2013 年11月 14日

FAST望远镜科学数据处理系统介绍

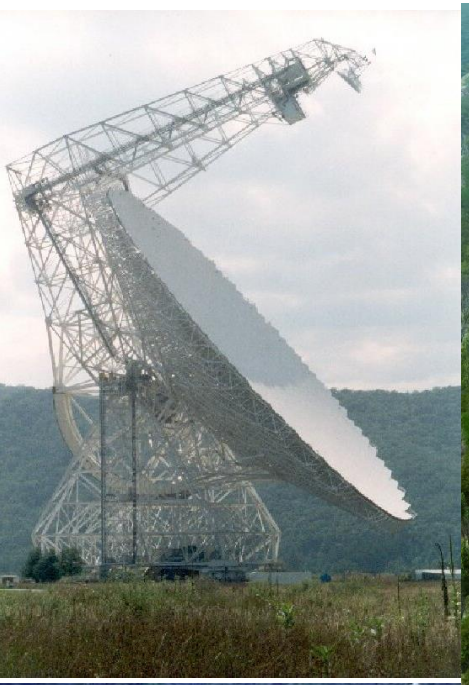
Five hundred meter Aperture Spherical radio Telescope

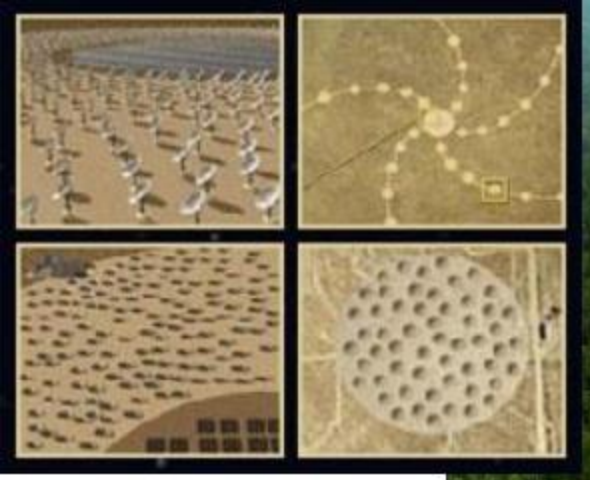
朱明
国家天文台

内容

- FAST望远镜介绍
- 射电数据处理的流程
- FAST数据储存需求，数据库

FAST望远镜模型

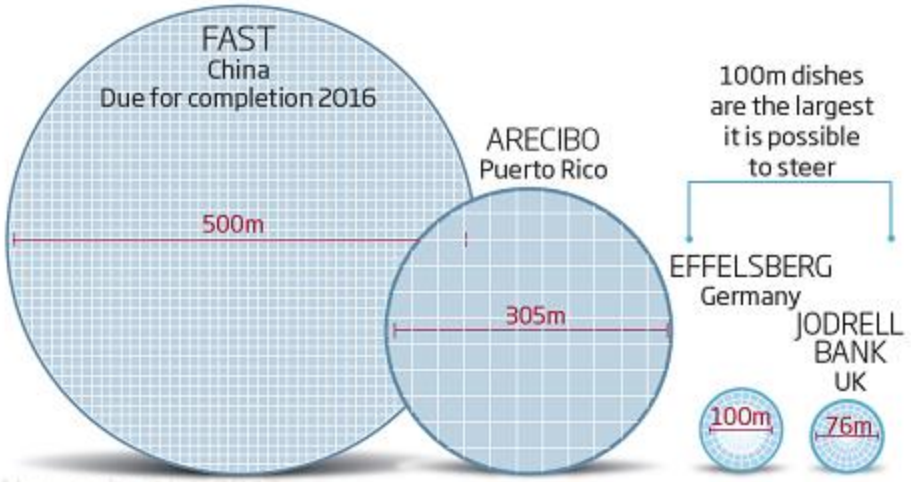




FAST的优势

- 高灵敏度
- 高巡天速度（多波束）

→作出新发现



New scientist 2011

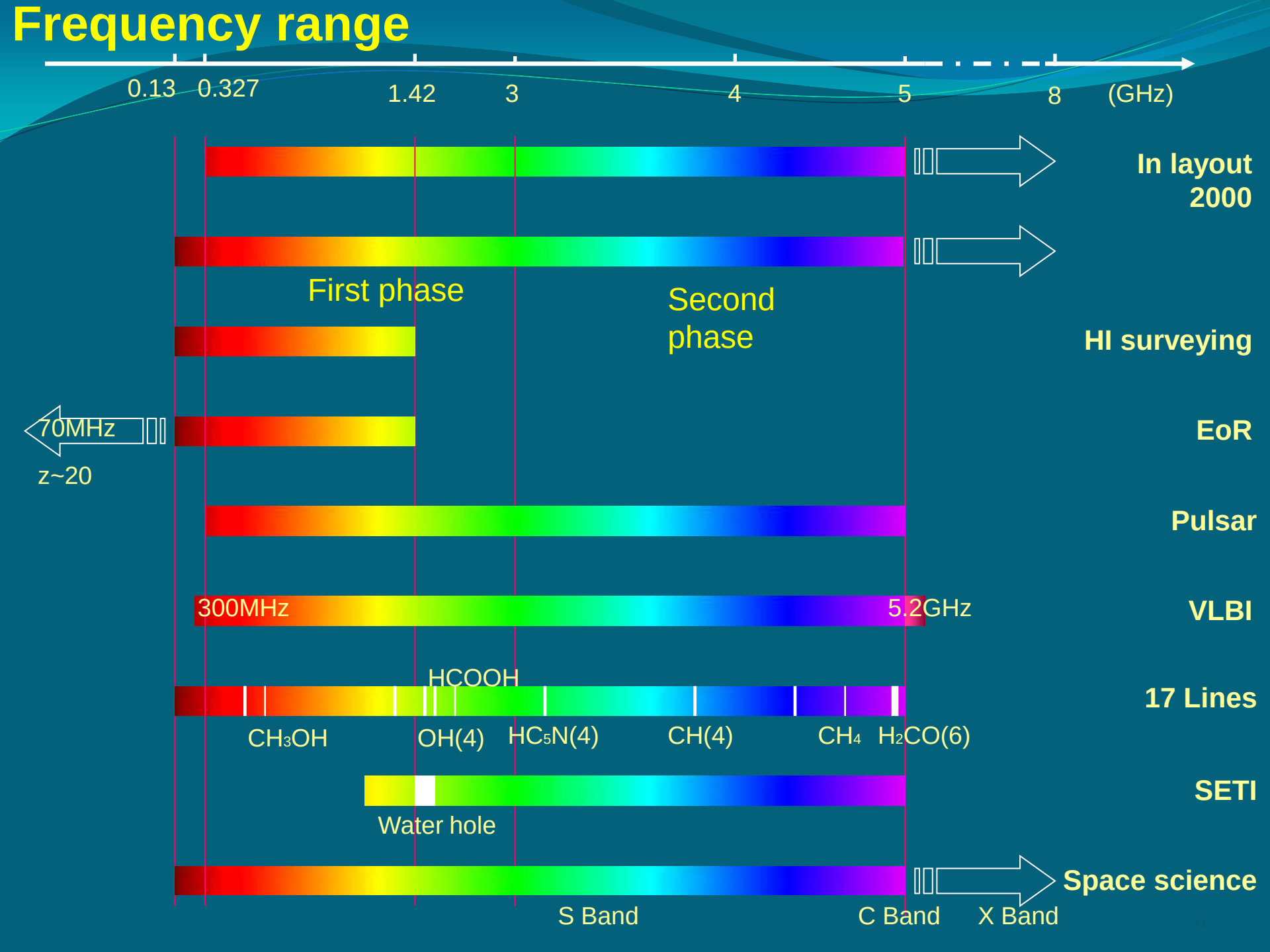
Water Tunnel 1.2km


Quick Bird Fly Oct. 6, 2005

2

2013-10-01

Frequency range



观测内容　　70 MHz到3GHz

- 脉冲星
- 原子氢（HI）21cm超精细结构谱线

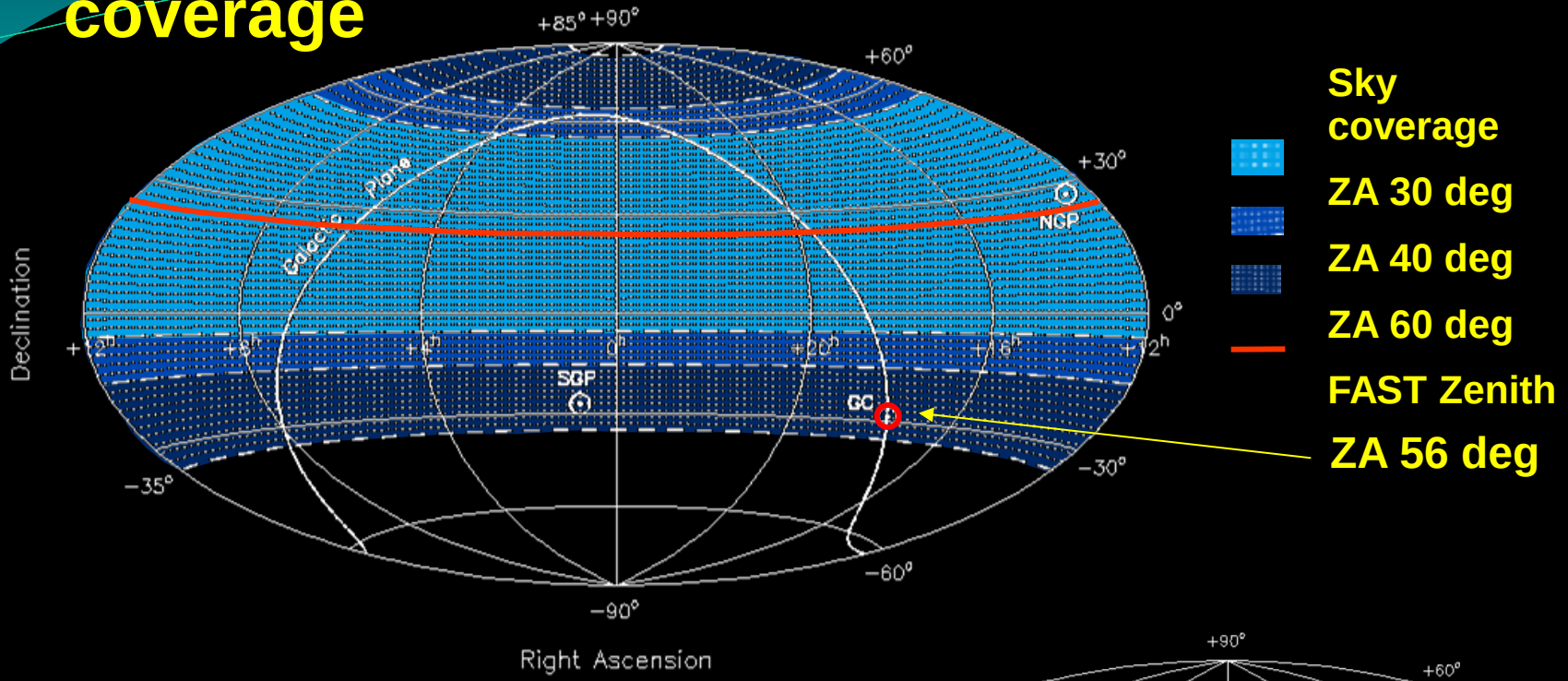
- 多种原子、多种分子谱线，射电复合线，脉泽
- 连续谱
- 行星射电辐射

9 sets of FAST receivers NAOC - JBO

No	Band (GHz)	Beams	Pol.	Cryo Tsys(K)	Science
1	0.07 – 0.14	1	RCP LCP	no 1000	High-z HI(EoR),PSR, VLBI, Lines
2	0.14 – 0.28	1	RCP LCP	no 400	High-z HI(EoR),PSR, VLBI, Lines
3	0.28 – 0.56	1 or multi	RCP LCP	no 150	High-z HI(EoR),PSR, VLBI, Lines Space weather, Low frequency DSN
4	0.56 – 1.02	1 or multi	RCP LCP	yes 60	High-z HI(EoR),PSR, VLBI, Lines Exo-planet science
5	0.320 – 0.334	1	RCP LCP	no 200	HI,PSR,VLBI Early sciences
6	0.55 – 0.64	1	RCP LCP	yes 60	HI,PSR,VLBI Early Sciences
7	1.15 – 1.72	1 L wide	RCP LCP	yes 25	HI,PSR,VLBI,SETI,Lines
8	1.23 – 1.53	19 Lnarrow multibeam	RCP LCP	yes 25	HI and PSR survey, Transients
9	2.00 – 3.00	1	RCP/ LCP	yes 25	PTA, DSN, VLBI, SETI

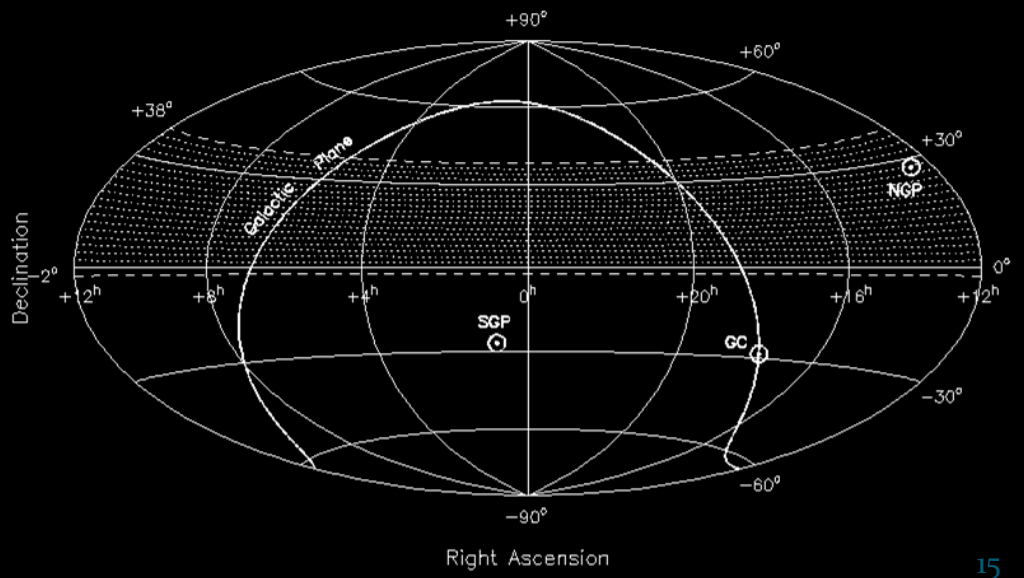
主动反射面	半径~300m，口径~500m， 球冠张角 110-120°
有效照明口径	$D_{照}$=300m
焦比	0.46-0.47
天空覆盖	天顶角 40°，跟踪 4-6h
工作频率	70MHz－3GHz
灵敏度(L 波段)	天线有效面积与系统噪声温度之比 A/T~2000 m^2/K 系统噪声温度 T~20K
偏振	全偏振（双圆或双线偏振），极化隔离度优于 30dB
分辨率(L 波段)	2.9′
多波束(L 波段)	19 个
观测换源时间	<10min
指向精度	8″

Opening angle - sky coverage



Sky coverage

FAST vs. Arecibo

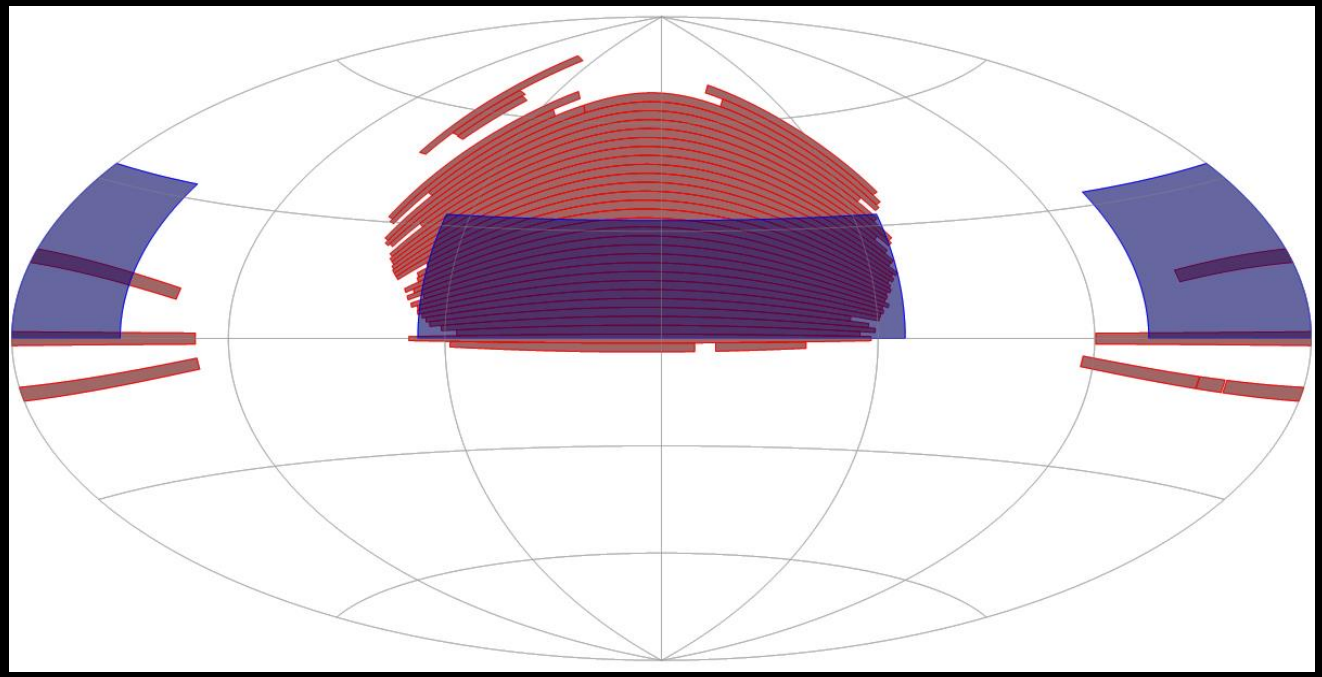


FAST的科学前沿课题

1. **宇宙起源:** 在低频射电波段对物质和能量进行精确测量, 检验暗物质模型
2. **星系演化和恒星形成**: 研究星系中星际介质和恒星形成区的物理／化学过程
3. **脉冲星及引力波**: 致密天体的状态，引力波探测，脉冲星时与导航
4. **地外行星及文明搜索**
5. **联网VLBI**: 提高VLBI灵敏度一个量级，高精度成像

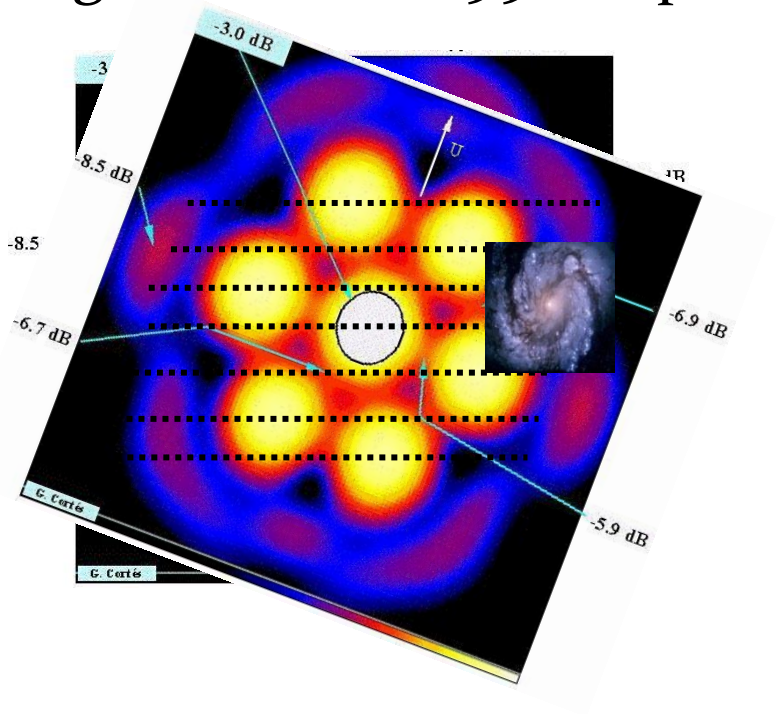
FAST all-sky HI survey

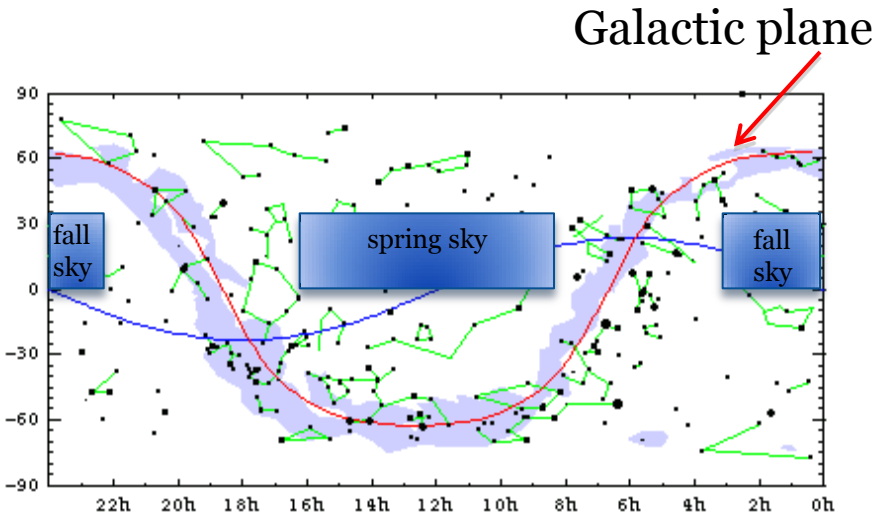
- Using a 19 beam L-band receiver to map 2.3π sterradians FAST sky at 40 sec per beam, doable in 1-2 yrs.
  - Expect about 1 million detections (Duffy et al. 2008) with $M_{HI} < 10^{11}$ $M_{\odot}$out to z ~ 0.15 in a range of environments including Coma, Hydra, Ursa Major, Persues-Pisces supercluster plus neighboring voids.



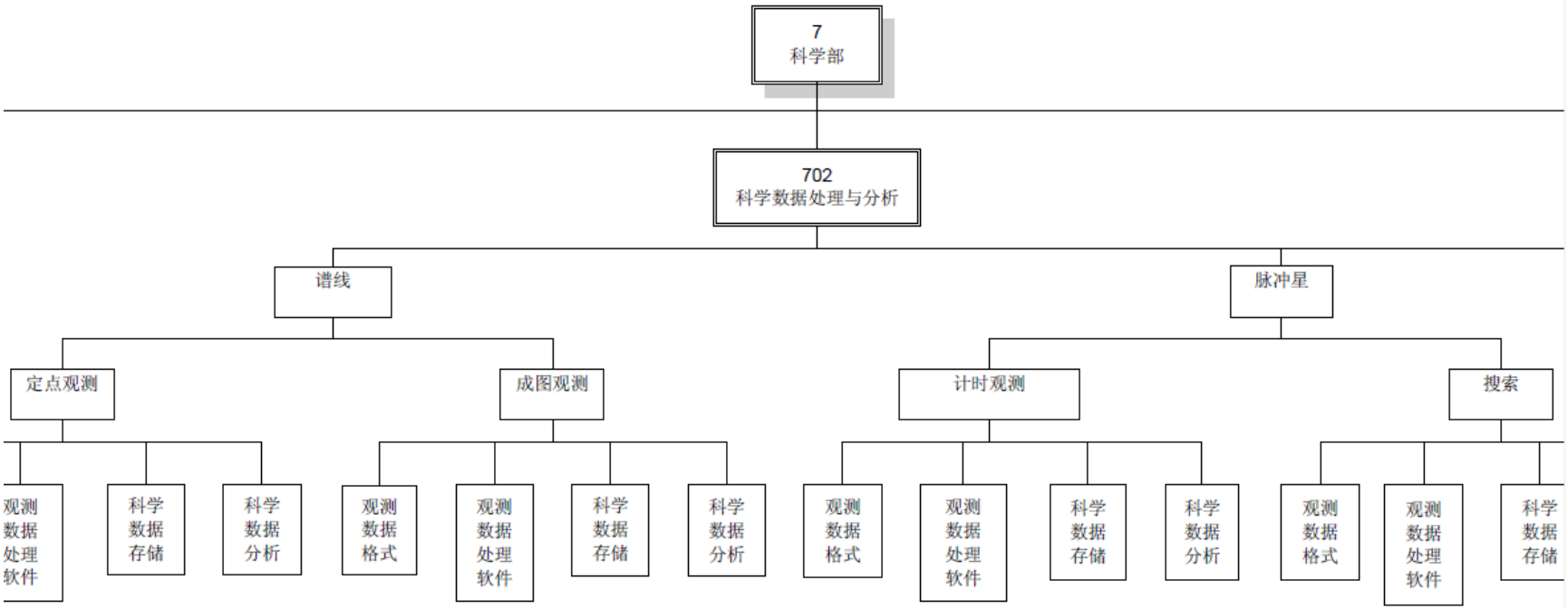
Survey strategy

- ALFA beam size: 3.3'x3.8'
- "Almost" fixed azimuth drift mode
- First pass beams spaced by 14.6'
- 2nd pass offset from the 1st to give 1.05' sampling
- Highlv efficient: 99% "open shutter" time



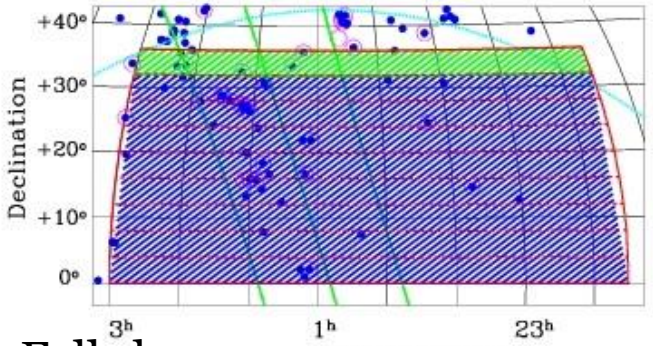


FAST科学数据处理子系统

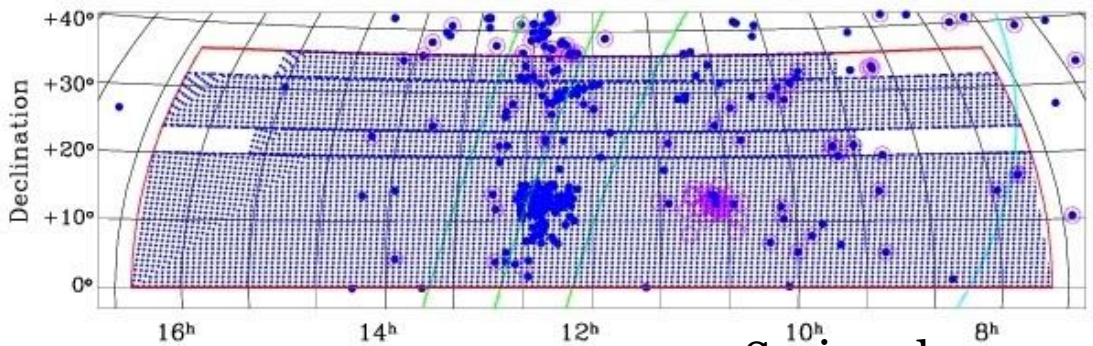


ALFALFA 进展 2012 – 05

- 779 observing runs, 4300 hours, most remotely
- ~20,000 good detections in 50% of final area



Fall sky

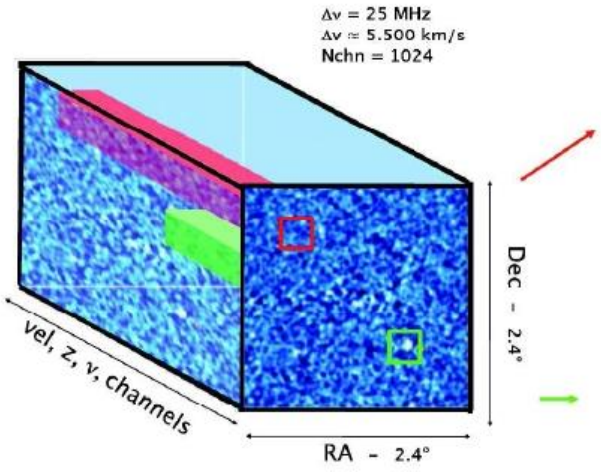


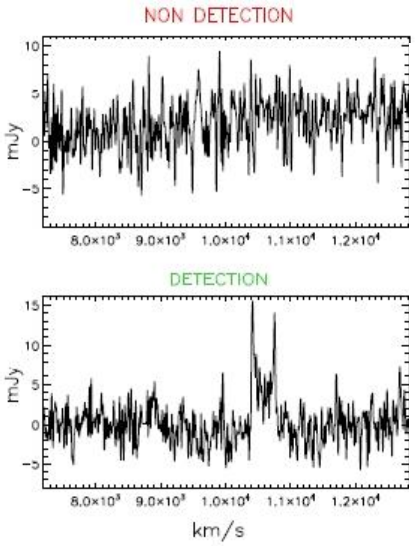
Spring sky

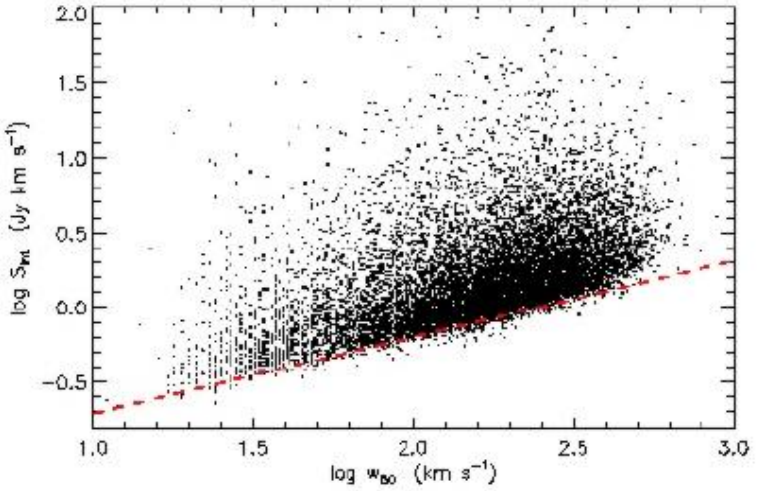
- Will complete in fall 2012
- Follow up pointed observations with LBW
  - Dark galaxies, OHM candidates, etc.

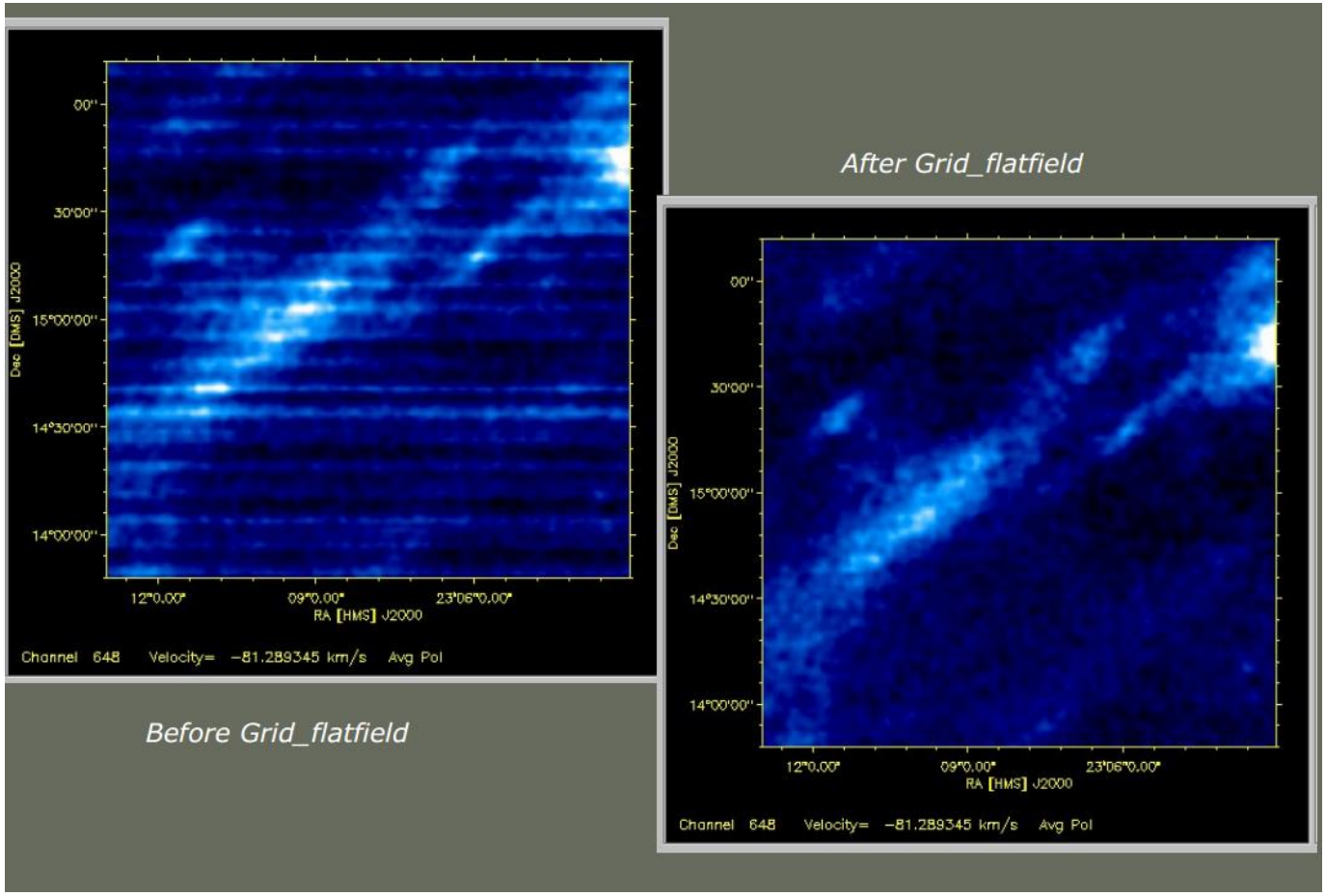
谱线数据处理系统

- 美国康乃尔大学ALFALFA星系巡天数据处理系统
- 3万个星系

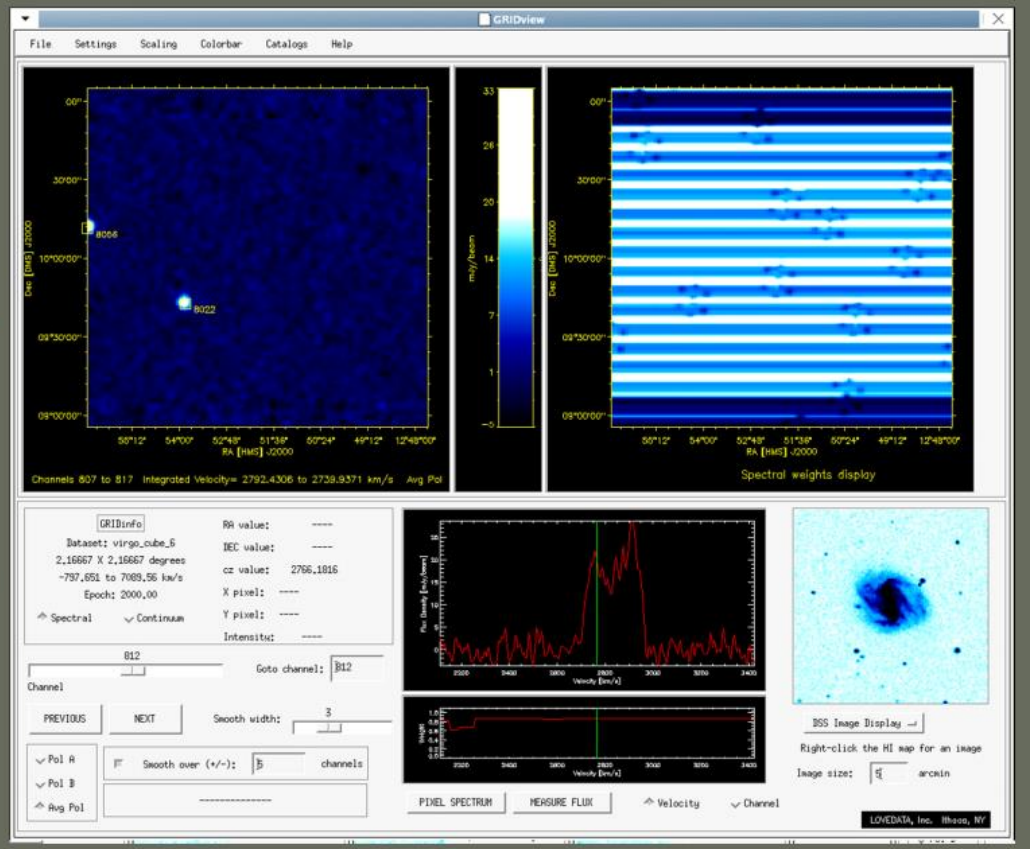


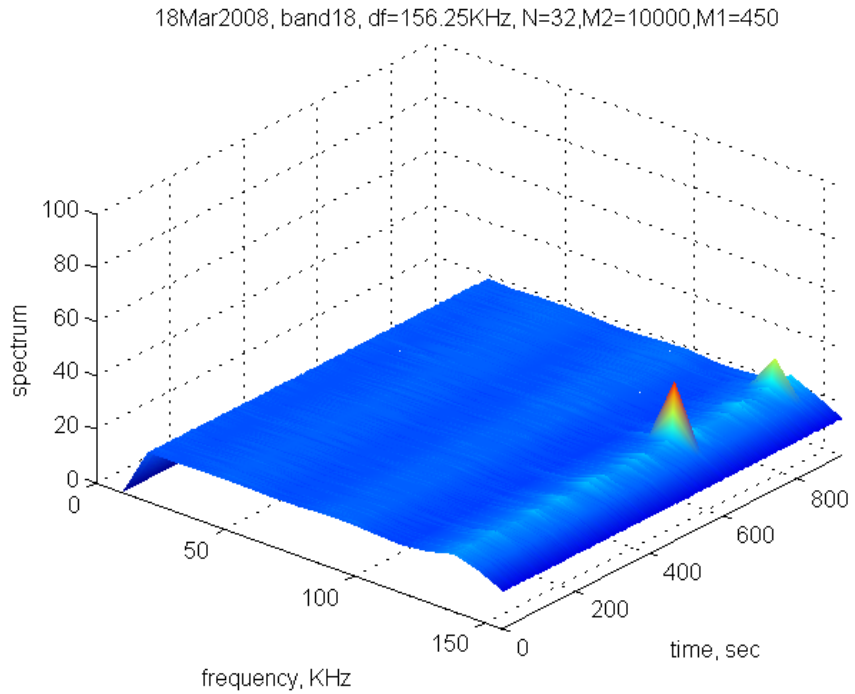


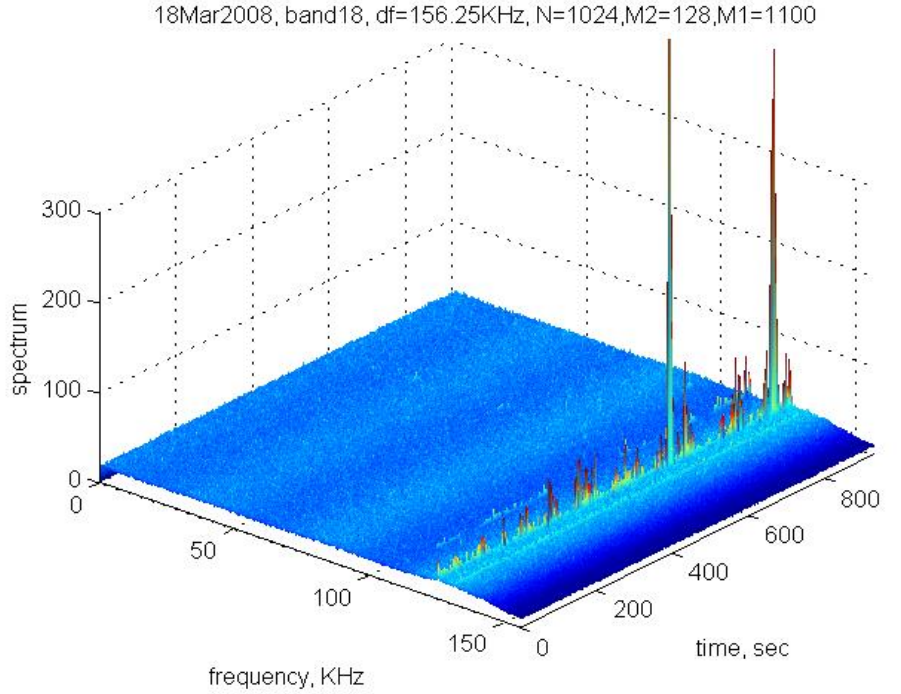


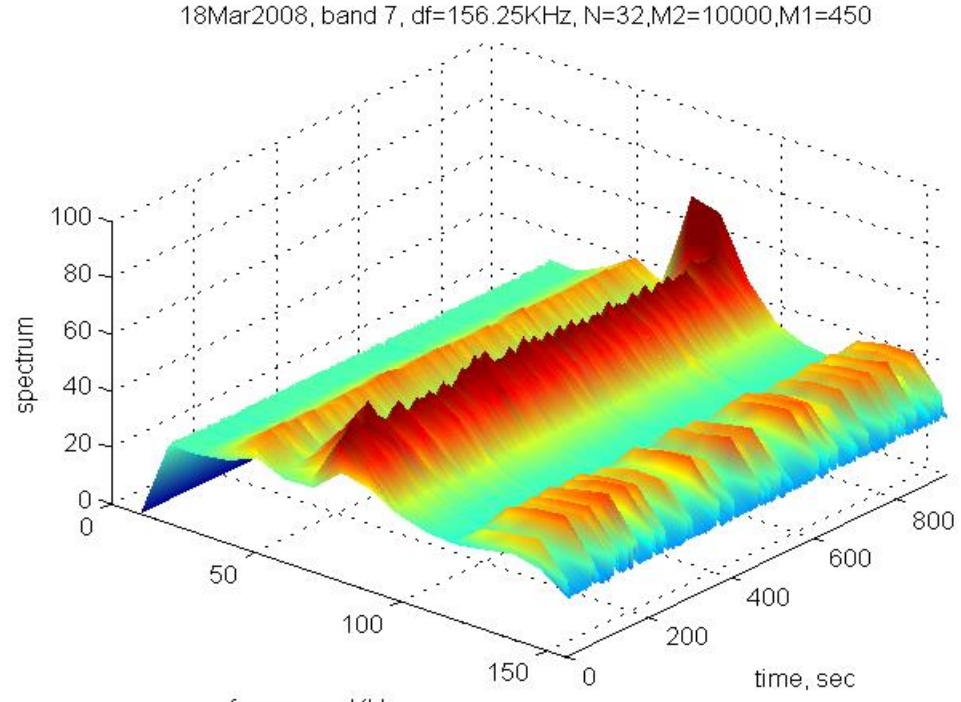

00''
30'00''
15°00'00''
14°30'00''
14°00'00''
Dec [DMS] J2000
12°0.00'
09°0.00'
23°06°0.00'
RA [HMS] J2000
Channel 648 Velocity= −81.289345 km/s Avg Pol
Before Grid_flatfield
After Grid_flatfield
00''
30'00''
15°00'00''
14°30'00''
14°00'00''
Dec [DMS] J2000
12°0.00'
09°0.00'
23°06°0.00'
RA [HMS] J2000
Channel 648 Velocity= −81.289345 km/s Avg Pol

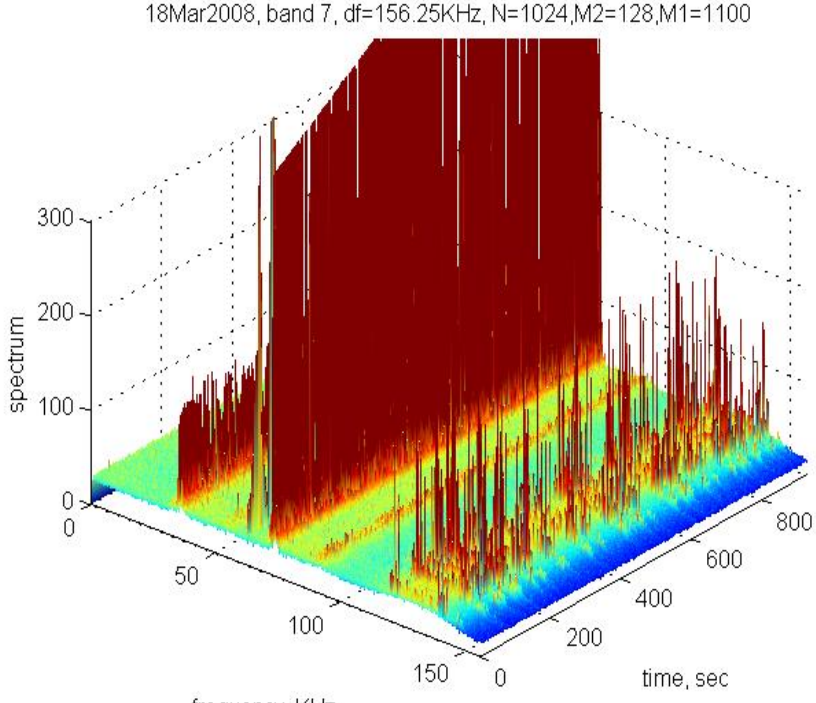
- *Data cubes and corresponding 3D catalogs are examined in GRIDview.*
- *The upper left display is a channel map; at upper right is the corresponding weights map.*
- *Controls allow user to view channel or integrated maps at different velocities.*
- *DSS, DSS2, Sloan, NVSS images can be fetched.*
- *NED and other on-line catalogs – including internal ones – can be accessed and overplotted*

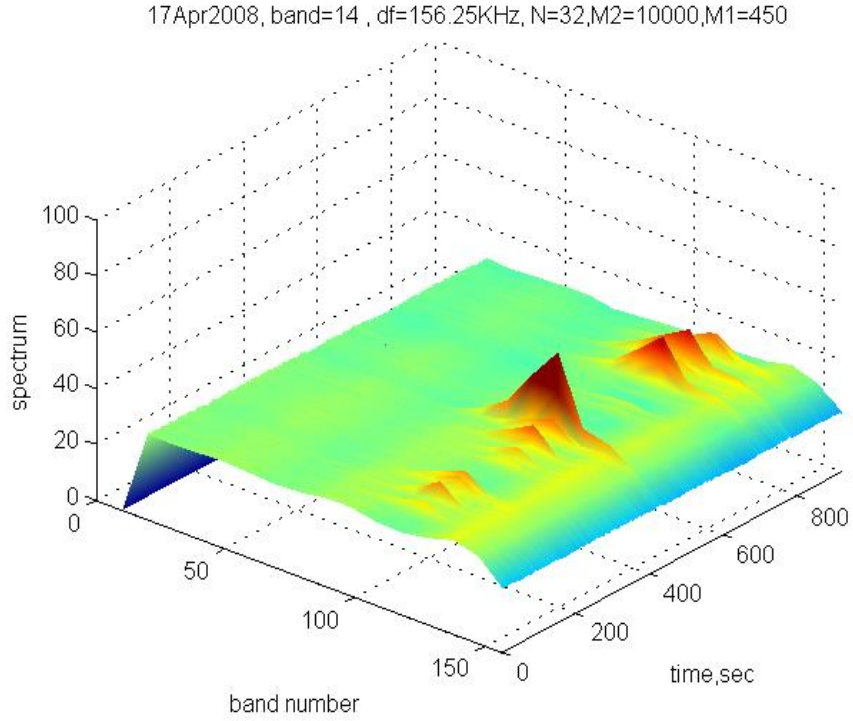


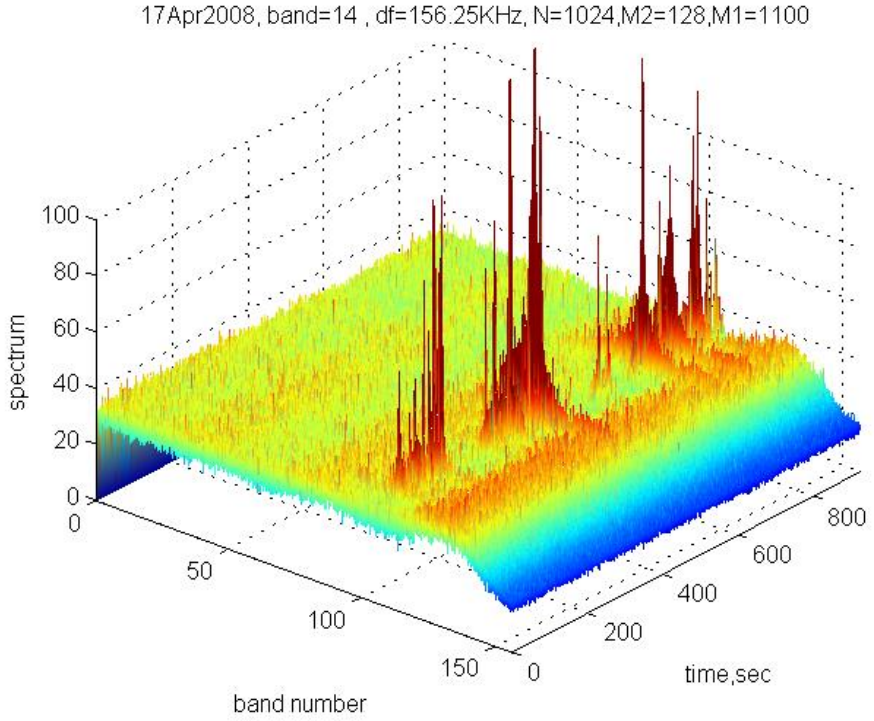

18Mar2008, band18, df=156.25KHz, N=32,M2=10000,M1=450
spectrum
frequency, KHz
time, sec

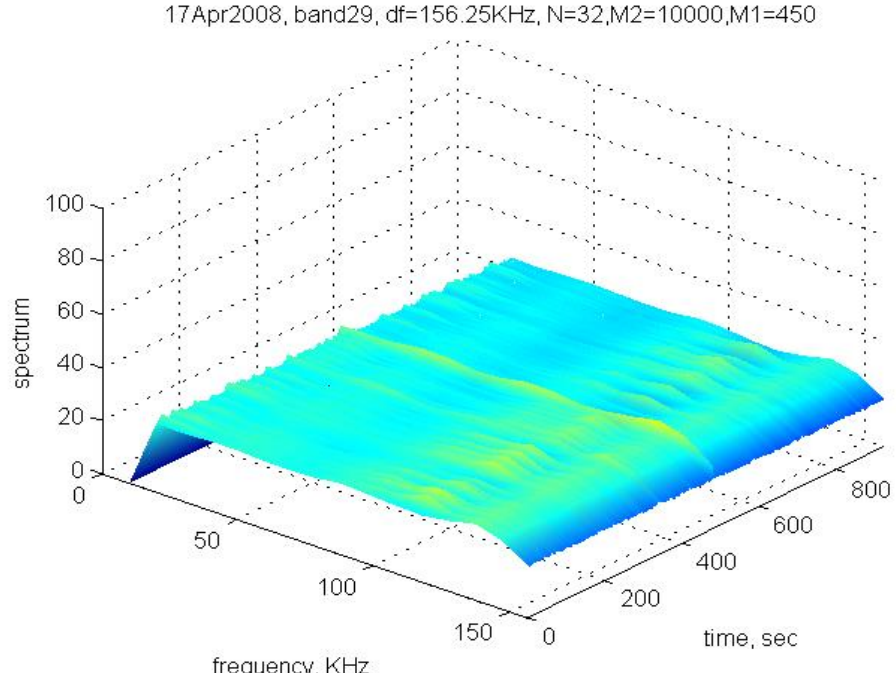

18Mar2008, band18, df=156.25KHz, N=1024,M2=128,M1=1100
spectrum
frequency, KHz
time, sec

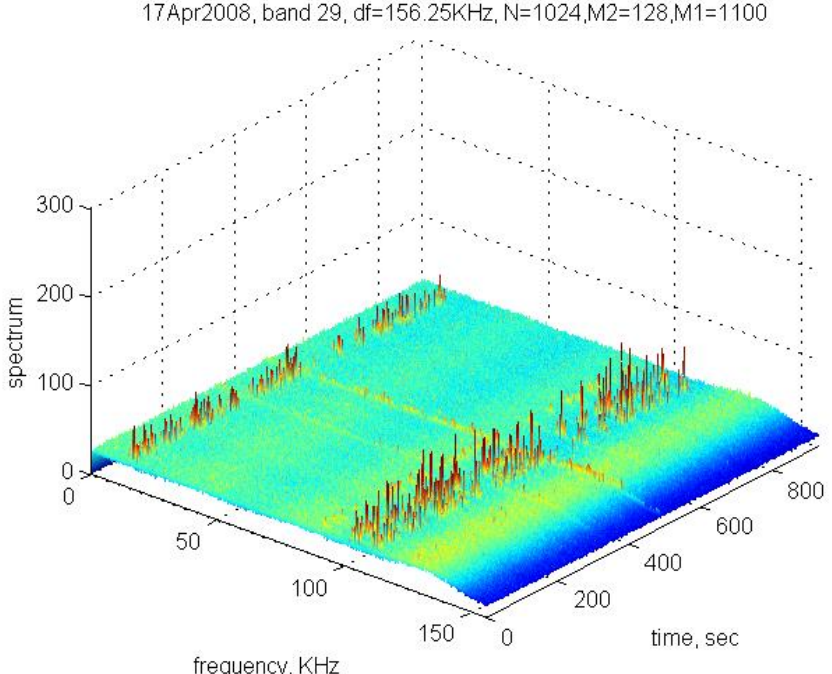

18Mar2008, band 7, df=156.25KHz, N=32,M2=10000,M1=450
spectrum
time, sec

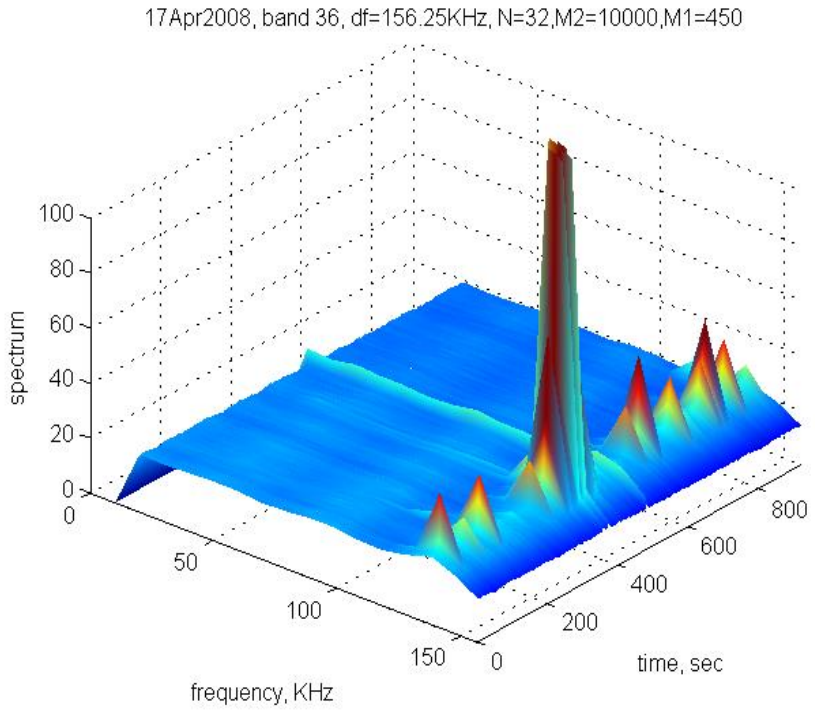

18Mar2008, band 7, df=156.25KHz, N=1024,M2=128,M1=1100
spectrum
time, sec

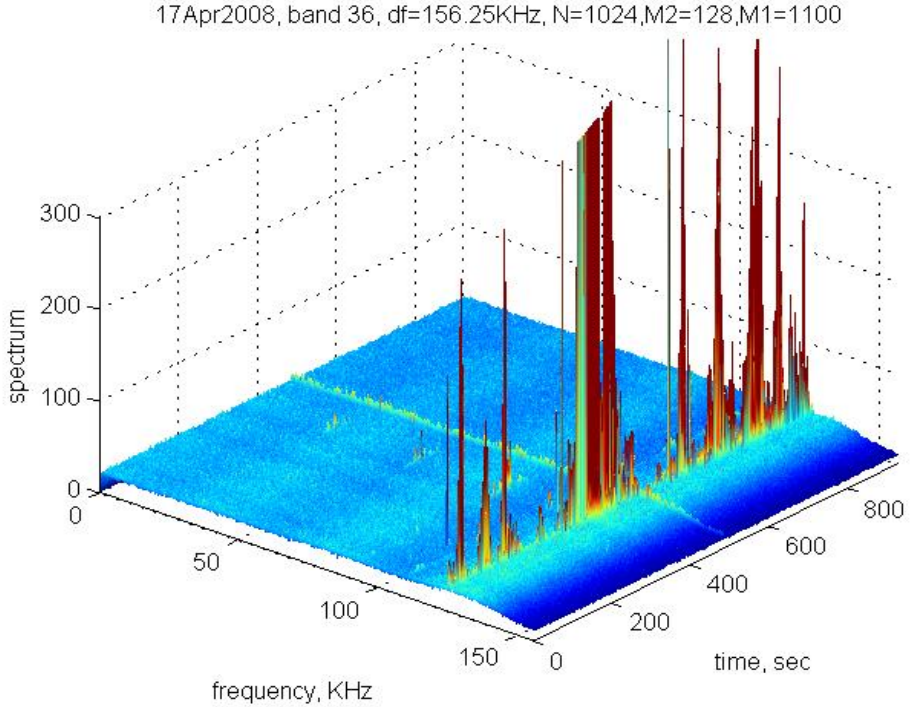

17Apr2008, band=14 , df=156.25KHz, N=32,M2=10000,M1=450
spectrum
band number
time,sec

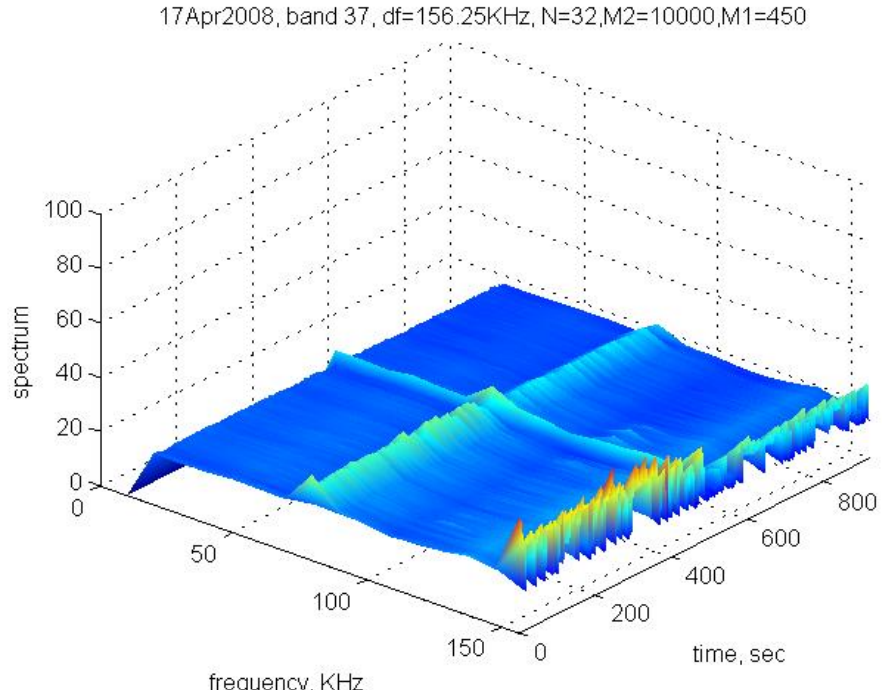

17Apr2008, band=14 , df=156.25KHz, N=1024,M2=128,M1=1100
spectrum
band number
time,sec

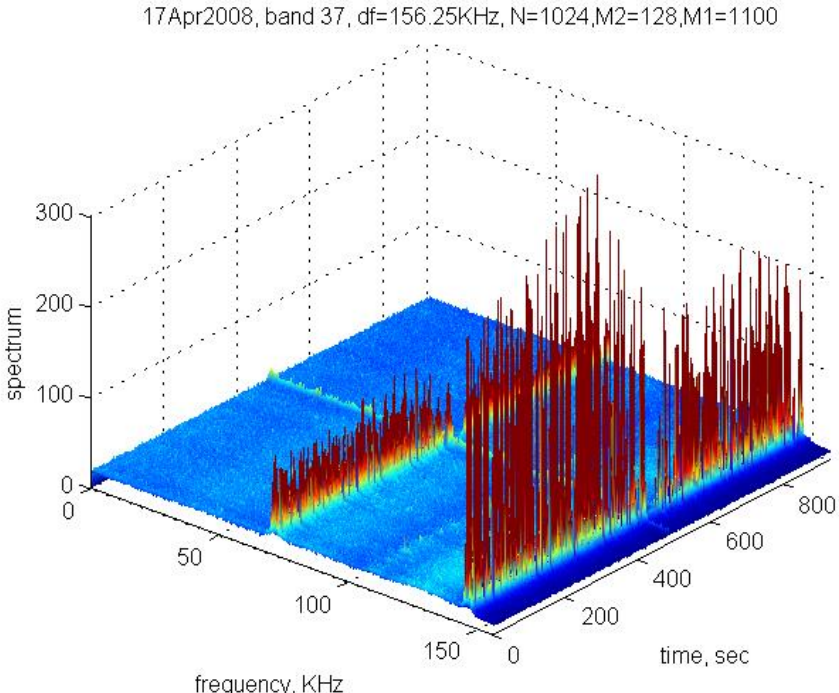

17Apr2008, band29, df=156.25KHz, N=32,M2=10000,M1=450
spectrum
frequency, KHz
time, sec

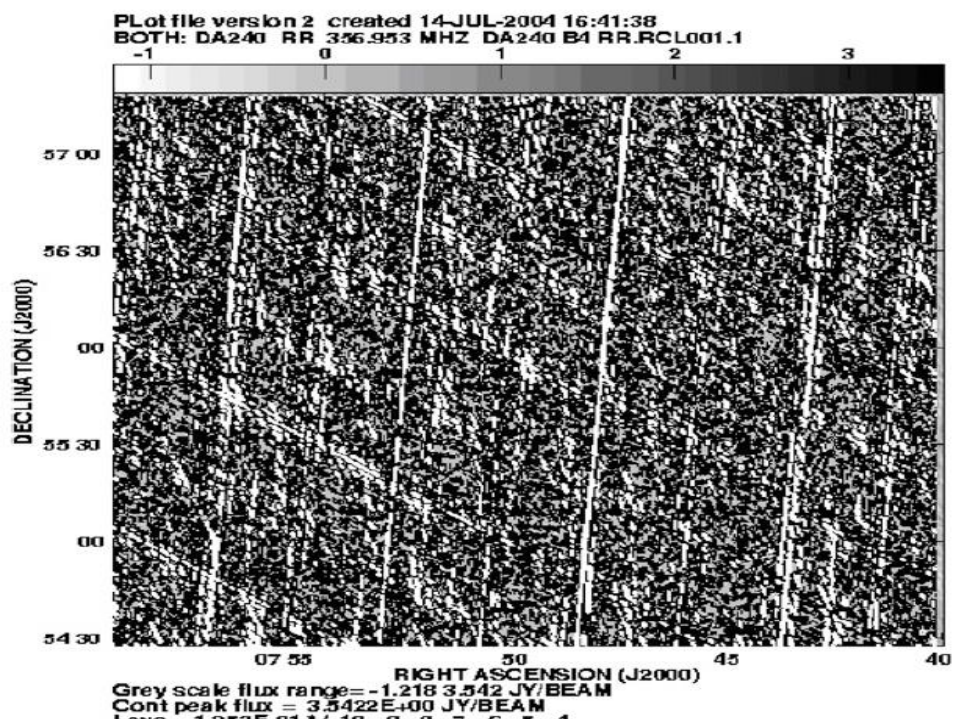

17Apr2008, band 29, df=156.25KHz, N=1024,M2=128,M1=1100
spectrum
frequency, KHz
time, sec

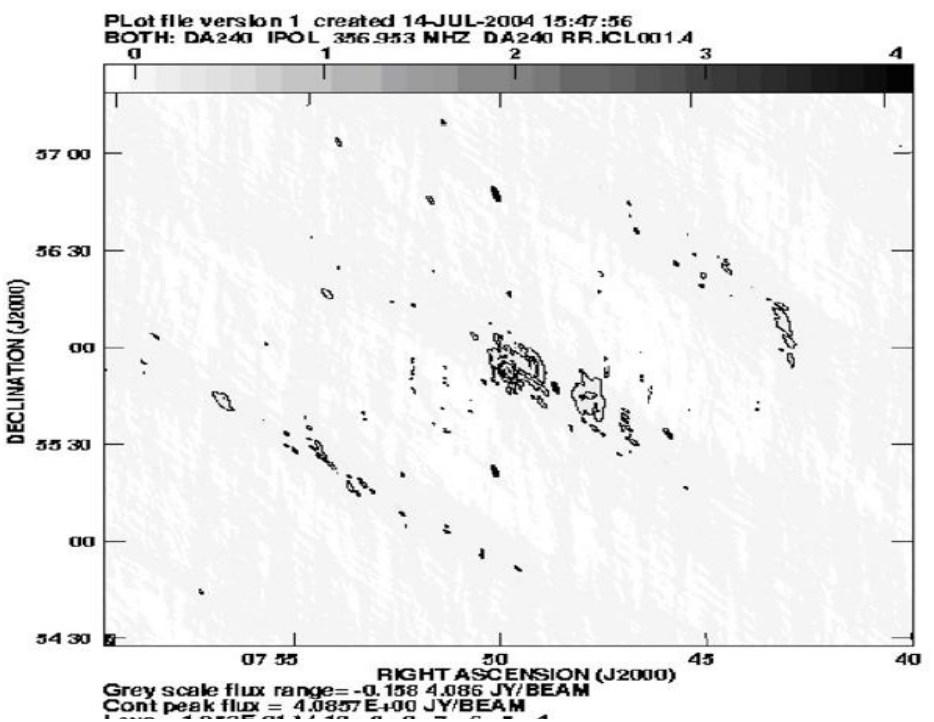

17Apr2008, band 36, df=156.25KHz, N=32,M2=10000,M1=450
spectrum
frequency, KHz
time, sec

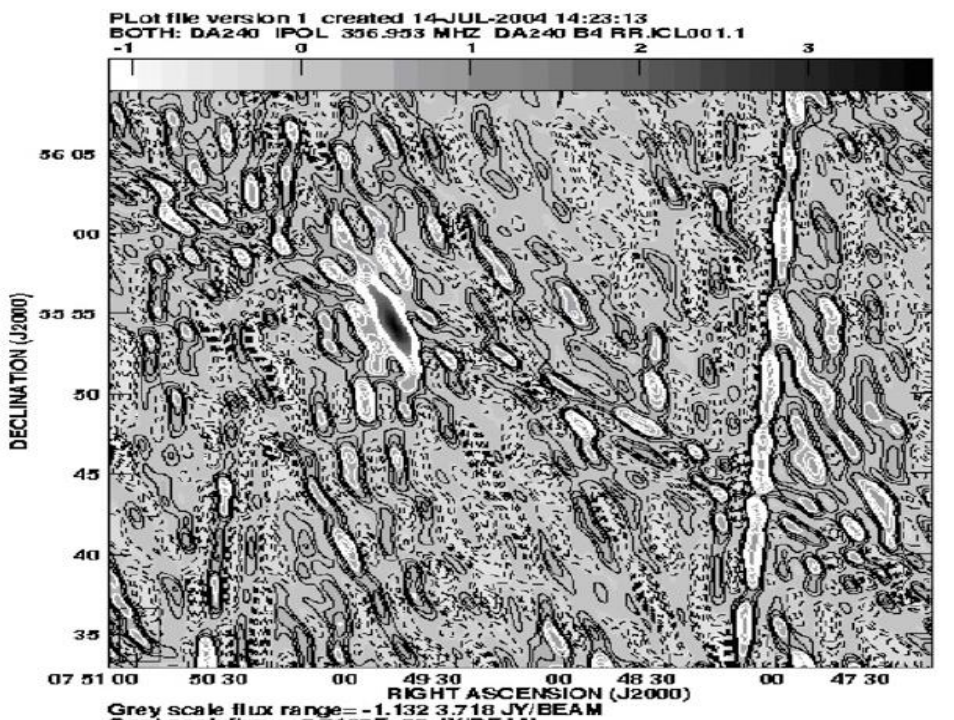

17Apr2008, band 36, df=156.25KHz, N=1024,M2=128,M1=1100
spectrum
frequency, KHz
time, sec

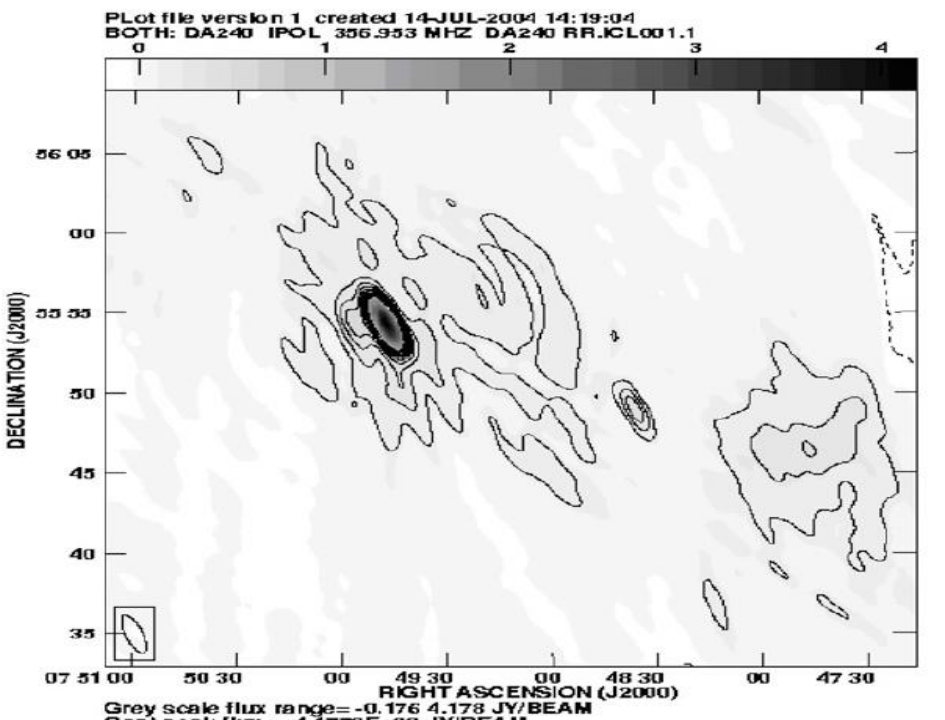

17Apr2008, band 37, df=156.25KHz, N=32,M2=10000,M1=450
spectrum
frequency, KHz
time, sec

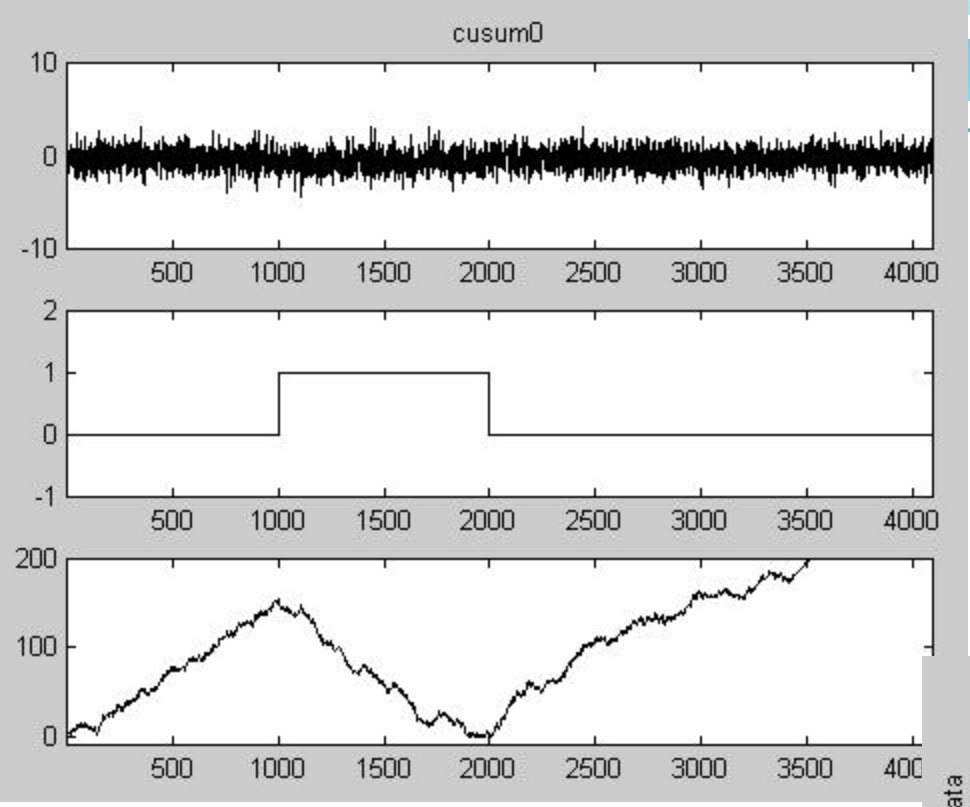

17Apr2008, band 37, df=156.25KHz, N=1024,M2=128,M1=1100
spectrum
frequency, KHz
time, sec


PLot file version 2 created 14-JUL-2004 16:41:38
BOTH: DA240 RR 356.953 MHZ DA240 B4 RR.RCL001.1
-1 0 1 2 3
57 00
56 30
00
55 30
00
54 30
DECLINATION (J2000)
07 55
50
45
40
RIGHT ASCENSION (J2000)
Grey scale flux range= -1.218 3.542 JY/BEAM
Cont peak flux = 3.5422E+00 JY/BEAM


PLot file version 1 created 14-JUL-2004 15:47:56
BOTH: DA240 IPOL 356.953 MHZ DA240 RR.ICL001.4
0 1 2 3 4
57 00
56 30
00
55 30
00
54 30
DECLINATION (J2000)
07 55
50
45
40
RIGHT ASCENSION (J2000)
Grey scale flux range= -0.158 4.086 JY/BEAM
Cont peak flux = 4.0857E+00 JY/BEAM


PLot file version 1 created 14-JUL-2004 14:23:13
BOTH: DA240 IPOL 356.953 MHZ DA240 B4 RR.ICL001.1
-1 0 1 2 3
56 05
00
55 55
50
45
40
35
DECLINATION (J2000)
07 51 00
50 30
00
49 30
00
48 30
00
47 30
RIGHT ASCENSION (J2000)
Grey scale flux range= -1.132 3.718 JY/BEAM


PLot file version 1 created 14-JUL-2004 14:19:04
BOTH: DA240 IPOL 356.953 MHZ DA240 RR.ICL001.1
0 1 2 3 4
56 05
00
55 55
50
45
40
35
DECLINATION (J2000)
07 51 00
50 30
00
49 30
00
48 30
00
47 30
RIGHT ASCENSION (J2000)
Grey scale flux range= -0.176 4.178 JY/BEAM



Change of mean

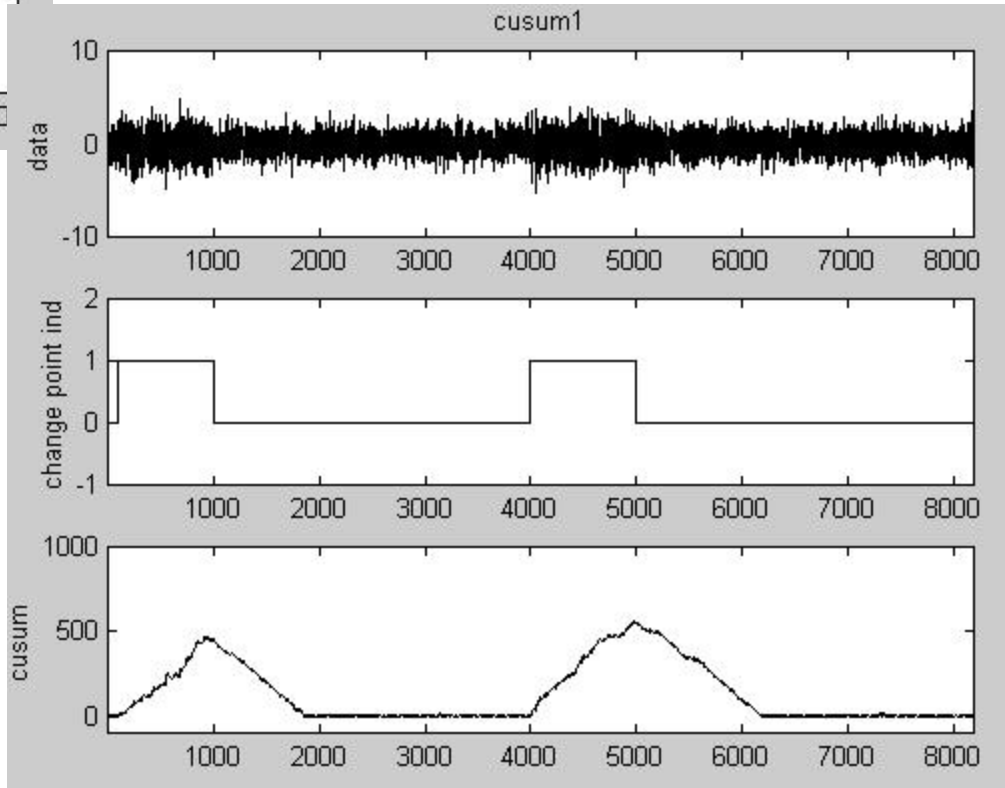
Cumulative sum

$$S_t = \max(0, S_{t-1} + x_t - \mu_0 - k)$$

$$S_0 = 0$$

signal if $S_t > h$

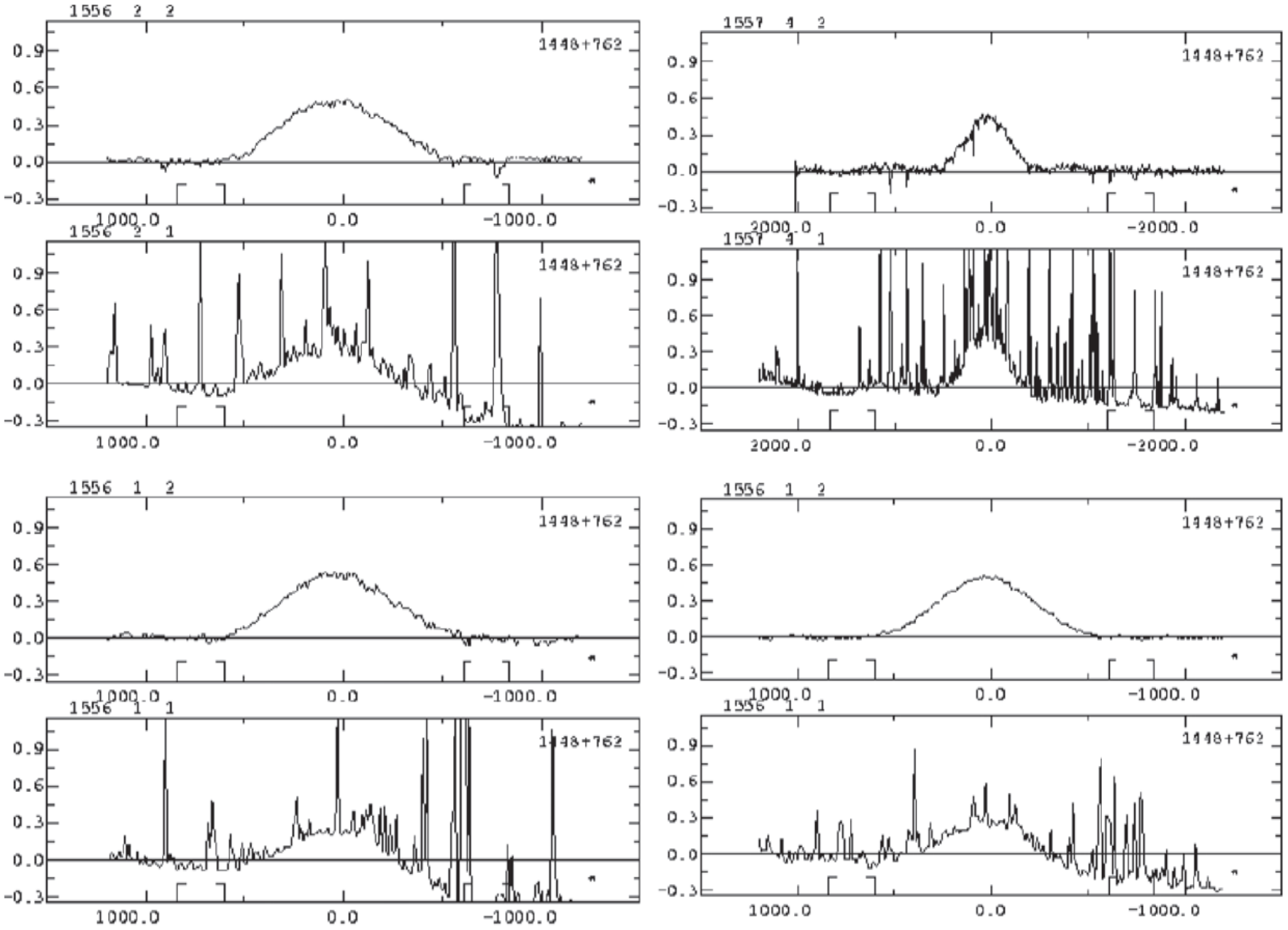
Change of variance

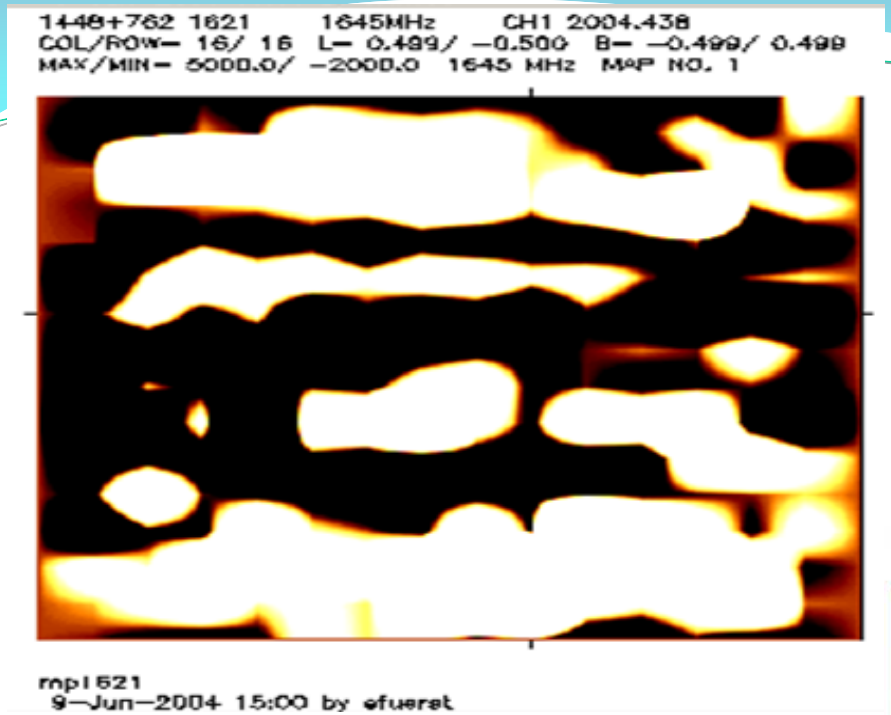


$$x_0, x_1, \ldots . x_r, \ldots . x_n$$

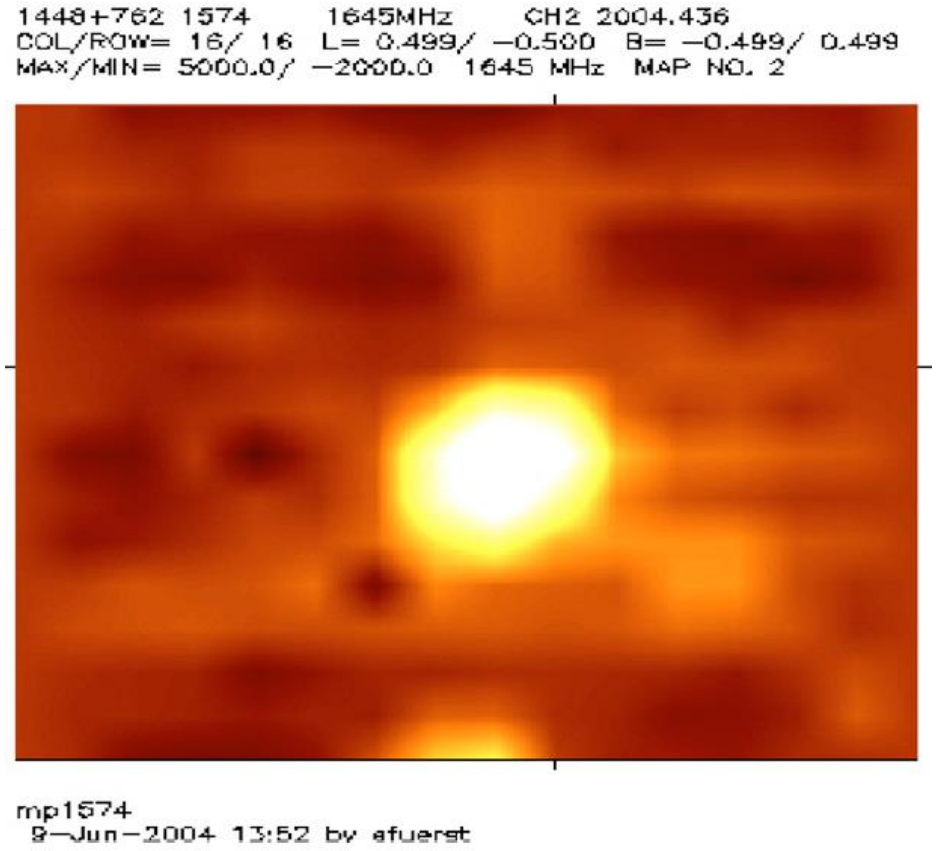
$$S_r^n(\mu_0, k) = \sum_{i=r}^{n} (x_i - \mu_0 - k/2),$$

$$k = \mu_0 - \mu_1$$


1556 2 2
1448+762
0.9
0.6
0.3
0.0
-0.3
1000.0
0.0
-1000.0
1556 2 1
1448+762
1557 4 2
1448+762
2000.0
0.0
-2000.0
1557 4 1
1448+762
1556 1 2
1448+762
1556 1 1
1448+762

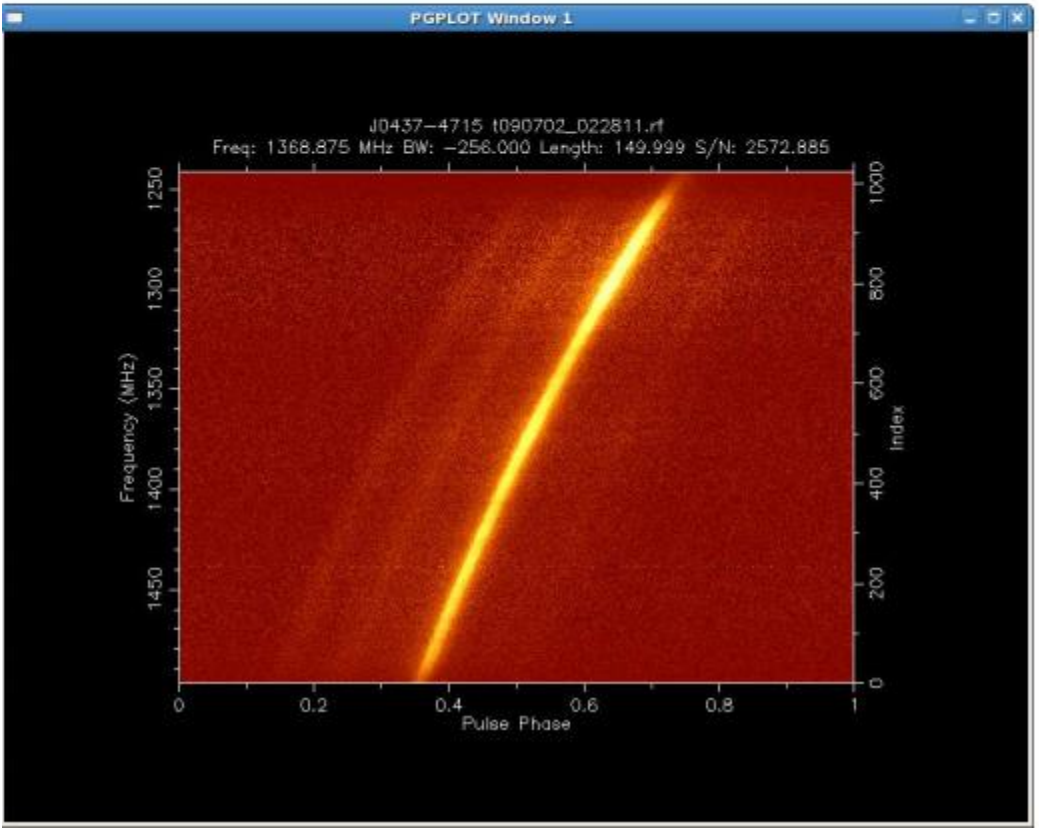


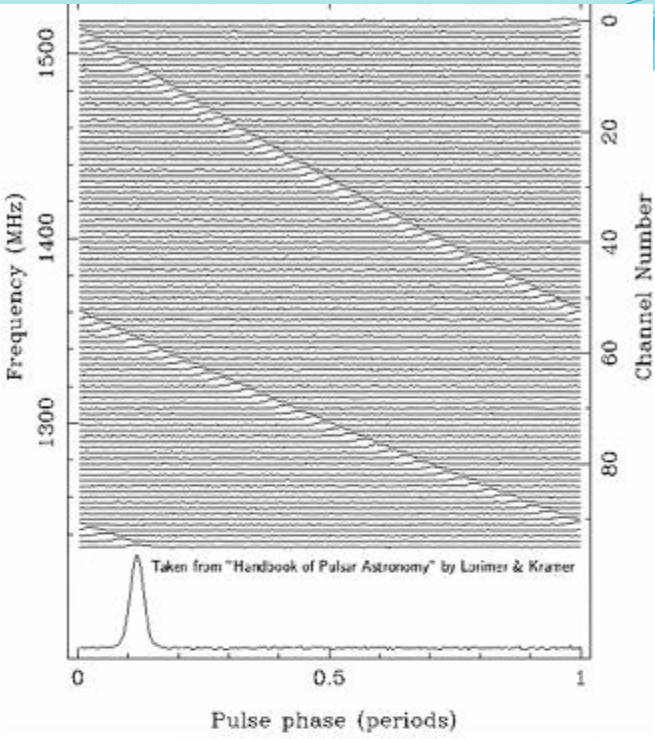
Source maps synthesized with the scans in the above shown figures: with RFI (top left) and after RFI mitigation (bottom right)

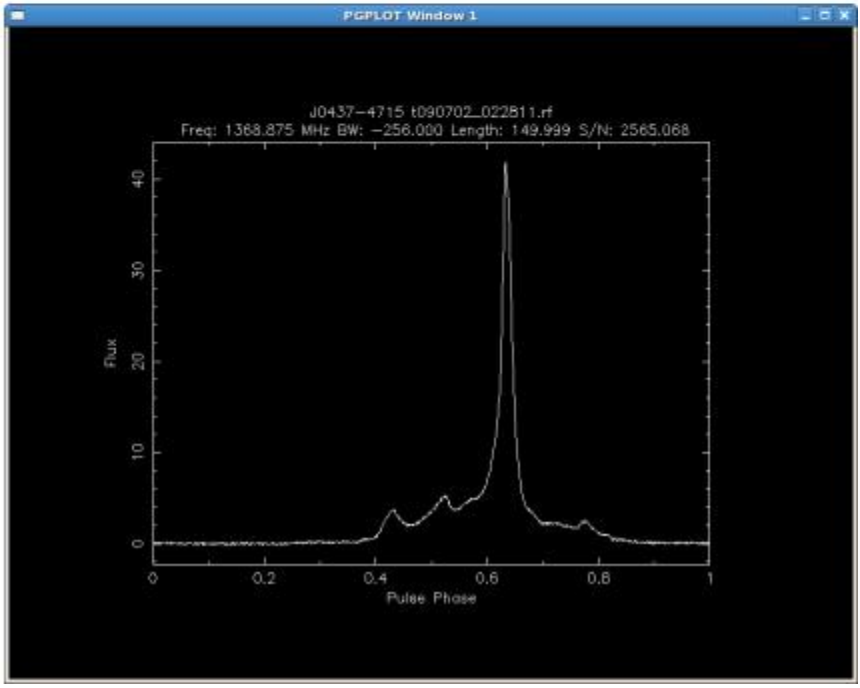


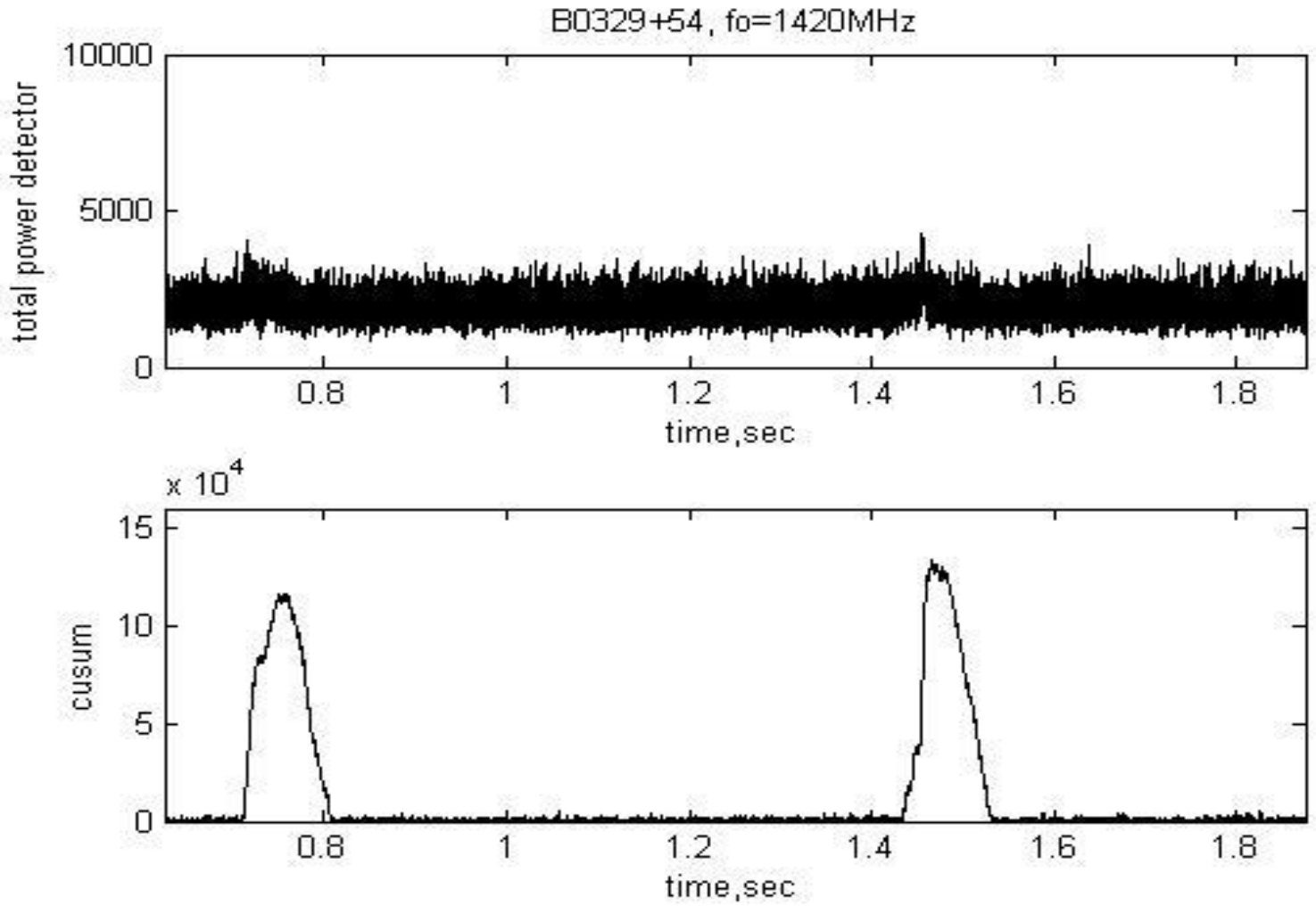
脉冲星信号特点

- 周期性时域信号
- 频域色散

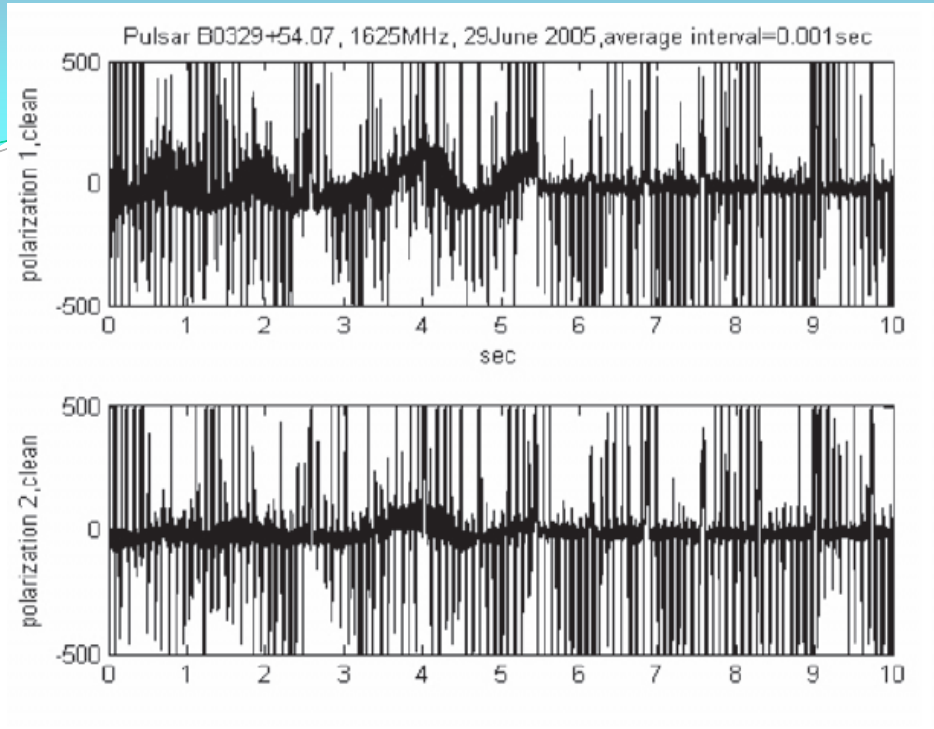


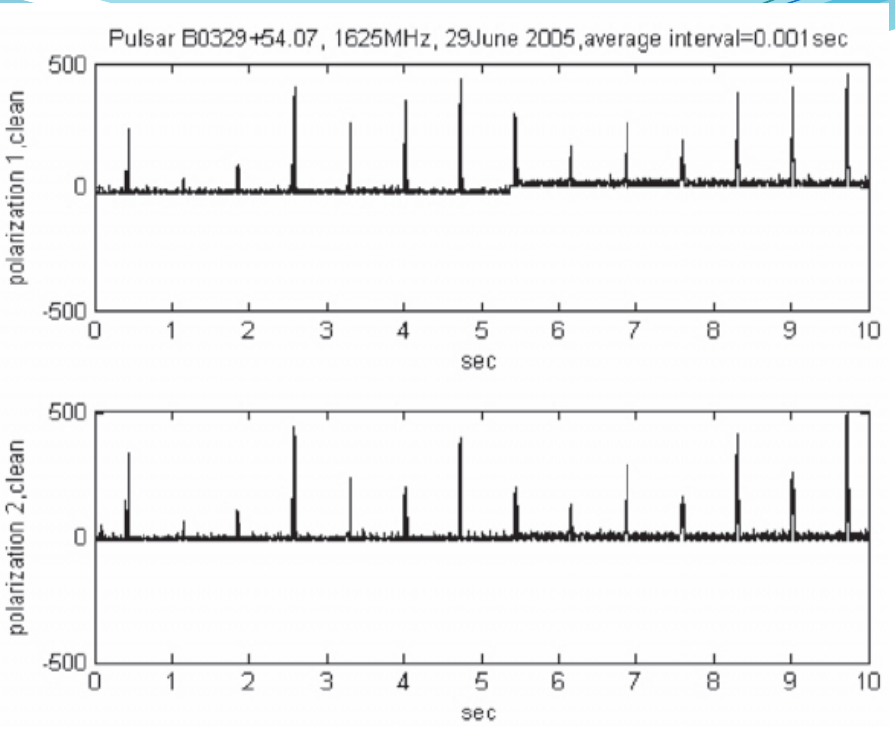


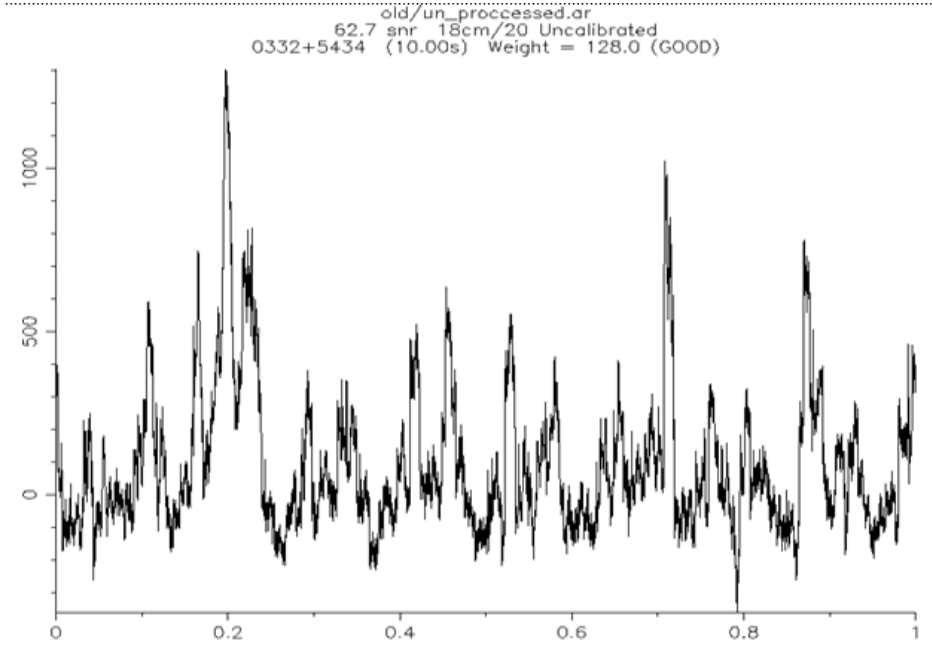


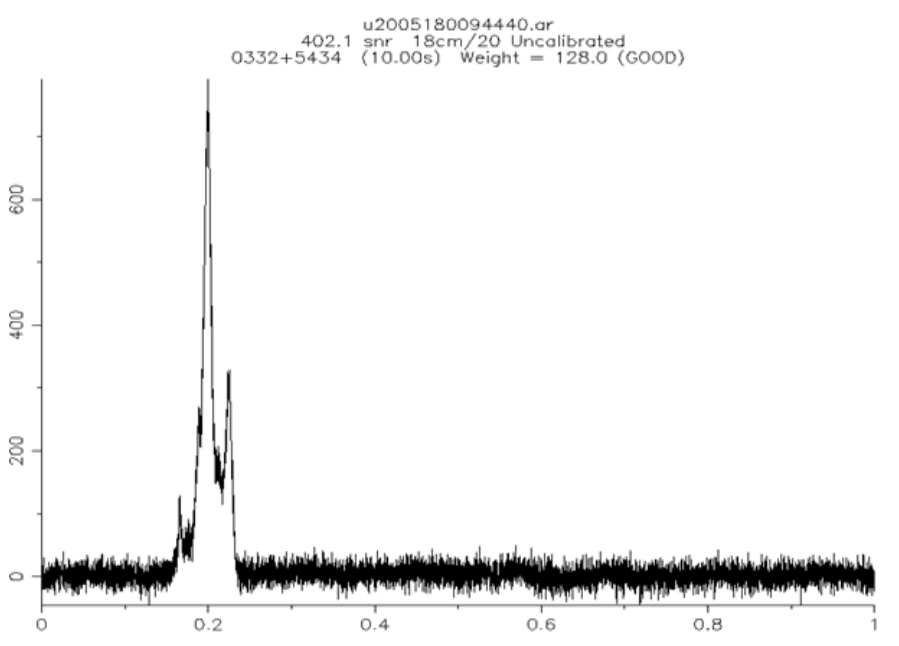


Upper panel: TPD data corresponding to the are aorund the second and the third pu
integration time = 1.25 x 10^(-6)s.
Lower panel: cusum calculated with samples od the data in the upper panel.

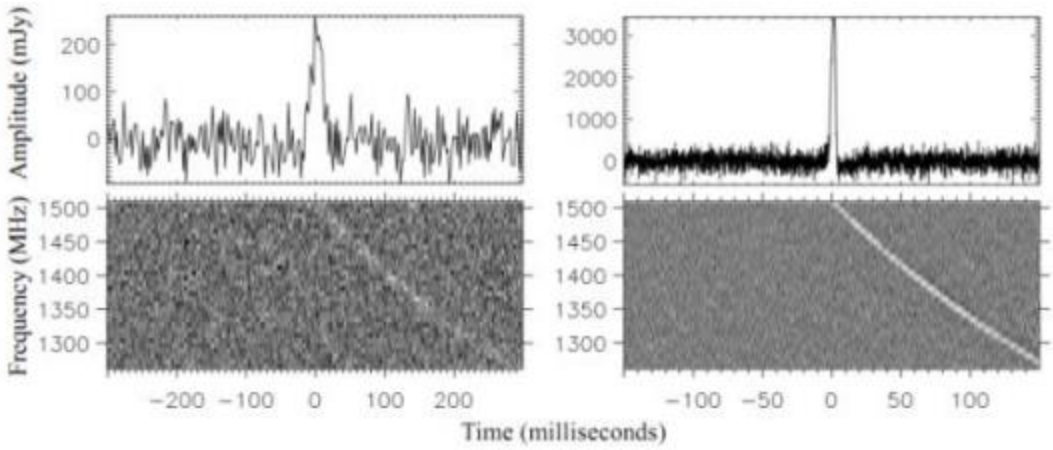

Pulsar B0329+54.07, 1625MHz, 29June 2005,average interval=0.001sec
polarization 1,clean
polarization 2,clean
sec


Pulsar B0329+54.07, 1625MHz, 29June 2005,average interval=0.001sec
polarization 1,clean
polarization 2,clean
sec

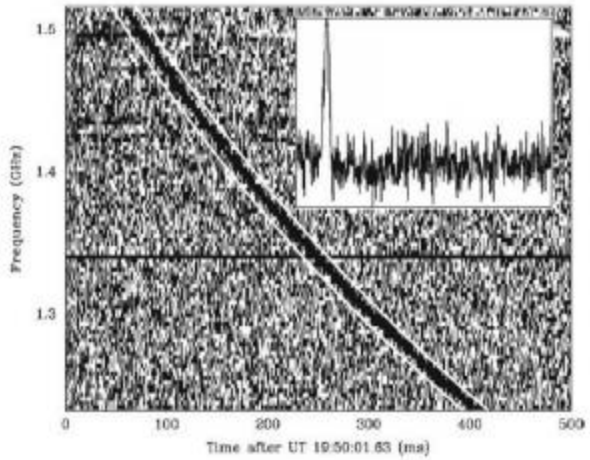

old/un_proccessed.ar
62.7 snr 18cm/20 Uncalibrated
0332+5434 (10.00s) Weight = 128.0 (GOOD)


u2005180094440.ar
402.1 snr 18cm/20 Uncalibrated
0332+5434 (10.00s) Weight = 128.0 (GOOD)

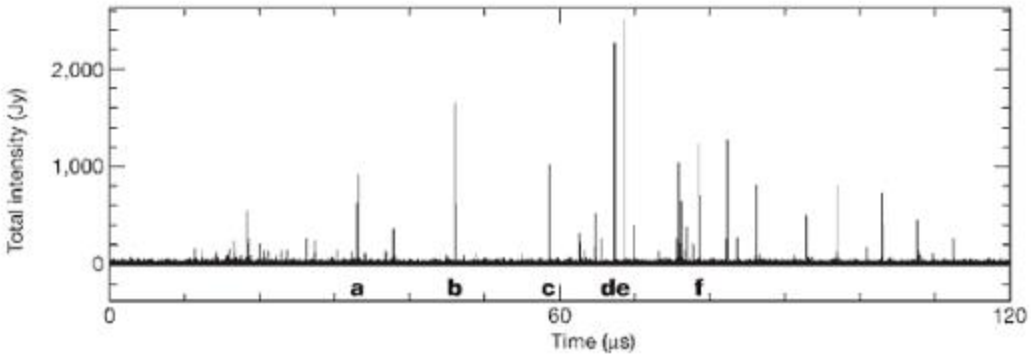
单个脉冲



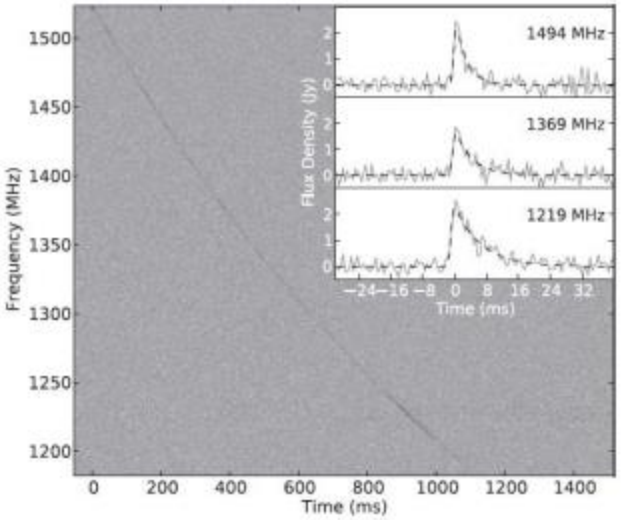
McLaughlin et al. 2006 RRAT



Lorimer et al. 2007

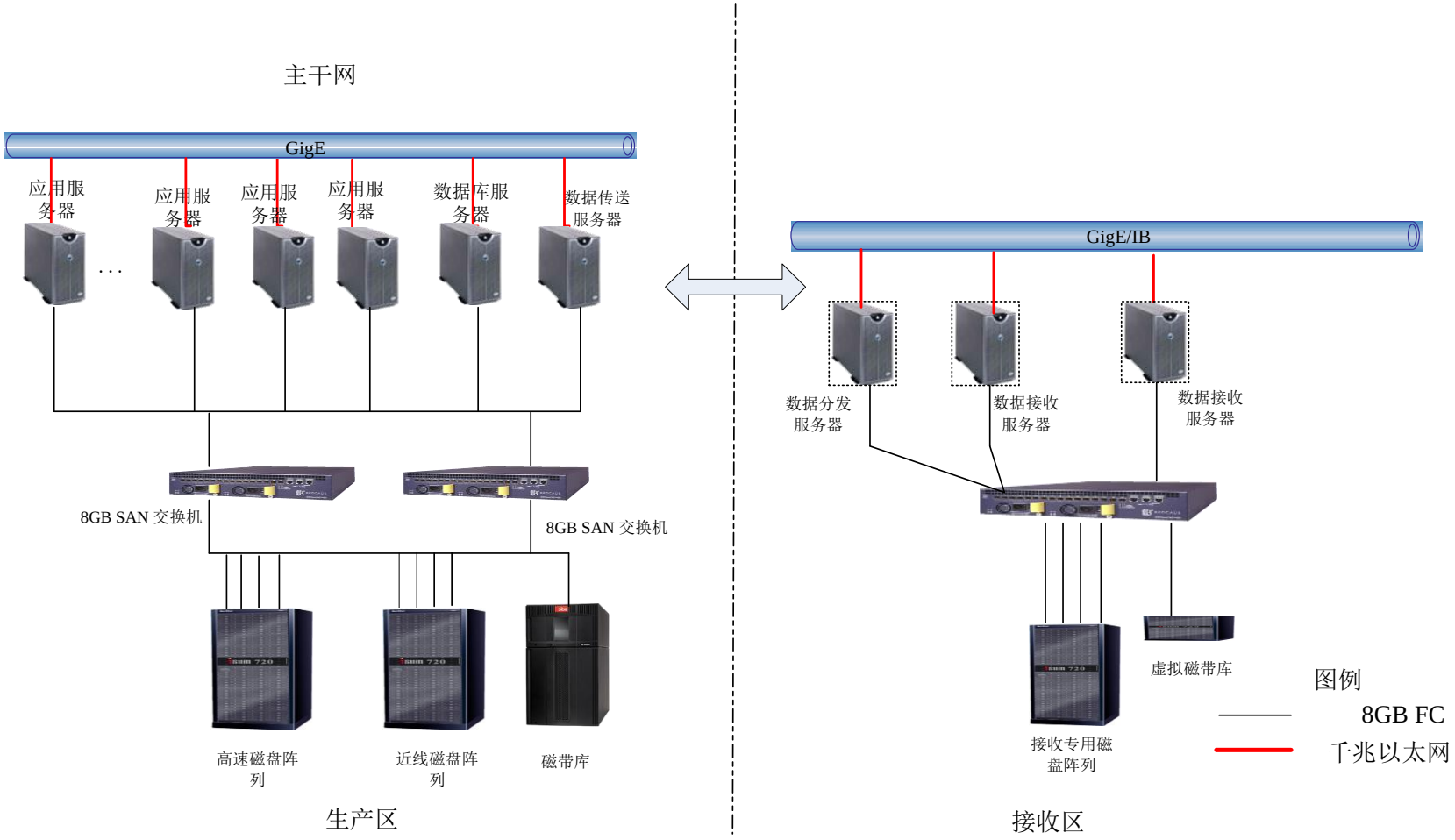


Hankins et al. 2003, ns pulse

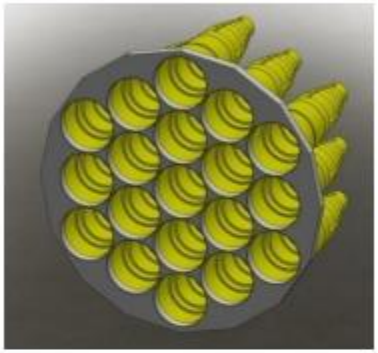


Thornton et al 2013

科学数据存储系统设计


主干网
GigE
应用服务器
应用服务器
应用服务器
应用服务器
数据库服务器
数据传送服务器
8GB SAN 交换机
8GB SAN 交换机
高速磁盘阵列
近线磁盘阵列
磁带库
生产区
GigE/IB
数据分发服务器
数据接收服务器
数据接收服务器
虚拟磁带库
接收专用磁盘阵列
接收区
图例
8GB FC
千兆以太网

FAST脉冲星存储需求



- 80MB/s*19 ~ 1.6GB/s~5TB/h
- 观测~8h/day, ~40TB/day
- 介质
  - 磁盘
  - 磁带
- 方案
  - 短期磁盘（1PB）+长期磁带（20PB）
  - 全部磁盘

- 5分钟观测单个beam数据
- 8k chan
- 0.1ms时间分辨率
- 8k*10000*8bit=80MB/s (160MB/s if 2pol)
- 8k * 3M * 8bit = 24GB
- 第一轮巡天银道面10 deg*70 deg
- 数据量 17000*19*24B ~ 8PB
- 总体需求 ~20PB或更多

最终取值在0.5-2倍范围内变化

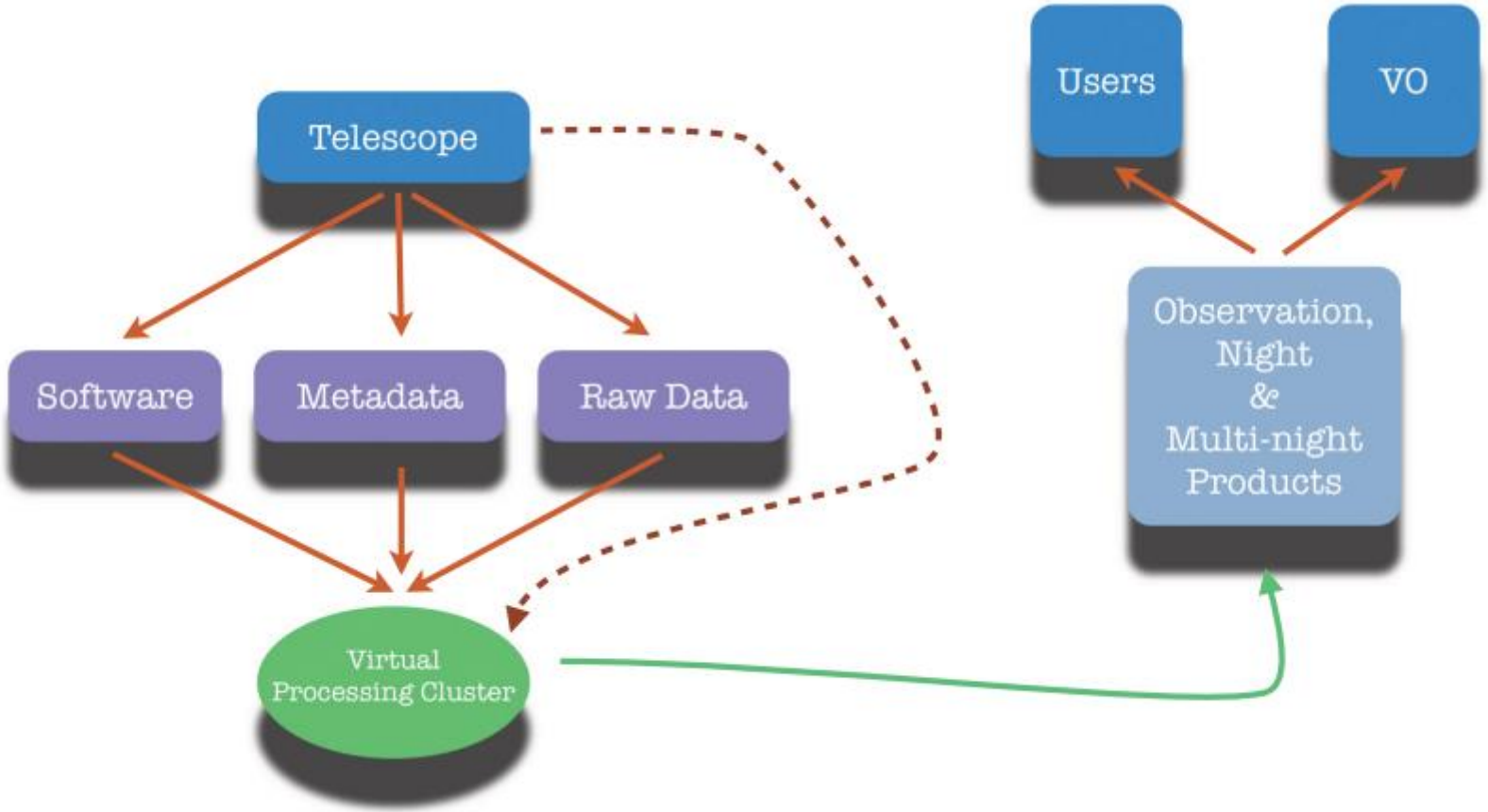
数据传输

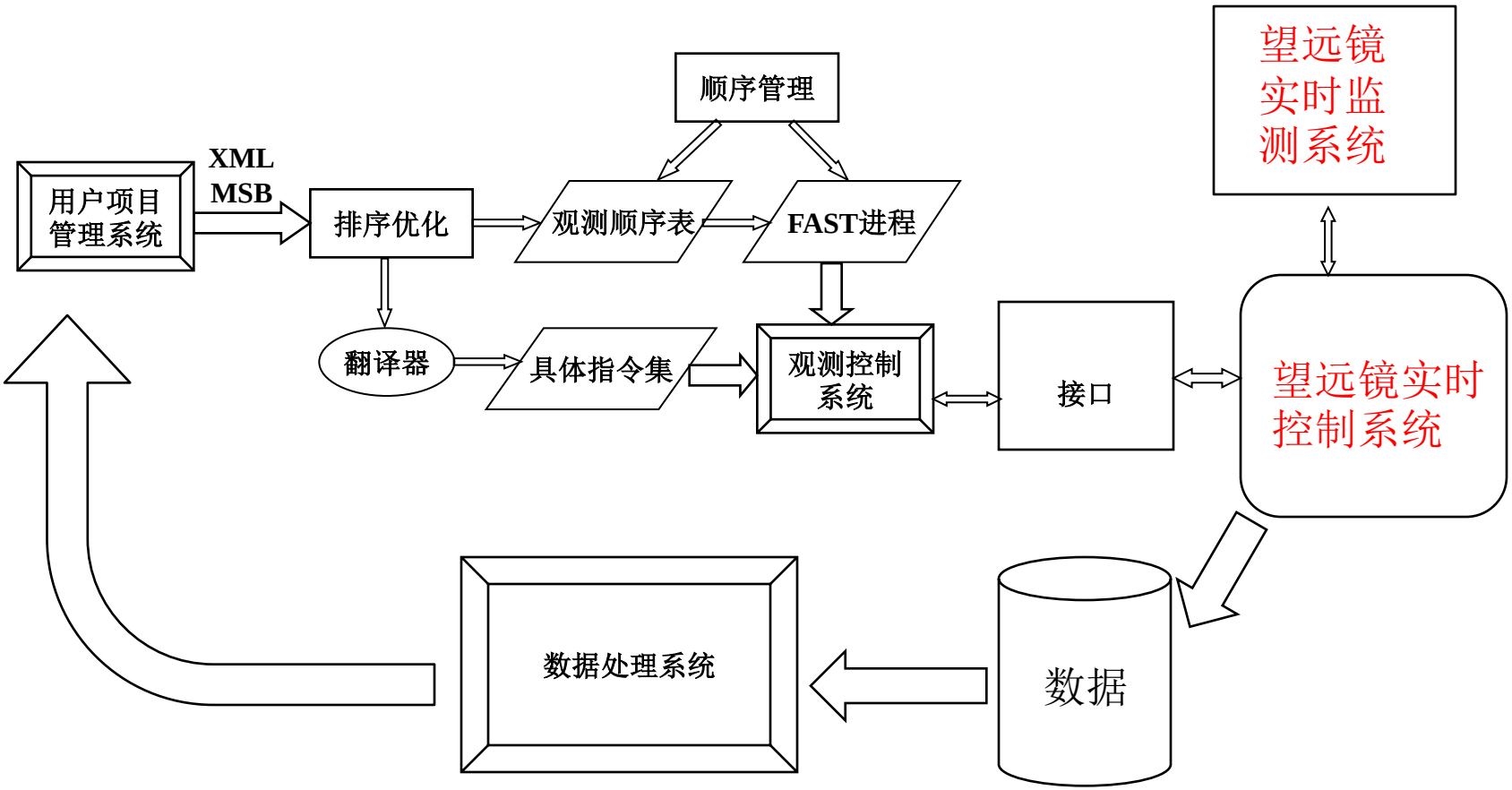
- 大窝凼至贵阳
- 原始数据
  - 每天进行8小时巡天
  - 2GB/s*19=38GB/s ~1PB/8h
- 脉冲星巡天
  - ~40TB/8h (降低~25倍)
- 传输 40TB/24h ~ 0.5GB/s

计算需求

- 主要来自脉冲星搜索
- 脉冲星搜索仍受硬件限制
  - 大量FFT运算
  - IO限制
- 优点：可简单并行化
- FAST基本需求 200T FLOPS
- ~1000 CPU ~500 2-socket server
- Assume GPU~10 CPU (?), then ~ 25 2-GPU server

Data archive system



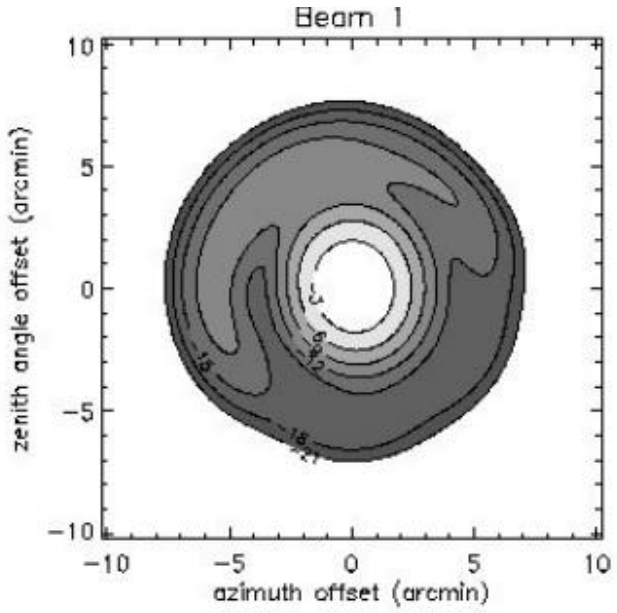

望远镜
实时监
测系统
顺序管理
XML
MSB
用户项目
管理系统
排序优化
观测顺序表
FAST进程
翻译器
具体指令集
观测控制
系统
接口
望远镜实时
控制系统
数据处理系统
数据

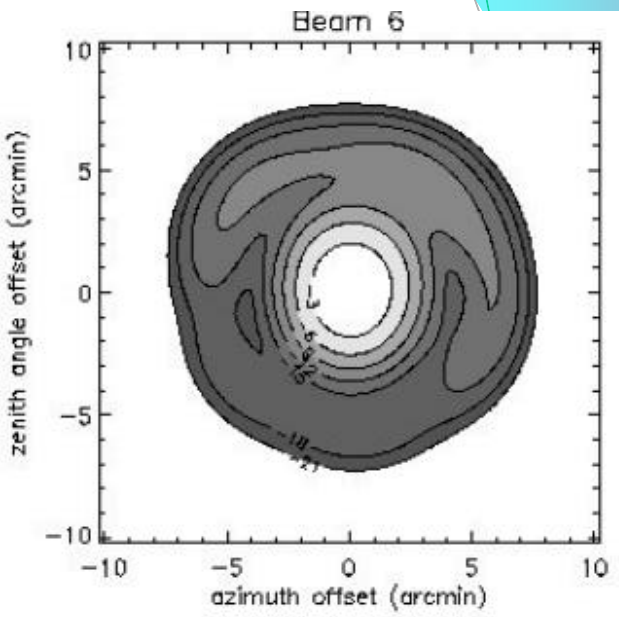
天文数据整合处理

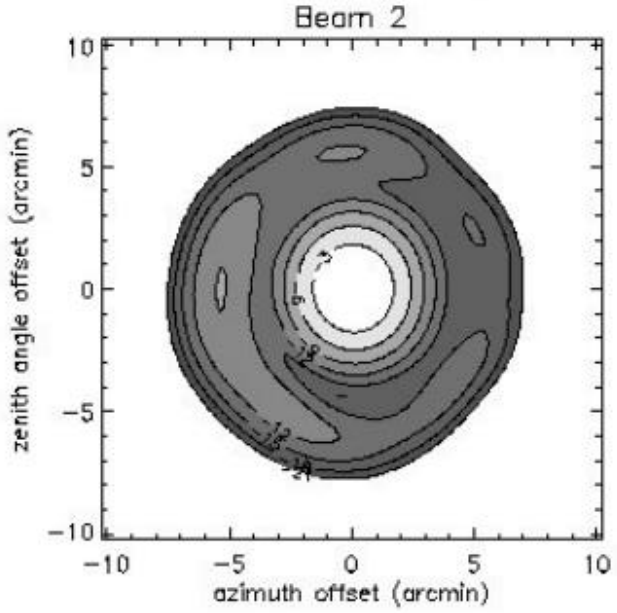
脉冲星数据

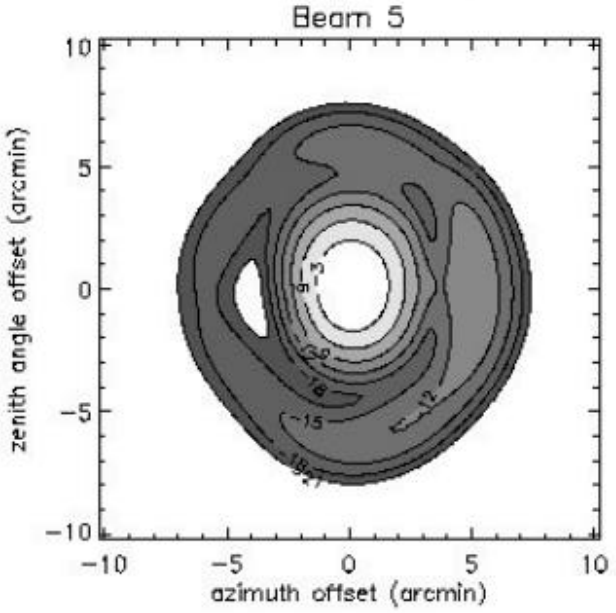
谱线数据： 3 D data cube，数据结构， 大图像处理， 重构 （FFT 卷积 相关等）

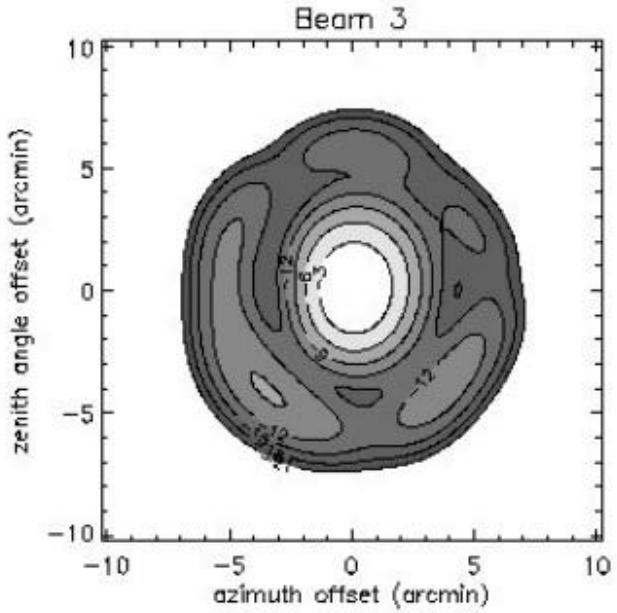
平行计算能力

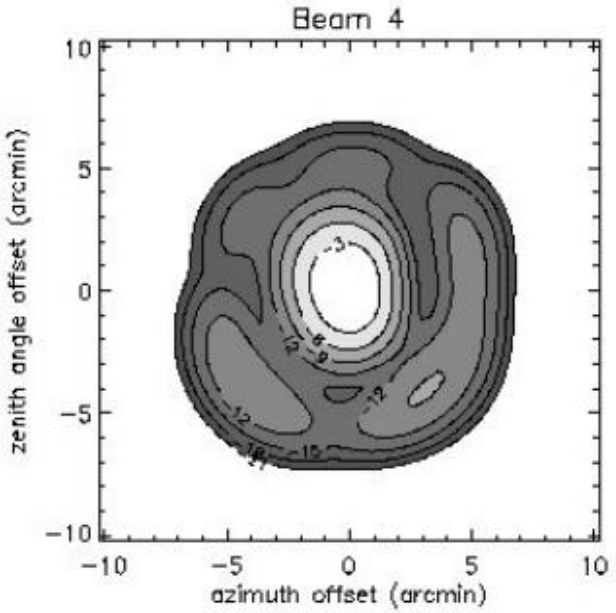

Beam 1
zenith angle offset (arcmin)
azimuth offset (arcmin)


Beam 6
zenith angle offset (arcmin)
azimuth offset (arcmin)

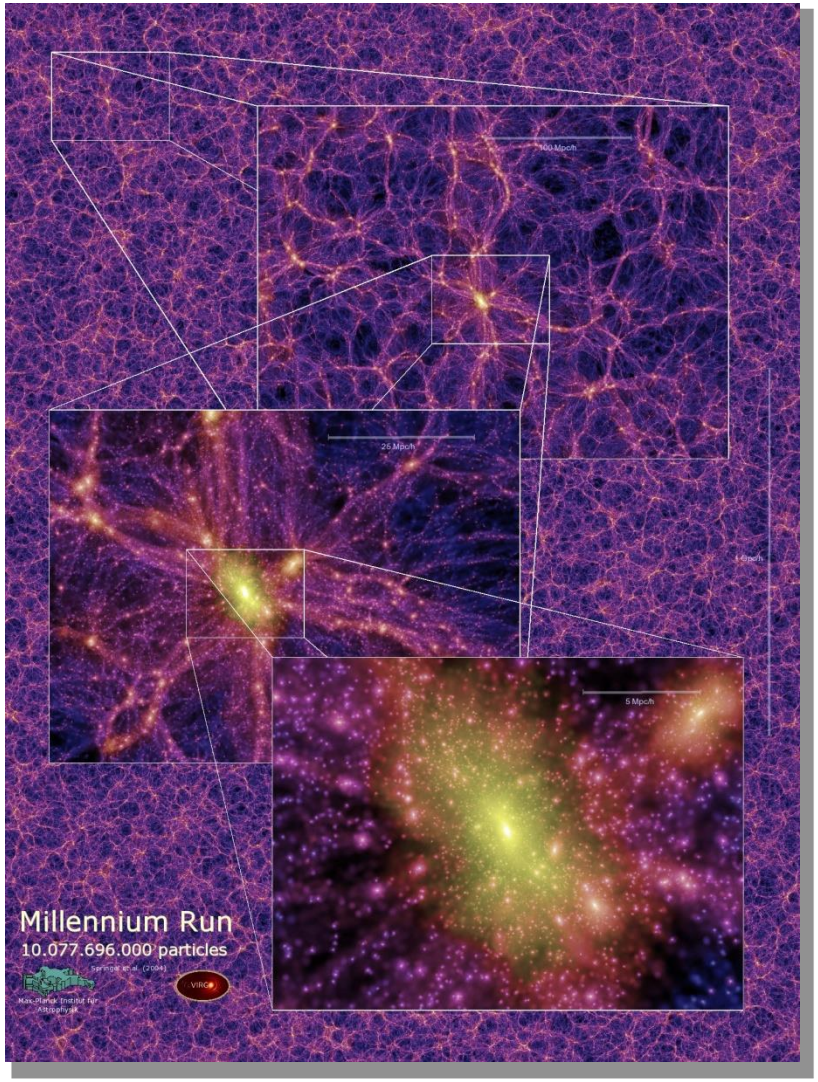

Beam 2
zenith angle offset (arcmin)
azimuth offset (arcmin)

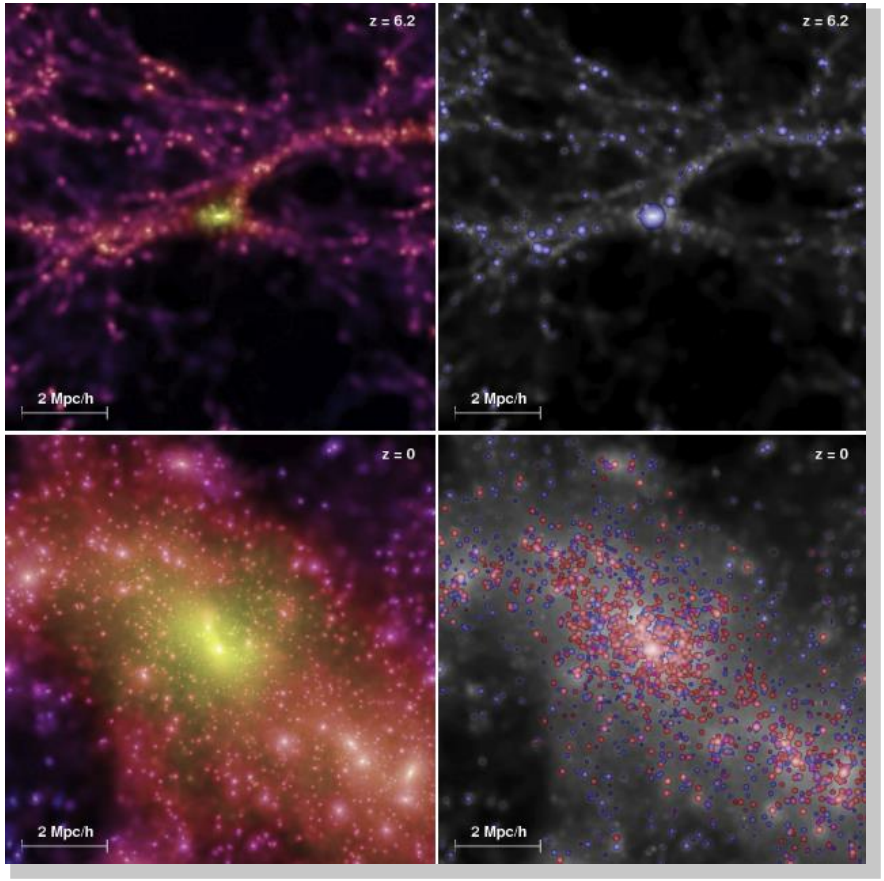

Beam 5
zenith angle offset (arcmin)
azimuth offset (arcmin)


Beam 3
zenith angle offset (arcmin)
azimuth offset (arcmin)


Beam 4
zenith angle offset (arcmin)
azimuth offset (arcmin)

宇宙的大尺度结构

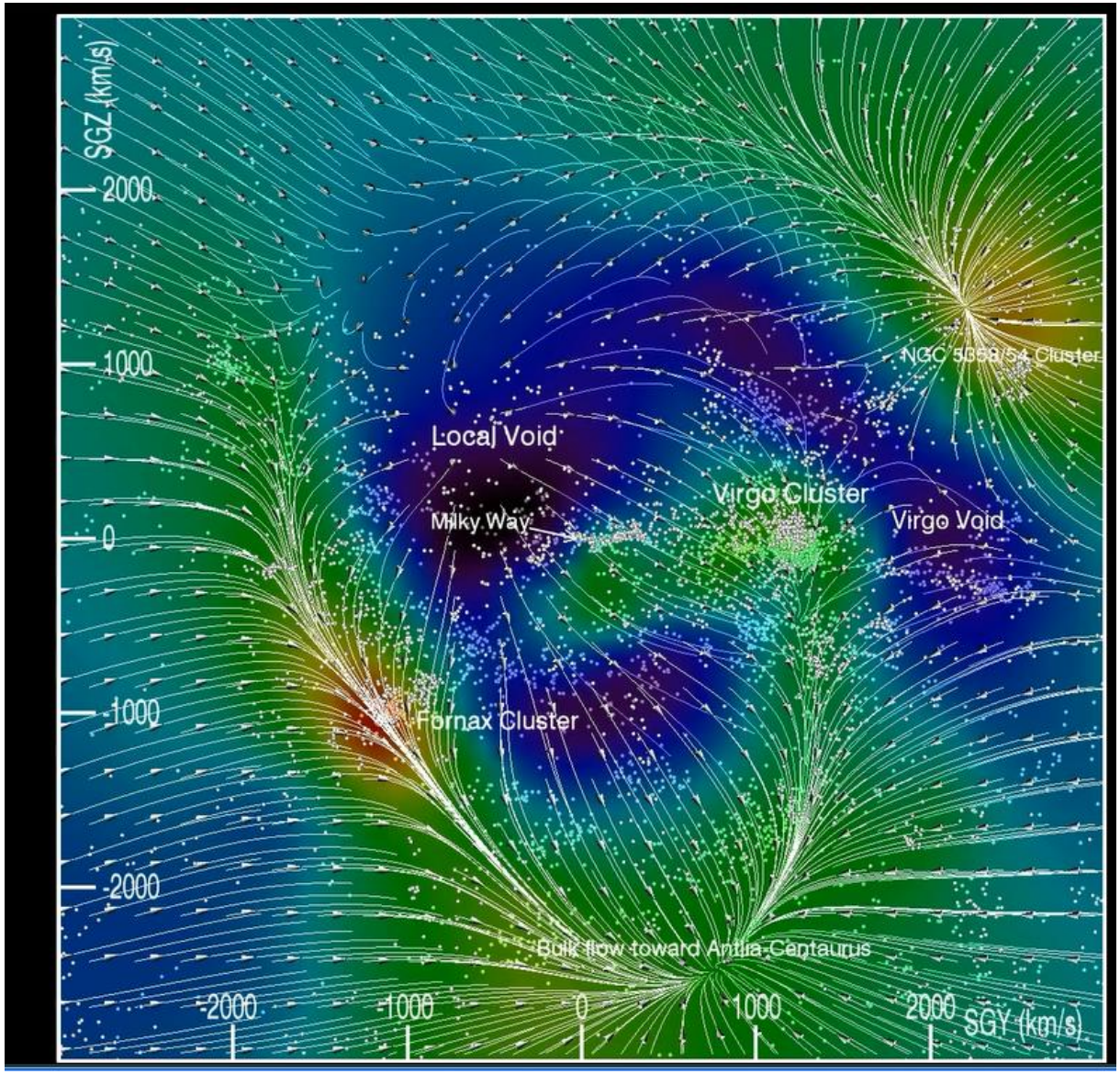


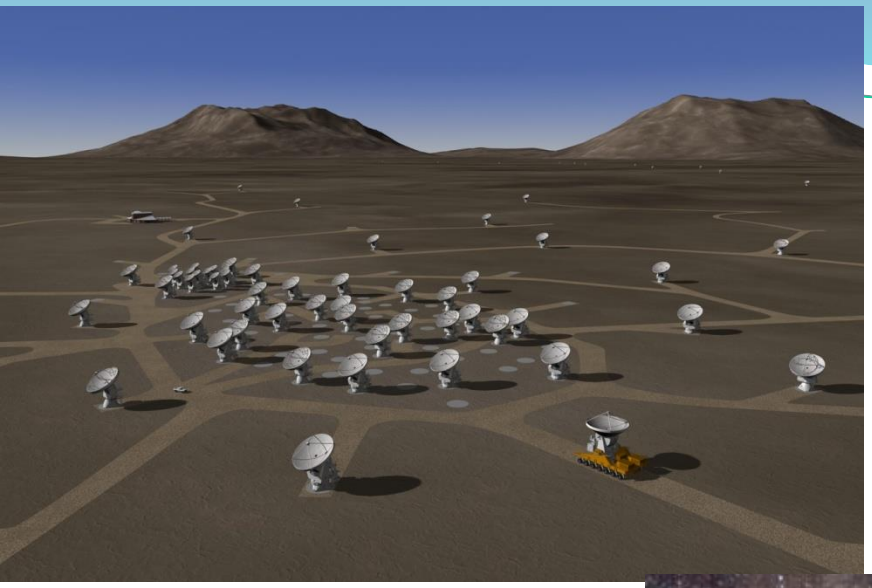


(Springel et al. 2005, Nature)

宇宙邻近速度场分布

Courtois et al. 2013

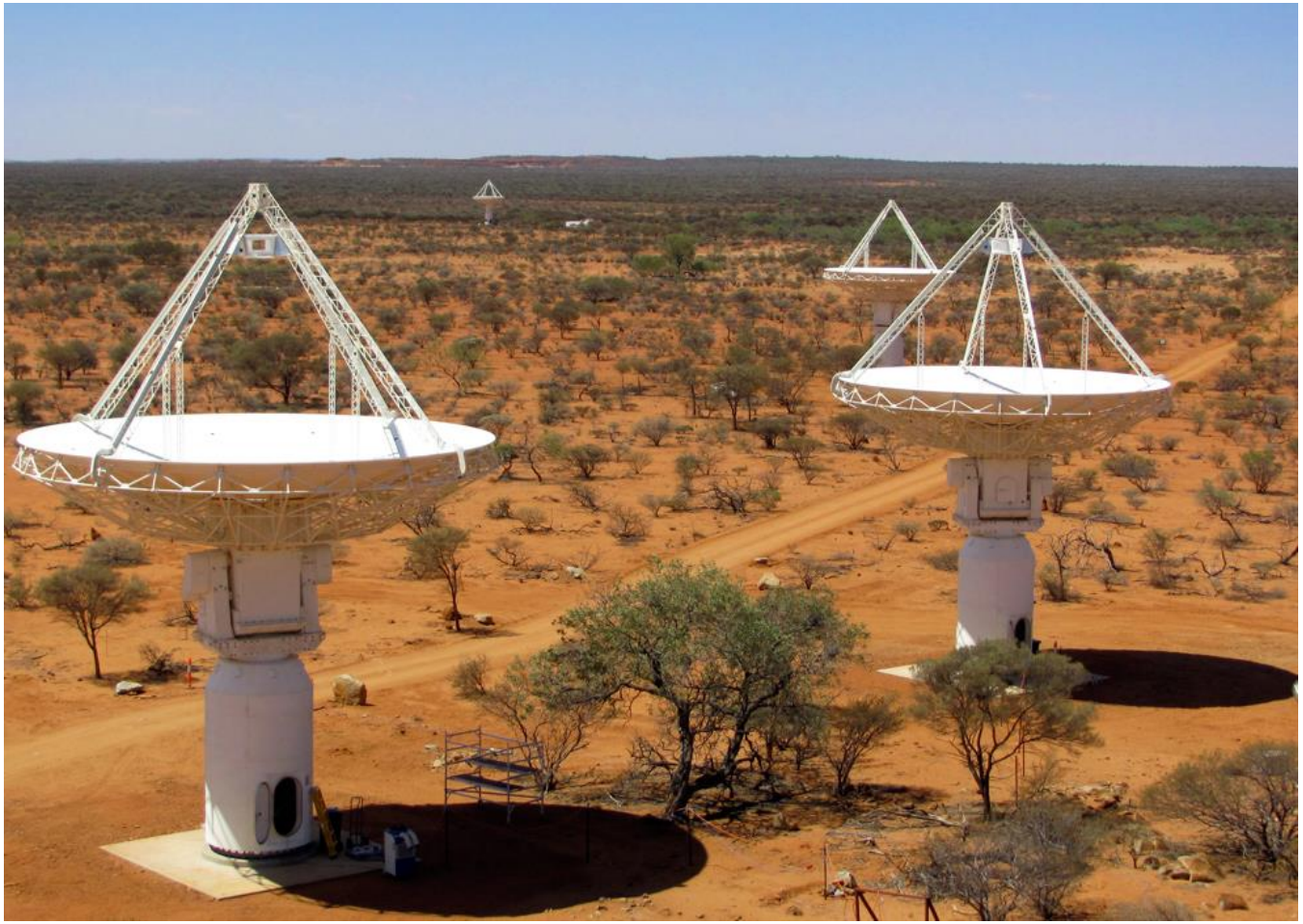




ALMA and JCMT,

亚毫米波望远镜

澳大利亚ASKAP (36x12m)—600PB




SKA
radio

谢谢