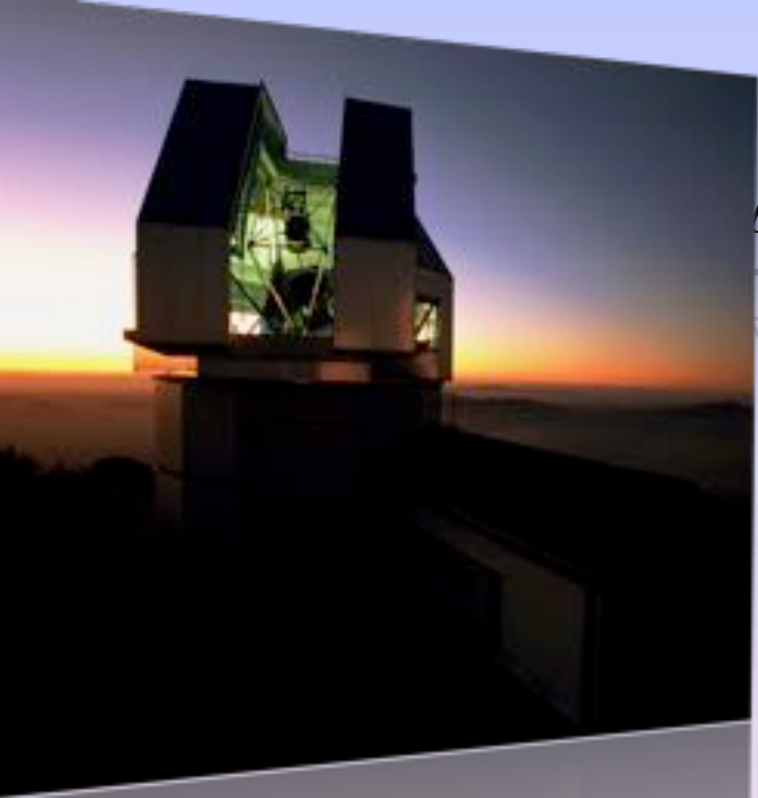
Astro 500



Techniques of Modern
Observational Astrophysics

Matthew Bershady
University of Wisconsin

Lecture Outline

Part I. Course Overview

- Regressions, error models and intrinsic scatter

Part II. Detectors

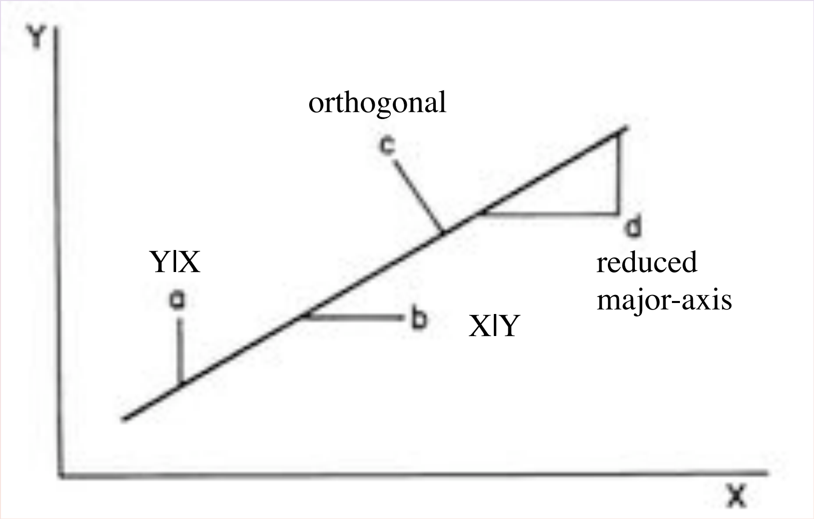
- CCDs: how they work, types, attributes & operation
- The digital unit: sampling, gain, and detector noise
- S/N formulation
- S/N regimes

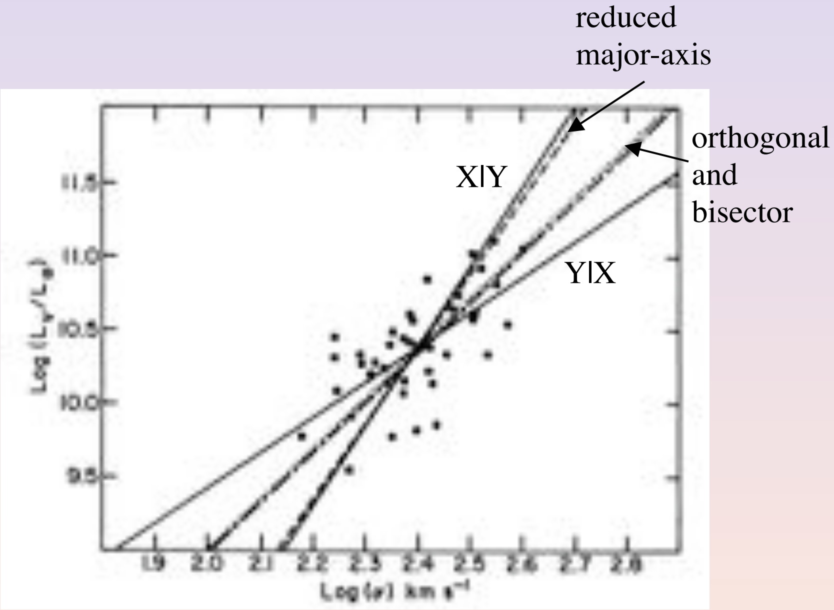
Linear Regressions

- Regressions are based on solving a set of linear equations based on different moments of the data and weighted by errors or priors.
- There are difference kinds of regression models (moments)
  - X|Y, Y|X, bisector, orthogonal
  - There are different assumptions to be made about errors that also lead to different moments and regressions
  - Is there an independent variable? (one variable with no errors)
  - Are the errors heteroscedastic or homoscedastic? (different or the same for all data)
  - Is there intrinsic scatter (usually other dimensions not known)?
- There is *no right regression model* (it depends what you want to learn), but there are correct and incorrect errors models and assumptions.
  - Social science analysis is plagued by systematic errors due to inaccurate models, but we're not free of such pitfalls because the universe is complicated.

Different Regressions

- X|Y, Y|X, bisector, orthogonal regressions
- Isobe et al. (1990, ApJ, 364, 104):
  - ordinary least squares (OLS) – no errors
- Akritas & Bershady (1996, ApJ, 470, 706)
  - bivariate correlated errors (heteroscedastic) and intrinsic scatter (BCES)





OLS Regression formulae

TABLE 1

Linear Regression Formulae for Slopes

Method	Expression for Slope	Estimate of the Variance of the Slope $\widehat{\mathrm{Var}}(\hat\beta_i)$
OLS(Y | X)	$\hat\beta_1 = \frac{S_{xy}}{S_{xx}}$	$\frac{1}{S_{xx}^2}\left[\sum_{i=1}^{n}(x_i-\bar x)^2(y_i-\hat\beta_1 x_i-\bar y+\hat\beta_1\bar x)^2\right]$
OLS(X | Y)	$\hat\beta_2 = \frac{S_{yy}}{S_{xy}}$	$\frac{1}{S_{xy}^2}\left[\sum_{i=1}^{n}(y_i-\bar y)^2(y_i-\hat\beta_2 x_i-\bar y+\hat\beta_2\bar x)^2\right]$
OLS bisector	$\hat\beta_3 = (\hat\beta_1+\hat\beta_2)^{-1}[\hat\beta_1\hat\beta_2-1+\sqrt{(1+\hat\beta_1^2)(1+\hat\beta_2^2)}]$	$\frac{\hat\beta_3^2}{(\hat\beta_1+\hat\beta_2)^2(1+\hat\beta_1^2)(1+\hat\beta_2^2)}[(1+\hat\beta_2^2)^2\,\widehat{\mathrm{Var}}(\hat\beta_1)$ $+2(1+\hat\beta_1^2)(1+\hat\beta_2^2)\,\widehat{\mathrm{Cov}}(\hat\beta_1,\hat\beta_2)+(1+\hat\beta_1^2)^2\,\widehat{\mathrm{Var}}(\hat\beta_2)]$
Orthogonal regression	$\hat\beta_4 = \frac{1}{2}[(\hat\beta_2-\hat\beta_1^{-1})+\mathrm{Sign}(S_{xy})\sqrt{4+(\hat\beta_2-\hat\beta_1^{-1})^2}]$	$\frac{\hat\beta_4^2}{4\hat\beta_1^2+(\hat\beta_1\hat\beta_2-1)^2}[\hat\beta_1^{-2}\,\widehat{\mathrm{Var}}(\hat\beta_1)+2\,\widehat{\mathrm{Cov}}(\hat\beta_1,\hat\beta_2)+\hat\beta_1^2\,\widehat{\mathrm{Var}}(\hat\beta_2)]$
Reduced major-axis	$\hat\beta_5 = \mathrm{Sign}(S_{xy})(\hat\beta_1\hat\beta_2)^{1/2}$	$\frac{1}{4}\left[\frac{\hat\beta_2}{\hat\beta_1}\,\widehat{\mathrm{Var}}(\hat\beta_1)+2\,\widehat{\mathrm{Cov}}(\hat\beta_1,\hat\beta_2)+\frac{\hat\beta_1}{\hat\beta_2}\,\widehat{\mathrm{Var}}(\hat\beta_2)\right]$

5

Errors on Regressions

- How do you estimate errors on slope and intercept?
- Resample your data:
  - Boot-strap – pick N data points out of sample of N, m times. Each pick is a random selection from N data points with equal probability of selecting i^{th} element.
  - Jack-knife – recalculate leaving out one datum, N times (N data)
- Monte Carlo simulation – artificial data

✧ When in doubt, “Monte Carlo” your data

✧ This applies not just to linear regressions but any modeling.

When in doubt….

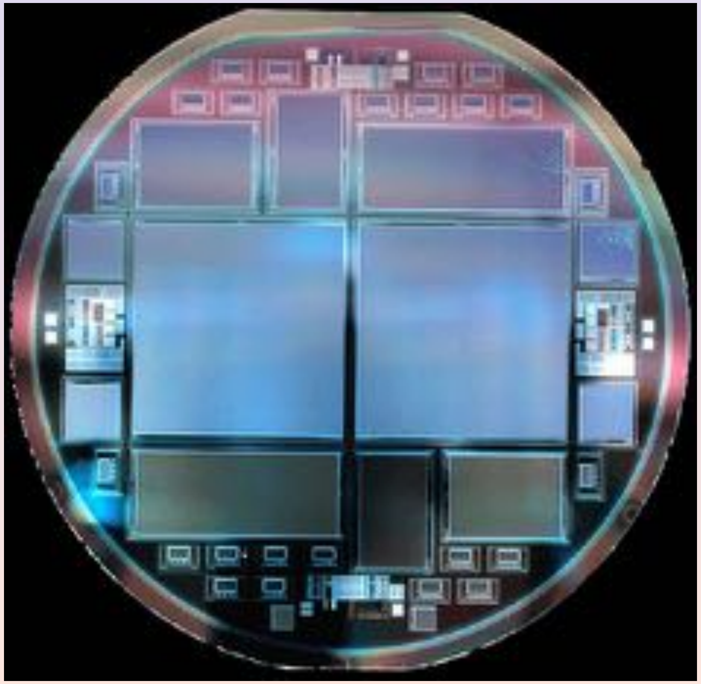
- "Monte Carlo (MC) your data"
- Monte Carlo: a town in Monaco (country in SE France) famous for gambling casinos
- What you need:
  - Model of data
  - Model of errors
  - Model of data sampling (range, censorship, incompleteness, spurious source (when applicable).
  - A good random-number generator
  - A modicum of computing skill and cpu time.
- How good is it?
- Only as good as your assumptions (i.e., model)
- Test your assumptions by comparing distributions (and their characterization) generated by MC against those from the data.

Miscellaneous topics in statistics

- Items to come back to:
  - How do distribution- and selection-functions bias regressions?
  - Famous biases in astronomy
    - Malmquist
    - Eddington
  - What linear regression should you use? …or:
  - Should you even use a linear regression?

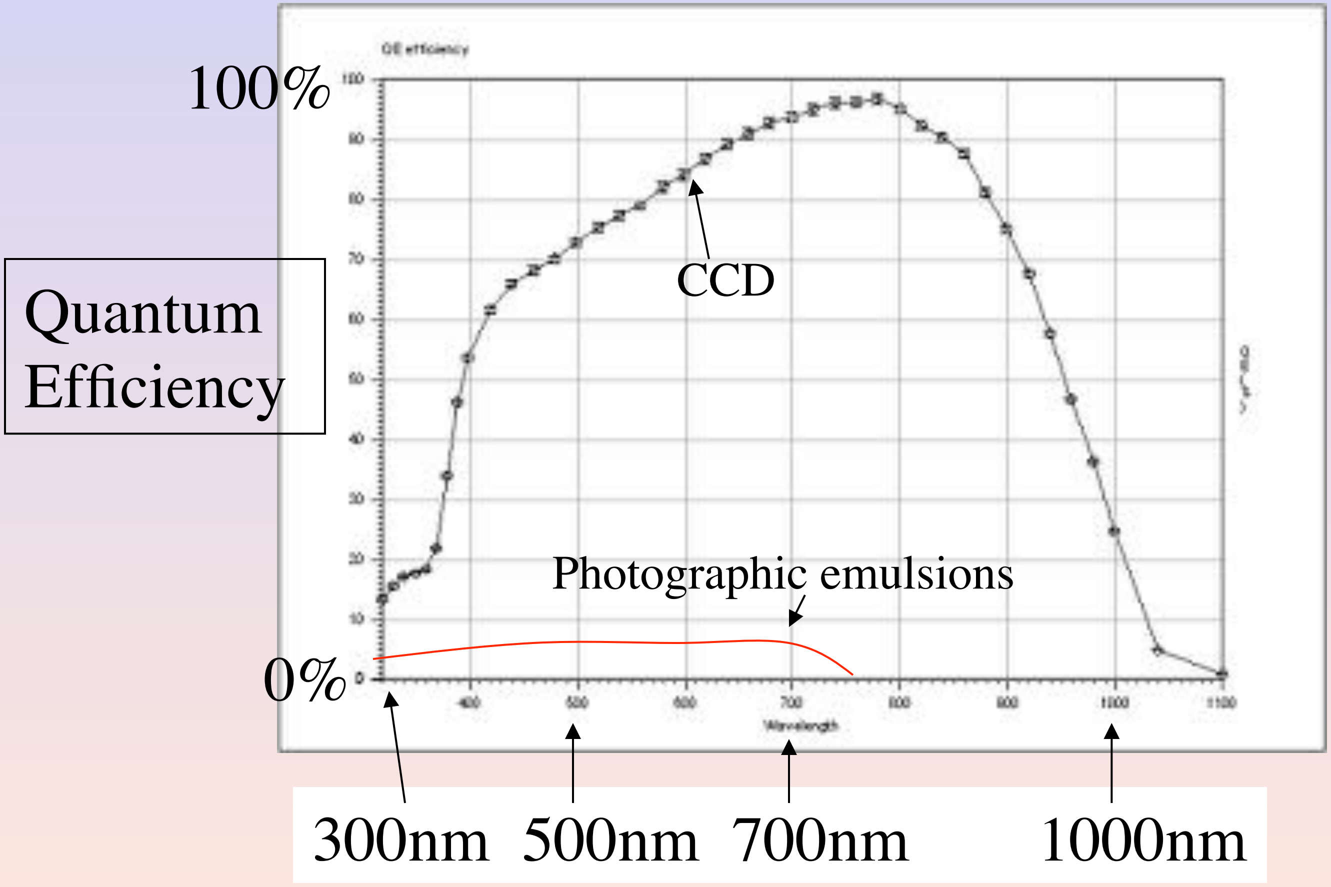
Digital Detectors

- By far the most common detector for wavelengths $300\text{nm} < \lambda < 1000\text{nm}$ is the CCD.

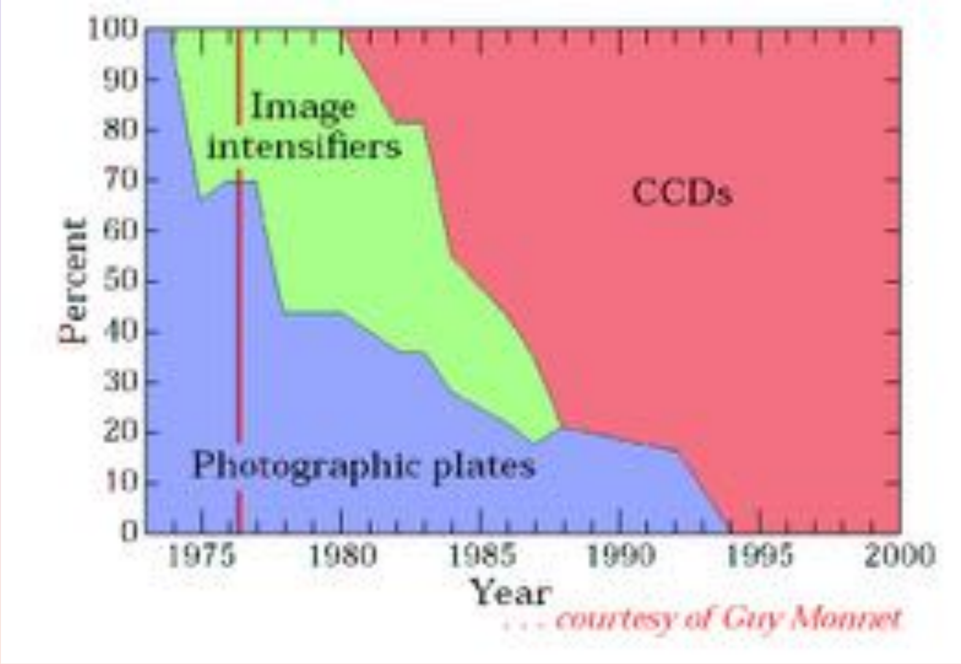


CCDs

1. Quantum efficiency is more than an order of magnitude better than photographic plates.



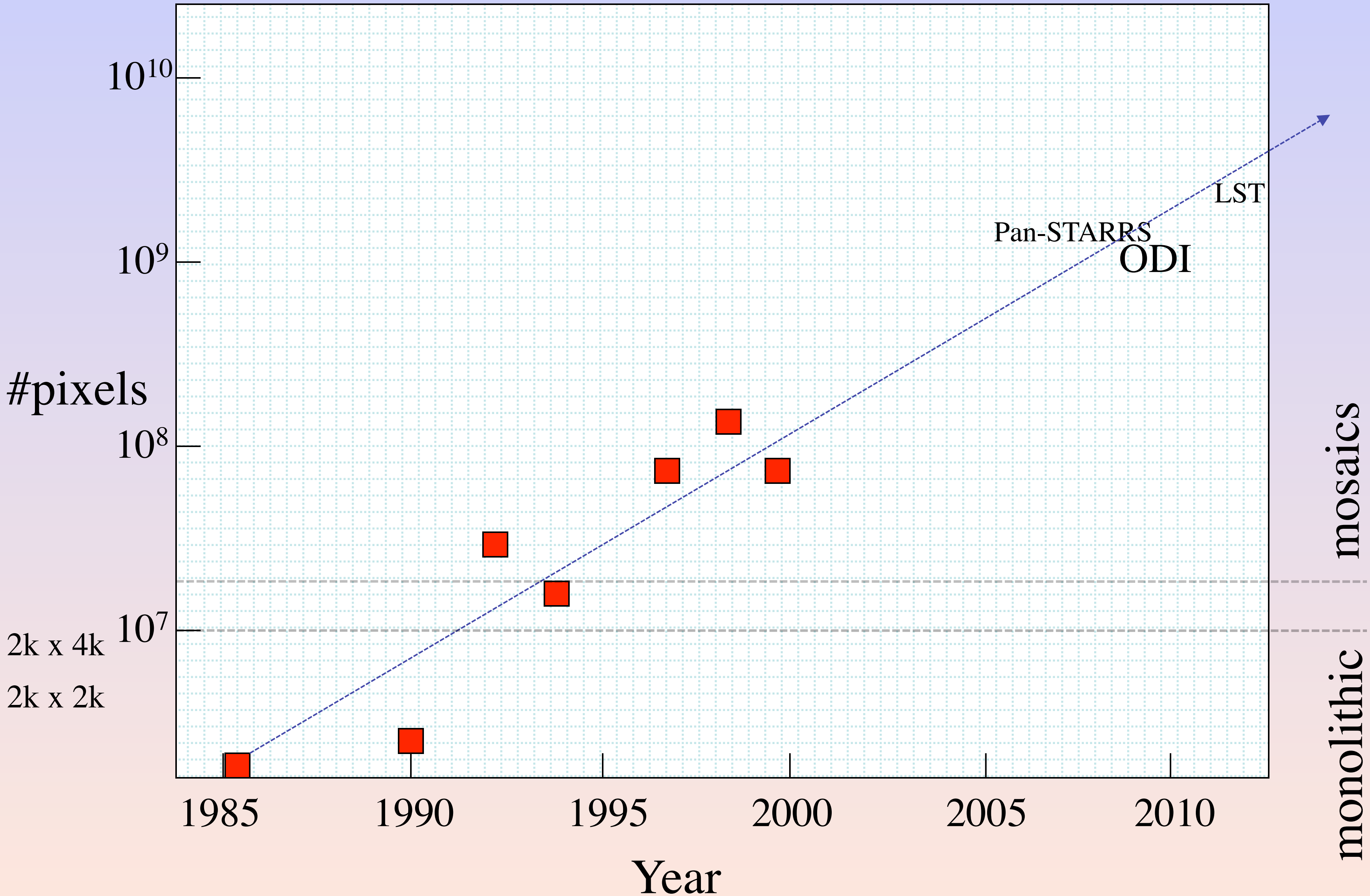
These are silicon fab-line devices and complicated to produce.



digital revolution

CCDs remain physically small compared to photographic plates, but they took over rapidly anyway.

CCD size



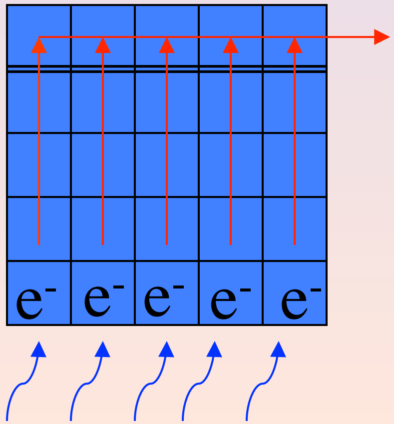
10^{10}
10^{9}
10^{8}
10^{7}
#pixels
2k x 4k
2k x 2k
LST
Pan-STARRS
ODI
mosaics
monolithic
1985
1990
1995
2000
2005
2010
Year

CCDs: How do they work?

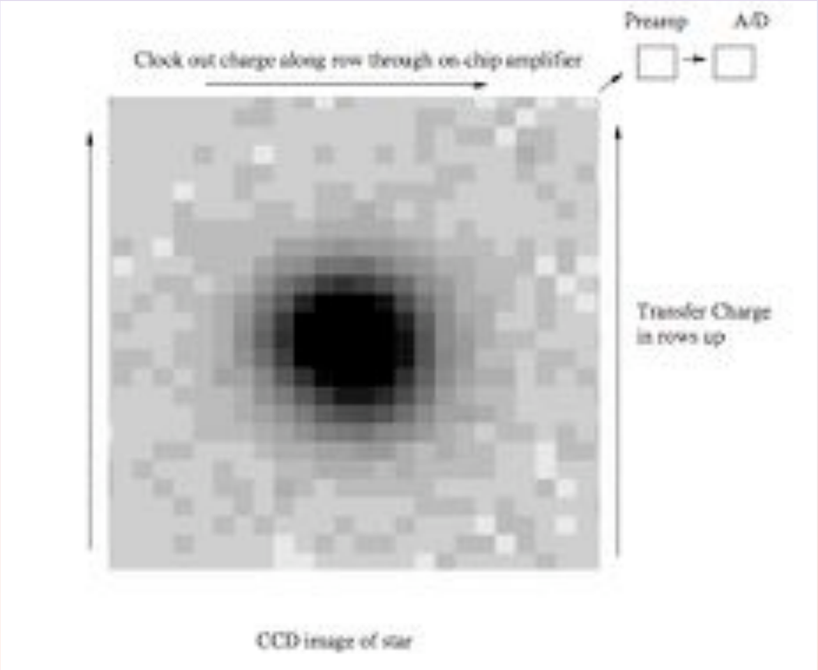
Photons → Electrons bumped into conduction band → Read out & amplify charge → A/D convert to DN

- Silicon semiconductors with “gate” structure to produce little potential corrals or wells.

serial register

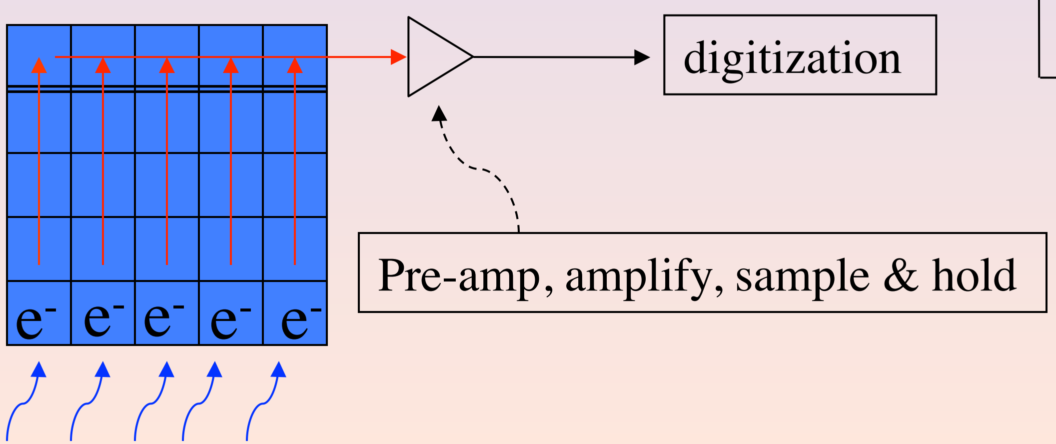


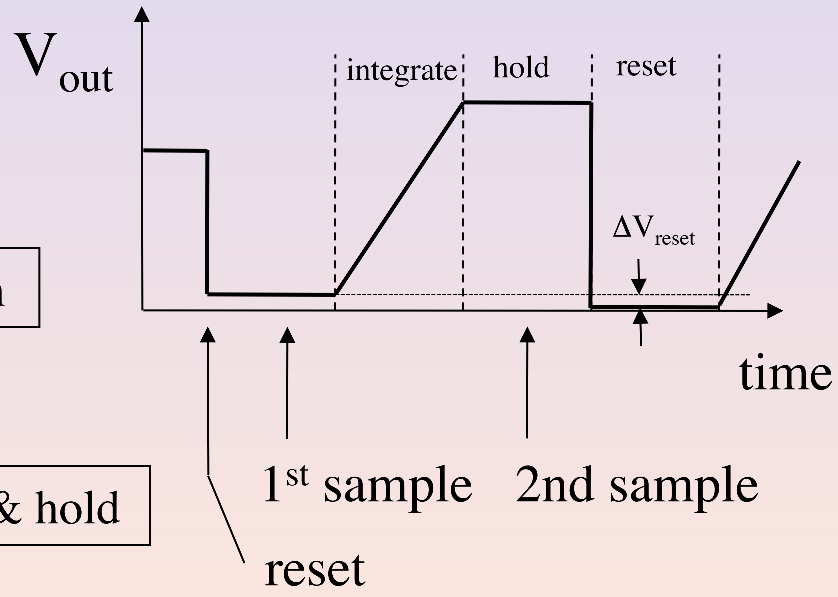
‘clock’ parallel and serial registers
“CTE” > 0.99999



Correlated double-sampling

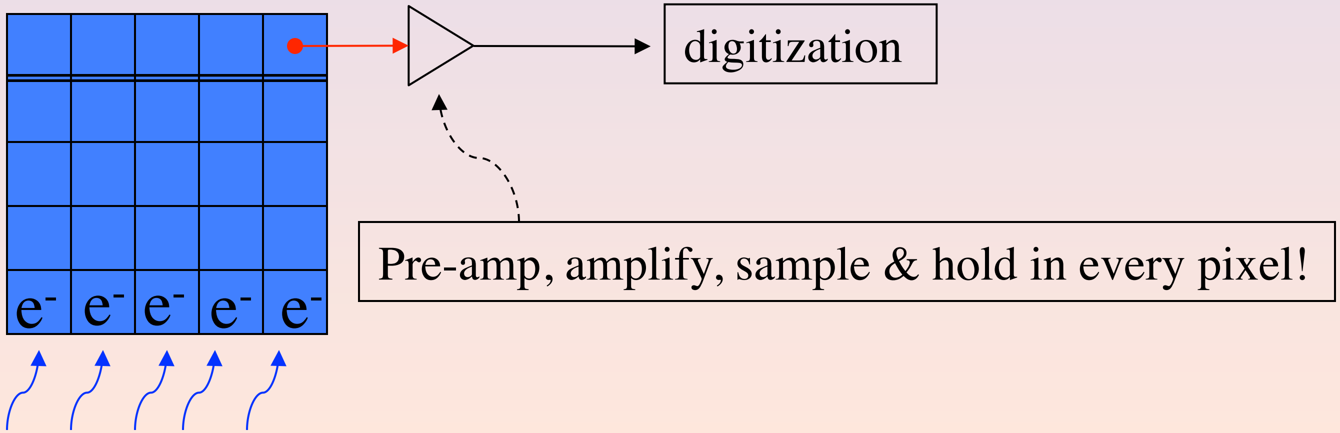
- After charge from each pixel is clocked out, amplified, and sampled, read-out amps are reset to a reference voltage.
- Reset has inherent (kT) noise.
- This is completely eliminated by measuring the voltage difference after reset and after integration (before next reset).





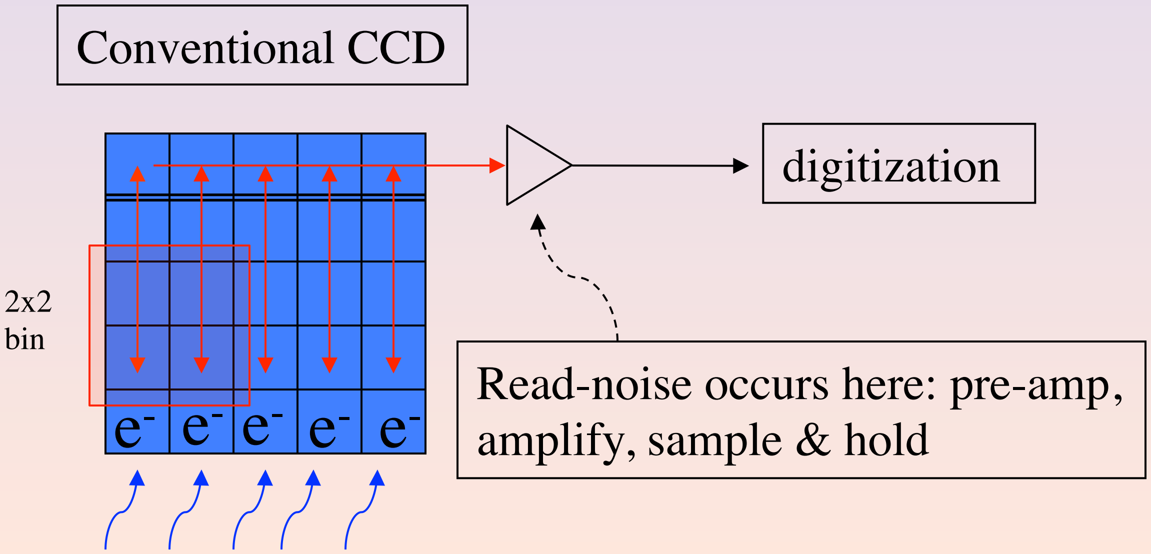
CMOS

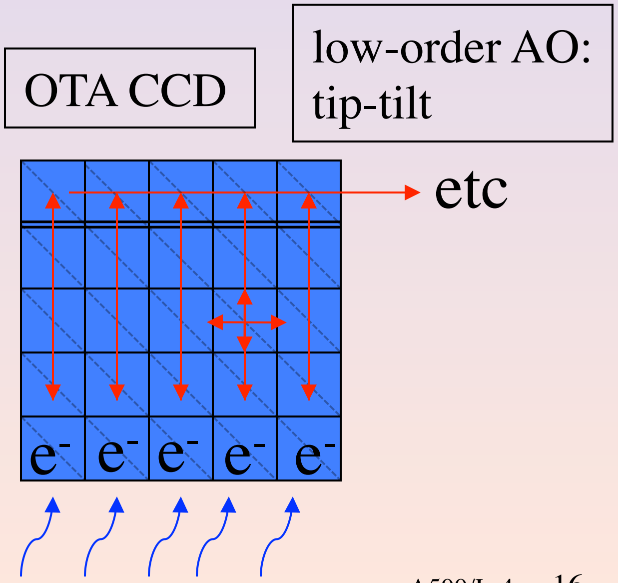
- Complementary Metal Oxide Semicondcutors – complementary pairs of p- and n-type MOSFETs.
- Advantages over CCD (with only p- or n-type): *low-power consumption*
- Allows additional circuitry to be placed in each pixel
  - Every pixel has its own R.O.E. and is directly addressable.
- Leads to < 100% fill-factor of light-sensitive region
  - Can be ameliorated somewhat by micro-lenses but these are lossy too, and scatter
- Gain, bias, and noise non-uniformity add additional calibration demands
  - e.g., fixed-pattern noise and more



CCDs: unusual features

- Non-destructive shifting of charge
  - Drift-scanning: optimizes flatness and efficiency (read-time)
  - Nod-and-shuffle: optimizes flatness and sky-subtraction
  - Frame-transfer: optimizes high-speed photometry
  - On-chip binning: optimizes read-noise
  - Orthogonal-transfer (OTA, e.g., ODI): optimizes image quality





CCD types

- Front-side vs back-side illuminated
- Thinned (back-side) illuminated
- Coated (UV enhanced)
- Deep-depletion (improved red response; decreased blue response)
- High vs low resistivity (improved red response)

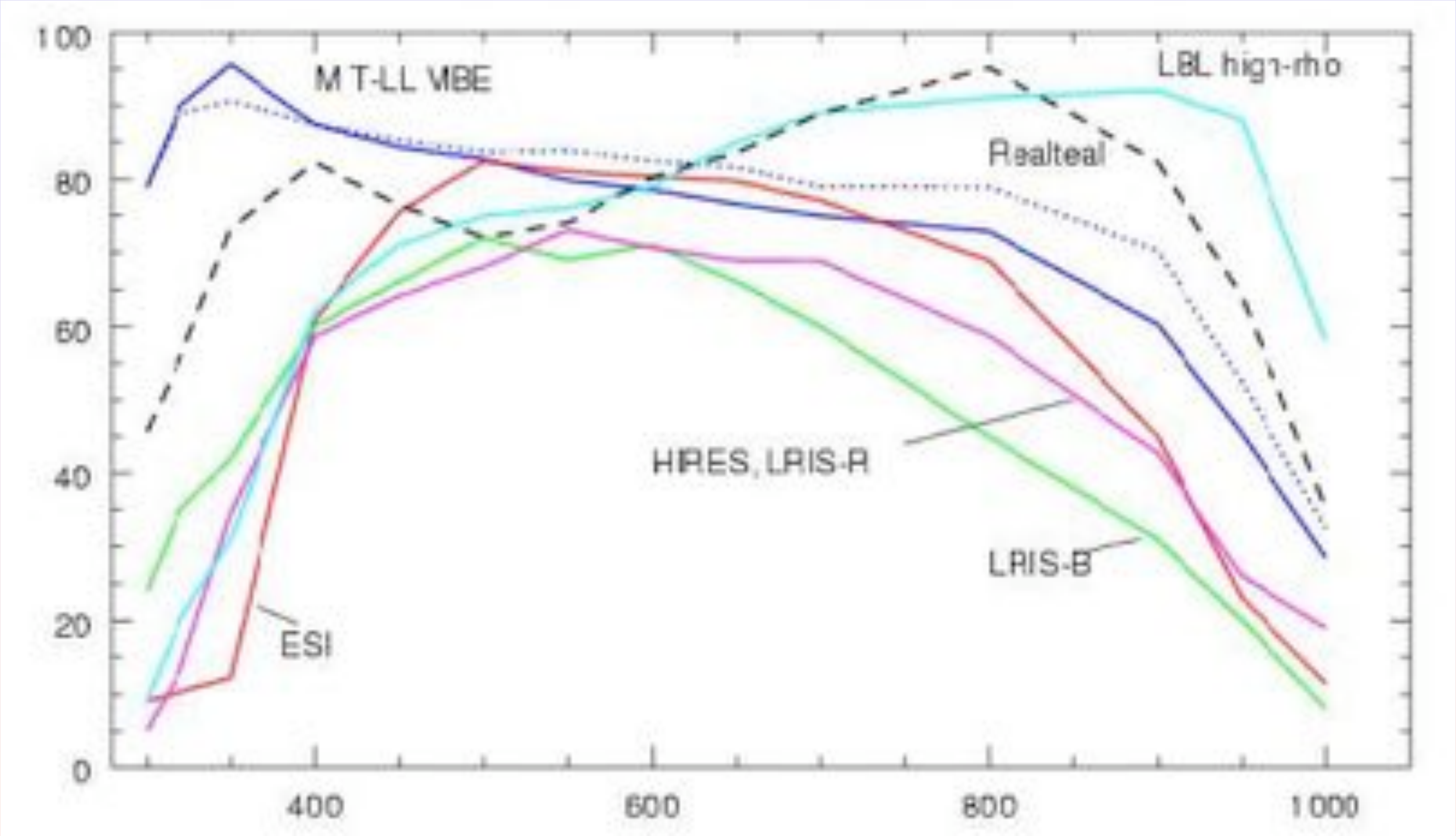
CCD attributes

- Pixel size
- Pixel fill-factor
- Array size
- Array flatness
- Quantum efficiency (QE_λ)
- Dark current
- Charge-transfer efficiency (CTE)
- Electron diffusion (MTF)
- Blooming
- Cosmetics / defects
  - Column defects
  - White and black spots
  - traps
- Amplifiers & electronics
  - How many
  - Read-noise
  - Noise uniformity (btwn amps)
  - Hysterisis / latency
  - Cross-talk and ghosting
  - System noise (RF) pickup
  - Stability (bias drift)

CCD QE_λ

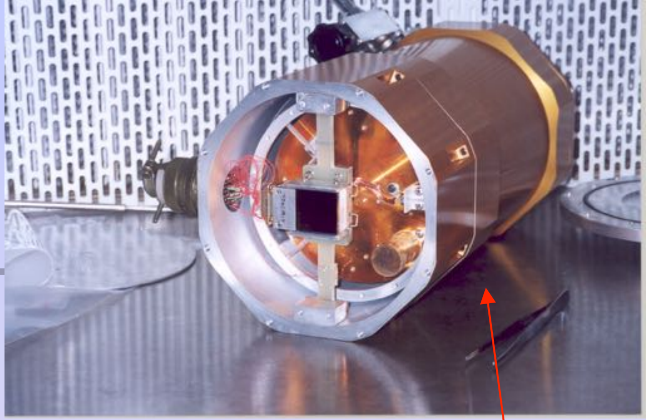
CCDs from Lick Observatory: present and future

QE [%]



wavelength [nm]

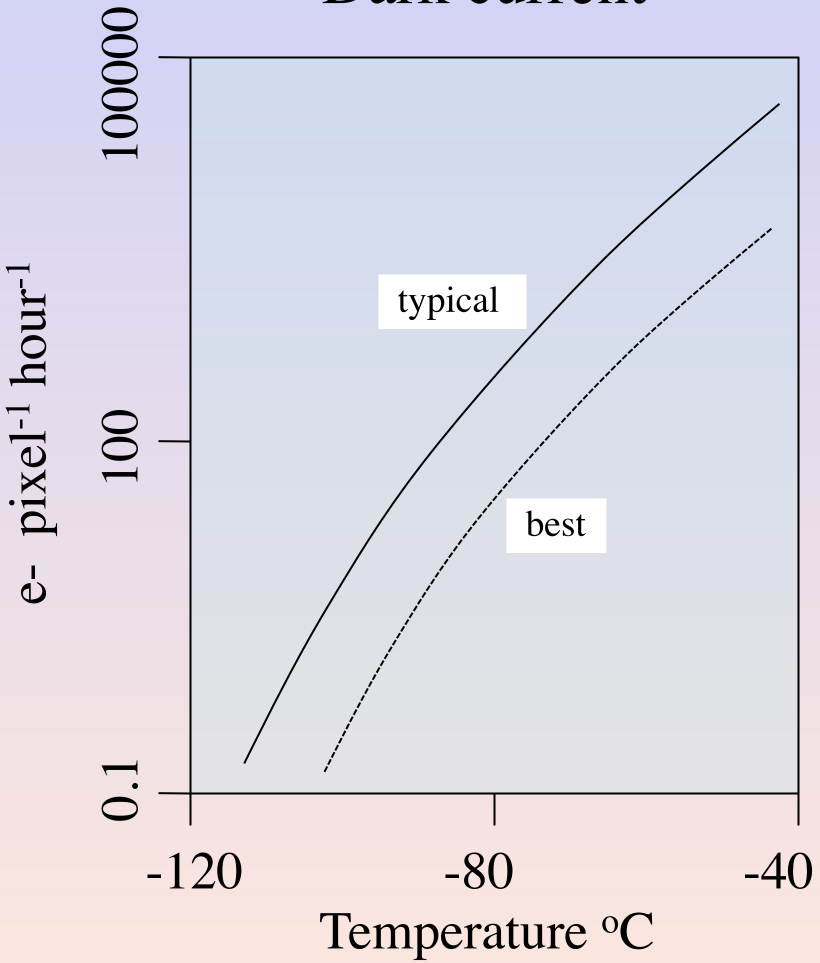
CCD operation



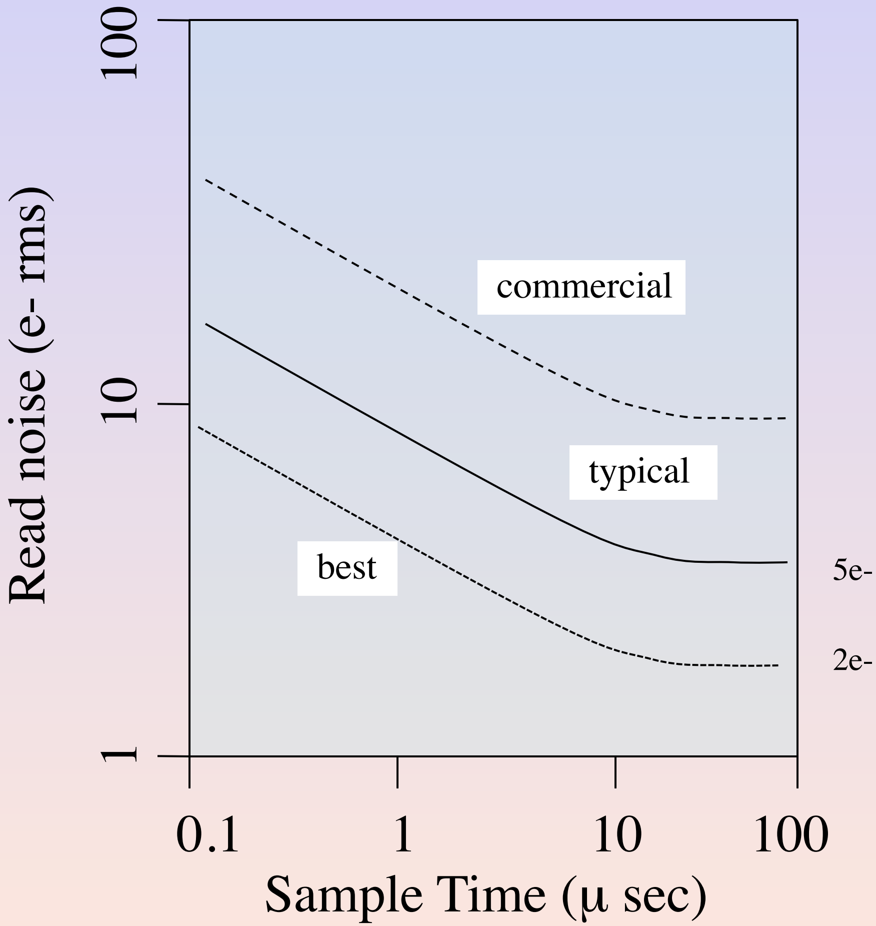
- At room temperature, electrons in high-energy thermal tail of the silicon spontaneously pop up into the conduction band: “dark current.” Cooling the detectors reduced the dark current although at about -120C the quantum efficiency starts to decrease.
- Therefore, CCDs usually are put into dewars with liquid nitrogen cold baths and heaters and the temperature is actively controlled to ~1C.
- Readout speed is typically adjustable--faster readout gives higher <u>readout noise</u> per pixel.

Dark current and Read-noise

Dark current



Read noise



Gain, linearity, and bias

- The potential corrals that define the pixels of the CCD start to flatten as e^- collect. This leads first to saturation, then to e^- spilling out along columns.
- The "inverse gain" is the number of e^- per final "count" post the A/D converter.
- One *very* important possibility for CCDs is to tune the response to be linear.
- An electronic pedestal voltage (bias) is introduced into the read-out electronics to ensure no negative data values occur due to noise. This pedestal has nothing to do with well-depth.

Digital Units

- "Counts'' = ADU = DN

DigitalUnit

Analogue-to-digital unit

- DN is not the fundamental unit, the # of detected electrons is. The "Gain" is set by the electronics.

- Most A/D converters use 16 bits.

  DN from: 0 to (2^{16} - 1) = 65535

  for unsigned, long integers

- Signed integers are nuts: -32735 to +32735

  +/-(2^{15} - 1)

What gain do you want?

Example: LRIS-R has a SITe 24μ-pixel CCD with pixel "wells" that hold ~350,000 e-

- 16-bit unsigned integer A/D saturates at 65525 DN
- Would efficiently maximize dynamic range by matching these saturation levels:

$$\frac{350{,}000}{65{,}535} = 5.3\frac{e-}{DN}$$

- Note, this under-samples the readout noise and leads to "digitization" noise.

Fundamental Performance trades

- Read-time vs Read-noise
- Dark-current vs QE
- Dynamic range vs Well-sampled noise

high-signal limit (→ Dynamic range)

low-signal limit (→ Well-sampled noise)

Signal

- Point source
  - We are measuring photon flux
  - $E_\gamma = f_\gamma A\ t$
- Resolved source
  - We are measuring surface brightness
  - $E_\gamma = I_\gamma A\ \Omega\ t$

S_γ

Signal-to-Noise (S/N)

- Signal = $S_{obj} \bullet \varepsilon \bullet t$

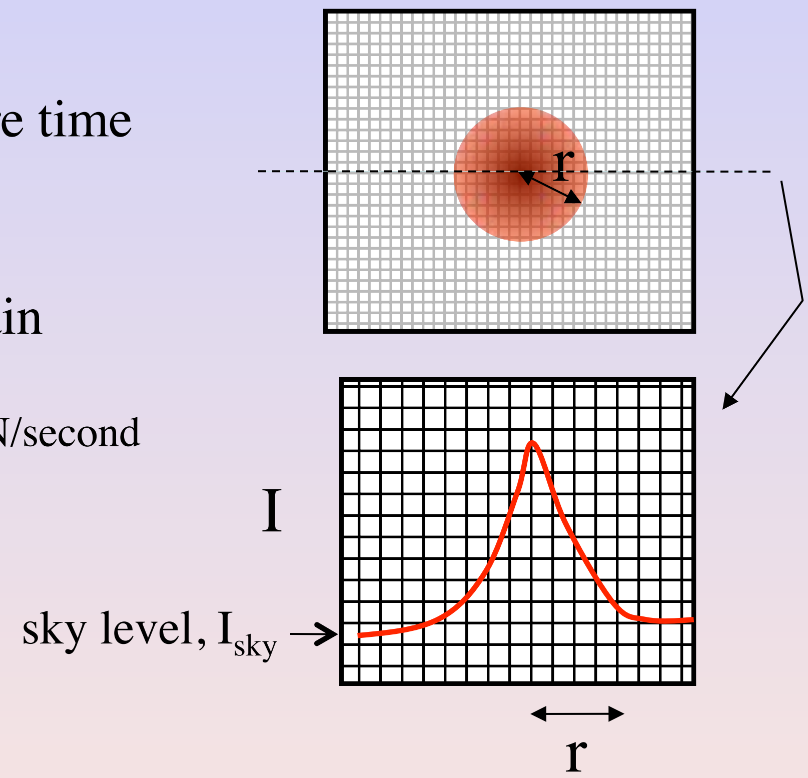
total system efficiency (ε)

exposure time (t)

detected e-/second: $S_{obj} = S_{DET} \bullet \text{gain}$

DN/second (S_{DET})

- Consider the case where we count all the detected e- in a circular aperture with radius r.



Aside: how big an area do we want to integrate over?

Noise Sources

$\sqrt{S_{obj} \cdot \varepsilon \cdot t}$ $\Rightarrow$ shot noise from source

$\sqrt{I_{sky} \cdot \varepsilon \cdot t \cdot \pi r^2}$ $\Rightarrow$ shot noise from sky in aperture

$\sqrt{RN^2 \cdot \pi r^2}$? $\Rightarrow$ readout noise in aperture

$\sqrt{\left[RN^2 + (0.5 \times \text{gain})^2\right]} \cdot \sqrt{\pi r^2}$ $\Rightarrow$ more general RN

$\sqrt{\text{Dark} \cdot t \cdot \pi r^2}$ $\Rightarrow$ shot noise in dark current in aperture

$S_{obj} \cdot \varepsilon = \text{e}^{-}/\text{sec}$ from the source

$I_{sky} \cdot \varepsilon = \text{e}^{-}/\text{sec/pixel}$ from the sky, $S_{sky} = I_{sky} \cdot \pi r^2 = I_{sky} \cdot n_{pix}$ $\longleftarrow$ NB

RN = read noise (as if RN^2 e^{-} had been detected)

Dark = e^{-} / second/pixel

S/N for object measured in aperture with radius r:

$n_{\rm pix}$ = # of pixels in the aperture = πr^2

Noise from the dark current in aperture

$$\frac{\text{Signal}}{\text{Noise}} = \frac{S_{obj}\cdot\varepsilon\cdot t}{\left[\underbrace{S_{obj}\cdot\varepsilon\cdot t}+\underbrace{I_{\rm sky}\cdot\varepsilon\cdot t\cdot n_{\rm pix}}+\underbrace{\left(RN+\frac{\rm gain}{2}\right)^2\cdot n_{pix}}+\overbrace{\text{Dark}\cdot t\cdot n_{\rm pix}}\right]^{\frac{1}{2}}}$$

$\sqrt{(S_{obj}\cdot\varepsilon\cdot t)^2}$

Noise from sky *e*- in aperture

Readnoise in aperture

All the noise terms added in quadrature

Note: always calculate in *e*- *why?*

What is ignored in this S/N eqn?

- Explicit inclusion of collecting aperture
- Break-out of terms that go into total system efficiency (starting from the top of the atmosphere)
- Bias level/structure correction and errors
- Flat-fielding correction and errors
- Charge Transfer Efficiency (CTE) 0.99999/pixel transfer
- Non-linearity when approaching full well
- Scale changes in focal plane
- Interpolation errors and correlation

S/N regimes

- Two basic regimes:
  1. Photon-limited (shot-noise from source+sky photons)
  2. Detector-limited (read-noise)
- In photon-limited case, two important sub-regimes
  a. Source-limited
  b. Sky-limited

S/N regimes: limiting cases

Let's assume CCD with Dark=0, well sampled read noise.

$$S/N = \frac{S_{obj} \cdot \varepsilon \cdot t}{\left[S_{obj} \cdot \varepsilon \cdot t + I_{\mathrm{sky}} \cdot \varepsilon \cdot t \cdot n_{\mathrm{pix}} + (RN)^2 \cdot n_{pix}\right]^{\frac{1}{2}}}$$

Note: seeing or source-size comes in with n_{pix} term

<u>1a. Bright Sources:</u> $(S_{obj}\,\varepsilon\, t)^{1/2}$ dominates noise term

$$\mathrm{S/N} \approx \frac{S_{obj}\,\varepsilon\, t}{\sqrt{S_{obj}\,\varepsilon\, t}} = \sqrt{S_{obj}\,\varepsilon\, t} \propto t^{\frac{1}{2}}$$

S/N limiting cases (*contd*)

$$S/N = \frac{S_{obj} \cdot \varepsilon \cdot t}{\left[S_{obj} \cdot \varepsilon \cdot t + S_{\text{sky}} \cdot \varepsilon \cdot t \cdot n_{\text{pix}} + (RN)^2 \cdot n_{pix}\right]^{\frac{1}{2}}}$$

<u>1b. Sky Limited:</u> $(\sqrt{I_{sky}\,\varepsilon\, t} > 3 \times RN)$

$$\text{S/N} \propto \frac{S_{obj}\,\varepsilon\, t}{\sqrt{n_{pix} I_{sky}\,\varepsilon\, t}} \propto t^{\frac{1}{2}}$$

<u>2. Read-noise Limited:</u> $(\sqrt{I_{sky}\,\varepsilon\, t} < 3 \times RN)$

$$\text{S/N} \propto \frac{S_{obj}\,\varepsilon\, t}{\sqrt{n_{pix} RN^2}} \propto t$$

Note: seeing comes in with n_{pix} term

What does this imply about exposure time?

DQE

- DQE is often defined as the *effective quantum efficiency* of a CCD relative to an ideal detector with no read-noise. In the source-limited regime, ignoring dark-current:

$$DQE = QE \,/ \left[1 + \frac{RN^2}{QE \cdot S_{obj} \cdot t}\right]$$

  where QE is the CCD quantum efficency.
- This can be generalized for any noise-regime, and including dark-current.
- A related concept is the *effective system efficiency*, DQE_{sys}, of which CCD QE is only one part.