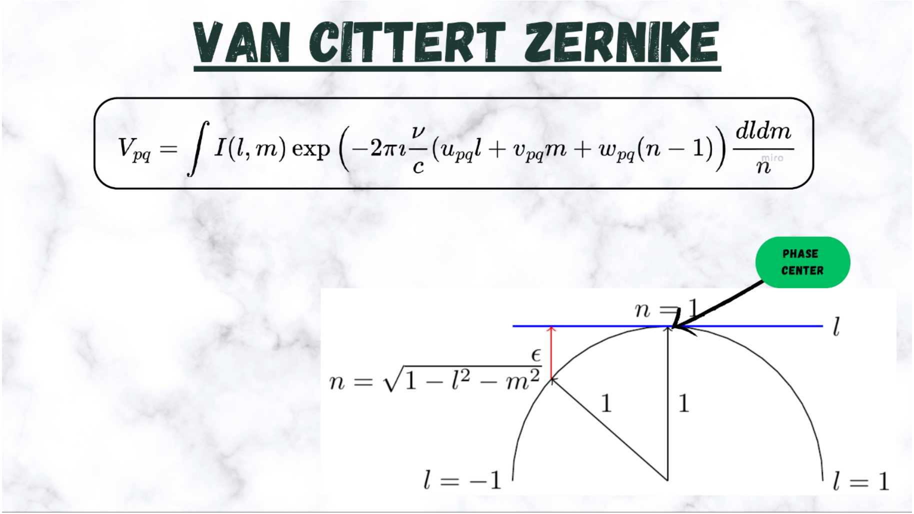
Another way to think about imaging

Tim Molteno

Elec Research
Department of Physics
University of Otago
Dunedin, New Zealand.
tim@elec.ac.nz

Is this the beginning?



Is this the end?

IMAGING IN A NUTSHELL

DISCRETIZED MEASUREMENT MODEL FOR IMAGING

$$y = Rx + \epsilon, \;\; \epsilon \sim \mathcal{N}(0, \Sigma)$$

DISTRIBUTION OF THE NOISE GIVES

$$\chi^2(x) \propto (y - Rx)^\dagger \Sigma^{-1} (y - Rx) \;\rightarrow\; \nabla_x \chi^2 = R^\dagger \Sigma^{-1} (y - Rx)$$

SETTING THE GRADIENT TO ZERO

$$\nabla_x \chi^2 = 0 \;\rightarrow\; R^\dagger \Sigma^{-1} y = R^\dagger \Sigma^{-1} R x \;\rightarrow\; I^D \approx I^{PSF} \star x$$

The challenge:

What is the simplest possible derivation?

Starting at the beginning

Yes, the radios.

No hidden assumptions

Make all assumptions clear. Make no assumptions that we don't need.

Make it fun

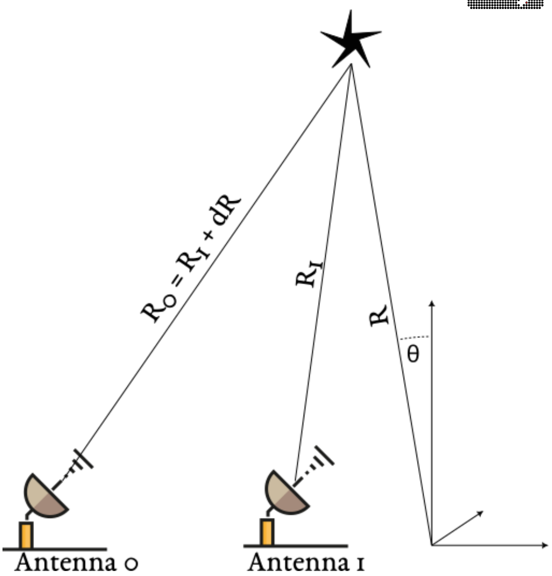
You're kidding right. It's Friday morning...

The beginning: Point Sources

E-field (θ_0, ϕ_0, R)

Assumptions:

- phase offset, ψ_0
- amplitude, $E_o(\omega, t) \sim$ constant
- R_i: from the antenna to source.



Point Source

For a point source at R, θ_0, ϕ_0, and an antenna located at (x, y, z), the distance to the point source from the antenna is:

$$\sqrt{R^2 \sin^2(\phi_0)\sin^2(\theta_0) + R^2\sin^2(\theta_0)\cos^2(\phi_0) + R^2\cos^2(\theta_0) - 2Rx\sin(\phi_0)\sin(\theta_0) - 2Ry\sin(\theta_0)\cos(\phi_0) -$$

The path length difference, δ_R, between the distance to the source, R, and the antenna.

$$\delta_R = R - \sqrt{R^2 - 2Rx\sin(\phi_0)\sin(\theta_0) - 2Ry\sin(\theta_0)\cos(\phi_0) - 2Rz\cos(\theta_0) + x^2 + y^2 + z^2$$

and take the limit as $R \to \infty$

$$\lim_{R\to\infty} \delta_R = x\sin(\phi_0)\sin(\theta_0) + y\sin(\theta_0)\cos(\phi_0) + z\cos(\theta_0)$$

cosines

If we define $l = \sin(\phi)\sin(\theta)$,$m = \cos(\phi)\sin(\theta)$ and $n = \cos(\theta)$, then

$$\delta_R = xl + ym + zn$$

Receivers

Antenna i has position (x_i, y_i, z_i), and operate at a single frequency ω.

Assumption

We are ignoring bandwidth and polarization here.

For a source E_0, R, θ_0, ϕ_0, the antenna will see a field,

$$
\begin{aligned}
E(x, y, z) &= E_0(\omega, t - \frac{\delta_R}{c}) e^{i(\omega t + k\delta_R)} \\
&= E_0(\omega, t) e^{i\left(\omega t + \frac{2\pi(x \sin(\phi_0) \sin(\theta_0) + y \sin(\theta_0) \cos(\phi_0) + z \cos(\theta_0))}{\lambda}\right)}
\end{aligned}
$$

Mixing

Receivers recover $E_0(\omega, t)$ using magic. Multiplication by $e^{-i\omega t}$.

$$
E_0(\omega, t) = E_0(\omega, t) e^{\frac{2i\pi(x \sin(\phi_0) \sin(\theta_0) + y \sin(\theta_0) \cos(\phi_0) + z \cos(\theta_0))}{\lambda}}
$$

Receivers...

Assumption

We have chosen the phase at the origin to be zero.

Assumption E_0 is almost constant

$$E_0(\omega, t - \frac{\delta_R}{c}) \sim E_0(\omega, t)$$

Radio Astronomy

Work out the properties, E_0, θ_0, ϕ_0, of the point source from the received field.

$$E_0(\omega, t) e^{\frac{2i\pi(x \sin(\phi_0) \sin(\theta_0) + y \sin(\theta_0) \cos(\phi_0) + z \cos(\theta_0))}{\lambda}}$$

Radio Astronomy

Work out the properties, E_0, θ_0, ϕ_0, of the point source from the received field.

$$E_0(\omega, t)e^{\frac{2i\pi(x\sin(\phi_0)\sin(\theta_0)+y\sin(\theta_0)\cos(\phi_0)+z\cos(\theta_0))}{\lambda}}$$

Two source sky:

$$\begin{aligned} E &= E_0(\omega, t)e^{i\left(\frac{2\pi(x\sin(\phi_0)\sin(\theta_0)+y\sin(\theta_0)\cos(\phi_0)+z\cos(\theta_0))}{\lambda}\right)} + \\ & E_1(\omega, t)e^{i\left(\frac{2\pi(x\sin(\phi_1)\sin(\theta_1)+y\sin(\theta_1)\cos(\phi_1)+z\cos(\theta_1))}{\lambda}\right)} \end{aligned}$$

Can this be done?

We don't know E. It has an average of 0. It is a complex quantity.

Visibility

To help with finding the sky from the measurements, we need something that doesn't average to zero.
The visibility is the product of the signals from two antennas. The antennas are located at (x_1, y_1, z_1), and (x_2, y_2, z_2).

$$V_{1,2} = \int_t E_1(t)\bar{E_2}(t)dt$$

Where $\bar{E}(t)$ is the complex conjugate.

Why the conjugate?

Single source sky

$$V_{1,2} = E_0(\omega, t)e^{\frac{2i\pi(x_1 \sin(\phi_0)\sin(\theta_0) - x_2 \sin(\phi_0)\sin(\theta_0) + y_1 \sin(\theta_0)\cos(\phi_0) - y_2 \sin(\theta_0)\cos(\phi_0) + z_1 \cos(\theta_0) - z_2 \mathrm{co}}{\lambda}}$$

Simplify with

$$u = \frac{(x_1 - x_2)}{\lambda}$$

$$v = \frac{(y_1 - y_2)}{\lambda}$$

$$w = \frac{(z_1 - z_2)}{\lambda}$$

$$V_{1,2} = E_0(\omega, t)e^{2i\pi(u \sin(\phi_0)\sin(\theta_0) + v \sin(\theta_0)\cos(\phi_0) + w \cos(\theta_0))}\overline{E_0(\omega, t)}$$

Two source sky

Receiver fields add...

$$
\begin{aligned}
E &= E_0(\omega,t)e^{i\left(\frac{2\pi(x\sin(\phi_0)\sin(\theta_0)+y\sin(\theta_0)\cos(\phi_0)+z\cos(\theta_0))}{\lambda}\right)} + \\
& \quad E_1(\omega,t)e^{i\left(\frac{2\pi(x\sin(\phi_1)\sin(\theta_1)+y\sin(\theta_1)\cos(\phi_1)+z\cos(\theta_1))}{\lambda}\right)}
\end{aligned}
$$

but the visibilities have cross terms

$$
\begin{aligned}
V &= \int (E_0 + E_1)\overline{(E_0 + E_1)} \\
&= \int \left(E_0\bar{E}_0 + E_1\bar{E}_1 + E_0\bar{E}_1 + E_1\bar{E}_0\right) \\
&= V_0 + V_1 + \int E_0\bar{E}_1 + \int E_1\bar{E}_0
\end{aligned}
$$

It gets worse with more sources

Two source vis

We have to assume that the sources are uncorrelated

$$\int E_i \overline{E_j} = 0$$

So the visibilities from a two source sky are the sum of the visibilities of each source.

Linearity

$$\begin{aligned} V(u, v, w) &= \sum_i (E_i \overline{E_j}) e^{2i\pi(u \sin(\phi_0) \sin(\theta_0) + v \sin(\theta_0) \cos(\phi_0) + w \cos(\theta_0))} \\ &= \int_{sky} I(\theta, \phi) e^{2i\pi(u \sin(\phi_0) \sin(\theta_0) + v \sin(\theta_0) \cos(\phi_0) + w \cos(\theta_0))} \end{aligned}$$

GRIME Gridded RIME

Ignoring most things (antenna properties, polarization, co-ordinates e.t.c.):

$$V_{ik} = \iint S(\theta, \phi) e^{-2\pi j (u_{ik} l + v_{ik} m + w_{ik} n} d\theta d\phi \equiv V(u_{ik}, v_{ik}, w_{ik}) \tag{1}$$

$$\begin{aligned} l &= \sin(\phi)\cos(\theta) \\ m &= \cos(\phi)\cos(\theta) \\ n &= \cos(\theta) \end{aligned} \qquad \tilde{\mathbf{u}} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

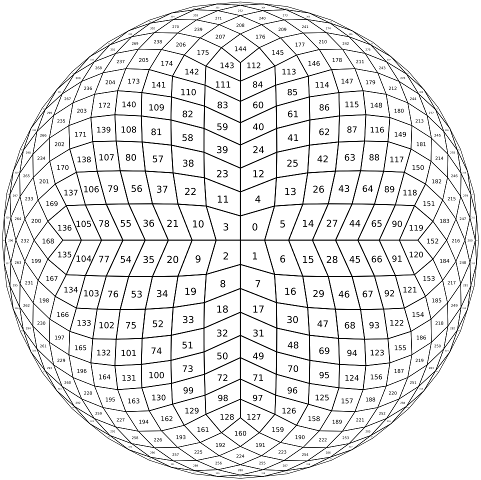
This is an *inner product* in a Hilbert space

$$V(u, v, w) = \langle S(\theta, \phi) |\, H(u, v, w, \theta, \phi) \rangle$$

Discretizing the sky

Discretize $S(\theta, \phi)$ as a piecewise-constant function directly in angular coordinates.

- The value of s_k, the kth element of the $\tilde{\mathbf{s}}$, is the brightness of the *pixel* with $(\theta, \phi) = (\theta_k, \phi_k)$ and area A_k.
- The sky $S(\theta, \phi)$ is represented as a vector, $\tilde{\mathbf{s}}$ in an N_s dimensional vector space $\mathcal{S}$.
- Can pixelize the entire sphere, or bits in any shape.



FFT! Aaaahhhh nooooooooo...

The expression for the visibility (Equation 1) in this representation is an inner-product in a finite vector space:

$$V(u,v,w) = \sum_{k}^{N_s} A_k s_k e^{-2\pi j(ul_k + vm_k + w(n_k - 1))} \tag{2}$$

$$= \tilde{\mathbf{s}} \cdot \hat{\tilde{h}}(u,v,w) \tag{3}$$

$\hat{\tilde{h}}(u,v,w)$ is a vector in the sky space, which we refer to as an *fringe* vector. There is one fringe per visibility (baseline).

Discrete Fringe Vectors

$$\hat{\tilde{h}}(u,v,w) = \begin{pmatrix} A_1 e^{2\pi j(ul_1+vm_1+(n_1-1)w)} \\ A_2 e^{2\pi j(ul_2+vm_2+(n_2-1)w)} \\ A_3 e^{2\pi j(ul_3+vm_3+(n_3-1)w)} \\ \vdots \\ A_k e^{2\pi j(ul_k+vm_k+(n_k-1)w)} \end{pmatrix} \tag{4}$$

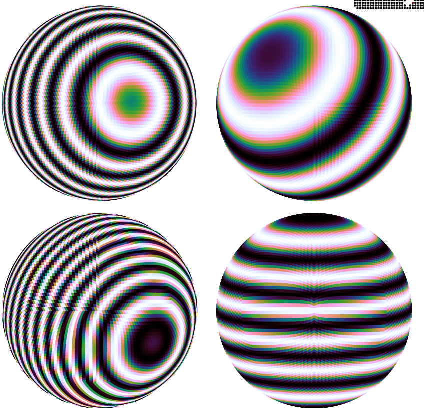
$$= \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_k \end{pmatrix} \odot \exp\left[2\pi j \begin{pmatrix} ul_1 + vm_1 + (n_1 - 1)w \\ ul_2 + vm_2 + (n_2 - 1)w \\ ul_3 + vm_3 + (n_3 - 1)w \\ \vdots \\ ul_k + vm_k + (n_k - 1)w \end{pmatrix} \right]$$

$$= \tilde{\mathbf{a}} \odot e^{2\pi j \mathbf{B}\tilde{\mathbf{u}}} \tag{5}$$

Choosing equal areas. Define:

$$\hat{\tilde{h}}_i = \frac{1}{\sqrt{N_s}} e^{2\pi j \mathbf{B}\tilde{\mathbf{u}}_\mathbf{i}}$$

Given a sky vector $\tilde{\mathbf{s}}$, each baseline produces a visibility $V_i = \tilde{\mathbf{s}} \cdot \hat{\tilde{h}}_i$, the dot-product of the sky vector with the associated fringe.



Note

Fringe vectors are sky vectors.

Properties of the Fringes

Normalized

Fringe vectors unit vectors. $\|\hat{\tilde{h}}_i\| = 1$.

Properties of the Fringes

Normalized

Fringe vectors unit vectors. $\|\hat{\tilde{h}}_i\| = 1$.

Orthogonality

If fringe vectors were orthogonal, we could image by expressing the sky as a sum of fringe vectors.

$$\tilde{\mathbf{s}} = \sum_i^{N_v} V_i \hat{\tilde{h}}_i \qquad S(l,m) = \iint V(u.v) e^{j \cdots} dl dm$$

Properties of the Fringes

Normalized

Fringe vectors unit vectors. $\|\hat{\tilde{h}}_i\| = 1$.

Orthogonality

If fringe vectors were orthogonal, we could image by expressing the sky as a sum of fringe vectors.

$$\tilde{\mathbf{s}} = \sum_i^{N_v} V_i \hat{\tilde{h}}_i \qquad S(l,m) = \iint V(u.v) e^{j\cdots} dl dm$$

Orthogonality

Fringe vectors are *not* orthogonal. $\hat{\tilde{h}}_i \cdot \hat{\tilde{h}}_j \neq 0$.

Telescope forward-map

The ith baseline has a fringe vector $\hat{\tilde{h}}_i$, and given a sky-brightness vector, $\tilde{\mathbf{s}}$, the corresponding visibility V_i is given by

$$V_i = \tilde{\mathbf{s}} \cdot \hat{\tilde{h}}_i \tag{6}$$

This is the *forward-map* for the telescope.

Imaging

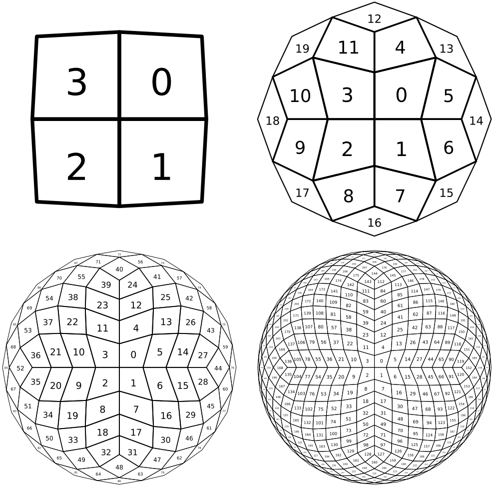
The purpose of an imaging algorithm is to calculate a sky-brightness vector $\tilde{\mathbf{s}}$ from a set of measured visibilities $\{V_i^m\}$. This is an *inverse problem.*

Practical discretizations of the sky: HEALPix

Requirements:

- Equal-area
- Isolateral
- Well understood

HEALPix fits the bill.
Available in C, C++, Fortran90, IDL, Java and Python!

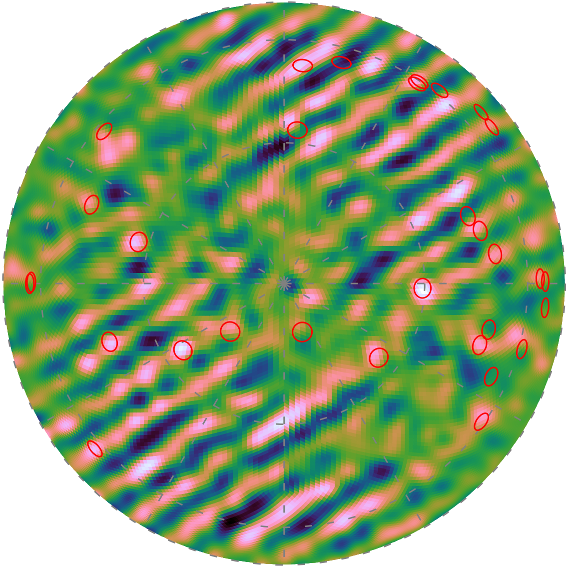


Least-squares sky

Find $\tilde{\mathbf{x}} \in \mathcal{S}$ that minimizes:

$$f(\tilde{\mathbf{x}}) = \sum_{i}^{N_v} \|V_i^m - \tilde{\mathbf{x}} \cdot \hat{\tilde{h}}_i\|^2$$

All-sky TART sky with $N_s = 12288$.

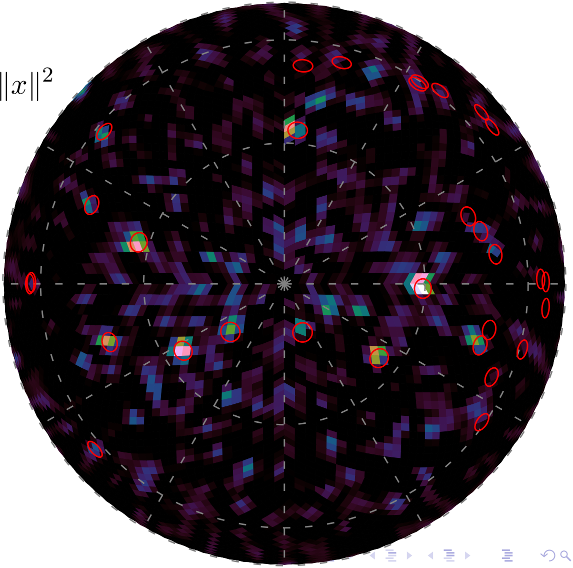


Tikhonov regularization

Find the sky, $\tilde{\mathbf{x}} \in \mathcal{S}$,
that minimizes

$$f(\tilde{\mathbf{x}}) = \sum_{i}^{N_v} \|V_i^m - \tilde{\mathbf{x}} \cdot \hat{\tilde{h}}_i\|^2 + \alpha \|x\|^2$$

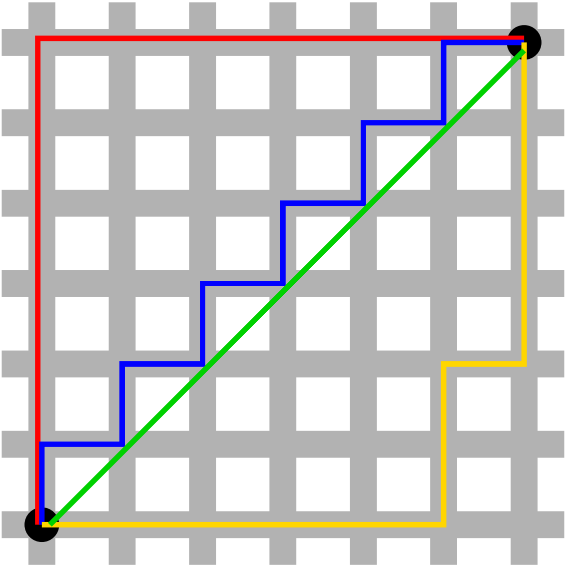
- Prefers a single sky solution (the shortest)
- Enforce positivity and reality.
- Bayesian choice $\alpha = \frac{\sigma_v}{\sigma_0}$.
- Most probable if $\tilde{\mathbf{x}} = \mathcal{N}(\tilde{\mathbf{0}}, \sigma_0 \mathbf{I})$ and $\tilde{\mathbf{v}} = \tilde{\mathbf{v}} + \mathbb{C}\mathcal{N}(\tilde{\mathbf{0}}, \sigma_v \mathbf{I})$.



L1: Taxicab Norm

Taxicab (L_1) norm

- All paths have the same distance (12)
- L_2 distance: $6\sqrt{2} \sim 8.5$.
- Nothing special about L_1 penalty. Just the easiest to work with

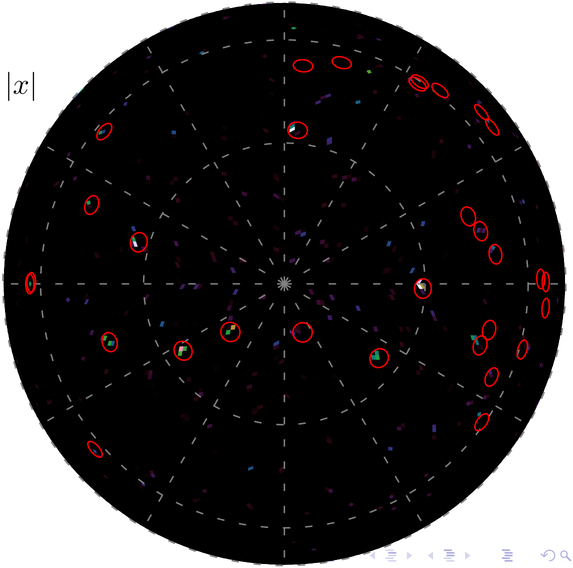


Lasso (L_1) regularization

Find the sky, $\tilde{\mathbf{x}} \in \mathcal{S}$, that minimizes

$$f(\tilde{\mathbf{x}}) = \sum_{i}^{N_v} \left| V_i^m - \tilde{\mathbf{x}} \cdot \hat{\tilde{h}}_i \right| + \alpha \left| x \right|$$

- Prefers sparsity (Fewest pixels)
- c.f. CLEAN
- Can enforce positivity and reality
- Optimum α?
- Cross Validation



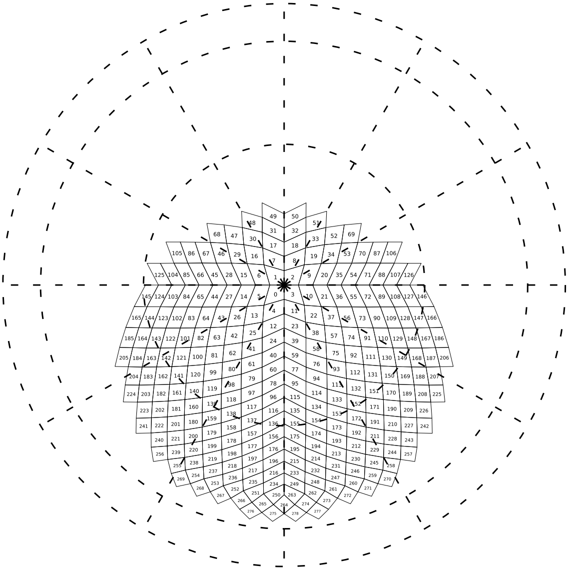
Round Images

The sky is not square

We have a framework for making any shape images we like.

What about beams?

Um, haven't got around to it yet.



Time for a breather.

Conclusion 1

Even a modest resolution discretized sky model is under-determined by snapshot visibility measurements. Even on MeerKET with $N_V \sim 4000$.

Time for a breather.

Conclusion 1

Even a modest resolution discretized sky model is under-determined by snapshot visibility measurements. Even on MeerKET with $N_V \sim 4000$.

Conclusion 2

In many cases, the null-space is far larger than the range space. Images then consist largely of your *prejudices about what should be there*, rather than what your instrument is telling you.

Time for a breather.

Conclusion 1

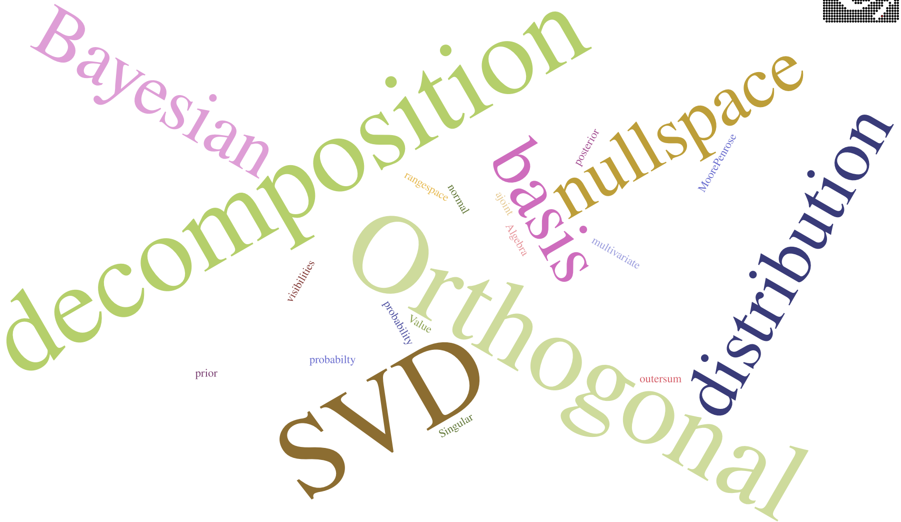
Even a modest resolution discretized sky model is under-determined by snapshot visibility measurements. Even on MeerKET with $N_V \sim 4000$.

Conclusion 2

In many cases, the null-space is far larger than the range space. Images then consist largely of your *prejudices about what should be there*, rather than what your instrument is telling you.

In the next part of this talk...

The range and null-space of a telescope operator are important. Explicitly expressing sky vectors in terms of their components in the range and null-space of $\mathbf{T}$ will allow Bayesian inference to be performed.


Bayesian
decomposition
nullspace
distribution
basis
Orthogonal
SVD
posterior
MoorePenrose
rangespace
normal
ajoint
Algebra
multivariate
visibilities
probability
Value
probabilty
prior
outersum
Singular

Linear Algebra

Given an N_v dimensional vector of visibilities $\mathbf{v}$, with one element for each baseline, the sky-vector, $\tilde{\mathbf{s}}$, satisfies the equation,

$$\mathbf{v} = \mathbf{T}\tilde{\mathbf{s}} \tag{7}$$

Where $\mathbf{T}$ is the *telescope operator*, an $(N_v \times N_s)$ matrix with the fringes as row vectors.

Invertibility

If $N_s > N_v$, $\mathbf{T}$ is not invertible. The sky can not be uniquely determined from visibilities.

From the rank-nullity theorem:

- rank of $\mathbf{T}$ must be smaller than the number of baselines, i.e, $\mathrm{rank}(\mathbf{T}) \leq N_v$
- the dimension of the space $\mathcal{S}$ is N_s
- null-space of $\mathbf{T}$ must be greater than $N_s - N_v$

Typically, $N_s \gtrsim 10^4$, and for TART, $N_v \sim 524$. The rank of the null-space is therefore far greater than the rank of the range space of $\mathbf{T}$.

Singular Value Decomposition

The telescope $\mathbf{T}$ is a map from sky vectors to visibility vectors. The SVD provides an explicit description of the range and null-space of this operator.

$$\mathbf{T} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \begin{pmatrix} \mathbf{U_1} & \mathbf{U_2} \end{pmatrix} \begin{pmatrix} \mathbf{\Sigma_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V_1}^H \\ \mathbf{V_2}^H \end{pmatrix} \tag{8}$$

where:

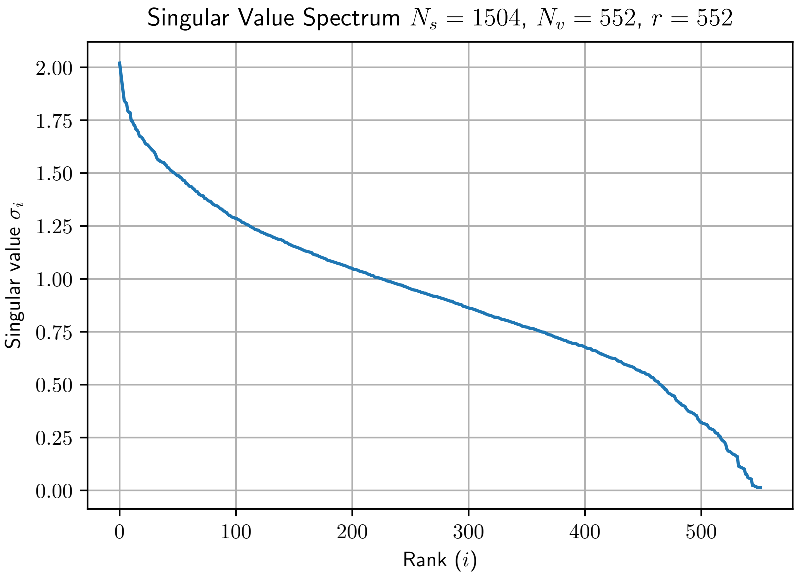
- $\mathbf{\Sigma_1} \in \mathbb{R}^{r \times r}$,
- $\mathbf{U_1} \in \mathbb{C}^{N_v \times r}$, $\mathbf{U_2} \in \mathbb{C}^{N_v \times (N_v - r)}$
- $\mathbf{V_1} \in \mathbb{C}^{N_s \times r}$, $\mathbf{V_2} \in \mathbb{C}^{N_s \times (N_v - r)}$,
- $r = \mathrm{rank}(\mathbf{T})$.

$$\mathbf{T} = \begin{pmatrix} \hat{\tilde{u}}_1 & \dots & \hat{\tilde{u}}_{N_v} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1^* & \dots \\ v_2^* & \dots \\ \vdots & \vdots \\ v_{N_v}^* & \dots \\ v_{N_{v+1}}^* & \dots \\ \dots & \dots \\ v_{N_s}^* & \dots \end{pmatrix}$$

where $\mathbf{U}$ is an $N_v \times N_v$ unitary matrix, $\mathbf{\Sigma}$ is an $N_v \times N_s$ diagonal matrix with non-negative real elements, and $\mathbf{V}$ is an $N_s \times N_s$ matrix.

- Columns of $\mathbf{U} = \begin{pmatrix} \hat{\tilde{u}}_1 & \dots & \hat{\tilde{u}}_{N_v} \end{pmatrix}$ are an orthonormal basis for range($\mathbf{T}$) (visibility space).
- The first N_v column-vectors of $\mathbf{V}$ are an orthonormal basis for range($\mathbf{T}^H$).
- The remaining $N_s - N_v$ column-vectors of $\mathbf{V}$ form an orthonormal basis for the null-space of $\mathbf{T}$.
- Non-degenerate singular values always have unique left- and right-singular vectors, up to a unit-phase factor $e^{j\Phi}$.

Spectrum of TART



Telescope operator sub-spaces

The fundamental theorem of algebra introduces four fundamental sub-spaces for a linear operator. The column space is the sky space $\mathcal{S}$, and can be expressed as the direct sum of two orthogonal spaces:

$$\mathcal{S} = \mathrm{range}(\mathbf{T}^H) \oplus \mathrm{null}(\mathbf{T})$$

This means that any sky vector $x \in \mathcal{S}$ can be expressed as

$$\tilde{\mathbf{x}} = \tilde{\mathbf{x_r}} \oplus \tilde{\mathbf{x_n}}$$

Null-space: null($\mathbf{T}$)

Sky vectors in here produce zero visibilities. They are invisible.

Range-space: range($\mathbf{T}^H$)

Sky vectors in here produce non-zero visibilities. This is the space of visible skies.

Orthogonal Projections

We can project sky vectors into the range and null spaces using the orthogonal projection operators $\mathbf{P_r}$ and $\mathbf{P_n}$ respectively where

$$\mathbf{P_r} = \mathbf{U_1}\mathbf{U_1}^H \tag{9}$$

$$\mathbf{P_n} = \mathbf{V_2}\mathbf{V_2}^H \tag{10}$$

where U_i and V_1 are the components of the SVD of $\mathbf{T}$ from Equation 8.

Question

What is the relationship between PSF, and projecting a single pixel sky onto the range space?

SVD and Natural Basis

$\mathbf{V}^H$ transforms $\tilde{\mathbf{s}} \in \mathcal{S}$ into a new basis factoring can easily be done

$$\tilde{\mathbf{x}} = \mathbf{V}^H \tilde{\mathbf{s}}$$

and as $\mathbf{V}$ is unitary $\tilde{\mathbf{s}} = \mathbf{V}\tilde{\mathbf{x}}$ and therefore

$$\begin{aligned} \mathbf{v} = \mathbf{T}\tilde{\mathbf{s}} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H\tilde{\mathbf{s}} && (11) \\ &= \mathbf{U}\boldsymbol{\Sigma}\tilde{\mathbf{x}} = \mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}(\tilde{\mathbf{x_r}} \oplus \tilde{\mathbf{x_n}}) = \mathbf{A}\tilde{\mathbf{x_r}} && (12) \end{aligned}$$

Where $\mathbf{A} \equiv \mathbf{U}\boldsymbol{\Sigma}$ is the telescope operator in the natural basis for this telescope. The operator $\mathbf{A}$ has the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_r} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{A_r}$ is an $r \times r$ square invertible matrix. In this case

$$\tilde{\mathbf{v}} = \mathbf{A_r}\tilde{\mathbf{x_r}} \tag{13}$$

This is the natural-basis telescope forward map.

Moore-Penrose Inverse

As $\mathbf{T}$ has independent rows, the Moore-Penrose right-inverse, $\mathbf{T}^+$ is given by,

$$\mathbf{T}^+ = \mathbf{V}\Sigma^+\mathbf{U}^*$$

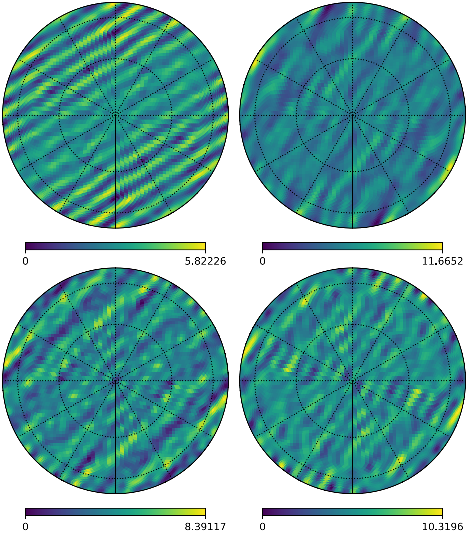
With $\mathbf{T}^+$ in hand, we can invert the telescope forward map as

$$\tilde{\mathbf{s}} = \mathbf{T}^+\mathbf{v}$$

This is clearly closely related to using the inverse of $\mathbf{A_r}$ to retrieve the sky from the visibilities.

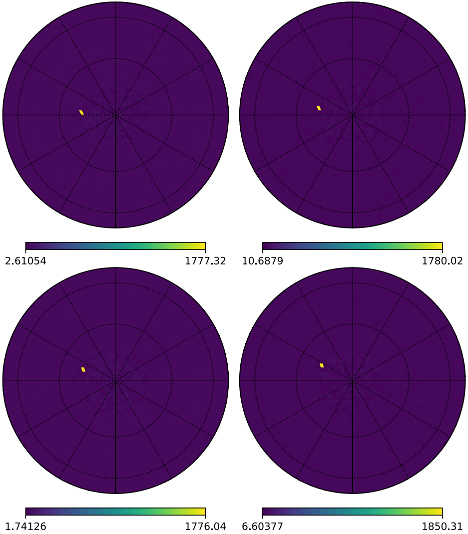
$$\tilde{\mathbf{x_r}} = \mathbf{A_r}^{-1}\tilde{\mathbf{v}}$$

Range-space basis vectors



Range-space basis vectors projected back into the sky space, corresponding to the four largest singular values.

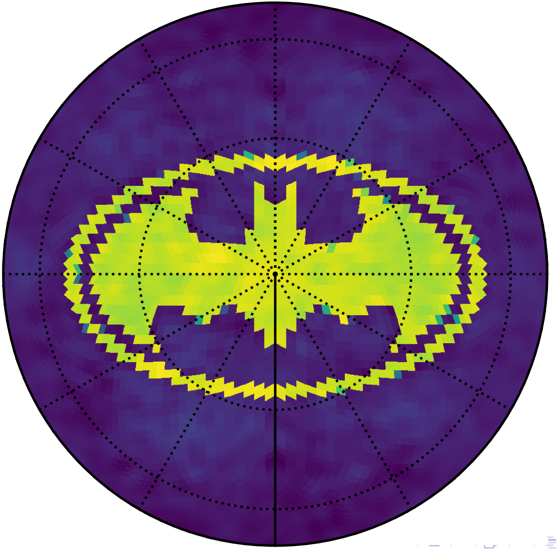
Null-space basis vectors



Some null-space basis vectors. Adding *any* linear combination of these null-space vectors to a sky will not change any telescope measurements.

Batcave projected with $\mathbf{P_n}$ into the null-space

This image can be added to *any* TART image, and the result will be consistent with observations.



Mini-summary

Pre-conceived ideas

The null space is where your prejudices and pre-conceived notions of the sky live.

Science

Science happens in the range space.

Astronomers

You can now give them lots of lovely discoveries, entirely consistent with measurements

Conclusion

Be careful making images. Project your image into the range space using $\mathbf{P_r}$ to check.

Bayesian Sky Inference

Meaure visibilities $\tilde{\mathbf{v}}$. Calculate the posterior sky probability

$$P(\tilde{\mathbf{x}}|\tilde{\mathbf{v}}) = \begin{pmatrix} P(\tilde{\mathbf{x}}_r|\tilde{\mathbf{v}}) \\ P(\tilde{\mathbf{x}}_n|\tilde{\mathbf{v}}) \end{pmatrix} = \begin{pmatrix} P(\tilde{\mathbf{x}}_r|\tilde{\mathbf{v}}) \\ P(\tilde{\mathbf{x}}_n) \end{pmatrix} = \begin{pmatrix} \frac{P(\tilde{\mathbf{v}}|\tilde{\mathbf{x}}_r)}{P(\tilde{\mathbf{v}})} P(\tilde{\mathbf{x}}_r) \\ P(\tilde{\mathbf{x}}_n) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{P(\tilde{\mathbf{v}}|\tilde{\mathbf{x}}_r)}{P(\tilde{\mathbf{v}})} \\ 1 \end{pmatrix} P(\tilde{\mathbf{x}}) \qquad (14)$$

where:

- $P(\tilde{\mathbf{x}})$ is the prior probability of the sky.
- GRIME gives $P(\tilde{\mathbf{v}}|\tilde{\mathbf{x}}_r) = \mathbb{C}\mathcal{N}(\tilde{\mathbf{v}} - A_r\tilde{\mathbf{x}}_r, \sigma_v\mathbf{I})$
- Covariance of v needs care.
- We need a result for Bayes' updates of complex multivariate normals.

Key Idea

In the null-space Prior = Posterior.

Measurement Likelihood: GRIME

Visibility measurements are subject to measurement uncertainty, we model this by additive N_v-variate circularly-symmetric complex normal distribution, i.e.,

$$\tilde{\mathbf{v}} = \mathbf{A_r}\tilde{\mathbf{x_r}} + \mathbb{C}\mathcal{N}(\tilde{\mathbf{0}}, \mathbf{\Sigma_v})$$

The likelihood is then,

$$\begin{aligned} P(\tilde{\mathbf{v}}|\tilde{\mathbf{x}}) &= \mathbb{C}\mathcal{N}\left(\tilde{\mathbf{v}} - \mathbf{A}\tilde{\mathbf{x}}, \mathbf{\Sigma_v}\right) \\ &= \mathbb{C}\mathcal{N}\left(\tilde{\mathbf{v}} - \mathbf{A_r}\tilde{\mathbf{x_r}}, \mathbf{\Sigma_v}\right) \equiv P(\tilde{\mathbf{v}}|\tilde{\mathbf{x_r}}). \end{aligned} \tag{15}$$

This is an assumption

We are assuming that the measurement errors are normal.

Using the likelihood from Equation 15, the posterior sky probability (from Equation 14) is then

$$P(\tilde{\mathbf{x}}|\tilde{\mathbf{v}}) = \begin{pmatrix} P(\tilde{\mathbf{x_r}}|\tilde{\mathbf{v}}) \\ P(\tilde{\mathbf{x_n}}|\tilde{\mathbf{v}}) \end{pmatrix} = \begin{pmatrix} \frac{P(\tilde{\mathbf{v}}|\tilde{\mathbf{x_r}})P(\tilde{\mathbf{x_r}})}{P(\tilde{\mathbf{v}})} \\ P(\tilde{\mathbf{x_n}}) \end{pmatrix} = \begin{pmatrix} \mathbb{C}\mathcal{N}(\mu_{\tilde{\mathbf{x_r}}|\mathbf{v}}, \mathbf{\Sigma}_{\mathbf{x_r}|\mathbf{v}}) \\ \mathbb{C}\mathcal{N}(\tilde{\mathbf{0}}, \sigma_s^2 \mathbf{I_n}) \end{pmatrix} \tag{16}$$

$$= \mathbb{C}\mathcal{N}\left(\tilde{\mu}, \mathbf{\Sigma}\right) \tag{17}$$

The vectors $\mu_{x_r|v}$ and covariance matrix $\mathbf{\Sigma}_{\mathbf{x_r}|\mathbf{v}}$ are found from the application of Bayes' rule to a multivariate normal distribution of likelihood and prior.

$$\tilde{\mu_{\mathbf{x_r}|\mathbf{v}}} = \frac{\mathbf{\Sigma}_{\mathbf{x_r}}}{\mathbf{\Sigma}_{\mathbf{x_r}} + \mathbf{\Sigma}_{\mathbf{v}}} \left(\tilde{\mathbf{v}} - \mathbf{A_r}\tilde{\mathbf{x_r}}\right) \tag{18}$$

$$\mathbf{\Sigma}_{\mathbf{x_r}|\mathbf{v}} = \frac{\mathbf{\Sigma}_{\mathbf{x_r}}\mathbf{\Sigma}_{\mathbf{v}}}{\mathbf{\Sigma}_{\mathbf{x_r}} + \mathbf{\Sigma}_{\mathbf{v}}} \tag{19}$$

where $\mathbf{\Sigma}_{\mathbf{x_r}} = \sigma_s^2\mathbf{I_r}$ is the prior covariance matrix (in the range space), and $\mathbf{\Sigma}_{\mathbf{v}}$ is the likelihood covariance matrix in the range space.
The posterior probability is therefore an N_s-variate complex normal distribution, $\mathbb{C}\mathcal{N}\left(\tilde{\mu}, \mathbf{\Sigma}\right)$, with means,

$$\tilde{\mu} = \begin{pmatrix} \tilde{\mu_{\mathbf{x_r}|\mathbf{v}}} \\ \tilde{\mathbf{0}} \end{pmatrix}$$

and covariance, a block-diagonal matrix,

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{\mathbf{x_r}|\mathbf{v}} & \mathbf{0} \\ \mathbf{0} & \sigma_s^2\mathbf{I}_n \end{pmatrix}.$$

Independent visibilities

If the errors in the measured visibilities are assumed independent, then the covariance matrix of the likelihood, $\mathbf{\Sigma_v}$ is diagonal. We will assume $\mathbf{\Sigma_v} = \sigma_v^2 \mathbf{I_r}$ and $\mathbf{\Sigma_{x_r}} = \sigma_s^2 \mathbf{I_r}$, where $\mathbf{I_r}$ is the identity matrix in the range-space. The mean and covariance of the posterior sky are then,

$$\mu_{\tilde{\mathbf{x}_\mathbf{r}}|\mathbf{v}} = \frac{\sigma_s^2 \mathbf{I_r}}{\sigma_s^2 \mathbf{I_r} + \sigma_v^2 \mathbf{I_r}} \left(\tilde{\mathbf{v}} - \mathbf{A_r} \tilde{\mathbf{x}_\mathbf{r}} \right) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2} \left(\tilde{\mathbf{v}} - \mathbf{A_r} \tilde{\mathbf{x}_\mathbf{r}} \right) \tag{20}$$

$$\mathbf{\Sigma}_{\mathbf{x_r}|\mathbf{v}} = \frac{\sigma_s^2 \mathbf{I_r} \sigma_v^2 \mathbf{I_r}}{\sigma_s^2 \mathbf{I_r} + \sigma_v^2 \mathbf{I_r}} = \frac{\sigma_s^2 \sigma_v^2}{\sigma_s^2 + \sigma_v^2} \mathbf{I_r}. \tag{21}$$

This isn't true

The visibilities are not independent. For example, the noise on v_{ij} and v_{ji} will be correlated.

The algorithm

For a simplified multivariate normal prior sky, the Bayesian inference of the posterior sky given a set of measurements of visibilities requires the following steps:

1. Assume a prior sky distribution with zero-mean and a variance of σ_s^2.
2. Estimate the measurement noise σ_v^2, of the visibility measurements.
3. Choose a pixel resolution on the sky, and based on antenna positions, generate the fringe vectors (Equation 5) and the telescope operator $\mathbf{T}$ (Equation 7).
4. Perform the singular value decomposition $\mathbf{T} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ to obtain the natural-basis telescope operator $\mathbf{A} \equiv \mathbf{U}\mathbf{\Sigma}$.
5. Generate the mean, $\tilde{\mu}$ and covariance, $\mathbf{\Sigma}$ in the natural-basis (Equations 21)
6. Transform back to the pixel basis for visualization to produce the posterior distribution $P(\tilde{\mathbf{s}}|\tilde{\mathbf{v}}) = \mathcal{N}\left(\mathbf{V}^H\tilde{\mu}, \mathbf{V}\mathbf{\Sigma}\mathbf{V}^H\right)$.

Conclusions

- We can make round images.
- Fourier Transforms are not used. I think they might even be a bad idea (c.f. non orthonormality of the basis set on the sphere).
- The operator sub-spaces are kinda nice, as a framework.
- The Bayes stuff seems good. Possibly too good to be true.
- We've done model-fitting in this framework for point-source deconvolution. Now just focussing on L_1 regularization as a CLEAN replacement.

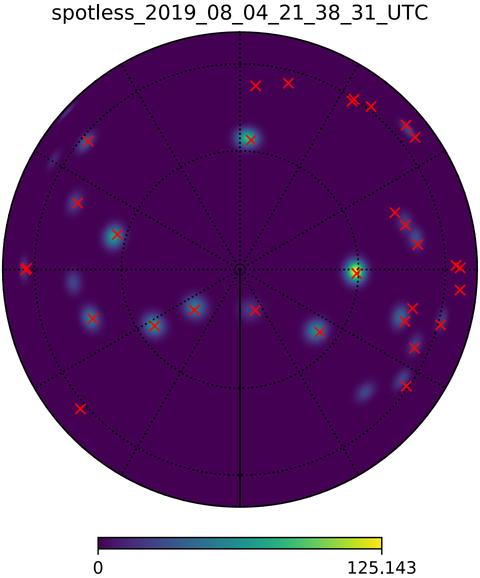
Software GPLv3

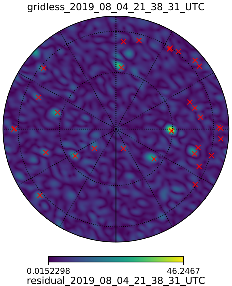
http://github.com/tmolteno/disko

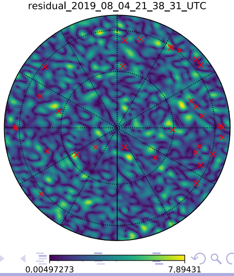
http://github.com/tmolteno/tart2ms

One more thing...

Spotless: Gridless CLEAN in visibility space:







What the future holds

- Can we use FFT/Spherical Harmonic Transforms to make things faster?
- Orthonormal functions on the sphere might replace complex exponentials?
- Million pixels: Out-of-core SVD and regularization.
- Apply to MeerKAT data. Switch to measurement sets.
- Sequential Inference instead of rotation synthesis.
- Using the SVD spectrum to evaluate arrays.
- Representation of posterior sky.
- Deal with frequency
- Tidy up connection to CLEAN
- Antenna response pattern & calibration into GRIME.
- Must actually project TART images into the range space to check.