# PolyChord & the future of nested sampling

## Sampling, Parameter Estimation and Bayesian Model Comparison

Will Handley
wh260@cam.ac.uk

Supervisors: Anthony Lasenby & Mike Hobson
Astrophysics Group
Cavendish Laboratory
University of Cambridge

March 8, 2016

Parameter estimation & model comparison

Metropolis Hastings

Nested Sampling

PolyChord

Applications

# What is nested sampling?

# What is nested sampling?

- Nested sampling is an alternative sampling approach.

# What is nested sampling?

- Nested sampling is an alternative sampling approach.
- Unlike traditional methods, it does not sample from the posterior directly.

# What is nested sampling?

- Nested sampling is an alternative sampling approach.
- Unlike traditional methods, it does not sample from the posterior directly.
- Instead it gradually compresses the prior onto the posterior.

# What is nested sampling?

- Nested sampling is an alternative sampling approach.
- Unlike traditional methods, it does not sample from the posterior directly.
- Instead it gradually compresses the prior onto the posterior.
- In doing so, it circumvents many issues (dimensionality, topology, geometry) that beset normal approaches.
- Similar to simulated annealing, but automatically picks the "correct" annealing schedule.

# Bayes' theorem

## Parameter estimation

# Bayes' theorem

## Parameter estimation

What does data tell us about the params $\Theta$ of our model $M$?

$$\mathrm{P}(\Theta|D, M) = \frac{\mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M)}{\mathrm{P}(D|M)}$$

# Bayes' theorem

## Parameter estimation

What does data tell us about the params $\Theta$ of our model $M$?

$$\mathrm{P}(\Theta|D, M) = \frac{\mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M)}{\mathrm{P}(D|M)}$$

$$\text{Posterior} \;= \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

# Bayes' theorem

## Parameter estimation

What does data tell us about the params $\Theta$ of our model $M$?

$$\mathrm{P}(\Theta|D, M) = \frac{\mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M)}{\mathrm{P}(D|M)}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

$$\mathcal{P}(\Theta) = \frac{\mathcal{L}(\Theta)\pi(\Theta)}{\mathcal{Z}}$$

# Bayes' theorem

## Model comparison

# Bayes' theorem

## Model comparison

What does data tell us about our model $M_i$ in relation to other models $\{M_1, M_2, \cdots\}$?

# Bayes' theorem

## Model comparison

What does data tell us about our model $M_i$ in relation to other models $\{M_1, M_2, \cdots\}$?

$$\mathrm{P}(M_i|D) = \frac{\mathrm{P}(D|M_i)\mathrm{P}(M_i)}{\mathrm{P}(D)}$$

# Bayes' theorem

## Model comparison

What does data tell us about our model $M_i$ in relation to other models $\{M_1, M_2, \cdots\}$?

$$\mathrm{P}(M_i|D) = \frac{\mathrm{P}(D|M_i)\mathrm{P}(M_i)}{\mathrm{P}(D)}$$

$$\mathrm{P}(M_i|D) = \frac{\mathcal{Z}_i \, \mu_i}{\sum_k \mathcal{Z}_k \, \mu_k}$$

# Parameter estimation & model comparison

## The challenge

# Parameter estimation & model comparison

## The challenge

Both of these are challenging things to compute.

# Parameter estimation & model comparison

## The challenge

Both of these are challenging things to compute.

- Markov-Chain Monte-Carlo (MCMC) can solve the first of these (kind of)

# Parameter estimation & model comparison

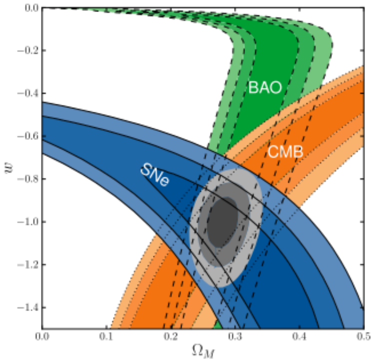
## The challenge

Both of these are challenging things to compute.

- Markov-Chain Monte-Carlo (MCMC) can solve the first of these (kind of)
- Nested sampling (NS) promises to solve both simultaneously.

# Parameter estimation & model comparison

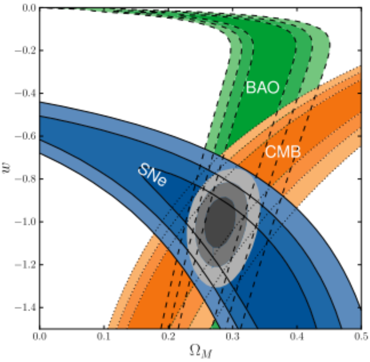
## Why is it difficult?



# Parameter estimation & model comparison

## Why is it difficult?

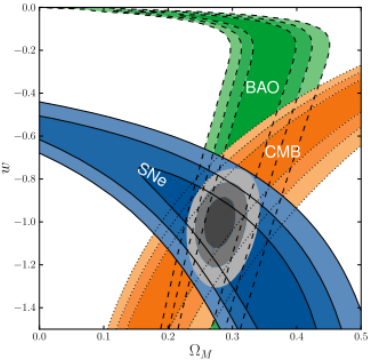
1. In high dimensions, posterior $\mathcal{P}$ occupies a vanishingly small region of the prior $\pi$.



# Parameter estimation & model comparison

## Why is it difficult?

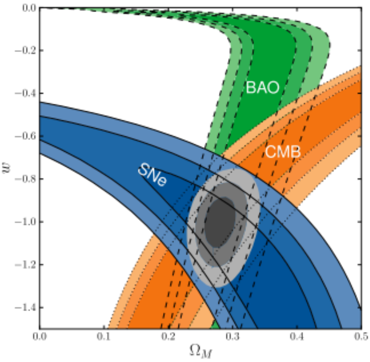
1. In high dimensions, posterior $\mathcal{P}$ occupies a vanishingly small region of the prior $\pi$.
2. Worse, you don't know where this region is.



# Parameter estimation & model comparison

## Why is it difficult?

1. In high dimensions, posterior $\mathcal{P}$ occupies a vanishingly small region of the prior $\pi$.
2. Worse, you don't know where this region is.

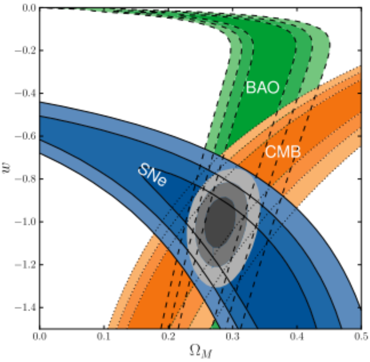



- Describing an $N$-dimensional posterior fully is impossible.

# Parameter estimation & model comparison

## Why is it difficult?

1. In high dimensions, posterior $\mathcal{P}$ occupies a vanishingly small region of the prior $\pi$.
2. Worse, you don't know where this region is.




- Describing an *N*-dimensional posterior fully is impossible.
- *Sampling* the posterior is an excellent compression scheme.

# Markov-Chain Monte-Carlo (MCMC)

## Metropolis-Hastings, Gibbs, Hamiltonian...

# Markov-Chain Monte-Carlo (MCMC)

## Metropolis-Hastings, Gibbs, Hamiltonian...

- Turn the $N$-dimensional problem into a one-dimensional one.

# Markov-Chain Monte-Carlo (MCMC)

## Metropolis-Hastings, Gibbs, Hamiltonian...

- Turn the $N$-dimensional problem into a one-dimensional one.
- Explore the space via a biased random walk.

# Markov-Chain Monte-Carlo (MCMC)

## Metropolis-Hastings, Gibbs, Hamiltonian...

- Turn the $N$-dimensional problem into a one-dimensional one.
- Explore the space via a biased random walk.
  1. Pick random direction

# Markov-Chain Monte-Carlo (MCMC)

## Metropolis-Hastings, Gibbs, Hamiltonian...

- Turn the $N$-dimensional problem into a one-dimensional one.
- Explore the space via a biased random walk.
  1. Pick random direction
  2. Choose step length

# Markov-Chain Monte-Carlo (MCMC)

## Metropolis-Hastings, Gibbs, Hamiltonian...

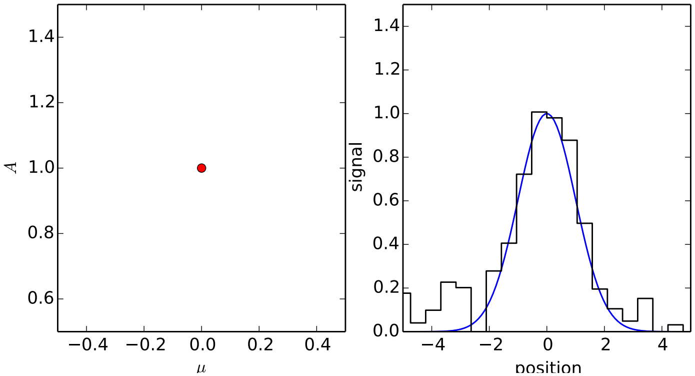
- Turn the $N$-dimensional problem into a one-dimensional one.
- Explore the space via a biased random walk.
  1. Pick random direction
  2. Choose step length
  3. If uphill, make step...

# Markov-Chain Monte-Carlo (MCMC)

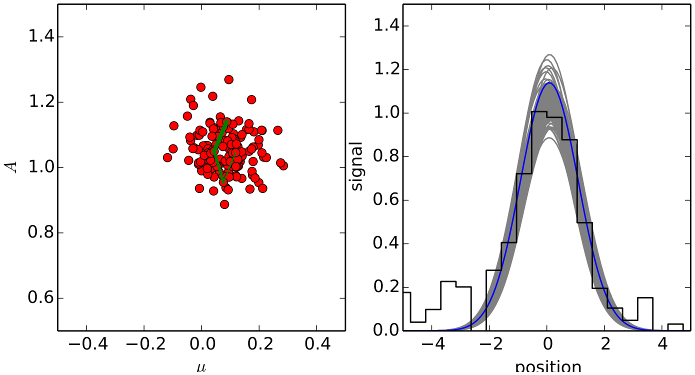
## Metropolis-Hastings, Gibbs, Hamiltonian...

- Turn the $N$-dimensional problem into a one-dimensional one.
- Explore the space via a biased random walk.
  1. Pick random direction
  2. Choose step length
  3. If uphill, make step...
  4. ...otherwise sometimes make step.

# MCMC in action

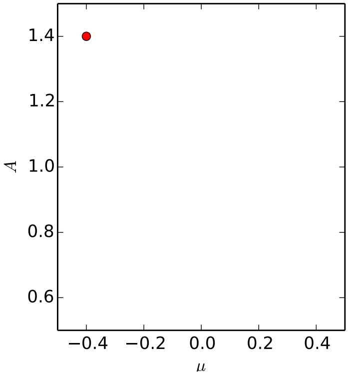


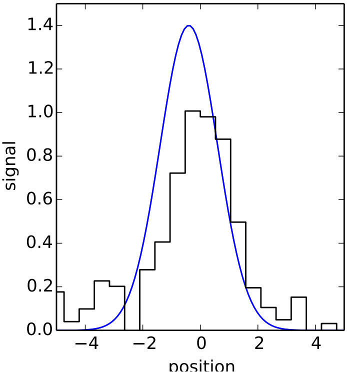
# MCMC in action



# When MCMC fails

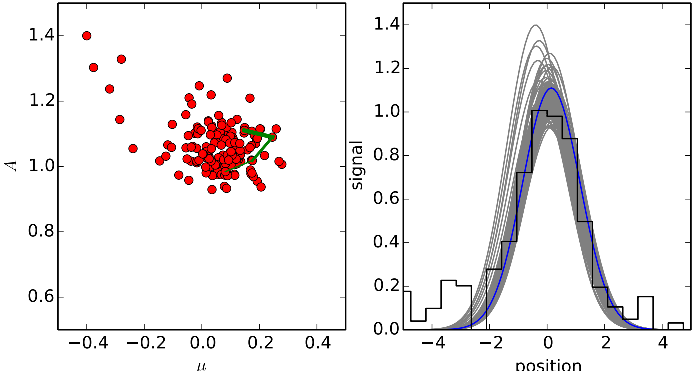
## Burn in



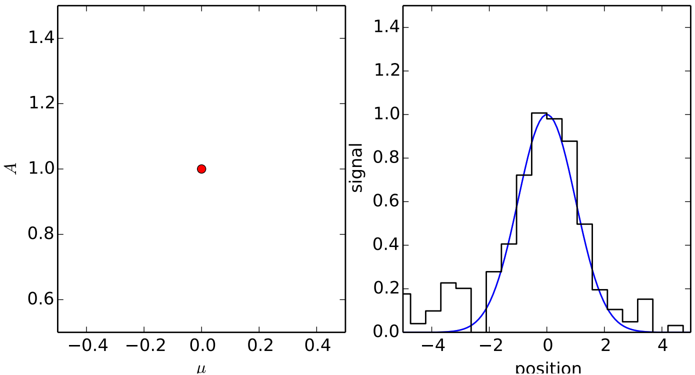
# When MCMC fails

## Burn in



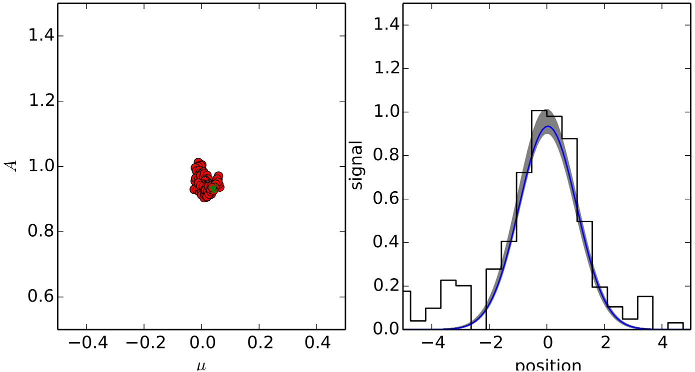
# When MCMC fails

## Tuning the proposal distribution



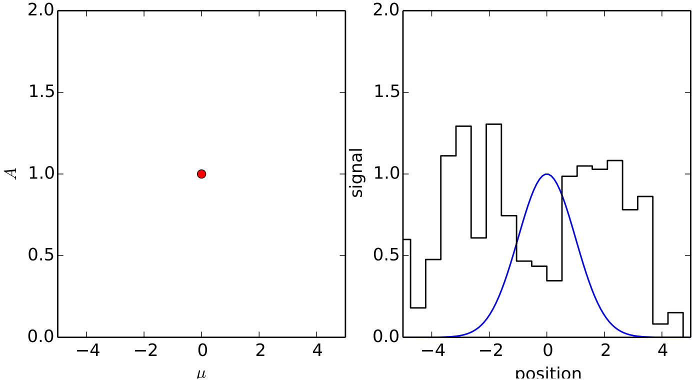
# When MCMC fails

## Tuning the proposal distribution



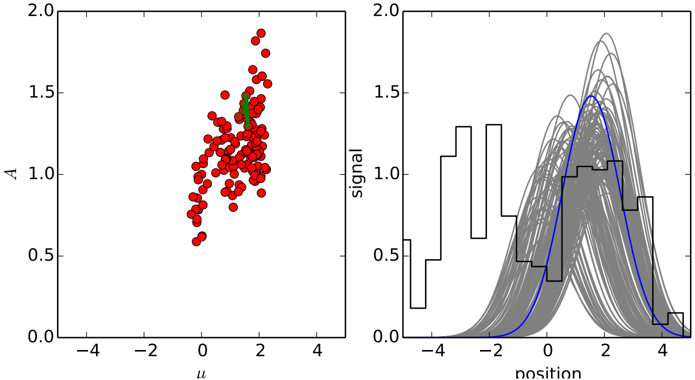
# When MCMC fails

## Multimodality



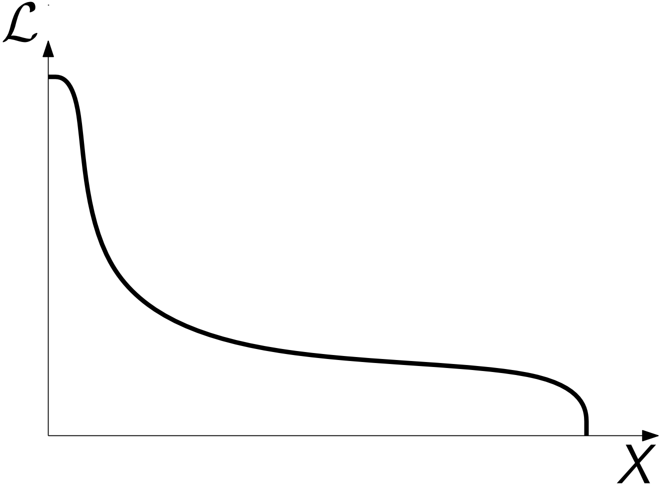
# When MCMC fails

## Multimodality



# When MCMC fails

## Phase transitions



# When MCMC fails

## The real reason. . .

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$\mathcal{Z} = \mathrm{P}(D|M)$$

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$
\begin{aligned}
\mathcal{Z} &= \mathrm{P}(D|M) \\
&= \int \mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M) d\Theta
\end{aligned}
$$

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$
\begin{aligned}
\mathcal{Z} &= \mathrm{P}(D|M) \\
&= \int \mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M)d\Theta \\
&= \int \mathcal{L}(\Theta)\pi(\Theta)d\Theta
\end{aligned}
$$

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$
\begin{aligned}
\mathcal{Z} &= \mathrm{P}(D|M) \\
&= \int \mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M) d\Theta \\
&= \int \mathcal{L}(\Theta)\pi(\Theta) d\Theta \\
&= \langle \mathcal{L} \rangle_\pi
\end{aligned}
$$

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$\begin{aligned}\mathcal{Z} &= \mathrm{P}(D|M) \\ &= \int \mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M)d\Theta \\ &= \int \mathcal{L}(\Theta)\pi(\Theta)d\Theta \\ &= \langle\mathcal{L}\rangle_\pi\end{aligned}$$

- MCMC fundamentally explores the posterior, and cannot average over the prior.

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$
\begin{aligned}
\mathcal{Z} &= \mathrm{P}(D|M) \\
&= \int \mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M) d\Theta \\
&= \int \mathcal{L}(\Theta)\pi(\Theta) d\Theta \\
&= \langle \mathcal{L} \rangle_\pi
\end{aligned}
$$

- MCMC fundamentally explores the posterior, and cannot average over the prior.
- Simulated annealing gives one possibility for computing evidences.

# When MCMC fails

## The real reason...

- MCMC does not give you evidences!

$$\begin{aligned}
\mathcal{Z} &= \mathrm{P}(D|M) \\
&= \int \mathrm{P}(D|\Theta, M)\mathrm{P}(\Theta|M) d\Theta \\
&= \int \mathcal{L}(\Theta)\pi(\Theta) d\Theta \\
&= \langle \mathcal{L} \rangle_\pi
\end{aligned}$$

- MCMC fundamentally explores the posterior, and cannot average over the prior.
- Simulated annealing gives one possibility for computing evidences.
  - Suffers from similar issues to MCMC, especially phase transitions

# Nested Sampling

## John Skilling's alternative to MCMC!

# Nested Sampling

## John Skilling's alternative to MCMC!

New procedure:

# Nested Sampling

## John Skilling's alternative to MCMC!

New procedure:
Maintain a set $S$ of $n$ samples, which are sequentially updated:

# Nested Sampling

## John Skilling's alternative to MCMC!

New procedure:

Maintain a set $S$ of $n$ samples, which are sequentially updated:

$S_0$: Generate $n$ samples from the prior $\pi$.

# Nested Sampling

John Skilling's alternative to MCMC!

New procedure:

Maintain a set $S$ of $n$ samples, which are sequentially updated:

$S_0$: Generate $n$ samples from the prior $\pi$.

$S_{n+1}$: Delete the lowest likelihood sample in $S_n$, and replace it with a new sample with higher likelihood

# Nested Sampling

## John Skilling's alternative to MCMC!

New procedure:

Maintain a set $S$ of $n$ samples, which are sequentially updated:

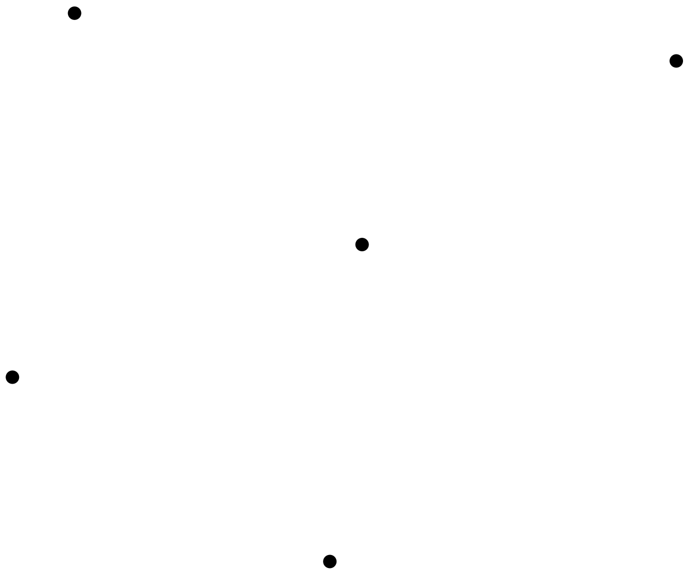
$S_0$: Generate $n$ samples from the prior $\pi$.

$S_{n+1}$: Delete the lowest likelihood sample in $S_n$, and replace it with a new sample with higher likelihood

Requires one to be able to sample from the prior, subject to a *hard likelihood constraint*.

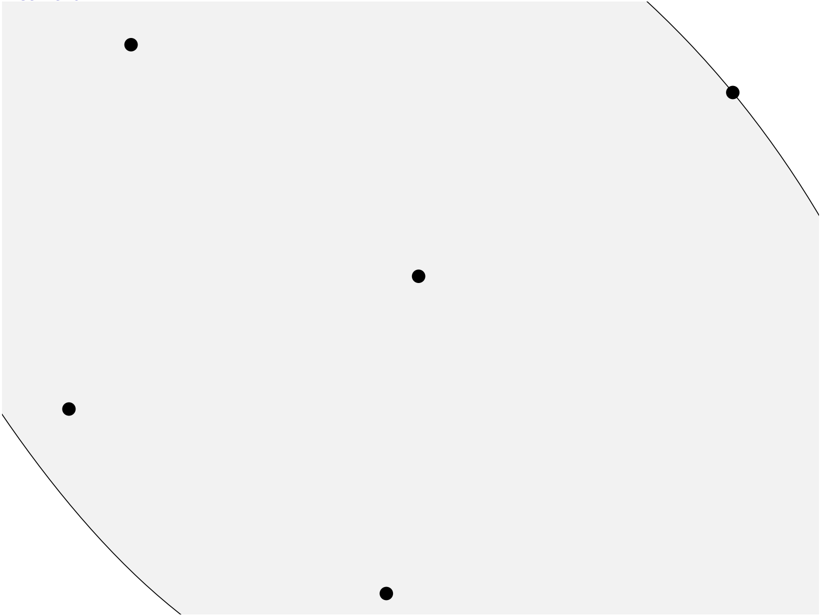
# Nested Sampling

## Graphical aid



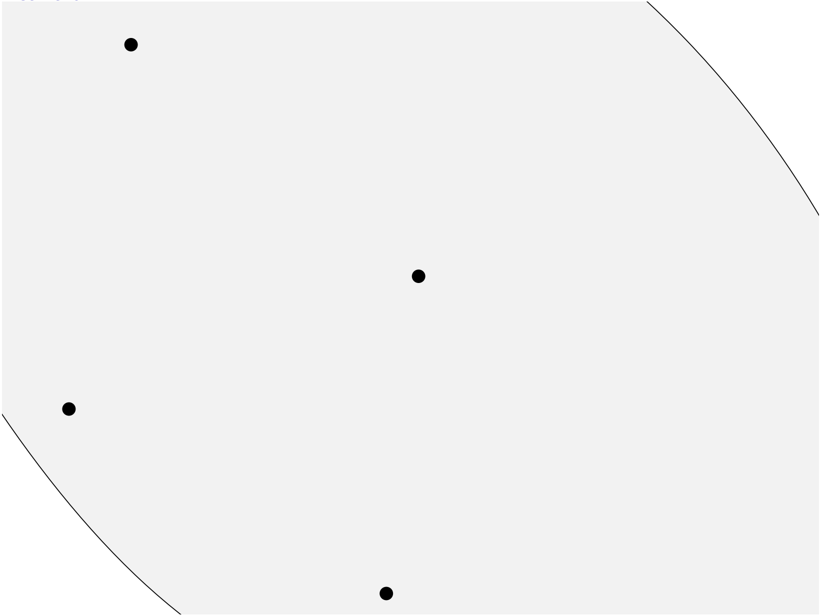
# Nested Sampling

## Graphical aid



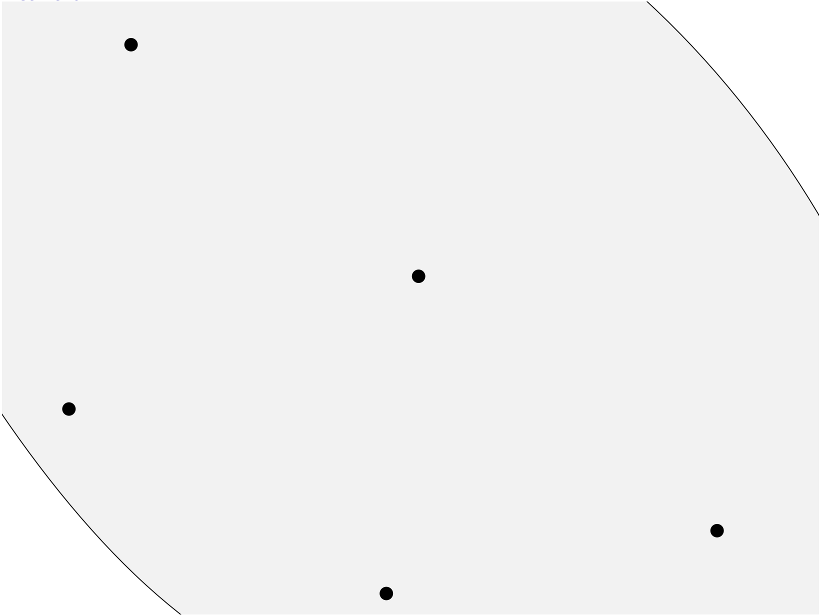
# Nested Sampling

## Graphical aid



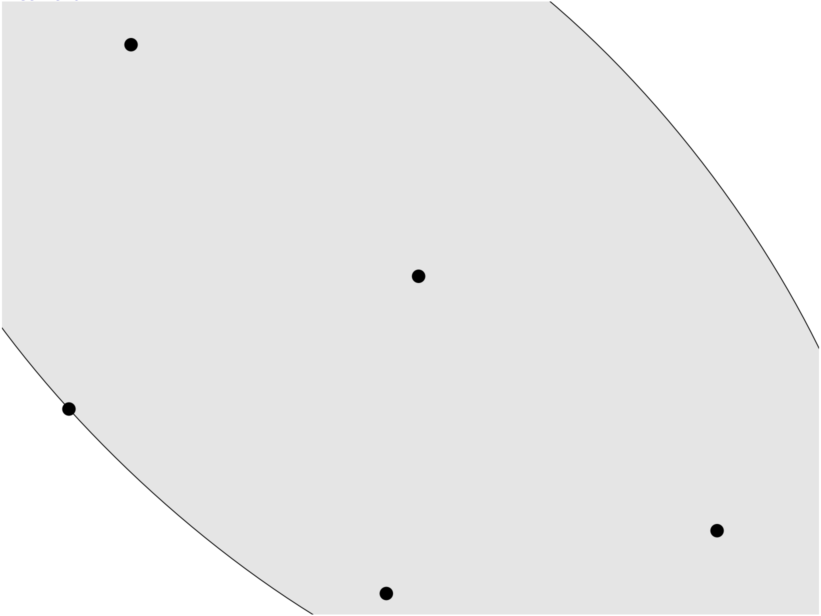
# Nested Sampling

## Graphical aid



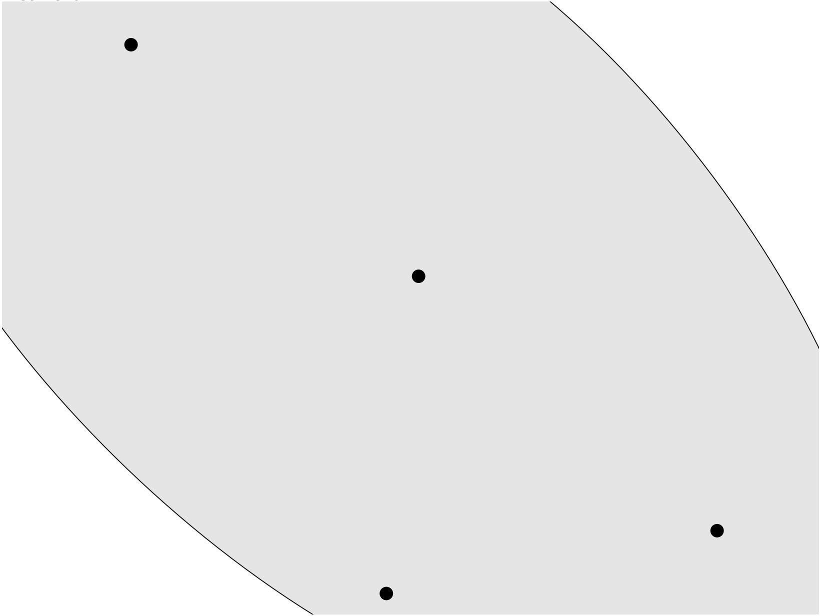
# Nested Sampling

## Graphical aid



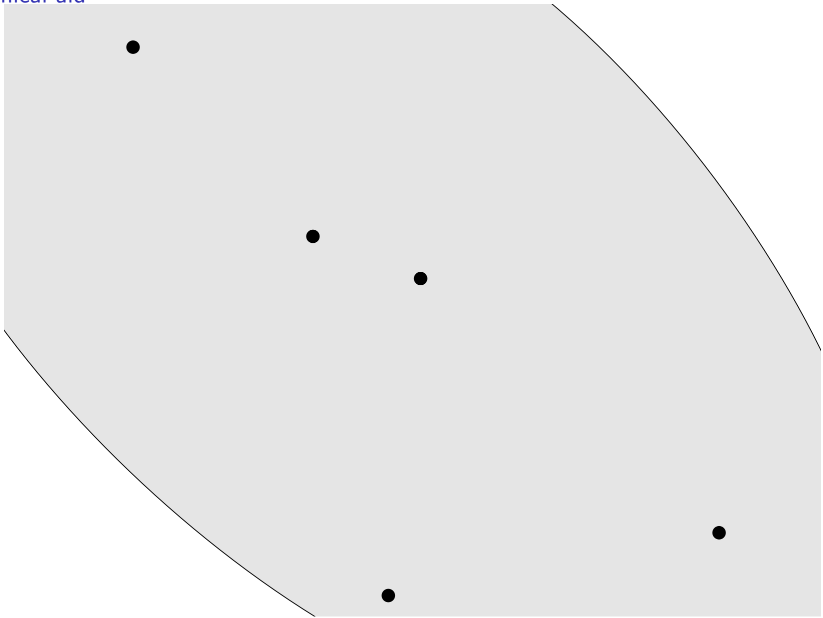
# Nested Sampling

## Graphical aid



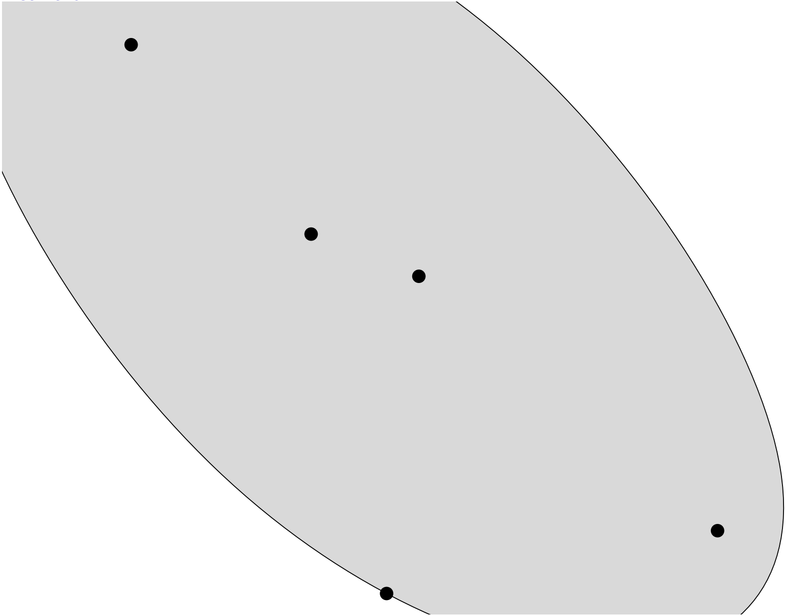
# Nested Sampling

## Graphical aid



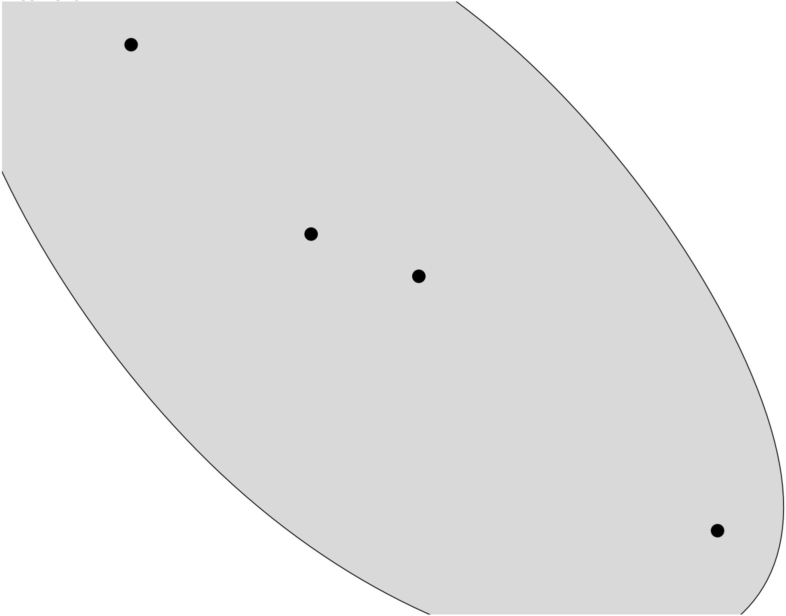
# Nested Sampling

## Graphical aid



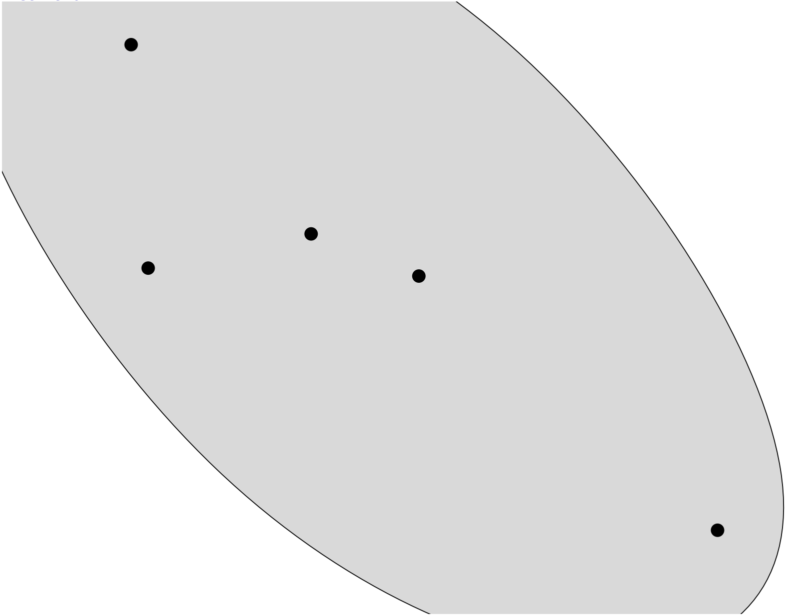
# Nested Sampling

## Graphical aid



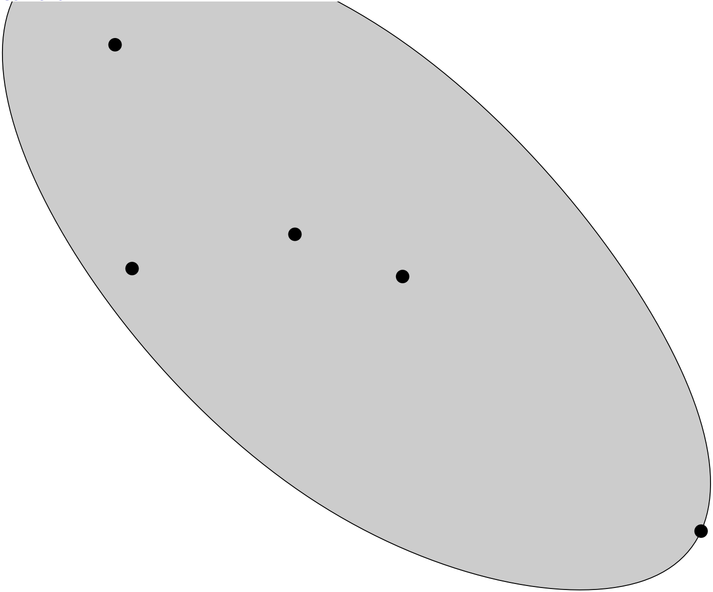
# Nested Sampling

## Graphical aid



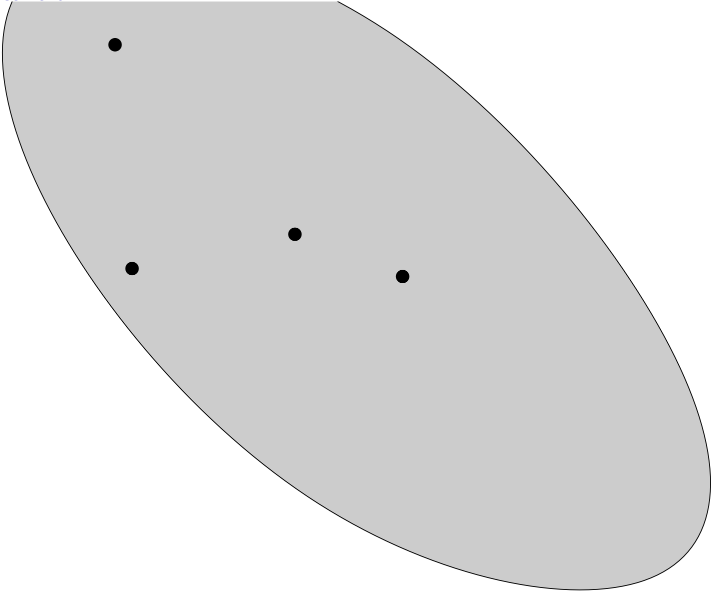
# Nested Sampling

## Graphical aid



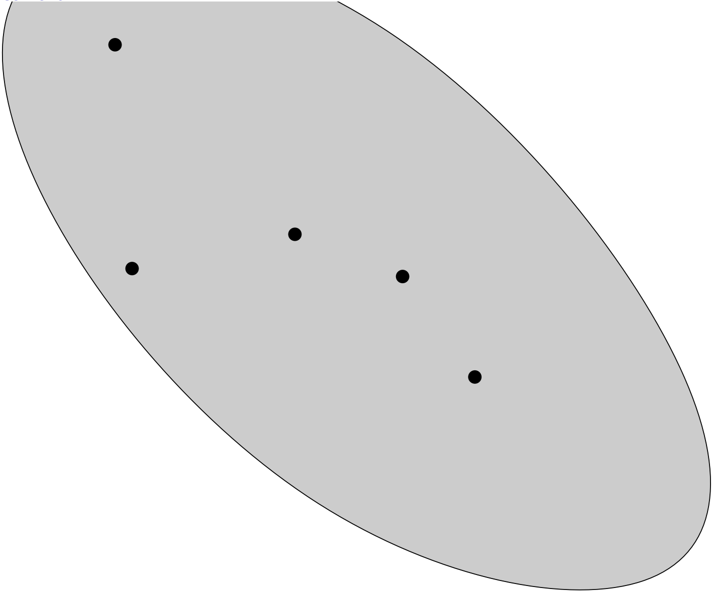
# Nested Sampling

## Graphical aid



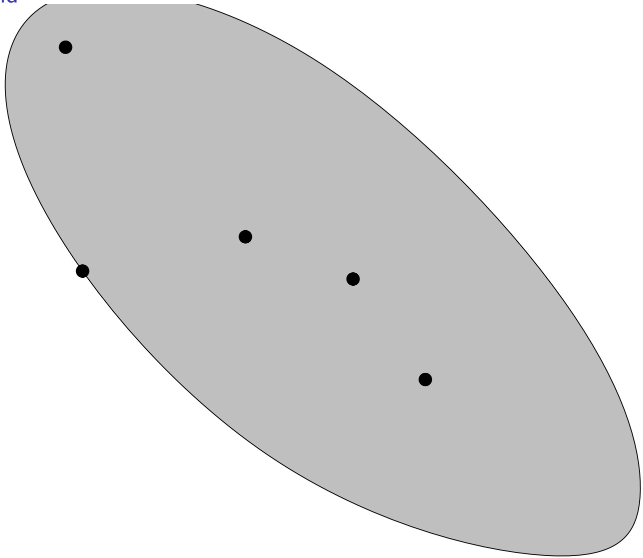
# Nested Sampling

## Graphical aid



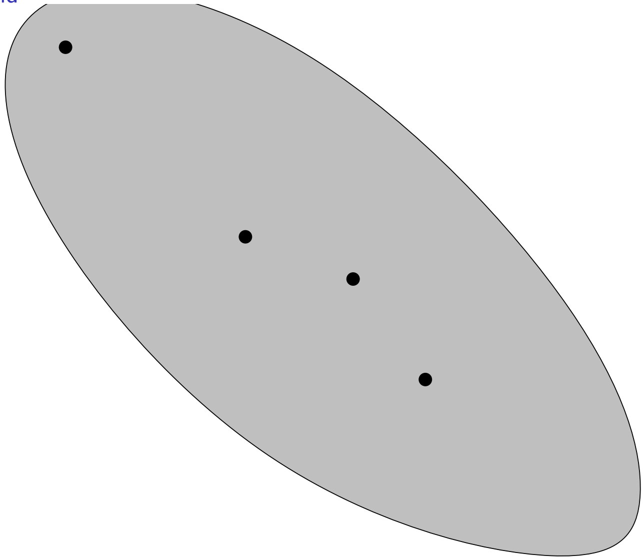
# Nested Sampling

## Graphical aid



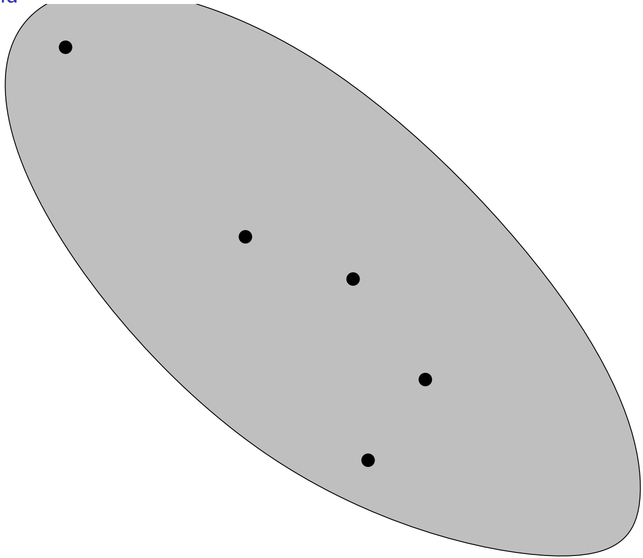
# Nested Sampling

## Graphical aid



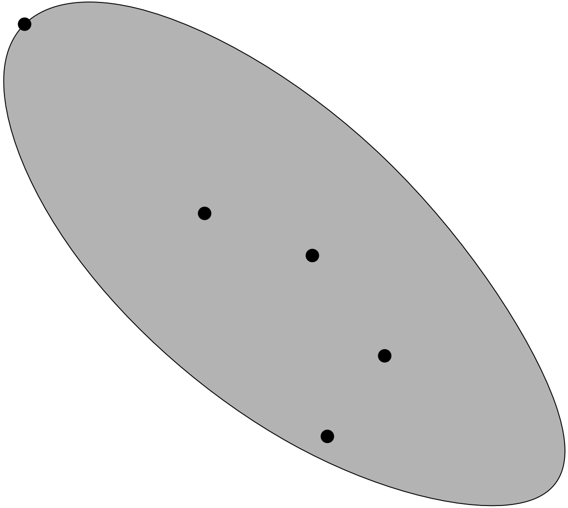
# Nested Sampling

## Graphical aid



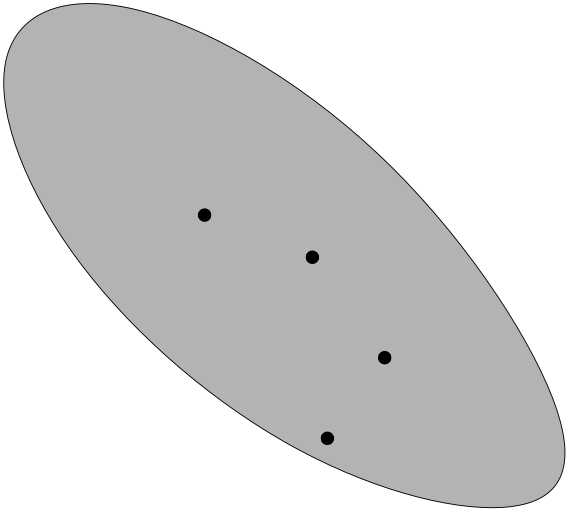
# Nested Sampling

## Graphical aid



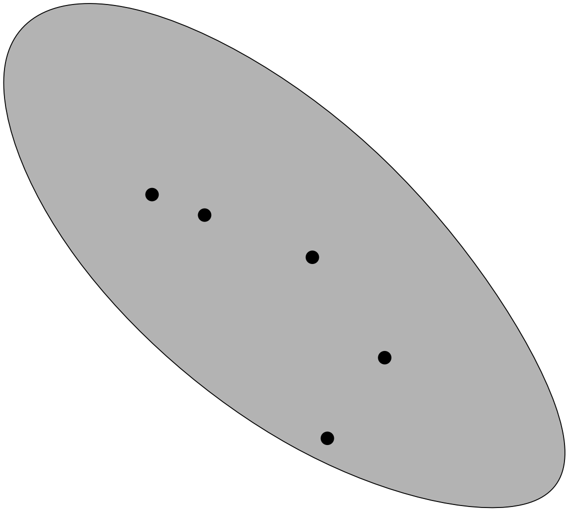
# Nested Sampling

## Graphical aid



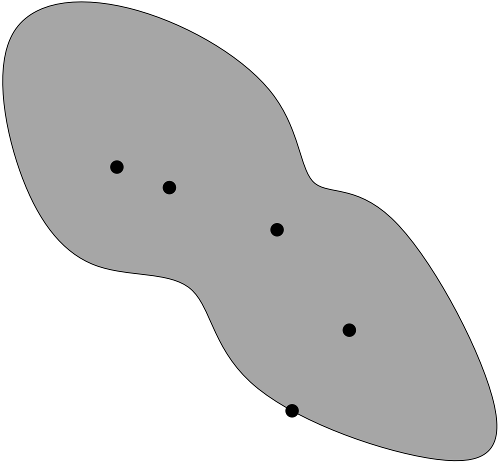
# Nested Sampling

## Graphical aid



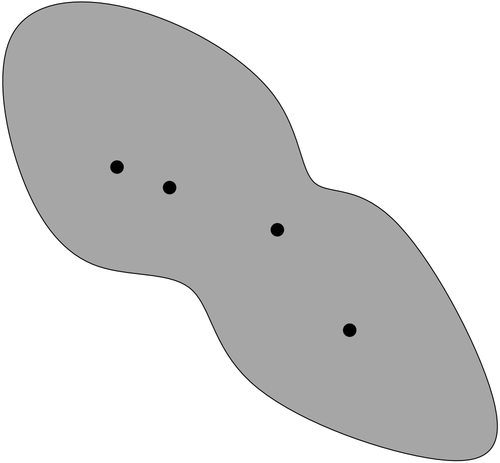
# Nested Sampling

## Graphical aid



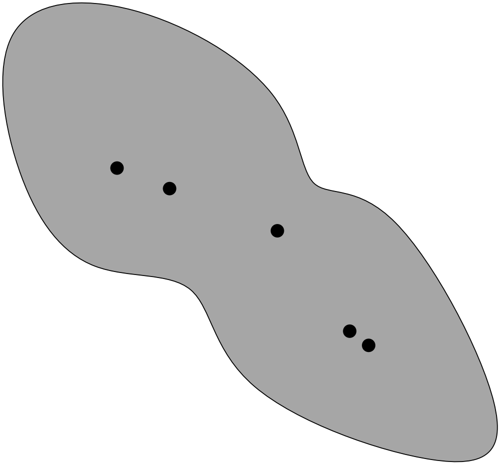
# Nested Sampling

## Graphical aid



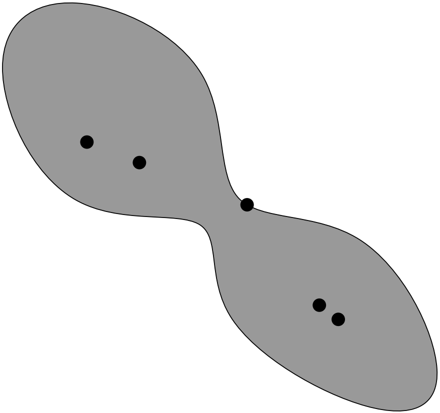
# Nested Sampling

## Graphical aid



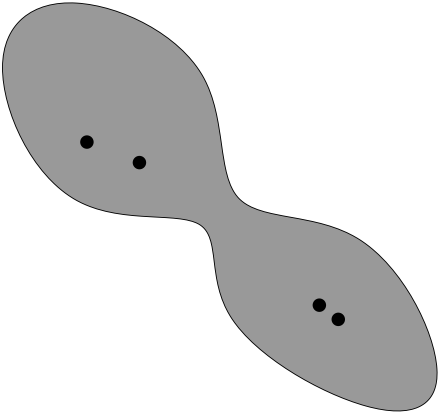
# Nested Sampling

## Graphical aid



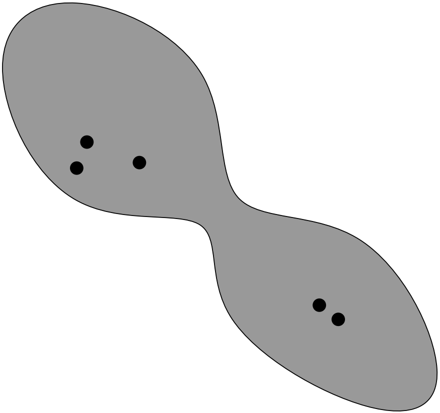
# Nested Sampling

## Graphical aid



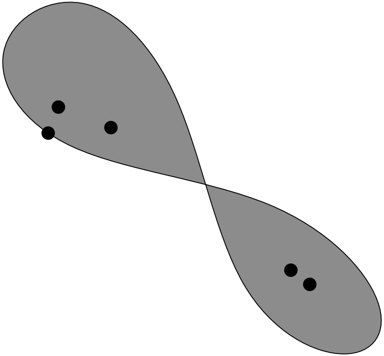
# Nested Sampling

## Graphical aid



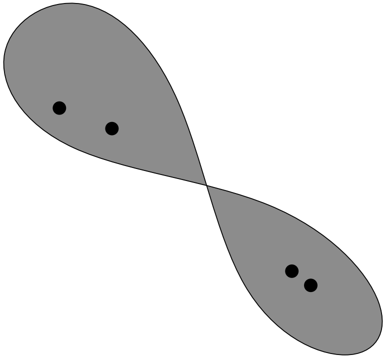
# Nested Sampling

## Graphical aid



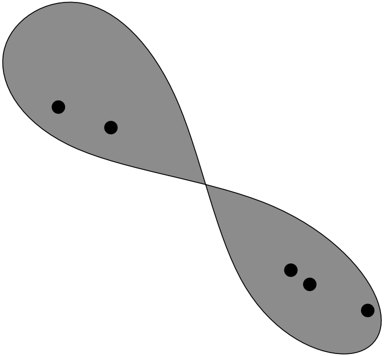
# Nested Sampling

## Graphical aid



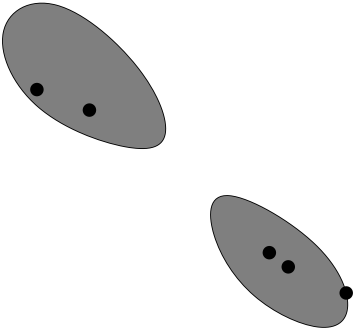
# Nested Sampling

## Graphical aid



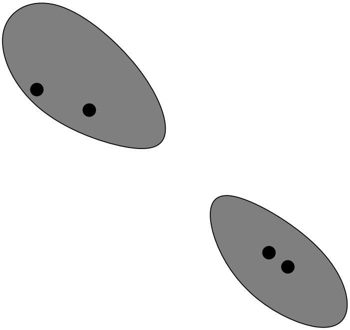
# Nested Sampling

## Graphical aid



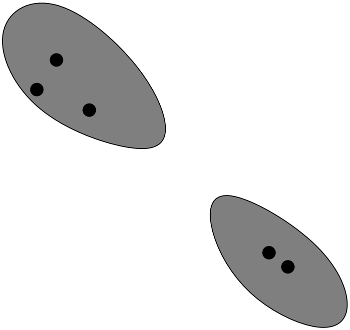
# Nested Sampling

## Graphical aid



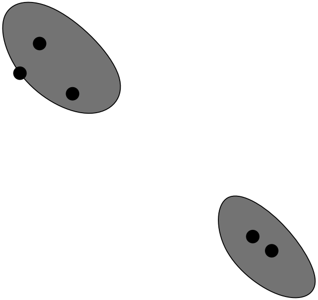
# Nested Sampling

## Graphical aid



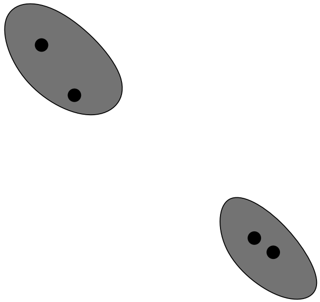
# Nested Sampling

## Graphical aid



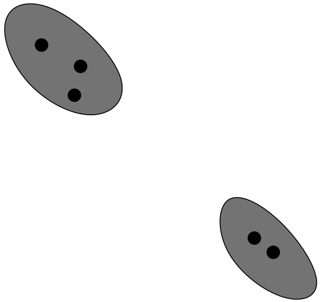
# Nested Sampling

## Graphical aid



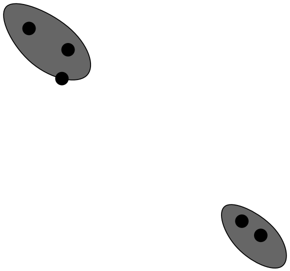
# Nested Sampling

## Graphical aid



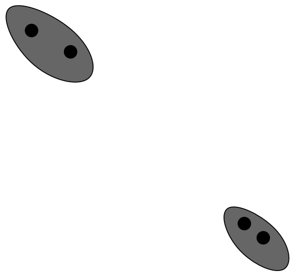
# Nested Sampling

## Graphical aid



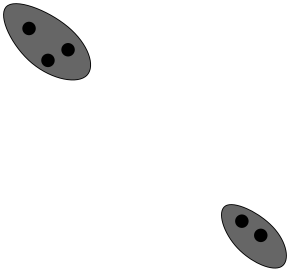
# Nested Sampling

## Graphical aid



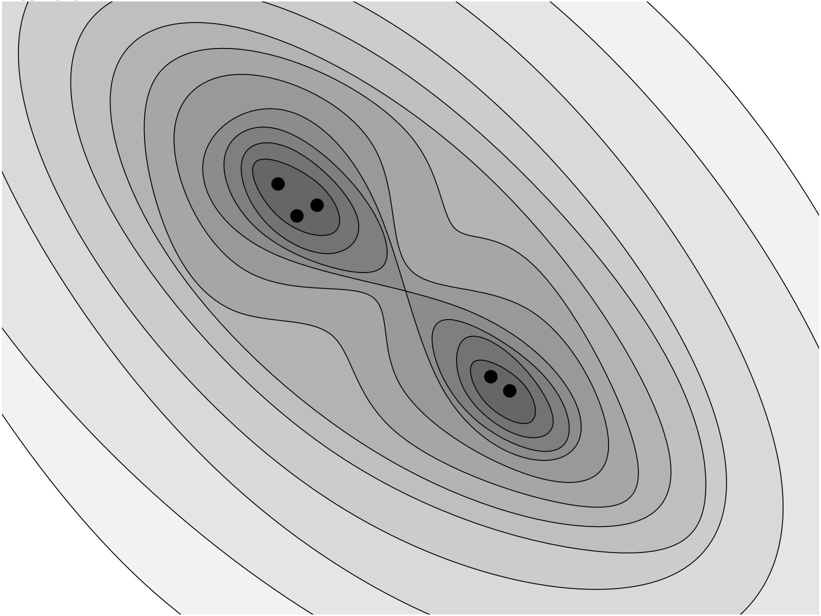
# Nested Sampling

## Graphical aid



# Nested Sampling

## Graphical aid



# Calculating evidences

# Calculating evidences

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta$$

# Calculating evidences

- Transform to 1 dimensional integral

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta$$

# Calculating evidences

- Transform to 1 dimensional integral $\pi(\theta)d\theta = dX$

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta$$

# Calculating evidences

- Transform to 1 dimensional integral $\pi(\theta)d\theta = dX$

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta = \int \mathcal{L}(X)dX$$

# Calculating evidences

- Transform to 1 dimensional integral $\pi(\theta)d\theta = dX$

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta = \int \mathcal{L}(X)dX$$

- $X$ is the *prior volume*

# Calculating evidences

- Transform to 1 dimensional integral $\pi(\theta)d\theta = dX$

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta = \int \mathcal{L}(X)dX$$

- $X$ is the *prior volume*

$$X(\mathcal{L}) = \int_{\mathcal{L}(\theta)>\mathcal{L}} \pi(\theta)d\theta$$

# Calculating evidences

- Transform to 1 dimensional integral $\pi(\theta)d\theta = dX$

$$\mathcal{Z} = \int \mathcal{L}(\theta)\pi(\theta)d\theta = \int \mathcal{L}(X)dX$$

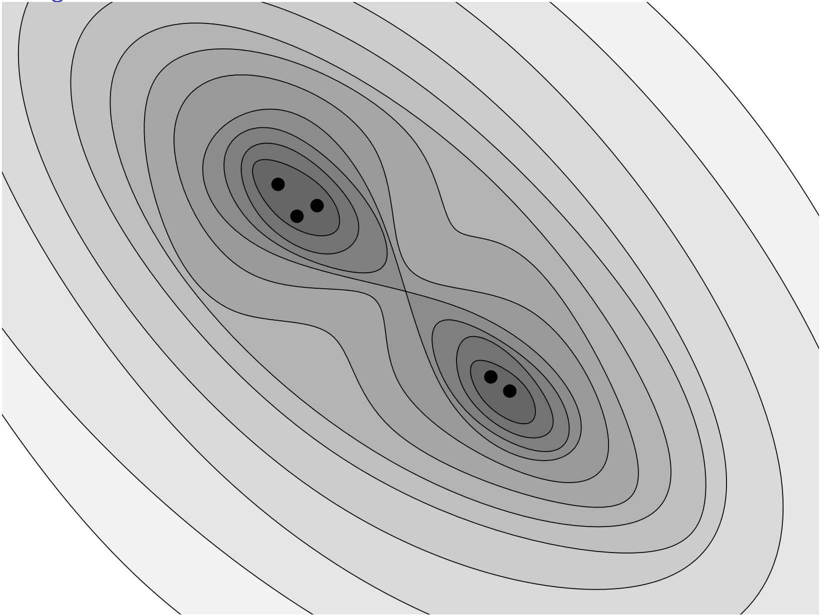
- $X$ is the *prior volume*

$$X(\mathcal{L}) = \int_{\mathcal{L}(\theta)>\mathcal{L}} \pi(\theta)d\theta$$

- i.e. the fraction of the prior which the iso-likelihood contour $\mathcal{L}$ encloses.

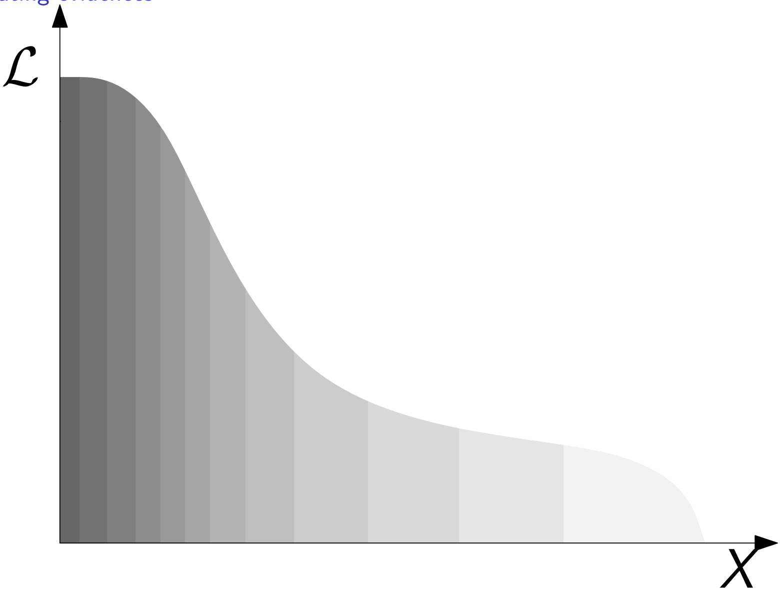
# Nested Sampling

## Calculating evidences



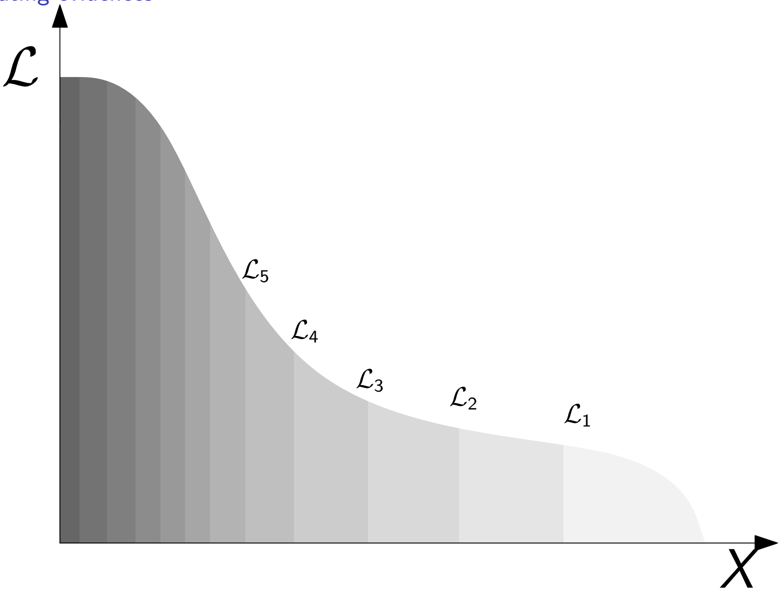
# Nested Sampling

## Calculating evidences



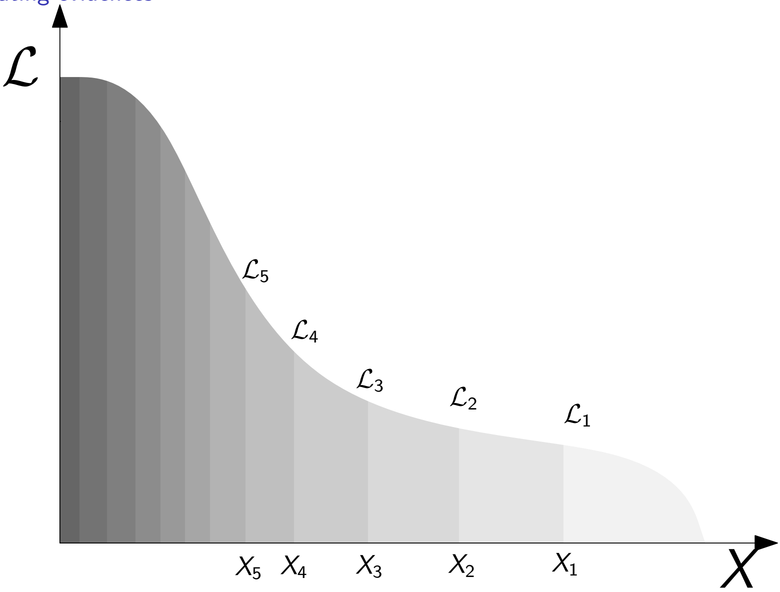
# Nested Sampling

## Calculating evidences



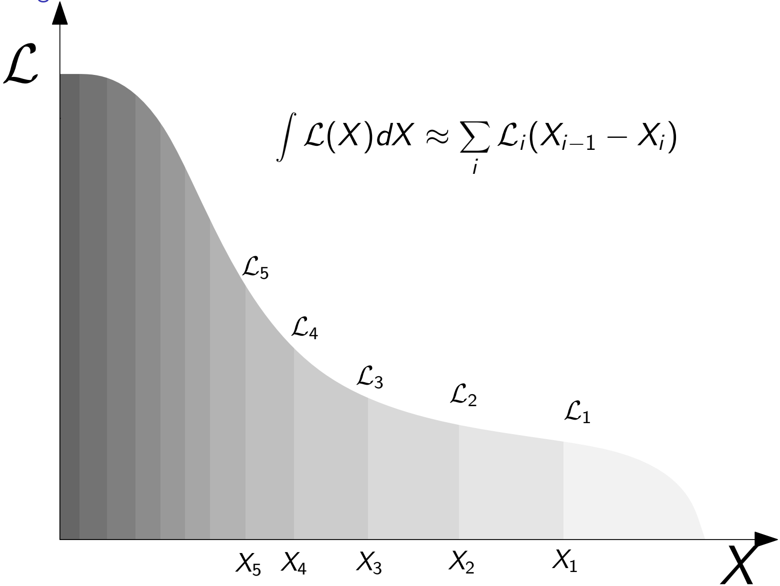
# Nested Sampling

## Calculating evidences



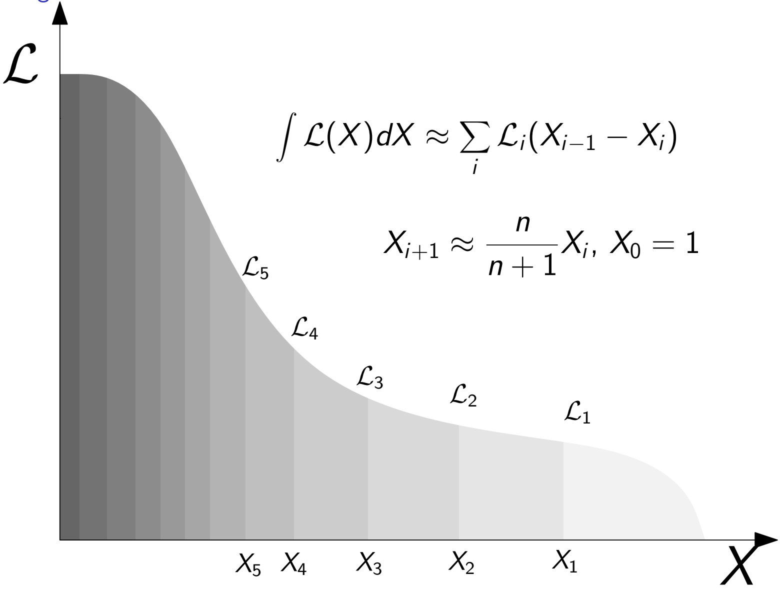
# Nested Sampling

## Calculating evidences



# Nested Sampling

## Calculating evidences




$$\int \mathcal{L}(X)dX \approx \sum_i \mathcal{L}_i(X_{i-1} - X_i)$$

$$X_{i+1} \approx \frac{n}{n+1}X_i,\ X_0 = 1$$

# Nested Sampling

## Exponential volume contraction

# Nested Sampling

## Exponential volume contraction

- At each iteration, the likelihood contour will shrink in volume by a factor of $\approx 1/n$.

# Nested Sampling

## Exponential volume contraction

- At each iteration, the likelihood contour will shrink in volume by a factor of $\approx 1/n$.
- Nested sampling zooms in to the peak of the posterior *exponentially*.

# Nested Sampling

## Exponential volume contraction

- At each iteration, the likelihood contour will shrink in volume by a factor of $\approx 1/n$.
- Nested sampling zooms in to the peak of the posterior *exponentially*.
- $$\mathcal{Z} \approx \sum_i \mathcal{L}_i (X_{i-1} - X_i) \tag{1}$$

# Nested Sampling

## Exponential volume contraction

- At each iteration, the likelihood contour will shrink in volume by a factor of $\approx 1/n$.
- Nested sampling zooms in to the peak of the posterior *exponentially*.
- $$\mathcal{Z} \approx \sum_i \mathcal{L}_i (X_{i-1} - X_i) \tag{1}$$
- $$X_{i+1} \approx \frac{n}{n+1} X_i, \qquad X_0 = 1 \tag{2}$$

# Nested sampling

## Parameter estimation

# Nested sampling

## Parameter estimation

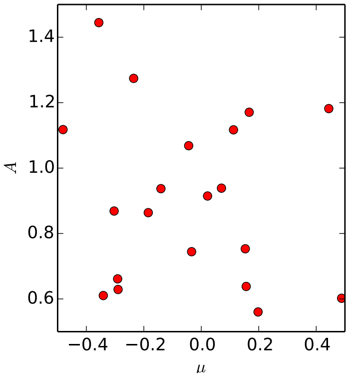
- NS can also be used to sample the posterior

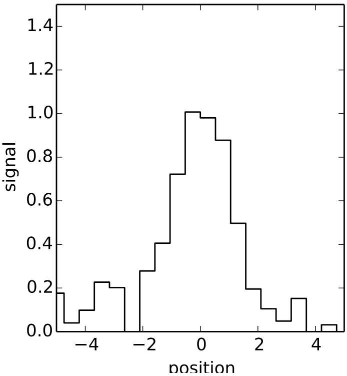
# Nested sampling

## Parameter estimation

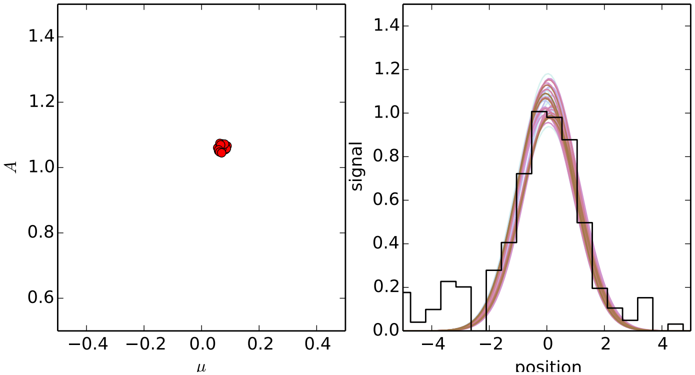
- NS can also be used to sample the posterior
- The set of dead points are posterior samples with an appropriate weighting factor

# When NS succeeds

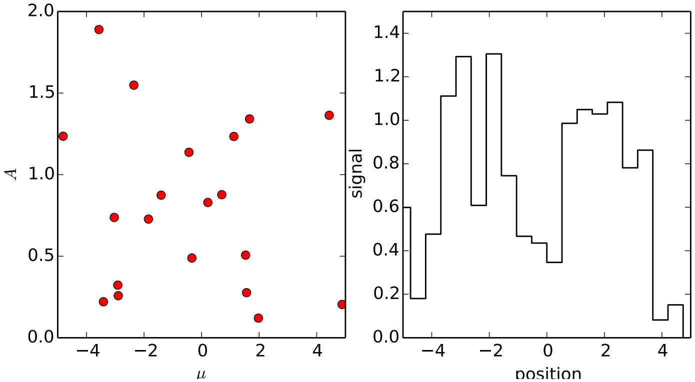


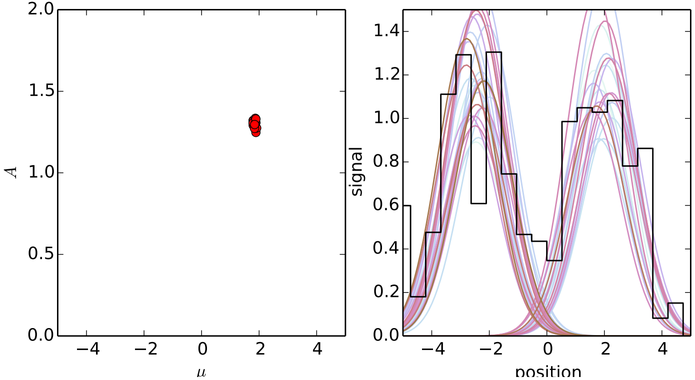
# When NS suceeds



# When NS succeeds



# When NS suceeds



# Sampling from a hard likelihood constraint

# Sampling from a hard likelihood constraint

> *"It is not the purpose of this introductory paper to develop the technology of navigation within such a volume. We merely note that exploring a hard-edged likelihood-constrained domain should prove to be neither more nor less demanding than exploring a likelihood-weighted space."*
>
> — *John Skilling*

# Sampling from a hard likelihood constraint

> *"It is not the purpose of this introductory paper to develop the technology of navigation within such a volume. We merely note that exploring a hard-edged likelihood-constrained domain should prove to be neither more nor less demanding than exploring a likelihood-weighted space."*
>
> — *John Skilling*

- Most of the work in NS to date has been in attempting to implement a hard-edged sampler in the NS meta-algorithm.

# Sampling within an iso-likelihood contour

## Previous attempts

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

- Suffers in high dimensions

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

- Suffers in high dimensions

Hamiltonian sampling F. Feroz & J. Skilling (2013).

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

- Suffers in high dimensions

Hamiltonian sampling F. Feroz & J. Skilling (2013).

- Requires gradients and tuning

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

- Suffers in high dimensions

Hamiltonian sampling F. Feroz & J. Skilling (2013).

- Requires gradients and tuning

Diffusion Nested Sampling B. Brewer et al. (2009).

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

- Suffers in high dimensions

Hamiltonian sampling F. Feroz & J. Skilling (2013).

- Requires gradients and tuning

Diffusion Nested Sampling B. Brewer et al. (2009).

- Very promising

# Sampling within an iso-likelihood contour

## Previous attempts

Rejection Sampling MultiNest; F. Feroz & M. Hobson (2009).

- Suffers in high dimensions

Hamiltonian sampling F. Feroz & J. Skilling (2013).

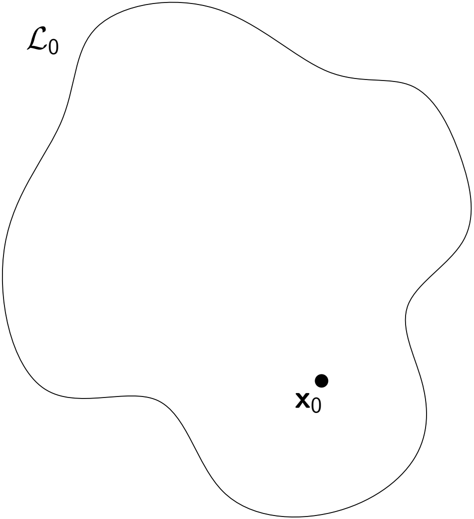
- Requires gradients and tuning

Diffusion Nested Sampling B. Brewer et al. (2009).

- Very promising
- Too many tuning parameters

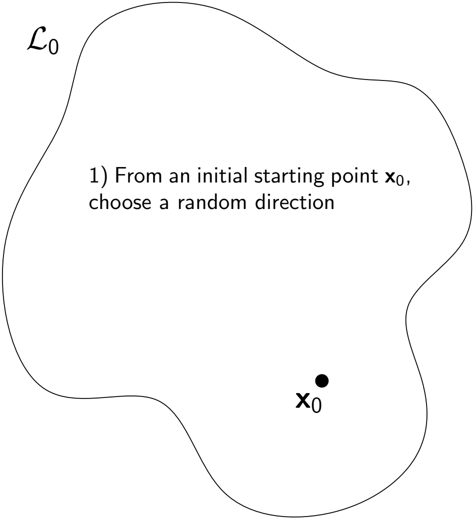
# PolyChord

## "Hit and run" slice sampling



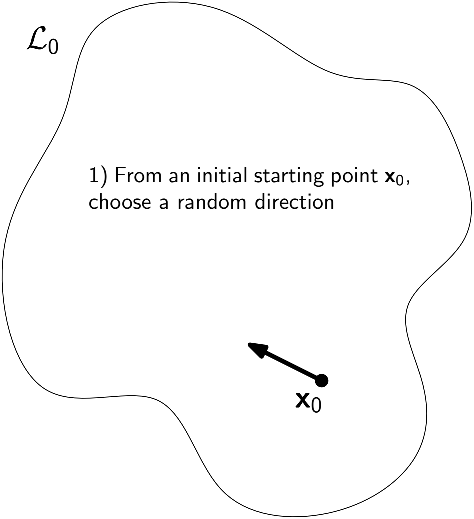
# PolyChord

## "Hit and run" slice sampling



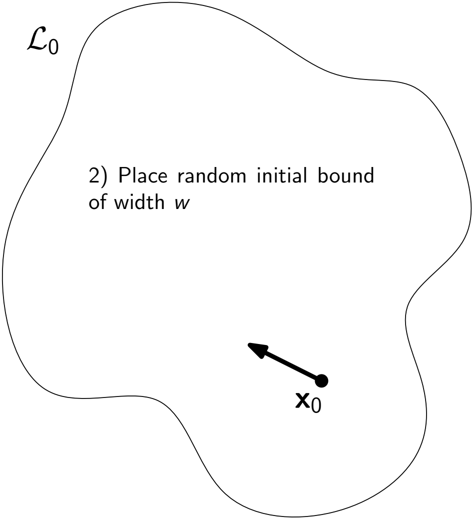
# PolyChord

## "Hit and run" slice sampling



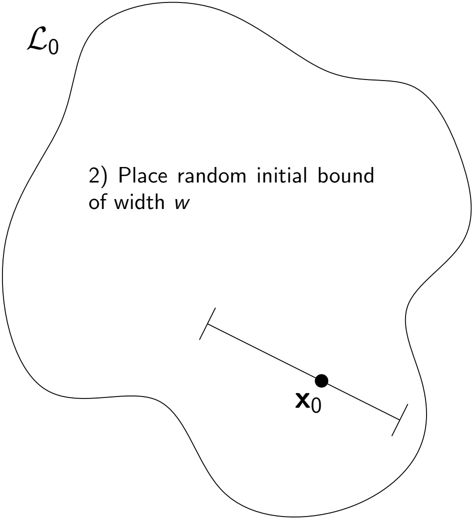
# PolyChord

## "Hit and run" slice sampling



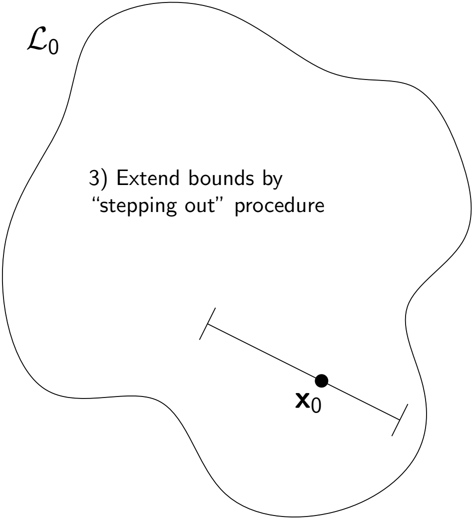
# PolyChord

## "Hit and run" slice sampling



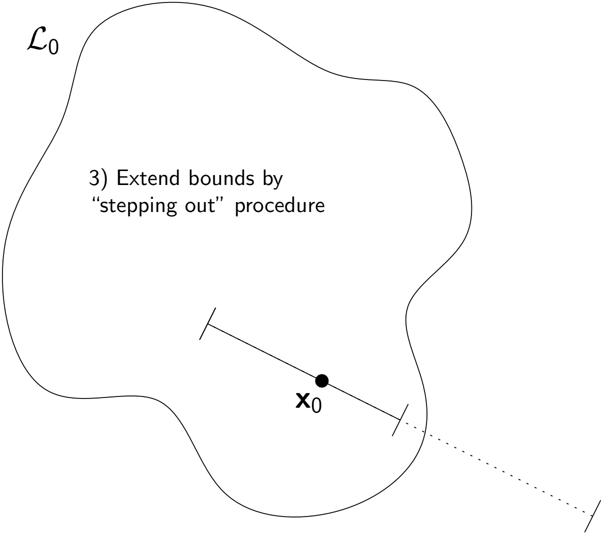
# PolyChord

## "Hit and run" slice sampling



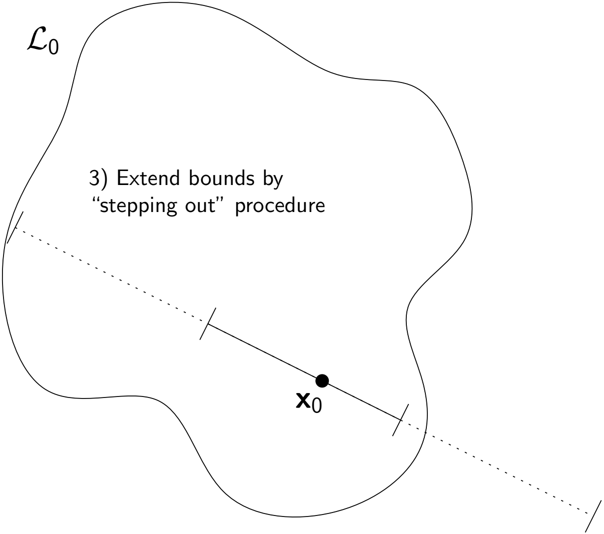
# PolyChord

## "Hit and run" slice sampling



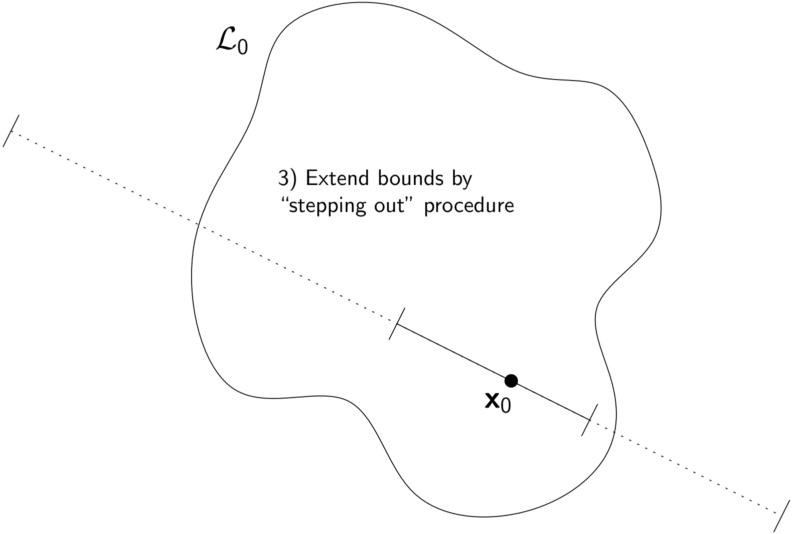
# PolyChord

## "Hit and run" slice sampling



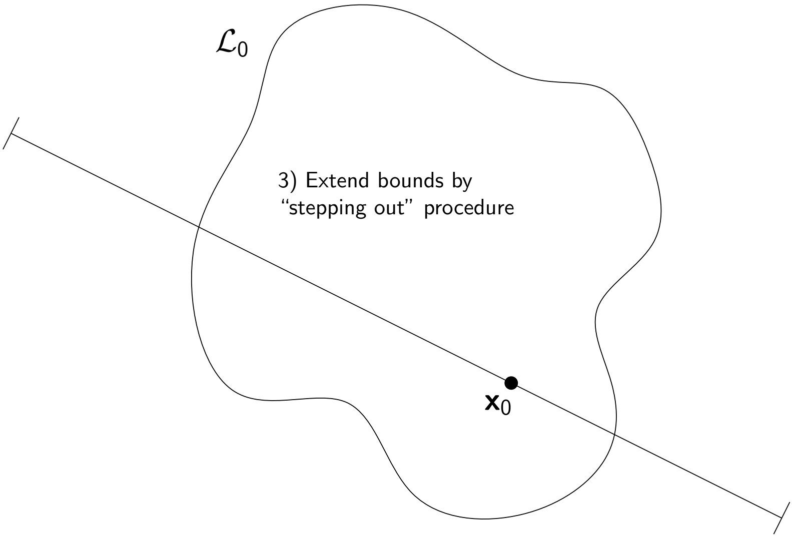
# PolyChord

## "Hit and run" slice sampling



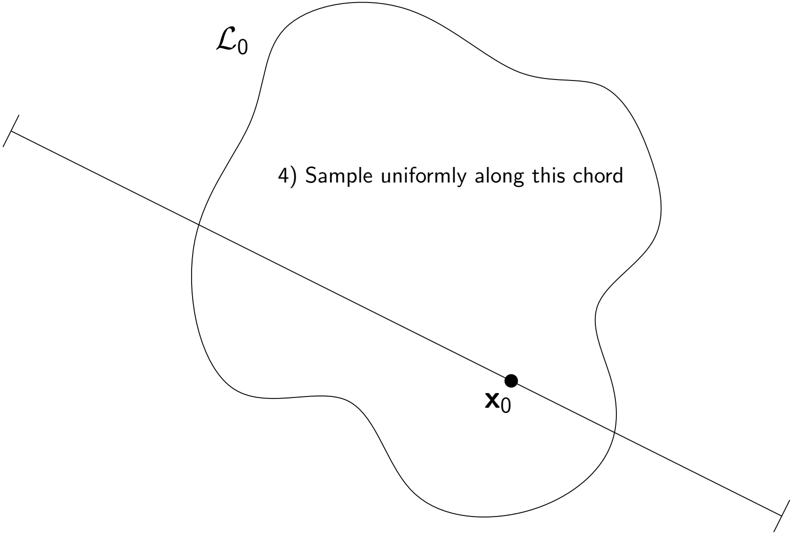
# PolyChord

## "Hit and run" slice sampling



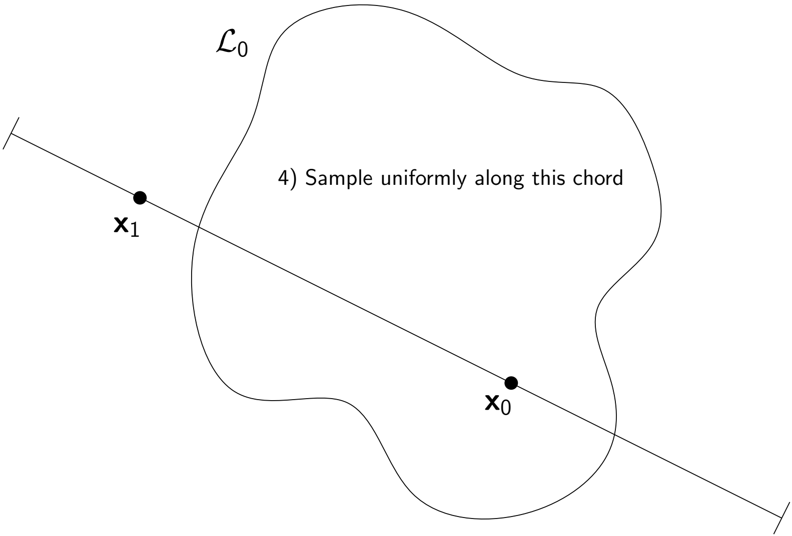
# PolyChord

## "Hit and run" slice sampling



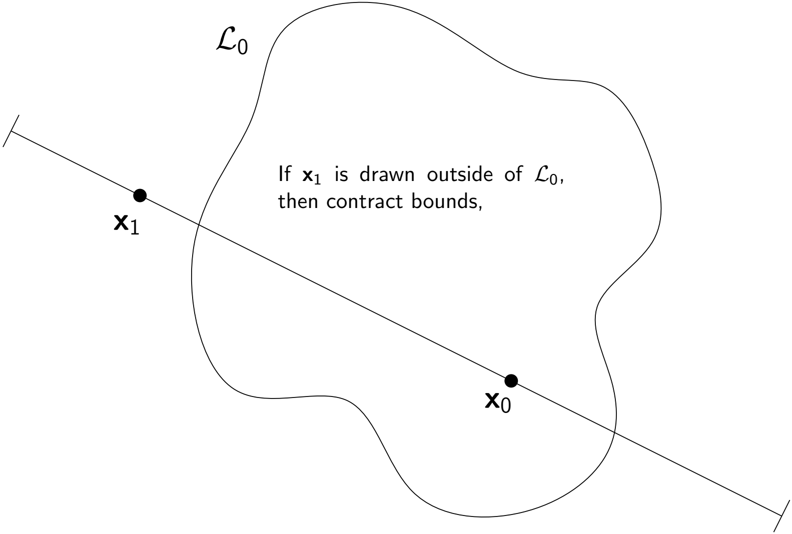
# PolyChord

## "Hit and run" slice sampling



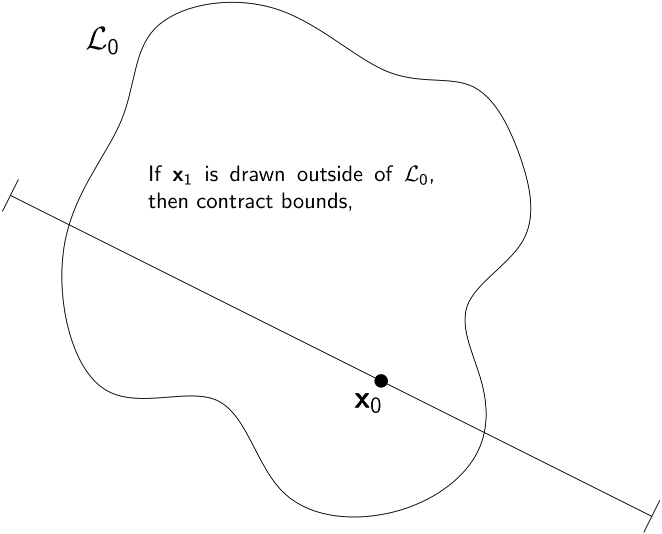
# PolyChord

## "Hit and run" slice sampling



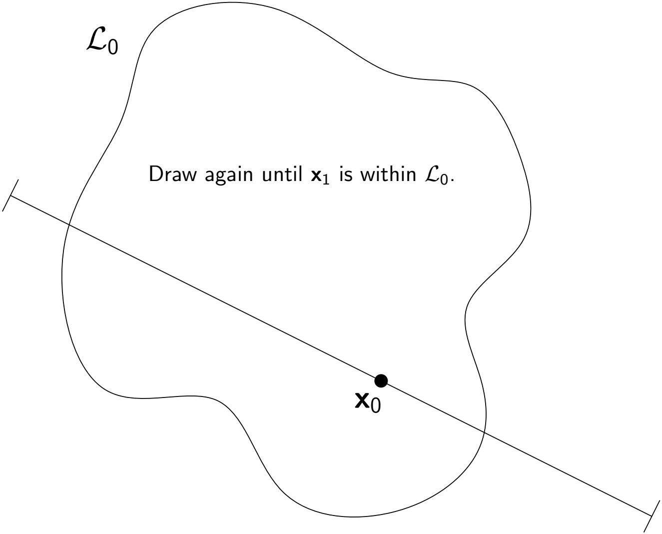
# PolyChord

## "Hit and run" slice sampling



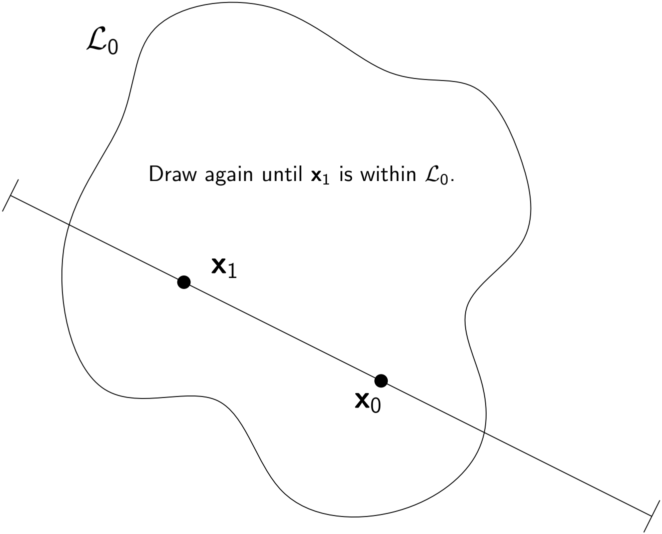
# PolyChord

## "Hit and run" slice sampling



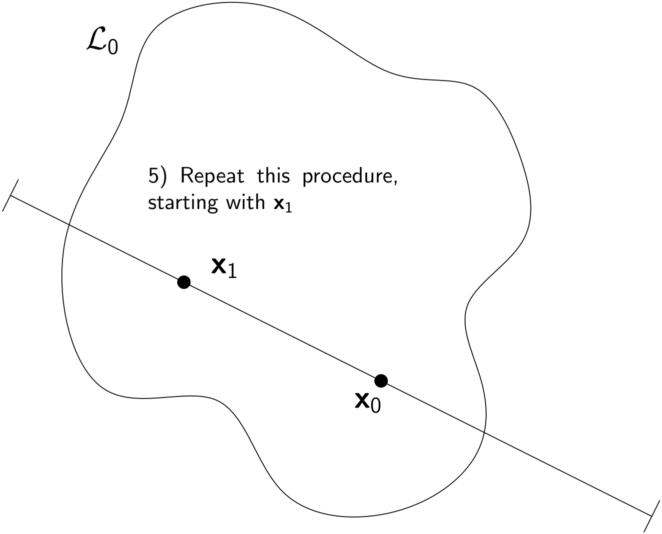
# PolyChord

## "Hit and run" slice sampling



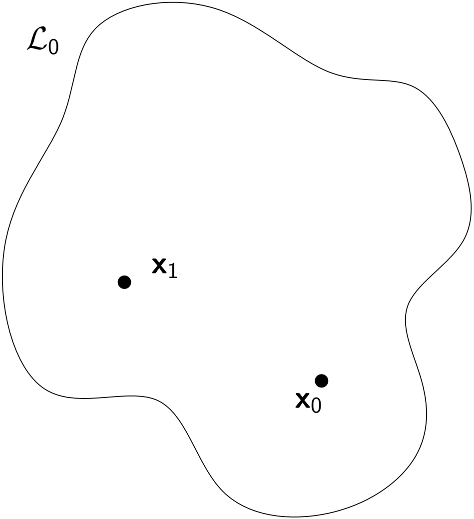
# PolyChord

## "Hit and run" slice sampling



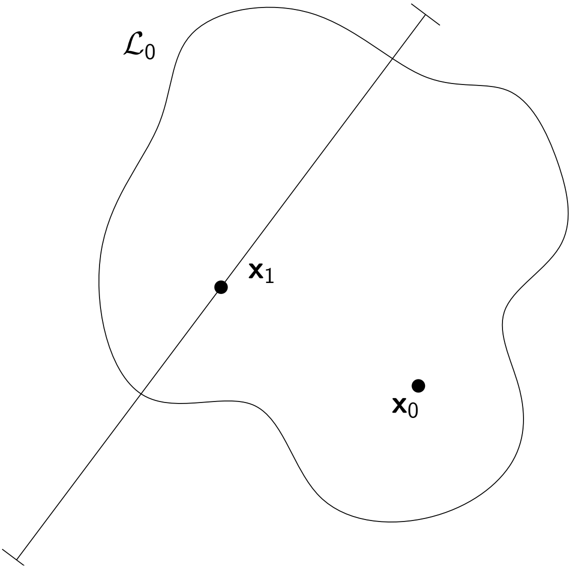
# PolyChord

## "Hit and run" slice sampling



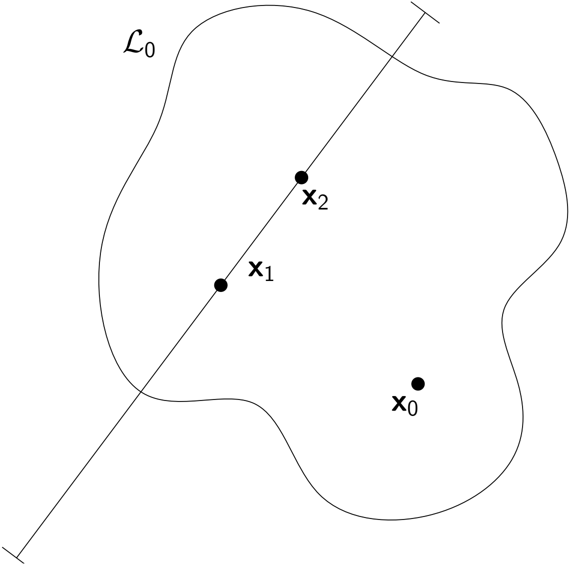
# PolyChord

## "Hit and run" slice sampling



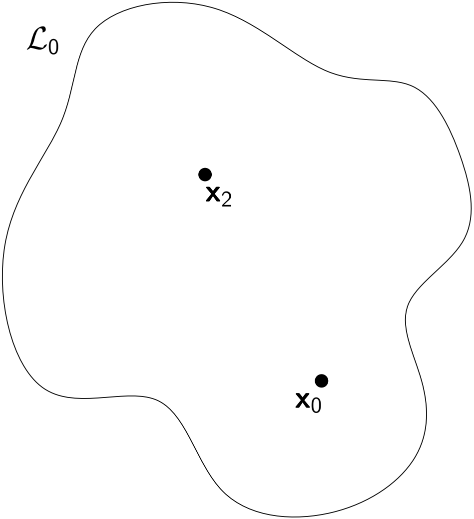
# PolyChord

"Hit and run" slice sampling



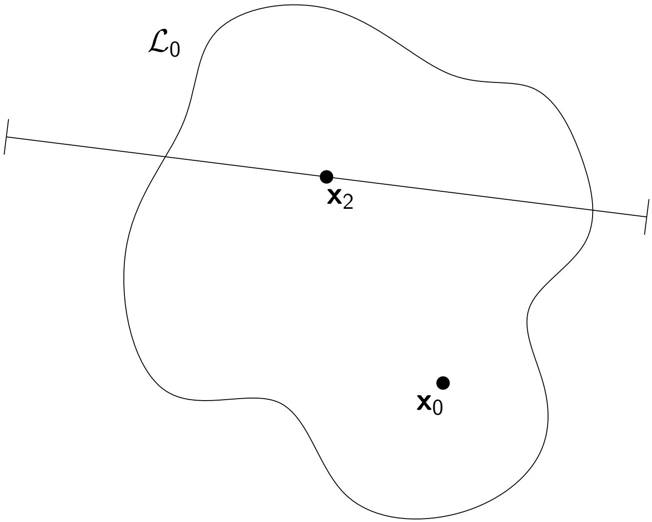
# PolyChord

## "Hit and run" slice sampling



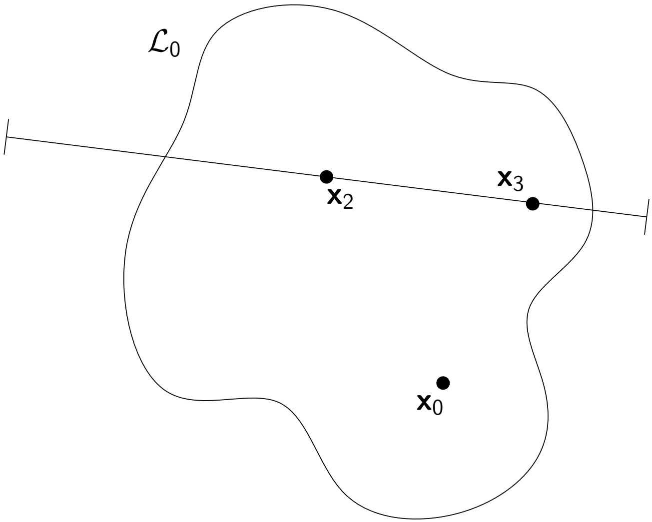
# PolyChord

## "Hit and run" slice sampling



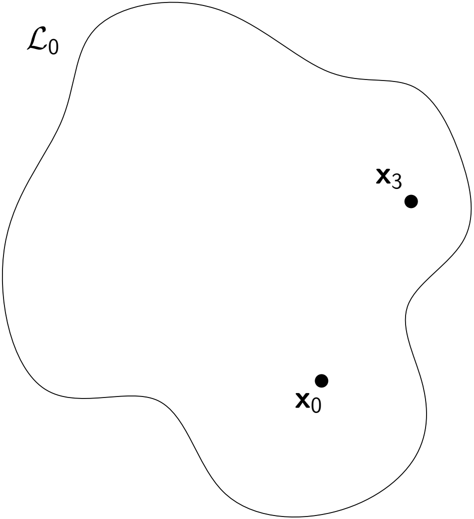
# PolyChord

## "Hit and run" slice sampling



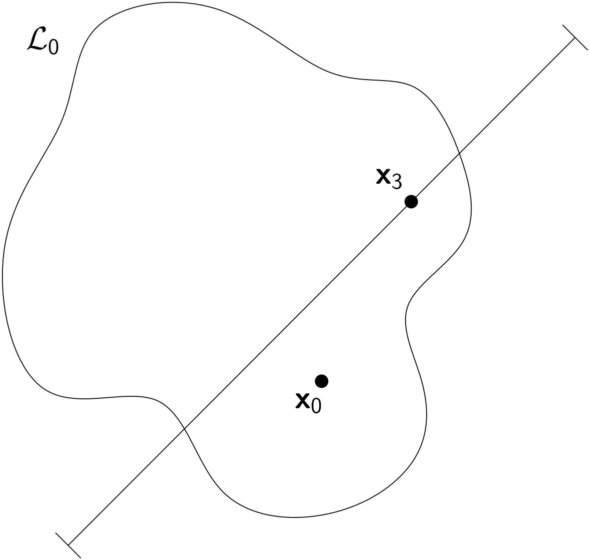
# PolyChord

## "Hit and run" slice sampling



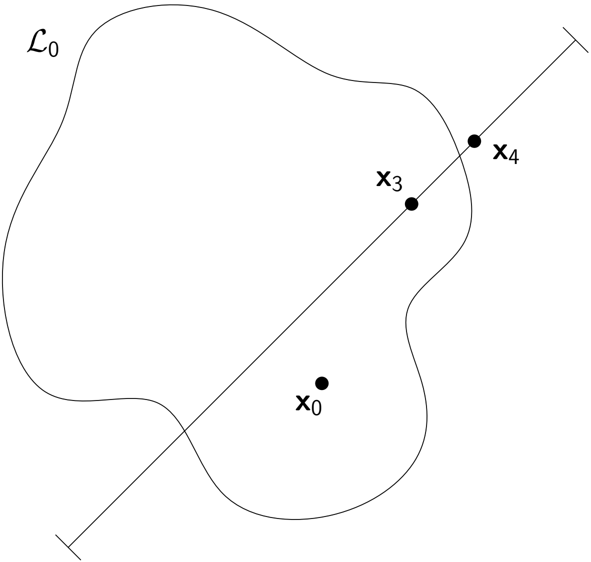
# PolyChord

## "Hit and run" slice sampling



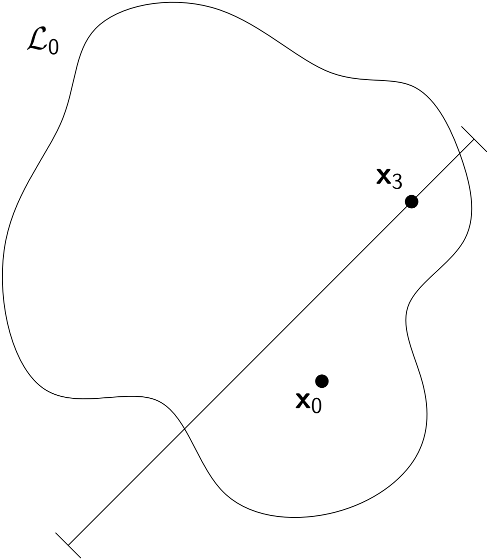
# PolyChord

## "Hit and run" slice sampling



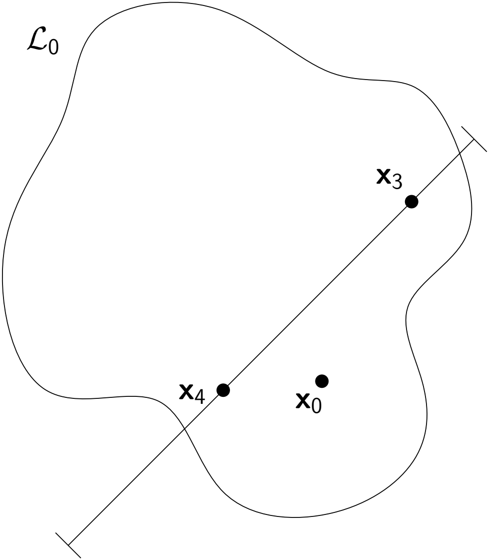
# PolyChord

## "Hit and run" slice sampling



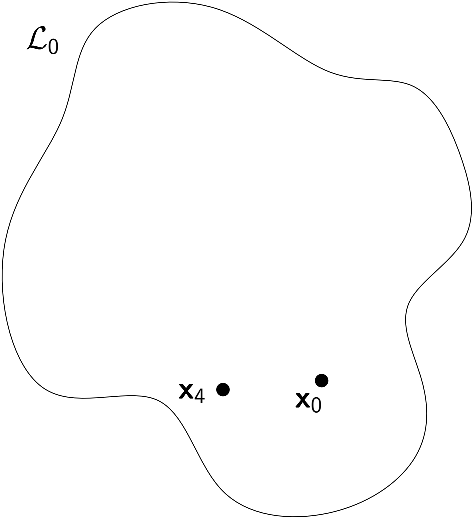
# PolyChord

## "Hit and run" slice sampling



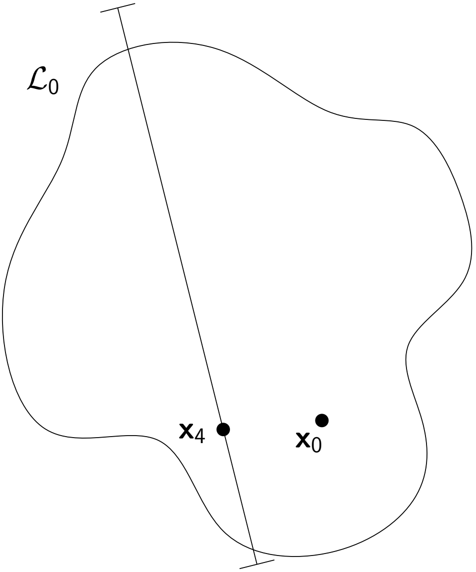
# PolyChord

## "Hit and run" slice sampling



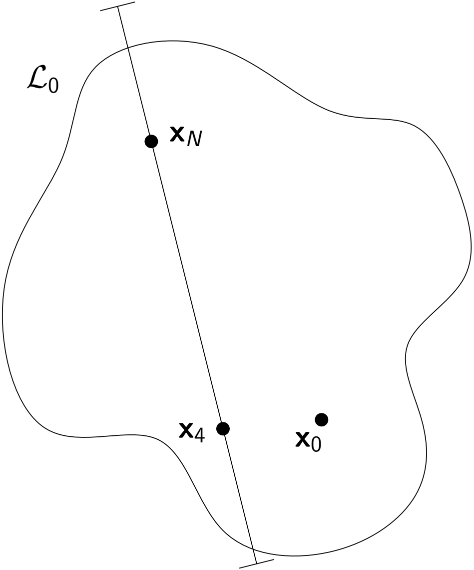
# PolyChord

"Hit and run" slice sampling



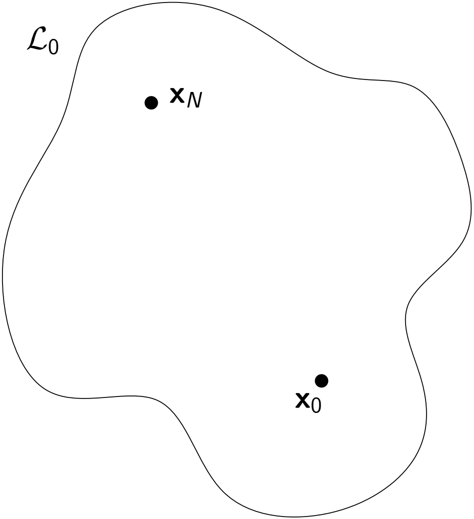
# PolyChord

## "Hit and run" slice sampling



# PolyChord

## "Hit and run" slice sampling



# PolyChord

## Key points

# PolyChord

## Key points

- This procedure satisfies detailed balance.

# PolyChord
## Key points

- This procedure satisfies detailed balance.
  - Works even if $\mathcal{L}_0$ contour is disjoint.

# PolyChord

## Key points

- This procedure satisfies detailed balance.
  - Works even if $\mathcal{L}_0$ contour is disjoint.
- Need $N$ reasonably large $\sim \mathcal{O}(n_{\mathrm{dims}})$ so that $x_N$ is de-correlated from $x_1$.

# Issues with Slice Sampling

# Issues with Slice Sampling

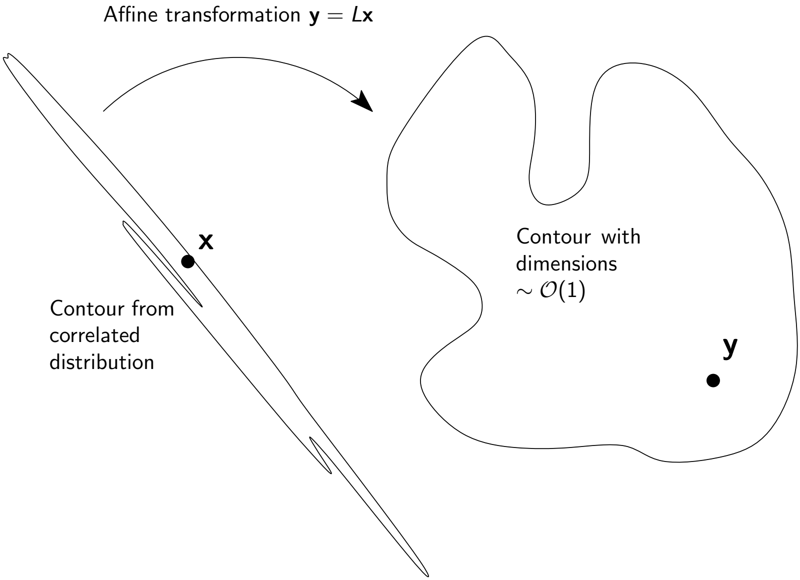
1. Does not deal well with correlated distributions.

# Issues with Slice Sampling

1. Does not deal well with correlated distributions.
2. Need to “tune” $w$ parameter.

# PolyChord's solutions

## Correlated distributions



# PolyChord's solutions

## Correlated distributions

# PolyChord's solutions

## Correlated distributions

- We make an affine transformation to remove degeneracies, and "whiten" the space.

# PolyChord's solutions

## Correlated distributions

- We make an affine transformation to remove degeneracies, and "whiten" the space.
- Samples remain uniformly sampled

# PolyChord's solutions

## Correlated distributions

- We make an affine transformation to remove degeneracies, and "whiten" the space.
- Samples remain uniformly sampled
- We use the covariance matrix of the live points and all inter-chain points

# PolyChord's solutions

## Correlated distributions

- We make an affine transformation to remove degeneracies, and "whiten" the space.
- Samples remain uniformly sampled
- We use the covariance matrix of the live points and all inter-chain points
- Cholesky decomposition is the required skew transformation

# PolyChord's solutions

## Correlated distributions

- We make an affine transformation to remove degeneracies, and "whiten" the space.
- Samples remain uniformly sampled
- We use the covariance matrix of the live points and all inter-chain points
- Cholesky decomposition is the required skew transformation
- $w = 1$ in this transformed space

# PolyChord's Additions

# PolyChord's Additions

- Parallelised up to number of live points with openMPI.

# PolyChord's Additions

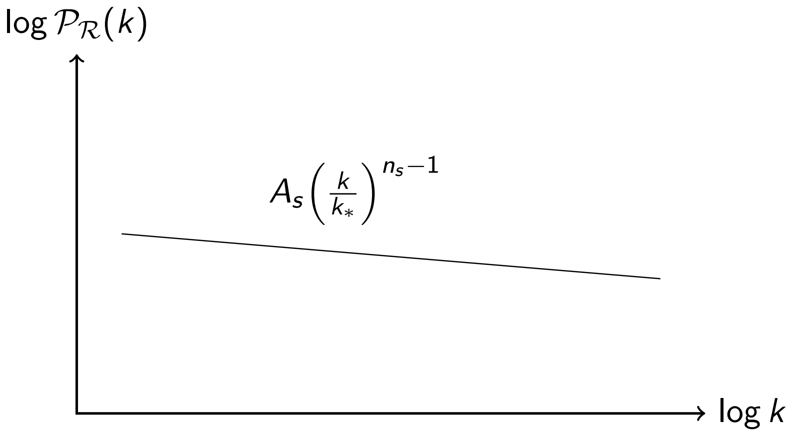
- Parallelised up to number of live points with openMPI.
- Novel method for identifying and evolving modes separately.

# PolyChord's Additions

- Parallelised up to number of live points with openMPI.
- Novel method for identifying and evolving modes separately.
- Implemented in CosmoMC, as "CosmoChord", with fast-slow parameters.

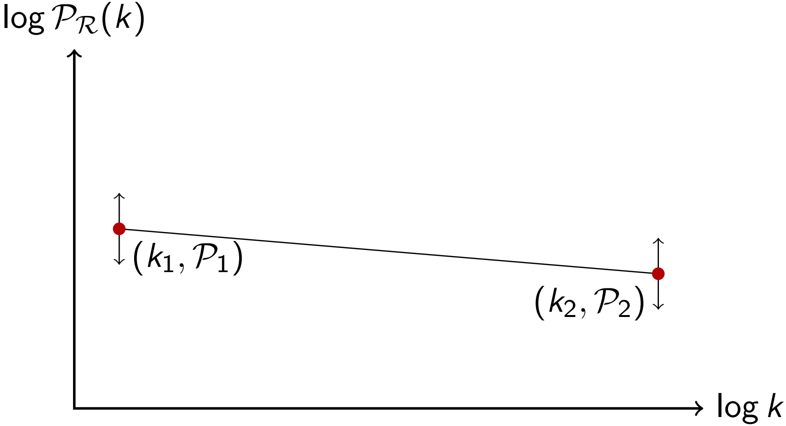
# PolyChord in action

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



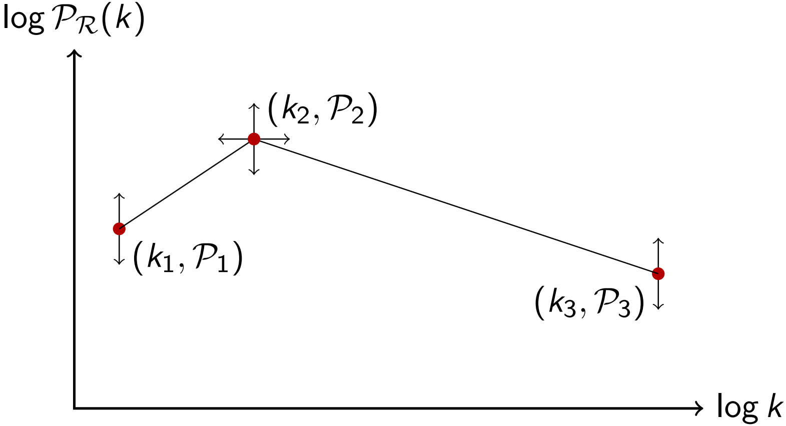
# PolyChord in action

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



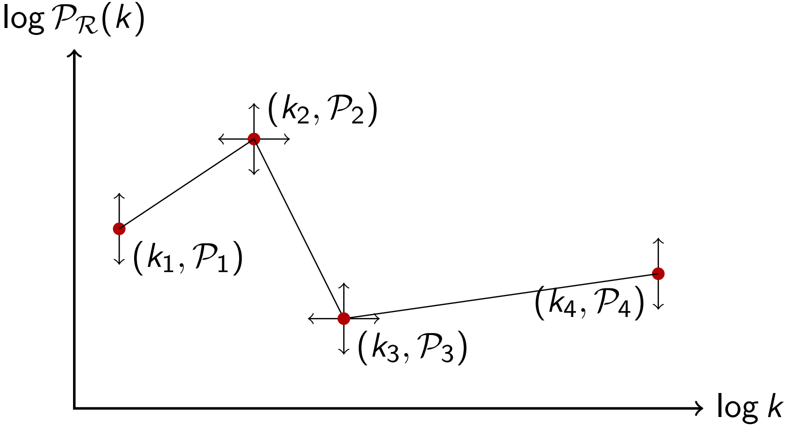
# PolyChord in action

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



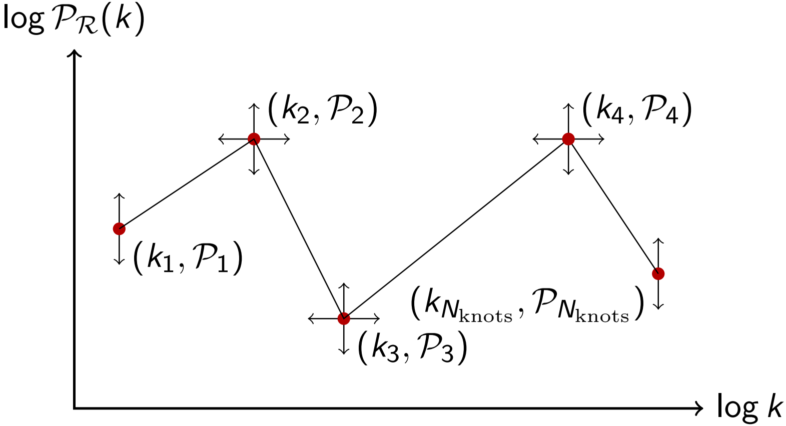
# PolyChord in action

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



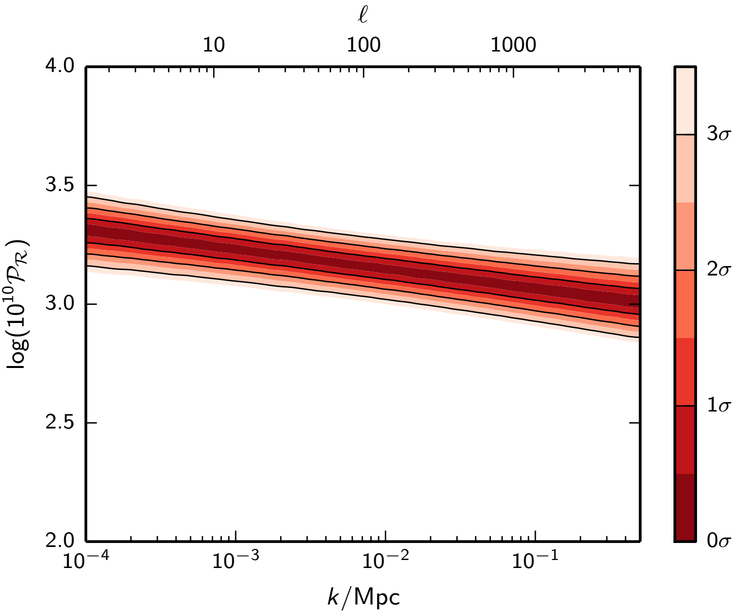
# PolyChord in action

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



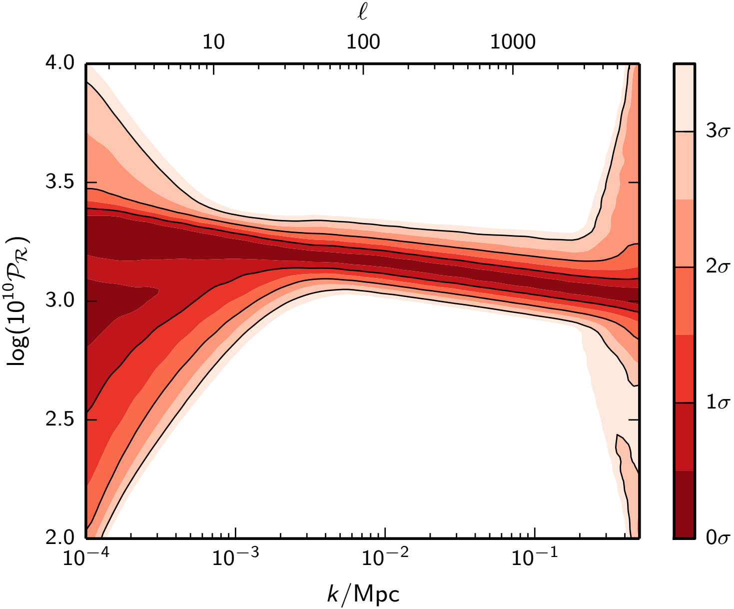
# 0 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



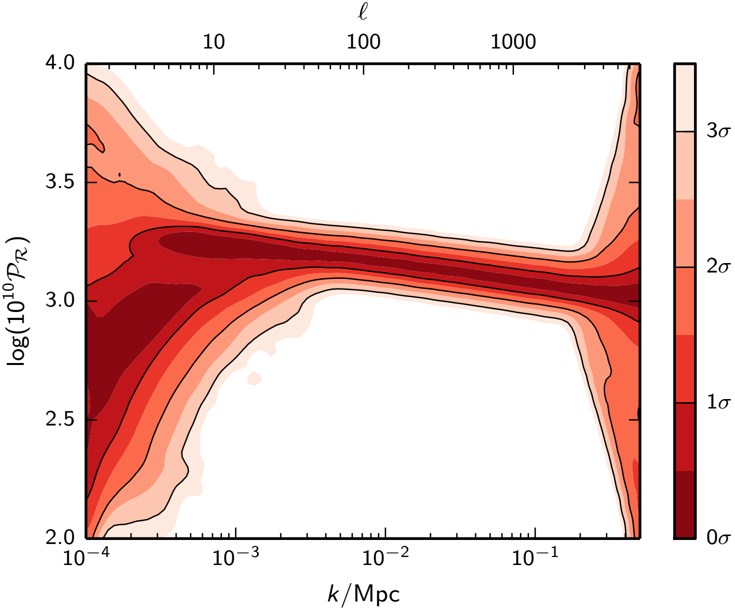
# 1 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



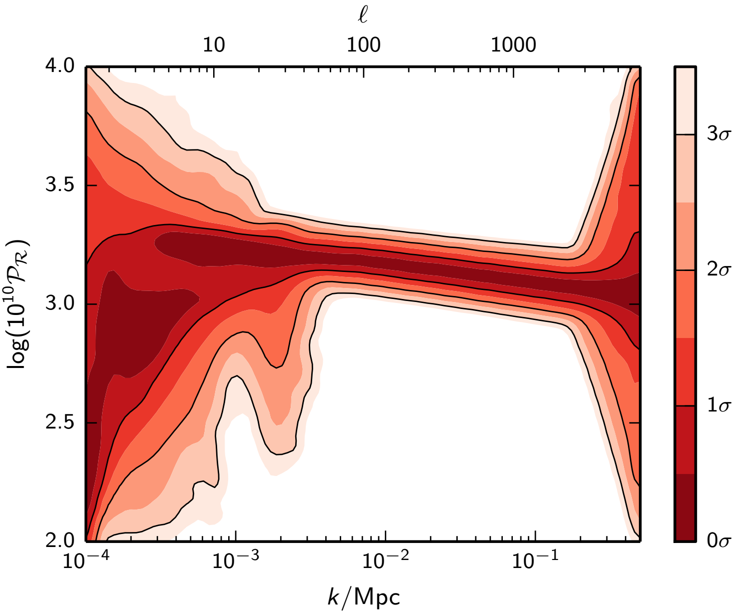
# 2 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



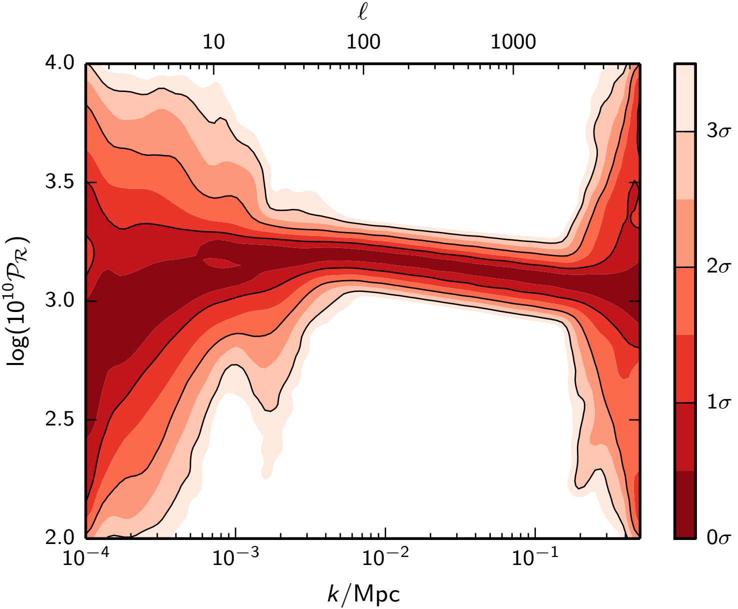
# 3 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



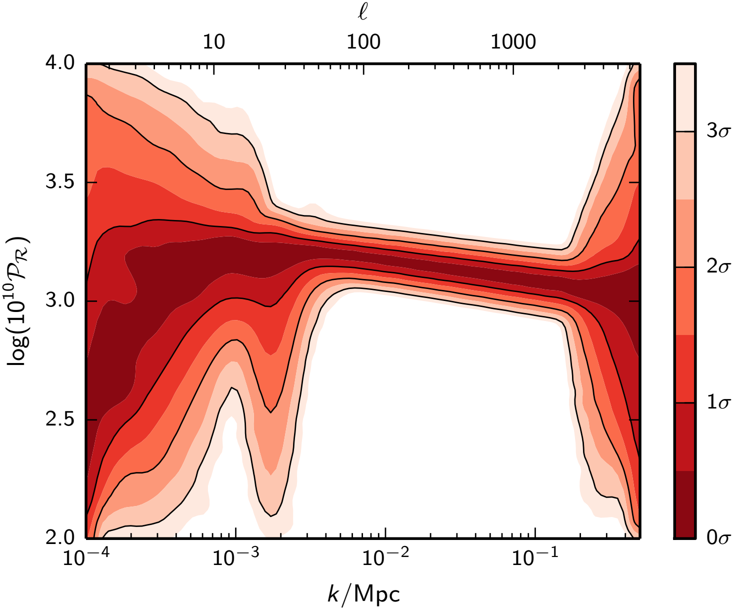
# 4 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



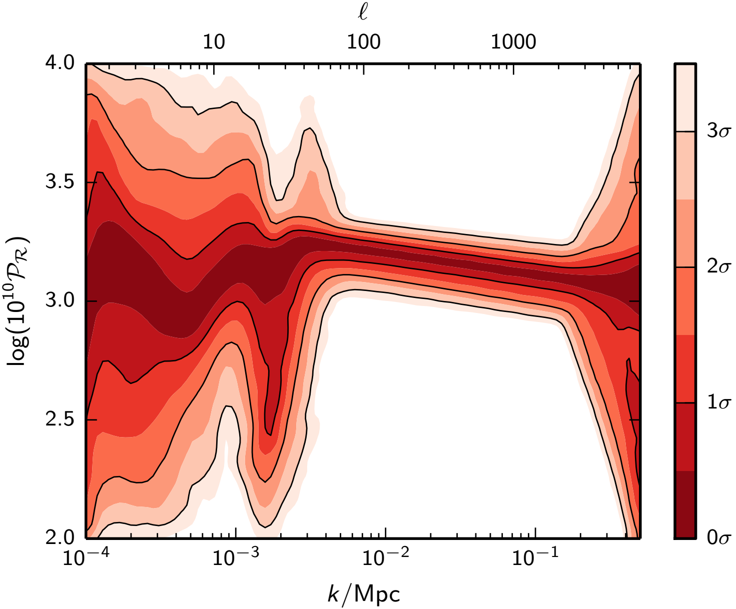
# 5 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



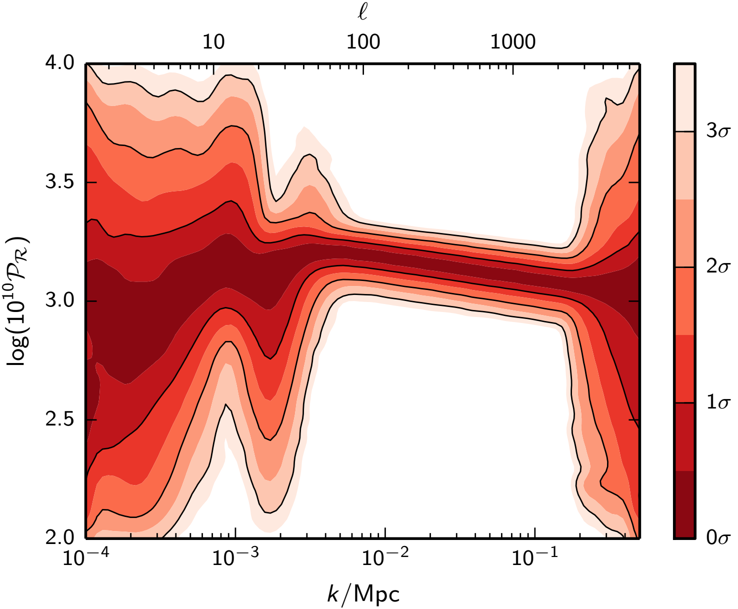
# 6 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



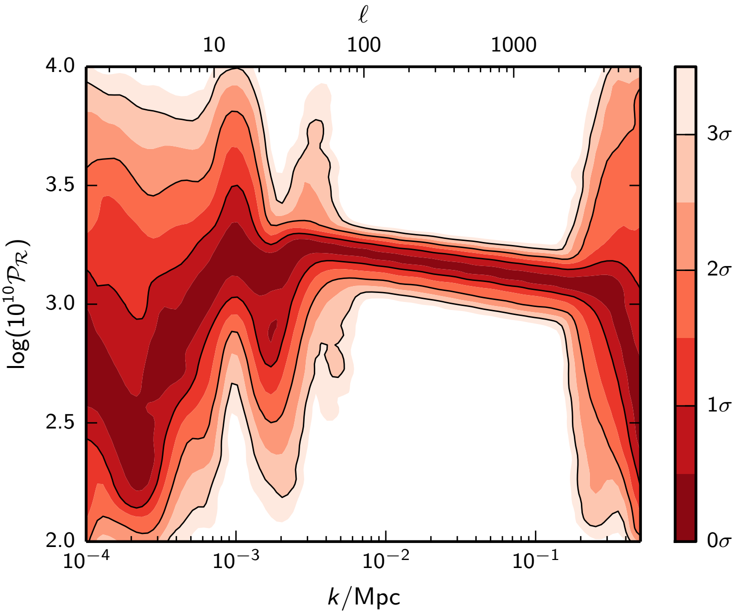
# 7 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



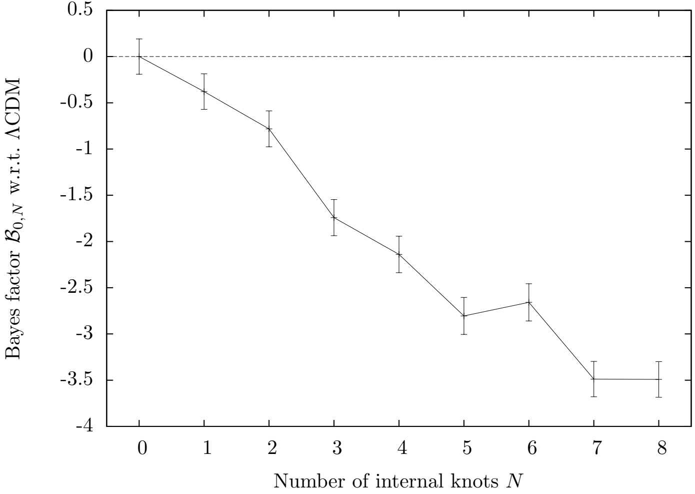
# 8 internal knots

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



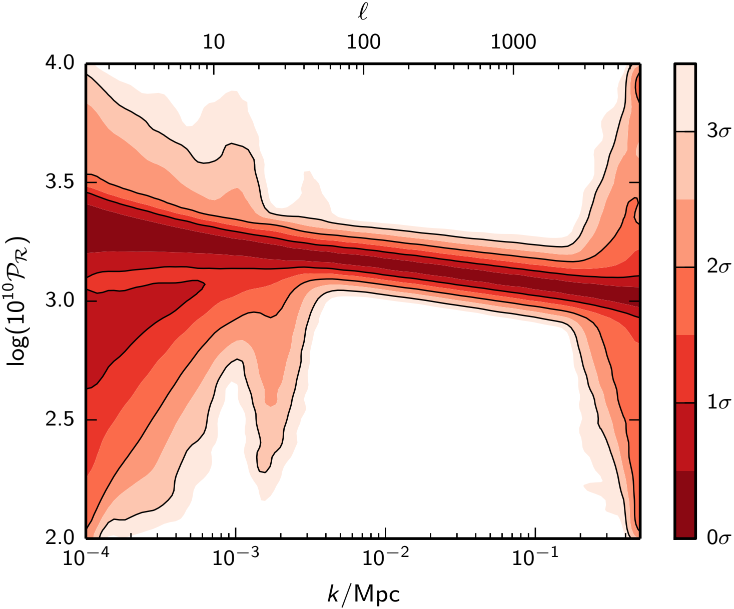
# Bayes Factors

## Primordial power spectrum $\mathcal{P}_\mathcal{R}(k)$ reconstruction



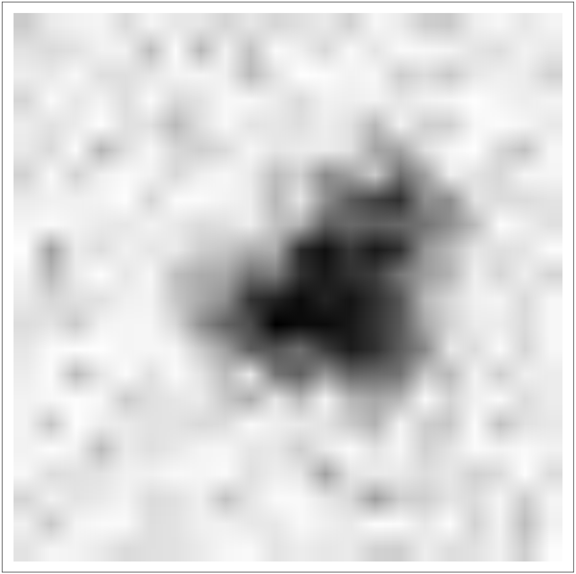
# Marginalised plot

## Primordial power spectrum $\mathcal{P}_{\mathcal{R}}(k)$ reconstruction



# Object detection

## Toy problem



# Object detection

## Evidences

# Object detection

## Evidences

- $\log \mathcal{Z}$ ratio: $-251 : -156 : -114 : -117 : -136$

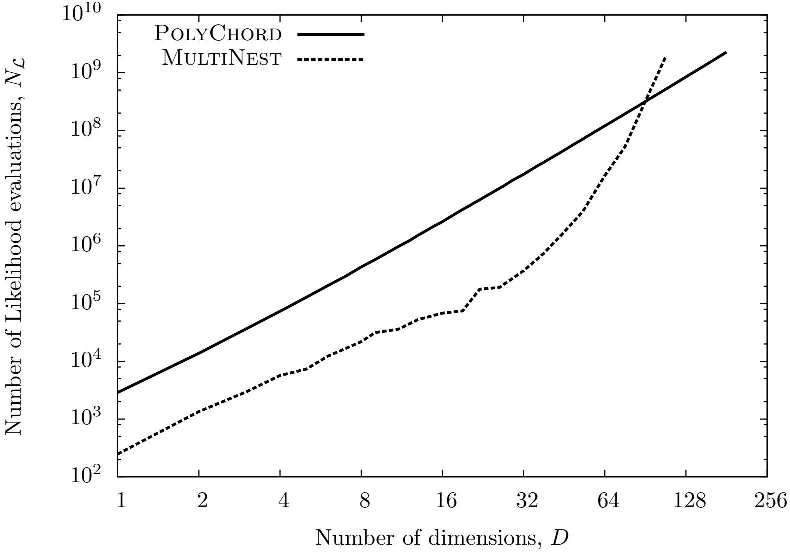
# Object detection

## Evidences

- $\log \mathcal{Z}$ ratio: $-251 : -156 : -114 : -117 : -136$
- odds ratio: $10^{-60} : 10^{-19} : 1 : 0.04 : 10^{-10}$

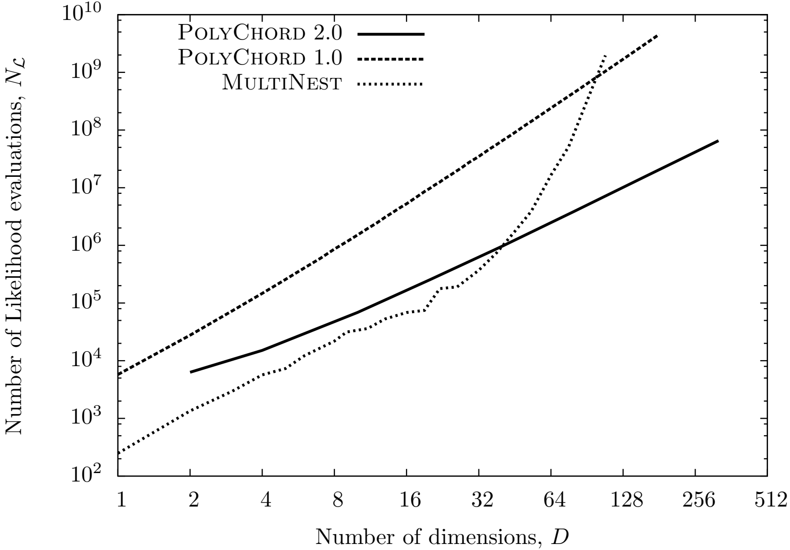
# PolyChord vs. MultiNest

## Gaussian likelihood



# PolyChord vs. MultiNest

## Gaussian likelihood



# The future of nested sampling

# The future of nested sampling

- All previous approaches have focussed on trying to implement John's theoretical algorithm (still some work to be done here).

# The future of nested sampling

- All previous approaches have focussed on trying to implement John's theoretical algorithm (still some work to be done here).
- Nested sampling is really the first in a new class of "probabilistic integration" algorithms.

# Adjusting the meta-algorithm

# Adjusting the meta-algorithm

- Need one sample from prior each time?

# Adjusting the meta-algorithm

- Need one sample from prior each time?
- Adaptive numbers of live points?

# Adjusting the meta-algorithm

- Need one sample from prior each time?
- Adaptive numbers of live points?
- Diffusive nested sampling?

# Adjusting the meta-algorithm

- Need one sample from prior each time?
- Adaptive numbers of live points?
- Diffusive nested sampling?
- Parallelisation?

# Conclusions

## The future of nested sampling

# Conclusions

## The future of nested sampling

- We are at the beginning of a new era of sampling algorithms

# Conclusions

## The future of nested sampling

- We are at the beginning of a new era of sampling algorithms
- Plenty of more work in to be done in exploring new versions of nested sampling

# Conclusions

## The future of nested sampling

- We are at the beginning of a new era of sampling algorithms
- Plenty of more work in to be done in exploring new versions of nested sampling
- Nested sampling is just the beginning

# Conclusions

## The future of nested sampling

- We are at the beginning of a new era of sampling algorithms
- Plenty of more work in to be done in exploring new versions of nested sampling
- Nested sampling is just the beginning
- arXiv:1506.00171

# Conclusions

## The future of nested sampling

- We are at the beginning of a new era of sampling algorithms
- Plenty of more work in to be done in exploring new versions of nested sampling
- Nested sampling is just the beginning
- arXiv:1506.00171
- http://ccpforge.cse.rl.ac.uk/gf/project/polychord/