Data Analysis II

Signal Templates, Un-modeled Signals, Stochastic Signals

Parameter Estimation

Neil J. Cornish

A unified approach to GW data analysis

- To detect a signal you need a model for the signal *and* a model for the noise
- The noise model defines the likelihood function
- The signal model defines priors on the wave amplitude in time/frequency

Prior, Likelihood, Posterior Redux

$$\mathbf{d} = \mathbf{h} + \mathbf{n}$$

$\mathbf{d} \rightarrow \text{data}$ $\qquad$ $\mathbf{h} \rightarrow \text{signal in the detector}$ $\qquad$ $\mathbf{n} \rightarrow \text{noise in the detector}$

Posterior distribution for the data $\quad p(\mathbf{d}|\mathbf{h}, \mathbf{n}) = \delta(\mathbf{h} + \mathbf{n} - \mathbf{d})$

Prior distributions for the signal and noise models $\quad p(\mathbf{h}), \quad p(\mathbf{n})$

Example: Correlated Gaussian Noise $\quad p(n_1, n_2, \dots, n_N) = \dfrac{1}{\sqrt{\det(2\pi C)}} e^{-\frac{1}{2}\, n_i C^{-1}_{ij} n_j}$

Noise Correlation Matrix $\quad C_{ij} = E[n_i\, n_j]$ $\qquad$ (Uncorrelated Noise has $C_{ij} = \delta_{ij}\, \sigma_i^2$)

Prior, Likelihood, Posterior Redux

Using the Identity

$$p(\mathbf{h},\mathbf{n}|\mathbf{d})p(\mathbf{d}) = p(\mathbf{d}|\mathbf{h},\mathbf{n})p(\mathbf{n})p(\mathbf{h})$$

Posterior for signal

$$p(\mathbf{h}|\mathbf{d}) = \int \frac{p(\mathbf{d}|\mathbf{h},\mathbf{n})p(\mathbf{n})p(\mathbf{h})}{p(\mathbf{d})}\, d\mathbf{n} = \frac{p(\mathbf{d}|\mathbf{h})p(\mathbf{h})}{p(\mathbf{d})}$$

Likelihood

$$p(\mathbf{d}|\mathbf{h}) = p(\mathbf{n} = \mathbf{d} - \mathbf{h})$$

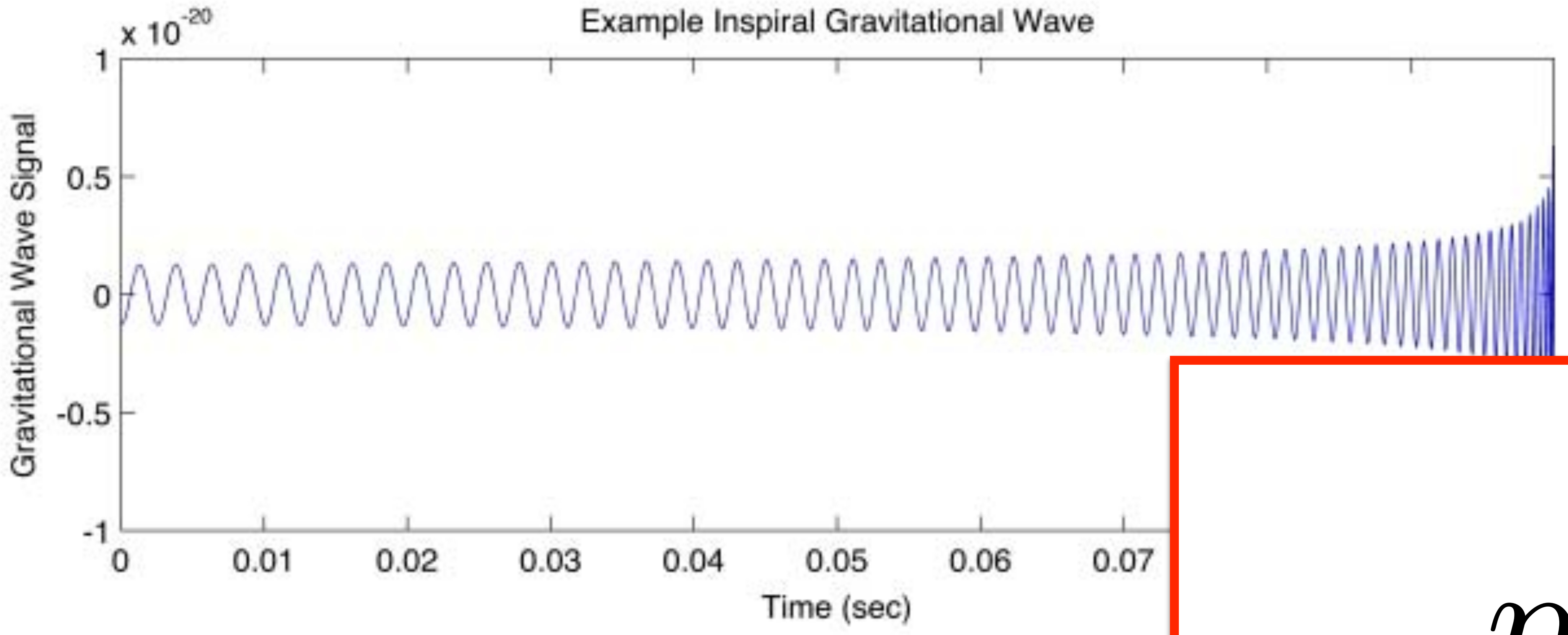
Likelihood = Noise Model

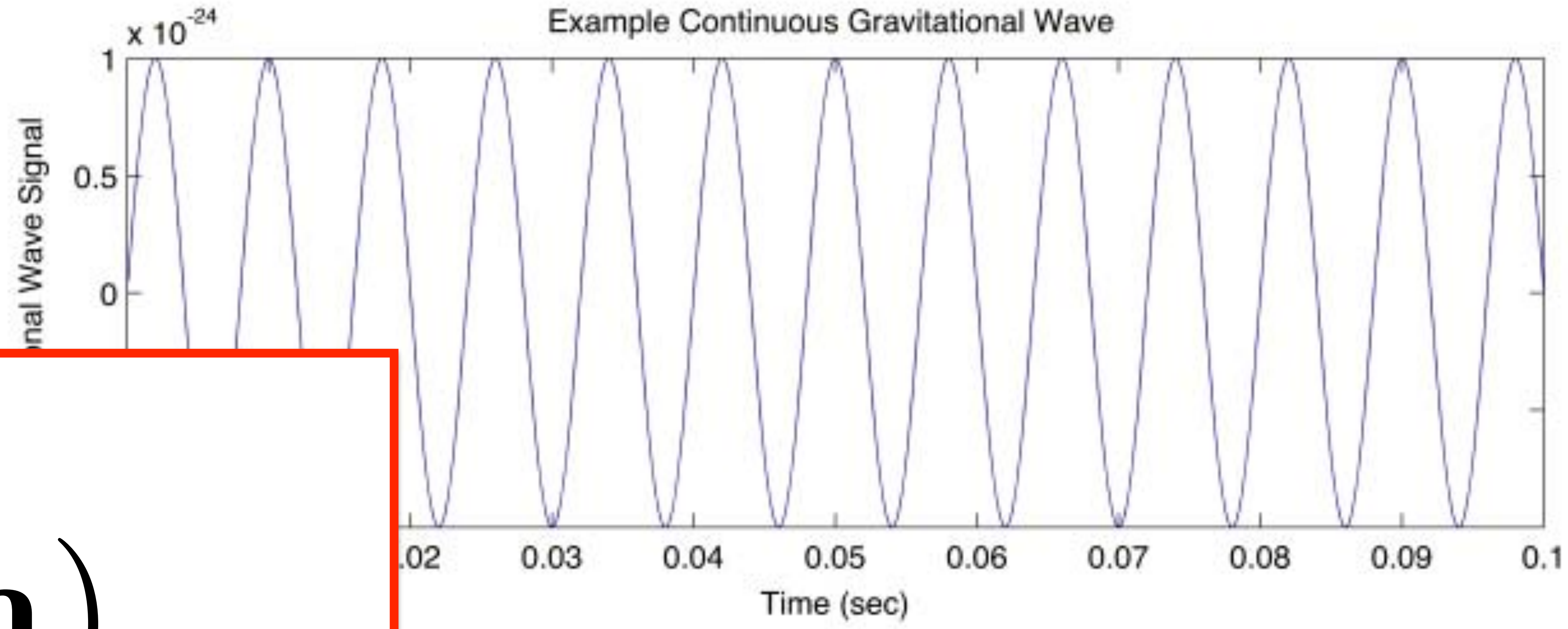
Example: Correlated Gaussian Noise

$$p(\mathbf{d}|\mathbf{h}) = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^\dagger\,\mathbf{C}^{-1}(\mathbf{d}-\mathbf{h})}$$

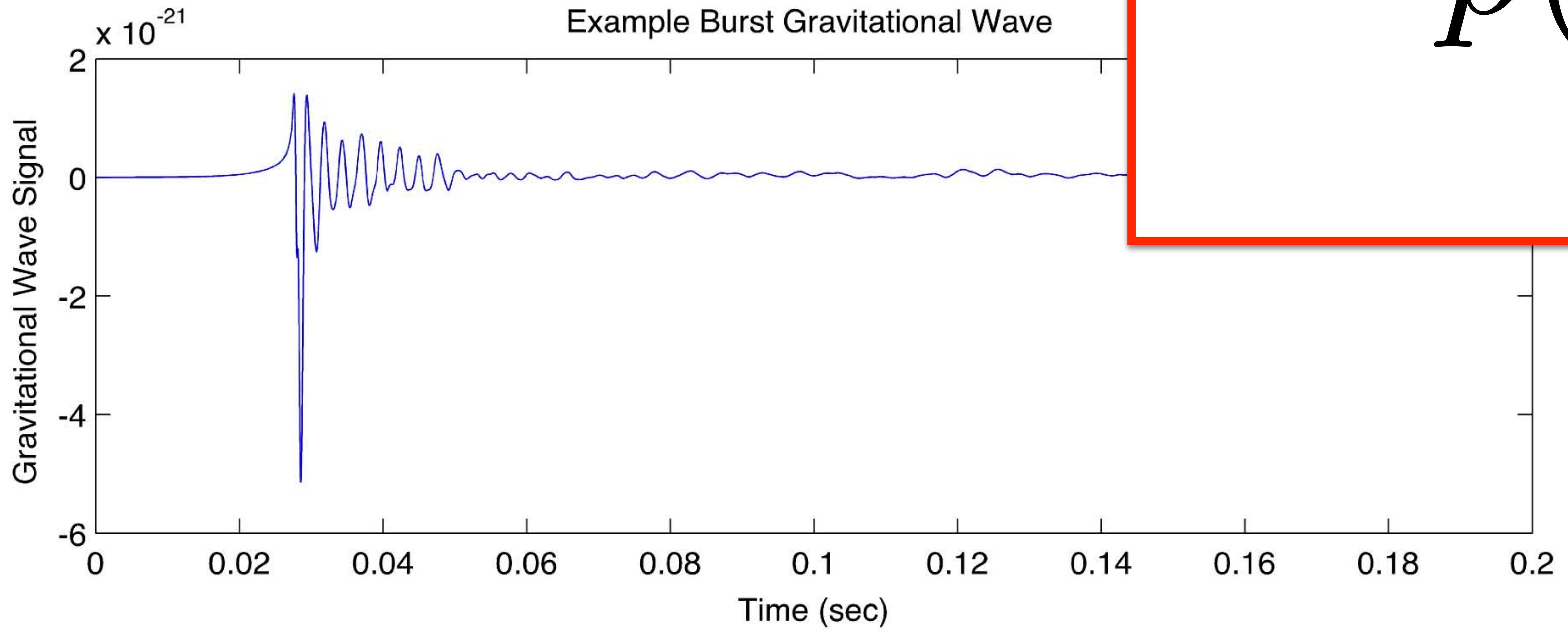
Gravitational wave signal types

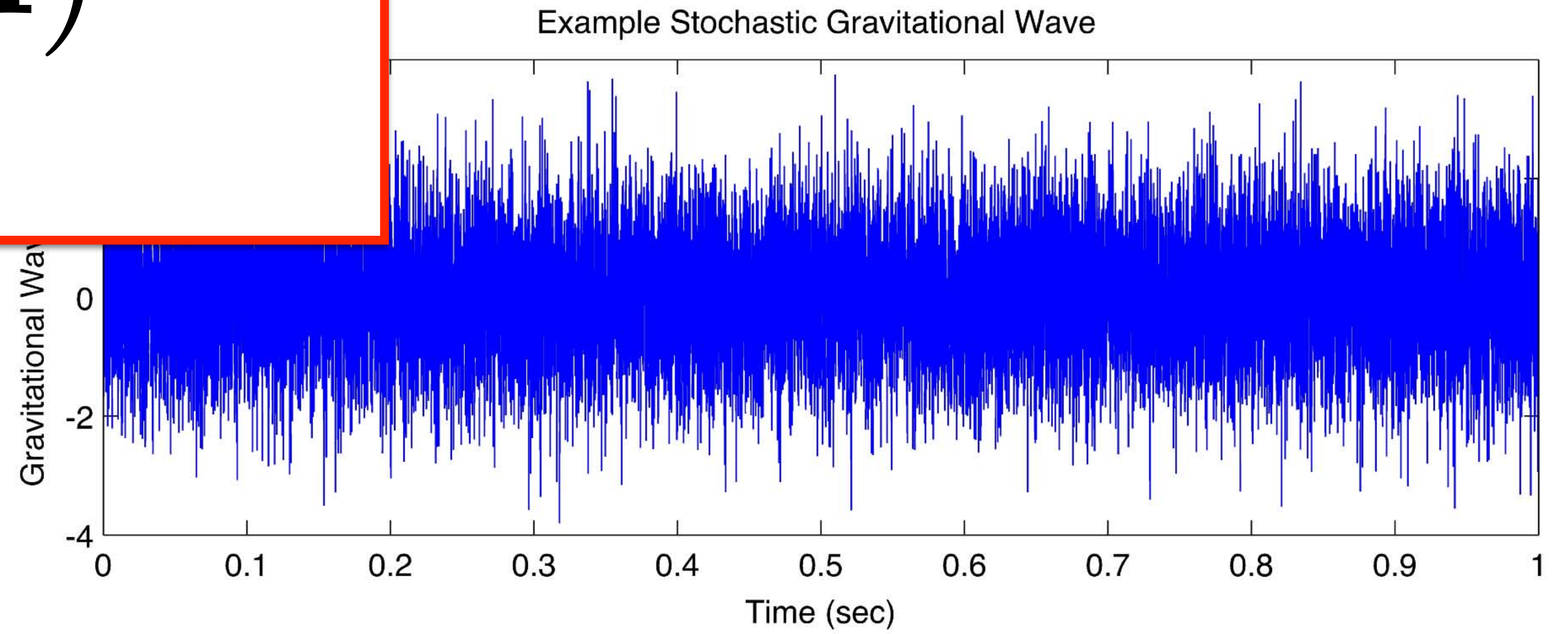
Well modeled - e.g. binary inspiral and merger





$$p(\mathbf{h})$$





Poorly modeled - e.g. core collapse supernovae

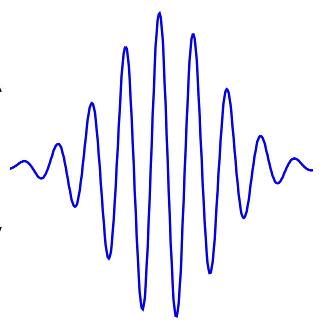
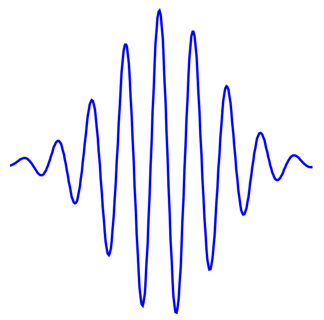
Stochastic- e.g. phase transition in early universe

Gravitational wave signal models

Template based

$$p(\mathbf{h}) = \delta(\mathbf{h} - \mathbf{h}(\vec{\lambda}))p(\vec{\lambda})$$

Burst signals

$p(\mathbf{h}) = \delta\Big(\mathbf{h} - \sum$  $\Big)p\Big($  $\Big)$

Stochastic signals

$$p(\mathbf{h}) = \frac{1}{\sqrt{\det(2\pi \mathbf{S_h})}} e^{-\frac{1}{2}(\mathbf{h}^\dagger \mathbf{S}_\mathbf{h}^{-1} \mathbf{h})}\, p(\mathbf{S_h})$$

Basic Noise Model: Colored, Stationary & Gaussian

$$p(\mathbf{d}|\mathbf{h}) = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^\dagger \mathbf{C}^{-1}(\mathbf{d}-\mathbf{h})}$$

If the noise is stationary, then the noise correlation matrix **C** only depends on the time difference. Implies that **C** is diagonal in the Fourier domain.

$$\log p(\mathbf{d}|\mathbf{h}) = -\frac{1}{2}(\mathbf{d}-\mathbf{h}|\mathbf{d}-\mathbf{h}) - \frac{1}{2}\int_0^\infty T_{\rm obs} \log[T_{\rm obs}\pi S_n(f)]\, df$$

$$(\mathbf{a}|\mathbf{b}) = 2\int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f) + \tilde{a}^*(f)\tilde{b}(f)}{S_n(f)}\, df$$

$$\mathbb{E}[\tilde{n}(f)\tilde{n}^*(f')] = \frac{1}{2}\delta(f-f')S_n(f)$$

$$\mathrm{SNR}^2 = 4\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)}\, df$$

Multiple Detectors: Uncorrelated, Colored, Stationary & Gaussian Noise

$$p(\mathbf{d}|\mathbf{h}) \propto e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h}|\mathbf{d}-\mathbf{h})}$$

$$(\mathbf{a}|\mathbf{b}) = 2\sum_{I=1}^{N}\int_0^\infty \frac{\tilde{a}_I(f)\tilde{b}_I^*(f) + \tilde{a}_I^*(f)\tilde{b}_I(f)}{S_{n,I}(f)}\,df$$

Stochastic Signals

$$p(\mathbf{h}) = \frac{1}{\sqrt{\det(2\pi \mathbf{S_h})}} \, e^{-\frac{1}{2}(\mathbf{h}^\dagger \mathbf{S_h}^{-1} \mathbf{h})} \, p(\mathbf{S_h})$$

We are not interested in the value of each GW signal sample $\tilde{h}(f)$. Want to infer the power spectrum $S_h(f)$

Marginalize over h: $$p(S_h|\mathbf{d}) = \int \frac{p(\mathbf{d}|\mathbf{h})p(\mathbf{h})}{p(\mathbf{d})} \, d\mathbf{h} = \frac{p(\mathbf{d}|S_h)p(S_h)}{p(\mathbf{d})}$$

Stochastic Signals

The integration over h is easy as it just involves Gaussians [Cornish & Romano, PRD 2013]

$$p(\mathbf{d}|S_h) \propto e^{-\frac{1}{2}(\mathbf{d}|\mathbf{d})_S}$$

Where
$$(\mathbf{a}|\mathbf{b})_S = 2\sum_{I,J}\int_0^\infty \left(\tilde{a}_I(f)\tilde{b}_J^*(f) + \tilde{a}_I^*(f)\tilde{b}_J(f)\right) S_{IJ}^{-1}(f)\, df$$

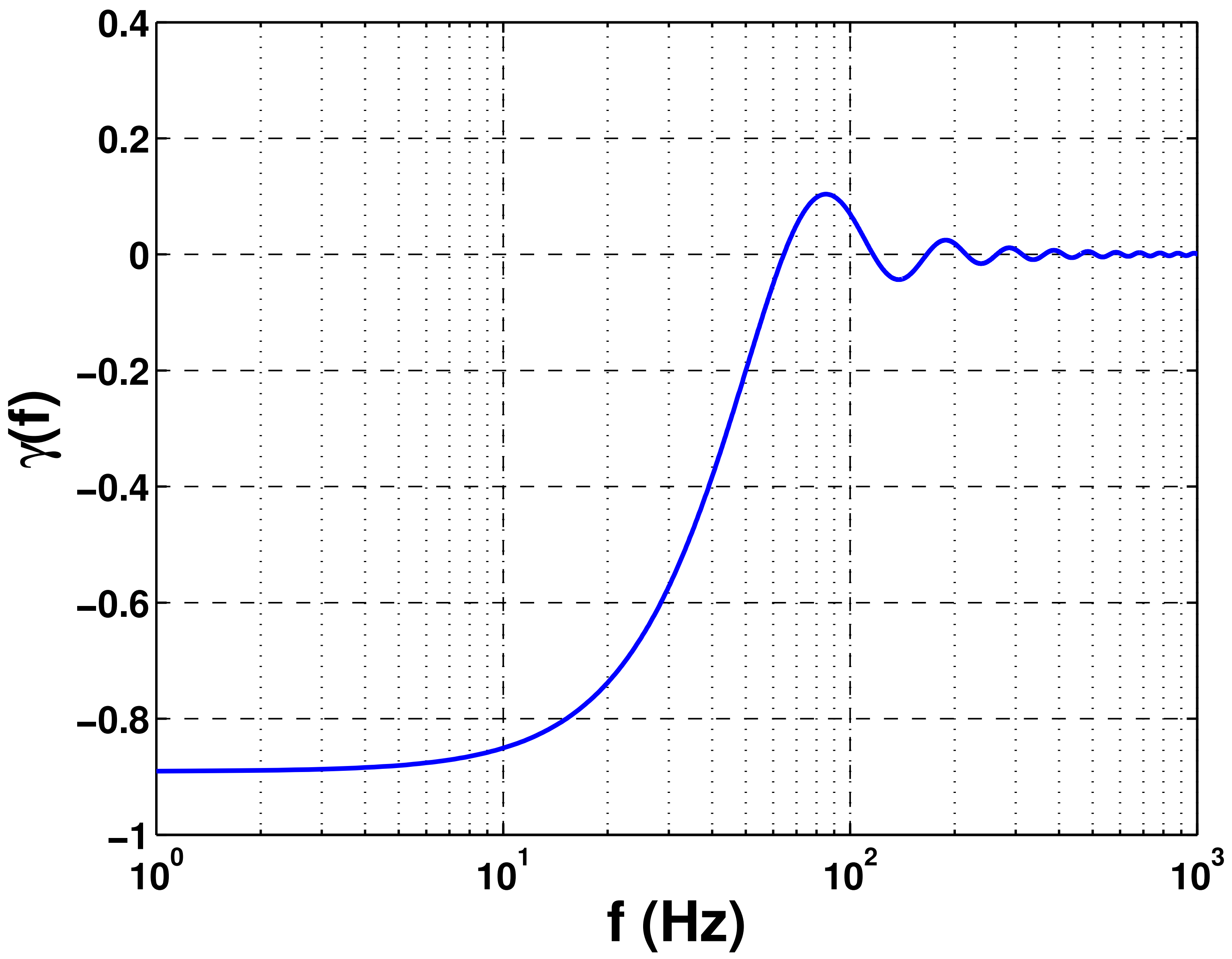
and
$$S_{IJ}(f) = S_{n,I}(f)\,\delta_{IJ} + S_h(f)\,\gamma_{IJ}(f)$$

Stochastic Signals

$$S_{IJ}(f) = S_{n,I}(f)\,\delta_{IJ} + S_h(f)\,\gamma_{IJ}(f)$$

The quantity $\gamma_{IJ}(f)$ is a geometrical factor that encodes the response of the detectors. In the long wavelength limit it is called the overlap reduction function

$$\gamma_{IJ}(f) = \frac{1}{4\pi}\int (F_I^+(\hat{n})F_J^+(\hat{n}) + F_I^+(\hat{n})F_J^+(\hat{n}))e^{2\pi i f(\vec{x}_I - \vec{x}_J)\cdot\hat{n}} d\Omega_{\hat{n}}$$



Normalized overlap for LIGO Hanford/Livngston (normalization factor is x 5)

Stochastic Signals

Stochastic signal detection is usually described in terms of cross correlation

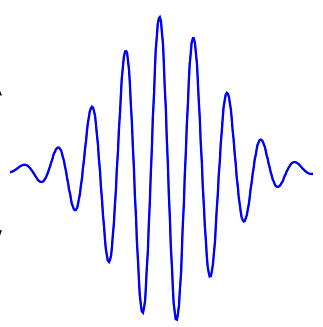
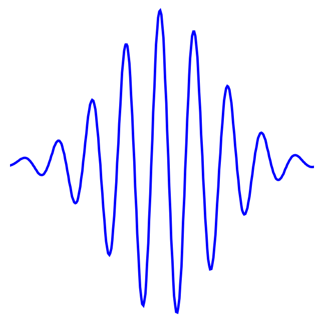
$$\mathbf{d}_I = \mathbf{n}_I + \mathbf{h}_I \qquad\qquad \mathbf{d}_J = \mathbf{n}_J + \mathbf{h}_J$$

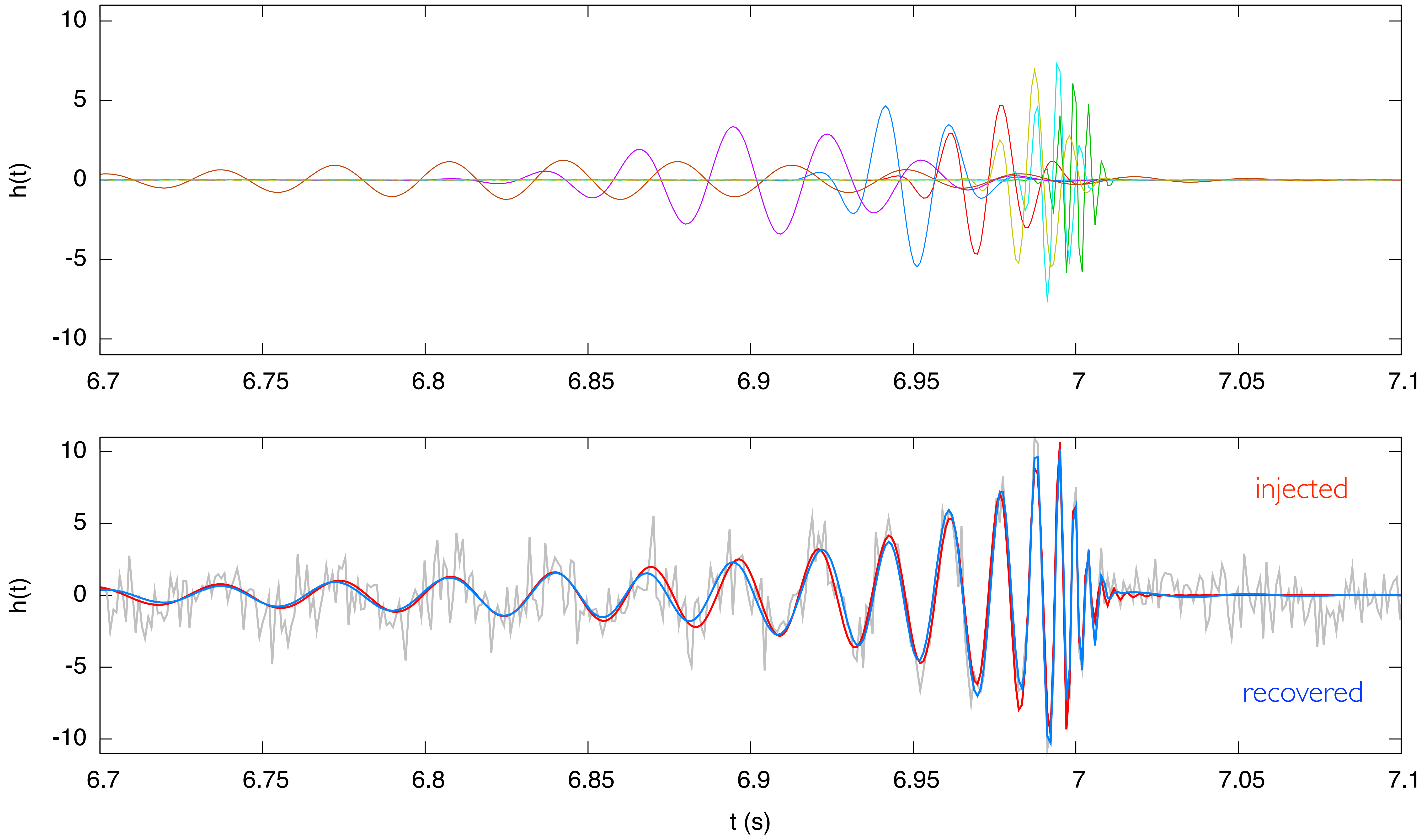
$$\langle \mathbf{d}_I \mathbf{d}_J \rangle = \underbrace{\langle \mathbf{n}_I \mathbf{n}_J \rangle}_{0} + \underbrace{\langle \mathbf{n}_I \mathbf{h}_J \rangle}_{0} + \underbrace{\langle \mathbf{h}_I \mathbf{n}_J \rangle}_{0} + \langle \mathbf{h}_I \mathbf{h}_J \rangle$$

In the weak signal limit $S_h(f) \ll S_n(f)$

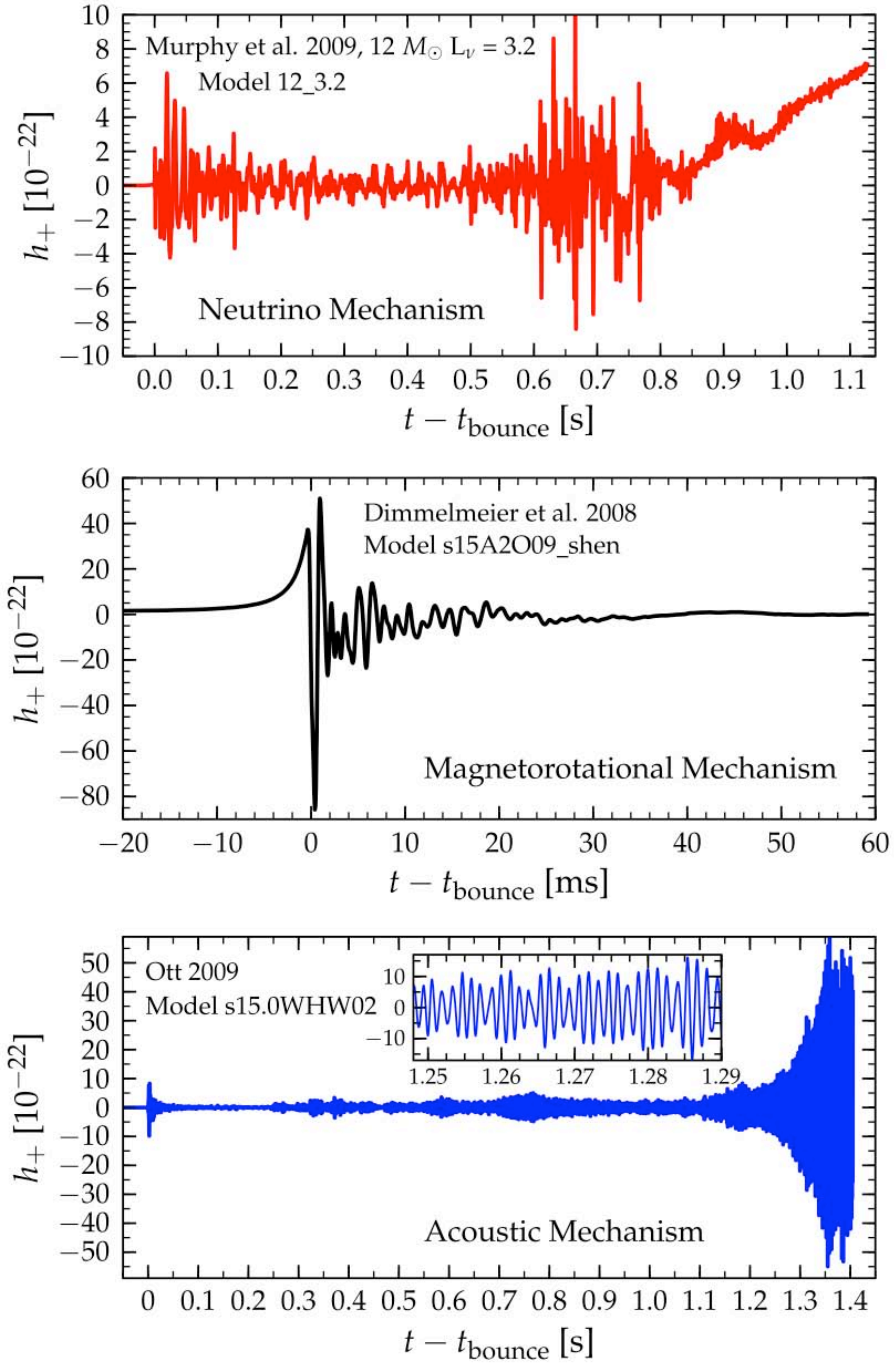
$$\log p(\mathbf{d}|S_h) = \text{const.} - \underbrace{\sum_{I \neq J} \int_0^\infty \frac{\left(\tilde{d}_I(f)\tilde{d}_J^*(f) + \tilde{d}_I^*(f)\tilde{d}_J(f)\right)\gamma_{IJ}(f) S_h(f)}{S_{n,I}(f) S_{n,J}(f)}\, df}_{\text{Optimal cross correlation statistic}}$$

Poorly Modelled Burst Signals

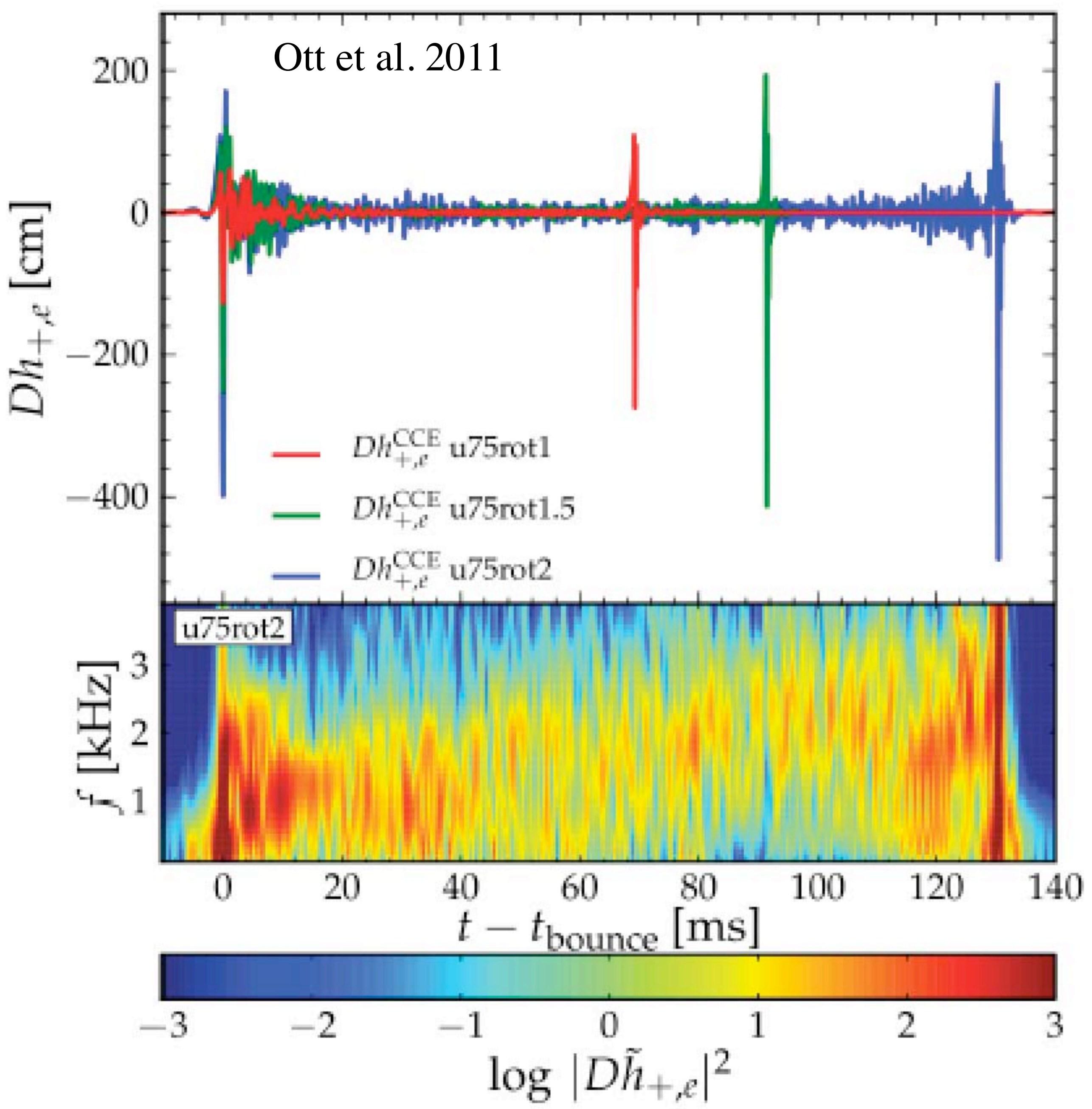
$p(\mathbf{h}) = \delta(\mathbf{h} - \sum$  $)p($  $)$



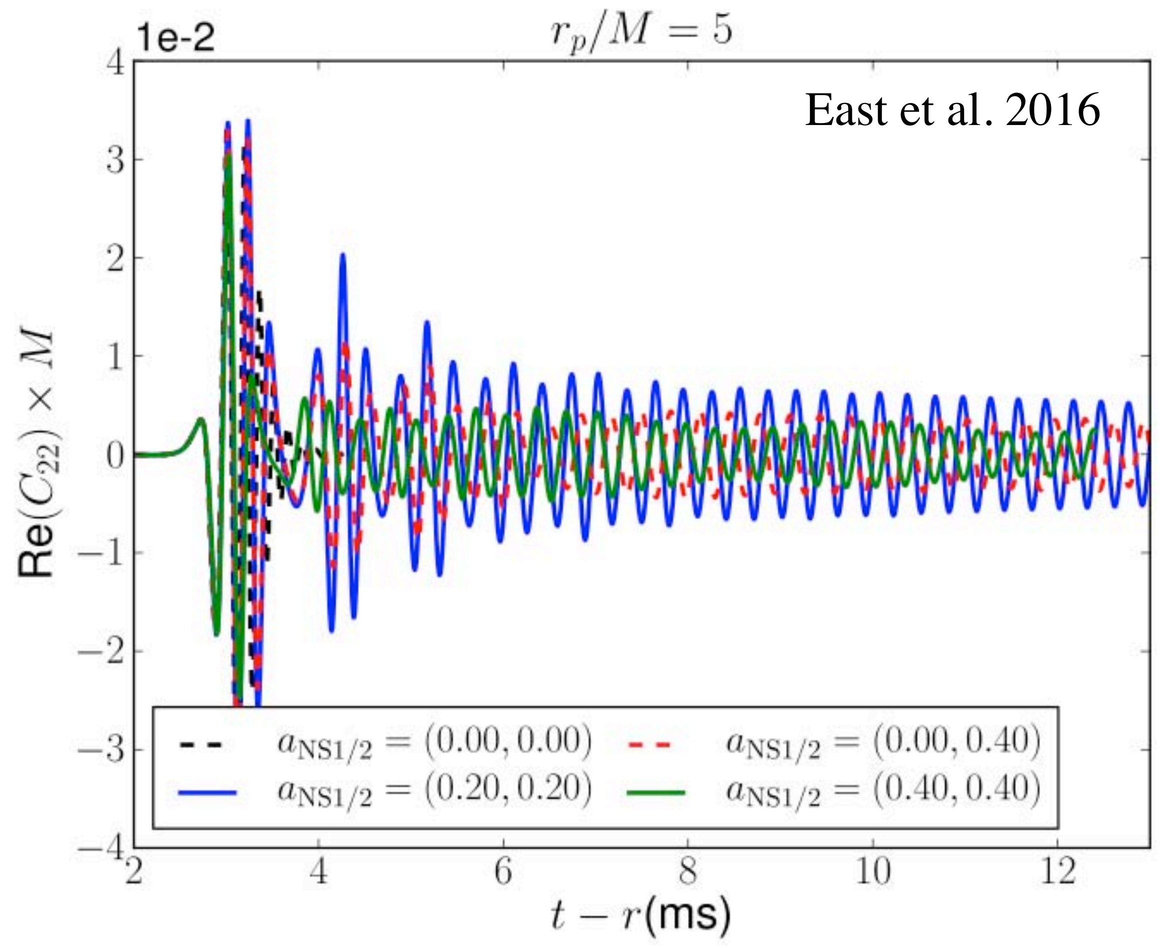
Things that go bump in the night



Core-collapse supernovae



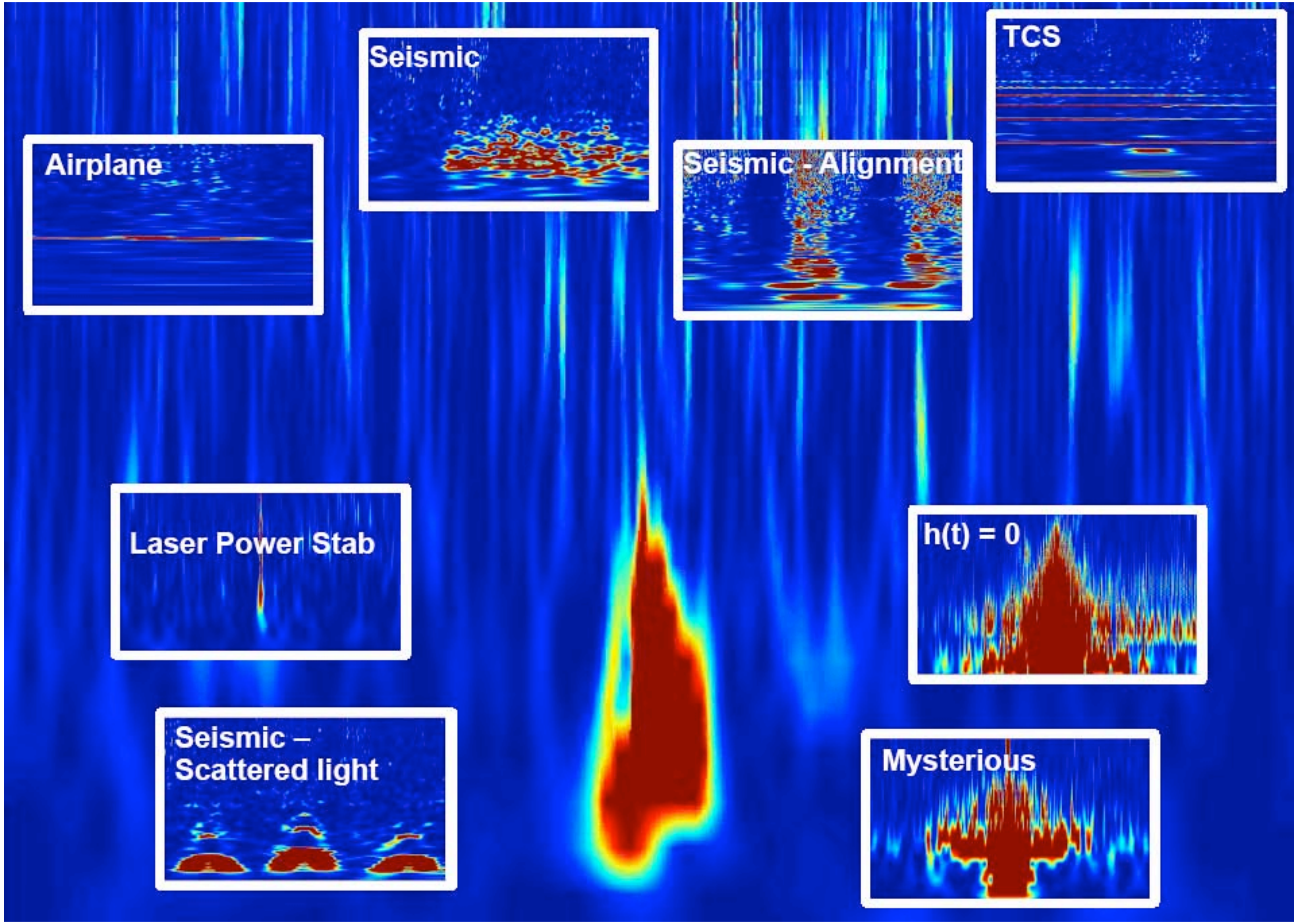
Collapsars



Dynamical capture NS mergers

Examples of poorly understood signals
Can't be used as templates

Things that go bump in the night



Template based analysis

These have the strongest priors and hence yield the most sensitive searches

$$p(\mathbf{h}) = \delta(\mathbf{h} - \mathbf{h}(\vec{\lambda}))p(\vec{\lambda})$$

Marginalizing over the individual signal samples yields the posterior on the model parameters

$$p(\vec{\lambda}|\mathbf{d}) = \int \frac{p(\mathbf{d}|\mathbf{h})p(\mathbf{h})}{p(\mathbf{d})}\, d\mathbf{h} = \frac{\overset{\text{Likelihood}}{p(\mathbf{d}|\vec{\lambda})}\,\overset{\text{Prior}}{p(\vec{\lambda})}}{\underset{\text{Evidence}}{p(\mathbf{d})}}$$

Techniques such as MCMC and Nested Sampling can be used to map out the full posterior distribution, allowing us to compute mean, median and mode and credible intervals. More on that in a bit.

But first we will look at using Maximum Likelihood (ML) and Maximum a Posteriori (MAP) estimators

Expansions about Maximum Likelihood and Maximum a Posteriori Points

Both the likelihood and the posterior can be expanded about (local) maxima. Usually done in terms of the logs

ML $\quad \partial_i \log p(\mathbf{d}|\vec{\lambda}) = 0$

MAP $\quad \partial_i \log p(\vec{\lambda}|\mathbf{d}) = \partial_i \log p(\mathbf{d}|\vec{\lambda}) + \partial_i \log p(\vec{\lambda}) = 0$

Fisher information matrix $\quad \Gamma_{ij} = -\mathbb{E}[\partial_i \partial_j \log p(\mathbf{d}|\vec{\lambda})|_{\mathrm{ML}}]$

"Augmented" Fisher information matrix $\quad K_{ij} = -\partial_i \partial_j \log p(\vec{\lambda}|\mathbf{d})|_{\mathrm{MAP}}$

Expansions about Maximum Likelihood and Maximum a Posteriori Points

$$p(\mathbf{d}|\vec{\lambda}) = \frac{1}{\sqrt{\det(2\pi\mathbf{\Gamma}^{-1})}}\, e^{-\frac{1}{2}\Gamma_{ij}(\lambda^i_{\rm ML}-\lambda^i)(\lambda^j_{\rm ML}-\lambda^j)}$$

$$p(\vec{\lambda}|\mathbf{d}) = \frac{1}{\sqrt{\det(2\pi\mathbf{K}^{-1})}}\, e^{-\frac{1}{2}K_{ij}(\lambda^i_{\rm MAP}-\lambda^i)(\lambda^j_{\rm MAP}-\lambda^j)}$$

The two distributions are very similar when the likelihood is more informative than the prior

Maximum Likelihood for Gaussian Noise

Typically the noise spectrum is assumed to be know and the normalization constant is ignored

$$\log p(\mathbf{d}|\vec{\lambda}) = -\frac{1}{2}(\mathbf{d} - \mathbf{h}(\vec{\lambda})|\mathbf{d} - \mathbf{h}(\vec{\lambda})) + \text{const.}$$

$$\text{Note, the first term is just } -\tfrac{1}{2} \text{ times } \chi^2 = (\mathbf{d} - \mathbf{h}|\mathbf{d} - \mathbf{h})$$

Expanding about the true signal: $\mathbf{h}(\vec{\lambda}) = \mathbf{h}_T + \partial_i \mathbf{h} \Delta\lambda^i + \ldots$

$$\log p(\mathbf{d}|\vec{\lambda}) = \text{const.} + (\mathbf{n}|\partial_i \mathbf{h})\Delta\lambda^i - \frac{1}{2}(\partial_i \mathbf{h}|\partial_j \mathbf{h})\Delta\lambda^i \Delta\lambda^j + \ldots$$

$$\partial_i \log p(\mathbf{d}|\vec{\lambda}) = 0 \qquad \Rightarrow \qquad (\partial_i \mathbf{h}|\partial_j \mathbf{h})\Delta\lambda^j = (\mathbf{n}|\partial_i \mathbf{h})$$

$$\Gamma_{ij} = -\mathbb{E}[\partial_i \partial_j \log p(\mathbf{d}|\vec{\lambda})|_{\text{ML}}] \qquad \Rightarrow \qquad \Gamma_{ij} = (\partial_i \mathbf{h}|\partial_j \mathbf{h})$$

Maximum Likelihood for Gaussian Noise

$$(\partial_i \mathbf{h}|\partial_j \mathbf{h})\Delta\lambda^j = (\mathbf{n}|\partial_i \mathbf{h})$$

$$\Gamma_{ij} = (\partial_i \mathbf{h}|\partial_j \mathbf{h})$$

$$\Rightarrow \quad \boxed{\Delta\lambda^i = (\mathbf{n}|\partial_j \mathbf{h})\,\Gamma^{ij}}$$

Noise covariance matrix

$$C^{ij} = \mathbb{E}[\Delta\lambda^i \Delta\lambda^j] = \mathbb{E}[(\mathbf{n}|\partial_k \mathbf{h})(\mathbf{n}|\partial_j \mathbf{h})]\Gamma^{ik}\Gamma^{jl}$$

$$= \Gamma_{kj}\Gamma^{ik}\Gamma^{jl} = \Gamma^{ij}$$

To leading order in the SNR, the noise covariance matrix is given by the inverse of the Fisher matrix

$$\boxed{C^{ij} = \Gamma^{ij}}$$

Simple Example

n = stationary, white Gaussian noise with unit variance (can work in time domain)

$$h(t) = A\cos(2\pi f t + \phi)$$ Use parameters $\{\ln A, \ln f, \phi\}$ $(\omega = 2\pi f)$

$$\mathrm{SNR}^2 = \int_0^T h(t)^2 dt = \frac{1}{2} A^2 T$$

$$\Gamma_{ij} = \mathrm{SNR}^2 \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{(\omega T)^2}{3} & \frac{\omega T}{2} \\ 0 & \frac{\omega T}{2} & 1 \end{pmatrix} \qquad C^{ij} \simeq \frac{1}{\mathrm{SNR}^2} \begin{pmatrix} 1 & \frac{6}{(\omega T)^2} & \frac{3}{(\omega T)} \\ \frac{6}{(\omega T)^2} & \frac{12}{(\omega T)^2} & -\frac{6}{(\omega T)} \\ \frac{3}{(\omega T)} & -\frac{6}{(\omega T)} & 4 \end{pmatrix}$$

(1) Parameter errors scale as 1/SNR

(2) Errors in $\ln f$ and ϕ are inflated due to correlation

(3) Error in $\ln f$ is tiny compared to the other parameters

Eigenvectors, Eigenvalues of Correlation Matrix

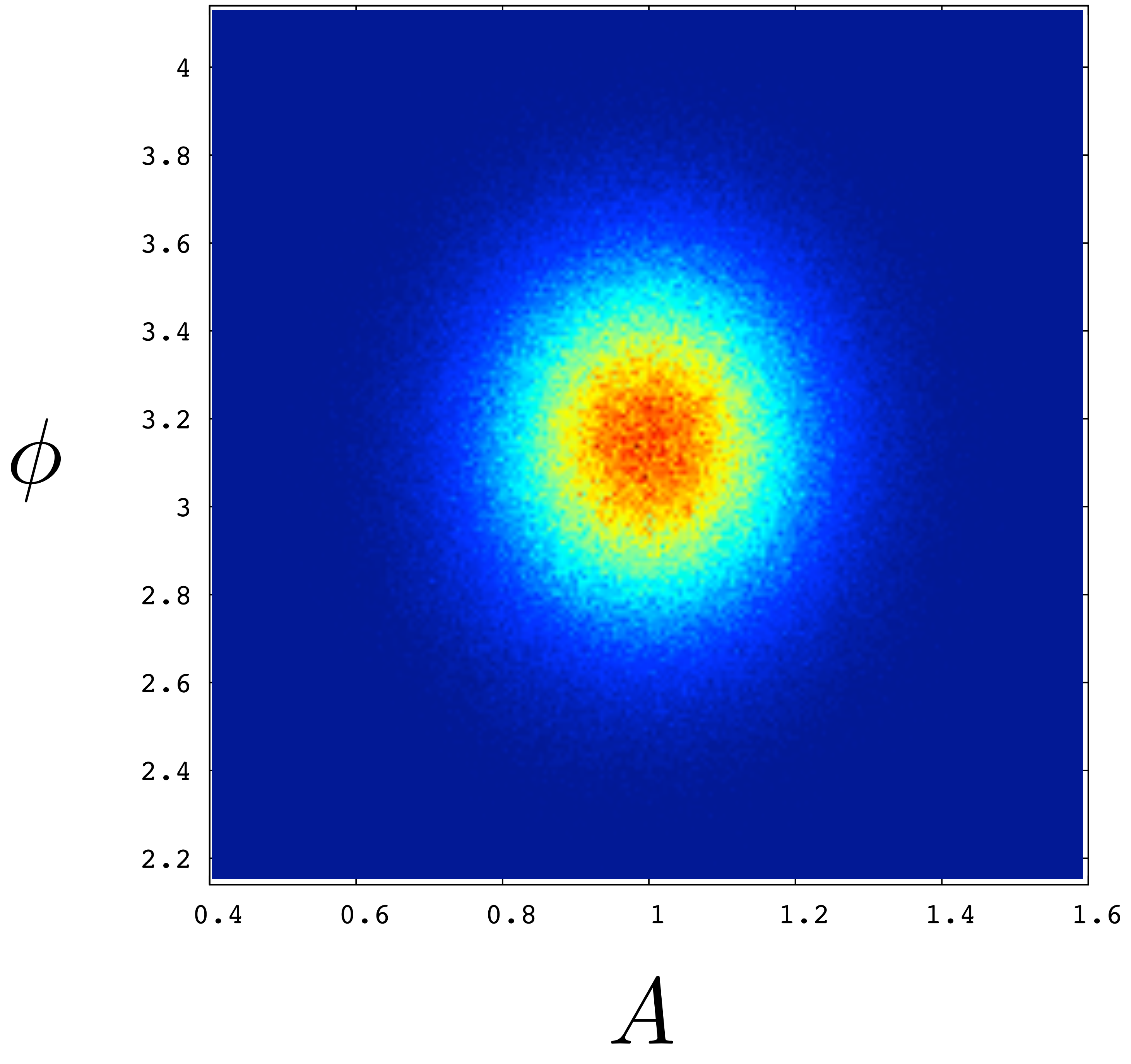
Parameters $\{\ln A, \ln f, \phi\}$

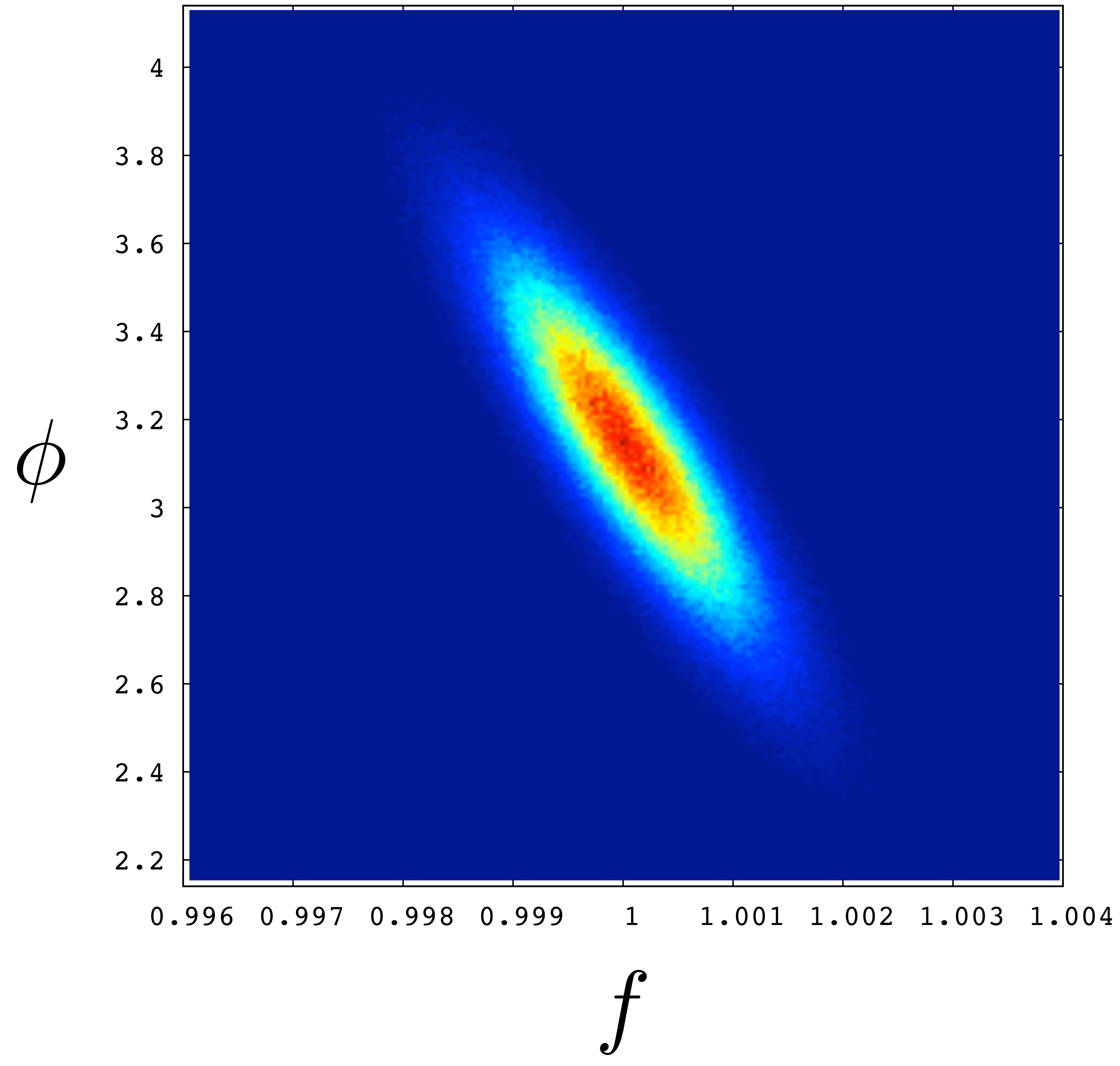
In the limit of many cycles $\omega T \gg 1$

Eigenvalues	Eigenvectors
$\frac{4}{\mathrm{SNR}^2}$	$\left\{0, \frac{-1}{\sqrt{37}}, \frac{6}{\sqrt{37}}\right\}$
$\frac{1}{\mathrm{SNR}^2}$	$\{1, 0, 0\}$
$\frac{3}{(\omega T)^2\,\mathrm{SNR}^2}$	$\{0, 1, 0\}$

Simple Example

$$\frac{\Delta A}{A} = \frac{1}{\mathrm{SNR}} \qquad \Delta\phi = \frac{2}{\mathrm{SNR}} \qquad \frac{\Delta f}{f} = \frac{\sqrt{12}}{\mathrm{SNR}} \frac{1}{(\omega T)} \quad \Rightarrow \quad \Delta f \approx \left(\frac{0.55}{\mathrm{SNR}}\right)\left(\frac{1}{T}\right)$$





$$p(\mathbf{d}|\vec{\lambda}) = \frac{1}{\sqrt{\det(2\pi\mathbf{\Gamma}^{-1})}} e^{-\frac{1}{2}\Gamma_{ij}(\lambda^i_{\mathrm{ML}}-\lambda^i)(\lambda^j_{\mathrm{ML}}-\lambda^j)}$$

Bayesian Parameter Estimation

Reverend Thomas Bayes

Sir Ronald Fisher

"there is a valid defense for using non-Bayesian methods, namely incompetence." - Skilling

Bayesian Inference

- Bayesian Probability Theory
- Bayesian Learning
- Model Selection (theory)
- Bayesian versus Frequentist
- Markov Chain Monte Carlo
- Return to our simple example
- Model Selection (implementation)

Bayesian Probability Theory

Degree of belief interpretation of probability - the natural expression of the scientific method (as opposed to the more binary view of Karl Popper that heavily influenced Fisher and other frequentists)

Initial Understanding	$\Rightarrow$	New Observations	$\Rightarrow$	Updated Understanding
$p(\vec{x})$		$p(d\|\vec{x})$		$p(\vec{x}\|d)$
Prior	$\Rightarrow$	Likelihood	$\Rightarrow$	Posterior

Bayes' Theorem $$p(\vec{x}|d) = \frac{p(\vec{x})p(d|\vec{x})}{p(d)}$$

Normalization factor is the marginal likelihood or ***evidence*** $$p(d) = \int p(\vec{x})p(d|\vec{x})\, d\vec{x}$$

Bayesian Probability Theory

The posterior distribution fully characterizes the model.

E.g. expectation values

$$E[x^i] = \int x^i \, p(\vec{x}|d) \, d\vec{x}$$

E.g. single parameter probability distributions

$$p(x^i|d) = \int p(\vec{x}|d) \, dx^1 dx^2 \ldots dx^{i-1} dx^{i+1} \ldots dx^D$$

E.g. quantile regions, such as 90%

$$\begin{aligned} 0.05 &= \int^{x_1} p(x|d) \, dx \\ 0.9 &= \int_{x_1}^{x_2} p(x|d) \, dx \end{aligned}$$

Bayesian Learning

"The (Bayesian) theory of probabilities is basically just common sense reduced to calculus" - Laplace

"Today's posterior is tomorrow's prior" - Lindley

The amount we learn from the data can be measured in bits, and can be computed in terms of the Kullback–Leibler divergence

$$D_{KL} = \int p(\vec{x}|d) \, \log_2 \left(\frac{p(\vec{x}|d)}{p(\vec{x})} \right) \, d\vec{x} \quad [\text{bits}]$$

Bayesian Model Selection

Probability of Model M: $p(M|d) \propto p(M)p(d|M)$

Prior Probability of M — Evidence for M

Odds Ratio:
$$
\begin{aligned}
O_{ij} &= \frac{p(M_i|d)}{p(M_j|d)} \\
&= \frac{p(M_i)}{p(M_j)} \frac{p(d|M_i)}{p(d|M_j)} \\
&= \text{Prior Odds Ratio} \times \text{Bayes Factor}
\end{aligned}
$$

More on how we compute the Bayes Factor later...

Bayesian Machinery: Markov Chain Monte Carlo

Bayes' Theorem $$p(\vec{x}|d) = \frac{p(\vec{x})p(d|\vec{x})}{p(d)}$$

Marginal likelihood or evidence $$p(d) = \int p(\vec{x})p(d|\vec{x})\, d\vec{x}$$

We know how to compute the prior and the likelihood. The difficulty lies in computing the evidence.

The MCMC technique, introduced by Metropolis and developed by Hastings, allows us to simulated samples from the posterior distribution directly, without having to compute the evidence.

It is possible to compute the evidence using augmented MCMC techniques - more on the later. Another powerful technique for computing the evidence and the posterior distributions is *Nested Sampling*

Bayesian Inference

Prior

$p(h|M)$

Likelihood

$p(d|h)$

$p(h|d, M)$

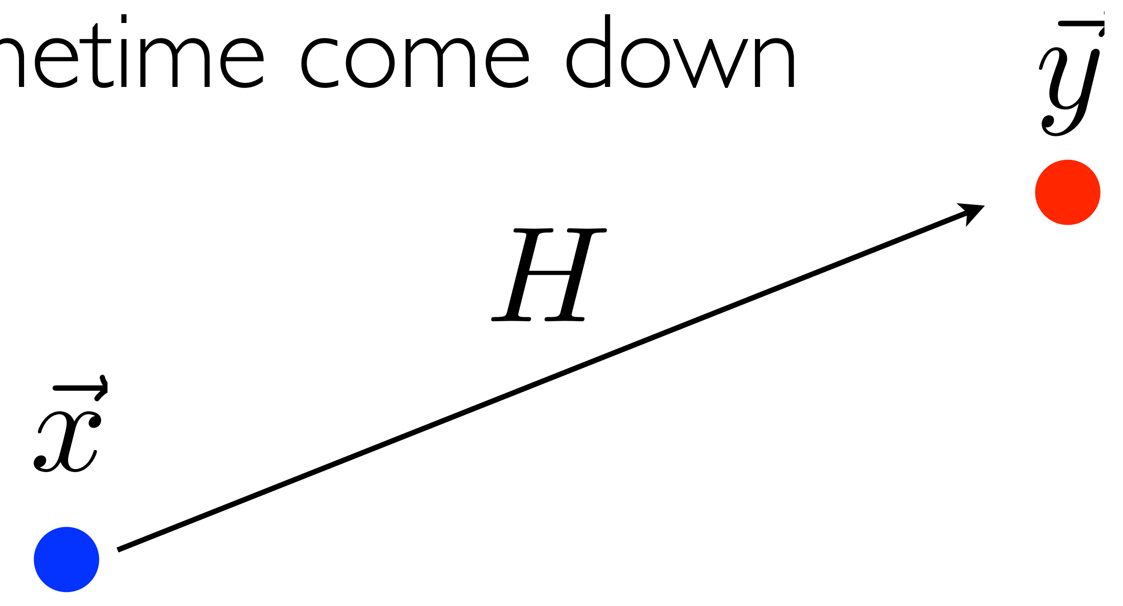
Posterior

$p(d|M)$

Evidence

Markov Chain Monte Carlo

Always go up,
Sometime come down



Yields PDF $p(\vec{x}|d)$ for parameters $\vec{x}$ given data d

$$H = \min\left(1, \underbrace{\frac{p(\vec{y})}{p(\vec{x})}}_{\text{Prior}} \underbrace{\frac{p(d|\vec{y})}{p(d|\vec{x})}}_{\text{Likelihood}} \underbrace{\frac{q(\vec{x}|\vec{y})}{q(\vec{y}|\vec{x})}}_{\text{Proposal}}\right)$$

Transition Probability
(Metropolis-Hastings)

Markov Chain Monte Carlo

The choice of jump proposal $q(\vec{y}|\vec{x})$ is key to convergence

Convergence to the target distribution has two facets:

- Burn-in (finding the dominant modes of the posterior)

- Mixing (exploring the dominant modes of the posterior)

The perfect proposal distribution is the posterior distribution itself, $q(\vec{y}|\vec{x}) = p(\vec{y}|d)$, since then

$$\begin{aligned} H &= \min\left(1, \frac{p(\vec{y})p(d|\vec{y})p(\vec{x}|d)}{p(\vec{x})p(d|\vec{x})p(\vec{y}|d)}\right) \\ &= 1 \end{aligned}$$

But if we knew the posterior distribution in advance there would be no need for the MCMC procedure! Instead we seek ways to approximate the posterior.

My favorite MCMC Recipe

Ingredients:

Fisher matrix proposals

Scouting proposals

Differential evolution proposals

Parallel tempering

Directions:

Mix all the proposals together. Check consistency by recovering the prior and diagonal PP plots. Results are ready when distributions are stationary.

Proposal Distributions

Fisher matrix proposals

Quadratic approximation to the posterior using the augmented Fisher Information Matrix

$$q(\vec{y}|\vec{x}) = \frac{1}{\sqrt{\det(2\pi \mathbf{K}^{-1})}} e^{-\frac{1}{2} K_{ij} (x^i - y^i)(x^j - y^j)}$$

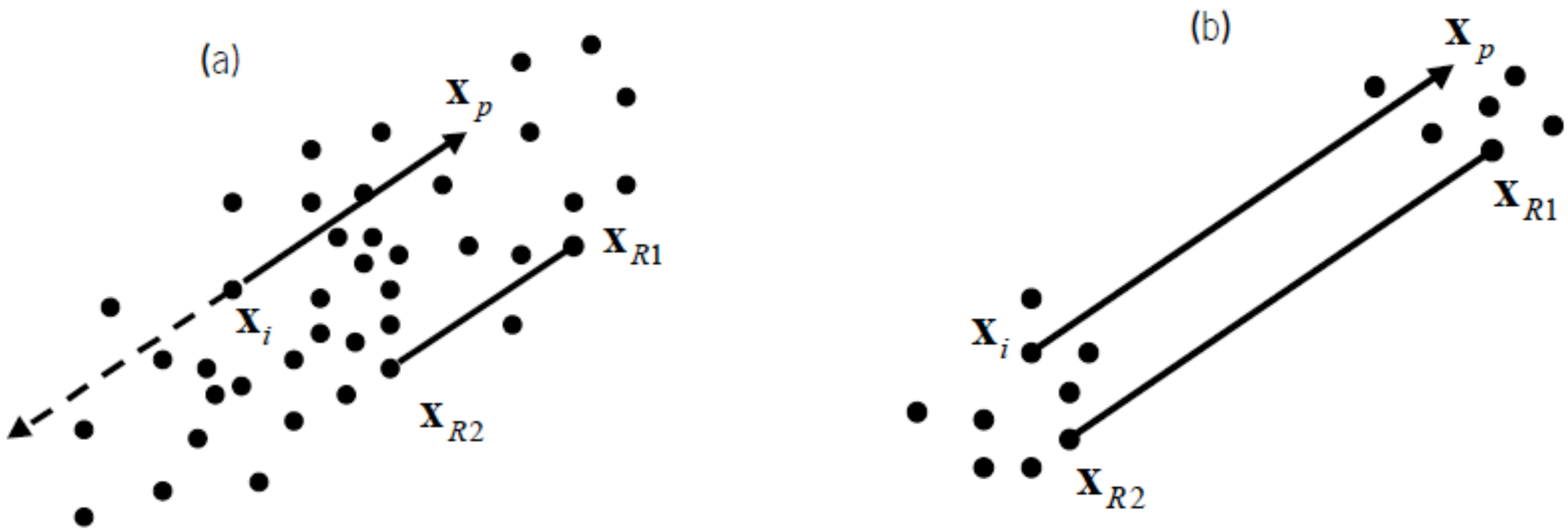
Propose jumps along eigendirections of **K**, scaled by eigenvalues

Scouting proposals

Use a Non-Markovian Pilot search (hill climbers, simulated annealing, genetic algorithms etc) to crudely map the posterior/likelihood and use this as a proposal distribution for a Markovian follow-up [Littenberg & Cornish, PRD 80, 063007, (2009)]

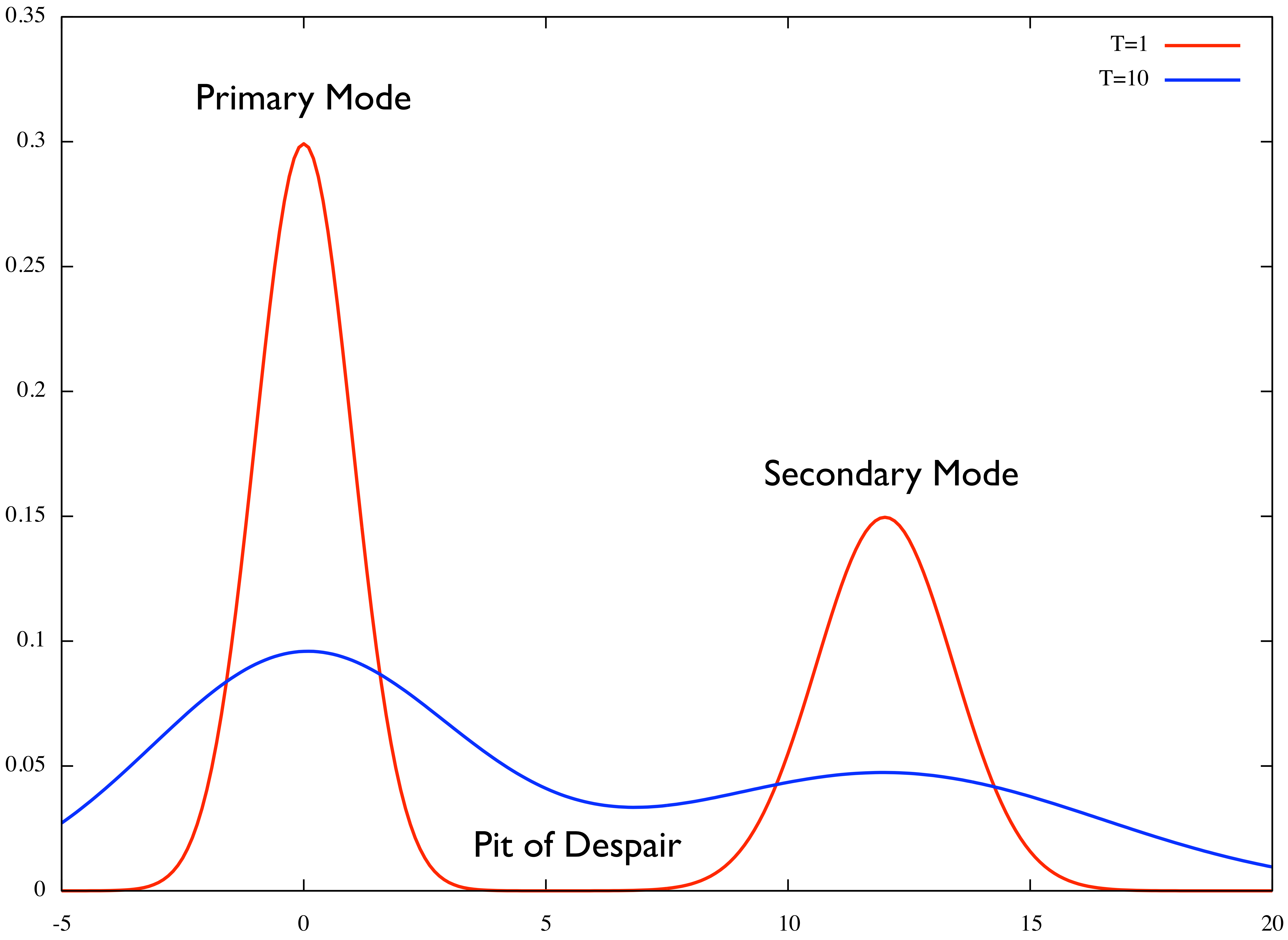
Proposal Distributions

Differential evolution [Braak (2005)]



Parallel Tempering

[Swendsen & Wang, 1986]



Ordinary MCMC techniques side-step the need to compute the evidence. PT uses multiple, coupled chains to improve mixing, and also allows the evidence to be computed.

Explore tempered posterior

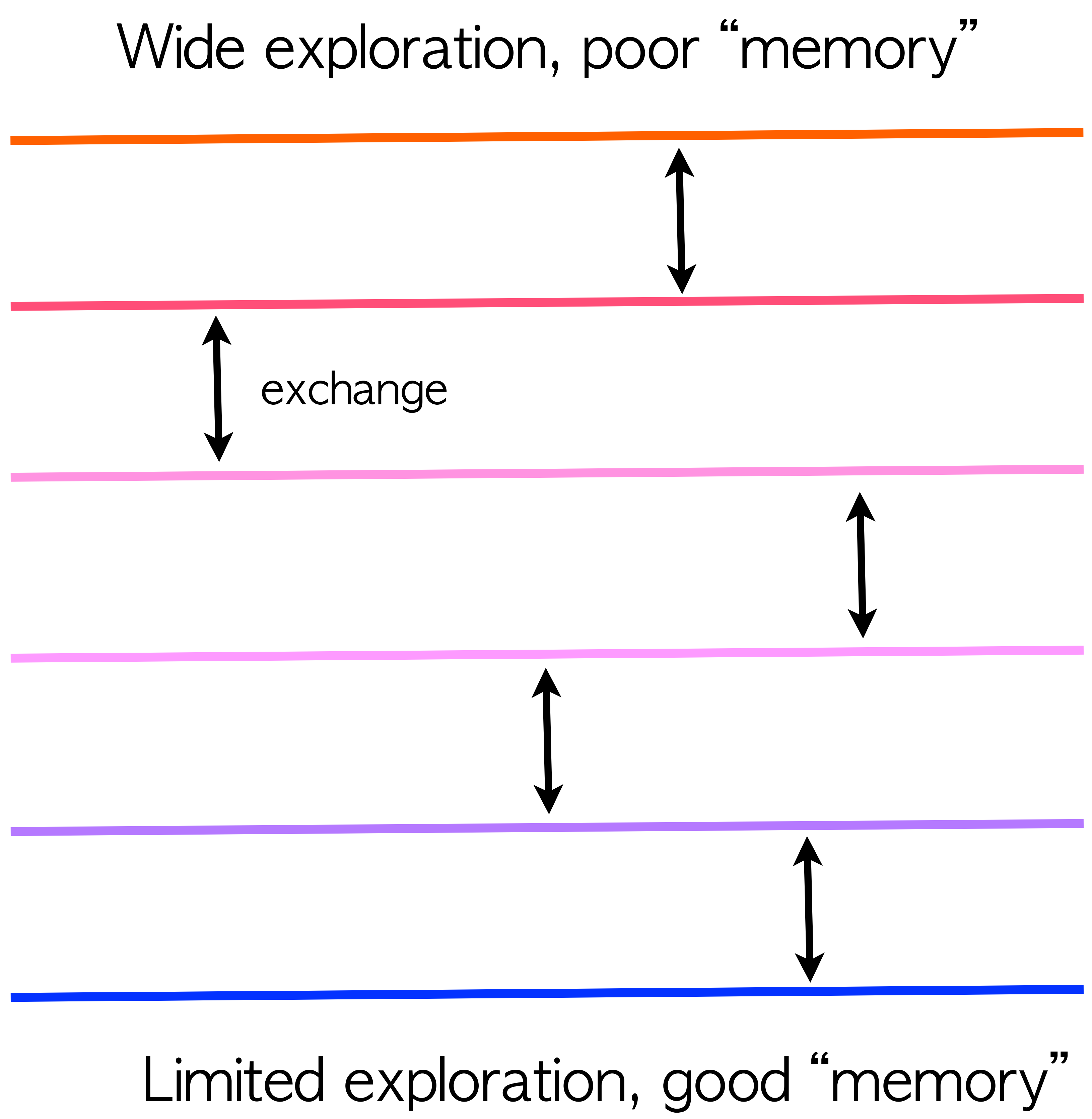
$$\pi(\vec{\lambda}|\mathbf{d})_T = p(\mathbf{d}|\vec{\lambda})^{1/T} p(\vec{\lambda})$$

Compute model evidence

$$\log p(\mathbf{d}) = \int_0^1 \mathbb{E}[\log p(\mathbf{d}|\vec{\lambda})]_\beta \, d\beta$$

(Here $\beta = \frac{1}{T}$)

Parallel Tempering



Inter-chain transition probability

$$H = \min\left(1, \frac{\pi(\vec{\lambda}_{j+1}|\mathbf{d})_{T_j}\,\pi(\vec{\lambda}_j|\mathbf{d})_{T_{j+1}}}{\pi(\vec{\lambda}_j|\mathbf{d})_{T_j}\,\pi(\vec{\lambda}_{j+1}|\mathbf{d})_{T_{j+1}}}\right)$$

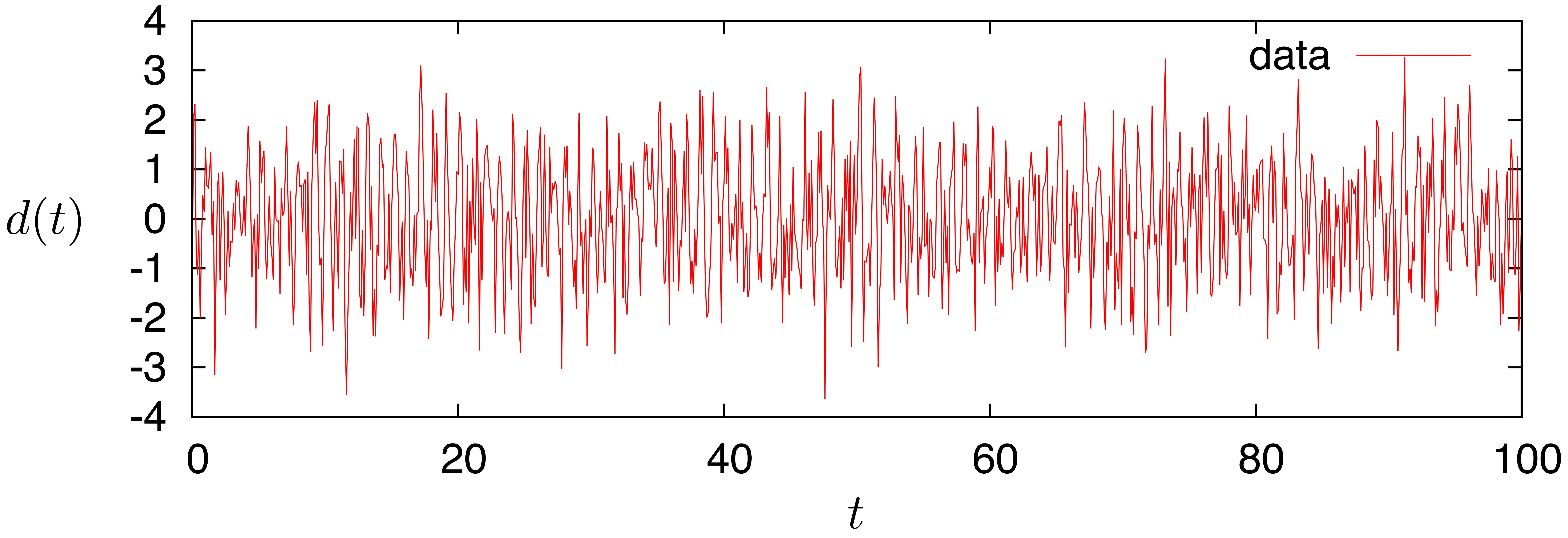
Good solutions sink down ladder and get stored

Simple Example: Redux

n = stationary, white Gaussian noise with unit variance

$h(t) = A\cos(2\pi f t + \phi)$

Parameters $\{\ln A, \ln f, \phi\}$

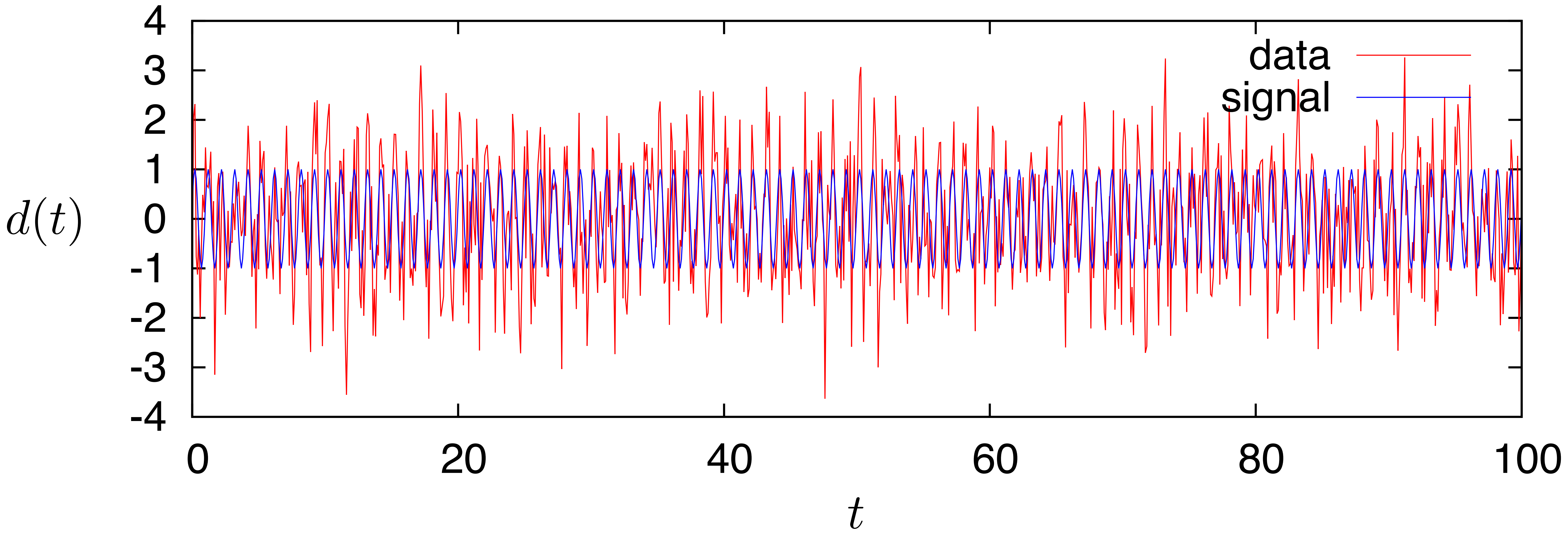


$(\mathrm{SNR} = 7.07)$

Simple Example: Redux

n = stationary, white Gaussian noise with unit variance

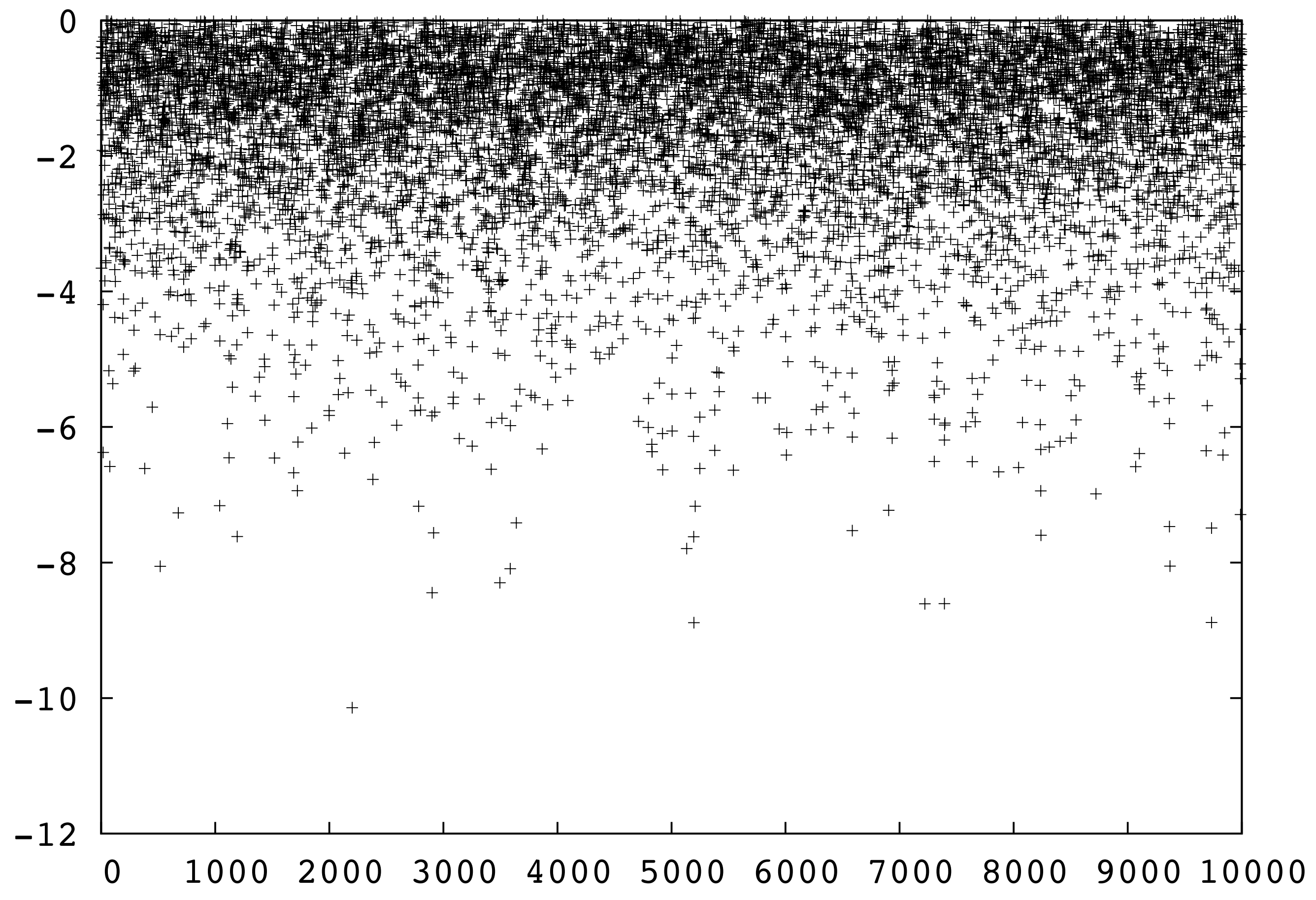
$h(t) = A\cos(2\pi f t + \phi)$ Parameters $\{\ln A, \ln f, \phi\}$



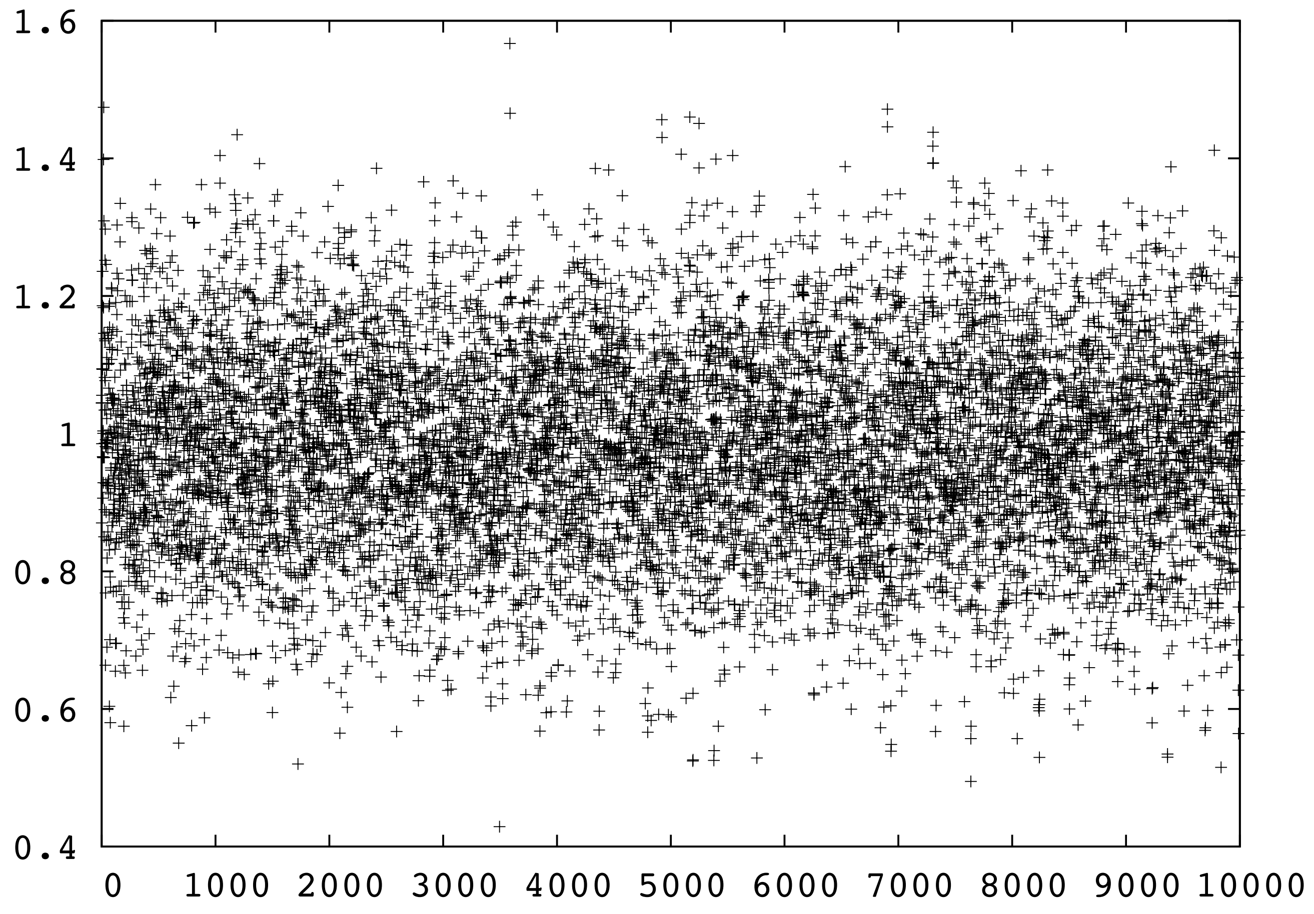
$(\mathrm{SNR} = 7.07)$

Markov Chains

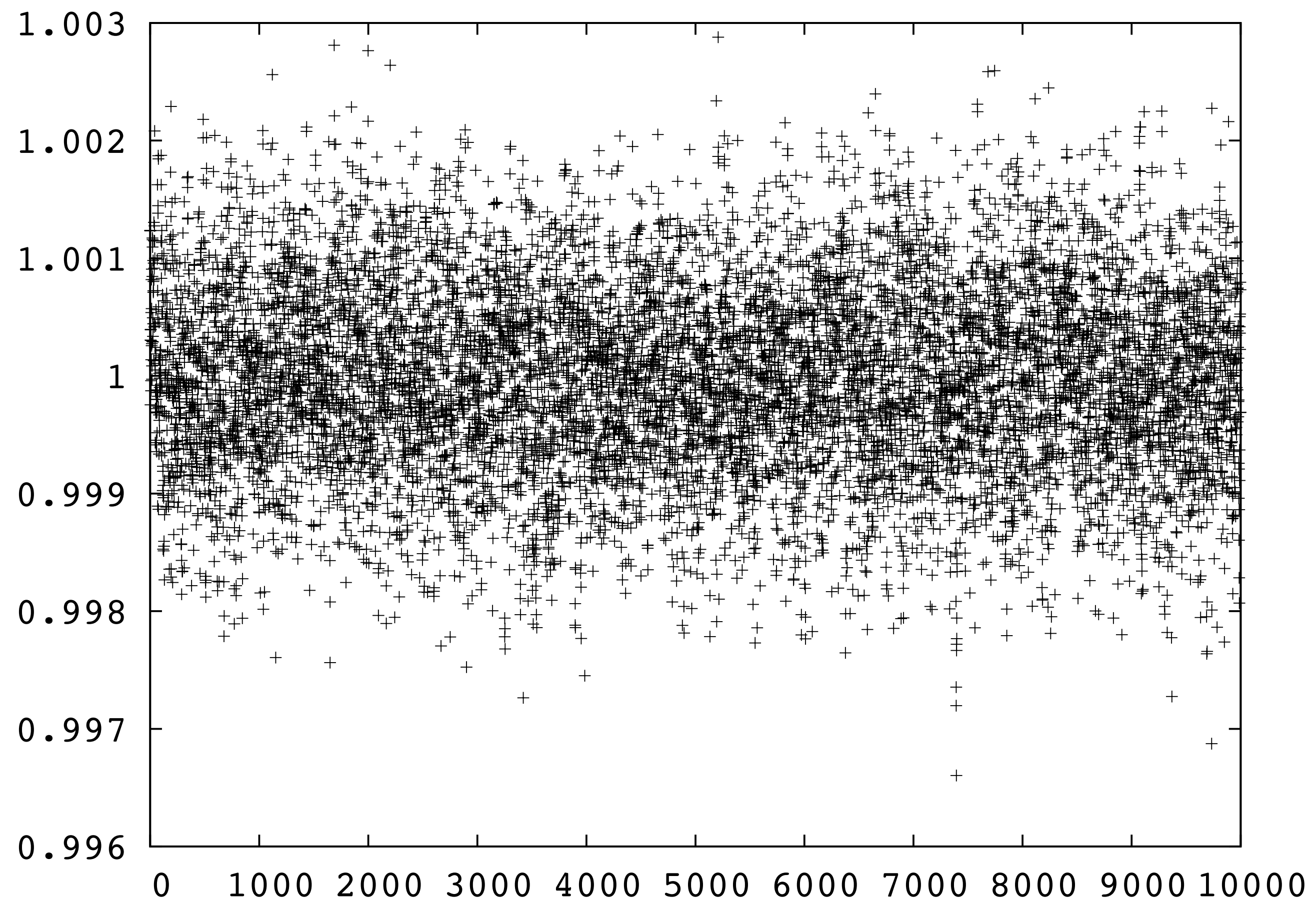
$\ln p(d|\vec{x})$



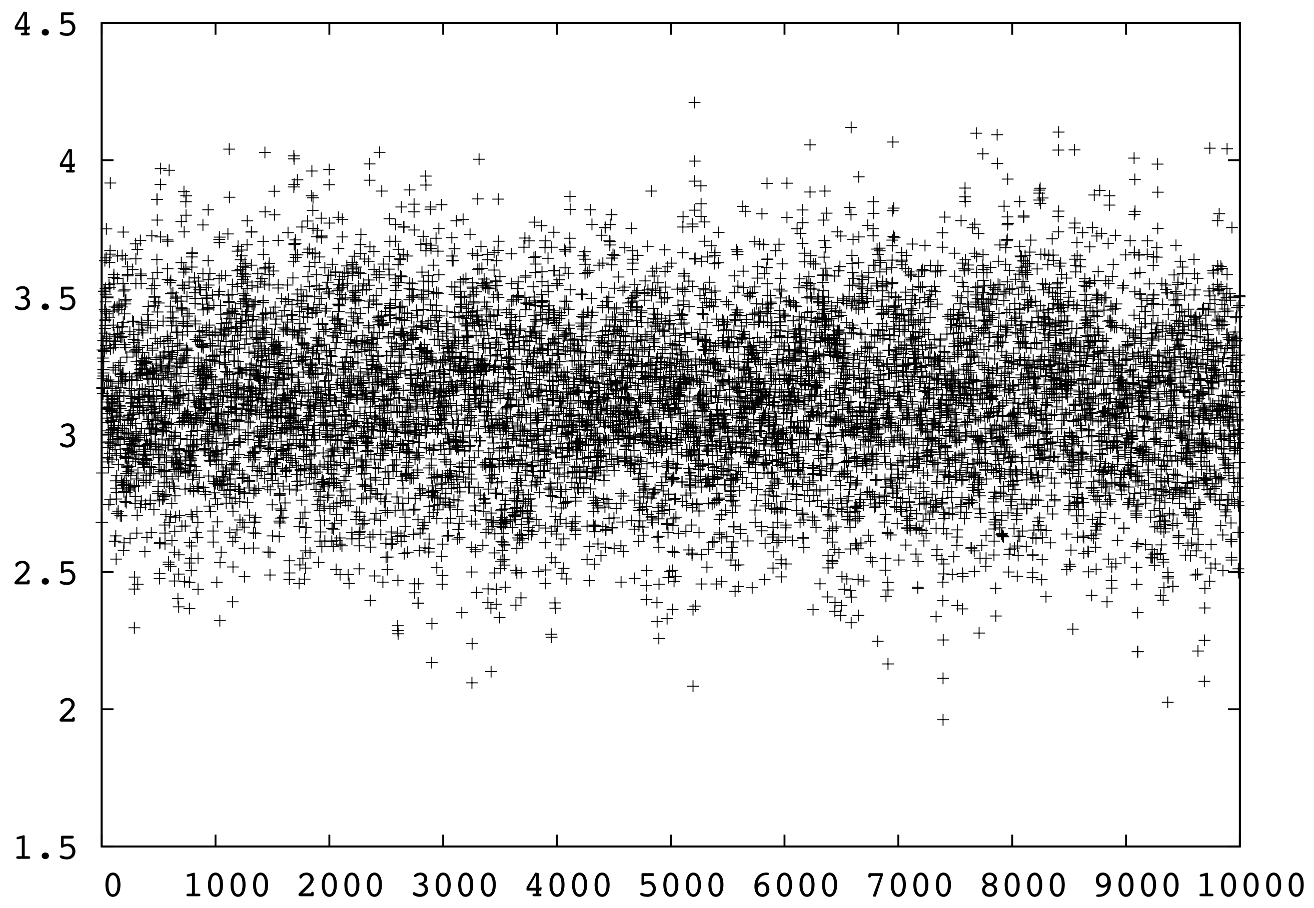
A



f

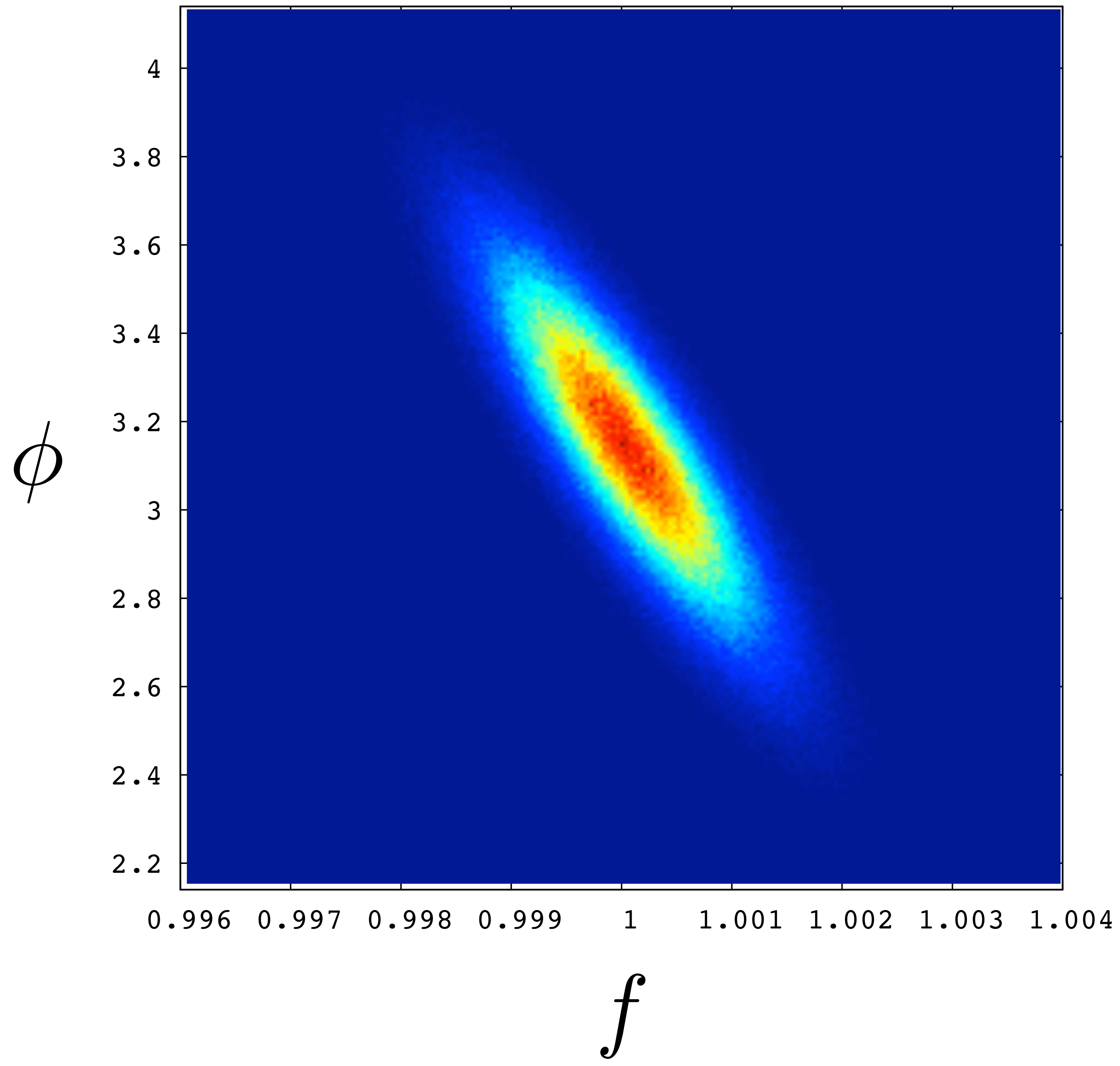


ϕ

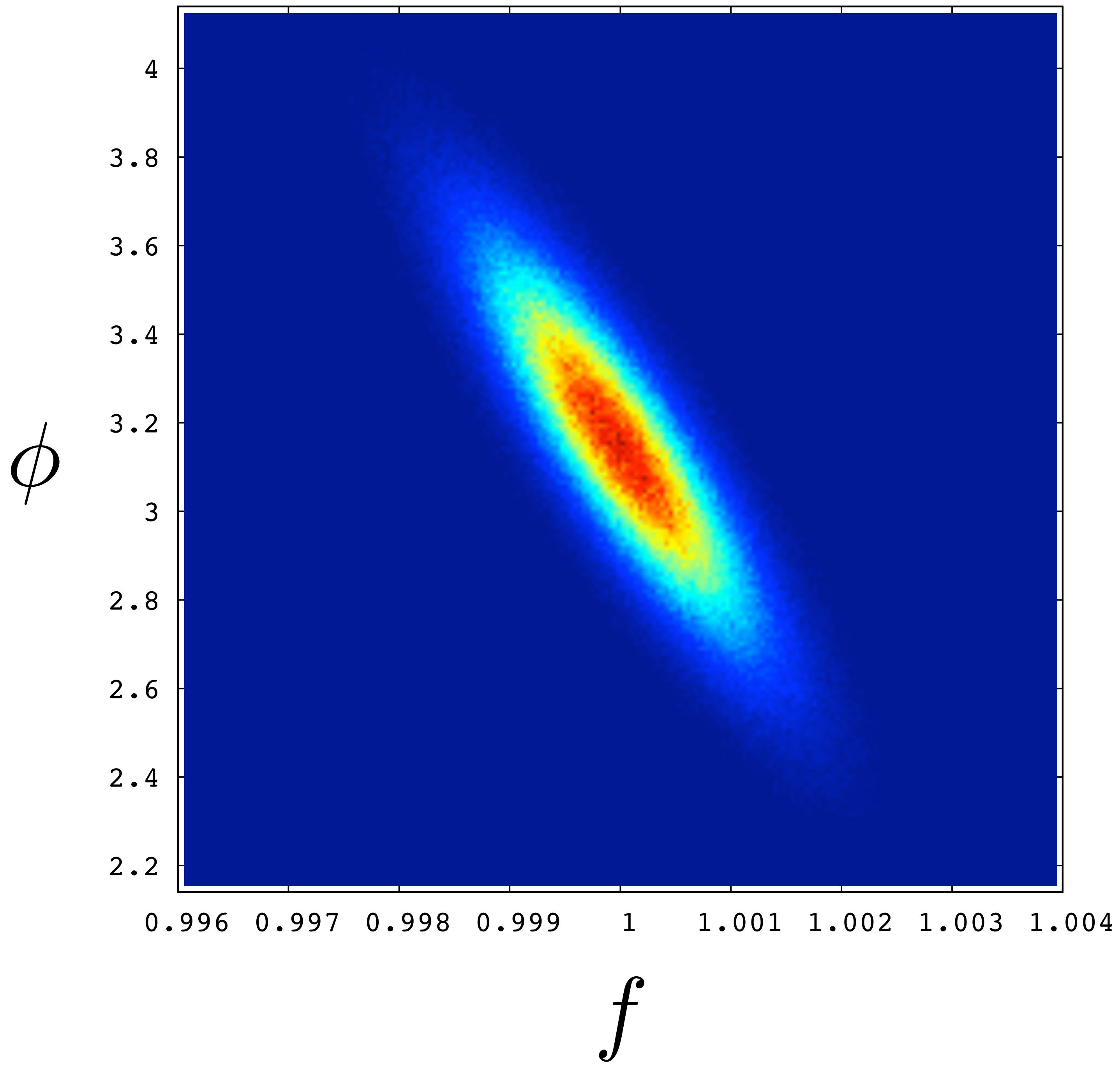


$(\mathrm{SNR} = 7.07)$

Fisher vs MCMC



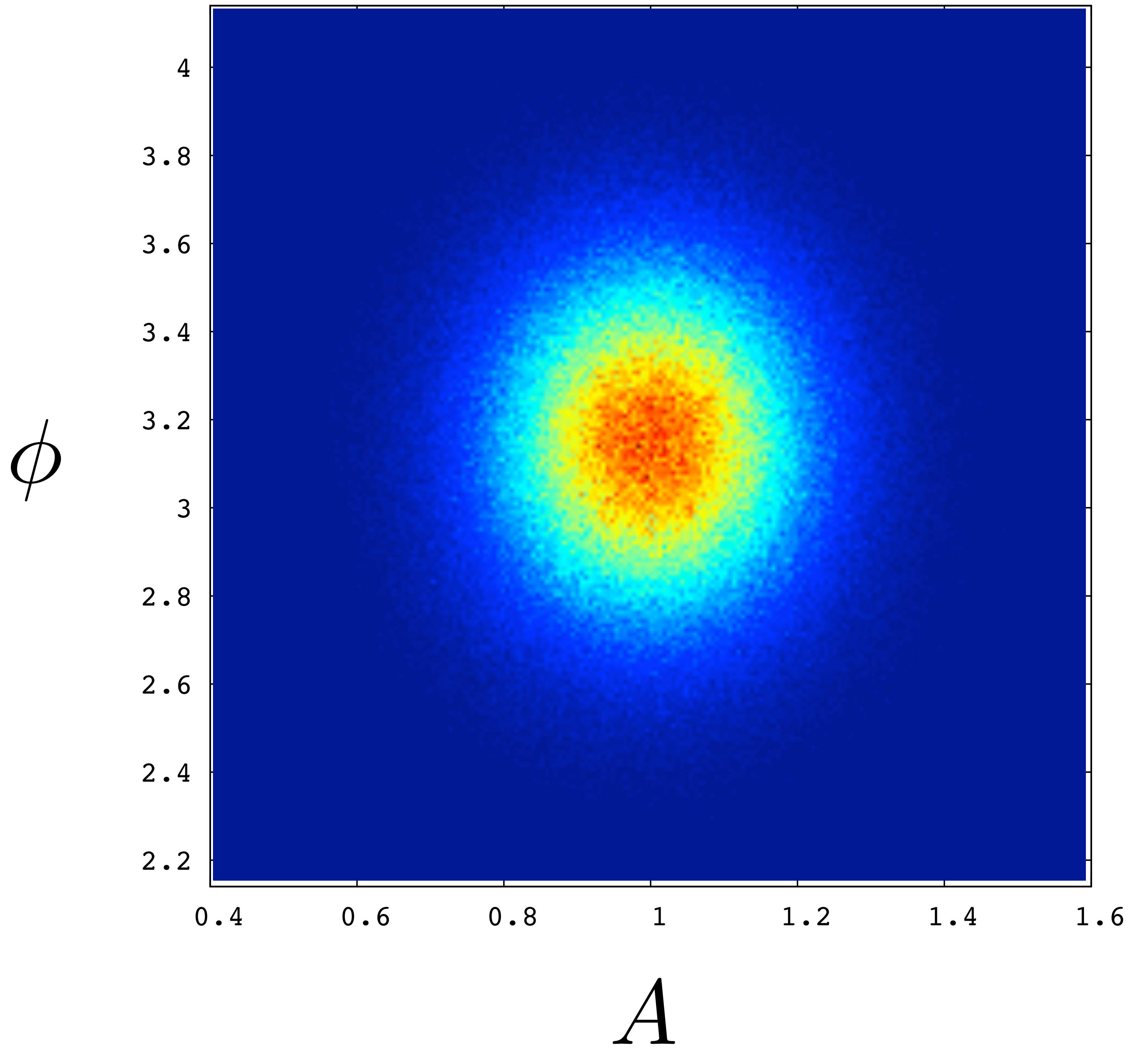
Fisher

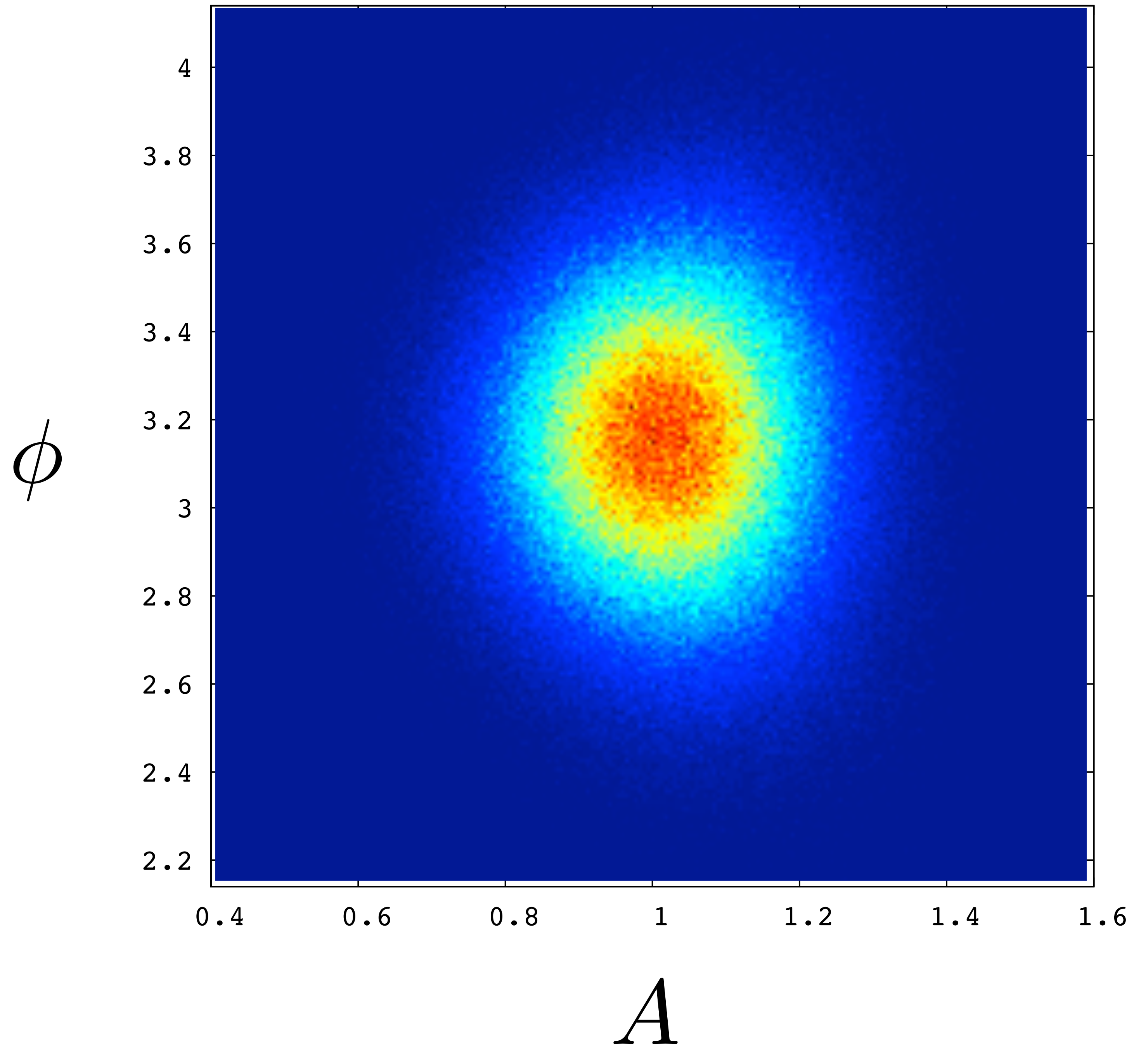


MCMC

$(\mathrm{SNR} = 7.07)$

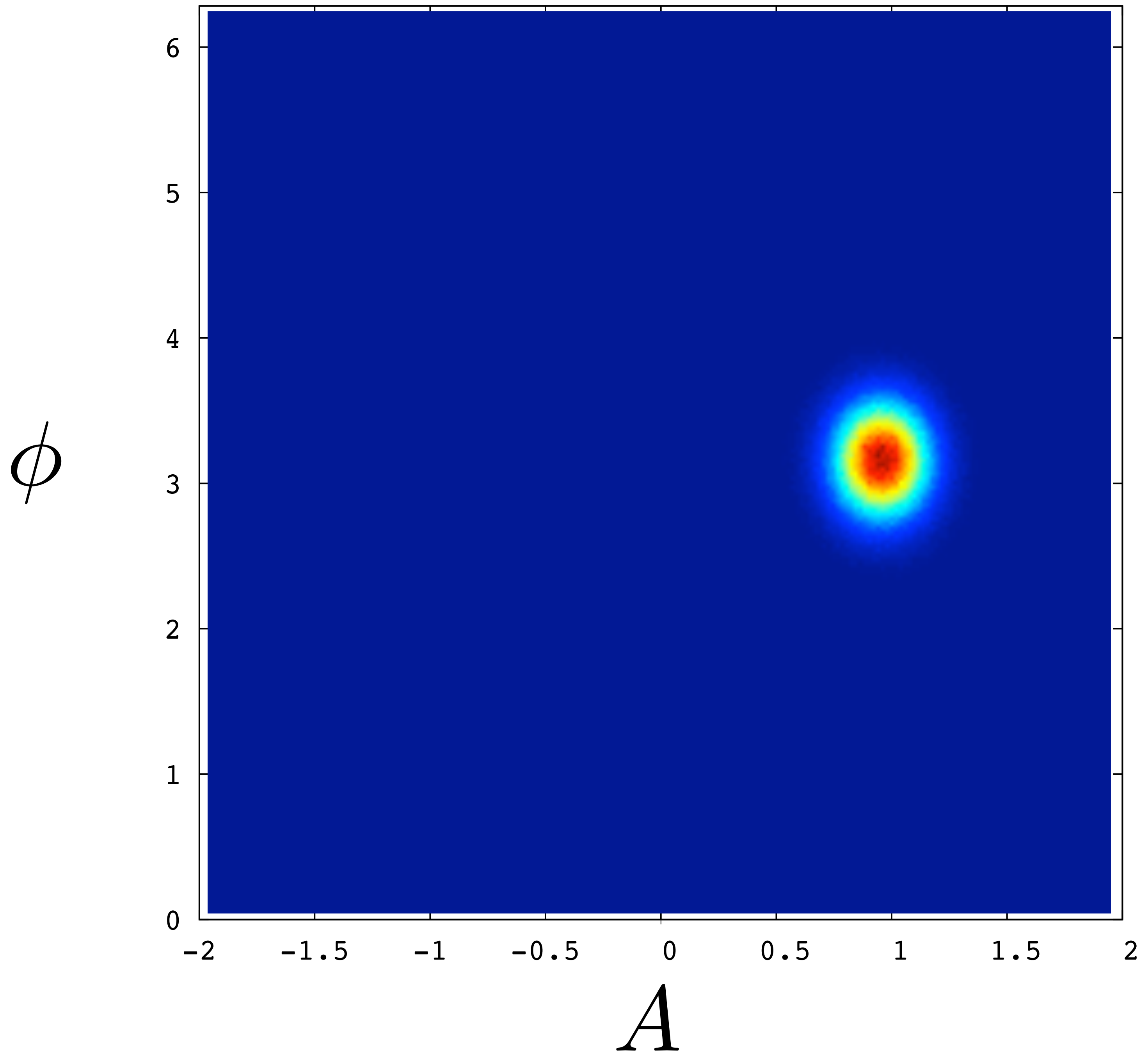
Fisher vs MCMC

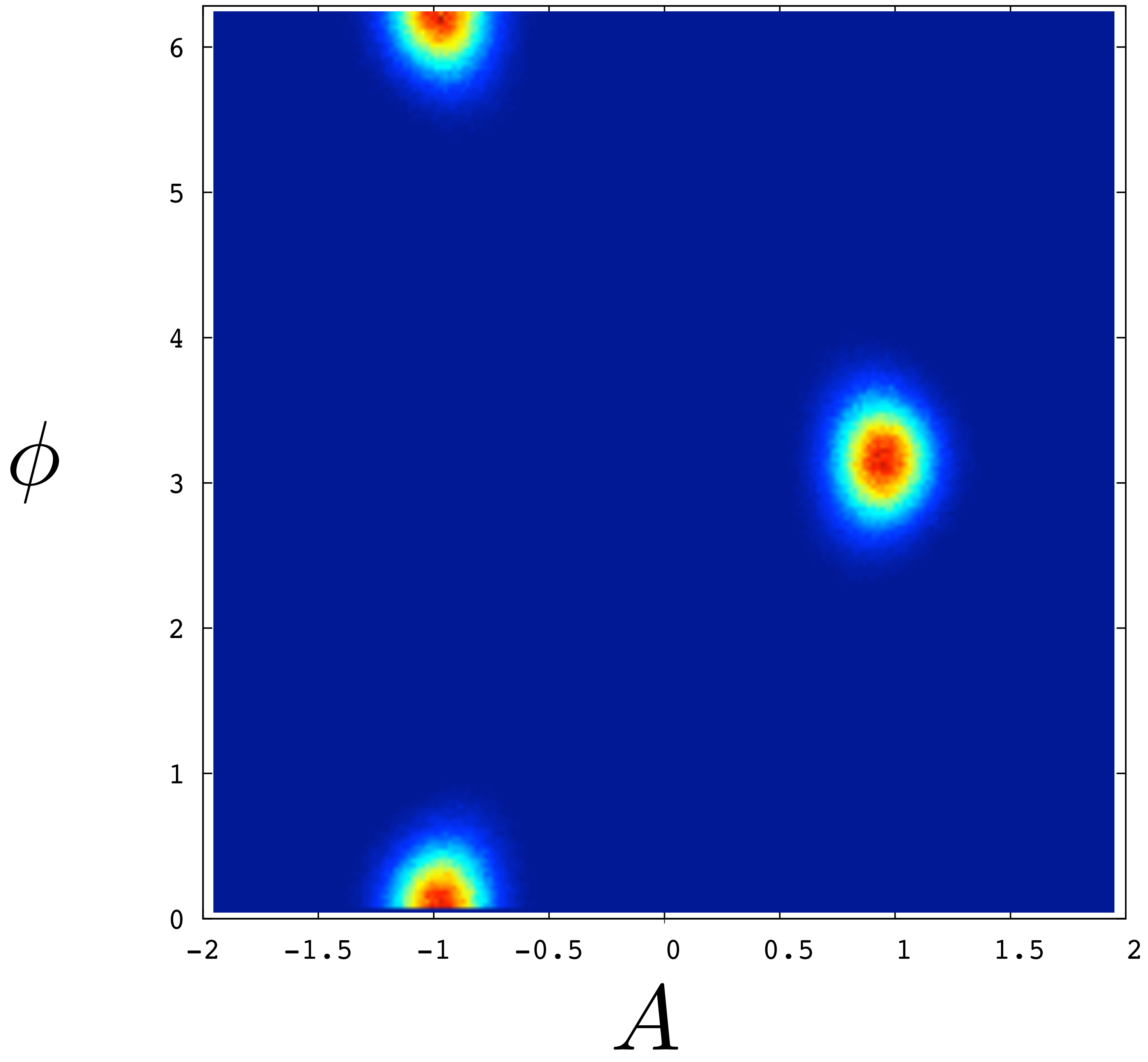




Fisher $(\mathrm{SNR} = 7.07)$ MCMC

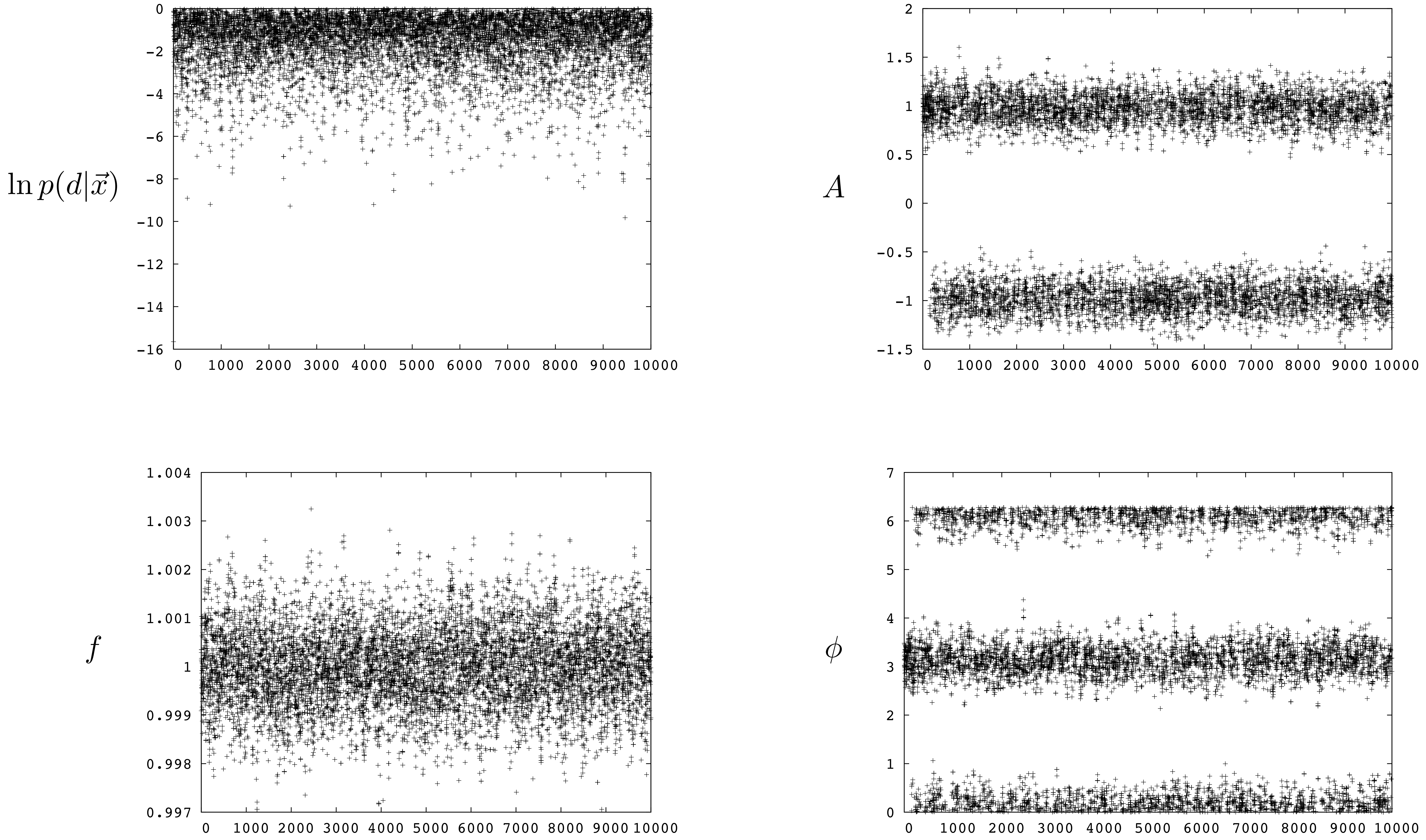
Fisher vs MCMC





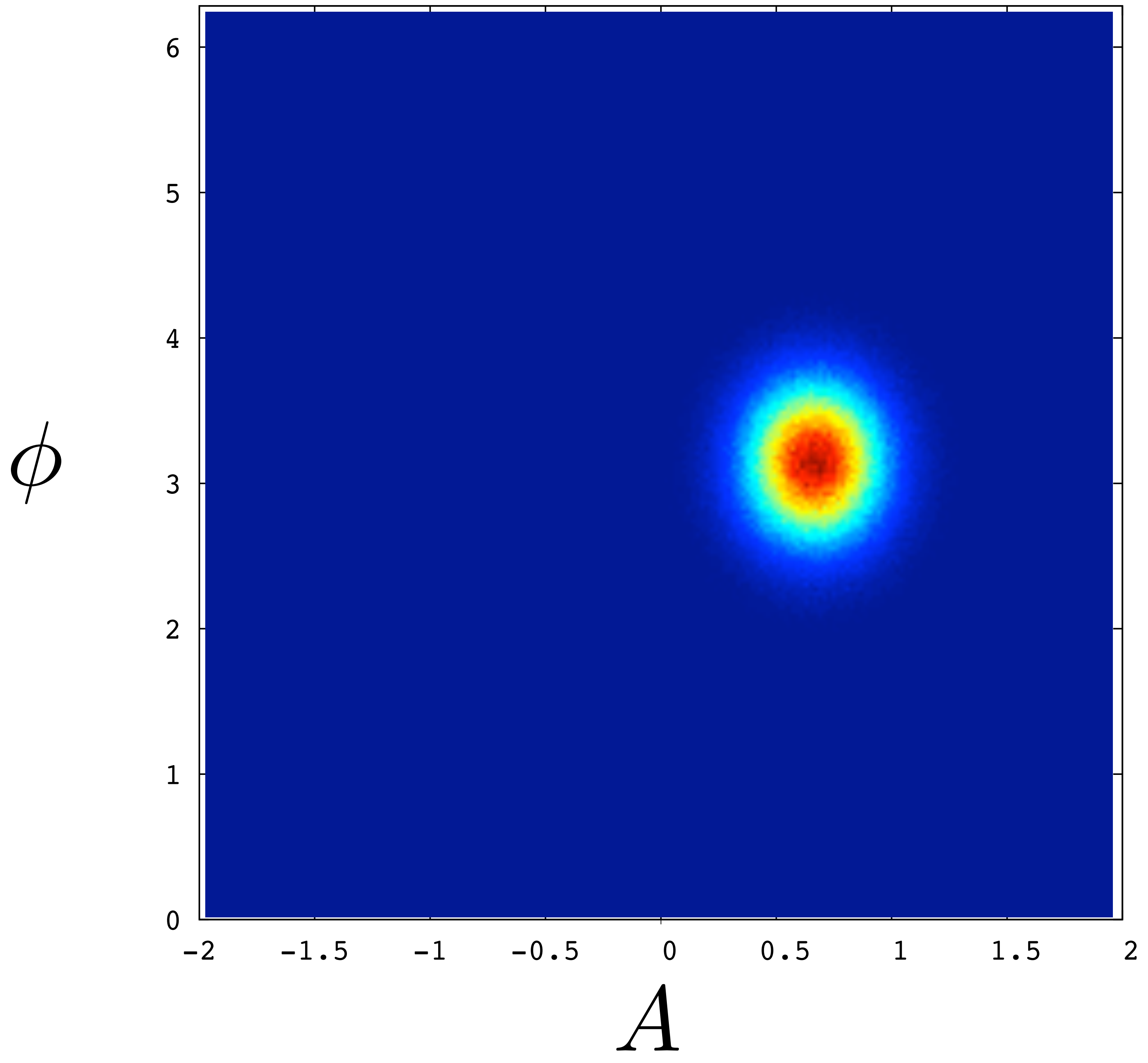
Fisher $(\mathrm{SNR} = 7.07)$ MCMC

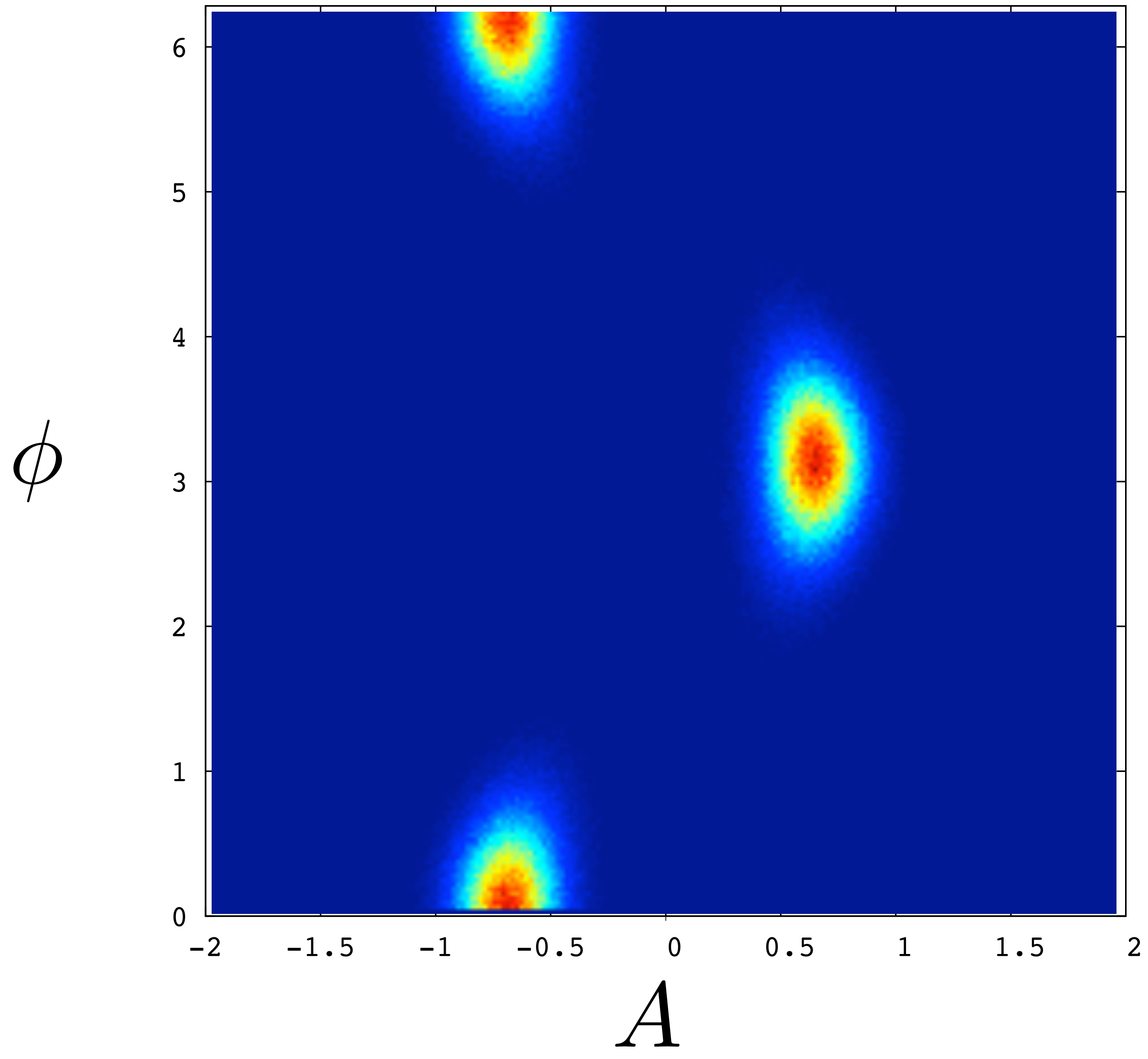
Markov Chains



(SNR = 7.07)

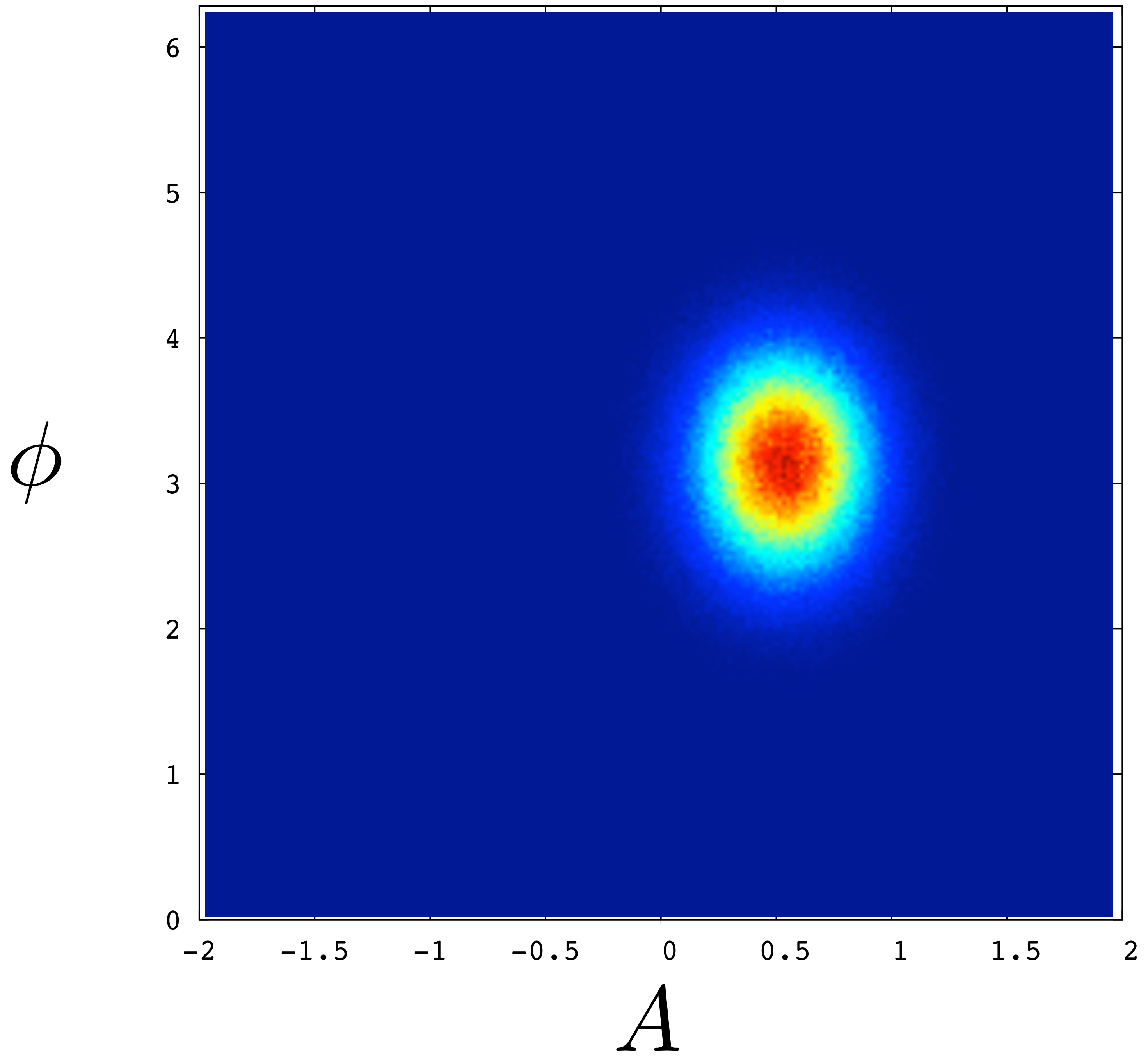
Fisher vs MCMC

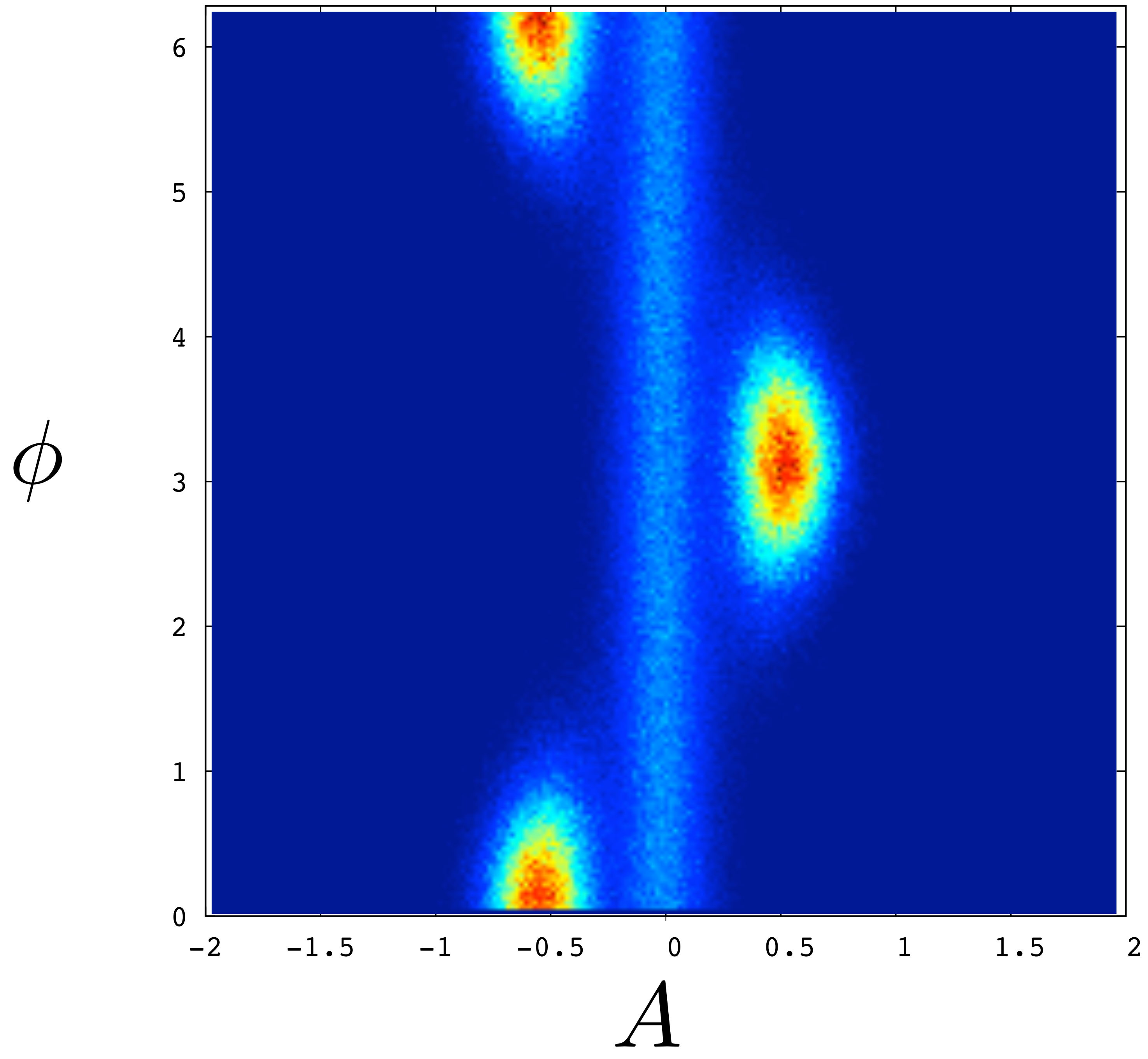




Fisher $(\mathrm{SNR} = 4.95)$ MCMC

Fisher vs MCMC





Fisher $(\mathrm{SNR} = 4.0)$ MCMC

Simple Example: Extended

n = stationary, white Gaussian noise with unit variance

$h(t) = A\cos(2\pi f(t - t_0) + \phi)$ Parameters $\{A, f, \phi, t_0\}$

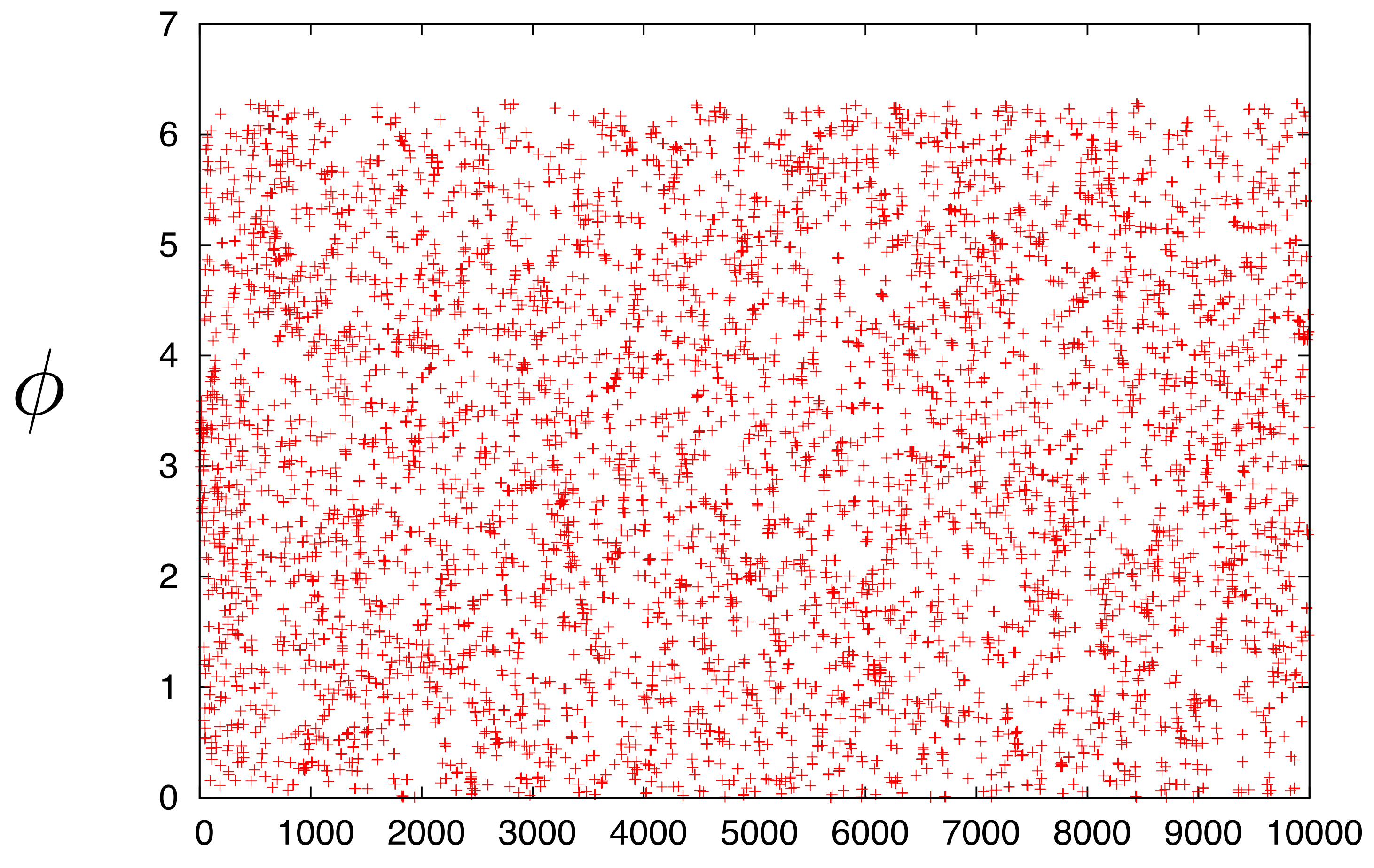
Now we have a problem......

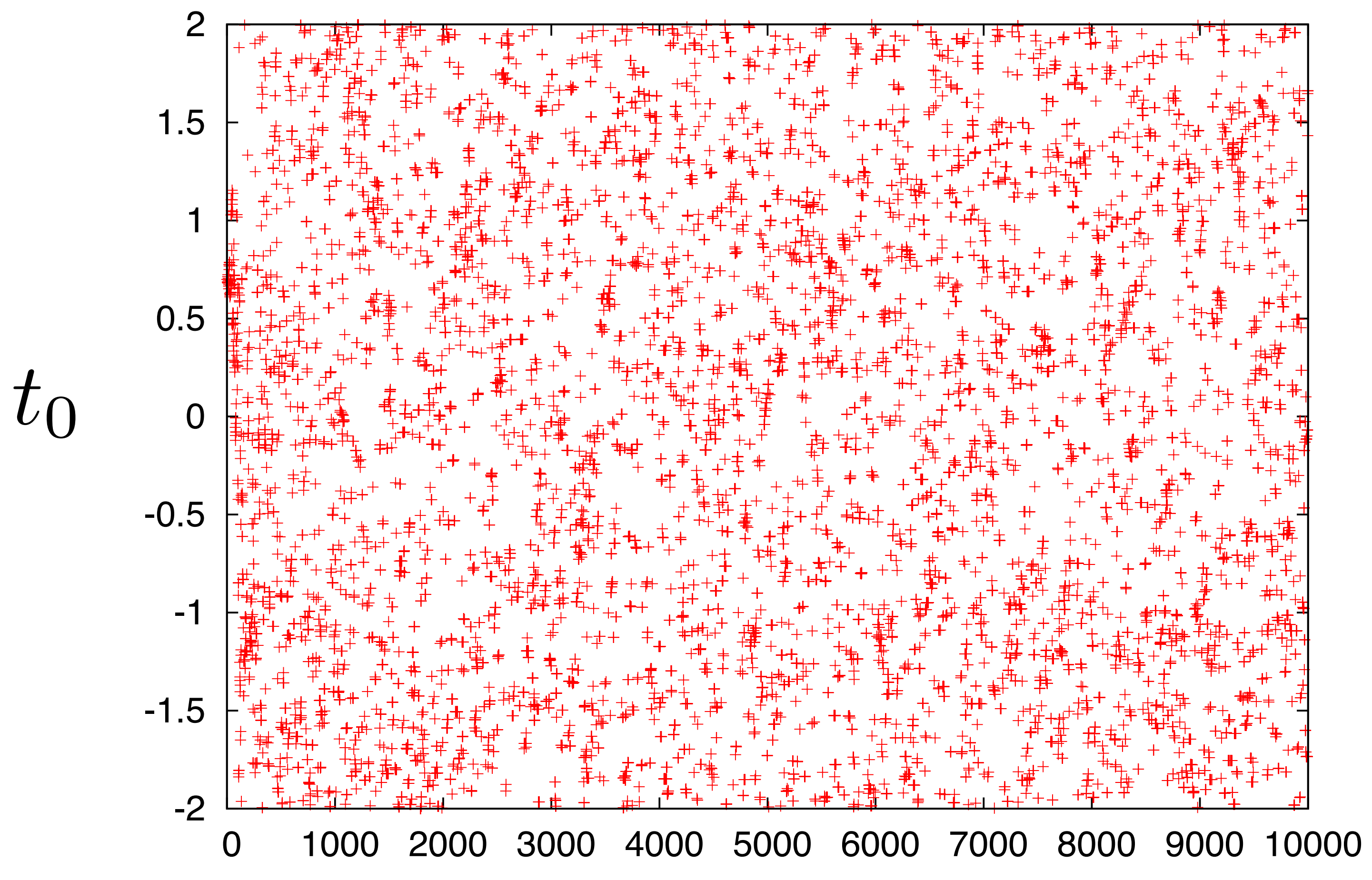
$$2\pi f t_0 \quad \rightarrow \quad 2\pi f t_0 + \alpha$$

$$\phi \quad \rightarrow \quad \phi + \alpha \qquad \Rightarrow \quad \Gamma_{ij} \quad \text{singular}$$

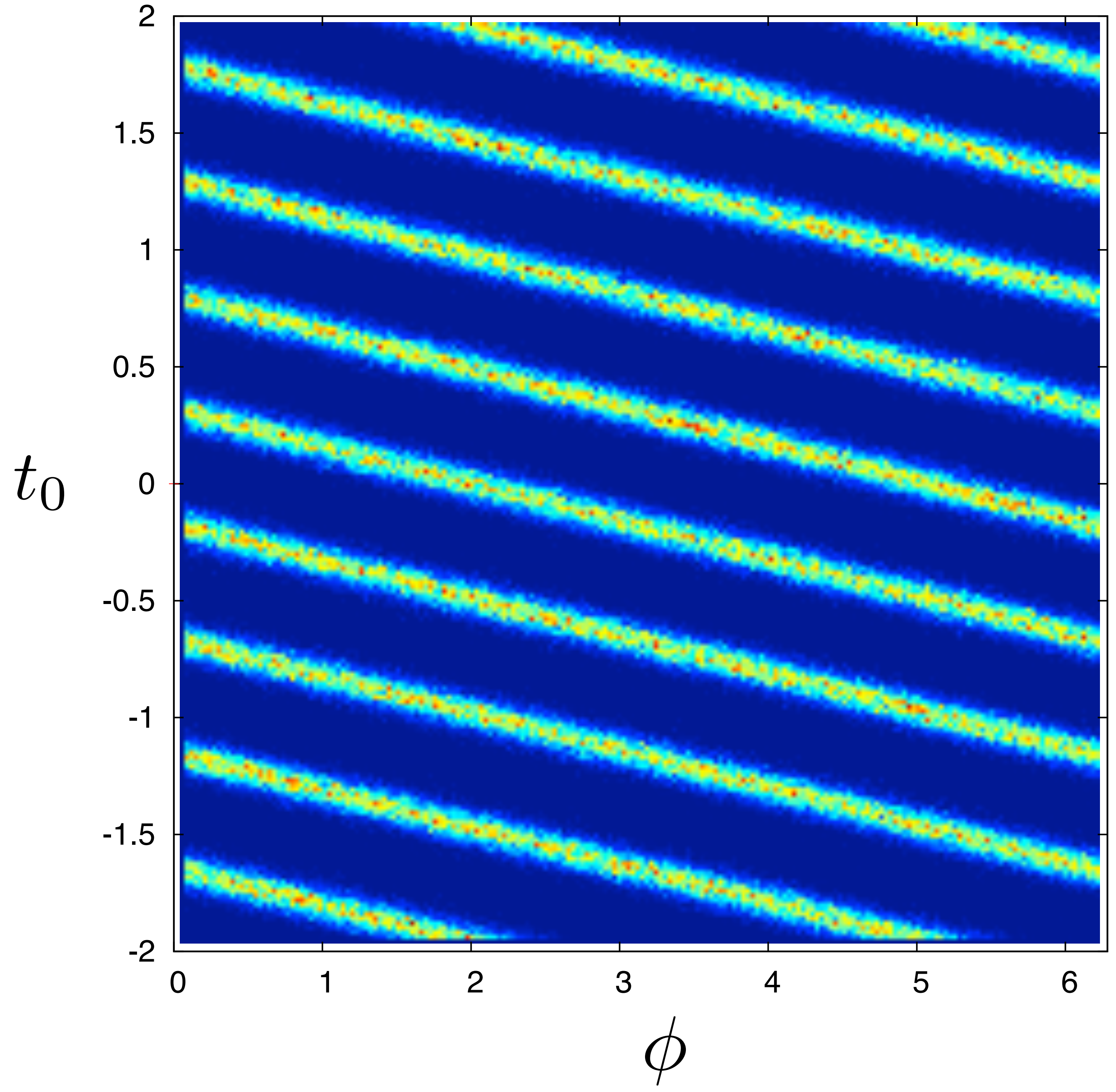
$$h(t) \quad \rightarrow \quad h(t)$$

Markov Chains





2d-PDF

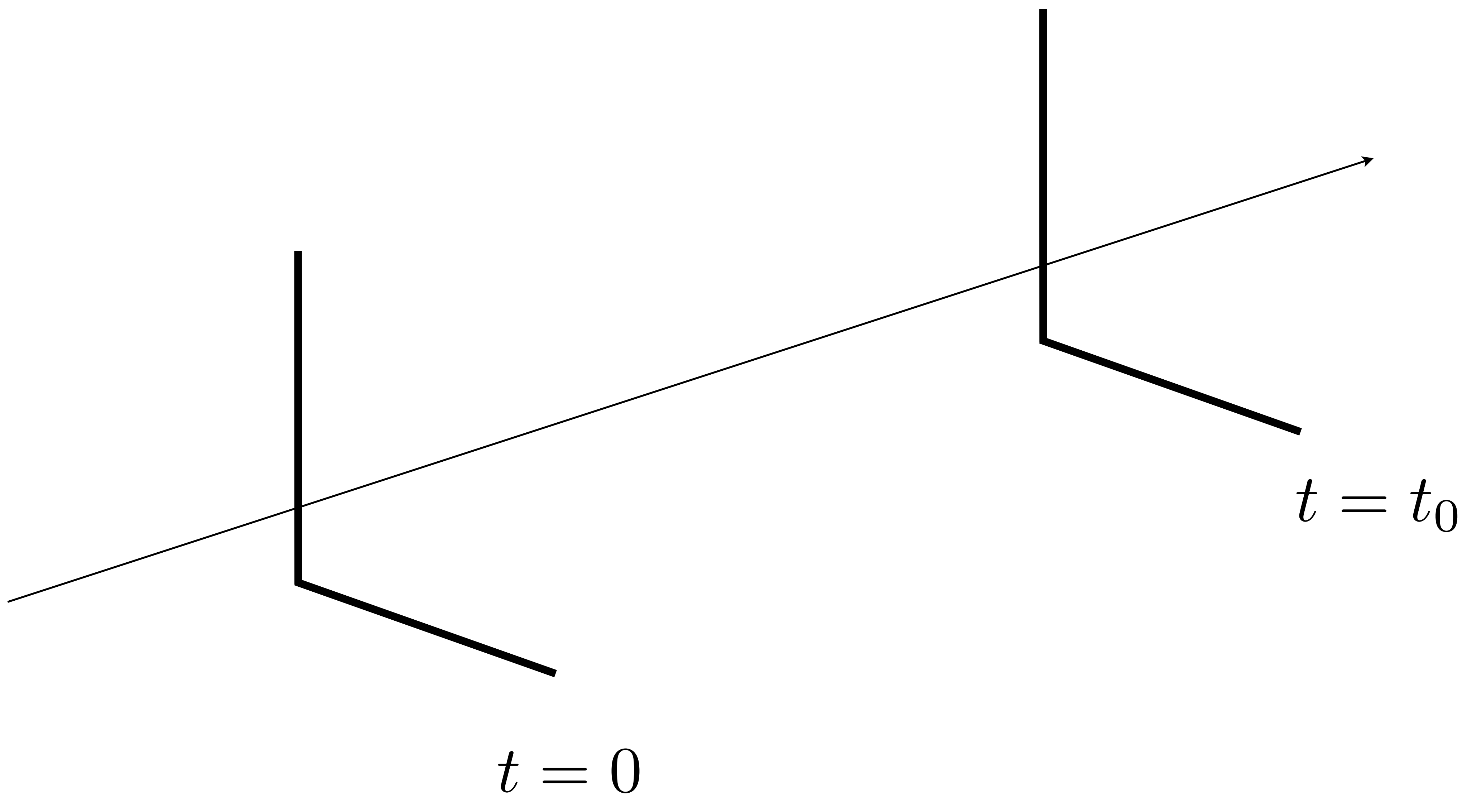


$(\mathrm{SNR} = 7.07)$

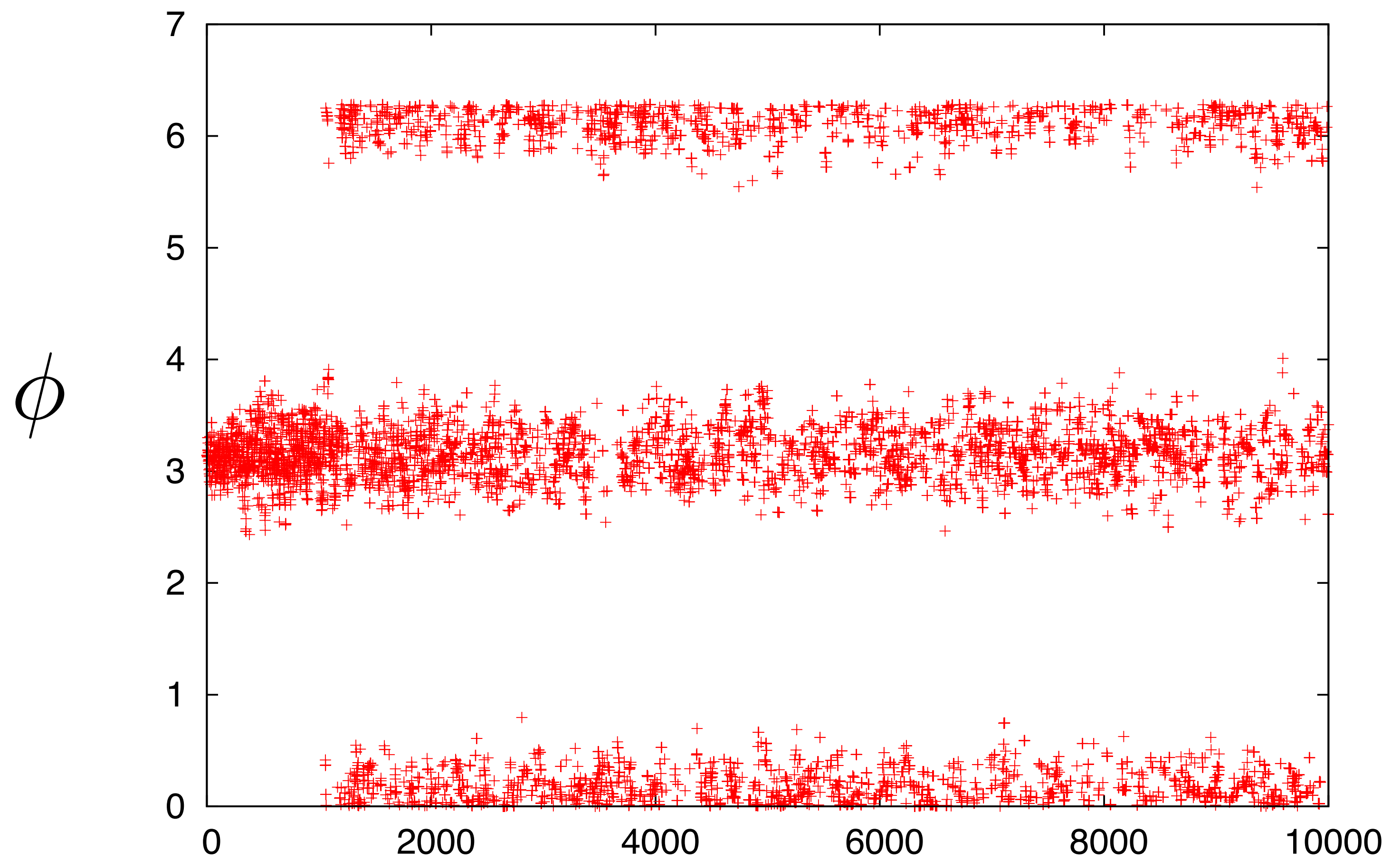
Now with two detectors...

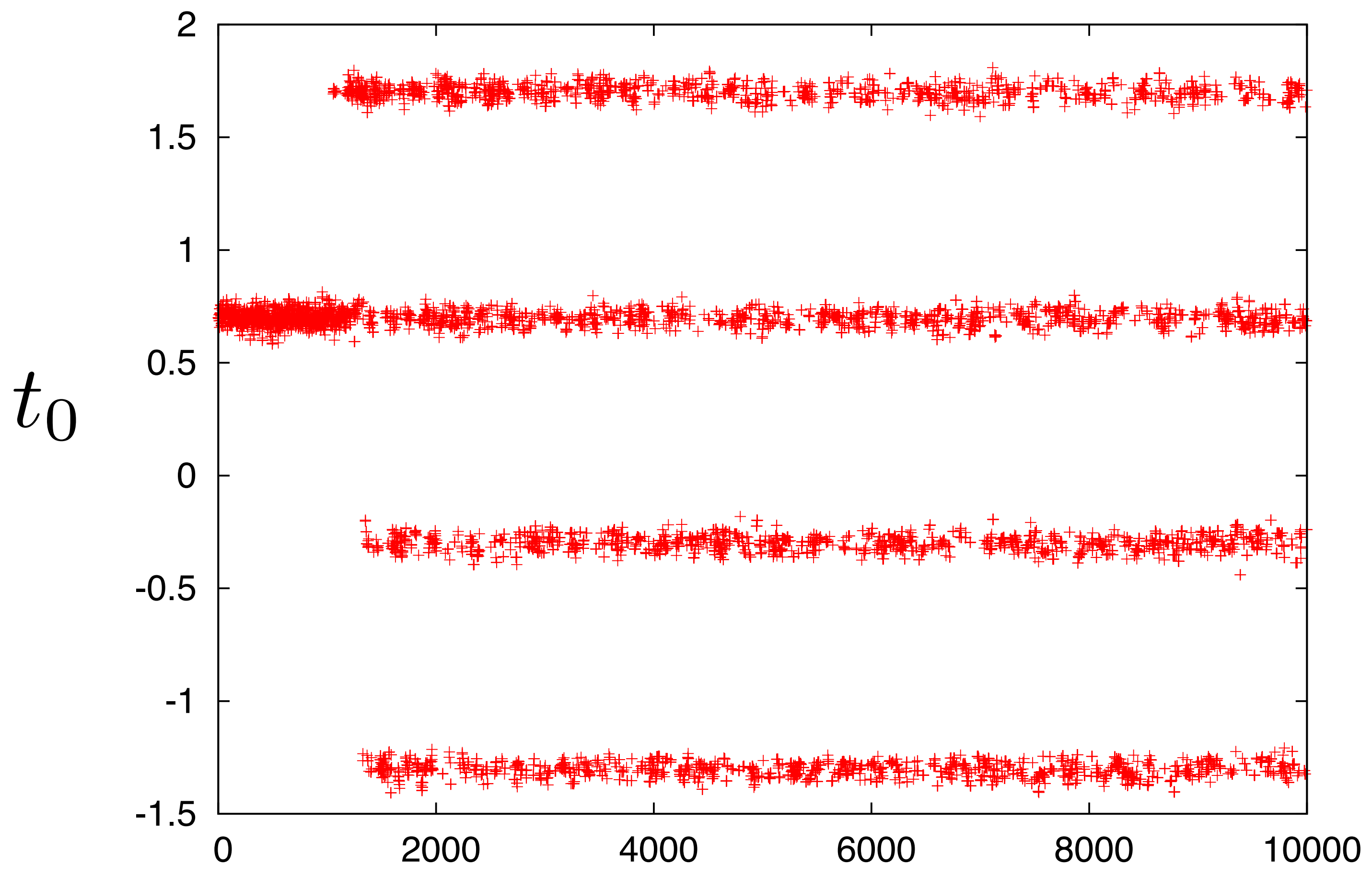
n = stationary, white Gaussian noise with unit variance

$h(t) = A\cos(2\pi f(t - t_0) + \phi)$ Parameters $\{A, f, \phi, t_0\}$

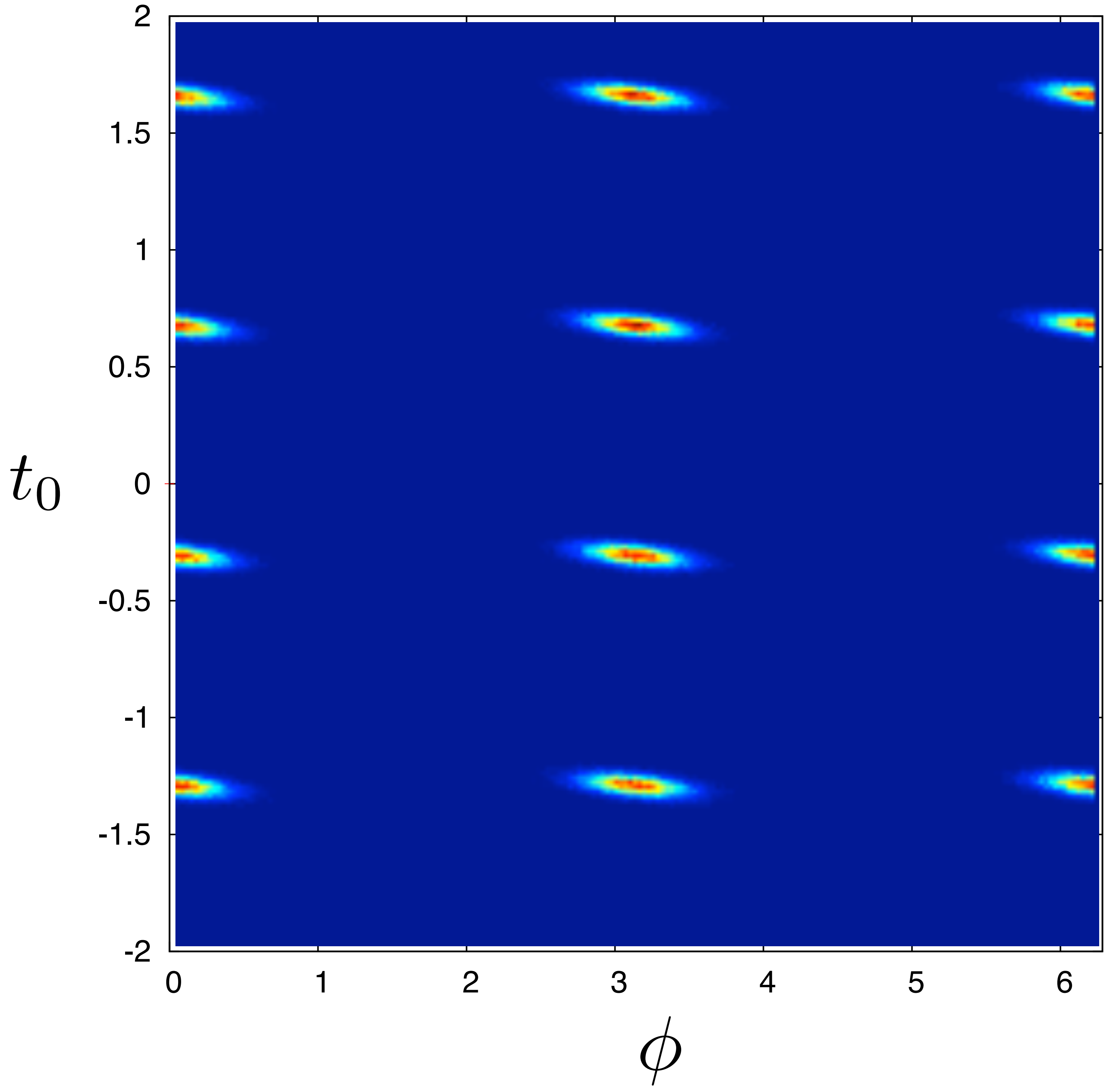


Markov Chains



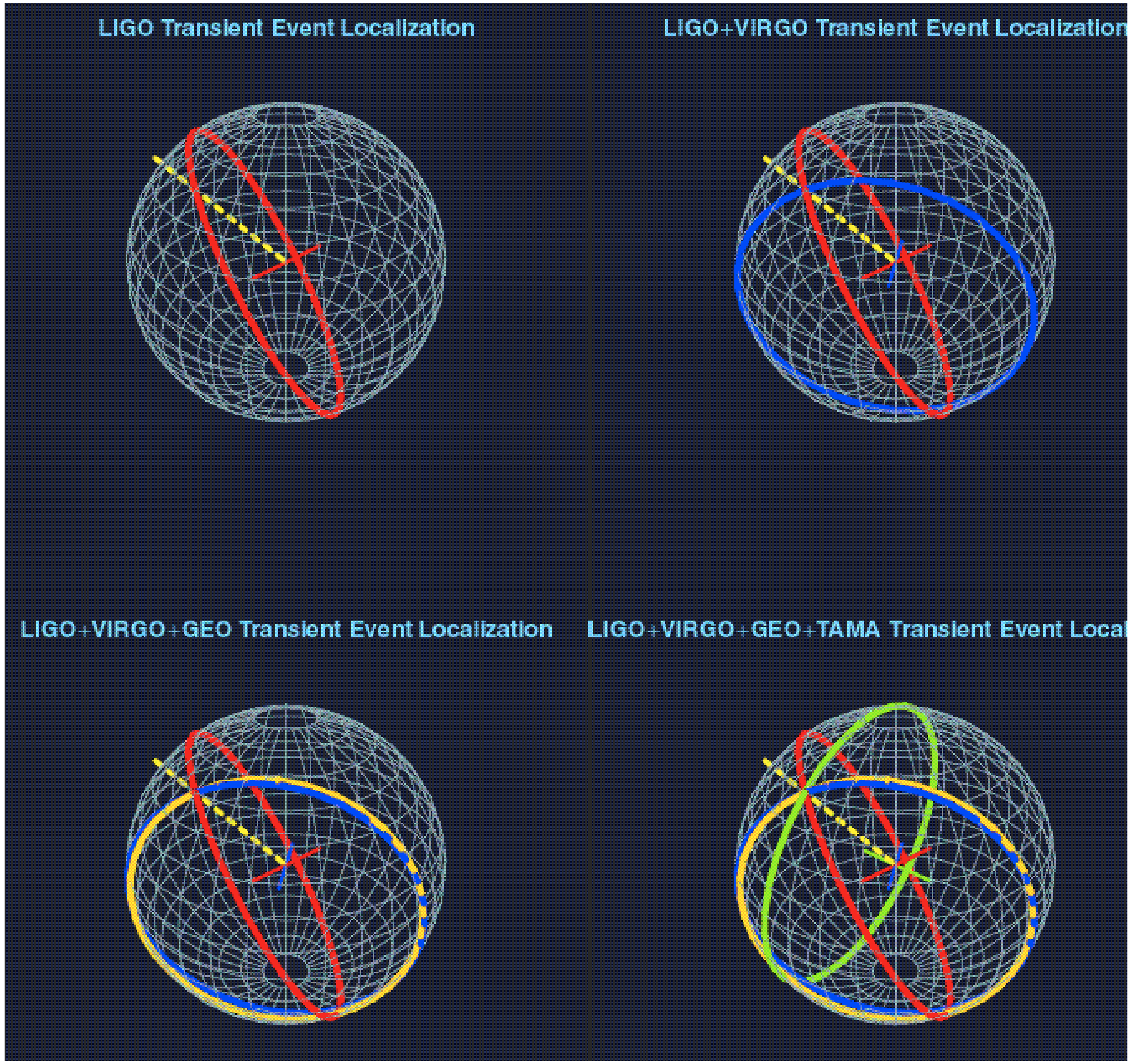


2d-PDF



(SNR = 10.0)

Time delay Measurement $\Rightarrow$ Triangulation



Two detectors

Note 1:

$$\Gamma_{ij} = \Gamma_{ij}^{(1)} + \Gamma_{ij}^{(2)}$$

Non-Singular Singular

Note 2: Bayesian analysis can be done simultaneously or sequentially

Starting with a uniform prior and the data from detector (1)

$$K_{ij}^{(1)} = \Gamma_{ij}^{(1)} - \cancelto{0}{\partial_i \partial_j \ln p(\vec{x})} = \Gamma_{ij}^{(1)}$$

Now update the posterior using the the data from detector (2)

$$K_{ij}^{(2)} = \Gamma_{ij}^{(2)} - \partial_i \partial_j \ln p(\vec{x}) = \Gamma_{ij}^{(2)} + \Gamma_{ij}^{(1)}$$

One GW detector + EM Prior

Suppose we have the sky location from an EM counterpart

Simulate this as a simple Gaussian prior on t_0

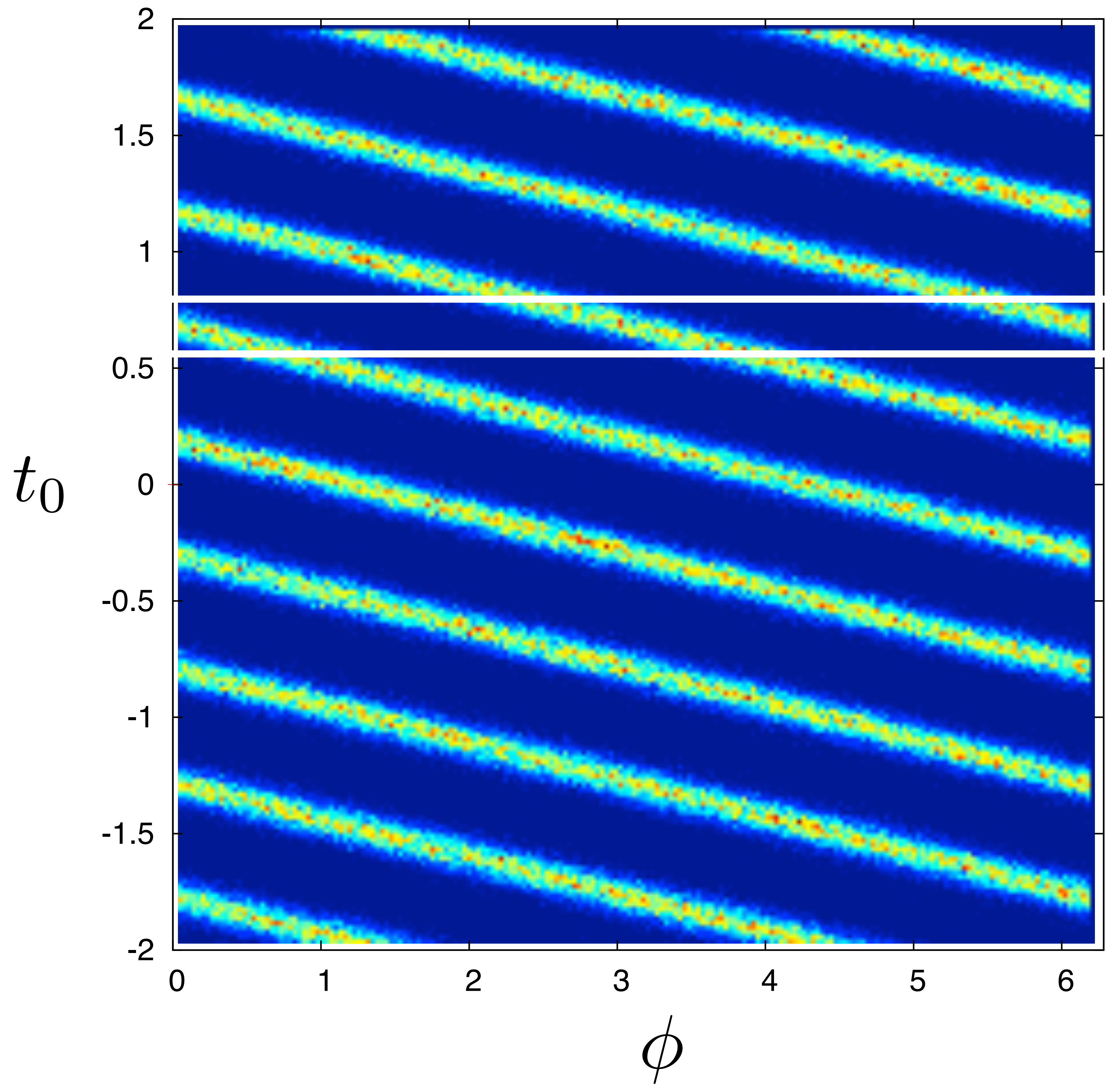
$$p(t_0) = \frac{1}{\sqrt{2\pi}\sigma_{\rm EM}} e^{-(t_0 - t_{\rm EM})^2/2\sigma_{\rm EM}^2}$$

Now the augmented FIM is no-singular

$$\underbrace{K_{ij}}_{\text{Non-Singular}} = \underbrace{\Gamma_{ij}}_{\text{Singular}} + \frac{1}{\sigma_{\rm EM}^2}\underbrace{\delta_{t_0 t_0}}_{\text{from prior}}$$

One GW detector + EM Prior

Without EM Prior



With EM Prior