Statistical inference for molecules: How many, When and Where?

Frank Werner

joint with

Katharina Proksch and Axel Munk

Statistical Inverse Problems in Biophysics Group
Max Planck Institute for Biophysical Chemistry, Göttingen

and

Felix Bernstein Institute for Mathematical Statistics in the Biosciences
University of Göttingen

Statistical inference for molecules:
Where?

Frank Werner

joint with

Katharina Proksch and Axel Munk

Statistical Inverse Problems in Biophysics Group
Max Planck Institute for Biophysical Chemistry, Göttingen

and

Felix Bernstein Institute for Mathematical Statistics in the Biosciences
University of Göttingen

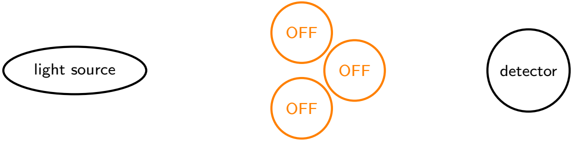
Outline

Outline

Scanning fluorescence microscopy in a nutshell

Pulsed scanning fluorescence microscopy

→ For each scanning position s repeat

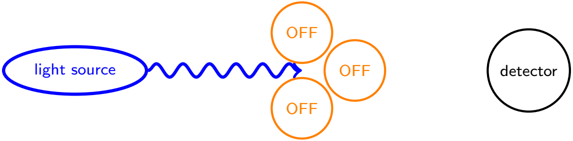


Scanning fluorescence microscopy in a nutshell

Pulsed scanning fluorescence microscopy

→ For each scanning position s repeat

1. Visible laser pulse focused to s (illumination)

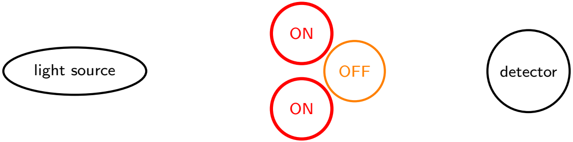


Scanning fluorescence microscopy in a nutshell

Pulsed scanning fluorescence microscopy

→ For each scanning position s repeat

1. Visible laser pulse focused to s (illumination)
2. Light is absorbed by fluorophores (excitation)

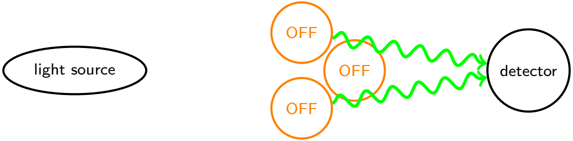


Scanning fluorescence microscopy in a nutshell

Pulsed scanning fluorescence microscopy

→ For each scanning position s repeat

1. Visible laser pulse focused to s (illumination)
2. Light is absorbed by fluorophores (excitation)
3. Fluorophores emit light of different wavelength (emission)

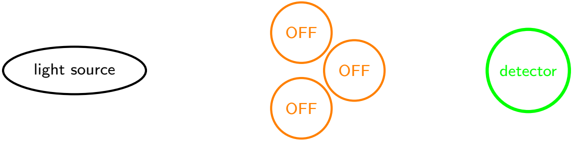


Scanning fluorescence microscopy in a nutshell

Pulsed scanning fluorescence microscopy

→ For each scanning position s repeat

1. Visible laser pulse focused to s (illumination)
2. Light is absorbed by fluorophores (excitation)
3. Fluorophores emit light of different wavelength (emission)
4. The emitted light is recorded (detection)



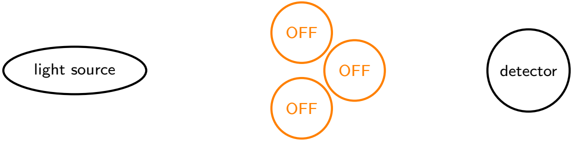
Scanning fluorescence microscopy in a nutshell

Pulsed scanning fluorescence microscopy

→ For each scanning position s repeat

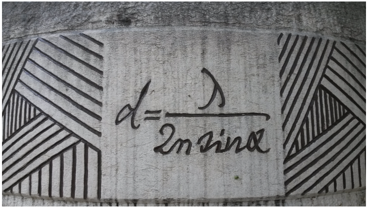
1. Visible laser pulse focused to s (illumination)
2. Light is absorbed by fluorophores (excitation)
3. Fluorophores emit light of different wavelength (emission)
4. The emitted light is recorded (detection)

→ Until the markers start to bleach (say t times)



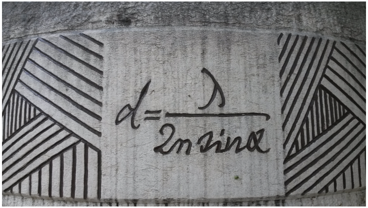
Resolution

- Abbe's diffraction limit: light cannot be focused to a point



Resolution

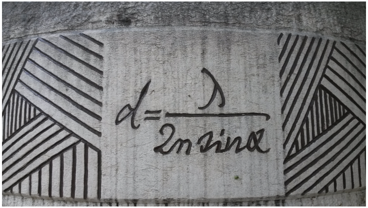
- Abbe's diffraction limit: light cannot be focused to a point



⇝ When scanning at s, also markers at neighboring positions are excited and may emit light

Resolution

- Abbe's diffraction limit: light cannot be focused to a point

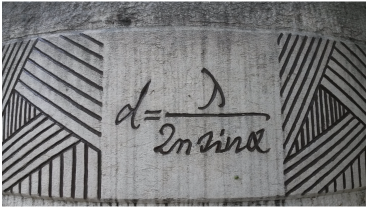


⇝ When scanning at s, also markers at neighboring positions are excited and may emit light

⇝ Recordings are blurred by the so-called **point-spread-function**, which limits the resolution to approximately half the wavelength of the excitation light.

Resolution

- Abbe's diffraction limit: light cannot be focused to a point



⇝ When scanning at s, also markers at neighboring positions are excited and may emit light

⇝ Recordings are blurred by the so-called **point-spread-function**, which limits the resolution to approximately half the wavelength of the excitation light.

- However, super-resolution is possible!

Super-Resolution

The Nobel Prize in Chemistry 2014

Photo: Matt Staley/HHMI

Eric Betzig

Prize share: 1/3

Photo: Wikimedia Commons, CC-BY-SA-3.0

Stefan W. Hell

Prize share: 1/3

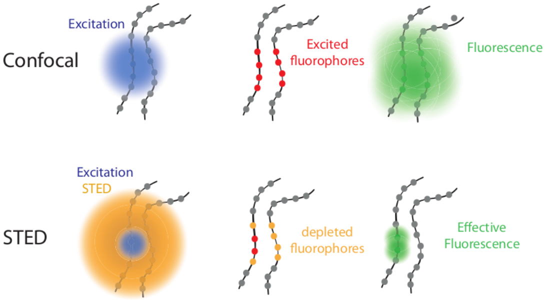
Photo: K. Lowder via Wikimedia Commons, CC-BY-SA-3.0

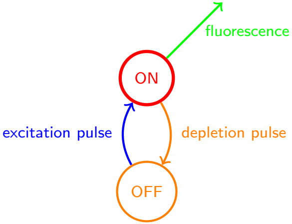
William E. Moerner

Prize share: 1/3

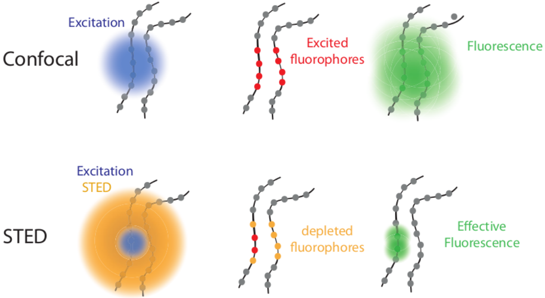
"for the development of super-resolved fluorescence microscopy"

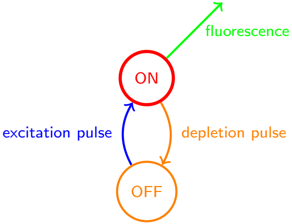
STimulated Emission Depletion (Hell & Wichmann '94)

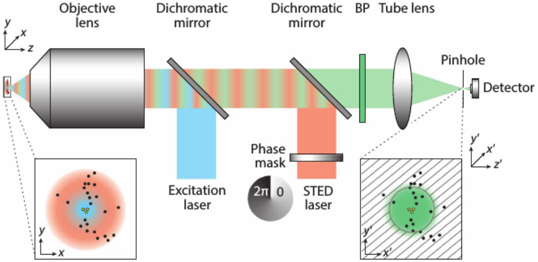




STimulated Emission Depletion (Hell & Wichmann '94)







Timo Aspelmeier, Alexander Egner, and Axel Munk
Modern Statistical Challenges in High-Resolution Fluorescence Microscopy
Ann. Rev. Statist. Appl. 2, 163–202 (2015).

Modeling

- Let k be the psf determined by the effective illumination pattern

Modeling

- Let k be the psf determined by the effective illumination pattern
- Denote by f the (unknown) fluorophore intensity

Modeling

- Let k be the psf determined by the effective illumination pattern
- Denote by f the (unknown) fluorophore intensity
- Observed data $Y(s) \sim \mathrm{Bin}(t, (k * f)(s))$ for any scanning position s,

$$(k * f)(x) = \int k(x - y) f(y) \, \mathrm{d}y$$

(with proper renormalization of k)

Modeling

- Let k be the psf determined by the effective illumination pattern
- Denote by f the (unknown) fluorophore intensity
- Observed data $Y(s) \sim \mathrm{Bin}(t, (k * f)(s))$ for any scanning position s,

$$(k * f)(x) = \int k(x - y) f(y) \, \mathrm{d}y$$

  (with proper renormalization of k)
- Typically $\mathcal{F}k > 0$, so f is uniquely determined by $k * f$

Modeling

- Let k be the psf determined by the effective illumination pattern
- Denote by f the (unknown) fluorophore intensity
- Observed data $Y(s) \sim \mathrm{Bin}(t, (k * f)(s))$ for any scanning position s,

$$(k * f)(x) = \int k(x - y) f(y) \, \mathrm{d}y$$

  (with proper renormalization of k)
- Typically $\mathcal{F}k > 0$, so f is uniquely determined by $k * f$
- But k is also smooth and hence the problem is ill-posed

Outline

What means Where? mathematically?

H_1 and H_2 be Hilbert-Spaces of functions, $T : H_1 \to H_2$ a bounded linear operator and let $f \in H_1$.

Given:

Observations

$$Y_{\mathbf{j}} = Tf(s_{\mathbf{j}}) + \xi_{\mathbf{j}}, \quad \mathbf{j} \in \{1, \dots, n\}^d.$$

- $s_{\mathbf{j}} \in \mathbb{R}^d, \mathbf{j} \in \{1, \dots, n\}^d$, are the sampling points.
- $\xi_{\mathbf{j}}, \mathbf{j} \in \{1, \dots, n\}^d$, are independent, centered random variables.

Aim:

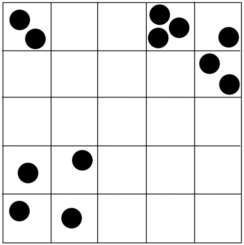
Identify regions B with positive intensity, i.e. $f_{|_B} \not\equiv 0$, at controlled family-wise error rate (FWER).

Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.

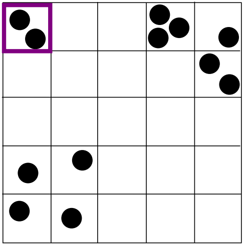
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



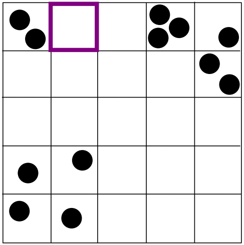
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



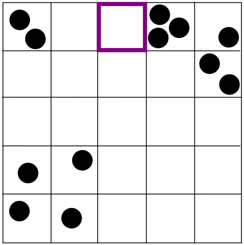
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



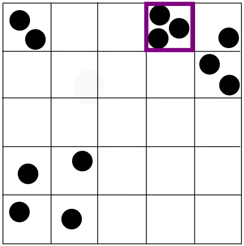
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



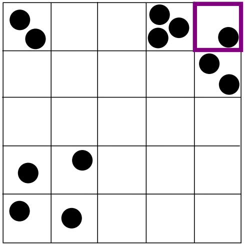
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



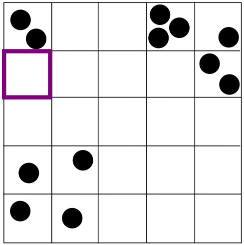
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



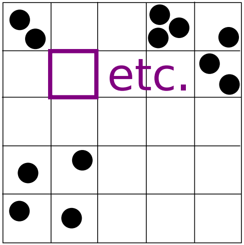
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



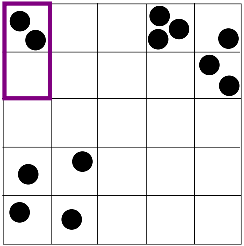
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



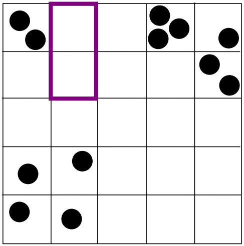
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



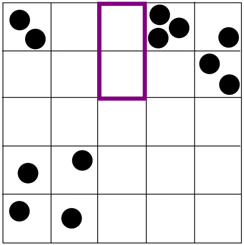
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



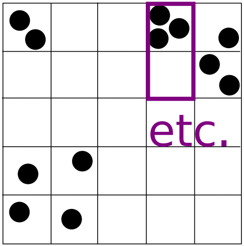
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



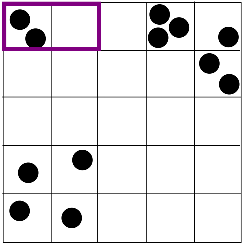
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



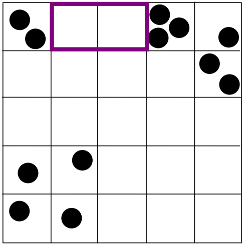
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



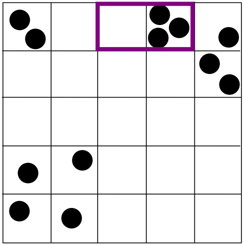
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



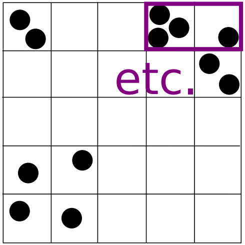
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



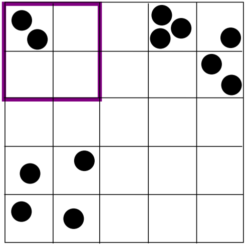
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



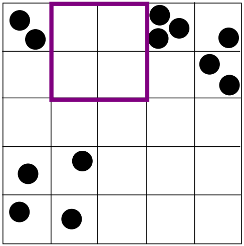
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



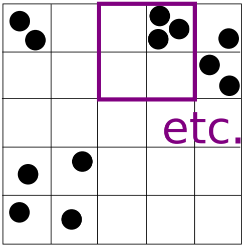
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



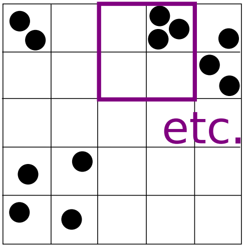
Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



Scanning

- A priori we do not know where markers are, if they cluster, and if so which size the clusters have.
- Therefore we want to **scan** the image with all possible translations of boxes of different sizes (scales):



But: We do not observe f directly, only data related to Tf available.

How to get rid of T?

Scanning in Inverse Problems

- For each box B, choose a suitable function φ_B with $\text{supp}(\varphi_B) \subset B$, $\varphi_B \geq 0$. Then it still holds

$$\langle \varphi_B, f \rangle > 0 \qquad \Rightarrow \qquad f_{|_B} \not\equiv 0$$

Scanning in Inverse Problems

- For each box B, choose a suitable function φ_B with $\operatorname{supp}(\varphi_B) \subset B$, $\varphi_B \geq 0$. Then it still holds

$$\langle \varphi_B, f \rangle > 0 \qquad \Rightarrow \qquad f_{|_B} \not\equiv 0$$

- If $\varphi_B = T^*\Phi_B$, then:

$$\langle \Phi_B, Tf \rangle = \langle T^*\Phi_B, f \rangle = \langle \varphi_B, f \rangle$$

Scanning in Inverse Problems

- For each box B, choose a suitable function φ_B with $\operatorname{supp}(\varphi_B) \subset B$, $\varphi_B \geq 0$. Then it still holds

$$\langle \varphi_B, f \rangle > 0 \qquad \Rightarrow \qquad f_{|_B} \not\equiv 0$$

- If $\varphi_B = T^* \Phi_B$, then:

$$\langle \Phi_B, Tf \rangle = \langle T^* \Phi_B, f \rangle = \langle \varphi_B, f \rangle$$

- The left-hand side can be estimated by $\langle \Phi_B, Y \rangle$ ($\rightsquigarrow$ local test statistic)

Scanning in Inverse Problems

- For each box B, choose a suitable function φ_B with $\operatorname{supp}(\varphi_B) \subset B$, $\varphi_B \geq 0$. Then it still holds

$$\langle \varphi_B, f \rangle > 0 \qquad \Rightarrow \qquad f_{|_B} \not\equiv 0$$

- If $\varphi_B = T^* \Phi_B$, then:

$$\langle \Phi_B, Tf \rangle = \langle T^* \Phi_B, f \rangle = \langle \varphi_B, f \rangle$$

- The left-hand side can be estimated by $\langle \Phi_B, Y \rangle$ ($\rightsquigarrow$ local test statistic)
- Consequently, we **scan over** f by means of $\{\varphi_B\}_B$ **by scanning over** Tf by means of $\{\Phi_B\}_B$

Multiscale Scanning in Inverse Problems

- Let $\mathcal{U} = \{\varphi_{i,n}\}_{1 \leq i \leq N} \subset H_1$ be a dictionary of scanning functions, $\varphi_{i,n} = T^* \Phi_{i,n}$ for all i and n.

Multiscale Scanning in Inverse Problems

- Let $\mathcal{U} = \{\varphi_{i,n}\}_{1 \leq i \leq N} \subset H_1$ be a dictionary of scanning functions, $\varphi_{i,n} = T^* \Phi_{i,n}$ for all i and n.
- Theoretical analysis depends on the "structure" of $\{\Phi_{i,n}\}_{i,n}$.

Multiscale Scanning in Inverse Problems

- Let $\mathcal{U} = \{\varphi_{i,n}\}_{1\leq i\leq N} \subset H_1$ be a dictionary of scanning functions, $\varphi_{i,n} = T^*\Phi_{i,n}$ for all i and n.
- Theoretical analysis depends on the "structure" of $\{\Phi_{i,n}\}_{i,n}$.

In this talk...

... $\mathrm{supp}(\Phi_{i,n}) \subseteq [0,1]^d$ for all i and n, and each index i belongs to position $\mathbf{t}_{i,n} = (t_{i,n,1},\ldots,t_{i,n,d})^T$ and scale $\mathbf{h}_{i,n} = (h_{i,n,1},\ldots,h_{i,n,d})^T$:

$$\Phi_{i,n}(\mathbf{z}) = \Phi_{\mathbf{h}_{i,n}}\left(\frac{\mathbf{t}_{i,n}-\mathbf{z}}{\mathbf{h}_{i,n}}\right).$$

Multiscale Scanning in Inverse Problems

- Let $\mathcal{U} = \{\varphi_{i,n}\}_{1 \leq i \leq N} \subset H_1$ be a dictionary of scanning functions, $\varphi_{i,n} = T^*\Phi_{i,n}$ for all i and n.
- Theoretical analysis depends on the "structure" of $\{\Phi_{i,n}\}_{i,n}$.

In this talk...

... $\mathrm{supp}(\Phi_{i,n}) \subseteq [0,1]^d$ for all i and n, and each index i belongs to position $\mathbf{t}_{i,n} = (t_{i,n,1}, \ldots, t_{i,n,d})^T$ and scale $\mathbf{h}_{i,n} = (h_{i,n,1}, \ldots, h_{i,n,d})^T$:

$$\Phi_{i,n}(\mathbf{z}) = \Phi_{\mathbf{h}_{i,n}}\left(\frac{\mathbf{t}_{i,n} - \mathbf{z}}{\mathbf{h}_{i,n}}\right).$$

... and we consider the multiscale scan statistic

$$T_n(Y) = \max_i \left[w_{i,n} \left(\frac{\langle \Phi_{i,n}, Y \rangle_{H_{2,n}}}{\sqrt{\mathrm{Var}\langle \Phi_{i,n}, Y \rangle_{H_{2,n}}}} - w_{i,n} \right) \right].$$

Outline

Assumptions

- Polynomial growth of the dictionary: For some $\kappa > 0$

$$\left|\{\varphi_{i,n}\}_{1 \leq i \leq N}\right| = N = O(n^{\kappa}) \quad \text{as} \quad n \to \infty.$$

- Scale restrictions:

$$h_{\min} \gtrsim n^{-1} \log(n)^{p} \quad \text{and} \quad h_{\max} = o\big(\log(n)^{-2}\big).$$

- Moment condition: Suppose that Bennett's Moment condition holds and assume further that

$$\max_{j} \mathbb{E}\xi_j^4 = O(1).$$

Gaussian Approximation

Gaussian Approximation (Proksch, W., Munk 2018)

If $\Phi_{i,n}/\|\Phi_{i,n}\|_2$ is uniformly bounded, the moment condition, the polynomial growth of the dictionary and under the scale restrictions, then the $\xi_{\mathbf{j}}$ can be replaced by an i.i.d. field of standard Gaussian rvs $\zeta = (\zeta_{\mathbf{j}})_{\mathbf{j}\in\{1,\dots,n\}^d}$.

More precisely, for

$$M_n(\zeta) = \max_i \left[w_{i,n} \left(\frac{\langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}{\sqrt{\mathrm{Var}\langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}} - w_{i,n} \right) \right]$$

it holds

$$\lim_{n\to\infty} \mathbb{P}_{f=0}\left(T_n(Y) > q_{1-\alpha}^{M_n(\zeta)} \right) \leq \alpha.$$

Victor Chernozhukov, Denis Chetverikov and Kengo Kato.
Gaussian approximation of suprema of empirical processes.
Ann. Statist., 2014.

Limiting distribution (log log-terms are important!)

Under further assumptions on the smoothness γ of $\Phi_{\mathbf{h}_i,n}$, the choice

$$w_{i,n} = \sqrt{2\log(C/\mathbf{h}^{\mathbf{1}}_{i,n})} + C_d \frac{\log(\sqrt{2\log(C/(\mathbf{h}^{\mathbf{1}}_{i,n}))})}{\sqrt{2\log(C/\mathbf{h}^{\mathbf{1}}_{i,n}))}},$$

with an explicit constant C_d depending on d, γ, and the number of scales leads to a Gumbel extreme value limit:

Limiting distribution (Proksch, W., Munk 2018)

For each suitable C there exists $D > 0$ such that for any x it holds

$$\lim_{n\to\infty} \mathbb{P}_{f=0}(M_n(\zeta) \leq x) = e^{-De^{-x}}.$$

Katharina Proksch, Frank Werner and Axel Munk.
Multiscale Scanning in Inverse Problems.
Ann. Statist., to appear.

Asymptotic power

Asymptotic power (Proksch, W., Munk 2018)

Under the above assumptions, the power $\mathbb{P}_f(T_n(Y) > q_{1-\alpha}^{M_n(\zeta)})$ is given by

$$\alpha + (1-\alpha)\,\Psi\left(\min_i \left(\sqrt{2\log\left(\frac{1}{\mathbf{h}_{i,n}^{\mathbf{1}}}\right)} - \frac{\langle \varphi_{i,n}, f\rangle_{H_1}}{\sqrt{\mathbf{Var}\left[\langle \Phi_{i,n}, Y\rangle_{H_{2,n}}\right]}}\right)\right) + o(1)$$

where Ψ is the tail function of the standard normal distribution.

Asymptotic power

Asymptotic power (Proksch, W., Munk 2018)

Under the above assumptions, the power $\mathbb{P}_f(T_n(Y) > q_{1-\alpha}^{M_n(\zeta)})$ is given by

$$\alpha + (1-\alpha)\,\Psi\left(\min_i\left(\sqrt{2\log\left(\frac{1}{\mathbf{h}_{i,n}^{\mathbf{1}}}\right)} - \frac{\langle \varphi_{i,n}, f\rangle_{H_1}}{\sqrt{\mathbf{Var}\left[\langle \Phi_{i,n}, Y\rangle_{H_{2,n}}\right]}}\right)\right) + o(1)$$

where Ψ is the tail function of the standard normal distribution.

Immediate consequences:

- Detection properties depend only on 'local' features $\langle \varphi_{i,n}, f\rangle_{H_1}$ of f

Asymptotic power

Asymptotic power (Proksch, W., Munk 2018)

Under the above assumptions, the power $\mathbb{P}_f(T_n(Y) > q_{1-\alpha}^{M_n(\zeta)})$ is given by

$$\alpha + (1-\alpha)\,\Psi\left(\min_i \left(\sqrt{2\log\left(\frac{1}{\mathbf{h}_{i,n}^{1}}\right)} - \frac{\langle \varphi_{i,n}, f\rangle_{H_1}}{\sqrt{\mathbf{Var}\left[\langle \Phi_{i,n}, Y\rangle_{H_{2,n}}\right]}}\right)\right) + o(1)$$

where Ψ is the tail function of the standard normal distribution.

Immediate consequences:

- Detection properties depend only on 'local' features $\langle \varphi_{i,n}, f\rangle_{H_1}$ of f
- When trying to detect a hypercube in a direct problem with i.i.d. Gaussian noise, the detection boundary is resembled

Asymptotic power

Asymptotic power (Proksch, W., Munk 2018)

Under the above assumptions, the power $\mathbb{P}_f(T_n(Y) > q_{1-\alpha}^{M_n(\zeta)})$ is given by

$$\alpha + (1-\alpha)\,\Psi\left(\min_i \left(\sqrt{2\log\left(\frac{1}{\mathbf{h}_{i,n}^1}\right)} - \frac{\langle \varphi_{i,n}, f\rangle_{H_1}}{\sqrt{\mathbf{Var}\left[\langle \Phi_{i,n}, Y\rangle_{H_{2,n}}\right]}}\right)\right) + o(1)$$

where Ψ is the tail function of the standard normal distribution.

Immediate consequences:

- Detection properties depend only on 'local' features $\langle \varphi_{i,n}, f\rangle_{H_1}$ of f
- When trying to detect a hypercube in a direct problem with i.i.d. Gaussian noise, the detection boundary is resembled
- Oracle optimality: the power does not improve by scanning only over the 'correct' **h** (asymptotically)

Outline

How to apply the method in practice

- Choose dictionary functions $\varphi_{i,n} \in R(T^*)$, compute $\Phi_{i,n}$

How to apply the method in practice

- Choose dictionary functions $\varphi_{i,n} \in R(T^*)$, compute $\Phi_{i,n}$
- Approximate the distribution of

$$M_n(\zeta) = \max_i \left[w_{i,n} \left(\frac{\langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}{\sqrt{\mathrm{Var} \langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}} - w_{i,n} \right) \right]$$

by Monte Carlo simulations to obtain the quantile $q_{1-\alpha}^{M_n(\zeta)}$

How to apply the method in practice

- Choose dictionary functions $\varphi_{i,n} \in R(T^*)$, compute $\Phi_{i,n}$
- Approximate the distribution of

$$M_n(\zeta) = \max_i \left[w_{i,n} \left(\frac{\langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}{\sqrt{\mathrm{Var} \langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}} - w_{i,n} \right) \right]$$

  by Monte Carlo simulations to obtain the quantile $q_{1-\alpha}^{M_n(\zeta)}$
- For each i compute $T_i(Y) := \langle \Phi_{i,n}, Y \rangle_{H_2,n}$

How to apply the method in practice

- Choose dictionary functions $\varphi_{i,n} \in R(T^*)$, compute $\Phi_{i,n}$
- Approximate the distribution of

$$M_n(\zeta) = \max_i \left[w_{i,n} \left(\frac{\langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}{\sqrt{\mathrm{Var} \langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}} - w_{i,n} \right) \right]$$

  by Monte Carlo simulations to obtain the quantile $q_{1-\alpha}^{M_n(\zeta)}$
- For each i compute $T_i(Y) := \langle \Phi_{i,n}, Y \rangle_{H_2,n}$
- Mark each i with

$$T_i(Y) > \left(\frac{q_{1-\alpha}^{M_n(\zeta)}}{w_{i,n}} + w_{i,n} \right) \sqrt{\mathbf{Var}[T_i(Y)]}$$

  as active

How to apply the method in practice

- Choose dictionary functions $\varphi_{i,n} \in R(T^*)$, compute $\Phi_{i,n}$
- Approximate the distribution of

$$M_n(\zeta) = \max_i \left[w_{i,n} \left(\frac{\langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}{\sqrt{\mathrm{Var} \langle \Phi_{i,n}, \zeta \rangle_{H_2,n}}} - w_{i,n} \right) \right]$$

  by Monte Carlo simulations to obtain the quantile $q_{1-\alpha}^{M_n(\zeta)}$
- For each i compute $T_i(Y) := \langle \Phi_{i,n}, Y \rangle_{H_2,n}$
- Mark each i with

$$T_i(Y) > \left(\frac{q_{1-\alpha}^{M_n(\zeta)}}{w_{i,n}} + w_{i,n} \right) \sqrt{\mathbf{Var}\left[T_i(Y)\right]}$$

  as active

$\rightsquigarrow$ due to FWER control, all active i are 'correct' with prob. $\geq 1 - \alpha$.

Computational issues

- As long as T and $\{\Phi_{i,n}\}_{i,n}$ are fixed, $q_{1-\alpha}^{M_n(\zeta)}$ can be precomputed.

Computational issues

- As long as T and $\{\Phi_{i,n}\}_{i,n}$ are fixed, $q_{1-\alpha}^{M_n(\zeta)}$ can be precomputed.

$\rightsquigarrow$ Applying MISCAT requires only a single evaluation of all $T_i(Y)$

Computational issues

- As long as T and $\{\Phi_{i,n}\}_{i,n}$ are fixed, $q_{1-\alpha}^{M_n(\zeta)}$ can be precomputed.

$\rightsquigarrow$ Applying MISCAT requires only a single evaluation of all $T_i(Y)$

- For fixed scale $\mathbf{h}_i = \mathbf{h}$, all corresponding translations can be computed using the convolution theorem:

$$(T_i(Y))_{\mathbf{h}_i=\mathbf{h}} = \mathcal{F}^{-1}(\mathcal{F}(Y)\cdot\mathcal{F}(\Phi_{\mathbf{h}})).$$

Computational issues

- As long as T and $\{\Phi_{i,n}\}_{i,n}$ are fixed, $q_{1-\alpha}^{M_n(\zeta)}$ can be precomputed.

$\rightsquigarrow$ Applying MISCAT requires only a single evaluation of all $T_i(Y)$

- For fixed scale $\mathbf{h}_i = \mathbf{h}$, all corresponding translations can be computed using the convolution theorem:

$$(T_i(Y))_{\mathbf{h}_i=\mathbf{h}} = \mathcal{F}^{-1}\left(\mathcal{F}(Y) \cdot \mathcal{F}(\Phi_{\mathbf{h}})\right).$$

- Consequently, MISCAT can be applied in

$$\mathcal{O}\left(\#\text{scales} \cdot \#\text{pixels} \log\left(\#\text{pixels}\right)\right)$$

Computational issues

- As long as T and $\{\Phi_{i,n}\}_{i,n}$ are fixed, $q_{1-\alpha}^{M_n(\zeta)}$ can be precomputed.

$\rightsquigarrow$ Applying MISCAT requires only a single evaluation of all $T_i(Y)$

- For fixed scale $\mathbf{h}_i = \mathbf{h}$, all corresponding translations can be computed using the convolution theorem:

$$(T_i(Y))_{\mathbf{h}_i=\mathbf{h}} = \mathcal{F}^{-1}\left(\mathcal{F}(Y) \cdot \mathcal{F}(\Phi_{\mathbf{h}})\right).$$

- Consequently, MISCAT can be applied in

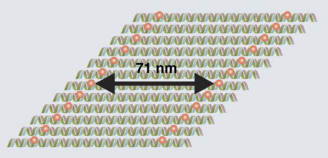
$$\mathcal{O}\left(\#\text{scales} \cdot \#\text{pixels} \log\left(\#\text{pixels}\right)\right)$$

- In the following examples we apply MISCAT for a dictionary with 28.100.601 elements within less than 20 seconds on a standard laptop (with precomputed $q_{1-\alpha}^{M_n(\zeta)}$).

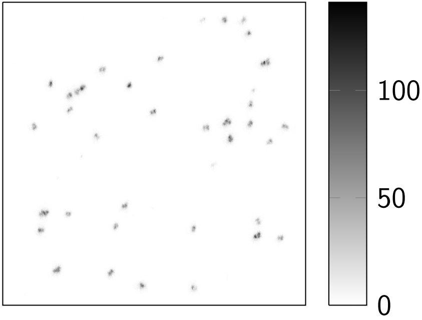
Outline

Real data example - Setup

- we analyze fluorescent dyes on single DNA Origami
- STED measurements
- each of the two strands can at most hold 12 markers



Observations (photon counts)



Data kindly provided by Haisen Ta, Hell Lab, Max Planck Institute for Biophysical Chemistry

Modeling (recap)

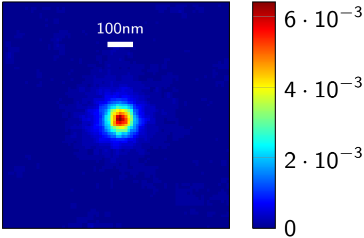
The observations can be modeled by

$$Y_{\mathbf{j}} \overset{\text{independent}}{\sim} \operatorname{Bin}\left(t, (f * k)\left(s_{\mathbf{j}}\right)\right), \qquad \mathbf{j} \in \{1, ..., n\}^2 .$$

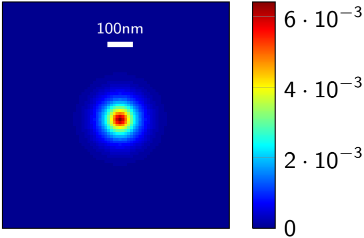
- $f(s)$ fluorophore intensity at s
- if $c \leq k * f \leq C$ for some constants $C, c > 0$, all assumptions on the noise are satisfied
- if the Fourier coefficients of k decay polynomially, then desired functions φ_B exist
- we approximate k in the family

$$(\mathcal{F} k_{a,b})(\xi) = (1 + b^2 \|\xi\|_2^2)^{-a}, \qquad \xi \in \mathbb{R}^2.$$

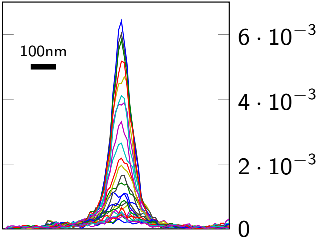
Modeling (cont')



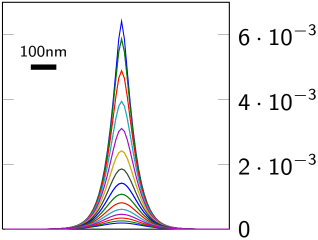
experimental kernel (top view)



$k_{2,0.016}$ (top view)

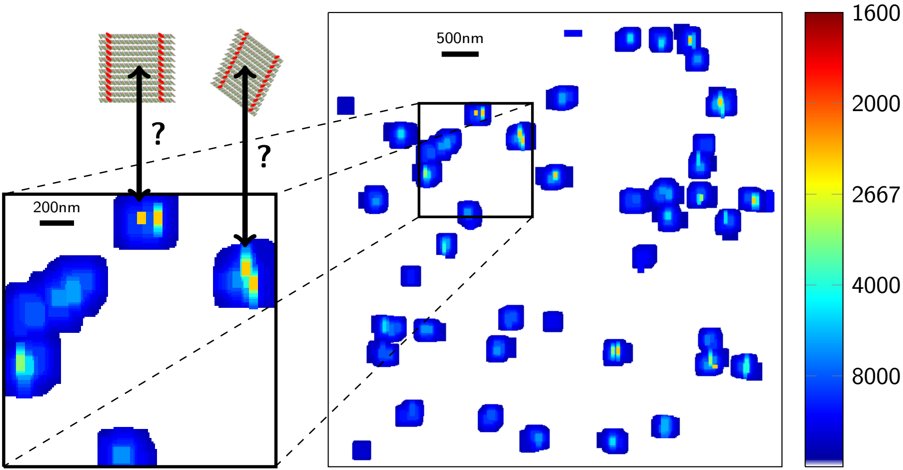


experimental kernel (slices)

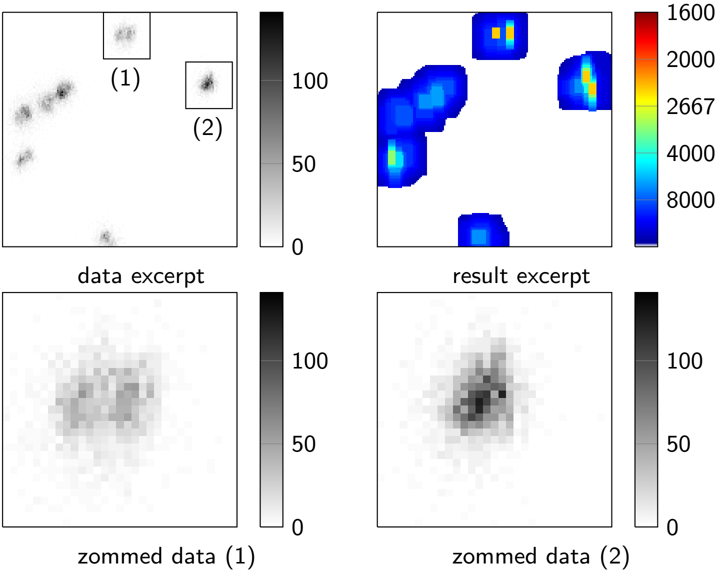


$k_{2,0.016}$ (slices)

Result



Comparison of the result with the data



Outline

Conclusion and outlook

- Methodology
    - inference on the support w.r.t. a dictionary via testing at controlled family-wise error rate
    - general limit theory, asymptotic expansion of the power

Conclusion and outlook

- Methodology
  - inference on the support w.r.t. a dictionary via testing at controlled family-wise error rate
  - general limit theory, asymptotic expansion of the power
- Statistical inference for molecules: Where?
  - application to a deconvolution problem in super-resolution microscopy
  - good performance in real data examples

Conclusion and outlook

- Methodology
  - inference on the support w.r.t. a dictionary via testing at controlled family-wise error rate
  - general limit theory, asymptotic expansion of the power
- Statistical inference for molecules: Where?
  - application to a deconvolution problem in super-resolution microscopy
  - good performance in real data examples
- Not shown here: How many?
  - estimation of molecule numbers inside a segment by additional measurements
  - asymptotic confidence intervals by central limit theorem

  $\rightsquigarrow$ construction of molecular map with controlled uncertainty

Conclusion and outlook

- Methodology
  - inference on the support w.r.t. a dictionary via testing at controlled family-wise error rate
  - general limit theory, asymptotic expansion of the power
- Statistical inference for molecules: Where?
  - application to a deconvolution problem in super-resolution microscopy
  - good performance in real data examples
- Not shown here: How many?
  - estimation of molecule numbers inside a segment by additional measurements
  - asymptotic confidence intervals by central limit theorem

  $\rightsquigarrow$ construction of molecular map with controlled uncertainty

Thank you for your attention!