Imaging EHT data with the sparse modeling

Kazunori Akiyama
(NAOJ —> MIT Haystack)

Collaborators
Fumie Tazaki, Kazuhiro Hada, Mareki Honma (NAOJ),
Shiro Ikeda (Institute of Statistical Mathematics),
Makoto Uemura (Hiroshima Univ.)

Outline of today's talk

- Our motivation to work on the imaging technique
- Application of the sparse modeling to the interferometric imaging
    - The basic idea
      (see §2 in Honma, ***KA,*** Uemura & Ikeda 2014, PASJ)
    - Mathematical description
      (see §3 in Honma, ***KA,*** Uemura & Ikeda 2014, PASJ)
    - Results of the sparse modeling on observational/simulated data
- Practical issue: Imaging pipeline with the sparse modeling

Angular diameter of super massive black holes

Source	BH Mass (M_{solar})	Distance (Mpc)	Angular radius of R_s (μas)
Sgr A* (Galactic Center)	4×10^6	0.008	10
M87 (Virgo A)	$3 - 6 \times 10^9$	17.8	3 - 7
M104 (Sombrero Galaxy)	1×10^9	10	2
Cen A	5×10^7	4	0.25

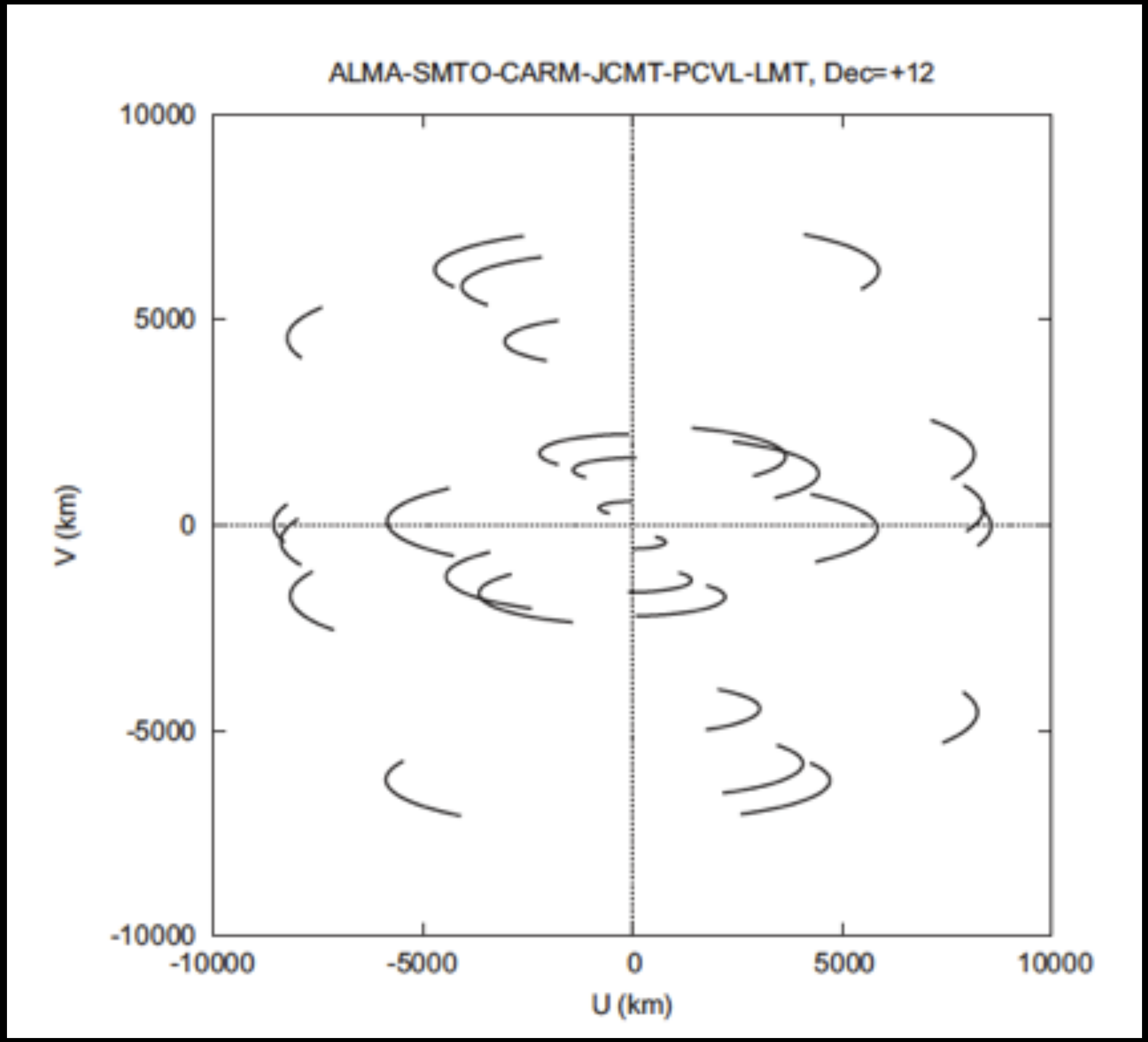
M 87 (NGC 4486) Ultra-high-sensitivity HDTV I.I. color camera (NHK)
Exp. 40 sec. (10 frames coadded) January 16, 1999
Subaru Telescope, National Astronomical Observatory of Japan

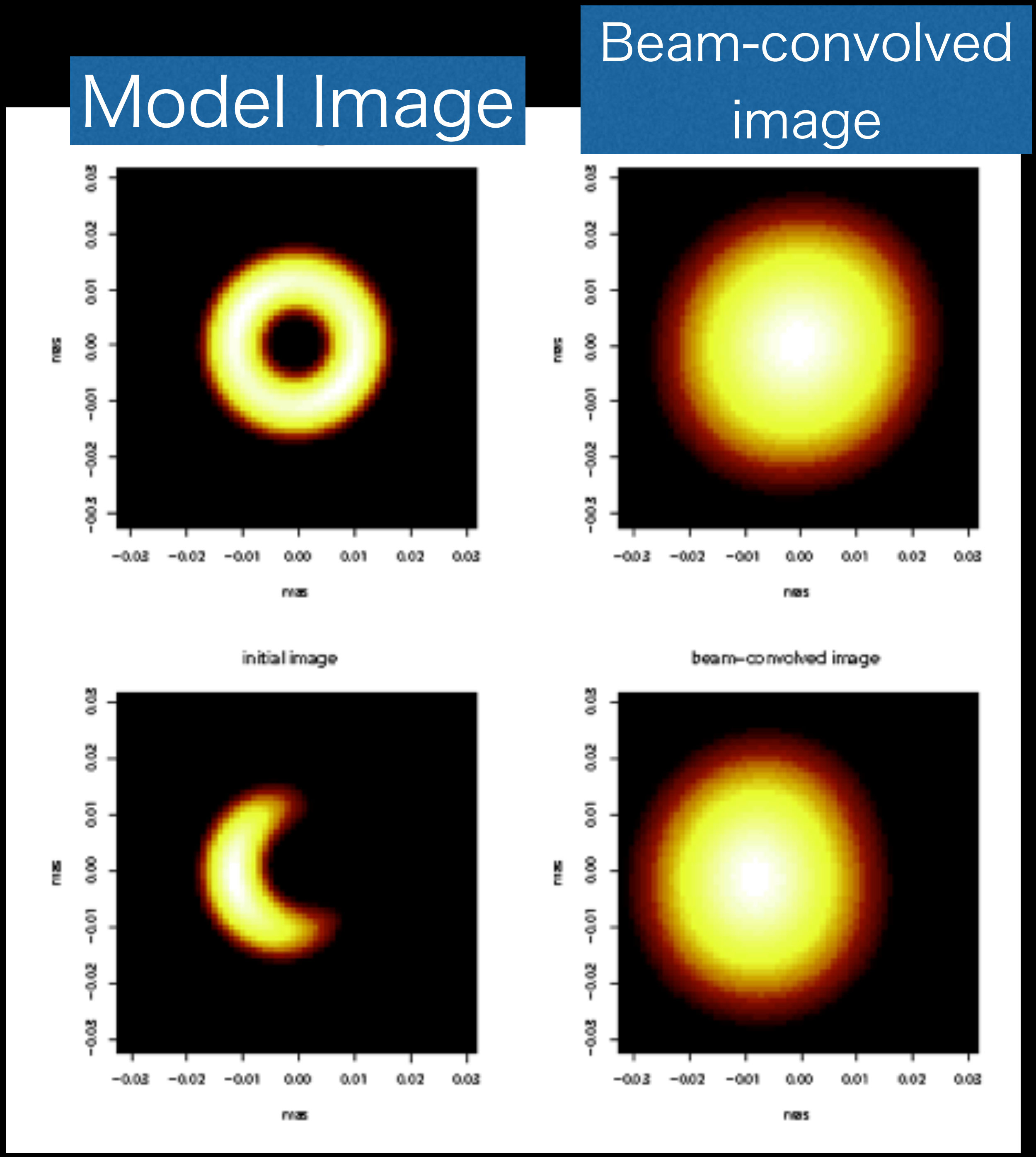
Photon sphere: (few - 3) x R_s (3 R_s for non-spinning BH)
ISCO size: (several - 10) x R_s

Event Horizon Telescope after 2015

Maximum Baseline length: ~10,000 km

Synthesized beam size:
1.3 mm/10,000 km ~ 27 μas





(Honma, ***KA***, Uemura & Ikeda 2014, PASJ)

Our Motivation

- Even with the full EHT, the size of the synthesized beam (~20 μas) will be comparable to expected shadow sizes for Sgr A* and M87
  - might be not enough? (particularly in low-mass case of M87)
  - require a shaper restoring beam for CLEAN (= super resolution)
- Is there a technique to enable robust high resolution imaging for ensuring a feasibility of EHT to take a picture of BH shadow
  - particularly, resolution higher than λ/D (= super resolution) equivalent to build up larger arrays.

Outline of today's talk

- Our motivation to work on the imaging technique
- Application of the sparse modeling to the interferometric imaging
  - the basic idea
    (see §2 in Honma, ***KA,*** Uemura & Ikea 2014, PASJ)
  - the mathematical description
    (see §3 in Honma, ***KA,*** Uemura & Ikea 2014, PASJ)
  - results of the sparse modeling on observational/simulated data
- Practical issue: Imaging pipeline with the sparse modeling

Imaging with the interferometer (I)

- **Basic Equation:** 2D Fourier Transform between the image and visibility

$$I_\nu(x, y) = \int\int dudv\, V_\nu(u, v) e^{-2\pi(ux+vy)}$$

- **Spatial frequency (u, v):** baseline vectors seen from the target source

- **What does interferometer observe?:**
  Fourier components at various baseline lengths (i.e. spatial frequencies)

- **How to Image:**
  In actual case, discrete Fourier transform of sampled visibility
  is performed to obtain images

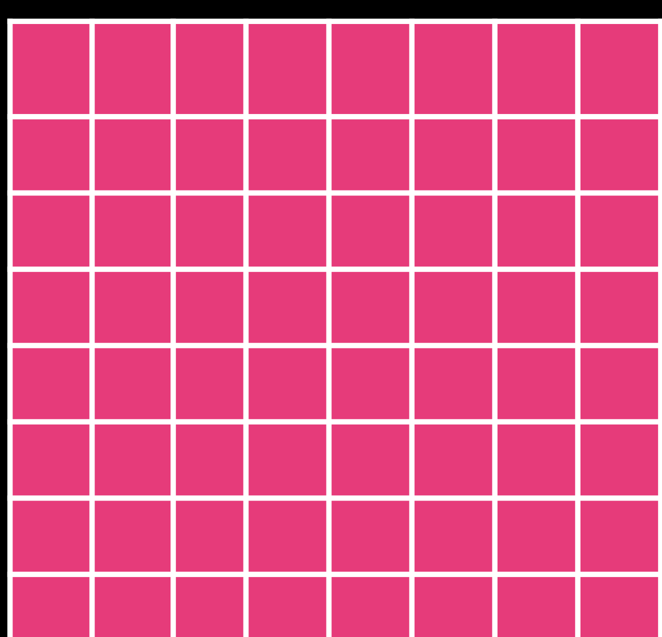
Imaging with the interferometer (II)

- In actual case：Imperfect sampling of Fourier components
  - 0-padding is used to obtain an image assuming visibilities of zero for unsampled Fourier components

- This cause finite resolutions and side lobes
  resolution: Θ ～ λ / B （λ：wavelength、B：baseline length）

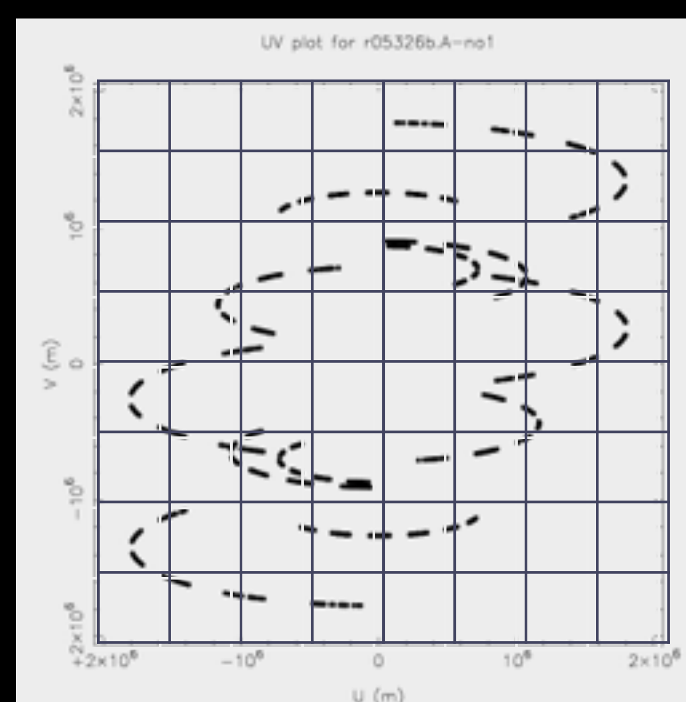
(e.g.) Point Source

visibility: **Uniform**



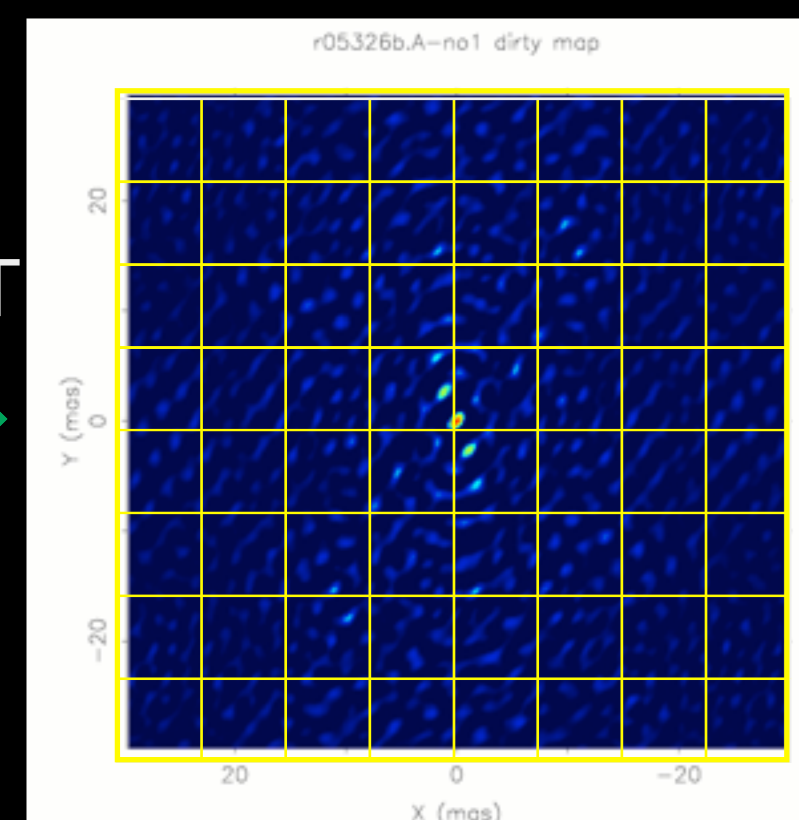
X

Sampled spatial frequencies
= ***uv* coverage**



2D DFT

Observed image:
Synthesized beam
= DFT of uv coverage



Imaging with the interferometer (II)

- In actual case：Imperfect sampling of Fourier components
  - 0-padding is used to obtain an image
    assuming visibilities of zero for unsampled Fourier components

- This cause finite resolutions and side lobes
  resolution: $\Theta \sim \lambda / B$ （λ：wavelength、B：baseline length）

(e.g.) Black hole shadow

convolution (Dirty image)

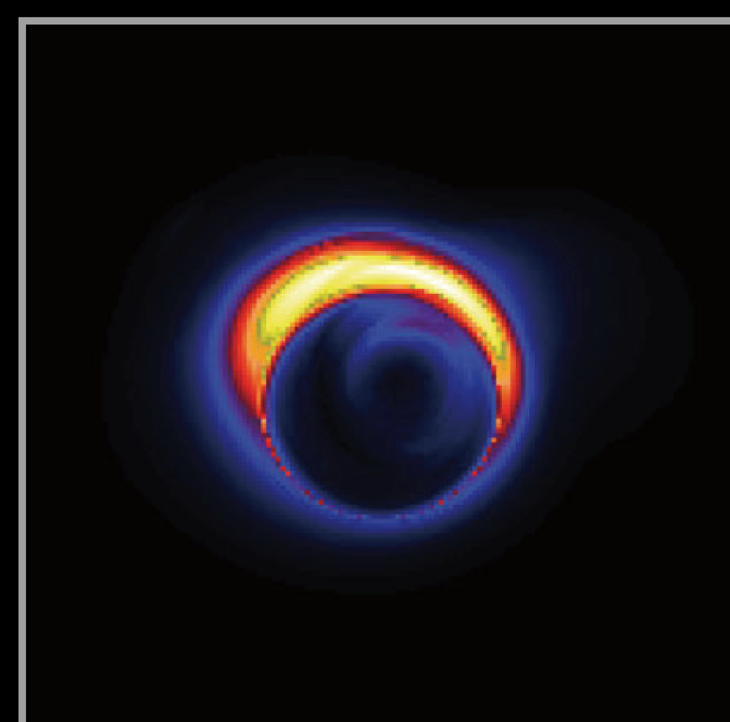
visibility on *uv* coverage

v

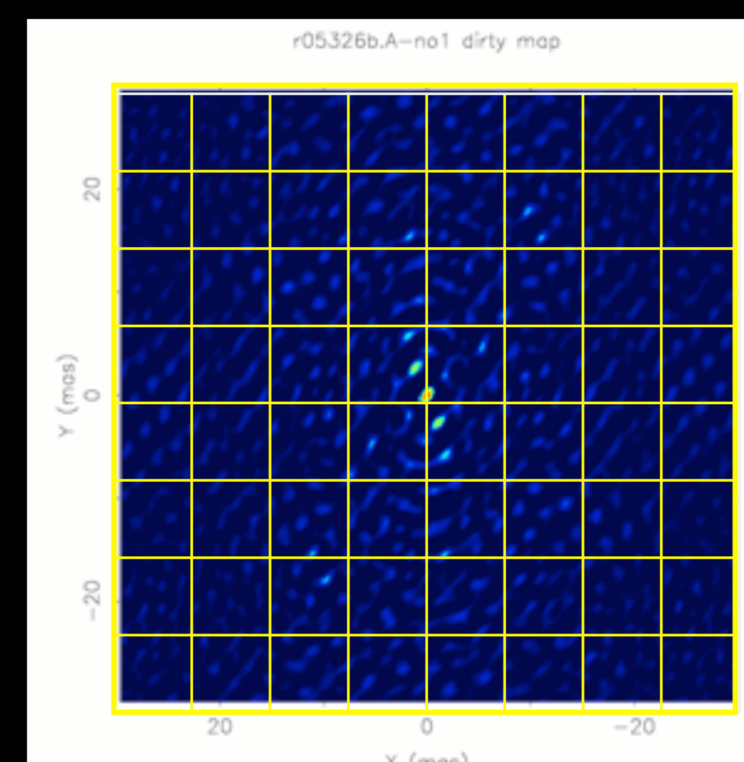
u

2D DFT

Real image



Synthesized beam
(DFT of *uv* coverage)



Imaging with the interferometer (II)

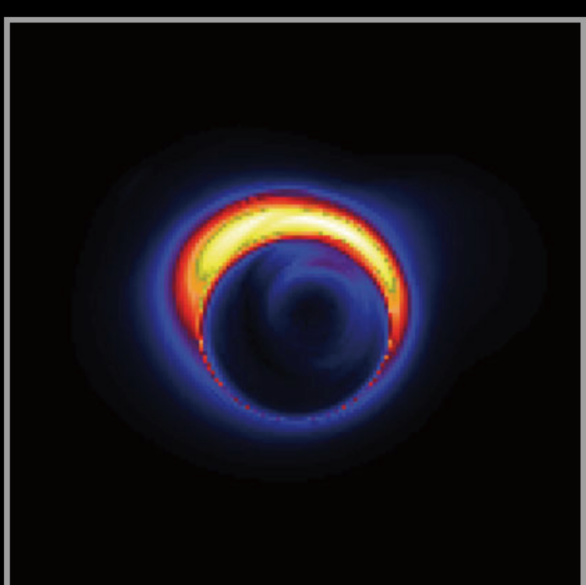
- In actual case：Imperfect sampling of Fourier components
  - 0-padding is used to obtain an image assuming visibilities of zero for unsampled Fourier components

- This cause finite resolutions and side lobes
  resolution: Θ ～ λ / B （λ：wavelength、B：baseline length）

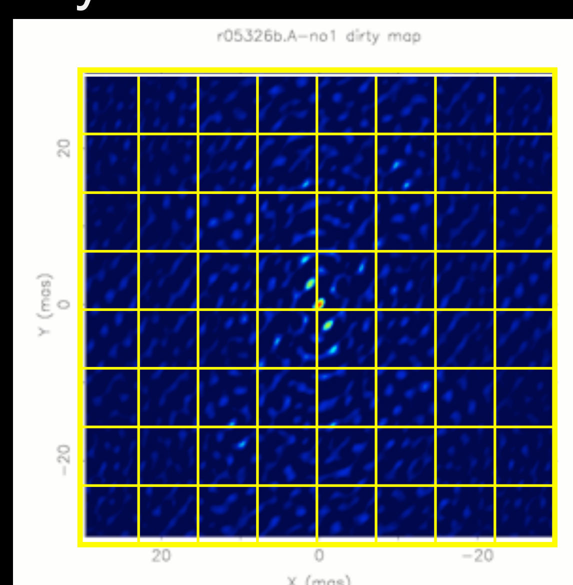
Traditional method (CLEAN = Matching Pursuit in Statistical Mathematics)

convolution (Dirty image)

Real

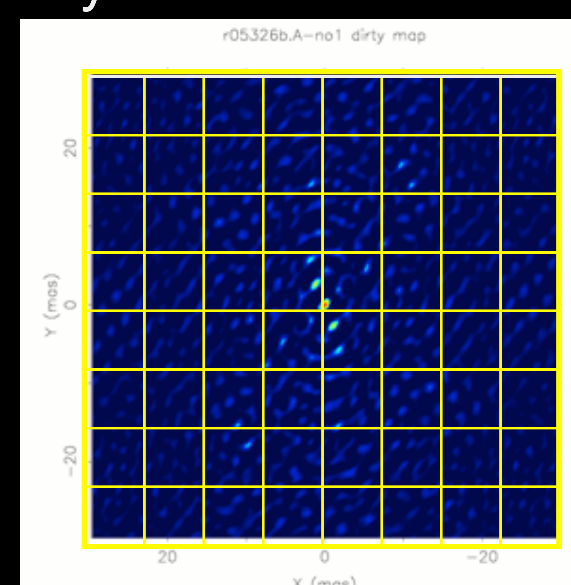


Synthesized beam



-

Synthesized beam



=>

Reconstructed Image:
consisting of
a minimum number
of point sources

reconstructing sparse images on the image plane

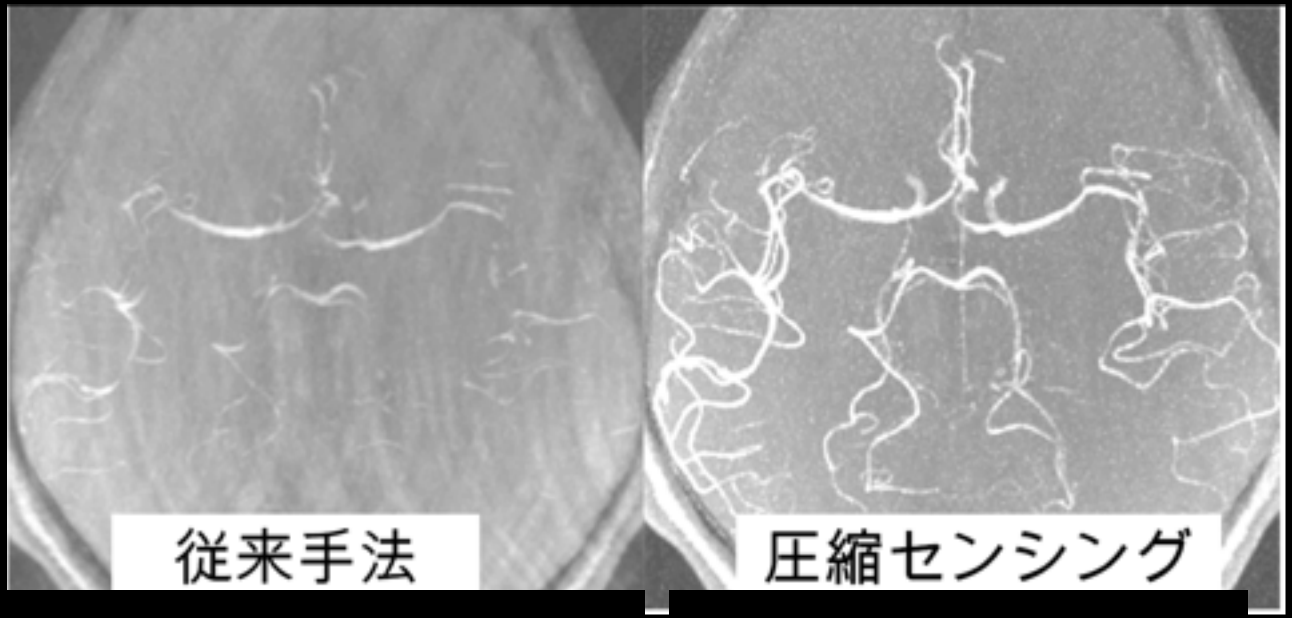
Sparse Modeling

Ill-posed problems

- Linear equations can be solved if number of equations *M* is larger than number of parameters *N* (i.e., requires $M > N$)
- Otherwise($M<N$), it becomes an ill-posed problem (can not be solved)

Idea of the sparse modeling to solve ill-posed problems

- If number of effective parameters (non-0 parameters) N' is smaller than M, equations can be solved (sparse solution)
- Mathematical background: (Donoho, Candes & Tao 2006; Compressive sensing)
- Compressing Sensing is now one of standard techniques for MRI (e.g. Lustig et al. 2008)



without Sparse Modeling with Sparse Modeling

MRI image of the cerebral blood vessel

Sparse modeling and interferometric imaging

- Observation Equation（2D DFT）can be written in a linear equation.

$$\mathbf{V} = \mathbf{A}\mathbf{I}$$

$$\begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_M \end{pmatrix} = \begin{pmatrix} e^{-2\pi i(u_1 x_1 + v_1 y_1)} & e^{-2\pi i(u_1 x_2 + v_1 y_2)} & \dots & e^{-2\pi i(u_1 x_{N^2} + v_1 y_{N^2})} \\ e^{-2\pi i(u_2 x_1 + v_2 y_1)} & e^{-2\pi i(u_2 x_2 + v_2 y_2)} & \dots & e^{-2\pi i(u_2 x_{N^2} + v_2 y_{N^2})} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-2\pi i(u_M x_1 + v_M y_1)} & e^{-2\pi i(u_M x_2 + v_M y_2)} & \dots & e^{-2\pi i(u_M x_{N^2} + v_N y_{N^2})} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_{N^2} \end{pmatrix}$$

- Observation Matrix A: dimension of M x N^2
  M: Number of visibility, N: Number of image grids

- ill-posed problem: In normal case, $M < N^2$: requiring 0 padding

- For the case of target sources of EHT, we can expect “sparse images”
  (the emission structure would be very compact compared with F.O.V)
  → We can apply the sparse modeling

Idea of imaging with CLEAN (Matching Pursuit)

$$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ \cdots \\ \cdots \\ \cdots \\ I_N \end{pmatrix} = \mathbf{A}^{-1} \times \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ \cdots \\ V_M \\ 0 \\ 0 \end{pmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \\ \end{matrix}\right\} \text{Actual Data} \\ \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{0 padding} \end{matrix}$$

- do 0 padding to equal numbers of data and image grids
- Try to find a sparse solution on the image plane

Idea of imaging with the sparse modeling

Actual Data

$$\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ \cdots \\ V_M \end{pmatrix} = \mathbf{A} \times \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ \cdots \\ \cdots \\ \cdots \\ I_N \end{pmatrix}$$

grids with brightness of zero

- ill-posed equations can be solved by focusing on “sparseness” of solutions.
- Try to find a sparse solution in the visibility plane
- Reconstructed image not affected by 0-padding
  → possibly we can get super-resolved image

Outline of today's talk

- Our motivation to work on the imaging technique
- Application of the sparse modeling to the interferometric imaging
  - The basic idea
    (see §2 in Honma, ***KA,*** Uemura & Ikeda 2014, PASJ)
  - Mathematical description
    (see §3 in Honma, ***KA,*** Uemura & Ikeda 2014, PASJ)
  - Results of the sparse modeling on observational/simulated data
- Practical issue: Imaging pipeline with the sparse modeling

Mathematical description of the sparse modeling (I)

- Problem without noise

Equation: $\mathbf{V} = A\mathbf{I}$

What to solve?

$$\min_{I} ||\mathbf{I}||_0 \text{ subject to } \mathbf{V} = A\mathbf{I}$$

0-dimensional norm:

$$||I||_0 = |\{I, Ii \neq 0 \text{ for } i = 1, 2, .., N\}|$$

= Number of non-zero parameters

solver: combinatorial optimization (CO)

- practically difficult to be solved for large N (say N~100)

Mathematical description of the sparse modeling (II)

- Problem without noise (Compressive Sensing)

Equation: $\mathbf{V} = A\mathbf{I}$

What to solve?

n-dimensional norm:

$$||I||_p = \left(\sum_i |I_i|^p\right)^{1/p}$$

$$\min_I ||\mathbf{I}||_1 \text{ subject to } \mathbf{V} = A\mathbf{I}$$

1-dimensional norm:

$$||I||_1 = \sum_i |I_i|$$

Equivalent for the sparse solution (Donoho et al. 2006)

solver: linear programming (LP)

- can be solved even for large N (say N>10,000)

$$\min_I ||\mathbf{I}||_0 \text{ subject to } \mathbf{V} = A\mathbf{I}$$

Mathematical description of the sparse modeling (III)

- Problem with noise (LASSO; Least Absolute Shrinkage and Selection Operator)

Equation: $\mathbf{V} = A\mathbf{I} + \mathbf{N}$

What to solve?

χ^2 term

$$\min_I ||\mathbf{V} - A\mathbf{I}||_2^2 \text{ subject to } ||\mathbf{I}||_1 \leq S$$

(Tibshirani 1996)

Equivalent

S determines the number of non-zero parameters
- Large S: $||\mathbf{I}||_0 = N$
- Small S: $||\mathbf{I}||_0 = 1$

$$\min_I \left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda ||\mathbf{I}||_1 \right)$$

(Tibshirani 1999)

solver: quadratic programming (QP)

- can be solved even for large N (say N>10,000)

LASSO and Bayesian Statistics

LASSO

$$\mathbf{I} = \mathrm{argmin}\left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda||\mathbf{I}||_1\right)$$

Chi-square + Regularization

Multiplying by -1/2 and then taking exponential

$$\mathbf{I} = \mathrm{argmax}\left(\exp(-||\mathbf{V} - A\mathbf{I}||_2^2/2)\exp(-\Lambda||\mathbf{I}||_1/2)\right)$$

Likelihood P(**V**|**I**) x Prior Prob. P(**I**)

∝ Posterior Prob. P(**I**|**V**)

Outline of today's talk

- Our motivation to work on the imaging technique
- Application of the sparse modeling to the interferometric imaging
    - The basic idea
      (see §2 in Honma, ***KA,*** Uemura & Ikeda 2014, PASJ)
    - The mathematical description
      (see §3 in Honma, ***KA,*** Uemura & Ikeda 2014, PASJ)
    - Results of the sparse modeling on observational/simulated data
- Practical issue: Imaging pipeline with the sparse modeling

Simulation with Sparse Modeling (I)

- Data set:
  Make visibilities from model images on actual uv-coverages

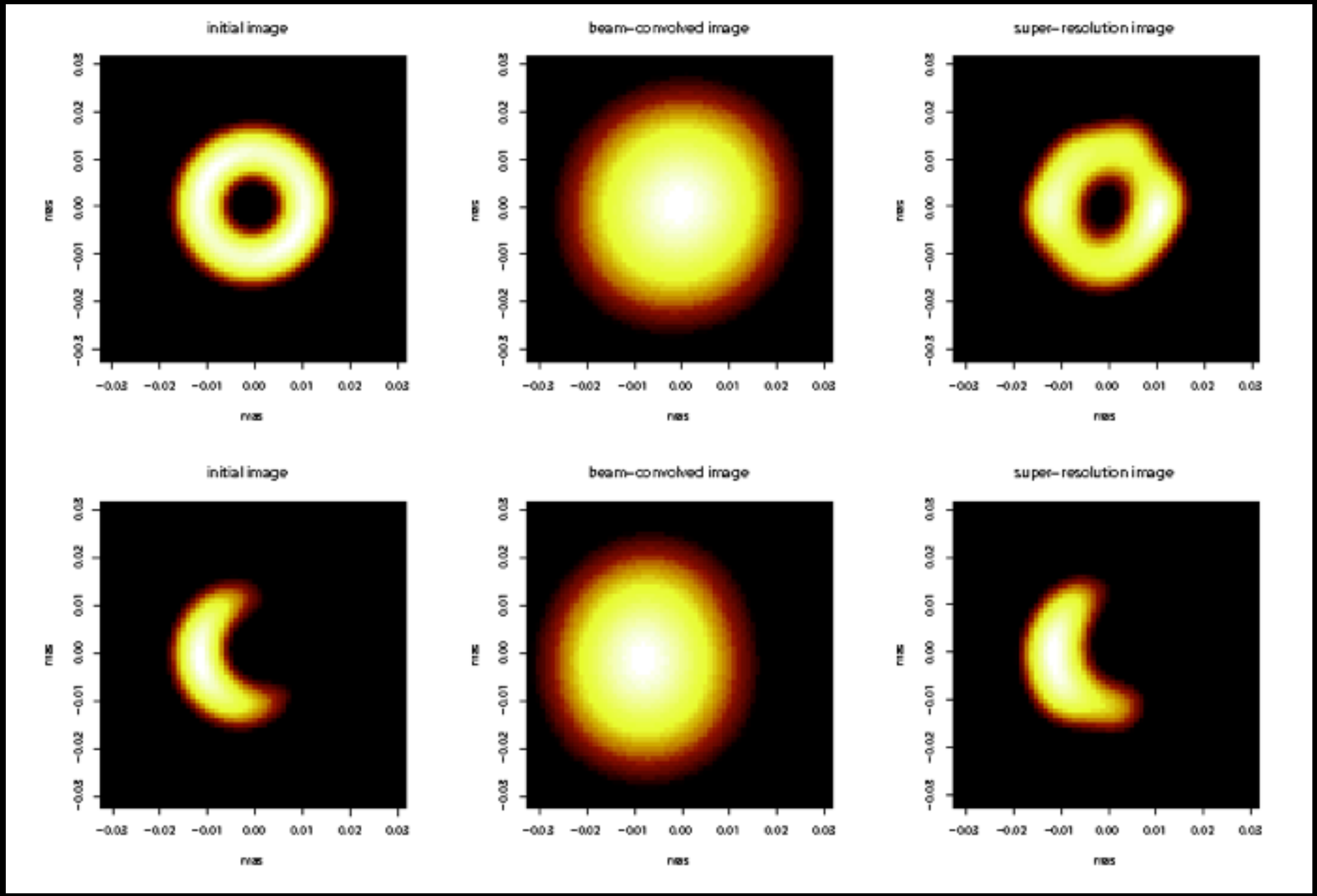
- Noise treatment: thermal, homogeneous
  thermal noise at a 5 %-level of the total flux
    (- SNR ~ 20 for the intra-site baselines, but no intra-site baselines)
    - much lower SNR for VLBI baselines
    - Similar to or worth than current observations

- Modeling method: LASSO + additional regularization term
  Solver : QP solver (original; MATLAB-based)

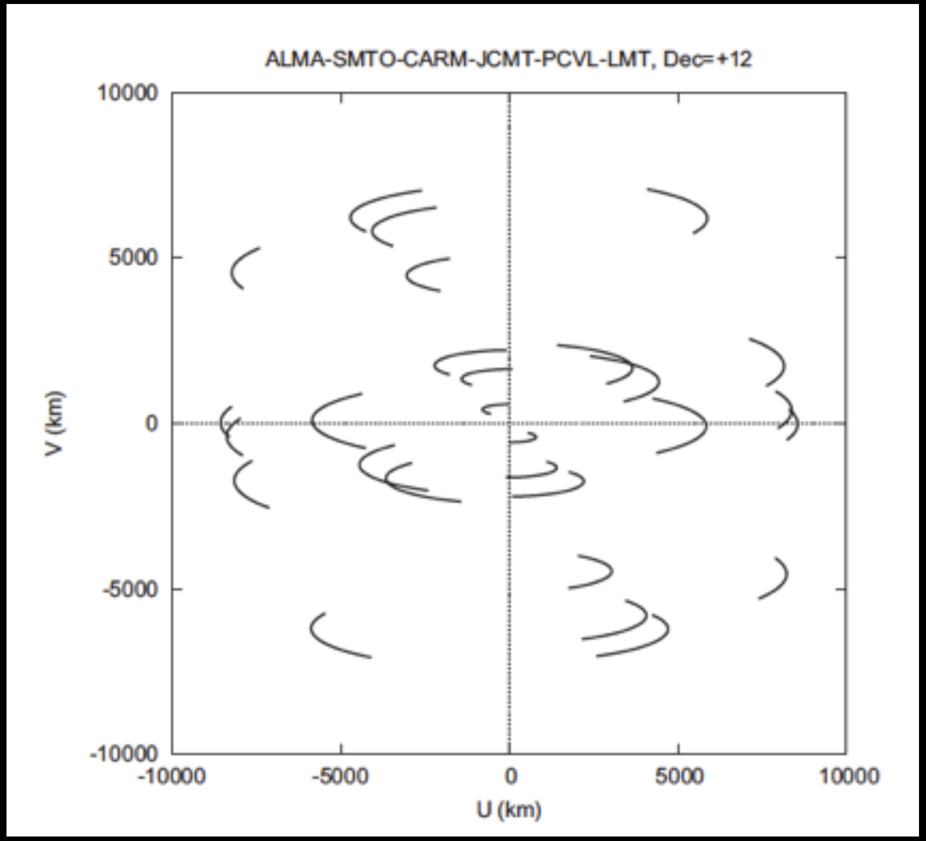
$$\mathbf{I} = \mathrm{argmin}\left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda||\mathbf{I}||_1\right) \quad \text{subject to } I_i \geq 0$$

Simulation with Sparse Modeling (I)

model beam-convolved image



uv-coverage

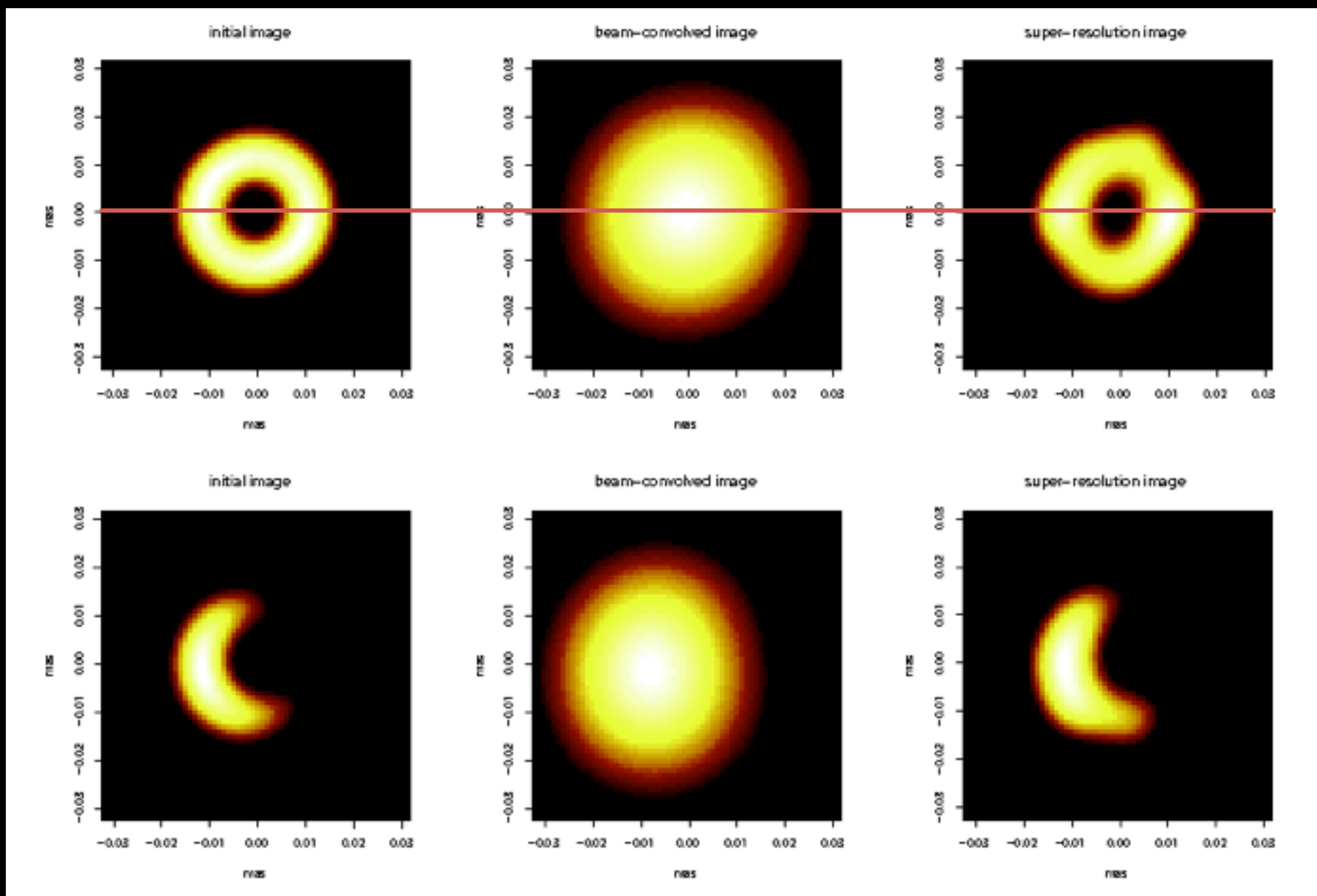


M87, Shadow Diameter ~ 20 uas (for the case of $M_{BH} = 3 \times 10^9\ M_{solar}$)

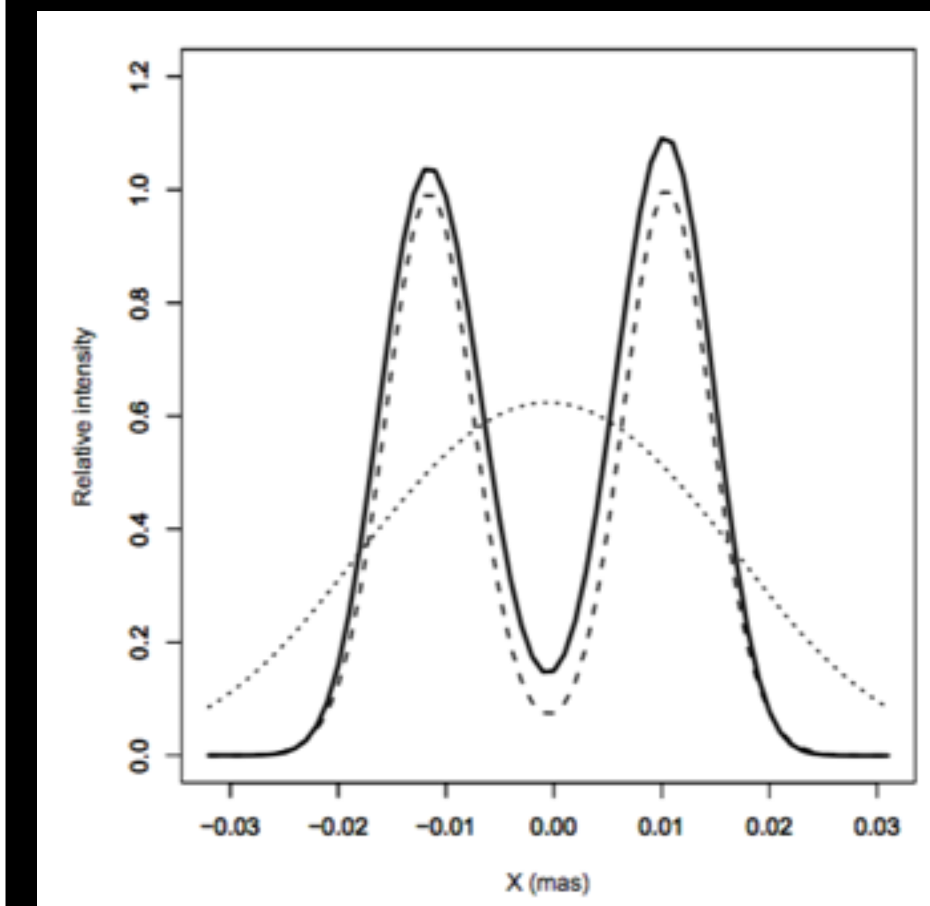
(Honma, ***KA***, Uemura & Ikeda 2014, PASJ)

Simulation with Sparse Modeling (I)

model beam-convolved image



Sliced image

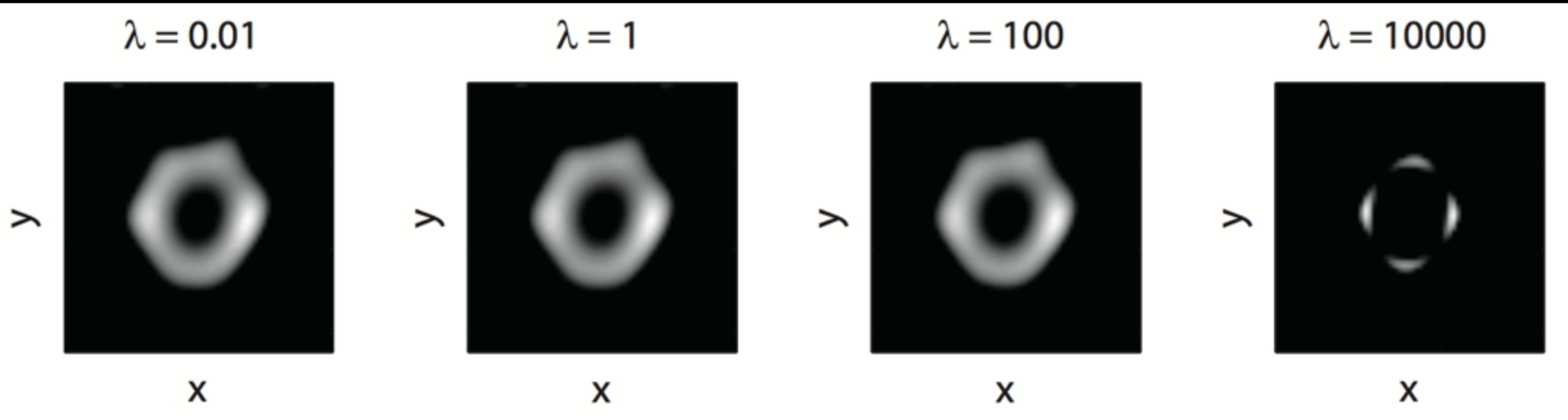


M87, Shadow Diameter ~ 20 uas (for the case of $M_{BH} = 3 \times 10^9\ M_{solar}$)

(Honma, ***KA***, Uemura & Ikeda 2014, PASJ)

Simulation with Sparse Modeling (I)

Choice of the Lambda



$$\mathbf{I} = \mathrm{argmin}\left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda||\mathbf{I}||_1\right)$$

M87, Shadow Diameter ~ 20 uas (for the case of M_{BH}=3x10^9 M_{solar})

(Honma, ***KA***, Uemura & Ikeda 2014, PASJ)

Simulation with Sparse Modeling (II)

- Data set: Simulated data on physically motivated models for M87
  MAPS Simulated Data-sets with parameters same to Lu et al. 2014, ApJ
  Models in Akiyama, Lu & Fish et al. 2014, ApJ, in press.
  approaching-jet-dominated type (Broderick+)
  counter-jet-dominated type (Dexter+)
  accretion-disk-dominated type (Dexter+)

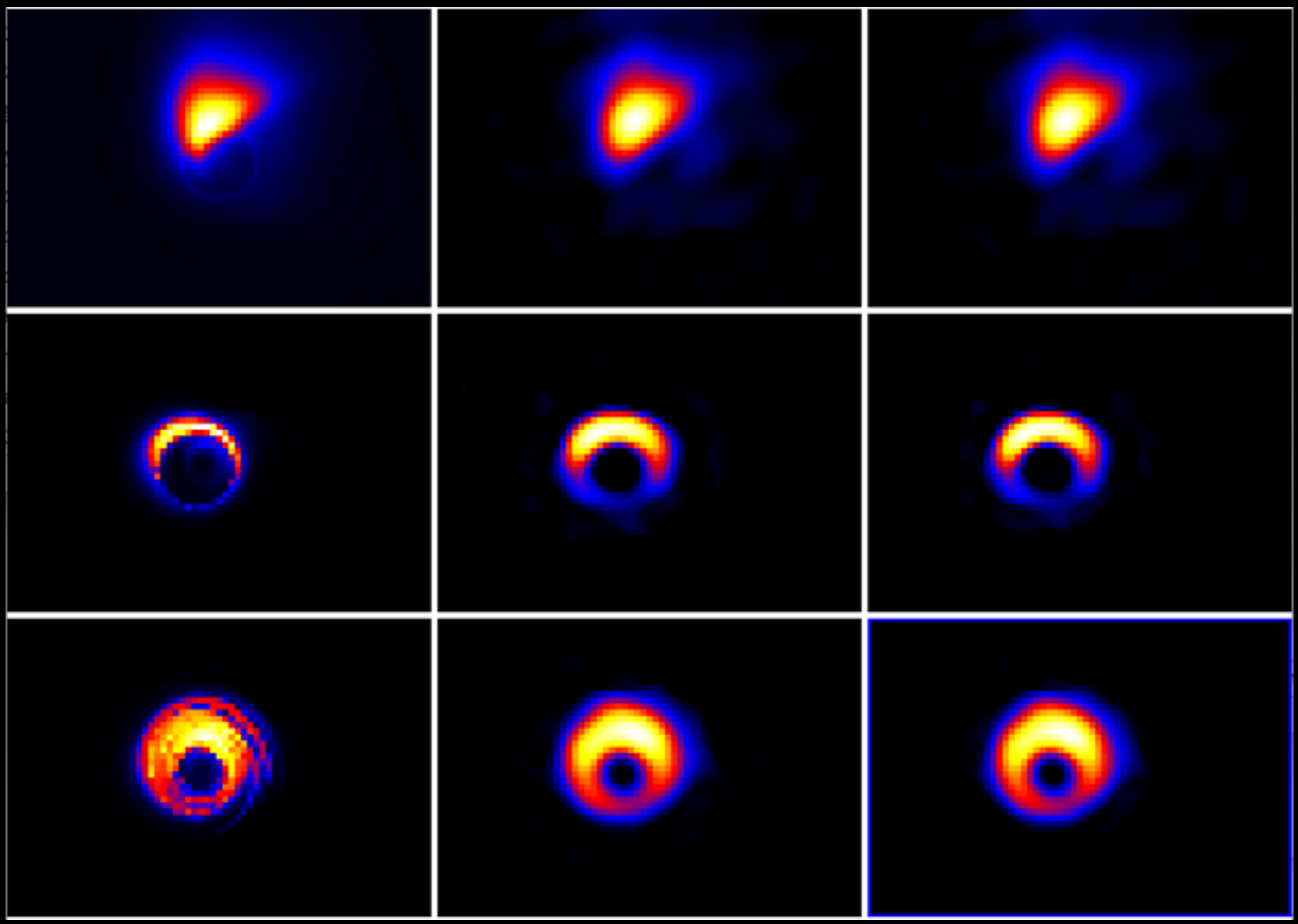
- Noise treatment: thermal, different by baselines
  Realistic thermal noises are included

- Modeling method: LASSO + additional regularization term
  Solver : QP solver (original; MATLAB-based); common threshold

$$\mathbf{I} = \mathrm{argmin}\left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda||\mathbf{I}||_1\right) \quad \text{subject to} \quad I_i \geq 0$$

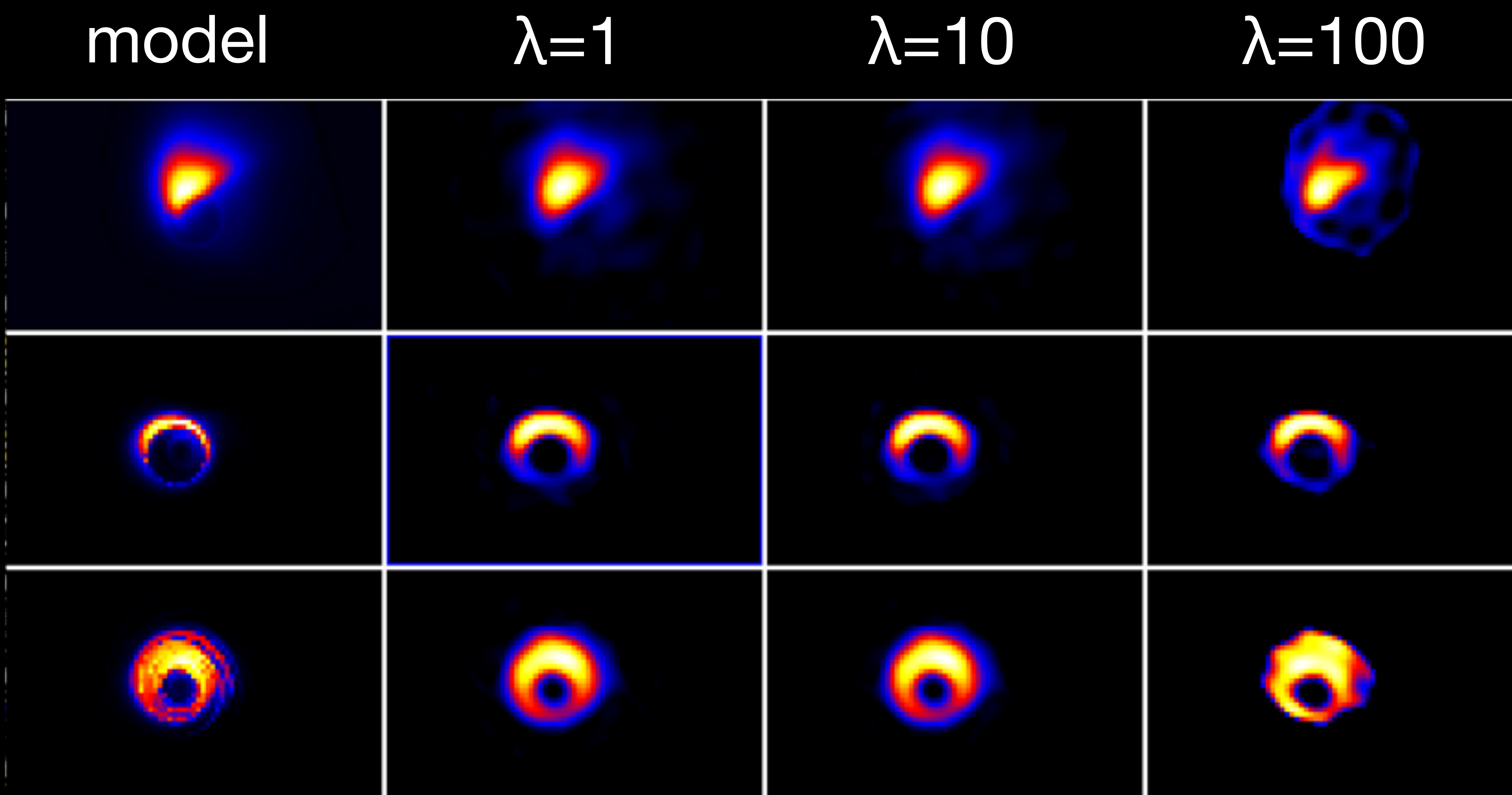
Simulation with Sparse Modeling (II)

model without noise with noise



M87, Shadow Diameter ~ 40 uas (for the case of $M_{BH} = 6 \times 10^9 M_{solar}$)

Simulation with Sparse Modeling (II)



$$\mathbf{I} = \mathrm{argmin}\left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda ||\mathbf{I}||_1\right)$$

M87, Shadow Diameter ~ 40 uas (for the case of $M_{BH} = 6 \times 10^9 M_{solar}$)

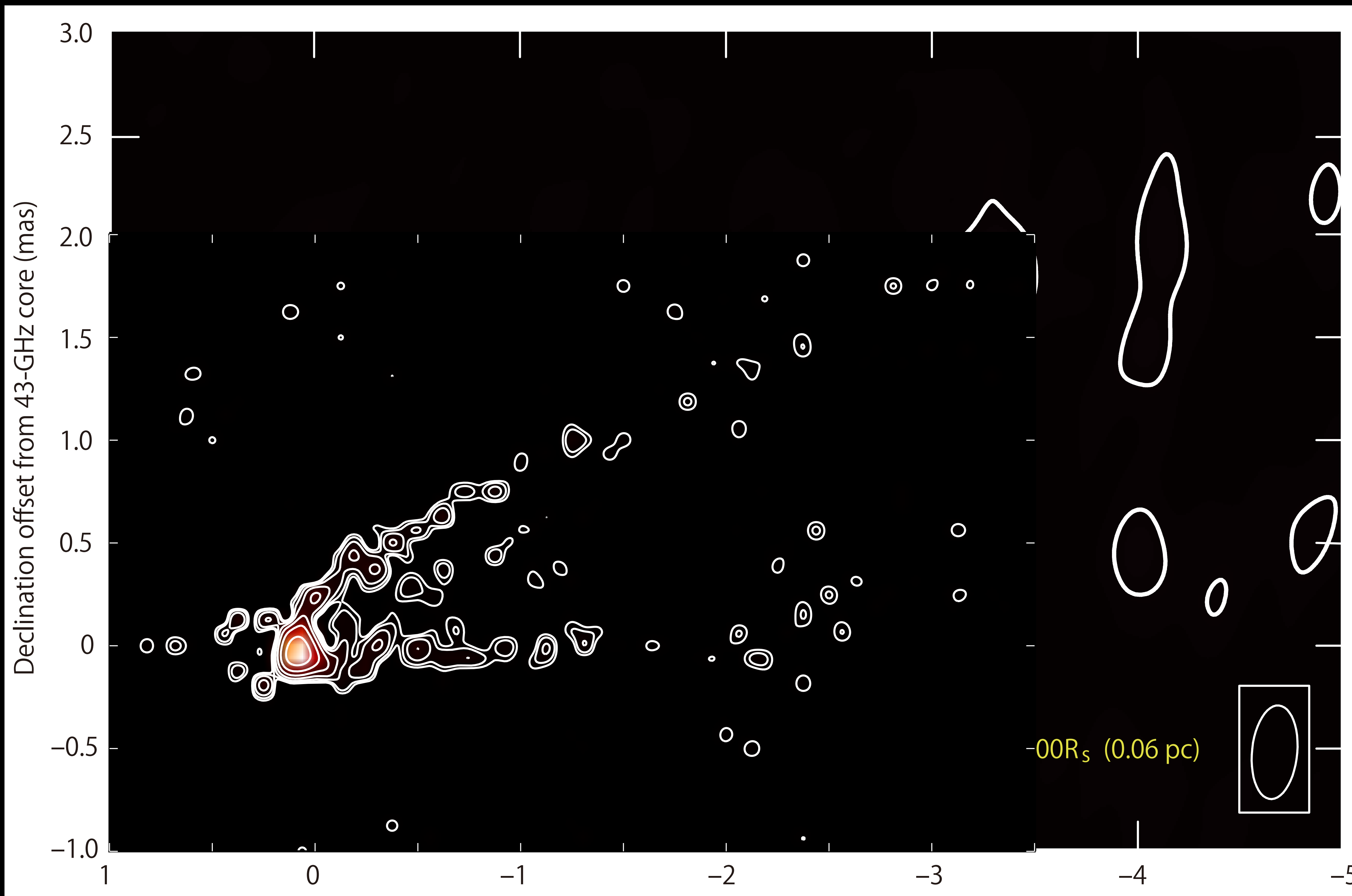
Application to observational data

- Data set: VLBA 43 GHz / 7mm data of M87
  (published in Hada et al. 2011, Nature)

- Modeling method: LASSO + additional regularization term
  Solver : QP solver (original; MATLAB-based); common threshold

$$\mathbf{I} = \mathrm{argmin}\left(||\mathbf{V} - A\mathbf{I}||_2^2 + \Lambda||\mathbf{I}||_1\right) \quad \text{subject to} \;\; I_i \geq 0$$

Application to observational data



Hada et al. 2011, Nature (43 GHz/7 mm)

Outline of today's talk

- Our motivation to work on the imaging technique
- Application of the sparse modeling to the interferometric imaging
    - The basic idea
      (see §2 in Honma, ***KA,*** Uemura & Ikea 2014, PASJ)
    - Mathematical description
      (see §3 in Honma, ***KA,*** Uemura & Ikea 2014, PASJ)
    - Results of the sparse modeling on observational/simulated data
- Practical issue: Imaging pipeline with the sparse modeling

Hybrid mapping with the sparse modeling (3 mm or longer-λ data)

Initially Calibrated Visibility
data fmt: UVFITS

Imaging with LASSO

Currently: MALAB-based
now translating to
Fortran 90 + python

Self-calibrated Visibility
data fmt: UVFITS

Self-calibration
correcting residual gains
using an edited image as model

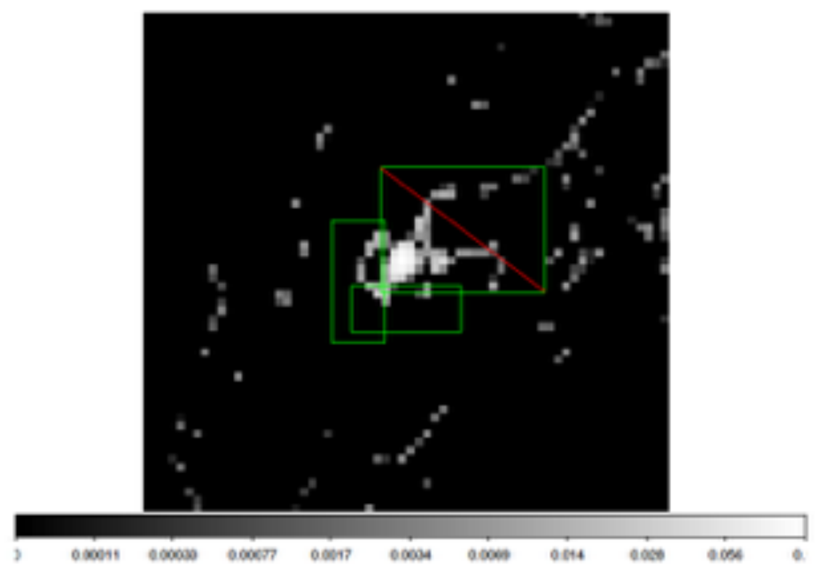
Python script
using Parsel Tongue (AIPS)

Model Image
data fmt: image FITS

Image Editing
removing noise features

Edited image
data fmt: image FITS

Final Products
Image and Calibrated visibility
(UV FITS + image FITS)



SAO DS9 (select area)
+
Python script (editing)

Hybrid mapping with the sparse modeling (3 mm or longer-λ data)

How to determine parameters?

the pixel size of the image (spatial resolution)

Λ-term (sparseness)

Evaluating goodness-of-fit with some information criterions

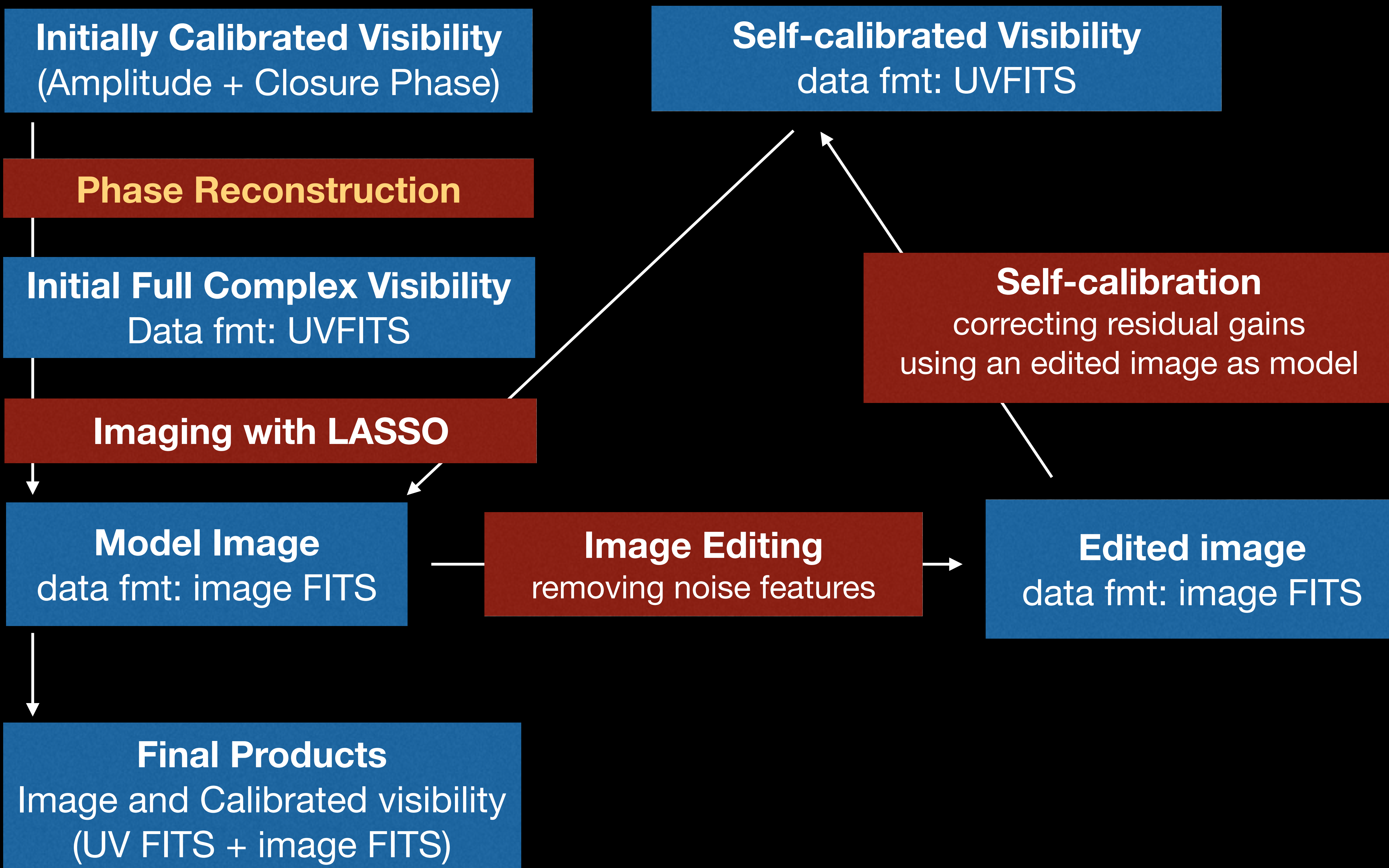
AIC (Akaike's Information Criterion)

$$\mathrm{AIC} = \chi^2 + 2\,||\mathbf{I}||_0$$

BIC (Bayesian Information Criterion)

$$\mathrm{BIC} = \chi^2 + (N_{\mathrm{data}} - ||\mathbf{I}||_0)\ \mathrm{ln}\ N_{\mathrm{data}}$$

Hybrid mapping with the sparse modeling (EHT data; Case 1)



Hybrid mapping with the sparse modeling (EHT data; Case 1)

Phased Reconstruction from the closure phase

Assumption: visibility phase is smoothly distributed

min $C(\phi, \xi)$ subject to $\psi = A\phi$

visibility phase

closure phase

$$C(\phi,\xi) = \frac{1}{2}\sum_{j\neq k} w_{jk}(\phi_j - \phi_k - \xi_{jk})^2$$

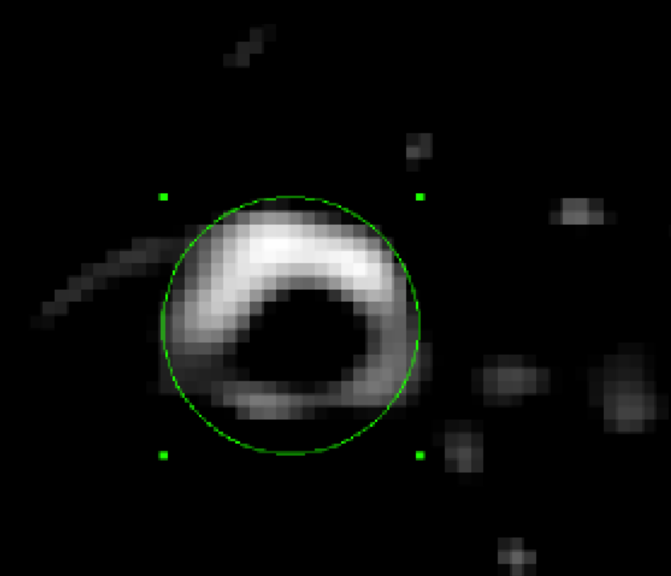
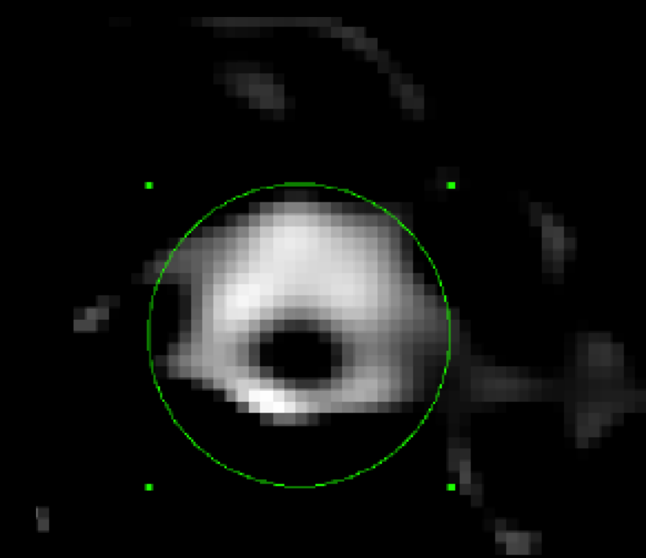
$$w_{jk} = \exp\left(-\lambda_r\sqrt{|r_j^2 - r_k^2|}\right)\exp\left(-\lambda_\theta|\theta_{jk}|\right)$$

This problem can be solved with quadratic programming (QP)

(Ikeda, Tazaki, KA et al. to be subm.)

Hybrid mapping with the sparse modeling (EHT data; Case 1)

Reconstructed Image from visibility amp. + reconstructed phase

Reconstructed Image from full-complex visibility



(Ikeda, Tazaki, KA et al. to be subm.)

Hybrid mapping with the sparse modeling (EHT data; Case 2)