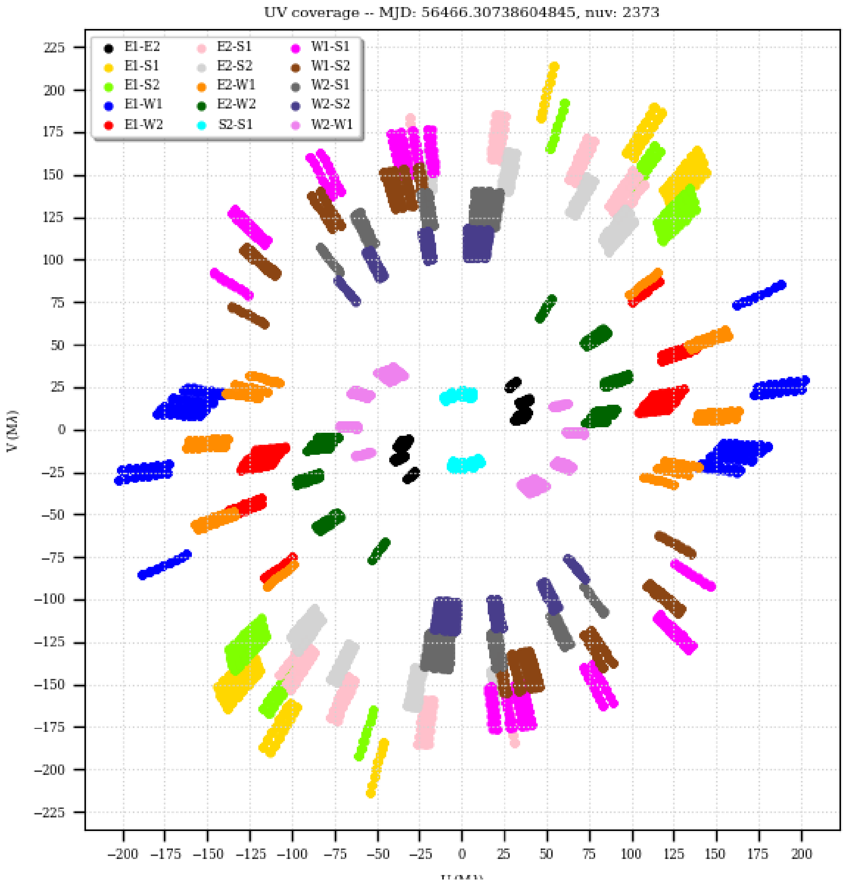
# Principles of Image Reconstruction

*Going from data to images…*

Fabien Baron
Associate Professor of Astronomy
*CHARA & Georgia State University*
*fbaron@gsu.edu → ask me anything!*
*github.com/fabienbaron → find code there to play with*

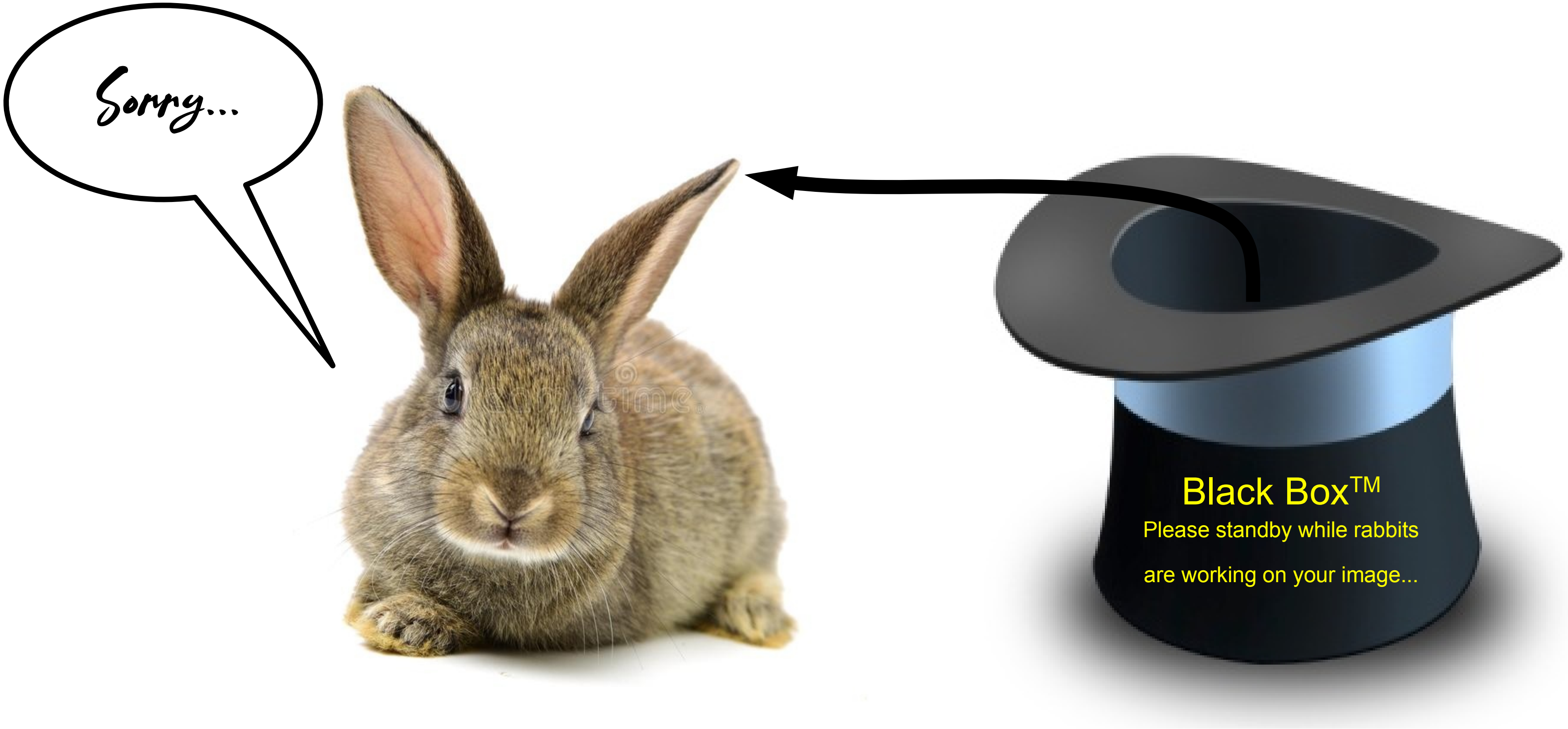
# Interferometric imaging is quite simple



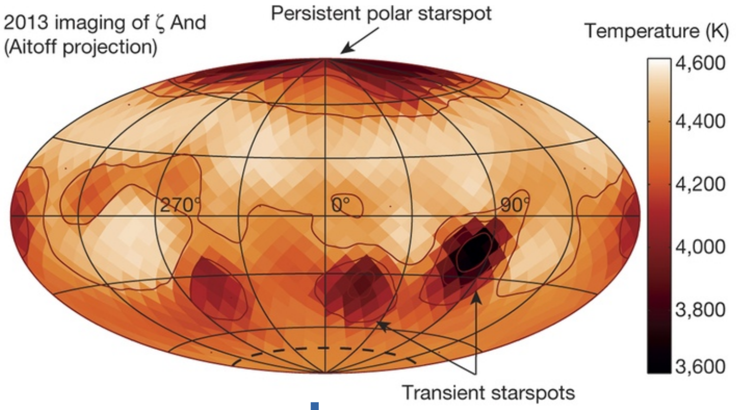

# Just do that


Black Box™
Please standby while rabbits
are working on your image...

# Sometimes it does not work.



# Sometimes it does.




Just add to your paper!


LETTER

doi:10.1038/nature17444

## No Sun–like dynamo on the active star ζ Andromedae from starspot asymmetry

R. M. Roettenbacher[1], J. D. Monnier[1], H. Korhonen[2,3], A. N. Aarnio[1], F. Baron[1,4], X. Che[1], R. O. Harmon[5], Zs. Kővári[6], S. Kraus[1,7], G. H. Schaefer[8], G. Torres[9], M. Zhao[1,10], T. A. ten Brummelaar[8], J. Sturmann[8] & L. Sturmann[8]

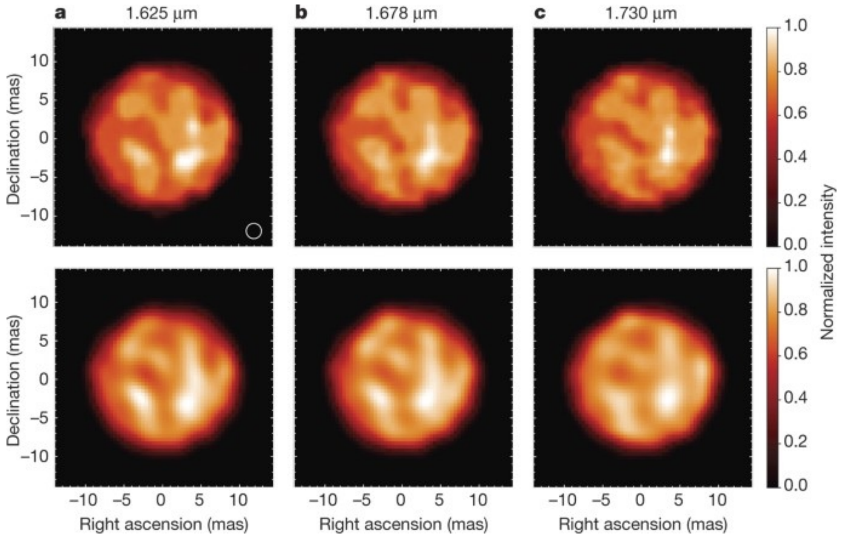
**Sunspots are cool areas caused by strong surface magnetic fields that inhibit convection[1,2]. Moreover, strong magnetic fields can alter the average atmospheric structure[3], degrading our ability to measure stellar masses and ages. Stars that are more active than**

The surface temperature maps for ζ And show peaks of 4,530 K and 4,550 K and minimum values of 3,540 K and 3,660 K in 2011 and 2013, respectively. The ~900-K range of temperatures we see across the surface is slightly larger than the ~700-K range found from recent

# The End?

# What do we want from imaging?

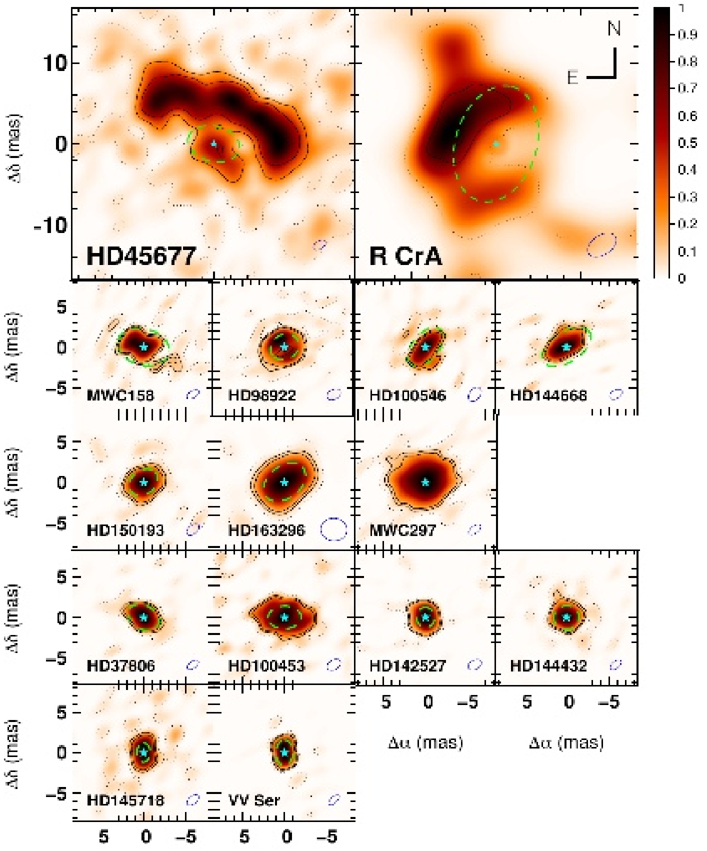
- We want to determine and characterize the most probable image(s) given a dataset **D**, as well as some other background info ***M***
  - Traditionally a single "most probable" image is given
  - Research is ongoing to determine effective algorithms giving the spread of potential images via error maps ("Is this spot real ?")
  - We don't want artefacts




*VLTI PIONIER images of Pi1 Gruis*
*Paladini et al. 2018*

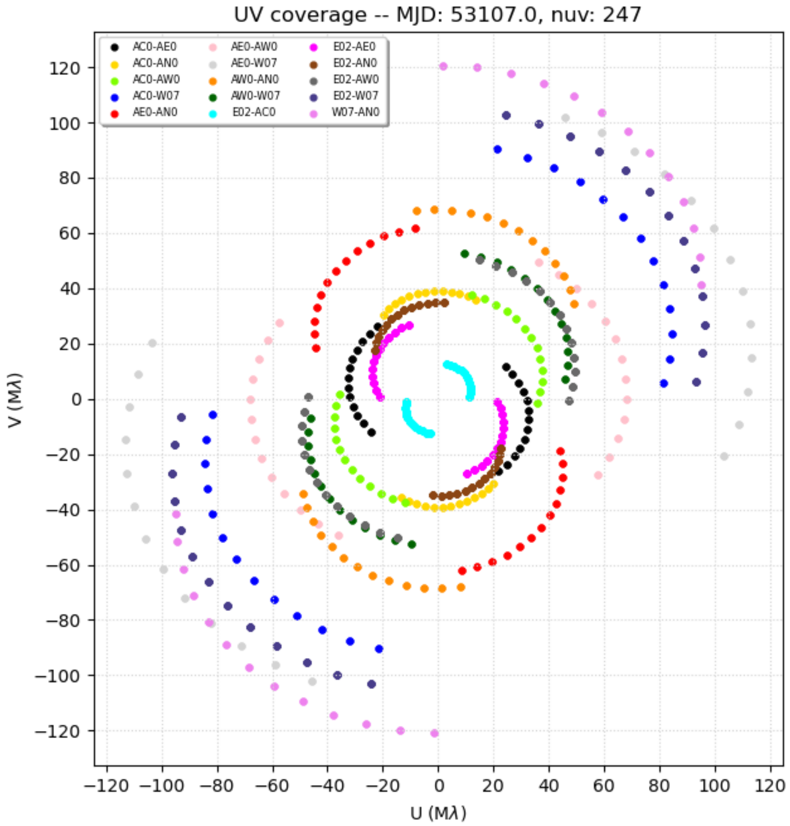
# What is an image?

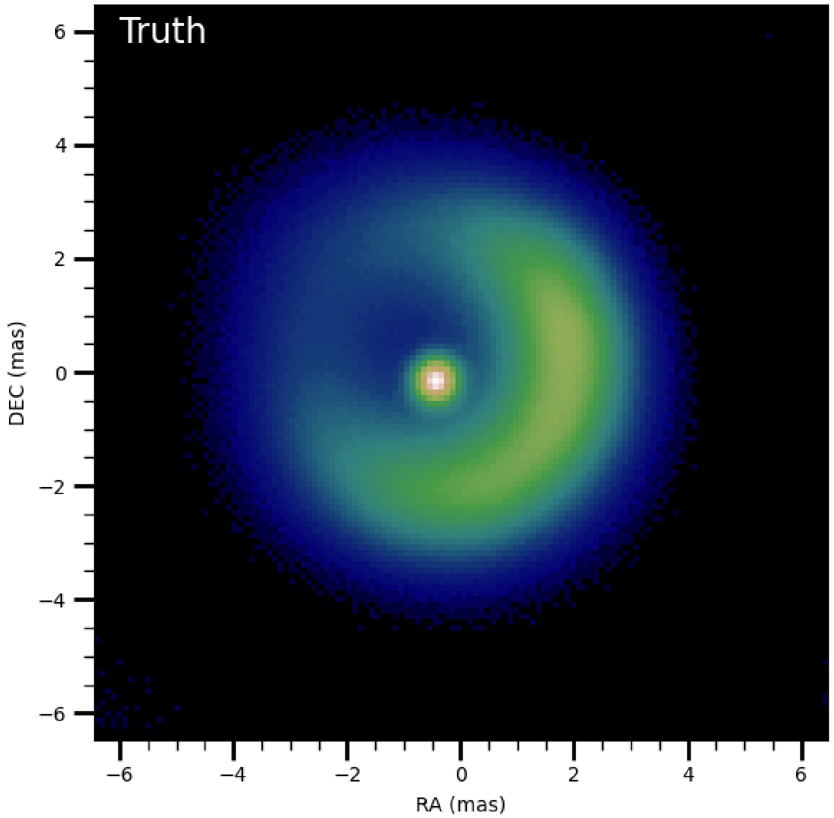
- Images can be:
  - Model-based: a geometrical or physical model is used to fit the data
  - Model-independent: the image is represented by **many identical model parameters**, such pixel intensities or a vector of stellar surface temperatures.
  - Mixed: a model is used to account for certain spatial components (unresolved star in YSO), spectral behavior ($\lambda^{-4}$) or temporal behavior (orbit)




*Herbig AeBe images from VLTI-PIONIER data*
*Kluska et al., 2020*

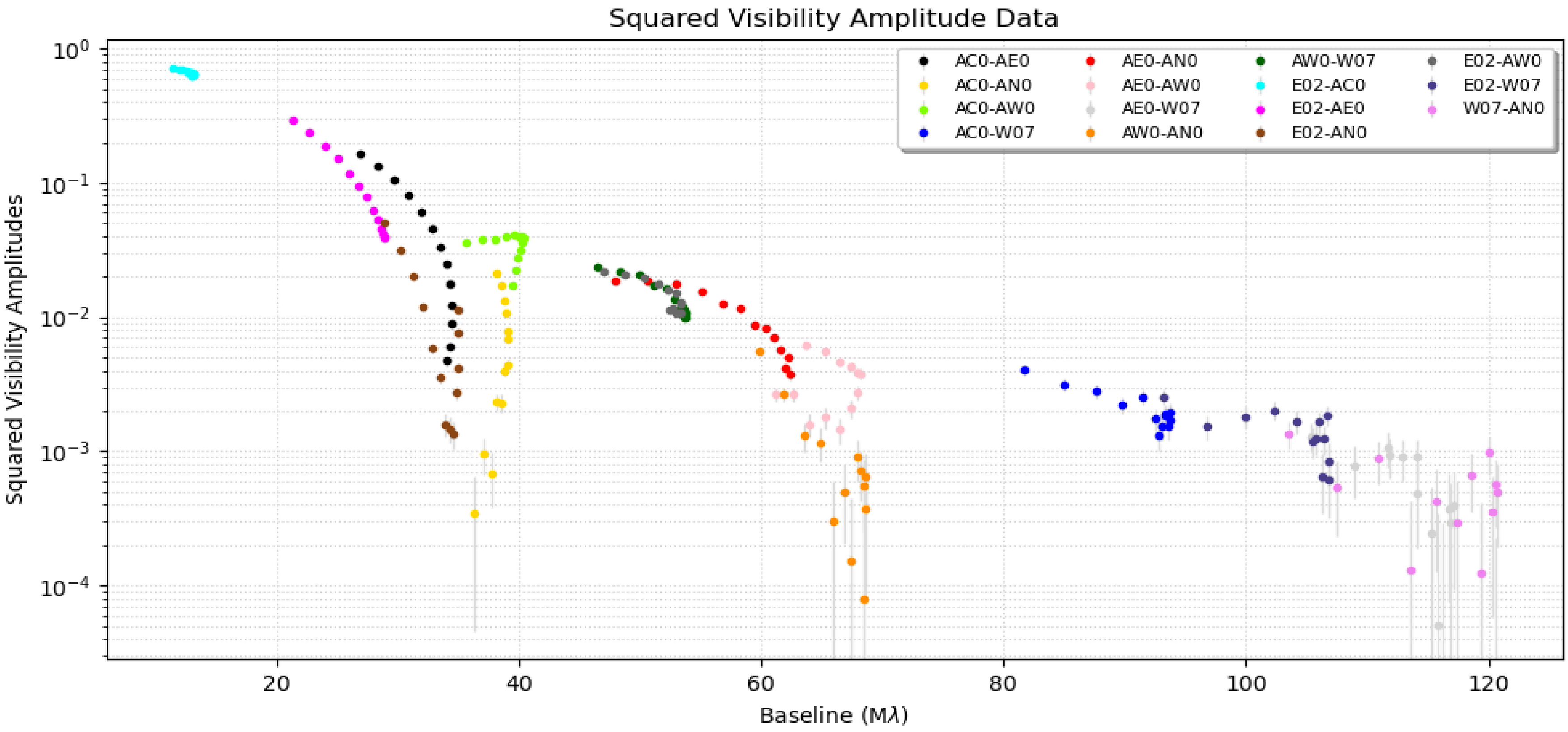
# A famous example: the 2004 Interferometry Beauty Contest data



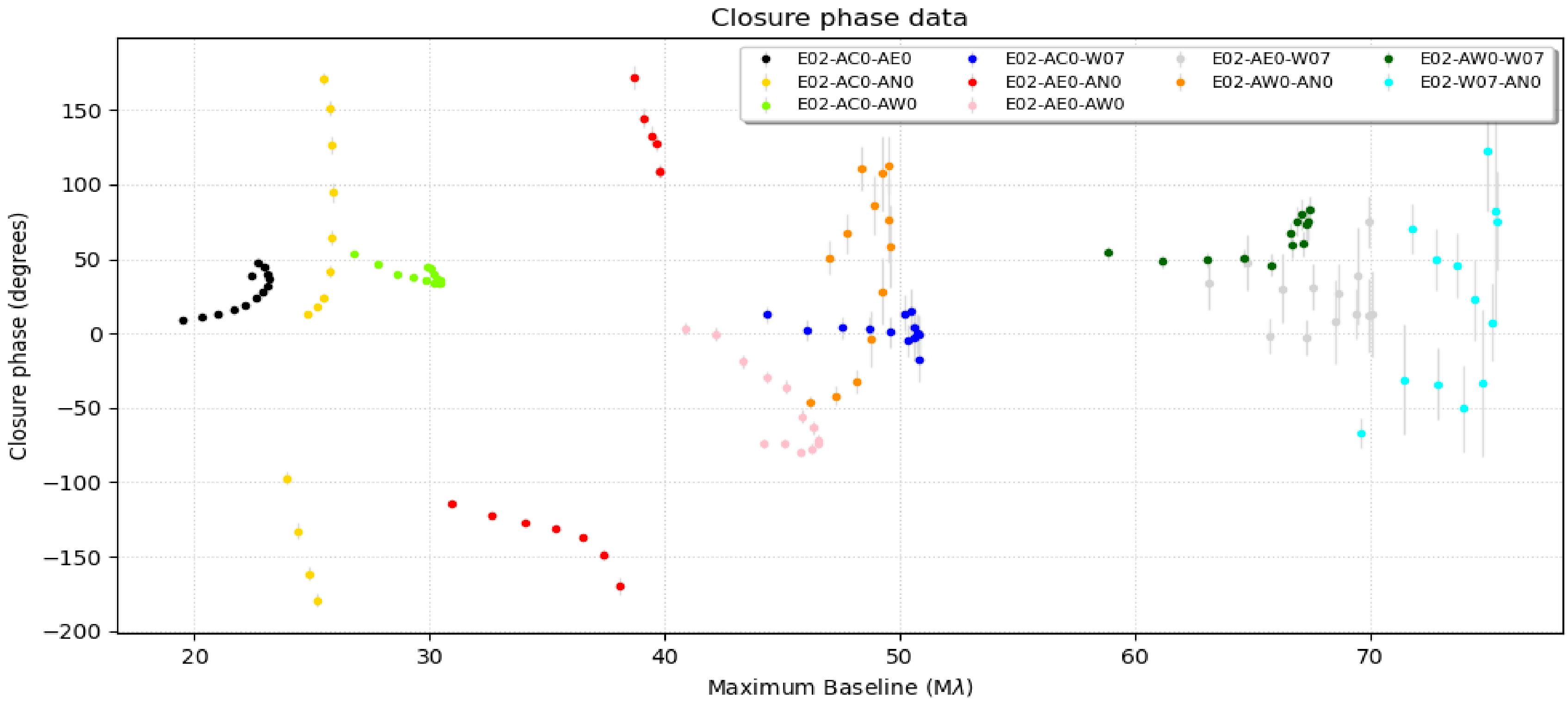

Most of the code used for this talk can be found at github.com/fabienbaron/OITOOLS.jl/demos

# A famous example: the 2004 Interferometry Beauty Contest data




Most of the code used for this talk can be found at github.com/fabienbaron/OITOOLS.jl/demos

# A famous example: the 2004 Interferometry Beauty Contest data




Most of the code used for this talk can be found at github.com/fabienbaron/OITOOLS.jl/demos

# Checklist of ingredients for image reconstruction

- A data set: choice of observables to take into account: powerspectra, closure phases, triple amplitudes, differential phases, fluxes (whatever is in your data)
- Angular size of a pixel (example: 0.5 mas/pixel). Typically $\Delta\theta \sim \frac{\lambda_{\min}}{4B_{\max}}$
- Field of view: image size (typically: 64x64, 128x128 pixels)
- A choice of Fourier Transform (exact DFT, NFFT, including or not bandwidth smearing) to define the $\chi^2$
- Choice of regularization functions and their weights relative to $\chi^2$
- An initial guess (example: point source at the center, random values, model fitted from the data)
- An algorithm to optimize the chi2 starting from the initial guess

# Reminder: the *forward model*: from image to observables

**D**iscrete
**F**ourier
**T**ransform
(slow, exact)

$\boldsymbol{V} = \boldsymbol{H}.\boldsymbol{x}$

or

**N**on-equispaced
**F**ast
**F**ourier
**T**ransform
(fast yet accurate)

$$\frac{\boldsymbol{F}(\boldsymbol{x}, (\boldsymbol{u}, \boldsymbol{v}))}{\sum_i x_i}$$

$\boldsymbol{x}$

image

$\mathbf{V}$

Complex visibilities

$|\boldsymbol{V}|^2$

Powerspectra

$\boldsymbol{T_3} = V_{ij} V_{jk} V_{ki}$

Bispectra

$|\boldsymbol{T_3}|$

Triple Amplitudes

$\arg \boldsymbol{T_3} = \boldsymbol{C\phi}$

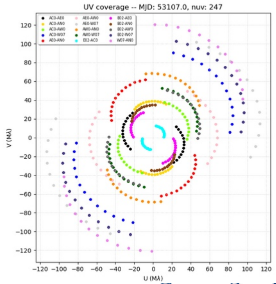
Closure phases (via bispectrum)

$$\phi = \operatorname{atan} \frac{\mathcal{I}\mathrm{m}(\boldsymbol{V})}{\mathcal{R}\mathrm{e}(\boldsymbol{V})}$$

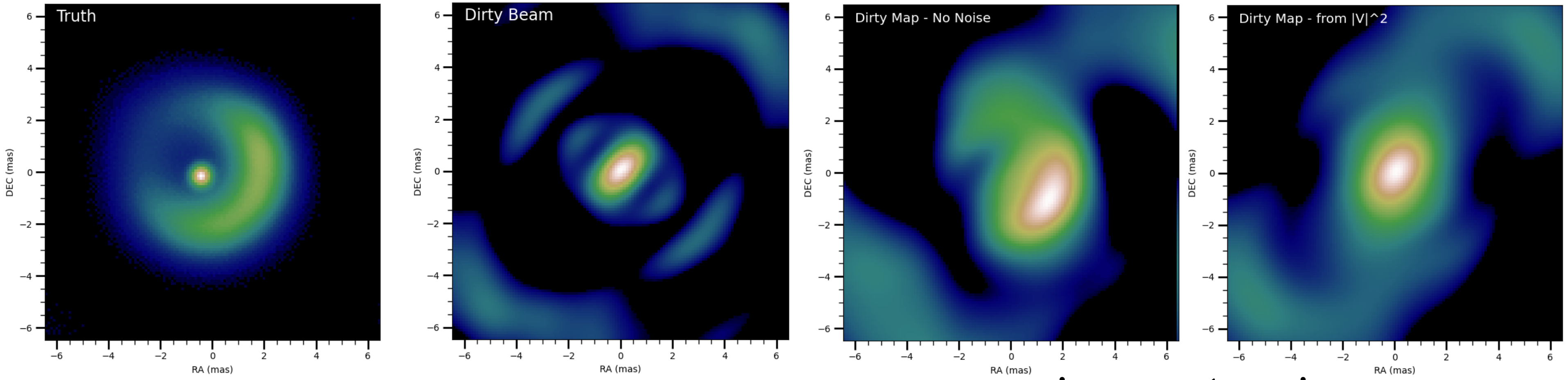
Phase of complex visibilities

$\boldsymbol{C\phi}, \boldsymbol{D\phi}$

Closure phases (via phases)
Differential phases



# Direct Inversion: dirty beam, dirty map




$$\boldsymbol{V} = \boldsymbol{H}.\boldsymbol{x} \implies \boldsymbol{x}_{\text{inverted}} = (\boldsymbol{H}^{\dagger}\boldsymbol{H})^{-1}\boldsymbol{H}^{\dagger}.\boldsymbol{V}$$

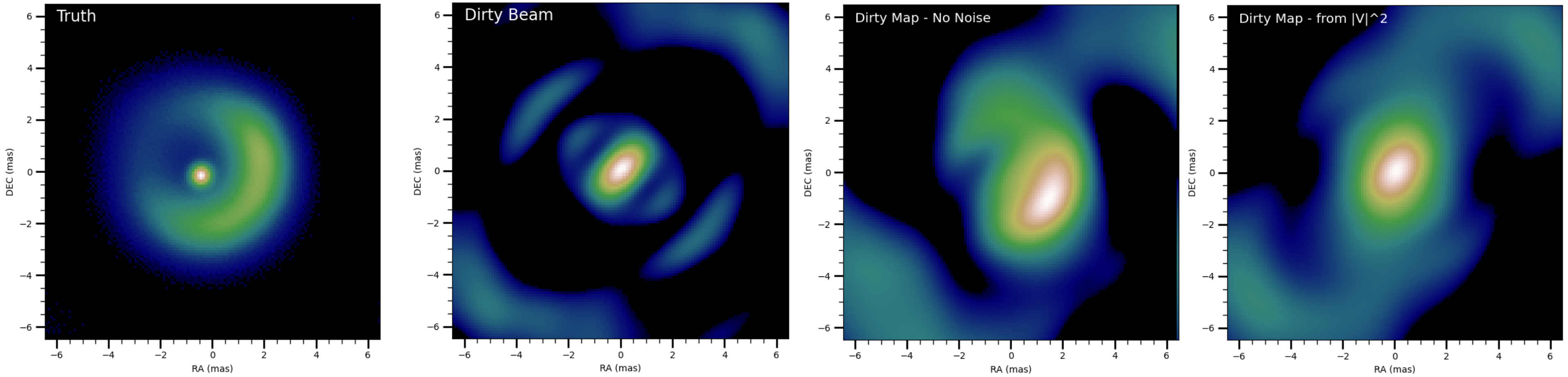
$\mathbf{H}$ only depends on the uv coverage + wavelengths + image size/field of view

Set $\mathbf{V} = 1$ for all uv points → **dirty beam** image, akin to the PSF of the observations

Use magic to measure noiseless complex visibilities $\mathbf{V}$ at each uv point → **dirty map** image

Use actual measured $|\boldsymbol{V}|^2$ data to infer $\mathbf{V}$ → dirty map from $|\boldsymbol{V}|^2$

# Direct Inversion: dirty beam, dirty map




Yet this data set is one of the easiest to image.

We often have more pixels in the sought image than measurements, so the image restoration problem is **ill-posed**.
Even when there are lots of data, the effective amount of information contained in them is still insufficient to properly recover all the image information.

We need a better approach to this ill-posed inverse problem.

# Bayes theorem to the rescue

We want to **compute** the probability of an image ***x***, knowing the data ***D*** and a model of image formation and background information ***M***

$$Pr(\boldsymbol{x}|\boldsymbol{D},\boldsymbol{M}) = \frac{Pr(\boldsymbol{D}|\boldsymbol{x},\boldsymbol{M})Pr(\boldsymbol{x}|\boldsymbol{M})}{Pr(\boldsymbol{D}|\boldsymbol{M})} \propto \mathcal{L}(\boldsymbol{x})\pi(\boldsymbol{x})$$

Posterior probability of the image

Evidence

Likelihood

Prior

$$Pr(\boldsymbol{D}|\boldsymbol{x},\boldsymbol{M}) = \mathcal{L}(\boldsymbol{x})$$

Likelihood of the image → how well the image fits the data

$$Pr(\boldsymbol{x}|\boldsymbol{M}) = \pi(\boldsymbol{x})$$

Our prior knowledge of the image → what we would expect the image to look like even without new data. ***M*** includes our choice of priors and how strongly we enforce them.

$$Pr(\boldsymbol{D}|\boldsymbol{M}) = \mathrm{constant}$$

Marginal likelihood or “evidence” = integrated likelihood & prior, how well the data can be described by our assumptions.
For a given dataset and choice of prior, this is constant.

# Regularized maximum likelihood

- Taking the log of the previous expressions

$$-\log Pr(\boldsymbol{x}|\boldsymbol{D},\boldsymbol{M}) = -\log \mathcal{L}(\boldsymbol{x}) - \log \pi(\boldsymbol{x}) = \frac{1}{2}\chi^2(\boldsymbol{x}) + \mu R(\boldsymbol{x})$$

$$\mathcal{L}(\boldsymbol{x}) \propto e^{-\frac{1}{2}\chi^2(\boldsymbol{x})}$$

Gaussian likelihood, normal error distribution on data

$$\pi(\boldsymbol{x}) \propto e^{-\mu R(\boldsymbol{x})}$$

Ad hoc expression, R is called a prior or regularization function

- The more likely an image, the lower the criterion $J(\boldsymbol{x}) = \frac{1}{2}\chi^2(\boldsymbol{x}) + \mu R(\boldsymbol{x})$

- The most likely image given ***D*** and ***M*** is found via ***maximum a posteriori***, also known as ***regularized maximum likelihood***:

$$x_{\mathrm{opt}} = \underset{\boldsymbol{x}\in\mathbb{R}^{N\times N}}{\operatorname{argmin}} J(\boldsymbol{x}) = \underset{\boldsymbol{x}\in\mathbb{R}^{N\times N}}{\operatorname{argmin}} \left\{\chi^2(\boldsymbol{x}) + \mu R(\boldsymbol{x})\right\}$$

# The Likelihood term: what is it?

$$\mathcal{L}(\boldsymbol{x}) \propto e^{-\frac{1}{2}\chi^2(\boldsymbol{x})}$$

- The likelihood measures how well the image fits the data
- Data is assumed with Normal distribution, as specified in the OIFITS standard
- Multiple observables in different tables (OI_VIS, OI_V2, OI_T3) have Gaussian distribution characterized by nominal value + error bar

$$\chi^2(\boldsymbol{x}) = \chi^2_{\mathrm{v2}} + \chi^2_{\mathrm{t3phi}} + \chi^2_{\mathrm{t3amp}} + \chi^2_{\mathrm{diffvis}}$$

- Current image $\boldsymbol{x}$ → current visibilities → current observables, to compare to actual data recorded

# Reminder: from model observables to $\chi^2$

| Computed | Measured (OIFITS data) |
|---|---|
| $\|V\|^2$ | $\boldsymbol{V2}, \sigma_{V2}$ |
| *Powerspectra* | |
| $\|\boldsymbol{T_3}\|$ | $\boldsymbol{T3amp}, \boldsymbol{\sigma}_{T3amp}$ |
| *Triple Amplitudes* | |
| $\arg \boldsymbol{T_3} = \boldsymbol{C\phi}$ | $\boldsymbol{T3phi}, \boldsymbol{\sigma}_{T3phi}$ |
| *Closure phases* | |
| $\boldsymbol{D\phi}$ | $\boldsymbol{VISphi}, \boldsymbol{\sigma}_{VISphi}$ |
| *Differential phases* | |

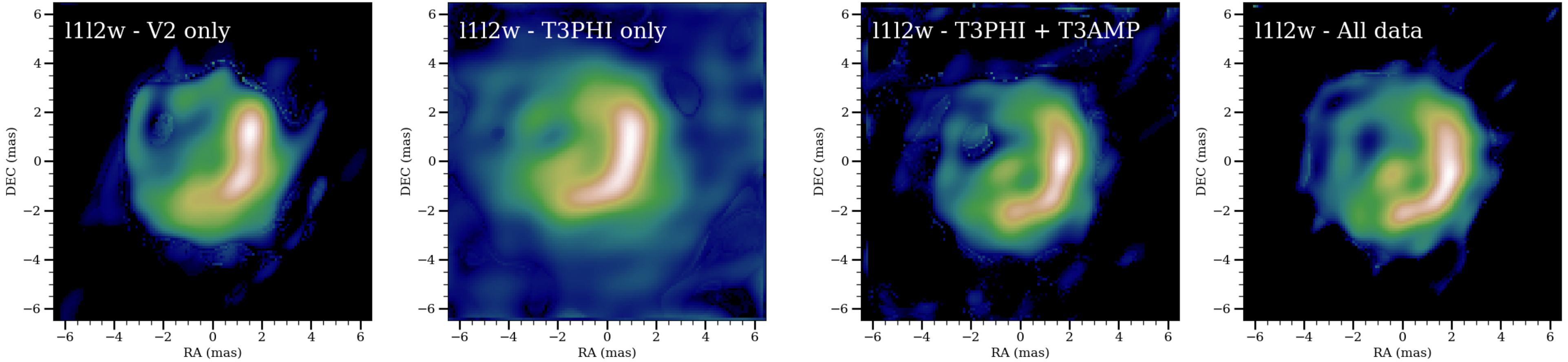
$$\boldsymbol{R} = \frac{\text{model}(\boldsymbol{x}) - \text{data}}{\sigma_{data}}$$

$$\chi^2 = ||\boldsymbol{R}||_2^2 = \sum_i R_i^2$$

$$\chi^2_{V2} = \sum_{i=1}^{n_{|V|^2}} \left( \frac{|V|_i^2 - V2_i}{\sigma_{i,V2}} \right)^2$$

Similar expressions for t3amp, t3phi, visphi

# Effect of the different observables on imaging




The powerspectrum data encode the Fourier radial profile of the target.

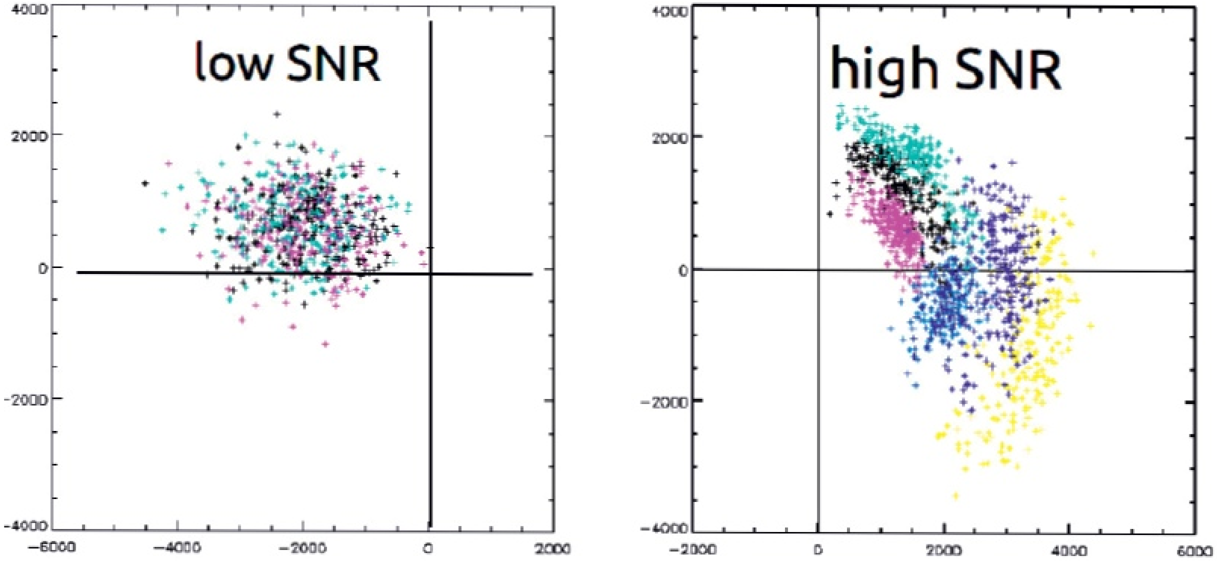
It is often said that “closure phases encode the flux asymmetries of the target” but not the location of said flux.

Reconstruction from bispectrum (initially what IRBiS did) look good.

Powerspectra and triple amplitudes provide different measurment of the visibility amplitudes.

# Common issues with likelihood

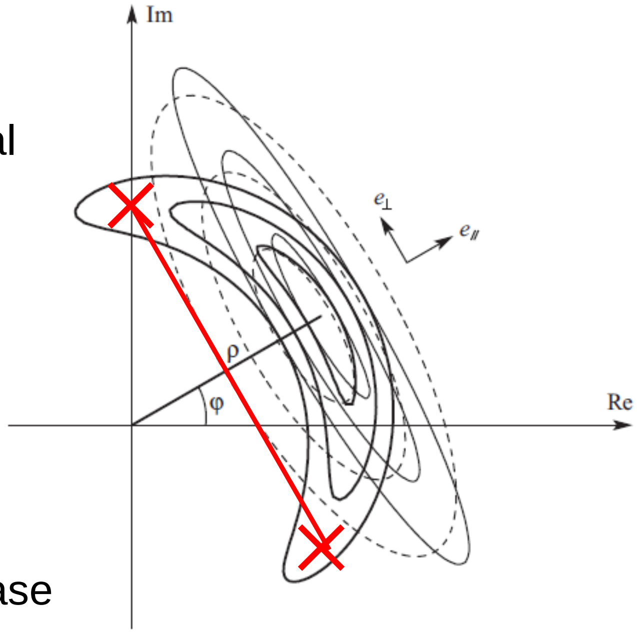
- Hidden correlations between observables, and between spectral channels (depends on instrument)
- Insufficient by itself
  - Previous images included regularization
- Hard to minimize
  - Multimodal due to missing phase (local minima)
  - Non-convex

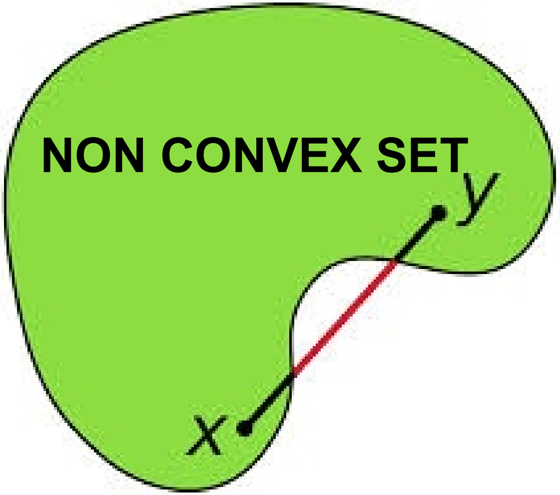



*Bispectrum noise distribution*

# Common issues with likelihood

- Hidden correlations between observables, and between spectral channels (depends on instrument)
- Insufficient by itself
  - Previous images included regularization
- Hard to minimize
  - Multimodal due to missing phase (local minima)
  - Non-convex



# Common issues with likelihood

- Hidden correlations between observables, and between spectral channels (depends on instrument)
- Insufficient by itself
  - Previous images included regularization
- Hard to minimize
  - Multimodal due to missing phase (local minima)
  - Non-convex

# Image reconstruction codes

| Software | MiRA | MiRA 3D | BSMEM (discontinued) | PAINTER (unused) | IRBiS | SQUEEZE (formely MACIM) | OITOOLS & ROTIR |
|---|---|---|---|---|---|---|---|
| Optimizer | VMLMB Semi-Newton method with inexact line search | Alternating Direction Method of Minimizers | Trust region method with Nonlinear Conjugate Gradient step | Alternating Direction Method of Minimizers | LBFGS | Simulated annealing, parallel tempering | VMLMB or Half quadratic + patch priors |
| Regularizers (not exhaustive) | Soft support, entropic priors, field of view, total variation, ... | MiRA's + group sparsity and other polychromatic ones | Smothness l2 and Maximum Entropy. | L1 norm of Wavelets | Tikhonov, edge preserving smoothness, maximum entropy | L0, Shannon entropy, total variation, laplacian, wavelets, group sparsity | All classic ones + patch priors |
| Polychromatic | No | Yes | No | Yes | Yes | Yes | Yes |
| Dynamical imaging | No | No | No | No | No | No | Yes |
| Simultaneous Model fitting | Yes | No? | No | No | No | Yes | Yes |

# Optimization methods: local optimization

- Local optimization: we start from an initial image with a given chi2 and we go down from there
  - Solution 1: gradient descent methods (MiRA, IRBIS, OITOOLS)
    - The **classic way of solving the minimization problem**, requires the analytic expressions of the gradient of the chi2 and of the regularizations with respect to vector **x**

Simple Tikohnov

$$R(\boldsymbol{x}) = ||\boldsymbol{x}||^2 = \sum_i x_i^2 \implies \frac{\partial R}{\partial \boldsymbol{x}} = 2\boldsymbol{x}$$

$$\frac{\partial \chi^2}{\partial \boldsymbol{x}}$$ involves more complex differentiation

# Optimization methods: local optimization

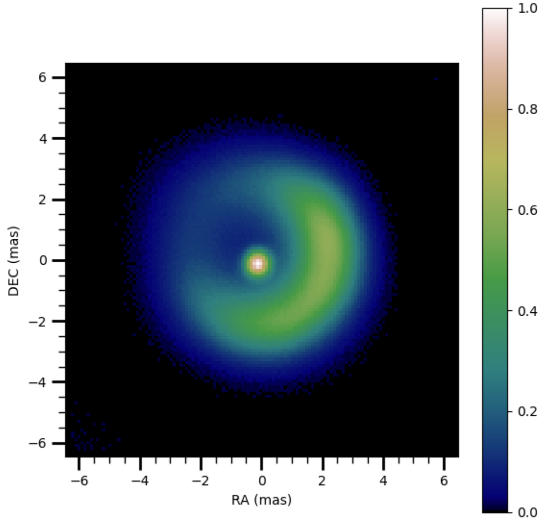
Example of gradient for the powerspectrum

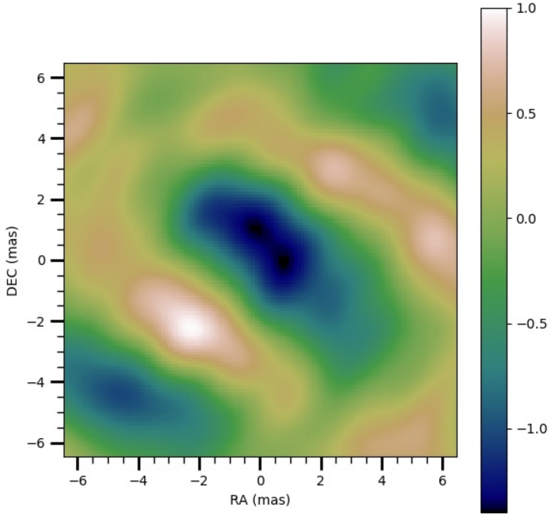
$$\chi^2 = \sum_k \frac{(|\boldsymbol{V}|^2_k - |V|^2_{k,\mathrm{data}})^2}{\sigma_k^2}$$

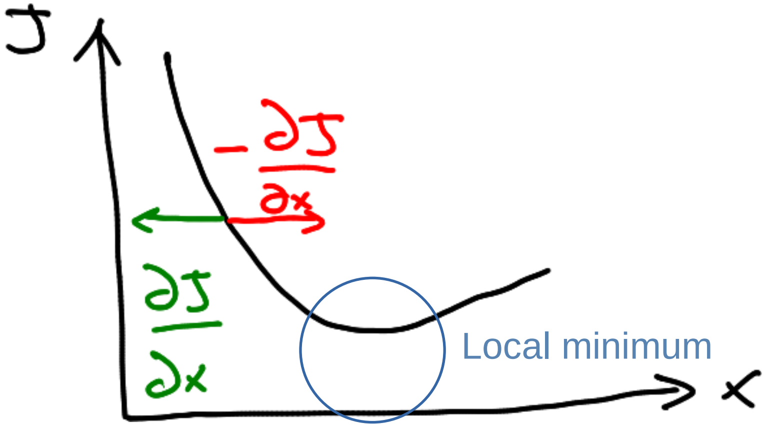
$$\frac{\partial \chi^2}{\partial x_i} = \sum_k 2\,\frac{|\boldsymbol{V}|^2_k - |V|^2_{k,\mathrm{data}}}{\sigma_k^2}\,\frac{\partial |\boldsymbol{V}|^2_k}{\partial x_i}$$

$$\frac{\partial |\boldsymbol{V}|^2_k}{\partial x_i} = \frac{\partial \boldsymbol{V}_k \boldsymbol{V}^*{}_k}{\partial x_i} = 2\,\mathcal{R}\mathrm{e}\left\{\boldsymbol{V}^*{}_k \frac{\partial \boldsymbol{V}_k}{\partial x_i}\right\} = 2\,\mathcal{R}\mathrm{e}(\boldsymbol{V}^*{}_k H_{ik})$$

# Optimization methods: local optimization




The gradient $\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}}$ has the same size as the current image.

$-\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}}$ is the **steepest descent direction**. Going this way will decrease the criterion.

Advanced codes like MiRA make better use of the gradient than the steepest descent and also approximate the second derivative as needed.

# Optimization methods: local optimization, ADMM

- Local optimization: we start from an initial image with a given chi2 and we go down from there
  - Solution 2: ADMM (MiRA-3D, PAINTER)
    - Alternating minimization scheme between likelihood minimization and regularization
    - **Ideal for polychromatic imaging**: very efficient parallelization to very large number of spectral channels
    - Availability of proximal operators limits regularization choices
    - **Finicky convergence** (hyperparameter need tweaking)
    - Has not been used yet “in the wild”

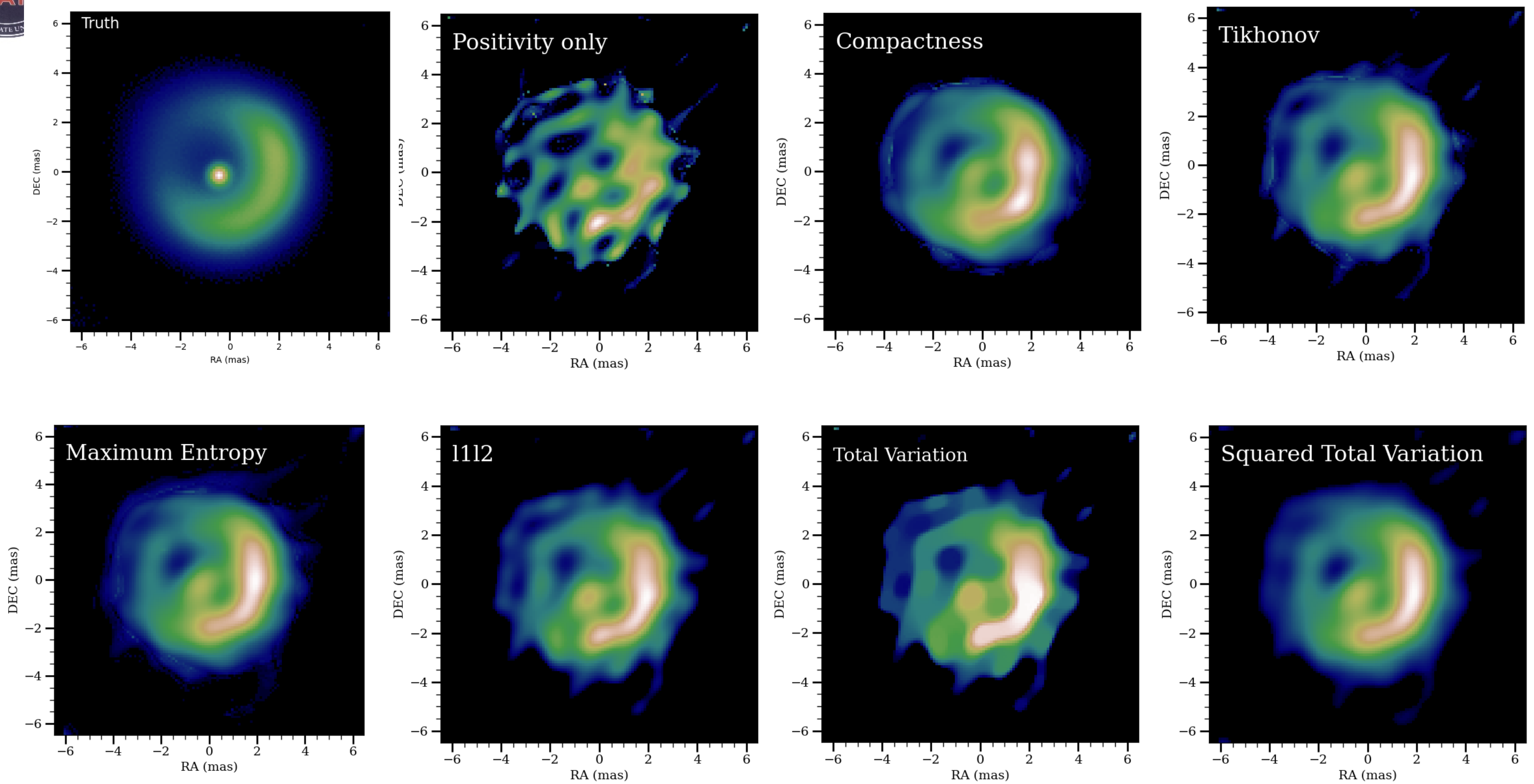
# Optimization methods: global optimization

- Global optimization: a strategy that attempts to explore the entire range of possible flux values for ***x****.*
  - Possible solution: stochastic methods (SQUEEZE)
  - A large number of flux elements are moving on an empty image
  - Each move made by these elements changes the criterion
  - SQUEEZE uses simulated annealing: moves improving the chi2 are accepted, moves worsening it are not always rejected
  - **Slow exploration**: unusable for polychromatic imaging with more than 10 spectral channels
  - **Avoids local minima**: somewhat verified in practice
  - **Flexible regularization**: no need for analytic gradients of regularizers

# Interferometry Beauty Contests


Organized every 2 years at SPIE Astronomical Instrumentation conference
**Everyone can participate. Next one in 2022!**
Constitute decent benchmarks of software capabilities… or of regularization strategies!

# Classic regularizers in OITOOLS.jl



# Choice of regularizer(s)

# Choice of regularizer(s)

# Compressed Sensing, Sparsity

- Sparsity: *an image is “sparse” in a basis if it can be expressed as a small number of non-zero coefficients in this basis*

- For a sparse image, optimal image reconstruction can be achieved (Candes 2007, Donoho 2008) by minimizing the number of non-zero coefficients in the sparsity basis

- This leads to regularizers based on the $\ell_0$ pseudonorm (= non-zero counts)

$$\ell_0(\boldsymbol{Ax}) = \sum_i \mathcal{I}(Ax)_i \qquad \mathcal{I}(\boldsymbol{z}) = \begin{cases} 1 \text{ where } z_i > 0 \\ 0 \text{ elsewhere} \end{cases}$$

non-convex
non-smooth
requires global optimization

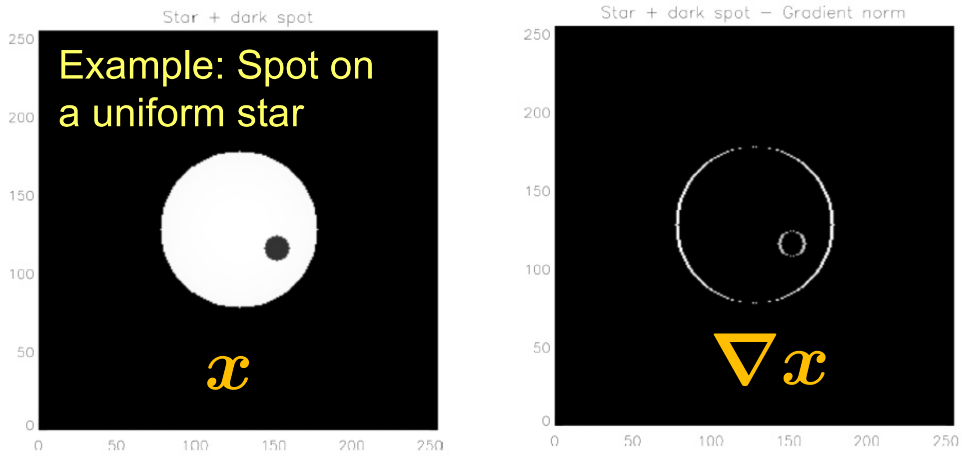
- Or based on the $\ell_1$ norm (sum of absolute values)

minimizes the number of lit-up pixels

$$\ell_1(\boldsymbol{Ax}) = \sum_i |Ax|_i$$


→ Convex

# A well known example: total variation = spatial gradient sparsity

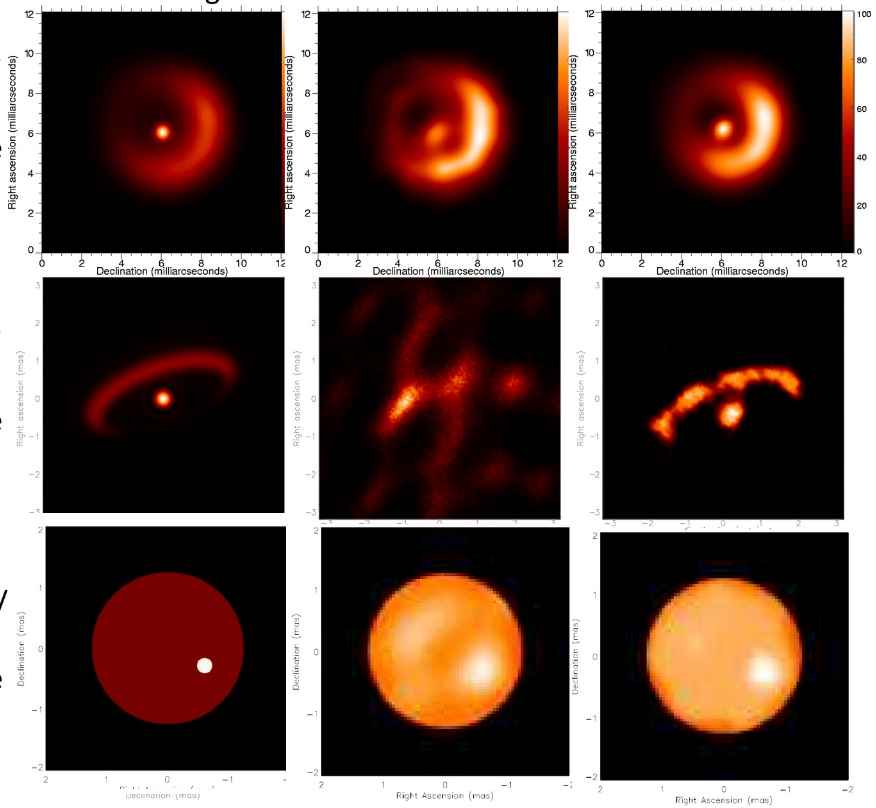



$$R_{\mathrm{TV}}(\boldsymbol{x}) = \ell_1 \left( \boldsymbol{\nabla x} \right) = \sum_i |\boldsymbol{\nabla x}|_i$$

WARNING: spatial gradient ≠ gradient of chi2 with respect to x

# Examples of Compressed sensing results

| | "Truth Image" | Positivity + entropy | Compressed Sensing Reconstruction | Sparsity basis |
|---|---|---|---|---|
| Case 1: YSO Disk w/ high-quality UV coverage | | | | Isotropic wavelets (BSWAVE) |
| Case 2: YSO Disk w/ only 4-T UV coverage | | | | Arclets (SQUEEZE) |
| Case 3: Star+Spot w/ only 4-T UV coverage | | | | Gradient (SQUEEZE UD) |




*Simple compressed sensing results*
*Baron et al. 2014*

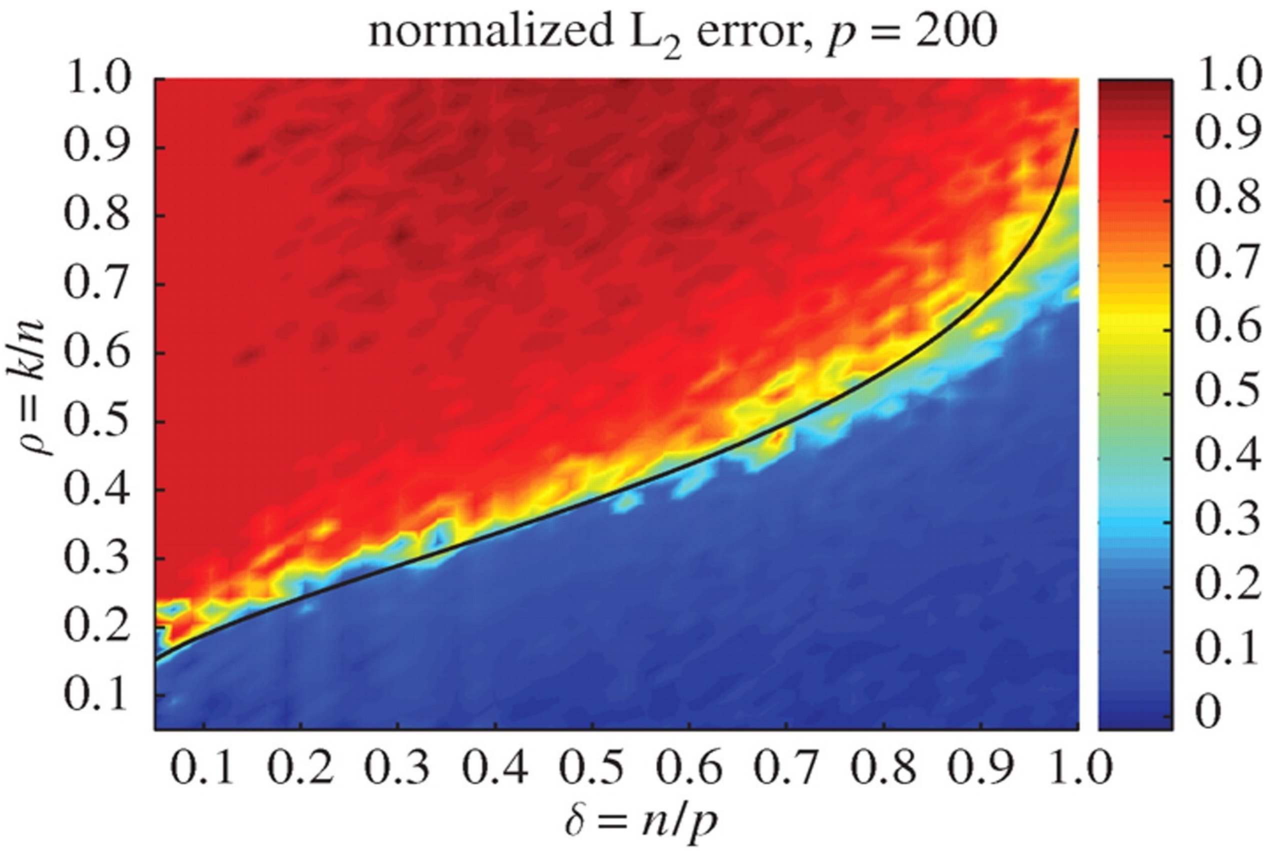
$\ell_p, p < 1$ are non-convex

$\ell_0$ cannot always be substituted by $\ell_1$, e.g. in image plane or any flux conservative wavelet transform.

# The limit of Compressed sensing

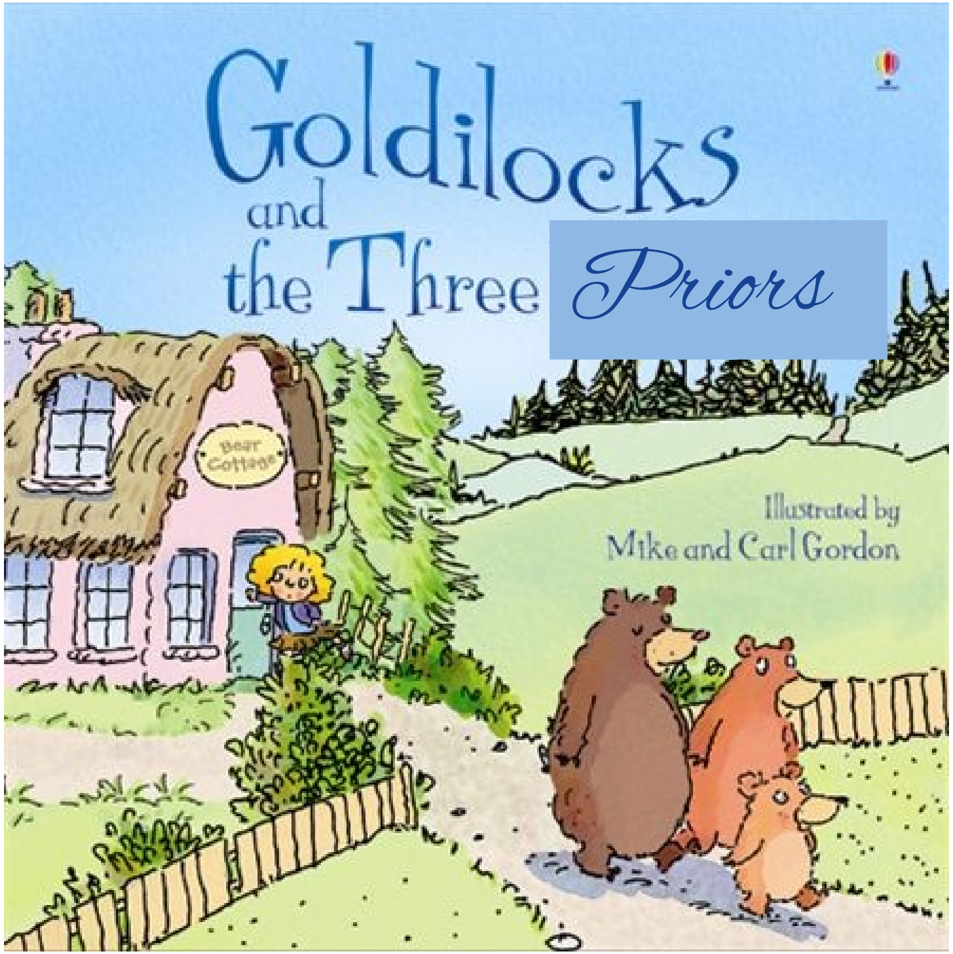
*Donoho-Tanner phase change:* below a certain number of data, probabilty of optimal recovery becomes too difficult

No magic this time!

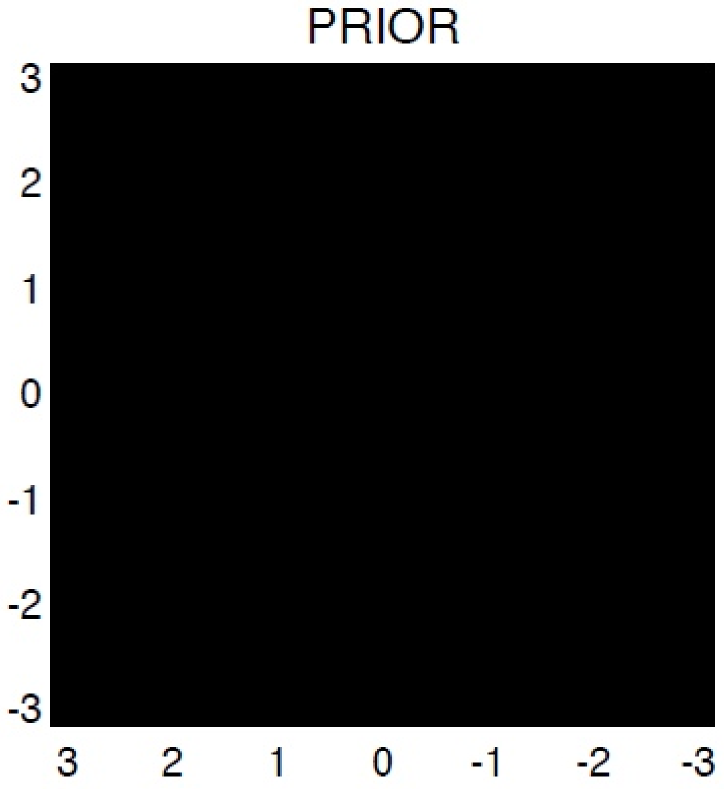


Support priors: dangerous. What happens if I use the wrong prior with too few data points?

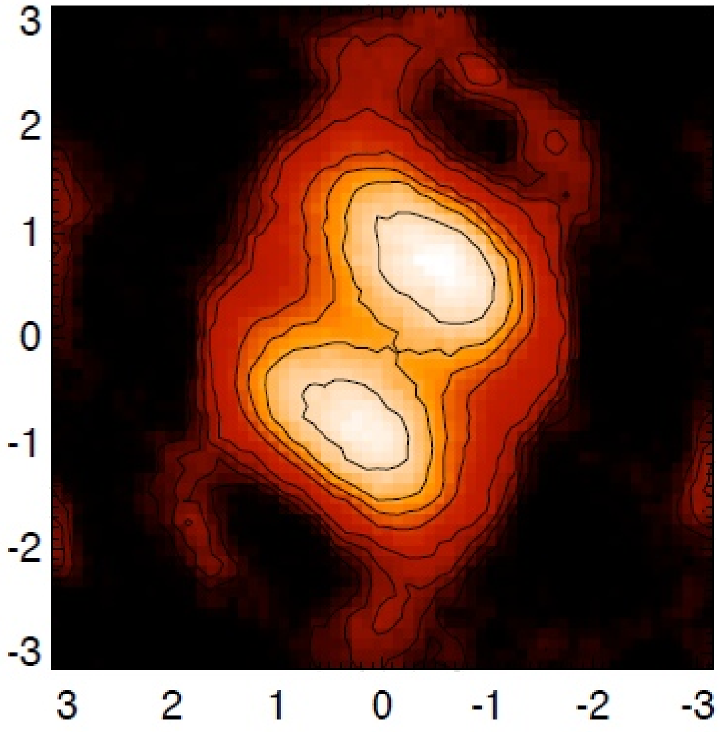
The example of using a prior image in MACIM (older version of SQUEEZE).



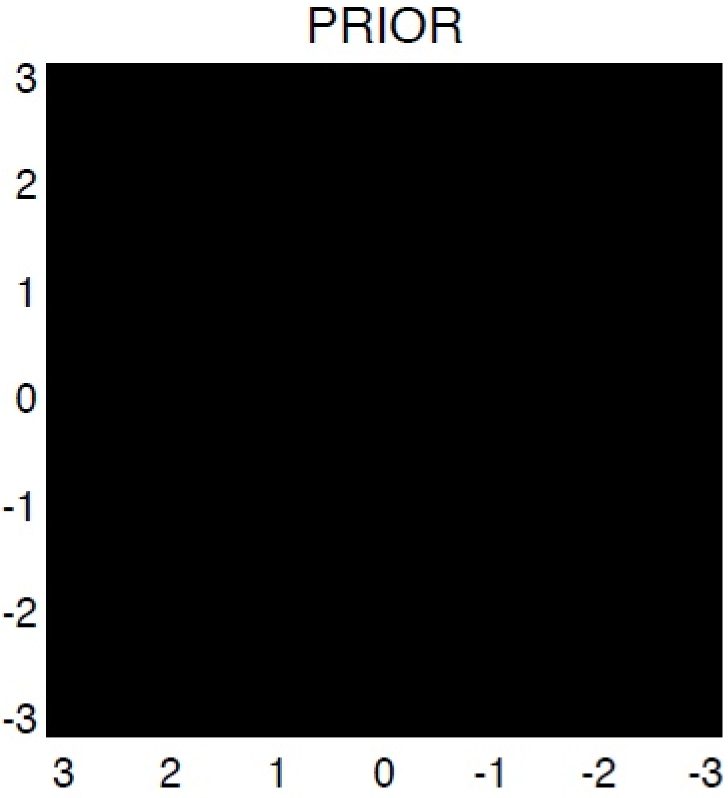
# Imaging with a wrong prior image: flat prior

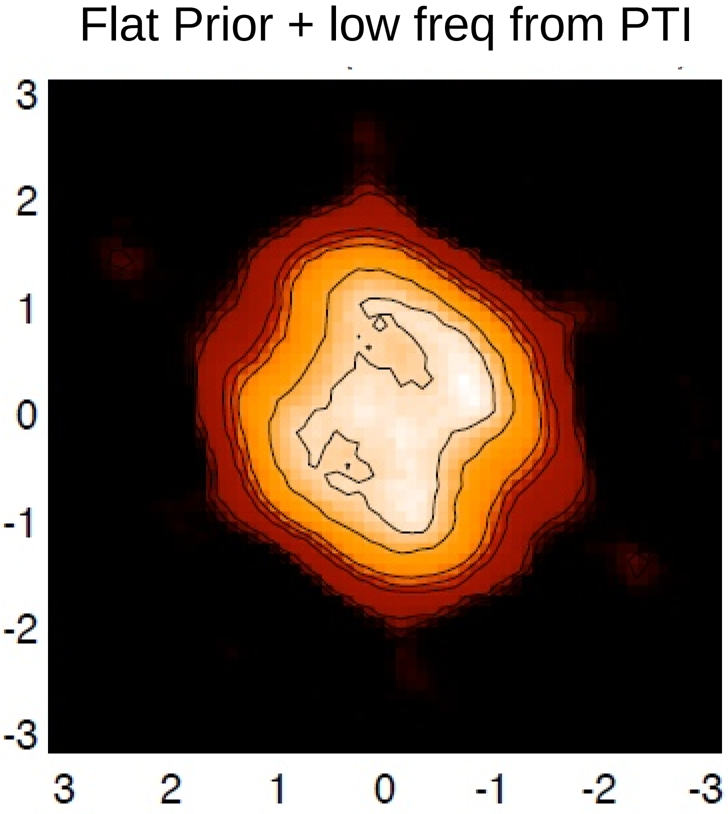



Flat Prior

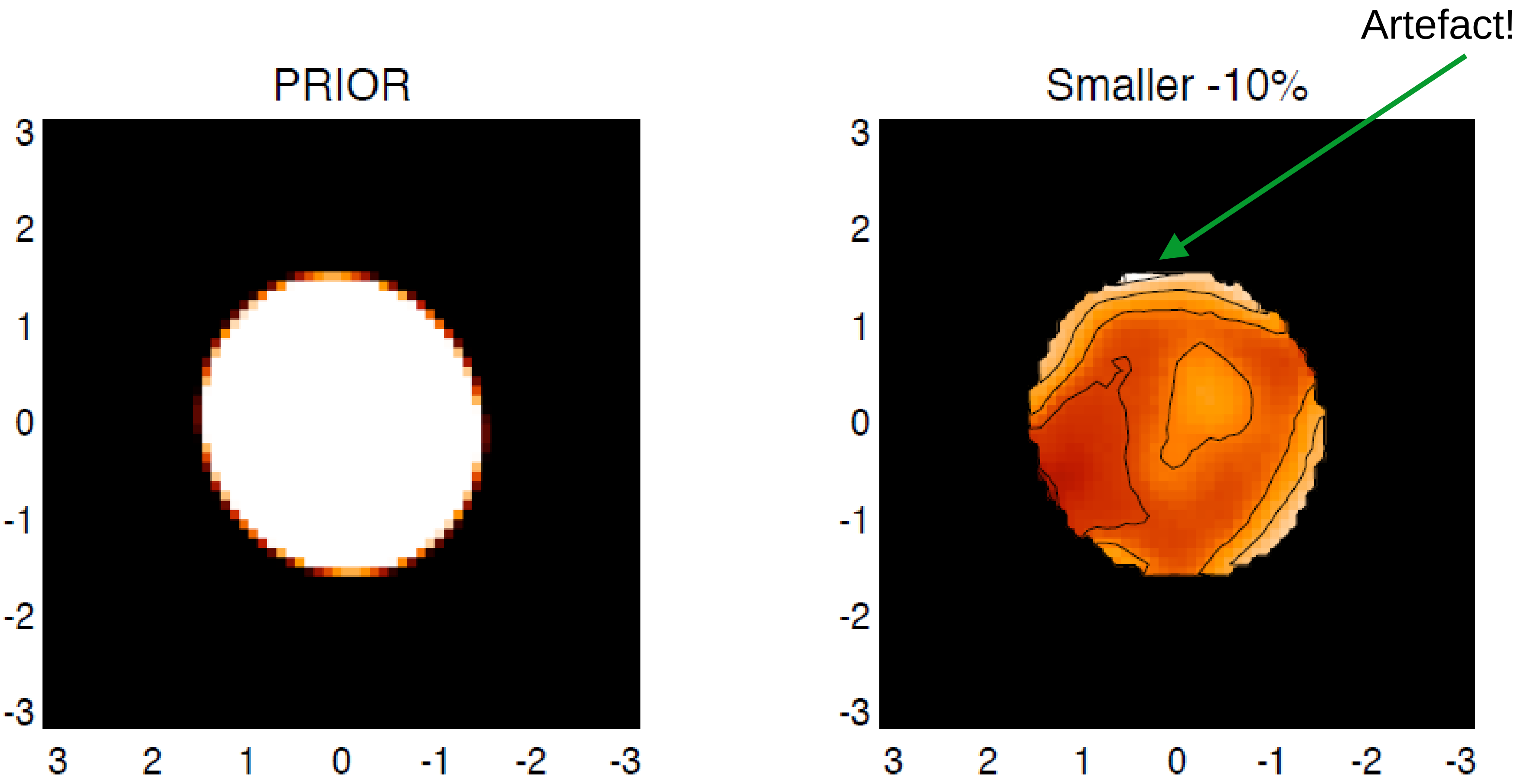


# Imaging with the wrong priors: flat prior, constrained short baselines

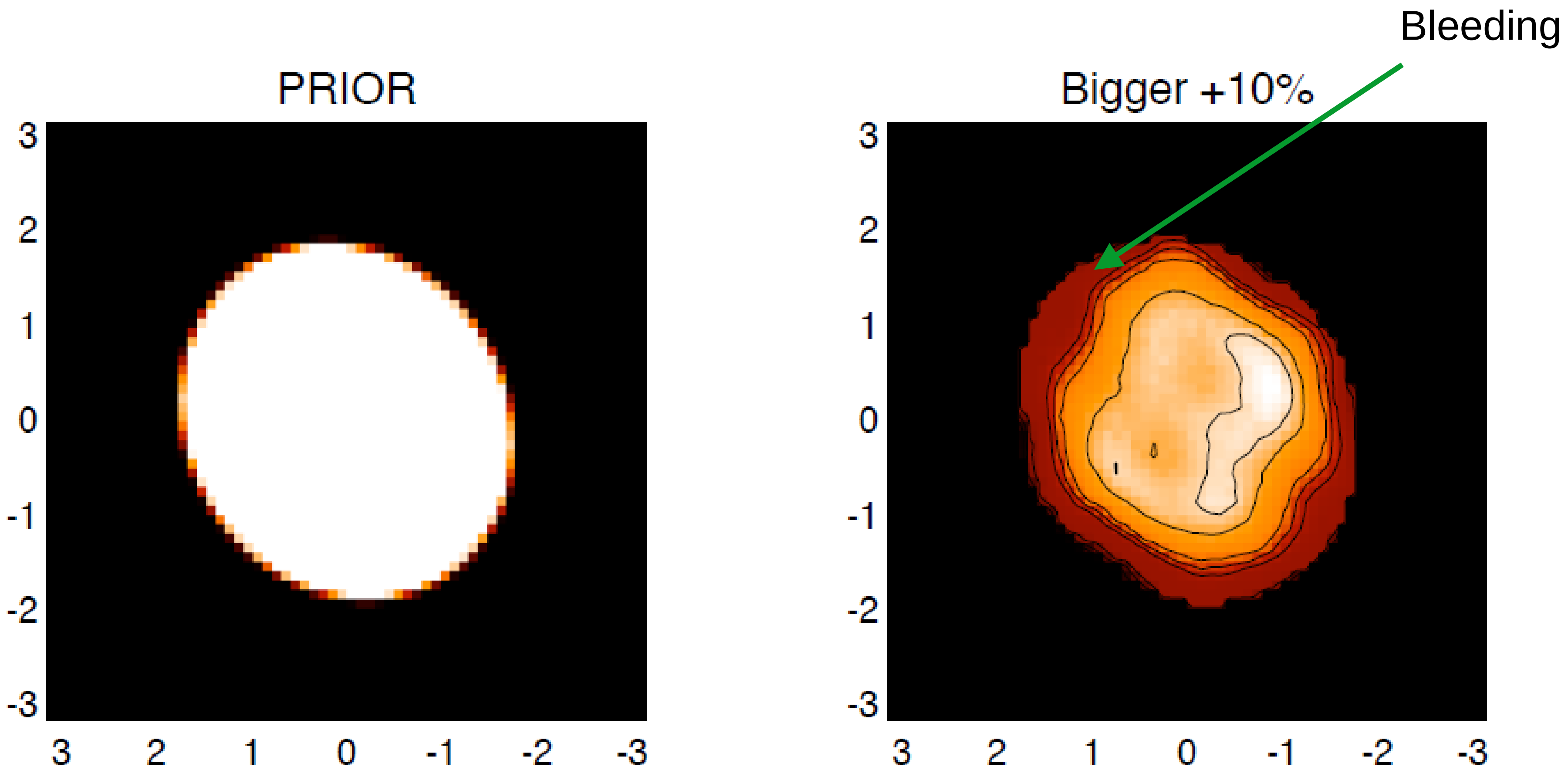


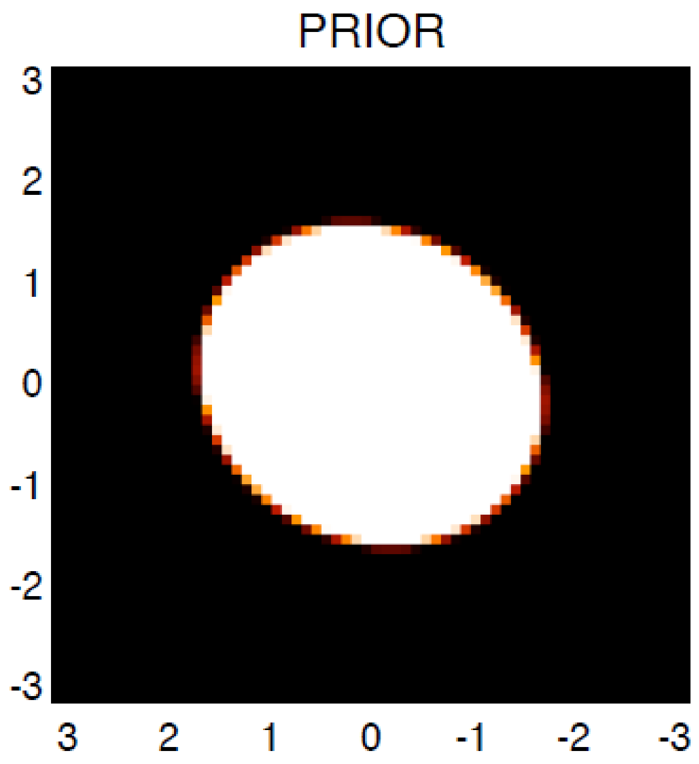
# Imaging with the wrong priors: elliptical prior, too small

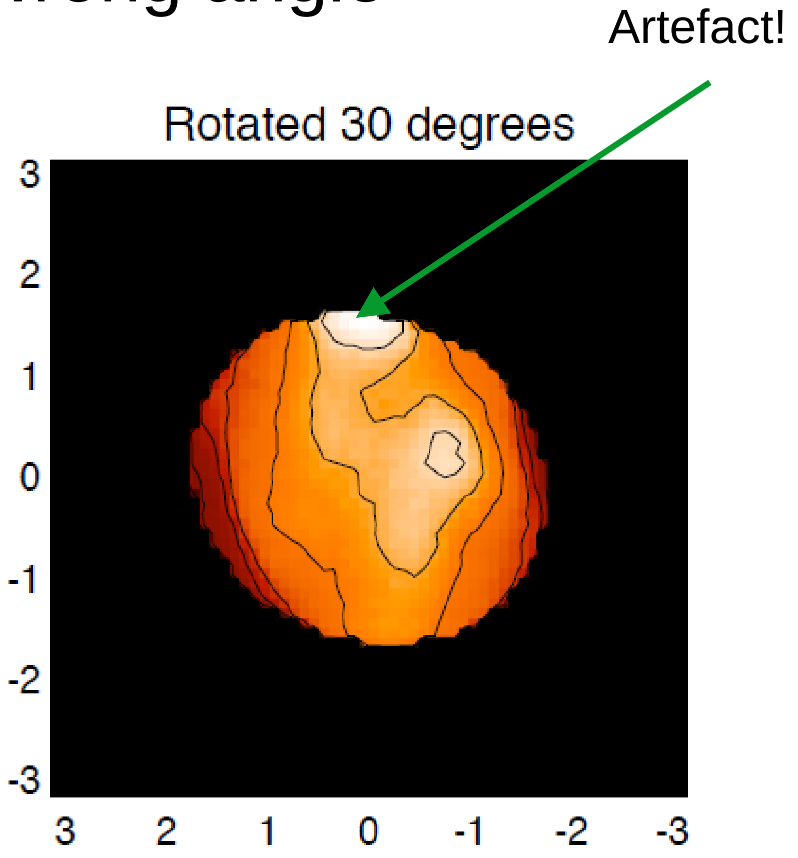


# Imaging with the wrong priors: elliptical prior, too large

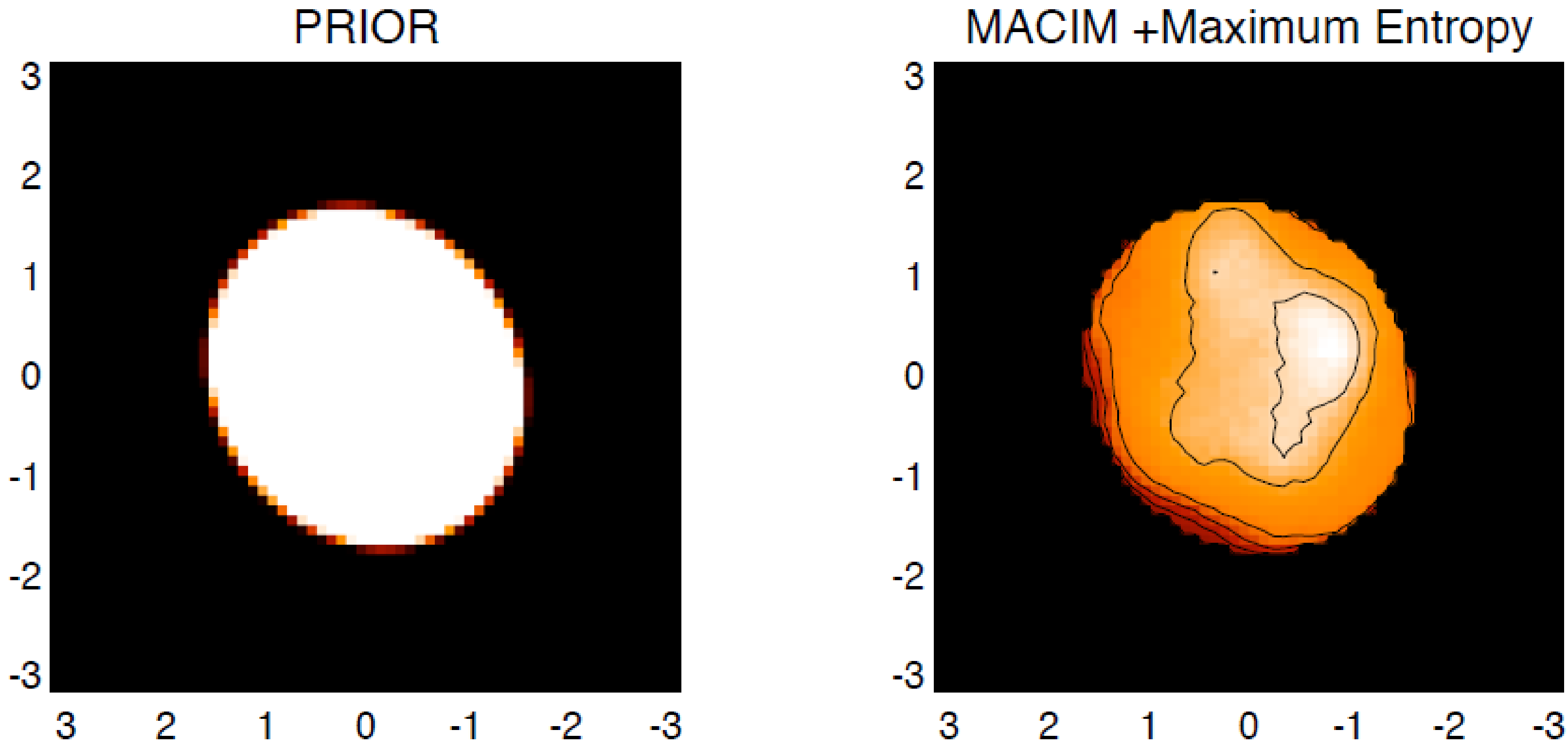


# Imaging with the wrong priors: elliptical prior, wrong angle



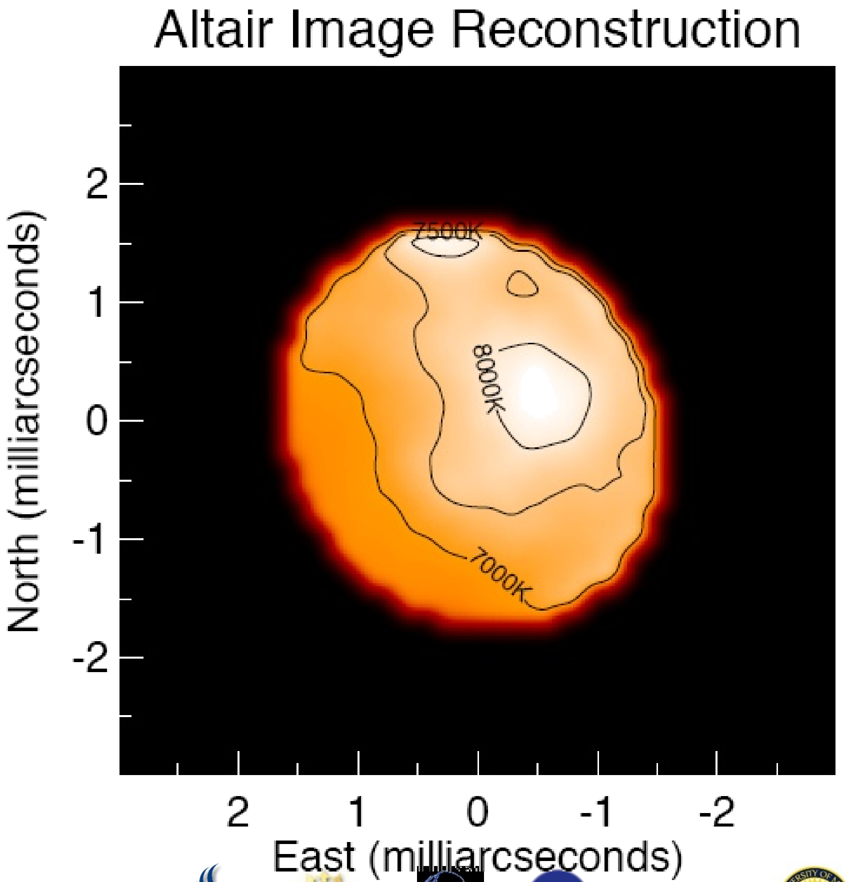



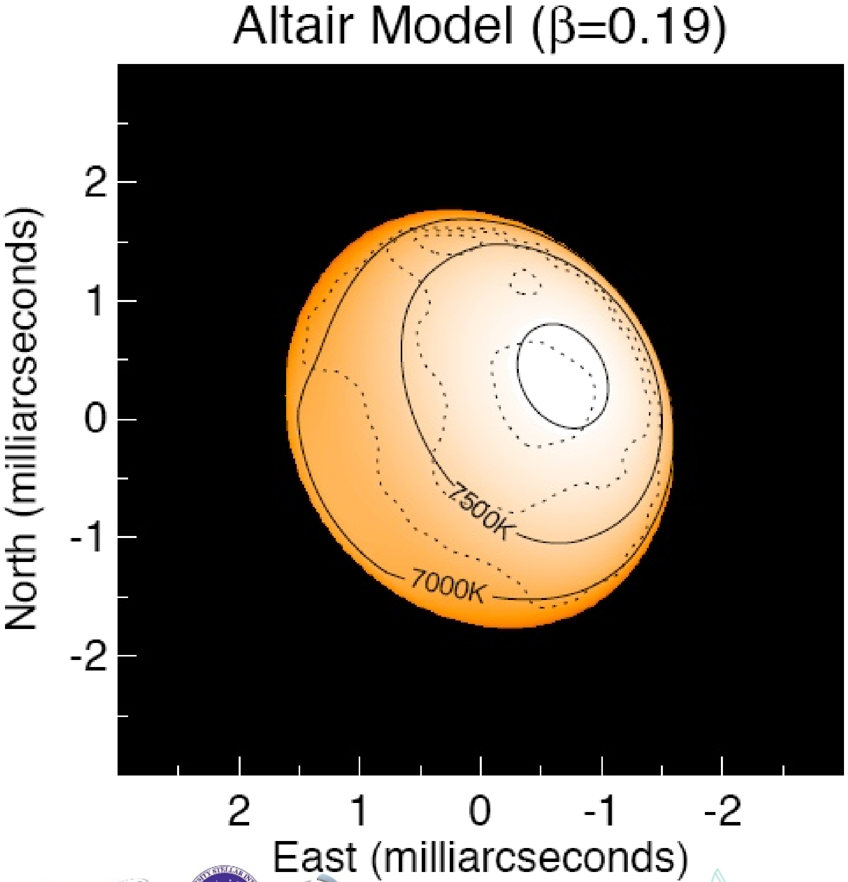
# Imaging with the wrong priors: elliptical prior, just right !



# First resolved image of a main sequence star (beyond Sun)

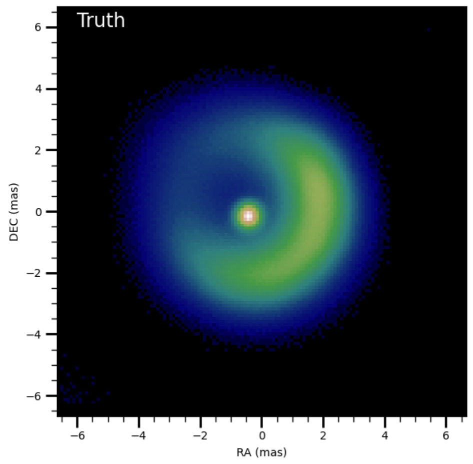
Monnier et al., *Science*, 2007

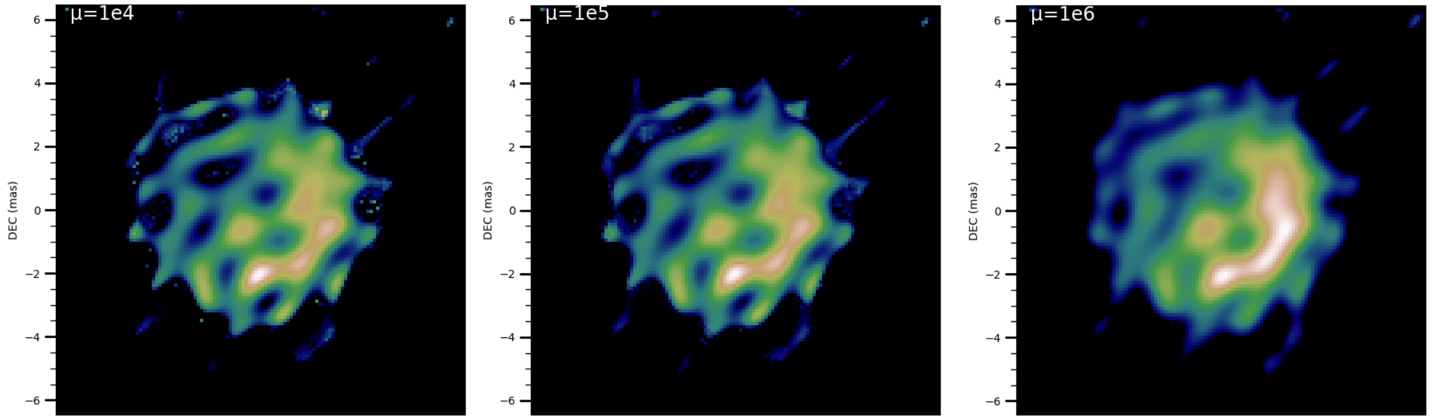


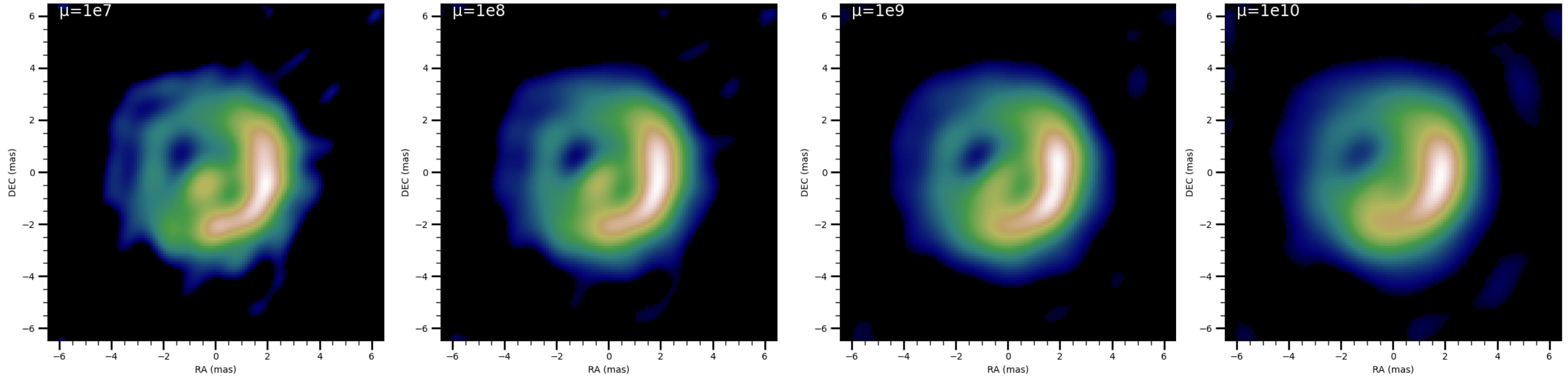



# Setting the regularization weight

# Effect of regularization weight hyperparameter μ



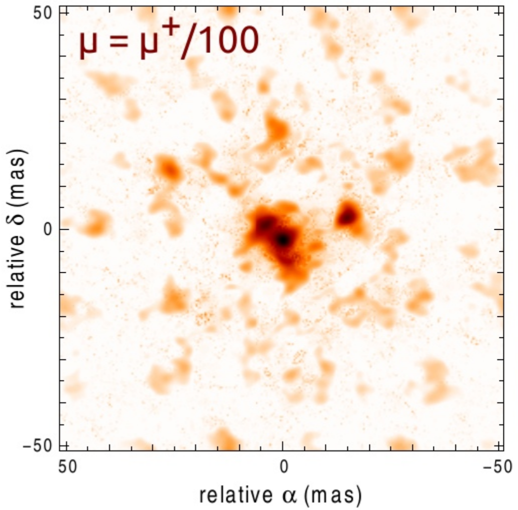



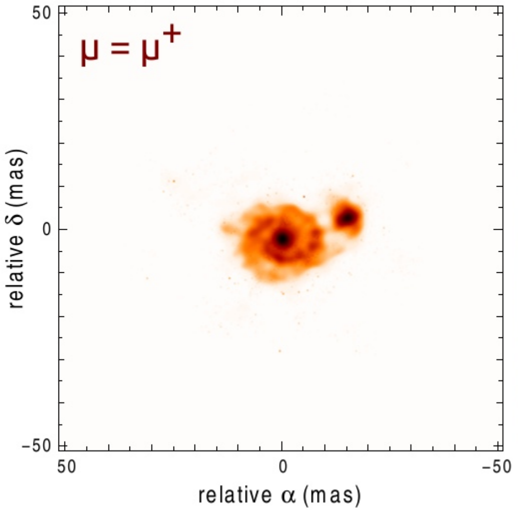



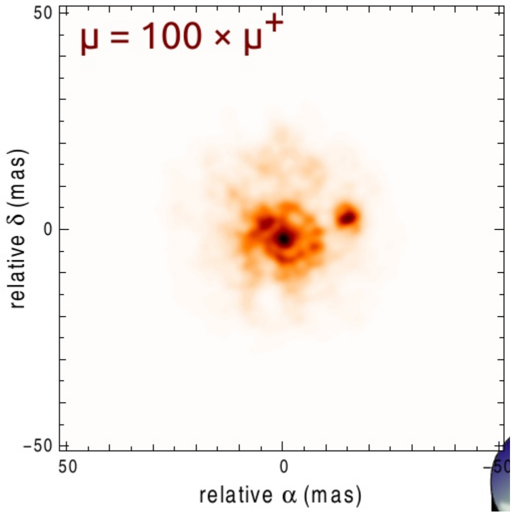

μ increases = more regularization = relative $\chi^2$ strength is lower

Georgia State University
NSF
NOAO
l'Observatoire de Paris
LESIA
NExScI
NASA Exoplanet Science Institute
SUSI
Observatoire de la CÔTE d'AZUR
Max-Planck-Institut für Radioastronomie

For some objects, one can more easily identify under and over regularization

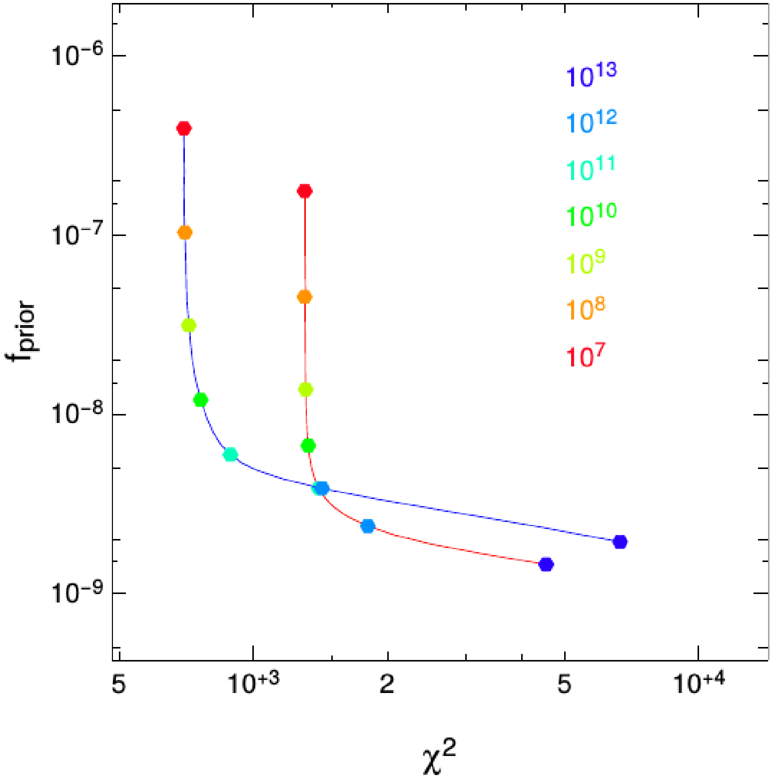



*Renard et al., 2011*

# The L-curve heuristic




*Kluska et al., 2016*
*MiRA-SPARCO with smoothness*

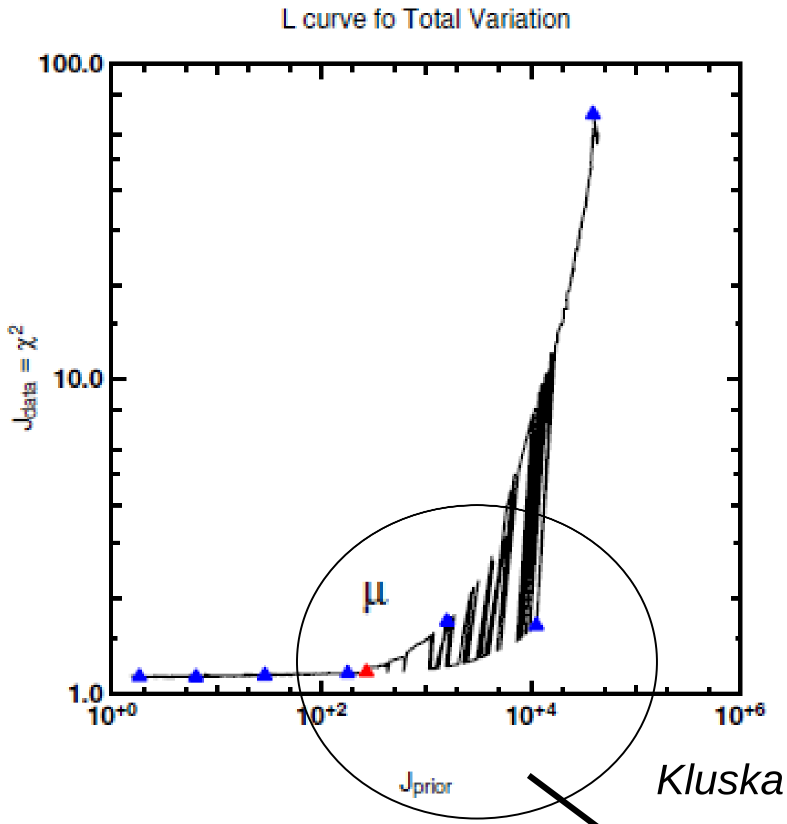
1. Run your reconstructions with different values for µ (logarithmic if no clue)

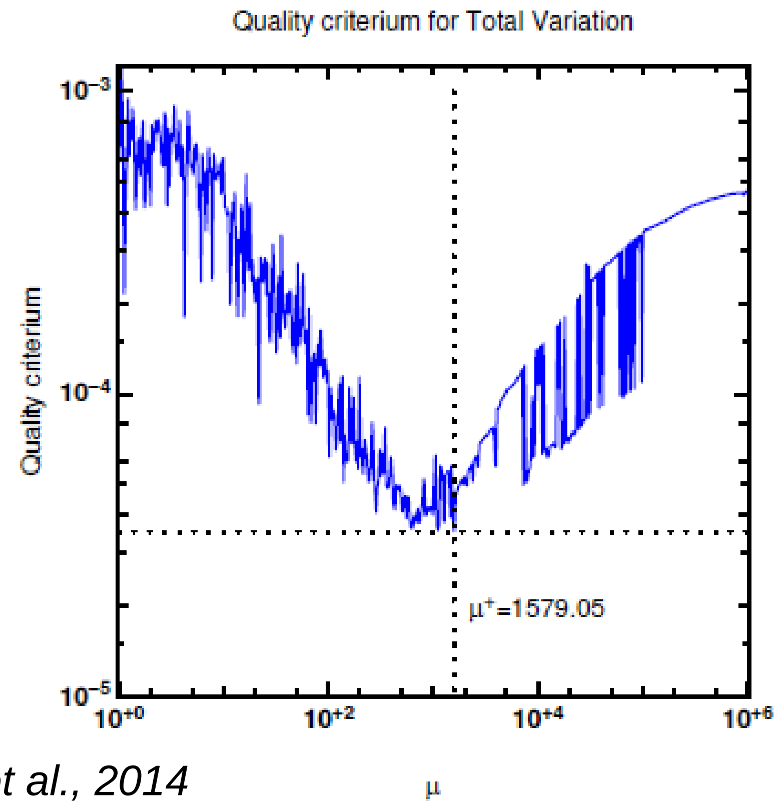
2. Plot chi2 vs regularization values in your best images

3. The best weight for µ is the one that ever slightly worsen your chi2. This corresponds to the elbow in the L-curve.

*Note: this assumes there is an optimal µ. The correct Bayesian approach would be to treat µ as another variable, described by P(µ), then marginalize over it.*

# Some issues with L-curve



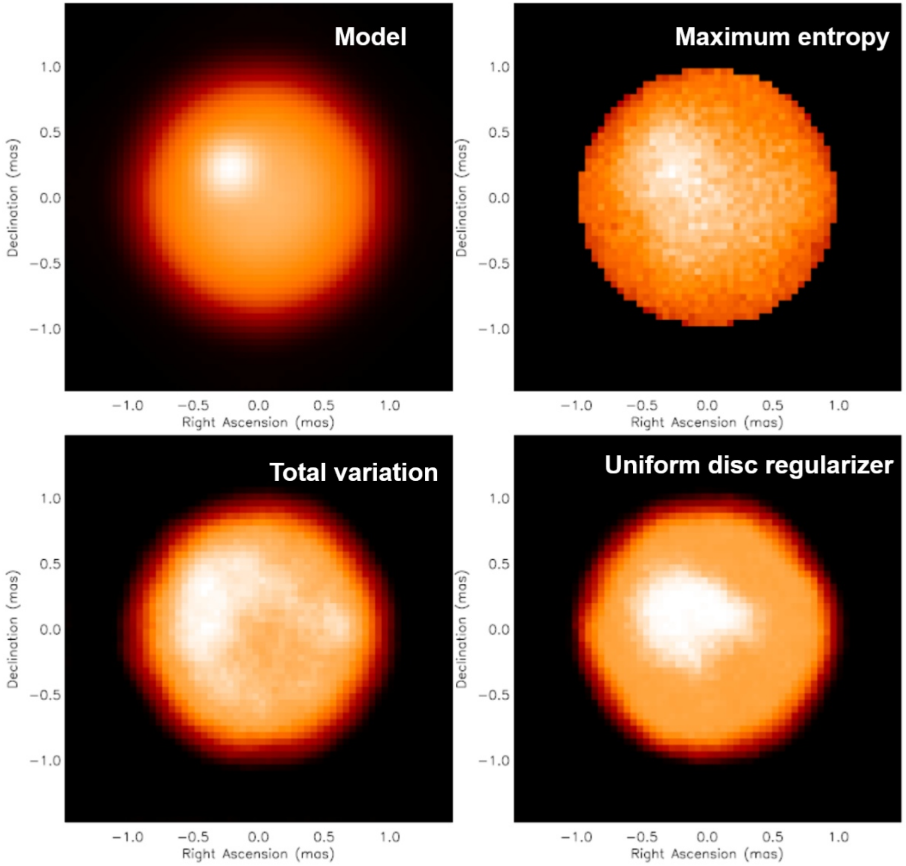




*Kluska et al., 2014*

Local minima may complicate the application of the L-curve method

# Advice on procedure

- Generate a model of the target you want to reconstruct
- Simulate the observations of this object, copying the uv coverage and signal to noise from the original data
- Reconstruct with various regularizations
- This allows to detect artefacts from the reconstruction process and to improve the regularization




*Simulations of reconstruction of large cells on red supergiants for CHARA*
*Baron et al. 2014*

# Polychromatic imaging

SPARCO
(polychromatic model)

A&A 564, A80 (2014)
DOI: 10.1051/0004-6361/201322926


Astronomy & Astrophysics

## SPARCO: a semi-parametric approach for image reconstruction of chromatic objects

### Application to young stellar objects

J. Kluska[1], F. Malbet[1], J.-P. Berger[2], F. Baron[3], B. Lazareff[1], J.-B. Le Bouquin[1], J. D. Monnier[4], F. Soulez[5], and E. Thiébaut[5]

[1] Institut de Planétologie et d'Astrophysique de Grenoble, UJF, CNRS, 414 rue de la piscine, 38400 Saint Martin d'Hères, France
e-mail: jacques.kluska@obs.ujf-grenoble.fr
[2] European Southern Observatory, Alonso de Cordóva 3107, Casilla 19001 Vitacura, Santiago, Chile
[3] Center for High Angular Resolution Astronomy, Georgia State University, PO Box 3969, Atlanta, GA 30302, USA
[4] University of Michigan Astronomy Department, 941 Dennison Bldg, Ann Arbor, MI 48109-1090, USA
[5] CRAL, Observatoire de Lyon, CNRS, Univ. Lyon 1, École Normale Supérieure de Lyon, 69364 Lyon, France

Received 28 October 2013 / Accepted 11 March 2014

MIRA 3D

## Multi-wavelength imaging algorithm for optical interferometry

Éric Thiébaut[a] & Ferréol Soulez[a]

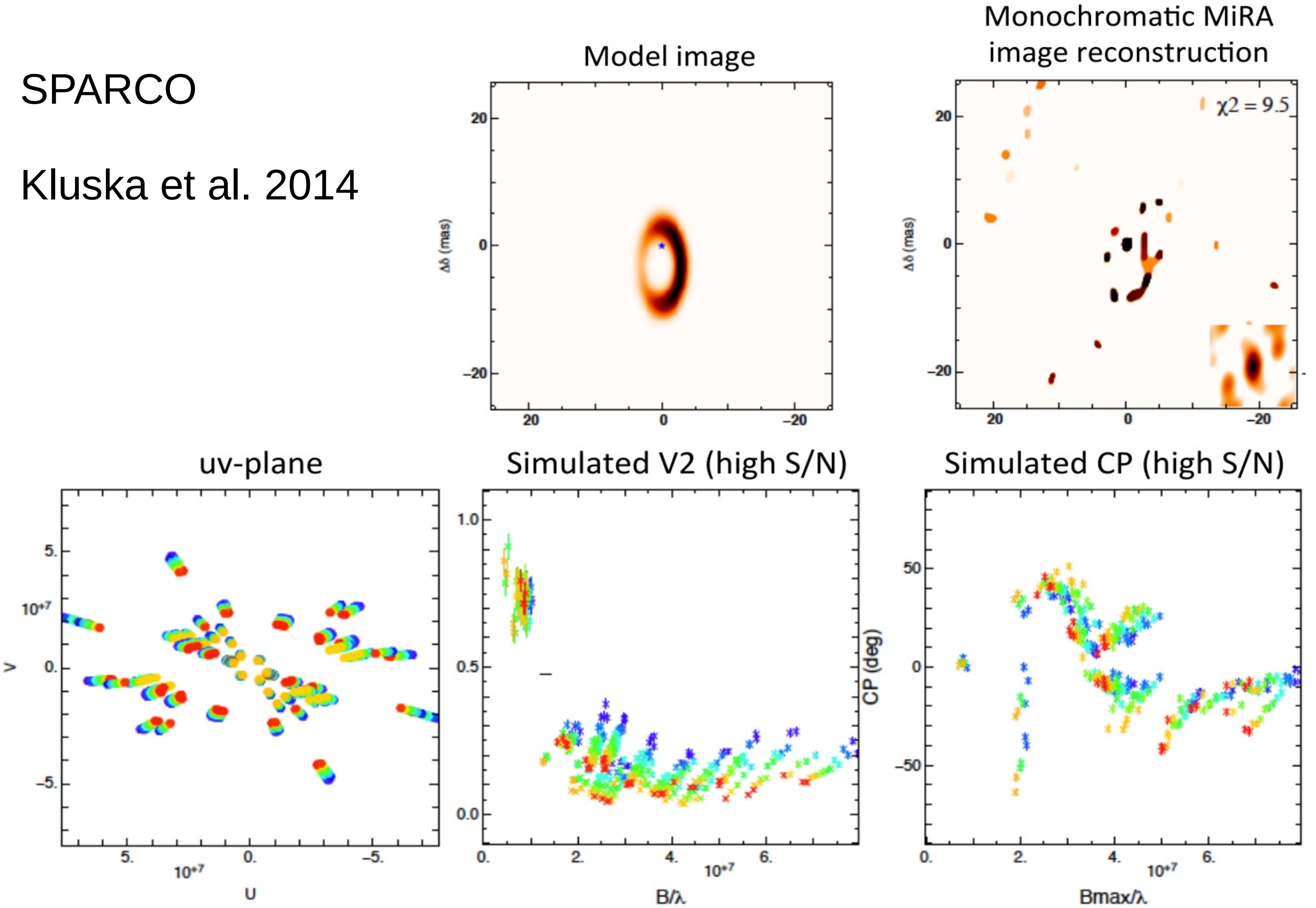
[a]Université de Lyon, Lyon, F-69003, France; Université Lyon 1, Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval, F-69230, France; CNRS, UMR 5574, Centre de Recherche Astrophysique de Lyon; École Normale Supérieure de Lyon, Lyon, F-69007, France.

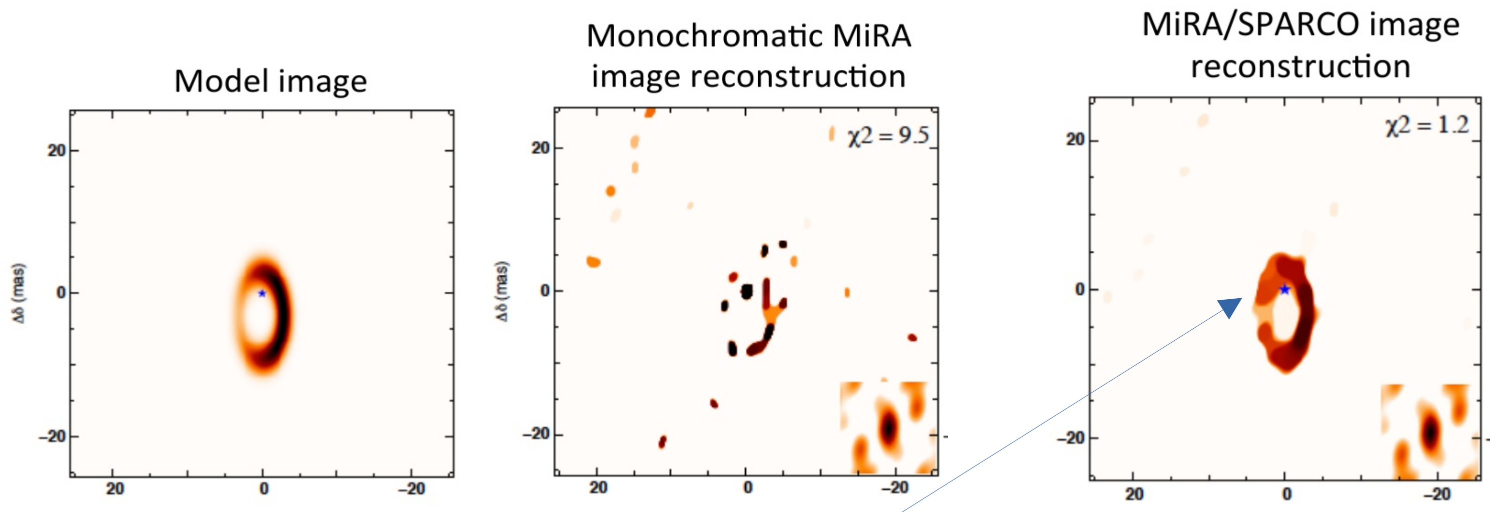
### ABSTRACT

Optical interferometers provide multiple wavelength measurements. In order to fully exploit the spectral and spatial resolution of these instruments, new algorithms for image reconstruction have to be developed. Early attempts to deal with multi-chromatic interferometric data have consisted in recovering a gray image of the object or independent monochromatic images in some spectral bandwidths. The main challenge is now to recover the full 3-D (spatio-spectral) brightness distribution of the astronomical target given all the available data. We

SPARCO

Kluska et al. 2014






$$\widetilde{V}_{\rm tot}\left(\frac{\boldsymbol{b}}{\lambda},\lambda\right)=\frac{\boxed{f_*^0\left(\frac{\lambda}{\lambda_0}\right)^{-4}+\left(1-f_*^0\right)\left(\frac{\lambda}{\lambda_0}\right)^{d_{\rm env}}\boxed{\widetilde{V}_{\rm env}\left(\frac{\boldsymbol{b}}{\lambda}\right)}}}{\boxed{f_*^0\left(\frac{\lambda}{\lambda_0}\right)^{-4}+\left(1-f_*^0\right)\left(\frac{\lambda}{\lambda_0}\right)^{d_{\rm env}}}}.$$

Unnormalized visibilities of point source and environment (= image). Both weighted by respective fluxes.

Flux normalization

Analytic model of a polychromatic point source (power law -4) + single image with spectral power law $d_{\rm env}$

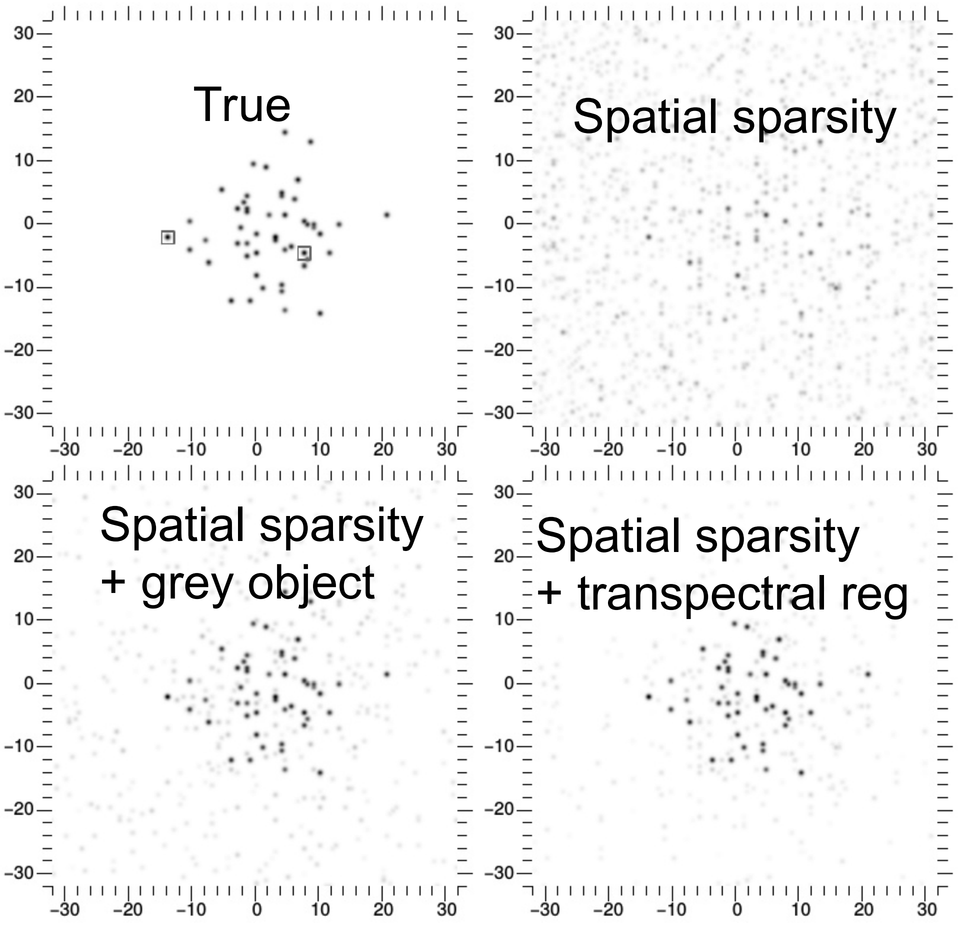
# MIRA 3D

Thiébaut et al., 2012 & 2013
Soulez et al., 2014 + ongoing work

Alternating direction method of multipliers (ADMM) –> sequence of closed form subproblems for regularization and likelihood terms

Transpectral regularization

$$R_{\text{joint}} = \sum_n \left( \sum_\lambda i_{n,\lambda}^2 \right)^{\frac{1}{2}}$$



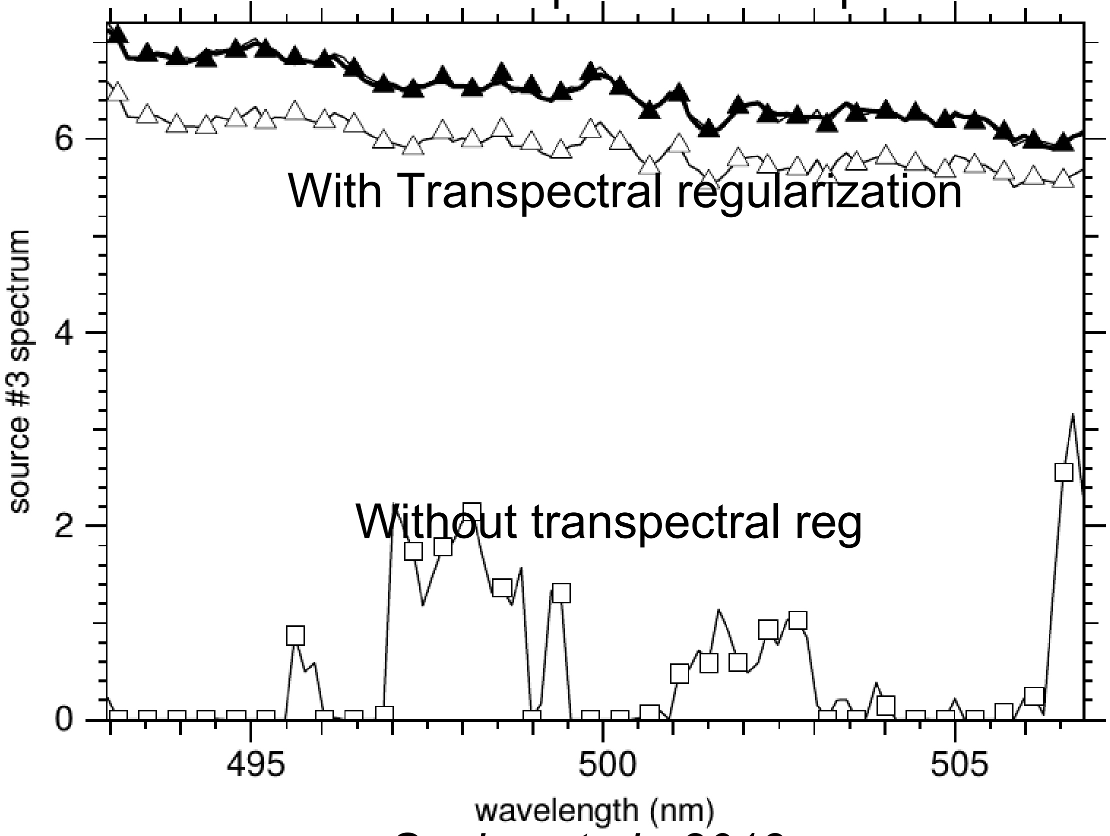
# MIRA 3D

**Transpectral continuity** is imposed in addition to transpectral sparsity

ADMM allows **trivial parallelization** of wavelength reconstructions

**Distributed computing**

Detailed view of the spectra of two point sources




*Soulez et al., 2013*

# Dynamical imaging (EHT)

Johnson et al. 2017

THE ASTROPHYSICAL JOURNAL, 850:172 (15pp), 2017 December 1 https://doi.org/10.3847/1538-4357/aa97dd


CrossMark

## Dynamical Imaging with Interferometry

Michael D. Johnson[1], Katherine L. Bouman[1,2], Lindy Blackburn[1], Andrew A. Chael[1], Julian Rosen[3], Hotaka Shiokawa[1], Freek Roelofs[4], Kazunori Akiyama[5], Vincent L. Fish[5], and Sheperd S. Doeleman[1]
[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA; mjohnson@cfa.harvard.edu
[2] Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge, MA 02139, USA
[3] University of Michigan, Department of Mathematics, 530 Church Street, Ann Arbor, MI 48109-1043, USA
[4] Department of Astrophysics, Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
[5] Massachusetts Institute of Technology, Haystack Observatory, Route 40, Westford, MA 01886, USA
Received 2017 July 30; revised 2017 October 27; accepted 2017 October 31; published 2017 November 30

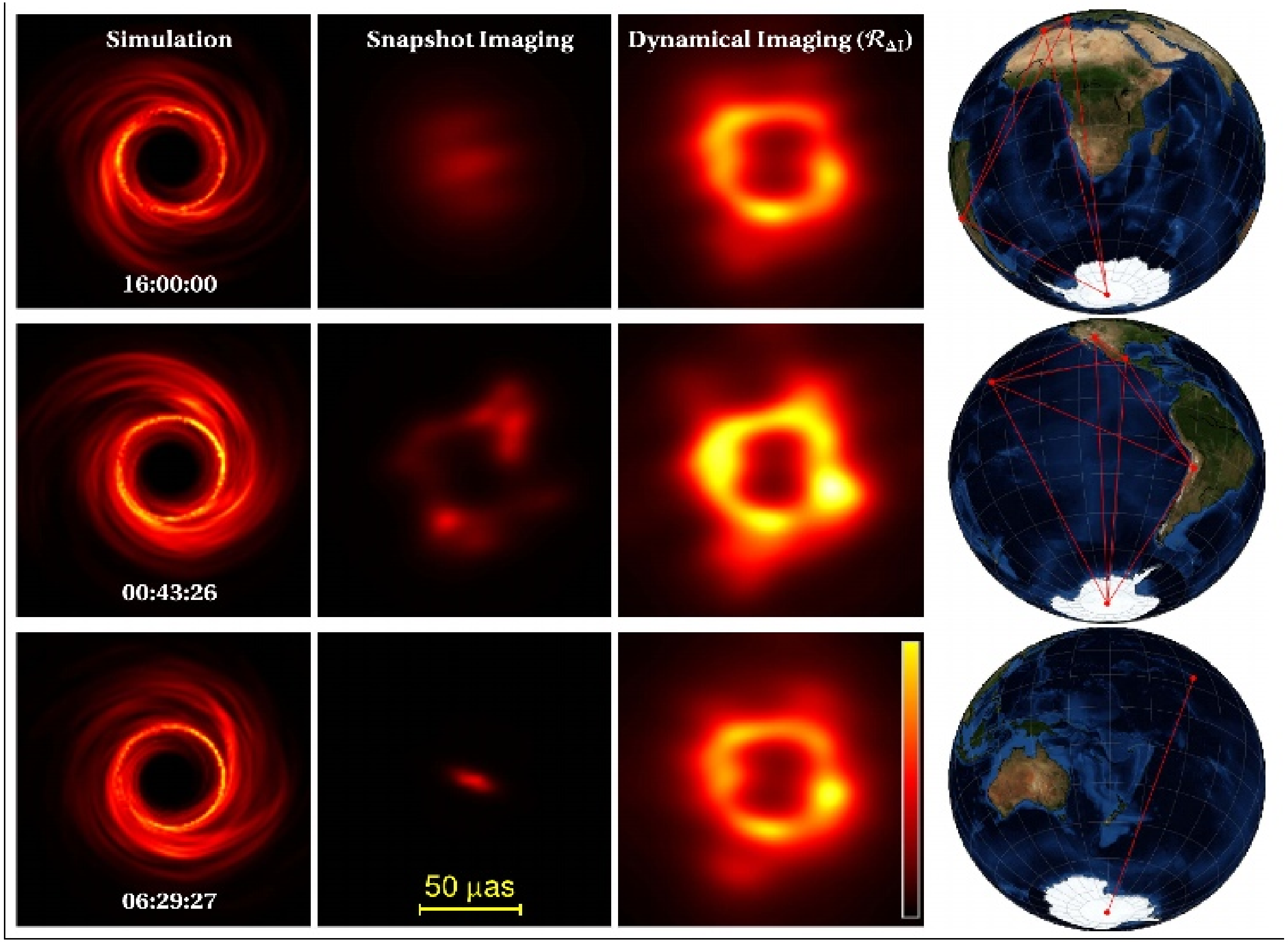
Bouman et al. 2018

Starwarps algorithm

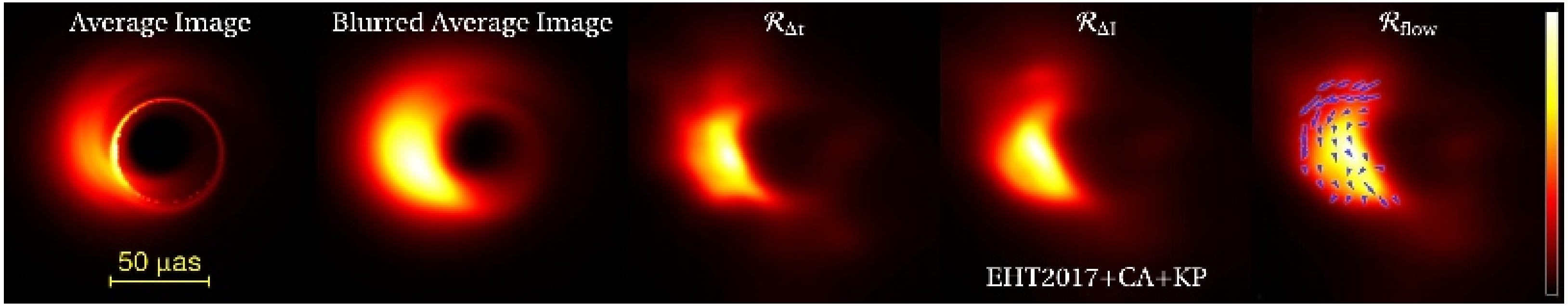
## Reconstructing Video from Interferometric Measurements of Time-Varying Sources

Katherine L. Bouman[1,2], Michael D. Johnson[2], Adrian V. Dalca[1,3], Andrew A. Chael[2], Freek Roelofs[4], Sheperd S. Doeleman[2], William T. Freeman[1,5]
[1]Massachusetts Institute of Technology, CSAIL [2]Harvard-Smithsonian Center for Astrophysics
[3]Massachusetts General Hospital, HMS [4]Radboud University [5]Google Research

# Johnson et al. 2017



# Johnson et al. 2017



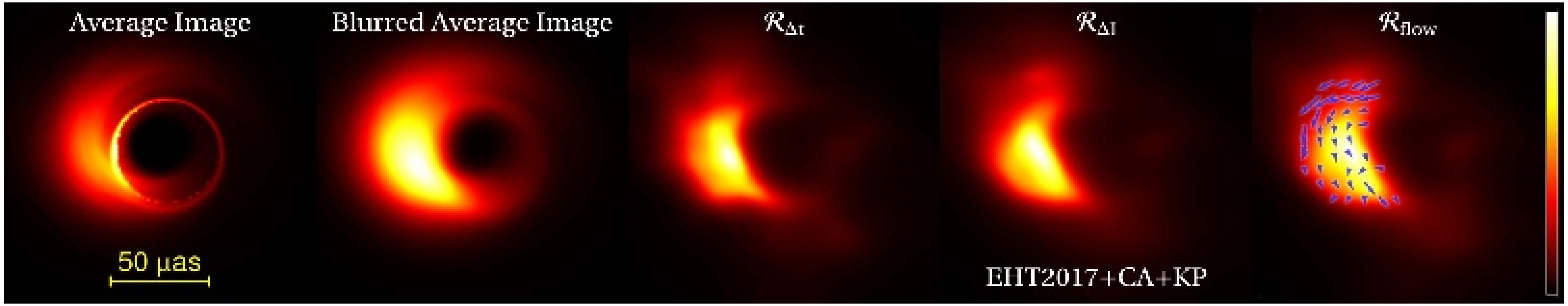

Three dynamical regularization terms

*Smoothly Varying Images over Time* $\mathcal{R}_{\Delta t}$

*A Stable Average Image with Small Perturbations* $\mathcal{R}_{\Delta I}$

*Time-variable Images with Regular Motion* $\mathcal{R}_{\rm flow}$

# Johnson et al. 2017




*Smoothly Varying Images over Time* $\mathcal{R}_{\Delta t}$

the summed difference among all adjacent images after blurring the frames $\boldsymbol{I}_j \longrightarrow B(\boldsymbol{I}_j)$ using a circular Gaussian kernel with standard deviation $\sigma_{\Delta t}$

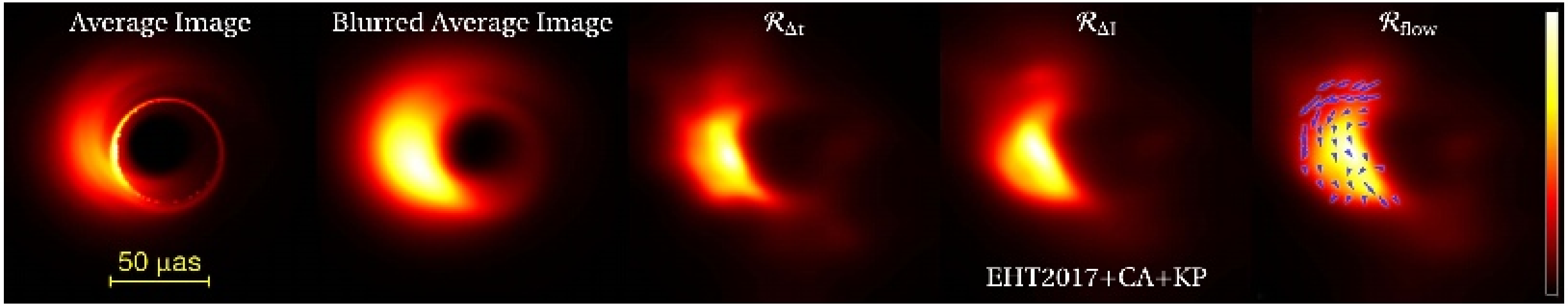
$$\mathcal{R}_{\Delta t}(\{\boldsymbol{I}_k\}) \equiv \sum_{j=1}^{N_t - 1} \mathcal{D}(B(\boldsymbol{I}_j),\, B(\boldsymbol{I}_{j+1})).$$

$$\mathcal{D}_{\rm KL}(\boldsymbol{I},\boldsymbol{I}') = D(\boldsymbol{I}' \| \boldsymbol{I}) \equiv \sum_{m,\ell} I'_{m,\ell} \ln\left(\frac{I'_{m,\ell}}{I_{m,\ell}}\right)$$

$$\mathcal{D}_2(\boldsymbol{I},\boldsymbol{I}') \equiv \|\boldsymbol{I} - \boldsymbol{I}'\|^2 = \sum_{m,\ell} (I_{m,\ell} - I'_{m,\ell})^2.$$

This regularizer penalizes changes between frames via a difference function. There is decreasing penalty for changes on scales smaller than $\sigma_{\Delta t}$
It does not favor stable “momentum” of features between frames.

# Johnson et al. 2017




*A Stable Average Image with Small Perturbations* $\mathcal{R}_{\Delta I}$

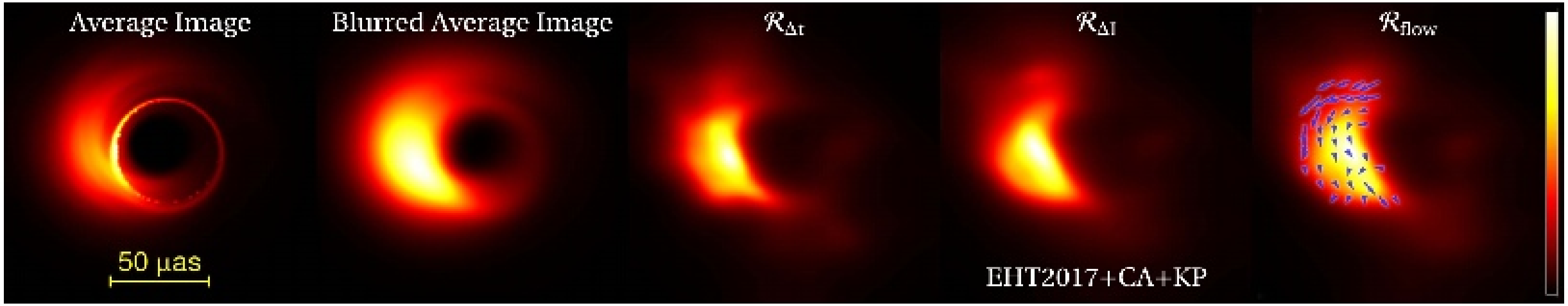
This regularization is adapted to the case where each frame can be described as a small perturbation from the time-averaged image.

Possible application: convection cells?

$$\boldsymbol{I}_{\rm avg} \equiv \frac{1}{N_{\rm t}} \sum_{j=1}^{N_{\rm t}} \boldsymbol{I}_j$$

$$\mathcal{R}_{\Delta I}(\{\boldsymbol{I}_k\}) = \sum_{j=1}^{N_{\rm t}} \mathcal{D}(\boldsymbol{I}_{\rm avg}, \boldsymbol{I}_j).$$

# Johnson et al. 2017




*Time-variable Images with Regular Motion* $\mathcal{R}_{\rm flow}$

$$I(x, y, t + \delta t) \approx I(x, y, t) + \delta t \frac{\partial I(x, y, t)}{\partial t}$$
$$= I(x, y, t) - \delta t \times (\boldsymbol{v} \cdot \nabla I + I \nabla \cdot \boldsymbol{v})$$

This reconstruction strategy must simultaneously estimate the flow vector field $\boldsymbol{m} = \boldsymbol{v}\delta t/\delta x$.

$$\mathcal{R}_{\rm flow}(\{\boldsymbol{I}_k\}, \boldsymbol{m}) = \sum_{j=1}^{N_t - 1} ||\boldsymbol{I}_{j+1} - (\boldsymbol{I}_j - \nabla \cdot [\boldsymbol{I}_j \boldsymbol{m}])||^2$$
$$\approx \sum_{j=1}^{N_t - 1} ||\boldsymbol{I}_{j+1} - \boldsymbol{I}_j + \boldsymbol{m} \cdot \nabla \boldsymbol{I}_j + (\nabla \cdot \boldsymbol{m})\boldsymbol{I}_j||^2$$

Possible application: YSO disks on long time scales, star spots on shorter ones

# Imaging on spheroids

SURFING
Roettenbacher et al., 2016

LETTER

doi:10.1038/nature17444

## No Sun-like dynamo on the active star ζ Andromedae from starspot asymmetry

R. M. Roettenbacher[1], J. D. Monnier[1], H. Korhonen[2,3], A. N. Aarnio[1], F. Baron[1,4], X. Che[1], R. O. Harmon[5], Zs. Kővári[6], S. Kraus[1,7], G. H. Schaefer[8], G. Torres[9], M. Zhao[1,10], T. A. ten Brummelaar[8], J. Sturmann[8] & L. Sturmann[8]

**Sunspots are cool areas caused by strong surface magnetic fields that inhibit convection[1,2]. Moreover, strong magnetic fields can alter the average atmospheric structure[3], degrading our ability to measure stellar masses and ages. Stars that are more active than**

The surface temperature maps for ζ And show peaks of 4,530 K and 4,550 K and minimum values of 3,540 K and 3,660 K in 2011 and 2013, respectively. The ~900-K range of temperatures we see across the surface is slightly larger than the ~700-K range found from recent

ROTIR
Martinez et al., submitted

Draft version November 30, 2020
Typeset using LATEX **twocolumn** style in AASTeX63

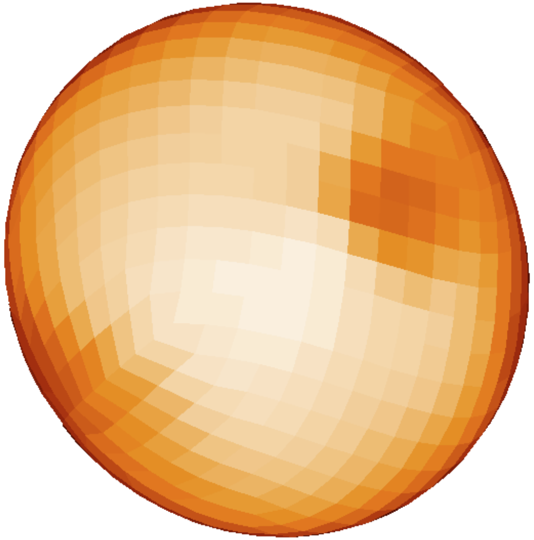
## Dynamical 3D Interferometric Imaging of λ Andromedae

Arturo O. Martinez,[1,2] Fabien R. Baron,[1,2] John D. Monnier,[3] Rachael M. Roettenbacher,[4] and J. Robert Parks[5]

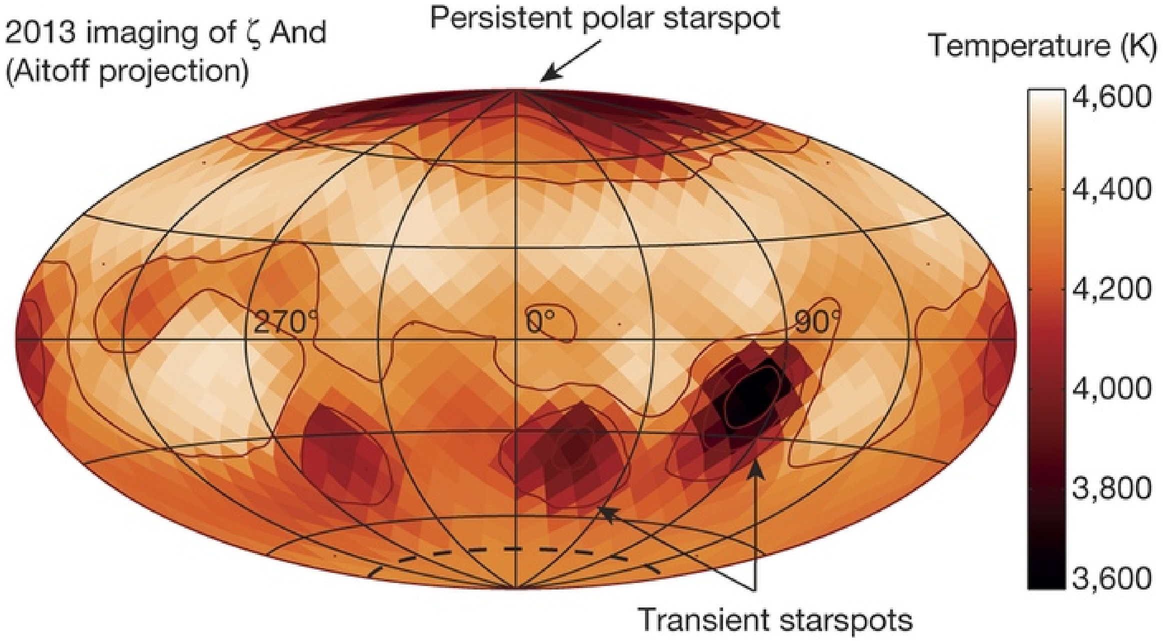
[1]Center for High Angular Resolution Astronomy, 25 Park Place NE #605, Atlanta, GA 30303-2911, USA
[2]Department of Physics and Astronomy, Georgia State University, 25 Park Place NE #605, Atlanta, GA 30303-2911, USA
[3]Department of Astronomy, University of Michigan, Ann Arbor, MI 48109-1090, USA
[4]Yale Center for Astronomy and Astrophysics, Department of Physics, Yale University, New Haven, CT 06520-8120, USA
[5]Department of Physics and Astronomy, George Mason University, 4400 University Drive Fairfax, VA 22030-4444, USA

# SURFING

Affine-invariant code developed by John Monnier




Credit: R. Roettenbacher (U. Michigan)



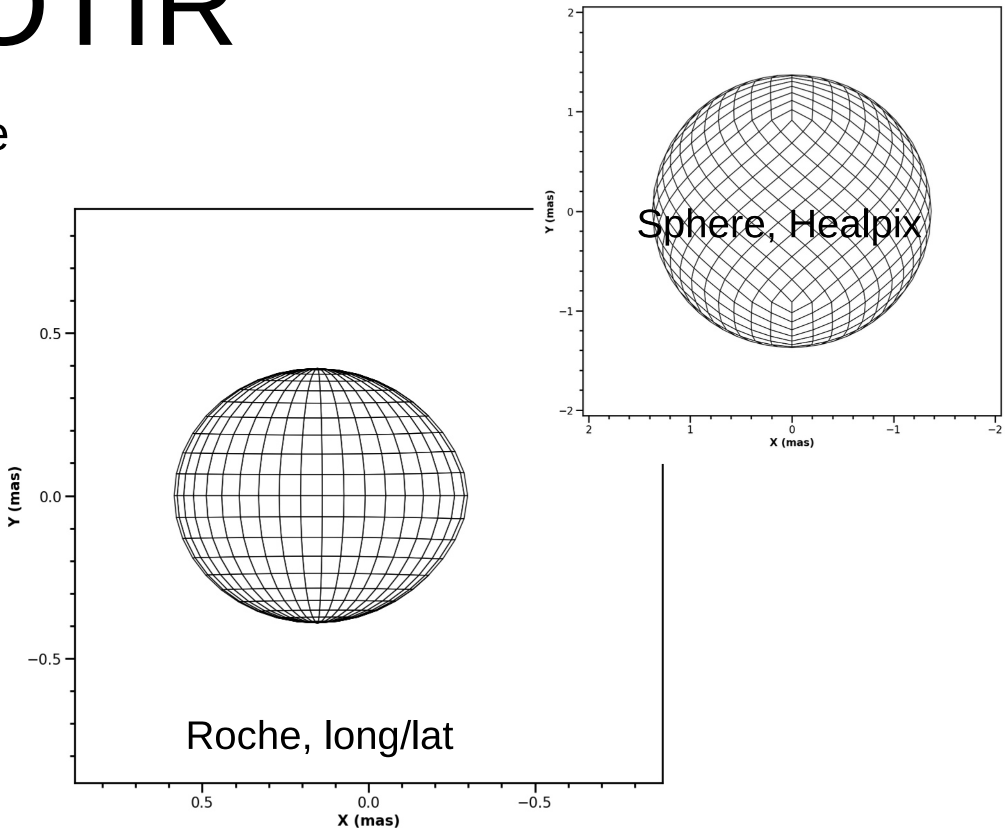

SURFING reconstruction of zet And
Roettenbacher et al., *Nature*, 2016

# ROTIR

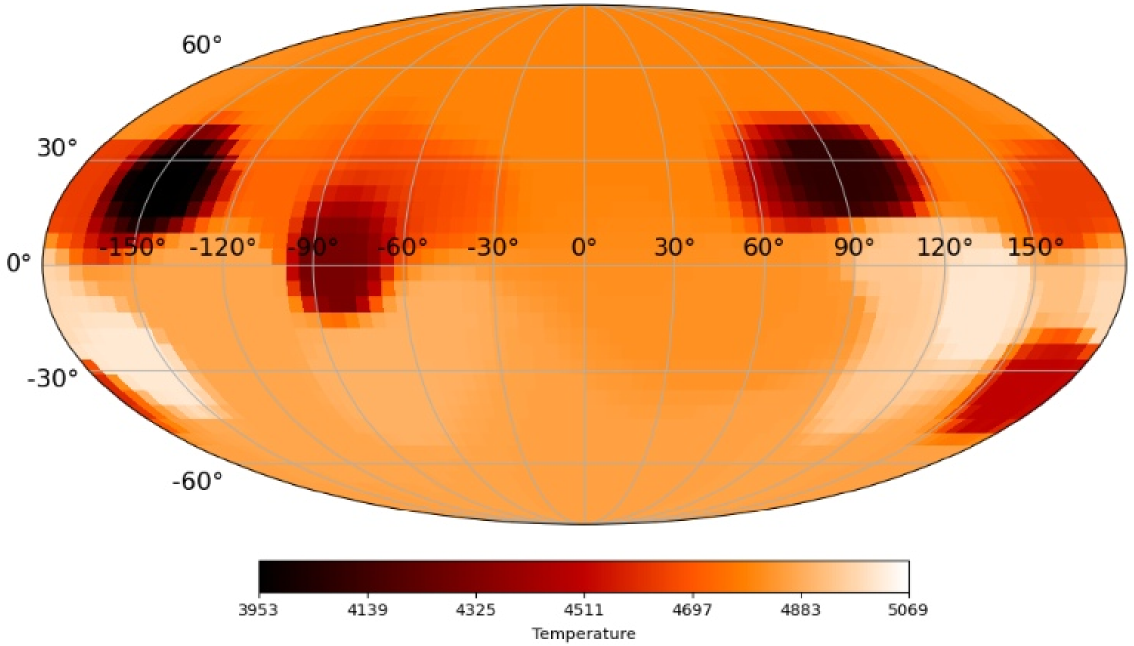
ROTIR can work on any tessellated surface defined by a vertex vector **v**, including binaries, Roche lobes, ...

The DFT matrix is replaced by:

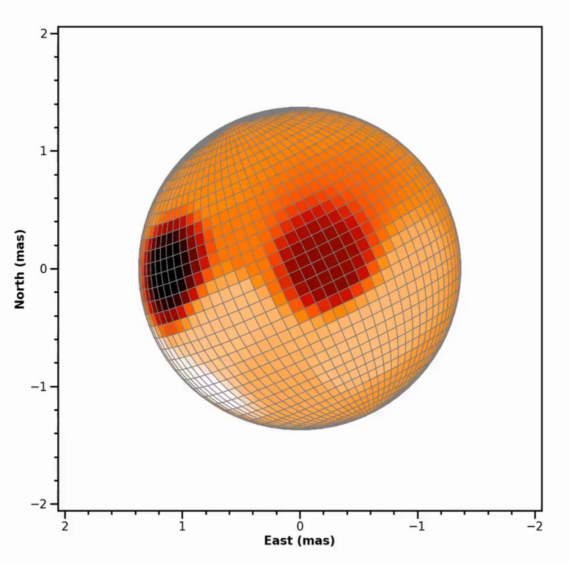
$$\widetilde{\mathbf{S}}_n = \sum_{j=1}^{m} \hat{\mathbf{z}} \cdot [(\mathbf{v}_{j+1} - \mathbf{v}_j) \times \mathbf{k}]$$
$$\exp[-i\pi\mathbf{k} \cdot (\mathbf{v}_{j+1} + \mathbf{v}_j)]$$
$$\mathrm{sinc}[\mathbf{k} \cdot (\mathbf{v}_{j+1} + \mathbf{v}_j)] \frac{L_n}{i2\pi|\mathbf{k}|^2}$$



# ROTIR




Map of lambda And in 2011
Martinez et al., 2020 (submitted)




ROTIR also does dynamical imaging, light curve inversion, and Doppler imaging is in the works.

# Machine learning: work in progress

Bouman et al., 2017

CHIRP patch prior algorithm

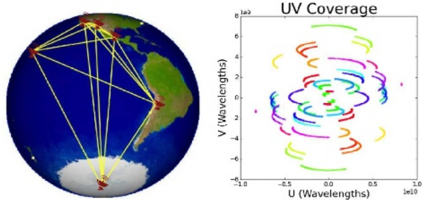
**Computational Imaging for VLBI Image Reconstruction**

Katherine L. Bouman[1] Michael D. Johnson[2] Daniel Zoran[1] Vincent L. Fish[3]
Sheperd S. Doeleman[2,3] William T. Freeman[1,4]

[1]Massachusetts Institute of Technology, CSAIL [2]Harvard, Center for Astrophysics [3]MIT Haystack Observatory [4]Google

**Abstract**

*Very long baseline interferometry (VLBI) is a technique for imaging celestial radio emissions by simultaneously observing a source from telescopes distributed across Earth. The challenges in reconstructing images from fine angular resolution VLBI data are immense. The data is extremely sparse and noisy, thus requiring statistical image models such as those designed in the computer vision community. In this paper we present a novel Bayesian approach for*




Claes et al., 2020

Generative Adversarial Networks

Paper 11446-110, *Neural network based image reconstruction with astrophysical priors, SPIE 2020.*

# Conclusion: the future of image reconstruction algorithms

- Global optimization, polychromatic and dynamic imaging, simultaneous image reconstruction and model-fitting
- Better error maps
- Use of machine learning for regularization and image comparison
- GPU acceleration (may be provided by the language, or hand-optimized)

# Read these!

*Principles of image reconstruction in optical interferometry: tutorial, Thiébaut, É and Young, J.S., Journal of the Optical Society of America A, vol. 34, issue 6, p. 904 (2017),* *https://ui.adsabs.harvard.edu/abs/2017JOSAA..34..904T.*

*Image reconstruction in optical interferometry, Thiébaut É. and Giovannelli J.-F. , IEEE Signal Process. Mag. 27(1), 97–109 (2010),* *https://ui.adsabs.harvard.edu/abs/2010ISPM...27...97T.*

*Image reconstruction in optical interferometry: Benchmarking the regularization*, Renard, S. ; Thiébaut, *É*. ; Malbet, F., vol. 533, p. A64 (2011), https://ui.adsabs.harvard.edu/abs/2011A%26A...533A..64R.