Bayesian Radio SED fitting in the era of Big Data

Emily F. Kerrison[1,2], Elaine M. Sadler[1,2] and Vanessa A. Moss[2,1]
1. Sydney Institute for Astronomy, School of Physics, University of Sydney
2. CSIRO Space & Astronomy
Image credit: ATNF

The radio spectral energy distribution (SED) of an AGN reveals important details about its **gaseous environment** (O'Dea & Baum 1997, Orienti & Dallacasa 2008), **emission mechanisms** (Tingay & de Kool 2003), and in certain cases its **sub-arcsecond morphology** (Chhetri et al. 2018). With new, all-sky surveys like RACS and GLEAM we have the data available to make and analyse radio SEDs for ~10^5 AGN or more, but there are two challenges faced in using this data:

1. **How do we build meaningful SEDs efficiently from surveys at different resolutions?**
2. **How do we extract information from a large collection of SEDs?**

I present here ongoing work on a Bayesian modelling framework which complements approaches by Callingham et al. 2017, Ross et al. 2021 and Vollmer et al. 2021. This framework is specifically tailored to broadband radio SEDs, and will be efficiently scalable to model many thousands of sources as we enter a new era of big data with the SKA.

1. Constructing radio SEDs from allsky surveys

We use RACS-LOW (Hale et al. 2021, 888 MHz, 25'' resolution) as our base survey and combine this with a subset of the radio surveys used in SPECFIND (Vollmer et al. 2021) which are selected for their reliability and coverage of the Southern Equatorial Sky between 80 MHz – 20 GHz.

Positional matching between RACS and each radio survey produces **~280,000 sources** with sufficient photometry to obtain a broadband SED (> 5 measurements at unique frequencies).

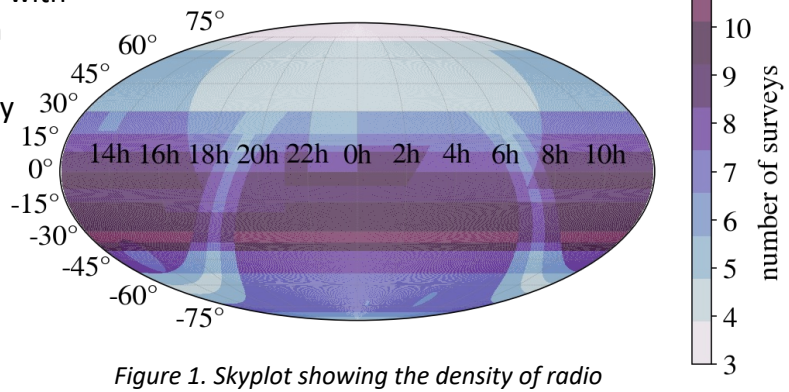
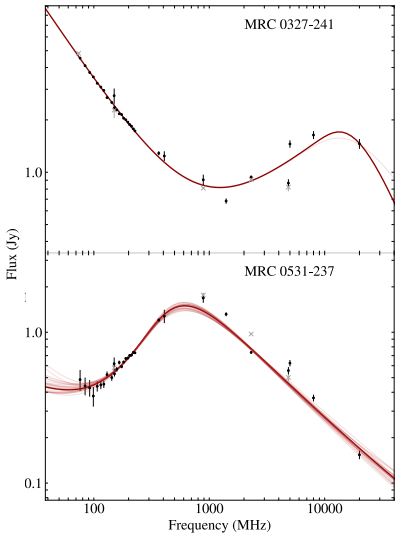


Figure 1. Skyplot showing the density of radio surveys used for our SEDs as a function of position.

2. Analysing SEDs using Bayesian modelling



We use the Bilby python package to obtain posterior probabilities for each input model given a single SED. By comparing these we can determine which model is *most likely given the data* and obtain *parameter estimates with uncertainties.*

If we use physically-motivated models, we can infer properties about the radio AGN. Currently our focus is on identifying **Peaked Spectrum** sources - a particular sub-class of radio AGN that are **compact, gas-rich and (likely) young** (O'Dea and Saikia 2021).

Figure 2. Two peaked spectrum AGN identified by our automated modelling procedure.

Why Bayesian modelling from survey data?

This approach is **uniform, flexible** (simple to incorporate new surveys) and **efficient** (SED classification is lightweight and parallelisable). It's going to make it easier to do science with Big Data!