Information Theory for Single Molecule Observation

Francisco Balzarotti

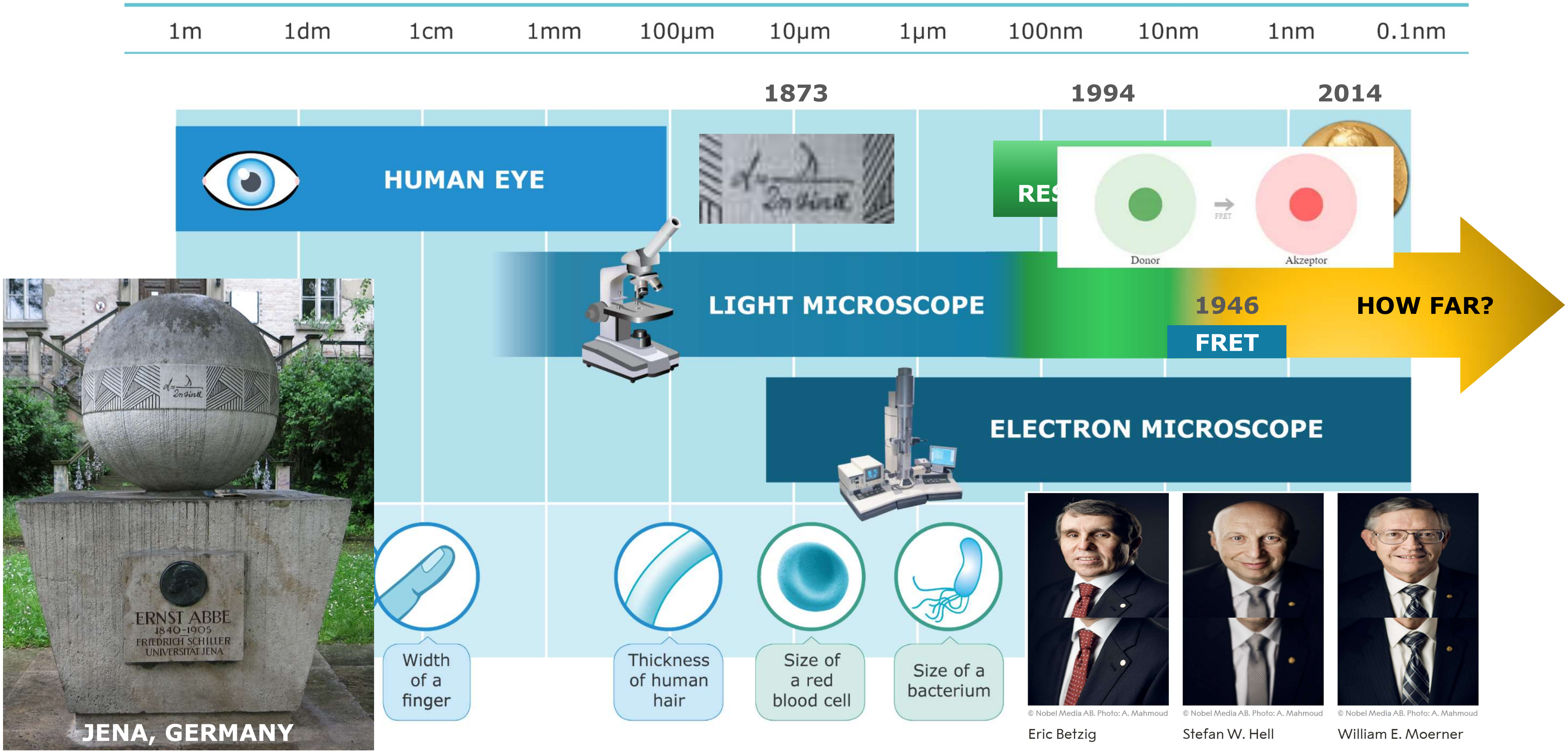
Likelihood functions, Parameter estimation, Fisher Information

12.07.2022

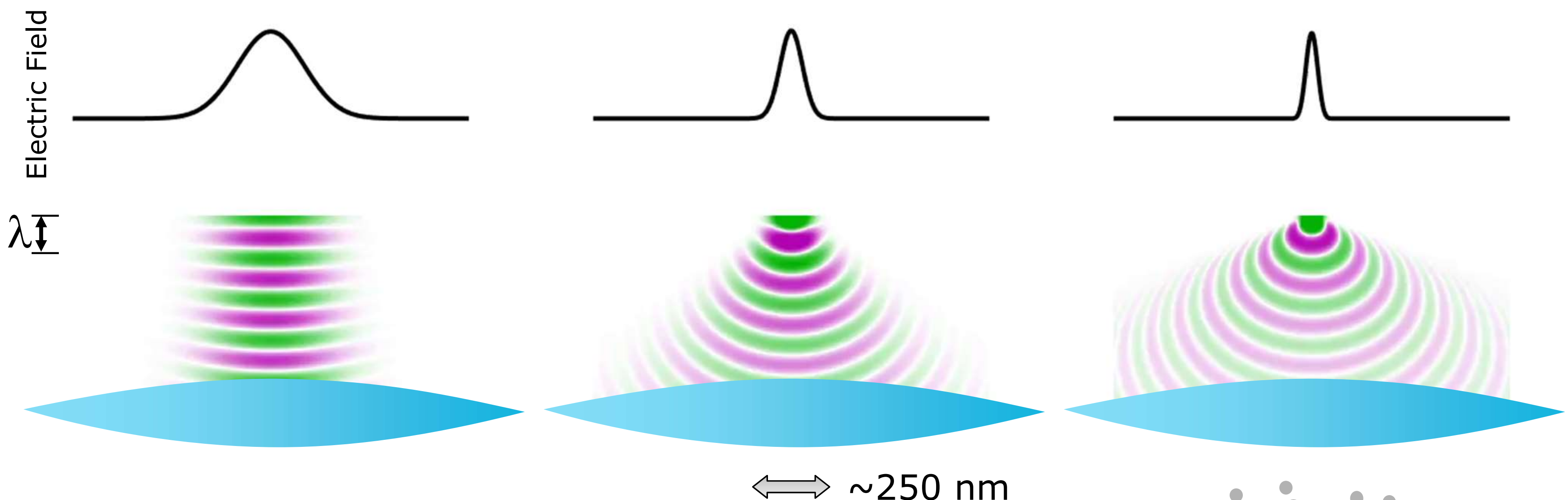
IMP Research Institute of Molecular Pathology

Part of Vienna BioCenter

LIGHT MICROSCOPY



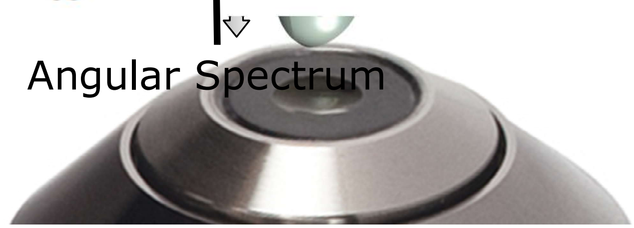
DIFFRACTION BARRIER

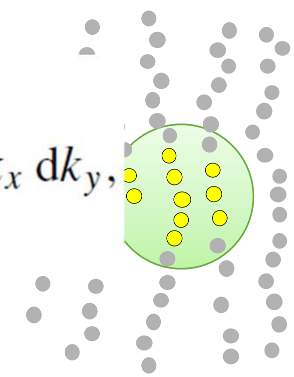


$$\mathbf{E}(x, y, z) = \iint_{-\infty}^{\infty} \hat{\mathbf{E}}(k_x, k_y; 0)\, e^{i[k_x x + k_y y \pm k_z z]}\, dk_x\, dk_y,$$

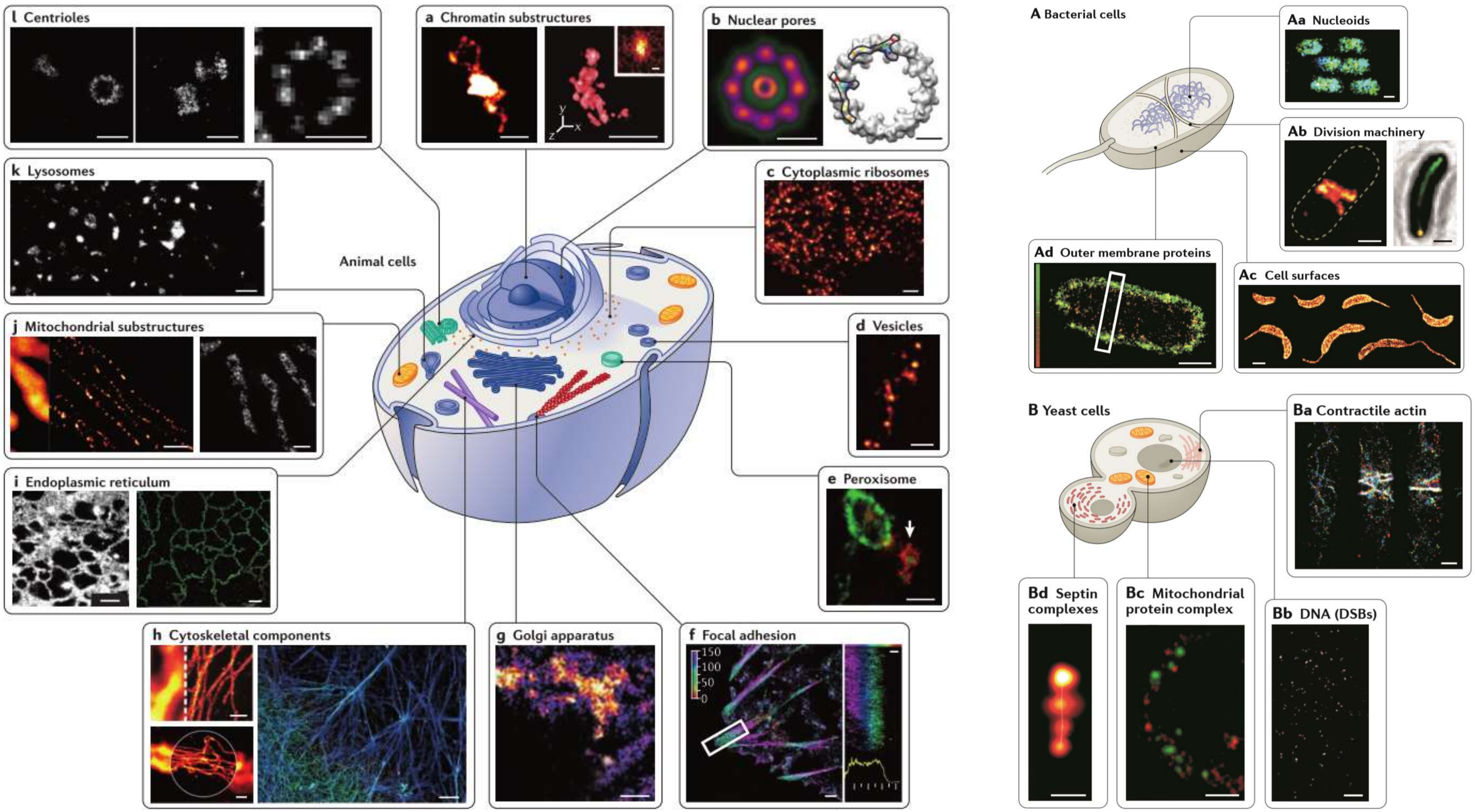
Electric field

Angular Spectrum

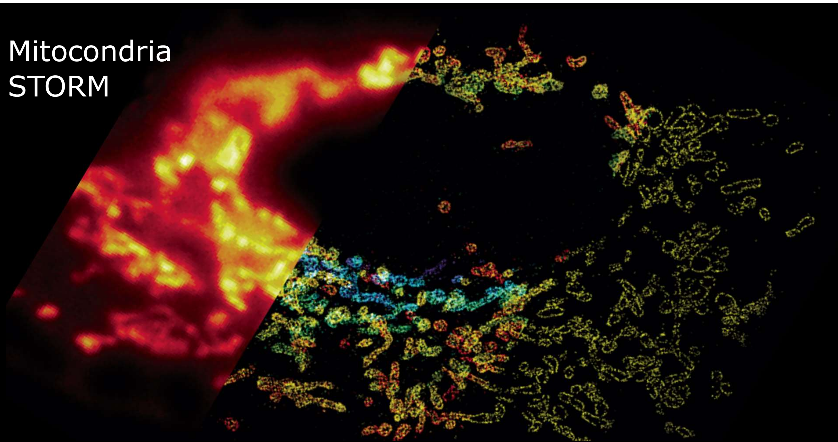




SUPER RESOLUTION IN CELL BIOLOGY

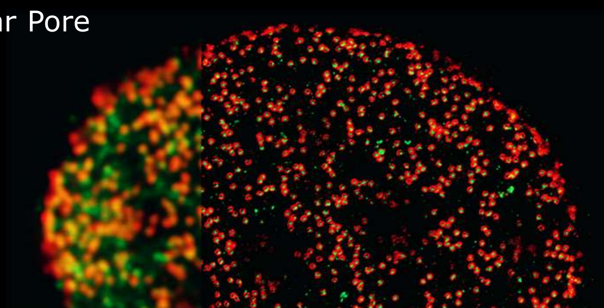


Mitocondria
STORM



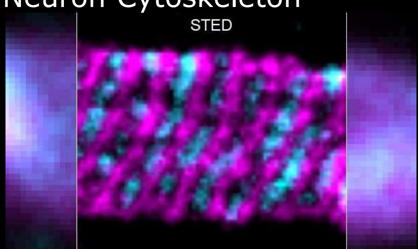
B. Huang et al. *Nat. Meth.* (2008)

Nuclear Pore
STED

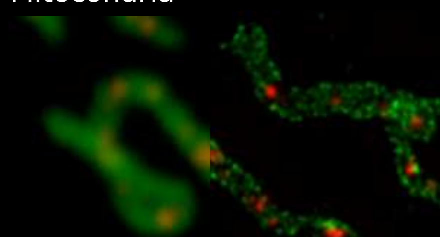


F. Göttfert *et al. Biophys. J.* (2013)

Neuron Cytoskeleton



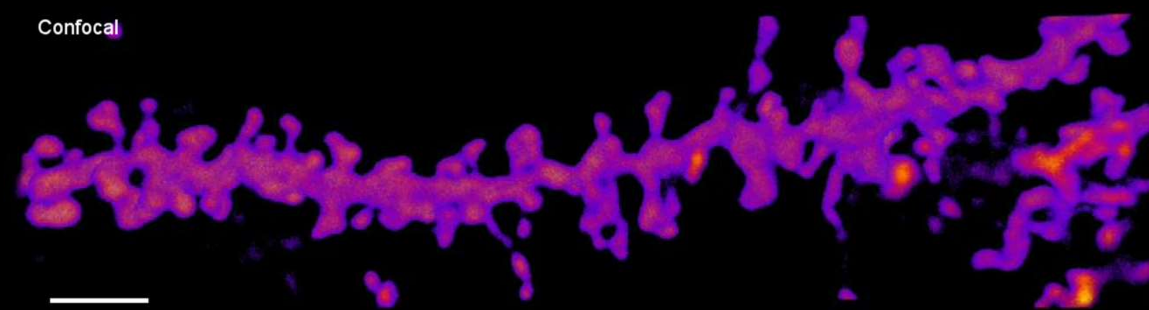
Mitocondria

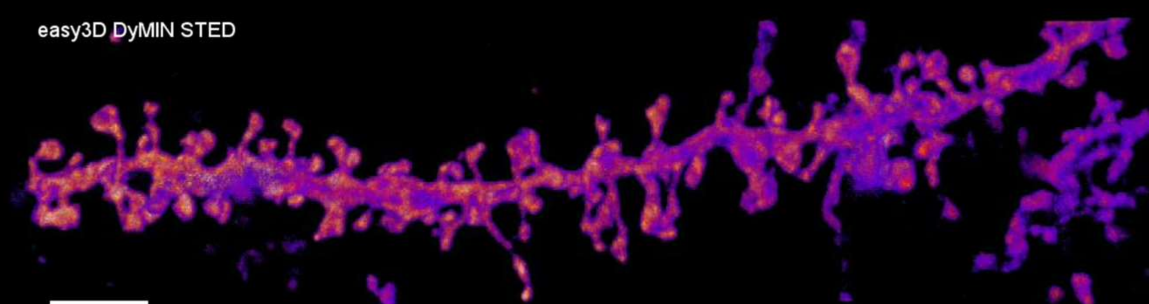


Abberior Instruments Gallery

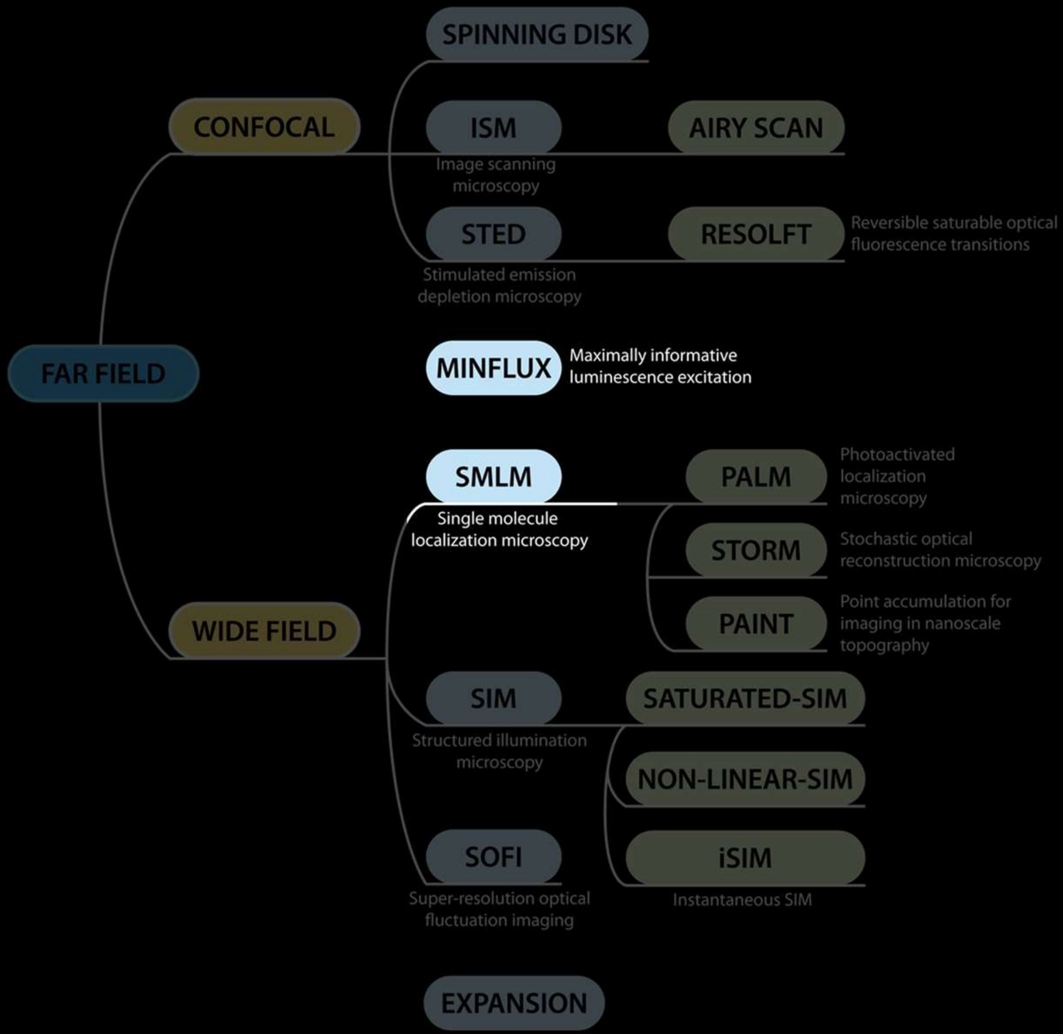
Cell cytoskeleton

Dendritic spines





SUPER RESOLUTION OPTICAL MICROSCOPY

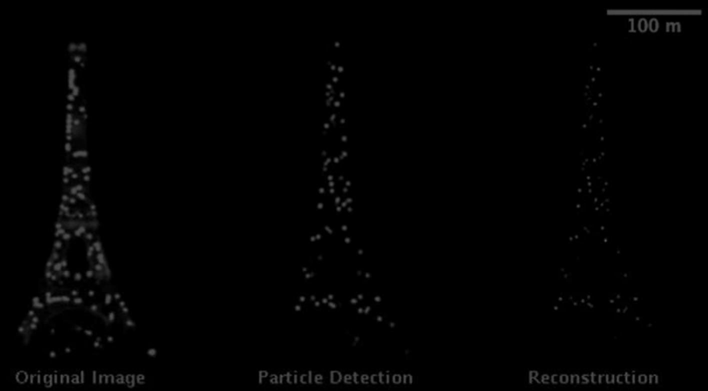


CONSIDERATIONS

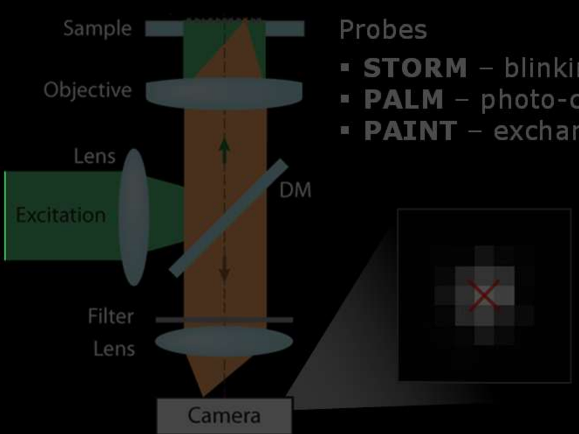
- Spatial resolution
- Time resolution
- Photo toxicity
- Long observation
- Probe availability
- Multicolor
- Live compatible
- Depth / tissue

SINGLE MOLECULE LOCALIZATION

SMLM – Super Resolution Microscopy



R. Henriques. *YouTube.com/paxcal* (2010)



Probes
- **STORM** – blinking dyes
- **PALM** – photo-convertible FPs
- **PAINT** – exchangeable reporter

Wide field precision

$$\sigma \propto \frac{\lambda}{NA}\frac{1}{\sqrt{N}}$$

Wavelength

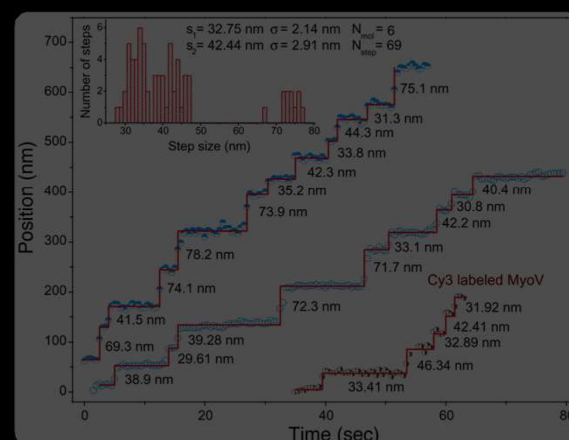
Numerical aperture

Photons

MINFLUX precision

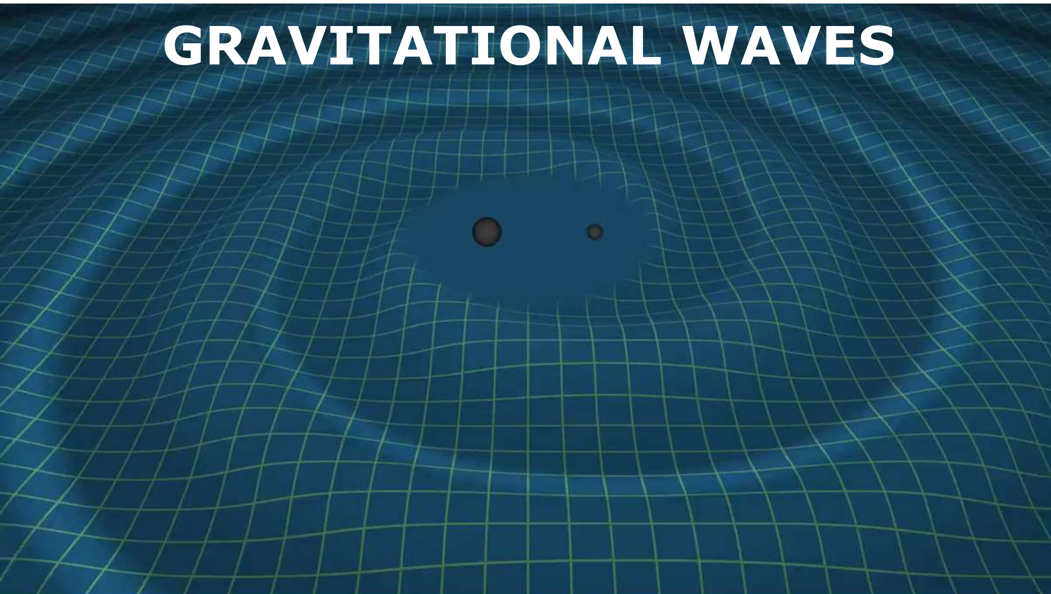
$$\sigma \propto \frac{L}{N^{K/2}}$$

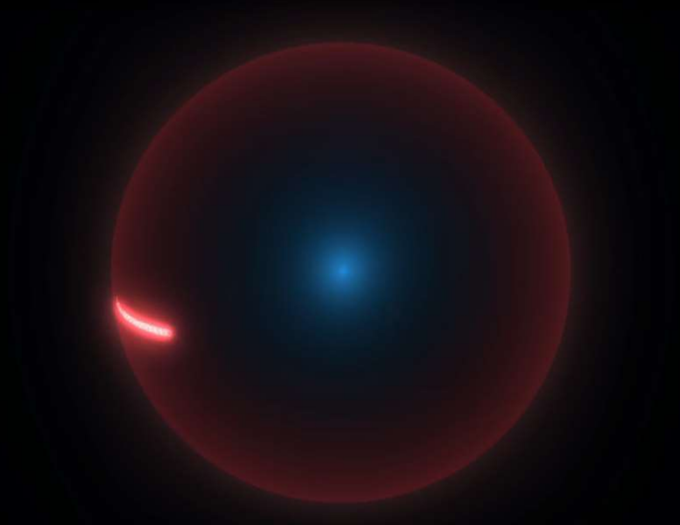
SMT – Single Molecule Tracking



A. Yildiz et al. *Science* (2003)

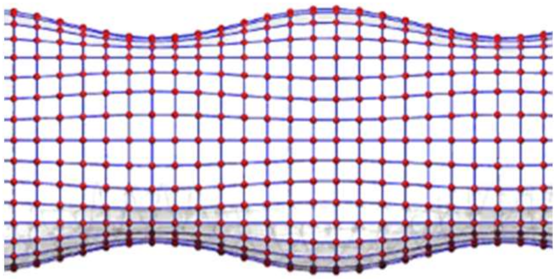
M.J. Rust et al. *Nat. Methods* (2006)
E. Betzig et al. *Science* (2006)
S.T. Hess et al. *Biophys. J.* (2006)

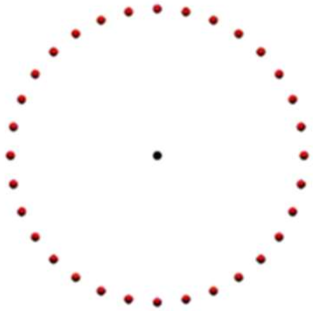




WHAT IS THE RULER?

Gravitational wave with linear polarization



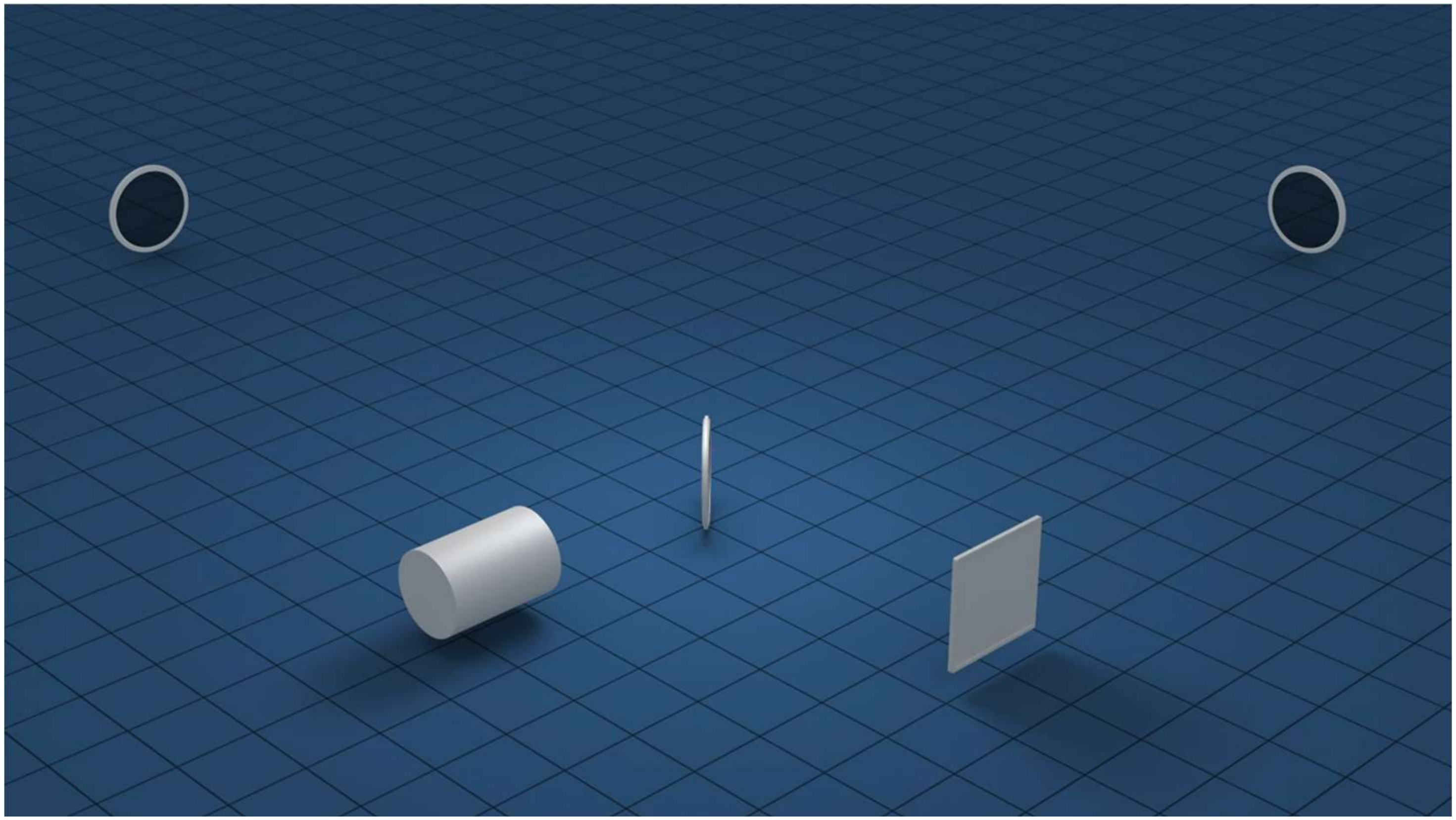


Michaelson Interferometer

$$E_T = E_1 e^{i(\phi_1 - \omega t)} + E_2 e^{i(\phi_2 - \omega t)}$$

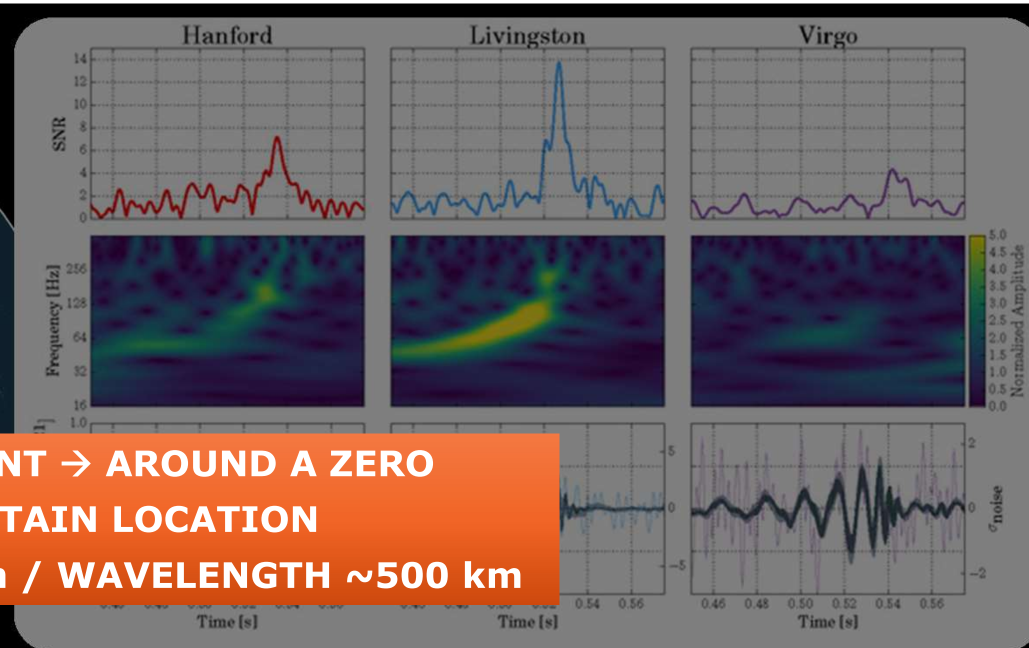
$$I_{det} = E_0^2\left[1 - \cos\left(2k(l_1 - l_2)\right)\right]$$

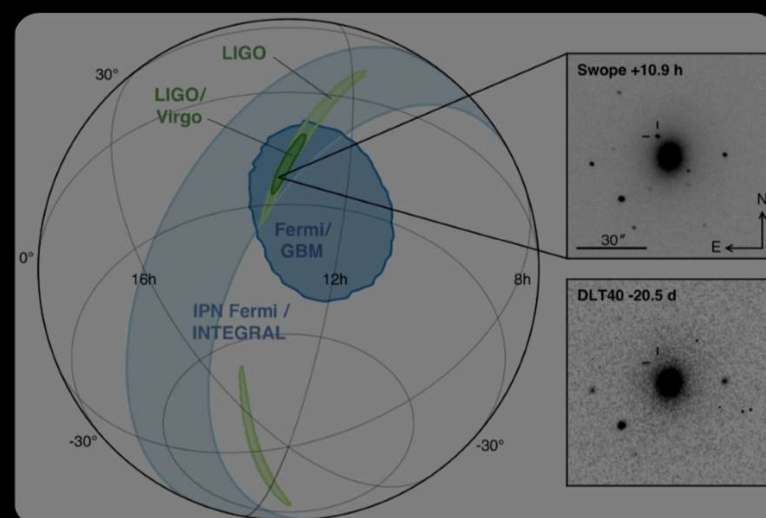
$$\boxed{I_{det} \propto (l_1 - l_2)^2}$$



www.ligo.caltech.edu

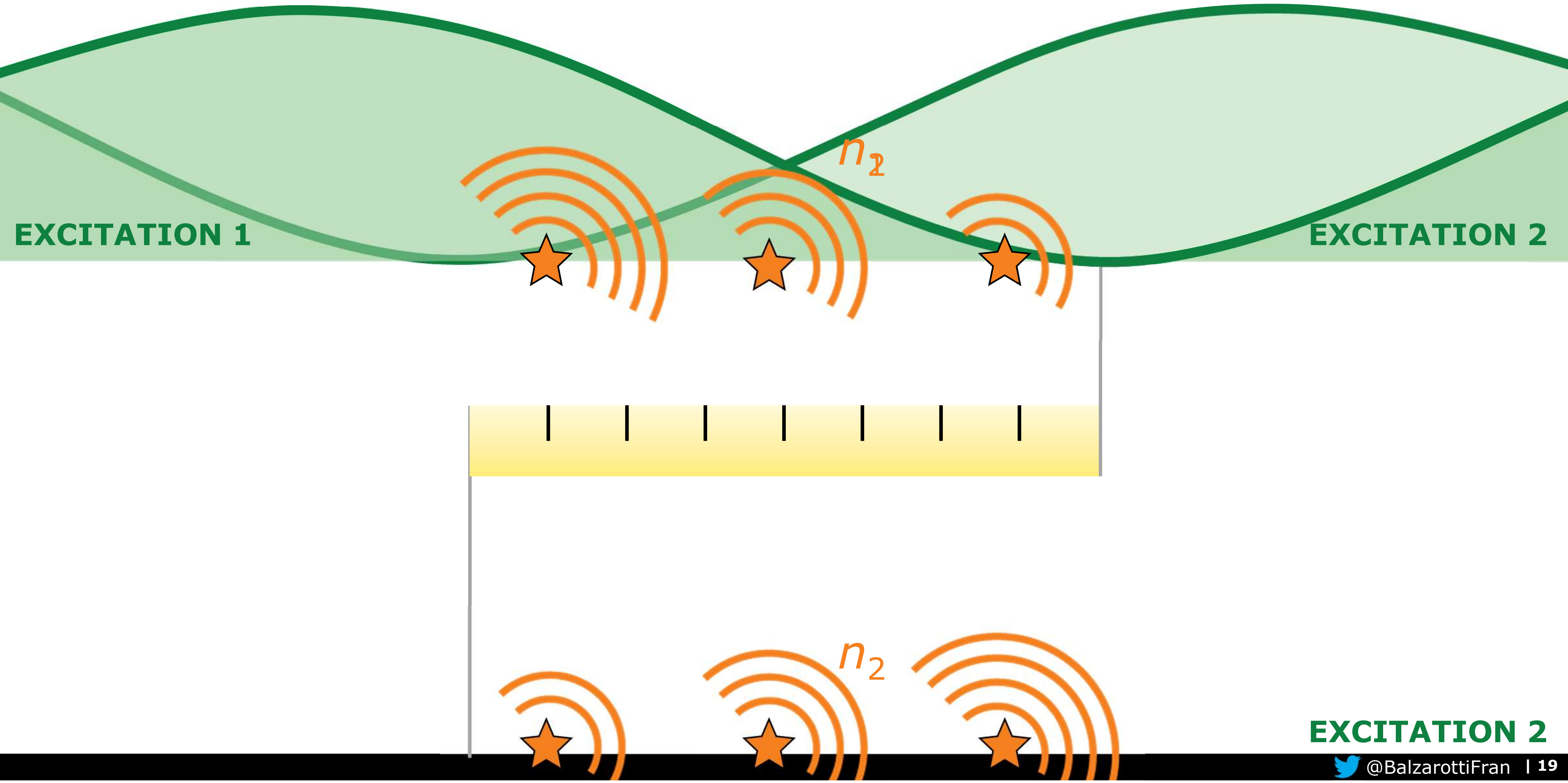
- **SENSITIVE MEASUREMENT → AROUND A ZERO**
- **MULTIPLE POINTS → OBTAIN LOCATION**
- **DISPLACEMENTS 10^{-18} m / WAVELENGTH ~500 km**





www.ligo.caltech.edu

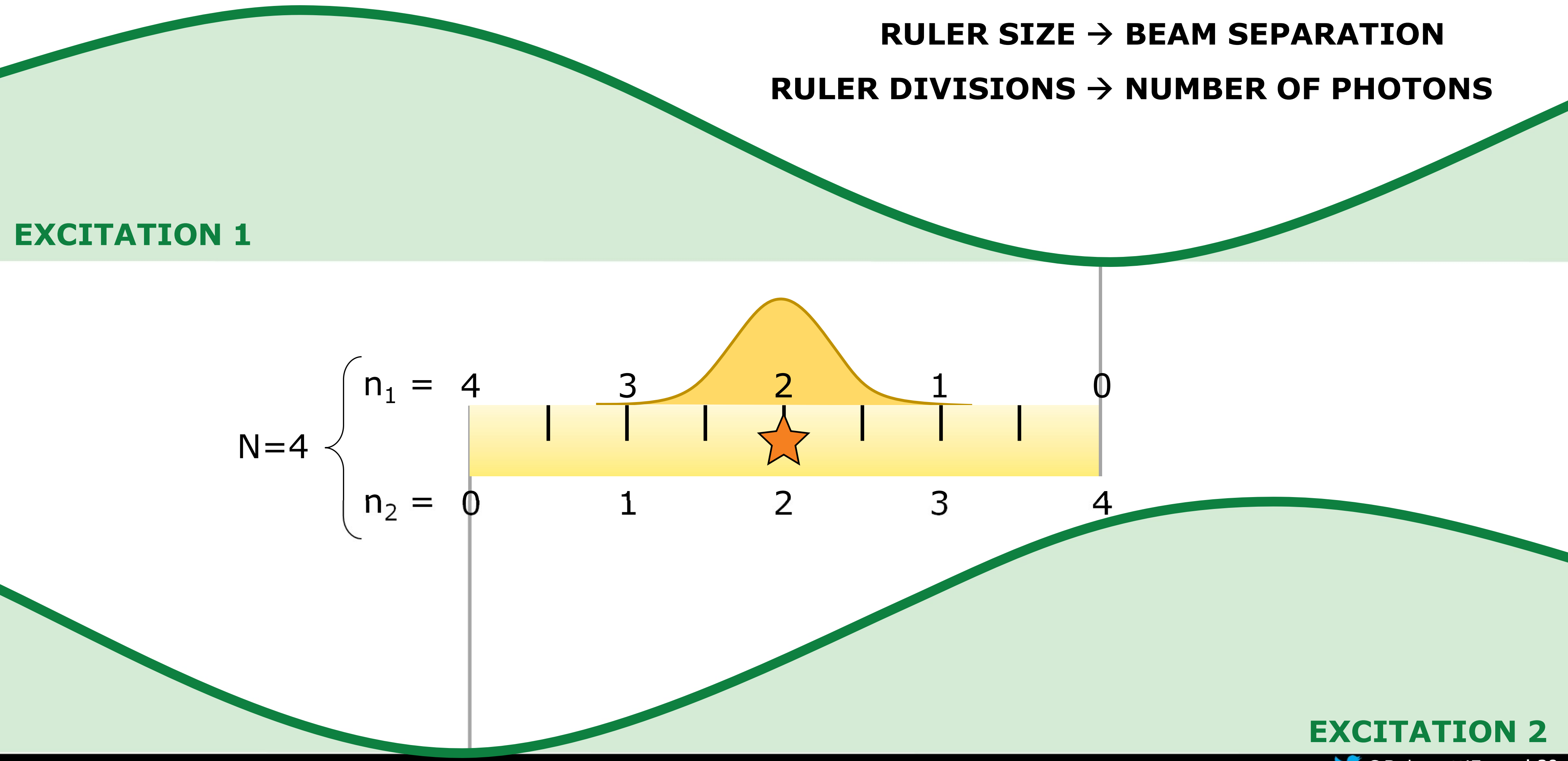
MINFLUX AS A RULER



MINFLUX AS A RULER

RULER SIZE → BEAM SEPARATION

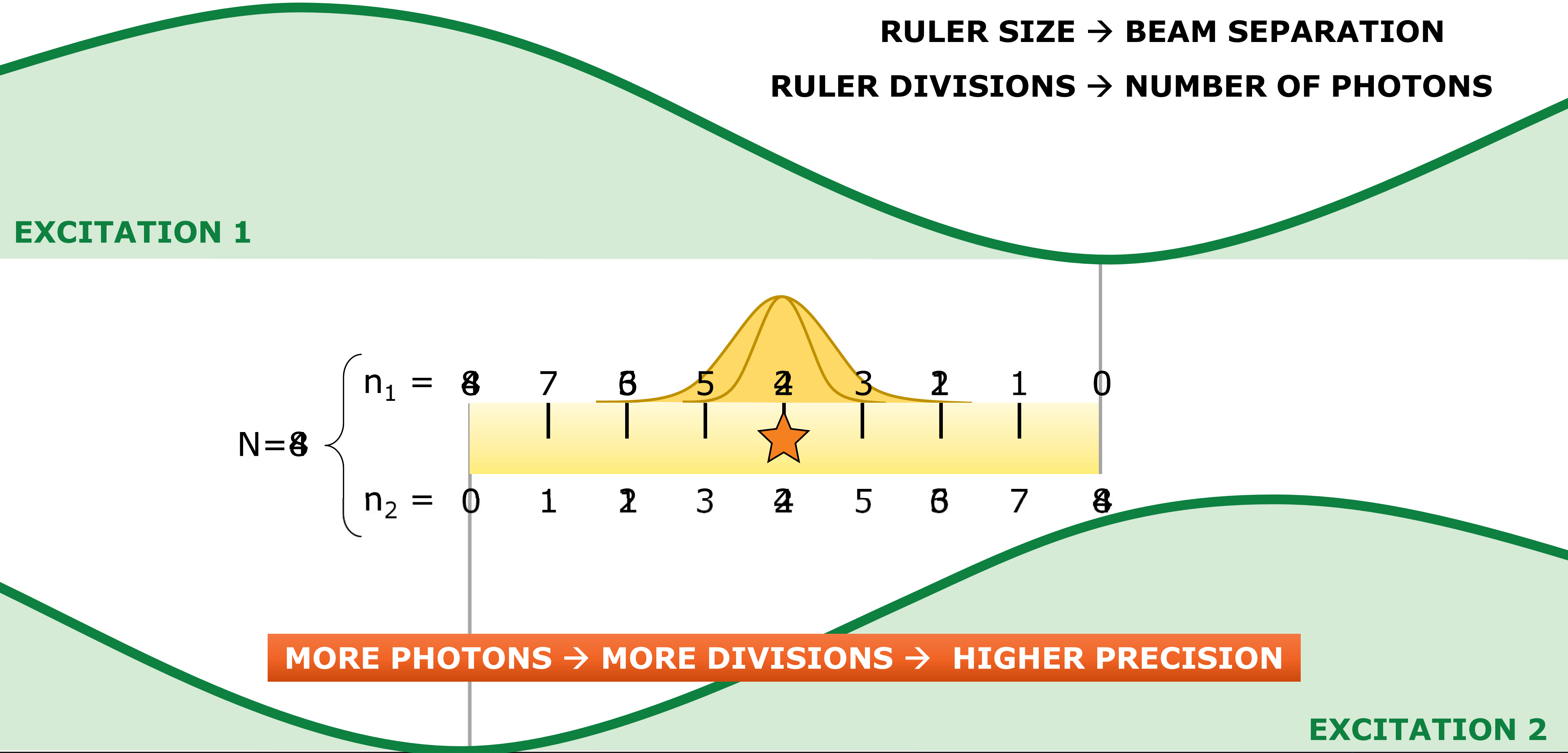
RULER DIVISIONS → NUMBER OF PHOTONS



MINFLUX AS A RULER

RULER SIZE → BEAM SEPARATION

RULER DIVISIONS → NUMBER OF PHOTONS

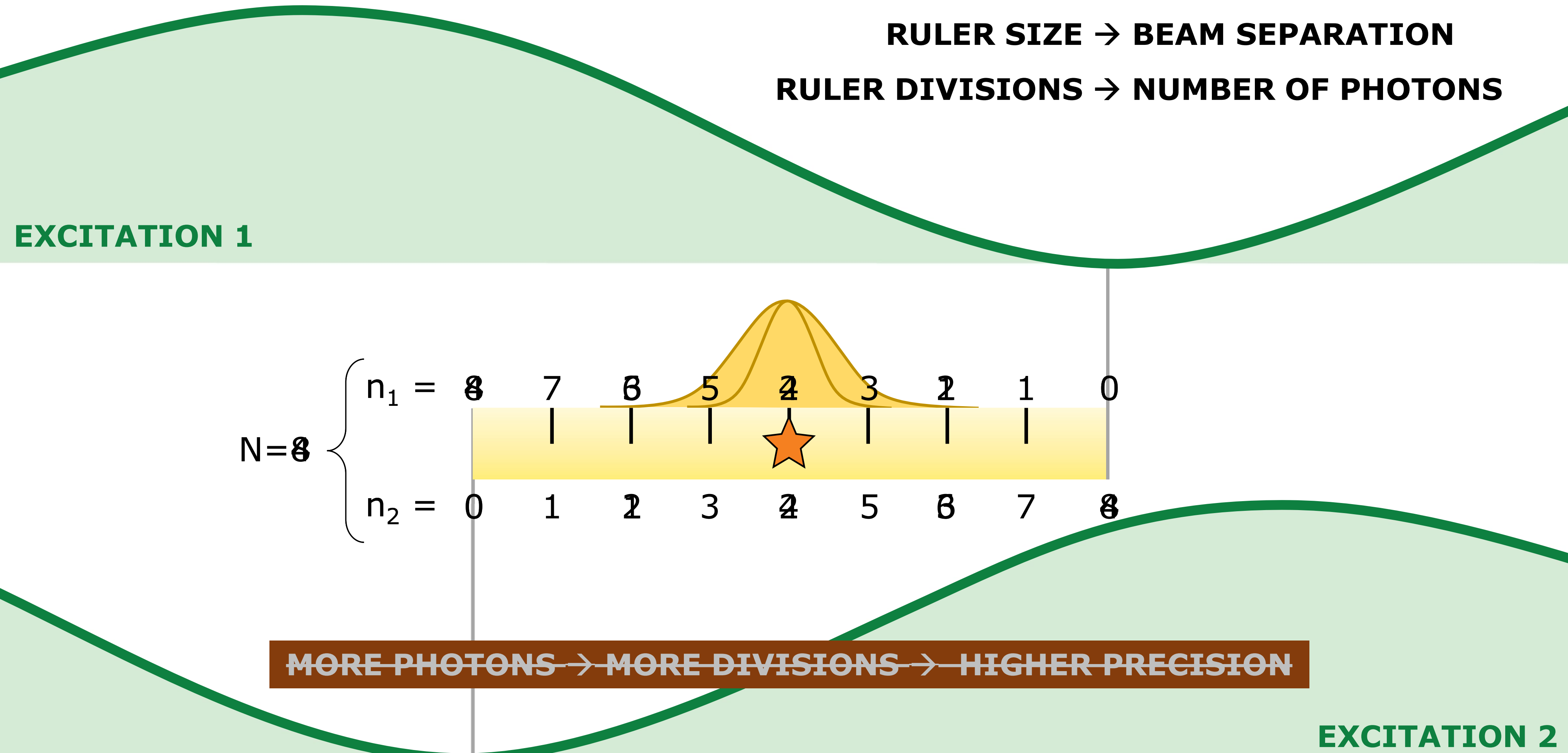


MORE PHOTONS → MORE DIVISIONS → HIGHER PRECISION

MINFLUX AS A RULER

RULER SIZE → BEAM SEPARATION

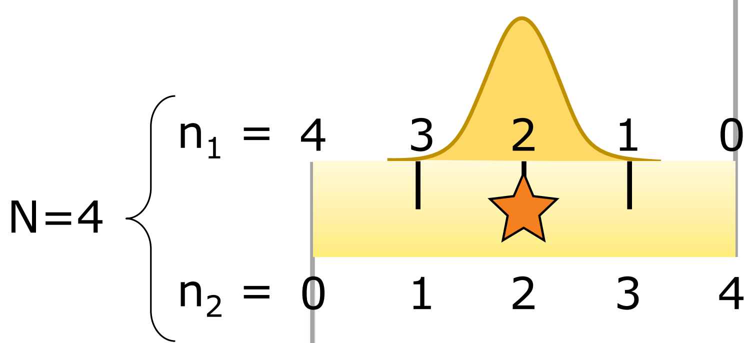
RULER DIVISIONS → NUMBER OF PHOTONS



MINFLUX AS A RULER

RULER SIZE → BEAM SEPARATION

RULER DIVISIONS → NUMBER OF PHOTONS



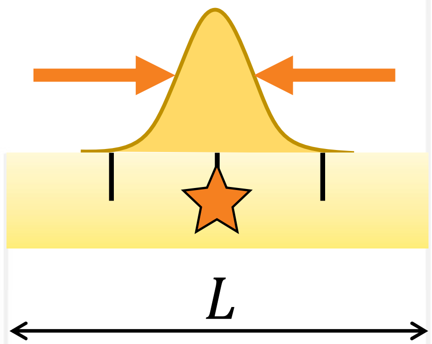
~~**MORE PHOTONS → MORE DIVISIONS → HIGHER PRECISION**~~

BEAMS CLOSER → SMALLER RULER → HIGHER PRECISION

MINFLUX AS MAXIMALLY INFORMATIVE

MINFLUX Resolution

$$\sigma \propto \frac{L}{\sqrt{N}}$$



MAXIMALLY
INFORMATIVE
LUMINESCENCE
EXCITATION

INFORMATION THEORY ELEMENTS

Likelihood Function

Parameter Estimation

Fisher Information

BINOMIAL DISTRIBUTION

Flip a coin n times each with success probablity p

Probability mass function:

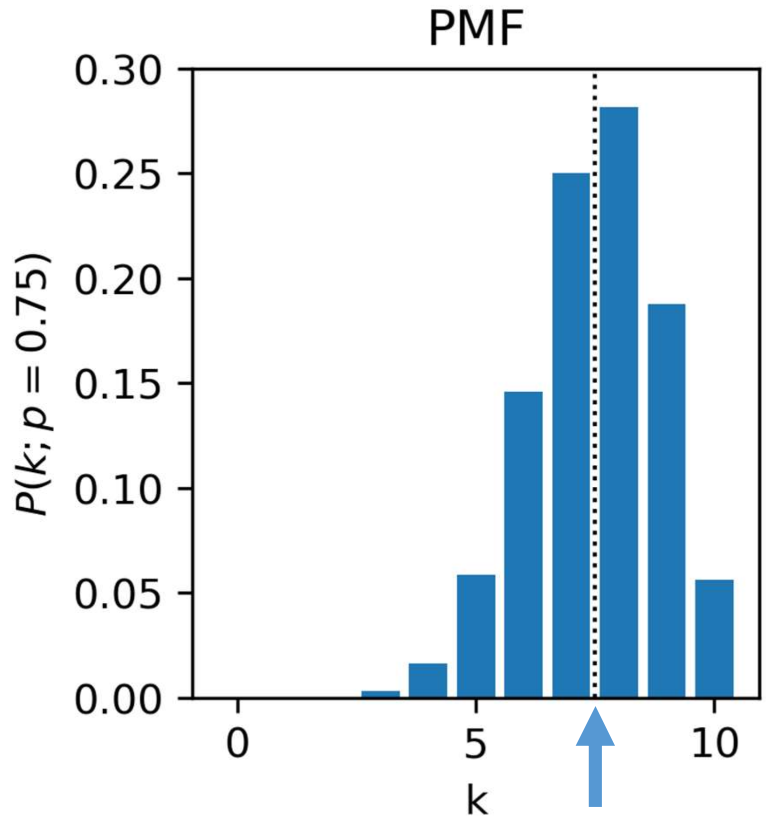
$$P_X(X = k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

Measure $k = 6$
→ How to obtain parameter p?

Notation abuses:

$P_X(X = k; n, p)$
$P(k; n, p)$
$P(k|n, p)$
$P(k)$

$n = 10;\ p = 0.75$



$np = 7.5$

LIKELIHOOD FUNCTION

Probability mass function:

$$P_X(X = k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Measure $k = 6$
→ How to obtain parameter p?

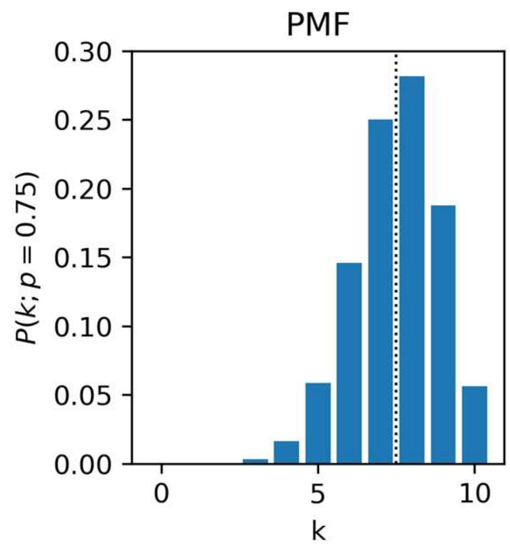
Likelihood function:

$$\mathcal{L}(k, p) = P_X(X = k; n, p)$$

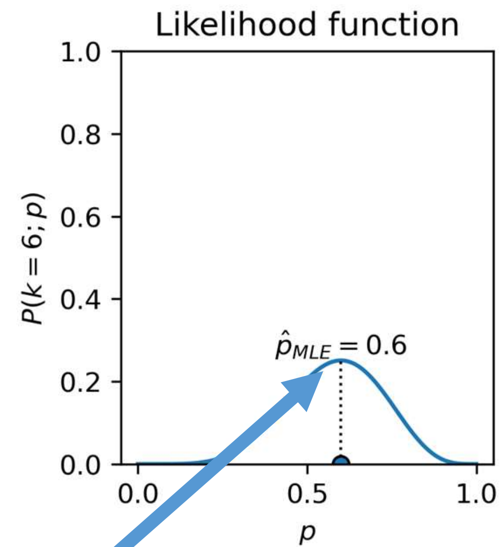
Which p makes this measurement the most likely outcome?

$$\frac{d\mathcal{L}(k,p)}{dp} = 0 \qquad \hat{p}_{MLE} = \frac{k}{n}$$

$n = 10;\ p = 0.75$

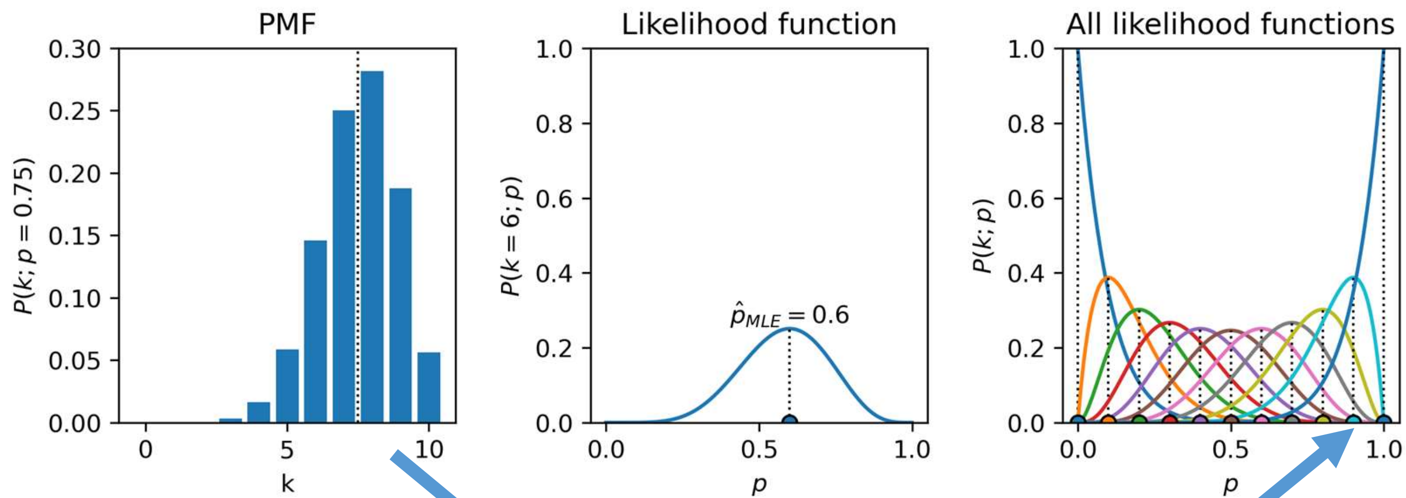


$n = 10; k = 6$



- Not dependent on the real value of p
- Only on the measurement k

LIKELIHOOD FUNCTION



Probability mass function:

$$P_X(X = k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

Likelihood function:

$$\mathcal{L}(k, p) = P_X(X = k; n, p)$$

Estimator:

$$\hat{p}_{MLE} = \frac{k}{n}$$

$$k \rightarrow \mathcal{L}(k, p) \rightarrow \hat{p}_{MLE}$$

Random variable

Random variable

MAXIMUM LIKELIHOOD ESTIMATOR

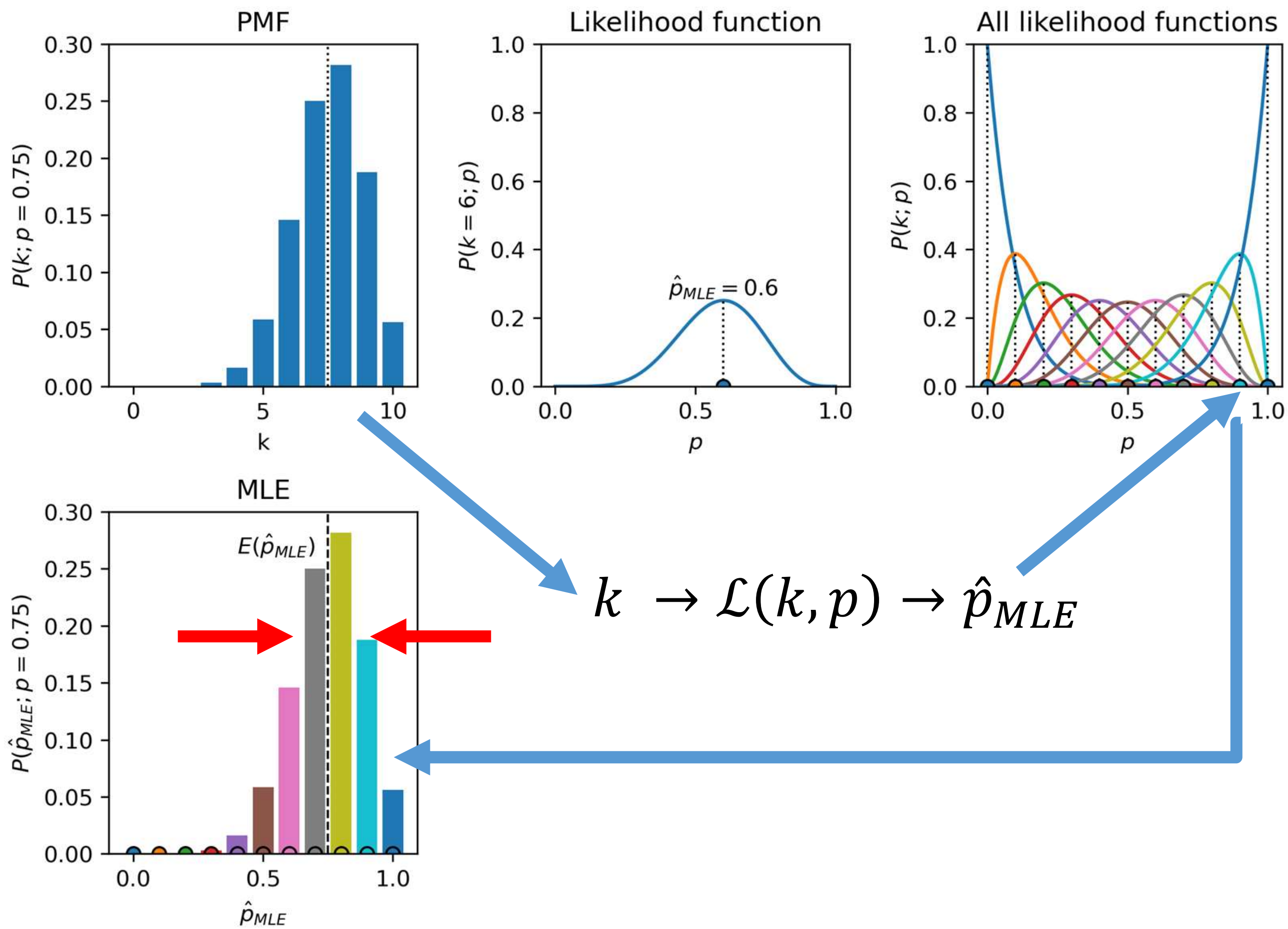
Estimator:

$$\hat{p}_{MLE} = \frac{k}{n}$$

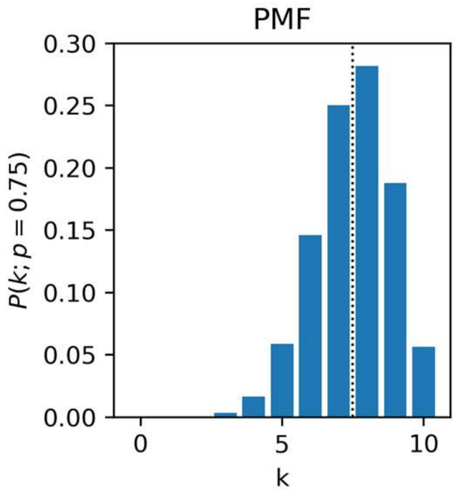
Characterize:

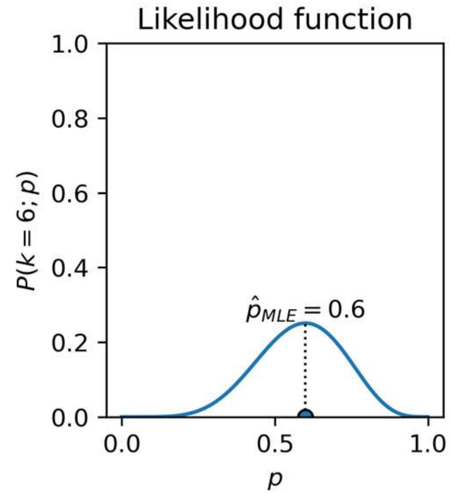
$E(\hat{p}_{MLE})$

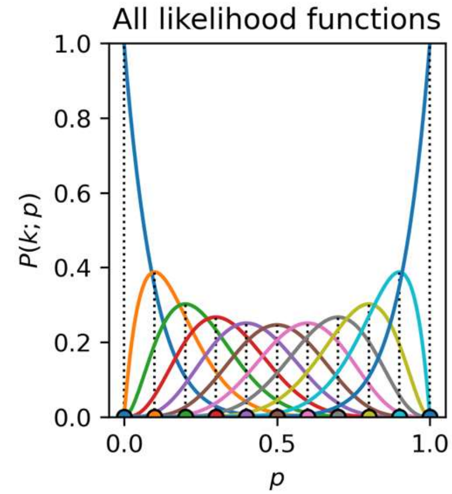
$\mathrm{var}(\hat{p}_{MLE})$



FISHER INFORMATION AND CRAMER-RAO BOUND

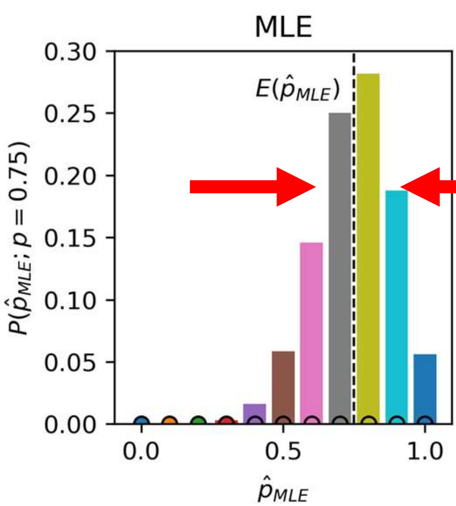






Estimator:

$$\hat{p}_{MLE} = \frac{k}{n}$$



Characterize:

$E(\hat{p}_{MLE})$

$\mathrm{var}(\hat{p}_{MLE})$

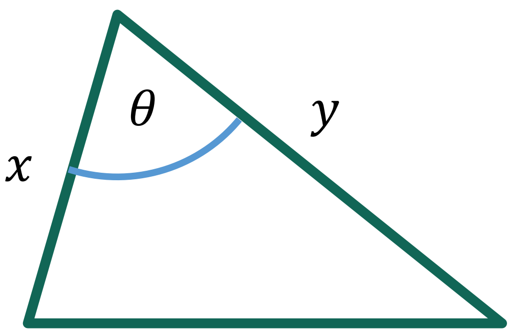
How sharp can this be?

$$\mathrm{var}(\hat{p}) \geq \frac{1}{F_p}$$

Cramer-Rao bound

Fisher Information

FISHER INFORMATION AND CRAMER-RAO BOUND



$$x^T y = |x||y|\cos(\theta) \leq |x||y|$$

Cauchy-Schwarz: $\langle x, y\rangle^2 \leq |x|^2|y|^2$

Fisher Information

$$\boxed{\frac{1}{E(y^2)} \leq \sigma_p^2}$$

Estimator $x = \hat{p}(k)$

Internal product

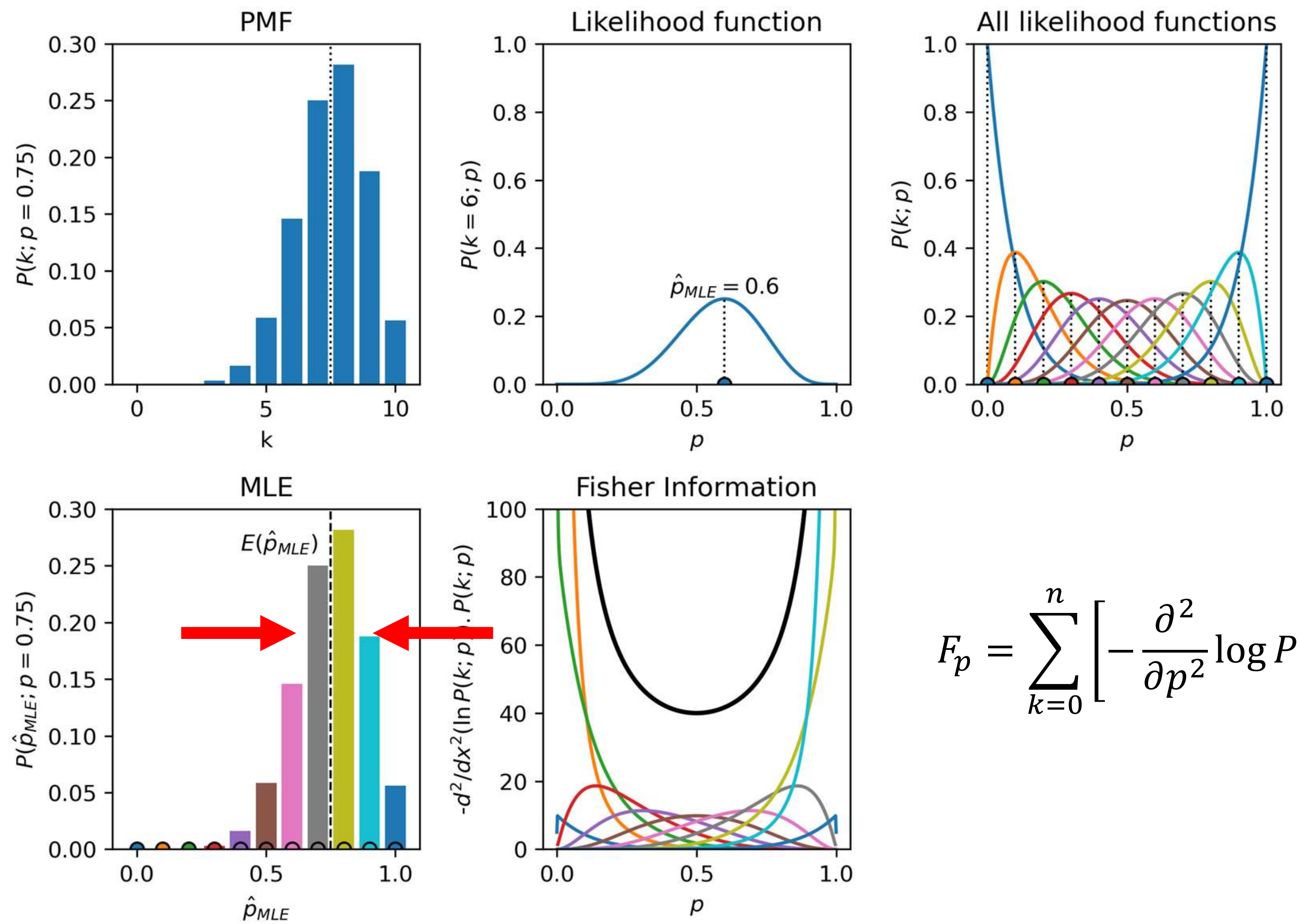
$$\langle x, y\rangle = \mathrm{cov}(x, y)$$
$$\langle x, y\rangle = \mathrm{E}([x - E(x)][y - E(y)])$$
$$\langle x, x\rangle = \mathrm{var}(\hat{p}\,)$$

Score function $y = \frac{\partial}{\partial p}\log P(k;p)$

$$F_p = E(y^2) = E\left(\left(\frac{\partial}{\partial p}\log P(k;p)\right)^2\right) = E\left(-\frac{\partial^2}{\partial p^2}\log P(k;p)\right)$$

Actual calculation → $F_p = \begin{cases} \sum_{k=0}^{\infty}\left[-\frac{\partial^2}{\partial p^2}\log P(k;p)\right]P(k;p) \\ \int\left[-\frac{\partial^2}{\partial p^2}\log p(k;p)\right]p(k;p)dk \end{cases}$

FISHER INFORMATION AND CRAMER-RAO BOUND



Estimator:

$$\hat{p}_{MLE} = \frac{k}{n}$$

Characterize:

$E(\hat{p}_{MLE})$

$\mathrm{var}(\hat{p}_{MLE})$

$$F_p = \sum_{k=0}^{n} \left[-\frac{\partial^2}{\partial p^2} \log P(k; p) \right] P(k; p)$$

FISHER INFORMATION AND CRAMER-RAO BOUND

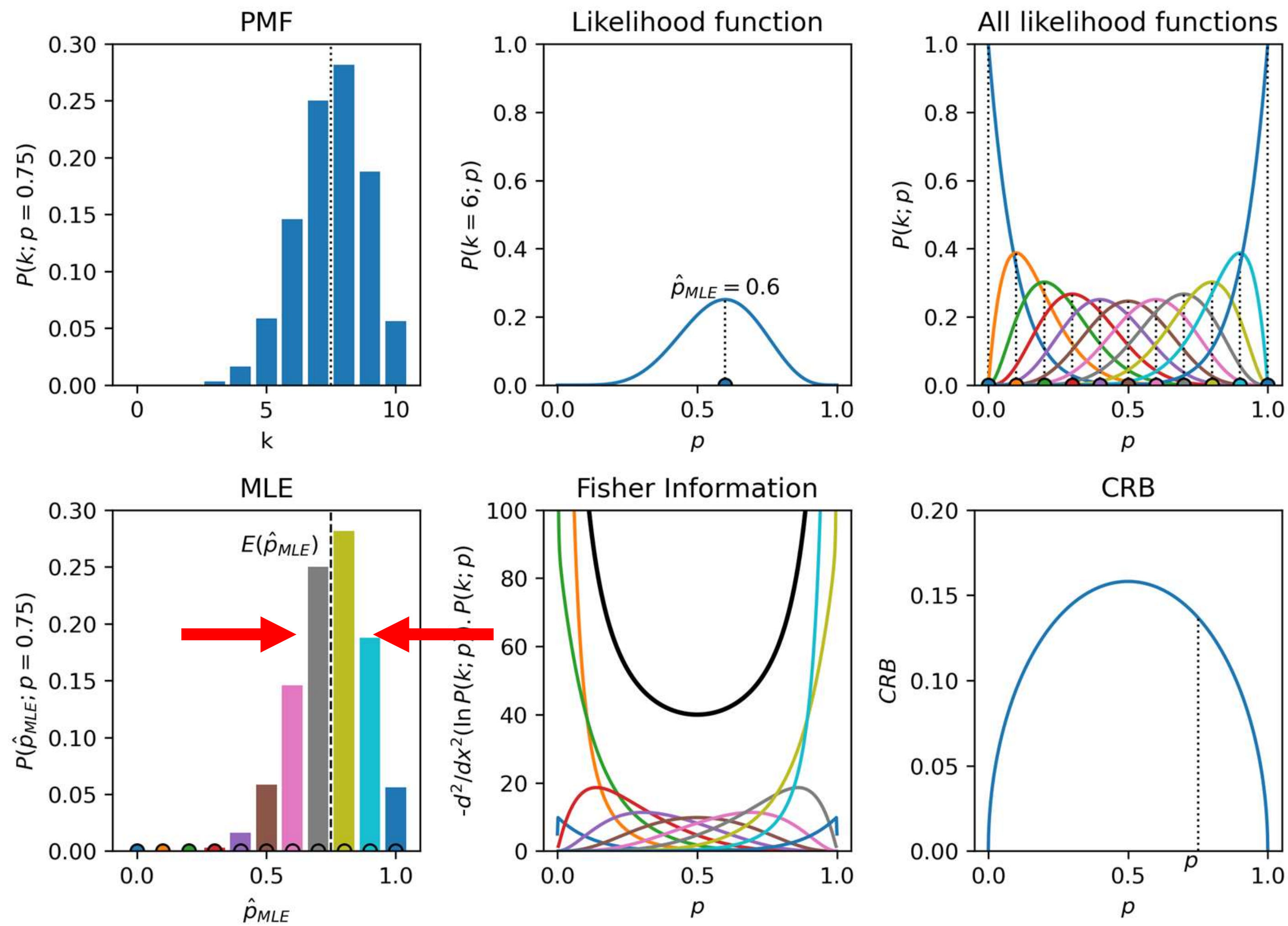
Estimator:

$$\hat{p}_{MLE} = \frac{k}{n}$$

Characterize:

$E(\hat{p}_{MLE})$

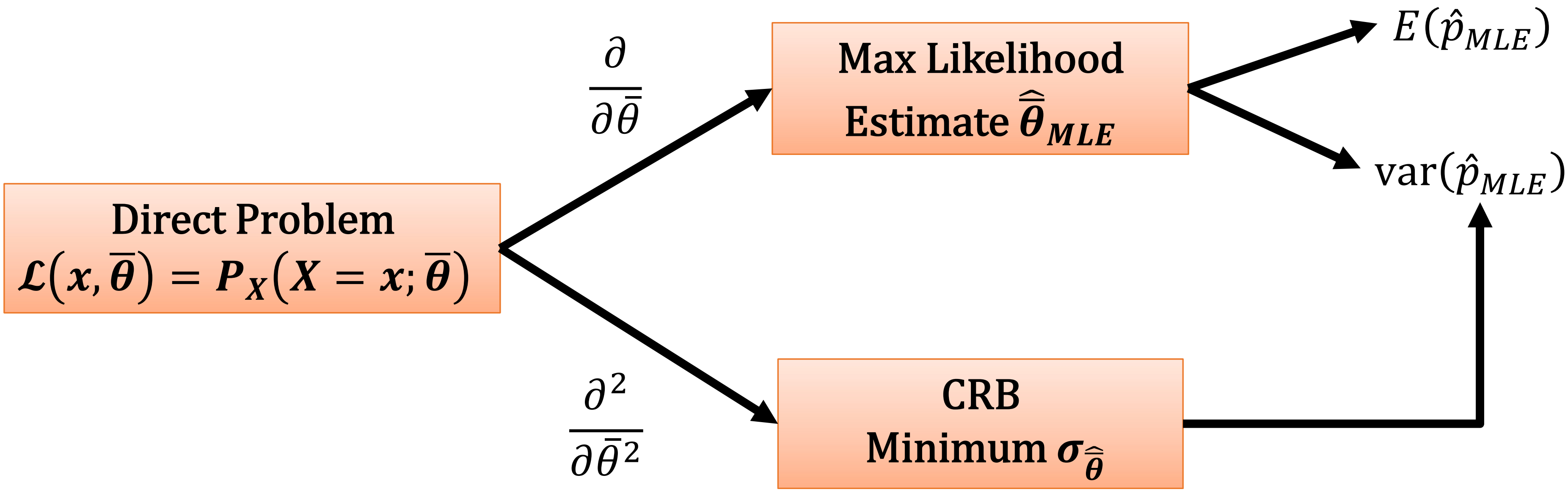
$\mathrm{var}(\hat{p}_{MLE})$



$$\mathrm{var}(\hat{p}) \geq \frac{1}{F_p}$$

$$\sigma_{CRB} = \sqrt{\frac{p(1-p)}{n}}$$

SUMMARY

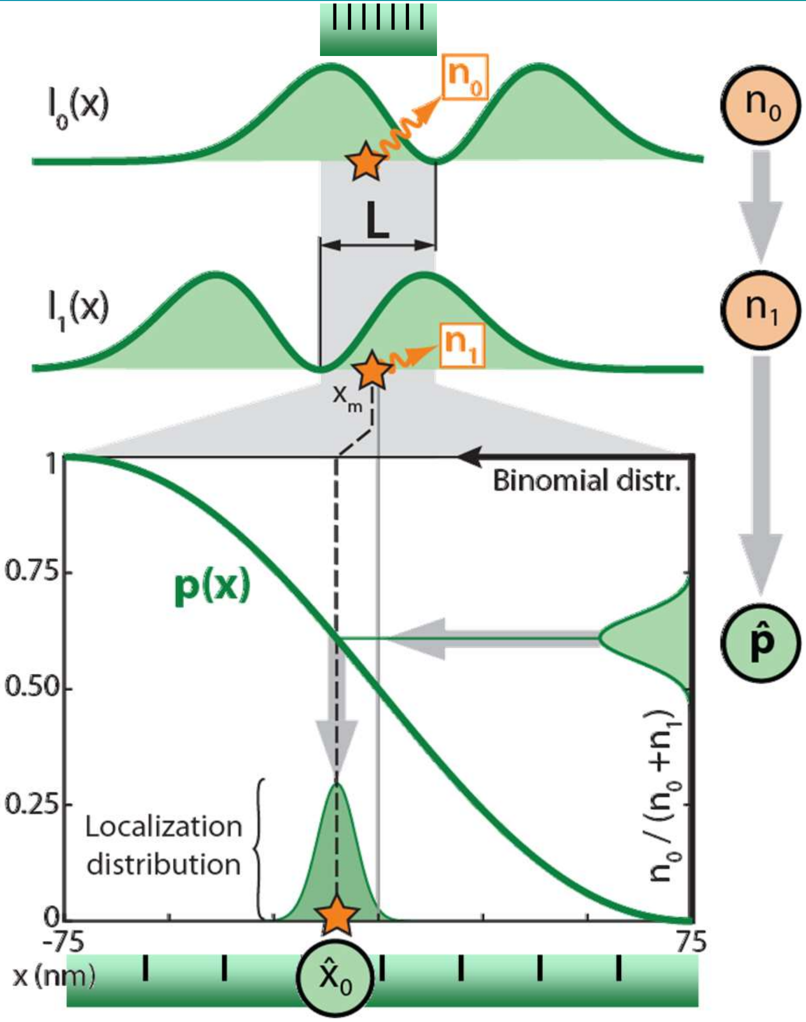


Direct Problem
$\mathcal{L}(x, \overline{\theta}) = P_X(X = x; \overline{\theta})$
$\frac{\partial}{\partial \overline{\theta}}$
Max Likelihood
Estimate $\widehat{\overline{\theta}}_{MLE}$
$\frac{\partial^2}{\partial \overline{\theta}^2}$
CRB
Minimum $\sigma_{\widehat{\overline{\theta}}}$
$E(\hat{p}_{MLE})$
$\text{var}(\hat{p}_{MLE})$

OTHER DISTRIBUTIONS

	Gaussian	Poissonian	Binomial
Variable	x	n	k
Parameters	$\vec{\theta} = [\mu, \sigma]$	$\vec{\theta} = [\lambda]$	$\vec{\theta} = [p]$
Distribution $p(x; \vec{\theta})$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{-\lambda} \frac{\lambda^n}{n!}$	$\binom{n}{k} p^k (1-p)^{n-k}$
Log-likelihood $\ln \mathcal{L}(x; \vec{\theta})$	$-\frac{(x-\mu)^2}{2\sigma^2} - \ln \sigma$	$-\lambda + n \ln \lambda - \ln n!$	$k \ln p + (n-k) \ln(1-p)$
$\frac{\partial}{\partial \theta} \ln \mathcal{L}(x; \vec{\theta})$	$\begin{bmatrix} -(x-\mu)/\sigma^2 \\ \frac{(x-\mu)^2}{2\sigma^3} - \frac{1}{\sigma} \end{bmatrix}$	$-1 + n/\lambda$	$\frac{k}{p} + \frac{n-k}{1-p}$
$\frac{\partial^2}{\partial \theta^2} \ln \mathcal{L}(x; \vec{\theta})$	$\begin{bmatrix} -1/\sigma^2 & 2\frac{(x-\mu)}{2\sigma^3} \\ 2\frac{(x-\mu)}{2\sigma^3} & -3\frac{(x-\mu)^2}{2\sigma^4} + \frac{1}{\sigma^2} \end{bmatrix}$	$-n/\lambda^2$	$-\frac{k}{p^2} + \frac{n-k}{(1-p)^2}$
MLE $\frac{\partial}{\partial \theta} \ln \mathcal{L}(x; \vec{\theta}) = 0$	$\hat{\mu}_{\mathrm{MLE}} = x$	$\hat{\lambda}_{\mathrm{MLE}} = n$	$\hat{p}_{\mathrm{MLE}} = k/n$
$\mathrm{E}(\hat{\theta})$	$\mathrm{E}(x) = \mu$	$\mathrm{E}(n) = \lambda$	$\mathrm{E}(k/n) = p$
$\mathrm{var}(\hat{\theta})$	$\mathrm{var}(x) = \sigma^2$	$\mathrm{var}(n) = \lambda$	$\mathrm{var}(k/n) = p(1-p)/n$
Fisher Information $E\left(-\frac{\partial^2}{\partial \theta^2} \ln \mathcal{L}(x; \theta)\right)$	$\begin{bmatrix} 1/\sigma^2 & 0 \\ 0 & 2/\sigma^2 \end{bmatrix}$	$1/\lambda$	$\frac{n}{p(1-p)}$
CRB	$\begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2/2 \end{bmatrix}$	λ	$\frac{p(1-p)}{n}$

MINFLUX LOCALIZATION ESTIMATION



n_0 ~ Poissonian $E(n_0) \propto I_0(x_m)$

n_1 ~ Poissonian $E(n_1) \propto I_1(x_m)$

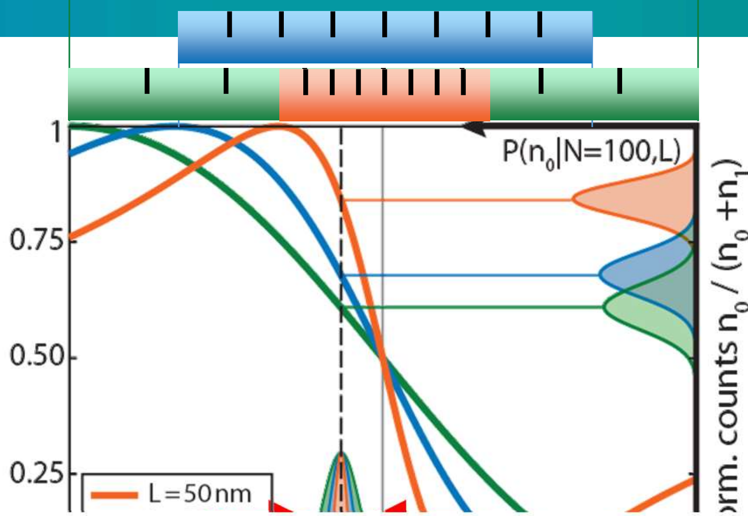
$n_0|n_0 + n_1 = N$ ~ Binomial $p(x) = \frac{I_0(x)}{I_0(x)+I_1(x)}; n = N$

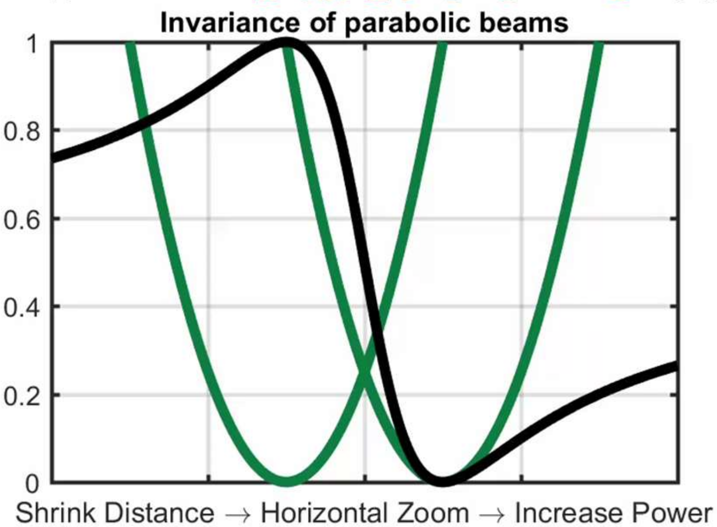
~~$k \rightarrow \mathcal{L}(k, p) \rightarrow \hat{p}_{MLE} = k/n$~~

$$n_0 \rightarrow \mathcal{L}(n_0, p(x)) \rightarrow \hat{p}_{MLE} = n_0/N$$

$\hat{x}_{MLE}$ such that $\hat{p}_{MLE} = p(\hat{x}_{MLE})$

$\frac{60}{60+40} = 0.6 = p(\mathrm{x})$ → find x





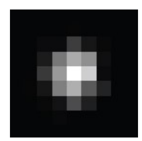
MAXIMALLY **INF**ORMATIVE **LU**MINESCENCE **EX**CITATION

$$\tilde{\sigma}^{quad}_{CRB}(x) = \frac{1}{\sqrt{N}}\frac{L}{4}\left[1+\left(\frac{x}{L/2}\right)^2\right]$$

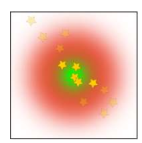
INFORMATION

$$\tilde{\sigma}^{quad}_{CRB}(0) = \frac{1}{\sqrt{N}}\frac{L}{4}$$

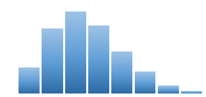
!!



$$\sigma_{CAM} \propto \frac{\lambda}{NA}\frac{1}{\sqrt{N}}$$

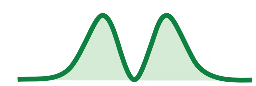


$$\sigma_{STED} \propto \frac{\lambda}{NA}\frac{1}{\sqrt{1+I/I_S}}$$



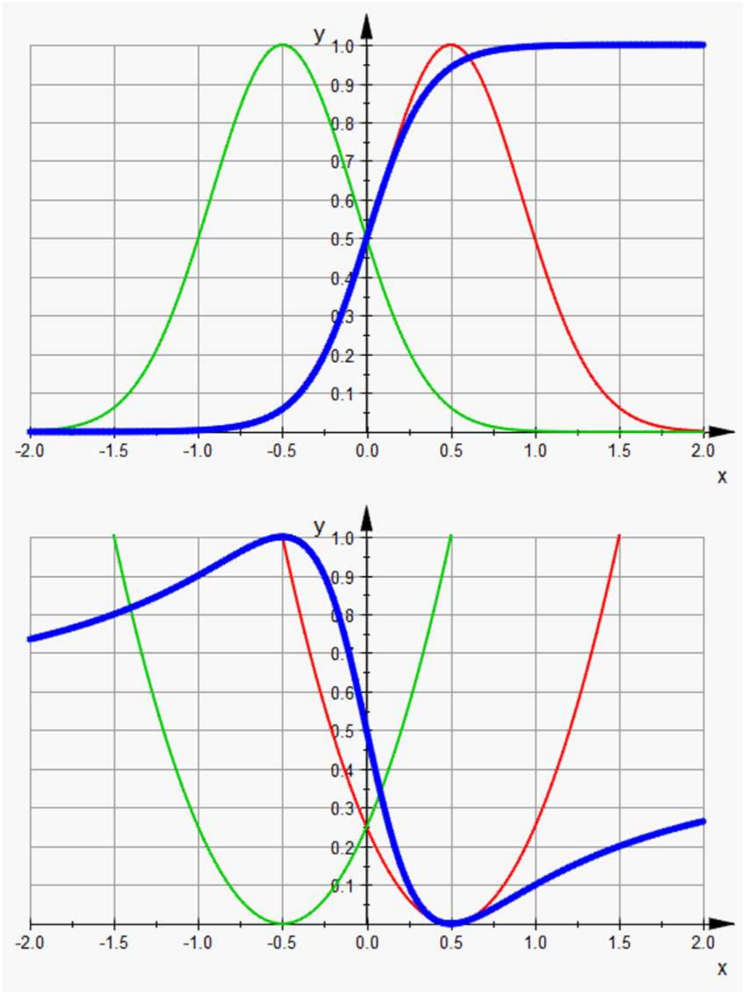
Poisson distribution

+



Donut beam shape

THE MAPPING OF STATISTICS VIA $p(x)$ ENABLES MANIPULATING THE INFORMATION



CRITICAL ELEMENTS:

1. POISSON STATISTICS

2. INHOMOGENEOUS ILLUMINATION