CRUSH: fast and scalable data reduction for imaging arrays

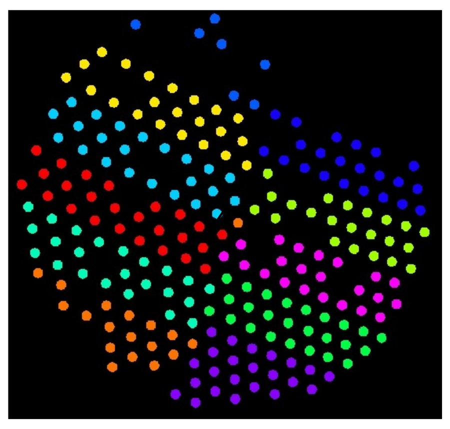
Attila Kovács (MPIfR)

Max-Planck-Institut für Radioastronomie

CRUSH is an approach to data analysis under noise interference, developed specifically for submillimeter imaging arrays. The method uses an iterated sequence of statistical estimators to separate source and noise signals. Its filtering properties are well-characterized and easily adjusted to preference. Implementations are well-suited for parallel processing and its computing requirements scale linearly with data size -- rendering it an attractive approach for reducing the data volumes from future large arrays.

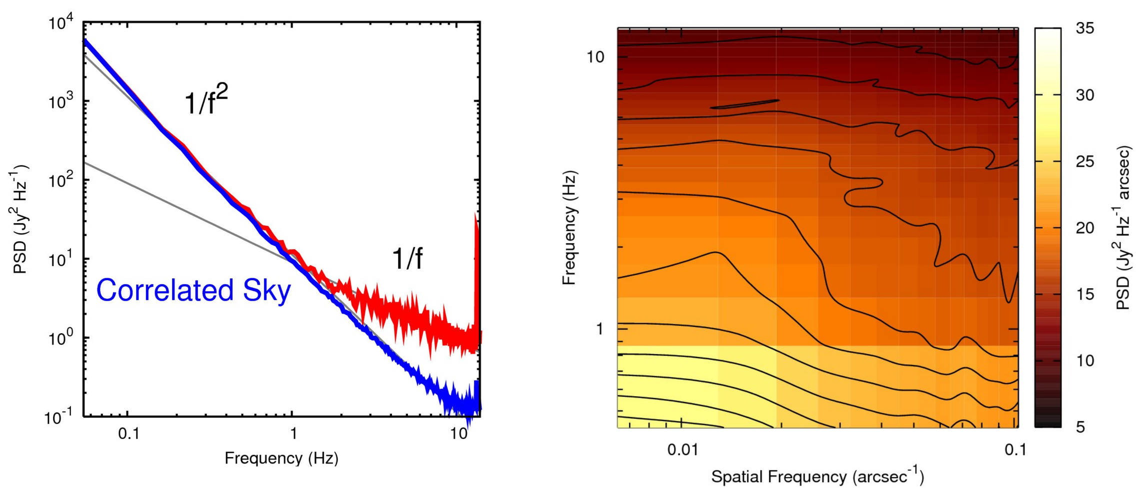
Correlate Noise

LABOCA band cables



The essence of CRUSH is in the way it deals with correlated signals. These may affect all detectors at once, such as atmospheric noise or temperature fluctuations, or smaller groups of detectors (e.g. On LABOCA's band cables, or SHARC-2 rows). The size of such correlated blocks largely determines the largest scales that can be sensitively recovered in the reduction.

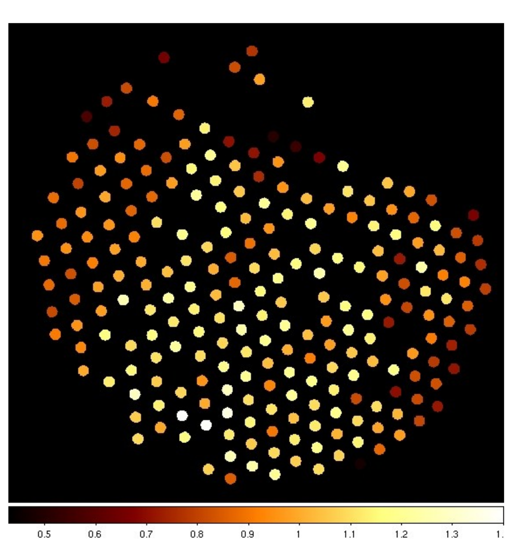
Getting Around 1/f Noise



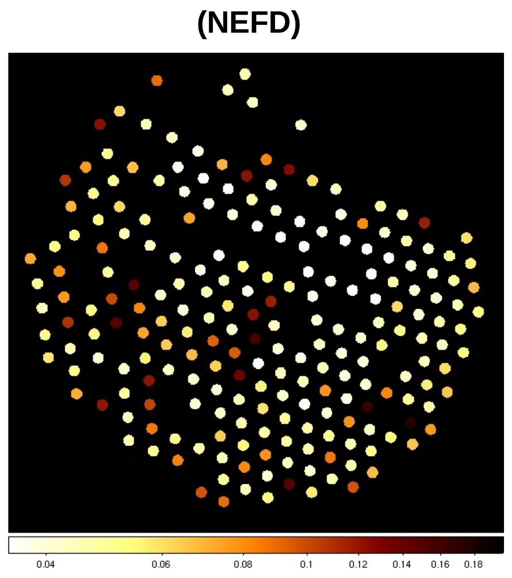
Ground-based bolometer arrays operate under a bright and highly variable atmosphere, which has a steep $1/f^2$ correlated power spectrum. CRUSH can model and remove the interfering noise, and filter uncorrelated noise components, to arrive at clean source signals for mapping.

Gains and Noise Weights

LABOCA atmospheric gains



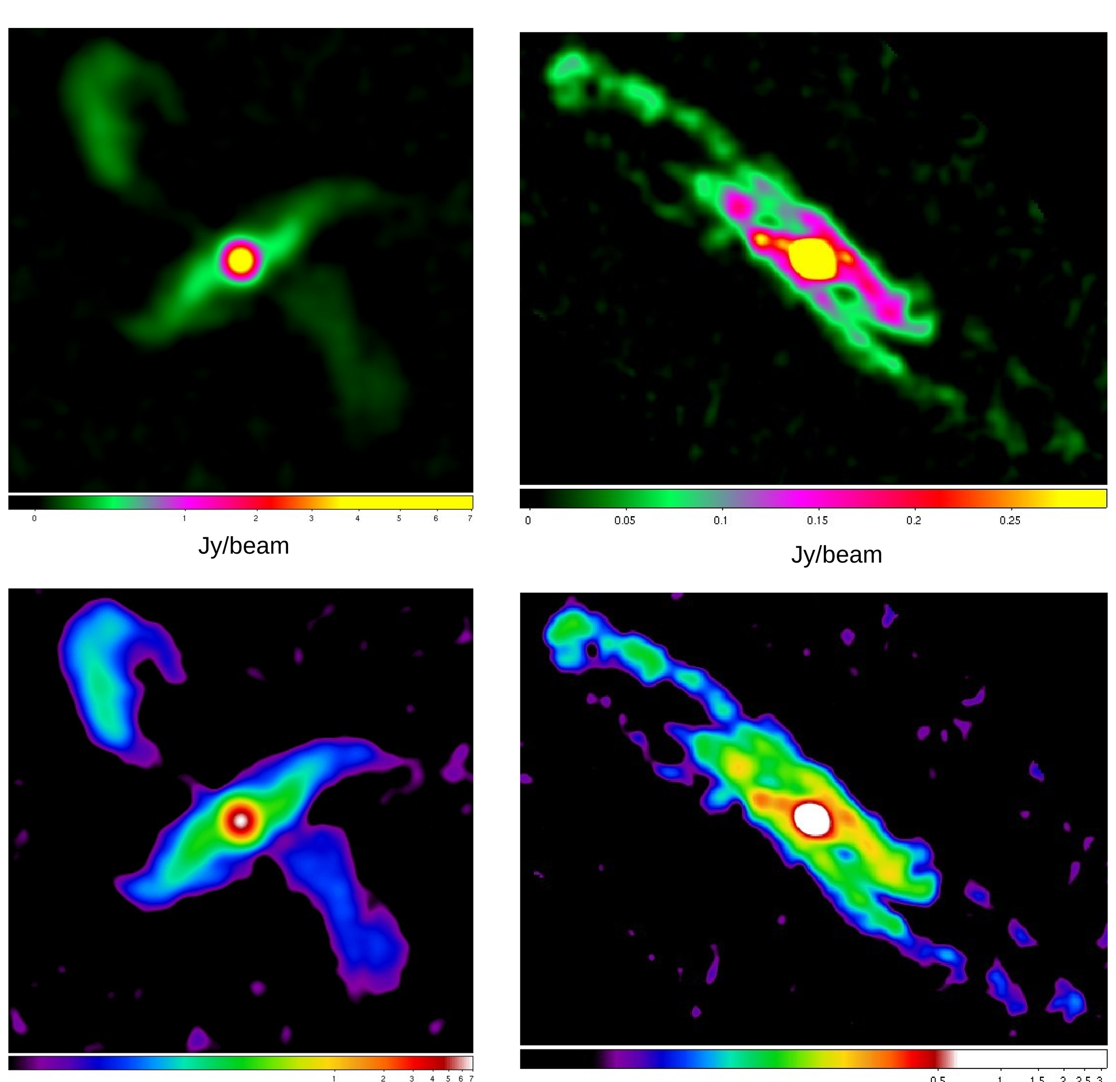
LABOCA Pixel Sensitivities (NEFD)



Jy $s^{1/2}$

The estimation of gains and noise weights forms an integral part of typical CRUSH pipelines. Precise gain knowledge is necessary for the accurate removal of the bright, correlated atmospheric noise. The noise weights derived during reduction are necessary to arrive at statistically sound estimates of signals (both source and noise).

Centaurus A **NGC 253**

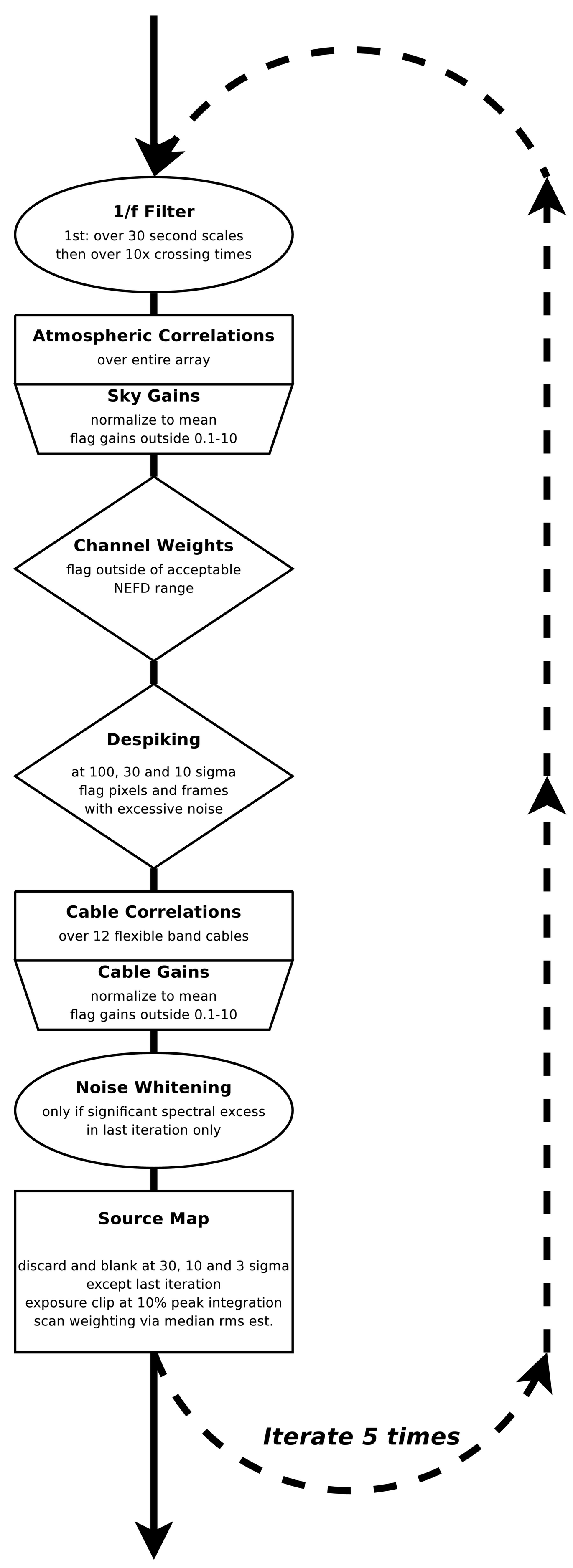


http://www.submm.caltech.edu/~sharc/crush

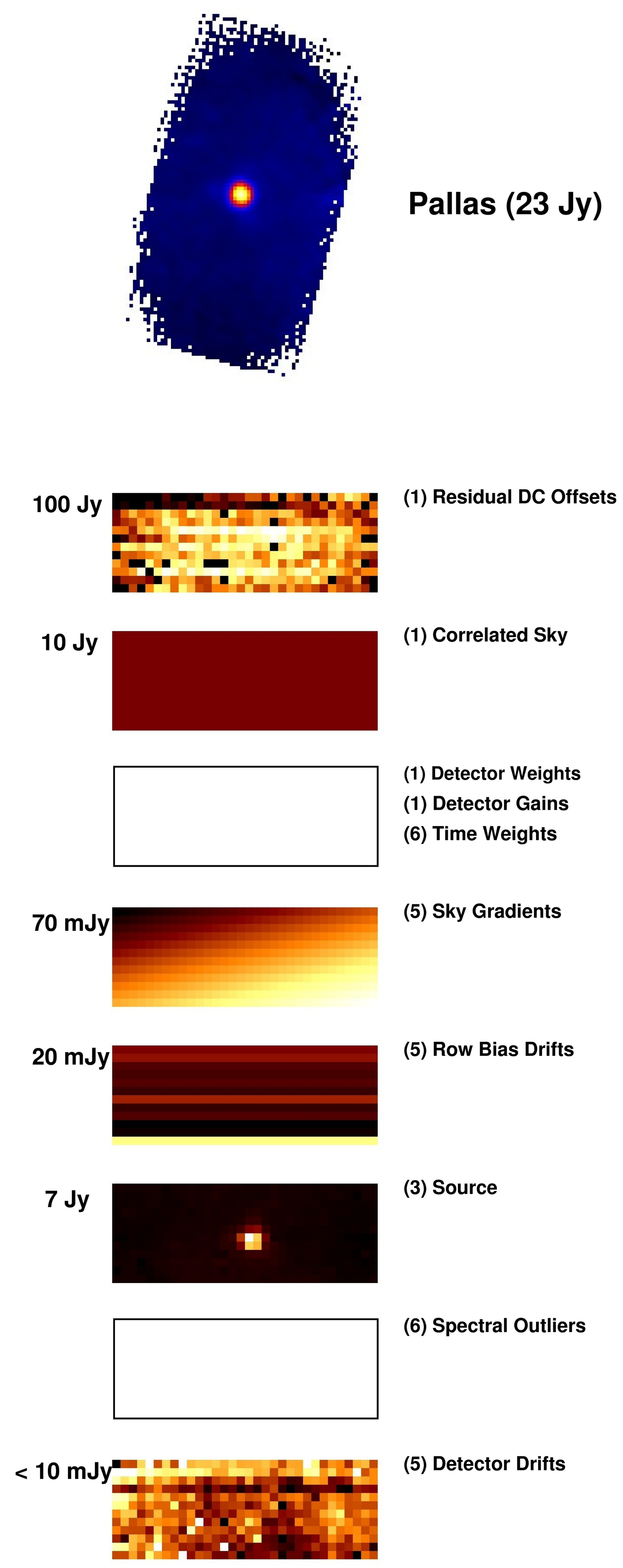
A Pipeline Approach

CRUSH uses an iterated sequence of statistical estimators (e.g. Maximum-likelihood, maximum-entropy or medians) to arrive at incremental models of the various signal components for source and noise. The ordering determines how fast convergence is achieved (typically in a handful of iterations), and the filtering properties of the reduction.

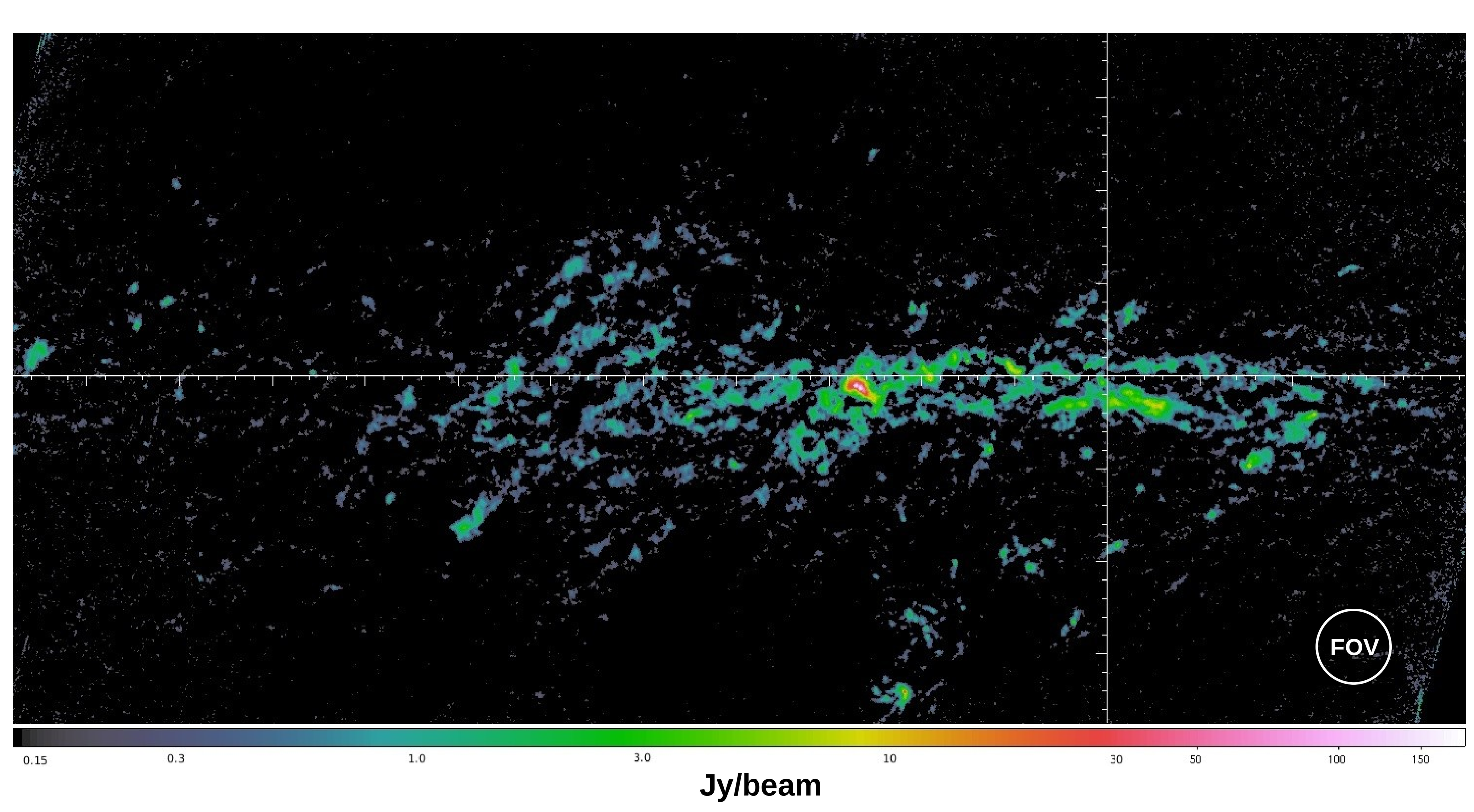
LABOCA pipeline



SHARC-2 pipeline

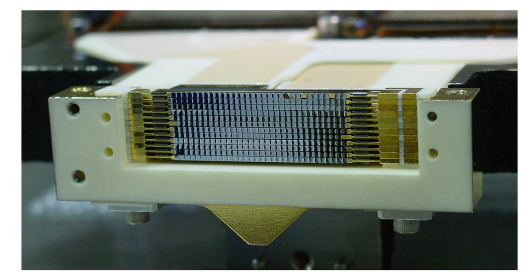


The Galactic Center Region at 870 µm with LABOCA

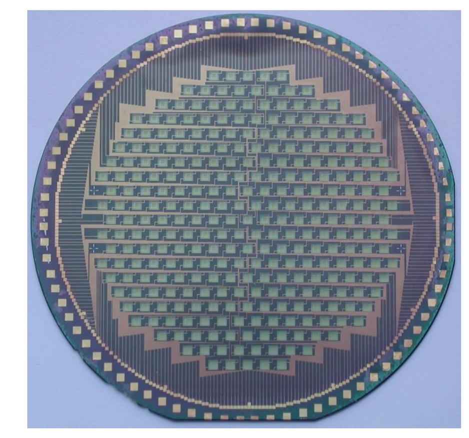


Applications

SHARC-2



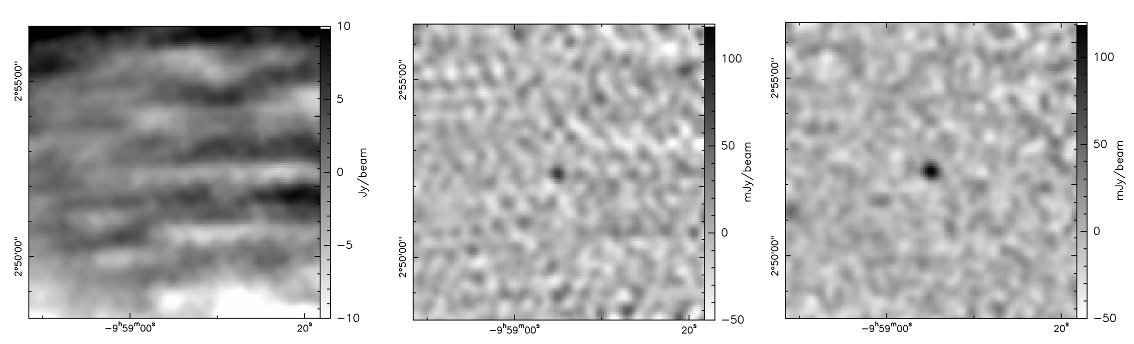
LABOCA



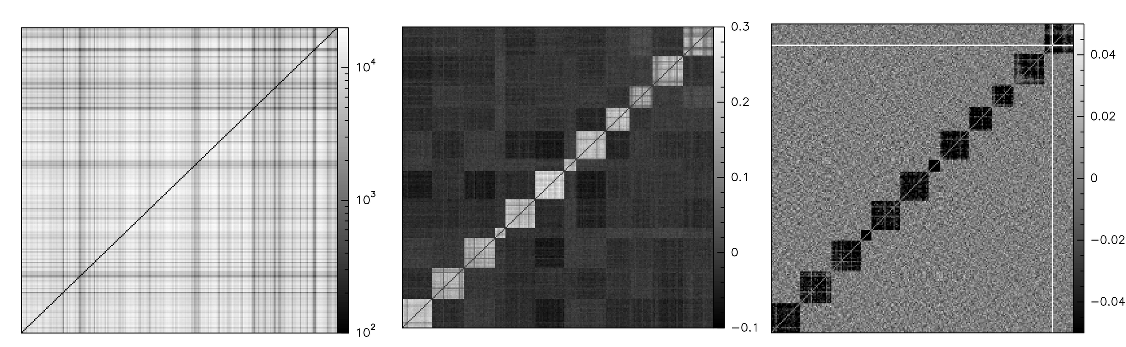
CRUSH (and its various implementations) provide the data reduction scheme for some of today's large bolometer arrays, like the 384-pixel SHARC-2 (350 µm) at the CSO, or the APEX bolometers: the 295-pixel LABOCA (870 µm), the 320-pixel ASZCA (2 mm). Its computing requirements are modest, grow linearly with data volume. CRUSH is also well-suited for distributed computing on clusters, allowing fast reductions of extremely large data sets. As such it is an ideal candidate for powering future kilo-pixel-scale instruments (e.g. SCUBA-2 or ArTeMiS), and adaptations for interferometry (ALMA) may also be possible.

Filtering of Large Scales

(esp. larger than the FOV)



The awkward choice between keeping more extended emission or paying the price of higher map noise: an example of a simulated 100 mJy point source implanted in a single 8-minute blank-field LABOCA scan and reduced three different ways. Shown are a direct map (top left), produced with signal centering only, a map with correlated sky removal (top center), and with additional band-cable decorrelation (top right) taking place before the mapping step. Below are the normalized residual pixel-to-pixel covariances after the reduction, for the 234 working channels in the array, here with the diagonal 1 values zeroed. The left map preserves source structures on all scales, but these would only be seen if are well in excess of the whopping 4 Jy/beam apparent noise level. After removal of the atmospheric noise, the image (top center) no longer contains scales >FOV (~11'), but the noise level drops over two orders of magnitude and the faint inserted source becomes visible. When the common-mode signals on the flexible band cables are also modeled prior to the map-making step, one is rewarded with a cleaner image. At this point, the covariances outside of the decorrelated cable blocks (bottom right) reveal no more correlated signals down to a few percent of the detector white noise levels. However, with the decorrelation of the cables go the scales above the typical footprint of detectors sharing a cable (i.e. >0.3–0.5 FOV).



Chandra Deep Field South at 870 µm with LABOCA