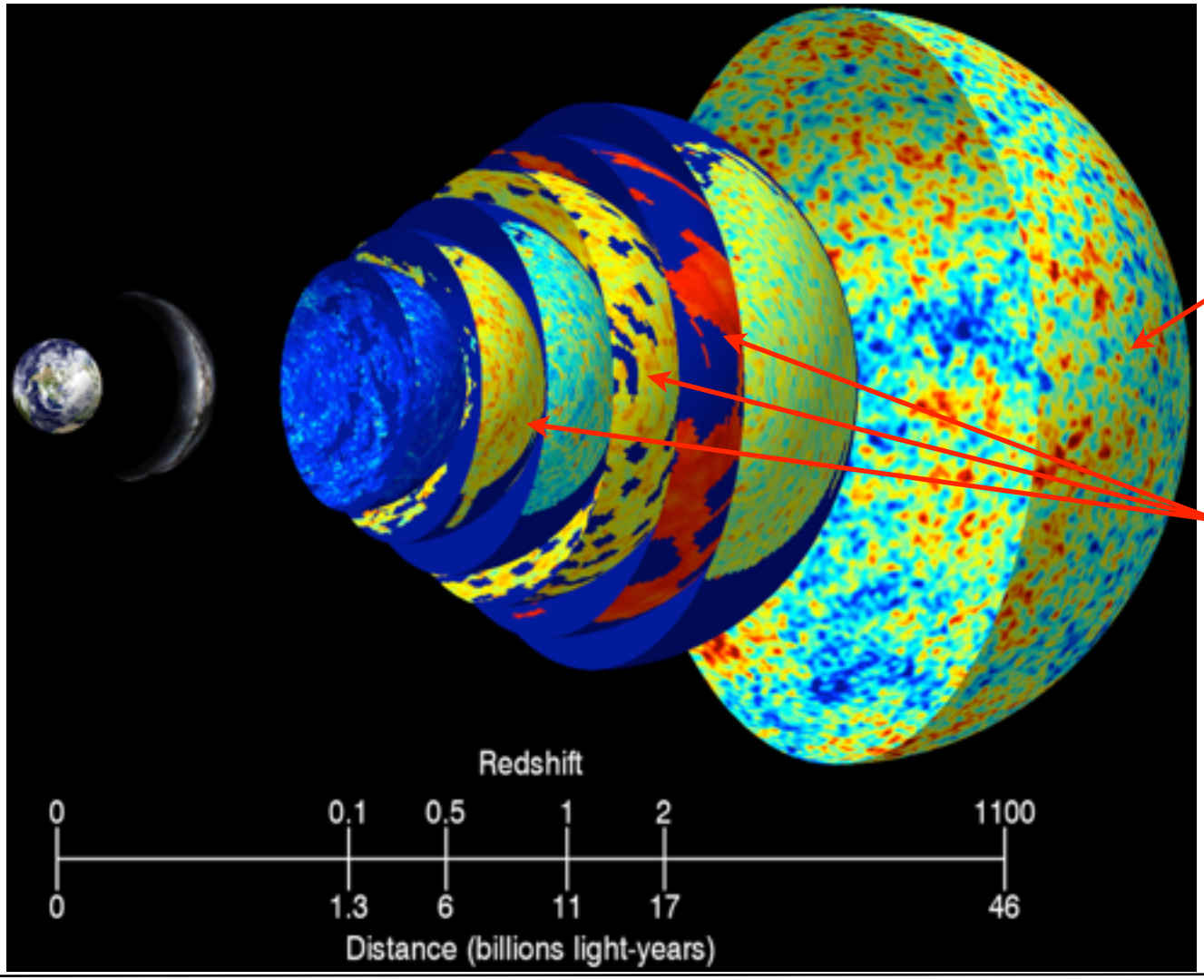
Cosmology on the Beach: Experiment to Cosmology

Monday: Fitting the universe, detectors + imaging maps

Tuesday: Spectroscopic maps

Thursday: SDSS, DESI, Euclid experiments

The age of optical maps



2-D map of CMB

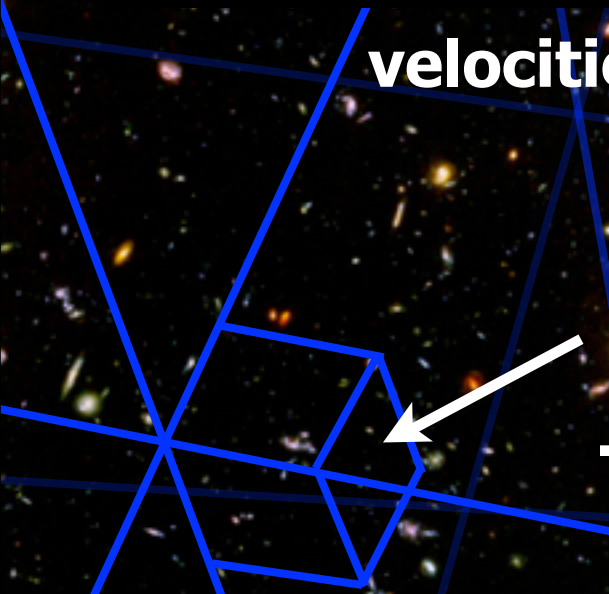
3-D map of galaxies, H gas

Maps of the Universe

Must fit:

* The initial conditions of the early Universe
* The laws of gravity + Dark Energy governing its evolution

$$\text{velocities} = \nabla\Phi$$

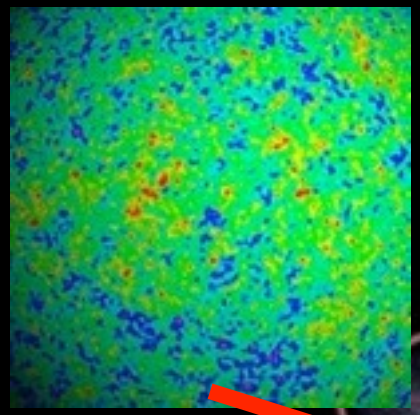


3-D map of "objects" (stars, galaxies, dust)
3-D velocities of those "objects"

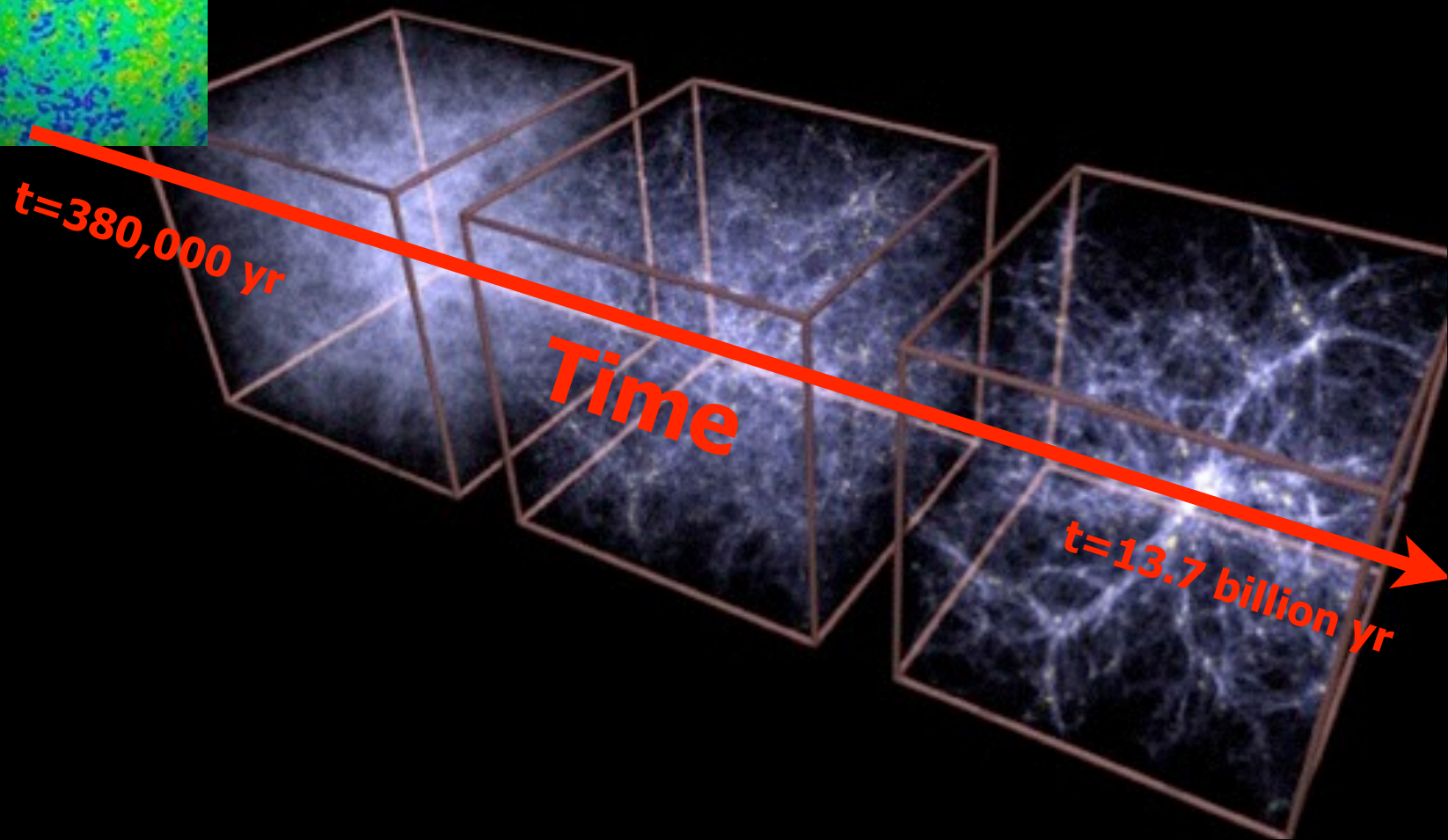
---

6 dimensions

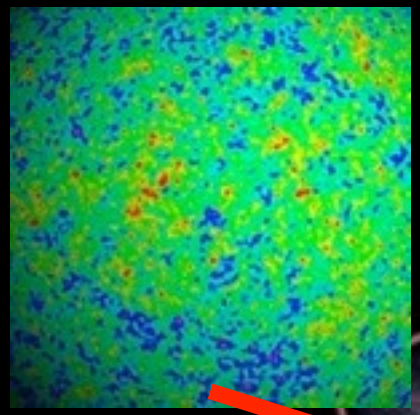
Fitting the Universe



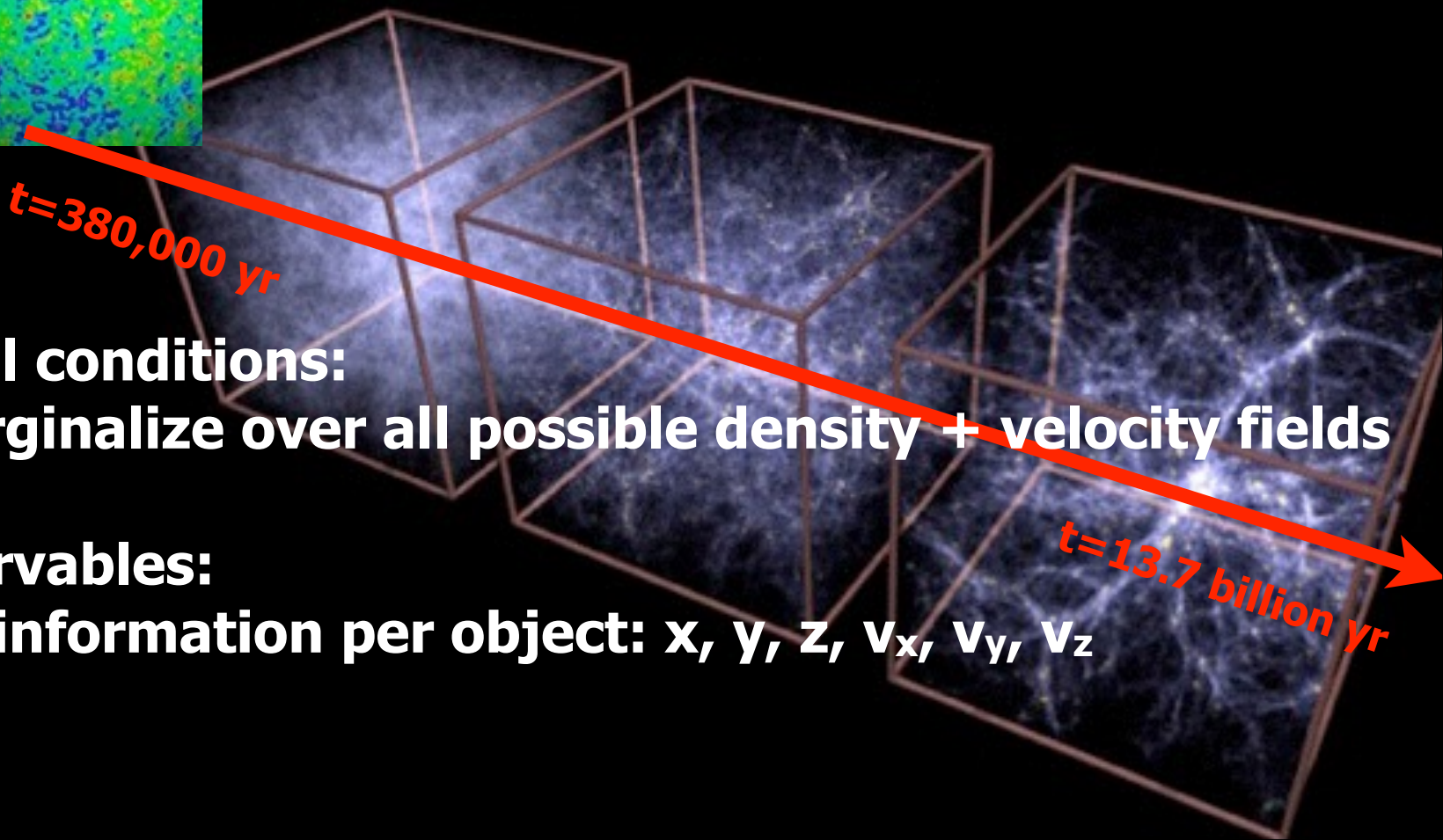
$\mathcal{L}$(Our Universe| initial conditions, forces)



Fitting the Universe



$\mathcal{L}$(Our Universe| initial conditions, forces)



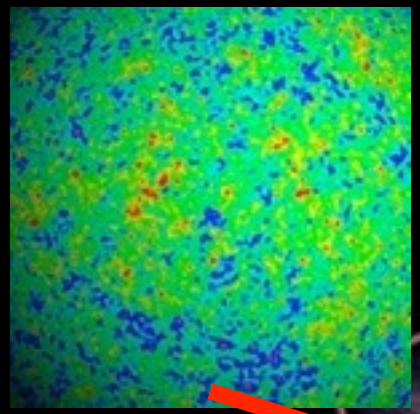
Initial conditions:

Marginalize over all possible density + velocity fields

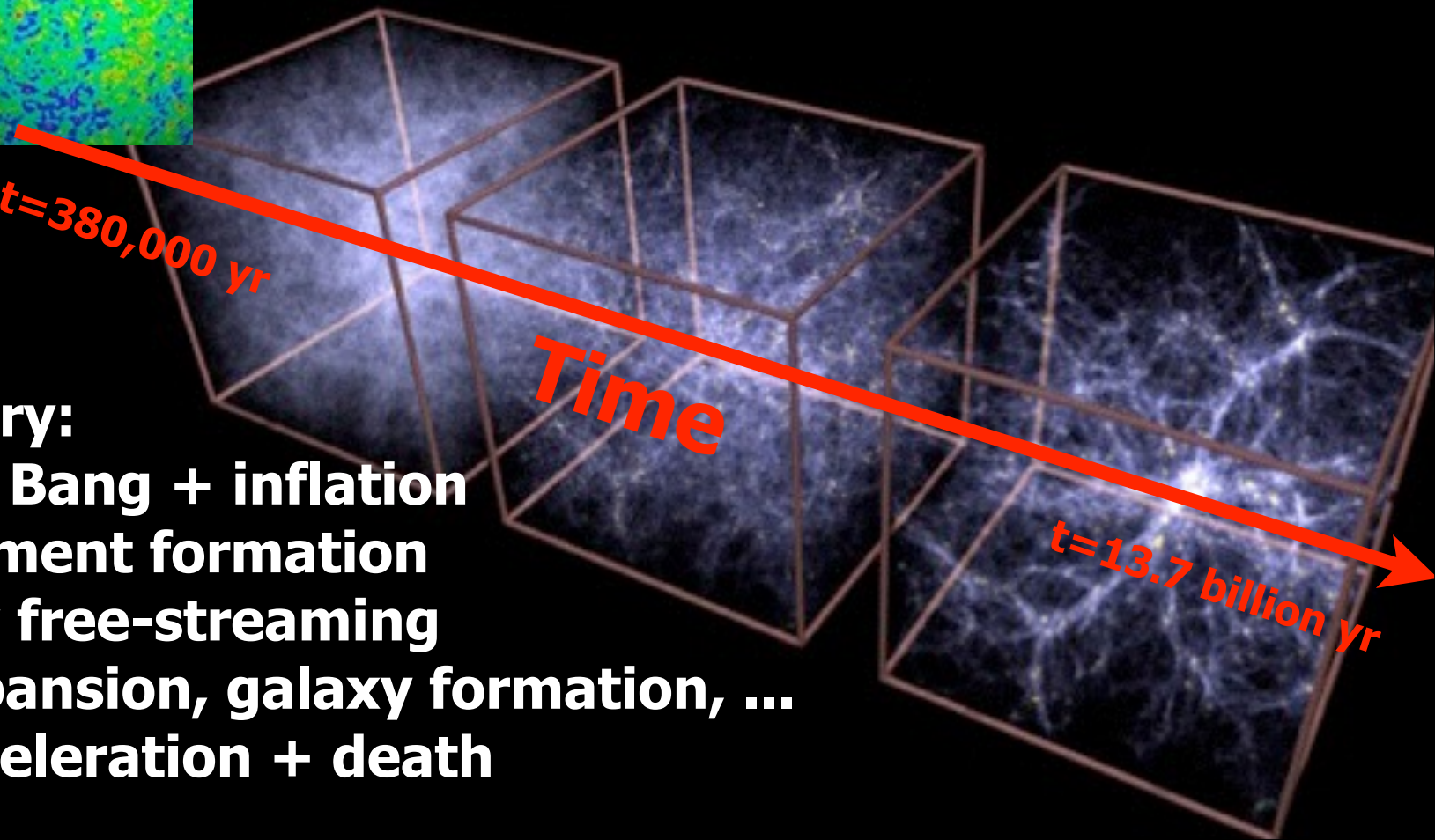
Observables:

6D information per object: x, y, z, v_x, v_y, v_z

Fitting the Universe



$\mathcal{L}$(Our Universe| initial conditions, forces)



History:

- Big Bang + inflation
- Element formation
- ν, γ free-streaming
- Expansion, galaxy formation, ...
- Acceleration + death

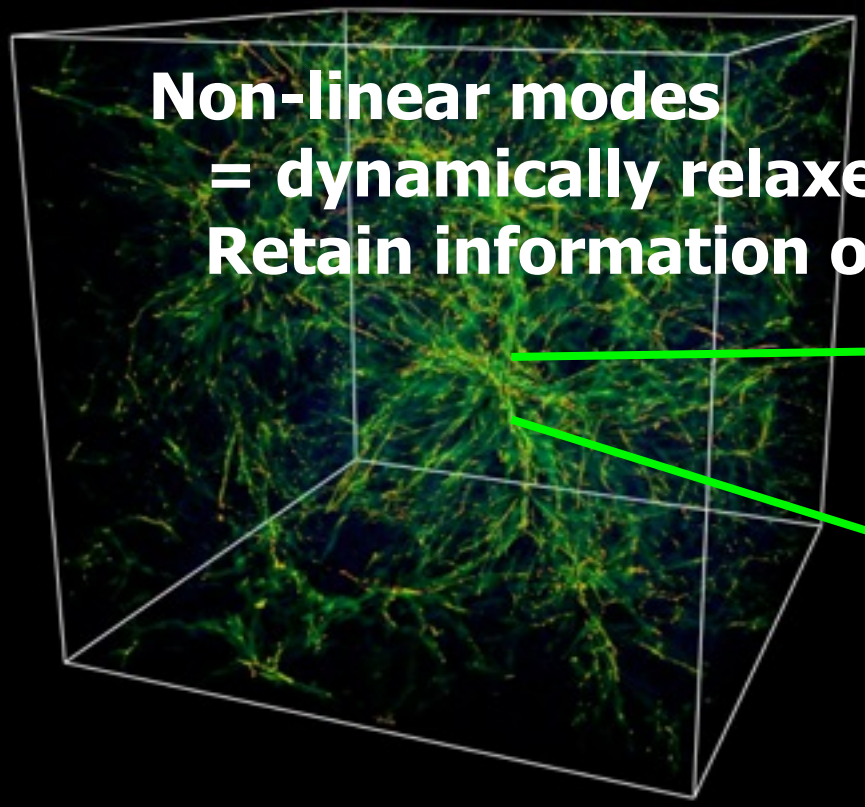
Fitting the Universe

Linear vs. non-linear

Early Universe fluctuations ↔ linear modes

How many linear modes?

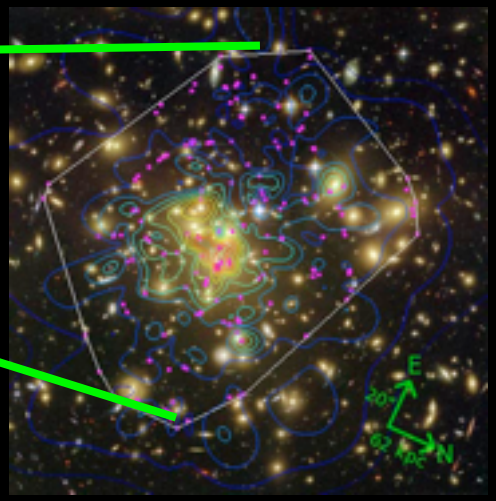
$\sim 4\pi(1000\ \mathrm{Mpc}/8\ \mathrm{Mpc})^3 \approx$ 100 million to z=1



Non-linear modes

= dynamically relaxed systems, e.g. galaxies

Retain information of init. cond. (statistically)



Fitting the Universe

Observables:

6D information per object: (x, y, z, v_x, v_y, v_z)

Imaging surveys

2-D images of the sky == (x, y) only

Constraints?

Observable universe is not infinite

Isotropy in init. conditions

Fitting the Universe

Observables:

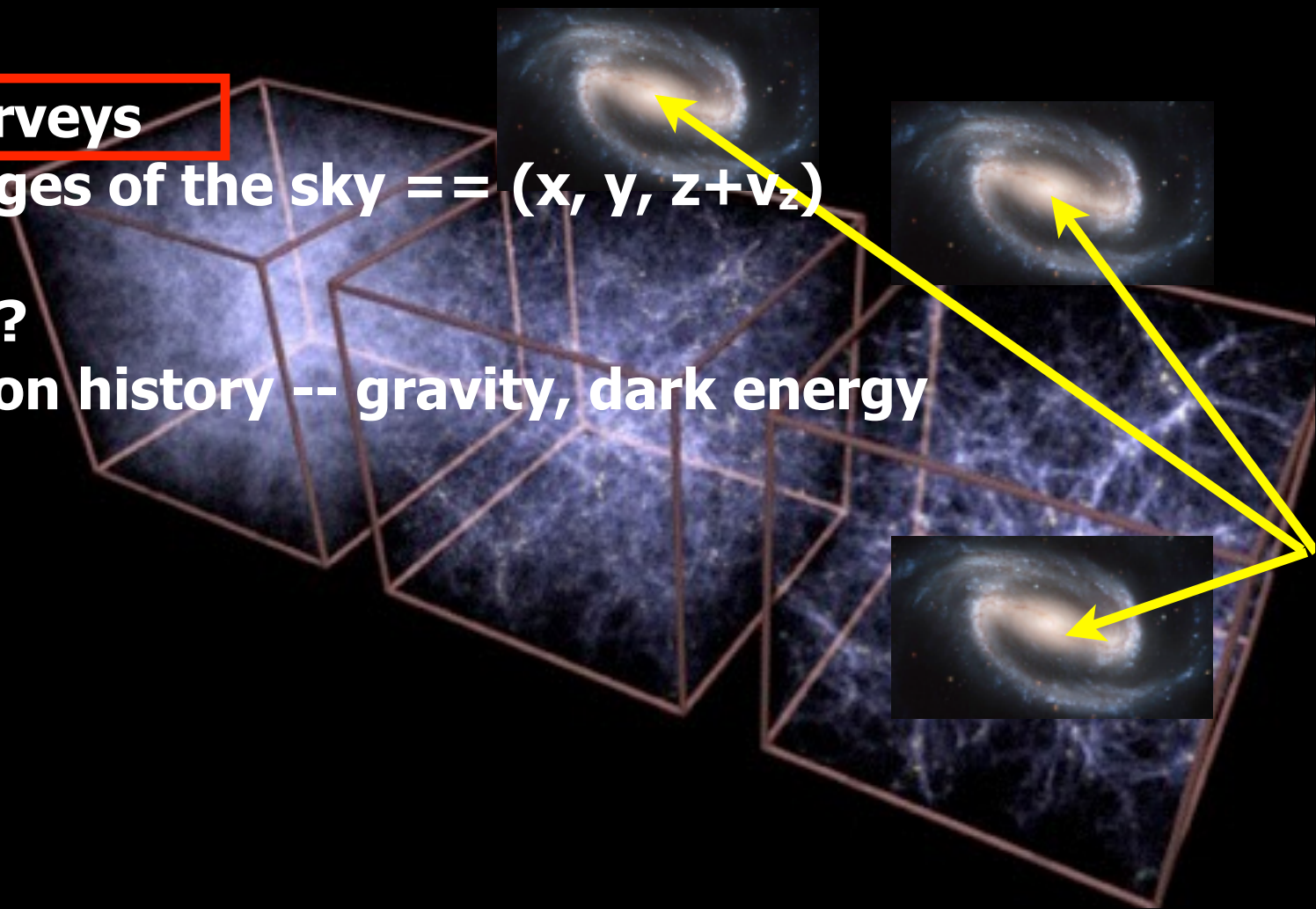
6D information per object: (x, y, z, v_x, v_y, v_z)

Redshift surveys

3-D images of the sky == (x, y, $z+v_z$)

Constraints?

Expansion history -- gravity, dark energy



Fitting the Universe

$\mathcal{L}$(Our Universe| initial conditions, forces)

Observables:

6D information per object: (x, y, z, v_x, v_y, v_z)

Even with arbitrary initial conditions

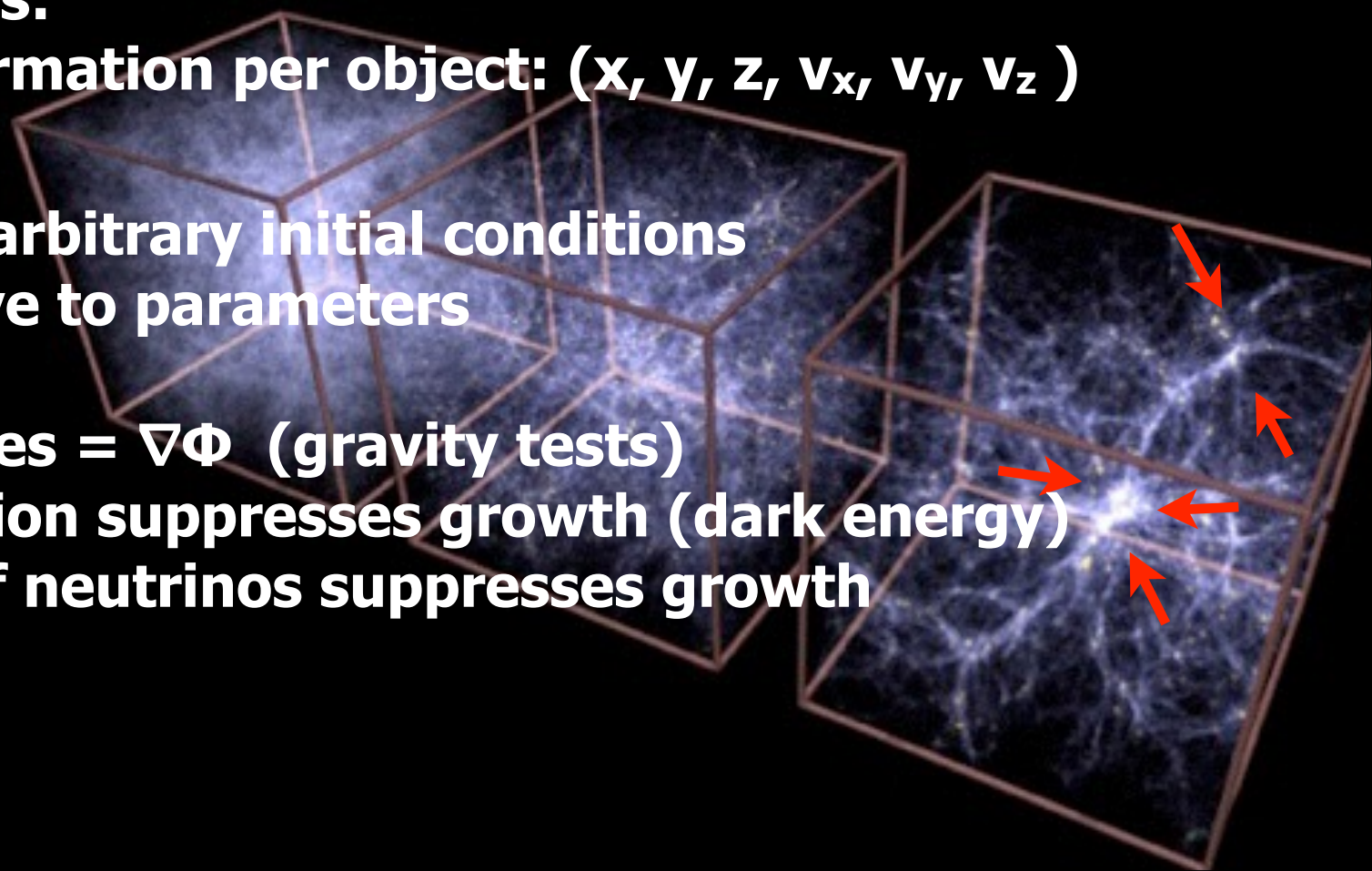
sensitive to parameters

For ex:

velocities = $\nabla\Phi$ (gravity tests)

expansion suppresses growth (dark energy)

mass of neutrinos suppresses growth



Mapping the sky: What's possible?

40,000 deg^2 X $(60\ arcmin/deg)^2$ X $(1000\ galaxies/arcmin^2)$

= 140 billion galaxies **(SDSS is 1 only million!)**

If we could map all of these?

~50 billion LSS modes

Also possible to measure...

Mass of each galaxy from kinematics

Projected foreground mass from galaxy shapes

What's stopping us?

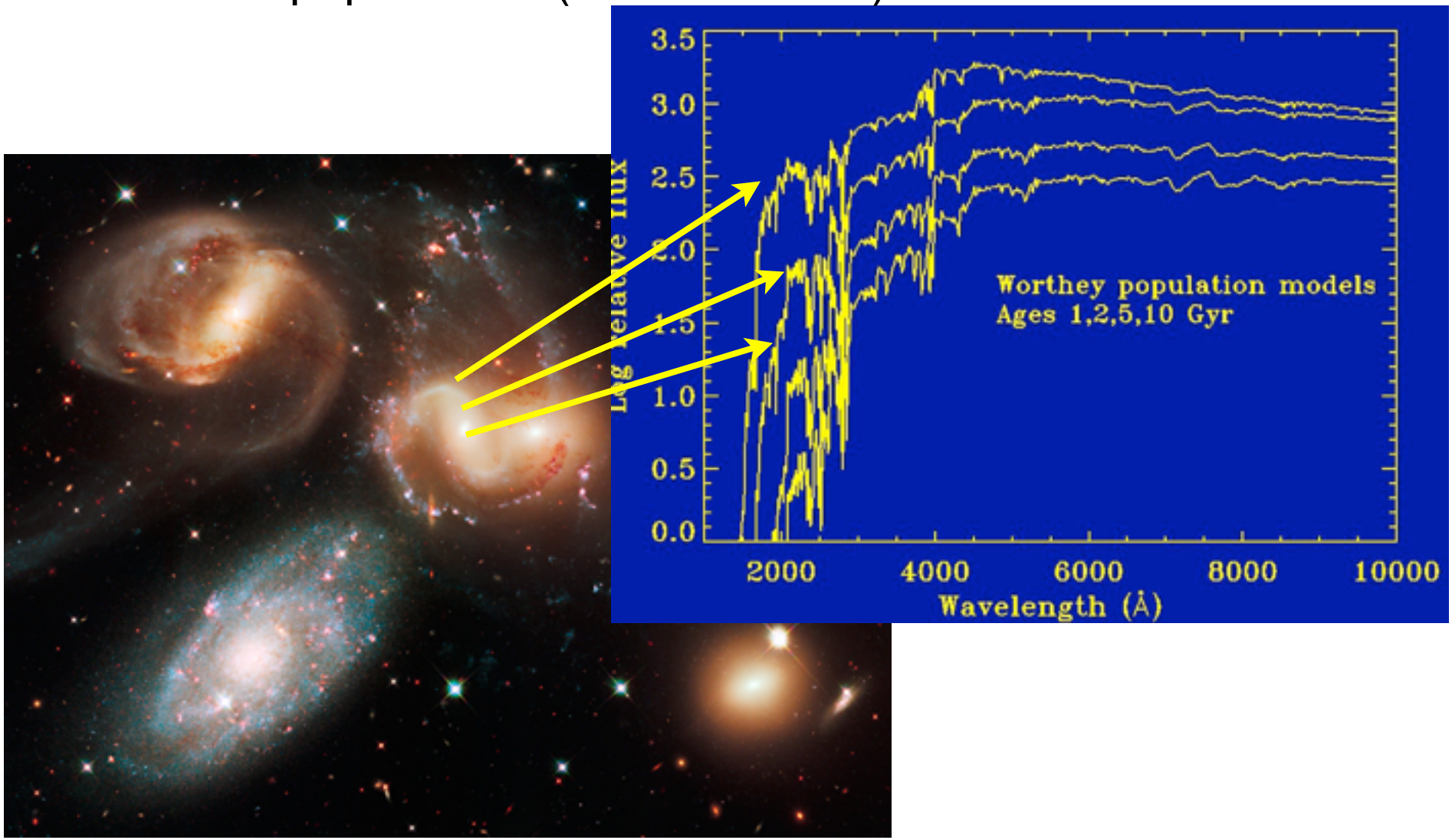
Possible with current technology w/optical cameras

"Only" telescope time :)

HST Ultra-Deep Field
10,000 galaxies / (11 arcmin^2)

Stephan's Quintet, HST

Measure spectrum in different parts of each galaxy
→ Stellar populations (star formation)

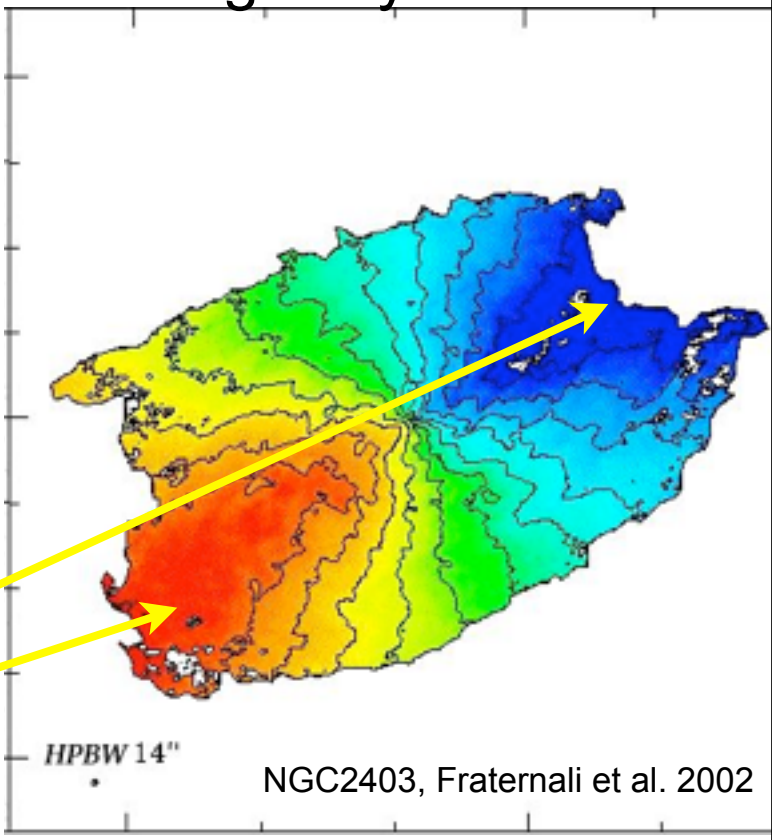


Stephan's Quintet, HST

Measure spectrum in different parts of each galaxy
→ Internal kinematics

Predict intrinsic shape of each galaxy from kinematics!

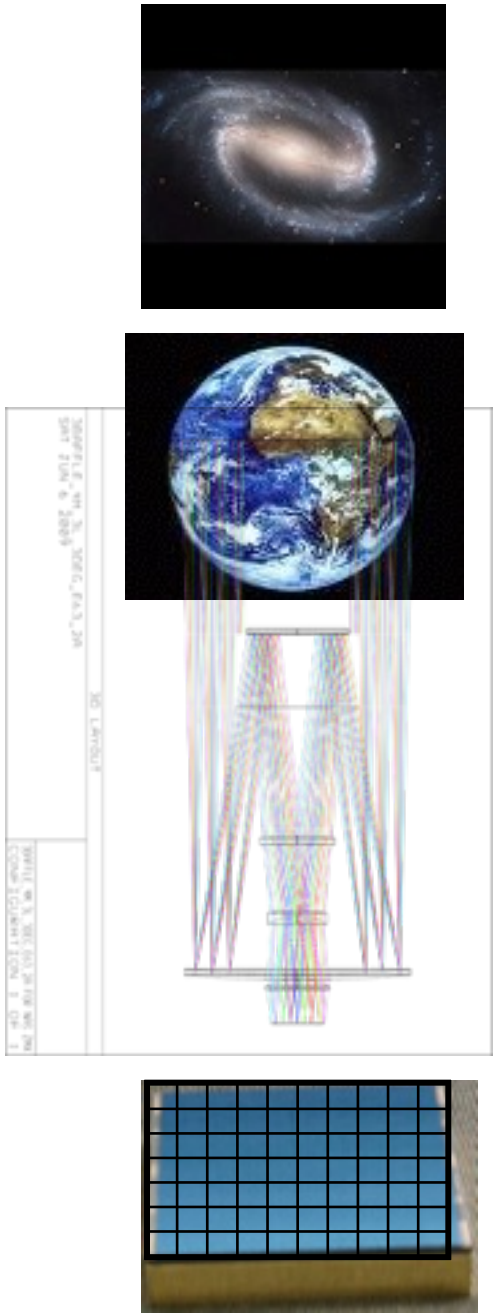
Simulation of the gravitational lens effect created by a massive object that passes in front of the Hubble Deep Field (HDF)
L.L. Christensen, ESA



NGC2403, Fraternali et al. 2002

This example: If no rotation, then elliptical shape due to gravitational lensing!

Telescope images + spectra
not a perfect representation of the sky



Galaxy

(convolve)

Atmospheric seeing

(convolve)

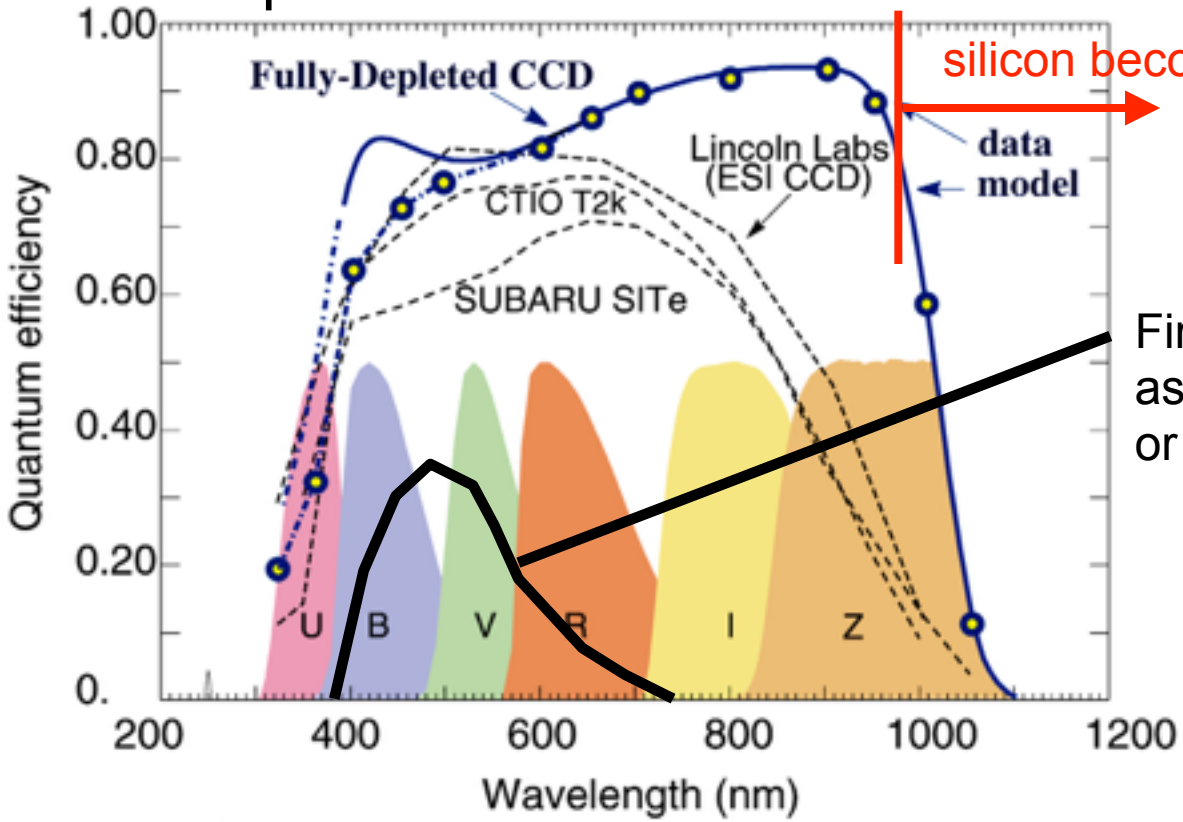
Telescope optics
+ diffraction

(convolve)

Pixelization

Detectors: CCDs in optical

Perfect detectors would have
- 100% Q.E. (quantum efficiency)
- Zero noise
- Independent pixels (no cross-talk)
- Well-sampled

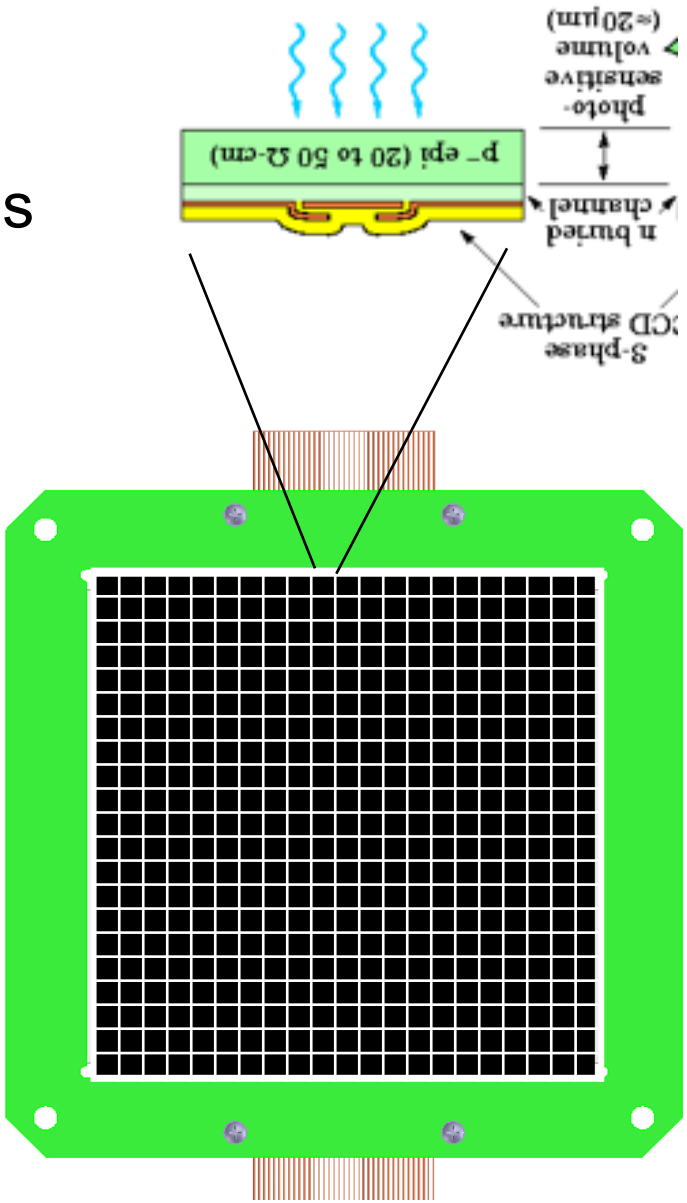


silicon becomes transparent

First CCDs had similar Q.E. as photographic film or the human eye

How do CCDs work?

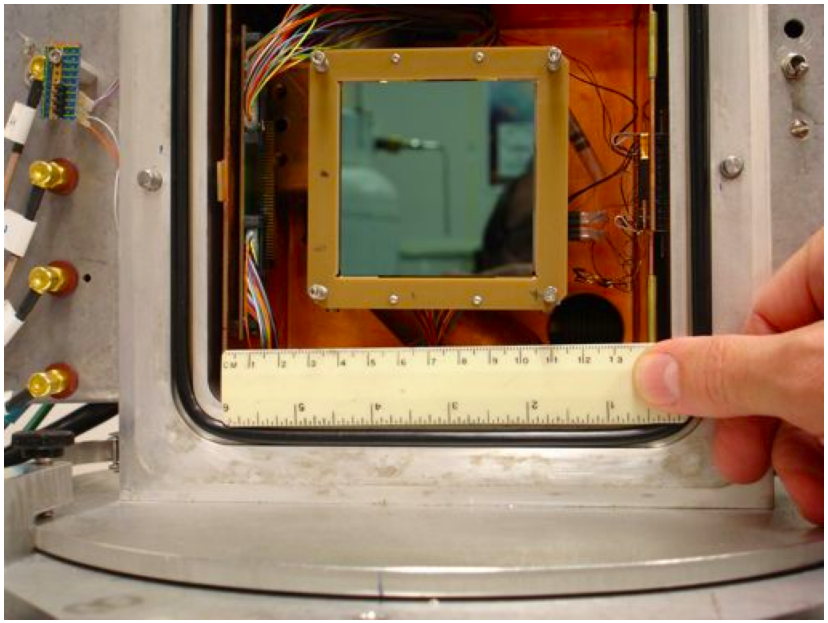
* Silicon detectors, same as computer memory
* Typically $(2048)^2$ or $(4096)^2$ pixels
* Active volume is thin



How do CCDs work?

* Mounted in “dewars”, keeping them in vacuum + cold (-120 C)

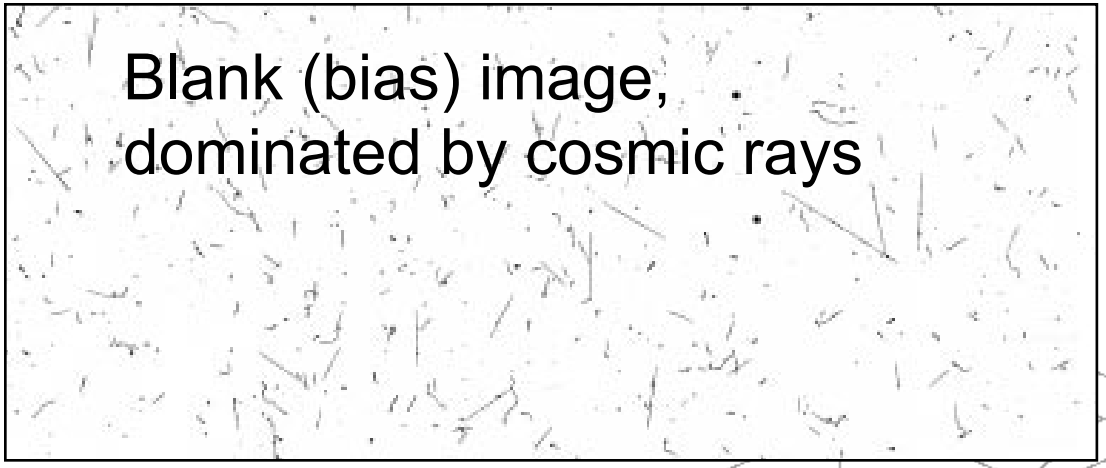
* Discrete counts in e- (digital), converted to current (analog), amplified, converted back to digital with A/D (analog-to-digital)

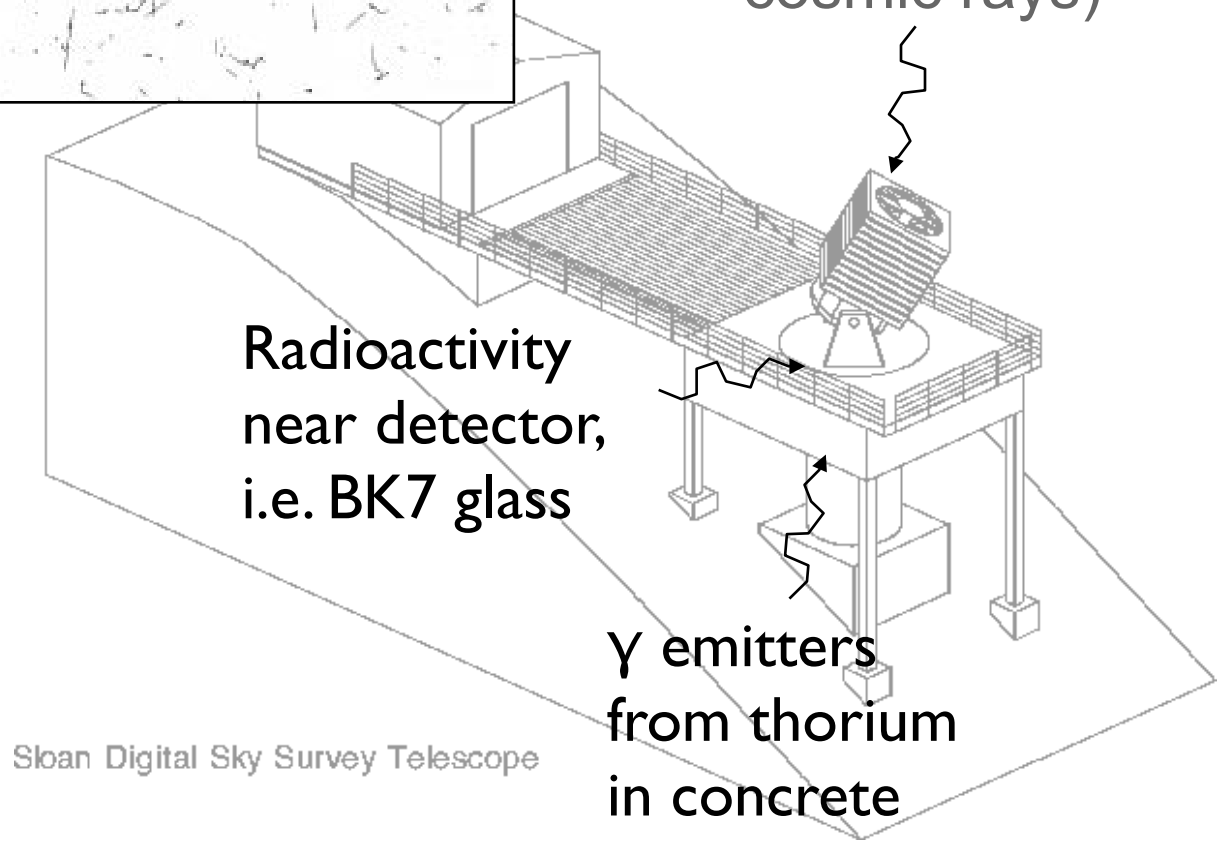


Problems with CCDs:

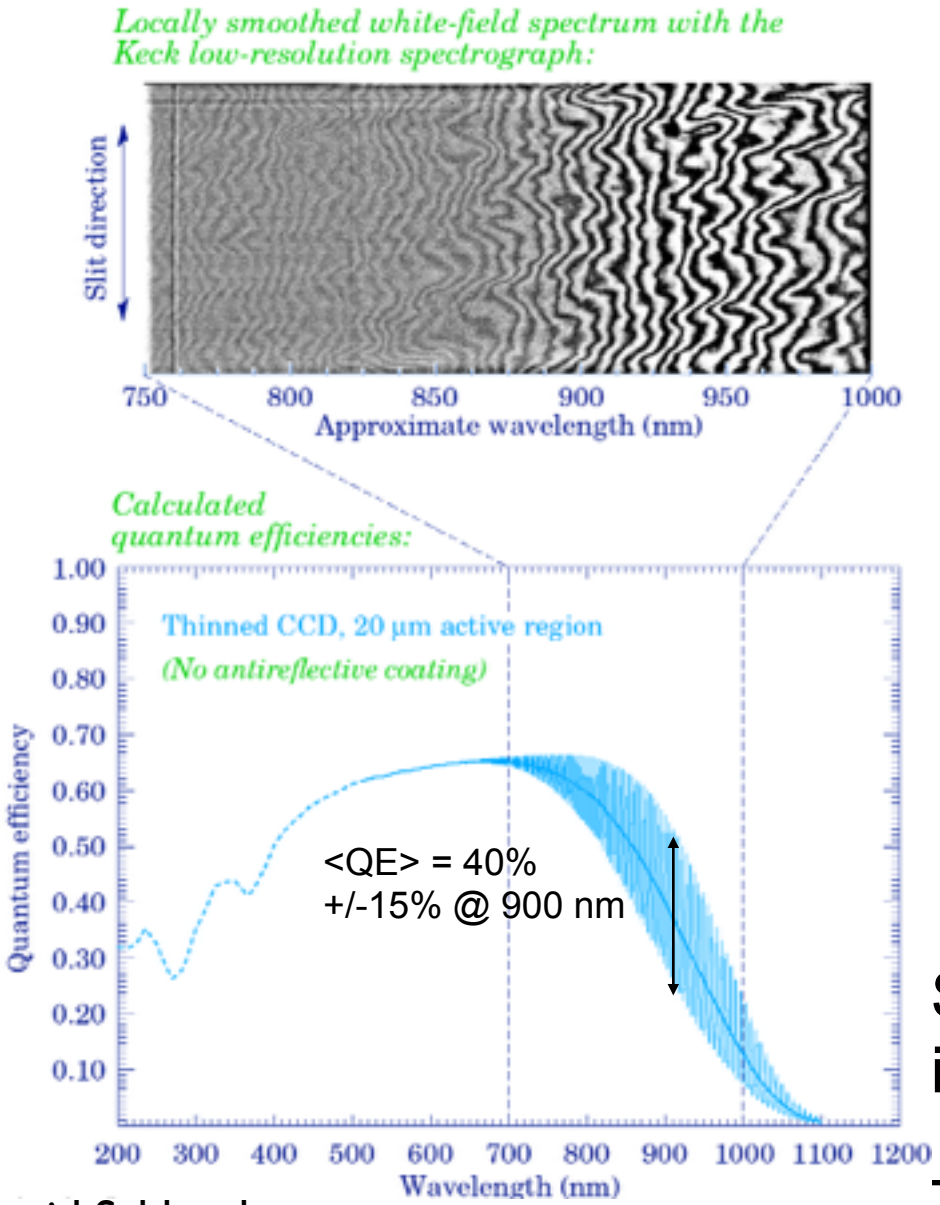
* Readout noise -- minimized by reading slower
* Dark current -- minimized by making colder
* Nonlinearity
* Gain uncertainty, drifts
* Saturation -- both in e- on the silicon, and A/D
* Persistence, especially for saturated pixels
* Cosmic rays
* Charge diffusion
* Fringing -- scattering in CCD surface, wavelength dependent

Problems with CCDs: Cosmic rays

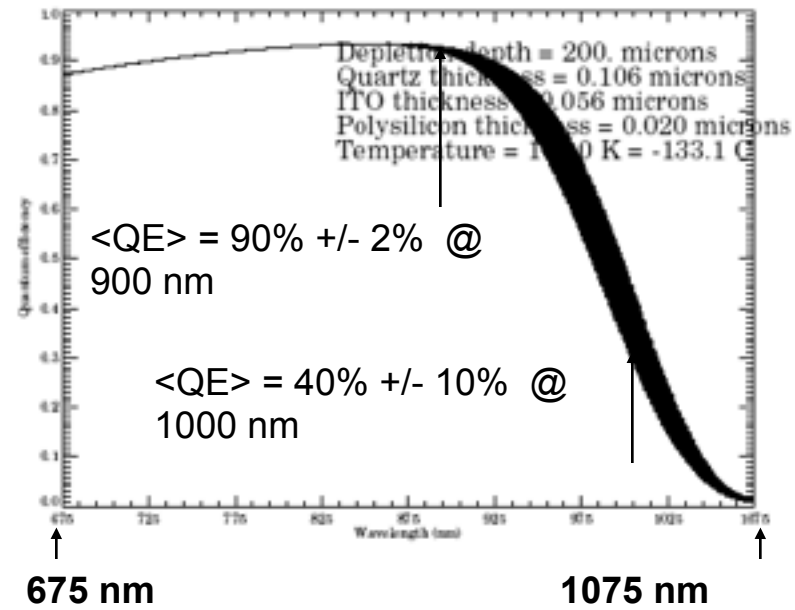




Problems with CCDs: Fringing



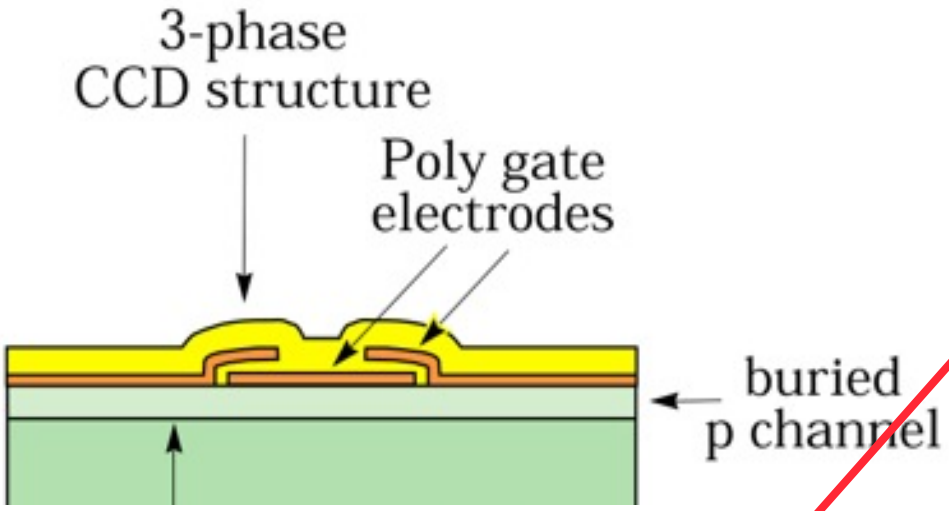
Simulation of fringing in 200 μm thick CCD

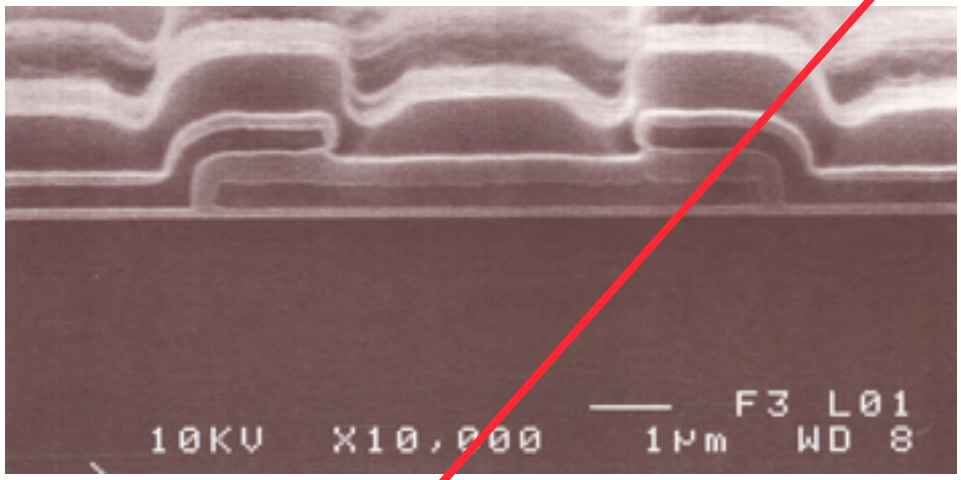


Simulation of fringing in 20 μm thick CCD

Problems with CCDs: internal scattering

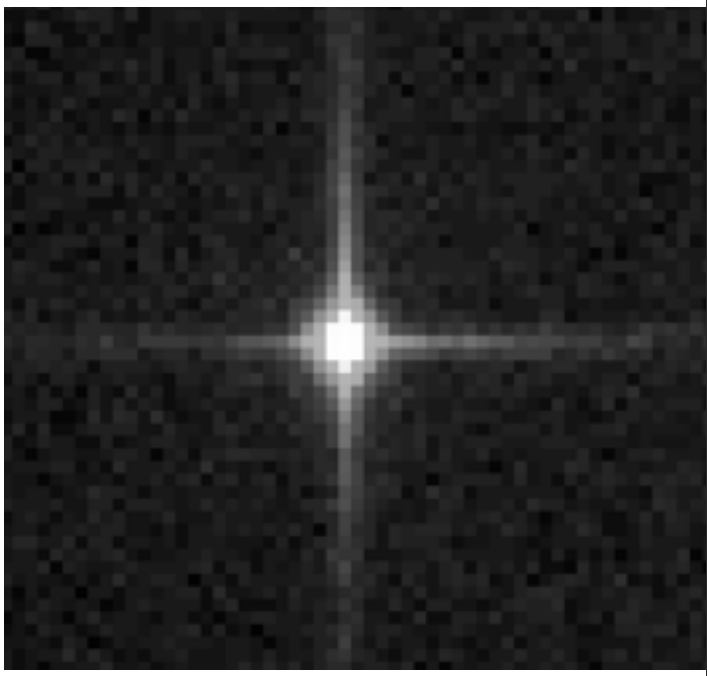
Example: 3-phase overlapping triple polysilicon electrode pixel





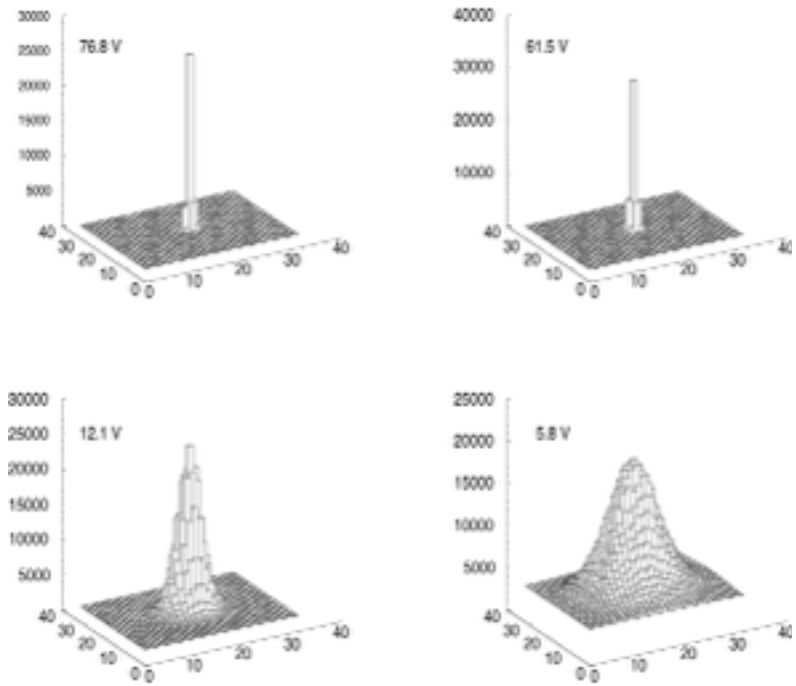
What happens to
this photon?

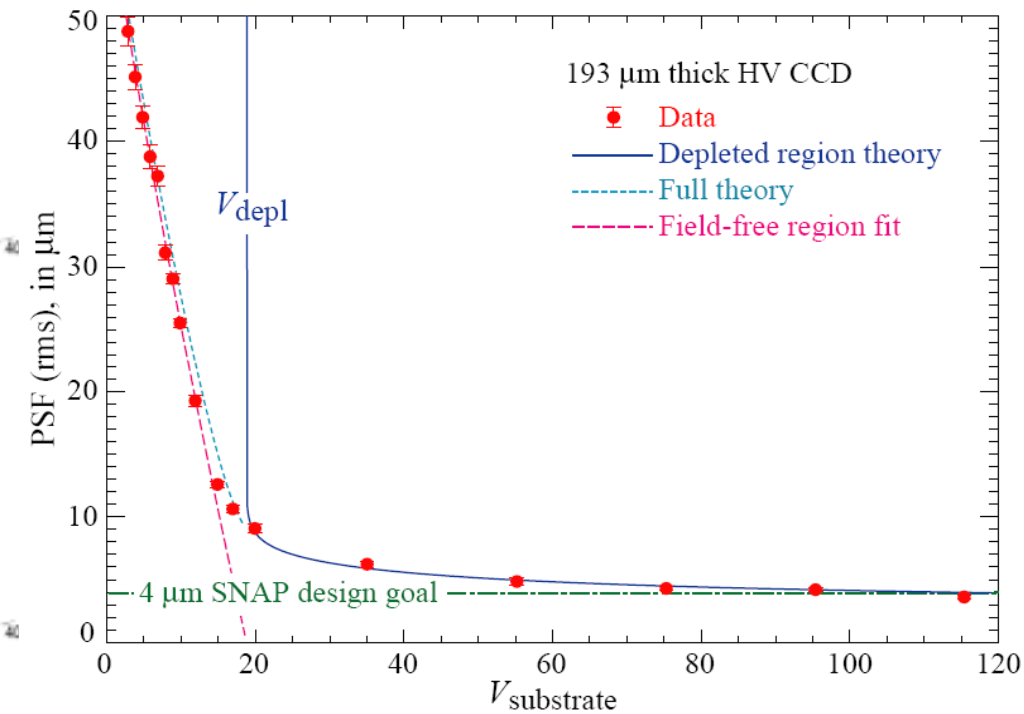
1 μm wavelength photons travel
through CCD, reflect off gates



Problems with CCDs: Charge diffusion

"Dial in" point spread function with bias voltage



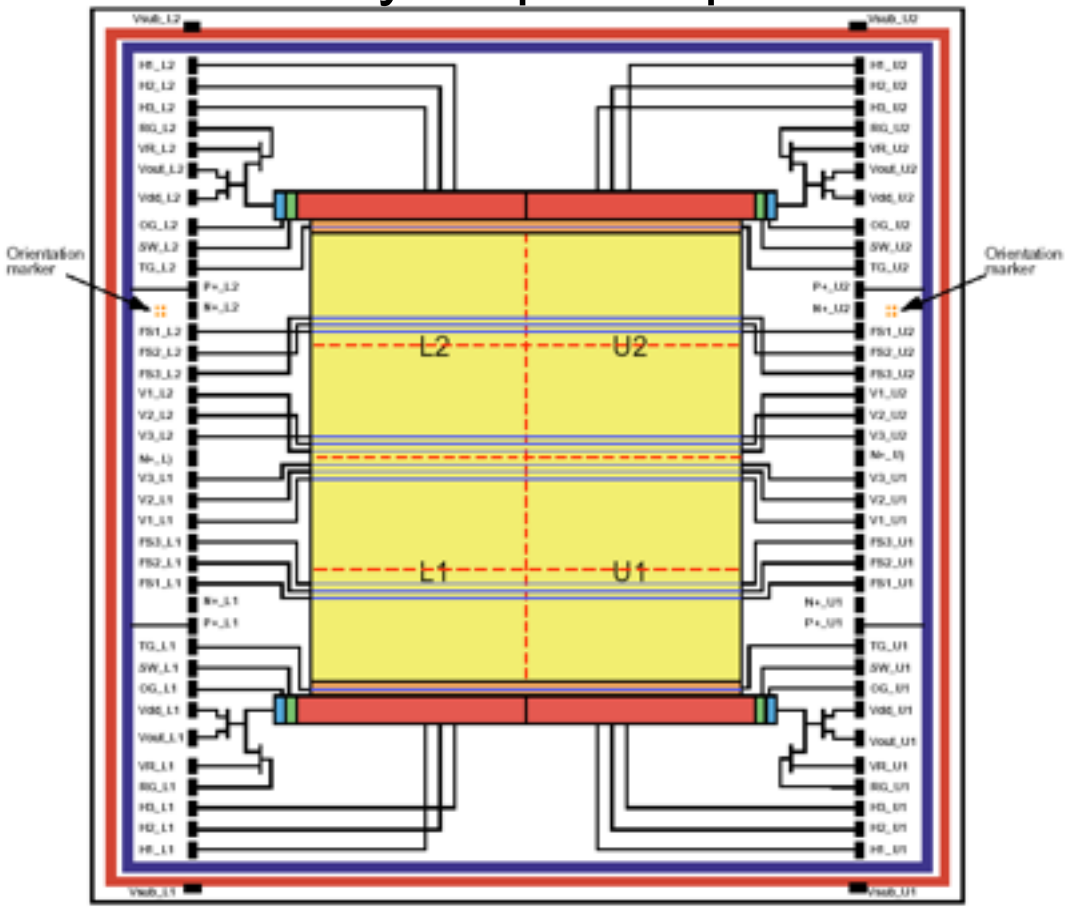


x-y axis: pixel number
z axis; arbitrary units
1100 x 800 back-illuminated CCD, 15 μm pixels

"Improved Spatial Resolution in Thick, Fully Depleted CCDs with Enhanced Red Sensitivity"
by J. A. Fairfield, presented at 2005 IEEE/NSS meeting

Problems with CCDs: Non-linearity, gain, A/D sat.

These problems are function of the amplifiers
Modern CCDs have many amplifers per CCD

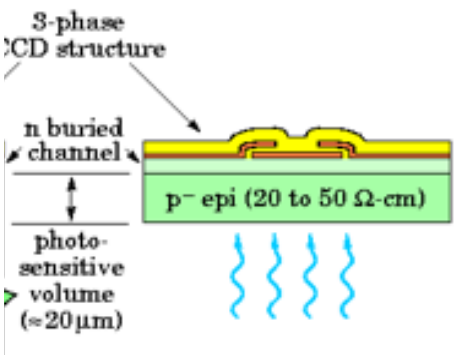


Frontside view

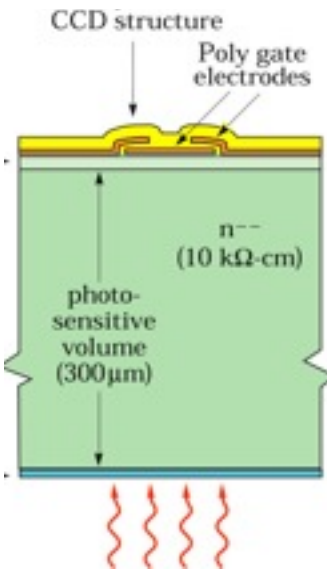
Figure 1. CCD pads and signals. Each end of the serial registers has 7 extended pixels. Serial channels are continuous and can be read at either end or both. Parallel channels are continuous and can be read at either end or both.

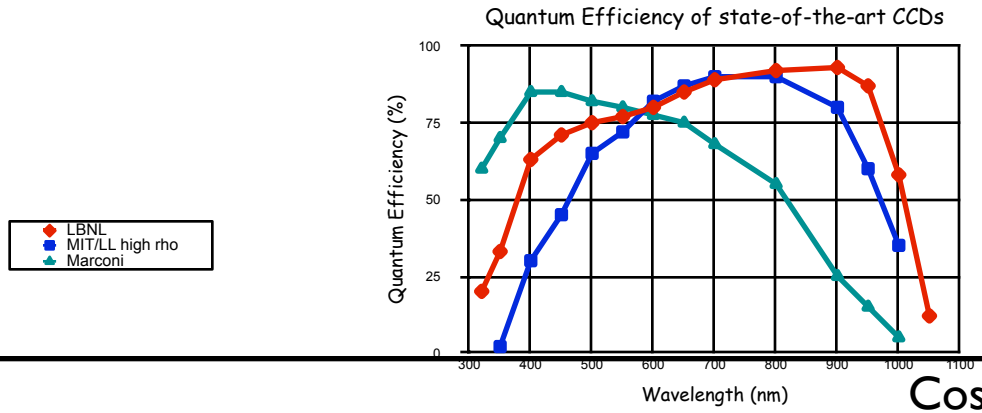
New red-sensitive CCDs:

- "Standard" CCDs used in astronomy are thinned, back-illuminated n-channel devices with ~10-20 um epitaxial depletion region



- LBNL CCDs are thick, back-illuminated fully-depleted p-channel devices
  - higher QE up to 1000 nm
  - no fringing at long wavelengths
  - small, controllable point spread function (PSF) through application of high bias voltage
  - p-channel has improved radiation tolerance
  - disadvantages: more sensitive to cosmic rays, background radiation

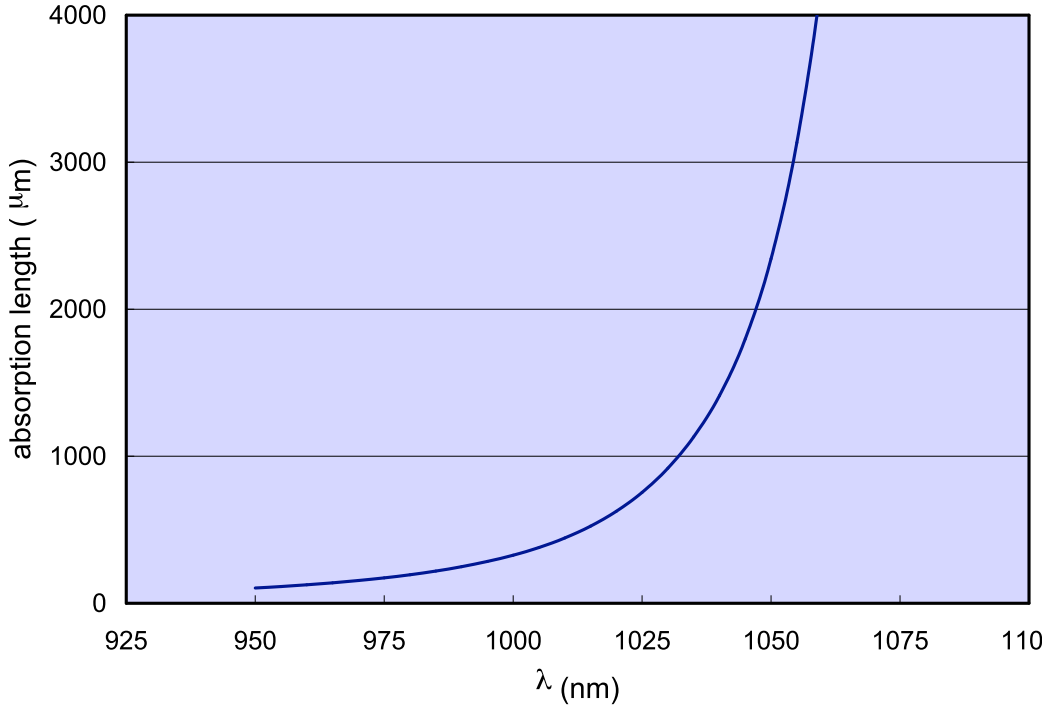




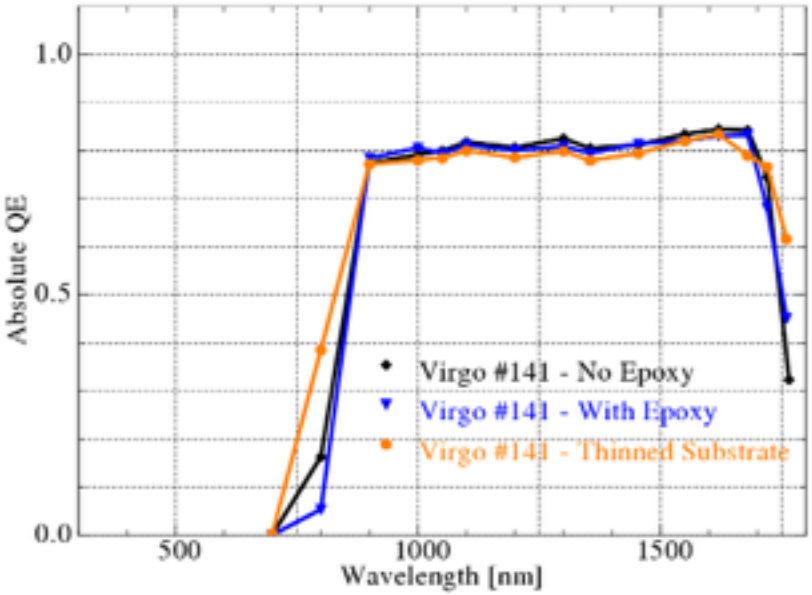
Infrared detectors:

CCDs become transparent at >1 micron

Silicon absorption length



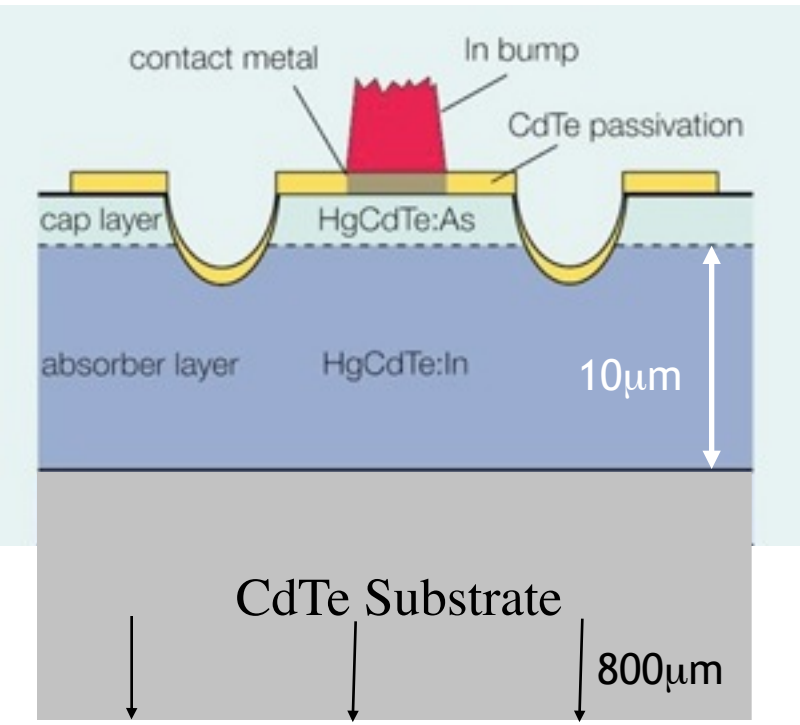
HgCd can see these photons

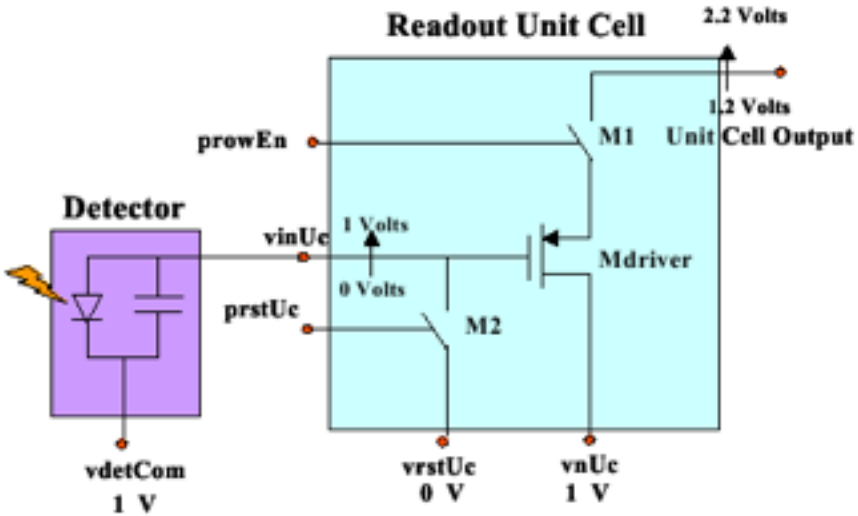


Infrared detectors:

CCDs become transparent at >1 micron
CMOS readout bump-bonded to HgCdTe diodes
Non-destructive readout -- read every ~few seconds,
reducing read noise, detect CRs in real time

Every pixel has this (expensive!)



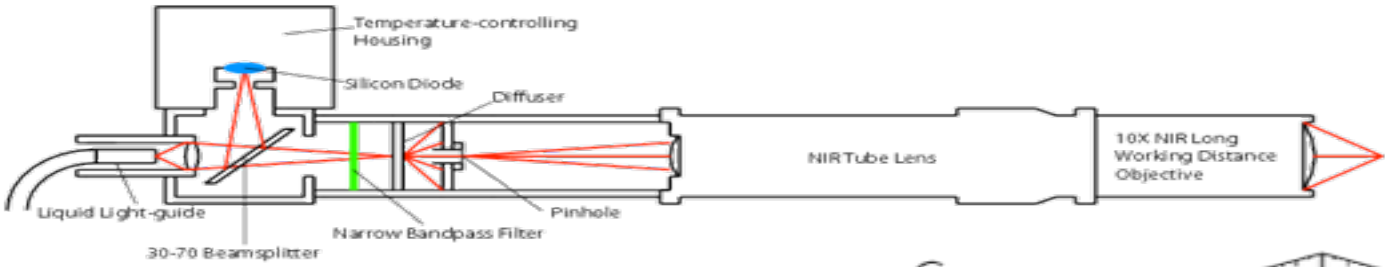


Problems with infrared detectors:

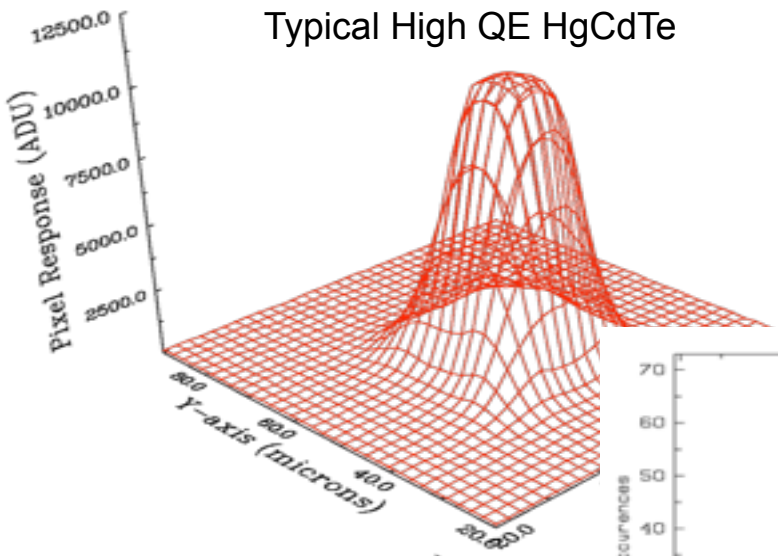
* Most of the same problems as with CCDs
* Dark current (worse)
* Intra-pixel: capacitive coupling between pixels
  -- pixels not independent!
* Inter-pixel: variation within a pixel
* “Reset anomaly” due to pixel self-heating
* Gain (+ noise) difficult to measure due to pixel correlations

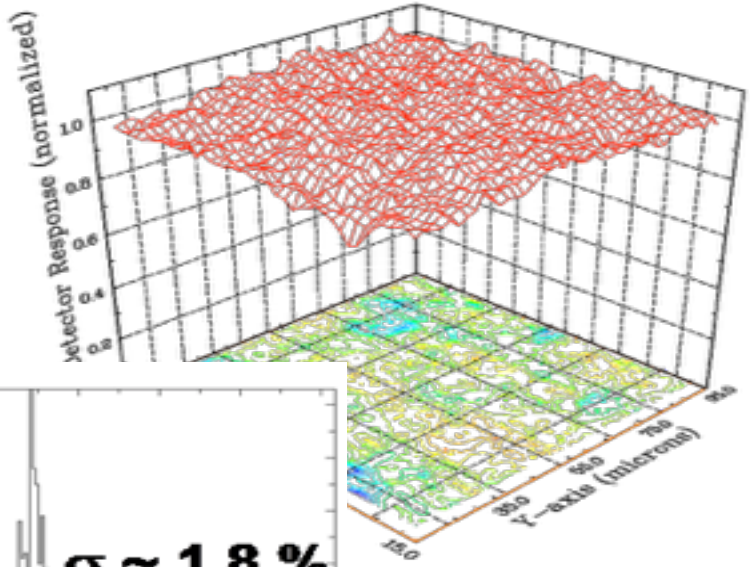
Problems with HgCdTe: Inter-pixel variation

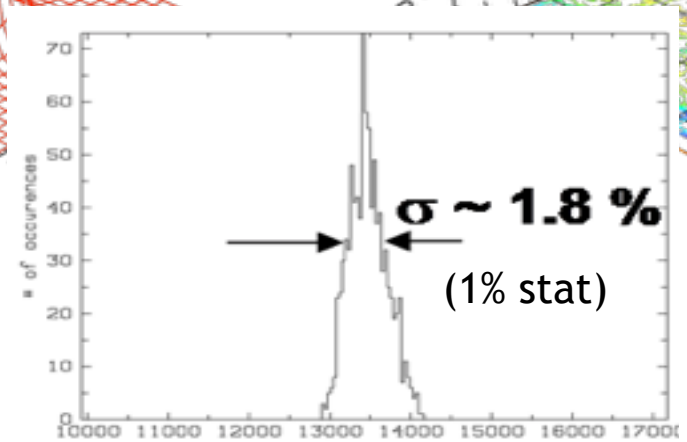
Micron-size NIR point projection system uncovers sub-pixel structure



Typical High QE HgCdTe







Problems with HgCdTe: Capacitive coupling

→ Cap. coupling occurs in mux and bump bond region

Average Correlation to neighboring pixels ~ 4%

(Nodal capacitance 32.2 fF 38.6 fF w/o IPC)

Average Correlation to neighboring pixels ~ 1% (rows), 0.5% (columns)

Nodal capacitance 75.1 fF

Average Correlation to neighboring pixels ~ 2.5% (rows), 1% (columns)

Nodal capacitance 77.7 fF

RSC Hawaii-2 HgCdTe with Epoxy Underfill

11	-3	0	-3	6	7	13
-7	14	4	23	-1	-9	-1
20	11	36	211	33	-4	-4
-3	12	265	5029	266	13	-3
-4	-5	34	210	37	9	22
0	-9	-1	22	7	14	-8
13	7	7	-3	1	-3	11

trace topology in multiplexer

RVS HgCdTe before Epoxy Underfill

0	-6	-6	10	8	4	3
18	2	1	-2	-3	-8	6
0	-3	-1	20	4	9	0
6	2	50	4552	50	3	5
-1	7	3	20	0	-4	0
9	-6	-2	-3	3	2	14
5	5	6	7	-4	-6	-1

RVS HgCdTe after Epoxy Underfill

11	-5	-1	-6	0	3	-3
12	-13	3	11	2	-8	-13
3	-2	28	43	17	3	-5
10	13	113	4136	114	14	11
-4	2	16	42	28	-1	2
-13	-7	4	9	5	-12	12
0	5	0	-5	0	-3	7

before after

epoxy underfill

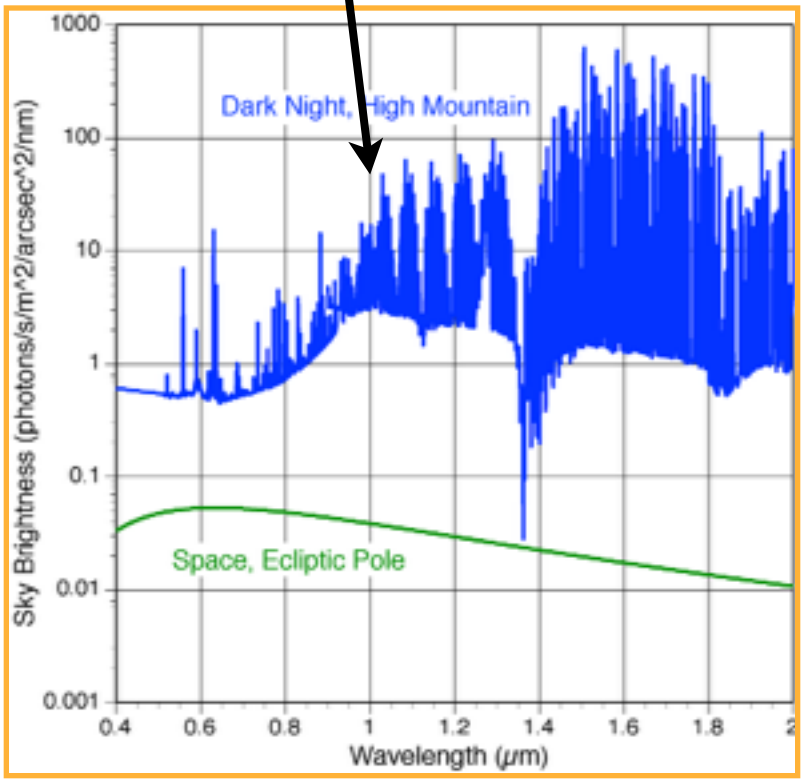
correlation increases by ~ 2x

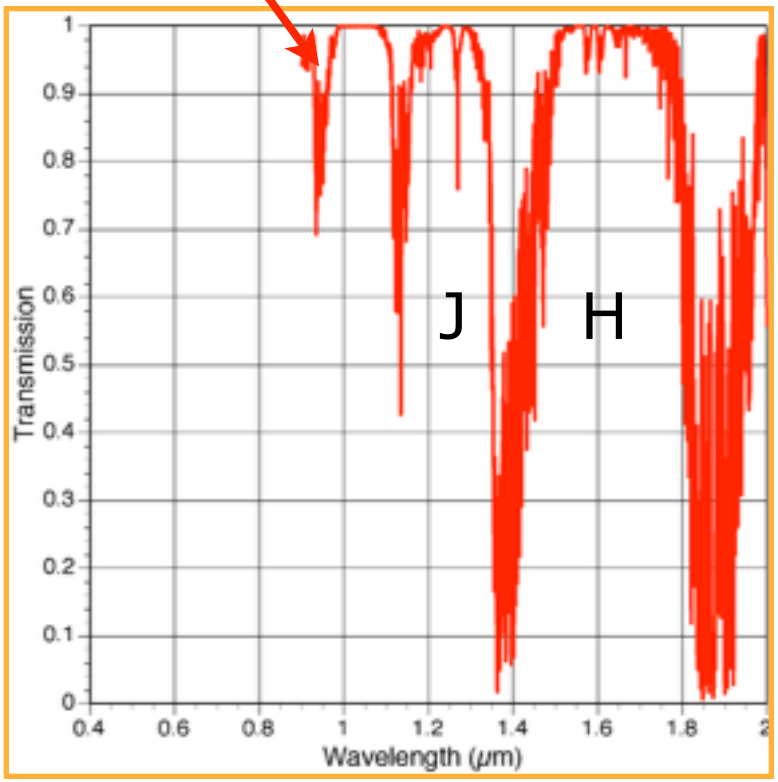
MGB, M. Schubnell, & G. Tarle, "Correlated Noise and Gain in Unfilled and Epoxy Under-filled Hybridized HgCdTe Detectors," *Published in PASP*, Sept. 2006.

Atmosphere:

Absorption, emission, seeing, refraction

- Sky is very bright in NIR: >100x brighter than in visible
- Sky is not transparent in NIR: absorption due to water is very strong and extremely variable





“Traditional” data reduction for imaging surveys

Steps in analyzing an image:

* Cosmetics
* Sky-subtraction
* Object-finding
* PSF
* Astrometry
* Photometric calibration: absolute, relative
* Measuring fluxes -- aperture, Petrosian, matched filter
* Measuring colors

↓

Object catalogs

Example image from SDSS

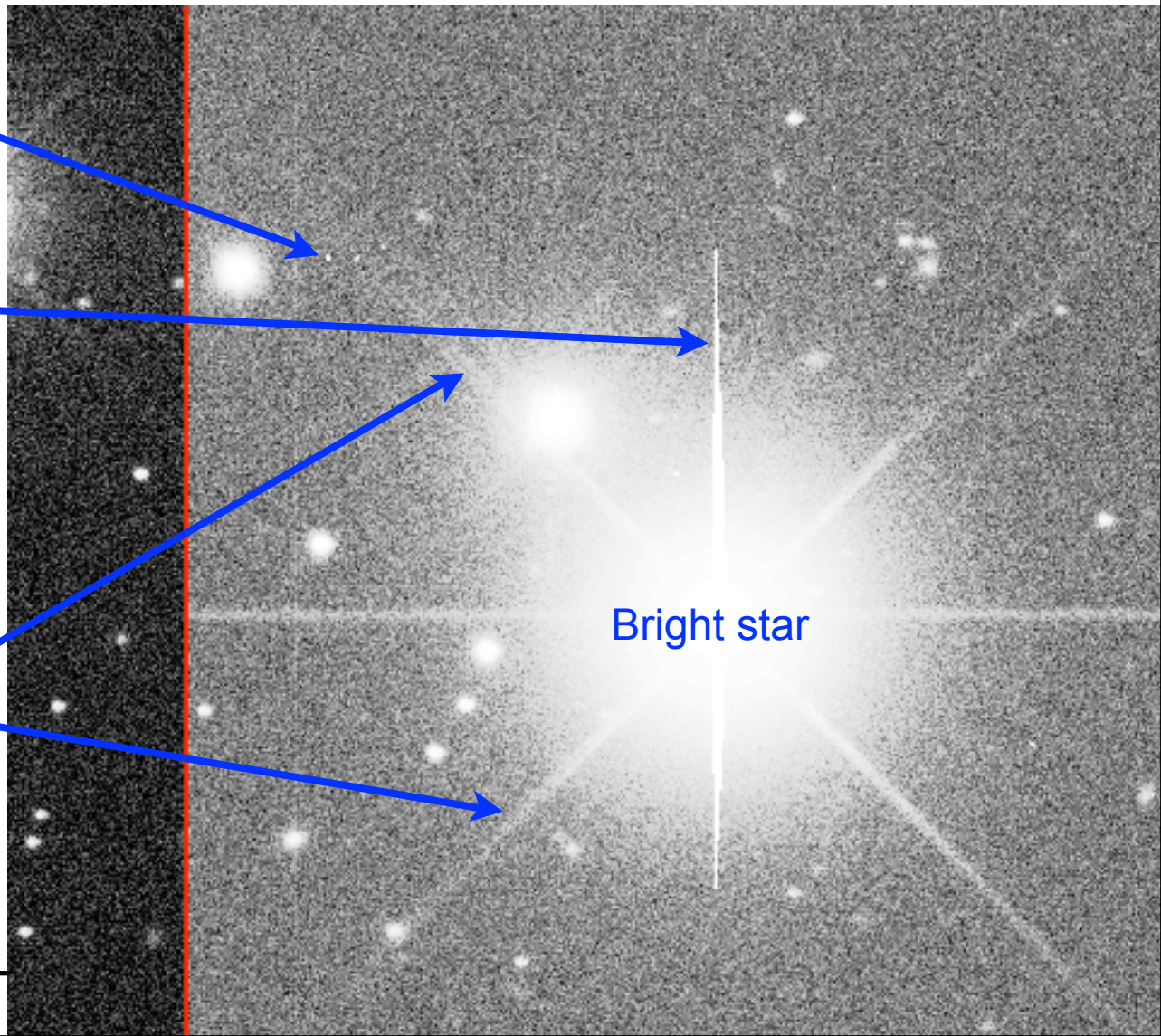
Raw image

2 amplifiers have different amplifier "gain" conversion from counts -> e-

Cosmic rays:
identified as too sharp,
violating sampling theorem

Saturation trails:
filled pixels bleed charge

Diffraction spikes from
telescope spider
holding secondary
mirror



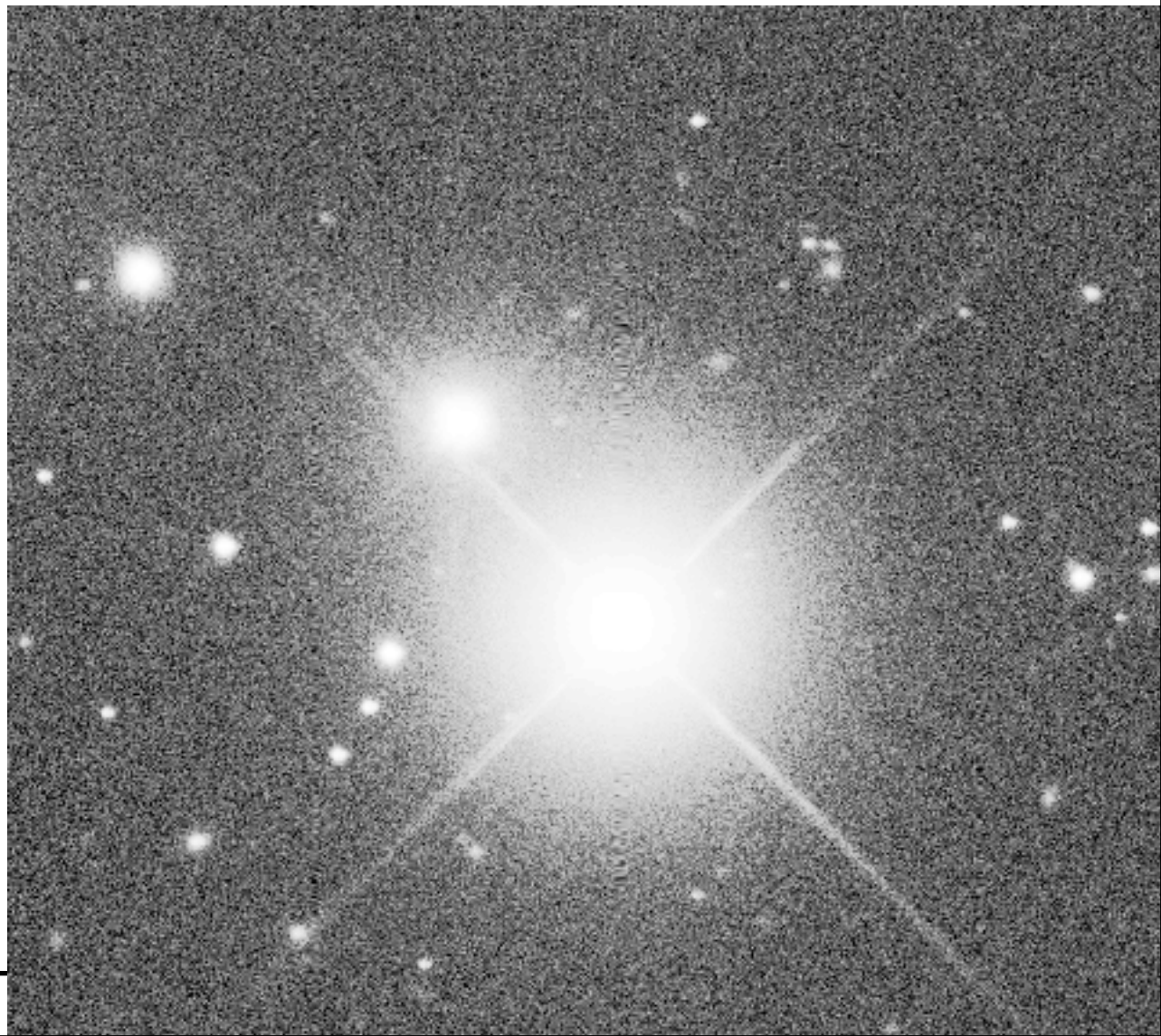
Example image from SDSS

Flat-fielded (divide by Q.E. of each pixel, amplifier gain)
Mask cosmic rays
Mask saturation
Mask bad columns

Example image from SDSS

Interpolate over bad pixels (but keep track of them!)

Noise/pixel = quad sum of read noise
+ Poisson noise
(estimated!!)



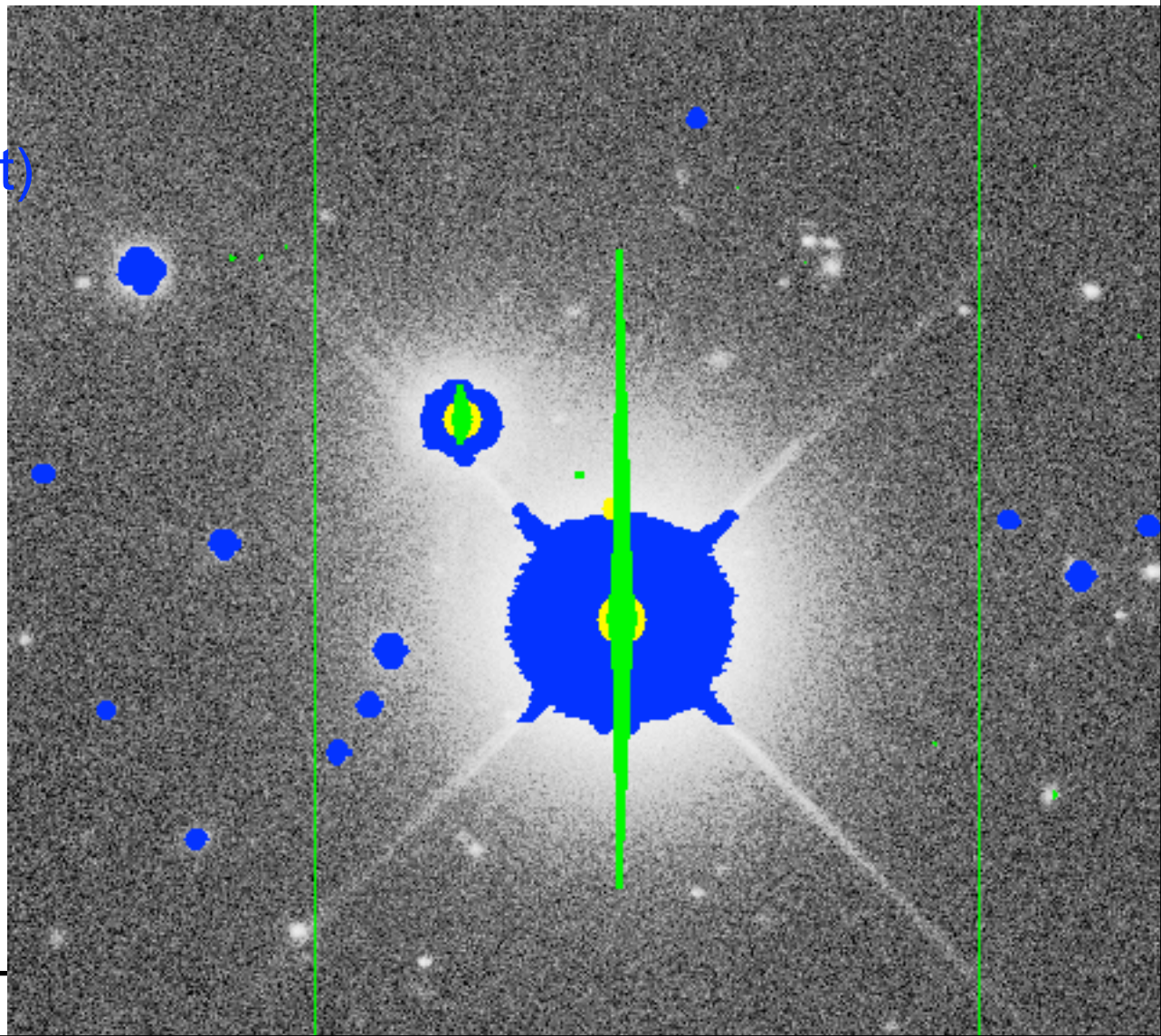
Example image from SDSS

Object detection

Two common approaches:

1) Matched filter

2) Peak-finding
   (PSF-smooth first)

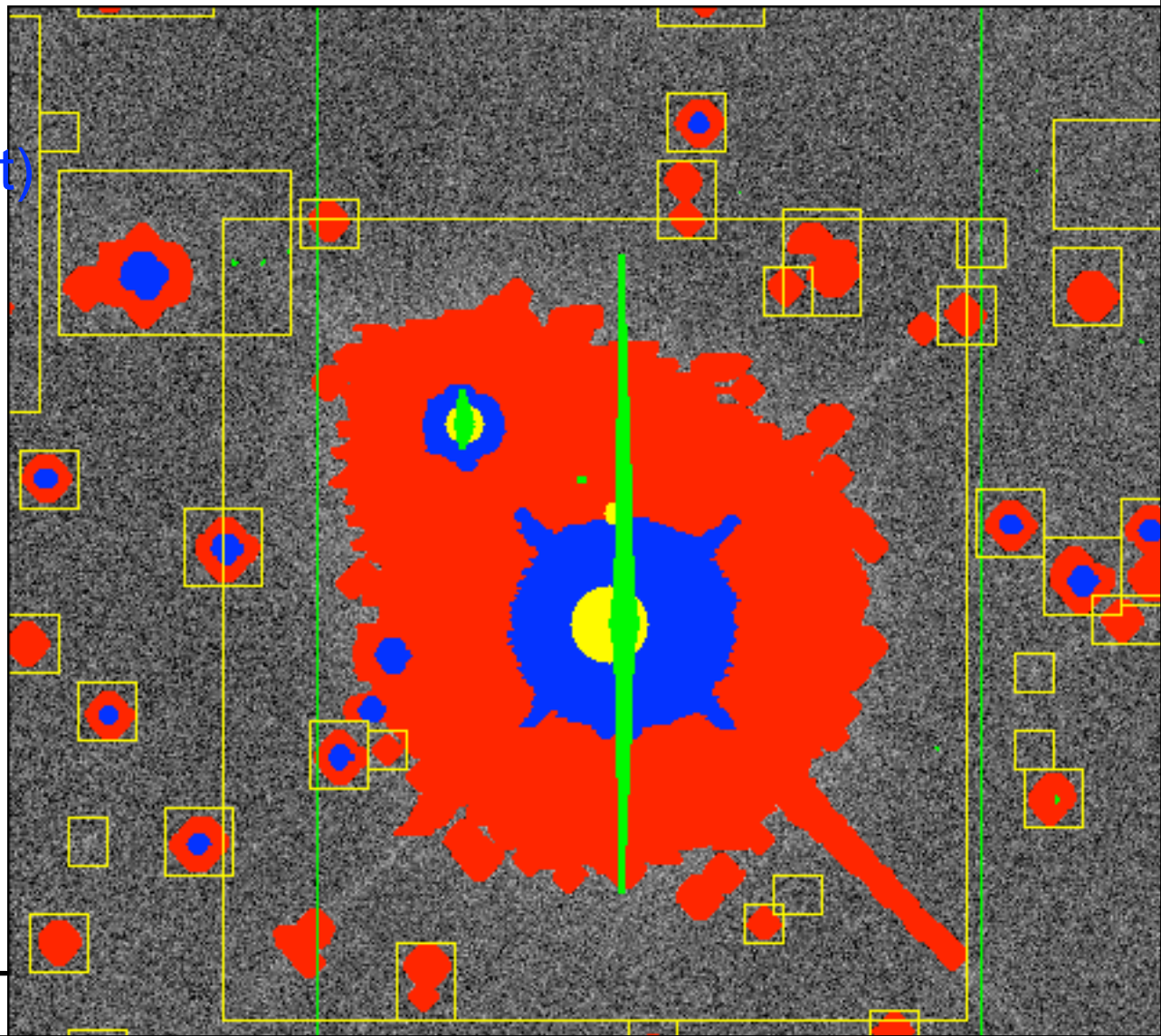


Example image from SDSS

Object detection -- ***iterative to find fainter objects***

Two common approaches:

1) Matched filter
2) Peak-finding
   (PSF-smooth first)



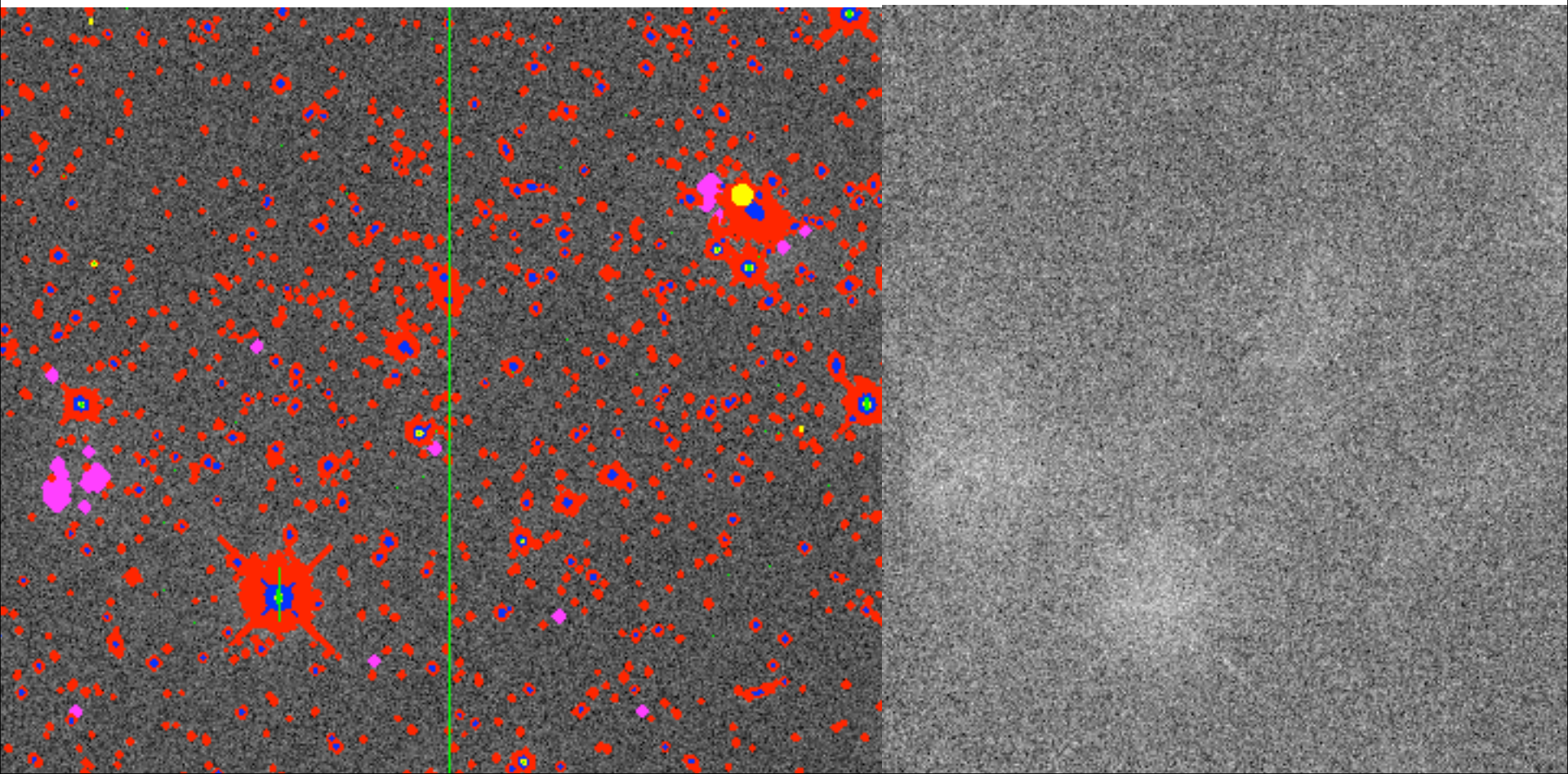
Example image from SDSS

What ***is*** the sky? Degeneracy sky ↔ faint, fuzzy objects

"True sky" = sky as seen above atmosphere,

possible to approach √N with multiple observations

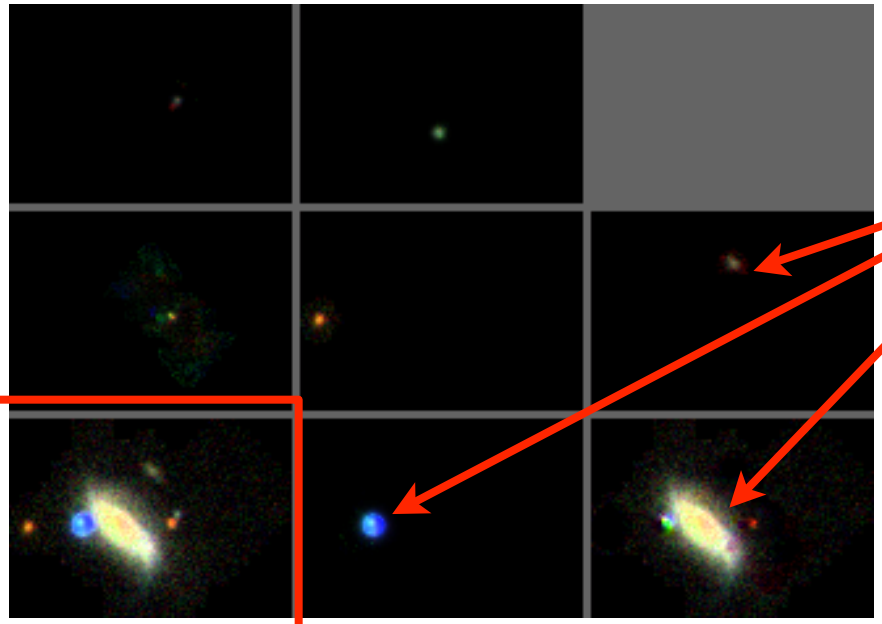
Mask objects when determining sky



Example image from SDSS

Object deblending

Not well-defined mathematically



Children objects

Parent image

SDSS algorithm: Assign flux to objects such that they are preferentially 180-deg symmetric

Courtesy: R. Lupton

Fig. 2.— The results of the SDSS deblender applied to the object `SDSS 756-3-356:1359,1051` (Lupton et al. 2005), using a `linear 0:100` stretch (see below). The bright star is heavily clipped. Figure 2b shows the same object with a `asinh:Q8 0:40` stretch to bring out the finer details of the object, and also the flaws in the deblended images.

Within each figure, the bottom left panel shows a *g*-*r*-*i* composite of the parent, prepared using the prescription of Lupton et al. (2004). The remaining panels each present a *g*-*r*-*i* composite of one of the children in approximate order of peak intensity. As discussed in the text, pixels outside the objects' atlas images are black.

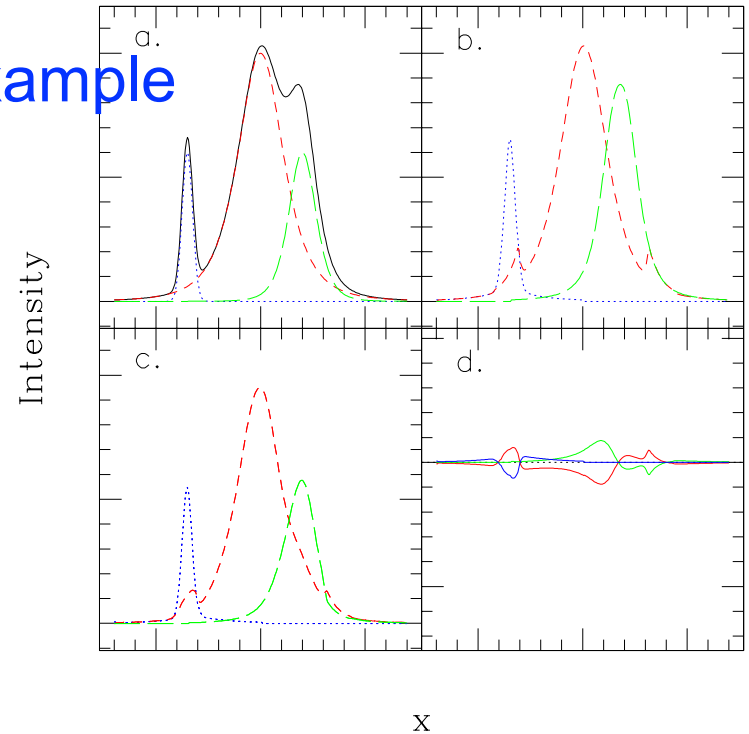
The stretches employed in this paper are `linear a:b` (a linear stretch from `a` counts to `b` counts) and `asinh:Qn a:b` (an inverse hyperbolic stretch with softening parameter `n`; the linear part of the transformation is equivalent to a `linear a:b` stretch).

Example image from SDSS

Object deblending

Not well-defined mathematically

1-D example



Courtesy: R. Lupton

Fig. 1.— A simple 1-dimensional example of deblending. Panel a shows the blended object (solid) and the three components that make it up; from left to right they are referred to as 'star', 'galaxy1', and 'galaxy2'. Panel b shows the corresponding *templates*; see text for details. Panel c shows the three deblended children; and panel d shows the difference between the three input components and the output of the deblender. The vertical scale is the same in all panels, but panel d has been shifted upwards to show negative as well as positive residuals.

PSF ↔ Object measurement

Not well-defined mathematically,
unless assumptions made

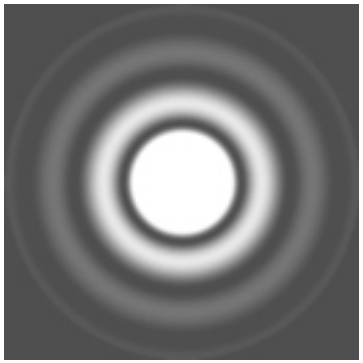
image = **PSF** * **galaxy** ?

image = **PSF** * **galaxy** ?

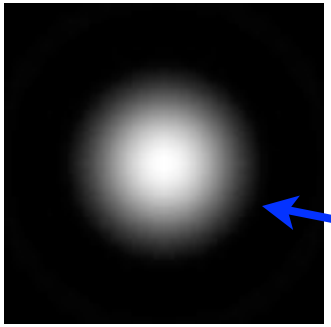
Assume the smallest objects are stars? Those can be CRs.

PSF:

There are lower bounds on the PSF size from:



Diffraction pattern from telescope
= Airy disk
(larger for longer λ)



Atmospheric blurring =
Kolmogorov turbulance

power-law wings

PSF not known from 1st principles

* Telescope optics not known to ~micron level
* Telescope optics change:
  temperature, gravity load (flexure)

DAOPHOT algorithm for PSF + object-finding

This is a common approach
Similar to "CLEAN" algorithm in radio astronomy

Assumptions: Every object is a star
Uniform PSF across an image

Algorithm:

Smooth image, find peaks
Solve PSF (often from bright objects) → linear equations
Solve flux of all stars → linear equations
Solve Δx, Δy to object positions → linear equations

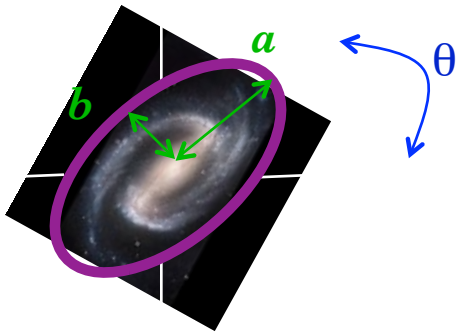
Equivalent to minimizing $\chi^2 = E\,(\text{flux})_i * (\text{PSF})$

→ number + position of stars is non-linear problem

Complications: Not all objects are stars! (galaxies at m > 20th)
PSF is function of x,y,time,λ → must model!

What object parameters are measured?

* Position on sky: α, δ
* Flux
* Colors == flux ratios in different filters
* Morphology: profile shape (PSF, deV, exp), ellipticity
* ***Errors*** on above



$e = (a^2-b^2)/(a^2+b^2)$
θ = Angle East of North

Model fitting

Star shapes:
* PSF

Galaxy shapes:
* Spirals: Flux = $\exp(-r/R_0)$
* Ellip: Flux = $\exp(-7.67\,(r/R_0)^{1/4})$
* More general profiles (Sersic,...)

Model fits never perfect due to galaxy substructure
→ Biased with S/N-dependence if usual χ^2 weighting

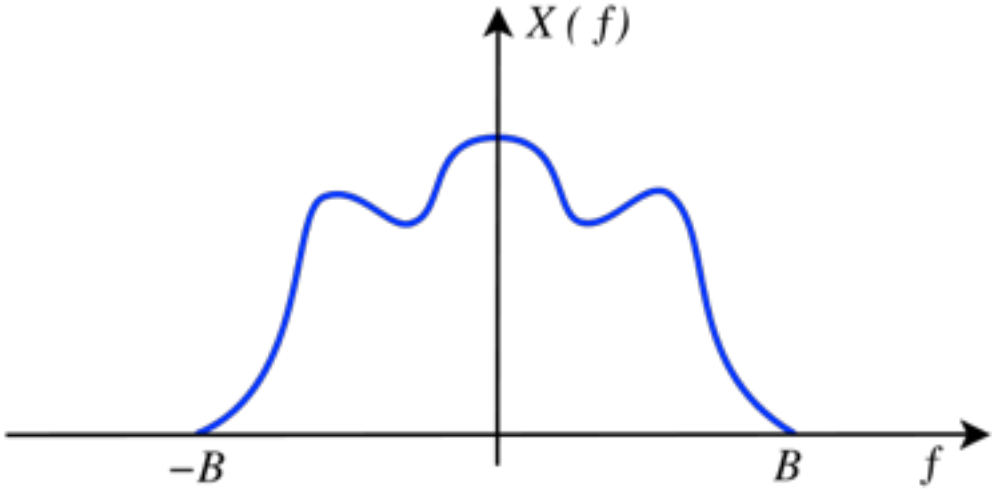
Well-sampled images -- what does that mean?

Nyquist-Shannon sampling theorem

"If a function x(t) contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced seconds apart."

$$\text{Image} = \Sigma\, A_f \sin(fx) + B_f \cos(fx)$$

(if sin,cos is the correct basis set)



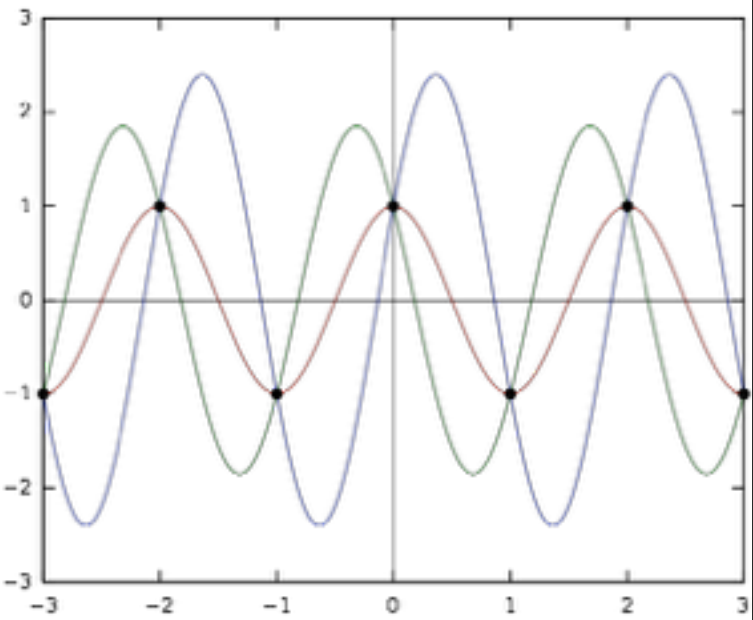


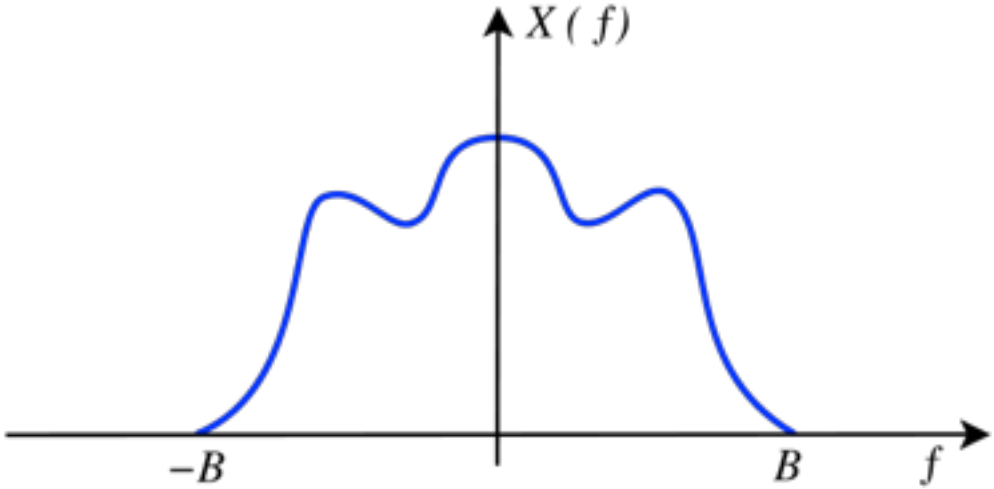
Image can be resampled (sync-shift) with no loss of information

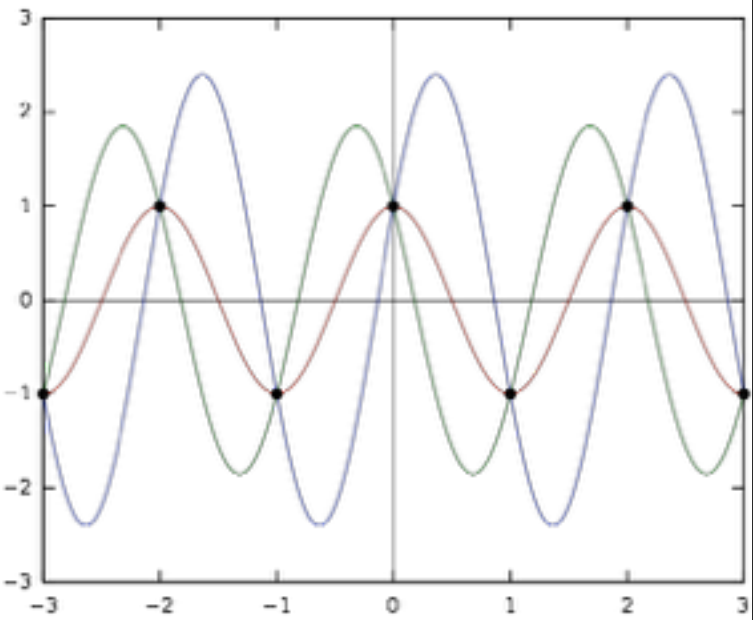
N.B.: ***Noise*** is never well-sampled in CCD images
(and therefore *aliases*)

Well-sampled images -- what if not well-sampled?

"Traditional" data reduction techniques don't work:

* Centroids
* Ellipticities
* Fluxes





Well-sampled images -- what if not well-sampled?

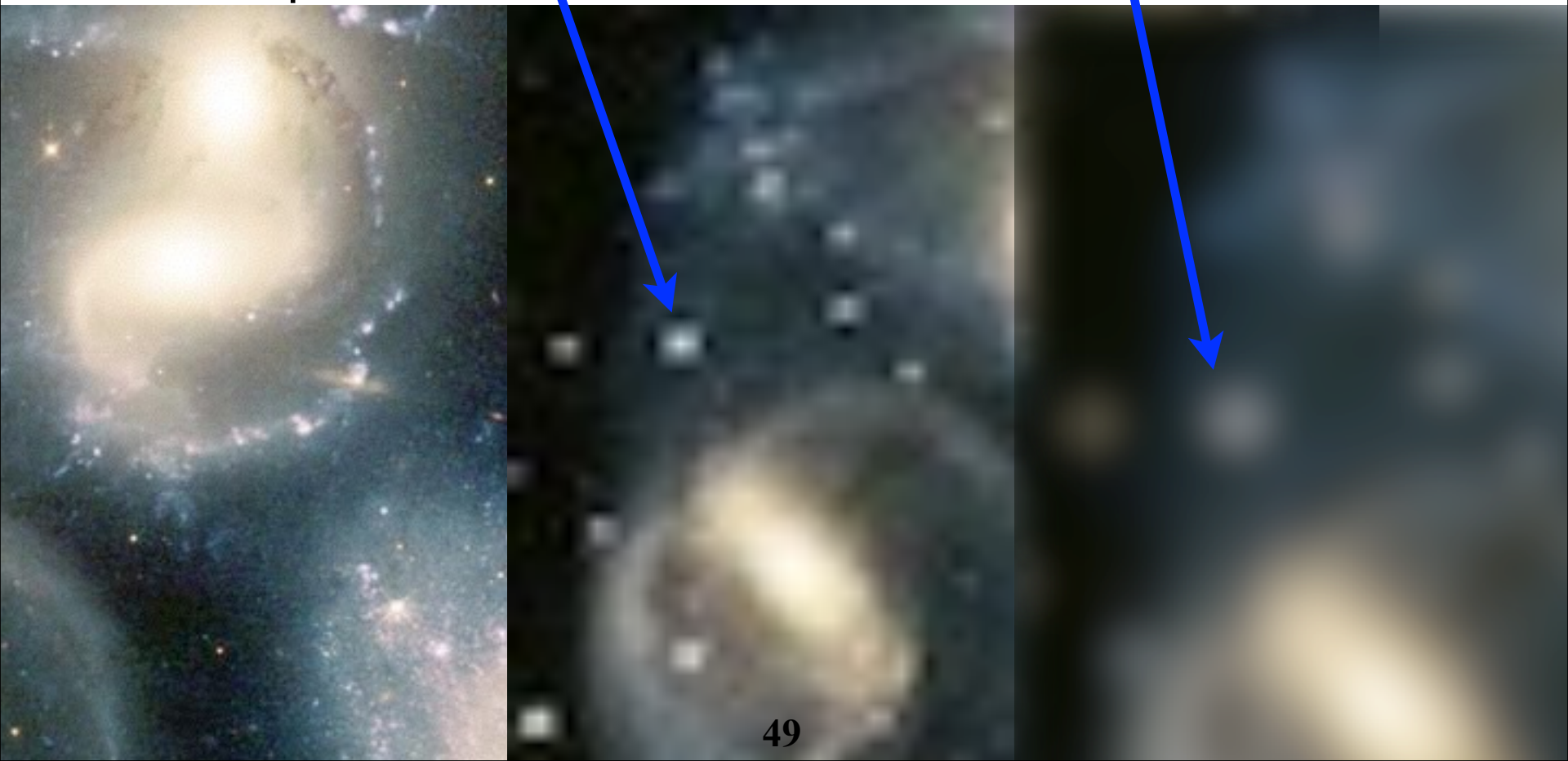
Centroids unknown to ~0.5 pix
Flux from neighboring objects aliased

Computable!

High-resolution well-sampled

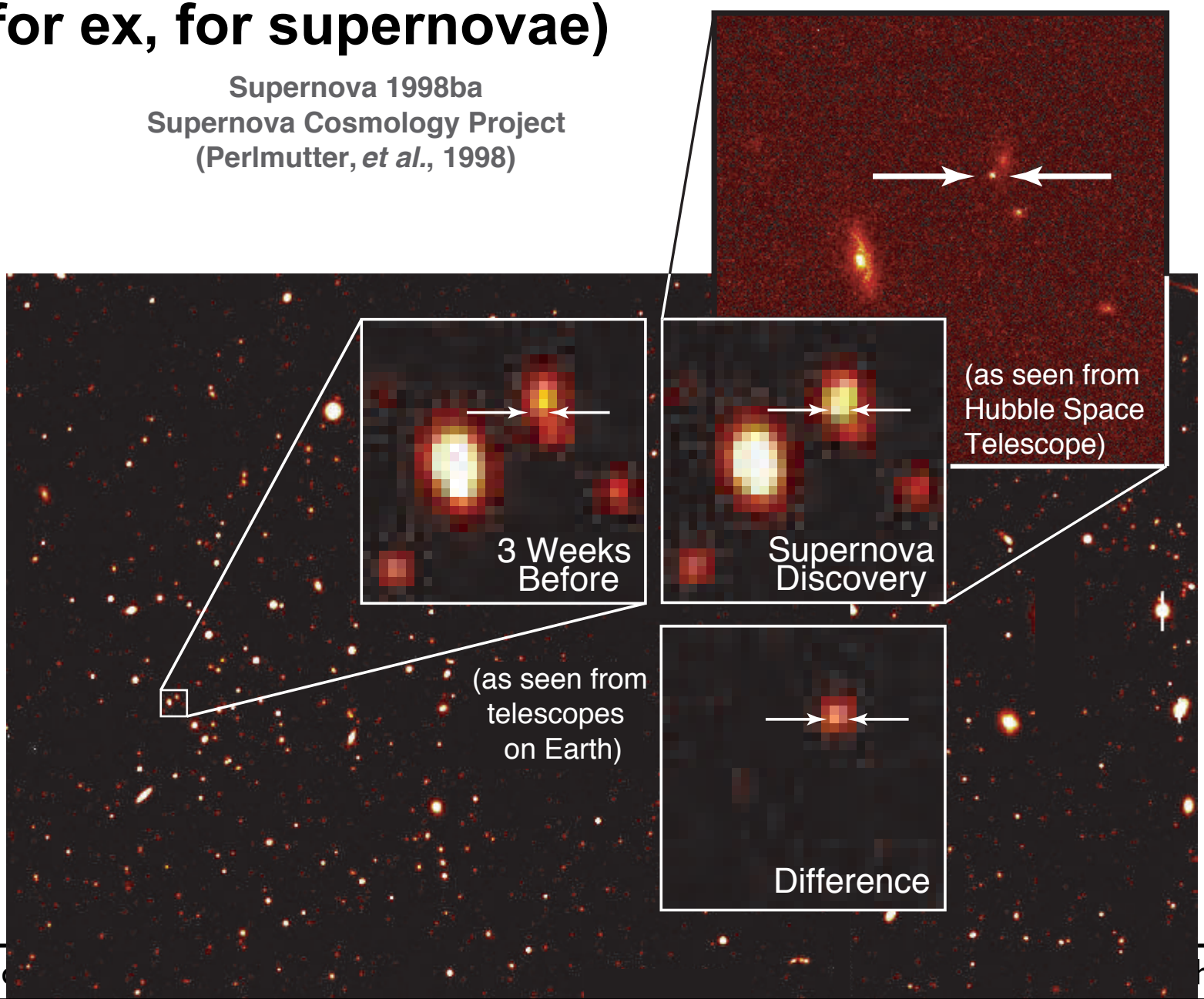
Undersampled

Well-sampled



Transient object detection from image differencing (for ex, for supernovae)

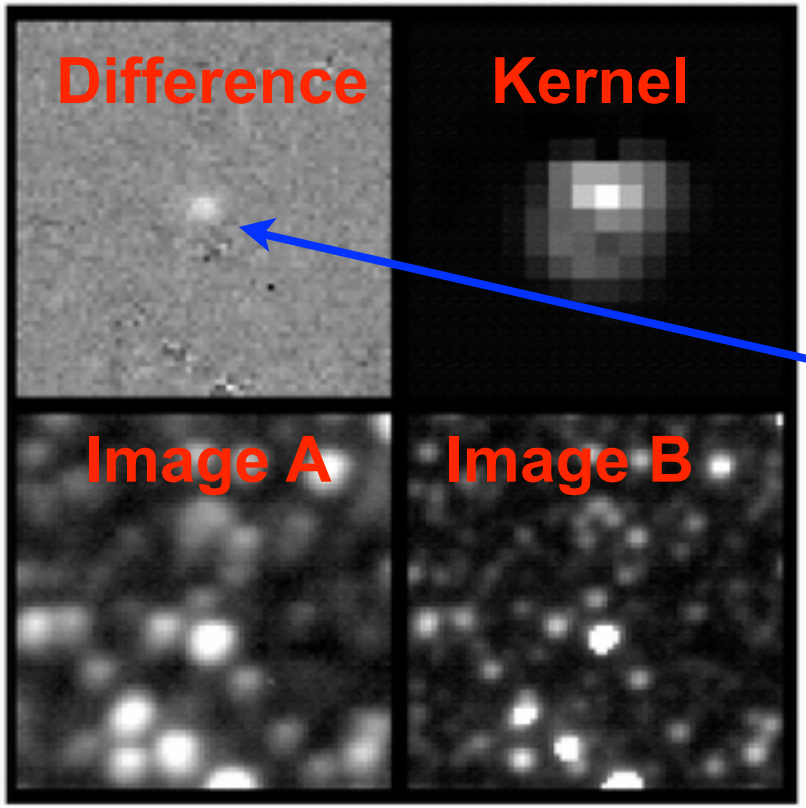
Supernova 1998ba
Supernova Cosmology Project
(Perlmutter, *et al.*, 1998)



Transient object detection from image differencing

The trick: PSF need not be known

Only need convolution kernel between images
→ More crowded images yield better convol. kernels



(Image A) = (Image B) * Kernel

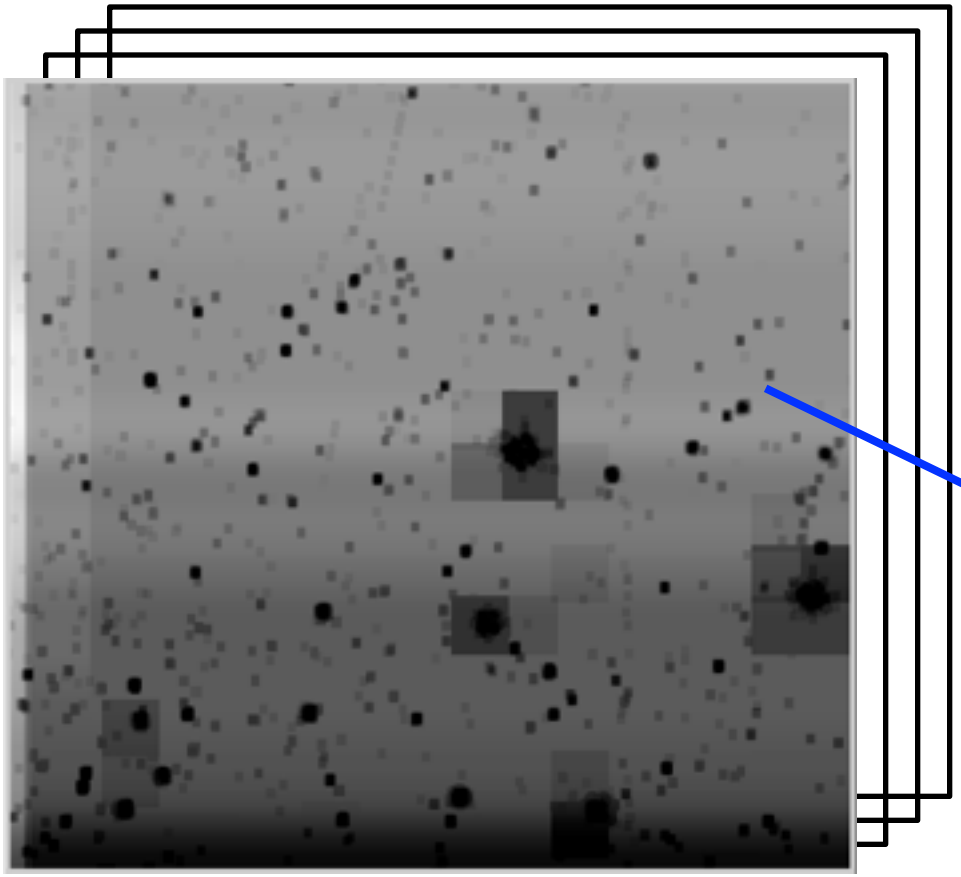
Difference = A - B * Kernel

High-z supernova

Learn more:
* Alard & Lupton (1998)
* Alard (2000)

Image stacking

Most imaging surveys (but not SDSS) combine multiple images
→ *Same issues of sampling*
If undersampled → what image consistent with all subimages?



Catalogs vs. Inference

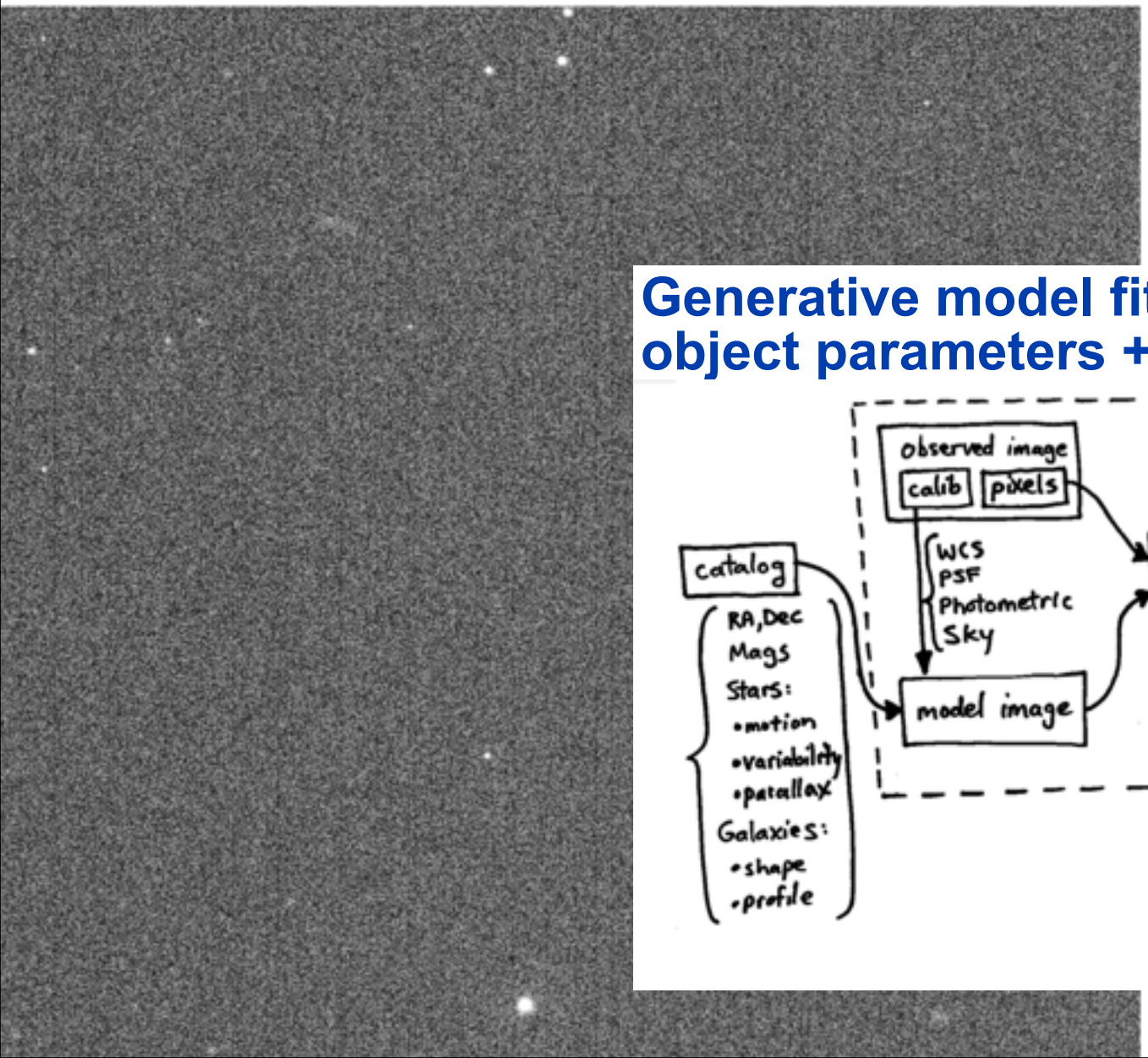
Catalogs: We currently live in the "age of catalogs"

This example:
Exactly 4 ± 0 galaxies
No covariances
Bad information if objects or separation smaller than pixels

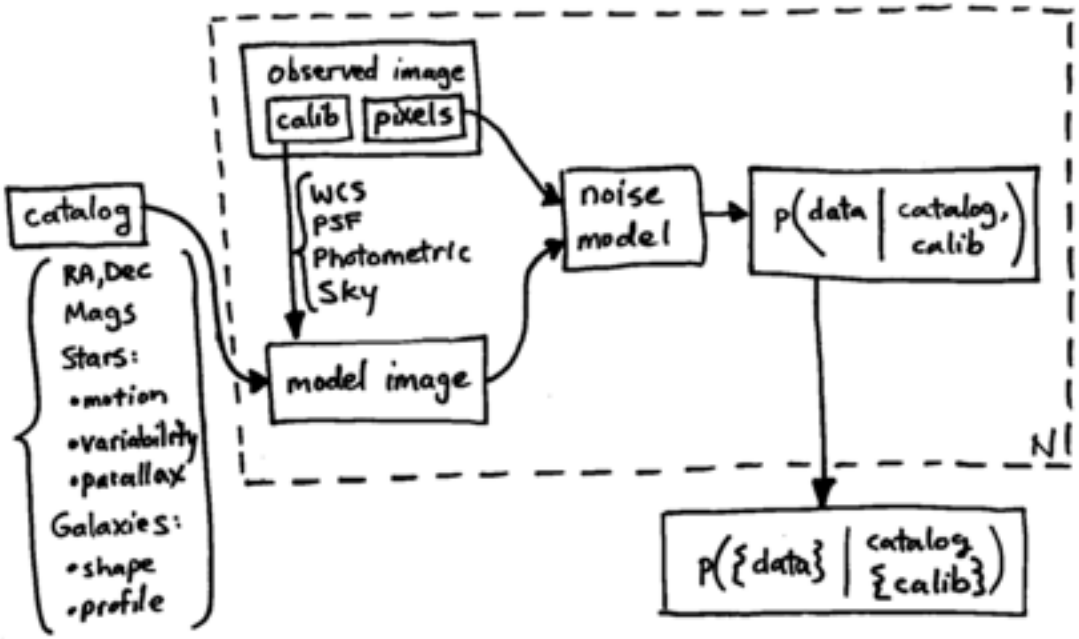
Inference: The future will be the age of "pixels"

Model the pixel-level data
No single "catalog"
Most easily applied to imaging data

Catalogs vs. Inference



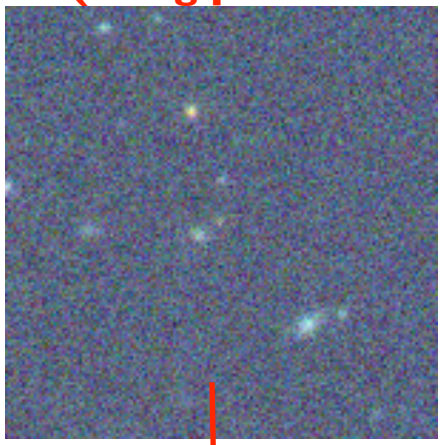
Generative model fits to raw data, object parameters + calibrations



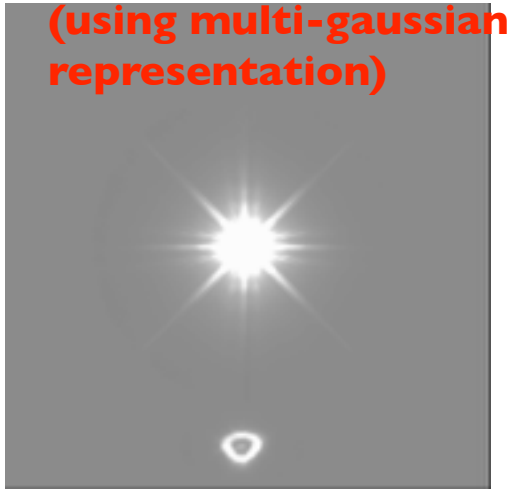
Inference: Dustin Lang's "Tractor"

Model space

SDSS catalog (using photo model params)



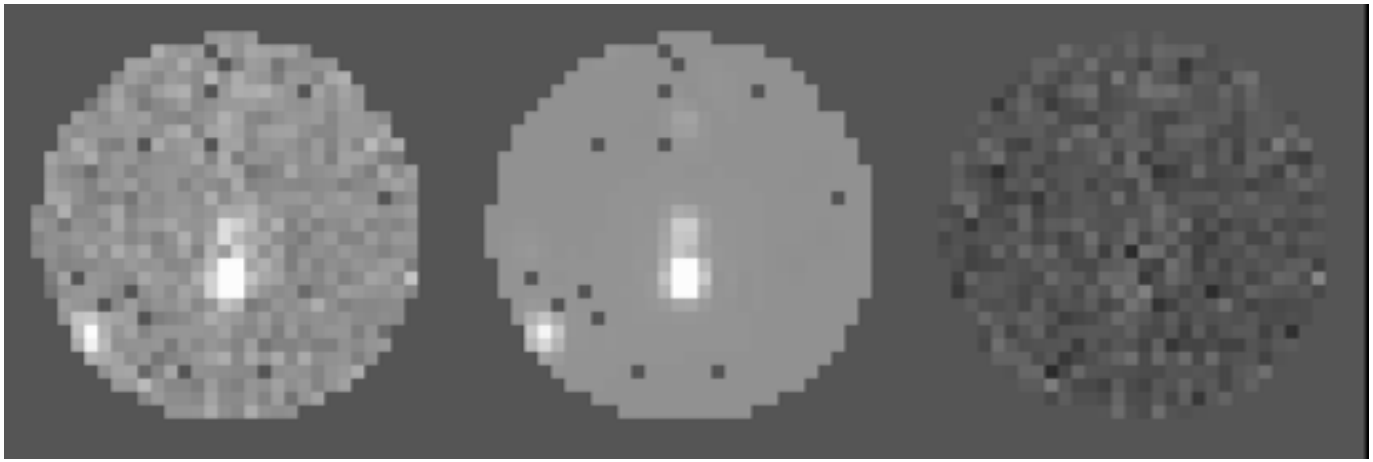
WISE PSF + ghosts, ... (using multi-gaussian representation)



Data space

WISE pixels **Model** **Residual**

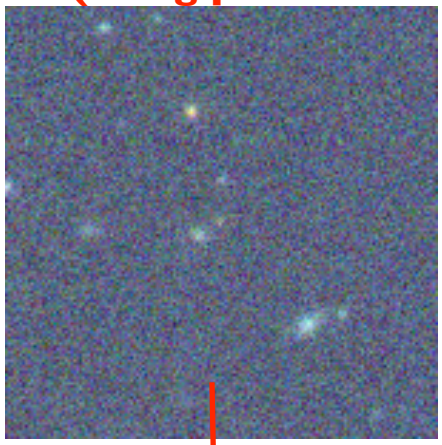
WISE single scan



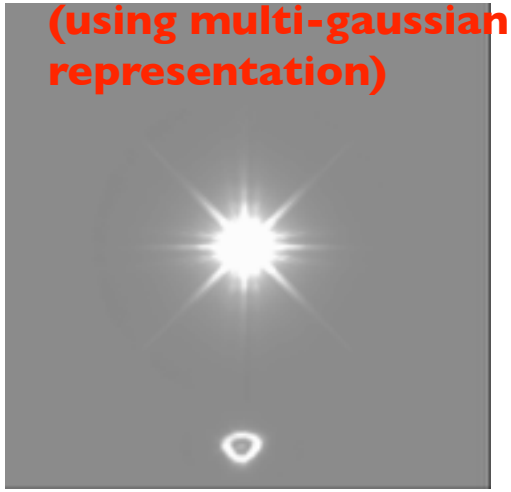
Inference: Dustin Lang's "Tractor"

Model space

SDSS catalog (using photo model params)



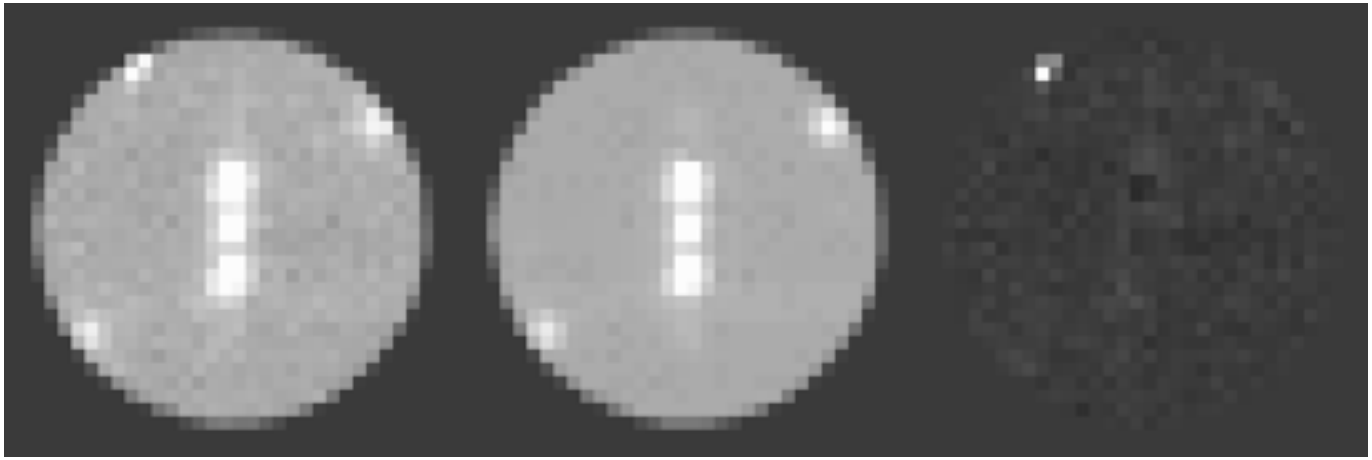
WISE PSF + ghosts, ... (using multi-gaussian representation)



Data space

WISE pixels **Model** **Residual**

WISE stacked



Large imaging surveys: What's the Future?

More étendue

Technology ➔ larger, close-packed CCD arrays

Proposed Large Synoptic Survey Telescope (LSST)

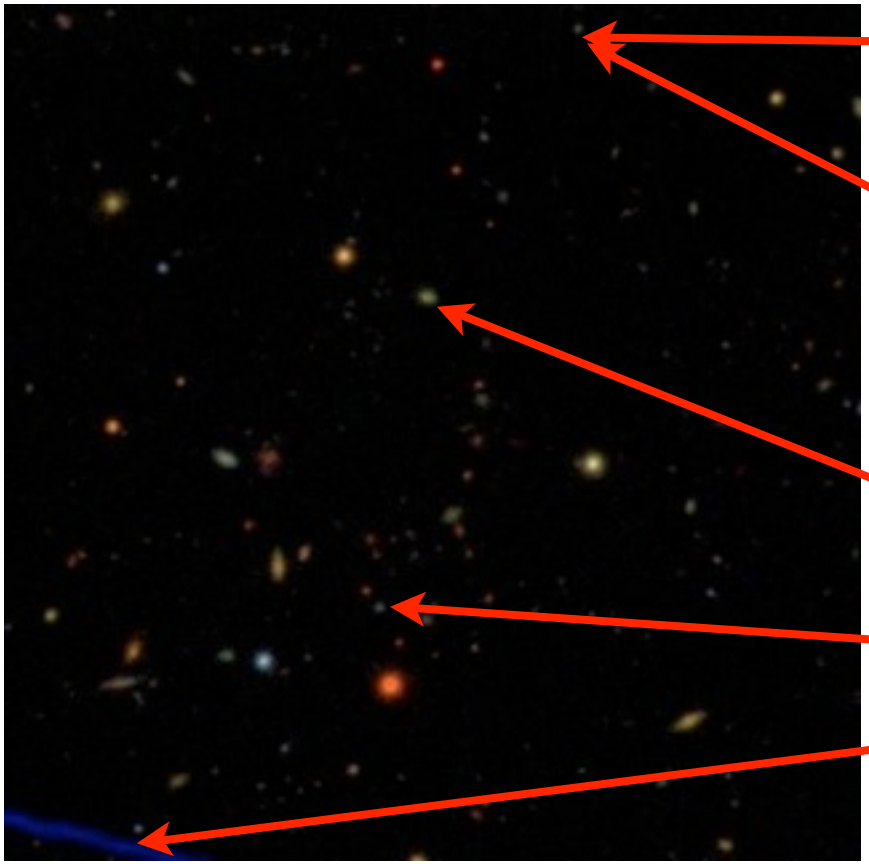
Scale of ***3-billion*** pixel camera for LSST

More volume

Time domain!

Large imaging surveys: What's the Future?

Time domain: Everything is variable!



All stars moving at ~0.001" per year relative to distant galaxies

Most stars vary in brightness by ~0.1% (sunspots,...)

Some stars explode as supernovae (briefly outshining the entire Universe)

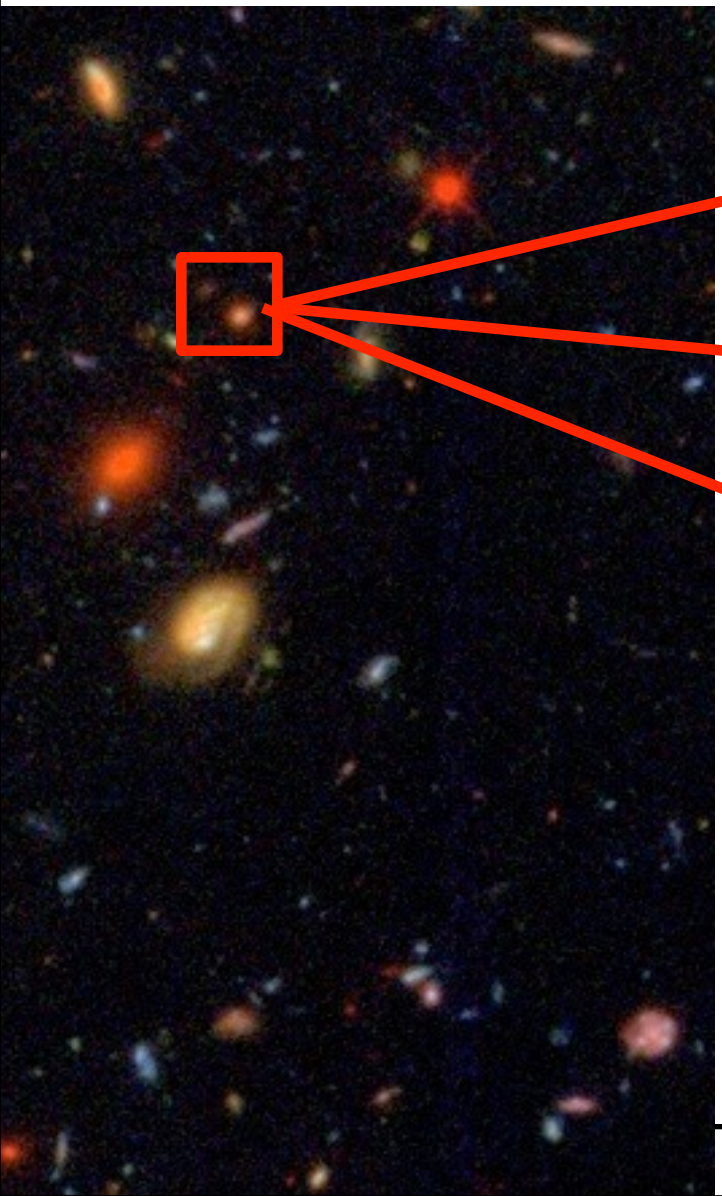
Most galaxies host massive black holes in their center (>10^6 x mass of the Sun), which vary

Distant QSOs vary by ~10% every ~month

Solar system objects move in 1 image (asteroids, outer solar system rocks,...)

"Static sky" is meaningless
Forward modeling is the only sensible approach

The future:
analyses using inference-based models



One galaxy

or

2 galaxies with covariant fluxes

or

3 galaxies with covariant fluxes

or …

~~"Catalog of catalogs" approach (Hogg 2011) - NO~~
~~Does not solve the detection problem~~

Likelihood functions for all model parameters - YES