Updating EMRIs detection rates and PE forecast

Manuel PIARULLI,
Danny LAGHI, Ollie BURKE, Shubam KEJRIWAL,
Christian CHAPMAN-BIRD, Federico POZZOLI, Nikos KARNESIS

EMRI Search and Inference within the LISA Global Fit - II, Singapore

EMRI populations

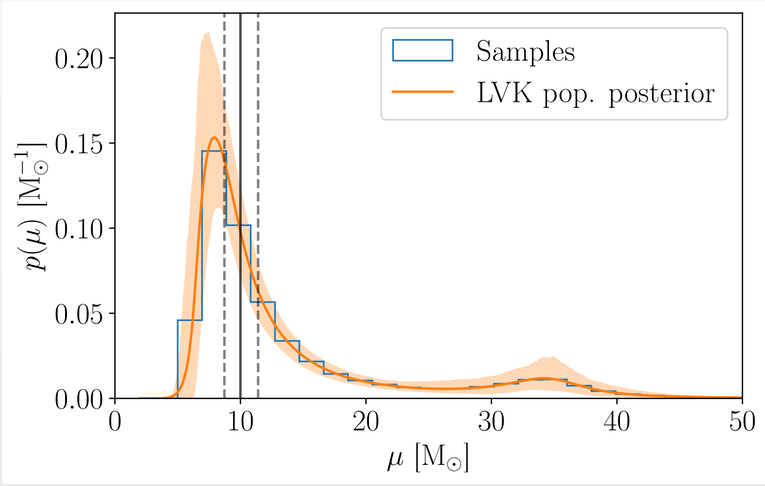
EMRI event rates are uncertain
expected between ~ 10 and ~ 20000 per year

Model	MBH Mass	MBH spin	Cusp Erosion	$M-\sigma$	N_p	Rate [yr^{-1}]
M1	Barausse12	a98	yes	Gultekin09	10	1600
M5	Gair10	a98	no	Gultekin09	10	140
M6	Barausse12	a98	no	Gultekin09	10	2080
M7	Barausse12	a98	yes	Gultekin09	0	15800
M11	Gair10	a0	no	Gultekin09	100	13

Babak et al. 2017 [arxiv:1703.09722]

we allow the mass of the small CO to vary:

`LVK - Power Law + Peak`



LVK Population analysis with GWTC-3 (2022)
[arxiv:2111.03634]

Each catalog contains EMRI populations across the Universe over 10 years

We focus on 5 EMRI catalogs
M1-M6 - M5-M11 - M7
intermediate - pessimistic - optimistic

What's the full process?

$$T_{obs} = 4\texttt{yrs},\ \texttt{dt} = 5\texttt{s},\ \ \texttt{2ndGenTDI}$$

	Computational Cost
1. **Evolve source parameters from plunge to initial values.** (FEW_backwards) We randomly sampled $t_{back} = t_{plunge} \in [0,\ 10]$yr	O(minutes)
2. **Compute signals and SNRs for each source and sum them together.** (FEW, lisa-on-gpu) $\sum_i \textbf{TDI}_k,\ k = A, E, T \Rightarrow\ S_{\text{EMRI},0}(f)$ is necessary as input for the **IFE algorithm**	O(hours - day)
3. **Iterative Foreground Estimation (IFE)** for more realistic detection rates. (IFE)	O(days - week/s)
4. **Fisher Estimates** on detectable sources (SNR > 20). (SEF)	O(days - 1week)
5. **Bayesian Inference** on a subset of detectable sources. (eryn)	1 day per source

2. Signals and SNRs

To obtain **detection rates** ⇛ we need SNRs ⇛ defining a noise PSD

Static PSD $S_n(f) = S_{\rm instr}(f) + S_{\rm gb}(f)$

$$\rho_i^{\rm iso}, \quad \sum_i {\rm TDI}_k,\ k = A, E, T \quad \Rightarrow \quad S_{\rm EMRI,0}(f) \Rightarrow S_{\rm EMRI,0}(f)$$

computationally efficient but neglects the self-generated EMRI confusion noise

Babak et al. 2017

Model	# before IFE ($T_{\rm obs} = 4$ yr)
M1	461
M5	32
M6	604
M7	4282
M11	1

Model	Rate [1/yr] (detected with AAK)	# after 4 years
M1	294	~1200
M5	47	~200
M6	479	~1900
M7	2712	~11000
M11	11	~45

3. Iterative Foreground Estimation

To obtain **detection rates** ⇛ we need SNRs ⇛ defining a noise PSD
Static PSD $S_n(f) = S_{\rm instr}(f) + S_{\rm gb}(f)$

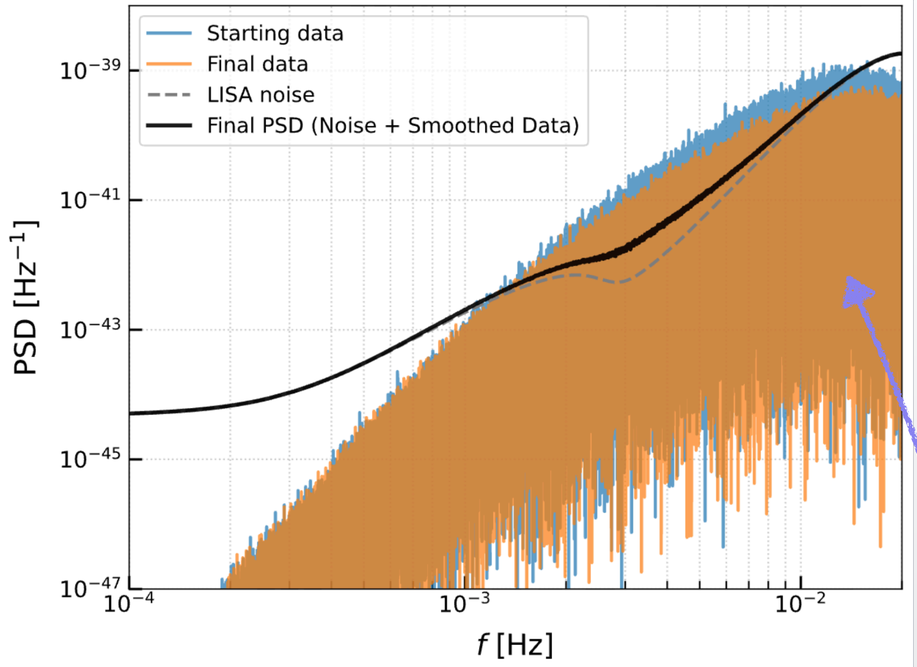
IFE steps:

1. **Initialization:** EMRI PSD at 0th iteration $S_{\rm EMRI,0}(f) = 2 \cdot df\,|\sum_{i=0}^{N} \mathrm{TDI}\,|^2$

2. **Noise Model Update:** At iteration k, the PSD is updated to include the unresolved sources below $\rho_{th} = 20$: $S_{n,k}(f) = S_{\rm instr}(f) + S_{\rm gb}(f) + S_{\rm EMRI,k}(f)$

3. **Classification and Subtraction:** SNRs $\rho_i^{(k)}$ are recalculated against the updated $S_{n,k}(f)$. Resolvable if $\rho_i^{(k)} > \rho_{\rm th}$.

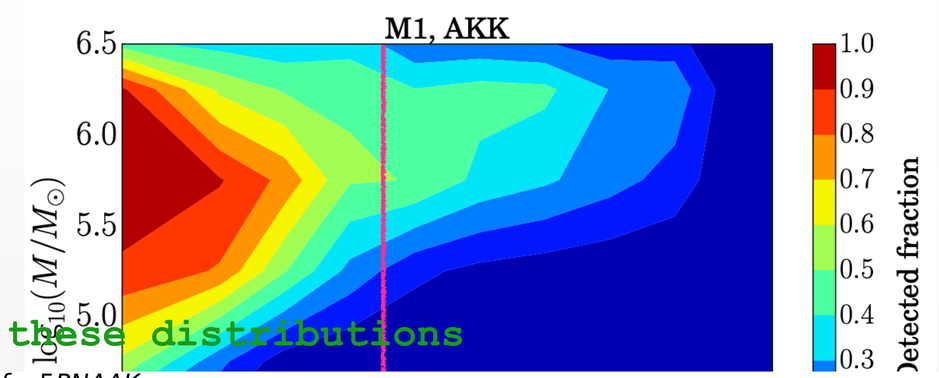
4. **Convergence:** This loop repeats until no additional sources cross the detection threshold (typically requiring 1 to 5 iterations)



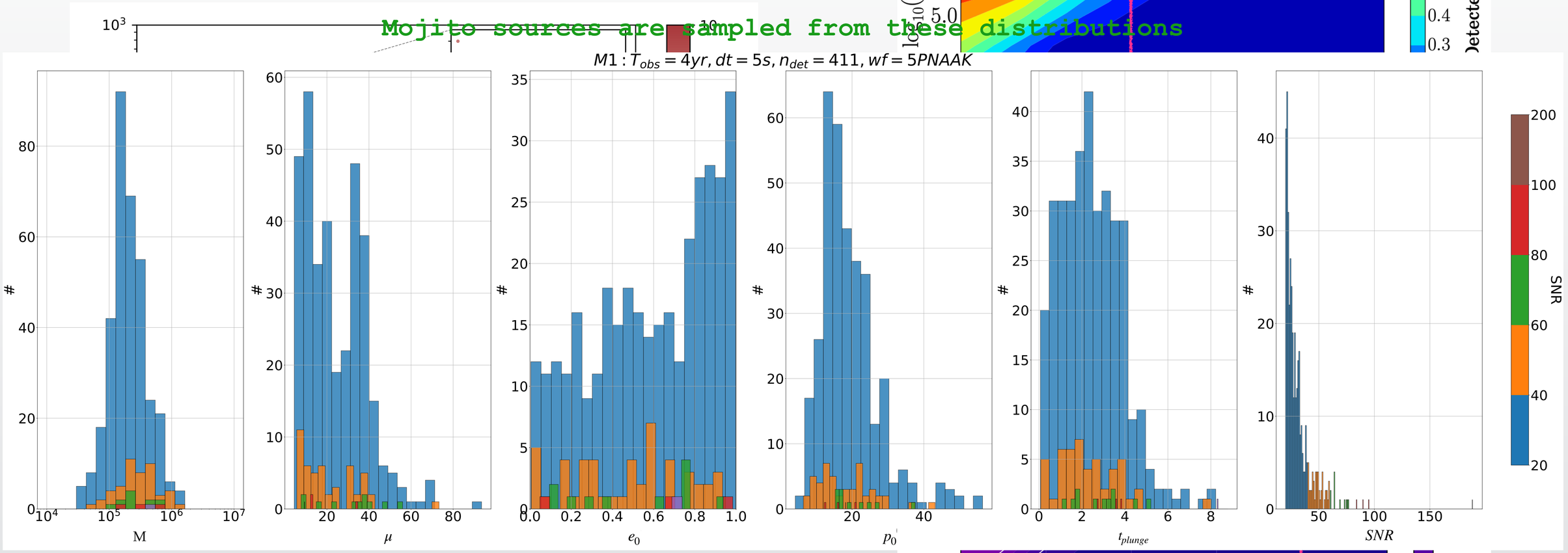
Model	# before IFE ($T_{\rm obs} = 4$ yr)	# after IFE at $T_{\rm obs}$ [yr]				Rate [yr^{-1}]
		1 yr	2 yr	3 yr	4 yr	
M1	461	71	192	310	411	102
M5	32	0	4	19	32	8
M6	604	92	227	390	527	132
M7	4282	413	870	1598	1962	491
M11	1	0	0	1	1	0.25

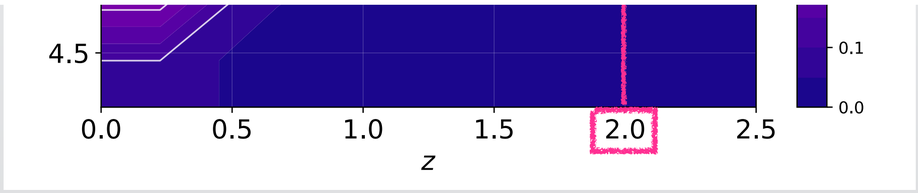
Detected Fraction: "Fiducial" Model (M1)

Detections with Tgw > 4 yrs = 48.
In good support for a longer LISA lifetime.



Mojito sources are sampled from these distributions





Not all nearby sources are detectable.

4. Fisher Estimates

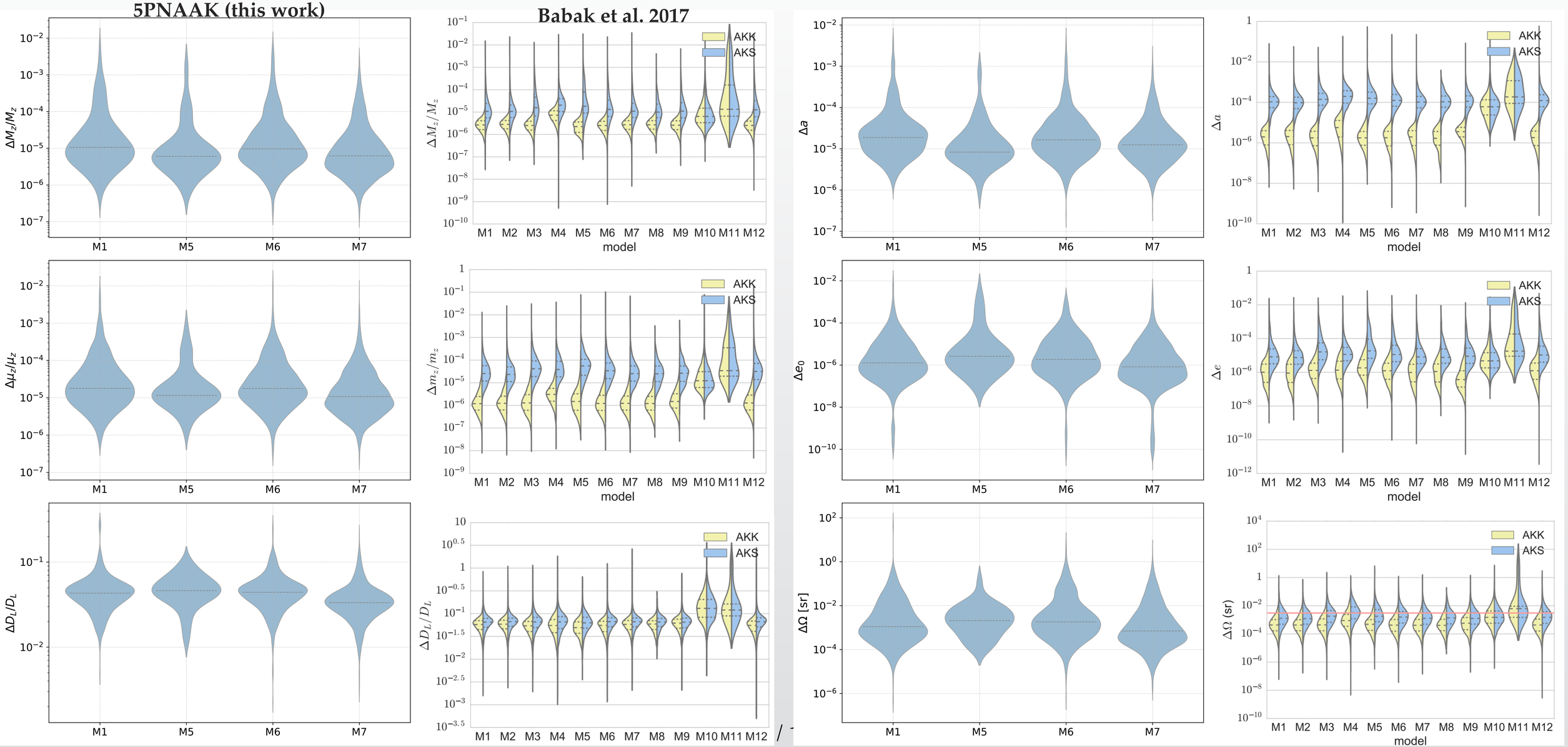
$< \Delta M/M > \sim 10^{-5}$ $< \Delta a > \sim 2 \cdot 10^{-5}$

$< \Delta \mu/\mu > \sim 2 \cdot 10^{-5}$ $< \Delta e_0 > \sim 10^{-6}$

$< \Delta D_L/D_L > \sim 4 \cdot 10^{-2}$ $< \Delta \Omega > \sim 10^{-3}$ $[sr]$

5PNAAK (this work)

Babak et al. 2017

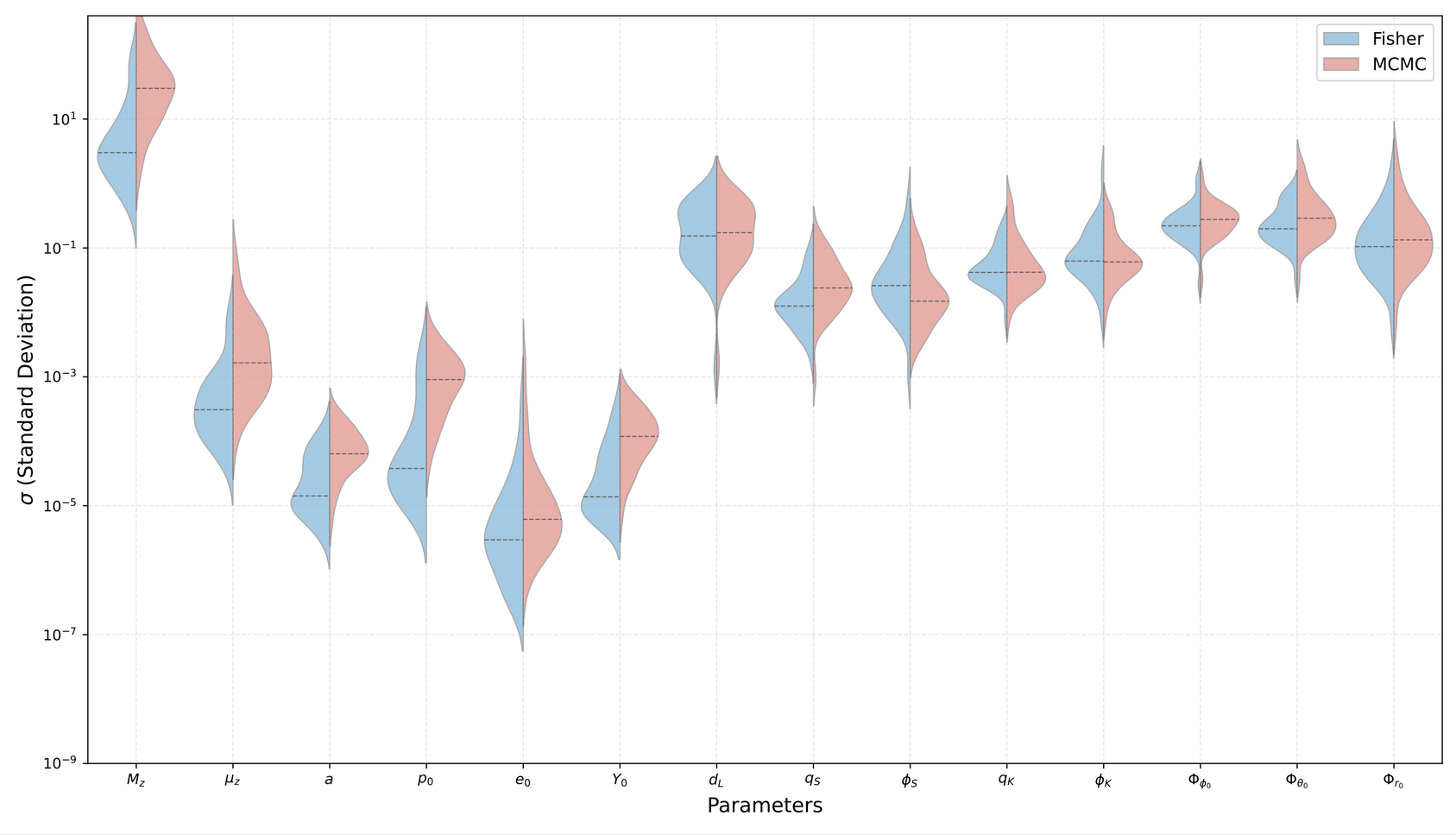


5. Bayesian Inference

Setup:
Sampler: `eryn`
$n_{iter} = 6000$
$n_{walkers} = 48$
$n_{temperature} = 1$
Uniform broad priors, with initialization close to the true values.

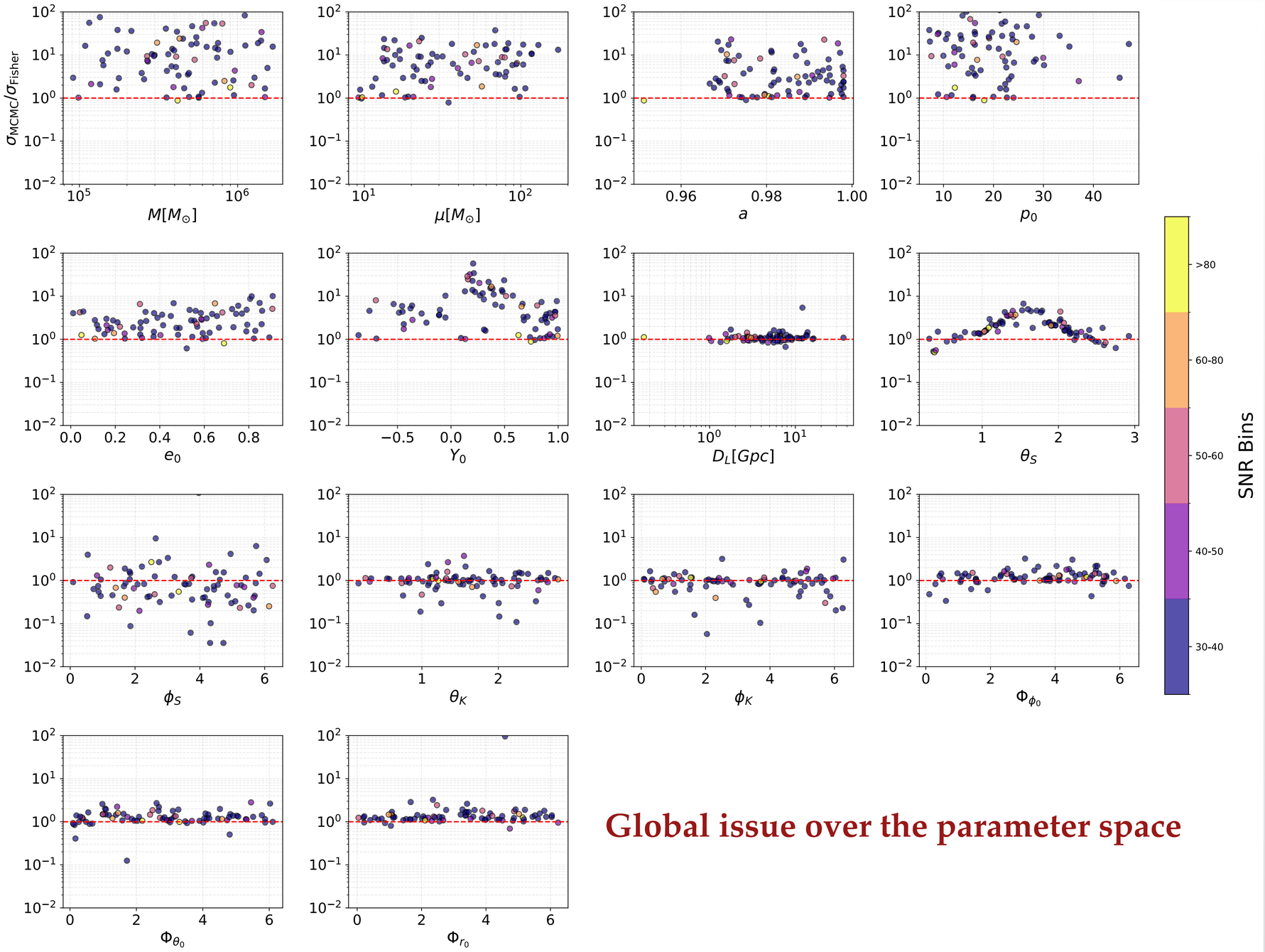
The comparison is made on ~80 sources

On average, Fisher seems to underestimate intrinsic parameters (M, μ, a, p_0, e_0)

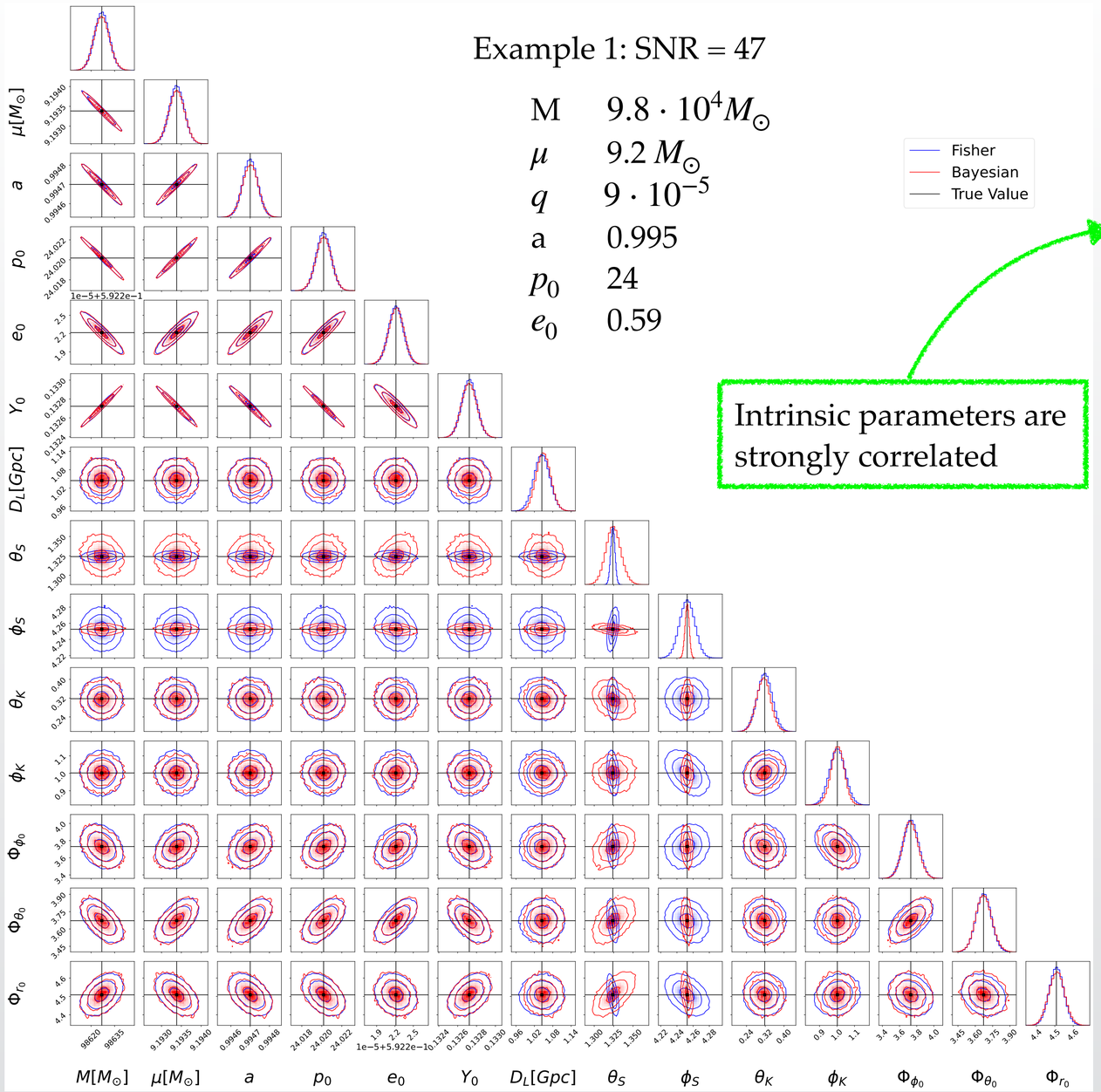


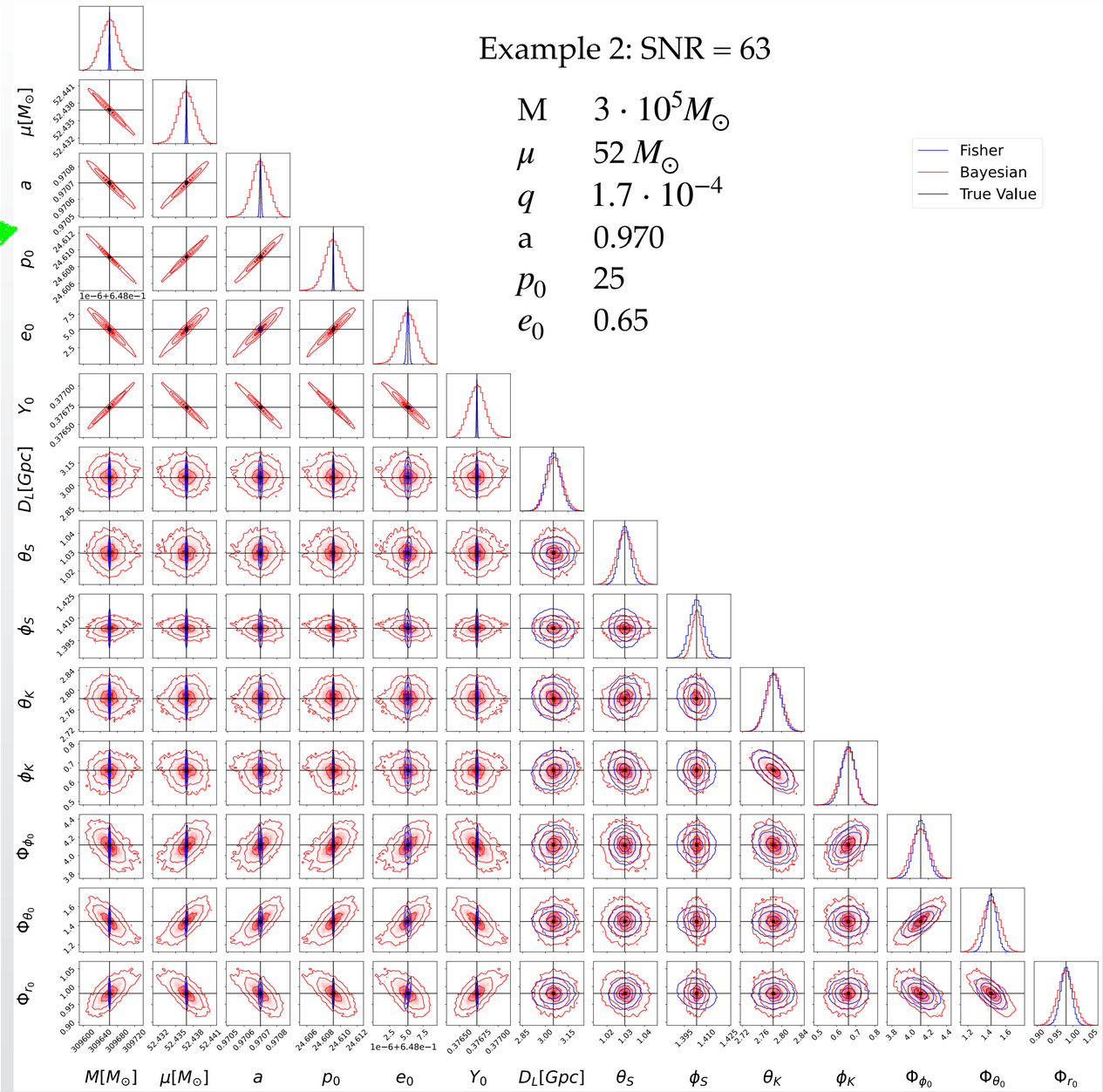
5. Bayesian Inference

1 by 1 comparison $\sigma_{MCMC}/\sigma_{Fisher}$ for each parameter.



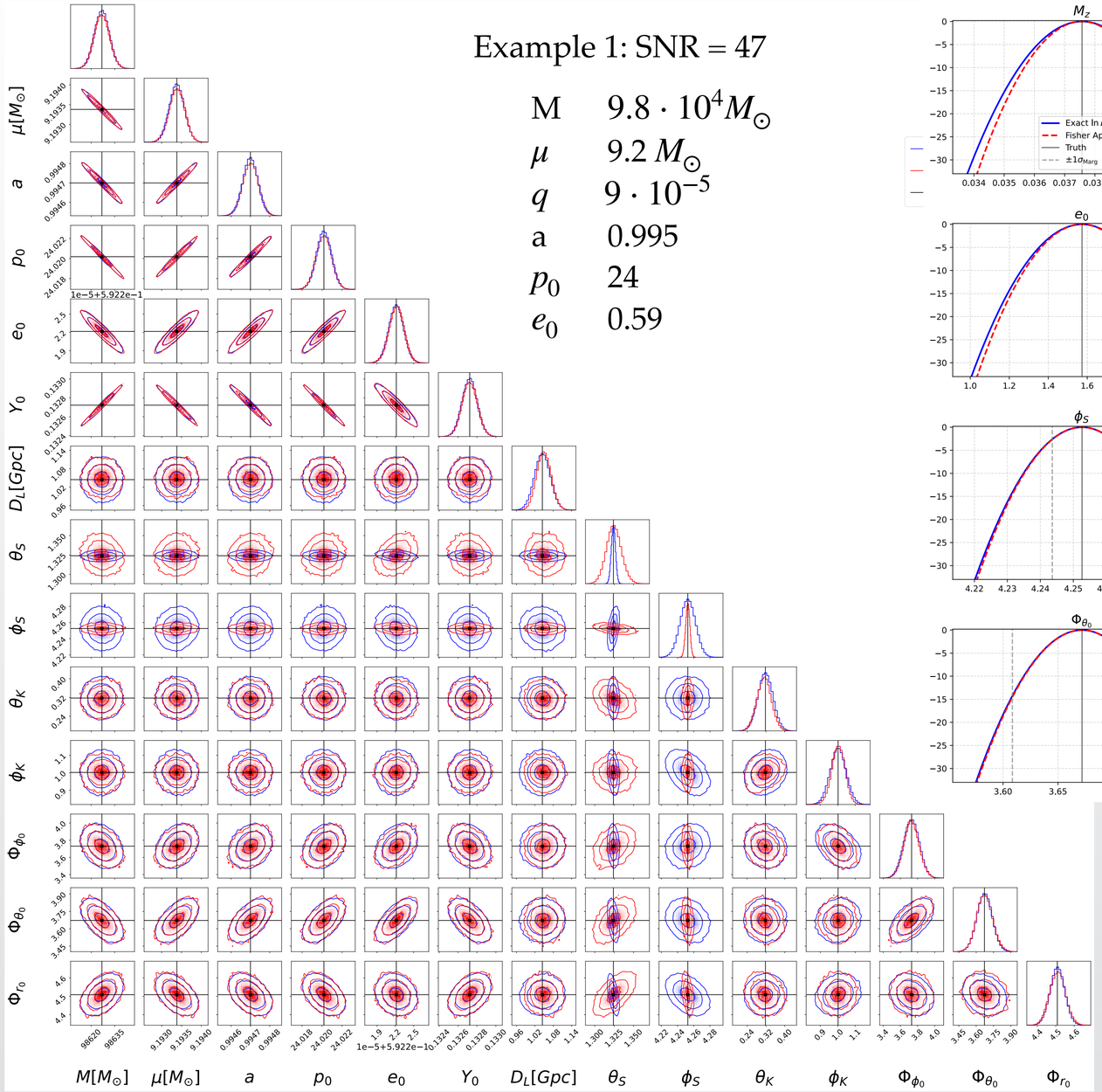
Global issue over the parameter space

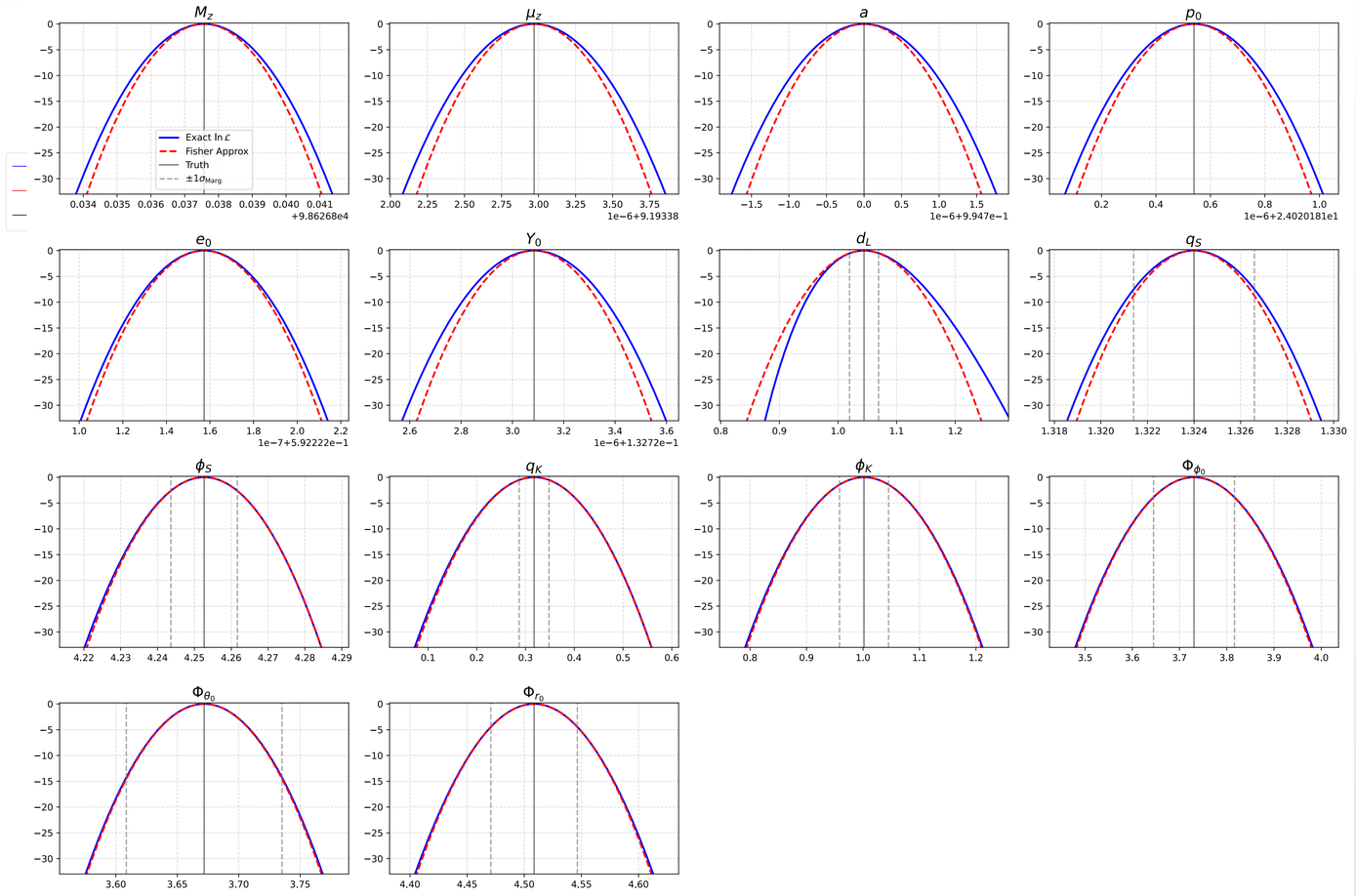

Example 1: SNR = 47
M $9.8 \cdot 10^4 M_\odot$
μ $9.2\, M_\odot$
q $9 \cdot 10^{-5}$
a 0.995
p_0 24
e_0 0.59
Fisher
Bayesian
True Value
Intrinsic parameters are strongly correlated
$M[M_\odot]$ $\mu[M_\odot]$ a p_0 e_0 Y_0 $D_L[Gpc]$ θ_S ϕ_S θ_K ϕ_K Φ_{ϕ_0} Φ_{θ_0} Φ_{r_0}


Example 2: SNR = 63
M $3 \cdot 10^5 M_\odot$
μ $52\, M_\odot$
q $1.7 \cdot 10^{-4}$
a 0.970
p_0 25
e_0 0.65
Fisher
Bayesian
True Value
$M[M_\odot]$ $\mu[M_\odot]$ a p_0 e_0 Y_0 $D_L[Gpc]$ θ_S ϕ_S θ_K ϕ_K Φ_{ϕ_0} Φ_{θ_0} Φ_{r_0}

Example 1: SNR = 47

M	$9.8 \cdot 10^4 M_\odot$
μ	$9.2\, M_\odot$
q	$9 \cdot 10^{-5}$
a	0.995
p_0	24
e_0	0.59

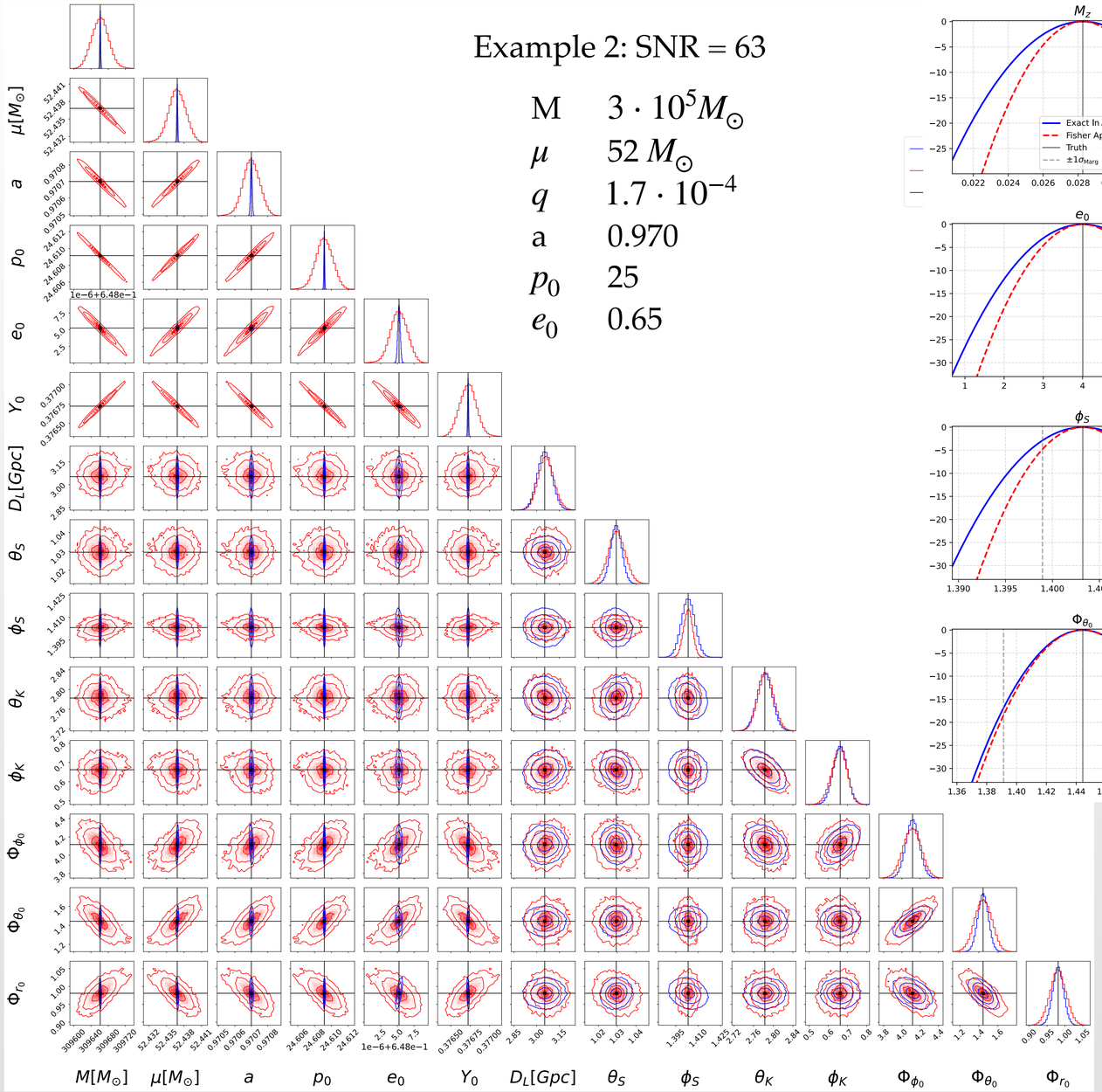


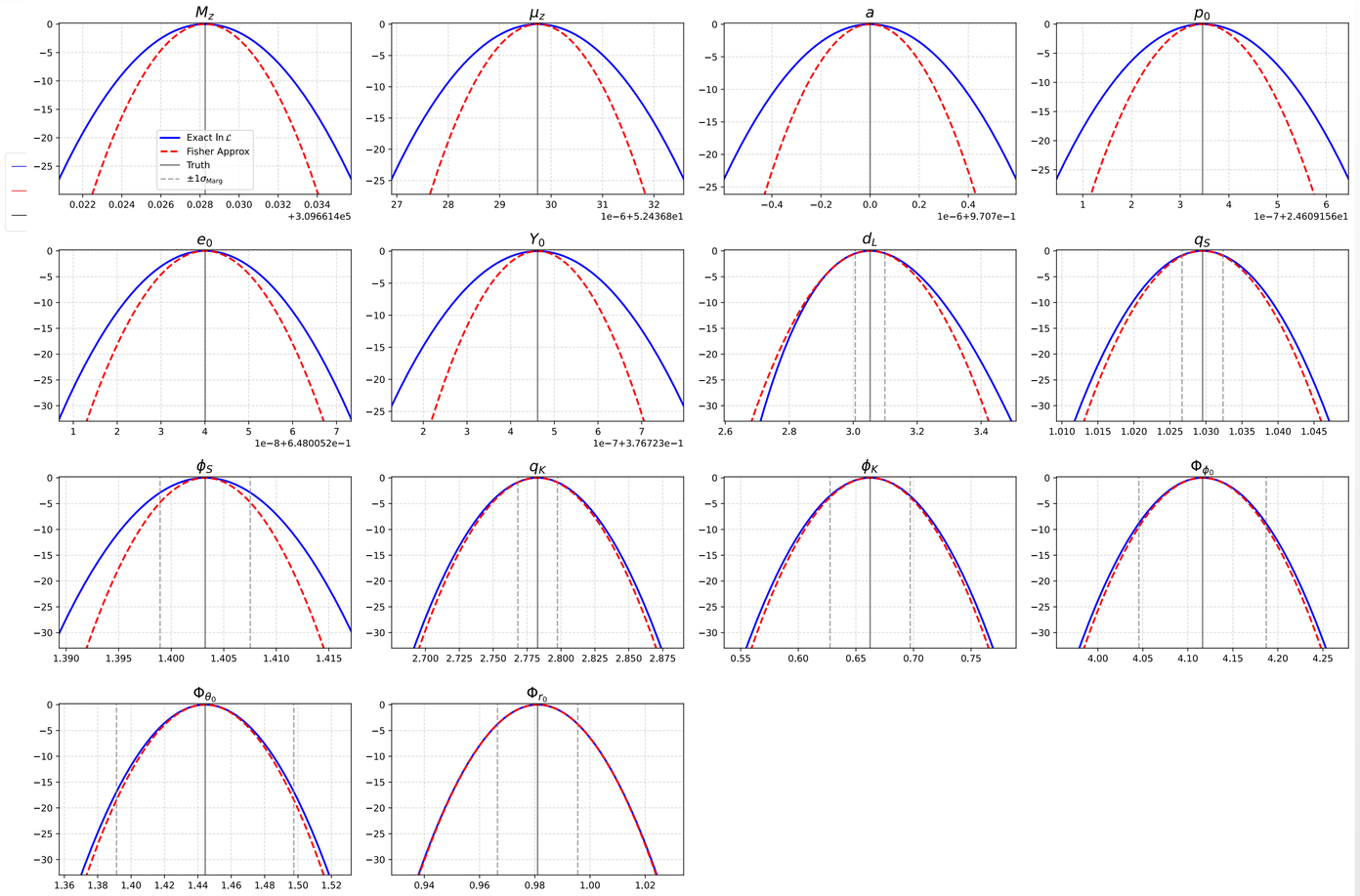


logL for intrinsic parameters drastically drops very close to the true value, and a subtle difference between the 2 may be the source of error.

Example 2: SNR = 63

M	$3 \cdot 10^5 M_\odot$
μ	$52\ M_\odot$
q	$1.7 \cdot 10^{-4}$
a	0.970
p_0	25
e_0	0.65

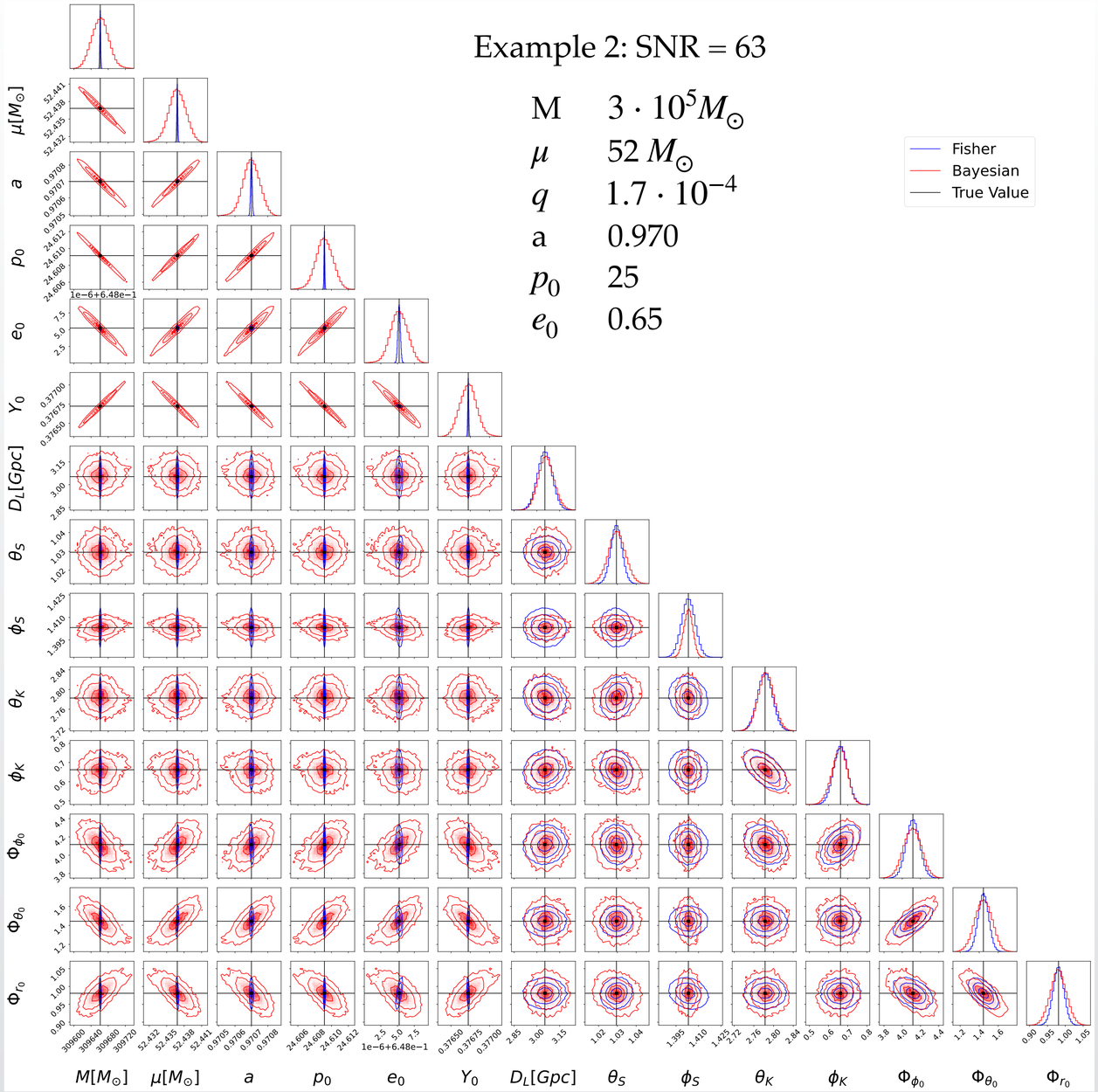




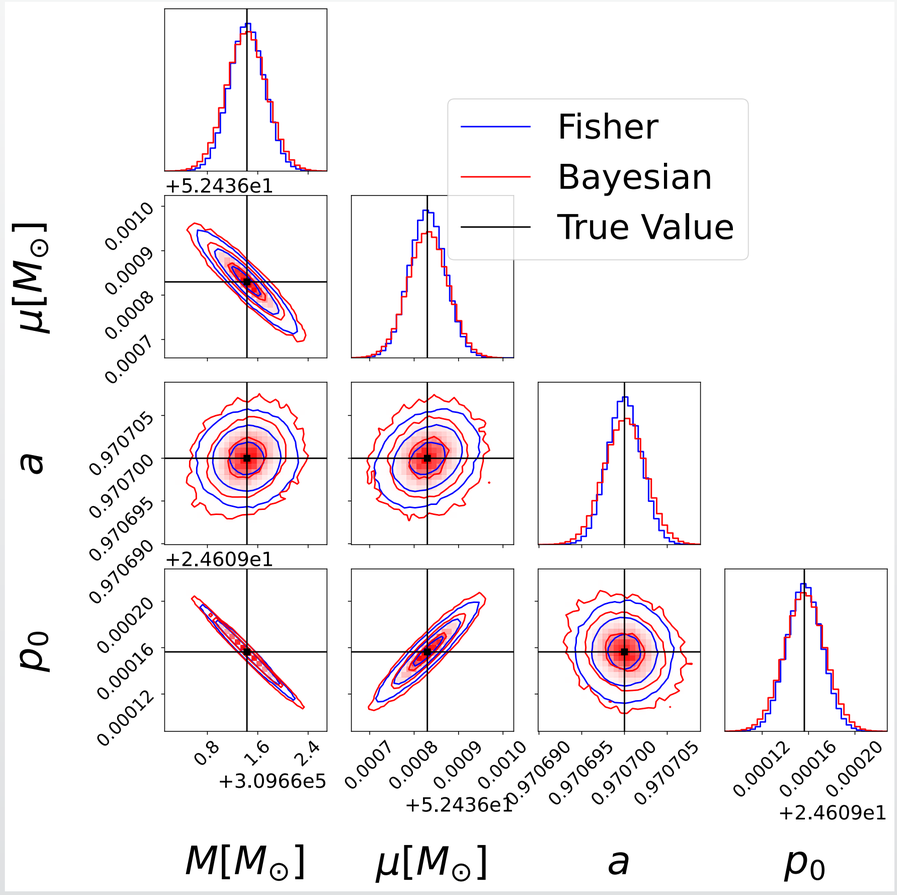
logL for intrinsic parameters drastically drops very close to the true value, and a subtle difference between the 2 may be the source of error.

Example 2: SNR = 63

M	$3 \cdot 10^5 M_\odot$
μ	$52\ M_\odot$
q	$1.7 \cdot 10^{-4}$
a	0.970
p_0	25
e_0	0.65

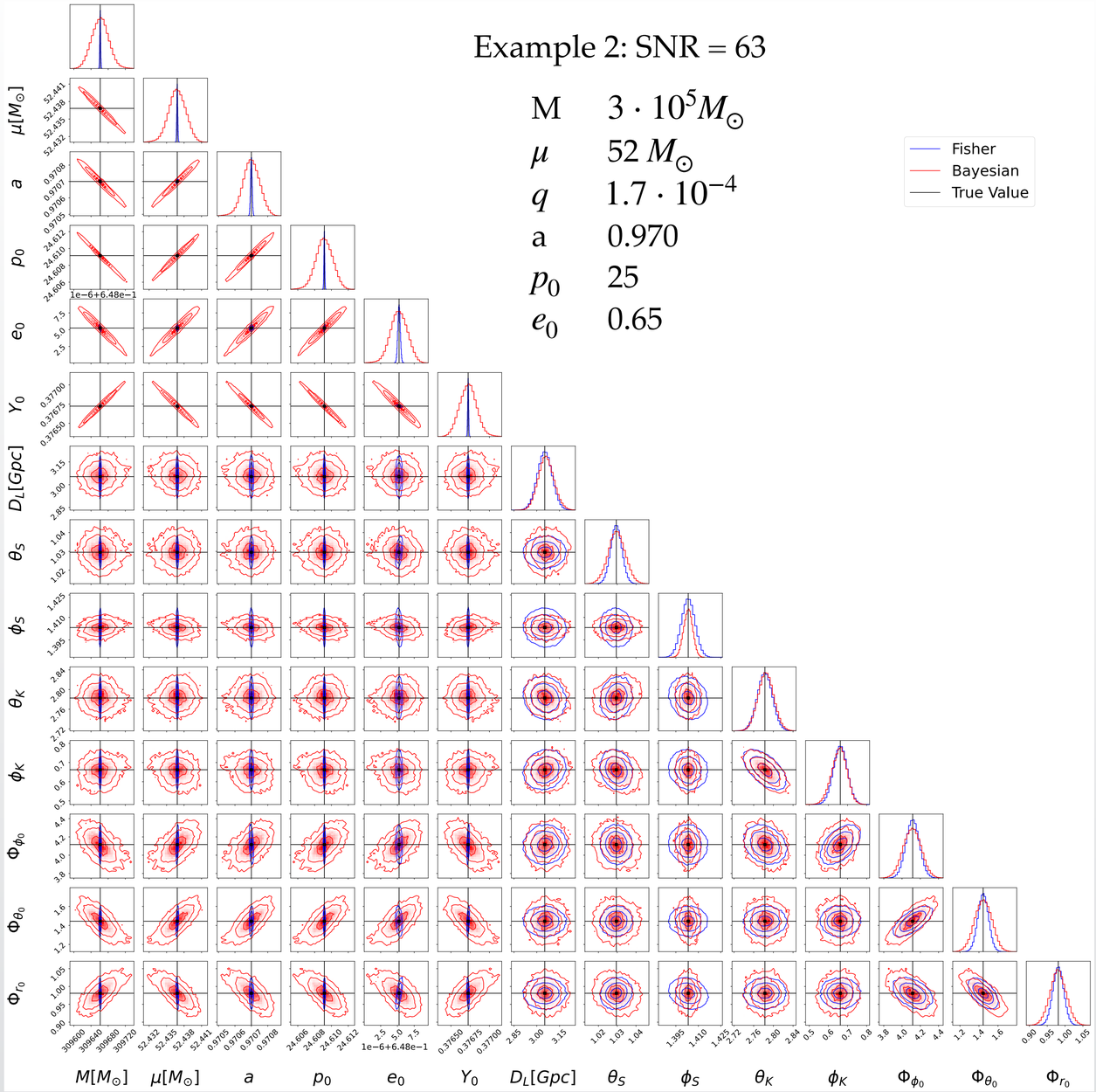


Running the analysis on a smaller subset of parameters could give us an idea if there are problematic parameters



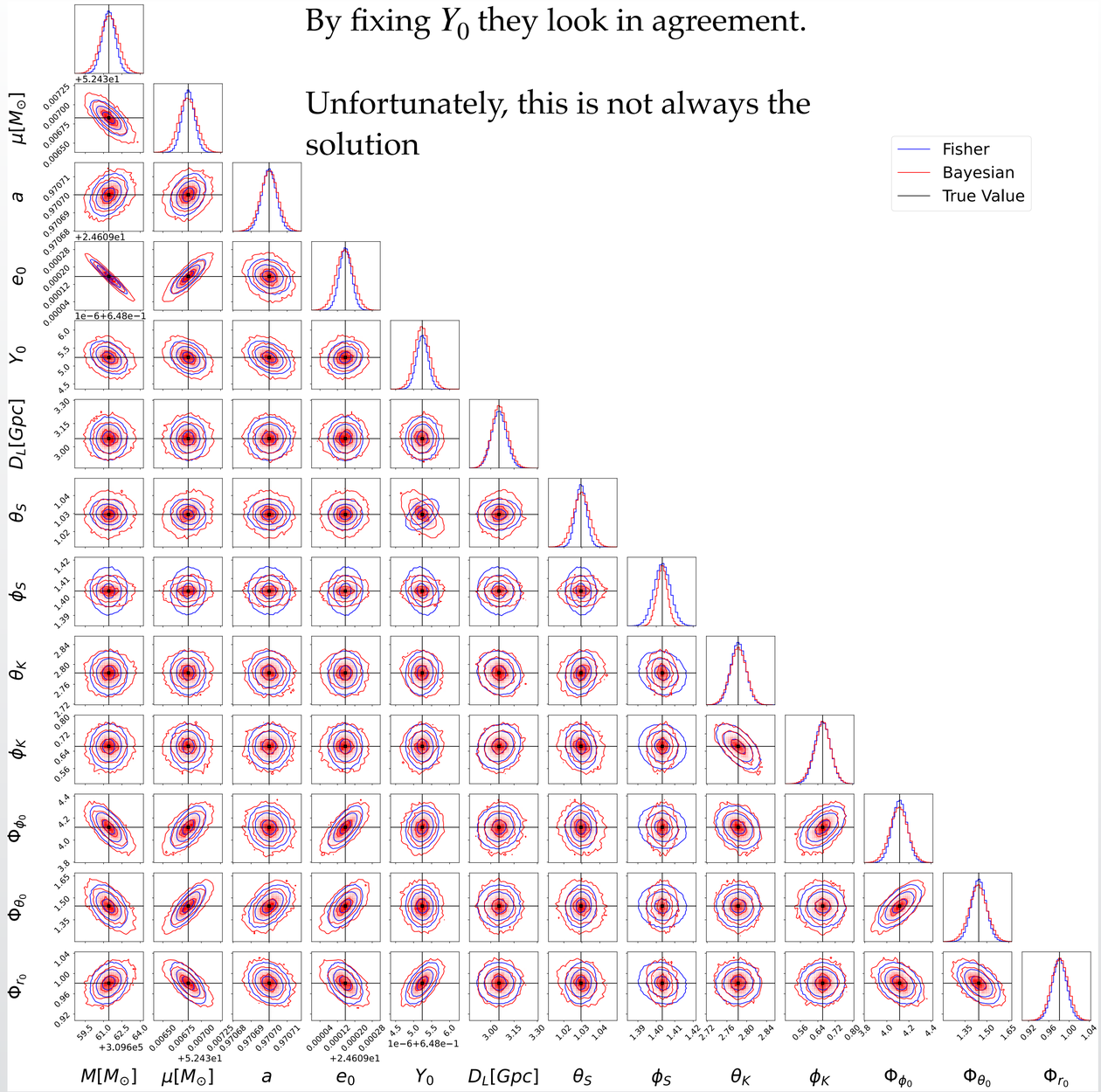
Example 2: SNR = 63

M	$3 \cdot 10^5 M_\odot$
μ	$52\ M_\odot$
q	$1.7 \cdot 10^{-4}$
a	0.970
p_0	25
e_0	0.65



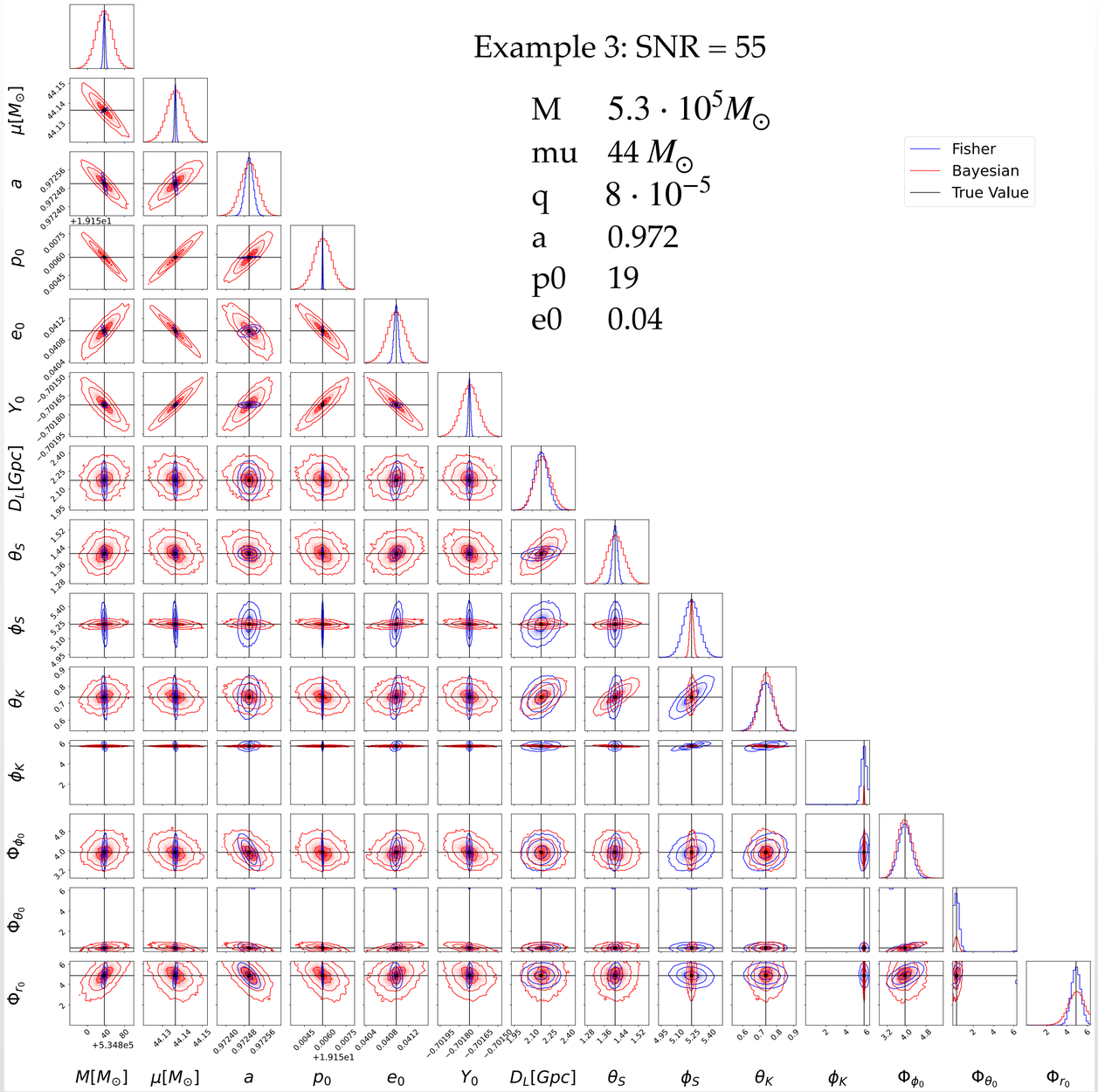
By fixing Y_0 they look in agreement.

Unfortunately, this is not always the solution

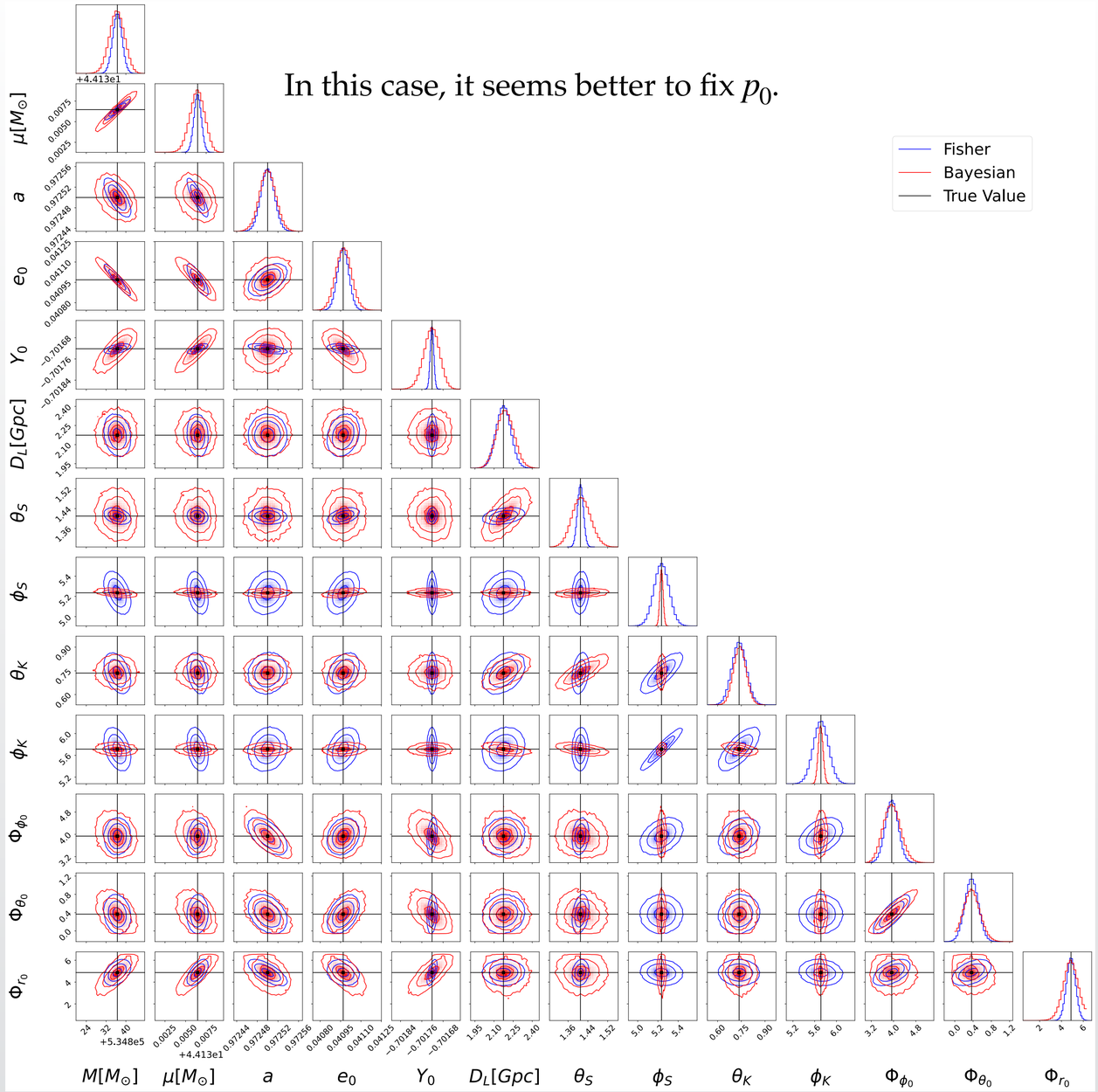


Example 3: SNR = 55

M	$5.3 \cdot 10^5 M_\odot$
mu	$44\ M_\odot$
q	$8 \cdot 10^{-5}$
a	0.972
p0	19
e0	0.04



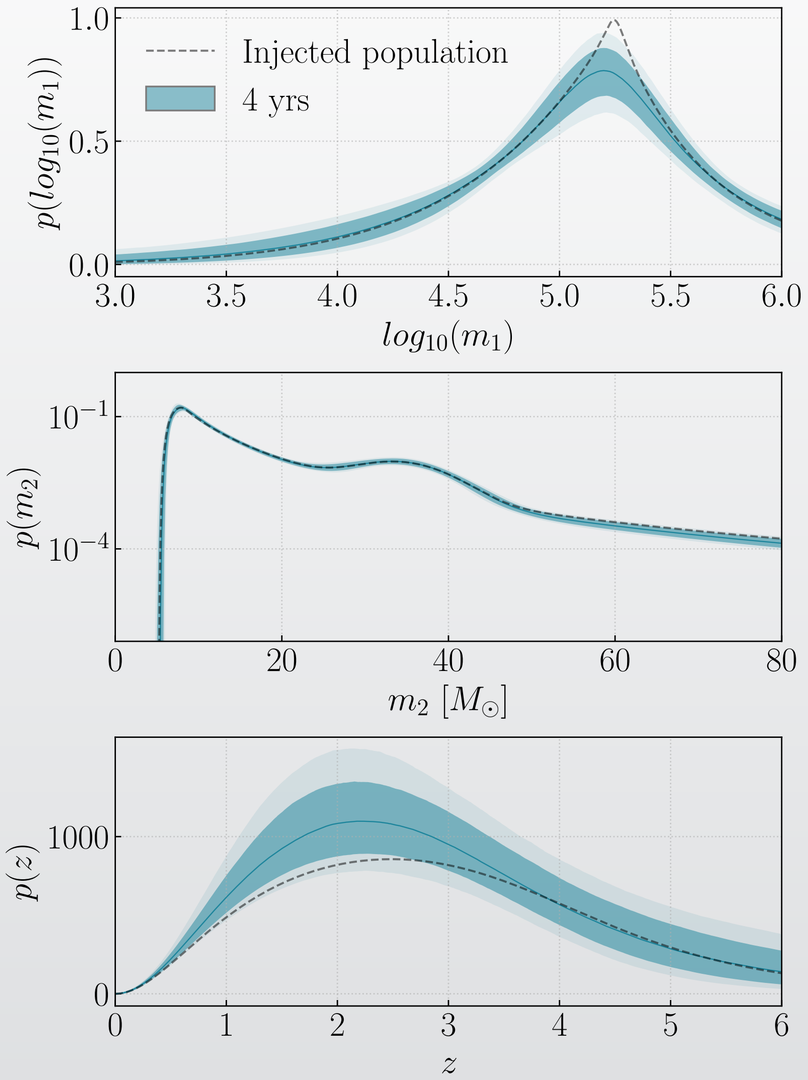
In this case, it seems better to fix p_0.



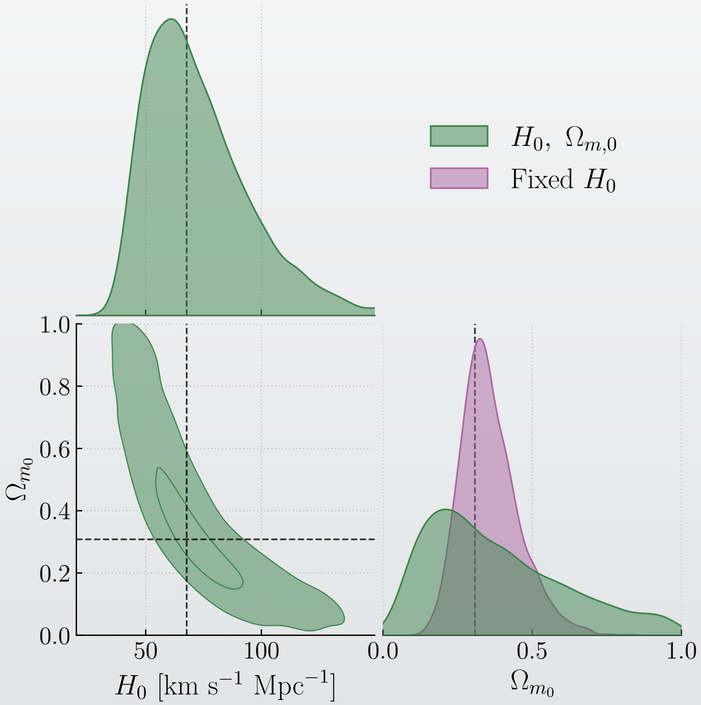
- Why are catalogs important?
  An application: **Spectral Sirens Cosmology with EMRIs** - CosmologyWG

Credit: Vasco Gennari

- **Population Inference First:** Phenomenological model fitting primary/secondary mass and redshift rate (includes selection effects; LVK-style).
- **Joint Cosmological Inference:** Cosmology parameters inferred simultaneously
- **The "Pink" Case:** H_0 is fixed (mimicking post-O5 knowledge before LISA), leaving other parameters free.
- **Key Primary Result:** EMRIs could measure $\Omega_{m,0}$ due to higher redshift reach compared to LVK.



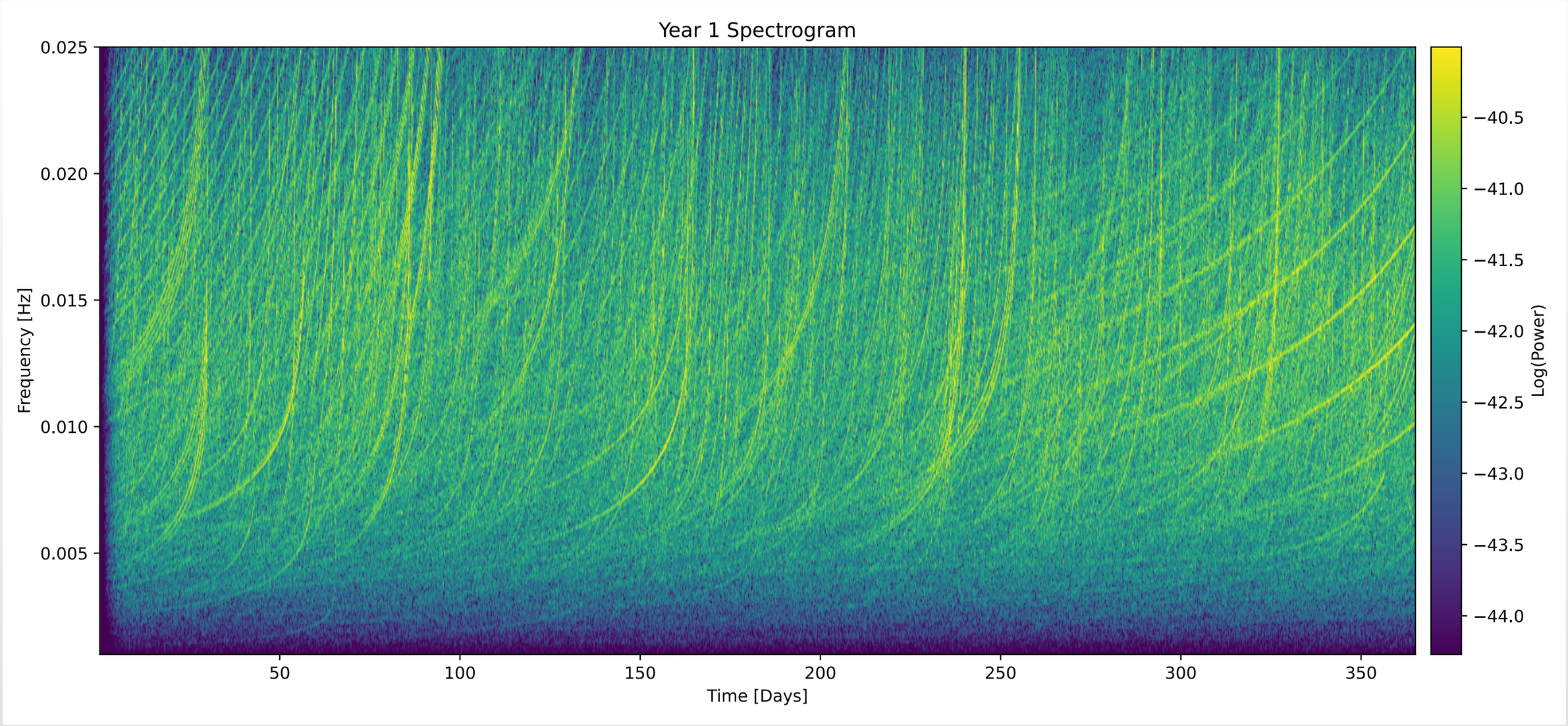
PRELIMINARY

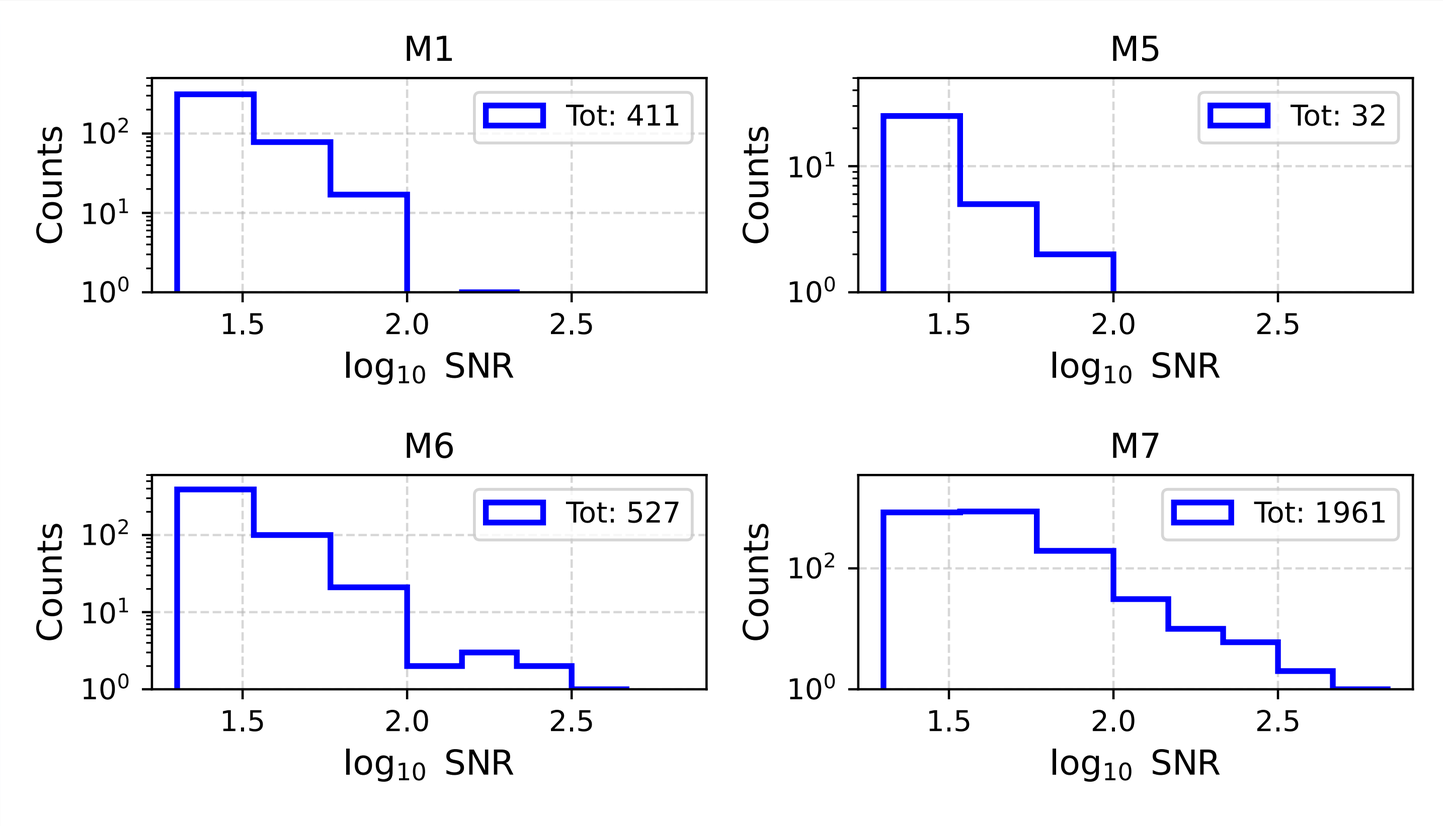


Take-home message and prospects

- Detection rates and horizons are lower than previous predictions.
- Fast tools like Fisher seem to be working only approximately. Parameters are too correlated / instability of the Fisher inversion
- We need new realistic EMRI populations (wet EMRIs, cliffanger EMRIs, etc..)
- How could we improve the efficiency of full MCMC runs?
- What are the most urgent tasks to focus on?

Good luck global fitters




M1
Tot: 411
Counts
10^2
10^1
10^0
1.5
2.0
2.5
$\log_{10}$ SNR
M5
Tot: 32
Counts
10^1
10^0
1.5
2.0
2.5
$\log_{10}$ SNR
M6
Tot: 527
Counts
10^2
10^1
10^0
1.5
2.0
2.5
$\log_{10}$ SNR
M7
Tot: 1961
Counts
10^2
10^0
1.5
2.0
2.5
$\log_{10}$ SNR

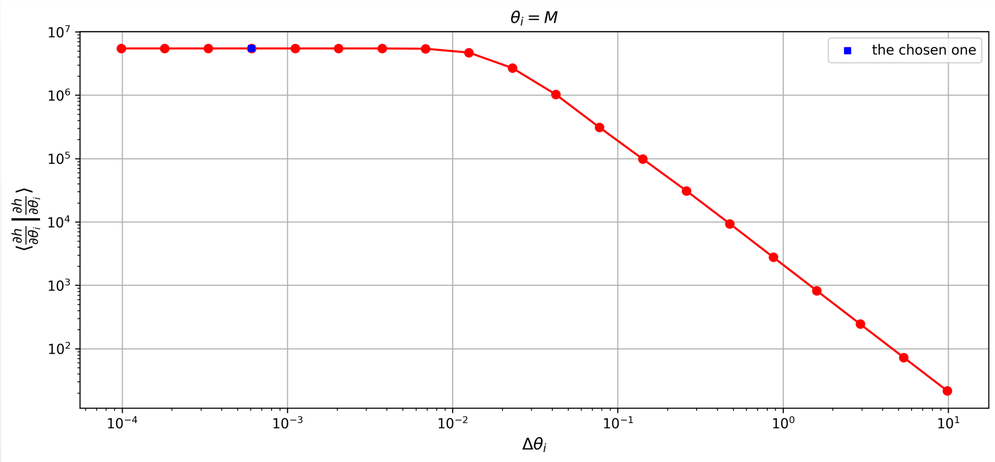
4. Fisher Estimates: Matrix inversion

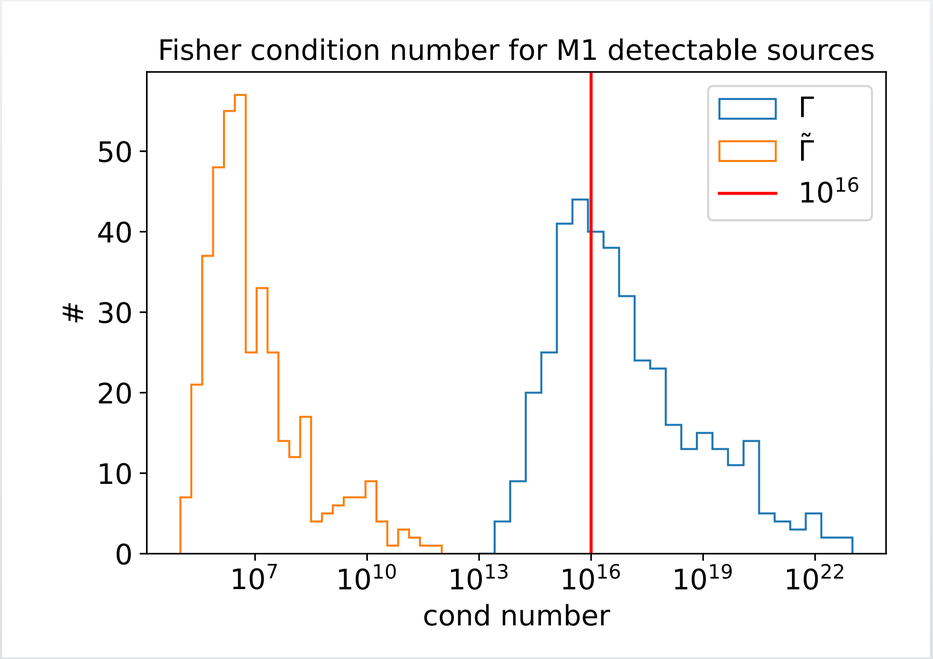
$$\Gamma_{ij} = \left(\frac{\partial h}{\partial \theta^i} \middle| \frac{\partial h}{\partial \theta^j} \right) \rightarrow \Delta\theta^i = \sqrt{(\Gamma^{-1})_{ii}}$$

Large dynamic range of EMRI parameters $\Longrightarrow$ we pre-condition before inversion

1. Pre-conditioning (Diagonal Scaling): $D_{ii} = \dfrac{1}{\sqrt{\Gamma_{ii}}}$
2. Dimensionless, well-conditioned Fisher matrix: $\tilde{\Gamma} = D\Gamma D$
3. Inversion & Rescaling, via Cholesky decomposition $\tilde{\Sigma} = \tilde{\Gamma}^{-1}$
   The physical covariance matrix is $\Sigma = D\tilde{\Sigma}D$
4. Further check on numerical stability: $\max \left| \Gamma\Sigma - I \right| < 10^{-3}$

Derivatives are evaluated at 20 different steps for each parameter to ensure stability.
High computational cost (~1 week for M7)

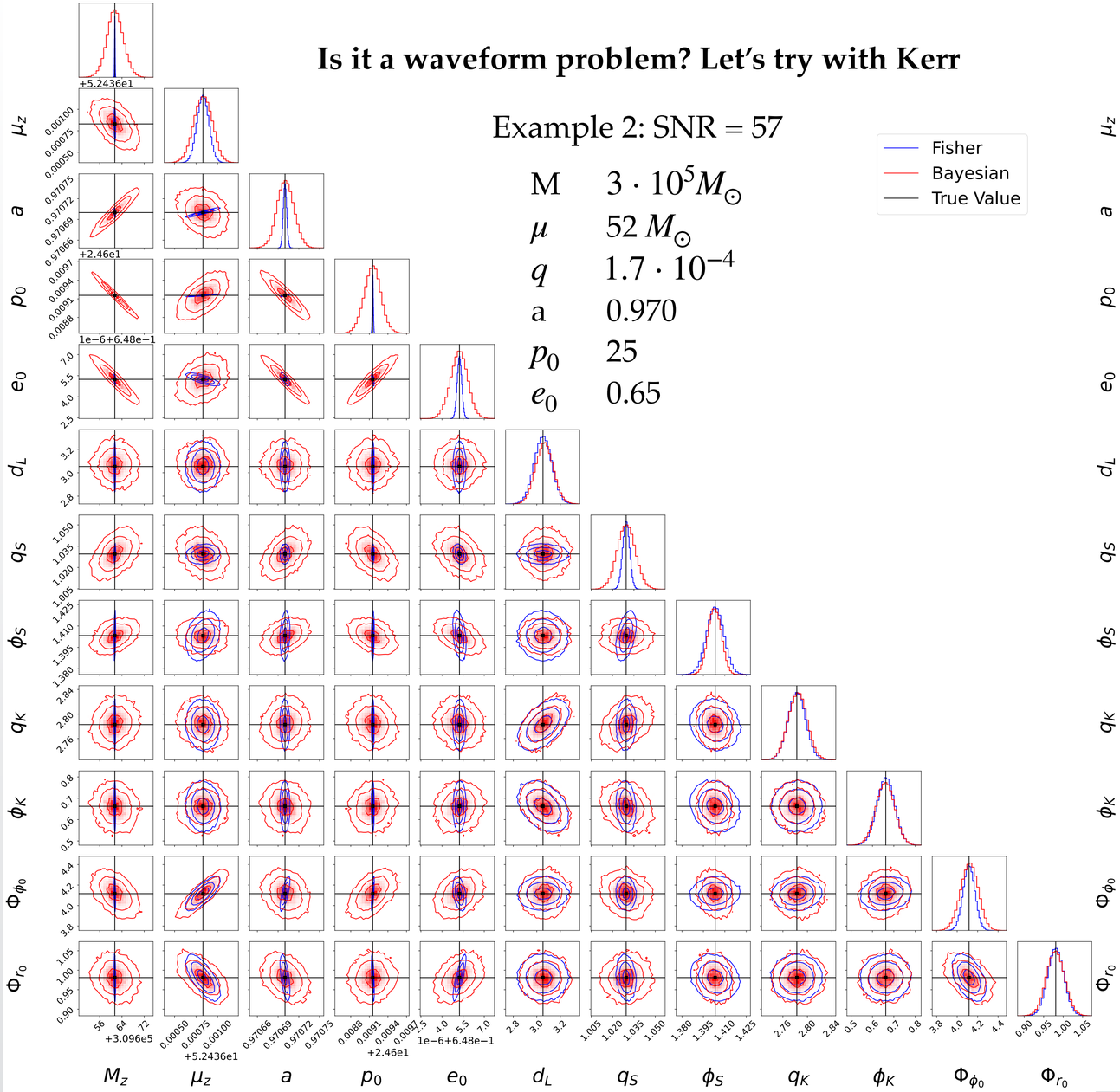




Is it a waveform problem? Let's try with Kerr

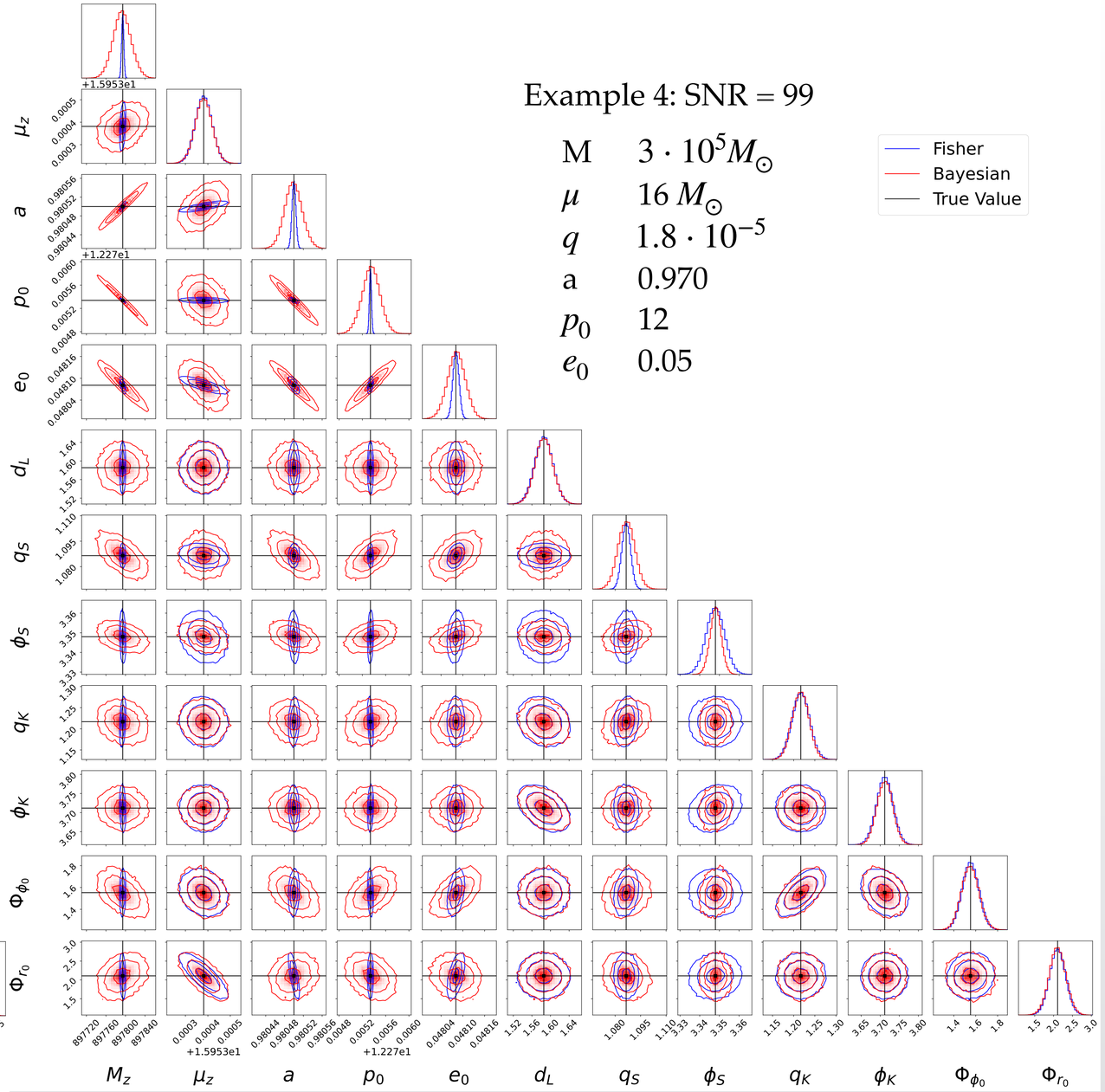
Example 2: SNR = 57

M	$3 \cdot 10^5 M_\odot$
μ	$52\ M_\odot$
q	$1.7 \cdot 10^{-4}$
a	0.970
p_0	25
e_0	0.65



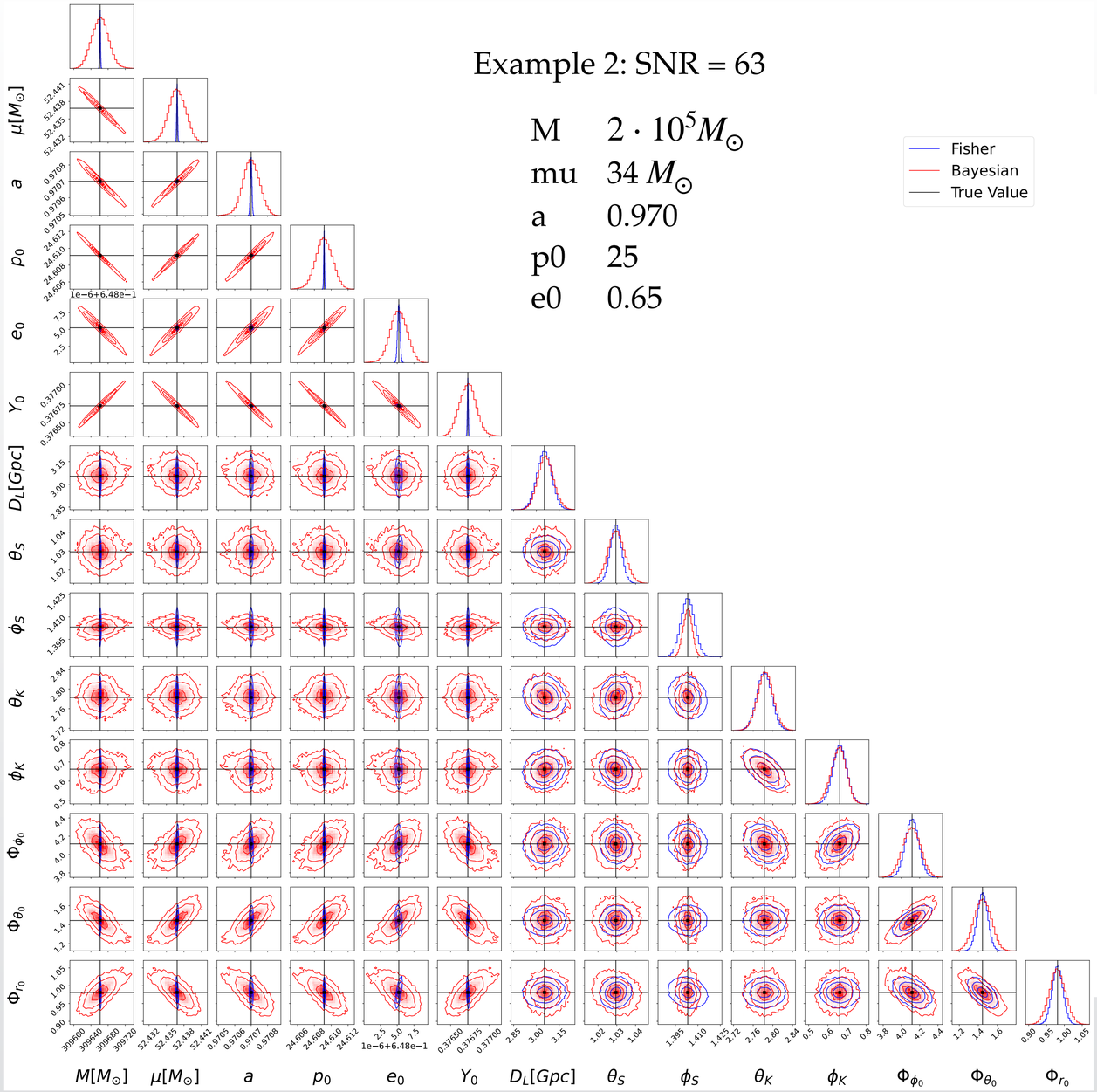
Example 4: SNR = 99

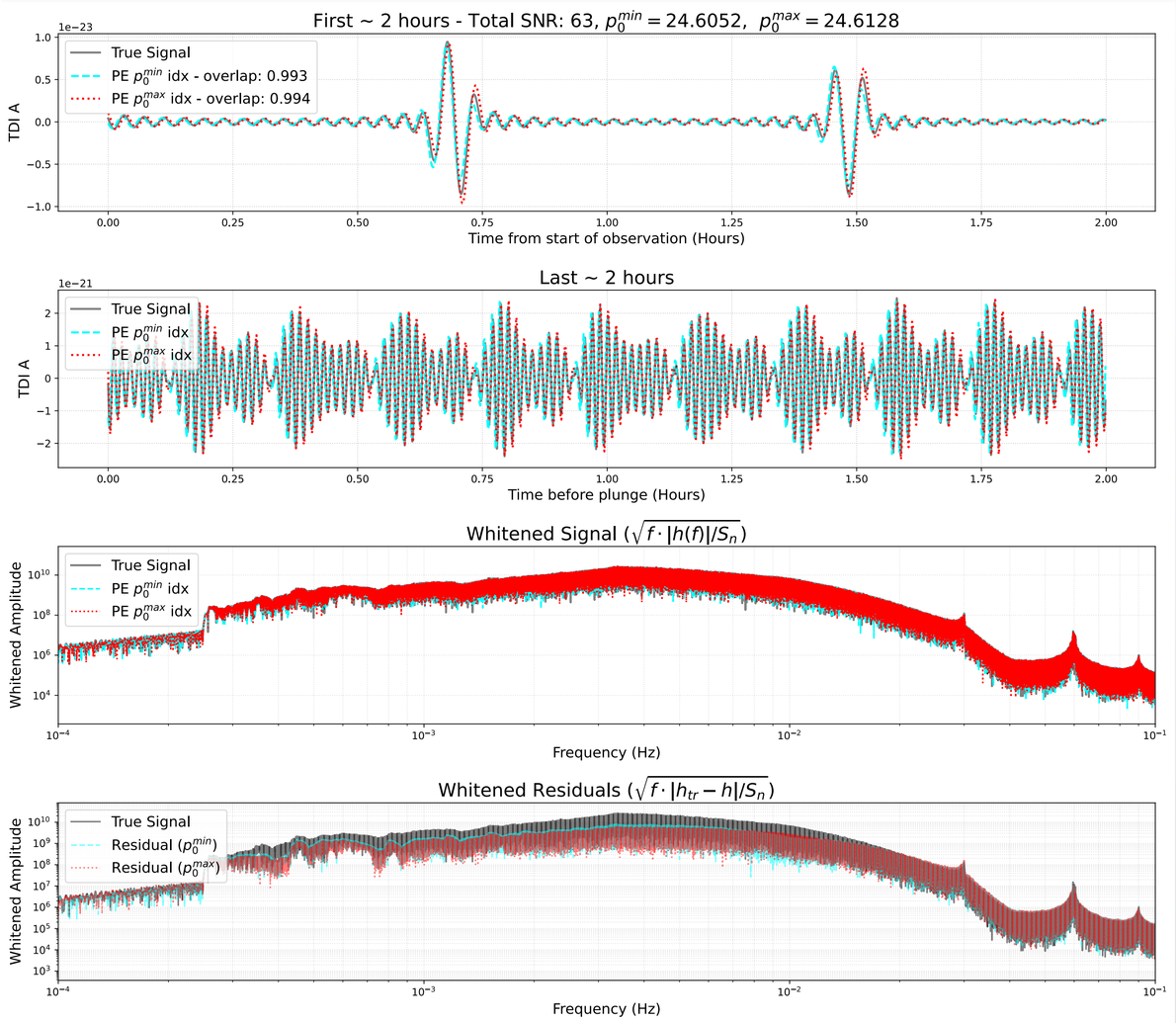
M	$3 \cdot 10^5 M_\odot$
μ	$16\ M_\odot$
q	$1.8 \cdot 10^{-5}$
a	0.970
p_0	12
e_0	0.05

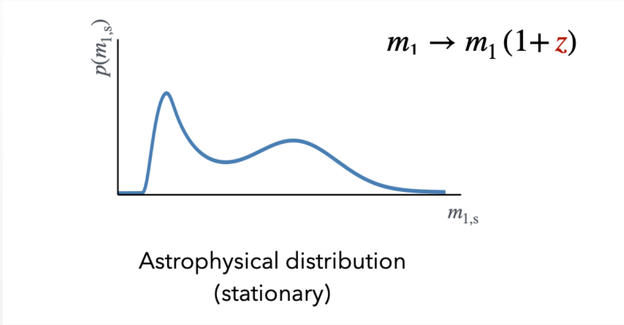


Example 2: SNR = 63

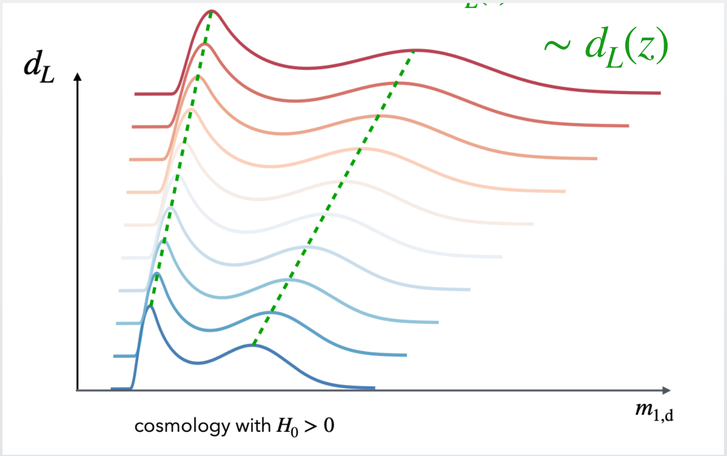
M	$2 \cdot 10^5 M_\odot$
mu	$34\ M_\odot$
a	0.970
p0	25
e0	0.65

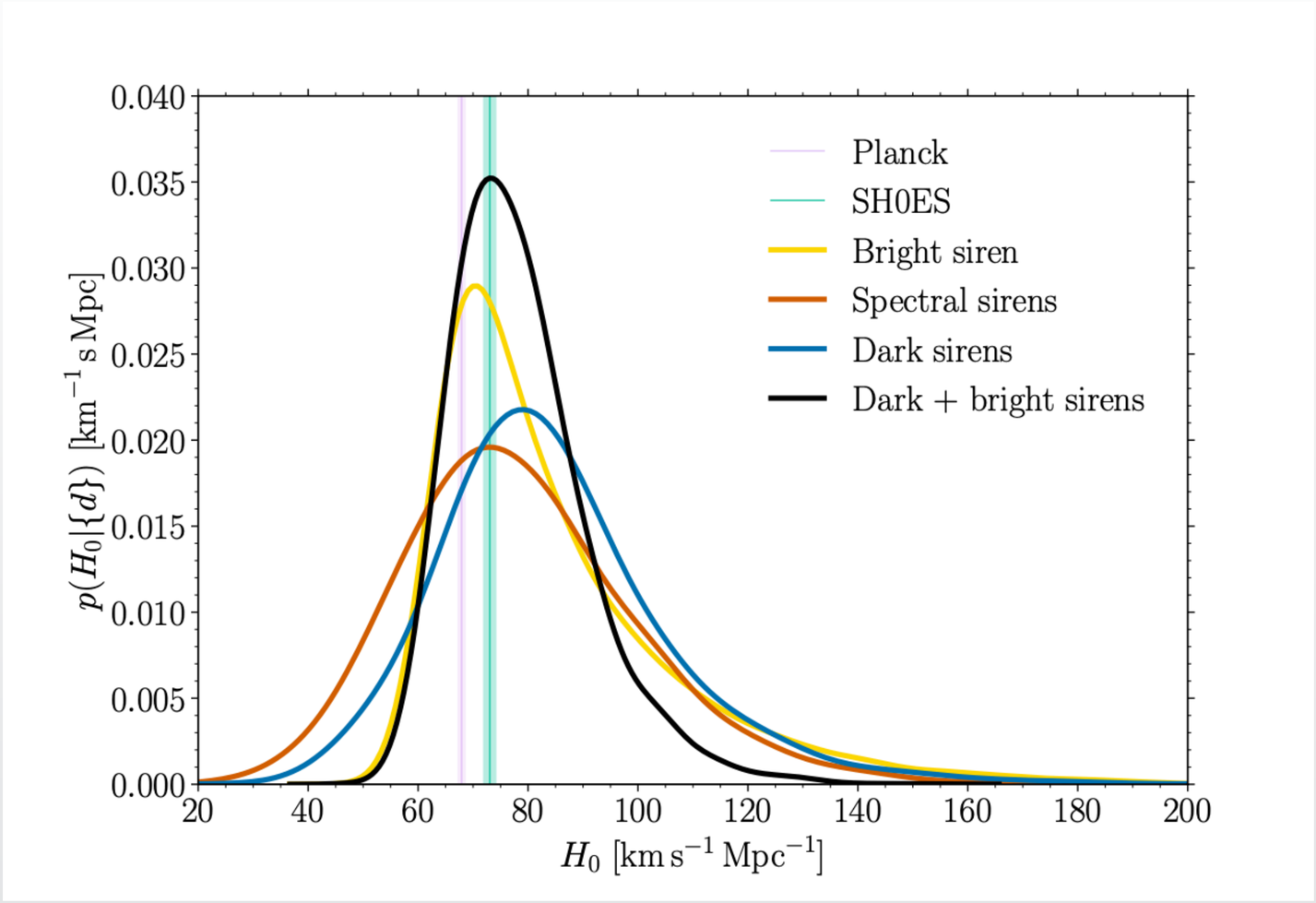






Phenomenological mass scales in BH distributions convey redshift information





GWTC-4.0: Constraints on the Cosmic Expansion Rate and Modified Gravity