EMRI Search and Inference within



A. Sasli, R. Buscicchio, N. Stergioulas

presence of noise non-Gaussianities

June 25, 2025

the LISA Global Fit - Part I

Nikos Karnesis based on works by



Robust EMRI parameter estimation in the (and monsters)

Motivation The likelihood function Applications



NK - EMRI Search and Inference within the LISA Global Fit - Part I - June 25, 2025

Motivation

The Why part

- Too many sources
- Overlap in time and frequency
- Correlations between them
- Noise is not fully known and also
- Non-stationary!
- EMRIs are Long-lived!





ithin the LISA Global Fit - Part I - June 25, 2025

Motivation

The Why part

- Too many sources
- Overlap in time and frequency
- Correlations between them
- Noise is not fully known and also
- Non-stationary!
- EMRIs are Long-lived!

Residuals!





Robustness against outliers The How part

- Certain outliers will be treated by fitting them out.
- Follow the example of BayesWave [Cornish and Littenberg, 2014]
- Still, there might be untreated signals, and residuals from other source types.
- Also slow unmodelled noise PSD variations.







Robustness against outliers The How part

- Certain outliers will be treated by fitting them out.
- Follow the example of BayesWave [Cornish and Littenberg, 2014]
- Still, there might be untreated signals, and residuals from other source types.
- Also slow unmodelled noise PSD variations.











- We can model or treat such situations at the likelihood level.





- We can model or treat such situations at the likelihood level.

$$\mathcal{GH}(x|\lambda,\alpha,\beta,\delta,\mu) = A(\lambda,\alpha,\beta,\delta,\mu) \left(\delta^2 + (x-\mu)^2\right)^{(\lambda-1)} \times K_{\lambda-1/2} \left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp\left(\frac{\lambda^2}{2}\right)$$

with

$$A(\lambda, \alpha, \beta, \delta, \mu) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi} \alpha^{\lambda - 1/2} \delta^{\lambda} K_{\lambda} \left(\delta \sqrt{\alpha^2 - \beta^2}\right)}$$



• Choose a heavier-tailed distribution, based on the Generalized Hyperbolic distribution [K. Prause, 1999].

 $(\frac{1}{2})/2$

 $\exp\left[\beta(x-\mu)\right],$

K_{λ} – modified Bessel function of the third kind





- We can model or treat such situations at the likelihood level.

$$\mathcal{GH}(x|\lambda,\alpha,\beta,\delta,\mu) = A(\lambda,\alpha,\beta,\delta,\mu) \left(\delta^2 + (x-\mu)^2\right)^{(\lambda-1)} \times K_{\lambda-1/2} \left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp\left(\frac{\lambda^2}{2}\right)$$

with

$$A(\lambda, \alpha, \beta, \delta, \mu) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi} \alpha^{\lambda - 1/2} \delta^{\lambda} K_{\lambda} \left(\delta \sqrt{\alpha^2 - \beta^2}\right)}$$



Choose a heavier-tailed distribution, based on the Generalized Hyperbolic distribution [K. Prause, 1999].

 $(\frac{1}{2})/2$

 $\exp\left[\beta(x-\mu)\right],$

Symmetric, $\beta = 0$

 K_{λ} – modified Bessel function of the third kind





- We can model or treat such situations at the likelihood level.

$$\begin{split} \Lambda_{\text{hyp}}(\alpha, \delta; f_i) = & n \sum_{i}^{N_f} \left[\left(\frac{d+1}{2} \log\left(\frac{\alpha}{\delta}\right) + \frac{1-d}{2} \log(2\pi) \right. \\ & \left. - \log(2\alpha) - \log\left(K_{(d+1)/2}(\delta\alpha)\right) \right) \\ & \left. - \alpha \sqrt{\delta^2 + (\tilde{d}_i - \tilde{h}_i)^2 / S_{n,i}} \right]_{K_\lambda - \text{modified Bessel function of the thir}} \end{split}$$

$$n \sum_{i}^{N_{f}} \left[\left(\frac{d+1}{2} \log\left(\frac{\alpha}{\delta}\right) + \frac{1-d}{2} \log(2\pi) - \log(2\alpha) - \log\left(K_{(d+1)/2}(\delta\alpha)\right) \right) - \alpha \sqrt{\delta^{2} + (\tilde{d}_{i} - \tilde{h}_{i})^{2}/S_{n,i}} \right]_{K_{\lambda} - \text{modified Bessel function of the thir}}$$







- We can model or treat such situations at the likelihood level.

$$\begin{split} \Lambda_{\text{hyp}}(\alpha, \delta; f_i) = &n \sum_{i}^{N_f} \left[\left(\frac{d+1}{2} \log \left(\frac{\alpha}{\delta} \right) + \frac{1-d}{2} \log(2\pi) \right. \\ &\left. - \log(2\alpha) - \log \left(K_{(d+1)/2}(\delta\alpha) \right) \right) \\ &\left. - \frac{\alpha}{\sqrt{\delta^2 + (\tilde{d}_i - \tilde{h}_i)^2 / S_{n,i}}} \right]_{K_\lambda - \text{modified Bessel function of the thir}} \end{split}$$









- We can model or treat such situations at the likelihood level.

$$\begin{split} \Lambda_{\text{hyp}}(\alpha, \delta; f_i) = & n \sum_{i}^{N_f} \left[\begin{pmatrix} d+1\\ 2 \log \left(\frac{\alpha}{\delta}\right) + \frac{1-d}{2} \log(2\pi) \\ & -\log(2\alpha) - \log\left(K_{(d+1)/2}(\delta\alpha)\right) \\ & -\alpha\sqrt{\delta^2 + (\tilde{d}_i - \tilde{h}_i)^2/S_{n,i}} \right] \\ & K_{\lambda} - \text{modified Bessel function of the third} \end{split}$$









- We can model or treat such situations at the likelihood level.

$$\Lambda_{\text{hyp}}(\alpha, \delta; f_i) = n \sum_{i}^{N_f} \left[\left(\frac{d+1}{2} \log \left(\frac{\alpha}{\delta} - \log(2\alpha) - \log\left(K_{(d-1)} - \alpha\sqrt{\delta^2 + (\tilde{d}_i - \tilde{h}_i)^2} \right) \right) \right] \right]$$











- Because it can be considered as the "mother" of exponential distributions.
- This is super useful and neat!

 \mathcal{N} if $\delta/\alpha \rightarrow$





• Thus, by fitting its parameters, we can "arrive" at the given distributions of the residual data!

• Because we can do parameter estimation AND get information about the statistics of the residuals!

$$S_n$$
 as $\alpha, \delta \to \infty$

- Because it can be considered as the "mother" of exponential distributions.
- This is super useful and neat!

$$\mathcal{N}$$
 if $\delta/\alpha \rightarrow$

Consider: we still have gaussian noise, but the model we have chosen is wrong. The relation above will still converge to the right answer.



• Thus, by fitting its parameters, we can "arrive" at the given distributions of the residual data!

• Because we can do parameter estimation AND get information about the statistics of the residuals!

$$S_n$$
 as $\alpha, \delta \to \infty$







[Sasli PhD Thesis, 2025]



• The "shape triangle" as a visual tool to aid with the interpretation of the results.

$$\begin{aligned} \zeta &= \delta \sqrt{\alpha^2 - \beta^2}, \quad \varrho = \beta / \alpha, \\ \xi &= (1 + \zeta)^{-1/2}, \quad \chi = \xi \varrho \end{aligned}$$







NK – EMRI Search and Inference w

vithin the LISA Global Fit - Part I - June 25, 2

NK – EMRI Search and Inference w

ithin the LISA Global Fit - Part I - June 25, 2025

 $\log \alpha_1$

thin the LISA Global Fit - Part I - June 25, 2025

NK – EMRI Search and Inference w

NK – EMRI Search and Inference w

$$\begin{aligned} \zeta &= \delta \sqrt{\alpha^2 - \beta^2}, \quad \varrho = \beta / \alpha, \\ \xi &= (1 + \zeta)^{-1/2}, \quad \chi = \xi \varrho \end{aligned}$$

ithin the LISA Global Fit - Part I - June 25, 2025

- Many different astrophysical population models predict confusion stochastic signal.
- Cyclo-stationary due to LISA motion.
- Gaussian? •

Frequency [Hz]

[NK et al, 2025]

- Do as before with the simple example.
- Divide the data into i segments.
- Fit \mathbf{a}_i and $\mathbf{\delta}_i$ parameters for each.
- Get an estimate of the level of the PSD, AND
- an estimate of their underlying statistical distribution.

NK et al, 2025

[NK et al, 2025]

[NK et al, 2025]

Application to LISA data B. Transient sources

Application to LISA data B. Transient sources

- For the case of Binary Black Holes, and
- Gaussian noise,
- Results of the Whittle and the hyperbolic likelihood match!
- Also, the hyperbolic converges to the Gaussian distribution! [Sasli PhD Thesis, 2025]

Application to LISA data C. Transient sources – EMRIs

	1
	_

Application to LISA data C. Transient sources – EMRIs

- Do a simple exercise.
- How does the noise knowledge affects the Parameter Estimation of EMRIs?
- Simulate Gaussian data and add "unknown noisy components" that disrupt the noise PSD.
- We can pretend that this mismatch can be due to slow PSD variations of the noise, glitches, imperfect residuals, un-subtracted Galactic Binaries. [...]
- 10^{-1} · Perform analyses with both Gaussian and Hyperbolic likelihoods, by assuming the nominal PSD value for the noise (black).

Application to LISA data C. Transient sources – EMRIs

	1
	_

Application to LISA data C. Transient sources – EMRIs -Hyperbolic —Gaussian ·?.... ${\mu/M_{\odot}\over [\times 10^{-4}+10^{1}]}$ ~;· 0.0 .' , ×, ? ,3,0 0.505 -0.500 D_s/Gpc 0.495 0.490 0.1253 ~~??~ θ_S ~;30 1.49 $0.16^{5} 0.19^{0} 0.19^{5} 0.50^{0} 0.50^{5}$ D_{s}/Gpc 39 3. 7. 0. 1. 3. ~^{1,20} ~^{1,50} ~^{1,51} ~^{1,51} x3 0 \$ M/M_{\odot} [+10⁶] θ_S μ/M_{\odot} [×10⁻⁴ + 10¹]

	1
	_

Application to LISA data C. Transient sources – EMRIs -Hyperbolic —Gaussian ~<u>~</u>) $\frac{\mu/M_{\odot}}{[\times 10^{-4}+10^1]}$ 0⁰. 7.3 ,3,0 0,505 -0.500 D_s/Gpc 0.495 0.490 0.263 ~~??~ θ_S 1.¹²⁹ 30 30 7, 00 1, 30 $0^{,10^{5}}$ $0^{,10^{9}}$ $0^{,10^{5}}$ $0^{,50^{9}}$ $0^{,50^{5}}$ D_{s}/Gpc ~^{1,20} ~^{1,50} ~^{1,51} ~^{1,51} ×? **\$**3 0 M/M_{\odot} [+10⁶] μ/M_{\odot} [×10⁻⁴ + 10¹]

Not entirely fair comparison though. Serves as an example application and motivation.

- PE and Search for Long-lived sources will depend on
 - Correlations to other sources (overlaps)
 - Noise non-stationarities (varying PSDs, glitches)
- A framework such as the one with the hyperbolic likelihood is flexible to
 - Unknown noise PSDs (gaussian data)
 - Glitches and residuals (departures from Gaussianity)
 - Data gaps treatment? •

Message to take home

- PE and Search for Long-lived sources will depend on
 - Correlations to other sources (overlaps)
 - Noise non-stationarities (varying PSDs, glitches)
- A framework such as the one with the hyperbolic likelihood is flexible to
 - Unknown noise PSDs (gaussian data)
 - Glitches and residuals (departures from Gaussianity)
 - Data gaps treatment? •

Message to take home

• PE and Search for Long-lived sources will depend on

	s (overlag
 Residuals become less of a problem. Departures from Gaussianity are more controlled. 	ying PSE
3.Easy to compute, the inner product is something we need anyways.	vith the h
4.Suitable for any kind of analysis with a likelihood.	an data)
	tures fro

Data gaps treatment? •

Message to take home

- ps)
- Ds, glitches)
- hyperbolic likelihood is flexible to

om Gaussianity)

• PE and Search for Long-lived sources will depend on

	s (overlap
 Residuals become less of a problem. Departures from Gaussianity are more controlled. 	ying PSE
3.Easy to compute, the inner product is something we need anyways.	/ith the ł
4.Suitable for any kind of analysis with a likelihood.	an data)
	tures fro

Data gaps treatment? •

Message to take home

- ps)
- Ds, glitches)
- hyperbolic likelihood is flexible to

om Gaussianity)

PE and Search for Long-lived sources will depend on

	s (overlap
 Residuals become less of a problem. Departures from Gaussianity are more controlled. 	ying PSE
3.Easy to compute, the inner product is something we need anyways.	/ith the ł
4.Suitable for any kind of analysis with a likelihood.	an data)
	tures fro

• Data gaps treatment?

Message to take home

ps)

Ds, glitches)

hyperbolic likeliho

om Gaussianity)

1. Sometimes flexibility is a problem: Sweep everything under the rug, or in this case under the heavy tails.

2.Penalize estimates across the duration of the measurement, even if outlier is localized in time (tackled with t-f analyses).

Flexible

Relevant to the discussion about EMRIs DA

- Provide a robust framework for searching for EMRIs in a Global Fit scheme.
- Incorporate uncertainties into the overall distribution of the residuals, and
- get an estimate for it as well.
- Hide biases under the rug!?
- Estimate properties of populations of sources (Gaussianity? Stochasticity? Anisotropy?) •
- Perhaps avoid the extra parametrisation steps during the GF procedure? (maybe for certain pipelines or certain purposes)
- Robust detection statistics and significance in the presence of other signals or unknown noise.
- Reason to keep the likelihood computations inside the Deep Analyses module of the DDPC.

