Astronomy & Astrophysics manuscript no. paper
(DOI: will be inserted by hand later)
November 18, 2014

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:

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

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

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

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:

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

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, $\boldsymbol{S}_\square$ 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:

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

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

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

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

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 ($[\boldsymbol{\mathcal{C}}_{\boldsymbol{S}_c}]_0$) 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 and computation cost

4.2. Quick derivation and computation cost

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

4.2.1. Averaging case

The approach is base on the evaluation of the band limited case of the following infinite signal:

$$\tilde{s}(\kappa) = \int_{-\infty}^{+\infty} exp(-j\kappa y)dy, \tag{25}$$

where κ is a real number and $\tilde{s}(\kappa)$ is the Fourier transform of a unitary function, $s(y) = 1$. We can also see $\tilde{s}(\kappa)$ as a delta function, $\delta(\kappa)$. If we band limited $\tilde{s}(\kappa)$ in a small finite phase range Δx_0 centered on the phase x_c, Eq.?? becomes:

$$\begin{aligned}(B_{\Delta x_0}s)(x) =&\frac{1}{\Delta x_0}\int_{x_c-\frac{\Delta x_0}{2}}^{x_c+\frac{\Delta x_0}{2}} \tilde{s}(\kappa)exp(jx\kappa)d\kappa\\ =&\frac{1}{\Delta x_0}\int_{x_c-\frac{\Delta x_0}{2}}^{x_c+\frac{\Delta x_0}{2}}\left(\int_{-\infty}^{+\infty} exp(-j\kappa y)dy\right)\\ &exp(jx\kappa)d\kappa\\ =&\int_{-\infty}^{+\infty}\left(\frac{1}{\Delta x_0}\int_{x_c-\frac{\Delta x_0}{2}}^{x_c+\frac{\Delta x_0}{2}} exp(j(x-y)\kappa)d\kappa\right)dy\\ =&\int_{-\infty}^{+\infty} K(x-y)dy\\ =&K(x)\circ s(x)\\ =&K(x)\end{aligned}$$

Where $s(x) = \mathcal{F}^{-1}\tilde{s}(\kappa)$ and,

$$\begin{aligned}K(x) =&\frac{1}{\Delta x_0}\int_{x_c-\frac{\Delta x_0}{2}}^{x_c+\frac{\Delta x_0}{2}} exp(jx\kappa)d\kappa\\ =&sinc\big(x\frac{\Delta x_0}{2}\big)exp(jxx_c)\end{aligned}$$

Then the result follows:

$$(B_{\Delta x_0}s)(x) =sinc\big(x\frac{\Delta x_0}{2}\big)exp(jxx_c). \tag{26}$$

The result for a two dimensional case follows:

$$\begin{aligned}(B_{\Delta x_0\Delta e_0}s)(x,e) =&sinc(x\frac{\Delta x_0}{2})sinc(e\frac{\Delta e_0}{2}) &(27)\\ &exp\big(j(xx_c+ee_c)\big) &(28)\end{aligned}$$

Prof:

$$\tilde{s}(\kappa,\gamma) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} exp\big(-j(\kappa y+\gamma h)\big)dydh, \tag{29}$$

where $\tilde{s}(\kappa,\gamma) = \delta(\kappa,\gamma)$. If the complex phase of $\tilde{s}(\kappa,\gamma)$ varies linearly in time within the range Δx_0 centered on x_c and in frequency within the range Δe_0 centered on e_c, then Eq.29 becomes:

$$\begin{aligned}(B_{\Delta x_0\Delta e_0}s)(x,e) =&\frac{1}{\Delta x_0\Delta e_0}\int_{x_c-\frac{\Delta x_0}{2}}^{x_c+\frac{\Delta x_0}{2}}\int_{e_c-\frac{\Delta e_0}{2}}^{e_c+\frac{\Delta e_0}{2}}\tilde{s}(\kappa,\gamma)\\ &exp(jx\kappa+je\gamma)d\kappa d\gamma\\ =&\frac{1}{\Delta x_0\Delta e_0}\int_{-\infty}^{+\infty}\int_{x_c-\frac{\Delta x_0}{2}}^{x_c+\frac{\Delta x_0}{2}} exp\big(j\kappa(x-y)\big)\\ &d\kappa dy\left[\int_{-\infty}^{+\infty}\int_{e_c-\frac{\Delta e_0}{2}}^{e_c+\frac{\Delta e_0}{2}} exp\big(j\gamma(e-h)\big)d\gamma dh\right]\\ =&K_1(e)K_2(x)\\ =&sinc(x\frac{\Delta x_0}{2})sinc(e\frac{\Delta e_0}{2})exp\big(j(xx_c+ee_c)\big)\end{aligned}$$

We can therefore approximate averaging by assuming that if Δt and $\Delta\nu$ are small enough that the amplitude of the visibility, V_{pq} of a baseline (p,q) remains constant while the complex phase varies linearly in time and frequency. In a two dimensional case, the approximation then follows:

$$V_{pq}^{avg}(t_c,\nu_c) \simeq sinc\frac{\Delta\Psi}{2}sinc\frac{\Delta\Phi}{2}V_{pq}(t_c,\nu_c). \tag{30}$$

Here, $\Delta\Psi$ and $\Delta\Phi$ are the phase changes in time and in frequency respectively, defined as:

$$\begin{aligned}\Delta\Phi =&argV_{pq}(t_c,\nu_e) - argV_{pq}(t_c,\nu_s) &(31)\\ \Delta\Psi =&argV_{pq}(t_e,\nu_c) - argV_{pq}(t_s,\nu_c) &(32)\end{aligned}$$

The high resolution data $V_{pq}(t,\nu)$ that is been average is given by:

$$V_{pq}(t,\nu) = exp\Big\{2j\pi\big(u_{pq}l + v_{pq}m + w_{pq}n\big)\Big\}, \tag{33}$$

where, u_{pq}, v_{pq} and w_{pq} are the baseline coordinates that causes V_{pq} to variate in time and frequency due to the Earth rotation.

The results of the instrument averaged visibilities for all baselines then follows:

$$\begin{aligned}V(t,\nu) \simeq& \sum_{pq} V_{pq}^{avg}(t_c,\nu_c) &(34)\\ \simeq&\sum_{pq} sinc\frac{\Delta\Psi}{2}sinc\frac{\Delta\Phi}{2}V_{pq}(t_c,\nu_c) &(35)\end{aligned}$$

For convenience let assume $V_{pq}(t_c,\nu_c)$ the visibility of a source at the phase centre (coordinates (l_0,m_0)) during the time t_c at ν_c. That said:

$$V_{pq}(t_c,\nu_c) = exp\Big\{2j\pi\big(u_{pq}l_0 + v_{pq}m_0 + w_{pq}n_0\big)\Big\}, \tag{36}$$

and $V_{pq}(t_c,\nu_e)$, $V_{pq}(t_c,\nu_s)$, $V_{pq}(t_e,\nu_c)$, $V_{pq}(t_s,\nu_c)$ the measurements data of a source locate at $(l,m) \neq (l_0,m_0)$ during t_c at ν_e, t_c at ν_s, t_e at ν_c and t_s at ν_c respectively.

Eq35 is inverse Fourier transform and from the convolution theorem, we have:

$$\begin{aligned}R(l,m) \simeq& \sum_{pq}\mathcal{F}^{-1}\Big\{sinc\frac{\Delta\Psi}{2}sinc\frac{\Delta\Phi}{2}\Big\} &(37)\\ &\circ R_{pq}(l_0,m_0). &(38)\end{aligned}$$

Here, $R(l,m)$ is the instrument response (PSF) for a source at location (l,m) and $R_{pq}(l_0,m_0)$ the response of the baseline (p,q) for a source at the phase centre.

Assuming that all the baselines are pointing at the same phase centre then the previous becomes:

$$\begin{aligned}R(l,m) \simeq&\Big(\sum_{pq}\mathcal{F}^{-1}\Big\{sinc\frac{\Delta\Psi}{2}sinc\frac{\Delta\Phi}{2}\Big\}\Big)\circ R(l_0,m_0) &(39)\\ \simeq&Tri(l,m)\circ R(l_0,m_0). &(40)\end{aligned}$$

With, $R(l_0,m_0)$ the instrument response for a source at the phase centre and the function, $Tri(l,m)$ is defined as:

$$Tri(l,m) = \begin{cases} a_{lm} \neq 1 & \text{if (l,m) exist}\\ 0 & \text{otherwise}\end{cases}$$

4.2.2. General case

From our previous work, we showed that averaging is similar to convolving the visibilities with a top-hat function, then we furthermore replaced the top-hat function by a function, f_w. With our terminology, if we tape the visibilities with a function f_w, a general case of Eq. 25 is derived as follows:

$$\tilde{s}(\kappa) = \int_{-\infty}^{+\infty} f_w(y) exp(-j\kappa y) dy \tag{41}$$

$$\tag{42}$$

In this case, $\tilde{s}(\kappa)$ is the Fourier transform of $f_w(\kappa)$. Following the derivation of Eq.**??**, a general case of the band-limited form of Eq.**??** is:

$$\begin{aligned}(B_{\Delta x_0} s)(x) &= \Big[sinc\big(x\frac{\Delta x_0}{2}\big) exp(jxx_c) \Big] \circ f_w(x) \\ &= \Big[sinc\big(x\frac{\Delta x_0}{2}\big) \circ f_w(x) \Big] exp(jxx_c)\end{aligned}$$

In a two dimensional case, the previous becomes:

$$\begin{aligned}(B_{\Delta x_0 \Delta e_0} s)(x, e) = &\Big[sinc(x\frac{\Delta x_0}{2}) sinc(e\frac{\Delta e_0}{2}) \circ f_w(x, e) \Big] \\ &exp\big(j(xx_c + ee_c)\big)\end{aligned}$$

Eq.30 can therefore be estimate in a general case as:

$$V_{pq}^{corr}(t_c, \nu_c) \simeq \Big[sinc\frac{\Delta\Psi}{2} sinc\frac{\Delta\Phi}{2} \circ f_w(t_c, \nu_c) \Big] V_{pq}(t_c, \nu_c) \tag{43}$$

The results of the instrument weighted average visibilities for all baselines then follows:

$$R(l, m) \simeq \sum_{pq} \mathcal{F}^{-1} \Big\{ sinc\frac{\Delta\Psi}{2} sinc\frac{\Delta\Phi}{2} \circ f_w(t_c, \nu_c) \Big\} \tag{44}$$

$$\circ R_{pq}(l_0, m_0) \tag{45}$$

$$\simeq \sum_{pq} \mathcal{F}^{-1} \Big\{ sinc\frac{\Delta\Psi}{2} sinc\frac{\Delta\Phi}{2} \Big\} \mathcal{F}^{-1} \Big\{ f_w(t_c, \nu_c) \Big\} \tag{46}$$

$$\circ R_{pq}(l_0, m_0) \tag{47}$$

$$\simeq \Big[Tri(l, m) F_w(l_0, m_0) \Big] \circ R(l_0, m_0), \tag{48}$$

where $\mathcal{F}^{-1}\Big\{ f_w(t_c, \nu_c) \Big\} = F_w(l_0, m_0)$.

5. Simulation and comparison

6. Discussion and conclusion

References