

Computational Imaging for VLBI Image Reconstruction

Supplemental Material

Katherine L. Bouman¹ Michael D. Johnson² Daniel Zoran¹ Vincent L. Fish³
Sheperd S. Doeleman^{2,3} William T. Freeman¹

¹Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology

²Harvard-Smithsonian Center for Astrophysics, Harvard University

³Haystack Observatory, Massachusetts Institute of Technology

Contents

1 Additional Results and Parameters	2
1.1 Method Parameters and Visual Results	2
1.2 EHT Telescope Parameters	26
1.3 Blind Test Data	26
2 Energy and Optimization	28
2.1 Optimization Method 1 - Using Gradients	28
2.2 Optimization Method 2 - Taylor Expansion	29
2.3 Derivatives	31
3 Interstellar Scattering Kernel	33
3.1 Gaussian Interstellar Scattering Kernel	33
4 Noise	34
4.1 Visibility Noise	34
4.2 Bispectrum Noise	34
4.3 Gaussian Noise Model for the Bispectrum Including Amplitude Error	34

1 Additional Results and Parameters

1.1 Method Parameters and Visual Results

We compare results from our algorithm, CHIRP, with the three state-of-the-art algorithms described in Section 3 of our paper. As with our algorithm, SQUEEZE [1] and BSMEM [2] use the bispectrum as input. To eliminate bias, images were obtained by asking either the authors of the competing algorithms or knowledgeable users for reconstruction parameters. CLEAN was run using the *calibrated* (eg. no phase error) visibilities in CASA [3]. In reality, these ideal calibrated visibilities would not be available, and the phase would need to be recovered through highly user-dependent self-calibration methods. However, in the interest of a fair comparison, we show the results of CLEAN in a “best-case” scenario.

For image results presented, we use synthetic data corresponding to parameters of the EHT telescope array (Refer to Section 1.2) when pointed towards the black hole in M87. The specific parameters used are:

- FOV Center Right Ascension (HH:MM:SS.SS): 12:30:49.423382
- FOV Center Declination(DD:MM:SS.SS): 12:23:28.04366
- FOV Size: 0.00018382 x 0.00018382 arcseconds
- Array: EHT (Parameters in Section 1.2)
- Center Frequency: 227297 MHz
- Bandwidth: 4096 MHz
- Integration Time: 12 seconds

```
fov = 0.00018382
```

```
numberOfPixels = 64
```

CLEAN command: clean(vis=msFile,imagingname=cleanFile, cell= str(fov/numberOfPixels*1000) + 'arcsec', threshold='0.001Jy', weighting='briggs', imsize=[numberOfPixels, numberOfPixels])

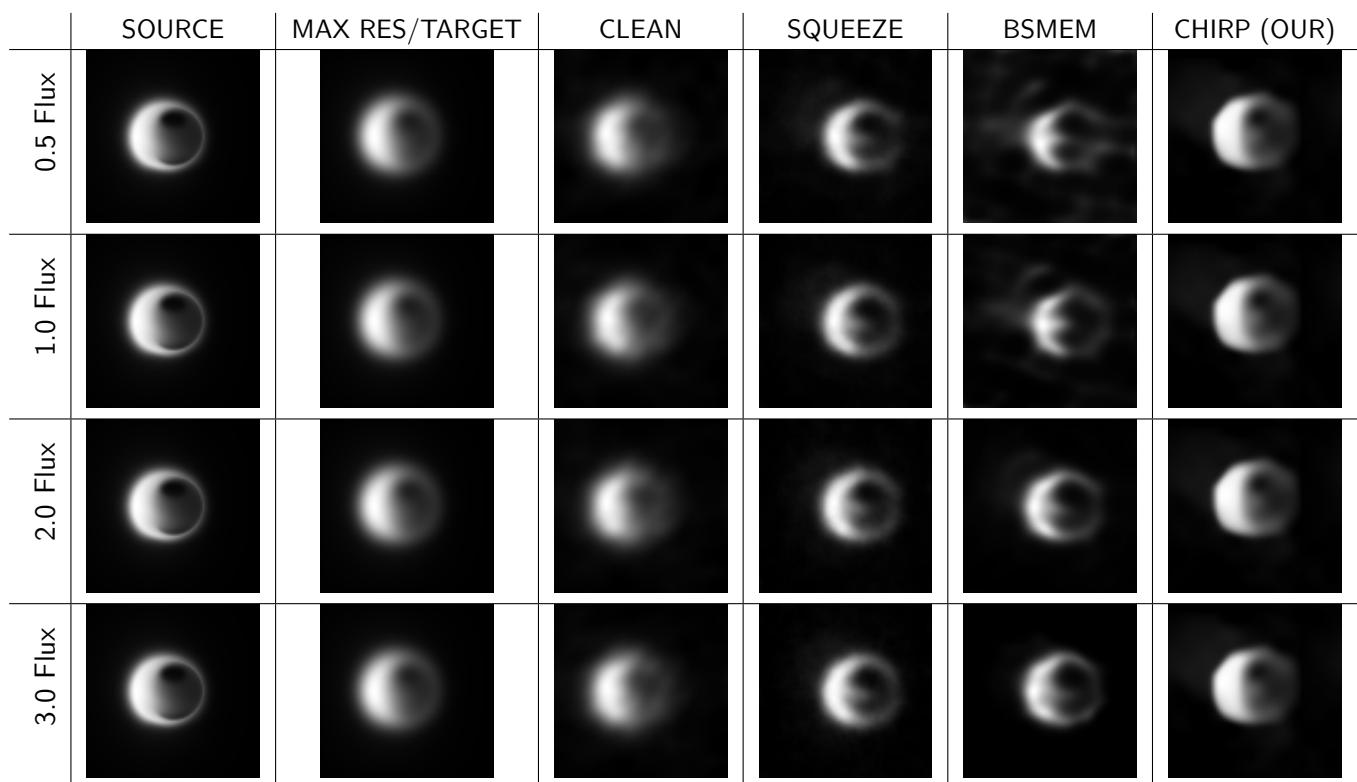
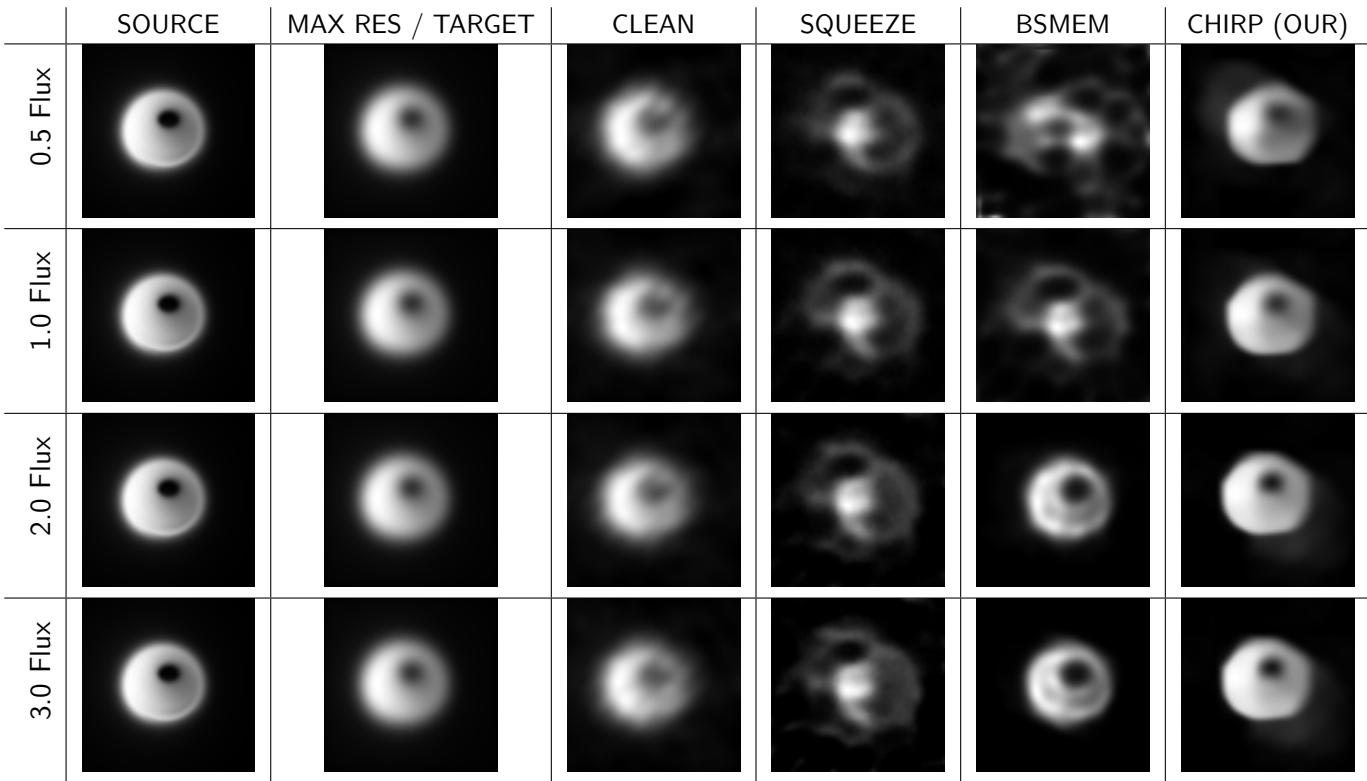
SQUEEZE command: './src/squeeze ' + dataPath + '/' + filename + ' -s ' + str(fov/numberOfPixels*1000) + ' -w 64 -fs ' + str(totalFlux) + ' -tv 3000 -e 10000 -novis'

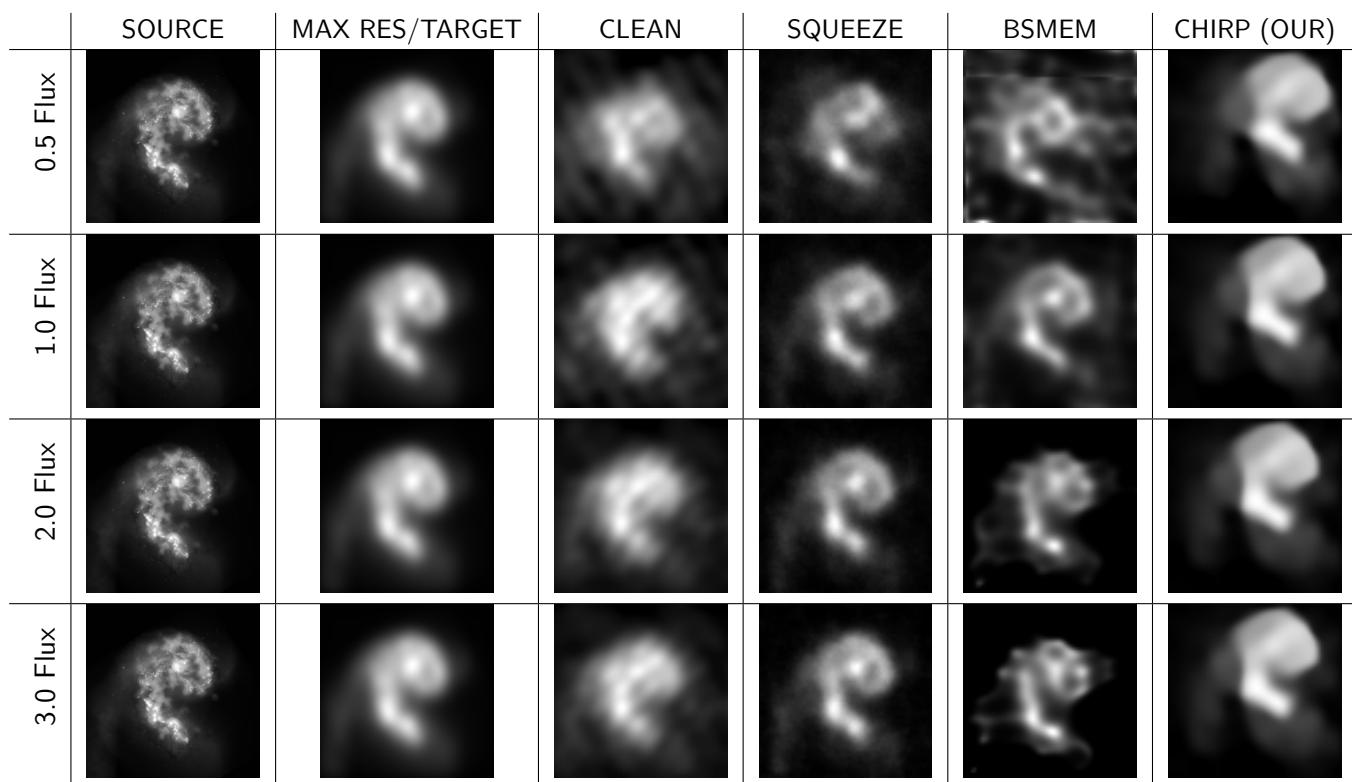
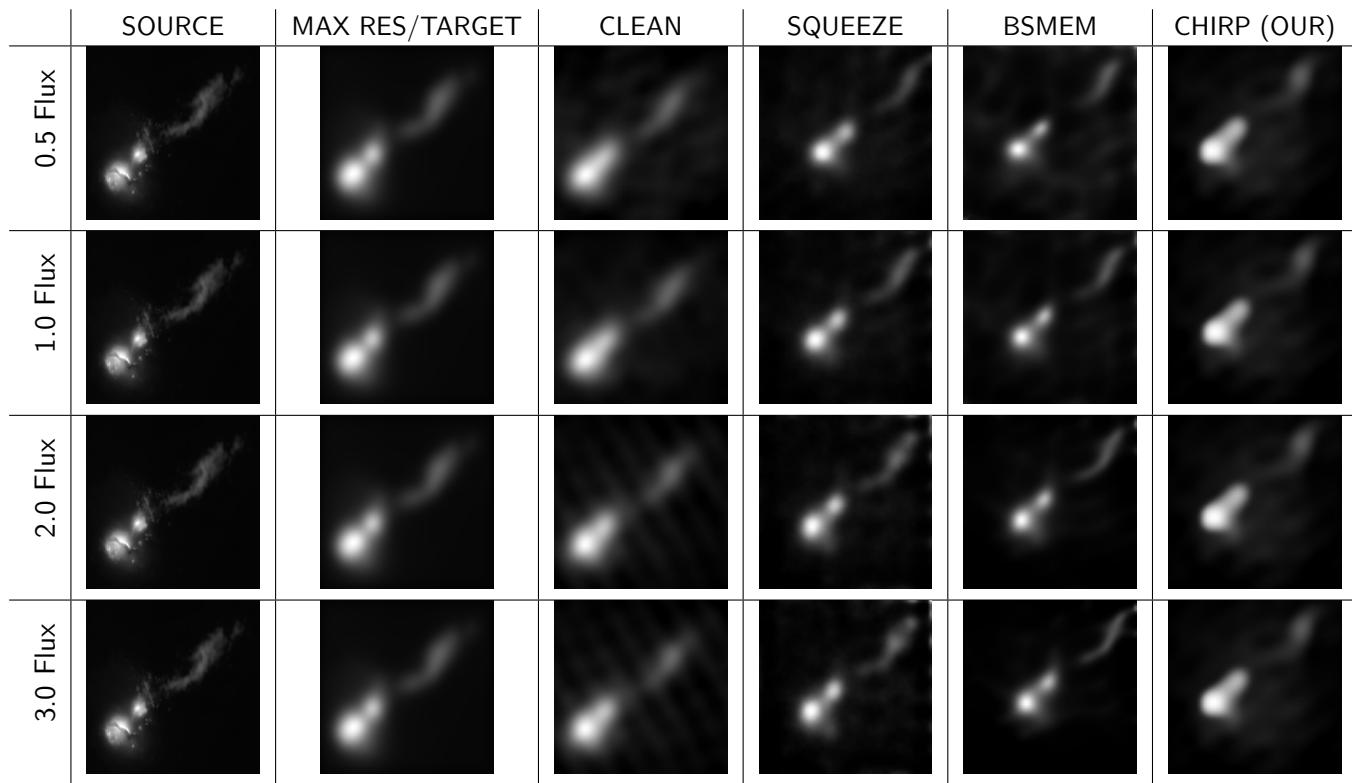
BSMEM command: './src/bsmem -d ' + dataPath + '/' + filename + ' -mt 0 -mf ' + str(totalFlux) + ' -p ' + str(fov/numberOfPixels*1000) + ' -w ' + str(numberOfPixels) + ' -wavmin 0 -wavmax 20000000 -it 20 < bsmem.batch'

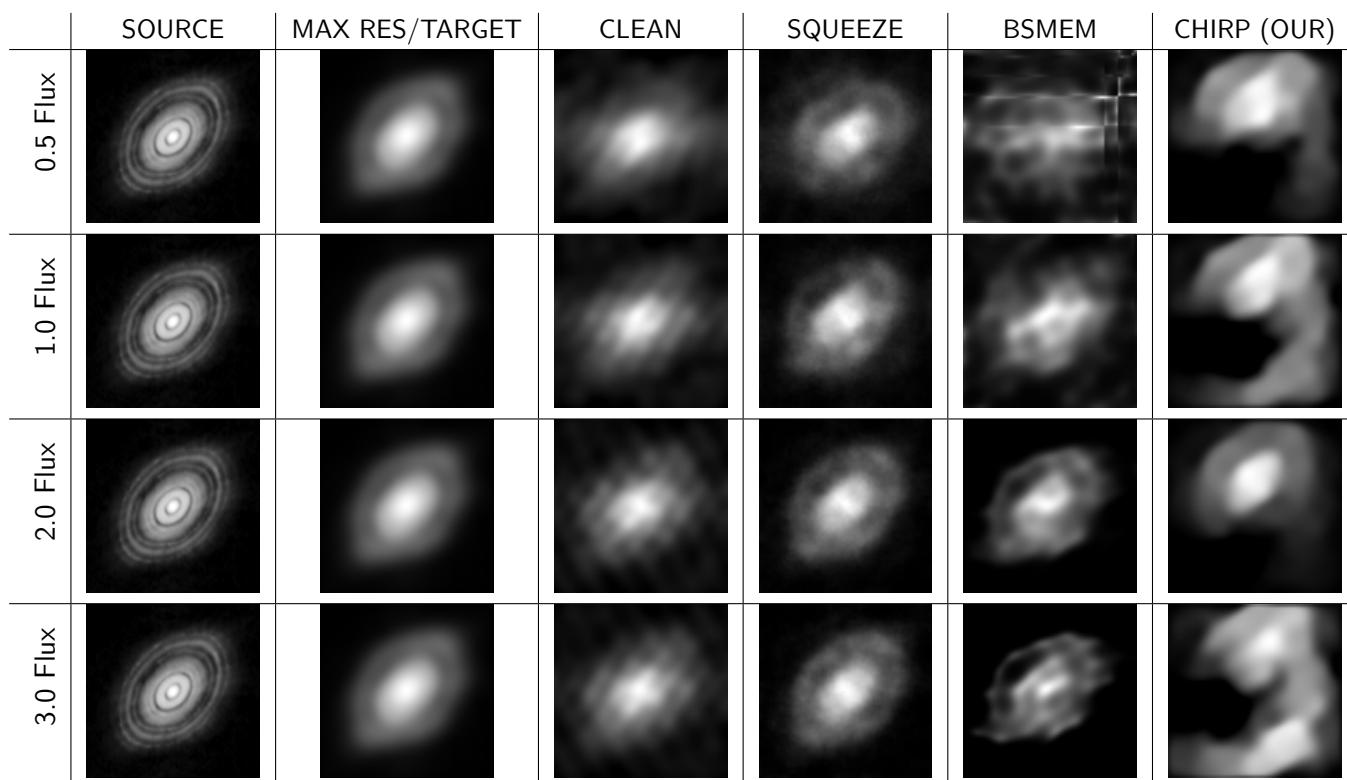
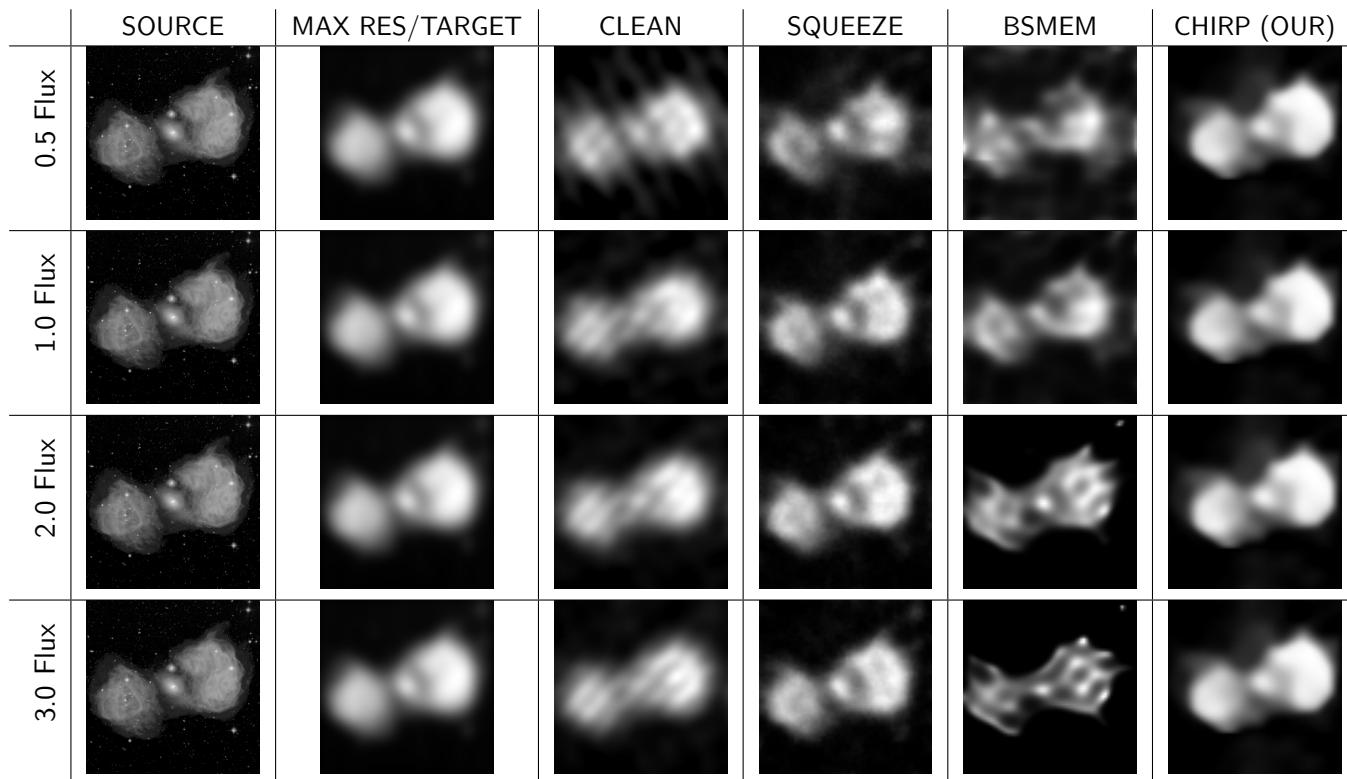
bsmem.batch file contents:

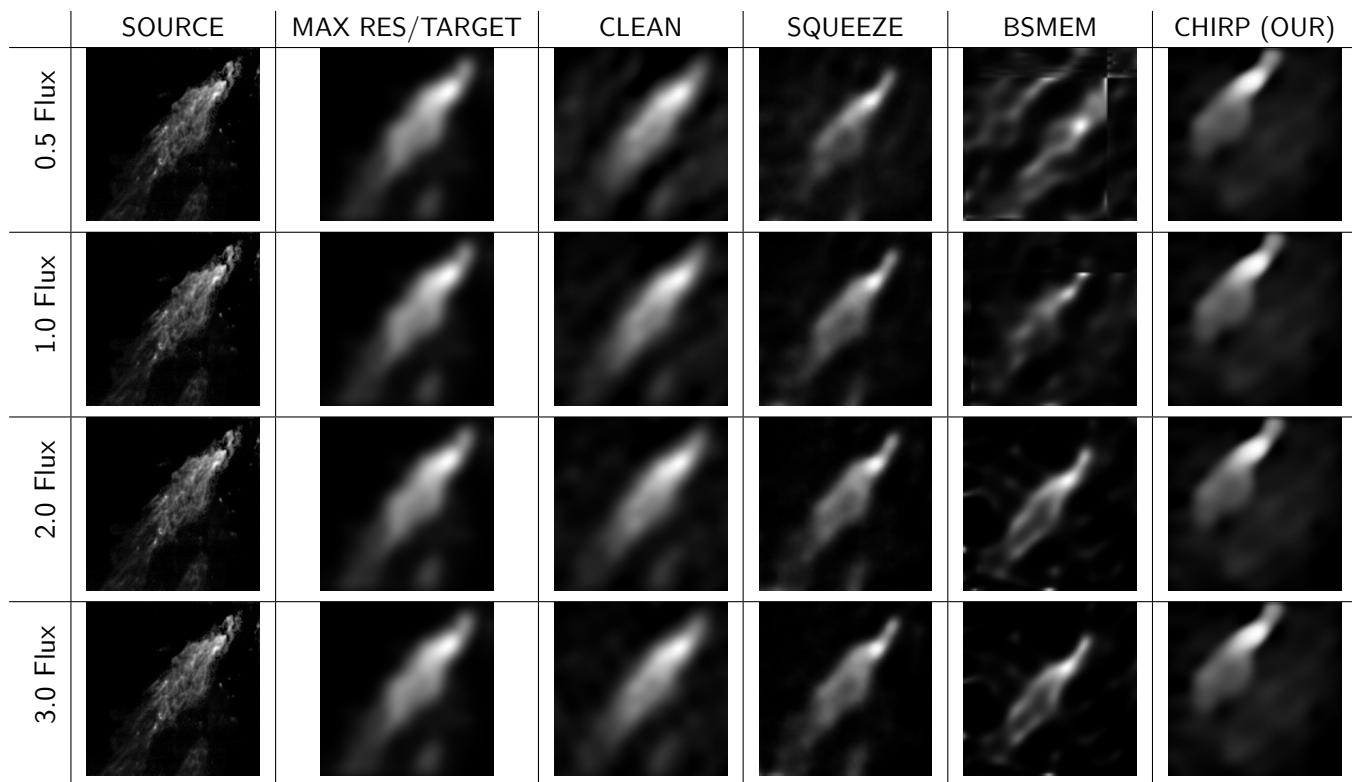
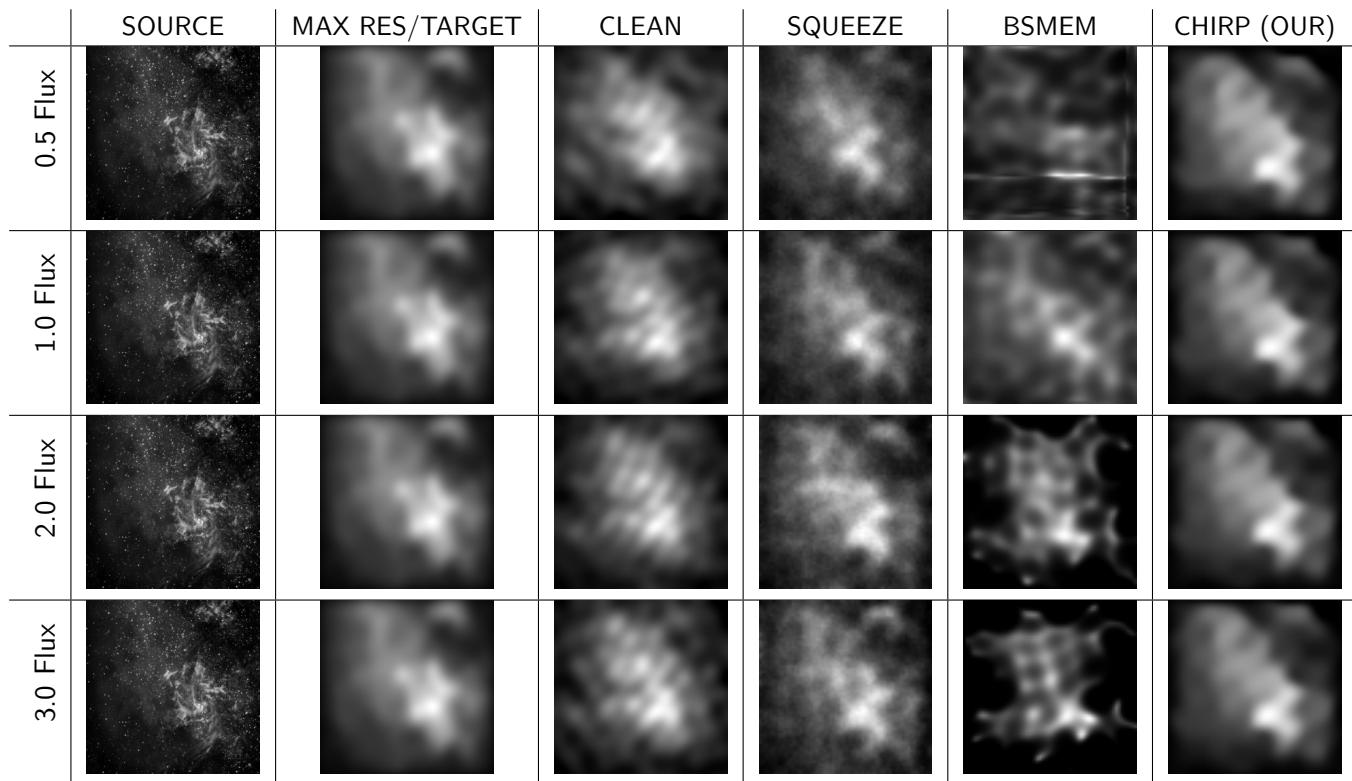
```
batch
do 10
center
do 20
writefits output.fits
exit
```

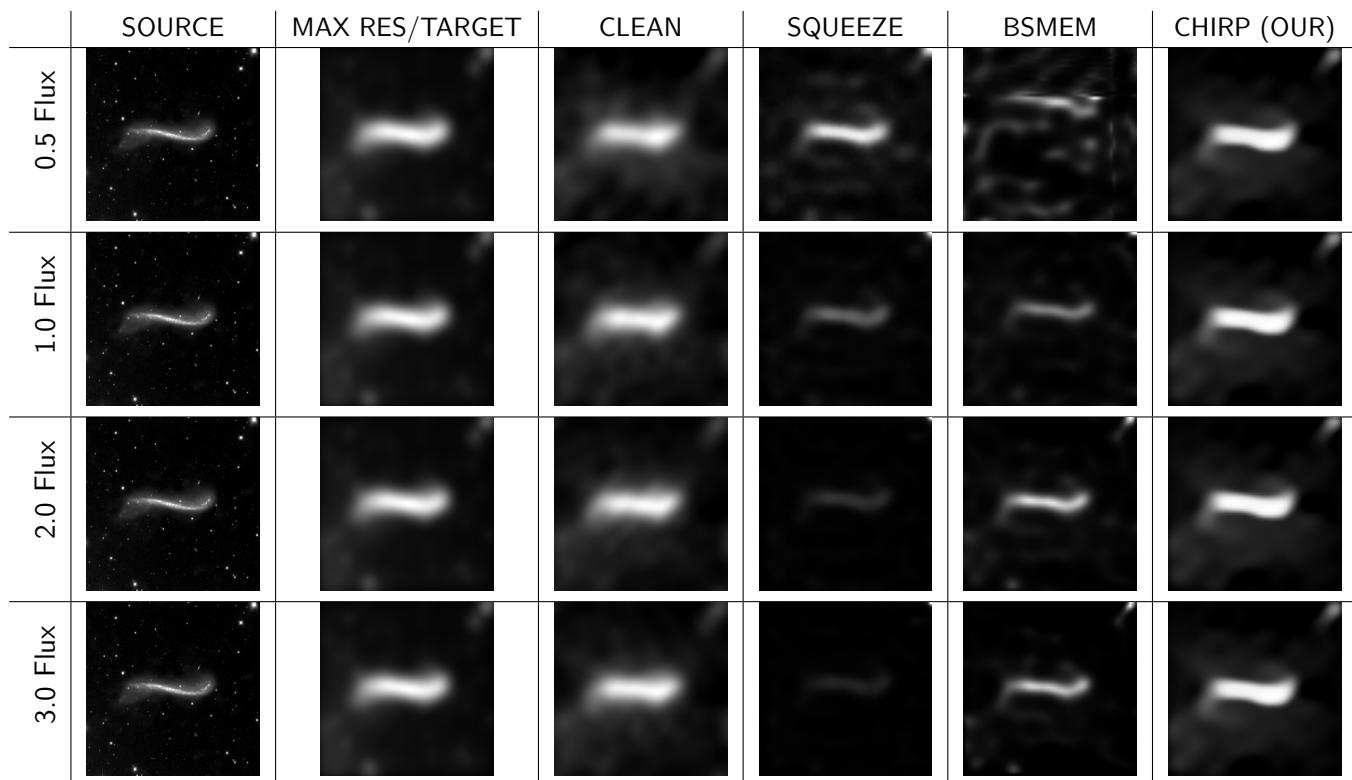
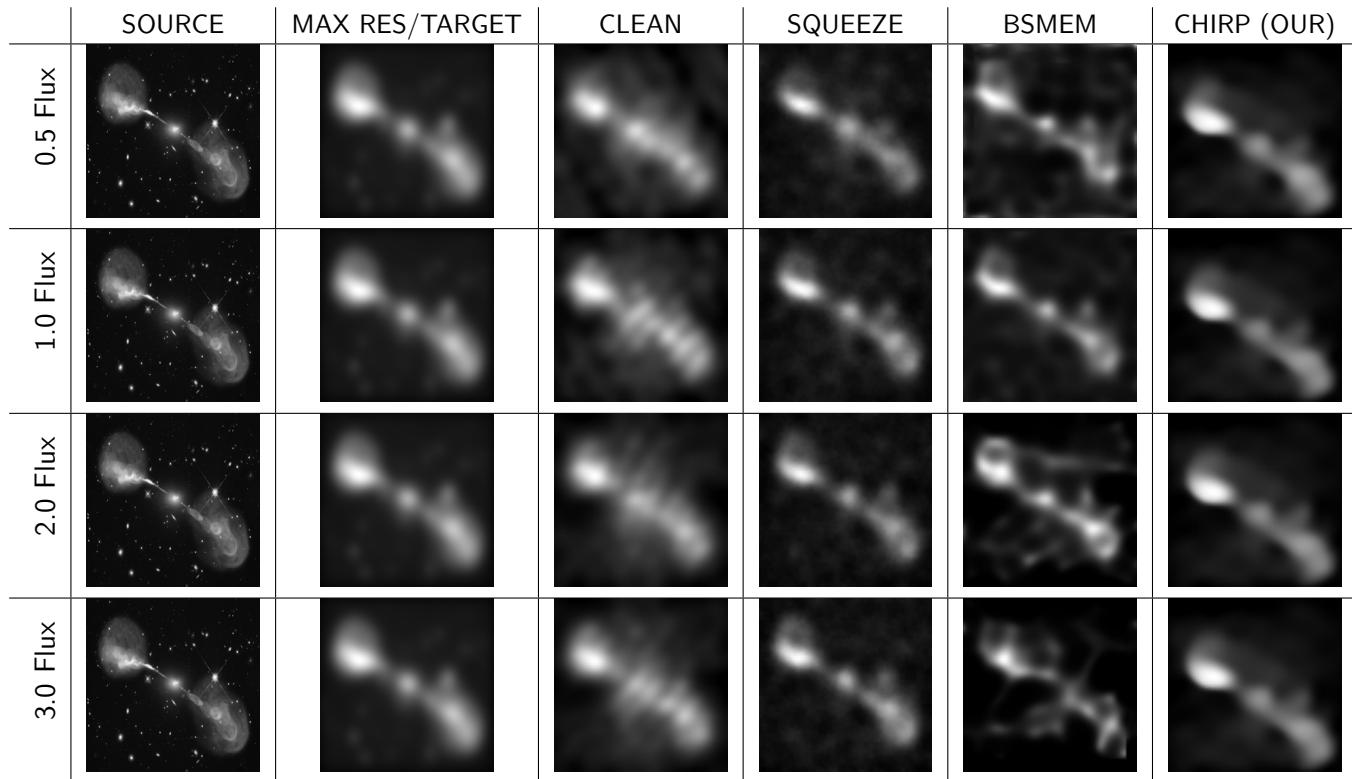
We show results for each image with varying levels of noise. The standard deviation of thermal noise introduced in each measured visibility is fixed based on measurement choices and the corresponding telescopes’ intrinsic properties. Consequently, a source emitting a lower total flux density will result in a signal with lower SNR.

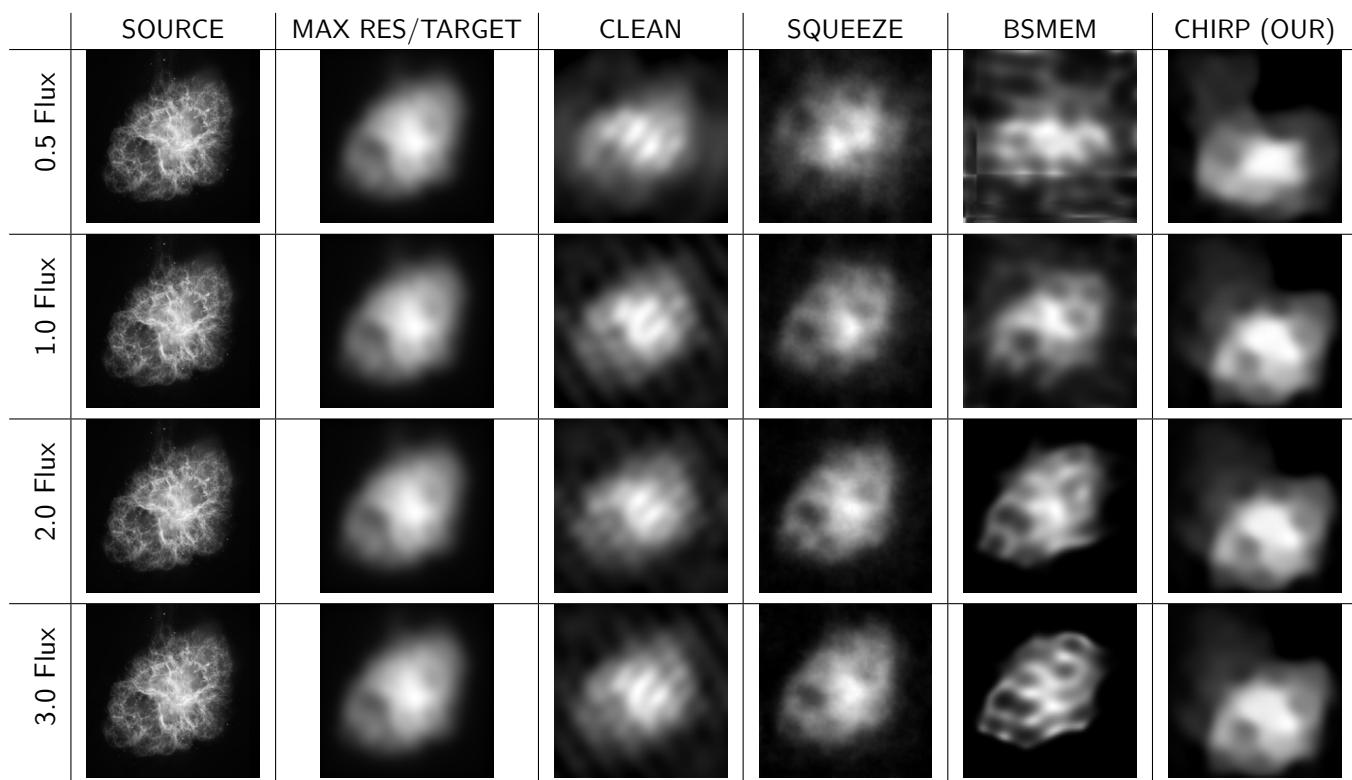
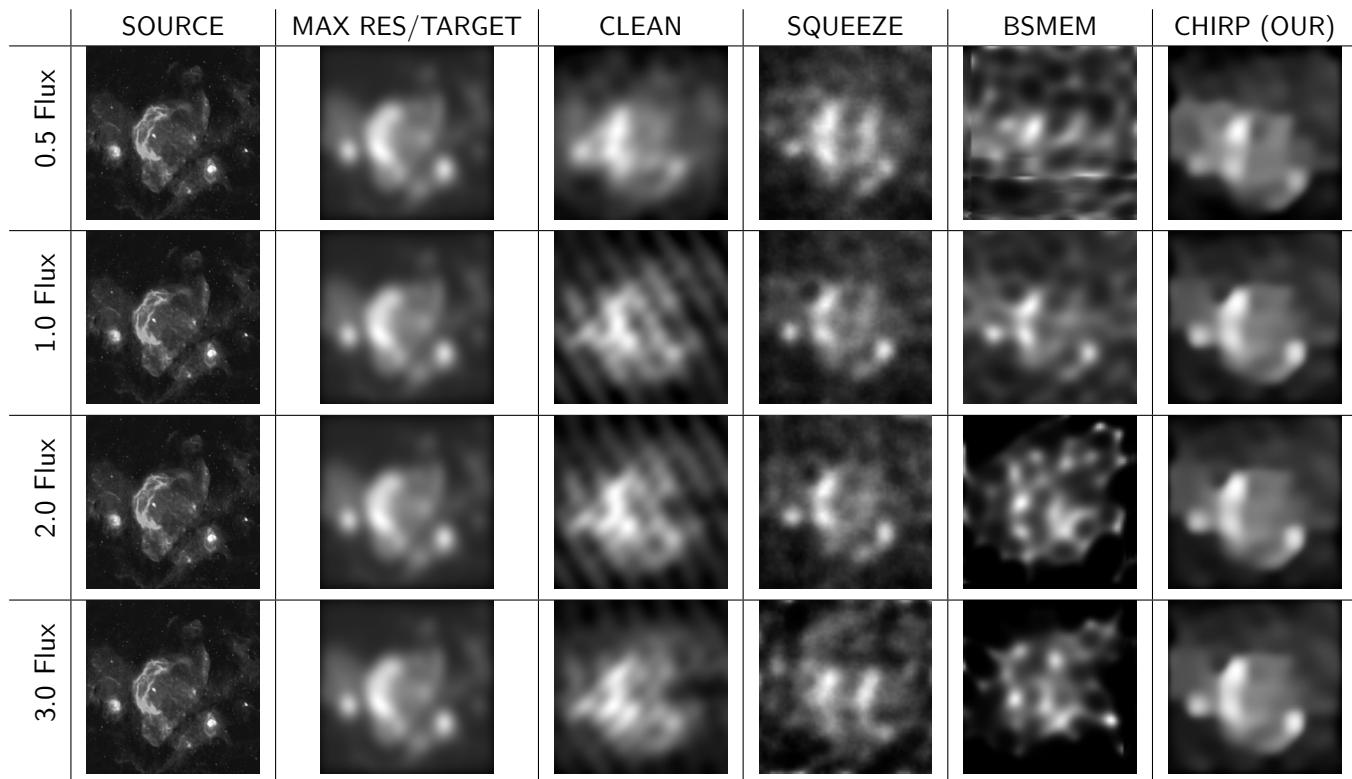


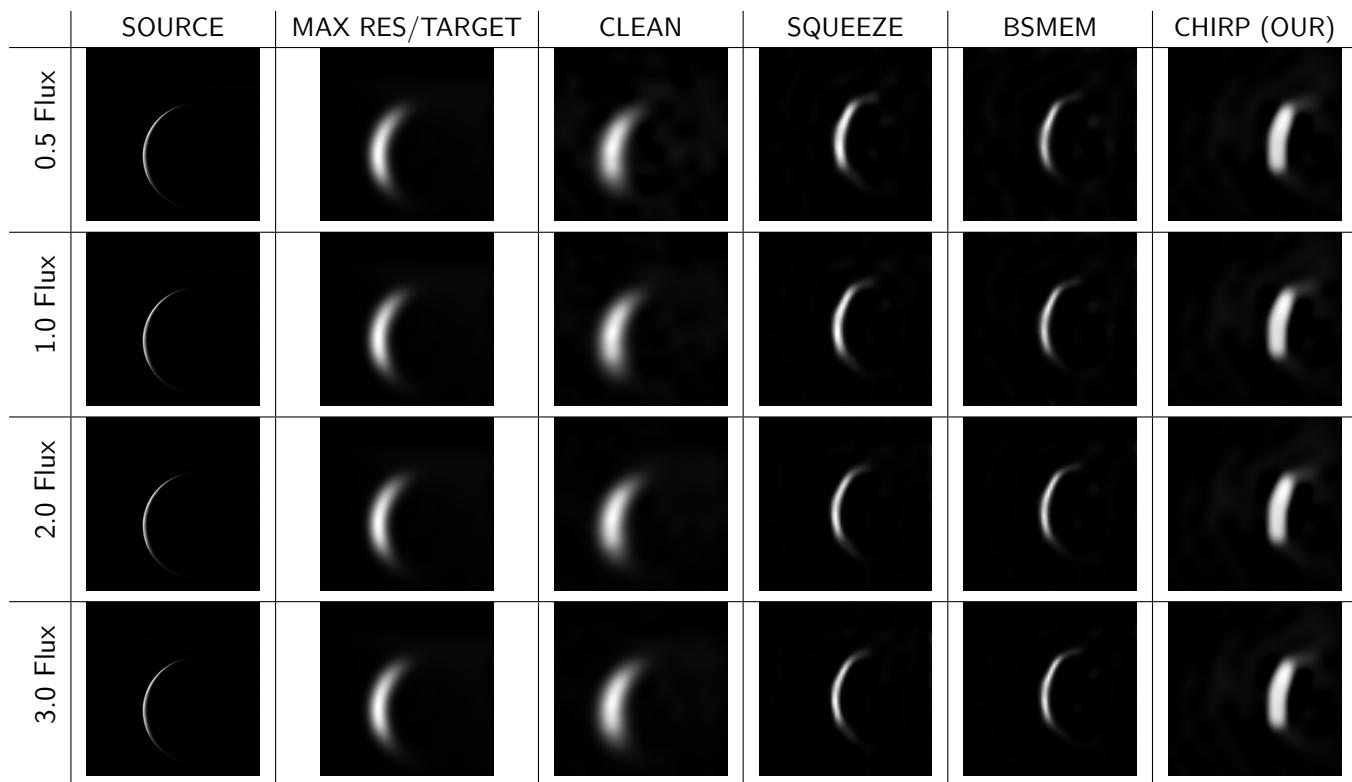
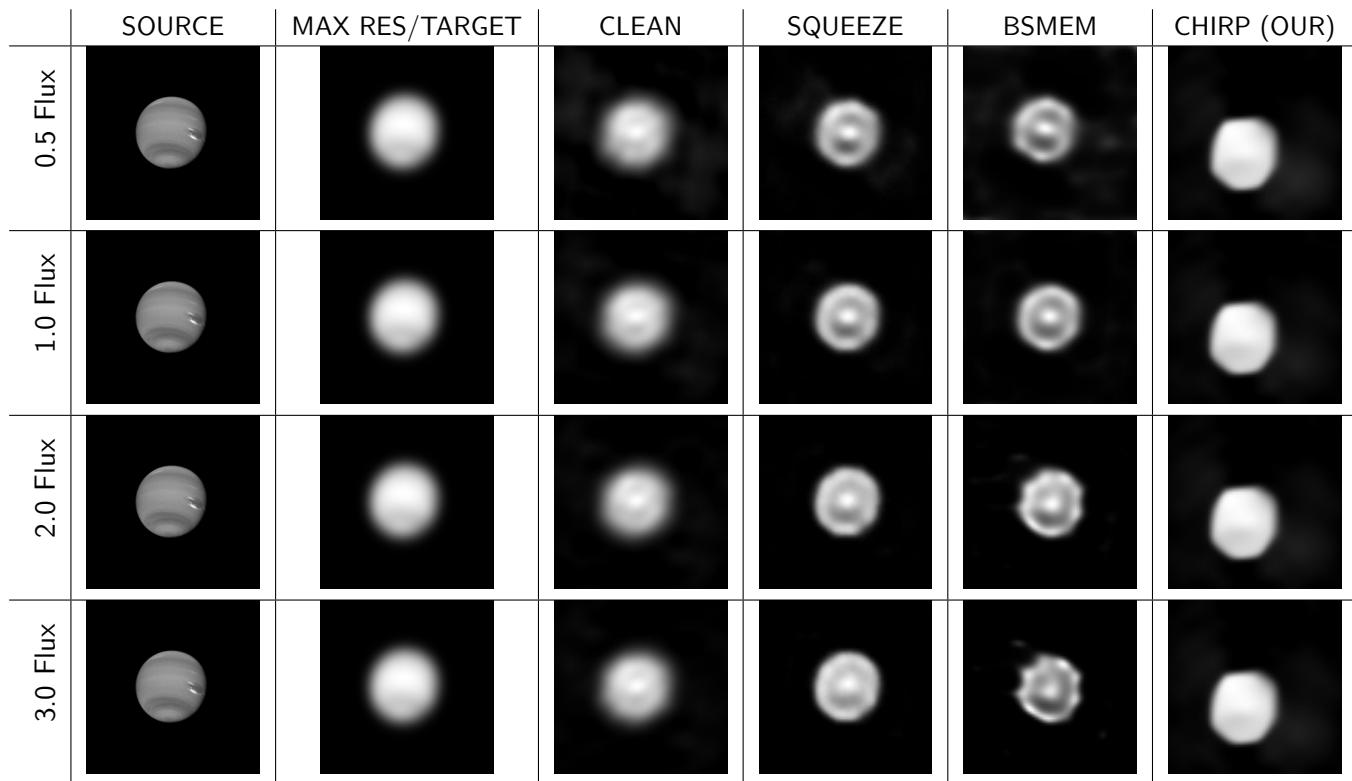


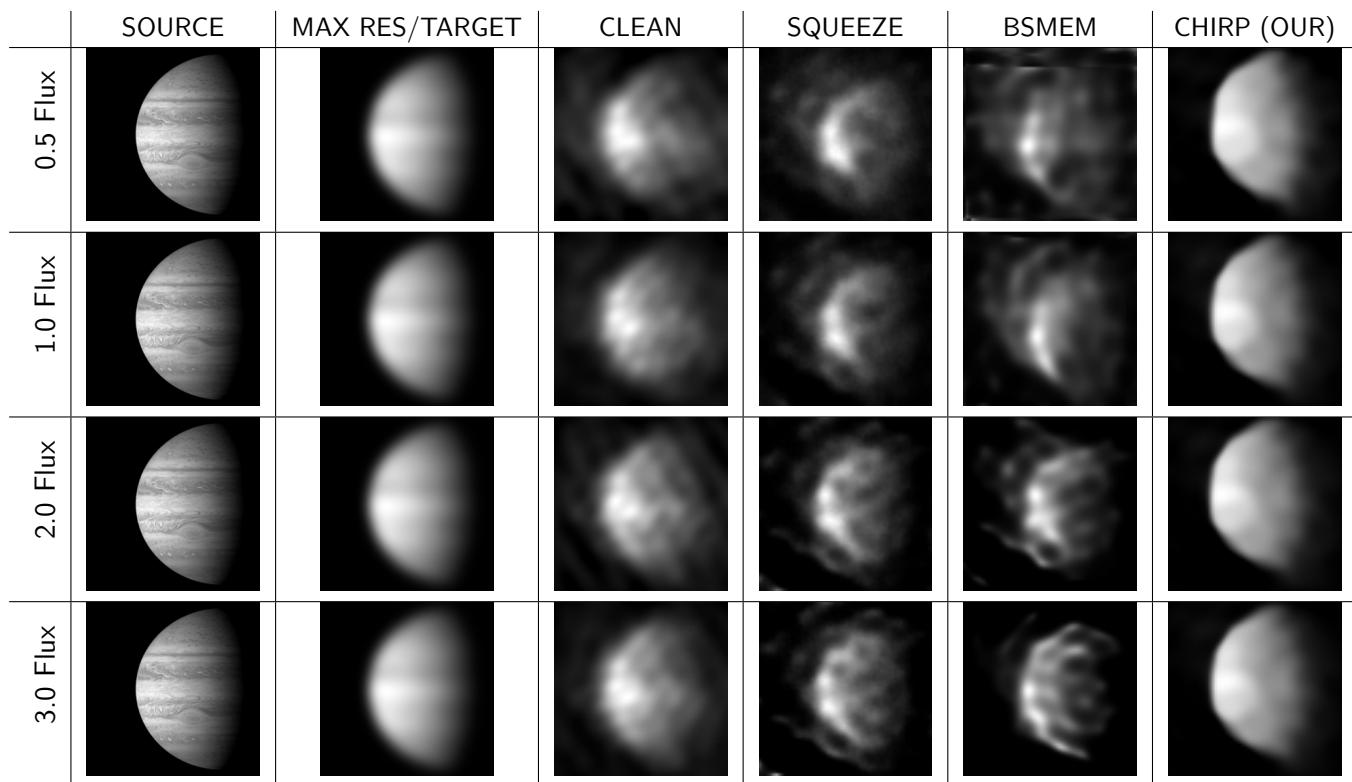
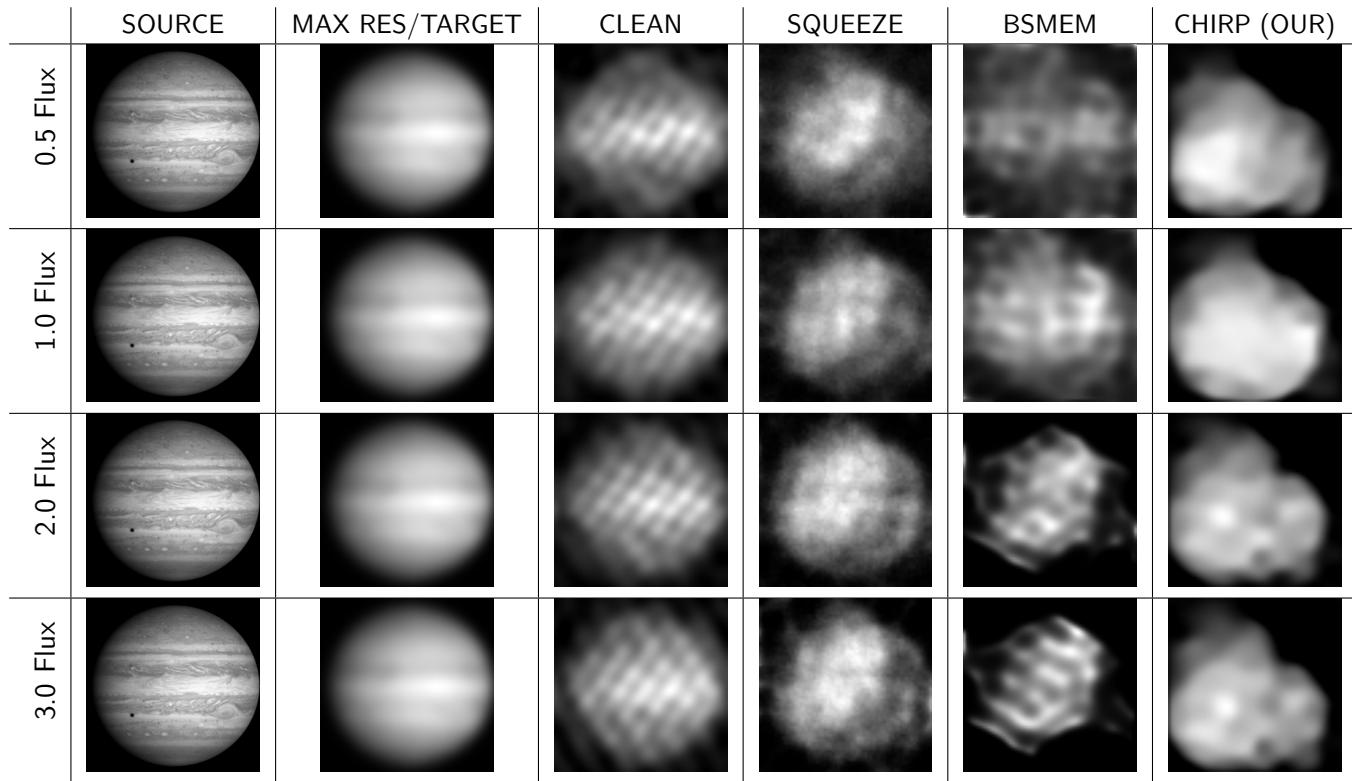


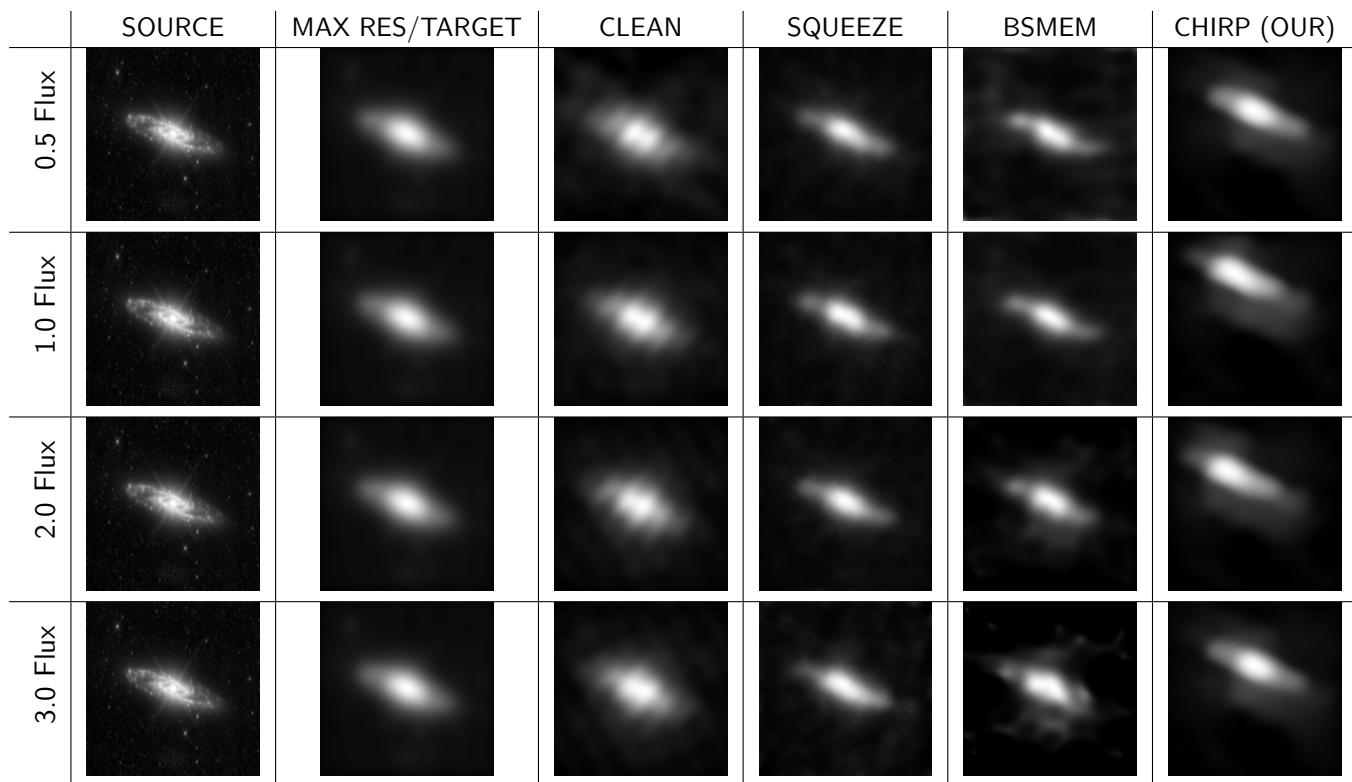
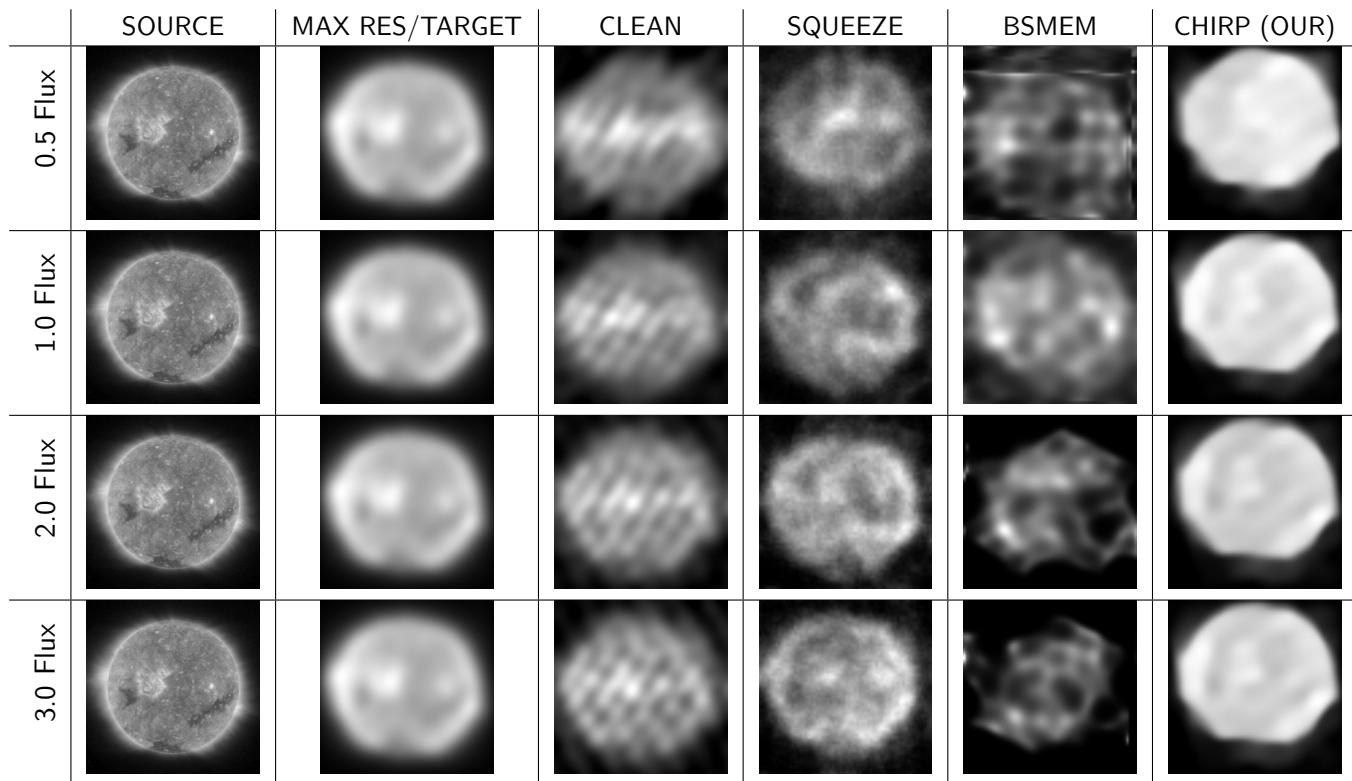


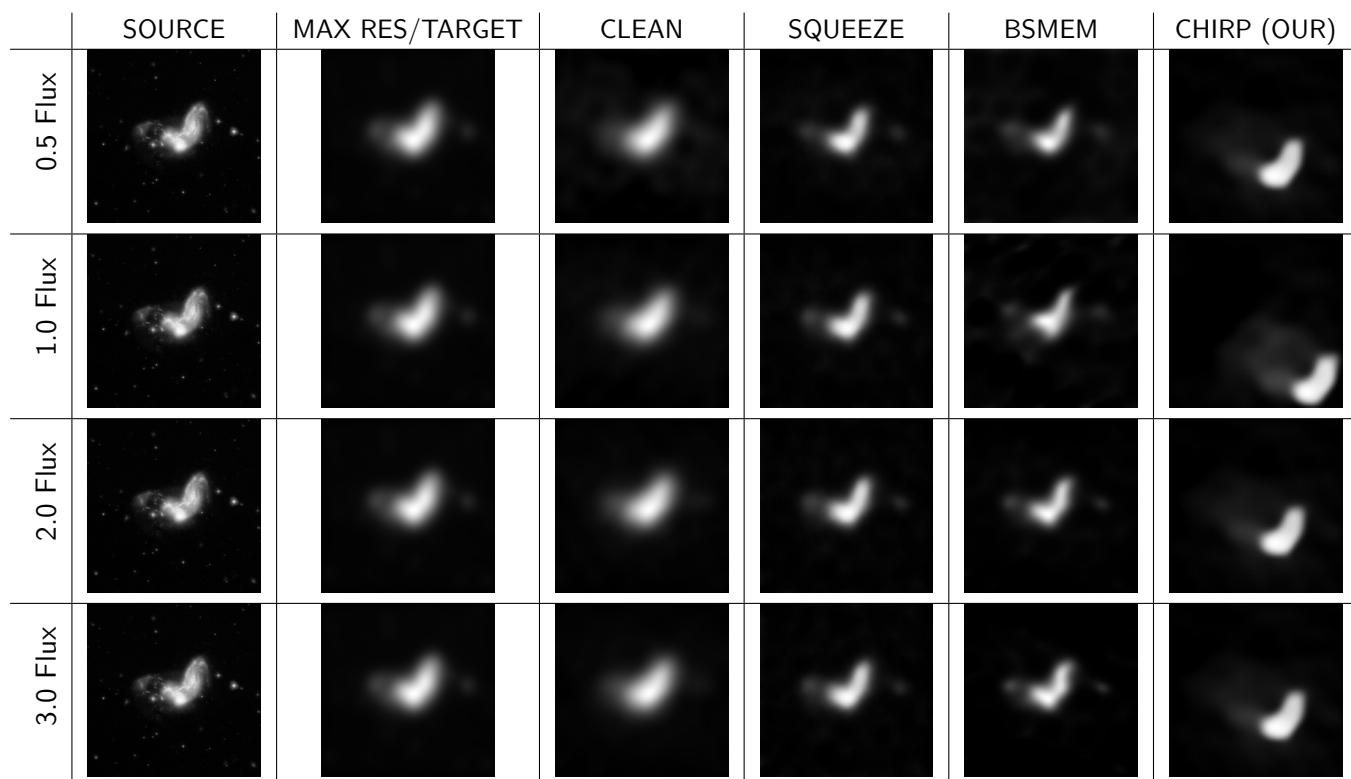
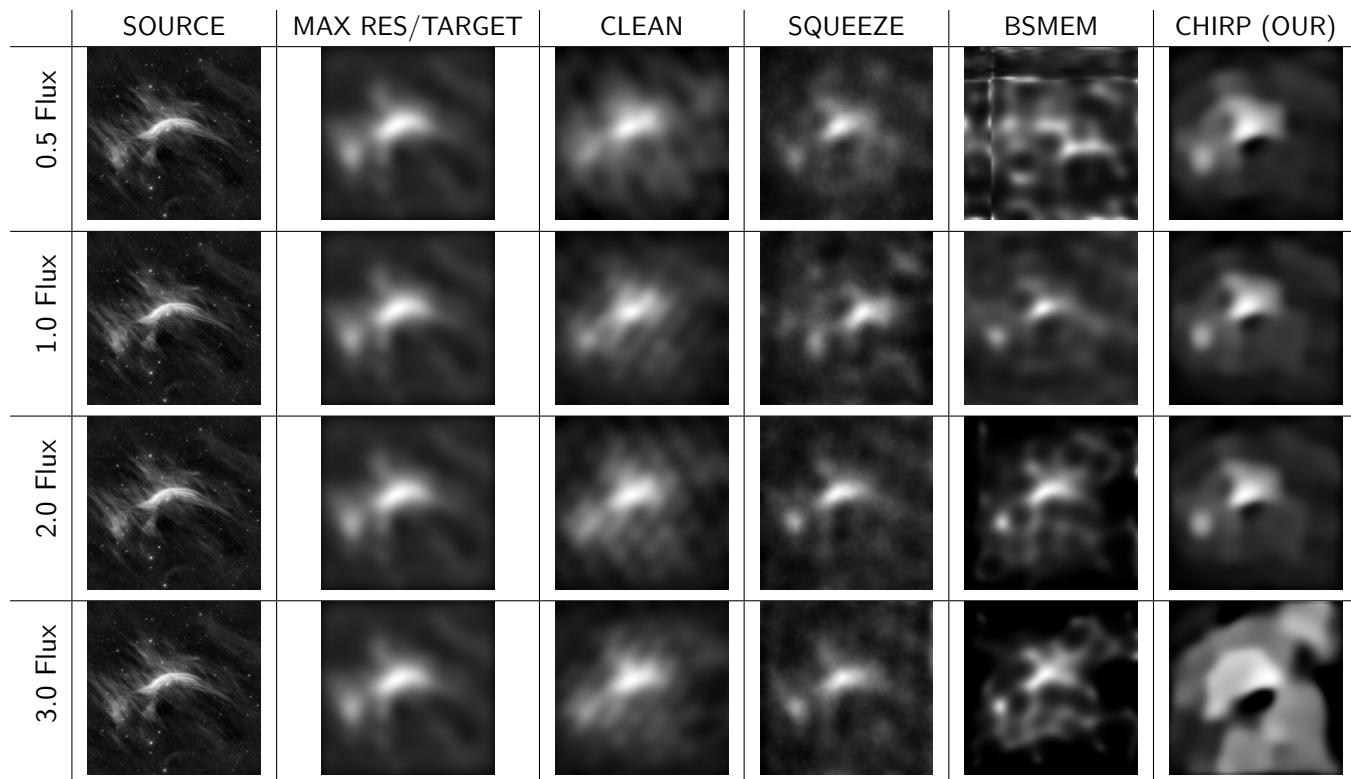


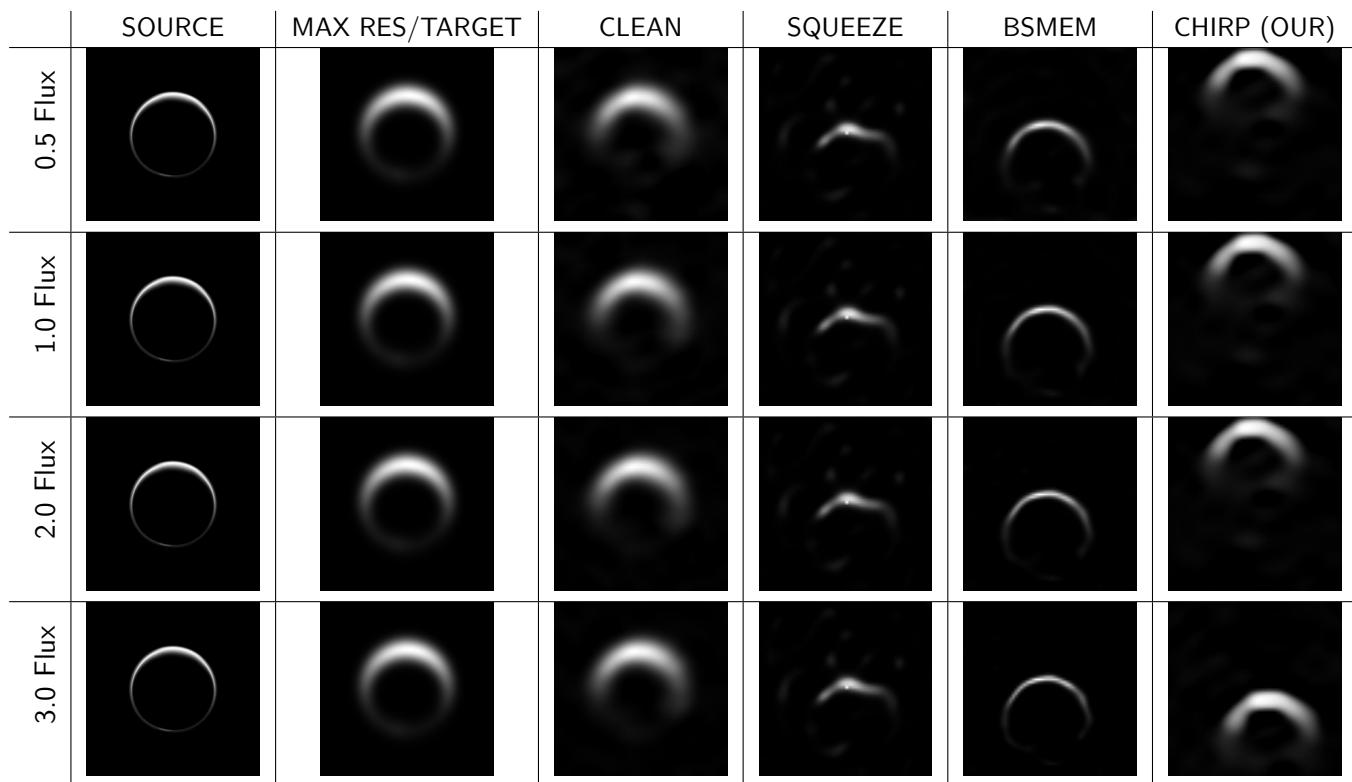
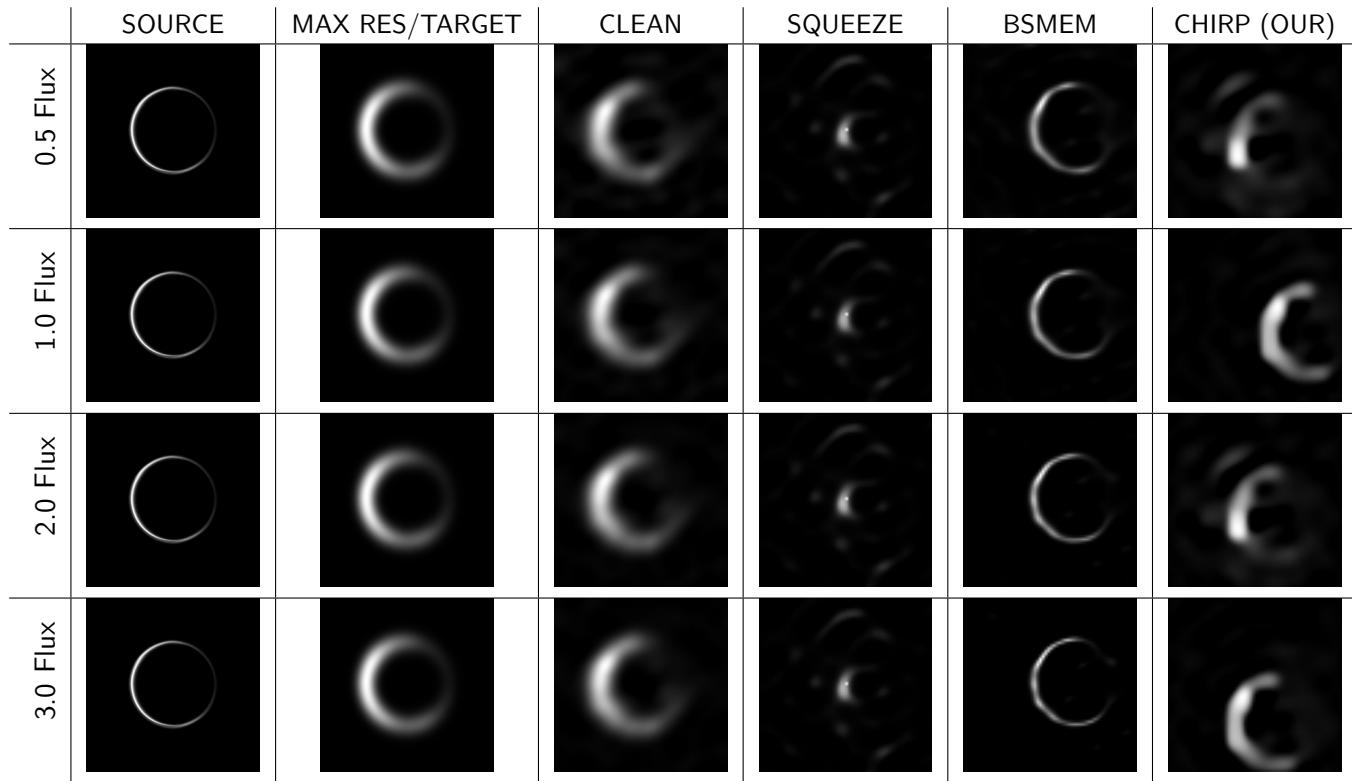


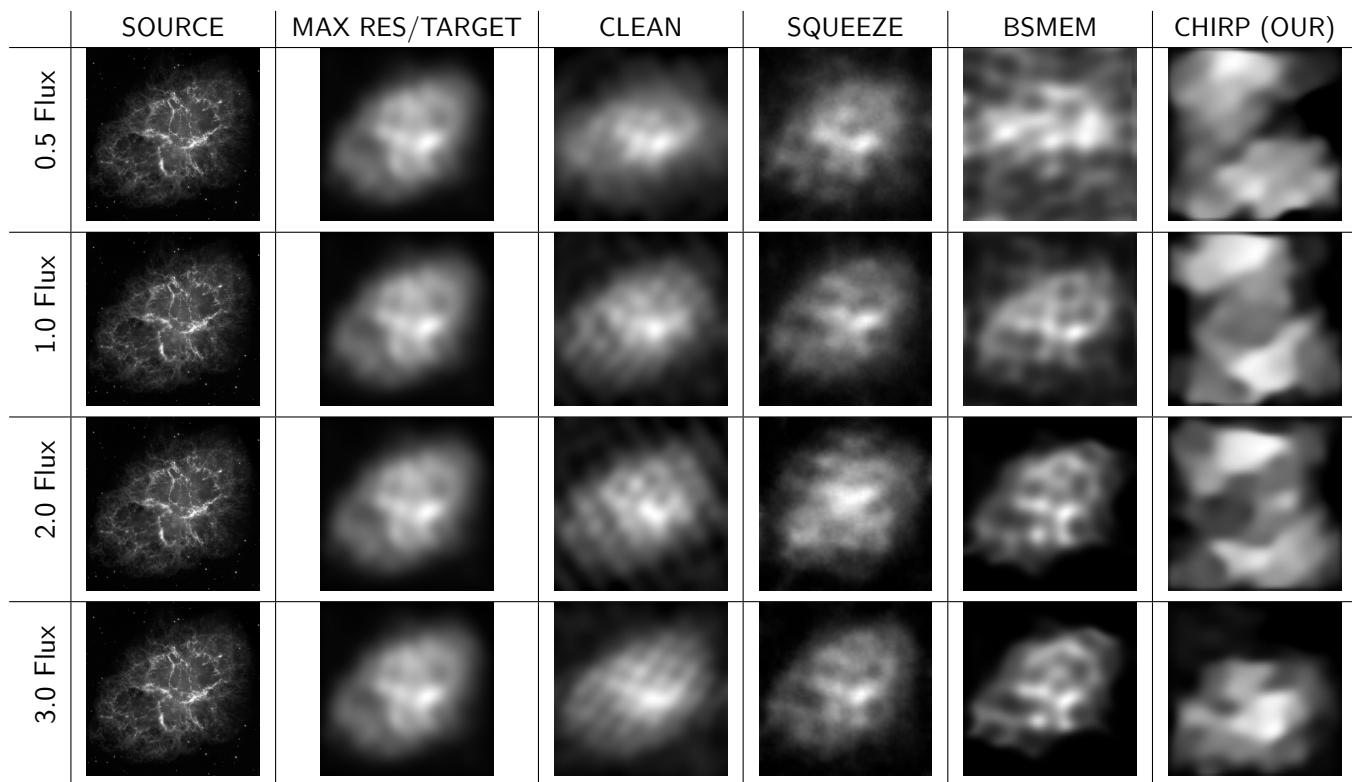
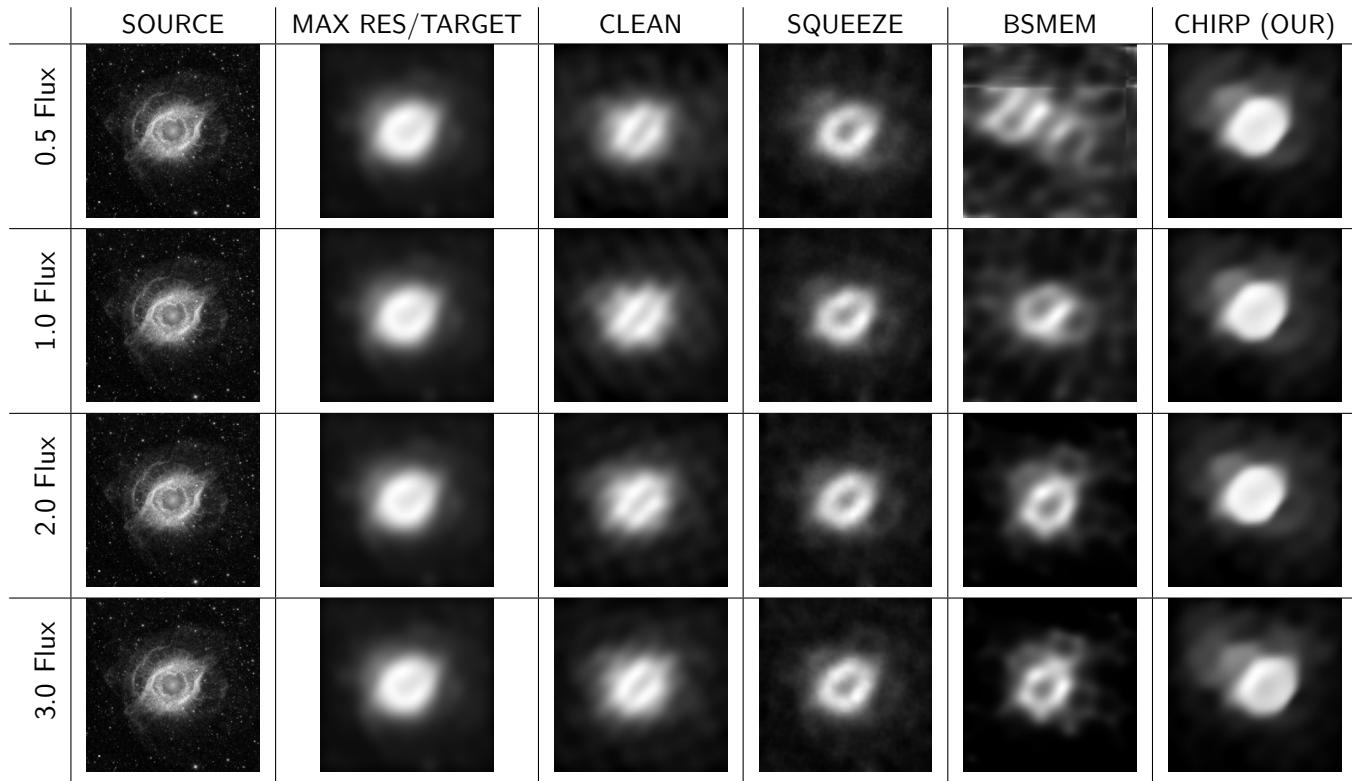


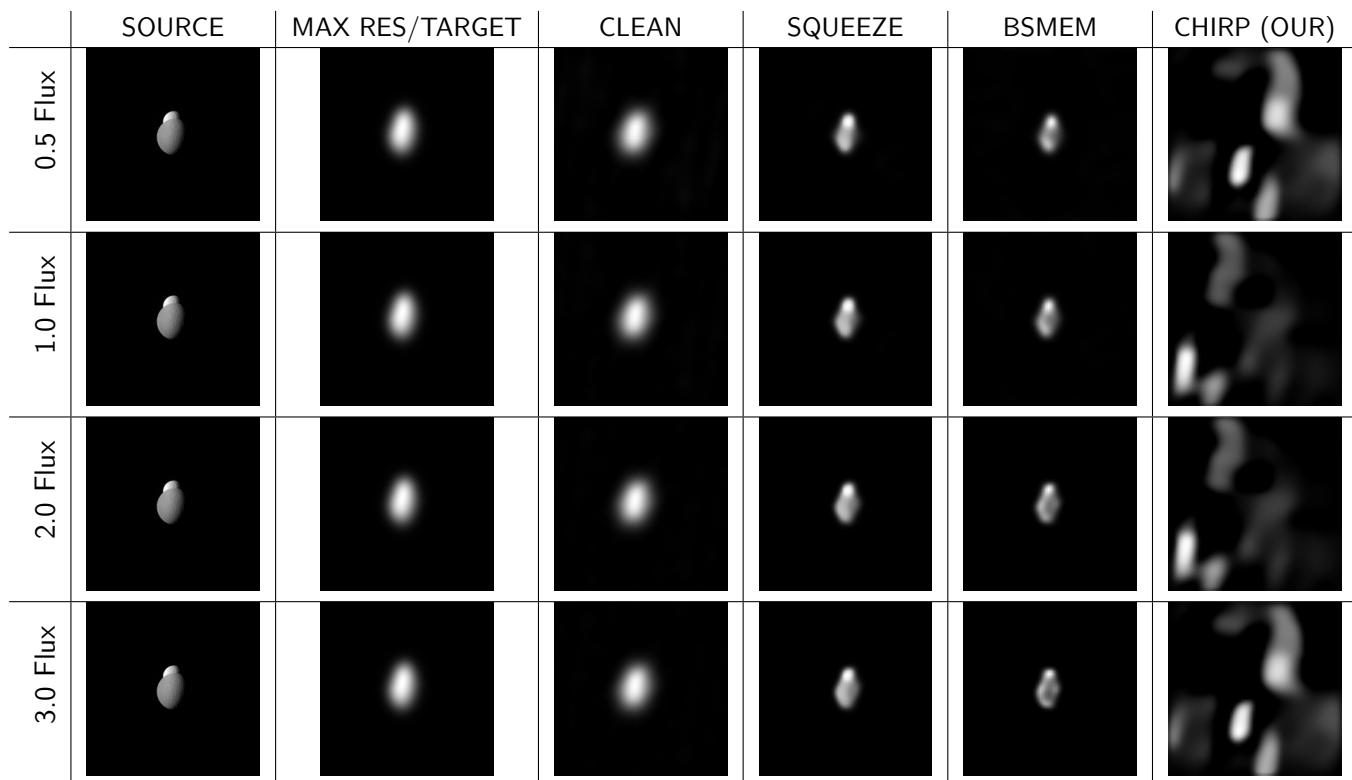
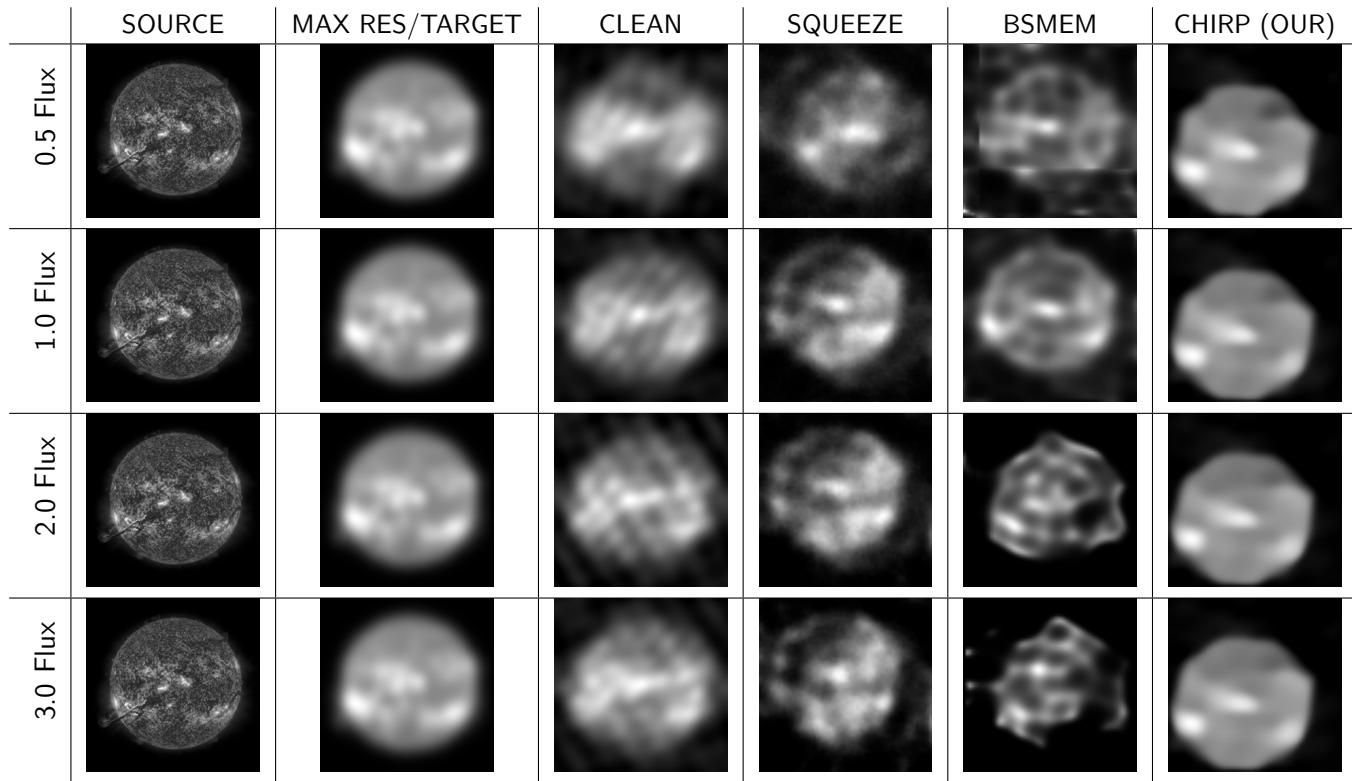


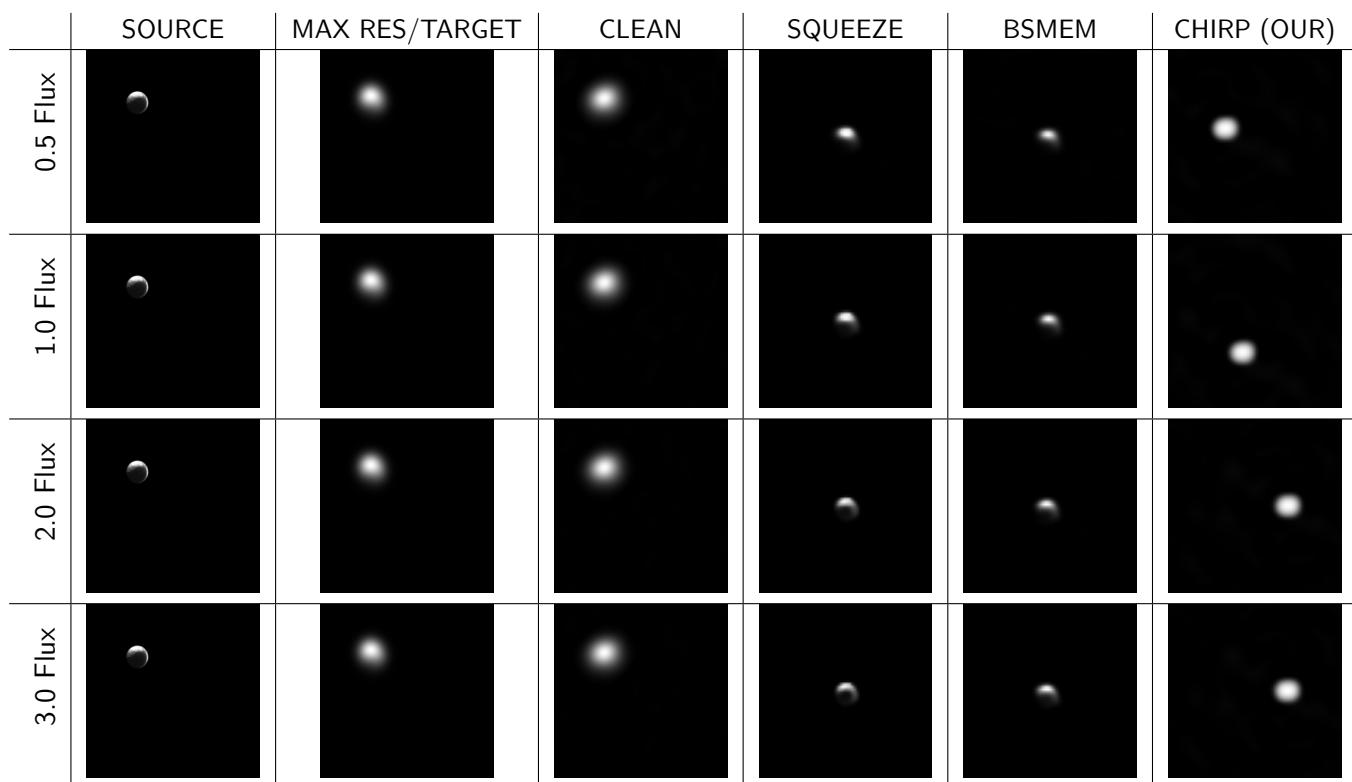
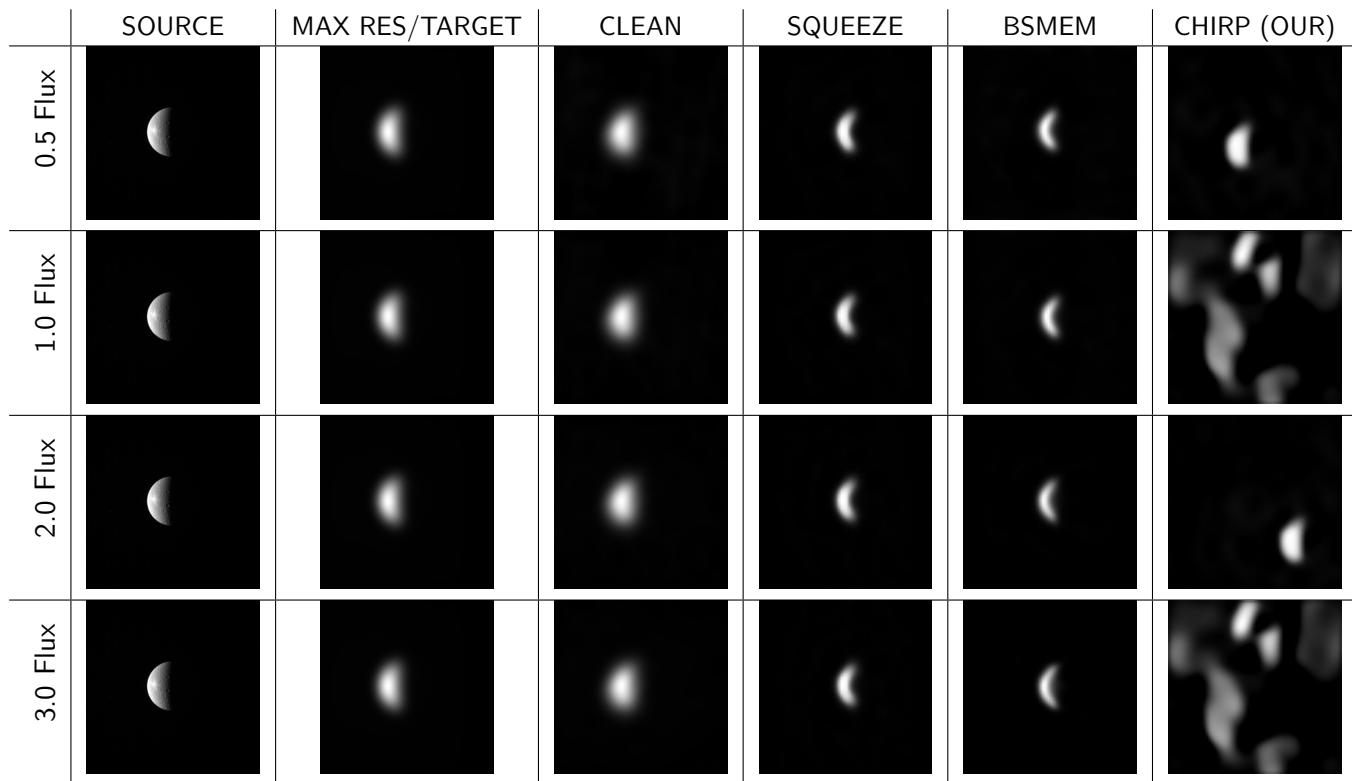


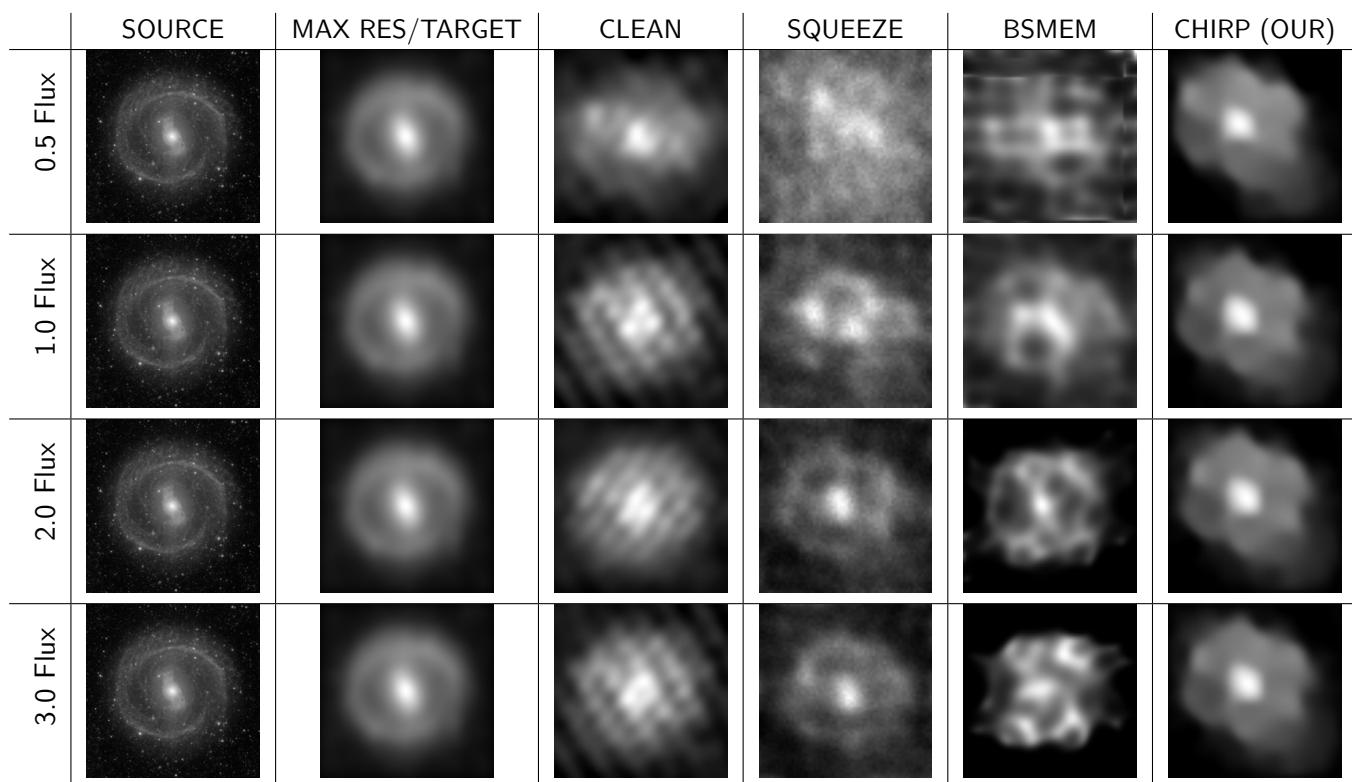
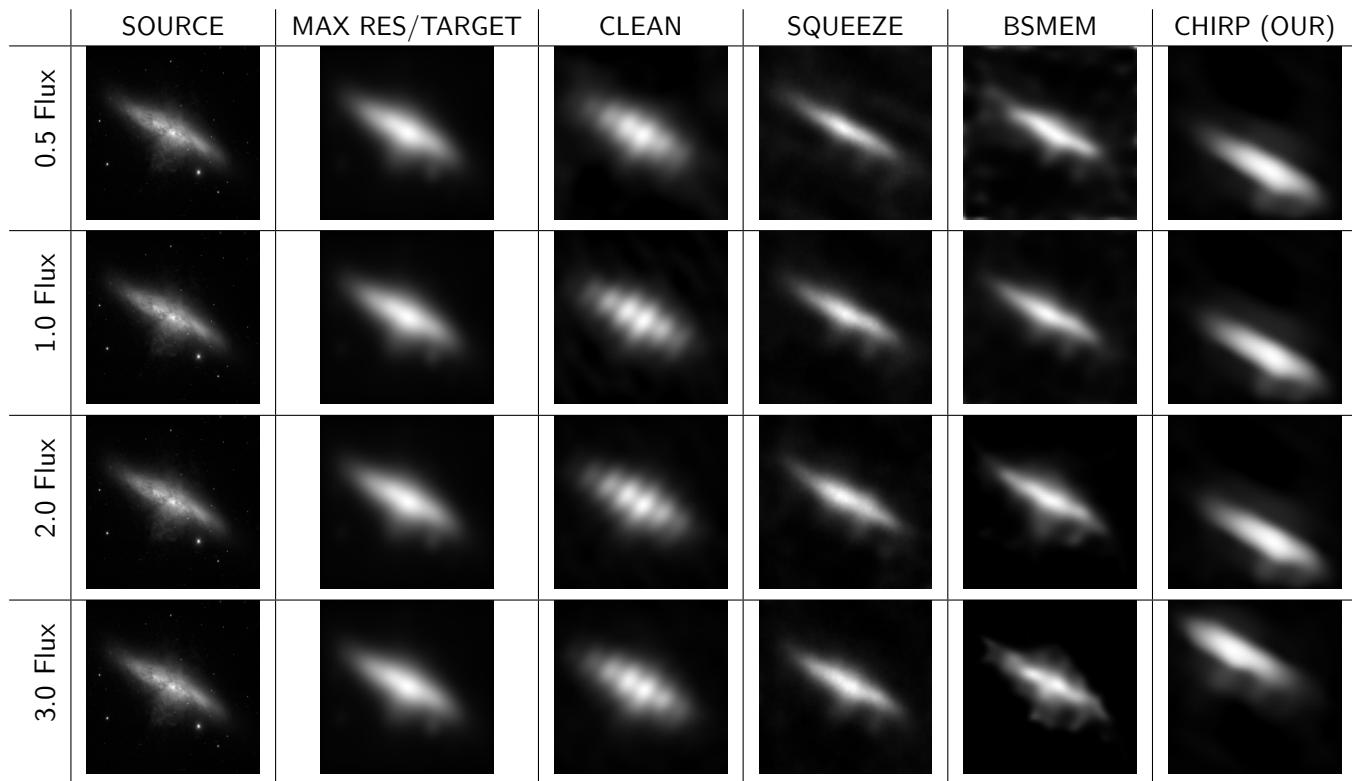


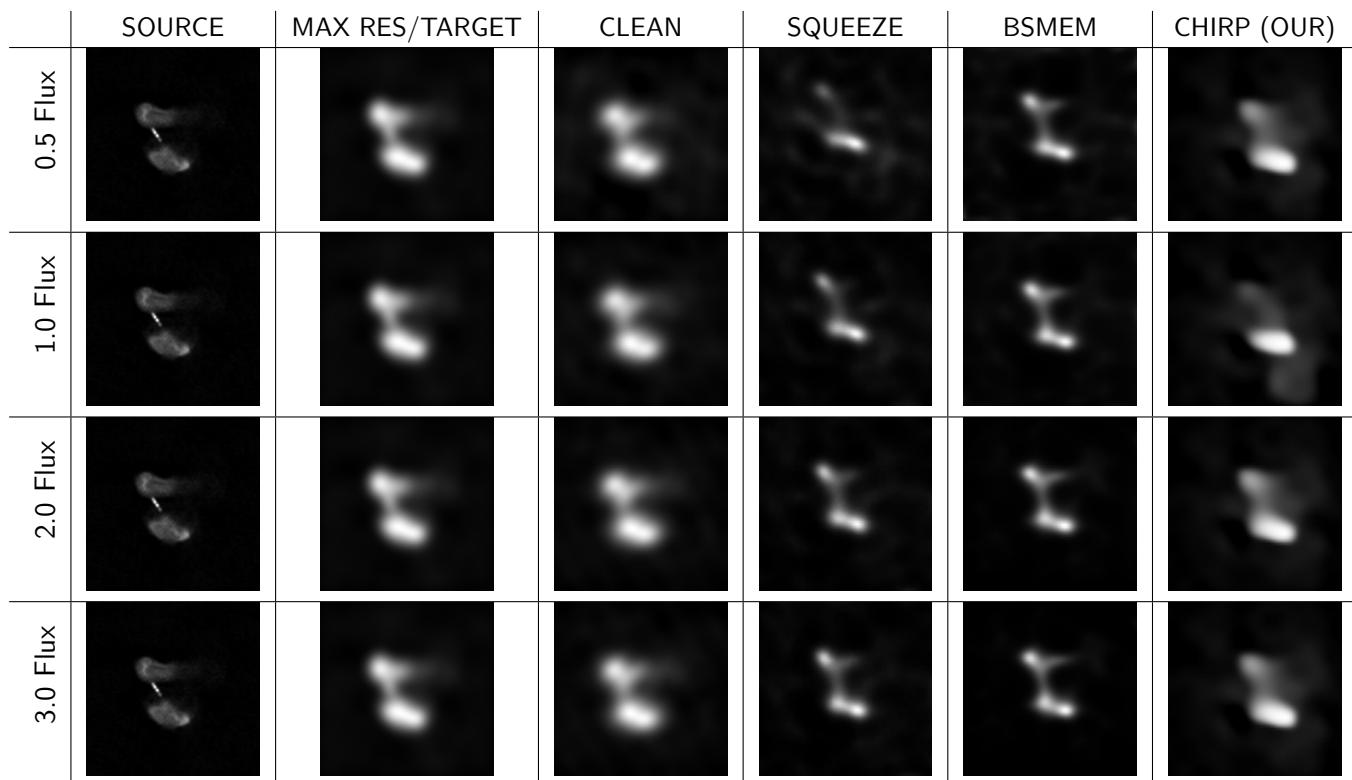
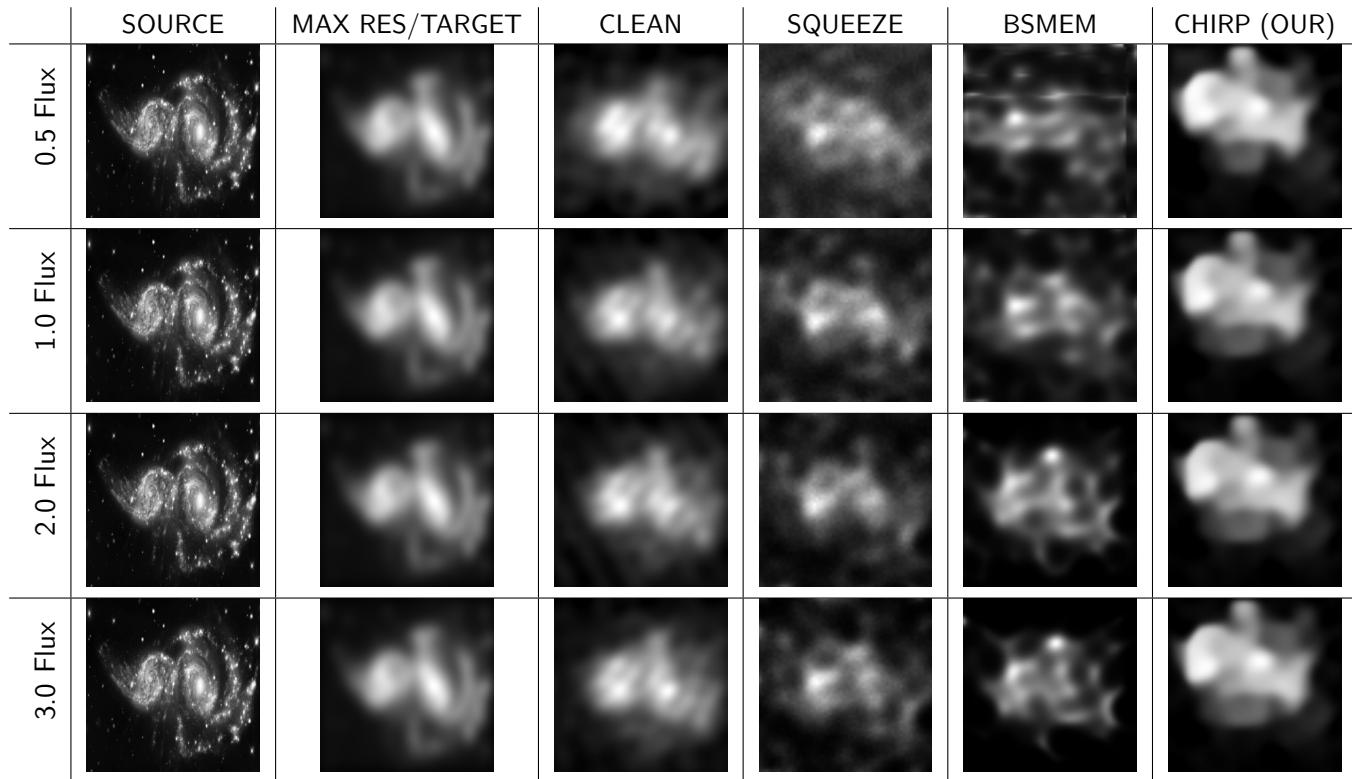


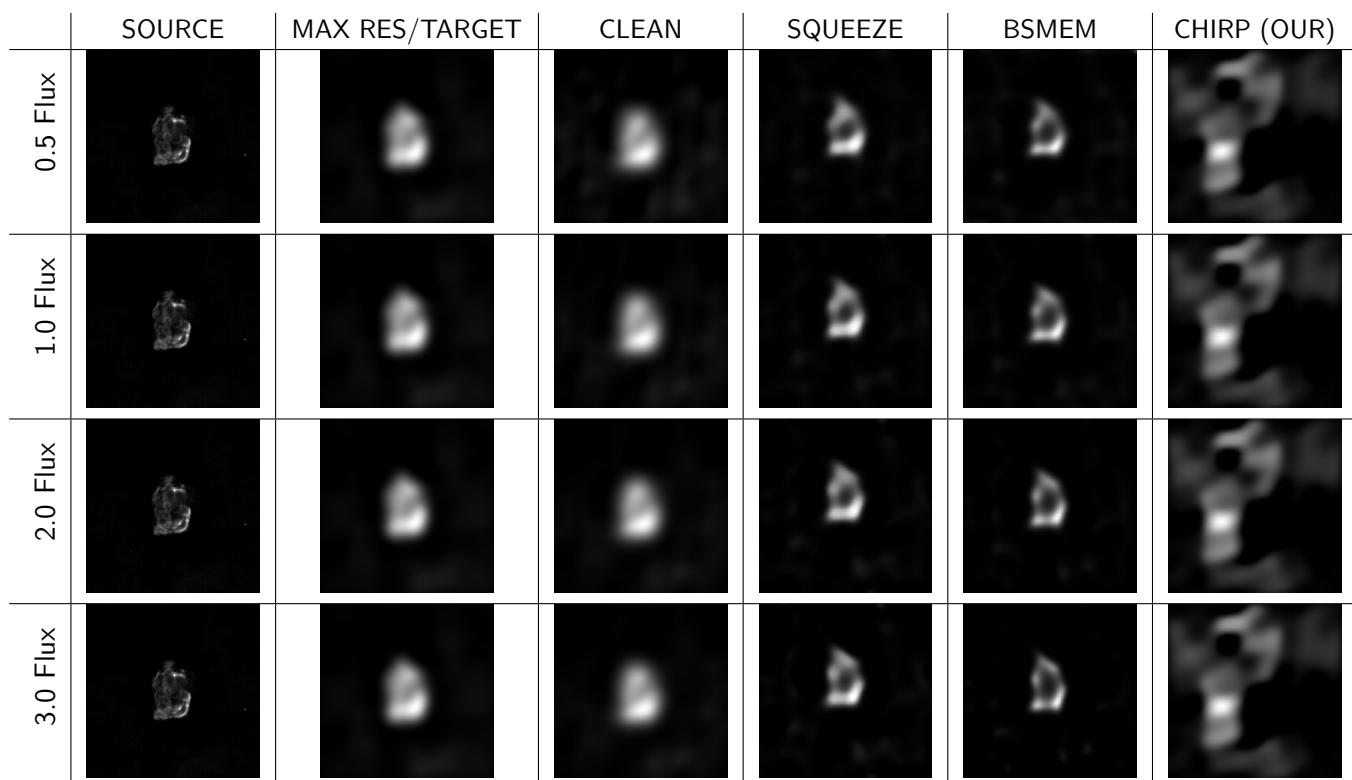
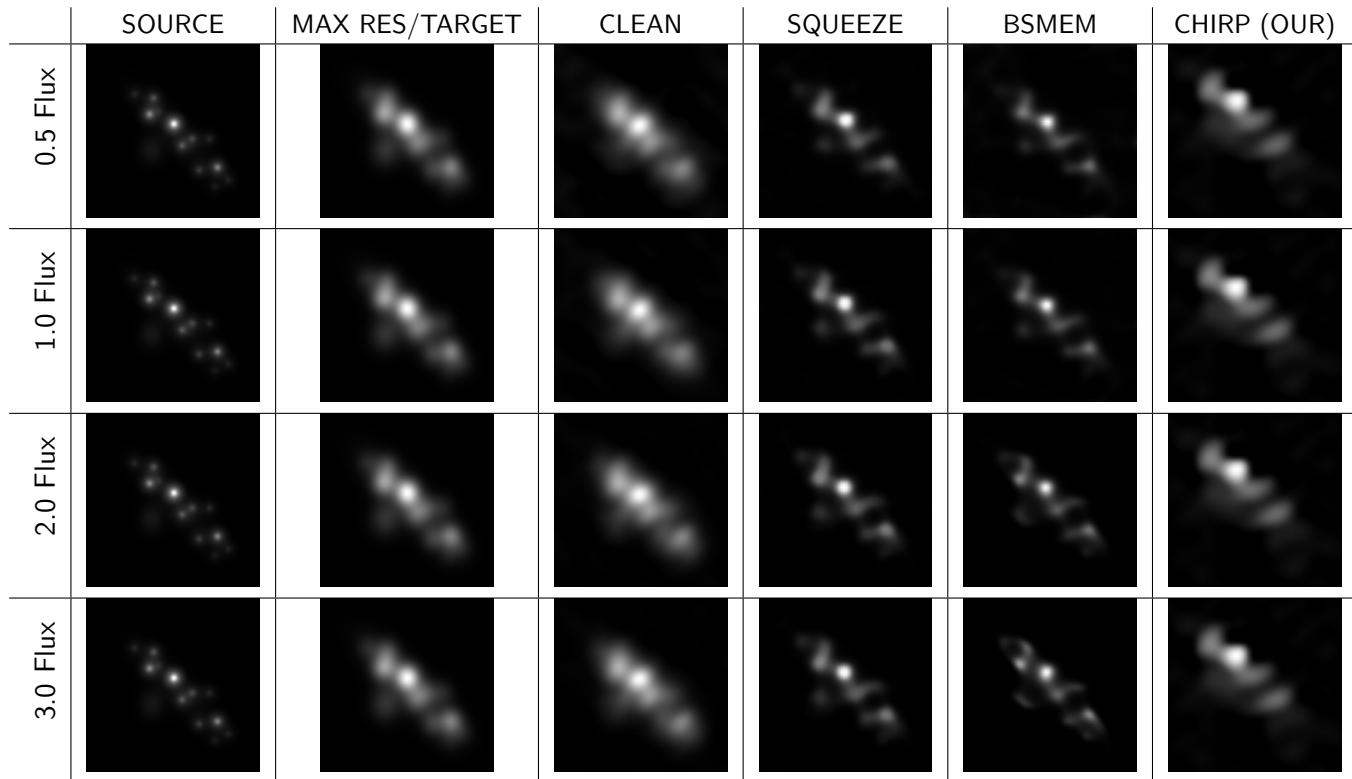


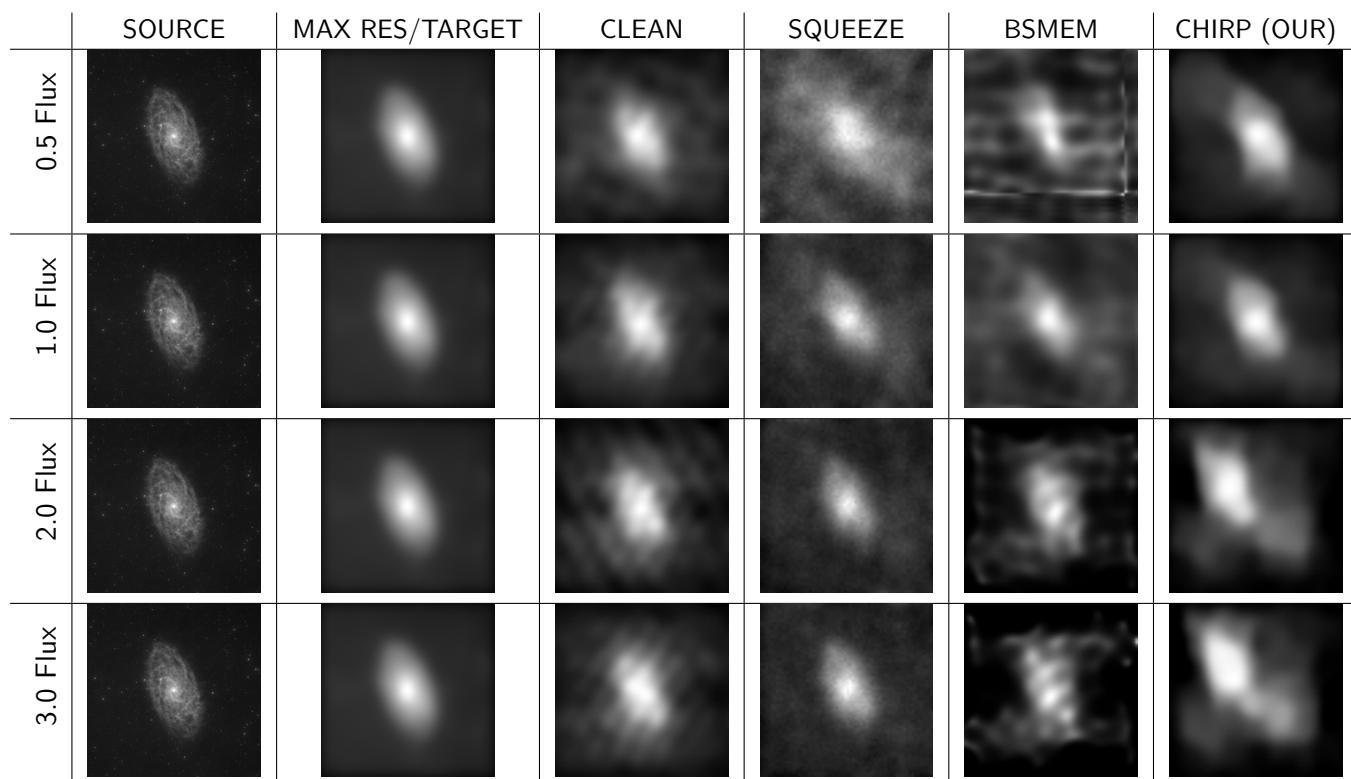
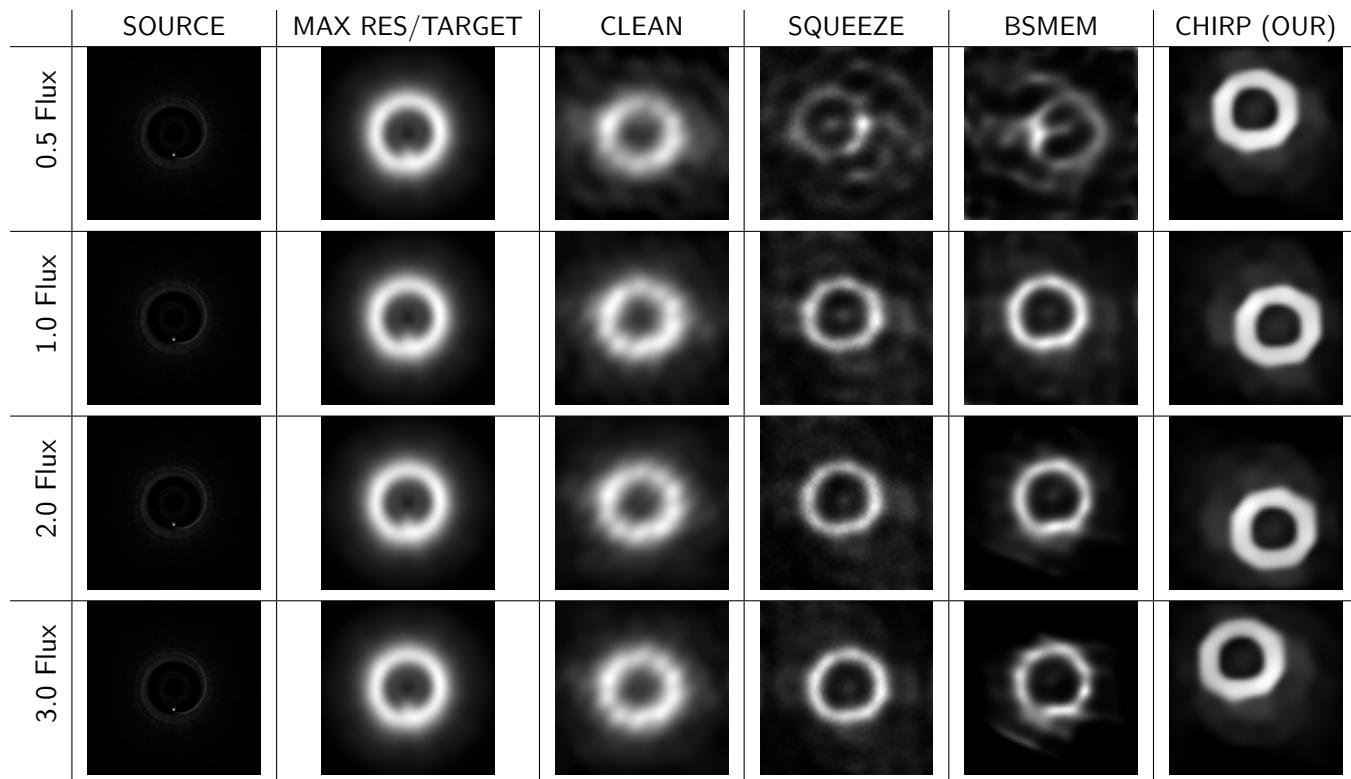


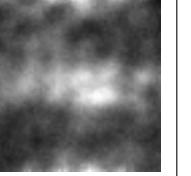
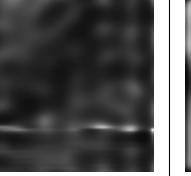
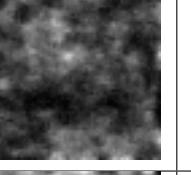
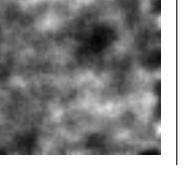


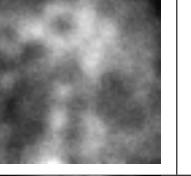
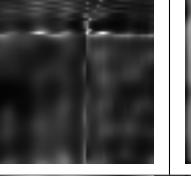
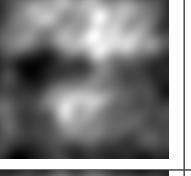
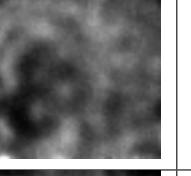
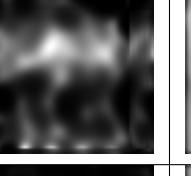
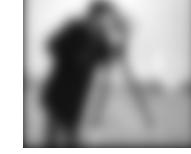


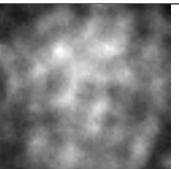
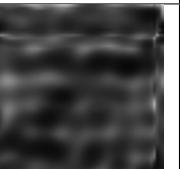
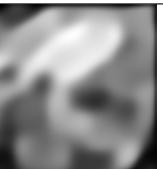
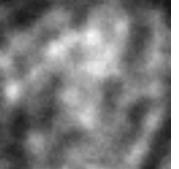
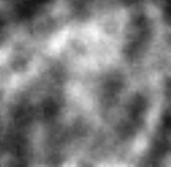
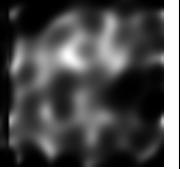
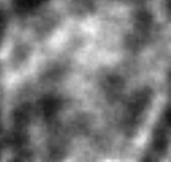
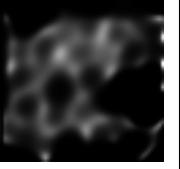


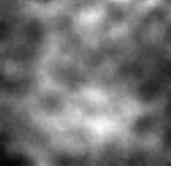
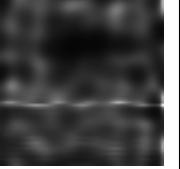
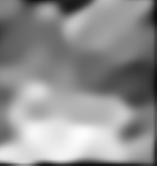
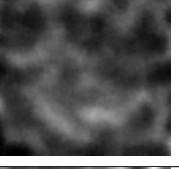
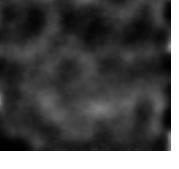


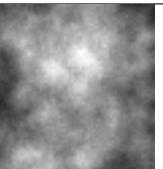
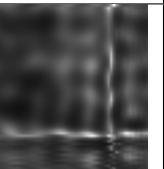
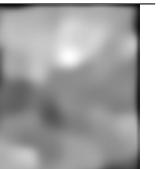
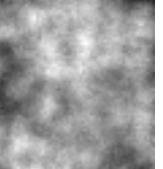
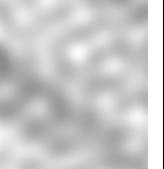
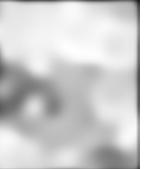
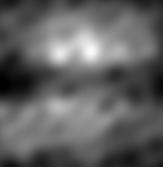
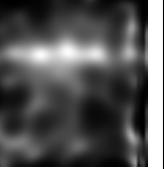
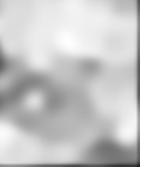
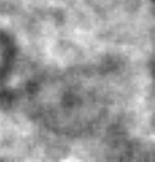
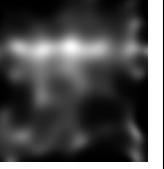


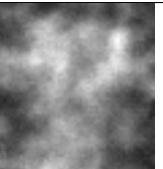
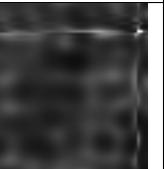
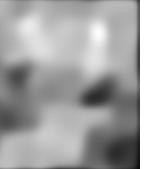
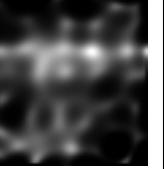
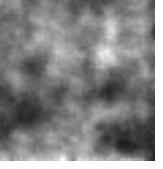
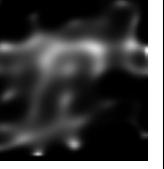
	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						

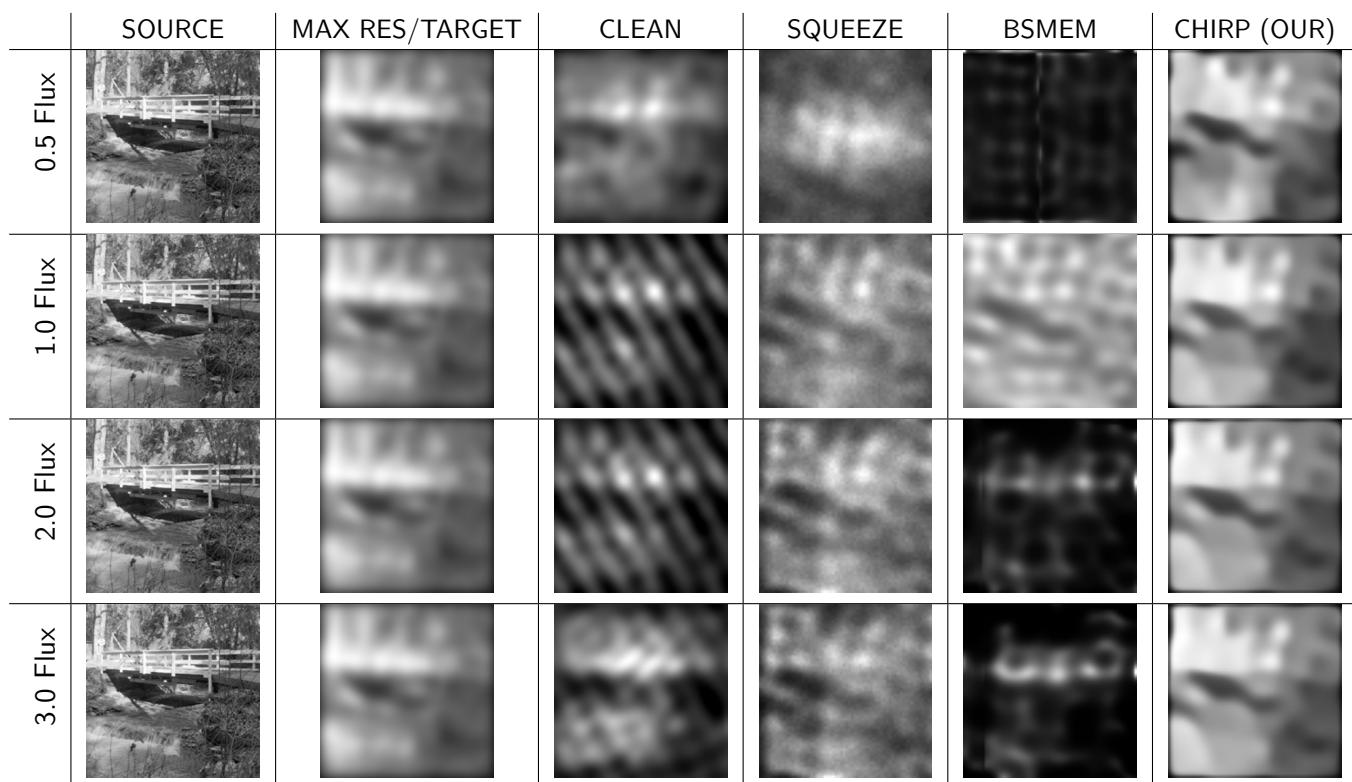
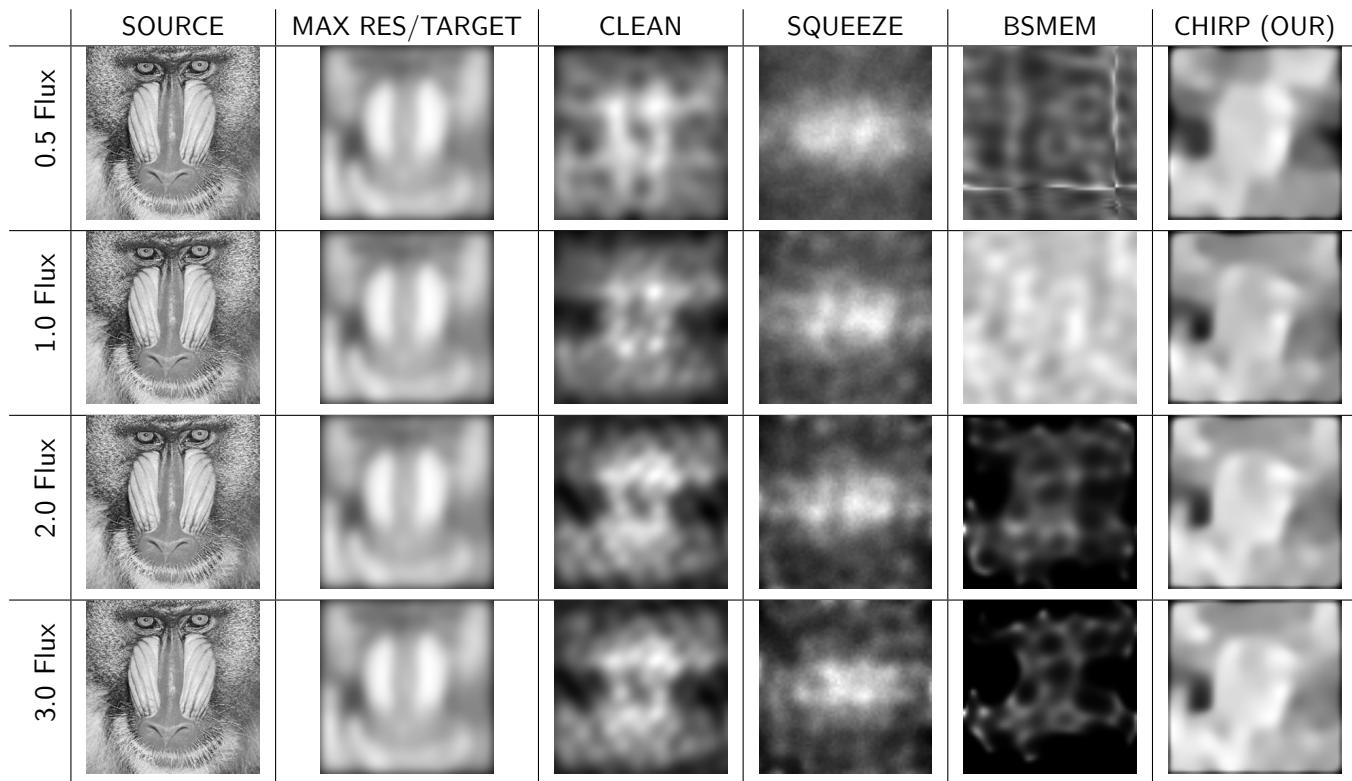
	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						

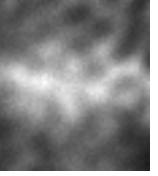
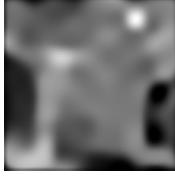
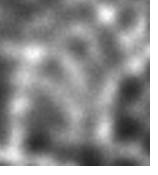
	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						

	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						

	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						

	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						



	SOURCE	MAX RES/TARGET	CLEAN	SQUEEZE	BSMEM	CHIRP (OUR)
0.5 Flux						
1.0 Flux						
2.0 Flux						
3.0 Flux						

1.2 EHT Telescope Parameters

Here we show the parameters corresponding to the telescopes in the EHT array. The locations of the telescopes determine what portions of the uv plane are sampled for a given source. the SEFD provides information about the noise introduced on each visibility.

NAME	ALMA	SMTO	LMT	HAWAII8	PV	PbBI	SPT	GLT	CARMA8
E. LONG	-67:45:11.4	-109:52:19	-97:18:53	-155:28:40.7	-3:23:33.8	05:54:28.5	-000:00:00.0	-38:25:19.1	-118:08:30.3
LAT	-23:01:09.4	32:42:06	18:59:06	19:49:27.4	37:03:58.2	44:38:02.0	-90:00:00	72:35:46.4	37:16:49.6
X-POS	2225037.1851	-1828796.2	-768713.9637	-5464523.4	5088967.9	4523998.4	0	1500692	-2397431.3
Y-POS	-5441199.162	-5054406.8	-5988541.7982	-2493147.08	-301681.6	468045.24	0	-1191735	-4482018.9
Z-POS	-2479303.4629	3427865.2	2063275.9472	2150611.75	3825015.8	4460309.76	-6359587.3	6066409	3843524.5
SEFD	110	11900	560	4900	2900	1600	7300	4744	3500

Table 1: Parameters corresponding to the telescopes in the EHT array: East Longitude, Latitude, X-Y-Z position (meters), and SEFD (System Equivalent Flux Density)

1.3 Blind Test Data

We introduce a blind test set of 20 challenging synthetic measurements. Bispectrum values along with their squared visibilities are generated using a variety of target sources and telescope parameters. The source/telescope parameters that we have used for the 20 test cases can be seen in Table 2. Source locations were taken from actual locations of black holes/blazars (M87, SgA*, 3C273, 3C279, OJ287, BL Lacertae). The EHT array's telescope parameters shown in Table 1 were used for simulation. The uv coverage for each of these source locations using the EHT array can be seen in Figure 1.

	Right Ascension	Declination	FOV (arcsec)	Center Freq (MHz)	Bandwidth (MHz)	Integration Time (s)
Test 1	12:30:49.423382	12:23:28.04366	0.00018382	227297	2048	10
Test 2	17:45:40.041	-29:00:28.118	0.000204595	227297	4096	10
Test 3	12:56:11.2	-05:47:21.5	0.00020264	227297	8192	10
Test 4	08:54:48.8	20:06:30	0.00018646	227297	4096	10
Test 5	22:02:43.2	42:16:40	0.00020280	227297	8192	10
Test 6	08:54:48.8	20:06:30	0.00018646	227297	2048	10
Test 7	22:02:43.2	42:16:40	0.00020280	227297	4096	10
Test 8	12:30:49.423382	12:23:28.04366	0.00018382	227297	4096	10
Test 9	12:29:6.6	02:03:08	0.00020075	227297	8192	10
Test 10	12:30:49.423382	12:23:28.04366	0.00018382	227297	8192	10
Test 11	17:45:40.041	-29:00:28.118	0.000204595	227297	4096	10
Test 12	12:29:6.6	02:03:08	0.00020075	227297	8192	10
Test 13	12:30:49.423382	12:23:28.04366	0.00018382	227297	4096	10
Test 14	22:02:43.2	42:16:40	0.00020280	227297	2048	10
Test 15	12:29:6.6	02:03:08	0.00020075	227297	8192	10
Test 16	08:54:48.8	20:06:30	0.00018646	227297	8192	10
Test 17	17:45:40.041	-29:00:28.118	0.000204595	227297	4096	10
Test 18	12:30:49.423382	12:23:28.04366	0.00018382	227297	8192	10
Test 19	17:45:40.041	-29:00:28.118	0.000204595	227297	8192	10
Test 20	12:30:49.423382	12:23:28.04366	0.00018382	227297	4096	10

Table 2: Parameters used for each of the 20 blind synthetic test data measurements. The FOV center is specified using Right Ascension (HH:MM:SS.SS) and Declination(DD:MM:SS.SS).

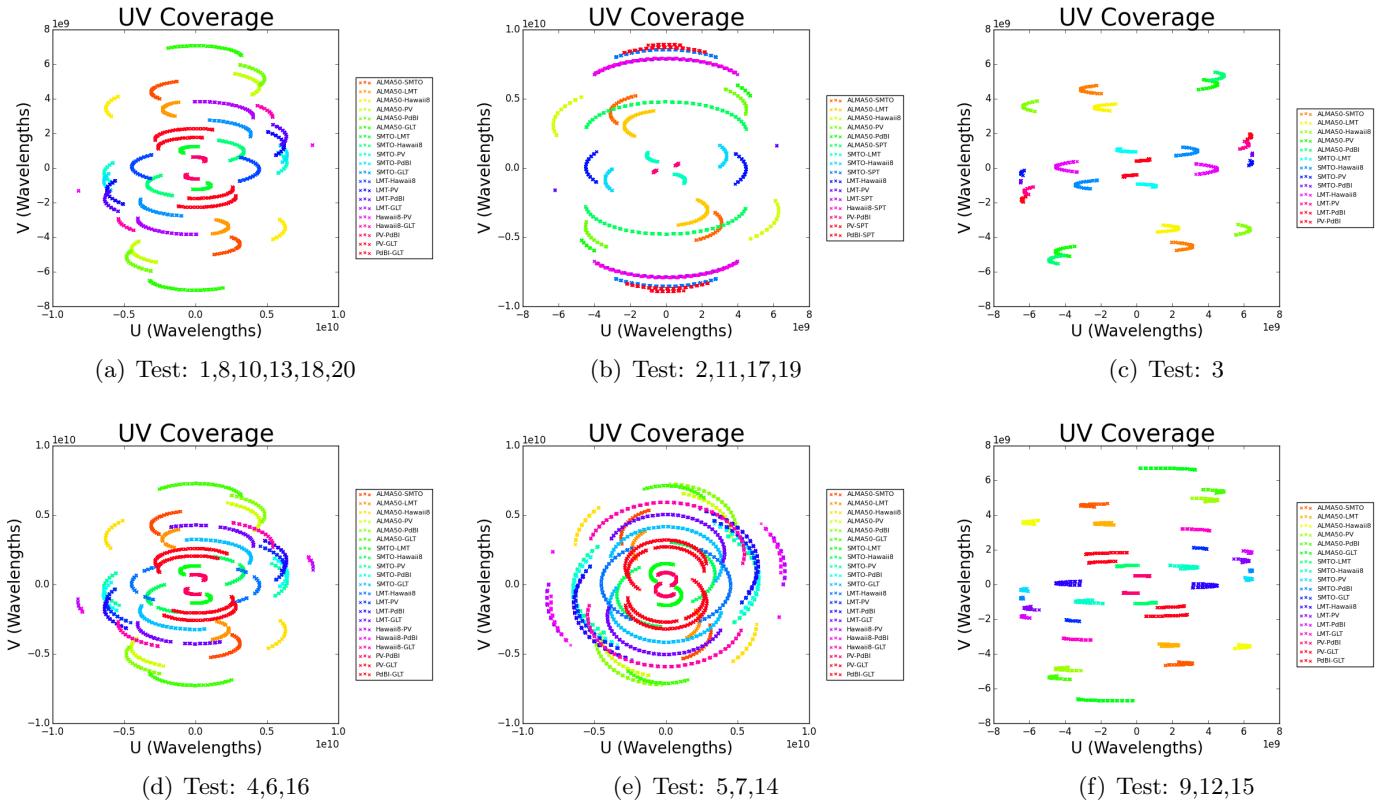


Figure 1: The uv coverage coresponding to the test data provided in Table 2. We have chosen to use a variety of source locations corresponding to very different uv coverage. Reconstruction of images using b, c, and f's uv coverage is very challenging.

2 Energy and Optimization

Following from Equations 9, 11, and 13 in our paper, we seek an approximate MAP reconstruction. We write our energy as having a distribution of Gaussian noise centered around each bispectrum in the complex plane. To optimize this function we use “Half Quadratic Splitting”. This method introduces a set of auxillary patches $\{z^i\}_1^N$, one for each overlapping patch $P_i X$ in the image. Using this technique, the problem we wish to solve is written as:

$$\begin{aligned} \hat{x} &= \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \sum_{i=1}^k \left[\frac{1}{2} \begin{pmatrix} \xi_i^R(x) - Y_i^R \\ \xi_i^I(x) - Y_i^I \end{pmatrix}^T \Sigma_i^{-1} \begin{pmatrix} \xi_i^R(x) - Y_i^R \\ \xi_i^I(x) - Y_i^I \end{pmatrix} \right] \right] \\ &= \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \sum_{i=1}^k \left[\frac{1}{2} \begin{pmatrix} \xi_i^R(x) - Y_i^R \\ \xi_i^I(x) - Y_i^I \end{pmatrix}^T \begin{pmatrix} \sigma_{i1} & \sigma_{i2} \\ \sigma_{i3} & \sigma_{i4} \end{pmatrix} \begin{pmatrix} \xi_i^R(x) - Y_i^R \\ \xi_i^I(x) - Y_i^I \end{pmatrix} \right] \right] \\ &= \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \sum_{i=1}^k \left[\frac{1}{2} \begin{pmatrix} \xi_i^R(x) - Y_i^R \\ \xi_i^I(x) - Y_i^I \end{pmatrix}^T \begin{pmatrix} \sigma_{i1}(\xi_i^R(x) - Y_i^R) + \sigma_{i2}(\xi_i^I(x) - Y_i^I) \\ \sigma_{i3}(\xi_i^R(x) - Y_i^R) + \sigma_{i4}(\xi_i^I(x) - Y_i^I) \end{pmatrix} \right] \right] \\ &= \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \frac{1}{2} \sum_{i=1}^k [\sigma_{i1}(\xi_i^R(x) - Y_i^R)^2 + \sigma_{i4}(\xi_i^I(x) - Y_i^I)^2 + (\sigma_{i2} + \sigma_{i3})(\xi_i^I(x) - Y_i^I)(\xi_i^R(x) - Y_i^R)] \right] \end{aligned} \quad (1)$$

for $\Sigma^{-1} = \begin{pmatrix} \sigma_{i1} & \sigma_{i2} \\ \sigma_{i2} & \sigma_{i3} \end{pmatrix}$, $\sigma_{i2} = \sigma_{i3}$ where

$$\begin{aligned} A_{i_{1,2}} X A_{i_{2,3}} X A_{i_{1,3}} X &= (A_{i_{1,2}}^R + i A_{i_{1,2}}^I) X (A_{i_{2,3}}^R + i A_{i_{2,3}}^I) X (A_{i_{1,3}}^R + i A_{i_{1,3}}^I) X \\ &= (A_{i_{1,2}}^R X + i A_{i_{1,2}}^I X) (A_{i_{2,3}}^R X + i A_{i_{2,3}}^I X) (A_{i_{1,3}}^R X + i A_{i_{1,3}}^I X) \\ &= A_{i_{1,2}}^R X A_{i_{2,3}}^R X A_{i_{1,3}}^R X - A_{i_{1,2}}^R X A_{i_{2,3}}^I X A_{i_{1,3}}^I X - A_{i_{1,2}}^I X A_{i_{2,3}}^R X A_{i_{1,3}}^I X - A_{i_{1,2}}^I X A_{i_{2,3}}^I X A_{i_{1,3}}^R X + \dots \\ &\quad i \left(A_{i_{1,2}}^R X A_{i_{2,3}}^R X A_{i_{1,3}}^I X + A_{i_{1,2}}^R X A_{i_{2,3}}^I X A_{i_{1,3}}^R X + A_{i_{1,2}}^I X A_{i_{2,3}}^R X A_{i_{1,3}}^R X - A_{i_{1,2}}^I X A_{i_{2,3}}^I X A_{i_{1,3}}^I X \right) \\ &= \xi_i^R(X) + i \xi_i^I(X) \end{aligned} \quad (2)$$

This can now be solved using an iterative framework: (1) solving for $\{z^n\}$ given X and (2) solving for X given $\{z^n\}$

(1) solving for $\{z^n\}$ given X : set $\{z^n\}$ to the most likely patch under the prior, given the corrupted measurements $P_n X$ and parameters β .

(2) solving for X given $\{z^n\}$: There are many different ways that we can go about solving this. In the subsections below we explore a few different ways of doing this

2.1 Optimization Method 1 - Using Gradients

We can attempt to solve for X given $\{z^n\}$ using gradient decent. This is guaranteed to find a local minimum of our energy.

$$\begin{aligned}
\frac{d}{dX} E &= \frac{d}{dX} \sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \frac{d}{dX} \frac{1}{2} \sum_{i=1}^k \left[\sigma_{i1} (\xi_i^R(x) - Y_i^R)^2 + \sigma_{i4} (\xi_i^I(x) - Y_i^I)^2 + (\sigma_{i2} + \sigma_{i3}) (\xi_i^I(x) - Y_i^I)(\xi_i^R(x) - Y_i^R) \right] \\
&= \frac{d}{dX} \sum_{n=1}^N \left[\frac{\beta}{2} (P_n X - z^n)^T (P_n X - z^n) - \log p(z^n) \right] + \dots \\
&\quad \frac{1}{2} \sum_{i=1}^k \left[\frac{d}{dX} \sigma_{i1} (\xi_i^R(x) - Y_i^R)^2 + \frac{d}{dX} \sigma_{i4} (\xi_i^I(x) - Y_i^I)^2 + \frac{d}{dX} (\sigma_{i2} + \sigma_{i3}) (\xi_i^I(x) - Y_i^I)(\xi_i^R(x) - Y_i^R) \right] \\
&= \frac{d}{dX} \sum_{n=1}^N \left[\frac{\beta}{2} (X^T P_n^T P_n X - (z^n)^T (P_n X - z^n)) - \log p(z^n) \right] + \dots \\
&\quad \frac{1}{2} \sum_{i=1}^k \left[2\sigma_{i1} (\xi_i^R(x) - Y_i^R) \frac{d\xi_i^R(x)}{dX} + 2\sigma_{i4} (\xi_i^I(x) - Y_i^I) \frac{d\xi_i^I(x)}{dX} + (\sigma_{i2} + \sigma_{i3}) \left[(\xi_i^I(x) - Y_i^I) \frac{d\xi_i^R(x)}{dX} + \frac{d\xi_i^I(x)}{dX} (\xi_i^R(x) - Y_i^R) \right] \right] \\
&= \frac{d}{dX} \sum_{n=1}^N \left[\frac{\beta}{2} (X^T P_n^T P_n X - X^T P_n^T z^n - (z^n)^T P_n X + (z^n)^T z^n) - \log p(z^n) \right] + \dots \\
&\quad \sum_{i=1}^k \left[\sigma_{i1} (\xi_i^R(x) - Y_i^R) \frac{d\xi_i^R(x)}{dX} + \sigma_{i4} (\xi_i^I(x) - Y_i^I) \frac{d\xi_i^I(x)}{dX} + \sigma_{i2} \left[(\xi_i^I(x) - Y_i^I) \frac{d\xi_i^R(x)}{dX} + \frac{d\xi_i^I(x)}{dX} (\xi_i^R(x) - Y_i^R) \right] \right] \\
&= \sum_{n=1}^N \left[\frac{\beta}{2} (2P_n^T P_n X - 2(z^n)^T P_n) \right] + \dots \\
&\quad \sum_{i=1}^k \left[\sigma_{i1} (\xi_i^R(x) - Y_i^R) \frac{d\xi_i^R(x)}{dX} + \sigma_{i4} (\xi_i^I(x) - Y_i^I) \frac{d\xi_i^I(x)}{dX} + \sigma_{i2} \left[(\xi_i^I(x) - Y_i^I) \frac{d\xi_i^R(x)}{dX} + \frac{d\xi_i^I(x)}{dX} (\xi_i^R(x) - Y_i^R) \right] \right]
\end{aligned}$$

(3)

Since we know the gradient of our energy function, conditioned on knowing z^n , we can solve for X using gradient decent. Refer to subsection 2.3 for the derivatives $\frac{d\xi_i^I(x)}{dX}$ and $\frac{d\xi_i^R(x)}{dX}$.

2.2 Optimization Method 2 - Taylor Expansion

We can attempt to solve for X given $\{z^n\}$ by doing a Taylor expansion our energy around the current estimate of X_0 and solving for X in closed-form. In order to solve this equation, we must write each of these terms as a quadratic equation of X .

Terms 1 and 2 For these terms we must approximate $\xi_i^R(X)$ and $\xi_i^I(X)$ as a function linear in X . To do this we linearize the interior of the quadat around a point X_0 to the first order term

$$\begin{aligned}
\xi_i(X_0) + \left(\frac{d\xi_i(x)}{dX}(X_0) \right) (X - X_0) &= \xi_i(X_0) - \left(\frac{d\xi_i}{dX}(X_0) \right) X_0 + \left(\frac{d\xi_i}{dX}(X_0) \right) X \\
&= \beta_i + \alpha_i X
\end{aligned}$$

(4)

Plugging this into its full equation we get a second order equation

$$\begin{aligned}
(\xi_i(X) - Y_i)^2 &\approx (\beta_i + \alpha_i X - Y_i)^2 \\
&= (\beta_i + \alpha_i X - Y_i)^T (\beta_i + \alpha_i X - Y_i)
\end{aligned}$$

(5)

$$= X^T \alpha_i^T \alpha_i X + 2(\beta_i - Y_i)^T \alpha_i X + (\beta_i - Y_i)^T (\beta_i - Y_i)$$

(6)

Term 3 $D(x) = (\xi_i^I(x) - Y_i^I)(\xi_i^R(x) - Y_i^R)$ is the third term. We must Taylor expand this equation around a point X_0 to the second order term to get a function quadratic in X .

$$\begin{aligned}
D(X_0) &+ \left(\frac{dD(x)}{dX}(X_0) \right) (X - X_0) + \frac{1}{2} (X - X_0)^T \left(\frac{dD(x)}{dX^2}(X_0) \right) (X - X_0) \\
&= D(X_0) - \left(\frac{dD(x)}{dX}(X_0) \right) X_0 + \left(\frac{dD(x)}{dX}(X_0) \right) X + \frac{1}{2} X^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X \dots \\
&\quad - \frac{1}{2} X^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X_0 + \frac{1}{2} X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X_0 - \frac{1}{2} X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X \\
&= \left(D(X_0) - \left(\frac{dD(x)}{dX}(X_0) \right) X_0 + \frac{1}{2} X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X_0 \right) \dots \\
&\quad + \left(\left(\frac{dD(x)}{dX}(X_0) \right) - \frac{1}{2} \left(\left(\frac{dD(x)}{dX^2}(X_0) \right) X_0 \right)^T - \frac{1}{2} X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) \right) X + \frac{1}{2} X^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X \\
&= \left(D(X_0) - \left(\frac{dD(x)}{dX}(X_0) \right) X_0 + \frac{1}{2} X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X_0 \right) \dots \\
&\quad + \left(\left(\frac{dD(x)}{dX}(X_0) \right) - X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) \right) X + \frac{1}{2} X^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X \\
&= D_0 + D_1 X + X^T D_2 X
\end{aligned} \tag{7}$$

where

$$\begin{aligned}
D_0 &= D(X_0) - \left(\frac{dD(x)}{dX}(X_0) \right) X_0 + \frac{1}{2} X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) X_0 \\
D_1 &= \left(\frac{dD(x)}{dX}(X_0) \right) - X_0^T \left(\frac{dD(x)}{dX^2}(X_0) \right) \\
D_2 &= \frac{1}{2} \left(\frac{dD(x)}{dX^2}(X_0) \right)
\end{aligned}$$

Since $D(x) = (\xi_i^I(x) - Y_i^I)(\xi_i^R(x) - Y_i^R) = \xi_i^I(x)\xi_i^R(x) + Y_i^I Y_i^R - Y_i^R \xi_i^I(x) - Y_i^I \xi_i^R(x)$ we know that

$$\frac{d}{dX} D = \frac{d\xi_i^I(x)}{dX} \xi_i^R(x) + \xi_i^I(x) \frac{d\xi_i^R(x)}{dX} - Y_i^R \frac{d\xi_i^I(x)}{dX} - Y_i^I \frac{d\xi_i^R(x)}{dX}$$

$$\begin{aligned}
\frac{d}{dX^2} D &= \frac{d}{dX} \left(\frac{d}{dX} D \right) = \frac{d}{dX} \left(\frac{d\xi_i^I(x)}{dX} \xi_i^R(x) + \xi_i^I(x) \frac{d\xi_i^R(x)}{dX} - Y_i^R \frac{d\xi_i^I(x)}{dX} - Y_i^I \frac{d\xi_i^R(x)}{dX} \right) \\
&= \frac{d\xi_i^I(x)}{dX^2} \xi_i^R(x) + \frac{d\xi_i^I(x)}{dX} \frac{d\xi_i^R(x)}{dX}^T + \xi_i^I(x) \frac{d\xi_i^R(x)}{dX^2} + \frac{d\xi_i^R(x)}{dX} \frac{d\xi_i^I(x)}{dX}^T - Y_i^R \frac{d\xi_i^I(x)}{dX^2} - Y_i^I \frac{d\xi_i^R(x)}{dX^2}
\end{aligned} \tag{8}$$

Solution Then, we plug this quadratic form into our optimization

$$\hat{x} = \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (||P_n X - z^n||^2) - \log p(z^n) \right] + \frac{1}{2} \sum_{i=1}^k \left[\sigma_{i1} (\beta_i^R + \alpha_i^R X - Y_i^R)^2 + \sigma_{i4} (\beta_i^I + \alpha_i^I X - Y_i^I)^2 + (\sigma_{i2} + \sigma_{i3}) (D_0 + D_1 X + X^T D_2 X) \right] \right] \tag{9}$$

Now, we can just solve for X

$$\begin{aligned}
\frac{d}{dX} E &= \frac{d}{dX} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \frac{d}{dX} \sum_{i=1}^k \frac{1}{2} [\sigma_{i1}(\beta_i^R + \alpha_i^R X - Y_i^R)^2 + \sigma_{i4}(\beta_i^I + \alpha_i^I X - Y_i^I)^2 + (\sigma_{i2} + \sigma_{i3})(D_0 + D_1 X + X^T D_2 X)] \right] \\
&= \sum_{n=1}^N \left[\frac{\beta}{2} (2P_n^{*T} P_n X - 2(z^n)^{*T} P_n) \right] + \frac{d}{dX} \sum_{i=1}^k \frac{1}{2} [\sigma_{i1}(\beta_i^R + \alpha_i^R X - Y_i^R)^2 + \sigma_{i4}(\beta_i^I + \alpha_i^I X - Y_i^I)^2] \dots \\
&\quad + \frac{d}{dX} \sum_{i=1}^k \frac{1}{2} [(\sigma_{i2} + \sigma_{i3})(D_0 + D_1 X + X^T D_2 X)] \\
&= \sum_{n=1}^N \left[\frac{\beta}{2} (2P_n^{*T} P_n X - 2(z^n)^{*T} P_n) \right] + \frac{d}{dX} \sum_{i=1}^k \frac{\sigma_{i1}}{2} [(\beta_i^R - Y_i^R)^2 + 2(\beta_i^R - Y_i^R)\alpha_R X + X^T \alpha^{RT} \alpha_i^R X] \dots \\
&\quad + \frac{d}{dX} \sum_{i=1}^k \frac{\sigma_{i4}}{2} [(\beta_i^I - Y_i^I)^2 + 2(\beta_i^I - Y_i^I)\alpha_I X + X^T \alpha^{IT} \alpha_i^I X] + \frac{d}{dX} \sum_{i=1}^k \frac{\sigma_{i2} + \sigma_{i3}}{2} [D_0 + D_1 X + X^T D_2 X] \\
&= \sum_{n=1}^N \left[\frac{\beta}{2} (2P_n^{*T} P_n X - 2(z^n)^{*T} P_n) \right] + \sum_{i=1}^k \frac{\sigma_{i1}}{2} [2(\beta_i^R - Y_i^R)\alpha_R + 2\alpha^{RT} \alpha_i^R X] \dots \\
&\quad + \sum_{i=1}^k \frac{\sigma_{i4}}{2} [2(\beta_i^I - Y_i^I)\alpha_I + 2\alpha^{IT} \alpha_i^I X] + \sum_{i=1}^k \frac{\sigma_{i2} + \sigma_{i3}}{2} [D_1 + 2D_2 X] \\
&= \left[\sum_{n=1}^N [\beta P_n^{*T} P_n] + \sum_{i=1}^k \sigma_{i1} [\alpha^{RT} \alpha_i^R] + \sum_{i=1}^k \sigma_{i4} [\alpha^{IT} \alpha_i^I] + \sum_{i=1}^k (\sigma_{i2} + \sigma_{i3}) D_2 \right] X \dots \\
&\quad - \left[\sum_{n=1}^N [\beta(z^n)^{*T} P_n] - \sum_{i=1}^k \sigma_{i1} [(\beta_i^R - Y_i^R)\alpha_R] - \sum_{i=1}^k \sigma_{i4} [(\beta_i^I - Y_i^I)\alpha_I] - \sum_{i=1}^k \frac{\sigma_{i2} + \sigma_{i3}}{2} D_1 \right]
\end{aligned}$$

(10)

Therefore,

$$\begin{aligned}
X &= \left[\sum_{n=1}^N [\beta P_n^{*T} P_n] + \sum_{i=1}^k [\sigma_{i1} \alpha^{RT} \alpha_i^R + \sigma_{i4} \alpha^{IT} \alpha_i^I + (\sigma_{i2} + \sigma_{i3}) D_2] \right]^{-1} \dots \\
&\quad \left[\sum_{n=1}^N [\beta(z^n)^{*T} P_n] - \sum_{i=1}^k [\sigma_{i1} (\beta_i^R - Y_i^R)\alpha_R + \sigma_{i4} (\beta_i^I - Y_i^I)\alpha_I + \frac{\sigma_{i2} + \sigma_{i3}}{2} D_1] \right]
\end{aligned}$$

(11)

$$\begin{aligned}
X &= \left[\sum_{n=1}^N [\beta P_n^{*T} P_n] + \sum_{i=1}^k [\sigma_{i1} \alpha^{RT} \alpha_i^R + \sigma_{i4} \alpha^{IT} \alpha_i^I + (\sigma_{i2} + \sigma_{i3}) D_2] \right]^{-1} \dots \\
&\quad \left[\sum_{n=1}^N [\beta(z^n)^{*T} P_n] - \sum_{i=1}^k [\sigma_{i1} (\beta_i^R - Y_i^R)\alpha_R + \sigma_{i4} (\beta_i^I - Y_i^I)\alpha_I + \frac{\sigma_{i2} + \sigma_{i3}}{2} D_1] \right]
\end{aligned}$$

(12)

2.3 Derivatives

We must find the derivative of ξ_i^R with respect to X

$$\begin{aligned}
\frac{d}{dX} \xi_i^R &= \frac{d}{dX} (A_{i1,2}^R X A_{i2,3}^R X A_{i1,3}^R X - A_{i1,2}^R X A_{i2,3}^I X A_{i1,3}^I X - A_{i1,2}^I X A_{i2,3}^R X A_{i1,3}^I X - A_{i1,2}^I X A_{i2,3}^I X A_{i1,3}^R X) \\
&= \frac{d}{dX} (A_{i1,2}^R X A_{i2,3}^R X A_{i1,3}^R X) - \frac{d}{dX} (A_{i1,2}^R X A_{i2,3}^I X A_{i1,3}^I X) - \frac{d}{dX} (A_{i1,2}^I X A_{i2,3}^R X A_{i1,3}^I X) - \frac{d}{dX} (A_{i1,2}^I X A_{i2,3}^I X A_{i1,3}^R X) \\
&= A_{i1,2}^R A_{i2,3}^R X A_{i1,3}^R X + A_{i2,3}^R A_{i1,2}^R X A_{i1,3}^R X + A_{i1,3}^R A_{i2,3}^R X A_{i1,2}^R X \dots \\
&\quad - (A_{i1,2}^R A_{i2,3}^I X A_{i1,3}^R X + A_{i2,3}^I A_{i1,2}^R X A_{i1,3}^I X + A_{i1,3}^I A_{i2,3}^R X A_{i1,2}^R X) \dots \\
&\quad - (A_{i1,2}^I A_{i2,3}^R X A_{i1,3}^I X + A_{i2,3}^R A_{i1,2}^I X A_{i1,3}^I X + A_{i1,3}^I A_{i2,3}^R X A_{i1,2}^I X) \dots \\
&\quad - (A_{i1,2}^I A_{i2,3}^I X A_{i1,3}^R X + A_{i2,3}^I A_{i1,2}^I X A_{i1,3}^R X + A_{i1,3}^I A_{i2,3}^I X A_{i1,2}^I X)
\end{aligned}$$

Now, we derive the second derivative of ξ_r^R with respect to X .

$$\begin{aligned}
\frac{d}{dX^2} \xi_i^R &= \frac{d}{dX} \left(\frac{d}{dX} \xi_i^R \right) \\
&= \frac{d}{dX} (A_{i_1,2}^R A_{i_2,3}^R X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_2,3}^R A_{i_1,2}^R X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_1,3}^R A_{i_2,3}^R X A_{i_1,2}^R X) \dots \\
&\quad - (\frac{d}{dX} (A_{i_1,2}^R A_{i_2,3}^I X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_2,3}^I A_{i_1,2}^R X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_1,3}^I A_{i_2,3}^I X A_{i_1,2}^R X) \dots \\
&\quad - (\frac{d}{dX} (A_{i_1,2}^I A_{i_2,3}^R X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_2,3}^R A_{i_1,2}^I X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_1,3}^R A_{i_2,3}^R X A_{i_1,2}^I X) \dots \\
&\quad - (\frac{d}{dX} (A_{i_1,2}^I A_{i_2,3}^I X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_2,3}^I A_{i_1,2}^I X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_1,3}^I A_{i_2,3}^I X A_{i_1,2}^I X)) \\
&= A_{i_1,2}^{RT} A_{i_2,3}^R A_{i_1,3}^R X + A_{i_1,2}^{RT} A_{i_2,3}^R A_{i_1,2}^R X + A_{i_2,3}^{RT} A_{i_1,2}^R A_{i_1,3}^R X + A_{i_2,3}^{RT} A_{i_1,2}^R A_{i_1,2}^R X + A_{i_1,3}^{RT} A_{i_1,2}^R A_{i_2,3}^R X \dots \\
&\quad - (A_{i_1,2}^{RT} A_{i_2,3}^I A_{i_1,3}^I X + A_{i_1,2}^{RT} A_{i_2,3}^I A_{i_1,2}^R X + A_{i_2,3}^{RT} A_{i_1,2}^R A_{i_1,3}^I X + A_{i_2,3}^{RT} A_{i_1,2}^R A_{i_1,2}^R X + A_{i_1,3}^{RT} A_{i_1,2}^R A_{i_2,3}^I X) \dots \\
&\quad - (A_{i_1,2}^{IT} A_{i_2,3}^R A_{i_1,3}^I X + A_{i_1,2}^{IT} A_{i_2,3}^R A_{i_1,2}^R X + A_{i_2,3}^{IT} A_{i_1,2}^R A_{i_1,3}^I X + A_{i_2,3}^{IT} A_{i_1,2}^R A_{i_1,2}^R X + A_{i_1,3}^{IT} A_{i_1,2}^R A_{i_2,3}^R X) \dots \\
&\quad - (A_{i_1,2}^{IT} A_{i_2,3}^I A_{i_1,3}^R X + A_{i_1,2}^{IT} A_{i_2,3}^I A_{i_1,2}^R X + A_{i_2,3}^{IT} A_{i_1,2}^I A_{i_1,3}^R X + A_{i_2,3}^{IT} A_{i_1,2}^I A_{i_1,2}^R X + A_{i_1,3}^{IT} A_{i_1,2}^I A_{i_2,3}^R X)
\end{aligned}$$

Similarly, the derivative of ξ_i^I with respect to X is

$$\begin{aligned}
\frac{d}{dX} \xi_i^I &= \frac{d}{dX} \left(A_{i_1,2}^R X A_{i_2,3}^R X A_{i_1,3}^I X + A_{i_1,2}^R X A_{i_2,3}^I X A_{i_1,3}^R X + A_{i_1,2}^I X A_{i_2,3}^R X A_{i_1,3}^R X - A_{i_1,2}^I X A_{i_2,3}^I X A_{i_1,3}^I X \right) \\
&= A_{i_1,2}^R A_{i_2,3}^R X A_{i_1,3}^I X + A_{i_2,3}^R A_{i_1,2}^R X A_{i_1,3}^I X + A_{i_1,3}^I A_{i_2,3}^R X A_{i_1,2}^R X \dots \\
&\quad + A_{i_1,2}^R A_{i_2,3}^I X A_{i_1,3}^R X + A_{i_2,3}^I A_{i_1,2}^R X A_{i_1,3}^R X + A_{i_1,3}^R A_{i_2,3}^I X A_{i_1,2}^R X \dots \\
&\quad + A_{i_1,2}^I A_{i_2,3}^R X A_{i_1,3}^R X + A_{i_2,3}^R A_{i_1,2}^I X A_{i_1,3}^R X + A_{i_1,3}^R A_{i_2,3}^R X A_{i_1,2}^I X \dots \\
&\quad - (A_{i_1,2}^I A_{i_2,3}^I X A_{i_1,3}^I X + A_{i_2,3}^I A_{i_1,2}^I X A_{i_1,3}^I X + A_{i_1,3}^I A_{i_2,3}^I X A_{i_1,2}^I X)
\end{aligned}$$

Now, we derive the second derivative of ξ_r^I with respect to X .

$$\begin{aligned}
\frac{d}{dX^2} \xi_i^I &= \frac{d}{dX} \left(\frac{d}{dX} \xi_i^I \right) \\
&= \frac{d}{dX} (A_{i_1,2}^R A_{i_2,3}^R X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_2,3}^R A_{i_1,2}^R X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_1,3}^I A_{i_2,3}^R X A_{i_1,2}^R X) \dots \\
&\quad + \frac{d}{dX} (A_{i_1,2}^R A_{i_2,3}^I X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_2,3}^I A_{i_1,2}^R X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_1,3}^R A_{i_2,3}^I X A_{i_1,2}^R X) \dots \\
&\quad + \frac{d}{dX} (A_{i_1,2}^I A_{i_2,3}^R X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_2,3}^R A_{i_1,2}^I X A_{i_1,3}^R X) + \frac{d}{dX} (A_{i_1,3}^R A_{i_2,3}^R X A_{i_1,2}^I X) \dots \\
&\quad - (\frac{d}{dX} (A_{i_1,2}^I A_{i_2,3}^I X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_2,3}^I A_{i_1,2}^I X A_{i_1,3}^I X) + \frac{d}{dX} (A_{i_1,3}^I A_{i_2,3}^I X A_{i_1,2}^I X)) \\
&= A_{i_1,2}^{RT} A_{i_2,3}^R A_{i_1,3}^I X + A_{i_1,2}^{RT} A_{i_2,3}^I A_{i_1,3}^R X + A_{i_2,3}^{RT} A_{i_1,2}^I A_{i_1,3}^R X + A_{i_2,3}^{RT} A_{i_1,2}^R A_{i_1,3}^R X + A_{i_1,3}^{RT} A_{i_1,2}^R A_{i_2,3}^R X \dots \\
&\quad + (A_{i_1,2}^{RT} A_{i_2,3}^I A_{i_1,3}^R X + A_{i_1,2}^{RT} A_{i_2,3}^R A_{i_1,3}^I X + A_{i_2,3}^{RT} A_{i_1,2}^I A_{i_1,3}^R X + A_{i_2,3}^{RT} A_{i_1,2}^R A_{i_1,3}^I X + A_{i_1,3}^{RT} A_{i_1,2}^R A_{i_2,3}^I X) \dots \\
&\quad + (A_{i_1,2}^{IT} A_{i_2,3}^R A_{i_1,3}^R X + A_{i_1,2}^{IT} A_{i_2,3}^R A_{i_1,2}^R X + A_{i_2,3}^{IT} A_{i_1,2}^R A_{i_1,3}^R X + A_{i_2,3}^{IT} A_{i_1,2}^R A_{i_1,2}^R X + A_{i_1,3}^{IT} A_{i_1,2}^R A_{i_2,3}^R X) \dots \\
&\quad - (A_{i_1,2}^{IT} A_{i_2,3}^I A_{i_1,3}^I X + A_{i_1,2}^{IT} A_{i_2,3}^I A_{i_1,2}^R X + A_{i_2,3}^{IT} A_{i_1,2}^I A_{i_1,3}^I X + A_{i_2,3}^{IT} A_{i_1,2}^I A_{i_1,2}^R X + A_{i_1,3}^{IT} A_{i_1,2}^I A_{i_2,3}^R X)
\end{aligned}$$

3 Interstellar Scattering Kernel

So far we have assumed that the measurements we obtain from the telescopes give us a noisy measurement of the frequency component of the true image. However, in reality, the frequency component we receive corresponds to an image that has been corrupted by interstellar scattering. Interstellar scattering essentially convolves the true, sharp image with a kernel. From now on we will call this kernel that interstellar scattering kernel $k(l, m)$. Suppose we know the Fourier transform of $k(l, m)$ at every observed frequency $K(u, v)$. Then, our actual measured frequency components are actually given by $K(u_{i_m, n}, v_{i_m, n})A_{i_m, n}X = K_{i_m, n}A_{i_m, n}X$ rather than just $A_{i_m, n}X$. Therefore, an optimization that removes this interstellar scattering can be written as

$$\begin{aligned}\hat{X} &= \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \sum_{i=1}^k \left[\frac{1}{2} \begin{pmatrix} \zeta_i^R(X) - Y_i^R \\ \zeta_i^I(X) - Y_i^I \end{pmatrix}^T \Sigma_i^{-1} \begin{pmatrix} \zeta_i^R(X) - Y_i^R \\ \zeta_i^I(X) - Y_i^I \end{pmatrix} \right] \right] \\ &= \operatorname{argmin}_{X \in \Omega} \left[\sum_{n=1}^N \left[\frac{\beta}{2} (\|P_n X - z^n\|^2) - \log p(z^n) \right] + \frac{1}{2} \sum_{i=1}^k [\sigma_{i1}(\zeta_i^R(X) - Y_i^R)^2 + \sigma_{i4}(\zeta_i^I(X) - Y_i^I)^2 + (\sigma_{i2} + \sigma_{i3})(\zeta_i^I(X) - Y_i^I)(\zeta_i^R(X) - Y_i^R)] \right]\end{aligned}\quad (13)$$

where the triple product is now written with the interstellar scattering kernel as

$$\begin{aligned}K_{i_1, 2} A_{i_1, 2} X K_{i_2, 3} A_{i_2, 3} X K_{i_1, 3} A_{i_1, 3} X &= K_{i_1, 2} K_{i_2, 3} K_{i_3, 1} (A_{i_1, 2}^R + i A_{i_1, 2}^I) X (A_{i_2, 3}^R + i A_{i_2, 3}^I) X (A_{i_1, 3}^R + i A_{i_1, 3}^I) X \\ &= K_{i_1, 2} K_{i_2, 3} K_{i_3, 1} (A_{i_1, 2}^R X + i A_{i_1, 2}^I X) (A_{i_2, 3}^R X + i A_{i_2, 3}^I X) (A_{i_1, 3}^R X + i A_{i_1, 3}^I X) \\ &= K_{i_1, 2} K_{i_2, 3} K_{i_3, 1} (A_{i_1, 2}^R X A_{i_2, 3}^R X A_{i_1, 3}^R X - A_{i_1, 2}^R X A_{i_2, 3}^I X A_{i_1, 3}^I X - A_{i_1, 2}^I X A_{i_2, 3}^R X A_{i_1, 3}^I X - A_{i_1, 2}^I X A_{i_2, 3}^I X A_{i_1, 3}^R X) + \dots \\ &\stackrel{i K_{i_1, 2} K_{i_2, 3} K_{i_3, 1}}{=} (A_{i_1, 2}^R X A_{i_2, 3}^R X A_{i_1, 3}^I X + A_{i_1, 2}^R X A_{i_2, 3}^I X A_{i_1, 3}^R X + A_{i_1, 2}^I X A_{i_2, 3}^R X A_{i_1, 3}^R X - A_{i_1, 2}^I X A_{i_2, 3}^I X A_{i_1, 3}^I X) \\ &= \zeta_i^R(X) + i \zeta_i^I(X)\end{aligned}\quad (14)$$

If we do not include the interstellar scattering kernel into our optimization, then ideally we could deconvolve its effects out in the end. However, this would require the result of the first stage of optimization to be a blurry image. This image would not necessarily score well on our metric for how much it looks like the true image, and as such may not be the best solution in the first stage of optimization. To achieve the best performance, we should include the interstellar scattering kernel into our optimization equation.

3.1 Gaussian Interstellar Scattering Kernel

The interstellar scattering introduced in the path to black hole SgrA* can be estimated by a Gaussian with covariance (radians) [4]

$$\Sigma_k = \begin{pmatrix} 2.00614 \times 10^{-21} & 3.21018 \times 10^{-22} \\ 3.21018 \times 10^{-22} & 5.64104 \times 10^{-22} \end{pmatrix}\quad (15)$$

Therefore, the kernel can be described using the following equation

$$k(x) = \frac{1}{2\pi} |\Sigma|^{-1/2} \exp \left[-\frac{1}{2} (x - \mu)^T \Sigma_k^{-1} (x - \mu) \right]\quad (16)$$

The continuous Fourier Transform of this kernel is given by

$$K(f) = \frac{1}{2\pi} |\Sigma|^{-1/2} \exp \left[-2\pi^2 (f - j\Sigma_k^{-1}\mu)^T \Sigma_k (f - j\Sigma_k^{-1}\mu) \right]\quad (17)$$

If we assume that our kernel is centered at 0 (shouldn't affect results at all), then since $\mu = 0$, $j\Sigma_k^{-1}\mu$ also is 0. Therefore,

$$K(f) = \frac{1}{2\pi} |\Sigma|^{-1/2} \exp \left[-2\pi^2 f^T \Sigma_k f \right]\quad (18)$$

where $f = (u, v)$. This equation can be used to scale each of the extracted bispectrum terms accordingly (as is done in Equation 14).

4 Noise

4.1 Visibility Noise

The noise model that is assumed for radio interferometry is given in Cornell and Wilkinson [5] as:

$$V_{jk} = \hat{V}_{jk}(1 + a_j)(1 + a_k) \exp[i(\phi_j - \phi_k)] + \epsilon_{jk} \quad (19)$$

For the true visibility \hat{V}_{jk} , zero-mean Gaussian amplitude error, a_j , phase error, ϕ_j , and thermal noise, ϵ_{jk} . The thermal noise, ϵ_{jk} , should be zero-mean and isotropic (circular) because the real and imaginary part of the visibility are correlated separately. So, we add Gaussian noise to both the real and imaginary part of the complex visibility. We ignore amplitude noise and only focus on phase and thermal noise:

$$V_{jk} = \hat{V}_{jk} \exp[i(\phi_j - \phi_k)] + \epsilon_{jk} \quad (20)$$

4.2 Bispectrum Noise

The bispectrum is the triple product of 3 visibilities. Let's take a look at the noise on one bispectrum:

$$V_{12}V_{23}V_{31} = (\hat{V}_{12} \exp[i(\phi_1 - \phi_2)] + \epsilon_{12}) (\hat{V}_{23} \exp[i(\phi_2 - \phi_3)] + \epsilon_{23}) (\hat{V}_{31} \exp[i(\phi_3 - \phi_1)] + \epsilon_{31}) \quad (21)$$

$$\begin{aligned} &= \hat{V}_{12}\hat{V}_{23}\hat{V}_{31} + \epsilon_{12}\hat{V}_{23}\exp[i(\phi_2 - \phi_3)]\hat{V}_{31}\exp[i(\phi_3 - \phi_1)] \dots \\ &\quad + \epsilon_{23}\hat{V}_{12}\exp[i(\phi_1 - \phi_2)]\hat{V}_{31}\exp[i(\phi_3 - \phi_1)] + \epsilon_{31}\hat{V}_{23}\exp[i(\phi_2 - \phi_3)]\hat{V}_{12}\exp[i(\phi_1 - \phi_2)] \dots \\ &\quad + \epsilon_{12}\epsilon_{23}\hat{V}_{31}\exp[i(\phi_3 - \phi_1)] + \epsilon_{23}\epsilon_{31}\hat{V}_{12}\exp[i(\phi_1 - \phi_2)] + \epsilon_{12}\epsilon_{31}\hat{V}_{23}\exp[i(\phi_2 - \phi_3)] \dots \\ &\quad + \epsilon_{12}\epsilon_{23}\epsilon_{31} \end{aligned} \quad (22)$$

$$\begin{aligned} &= \hat{V}_{12}\hat{V}_{23}\hat{V}_{31} + \epsilon_{12}\hat{V}_{23}\hat{V}_{31}\exp[i(\phi_2 - \phi_1)] + \epsilon_{23}\hat{V}_{12}\hat{V}_{31}\exp[i(\phi_3 - \phi_2)] + \epsilon_{31}\hat{V}_{23}\hat{V}_{12}\exp[i(\phi_1 - \phi_3)] \dots \\ &\quad + \epsilon_{12}\epsilon_{23}\hat{V}_{31}\exp[i(\phi_3 - \phi_1)] + \epsilon_{23}\epsilon_{31}\hat{V}_{12}\exp[i(\phi_1 - \phi_2)] + \epsilon_{12}\epsilon_{31}\hat{V}_{23}\exp[i(\phi_2 - \phi_3)] + \epsilon_{12}\epsilon_{23}\epsilon_{31} \end{aligned} \quad (23)$$

Like we expect, by multiplying the visibilities we remove the effect of phase noise on $\hat{V}_{12}\hat{V}_{23}\hat{V}_{31}$.

4.3 Gaussian Noise Model for the Bispectrum Including Amplitude Error

Assume that the noise model for a visibility is given by:

$$\begin{aligned} \Gamma'_{12} &= [(\Gamma_{12}^R + i\Gamma_{12}^I) + (n_{12}^R + in_{12}^I)] \sqrt{(\mu'_1 + \epsilon'_1)} \sqrt{(\mu'_2 + \epsilon'_2)} \\ &\approx [(\Gamma_{12}^R + i\Gamma_{12}^I) + (n_{12}^R + in_{12}^I)] (\mu_1 + \epsilon_1)(\mu_2 + \epsilon_2) \end{aligned} \quad (24)$$

Therefore, the noise on the amplitude of the visibility, and thus the amplitude of the bispectrum is given by:

$$\begin{aligned} |\Gamma_{12}|' &= [|\Gamma_{12}| + n_{12}] (\mu_1 + \epsilon_1)(\mu_2 + \epsilon_2) \\ |\Gamma_{12}|' &= V_{12}A_1A_2 \\ |\Gamma_{12}\Gamma_{23}\Gamma_{31}|' &= |\Gamma_{12}|'|\Gamma_{23}|'|\Gamma_{31}|' = V_{12}V_{23}V_{31}A_1^2A_2^2A_3^2 \end{aligned}$$

where

$$\begin{aligned} V_{ij} &= \mathcal{N}(\Gamma_{ij}, \sigma_{n_{ij}}) \\ A_i &= \mathcal{N}(\mu_i, \sigma_{\epsilon_i}) \end{aligned} \quad (25)$$

We know that $Var[X] = E[X^2] - E[X]^2$ and for Gaussian random variables the moments are:

$$\begin{aligned}
E[X] &= \mu \\
E[X^2] &= \mu^2 + \sigma^2 \\
E[X^3] &= \mu^3 + 3\mu\sigma^2 \\
E[X^4] &= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4
\end{aligned} \tag{26}$$

Therefore,

$$\begin{aligned}
Var[|\Gamma_{12}|'|\Gamma_{23}|'|\Gamma_{31}|'] &= E[|\Gamma_{12}|'^2|\Gamma_{23}|'^2|\Gamma_{31}|'^2] - E[|\Gamma_{12}|'|\Gamma_{23}|'|\Gamma_{31}|']^2 \\
&= E[V_{12}^2 V_{23}^2 V_{31}^2 A_1^4 A_2^4 A_3^4] - E[V_{12} V_{23} V_{31} A_1^2 A_2^2 A_3^2]^2 \\
&= E[V_{12}^2] E[V_{23}^2] E[V_{31}^2] E[A_1^4] E[A_2^4] E[A_3^4] - E[V_{12}] E[V_{23}] E[V_{31}] E[A_1^2] E[A_2^2] E[A_3^2]^2 \\
&= (|\Gamma_{12}|^2 + \sigma_{n_{12}}^2)(|\Gamma_{23}|^2 + \sigma_{n_{23}}^2)(|\Gamma_{31}|^2 + \sigma_{n_{31}}^2)(\mu_1^4 + 6\mu_1^2\sigma_{\epsilon_1}^2 + 3\sigma_{\epsilon_1}^4)(\mu_2^4 + 6\mu_2^2\sigma_{\epsilon_2}^2 + 3\sigma_{\epsilon_2}^4)(\mu_3^4 + 6\mu_3^2\sigma_{\epsilon_3}^2 + 3\sigma_{\epsilon_3}^4) \\
&\quad - |\Gamma_{12}|^2 |\Gamma_{23}|^2 |\Gamma_{31}|^2 (\mu_1^2 + \sigma_{\epsilon_1}^2)^2 (\mu_2^2 + \sigma_{\epsilon_2}^2)^2 (\mu_3^2 + \sigma_{\epsilon_3}^2)^2
\end{aligned} \tag{27}$$

Gain Error Distribution Approximation Given we know the gain error parameters, μ' and ϵ' , how should we estimate μ and ϵ ? Find these parameters such that

$$\begin{aligned}
X &\sim \mathcal{N}(\mu', \sigma_{\epsilon'}^2) \\
Y &\sim \mathcal{N}(\mu, \sigma_{\epsilon}^2) \\
X &\approx Y^2
\end{aligned} \tag{28}$$

If $Y = \frac{1}{\sqrt{\mu'}}(\mu' + \epsilon'/2)$ then $X = \frac{1}{\mu'}(\mu' + \epsilon'/2)^2 = \mu' + \epsilon' + \frac{\epsilon'^2}{4\mu'}$. If $\epsilon'^2 \ll \mu'$, then $X \approx \mu' + \epsilon'$. This implies that if

$$\begin{aligned}
X &\sim \mathcal{N}(\mu', \sigma_{\epsilon'}^2) \\
Y &\sim \mathcal{N}(\sqrt{\mu'}, \frac{1}{4\mu'}\sigma_{\epsilon'}^2)
\end{aligned} \tag{29}$$

References

- [1] Baron, F., Monnier, J.D., Kloppenborg, B.: A novel image reconstruction software for optical/infrared interferometry. In: SPIE Astronomical Telescopes+ Instrumentation, International Society for Optics and Photonics (2010) 77342I–77342I
- [2] Buscher, D.: Direct maximum-entropy image reconstruction from the bispectrum. In: Very High Angular Resolution Imaging. Springer (1994) 91–93
- [3] Jaeger, S.: The common astronomy software application (casa). In: Astronomical Data Analysis Software and Systems XVII. Volume 394. (2008) 623
- [4] Bower, G.C., Goss, W., Falcke, H., Backer, D.C., Lithwick, Y.: The intrinsic size of sagittarius a* from 0.35 to 6 cm. *The Astrophysical Journal Letters* **648**(2) (2006) L127
- [5] Cornwell, T., Wilkinson, P.: A new method for making maps with unstable radio interferometers. *Monthly Notices of the Royal Astronomical Society* **196**(4) (1981) 1067–1086