Extending the Fisher matrix formalism towards the edges

Francesco Iacovelli

Based on arXiv:2405.02286

In collaboration with: Michele Mancarella, Stefano Foffa, Niccolò Muttoni, Michele Maggiore

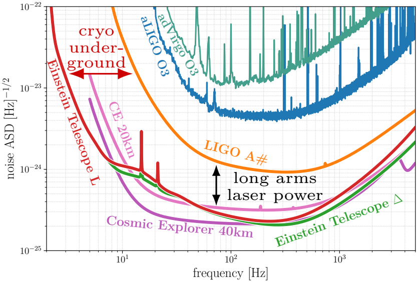
University of Geneva (UNIGE) – Department of Theoretical Physics

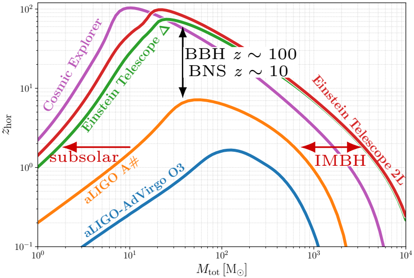
XIV ET Symposium – Maastricht 2024

Introduction: 3G GW detectors

2G detectors offer outstanding possibilities...
...but the potential of 3G detectors is unprecedented

Thanks to their technological advancements and the bigger facilities, ET and CE will have a broader frequency range and sensitivities improved more than 10 times compared to LVK





Assessing the capabilities of 3G detectors is fundamental to take informed decisions!

Introduction: challenged by the numbers

One of the key challenges when performing studies for ET and CE that emerged in recent years is the number of detectable sources

Network	BBH/yr	BNS/yr	NSBH/yr
LVK–O4	$\mathcal{O}(10^2)$	$\mathcal{O}(1-10)$	$\mathcal{O}(1-10)$
ET	$\mathcal{O}(10^4)$	$\mathcal{O}(10^3-10^5)$	$\mathcal{O}(10^3-10^4)$
ET+2CE	$\mathcal{O}(10^4-10^5)$	$\mathcal{O}(10^4-10^5)$	$\mathcal{O}(10^3-10^5)$

Currently used Bayesian parameter estimation codes, like `bilby`, can take $\mathcal{O}(1\,\mathrm{day/ev})$ to perform the analysis...

...and we do not have 10^5 days 🥲

Introduction: Fisher codes

Various groups all across the world started to tackle the problem, and by now there are three public codes that can perform such a complex analysis exploiting the Fisher matrix formalism:

GWBENCH: a novel Fisher information package for gravitational-wave benchmarking

S. Borhanian[1,2]

[1] *Institute for Gravitation and the Cosmos, Department of Physics, Pennsylvania State University, University Park, PA 16802, USA*
[2] *Theoretisch-Physikalisches Institut, Friedrich-Schiller-Universität Jena, 07743, Jena, Germany*[*]
(Dated: August 31, 2021)

GWFISH: A simulation software to evaluate parameter-estimation capabilities of gravitational-wave detector networks

Jan Harms[1,2], Ulyana Dupletsa[1,2], Biswajit Banerjee[1,2], Marica Branchesi[1,2], Boris Goncharov[1,2], Andrea Maselli[1,2], Ana Carolina Silva Oliveira[3], Samuele Ronchini[1,2], and Jacopo Tissino[1,2]
[1] *Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy*
[2] *INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi, Italy and*
[3] *Department of Physics, Columbia University in the City of New York, New York, NY 10027, USA*
(Dated: May 6, 2022)

GWFAST: a Fisher information matrix Python code for third–generation gravitational–wave detectors

Francesco Iacovelli,[1] Michele Mancarella,[1] Stefano Foffa,[1] and Michele Maggiore[1]

[1] *Département de Physique Théorique, Université de Genève, 24 quai Ernest Ansermet, 1211 Genève 4, Switzerland*

see also TiDoFM, Li et al. (2022) and Pieroni et al. (2022)

Introduction: GW parameter estimation

A GW signal as observed by a detector can be expressed as

$$\boxed{s(t) = h_0(t) + n(t)\,.}$$

Defining the inner product for any two time–domain signals as

$$\boxed{(a\,|\,b) = 4\,\mathrm{Re}\left\{\int_0^\infty \mathrm{d}f\,\frac{\tilde{a}^*(f)\,\tilde{b}(f)}{S_n(f)}\right\} \implies \mathrm{SNR} = (h_0\,|\,h_0)^{1/2}\,,}$$

we have for the GW likelihood, choosing a waveform model $h(\boldsymbol{\theta})$,

$$\boxed{\mathcal{L}(s\,|\,\boldsymbol{\theta}) \propto \exp[-\left(s - h(\boldsymbol{\theta})\,|\,s - h(\boldsymbol{\theta})\right)/2]\,.}$$

Introduction: Fisher matrix

To approximate this likelihood it is possible to expand the template signal around the true waveform as

$$h(\boldsymbol{\theta}) \approx h_0 + h_i \delta\theta^i + \ldots, \quad \text{with} \quad \delta\theta^i \equiv \theta^i - \theta^i_0 \,.$$

The approximation where only first derivatives of the signal are included is known as the *linearized signal approximation* (LSA), and plugging it in the likelihood we then obtain

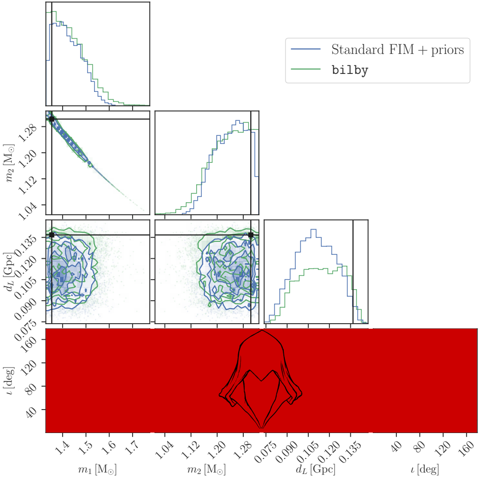
$$\mathcal{L}(s \,|\, \boldsymbol{\theta}) \propto \exp\left[-\frac{1}{2}\,(n\,|\,n) + \delta\theta^i\,(n\,|\,h_i) - \frac{1}{2}\delta\theta^i\delta\theta^j\,(h_i\,|\,h_j) \right],$$

which, neglecting noise–dependent factors, is a multivariate Gaussian with inverse covariance $\Gamma_{ij} = (h_i|h_j)$.

Introduction: Fisher matrix validity

It can be shown that the LSA is equivalent to the high–SNR limit, and we can thus expect the Fisher to be valid in this regime, but degeneracies can spoil the validity irrespectively of the SNR! Vallisneri (2008)

One can try to partially "help" the Fisher e.g. by adding physical priors (see Ulyana's talk!), but in some cases this is not enough.

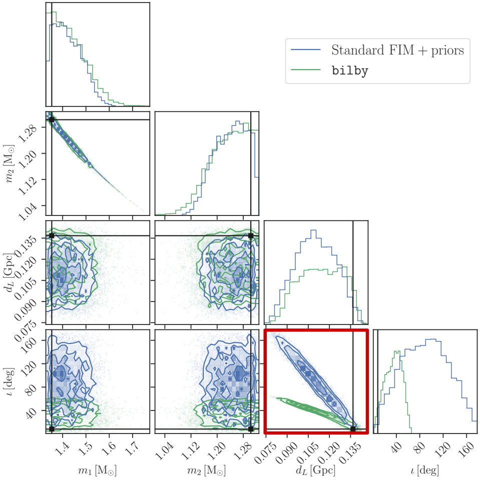


face–on BNS, SNR$\sim$45 at LVKI O5

Introduction: Fisher matrix validity

It can be shown that the LSA is equivalent to the high–SNR limit, and we can thus expect the Fisher to be valid in this regime, but degeneracies can spoil the validity irrespectively of the SNR! Vallisneri (2008)

One can try to partially "help" the Fisher e.g. by adding physical priors (see Ulyana's talk!), but in some cases this is not enough.

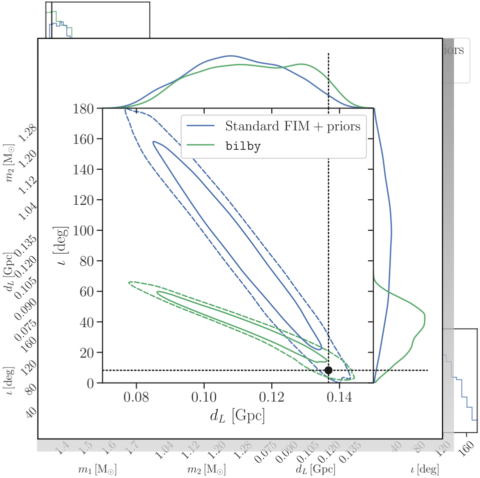


face–on BNS, SNR∼45 at LVKI O5

Introduction: Fisher matrix validity

It can be shown that the LSA is equivalent to the high–SNR limit, and we can thus expect the Fisher to be valid in this regime, but degeneracies can spoil the validity irrespectively of the SNR! Vallisneri (2008)

One can try to partially "help" the Fisher e.g. by adding physical priors (see Ulyana's talk!), but in some cases this is not enough.



face–on BNS, SNR∼45 at LVKI O5

Extending the Fisher likelihood: general idea

Luckily, the dependence of the likelihood on some parameters is fairly simple, so why bothering to do the gaussian approximation at all in these directions?

$$h(\bar{\boldsymbol{\theta}}, d_L, \iota) = \frac{1}{d_L}\left[h_+(\bar{\boldsymbol{\theta}})\left(\frac{1+\cos^2\iota}{2}\right) + i\, h_\times(\bar{\boldsymbol{\theta}})\cos\iota\right] .$$

We can split $\boldsymbol{\theta}$ in a set $\bar{\boldsymbol{\theta}}$, for which the LSA is made, and $\boldsymbol{\beta}$, on which the dependence of $\mathcal{L}$ is taken exact

$$h(\boldsymbol{\beta}\,,\bar{\boldsymbol{\theta}}) \approx \bar{h}(\boldsymbol{\beta}) + \delta\bar{\theta}^i\; h_i(\boldsymbol{\beta})\,, \quad \bar{h}(\boldsymbol{\beta}) \equiv h(\boldsymbol{\beta}\,,\bar{\boldsymbol{\theta}}_0)\,,$$
$$h_i(\boldsymbol{\beta}) \equiv \partial_{\bar{\theta}^i} h(\boldsymbol{\beta}\,,\bar{\boldsymbol{\theta}})\Big|_{\bar{\boldsymbol{\theta}}=\bar{\boldsymbol{\theta}}_0}\,, \quad \delta\boldsymbol{\theta} \equiv \bar{\boldsymbol{\theta}} - \bar{\boldsymbol{\theta}}_0\,.$$

Extending the Fisher likelihood: general idea

Defining $\Delta h(\boldsymbol{\beta}) \equiv \bar{h}(\boldsymbol{\beta}) - h_0$ the likelihood then simply becomes (neglecting noise terms)

$$\boxed{\begin{aligned} -\log \mathcal{L}(s|\boldsymbol{\beta}\,,\bar{\boldsymbol{\theta}}) \propto & \frac{1}{2}(\Delta h(\boldsymbol{\beta})|\Delta h(\boldsymbol{\beta})) \\ & + \delta\bar{\theta}^i(\Delta h(\boldsymbol{\beta})|h_i(\boldsymbol{\beta})) + \frac{1}{2}\delta\bar{\theta}^i\Gamma_{ij}(\boldsymbol{\beta})\delta\bar{\theta}^j\,. \end{aligned}}$$

It might seem that the $\boldsymbol{\beta}$–dependence of Γ_{ij} and of $(\Delta h|h_i)$ would make it problematic to compute them numerically, but in the most interesting cases this does not happens because the $\boldsymbol{\beta}$–dependence of $h(\boldsymbol{\beta}\,,\bar{\boldsymbol{\theta}})$ is simple!

Extending the Fisher likelihood: warm-up with d_L

In the simple case $\boldsymbol{\beta} = d_L$ case we have

$$\boxed{\Delta h(d_L) = -\frac{\delta d_L}{d_L}\, h_0 = \frac{\delta d_L}{d_L}\, d_{L\,0}\, \partial_{d_L} h\Big|_{d_L = d_{L\,0}} ,}$$

so the likelihood reads (with $\Gamma_{ij} = \Gamma_{ij}(d_L = d_{L\,0})$)

$$\boxed{\begin{aligned} -2\log\mathcal{L}(s|d_L\,, \bar{\boldsymbol{\theta}}) \propto \left(\frac{d_{L\,0}}{d_L}\right)^2 & [(\delta d_L)^2 \Gamma_{d_L\, d_L} \\ & + 2\delta\bar{\theta}^i \Gamma_{i\, d_L} \delta d_L + \delta\bar{\theta}^i \Gamma_{ij} \delta\bar{\theta}^j]\,, \end{aligned}}$$

which is close to the standard Fisher but already suppresses $d_L \to 0$!

Extending the Fisher likelihood: infamous $d_L - \iota$

We now turn to the more interesting case $\boldsymbol{\beta} = \{d_L, \iota\}$. The dependence on ι is an overall factor for the two polarizations separately, so defining $c_+(\iota) = (1+\cos^2\iota)/2$ and $c_\times(\iota) = \cos\iota$ we have

$$\boxed{\Delta h(d_L\,,\iota) = \frac{h^+_{d_L}}{\bar{c}_+}\frac{d_{L\,0}}{d_L}\left[\bar{c}_+\delta d_L - d_{L\,0}\,\Delta c_+\right] + (+\to\times)\,,}$$

where $\Delta c_{+,\times} \equiv c_{+,\times}(\iota) - \bar{c}_{+,\times}$ and $\bar{c}_{+,\times} = c_{+,\times}(\iota_0)$, and for $\mathcal{L}$

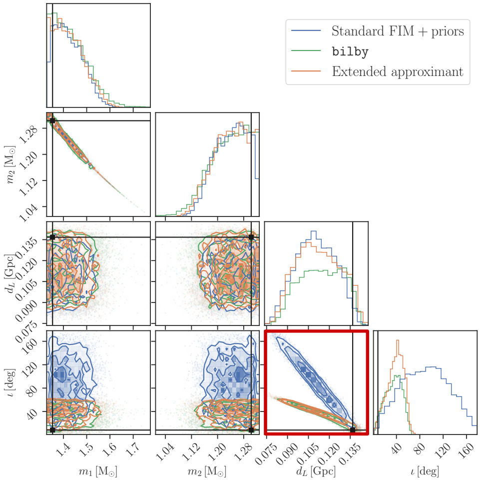
$$\boxed{-2\log\mathcal{L}(s|d_L\,,\iota\,,\bar{\boldsymbol{\theta}}) \propto \left(\frac{d_{L\,0}}{d_L}\right)^2\left[\delta\hat{\hat{\theta}}^a_I\Gamma^{ab}_{IJ}\delta\hat{\hat{\theta}}^b_J\right]\,, \quad \begin{array}{c} a,b=\{+,\times\} \\ I\,(J)=\{d_L,\, i(j)\}\end{array}}$$

in which $\delta\hat{d}_L^{+,\times} \equiv \bar{c}_{+,\times}\delta d_L - d_{L\,0}\Delta c_{+,\times}\,,\ \delta\hat{\theta}_i^{+,\times} \equiv c_{+,\times}\delta\theta_i.$

Extending the Fisher likelihood: infamous $d_L - \iota$

So we just need to compute the two sets of derivatives separately for the two polarizations (or, if we are lazy, build them from three Fishers evaluated at specific values of ι).

Thanks to this extension now the agreement in the $d_L - \iota$ plane is perfect for any value of the inclination, and for the rest of the parameters it works as before!

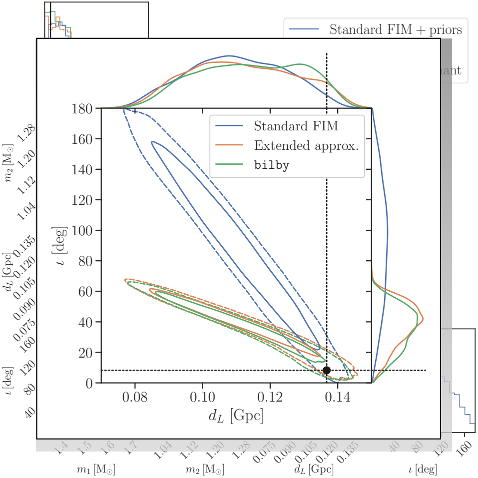


face–on BNS, SNR∼45 at LVKI O5

Extending the Fisher likelihood: infamous $d_L - \iota$

So we just need to compute the two sets of derivatives separately for the two polarizations (or, if we are lazy, build them from three Fishers evaluated at specific values of ι).

Thanks to this extension now the agreement in the $d_L - \iota$ plane is perfect for any value of the inclination, and for the rest of the parameters it works as before!



face–on BNS, SNR∼45 at LVKI O5

Extending the Fisher likelihood: other parameters

This formalism can be easily extended to other parameters, e.g. for $\boldsymbol{\beta} = \{d_L, \iota, \psi\}$ we simply gain an index and have to compute a total of four sets of derivatives

$$\boxed{-2\log\mathcal{L}(s|d_L\,,\iota\,,\bar{\boldsymbol{\theta}}) \propto \left(\frac{d_{L\,0}}{d_L}\right)^2 [\delta\hat{\bar{\theta}}_I^{a\alpha}\Gamma_{IJ}^{a\alpha,\,b\beta}\delta\hat{\bar{\theta}}_J^{b\beta}]\,. \quad \begin{matrix} a, b = \{+, \times\} \\ I\,(J) = \{d_L, i(j)\} \\ \alpha, \beta = \{\cos\psi, \sin\psi\} \end{matrix}}$$

As far as the dependence on a parameter is easy, this formalism can be applied, and lead to exact results at the same price of a bunch of Fishers!

Including the noise

The effect of noise fluctuations, which is needed in order to avoid biases when performing hierarchical inference, can be included in the above likelihood without resorting to the Fisher approximation.

Generate signal and noise
As for an injection, we generate the signal from a waveform with parameters $\boldsymbol{\theta}_0$ and an explicit noise realization from the detector PSD

$\longrightarrow$

Find the ML
For each event, we minimize the negative log-likelihood using a standard `scipy` minimizer initialized close to $\boldsymbol{\theta}_0$

$\longrightarrow$

Expand around ML
Compute the likelihood expanding around the ML point $\hat{\boldsymbol{\theta}}$

which works since at first order ($\hat{h} = h(\hat{\boldsymbol{\theta}})$, $\delta\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$)

$$\boxed{-2\log\mathcal{L}(\mathcal{D}_{\rm GW}\,|\,\boldsymbol{\theta}) \propto -\delta\hat{\theta}^i\delta\hat{\theta}^j\left(\hat{h}_i\,|\,\hat{h}_j\right) + 2\delta\hat{\theta}^i\left(h_0 + n - \hat{h}\,|\,\hat{h}_i\right)\,.}$$

Including the noise

The effect of noise fluctuations, which is needed in order to avoid biases when performing hierarchical inference, can be included in the above likelihood without resorting to the Fisher approximation.

Generate signal and noise

As for an injection, we generate the signal from a waveform with parameters $\boldsymbol{\theta}_0$ and an explicit noise realization from the detector PSD

$\longrightarrow$

Find the ML

For each event, we minimize the negative log-likelihood using a standard `scipy` minimizer initialized close to $\boldsymbol{\theta}_0$

$\longrightarrow$

Expand around ML

Compute the likelihood expanding around the ML point $\hat{\boldsymbol{\theta}}$

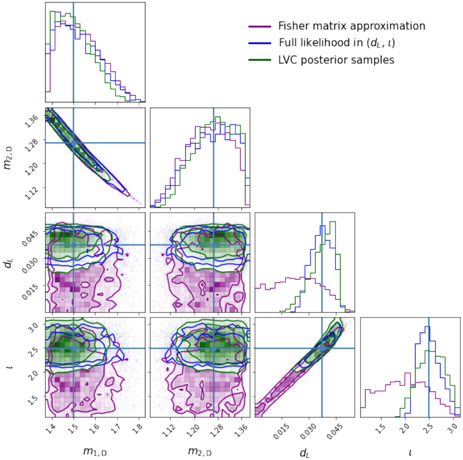
which works since at first order ($\hat{h} = h(\hat{\boldsymbol{\theta}})$, $\delta\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$)

≈ 0 at ML by definition!

$$\boxed{-2\log\mathcal{L}(\mathcal{D}_{\rm GW}\,|\,\boldsymbol{\theta}) \propto -\delta\hat{\theta}^i\delta\hat{\theta}^j\left(\hat{h}_i\,|\,\hat{h}_j\right) + 2\delta\hat{\theta}^i\left(h_0 + n - \hat{h}\,|\,\hat{h}_i\right)\,.}$$

Conclusions and future work

- The proposed formalism can be a powerful tool to overcome some known limitations of the Fisher matrix approximation, while keeping a low computational cost;
- It can be extended to other parameters and other expressions of the likelihood (e.g. the one marginalized over Φ_c);
- It can be complemented to include the effect of noise fluctuations.



GW170817

Thanks for your attention...questions?

I am also available at Francesco.Iacovelli@unige.ch

Slides available on ET-TDS at ET-0218A-24