Astronomy & Astrophysics manuscript no. paper March 3, 2015
(DOI: will be inserted by hand later)

Tapering, DDE, and PSFs

Address(es) of author(s) should be given

Received ¡date¿ / Accepted ¡date¿

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:

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

Noting the convolution product is linear, we can reexpress the convolution product and associated theorem asing linear transformations:

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

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 $\boldsymbol{\mathcal{C}}_{\boldsymbol{A}}$ is the convolution matrix of size $n_{uv} \times n_{uv}$. There is a bijective relation

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

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

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

The matrix $\boldsymbol{\mathcal{C}}_{\boldsymbol{A}}$ being Toeplitz, each row $[\boldsymbol{\mathcal{C}}_{\boldsymbol{A}}]_l$ 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:

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

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

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

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

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

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:

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

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:

$$\boldsymbol{x}^D_{ij} = \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\boldsymbol{T}\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{\Box,ij}}\boldsymbol{F}^H\boldsymbol{F}\boldsymbol{A}_{ij}\boldsymbol{x} \tag{9}$$

$$= \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\boldsymbol{T}\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{\Box,ij}}\boldsymbol{A}_{ij}\boldsymbol{x} \tag{10}$$

$$\sim \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\boldsymbol{T}\boldsymbol{A}_{ij}\boldsymbol{x} \tag{11}$$

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:

$$\boldsymbol{x}^D = \boldsymbol{\mathcal{C}}_{STA}\boldsymbol{x} \tag{12}$$

$$\text{with } \boldsymbol{\mathcal{C}}_{STA} = \sum_{ij} \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{T}\boldsymbol{A}_{ij} \tag{13}$$

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:

$$[\boldsymbol{\mathcal{C}}]_l = \left[\sum_{ij} \boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}} \boldsymbol{A}_{ij}\right]_l \tag{14}$$

$$= \sum_{ij} a_{ij}^l \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_l \tag{15}$$

$$\text{with } a_{ij}^l = \boldsymbol{A}_{ij}(l) \tag{16}$$

$$= \sum_{ij} \Delta_l \left\{a_{ij}^l \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_0\right\} \tag{17}$$

$$= \sum_{ij} \Delta_l \left\{\boldsymbol{F}^H a_{ij}^l \operatorname{diag}\left(\boldsymbol{S}_{c,ij}\right)\right\} \tag{18}$$

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:

$$[\boldsymbol{\mathcal{C}}]_l = \sum_{ij} \Delta_l \left\{\boldsymbol{F}^H \boldsymbol{A}^l \operatorname{diag}\left(\boldsymbol{S}_{c,ij}\right)\right\} \tag{19}$$

$$= \sum_{ij} \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^l}\ \boldsymbol{F}^H \operatorname{diag}\left(\boldsymbol{S}_{c,ij}\right)\right\} \tag{20}$$

$$= \sum_{ij} \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^l}\ \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_0\right\} \tag{21}$$

$$= \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^l} \sum_{ij} \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_{c,ij}}\right]_0\right\} \tag{22}$$

$$= \Delta_l \left\{\boldsymbol{\mathcal{C}}_{\boldsymbol{A}^l}\ \left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_c}\right]_0\right\} \tag{23}$$

$$\tag{24}$$

The approximate observed Pseudo-PSF is the convolution of the PSF at the phase center ($\left[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_c}\right]_0$) 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

We will now show how to derived a pseudo PSF which is based and resolved on the nominal PSF but labelled by a set of band-limited integration.

4.2.1. Averaging case

In order to further optimize the slow derivation of the PSF described above to particularly reduce the computational cost, we will need to understand the concept and theory of signal correlation in aperture synthesis. It is worth noting in Radio Astronomy community that 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 $d\Omega$ (see Thompson, annnff, ahhhh, jjdkk)

$$V_\nu(\mathbf{b}) = \int_\Omega I_\nu(\mathbf{s}) e^{\frac{-2\pi\nu i \mathbf{bs}}{c}} d\Omega, \tag{25}$$

where $\mathbf{b}$ is the so called "*baseline vector*" with module the distance between the two elements interferometer and c is the speed of the light. The measurement in Eq.25 is over the entire surface of the celestial sphere, in practice 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. Assuming that the frequency ν and the time t change small enough so that $I_\nu(\mathbf{s})$ remains constant while the complex phase, $\frac{-2\pi\nu i \mathbf{bs}}{c}$ varies linearly. Eq.25 becomes:

$$V_\nu(\mathbf{b}) = \frac{1}{\Delta t \Delta\nu} \int_{\Delta\nu}\int_{\Delta t} \left[\int_\Omega I_\nu(\mathbf{s}) e^{\frac{-2\pi\nu i \mathbf{bs}}{c}} d\Omega \right] dt d\nu \tag{26}$$

$$= \iint_{-\infty}^{+\infty} f_{\mathbf{s}}(\mathbf{b}) \left[\int_\Omega I_\nu(\mathbf{s}) e^{\frac{-2\pi\nu i \mathbf{bs}}{c}} d\Omega \right] dt d\nu. \tag{27}$$

Here, $f_{\mathbf{s}}(\mathbf{b})$ is the visibilities plane tapering kernel, theoretically it is an elliptical function that maps the ellipse formed by the baseline. Eq.27 allows us to completely understand that in theory, the integration is over a continuous elliptical arc both in frequency and time. In the rest of this paper, Eq.27 is restricted to the complex visibility measured by the two elements interferometer for a point source located towards the direction $\mathbf{s}$ with unitary brightness. A delay correction of $\frac{\mathbf{bs_0}}{c}$ is generally applied to the signals from the antennas array to steer towards the direction $\mathbf{s_0}$. Therefore, the scalar product $\frac{\nu \mathbf{b}(\mathbf{s}-\mathbf{s_0})}{c} = u_{t\nu} l + v_{t\nu} m + w_{t\nu}(\sqrt{1-l^2-m^2}-1)$ describes the time difference between the two incoming signals. Note that, $u_{t\nu}, v_{t\nu}, w_{t\nu}$ are the components of $\mathbf{b}$ given in wavelength and $l, m, \sqrt{1-l^2-m^2}-1$ are the components of $\mathbf{s}-\mathbf{s_0}$ given in radian, describing the three directions cosine. We can write:

$$V_\nu(\mathbf{b}) = \iint_{-\infty}^{+\infty} f_{lm}(u_{t\nu}, v_{t\nu}) e^{-2\pi i(u_{t\nu} l + v_{t\nu} m + w_{t\nu}(n-1))} dt d\nu \tag{28}$$

$$= \check{f}_{lm}(u_{t\nu}, v_{t\nu}) \circ \delta(u_{t\nu}, v_{t\nu}) \tag{29}$$

where $\check{f}_{lm}$ is the Fourier transform of f_{lm} and $\delta(u_{t\nu}, v_{t\nu})$ the delta Dirac. In practice, the variation is limited to a small finite frequency range, $\Delta\nu$ centered at ν_c and time range Δt centered at t_c, which implies that we are not integrating over continuous elliptical arcs but on shorts discrete segments across time and frequency. In this case, $f_{lm}(u_{t\nu}, v_{t\nu})$ is trouncated over a small region, $\Pi_{lm}(u_{t\nu}, v_{t\nu})$ of size $\Delta t \times \Delta\nu$ defined as

$$\Pi_{lm}(u_{t\nu}, v_{t\nu}) = \begin{cases} 1 & \text{for } t \times \nu \in [t_s, t_e] \times [\nu_s, \nu_e], \\ 0 & \text{for } t \times \nu \neq [t_s, t_e] \times [\nu_s, \nu_e] \end{cases} \tag{30}$$

where t_s, t_e, ν_s and ν_e are the starting times, ending time, starting frequency and ending frequency respectively. The Fourier transform, $\check{\Pi}_{lm}(u_{t\nu}, v_{t\nu})$ of Eq.30 is given by

$$\check{\Pi}_{lm}(u_{t\nu}, v_{t\nu}) = sinc\frac{-2\pi t \Delta\nu}{2} sinc\frac{-2\pi\nu\Delta t}{2} \check{\delta}_{lm}(u_{t\nu} + u_{t_c\nu_c}, v_{t\nu} + v_{t_c\nu_c}). \tag{31}$$

Let supose that

$$f^{\Pi}_{lm}(u_{t\nu}, v_{t\nu}) = \Pi_{lm}(u_{t\nu}, v_{t\nu}) w(u_{t\nu}, v_{t\nu}) f_{lm}(u_{t\nu}, v_{t\nu}),$$

is the trouncated version of $f_{lm}(u_{t\nu}, v_{t\nu})$ including a weighting kernel $w(u_{t\nu}, v_{t\nu})$ with Fourier transform $\check{w}(u_{t\nu}, v_{t\nu})$. Following the analogy, $f^{\Pi}_{lm}(u_{t\nu}, v_{t\nu})$ is of size $\Delta t \times \Delta\nu$. We have:

$$V_\nu(\mathbf{b}) = \check{f}^{\Pi}_{lm}(u_{t\nu}, v_{t\nu}) \circ \delta(u_{t\nu}, v_{t\nu}) \tag{32}$$

$$= \left[\check{\Pi}_{lm}(u_{t\nu}, v_{t\nu}) \circ \check{w}(u_{t\nu}, v_{t\nu}) \circ \check{f}_{lm}(u_{t\nu}, v_{t\nu}) \right] \circ \delta(u_{t\nu}, v_{t\nu}) \tag{33}$$

It is then observed that for finite frequency and time range, the complex phase varies linearly within the range $\Delta\Phi = 2\pi t \Delta\nu$ and $\Delta\Psi = 2\pi\nu\Delta t$ in frequency and time respectively.

O.M. Smirnov (cite, Jan 2011) demonstrated that in the case of averaging, natural weighting where $f^{\Pi}_{lm}(u_{t\nu}, v_{t\nu}) = \Pi_{lm}(u_{t\nu}, v_{t\nu})$, the bove Equation is approximated in term of the phase changing in time $\Delta\Psi$ and frequency $\Delta\Phi$ for the case of smearing as:

$$V_\nu(\mathbf{b})^{\Pi}_{t_c\nu_c} \simeq sinc\frac{\Delta\Psi}{2} sinc\frac{\Delta\Phi}{2} \check{\delta}_{lm}(u_{t\nu} + u_{t_c\nu_c}, v_{t\nu} + v_{t_c\nu_c}) \circ \delta(u_{t_c\nu_c}, v_{t_c\nu_c}), \tag{34}$$

with,

$$\Delta\Psi = 2\pi\Big[(u_{t_s\nu_c} - u_{t_e\nu_c})l + (v_{t_s\nu_c} - v_{t_e\nu_c})m + (w_{t_s\nu_c} - w_{t_e\nu_c})(n-1)\Big]$$

$$\Delta\Phi = 2\pi\Big[(u_{t_c\nu_s} - u_{t_c\nu_e})l + (v_{t_c\nu_s} - v_{t_c\nu_e})m + (w_{t_c\nu_s} - w_{t_c\nu_e})(n-1)\Big]$$

We then generalized the approximation of smearing for a random taper $f_{lm}(u_{t\nu}, v_{t\nu})$ as follows:

$$V_\nu(\mathbf{b})^{window}_{t_c\nu_c} \simeq \left[sinc\frac{\Delta\Psi}{2} sinc\frac{\Delta\Phi}{2} \check{\delta}_{lm}(u_{t\nu} + u_{t_c\nu_c}, v_{t\nu} + v_{t_c\nu_c}) \right] \circ \check{w}(u_{t_c\nu_c}, v_{t_c\nu_c}) \circ \check{f}_{lm}(u_{t_c\nu_c}, v_{t_c\nu_c}) \circ \delta(u_{t_c\nu_c}, v_{t_c\nu_c}). \tag{35}$$

Assuming that all the baselines are pointing at the same phase tracking centre, during conventional interferometric imaging, the baseline visibilities are mapped on a uv-plane, and the result is inverse Fourier transformed:

$$PSF(\mathbf{s}) \simeq \mathcal{F}^{-1}\left\{ \sum_{i=1,j=1}^{n_v \times n_{bl}} V_\nu(\mathbf{b})^{window}_{t_i\nu_j} \right\} \tag{36}$$

$$\simeq C(\mathbf{s}) PSF(\mathbf{s_0}), \tag{37}$$

where $PSF(\mathbf{s}_0)$ is the PSF at the phase tracking center and $C(\mathbf{s})$ is the image plane smearing response for a source located toward the direction $\mathbf{s} \neq \mathbf{s_0}$, defined as:

$$C(\mathbf{s}) = \sum_{i=1,j=1}^{n_v \times n_{bl}} w(u_{t_c\nu_c}, v_{t_c\nu_c}) f_{lm}(u_{t_c\nu_c}, v_{t_c\nu_c}) \mathcal{F}^{-1} \left\{ sinc\frac{\Delta\Psi}{2} sinc\frac{\Delta\Phi}{2} \check{\delta}_{lm}(u_{t\nu} + u_{t_c\nu_c}, v_{t\nu} + v_{t_c\nu_c}). \right\} \tag{38}$$

Here, $n_v \times n_{bl}$ is the array total number of visibilities with n_v and n_{bl} the number of visibilities per baseline after averaging and the number of baselines respectively.

The smeared unnormalized PSF of a source, $PSF(\mathbf{s})$ located toward the direction $\mathbf{s}$ was derived and shown that it is proportional to the PSF, $PSF(\mathbf{s}_0)$ of a source at the phase centre.

References