SHARC II Data Reduction Workshop
November 8, 2004

09:05 Colin Borys: Intoruduction

Welcome. Agenda.
Goals:
1. Transfer expertise from Caltech to community
2. Transfer expertise fro community to Caltech
3. Learn about difficulties with hardware/software
4. Improve data acquisition techniques

09:10 Darren Dowell: Brief update

1/3 of array died. Fixing it now, Will be done soon. Nasymth move worked.

09:10 Colin Borys: Caltech success stories

SLUGS sources: SHARC @ CSO good at 350μm. Sources at z<0.05; are bright enough to do as a backup project. Takes 0.5h in tau~0.06, 1.0 in ta=0.08; sources are between 0.5 and 3 Jy.
High redshift Spitzer HLIRGs being observed (Soifer's Bootes field). First one if 226±45mJy.
Min Yang looking at Stanford 0.1<z<1 galaxies; seen 20 of them so far. Typical time of ~2h in best weather; <100mJy typ. Flux.
Darren Dowell has beautiful map of Fomalhaut disk; 150mJy/beam, 1.2Jy total. Observed with tau=0.03 for 3h! Publication in prep (Ken Karsh). Used 'crush – deep' reduction.
Orion: 4h observation over almost 1° strip. RMS of 300mJy to 1Jy/beam. Observed with Box scans and mosaiced. SCUBA took 18h to get similar map at 450μm. Published by Martin Houde.
Low-z interacting galaxy survey: Antennae & Arp 299. Observed 14 of 42 sources; work by Jonathan Bird & Darren Dowell.
Deep field toward Abell 370 observed in Chopped mode with slow sweep. 10h in good weather (tau=0.044). Two sources at ~80mJy detected, which were known from SCUBA.
Darek Lis is working on Galactic molecular clouds. SgrB2, MonR2, W51, S68, HH24, Box scans for 0.5h in good weather; 50Jy sources.
Attila Kovacs: high-z SCUBA sources. Detected Lockman Hole sources with 3-4h at tay-0.06 to yield rms=5Mjy. LE12, LE21, LE31, LE2 all detected.
Other things: Moon image, Saturn, Titan, E-Eridani. Titan is a poor calibrator! SgrA*, Horsehead. Nebulae.

09:25 Colin Borys: Software and Requirements

CRUSH by Attila Kovacs – supported

SHARCSolve by Darren Dowell – trying to phase out

Ancillary stuff by Colin Borys. – Some being phased out for upcoming ImageTool by Attila Kovacs.

Ancillary items:

Header_update: alters header keywords

Boxscan: calculates box_scan parameters

SHARCCal: applies cal to reduced map; can take out offsets.

SHARCGap: tests file for timing gaps.

SHARCSmooth: Performs a PSF fit to reduced SHARC2 map.

SHARCLog: gets info out of headers

SHARCStat: computes statistics on reduced maps.

SHARCTau: Uses Jon Bird's tau fits to estimate tau; requires taufit file

Presentation: IDL, GRAPHIC, GAIA, DS9.

Software installation: set up CRUSH directory; install utilities anywhere.

For convenience with CRUSH, use the '-outpath=' argument to move output files; have script move into CRUSH's directory.

For maintaining log files, use ">!" to pipe file out. CRUSH can now do ranges (e.g., 10417-10426).

Publishable images: IDL AstroLib routines. Colin Borys has a program to do SHARC plotting.

GRAPHIC is available for many platforms; also produces publication-quality images.

09:40 Attila Kovacs: CRUSH reduction

Overview: how data reduction should work. How CRUSH works internally; CRUSH options. How to handle data.

Why submm is so difficult: Sky us around 100,000Jy; sky noise is 100-1000Jy; faint things are only 0.01Jy. Dynamic range is thus 10^7 – analogous to 16^{th} mag star in the day!

Atmospheric power variation is very 1/f, turning over at 10Hz or so!

Signals must be differenced; Lissajous sweep or chopping. Simulated chopped noise is 2-3 times higher than for Lissajous, without deconvolution noise! Chopped data is not fully understood; actual measurements seem to match to within 30% (according to Darren).

Sweep speed: want optimal interconnections to calibrate pixels against each other; calibrate as many as possible. Should not be periodic.

Reduction by Singular Value Decomposition, a mathematically rigorous maximum entropy solution.

SVD is difficult because it's computationally costly, isn't good with nonlinearities, has degeneracis/singularities; time=dependent noise

Parallel effort at GSFC produced some results.

CRUSH approach: SVD by iteration. At each stage, assume one thing dominates; make an estimate by maximum likelihood; remove it; go to next thing. It does converge, and has some advantages: it's intuitive in approach; is fast (linear with data size); can deal with nonlinearities; is easy to manipulate/change.

Example: Pixel offsets fitted, then Source model,then row drifts, then detector 1/f, then acceleration response.

Models in order of decreasing typical brightness: pixel offsets, correlated noise, gain model, pixel weights, chopping residuals, temperature graditents, row drifts, 1/f drifts, time weights, residual spikes, regional correlations, acceleration, temporal, spectral.

Preliminary simulations showed that this iterative approach does converge.

Recent simulations by Tom Tyranowksi showed that cleaning would work to recover very faint sources. However, it was difficult to clean ripples that are large compared to array size. This comes from certain parameters: temperature gradients, row drifts, and 1/f drifts; gain modeling affects them too.

Gain fitting and correlated noise are difficult. Linear and nonlinear responses. Overall gain drift is around 30% over full range.

Weighting by both pixel-only and time-only.

Anomalous behavior for pixels is important; can have strange noise, pickup. Identified in time-stream and FFT space; looking for unreasonable fits.

Deep images of Lockman Hole show it is possible to get good performance. Map residual histogram is very Gaussian up to about 10sigma (may be source flux). However, noise is 2x wider than expected from statistically independent pixel noise.

Looking at pixel-pixel correlations, the covariance shows that pixels within around a 35'' FWHM Gaussian are correlated.

10:10 Attila Kovacs: CRUSH options

Three categories: reduction options (change pipeline), scan options, and model options.

All options in "crush –help" or on web.

Principal reduction options:

Brightness-related:	-faint and –deep
Size-related:	-compact

Scan-specific options:

Average:	downsamples data
Chopped:	if you chopped
Scale:	multiplies scans by some factor (e.g. "=scale=1.23 13852 13853)
Pointing	FAZO/FZAO adjustments
Tau:	using network server is optimal; can be overridden (e.g., "tau=225GHz:0.046 15224")

Generic options:

Activation iteration	turns on model at some iteration
Time constant	what time period applies to many models. Large is robust, while small is aggressive.

Source model options:

Fidelity	when the fainter-than-source models are allowed
Extended	For large scale structures, it tried not to remove models that could look like extended structure

Exposure clipping edges
Convolve Makes reasonable pixel size.

Gain fitting and adjustment:
GainRounds: number of iterations
GainGoal: convergence criterion. For faint sources, needs better convergence.
Opacity adjustment: real-time tau fitting.

Weighting/flagging:
Weight Time: should be smaller than beam crossing time – don't want real structure!
Degrees of Freedom: if not enough DOF, pixel will be bad.

Output options:
Outpath, name: file locations
Precess: moves epoch
Resolution: map pixellization. Default is 1.62''.

Tips:
Reduce all data together
Reduce data at different scan directions
If reduction is weird:
Use different brightness
Change gain convergence
Use –compact
Adjust timescales
Try to identify problematic scans
Trade extended flux for flatter baselines
Check pointing
Use Mai-Tau

Observing tips:
Use proper scanning pattern
Point often!
Calibrate often!
Use DSOS
Keep logs!

10:30 Attila Kovacs: Data Structure

What formulae you want to use.
Using FV to look at FITS files will help. Investigate scans to see contents:
Primary image
Coverage
RMS
Signal/Noise
Single scan data; histograms etc.

Primary image: flux value I at each point is the "Measurement Flux" as would be seen by detector with no atmosphere. Natural units are nV. Calibrate by response to known source to produce pseudo-flux units such as Jy/beam.

Aperture fluxes:

Flux = Apix/(4.85”*4.77”)*sum(pixels) [Jy/I]

Flux=1/9 *sum(pixels) [Jy/I] with default pixellization

Peak fluxes: good is faint

Flux=Abolo/Abeam*I [Jy/I]

=4.85*4.77/2π/s^2 s=8.5’’*FWHM/2.35

Smoothing with beam:

Typical smoothing is Gaussian with beam.

RMS images:

Measurement uncertainty in map. For default pixels & Gaussian, just simple equation: sigma in map is 1.83*FWHM*sigma

Flux uncertainty in an aperture is 1/9 * sum of sigma^2.

Excess noise: factor is (1+C) higher than expected. Is sum of correlated pixel noise, and C~1 typically.

CRUSH suite of utilities:

ImageTool:	image view/manipulation
Show:	image view portion of ImageTool
CoAdd:	adds maps
Jiggle:	Allows you to shift-on-the-fly
CoVarSee:	Visualization for covariance matrices
Histogram:	useful for determining excess noise
DeConvolve:	superresolution

(Break)

11:15 Jon Bird: Calibration

Tau measured by tippers at 225GHz and 350μm every 10 minutes. There’s a lot of noise in the measurements. Least-squares fitting smooths out the tau measurements; available online.

Fit is in UT fraction of day; typically between 2h and 20h UT, and covers most observations.

Fits are NOT always reliable: e.g., 2004-09-25 shows difference between 350μm and 225GHz. Need to choose portion of UT day when you believe tau fits

Mai-Tau server parses fit table & produces fitted tau value. Uses 350μm by default.

Calibration: “known” fluxes from HORIZONS database; want to convert CRUSH nV output into Jy.

Aperture is FWHM+20’’ radius. Aperture size not important, but should be the same for all cases (same for all calibrators & science target).

Known sources: Mars, Uranus, Neptune, Callisto, Ceres, Vesta, Pallas, Ganymede.

In 2004 Aug/Sep run, calibration came to nearly the same value (21.3% for all sources over time).

Tau fitting is good and getting better – gets to 20% absolute!

11:35 Colin Borys: Calibrators

Availability of sources is challenging. Blazars can't be used. Solar system objects are only good ones, but are variable in a predictable way.
Primary calibrators are bootstrapped the secondary calibrators.
List of many calibrators is available on the web.
There's a UIP catalog for use when observing with SHARC2: please use these!
PSF calibration: PSF photometry is most commonly used for point source extraction. Equivalent to "convolving with the beam". SHARCSmooth does this. Answers the question: "what is the best fit amplitude of a Gaussian at the center pixel?"
Aperture photometry: different from Attila. Use a radius for source, and an annulus for sky estimation. Important caveat: when convolving, RMS will be lower than it's supposed to be!
What do you use: point sources are best with aperture photometry; extended sources use PSF photometry. For deep integrations, correlated sky signal is still in the map!
How-to of calibrating data: get good estimates of atmosphere, a decent collection of calibrators, and a choice of flux extraction.
Example: ULIRG reduction
MaiTau shows good tau fit to the time of observation
Use CRUSH with faint and compact with no convolving
Then SHARCSmooth to produce a good smoothing.
Use a calibrator in the same way.
Use GAIA to find Signal and RMS maps. Brightest pixel is 1.80 ± 0.02 Jy.
Calibration uncertainty is probably 10%.

11:50 Darren Dowell: Tweaks

Pointing Correction: -FAZO and –FZAO can adjust prior to reduction
Header_Update utility can be used afterwards, e.g. before CoAdding.
Pointing residuals (T-Terms) are at about 2.1" FAZO. Can be adjusted per night.
Hiroko has been studying temperature variations: a whole beam for different FAZO vs. ZA.
Darren's crop tweak uses pushd and imagetool and popd. Use –out option to avoid overwriting file.
Mosaics… leave for later.

(Lunch)

1:10 Colin Borys: user questions

Velu: working with resolution for Spitzer. Has done Orion, Fomalhaut. Deconvolution is good! Tested Lucy algorithm with modification (by IPAC for IRAS). Takes advantage of oversampling, using the final output of CRUSH and the PSF to deconvolve. Runs as a standalone.
Ken: Can it be done as a part of CRUSH? Need to know the PSF well.
Attila: PSF is small, but variable. Depends on weather, elevation.
Dominic: We get 8.6±0.45'' for a set of calibrations.
Ken: Fomalhaut PSF is estimated from CRUSH…

Attila: Smoothing is only 4'' by default, more in deep mode. Probably want to turn off convolution.

Jingwen: CRUSH reduction of a galaxy (RXJ094144). SHARCSolve finds the source, CRUSH does not!
Attila: Apologies. V1.34 has a lot of time constant options tweaked for certain maps, and turned out to be too long for certain other projects. For Lockman Hole, reducing the time constants – especially drift to ~100 – helped. When you don't see a source you know is there, there must be a model absorbing it, such as 2D gradients, rows, and drifts. Haven't done much chopping data, so CRUSH doesn't fully replace SHARCSolve.

Alexandre Beelen: We integrated for 7h on a faint object, but it can no longer be reduced by CRUSH without out-of-memory problems. It's 2GB of data. With early versions, there was a big field curvature. In 1.3b2, it was flatter but fainter?

Eric Murphy: Variety of map sizes made; can CRUSH reduce without edge effects?
Attila: Yes, but there was a problem in CRUSH that forgot to downweight edges properly.
Colin: Lots of SINGS and similar sources.
Attila: Always a tradeoff: flat maps or large-scale variations?
Darek Lis: Is that intrinsic to data acqiusition, or something else?
Attila: Limiting factor is speed of sky variation. How fast can you sweep in 1/f time of sky, which is 100ms.

Dominic: We have gotten good images on extended sources, with some negatives.
Attile: That's because of poor gain fitting.
Dominic: Weird wiggles (worms). Where from?
Colin: Try one pass with convolve=-1. This traces out the one really bad pixel. Want to find & remove.
Eric: Can you mark up a file with bad pixels?
Colin: Yes; CRUSH can write out pixel gains file & use iteratively.
Attila: "-gains" option.
Colin: we need to make a tool for this or change in new versions.
Dominic: M51 has major baseline problems
Attila: Try compact flag for some things.
Darek: Should try the Ring source simulation with that.
Colin: Could we release the Simulator to the public?
Attila: It's really a good idea to find a single characteristically-bad scan and fix that one, then go through full reduction.
Dominic: Nonconvergence problems later in iterations.
Attila: There's a divide-by-zero problem that produces infinities. There's an option in 1.33+ with a –debug option that checks models for infinities. These are software bugs.
Colin: Things that seem to fail to converge are when sky is highly variable.

Attila: There's a time constant issue as well. Sky is not stable for an hour, so different scans may be very different.
Colin: For SLUGS, would remove one scan at a time until it works.
Attila: Reduce a basket of representative scans to see what models correlate with background noise, which is the biggest thing to fit. It is sky variability that really screws things up.
Dominic: Johannes' question about reducing SgrA* with highly varying taus. SHARCSolve doesn't handle well, and CRUSH removes extended structures.
Attila: Try using a version after 1.3 – extended sources are more preserved.
Darren: Need to understand how the data is reduced & mosaiced.

Kristen: SXDF map has several negative and positive bumps. SCUBA tells you where to look, and you don't get the sources in the right places.
Attila: Previous CRUSH versions have had trouble with this. New version better,
Colin: Try reducing individual scans, and throw away really bad ones.

(Break)

03:20 Alexander van Engelen

Working on SCUBA II algorithms, specifically for scan maps.
Atmosphere is hardest part to work with; simulations being developed using a model based on SHARC2 data.
Model is Kolmogorov spectrum fluid dynamics model. Emission is 2D screen at 800m altitude, Gaussian random screen blowing past observatory at constant wind speed. Effective speed is 5000''/sec = 15m/s.
SCUBA II raster scan at an angle to fill in array (450μm array is undersampled).
Reprojected source+atmosphere gives a source with big streaks from atmosphere.
Questions remain about assumed structure, power spectrum, motion. Is array FOV too big to assume constant atmosphere?
SHARC2 data on MS0451 field, using Lissajous scan. Data is overwhelmingly common-mode. Fit linear gains only on common-mode signal.
Model predicted a Kolmogorov $P \propto 1+(\omega_0/\omega)^{(8/3)}$ spectrum. SHARC2 data is pretty close.
Animation on uncalibrated data with modeled screen. Animated gif made by nifty C code.
After removing common-mode signal, there's a lot of structure that's mostly instrumental. However, there's still a strong correlation with nearby pixels, which appears only when sky is not opaque.
Look at derivatives in signal to find slope of fit plane across the array; positive correlation would indicate a confirmation of Kolmogorov spectrum. Unfortunately, SHARC2 drift overhwlems – fitting is too difficult.
Since data is so common-mode, a simple mean should be sufficient even for SCUBA-II. Some small-amplitude correlated structure remains.

03:35 More user concerns

Attila: histogram of points in Lockman Hole field. Scaled to sigma, it makes sense. You can trust anything about 3.5sigma (for small fields) to be real, not statistical occurrence of faux sources.
Attila: NGC891 scan with new reduction improves mosaic, but lissajous scan pattern was nonoptimal for this source.

03:50 Colin: summary of issues

Alexandre Beelen's high-z quasar issue: new version of CRUSH (1.34) runs out of memory. Also, different versions of CRUSH changes negative bowling around sources.
Answers: break data in half, reduce, coadd.
Eric's galaxy: observing mode was poor, and coadding was problematic. Maps look good with new CRUSH/coadd.
Kristen: high-z galaxy survey with false negative/positive bumps. Posited that this is from bad pixels. Colin's solution: use "-convolve=-1" to look at individual scans to find bad pixels. New version may fix.
Failure to converge / NaN being worked. One bad scan usually the culprit. Attila wants "-debug" output for bad scans.
Jingwen: chopped observation works in SHARCSolve but not CRUSH. Attila working this out. L1014 has 4 sets of scans, each of which can be reduced individually but with drifting fluxes. Tau changed during observation, so using taufits helps.
Dominic: baseline fitting?
Attila: removed from CRUSH, can be put in imagetool.
Neal Evans: Is it known where this comes from?
Attila: Combination of gain problems. Measure illumination vs. beam efficiency improperly, so residual bright sources remain. Proposed test to figure this out: scan a point source across the array at different elevations to figure out how its brightness correlates with sky response.

04:05 Darren: Prosaic Things

Lessons Learned:
Don't use extended sources for focusing (e.g., NGC2071, IRAS16293-2422, Jupiter, Venus). Use small things (Mars, Uranus, Neptune, Callisto, Ganymede, Ceres, Vesta, Pallas, Crl618, CRL2688, IRC10216, OH231.8, Arp220).
Don't point on NGC2071 (use CRL618, HLTau, OH231.8 instead)
Don't calibrate on NGC2071, blazers.
Don't use Titan since Saturn is in sidelobes.
For big fields, map full area in as short a time as possible. Mosaics tend to have stitches.
Difficult projects:
High dynamic range observations (negative artifacts and sidelobes are both at 4% level)
Faint, widely extended emission

Integrating total emission. Must use same procedure for source & calibrator, using sky aperture.

Tips:

Use DSOS; check that it's working (monitor screen); allow settling time for ZA slews; turn off at end of night.

Check source catalogs before going to CSO.

Use CAL_* sources for Extrasolar calibrators.

Check pointing/focus every 45 minutes until 8pm; then check pointing every hour and focus every 2 hours thereafter.

Report every problem, even minor ones.

New things (Colin):

DSOS is tricky to use, but very useful (especially at high elevations)

Now on Nasmyth. Practically, means that SHARC will be always nearly-on and can be started up quickly. Thus, SHARC/Bolocam will be flexibly scheduled.

Scanning a bit more complicated (Attila); array rotates on sky. Therefore, can accidentally scan along the array when close to zenith & scanning in AltAz.

Dominic: what is subscription rate like for SHARC? Large projects?

Colin: last semester was low. Looks high now. About 2/3 got time. SHADES is a big project, but is only 8 nights. Bolocam asked for 80 nights. May be possible that more observing time in the future will be available, and other groups may collaborate.

Colin: CRUSH exploder may be formed for SHARC data reduction purposes.

Colin: Thanks for coming!