# Clustering compact-binary objects in the parameter space through Probabilistic Hough Transform

Marco Canducci, Ilya Mandel and Peter Tiňo

mxc743@cs.bham.ac.uk   Ilya.Mandel@monash.edu   P.Tino@cs.bham.ac.uk
School of Computer Science, University of Birmingham, Birmingham B15 2TT, UK
School of Physics and Astronomy, Monash University, Melbourne, Asutralia

SUNDIAL
H2020 Innovative Training Network

@SUNDIALITN
@unibirmingham
@MonashUni
@CanducciMarco

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 721463

## Motivation

### Masses and Spins distributions of simulated compact-binaries.

N = 1000 simulated binaries via population synthesis code, corrupted by realistic noise. These observations are mock posteriors on future LIGO data sets.

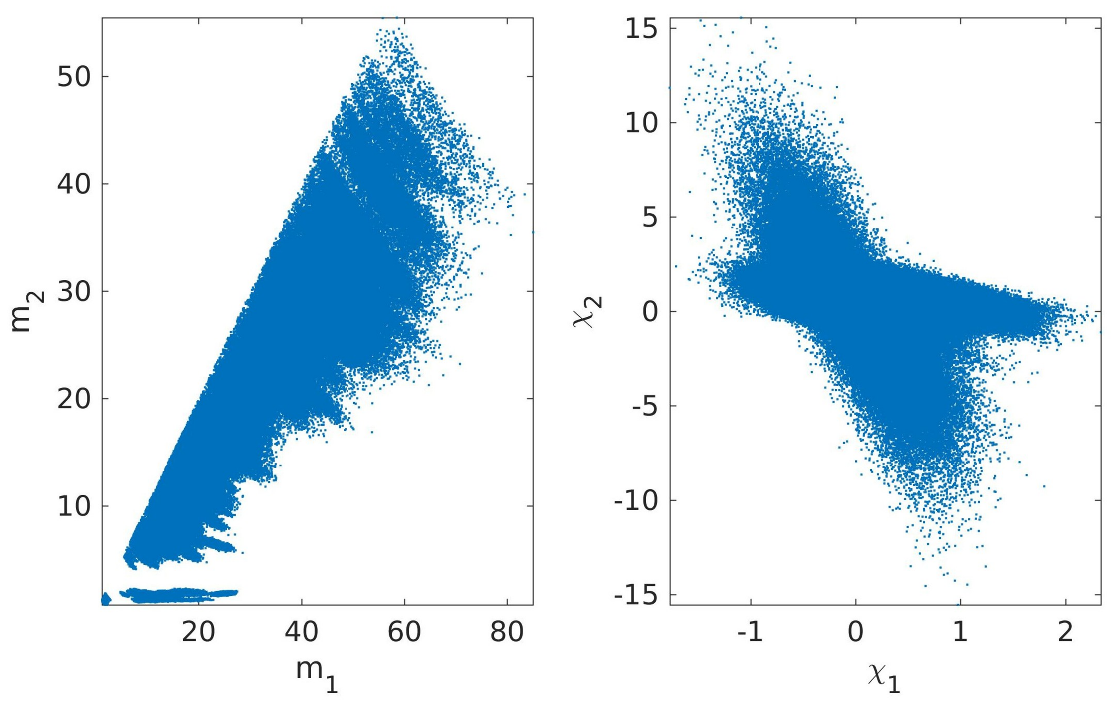


FIGURE 1: Posteriors on future LIGO datasets in masses (left) and spins (right) spaces.

Finding prototypical systems in the parameter space and clustering the models w.r.t. the prototypes could provide useful model-independent information about the population [2].

### Building a probabilistic model for each simulated binary.

Each model is a mixture of five gaussians with means aligned along the principal direction and full covariance matrices.

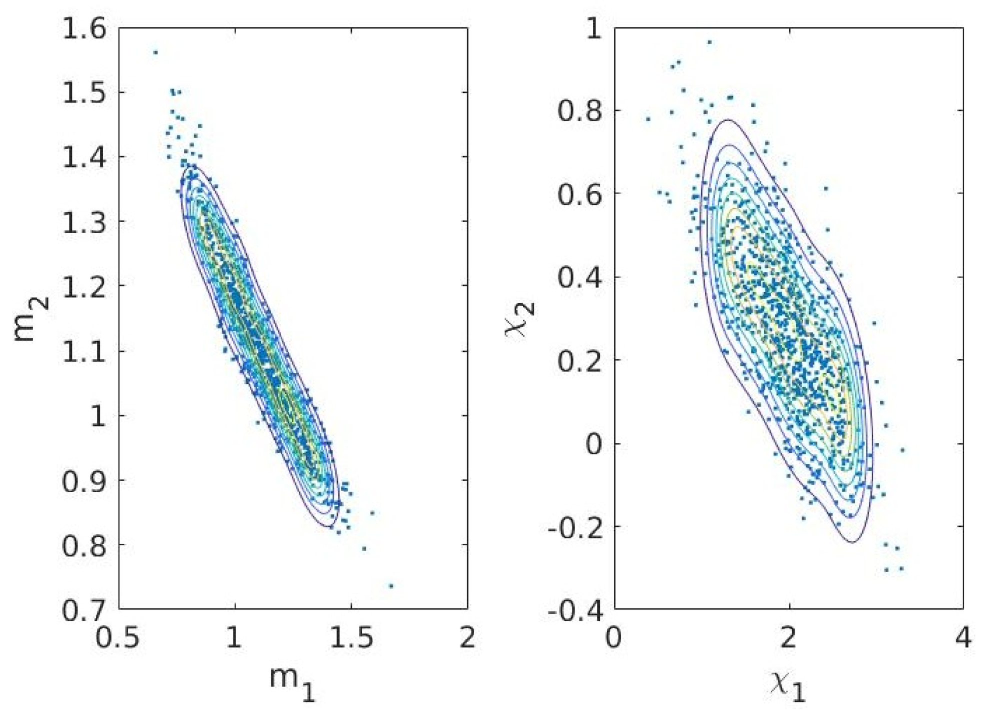


FIGURE 2: Gaussian mixture contours for a single posterior in both parameter spaces.

For each observation we then have a probabilistic model of the form:

$$q_i = \sum_{k=1}^{5} p_k \mathcal{N}(\mu_k, \Sigma_k) \tag{1}$$

## Methodology

### Probabilistic Hough-Transform

Following the work by [1], each parameter space is covered by a uniform grid and each cell (identified by the vector $\vec{x}$) contains the summation of the responsibilities of all the 1000 models $\mathcal{O}_{i=1}^{N}$ for that cell:

$$P(\vec{x}) = \sum_{i=1}^{N} p(i) q(\vec{x}|\mathcal{O}_i) \tag{2}$$

The prior $p(i)$ is assumed to have a flat distribution and is fixed to $1/N$ for every $\mathcal{O}_i$.

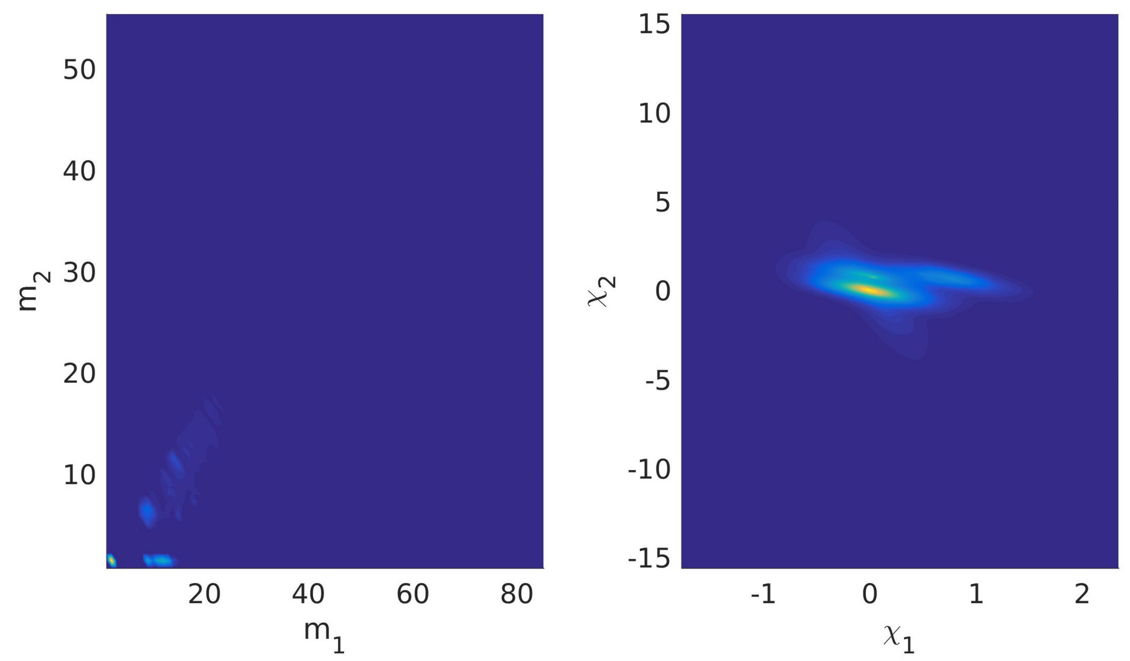


FIGURE 3: PHT for both parameter spaces: masses (left) and spins (right).

The obtained map of the reponsibilities on each parameter space is the Probabilistic Hough-Transform (PHT, [3]).

### Peak detection and optimal number of clusters

Treating the PHT as a grey-scale image, the number of connected components in it, are the number of peaks ($c_j$) in the map and thus the clusters prototypes. We can estimate the number of relevant peaks by letting a threshold T vary from the maximum to the minimum of the PHT and looking for a 'knee' in the $T - N_{Peaks}$ curve.

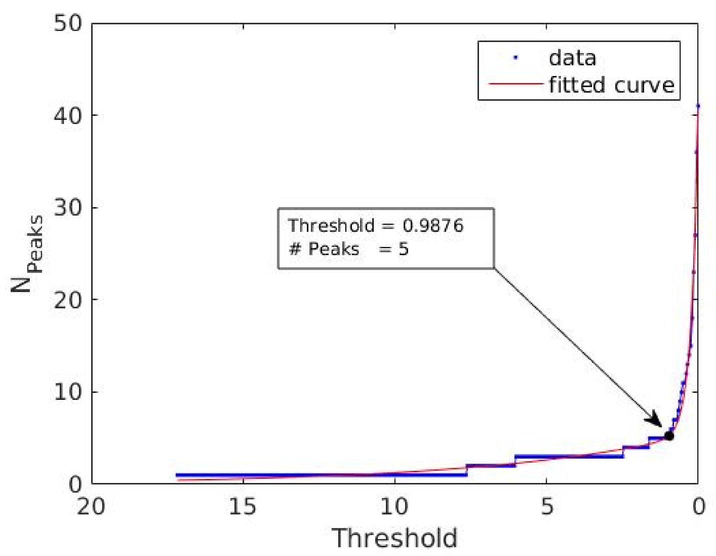


FIGURE 4: $T - N_{Peaks}$ curve for $m_1 - m_2$ parameter space, showing a knee at Threshold $\approx 1$, corresponding to 5 detected peaks.

## PHT Peaks and associated clusters

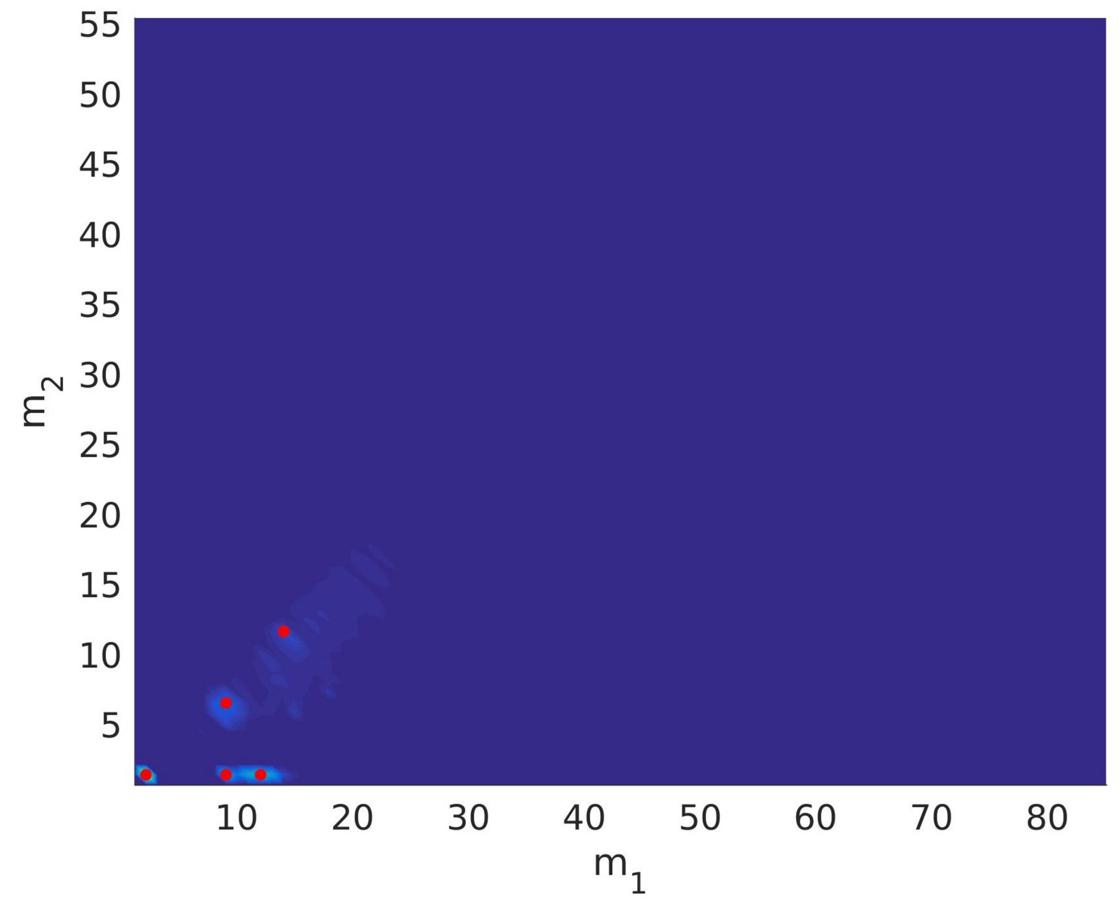


FIGURE 5: PHT with overposed detected optimal peaks.

For each peak we can now compute its responsibility w.r.t. every model:

$$p(c_j|\mathcal{O}_i) = \frac{\frac{q(c_j|Oi)}{\sum_{k=1}^{N} q(c_j|\mathcal{O}_k)}}{\sum_{l=1}^{M} \frac{qc(c_l|\mathcal{O}_i)}{\sum_{a=1}^{M} q(c_l|\mathcal{O}_a)}} \tag{3}$$

## Results

The clusters are formed by assigning the models to the peak $c_j$ for which $p(c_j|\mathcal{O}_i)$ is maximum, obtaining:

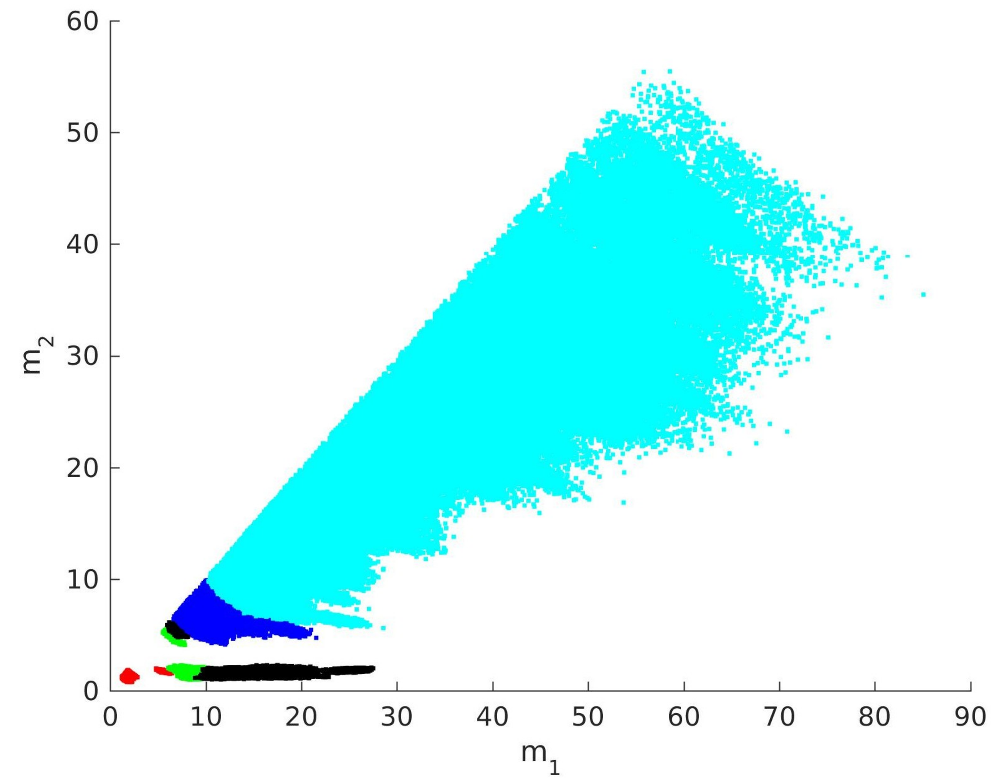


FIGURE 6: Posteriors clustered w.r.t. the detected peaks.

## References

[1] R. T. Ibrahem, P. Tino, R. J. Pearson, T. J. Ponman, and A. Babul. Automated detection of galaxy groups through probabilistic hough transform. In S. Arik, T. Huang, W. K. Lai, and Q. Liu, editors, *Neural Information Processing*, pages 323–331, Cham, 2015. Springer International Publishing.

[2] I. Mandel, W. M. Farr, A. Colonna, S. Stevenson, P. Tiňo, and J. Veitch. Model-independent inference on compact-binary observations. *MNRAS*, 465(3):3254–3260, Mar 2017.

[3] R. Stephens. Probabilistic approach to the hough transform. *Image and Vision Computing*, 9(1):66 – 71, 1991. The first BMVC 1990.