Population Analysis of Massive Black hole Binaries with LISA using Iterative weighted KDE

Jam Sadiq

In collaboration with Enrico Barausse (SISSA, Italy), Thomas Dent (IGFAE, Spain)

LISA Spain Meeting 2024

October 16-17, 2024

Institute of Space Sciences (ICE, CSIC and IEEC)

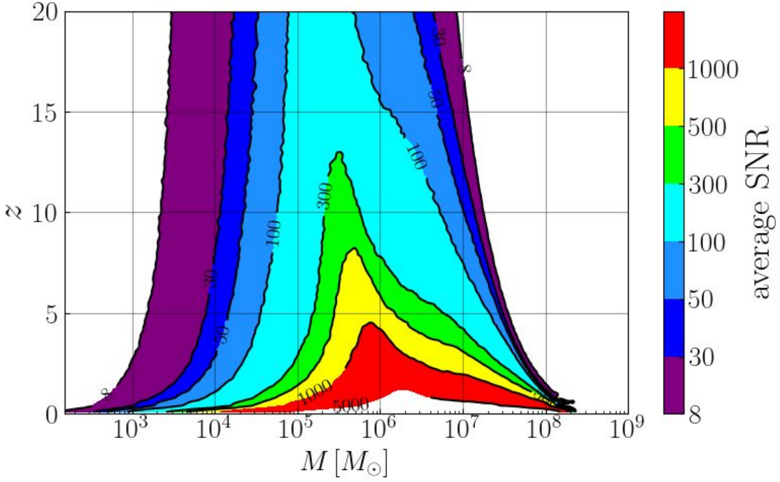
UAB, Barcelona, Spain

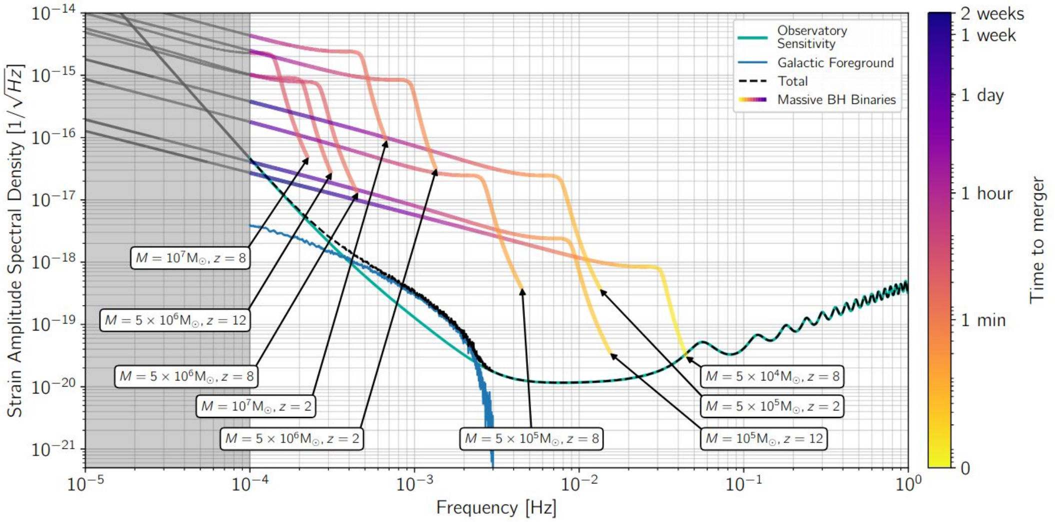
Overview

- Motivation
- Method of KDE with Iterative Reweighting technique
- Applications on Massive BHBs evolve from Pop III star (“light seed”) for LISA
- Future Work

Motivation

- LISA will see Massive Black hole Binaries (MBHBs), crucial for true understanding of formation, evolution of such systems as well as Galaxy evolution

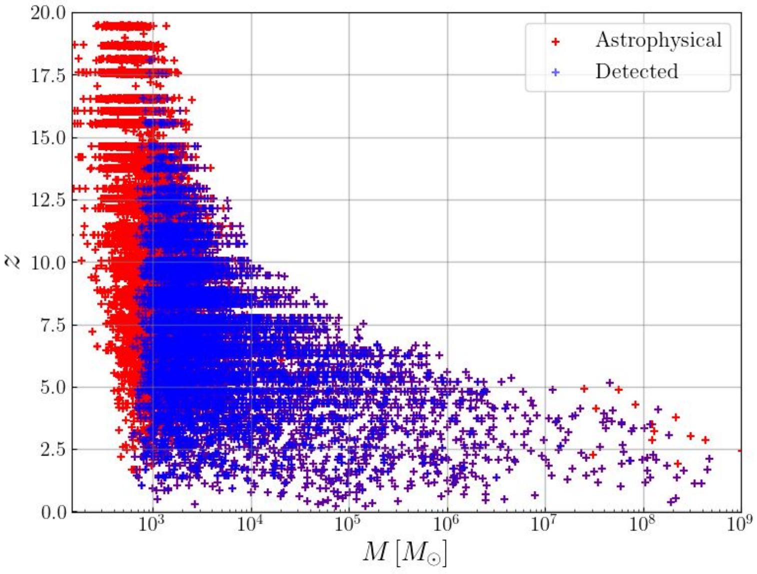




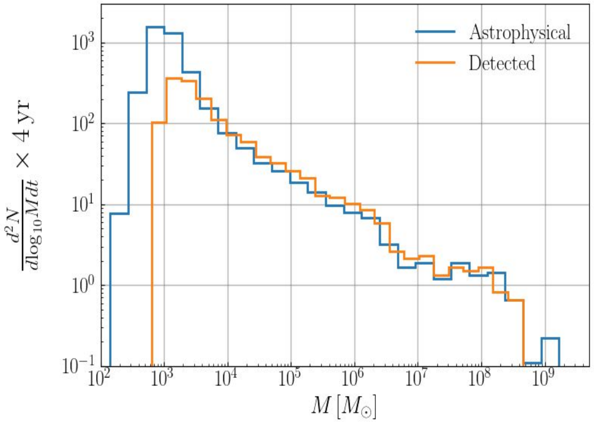
Credit: ESA-SCI-DIR-RP-002, Sep, 2023

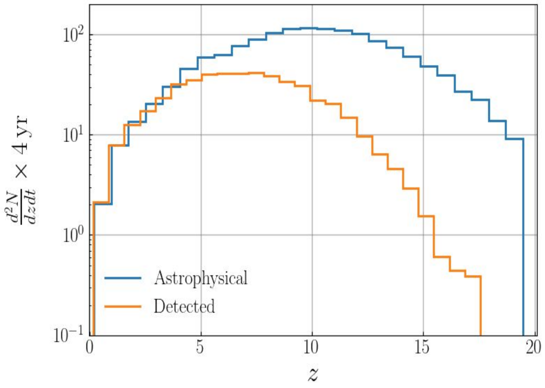
Pop-III ("light seed") Remnant Binary Black holes

- Low Mass High Redshift events
- Selection effects are crucial



Data from Phys. Rev. D 93, 024003 (2016)





Our Goal: Non parametric Method with Selection Effects to Reconstruct Astrophysical Rates

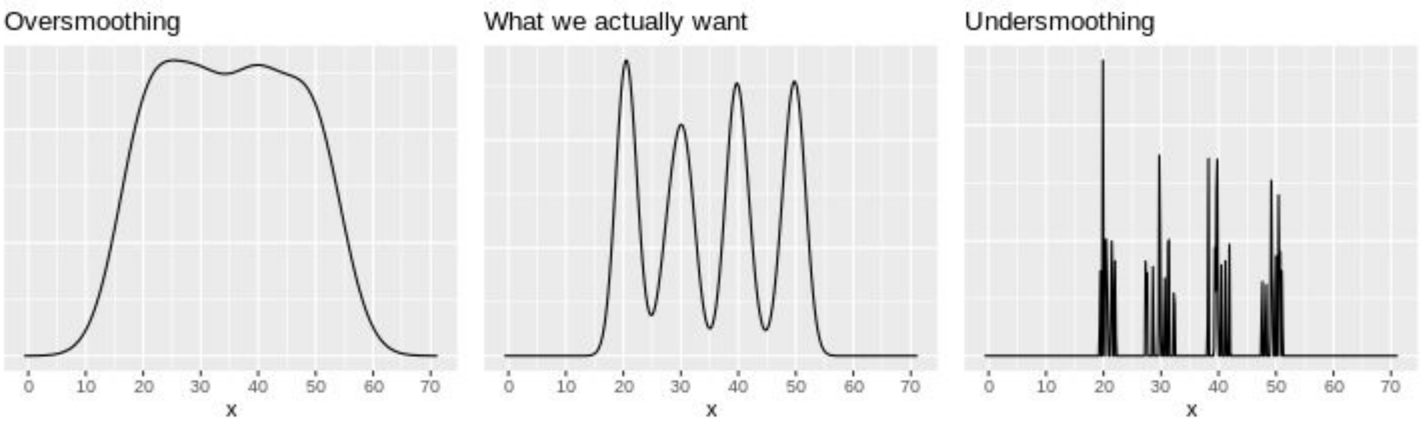
- *Phys.Rev.D* 105 (2022)
- *Astrophys.J.* 960 (2024)

Kernel Density Estimation (KDE)

- **Objective:**
  - Flexible, and, non parametric method
  - Prioritize computational simplicity and speed
  - Go from a set of detected events to a rate density with small/quantified uncertainty
  - Addresses same questions as other methods
    - stellar evolution, Black hole and Binary formation channel

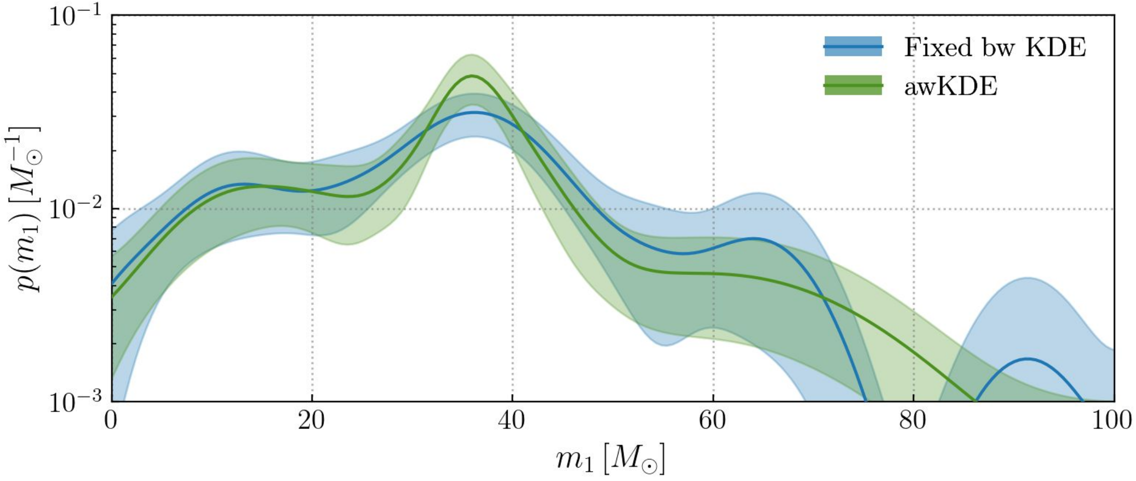
Kernel Density Estimation (KDE)

- Put a blob of density at each event. Ensures data have "fairly high" likelihood
  - but what size of blob (bandwidth choice)?



Adaptive Width Kernel Density Estimation (awKDE)

- Under-smoothing makes estimate biased Over-smoothing makes estimate (too) uncertain
- Optimized adaptive bandwidth: Adjust local bw ∝ to typical distance between events



Adaptive Width Kernel Density Estimation (awKDE)

- Using an initial global bandwidth choice to make a pilot Gaussian KDE

$$\hat{f}(x) = \frac{1}{\sum_i W_i} \sum_{i=1}^{n} \frac{W_i}{h\lambda_i} K\left(\frac{x - X_i}{h\lambda_i}\right). \qquad \lambda_i = 1 \quad W_i = 1$$

- Derive local (adaptive) bandwidth from pilot KDE & sensitivity parameter, $0 \leq \alpha \leq 1$

$$\lambda_i = \left(\frac{\hat{f}_0(X_i)}{g}\right)^{-\alpha}, \quad \log g = n^{-1} \sum_{i=1}^{n} \log \hat{f}_0(X_i)$$

- Final KDE uses this local bandwidth
- For multi-dimensional case linearly transform data to have zero mean and unit covariance

Choice of KDE Hyper-parameters

- Determine optimum initial bandwidth and α by grid search using maximum likelihood as a figure of merit on leave-one-out cross validation with

$$\log \mathcal{L}_{LOO} = \sum_{i=1}^{n} \log \hat{f}_{LOO,i}(X_i)$$

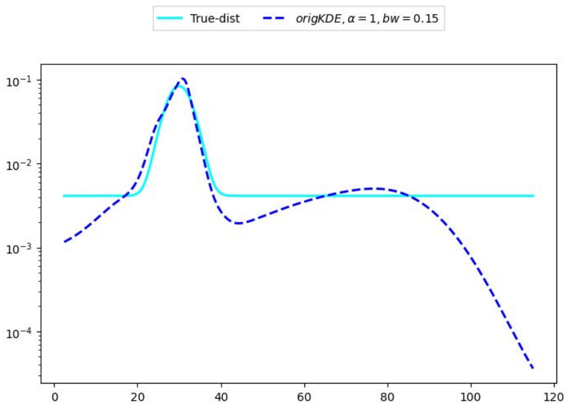
where $\hat{f}_{LOO,i}$ is the KDE constructed from all samples except X_i

- Linear in logarithm of estimate at observed values and will penalize relative errors

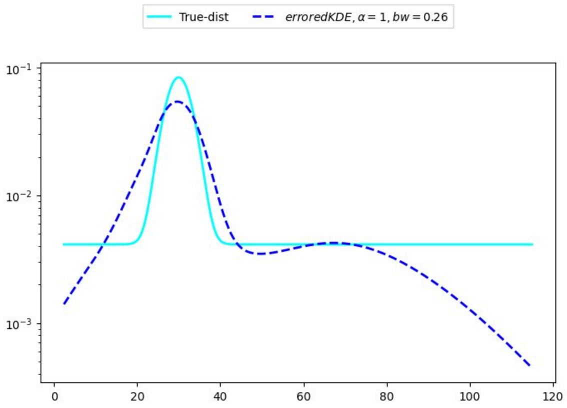
- For large number of samples we use k-fold cross validation with 5 folds with same figure of merit

Parameter Estimation Uncertainties

- Event parameters can't be measured exactly
- Posterior samples are only accurate under default PE-prior population
- Expected to over-disperse population feature



Data without uncertainty



Data with uncertainty

'Deconvolving' PE uncertainty via Reweighting

- Accurate population model requires accurate event values

- Accurate estimate of event value requires population model

- Address by iteration : use previous rate estimate for reweighting PE samples

Expectation-Maximization Algorithm

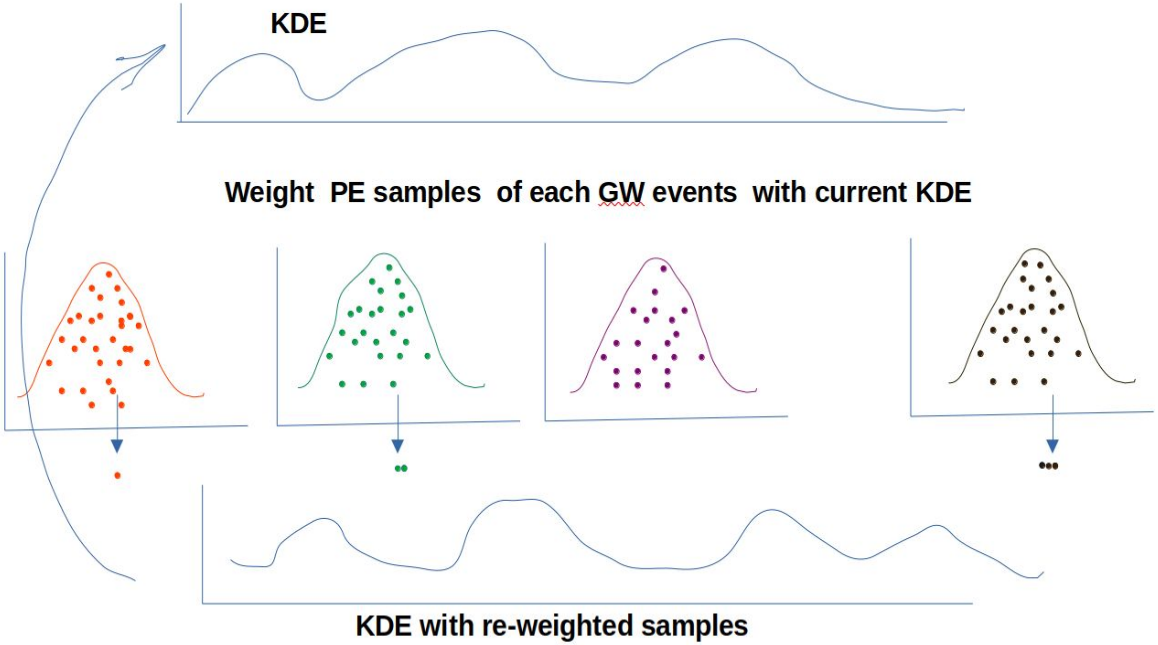
Astrophys.J. 960 (2024)

1. Draw Poisson(1) PE samples per event weighted by current estimate of population rate density

2. Optimize an adaptive KDE trained with this sample set

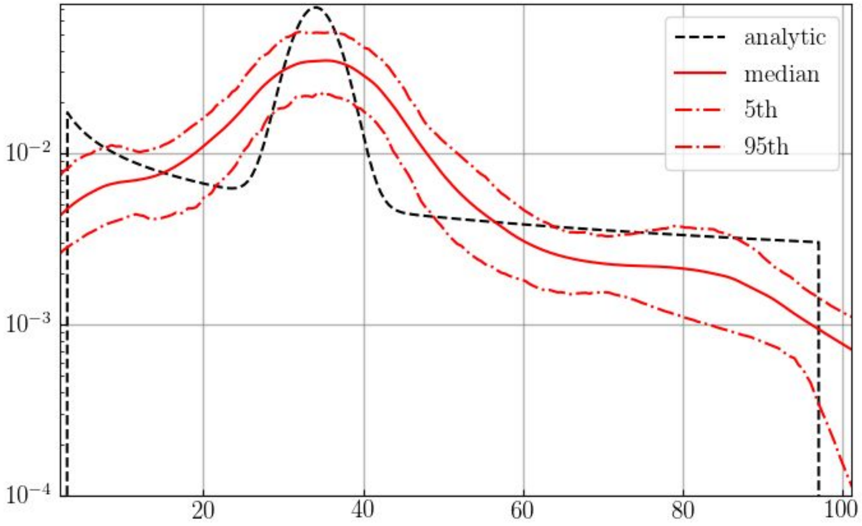
3. Update current rate estimate using the KDE

4. Go to step 1

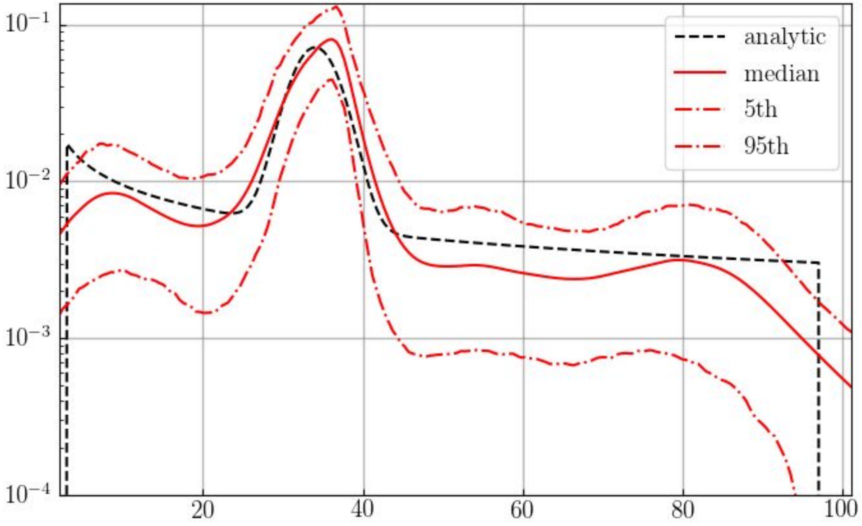


Test: Expectation-Maximization Algorithm

- 50% of events in Gaussian peak, peak σ = 3, PE error & sample uncertainty 5



Without iterative reweighting



With iterative reweighting

Selection Effects

- To get the rates from density estimate, account for selection bias

$$p_{\rm det}(x,\theta) = \int p(\rho > \rho_{\rm th}|\bar{\rho}(x,\theta,\psi))p(\psi)\, d^n\psi\,,$$

- KDE parameters $x:\ (M, z)$
- Other intrinsic parameters $\theta:\ (q,\ \chi, \dots)$
- Uniform distribution over extrinsic parameters and time of observation $p(\psi)$

Rate Estimates including Selection effects

- Astrophysical rate density

$$R_{\rm det}(x) = \int R(x,\theta) p_{\rm det}(x,\theta)\, d^n\theta\,.$$

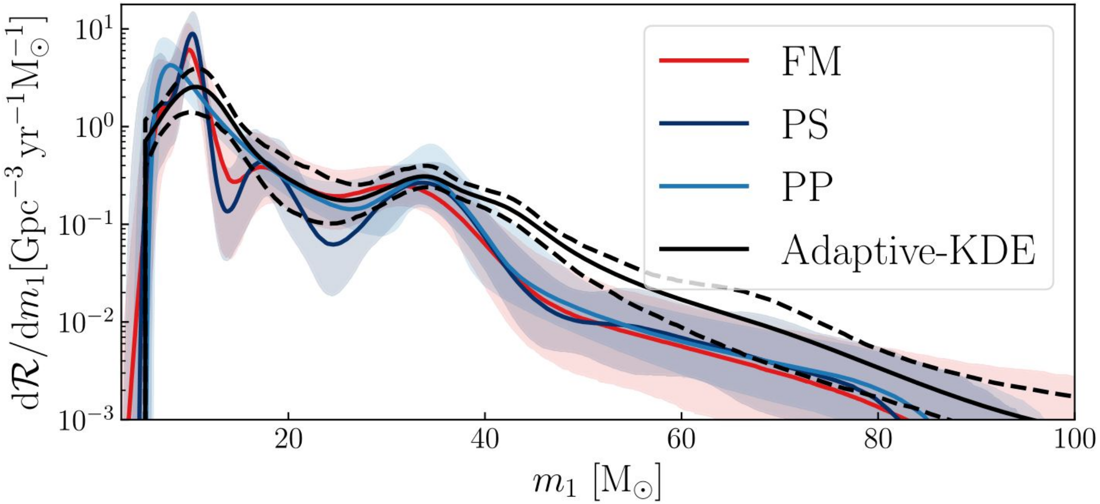
$$R(x,\theta) = R(x) p_{\rm pop}(\theta)$$

$$\begin{aligned} R_{\rm det}(x) &= R(x) \int p_{\rm pop}(\theta) p_{\rm det}(x,\theta)\, d^n\theta \\ &\equiv R(x) p_{\rm det}(x; p_{\rm pop})\,, \end{aligned}$$

$$R(x) = \frac{R_{\rm det}(x)}{p_{\rm det}(x; p_{\rm pop})}$$

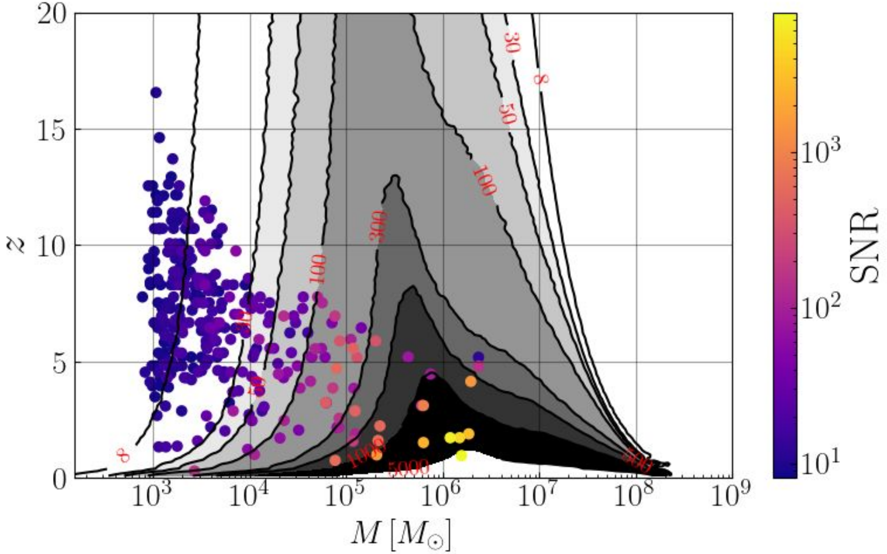
Application on GWs

- Applied on GWTC-3



- *Phys.Rev.D* 105 (2022)
- *Astrophys.J.* 960 (2024)

Data: light-seed model popIII (K+16)



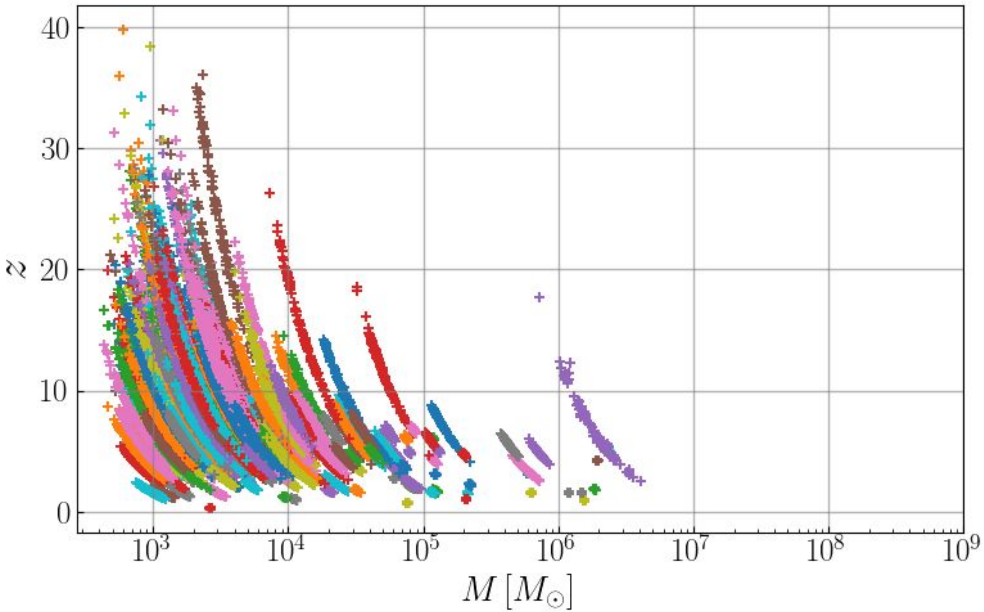
Data from Phys. Rev. D 93, 024003 (2016)

Parameter Estimation

Light-seed model: popIII-d (K+16)

- Posterior samples (100 per event)
- Included weak lensing effects

$$\sigma_{\rm wl}(z) = D_L \times 0.066 \left(\frac{1-(1+z)^{-0.25}}{0.25} \right)^{1.8}$$



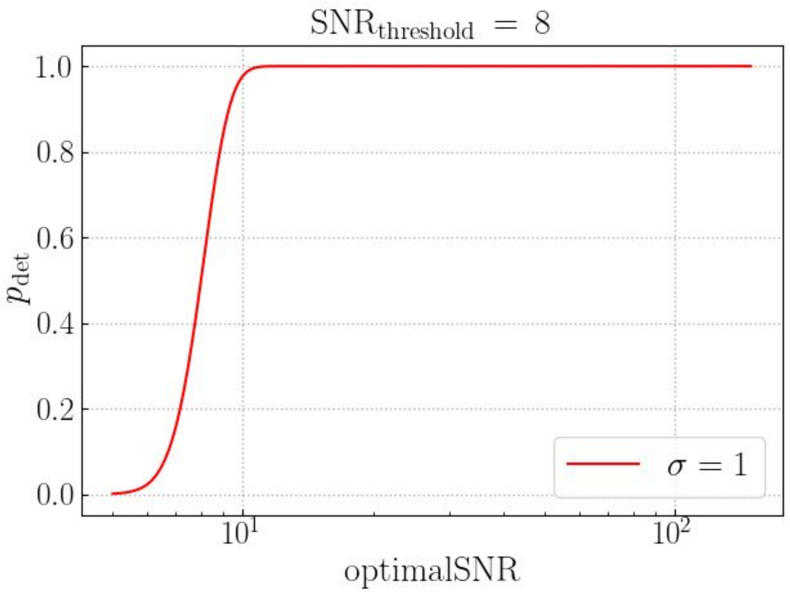
Phys. Rev. D.108 103034 (2023)

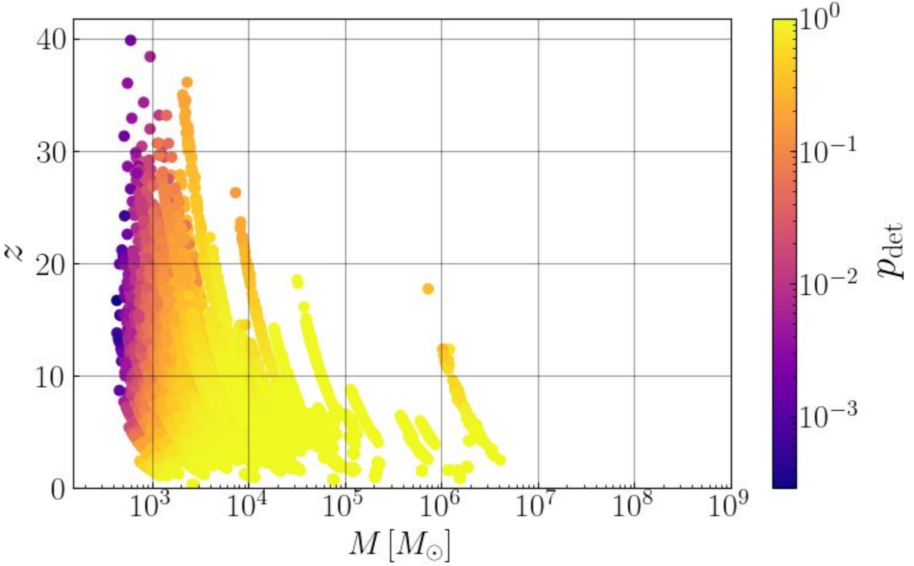
Selection Effects

- Matched-filter SNR (standard Gaussian noise) to get Pdet

$$p(\rho|\bar{\rho})d\rho = \frac{1}{\sqrt{2\pi}} e^{-(\rho-\bar{\rho})^2/2}\, d\rho\,.$$

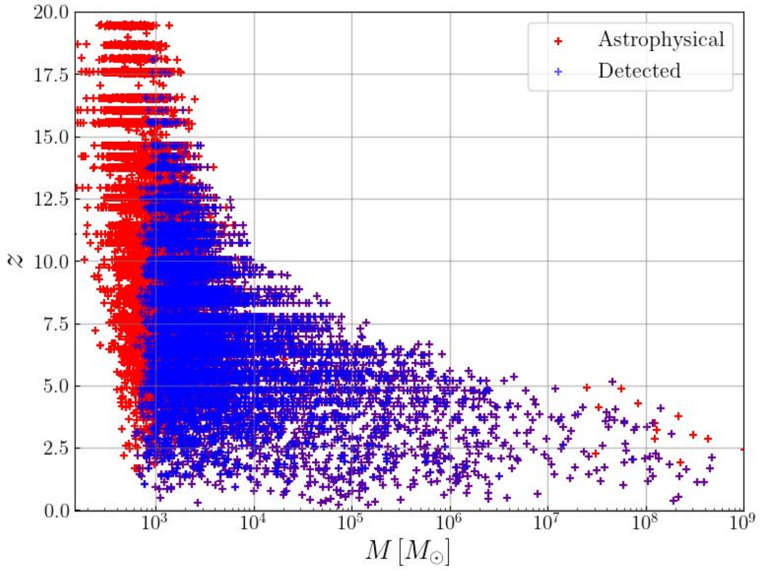
cap: pdet < 1e-3



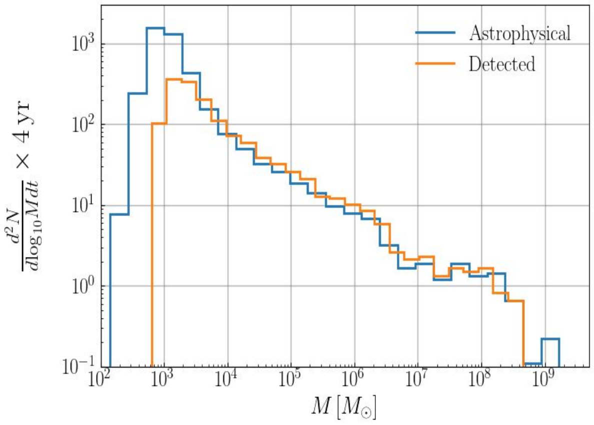


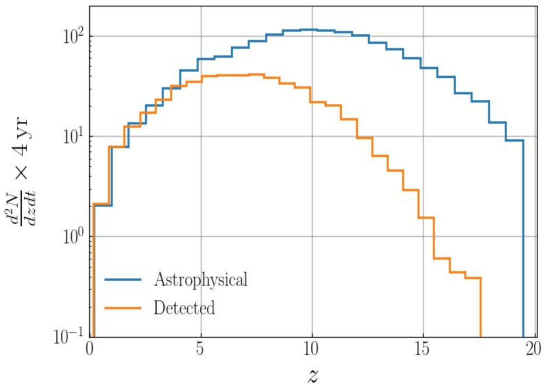
Testing methods: Model

Light-seed model: popIII-d (K+16)



From Phys. Rev. D 93, 024003 (2016)

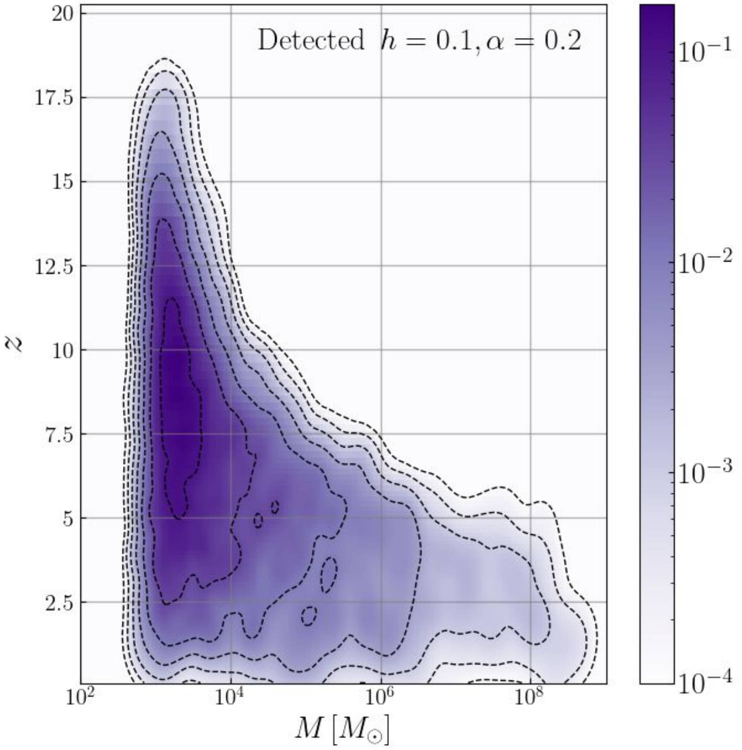




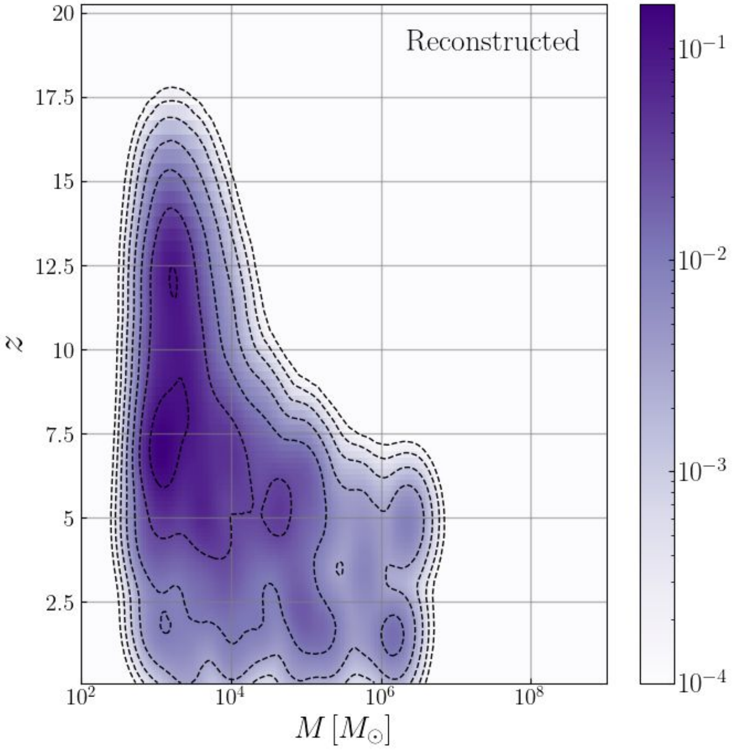
Results: KDE reconstruction

Without applying selection effects

Model

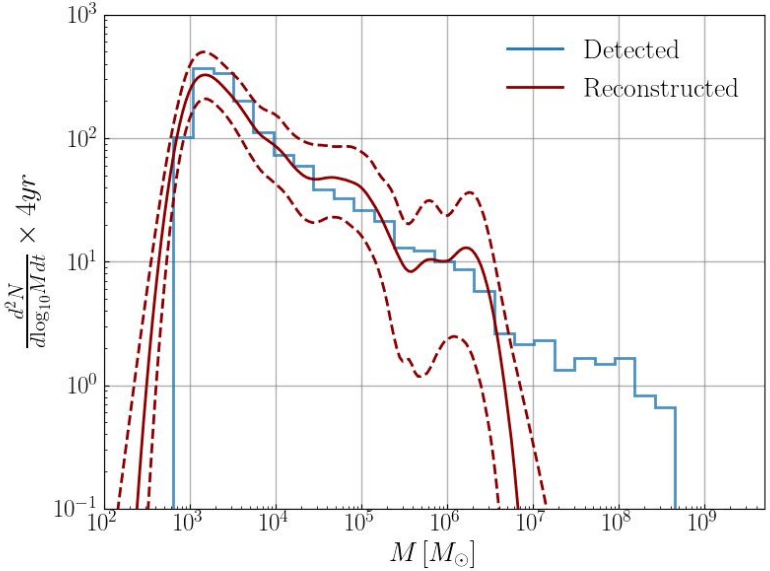


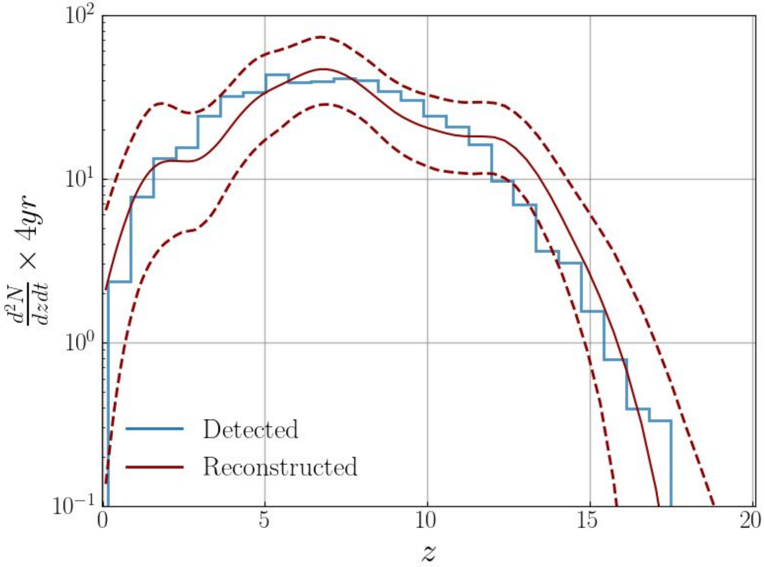
Posteriors



Results: Rates without Selection effects

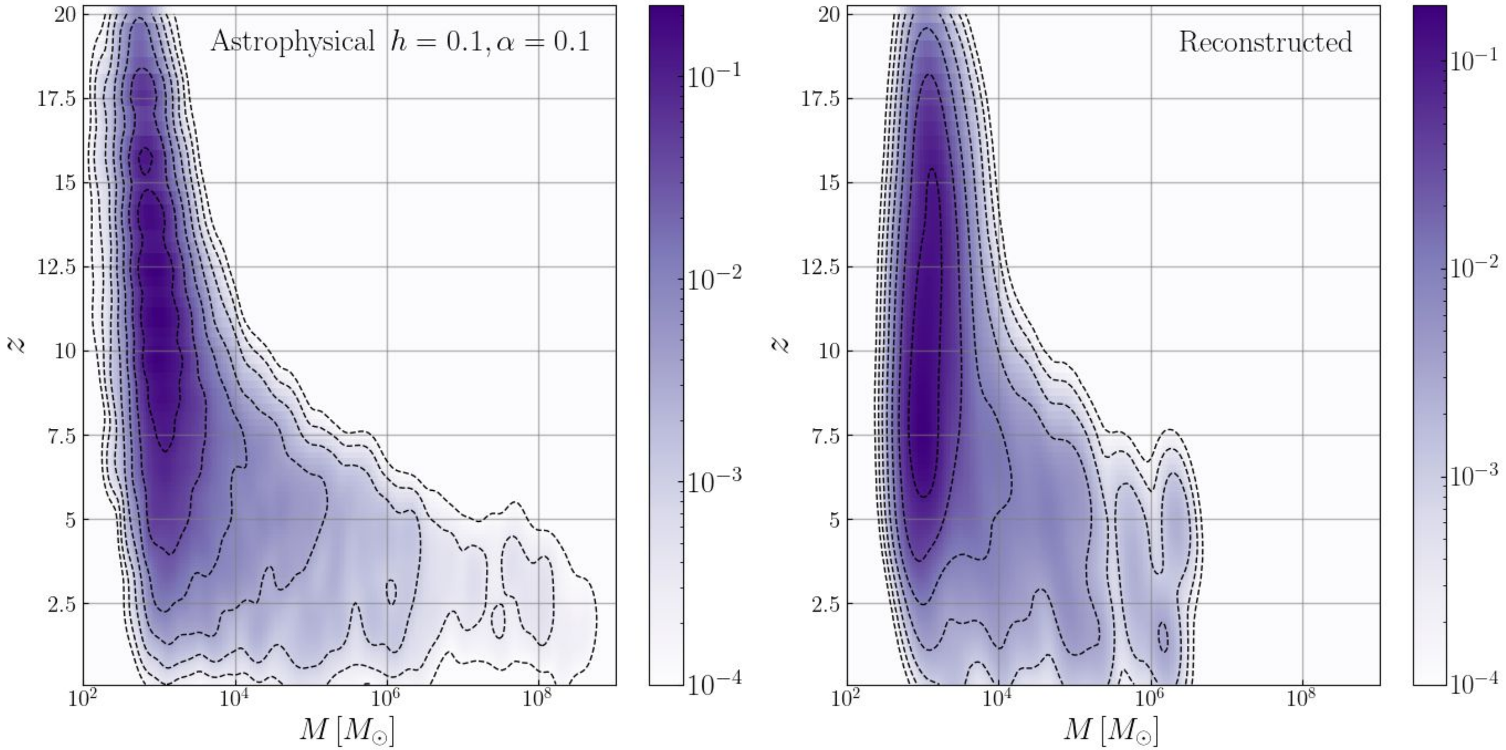
Rate = KDE*Number of events



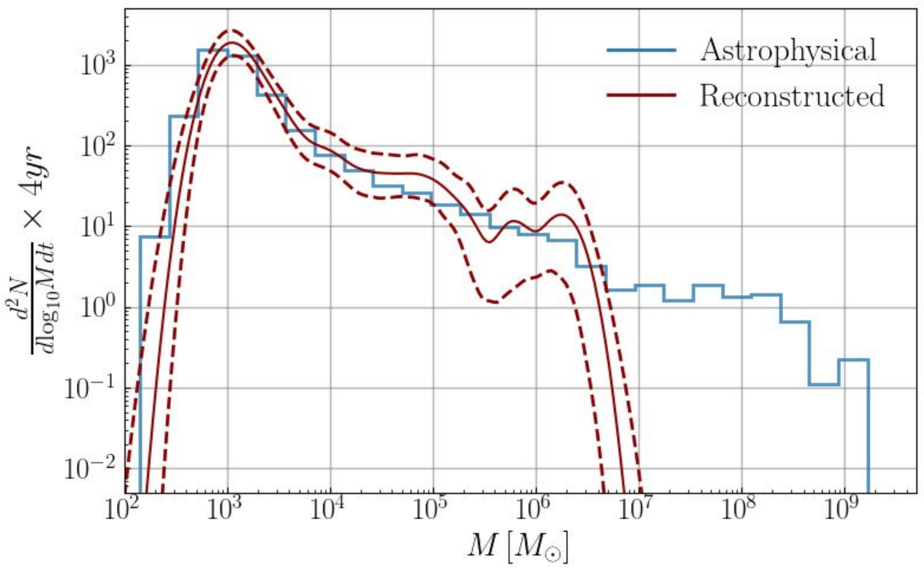


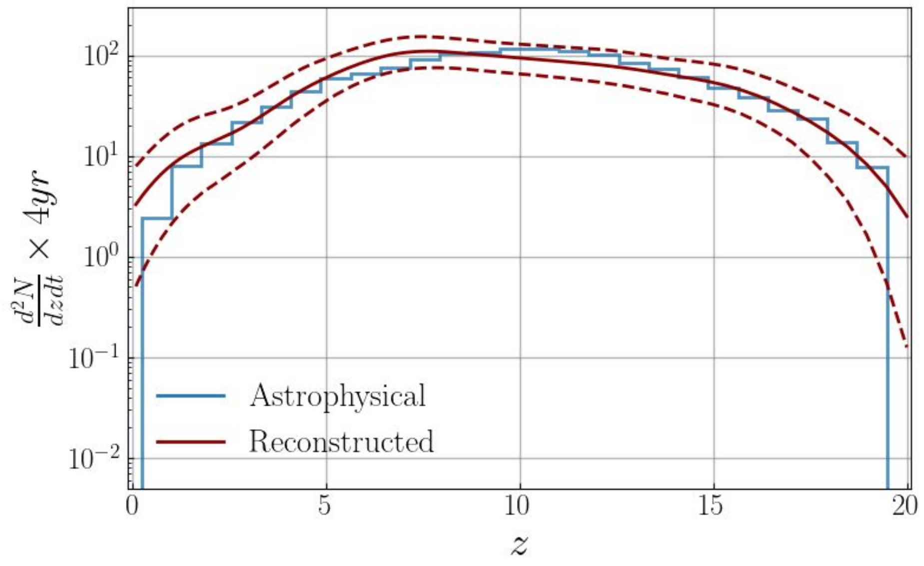
Results: KDE reconstruction

Applying selection effects: weights = 1/pdet

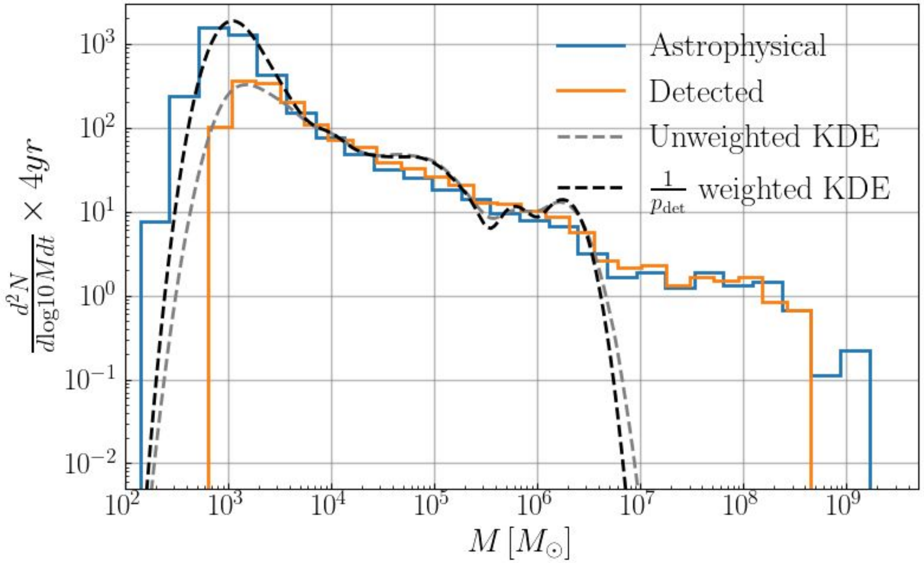


Results: Rates from KDE with Selection Effects

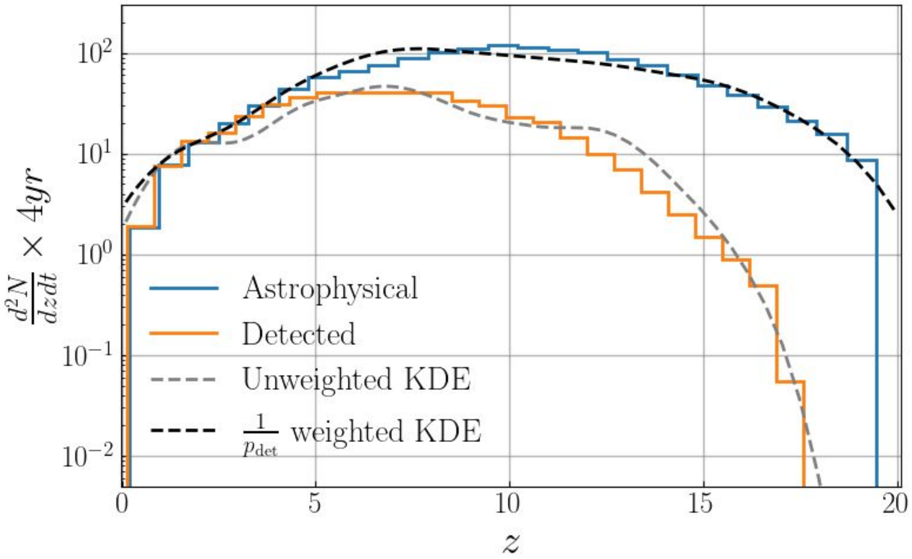




Results



Astrophysical
Detected
Unweighted KDE
$\frac{1}{p_{\rm det}}$ weighted KDE
$\frac{d^2N}{d\log 10 M dt} \times 4yr$
$M\,[M_\odot]$



Astrophysical
Detected
Unweighted KDE
$\frac{1}{p_{\rm det}}$ weighted KDE
$\frac{d^2N}{dzdt} \times 4yr$
z

Future: Test Adaptive KDE with Selection Effects For other light seed models

Quick Recap

1) KDE as non parametric model for population analysis
   a) Adaptive KDE without selection effects
   b) Weighted KDE with selection effect

2) Iterative Reweighting to reduce PE uncertainty

3) Fast, Flexible method for quick Population Analysis

4) Future: Weighted Adaptive KDE fixing selection effects for small values

Questions!

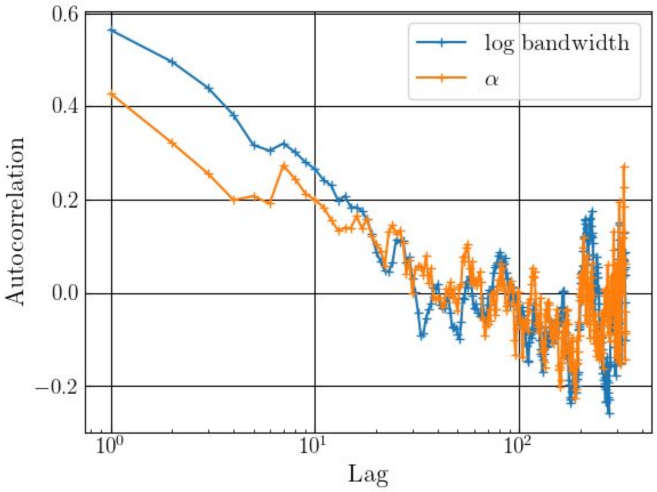
Future Work

- Extension of multidimensional iterative awKDE including component spins, mass-ratio
  - Technical issue in optimizing the Gaussian kernel for a multidimensional data set, where it will not be appropriate (or even meaningful, given the different units) to impose equal variances over different parameters as we currently do for (log) m1 and m2
  - For more than two dimensions a grid search may not be practicable; more sophisticated methods may be required in order to realize the potential of iterative KDE over a full set of population parameters

Autocorrelation of KDE ‘hyper’parameters

- KDE has global bandwidth (bw) and adaptive (α) parameters
- Optimized via CV max likelihood at each iteration
- Monitor evolution to characterize the process

- Autocorrelation close to 0 after ~30 iterations



Rates and Population Analysis

Why:

- From features in population study, understand the astrophysical origin of detected events
- Improve theoretical models, learn about outliers ,....

How:

<u>Parametric Method</u>:	**<u>Non Parametric Method</u>:**
Have functional form Good: extrapolation, robust even with few events, Bayes factor, Interpretable Issues: Bound to specific functional form, New features to be added by hand	Good: Data based, flexible, can be computationally efficient, provide insights for parametric models Issues: no extrapolation, no Bayes factor, no direct connection with physics

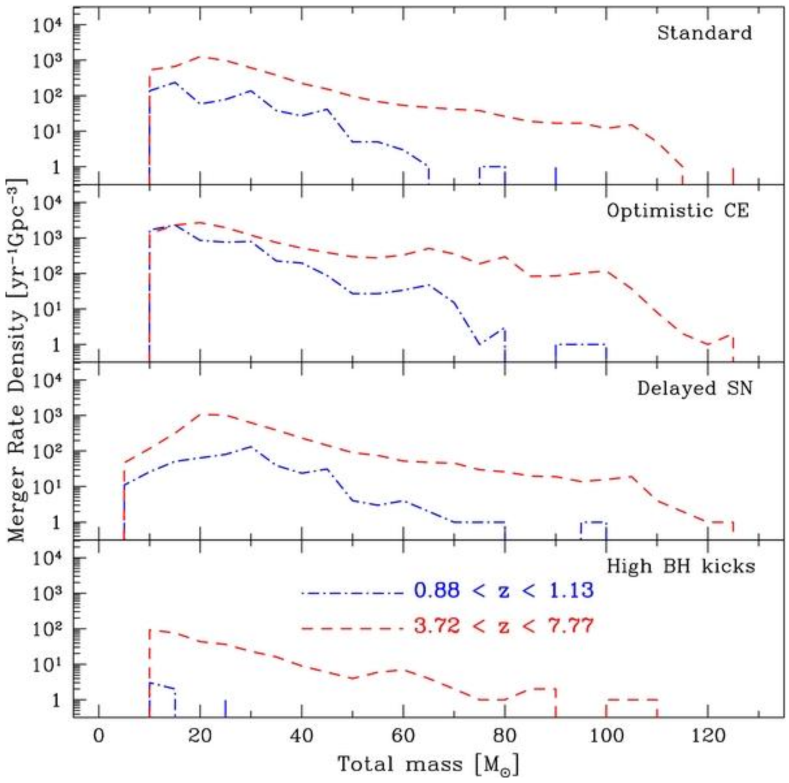
Population Study

- Population analysis to understand individual events
    - New events exciting as an outlier in population

- From features in population study, understand the astrophysical origin of detected events
    - Stellar evolution, formation channel, cosmology

Astrophysical Models vs Gravitational Wave Detections

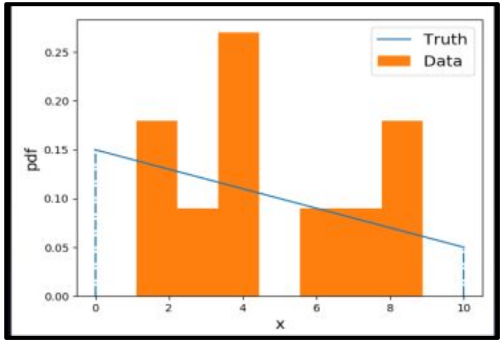
- Astrophysics modelling
  ⇒ expected merger distribution over redshift, masses, spins ...
- Models do not predict individual merger parameters
- GW detections ⇒distribution "samples"

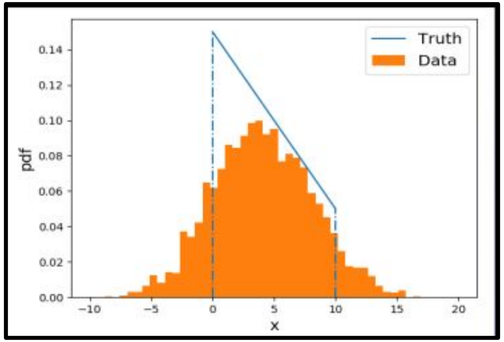


Dominik et al. Astrophys.J. 779 (2013) 72

Hazards of Gravitational Wave Population Analysis

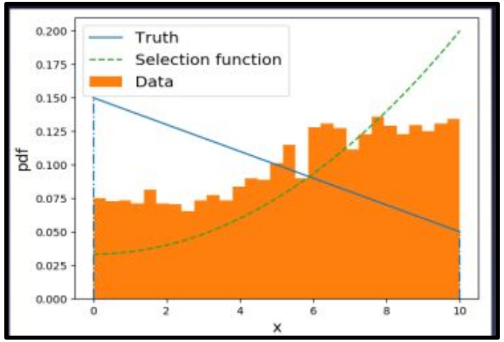
Low event Statistics

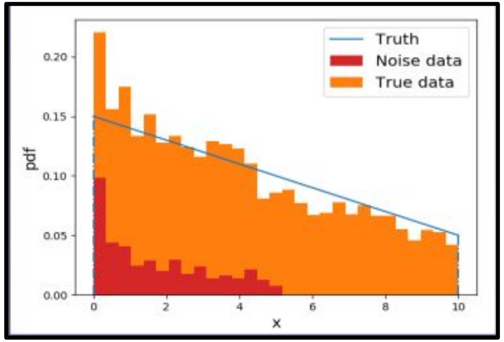




Measurement error

Selection bias





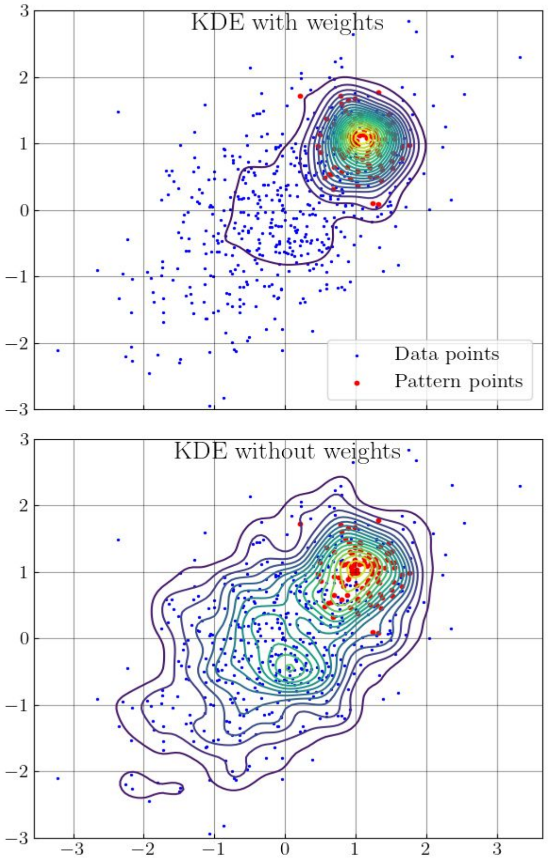
Noise contamination

Weighted Kernel Density Estimation (Weighted-KDE)

- Using a global bandwidth for a Gaussian KDE with weights based on data points

$$\hat{f}(x) = \frac{1}{\sum_i W_i} \sum_{i=1}^{n} \frac{W_i}{h\lambda_i} K\left(\frac{x - X_i}{h\lambda_i}\right).$$

$$\sum_i W_i = n, \quad W_i \iff X_i, \quad \lambda_i = 1$$



Selection Effects

- To Get the Rates from density estimate, account for selection bias

$$p_{\rm det}(x,\theta) = \int p(\rho > \rho_{\rm th}|\bar{\rho}(x,\theta,\psi))p(\psi)\, d^n\psi\,,$$

$$R_{\rm det}(x) = \int R(x,\theta) p_{\rm det}(x,\theta)\, d^n\theta\,.$$

$$R(x,\theta) = R(x) p_{\rm pop}(\theta)$$

$$R_{\rm det}(x) = R(x) \int p_{\rm pop}(\theta) p_{\rm det}(x,\theta)\, d^n\theta$$

$$\equiv R(x) p_{\rm det}(x; p_{\rm pop})\,,$$