(APPENDIX 4 to "CASA_Design_static_Analysis_Notes1")

Algorithm/Pseudo-Code : CLEAN w/ Visibility-Domain Corrections for DDEs

Data : Calibrated Visibilities : $\vec{V}^{corr}_{n\times 1} = [K^{Vis}_{n\times n}{}^{+}]\vec{V}^{obs}_{n\times 1} = [K^{Vis}_{n\times n}]^{-1} V^{obs}_{n\times 1}$ (corrected)

Data : Gridding convolution function : $K^{dd^{-1}}_{ij} \star \vec{P}_s$

(Convolution Kernel) (prolate spheroidal function)

→ GCF that corrects for DDEs for baseline ij during gridding. Several DDEs can be combined to form a single GCF.

(transfer function)

Data : uv-sampling function : $S_{n\times m}$

Data : Image noise threshold & loop gain σ_{thr}, g_s

Result : Model Image : $\vec{I}^{model}_{m_I \times 1}$ → reconstructed estimate of $\vec{I}^{sky}$

'known' model of the source

sky brightness distribution

① Compute the PSF : $\vec{I}^{psf}$ i.e.,

$\rightarrow \vec{I}^{psf}_{m_I\times 1} = W^{-1}_{sum}[I^{ps-1}][F^{+}RG^{ps}S^{+}W]\vec{1}_{n\times 1}$ (using only prolate spheroidal $\vec{P}_s$ as GCF and no DDEs).

→ (PSF is spatially variant)

② Compute the weight image : $\vec{I}^{wt}$, using $[S^{dd}]$ & $[S^{dd^{+}}]$,

$\vec{I}^{wt} = [F^{+}]\sum_{ij} \vec{K}^{dd*}_{ij} W_{ij} \vec{K}^{dd}_{ij}$, where $\boxed{\vec{K}^{dd} \text{ is u-v domain convolution kernel for DDEs}}$

$\rightarrow I^{wt}_{dd}$

$[S^{dd}]$ = sampling matrix/function that includes a baseline-based convolution of visibility function w/ $\vec{K}^{dd}_{ij}$.

Contains Vector $\vec{K}^{dd}_{ij}$

$= [S][G^{Pb}] = [S][FP_bF^{+}]$, where $[P_b] = diag(\vec{P}_b)$ and

(convolution operator)

$\vec{P}_b = [F^{+}]K^{Pb} = \sqrt{\vec{I}^{wt}/W_{sum}}$

③ Compute PB $\vec{P}_b$ from $\vec{I}^{wt}$.

- When antenna PB is dominant DDE, an average PB is computed via

$\vec{P}_b = \sqrt{\vec{I}^{wt}/W_{sum}}$

Just getting to the start or setting up for major **cycle**

(4) Compute the dirty image: $\vec{I}^{dirty}$ using $[S^{dd^{\dagger}}]$

$$\vec{I}^{dirty} = \vec{I}^{res}_{m_i \times 1} = [F^{\dagger} R S^{dd^{\dagger}} W^{im}] \vec{V}^{res}_{n \times 1}$$

(UV-domain to image domain) Reverse transform step: (FFT) $\vec{V}^{res} = \vec{V}^{corr}$ here

Note: When this step is combined w/ forward transform step defined in step 16, dirty/residual image is equivalent to computing $\nabla \chi^2$ (Newton-Raphson), i.e.,

$$\vec{I}^{D}_{s,k} = -\underbrace{\left[\frac{\partial^2 \chi^2}{\partial I_{s,k}\, \partial I_{s,k}}\right]^{-1}}_{\text{Hessian (pre-computed)}} \underbrace{\frac{\partial \chi^2}{\partial I_{s,k}}}_{\nabla \chi^2} \Bigg|_{\vec{I}_{s,k}=0} \quad \text{where } \vec{I}_{s,k} = \{\text{model image}\}$$

$\vec{I}^{D}_{s,k}$: polarized sky brightness, residual image

(5) Measure the peak PSF sidelobe: $f_{sidelobe} = \max\{\text{sidelobe}(\vec{I}^{psf})\}$

Actual start of major cycle here

(6) Initialize model image & residual image: Model image $\vec{I}^{model}$ is initialized to zero or to an a-priori model, i.e.,
$\vec{I}^{model} = 0$, $\vec{I}^{res} = \vec{I}^{dirty}$

(7) Repeat

(8) Normalize residual image: $\vec{I}^{res} = [\vec{P}_b]^{-1} \vec{I}^{dirty}$

- There are 2 ways of normalizing I^{res} before ~~beginnin~~ beginning minor cycle iterations:
  - (1) Calculate principle solution by dividing RHS by $\vec{I}^{wt}$, &
  - (2) divide RHS by estimate of average $\vec{P}_b$ and sum of weights

Flat noise: Normalization by $\vec{P}_b W_{sum}$ before deconvolution will result in model image described by $\boxed{\vec{I}^{model} = \vec{I}^{sky} \cdot \vec{P}_b}$

$$\vec{I}^{model}_{(i)} = T\left(\vec{I}^{res}, \vec{I}^{PSF}\right)$$ Computed earlier

$\nabla\chi^2$ image (Newton Raphson), (Hessian)

Minor cycle starts here

(9) Compute a flux limit: $f_{limit} = \max\left\{\sigma_{thr}, f_{sidelobe} * \max\{\vec{I}^{res}\}\right\}$

(10) Repeat

(11) Find location & amplitude of peak: $\delta\vec{I}^{model} = \text{peak}(\vec{I}^{res})$

Note: $\delta\vec{I}^{model}$ (= update to existing model image) $= \max\left[C\nabla\chi^2\right]$

$$= \max\left[C\frac{\partial\chi^2}{\partial I^{model}}\right]$$

$$= \max\left\{-\left[\frac{\partial^2\chi^2}{\partial I^{model}\partial I^{model}}\right]^{-1}\frac{\partial\chi^2}{\partial I^{model}}\right\}$$

$$= \max\left(\Delta I_i^{dirty}\right)$$

(12) Update model image: $\vec{I}^{model} = \vec{I}^{model} + g_s\,\delta\vec{I}^{model}$

$\to$ loop gain ($0<g_s<1$)

$\to$ Accumulate flux components from iteration i onto a model image

(13) Update residual image: $\vec{I}^{res} = \vec{I}^{res} - g_s\left[\delta\vec{I}^{model} \star \vec{I}^{PSF}\right]$

i.e., first update RHS:

Residual image, $\vec{I}^{res}_{m_i \times 1} = \left[F^{\dagger} R S^{dd^{\dagger}} W^{im}\right]\vec{V}^{res}_{n\times 1}$, is updated

by subtracting out contribution of flux components found in iteration i, damped by loop gain:

$$\vec{I}^{res} = \vec{I}^{res} - g\left(\vec{I}^{PSF} \star \vec{I}^{model}_i\right)$$

(14) Until $\max\{\vec{I}^{res}\} < f_{limit}$, i.e., until minor cycle flux limit is reached $\to$ until some termination criterion is satisfied (usually when T can no longer reliably extract any flux from I^{res})

END OF MINOR CYCLE: ($\vec{I}^{model}$ is obtained)

MAJOR CYCLE

(15) Correct for PB: Divide model image by PB: $\vec{I}^{model} = \vec{I}^{model}/P_b$

$\to$ Depending on choice of normalization, model image @ end of minor cycle has to be further processed before model visibilities can be predicted from it

- Flat noise : New model image is computed as $\vec{I}^{model} / \vec{P}_b$, i.e., division of model image by PB (known a-priori) will give an image of reconstructed sky brightness distribution.

MAJOR CYCLE (Continued)

(16) Predict model visibilities $\vec{V}^{model}$ from $\vec{I}^{model}$ using $[S^{dd}]$.

Forward Transform (FFT) (image-domain to uv-domain).

GCF is used

- Prediction step of computing model visibilities from current sky model needs to re-introduce all DDEs that are being corrected for during gridding, before model can be compared w/ data for χ^2 computation, (i.e., visibilities that would be measured for current sky model $\vec{I}^{model}$ are computed so that model can be compared w/ data $\vec{I}^{corr}$ & new residual visibilities computed

$$\vec{V}^{model}_{n\times 1} = [\underbrace{S^{dd}}_{\text{includes } G^{dd}} R^{\dagger} F][I^{PS}]^{-1} \vec{I}^{model}_{m_I \times 1}$$

(17) Compute a new residual image $\vec{I}^{res}$ from $\vec{V}^{res} = \vec{V}^{corr} - \vec{V}^{model}$ using $[S^{dd}]$:

$$\vec{I}^{res}_{m_I \times 1} = [F^{\dagger} R \underbrace{S^{dd\dagger}}_{\text{includes } G^{dd}} W^{im}] \vec{V}^{res}_{n\times 1}$$

(18) until $\max\{\vec{I}^{res}\} < \sigma_{thr}$, i.e., final convergence criterion is satisfied (usually, when $\vec{V}^{res}$ and $\vec{I}^{res}$ are noise-like)

(19) Restore the final model image:

Final model image $\vec{I}^{model}$ is restored by smoothing it w/ restoring beam, & adding back residuals computed via step (17) above, and then normalizing by $\vec{I}^{wt}$ (principal solution):

Convolution ~~Conv~~ kernel

restoring beam (Gaussian whose width is chosen as width of central lobe of PSF).

$$\vec{I}^{wt} = [F^{\dagger}] \sum_{ij} \vec{K}_{ij}^{dd*} \, W_{ij} \, \vec{K}_{ij}^{dd}$$

$\Rightarrow$ Result : $\vec{I}^{restored} = \vec{I}^{model}_{(final)} \star \vec{I}^{beam} + \vec{I}^{res}_{M_I \times 1}$

and, normalize by/with $\vec{I}^{wt}$

Additional Notes

$[S^{dd}]$: sampling matrix that includes a baseline-based convolution of visibility function with $\vec{K}_{ij}^{dd}$.

$$[S^{dd}] = [S][G^{Pb}] = [S][FP_b F^{\dagger}] \equiv [S_{n\times m}][G^{dd}_{m\times m}]$$

where $[G^{Pb}] = [FP_b F^{\dagger}]$ = gridding convolution operator,

$\vec{P}_b = [F^{\dagger}]\vec{K}^{Pb} = \sqrt{\vec{I}^{wt}/W_{sum}}$,

$\vec{K}_{ij}^{Pb} = \vec{J}_i \star \vec{J}_j$ = convolution of 2 AIFs for 1 polarization pair

$[G_{ij}^{dd}]_{1\times m} = diag([FK_{ij}^{sky}]_{m\times m})$

- Note that in the Forward Fourier Transform steps (FFT),

$$\vec{V}^{model}_{n\times 1} = [S^{dd}R^{\dagger}F][I^{PS}]^{-1}\vec{I}^{model}_{M_I \times 1},$$

$[I^{PS}]^{-1}$ is the grid correction step, and

$[R^{\dagger}]$ maps the model visibility function from coarse grid to fine grid, before interpolating across a fine grid via a convolution, to evaluate model visibilities at sampled spatial frequencies.

Additional Notes /List of symbols

(http://www.aoc.nrao.edu/~rurvashi/DataFiles/UrvashiRV_PhdThesis.pdf)

$[F_{m\times m}]$	matrix operator : Discrete Fourier Transform .

$[F]$ represents the forward transform (image-domain to uv-domain). $[F^\dagger]$ represents the reverse transform (uv-domain to image-domain). $[F^\dagger F] = m1_m$ where 1_m is the $m \times m$ Identity matrix. $[F^\dagger]$ gives an un-normalized Fourier inverse. $[F^{-1}] = \frac{1}{m}[F^\dagger]$ gives a normalized Fourier inverse.

$[G_{m\times m}]$	gridding convolution operator $[G] = [FXF^\dagger]$ .... (uv-domain convolution operator with kernel given by $[F^\dagger]\vec{X}$) (image-domain element-by-element multiplication with $\vec{X}$)
$[G^{dd}]$	gridding convolution with K^{dd} ....
$[G^{pb}]$	gridding convolution with K^{pb} ....
$[G^{ps}]$	gridding convolution with the prolate spheroidal ....
$\vec{I}^{dirty}$	dirty image
$\vec{I}^{model}$	model image, reconstructed estimate of $\vec{I}^{sky}$
$\vec{I}^{psf}$	point spread function
$\vec{I}^{sky}$	sky brightness distribution
$\vec{I}^{wt}$	weight image
$[J_{2\times 2}]$	Jones matrix ....
$\vec{J}$	uv-plane aperture illumination function ....
$[J_{m\times m}]$	uv-plane aperture illumination function in matrix form $[J] = diag(\vec{J})$
$[K_{4\times 4}]$	outer product of two Jones matrices ($[J_{2\times 2}] \otimes [J_{2\times 2}]$) ....
$[K_{ij}(u,v)]$	$[K_{4\times 4}]$ for baseline ij and 2-D spatial frequency u, v ....
$\vec{K}$	$\vec{K} = \vec{J} \star \vec{J}$ is a uv-plane convolution kernel ....
$[K_{m\times m}]$	uv-plane convolution kernel $[K] = diag(\vec{K})$ ....
$\vec{K}^{dd}$	uv-domain convolution kernel for direction-dependent effects .....
$\vec{K}^{mos}$	convolution function for mosaicing ....
$\vec{K}^{pb}$	convolution of two aperture illumination functions ....
$\vec{P}_s$	prolate spheroidal function ....
$\vec{P}_b$	$\vec{P}_b = [V_p^\dagger][V_p] = [F^\dagger]\vec{K}^{pb}$ is the antenna primary beam
$[P_b]$	antenna primary beam in matrix form $[P_b] = diag(\vec{P}_b)$
$[R_{m_l\times m}]$	projection matrix, resampling operator m to m_l pixels ....
$[S]$	matrix operator : uv-coverage, sampling function, transfer function
$[S^{dd}]$	sampling function with baseline-based convolution ....

$\vec{V}^{corr}$	corrected/calibrated visibilities
$\vec{V}^{model}$	model visibilities . . . .
$\vec{V}^{obs}$	observed visibilities . . . .
$\vec{V}^{res}$	residual visibilities . . . .
$\vec{V}_p$	$\vec{V}_p = [F^{+}]\vec{J}$ is the antenna voltage pattern . . . .
$[V_p]$	antenna coltage pattern in matrix form $[V_p] = diag(\vec{V}_p)$
$[W]$	diagonal matrix : measurement or visibility weights . . .
$[W^G]$	gridded weights $[W^G] = [S^{\dagger}WS]$ . . . .
$[W^{im}]$	imaging weights