Parallelizing radio-interferometric image reconstruction by baselines

Sunrise Wang, Simon Prunet, Shan Mignot, André Ferrari

Laboratoire ECLAT
Atelier technique Nov 2024

Outline

Introduction

Parallelization Framework

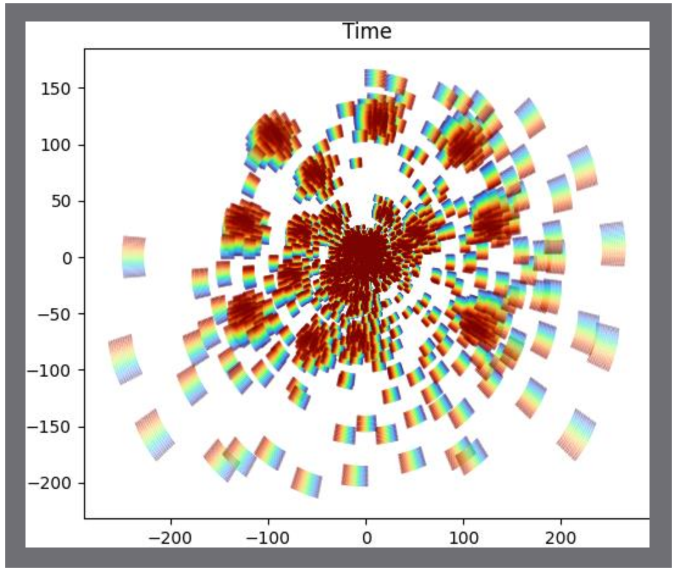
Applying our Framework to existing reconstruction methods

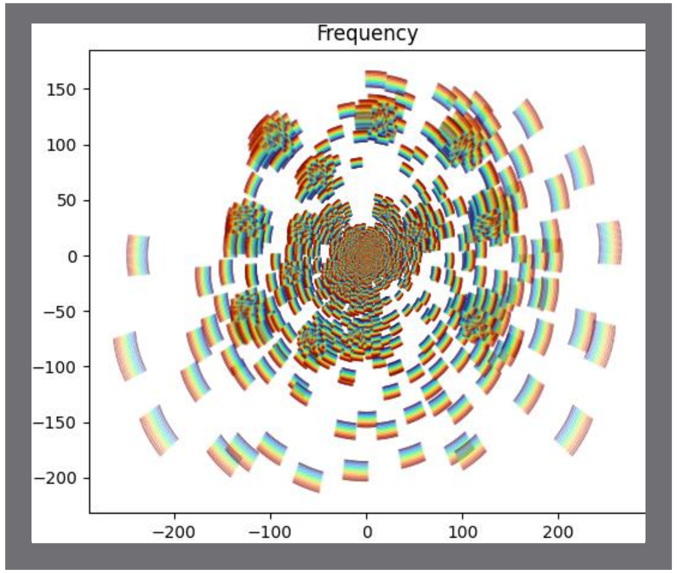
Results

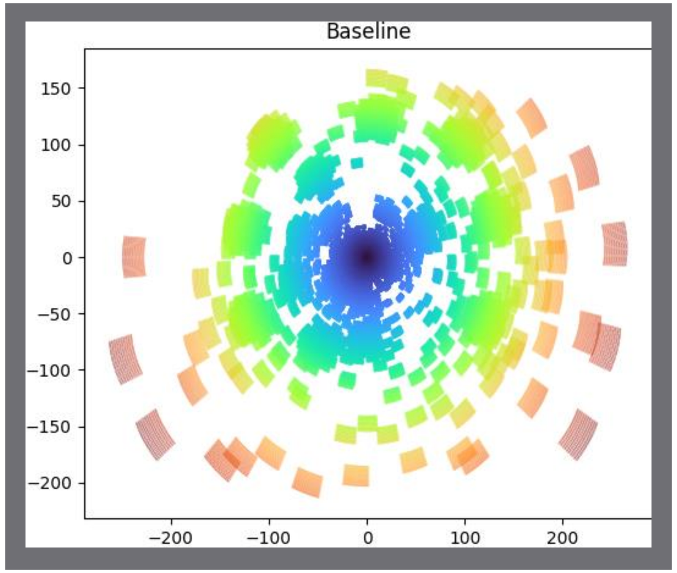
Introduction

image sources: [3, 4]
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/0
[4] http://www.idris.fr/







Scaling the computation

Need to scale the computation

Partition visibilities, process separately

Time and Frequency partitioning relatively trivial, baseline-partitioning more complicated as not all frequencies are available

Distributed and parallel sparse convex optimization for radio interferometry with PURIFY

Luke Pratley[a], Jason D. McEwen[a], Mayeul d'Avezac[b], Xiaohao Cai[a], David Perez-Suarez[b], Ilektra Christidi[b], Roland Guichard[b]

[a] *Mullard Space Science Laboratory (MSSL), University College London (UCL), Holmbury St Mary, Surrey RH5 6NT, UK*
[b] *Research Software Development Group, Research IT Services, University College London (UCL), London WC1E 6BT, UK*

Scalable splitting algorithms for big-data interferometric imaging in the SKA era

Alexandru Onose,[1★†] Rafael E. Carrillo,[2†] Audrey Repetti,[1] Jason D. McEwen,[3] Jean-Philippe Thiran,[2] Jean-Christophe Pesquet[4] and Yves Wiaux[1]

Previous related work

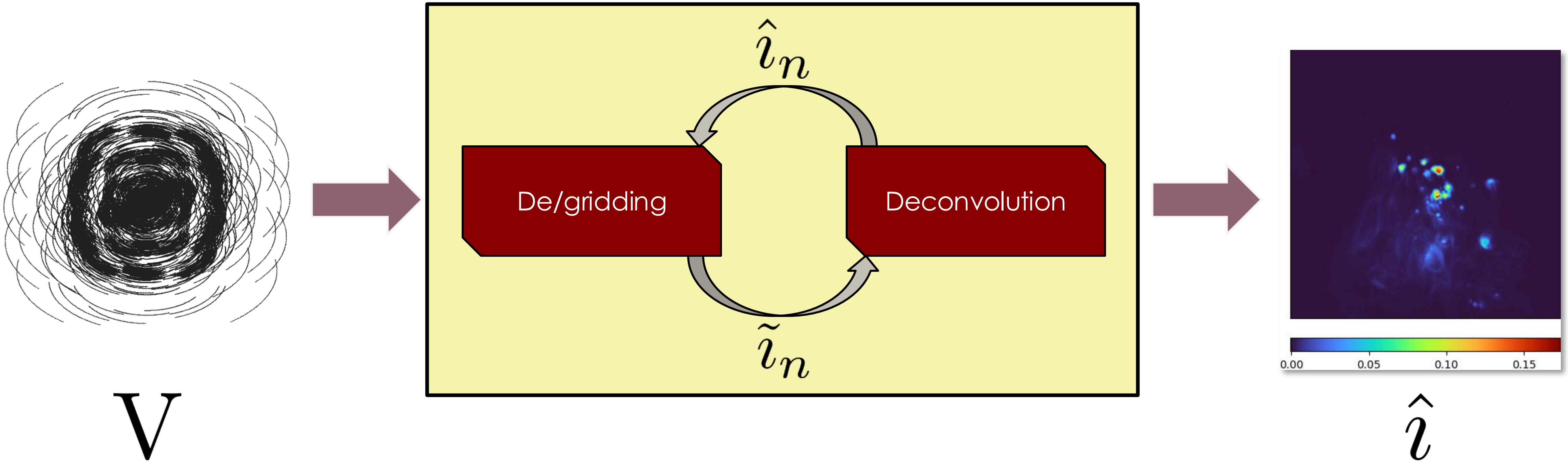
Both works are primal-dual methods that look to solve the measurement operator directly

Implemented in PURIFY[1] framework

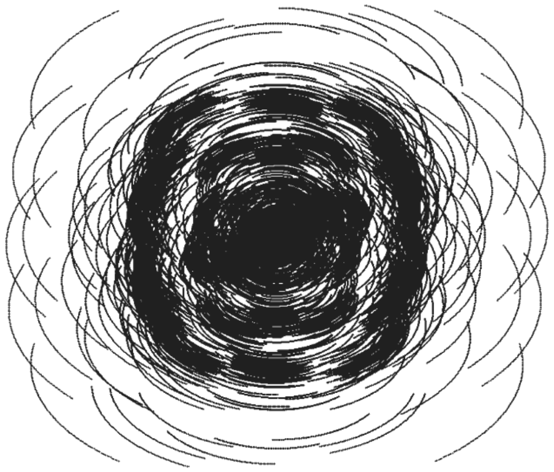
[1] Carrillo, Rafael E., Jason D. McEwen, and Yves Wiaux. "PURIFY: a new approach to radio-interferometric imaging." *Monthly Notices of the Royal Astronomical Society* 439.4 (2014): 3591-3604.

Parallelizing image reconstruction by baseline length

Major-Minor Loop Reconstruction



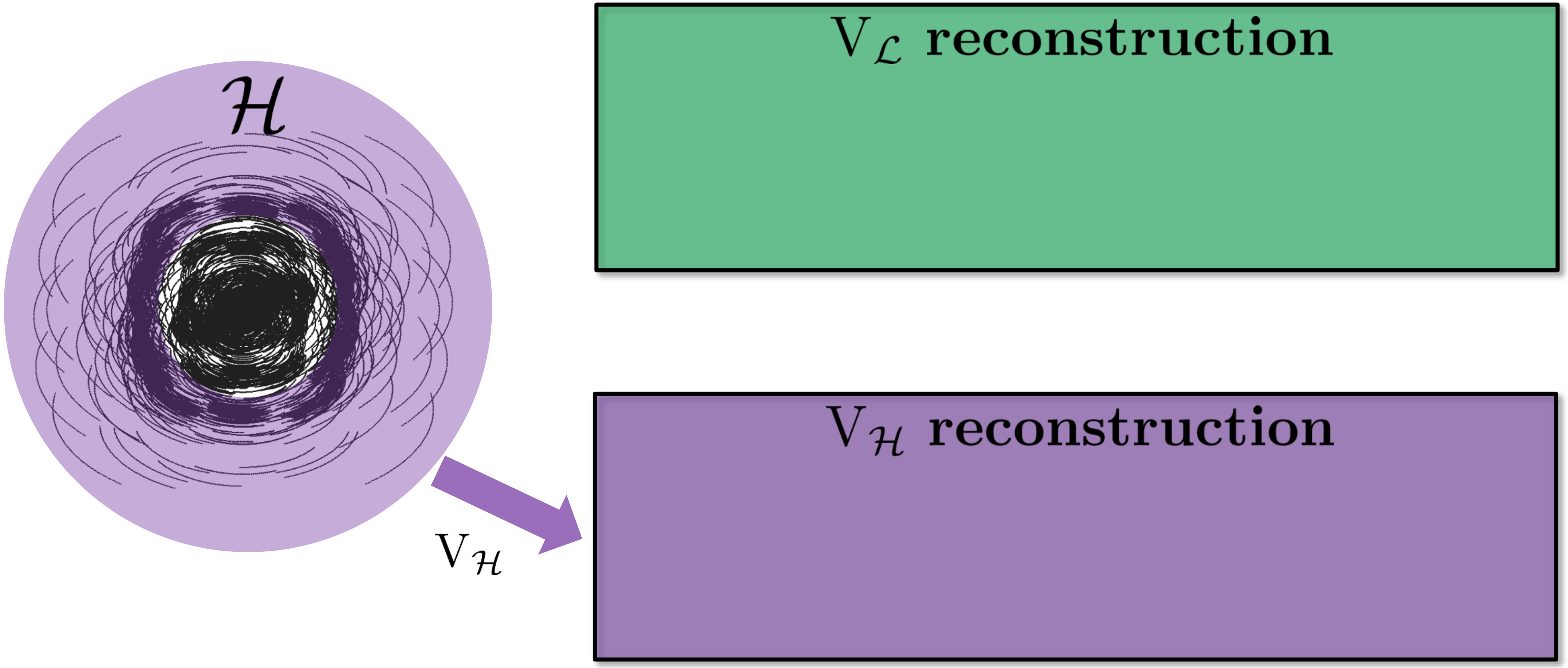
Parallelization Framework



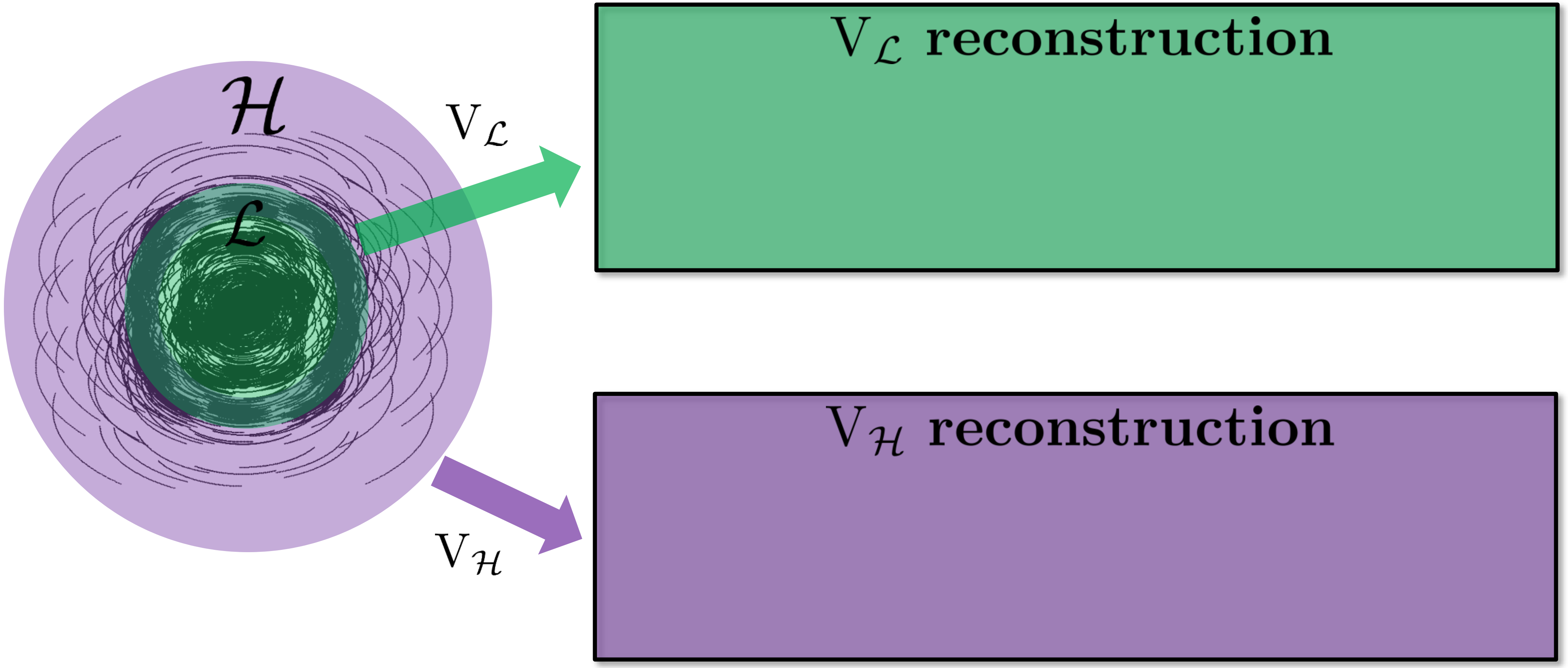
$V_{\mathcal{L}}$ **reconstruction**

$V_{\mathcal{H}}$ **reconstruction**

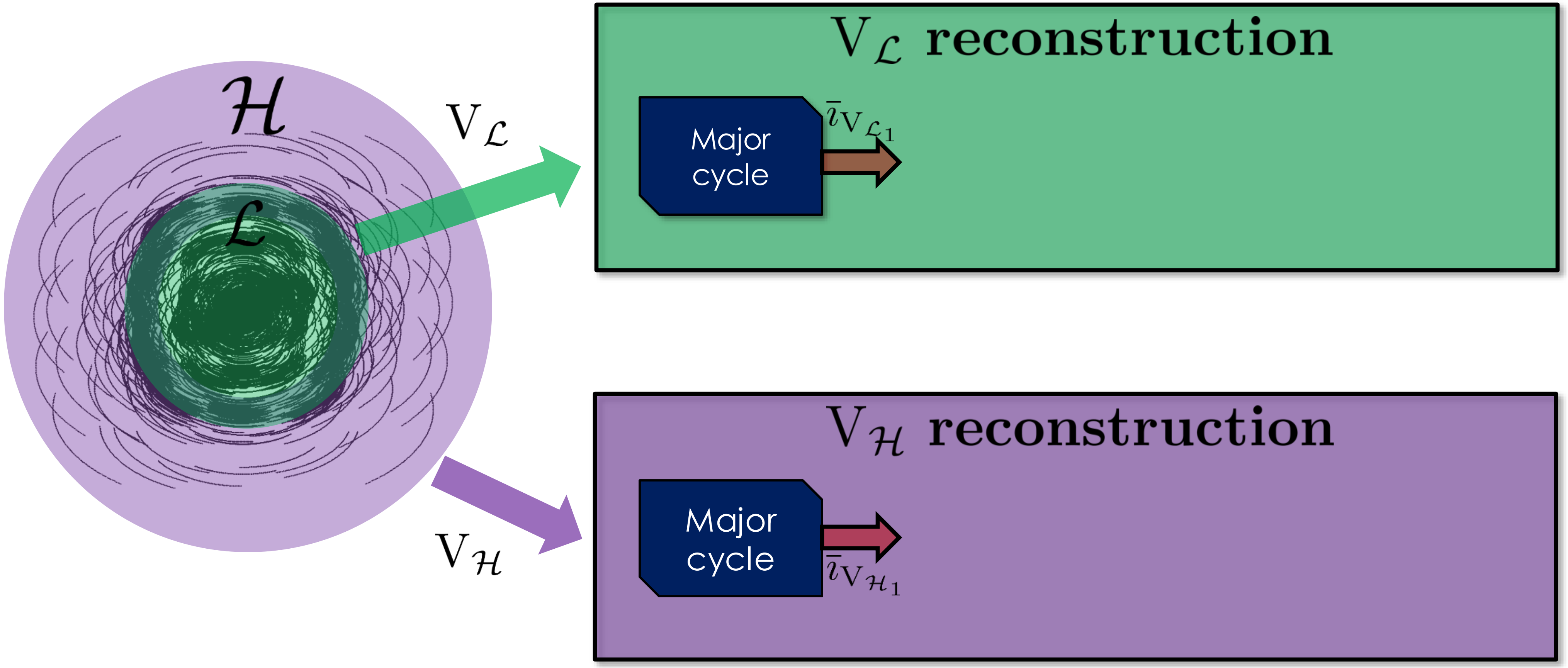
Parallelization Framework



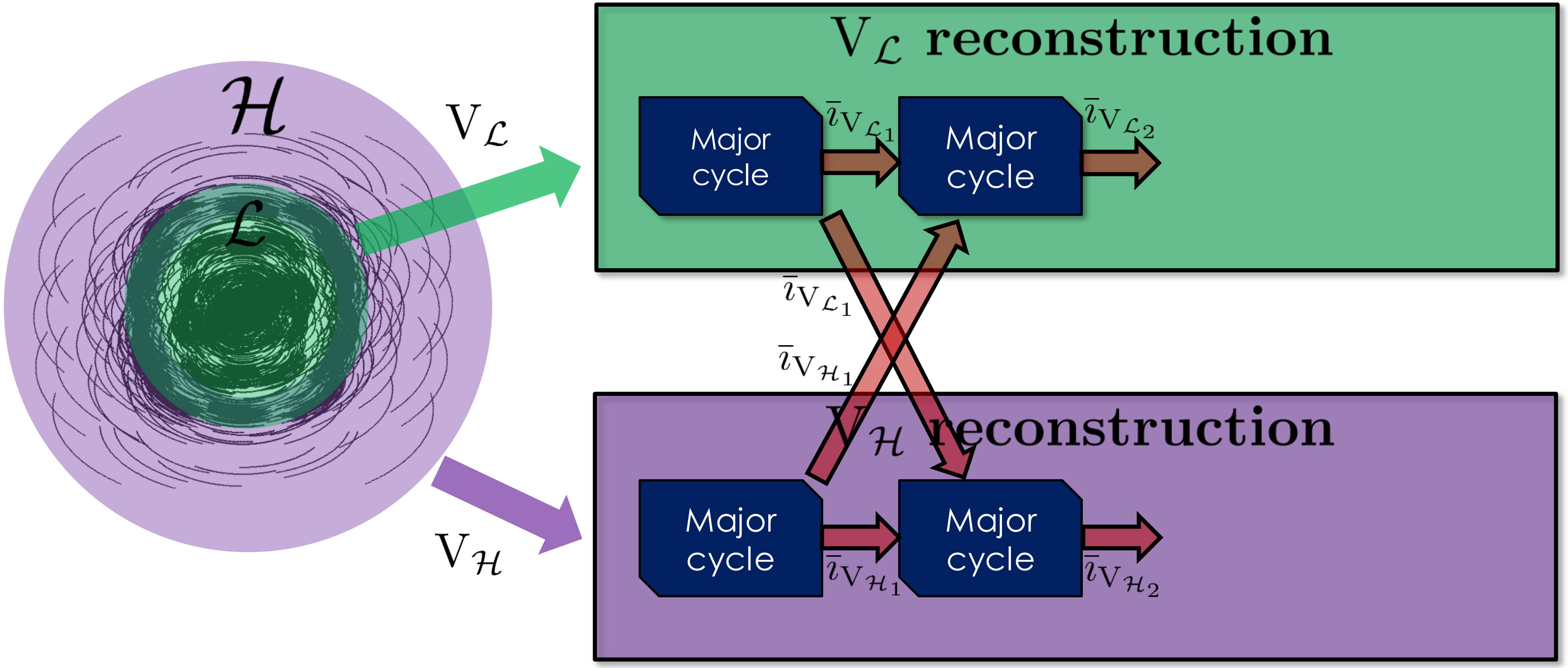
Parallelization Framework



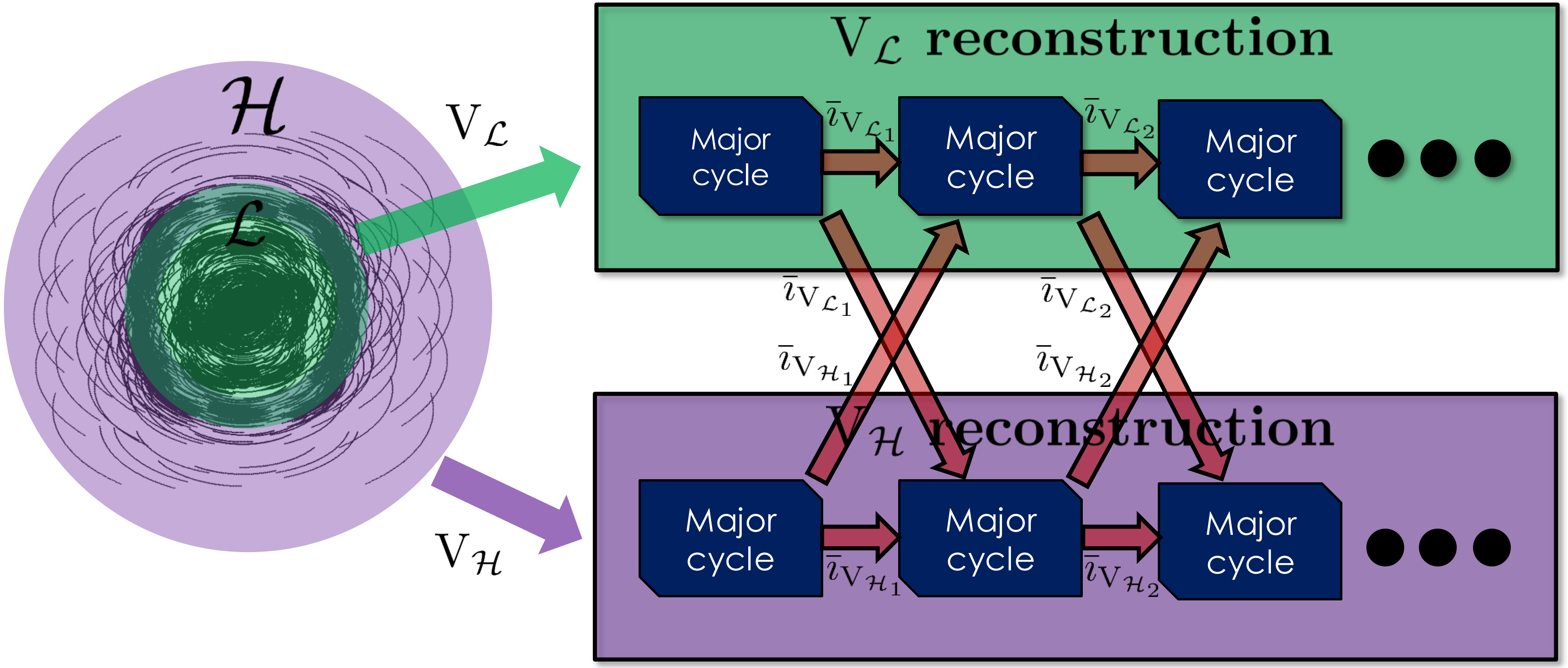
Parallelization Framework



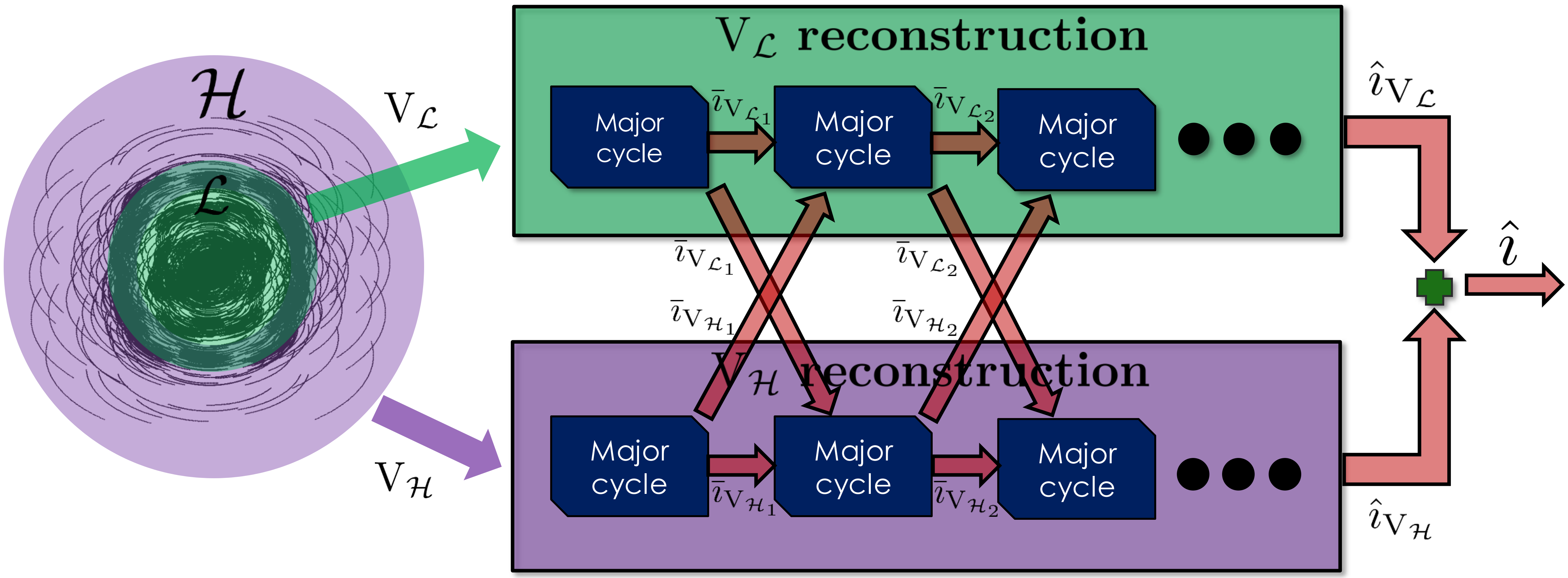
Parallelization Framework



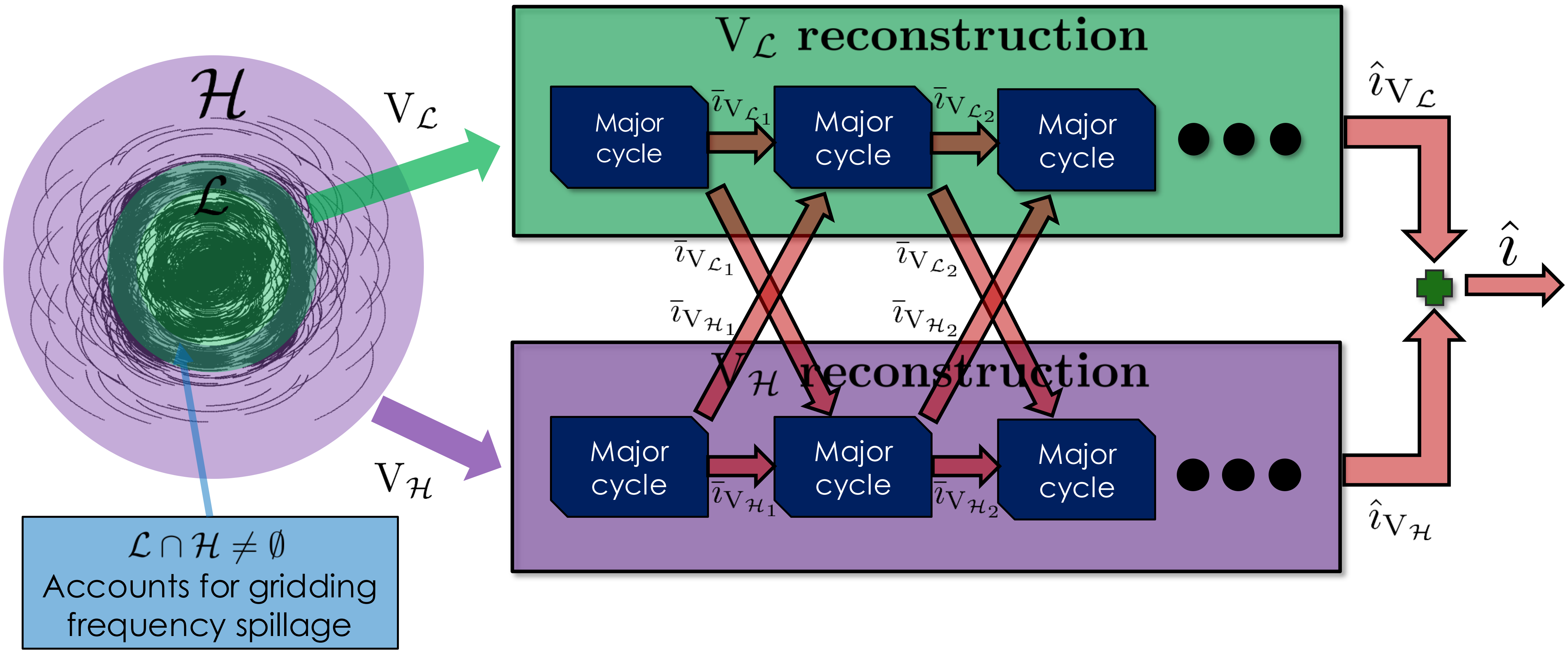
Parallelization Framework



Parallelization Framework



Parallelization Framework



Applying our Framework to existing reconstruction methods

Parallelized L1 reconstruction

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

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

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

[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.

Parallelized L1 reconstruction

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

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

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

Convolution by PSF operator

Wavelet transform operator (Daubechies 1-8)

[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.

Parallelized L1 reconstruction

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

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

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

Convolution by PSF operator

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.

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.

Parallelized L1 reconstruction

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

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

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

Visibility data-fidelity term

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

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

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

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

[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.

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)

$V_{\mathcal{L}}$ deconvolution

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

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

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.

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:

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

[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.

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:

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

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

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

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

[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.

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:

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

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

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

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

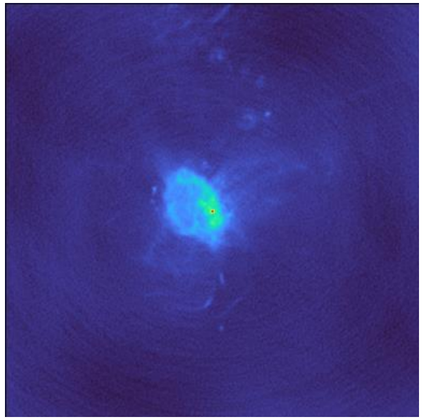
Constants ensure PSFs and dirty images are normalized correctly

[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.

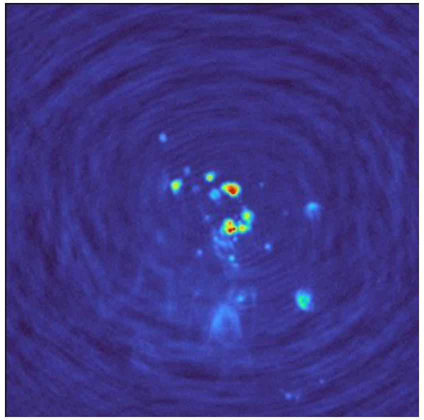
Results (Preliminary)

Results – datasets

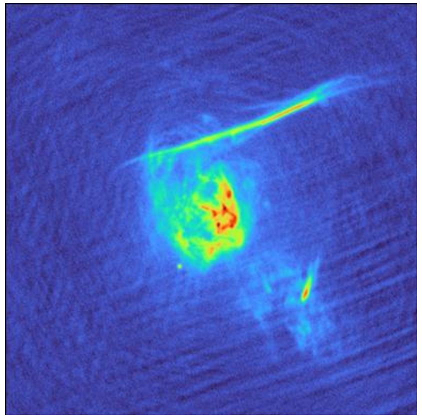
Simulated



Sgr A



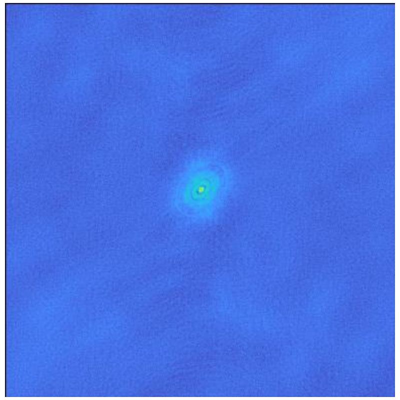
Sgr B2



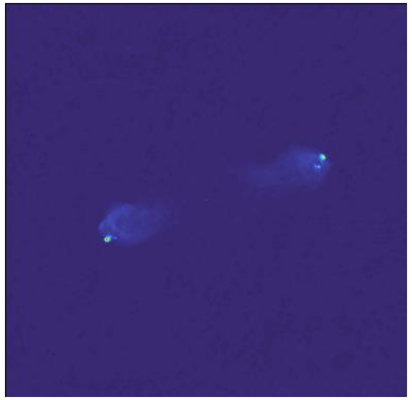
Sgr C

- ➢ Initial images tapered and cutout from 1.28GHz mosaic produced in [1]
- ➢ Visibilities generated with Meerkat, SKA-Mid AA4, and SKA-Low AA4 configurations
- ➢ Observation time of HA=[-2,2] with integration times of 120s, 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 3.815", 0.18", 0.429" respectively
- ➢ Pixel resolutions of 512x512
- ➢ Pseudo declinations also used to vary uv-coverages (-40, -35, -50 respectively)

Real



HL Tau



Cygnus A

- ➢ 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	ℓ	$V_{\mathcal{L}}$	$V_{\mathcal{H}}$	$V_{\mathcal{L}\cap\mathcal{H}}$
Sgr A	30	132648	121140	4188
Sgr B2	20	6989812	2532615	160987
Sgr C	25	9053277	6811266	105183
HL Tau	60	39864808	46243732	605916
Cygnus A	40	41645824	41828032	1227264

Results – partitioned datasets

- Ideally want to create even-sized partitions
- Difficult for Sgr B2 dataset due to SKA-Mid AA4 array (BDA[1] or something similar may be needed)
- Mostly even partitionings for rest of datasets, with the exception of Sgr C which was partitioned slightly more unevenly for testing purposes

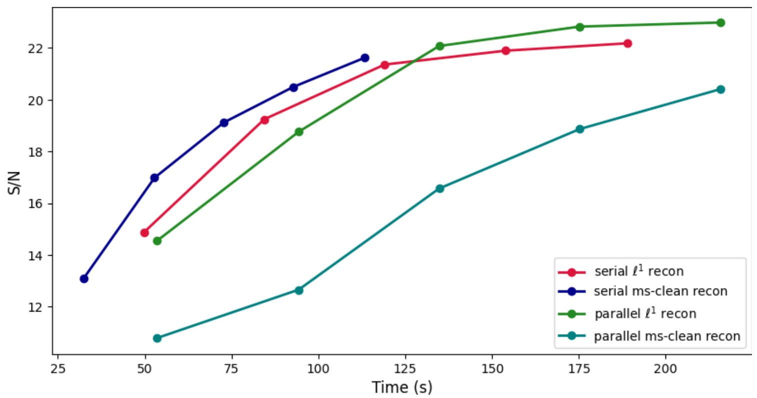
[1] Wijnholds, S. J., A. G. Willis, and S. Salvini. "Baseline-dependent averaging in radio interferometry." *Monthly Notices of the Royal Astronomical Society* 476.2 (2018): 2029-2039.

PS: Parallel ms-clean times are fake!

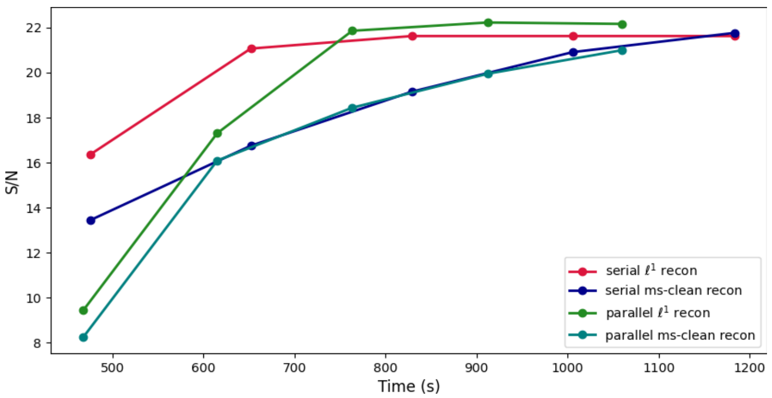
Results – Simulated

S/N compared to ground truths

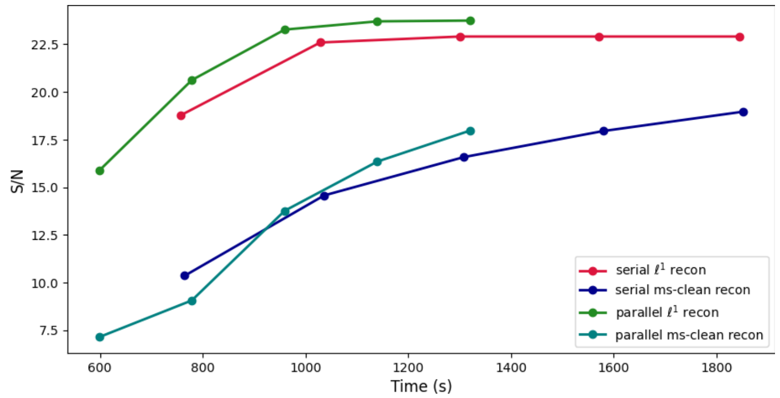
Sgr A

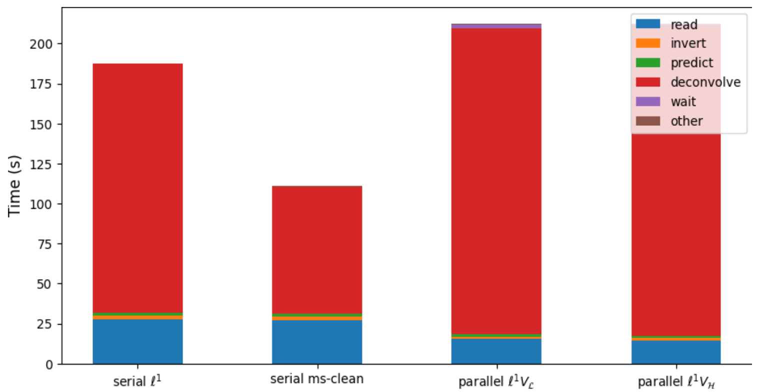


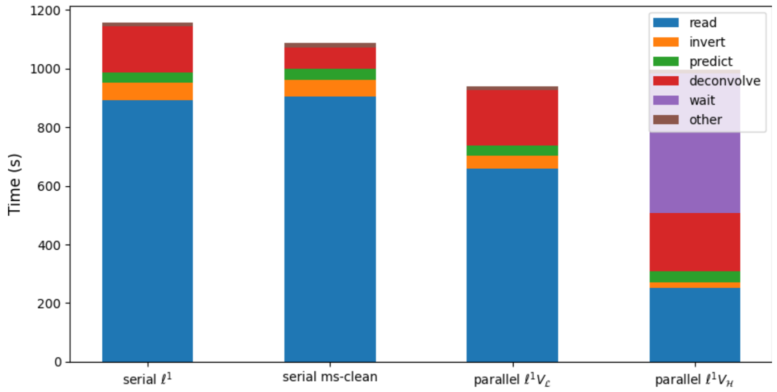
Sgr B2

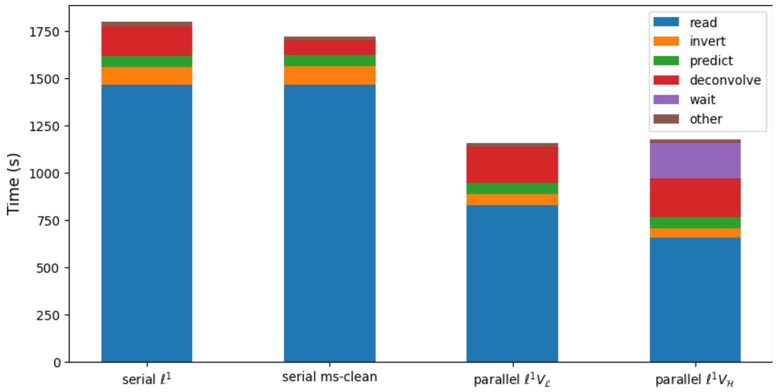


Sgr C





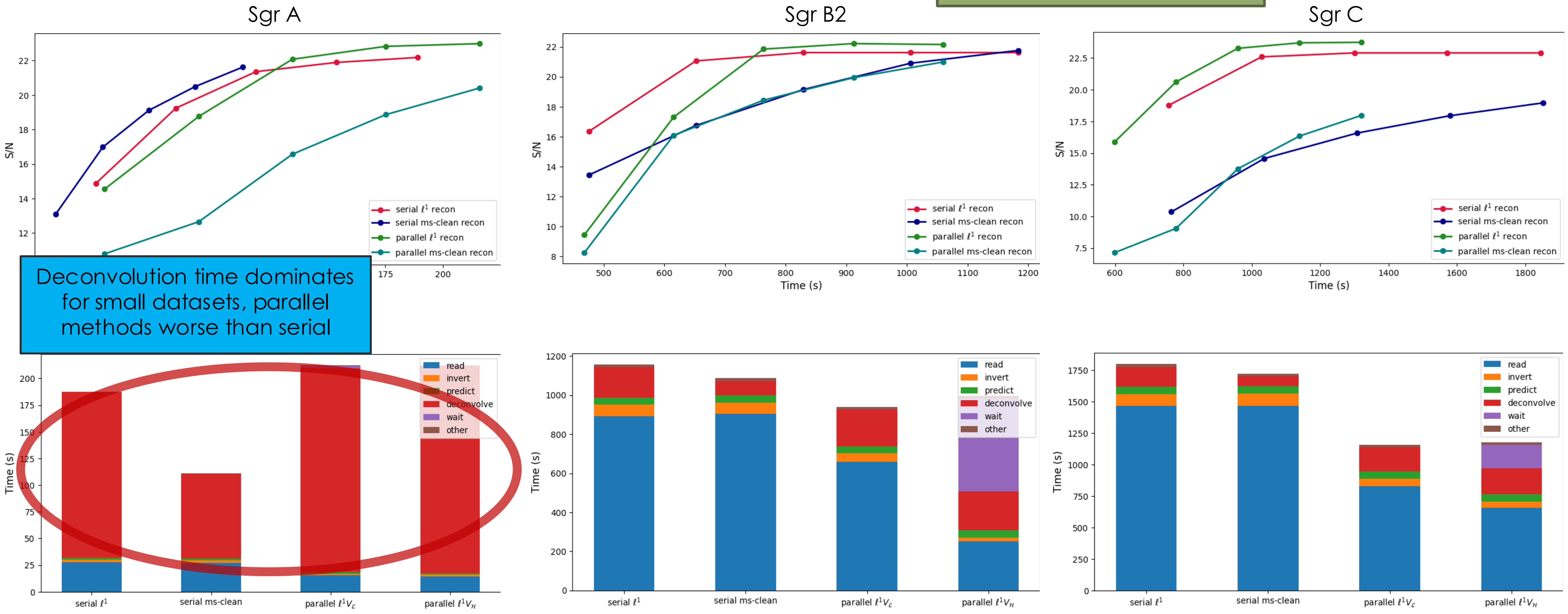




Results – Simulated

S/N compared to ground truths

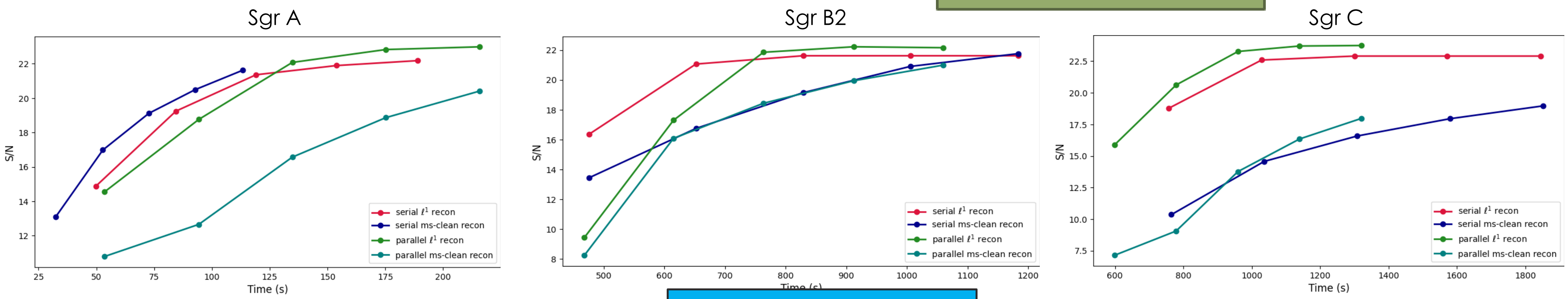
Deconvolution time dominates for small datasets, parallel methods worse than serial



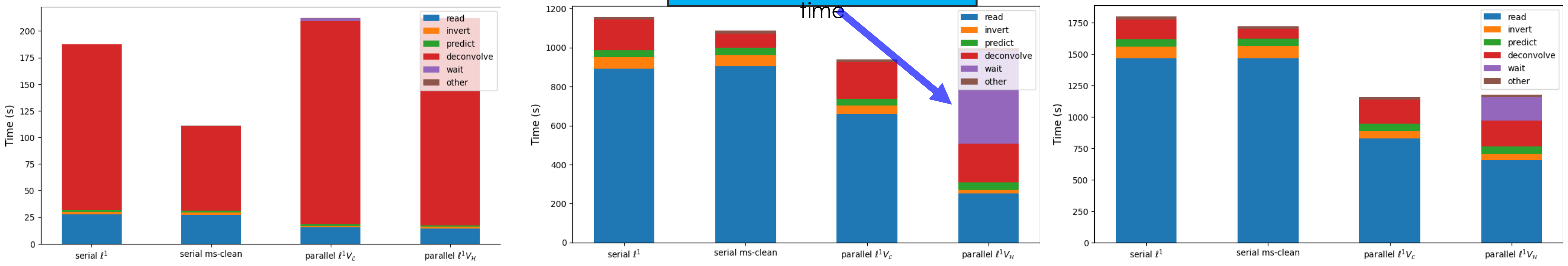
PS: Parallel ms-clean times are fake!

Results – Simulated

S/N compared to ground truths



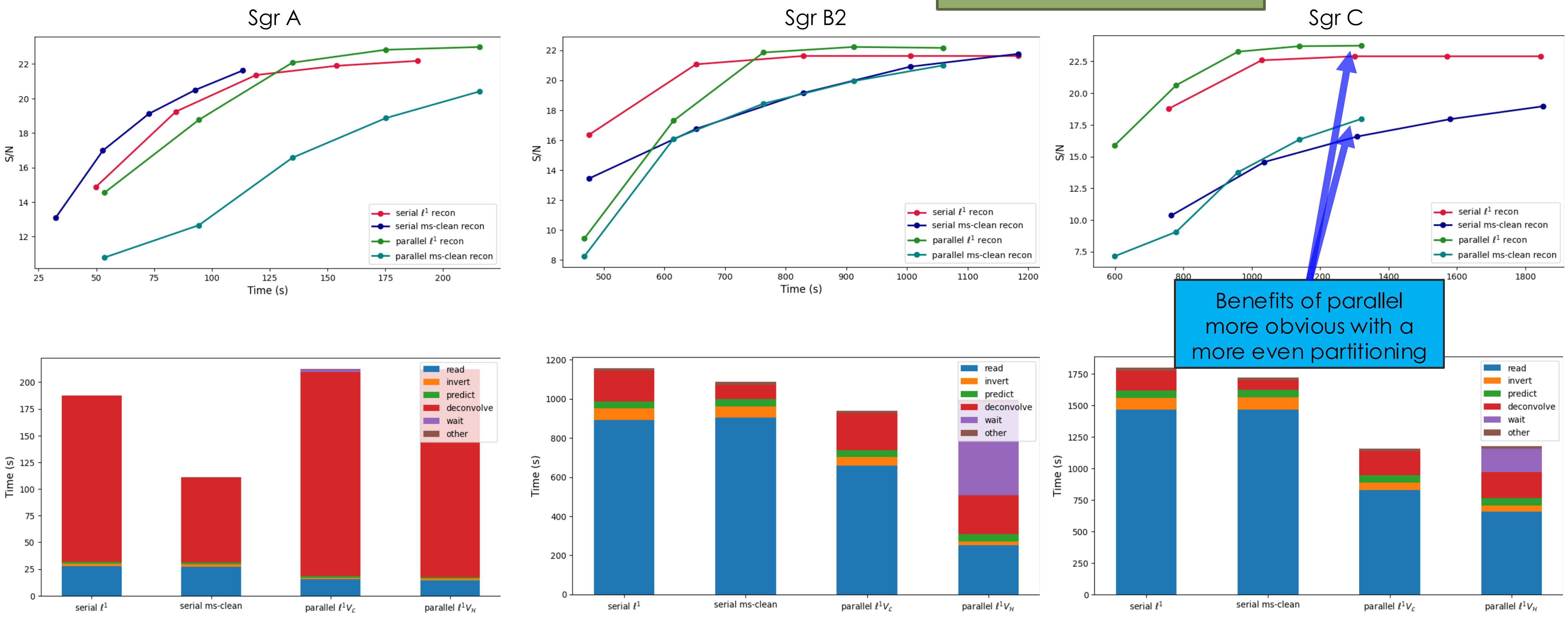
Non-even partitioning results in lots of waiting time



PS: Parallel ms-clean times are fake!

Results – Simulated

S/N compared to ground truths

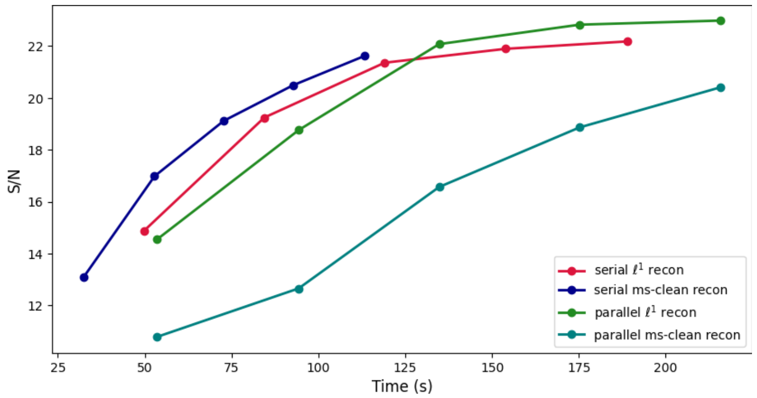


PS: Parallel ms-clean times are fake!

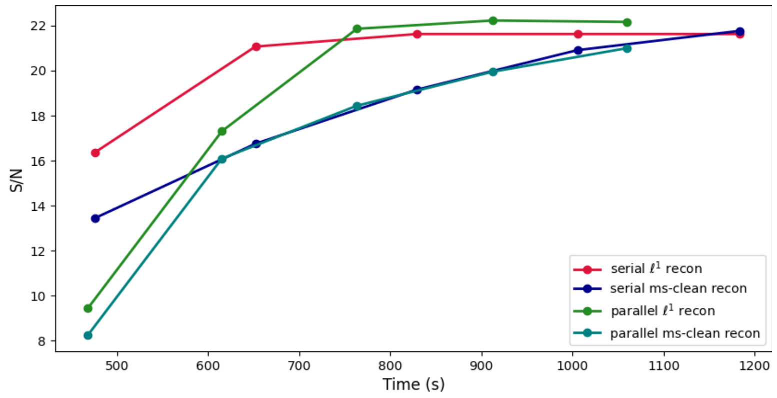
Results – Simulated

S/N compared to ground truths

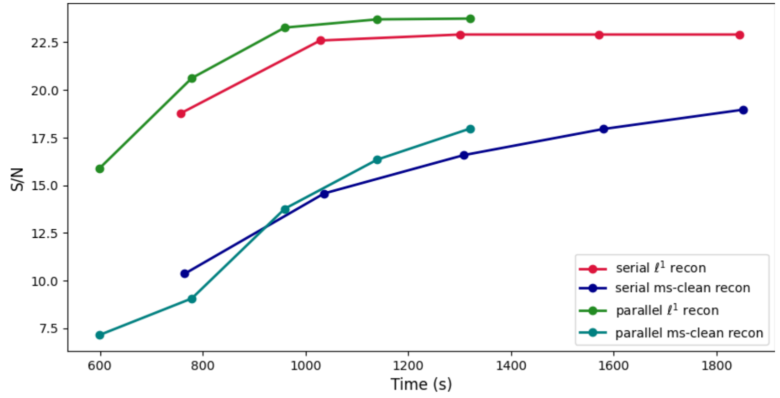
Sgr A

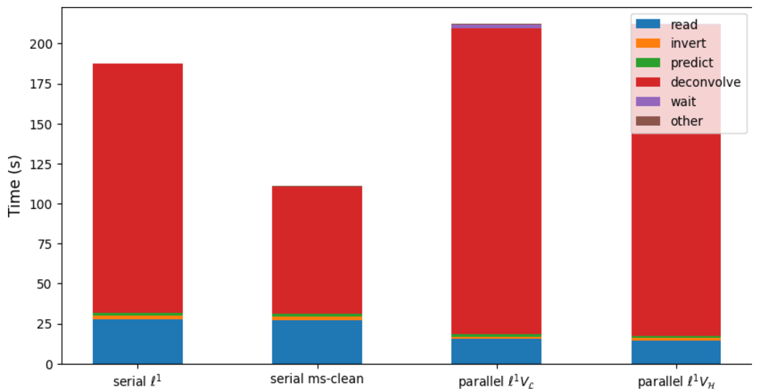


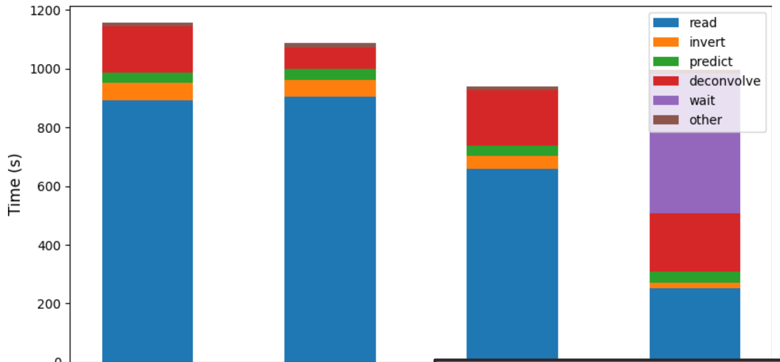
Sgr B2

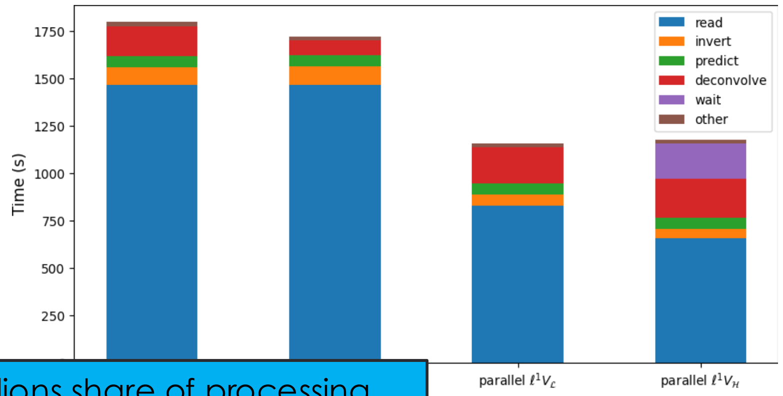


Sgr C





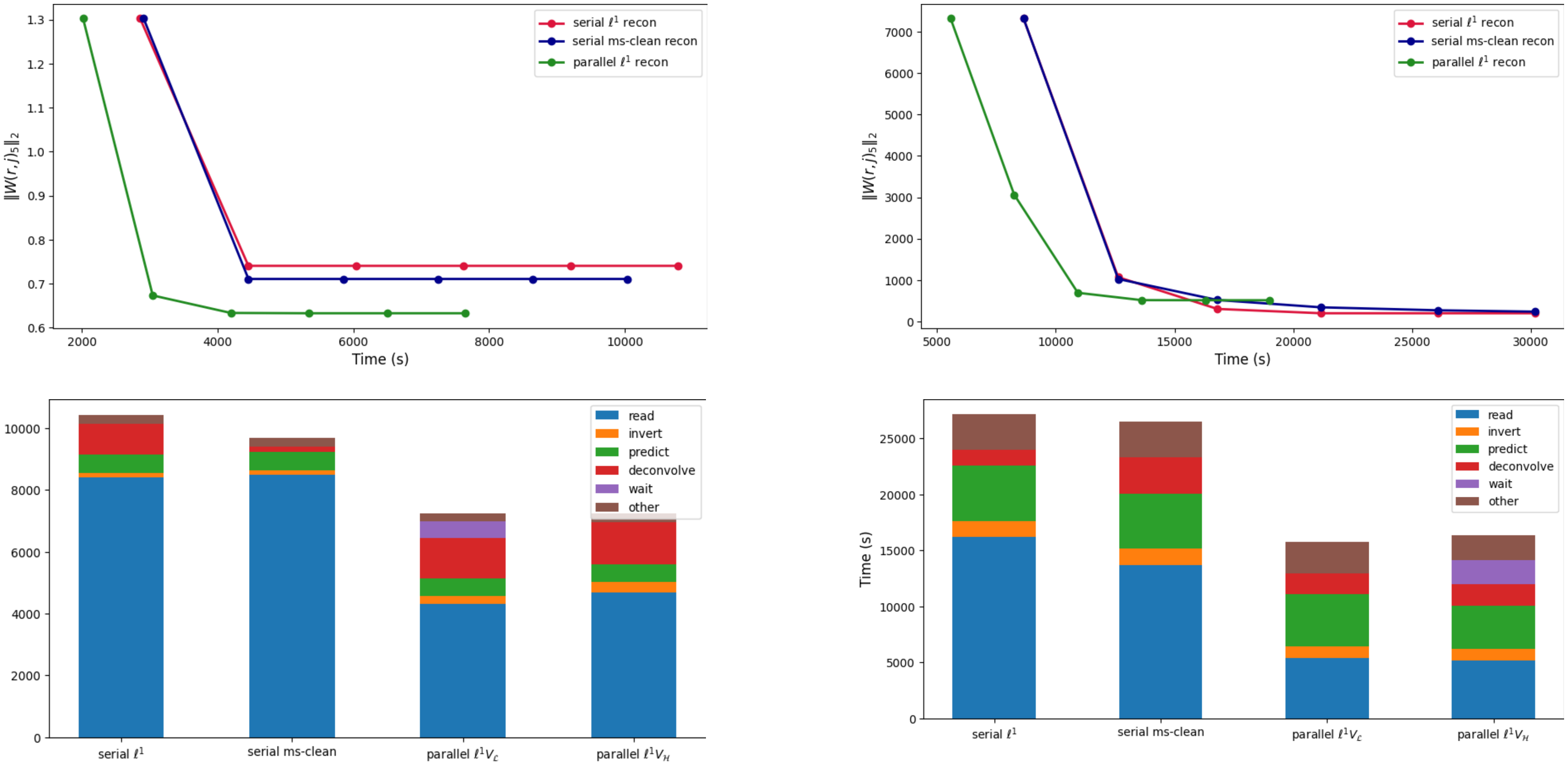




read takes up the lions share of processing time. All visibilities loaded every major cycle. Additional inefficiencies with RASCIL

Results – Real

Compare pixel distributions of residual to reference noise obtain using jackknifing as in[4, 5, 6] in a 5x5 sliding window (as we only have one realization), can use various different statistical tests e.g. [1, 2]. We use Wasserstein difference[3].

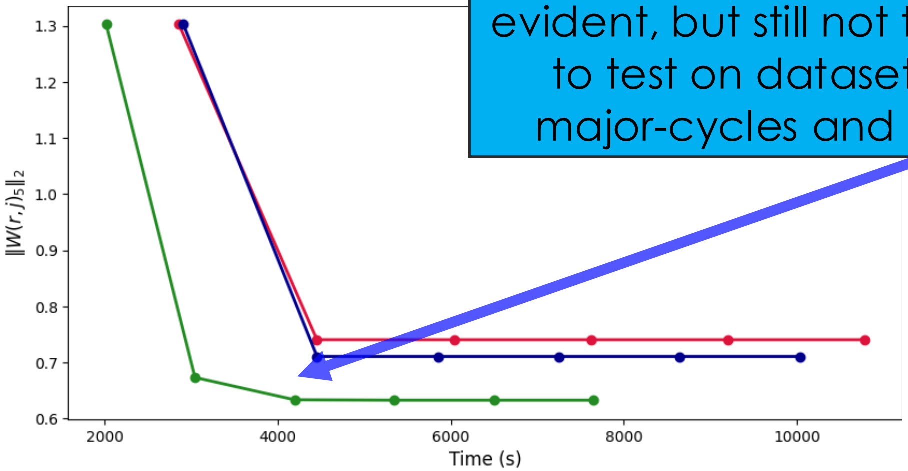


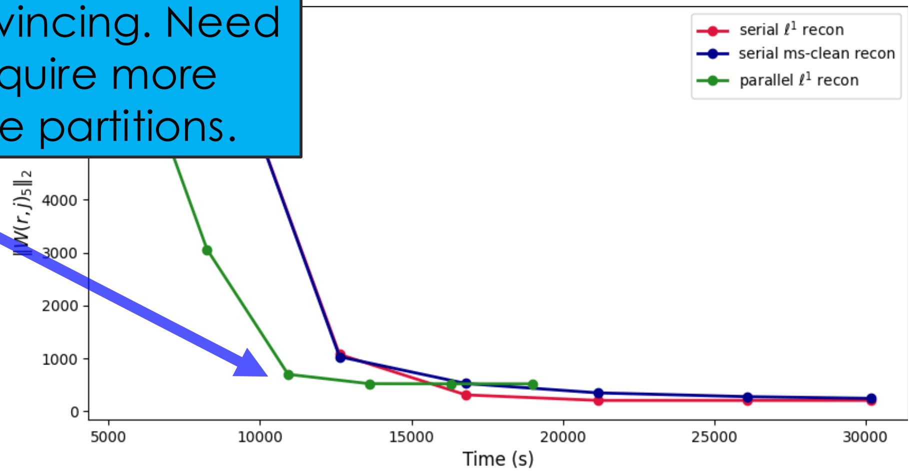
[1] McKnight, Patrick E., and Julius Najab. "Mann-Whitney U Test." *The Corsini encyclopedia of psychology* (2010): 1-1.
[2] Scholz, Fritz W., and Michael A. Stephens. "K-sample Anderson–Darling tests." *Journal of the American Statistical Association* 82.399 (1987): 918-924.
[3] Vaserstein, Leonid Nisonovich. "Markov processes over denumerable products of spaces, describing large systems of automata." *Problemy Peredachi Informatsii* 5.3 (1969): 64-72.
[4] González-López, Jorge, et al. "The ALMA Spectroscopic Survey in the HUDF: Deep 1.2 mm continuum number counts." *The Astrophysical Journal* 897.1 (2020): 91.
[5] Di Mascolo, Luca, et al. "Forming intracluster gas in a galaxy protocluster at a redshift of 2.16." *Nature* 615.7954 (2023): 809-812.

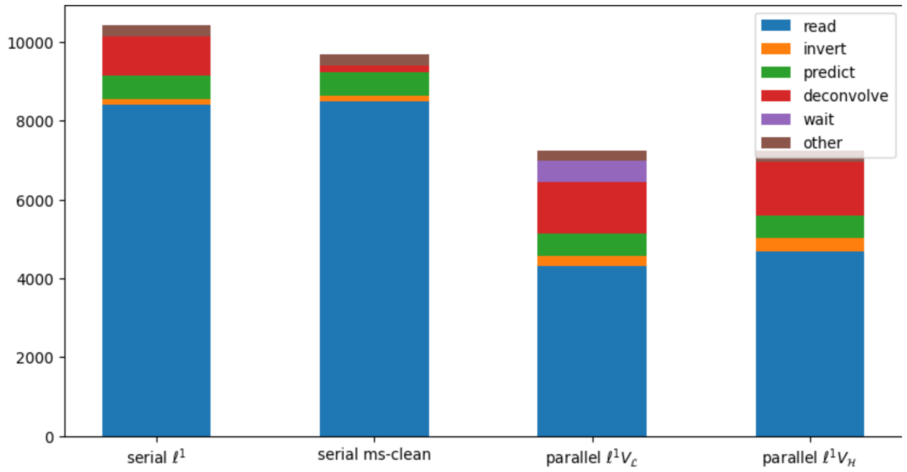
Results – Real

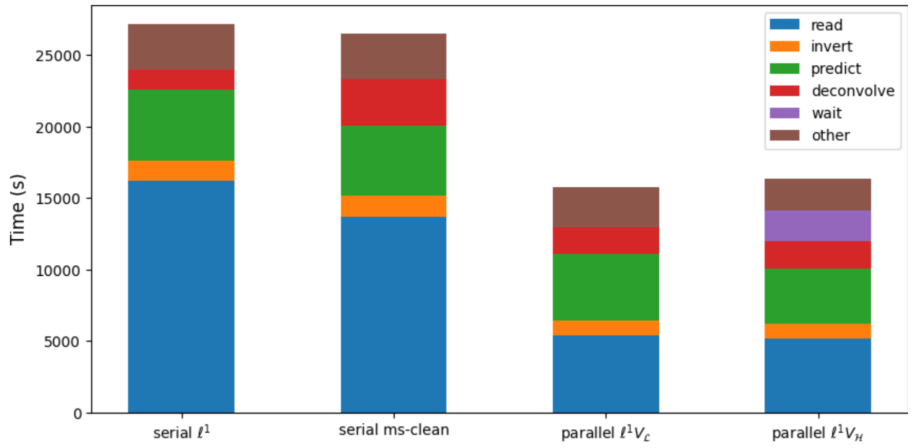
Compare pixel distributions of residual to reference noise obtain using jackknifing as in[4, 5, 6] in a 5x5 sliding window (as we only have one realization), can use various different statistical tests e.g. [1, 2]. We use Wasserstein difference[3].

Benefits of parallel framework more evident, but still not that convincing. Need to test on datasets that require more major-cycles and with more partitions.





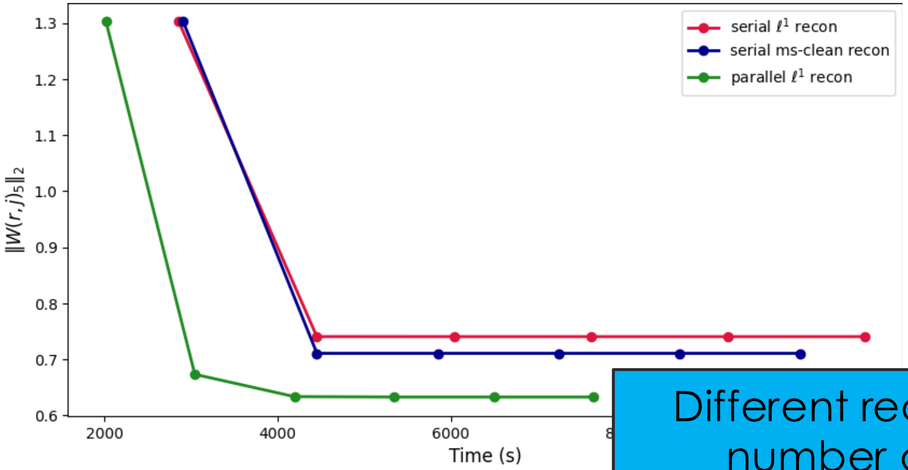


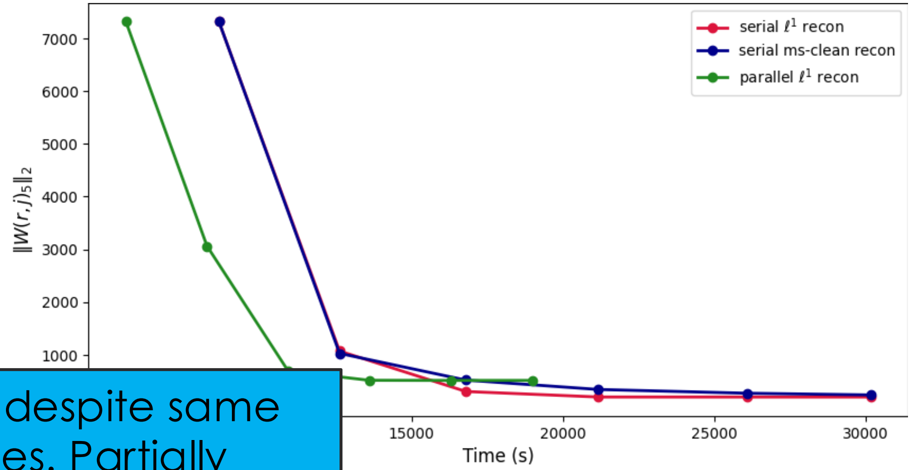


[1] McKnight, Patrick E., and Julius Najab. "Mann-Whitney U Test." *The Corsini encyclopedia of psychology* (2010): 1-1.
[2] Scholz, Fritz W., and Michael A. Stephens. "K-sample Anderson–Darling tests." *Journal of the American Statistical Association* 82.399 (1987): 918-924.
[3] Vaserstein, Leonid Nisonovich. "Markov processes over denumerable products of spaces, describing large systems of automata." *Problemy Peredachi Informatsii* 5.3 (1969): 64-72.
[4] González-López, Jorge, et al. "The ALMA Spectroscopic Survey in the HUDF: Deep 1.2 mm continuum number counts." *The Astrophysical Journal* 897.1 (2020): 91.
[5] Di Mascolo, Luca, et al. "Forming intracluster gas in a galaxy protocluster at a redshift of 2.16." *Nature* 615.7954 (2023): 809-812.

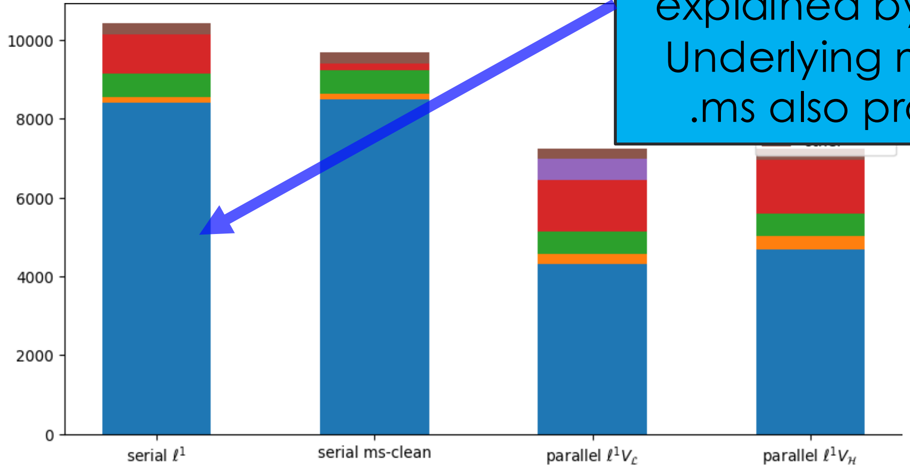
Results – Real

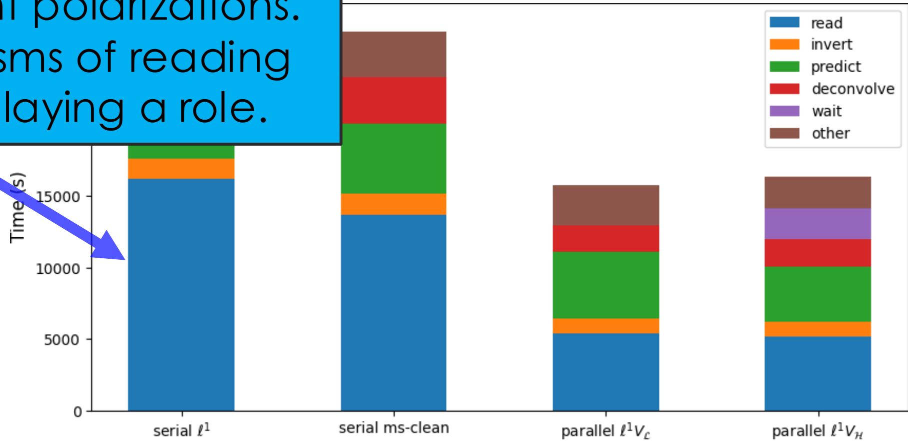
Compare pixel distributions of residual to reference noise obtain using jackknifing as in[4, 5, 6] in a 5x5 sliding window (as we only have one realization), can use various different statistical tests e.g. [1, 2]. We use Wasserstein difference[3].





Different read times despite same number of visibilities. Partially explained by different polarizations. Underlying mechanisms of reading .ms also probably playing a role.

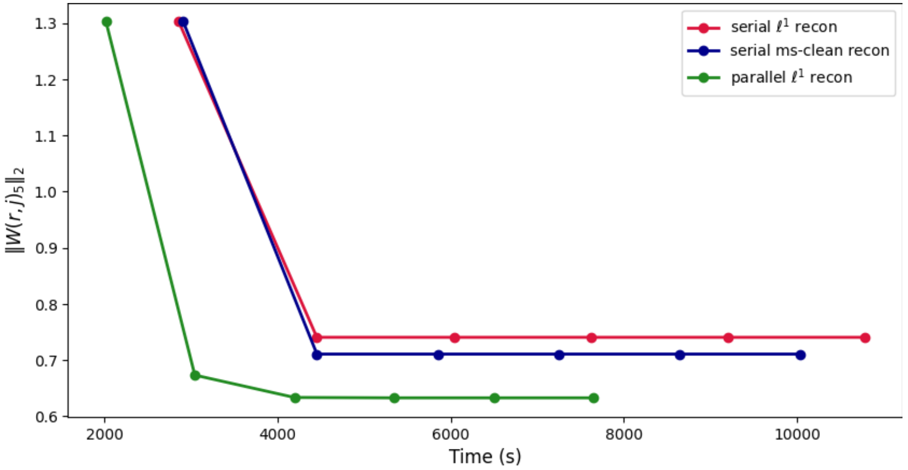


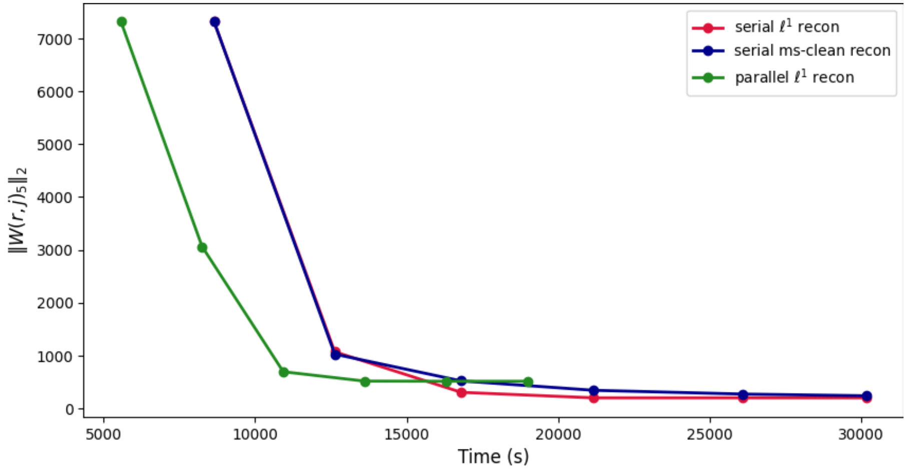


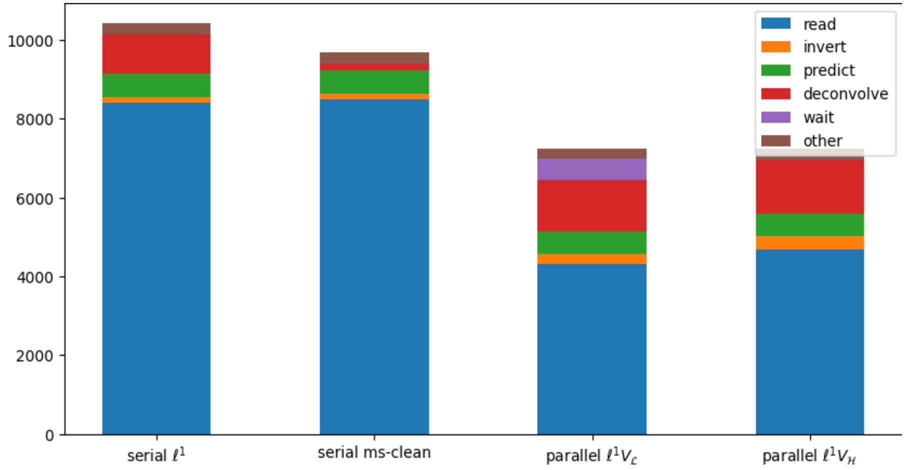
[1] McKnight, Patrick E., and Julius Najab. "Mann-Whitney U Test." *The Corsini encyclopedia of psychology* (2010): 1-1.
[2] Scholz, Fritz W., and Michael A. Stephens. "K-sample Anderson–Darling tests." *Journal of the American Statistical Association* 82.399 (1987): 918-924.
[3] Vaserstein, Leonid Nisonovich. "Markov processes over denumerable products of spaces, describing large systems of automata." *Problemy Peredachi Informatsii* 5.3 (1969): 64-72.
[4] González-López, Jorge, et al. "The ALMA Spectroscopic Survey in the HUDF: Deep 1.2 mm continuum number counts." *The Astrophysical Journal* 897.1 (2020): 91.
[5] Di Mascolo, Luca, et al. "Forming intracluster gas in a galaxy protocluster at a redshift of 2.16." *Nature* 615.7954 (2023): 809-812.

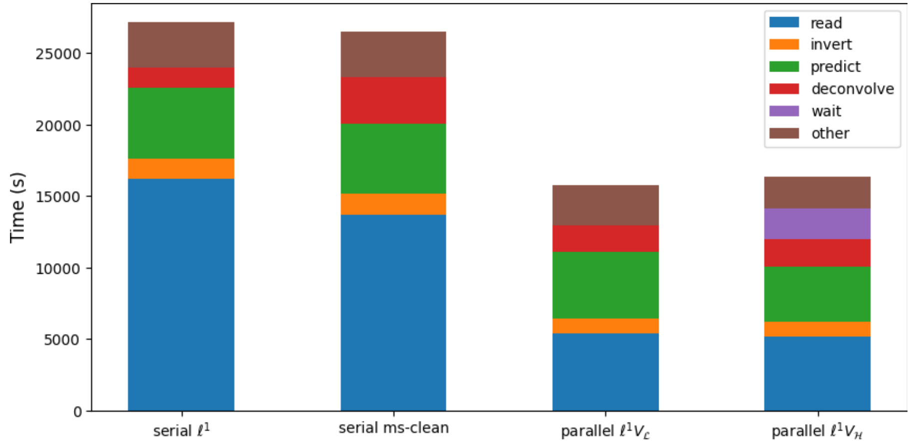
Results – Real

Compare pixel distributions of residual to reference noise obtain using jackknifing as in[4, 5, 6] in a 5x5 sliding window (as we only have one realization), can use various different statistical tests e.g. [1, 2]. We use Wasserstein difference[3].





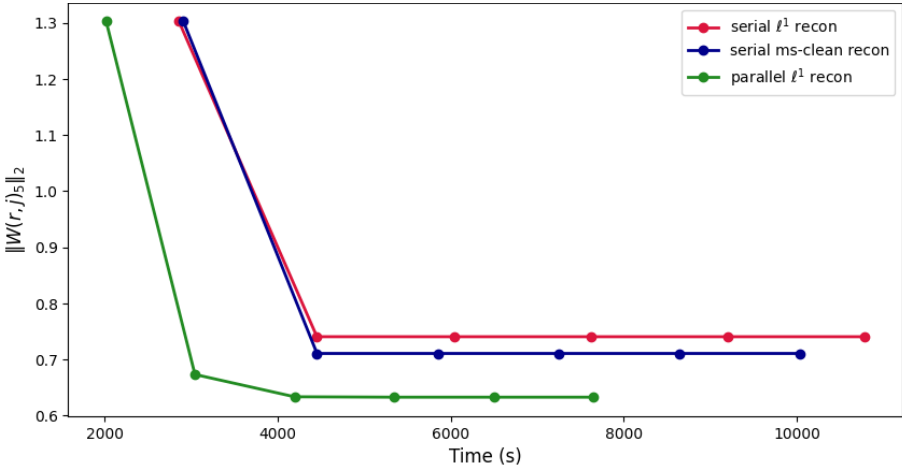


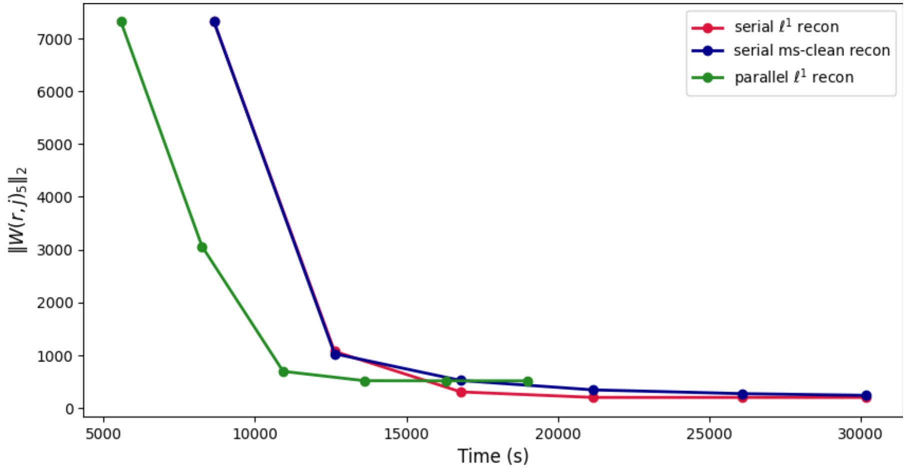


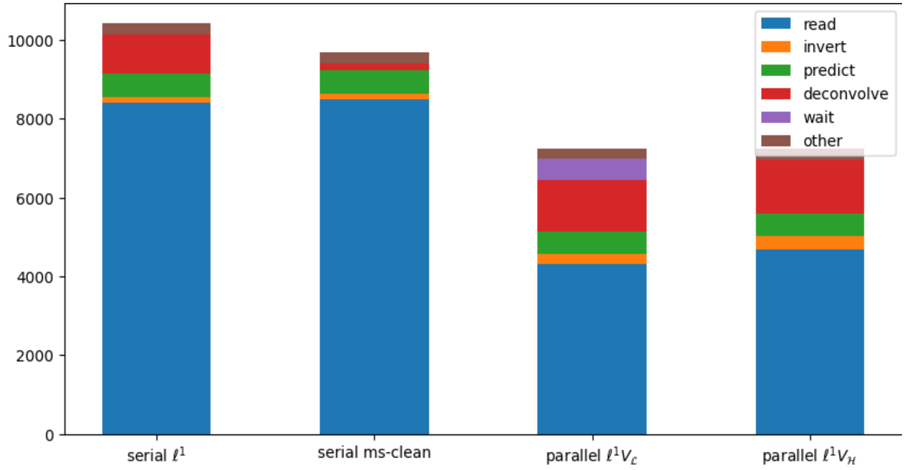
[1] McKnight, Patrick E., and Julius Najab. "Mann-Whitney U Test." *The Corsini encyclopedia of psychology* (2010): 1-1.
[2] Scholz, Fritz W., and Michael A. Stephens. "K-sample Anderson–Darling tests." *Journal of the American Statistical Association* 82.399 (1987): 918-924.
[3] Vaserstein, Leonid Nisonovich. "Markov processes over denumerable products of spaces, describing large systems of automata." *Problemy Peredachi Informatsii* 5.3 (1969): 64-72.
[4] González-López, Jorge, et al. "The ALMA Spectroscopic Survey in the HUDF: Deep 1.2 mm continuum number counts." *The Astrophysical Journal* 897.1 (2020): 91.
[5] Di Mascolo, Luca, et al. "Forming intracluster gas in a galaxy protocluster at a redshift of 2.16." *Nature* 615.7954 (2023): 809-812.

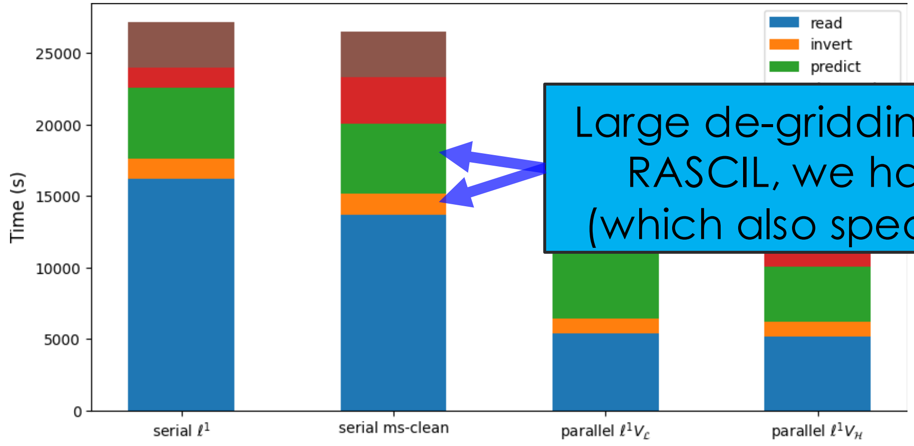
Results – Real

Compare pixel distributions of residual to reference noise obtain using jackknifing as in[4, 5, 6] in a 5x5 sliding window (as we only have one realization), can use various different statistical tests e.g. [1, 2]. We use Wasserstein difference[3].









Large de-gridding times due to RASCIL, we have fixed this (which also sped up the read)

[1] Mc... and Julius Najab. "Mann-Whitney U Test." *The Corsini encyclopedia of psychology* (2010): 1-1.
[2] Sc... Michael A. Stephens. "K-sample Anderson–Darling tests." *Journal of the American Statistical Assoc...*: 918-924.
[3] Vaserstein, Leonid Nisonovich. "Markov processes over denumerable products of spaces, describing large systems of automata." *Problemy Peredachi Informatsii* 5.3 (1969): 64-72.
[4] González-López, Jorge, et al. "The ALMA Spectroscopic Survey in the HUDF: Deep 1.2 mm continuum number counts." *The Astrophysical Journal* 897.1 (2020): 91.
[5] Di Mascolo, Luca, et al. "Forming intracluster gas in a galaxy protocluster at a redshift of 2.16." *Nature* 615.7954 (2023): 809-812.

Conclusions and Future Work

To conclude:

- Parallelization framework for reconstructing radio-interferometric images by baseline length
- Applied to both ms-clean and L1 regularized methods
- Showing promising results but needs further testing
- Some drawbacks include the methods being more tricky to regularize, and needing at least two major-cycles.

Future work:

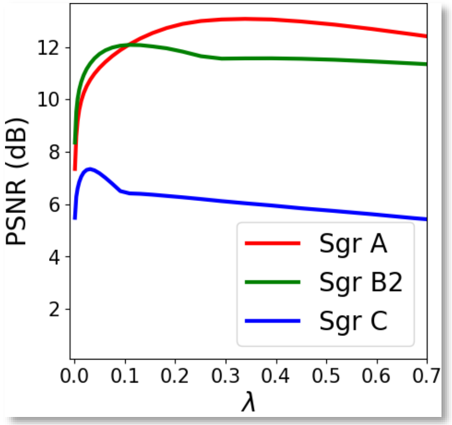
- Investigate more partitions and datasets that require more major-cycles to reconstruct
- More optimized implementation with profiling, preferably in a more performant language and not being based on RASCIL.
- Investigate possible visibility reduction techniques for better partitioning for SKA-Mid.

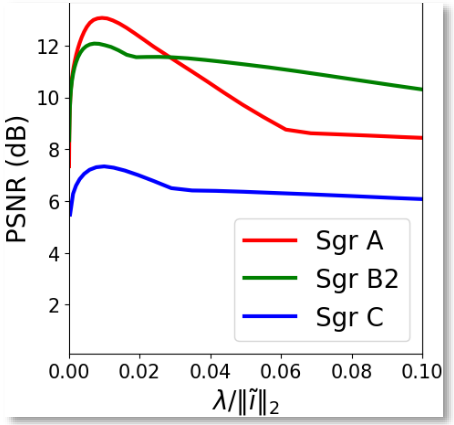
Thank you! Questions?

Appendices

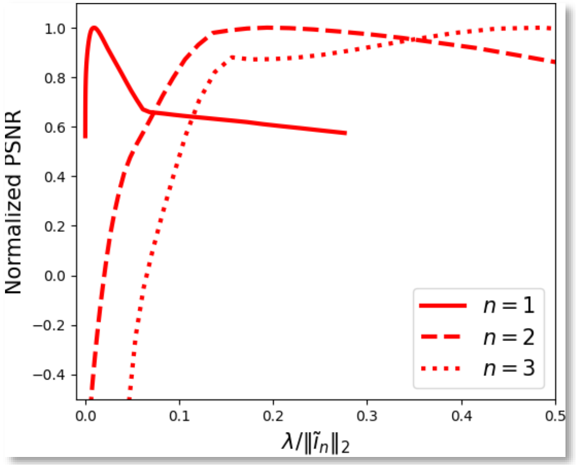
Selection of λ

Across different datasets for first major-cycle:



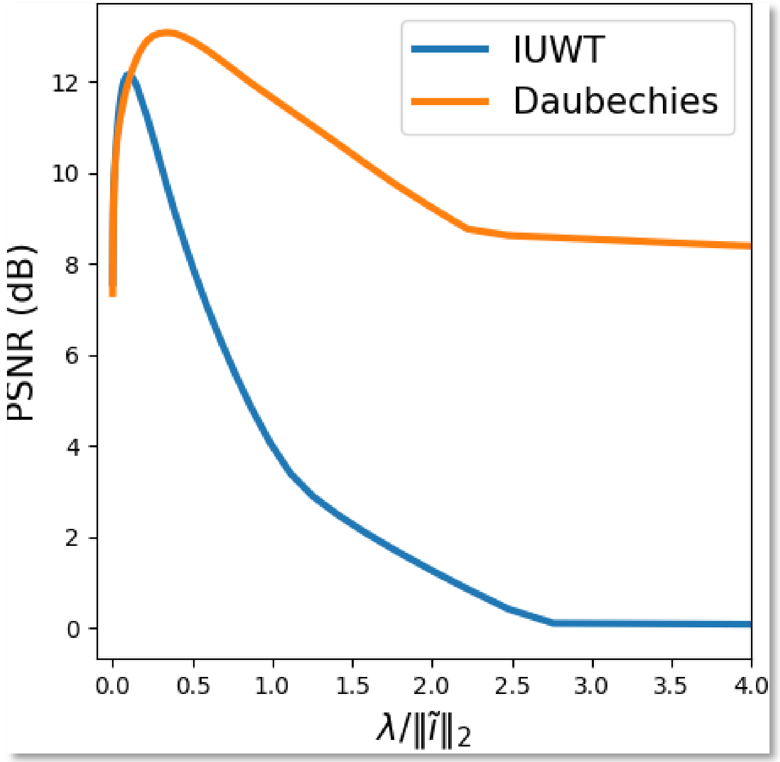


Across first three major-cycles for Sgr A dataset:

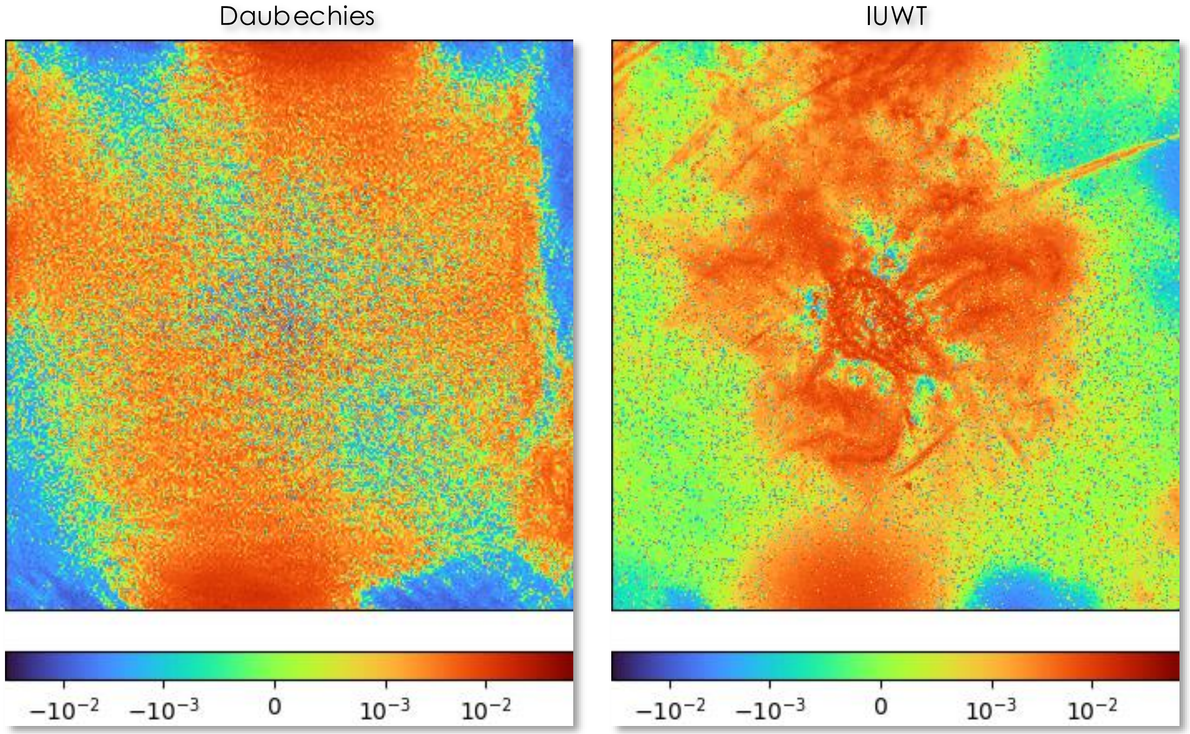


Results suggest that lambda should be normalized by the norm of the image, and be increased as the major-cycles progress to maximize RMSE/PSNR.

IUWT vs Daubechies



First major-cycle residuals for Sgr A

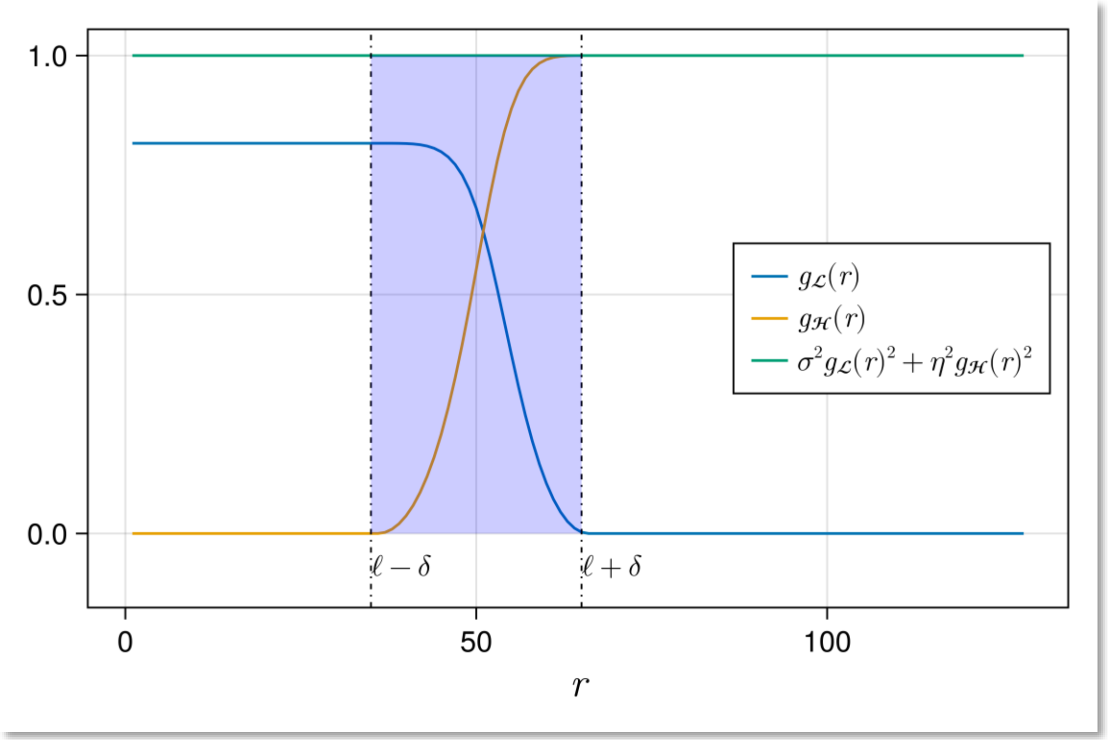


IUWT seems worse at reconstructing large-scale extended emissions, possibly due to its isotropic nature.

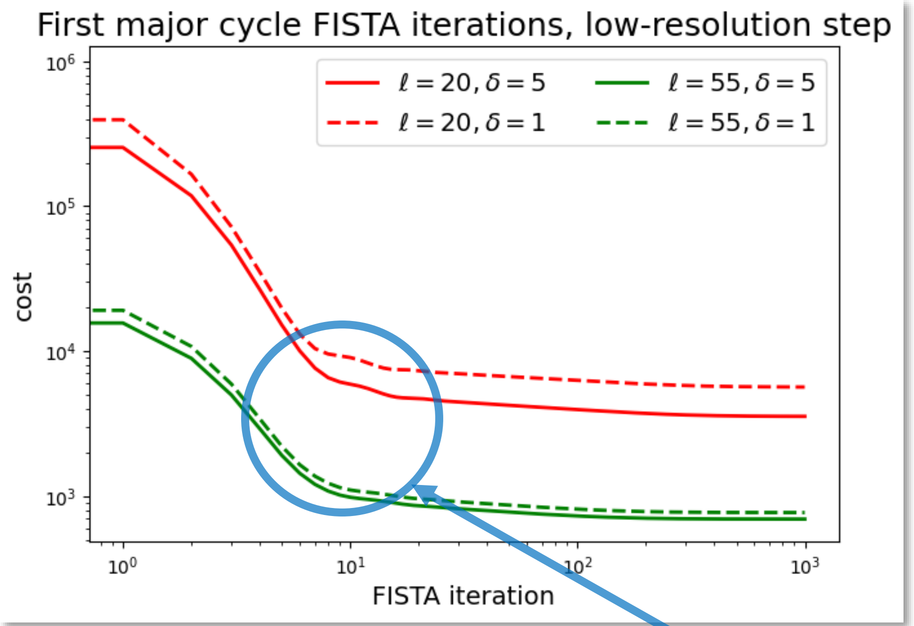
Filters

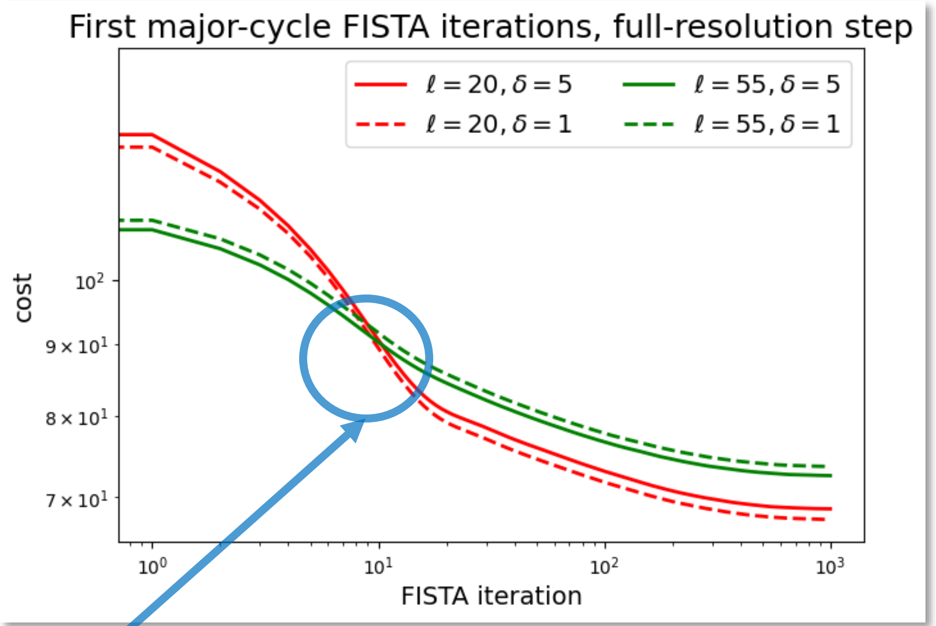
$$r > \ell + \delta : \quad |g_{\mathcal{H}}(r)|^2 = 1/\sigma^2, \; g_{\mathcal{L}}(u) = 0$$
$$r < \ell - \delta : \quad g_{\mathcal{H}}(r) = 0, \; |g_{\mathcal{L}}(r)|^2 = 1/\eta^2$$
$$\ell - \delta < r < \ell + \delta : \quad \sigma^2 |g_{\mathcal{H}}(r)|^2 + \eta^2 |g_{\mathcal{L}}(r)|^2 = 1$$

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



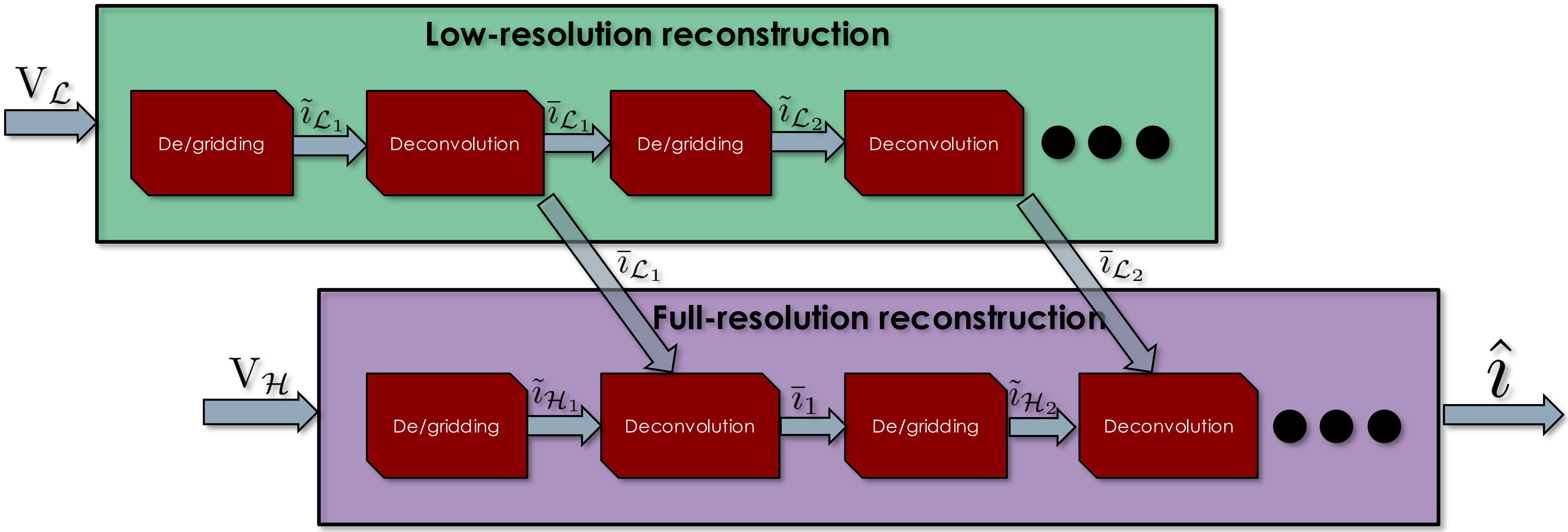
Results: Partition configuration affect on reconstruction accuracy and speed, FISTA





Convergence roughly the same, gradient of cost function remains very similar between partition configurations and starts tapering off around 10-25 iterations.

Pipelined parallelization strategy



Pipelined parallelization strategy

- Low-resolution step remains the same
- For the full-resolution step:

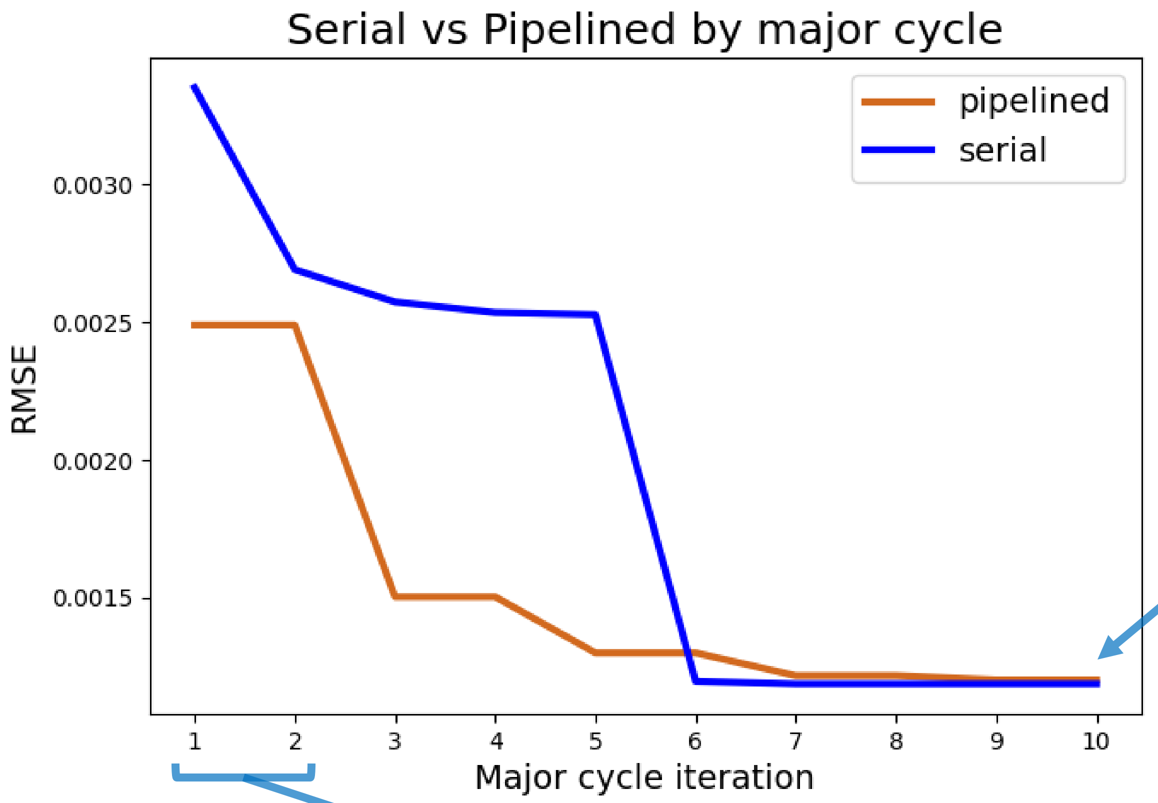
$$l_n = \sum_{j=1}^{N} \bar{\imath}_{\mathcal{L}_j} - \sum_{j=1}^{n-1} \bar{\imath}_j = \hat{\imath}_{\mathcal{L}} - \hat{\imath}_{n-1}$$

changes to:

$$l_n = \sum_{j=1}^{n} \bar{\imath}_{\mathcal{L}_j} - \sum_{j=1}^{n-1} \bar{\imath}_j = \hat{\imath}_{\mathcal{L}_{n-1}} - \hat{\imath}_{n-1} + \bar{\imath}_{\mathcal{L}_n}$$

- Can result in waiting if computation costs of each step not similar
  - Asynchronous strategy can alleviate this somewhat
- Quality upper bound of serial (but most likely slightly worse)
- One image transmitted per major cycle

Results: Reconstruction quality of pipelined parallel method

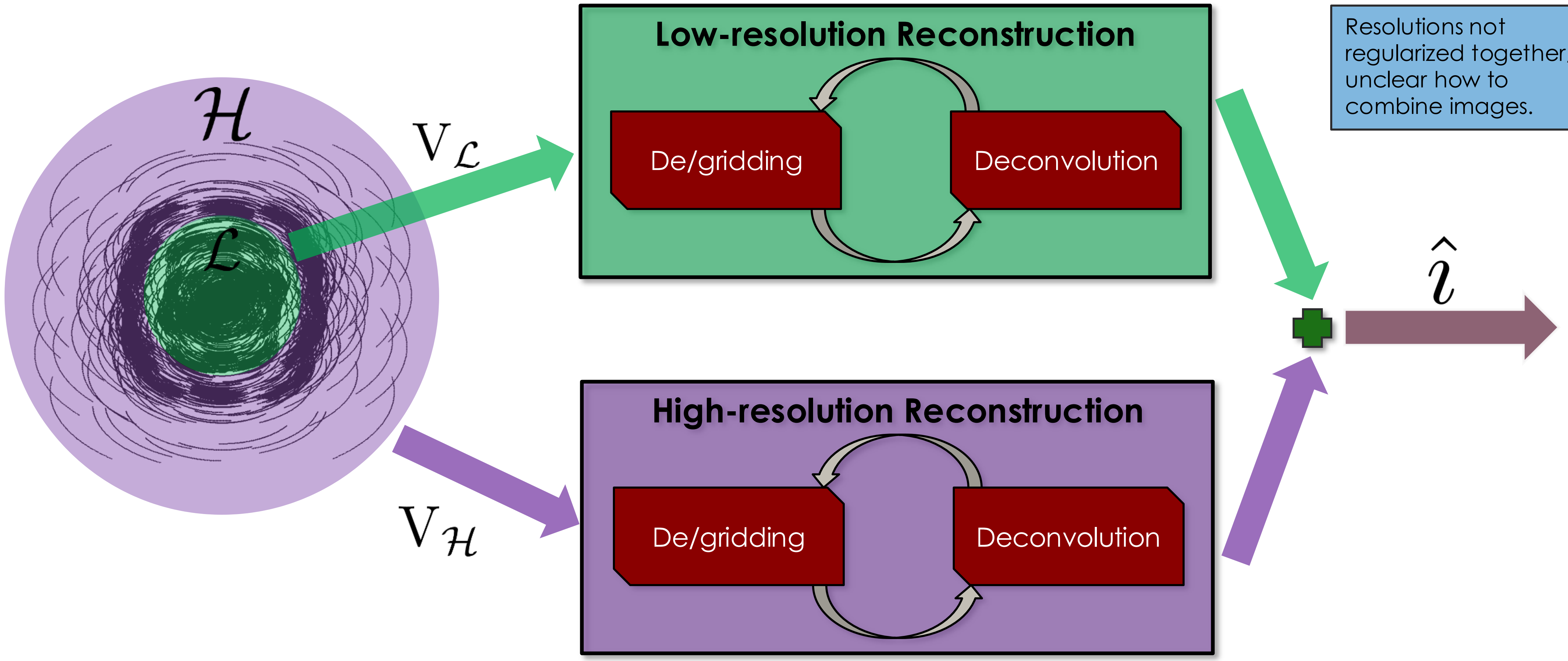


As expected, converges to similar result but slightly worse

For pipelined, only showing full-resolution reconstruction after 1 low and 1 full-resolution major-cycle

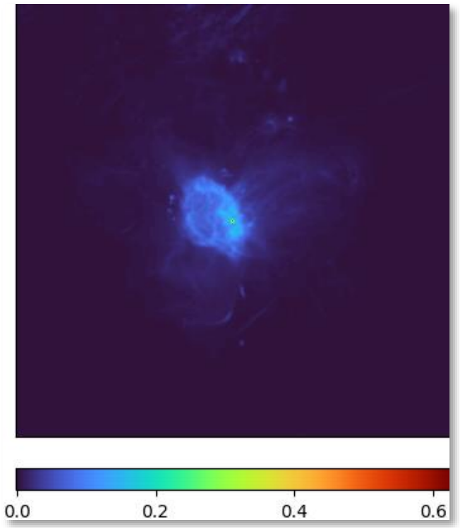
Trading off quality for parallelism

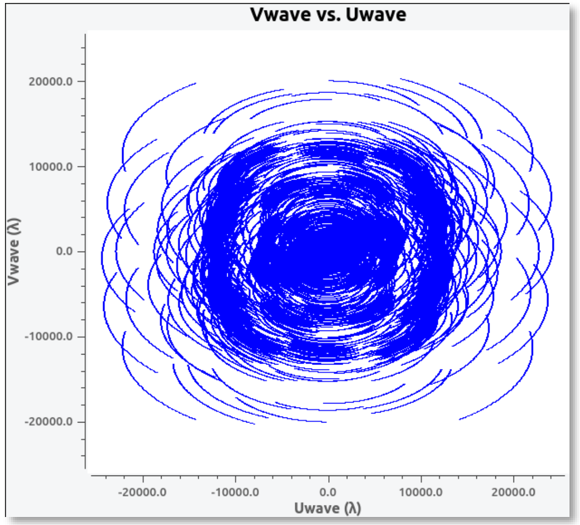
Parallelization by baseline length – A naïve approach

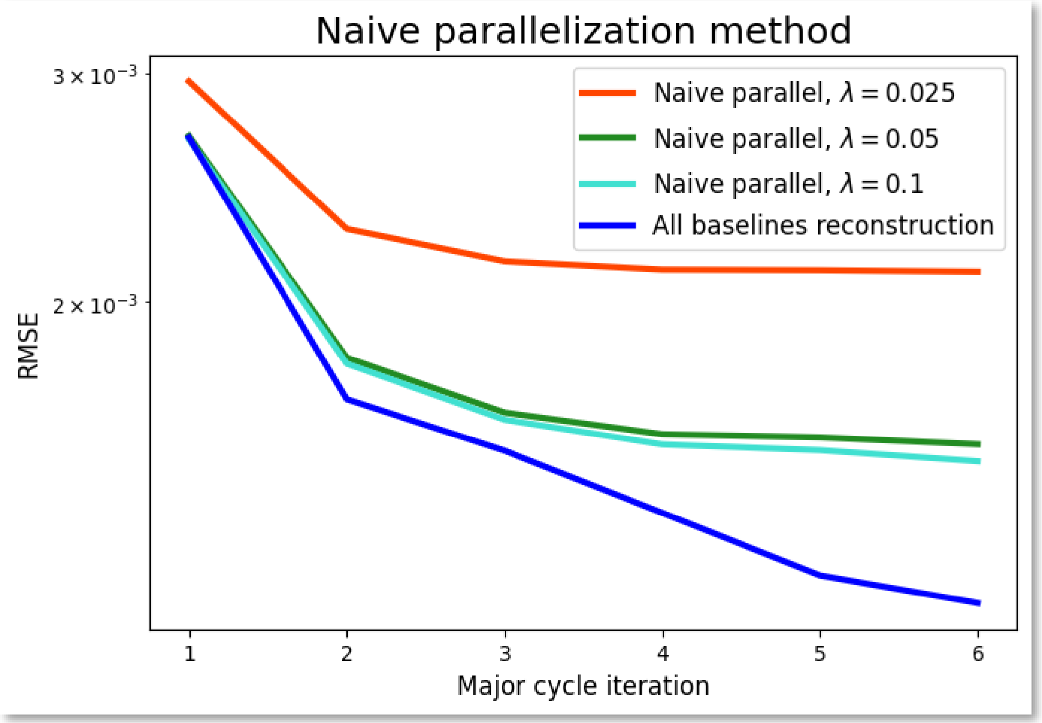


Resolutions not regularized together, unclear how to combine images.

Just adding separately deconvolved images





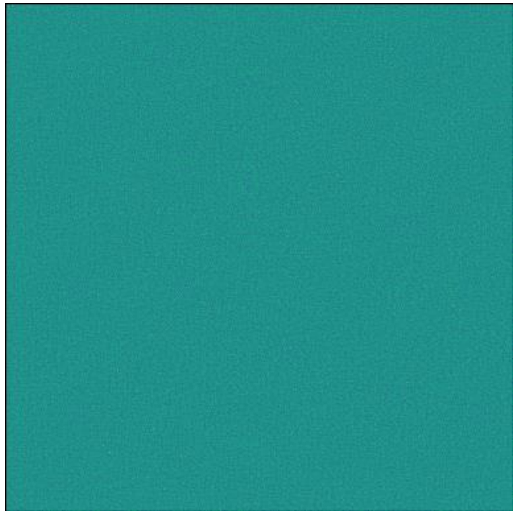


- Initial images tapered and cutout from 1.28GHz mosaic produced in [1]
- Visibilities generated with Meerkat configuration
- Exposure time of 4h, samples every 120s for to generate visibility positions
- Degrid to get visibility values
- Visibility noise artificially added

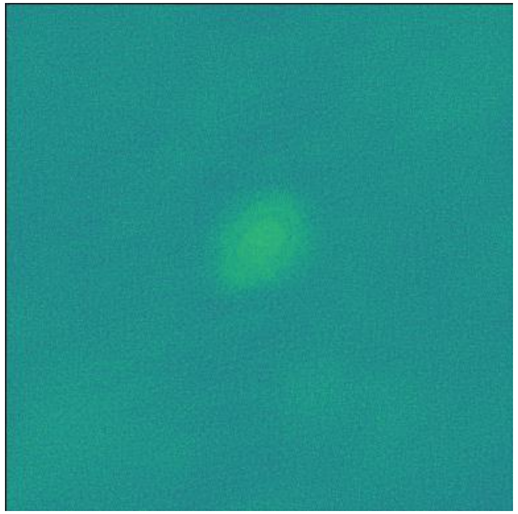
Naïve parallel reconstructions seem always worse. Maybe can regularize and combine to achieve similar quality but unclear how.

Evaluating reconstructions on real datasets

Noise reference image
(via jackknife)



Residual



Wasserstein Distance in 5x5 sliding window