Section A. Overview of Research Project

1. **Title:** Novel data correlation methods for Very Long baseline Long Interferometry (VLBI) with next generation tensor-core GPUs

2. **Area of research:** Engineering

3. **Academic level:** Doctoral

4. **Abstract:** Observations of radio galaxies with digitally-connected radio telescopes from around the world (termed *Very Long Baseline Interferometry*; VLBI), including MeerKAT and those at HartRAO, can provide the highest angular resolution in astronomy. This observing mode uniquely allows the fine details from black holes, cosmic explosions and star-formation processes to be imaged on the important small-scales needed for astronomers to test their emission models. However, this process results in narrowing the area of sky that can be imaged, which has motivated the development of novel signal-processing techniques that allow many parts of the sky to be imaged at once. These methods are limited by the computational power of current technology. However, recent advances in FPGAs and the emergence of tensor-core GPUs has dramatically increased the number of computations that can be done, while lowering the amount of power needed. In this project, a PhD student will investigate methods that apply next generation hardware to wide-field VLBI data. This will involve developing signal-processing algorithms and benchmarking the various technologies to determine which is most efficient. The final goal is to present a method for correlating VLBI data from experiments using MeerKAT and HartRAO with European and Australian telescopes in the short term, while providing a viable methodology that can be applied to the African VLBI Network in the future.

5. **Primary supervisor:**

   a. Prof. John McKean
   b. mckean@astro.rug.nl
   c. South African Radio Astronomy Observatory (SARAO) / University of Pretoria

6. **Co-supervisor / Research Supervisor:**

   a. Dr. Johan Schoeman
   b. j.schoeman@up.ac.za
   c. University of Pretoria

Section B. Details of Research Project

1. **Engineering merit:**

Our ability to image the structure of celestial objects with radio telescopes is limited by the sensitivity and resolution of current facilities. Building extremely large (continent-scale) dish antennas is clearly not practical, therefore, astronomers use a process called interferometry, which combines the digitised signals from radio telescopes that are separated by vast distances to form a multi-element array. The more radio telescopes that are connected, the greater the sensitivity of the observations; the larger the distance between them, the better the angular-resolution for identifying small-scale structure in the objects. This technique relies on being able to combine the signals in phase, which involves a process called correlation (multiplying and time-averaging the signals after using an appropriate time-delay). This process is the foundation of modern radio astronomy. However, it also comes at a cost, since due to the finite bandwidth of the observations, the resulting effective area of the sky that can be used becomes extremely narrow, limiting the amount of sky that can be imaged at high angular resolution. To overcome this, a signal-processing technique that involves making many zoomed-in datasets from a much larger dataset was introduced (e.g., DiFX; Deller et al. 2011; termed multiphase centre correlation). Even though this allows many more objects to be imaged at high angular resolution, the computational requirements still limit the numbers that can be effectively imaged, and the time needed to correlate the data can be prohibitive.

The development of next generation FPGA and tensor-core GPUs (Broekema et al. 2018) can potentially revolutionise this field by allowing several hundred radio sources to be imaged at extremely high angular-resolution from a single VLBI observation. This would provide a new avenue for research since the numbers of objects that can be studied by astronomers can quickly move into the tens of thousands. However, it is not clear which of FPGA or tensor-core GPU-based correlation offers the best value in terms of cost, power, efficiency and availability. In this PhD project, the student will investigate recent progress and advances in tensor-core GPU correlation implementation techniques, and based on this, develop the required software needed to build a next generation correlator for VLBI in South Africa; currently such datasets are correlated by colleagues in other parts of the world. A similar project is being conducted on a hybrid GPU/FPGA architecture to leverage the power of recently released RFSoC technology by another PhD student at the University of Pretoria, which will serve as a useful comparison. The student will apply their methodology to test data taken with the European VLBI Network (EVN) and/or the Long Baseline Array (LBA), both of which includes HartRAO and eventually MeerKAT. The impact of the project is that wide-field VLBI projects (led by astronomers at SARAO/UP) can be correlated locally, developing local expertise for the AVN and facilitating high-impact science by the South African community

2. Feasibility:

As a starting point, there are several open-source codes available for multi-phase centre correlation (SFXC; DiFX) that are based on CPU (~1000 cores) technology. These are capable of making 100s of phase centres per dataset, albeit with a runtime-factor of ~12 between the total observing time and the total correlation time for large EVN datasets (with ~130 phase centres). Therefore, this is ultimately a “big data” challenge that involves processing the raw voltage data taken at 20 to 25 radio telescopes around the world, operating for about 12 hours per experiment and with data-rates of 1 Gbit/s. This results in typical raw datasets of about 80 TB that need to be correlated and a total data-rate of 25 Gbit/s that needs to be read into the correlator. Our goal is to have a runtime factor of 1 to keep up with the data-rates (real-time correlation), but a runtime factor of 4 would also be acceptable. This requires 39.6 TFLOPS, which could be achieved with 2 x NVIDIA A100 PCIe 80 GB GPU (assuming 32 bit floats; max bandwidth 16 Tbits/s). The UP compute cluster (expected delivery mid-2023) is specified to operate with 100 TB SSD “Flash” memory, 6 TB of RAM and will have 4 x A100 GPUs. This cluster will be used for testing correlation methods, and eventually, to correlate wide-field VLBI datasets that include HartRAO and then MeerKAT.

The overall project plan is:

Year 1: Carryout literature review, install CPU-based correlation codes on cluster (CPU compute nodes), and run correlation of wide-field VLBI experiments led by UP astronomers to gain experience (and provide science datasets for other PhD and MSc students in the group).

Year 2: Port standard methods to GPU technology (or use GPU-based modifications to existing SFXC and DiFX, if available), determine run-times and bottlenecks in the analysis. Develop solutions to these issues and create an efficient workflow for VLBI data processing. Apply to single-field correlation of MeerKAT-HartRAO-Ghana(+EVN, LBA elements) commissioning datasets.

Year 3: Develop and apply GPU-based correlation methods to wide-field VLBI experiments to provide science quality dataset and identify areas of improvement for future work. Conclude thesis.

3. SARAO research priority areas:

The project directly ties in with the following main SARAO postgraduate research focus areas in 2022:

> *Real-time digital signal processing instrumentation for radio astronomy, specifically using FPGA and GPU platforms.*

The proposed techniques, if successful, will lead to local expertise in the area of data correlation and signal processing, which is vital for realising the AVN.

4. Qualifications, academic abilities, skills and/or experience:

The PhD student should have a signal processing background with an interest in digital systems and algorithmic design in CUDA (Nvidia GPU).

Section C. Curriculum Vitae

Personal Details

Name: Prof. John McKean
Position: (Incoming) SARChI in Very Long Baseline Interferometry (VLBI)
Institution: SARAO / University of Pretoria
Email: mckean@astro.rug.nl

Education

- PhD Radio Astronomy, 2004, University of Manchester
- MSci (Hons) Physics and Astronomy, 1999, University of Glasgow

Awards, Distinctions, Fellowships

- Marie Cure Research Fellowship (2005–2007), Max Planck Institute for Radio Astronomy
- Postdoctoral Scholarship (2003–2005), University of California, Davis
- PPARC (UK) PhD Studentship (1999–2002), University of Manchester

Leadership Positions

- National Facilities (HartRAO) SARChI Chair Holder in VLBI, *"Probing the Nature of Dark Matter with Very Long Baseline Interferometry",* (R11M)
- NWO-CAS, "*Testing galaxy formation on the smallest sales with gravitational lensing",* (R9M)

Employment History

- Associated Professor, University of Groningen, 2013—present
- Associated Scientist, Netherlands Institute of Radio Astronomy, 2013—present
- Institute Fellow, Netherlands Institute of Radio Astronomy, 2009—2013
- Institute Fellow, Max Planck Institute for Radio Astronomy, 2005—2008
- Postdoctoral Scholar, University of California

Teaching

- *Introduction to Radio Astronomy,* University of Groningen, BSc Astronomy, 2014—present

Research Interests

- Continuum studies (LOFAR, VLBI)
- Active Galactic Nuclei and star-formation activity (triggering, feedback)
- Gravitational lensing (surveys, dark matter, high redshift Universe)
- Machine learning (source detection and characterisation)

Supervision

- 3 Postdoctoral Fellows (since 2018)
- 11 PhD Students (since 2008)
- 7 MSc Students (since 2016)
- 25 BSc Students (since 2015)

Publication Metrics

- 130 refereed papers (4 in *Nature*, 1 in *Science*)
- 7825 citations
- *h*-index = 45

Select Publications

1. *"A machine learning based approach to gravitational lens identification with the International LOFAR Telescope"*, Rezaei et al., 2021, MNRAS, 517, 1156
2. *"Gravitational lensing in LoTSS DR2: extremely faint 144-MHz radio emission from two highly magnified quasars"*, McKean et al., 2021, MNRAS, 505, L36
3. *"LOFAR imaging of Cygnus A - direct detection of a turnover in the hotspot radio spectra"*, McKean et al., 2016, MNRAS, 463, 3143
4. "*LOFAR: The LOw-Frequency ARray"*, van Haarlem et al., 2013, A&A, 556, A2
5. *"Gravitational detection of a low-mass dark satellite galaxy at cosmological distance"*, Vegetti et al., 2012, Nature, 7381, 341
6. *"High-resolution imaging of the anomalous flux ratio gravitational lens system CLASS B2045+265: dark or luminous satellites?",* McKean et al., 2007, MNRAS, 378, 109
7. *"The Cosmic Lens All-Sky Survey - II. Gravitational lens candidate selection and follow-up",* Browne et al., 2003, MNRAS, 341, 13

Dr Johan Schoeman
Pr.Eng, PhD(Eng)(UP), SMIEEE

Personal details

Gender: Male
Nationality: South African
Current residence: Pretoria, South Africa
Contact number: +27 12 420 2955
Contact e-mail: j.schoeman@up.ac.za

Education
- **PhD, Electronic Engineering**
  University of Pretoria, South Africa, 2018.
- **M.Eng, Electronic Engineering**
  University of Pretoria, South Africa, 2011
- **B.Eng (Hons), Electronic Engineering**
  University of Pretoria, South Africa, 2003
- **B.Eng, Electronic Engineering**
  University of Pretoria, South Africa, 2002

Awards, Distinctions and Fellowships
- SMEOS Best student paper runner up (2018)

Leadership positions
- Principle investigator, NRF Thuthuka Research Project "Compact In-Line Holographic Microscopy for Particulate Matter Measurement".
- Principle investigator, NRF Research Project "Characterisation and modelling of a novel dual element uncooled MEMS IR sensor".

Professional Activities
- Registered as Professional Engineer with the Engineering Council of South Africa
- Senior member of the IEEE
- Consultant to AMTS on project "Uncooled MEMS IR microbolometers" (2008 - 2012)
- Member of the Technical Programme Committees: IEEE AFRICON 2007 and International Conference on Telecommunications ICT 2005
- Journal Reviewer: ELSEVIER Vacuum
- External Examiner: UNISA

Employment History
- **Senior Lecturer**
  Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa, January 2020 – present.
- **Lecturer**
  Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa, January 2003 – 2019.
- **Assistant Lecturer**
  Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa, January 2002 – 2003.

Teaching Activities
Postgraduate:
- Research project: Theory EPT732, UP, since 2018
- Research project: Design and laboratory EPT733, UP, since 2018
- Digital Electronics EDG780, UP, 2006-2011

Undergraduate:
- Research project EES424, UP, since 2016
- Analogue electronics ENE310, since 2014
- Specialization (Advanced Digital Design) EES424, 2011-2015
- Specialization (VHDL for Engineers) EES423, 2009-2010
- Advanced Electronics ENE410, UP, 2004-2013
- Electronic Components ELK220, UP, 2004-2008
- Computer Architecture COS284, UP, 2003
- Modulation Systems EMS310, UP, 2003

Study leader:
- Postgraduate: Introduction to research EIN732, UP, since 2016
- Undergraduate: Project EPR400/402, UP, since 2003

Research Interests
- Digital signal processing for holographic microscopy.
- Digital signal processing for terrestrial communications.
- Digital signal processing for radio astronomy.
- Additive manufacturing for microfluidics and in-line digital microscopy components and packaging.
- Uncooled MEMS IR sensor (bolometer) characterisation and modelling.
- Digital signal processing for wireless communications.

Research Activities

Current Student Supervision
Supervisor or co-supervisor to 6 postgraduate students (M.Eng and PhD)

Current Research Grants
- NRF Thuthuka Grant (2021 – 2023)

Publication Metrics
- Total journal papers: 8
- Total international conference papers: 17
- Total national conference papers: 8
- Total citations in Scopus: 54
- h-index: 4

Top Publications

1. M. G. Maritz and J. Schoeman, "Programmable Aperture Using a Digital Micromirror Device for In-Line Holographic Microscopy," in IEEE Journal of Quantum Electronics, vol. 58, no. 5, pp. 1-8, Oct. 2022, Art no. 5700108, doi: 10.1109/JQE.2022.3190501.
2. Schoeman, J. and du Plessis, M. "A two-port electrothermal model for suspended MEMS device structures with multiple inputs", J. Sens. Sens. Syst., 8, 293–304, https://doi.org/10.5194/jsss-8-293-2019, 2019.
3. Schoeman J. and Du Plessis M., "An analytic model employing an elliptical surface area to determine the gaseous thermal conductance of uncooled VOx microbolometers", Sensors and Actuators A: Physical Volume 250, 15 October 2016, pp. 229-236, http://dx.doi.org/10.1016/j.sna.2016.09.033
4. Schoeman J. and Du Plessis M., "Characterisation of the electrical response of a novel dual element thermistor for low frequency applications", SAIEE Africa Research Journal, Vol. 103 (1), March 2012, pp. 9-13, http://www.saiee.org.za/
5. Maclean W., Du Plessis M. and Schoeman J., "Optimization of CMOS compatible microbolometer device performance", SAIEE Africa Research Journal, Vol. 103 (1), March 2012, pp. 3-8, http://www.saiee.org.za/
6. Du Plessis M., Schoeman J., Maclean W. and Schutte C, "The electro-thermal properties of integrated circuit microbolometers", SAIEE Africa Research Journal, Vol. 102 (2), June 2011, pp. 40-48, http://www.saiee.org.za/
7. Schoeman J. and Linde L.P., "Employing a measure of sparseness to investigate sparse data compression in AWGN conditions", SAIEE Africa Research Journal (Africon '04 – Special Issue 1: Towards Next Generation Wireless Communication Systems) Sept. 2006, Vol. 97, No. 2, pp. 157-161, http://www.saiee.org.za/
8. M. E. Goosen, P. J. Venter, N. M. Faure, P. N. Msomi, J. Schoeman and T-H. Joubert, "Hot Carrier Degradation of Mixed-mode Polysilicon Light Emitting Diodes", accepted for publication, Materials Science & Engineering B, Febr. 2023