# Understanding The Impact Of Beamshapes On Radio Interferometer Imaging Performance

O.M. Smirnov[1,3]        B.S. Frank[2,3]        I.P. Theron[4]        I. Heywood[5]

1       Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa
            e-mail: o.smirnov@ru.ac.za
2       Astrophysics, Cosmology and Gravity Centre (ACGC), Department of Astronomy, University of Cape Town,
            Private Bag X3, Rondebosch 7701, South Africa
            e-mail: bradley@ast.uct.ac.za
3       SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands 7405, South Africa
            e-mail: oms@ska.ac.za
4       EMSS Antennas, PO Box 492, Stellenbosch 7599, South Africa
            e-mail: iptheron@emss.co.za
5       Astrophysics, Department of Physics, University of Oxford, Keble Road, Oxford, OX1 3RH, UK
            e-mail: ianh@astro.ox.ac.uk

***Abstract*** Primary beam shapes are emerging as a vital consideration both in the calibration of the new generation of radio interferometers, and in the design of future instruments such as the SKA. However, their impact on interferometer performance is far from obvious, especially since calibration techniques addressing these effects are still in their relative infancy. We present the results of recent work aimed at developing a comprehensive interferometric simulations methodology which can be used to quantify the impact of beamshapes on radio interferometer sensitivity. We show the results of applying this methodology to a comparative study of several possible MeerKAT dish designs.
In particular, two beamshape-related effects that have been virtually ignored in older (lower-sensitivity) telescopes are far sidelobe confusion noise (FSCN), and calibration artefacts due to pointing errors (PEs). These effects are quite subtle and often completely non-intuitive, so exploring them via simulations can lead to interesting new insights into radio telescope design.

# 1   INTRODUCTION

The primary beam (PB) pattern, or beamshape, of a radio interferometer element (i.e. antenna or station) determines that interferometer's field of view (FoV): the instrument samples the sky attenuated by the square of the PB amplitude. Traditional radio interferometry implicitly assumes stable beamshapes that are identical from element to element, and are significantly non-zero over only a limited solid angle.

Under such conditions, the flux sensitivity (i.e. the faintest sky source that can be reliably detected) of the interferometer is limited by either *thermal noise* or *classical confusion limit*. The properties of thermal noise are well understood [1]: it is (usually correlated) Gaussian, and, since it scales as

$$\sigma_{\text{thermal}} \sim (T_{\text{sys}}/A_{\text{eff}})(\Delta \nu \, \Delta t)^{-1/2}, \tag{1}$$

it can be reduced by either lowering the system temperature ($T_{\text{sys}}$), or building more collecting area (effective aperture size $A_{\text{eff}}$), or providing more bandwidth $\Delta \nu$, or observing for a longer integration time $\Delta t$. The last term in particular suggests that we may reach arbitrary high sensitivity by "simply" observing for longer. In fact this is not the case – the lower in flux we go, the more crowded the sky becomes, until we hit the *classical confusion limit* – i.e. that flux level at which sky sources become too dense to be resolved individually given the spatial resolution of the interferometer – which, in turn, is determined by the maximum baseline length $B_{\text{max}}$. Modern interferometers such as MeerKAT are designed to go deeper in flux than ever before; source densities at such faint fluxes have not been established observationally, so we must rely on extrapolations and models [2] to estimate confusion limit; an L-band limit of 0.1 μJy has been used as a rule of thumb (although recent work [3] suggests that faint source counts have been underestimated until now, so the real confusion limit is actually higher). Note that according to specifications, MeerKAT can reach 0.1 μJy thermal noise in L-band after about 5000 hours of observation.

At such high sensitivities, instrumental effects hitherto ignored become important enough to be potentially limiting. Two of these, both intimately connected to beamshapes, and are the subject of the present work: *far sidelobe confusion noise* (FSCN), and *pointing error* (PE).

## 1.1   Far sidelobe confusion noise

The point spread function (PSF) of an aperture synthesis radio interferometer is determined by (being a Fourier transform of) its *uv*-coverage, i.e. the subset of Fourier components actually measured. A typical PSF will have a strong main lobe, with a characteristic size determined by the maximum baseline, and fainter but very extensive sidelobes which fundamentally never go to zero, even asymptotically, due to the fact that our *uv*-coverage is always incomplete. This means that the interferometric response in any given direction ("image pixel", if we think in terms of radio maps) is contaminated by signal from all other directions in the sky, because the observed sky is *convolved* with a PSF of infinite extent.

Various *deconvolution* algorithms exist that can disentangle these contributions with varying degrees of success. All of them are fundamentally limited by the area of the sky actually imaged. Traditionally, one would image only a limited FoV, and assume that the signal from sources outside the FoV is negligible (with the possible exception of a handful of exceptionally bright sources – the so-called "A-team" – which can be subtracted individually during calibration). This neglected signal produces an additional noise-like contribution in the image, by virtue of being the sum of an essentially random distribution of a "foam" of faint sources, modulated by three attenuating factors:

- the far sidelobe level of the PB, squared – this being the dominant factor,
- the far sidelobe level of the PSF,
- time averaging (smearing) in the correlator, which increases with distance from phase centre.

This signal is called the far sidelobe confusion noise (FSCN). For any given image size, the FSCN level provides a fundamental "astrophysical" noise floor. On the other hand, FSCN can be driven arbitrarily low by imaging and deconvolving[6] larger and larger areas of the sky, with the full sky as the limit. Doing this at the full resolution required is computationally expensive to the point of being prohibitive – there is

---

6 Assuming an infinitely powerful deconvolution algorithm. No such thing actually exists – thus real-life images may often be limited by deconvolution artefacts rather than noise – but this is not the subject of the present work.

then a clear trade-off between imaging cost (or the area of the sky being imaged) and attainable FSCN level. Older instruments were generally limited by thermal noise and classical confusion long before they hit FSCN. However, at the sensitivities that MeerKAT (and, in the future, the SKA) is expected to attain, the FSCN limit becomes important.

## 1.2 Pointing errors and calibration "noise"

Mechanical and gravitational deformation, wind load, and other environmental effects prevent beamshapes from being perfectly stable and identical. One of the more significant effects is pointing error (PE) – rather than being centred at the nominal pointing direction, each antenna's beamshape is unavoidably offset by some amount of PE, which varies in time and from antenna to antenna. In particular, the MeerKAT specification calls for 25" pointing accuracy under normal conditions, and 5" under optimal (low wind) conditions.

PE introduces time-variable and antenna-dependent gain variations towards each source, which an interferometer ultimately interprets as structure, thus introducing artefacts into the final image. Besides distorting the image of the source itself, these artefacts propagate (modulo the PSF) across the entire field, and cannot be removed by normal deconvolution. A single sufficiently bright source can thus potentially inject sensitivity-limiting artefacts across the entire field.

Two broad approaches of calibrating for PEs are being developed, the direction-dependent (DD) gain approach (differential gain solutions in MeqTrees [4], SAGECAL [5]), and pointing selfcal [6]. To date, only the DD methods have managed to yield noise-limited (rather than artefact-limited) maps, with pointing selfcal apparently being handicapped by the fidelity of PB models.

Whichever method is being employed, the ultimate result of successfully calibrating a single observation is a noise-limited image. Any remaining calibration "noise" is buried in the thermal noise. Deep surveys involve repeated observations of the same field, which averages down the thermal noise following Eq. (1). The issue we raise in the present work is, does PE-related calibration "noise" average down in the same way, or does it eventually impose a sensitivity limit of its own – and if so, how does this depend on beamshapes?

## 2    THE BEAMSIMS FRAMEWORK

The difficulty of addressing such issues analytically calls for  simulations-based approach. We have therefore developed a simulations framework called BeamSims, based around the MeqTrees system [8]. Given the following inputs:

- a set of beamshapes, i.e. the complex PB response as a function of direction,
- a sky model (essentially, a list of sources),
- degree of pointing error, if any,
- interferometer configuration, time and frequency grid (in the form of a CASA MS),

the BeamSims code can completely simulate an interferometric observation, and write the simulated visibilities to an MS. This includes interpolating the beamshape in the direction of each source to determine PB gain, while taking into account parallactic rotation and (optionally) pointing errors. The resulting visibilities can then be imaged using standard interferometric tools such as CASA or lwimager.

## 2.1 Simulated beamshapes

A full beamshape specification consists of the 4 complex elements of the 2×2 E-Jones matrix, as a function of direction on the sky. Beamshapes are specified as either a set of FITS files (rectangular *lm* grid), or ASCII tables (spherical $\theta\varphi$ grid). In the latter case, variable grid steps are supported, with e.g. finer steps used for the main lobe, and progressively coarser steps for the sidelobes. This allows an all-sky beamshape to be economically represented.
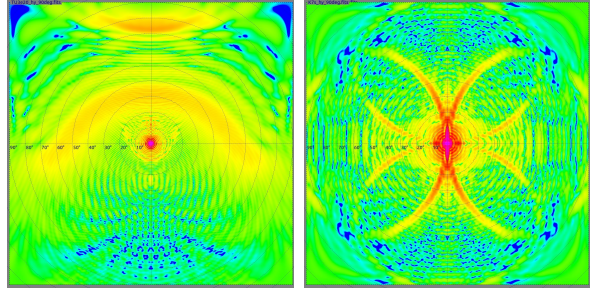


Figure 1. Simulated beamshapes for the MeerKAT offset  Gregorian (OG, left) and KAT-7 prime focus (PF, right) dish designs – complex amplitude.

The beamshapes used in this study were produced using the EMSS FEKO software. Figure 1 shows two characteristic beamshapes for the OG and PF dishes. These are plotted out to θ=90º. Both images are rendered in log scale and with the same colour range. Only the main lobe (the bright spot in the centre) actually has any appreciable amplitude; the amplitude of the far sidelobes varies in the range 0.001 to 0.005.

## 2.2 Simulated skies

To obtain statistically realistic all-sky source distributions for the FSCN study, we extracted hemispherical (i.e. out to 90º) samplings from the NVSS catalogue [9], centred at various points on the sky, and translated these to our nominal pointing centre. We imposed an upper flux cut-off of 3 Jy (assuming that anything brighter gets special

treatment during calibration), and a lower cut-off of 0.5 Jy (for performance reasons, to limit the number of sources needing to be simulated). Using the simulated source counts from [2], we estimate that omitting sources in the 0–0.5 Jy range causes us to underestimate FSCN by a factor of ~3; this correction factor is applied to the results after simulation.
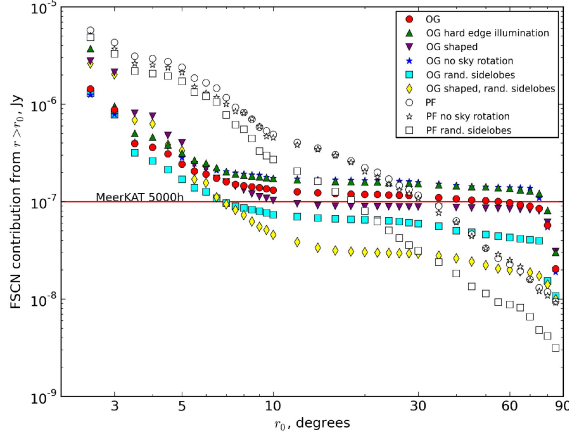
## 3 FSCN RESULTS



Figure 2. FSCN cost curve for MeerKAT64, using natural weighting.

Our interest, ultimately, is to obtain a "cost curve" of flux level versus the area of the sky needing to be imaged and deconvolved to get FSCN down to that level. In other words, for any given angular distance $r_0$, the cost curve gives us the FSCN contribution from all sources at $r \geq r_0$ from field centre.

To simulate this, we split our sky models into concentric "rings" of $r_i \leq r < r_{i+1}$, and simulate the visibilities corresponding to sources within each ring. We then image the (nominally empty) patch of sky in the middle, and measure the r.m.s. pixel values across that image. These pixels contain nothing more and nothing less than the FSCN contribution from sources in the current "ring". For each radius we then add (in quadrature) the FSCN contribution from rings at that radius and beyond.

This procedure results in cost-curves of the kind shown in Fig. 2. This summarizes FSCN level as a function of imaging radius, for a number of designs:

- OG and PF dishes on alt-az mounts (the nominal MeerKAT/KAT-7 design)
- OG and PF dishes on equatorial or three-axis mounts (thus no sky rotation)
- rot-randomized (RR) configurations, where each dish is rotated through a random angle, in order to randomize far sidelobes. Note that for mechanical reasons this is physically feasible for a PF design only, but we simulate an OG version for the sake of

completeness. Note also that some degree of far sidelobe randomization is introduced anyway, via within-tolerance differences in dish surfaces.
- Two OG variants: a shaped main reflector dish, and the default dish with a harder-illuminating feed.

These curves suggest that OG and PF designs have totally different performance characteristics with respect to FSCN. Offset Gregorian variations have lower FSCN at small radii, and have an optimum radius (~10º) beyond which increasing the image size brings very little improvement until about ~70º, where another order of magnitude is reluctantly gained. In other words, an OG picks up most of its FSCN out at its far edge, and within the inner 10º. A prime-focus dish, on the other hand, exhibits a steadily decreasing FSCN as a function of radius.

These differences can be understood by studying the beamshapes in Fig. 1. For the OG, the jump in FSCN between 70º~90º is probably associated with the elongated "hot spot" at the top of the pattern, which is caused by illumination spillover inherent to the design. For the PF dish, most of FSCN comes in via the strong X-shaped sidelobes associated with diffraction off the feed box struts, which remain significant at all radii.

There is a cross-over point at ~30º, beyond which PF dishes start outperforming OGs in terms of FSCN; RR designs reach this cross-over somewhat earlier. If the required FSCN level can be achieved before cross-over, OG designs can be considerably cheaper in terms of imaging cost. The default OG design (almost) reaches MeerKAT's .1 µJy threshold by 10º; the default PF reaches it only at 30º. Since imaging costs scale as the square of the radius, this represents almost an order of magnitude difference in computing cost. Beyond that, the OG curve plateaus: going deeper, one is therefore better off with a PF.

Considerations like this ought to inform SKA dish design choices. Note, however, that the curves given here only apply to the one specific interferometer configuration of MeerKAT. Since FSCN scales with PSF sidelobe level (which decreases with bigger arrays/better $uv$-coverage) and also decreases on longer baselines due to averaging, an SKA-sized interferometer will certainly show a different set of curves, and thus requires separate simulations.

A detailed analysis of FSCN curves will be presented in a follow-up work: the aim of the present paper is only to highlight the methodology used.

## 4 POINTING ERROR RESULTS

To isolate PE-related calibration artefacts, we proceed as follows. We simulate a deep 2º×2º field containing a "known" bright sky (for which we use a

full NVSS sampling, down to its limiting flux of ~2.7 mJy), an "unknown" faint sky (for which we use the S³-SEX simulation [2]), and two "contaminator" sources of 1 Jy each, at approximately the half-power point of the PB, one due South, and one due NE of the field centre. Bright sources on the half-power point are a known worst-case scenario for PE.

We then do a complete interferometric simulation of the field, with realistic PEs, thermal noise, and antenna gain (G-Jones [4]) variations. We call the resulting visibilities the *full data*. We also do a simulation of the contaminator sources by themselves, without noise, but with exactly the same values for pointing errors and gains. This we call the *contaminator-only* data.

Next, we do a standard G-Jones solution (i.e. regular selfcal) on the full data, using the "known" sky plus the contaminators as the prior sky model. The resulting G solutions are somewhat different from the "true" G-Jones that went into the simulation due to the effect of pointing error, the unknown faint sky, and thermal noise. Applying these solutions while subtracting the sky model results in an imperfect subtraction. The residual dirty map then contains the faint sky, calibration artefacts left over from the subtraction, and thermal noise.

We can now "distil out" the calibration artefacts associated with the two contaminator sources as follows: we take the contaminator-only data, and subtract a model of the two contaminators while applying our G solutions. The resulting image then contains the "pure" artefacts (Fig. 3a).

Next, we run a dE-Jones solution [4] on the contaminator sources, and subtract them from the full data using both the G and the dE solutions. This produces a much better residual dirty map: in fact only the faint sky and the thermal noise is apparent; any artefacts associated with the contaminators are not visible above the noise. We can still "distil out" these artefacts though, by repeating the procedure above. This results in an image that appears to be pleasantly noise-like, and much-improved compared to the one after G-Jones calibration only (Fig. 3b).
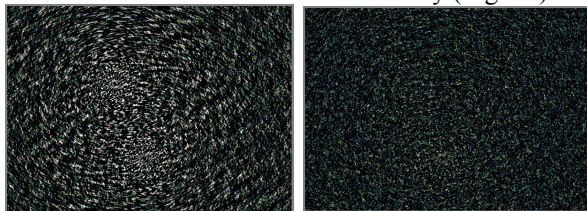


Figure 3a,b. Distilled calibration artefacts after (a, left) G-Jones and (b, right) dE-Jones solutions. OG beamshape, 10" pointing error. Intensity scale is from 0 to 50 μJy.

The crucial question is the following: with repeated observations of the field, how do the artefacts of Fig. 3b add up statistically? To answer this, we simply repeat the whole procedure many times with different realizations of antenna gains and pointing errors, and average up the artefact maps that this produces.
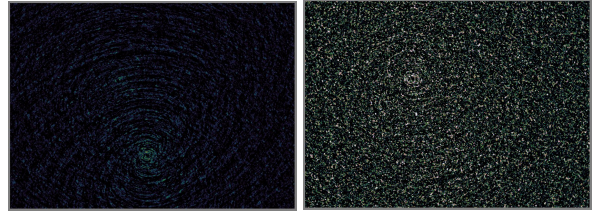


Figure 4a,b: Mean dE artefacts after 10 observations. (a, left) OG beamshape, (b, right) PF beamshape.

The results of this are somewhat alarming (Fig. 4). The artefacts add up coherently, and do not look like noise at all any more. This implies a real dynamic range limit due to pointing error, which in the OG case works out to about $10^6$:1. Interestingly, PF dishes perform roughly a factor of three worse (Fig. 4b), which is probably explained by the fact that their beamshape is less smooth.

In the case of MeerKAT, such dynamic range limits are really only a worry at the deep end of operations. The direct implication is that deep surveys should avoid fields with strong off-axis sources, and should be preferentially scheduled for times of low wind. At SKA sensitivities, however, PE can become an everyday dominant limiting factor, so further study of this problem is urgently needed.

## Acknowledgments

## References

[1] A.R. Thompson, J.M. Moran, and G.W. Swenson Jr, "Interferometry and Synthesis in Radio Astronomy", Wiley, 2nd ed., 2001, Sect. 6.2
[2] R.J. Wilman et al., Mon. Not. R. Astron. Soc., vol. 388, 2008, p. 1335
[3] M.J. Jarvis, in preparation
[4] O.M. Smirnov, Astron. Astrophys., vol. 527, 2011, id. A108
[5] S. Kazemi et al. , Mon. Not. R. Astron. Soc., vol. 414, 2011, p. 1656
[6] S. Bhatnagar, T.J. Cornwell and K. Golap, NRAO, EVLA Memo 84, 2004
[7] S. Bhatnagar et al., Astron. Astrophys., vol. 487, 2008, p. 419
[8] J.E. Noordam and O.M. Smirnov, Astron. Astrophys., vol. 524, 2010, id. A61
[9] J.J. Condon et al., Astron. J., vol. 115, 1998, p. 1693