VLA Memo

Sporadic Localised High-noise Channels in Target Observations

Riccardo Pavesi,[*] Chelsea Sharon,[†] Dominik Riechers[‡]

April 21, 2014

1 INTRODUCTION

The EVLA was used to observe the COSMOS extragalactic deep-field during the spring of 2013 (project VLA/13A-398; PI: Riechers). The purpose of these observations is a blind search for CO emission lines at high redshift. A seven pointing mosaic was obtained with 8 GHz of total bandwidth in the Ka-band, the radio quasar J1041+0610 was used as phase calibrator and 3C286 for bandpass and flux calibration. The correlator is set up to provide 64 spectral windows of 128 MHz bandwidth each, divided into 2 MHz wide channels, covering a frequency range of 30.988–39.036 GHz. Three frequency tunings offset in steps of 12 MHz were adopted to cover the gaps between spectral windows and obtain uninterrupted bandwidth. Our dataset is composed of 46 dynamically scheduled observing tracks between 2013, January 26 and May 14, each about 3 hours in duration. The calibration is carried out using the EVLA data reduction pipeline (v.1.2.0) developed using CASA 4.1.0. The pipeline RFI flagging, which uses CASA `rflag` to identify transient lines, was switched off as recommended by the developers because it can potentially remove narrow spectral lines and because there is little RFI in the Ka-band (with

[*]rp462@cornell.edu

[†]csharon@astro.cornell.edu

[‡]riechers@astro.cornell.edu

the exception of the 31.487–31.489 GHz range which we flag prior to running the CASA pipeline). The pipeline was further modified to only flag the first and last channel of each spectral window (instead of 3 channels) regardless of proximity to baseband edges to minimize the gap between bands. We find that the bandpass is sufficiently flat that this choice gives the best trade-off between bandwidth and introduced distortion. After executing the pipeline we visually inspect the calibrator fields to identify any necessary additional flagging, and then re-execute the final calibrations.

Since the purpose of our study is a blind search for spectral line emission, we performed a careful inspection of the noise as a function of frequency on the calibrators and source since significant variations would affect the identification of faint lines.

2 The Identification of High-noise Channels

We report a significant number of noisy spectral channels in our observations that are not removed by the calibration pipeline. The noisy channels were initially identified as narrow spikes of a small number of channels in amplitude vs. frequency plots of visibilities from the science target fields (Figs. 2.1, 2.2) and are mostly associated with single antennas (Fig. 2.3).

In the example shown in Fig. 2.4 we can see that the problematic channel has higher noise than the neighbouring channel throughout the observation. However, the difference in noise for the affected channels decreases during observations of calibration targets.

Being very narrow in frequency, the noise spikes are not significantly reduced by the statistical weights obtained from `statwt`, which minimizes the effects of all other noise features since the weights are computed per spectral window.

Including one of these noisy channels for the affected antenna during the imaging of a single pointing of the mosaic from a single observation track increases the rms noise by $\sim 20\%$ in that frequency channel. Depending on how frequently these high-noise channels occur, they may significantly affect our ability to perform blind line searches.

3 Characterisation and Removal of the High-noise Channels

In order to better characterise the noise spikes, we compare channel-to-channel variations in the noise within each spectral window for each antenna, treating each field separately. We select the noise spikes in the statistically weighted calibrated data, but the problematic channels are already apparent in the raw data.

For our comparison we select problematic channels as those whose standard deviation of the visibility amplitudes is more than 3σ greater than the mean standard deviation in that antenna for that spectral window. Fig. 3.1 shows a histogram of occurrences, with the abscissa given by the number of observations (out of 46) where a particular channel had a noise spike in a particular antenna. We exclude the first channel for all spectral windows since it is always more noisy, and ignore noise spikes that occur only once. This

plot shows that many noisy channels in the same antenna repeat over time, and would therefore be problematic if left in the datacube.

In Fig. 3.2 we display the spectral window distribution of those same noise spikes, selected as described above. We see a concentration of noise spikes in roughly four peaks, which correlate with peaks in the weighted calibrated amplitudes as a function of frequency. We consider this to be indicative of random electronic problems that manifest as increased noise and thus are more prevalent in certain hardware components of the correlator than others. The presence of four peaks is likely associated with the underlying basebands, there appears to be one peak in each of these, even though no precise correlation has been identified.

In order to achieve low variation as a function of frequency, our recommendation is to flag out these particularly noisy channels for the antennas in which they occur, prior to imaging stage, but after calibration. Selecting channels whose standard deviation exceeds the mean standard deviation in that spectral window for that antenna by 3σ is a sufficient criterion to exclude most of the problematic noise spikes. This method is partially redundant to the algorithms in `rflag` (which we did not execute as part of the pipeline), but reduces the risk of removing real spectral lines since the noisy channels are selected within individual antennas.

4 APPENDIX

4.1 SCRIPT FOR THE IDENTIFICATION OF THE HIGH NOISE CHANNELS

```
import numpy as np
import operator
execfile('\begin/data/common/COdeep_cosmos/CASA_VLApipe_v1.2.0/
    EVLA_pipe_restore.py')
vis=ms_active
tb.open(vis+'/ANTENNA')
antdict=dict(enumerate(tb.getcol("NAME").tolist()))
tb.close()
for field in range(6,10):
        f=open('deviating3sigfield'+str(field),'w')
        for j in range(0,66):
          for antenna in range(0,27):
            ampantenna=[]
            meanant=(0.0,0.0)
            sqant=(0.0,0.0)
            ms.open(vis)
            ms.selectinit(datadescid=j)
            ms.selectchannel(64,0,1,1)
            ms.select({'field_id':[field]})
            ms.select({'antenna1':[antenna]})
            data1=ms.getdata(['CORRECTED_DATA','WEIGHT'])
            flag1=ms.getdata('FLAG')
            ms.close()
```

```
ms.open(vis)
ms.selectinit(datadescid=j)
ms.selectchannel(64,0,1,1)
ms.select({'field_id':[field]})
ms.select({'antenna2':[antenna]})
data2=ms.getdata(['CORRECTED_DATA','WEIGHT'])
flag2=ms.getdata('FLAG')
ms.close()
for i in range(0,64):
  if data1!={}:
    RR1=np.array(map(operator.mul,data1['weight
        '][0,:][np.where(flag1['flag'][0,i,:]==False)
        [0]],data1['corrected_data'][0,i,:][np.where(
        flag1['flag'][0,i,:]==False)[0]] ))
    LL1=np.array(map(operator.mul,data1['weight
        '][3,:][np.where(flag1['flag'][3,i,:]==False)
        [0]],data1['corrected_data'][3,i,:][np.where(
        flag1['flag'][3,i,:]==False)[0]] ))
  else:
    RR1=np.array([])
    LL1=np.array([])
  if data2!={}:
    RR2=np.array(map(operator.mul,data2['weight
        '][0,:][np.where(flag2['flag'][0,i,:]==False)
        [0]],data2['corrected_data'][0,i,:][np.where(
        flag2['flag'][0,i,:]==False)[0]] ))
    LL2=np.array(map(operator.mul,data2['weight
        '][3,:][np.where(flag2['flag'][3,i,:]==False)
        [0]],data2['corrected_data'][3,i,:][np.where(
        flag2['flag'][3,i,:]==False)[0]] ))
  else:
    RR2=np.array([])
    LL2=np.array([])
  RR=np.concatenate((RR1,RR2))
  LL=np.concatenate((LL1,LL2))
  RRampt= np.absolute(RR).std()
  LLampt= np.absolute(LL).std()
  ampantenna.append((RRampt,LLampt))
  meanant=tuple(map(operator.add,meanant,(RRampt,
      LLampt)))
  sqant=tuple(map(operator.add,sqant,(RRampt**2,LLampt
      **2)))
meanant=tuple(map(operator.div,meanant,(64,64)))
sqant=tuple(map(operator.div,sqant,(64,64)))
stdant=tuple(map(np.sqrt,tuple(map(operator.sub,sqant,
    tuple(map(operator.mul,meanant,meanant))))))
for i in range(0,64):
            if ampantenna[i][0]>meanant[0]+3*stdant[0]
                or ampantenna[i][1]>meanant[1]+3*
```

```
                          stdant [1]:
                              f.write('field=\''+str(field)+'\'
                                 spw=\''+str(j)+':'+str(i)+'\'
                                 antenna=\''+antdict[antenna
                                 ]+'\'\n')
          f.close()
```
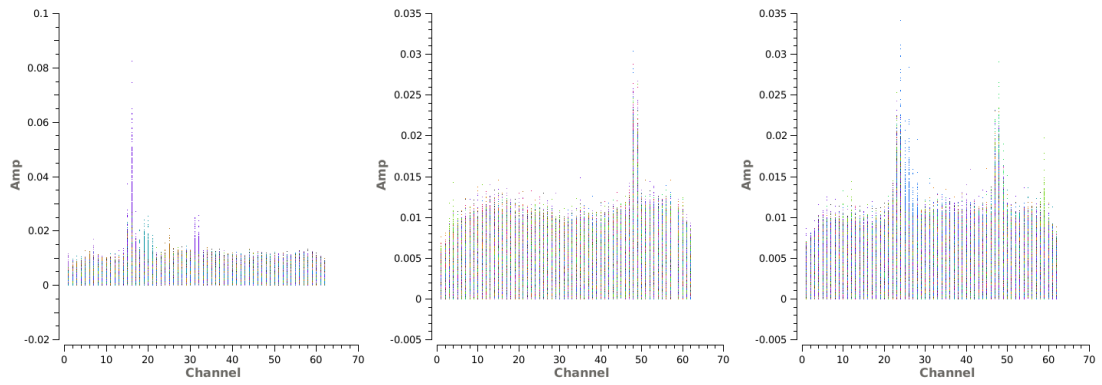
Figure 2.1: Calibrated amplitudes for one spectral window for one antenna, colorized by baseline. First: a spike that only significantly affects one channel. Second: a two-channel spike. Third: a multiple channel spike
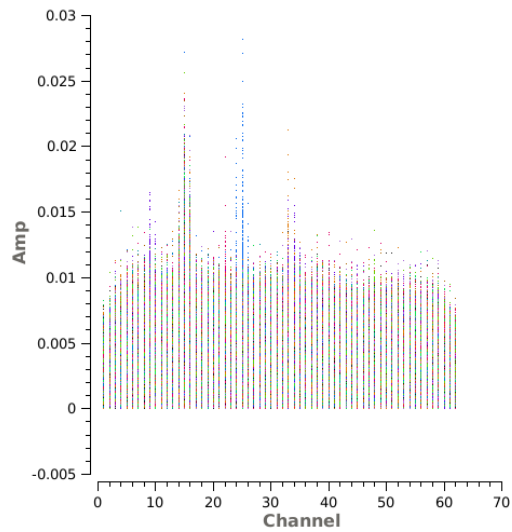
Figure 2.2: Same as Fig 2.1. Only the leftmost (channel 15) is selected as a spike by our criterion because not enough datapoints are deviant in the second one; the standard deviation of the channel is thus not high enough
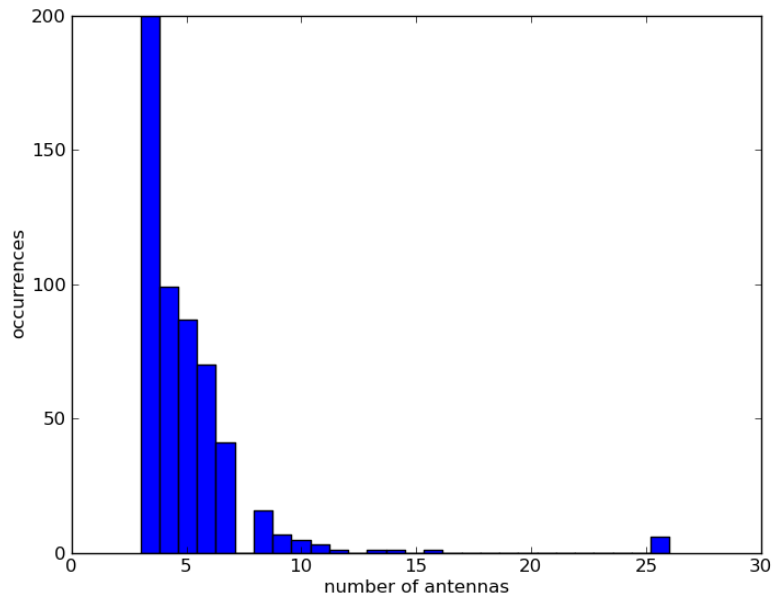
Figure 2.3: Frequency distribution for the number of antennas (x-axis) where the same
channel has high noise ($> 3\sigma$) relative to its spectral window (evaluated for
each field in each track) for all observations. The columns corresponding to a
noisy channel appearance in 1 antenna and 2 antennas only are off the scale,
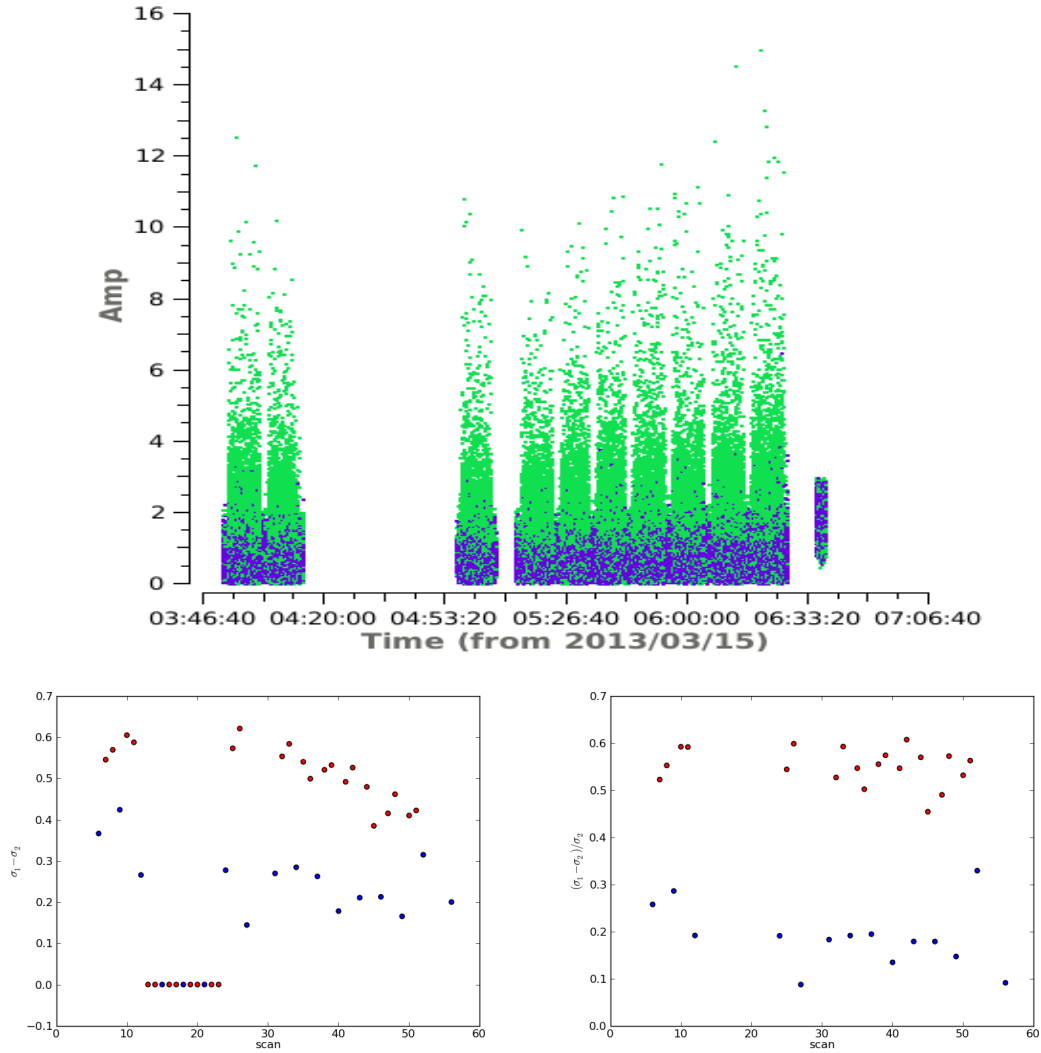the values are ~ 24000 and ~ 2200 respectively.

Figure 2.4: Calibrated amplitude versus time plot. Top: Amplitude as a function of time for two channels, one with excess noise (green) and one without (blue), for a single observing track. Bottom: $\sigma_1 - \sigma_2$ and $(\sigma_1 - \sigma_2)/\sigma_2$ for the two channels plotted above, calculated in each scan, where scans of the calibrator field are blue and scans of science fields are red.
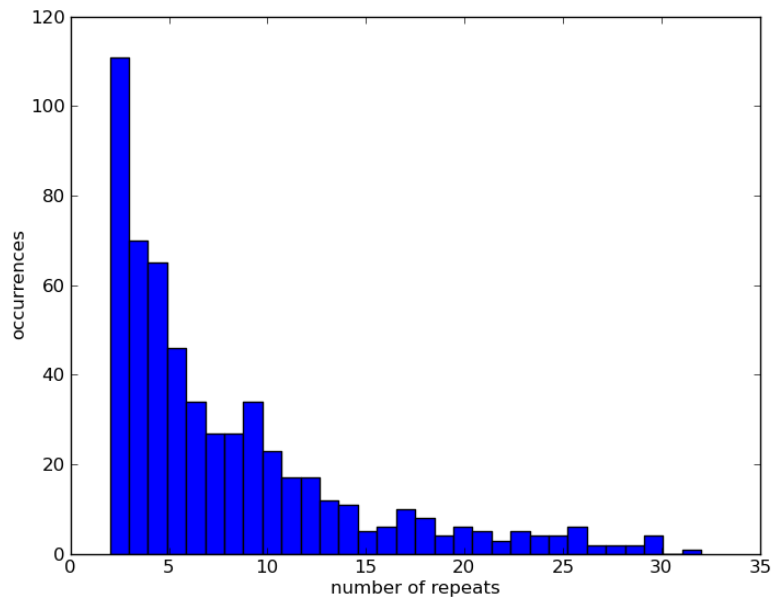
Figure 3.1: Frequency distribution for the number of observations (out of 46 possible; x-axis) where the same antenna-channel pair has high noise ($> 3\sigma$) relative to its spectral window
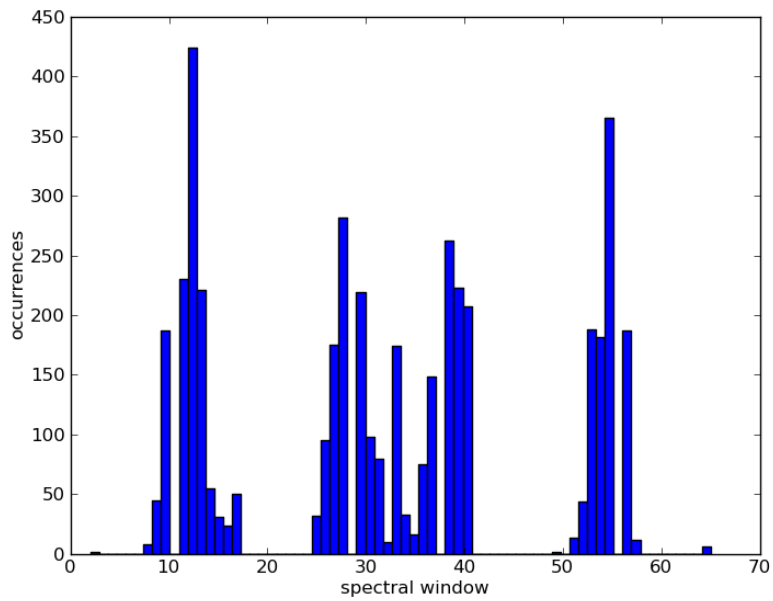

Figure 3.2: Spectral window distribution of the antenna-channel noise spikes.