Exploring the past with future radio telescopes

When Big Bang meets Big Data

Dr. Ir. Shahrzad Naghibzadeh and Prof. Dr. Ir. Alle-Jan van der Veen

It all started with the Big Bang! After the explosion about 13.8 billion years ago, in a fraction of a second, the universe started to expand and gradually cool down to form the first particles. After about 300,000 years, the first atoms, mainly hydrogen and helium, were formed. In the subsequent 6 to 7 hundred million years, galaxies and quasars began to form in an otherwise dark universe. About a billion years ago, the first stars began to shine and the cosmic dawn started. Fast forward to today, and astronomers and engineers work hand-in-hand to build powerful telescopes that look back in time, to unravel the history of the universe with increasingly more observational evidence. However, the increasing volumes of data poses many processing challenges. We aim to address these challenges by developing highly efficient imaging algorithms.

In observational astronomy, electromagnetic waves originating from cosmic sources, impinging on the earth, are measured with the help of telescopes. Astronomers study the cosmic objects and phenomena based on these measurements. While optical telescopes can only observe the light emitted from the stars, highly sensitive radio telescopes can look back in time. This is done by observing weak radio emissions originating from the hydrogen line and ionized gases, even before the first stars were created. These observations provide us with valuable evidence to study the creation of the universe and formations of the stars. Figure 1 shows the different views of the Centaurus A galaxy at radio versus optical frequencies. Clearly, radio frequencies provide an entirely independent view of the galaxy. During the last century, radio astronomy has been vastly advancing. Important discoveries on the formation of various celestial objects such as pulsars, neutron stars, black holes, radio galaxies and quasars are the result of radio astronomical observations.


Figure 1. *From left to right: Centaurus A galaxy at optical frequencies, at radio frequencies and superposition of the two views. (Image credit: NASA and ESA)*

To study celestial objects and the astrophysical processes that are responsible for their radio emissions, images must be formed. Proper astrophysical interpretations require the image to have a high resolution and dynamic range. Radio waves are about 1 million times longer than optical waves. Therefore, to attain a similar resolution as optical telescopes, very large radio telescopes are required. Furthermore, radio telescopes must be able to detect incredibly faint signals from very distant objects, i.e. look back in time. Consequently, sensitivity is a defining factor for radio telescopes. Sensitivity of a radio telescope is proportional to its total collecting area. Hence, to make high-resolution and high dynamic range images of the radio sky, radio telescopes with large apertures and large collecting areas are required. To make the building of extremely large telescopes more practical, a large telescope is synthesized by combining the received signals from an array of radio telescopes with smaller apertures as shown in Figure 2. Given the noisy, incomplete and indirect observations, the radio interferometric imaging problem is to invert the measurement process in such a way that an estimate of the intensity of the measured celestial sources are attained. The essence of the inversion problem consists of a (non-uniform) multidimensional inverse Fourier transform, between the radio telescope measurement space and the image space, followed by a deconvolution to correct for the (non-uniform) sampling effects. Similar problems occur very often in computational imaging, e.g. MRI, ultrasound echoscopy and radar image formation .

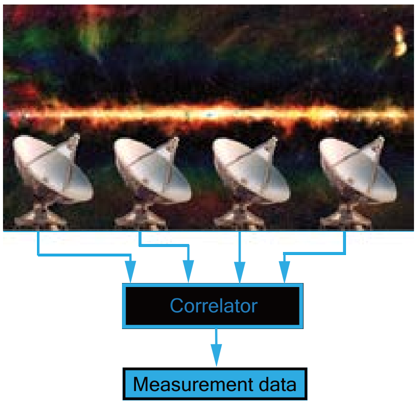


Figure 2. *Interferometry*

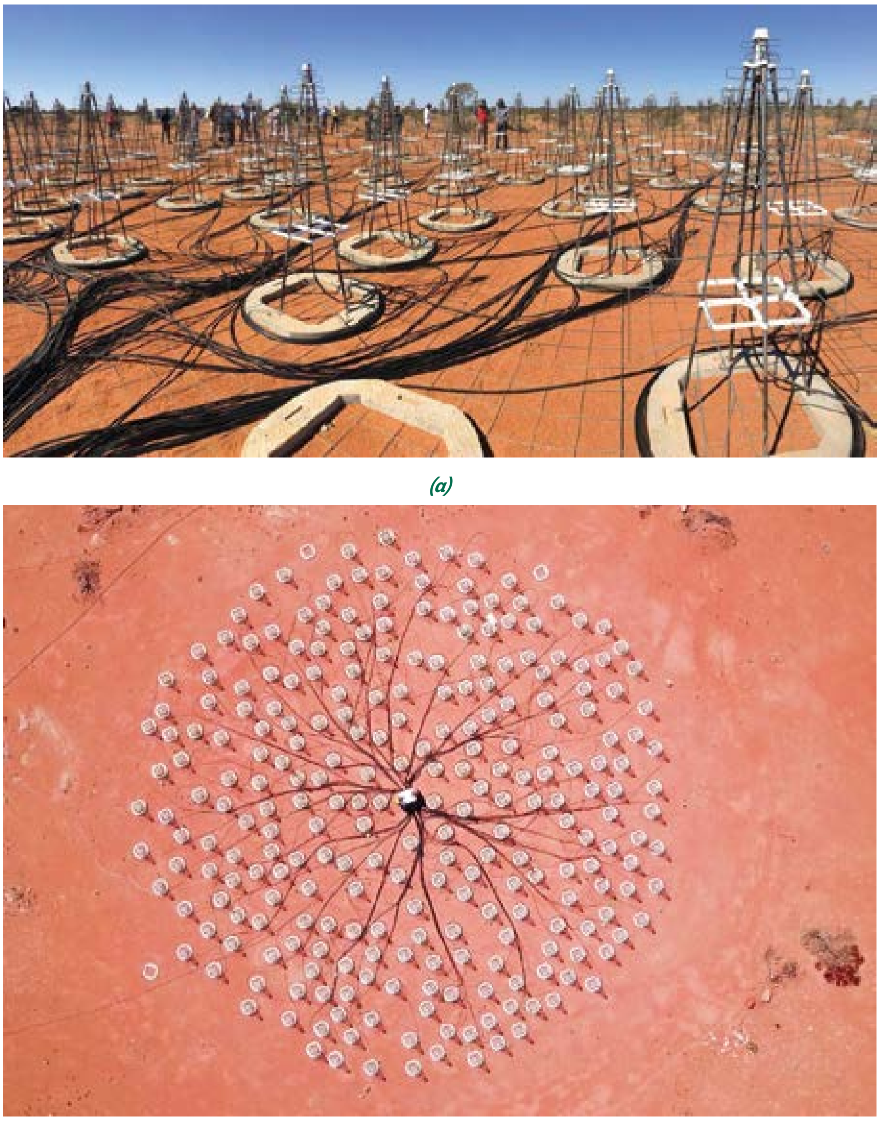
Next generation radio telescopes

To unravel the origins of the universe, engineers are building the largest science facility ever built by mankind, the Square Kilometre Array (SKA). The SKA is designed to satisfy highly ambitious scientific goals, by generating very high resolution and high fidelity images. To achieve the high-resolution requirements, the SKA will be composed of a large aperture consisting of millions of coherently connected antennas extended over an area of about 3000 kilometres. Furthermore, the SKA has a large total collecting area of about one square kilometre to increase sensitivity about two orders of magnitude relative to the current radio telescopes. The phase-one SKA (SKA1) design has been largely completed and its construction is expected to start in 2019, while the complete instrument should be ready by 2023 (provided funding becomes available). Some of the antennas constituting the SKA1 are currently under construction in the designated sites (radio-quiet zones) in Western Australia and South Africa. One of the currently operational prototype stations of the SKA is shown in Figure 3.

Nevertheless, the introduction of the next generation radio telescopes, and in particular the SKA, are bringing about many new challenges. Image formation from measurement data has become increasingly more difficult. This is due to three reasons. Firstly, the amount of measurement data is enormous: the raw data is more than 260 terabytes per second, and a single measurement session lasts hours to weeks. This data volume is beyond the available storage facilities and requires quasi-real-time processing and reduction. Secondly, to make images in the current way, supercomputers are needed that can handle more than 350 peta operations per second (peta = 10^{15}; about one million PCs). Lastly, the increased sensitivity, resolution and sky coverage, as well as the ambitious science cases of the new instruments, are beyond the capability of the current imaging methods and ask for highly accurate new imaging algorithms.

Our mission

In a shared project called DOME, ASTRON and IBM have been involved in developing a computing system that is expected to become the IT-backbone of the SKA. During the past 4 years, we have been involved in DOME to help overcome the imaging hurdles of the SKA by developing novel imaging algorithms. Our aim has been to create images that are good enough for the SKA, yet certainly do not require more computing power than current techniques do. Furthermore, the imaging algorithms have to be feasible at the scales



(a)

(b)

Figure 3. *(a) A full station of 256 low-frequency antennas at the Murchison Radio-astronomy Observatory (MRO) in outback Western Australia, (b) top view (images courtesy of www.skatelescope.org)*

required by the SKA. A major issue is that the imaging problem is numerically very sensitive; a small disruption in the data could yield very different results. To solve this issue, prior knowledge (i.e. what the sky should look like), has to be inputted to the problem. Until now, this prior knowledge consisted of modelling the sky as a collection of point sources. However, at the higher sensitivities offered by the SKA, radio emissions also appear as distributed radiation from diffuse media. In order to achieve the sensitivity of SKA, different and more general prior knowledge is required.

We have introduced a technique to efficiently model prior knowledge and to introduce it into the imaging problem. Moreover, we have developed numerical techniques to solve the problem efficiently. The proposed algorithm is named PRIFIRA [3]. The first results on small-scale test data are promising (see Figure 4): the proposed algorithm is faster than the algorithms that are currently in use, namely the methods CLEAN [2] and SARA [1]. PRIFIRA also seems to be better than the high-resolution algorithms that others recently have proposed. The next step is to test this algorithm on large-scale astronomical data and to check whether astronomers find the resulting images reliable.

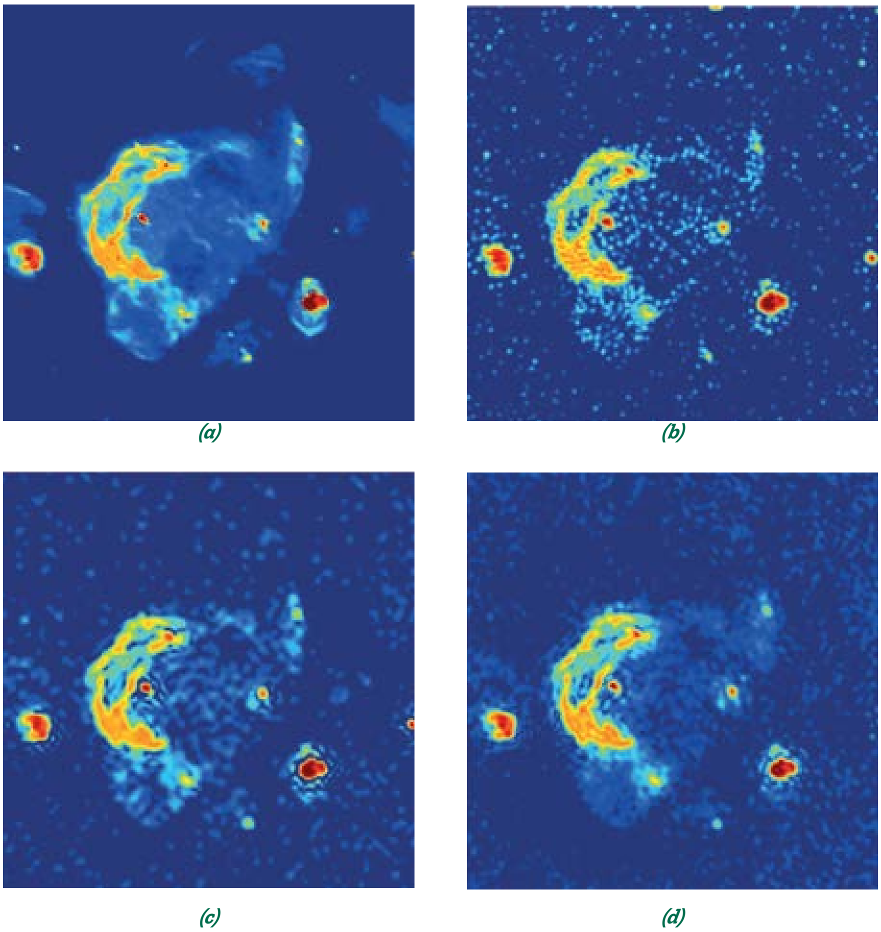


Figure 4. *(a) Original image (the measurements are simulated by transforming this to the measurement domain, by sparsely sampling based on telescope locations which are further distortted by noise.), (b) Current standard algorithm (CLEAN), 1 min computation time, (c) current best algorithm (SARA), 6 min, (d) Proposed PRIFIRA algorithm, 35 sec*

To enable the efficient use of large resources, algorithms are required to be highly scalable. Luckily, modern convex optimization problems can be formulated in a distributed fashion. Leveraging the distributed algorithmic structures, both the data and the image can be split into an arbitrary number of blocks which can be handled in parallel by splitting over multiple processors. We have devised highly scalable algorithmic structures for radio interferometric imaging by applying distributed processing in the image and data domain. We do so by dividing the large images over source occupancy regions and evenly distributing the telescope data over blocks. We have concluded, based on realistic simulations, that this scheme provides considerable memory and computational savings.

Conclusions

All in all, the SKA is one of the most ambitious projects where engineering and scientific expertise come together to make one of the long-standing human dreams come true: to explore the origin of the universe. Since the start of the project, many of the engineering obstacles have been overcome and valuable research has been conducted, whose impact goes beyond the SKA alone. For the instrument to become fully operational, there are still a lot of engineering efforts required to overcome the foreseen and unforeseen challenges.

[1] Rafael E Carrillo, JD McEwen, and Yves Wiaux. Sparsity averaging reweighted analysis (SARA): a novel algorithm for radio-interferometric imaging. Monthly Notices of the Royal Astronomical Society, 426(2):1223–1234, 2012.

[2] JA Högbom. Aperture synthesis with a non-regular distribution of interfer-ometer baselines. Astronomy and Astrophysics Supplement Series, 15:417, 1974.

[3] Shahrzad Naghibzadeh and Alle-Jan van der Veen. PRIFIRA: General regularization using prior-conditioning for fast radio interferometric imaging. Monthly Notices of the Royal Astronomical Society, 2018.