Dipartimento di Fisica e Astronomia Galileo Galilei

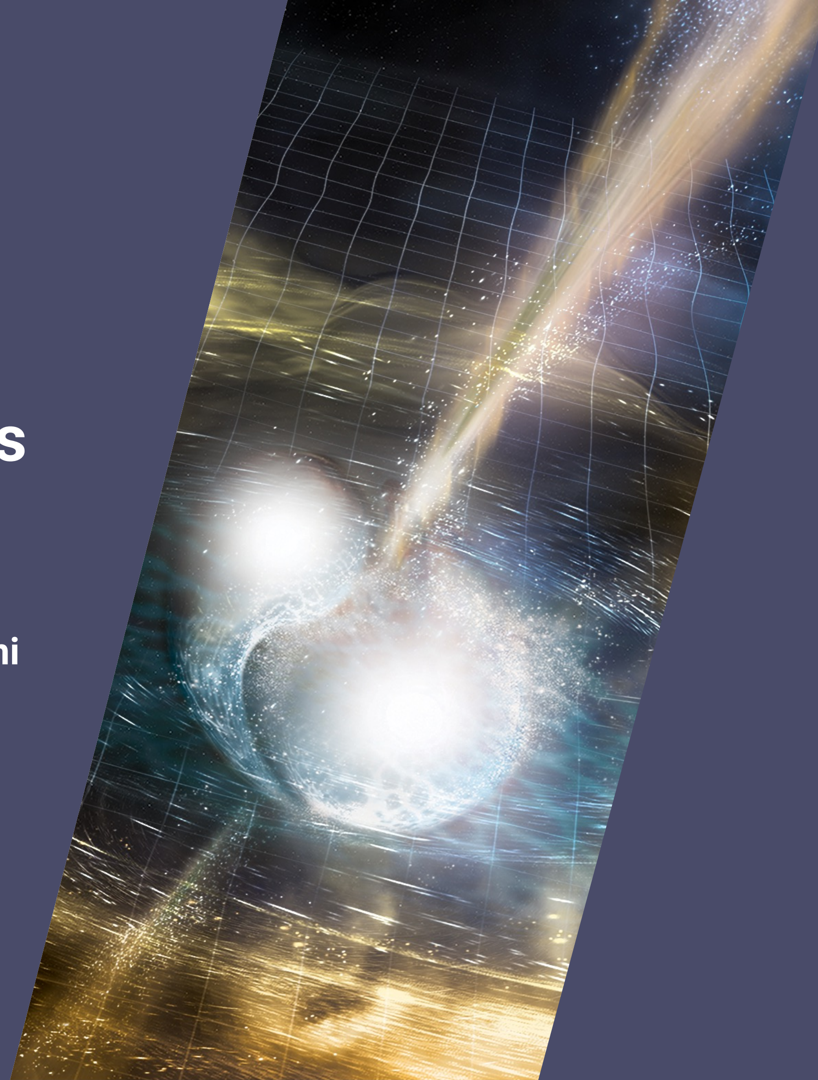
Università degli Studi di Padova

Distribution of a Population of Gravitational Waves Sources

By Giulia Doda & Laura Ravagnani

Advanced Statistics for Physics Analysis – Final Project

4th July 2023

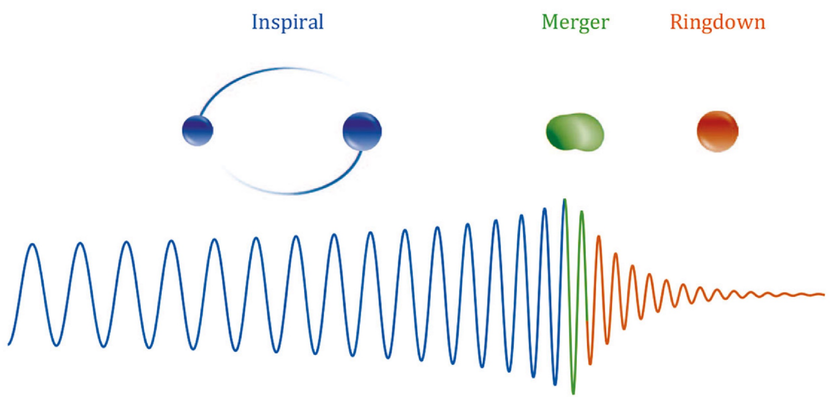


Outline

- Introduction
- Statistical model for the number of GW sources
- Two different models for the population density
- Simulations
- Results
- Conclusions

Introduction

Gravitational waves (GW) are ripples in space-time traveling at the speed of light. The most common **sources** of GWs are compact binary coalescences (**CBCs**). Compact binary systems are binary black holes (BBHs), binary neutron stars (BNSs), and black hole/neutron star binaries (NS-BHs).



The **interaction** between the two objects and their subsequent **merger** result in a **GW emission**.

The Physical Problem

Why do we care about the population distribution of GW sources?

The population distribution of GW sources can answer some questions about the **evolution of the Universe and of the GW sources** themselves:

- *How many GW sources are there in the Universe?*
- *Has the population of compact objects changed as the Universe has evolved?*
- *How do compact binaries form?*

Each possible answer is associated with a different population distribution of GW sources.

Our Problem

How can we understand the population distribution of GW sources from GW observations?

Interferometers such as LIGO and Virgo observe GWs from CBCs, and from these events we can obtain information about the population distribution of GW sources. More specifically, in this work **we want to understand from some simulated observations whether the population density follows one distribution or another**.

Our Problem

How can we understand the distribution of the population of GW sources from GW observations?

Interferometers such as LIGO and Virgo observe GWs from CBCs, and from these events we can obtain information about the distribution of the population of GW sources. More specifically, in this work **we want to understand from some simulated observations whether the population density follows one distribution or another**.

→ **we will focus on observations from BNS mergers**

Our Analysis

1

Statistical model for the number of sources N

Our Analysis

1

Statistical model for the number of sources N

2

Simple detection efficiency model ε

Our Analysis

1. Statistical model for the number of sources N

2. Simple detection efficiency model ε

3. Simulated GW observations from two different population models

Our Analysis

1. Statistical model for the number of sources N
2. Simple detection efficiency model ε
3. Simulated GW observations from two different population models
4. Analysis of the population density n

Our Analysis

1. Statistical model for the number of sources N
2. Simple detection efficiency model ε
3. Simulated GW observations from two different population models
4. Analysis of the population density n
5. n estimate dependencies on the observation time T and the source distance r

Our Analysis

1

Statistical model for the number of sources N

2

Simple detection efficiency model ε

3

Simulated GW observations from two different population models

4

Analysis of the population density n

5

n estimate dependencies on the observation time T and the source distance r

A Statistical Model for N

Hypotheses

- cosmological redshift $z \ll 1$
- numerical density of the sources population $n \ll 1\ Mpc^{-3} yrs^{-1}$
- volume $V = 4\pi R^2 \Delta R$
- statistically independent sources positions

average number of sources in a time interval T :

$\langle N \rangle = nTV = nT4\pi R^2 \Delta R$

the number of sources N follows a **Poisson probability distribution** with rate $\lambda = \langle N \rangle$

A Statistical Model for N

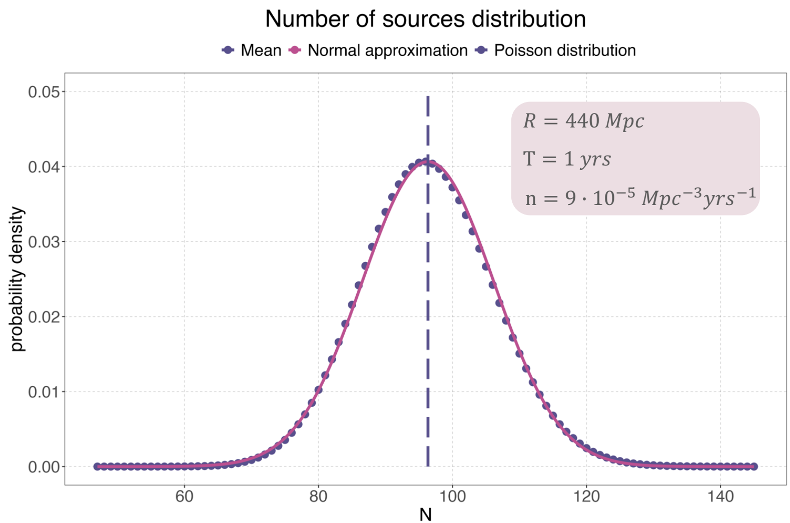
Hypotheses

- cosmological redshift $z \ll 1$
- numerical density of the sources population $n \ll 1\ Mpc^{-3}yrs^{-1}$
- volume $V = 4\pi R^2 \Delta R$
- statistically independent sources positions

average number of sources in a time interval T :

$$\langle N \rangle = nTV = nT4\pi R^2 \Delta R$$

the number of sources N follows a **Poisson probability distribution** with rate $\lambda = \langle N \rangle$



Our Analysis

1

Statistical model for the number of sources N

2

Simple detection efficiency model ε

3

Simulated GW observations from two different population models

4

Analysis of the population density n

5

n estimate dependencies on the observation time T and the source distance r

Detection Efficiency

N is not the actual number of GW sources we can observe with our detectors

we must consider the **detection efficiency** $\boldsymbol{\varepsilon(h)}$, a function of the **GW strain** $\boldsymbol{h}$

physical quantity measured by the interferometers: relative displacement between test masses during the passage of a GW

Detection Efficiency

N is not the actual number of GW sources we can observe with our detectors

we must consider the **detection efficiency** $\boldsymbol{\varepsilon(h)}$, a function of the **GW strain** $\boldsymbol{h}$

physical quantity measured by the interferometers: relative displacement between test masses during the passage of a GW

we consider a simple phenomenological model for the detection efficiency, characterized by two parameters h_0, w

$$\varepsilon(h) = \int_{-\infty}^{h} \frac{1}{\sqrt{2\pi w^2}} exp\left[-\frac{(h' - h_0)^2}{2w^2}\right] dh'$$

Detection Efficiency

N is not the actual number of GW sources we can observe with our detectors

we must consider the **detection efficiency $\boldsymbol{\varepsilon(h)}$**, a function of the **GW strain $\boldsymbol{h}$**

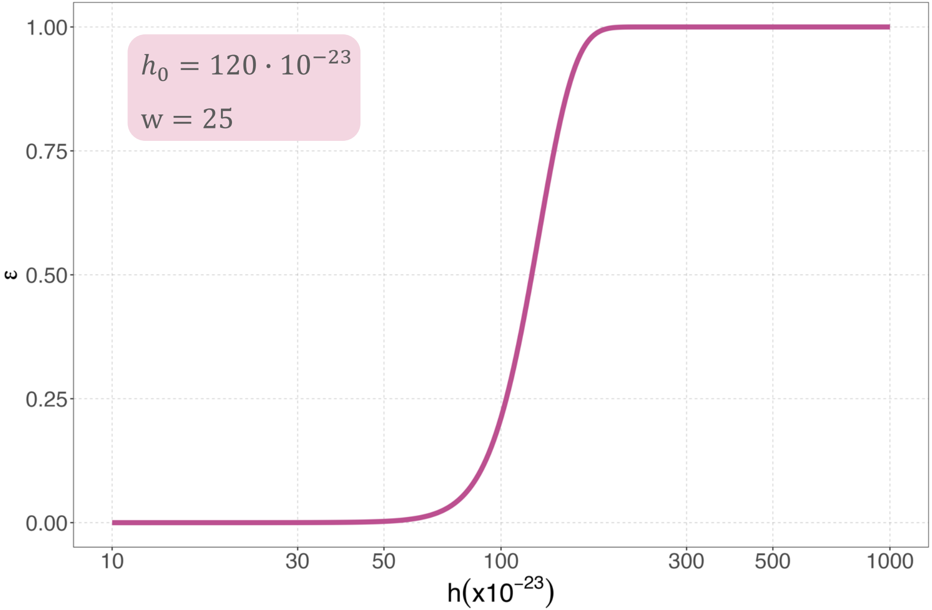
physical quantity measured by the interferometers: relative displacement between test masses during the passage of a GW

we consider a simple phenomenological model for the detection efficiency, characterized by two parameters h_0, w

$$\varepsilon(h) = \int_{-\infty}^{h} \frac{1}{\sqrt{2\pi w^2}} exp\left[-\frac{(h' - h_0)^2}{2w^2}\right] dh'$$

Then, the **number observed sources** results in $N_{obs} = N\varepsilon(h) = nTV\varepsilon(h)$

Detection Efficiency $\varepsilon(h)$



$\boldsymbol{h_0}$

strain produced by a source positioned at distance r_0 such that $\varepsilon(h_0) = 0.5$

$\boldsymbol{w}$

adjusts the sigmoid slope around h_0

$$w \ll 1 \leftrightarrow \frac{\partial \varepsilon}{\partial h}\Big|_{h_0} \gg 1$$

Detection Efficiency $\hat{\varepsilon}(r)$

For the purpose of the analysis, we ignore distinctions due to:

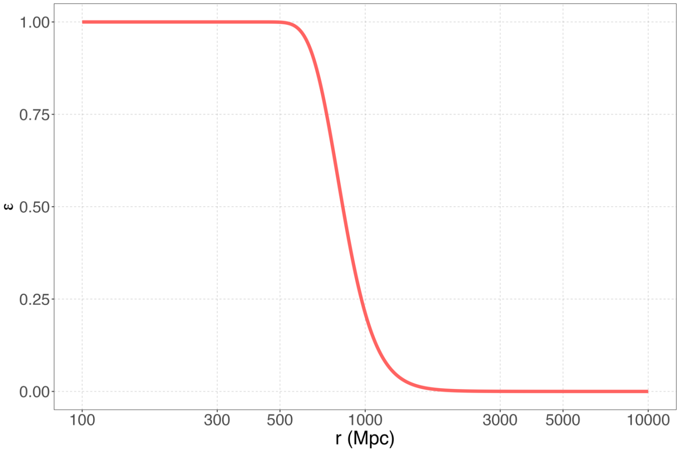
- sources **polarization**
- sources **orientation**

$h \propto \frac{1}{r}$

We also assume that the source **distance** is known with **no error**.

We can then express the detection efficiency as a function of the distance r :

$$\hat{\varepsilon}(r) = \varepsilon\big(r(h)\big) \propto \varepsilon\left(\frac{1}{h}\right)$$



Our Analysis

1 Statistical model for the number of sources N

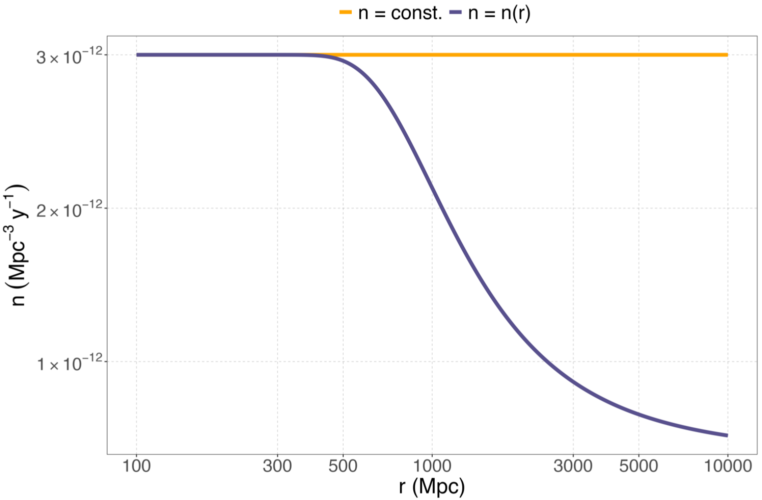
2 Simple detection efficiency model ε

3 **Simulated GW observations from two different population models**

4 Analysis of the population density n

5 n estimate dependencies on the observation time T and the source distance r

Models for n



$h_{0,1} = 67$
$w_1 = 60$
$n_0 = 3 \cdot 10^{-12}\ Mpc^{-3}s^{-1}$

1^{st} model: $n = const.$

the population density is constant over the source distance r:

$$\mathrm{n} = 3 \cdot 10^{-12}\ Mpc^{-3}s^{-1}$$

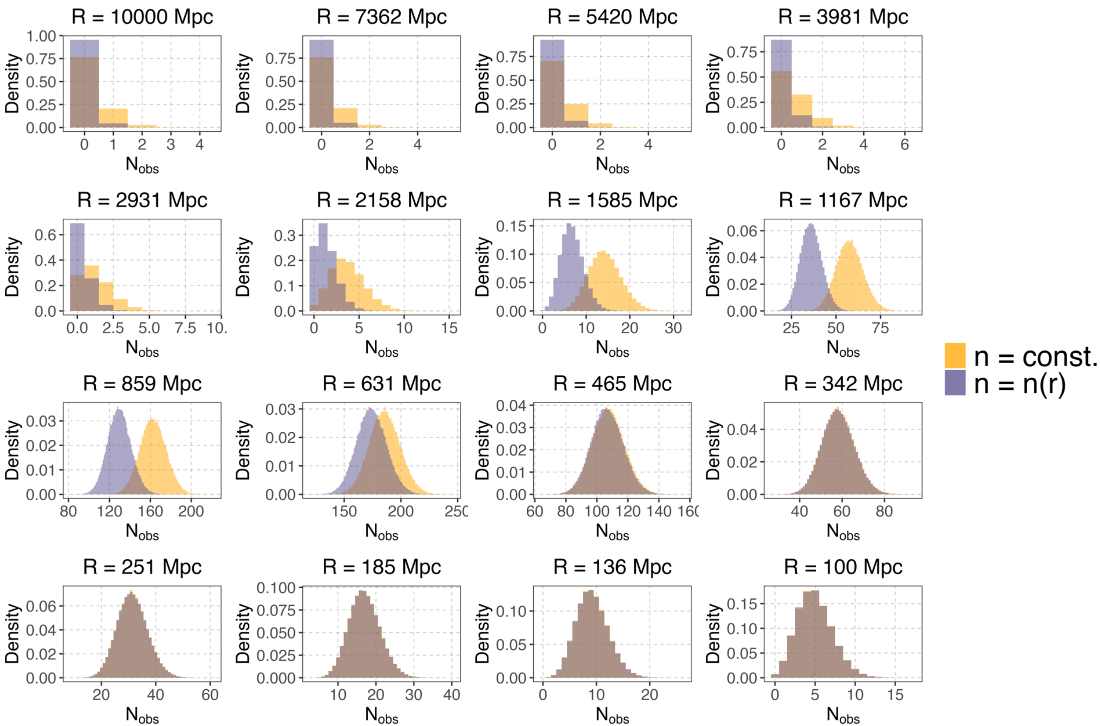
2^{nd} model: $n = n(r)$

the population density decreases with the source distance r following a sigmoid function:

$$n(h) = n_0 \int_{-\infty}^{h} \frac{1}{\sqrt{2\pi w_1^2}} exp\left[-\frac{\left(h' - h_{0,1}\right)^2}{2w_1^2}\right] dh'$$

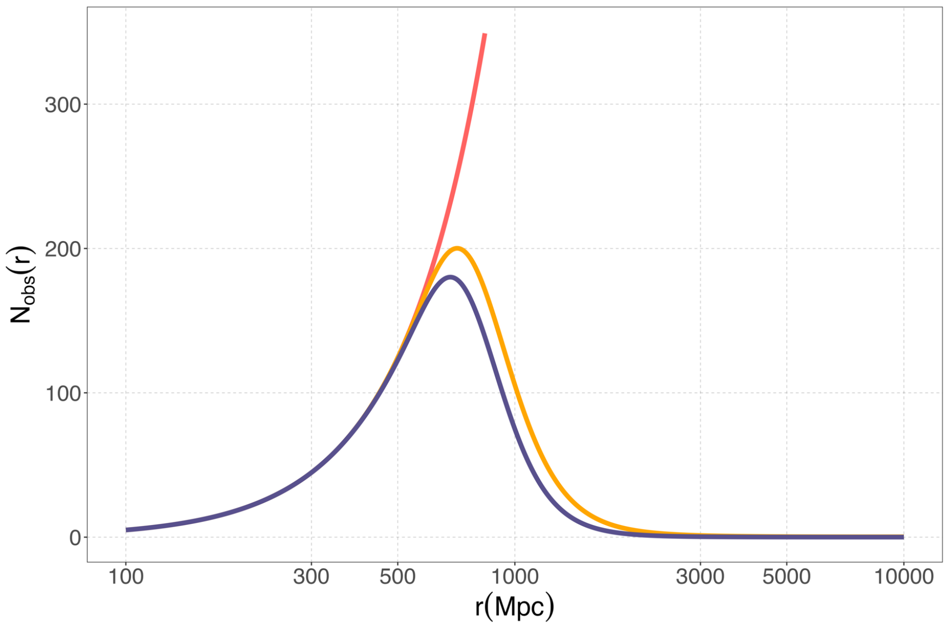
Simulations

We simulate observations from both models following a **Poisson distribution with rate** $\lambda = n(r)TV$ keeping **fixed** the observation time T but changing the shell volume V by **changing** the source distance r



N vs N_{obs}

We compare also the number of observation N_{obs} as a function of r expected from the two different models for n and with the **number of sources N** if the population density is constant



Our Analysis

1

Statistical model for the number of sources N

2

Simple detection efficiency model ε

3

Simulated GW observations from two different population models

4

Analysis of the population density n

5

n estimate dependencies on the observation time T and the source distance r

Bayesian inference on n

We now assume our simulated data of N_{obs} as **real data**

we perform Bayesian inference to find an estimate of $n(r)$ from the data

as the source distance r varies, for both datasets

Bayes' theorem $$P(\lambda|\{N\}) \propto f(\{N\}|\lambda) \times g(\lambda)$$

- Prior on n: $Gamma(1, 0)$
- Likelihood: $\prod Pois(\lambda = nTV)$
- Posterior: $Gamma(\alpha_{post}, \beta_{post})$

$$\alpha_{post} = 1 + \sum_{i=1}^{n_{sample}} N_{obs,i}$$

$$\beta_{post} = 0 + n_{sample}$$

Bayesian inference on n

We now assume our simulated data of N_{obs} as **real data**

we perform Bayesian inference to find an estimate of $n(r)$ from the data

as the source distance r varies, for both datasets

Bayes' theorem $P(\lambda|\{N\}) \propto f(\{N\}|\lambda) \times g(\lambda)$

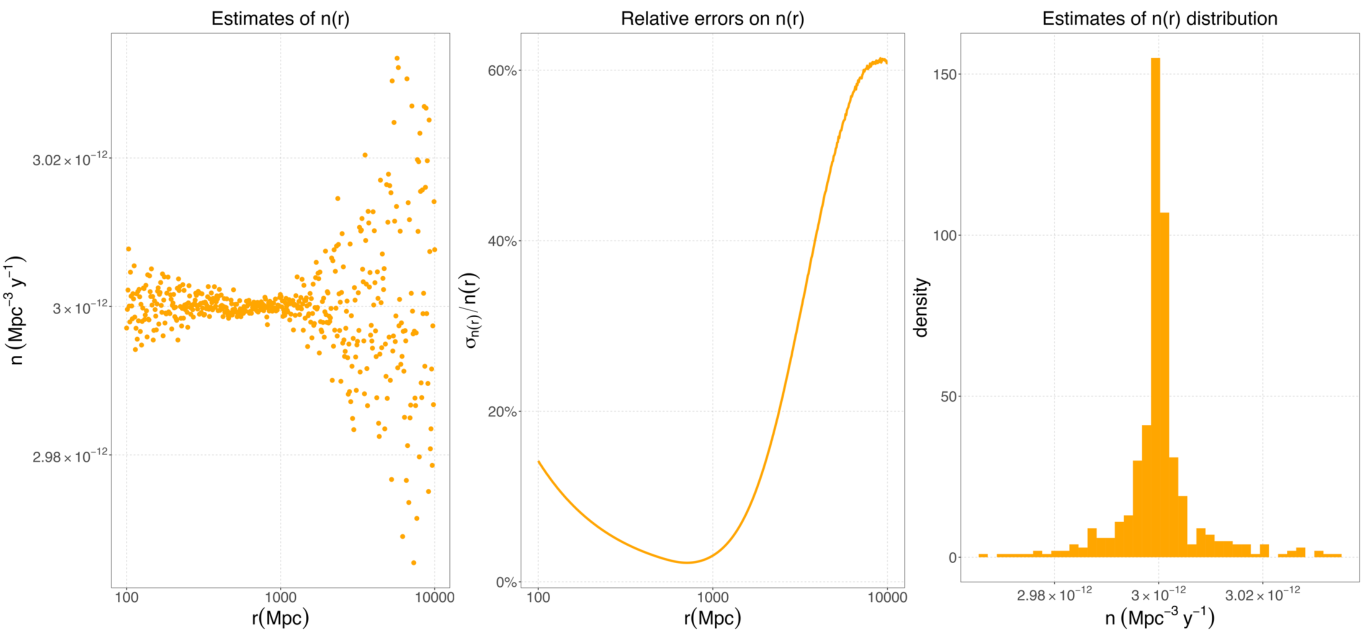
We first infere on $\lambda = \widehat{N}_{obs}$ and then we find an estimate and a reliability for $n(r)$ as

$$n(r) = \frac{\widehat{N}_{obs}}{4\pi r^2 \Delta r T \varepsilon(r)} \quad \text{and} \quad \sigma_{n(r)} = \frac{\sigma_{\widehat{N}_{obs}}}{4\pi r^2 \Delta r T \varepsilon(r)}$$

where $\widehat{N}_{obs}$ and $\sigma_{\widehat{N}_{obs}}$ are respectively our best estimate and the reliability for N_{obs}

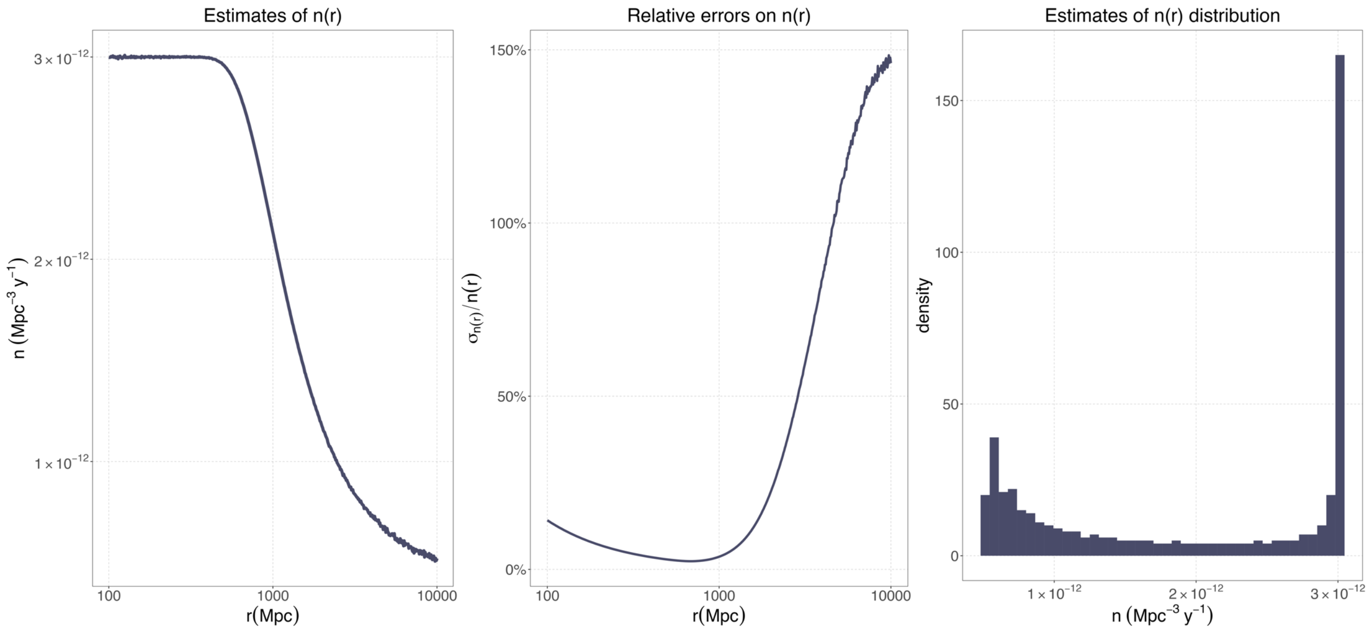
Bayesian inference on n

1^{st} dataset – $n = const.$



Bayesian inference on n

2^{nd} dataset – $n = n(r)$



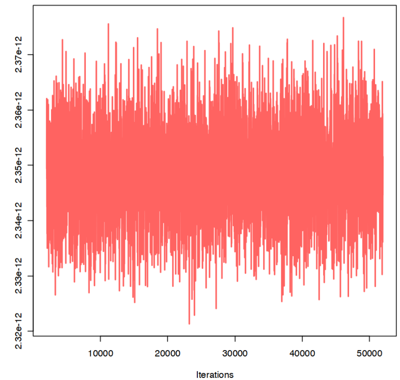
Bayesian inference on $n(r)$ with JAGS

We now simulate data at $r = 10^3\ Mpc$ from the **second model** and we assume the data as real: we want to find the value for the population density n

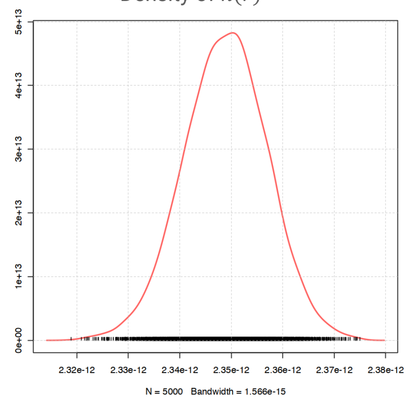
we sample from the posterior distribution of N_{obs} with a MCMC using **JAGS** and we perform Bayesian inference on $n(r)$

we find $n = (2.349 \cdot 10^{-12} \pm 8 \cdot 10^{-15})\ Mpc^{-3} s^{-1}$

Trace of $n(r)$



Density of $n(r)$

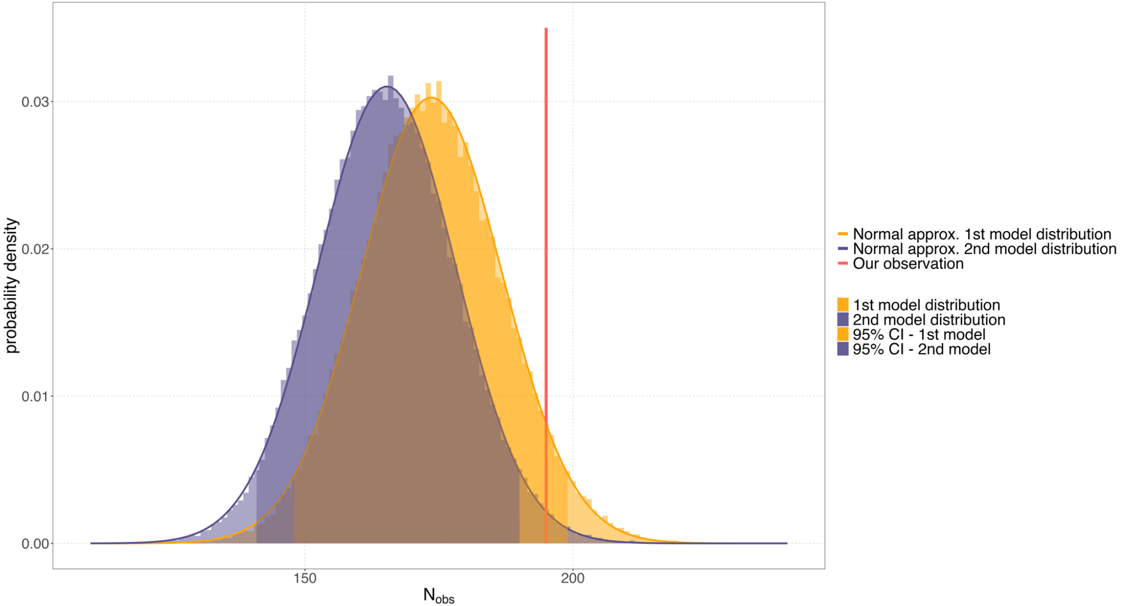


Bayesian HT

We now imagine to observe $N_{obs} = 195$ sources at $r = 600\ Mpc$:

we want to know if the correspondent population density belongs to the first or to the second model

Bayesian HT at $\alpha = 5\%$ level of significance



Our Analysis

1

Statistical model for the number of sources N

2

Simple detection efficiency model ε

3

Simulated GW observations from two different population models

4

Analysis of the population density n

5

n estimate dependencies on the observation time T and the source distance r

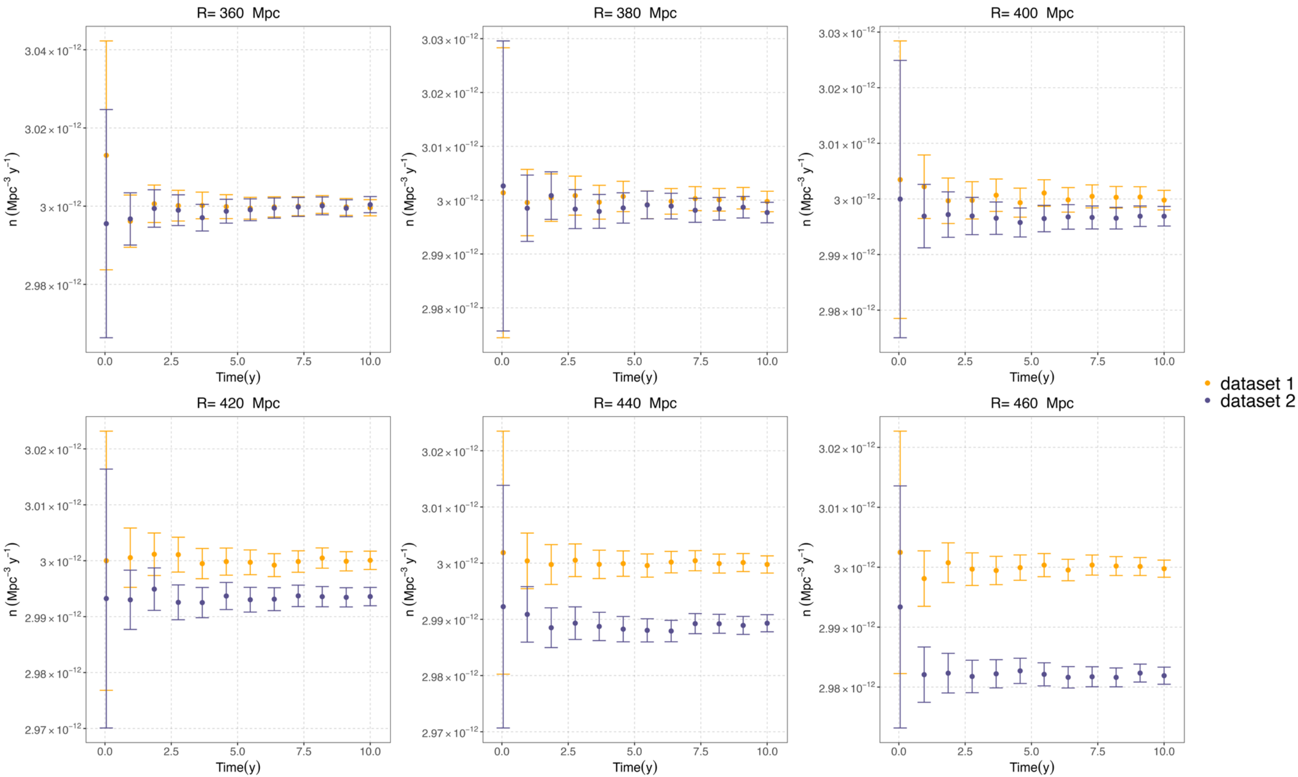
n estimates dependencies from T and r

We want to understand in which range of observation time T we can discriminate the population density models, also with respect to the detection efficiency volume

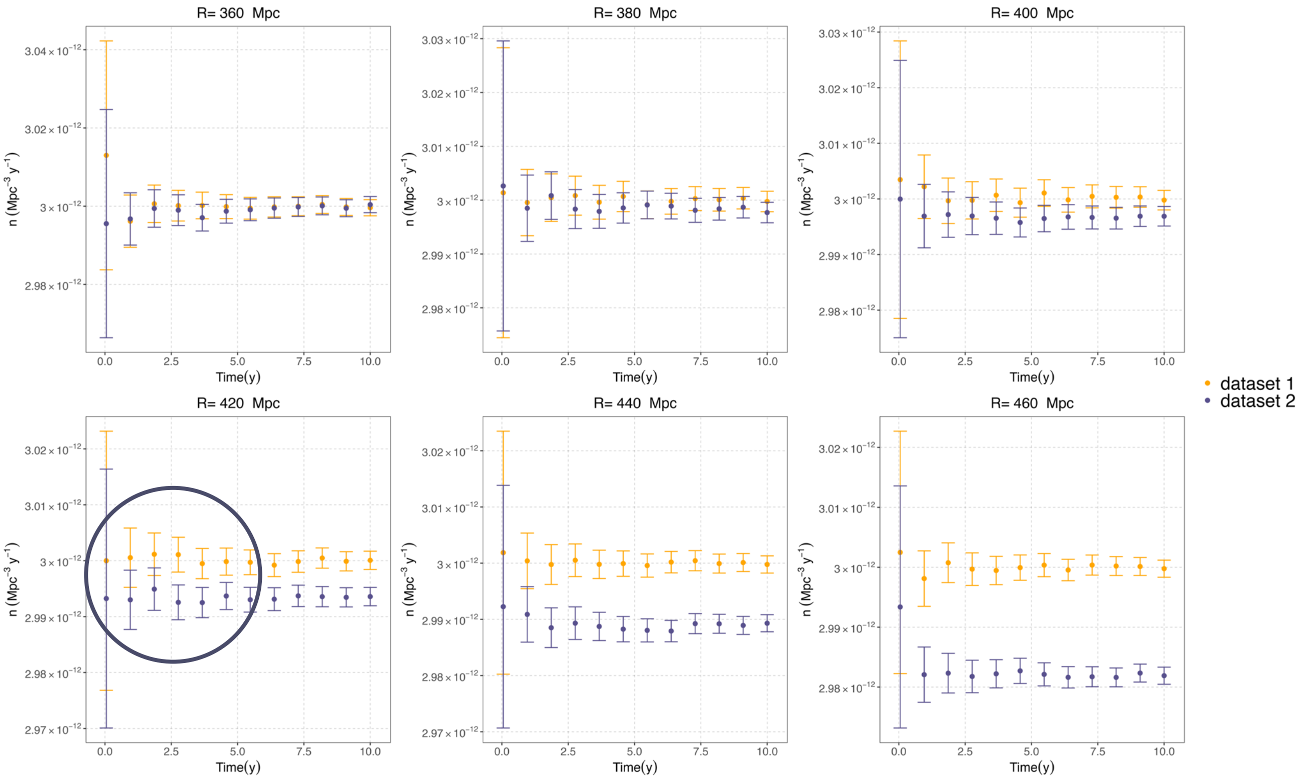
we perform the same Bayesian analysis as before, also considering the observation time T to study how it affects the estimates for n

we compare n estimates to see if they are compatible at 5σ

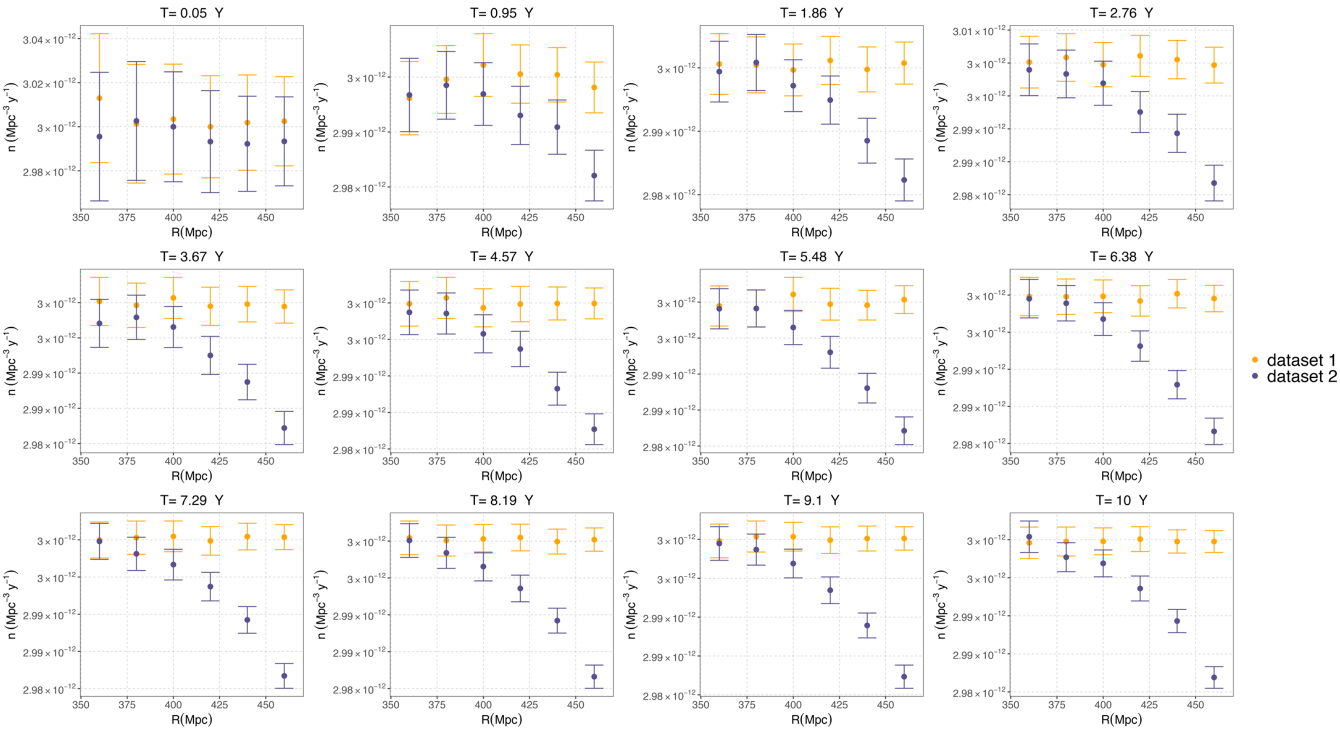
n estimates dependencies from T and r



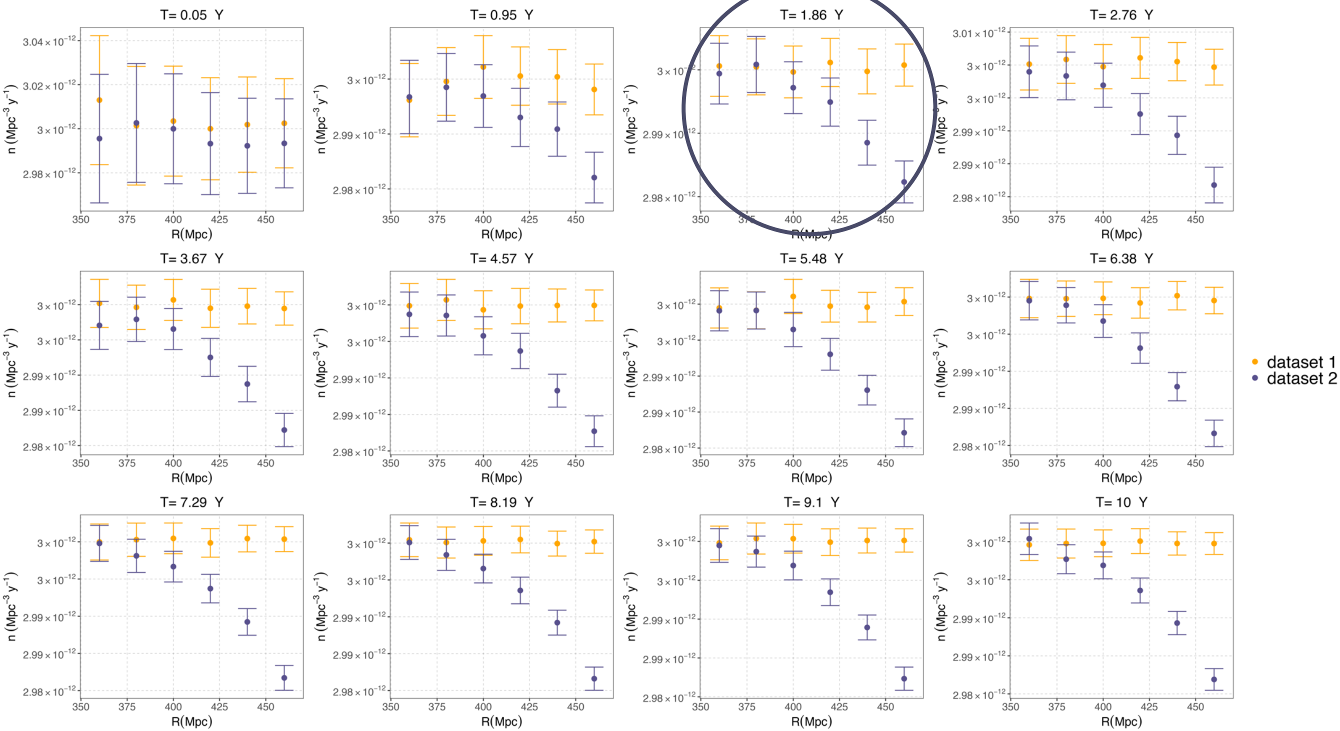
n estimates dependencies from T and r



n estimates dependencies from T and r



n estimates dependencies from T and r



Conclusions

- **Introduction** to the real analysis done by the LIGO and Virgo collaboration

- Analysis of the **distribution of GW sources** with a simplified model for the detection efficiency

- Focus on **Bayesian inference** to study the sources population density

THANK YOU !