AWAIC: A WISE Astronomical Image Co-adder

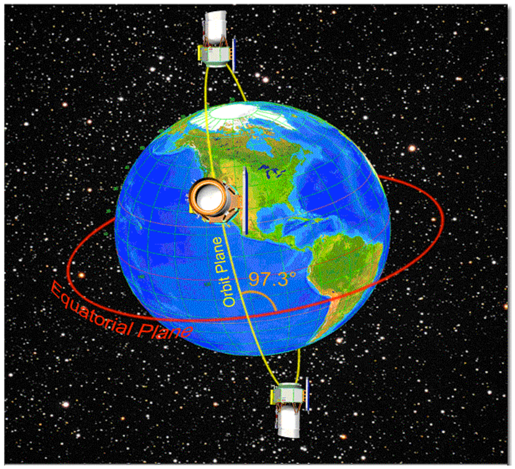
Frank Masci & John Fowler
Infrared Processing and Analysis Center / Caltech

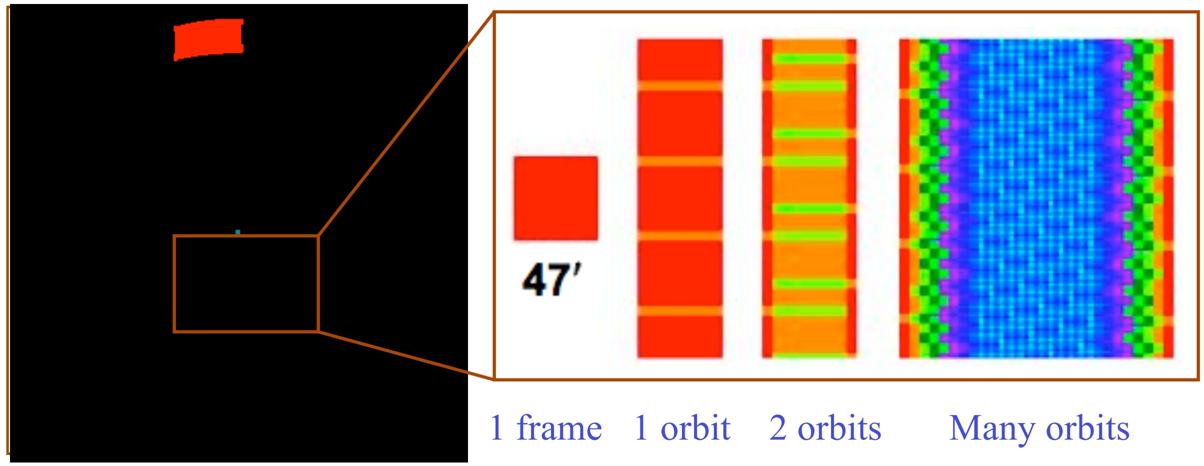
What Is WISE?

- A NASA Medium Explorer (MIDEX) Mission
- P.I. - Ned Wright (UCLA)
- Scheduled for launch in November 2009

- The **W**ide-field **I**nfrared **S**urvey **E**xplorer (WISE):
  - Perform an all-sky survey at 3.3, 4.7, 12 & 23 μm with up to 3 orders of magnitude more sensitivity than previous surveys
  - A cold 40 cm telescope in a sun-synchronous low-Earth orbit
  - Image quality ≈ 6″ FWHM at wavelengths 3.3 - 12 μm; ≈ 12″ at 23 μm
  - 1024 × 1024 pixel infrared detector arrays, at 2.75″/pixel

Simple Mission Design



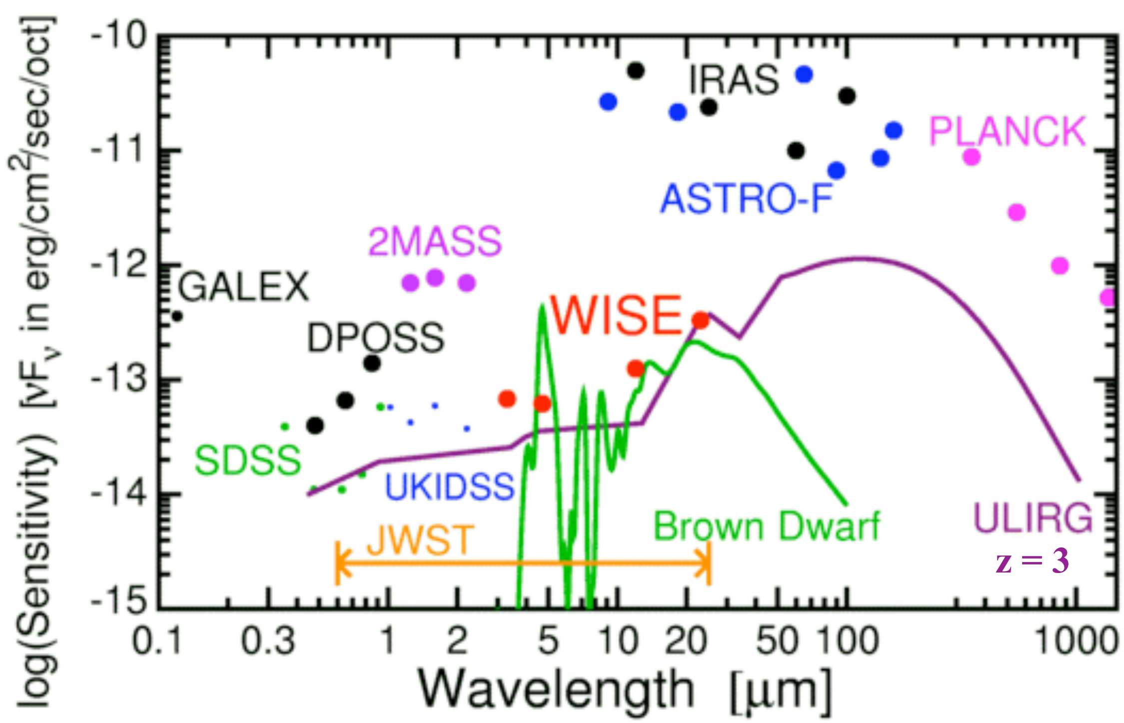


- 523 km, circular, polar sun-synchronous orbit
  - One month of checkout
  - 6 months of survey ops
- One simple observing mode - half-orbit scans

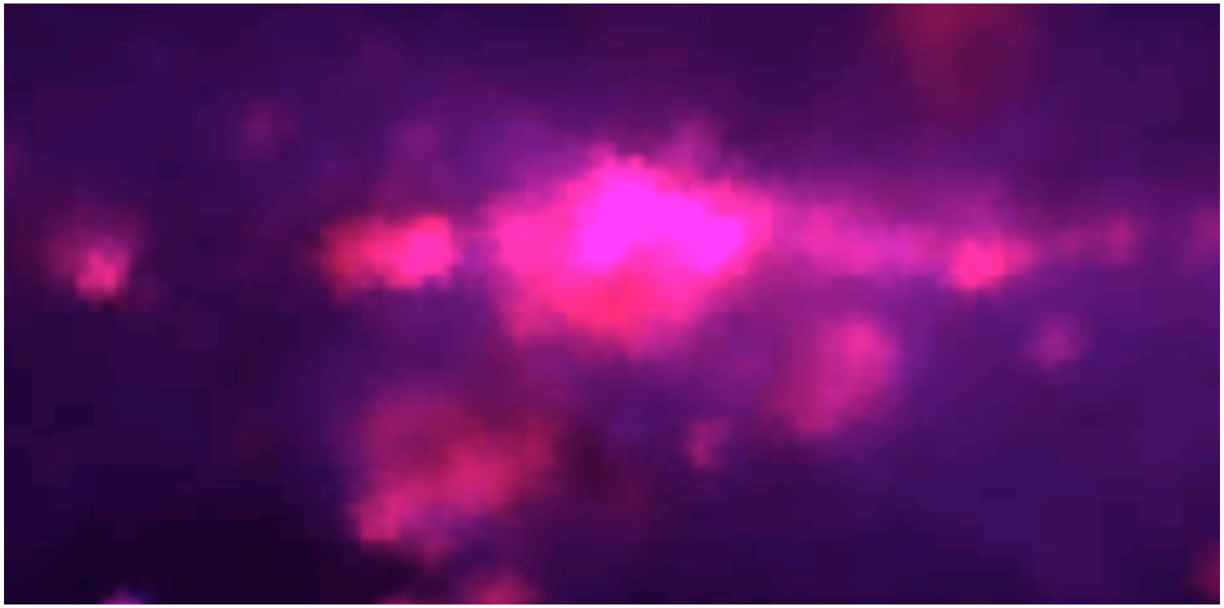
- Scan mirror "freezes" orbital motion ⇒ efficient mapping
  - 8.8-s exposure per frame
  - 10% frame to frame overlap (in-scan)
  - 90% orbit to orbit overlap (cross-scan)
- Expect to achieve a median of 8 exposures/position on the ecliptic equator, > 1000 exposures at poles
- Requirement is to have >95% of sky with ≥4 exposures
- Uplinks, downlinks and calibrations occur at poles

Science Goals

- Find the most luminous galaxies in the Universe
- Find the closest stars to the Sun
- Detect most main belt asteroids larger than ~3km
- Extend the 2MASS Survey into the thermal (mid) infrared
- Provide the essential catalog for the James Webb Space Telescope (JWST)



IRAS versus WISE



- 20 years ago, IRAS gave us this view of the galactic center
- Still our best view of the *whole* sky in the mid-IR

- Same region as expected from WISE. This is a MSX-2MASS composite

WISE Products

WISE will deliver to the scientific community:

- A digital Image Atlas containing ~220,000 calibrated images, or co-adds of the survey frame exposures covering the whole sky in 4 mid-IR bands
- Ancillary co-add products: depth-of-coverage maps (from all good pixels) and uncertainty maps
- Atlas Image tiles are ≈ 1.5° × 1.5° re-sampled at 1.375″/pixel
- A Source Catalog of ≈ 5×10^8 objects merged across all 4 bands to photometric S/N = 5. All sources will be astrometrically and photometrically calibrated
- All processing will occur at the WISE Science Data Center at IPAC

This Presentation

- Describe image co-addition framework as implemented at the WISE Science Data Center, including preparatory steps:
  - outlier detection and masking
  - background-level matching

- Describe algorithms implemented in AWAIC - A WISE Astronomical Image Co-adder
  - Interpolation using the detector's Point Response Function (PRF)
  - How this compares to other interpolation methods

- Methods to assess statistical robustness of co-add fluxes (uncertainty estimation)

- Extension of AWAIC to resolution enhancement (HiRes):
  - Describe the Maximum Correlation Method (MCM) for HiRes
  - Associated diagnostics and uncertainties in HiRes'd products (received little attention in the past)
  - HiRes is not in WISE automated pipeline. Implemented to support offline research

Co-addition Pipeline Overview

INPUTS: (images in FITS format)
- instrumentally/astrometrically calibrated image frames
- bad-pixel masks
- uncertainty (sigma) images from prior noise model

Throughput matching (*multiplicative*) to scale input frames to a common photometric zero-point

Frame background-level matching (*additive*)

Outlier detection & masking

Interpolate frames onto a common grid; use kernel optimized for outlier detection

Outlier detection on interp. pixel stacks using robust statistics. *Masks updated*

Co-addition of all unmasked pixels using **AWAIC:**
- fast re-projection and distortion correction
- interpolation: PRF weighted averaging with optional inverse variance weighting

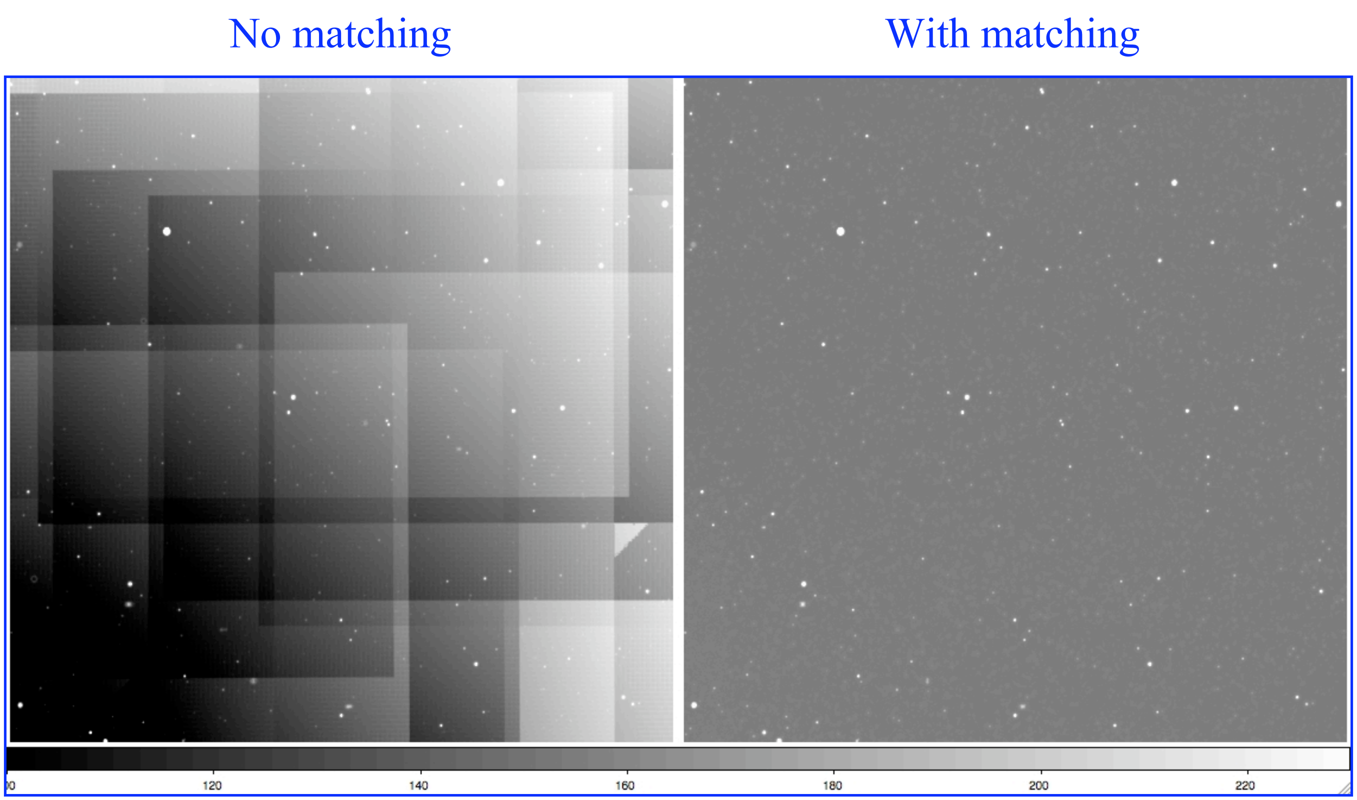
Quality Assurance: backgrounds, noise, coverage/outlier stats, frame-stack χ^2 stats for uncertainty verification

OUTPUTS:
- main co-add intensity images
- depth-of-coverage maps
- uncertainty images
- all 4096 x 4096; ≤1.375″/pixel

Final Product Generator. Deliver to public through IRSA at IPAC

Background-level Matching

- Instrumental transients lead to varying background levels between frames
- **Goal:** obtain seamless (or smooth) transitions between frames across overlaps but preserve natural background variations as much as possible
- **Simple method:** fit a “robust” plane to each frame, subtract to equalize frames, then add back a common plane or level to all frames computed from a median over all the fits

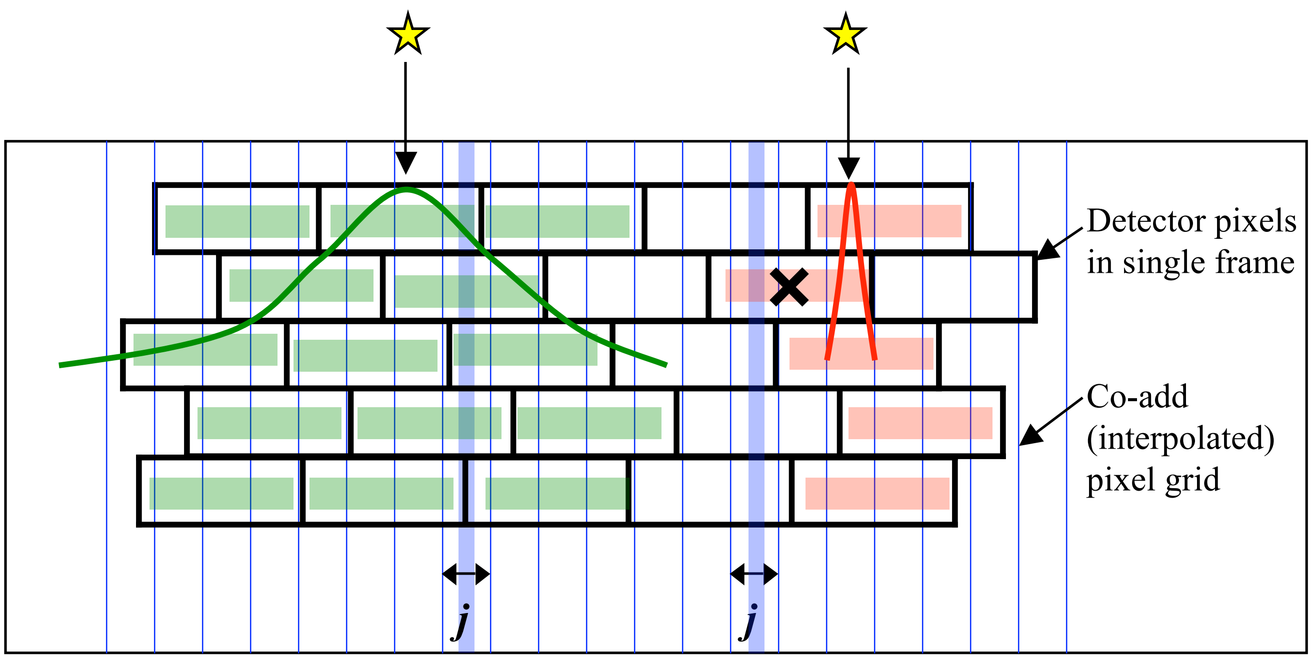


Outlier Detection

- Take advantage of the redundancy in multiple frame exposures and flag outlying measurements
- Project and interpolate frames onto a common grid, apply an outlier identification algorithm to pixel stacks:

  - flag in mask if : $p_i > median\{p_i\} + t_{thres}\sigma_j$ or $p_i < median\{p_i\} - b_{thres}\sigma_j$
  - where σ_j is a robust measure of spread, e.g., via percentiles : $\sigma_j \approx 0.5(p_{84} - p_{16}) \approx (p_{50} - p_{16})$

- It helps to have good sampling of the PSF for method to be reliable! WISE bands: >~ critically sampled



≥ Critically sampled PSF case *Under-sampled PSF case*

Image Co-addition in AWAIC

- **Goal:** want to optimally combine all measurements into a *faithful* representation of the sky given all the instrumental systematics, cosmic rays etc. “Optimality criterion” defined later
- AWAIC uses the detector’s **P**oint **R**esponse **F**unction (PRF) as the interpolation kernel
- **PRF** = *Point Spread Function* (*PSF*) ⊗ *pixel response;* response is usually a top hat
  - represents the end-to-end transfer function from sky to measurement pixels
  - each pixel collects light from its vicinity with an efficiency described by the PRF

- Flux in a co-add pixel *j* is estimated using PRF and inverse-variance weighted averaging:

$$\langle f_j \rangle = \frac{\sum_i \frac{r_{ij}}{\sigma_i^2} D_i}{\sum_i \frac{r_{ij}}{\sigma_i^2}}$$

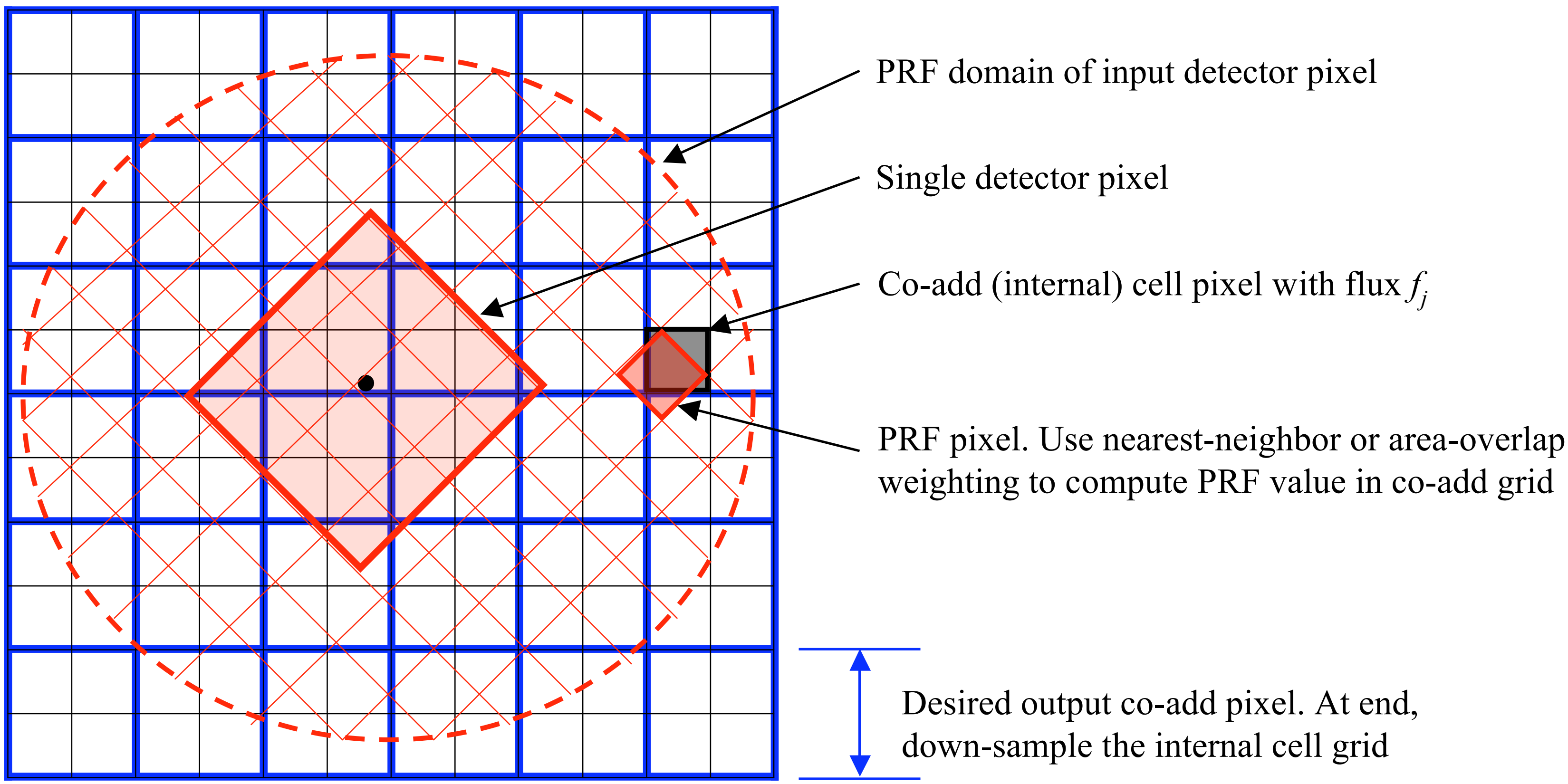
Input detector (pixel) measurements

PRF (volume normalized to unity)

Variance from propagated noise model (optional)

- Some popular interpolation methods:
  - *Overlap-area weighted averaging*: interpolation weights are pixel overlap areas $r_{ij} = a_{ij}$. PRF ≡ top hat
  - *Drizzle*: extension of overlap-area that includes shrinkage of input pixels
  - *Tapered sync interpolation*: optimal for band-limited data sampled at or better than Nyquist. Missed cosmic rays, noise spikes can mess up a large region and lead to severe ringing

PRF Interpolation Schematic



Why PRF as Interpolation Kernel?

Pros:

- Reduces impact of bad/masked pixels if the data are well sampled (even close to critical). Leads to effectively non-zero coverage at the bad pixel locations on co-add due to the overlapping PRF tails of 'good' pixels:

- Defines a linear matched filter optimized for point source detection
  - High frequency noise is smoothed out without affecting point source signals ⇒ peak S/N maximized
  - Process is effectively a cross-correlation of a point source template (the PRF) with input data
  - This will benefit processing at the WSDC since a source catalog is one of its release products
  - Weighted average also ensures S/N is maximized ⇒ maximum likelihood estimator for 'Normal' data

- The big one: allows for resolution enhancement (HiRes): PRF can be "deconvolved" - more later

Cons:

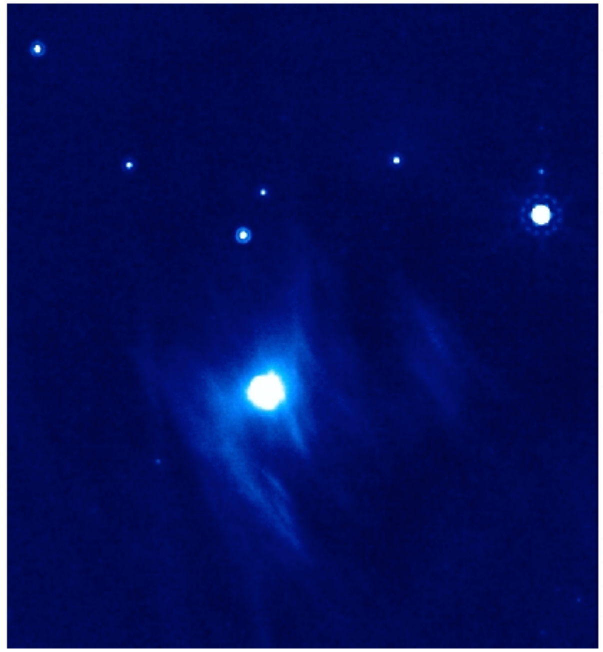
- Noise is correlated on *larger* spatial scales in the co-add when a broad kernel is used
- Smoothing operation ⇒ "flux smearing". Cosmic rays can masquerade as real sources if not masked
- Both these must be accounted for in photometry off co-adds: in flux and uncertainty estimation (e.g., PTO)

Area Overlap vs PRF Interp.

A field in Taurus: *Spitzer* 24μm

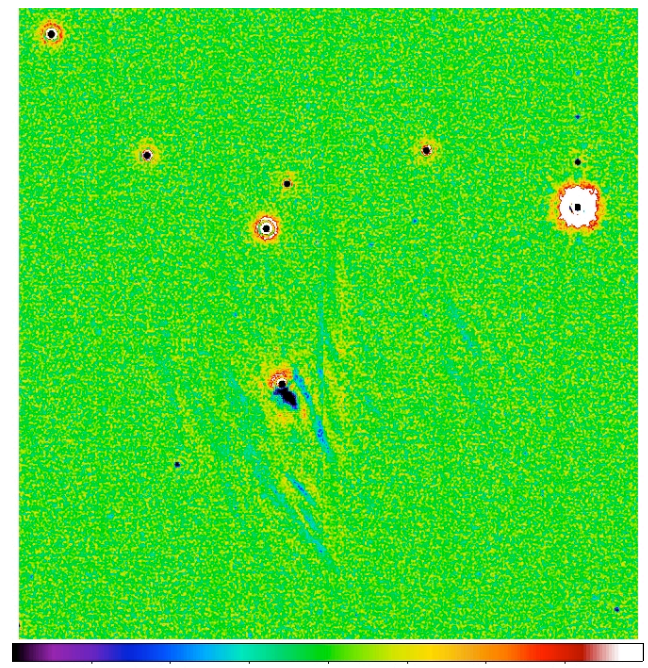
PRF-weighted averaging
(PRF interp kernel)

÷

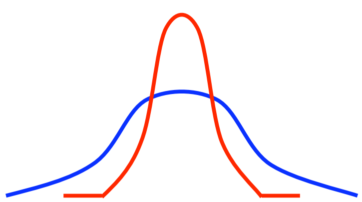


area-overlap weighting
(top-hat interp kernel)

=



±2%



⇒ PRF interpolation “smears” flux on small scales
⇒ photometry with small apertures must use appropriate aperture correction

Other Features in AWAIC

- Allows for a spatially varying PRF. Usually non-isoplanatic over the focal plane for large detector arrays

- Uncertainties in co-add pixel fluxes
  - Stored as 1-sigma values in separate image products
  - Based on input priors: combines input measurement uncertainties propagated from a noise model

- Ancillary products: depth-of-coverage maps and images of outlier locations (some examples later)

- Quality Assurance: e.g., statistics on depth-of-coverage, sky-backgrounds, outliers. Metrics to check that co-add uncertainties (based on priors) are statistically compatible with the input data:
  - e.g., compare with *a posteriori* data-derived variances using χ^2 :

$$\chi^2 = \sum_{pixel\,j}^{N} \frac{\left[p_j - \langle p \rangle\right]^2}{\sigma_j^2}$$

⇒ Applied spatially on uniform *sky* pixels in co-add, or on input image stacks to quantify systematics

Co-add pixel uncertainties propagated from noise model

- Supports FITS standard, WCS standards with distortion, and five commonly used projections (TAN, SIN, ZEA, STG, ARC) implemented in a fast re-projection library

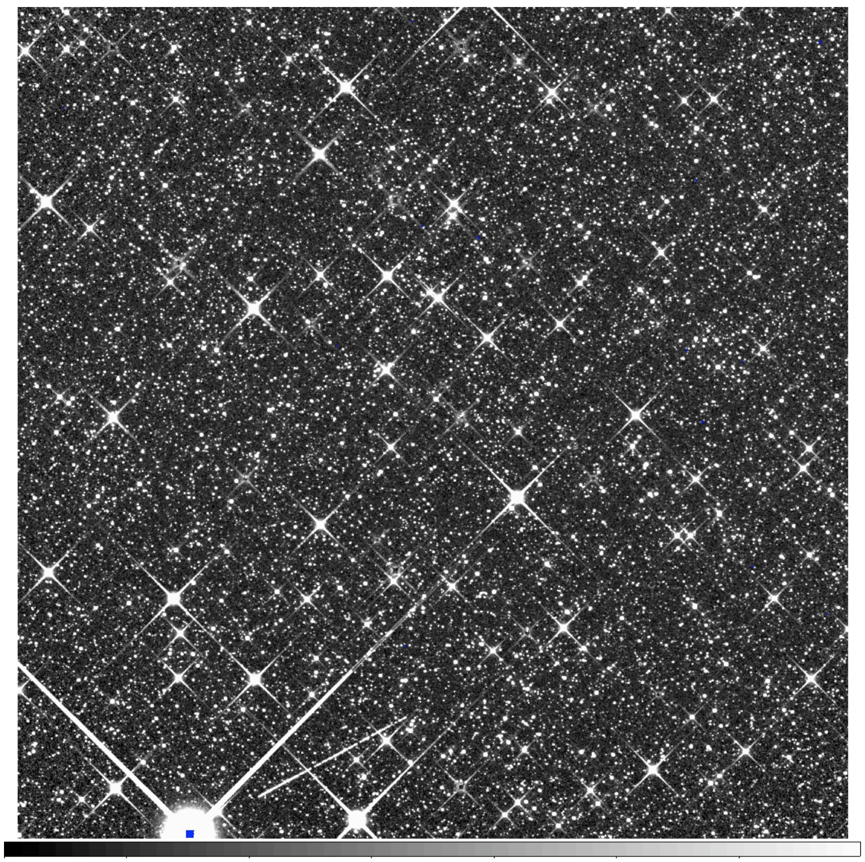
- Generic enough for use on non-WISE image data: e.g., exercised on *Spitzer* and *HST* data

Example of WISE Atlas Images

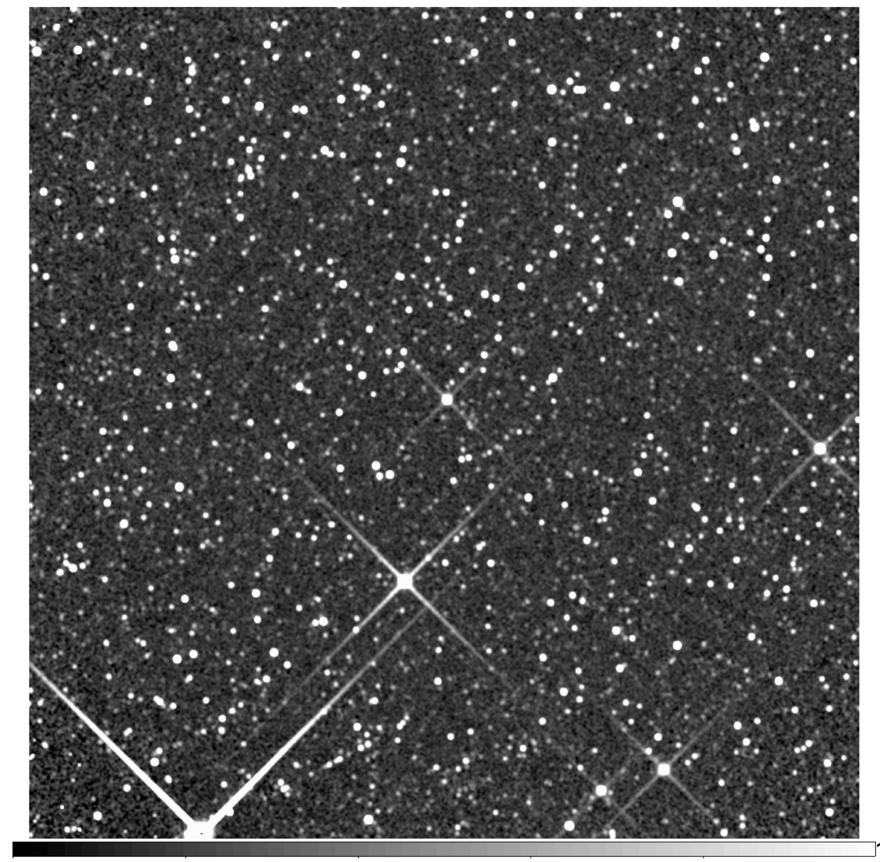
- Simulated frames provided by Ned Wright (P.I.): used seed sources from 2MASS catalog
- Then co-added with AWAIC
- Mid-ecliptic latitude field (β ≈ +30°) - example of what WISE *may* see

1.56°

3.3 μm

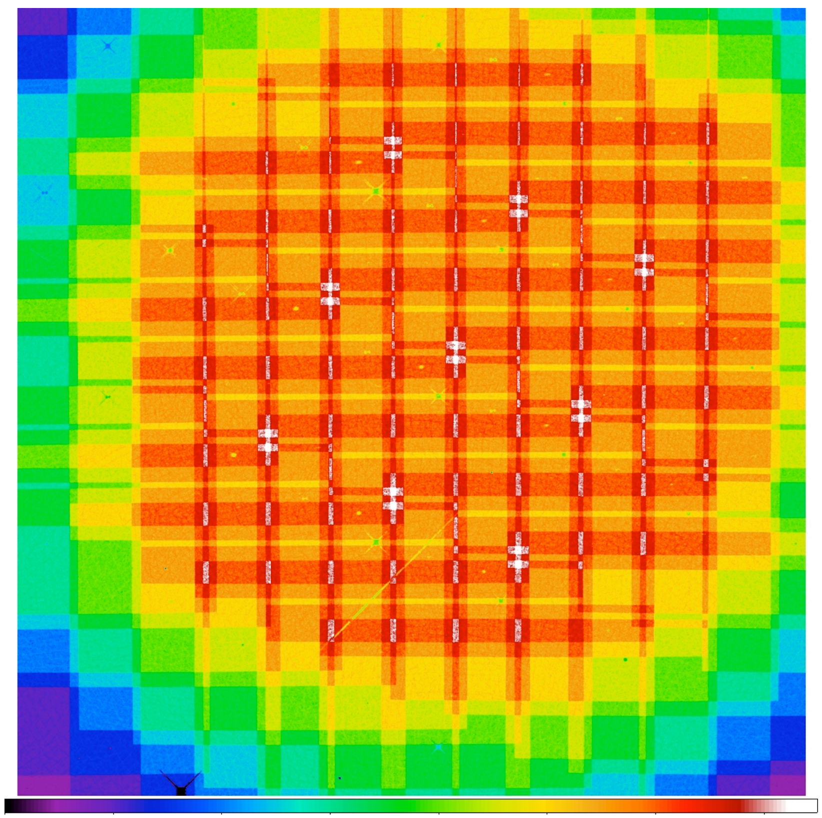


23 μm



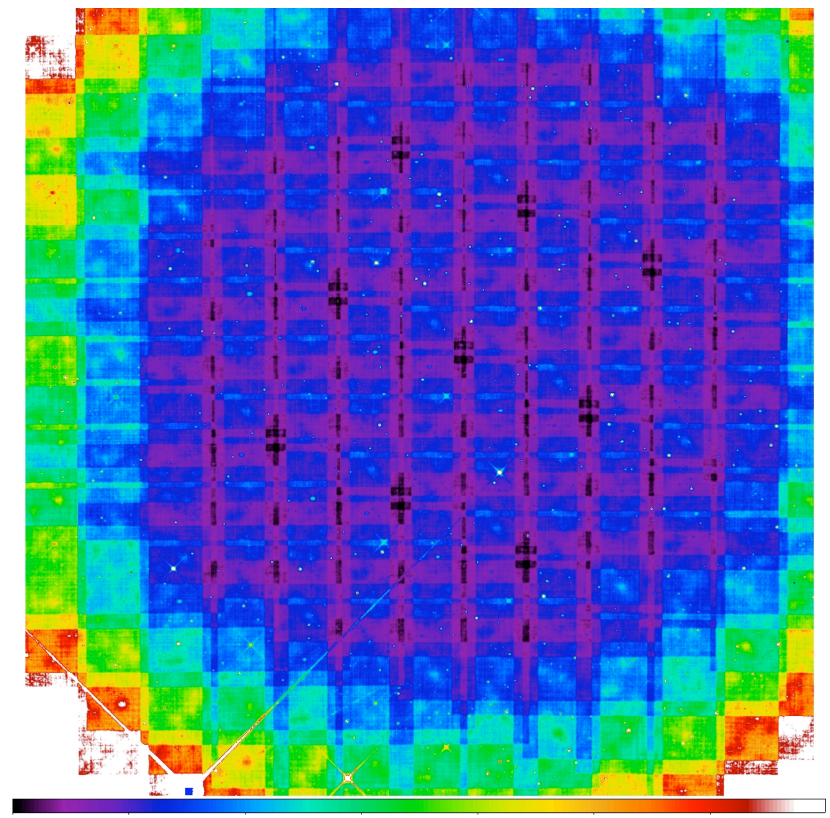
Depth-of-coverage and σ maps

- Depth-of-coverage map: effective number of repeats from all *unmasked* pixels at each location
- σ-map: 1-sigma uncertainty for each pixel propagated from a noise model



4 14.5

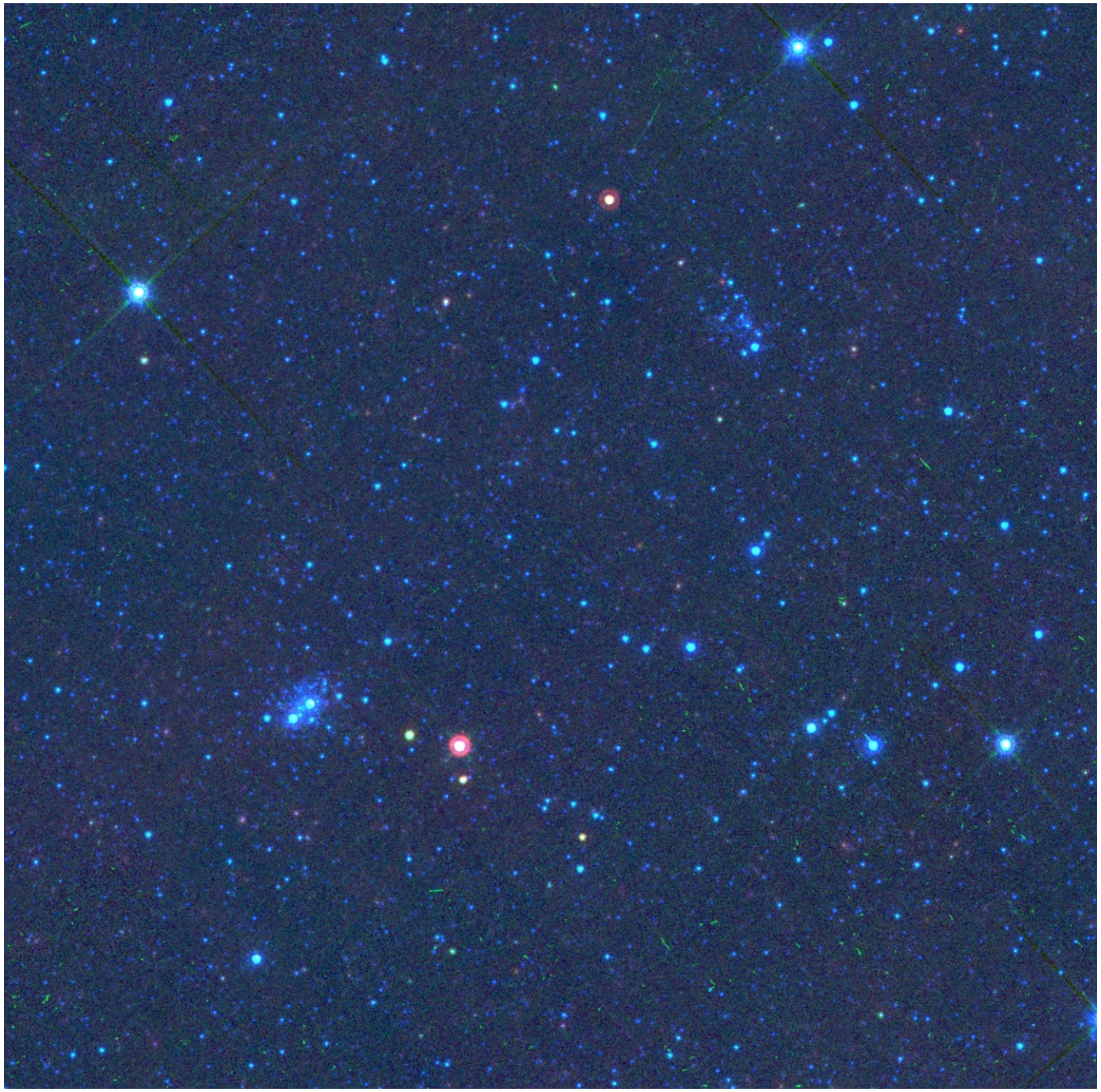
coverage



0.17 0.3

$$\sigma \propto \frac{1}{\sqrt{\text{coverage}}}$$

South Ecliptic Pole (near LMC)



WISE "Touchstone field"
Combines AWAIC mosaics in *Spitzer* bands:
4.5μm (blue)
8μm (green)
24μm (red)

⇒ Proxy for WISE bands 2, 3, 4

~ 20′ ~ 1/5 of WISE Atlas Image

HiRes: Maximum Correlation Method (MCM)

- Originally implemented to operate on data from the InfraRed Astronomical Satellite (IRAS) ~ 20 years ago

- Earlier we discussed combining images to create a co-add, **MCM** asks the reverse:
  - what model or representation of the sky propagates through the measurement process to yield the observations within measurement error?

- Measurement process is a filtering operation performed by the instrument's Point Response Function (PRF):

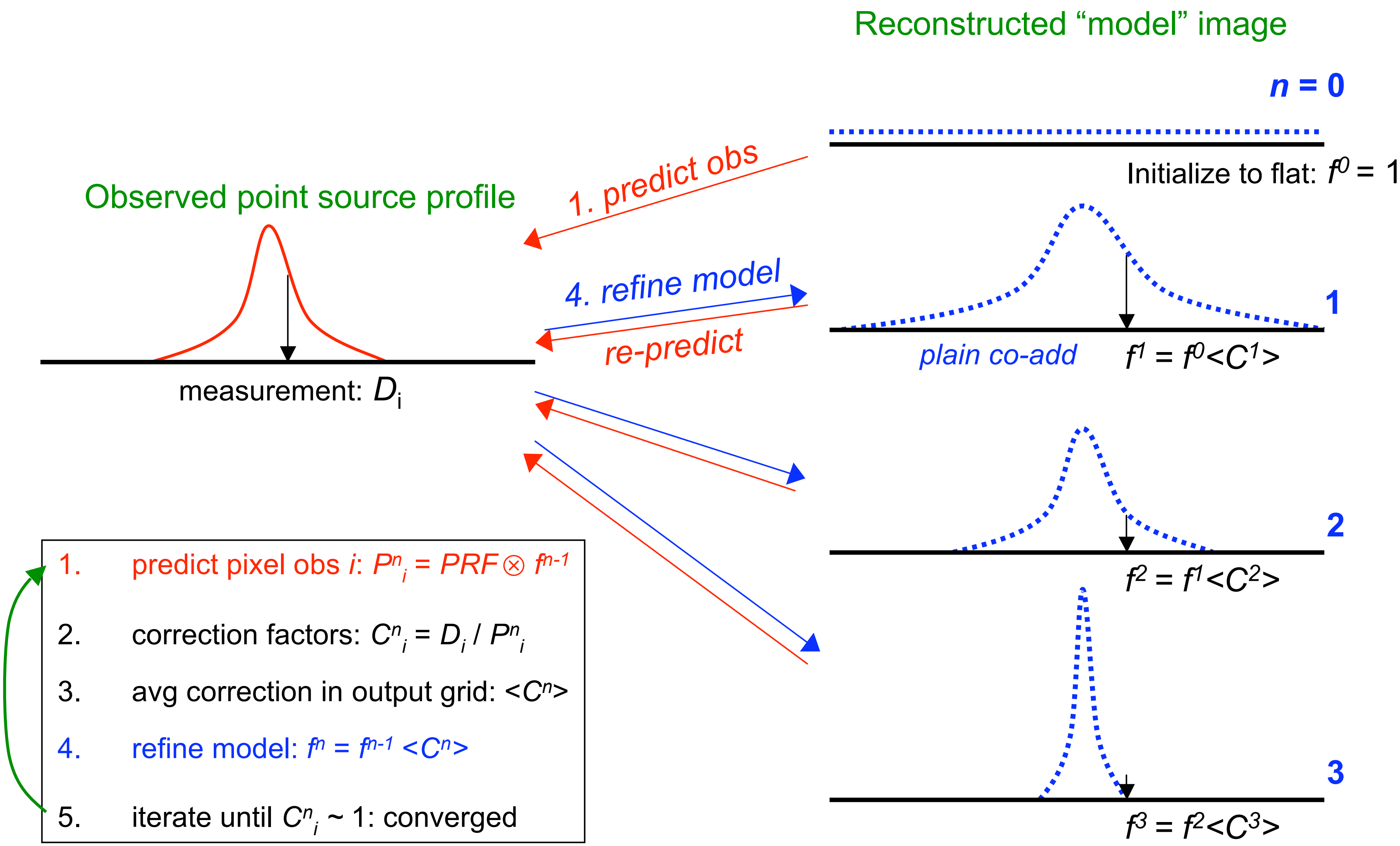
$$\boxed{\underset{?}{Sky\ "truth"} \quad \otimes \quad \underbrace{PSF \otimes \Pi}_{PRF} \quad \otimes \quad sampling \quad \rightarrow \quad measurements}$$

- MCM starts with a "maximally correlated" image - a flat model image and modifies (or de-correlates) it to the extent necessary to make it reproduce the measurements to within the noise
  - Instead of a flat model image, can also use prior information as starting model

- MCM *implicitly* gives a solution which is the "smoothest" possible, i.e., has maximal entropy
  - c.f. to Maximum Entropy Methods: smoothness built in explicitly as a constraint in cost function

- In general, noisy data ⇒ solution to the deconvolution problem is not unique. Some methods give more structure or detail than necessary to satisfy the data ⇒ no guarantee that structure is genuine
  - with input data as only constraint, MCM gives the "simplest" solution - the smoothest

MCM Process



1. predict pixel obs i: $P^n_i = PRF \otimes f^{n-1}$
2. correction factors: $C^n_i = D_i / P^n_i$
3. avg correction in output grid: $\langle C^n \rangle$
4. refine model: $f^n = f^{n-1} \langle C^n \rangle$
5. iterate until $C^n_i \sim 1$: converged

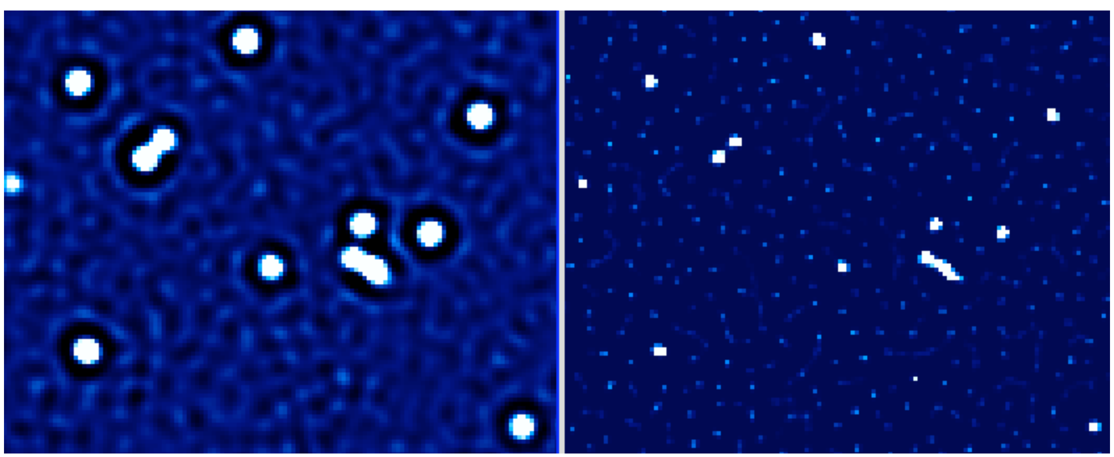
MCM Details

- MCM reduces to the classic Richardson-Lucy method if:
  - PRF is isoplanatic. Constant kernel $\Rightarrow$ allows use of Fourier de-convolution methods
  - Inverse variance weighting is disabled from the PRF-weighted averaging of input data
  - Prediction (simulator) step to check for data consistency and terminate iterations is removed

- MCM does not alter *information* content of an image. Is reversible within measurement error
  - Process re-emphasizes different parts of the frequency spectrum to allow detection of unresolved objects

- Includes a ringing suppression algorithm
  - Ringing is common to all deconvolution methods and limits super-resolution
  - Due to band-limited nature of input data, information beyond some high freq. cutoff cannot be recovered
  - Method: separate background and “source” flux, run MCM on source images and recombine at end. Enforces a positivity constraint - source flux won’t ring against a zero background

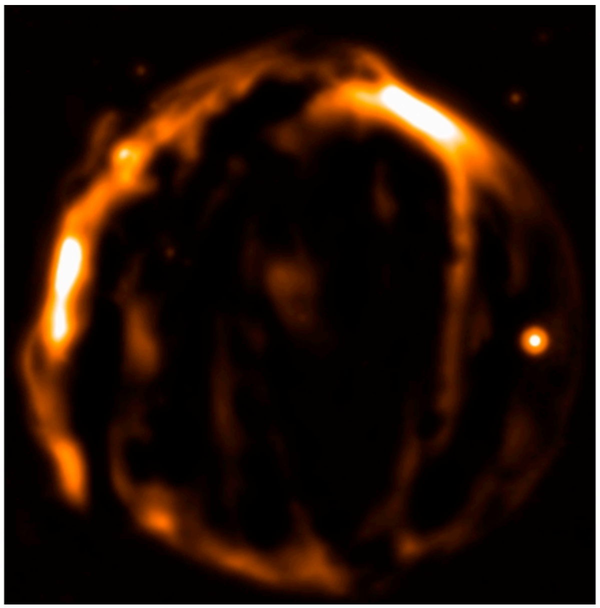
with ringing



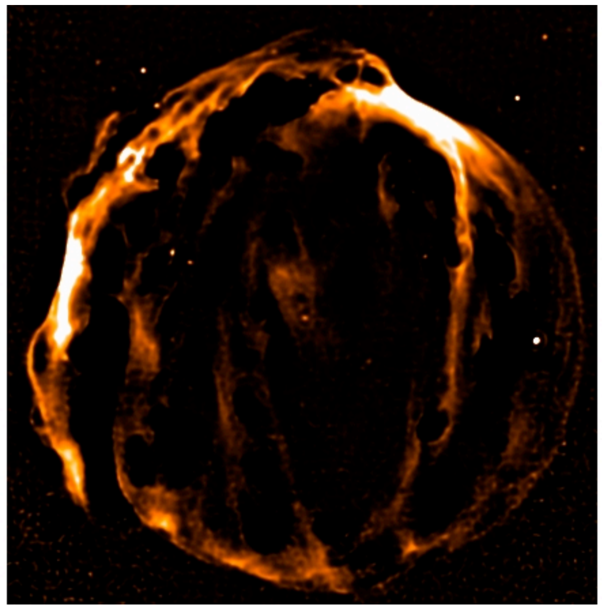
ringing suppressed

Tycho's Supernova Remnant

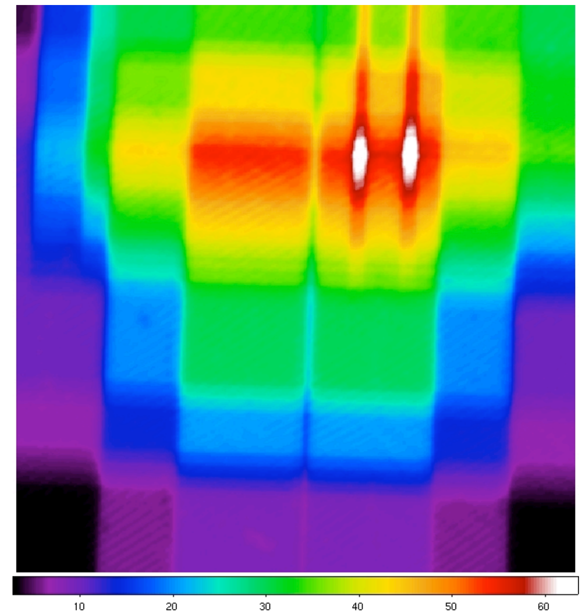
Spitzer-MIPS 24 μm



Co-add (1st MCM iteration)



HiRes: 40 MCM iterations



5 65

depth-of-coverage map

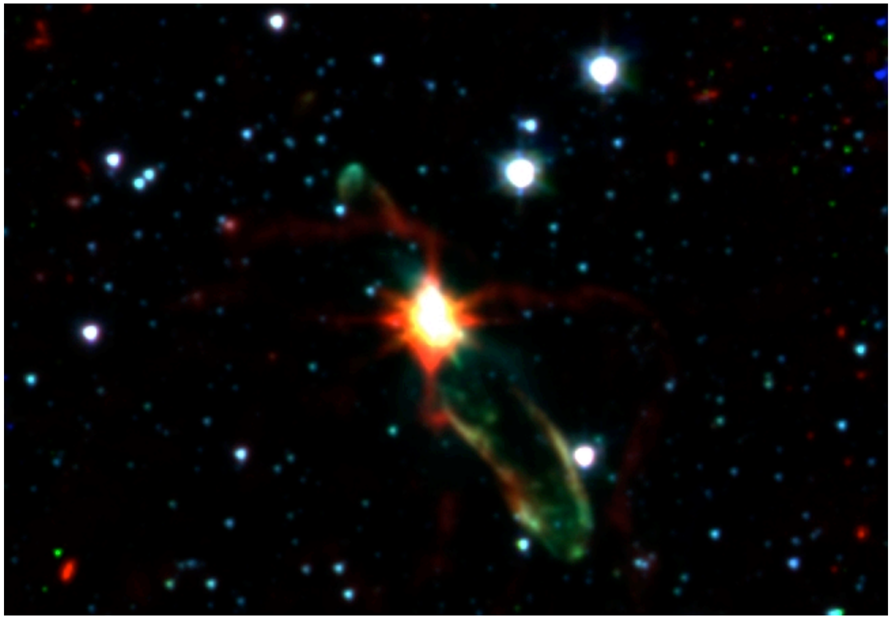
FWHM of effective PRF: went from ~5.8″ (native) to ~1.9″

⇒ ×3 gain in resolution per axis

Herbig-Haro 46-47

Spitzer-IRAC composite:

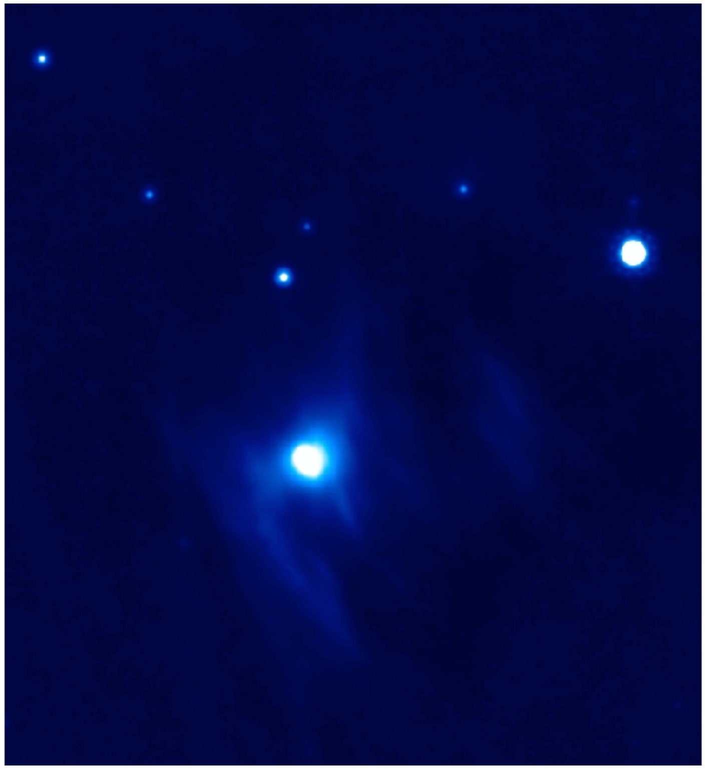
3.6 μm, 4.5 μm, 8 μm



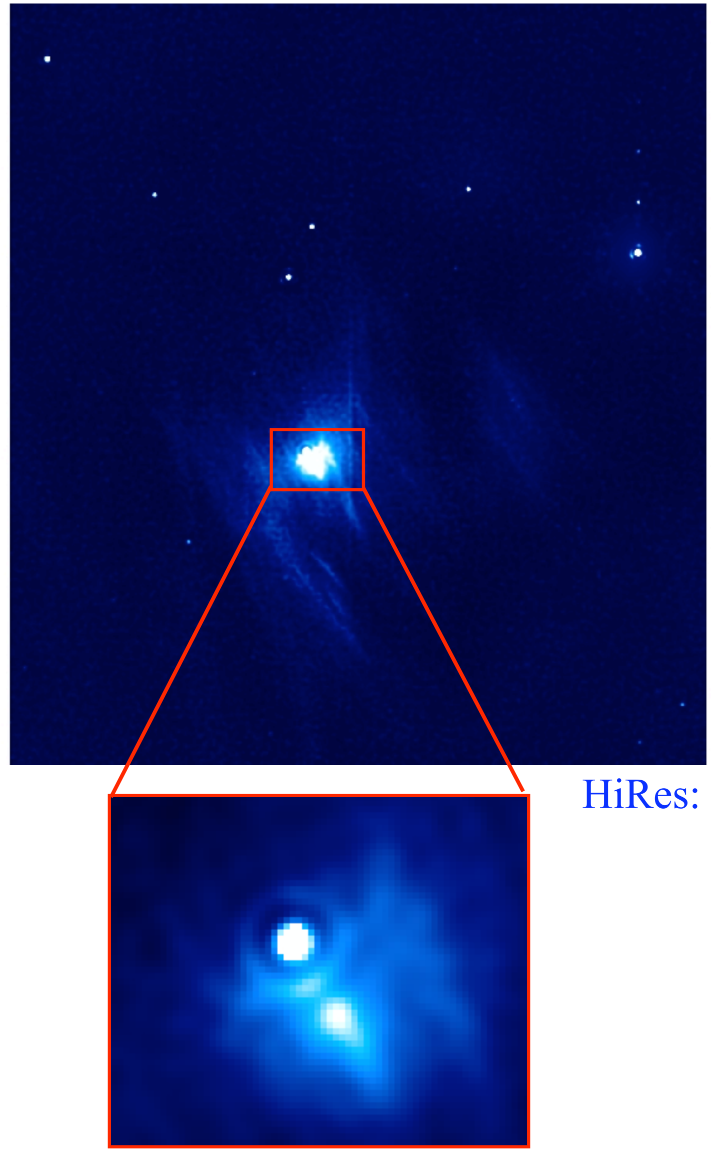
Co-add (1st MCM iteration)

HiRes: 20 MCM iterations

SF Region in Taurus



Co-add (1st iteration)



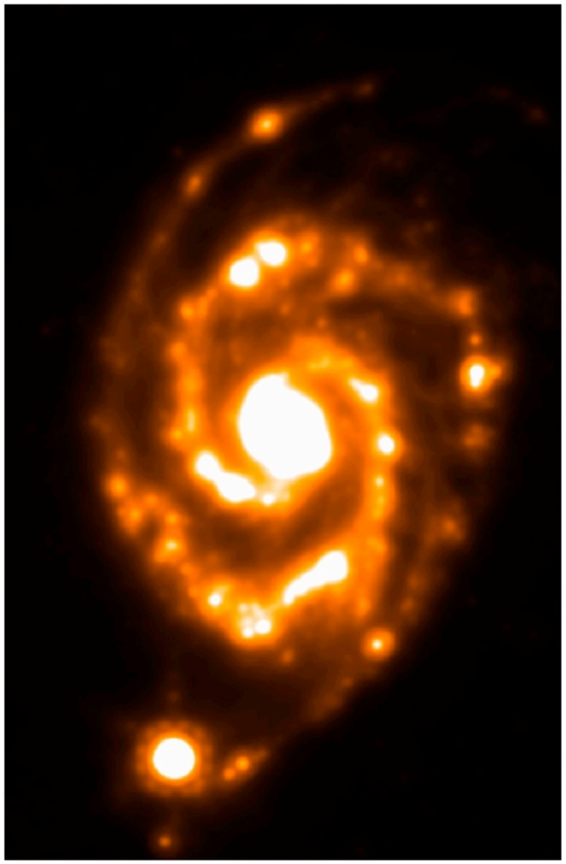
HiRes: 40 iterations

Spitzer-MIPS 24 μm from
Taurus-2 Legacy Program

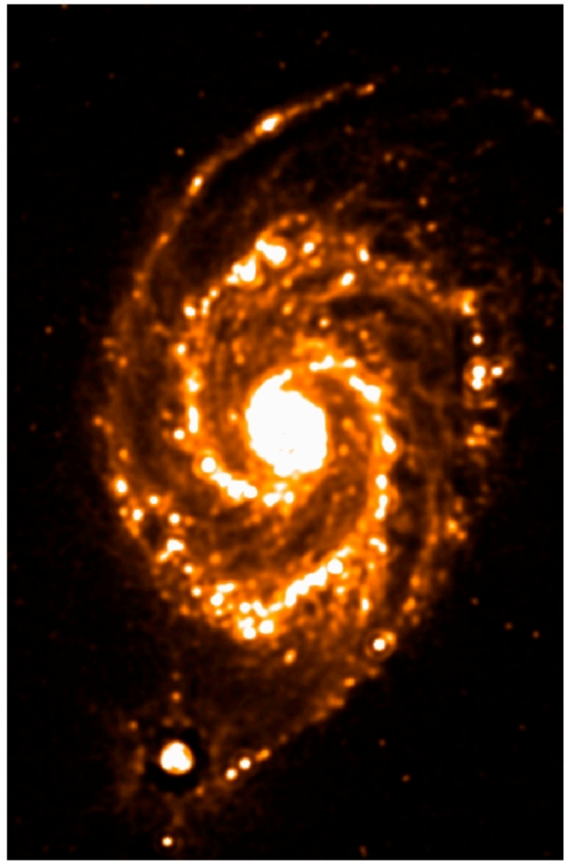
M51 or NGC 5194/95

"Whirlpool Galaxy"

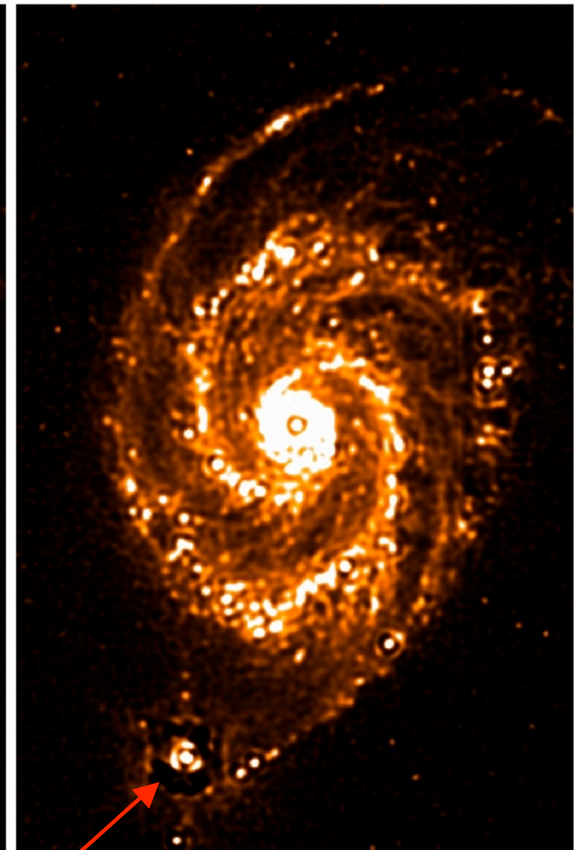
Spitzer-MIPS 24 μm



Co-add (1st iteration)



HiRes: 10 iterations



HiRes: 40 iterations

profile saturated!

M51 or NGC 5194/95

Spitzer-IRAC 5.8 μm
1st iteration Co-add from AWAIC

HST composite - NOT from AWAIC

CFV Diagnostic

- **C**orrection **F**actor **V**ariance (**CFV**) is an ancillary image product from MCM-HiRes algorithm

- Recall: correction factor for input pixel *i* at any MCM iteration:

$$C_i = \frac{\textit{measured flux}}{\textit{predicted flux}: \ \textit{PRF} \otimes \ \text{hires model}}$$

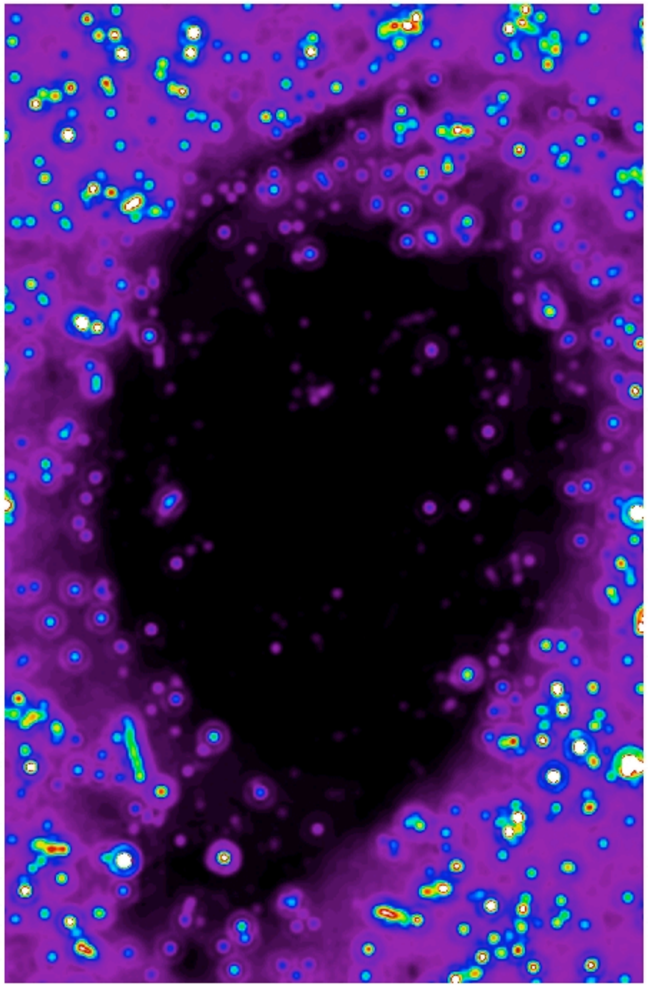
- Variance in PRF-weighted avg correction factors from all input pixels at a location in output grid

$$CFV = \langle C_i^2 \rangle - \langle C_i \rangle^2$$

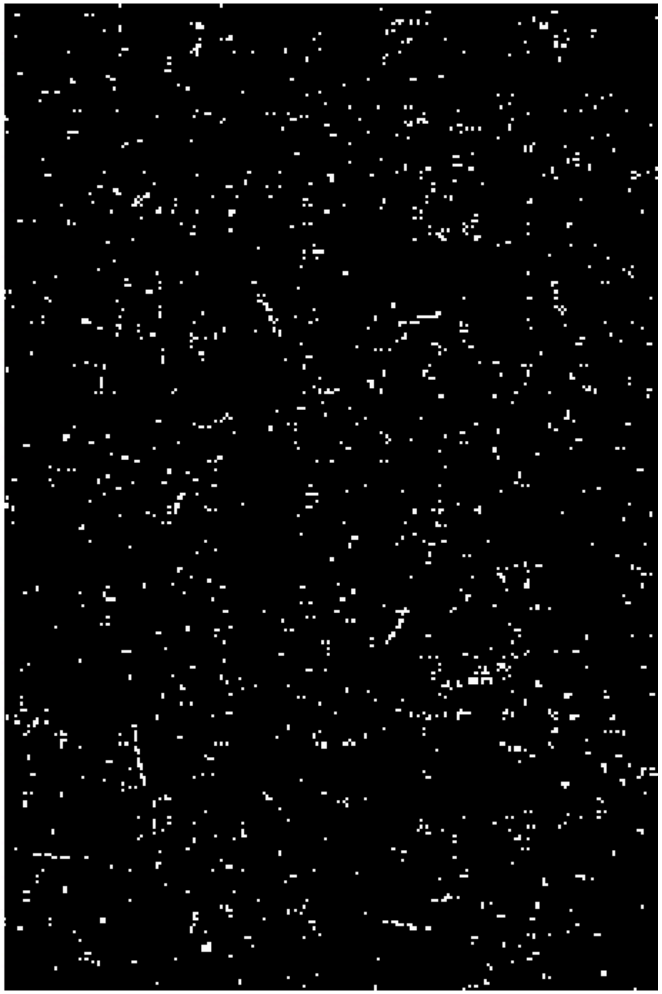
- At early iterations, CFV is everywhere high ⇒ HiRes not yet converged
  - After convergence (i.e., all $C_i \sim 1$), expect CFV ~ 0 everywhere: "spatial resolution error" minimized
  - Any remaining high values of CVF ⇒ inconsistency of input measurements at that location, e.g., outliers

- Qualitative diagnostic to indicate (i) locations in HiRes image where measurements disagree, and (ii) locations where input PRF is not a good match to the data
- Quantitative metric for computing an *a posterior* (data-derived) uncertainty for HiRes fluxes

M51: CFV and Outlier Map



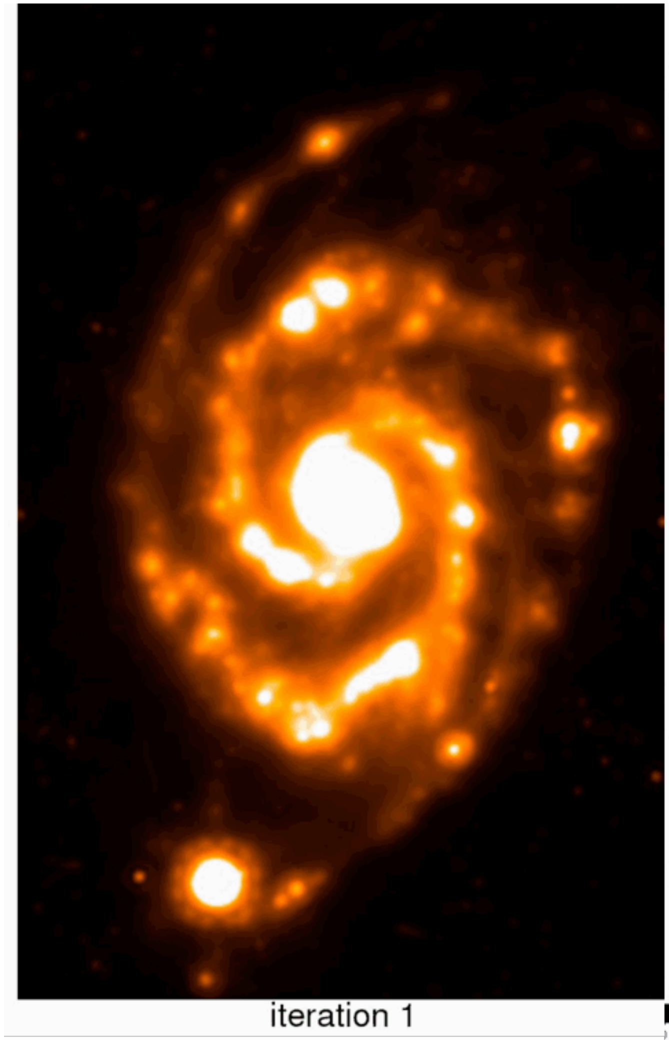
CFV after 40 iterations



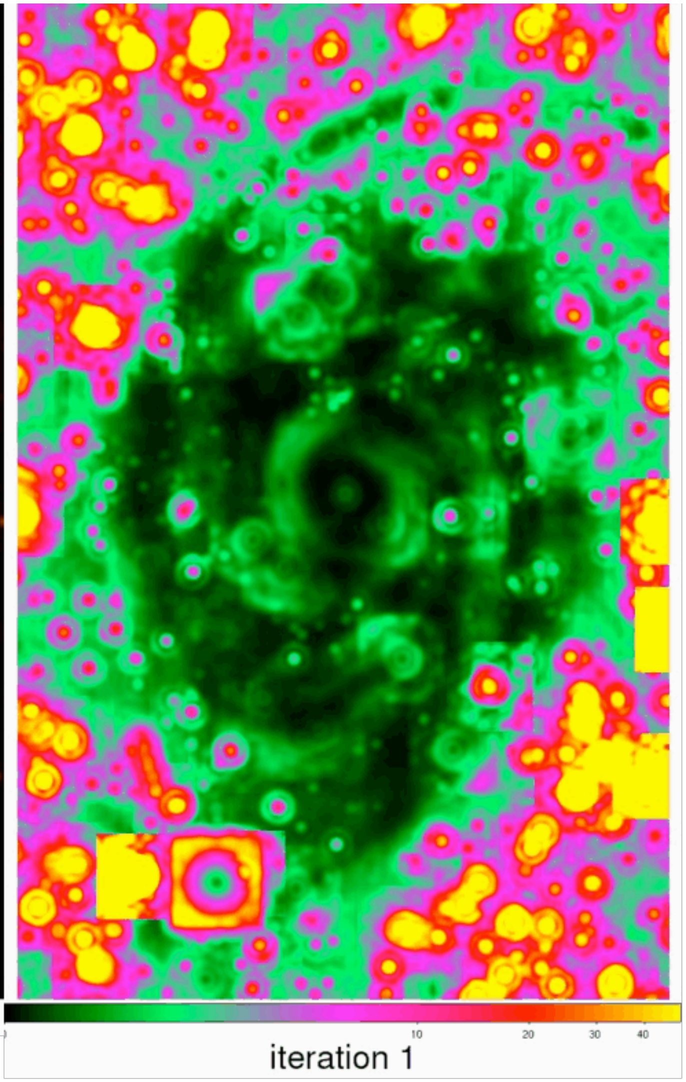
Outlier location map from stacking method

M51 movie - outliers retained

HiRes image



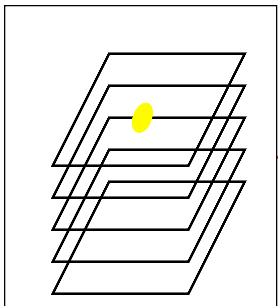
CFV image



Simulation: S/N Check in HiRes

What does HiRes do the image noise, and signal-to-noise ratio for source detection? Monte Carlo:

Simulate an ensemble of 10 images with Poisson noise
and a point source with well sampled PSF at the center

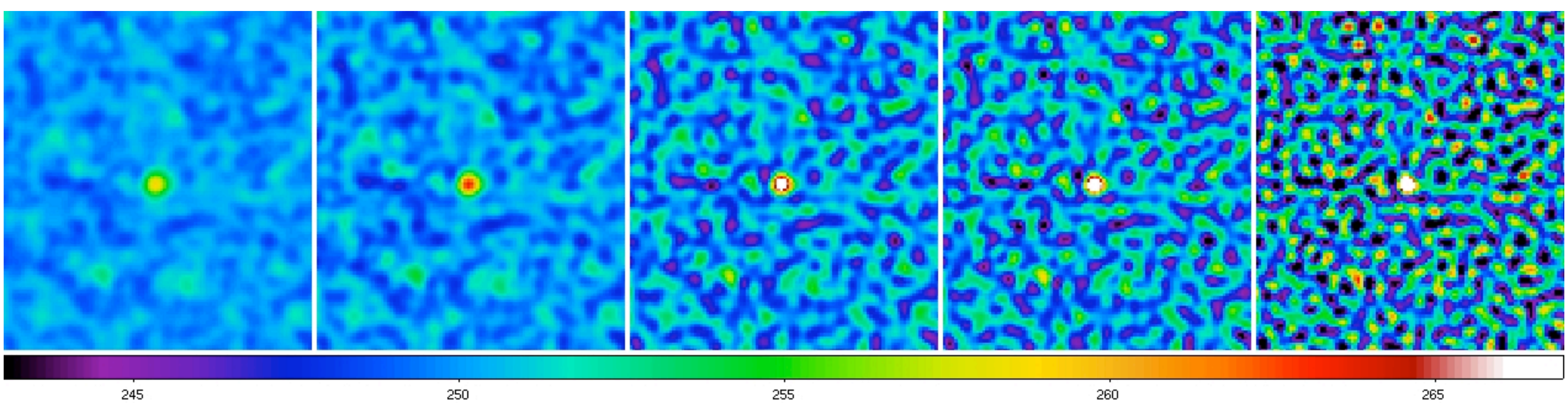


HiRes to N = 1, 2, 3… 30 iterations

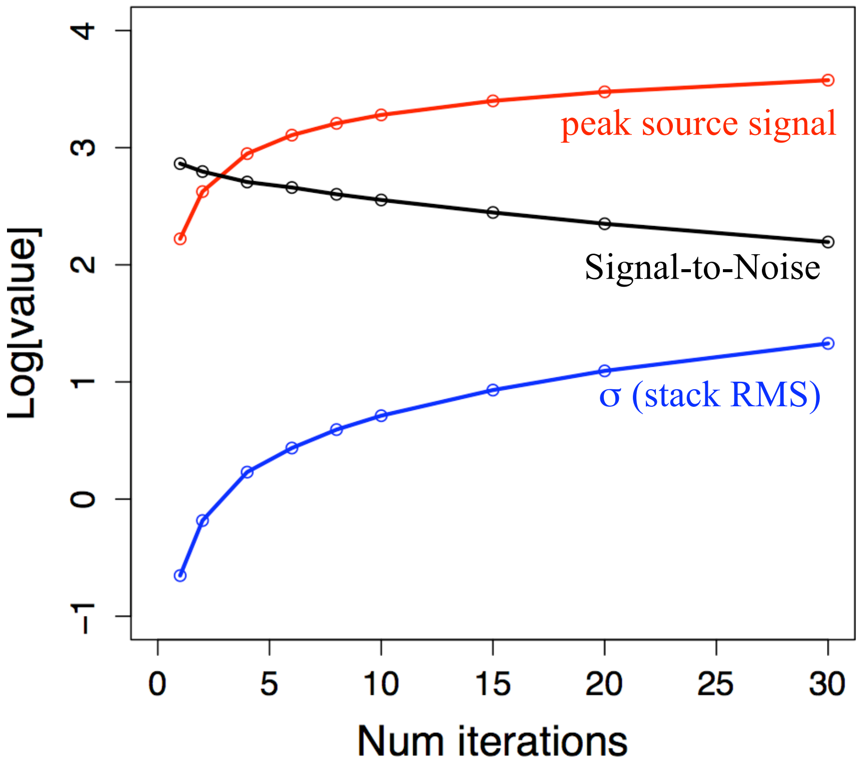
After 100 independent trials (ensembles), compute noise statistics across ensembles for each HiRes iteration

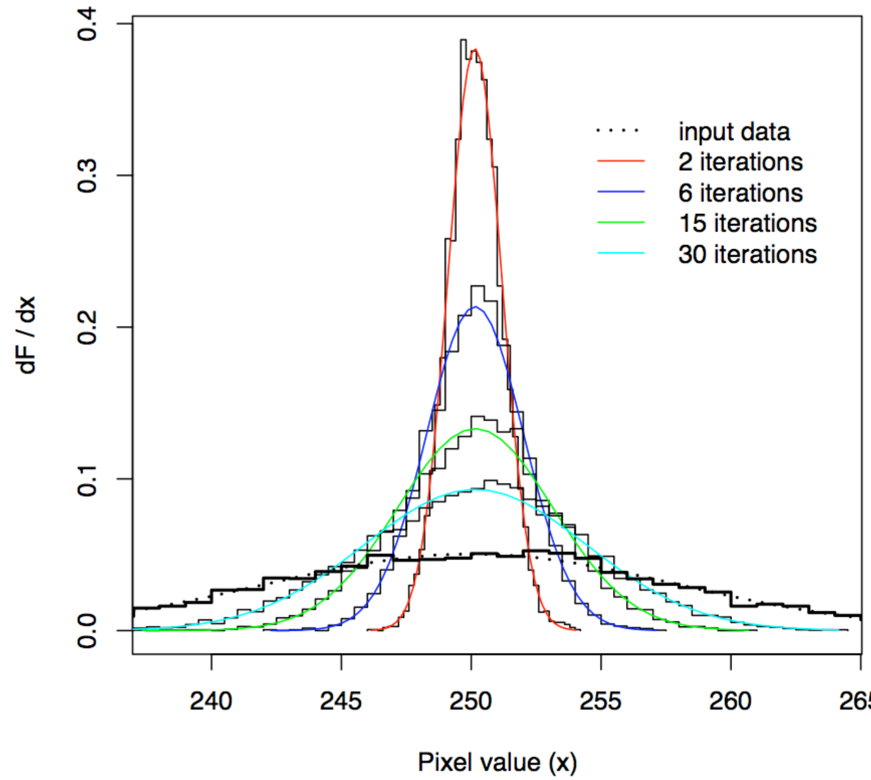
Re-simulate another trial ensemble

#iterations: 1 2 6 10 30



Simulation: S/N in HiRes





- At low iterations, power at low frequency is relatively high, i.e., noise is correlated across pixels
- With more iterations, power moves to high frequencies ⇒ de-correlation process at work
- Noise power spectrum approaches that of the input data depending on PRF accuracy

Summary and Future Plans

- Described co-addition framework for WISE with extension to resolution enhancement
- Provides a generic tool for use on any image data that conforms with FITS/WCS standards
- Goal is to produce high-fidelity, science quality image products for accurate photometry with quantifiable uncertainties

- Currently AWAIC is a suite of modules implemented in ANSI C and wrapped into a Perl script
  - Runs under Linux in WISE processing environment
  - Implement a platform independent version for portability to the community

- Explore methods for accelerating convergence in MCM (currently converges logarithmically)

- Extend to handle time dependent PSFs (e.g., adapted to seeing). This has applications for ground based projects, e.g., LSST. PSF matching is important for time-domain studies

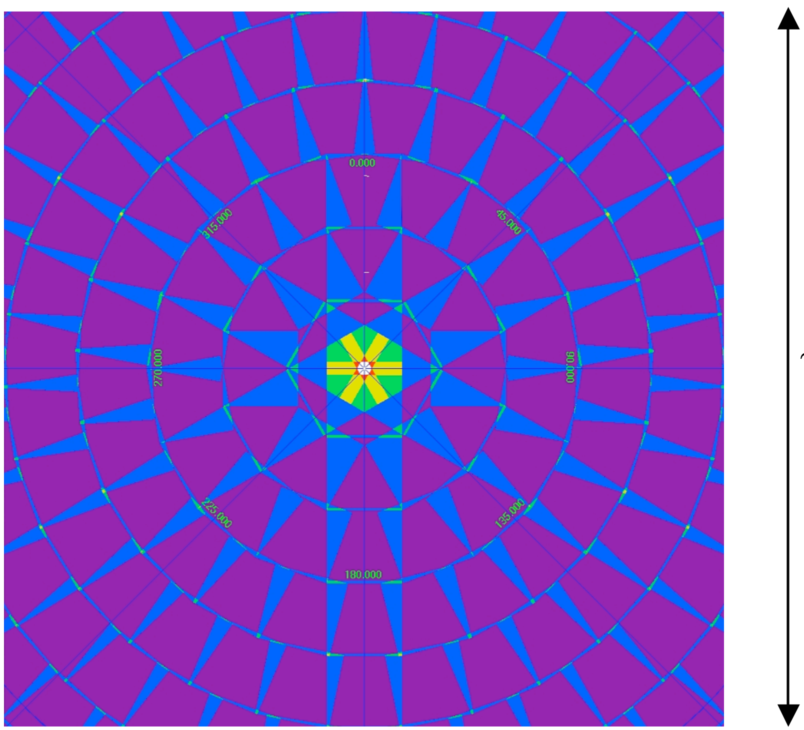
- Explore performance of MCM on confusion limited observations: how far below the native confusion limit can we go and reliably detect sources?

- More thorough explanation of all algorithms can be found at:
  http://web.ipac.caltech.edu/staff/fmasci/home/wise/awaic.html

Backup Slides

Atlas Image Tiling Geometry

Example of tiling pattern (or co-add image footprints) over an equatorial pole:



~15°

Tile overlaps:

- Purple ⇒1
- Blue ⇒ 2
- Green ⇒ 3
- Yellow ⇒ 4
- Red ⇒ 5
- White ⇒ 6 (on pole)

Background-level Matching

- Instrumental and detector transients lead to varying background levels between frames
- **Goal:** obtain seamless (or smooth) transitions between frames across overlaps in a co-add
- Want to equalize background levels but preserve natural background variations if possible
- Make each Atlas Image co-add self-consistent for scientific purposes
- Later tie together and match levels in co-adds across sky if needed

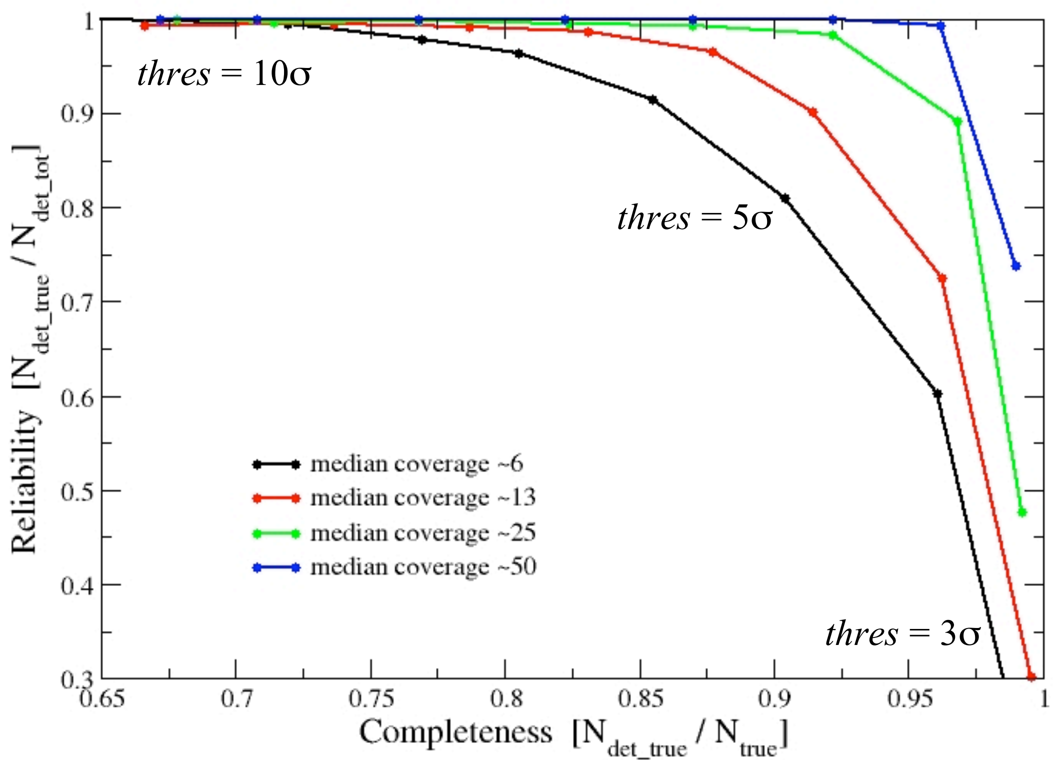
Simple Method:

1. Fit a plane to each input frame that overlaps with co-add footprint to capture "global" level
   - Fitting is done "robustly", i.e., ~ immune to presence of bright sources and extended structure
2. Subtract robust planar fits from each respective frame ⇒ places frames on a zero baseline
3. Compute a global median (or modal) plane from all fits and extend over co-add footprint
4. Add this "common plane" to all the input frames

⇒ Ensures continuity of background across footprint region after co-addition

Outlier Detection

- Performed a simulation containing known cosmic ray hits and noise to explore completeness and reliability as a function of depth-of-coverage and outlier detection threshold
- For depths-of-coverage >~ 10, completeness and reliability are reasonable for a threshold of ~5σ



- Moving objects, e.g., asteroids and highly variable sources will be flagged as outliers in WISE co-adds unless they're moving (or varying) slowly across overlapping frames

  ⇒ co-adds will represent the "static" inertial sky

Ringing Suppression Algorithm

Input image frames

Each frame: compute background using block median or mode filter over N x N grid w/ Gaussian smoothing

Mosaic the background frames

- subtract background frames from input frames
- reset all negative pixel values to zero
- add small positive offset $\delta \sim 10^{-20}$

($\Rightarrow$ positivity constraint with a "flux bias")

Using a flat model (starting) image, HiRes bckgnd-sub to *N* iterations until convergence

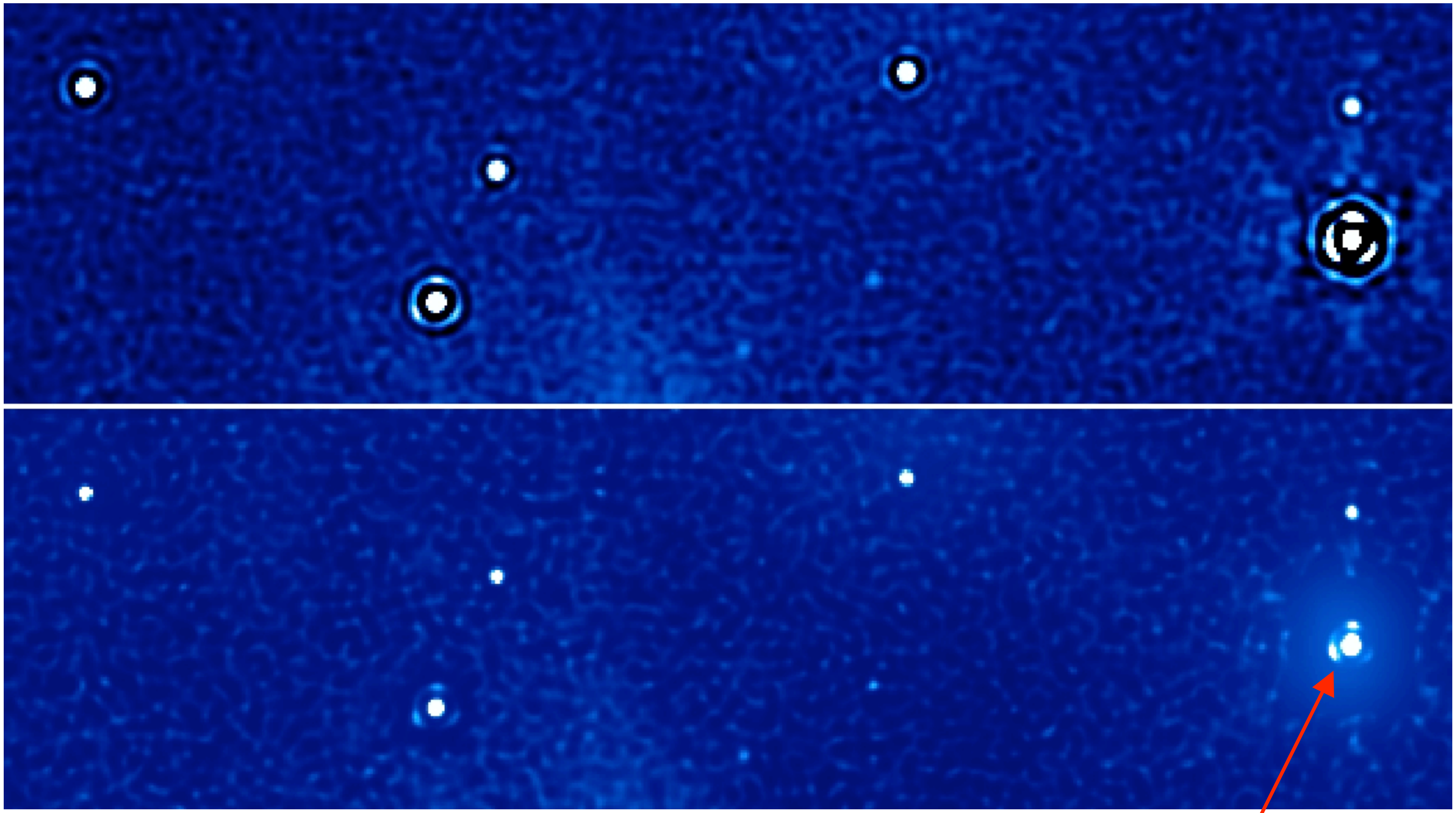
Add background mosaic to HiRes products

Using new HiRes as starting image and original frames (w/ bckgnds), HiRes again to 4-10 iterations to restore (de-bias) noise structure and reproduce measurements

Output HiRes'd mosaic with ringing suppressed. Residual ringing depends on complexity of background

Ringing Suppression (Field in Taurus)

suppression off

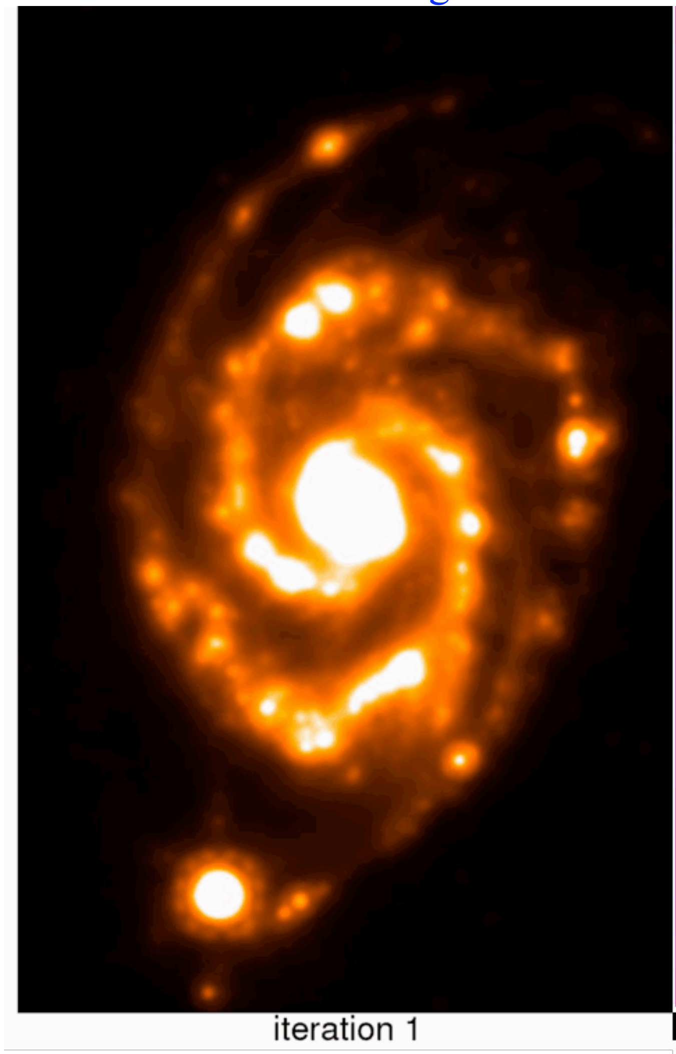


suppression on

profile saturated

M51 movie - outliers first rejected

HiRes image



CFV image