HierarchMC

A Bayesian Framework for Hierarchical All-Sky Searches for Continuous Gravitational Waves

Jasper Martins

Maria Alessandra Papa, Benjamin Steltner

HierarchMC

A Bayesian Framework for Hierarchical All-Sky Se … Gravitational Waves

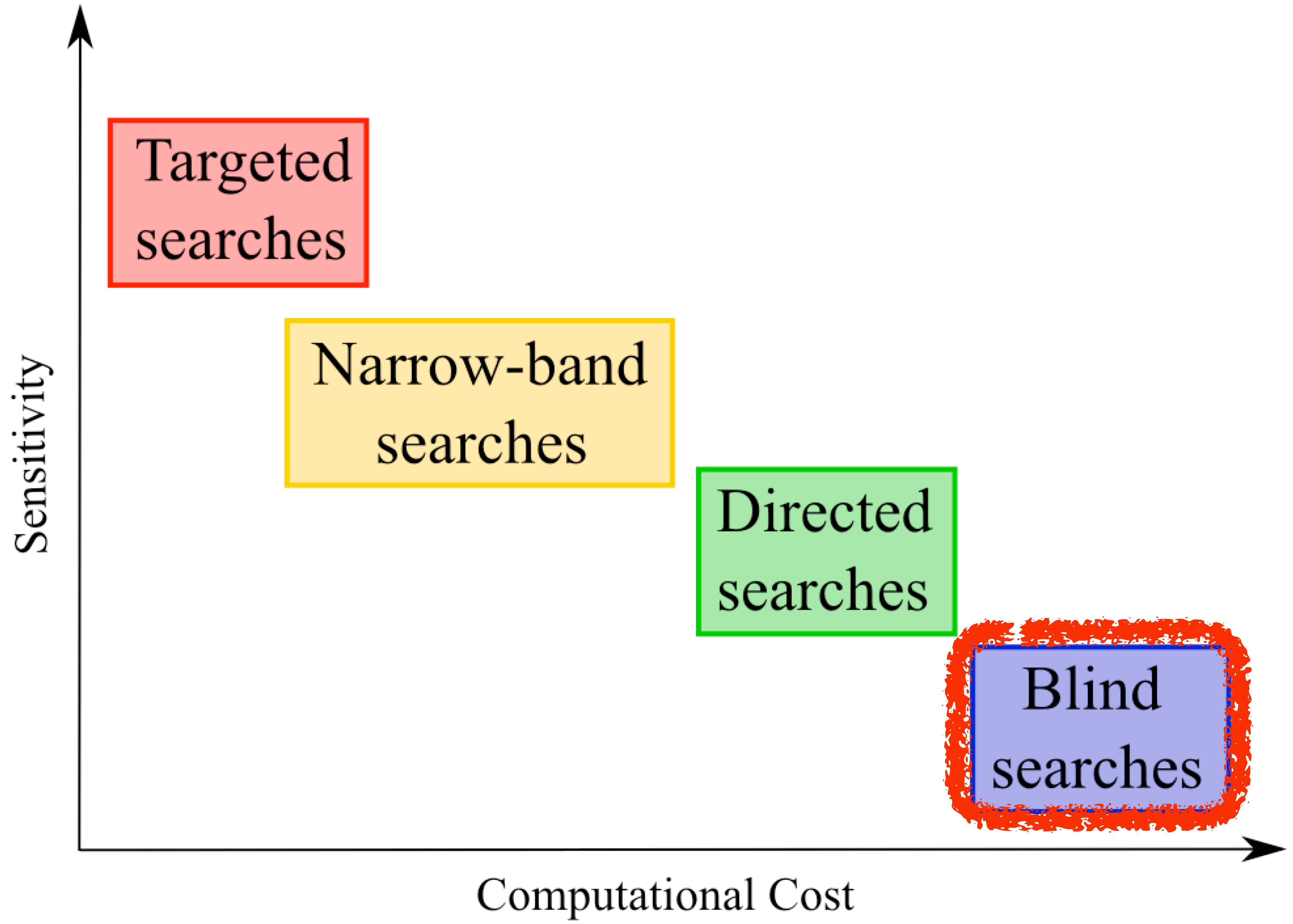
Jasper Martins

Maria Alessandra Papa, Benjamin Steltner

Where are we going in this talk?

Earlier (later) talks:

- Different detection techniques:
  - Hough-Transform (Badri's Talk)
  - $\mathscr{F}-$Statistic (Reinhard's Talk)
  - Falcon-Pipeline (Vladimir's Talk)
  - Many more....
- Different search methods:
  - Grid-based / template-bank based searches
  - Bayesian Parameter Estimation
  - Neural networks / machine learning



Source: Sieniawska and Berger, Universe 5, 217 (2019)
(and Karl's talk)

All-Sky Searches

- Significant computational effort

- Semi-coherent methods reduce the computational cost, but have significant false-alarm rates!

➡ Hierarchical Searches

Initial Search Stage

- Most computational expensive part of the search
- Sets baseline sensitivity:
  - What was not found can never be recovered later

Follow-up Stages

- Constrain overall sensitivity:
  - What was not followed-up can not be detected

➡ More efficient pipelines: Follow-up more candidates and increase sensitivity

➡ Bayesian pipelines promising / successful alternatives

Bayesian Parameter Estimation

Bayes' Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Prior
- Likelihood
- Evidence Z
- Posterior

Bayesian Parameter Estimation

Bayes' Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Prior
- Likelihood
- Evidence Z
- Posterior

Credible Regions

$$\mathrm{CR}(\epsilon) := \mathrm{argmin}_{\Theta}\, \mathrm{Vol}\left(\Theta \subset \mathscr{R} \mid \int_{\Theta} \mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) \mathrm{d}\vartheta = \epsilon \right)$$

Bayesian Parameter Estimation

Bayes' Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Prior
- Likelihood
- Evidence Z
- Posterior

Credible Regions

$$\mathrm{CR}(\epsilon) := \mathrm{argmin}_{\Theta}\, \mathrm{Vol}\left(\Theta \subset \mathscr{R} \mid \int_{\Theta} \mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) \mathrm{d}\vartheta = \epsilon \right)$$

Posterior describes where to allocate follow-up resources efficiently!

Evidence (Bayes factor) can be used like a detection statistic!

Posterior-based search regions are the smallest possible at given credible level!

Bayesian Parameter Estimation

Bayes' Theorem

$$P(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M}) d\vartheta}$$

- Prior ⇒ Chosen via Jeffreys invariance principle
- Likelihood ⇒ $\exp(2\mathscr{F})$
- Evidence Z
- Posterior ⇒ Monte Carlo sampling

Credible Regions

$$CR(\epsilon) := \mathrm{argmin}_{\Theta} \mathrm{Vol}\left(\Theta \subset \mathscr{R} \mid \int_{\Theta} P(\vartheta \mid \mathscr{D}, \mathscr{M}) d\vartheta = \epsilon \right)$$

Posterior describes where to allocate follow-up resources efficiently!

Evidence (Bayes factor) can be used like a detection statistic!

Posterior-based search regions are the smallest possible at given credible level!

Bayesian Parameter Estimation

Bayes' Theorem

$$P(\vartheta \mid \mathcal{D}, \mathcal{M}) = \frac{P(\mathcal{D} \mid \vartheta, \mathcal{M}) P(\vartheta \mid \mathcal{M})}{\int_{\mathcal{R}} \ldots}$$

- Posterior ⇒ Obtain throu...
- Prior ⇒ Chosen via Jeffre...
- Likelihood ⇒ $\mathcal{F}$ − Stati...
- Evidence Z

Credible Regions

$$\mathrm{CR}(\epsilon) := \mathrm{argmin}_{\Theta} \mathrm{Vol}\left(\Theta \subset \mathcal{R} \mid \int_{\Theta} P(\vartheta \mid \mathcal{D}, \mathcal{M}) \mathrm{d}\vartheta = \epsilon \right)$$

Core idea: Use posteriors directly as the prior distribution for a MC-run at a higher coherence time!

Posterior describes where to allocate follow-up resources efficiently!

Evidence (Bayes factor) can be used like a detection statistic!

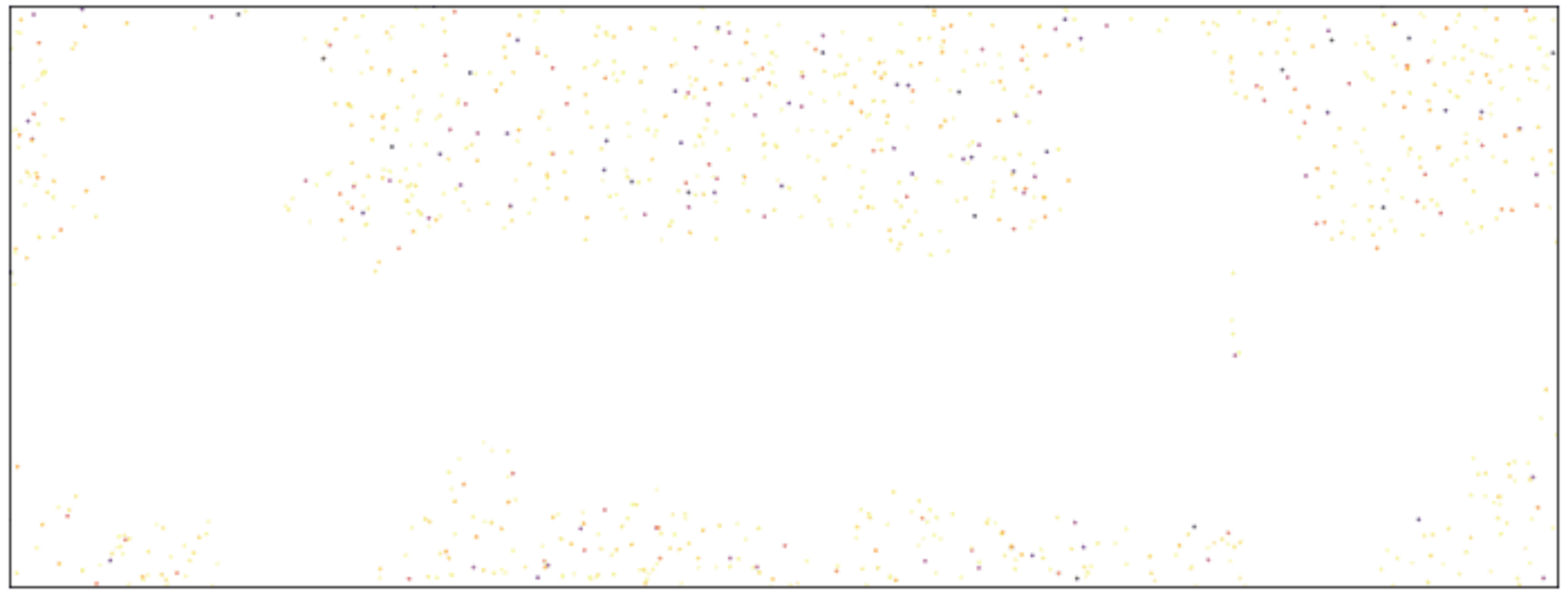
Posterior-based search regions are the smallest possible at given credible level!

Implementation I - Sampler

- There is a vast library of Monte-Carlo Sampling algorithms. Which is the best for the task?

Nested Sampling

- Originally designed to compute evidences
  - Focus on exploration of the prior volume
  - Robust for multimodal likelihoods
  - Predictable convergence
- One key parameter for convergence rate and accuracy: n_{live}
- Interactive Demo: here
- We use dynesty (through Bilby and SWIGLAL)



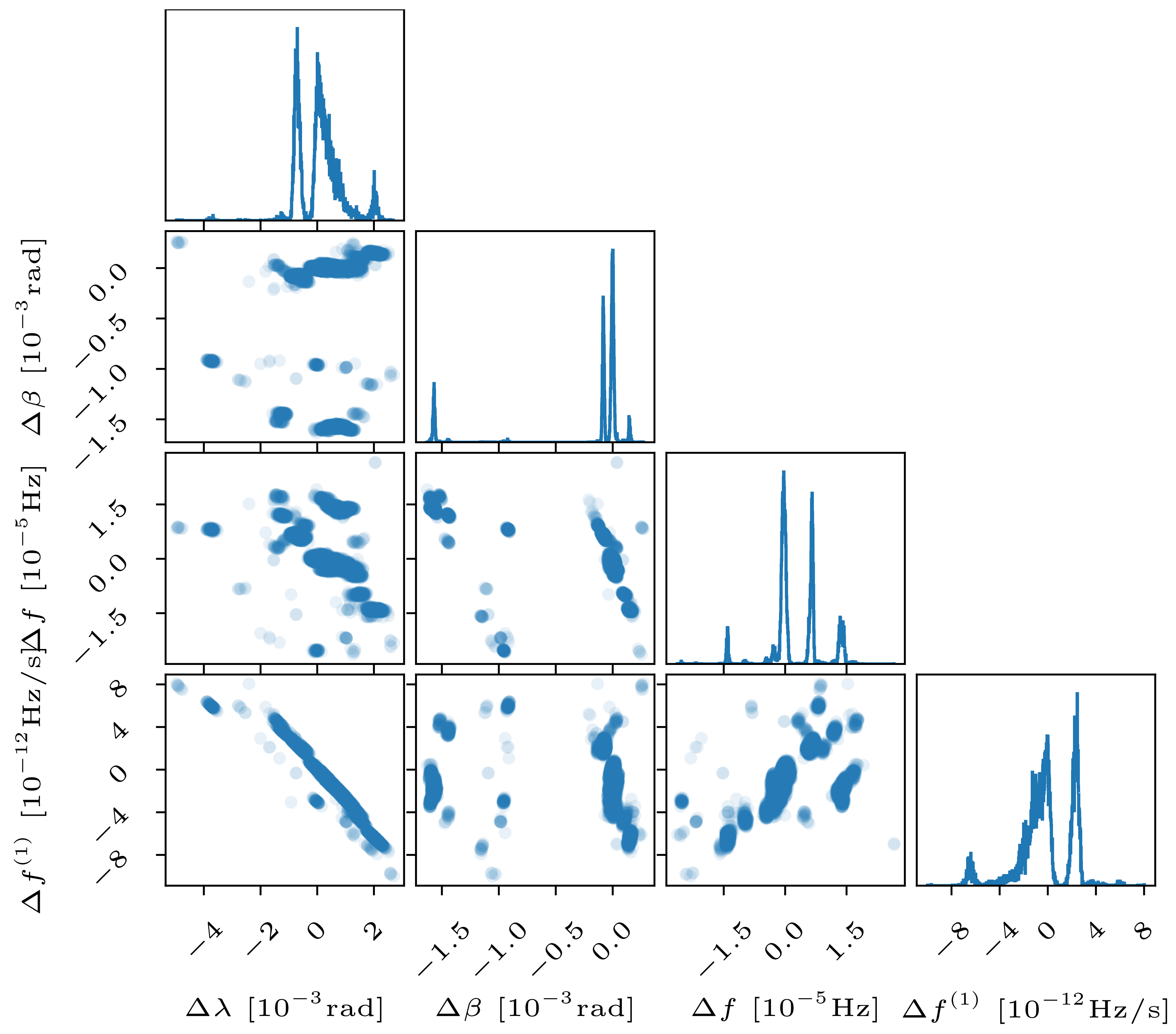
Source: https://dynesty.readthedocs.io/en/stable/_images/title.gif

Implementation II — A Posterior Model

Bayes Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!

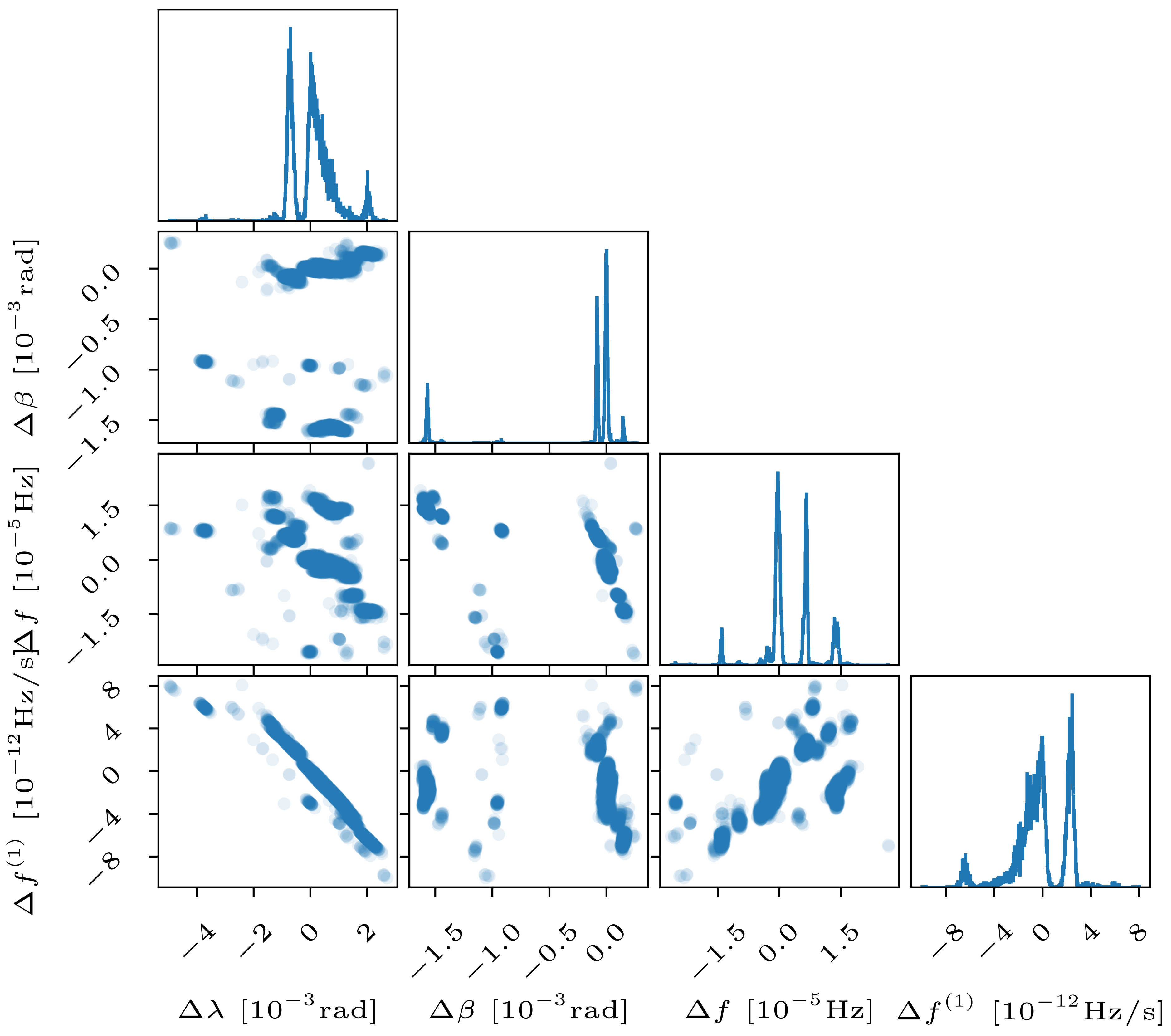


Implementation II — A Posterior Model

Bayes Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\,\mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\,\mathrm{P}(\vartheta \mid \mathscr{M})\mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples



Implementation II — A Posterior Model

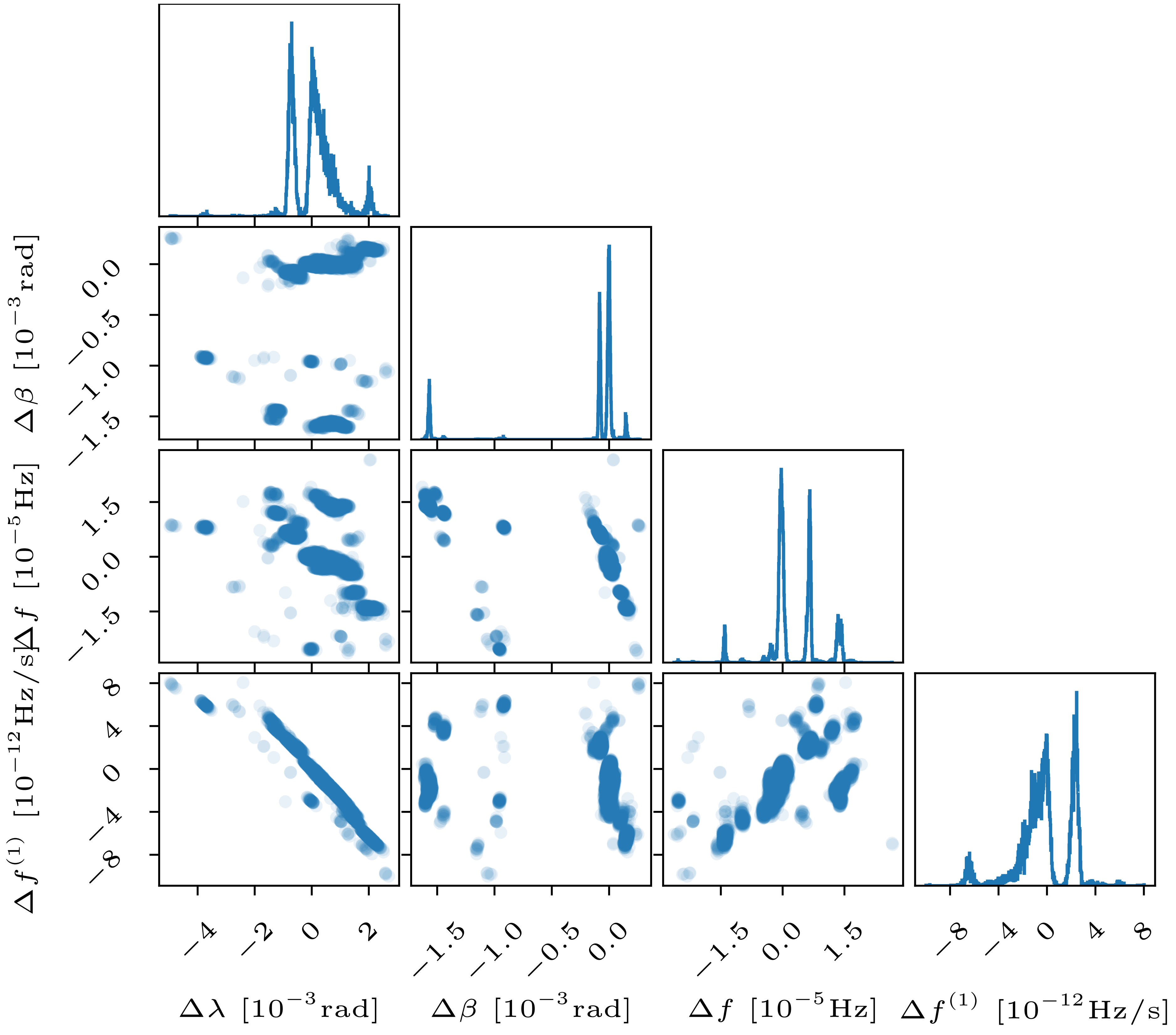
Bayes Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples

Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering



Implementation II — A Posterior Model

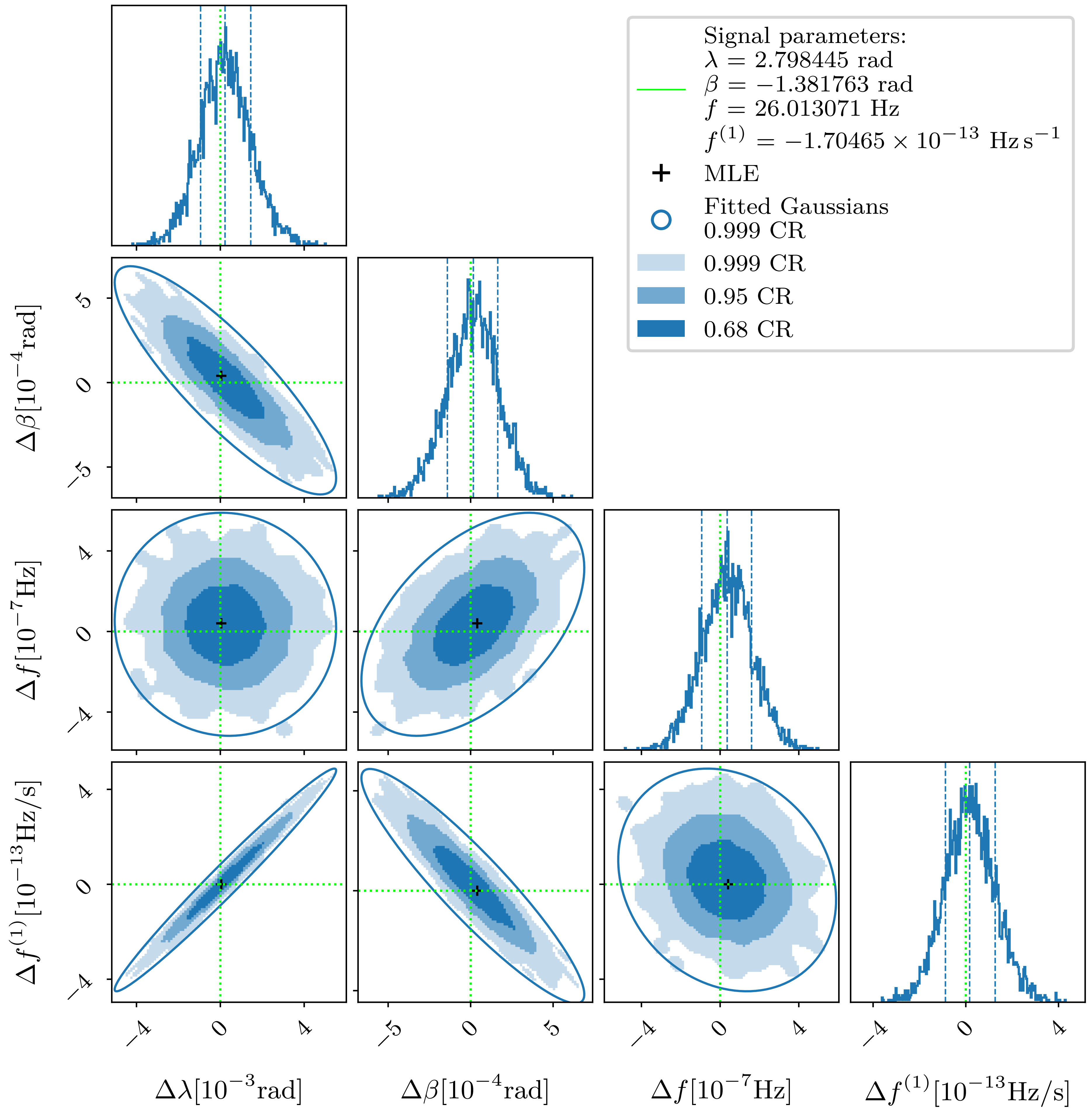
Bayes Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples

Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering

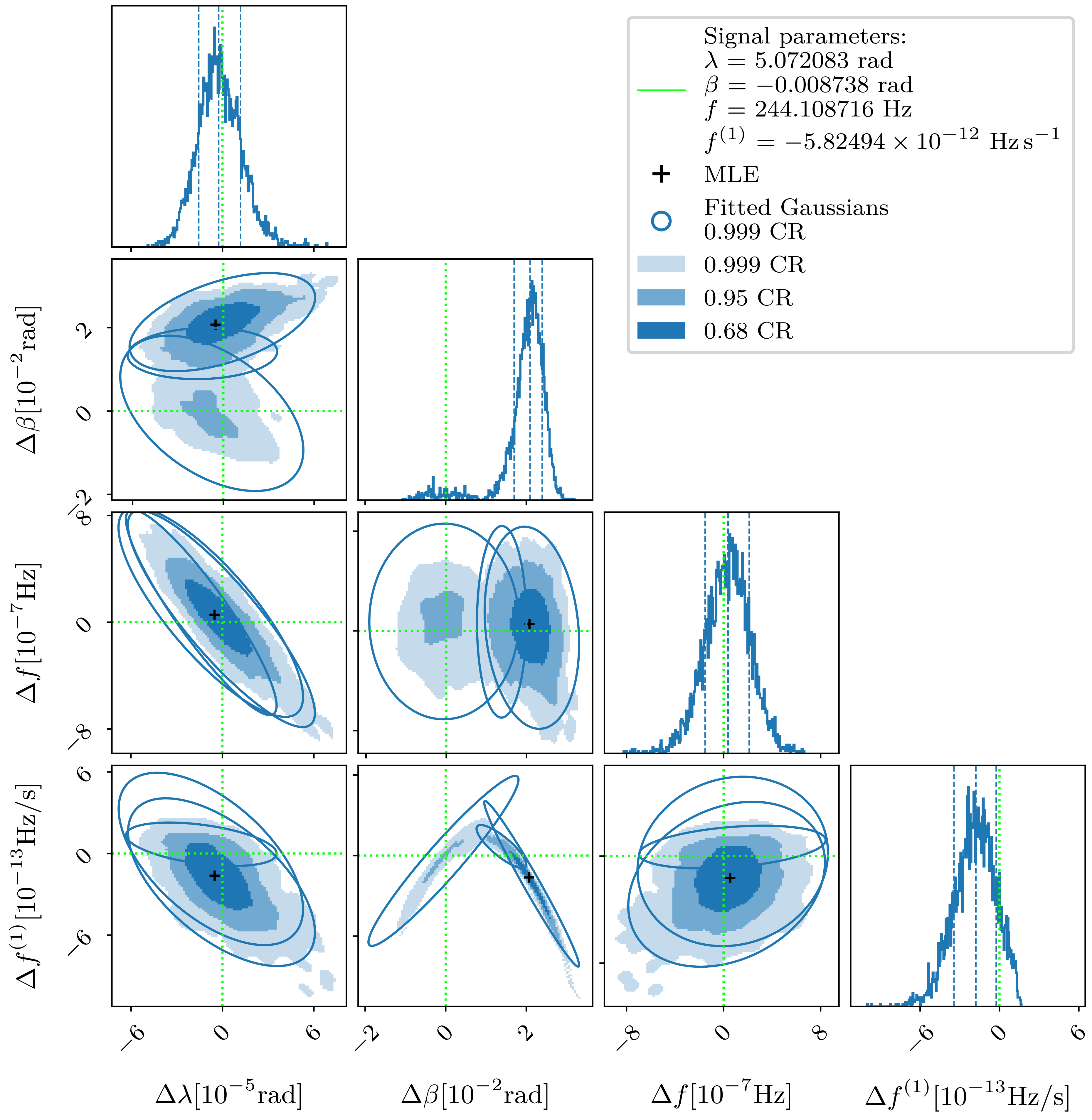


Implementation II — A Posterior Model

Bayes Theorem

$$P(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M}) d\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples



Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering

Implementation II — A Posterior Model

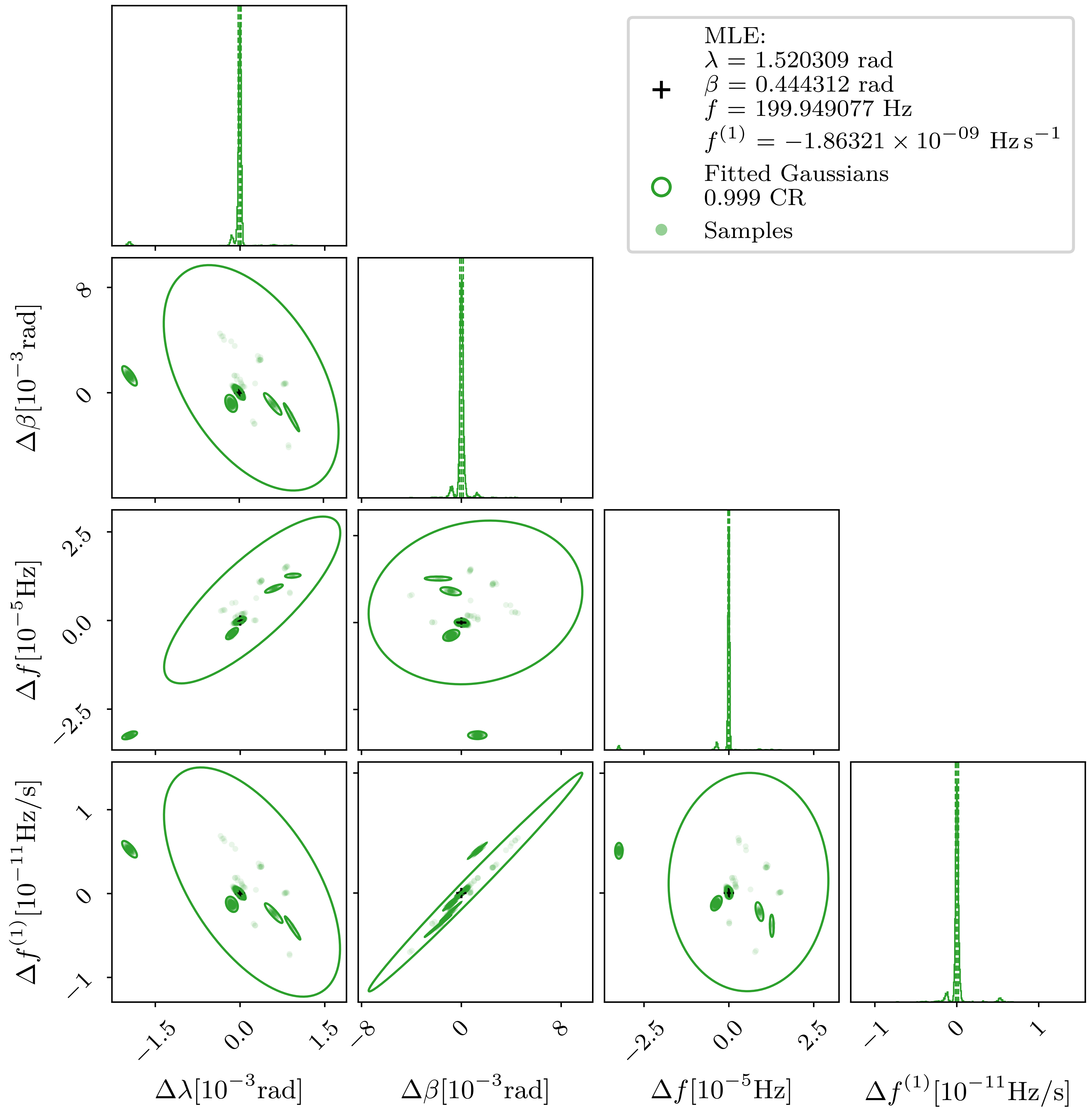
Bayes Theorem

$$\mathrm{P}(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{\mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} \mathrm{P}(\mathscr{D} \mid \vartheta, \mathscr{M})\, \mathrm{P}(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples

Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering



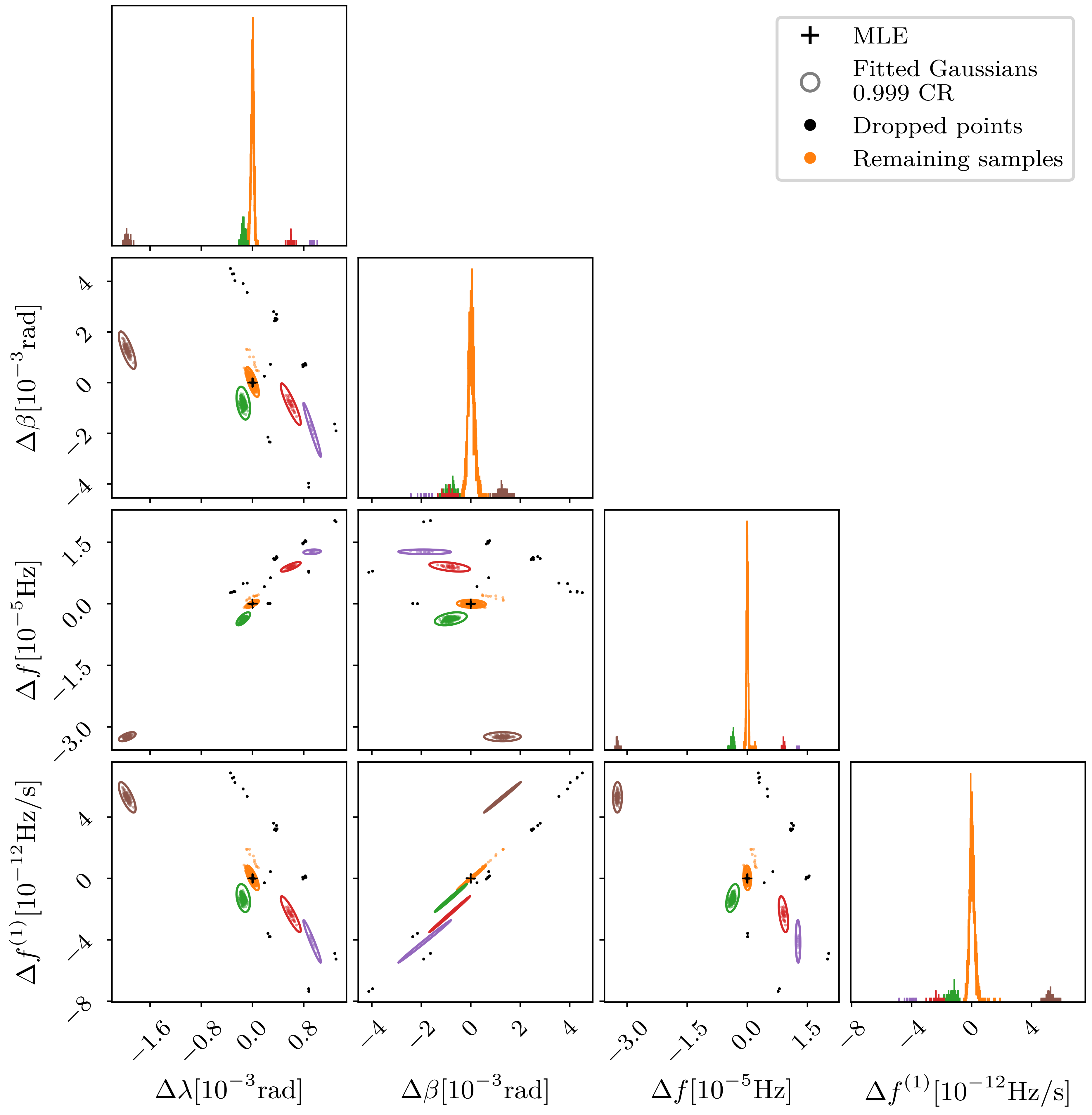
Implementation II — A Posterior Model

Bayes Theorem

$$P(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!

- Need a model to describe the posterior that can be used to propose new (prior) samples



Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering

Implementation II — A Posterior Model

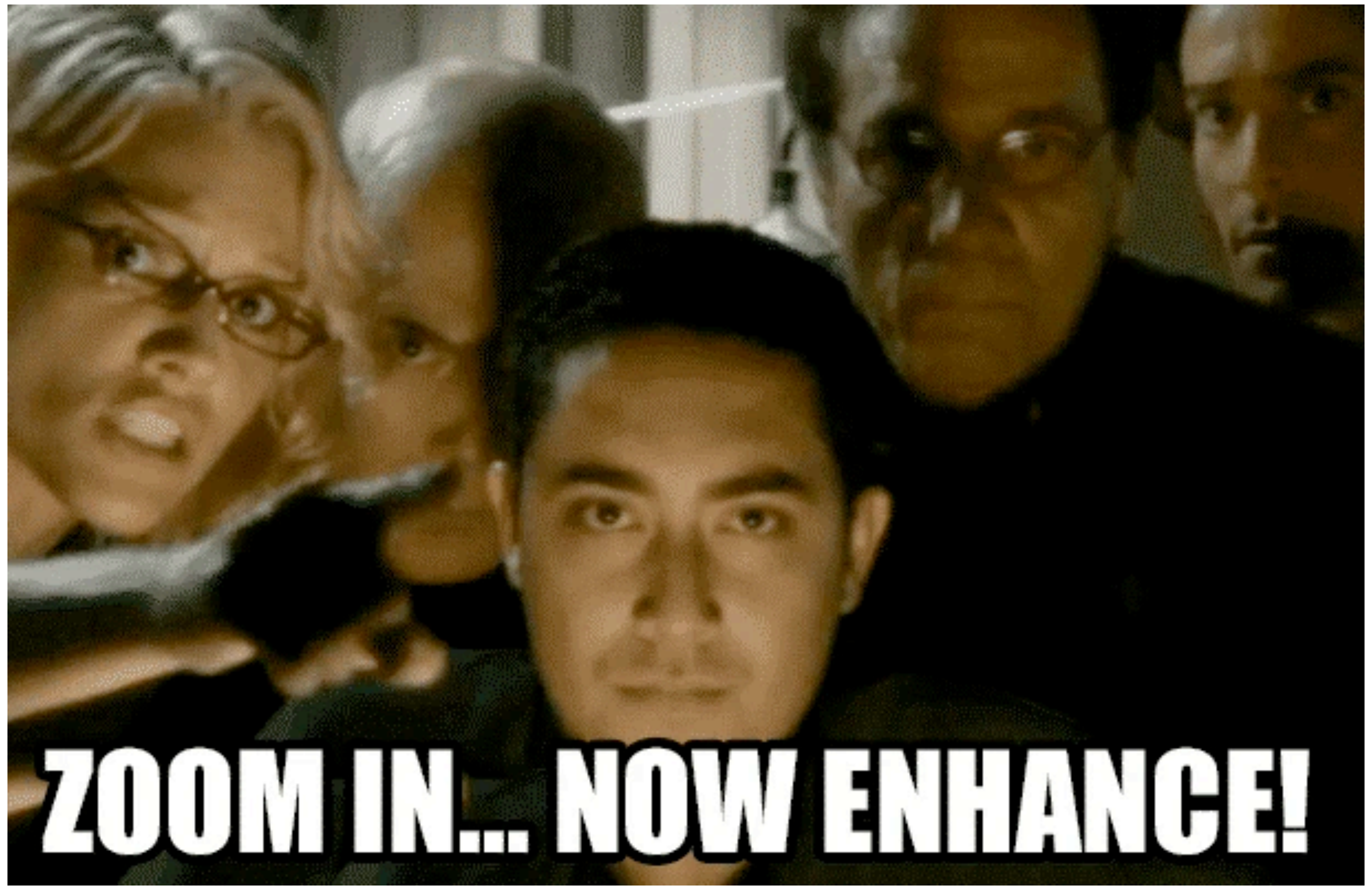
Bayes Theorem

$$P(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M}) d\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples

Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering

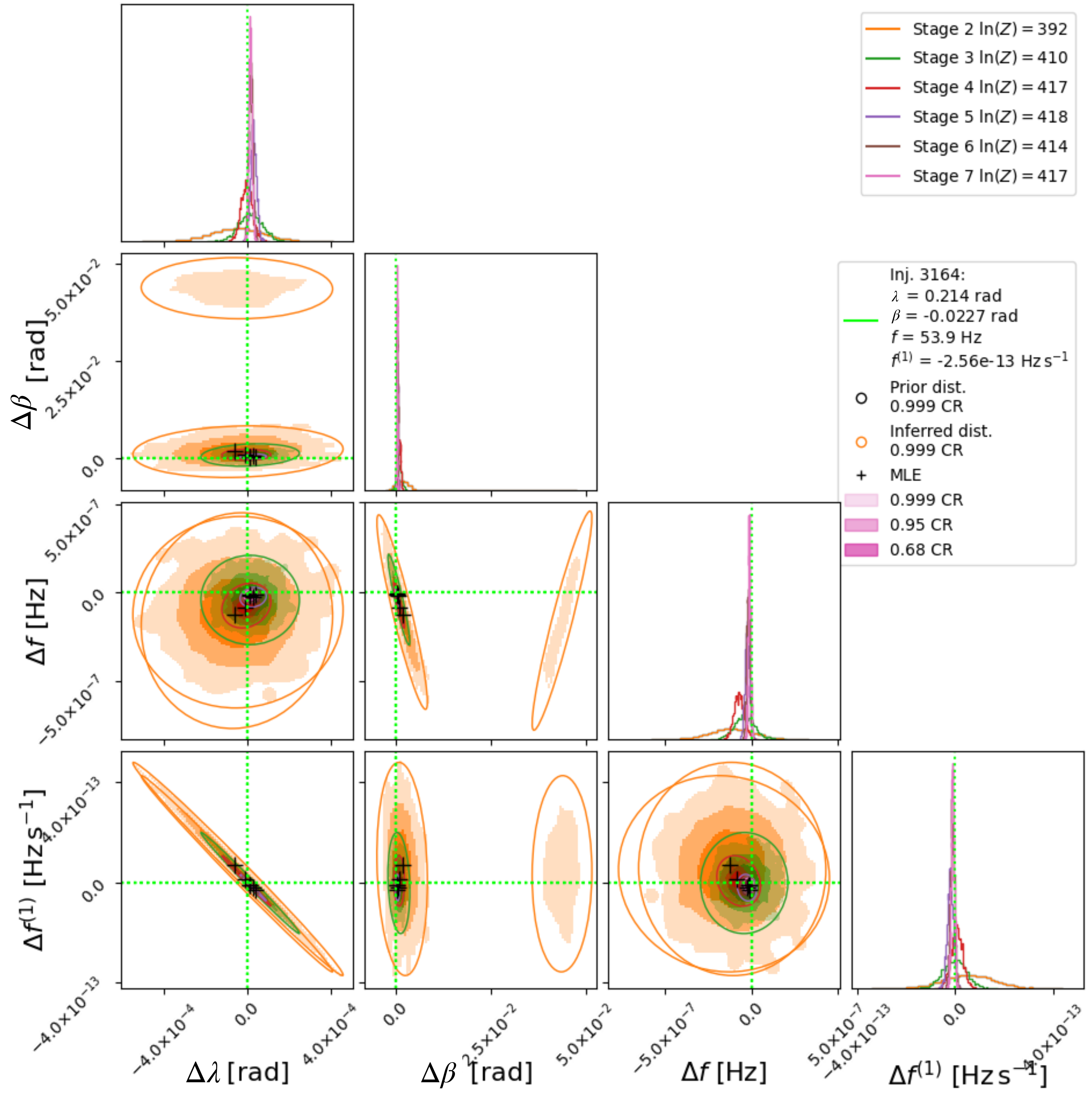


Implementation II — A Posterior Model

Bayes Theorem

$$P(\vartheta \mid \mathscr{D}, \mathscr{M}) = \frac{P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M})}{\int_{\mathscr{R}} P(\mathscr{D} \mid \vartheta, \mathscr{M})\, P(\vartheta \mid \mathscr{M}) \mathrm{d}\vartheta}$$

- Core Idea of HierarchMC: Use posterior as prior, but: MC-sampling only yields posterior samples!
- Need a model to describe the posterior that can be used to propose new (prior) samples



Solution: Gaussian Mixture Models (core: scikit-learn)

- Do not know the number of modes
  ➡ Bisecting, recursive algorithm
- Spurious Samples
  ➡ DBSCAN pre-clustering

Results

Results - Performance Evaluation

- Question 1: In standard follow-up procedures, at which stage **can** HierarchMC be used?
- Question 2: Once HierarchMC can be used, is it competitive?

Results - Performance Evaluation

- Question 1: In standard follow-up procedures, at which stage **can** HierarchMC be used?
- Question 2: Once HierarchMC can be used, is it competitive?

The "Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data" — Steltner et al. (2023)

- Most sensitive all-sky search for isolated sources for the target parameter space published
- Conducted on LIGO O3a data (+ LIGO O3b)
- Reference signal population: ≈ 1600 signals with sensitivity depths $\mathrm{Unif}(50,65)\ [1/\sqrt{\mathrm{Hz}}]$
- Recovered 6 (out of 7 possible) hardware injections

The "Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data"

Target:
$f \in [20\,\mathrm{Hz}, 800\,\mathrm{Hz}], \dot{f} \in [-2.6 \times 10^{-9}\,\mathrm{Hz/s}, 2.6 \times 10^{-10}\,\mathrm{Hz/s}]$

Overview of the Full Hierarchy of Searches

Search	T_{coh} (hr)	N_{seg}	δf (μHz)	$\delta \dot{f}$ (10^{-14} Hz s^{-1})	m_{sky}	$\langle\mu\rangle$ (10^{-2})	Δf (μHz)	$\Delta \dot{f}$ (10^{-14} Hz s^{-1})	$\frac{r_{sky}}{d(0.002)}$	N_{in}	N_{out}
Stage 0	120	37	2	60	0.002	56	Full range	Full range	All sky	6.7×10^{18}	3513855
Stage 1	120	37	1	15	0.0002	31	1000	11,250	10.0	3,513,855	386,429
Stage 2	120	37	1	2	2×10^{-6}	2.5	50	1200	1.0	386,429	35,635
Stage 3	240	19	0.1	0.2	1.0×10^{-8}	0.06	5	200	0.2	35,635	5116
Stage 4	490	9	0.013	0.064	3.2×10^{-10}	<0.01	0.5	35	0.04	5116	1387
Stage 5	1100	4	0.001	0.032	1.4×10^{-10}	<0.01	0.2	20	0.017	1387	310
Stage 6	2200	2	0.001	0.009	2.8×10^{-11}	<0.01	0.1	8.5	0.008	310	54
Stage 7	Coherent	1	0.001	0.0063	1×10^{-11}	<0.01	0.07	6	0.0057	54	12
Stage 8	O3b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	12	6
Stage 9[b]	O3a+b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	6	6

Table 1 in Stelter et al. (2023)

The "Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data"

Target:
$f \in [20\,\mathrm{Hz}, 800\,\mathrm{Hz}], \dot{f} \in [-2.6 \times 10^{-9}\,\mathrm{Hz/s}, 2.6 \times 10^{-10}\,\mathrm{Hz/s}]$

Overview of the Full Hierarchy of Searches

Search	T_{coh} (hr)	N_{seg}	δf (μHz)	$\delta\dot{f}$ (10^{-14} Hz s^{-1})	m_{sky}	$\langle\mu\rangle$ (10^{-2})	Δf (μHz)	$\Delta\dot{f}$ (10^{-14} Hz s^{-1})	$\frac{r_{sky}}{d(0.002)}$	N_{in}	N_{out}
Stage 0	120	37	2	60	0.002	56	Full range	Full range	All sky	6.7×10^{18}	3513855
Stage 1	120	37	1	15	0.0002	31	1000	11,250	10.0	3,513,855	386,429
Stage 2	120	37	1	2	2×10^{-6}	2.5	50	1200	1.0	386,429	35,635
Stage 3	240	19	0.1	0.2	1.0×10^{-8}	0.06	5	200	0.2	35,635	5116
Stage 4	490	9	0.013	0.064	3.2×10^{-10}	<0.01	0.5	35	0.04	5116	1387
Stage 5	1100	4	0.001	0.032	1.4×10^{-10}	<0.01	0.2	20	0.017	1387	310
Stage 6	2200	2	0.001	0.009	2.8×10^{-11}	<0.01	0.1	8.5	0.008	310	54
Stage 7	Coherent	1	0.001	0.0063	1×10^{-11}	<0.01	0.07	6	0.0057	54	12
Stage 8	O3b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	12	6
Stage 9[b]	O3a+b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	6	6

Table 1 in Stelter et al. (2023)

Stage 0: Einstein@Home

Stage 1-9: Atlas

Stage 1: Clustering + Refinement

Stage 2: Refinement

The “Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data”

Target:
$f \in [20\,\mathrm{Hz}, 800\,\mathrm{Hz}], \dot{f} \in [-2.6\times10^{-9}\,\mathrm{Hz/s}, 2.6\times10^{-10}\,\mathrm{Hz/s}]$

Overview of the Full Hierarchy of Searches

Search	T_{coh} (hr)	N_{seg}	δf (μHz)	$\delta \dot{f}$ (10^{-14} Hz s^{-1})	m_{sky}	$\langle\mu\rangle$ (10^{-2})	Δf (μHz)	$\Delta \dot{f}$ (10^{-14} Hz s^{-1})	$\frac{r_{sky}}{d\,(0.002)}$	N_{in}	N_{out}
Stage 0	120	37	2	60	0.002	56	Full range	Full range	All sky	6.7×10^{18}	3513855
Stage 1	120	37	1	15	0.0002	31	1000	11,250	10.0	3,513,855	386,429
Stage 2	120	37	1	2	2×10^{-6}	2.5	50	1200	1.0	386,429	35,635
Stage 3	240	19	0.1	0.2	1.0×10^{-8}	0.06	5	200	0.2	35,635	5116
Stage 4	490	9	0.013	0.064	3.2×10^{-10}	<0.01	0.5	35	0.04	5116	1387
Stage 5	1100	4	0.001	0.032	1.4×10^{-10}	<0.01	0.2	20	0.017	1387	310
Stage 6	2200	2	0.001	0.009	2.8×10^{-11}	<0.01	0.1	8.5	0.008	310	54
Stage 7	Coherent	1	0.001	0.0063	1×10^{-11}	<0.01	0.07	6	0.0057	54	12
Stage 8	O3b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	12	6
Stage 9[b]	O3a+b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	6	6

Table 1 in Stelter et al. (2023)

Stage 0: Einstein@Home

Stage 1-9: Atlas

Stage 1: Clustering + Refinement

Stage 2: Refinement

The "Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data"

Target:
$f \in [20\,\text{Hz}, 800\,\text{Hz}], \dot{f} \in [-2.6\times10^{-9}\,\text{Hz/s}, 2.6\times10^{-10}\,\text{Hz/s}]$

Overview of the Full Hierarchy of Searches

Search	T_{coh} (hr)	N_{seg}	δf (μHz)	$\delta\dot{f}$ (10^{-14} Hz s^{-1})	m_{sky}	$\langle\mu\rangle$ (10^{-2})	Δf (μHz)	$\Delta\dot{f}$ (10^{-14} Hz s^{-1})	$\frac{r_{sky}}{d(0.002)}$	N_{in}	N_{out}
Stage 0	120	37	2	60	0.002	56	Full range	Full range	All sky	6.7×10^{18}	3513855
Stage 1	120	37	1	15	0.0002	31	1000	11,250	10.0	3,513,855	386,429
Stage 2	120	37	1	2	2×10^{-6}	2.5	50	1200	1.0	386,429	35,635
Stage 3	240	19	0.1	0.2	1.0×10^{-8}	0.06	5	200	0.2	35,635	5116
Stage 4	490	9	0.013	0.064	3.2×10^{-10}	<0.01	0.5	35	0.04	5116	1387
Stage 5	1100	4	0.001	0.032	1.4×10^{-10}	<0.01	0.2	20	0.017	1387	310
Stage 6	2200	2	0.001	0.009	2.8×10^{-11}	<0.01	0.1	8.5	0.008	310	54
Stage 7	Coherent	1	0.001	0.0063	1×10^{-11}	<0.01	0.07	6	0.0057	54	12
Stage 8	O3b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	12	6
Stage 9[b]	O3a+b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	⩾0.07[a]	6	0.0057	6	6

Stage 0: Einstein@Home

Stage 1-9: Atlas

Stage 1: Clustering + Refinement

Stage 2: Refinement

Table 1 in Stelter et al. (2023)

The "Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data"

Target:
$f \in [20\,\mathrm{Hz}, 800\,\mathrm{Hz}], \dot{f} \in [-2.6 \times 10^{-9}\,\mathrm{Hz/s}, 2.6 \times 10^{-10}\,\mathrm{Hz/s}]$

Overview of the Full Hierarchy of Searches

Search	T_{coh} (hr)	N_{seg}	δf (μHz)	$\delta\dot{f}$ (10^{-14} Hz s^{-1})	m_{sky}	$\langle\mu\rangle$ (10^{-2})	Δf (μHz)	$\Delta\dot{f}$ (10^{-14} Hz s^{-1})	$\frac{r_{sky}}{d(0.002)}$	N_{in}	N_{out}
Stage 0	120	37	2	60	0.002	56	Full range	Full range	All sky	6.7×10^{18}	3513855
Stage 1	120	37	1	15	0.0002	31	1000	11,250	10.0	3,513,855	386,429
Stage 2	120	37	1	2	2×10^{-6}	2.5	50	1200	1.0	386,429	35,635
Stage 3	240	19	0.1	0.2	1.0×10^{-8}	0.06	5	200	0.2	35,635	5116
Stage 4	490	9	0.013	0.064	3.2×10^{-10}	<0.01	0.5	35	0.04	5116	1387
Stage 5	1100	4	0.001	0.032	1.4×10^{-10}	<0.01	0.2	20	0.017	1387	310
Stage 6	2200	2	0.001	0.009	2.8×10^{-11}	<0.01	0.1	8.5	0.008	310	54
Stage 7	Coherent	1	0.001	0.0063	1×10^{-11}	<0.01	0.07	6	0.0057	54	12
Stage 8	O3b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	$\geqslant$0.07[a]	6	0.0057	12	6
Stage 9[b]	O3a+b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	$\geqslant$0.07[a]	6	0.0057	6	6

Stage 0: Einstein@Home

Stage 1-9: Atlas

Stage 1: Clustering + Refinement

Stage 2: Refinement

Table 1 in Stelter et al. (2023)

The "Deep Einstein@Home All-sky Search for Continuous Gravitational Waves in LIGO O3 Public Data"

Target:
$f \in [20\,\mathrm{Hz}, 800\,\mathrm{Hz}], \dot{f} \in [-2.6 \times 10^{-9}\,\mathrm{Hz/s}, 2.6 \times 10^{-10}\,\mathrm{Hz/s}]$

Overview of the Full Hierarchy of Searches

Search	T_{coh} (hr)	N_{seg}	δf (μHz)	$\delta\dot{f}$ (10^{-14} Hz s^{-1})	m_{sky}	$\langle\mu\rangle$ (10^{-2})	Δf (μHz)	$\Delta\dot{f}$ (10^{-14} Hz s^{-1})	$\frac{r_{sky}}{d(0.002)}$	N_{in}	N_{out}
Stage 0	120	37	2	60	0.002	56	Full range	Full range	All sky	6.7×10^{18}	3513855
Stage 1	120	37	1	15	0.0002	31	1000	11,250	10.0	3,513,855	386,429
Stage 2	120	37	1	2	2×10^{-6}	2.5	50	1200	1.0	386,429	35,635
Stage 3	240	19	0.1	0.2	1.0×10^{-8}	0.06	5	200	0.2	35,635	5116
Stage 4	490	9	0.013	0.064	3.2×10^{-10}	<0.01	0.5	35	0.04	5116	1387
Stage 5	1100	4	0.001	0.032	1.4×10^{-10}	<0.01	0.2	20	0.017	1387	310
Stage 6	2200	2	0.001	0.009	2.8×10^{-11}	<0.01	0.1	8.5	0.008	310	54
Stage 7	Coherent	1	0.001	0.0063	1×10^{-11}	<0.01	0.07	6	0.0057	54	12
Stage 8	O3b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	$\geqslant$0.07[a]	6	0.0057	12	6
Stage 9[b]	O3a+b coh.	1	0.001	0.0063	1×10^{-11}	<0.01	$\geqslant$0.07[a]	6	0.0057	6	6

Stage 0: Einstein@Home

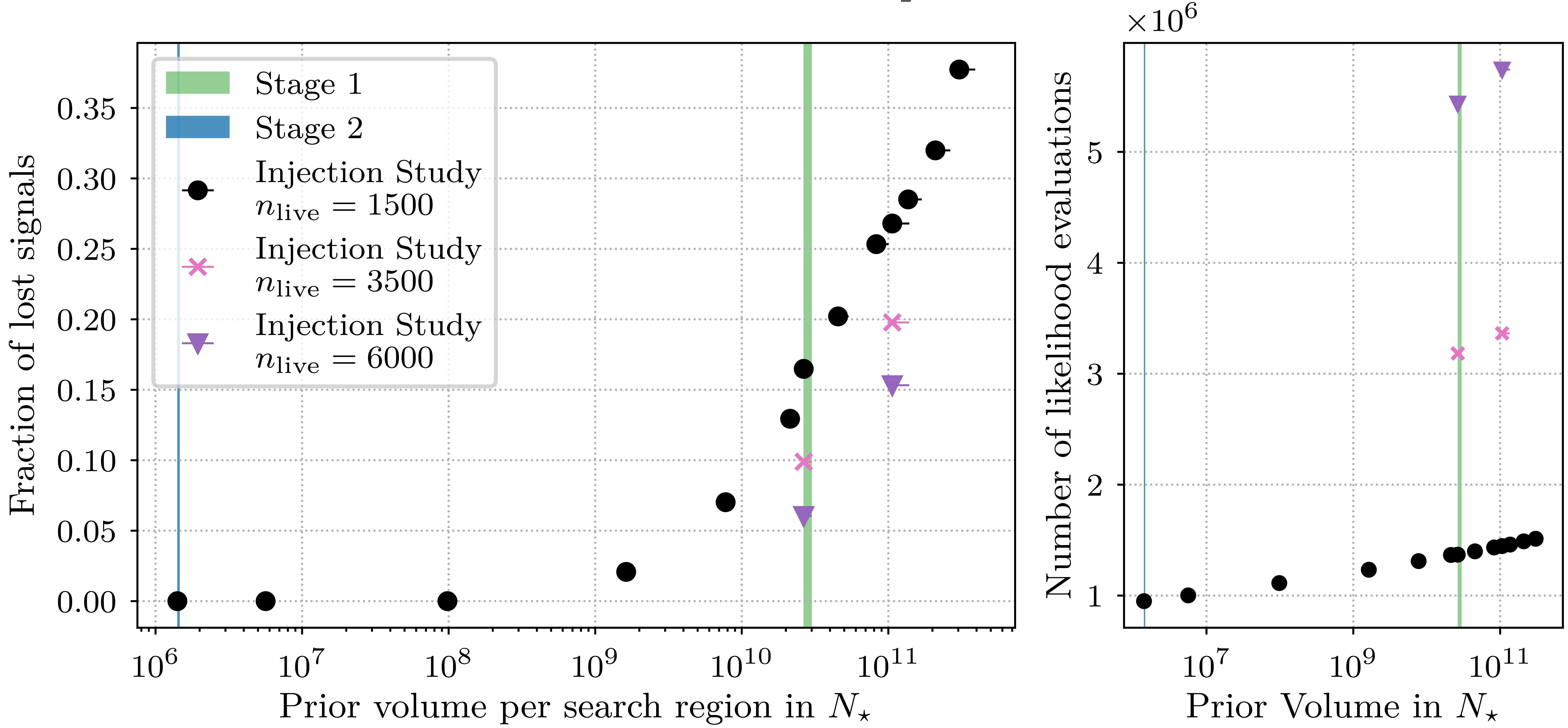
Stage 1-9: Atlas

Stage 1: Clustering + Refinement

Stage 2: Refinement

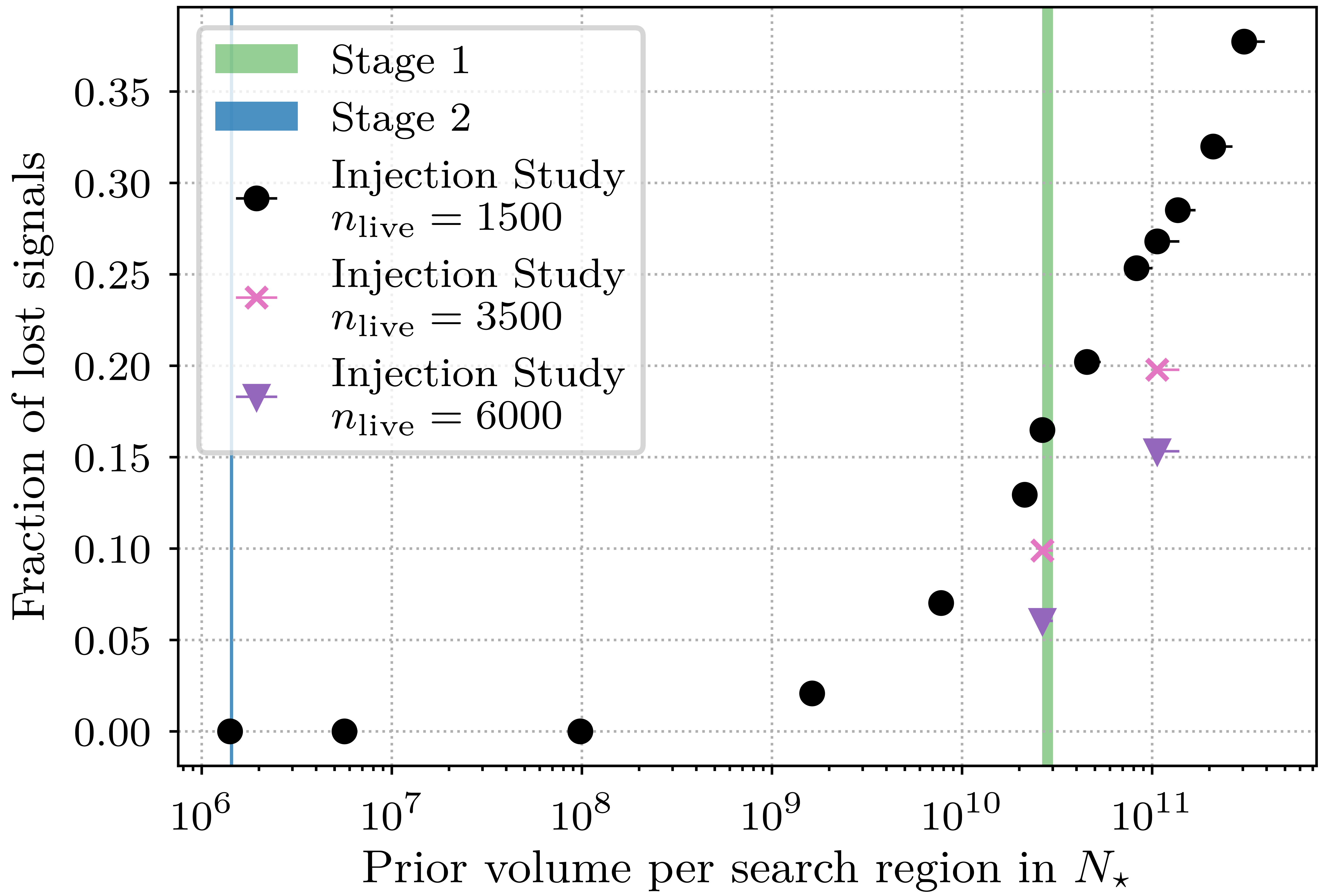
Table 1 in Stelter et al. (2023)

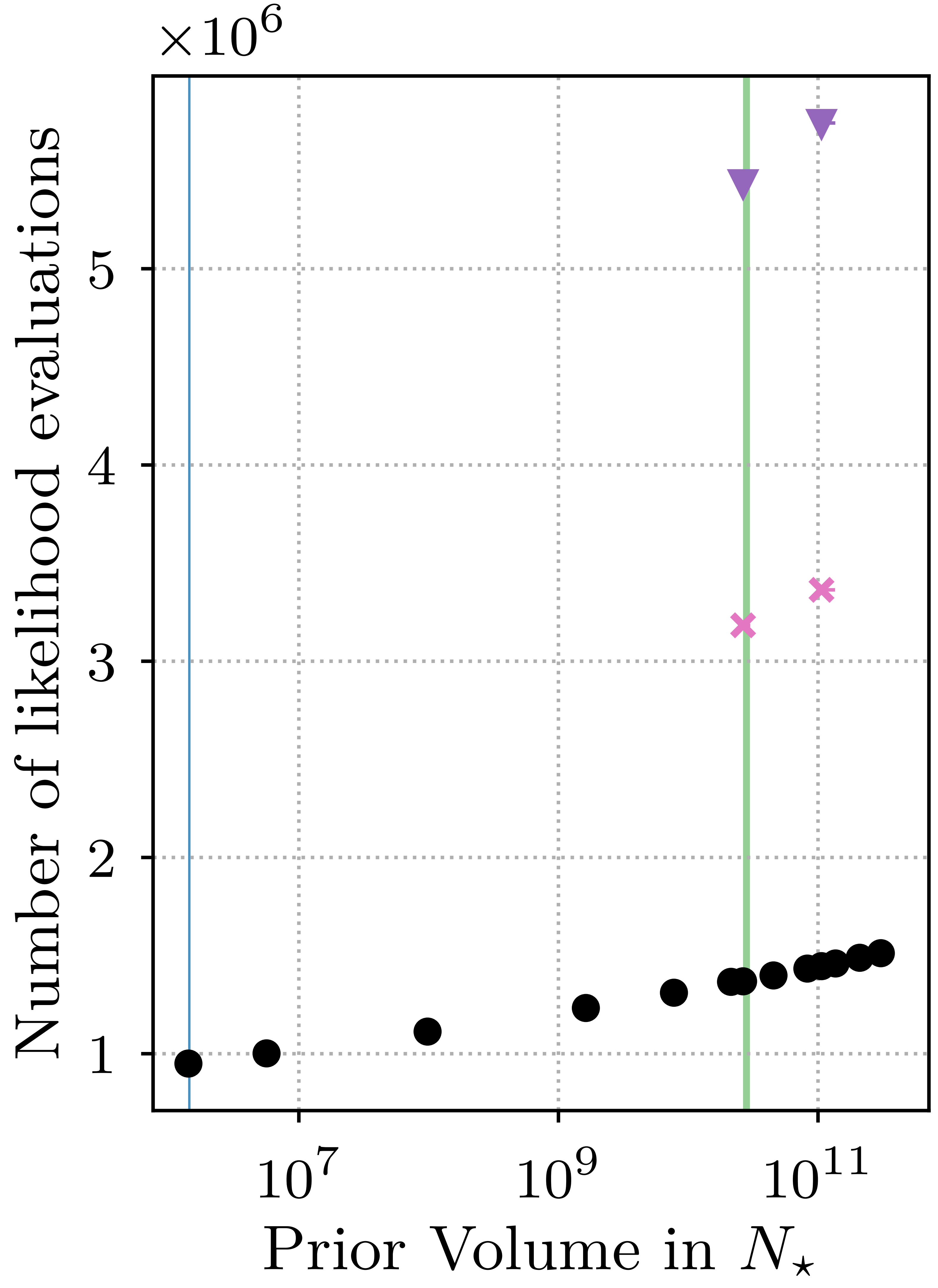
Results — Initial Viability



Results — Initial Viability

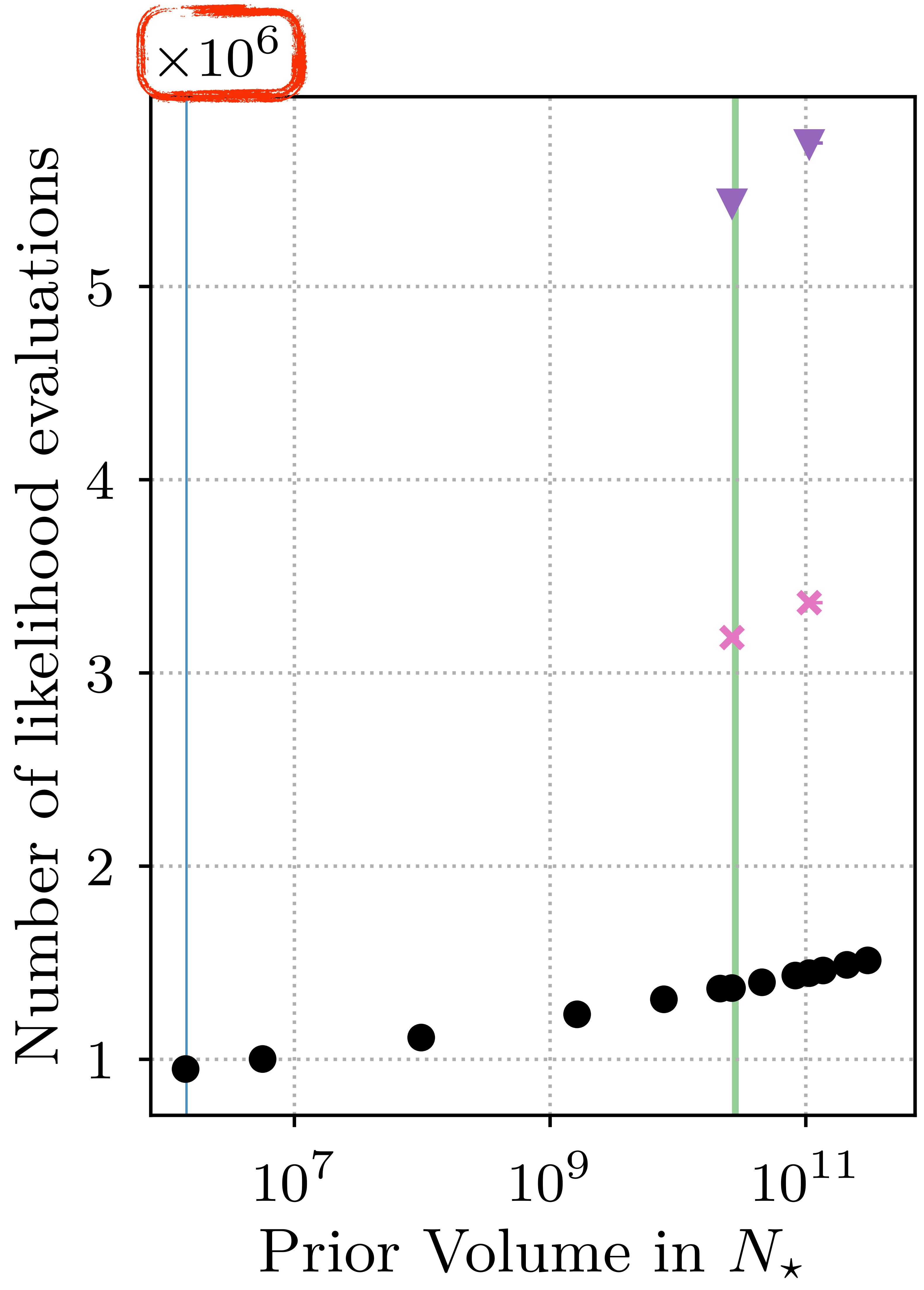
Consistent with recent results (Covas, Prix and Martins, 2024)





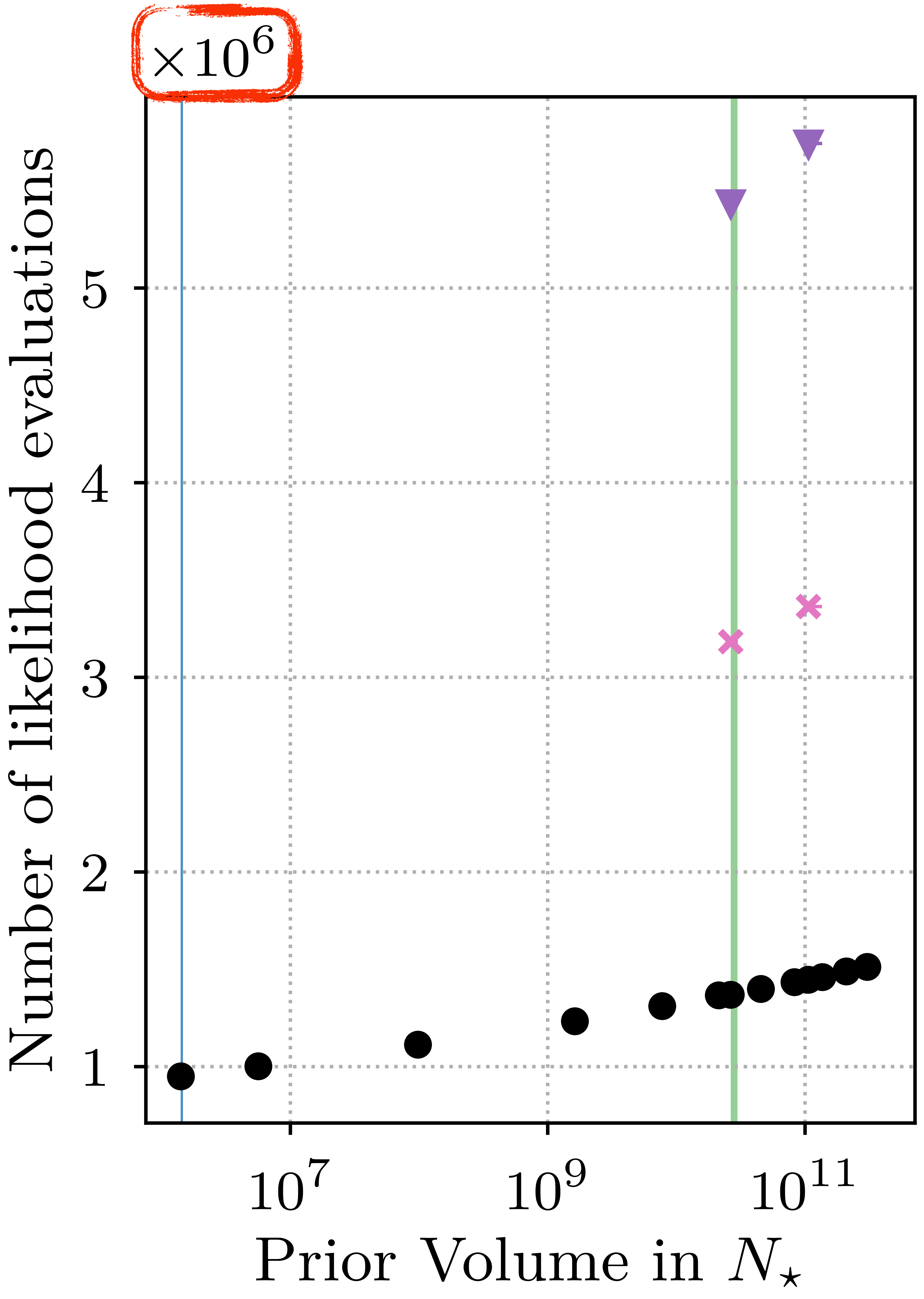
Results — Computational Cost

- HierarchMC is only useful if it is fast
- Stage 2 req. $\approx 10^6$ likelihood evaluations



Results — Computational Cost

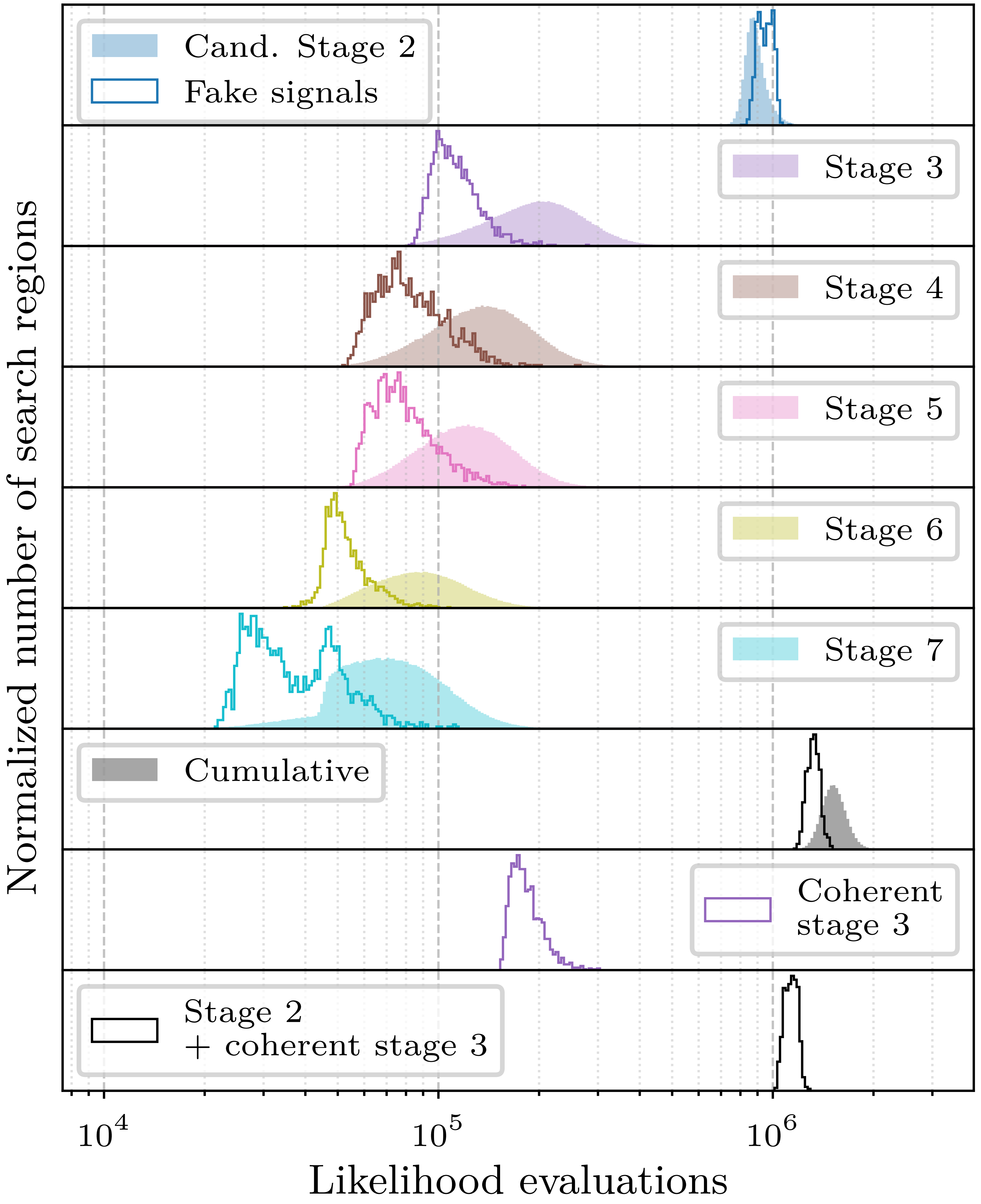
- HierarchMC is only useful if it is fast
- Stage 2 req. $\approx 10^6$ likelihood evaluations
  - 10^5Cand., 10^{-2}s per likelihood evalution $\Rightarrow 10^9$s Runtime
  - Atlas: 10^4 CPU cores $\Rightarrow 10^5$s Cluster-time
  - Cluster time $\mathcal{O}$(days)
- What about subsequent stages?



Results — Computational Cost

- HierarchMC is only useful if it is fast
- Stage 2 req. $\approx 10^6$ likelihood evaluations
  - 10^5Cand., 10^{-2}s per likelihood evalution $\Rightarrow 10^9$s Runtime
  - Atlas: 10^4 CPU cores $\Rightarrow 10^5$s Cluster-time
  - Cluster time $\mathcal{O}$(days)
- What about subsequent stages?
  - Do not sig. increase runtime!

A fully coherent follow-up of **all** candidates is possible

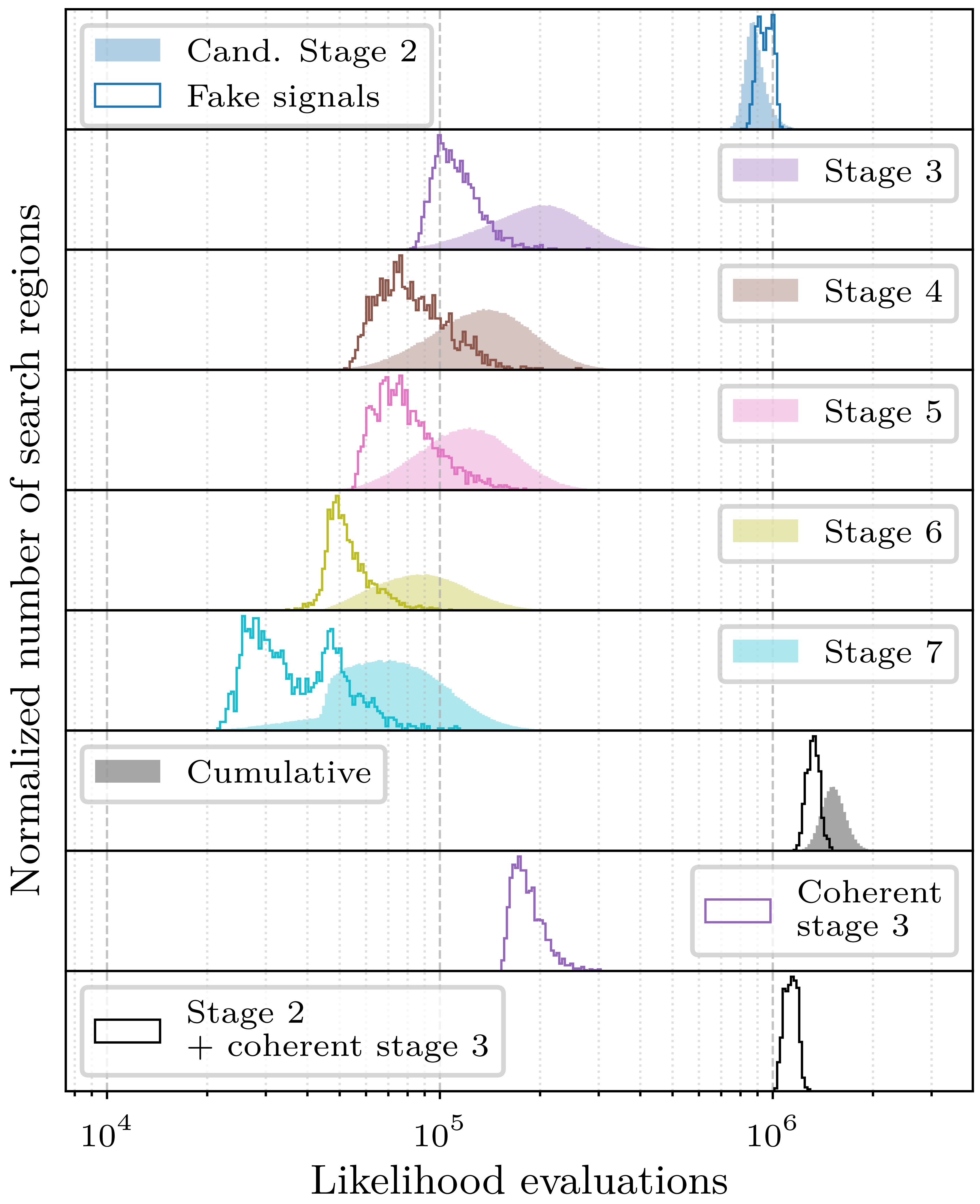


Results — Computational Cost

- HierarchMC is only useful if it is fast
- Stage 2 req. $\approx 10^6$ likelihood evaluations
  - 10^5Cand., 10^{-2}s per likelihood evalution $\Rightarrow 10^9$s Runtime
  - Atlas: 10^4 CPU cores $\Rightarrow 10^5$s Cluster-time
  - Cluster time $\mathcal{O}$(days)
- What about subsequent stages?
  - Do not sig. increase runtime!

A fully coherent follow-up of **all** candidates is possible

- Original Follow-up: Stage 2 and beyond took $\mathcal{O}$(weeks) in human and cluster time

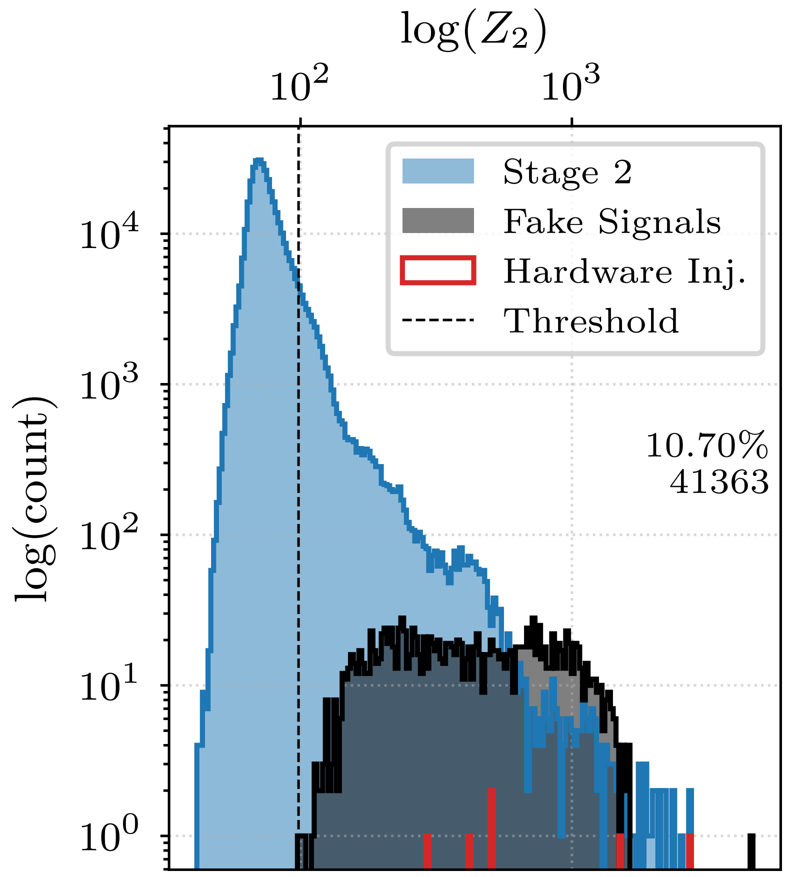


Results - Follow-up of 386,429 Candidates

- Found viable configuration at the second stage

  ➡ Apply Hierarch MC to candidates from that stage

- Depth-first approach: All stages run back-to-back, results are collected at the end

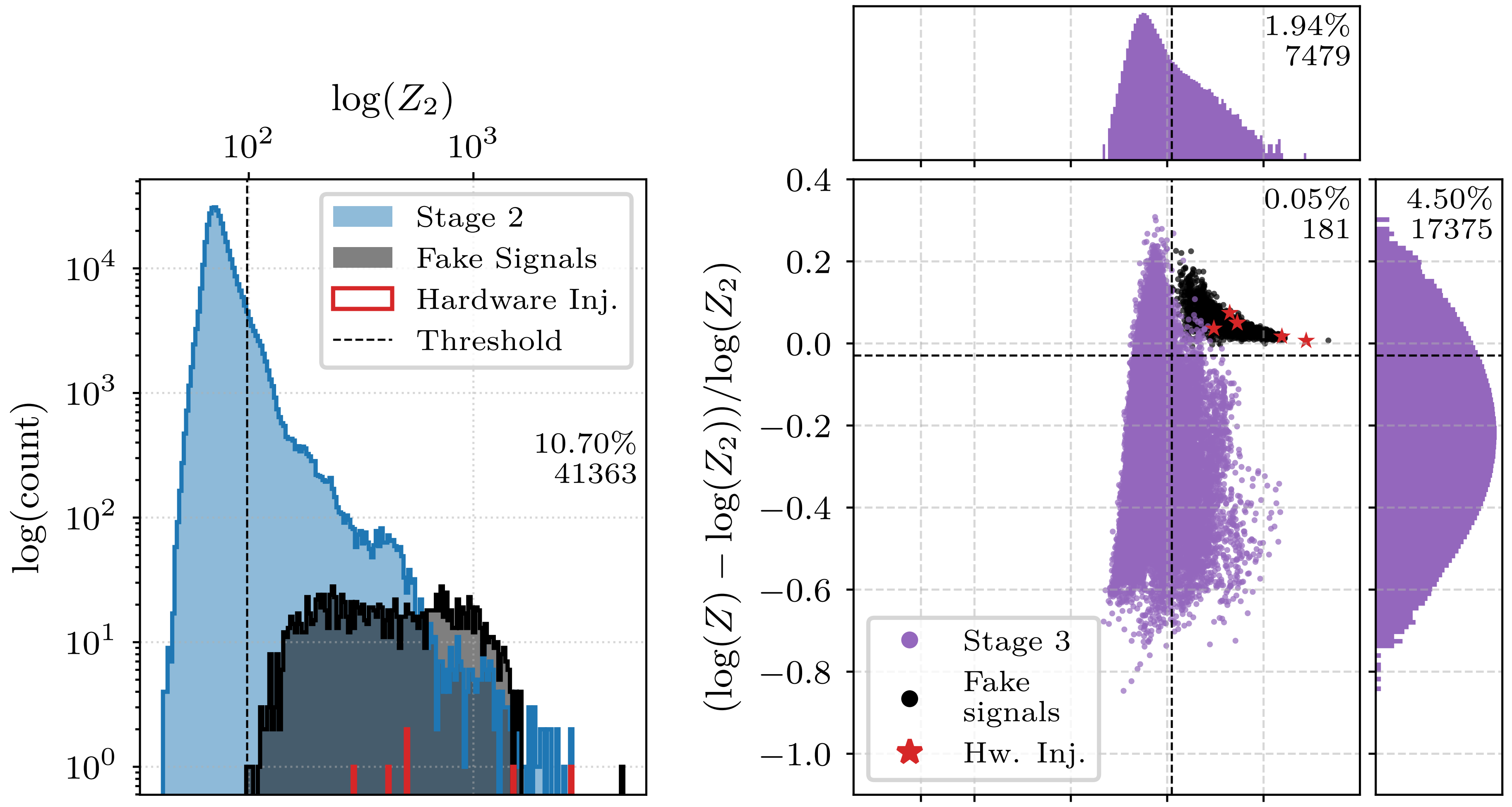


Results - Follow-up of 386,429 Candidates

- Found viable configuration at the second stage

  ➡ Apply Hierarch MC to candidates from that stage

- Depth-first approach: All stages run back-to-back, results are collected at the end

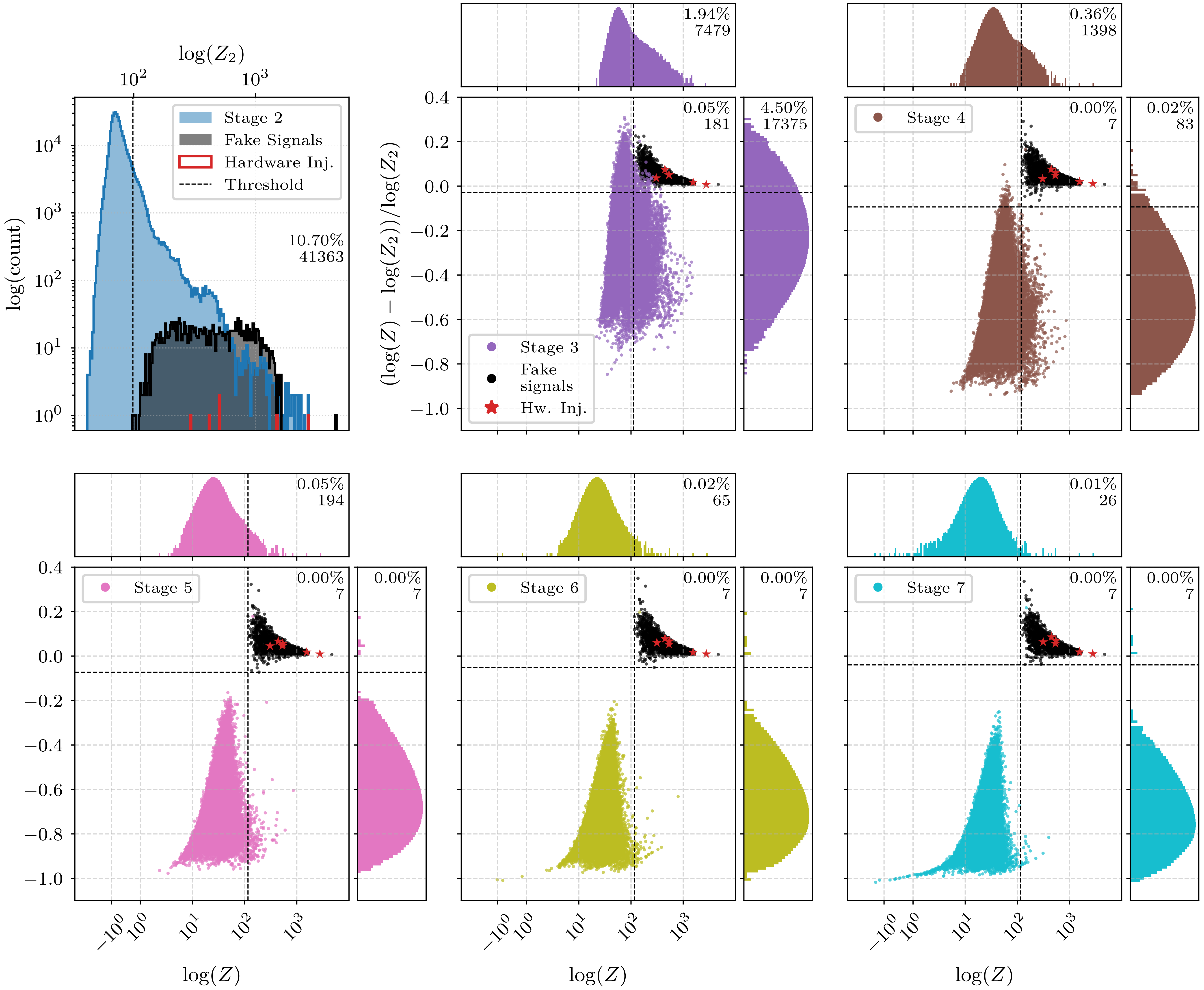


Results - Follow-up of 386,429 Candidates

- Found viable configuration at the second stage

  ➡ Apply Hierarch MC to candidates from that stage

- Depth-first approach: All stages run back-to-back, results are collected at the end



Results - Follow-up of 386,429 Candidates

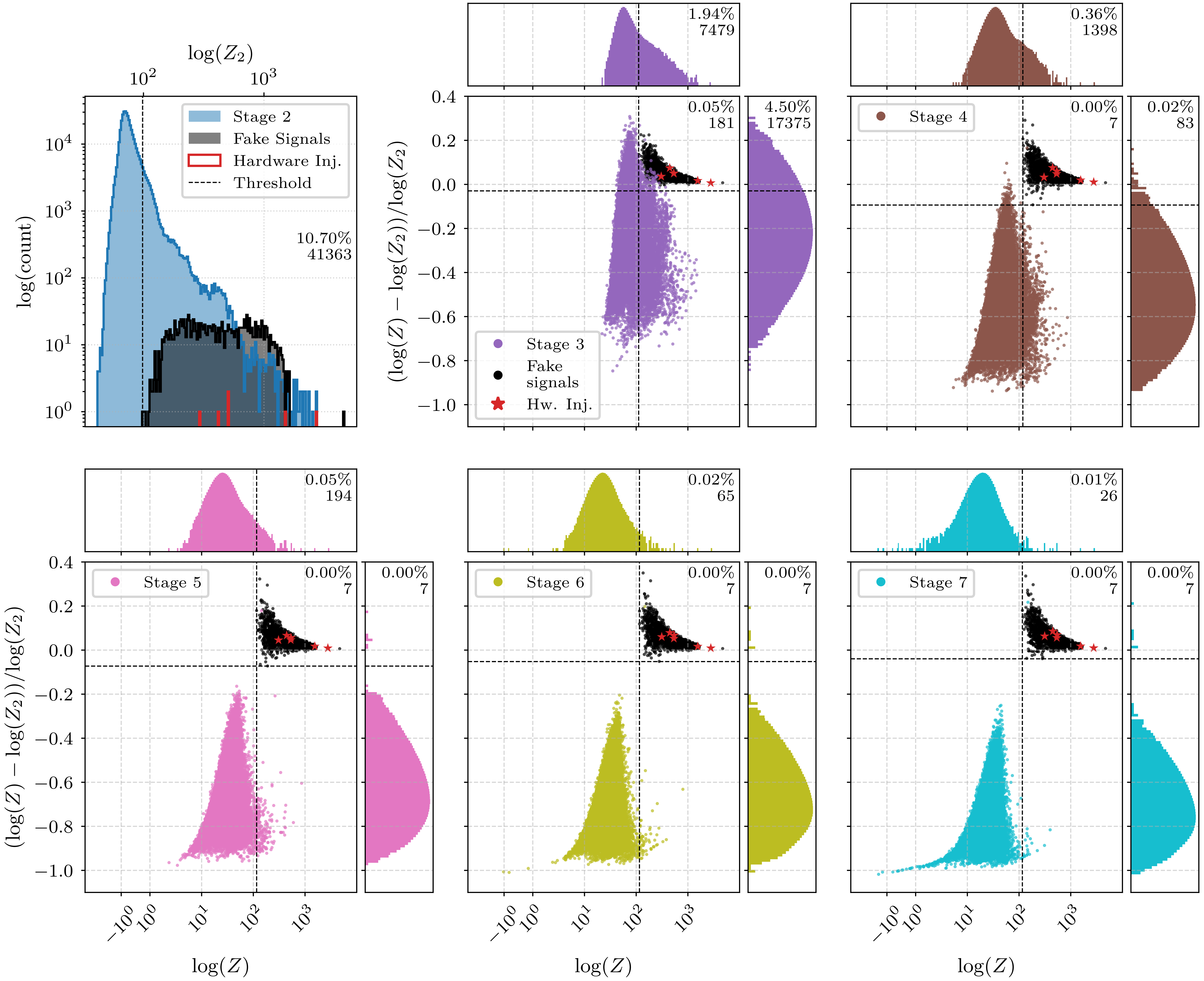
- Found viable configuration at the second stage

  ➡ Apply Hierarch MC to candidates from that stage

- Depth-first approach: All stages run back-to-back, results are collected at the end

- Recover all six hardware injections in band

- One remaining candidate



Results - Follow-up of 386,429 Candidates

Wrong RefTime	Late to the Meeting	Sampler converges on side peaks	"watch condor_release"
Having remaining candidates	That one cluster job that keeps running forever....	Crashing a headnode	"...of course it would be nice if these plots had meaningful titles"
Forgetting the mystery factor	Finding a CW	Thinking you found a CW....but you really haven't	All jobs going on hold immediately
Really understanding GCT	Wrong RefTime... again	Installing LALSuite	Forgetting the mystery factor...again
Cluster Jobs not starting due to missing log directory	log10B...tS? GL...t...L!?	"These hardware injections are quite loud"	Wildcard

The CW Bingo

Results - Follow-up of 386,429 Candidates

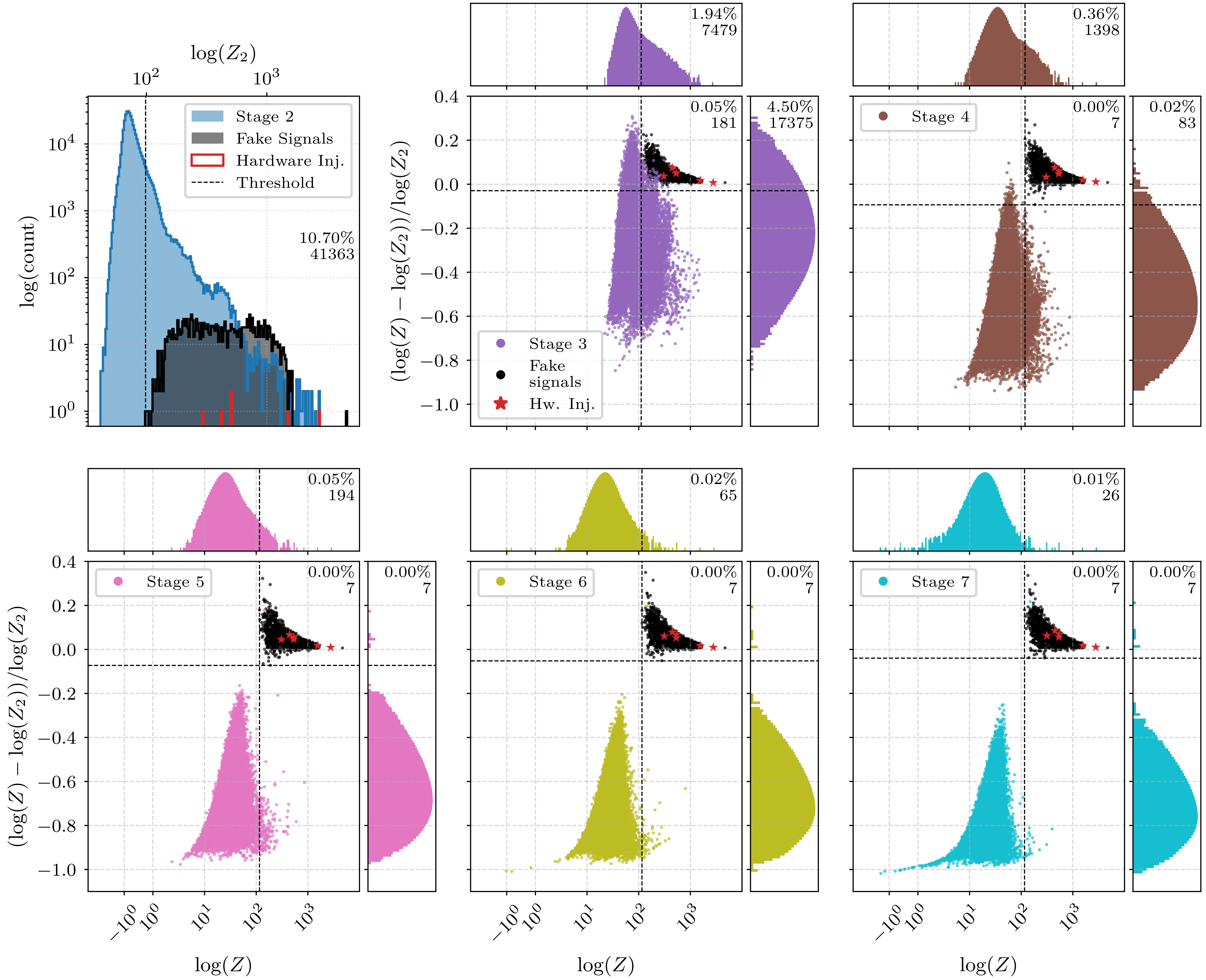
- Found viable configuration at the second stage

  ➡ Apply Hierarch MC to candidates from that stage

- Depth-first approach: All stages run back-to-back, results are collected at the end

- Recover all six hardware injections in band

- One remaining candidate ⇒ Uncleaned detector line in both



Results - Follow-up of 386,429 Candidates

Wrong RefTime	Late to the Meeting	Sampler converges on side peaks	"watch condor_release"
Having remaining candidates	That one cluster job that keeps running forever....	Crashing a headnode	"...of course it would be nice if these plots had meaningful titles"
Forgetting the mystery factor	Finding a CW	Thinking you found a CW....but you really haven't	All jobs going on hold immediately
Really understanding GCT	Wrong RefTime... again	Installing LALSuite	Forgetting the mystery factor...again
Cluster Jobs not starting due to missing log directory	log10B...tS? GL...t...L!?	"These hardware injections are quite loud"	Wildcard

The CW Bingo

Summary

- HierarchMC is a novel, Bayesian Framework for **large-scale**, **depth-first** Monte Carlo follow-ups
- At it's core, HierarchMC implements a posterior-to-prior conversion
- With the examined configuration, HierarchMC was applicable to per-candidate SRs as large as $10^8 N_\star$
- We followed up 386,429 candidates from the reference search and get:
  - Significantly reduced computational and human cost
  - Coherent results on all candidates
  - Fewer false alarms

Future Plans

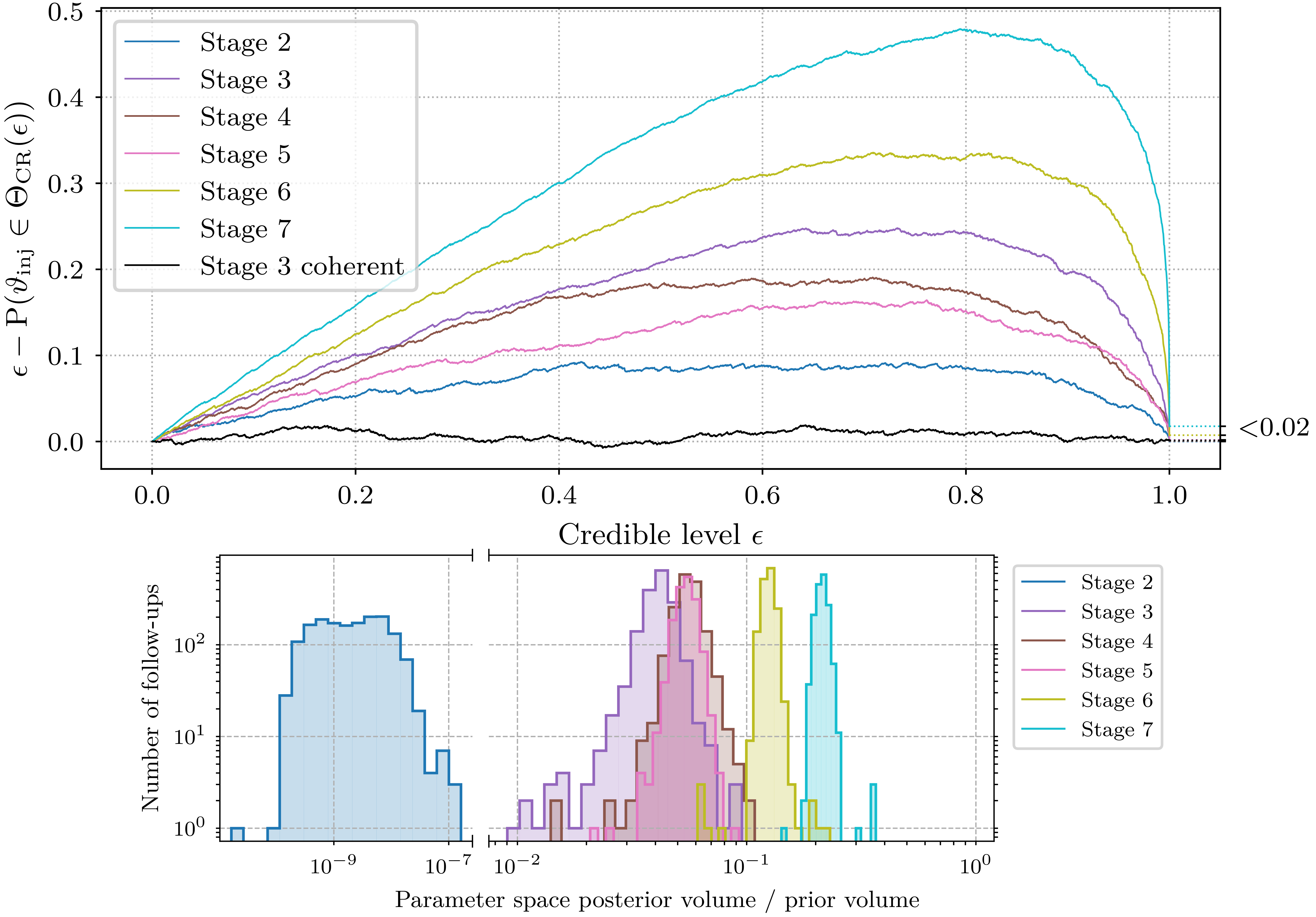
- Write the Paper!
- Application to earlier stages of all-sky searches ⇔ Push towards **higher** $N_\star$
  - Other sampling algorithms / proposal methods
  - Exploit caching optimizations from Grid-based searches
- Application to other searches
  - Binary systems
  - Directed Searches

Future Plans

- Write the Paper!
- Application to earlier stages of all-sky searches ⇔ Push towards **higher** $N_\star$
  - Other sampling algorithms / proposal methods
  - Exploit caching optimizations from Grid-based searches
- Application to other searches
  - Binary systems
  - Directed Searches

Dream: Run HierarchMC directly on Einstein@Home

Backup Slides: Parameter Estimation



Backup Slides: PyFstat

- Based on ptemcee
  - Harder to assess convergence in the **absence** of signals: When did we sample enough?
  - May have trouble converging on multimodal posteriors
- Uses posterior samples as starting points at higher T_{coh}, but the **same** priors

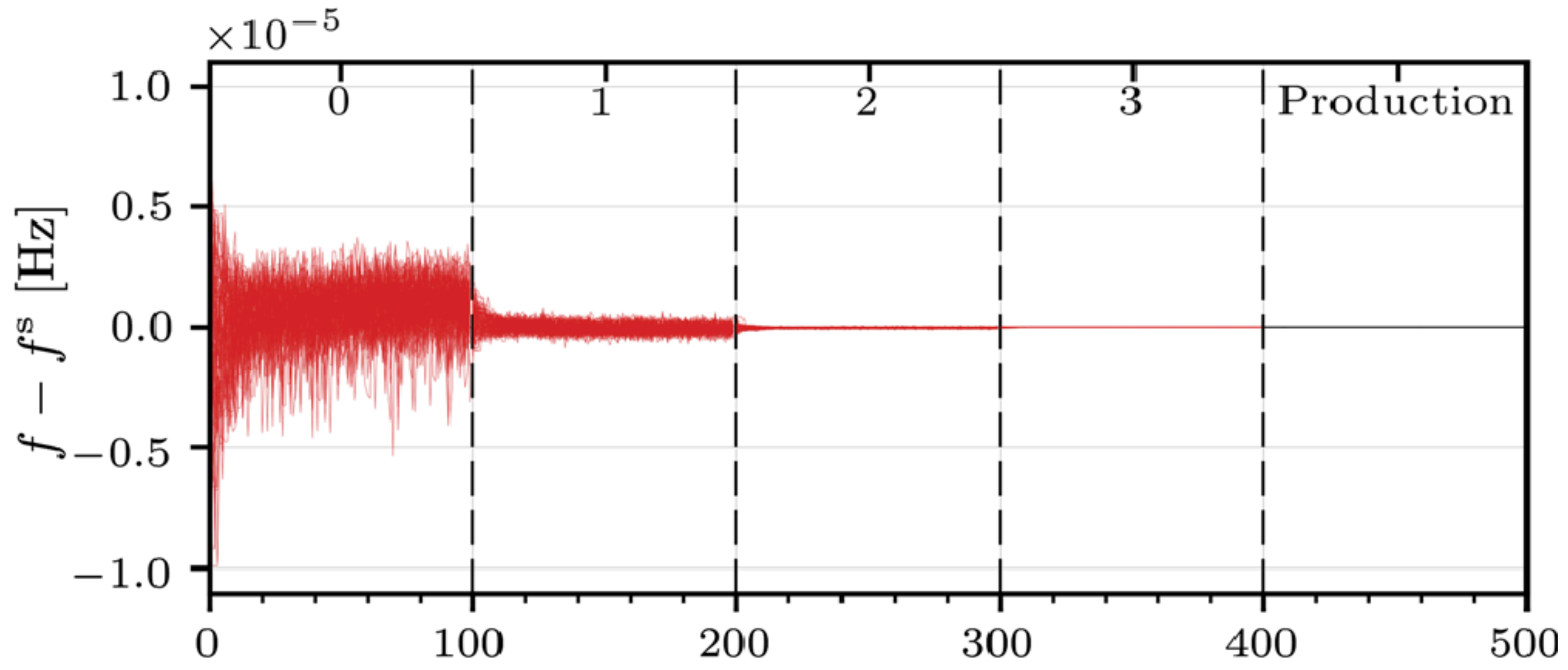
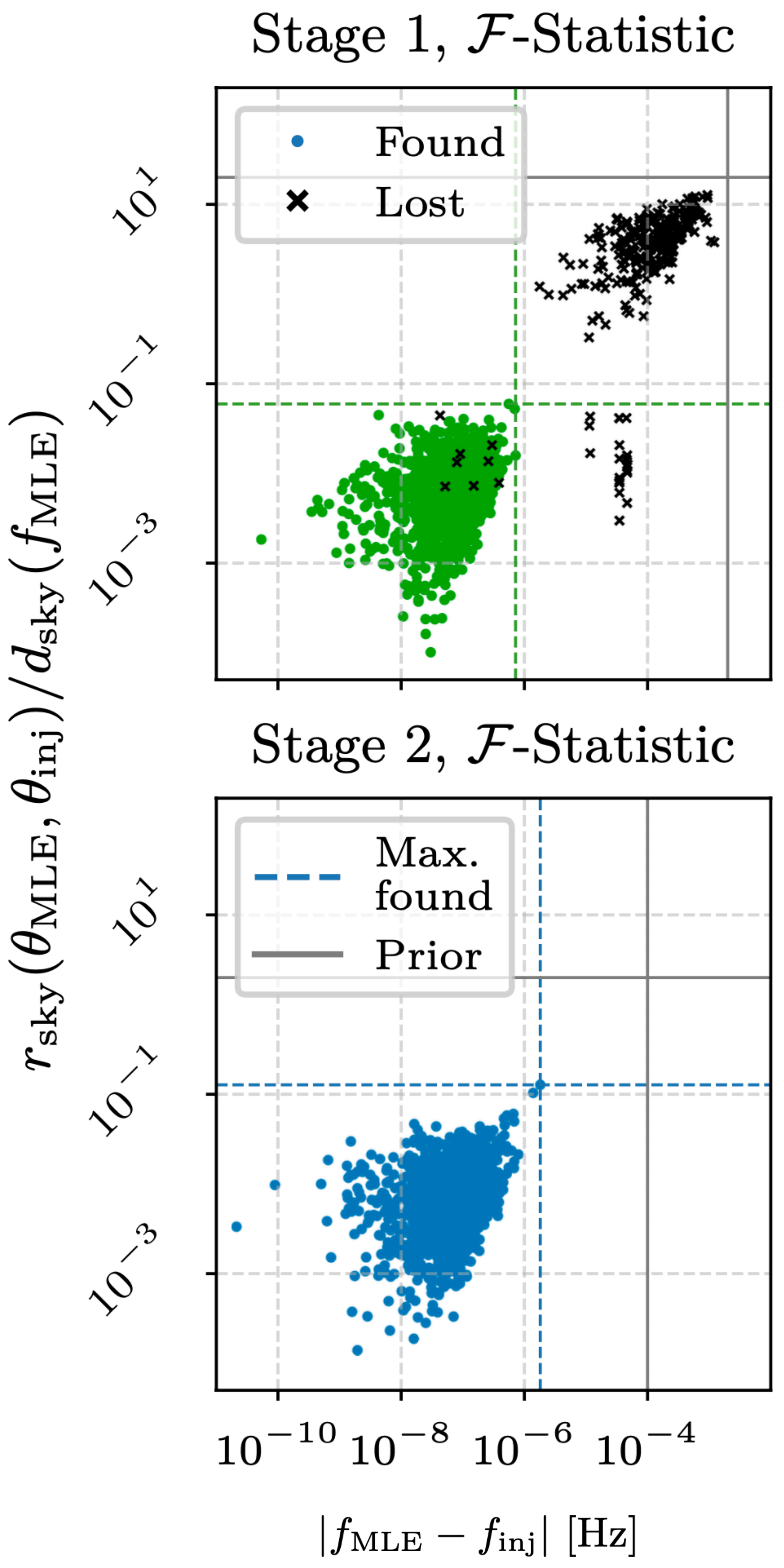


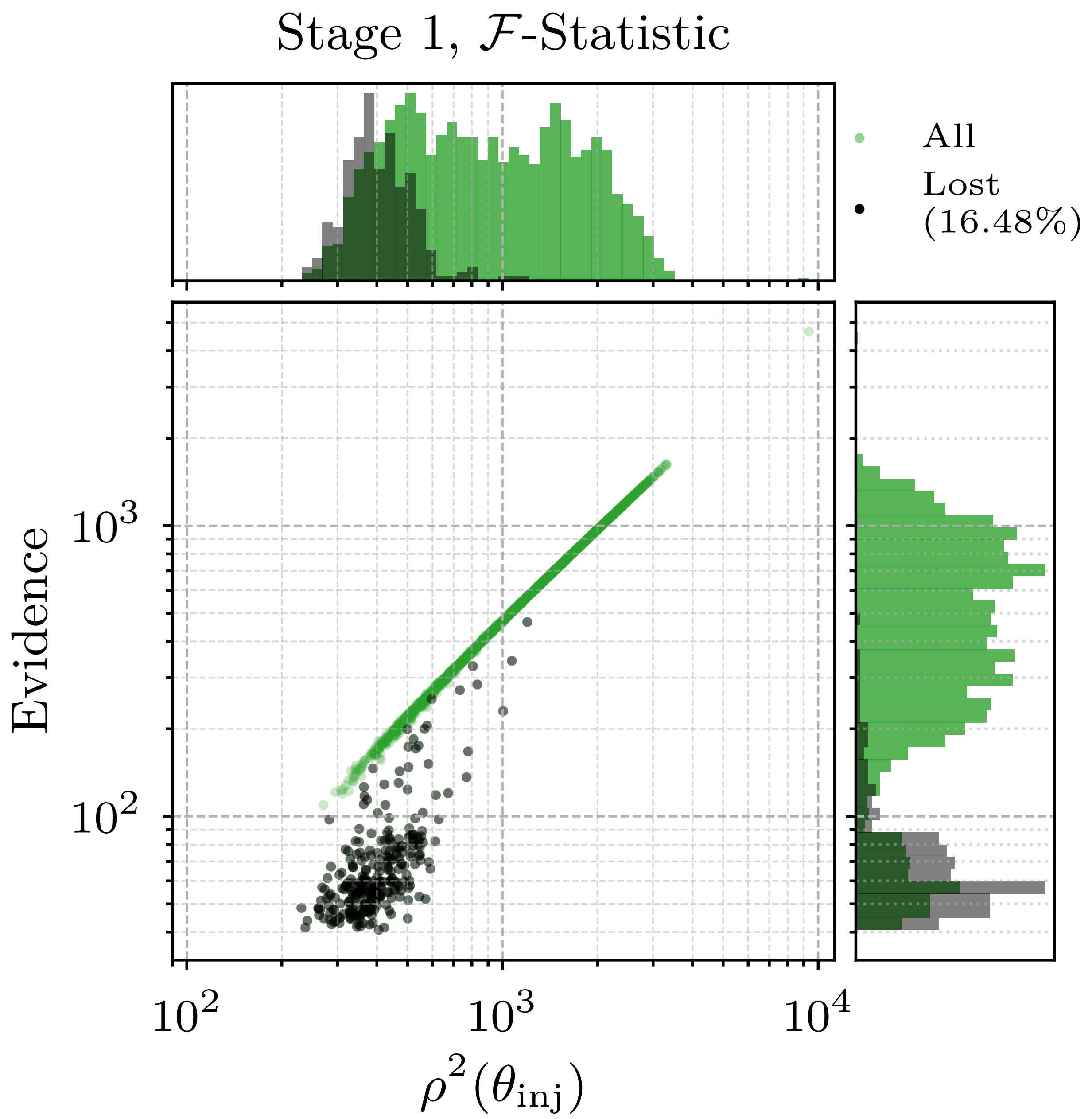
Figure 7 in Ashton and Prix (2018)

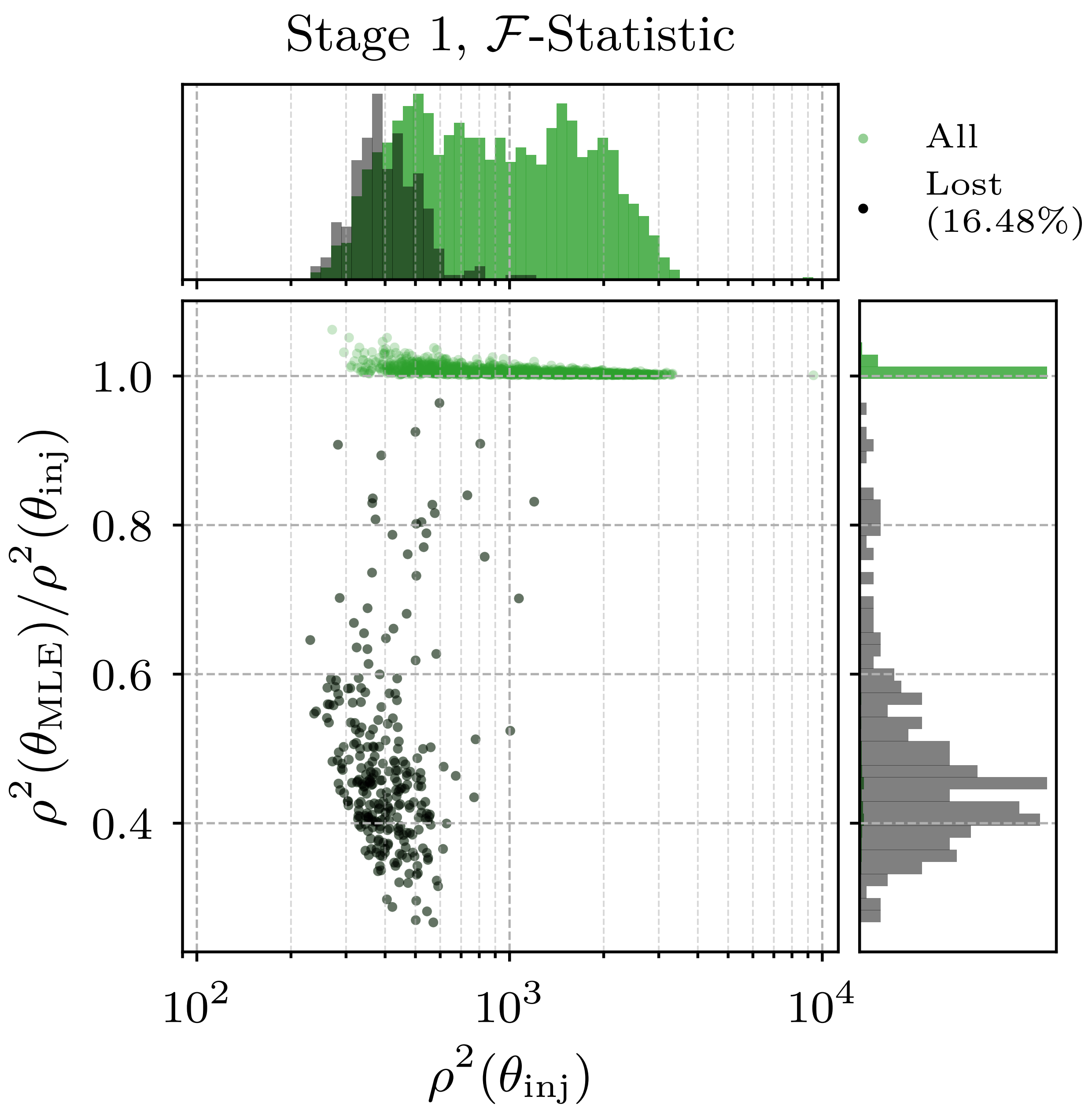
Backup Slides: Hardware Injections

$\mathrm{ID}_{\mathrm{inj}}$	f_{inj} [Hz]	$f^{(1)}_{\mathrm{inj}}$ [Hz s^{-1}]	α_{inj} [h:m:s]	δ_{inj} [deg:m:s]
Dist. to MLEs:	Δf [Hz]	$\Delta \dot{f}$ [Hz s^{-1}]		Sky distance [deg:m:s]
Steltner et al. [81]	Δf [Hz]	$\Delta \dot{f}$ [Hz s^{-1}]		Sky distance [deg:m:s]
0	265.57505348	-4.15×10^{-12}	4:46:12.4628	−56:13:02.9490
	-4.6×10^{-09}	9.0×10^{-15}		0:00:00.4012
	-4.7×10^{-11}	9.5×10^{-16}		0:00:00.4011
2	575.16350527	-1.37×10^{-13}	14:21:01.4800	3:26:38.3626
	3.4×10^{-09}	4.3×10^{-15}		0:00:00.2198
	-1.1×10^{-09}	-8.8×10^{-16}		0:00:00.0955
3	108.85715939	-1.46×10^{-17}	11:53:29.4178	−33:26:11.7687
	-4.6×10^{-09}	2.4×10^{-14}		0:00:02.9579
	-6.7×10^{-10}	-5.8×10^{-16}		0:00:00.3080
5	52.80832436	-4.03×10^{-18}	20:10:30.3939	−83:50:20.9036
	1.9×10^{-09}	6.0×10^{-15}		0:00:01.6488
	-6.2×10^{-10}	-4.2×10^{-16}		0:00:00.2212
9	763.84731649	-1.45×10^{-17}	13:15:32.5397	75:41:22.5205
	-6.0×10^{-09}	-9.48×10^{-15}		0:00:00.1649
	9.4×10^{-10}	-5.60×10^{-17}		0:00:00.0023
10	26.33209638	-8.50×10^{-11}	14:46:13.3549	42:52:38.2953
	-6.7×10^{-09}	8.3×10^{-15}		0:00:03.9060
	-8.3×10^{-11}	2.4×10^{-16}		0:00:00.3109

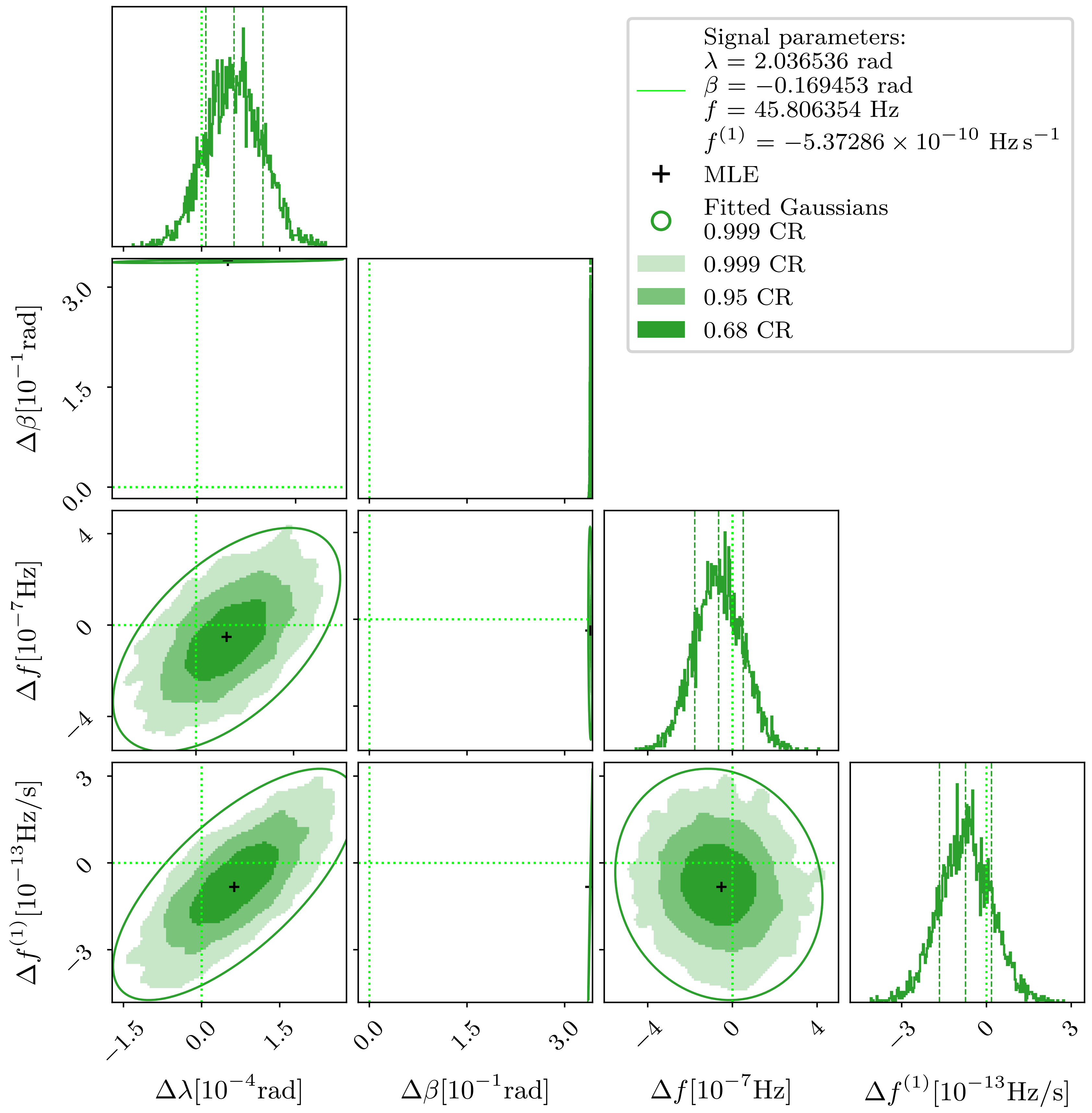
Backup Slides: Failed Recovery

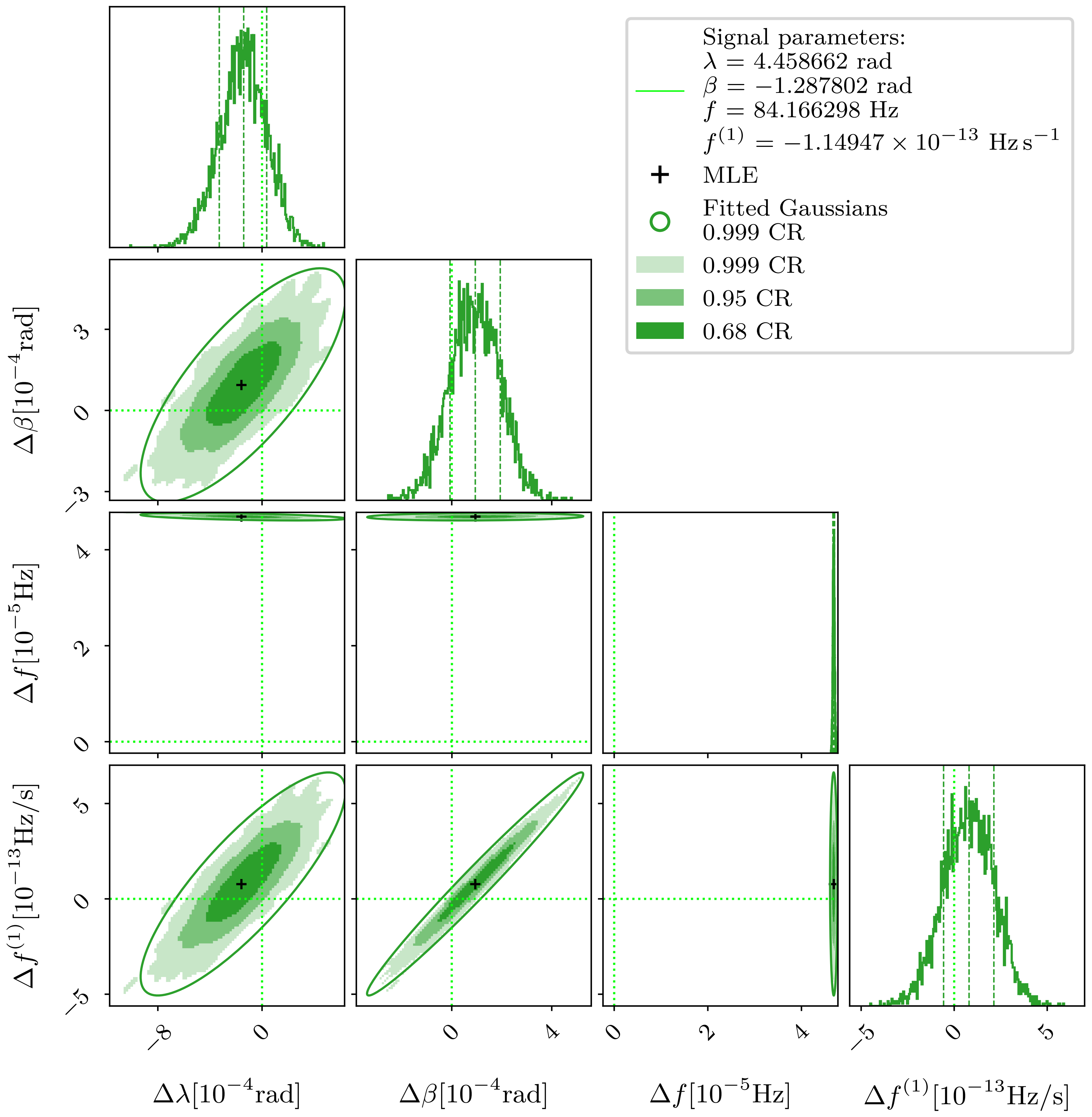




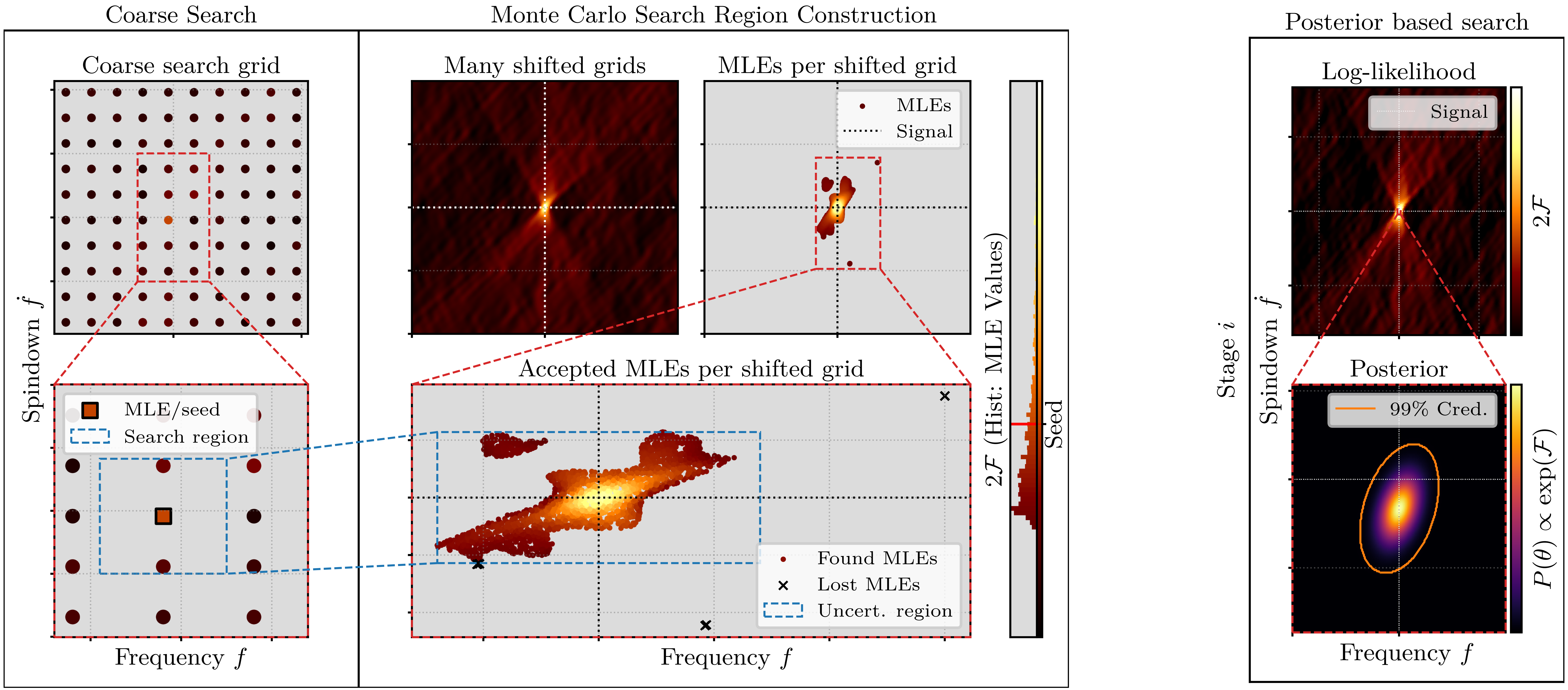


Backup Slides: Side Peaks





Backup Slides: Grids vs. Posteriors



Backup Slides: Sampler Configuration

Option	Setting
n_{live}	1500
$\Delta \log Z$	0.1
Bounding	Multi-ellipsoids
Sampler	Random walk
Walks per iter.	25
Update interval	$2n_{\text{live}}$
Min. eff.	5%
Bootstrap	25
Enlarge	1

Backup Slides: The $\mathscr{F}$-Statistic Metric

- Second-order Taylor expansion of mismatch μ:

$$\mu(\vartheta_S, \vartheta_S + \Delta\vartheta) := 1 - \frac{\rho^2(\vartheta_S + \Delta\vartheta)}{\rho^2(\vartheta_S)} \approx g_{ij}\Delta\vartheta_i\Delta\vartheta_j$$

- Defines natural unit for search region volume

$$\mathrm{Vol}(\mathcal{R})[N_\star] := \sqrt{\det g_{ij}}\,\mathrm{Vol}(\mathcal{R})$$

- Approximately flat in frequency, spindown and sky position when projected to ecliptic plane

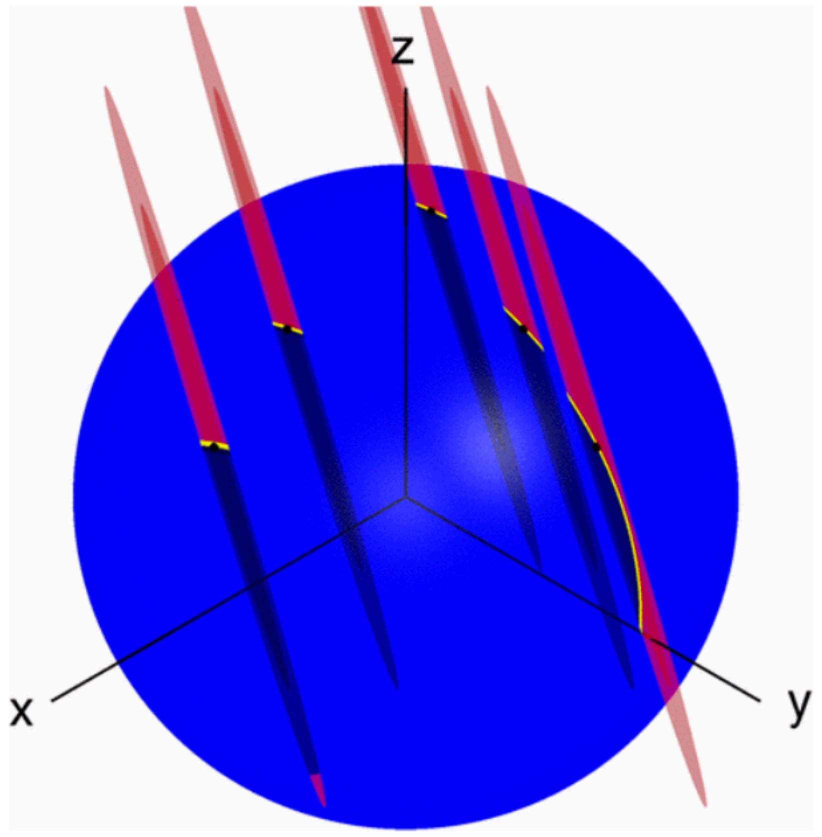


Figure 4 in Wette and Prix (2013)

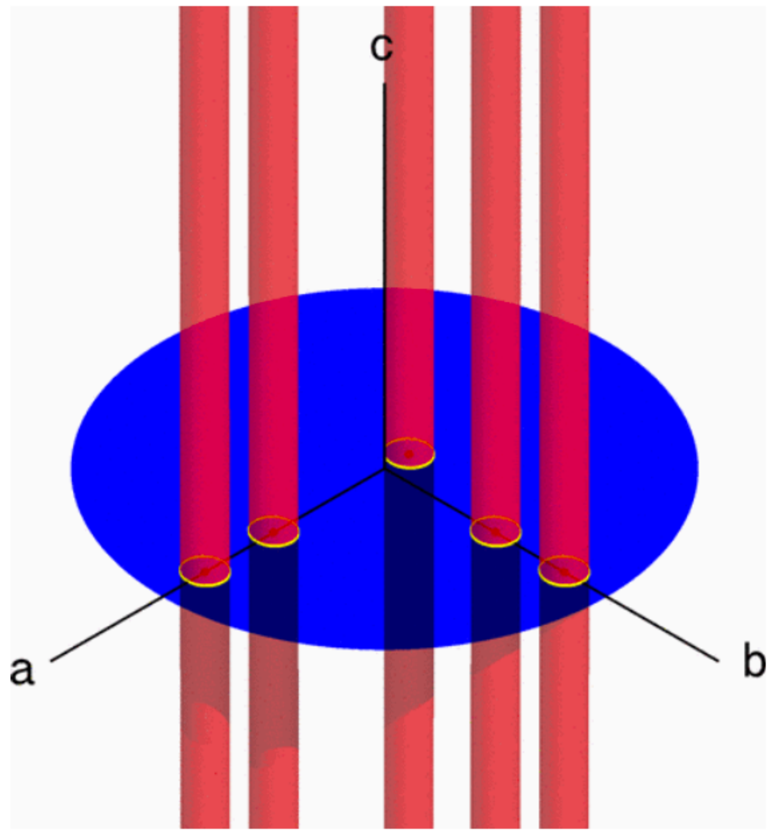


Figure 10 in Wette and Prix (2013)