

Lecture 14: Noise and Dynamic Range

Yuhao Zhu

<http://yuhaozhu.com>
yzhu@rochester.edu

CSC 259/459, Fall 2024
Computer Imaging & Graphics

The Roadmap

Theoretical Preliminaries

Human Visual Systems

Digital Camera Imaging

Modeling and Rendering



Photographic Optics

Image Sensor

Image Signal Processing

Image/Video Compression

Immersive Content

Noise



Clean



Noisy



Noise

Without noise, raw pixel value is proportional to light luminance

- ignore ADC quantization error, which itself also introduces noise

Two forms of noise:

- **Fixed Pattern Noise**; doesn't vary from capture to capture but varies from pixel to pixel
- **Temporal Noise**; does vary from capture to capture

Temporal noise:

- Charge domain noise: the number of collected charges (electrons) is noisy
- Voltage domain noise: the voltage reading (converted from charges) is noisy

Any amplification (gain) **amplifies the noise** accumulated before the gain.

Temporal Noises and SNR

Temporal Noise

Varies from shot to shot, but could be observed from neighboring pixels in one shot if scene has relatively uniform luminance.

Charge-domain noise: the noise is represented in the collected charges

- **Photon Shot Noise**: due to random arrival of photons.
- **Dark Current Noise** (electronic shot noise): random arrival of (temperature-dependent) thermal generation of electrons even without any incident photons.

Voltage-domain noise: noise in circuits that convert charges to voltage; represented in the voltage measurement

- a.k.a. **Read Noise** (as the noise is generated in the process of reading the charges)

Observing Temporal Noise Spatially

Technically, temporal noise should be observed by comparing consecutive frames, which requires taking multiple frames.

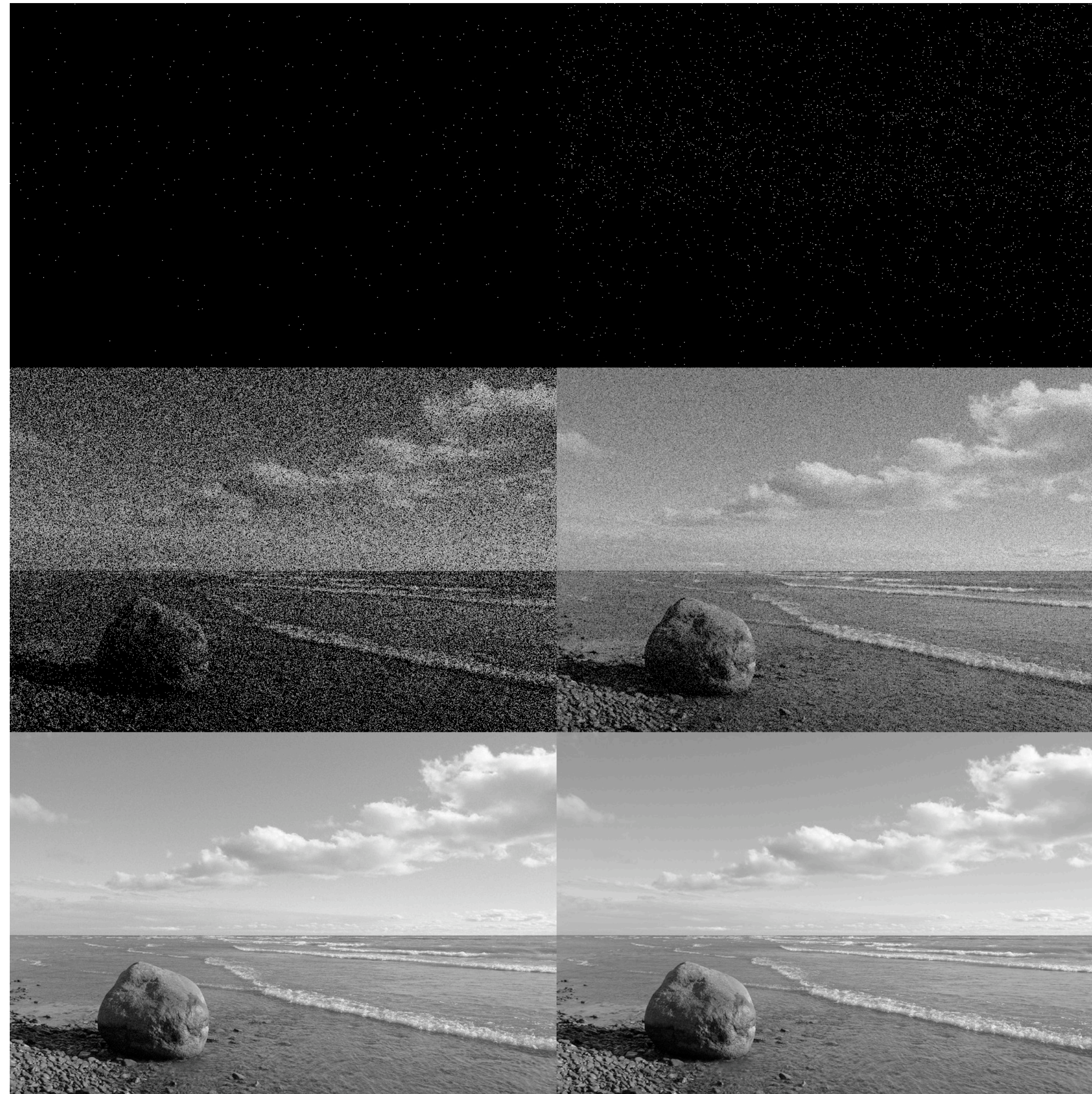
But if the scene has ~uniform illumination we can observe temporal noise **spatially** by analyzing pixels in a single image.

- Think of M pixels in an image as taking M captures at a single pixel

Or, for a real scene which is unlikely uniformly illuminated, we can analyze *neighboring* pixels, which likely are similarly illuminated.

- such as the example on the next slide

Photon Shot Noise Example



Photon shot noise is less significant with longer exposure time (or equivalently brighter scenes)



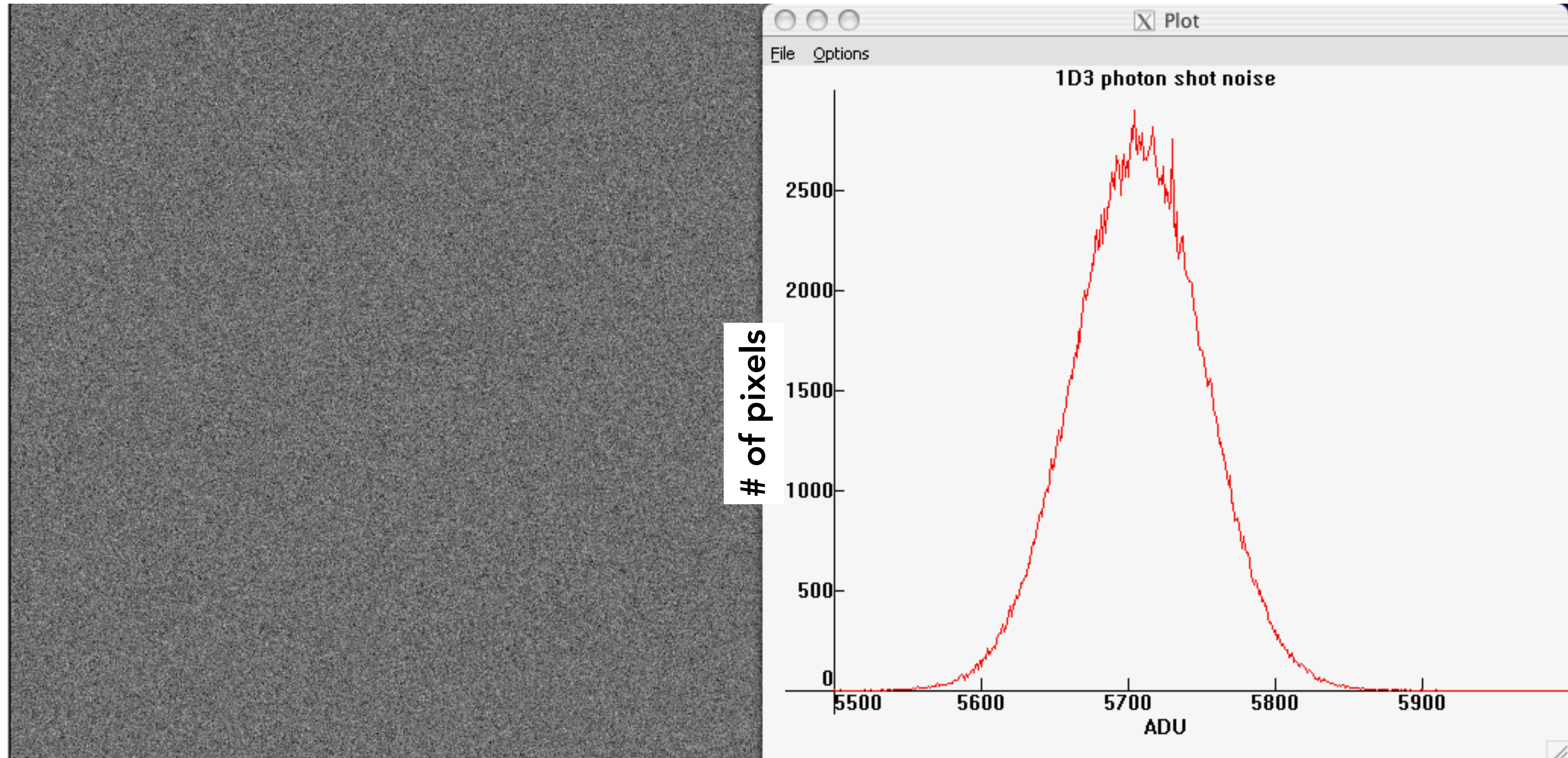
Photon Shot Noise

Photon emission is a random process. A light source with constant power **on average** emits the same amount of photons per second, but in any given constant period, the **absolute** amount of photons emitted will vary.

The relative deviation away from average is less significant for bright scenes (lots of photons on average) and is more significant for dark scenes (few photons on average).

Photon Shot Noise Example

Histogram of raw pixel value distribution in an image of sky (Canon 1D3, green channel).

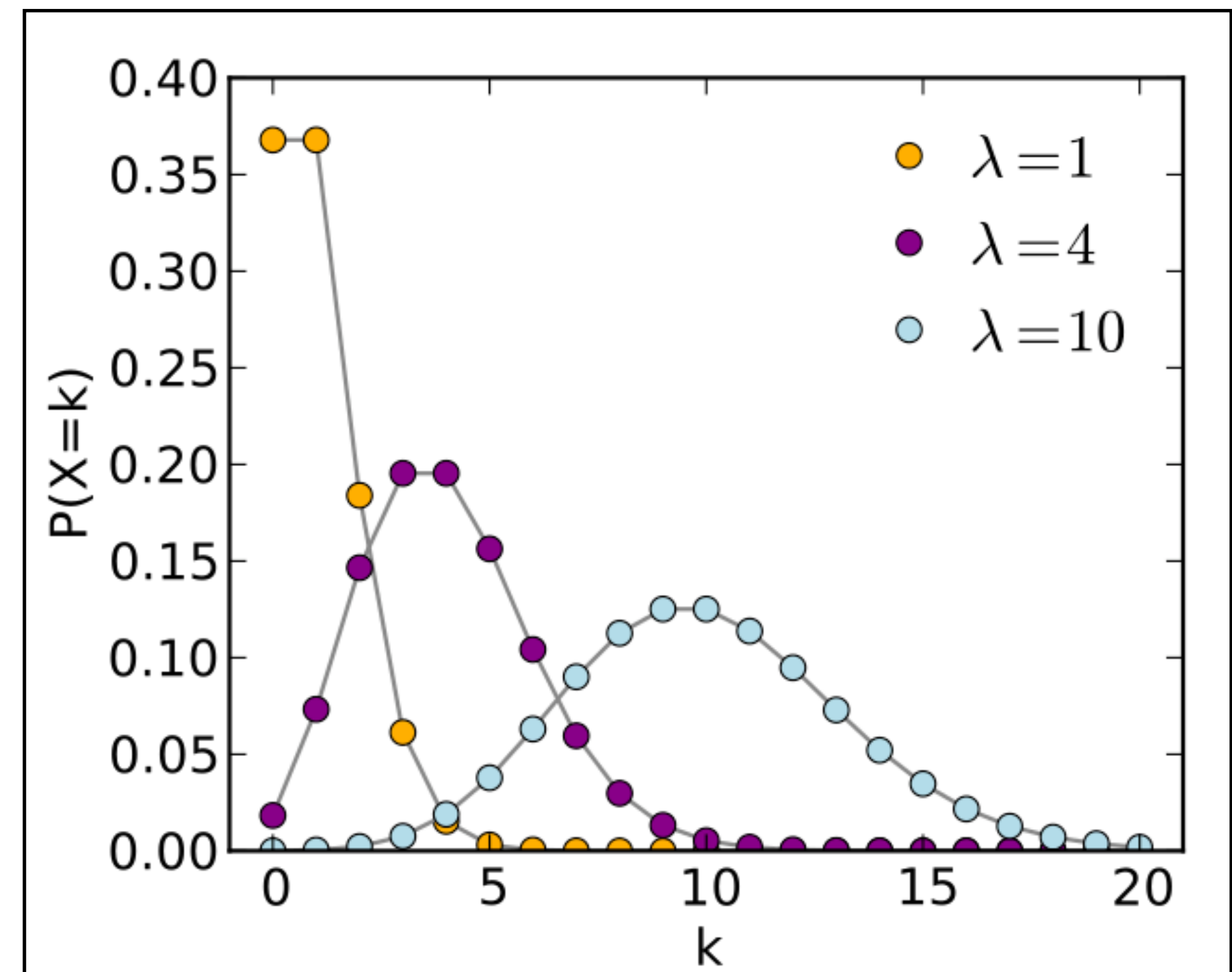


Modeling Photon Shot Noise

The distribution of the number of photons received by a pixel is governed by the **Poisson distribution**.

The probability that k photons actually arrive at a pixel when on average λ photons arrive (per unit time).

$$p(k; \lambda) = \frac{\lambda^k \times e^{-\lambda}}{k!}$$

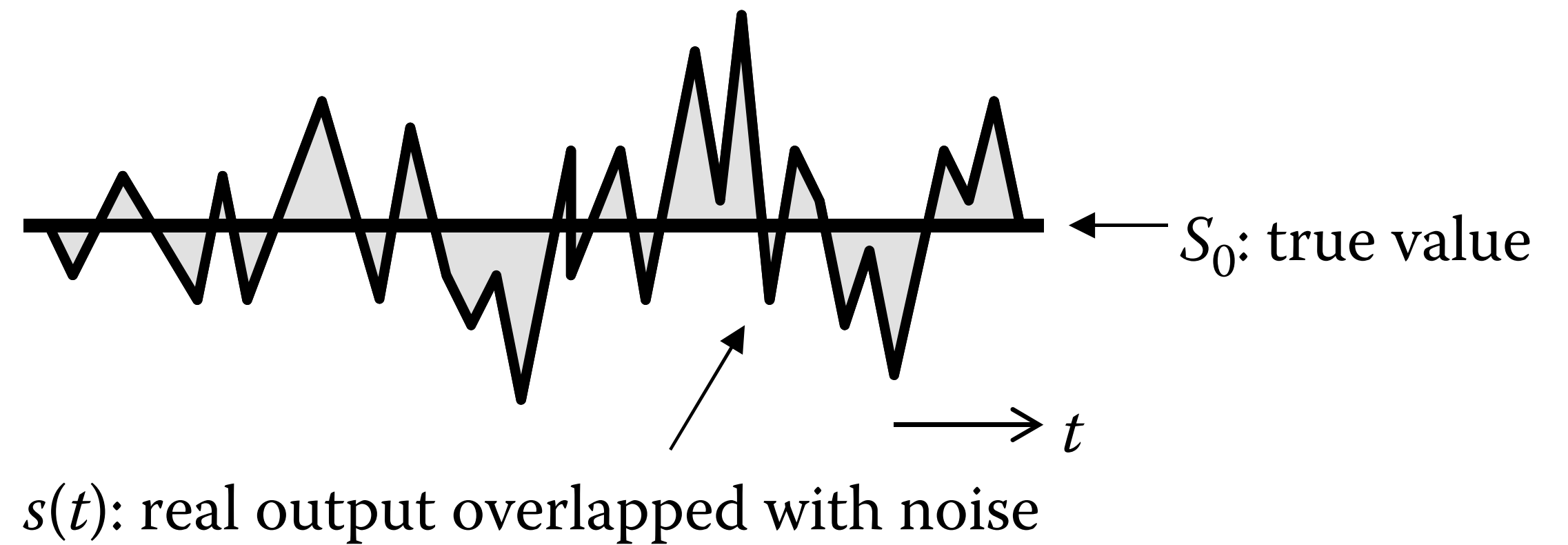


Signal to Noise Ratio (SNR)

Signal: mean value

Noise: Std. of value

SNR = Signal / Noise



Signal to Noise Ratio (SNR)

Signal: mean value

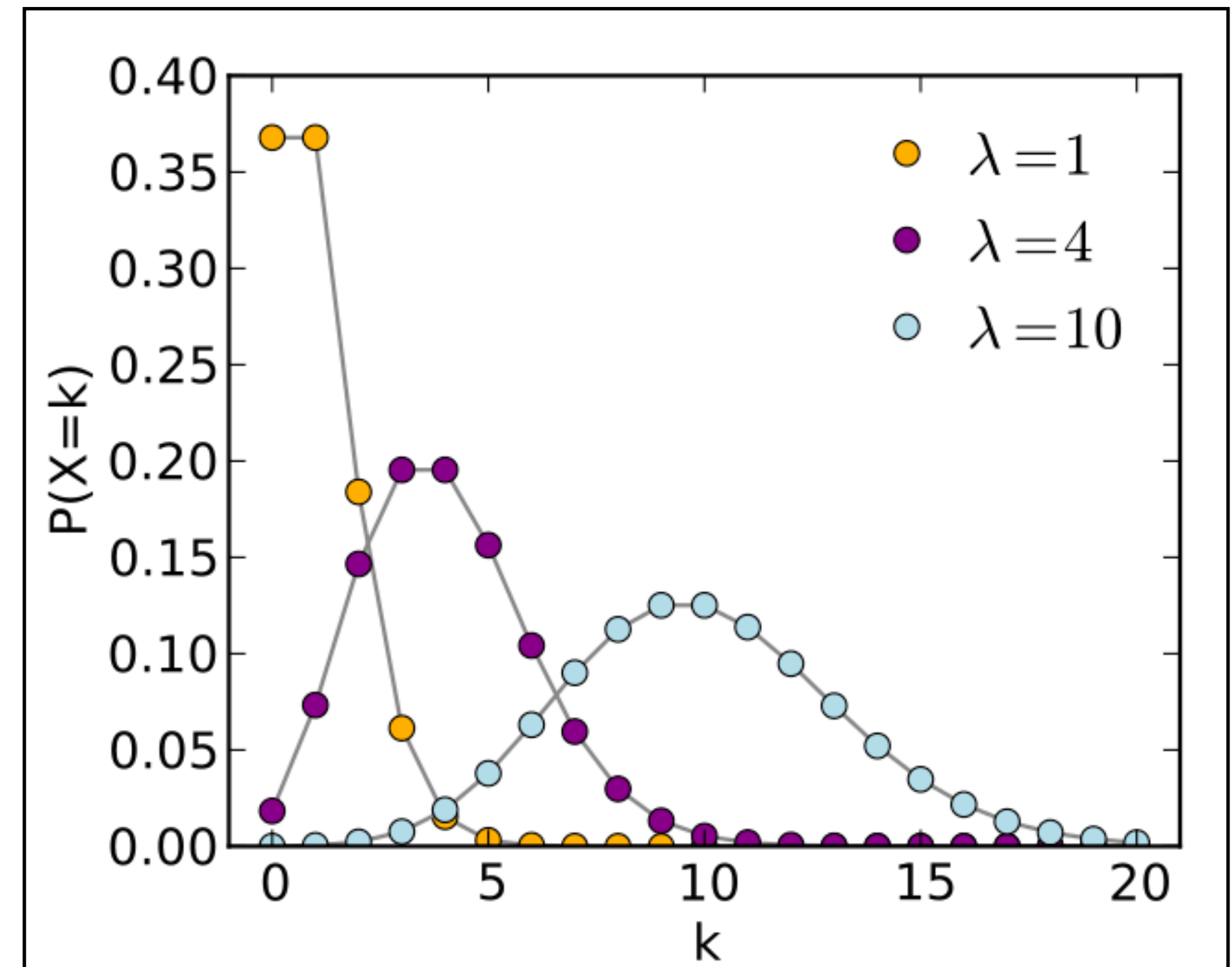
Noise: Std. of value

For a Poisson distribution:

- Mean μ is λ
- Variance is λ
- Standard deviation σ is $\sqrt{\lambda}$

$$SNR = \frac{\mu}{\sigma} = \sqrt{\lambda}$$

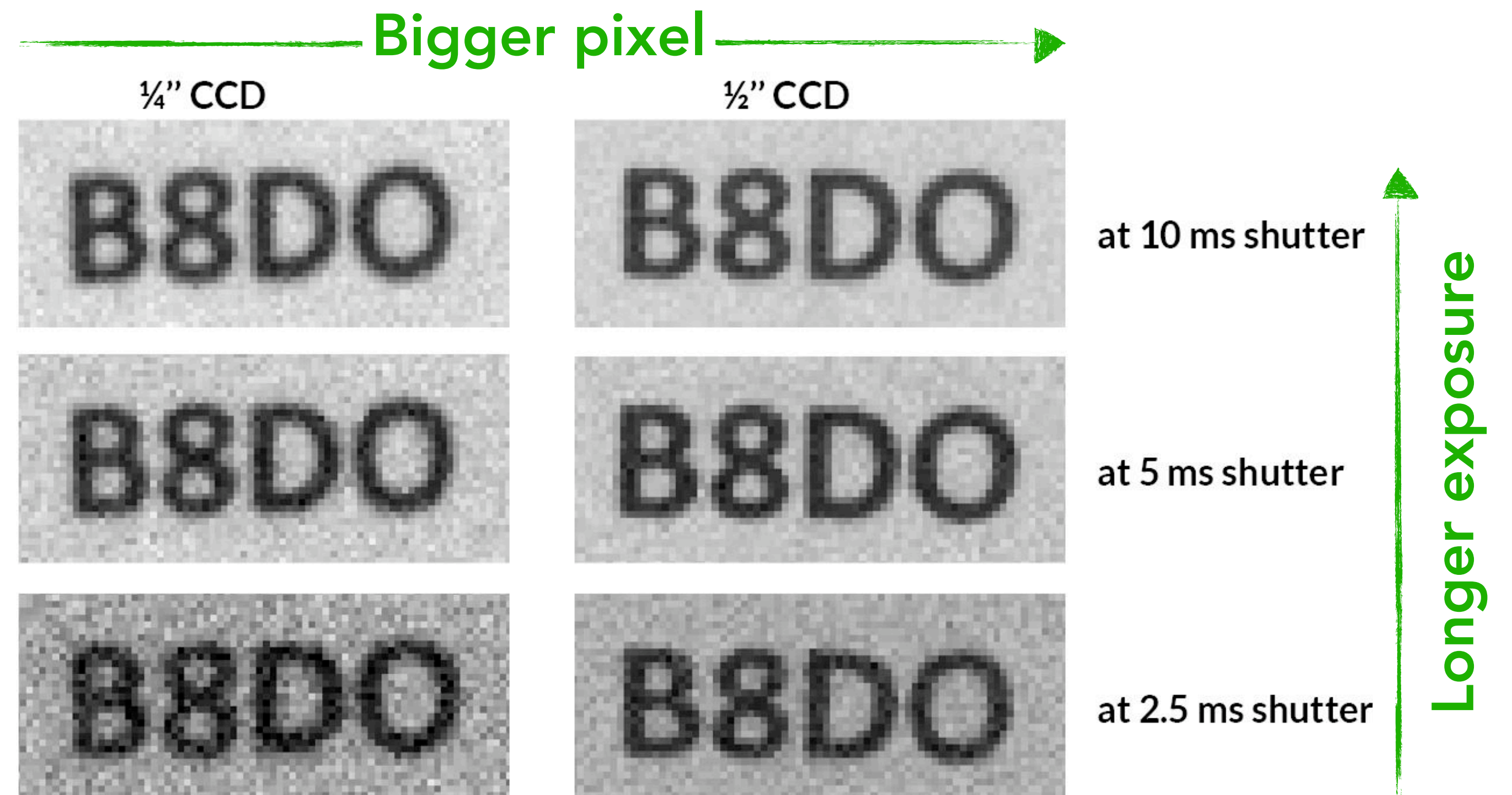
$$p(k; \lambda) = \frac{\lambda^k \times e^{-\lambda}}{k!}$$



What Dictates SNR?

When a pixel receives more photons (e.g., brighter scene, bigger pixel, longer exposure), the SNR is higher, i.e., less noisy.

$$SNR = \frac{\mu}{\sigma} = \sqrt{\lambda}$$

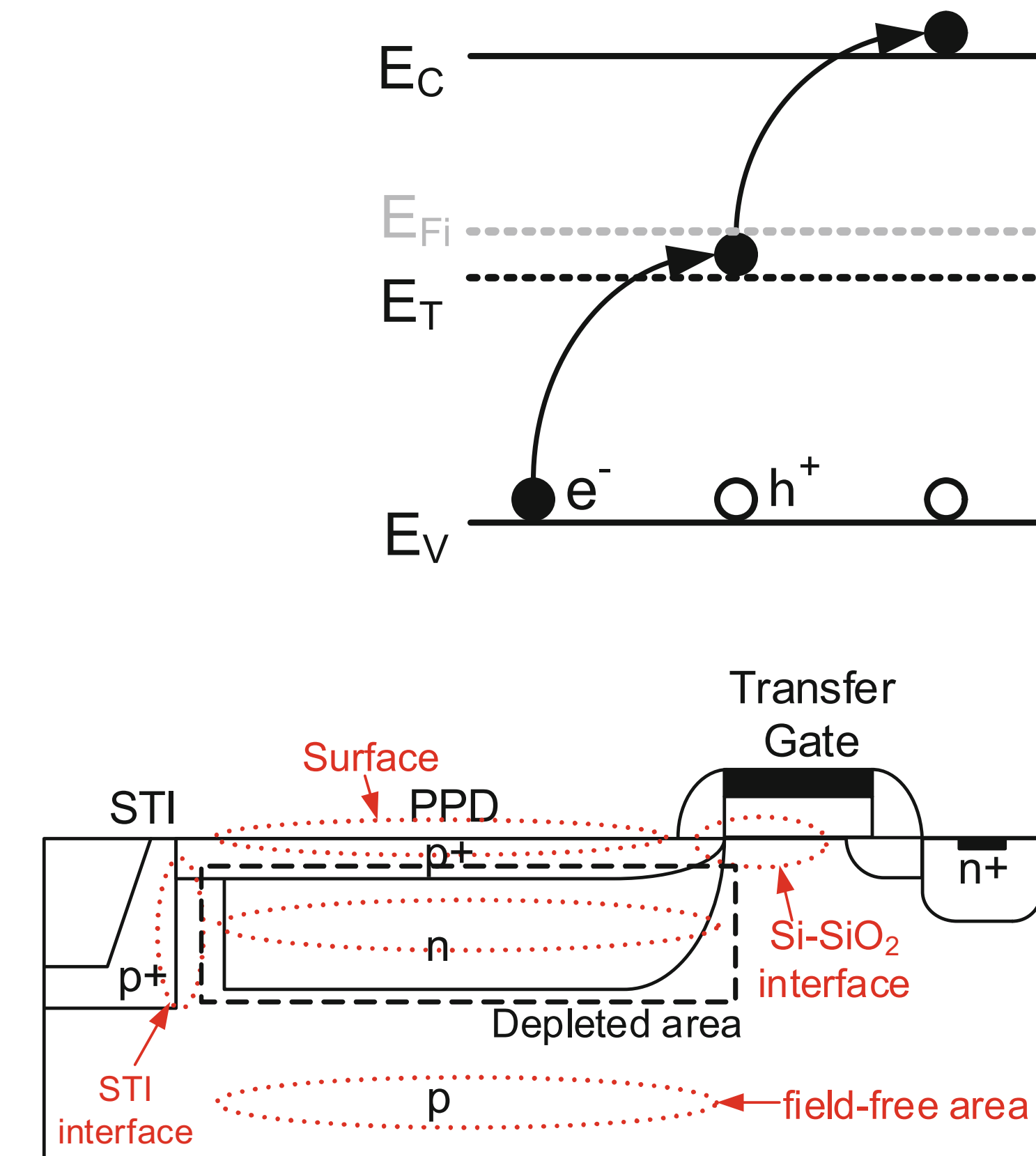


Dark Current Noise (Electronic Shot Noise)

As temperature increases, electrons are dislodged even without any incident photon (unless sensor cooled to absolute zero temperature).

Dark current adds a pedestal offset to the actual electron read-out. But the offset is not constant across time — hence “noise”.

- If the offset was constant, the readings would still be incorrect but won't constitute noise (since the offset is not random).



Dark Current Noise (Electronic Shot Noise)

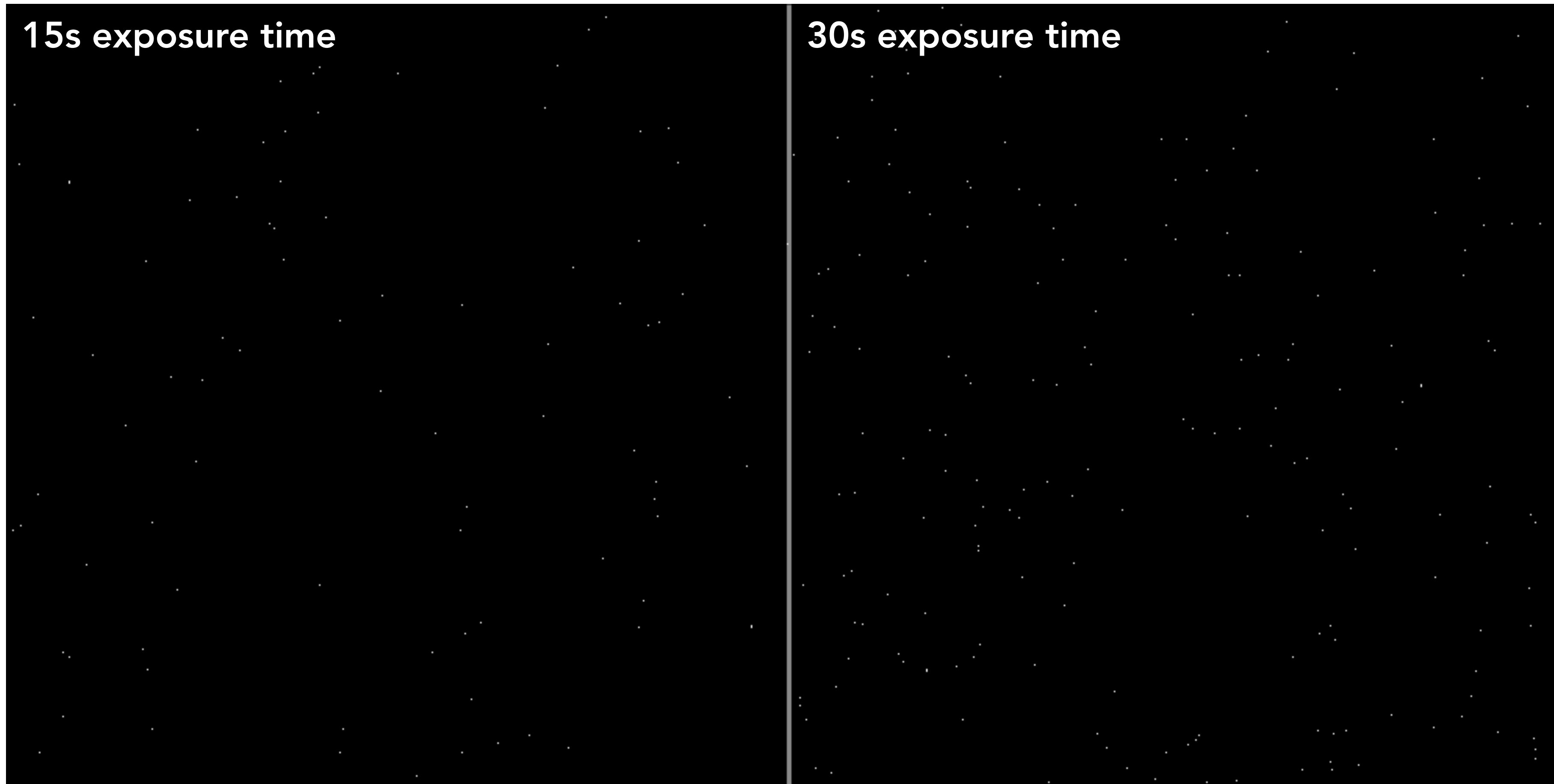
Dark current I_d represents the **average** (expected value) # of dark electrons/sec/pixel, which follows Poisson distribution (similar to photon shot noise).

Over an exposure time t , average # of electrons (collected at each pixel) contributed by dark current is $I_d \times t$.

So the dark current **noise** is ***sqrt*** ($I_d \times t$)

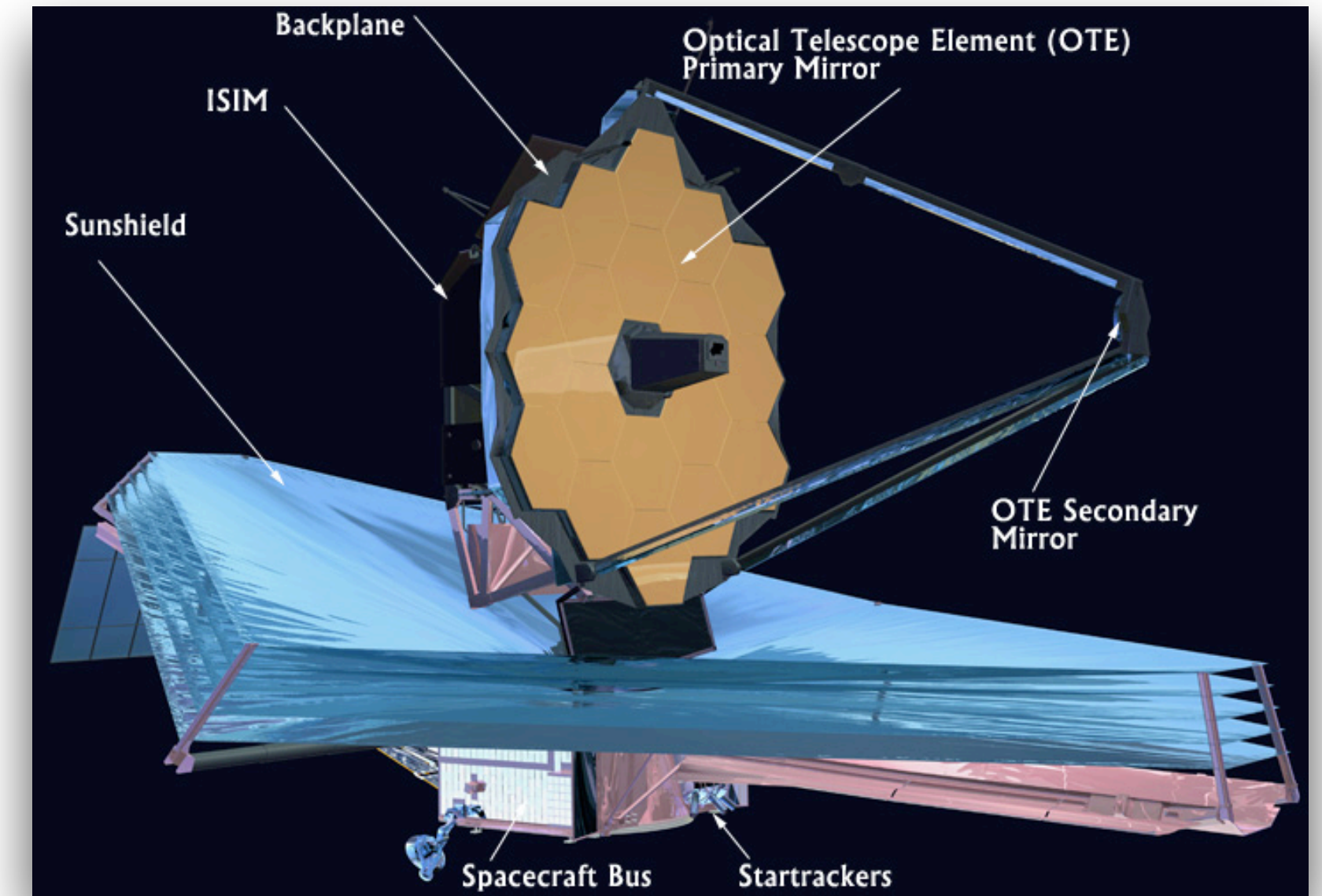
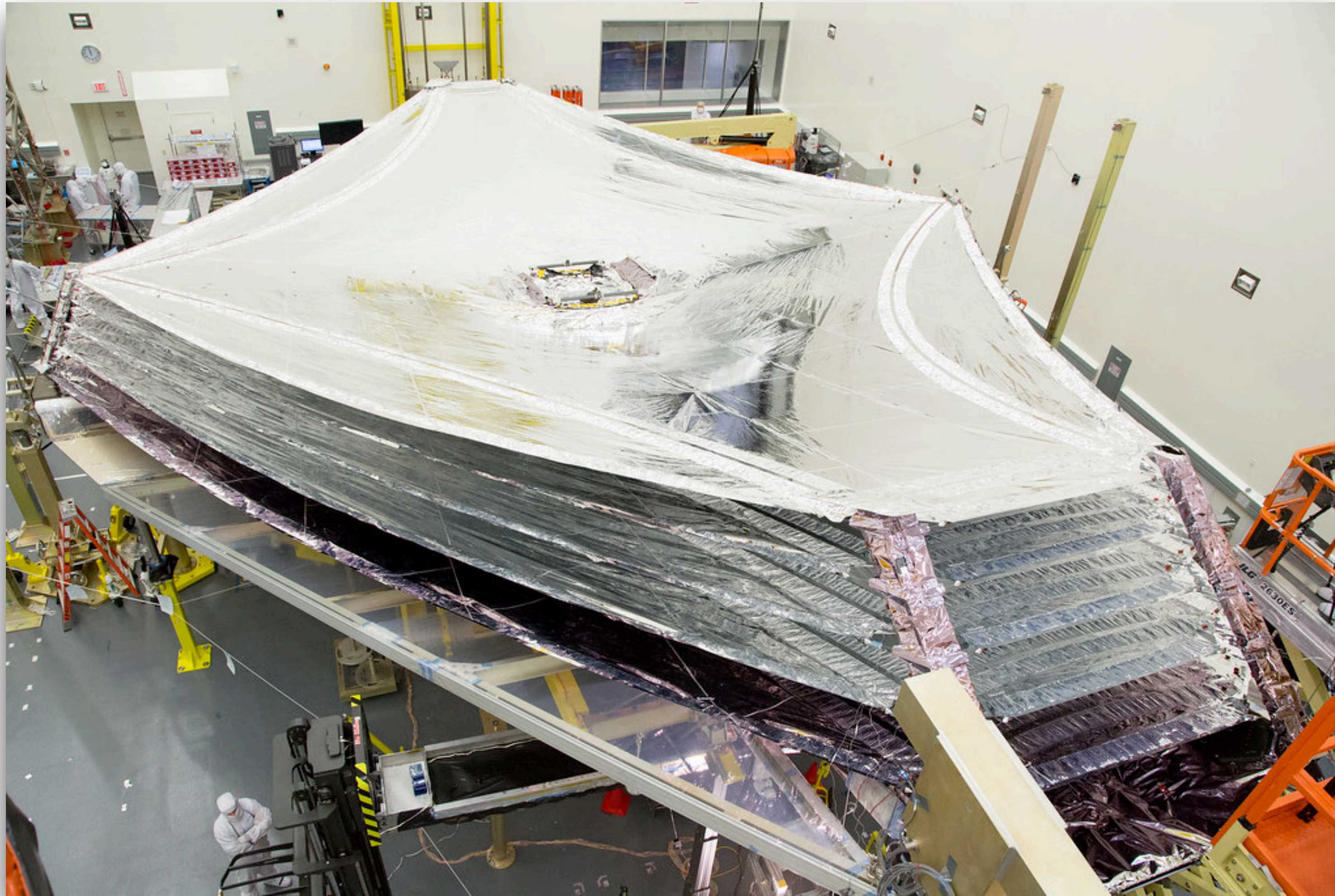
- because dark current follows Poisson distribution
- I_d increases with the temperature
- therefore dark current noise increases with temperature and t .
- I_d is also spatially non-uniform: a source of FPN.

"Hot" Pixels From Dark Current Noise



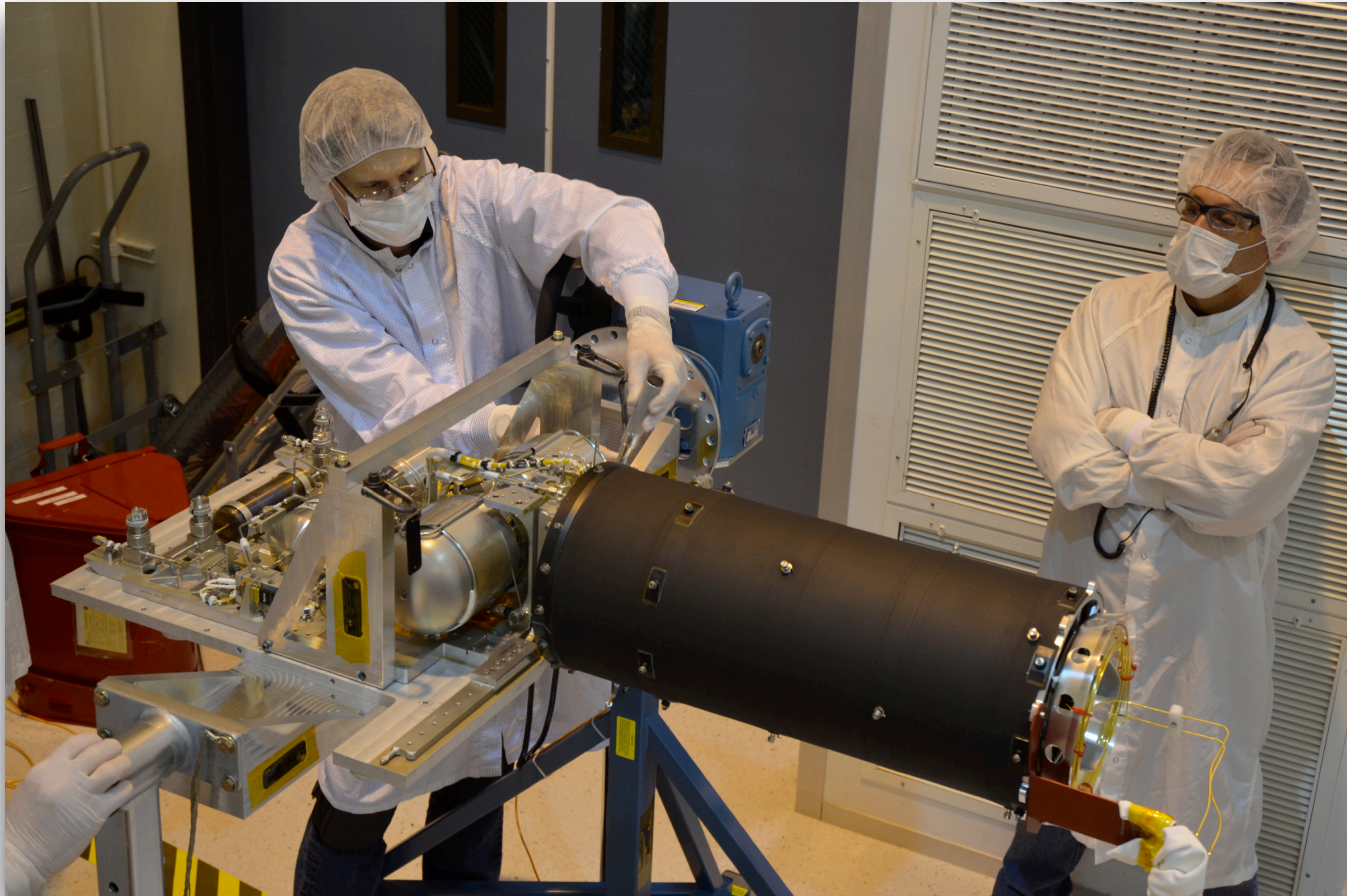
Usually dark current noise is negligible, since dark current is very low (e.g., 50 e⁻/s; c.f., full well capacity > 10ke⁻), but can be significant if long exposure time is required, e.g., astrophysical imaging.

Cooling in James Webber Space Telescope



Passive cooling: sunshield to cool three of the four cameras (37 K)

Cooling in James Webb Space Telescope



Active cooling: cryocooler (7 K)

Read Noise

The circuits that convert photons to digital values suffer voltage fluctuation

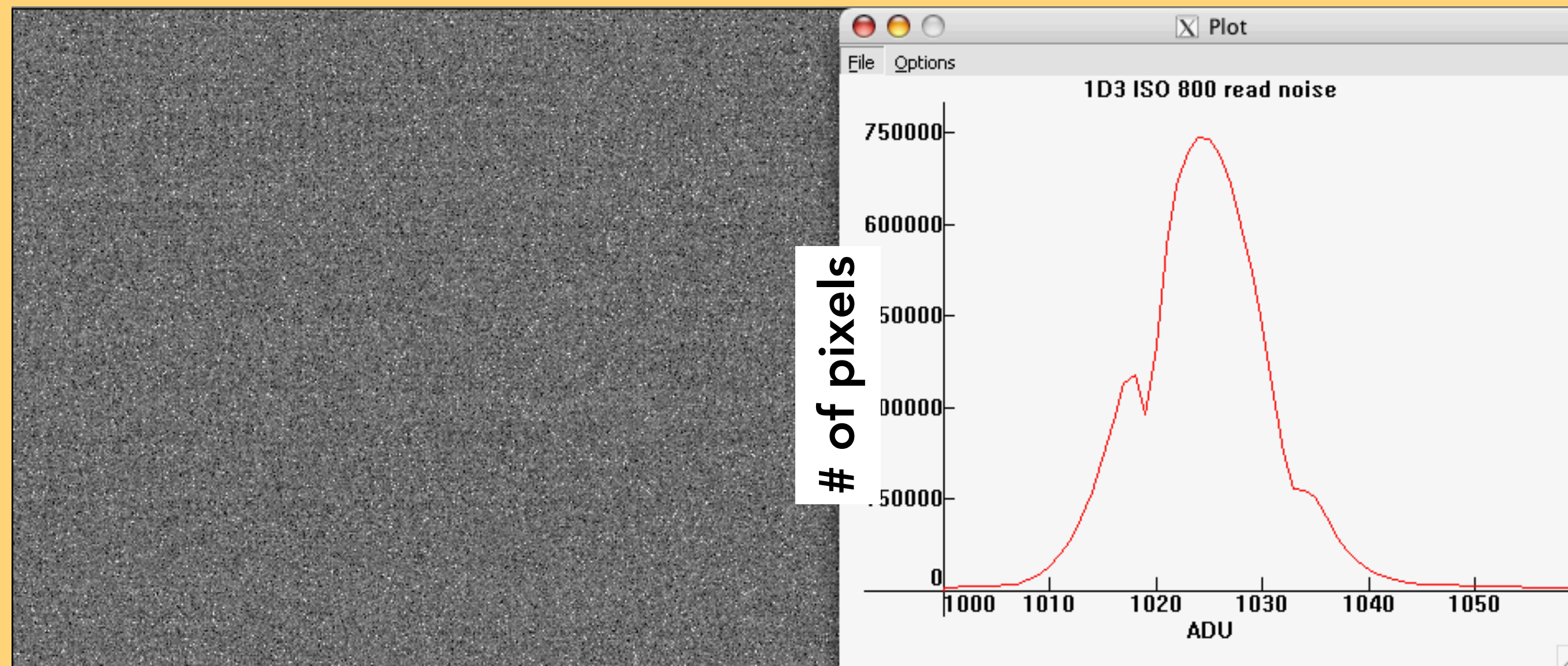
- Remember eventually we read the voltage, not the photon count, so voltage fluctuation in the circuit introduces noise.
- Could be thermal-induced (Johnson–Nyquist noise, which is fundamental to all circuits and is called kTC noise when manifested on capacitors), 1/f noise, or burst noise.
- Sources: various amplifiers (e.g., FD, SF, other gain controls), ADC (with additional quantization errors), reset (of FD; can be suppressed by CDS), CDS circuit itself, etc.

Read noise can be both positive and negative, since voltage fluctuation can be both positive and negative.

Read noise is modeled as a **0-mean Gaussian distribution**.

Estimating Read Noise

Could be measured by reading a bias frame, which is an image captured at dark (lens cap on or shutter closed) with 0 exposure time (or use highest possible shutter speed), i.e., no light gets to the photodiodes and dark current is negligible, so the major noise source is read noise (ignore FPN for now).



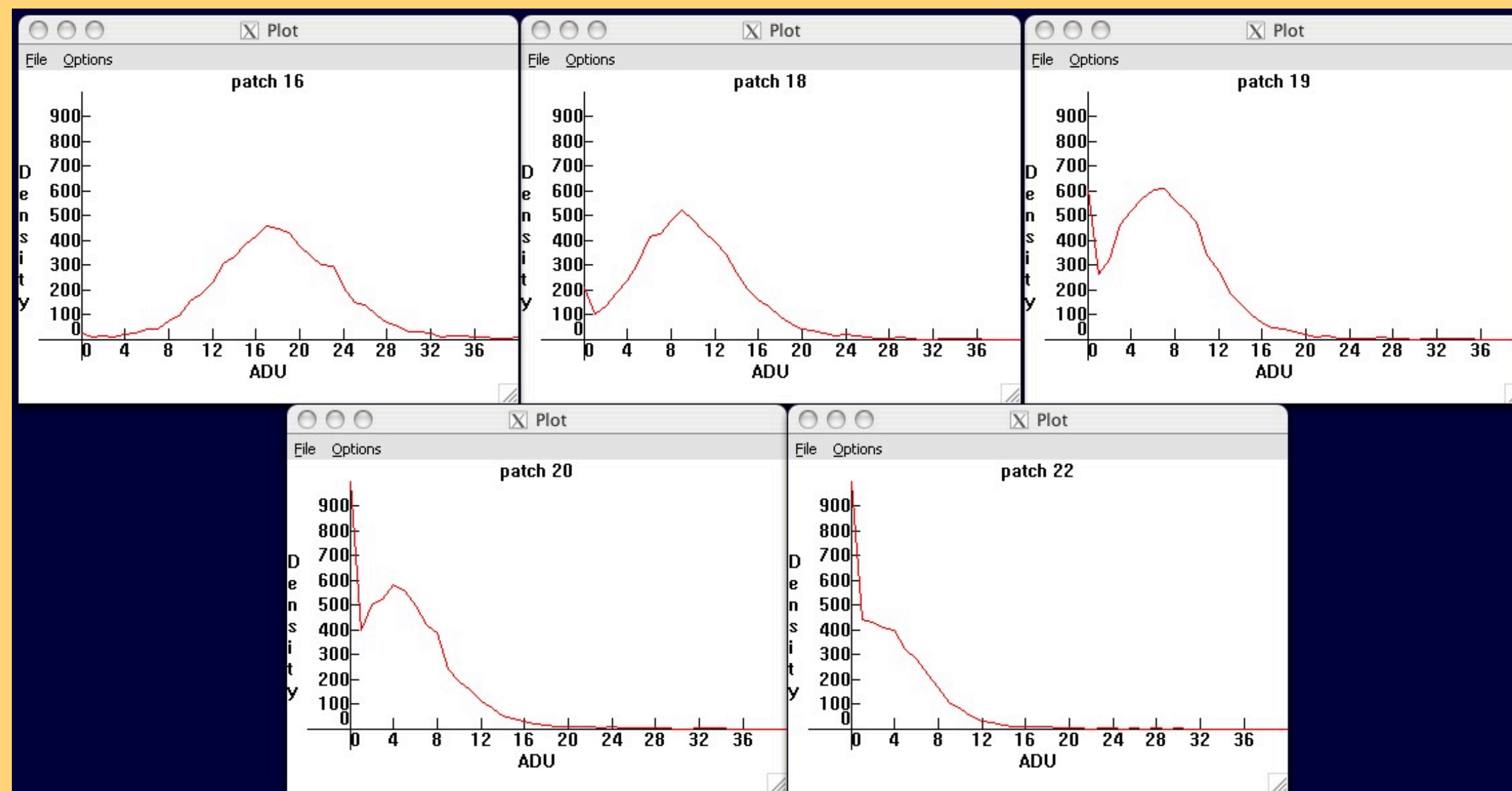
Raw pixel value histogram of Canon 1D3 at ISO 800.

* Again, read noise should technically be observed temporally, but since scene is uniform (dark), noise across pixels spatially equates temporal noise, i.e., standard deviation of pixel values is the σ of read noise.

Three Slight Issues

First, signal voltage could be negative (recall: read noise is a zero-mean Gaussian), which will be clipped to 0 by the ADC, contaminating analysis.

- A bias voltage is usually added to the signal voltage (e.g., in the previous slide). The bias is one reason why dark frame pixels are not zero (even without any noise present).



Pixel histogram at different illumination levels of Nikon D300, which doesn't add ADC bias.

Three Slight Issues

First, signal voltage could be negative (recall: read noise is a zero-mean Gaussian), which will be clipped to 0 by the ADC, contaminating analysis.

- A bias voltage is usually added to the signal voltage (e.g., in the previous slide). The bias is one reason why dark frame pixels are not zero (even without any noise present).
- Black level raw value is stored as metadata (part of EXIF tag) of a raw image. It can be read through raw processing APIs (e.g., rawPy) and must be subtracted from the raw value of each pixel by the raw image signal processing pipeline (usually called black-level correction/compensation). We will see this in the ISP programming assignment.

Three Slight Issues

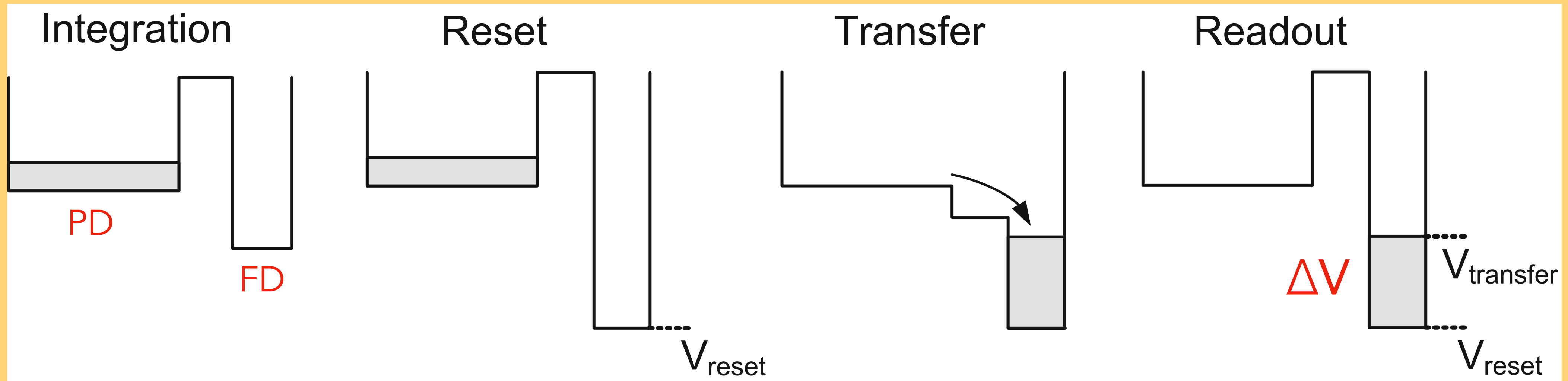
First, signal voltage could be negative (recall: read noise is a zero-mean Gaussian), which will be clipped to 0 by the ADC, contaminating analysis.

Second, the read noise distribution should be measured for each ISO setting, which affects gains, which amplify noise.

Third, it also assumes that no FPN is present, which is not true.

- We will later see a more sophisticated method to estimate read noise isolated from FPN.

Reset Noise: a Special Read Noise



Reset noise: V_{reset} is not the same for each capture, usually thermal induced. Eliminate reset noise by reading V_{reset} and V_{transfer} and taking the delta; this is what Correlated Double Sampling does, among other things (in a few slides). Without CDS, we read only V_{transfer} and the ADC circuit needs to be designed with a fixed, offline-determined V_{reset} in mind, introducing reset noise.

Gains and ISO

Signal Gain: ISO

A scaling factor (gain) applied to the actual measured voltage of a pixel.

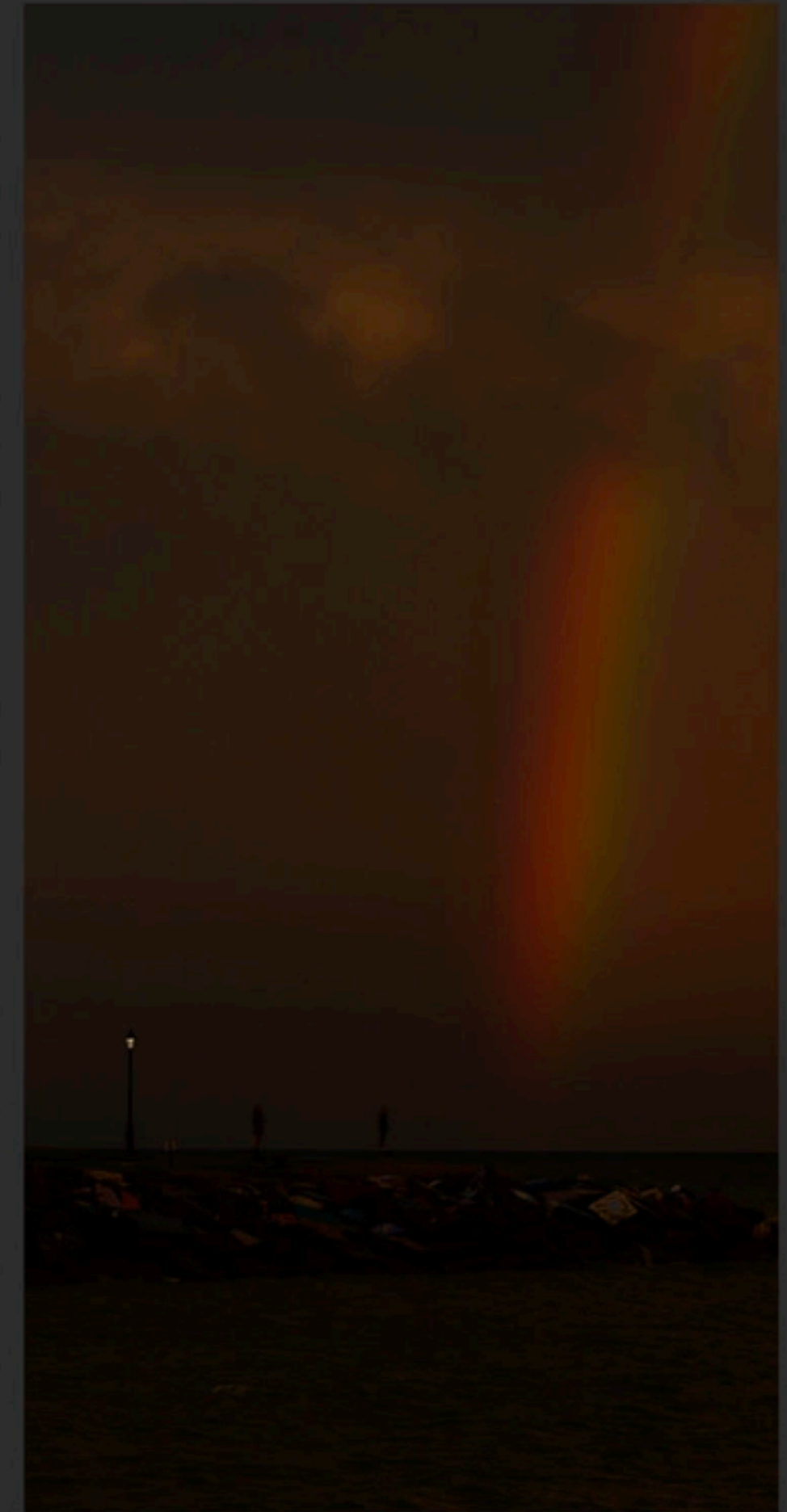
- Each camera has a base ISO value (e.g., ISO 400). Doubling ISO effectively doubles the measured voltage — equivalent to doubling exposure time.
- ISO is informally called camera sensitivity, but don't confuse it with the spectral sensitivity of the sensor.

Why use ISO gain?

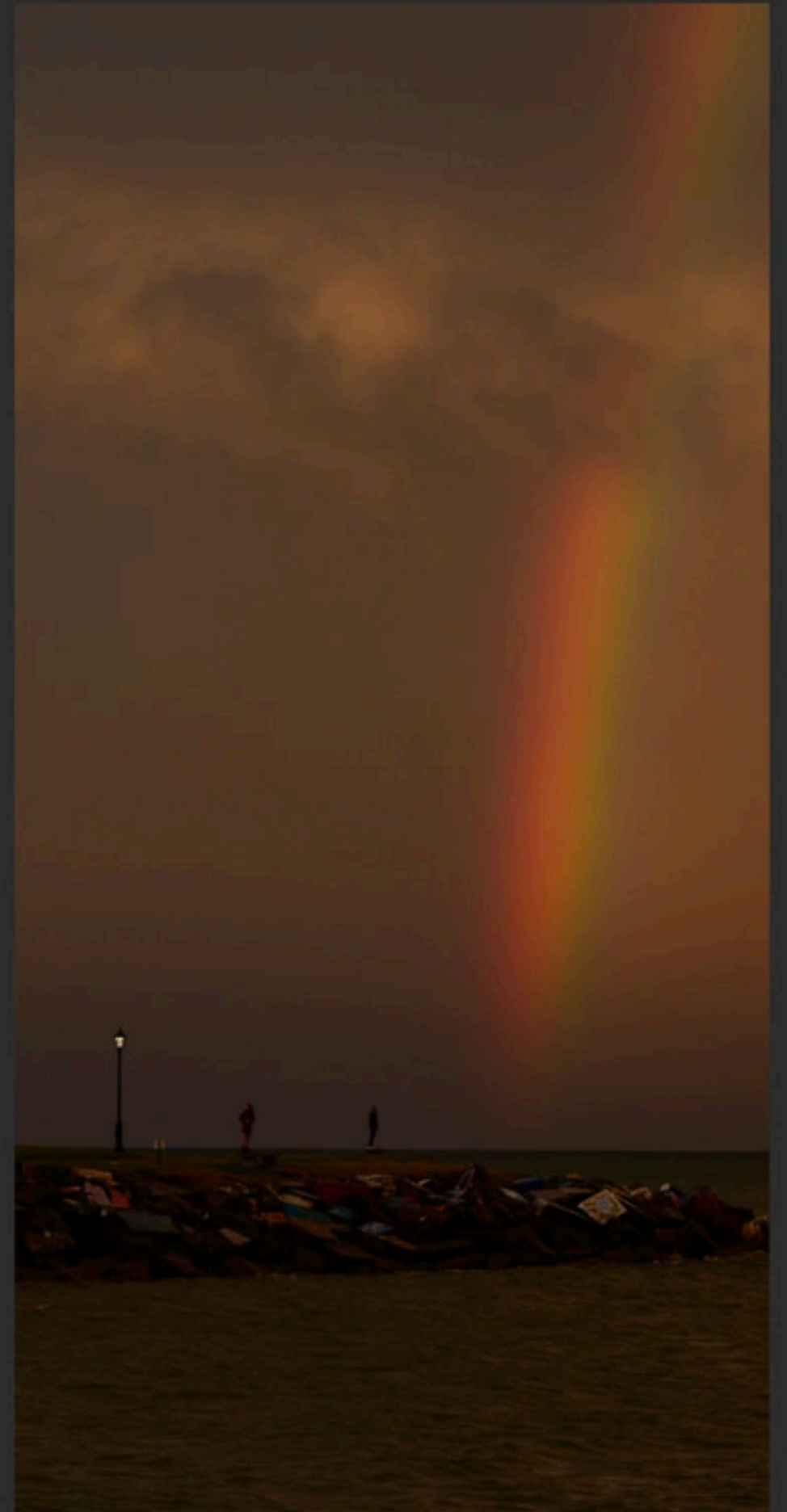
- Boost image brightness at low light: increasing effective exposure time without increasing the actual exposure time
- Reduce the exposure time, which reduces motion blur (if one were to increase the exposure time) and increases the frame rate
- **Big caveat:** shorter exposure time also reduces the SNR: noisy photo!



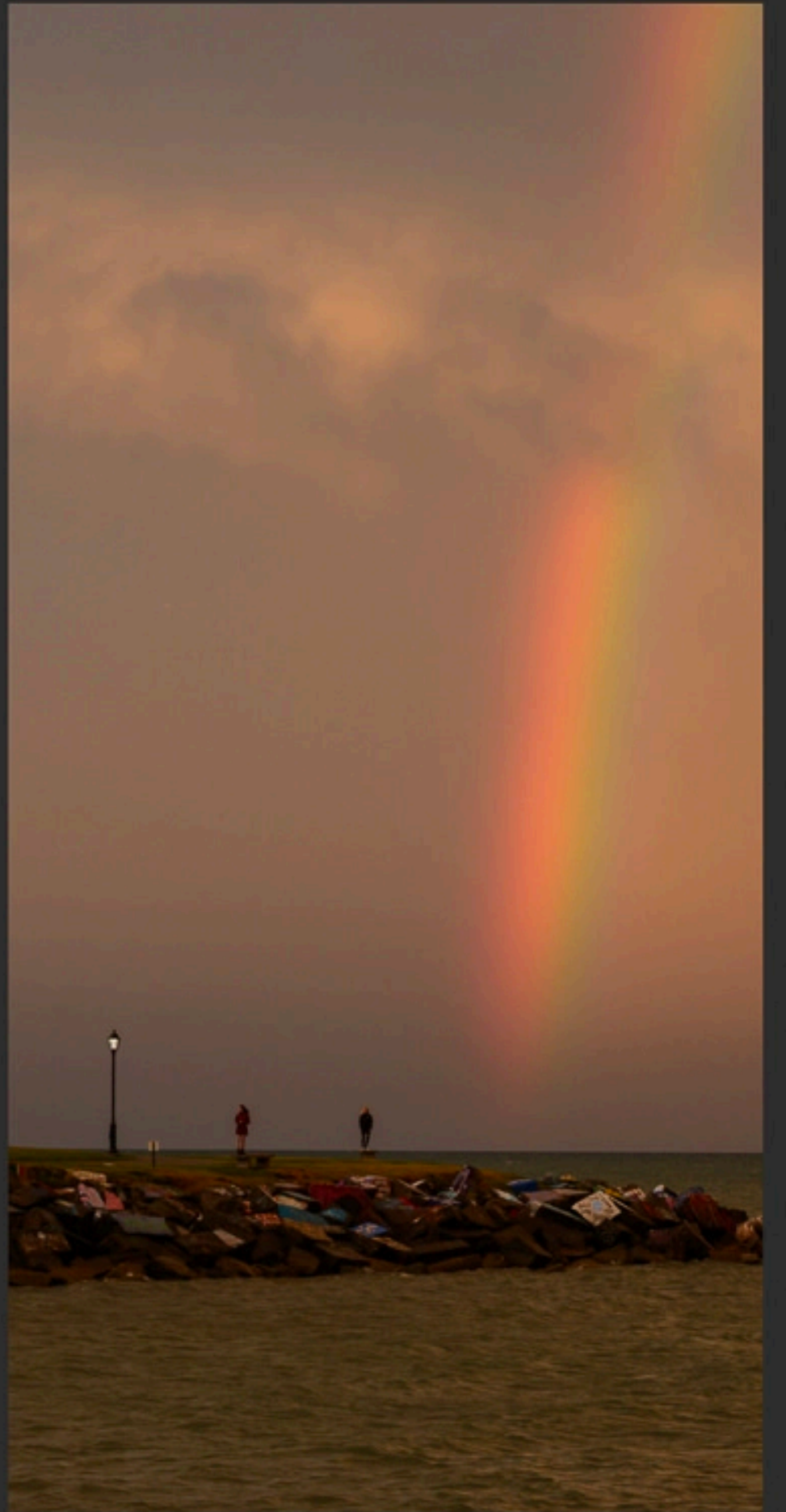
ISO 100



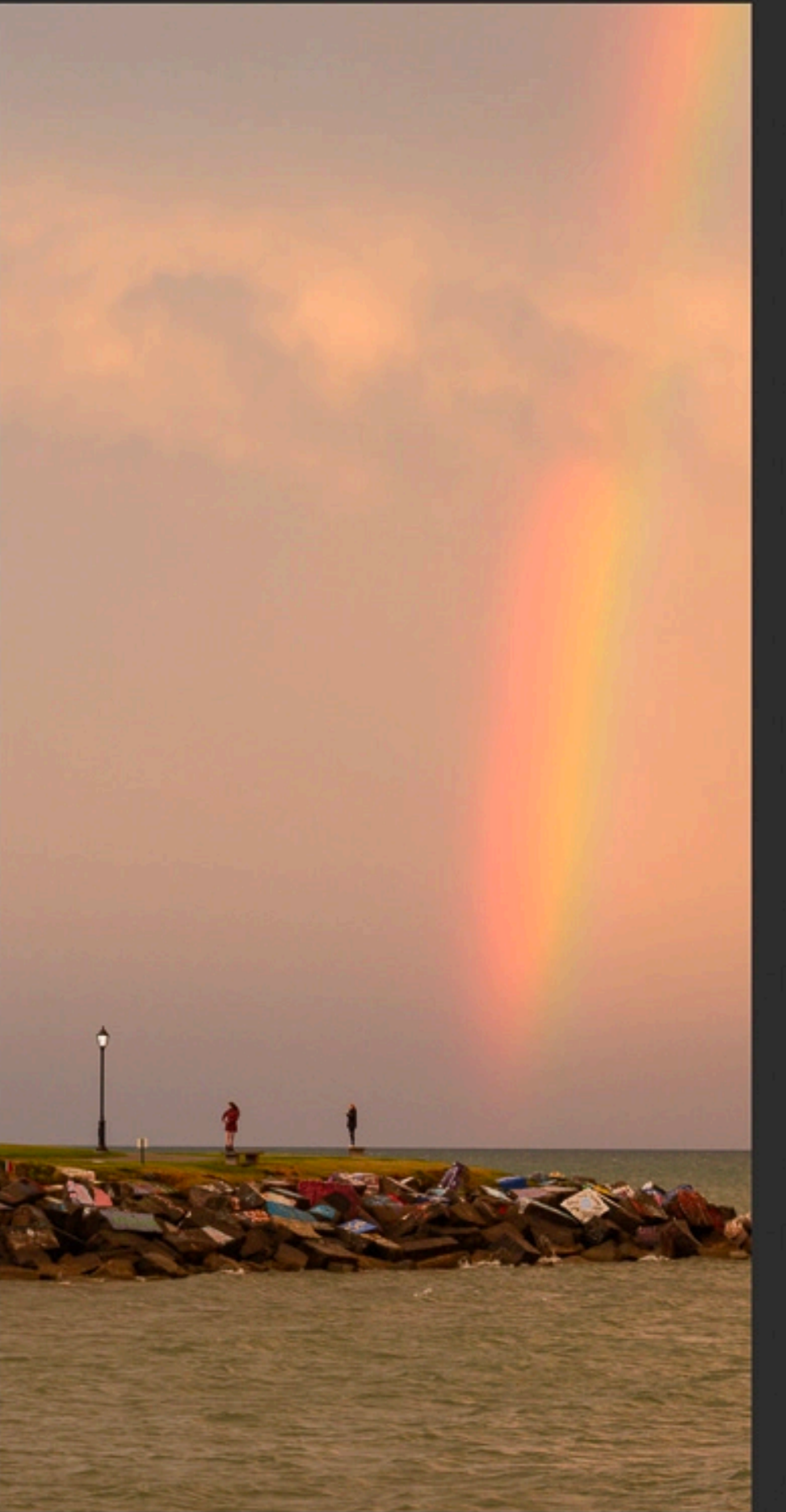
ISO 200



ISO 400



ISO 800



ISO 1600

ISO 800 and 1/2000s exposure



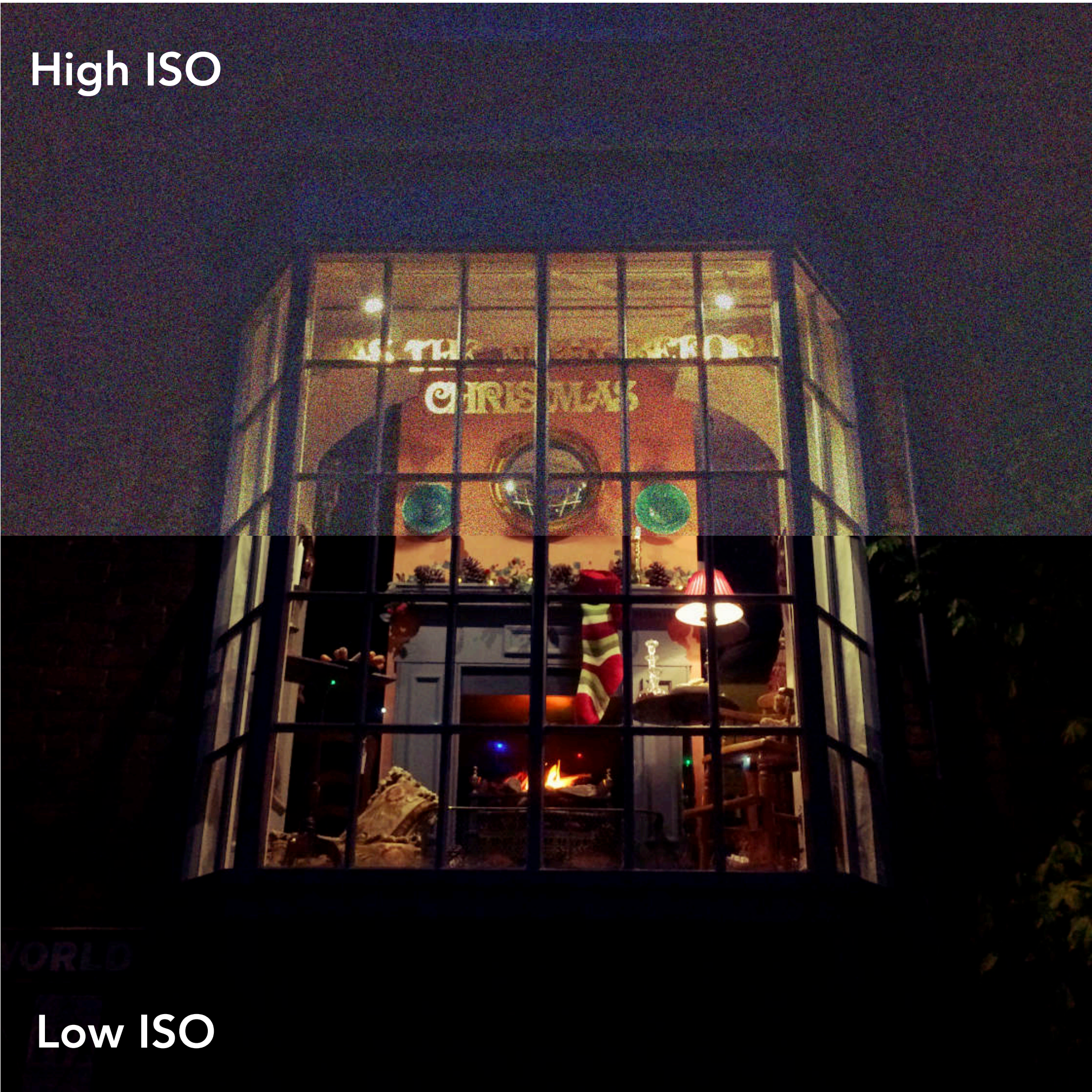
ISO 1800 and 60s exposure on Pentax KP



Noise From High ISO



Noise From High ISO

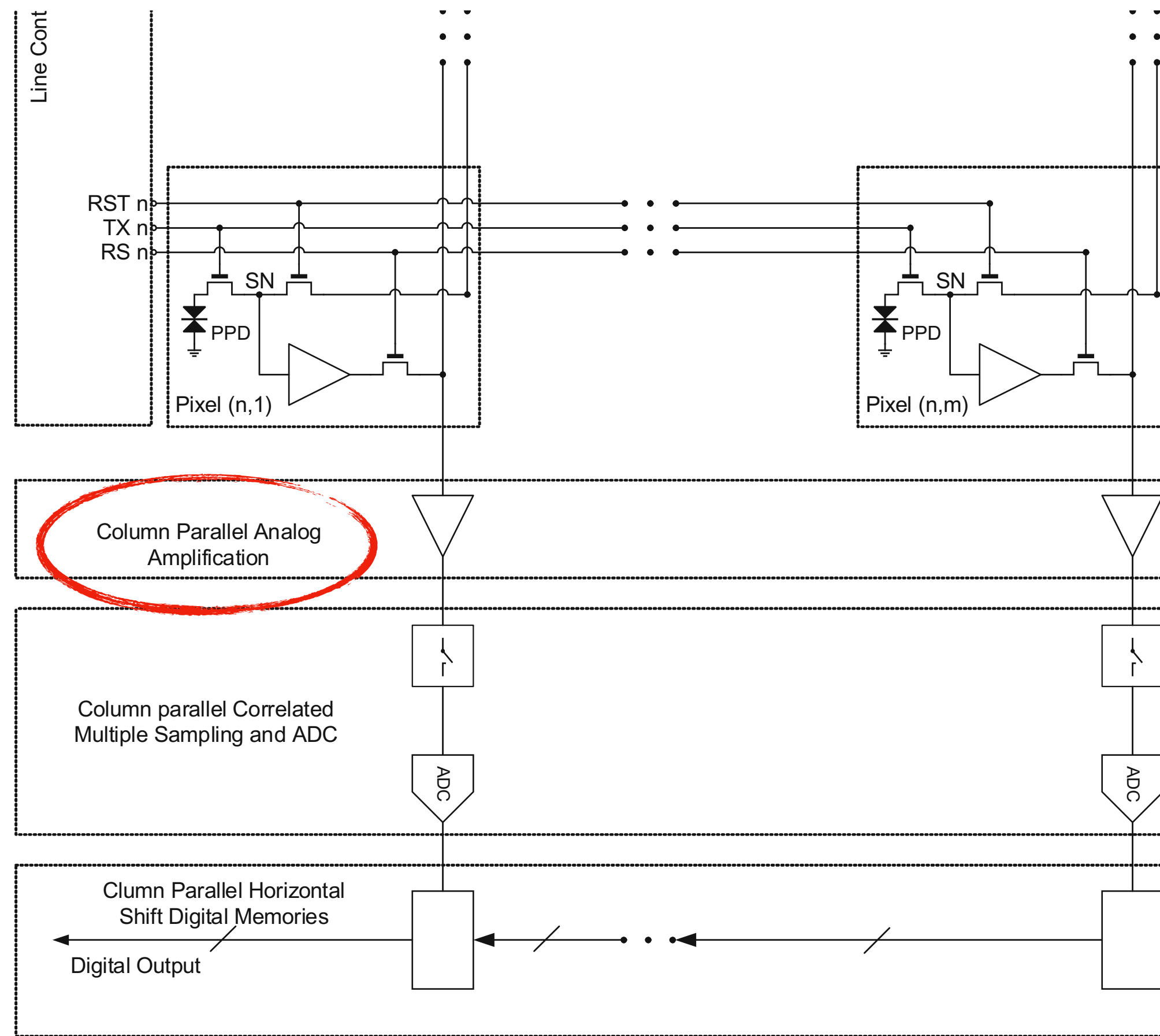


Camera+2 app on iPhone

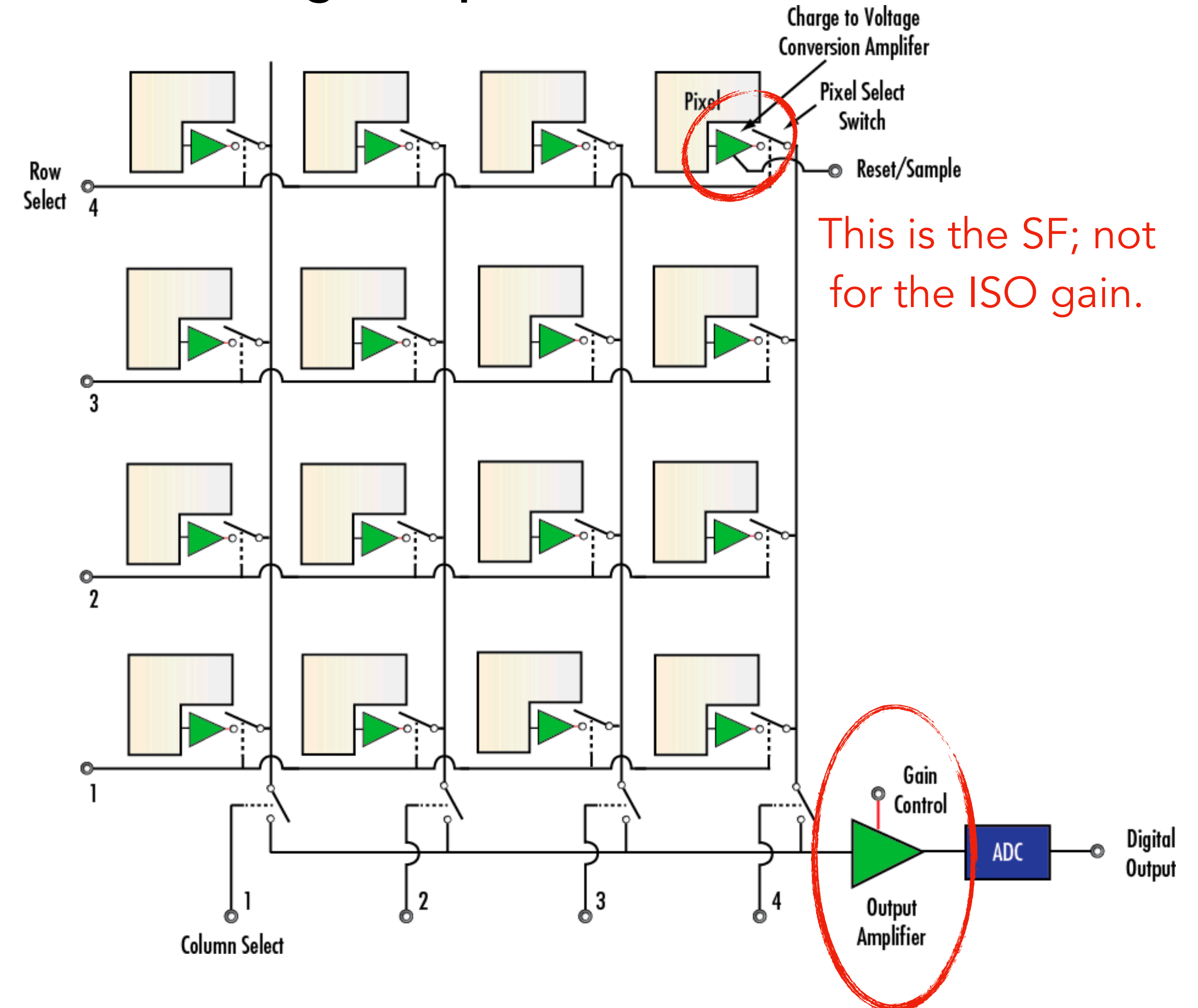


Where to Apply ISO Gain?

Column amplifiers

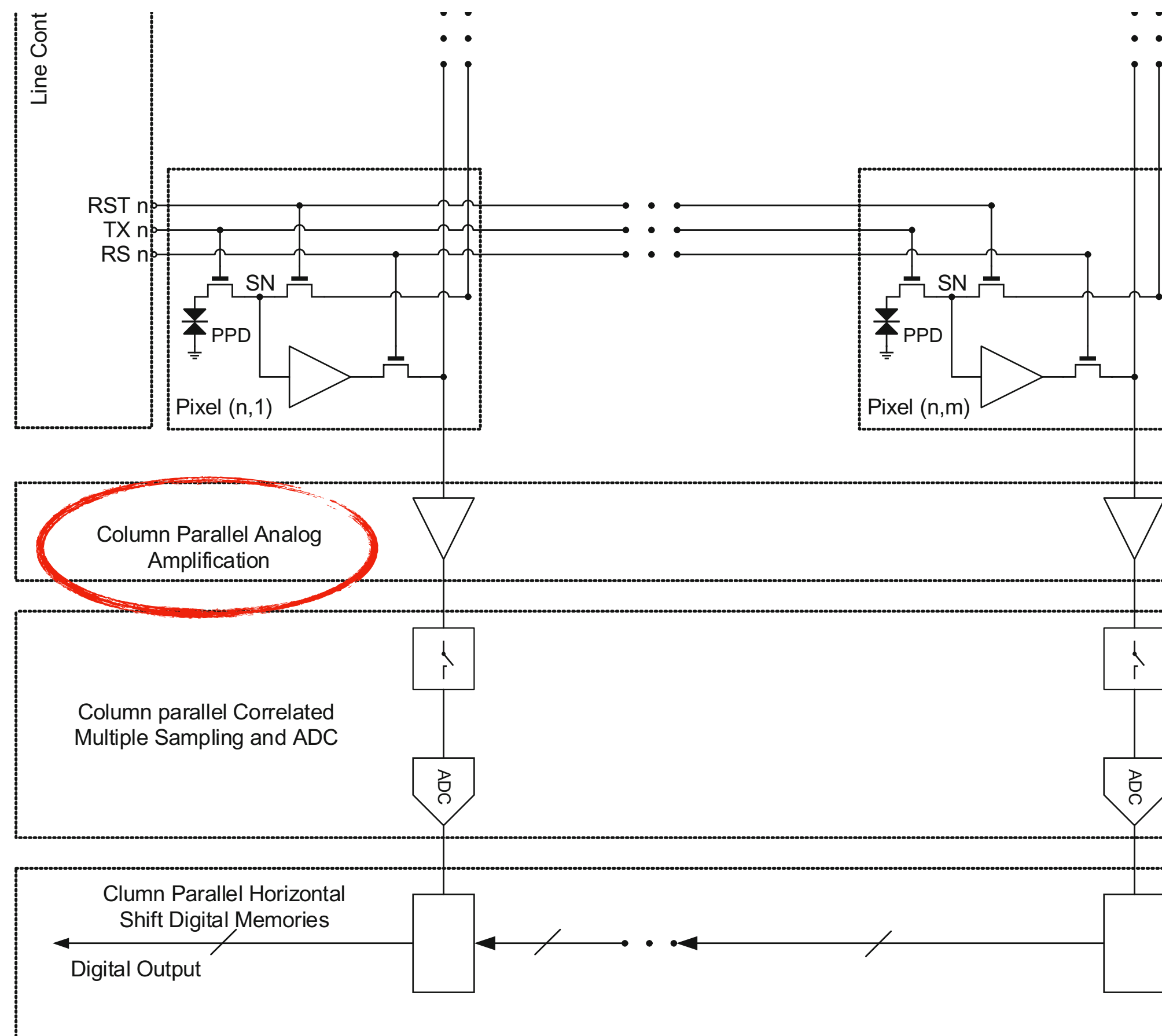


One single amplifier at the end



Where to Apply ISO Gain?

Column amplifiers



Applying the gain to the signal applies the gain to the noise too! Applying the gain early before much noise trickles in helps SNR.

- It's common to use column-level amplification. Some cameras have digital gains (post ADC).
- We could also use an in-pixel amplifier.
- ISO gain circuitry is usually called **A**daptive **G**ain **C**ontrol (AGC) or **P**rogrammable **G**ain **A**mplifier (PGA).

Apply ISO Gain After ADC

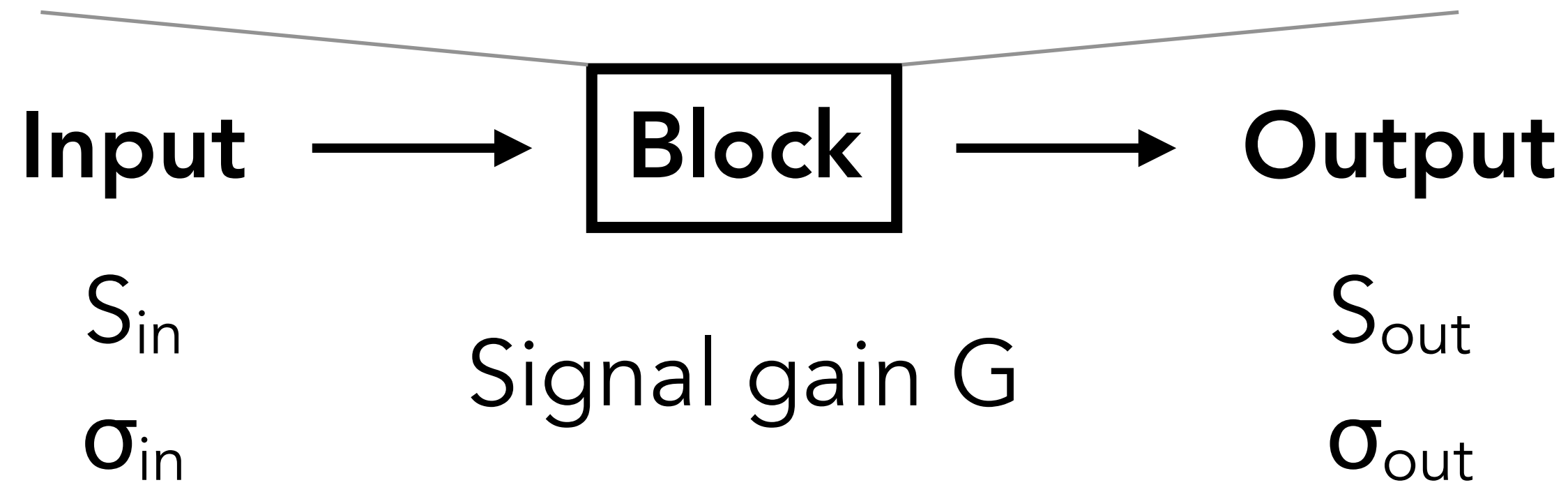
Applying ISO gains after ADC quantization, i.e., digital gain, magnifies the quantization artifact (contouring).



Before quantization: **3.4, 4.1**
After quantization: **3, 4**
Increasing ISO by 10X: **30, 40**

Gain Blocks

- Photon to electron conversion in PD
- electron charge to voltage conversion in FD
- SF in read-out
- other amplifiers (e.g., for ISO control)
- ADC, etc.



$$S_{out} = GS_{in}$$

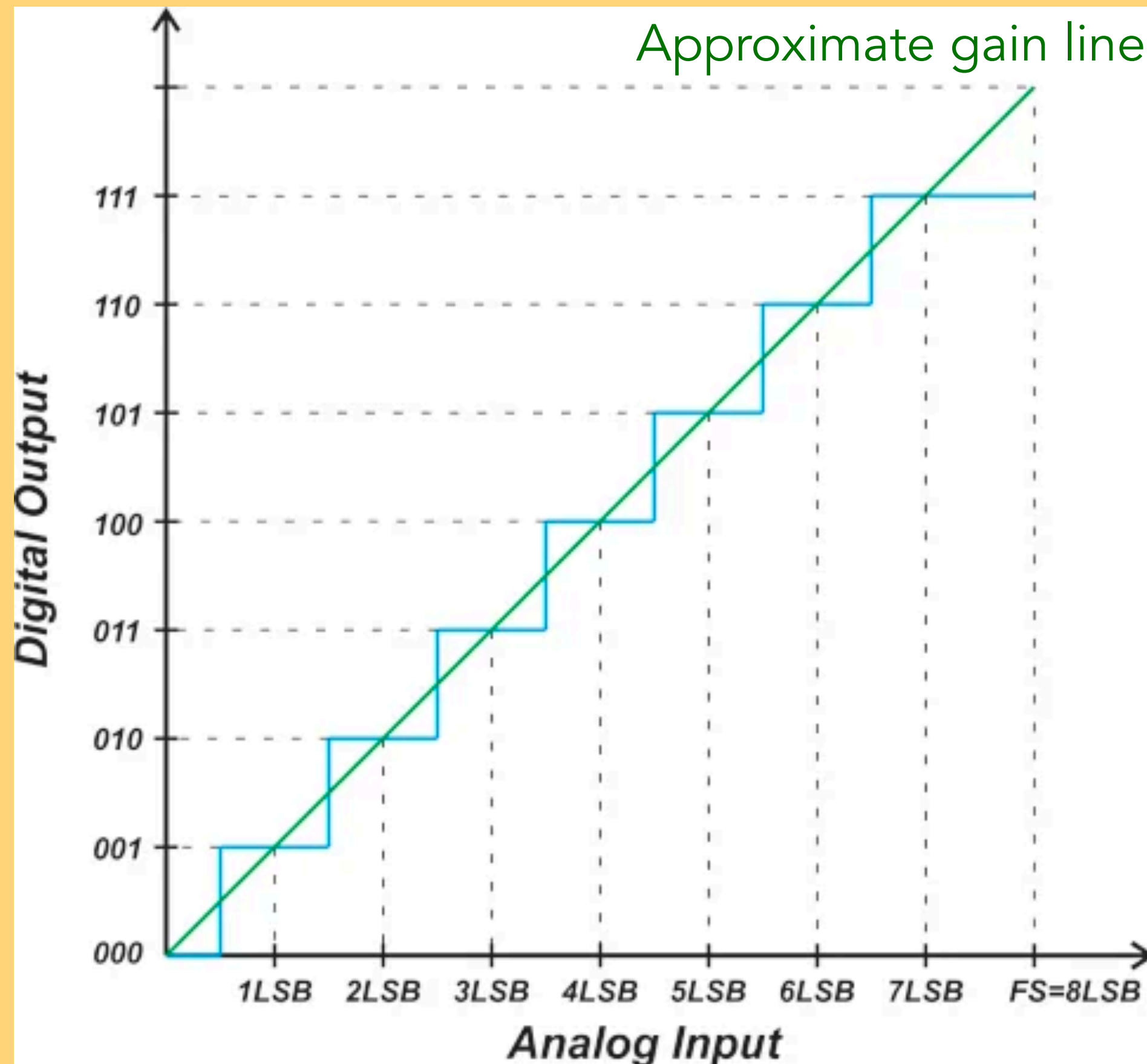
$$\sigma_{out} = G\sigma_{in}$$

Many blocks in signal chain can be seen as applying gains even though their goals are not to amplify signals.

- Photodiodes apply a photon to electron conversion; unit of gain is photon/electron.
- FD converts charges to voltage, and has a conversion gain with a unit of Volt/electron.
- ADCs convert voltages to raw pixel values, a.k.a., **ADUs** (analog-to-digital units) or **DNs** (data numbers). ADC gains have units of ADU/volt.

They also amplify noises.

ADC Gain



ADC gain is technically non-linear, but for the simplicity of noise analysis we assume it's linear.

Full-scale range [V] $FSR = V_{max} - V_{min}$ ADC input voltage range

ADC Gain $G_{ADC} = \frac{2^N}{FSR}$ Resolution

Voltage equivalent to one LSB $\Delta V_{LSB} = \frac{FSR}{2^N}$

ADC Quantization Error

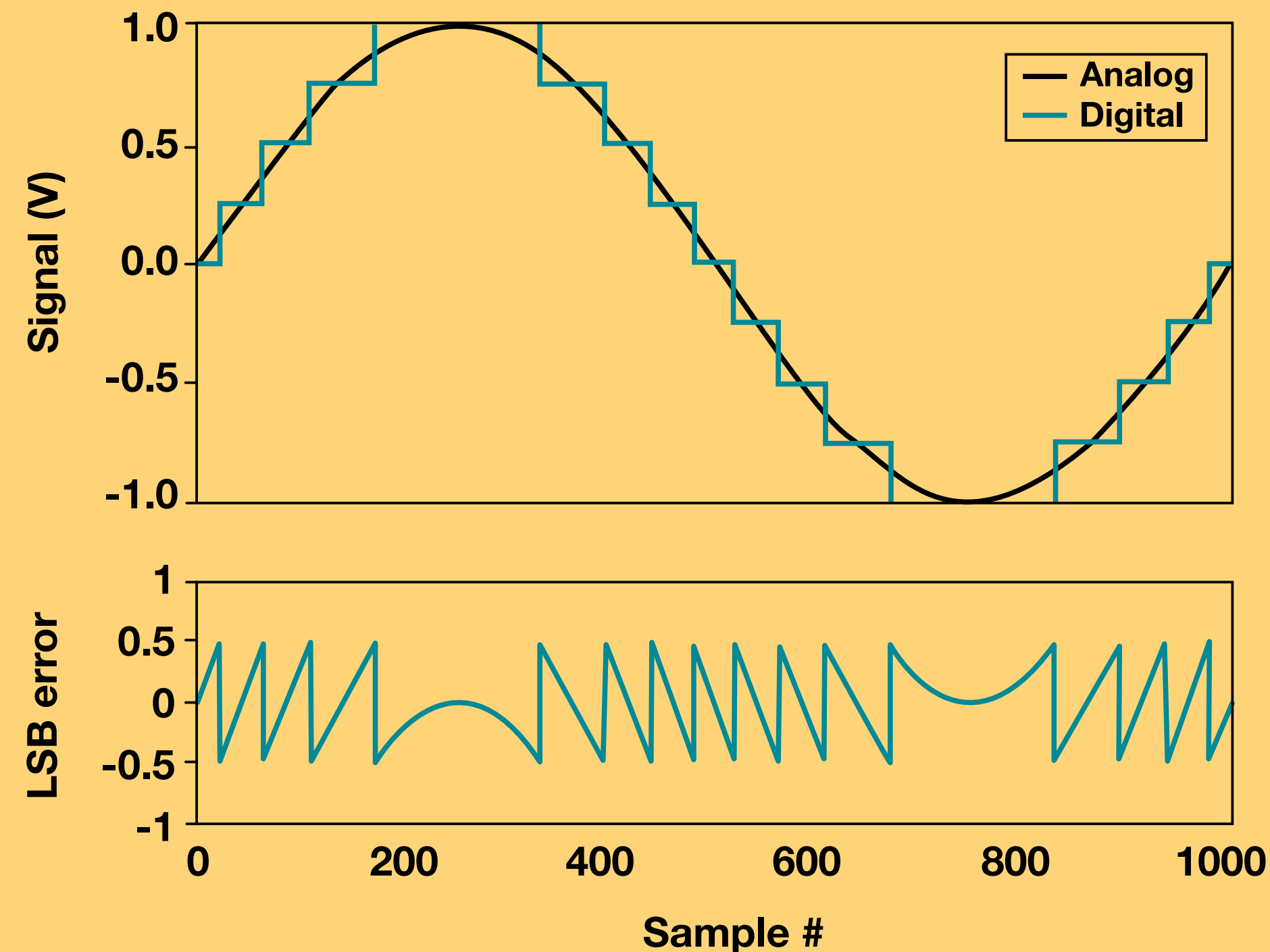


Figure 4. Analog input, digital output and LSB error waveforms.

ADC quantization error can be modeled as a “noise” source too.

- All voltages v within a LSB are replaced with the same level V_d .

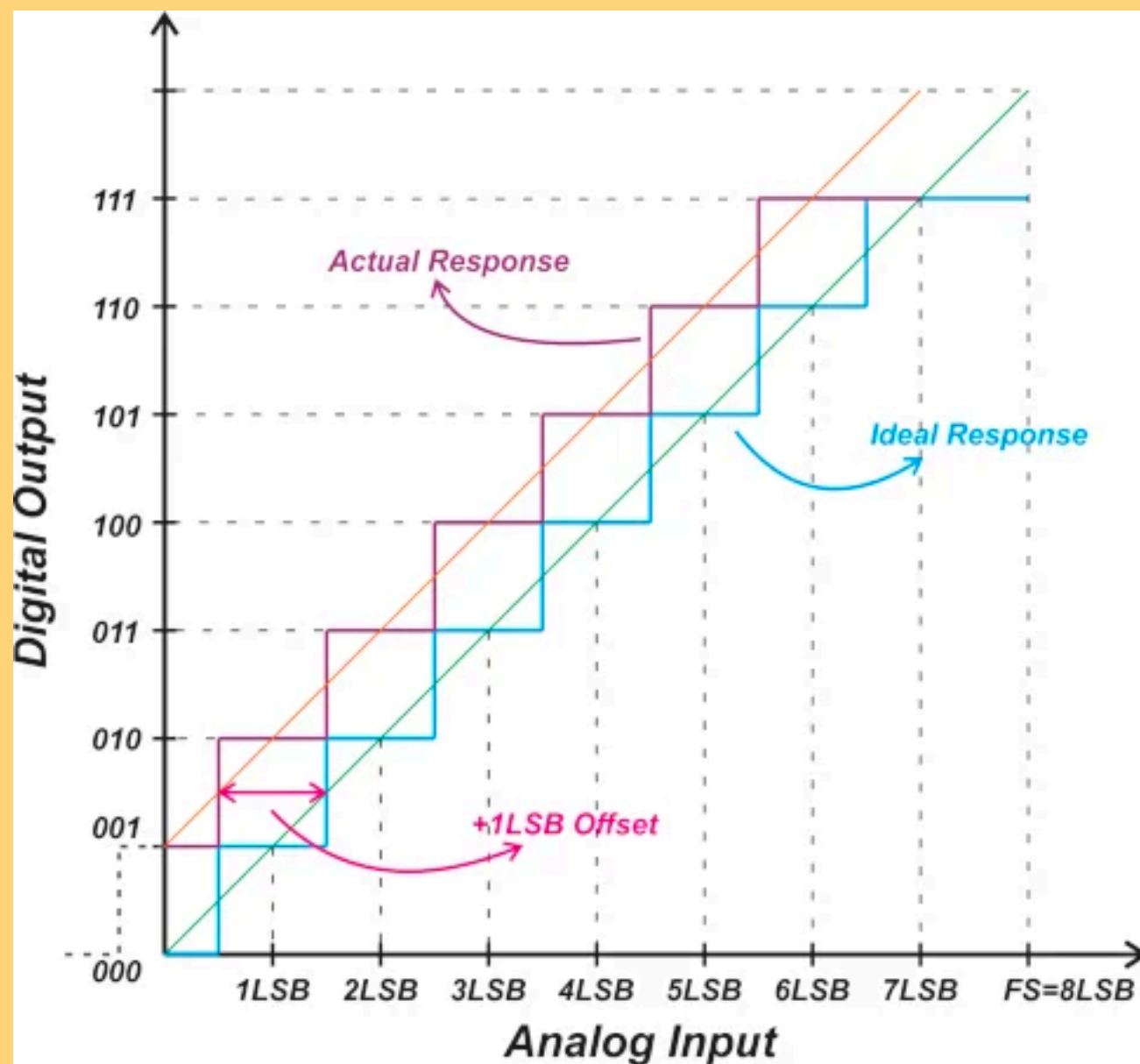
Assuming analog values are uniformly distributed, the variance introduced by quantization is:

$$\sigma_{ADC-q}^2 = \int_{V_d - \frac{1}{2}\Delta V_{LSB}}^{V_d + \frac{1}{2}\Delta V_{LSB}} (v - V_d)^2 \frac{1}{\Delta V_{LSB}} dv = \frac{1}{12} \Delta V_{LSB}^2$$

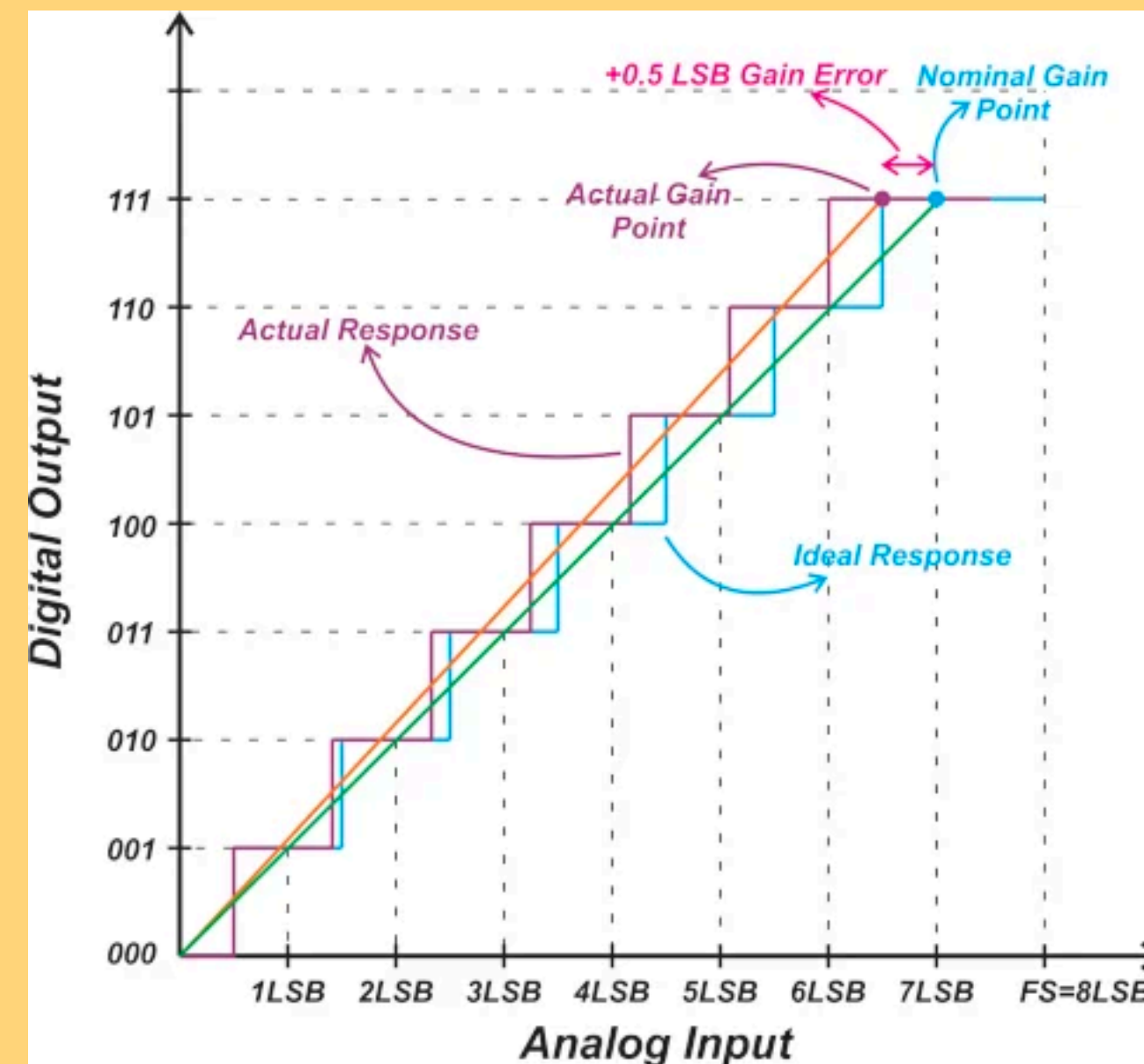
Other ADC Noises/Errors

Offset error and gain error are fixed for an ADC and can be compensated once characterized. They are technically errors, not noises, just like quantization errors, but can be modeled as noises. Usually in modeling we lump all ADC noise sources together as a single ADC noise.

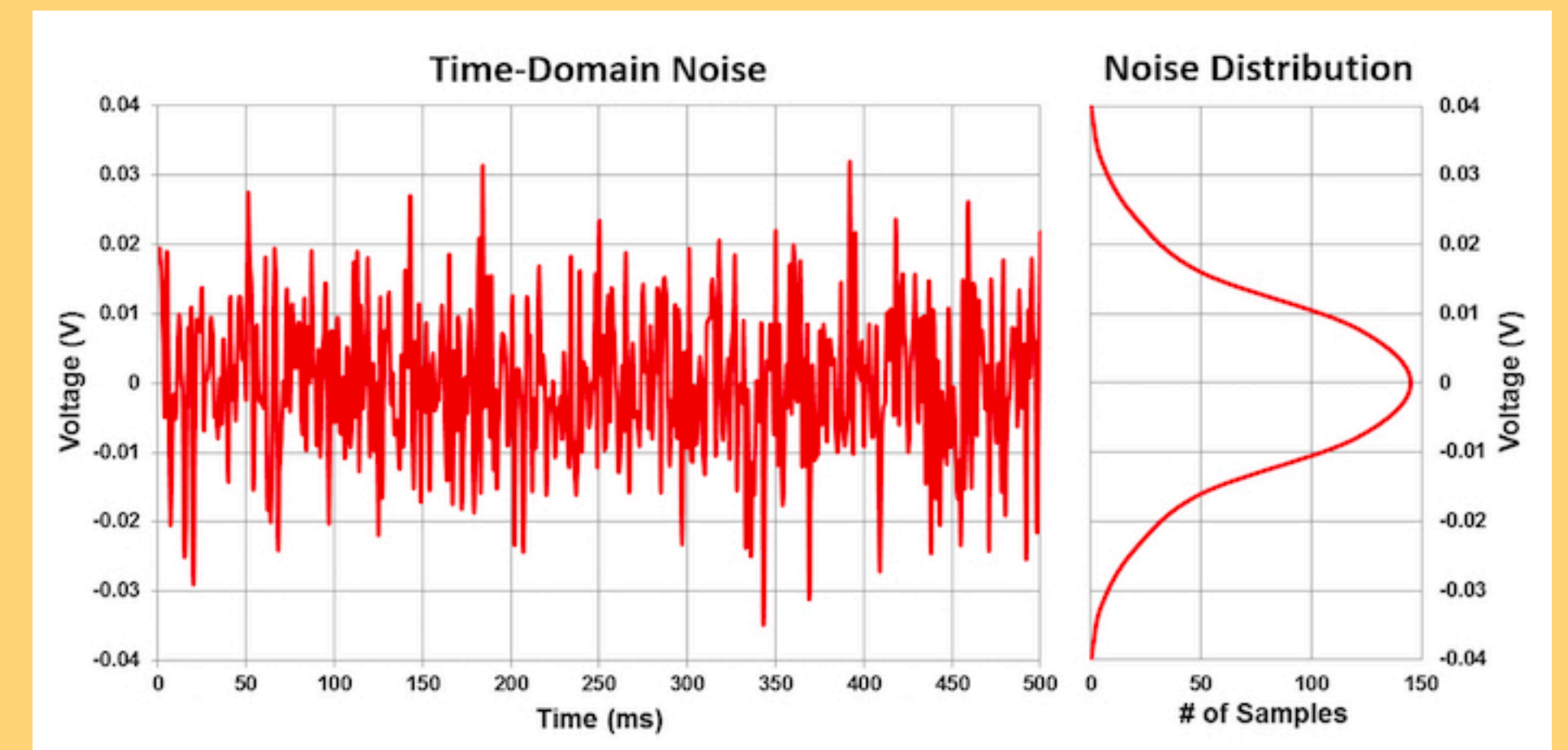
Offset error
(ADC specific)



Gain error
(ADC specific)



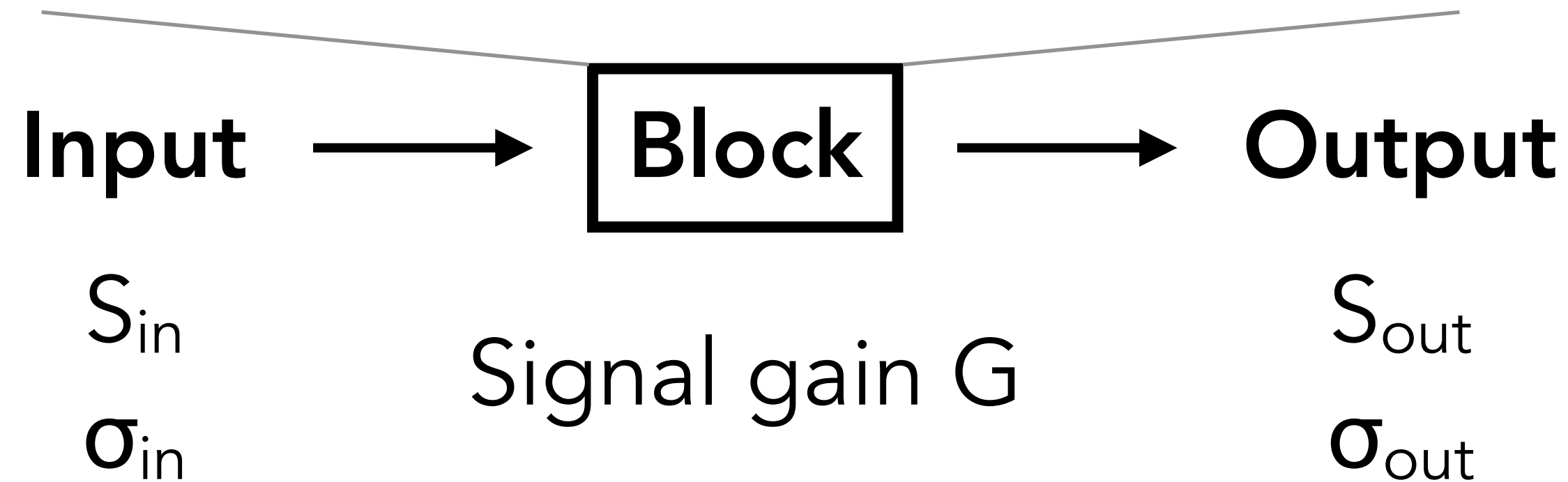
Read noise
(General to any circuit)



Noise Propagation

Noise Amplification in the Signal Chain

- Photon to electron conversion in PD
- electron charge to voltage conversion in FD
- SF in read-out
- other amplifiers (e.g., for ISO control)
- ADC, etc.



$$S_{out} = G S_{in}$$

$$\sigma_{out} = G \sigma_{in}$$

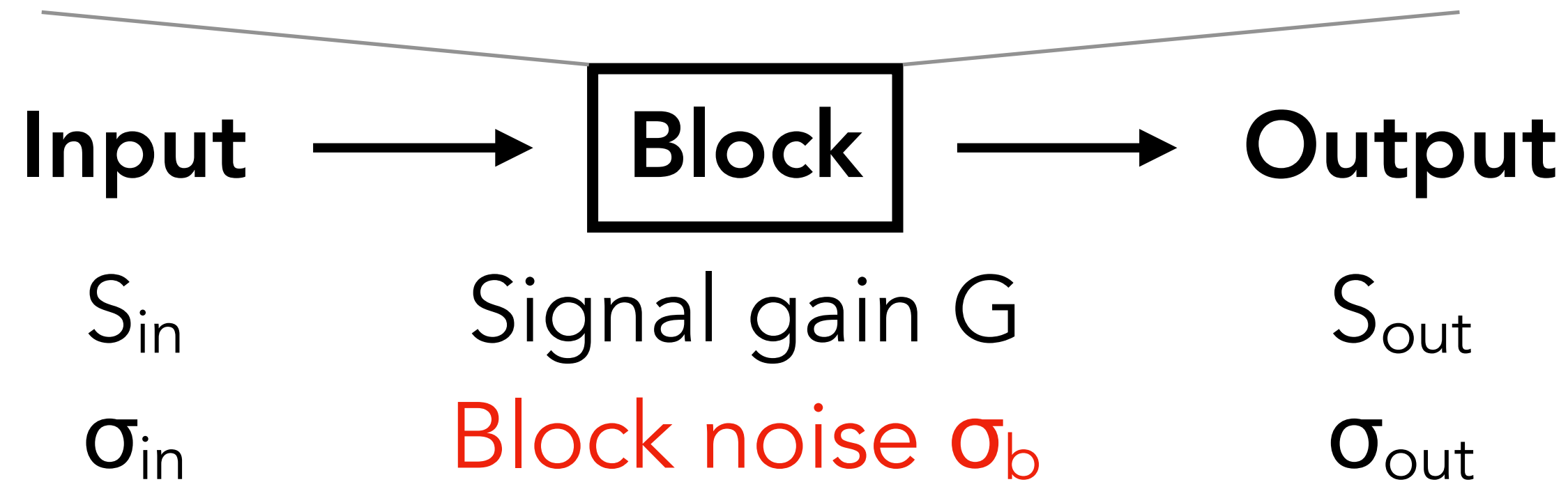
Recall that scaling a random variable by a constant scales the variance by the square of that constant.

- $\text{Var}(a X) = a^2 \text{Var}(X)$
- $\sigma(a X) = a \sigma(X)$

Amplifying a signal amplifies the noise as well; SNR is unchanged (assuming the block has no inherent noise)!

Noise Amplification in the Signal Chain

- Photon to electron conversion in PD
- electron charge to voltage conversion in FD
- SF in read-out
- other amplifiers (e.g., for ISO control)
- ADC, etc.



$$S_{out} = GS_{in}$$

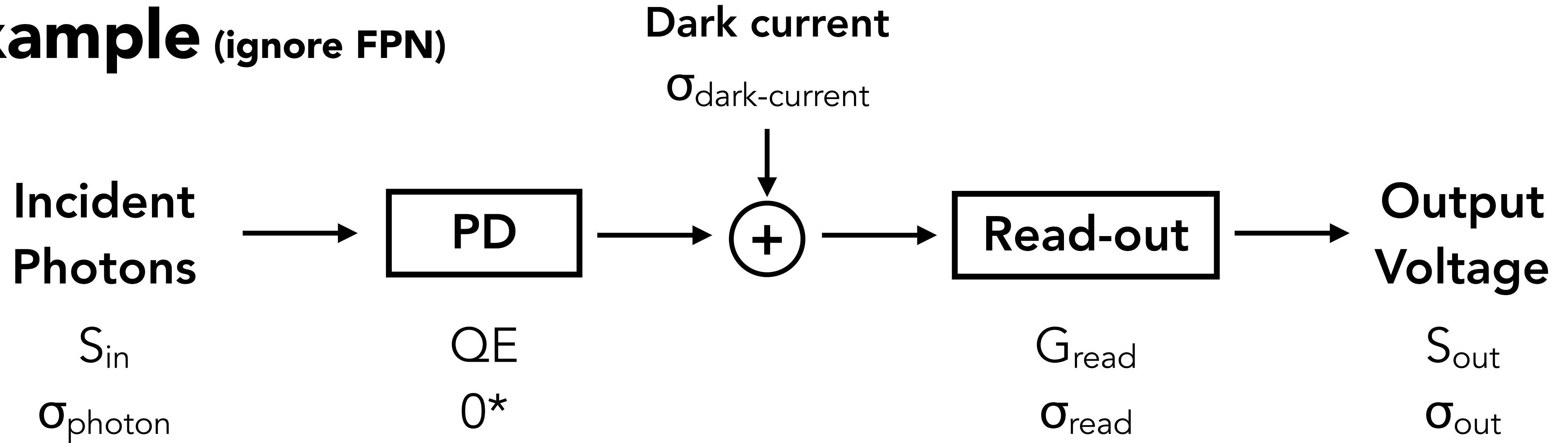
$$\sigma_{out} = \sqrt{G^2 \sigma_{in}^2 + \sigma_b^2}$$

Each block in the signal processing chain introduces noises, which can be modeled as **independent** random variables. The variance of their sum is the sum of their variances.

$$\bullet \text{ Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$

In general, SNR is worsened by the signal processing chain.

Example (ignore FPN)

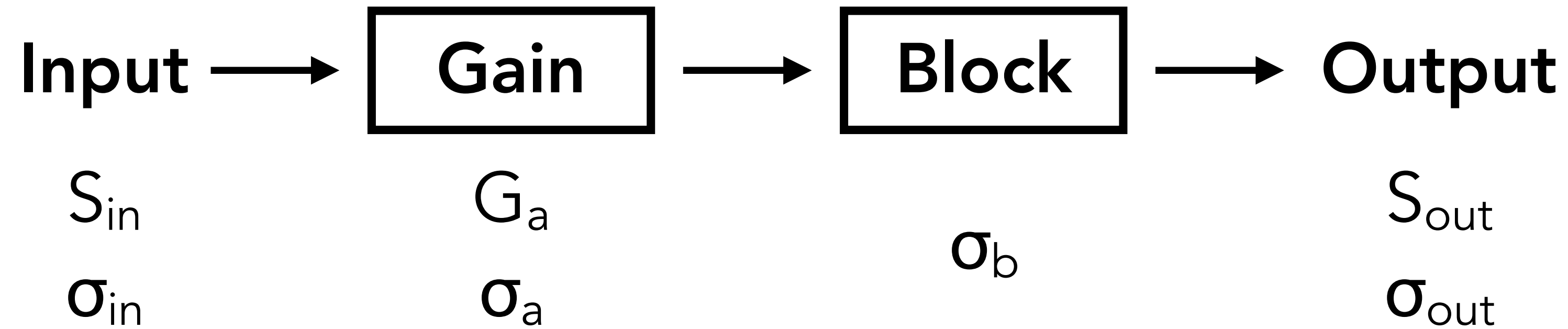


$$S_{out} = QE \times G_{read} \times S_{in}$$

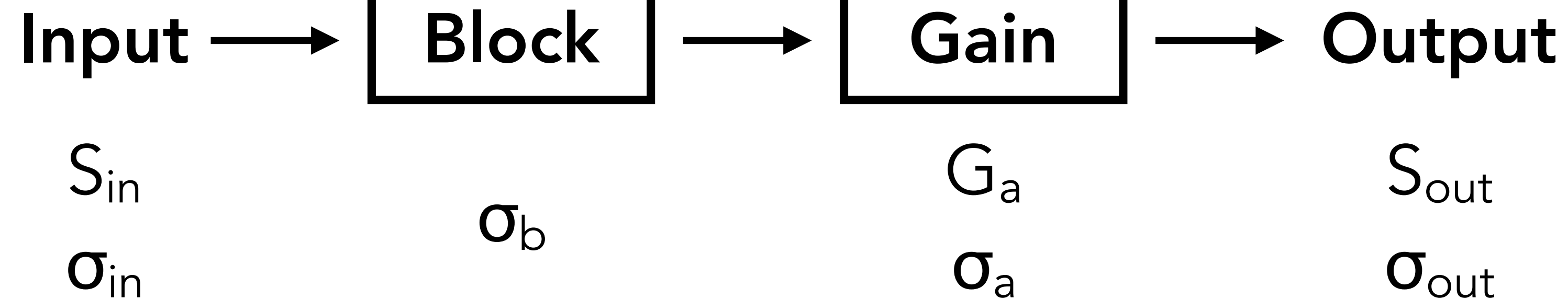
$$\sigma_{out} = \sqrt{QE^2 G_{read}^2 \sigma_{photon}^2 + G_{read}^2 \sigma_{dark-current}^2 + \sigma_{read}^2}$$

* assuming no noise in photon to electron conversion (QE is a constant)

Apply Gains Earlier Improves SNR

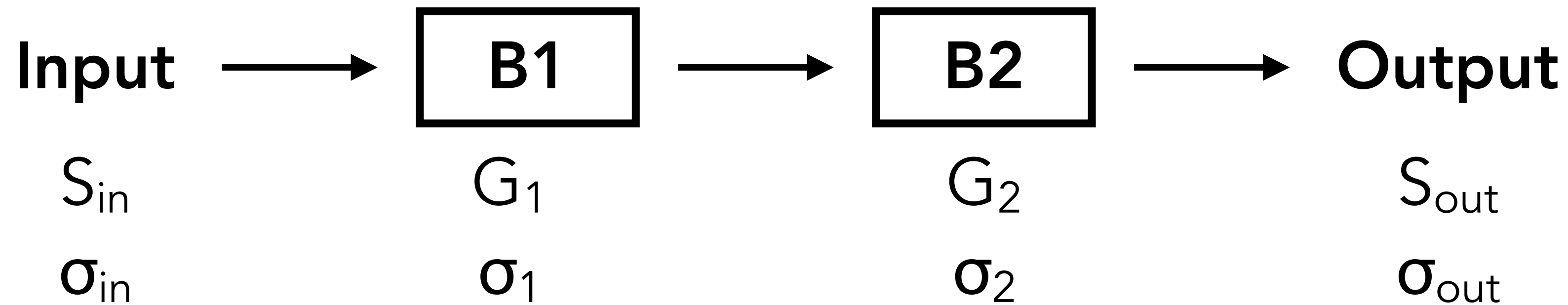


$$\sigma_{out} = \sqrt{G_a^2 \sigma_{in}^2 + \sigma_a^2 + \sigma_b^2}$$



$$\sigma_{out} = \sqrt{G_a^2 (\sigma_{in}^2 + \sigma_b^2) + \sigma_a^2}$$

Modeling Multiple Blocks as One Block



$$S = S_{in} G_1 G_2 = S_{in} G$$

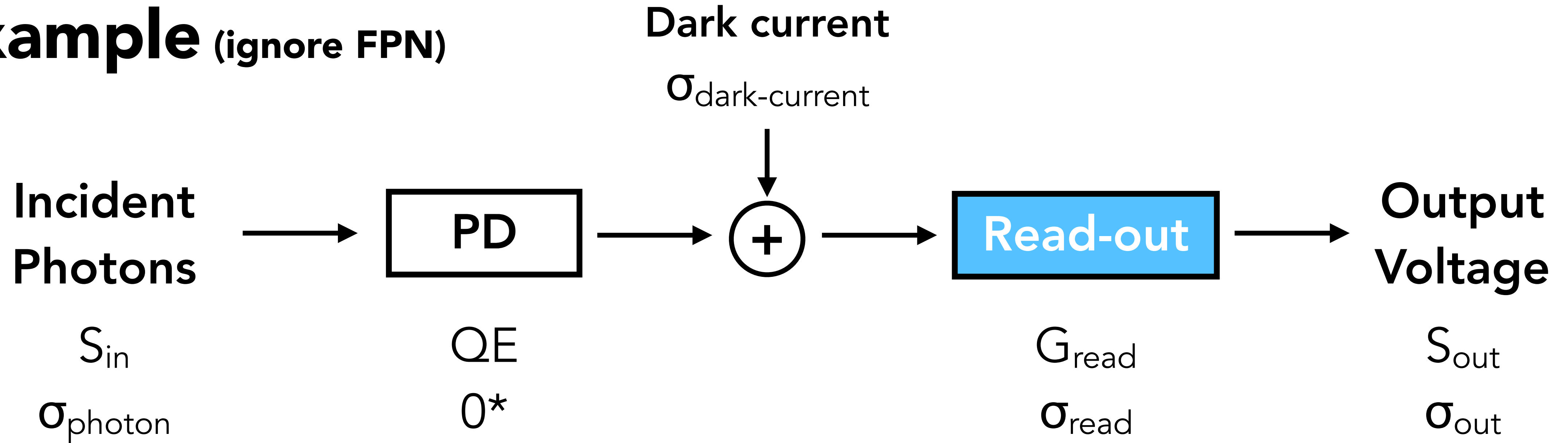
$$\sigma = (\sigma_{in}^2 G_1^2 + \sigma_1^2) G_2^2 + \sigma_2^2 = \sigma_{in}^2 G_1^2 G_2^2 + \sigma_1^2 G_2^2 + \sigma_2^2 = \sigma_{in}^2 (G_1 G_2)^2 + \sigma^2$$

$$\longrightarrow \begin{cases} G = G_1 G_2 \\ \sigma = \sqrt{\sigma_1^2 G_2^2 + \sigma_2^2} \end{cases}$$

If three gains are
combined:

$$\begin{cases} G = G_1 G_2 G_3 \\ \sigma = \sqrt{\sigma_1^2 G_2^2 G_3^2 + \sigma_2^2 G_3^2 + \sigma_3^2} \end{cases}$$

Example (ignore FPN)



If the read-out chain consists of FD, SF, and an Adaptive Gain Control (e.g., column-level amp):

$$\begin{cases} S_{\text{out}} = QE \times G_{\text{read}} \times S_{\text{in}} \\ \sigma_{\text{out}} = \sqrt{QE^2 G_{\text{read}}^2 \sigma_{\text{photon}}^2 + G_{\text{read}}^2 \sigma_{\text{dark-current}}^2 + \sigma_{\text{read}}^2} \end{cases} \quad \begin{cases} G_{\text{read}} = G_{\text{FD}} G_{\text{SF}} G_{\text{AGC}} \\ \sigma_{\text{read}} = \sqrt{\sigma_{\text{FD}}^2 G_{\text{SF}}^2 G_{\text{AGC}}^2 + \sigma_{\text{SF}}^2 G_{\text{AGC}}^2 + \sigma_{\text{AGC}}^2} \end{cases}$$

* assuming no noise in photon to electron conversion

SNR and Input- vs. Output-Referred Noise

$$\begin{aligned}
 SNR &= \frac{QE \times G_{read} \times S_{in}}{\sqrt{QE^2 G_{read}^2 \sigma_{photon}^2 + G_{read}^2 \sigma_{dark-current}^2 + \sigma_{read}^2}} \\
 &= \frac{S_{in}}{\sqrt{\sigma_{photon}^2 + \sigma_{dark-current}^2 / QE^2 + \sigma_{read}^2 / QE^2 / G_{read}^2}} \\
 &= \frac{QE \times S_{in}}{\sqrt{QE^2 \sigma_{photon}^2 + \sigma_{dark-current}^2 + \sigma_{read}^2 / G_{read}^2}}
 \end{aligned}$$

Output-referred noise
(expressed in Volts)

Input-referred noise
(expressed in photon counts)

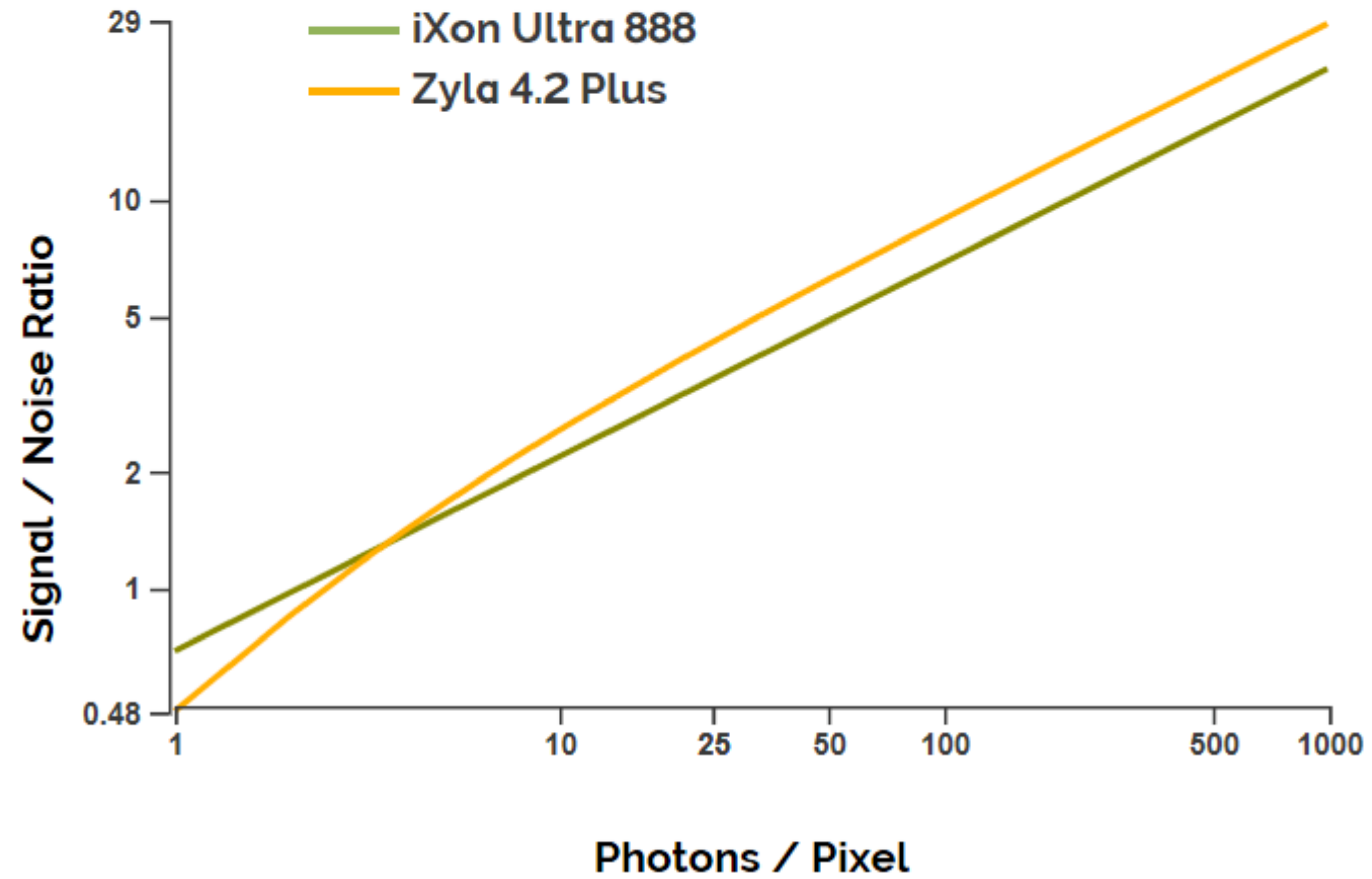
Sometimes input-referred noise is also expressed in electron counts

SNR vs. Exposure Time

$$\begin{aligned}
 SNR &= \frac{QE \times S_{in}}{\sqrt{QE^2 \sigma_{photon}^2 + \sigma_{dark-current}^2 + \sigma_{read}^2 / G_{read}^2}} \\
 &= \frac{P \times QE \times t}{\sqrt{\underbrace{P \times QE \times t}_{\text{Recall: photon shot variance = mean}} + \underbrace{I_{dark} \times t}_{\text{Dark current: \# of electrons/sec/pixel}} + \sigma_{read}^2 / G_{read}^2}}
 \end{aligned}$$

of photons/sec/pixel
Exposure time

SNR Examples



A Sensor Model with Temporal Noise

of incident photons at pixel (x, y); varies with λ and t. Subject to photon-shot noise.

$$Q(x, y) = \int_t \int_\lambda \int_y^{y+v} \int_x^{x+u} (Y(x, y, \lambda, t) QE(\lambda) dx dy d\lambda + I_d) q dt$$

Dark current / Elementary charge

$$\Delta V(x, y) = \frac{Q(x, y)}{C} G + \eta_{read}$$

Read noise / Gain (SF, ISO gain, etc.)

$$n(x, y) = \left\lfloor \frac{\Delta V(x, y)}{V_{max}} (2^N - 1) \right\rfloor$$

Raw pixel value

Max voltage corresponding to max raw value

Fixed Pattern Noise

Fixed Pattern Noise (FPN): Spatial Non-Uniformity

Better be called spatial non-uniformity; due to manufacturing variation of circuitry components across pixels, columns, etc. The variation is spatially different ("noise") but doesn't vary with shots ("fixed").

Two FPN components:

- Dark signal non-uniformity (**DSNU**): non-uniformity without illumination
- Photo response non-uniformity (**PRNU**): non-uniformity under illumination

FPN can be removed/reduced by subtracting the pattern (at each exposure time, illumination level, and/or ISO setting) if the pattern is known offline.

- ISO setting affects gains, which amplify noise.

Two Components of DSNU

DCNU (dark current non-uniformity):

- Think of DCNU as the fact that each pixel has a different (but fixed) mean dark current. Since actual dark current is inherently temporally varying, the effect of DCNU is not a constant voltage offset, but we can average multiple frames to get the average dark current of each pixel, which can be seen as the pattern of DCNU.

The effects of DCNU:

- increase with exposure time
 - unaffected by irradiance
 - increase with ISO setting
- } **Illumination level**

Two Components of DSNU

SONU (spatial offset non-uniformity):

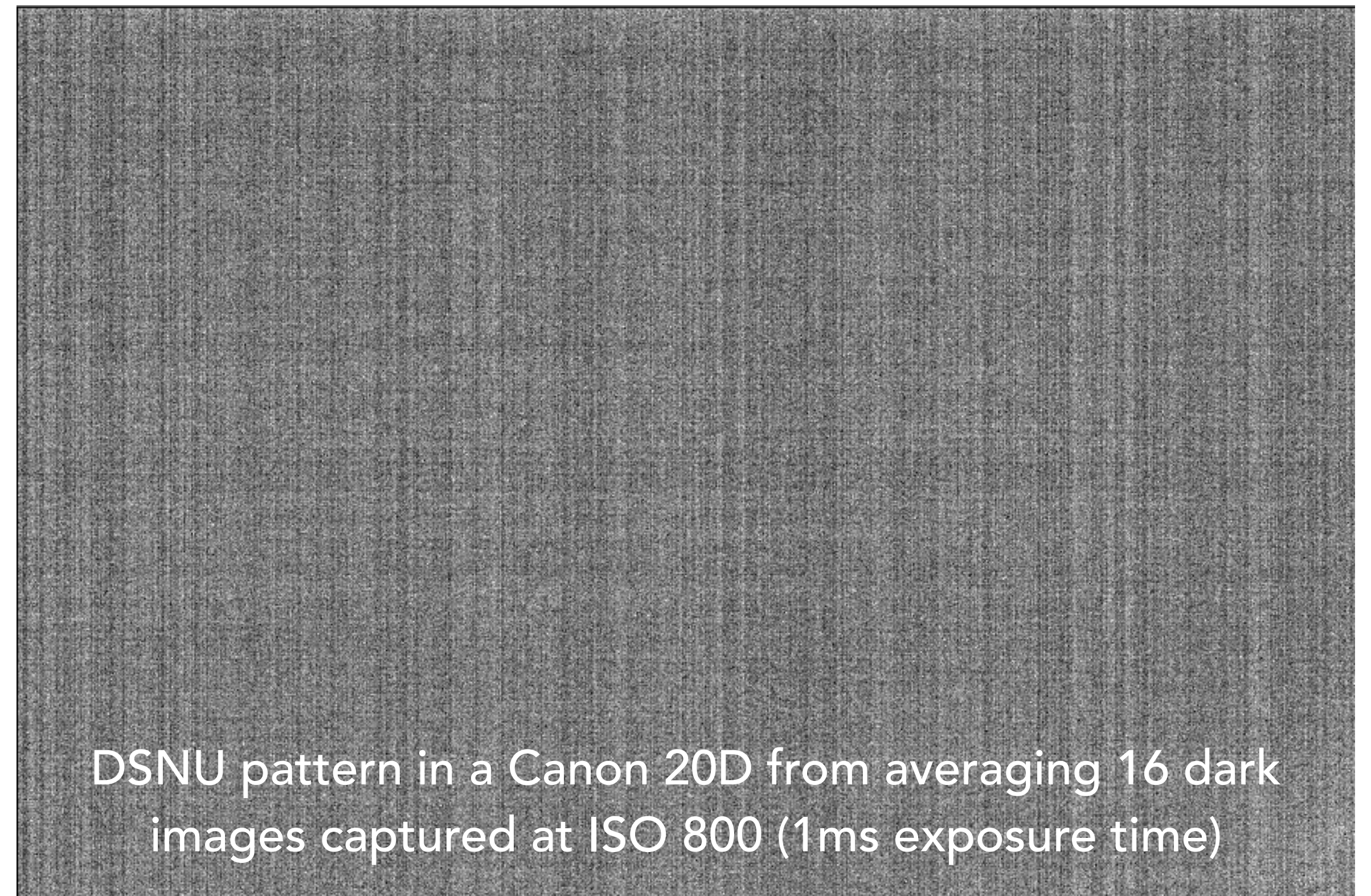
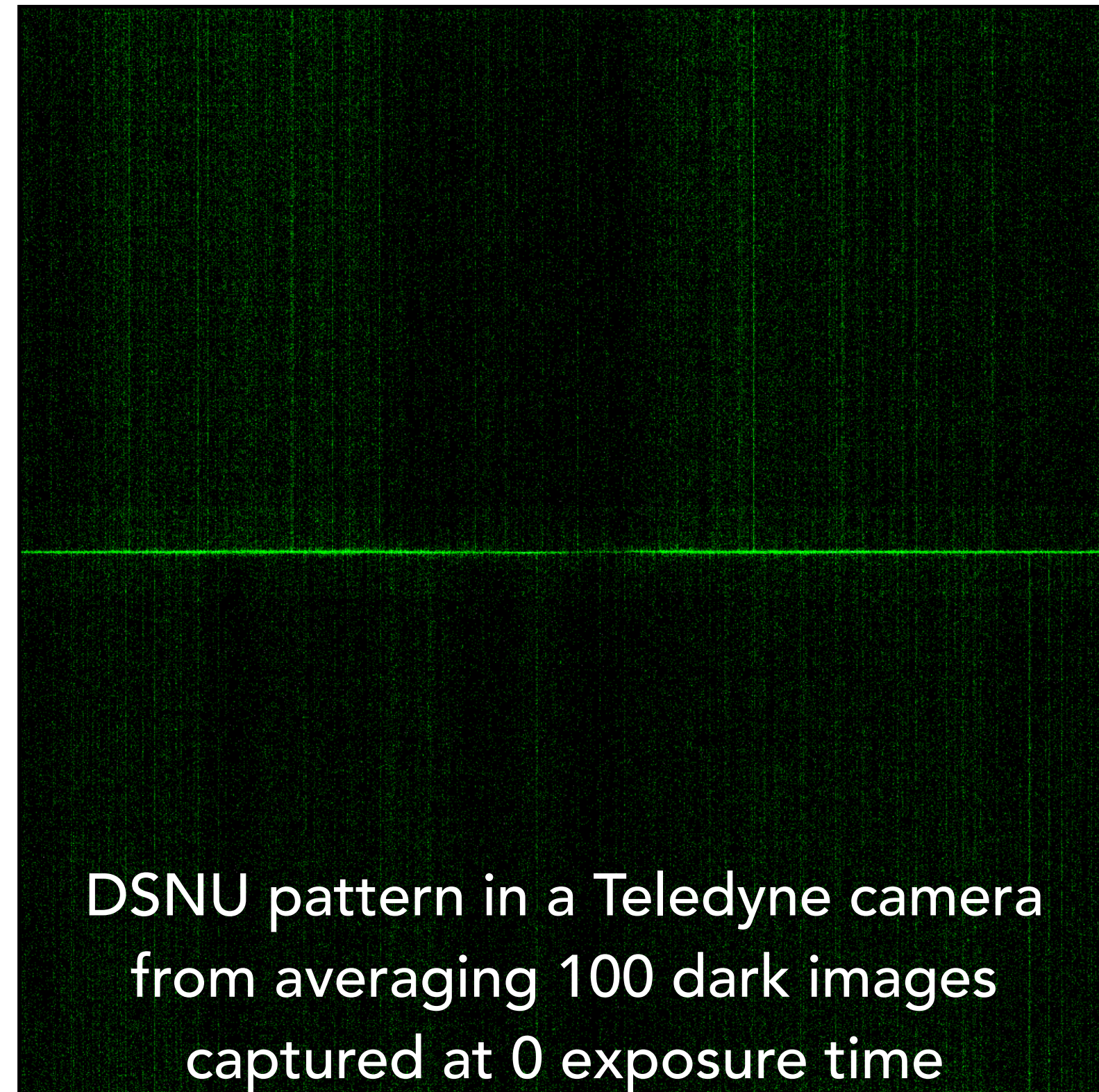
- Many circuit blocks add a voltage offset/bias (e.g., in-pixel amplifiers/column ADC) or inherently introduce a voltage offset (e.g., charge injection on FD after PD to FD transfer). The offset voltage, while a constant for a pixel/column, varies across pixels.
- Think of SONU as adding a constant voltage offset to each pixel/column.

The effects of SONU:

- unaffected by exposure time
 - unaffected by irradiance
 - increase with ISO setting
- } Illumination level

DSNU Pattern (Dark Frame) Examples

DSNU pattern under an exposure time T and ISO setting is obtained by averaging multiple dark frames.



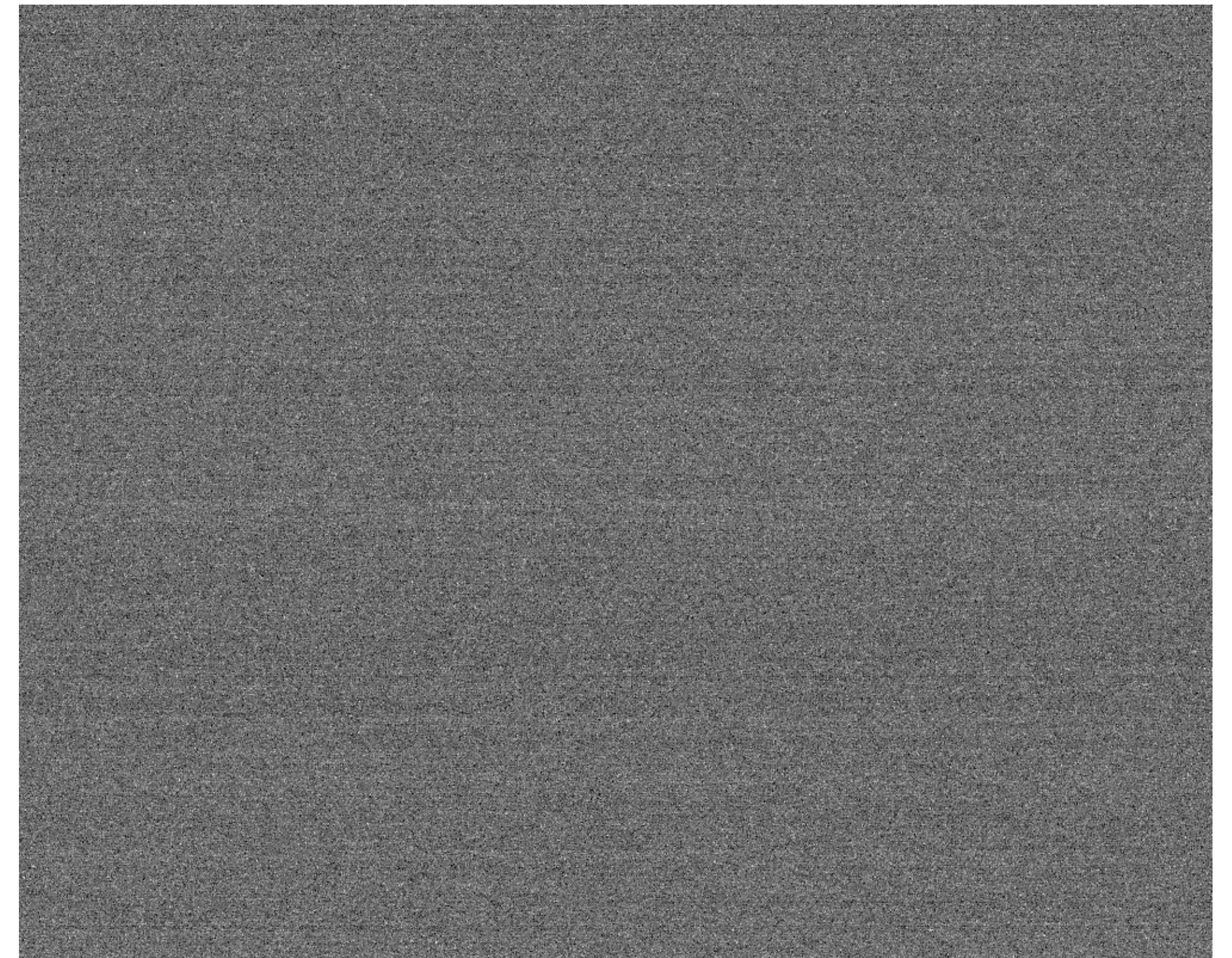
Removing DSNU Using Dark-Frame Subtraction

Capture N dark frames, each captured at dark (with lens cap on; 0 incident photon) with an exposure time T and an ISO setting.

Average the N dark frames to obtain an averaged dark frame at $\langle T, \text{ISO} \rangle$.

For an actual capture with $\langle T, \text{ISO} \rangle$, subtract the corresponding averaged dark frame.

AThewcAlborate20Dningagaft@rSLarBfDa&dsubtractionsome



PRNU

Think of/model PRNU as gain variations across pixels/columns.

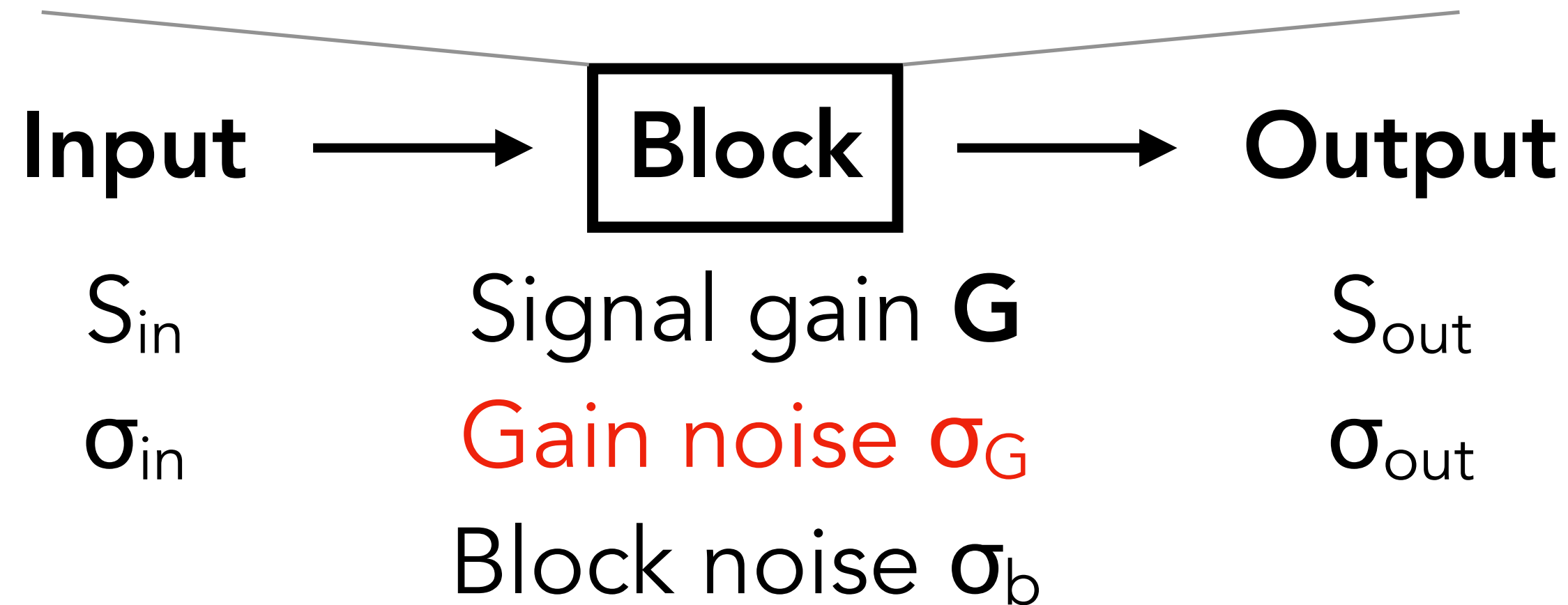
- Under illumination, electrical output given the same amount of incident photons isn't uniform across pixels/columns (e.g., variation in full well capacity, quantum efficiency, capacitances and transistors at each pixel/column, etc.).

The effects of PRNU:

- increase with exposure time
 - increase with irradiance
 - increase with ISO setting
- } **Illumination level**

Noise Amplification With PRNU

- Photon to electron conversion in PD
- electron charge to voltage conversion in FD
- SF in read-out
- other amplifiers (e.g., for ISO control)
- ADC, etc.

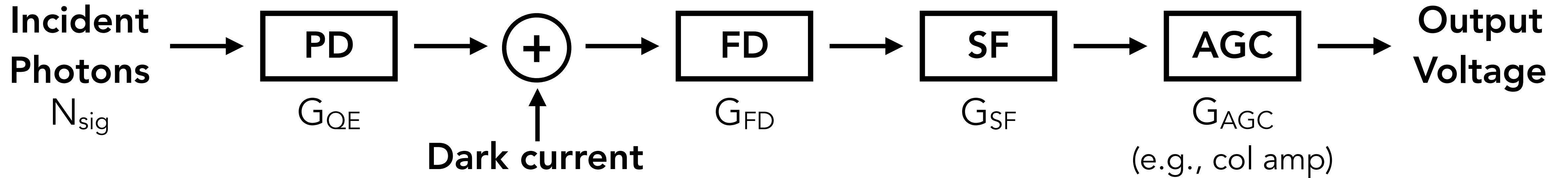


$$\sigma_{out} = \sqrt{G^2 \sigma_{in}^2 + S_{in}^2 \sigma_G^2 + \sigma_b^2}$$

Ignoring second-order terms:

$$\text{Var}(XY) = \text{Var}(X)\text{Var}(Y) + \text{Var}(X)(E(Y))^2 + \text{Var}(Y)(E(X))^2 \approx \text{Var}(X)(E(Y))^2 + \text{Var}(Y)(E(X))^2$$

A Sensor Model with Temporal Noise and FPN



The output voltage of an arbitrary pixel in an arbitrary frame (assuming reset noise is taken care of)

$$\left(\left(\left(X_{N_{sig}}^t X_{G_{QE}}^s + X_{N_{dc}}^t \right) X_{G_{FD}}^s + X_{V_{read-FD}}^t \right) X_{G_{SF}}^s + X_{V_{read-SF}}^t + X_{V_{pix-ofs}}^s \right) X_{G_{AGC}}^s + X_{V_{read-AGC}}^t + X_{V_{col-ofs}}^s$$

Temporal noise (expressed as random variables (r.v.) whose values vary across captures)

FPN (expressed as r.v. whose values are fixed across frames but vary across pixels; *see caveats next slide)

Photon shot noise	{	$X_{N_{sig}}^t \sim \text{Poisson}(\mu = N_{sig}, \sigma = \sigma_{sig} = \sqrt{N_{sig}})$
Dark current noise		$X_{N_{dc}}^t \sim \text{Poisson}(\mu = X_{N_{dc}}^s, \sigma = \sigma_{dc} = \sqrt{X_{N_{dc}}^s})$
Read noise		$X_{V_{read-FD/read-SF/read-AGC}}^t \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-FD/read-SF/read-AGC}})$
PRNU		$X_{G_{QE/FD/SF/AGC}}^s \sim \text{N}(\mu = G_{QE/FD/SF/AGC}, \sigma = \sigma_{G_{QE/FD/SF/AGC}})$
SONU	{	$X_{V_{pix-ofs/col-ofs}}^s \sim \text{N}(\mu = V_{pix-ofs/col-ofs}, \sigma = \sigma_{V_{pix-ofs/col-ofs}})$
DCNU		$X_{N_{dc}}^s \sim \text{N}(\mu = N_{dc}, \sigma = \sigma_{N_{dc}})$

A Note About Probabilistic Sensor Modeling

It's correct to model temporal noises using random variables. When calculate the mean and std across pixels (under uniform illumination) or frames, the goal is to use the sample mean/std to approximate the expected value/std of the underlying distribution, which is what we care about.

It's technically **incorrect** to model FPN of a given sensor using random variables. It's true that there probably is an underlying distribution of pixel manufacturing variations (which could be Gaussian), but the spatial distribution on a given sensor is fixed once manufactured, which gives a **distribution of samples**.

- Importantly, when we calculate mean/std of spatial non-uniformities over pixels, we are calculating the **sample mean/std**, which are indeed what we care about, because they are statistics specific to the sensor of interest. We are **not** using sample mean/std to approximate population mean/std (i.e., underlying distribution). For instance, G_{FD} represents the sample mean of FD gain across the pixels on a particular sensor, not the expected value of the FD gain for a given manufacturing facility.

Also note that the spatial distribution on a given sensor is most likely not perfectly Gaussian, although close; see https://www.emva.org/wp-content/uploads/EMVA1288General_4.0Release.pdf, Figure 13.

Reducing FPN Using CDS

The SONU portion of FPN can be eliminated by Correlated Double Sampling (CDS), which, for each pixel during each capture, reads twice — first the reset voltage and second the actual voltage signal under exposure — and then subtracts the two.

CDS can't remove PRNU/DCNU, because the effects of dark current and gain variations are not simply adding a constant voltage offset to each reading.

CDS can also suppress:

- reset noise (part of thermal-induced read noise).
- some read noise (the $1/f$ part, which is correlated between the two samples).

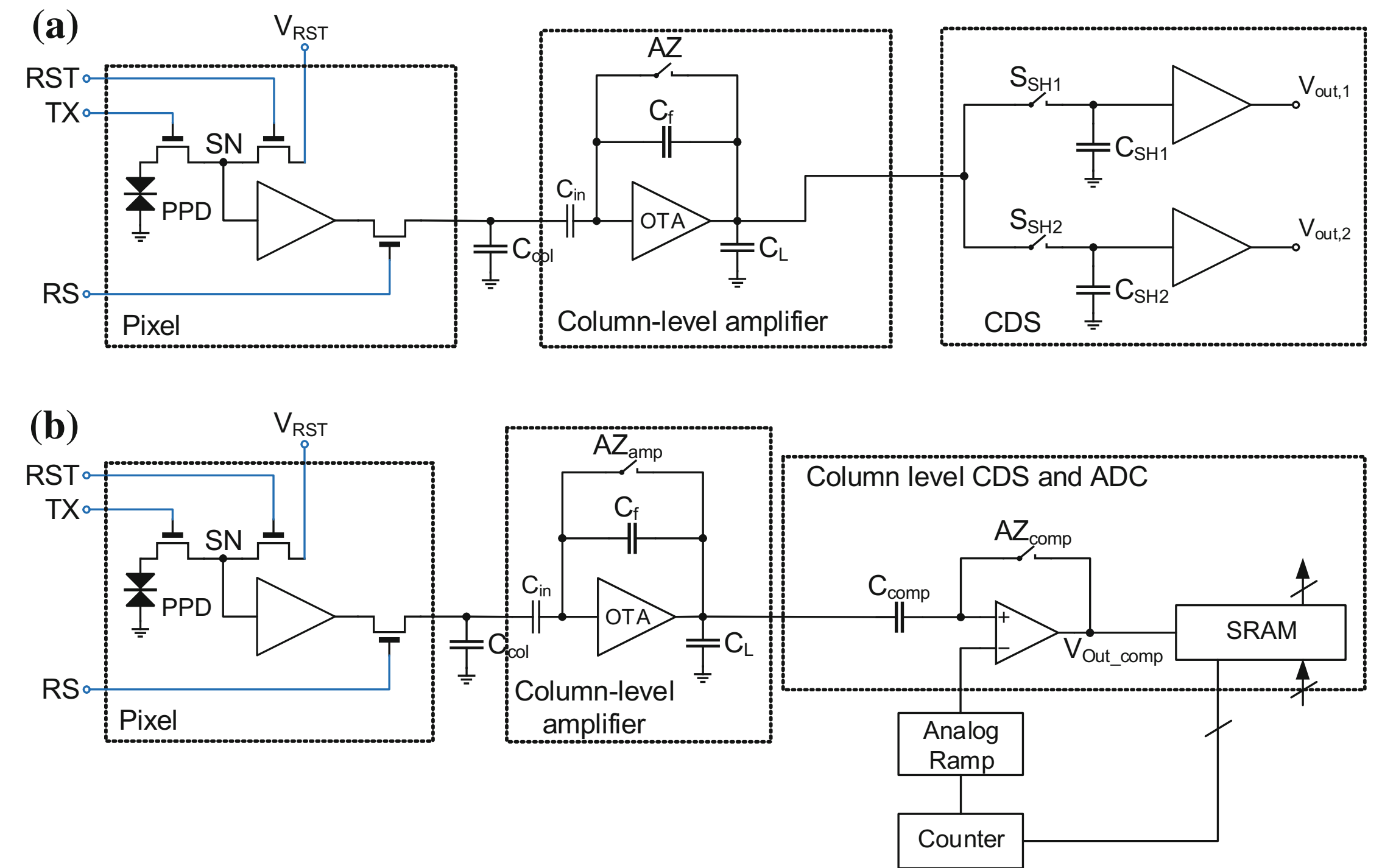
CDS Circuits

Usually sits right before column ADC

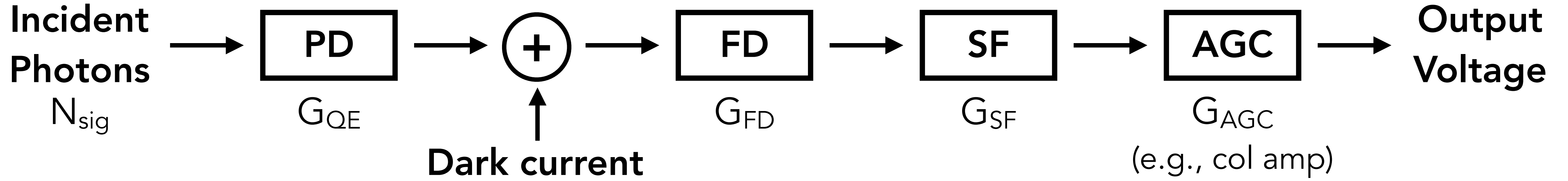
- Therefore after FD, SF, and column amp.

Two implementations:

- two sampling caps for reset and signal, and then ADC takes differential input.
- one sampling cap, and the subtraction happens on that cap by sequentially sampling the reset and signal; then ADC takes the subtraction result as input.



Effect of CDS



First sample:

$$((X_{N_{rst}} G_{FD} + X_{V_{read-FD}}) G_{SF} + X_{V_{read-SF}} + V_{pix-offset}) G_{AGC} + X_{V_{read-AGC}} + V_{col-offset}$$

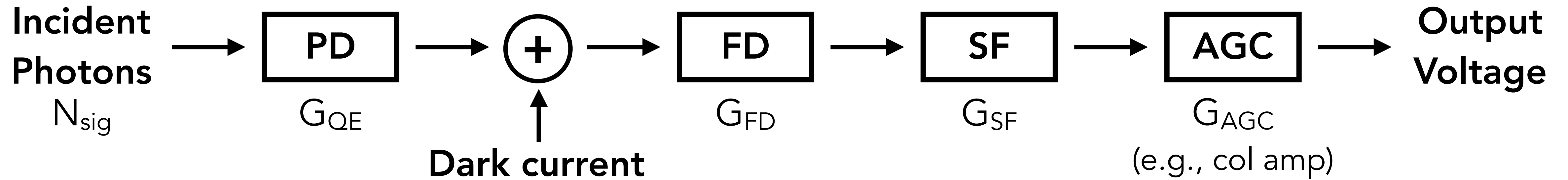
Second sample (**subtract** N_{sig} and N_{dc} because charge transfer **discharges** FD):

$$(((X_{N_{rst}} - X_{N_{sig}} G_{QE} - X_{N_{dc}}) G_{FD} + Y_{V_{read-FD}}) G_{SF} + Y_{V_{read-SF}} + V_{pix-offset}) G_{AGC} + Y_{V_{read-AGC}} + V_{col-offset} + Y_{V_{read-CDS}}$$

For a specific pixel only temporal noises need to be considered

$$\left\{ \begin{array}{l} X_{N_{sig}} \sim \text{Poisson}(\mu = N_{sig}, \sigma = \sigma_{sig} = \sqrt{N_{sig}}) \\ X_{N_{dc}} \sim \text{Poisson}(\mu = N_{dc}, \sigma = \sigma_{dc} = \sqrt{N_{dc}}) \\ X_{N_{rst}} \sim \text{N}(\mu = N_{rst} = V_{rst} C_{FD}, \sigma = \sigma_{rst}) \end{array} \right. \quad \left\{ \begin{array}{l} X_{V_{read-FD}}, Y_{V_{read-FD}} \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-FD}}) \\ X_{V_{read-SF}}, Y_{V_{read-SF}} \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-SF}}) \\ X_{V_{read-AGC}}, Y_{V_{read-AGC}} \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-AGC}}) \\ X_{V_{read-ADC}}, Y_{V_{read-ADC}} \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-ADC}}) \end{array} \right.$$

Effect of CDS



First sample:

$$((X_{N_{rst}} G_{FD} + X_{V_{read-FD}}) G_{SF} + X_{V_{read-SF}} + V_{pix-offset}) G_{AGC} + X_{V_{read-AGC}} + V_{col-offset}$$

Second sample (**subtract** N_{sig} and N_{dc} because charge transfer **discharges** FD):

$$(((X_{N_{rst}} - X_{N_{sig}} G_{QE} - X_{N_{dc}}) G_{FD} + Y_{V_{read-FD}}) G_{SF} + Y_{V_{read-SF}} + V_{pix-offset}) G_{AGC} + Y_{V_{read-AGC}} + V_{col-offset} + Y_{V_{read-CDS}}$$

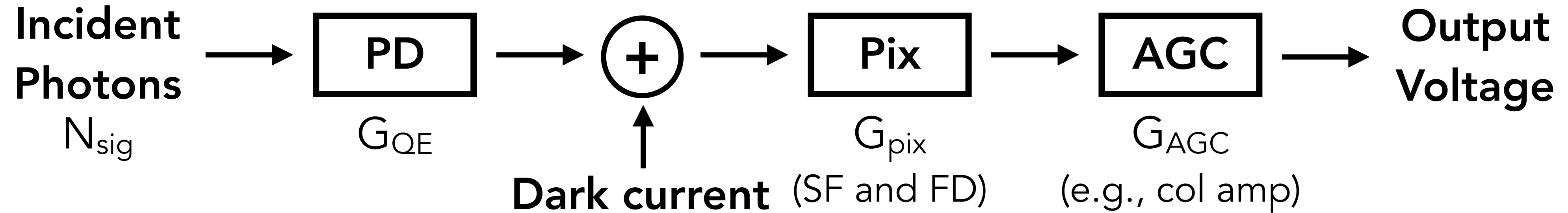
After subtraction:

$$(((X_{N_{sig}} G_{QE} + X_{N_{dc}}) G_{FD} + X_{V_{read-FD}} - Y_{V_{read-FD}}) G_{SF} + X_{V_{read-SF}} - Y_{V_{read-SF}}) G_{AGC} + X_{V_{read-AGC}} - Y_{V_{read-AGC}} - Y_{V_{read-CDS}}$$

- **Eliminated:** the offset portion of FPN, reset noise (kTC-induced), and some amount of read noise (1/f component)
- **Retained:** PRNU/DCNU, other temporal noise (in fact, variance of non-reset/non-1/f read noise is doubled)
- **Introduced:** read noise of the CDS circuitry itself

Signal Processing For Noise Removal/Estimation

Recovering Noise-Free Images



1. Under the same illumination level D , take N captures and average them pixel-wise; this effectively calculates the expected value of a composite random variable *for each pixel*:

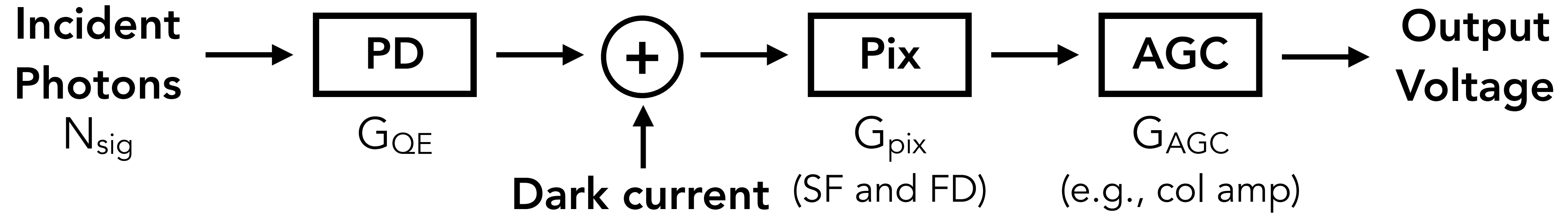
$$E[(X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t) G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p] G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p = \underbrace{((N_{sig} G_{QE}^p + N_{dc}^p) G_{pix}^p + V_{pix-ofs}^p)}_{\text{^p terms are pixel-specific constants}} G_{AGC}^p + \underbrace{V_{col-ofs}^p}_{\text{^p terms are pixel-specific constants}}$$

$$\left\{ \begin{array}{l} X_{N_{sig}}^t \sim \text{Poisson}(\mu = N_{sig}, \sigma = \sigma_{sig} = \sqrt{N_{sig}}) \\ X_{N_{dc}}^t \sim \text{Poisson}(\mu = N_{dc}^p, \sigma = \sigma_{dc} = \sqrt{N_{dc}^p}) \\ X_{V_{read-FD/read-SF/read-AGC}}^t \sim N(\mu = 0, \sigma = \sigma_{V_{read-FD/read-SF/read-AGC}}) \end{array} \right.$$

$$E[XY] = E[X]E[Y]$$

$$E[X + Y] = E[X] + E[Y]$$

Recovering Noise-Free Images



1. Under the same illumination level D , take N captures and average them pixel-wise; this effectively calculates the expected value of a composite random variable *for each pixel*:

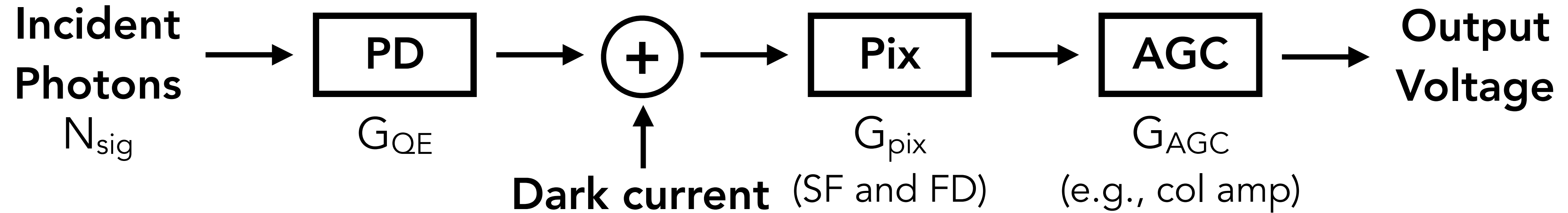
$$E[(X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t) G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p] G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p] = ((N_{sig} G_{QE}^p + N_{dc}^p) G_{pix}^p + V_{pix-ofs}^p) G_{AGC}^p + V_{col-ofs}^p$$

2. Average across all the pixels in the averaged frame obtained before. This effectively calculates the expected value of a composite random variable:

$$E[(N_{sig} X_{G_{QE}}^s + X_{N_{dc}}^s) X_{G_{pix}}^s + X_{V_{pix-ofs}}^s] X_{G_{AGC}}^s + X_{V_{col-ofs}}^s] = ((N_{sig} G_{QE} + N_{dc}) G_{pix} + V_{pix-ofs}) G_{AGC} + V_{col-ofs}$$

$$\begin{cases} X_{G_{QE/pix/AGC}}^s \sim N(\mu = G_{QE/pix/AGC}, \sigma = \sigma_{G_{QE/pix/AGC}}) \\ X_{V_{pix-ofs/col-ofs}}^s \sim N(\mu = V_{pix-ofs/col-ofs}, \sigma = \sigma_{V_{pix-ofs/col-ofs}}) \\ X_{N_{dc}}^s \sim N(\mu = N_{dc}, \sigma = \sigma_{N_{dc}}) \end{cases}$$

Recovering Noise-Free Images



1. Under the same illumination level D , take N captures and average them pixel-wise; this effectively calculates the expected value of a composite random variable *for each pixel*:

$$E[(X_{N_{sig}}^t G_{QE}^P + X_{N_{dc}}^t) G_{pix}^P + X_{V_{read-pix}}^t + V_{pix-ofs}^P] G_{AGC}^P + X_{V_{read-AGC}}^t + V_{col-ofs}^P] = ((N_{sig} G_{QE}^P + N_{dc}^P) G_{pix}^P + V_{pix-ofs}^P) G_{AGC}^P + V_{col-ofs}^P$$

2. Average across all the pixels in the averaged frame obtained before. This effectively calculates the expected value of a composite random variable:

$$E[(N_{sig} X_{G_{QE}}^S + X_{N_{dc}}^S) X_{G_{pix}}^S + X_{V_{pix-ofs}}^S] X_{G_{AGC}}^S + X_{V_{col-ofs}}^S] = ((N_{sig} G_{QE} + N_{dc}) G_{pix} + V_{pix-ofs}) G_{AGC} + V_{col-ofs}$$

3. We subtract the two above arithmetically, we get the expected value of FPN at D for a pixel:

$$FPN = ((N_{sig} G_{QE}^P + N_{dc}^P) G_{pix}^P + V_{pix-ofs}^P) G_{AGC}^P + V_{col-ofs}^P - ((N_{sig} G_{QE} + N_{dc}) G_{pix} + V_{pix-ofs}) G_{AGC} - V_{col-ofs}$$

Recovering Noise-Free Images

4. We take a capture of a scene of interest, and for a pixel whose illumination level is D , we subtract its voltage reading (a composite random variable itself) with the FPN at D corresponding to that pixel (we do this subtraction for all pixels):

$$\begin{aligned}
 V^p &= ((X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t) G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p) G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p - FPN \\
 &= ((X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t) G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p) G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p \\
 &\quad - ((N_{sig} G_{QE}^p + N_{dc}^p) G_{pix}^p + V_{pix-ofs}^p) G_{AGC}^p - V_{col-ofs}^p + ((N_{sig} G_{QE} + N_{dc}) G_{pix} + V_{pix-ofs}) G_{AGC} + V_{col-ofs} \\
 &= (X_{N_{sig}}^t - N_{sig}) G_{QE}^p G_{pix}^p G_{AGC}^p + (X_{N_{dc}}^t - N_{dc}^p) G_{pix}^p G_{AGC}^p + X_{V_{read-pix}}^t G_{AGC}^p + X_{V_{read-AGC}}^t \\
 &\quad + ((N_{sig} G_{QE} + N_{dc}) G_{pix} + V_{pix-ofs}) G_{AGC} + V_{col-ofs}
 \end{aligned}$$

Given this, there are a few interesting conclusions we can draw. **4.1:** the expected value of each so-calculated pixel is the true signal of that pixel, which means if we **average over many frames**, we will get a **completely noise-free image**.

$$E[V^p] = ((N_{sig} G_{QE} + N_{dc}) G_{pix} + V_{pix-ofs}) G_{AGC} + V_{col-ofs}$$

The dark current offset and pixel/column voltage offsets are still there, but they are constant across pixels (hence not noise). Consequently, the black level voltage won't be 0. **If the sensor first performs CDS** before going through our subtraction procedure, the constant voltage offsets will be eliminated, but dark current offset is still there.

Recovering Noise-Free Images

4.2: each pixel itself, without averaging across frames, has a noise of the following form. We can see that the SONU is gone, but the PRNU and DSNU, although fixed for the pixel, is still there in addition to temporal noises. So the effect is similar to using CDS (which removes 1/f read noise and amplifies other read noises; see before).

$$\begin{aligned} \text{Var}[V^P] &= \text{Var}[(X_{N_{sig}}^t - N_{sig})G_{QE}^P G_{pix}^P G_{AGC}^P + (X_{N_{dc}}^t - N_{dc}^P)G_{pix}^P G_{AGC}^P + X_{V_{read-pix}}^t G_{AGC}^P + X_{V_{read-AGC}}^t] \\ &= \sigma_{sig}^2 (G_{QE}^P G_{pix}^P G_{AGC}^P)^2 + N_{dc}^P (G_{pix}^P G_{AGC}^P)^2 + \sigma_{V_{read-pix}}^2 (G_{AGC}^P)^2 + \sigma_{V_{read-AGC}}^2 \end{aligned}$$

4.3: if we want to show the noise both spatially and temporally (i.e., capture both FPN and temporal noises), then the voltage value at each pixel is expressed as a random variable below:

$$(X_{N_{sig}}^t - N_{sig})X_{G_{QE}}^S X_{G_{pix}}^S X_{G_{AGC}}^S + (X_{N_{dc}}^t - N_{dc}^P)X_{G_{pix}}^S X_{G_{AGC}}^S + X_{V_{read-pix}}^t X_{G_{AGC}}^S + X_{V_{read-AGC}}^t$$

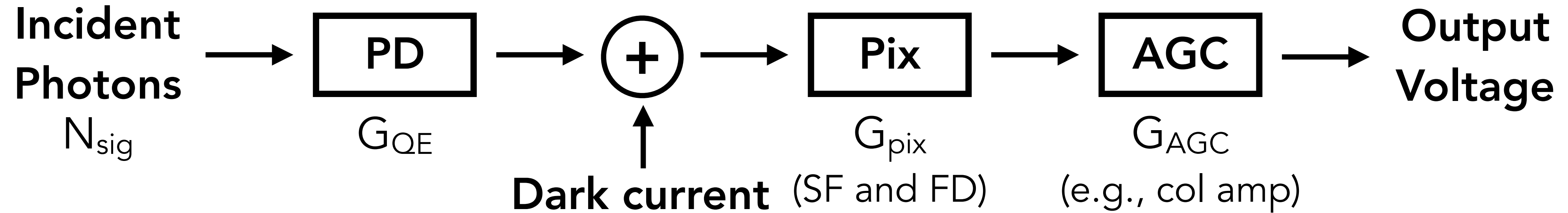
$$\text{Var}(X_1 X_2 \cdots X_n) = \prod_{i=1}^n (\sigma_i^2 + \mu_i^2) - \prod_{i=1}^n \mu_i^2$$

The variance of this random variable is:

$$\begin{aligned} \text{Var} &= (\sigma_{sig}^2 G_{QE}^2 G_{pix}^2 + \sigma_{sig}^2 \sigma_{G_{QE}}^2 G_{G_{pix}}^2 + \sigma_{sig}^2 G_{QE}^2 \sigma_{G_{pix}}^2 + \sigma_{sig}^2 \sigma_{G_{QE}}^2 \sigma_{G_{pix}}^2 + N_{dc}^P \sigma_{G_{pix}}^2 + N_{dc}^P G_{pix}^2 + \sigma_{read-pix}^2) G_{AGC}^2 + \sigma_{read-AGC}^2 \\ &\approx (\sigma_{sig}^2 G_{QE}^2 G_{pix}^2 + N_{dc}^P G_{pix}^2 + \sigma_{read-FD}^2) G_{AGC}^2 + \sigma_{read-AGC}^2 \end{aligned}$$

We can see that when omitting high-order terms, the noise is the same as that without any FPN. So in general we can say that the so-obtained image is **FPN-free**. Alternatively, we can say that given a uniform illumination D, the pixel value differences in the image are due only to temporal noises.

Dark-Frame Subtraction to Remove DSNU



1. Under the same illumination level D , take N captures and average them pixel-wise; this effectively calculates the expected value of a composite random variable *for each pixel*:

$$E[(X_{N_{sig}}^t G_{QE}^P + X_{N_{dc}}^t) G_{pix}^P + X_{V_{read-pix}}^t + V_{pix-ofs}^P] G_{AGC}^P + X_{V_{read-AGC}}^t + V_{col-ofs}^P = ((N_{sig} G_{QE}^P + N_{dc}^P) G_{pix}^P + V_{pix-ofs}^P) G_{AGC}^P + V_{col-ofs}^P$$

When the illumination is 0, the frame is called a **dark frame**. After averaging above what we get is essentially a **DSNU pattern** (including both DCNU and SONU) under a given exposure time. Figure 3 of [this article](#) obtains the DSNU pattern under an 1ms exposure time by averaging 16 dark frames.

$$(N_{dc}^P G_{pix}^P + V_{pix-ofs}^P) G_{AGC}^P + V_{col-ofs}^P$$

If the exposure time is set to 0, each capture is called a **bias frame**. Figure 2 of [this article](#) is obtained by averaging 100 bias frames. In an averaged bias frame, the voltage reading of each pixel is (no dark current):

$$V_{pix-ofs}^P G_{AGC}^P + V_{col-ofs}^P$$

Dark-Frame Subtraction to Remove DSNU

2. We take a capture of a scene of interest with an exposure time T, and subtract, pixel-wise, the averaged dark frame obtained at T. The voltage reading of each pixel is a random variable:

$$\begin{aligned}
 V^p &= ((X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t) G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p) G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p - (N_{dc}^p G_{pix}^p + V_{pix-ofs}^p) G_{AGC}^p - V_{col-ofs}^p \\
 &= X_{N_{sig}}^t G_{QE}^p G_{pix}^p G_{AGC}^p + (X_{N_{dc}}^t - N_{dc}^p) G_{pix}^p G_{AGC}^p + X_{V_{read-pix}}^t G_{AGC}^p + X_{V_{read-AGC}}^t
 \end{aligned}$$

The expected value of this random variable (which can be approximated by averaging multiple frames) is the true signal without dark current and voltage offsets, but PRNU is still there. In other words, the noise in the frame is due only to PRNU.

$$E[V^p] = N_{sig} G_{QE}^p G_{pix}^p G_{AGC}^p$$

It's called dark-frame subtraction, which eliminates dark current noise and SONU. In astrophotography, it's called dark current and bias correction (bias being the constant voltage offsets in our language). Note that it removes the impact of dark current altogether (both DCNU and the shot noise of dark current). That's why it's often used in astrophotography, where the exposure time is usually excessively long.

Two things need to be noted. First, the dark frame to be subtracted must be taken under the same exposure time as the intended capture. Second, dark-frame subtraction does not eliminate FPN (which is a common misconception in online articles) because PRNU is still there.

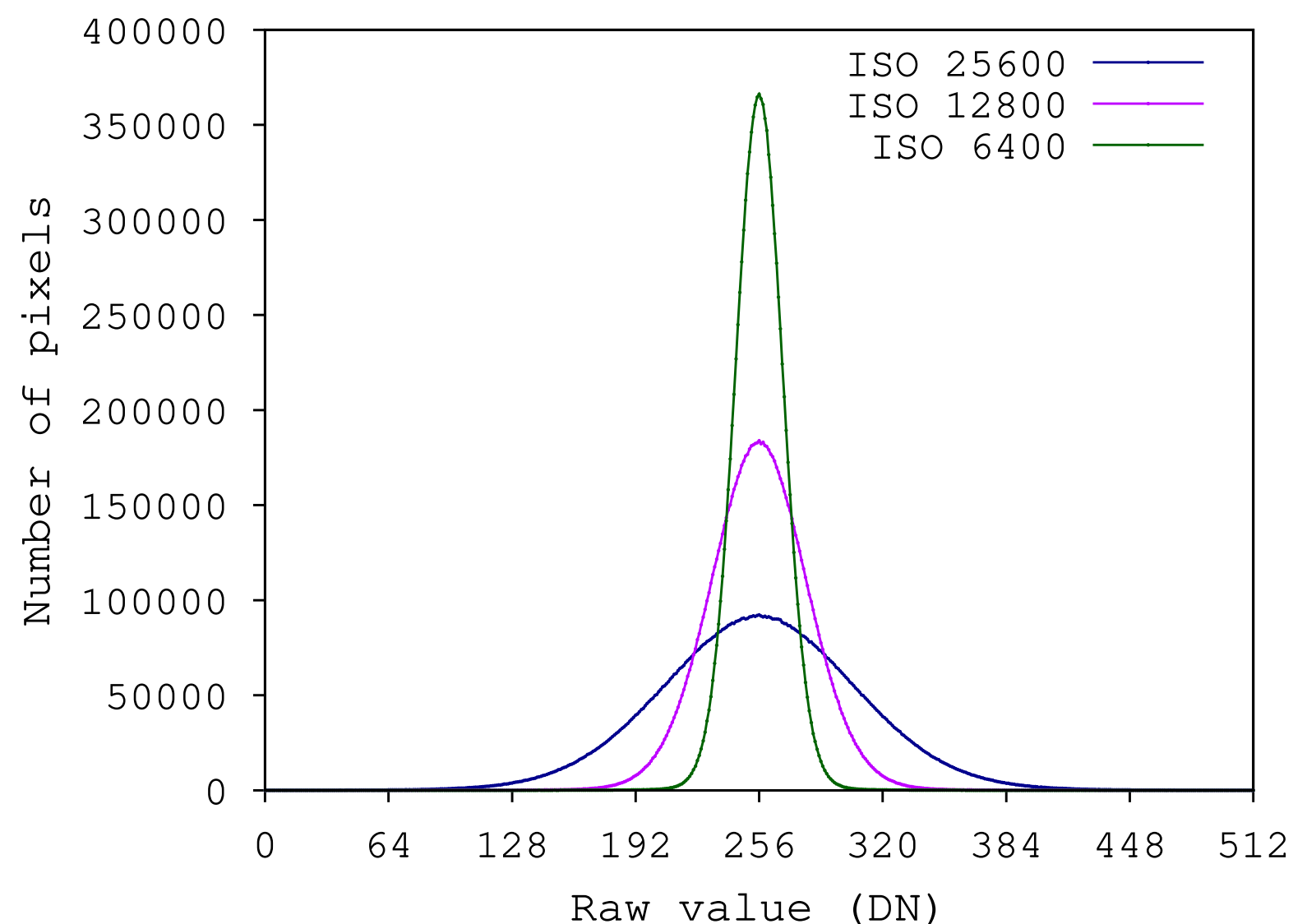
Noise Analyses Using Dark Frames and Bias Frames

Here are a few interesting things we can do now that we have dark and bias frames.

1: If we subtract an averaged bias frame from a single bias frame and plot the pixel value histogram (or plot the histogram over multiple frames), what we get is the output-referred read noise distribution of the sensor.

$$(X_{V_{read-pix}}^t + V_{pix-ofs}^P)G_{AGC}^P + X_{V_{read-AGC}}^t + V_{col-ofs}^P - (V_{pix-ofs}^P G_{AGC}^P + V_{col-ofs}^P) = X_{V_{read-pix}}^t G_{AGC}^P + X_{V_{read-AGC}}^t$$

Alternatively, we could also just plot the histogram of a single bias frame, in which case we will see a Gaussian with a **non-zero** mean, where the mean is due to the constant pixel/column voltage biases, but the distribution still shows read noise.



- Gaussian read noise distribution measured in raw values of a single bias frame for the Olympus E-M1 at a selection of high ISO settings.
- The distributions are centered at the DN = 256, which shows the bias offset added by the sensor (ADC bias or other pixel/column offsets). Without no bias/offset, the negative values will be clipped to 0.
- The std increases with the ISO setting, because the ISO setting dictates the gain, which amplifies the variance (noise).
- Strictly speaking, the distribution is not purely from read noise, as PRNU/SONU are still there. Subtracting the averaged bias frame eliminates SONU.

Noise Analyses Using Dark Frames and Bias Frames

2: If we subtract an averaged bias frame from an averaged dark frame, what we get is the dark current component of each pixel or, more specifically, the expected value of output-referred dark current-induced voltage of a pixel:

$$(N_{dc}^p G_{pix}^p + V_{pix-ofs}^p) G_{AGC}^p + V_{col-ofs}^p - (V_{pix-ofs}^p G_{AGC}^p + V_{col-ofs}^p) = N_{dc}^p G_{pix}^p G_{AGC}^p$$

3: Take a look at the expected value of a pixel after dark-frame subtraction (i.e., averaging over multiple frames with dark frame subtraction), what we get is the PRNU distribution of the sensor, because the noise in the frame is due only to PRNU.

$$E[V^p] = N_{sig} G_{QE}^p G_{pix}^p G_{AGC}^p$$

We can see that the affect of PRNU is dependent on the illumination level (incident irradiance and exposure time).

Flat-Field Correction to Remove Noise

1. Take a raw capture of a scene of interest

$$R = ((X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t) G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p) G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p$$

2. Obtain an averaged dark frame

$$D = (N_{dc}^p G_{pix}^p + V_{pix-ofs}^p) G_{AGC}^p + V_{col-ofs}^p$$

3. Obtain an averaged flat-field frame, i.e., averaging frames under uniform illumination

$$F = ((N_{sig} G_{QE}^p + N_{dc}^p) G_{pix}^p + V_{pix-ofs}^p) G_{AGC}^p + V_{col-ofs}^p$$

4. Obtain dark noise and bias-corrected flat-field frame

$$F - D = N_{sig} G_{QE}^p G_{pix}^p G_{AGC}^p$$

5. Calculate the pixel average of the corrected flat-field frame

$$m = N_{sig} G_{QE} G_{pix} G_{AGC}$$

6. Obtain dark noise and bias-corrected raw frame

$$R - D = X_{N_{sig}}^t G_{QE}^p G_{pix}^p G_{AGC}^p + (X_{N_{dc}}^t - N_{dc}^p) G_{pix}^p G_{AGC}^p + X_{V_{read-pix}}^t G_{AGC}^p + X_{V_{read-AGC}}^t$$

7. Calculate the final, calibrate image. After frame averaging, the only noise is photon shot noise

$$\begin{aligned} \frac{(R - D)m}{F - D} &= \frac{X_{N_{sig}}^t G_{QE}^p G_{pix}^p G_{AGC}^p + (X_{N_{dc}}^t - N_{dc}^p) G_{pix}^p G_{AGC}^p + X_{V_{read-pix}}^t G_{AGC}^p + X_{V_{read-AGC}}^t}{N_{sig} G_{QE}^p G_{pix}^p G_{AGC}^p} \times N_{sig} G_{QE} G_{pix} G_{AGC} \\ &= X_{N_{sig}}^t G_{QE} G_{pix} G_{AGC} + (X_{N_{dc}}^t - N_{dc}^p) \frac{G_{QE} G_{pix}}{G_{QE}^p G_{pix}^p} G_{AGC} + X_{V_{read-pix}}^t \frac{G_{QE} G_{pix}}{G_{QE}^p G_{pix}^p} G_{AGC} + X_{V_{read-AGC}}^t \frac{G_{QE} G_{pix} G_{AGC}}{G_{QE}^p G_{pix}^p G_{AGC}^p} \end{aligned}$$

$$E\left[\frac{(R - D)m}{F - D}\right] = X_{N_{sig}}^t G_{QE} G_{pix} G_{AGC}$$

Estimating Temporal Noise and FPN

Described by Emil Martinec [here](#), which is very similar to the [EMVA Standard 1288](#) (Release 4.0), which is a black-box model that doesn't specify the exact analytical form of the noise being measured, which we show here. Ignore dark current and DCNU.

1. Take two captures under uniform illumination D ; voltages of the same pixel in the two images:

$$\begin{aligned} & (X_{N_{sig}}^t G_{QE}^p G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p) G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p \\ & (Y_{N_{sig}}^t G_{QE}^p G_{pix}^p + Y_{V_{read-pix}}^t + V_{pix-ofs}^p) G_{AGC}^p + Y_{V_{read-AGC}}^t + V_{col-ofs}^p \end{aligned} \quad \left\{ \begin{array}{l} X_{N_{sig}}^t, Y_{N_{sig}}^t \sim \text{Poisson}(\mu = N_{sig}, \sigma = \sigma_{sig} = \sqrt{N_{sig}}) \\ X_{V_{read-FD/read-SF/read-AGC}}^t, Y_{V_{read-FD/read-SF/read-AGC}}^t \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-FD/read-SF/read-AGC}}) \end{array} \right.$$

2. Subtract the two images pixel-wise; each pixel is expressed as:

$$((X_{N_{sig}}^t - Y_{N_{sig}}^t) G_{QE}^p G_{pix}^p + (X_{V_{read-pix}}^t - Y_{V_{read-pix}}^t)) G_{AGC}^p + (X_{V_{read-AGC}}^t - Y_{V_{read-AGC}}^t)$$

3. Now we want to calculate the variance across all pixels in this residual image, in which case FPN terms become random variables. So the voltage of each pixel in the residual image is expressed as a new random variable:

$$\begin{aligned} & ((X_{N_{sig}}^t - Y_{N_{sig}}^t) X_{G_{QE}}^s X_{G_{pix}}^s + (X_{V_{read-pix}}^t - Y_{V_{read-pix}}^t)) X_{G_{AGC}}^s + (X_{V_{read-AGC}}^t - Y_{V_{read-AGC}}^t) \\ & X_{G_{QE/pix/AGC}}^s \sim \text{N}(\mu = G_{QE/pix/AGC}, \sigma = \sigma_{G_{QE/pix/AGC}}) \end{aligned}$$

Note that it's incorrect to use the following random variable, which assumes we sample both temporally **AND** spatially before subtraction; instead, we sample temporally, subtract, and then sample spatially.

$$(X_{N_{sig}}^t X_{G_{QE}}^s X_{G_{pix}}^s + X_{V_{read-pix}}^t + V_{pix-ofs}^p) X_{G_{AGC}}^s + X_{V_{read-AGC}}^t + V_{col-ofs}^p - ((Y_{N_{sig}}^t X_{G_{QE}}^s X_{G_{pix}}^s + Y_{V_{read-pix}}^t + V_{pix-ofs}^p) X_{G_{AGC}}^s + Y_{V_{read-AGC}}^t + V_{col-ofs}^p)$$

Estimating Temporal Noise and FPN

4. Now calculate the variance across N pixels. Assuming N is sufficiently large, the sample variance approaches the population variance, which is expressed as:

$$\text{Var}[\left((X_{N_{sig}}^t - Y_{N_{sig}}^t)X_{G_{QE}}^s X_{G_{pix}}^s + (X_{V_{read-pix}}^t - Y_{V_{read-pix}}^t)X_{G_{AGC}}^s + (X_{V_{read-AGC}}^t - Y_{V_{read-AGC}}^t)\right)]$$

5. Ignore high-order terms, the variance turns out to be exactly the output-referred temporal noise without any FPN:

$$\begin{aligned} &\approx (2\sigma_{sig}^2 G_{QE}^2 G_{pix}^2 + 2\sigma_{read-pix}^2)G_{AGC}^2 + 2\sigma_{read-AGC}^2 \\ &= 2\sigma_{output-referred-temporal-noise}^2 \end{aligned}$$

$$\boxed{\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)}$$

It is worth noting that this methodology to estimate temporal noise is almost exactly the same as that used in the EMVA Standard 1288 (see Equation 18) with one caveat: the EMVA standard also subtracts the difference of the mean values of the two images, which in the ideal case should be exactly the same. The standard does this correction to give *“an unbiased estimate even if the mean values are slightly different by a temporal noise source that causes all pixels to fluctuate in sync.”*

Estimating Temporal Noise and FPN

6. Look at the temporal noise expression, we can see that the temporal noise scales linearly with the illumination level (N_{sig}) and has a constant intercept, which is the read noise.

$$\begin{aligned}\sigma_{output-referred-temporal-noise}^2 &\approx \sigma_{sig}^2 G_{QE}^2 G_{pix}^2 G_{AGC}^2 + \sigma_{read-pix}^2 G_{AGC}^2 + \sigma_{read-AGC}^2 \\ &= \underbrace{N_{sig} G_{QE}^2 G_{pix}^2 G_{AGC}^2}_{\text{Output-referred photon shot noise; increases linearly with } \sigma_{sig}^2=N_{sig}, \text{ which is the incident photon count}} + \underbrace{\sigma_{read}^2}_{\text{Output-referred read noise; a constant w.r.t. to incident photon count}}\end{aligned}$$

Output-referred photon shot noise; increases linearly with $\sigma_{sig}^2=N_{sig}$, which is the incident photon count

Output-referred read noise; a constant w.r.t. to incident photon count

But ultimately what we can measure is in ADU/DN, so we have to consider ADC and its gain, in which case the output-referred temporal noise, now in the unit of ADU/DN, is expressed as:

$$\begin{aligned}\sigma_{output-referred-temporal-noise}^2 &\approx \sigma_{sig}^2 G_{QE}^2 G_{pix}^2 G_{AGC}^2 G_{ADC}^2 + (\sigma_{read-pix}^2 G_{AGC}^2 G_{ADC}^2 + \sigma_{read-AGC}^2 G_{ADC}^2 + \sigma_{ADC}^2) \\ &= N_{sig} G_{QE}^2 G_{pix}^2 G_{AGC}^2 G_{ADC}^2 + \sigma_{read}^2 = \underbrace{(N_{sig} G_{QE} G_{pix} G_{AGC} G_{ADC})}_{N_{sig} \mathbf{G} = N_{ADU}} \underbrace{G_{QE} G_{pix} G_{AGC} G_{ADC}}_{\mathbf{G}; \text{ unit is ADU/photon}} + \sigma_{read}^2 \\ &= N_{ADU} \mathbf{G} + \sigma_{read}^2\end{aligned}$$

Estimating Temporal Noise and FPN

To estimate the read noise, we regress a linear function of output-referred temporal noise vs. ADU under different illumination levels. The ADU is calculated by first adding the two frames together, calculating the mean raw value, and divide that by two.

The intercept of the linear function is the best-fit read noise. The slope of the linear function has a unit of ADU/photon, i.e., how many ADUs are increased for one photon, which is sometimes called the “conversion factor” and is proportional to the ISO and inversely proportional to the overall gain (how many incident photons are needed to increase ADU by 1).

Note that $N_{sig} G = N_{ADU}$ is true only when all noises are eliminated. This is not achieved in what’s described above, which includes the voltage offset (below), and that’s why Martinec’s article says we need to subtract the bias when calculating N_{ADU} .

$$N_{sig} G + V_{pix-ofs} G_{AGC} G_{ADC} + V_{col-ofs} G_{ADC}$$

$$\begin{aligned} \sigma_{output-referred-temporal-noise}^2 &\approx N_{sig} G_{QE}^2 G_{pix}^2 G_{AGC}^2 G_{ADC}^2 + \sigma_{read}^2 = \underbrace{(N_{sig} G_{QE} G_{pix} G_{AGC} G_{ADC})}_{N_{sig} \mathbf{G} = N_{ADU}} \underbrace{G_{QE} G_{pix} G_{AGC} G_{ADC}}_{\mathbf{G}; \text{ unit is ADU/photon}} + \sigma_{read}^2 \\ &= N_{ADU} G + \sigma_{read}^2 \end{aligned}$$

Estimating Temporal Noise and FPN

7: To calculate FPN, take a capture and calculate its the pixel variance, and subtract the temporal noise calculated before. Intuitively, the capture has FPN+temporal noise, from which we subtract the temporal noise. What we are calculating is:

$$\text{Var}\{(Z_{N_{sig}}^t Z_{G_{QE}}^s Z_{G_{pix}}^s + Z_{V_{read-pix}}^t + Z_{V_{pix-ofs}}^s) Z_{G_{AGC}}^s + Z_{V_{read-AGC}}^t + Z_{V_{col-ofs}}^s\} - \frac{\text{Var}[(X_{N_{sig}}^t - Y_{N_{sig}}^t) X_{G_{QE}}^s X_{G_{pix}}^s + (X_{V_{read-pix}}^t - Y_{V_{read-pix}}^t) X_{G_{AGC}}^s + (X_{V_{read-AGC}}^t - Y_{V_{read-AGC}}^t)]}{2}$$

$$= N_{sig}^2 \sigma^2(Z_{G_{QE}}^s Z_{G_{pix}}^s Z_{G_{AGC}}^s) + \sigma^2(Z_{pix-ofs}^s Z_{G_{AGC}}^s) + \sigma_{col-ofs}^2$$

$$\left\{ \begin{array}{l} Z_{N_{sig}}^t \sim \text{Poisson}(\mu = N_{sig}, \sigma = \sigma_{sig} = \sqrt{N_{sig}}) \\ Z_{V_{read-FD/read-SF/read-AGC}}^t \sim \text{N}(\mu = 0, \sigma = \sigma_{V_{read-FD/read-SF/read-AGC}}) \\ X_{G_{QE/pix/AGC}}^s \sim \text{N}(\mu = G_{QE/pix/AGC}, \sigma = \sigma_{G_{QE/pix/AGC}}) \\ X_{V_{pix-ofs/col-ofs}}^s \sim \text{N}(\mu = V_{pix-ofs/col-ofs}, \sigma = \sigma_{V_{pix-ofs/col-ofs}}) \end{array} \right.$$

In fact, we get the overall output-referred FPN, not just PRNU as the article describes, and the results are exact, i.e., no omission of high-order terms.

This method to estimate FPN is exactly the same as that used in the EMVA Standard 1288 (see Equations 16, 17, 18, 27).

We can report the FPN directly using the std (e- for input-referred noise or mV for output-referred noise) or as the ratio of std/full-well capacity or std/voltage swing, which then is a % number.

$$FPN = \frac{\sqrt{N_{sig}^2 \sigma^2(Z_{G_{QE}}^s Z_{G_{pix}}^s Z_{G_{AGC}}^s) + \sigma^2(Z_{pix-ofs}^s Z_{G_{AGC}}^s) + \sigma_{col-ofs}^2}}{FWC \times G_{Pix} G_{AGC}}$$

Another FPN Estimation

1. Under the same illumination level D , take N captures and average them pixel-wise; this effectively calculates the expected value of a composite random variable *for each pixel*:

$$E[((X_{N_{sig}}^t G_{QE}^p + X_{N_{dc}}^t)G_{pix}^p + X_{V_{read-pix}}^t + V_{pix-ofs}^p)G_{AGC}^p + X_{V_{read-AGC}}^t + V_{col-ofs}^p] = ((N_{sig}G_{QE}^p + N_{dc}^p)G_{pix}^p + V_{pix-ofs}^p)G_{AGC}^p + V_{col-ofs}^p$$

2. Calculate the standard deviation/variance across all the M pixels in the averaged frame obtained before. Given that M and N are sufficiently large, this effectively calculates the standard deviation/variance of a composite random variable:

$$\begin{aligned} & Var[((N_{sig}X_{G_{QE}}^s + X_{N_{dc}}^s)X_{G_{pix}}^s + X_{V_{pix-ofs}}^s)X_{G_{AGC}}^s + X_{V_{col-ofs}}^s] \\ &= N_{sig}^2 \sigma^2(X_{G_{QE}}^s X_{G_{pix}}^s X_{G_{AGC}}^s) + \sigma^2(X_{N_{dc}}^s X_{G_{pix}}^s X_{G_{AGC}}^s) + \sigma^2(X_{V_{pix-ofs}}^s X_{G_{AGC}}^s) + \sigma_{V_{col-ofs}}^2 \end{aligned} \quad \left\{ \begin{array}{l} X_{G_{QE/FD/SF/AGC}}^s \sim N(\mu = G_{QE/FD/SF/AGC}, \sigma = \sigma_{G_{QE/FD/SF/AGC}}) \\ X_{V_{pix-ofs/col-ofs}}^s \sim N(\mu = V_{pix-ofs/col-ofs}, \sigma = \sigma_{V_{pix-ofs/col-ofs}}) \\ X_{N_{dc}}^s \sim N(\mu = N_{dc}, \sigma = \sigma_{N_{dc}}) \end{array} \right.$$

This is the same procedure as a method described in EMVA Standard 1288 (see Equations 28). The standard isn't very clear about the difference in applicability of the two methods (subtracting two images vs. averaging over multiple images).

Dynamic Range

Dynamic Range of a Sensor

$$DR = \frac{\textit{Maximal signal}}{\textit{Noise floor}} = \frac{\textit{Full well capacity}}{\sqrt{I_d \times t + \sigma_{read}^2}}$$

The max signal record when a pixel is saturated

of electrons sensed in the dark, i.e., when no incident photon

Dark noise

Std. of read noise

DR is a single number characterizing the ability of a camera to simultaneously sense the brightest and the darkest points of a scene.

Ideally: maximal signal is full-well capacity; minimal signal: 0

Bigger sensors improve DR, as we saw before.

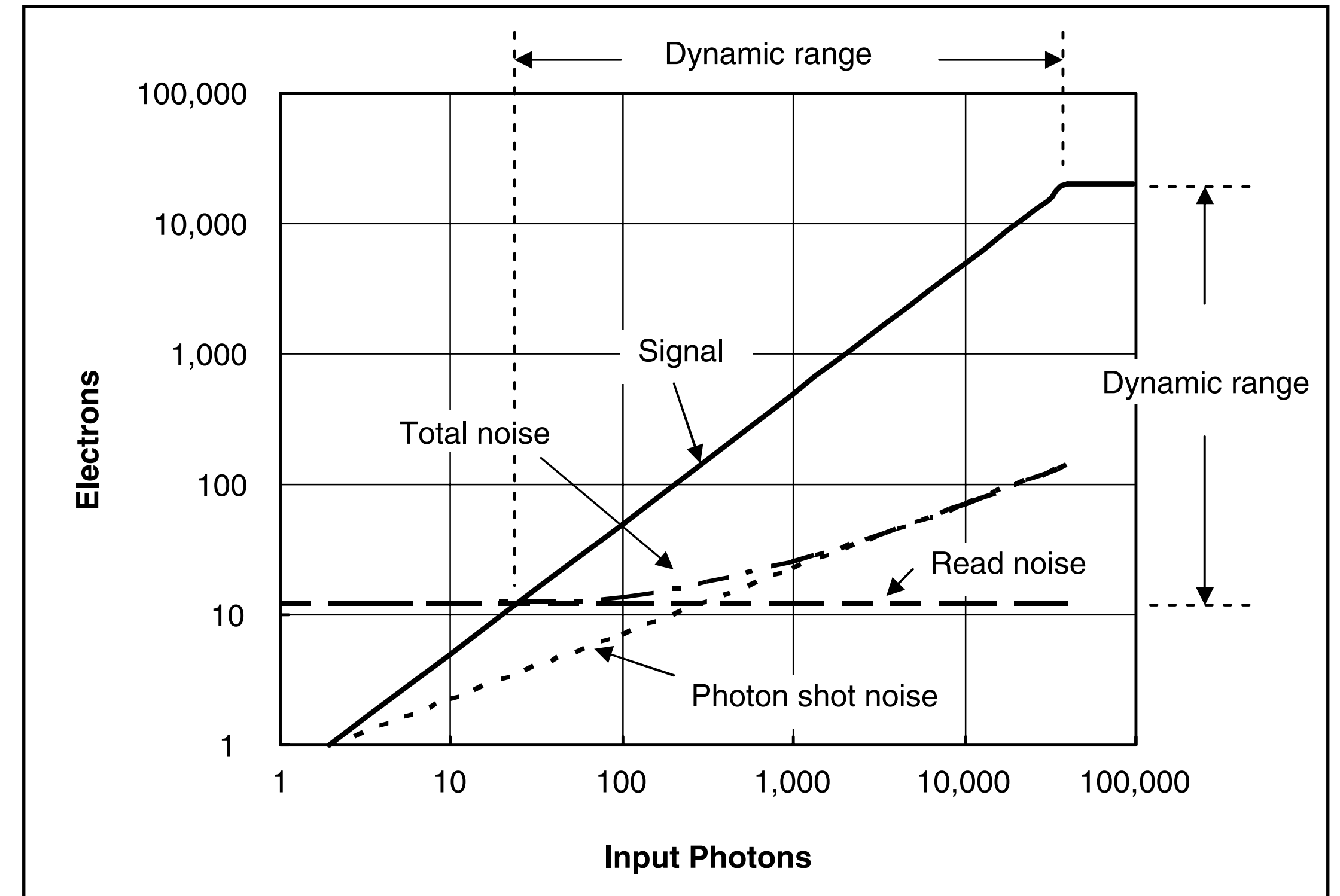
Dynamic Range and Signal to Noise Ratio

High DR doesn't mean less noisy.

DR is a single number characterizing the ability of a camera to simultaneously sense the brightest and the darkest points of a scene.

SNR varies with light levels. At a given light level, the SNR characterizes "how noisy" the image it.

- Higher light level leads to higher SNR.



Consider only photon shot noise and read noise and ignore dark-current noise here.

DR and SNR Units

Often DR and SNR are expressed in decibel or "stop".

$$SNR_{dB} = 20 \log_{10}(SNR)$$

$$DR_{dB} = 20 \log_{10}(DR)$$

$$SNR_{stop} = \log_2(SNR)$$

$$DR_{stop} = \log_2(DR)$$

Factor (power)	Decibels	Stops
1	0	0
2	3.01	1
3.16	5	1.66
4	6.02	2
5	6.99	2.32
8	9.03	3
10	10	3.32
16	12.0	4
20	13.0	4.32
31.6	15	4.98
32	15.1	5
50	17.0	5.64
100	20	6.64
1000	30	9.97
1024	30.1	10
10000	40	13.3
100000	50	16.6
1000000	60	19.9
1048576	60.2	20
100000000	80	26.6
1073741824	90.3	30
10000000000	100	33.2

Typical Dynamic Range

Input "device": (maximal signal/noise floor)

- Smartphone camera: 10 stops (iPhone 11)
- DSLR: 14.8 stops (Sony A7R4)
- Human eyes: 10-14 stops (instantaneous, i.e., when the pupil size doesn't change)

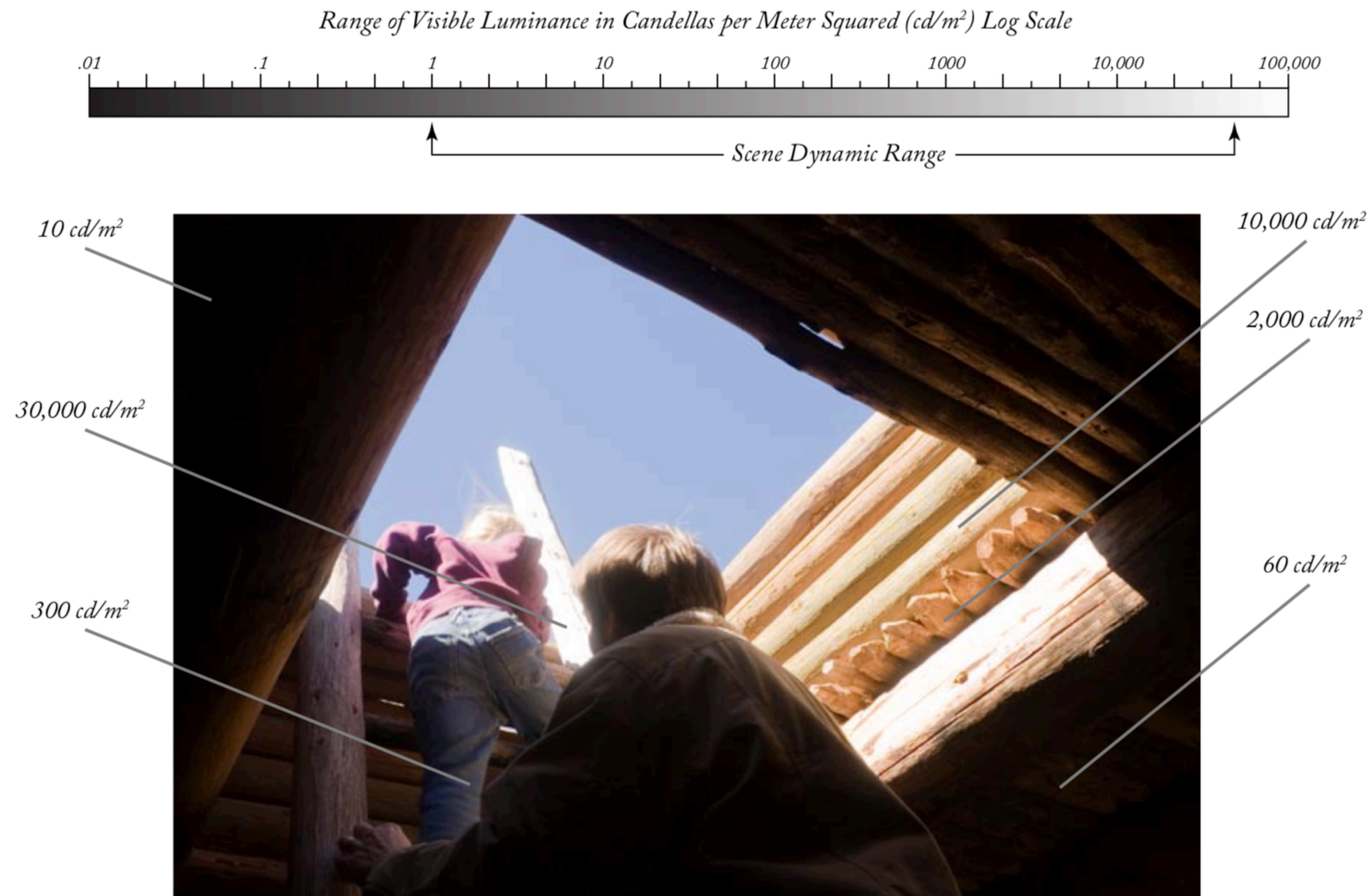
Dynamic Range of "Output Devices"

DR of light-emitting output "devices" is defined slightly differently.

DR = highest illumination/lowest illumination

- Natural scenes: ~20 stops (measurement)
- Typical Display: ~10 stops (LG 27UK850); of course every display you buy touts HDR...
- HDR Display: 14.2 stops (UHD Alliance Certified HDR LCD display)
- Paper: 6-8 stops (typical value); related to pigment density

Output "Device" Illumination



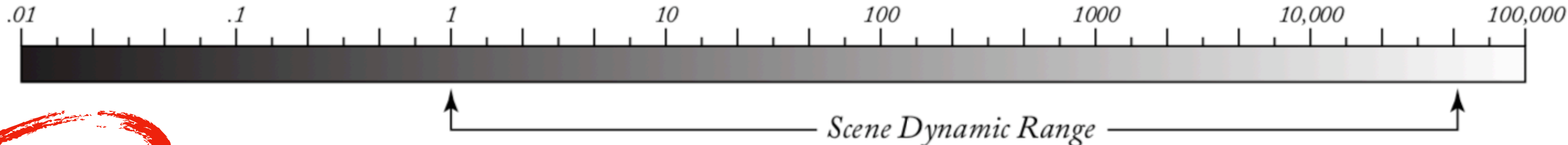
LG 27UK850-W Brightness: 350 cd/m^2

Displays will never reproduce the same luminance in a scene (power of the sun!), but tries to maintain the **relative ratio**.

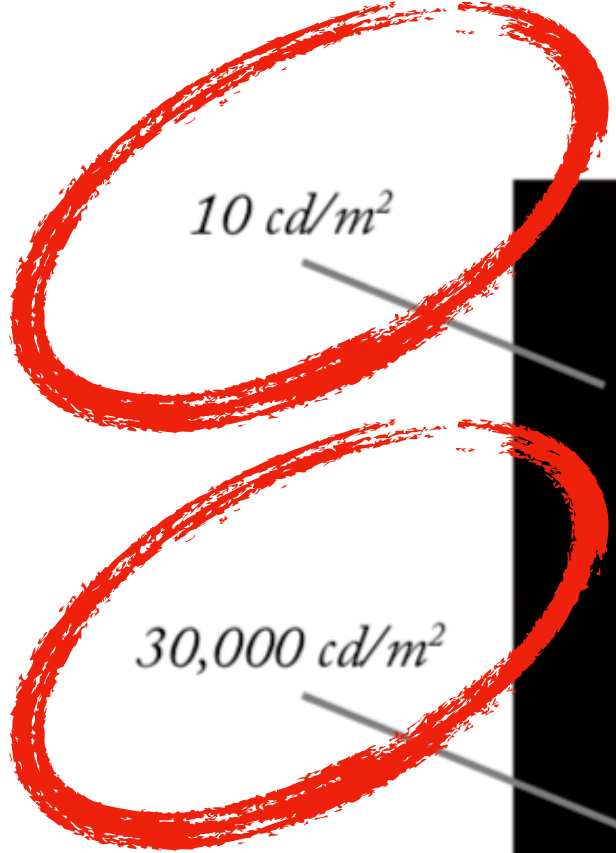
Natural Scenes Have High Dynamic Range

Scene

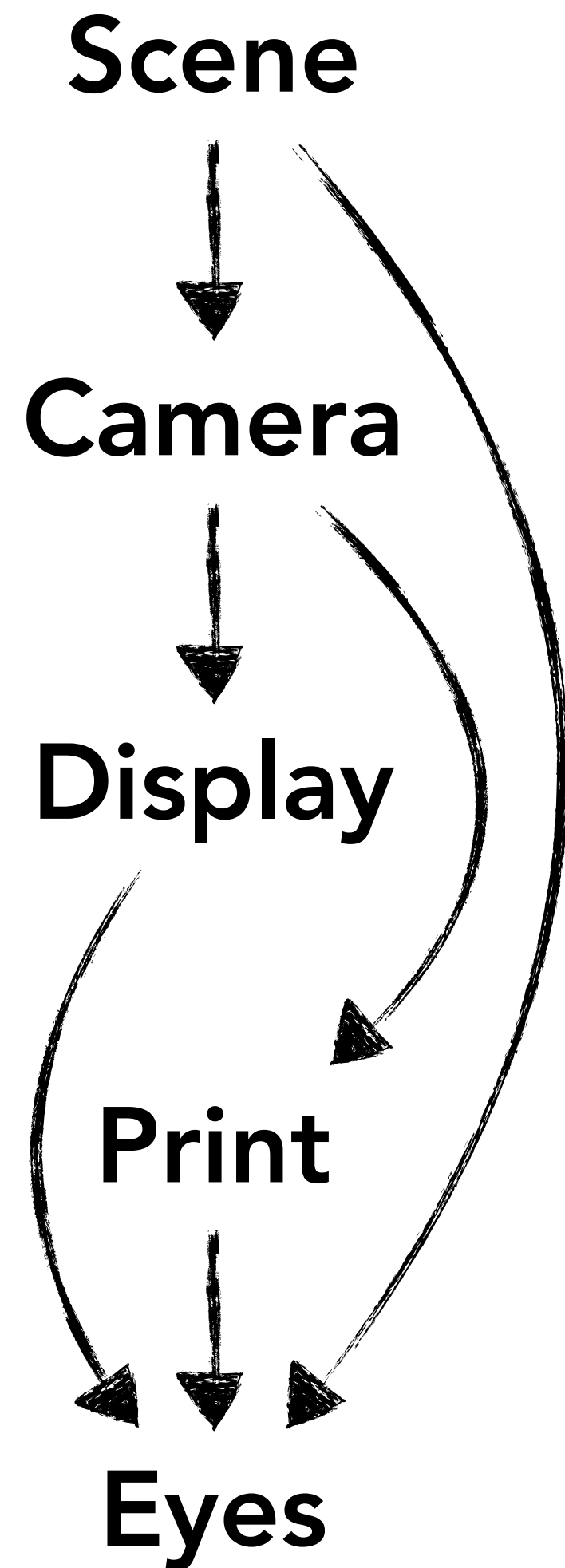
Range of Visible Luminance in Candellas per Meter Squared (cd/m^2) Log Scale



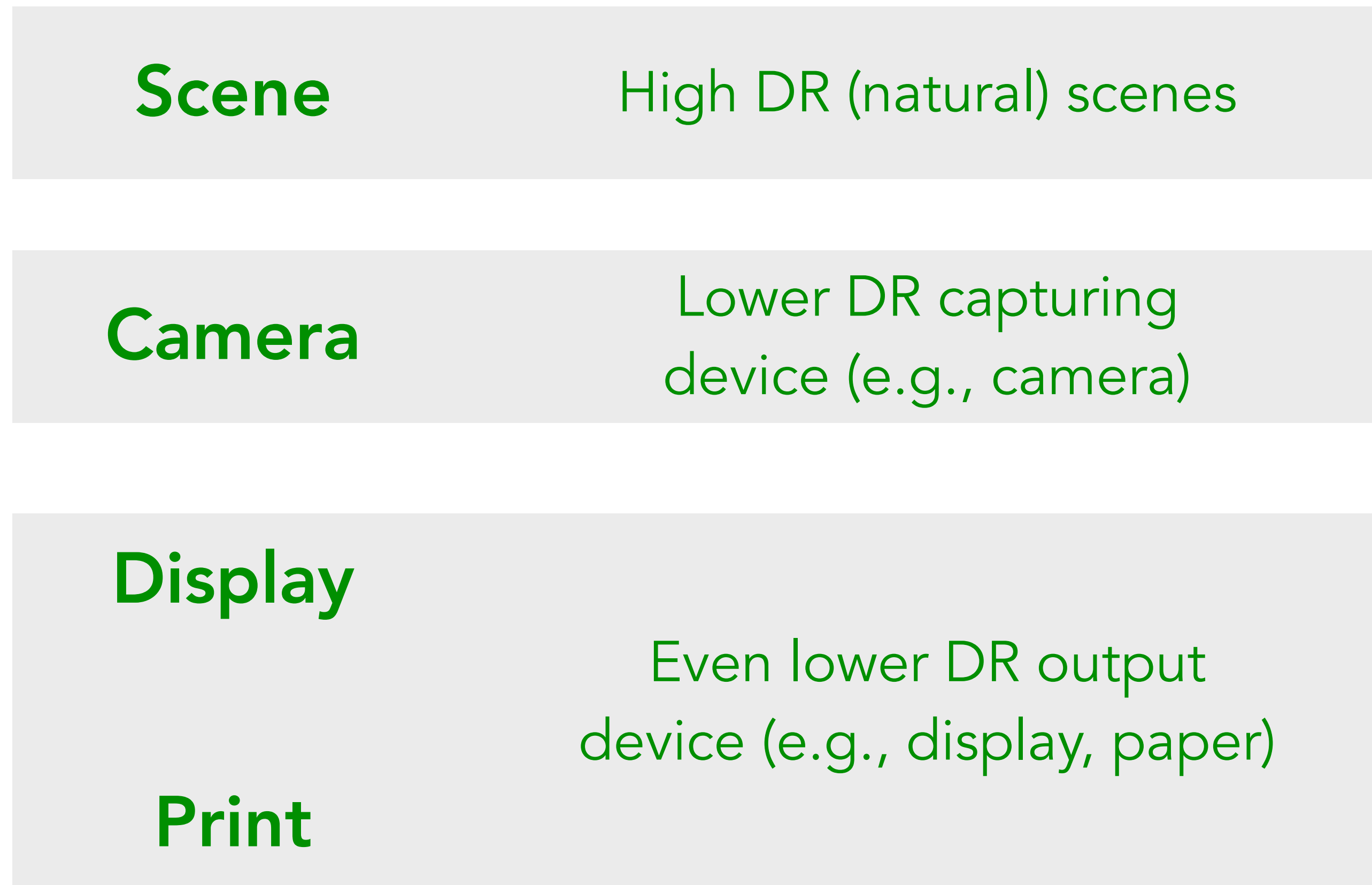
Eyes



The Problem



Two Related Tasks in Sensor Signal Processing



- 1. HDR Imaging:** how to capture an HDR scene with a lower DR capturing device. Limited by the camera.
- 2. Tone Mapping:** how to display HDR images on lower DR display devices. Limited by the display medium.