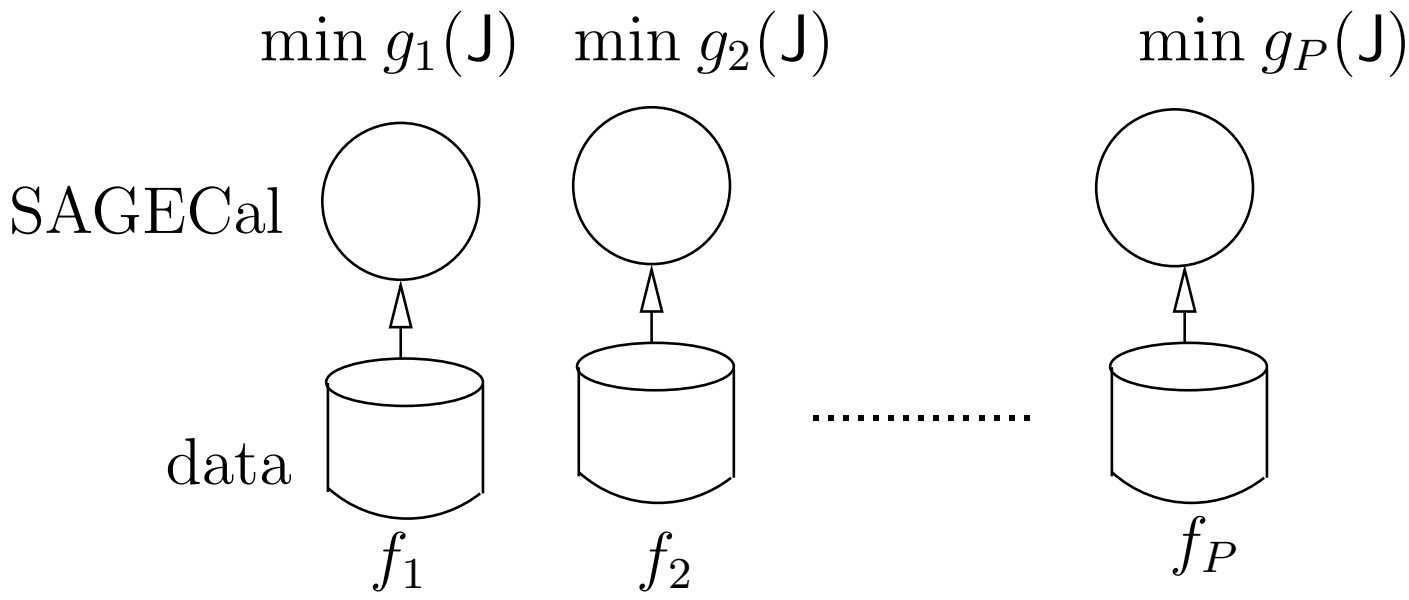
Distributed SAGECal

Sarod Yatawatta

ASTRON

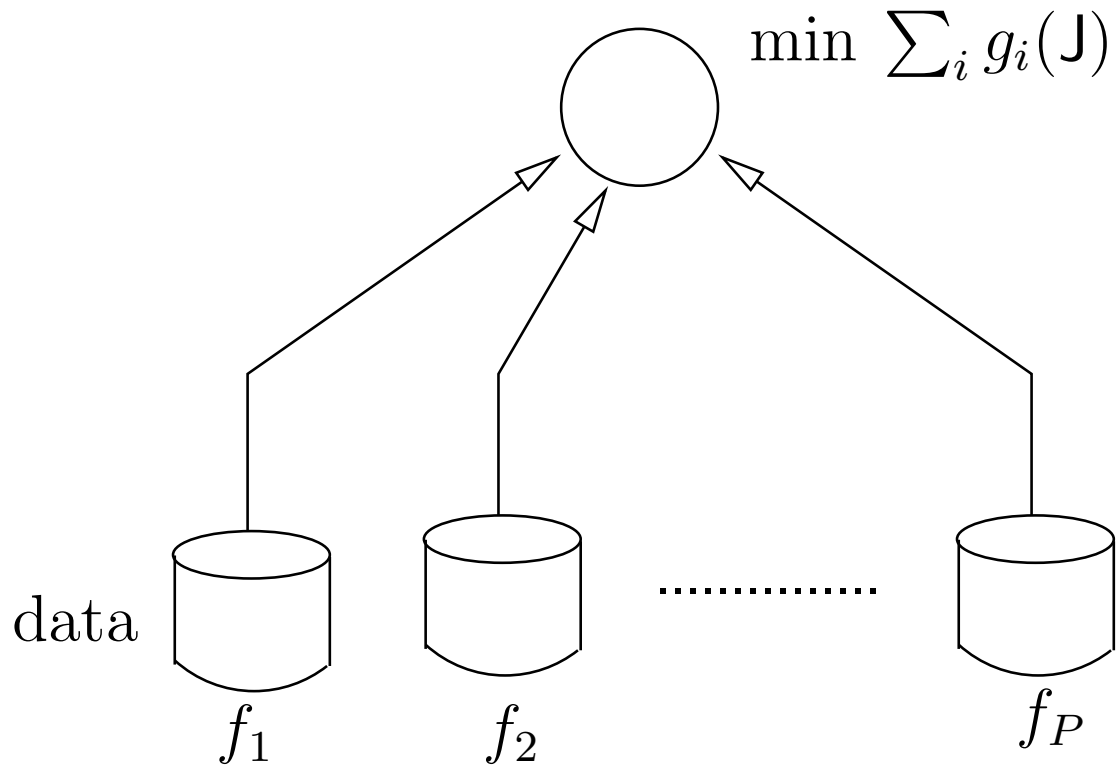
The Netherlands

Normal Calibration



Each SAGECal operates independently on data at different frequencies f_i. Solutions are only later interpolated.

What We Want



We want a unified solution exploiting smoothness in frequency.
But this does not work in practice: too much data, not enough memory, no accurate model to parametrize.

Distributed Calibration

- Normal calibration: each SAGECal works independently

$$\mathsf{J}_{f_i} = \arg\min_{\mathsf{J}} \; g_{f_i}(\mathsf{J})$$

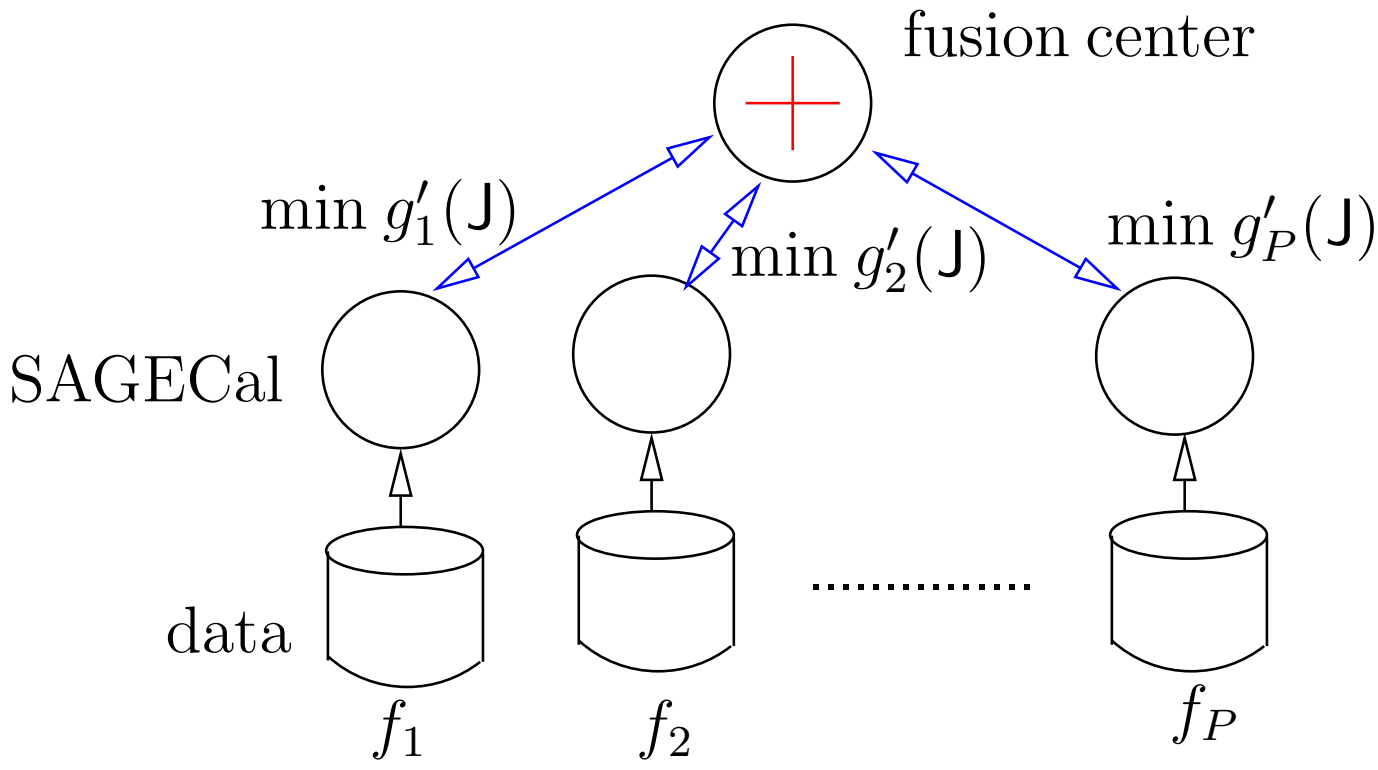
- Distributed calibration: each SAGECal appears to work independently, but actually solves

$$\{\mathsf{J}_{f_1}, \mathsf{J}_{f_2}, \ldots, \mathsf{Z}\} = \arg\min_{\mathsf{J}_{f_i},\ldots,\mathsf{Z}} \sum_i g_{f_i}(\mathsf{J}_{f_i})$$

$$\text{subject to} \;\; \mathsf{J}_{f_i} = \mathsf{B}_{f_i}\mathsf{Z}, \;\; i \in [1, P]$$

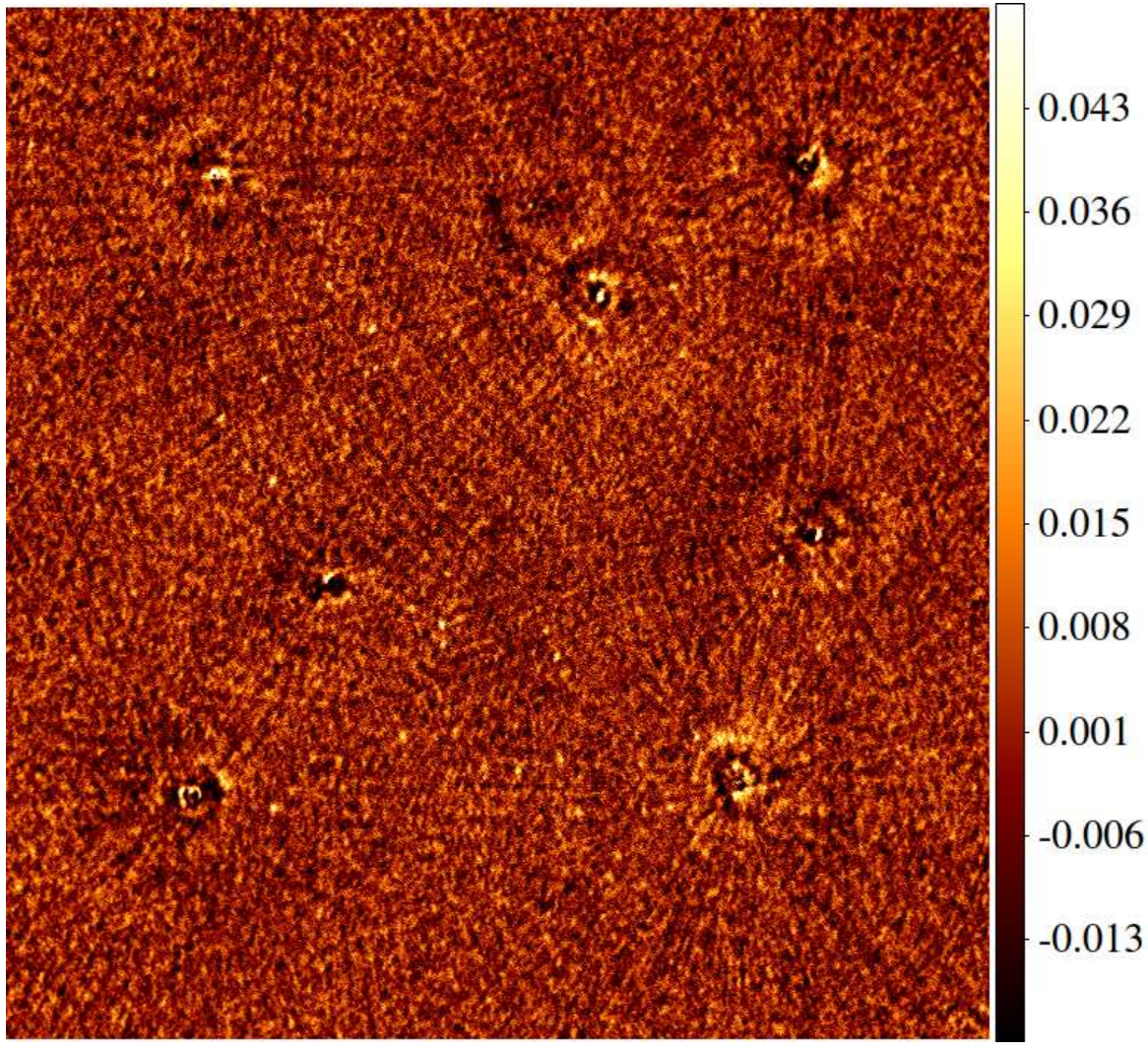
- Basic principle is consensus optimization : details [Tsitsiklis,1984], [Boyd et al.,2011], [Yatawatta,2015]

Distributed Calibration



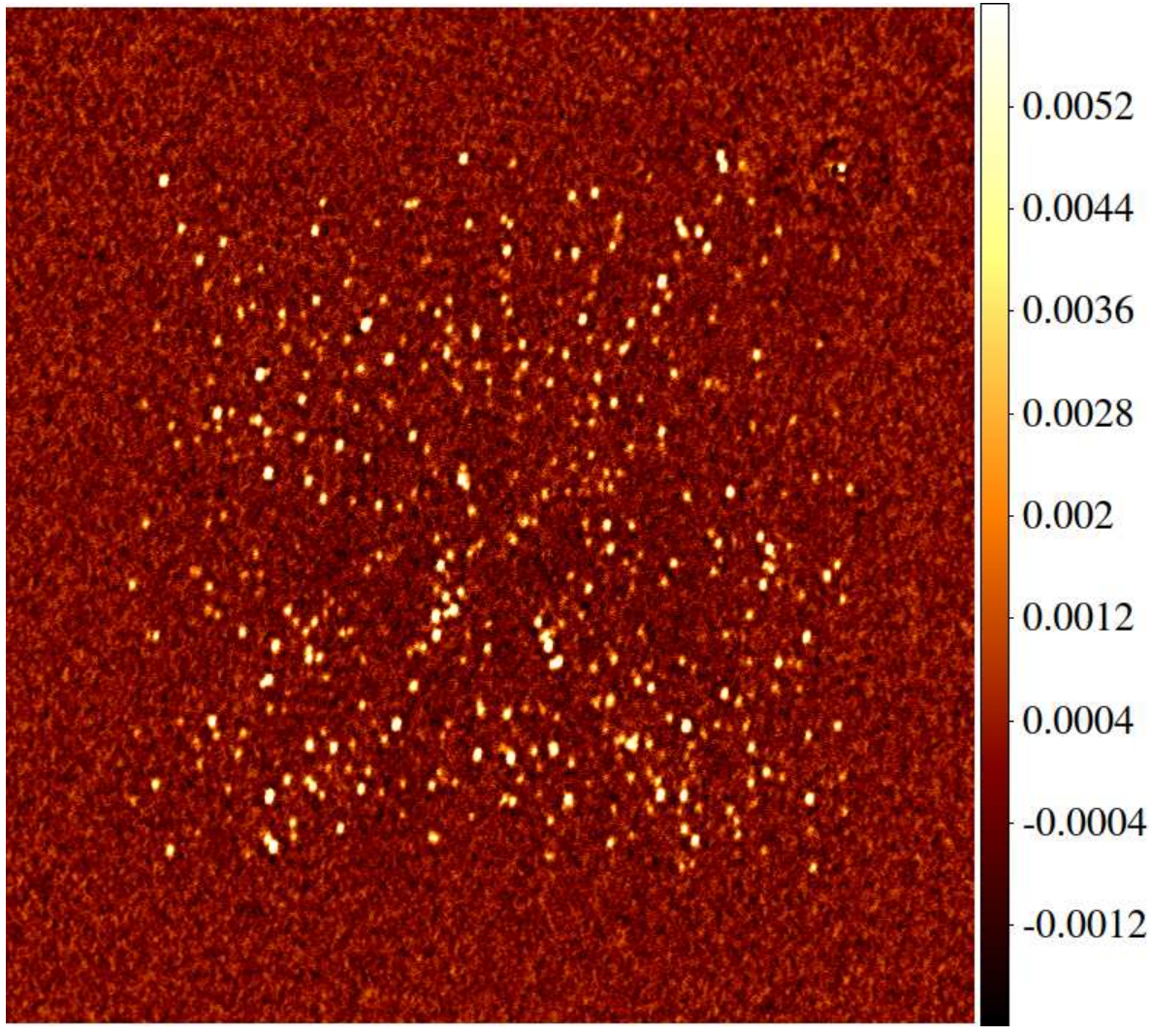
Information passed is much less than actual data calibrated. Only minor modifications to software needed, not running out of memory.

Example I



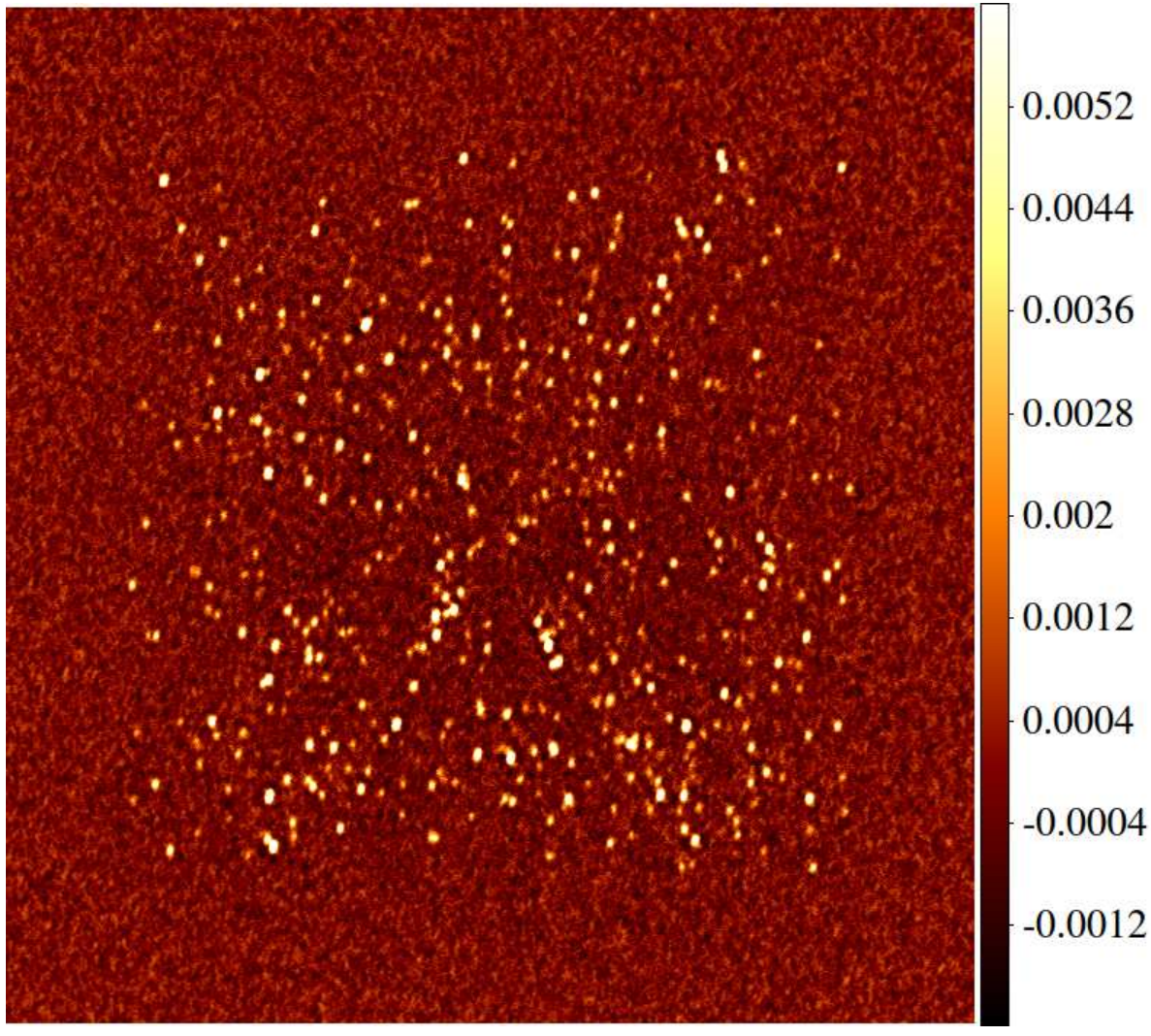
16 clusters calibrated using 1 time sample, 16 subbands 115-185 MHz

Normal Calibration



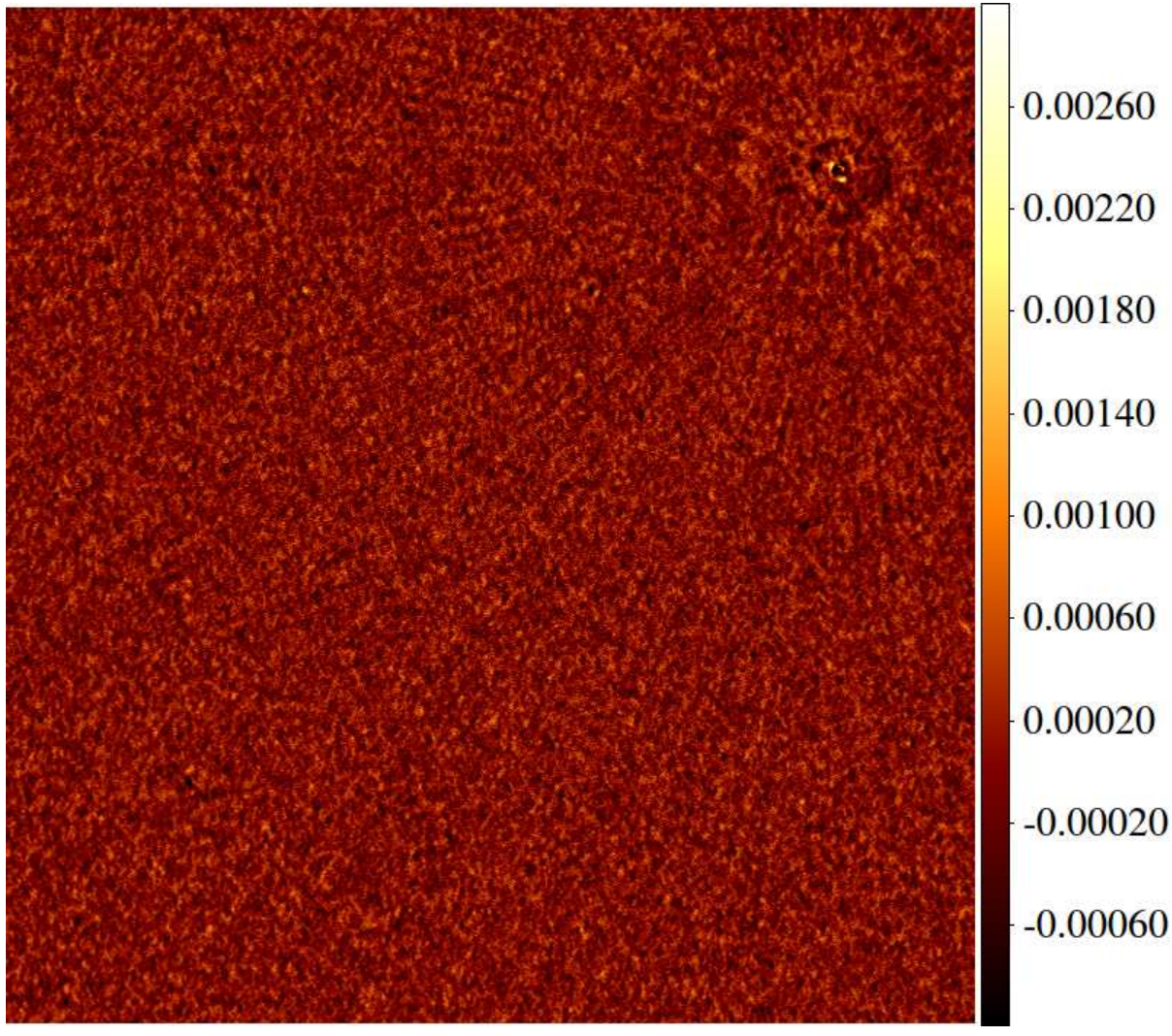
residual continuum image 16 subbands

Distributed Calibration



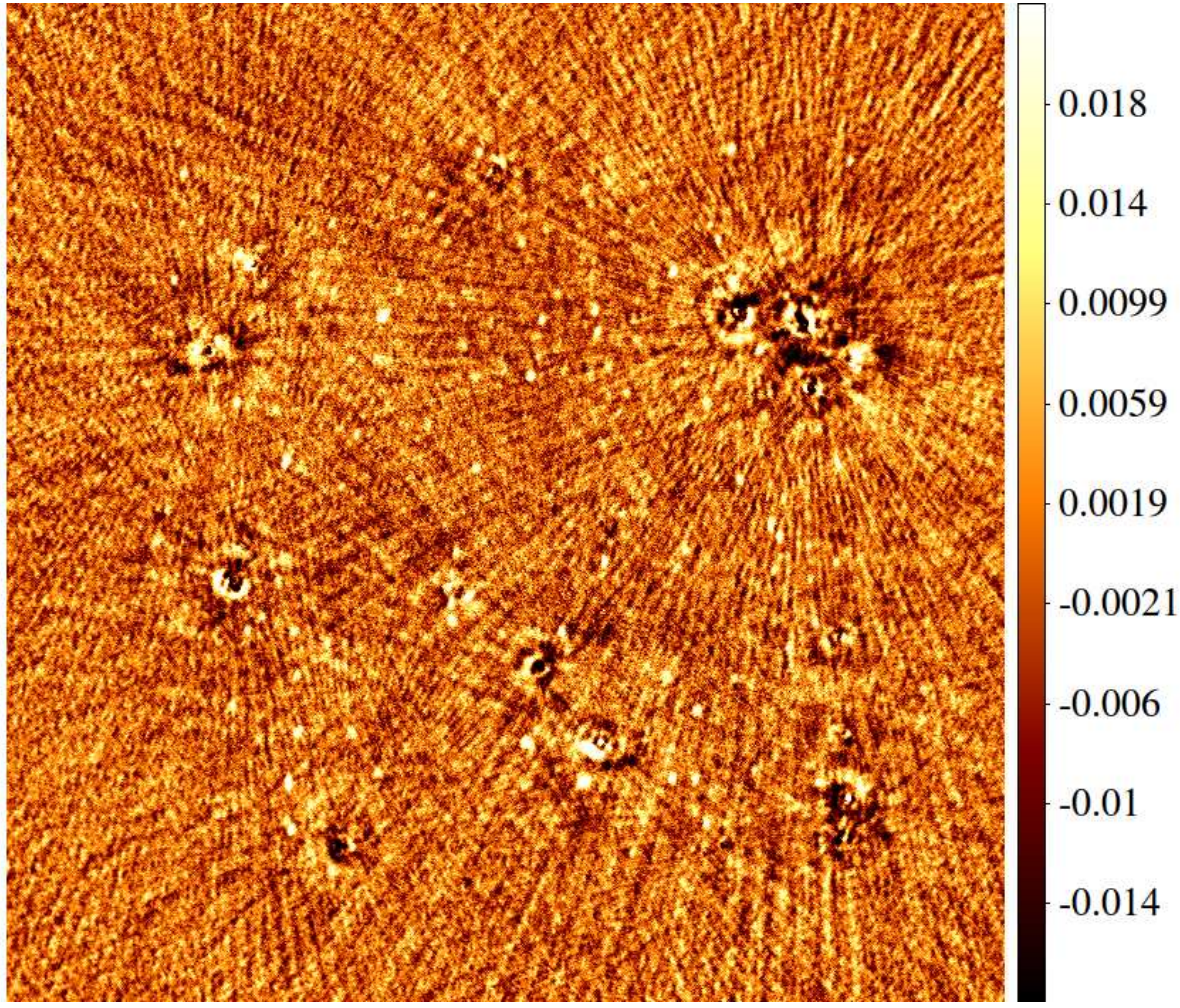
residual continuum image 16 subbands (noise lower)

Difference



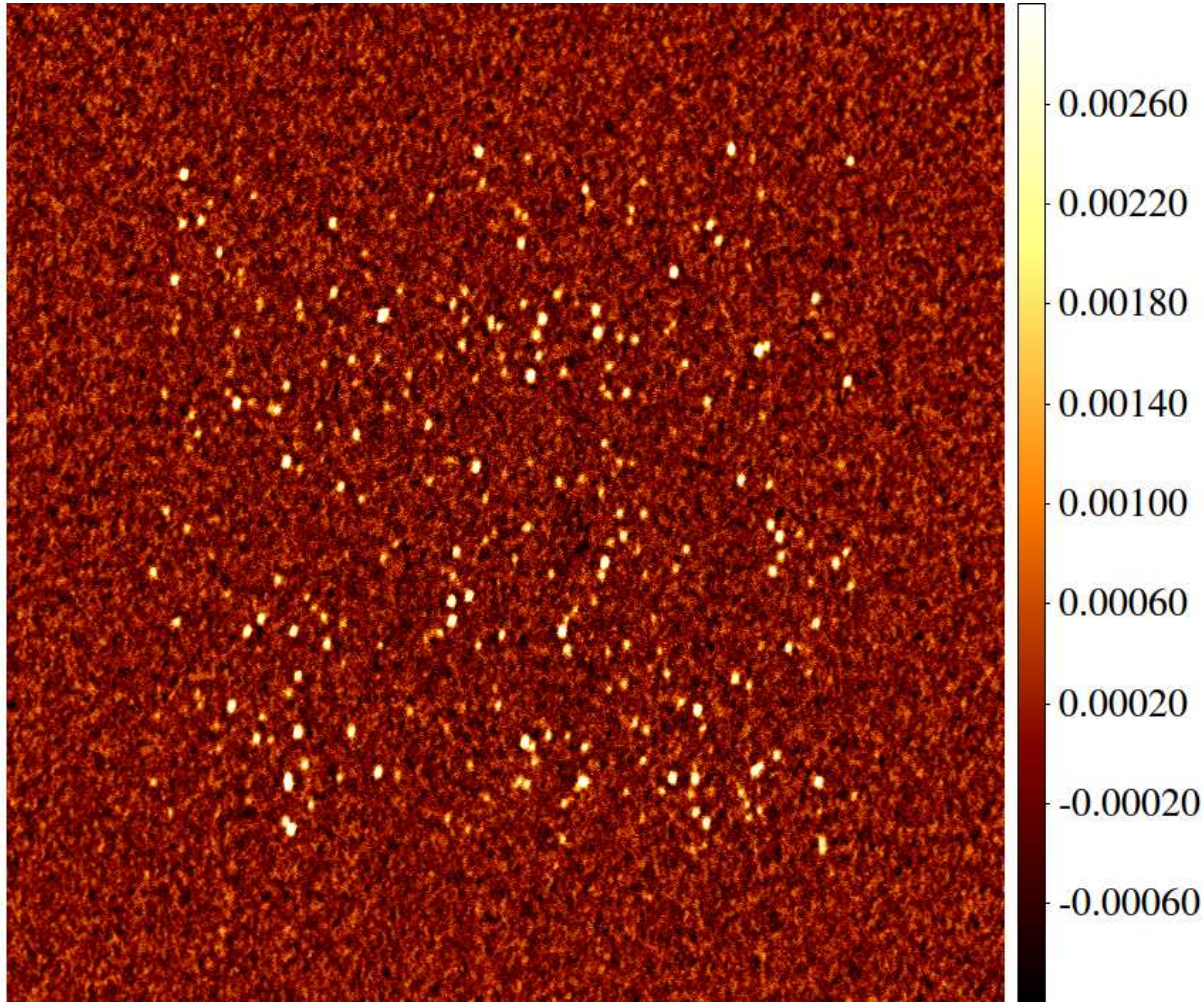
normal-distributed: calibration failed for one (weak) cluster with normal calibration

Example II



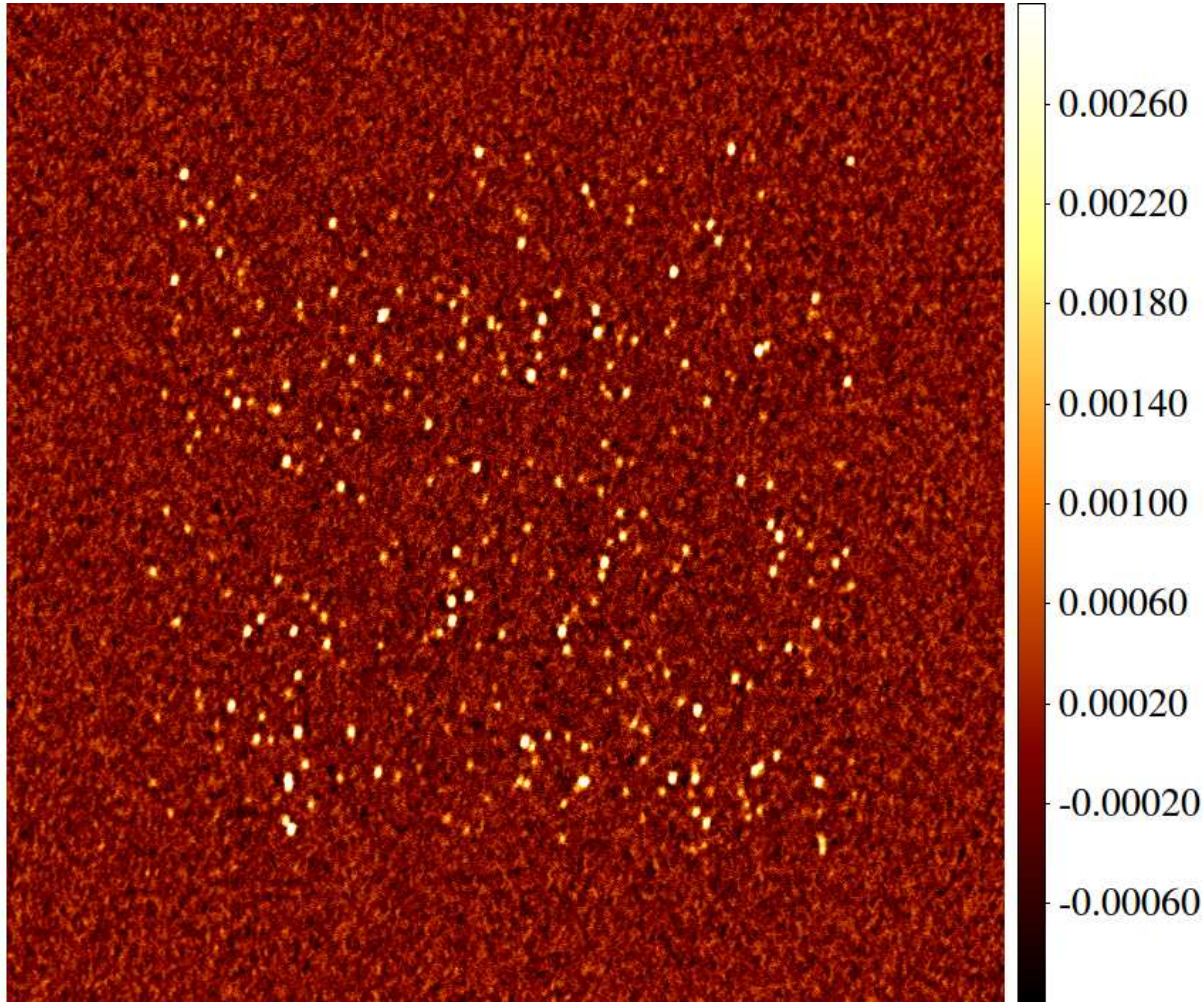
25 clusters calibrated using 1 time sample, 32 subbands 115-185 MHz,
unknowns > constraints

Normal Calibration



noise is higher

Distributed Calibration



noise is lower

Conclusions

Distributed calibration

- Not much additional computational cost.
- No need to access full dataset at each node.
- Robust to model errors.
- Calibration less likely to get stuck in local minima.
- Cons?? Theory is more complex.