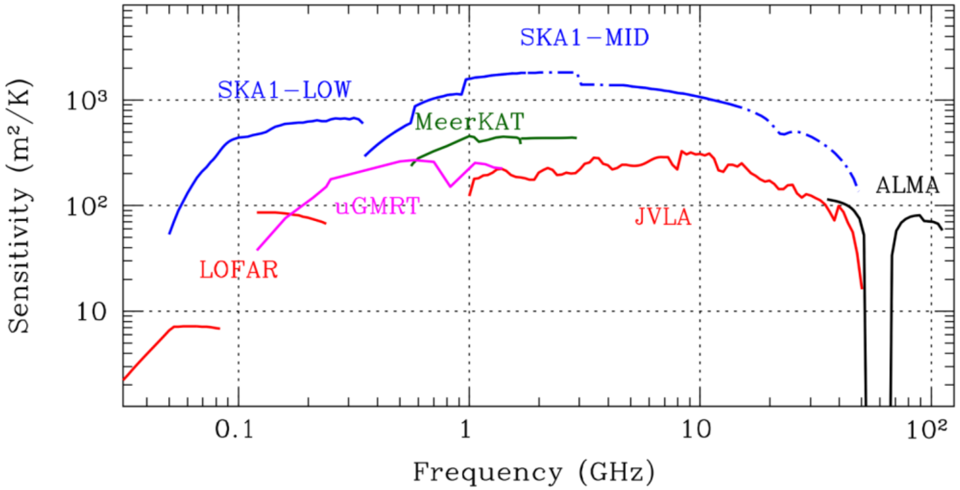
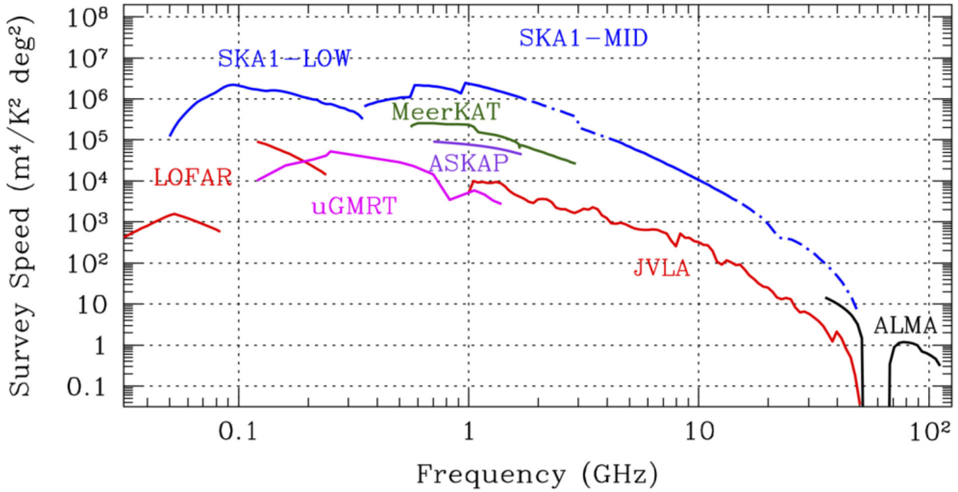
Decentralizing radio-interferometric image reconstruction by spatial frequency

Sunrise Wang, Simon Prunet, Shan Mignot André Ferrari

OBSERVATOIRE DE LA CÔTE D'AZUR
UNIVERSITÉ CÔTE D'AZUR
RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité
LAGRANGE
DARK ERA

Introduction





SKA-Low

SKA-Mid

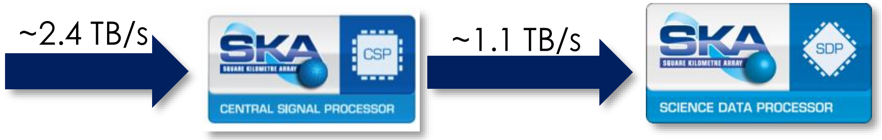
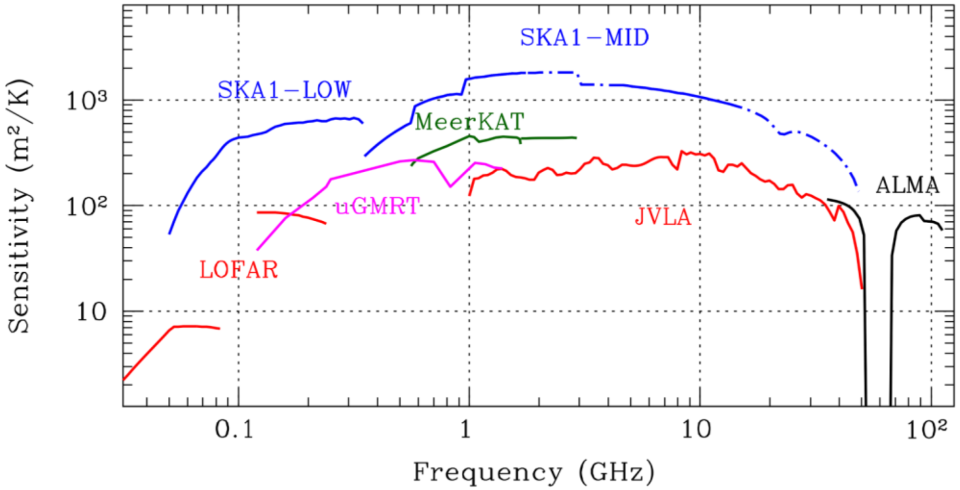
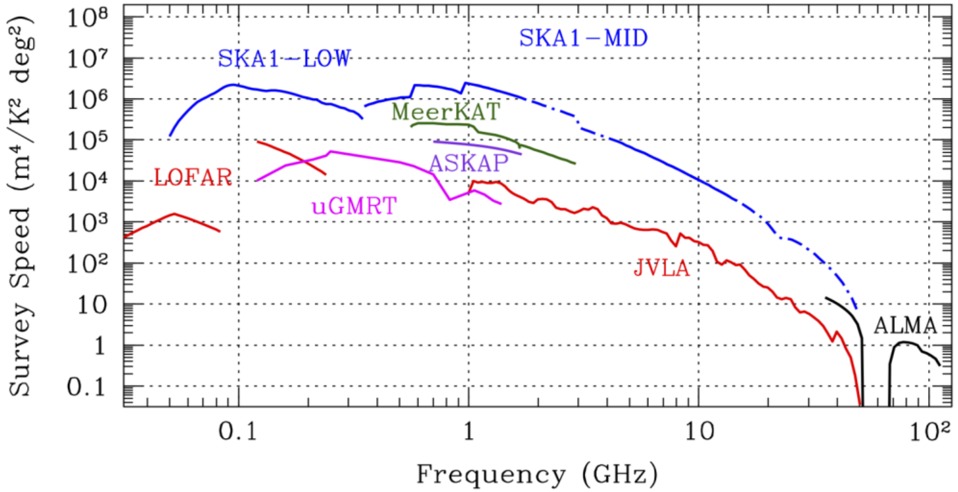


image sources: [3]
data source: [1, 2]

[1]Labate, Maria G., et al. "Highlights of the square kilometre array low frequency (SKA-LOW) telescope." Journal of Astronomical Telescopes, Instruments, and Systems 8.1 (2022): 011024-011024.
[2] Swart, Gerhard P., Peter E. Dewdney, and Andrea Cremonini. "Highlights of the SKA1-Mid telescope architecture." Journal of Astronomical Telescopes, Instruments, and Systems 8.1 (2022): 011021-011021.
[3] https://www.skao.int

Introduction





SKA-Low

~0.7 TB/s → SKA CSP CENTRAL SIGNAL PROCESSOR → ~0.3 TB/s → SKA SDP SCIENCE DATA PROCESSOR

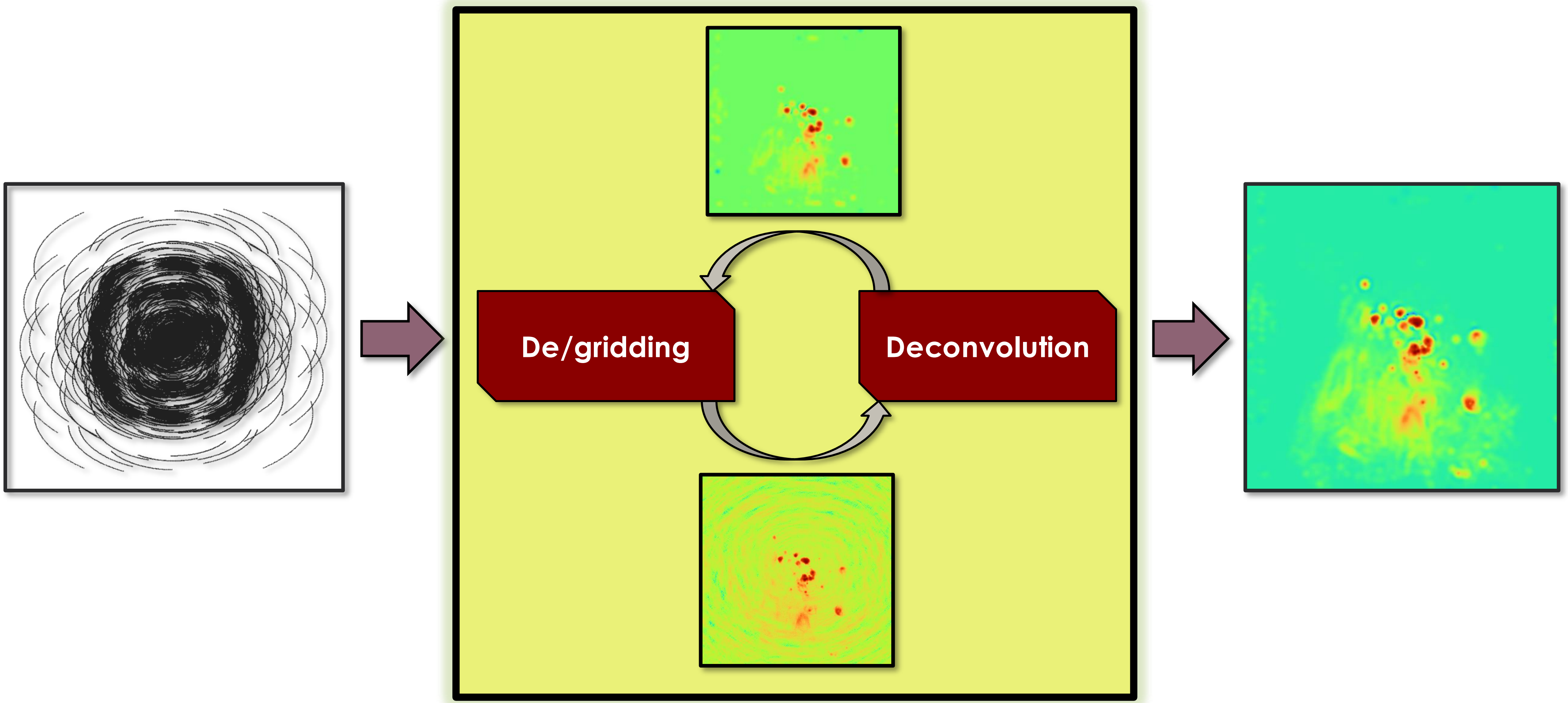
SKA-Mid

~2.4 TB/s → SKA CSP CENTRAL SIGNAL PROCESSOR → ~1.1 TB/s → SKA SDP SCIENCE DATA PROCESSOR

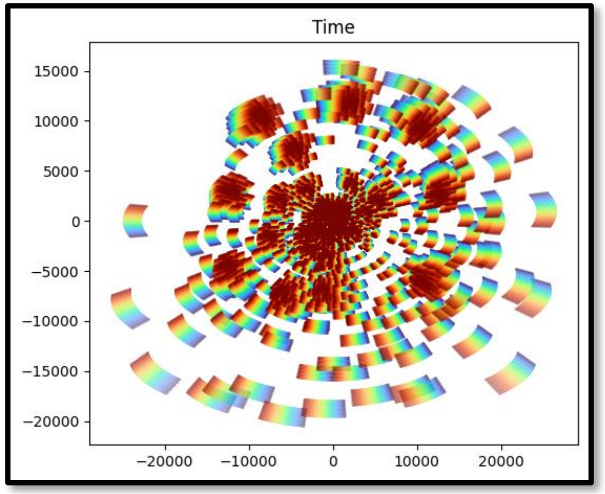
image sources: [3]
data source: [1, 2]

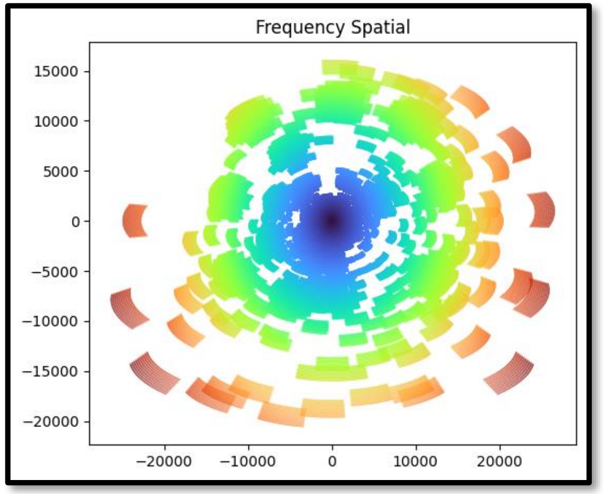
[1] Labate, Maria G., et al. "Highlights of the square kilometre array low frequency (SKA-LOW) telescope." Journal of Astronomical Telescopes, Instruments, and Systems 8.1 (2022): 011024-011024.
[2] Swart, Gerhard P., Peter E. Dewdney, and Andrea Cremonini. "Highlights of the SKA1-Mid telescope architecture." Journal of Astronomical Telescopes, Instruments, and Systems 8.1 (2022): 011021-011021.
[3] https://www.skao.int

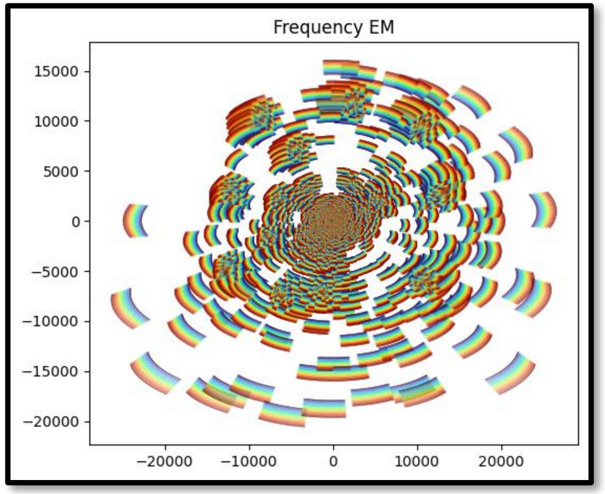
Major-Minor Loop Reconstruction



Scaling the computation

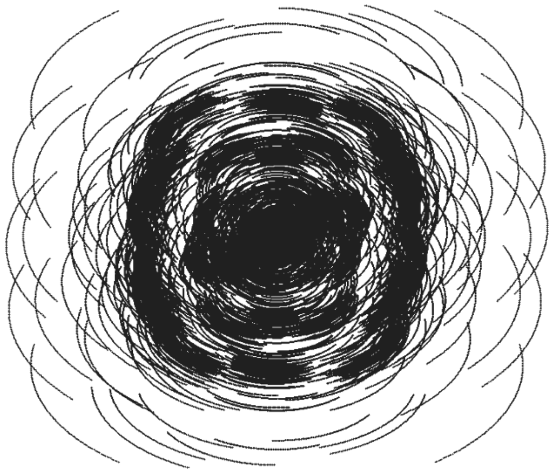






- Partition visibilities, process separately
- Commonly by time and em frequency, partitioning relatively trivial
- We look at spatial frequency, more complicated as not all frequencies are available for deconvolution

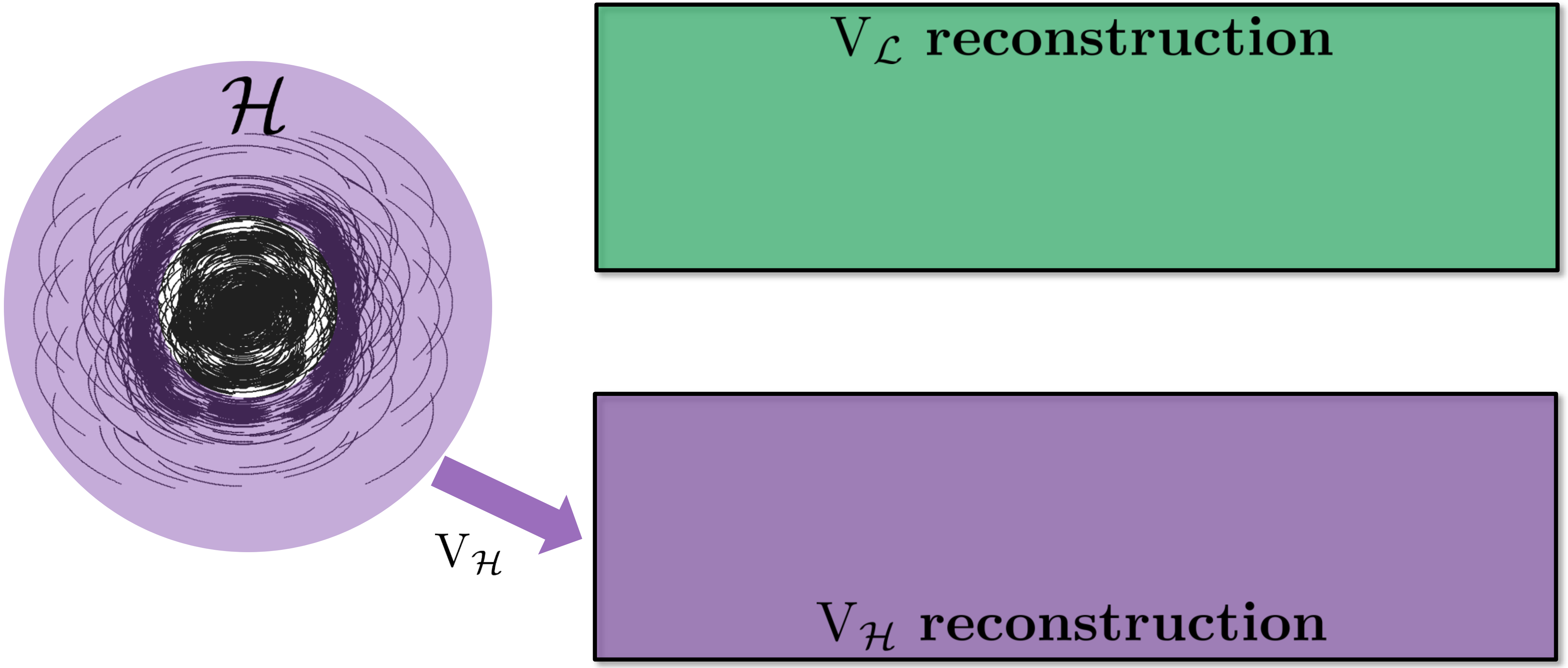
Parallelization Framework



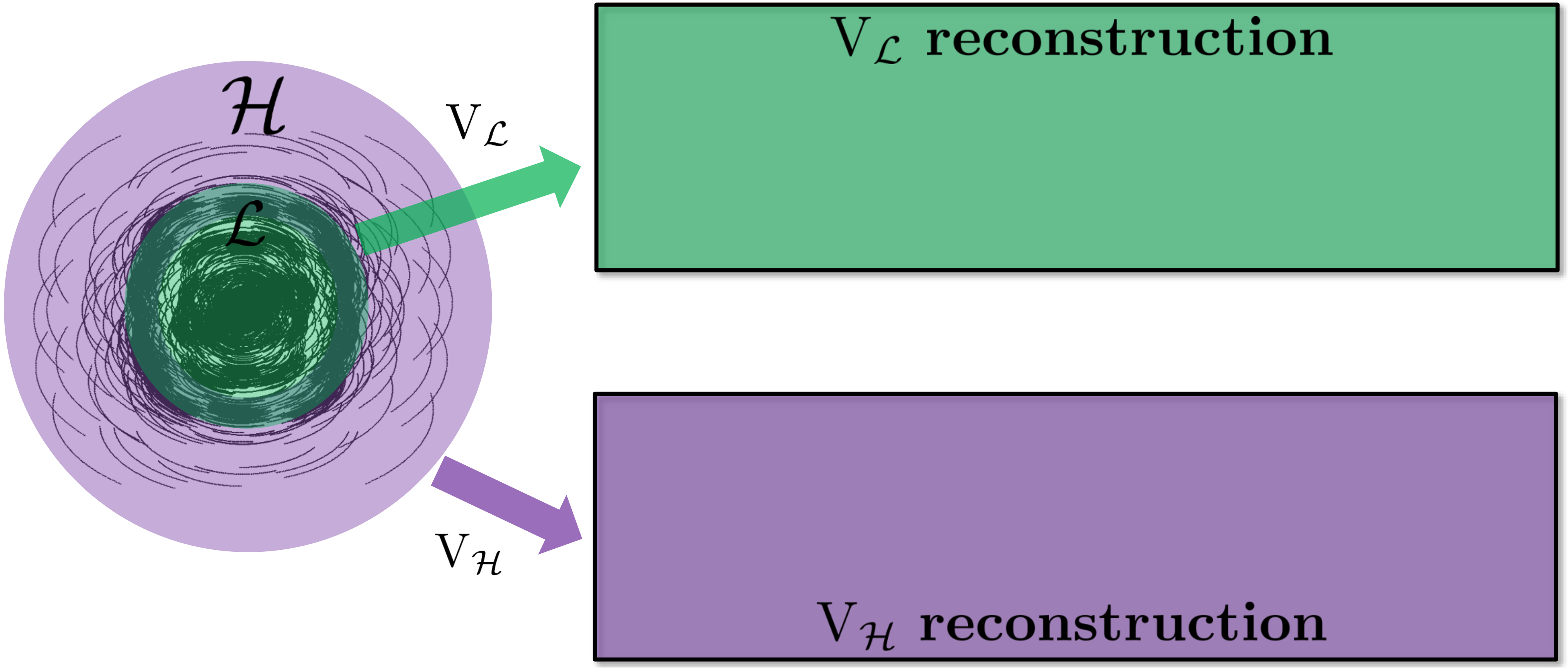
$\mathrm{V}_{\mathcal{L}}$ **reconstruction**

$\mathrm{V}_{\mathcal{H}}$ **reconstruction**

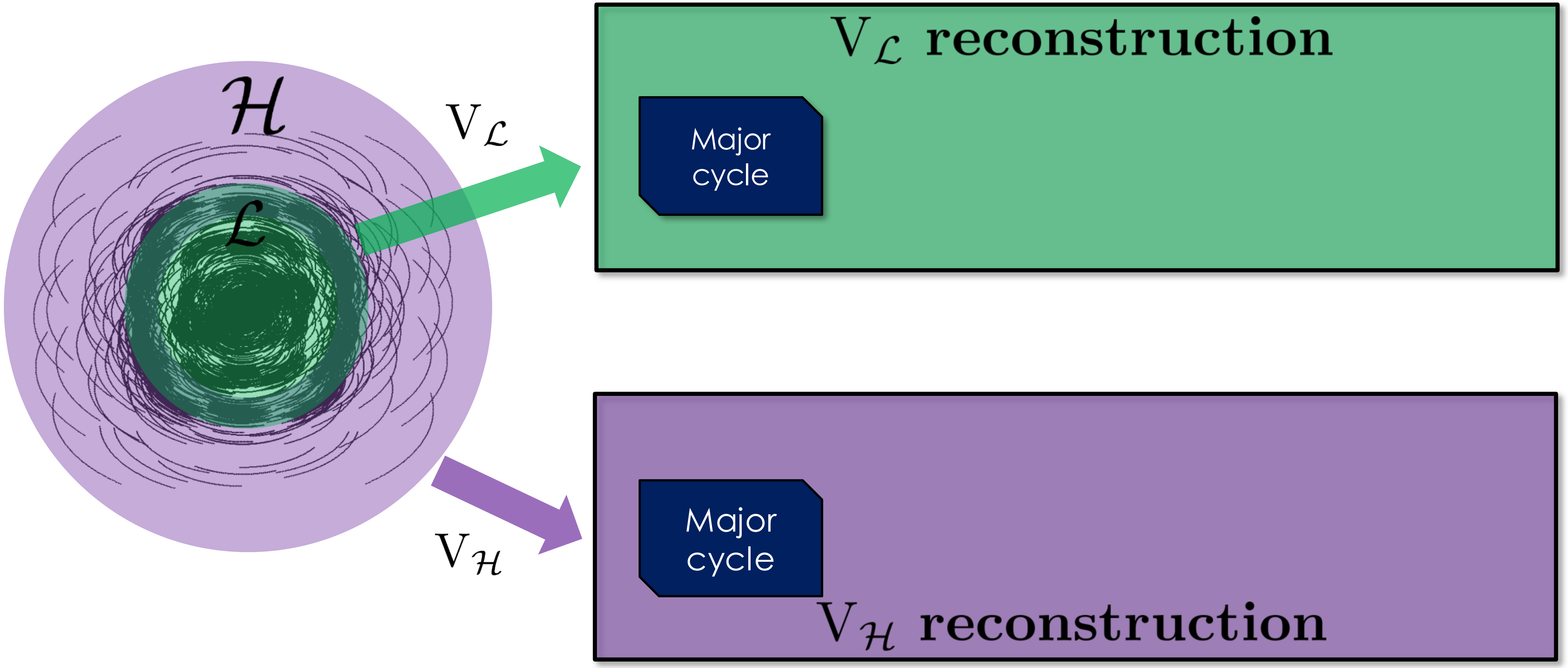
Parallelization Framework



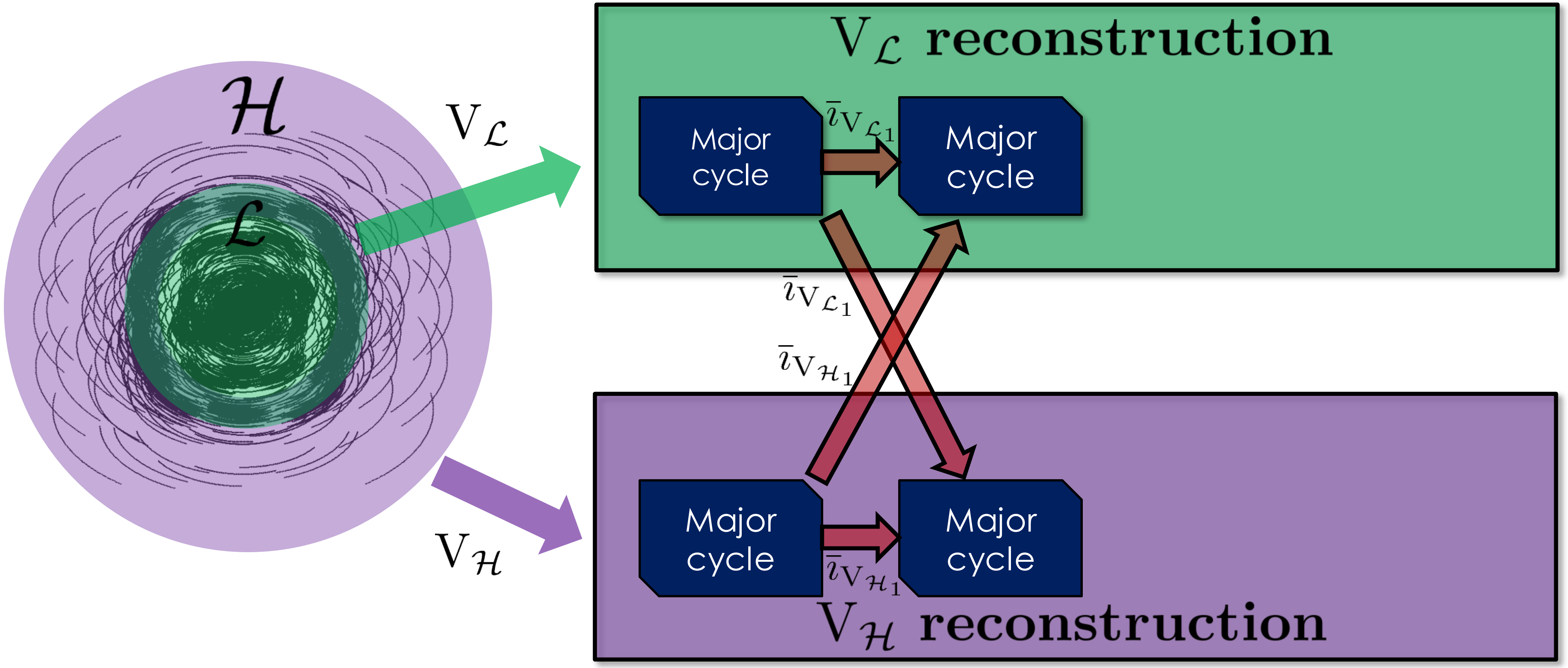
Parallelization Framework



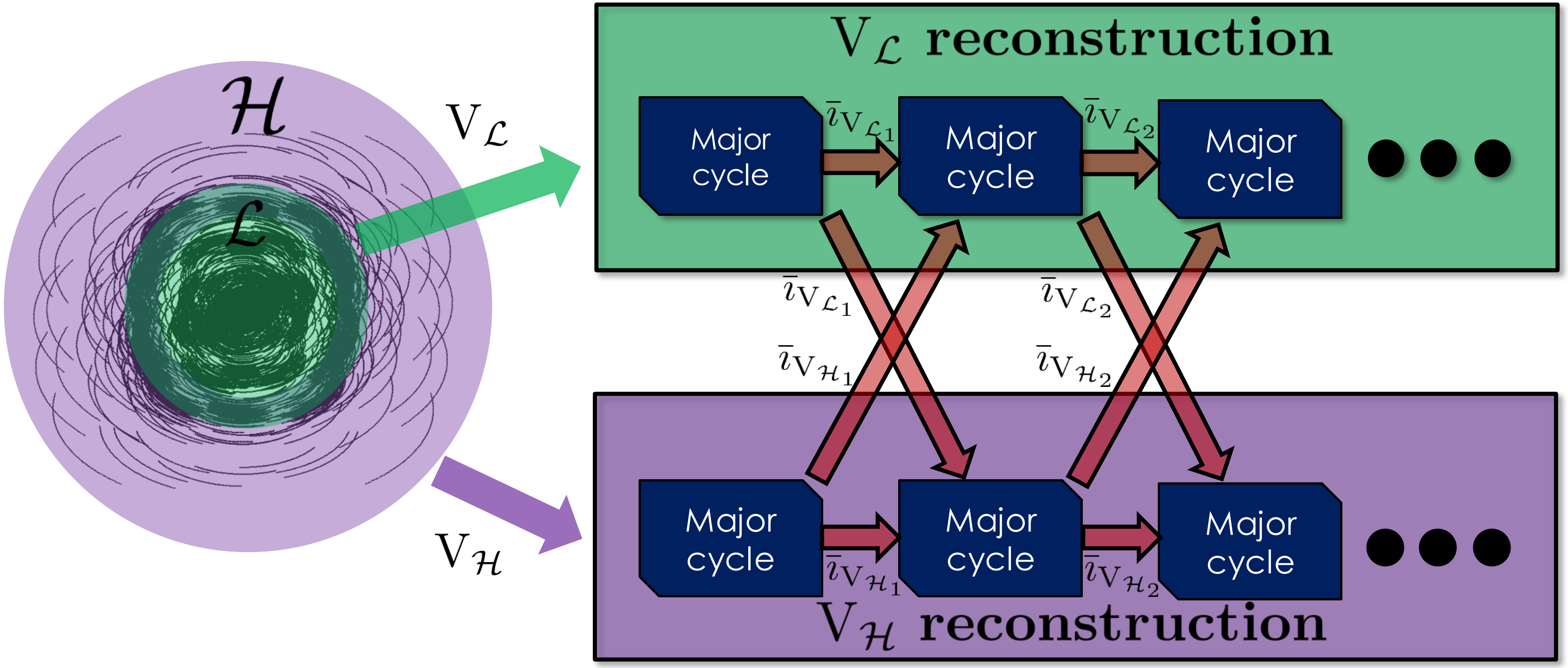
Parallelization Framework



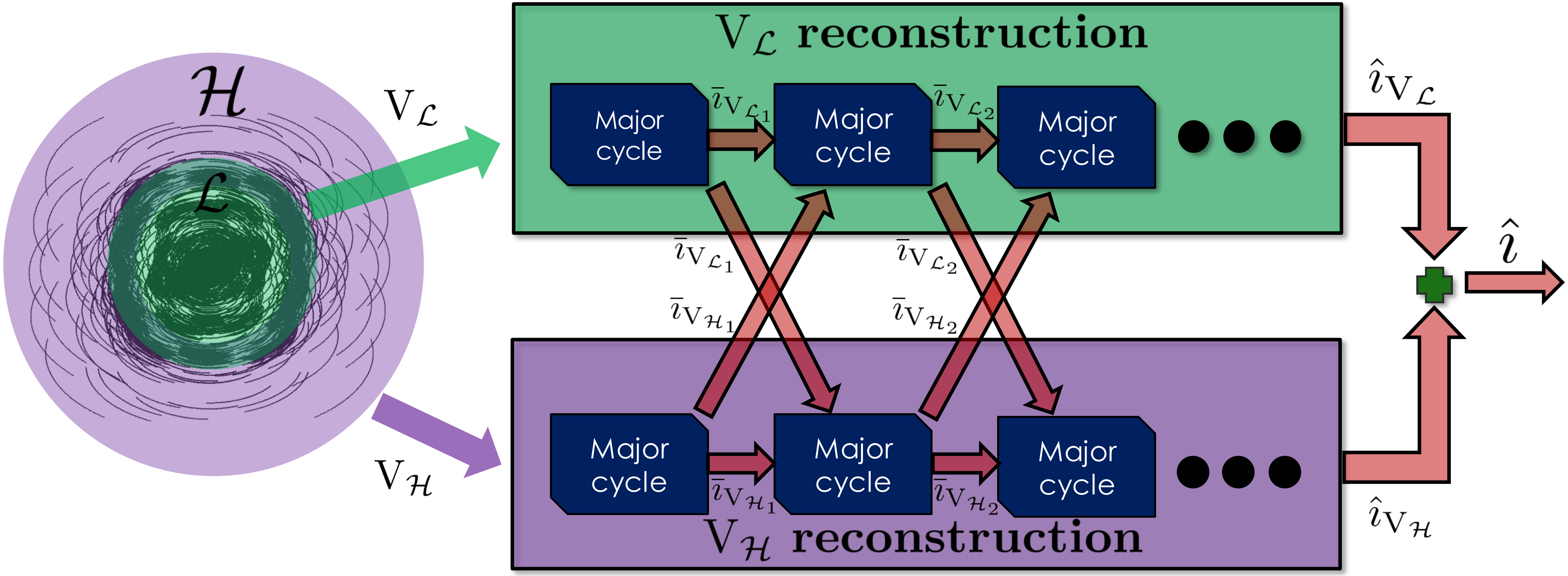
Parallelization Framework



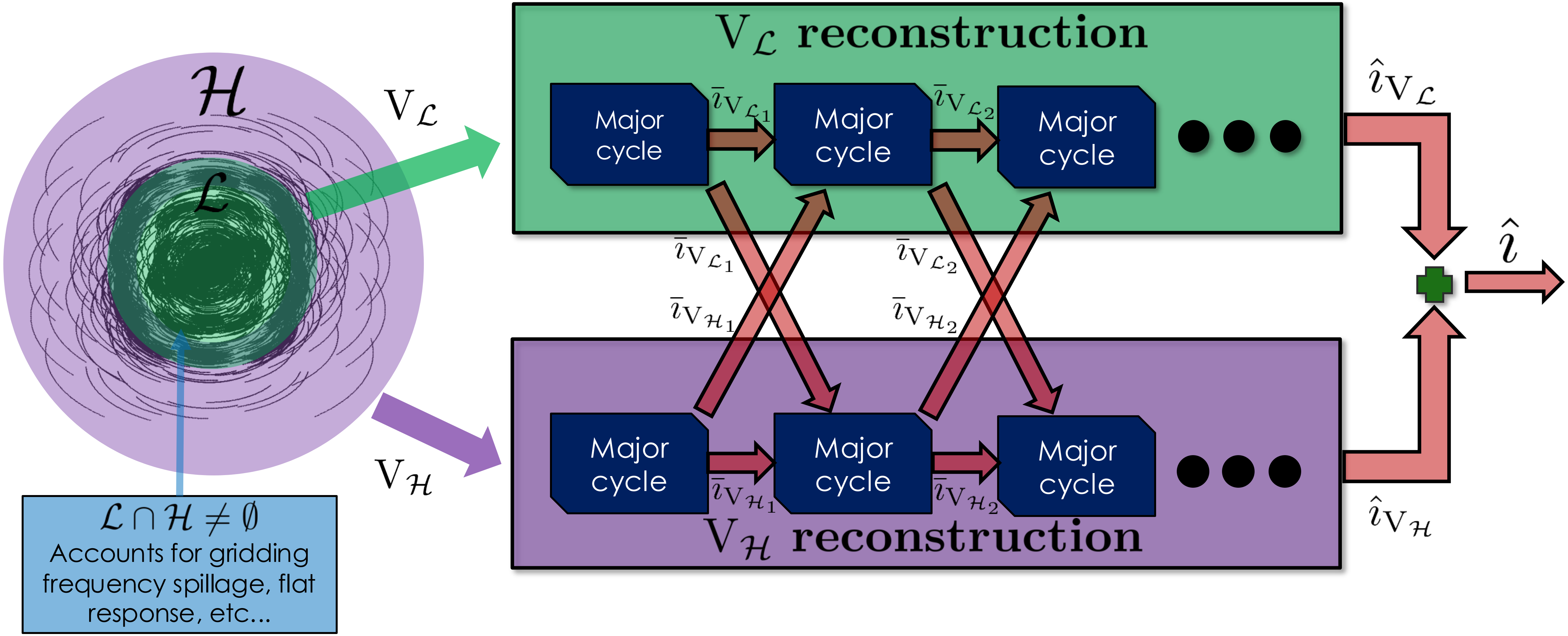
Parallelization Framework



Parallelization Framework



Parallelization Framework



Example 1: Parallelized L1 reconstruction

Deconvolution framework for every major cycle n, similar to [1, 2]

$$\alpha_n = \arg\min_{\alpha} \|\tilde{\imath}_n - HW\alpha\|_2^2 + \lambda_n \|\alpha\|_1$$

$$\bar{\imath}_n = W\alpha_n$$

[1] Wiaux, Yves, et al. "Compressed sensing imaging techniques for radio interferometry." Monthly Notices of the Royal Astronomical Society 395.3 (2009): 1733-1742.
[2] Garsden, Hugh, et al. "LOFAR sparse image reconstruction." Astronomy & astrophysics 575 (2015): A90.

Example 1: Parallelized L1 reconstruction

Convolution by PSF operator

Deconvolution framework for every major cycle n, similar to [1, 2]

$$\alpha_n = \arg\min_{\alpha} \|\tilde{\imath}_n - HW\alpha\|_2^2 + \lambda_n \|\alpha\|_1$$

$$\bar{\imath}_n = W\alpha_n$$

Wavelet transform operator (Daubechies 1-8)

Used over $\|v - G^{\dagger}FW\alpha\|_2^2$ for efficiency
Assumes $Hi \approx F^{\dagger}GG^{\dagger}Fi$
Errors corrected in major cycle

[1] Wiaux, Yves, et al. "Compressed sensing imaging techniques for radio interferometry." Monthly Notices of the Royal Astronomical Society 395.3 (2009): 1733-1742.
[2] Garsden, Hugh, et al. "LOFAR sparse image reconstruction." Astronomy & astrophysics 575 (2015): A90.

Example 1: Parallelized L1 reconstruction

Deconvolution framework for every major cycle n, similar to [1, 2]

$$\alpha_n = \arg\min_{\alpha} \|\tilde{\imath}_n - HW\alpha\|_2^2 + \lambda_n \|\alpha\|_1$$

$$\bar{\imath}_n = W\alpha_n$$

$\mathrm{V}_{\mathcal{L}}$ deconvolution

$$\alpha_{\mathrm{V}_{\mathcal{L}_n}} = \arg\min_{\alpha} \|G_{\mathcal{L}}(\tilde{\imath}_{\mathcal{L}_n} - H_{\mathcal{L}}W\alpha)\|_2^2 + \|G_{\mathcal{H}}(h_n - W\alpha)\|_2^2 + \lambda_{\mathrm{V}_{\mathcal{L}_n}} \|\alpha\|_1$$

$$\bar{\imath}_{\mathrm{V}_{\mathcal{L}_n}} = W\alpha_{\mathrm{V}_{\mathcal{L}_n}}, h_n = \sum_{j=1}^{n-1} \bar{\imath}_{\mathrm{V}_{\mathcal{H}_j}} - \sum_{j=1}^{n-1} \bar{\imath}_{\mathrm{V}_{\mathcal{L}_n}} = \hat{\imath}_{\mathrm{V}_{\mathcal{H}_{n-1}}} - \hat{\imath}_{\mathrm{V}_{\mathcal{L}_{n-1}}}$$

[1] Wiaux, Yves, et al. "Compressed sensing imaging techniques for radio interferometry." Monthly Notices of the Royal Astronomical Society 395.3 (2009): 1733-1742.
[2] Garsden, Hugh, et al. "LOFAR sparse image reconstruction." Astronomy & astrophysics 575 (2015): A90.

Example 1: Parallelized L1 reconstruction

Deconvolution framework for every major cycle n, similar to [1, 2]

$$\alpha_n = \arg\min_{\alpha} \|\tilde{i}_n - HW\alpha\|_2^2 + \lambda_n \|\alpha\|_1$$

$$\bar{i}_n = W\alpha_n$$

Visibility data-fidelity term

Additional data-fidelity term for rest of frequency information (only from 2nd major cycle onwards)

$\mathrm{V}_{\mathcal{L}}$ deconvolution

$$\alpha_{\mathrm{V}_{\mathcal{L}_n}} = \arg\min_{\alpha} \|G_{\mathcal{L}}(\tilde{i}_{\mathcal{L}_n} - H_{\mathcal{L}}W\alpha)\|_2^2 + \|G_{\mathcal{H}}(h_n - W\alpha)\|_2^2 + \lambda_{\mathrm{V}_{\mathcal{L}_n}} \|\alpha\|_1$$

$$\bar{i}_{\mathrm{V}_{\mathcal{L}_n}} = W\alpha_{\mathrm{V}_{\mathcal{L}_n}}, h_n = \sum_{j=1}^{n-1} \bar{i}_{\mathrm{V}_{\mathcal{H}_j}} - \sum_{j=1}^{n-1} \bar{i}_{\mathrm{V}_{\mathcal{L}_n}} = \hat{i}_{\mathrm{V}_{\mathcal{H}_{n-1}}} - \hat{i}_{\mathrm{V}_{\mathcal{L}_{n-1}}}$$

High and low resolution filters
Normalizes data fidelity terms
Ensures visibilities in $\mathcal{L} \cap \mathcal{H}$ sum to 1

[1] Wiaux, Yves, et al. "Compressed sensing imaging techniques for radio interferometry." Monthly Notices of the Royal Astronomical Society 395.3 (2009): 1733-1742.
[2] Garsden, Hugh, et al. "LOFAR sparse image reconstruction." Astronomy & astrophysics 575 (2015): A90.

Example 2: Parallelized MS-CLEAN reconstruction

CLEAN iteratively removes the brightest source at the most relevant scale convolved by the PSF from the residual[1]. We can denote this as:

$$\bar{\imath} = \text{MS-CLEAN}(\tilde{\imath}, H)$$

[1] [2]Cornwell, Tim J. "Multiscale CLEAN deconvolution of radio synthesis images." *IEEE Journal of selected topics in signal processing* 2.5 (2008): 793-801.

Example 2: Parallelized MS-CLEAN reconstruction

CLEAN iteratively removes the brightest source at the most relevant scale convolved by the PSF from the residual[1]. We can denote this as:

$$\bar{\imath} = \text{MS-CLEAN}(\tilde{\imath}, H)$$

$\mathrm{V}_{\mathcal{L}}$ deconvolution

$$\bar{\imath}_{\mathrm{V}_{\mathcal{L}n}} = \text{MS-CLEAN}(\alpha G_{\mathcal{L}} \tilde{\imath}_{\mathcal{L}_n} + \beta G_{\mathcal{H}} H_{\mathcal{H}} h_n, \alpha G_{\mathcal{L}} H_{\mathcal{L}} - \beta G_{\mathcal{H}} H_{\mathcal{H}})$$

$$h_n = \sum_{j=1}^{n-1} \bar{\imath}_{\mathrm{V}_{\mathcal{H}}} - \sum_{j=1}^{n-1} \bar{\imath}_{\mathrm{V}_{\mathcal{L}}}$$

[1] [2]Cornwell, Tim J. "Multiscale CLEAN deconvolution of radio synthesis images." *IEEE Journal of selected topics in signal processing* 2.5 (2008): 793-801.

Example 2: Parallelized MS-CLEAN reconstruction

CLEAN iteratively removes the brightest source at the most relevant scale convolved by the PSF from the residual[1]. We can denote this as:

$$\bar{\imath} = \text{MS-CLEAN}(\tilde{\imath}, H)$$

Pseudo full-resolution dirty

Pseudo full-resolution PSF

$\mathrm{V}_{\mathcal{L}}$ deconvolution

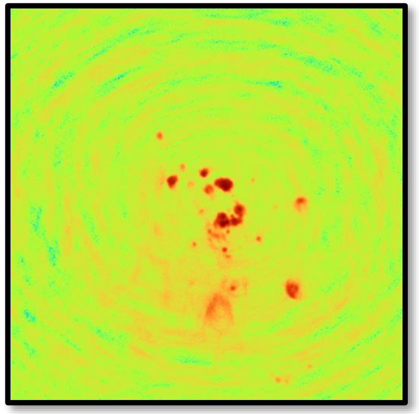
$$\bar{\imath}_{\mathrm{V}_{\mathcal{L}n}} = \text{MS-CLEAN}(\alpha G_{\mathcal{L}} \tilde{\imath}_{\mathcal{L}_n} + \beta G_{\mathcal{H}} H_{\mathcal{H}} h_n, \alpha G_{\mathcal{L}} H_{\mathcal{L}} - \beta G_{\mathcal{H}} H_{\mathcal{H}})$$

$$h_n = \sum_{j=1}^{n-1} \bar{\imath}_{\mathrm{V}_{\mathcal{H}}} - \sum_{j=1}^{n-1} \bar{\imath}_{\mathrm{V}_{\mathcal{L}}}$$

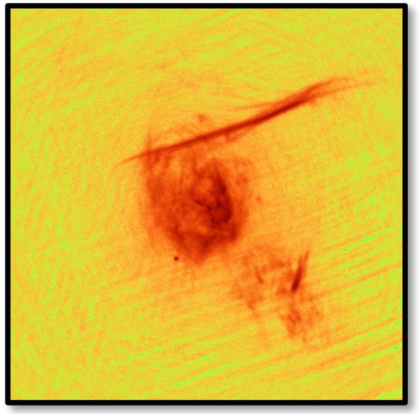
[1] [2]Cornwell, Tim J. "Multiscale CLEAN deconvolution of radio synthesis images." *IEEE Journal of selected topics in signal processing* 2.5 (2008): 793-801.

Experiment Datasets

Simulated

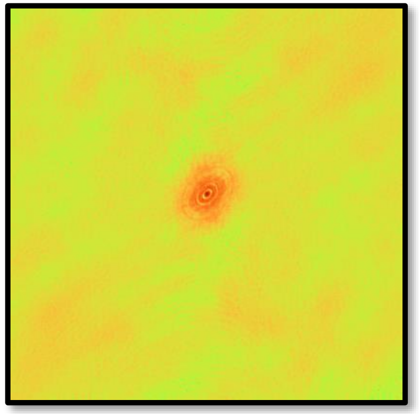


Sgr B2



Sgr C

Real



HL Tau

Cygnus A

- Initial images tapered and cutout from 1.28GHz mosaic produced in [1]
- Visibilities generated with SKA-Mid AA4, and SKA-Low AA4 configurations
- Observation time of HA=[-2,2] with integration times of 30s, and 120s respectively, 1 channel at a pseudo-frequency of 1GHz
- Degrid to get visibility values
- Visibility noise artificially added (to be ~2% of average signal)
- Angular resolutions of 0.18" and 0.429" respectively
- Pixel resolutions of 512x512
- Pseudo declinations also used to vary uv-coverages (-35 and -50 respectively)

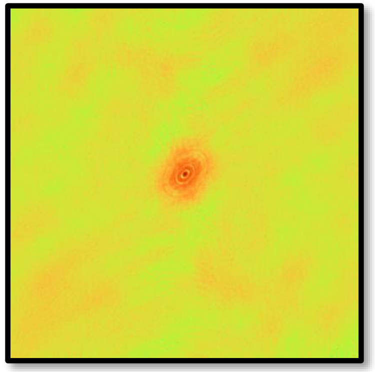
- Datasets taken from ALMA long baselines survey[2] and the VLA observation described in [3] for HL Tau and Cygnus A respectively
- ALMA Band 6 observation used for HL Tau (224.750GHz - 228.750GHz, 239.250 - 243.250 GHz, 4 spectral windows, 4 channels per spectral window, configuration 10)
- First spectral window (of 8) of VLA S-band used for Cygnus A (64 channels @ 1988.5 MHz – 2020.5 MHz, all 4 configurations)
- Angular resolutions of 0.005" and 0.125" respectively
- Pixel resolutions of 1500x1500 and 1728x1728 respectively

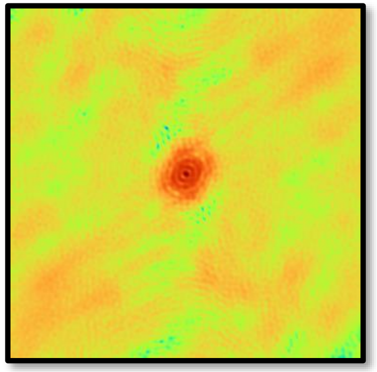
[1] Heywood, I., et al. "The 1.28 GHz MeerKAT galactic center mosaic." The Astrophysical Journal 925.2 (2022): 165.
[2] https://casaguides.nrao.edu/index.php/ALMA2014_LBC_SVDATA
[3] Sebokolodi, M. Lerato L., et al. "A Wideband Polarization Study of Cygnus A with the Jansky Very Large Array. I. The Observations and Data." The Astrophysical Journal 903.1 (2020): 36.

Dataset partitioning

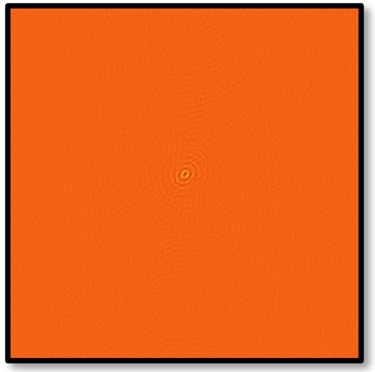
Example partition images are of the HL-Tau B6 dataset

$\tilde{\imath}$

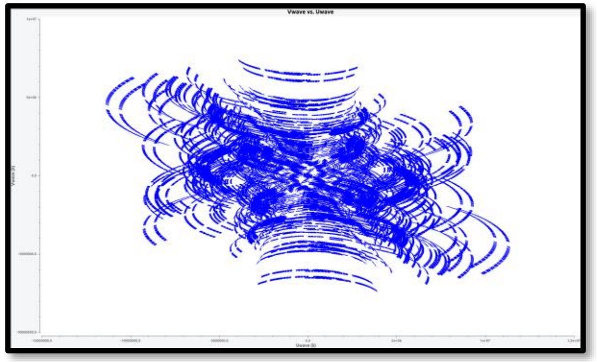




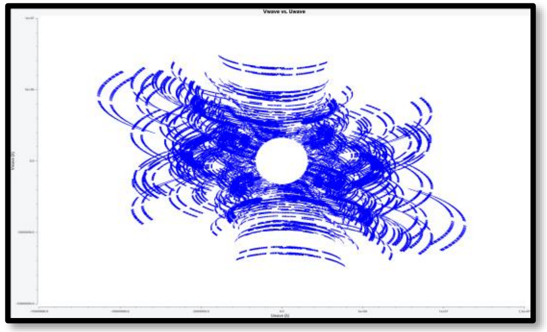
$\tilde{\imath}_{\mathcal{L}}$



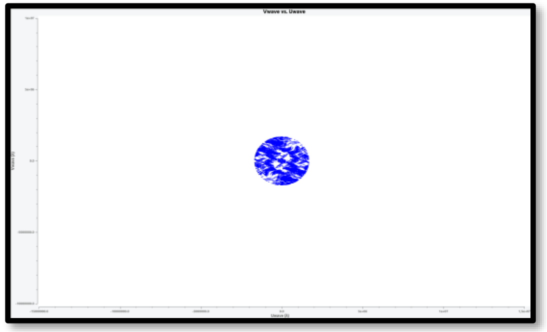
$\tilde{\imath}_{\mathcal{H}}$



V



$V_{\mathcal{H}}$



$V_{\mathcal{L}}$

Dataset	ℓ	$V_{\mathcal{L}}$	$V_{\mathcal{H}}$	$V_{\mathcal{L}\cap\mathcal{H}}$
Sgr B2	20	6.99M(.75)	2.53M(.27)	0.16M(.02)
Sgr C	25	9.05M(.57)	6.81M(.43)	0.11M(.01)
HL Tau	60	39.86M(.47)	46.24M(.54)	0.61M(.01)
Cygnus A	40	41.65M(.51)	41.83M(.51)	1.22M(.01)

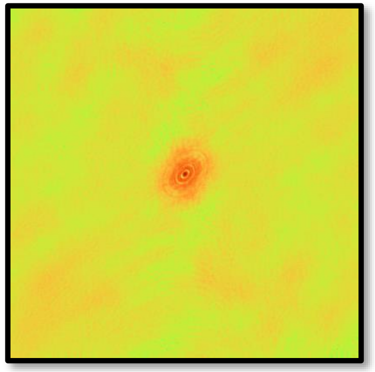
- Ideally want to create partitions with equal number visibilities and similar amounts of spatial freq info
- Difficult for Sgr B2 dataset due to SKA-Mid AA4 array (BDA[1] or something similar may be needed)

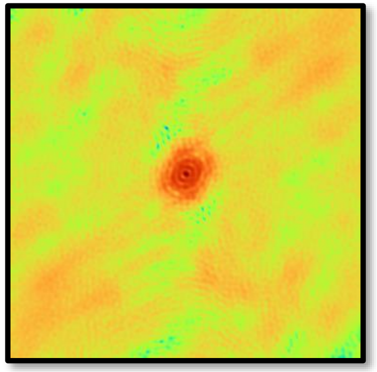
[1] Wijnholds, S. J., A. G. Willis, and S. Salvini. "Baseline-dependent averaging in radio interferometry." *Monthly Notices of the Royal Astronomical Society* 476.2 (2018): 2029-2039.

Dataset partitioning

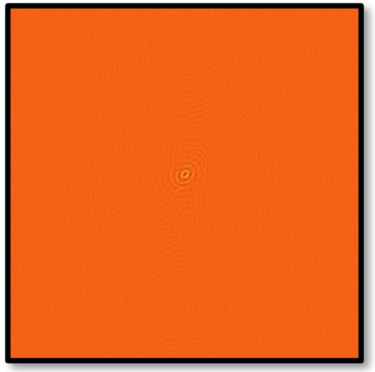
Example partition images are of the HL-Tau B6 dataset

$\tilde{\imath}$

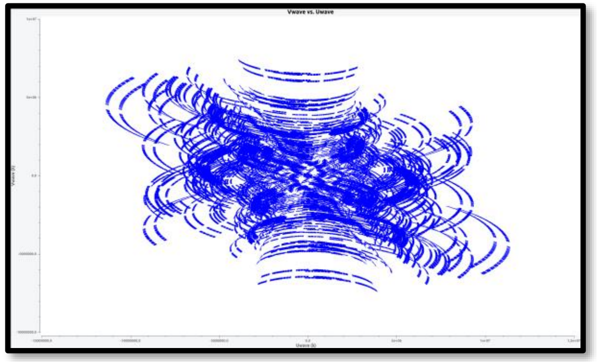




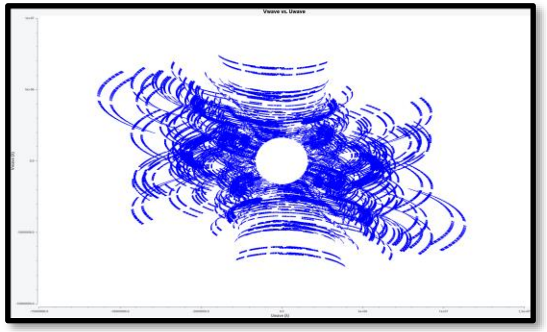
$\tilde{\imath}_{\mathcal{L}}$



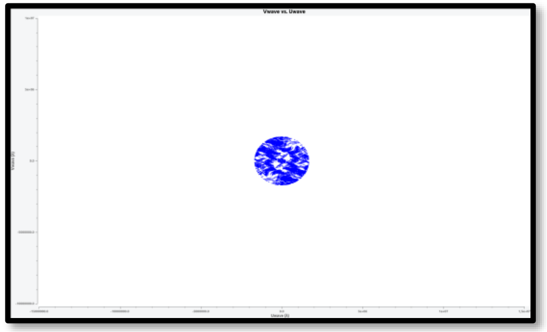
$\tilde{\imath}_{\mathcal{H}}$



V



$\mathrm{V}_{\mathcal{H}}$



$\mathrm{V}_{\mathcal{L}}$

Dataset	ℓ	$\mathrm{V}_{\mathcal{L}}$	$\mathrm{V}_{\mathcal{H}}$	$\mathrm{V}_{\mathcal{L}\cap\mathcal{H}}$
Sgr B2	20	6.99M(.75)	2.53M(.27)	0.16M(.02)
Sgr C	25	9.05M(.57)	6.81M(.43)	0.11M(.01)
HL Tau	60	39.86M(.47)	46.24M(.54)	0.61M(.01)
Cygnus A	40	41.65M(.51)	41.83M(.51)	1.22M(.01)

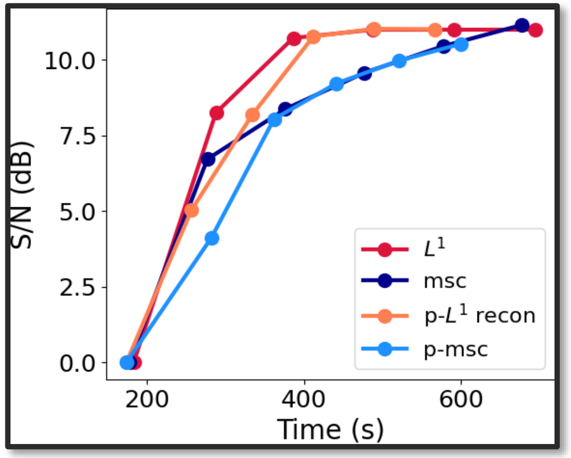
- Ideally want to create partitions with equal number visibilities and similar amounts of spatial freq info
- Difficult for Sgr B2 dataset due to SKA-Mid AA4 array (BDA[1] or something similar may be needed)

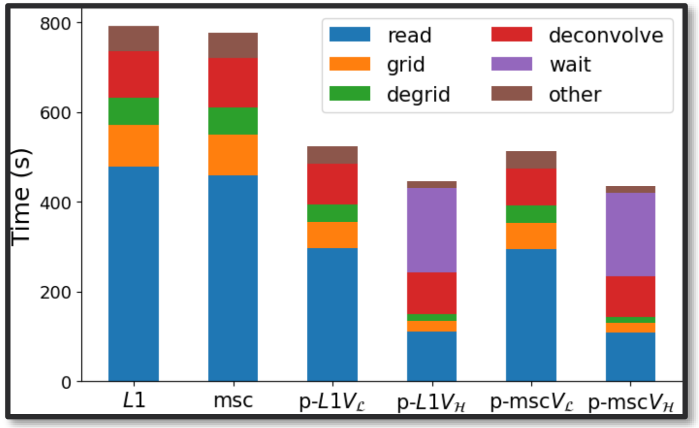
[1] Wijnholds, S. J., A. G. Willis, and S. Salvini. "Baseline-dependent averaging in radio interferometry." *Monthly Notices of the Royal Astronomical Society* 476.2 (2018): 2029-2039.

Results – Simulated

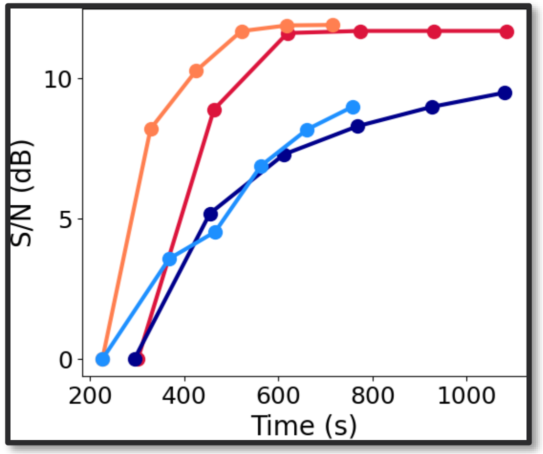
S/N obtained with comparison to ground truth

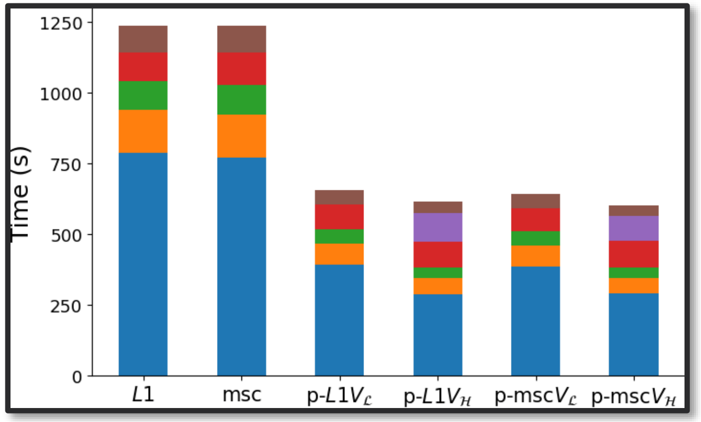
Sgr B2





Sgr C

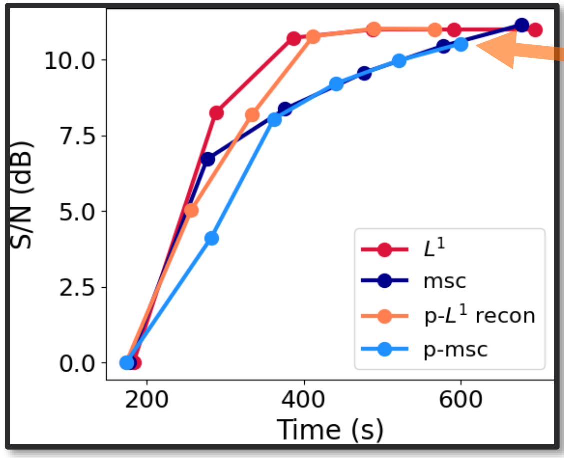


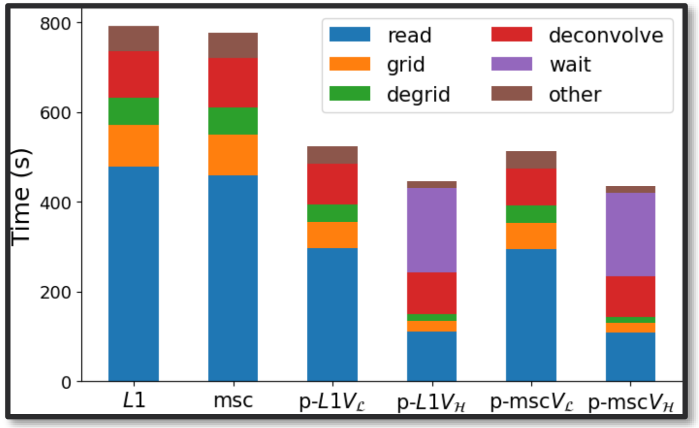


Results – Simulated

S/N obtained with comparison to ground truth

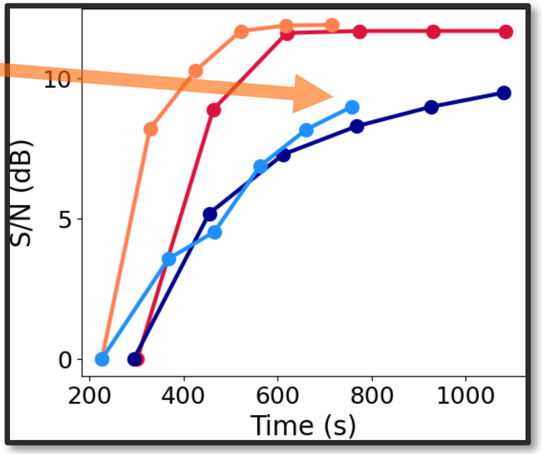
Sgr B2

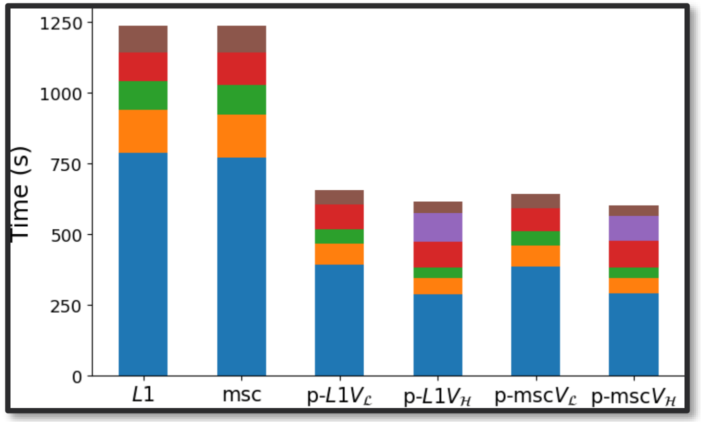




Poor parallelization for Sgr B2, much better for Sgr C. Mainly due to uneven partitioning.

Sgr C

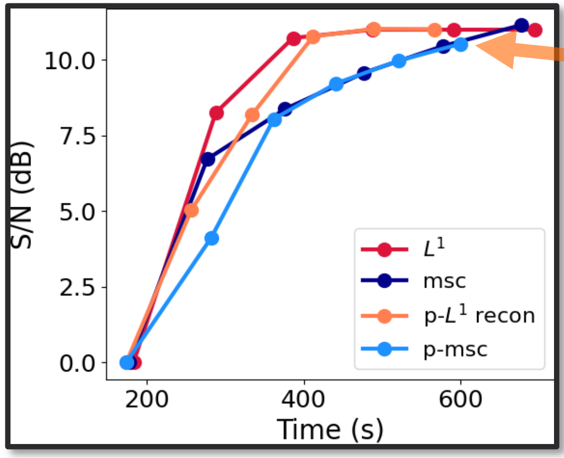




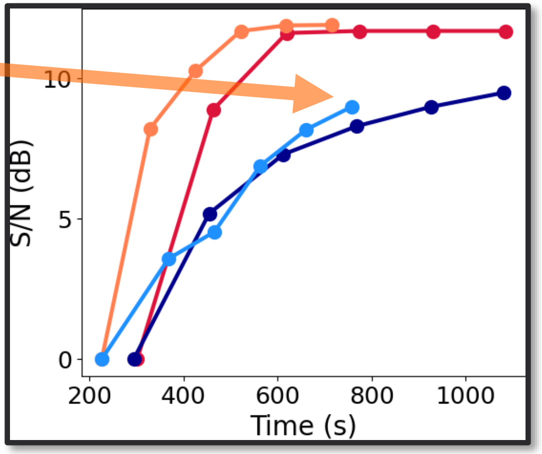
Results – Simulated

S/N obtained with comparison to ground truth

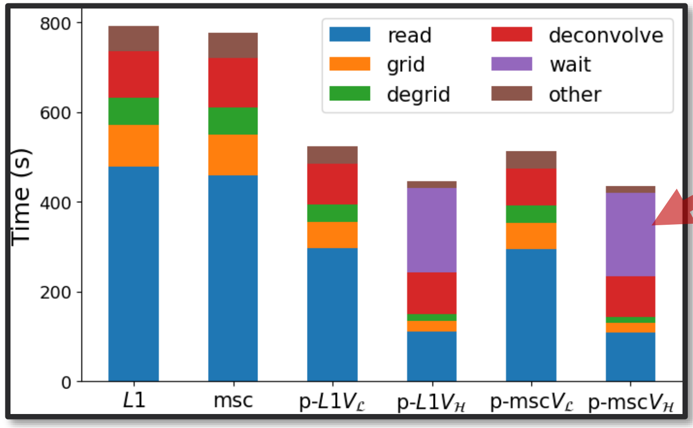
Sgr B2



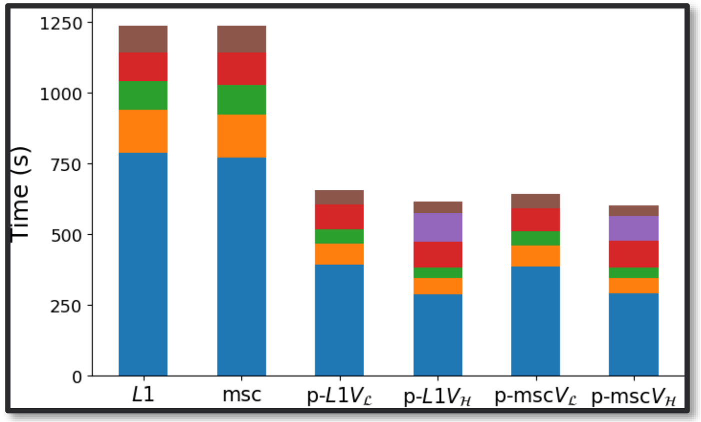
Sgr C



Poor parallelization for Sgr B2, much better for Sgr C. Mainly due to uneven partitioning.



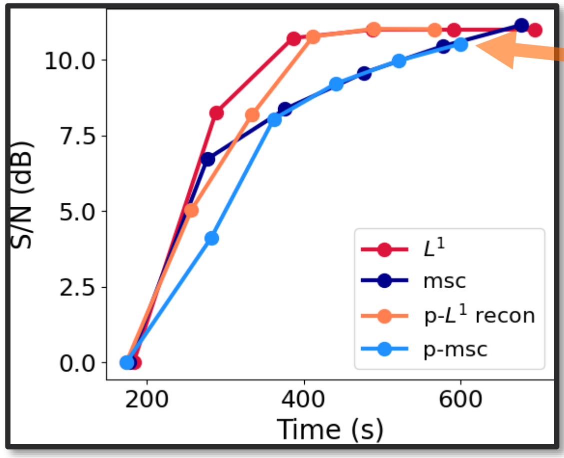
Poor partitioning causes a lot of idle time for node with less visibilities.



Results – Simulated

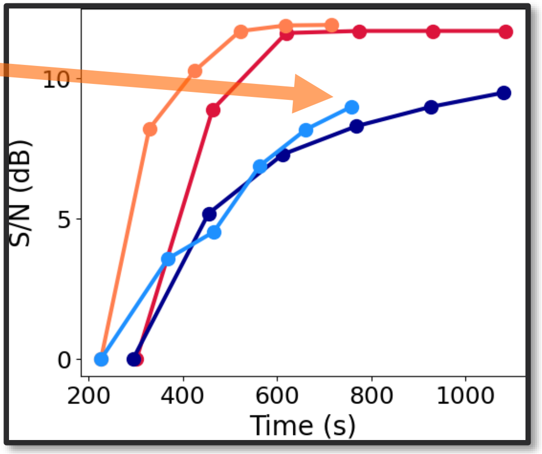
S/N obtained with comparison to ground truth

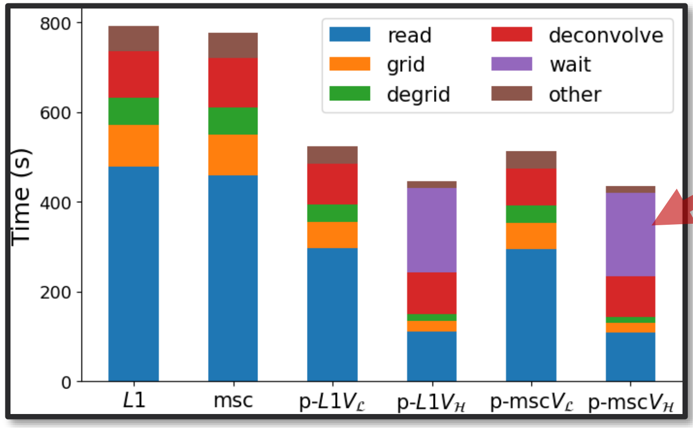
Sgr B2



Poor parallelization for Sgr B2, much better for Sgr C. Mainly due to uneven partitioning.

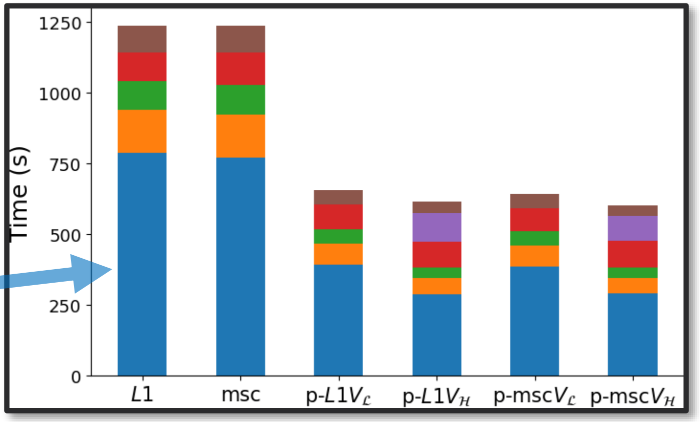
Sgr C





Poor partitioning causes a lot of idle time for node with less visibilities.

Read uses a lot of time due to implementation and RASCIL overhead

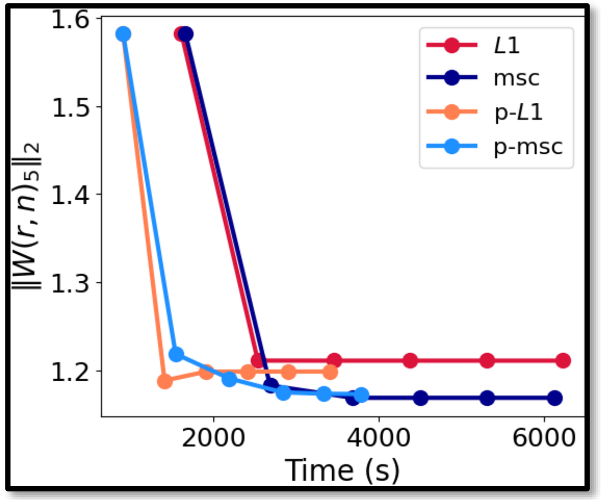


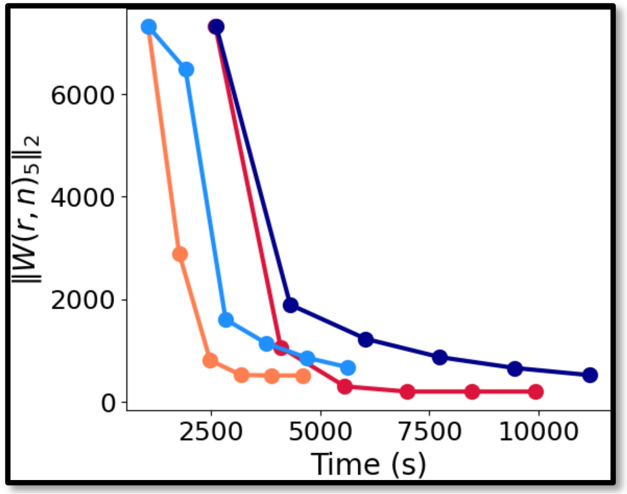
Results – Real

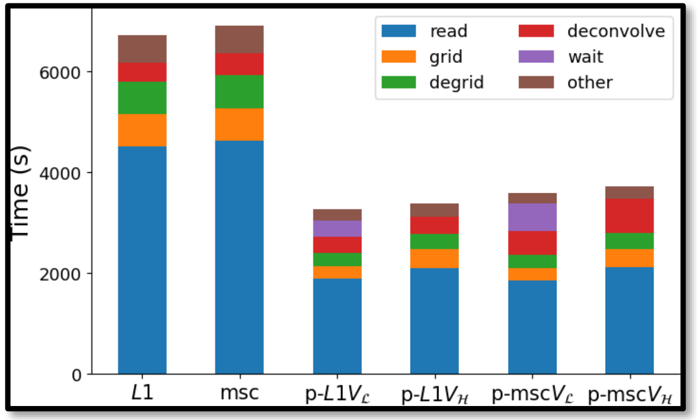
HL Tau

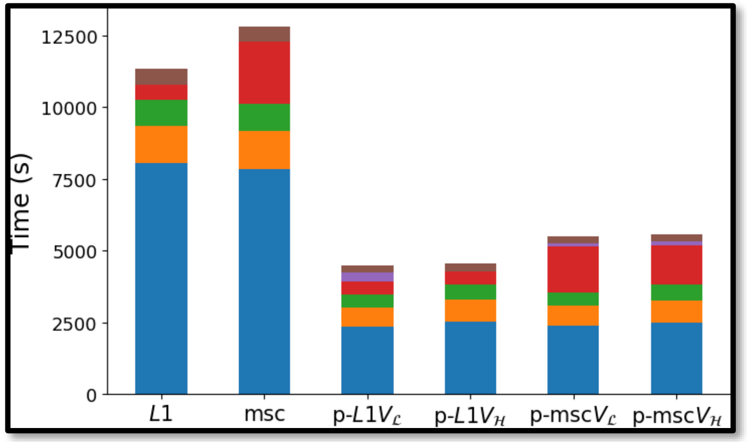
Wasserstein-1 distance between residual and ideal residual obtained via visibility negation. Computed per pixel (windowed), with L2 being plotted.

Cygnus A







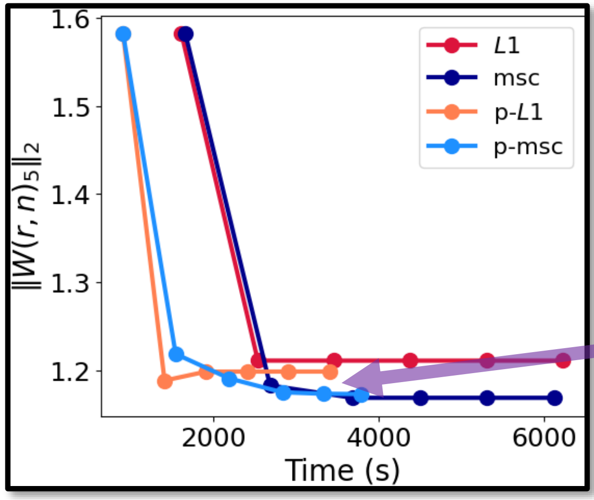


Results – Real

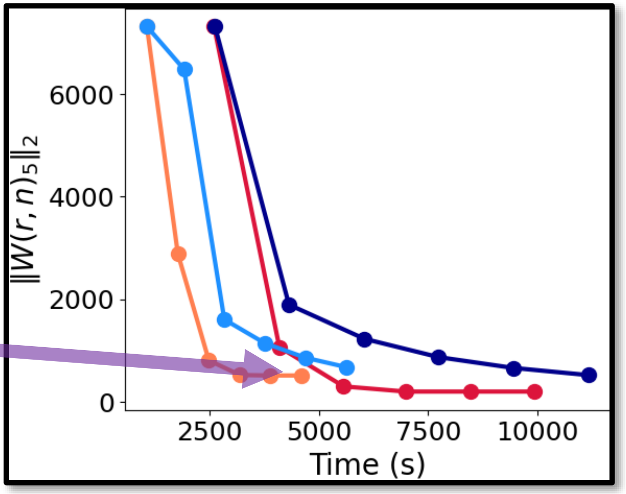
HL Tau

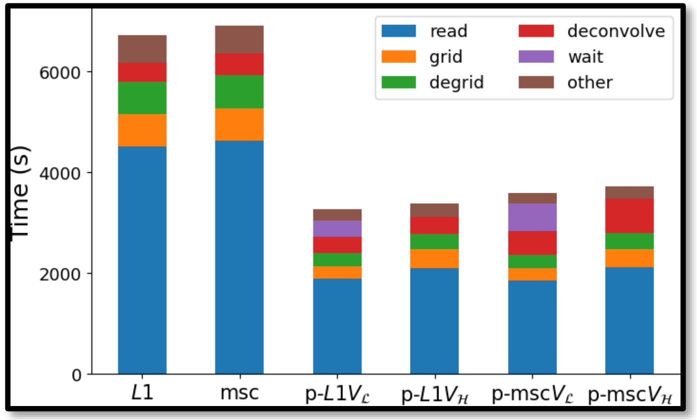
Wasserstein-1 distance between residual and ideal residual obtained via visibility negation. Computed per pixel (windowed), with L2 being plotted.

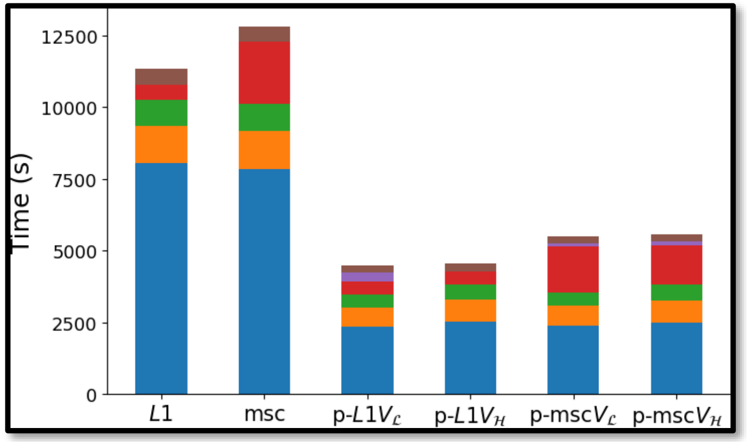
Cygnus A



Speedup much better for real datasets, close to optimal 2x

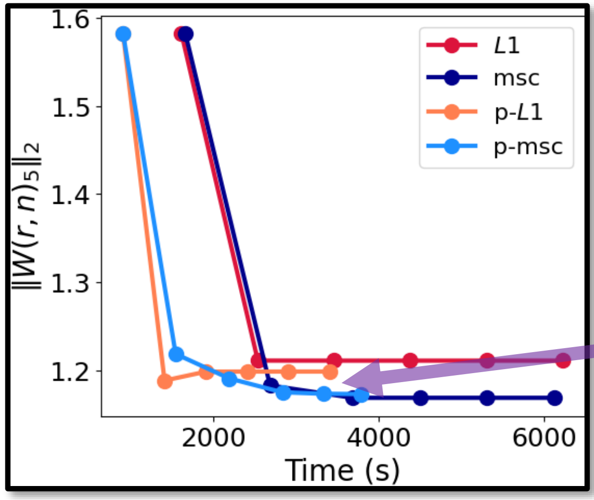






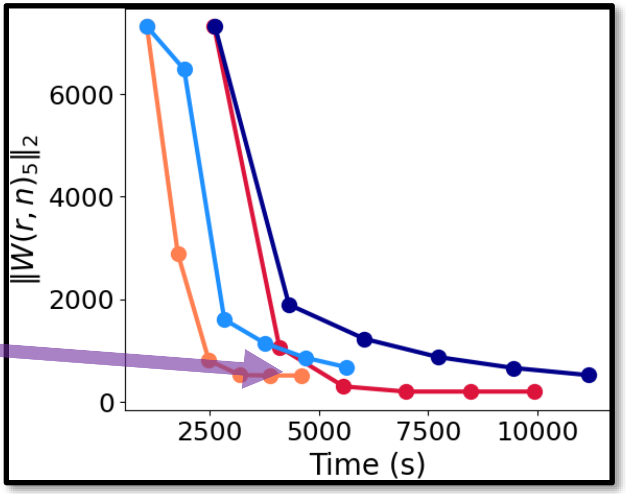
Results – Real

HL Tau

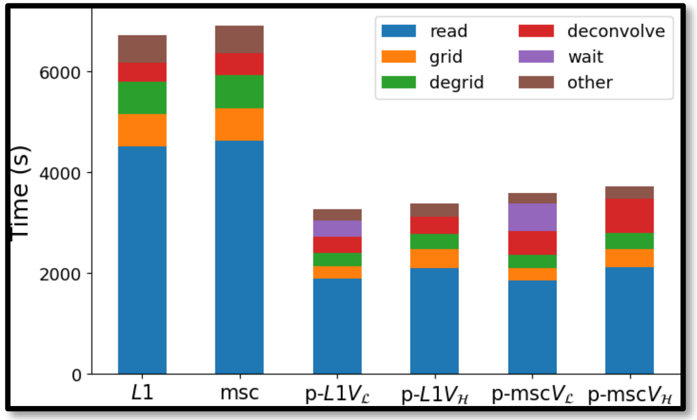


Wasserstein-1 distance between residual and ideal residual obtained via visibility negation. Computed per pixel (windowed), with L2 being plotted.

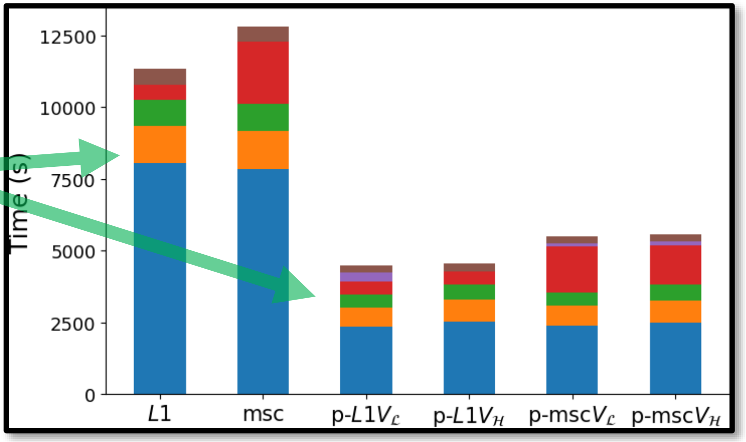
Cygnus A



Speedup much better for real datasets, close to optimal 2x



Over 2x speedup can be due to RASCIL overheads.



Results – Scaling to larger image sizes

Average processing times for Cygnus A dataset per major cycle

Alg.	Node	Pix. res.	Deconv.	Degrid	Grid	Disk I/O	Transf.	Other
p-msc	$V_{\mathcal{L}}$	1728 × 1728	354.70s	92.27s	117.21s	342.12s	0.01s	13.51s
	$V_{\mathcal{H}}$	1728 × 1728	288.83s	109.84s	129.62s	353.67s	0.01s	11.88s
	$V_{\mathcal{L}}$	10k × 10k	17778.44s (× 50)	1450.58s (× 16)	2519.88s (× 21)	358.31s (× 1)	0.45s (× 52)	21.35s (× 2)
	$V_{\mathcal{H}}$	10k × 10k	18014.80s (× 62)	2159.66s (× 20)	2533.84s (× 20)	362.50s (× 1)	0.67s (× 51)	21.11s (× 2)
p-L1	$V_{\mathcal{L}}$	1728 × 1728	91.66s	92.44s	111.52s	332.31s	0.02s	11.52s
	$V_{\mathcal{H}}$	1728 × 1728	89.33s	108.17s	126.85s	359.43s	0.02s	13.49s
	$V_{\mathcal{L}}$	10k × 10k	3595.75s (× 39)	1449.08s (× 16)	2457.80s (× 22)	365.53s (× 1)	0.59s (× 30)	20.16s (× 2)
	$V_{\mathcal{H}}$	10k × 10k	3573.73s (× 40)	2173.30s (× 20)	2555.44s (× 20)	363.62s (× 1)	0.60s (× 25)	20.22s (× 1)

Results – Scaling to larger image sizes

Average processing times for Cygnus A dataset per major cycle

Alg.	Node	Pix. res.	Deconv.	Degrid	Grid	Disk I/O	Transf.	Other
p-msc	$V_{\mathcal{L}}$	1728 × 1728	354.70s	92.27s	117.21s	342.12s	0.01s	13.51s
	$V_{\mathcal{H}}$	1728 × 1728	288.83s	109.84s	129.62s	353.67s	0.01s	11.88s
	$V_{\mathcal{L}}$	10k × 10k	17778.44s (× 50)	1450.58s (× 16)	2519.88s (× 21)	358.31s (× 1)	0.45s (× 52)	21.35s (× 2)
	$V_{\mathcal{H}}$	10k × 10k	18014.80s (× 62)	2159.66s (× 20)	2533.84s (× 20)	362.50s (× 1)	0.67s (× 51)	21.11s (× 2)
p-L1	$V_{\mathcal{L}}$	1728 × 1728	91.66s	92.44s	111.52s	332.31s	0.02s	11.52s
	$V_{\mathcal{H}}$	1728 × 1728	89.33s	108.17s	126.85s	359.43s	0.02s	13.49s
	$V_{\mathcal{L}}$	10k × 10k	3595.75s (× 39)	1449.08s (× 16)	2457.80s (× 22)	365.53s (× 1)	0.59s (× 30)	20.16s (× 2)
	$V_{\mathcal{H}}$	10k × 10k	3573.73s (× 40)	2173.30s (× 20)	2555.44s (× 20)	363.62s (× 1)	0.60s (× 25)	20.22s (× 1)

Primary bottlenecks seem to be deconvolution and de/gridding (to a lesser extent).

Results – Scaling to larger image sizes

Average processing times for Cygnus A dataset per major cycle

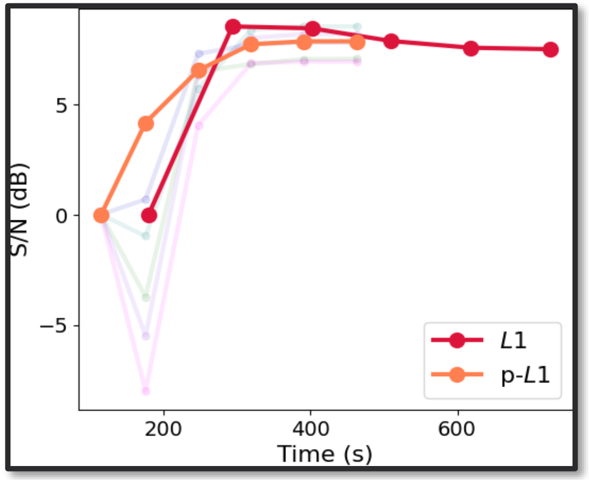
Alg.	Node	Pix. res.	Deconv.	Degrid	Grid	Disk I/O	Transf.	Other
p-msc	$V_{\mathcal{L}}$	1728 × 1728	354.70s	92.27s	117.21s	342.12s	0.01s	13.51s
	$V_{\mathcal{H}}$	1728 × 1728	288.83s	109.84s	129.62s	353.67s	0.01s	11.88s
	$V_{\mathcal{L}}$	10k × 10k	17778.44s (× 50)	1450.58s (× 16)	2519.88s (× 21)	358.31s (× 1)	0.45s (× 52)	21.35s (× 2)
	$V_{\mathcal{H}}$	10k × 10k	18014.80s (× 62)	2159.66s (× 20)	2533.84s (× 20)	362.50s (× 1)	0.67s (× 51)	21.11s (× 2)
p-L1	$V_{\mathcal{L}}$	1728 × 1728	91.66s	92.44s	111.52s	332.31s	0.02s	11.52s
	$V_{\mathcal{H}}$	1728 × 1728	89.33s	108.17s	126.85s	359.43s	0.02s	13.49s
	$V_{\mathcal{L}}$	10k × 10k	3595.75s (× 39)	1449.08s (× 16)	2457.80s (× 22)	365.53s (× 1)	0.59s (× 30)	20.16s (× 2)
	$V_{\mathcal{H}}$	10k × 10k	3573.73s (× 40)	2173.30s (× 20)	2555.44s (× 20)	363.62s (× 1)	0.60s (× 25)	20.22s (× 1)

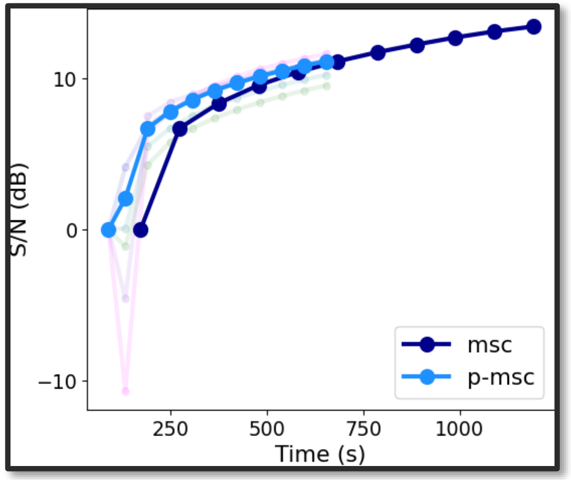
Primary bottlenecks seem to be deconvolution and de/gridding (to a lesser extent).

Transfer time also increases similarly to deconvolution, but cost negligible. Even for 100kx100k images, with the current cost increases, a transfer only takes ~72s which is substantially less than even the 10kx10k deconvolution.

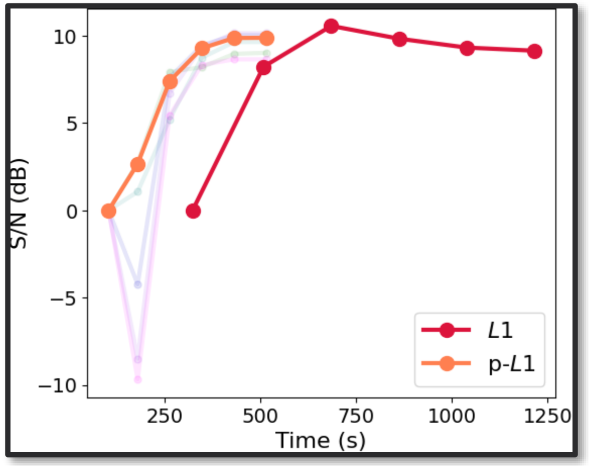
Results – 5 partitions on simulated

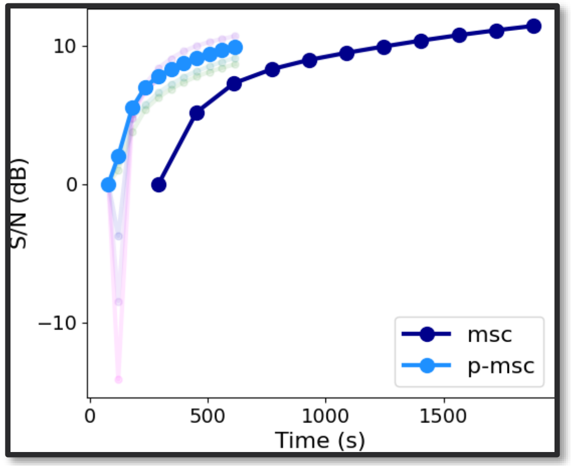
Sgr B2





Sgr C

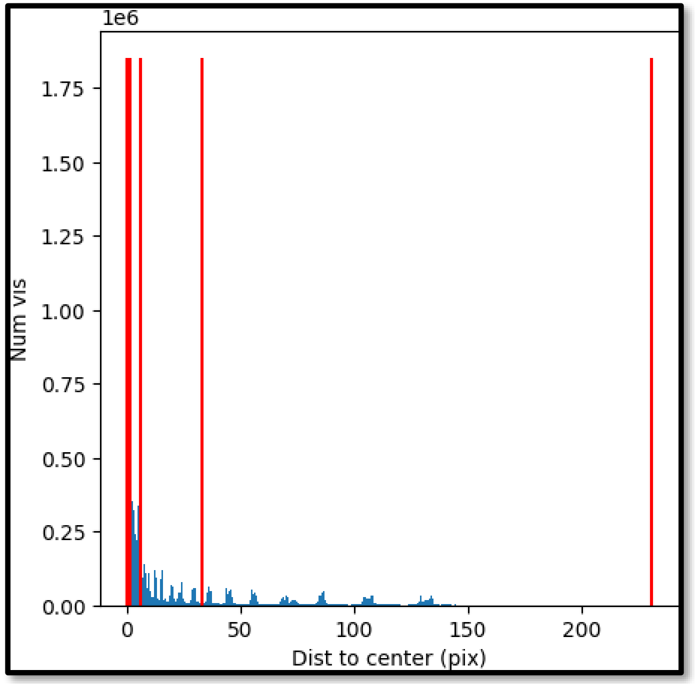




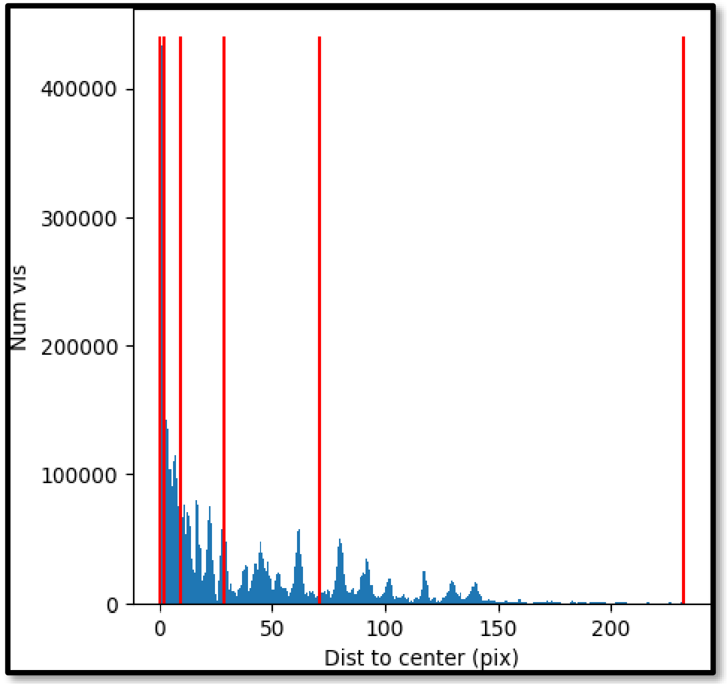
- Partition configurations
  - Cuts made at <4, 2-11, 9-36, 34-71, >69
  - Visibilities per partition for Sgr B2: 4710396,2569876,1565927,922346,942517
  - Visibilities per partition for Sgr C: 4196451,3975415,3910825,2742692,3152216
- Total processing times:
  - Sgr B2:
    - 726.66s (serial) vs 462.46s (parallel) for l1 after 5 major cycles
    - 1192.22s (serial) vs 654.69s (parallel) for msc after 10 major cycles
  - Sgr C:
    - 1215.17s (serial) vs 514.02s (parallel) for l1 after 5 major cycles
    - 1880.88s (serial) vs 615.99s (parallel) for msc after 10 major cycles
- 1.57x (l1) and 1.82x (msc) speedup for Sgr B2
- 2.36x (l1) and 3.05x (msc) speedup for Sgr C

Short-baselines density problem

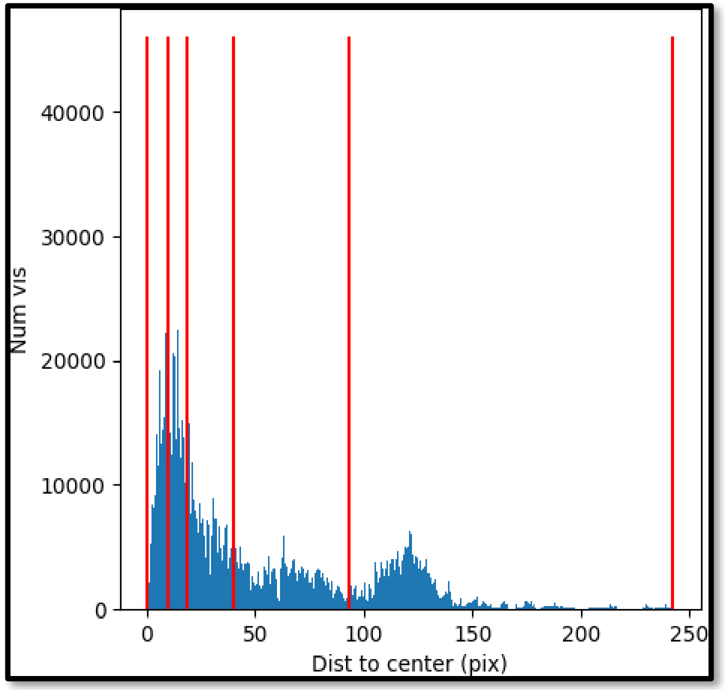
SKA AA4 Mid



SKA AA4 Low



Meerkat



Baseline Dependent Averaging

Baseline dependent averaging (BDA)[1] averages visibilities based on decorrelation.

- Shorter baselines -> less decorrelation
- Evens out visibility distributions

Time and frequency decorrelation

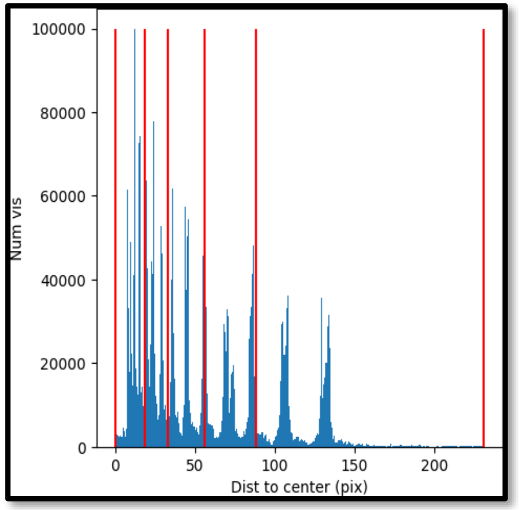
$$\rho_{total} = \rho_t \times \rho_f$$

$$\rho_t = 1 - \frac{\pi^2 T^2}{6}\left(\frac{\mathrm{d}u}{\mathrm{d}t} l + \frac{\mathrm{d}v}{\mathrm{d}t} m\right)^2$$

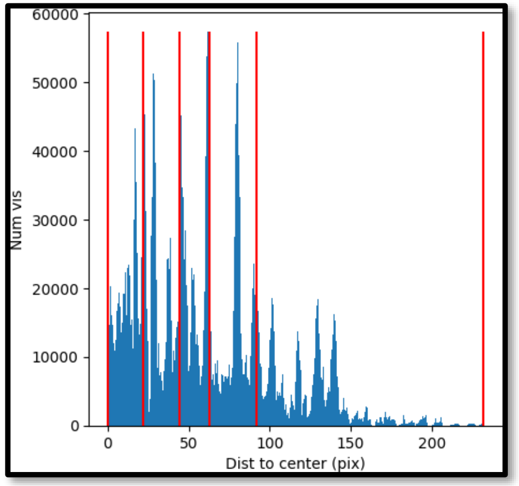
$$\rho_f = \| sinc(\frac{\pi \Delta \nu \tau_g}{2}) \|$$

Visibilities averaged with $\max(\rho_{total}) = 0.99$

SKA AA4 Mid

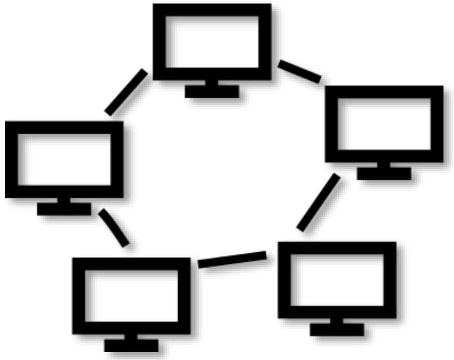


SKA AA4 Low

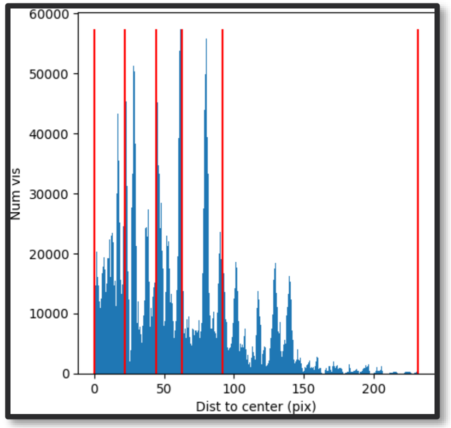


[1] Wijnholds, S. J., A. G. Willis, and S. Salvini. "Baseline-dependent averaging in radio interferometry." *Monthly Notices of the Royal Astronomical Society* 476.2 (2018): 2029-2039.

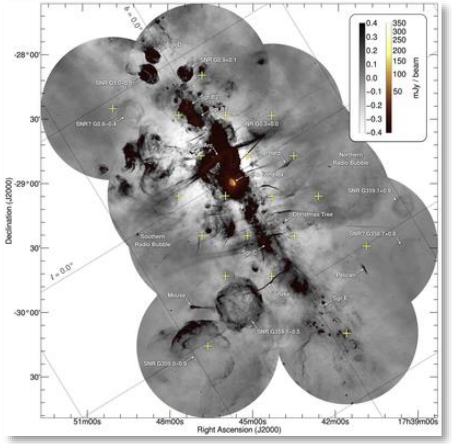
Future Work



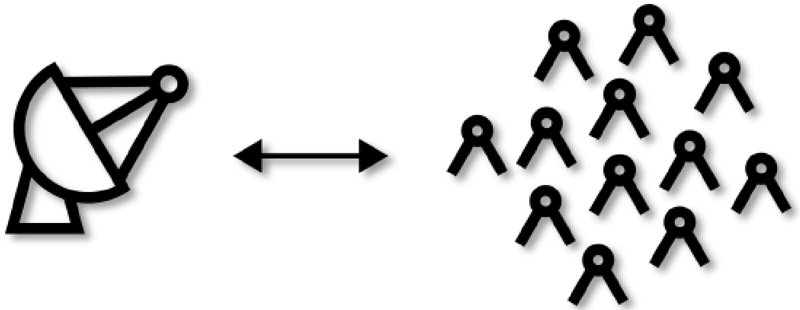
More complex network topologies



Incorporate BDA



Reconstruct very large images (32k x 32k – 100k x 100k), image source[1]

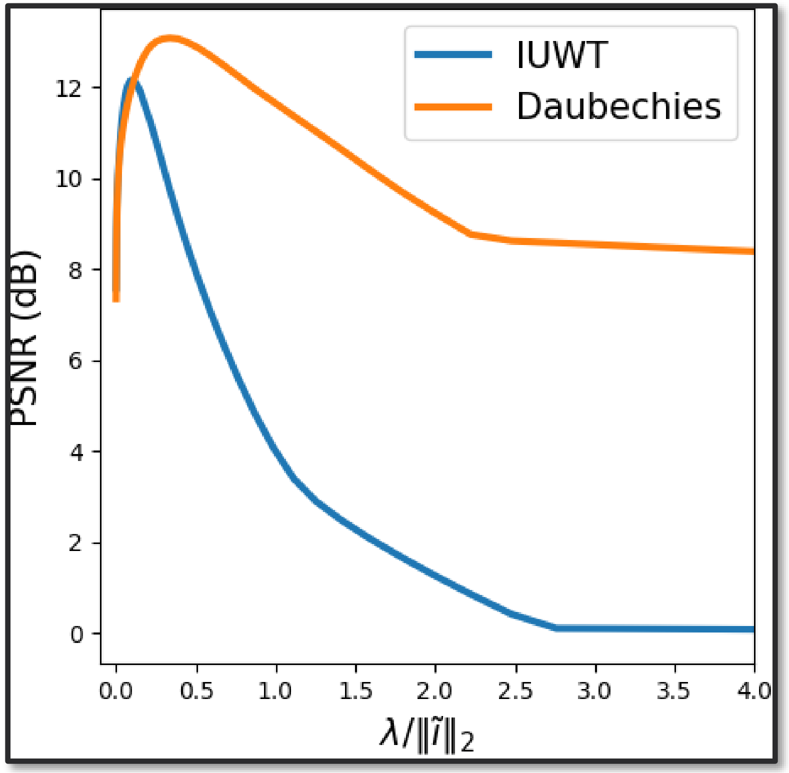


Reconstruct single-dish + interferometric

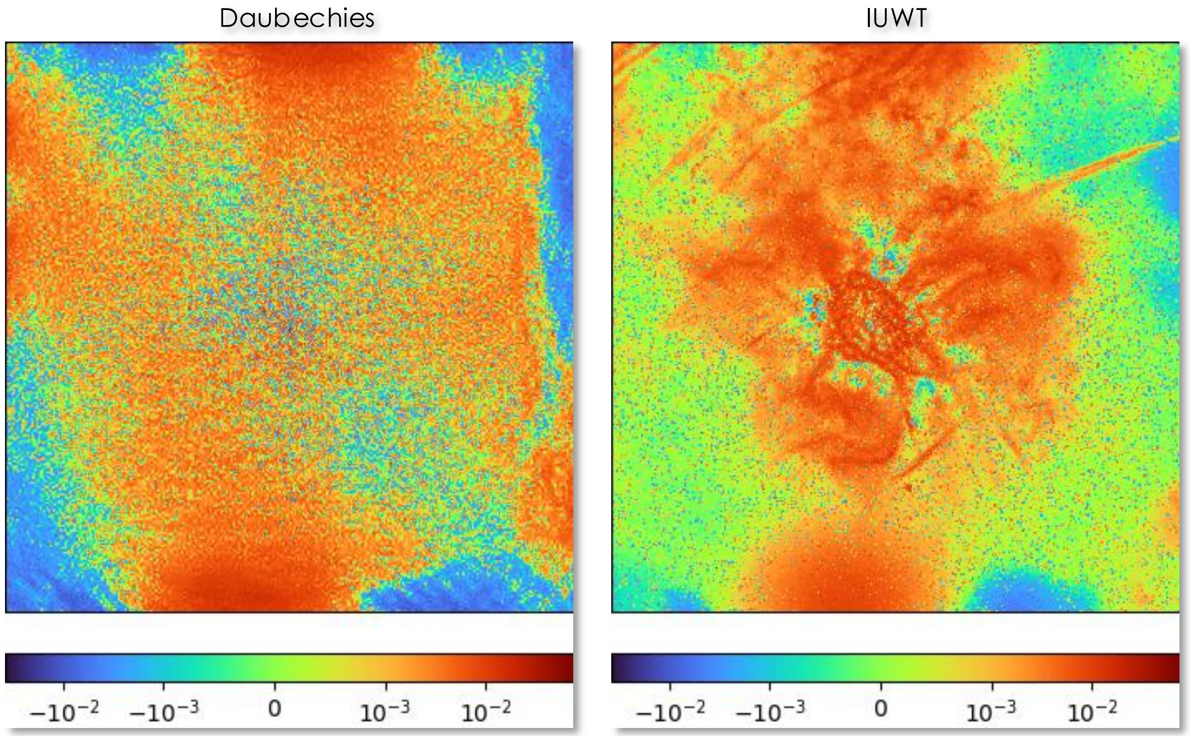
[1] Heywood, I., et al. "The 1.28 GHz MeerKAT galactic center mosaic." The Astrophysical Journal 925.2 (2022): 165.

Appendices

IUWT vs Daubechies

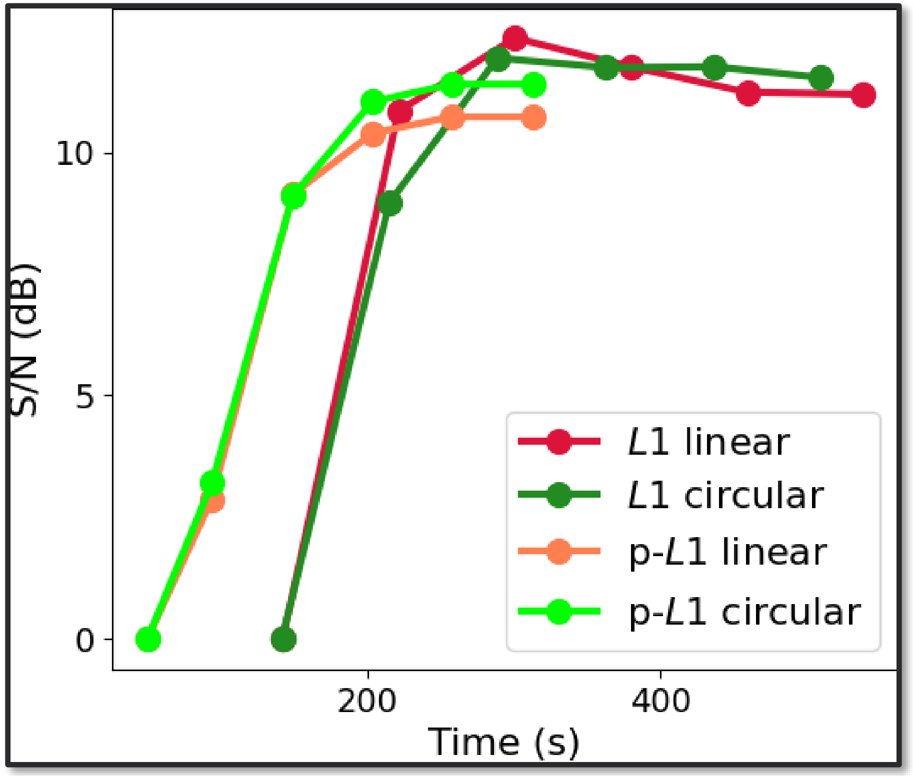


First major-cycle residuals for Sgr A



IUWT seems worse at reconstructing large-scale anisotropic extended emissions.

Linear vs circular convolution



- Using linear convolution instead of circular can be desired so that bright extended sources don't wrap around.
- More complicated to find the step size as operator does not diagonalize with Fourier transform, have to rely on something like power iteration.
- Results don't necessarily seem better as shown on the left.
- More expensive to compute

Filters

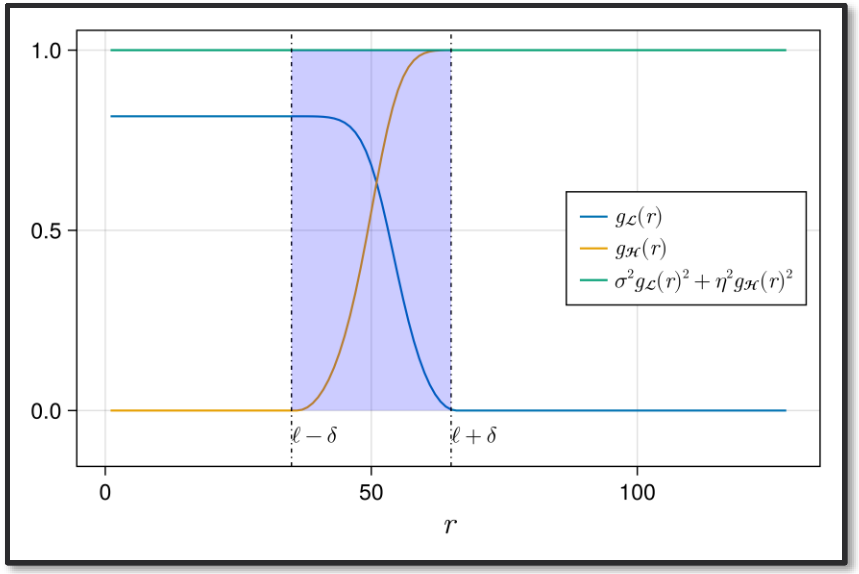
$$r > \ell + \delta : \quad |g_{\mathcal{H}}(r)|^2 = 1/\sigma^2, \; g_{\mathcal{L}}(u) = 0$$

$$r < \ell - \delta : \quad g_{\mathcal{H}}(r) = 0, \; |g_{\mathcal{L}}(r)|^2 = 1/\eta^2$$

$$\ell - \delta < r < \ell + \delta : \quad \sigma^2 |g_{\mathcal{H}}(r)|^2 + \eta^2 |g_{\mathcal{L}}(r)|^2 = 1$$

$$g_{\mathcal{L}}(r) = \alpha(r) \left(1 - \sin\left(\frac{\pi}{2\delta}(r - \ell)\right)\right)$$

$$g_{\mathcal{H}}(r) = \alpha(r) \left(1 + \sin\left(\frac{\pi}{2\delta}(r - \ell)\right)\right)$$



- 1-D filters used as distance in 2-D, resulting in an annulus
- Filters have hard cut-off in Fourier domain, may result in either oscillations or infinite support in spatial domain
- Testing with windowed-sinc and Parks-McClellan, no conclusive results yet, but preliminary experiments suggest no large difference.

Selection of λ

Preliminary results suggested that lambda should be normalized by the norm of the image, and be increased as the major-cycles progress to maximize S/N. We use:

$$\lambda_{\mathcal{L}_n} = 0.05\|\tilde{\imath}_{\mathcal{L}_n}\|_2 \times 2^n$$
$$\lambda_{\mathcal{H}_n} = 0.05\|\tilde{\imath}_{\mathcal{H}_n}\|_2 \times 2^n$$
$$\lambda_n = 0.01\|\tilde{\imath}_n\|_2 \times 2^n$$

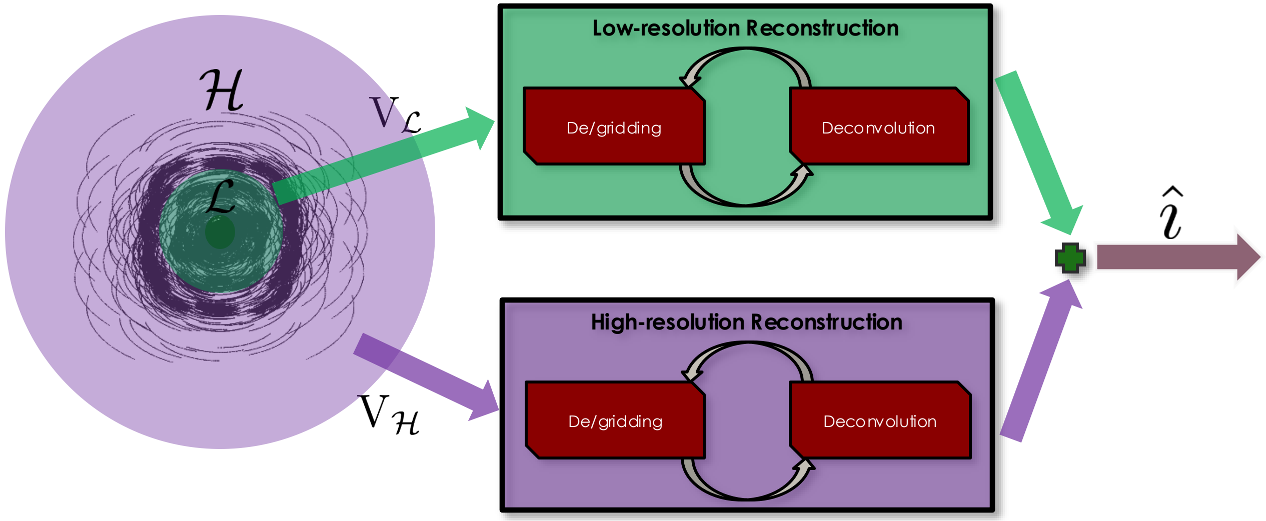
The above strategy no longer works well for larger numbers of partitions. For this case, and onwards, For the multipartition problem:

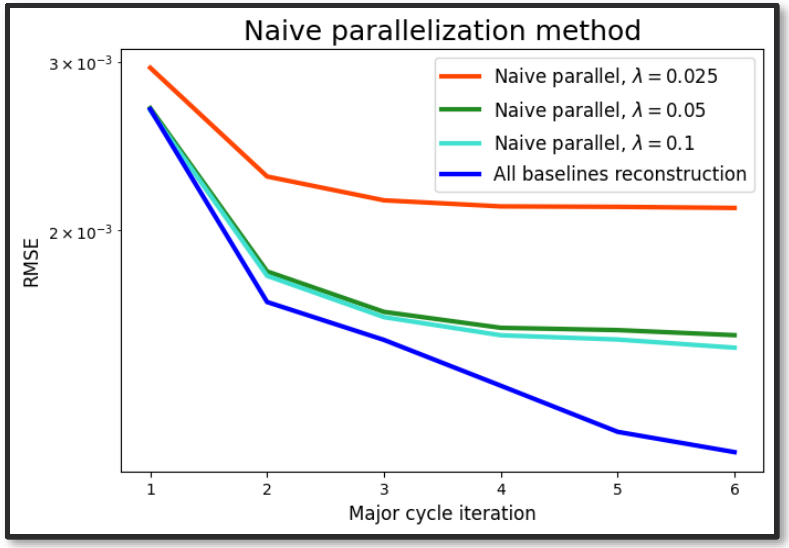
$$\|G_j(\tilde{\imath} - HW\alpha)\|^2 + \sum_{\substack{i=1 \\ i\neq j}} \|G_i C_i - W\alpha)\|^2 + \lambda\|W\alpha\|_1$$

We use:

$$\lambda_n = \eta_n \lambda_{max_n}$$
$$\eta_n = \alpha + (1-\alpha)\frac{e^{\beta t_n} - 1}{e^{\beta} - 1}$$
$$t_n = \frac{n}{N-1}$$
$$\lambda_{max_n} = 2\|W^\dagger(H_j^\dagger G_j^\dagger G_j \tilde{\imath}_{n_j} + \sum_{i=1, i\neq j} G_i^\dagger G_i C_{n_i})\|_\infty$$

Naively adding separately deconvolved images



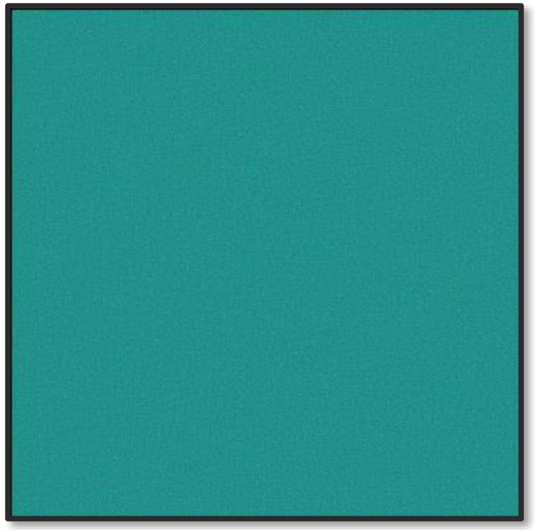


- Naïve parallel reconstructions seem always worse.
- Possibly due to terms not regularized together, which introduces some assumptions on sparsity.
- Can probably tune lambda so that the same result is obtained, but unclear how.

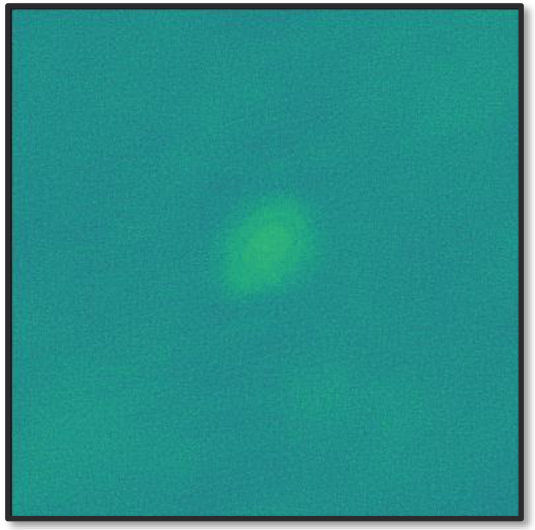
Evaluating on real datasets

- Reference residual obtained by randomly negating half of the visibilities
- Residual after imaging compared to reference statistically in a per-pixel manner
- 5x5 sliding window used as only 1 realization of imaging residual
- Wasserstein-1 distance used for statistical test
- L2 norm of Wasserstein distances used as final metric

Reference residual



Imaging residual



Wasserstein-1 distances using 5x5 sliding window