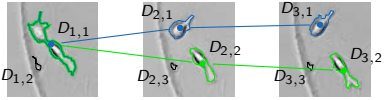
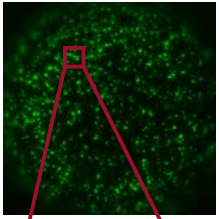
Cell Detection by Inverse Diffusion and Non-negative Group Sparsity

One World SP Seminar Series, June 26, 2020

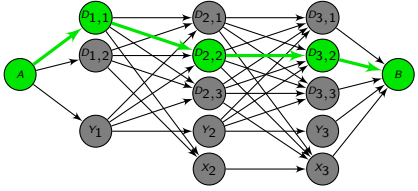
Joakim Jaldén, jalden@kth.se

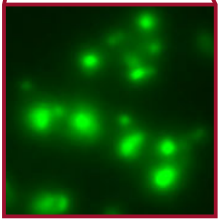
Background

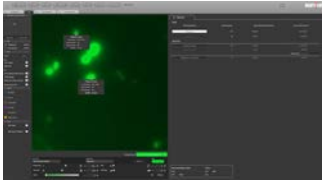




2011 2015 2018







Acknowledgements

Pol del Aguila Pla

Klas Magnusson

MABTECH

qamcom

Vetenskapsrådet

Immunoassays 101

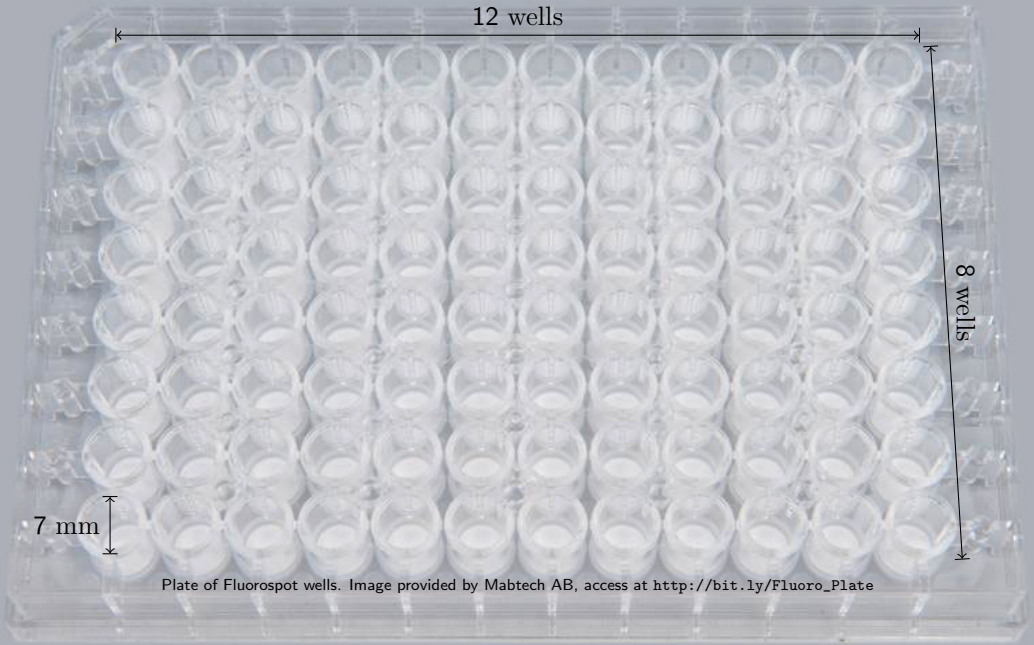
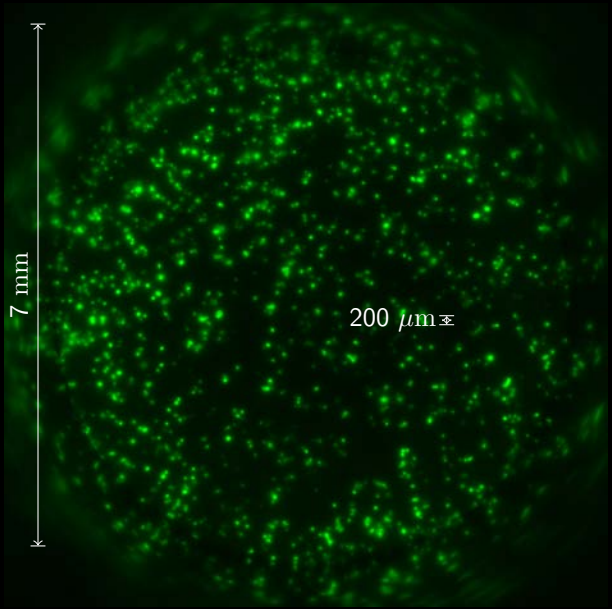
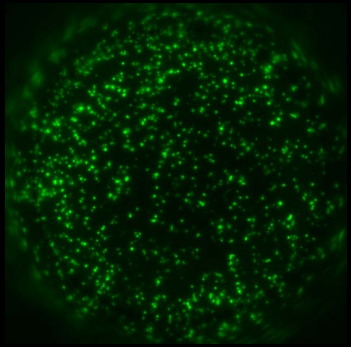


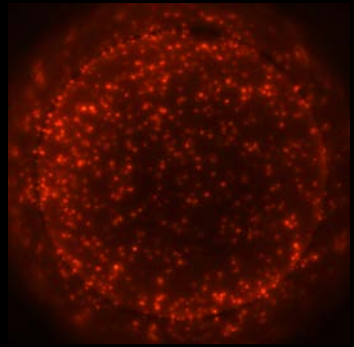
Plate of Fluorospot wells. Image provided by Mabtech AB, access at `http://bit.ly/Fluoro_Plate`



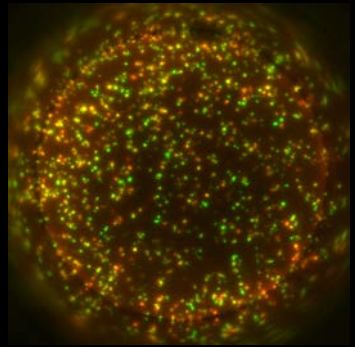
Fluorospot image, provided by Mabtech AB



Fluorescein Isothiocyanate (FITC)

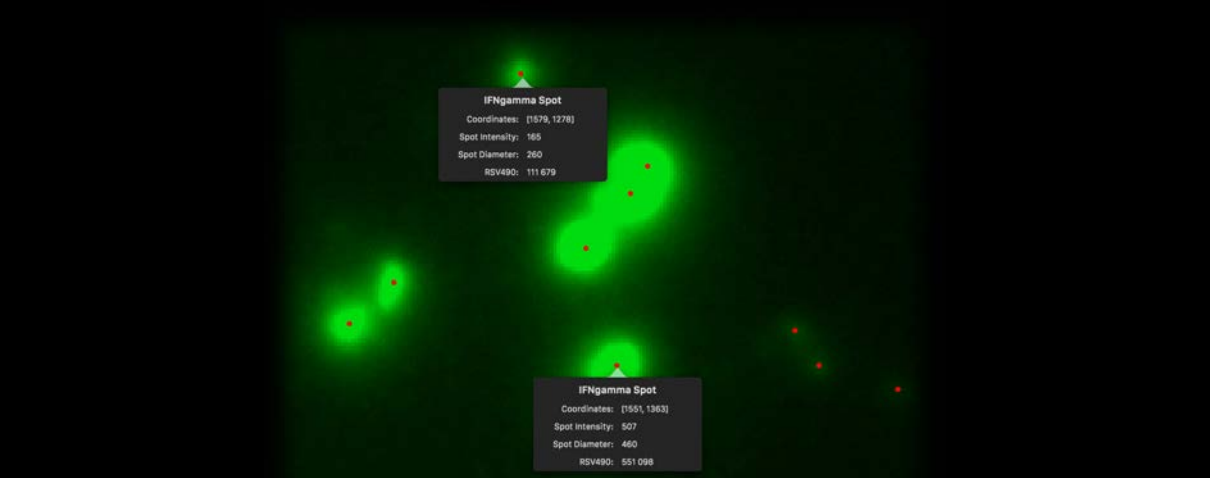


Cyanine 3 (Cy3)



FITC+Cy3

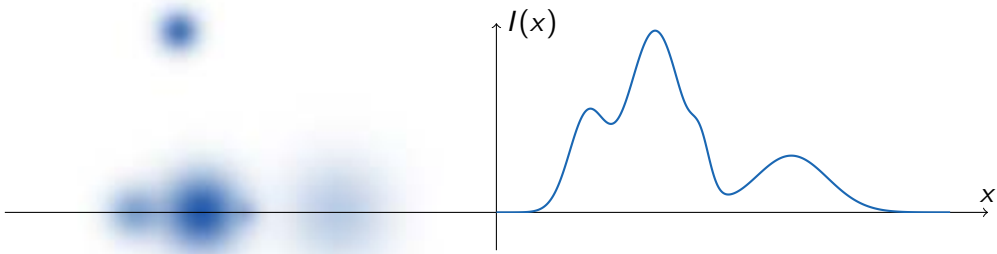
FluoroSpot Multiplex Assay

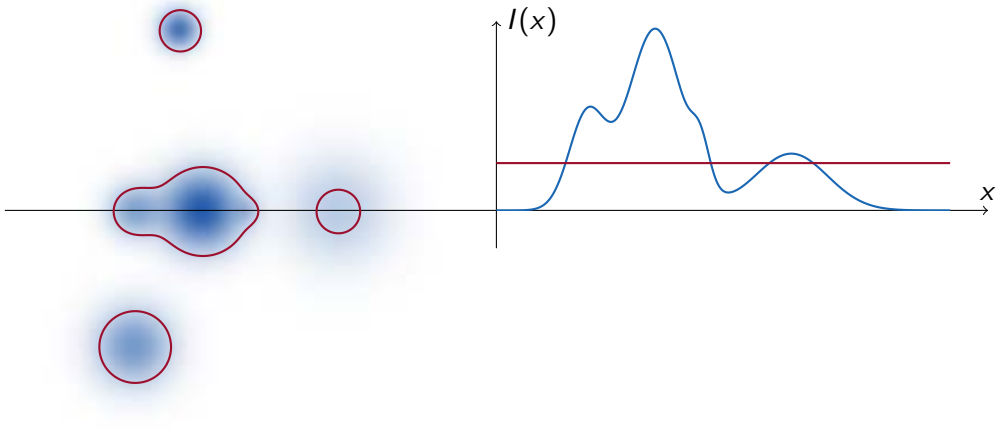


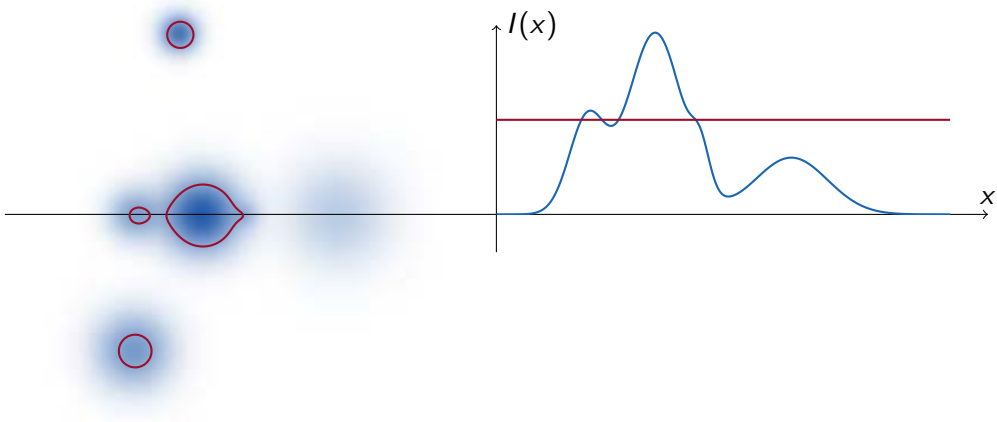
Data Analysis and Labelling

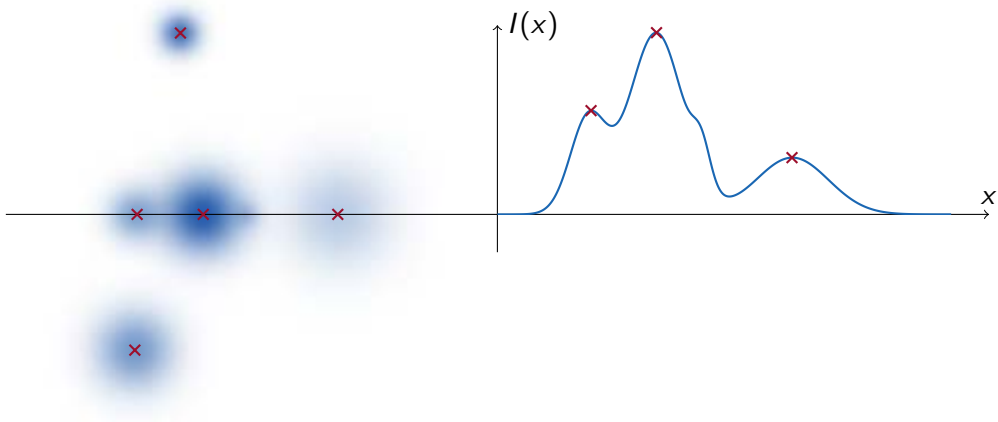
(IFNγ, or type II interferon, is a cytokine that is critical for innate and adaptive immunity against viral, some bacterial and protozoal infections - Wikipedia)

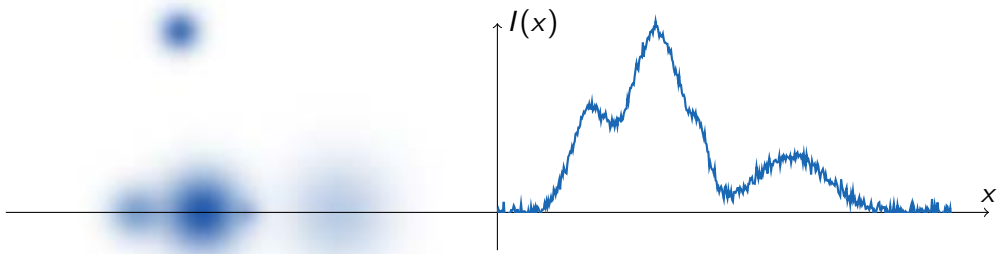
Why is the Problem Challenging?


I(x)
x


I(x)
x


I(x)
x


I(x)
x


I(x)
x

Challenges for spot detection in immunoassays

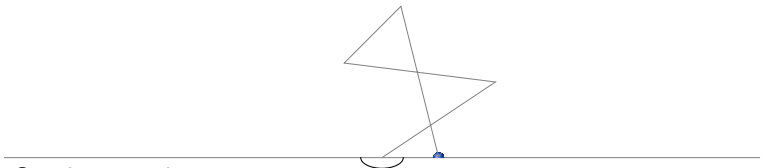
- Spots vary in size, shape, and intensity
- Image noise adversely affects performance
- Strong spots may partially occlude or mask weak spots

The proposed solution

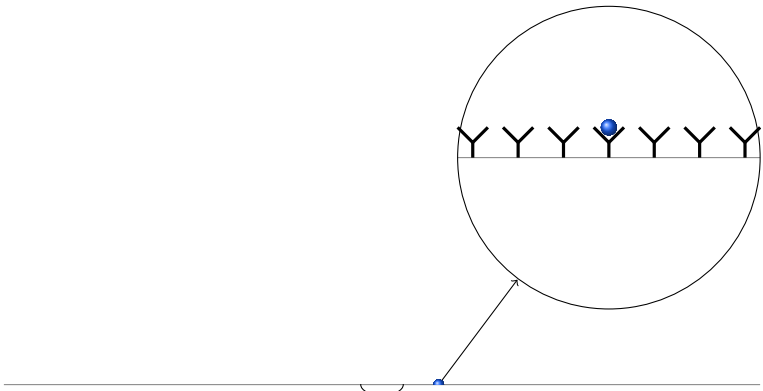
- A model based source point localization algorithm
  - Relying on a physically motivated observation model
  - An inverse problem formulation for source localization
  - Large scale numerical optimization

Spot Formation

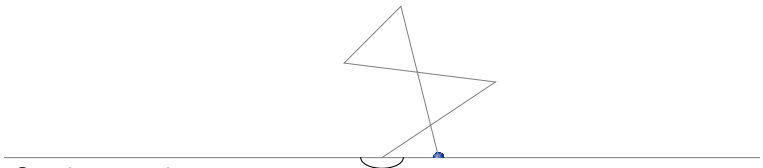
Single particle



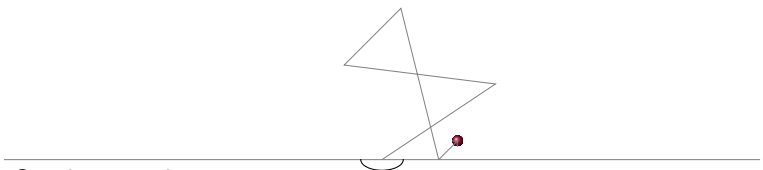
Single particle



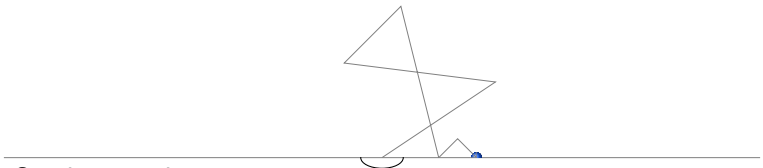
Single particle



Single particle

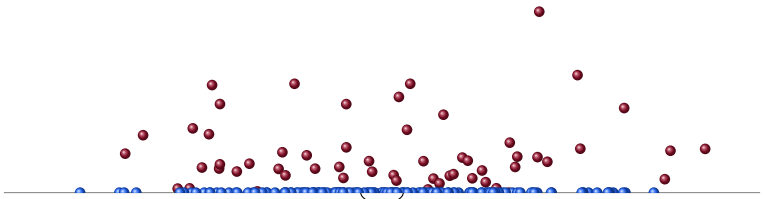


Single particle



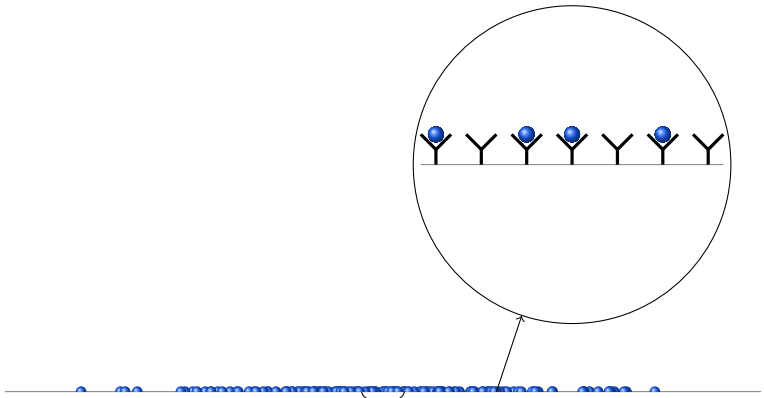
Single particle

Multiple particles

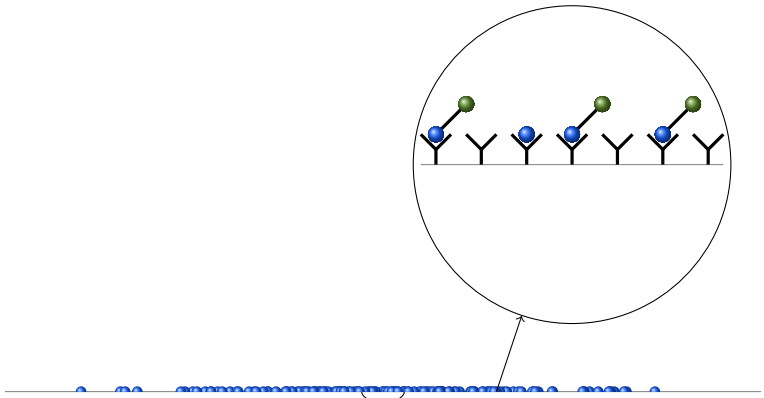


Multiple particles

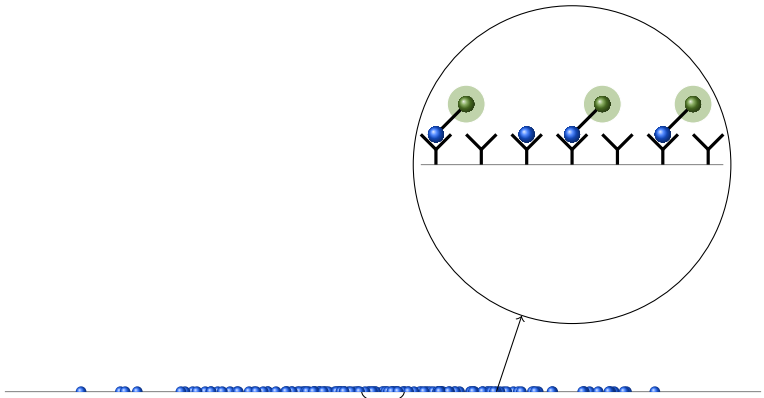
Multiple particles



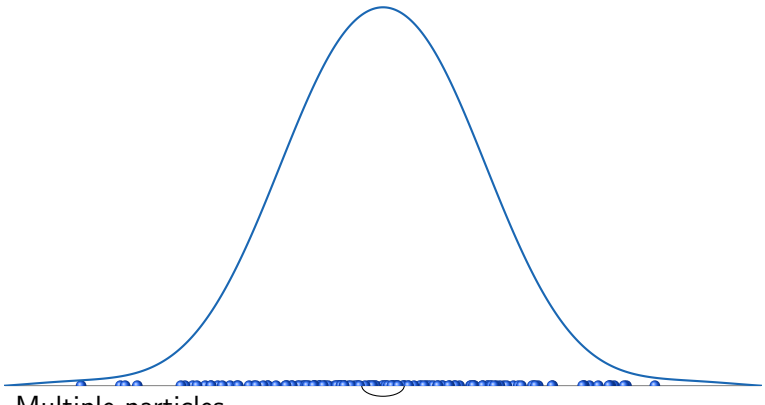
Multiple particles



Multiple particles



Multiple particles



Multiple particles

Spot characteristics

The shape of any particular spot depends on

- when and how many particles are released during incubation
- the diffusivity of the liquid medium
- the capture affinity of the antibodies
- the disassociation probability of the antibodies
- . . . and fluorescence strength, optics, camera exposure, etc.

Observation Modeling

A Physical Model for Biomedical Assays

Relevant quantities for the assay are

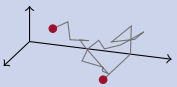
- the density of bound particles $d(x, y, t) \geq 0$, where the image will be (proportional to) $d_{\text{obs}}(x, y) = d(x, y, T)$, which evolves coupled to
- the 3D density of free particles $c(x, y, z, t) \geq 0$ on $z \geq 0$, and to
- the source density rate of new particles $s(x, y, t) \geq 0$, that is spatially sparse and reveals the cell locations and characterization.

A Physical Model for Biomedical Assays

Relevant quantities for the assay are

- the density of bound particles $d(x, y, t) \geq 0$, where the image will be (proportional to) $d_{\mathrm{obs}}(x, y) = d(x, y, T)$, which evolves coupled to
- the 3D density of free particles $c(x, y, z, t) \geq 0$ on $z \geq 0$, and to
- the source density rate of new particles $s(x, y, t) \geq 0$, that is spatially sparse and reveals the cell locations and characterization.

$$\frac{\partial}{\partial t} c = D \Delta c \,,$$



$$\frac{\partial}{\partial t} d = \kappa_{\mathrm{a}} c\big|_{z=0} - \kappa_{\mathrm{d}} d \,,$$

$$-D \frac{\partial}{\partial z} c\big|_{z=0} = s - \frac{\partial d}{\partial t} \,.$$

This physical model was presented before, also for ELISPOT and Fluorospot.

An Observation Model for Biomedical Assays

We consider the image observation $d_{\text{obs}} \in \mathcal{D}_+$, with $\mathcal{D}$ a weighted $\mathrm{L}^2\left(\mathbb{R}^2\right)$ space and prove that

$$d_{\text{obs}}(x,y) = \int_0^{\sigma_{\max}} \left(g_\sigma(\bar{x},\bar{y}) * a(\bar{x},\bar{y},\sigma)\right)(x,y)\,\mathrm{d}\sigma = \int_0^{\sigma_{\max}} G_\sigma a_\sigma \,\mathrm{d}\sigma\,,$$

with $a \in \mathcal{A}_+$ and $\mathcal{A}$ of bounded support μ on $\mathrm{L}^2\left(\mathbb{R}^2 \times \mathbb{R}_+\right)$ space, where $g_\sigma(x,y)$ is a two dimensional Gaussian kernel of width $\sigma \geq 0$, where $\sigma_{\max} = \sqrt{2DT}$, and where

$$a(x,y,\sigma) = \frac{\sigma}{D} \int_{\frac{\sigma^2}{2D}}^{T} s(x,y,T-\eta)\,\varphi\left(\frac{\sigma^2}{2D},\eta\right)\mathrm{d}\eta\,,$$

where $\varphi(\tau,t)$ expresses the probability density for time in free motion.

Theorem

The distribution of time in free motion $\varphi(\tau, t)$, is given by

$$\varphi(\tau, t) = i_{[0,t)}(\tau) \sum_{j=1}^{\infty} \phi^{j*}(\tau) p\left[j-1; \kappa_{\mathrm{d}}(t-\tau)\right] ,$$

where $\phi^{j*}(\tau)$ is the j-th convolutional power of

$$\phi(\tau) = \frac{\kappa_{\mathrm{a}}}{\sqrt{\pi D \tau}} - \frac{\kappa_{\mathrm{a}}^2}{D} \mathrm{erfcx}\left(\kappa_{\mathrm{a}} \sqrt{\frac{\tau}{D}}\right) ,$$

where

$$\mathrm{erfcx}(x) = e^{x^2} \mathrm{erfc}(x) , \quad \mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} \mathrm{d}t ,$$

and where $p[j; \lambda]$ is the Possion distribution with intensity λ.

An Observation Model for Biomedical Assays

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ is

$$d_{\mathrm{obs}} = \int_0^{\sigma_{\max}} G_\sigma a_\sigma \mathrm{d}\sigma, \text{ with } a(x,y,\sigma) = \frac{\sigma}{D} \int_{\frac{\sigma^2}{2D}}^{T} s(x,y,T-\eta)\, \varphi\left(\frac{\sigma^2}{2D}, \eta\right) \mathrm{d}\eta\,.$$

How?

- Independence of Brownian motion in x, y and z.
- Adsorption (κ_{a}) and desorption (κ_{d}) only regulated by z-movement.
- x- and y-movements only depend on τ, total time in Brownian motion. In particular, according to Green function for 2D diffusion, $g_{\sqrt{2D\tau}}(x,y)$.
- $\varphi(\tau, t)$ summarizes the effect of adsorption and desorption onto the time in free motion τ for each time of final adsorption t.
- Change variables to those significative to x- and y-movement, $\sigma = \sqrt{2D\tau}$.

Crucial Observations

- The mapping $s \to a$ given by

$$a(x, y, \sigma) = \frac{\sigma}{D} \int_{\frac{\sigma^2}{2D}}^{T} s(x, y, T - \eta)\, \varphi\left(\frac{\sigma^2}{2D}, \eta\right) \mathrm{d}\eta$$

  does not act on spatial coordinates $\mathbf{r} = (x, y)$.
- The non-negativity of $s \geq 0$ is retained by $a \geq 0$.
- While the mapping from $s \to a$ depends on D, κ_{a}, and κ_{d}, the mapping from a to d_{obs} does not.

Proposed Methodology

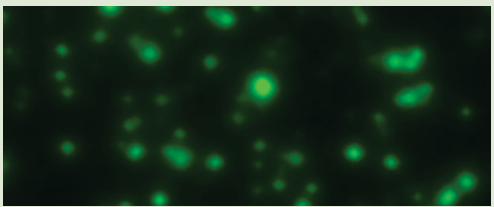
Recover the post adsorption-desorption source density rate (PSDR) a in place of the source density rate (SDR) s, obviating the need to explicitly obtain D, κ_{a}, and κ_{d}.

An Observation Model for Biomedical Assays

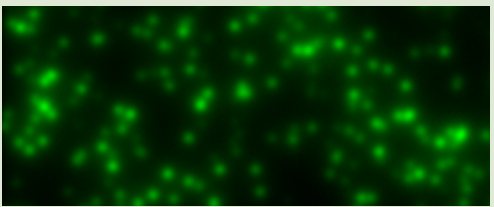
The modeling result: The image $d_{\text{obs}} \in \mathcal{D}_+$ is

$$d_{\text{obs}} = \int_0^{\sigma_{\max}} G_\sigma a_\sigma \mathrm{d}\sigma = Aa\,.$$

Consequences



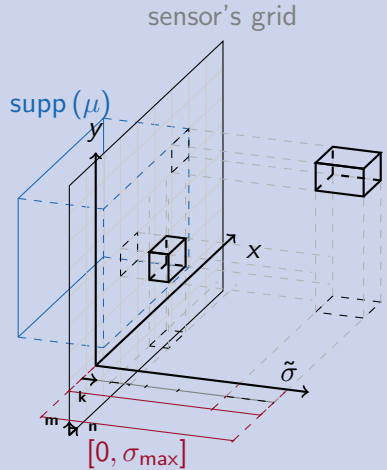
Real observation (section)



Simulated observation (section)

- Ability to generate synthetic data
- A workable observation model for inverse problem

Discretization



- Spatial grid given by camera sensor
- σ-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators to ensure norm equivalence
- The typical size of the variable a to recover will be $9 \times 2048^2 \approx 40 \times 10^6$
- Different kernel approximations are considered

Continuous observation mode

$$d_{\text{obs}} = \int_0^{\sigma_{\max}} G_\sigma a_\sigma \, \mathrm{d}\sigma = Aa$$

- $d_{\text{obs}} \in \mathcal{D} \subset \mathrm{L}^2\left(\mathbb{R}^2\right)$
- $a \in \mathcal{A}_+ \in \mathrm{L}^2\left(\mathbb{R}^3\right), \qquad a_\sigma \in \mathrm{L}^2\left(\mathbb{R}^2\right)$

Discrete observation mode

$$d_{\text{obs}} \approx \sum_{k=1}^{K} g_k \circledast a_k = Aa$$

- $d_{\text{obs}} \in \mathcal{D} \subset \mathbb{R}^{M \times N}$
- $a \in \mathcal{A}_+ \subset \mathbb{R}^{M \times N \times K}, \qquad a_k \in \mathbb{R}^{M \times N}$

Algorithmic Solutions

Naive Inverse Problem Formulation (Discrete Case)

$$a^\star = \arg\min_{a \geq 0} \|Aa - d_{\text{obs}}\|_2^2$$

A Problem with Observability

$$a^\star(d) : \mathbb{R}^{M \times N} \rightarrow \mathbb{R}^{M \times N \times K}$$

The solution is regularization (group sparsity)!

Non-negative Group Sparsity Regularized Inverse Problem

We have $d_{\text{obs}} \in \mathcal{D}_+$ and want to recover $a \in \mathcal{A}_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{a \geq 0} \left\{ \|Aa - d_{\text{obs}}\|_2^2 + \lambda \sum_{m,n} \|a_{m,n}\|_2 \right\}$$

$$a_{m,n} \triangleq \{a(m,n,k)\}_k \in \mathbb{R}^K$$

Functional Non-negative Group Sparsity Regularized Inverse Problem

We have $d_{\text{obs}} \in \mathcal{D}_+$ and want to recover $a \in \mathcal{A}_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{a \in \mathcal{A}} \left\{ \|Aa - d_{\text{obs}}\|_{\mathcal{D}}^2 + \delta_{\mathcal{A}_+}(a) + \lambda \underbrace{\int_{\mathbb{R}^2} \left(\int_0^{\sigma_{\max}} a^2(x, y, \sigma) \, \mathrm{d}\sigma \right)^{\frac{1}{2}} \mathrm{d}x\mathrm{d}y}_{\left\| \|a_{\mathbf{r}}\|_{\mathrm{L}^2(\mathbb{R}_+)} \right\|_{\mathrm{L}^1(\mathbb{R}^2)}} \right\}$$

Non-negative Group Sparsity Regularized Inverse Problem

$$\min_{a \geq 0} \left\{ \|Aa - d_{\text{obs}}\|_2^2 + \lambda \sum_{m,n} \|a_{m,n}\|_2 \right\}$$

- Still an optimization problem over 40×10^6 variables...
- Active research on first order (matrix free) methods
  - operator view instead of matrix vector multiplication
  - decoupled and closed form proximal operators
- We have considered
  - accelerated proximal gradient methods, and
  - multiplicative update rules

Proximal Gradient Methods

Nesterov Accelerated Proximal Gradient Algorithm (Fluorospot)

Require: Initial $a^{(0)} \in \mathcal{A}_+$, image observation $d_{\text{obs}} \in \mathcal{D}_+$

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** *I* **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast b_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\text{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[b_k^{(i-1)} - \eta\, g_k \circledast r^{(i)} \right]_+ \quad , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ \quad , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} \quad , \quad k = 1, \ldots, K$

6: $\quad b^{(i)} \leftarrow a^{(i)} + \alpha_i \left(a^{(i)} - a^{(i-1)} \right)$

7: **end for**

8: $a_{\text{opt}} \leftarrow a^{(i)}$

Proximal Gradient Algorithm (Simplified)

Require: Initial $a^{(0)} \in \mathcal{A}_+$, image observation $d_{\text{obs}} \in \mathcal{D}_+$

1: **for** $i = 1$ **to** I **do**

2: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\text{obs}}$

3: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta \, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

4: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} \quad , \quad k = 1, \ldots, K$

5: **end for**

6: $a_{\text{opt}} \leftarrow a^{(i)}$

Model prediction $d = Aa$

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\text{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta\, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2}\lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \quad k = 1, \ldots, K$

6: **end for**

7: $a_{\text{opt}} \leftarrow a^{(i)}$

Residual calculation $r = Aa - d_{\text{obs}}$

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \qquad r^{(i)} = d^{(i)} - d_{\text{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta\, g_k \circledast r^{(i)} \right]_+ , \qquad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \qquad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \qquad k = 1, \ldots, K$

6: **end for**

7: $a_{\text{opt}} \leftarrow a^{(i)}$

Gradient calculation $\nabla f = A^* r = A^*(Aa - d_{\text{obs}})$

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\text{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta \, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \quad k = 1, \ldots, K$

6: **end for**

7: $a_{\text{opt}} \leftarrow a^{(i)}$

Projected gradient step $a \leftarrow [a - \eta \nabla f]_+$

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\text{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta \, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \quad k = 1, \ldots, K$

6: **end for**

7: $a_{\text{opt}} \leftarrow a^{(i)}$

Shrinkage factor due to $\lambda \|a_{m,n}\|_2$

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\text{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta \, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \quad k = 1, \ldots, K$

6: **end for**

7: $a_{\text{opt}} \leftarrow a^{(i)}$

Variable update of a

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\mathrm{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta\, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \quad k = 1, \ldots, K$

6: **end for**

7: $a_{\mathrm{opt}} \leftarrow a^{(i)}$

Complexity bottlenecks ($2K$ convolutions)

1: $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$

2: **for** $i = 1$ **to** I **do**

3: $\quad d^{(i)} \leftarrow \sum_{k=1}^{K} g_k \circledast a_k^{(i-1)}, \quad r^{(i)} = d^{(i)} - d_{\mathrm{obs}}$

4: $\quad a_k^{(i)} \leftarrow \left[a_k^{(i-1)} - \eta\, g_k \circledast r^{(i)} \right]_+ , \quad k = 1, \ldots, K$

5: $\quad p \leftarrow \left(1 - \frac{\eta}{2} \lambda \left[\sqrt{\sum_{k=1}^{K} \left(a_k^{(i)} \right)^2} \right]^{-1} \right)_+ , \quad a_k^{(i)} \leftarrow p \odot a_k^{(i)} , \quad k = 1, \ldots, K$

6: **end for**

7: $a_{\mathrm{opt}} \leftarrow a^{(i)}$

Convexity guarantees convergence of the algorithm...

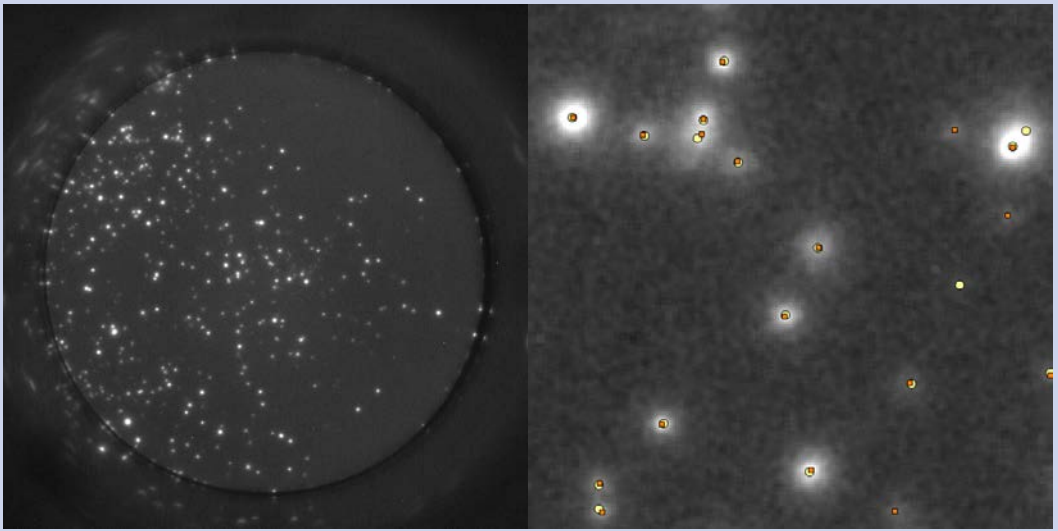
Results and Performance

Evaluation framework and metrics

- Pseudo-likelihood for source localization given by $p = \sqrt{\sum_{k=1}^{K} a_k^2}$
  - candidate cell locations obtained by local maxima of p
  - candidates pruned based on an optimized pseudo-likelihood threshold
- Detection considered correct within 3 pixels
- Evaluation metrics

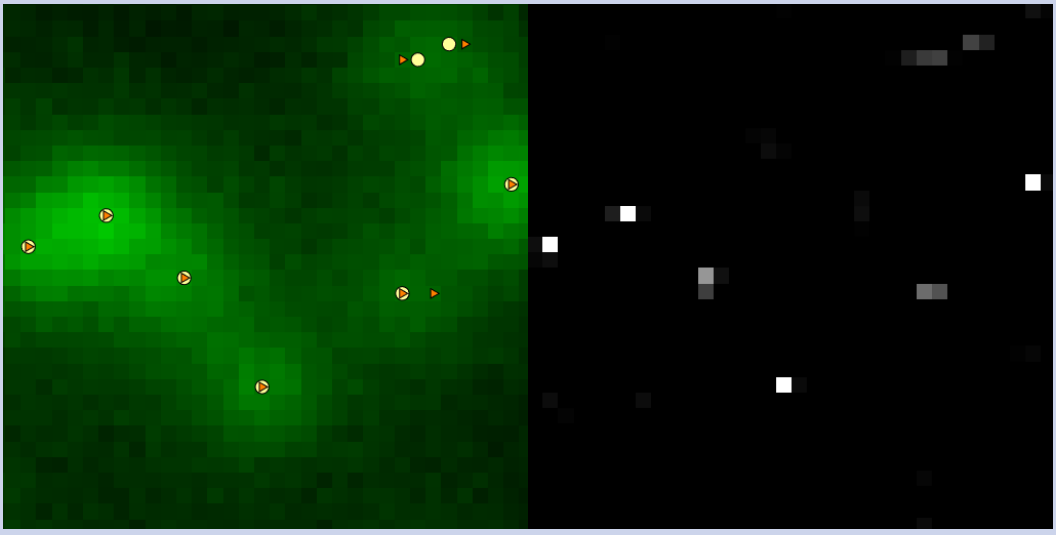
$$\mathrm{pre} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}, \quad \mathrm{rec} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}, \text{ and } \quad \mathrm{F1} = \frac{2\,\mathrm{pre}\cdot\mathrm{rec}}{\mathrm{pre}+\mathrm{rec}}$$

Example Results for Real Data (F1-Score of 0.9)



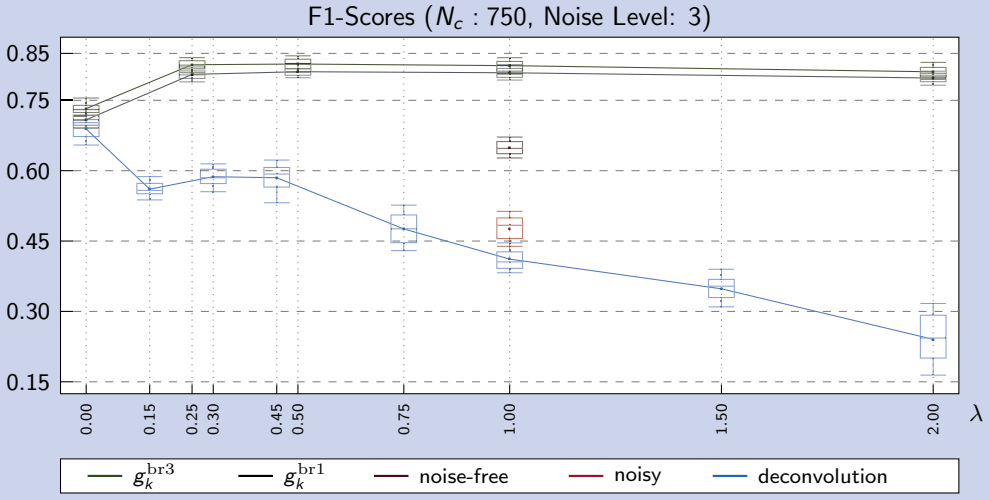
Detection results (yellow circles) and human expert labeling (orange squares).

Example Results for Synthetic Data ($N_c = 1250$)

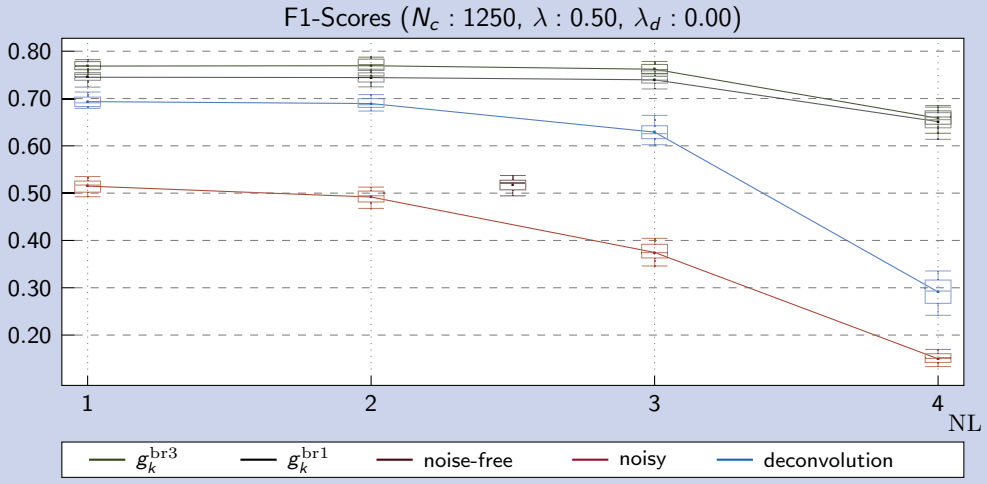


Detection results (yellow circles) and true positions (orange squares).

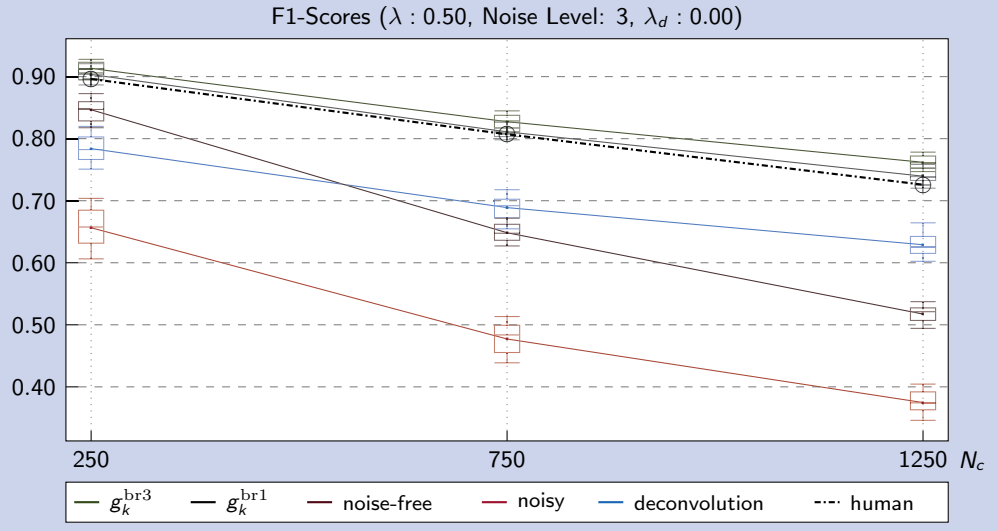
Results on Synthetic Data (F1 vs. Regularization λ)



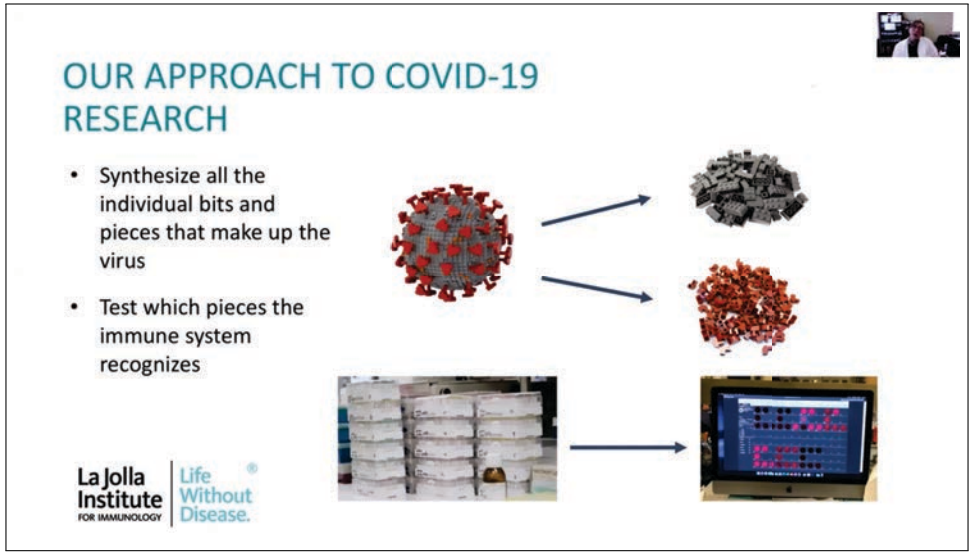
Results on Synthetic Data (F1 vs. Noise Level)



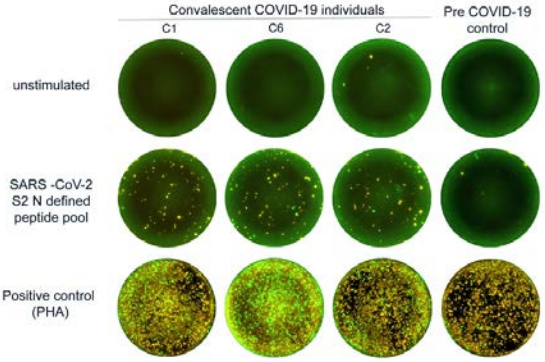
Results on Synthetic Data (F1 vs. Spot Density)



Epilogue



La Jolla Institute for Immunology - Coronavirus Update, May 14, 2020 (YouTube)



Peptide pool based on – Ahmed, Syed Faraz, Ahmed A. Quadeer, and Matthew R. McKay. "Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies." *Viruses* (2020)

Summary

- We have solved an image processing problem in immunology
  - ELISPOT and FluoroSpot
- How:
  - formulation of an over parameterized linear observation model
  - spatial group sparsity regularization
  - GPU accelerated first order (matrix free) optimization methods
- What:
  - state of the art analysis of said assays
  - capable of breaking clusters of spots
  - unparalleled positioning of spot centers
  - individual relative spot volume estimates (RSV) – even in clusters

References

- Pol del Aguila Pla and Joakim Jaldén, "Cell detection by functional inverse diffusion and group sparsity – Part I: Modeling and Inverse Problems," IEEE Transactions on Signal Processing, vol 66 , no 20, pp. 5407–5421, Sept. 2018
  `https://dx.doi.org/10.1109/TSP.2018.2868258`
- Pol del Aguila Pla and Joakim Jaldén, "Cell detection by functional inverse diffusion and group sparsity – Part II: Proximal optimization and Performance evaluation," IEEE Transactions on Signal Processing, vol 66 , no 20, pp. 5422–5437, Sept. 2018
  `https://dx.doi.org/10.1109/TSP.2018.2868256`
- Popular Scientific Descriptions and Visuals at `http://mabtech.com/iris`