Shaky, Spiky, STORMy & Sharp: Statistical analysis of super-resolution and hyperacuity from microscopes to retinas

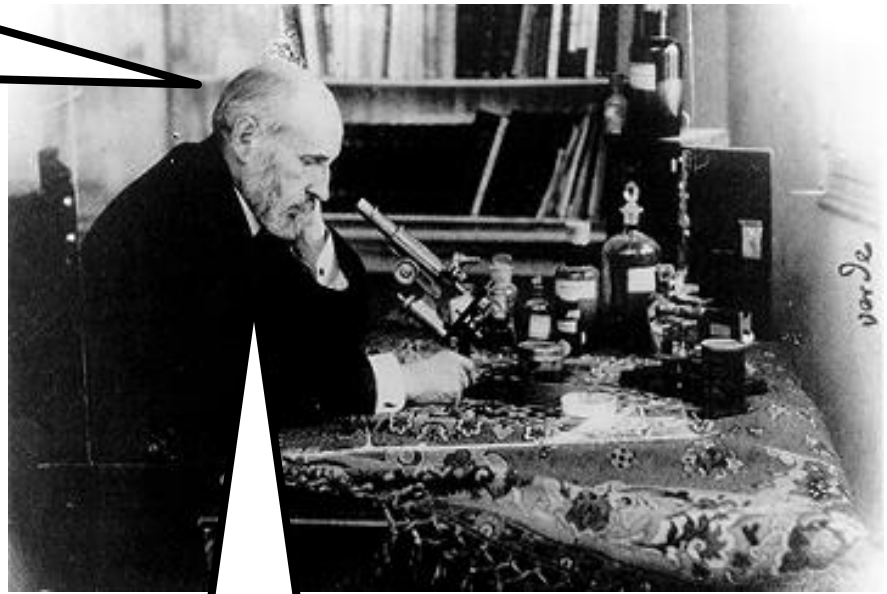
Eran Mukamel

Swartz Program in Theoretical Neuroscience
Harvard Center for Brain Science

Questions that (should have) kept Cajal up at night

What limits optical resolution?

Santiago Ramón y Cajal, ca. 1900



What limits visual acuity?

Outline: Acuity and Resolution

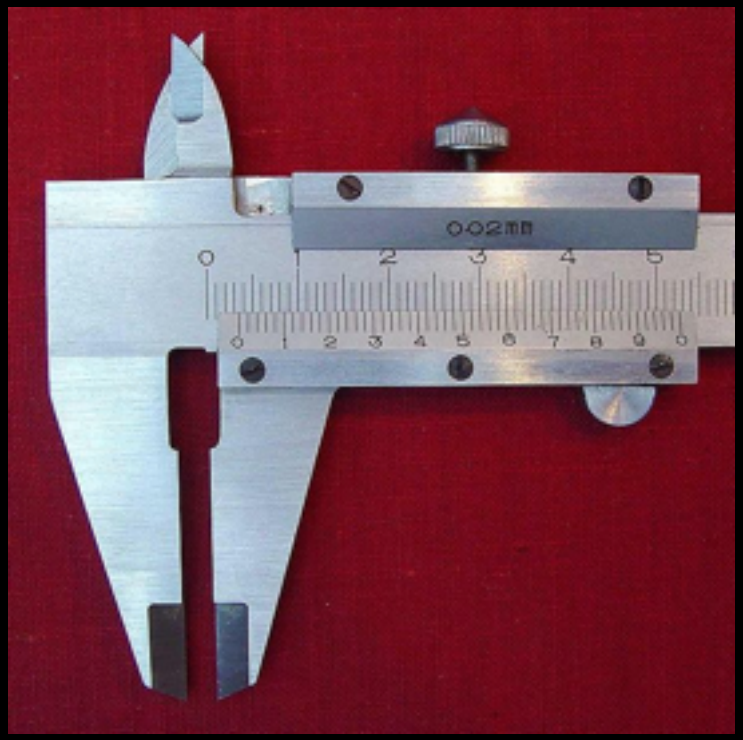
I. Visual acuity

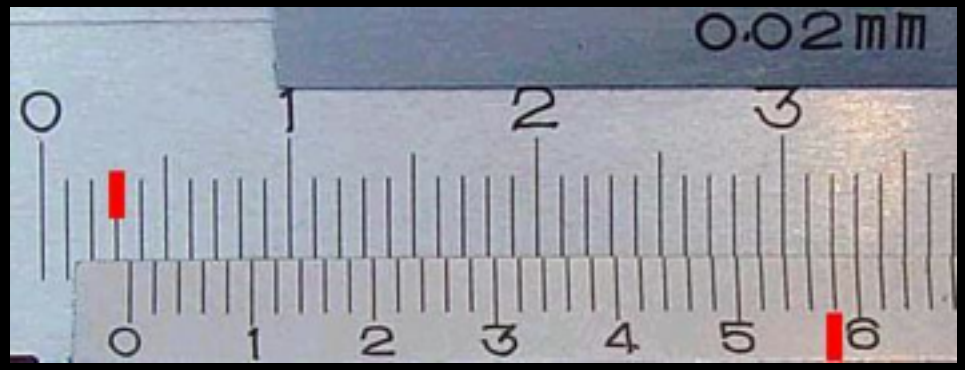
- What are the fundamental, statistical limits on performance in visual hyperacuity amid fixational eye movements?
- A biologically plausible neural network decoder for fine scale vision.

II. Super-resolution microscopy

- What fundamental limits apply to conventional and super-resolution techniques?

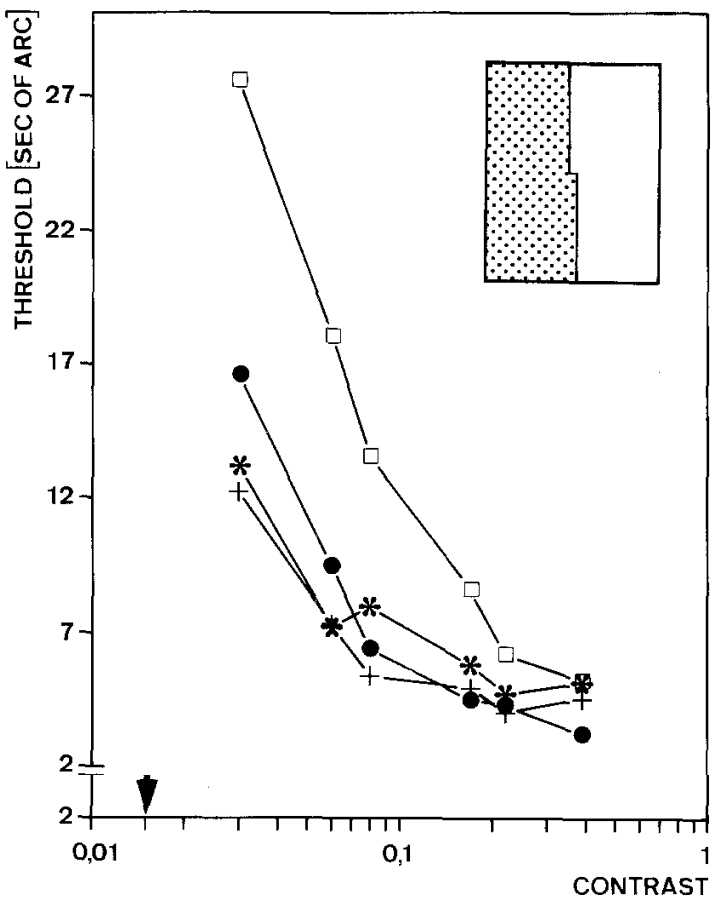
Vernier calipers





Humans achieve hyperacuity in Vernier tasks

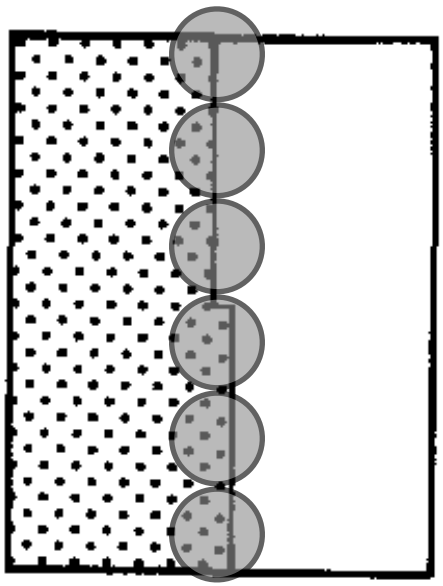
Photoreceptor/ganglion cell spacing in the fovea (30")



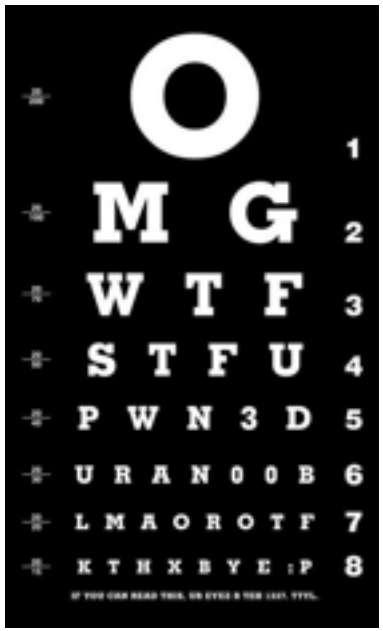
Wehrhahn and Westheimer (1990)

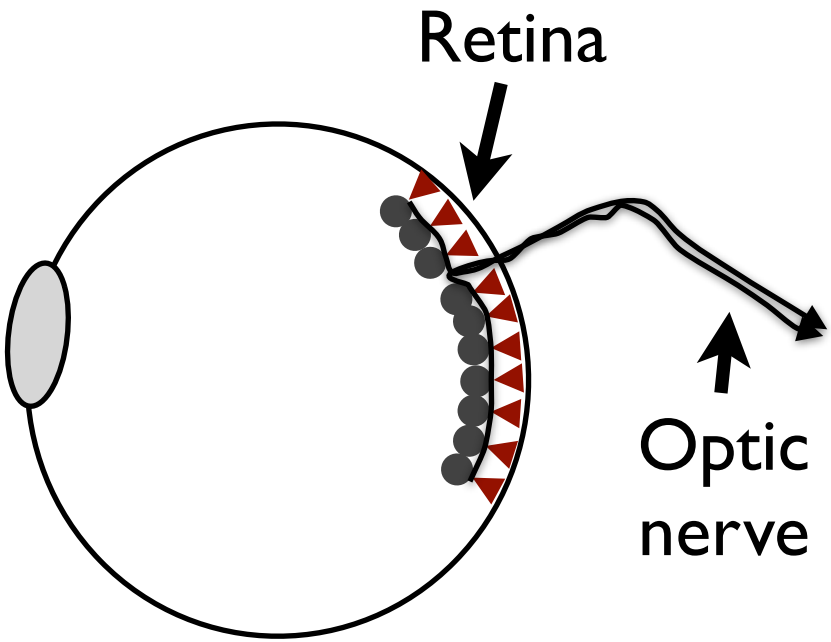
Hering's theory of Vernier acuity

- Ewald Hering (1899) proposed integration of information from multiple photoreceptors along the length of the bar

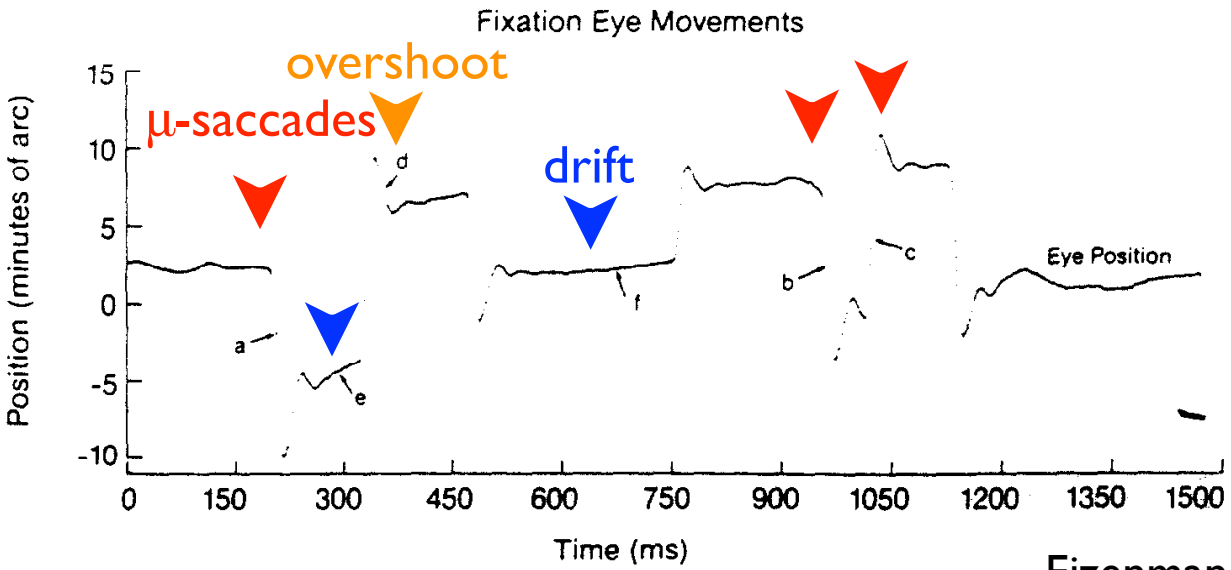


Eyes are never still





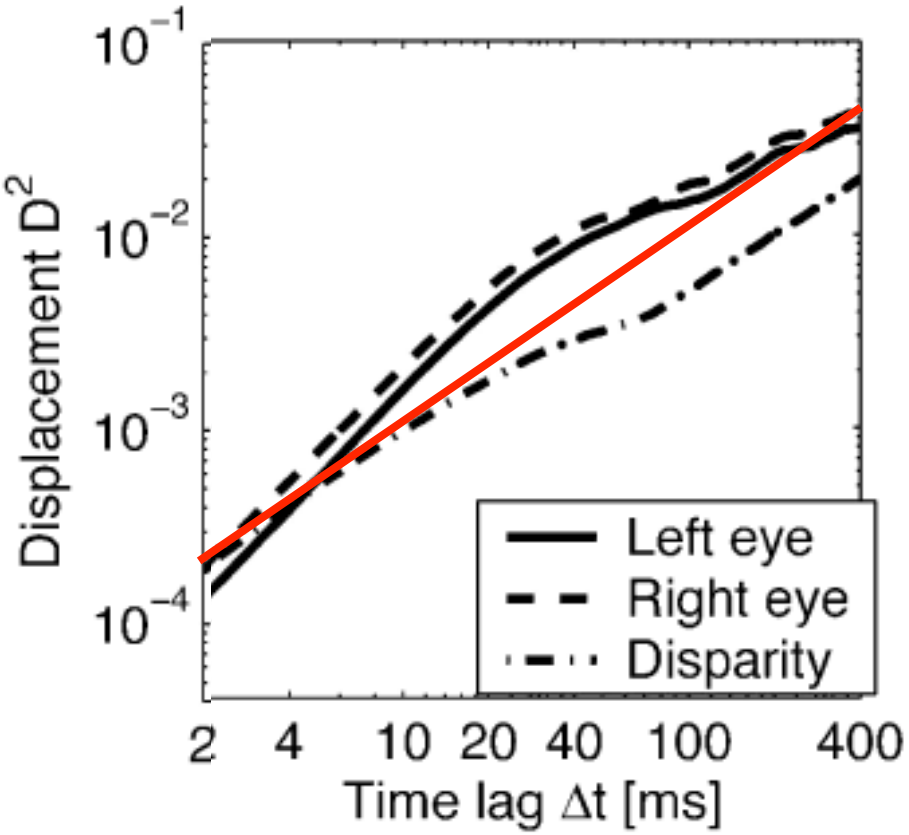
Small eye movements are significant



Eizenman *et al.* (1985)

- Peak amplitude of 40-100 Hz power (with μ-saccades removed) is ~6 arc-sec, similar to Vernier acuity threshold

Eye movements approximate a random walk



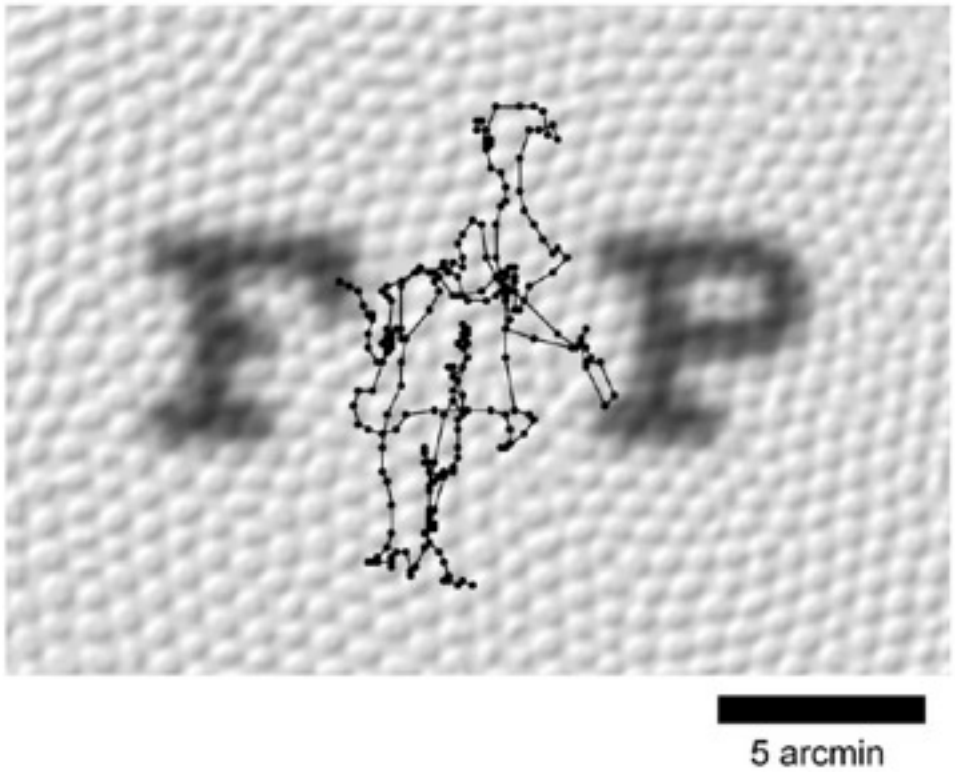
$$\left\langle \Delta x^2 \right\rangle \approx D\Delta t$$

Effective diffusion coefficient:

$$D \approx 100 \frac{\text{arc-min}^2}{s}$$

Engbert and Kliegl (2004)

Fixational eye movements are comparable to photoreceptor spacing



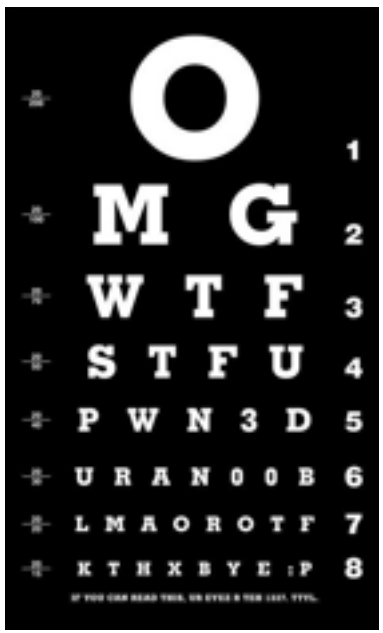
Fixational eye movement trace lasting 500 ms, sampled every 2 ms

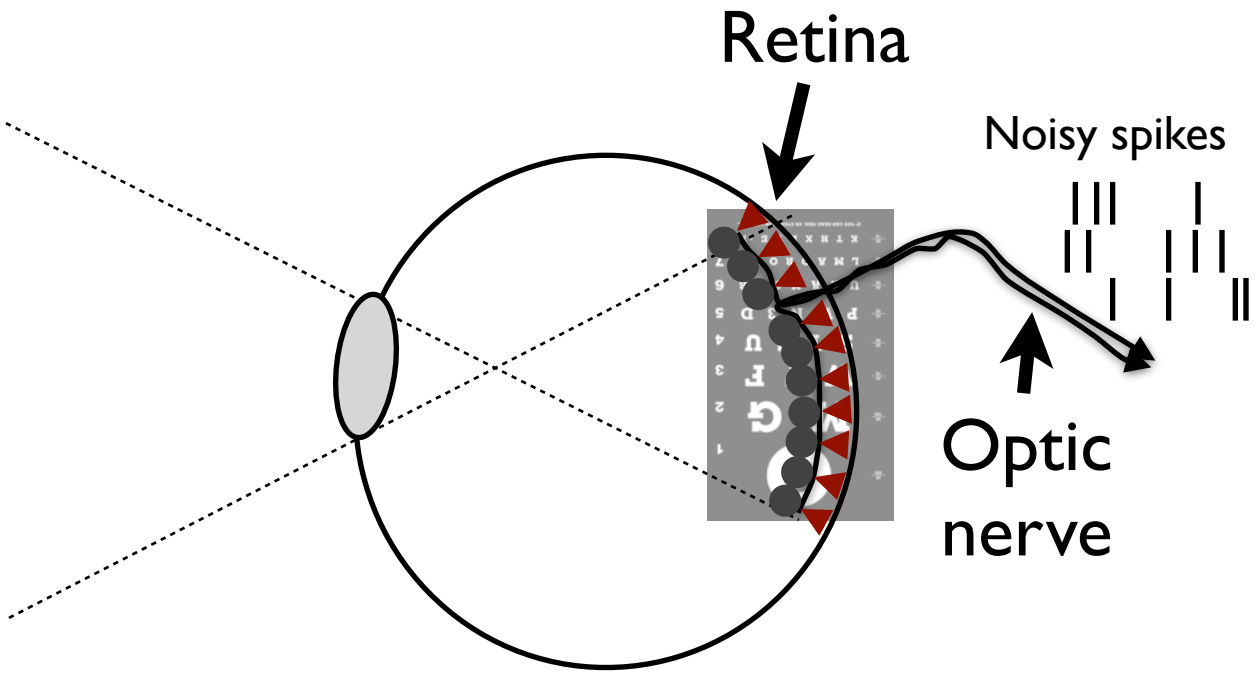
Pitkow, Sompolinsky and Meister, 2007

Open problems in visual hyperacuity

1. How does fine-scale visual acuity depend on:
   - Photoreceptor spacing?
   - Fixational eye movement amplitude?
   - Spike train statistics?
2. What is the optimal decoder of retinal spike trains?
3. Is there a near-optimal, biologically feasible decoder?

A statistical model of neural encoding for fine vision

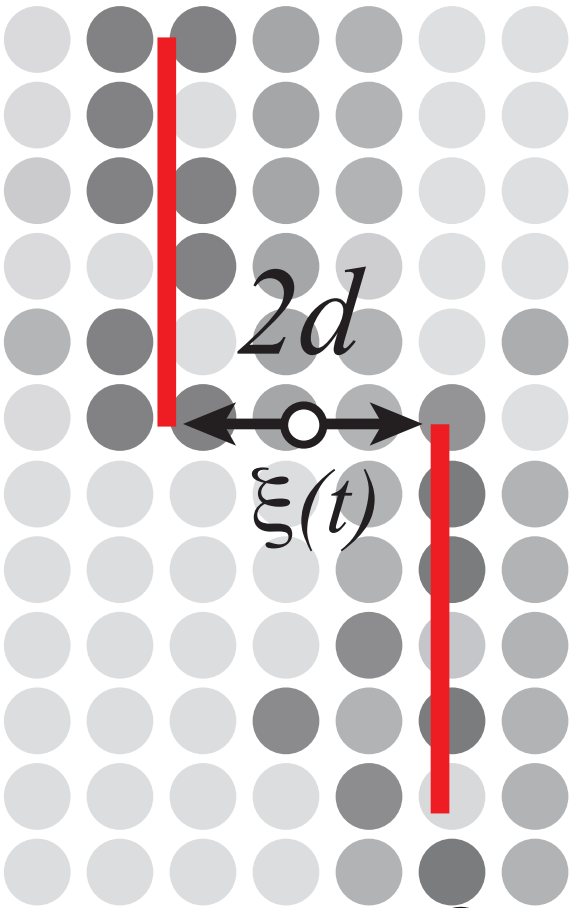




Key assumptions:

1. Linear, Gaussian receptive fields
2. Poisson spiking in retinal ganglion cells
3. Diffusive (random walk) eye movements

Parameters of the model for the Vernier task



1. Receptive field spread

$$\sigma \approx 2.5 - 3.5 \text{ arc-min}$$

2. Eye movement amplitude

$$D \approx 100 \frac{\text{arc-min}^2}{s}$$

3. Ganglion cell spike rate

$$\rho_{max} \leq 100 \text{ Hz}$$

4. Number of ganglion cells

$$N \approx 1$$

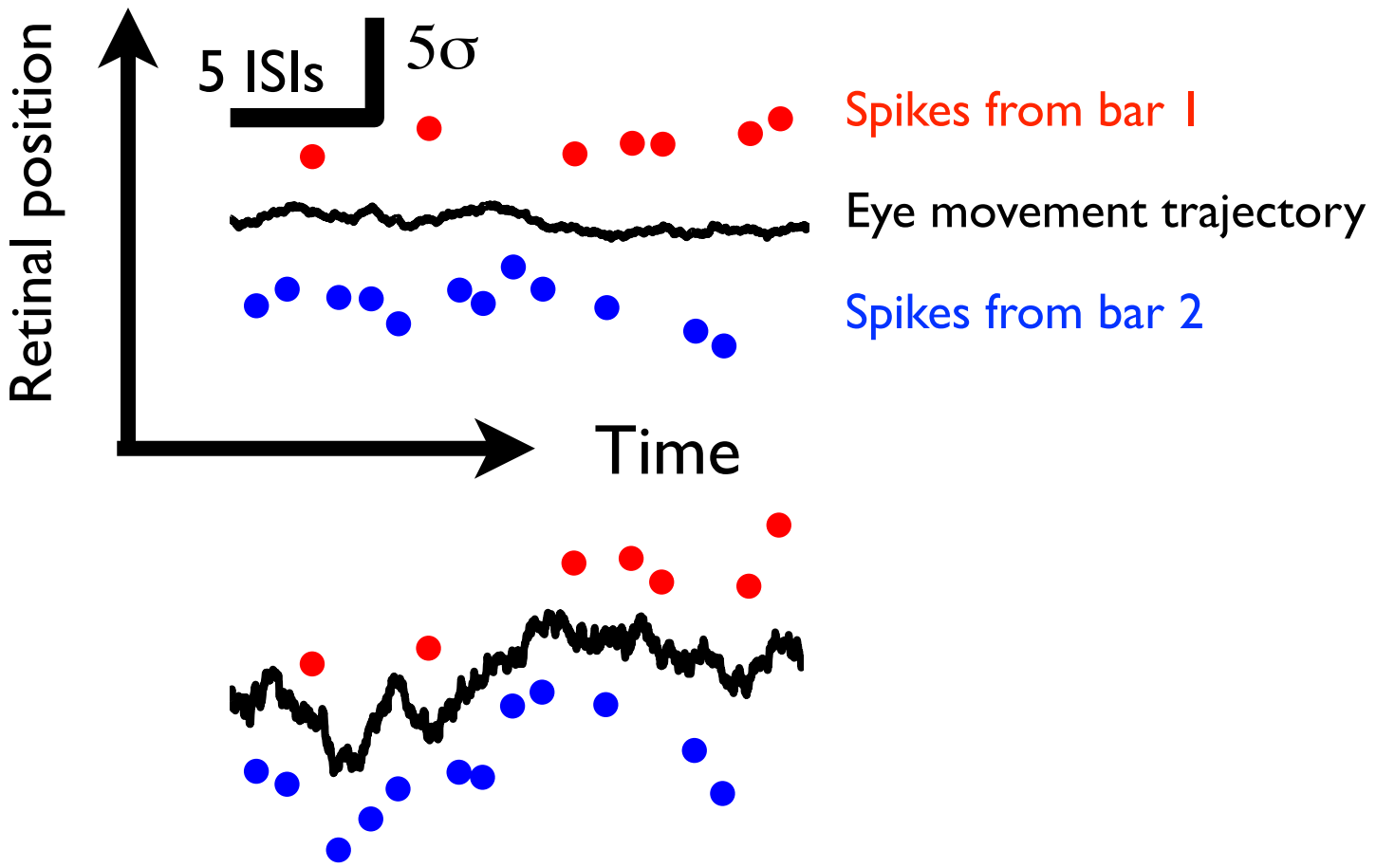
A single dimensionless parameter sets the difficulty of Vernier estimation

$$\varepsilon = \frac{D}{N\rho\sigma^2} \approx 0.1 - 1$$

The mean squared displacement of the eyes between individual spikes, in units of the receptive field size

Simulated trajectories and spike trains

Slow eye movements:
$\varepsilon = 0.1$



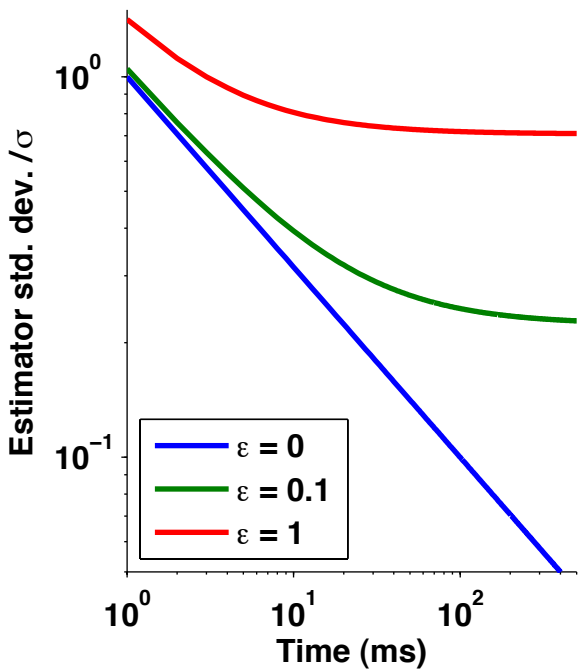
Fast eye movements
$\varepsilon = 10$

A simple averaging estimator fails to achieve hyperacuity

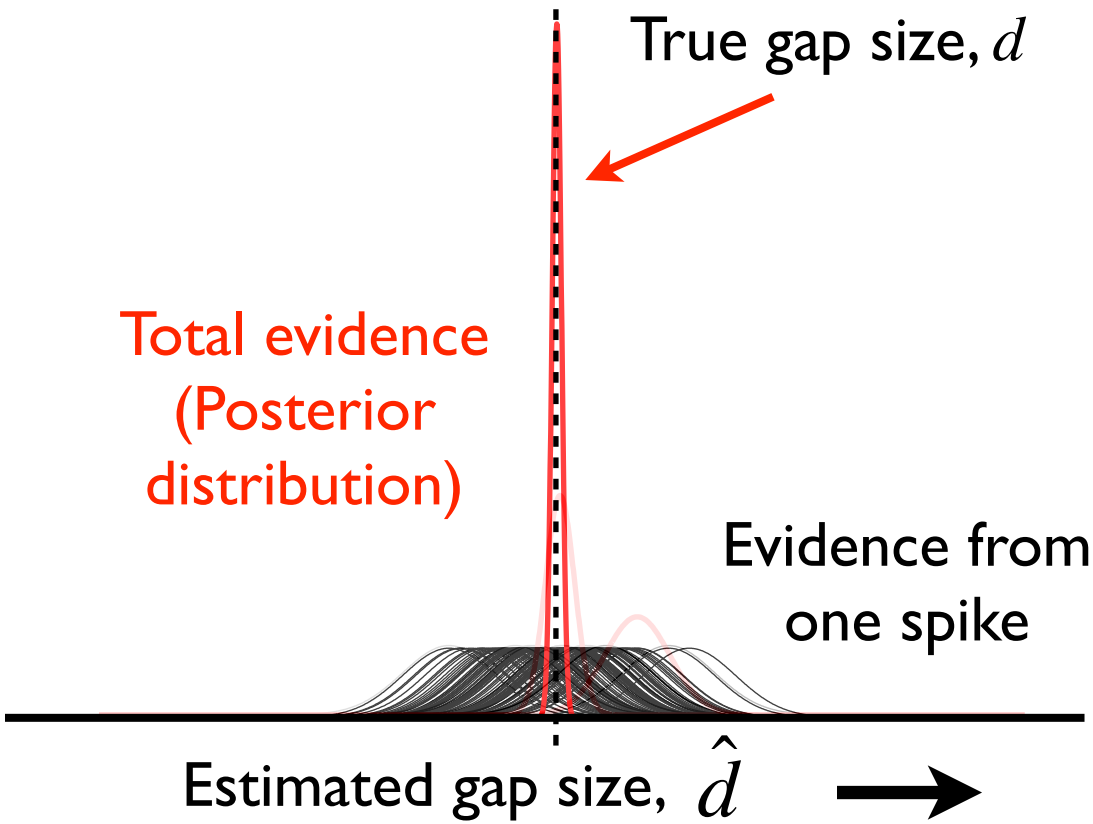
$$\hat{d}_{Naive} = \langle x_i \rangle_{Bar1} - \langle x_i \rangle_{Bar2}$$

$$\mathrm{var}(\hat{d}_{Naive}) \approx \frac{\sigma^2}{\rho T}\left[1 + \frac{\varepsilon}{2}\right] + \sigma^2 \frac{\varepsilon}{2}$$

Minimum error

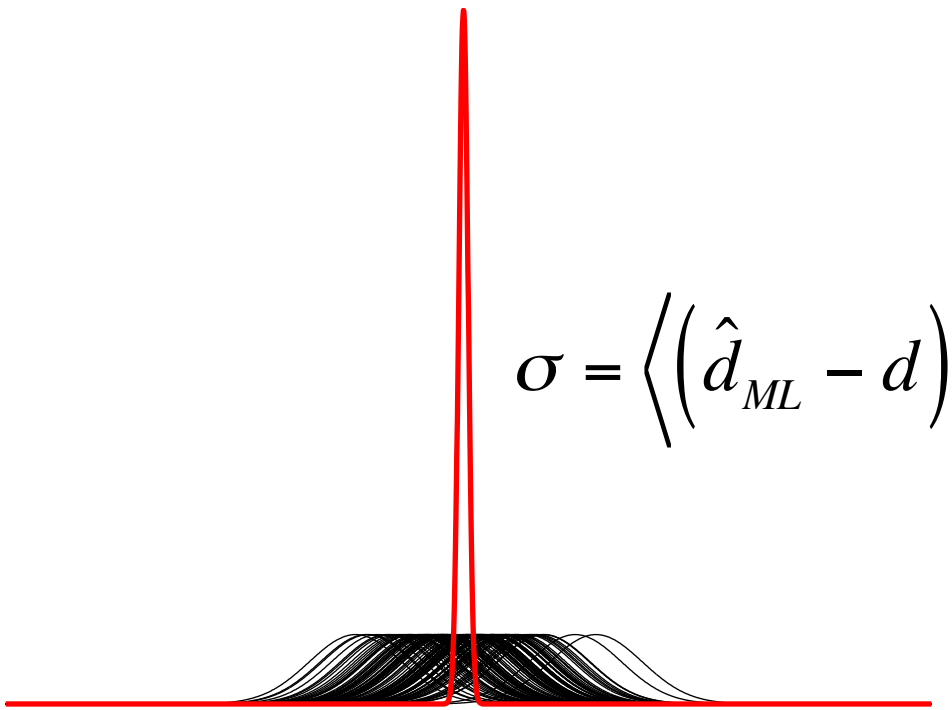


Evidence from individual spikes accumulates over time



- Gaussian evidence from each spike leads to a Gaussian posterior distribution.

Posterior distribution for the gap parameter is Gaussian



σ Estimator uncertainty

$$\sigma = \left\langle \left(\hat{d}_{ML} - d \right)^2 \right\rangle^{1/2} \propto$$ *Psychophysical threshold*

$\hat{d}_{ML}$ Maximum likelihood (optimal) estimate of gap size:

The optimal estimator is a function of all the spike times

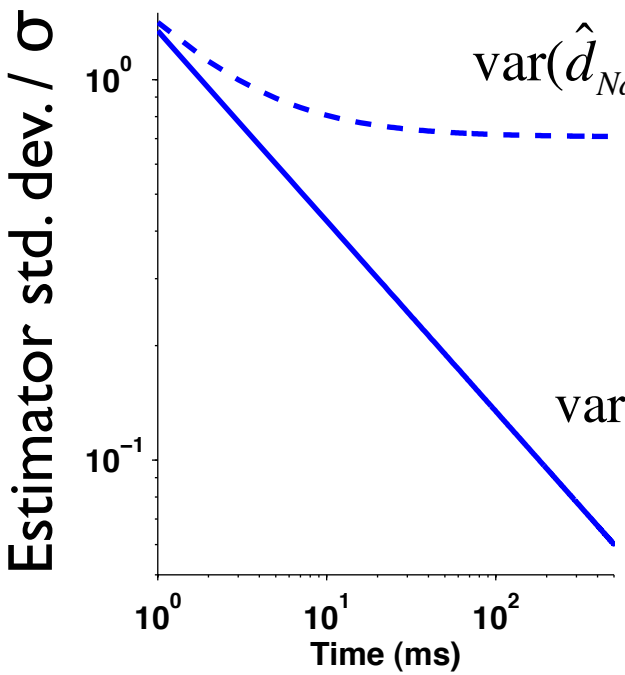
$$\hat{d}_{ML} = 2\frac{\mathbf{s}^T\Sigma^{-1}\mathbf{x}}{\mathbf{s}^T\Sigma^{-1}\mathbf{s}}$$

where the **decoder matrix** is:

$$\Sigma = \sigma^2\left[\mathbf{Id} + \varepsilon\mathbf{A}\Delta\mathbf{A}^{\mathbf{T}}\right]$$

$$\mathbf{A} \equiv \begin{bmatrix} 1 & 0 & 0 & \\ 1 & 1 & 0 & \\ 1 & 1 & 1 & \\ & & & \ddots \end{bmatrix}, \quad \Delta = \begin{bmatrix} \ddots & & \\ & t_{i+1} - t_i & \\ & & \ddots \end{bmatrix}.$$

The optimal decoder achieves hyperacuity amid eye movements



$$\mathrm{var}(\hat{d}_{Naive}) \approx \frac{\sigma^2}{\rho T}\left[1+\frac{\varepsilon}{2}\right]+\sigma^2\frac{\varepsilon}{2}$$

$$\mathrm{var}(\hat{d}_{ML}) = \frac{\sigma^2}{\rho T}\frac{\sqrt{4+\varepsilon}}{\sqrt{4+\varepsilon}-\sqrt{\varepsilon}}$$

Optimal decoders for slow and fast eye movements

- In the limit of slow eye movements ($\varepsilon << 1$) the optimal decoder simply averages each population:

$$\hat{d}_{ML} \approx \mathbf{s}^T \mathbf{x} / N = \left(\sum x_i^+ - \sum x_i^- \right) / N$$

- In the limit of fast eye movements ($\varepsilon >> 1$) use only pairs of nearly-coincident spikes in opposite populations, weighted by the inverse inter-spike interval:

$$\hat{d}_{ML} \approx \left(\sum_{+\,switches} \frac{\Delta x_i}{\Delta t_i} - \sum_{-\,switches} \frac{\Delta x_i}{\Delta t_i} \right) \Bigg/ \sum_{switches} \frac{1}{\Delta t_i}$$

Outline: Acuity and Resolution

I. Visual acuity

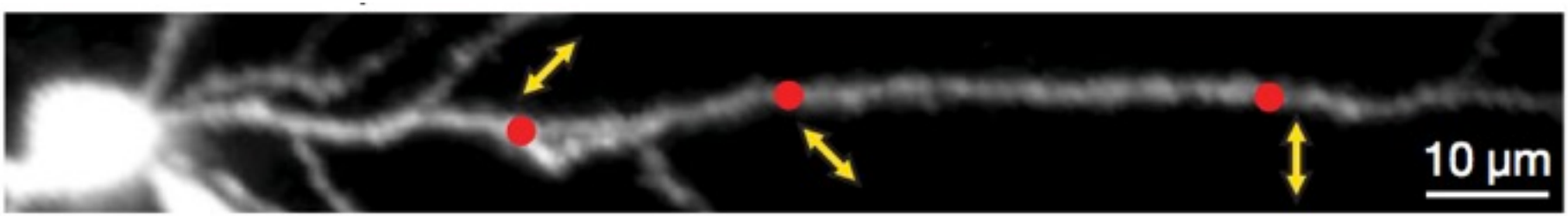
- What are the fundamental, statistical limits on performance in visual hyperacuity amid fixational eye movements?
- A biologically plausible neural network decoder for fine scale vision.

II. Super-resolution microscopy

- What fundamental limits apply to conventional and novel techniques?

Diffraction-limited images are blurry at the nano-scale

Two-photon Ca^{2+}-imaging of dendritic "hot-spots" *in vivo*

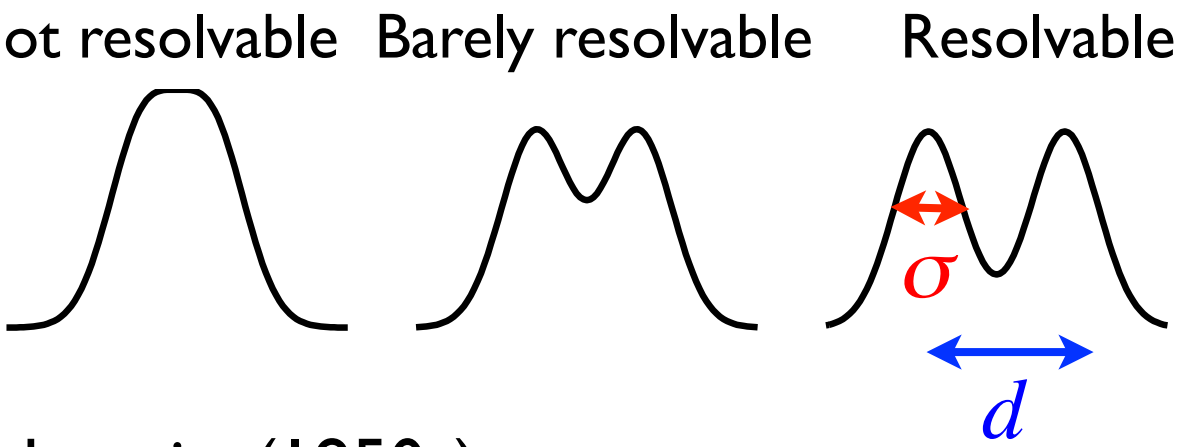


Beyond the limit: spines? vesicles? synapses? receptors?

Jia et al., Nature (2010)

The classical diffraction limit characterizes optical blurring

Real-space: Lord Rayleigh, Ernst Abbe (1870s)



Frequency domain: (1950s)

$$A(\mathbf{x}) = \int I(\mathbf{x})h(\mathbf{x}-\mathbf{x}')d\mathbf{x}' \Rightarrow A(\mathbf{k}) = I(\mathbf{k})h(\mathbf{k})$$

$$P(\mathbf{k}) = \left|h(\mathbf{x})\right|^2 \left|I(\mathbf{x})\right|^2$$

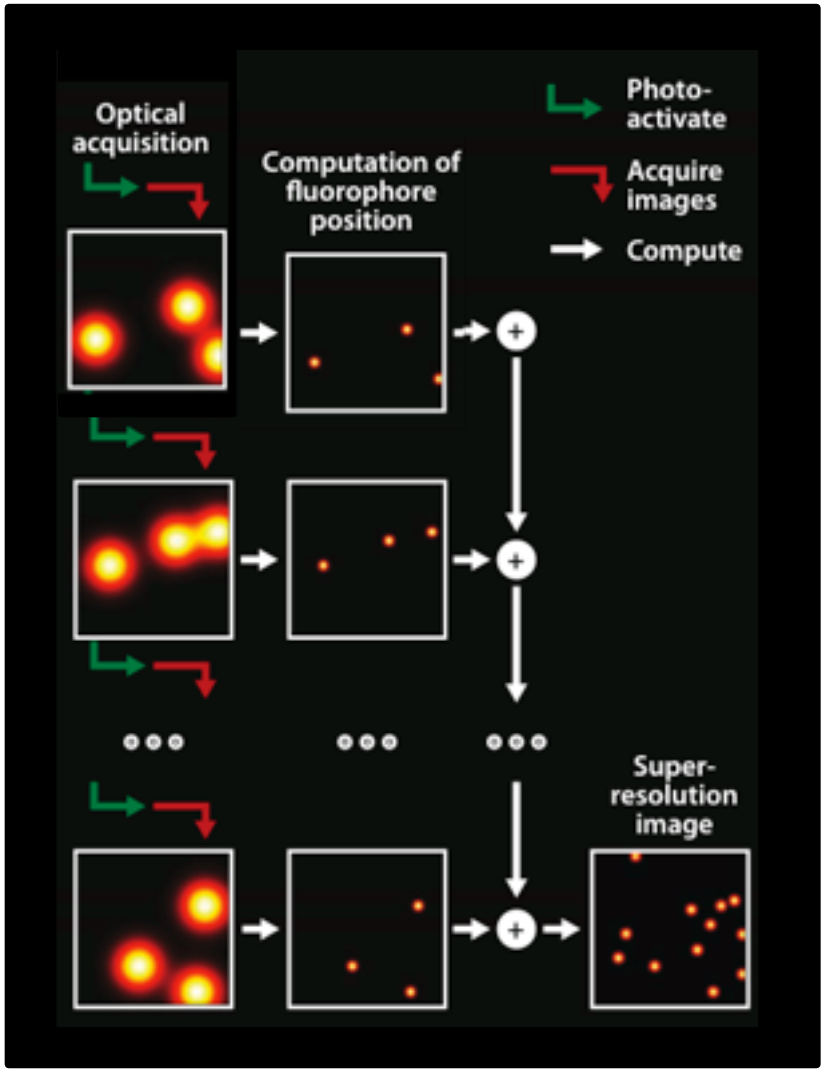
Point-spread function
(incoherent, translation-invariant imaging)

Modulation transfer function

Optical imaging at the nano-scale

- Ways to beat the diffraction limit:
  1. Shorter wavelength (X-ray, EM)
  2. Near-field optics
  3. Multiphoton microscopy (e.g. two-photon, STED)
  4. Stochastic localization microscopy (e.g. STORM, PALM, etc.)

Stochastic localization is a new imaging paradigm



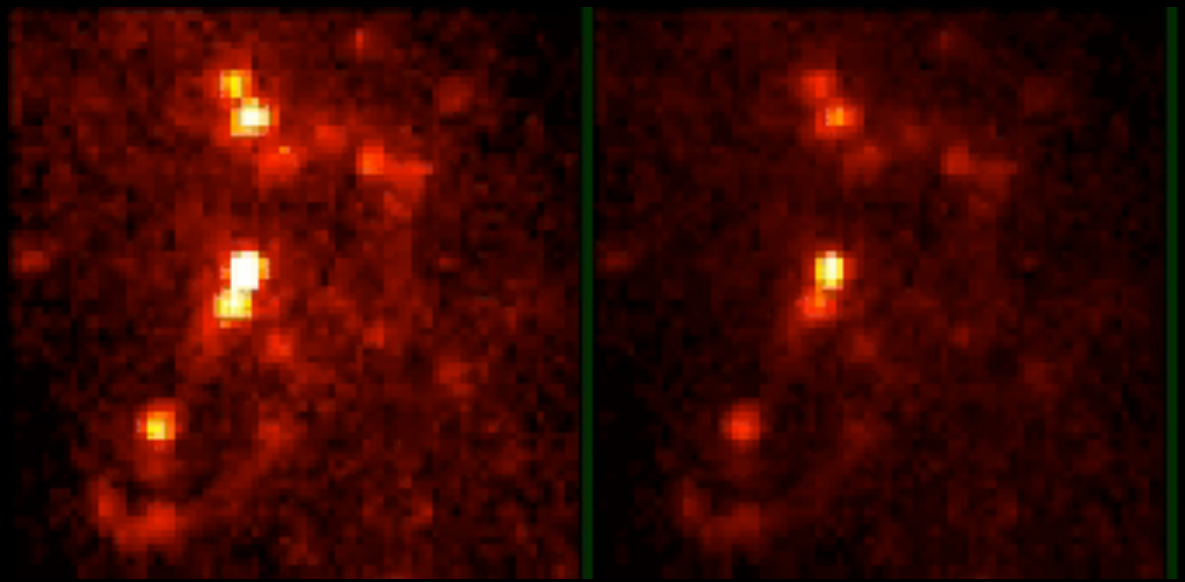
Wilt *et al.*, Ann. Rev. Neurosci. (2009)

Example of stochastic localization data

Raw data

Conventional image

Super-resolution

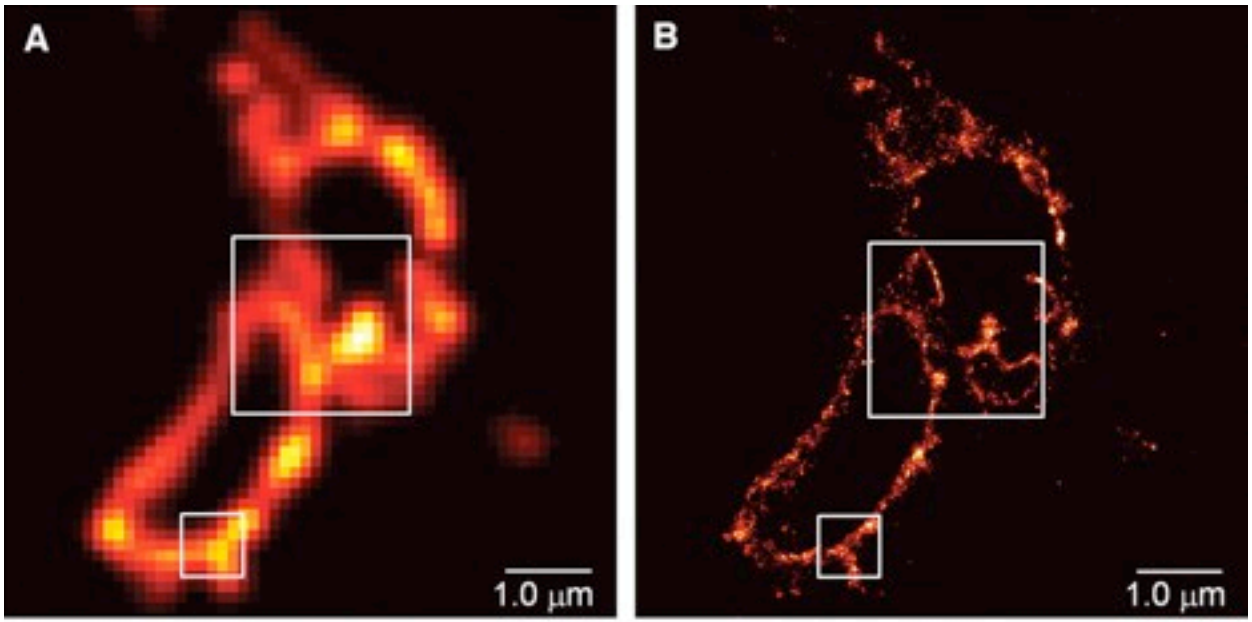


0.5 μm

Beating the diffraction limit by localization microscopy

Summed fluorescence

Single-molecule reconstruction



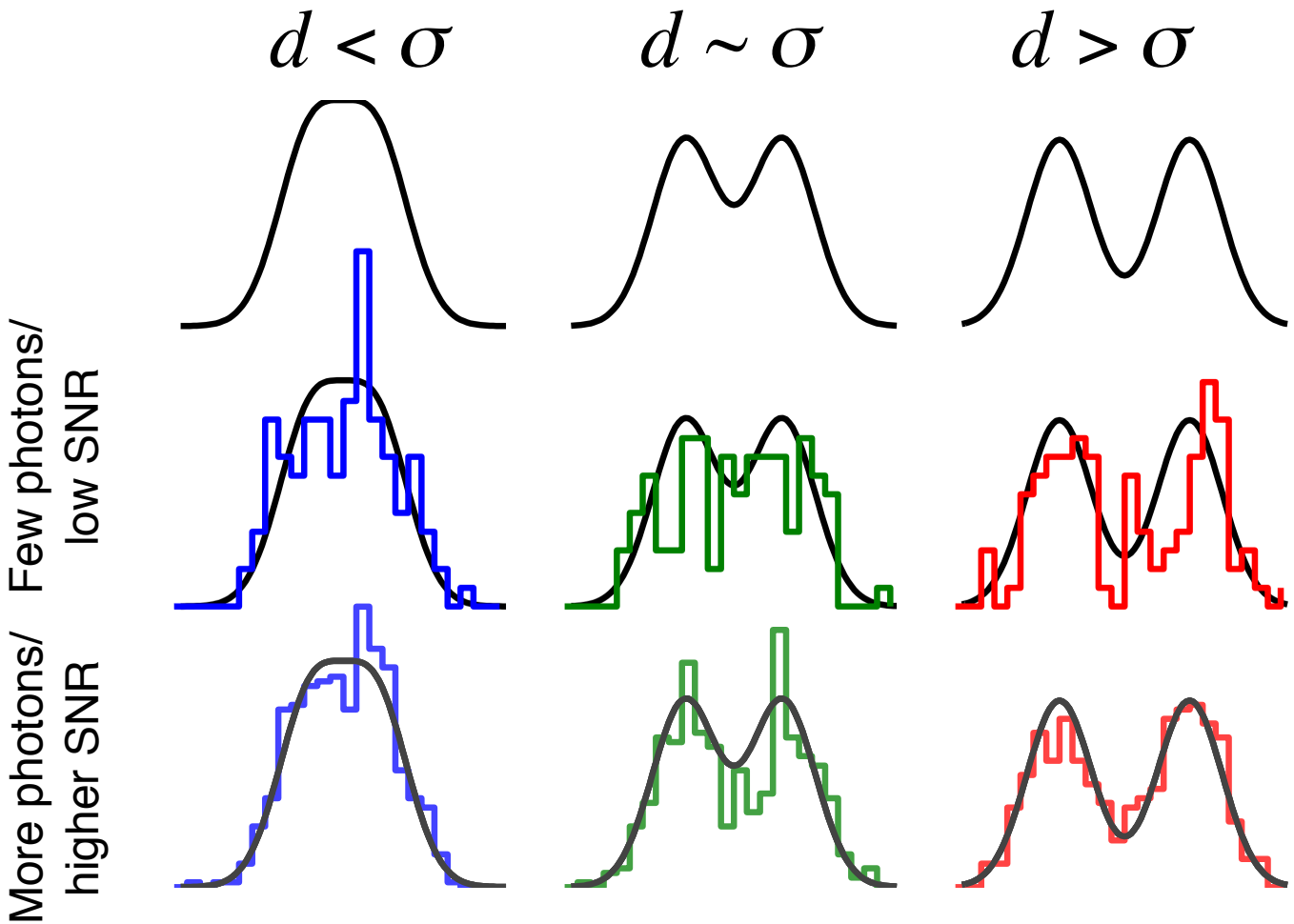
COS-7 cell expressing FP-tagged lysosomal trans-membrane protein membrane

E. Betzig et al., Science 313, 1642 -1645 (2006)

Where did Rayleigh and Abbe (and Born, Wolf, *et al.*) go wrong?

- If diffraction is a fundamental "limit," how does localization transcend it?

Rayleigh's criterion ignores signal/noise ratio



Estimation theory bounds localization performance

- Start with a statistical model of measurement:

Observed Hidden

$$P(\{\vec{x}_1, \vec{x}_2, \ldots\} \mid \theta)$$

  - For example, a good model of stochastic localization microscopy:

Photons Fluorophore location

$$P(\{x_1, x_2, \ldots\} \mid \theta) = \prod_i h(x_i - \theta)$$

Microscope point spread function

Fundamental theorem of estimation theory bounds error

$$Bias = 0 \quad \Rightarrow \quad Var[\hat{\theta}] \geq J_\theta^{-1}$$

The **Fisher information,** J, measures the sensitivity of the observations to the parameter:

$$J_\theta = \left\langle \left[\frac{d}{d\theta} \log P(\{\vec{x}_i\} \mid \theta) \right]^2 \right\rangle_{\{\vec{x}_i\}}$$

In general, the maximum likelihood estimator achieves the fundamental bound in the limit of large N.

An example of the fundamental bound

Gaussian noise in 1D:

$$P(x \mid \theta) \propto \exp\left[-\frac{(x-\theta)^2}{2\sigma^2}\right]$$

Fisher information:

$$J_\theta = \frac{1}{\sigma^2}$$

Fundamental bound:

$$Var(\hat{\theta}) \geq \sigma^2$$

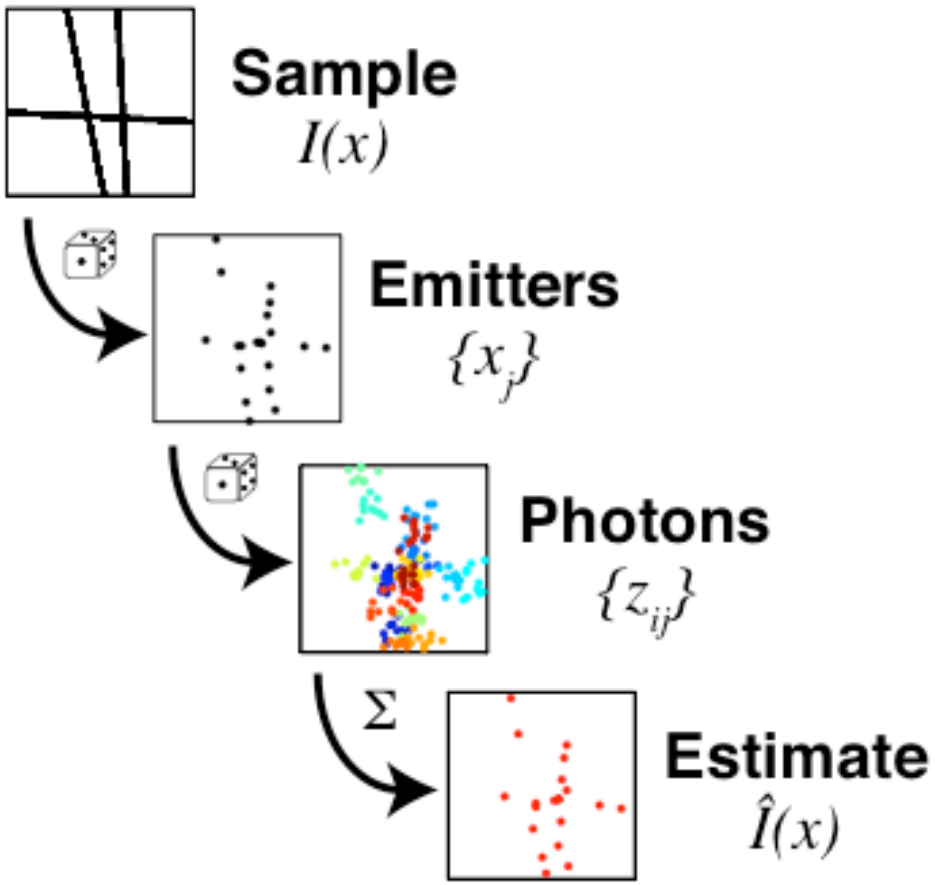
Examples of estimators:

$$\hat{\theta}_A = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\hat{\theta}_B = Median[\{x_i\}]$$

$$\hat{\theta}_C = x_1$$

Single-molecule localization techniques are stochastic



- The biological tissue is stochastically labeled by fluorophores
- Each fluorophore stochastically generates photons
- The image estimate reconstructs the sample from the photons

Information transfer function (ITF) bounds image estimation

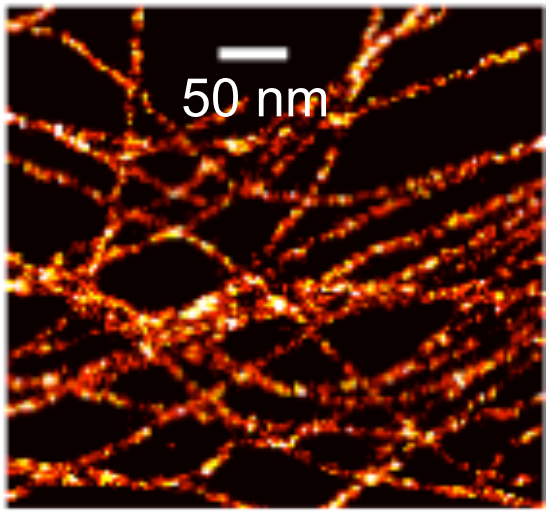
$$\left\langle \left[\hat{I}(k) - I(k) \right]^2 \right\rangle \geq [\mathbf{J}^{-1}]_{k,k}$$

Information transfer function (ITF):

$$F(k) = 1 / [\mathbf{J}^{-1}]_{k,k}$$

The effect of imaging parameters on a biological image estimate

Microtubules (True image)

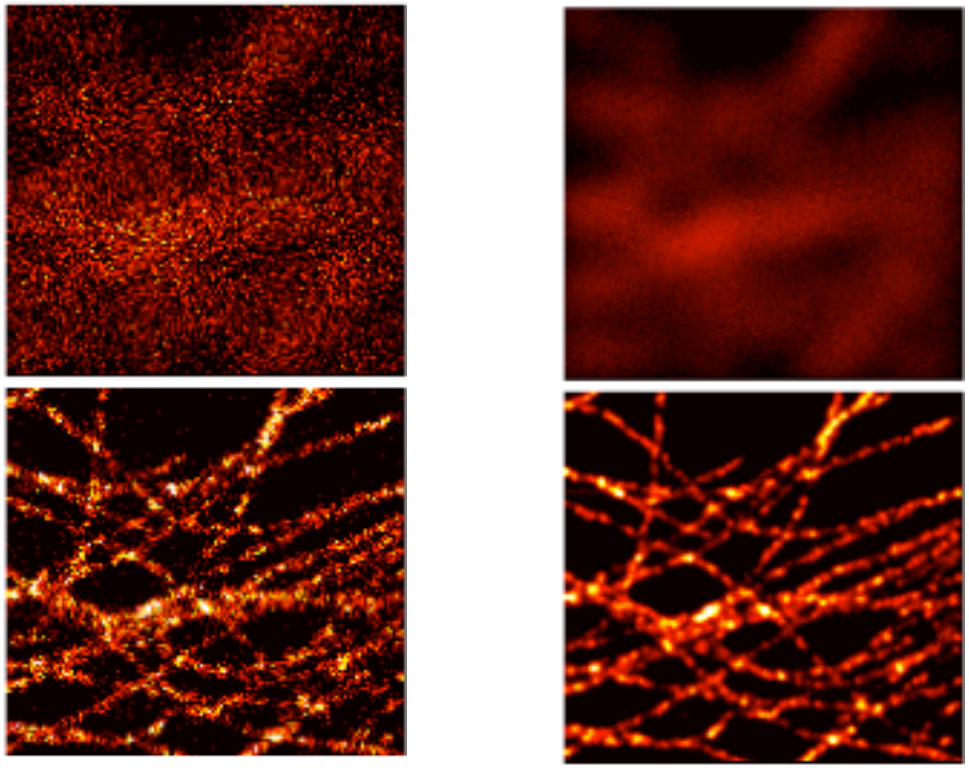


X. Zhuang *et al.* Curr. Op. Chem. Bio. (2008)

Estimates

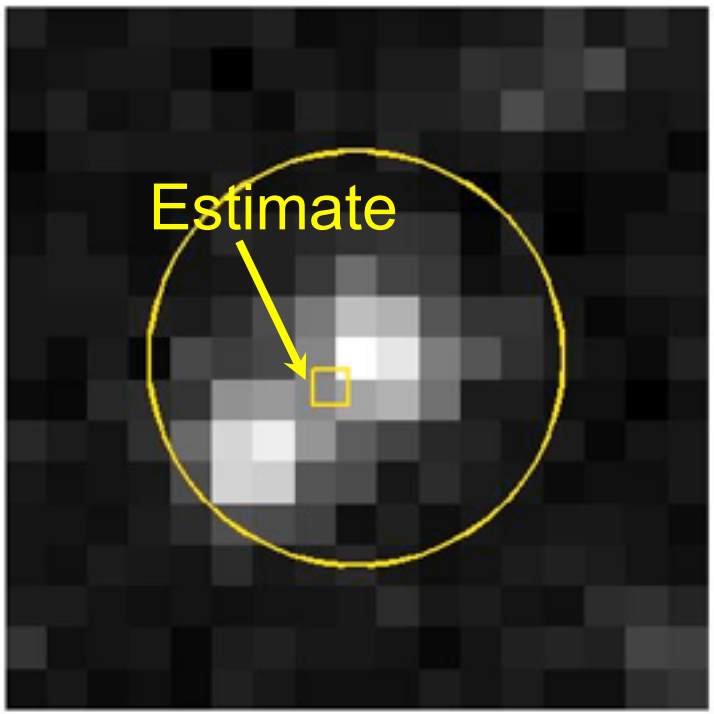
More emitters

More photons/emitter



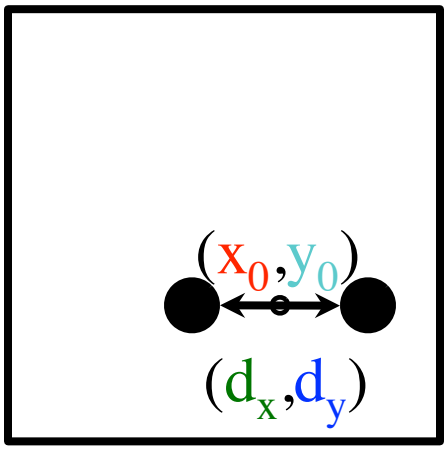
Simultaneously active emitters degrade estimation

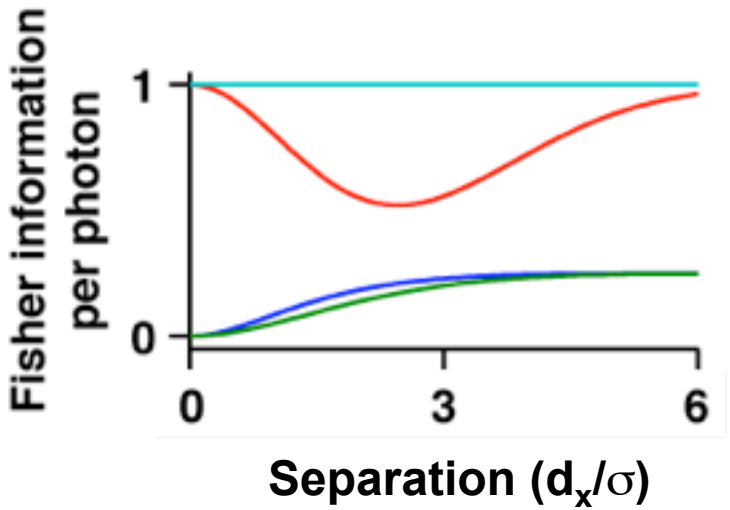
Fluorescence data

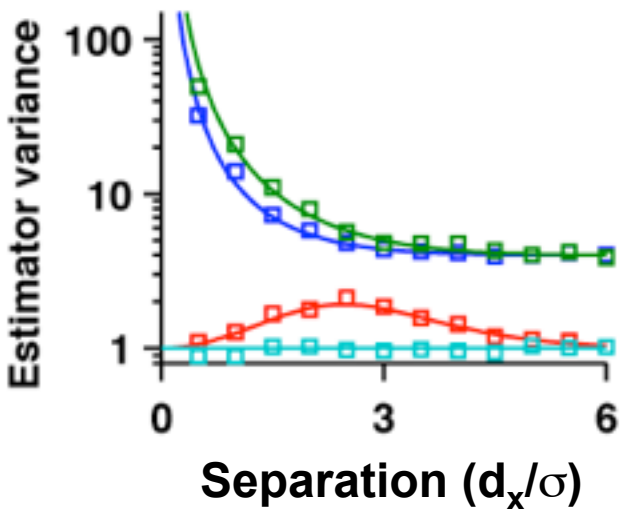


STORM Data courtesy of Babcock, Zhuang et al.

Localization of two emitters depends on their separation

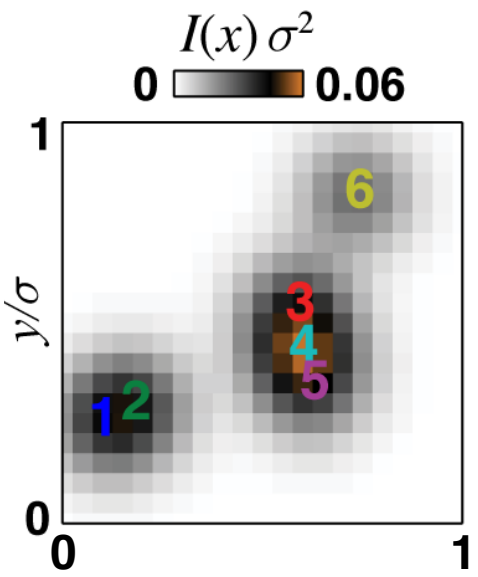




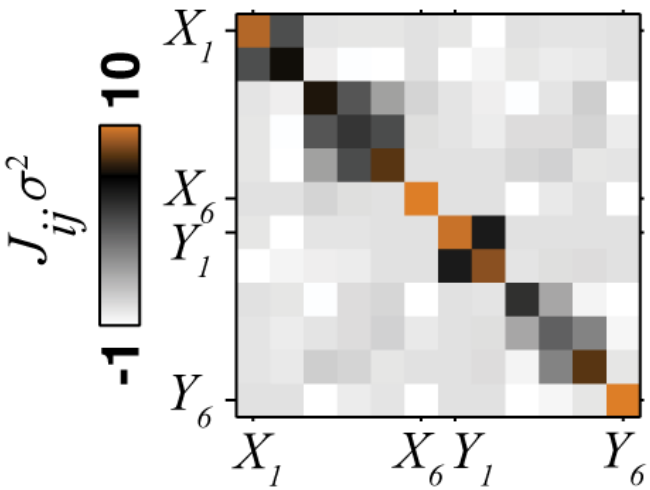


Fisher information bounds estimation for multiple emitters

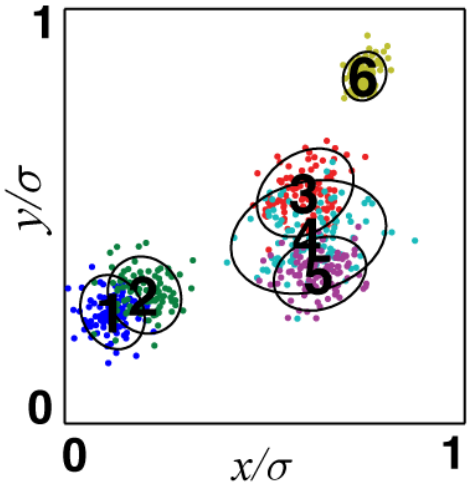
Observed image



Fisher information matrix



Emitter location estimates



- However, Fisher information for localization is inconvenient:
  - A matrix quantity
  - Depends on the configuration of emitters, not just density

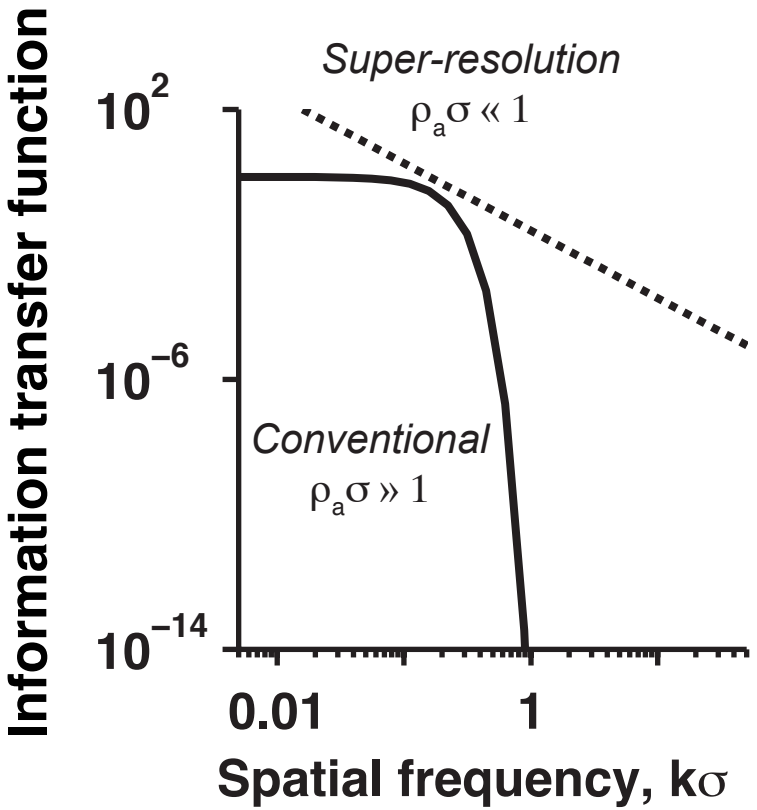
The information transfer function measures resolution for each spatial frequency

Limit 1: Conventional imaging

$$\lim_{\rho_a \sigma \gg 1} F(k) \approx N \exp\left[-\frac{1}{2}(2\pi k\sigma)^2\right]$$

Limit 2: Super-resolution

$$\lim_{\rho_a \sigma \ll 1} F(k) = \frac{N}{(2\pi k\sigma)^2}$$



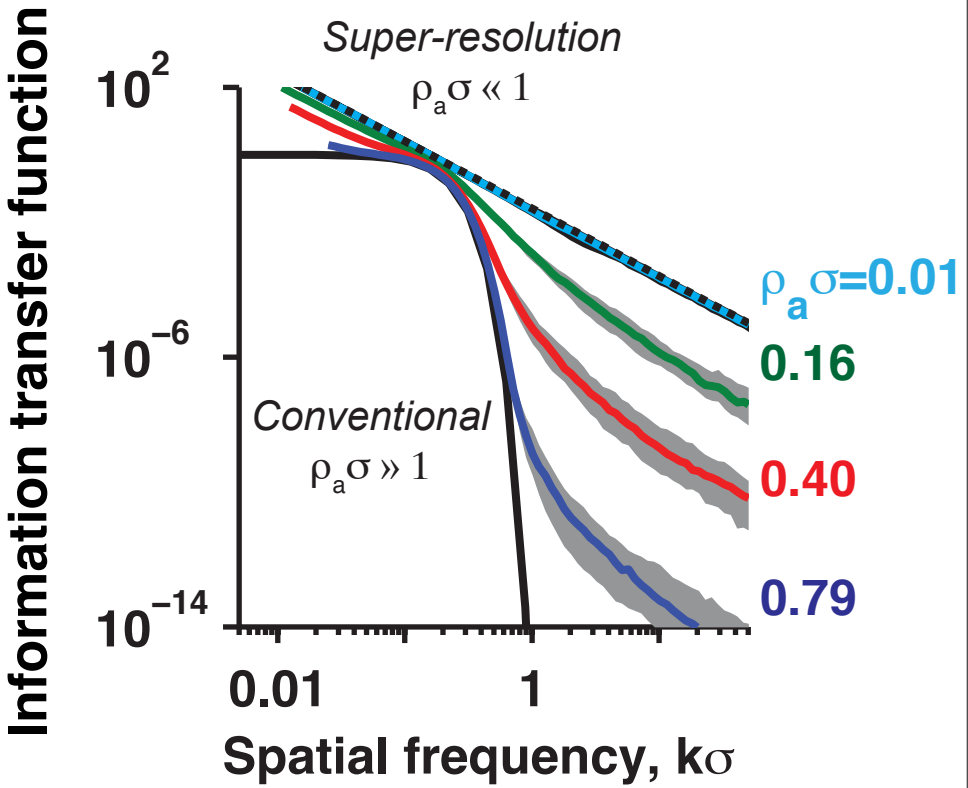
Estimation accuracy for image spatial frequencies depends on density

Limit 1: Conventional imaging

$$\lim_{\rho_a\sigma\gg 1} F(k) \approx N \exp\left[-\frac{1}{2}(2\pi k\sigma)^2\right]$$

Limit 2: Super-resolution

$$\lim_{\rho_a\sigma\ll 1} F(k) = \frac{N}{(2\pi k\sigma)^2}$$

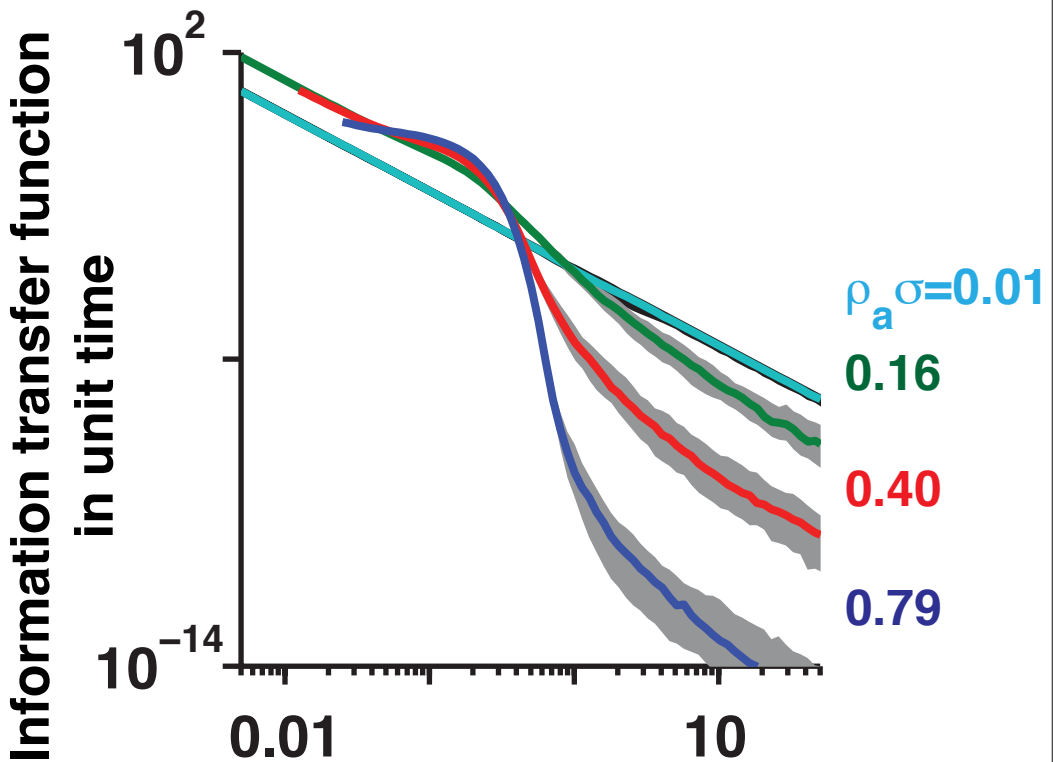


Time is a key constraint for live-cell and *in vivo* imaging

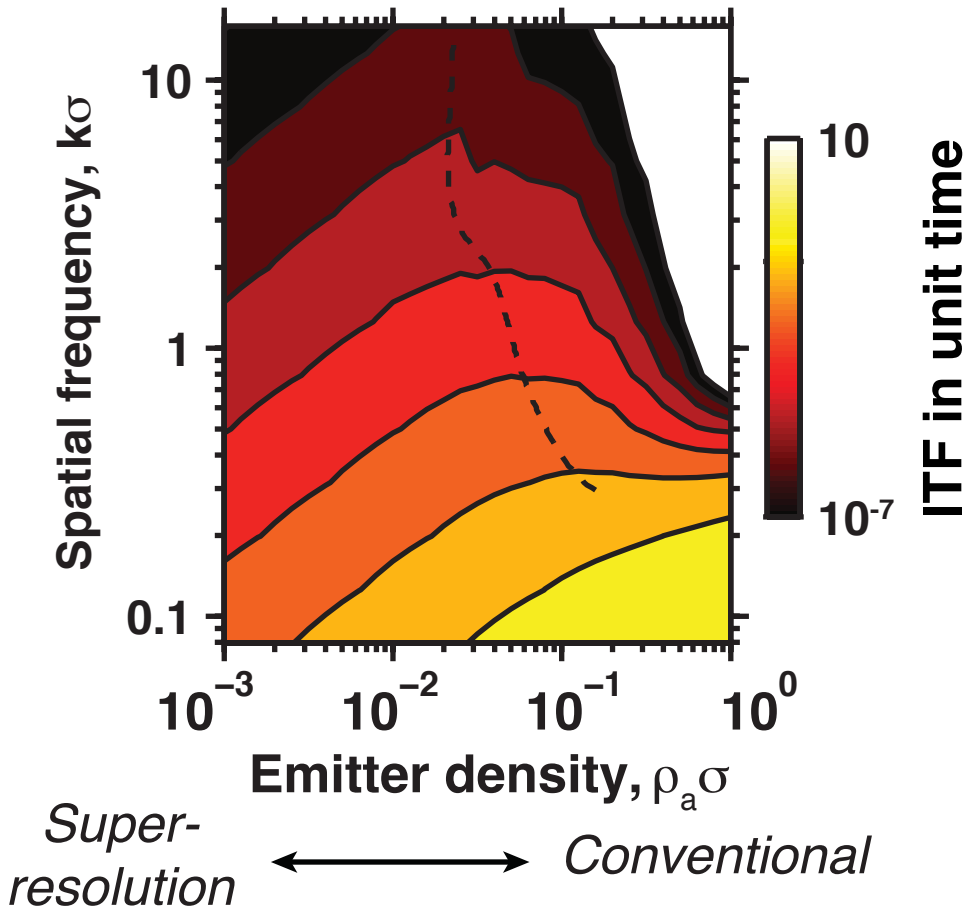
Photons increase with time and emitter density $N \propto \rho_a T$

Define the ITF in unit time:

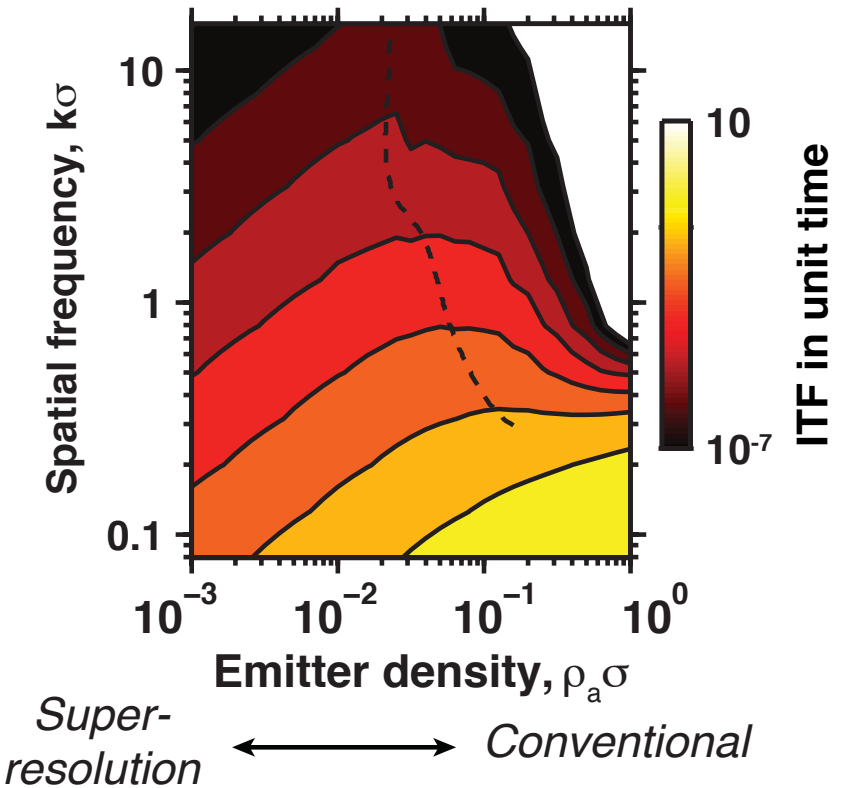
$$F_T(k) = F(k) \times \frac{\rho_a}{N}$$

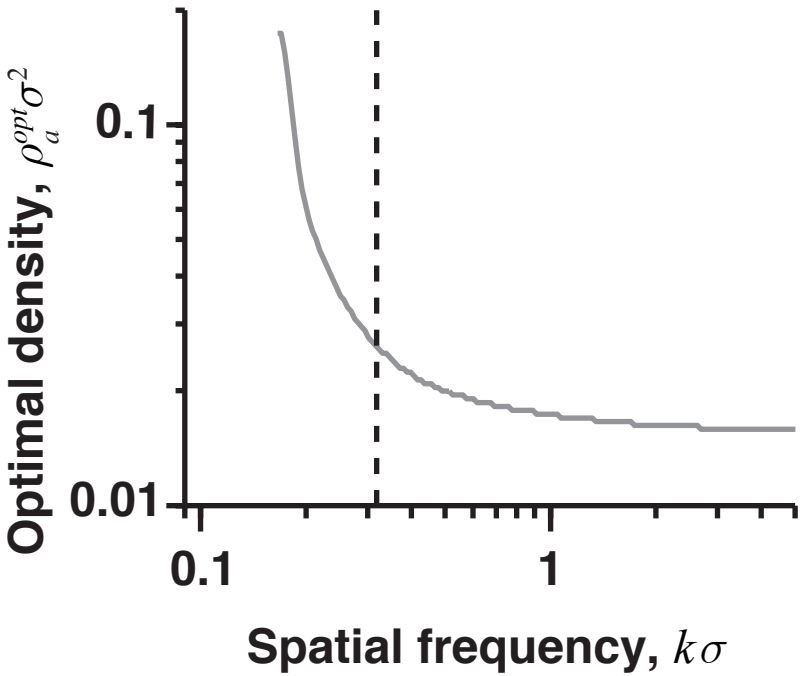


Optimal emitter density depends on spatial frequency



Optimal emitter density depends on spatial frequency





Shaky, Spiky, STORMy & Sharp: Conclusions

Visual acuity:

- Optimal decoder of retinal spiking must account for eye movements.
- A biologically feasible, Reichardt-style decoder achieves near-optimal performance.
- Temporal filtering in the retinal circuit requires lengthening the decoder time constant.

Super-resolution microscopy:

- Estimation accuracy depends on spatial frequency, number of photons, and density of emitters
- Speed-accuracy tradeoff sets the optimal emitter density.

Acknowledgments

Visual hyper-acuity:

- Yoram Burak
- Markus Meister
- Haim Sompolinsky
- Ofer Mazor

Super-resolution microscopy:

- Mark Schnitzer (Stanford)