

## Figures of merit for SKA scale configurations

R.C. Bolton<sup>1</sup>, A. Scaife<sup>1</sup>, G. Grigorescu<sup>1</sup>, R.P. Millenaar<sup>2</sup> and A. Lobanov<sup>3</sup>

<sup>1</sup> Cavendish Astrophysics Group, Cavendish Laboratory, University of Cambridge

<sup>2</sup> SKA Program Development Office, University of Manchester

<sup>3</sup> Max Planck Institute for Radio Astronomy, Bonn

**Abstract.** We present algorithms and examples for two uv-plane figures of merit that we are using to compare different SKA configurations. These are the uvgap figure of merit (which is related to the range of observable spatial scales that an array provides) and the “psfrms”, related to the efficiency with which the array samples the uv-plane on a cartesian grid. We also compare costs of data transport for different arrays using Grigorescu’s TrenchCOAT tool which optimises for trench and fibre costs (see Grigorescu et al, these proceedings).

We find that these algorithms can be used to compare layouts with  $O(2000)$  dish positions in a few minutes. Our results show that, as one would expect, grouping the dishes together on intermediate distance (20-180km) is detrimental to the uvgap figure of merit but that it has no effect on the “psfrms” calculation. We also shown that neither figure of merit is improved by having randomised rather than logarithmically spaced spiral configurations, but that such random configurations are much more expensive since so much more trench must be dug to connect the dishes to the central processing area.

### 1. Introduction

In order to quickly assess initial SKA configurations we need to develop figures of merit and supporting software, which can be applied to the uv points generated by an array. Here we describe two scientific figures of merit (PSFRMS and UVGAP) and give some examples of how we have used these, in combination with cost considerations, to assess layouts and make some preliminary decisions. Our code is quick to run: layouts of 2000 dishes take only a few minutes to assess for uvgap and psfrms.

### 2. The “PSFRMS” figure of merit

We calculate the psfrms figure of merit using the method proposed in Cornell (1984). This is implemented in Fortran and also in Mattieu de Villiers’ code iAntConfig (de Villiers 2009). The PSFRMS is a measure of the expected side lobe level in the synthesized beam. Applying Parseval’s theorem, by assessing the variation in filling factor in uv cells one obtains a measure of the rms side-lobe level in the synthesized beam. Thus, to obtain the PSFRMS quantity from a sample of uv-points the following steps are taken:

1) Divide the uv-plane into cells, sufficiently small to oversample the synthesized beam in the spatial domain (i.e. cell width = dish diameter / 2.4, so the number of cells in the uv plane is given by  $(2.4 \times \text{maximum baseline length} / \text{dish diameter})^2$ ). For example, the uv plane for a layout of 15m dishes with a maximum baseline of 2,000 km would have 320,000 cells on a side, giving a total of  $10^{11}$  uv cells.

2) Into these uv cells place the uv points. Let the number of uv points be  $N_{uv}$ . For typical SKA configurations out to 3,000km baselines, the vast majority of these uv cells will be empty, however, near to the core there may be a large number of uv cells that contain more than one uv point. Let the number of uv points in uv cell  $i,j$  be denoted  $n_{ij}$ .

3) Following ALMA Memo 18, find the standard deviation of the occupancy ( $n_{ij}$ ) of the uv cells, this is the PSFRMS figure of merit value:

$$PSFRMS = \frac{\sqrt{\sum_{ij} n_{ij}}}{N_{uv}} \quad (1)$$

Clearly, lower values are good, large values are bad. In the limit where each uv point falls into a different uv cell, the PSFRMS becomes  $1/\sqrt{N_{uv}}$ , giving the classic “improving like root N” performance.

4) When we assess configurations we take only one sample every ten minutes, which is not a realistic SKA scenario. This means that the PSFRMS value is not a good approximation to the actual power in the synthesized beam, but in the context of comparing different configurations the PSFRMS is still minimised if the number of cells sampled in the uv-plane is maximised for fixed  $N_{uv}$  (so that all the  $n_{ij}$  are unity or zero), so it remains a useful relative measure.

### 3. The “UVGAP” figure of merit

The general uvgap algorithm is presented in Lal (2009), and here we describe our new version which differs from the general algorithm in that there is no binning of uv-points in the U direction. The approach that we use is as follows:

1) Calculate the uv-points for a given layout and observation set-up (again, as with the PSFRMS, we use 1 sample every 10 minutes). Assign each uv point a value of  $q = (u^2 + v^2)^{1/2}$  and polar angle,  $\theta$  (in the range  $0-\pi$ , using conjugate symmetry to re-map points that initially lie between  $\pi$  and  $2\pi$ ).

2) Bin these  $q, \theta$  points in  $\theta$ , with the bin size in  $\theta$  chosen to correspond to the sampling time for example a sampling time of 10 minutes would require 72 bins in 12 hours, corresponding to a bin width of  $\pi/72$  radians or 2.5 degrees.

3) Now in each angular bin,  $i$ , sort the uv points into increasing value of  $q$ . Calculate the unbinned uv-gap value from  $q=0$  out to a  $q$  value of  $q_{max}$ . This might correspond to some baseline range that we are interested in such as 3 km, 30 km or 3,000 km, for example.

4) Let there be  $N$  points in the angular bin which have  $q$  values less than  $q_{max}$ . Calculate the fractional change in  $q$  for each of these  $N$  points: for the  $k^{th}$  point,  $\delta q_k = \Delta q_k / q_k = (q_k - q_{k-1}) / q_k$ . Then calculate the integral average of  $\delta q$  by assigning  $\delta q(q_{k-1}, q_k) = \delta q_k$ . We take only the uv points that have  $q < q_{max}$ , so if  $q_N$  is the last uv-point within a given sector we take  $q_{N+1} = q_{max}$  and  $\delta q_{N+1} = 1$ . Then, at the low  $q$  end of the samples, we assign  $q_0 = 0$ . In other words, for the portions of the  $q$  wedge that are before the first and after the last uv point, the  $\delta q$  value is set to unity. The integrated mean in the  $i^{th}$  angular wedge is then calculated for  $\delta q$  between  $q = 0$  and  $q = q_{max}$ .

5) The average over all angular bins of the  $\langle \delta q \rangle_i$  values is then taken to generate a single number for the uv-gap value of an array, for a particular value of  $q_{max}$  and for a given observational set up.

#### 4. Testing example configurations

In order to test the usefulness of these figures of merit and to take some initial steps towards finding an optimal, ideal SKA configuration we generated a suite of layouts of 2250 dishes. 1500 of these dishes were placed within a 2.5km radius central core and the remainder were spread logarithmically out to 180km away from the core. We tested layouts with 5 spiral arms and randomised layouts with no spiral arms (see figure 1). In the random layout, a probability distribution which was uniform in  $\log(\text{radius})$  was used to select the distance away from the core, and then the polar angle was chosen from a random uniform distribution. The spiral layouts were made following a log-spiral pattern ( $r = ae^{b\theta}$ ), where three different values of the spiral arm parameter,  $b$ , were used: 0.5, 1.0 and 1.5. The radius (away from the core) position for each dish was chosen so that the dishes outside the core were spread out by a constant ratio in  $r$ , and then the correct value of  $\theta$  was selected to place the dish on one of the 5 spiral arms. In all of these layouts we have left two gaps in the distribution, which are circles 2.5km in radius. These have been left empty to represent the forbidden regions that would be imposed by aperture array cores. In a separate investigation we have shown that these gaps do not affect either figure of merit significantly.

In addition to these layouts we also investigated the effect of grouping dishes together into clumps, for all dishes beyond a certain distance from the core (the *clumping radius*). We did this because in remote areas of the desert there will be significant costs associated with establishing infrastructure and these costs will be lowered by reducing the number of distinct locations that require power, water & access roads etc. Being aware of this natural driving factor we can use the figures of merit to ask whether such a process is likely to reduce the imaging ability of the array. An example of an array with clumping is shown in figure 4.

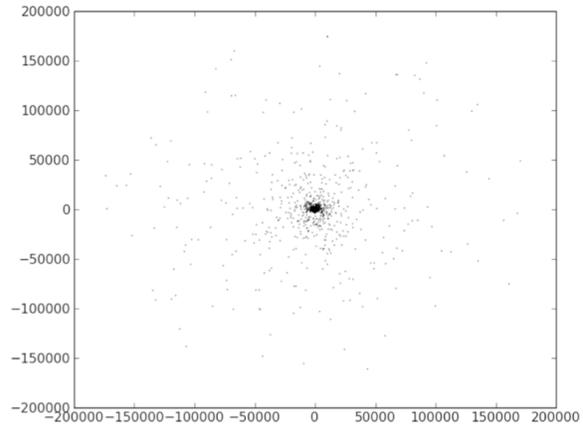


Fig. 1: Example of a randomised layout following a logarithmic probability distribution. X,Y axes in metres.

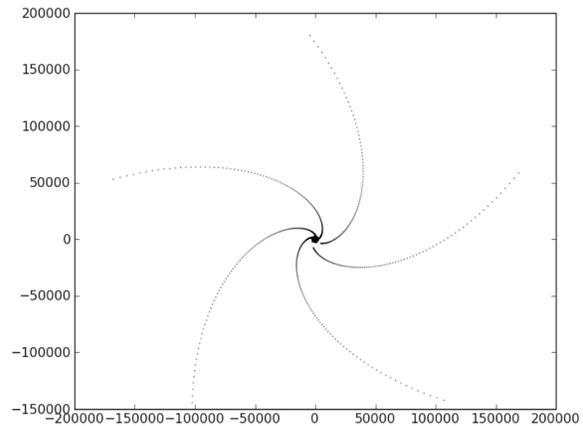


Fig. 2: Example of regular spiral layout following a logarithmic distribution. Spiral arm parameter = 1.5. X,Y axes in metres.

Figures 5 and 6 show how much clumping the dishes affects the uv coverage (for a snapshot observation made with the source at zenith). We can now quantify this effect by showing the trend in uvgap figure of merit with increasing clumping radius. Larger values for the clumping radius mean that fewer dishes are clumped, giving better uv coverage. With a clumping radius of 200km, no dishes are clumped & we return to the unclumped layout. The trend is shown in figure 7, which confirms numerically what we would assume to be the case: that clumping reduces the area of the uv plane that has uv points in it. However, results for the psfrms figure of merit show no variation with clumping. This is because relatively few of the uv points come from outer-dish to outer-dish baselines. Moving the dishes into groups changes the way that the uv points are spread over the uv plane but had little effect on the number of Cartesian uv cells that are filled (i.e. we don't change the over-

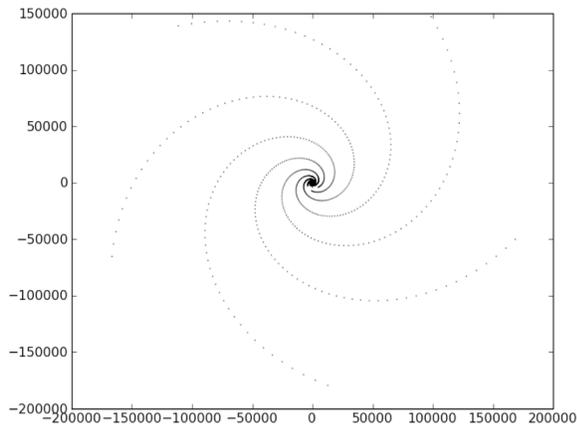


Fig. 3: Example of regular spiral layout following a logarithmic distribution. Spiral arm parameter = 0.5. X,Y axes in metres.

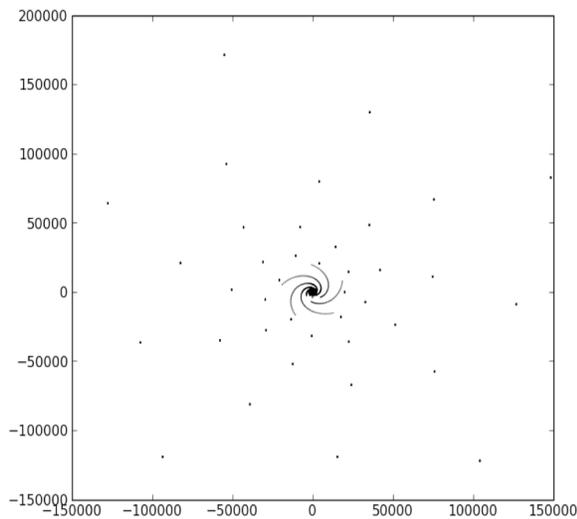


Fig. 4: Example of regular spiral layout following a logarithmic distribution. Spiral arm parameter = 0.5. Dishes beyond 20km from the core are grouped into clumps of 10. These clumps are  $\approx 200\text{m}$  in size and look like single points in this image. X,Y axes in metres.

all number of baselines), thus the psfrms is not affected strongly in this example.

We also find that spiral arms and random configurations (with the same logarithmic distribution in distance) give equivalent uvgap and psfrms results.

We use the TrenchCOAT software (see Grigorescu et al. 2009) to compare the cost of trenching and cabling for the different layouts, assuming 80 Gbits/s from each dish.

The cost analysis shows that random configs are much more expensive than spiral configs, largely because so much extra trenching must be used to connect each dish to the correlator

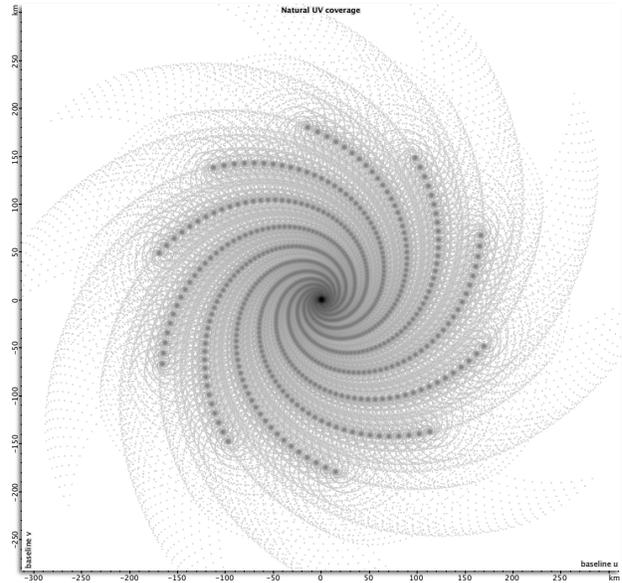


Fig. 5: Example snapshot uv coverage from an unclumped regular spiral layout similar to that shown in figure 3. These uv plots have been produced in iAntConfig.

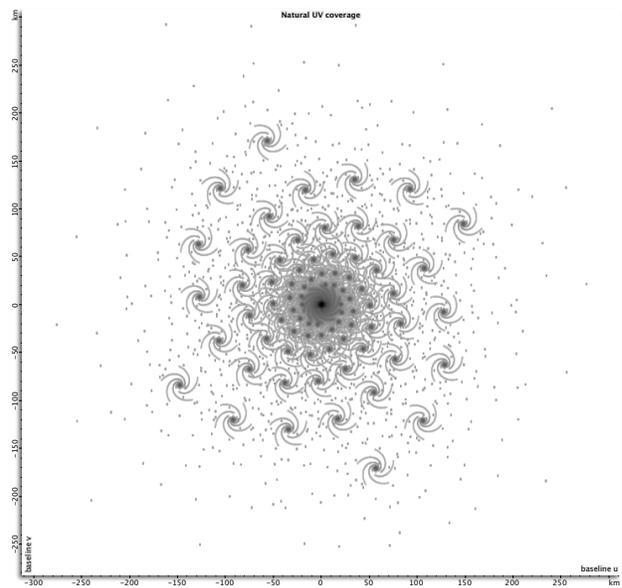


Fig. 6: Example snapshot uv coverage from a layout similar to that in 4. These uv plots have been produced in iAntConfig.

(see figure 8). In figure 9 we plot the cost from TrenchCOAT against the clumping radius for a regular tidy layout and a fully random layout. This clearly shows that for a random layout the network cost increases rapidly if less grouping of dishes occurs (since more individual, randomly placed positions then need to be added to the network), however, for layouts on spiral arms clumping the dishes together does not save any money as far as the network cost is concerned. This is because the trenches follow the spiral arms and so adding more positions

POS(SKA)DS 20091024

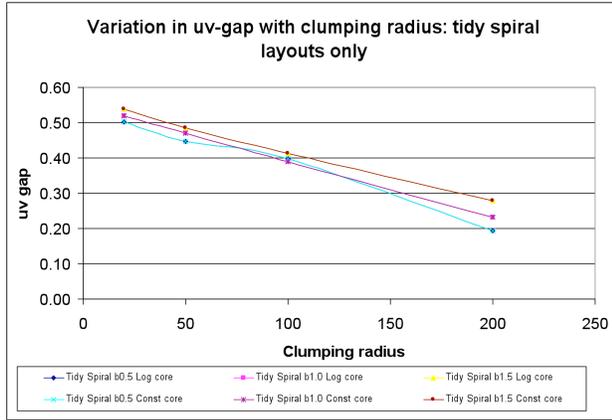


Fig. 7: Effect of clumping on the uvgap figure of merit.

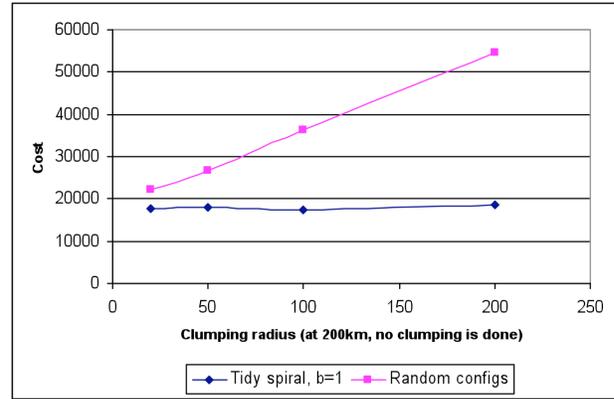


Fig. 9: Effect of clumping on the cost.

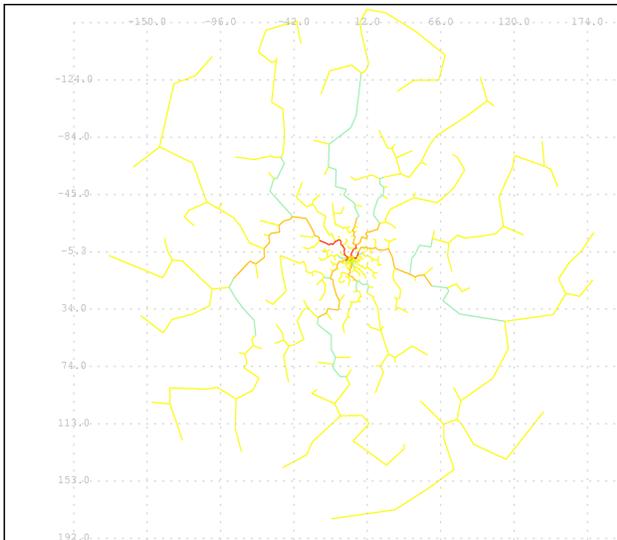


Fig. 8: Optimised trenching and fibre network for a random layout, from Grigorescu's TrenchCOAT tool.

that also sit on the spiral arms does not require any new trenching. These results, when combined with the uvgap and psfrms results, clearly suggest that we should use spiral arms with as little clumping as possible. We can therefore rule out random layouts and continue our investigations by considering spirals only.

## 5. Conclusions

We have developed some algorithms for assessing SKA-scale configurations that can be applied to the uv-plane (in the case of the uvgap and psfrms parameters) or to the layout itself (in the case of the networking cost). These can be used to test different configurations quickly and are able to distinguish different layouts quite well.

Based on the results of these initial investigations we can rule-out fully randomised arrays since these are more expen-

sive whilst offering no scientific advantages as measured by our figures of merit.

We have established that grouping the dishes together has a strongly detrimental effect on the uvgap figure of merit, whilst offering no cost savings in the network analysis. There may be other cost savings associated with the grouping of dishes on intermediate scales (e.g. from access roads etc), but these savings will have to be carefully traded-off against their associated reduction in imaging capability.

*Acknowledgements.* This work was supported by the European Commission Framework Program 6, Project SKADS, Square Kilometre Array Design Studies (SKADS), contract no 011938.

AntConfig has been developed by Mattieu de Villiers, we are very grateful for the time he has spent developing this software.

## References

- Cornwell, T 1984, MMA Memo 18
- de Villiers, M., AntConfig and iAntConfig are available to download at: <http://www.kat.ac.za/public/wiki/iAntConfig02>
- Grigorescu, G, Alexander, P, Bolton, R and McCool, R. 2009, Experimental Astronomy, article 9177.
- Lal et al., 2009, SKA Memo 107.

POS (SKADS 2009) 024