MEMoirs

Rick Pernak[1,2]

[1]*Solar Physics Laboratory (Code 671), Heliophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771*

[2]*Institute for Astrophysics and Computational Sciences, The Catholic University of America, Washington, DC 20064*

1. Introduction

My first project at Goddard Space Flight Center (GSFC) involved testing a new Maximum Entropy Method (MEM) with RHESSI data. The algorithm was developed by Su-Chan Bong – in his thesis, he describes the Spatio-Spectro Maximum Entropy Method (SSMEM) which works for both imaging *and* spectroscopy – and Jeongwoo Lee, who adapted the technique for RHESSI imaging and named the altered code MEM_NJIT. Schmahl et al. (2007) showed how this imaging algorithm, which was developed for data from the Owens Valley Solar Array (OVSA), performed with RHESSI visibilities.

I'm too lazy to reference papers correctly in something that I have no intention of publishing. From now on, I'll refer to the Narayan and Nityananda (1986) paper entitled,"Maximum Entropy Image Restoration in Astronomy," as NN, and the Cornwell and Evans (1984, "A Simple Maximum Entropy Deconvolution Algorithm") as CE. The Bong et al. (2004, ApJ and JKAS) papers are also useful for understanding the SSMEM program, but I do not refer to them beyond this intro. Bong's SSMEM is an extension of the CE MEM.

Though Ed Schmahl and I tested MEM_NJIT and confirmed its viability with RHESSI data in Schmahl et al. 2007 (leading to its implementation into the RHESSI software), there were some lingering issues that proved to be a cause for concern for some on the team. I've diagnosed a couple of problems and devoted a section to each in these MEMoirs in an effort to explain why they are happening and what, if anything, can be done to address them.

2. Input Flux

As NN allude to in their paper, MEM is "supplemented by prior knowledge about the image" which, in their case, is positivity of intensity (pixel fluxes greater than zero). CE take it one step further by adding a constraint to their objective function:

$$F = \sum_i b_i = F_{observed}, \tag{1}$$

where we just sum over all pixels, so the flux in the map should be consistent with the flux we observe. $F_{observed}$ is obtained from the data and used as an input into MEM_NJIT.

Originally, we set the input flux to the maximum amplitude of the input visibilities after (arbitrarily) eliminating outliers. This method of flux determination is automatically achieved in the RHESSI code. Resulting maps seem to have a lot of low level ($\sim$ 1%) flux distributions (one could say extended sources), and since all pixels in MEM maps are positive, the flux when integrating over the entire map is relatively large, though consistent with the input flux. By "relatively large," I'm referring to MEM_NJIT flux integrated over the entire map compared with Pixon flux integrated over its entire map. I've investigated a couple of cases – an X class flare from March 18, 2003, and a C class flare from September 20. For March 18, 2003, our "default" MEM_NJIT input flux is $\sim$28.1, which results in a total integrated map flux of $\sim$28.4. A Pixon map for this event yields a total map flux of 17.35. As for the September 20 event, the default flux is $\sim$45.45, which produces a map with a total flux of $\sim$46.8. All flux units are *photons* cm^{-2} s^{-1}. All these numbers are summarized in the tables below.

Another approach to flux is to use what is tagged in the RHESSI visibility structure arrays as "totflux" and documented as the DC term of a given visibility. Physically, one can look at this as the unmodulated RHESSI photon flux seen at a given detector. Graphically, it represents the y-offset of the sinusoid that is fitted in the visibility calculation. Ideally, the DC term averaged over all roll bins in a given detector is the same for each detector. To achieve uniform total fluxes, though, correction factors need to be applied to every visibility. The correction factors can be determined and applied to the bag of visibilities with the SSW program hsi_vis_normalize.pro. Once the visibilities are normalized,

$$DC_{all} = \sum_{i=1}^{9} DC_i/9, \tag{2}$$

where DC_i is the DC term for the ith detector (averaged over all roll bins for the detector) and we sum from 1 to 9 because that's how many subcollimators are on RHESSI. Since, after visibility normalization the average DC term for each detector should be equal to that of any other detector, DC_{all} should be equal to any DC_i.

Below is a table of source fluxes for the different input flux cases into MEM_NJIT, and for comparison, from Pixon. Also included are total flux of the sources defined by the user,

the flux for the entire map (integrated over all pixels), and the "noise" level, which is the map flux minus the total source flux. Again, all fluxes are in units of $photons\ cm^{-2}\ s^{-1}$. We'll start with the September 20, 2002 event:

Sep. 20, 2002 09:26:08-09:27:08 25-50 keV	Source 1 Flux	Source 2 Flux	Total Source Flux	Map Total Flux	Noise
MEM_NJIT Default	14.28	7.75	22.03	46.79	24.76
MEM_NJIT TOTFLUX	13.93	7.56	21.50	31.80	10.30
Pixon	10.16	6.44	16.60	33.30	16.70

Table 1: nj_influx = default ≈ 45.4; nj_influx = $DC_{all} \approx 31.9$

So for this event, the two MEM cases yield similar results for the two sources, but not the total map flux. Similar source fluxes result despite a default flux that is ∼1.5 times the input DC term. The "extra" flux is thrown into the background at low levels and contributes to the higher noise level. Presumably, telling MEM_NJIT to converge on a map with a flux close to the DC term will return a less noisy map, and one that is even closer photometrically to Pixon.

The same study can be carried out for the X-Class flare from March 18, 2003:

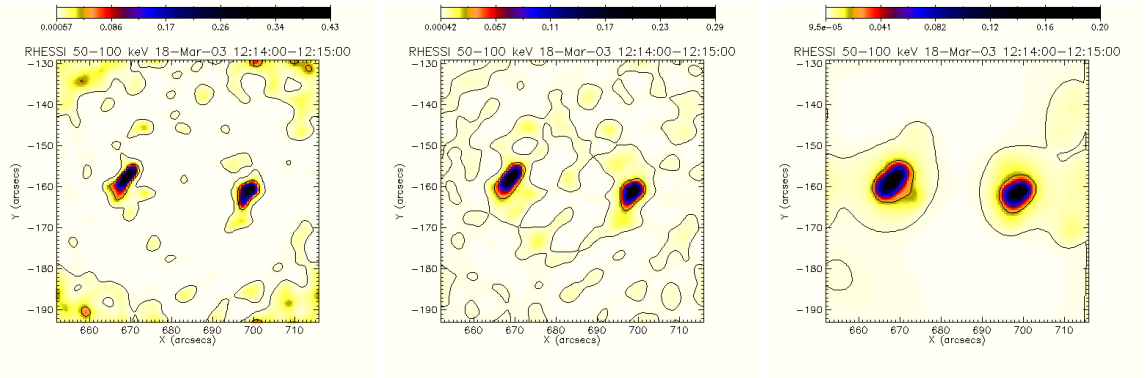
Mar. 18, 2003 12:14:00-12:15:00 50-100 keV	Source 1 Flux	Source 2 Flux	Total Source Flux	Map Total Flux	Noise
MEM_NJIT Default	6.56	5.69	12.25	28.36	16.11
MEM_NJIT TOTFLUX	6.24	5.63	11.87	18.39	6.52
Pixon	7.66	6.49	14.15	17.35	3.20

Table 2: nj_influx = default ≈ 28.1; nj_influx = $DC_{all} \approx 17.9$

So we're left with similar results – noise levels that are ∼2-3 higher in MEM_NJIT maps given the default flux value than in MEM_NJIT maps with an input flux equal to the average DC component, though total source flux is within 5%. It looks like we need a new default flux.

For a qualitative look, I've included color images for the March 18 event in Figure 1.

The images were plotted with IDL colortable number 5 (zero offset) and a gamma correction of 4. The resulting colortable was reversed in the plots so the darker pixels are more



(a) MEM_NJIT Default (b) MEM_NJIT TOTFLUX (c) Pixon

Fig. 1.— Image comparison

intense. Gamma correction was applied simply to highlight the lower level flux contributions. Contour levels are 1% and 10% of the map maximum. While the flux contributions are still more diffuse in the MEM_NJIT images than in Pixon, the pixels in Figure 1(b) are fainter than those in Figure 1(a). In addition, the color scales for Figures 1(b) and 1(c) are a better "match." For example, the maximum for the map in Figure 1(b) is only 50% higher than the Pixon maximum, whereas the Figure 1(a) maximum is more than twice as much as Pixon's.

3. Super-resolution

MEM_NJIT has an inherent property, as do all Maximum Entropy Methods, of super-resolution – going beyond the resolving power of an instrument. This super-resolution phenomenom exists because of the nature of Fourier Imagers, whether they be Radio Interferometers or rotating subcollimators that have been used with Yohkoh and RHESSI. In any case, the UV coverage with these instruments is limited and Fourier inter/extrapolation is required to get anything beyond the dirty map. Another reason for super-resolution is the fact that any MEM is nonlinear with the a priori knowledge of positivity, i.e. all pixels in a map have intensities that are greater than zero. (NN)

Fourier extrapolation leads to visibilities being "measured" at UV coordinates beyond the resolution of a given instrument. In radio jargon, visibilities are measured at higher spatial fequencies than what is allowed by the instrument. Positivity opens the door to nonzero weighting of these visibilities. One might ask why this is so, and the reason is the relationship between visibilities and maps.

Mathematically, visibilities can be computed from a map with the equation $V_j = A_j e^{i\phi_j}$, where A_j is the amplitude of one visibility and the phase ϕ_j is a vector product of the Fourier coordinates and the Cartesian coordinates. More precisely, since one visibility is measured at each UV point (or at each roll angle θ_j, since $\theta_j = \arctan \frac{v_j}{u_j}$), and for that roll angle the visibility is summed over the entire map, the visibility vector component can be expressed as:

$$V_j = \sum_m \sum_n F_{mn} e^{2\pi i (u_j * x_m + v_j * y_n)}, \tag{3}$$

with F_{mn} being the flux of one pixel at position (m, n). x_m and y_n are just Cartesian points of a map (so the visibility component is summed over all x pixels and y pixels). For RHESSI, the Fourier (UV) coordinates lie along circles with radius $k_i = \frac{1}{pitch_i}$, with $pitch_i$ corresponding to the angular resolution of the i^{th} detector. The range of k_i is from $\frac{1}{183.323} \approx 0.005\ arcsec^{-1}$ to $\frac{1}{2.26} \approx 0.4\ arcsec^{-1}$. Values for k_i can then be used for the individual UV values along the i^{th} UV circle:

$$u_j = k_i \cos\theta_j, v_j = k_i \sin\theta_j \tag{4}$$

θ_j is the aforementioned roll angle, which is just the orientation of the instrument, so u and v values are instrument-dependent.

Fourier extrapolation will extend the UV circles beyond the largest RHESSI circle, thus contributing to the complex exponential of equation 3 , while positivity keeps F_{mn}, well, positive. Thus, we have nonzero values for V_j beyond the allowed spatial frequencies of the instrument, providing higher pseudo-resolution, which we call super-resolution. The result is a MEM image that is comparable to a Clean image using data from an adjacent array configuration with higher resolution (CE), or for RHESSI, adding a finer grid.

3.1. Analysis With a Simulated Source

Consider a Dirac δ-function source – infinite flux that is infinitely small in width. If a map was composed of only the δ-function, amplitudes of calculated visibilities would be constant over all frequency space. Computational constraints in IDL inhibit the use of a δ-function (because of the infinite flux), but we can construct a point source of a finite flux and expect similar results. The model source map is a map of pixels with zero flux except for one pixel near the center. Model visibility amplitudes, calculated with Equation 3, are

displayed in Figure 2 – they form a straight line over all detectors until the resolution of the detector approaches the pixel size. Pixel sizes for this map are 0.1", far beyond the resolution of RHESSI, so we only see a straight line.

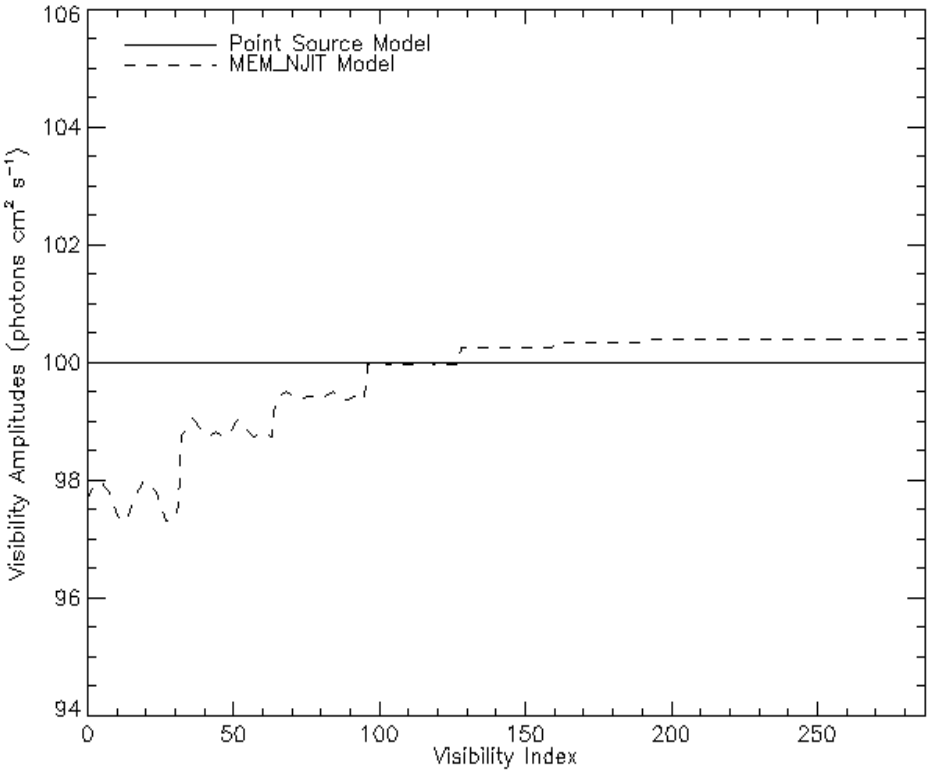


Fig. 2.— Model visibilites for the point source and MEM_NJIT output, both calculated with Equation 3. Here, finer grids have lower indices.

When running MEM_NJIT, these model visibilities are used as input, as are the flux of the pixel (which is defined by the user; in my case, I chose a flux of 100 *photons* cm^{-2} s^{-1}), and the Fourier components. My UV coordinates were uniformly distributed (i.e. equal angular spacing of say, 10 degrees, between UV points) onto 9 circles with the aforementioned radii, k_i. The resulting MEM_NJIT solution, Figure 3, is *almost* a point source (if a 2D Gaussian is fitted to the MEM model, it's y axis, σ_2, is slightly larger than it's x axis, σ_1 – something like 0.8 pixels to 0.7 pixels) with a flux less than 1% larger than the theoretical, user defined flux (MEM's map returns a flux of 100.4). Model visibility amplitudes of the MEM_NJIT map overplotted in Figure 2 exhibit a little bit of modulation in the finer grids since the source is a very small Gaussian, and the minimum amplitude is 97.3. The reason for the divergence is the uncertainty in the visibilities, which by default is the MEM_NJIT

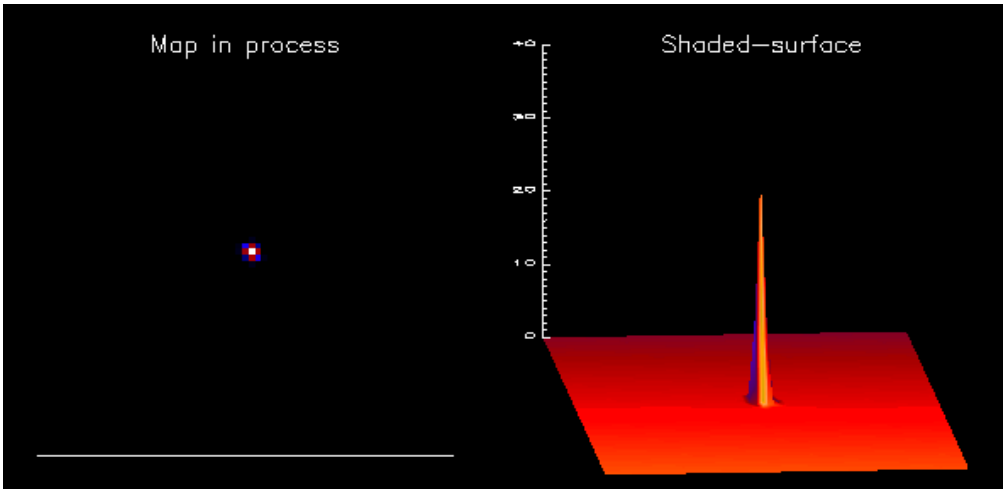


Fig. 3.— Point source map and shaded surface

tolerance (0.03 by default) multiplied by the input visibility amplitudes (all equal to 100), so there is an uncertainty of ± 3 *photons* cm^{-2} s^{-1} at each roll angle.

Now onto the super-resolution: to prove Fourier extrapolation is occuring in MEM_NJIT, I choose to just calculate model visibilities from the MEM_NJIT map as if RHESSI had grids with finer resolutions, i.e. I'm adding larger UV circles to my sample. The goal is to see if the visibility amplitudes drop off with finer resolutions (of which RHESSI isn't capable), and if they do, how quickly they drop. Since higher spatial frequencies are expected to be weighted less, possible trends could be exponential decay or something like a "Gaussian drop off."

Indeed, I have observed the Gaussian drop off starting at around the 2^{nd} grid. Again, the change in fluxes measured by RHESSI detectors is still within the default uncertainty. Significant drop off starts to happen once we add one or two more grids with finer resolution. Figure 4 illustrates the drop off with added detectors. Note we can only add detectors that have resolutions greater than the pixel size. Full disclosure: Figure 4 is not a plot of *every* amplitude at *every* roll angle. Rather, each data point represents the *average* amplitude of the visibilities in a given detector (amplitudes for an elliptical Gaussian will oscillate around the average).

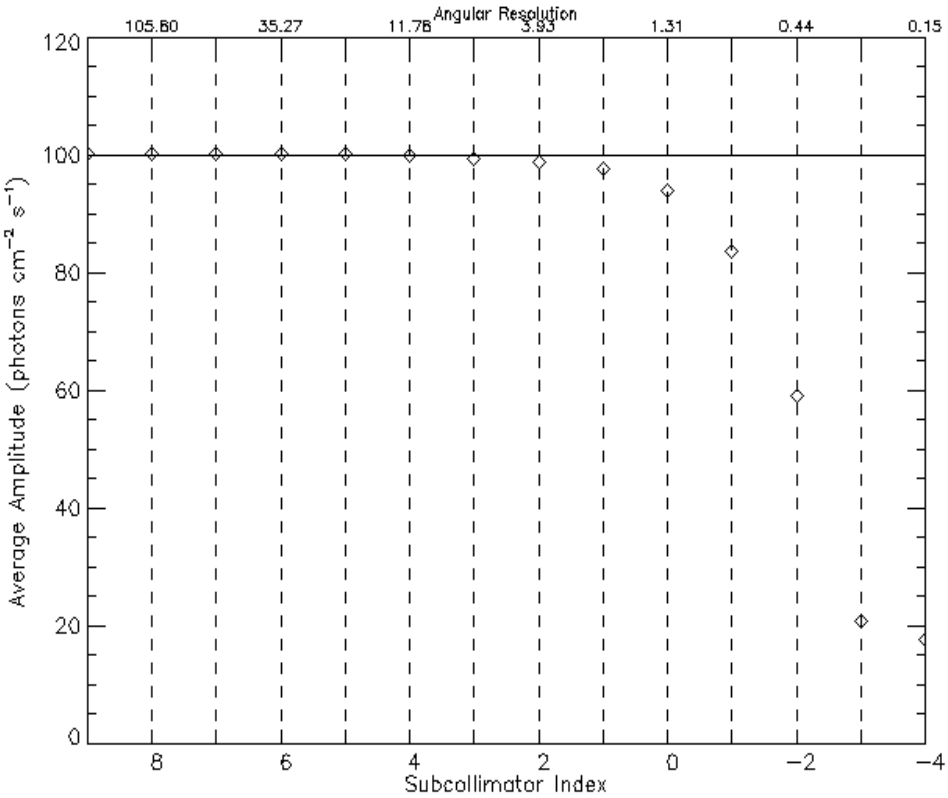


Fig. 4.— Average amplitudes of model visibilities, which include Fourier extrapolation. The coarsest grids are plotted first (flipping the X axis from Figure 2). Resolution can only be extended to ≈0.15" since the pixel size of the map is 0.1". Subcollimator indices <1 have angular resolutions less than RHESSI's finest grid pair.