Stacking of interferometric data: new tools for stacking of ALMA data

Kirsten K. Knudsen,[1] Lukas Lindroos,[1] Wouter Vlemmings,[1] John Conway,[1], and Iván Martí-Vidal[1]

[1]*Department of Earth and Space Sciences, Chalmers University of Technology, Onsala Space Observatory, SE-43992 Onsala, Sweden;*
`kirsten.knudsen@chalmers.se,lindroos@chalmers.se`

Abstract. Radio and mm observations play an important role in determining the star formation properties of high-redshift galaxies. With the unprecedented sensitivity, ALMA now enable studies of faint, distant star-forming galaxies. However, most galaxies with low star formation rates at high redshift are too faint to be detected individually at these wavelengths. A way to study such galaxies is to use stacking. By averaging the emission of a large number of galaxies detected in optical or near-infrared surveys, we can achieve statistical detection. We investigate methods for stacking data from interferometric surveys. Interferometry poses unique challenges in stacking due to the nature of this data. We have compared stacking of *uv*-data with stacking of imaged data, the latter being the commonly used approach. Using simulated data, we find that *uv*-stacking may provide up to 50% less noise and that image based stacking systematically loses around 10% of the flux. More importantly, we find that the *uv*-stacking yield more robust results, especially in the case of (marginally) resolved sources and mosaicked data.

1. Introduction

We have investigated a different method for stacking of interferometric data, namely stacking of the calibrated visibilities (*uv* data) as an alternative to the traditional image-stacking. In our comparative study we find that stacking in the *uv*-plane yields a more accurate and robust result as it is less sensitive to artifacts introduced when imaging interferometric data. We present both simulated and real data. The algorithm and code is available for usage.

Working directly on the calibrated visibility data (i.e., *uv*-stacking), the advantages are in particular (i) working directly on the visibility means noise on data is statistically independent, and (ii) model fitting on stacked data allows to reduce side-lopes and give reliable estimates on source size. The challenges are especially (i) model and subtraction of contribution from bright sources, and (ii) time consuming for large data sets (though if using a tailored GPU version of the stacker this can be improved by a factor 50-100). In comparision, the advantage of stacking in the image-plane is a faster algorithm and it is possibly more intuitive as it is similar to stacking at other wavelengths. The challenges of image-stacking are, however, (i) very sensitive to artifacts introduced in imaging, and (ii) does not allow for any reduction of the side-lopes due to stacked sources.

2. Algorithm for *uv*-stacking

In brief the uv-stacking algorithm work by recalculating the visibilities for sources within a single pointing using

$$V_{\mathrm{stack}}(u,v,w) = V(u,v,w)\frac{\Sigma_{k=1}^{N} W_k \frac{1}{A_N(\mathbf{\hat{S}_k})} \mathrm{e}^{\frac{2\pi}{\lambda}\mathrm{i}\mathbf{B}\cdot\left(\mathbf{\hat{S}_0}-\mathbf{\hat{S}_k}\right)}}{\Sigma_{k=1}^{N} W_k} \quad , \qquad (1)$$

where $\mathbf{B}$ is the baseline of the visibility, $\mathbf{\hat{S}_0}$ and $\mathbf{\hat{S}_k}$ are the unit vectors pointing to the phase center and the stacking position, respectively, W_k is weight of the stacking position, and A_N is primary beam attenuation.

The algorithm is designed to preserve the size of the data set and thus does not increase it. Computation is done for each visibility individually and the code can be [is] parallelized.

3. Comparing *uv*- and image-stacking

We have carried out an extensive comparative analysis of the uv- and image-stacking (details in Lindroos et al. 2015 (L15); additional discussion on implications for future facilities like the SKA in Knudsen et al. 2015). Simulated data mimicking ALMA and VLA data, both single fields and mosaics (contiguous and non-contiguous), were used. The simulations included bright foreground sources (in some cases extended), which added to the noise, and faint target sources that would be too faint to detect individually. We also simulated cases where the target sources were extended (e.g., see Fig. 1).

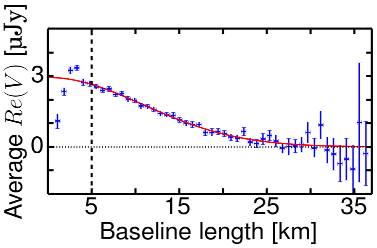

Fig. 1: Example from the *uv*-stacking (from L15): The amplitude as function of the baseline length for marginally extended sources (1.5″). Artifacts from the removal of bright sources at the short baselines causes biased results. The red line shows the average fitted size and flux densities when excluding the baselines < 5000 m.

4. Availability

The 'Stacker' code is available via `http://nordic-alma.se/support/software-tools` and runs under CASA. The code can do: (i) *uv*-stacking on single fields, contiguous mosaics and non-contiguous mosaics; (ii) image-stacking; (iii) Monte Carlo simulations; (iv) works for all radio and mm interferometric data sets incl. ALMA and VLA.

Acknowledgments. KK acknowledges support from the Swedish Research Council.

References

Knudsen, K.K., et al., 2014, in 'Advancing Astrophysics with the Square Kilometre Array', Proceedings of Science (arxiv:1501.05643)

Lindroos, L., et al., 2015, MNRAS, 446, 3502