection
  - Characterization



Textures in CMB

Bayesian Parameter Estimation

- Collect a set of N data points D_i (i = 1, 2, …, N), denoted collectively as data vector **D**.

- Propose some model (or hypothesis) H for the data, depending on a set of M parameter θ_i (i = 1, 2, …, N), denoted collectively as parameter vector **θ**.

- Bayes' Theorem:

Likelihood

Prior

$$\mathrm{P}(\boldsymbol{\theta} \mid \mathbf{D}, H) = \frac{\mathrm{P}(\mathbf{D} \mid \boldsymbol{\theta}, H)\,\mathrm{P}(\boldsymbol{\theta} \mid H)}{\mathrm{P}(\mathbf{D} \mid \mathrm{H})} \longrightarrow \mathrm{P}(\boldsymbol{\theta}) = \frac{\mathrm{L}(\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\mathrm{Z}}$$

Posterior

Evidence

- Parameter Estimation: $\mathrm{P}(\boldsymbol{\theta}) \propto \mathrm{L}(\boldsymbol{\theta})\pi(\boldsymbol{\theta})$

  posterior α likelihood x prior

Bayesian Model Selection

$$P(\boldsymbol{\theta} \mid \mathbf{D}, H) = \frac{P(\mathbf{D} \mid \boldsymbol{\theta}, H) P(\boldsymbol{\theta} \mid H)}{P(\mathbf{D} \mid \mathrm{H})} \rightarrow P(H \mid \mathbf{D}) = \frac{P(\mathbf{D} \mid H) P(H)}{P(\mathbf{D})}$$

- Consider two models H_0 and H_1

$$R = \frac{P(H_1 \mid \mathbf{D})}{P(H_0 \mid \mathbf{D})} = \frac{P(\mathbf{D} \mid H_1) P(H_1)}{P(\mathbf{D} \mid H_0) P(H_0)} = \frac{Z_1}{Z_0} \frac{P(H_1)}{P(H_0)}$$

Bayes Factor

Prior Odds

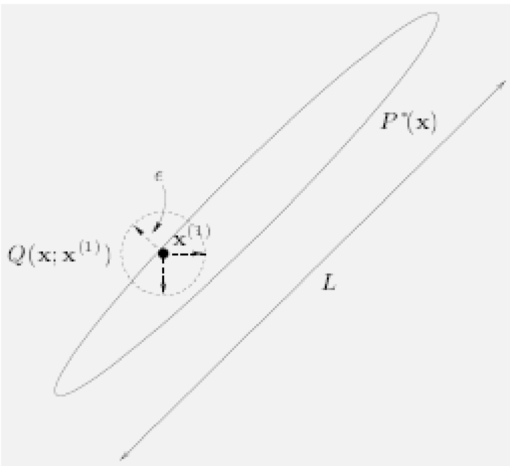
- Bayesian Evidence $Z = P(\mathbf{D}|H) = \int L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ plays the central role in Bayesian Model Selection.

- Bayesian Evidence rewards model preditiveness.
  - Sets more stringent conditions for the inclusion of new parameters

Computation of Bayesian Evidence

- Evidence = $Z = \int L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$
- Evaluations of the *n*-dimensional integral presents great numerical challenge
- If dimension n of parameter space is small, calculate unnormalized $\bar{P}(\boldsymbol{\theta}) = L(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ over grid in parameter space → get evidence trivially
- For higher-dimensional problems, this approach rapidly becomes impossible
  - Need to find alternative methods
  - Gaussian approximation, Savage-Dickey ratio (see Trotta, 2007, MNRAS, 378, 72)
- Evidence evaluation at least an order of magnitude more costly than parameter estimation.

Metropolis Hastings Algorithm



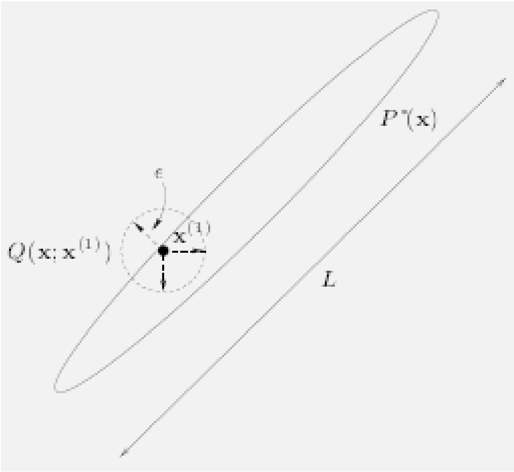
- Metropolis-Hastings algorithm to sample from $P(\theta)$
  - Start at an arbitrary point θ_0
  - At each step, draw a trial point, θ', from the proposal distribution $Q(\theta' \mid \theta_0)$
  - Calculate ratio $r = P(\theta')\, Q(\theta_n \mid \theta') \,/\, P(\theta_n)\, Q(\theta' \mid \theta_n)$
  - accept $\theta_{n+1} = \theta'$ with probability max(1, r) else set $\theta_{n+1} = \theta_n$

- After initial burn-in period, any (positive) proposal $Q \longrightarrow$ convergence to $P(\theta)$
- Common choice of Q, multivariate Gaussian centred on θ_n but many others
- Inferences wrt posterior can be obtained easily from converged chains

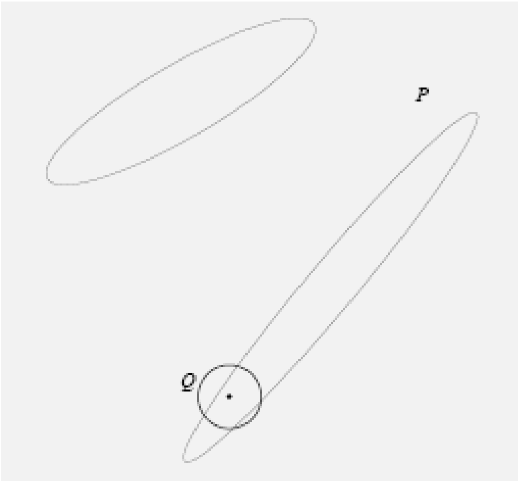
$$\langle \theta \rangle = \int \theta P(\theta) d\theta \approx \frac{1}{N} \sum_i \theta_i$$

$$\langle f(\theta) \rangle = \int f(\theta) P(\theta) d\theta \approx \frac{1}{N} \sum_i f(\theta_i)$$

Metropolis Hastings Algorithm – Some Problems

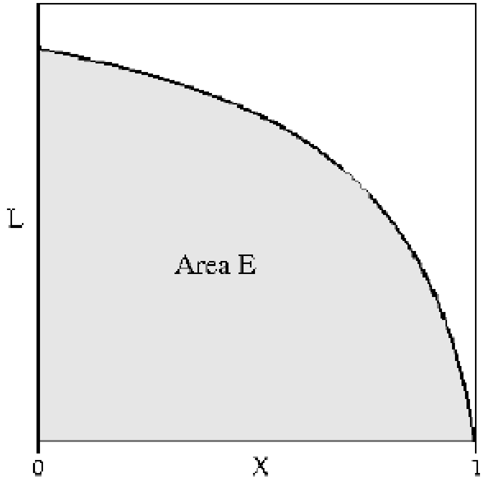


- Choice of proposal Q strongly affects convergence rate and sampling efficiency
  - large proposal width $\varepsilon \rightarrow$ trial points rarely accepted
  - small proposal width $\varepsilon \rightarrow$ chain explores $P(\theta)$ by a random walk $\rightarrow$ very slow
- If largest scale of $P(\theta)$ is L, typical diffusion time $t \sim (L/\varepsilon)^2$
- If smallest scale of $P(\theta)$ is l, need $\varepsilon \sim l$, diffusion time $t \sim (L/l)^2$



- Particularly bad for multi-modal distributions
  - Transitions between distant modes very rare
  - No one choice of proposal width ε works
  - Standard convergence tests will suggest convergence, but actually only true in a subset of modes

Nested Sampling



- Introduced by John Skilling
  (AIP Conference Proceedings, Volume 735, pp. 395-405, 2004).

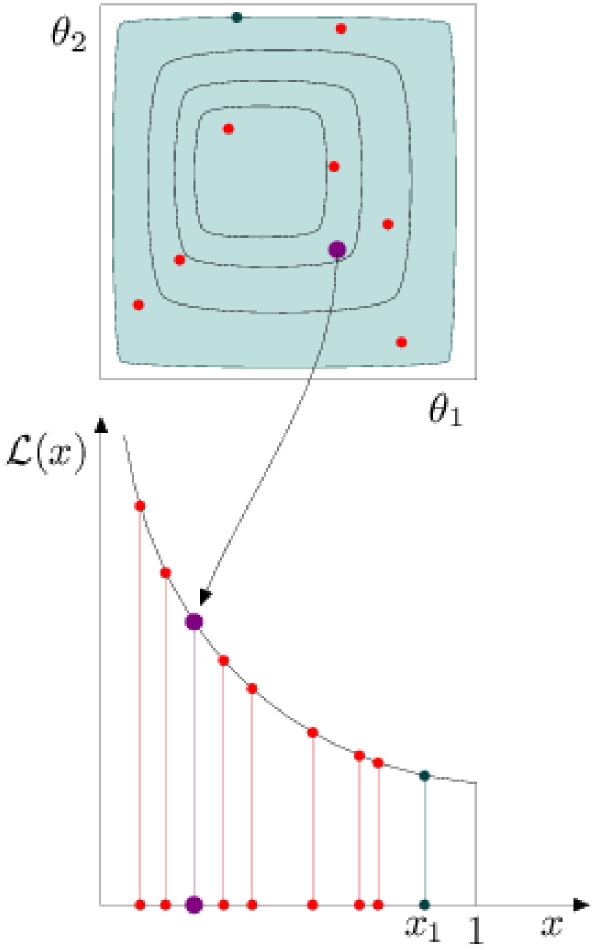
- Monte Carlo technique for efficient evaluation of the Bayesian Evidence.

- Re-parameterize the integral with the prior mass X

$$X(\lambda) = \int_{L(\theta)>\lambda} \pi(\theta) d^n\theta$$

$$Z = \int L(\theta)\pi(\theta) d^n\theta = \int_0^1 L(X) dX$$

Nested Sampling: Algorithm



1. Sample N 'live' points uniformly inside the initial prior space ($X_0 = 1$) and calculate their likelihoods

2. Find the point with the lowest L_i and remove it from the list of 'live' points

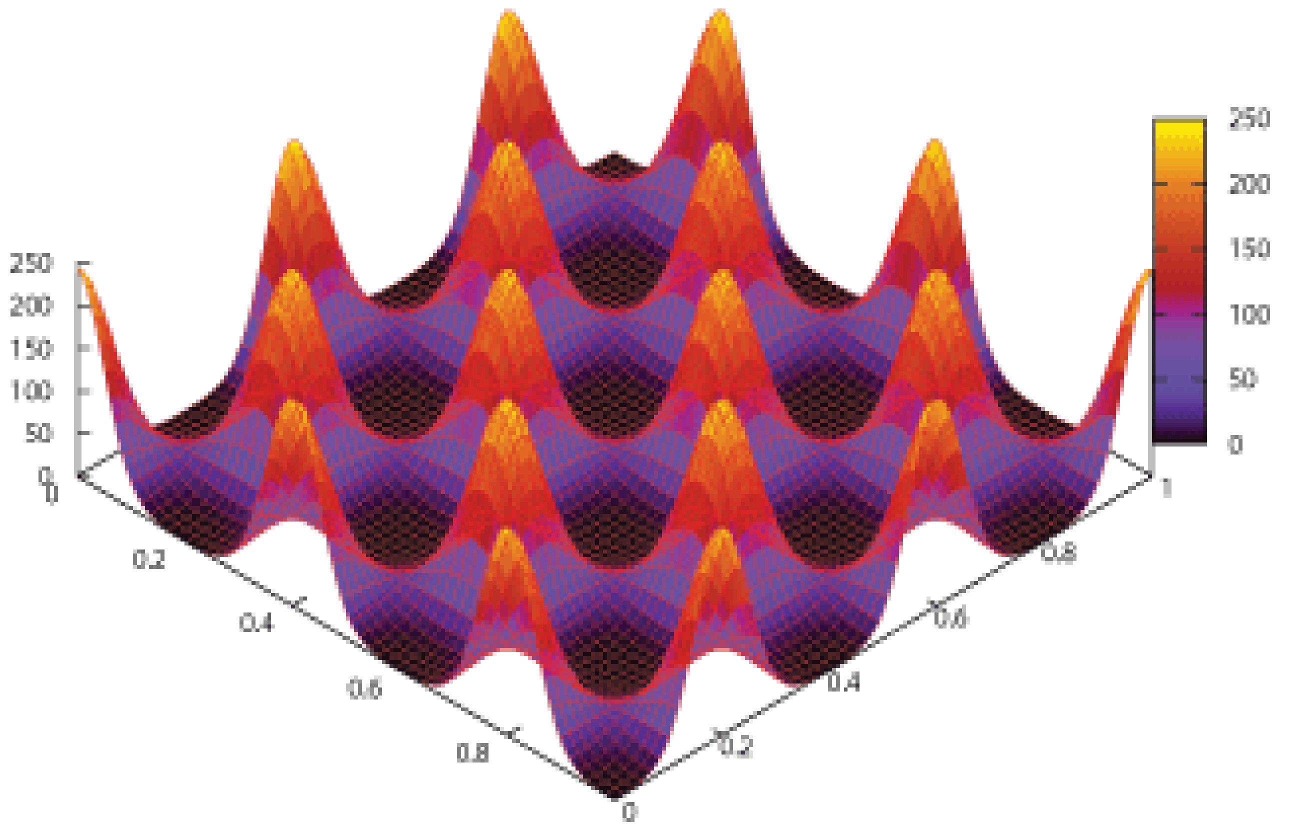
3. Increment the evidence as $Z = Z + L_i\,(X_{i-1} - X_{i+1})/2$

4. Reduce the prior volume $X_i / X_{i-1} = t_i$ where $P(t) = N\,t^{N-1}$

5. Replace the rejected point with a new point sampled from $\pi(\theta)$ with constraint $L > L_i$

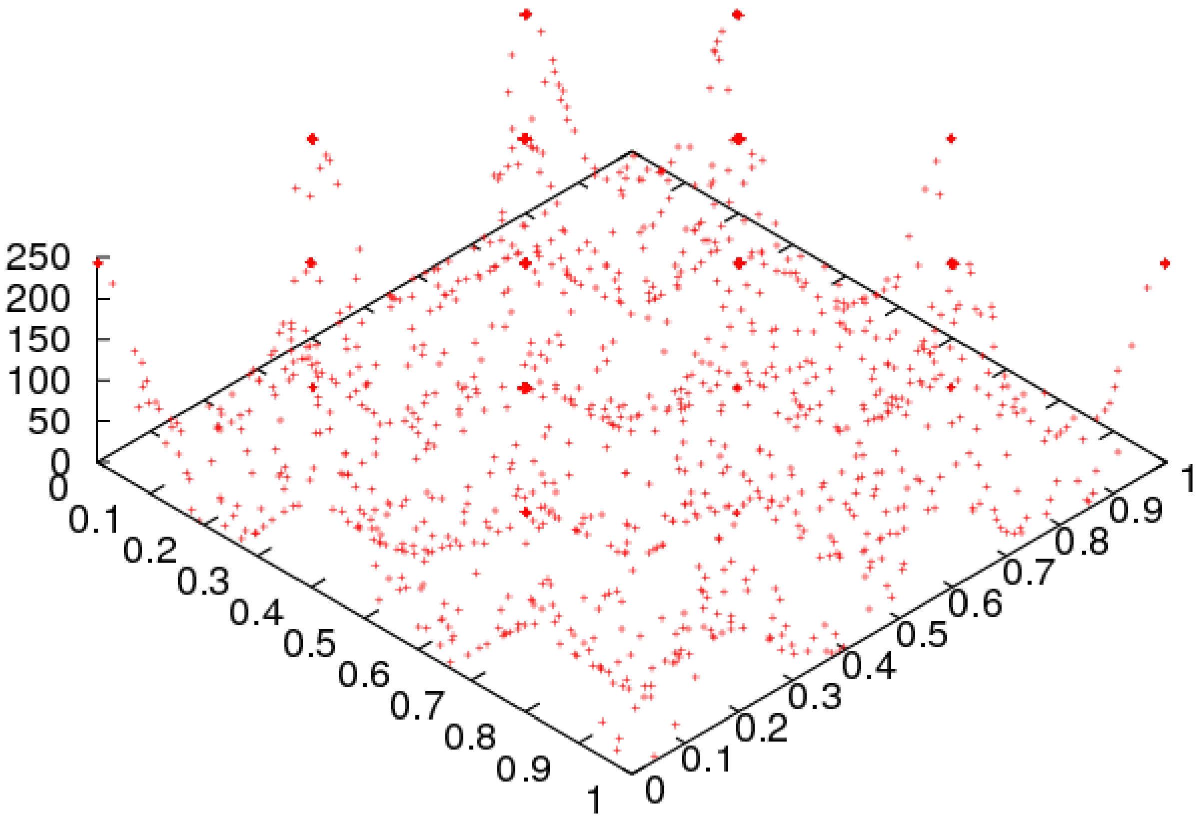
6. If $L_{\max} X_i < \alpha Z$ then stop else goto 3

Nested Sampling: Demonstration



Egg-Box Posterior

Nested Sampling: Demonstration

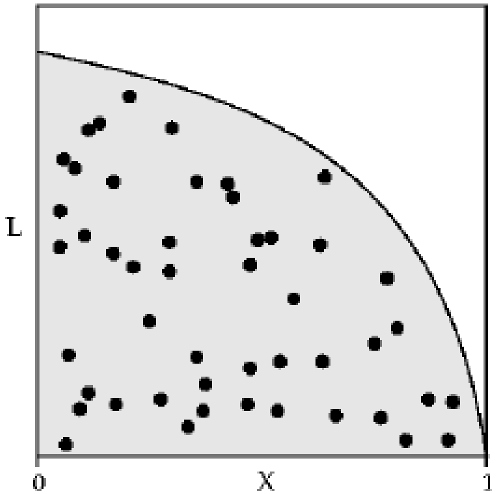


Egg-Box Posterior

Nested Sampling

- Advantages:
  - Typically requires around 100 times fewer samples than standard MCMC methods
  - Proceeds exponentially towards high likelihood regions
    - Prior volume shrinks by $\exp(-1/N_{\mathrm{live}})$ at each iteration
  - Parallelization easily done

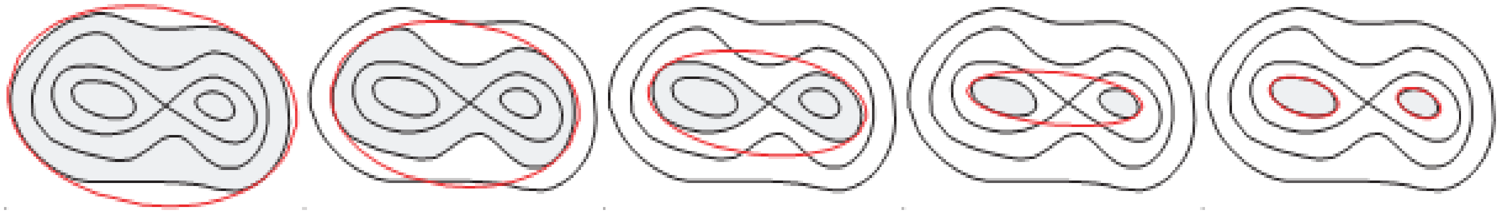
- Bonus: posterior samples easily obtained as by-product. Take full sequence of rejected points, θ_i & weigh i^{th} sample by $p_i = L_i w_i / Z$



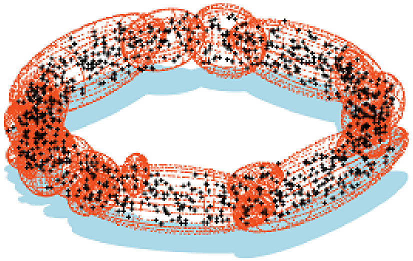
- Problem: must sample efficiently from prior within complicated, hard-edged likelihood constraint.
  - Possible solutions:
    - MCMC (can be inefficient)
    - ellipsoidal rejection sampling (MultiNest)
    - Galilean Monte Carlo (GMC)

Multi-modal Nested Sampling (MultiNest)

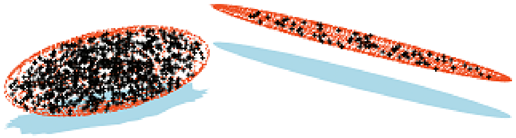
- Introduced by Feroz & Hobson (2008, MNRAS, 384, 449, arXiv:0704.3704), refined by Feroz, Hobson & Bridges (2009, MNRAS, 398, 1601, arXiv:0809.3437)



Ellipsoidal Rejection Sampling



Uni-Modal Distribution



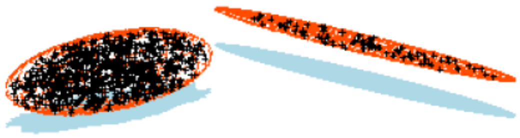
Multi-Modal Distribution

Optimal Ellipsoidal Decomposition

Feroz, Hobson & Bridges, 2009, MNRAS, 398, 1601, arXiv:0809.3437

Optimize $F(S) \equiv \frac{1}{V(S)} \sum_k V(E_k)$, subject to $F(S) \geq 1$

S = collection of live points, $V(S)$ = prior (target) volume, E_k = k^{th} ellipsoid



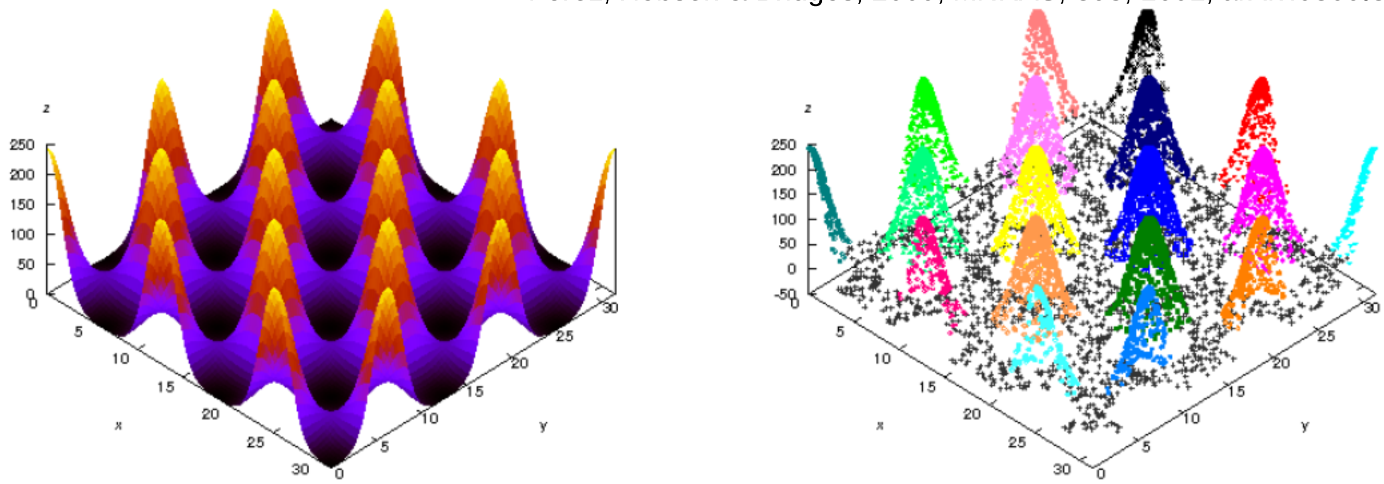
1000 points drawn from two ellipsoids

1000 points drawn from a torus

1. For S, calculate bounding ellipsoid E and $V(E)$
2. Enlarge E so that $V(E) = \max[V(E), V(S)]$
3. Partition S into S_1 and S_2 containing n_1 and n_2 points using $k-$means with $K = 2$
4. Calculate E_1, E_2 and volumes $V(E_1)$, $V(E_2)$
5. Enlarge E_k $(k = 1, 2)$ so that $V(E_k) = \max[V(E_k), V(S_k)]$.
6. For all $\boldsymbol{u} \in S$, assign $\boldsymbol{u}$ to S_k such that $h_k(\boldsymbol{u}) = \min[h_1(x), h_2(x)]$
7. If no point reassigned goto 8; else goto 4
8. If $V(E_1) + V(E_2) < V(E)$ or $V(E) > 2V(S)$
   – partition S into S_1 and S_2
   – repeat entire algortihm for each subset S_1 and S_2
   else
   – return E as the optimal ellipsoid of the point set S

Identification of Posterior Modes

Feroz, Hobson & Bridges, 2009, MNRAS, 398, 1601, arXiv:0809.3437



- For multi-modal posteriors, useful to identify which samples ‘belong’ to which mode

- For well-defined ‘isolated’ modes:
  - can make reasonable estimate of posterior mass each contains (‘local’ evidence)
  - can construct posterior parameter constraints associated with each mode

- Once NS process reached likelihood such that ‘footprint’ of mode well-defined ⟶ identify at each subsequent iteration the points in active set belonging to mode

- Partitioning and ellipsoids construction algorithm described above provides efficient and reliable method for performing this identification

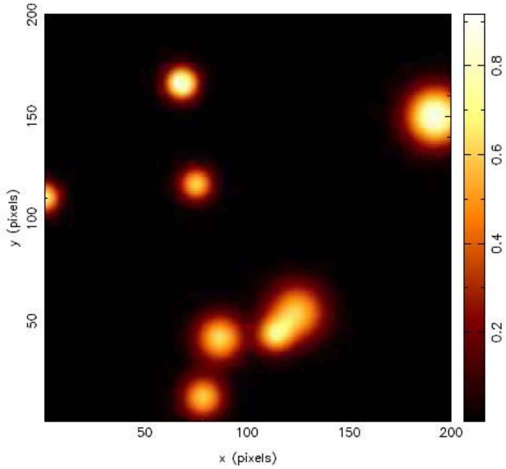
Bayesian Object Detection: Variable Source Number Model

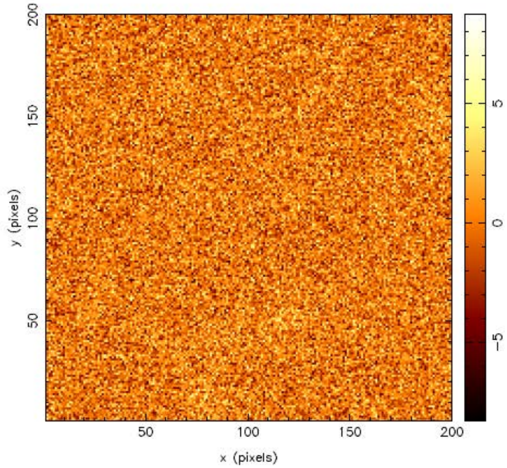
- Bayesian Purist Gold Standard: detect and characterize all sources in the data simultaneously ⇒ infer full parameter set $\boldsymbol{\theta} = \{N_s, p_1, p_2, \ldots, p_{N_s}, q, r\}$
- Allows straight-forward inclusion of prior information on number of sources, N_s.
- Complication
  - Length of parameter vector, $\boldsymbol{\theta}$, is variable
  - Requires reversible-jump MCMC (see Green, 1995, *Biometrika*, V. 82)
  - Counting degeneracy when assigning source parameters in each sample to sources in image ⇒ at least $N_s!$ modes
- Practical Concern: If prior on N_s remains non-zero at large N_s
  - Parameter space to be explored becomes very large
  - Slow mixing, can be very inefficient

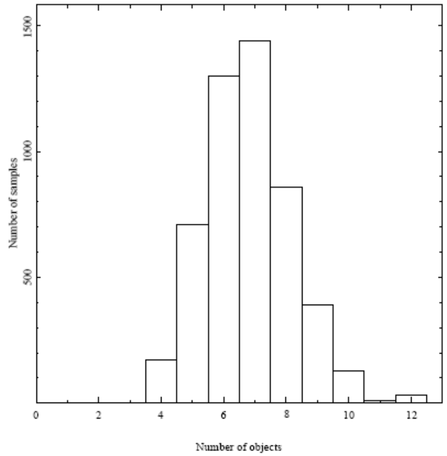
Bayesian Object Detection: Variable Source Number Model

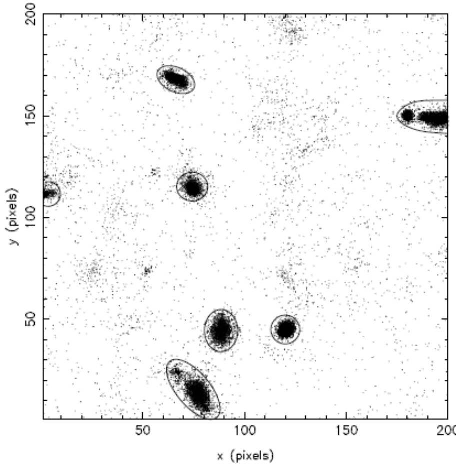
Hobson & McLachlan, 2002, astro-ph/0204457

- 8 Gaussian sources, with variable scale and amplitude, in Gaussian noise
- Analysis done with BayeSys (http://www.inference.phy.cam.ac.uk/bayesys/)
  - Runtime: 17 hours CPU time









Bayesian Object Detection: Fixed Source Number Model

- Poor man's approach to Bayesian gold standard
- Consider series of models H_{N_s}, each with fixed N_s, where N_s goes from say 0 to N_{max}
  - Length of parameter space is fixed for each model
  - Can use standard MCMC or nested sampling
- Determine preferred number of source using Bayesian model selection
- See e.g. Feroz, Hobson & Balan, 2011, arXiv:1105.1150
  - Detection of second companion orbiting HIP 5158

N_p	$\ln \mathcal{Z}$	s (m/s)
1	30.05 ± 0.14	10.31 ± 1.09
2	78.19 ± 0.15	2.28 ± 0.31
3	70.68 ± 0.16	2.28 ± 0.28

Table 2. The evidence and jitter values for the system HIP 5158. The null evidence (-255.3) has been subtracted from each $\ln \mathcal{Z}$ value.

Bayesian Object Detection: Single Source Model

- Special case of fixed source number model, simply set $N_s = 1$
- Not restricted to detecting just one source in the data
  - Trade-off high dimensionality with multi-modality
  - Posterior will have numerous modes
  - Each corresponding to a either real or spurious source
- Fast and reliable method when sources (effects) are non-overlapping
- Use local evidences for distinguishing between real and spurious sources

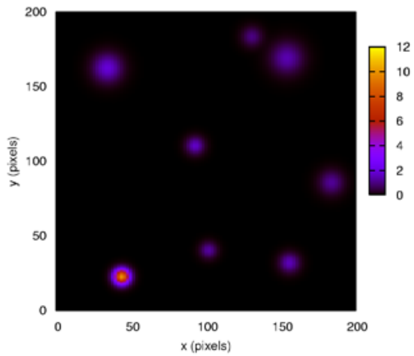
Quantifying Object Detection: Single Source Model

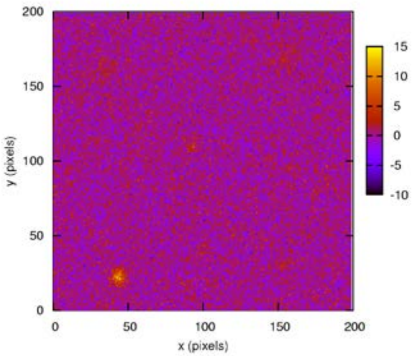
- $$p_{TP} = \frac{R}{1+R}$$
- $$R = \frac{\Pr(H_1 \mid \mathrm{D})}{\Pr(H_0 \mid \mathrm{D})} = \frac{\Pr(\mathrm{D} \mid H_1)\Pr(H_1)}{\Pr(\mathrm{D} \mid H_0)\Pr(H_0)} = \frac{Z_1 \Pr(H_1)}{Z_0 \Pr(H_0)}$$
- H_0 = "there is no object with its centre lying in the region S"
- H_1 = "there is one object with its centre lying in the region S"
- $$Z_0 = \frac{1}{|S|}\int_s L_0 dX = L_0$$
- For objects distributed according to Poisson distribution

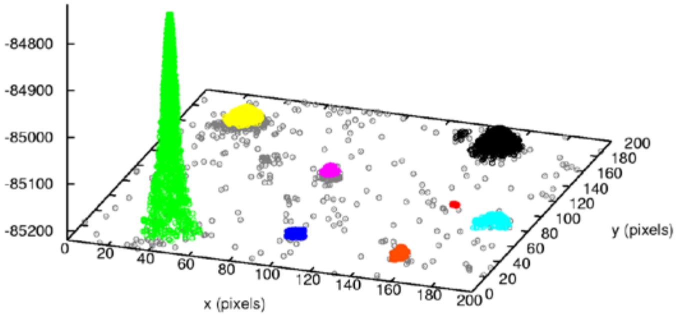
$$\frac{\Pr(H_1)}{\Pr(H_0)} = \mu_s, \quad \therefore R = \frac{Z_1 \mu_s}{L_0}$$

How Many Sources? Bayesian Solution

Feroz & Hobson, 2008, MNRAS, 384, 449, arXiv:0704.3704







MultiNest

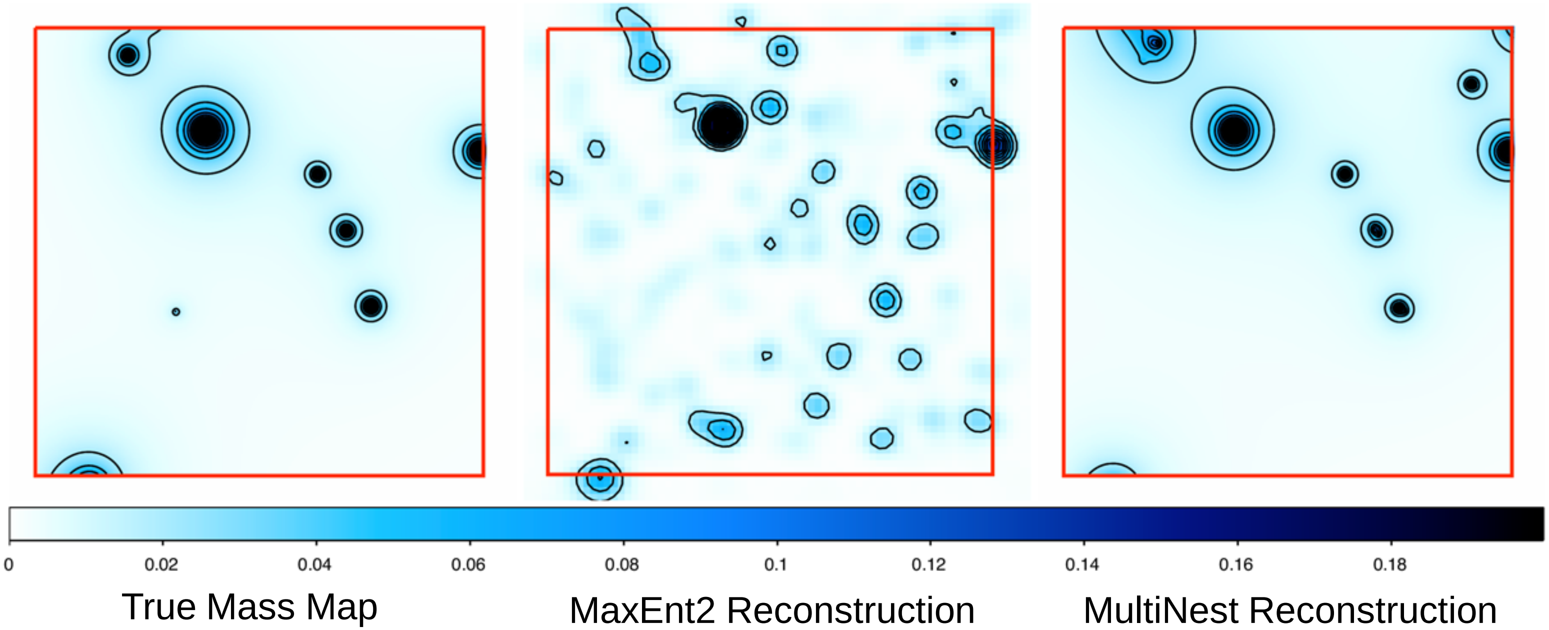
- 7 out of 8 objects identified
  - missed 1 object because 2 objects are very close
- *runtime* = 2 *min* on a normal desktop

Thermodynamic Integration

- Solution possible only through iterative sampling (see McLachlan & Hobson, 2002)
- *runtime* > 16 *hours* on a normal desktop

Applications: Clusters in Weak Lensing

Feroz, Marshall & Hobson, 2008, arXiv:0810.0781

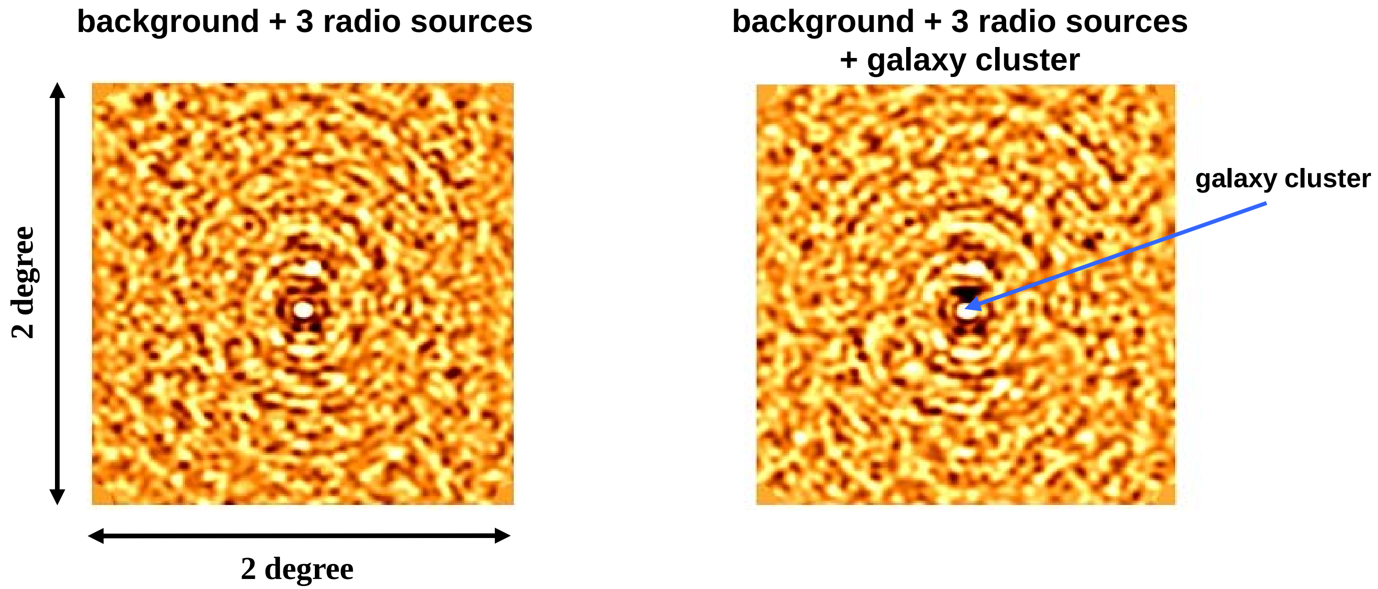


- 0.5 x 0.5 degree^2, 100 gal per arcmin^2 & $\sigma = 0.3$
- Concordance ΛCDM Cosmology with cluster mass & redshifts drawn from Press-Schechter mass function
- $p_{\mathrm{th}} = 0.5$

Applications: Clusters in Sunyaev Zel'dovich (SZ)

Feroz et al., 2009, arXiv:0811.1199

Galaxy cluster (and radio sources) in interferometric SZ data

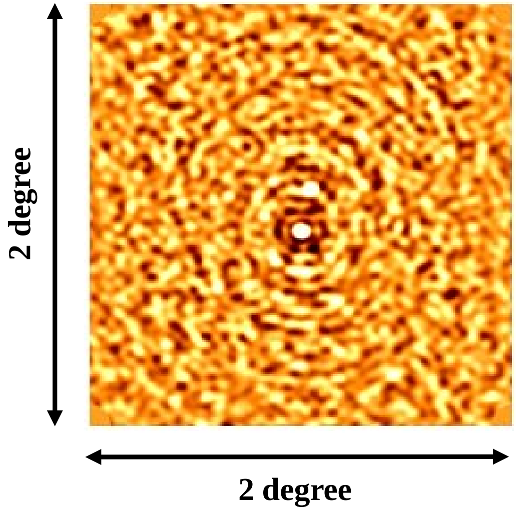


Applications: Clusters in Sunyaev Zel'dovich (SZ)

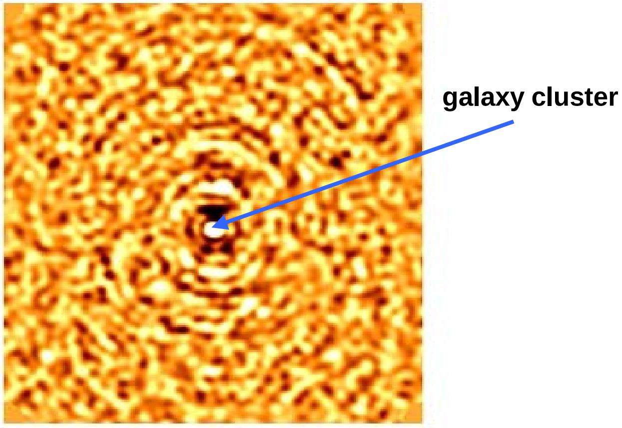
Feroz et al., 2009, arXiv:0811.1199

Galaxy cluster (and radio sources) in interferometric SZ data

background + 3 radio sources



background + 3 radio sources + galaxy cluster



Bayesian Model Comparison

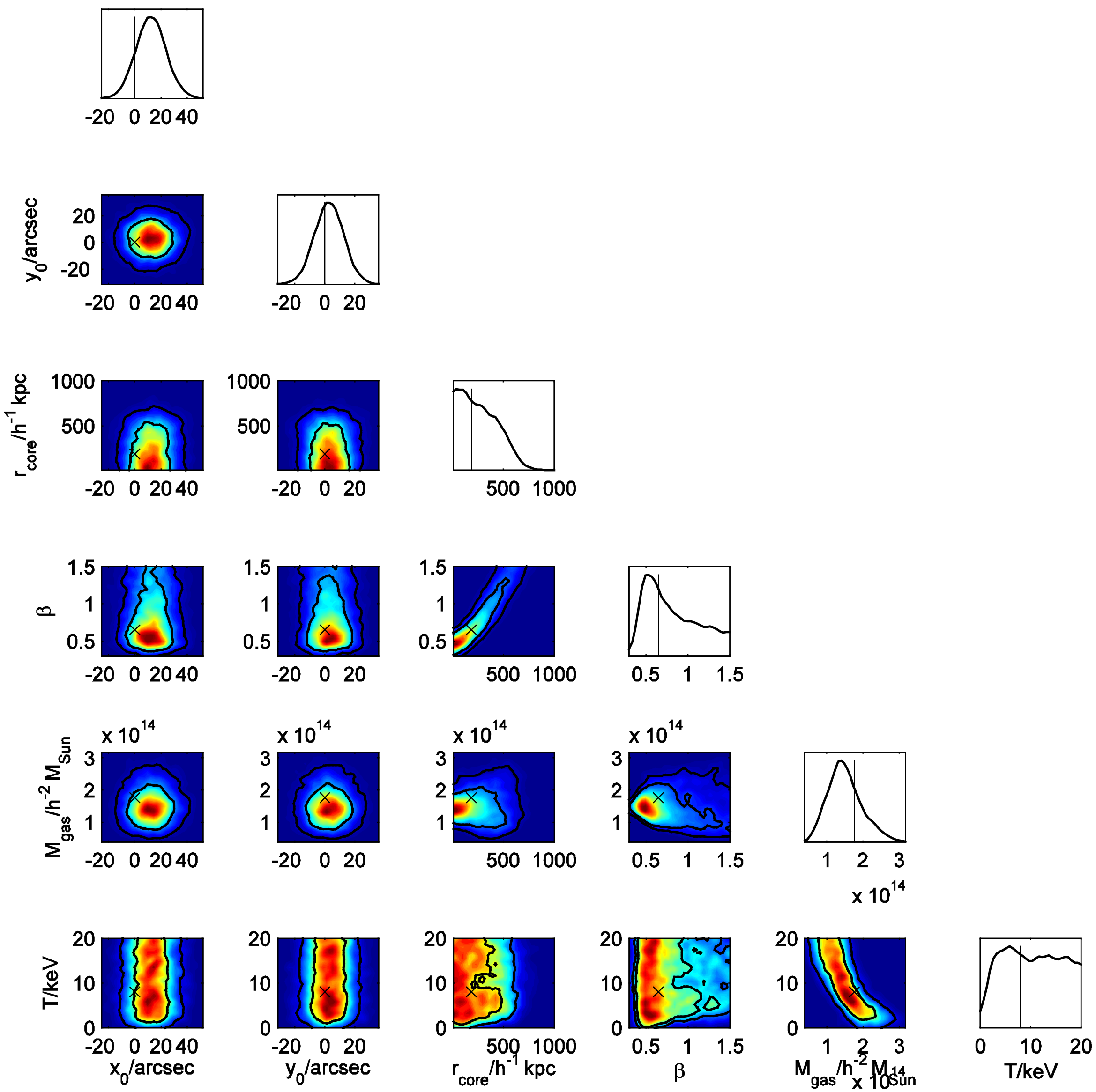
R = P(cluster | data)/P(no cluster | data)

R = 0.35

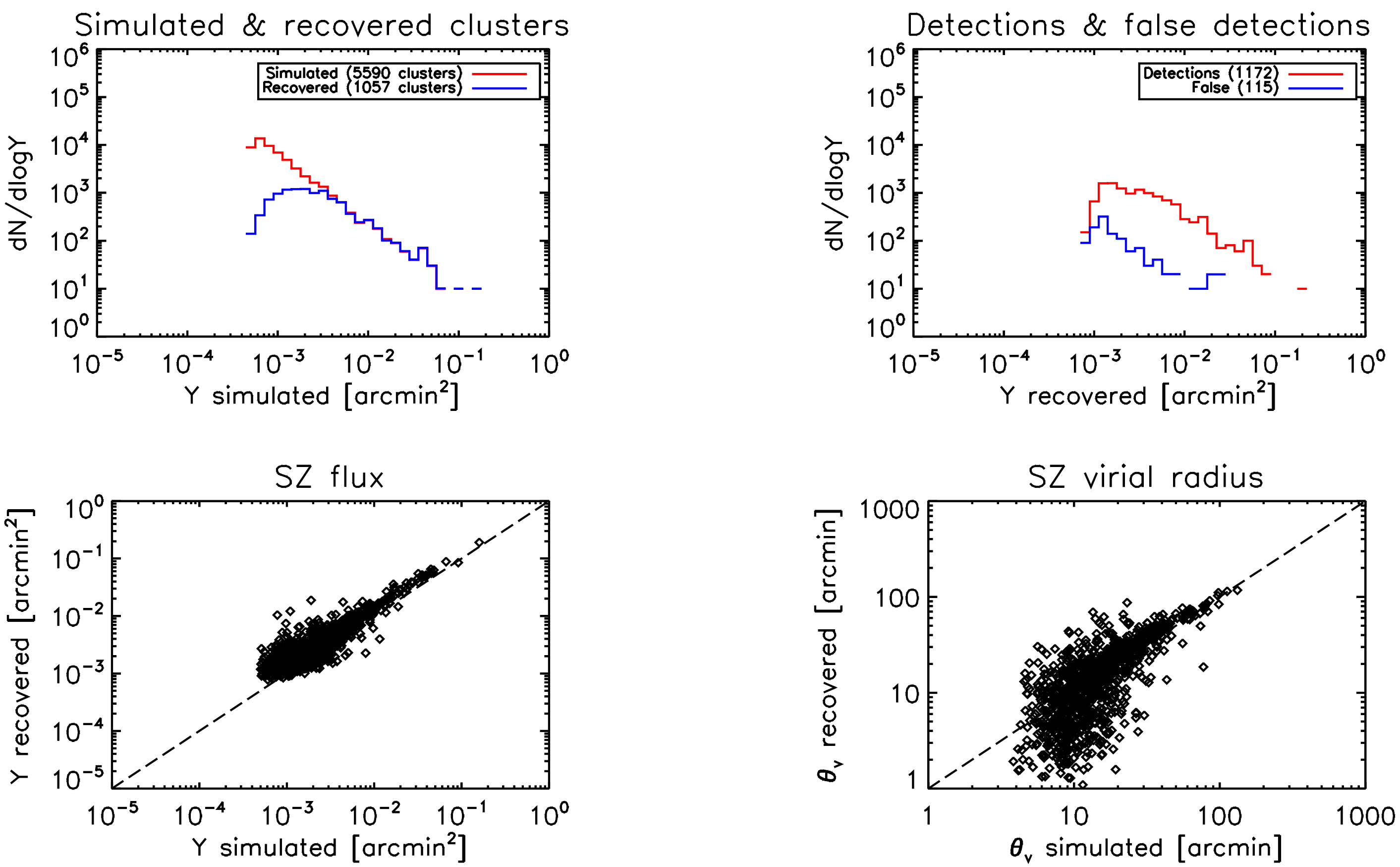
R ~ 10^{33}

Clusters in SZ – Parameter Constraints

Feroz et al., 2009, arXiv:0811.1199



Planck SZ Challenge II – Results with MultiNest



- 50×10^6 pixels, ~ 1000 recovered clusters, ~ 3 CPU hours

Applications: Gravitational Waves

- Simulated LISA data contatining two signals from non-spinning SMBH mergers. Each source has antipodal degeneracy ⇒ at least 4 modes
- All identified and well characterized by MultiNest
  (Feroz et al., 2009, arXiv:0904.1544)

		M_c	μ	t_c	θ	ϕ
1	$\sigma_{\rm FIM}$	1.289×10^3	4.719×10^3	1.209×10^{-5}	6.315×10^{-3}	1.159×10^{-2}
	$\Delta\lambda$	0.1164	0.0763	0.0511	0.0019	0.2036
1A	$\sigma_{\rm FIM}$	1.198×10^3	4.405×10^3	1.128×10^{-5}	9.683×10^{-3}	8.529×10^{-3}
	$\Delta\lambda$	0.1336	0.0795	0.9247	0.1532	0.6597
2	$\sigma_{\rm FIM}$	6.986×10^2	8.642×10^3	3.292×10^{-5}	6.283×10^{-3}	7.854×10^{-3}
	$\Delta\lambda$	0.7587	1.198	1.1562	0.9742	1.0675
2A	$\sigma_{\rm FIM}$	7.025×10^2	8.683×10^3	3.301×10^{-5}	6.446×10^{-3}	7.631×10^{-3}
	$\Delta\lambda$	3.3452	4.0735	2.7418	0.3818	1.0907

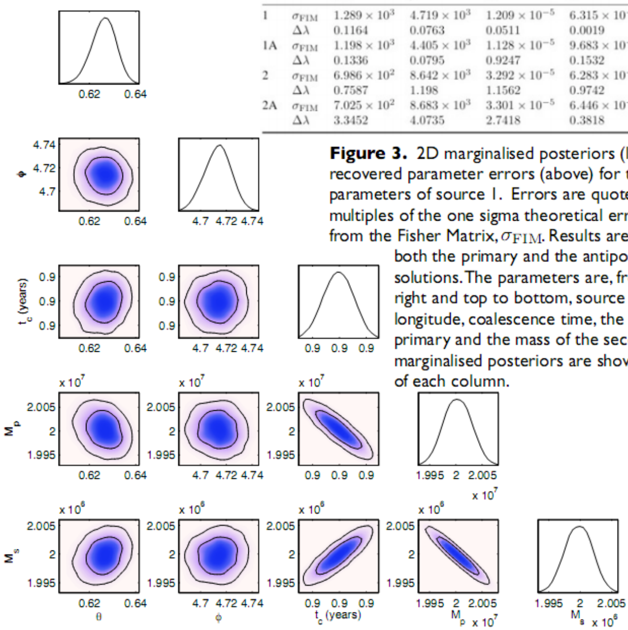


Figure 3. 2D marginalised posteriors (left) and recovered parameter errors (above) for the intrinsic parameters of source 1. Errors are quoted as multiples of the one sigma theoretical error computed from the Fisher Matrix, $\sigma_{\rm FIM}$. Results are reported for both the primary and the antipodal ('A') sky solutions. The parameters are, from left to right and top to bottom, source latitude and longitude, coalescence time, the mass of the primary and the mass of the secondary. 1D marginalised posteriors are shown at the top of each column.

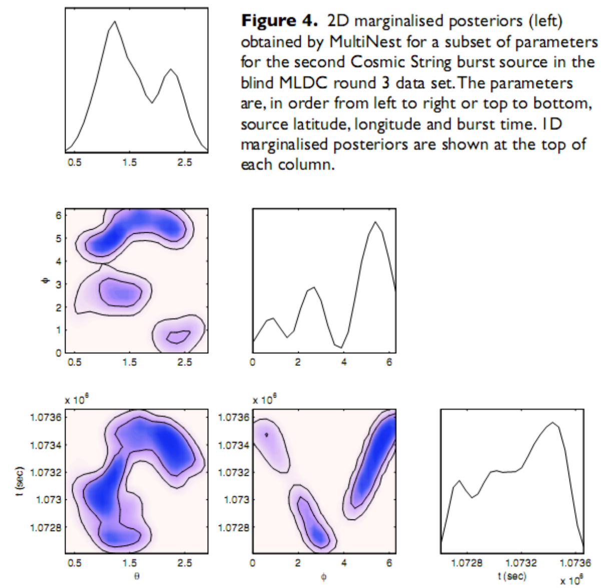


Figure 4. 2D marginalised posteriors (left) obtained by MultiNest for a subset of parameters for the second Cosmic String burst source in the blind MLDC round 3 data set. The parameters are, in order from left to right or top to bottom, source latitude, longitude and burst time. 1D marginalised posteriors are shown at the top of each column.

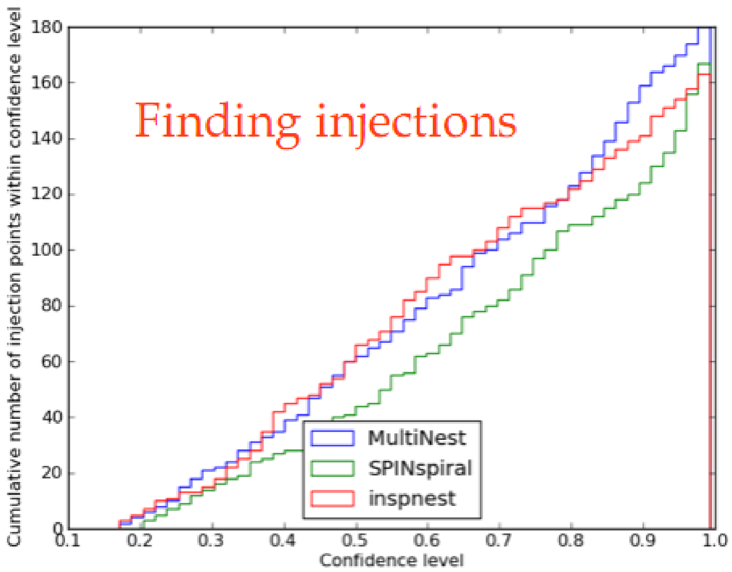
- Also applied successfully in Mock LISA Data Challenge Round 3 to simulations of 5 spinning SMBH binary inspirals and 3 cosmic strings

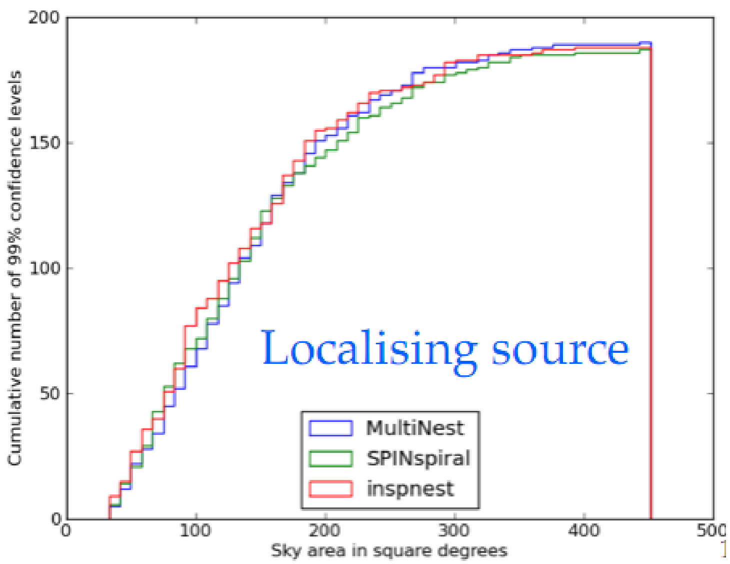
  (Feroz et al., 2010, arXiv:0911.0288)

Applications: Gravitational Waves

- MultiNest integrated into data analysis pipeline of LIGO:
  - Fully coherent analyses for follow-up of events using the network of detectors.
  - Infer physical parameters of the waveforms.

- Results from the sky localization study, 190 compact coalescing binaries injected





Bayesian Object Detection: Iterative Approach

- Can be used when single-object model is not valid
  - Overlapping/correlated (in terms of data) sources
- Fit *n*-source model and determine the distribution of residual data
  - $$\Pr(\mathbf{R}_n|\mathbf{D}, H_n) = \int \Pr(\mathbf{R}_n|\Theta, H_n) \Pr(\Theta|\mathbf{D}, H_n)\, d\Theta$$
- Analyse residual data and compare between:
  - H_0 = "there is no additional object, residual data is due to noise only"
  - H_1 = "there is an additional object present"
- If H_1 is preferred then fit for n+1 sources and repeat the procedure
- Example: Extra-solar planet detection
  - See Feroz, Balan & Hobson, 2011, arXiv:1012.5129

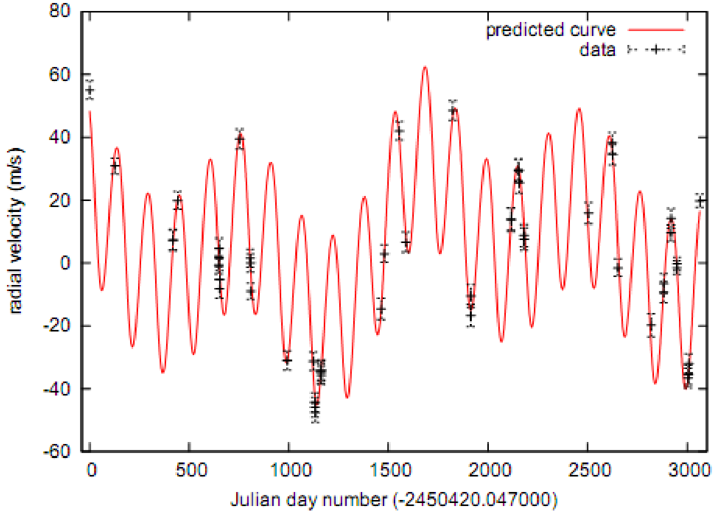
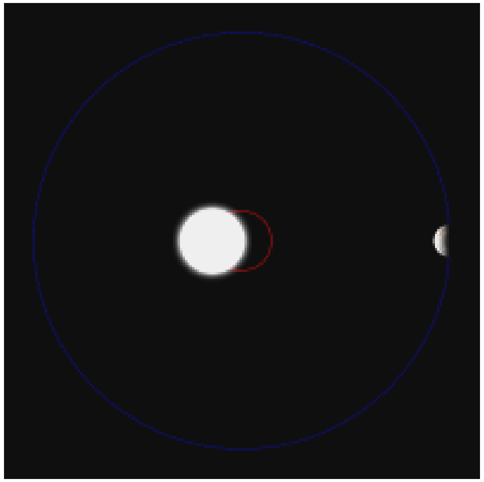


Figure 1. Radial velocity measurements, with 1σ errorbars, and the mean fitted radial velocity curve with three planets for HD 37124.

Applications: Exoplanet Detection

$$v(t_i, j) = V_j - \sum_{p=1}^{N_p} K_p \left[\sin(f_{i,p} + \omega_p) + e_p \sin(\omega_p) \right]$$



where

V_j = systematic velocity with reference to j^{th} observatory

K_p = velocity semi-amplitude of the p^{th} planet

ω_p = longitude of periastron of the p^{th} planet

$f_{i,p,}$ = true anomaly of the p^{th} planet

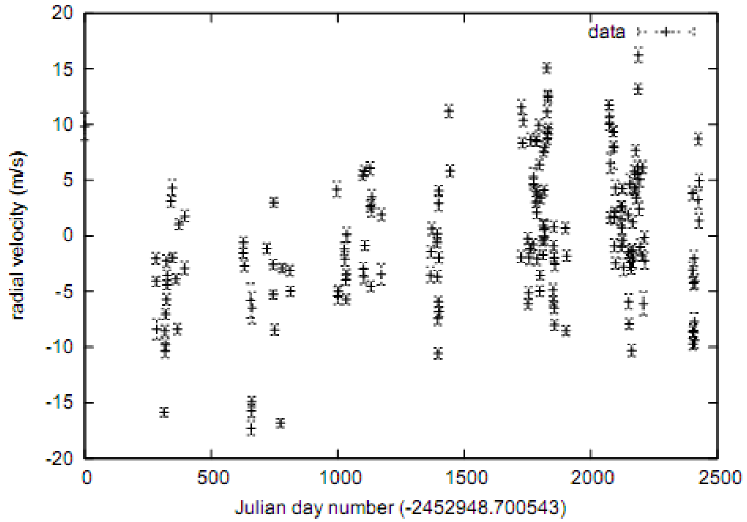
e_p = orbital eccentricity of the p^{th} planet

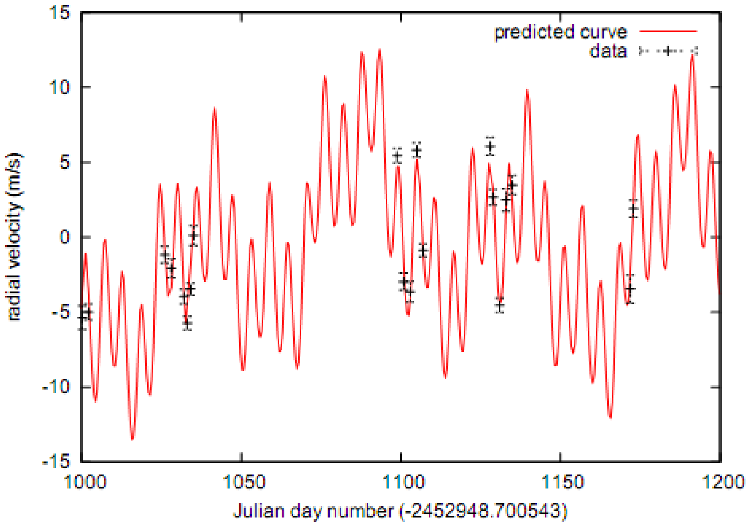
e_p = orbital period of the p^{th} planet

e_p = fraction of an orbit of the p^{th} planet, prior to the start of data taking at which periastron occurred

Applications: Exoplanet Detection – HD 10180

Feroz, Balan & Hobson, 2011, arXiv:1012.5129



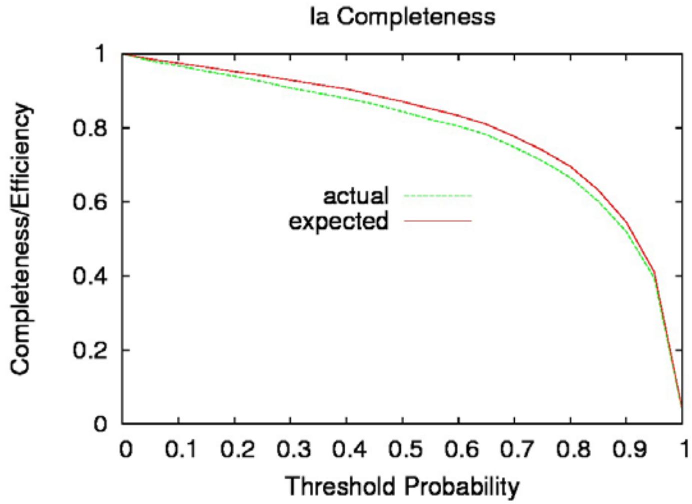


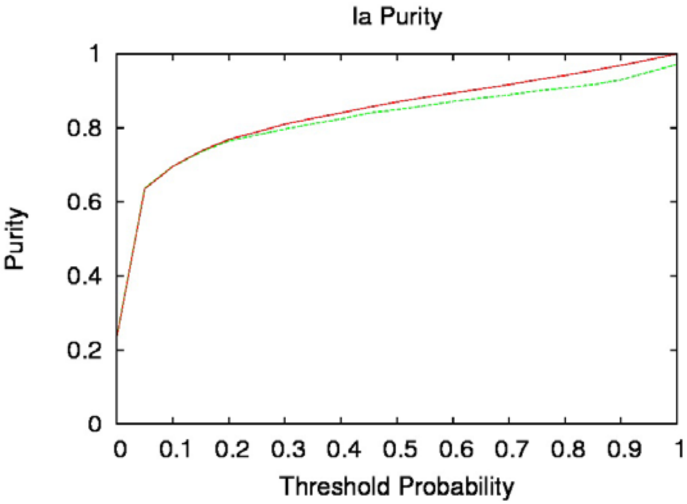
Parameter	HD 10180 b	HD 10180 c	HD 10180 d	HD 10180 e	HD 10180 f	HD 10180 g
P (days)	5.76 ± 0.02 (5.76)	16.35 ± 0.05 (16.36)	49.74 ± 0.20 (49.74)	122.75 ± 0.54 (122.69)	600.17 ± 13.75 (601.88)	2266.22 ± 412.42 (2231.44)
K (m/s)	4.54 ± 0.12 (4.63)	2.89 ± 0.13 (2.94)	4.28 ± 0.14 (4.25)	2.91 ± 0.14 (2.70)	1.43 ± 0.20 (1.79)	3.06 ± 0.16 (2.98)
e	0.07 ± 0.03 (0.08)	0.13 ± 0.04 (0.12)	0.03 ± 0.02 (0.03)	0.09 ± 0.04 (0.08)	0.15 ± 0.09 (0.25)	0.09 ± 0.05 (0.05)
ϖ (rad)	2.60 ± 0.38 (2.51)	2.62 ± 0.35 (2.49)	2.56 ± 0.16 (5.12)	2.65 ± 0.53 (2.95)	3.08 ± 0.97 (2.43)	2.89 ± 2.60 (5.98)
χ	0.22 ± 0.06 (0.24)	0.35 ± 0.06 (0.37)	0.43 ± 0.27 (0.83)	0.23 ± 0.11 (0.16)	0.31 ± 0.28 (0.27)	0.67 ± 0.10 (0.73)
$m \sin i$ (M_J)	0.04 ± 0.00 (0.04)	0.04 ± 0.00 (0.04)	0.08 ± 0.00 (0.08)	0.07 ± 0.00 (0.07)	0.06 ± 0.00 (0.07)	0.20 ± 0.01 (0.20)
a (AU)	0.06 ± 0.00 (0.06)	0.13 ± 0.00 (0.13)	0.27 ± 0.00 (0.27)	0.49 ± 0.00 (0.49)	1.42 ± 0.03 (1.42)	3.45 ± 0.16 (3.40)

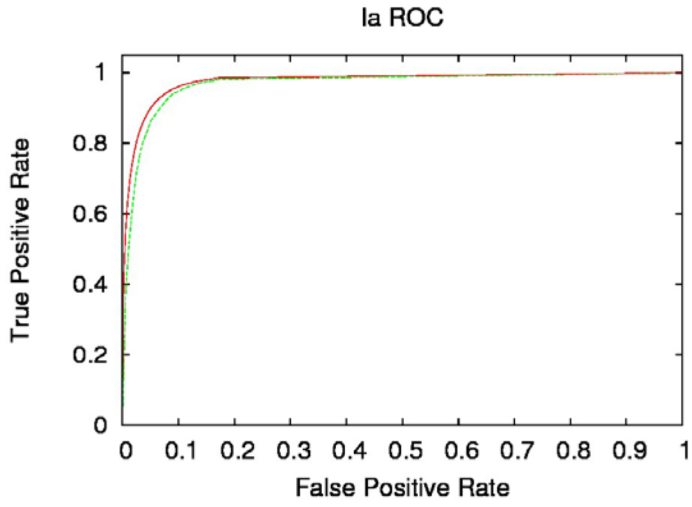
Purity, Completeness & Threshold Probability

- H_0 = "there is no object with its centre lying in the region S"
- H_1 = "there is one object with its centre lying in the region S"
- $R = \frac{\Pr(H_1 \mid \mathrm{D})}{\Pr(H_0 \mid \mathrm{D})} = \frac{\Pr(\mathrm{D} \mid H_1)\Pr(H_1)}{\Pr(\mathrm{D} \mid H_0)\Pr(H_0)} = \frac{Z_1 \Pr(H_1)}{Z_0 \Pr(H_0)}$
- $p_{TP} = R/(1+R)$
- Expected number of objects $\hat{N}_{tot} = \sum_{i=1}^{N} p_{TP,i}$
- Expected number of *true positives* $\hat{N}_{TP} = \sum_{i=1,\, p_{TP,i} > p_{th}}^{N} p_{TP,i}$
- Expected number of *false positives* $\hat{N}_{FP} = \sum_{i=1,\, p_{TP,i} > p_{th}}^{N} (1 - p_{TP,i})$
- Expected *completeness* $\hat{\varepsilon} = \hat{N}_{TP} / \hat{N}_{tot}$
- Expected *purity* $\hat{\tau} = \hat{N}_{TP} / (\hat{N}_{TP} + \hat{N}_{FP})$

Purity, Completeness & Threshold Probability



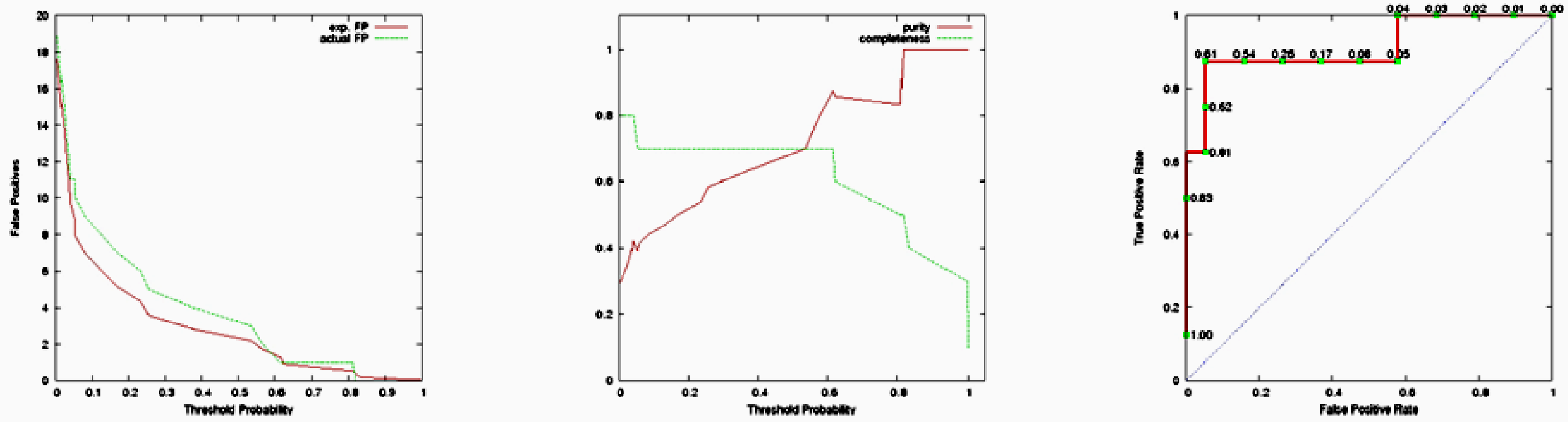




- **Probabilistic classification of type-Ia Supernovae using Neural Network**

- Karpenka, Feroz & Hobson, 2012, arXiv:1208.1264

Purity, Completeness & Threshold Probability



- **Galaxy cluster detection in weak lensing surveys**
- Feroz, Marshall & Hobson, 2008, arXiv:0810.0781

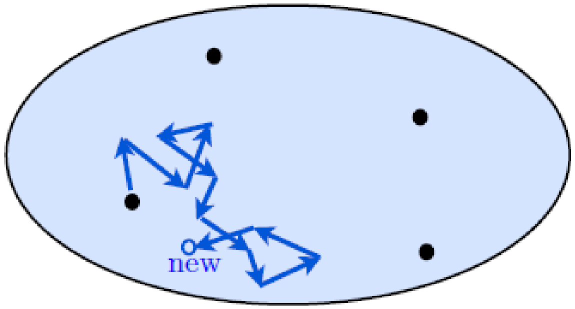
Conclusions

- Bayesian statistics provide rigorous approach to astrophysical object detection
  - Use Bayesian model selection to distinguish real objects from spurious ones
- Efficient and robust object detection can be done using nested sampling
  - MultiNest allows sampling from multimodal/degenerate posteriors
  - local and global evidences and parameter constraints
  - typically ~ 100 times more efficient than standard MCMC
- Probabilistic object detection removes arbitrariness in choice of detection criterion
  - allows calculation of expected purity and completeness
- MultiNest publicly available
  - with SuperBayeS for SUSY phenomenology (www.superbayes.org)
  - as a standalone inference engine (www.mrao.cam.ac.uk/software/multinest)

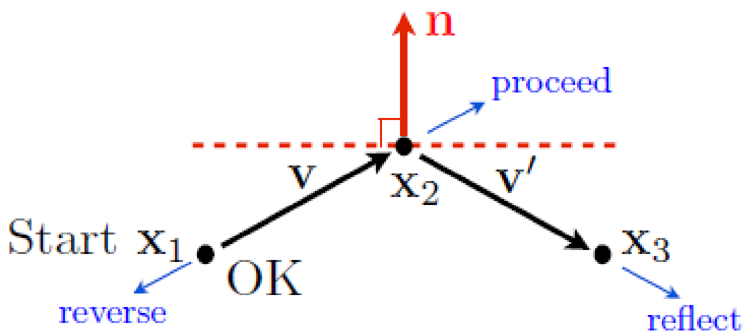
Supplementary Slides

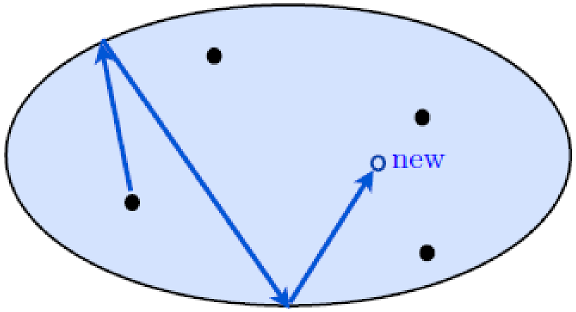
Galilean Monte Carlo

Nested Sampling needs to generate a new point from constrained prior



Generating new point using MCMC
Problem: Diffusion time very long

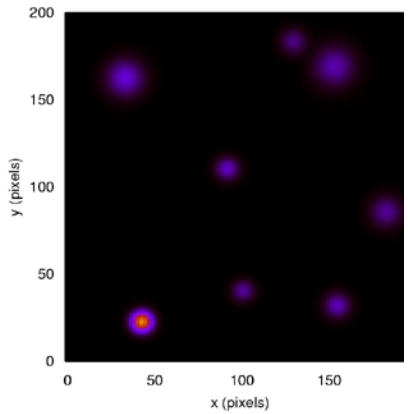


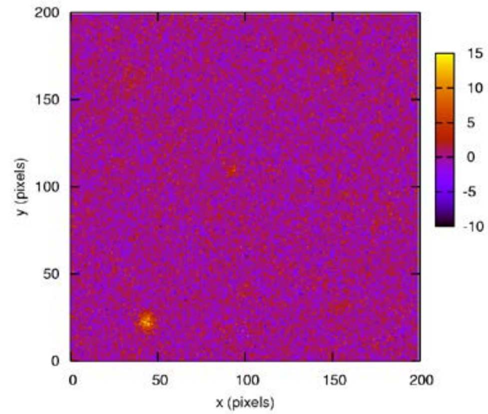


Generating new point using Hamiltonian Monte Carlo
(Reflective Slice Sampling of Radford Neal)
Problem: Don't know the edge of constrained prior

1. Start at x_1 where $L(x_1) > L'$
2. Propose $x_2 = x_1 + v$
3. If $(L(x_2) > L')$
        Proceed to x_2
    Else
        Reflect to x_3
        If $(L(x_3) > L')$
            Proceed to x_3
        Else
            Reverse

Object Detection Toy Problem





- Signal from a circularly symmetric Gaussian shaped object

$$s(a) = A\exp\left[-\frac{(x-X)^2+(y-Y)^2}{2R^2}\right], \text{ so that } a=\{X,Y,A,R\}$$

- With k such discrete objects and the generalized noise contribution n, the data

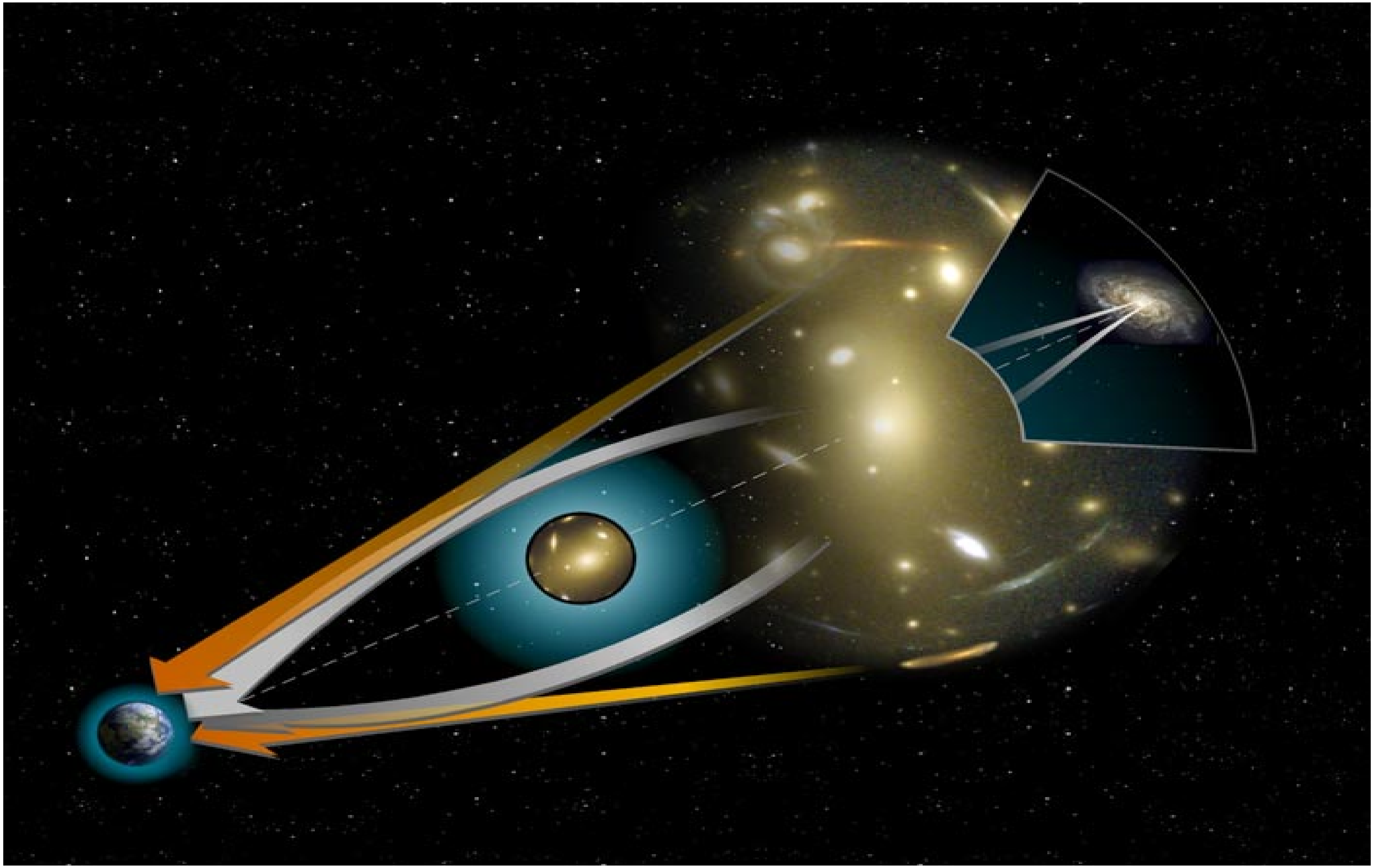
$$D = n + \sum_{i=1}^{k} s(a_i)$$

- Likelihood function takes the form

$$P(D|\theta) = \frac{\exp\left\{-\tfrac{1}{2}\left[D-s(a)\right]^t N^{-1}\left[D-s(a)\right]\right\}}{\left|2\pi N\right|^{1/2}}, \text{ where}$$

N = $<nn^t>$ is the noise covariance matrix

Applications: Clusters in Weak Lensing



Applications: Clusters in Weak Lensing