Request for (no) Change by the CD/EoR SKA SWG, plus supporting material

Oct. 5, 2013

CD/EoR Science Team: L.V.E. Koopmans (chair), F. Abdalla, J. Aguirre, K. Ahn, A. Parsons, I. van Bemmel, G. Bernardi, F. Briggs, A.G. de Bruyn, T.-C. Chang, L. Greenhill, G. Mellema, M. Morales, U.-L. Pen, J. Pritchard (vice-chair), M. Santos, B. Semelin, T. Tsutomu, R.L. Webster

Study of the Cosmic Dawn and Epoch of Reionization, or HI at high redshifts more generally, is one of the two main science divers of SKA in phase 1 (SKA memo 125). The current SKA1 baseline design (SKA-TEL-SKO-DD-001-1_BaselineDesign1) already closely matches the science requirements needed to successfully detect and study the redshifted 21-cm emission of neutral hydrogen over a redshift range of ~6 to ~25, as extensively discussed in Mellema, Koopmans et al. (2013, Exp.Ast. 36, 235).

In March 2013 a science-assessment workshop was held at the SKA office to discuss the baseline-design for SKA1 with many of its stakeholders. A number of questions were raised during the workshop on changes and/or modifications to the current baseline design, as summarized and listed in Appendix A.

Based on extensive discussions during and since the workshop, and during subsequent telecons by the CD/EoR sciene team, we conclude the following:

Recommendations and Request(s) for Change:

Generally the CD/EoR science team is content with the current SKA1-low baseline design, but we make several recommendations and request a number of modifications to the current baselines design. We do not propose a detailed design, but some general principles that we think should be included in the detailed design that will emerge over the coming year(s).

(i) The CD/EoR SWG recommend no change in the frequency range and critical frequency to either lower or high frequencies, nor major changes in the current sensitivity (A/T) of the current baselines design as function of baseline-length, the latter both of which could be detrimental to SKA's ability to do tomography of the EoR by lower the required sensitivity on shorter baselines. Although the SKA will be the only instrument in the near future that can detect 21-cm emission from the Cosmic Dawn (50-100MHz) it is felt that the impact of shifting the optimal frequency from 111 MHz to much lower frequencies (e.g. 80 MHz), which could increase SKA's effective collecting area at its lowest frequencies if accompanied by a larger collecting area, is not sufficiently strong at this point to warrant such a request. It also poses the potential thread of lowering the effective collecting area above the optimal frequency (if not cost-neutral), leading to a decreased sensitivity at redshifts/frequencies of the Epoch of Reionization. During the latter phase, imaging/tomography can be done at the >3-sigma level with the current design and any loss of sensitivity could be detrimental to the success of EoR studies. In addition, we conclude that shifting the optimal frequencies to higher frequencies should not be done either since it inevitably leads to a severe loss of effective collecting area at lower frequencies[1]. Finally, current understanding places the peak of the EoR (z~10; see e.g. Mellema et al. 2013), where SKA1-low's sensitivity is optimized in the current baselines design.

(ii) The CD/EoR SWG strongly recommend to retain the currently recommended long baselines up ~90 km, a length which for SKA's pathfinder LOFAR has proven extremely useful in calibrating the instrument's directionally (in)dependent gains and ionospheric effects. The increased dynamic range required by SKA, for tomography to ~1 mK level, is unlikely to be reached without the use and diagnostics provided by the longer baselines on the instrument, ionopshere and sky. Despite the higher costs (also computational) of longer baselines, the CD/EoR science team concludes that much shorter baselines could seriously impede the ability of calibrating SKA-AA-low and strongly increase degeneracies between sky, instrument and ionospheric parameters – because of a severe lack of independent measurements at shorter baselines – all three of which need to be determined to levels of $<10^{-7}$. In addition to these arguments, we also foresee that low-frequency science cases (e.g. continuum surveys), other than the CD/EoR case, would severely suffer from decreasing the longest baselines below the currently recommended ~90 km lengths in the the baseline design, especially at lower frequencies where the confusion level is reached rather rapidly. Whereas longer baselines could be considered, it should not go at the cost of decreases sensitivity on the shorter baselines needed to reach ~1 mK brightness temperatures from scales of 5' to a ~1 degree from z~6 to z~25, respectively.

(iii) We request that the elements in the inner core region (inside a 400 m diameter area with filling factor f=1) are be placed on a regular grid to create redundant baselines, which will help speed up instantaneous calibration and instrument diagnostics. We note that a filling factor of ~1 already demands some level of redundancy, hence we expect little difference in the uv-coverage between such a redundant part of the core and a core where elements are (semi)randomly placed. We also request that the clusters of elements ("stations"), outside the core, are build with internal redundancy to allow rapid station calibration and diagnostics and hierarchical beam forming (see below). For radii >3 km (the current "long-baslines" stations), this could be done by placing elements on a regular grid of say 70m x 70m of 105m x 105m stations inside which at least 35m x 35m (or smaller) station-beams are formed to ensure the required FoV for CD/EoR science. For calibration or other purposes, these clusters could form larger stations when needed for S/N reasons. However, inside the core area of 3 km radius as current foreseen in the baseline design, we recommend to keep ~35 m diameter stations rather than ~105 m diameter stations, opposed to the suggested modifications to the station layout (see "SKA1-Low: Station Size and Core Configuration" and the "Imaging Science Performance Document" by R. Braun). Within the 3-km core this will improve the *instantaneous* uv coverage compared to the use of fewer clustered stations, which will be advantageous in short time-interval (ionospheric and instrument) calibration and imaging, all needed to reach dynamic ranges of $>10^7$.

---

[1] Shifting the optimal frequency from the current 111 MHz (z~12) to for example 150 MHz leads to a factor ~1.8 smaller effective collecting area (for fixed number of receivers) at lower frequencies and a factor 3.3 in integration time for CD/EoR power-spectrum and tomography, which is very substantial.

(iv) We request that (hierarchical) beam forming can be done inside a central bunker, with each core element having a separate signal path to a single central beam-former, allowing *any* combination of elements to be combined in a beam-formed signal, not exceeding the current (~866) signal paths in the core; the 45 outer stations have a single fiber connection possibly transporting multiple station beam signals). This allows for maximal flexibility in the system (e.g. beam-forming of smaller stations can be done where more FoV and less collecting area is needed, or visa versa) and allows for a much easier path to upgrading the beam-former and data-processing when processing cost go down in the future. In addition, it allows in combination with elements placed in a regular grid for novel data-processing techniques (e.g. FFT-telecope) to be tested at later stages of the project.

In summary, we request minor changes and additions to the current baseline design and remain strongly supportive of the current SKA1-AA-low baseline design recommending no major departures from this design in its frequency range, collecting area, radial antenna distributions, etc., which could all be detrimental to the CD/EoR science case.

(A) Cosmic Dawn-Epoch of Reionization Science Assessment Workshop Summary

Summary: A science assessment workshop focused on the SKA1 Baseline Design and its ability to conduct science observations relevant to Cosmic Dawn and the Epoch of Reionization was hosted by the SKA Office on 2013 March 26–28. There was also discussion of the path toward SKA2. This document summarizes the main conclusions and where work is needed over the coming months by the SKA EoR Science Team, possibly leading to requests for changes to the Baseline Design before 2013 October.

Frequency Coverage: There was good agreement on the Baseline Design frequency coverage. Reducing the lowest operational frequency to 50 MHz was seen as a very positive development. The higher frequency limit of 300–350 MHz is sufficient to cover the end of the EoR and might even enable some intensity mapping after the EoR. The current design was also recognised as sufficient for high-z HI absorption studies. No major changes are likely be recommended, but the Science Team will investigate whether the dense-sparse “transition” frequency should be shifted somewhat below 100 MHz in order to optimize for sensitivity at high z. The Team recognises the potential impact for the sensitivity of the SKA1 at higher frequencies and that such a change might require an increase in physical collecting area in order to compensate for the loss in A_{eff}/T_{sys} above the “transition” frequency. The Team will also assess whether a small frequency buffer below 50 MHz should be recommended. No recommendations for change will be made at this time.
Required A/T, sensitivity: A brightness temperature sensitivity limit of 1 mK in 1000 hrs on 1’ scales is broadly agreed upon as stated in the baseline design and DRM, but it was recognized that this would only be achieved in an rms sense in power-spectra measurements during the EoR, not during the CD. Tomography can be done on scales ~5’ at 200 MHz, increasing to scales ~1° at 50 MHz inside neutral patches. It was recognised that this is unavoidable without increasing the collecting area of SKA by 1–2 orders of magnitude. However, it was also noted that 1‘ tomography can be

done during the EoR to map out ionized bubbles that have a contrast of 20–30 mK with respect to neutral patches. Overall the current Baseline Design was agreed as being close to optimal within given boundaries. No major changes are likely to be recommended, but the Science Team will look in to minor adjustments of the core size, to gain a little in angular resolution without impacting the brightness temperature sensitivity too much due to sparsening of the array. No recommendations for change will be made at this time.

Station-size/FoV/Multi-beaming: Interesting angular scales during the CD-EoR range from arcminute to ~1°. Whereas there was agreement that one should image these scales, there was some disagreement on whether this can be done though mosaicing the images produced from smaller beams/beamlets (i.e., from larger stations) or whether it is better to produce these images using a single beam (i.e., from a smaller station). The disagreement centered less on whether mosaicing can be done in principle, but more around whether the resulting images would be reliable, especially since the CD-EoR H i intensity levels would be at a level of ~10-4 from foregrounds that show structure on similar scales. The argument in favour of larger stations would be, in part, a reduced computational burden (correlation and imaging) and possibly improved calibratability. Overall this was recognised that the current station-size/field-of-view of > 4° is sufficient to reach all stated science goals, but the Science Team will study the feasibility of mosaiced multi-beam images versus large single beam images of similar field-of-view and the impact of mosaicking on sensitivity for tomography and power-spectra. It will also look into the hierarchy of beamforming/correlation inter- versus intra-station. A CD-EoR Science memo should be written on this topic.

Station distribution/u-v coverage: The current Baseline Design (core) will be able to reach the 1 mK power spectrum limit within ~1000 hrs, and there were no major concerns identified. The core size seems matched well with tomography, i.e., ~5‘ resolution at 200 MHz, which would allow imaging to the 1 mK level. The FWHM angular scale increases roughly to match the tomography requirement. The Science Team will look into the impact on the CD-EoR science case of having baseline redundancy in the core, although the high filling factor will almost guarantee this. In addition, it will look into a slight expansion of the core. No recommendations for change will be made at this time.

Long Baselines: Arguments were made in favour of having longer baselines: calibration, information content, ionosphere, reducing confusion noise, but what maximum baseline length is required remains open, and was recognised as a very important issue that needs to be addressed quickly. Longer baselines, it was argued, could be essential to keep the interest of a wider (non-EoR) community, since the telescope can only do EoR science a fraction of the year. Having confusion noise-limited short-baseline images will be of less interest to the wider community. The Science Team will assess what the minimum maximum-baseline length should be for CD-EoR science and how many stations are needed at longer (outside core) baselines. A CD-EoR science-team memo should be written on this topic.
Log-Periodic Elements: It was recognised that the drop-off in sensitivity with zenith angle (which scales as cos q) is acceptable for CD-EoR science given that one wants to observe around the zenith and has a limited observing window, due to a desire to avoid observing close to the Galactic plane. Questions were raised regarding the

beam pattern and whether this receptor behaves well over the desired wide frequency range especially if the optimal frequency is in the lower part of the frequency range. It was recognised that using circular polarisation feeds will not help. The Science Team will look into the impact of the frequency dependent beam pattern on the ability to do tomography over the full frequency range (w/varying resolution) and look into polarization purity and the ability to measure Faraday rotation. No recommendations for change will be made at this time.

Computational Costs and Data Products: It is recognised that the computational costs of a CD-EoR experiment can be high and that long-baselines and station size play a prominent role here. Overall the EoR community would like to work with some level of preprocessed visibilities. At the moment the Science Team makes no recommendations, but it will try to scale LOFAR/MWA/PAPER/GMRT computing costs to SKA1.

Path to SKA2: Overall it was recognised that SKA_Low should also excite a wider community and that the inclusion of longer baselines would be a requirement for doing so. Important (future) upgrades from the current baseline design could be multi-beaming and an increased A/T at either very low frequencies (< 100 MHz) and/or higher frequencies (> 200 MHz) for CD and/or intensity mapping studies, respectively. Finally, baselines longer than the currently recommended 100 km could be added to attract a wider science community. More general discussions focussed on that any major design change should focus on science goals where SKA (1 & 2) can be unique, and maybe also on exciting the physics community. A deeper thinking on complementarity with other instruments/surveys should also be encouraged in this process.

Participants: J. Aguirre, B. Semelin, A.G. de Bruyn, G. Mellema, J. Pritchard (vice-chair SKA EoR Science Team), K. Ahn, L. Greenhill, L.V.E. Koopmans (Chair SKA EoR Science Team), M. Santos, T.-C. Chang, U.-L. Pen, F. Abdalla, I. van Bemmel, R.L. Webster, R. Braun, F. Briggs, M. Huynh (SKA Deputy Project Scientist), SKA Office

(B) SKA1-Low Core Antennas on a Redundant Grid

Introduction: This short memo contains arguments for a redundant position grid for the antenna elements in the SKA1-AA-low. It grew out of a brief discussion at the SWG EoR meeting in Manchester on 26/27 March 2013. A 2-dimensional redundant array permits a sky-model independent calibration[2] and very fast system diagnostics without the need to self-calibrate, grid and image the data. It also provides an accurate starting image for subsequent high-level calibration. Previous work on this topic has been summarized in a few documents that are enclosed. The first is a presentation by an ASTRON-group on the pros (many) and cons (hardly). The second is, we believe, by Rosie Bolton from the (then) SPDO. Both documents were drawn up in Jan-Mar 2010 in preparation for the SKA-science meeting in Manchester in March 2010. For some studies on the use of redundant spacings calibration we also

---

[2] The level to which this is possible depends on the level to which synthesized/receiver beams (including mutual coupling) and the ionosphere in the core are identical since redundancy assumes that similar baselines see an identical sky.

refer to Wieringa (1992, Exp Astronomy 2, 203), Liu et al (2010, MNRAS 408, 1029) and Noorishad et al (2012, A&A 545, A108)

Why redundancy, or why not? A commonly heard argument against redundancy is the loss of uv-coverage and the associated increased synthesis PSF sidelobe level, including possible grating lobes due to the regularity in the array. For arrays with small number of telescopes/stations this is an obvious and valid concern. However, with the very large number of elements, grouped in stations, in SKA1-AA-low, this argument against redundancy vanishes while the arguments in favour of a redundant grid increase as we will argue below. A low array-PSF sidelobe level is obviously important but certainly not the whole story. The reliability and stability (the complex gain) of the elements/stations is probably equally, if not more, important in determining the effective PSF sidelobe level. Redundancy can help diagnose these in a very simple and also very fast way. This would also allow real-time calibration, needed for detecting and locating transients, and reduce the enormous computational burden in creating an initial SKA instrument model and multi-frequency sky model.

There are additional considerations related to the array configuration of which a large fraction of elements are located on a regular grid. Among these are potential savings on roll out costs (including digging trenches and cabling), and the computational costs for system diagnostics, maintenance and calibration. These will not be discussed below, but should be seriously considered.

A redundant grid in the SKA1-AA-low core: a simple proposal: For the sake of the argument, although this is not necessarily the configuration we strive for, we assume here a core of 3 km radius containing a total of say 175,000 elements (70% of the ~250,000 assumed in the current baseline design). The remaining 30% of the elements are in clusters/stations along 3 spiral arms extending out to about 45 km, i.e. creating up to 90-km baselines. Here we assume the original 'sea-of-elements' concept for the inner part of the core, hence the core consists of two parts: (1) a fully filled (1.5-m grid) inner part (say 400m diameter = 70,000 elements) and (2) another N_s stations of each N_e elements in an area of about 2 km diameter (hence $N_s \times N_e$= 105,000). It seems reasonable, and probably unavoidable given the filling factor near unity, to put the central 70,000 elements on a regular grid. Outside this central part of the core, the filling factor will decrease and elements need to be clustered, which can also be done on a regular grid if needed (Liu et al., 2010, MNRAS 408, 1029; Tegmark & Zaldarriaga, 2010, Physical Review D 82, 103501).

It has been argued that all digitized element signals are available in the 'bunker' (AA-low/mid WPs). We emphatically support this. The SKA core would then become an extremely flexible array (see Appendix D) that could be optimized for the application of choice, as well as capitalize on future increase in computing power. It would, at some point, also allow the use of novel ideas and data-processing, such as the FFT telescope (Tegmark & Zaldarriaga, 2009, Physical Review D 79, 083530). Within a central 'bunker' one can form virtual stations, of any size, that can be combined to form interferometers, tied-array beams, etc. Such a core would offer a very wide range of scientific applications ranging from the Cosmic Dawn at 60 MHz, Epoch-of-Reionization at 180 MHz, HI Baryonic Accoustic Oscillations at 350 MHz, Pulsars at 500 MHz etc. We think it would be a missed opportunity if this flexibility would not be built in to the SKA1-AA-low core. Elements on a regular grid in the core and elements on a regular grid inside clusters/station outside this core would also allow `scale-free' beam-forming or beam-forming that compensates for the change in FoV as function of frequency. This is much harder when elements are randomly placed.

[NB: For the elements outside the 'solid' core, R. Braun recently suggested the concept of a 'superstation', consisting of 8 close-packed 45-m sized substations, where each superstation would contain 7000 elements assuming they are 1.5 m apart. Because of their size there would only be a very limited number of such superstations. It is also argues that these sub-station locations should be randomized. In view of the small number of superstations such a configuration would obviously have very limited use for redundancy and we do not consider it any further here.]

Now let us assume a (smallish) 35m diameter (virtual) substation containing about N_e=525 elements. There would then be N_s=200 such virtual stations yielding nearly 20,000 baselines. Surely these provide excellent instantaneous uv-coverage and we can afford to put them on a regular grid (Tegmark & Zaldarriaga, 2010, Physical Review D 82, 103501) without sacrificing significant uv-coverage. The array regularity of course increases grating (side) lobes levels but tapering of the array can be used to limit these to any desired level with only a small sensitivity penalty (e.g. "Imaging Science Performance Document" by R. Braun). The 35-m stations could still be co-located in clusters of (say) 9 x 35m stations in case the very low frequency work is limited by S/N. However, also here, placing elements on a regular pattern inside these (sub)stations would allow for scale-free beams to be formed in a much more flexible manner than randomized elements.

Advantages of a (partially) redundant array are: This is not the place to give an extensive discussion of the pros and cons of a redundant grid. However, we do want to list the most important ones which mostly fall in the category of (real-time) calibration, and array performance. Redundancy provides:

- instantaneous element/station performance data including their complex gains
- optimal input selection in hierarchical beamforming (tied arrays)
- speeding up sky-model building in complicated (e.g. Galactic plane) fields and subsequent direction-dependent higher-level calibration

Redundancy is only useful for the core. Beyond that the loss of uv-coverage makes it much less attractive. For further arguments and some subtleties we refer to the summary of de Bruyn et al (2010).

Questions and Work To-Do: There are still many open questions : (1) How planar is the core area? Can it be made into one plane, including Earth curvature (~30cm for 2 km core)? How planar does is have to be (see also Liu et al, 2010)? (2) How do direction-dependent effects (mostly instrumental, i.e. not-ionospheric) limit the diagnostic value of redundancy? (3) At which frequencies is redundancy most useful? (4) Can we live with a small fraction of the array not being redundantly calibrated? E.g. can they be brought in a second (self-cal) phase (5) Is there an optimum (virtual) station size or does this depend on application? (6) How do we monitor element, as opposed to station, gains? **To-do's:** (a) Calculate PSF levels for various arrays, with regular and randomized grids. (b) Assess robustness of redundant solutions against failing stations.

(C) The shortest minimum baseline of SKA1-AA-low

Introduction: Two key science drivers for the AA-low configuration of SKA1 are the study of neutral hydrogen during the phases known as Cosmic Dawn (CD) and the Epoch of Reionization (EoR). Key astrophysical parameters in this study are the spectral and angular structure of the emission and/or absorption signals. These dictate the instrumental frequency range and resolution on the one hand, and the required spatial distribution of the collecting area on the other hand. The frequency range required for this study is not a subject of great controversy although the lowest frequency is still being debated. However, the angular resolution required for the detection and characterization of redshifted HI signals is less clear.

The isotropic character of the brightness fluctuations of HI at high redshift implies that, in principle, an CD/EoR experiment could be conducted with a very small array. This is because to first order (i.e. disregarding redshift space distortions, RSD) the spatial fluctuations are encoded in both the spatial and frequency (depth) dimensions. It is also obvious that the very low surface brightness of the diffuse redshifted 21cm signals requires a centrally concentrated array with most of its collecting area located in a few km diameter core (i.e. few arcmin resolution).

It has therefore been argued that the EoR experiment does not require a physically extended array to reach their science goals, especially also because longer baselines increase costs. This approach is current being adopted by both PAPER and the MWA which have maximum baselines of about 300 m and 3 km, respectively. On the other hand, the GMRT and LOFAR arrays have telescopes/stations located very far from their dense cores. They have the potential for much higher angular resolution observations; these arrays were designed to permit a broader spectrum of astrophysical applications where angular resolution is a key asset. Although the latter is also true for SKA1-AA-low, the question has been posed what would be the minimum longest baseline to successfully conduct studies of CD and EoR: i.e. how large and array should the AA-low array really be? Conducting CD/EoR studies requires very high ($>10^7$) dynamic range imaging of the sky, and hence also a similar dynamic range modeling of the instrument. It is the latter, as we will argue, that demands longer baselines for CD/EoR studies.

Because the cost of the array, and probably the signal processing, calibration and imaging requirements are a steep function of baseline length this question has a broader impact. None of the current EoR experiments has reached the stage that they can claim to have the answer. However, here we want to argue the point of view taken by the LOFAR EoR key science project, where baselines up to ~100 km length are being utilized to enable the required instrument calibration level (i.e. dynamic ranges exceeding a million to one) to calibrate the shorter baselines on which to detect the EoR signals.

The use of long baselines: Without going into much detail we can see the following arguments in favour of a significant minimum baseline:

a) Long baselines help to build very high quality and high-resolution models of the low frequency sky. These require information on the structure, position, spectra and polarization. Depending on the element or station size, dictating the FOV, we expect there will be hundreds, if not thousands, of very bright radio sources. Good source models will also allow one to accurately remove the (chromatic!) response of these sources from the visibility data.

b) The calibration of direction–dependent element and station gains, which are time dependent, is greatly facilitated via an accurate source model. A relevant quantity in this regard is the number of constraints required to solve the many unknowns. The three classes of unknowns in CD/EoR experiments are sky, telescope and ionosphere. The longest baselines provide the largest number of independent constraints on these unknowns[3]. Calibrating the core stations using data on baselines to outer stations is indeed an attractive possibility with which we are currently experimenting using LOFAR data.

c) The ionosphere is still somewhat of a 'dark horse' in this discussion. The longer the baseline the larger the ionospheric (differential) refraction and differential Faraday rotation. During the current Solar maximum the ionosphere has been very benign over the Netherlands, which is arguably the best site of the four mentioned in terms of its spatial and temporal ionospheric structure. A spatially extended array offers the possibility to embark on 3D ionospheric tomography studies and solving for ionospheric delay screens. We have come along way in this regard using LOFAR but do not yet have convincing arguments to show that these are required to remove the much milder effects on short baselines. However, given the structure function of the ionosphere, any absence of its effect on longer baselines would imply that on the shorter baselines, where CD/EoR science is conducted, its impact is even smaller. This is indeed what we see in LOFAR images, where some remaining effects of the ionosphere on longer baselines is absent on the shorter baselines.

d) A scientific argument for longer baseline is the detection of redshift-space distortions on the cosmic HI signals which requires the separate detection of longitudinal and depth dimensions.

If the issues discussed above are accepted to some degree, the next question is how long these baselines should be. LOFAR has baselines up to about 100km, with a further array of 8 (and in the future 11) international stations extending over more than 1000 km. To achieve the exquisite spatial and spectral dynamic range to remove extended sources of 100 Jy (more than one million times as bright as the required thermal image noise) requires baselines at least 10x larger than the baselines that will ultimately be used to extract the power spectra.

(D) Flexible beamforming for SKA-low

Current baseline design: The layout of SKA-AA-low in the baseline design puts a significant fraction of the collecting area in a central core. In a LOFAR LBA-like aperture array system, the elements are first ordered into stations, and stations feed their signal into the correlator. The current baseline design suggests to build 911

---

[3] The LOFAR, but also SKA-AA-low array, baselines cover ~10x uv cells with their long baselines, than they do with only their core baselines (despite there being more visibilities per uv-cell on shorter baselines to build S/N). Combined with the the 'simpler' sky, these longer baselines are therefore very useful in calibration of the instrument and ionosphere, significant reducing the covariance between sky, instrument and ionospheric model parameters.

stations (866 in the core area) with a diameter of ~35 m and ~250 elements each. The main argument for the station size is the requirement that structures of order a few degrees need to be captured, and cannot be recovered by stitching or mosaicing (see "SKA1-Low: Station Size and Core Configuration" document by R. Braun). This poses constraints on:

- correlator load
- associated costs (multi-beaming is cheaper for fixed sky coverage)
- operational flexibility (multi-beaming enables multi-science?)
- ionospheric calibration (requires larger stations to detect enough calibrators)
- survey speed

Beamformer basics: LOFAR hierarchical beamforming

Aperture arrays, such as LOFAR and MWA, know different types of beamforming:

- Tied array: each station forms multiple beams to a slightly different part of the sky. Together all tied-array beams fill the station field-of-view.
- Phased array: combines multiple stations to increase sensitivity, but the field-of-view will be smaller. For example, the super-terp and LOFAR core can be used in phased array mode as they are on a single clock.
- Coherent: add the signals coherently will boost sensitivity but limit the FoV to the size of the total phased array.
- Incoherent: adding incoherently will increase fov to the fov of the smallest element in the phased array, but limits sensitivity
- Interferometric: the correlator multiplies the signals to produce the high resolution interferometric beam

In the case of the SKA-low pathfinders LOFAR and MWA, beamforming occurs at different levels (PAPER and the GMRT only synthesize images). For the LOFAR-HBA and MWA, a tile is analogue beamformed, combining 16 dipoles signals into 1 signal. This tile beam is used as input for the station beam for LOFAR and for MWA forms the "station" itself. For the LOFAR-LBA, the station beam is formed out of individual elements. The station/tile beams are the input for several options of additional beamforming. Within the LOFAR for example the number of stations beams (inside the tile-beam) and frequency coverage can be exchanged (currently limited to maximally 96 MHz x beams). The latter is mostly limited by the beam-former and correlator capacities, although its also limited by the data-rate from the receivers to the beamformer and/or correlator.

Operational modes for CD/EoR with SKA-low

From the science cases in the DRM for phase 1 and 2, there are several main operational modes that can be identified. These modes are described in this section and will set the boundary conditions for the flexible beamforming strategy. They are set by a number of science cases such as pulsar surveys, pulsar monitoring and timing, continuum surveys and deep fields, monitoring of the variable sky, fast transients, and the CD/EoR. Each require large flexibility in their beamforming capabilities. We concentrate here on the CD/EoR case, but note that the other science cases will dramatically benefit from the ability to flexibly form beams.

Epoch of Reionization beam-forming

The EoR will observe several fields to enormous depth. To detect structures up to several degrees in size, the instantaneous FoV needs to be at least 5 degrees (Mellema, Koopmans et al., 2013, Exp.Ast. 36, 235). This corresponds to station sizes of ~35m. The current baseline design provides for this, however, in some cases larger stations (smaller beams) or smaller stations (larger beams) could be advantageous or station beams size could grow proportional to wavelength allowing for a more uniform beamshape as function of frequency. The latter could benefit from placing elements on a regular grid (see Appendix B).

The CD/EoR science case also requires small stations to cover a large FoV (note that power-spectrum sensitivity scale with the square root of the FoV for a fixed collecting area of the array). The smaller the better, although this is ultimately limited by signal processing (911 input channels in the current baseline design). For the continuum surveys a range of station sizes may apply. When the system meets the demands for the CD/EoR science case, it often automatically meet the demands of the surveys' beam forming requirements.

Both cases require a proper calibration of ionospheric effects, especially when the SKA expands to the very long baselines. Stations of 35m may well be too small to calibrate for ionospheric effects, especially at the lowest SKA-low frequencies (i.e. below ~100 MHz; see also Appendix B). This requires at least a few sources to be detected. Given the source density in the Universe, there is a natural limit to the number of sources at any given flux. To detect enough sources, the stations need to be more of order ~100m in size. This can be part of the flexible beamforming strategy, where station sizes are larger beyond the core area (see Appendix B).

Several options are possible:

- simultaneous beam forming at a small station size and a larger one. This requires two independent beam forming paths for the same signal. EMBRACE does this already at an analogue level, SKA-low is fully digital and will be capable of even more than two independent paths.
- hierarchical station sizes/onion shells: small stations in the center of the array, surrounded by rings of ever larger stations
- frequency bootstrapping: in alternating subbands the stations are large or small, and the calibration of the large station frequency bands in interpolated to correct the small stations
- frequency scaling: the stations can be sized as a function of frequency to keep the station beam at a fixed size as a function of frequency

Architecture and requirements

In an ideal situation, this architecture would allow correlation of individual elements, which is obviously not possible within the current budget and with current technologies. Hence, *beamforming is necessary to reduce the correlator input*. From the analysis of the use cases for SKA-low, there are clearly two main modes: tied-array beam former mode and imaging mode.

One proposed design for SKA-low (see the AA-low/mid WPs) envisions all signals of individual elements in the core to be transported to one central processing bunker. This enables the implementation of a flexible beamforming architecture without physical limitations. A station is no longer necessarily a physical unit. Any size is possible within the correlator limitation. A station size can also be set independent of the frequency (subband). Stations can 'lend' elements from one another to optimize the shape of the beam, and the signal can be processed independently of frequency and station size. For CD/EoR imaging the smallest station size sets the limitations for the correlator. As correlator capacity tends to increase in time, this will enable smaller stations as time progresses. However, for calibration purposes there must always be larger stations included, as is currently already the case for the larger outer LOFAR stations (both Dutch and European).