Dynamic Nested Sampling with *dynesty*

Josh Speagle

jspeagle@cfa.harvard.edu

CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

Banting
Postdoctoral Fellowships

UNIVERSITY OF TORONTO

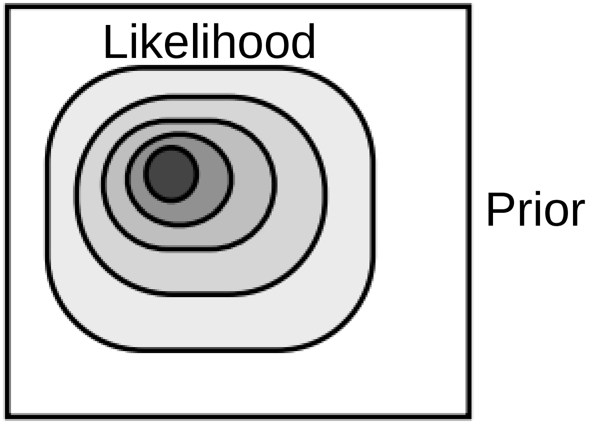
DUNLAP INSTITUTE for ASTRONOMY & ASTROPHYSICS

Background

$$\underbrace{\Pr(\boldsymbol{\Theta}|\mathbf{D},\mathrm{M})}^{\text{Posterior}} = \frac{\overbrace{\Pr(\mathbf{D}|\boldsymbol{\Theta},\mathrm{M})}^{\text{Likelihood}}\ \overbrace{\Pr(\boldsymbol{\Theta}|\mathrm{M})}^{\text{Prior}}}{\underbrace{\Pr(\mathbf{D}|\mathrm{M})}_{\text{Evidence}}}$$

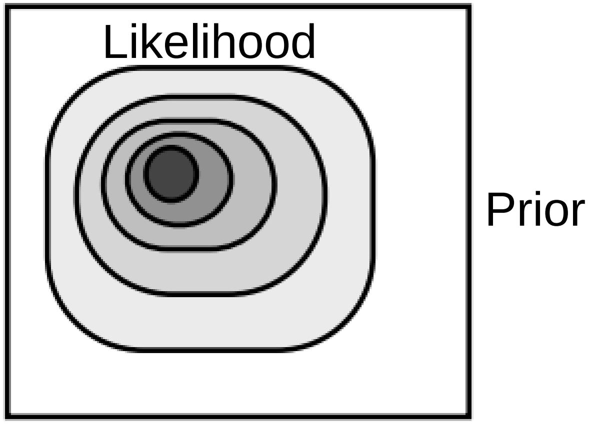
Bayes' Theorem

Motivation: Sampling the Posterior



Pictures adapted from [this 2010 talk](#) by John Skilling.

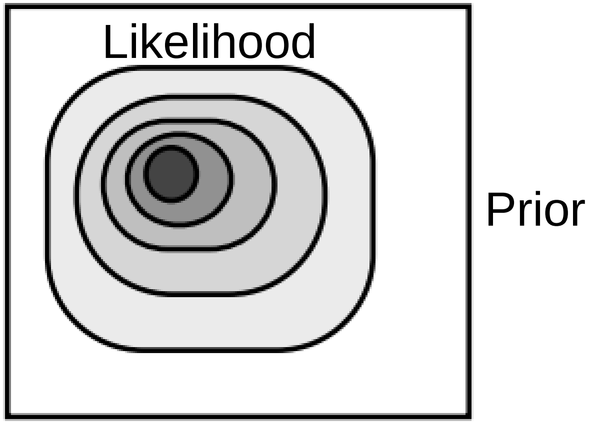
Motivation: Sampling the Posterior



Sampling directly from the likelihood $\mathcal{L}(\boldsymbol{\Theta})$ is **hard**.

Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior



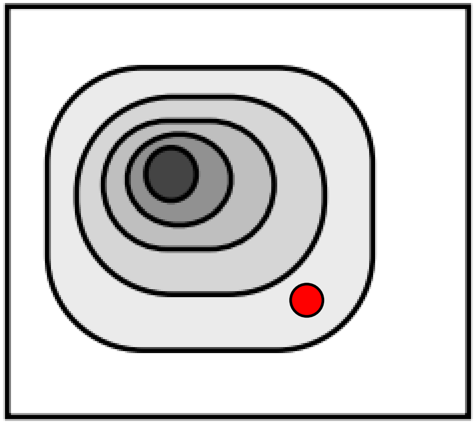
Sampling directly from the likelihood $\mathcal{L}(\mathbf{\Theta})$ is **hard**.

MCMC: Solving a Hard Problem once.

(Markov Chain Monte Carlo)

Pictures adapted from [this 2010 talk](#) by John Skilling.

Motivation: Sampling the Posterior



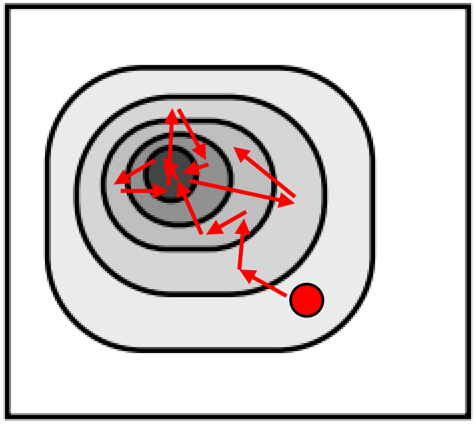
Sampling directly from the likelihood $\mathcal{L}(\boldsymbol{\Theta})$ is **hard**.

MCMC: Solving a Hard Problem once.

(Markov Chain Monte Carlo)

Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior



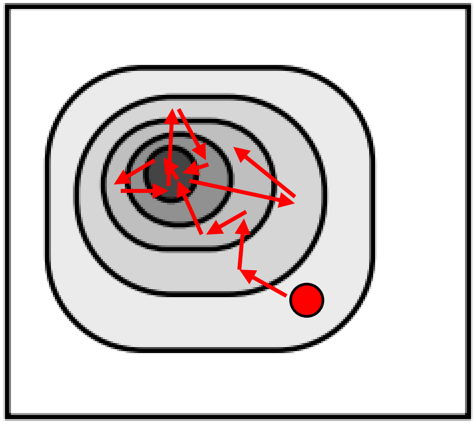
Sampling directly from the likelihood $\mathcal{L}(\boldsymbol{\Theta})$ is **hard**.

MCMC: Solving a Hard Problem once.

(Markov Chain Monte Carlo)

Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior

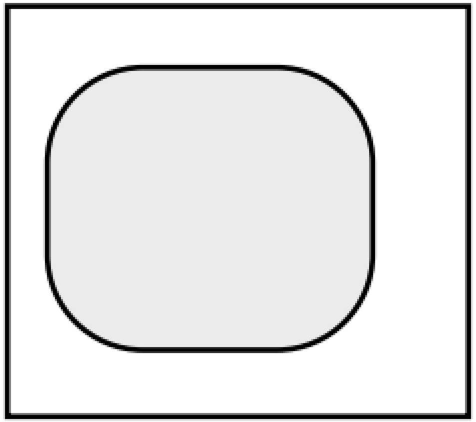


MCMC: Solving a Hard Problem once.

vs

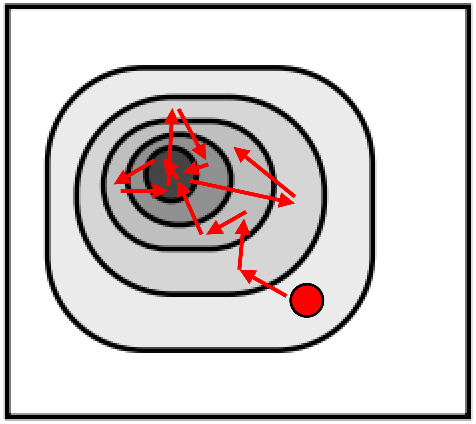
Nested Sampling: Solving an Easier Problem many times.

Sampling uniformly within bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.



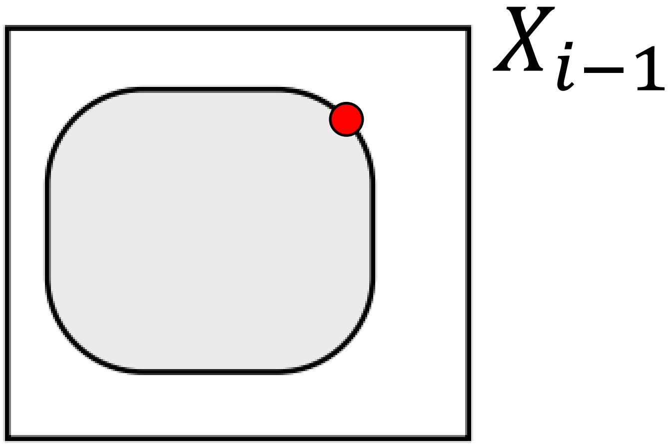
Pictures adapted from [this 2010 talk](#) by John Skilling.

Motivation: Sampling the Posterior



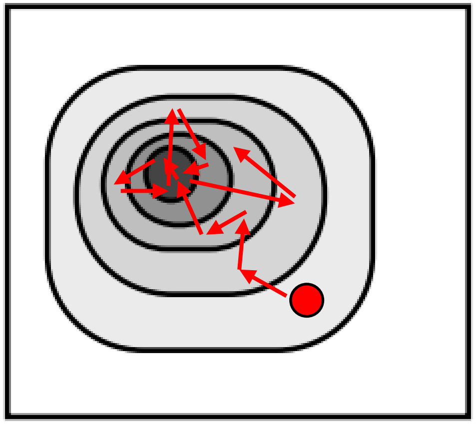
MCMC: Solving a Hard Problem once.
vs
Nested Sampling: Solving an Easier Problem many times.

Sampling uniformly within bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.



Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior



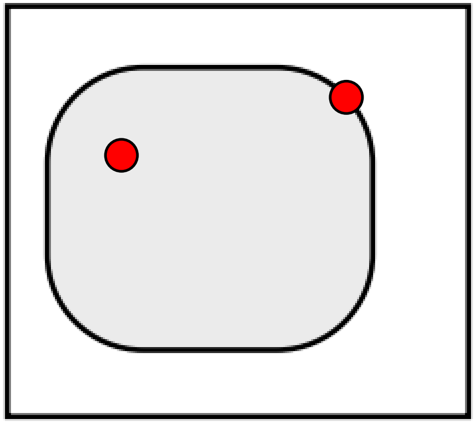
MCMC: Solving a Hard Problem once.

vs

Nested Sampling: Solving an Easier Problem many times.

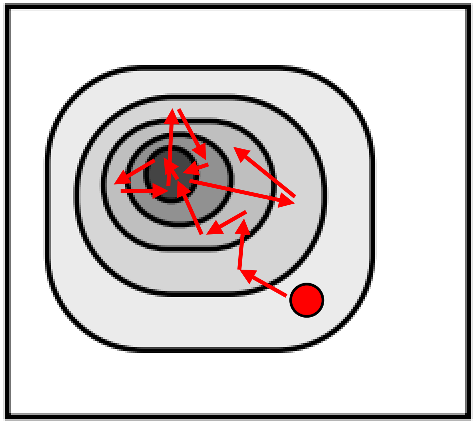
Sampling uniformly within bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.

X_{i-1}



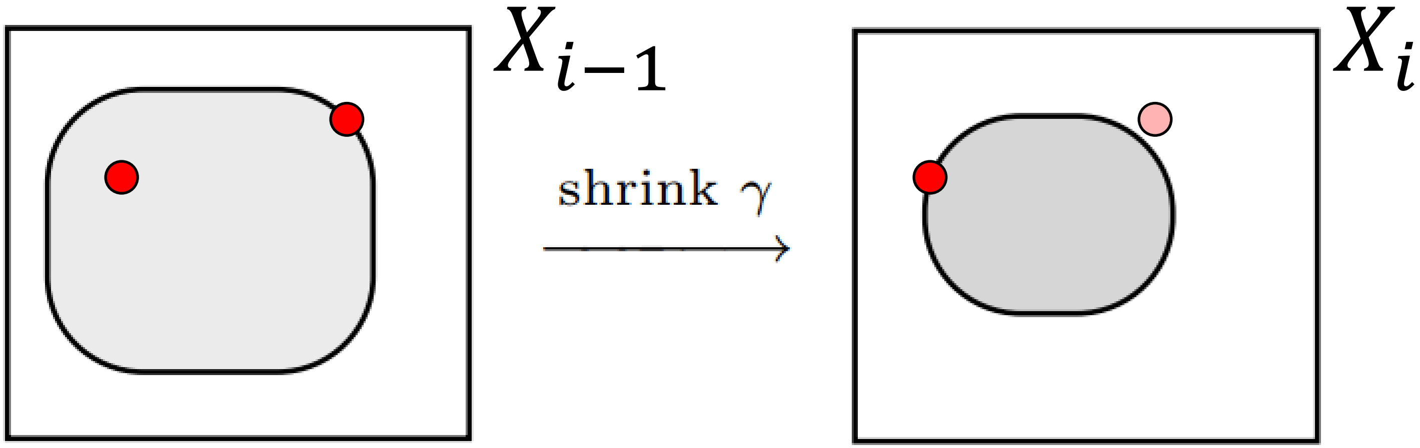
Pictures adapted from [this 2010 talk](#) by John Skilling.

Motivation: Sampling the Posterior



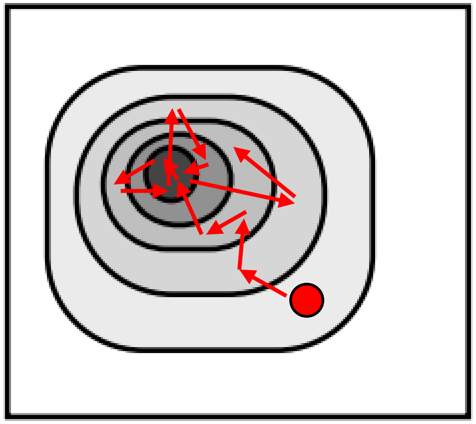
MCMC: Solving a Hard Problem once.
vs
Nested Sampling: Solving an Easier Problem many times.

Sampling uniformly within bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.



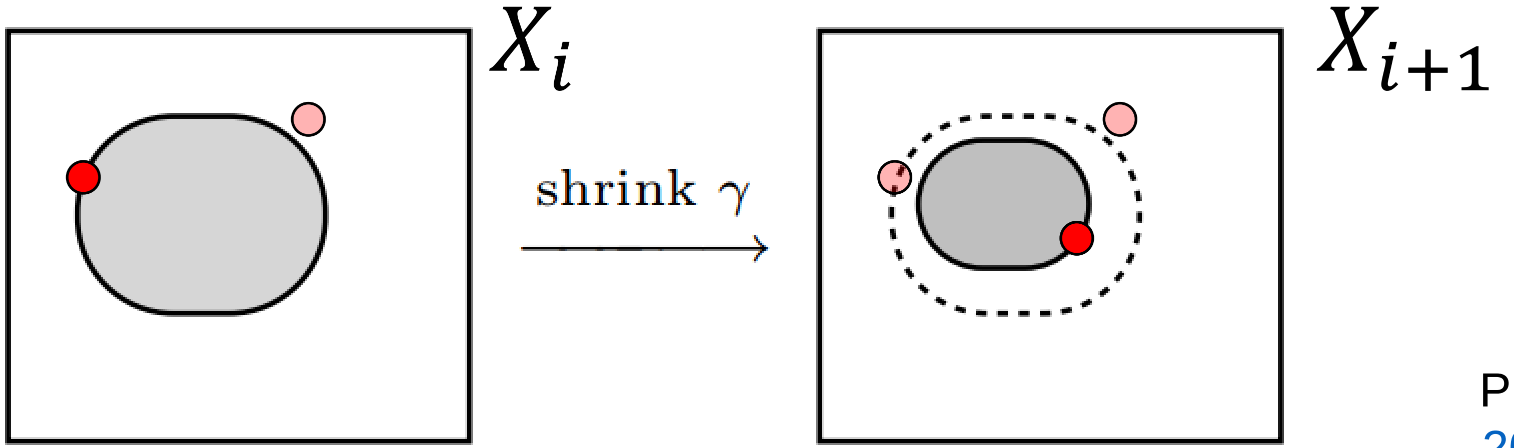
Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior



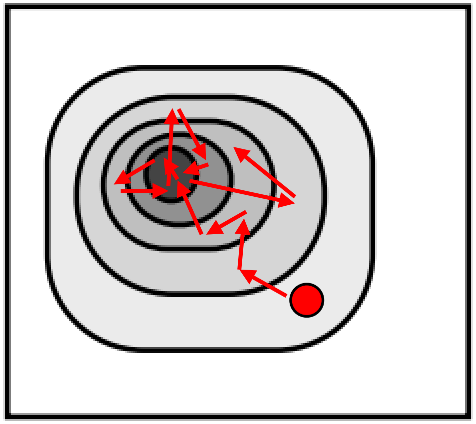
MCMC: Solving a Hard Problem once.
vs
Nested Sampling: Solving an Easier Problem many times.

Sampling uniformly within bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.



Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior

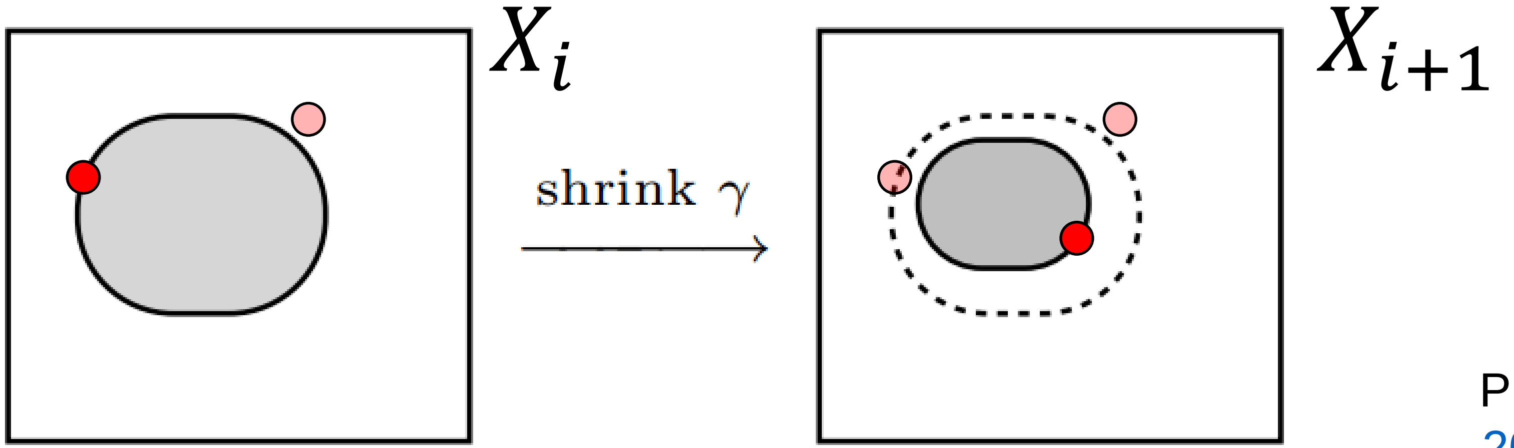


MCMC: Solving a Hard Problem once.

vs

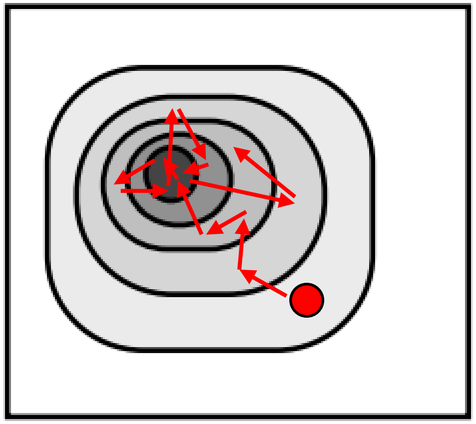
Nested Sampling: Solving an Easier Problem many times.

Sampling $\boxed{\text{uniformly within}}$ bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.



Pictures adapted from this 2010 talk by John Skilling.

Motivation: Sampling the Posterior



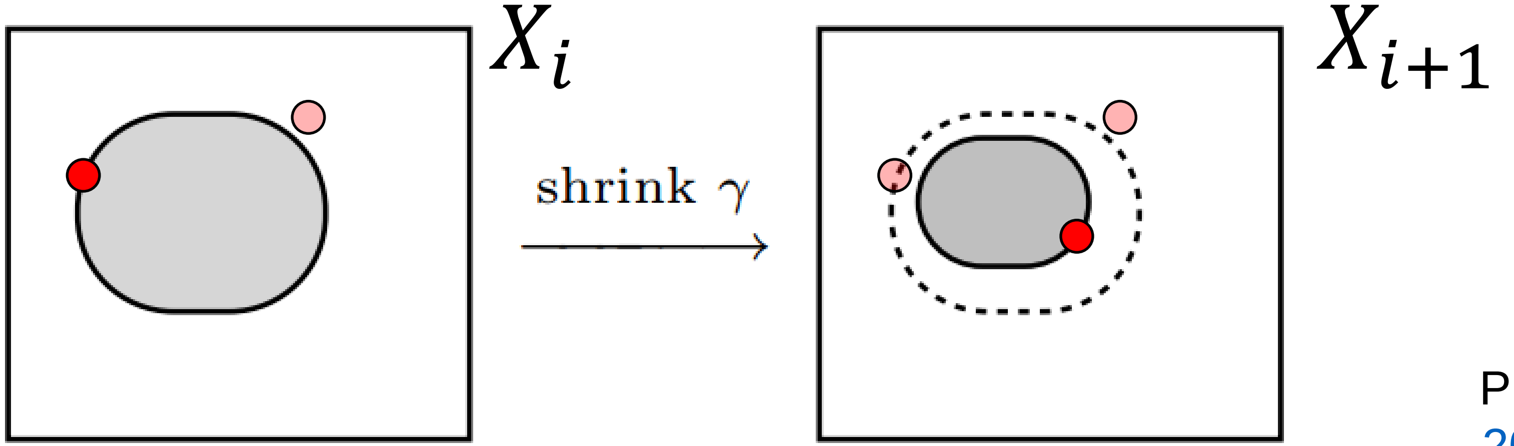
MCMC: Solving a Hard Problem once.

vs

Nested Sampling: Solving an Easier Problem many times.

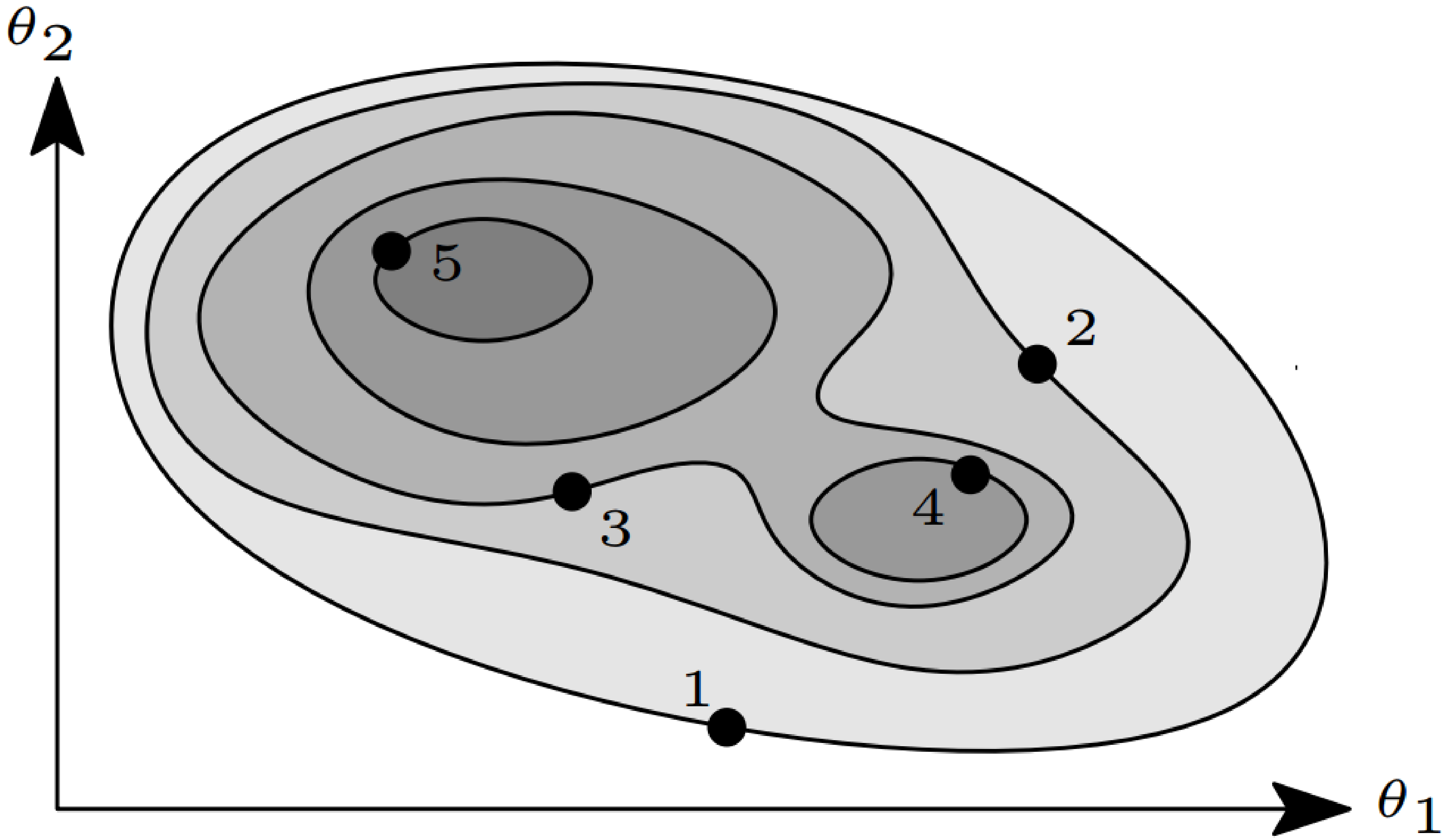
If you have a **prior transform** that converts your priors to look uniform, then this case is equivalent.

Sampling uniformly within bound $\mathcal{L}(\boldsymbol{\Theta}) > \lambda$ is **easier**.



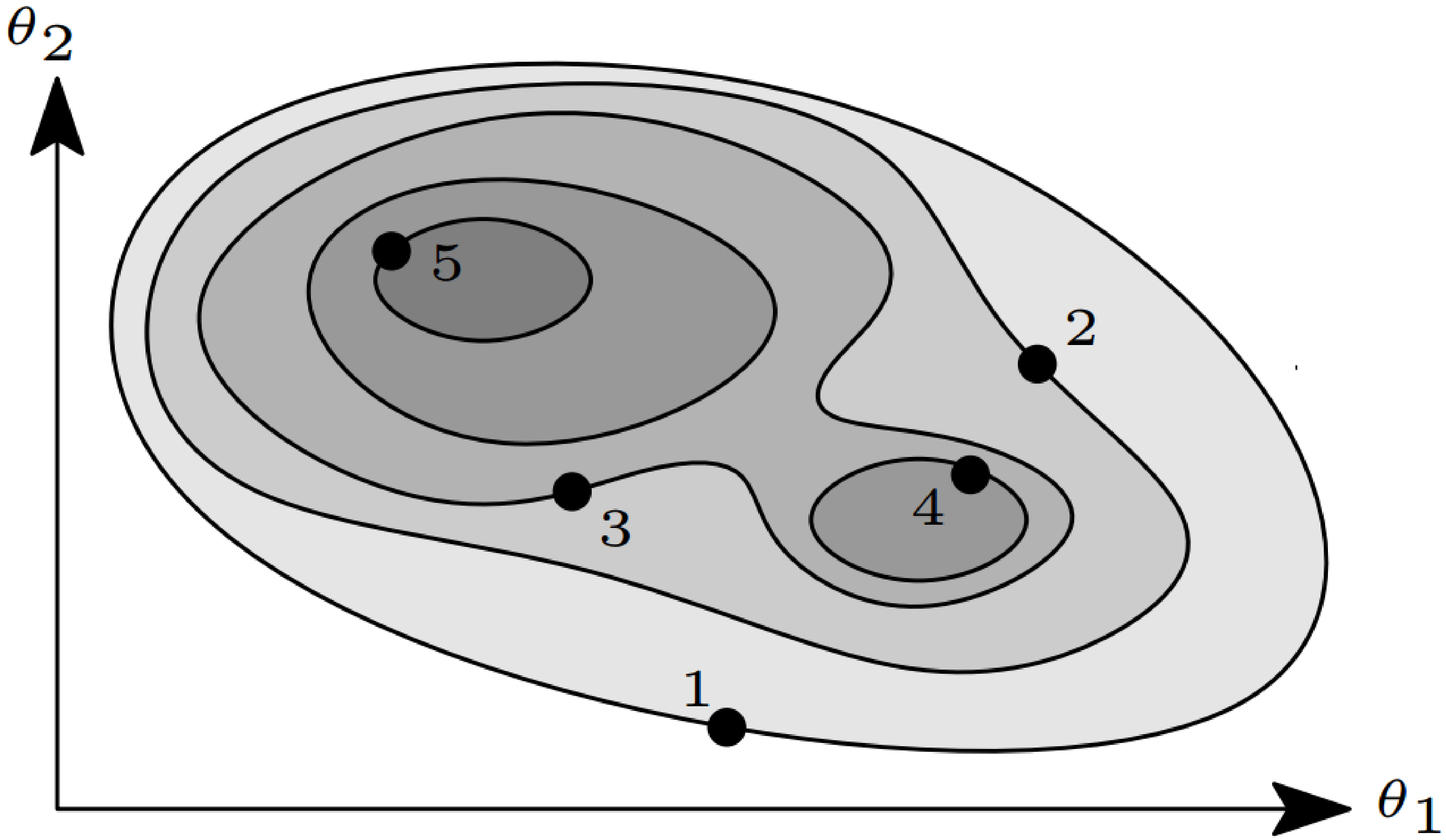
Pictures adapted from [this 2010 talk](#) by John Skilling.

Nested Sampling In Practice



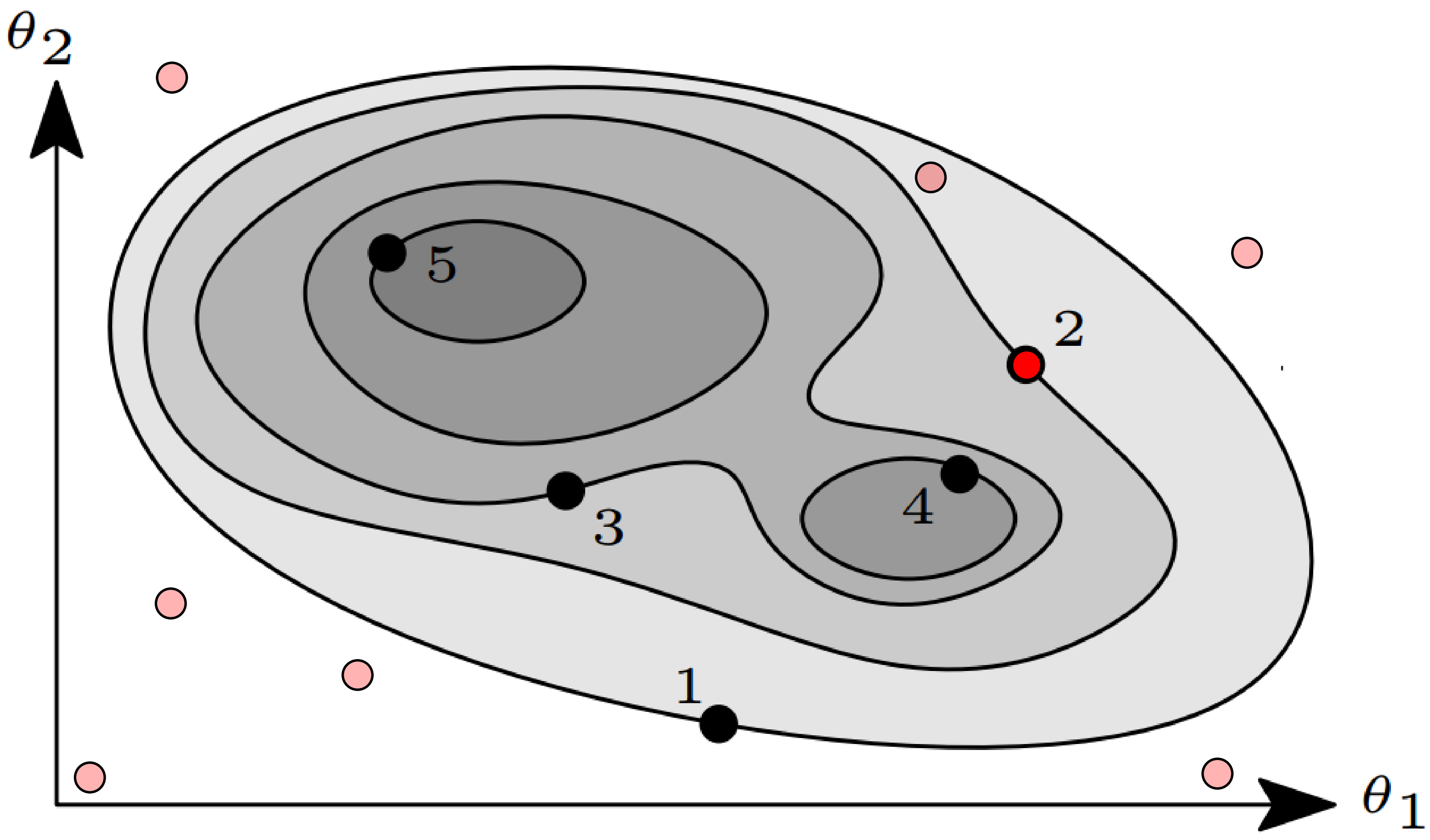
Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

Naïve Approach: Sampling from the Prior



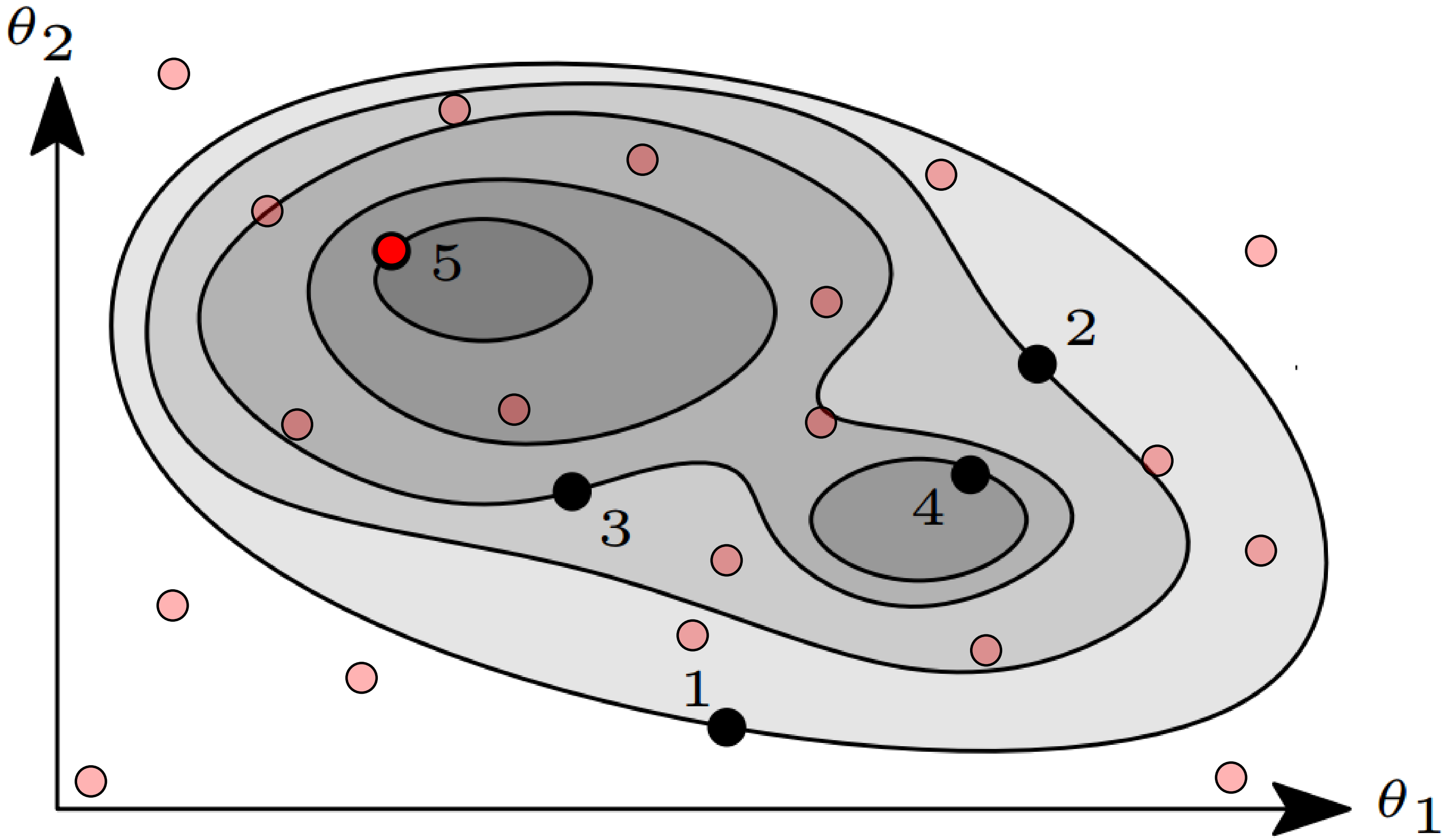
Higson et al. (2017)
arxiv:1704.03459

Naïve Approach: Sampling from the Prior



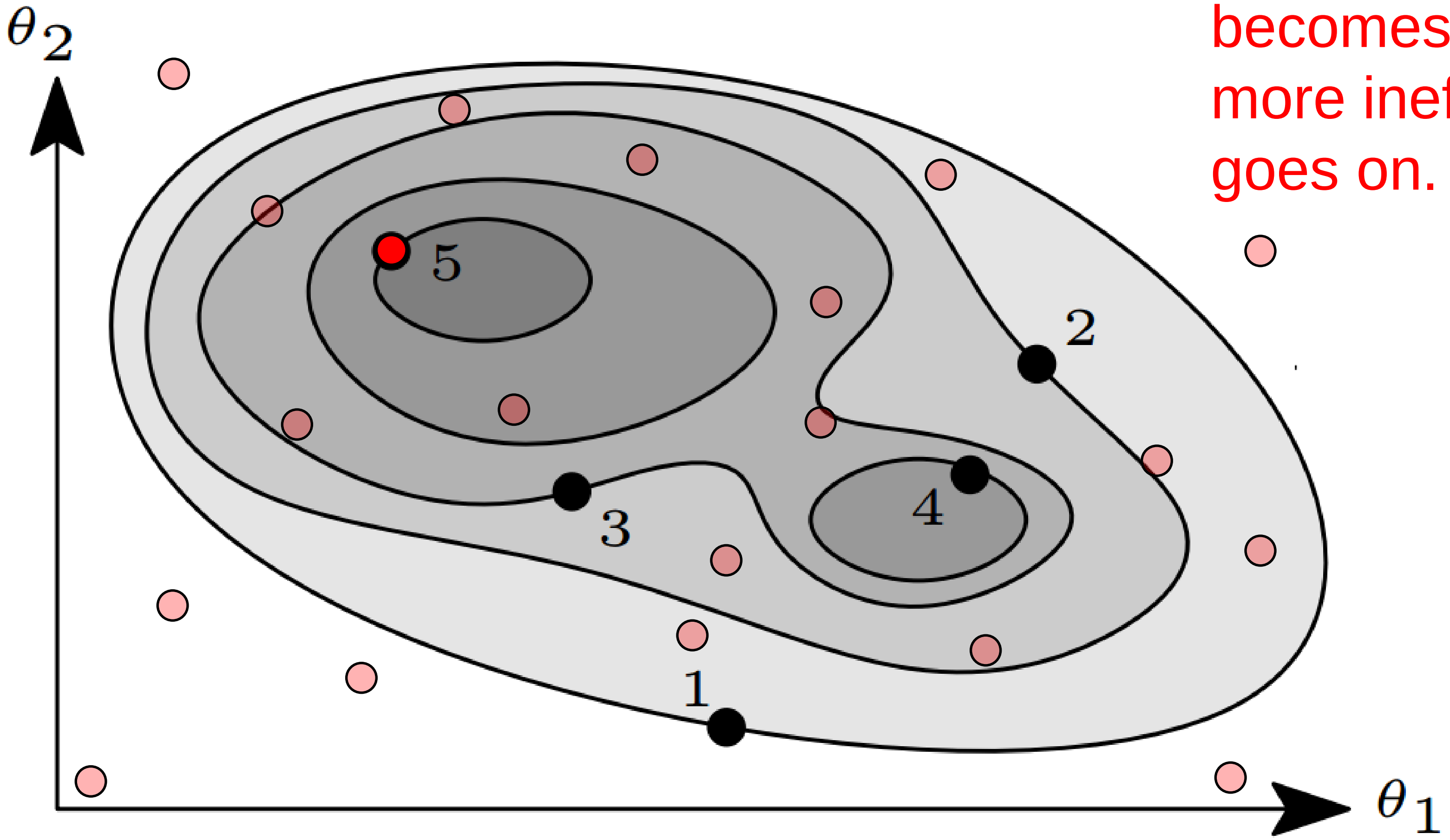
Higson et al. (2017)
arxiv:1704.03459

Naïve Approach: Sampling from the Prior



Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

Naïve Approach: Sampling from the Prior

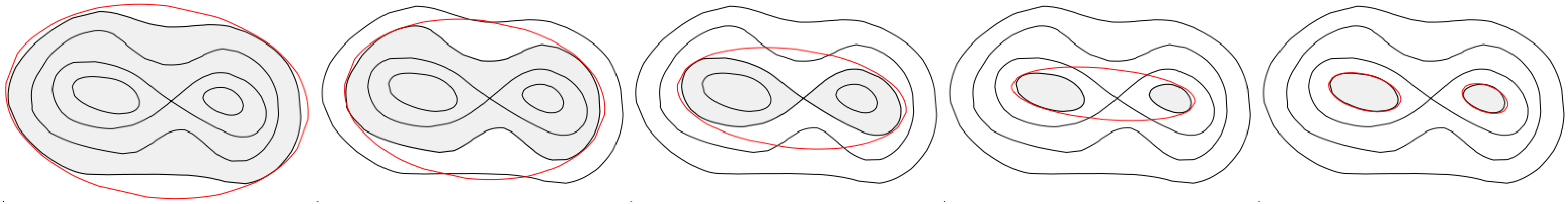


Sampling from the prior becomes **exponentially** more inefficient as time goes on.

Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

Sampling from the Constrained Prior

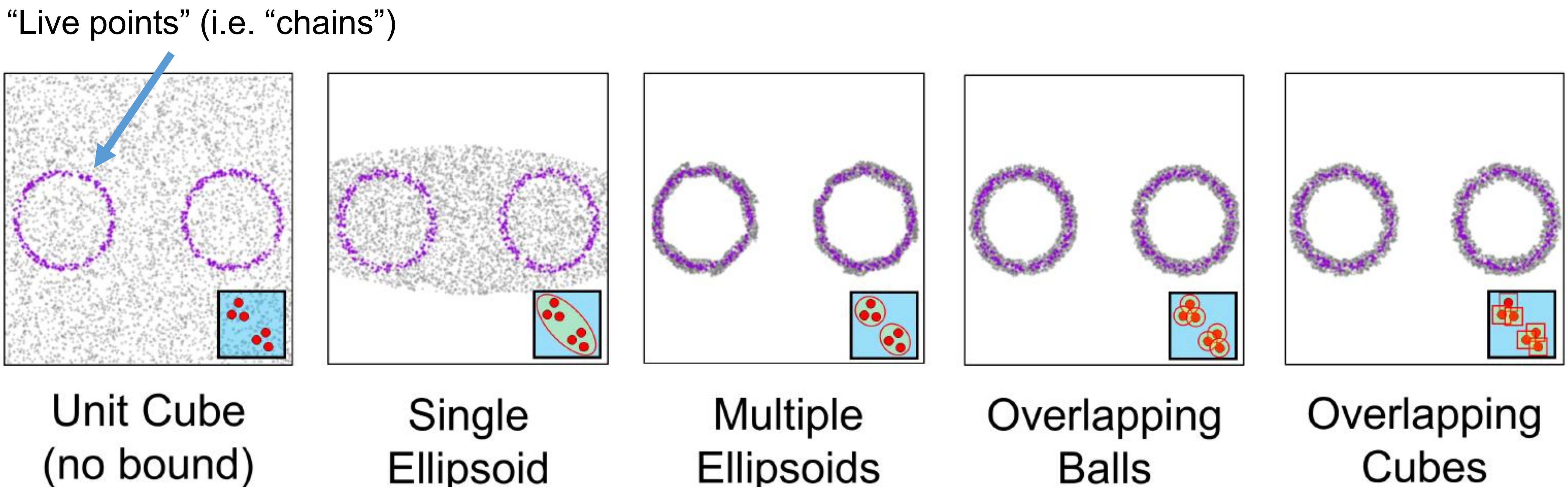
Proposal:
Try to bound the iso-likelihood contours in real time.



(a) (b) (c) (d) (e)

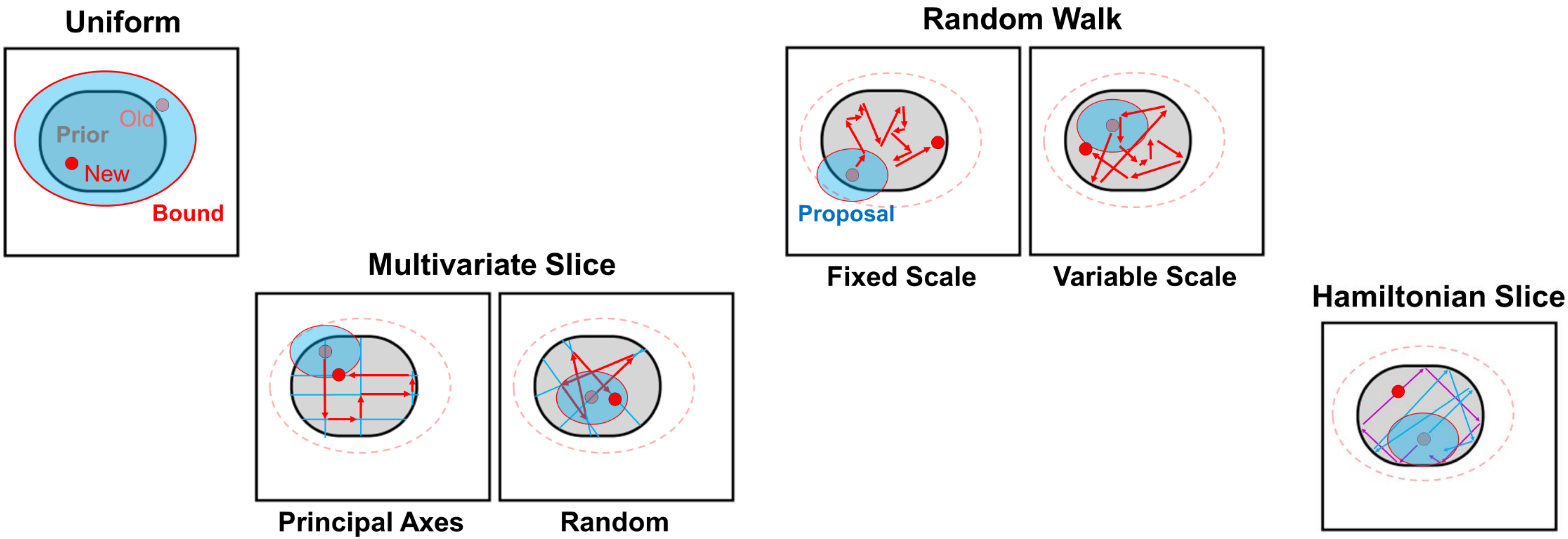
Feroz et al. (2009)

Examples of Bounding Strategies



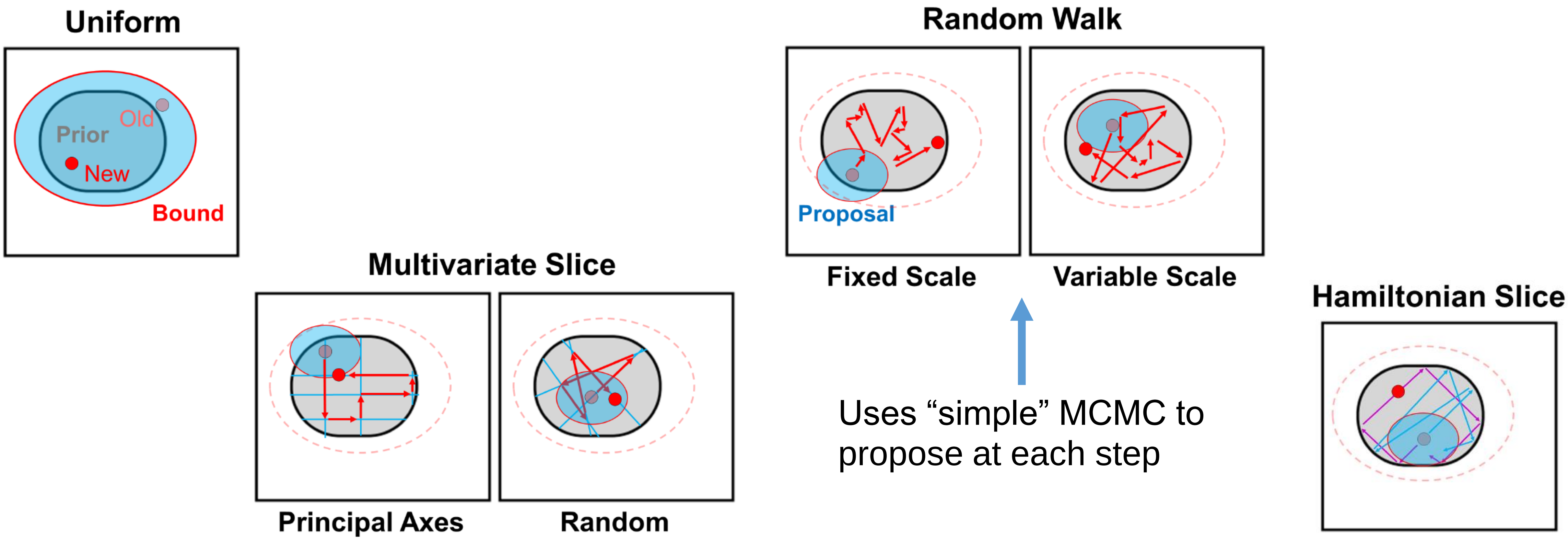
Bounding Method	**Relevant Papers**
Single ellipsoid (§4.1.2)	Mukherjee et al. (2006)
Multiple ellipsoids (§4.1.3)	Shaw et al. (2007), Feroz & Hobson (2008), Feroz et al. (2009, 2013), Handley et al. (2015)
Overlapping balls/cubes (§4.1.4, §4.1.5)	Buchner (2016, 2017)

Examples of Sampling Strategies



Sampling Method	
Uniform Sampling (§4.2.1)	Mukherjee et al. (2006), Feroz & Hobson (2008), Feroz et al. (2009), Buchner (2016)
Random walks (§4.2.2)	Skilling (2004), Skilling (2006)
Slice sampling (§4.2.3)	Neal (2003), Handley et al. (2015)
Hamiltonian slicing (§4.2.4)	Neal (2003), Skilling (2012), Feroz & Skilling (2013), this work

Examples of Sampling Strategies



Sampling Method	
Uniform Sampling (§4.2.1)	Mukherjee et al. (2006), Feroz & Hobson (2008), Feroz et al. (2009), Buchner (2016)
Random walks (§4.2.2)	Skilling (2004), Skilling (2006)
Slice sampling (§4.2.3)	Neal (2003), Handley et al. (2015)
Hamiltonian slicing (§4.2.4)	Neal (2003), Skilling (2012), Feroz & Skilling (2013), this work

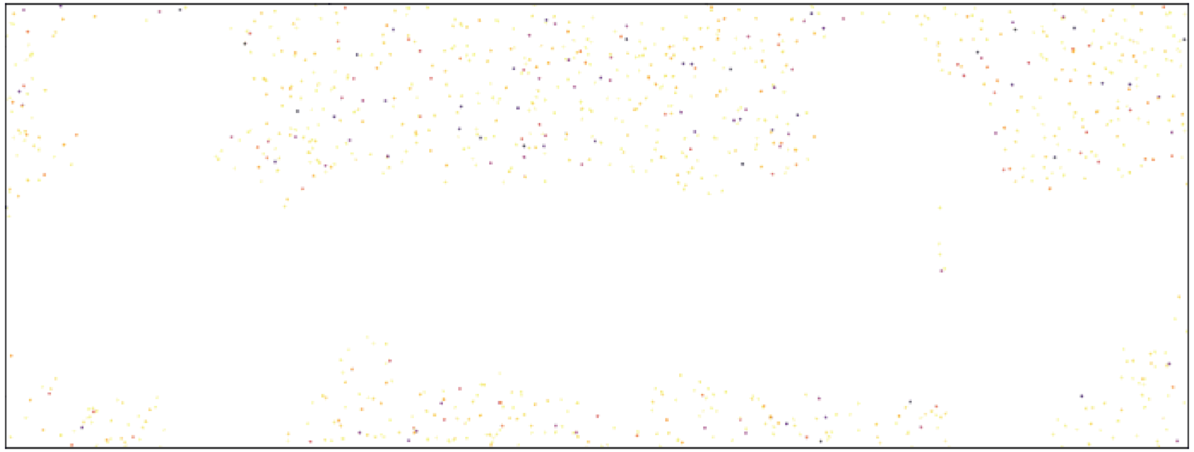
dynesty

Inspired by *emcee* (Foreman-Mackey et al. 2013)

- **Public, open source Python package** designed to make (Dynamic) Nested Sampling easy to use but also easy to customize.

- Designed to be **highly modular** and can mix-and-match methods.

- Includes **built-in plotting utilities** and post-processing tools.



Speagle (2020)

https://dynesty.readthedocs.io

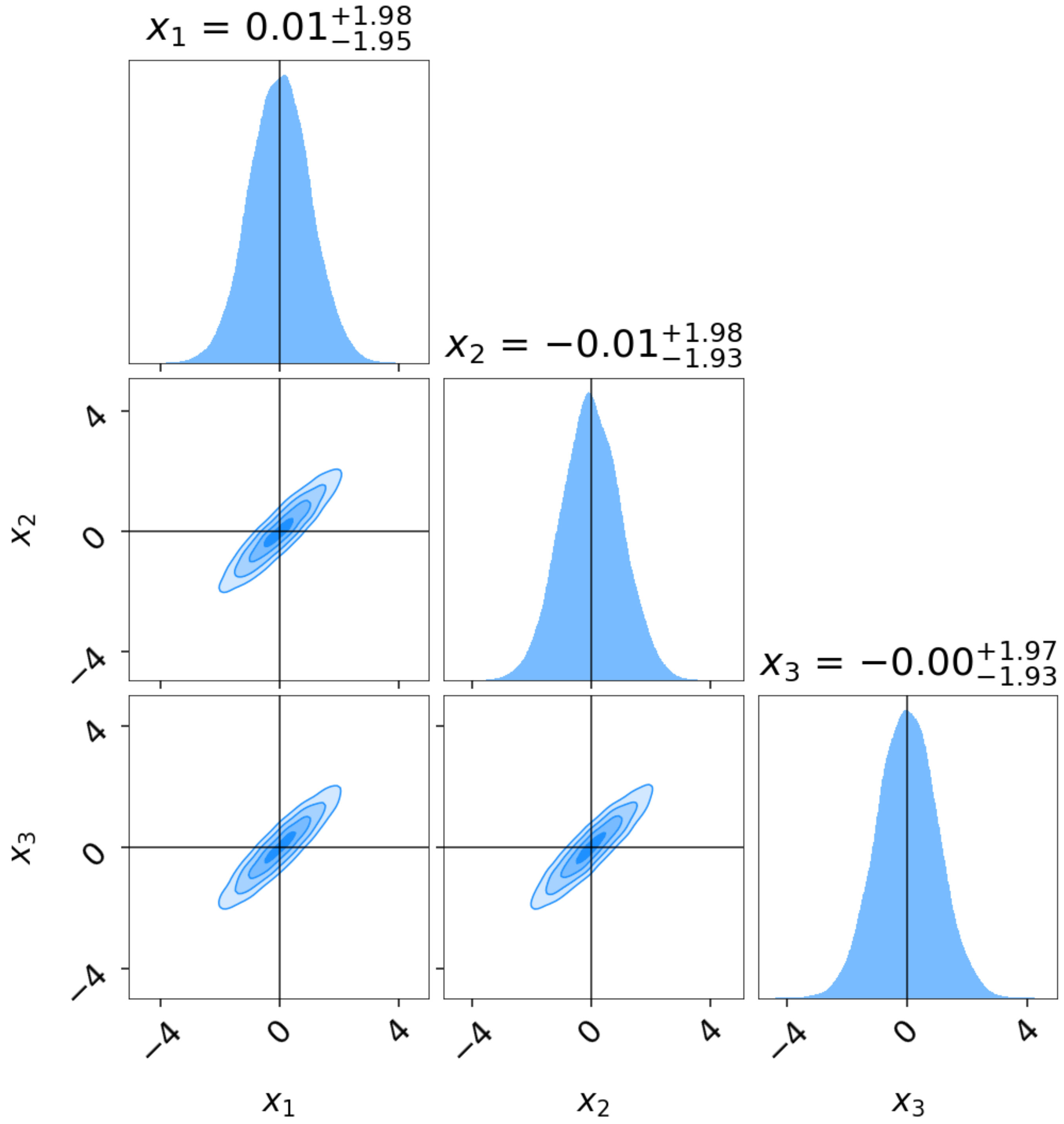
dynesty

```
import dynesty

# "Static" nested sampling.
sampler = dynesty.NestedSampler(loglike, ptform, ndim)
sampler.run_nested()
sresults = sampler.results

# "Dynamic" nested sampling.
dsampler = dynesty.DynamicNestedSampler(loglike, ptform, ndim)
dsampler.run_nested()
dresults = dsampler.results
```

dynesty

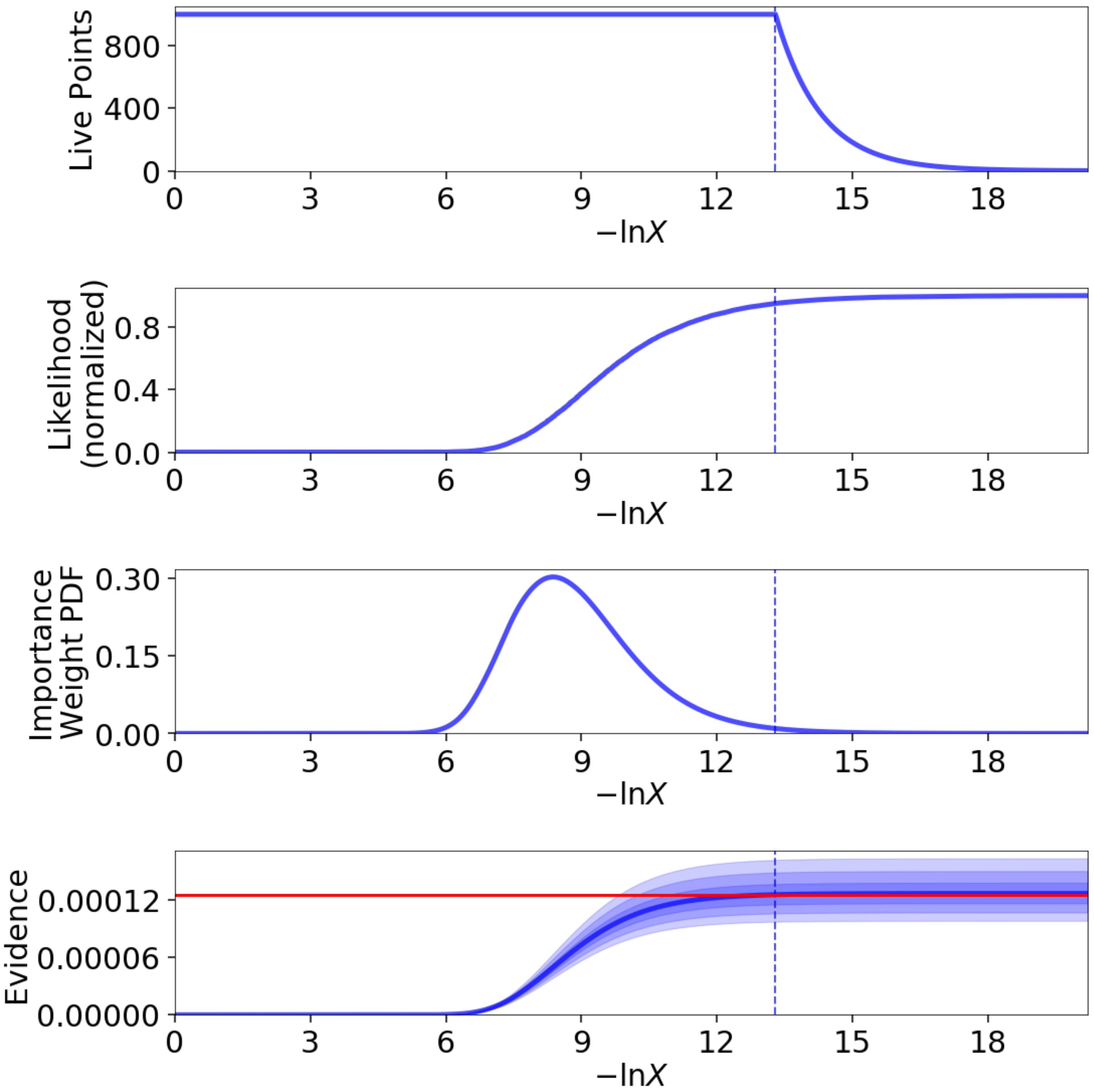
```python
from dynesty import plotting as dyplot

# Plot a summary of the run.
rfig, raxes = dyplot.runplot(results)

# Plot traces and 1-D marginalized posteriors.
tfig, taxes = dyplot.traceplot(results)

# Plot the 2-D marginalized posteriors.
cfig, caxes = dyplot.cornerplot(results)
```

dynesty



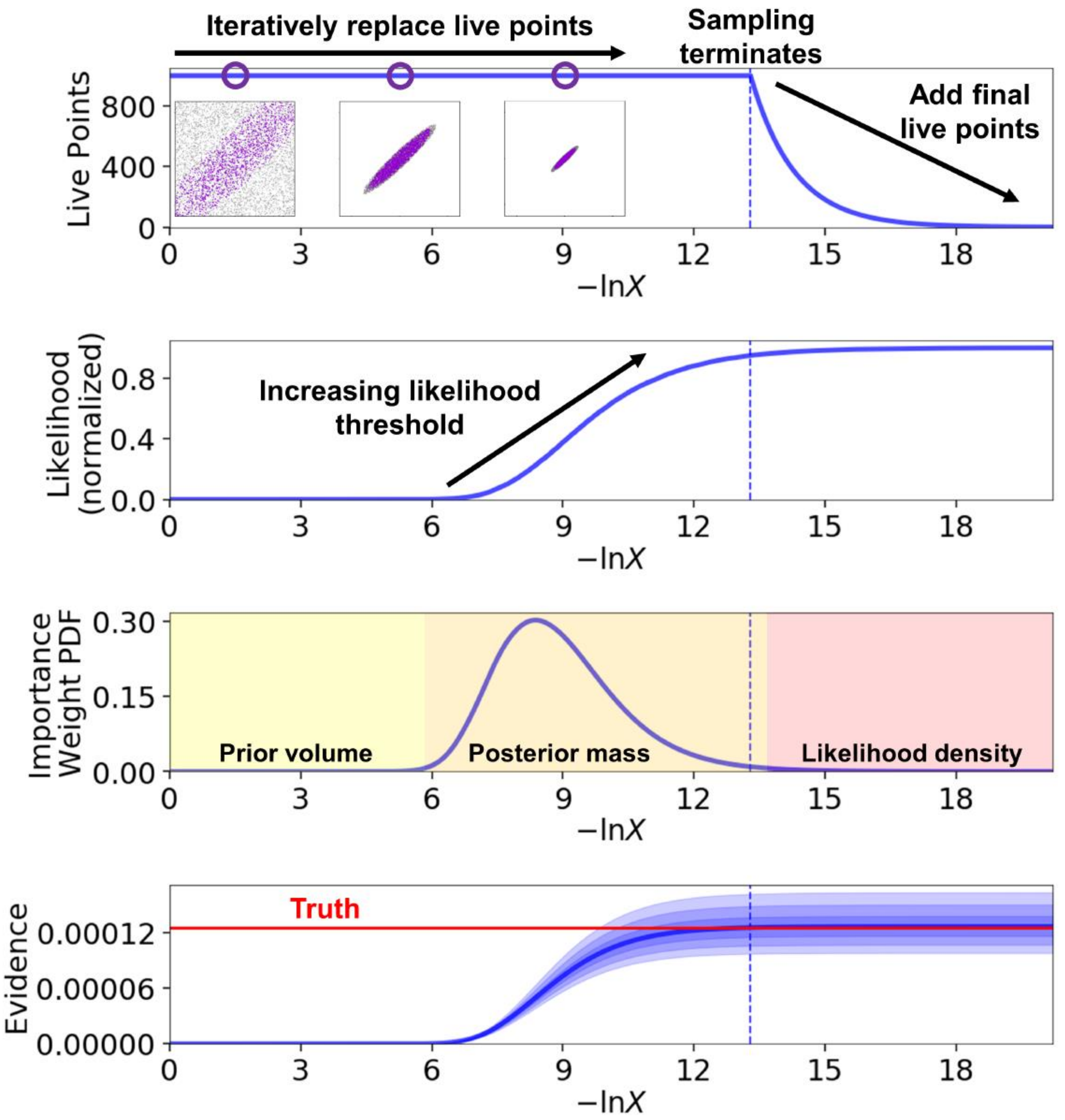
dynesty

```python
from dynesty import plotting as dyplot

# Plot a summary of the run.
rfig, raxes = dyplot.runplot(results)

# Plot traces and 1-D marginalized posteriors.
tfig, taxes = dyplot.traceplot(results)

# Plot the 2-D marginalized posteriors.
cfig, caxes = dyplot.cornerplot(results)
```



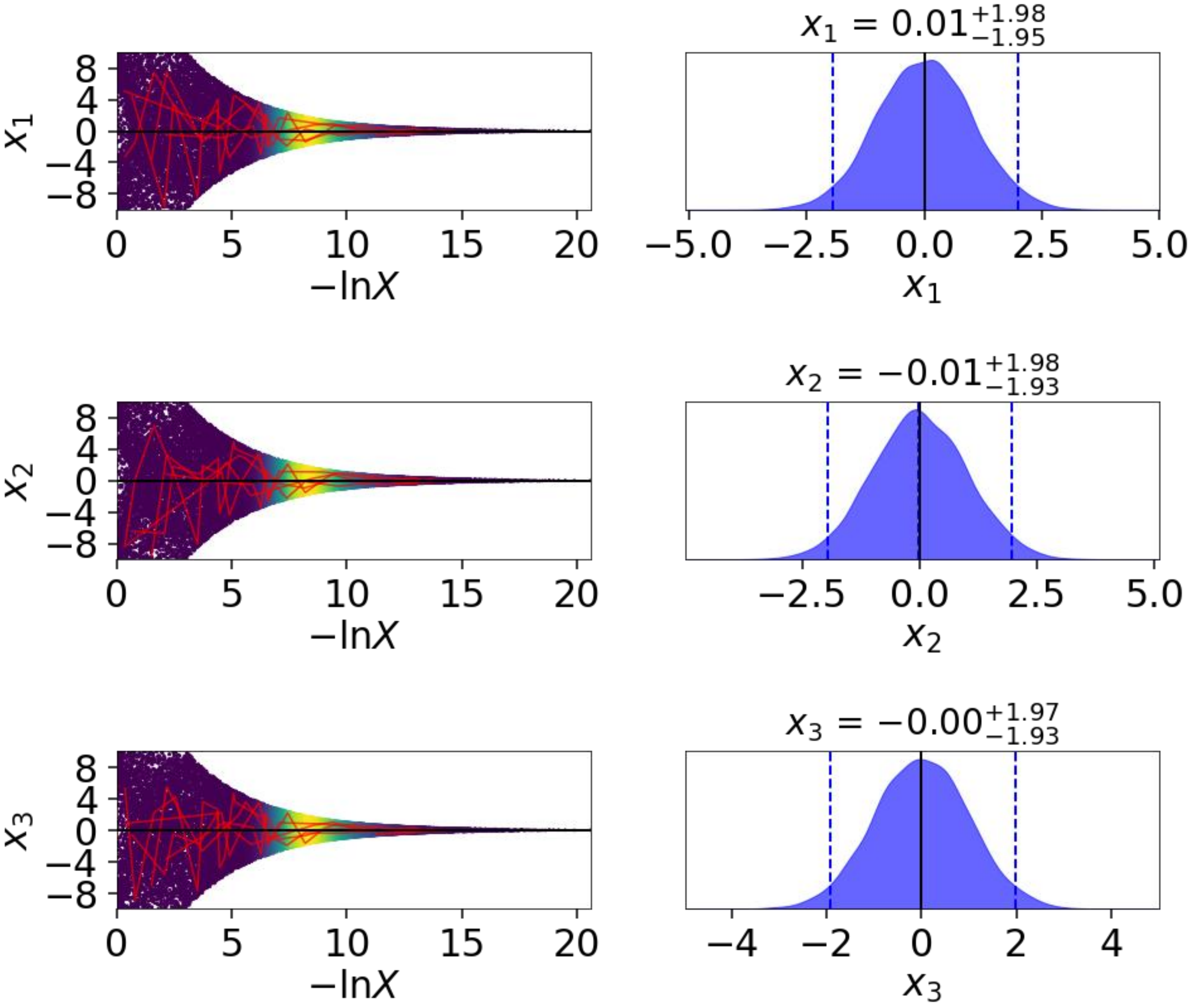
dynesty

```
from dynesty import plotting as dyplot

# Plot a summary of the run.
rfig, raxes = dyplot.runplot(results)

# Plot traces and 1-D marginalized posteriors.
tfig, taxes = dyplot.traceplot(results)

# Plot the 2-D marginalized posteriors.
cfig, caxes = dyplot.cornerplot(results)
```



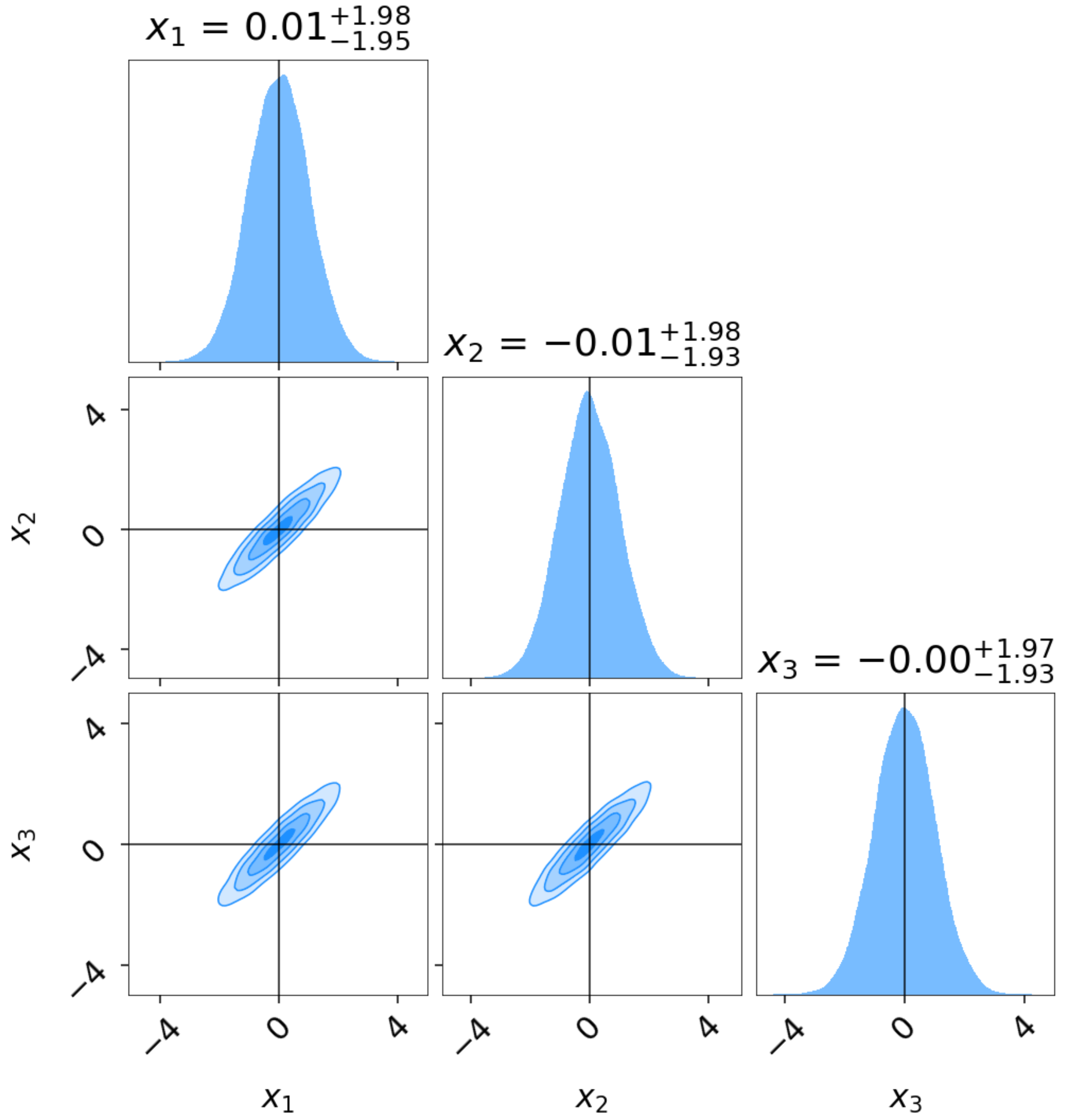
dynesty

```
from dynesty import plotting as dyplot

# Plot a summary of the run.
rfig, raxes = dyplot.runplot(results)

# Plot traces and 1-D marginalized posteriors.
tfig, taxes = dyplot.traceplot(results)

# Plot the 2-D marginalized posteriors.
cfig, caxes = dyplot.cornerplot(results)
```

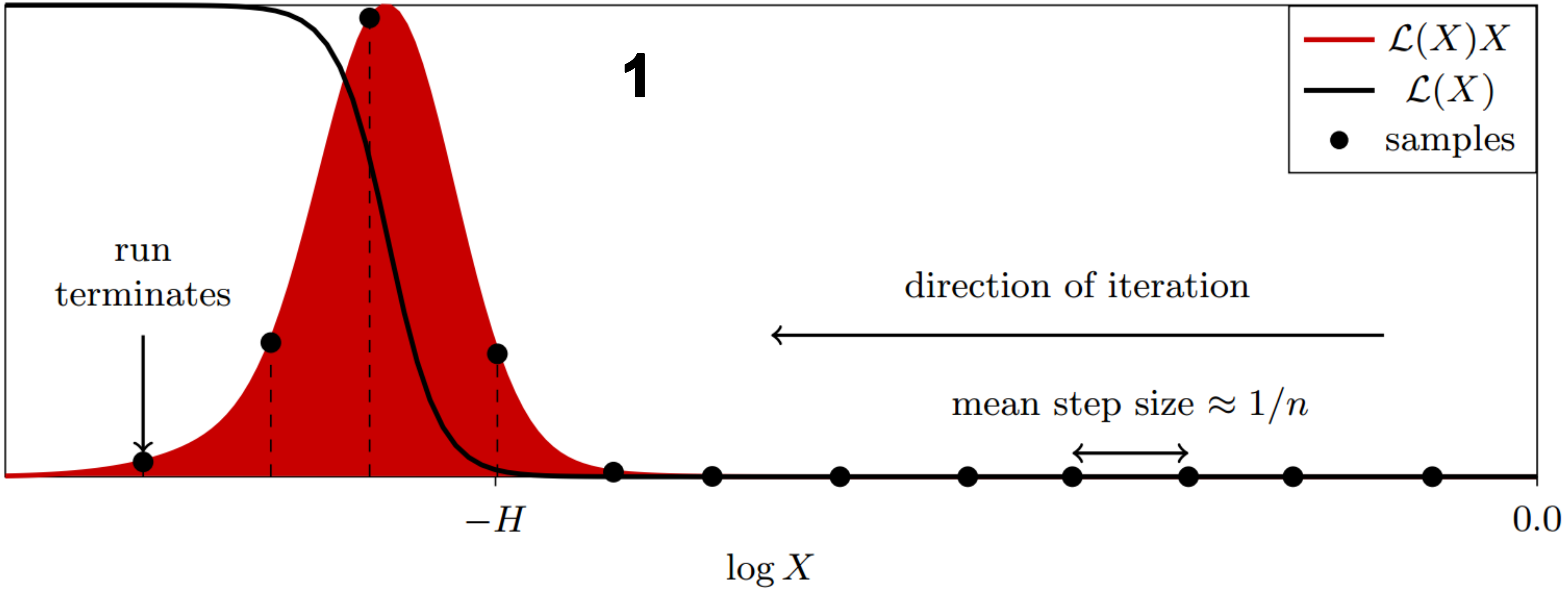


dynesty

```python
from dynesty import plotting as dyplot

# Plot a summary of the run.
rfig, raxes = dyplot.runplot(results)

# Plot traces and 1-D marginalized posteriors.
tfig, taxes = dyplot.traceplot(results)

# Plot the 2-D marginalized posteriors.
cfig, caxes = dyplot.cornerplot(results)
```

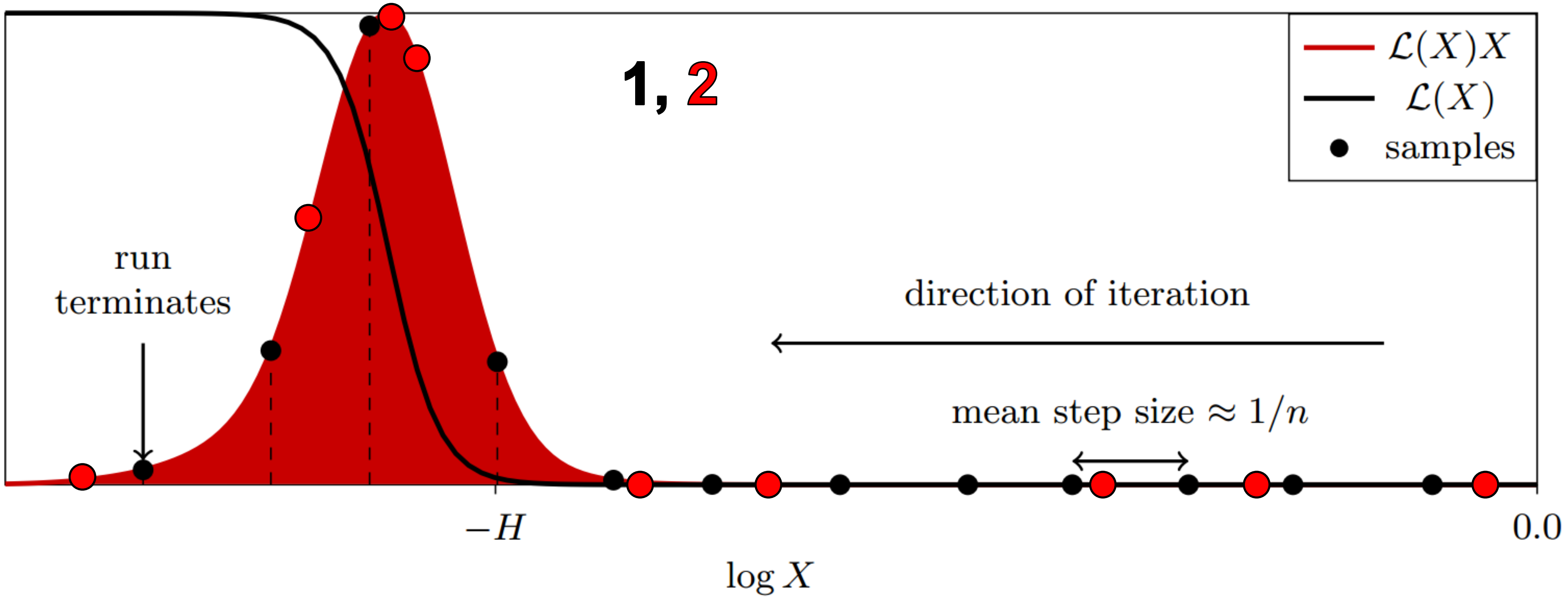


Dynamic Nested Sampling



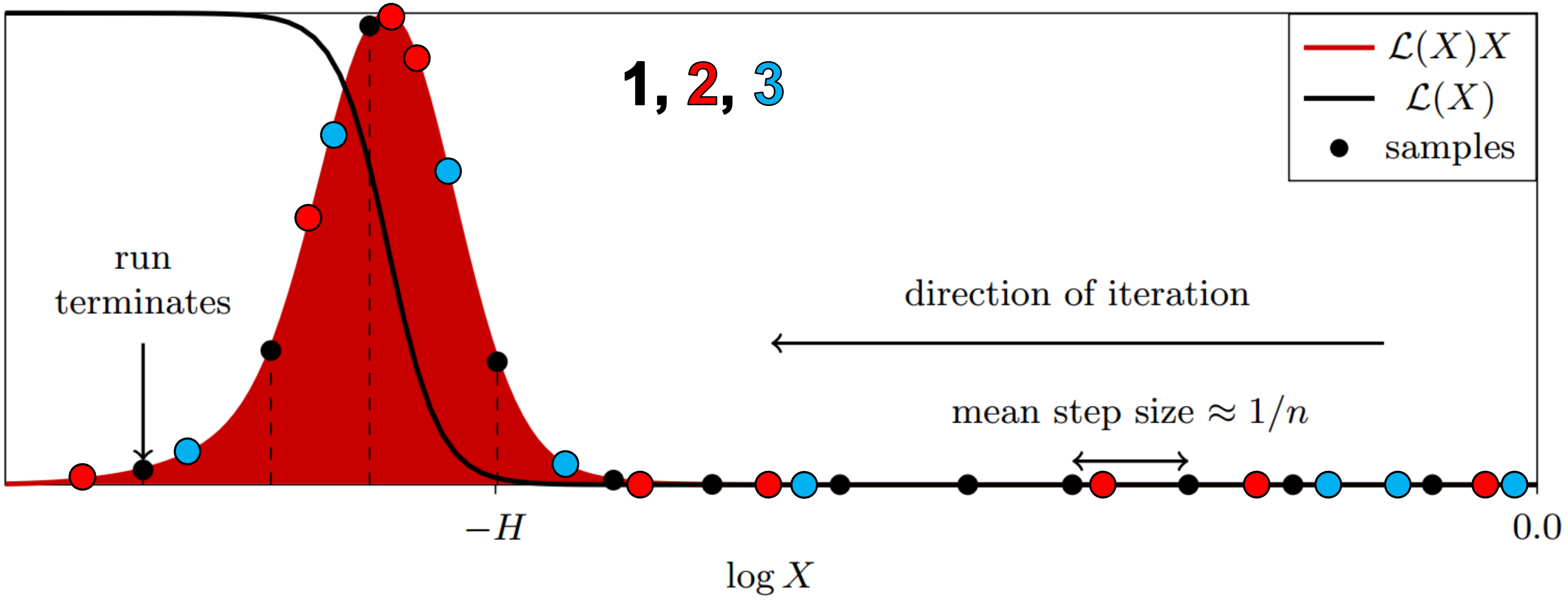
Higson et al. (2017)
arxiv:1704.03459

Dynamic Nested Sampling



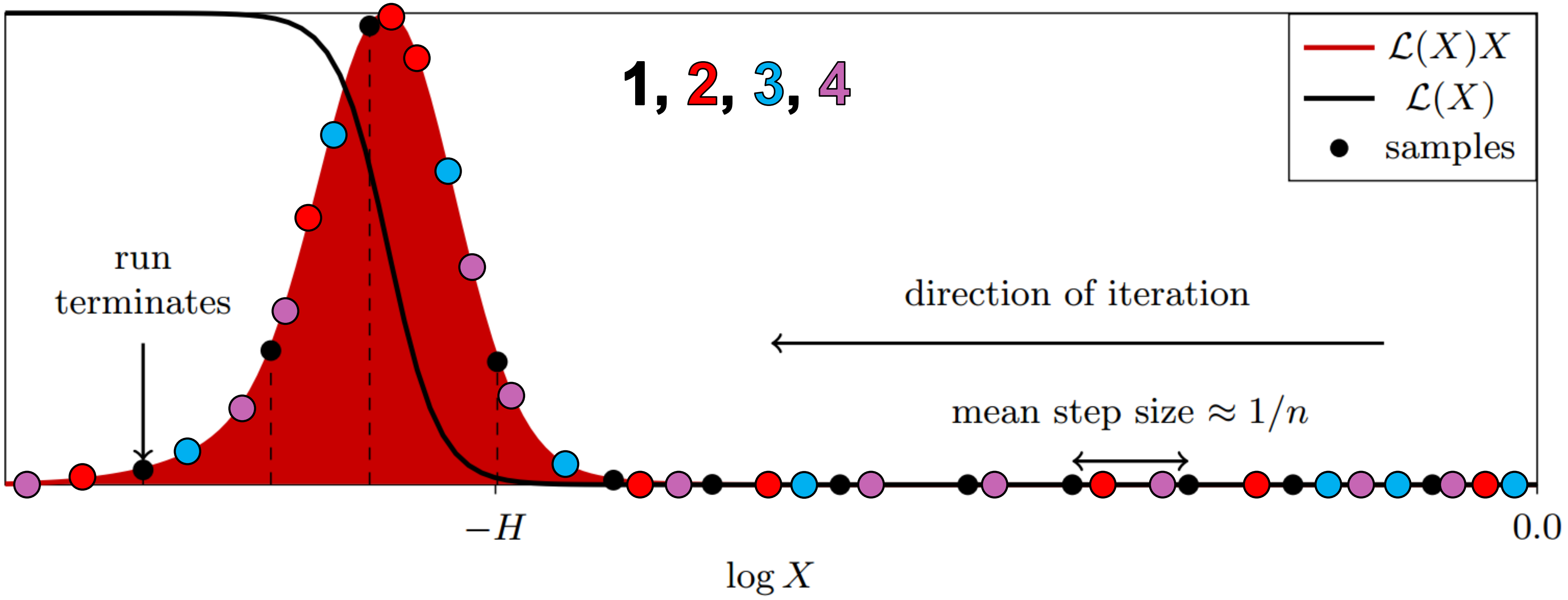
Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

Dynamic Nested Sampling



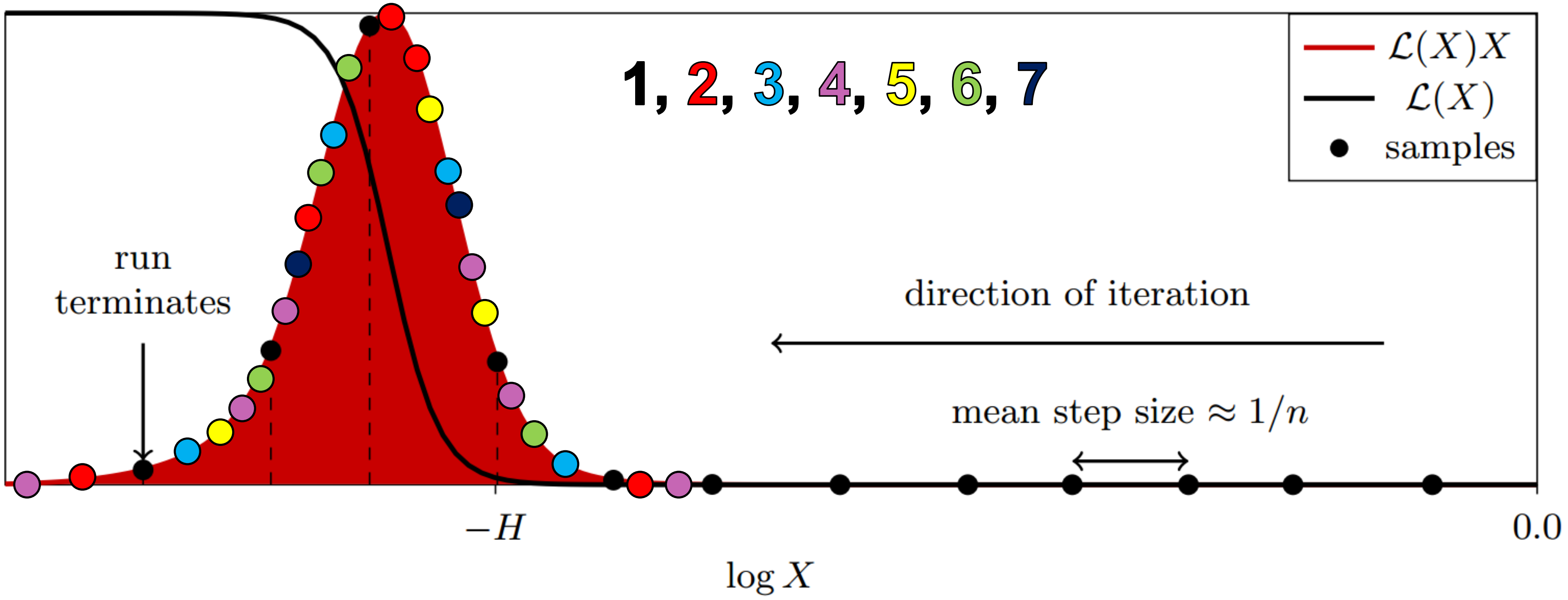
Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

Dynamic Nested Sampling



Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

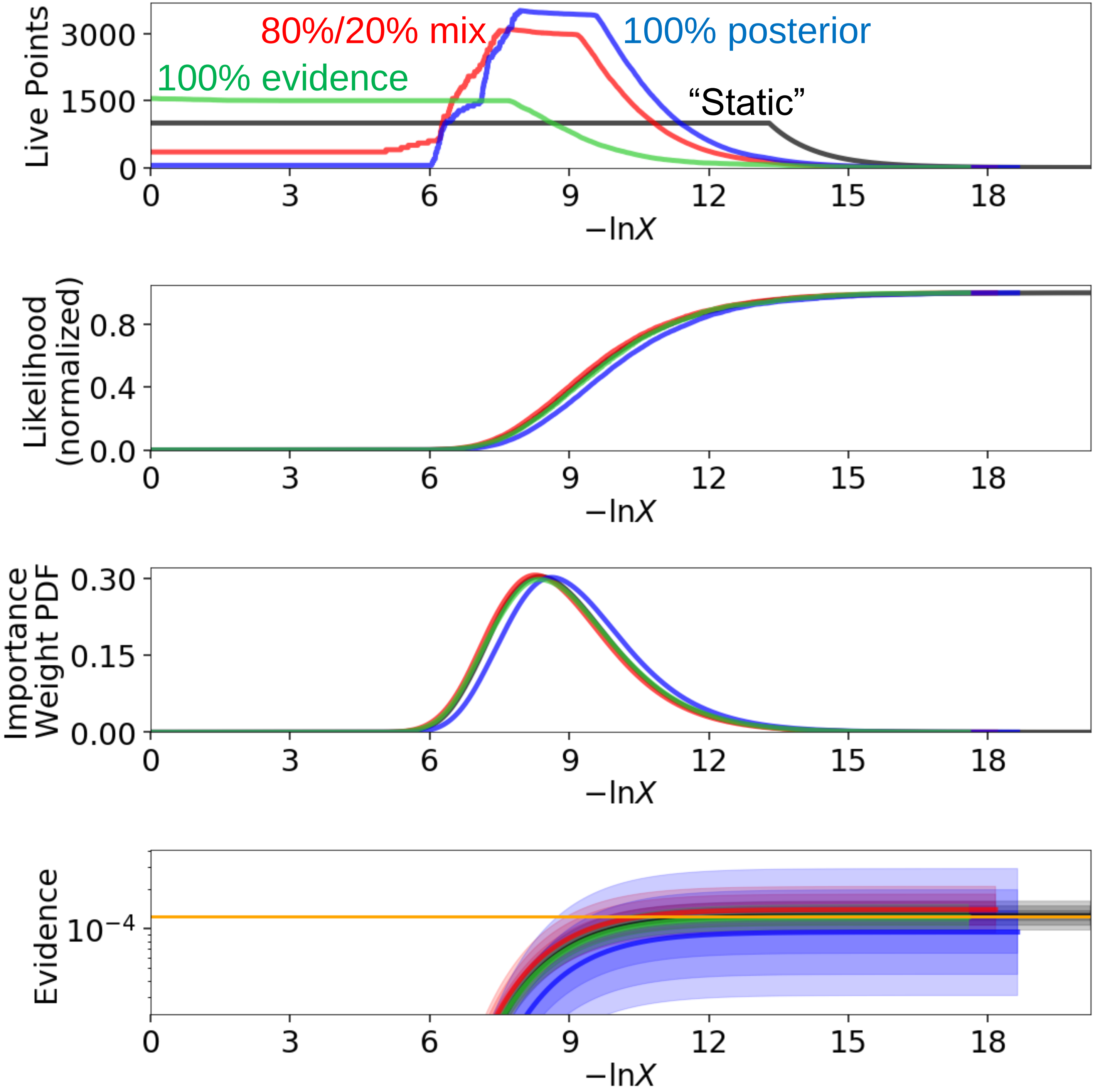
Dynamic Nested Sampling



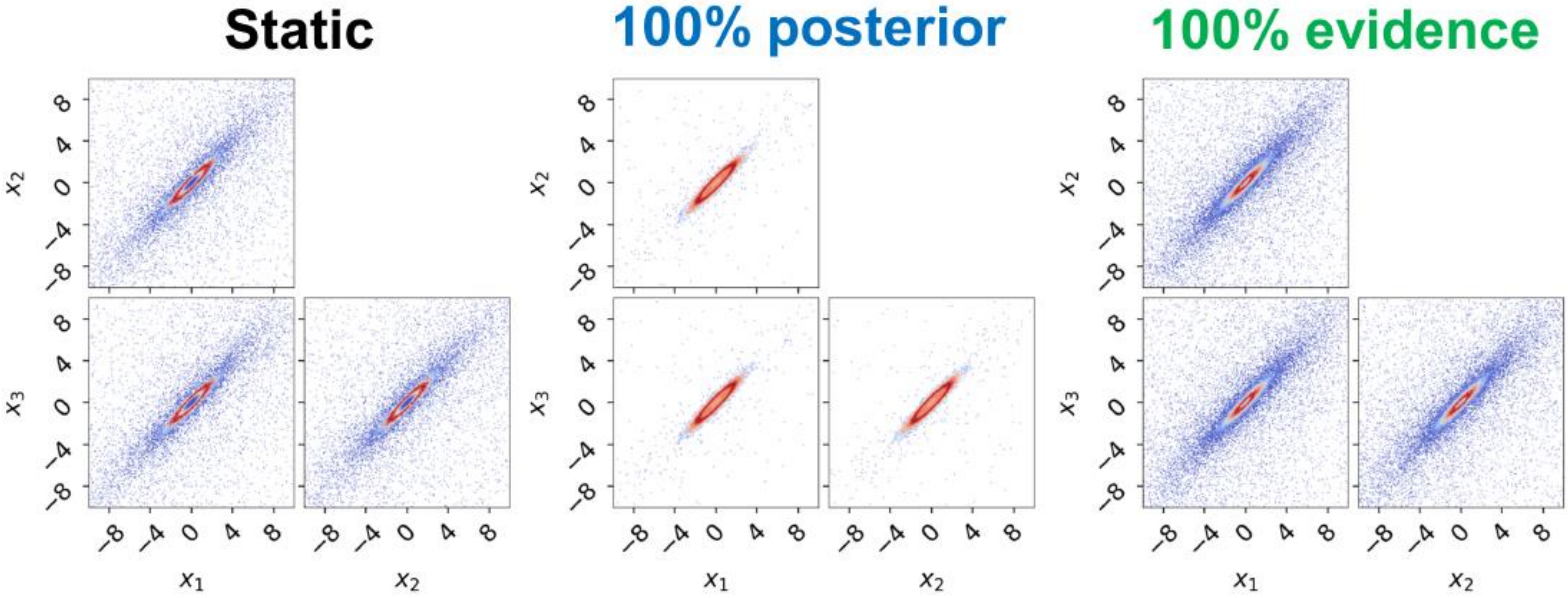
Higson et al. (2017)
[arxiv:1704.03459](https://arxiv.org/abs/1704.03459)

Comparisons

- Same 2-D Gaussian example as before with **fixed number of samples**.
- Only change is in overall Dynamic Nested Sampling **strategy**.



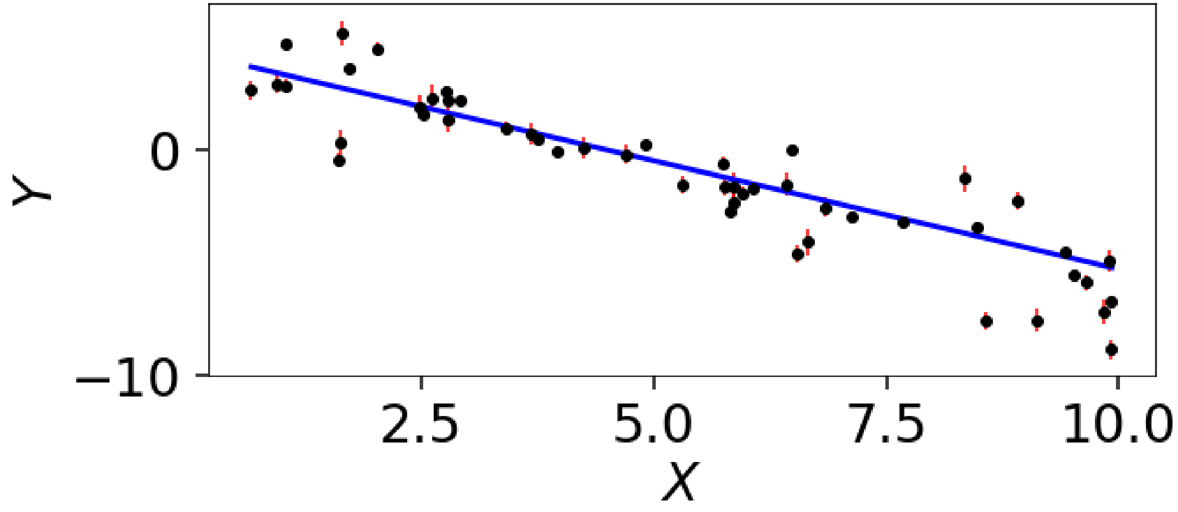
Comparisons

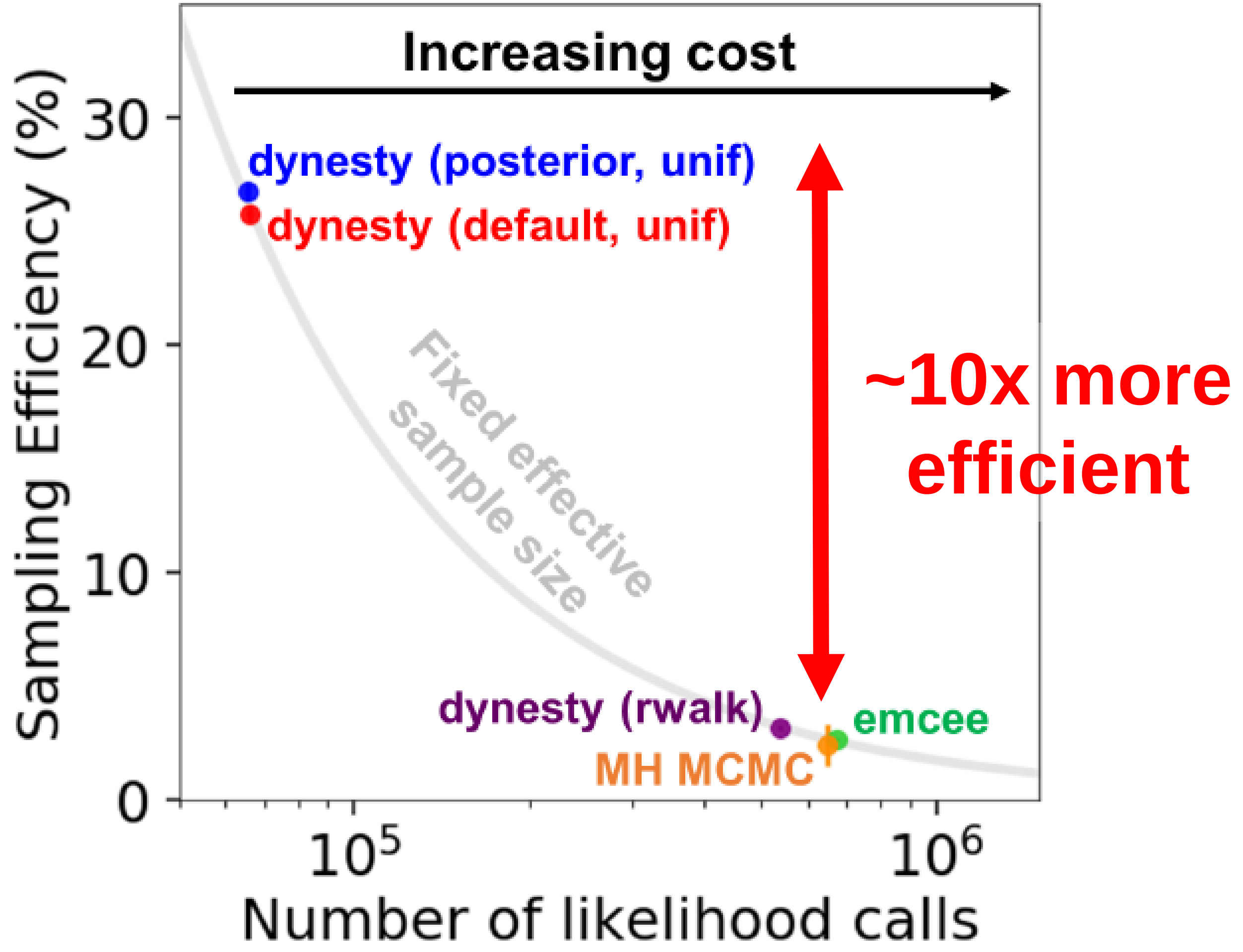


Speagle (2020)

Comparisons with emcee

- Simple 3-parameter linear regression problem with reasonable prior constraints.

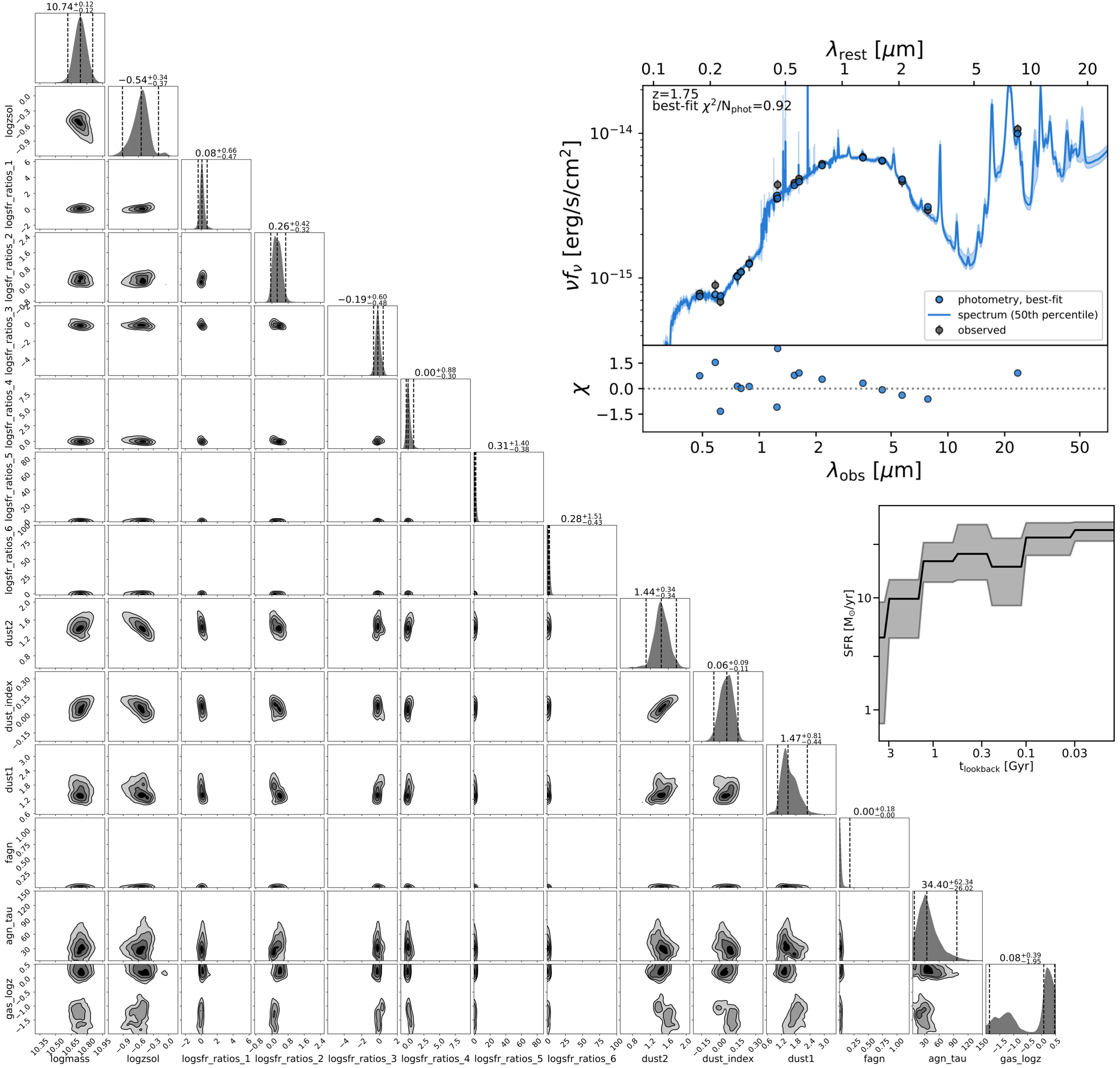




Speagle (2020)

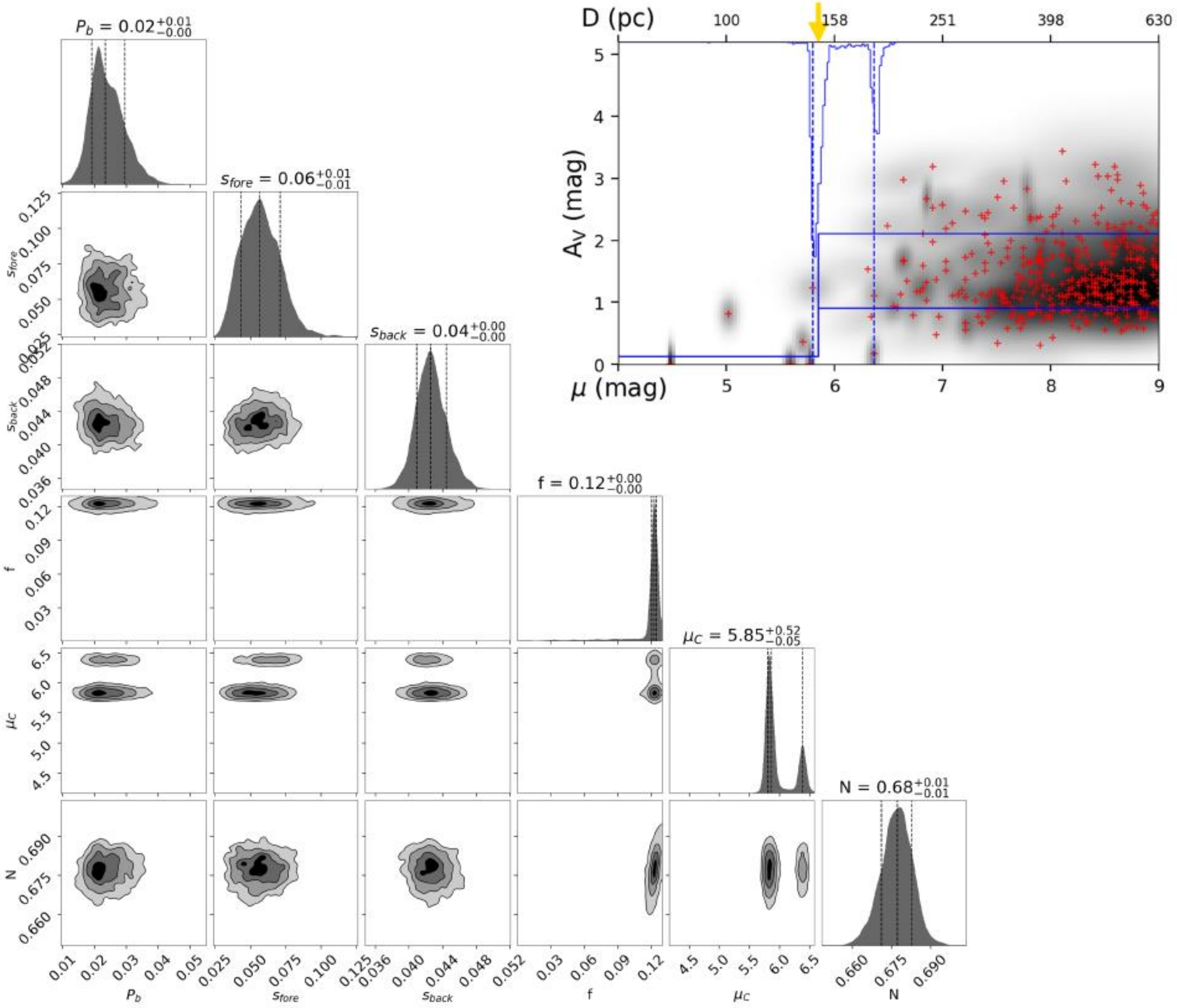
Applications

- Galaxy SED modeling with *Prospector* (Leja et al. 2019, Johnson et al. in prep.).



Applications

- 3-D dust mapping (Zucker & Speagle et al. 2019)



Advantages and Disadvantages

Advantages to Nested Sampling:

1. Can characterize complex uncertainties in real-time.
2. Can allocate samples much more efficiently in some cases.
3. Possesses well-motivated stopping criteria (Skilling 2006; Speagle 2020).

Advantages and Disadvantages

Advantages to Nested Sampling:

1. Can characterize complex uncertainties in real-time.
2. Can allocate samples much more efficiently in some cases.
3. Possesses well-motivated stopping criteria (Skilling 2006; Speagle 2020).

Disadvantages to Nested Sampling:

1. Implementations require a prior transform.
2. Runtime sensitive to size of prior.
3. Overall approach can sometimes miss certain types of solutions.
4. Sampling is more involved.

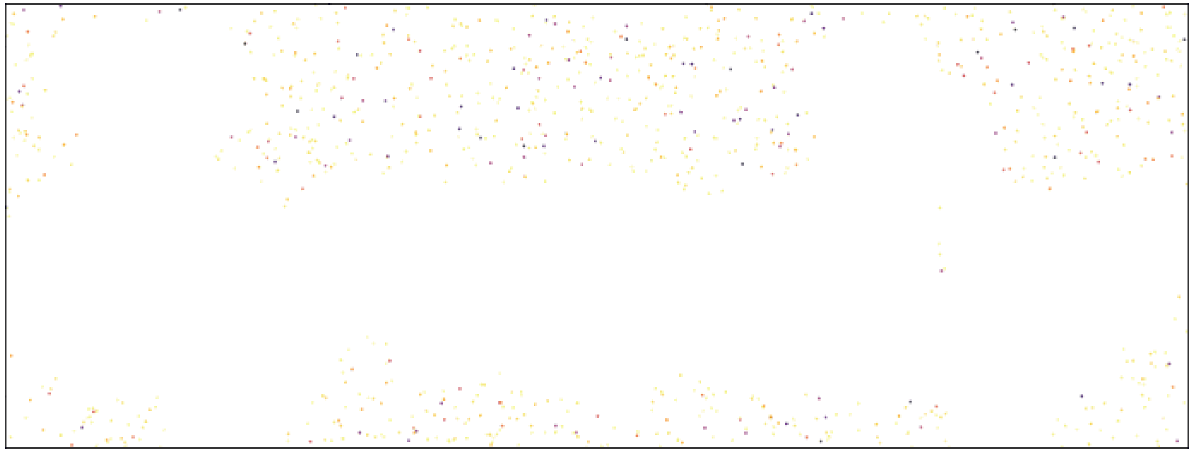
dynesty

Inspired by *emcee* (Foreman-Mackey et al. 2013)

- **Public, open source Python package** designed to make (Dynamic) Nested Sampling easy to use but also easy to customize.

- Designed to be **highly modular** and can mix-and-match methods.

- Includes **built-in plotting utilities** and post-processing tools.



Speagle (2020)

https://dynesty.readthedocs.io