Astronomy & Astrophysics manuscript no. paper
(DOI: will be inserted by hand later)
April 22, 2015

Tapering, DDE, and PSFs

[1] Department of Physics & Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa
[2] SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands, 7405, South Africa
[3] GEPI, Observatoire de Paris, CNRS, Université Paris Diderot, 5 place Jules Janssen, 92190 Meudon, France

Received / Accepted

Abstract.

1. Introduction

2. Motivation

3. Matrix formulation of the problem - 1D interferometer

Here I intend to use convolution matrices properties to *qualitatively* study how "pseudo-PSF" vary as a function of source location. Here I limit myselves to a 1-dimensional interferometer (scalar only), so that Convolution matrices are Toeplitz-symetyric (see bellow). In a more general case, (along my intuition - but should be thought more carefully), convolution matrices should be block-Toeplitz (each block is a Toeplitz), while symetricity should still be true.

3.1. Remarks on the convolution and linear algebra

In functional form the convolution theorem can be written as follows:

$$\mathcal{F}\left\{a.b\right\} = \mathcal{F}\left\{a\right\} * \mathcal{F}\left\{b\right\} \tag{1}$$

Noting the convolution product is linear, we can re-express the convolution product and associated theorem asing linear transformations:

$$\boldsymbol{F}\boldsymbol{A}\boldsymbol{b} = \boldsymbol{\mathcal{C}_A}\boldsymbol{F}\boldsymbol{b} \tag{2}$$

where $\boldsymbol{F}$ is the Fourier operator of size $n_{uv} \times n_{lm}$ ($\boldsymbol{F}$ is unitary $\boldsymbol{F}^H\boldsymbol{F} = \boldsymbol{1}$), $\boldsymbol{b}$ is a vector with size n_{lm}. The matrix $\boldsymbol{A}$ models the scalar multiplication of each point in $\boldsymbol{b}$, and is therefore diagonal of size $n_{lm} \times n_{lm}$, and $\boldsymbol{\mathcal{C}_A}$ is the convolution matrix of size $n_{uv} \times n_{uv}$. There is a bijective relation

$$\boldsymbol{A} \longleftrightarrow \boldsymbol{\mathcal{C}_A} \tag{3}$$

in the sense that a scalar multiplucation defines a convolution function and conversely. The matrices $\boldsymbol{A}$ and $\boldsymbol{\mathcal{C}_A}$ always have the following properties:

- $\boldsymbol{A}$ is diagonal
- In the 1D case
  - $\boldsymbol{\mathcal{C}_A}$ is Toeplitz
  - In addition, for radiointerferometry, because the uv plane is symetric, $\boldsymbol{\mathcal{C}_A}$ is symetric

The matrix $\boldsymbol{\mathcal{C}_A}$ being Toeplitz, each row $[\boldsymbol{\mathcal{C}_A}]_l$ with sky coordinate l can be built using a rolling operator Δ_l that shifts the first row (the PSF at the field center for example) to location of row l:

$$[\boldsymbol{\mathcal{C}_A}]_l = \Delta_l \left\{[\boldsymbol{\mathcal{C}_A}]_0\right\} \, and \tag{4}$$

$$[\boldsymbol{\mathcal{C}_A}]_0 = \boldsymbol{F}^H \ \mathrm{diag}(\boldsymbol{A}) \tag{5}$$

The rolling operator is essentially just a reindexing, and has the following properties:

$$\Delta_l \left\{a\boldsymbol{x}\right\} = a\Delta_l \left\{\boldsymbol{x}\right\} \tag{6}$$

$$\Delta_l \left\{\sum_i \boldsymbol{x}_i\right\} = \sum_i \Delta_l \left\{\boldsymbol{x}_i\right\} \tag{7}$$

3.2. PSF behaviour

If $\boldsymbol{X}$ is the true sky, then the dirty image $\boldsymbol{X}^D_{ij}$ of baseline (ij) can be written as:

$$\boldsymbol{x}^D_{ij} = \boldsymbol{F}^H \boldsymbol{S}_{c,ij} \boldsymbol{\mathcal{C}_T} \boldsymbol{S}_{\Box,ij} \boldsymbol{F} \boldsymbol{A} \boldsymbol{x} \tag{8}$$

where $\boldsymbol{A}_{ij}$ models the DDE effets and is an $n_{pix} \times n_{pix}$ diagonal matrix (taking polarisation into account it is an $4n_{pix} \times 4n_{pix}$ block diagonal matrix), $\boldsymbol{T}$ is the tapering/averaging function, $\boldsymbol{S}_\Box$ samples the region over which the tapering/averaging is made, and $\boldsymbol{S}_{c,ij}$ selects the central point of the averaged/tapered visibility set. Using Eq. 4, we have:

$$\begin{aligned}
\boldsymbol{x}_{ij}^{D} &= \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{T} \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{\square,ij}} \boldsymbol{F}^{H} \boldsymbol{F} \boldsymbol{A}_{ij} \boldsymbol{x} && (9)\\
&= \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{T} \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{\square,ij}} \boldsymbol{A}_{ij} \boldsymbol{x} && (10)\\
&\sim \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{T} \boldsymbol{A}_{ij} \boldsymbol{x} && (11)
\end{aligned}$$

where Eq. 11 is true when the support of the function T is smaller than the sampling domain of $\boldsymbol{S}_{\square}$.

Averaged over all baselines, the dirty image becomes:

$$\begin{aligned}
\boldsymbol{x}^{D} &= \boldsymbol{\mathcal{C}}_{STA} \boldsymbol{x} && (12)\\
\text{with } \boldsymbol{\mathcal{C}}_{STA} &= \sum_{ij} \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{T} \boldsymbol{A}_{ij} && (13)
\end{aligned}$$

3.3. Deriving the Pseudo-PSF

3.3.1. PSF and Pseudo-PSF

We can already see that $\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{T} \boldsymbol{A}_{ij}$ in Eq. 11 is NOT Toeplitz anymore because each colunm is multiplied by a different value (DDE muliplied by the tapering function). The dirty sky is therefore not anymore the convolution of the true sky by the psf *ie* the PSF varies accross the field of view.

3.3.2. Slow way

Calculate the psf estimating $\mathcal{C}$ from direct calculation. Eventually at discrete locations on a grid.

3.3.3. Quickly deriving the Pseudo-PSF

This is tricky part. The problem amount to finding any column l of $\boldsymbol{\mathcal{C}}$ on demand. For notation convenience, we merge $\boldsymbol{T}$ and $\boldsymbol{A}_{ij}$ together in $\boldsymbol{A}_{ij}$. Operator $[\boldsymbol{M}]_l$ extracts column l from matrix $\boldsymbol{M}$, and using Eq. 6, 7 and 12:

$$\begin{aligned}
[\boldsymbol{\mathcal{C}}]_l &= \left[\sum_{ij} \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{A}_{ij}\right]_l && (14)\\
&= \sum_{ij} a_{ij}^{l} \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_l && (15)\\
&\quad \text{with } a_{ij}^{l} = \boldsymbol{A}_{ij}(l) && (16)\\
&= \sum_{ij} \Delta_l \left\{a_{ij}^{l} \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_0\right\} && (17)\\
&= \sum_{ij} \Delta_l \left\{\boldsymbol{F}^{H} a_{ij}^{l} \operatorname{diag}(\boldsymbol{S}_{c,ij})\right\} && (18)
\end{aligned}$$

If we now assume that at any given location l, the scalar a_{ij}^{l} can be described by a smooth *function* of the uv coordinates ((ij)-indices), then we can write:

$$\begin{aligned}
[\boldsymbol{\mathcal{C}}]_l &= \sum_{ij} \Delta_l \left\{\boldsymbol{F}^{H} \boldsymbol{A}^{l} \operatorname{diag}(\boldsymbol{S}_{c,ij})\right\} && (19)\\
&= \sum_{ij} \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^{l}}\ \boldsymbol{F}^{H} \operatorname{diag}(\boldsymbol{S}_{c,ij})\right\} && (20)\\
&= \sum_{ij} \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^{l}}\ \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_0\right\} && (21)\\
&= \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^{l}} \sum_{ij} \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_0\right\} && (22)\\
&= \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^{l}}\ \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c}}\right]_0\right\} && (23)\\
& && (24)
\end{aligned}$$

The approximate observed Pseudo-PSF is the convolution of the PSF at the phase center ($[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_c}]_0$) and the fourier transform of the uv-dependent tapering function at given lm ($\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^{l}}$).

In other words, to compute the PSF at a given location (lm):

- Find $\boldsymbol{A}$:
  - Compute weight w_{ij} for each baseline (ij)
  - Fit the uv-dependent weight by (for example), a Gaussian function $w_{ij} \sim w(u,v) = \boldsymbol{\mathcal{G}}(u,v)$
- Compute the PSF_{lm} at (lm) from the PSF at the phase center PSF_0 as $PSF_{lm} = \mathcal{F}^{-1}(w) * PSF_0$

For example if the long baselines are more tapered, they are "attenuated". The effective PSF on the edge of the field will get larger by the convolution... Something like that...

4. Numerical Experiments

We demonstrate the computational complexity of the quick, the slow derived PSF as a function of sky coordinates and perform a direct numerical results.

4.1. Slow derivation

4.2. Quick derivation

In order to further optimize the slow derivation of the PSF described above to particularly reduce its computational cost, we will need to understand the concept and theory of signal correlation in aperture synthesis. It is worth noting that in the Radio Astronomy community the cross-correlator output of two elements interferometer in response to a source with spectral brightness distribution $I_\nu(\mathbf{s})$ as a function of the pointing direction $\mathbf{s}$ is the visibility function defined in Eq.25 and obtained by integrating over the solid angle $d\Omega$ (Thompson et al. 2008; Taylor et al. 1999)

$$V(\mathbf{b}) = \int_{\Omega} I_\nu(\mathbf{s}) e^{-2\pi i \mathbf{b}\mathbf{s}} d\Omega, \qquad (25)$$

where $\mathbf{b} = \mathbf{b}(t,\nu) = (u_{t\nu}, v_{t\nu}, w_{t\nu})$ is the so-called "*baseline vector*" in wavelength and with modulus the distance between the two elements interferometer. The measurement in Eq.25 is over the entire surface of the celestial sphere, practically the measurement is generally taken over a finite surface area of the celestial sphere due to the finite nature of the tracking source and other effects. Furthermore, the result is averaged over finite time/frequency bins. Suppose Δt centered at t_0 and $\Delta\nu$ centered at ν_0 the time and frequency sampling intervals respectively. Assuming that Δt and $\Delta\nu$ are small enough so that $I_\nu(\mathbf{s})$ remains constant while the complex phase, $-2\pi i\mathbf{bs}$ varies linearly. Eq. 25 becomes:

$$V(t_0,\nu_0) = \frac{1}{\Delta t\Delta\nu}\iint\limits_{\Delta t\Delta\nu}\left[\int\limits_{\Omega} I_\nu(\mathbf{s})e^{-2\pi i\mathbf{bs}}d\Omega\right]dtd\nu. \tag{26}$$

However, performing the integration over a surface of size, $\Delta t\times\Delta\nu$ yields a nonzero value only in this surface area, which is recognized as integrating of a windowing function $\Pi_{t_0,\nu_0}(t,\nu)$ of size $\Delta t\times\Delta\nu$. Let $\mathbf{b}_0 = \mathbf{b}(t_0,\nu_0) = (u_{t_0\nu_0}, v_{t_0\nu_0}, w_{t_0\nu_0})$ be the baseline vector at the centre of the sampling intervals and $W_{t_0,\nu_0}(t,\nu)$ a weighting sampling function. A generalized sampling kernel for a given (u,v) track, $W_{\mathbf{b}_0}(\mathbf{b})$ associated with the windowing function can be defined as

$$W_{\mathbf{b}_0}(\mathbf{b}) = \Pi_{t_0,\nu_0}(t,\nu)W_{t_0,\nu_0}(t,\nu) \tag{27}$$

$$= \Pi(\mathbf{b}-\mathbf{b}_0)W(\mathbf{b}-\mathbf{b}_0). \tag{28}$$

We can pose, $\frac{dtd\nu}{dudv} = \frac{\Delta t\Delta\nu}{\Delta u\Delta v}$ then Eq.26 can be written in terms of $\mathbf{b}_0$ and can thus be expressed as an integration over uv-bins. We can write

$$V_{\mathbf{b}_0}(\mathbf{x}) = \frac{1}{\Delta u\Delta v}\iint\limits_{uv} W_{\mathbf{b}_0}(\mathbf{b})\left[\int\limits_{\Omega} I_\nu(\mathbf{s})e^{-2\pi i\mathbf{bs}}d\Omega\right]dudv, \tag{29}$$

where, $\mathbf{x} = (u,v,w)$ the averaged track. As previously mentioned, we are interested in PSF response. Therefore, Eq.29 is restricted to the complex visibility measured by the two-element interferometer for a point source located towards the direction $\mathbf{s}$ with unit brightness. That said, $I_\nu(\mathbf{s}) = \delta(\mathbf{b}_0)$ with δ the Dirac delta. It then follows that Eq29 can be written as

$$V_{\mathbf{b}_0}(\mathbf{x}) = \frac{1}{\Delta u\Delta v}\iint\limits_{uv} W_{\mathbf{b}_0}(\mathbf{b})e^{-2\pi i(u_{t\nu}l+v_{t\nu}m+w_{t\nu}(n-1))}dudv. \tag{30}$$

Note that Eq.30 is obtained after a delay correction of $\mathbf{bs_0}$ is applied to the signals from the antennas array to steer towards the direction $\mathbf{s}_0$ and $\mathbf{b}(\mathbf{s}-\mathbf{s_0}) = u_{t\nu}l + v_{t\nu}m + w_{t\nu}(n-1)$ describes the time difference between the two incoming signals. The three direction cosines l, m and n are components of $\mathbf{s}-\mathbf{s_0}$ in radians with $n = \sqrt{1-l^2-m^2}$. For an extensive discussion, see (Thompson et al. 2008; Taylor et al. 1999; Leshem & van der Veen 2000). Let $G_{\mathbf{b}}(l_0,m_0) = e^{-2\pi i[w_{t\nu}(n-1)]}$ the fringe from the w-term. The w-term depends on the variation of u and v. In other words, $w_{t\nu} = (\gamma + \alpha du + \beta dv)$. In this paper, we consider the w-term in the general case where the array is coplanar ($w_{t\nu} = 0$) or not ($w_{t\nu} \neq 0$) and/or when the array is tracking a small field ($n \approx 1$) or a wide field ($n \ll 1$). The fringe from the w-term is separated; the measurement in Eq30 becomes a two dimensional Fourier transform. Making use of the convolution theorem which states that the Fourier transform of the product of two functions result in a convolution and vice-versa [1], we have

$$V_{\mathbf{b}_0}(\mathbf{x}) = \widetilde{G}_{\mathbf{b}}(l,m)\circ\iint\limits_{uv} W_{\mathbf{b}_0}(\mathbf{b})e^{-2\pi i(u_{t\nu}l+v_{t\nu}m)}dudv. \tag{31}$$

Also, $W_{\mathbf{b}_0}(\mathbf{b})$ is defined as a product of two functions (refer to Eq. 28), making use of the shifted Fourier transform properties and the convolution theorem, the following expansion is valid

$$V_{\mathbf{b}_0}(\mathbf{x}) = \left(\widetilde{G}_{\mathbf{b}}\circ\left[C_{\mathbf{b}_0}\widetilde{\Pi}_{\mathbf{b}}\right]\circ\left[C_{\mathbf{b}_0}\widetilde{W}_{\mathbf{b}}\right]\right)(l,m) \tag{32}$$

The symbols $\circ$ and $\tilde{}$ denote the convolution operator and the Fourier transform respectively. The phase gradient, $C_{lm}(\mathbf{b}_0)$ is defined as

$$C_{lm}(\mathbf{b}_0) = e^{-2\pi i(u_{t_0\nu_0}l+v_{t_0\nu_0}m)} \tag{33}$$

If it is assumed that $\Pi(\mathbf{b})$ is the top-hat windowing function, then it can be demonstrated that its Fourier transform is (see Appendix A)

$$\widetilde{\Pi}(\mathbf{b}) = sinc\Big(\frac{-2\pi t\Delta\nu}{2}\Big)sinc\Big(\frac{-2\pi\nu\Delta t}{2}\Big) \tag{34}$$

It is demonstrated in (Smirnov 2011) that for natural weighting, Eq.31 can be approximated in term of the phase change in time, $\Delta\Psi$ and in frequency $\Delta\Phi$ for the case of smearing as:

$$V_{\mathbf{b}_0}(\mathbf{x}) \simeq \left[C_{\mathbf{b}_0}(l,m)sinc\frac{\Delta\Psi}{2}sinc\frac{\Delta\Phi}{2}\right]\circ\left(\widetilde{G}_{\mathbf{b}_0}\circ\left[C_{\mathbf{b}_0}\widetilde{W}_{\mathbf{b}_0}\right]\right)(l,m) \tag{35}$$

where $\Delta\Psi$ and $\Delta\Phi$ are defined as

$$\begin{aligned}\Delta\Psi = & 2\pi\Big[(u_{t_s\nu_0}-u_{t_e\nu_0})l + (v_{t_s\nu_0}-v_{t_e\nu_0})m \\ & + (w_{t_s\nu_0}-w_{t_e\nu_0})(n-1)\Big],\end{aligned} \tag{36}$$

$$\begin{aligned}\Delta\Phi = & 2\pi\Big[(u_{t_0\nu_s}-u_{t_0\nu_e})l + (v_{t_0\nu_s}-v_{t_0\nu_e})m \\ & + (w_{t_0\nu_s}-w_{t_0\nu_e})(n-1)\Big],\end{aligned} \tag{37}$$

where t_s, t_e, ν_s and ν_e are the starting time, ending time, starting frequency and ending frequency respectively of

[1] *http://mathworld.wolfram.com/ConvolutionTheorem.html*
http://en.wikipedia.org/wiki/Convolution_theorem

the sampling intervals.

We then generalized the approximation of smearing where $\Pi(\mathbf{b})$ is a random windowing function as follows:

$$V_{\mathbf{b}_0}(\mathbf{x}) \simeq \left(\widetilde{G}_{\mathbf{b}_0} \circ \left[C_{\mathbf{b}_0} \widetilde{\Pi}_{\mathbf{b}_0} \right] \circ \left[C_{\mathbf{b}_0} \widetilde{W}_{\mathbf{b}_0} \right] \right) (l, m) \quad (38)$$

4.3. Imaging

During conventional interferometric imaging, the baseline visibilities are mapped on a uv-plane and the result is inverse Fourier transformed

$$I_D(l,m) = I_D(l - l_0, m - m_0) \quad (39)$$

$$= \mathcal{F}^{-1} \left\{ \sum_{\mathbf{b}_0} S_{\mathbf{b}_0} V_{\mathbf{b}_0} \right\}, \quad (40)$$

where $\mathcal{F}^{-1}$ represents the inverse Fourier transform and $S_{\mathbf{b}_0}$ is the sampling function of the baseline pq. The sampling function is only nonzero in the neighborhood of the track described by $\mathbf{b}_0$ and account for the interpolation coefficients of the gridding process.

$$V_{\mathbf{b}_0} = e^{-2i\pi \mathbf{b}_0 \mathbf{s}_0} \quad (41)$$

for a source at $\mathbf{s}_0$.

$$I_D(l,m) = \mathcal{F}^{-1} \left\{ \sum_{\mathbf{b}_0} S_{\mathbf{b}_0} \right\} \circ \delta \quad (42)$$

$$= P(l,m) \circ \delta \quad (43)$$

In the above section, an approximate visibility, $V_{\mathbf{b}_0}(\mathbf{x}_0)$ is predicted for each baseline corresponding to its averaged uv track and account in our case for both the observed *source location*, the windowing kernel and the imaging weights.

$$V_{\mathbf{b}_0} = f(\mathbf{b}_0) \quad (44)$$

$$I_D(l,m) = \mathcal{F}^{-1} \left\{ \sum_{\mathbf{b}_0} S_{\mathbf{b}_0} V_{\mathbf{b}_0} \right\} \quad (45)$$

$$= \sum_{\mathbf{b}_0} \mathcal{F}^{-1} \left\{ S_{\mathbf{b}_0} f(\mathbf{b}_0) \right\} \quad (46)$$

4.3.1. Computational cost

5. Simulation and comparison

6. Discussion and conclusion

Acknowledgements. No thanks

References

Leshem A., van der Veen A.-J., 2000, Information Theory, IEEE Transactions on, 46, 1730

Smirnov O. M., 2011, Astronomy & Astrophysics, 527, A106

Taylor G. B., Carilli C. L., Perley R. A., 1999, in Synthesis Imaging in Radio Astronomy II Vol. 180, Synthesis imaging in radio astronomy ii

Thompson A. R., Moran J. M., Swenson Jr G. W., 2008, Interferometry and synthesis in radio astronomy. John Wiley & Sons

Appendix A: The Fourier Transform of a top-hat windowing function

A two dimensional top-hat window is defined as

$$\mathbf{\Pi}(\mathbf{b}) = \begin{cases} 1 & \text{for } t \times \nu \in [t_s, t_e] \times [\nu_s, \nu_e], \\ 0 & \text{for } t \times \nu \notin [t_s, t_e] \times [\nu_s, \nu_e] \end{cases} \quad (A.1)$$

where $[t_s, t_e] \times [\nu_s, \nu_e]$ is the passband region.

$$\widetilde{\Pi}(\mathbf{b}) = \frac{1}{\Delta t \Delta \nu} \int_{t_s}^{t_e} \int_{\nu_s}^{\nu_e} e^{-2\pi i t \nu} dt d\nu$$

Let suppose that

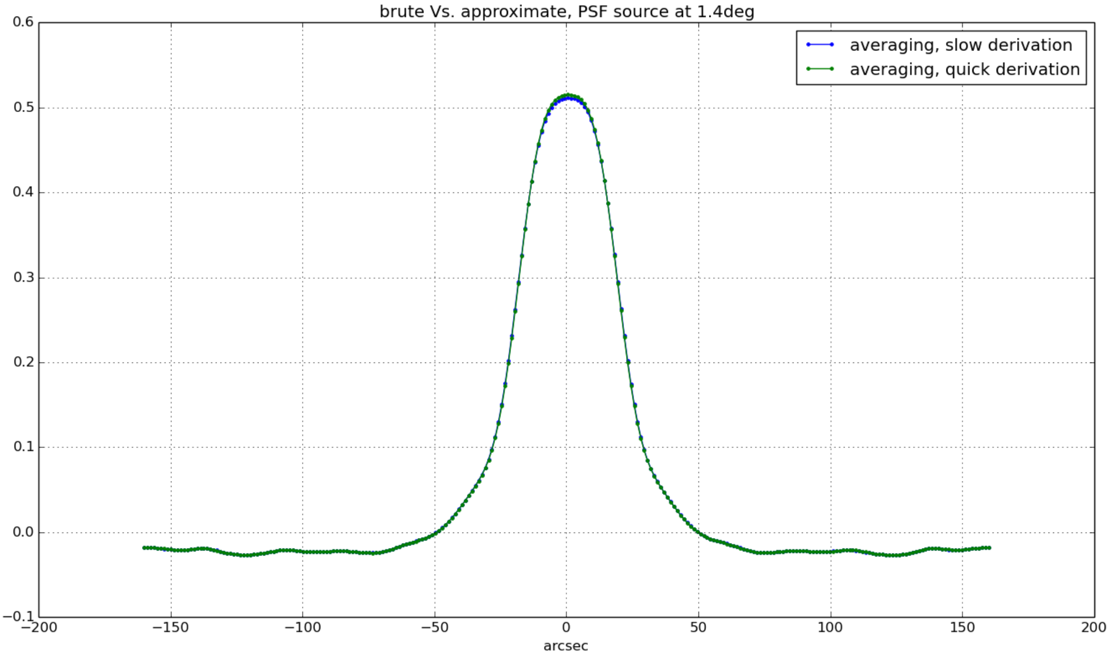


Fig. 1. Brute Vs approximate PSF for a source at 1.4deg

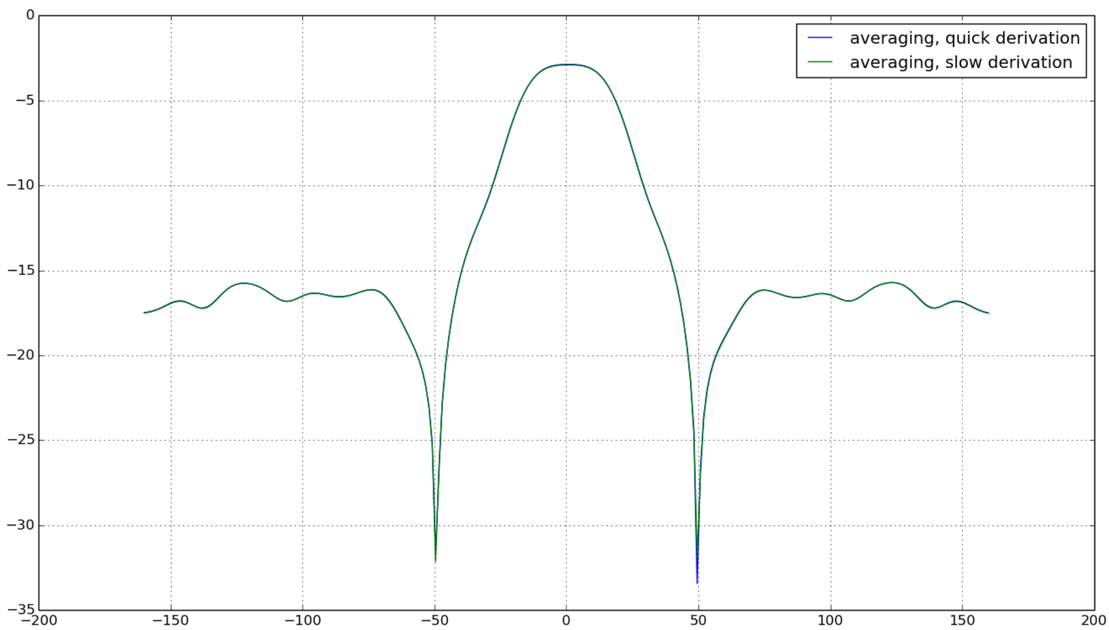
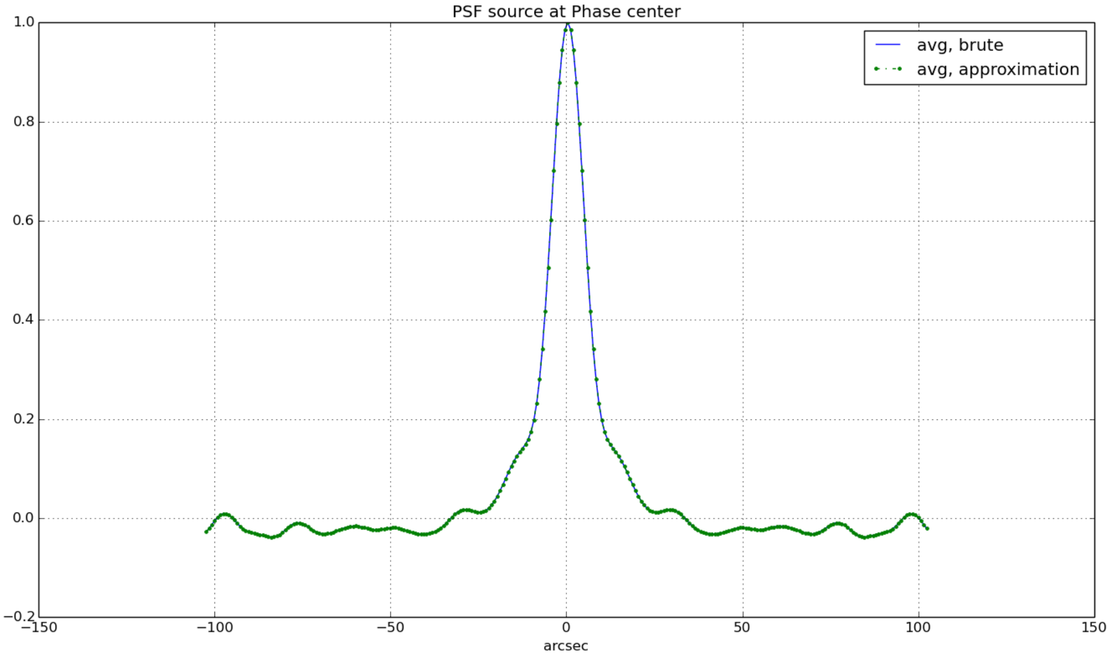


Fig. 2. Brute Vs approximate PSF for a source at 1.4deg compare in log space


PSF source at Phase center
avg, brute
avg, approximation
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−150
−100
−50
0
50
100
150
arcsec