Cell detection by functional inverse diffusion and non-negative group sparsity

Pol del Aguila Pla, Ph.D. Candidate
https://people.kth.se/~poldap
Division of Information Science and Engineering
School of Electrical Engineering and Computer Science

January 14, 2019 at CosmoStat
Acknowledgements


Royal Swedish Academy of Sciences

Knut and Alice Wallenberg foundation

MabTech AB

Swedish Research Council

KTH Opportunities and EECS school


Plate of Fluorospot wells. Image provided by Mabtech AB, access at http://bit.ly/Fluoro_Plate
Plate of Fluorospot wells. Image provided by Mabtech AB, access at http://bit.ly/Fluoro_Plate
\[ \approx 10 \mu m \]

\[ \approx 1 \text{ nm} \]
A Physical Model for Biomedical Assays (Modeling I)

Relevant quantities for the assay are

- A density of bound particles $d(x, y, t) \geq 0$, where the image will be $d_{\text{obs}}(x, y) = d(x, y, T)$, which evolves coupled to
A Physical Model for Biomedical Assays (Modeling I)

Relevant quantities for the assay are

- A density of bound particles \( d(x, y, t) \geq 0 \), where the image will be \( d_{\text{obs}}(x, y) = d(x, y, T) \), which evolves coupled to
- the 3D density of free particles \( c(x, y, z, t) \geq 0 \) on \( z \geq 0 \), and to
Relevant quantities for the assay are

- A density of bound particles $d(x, y, t) \geq 0$, where the image will be $d_{\text{obs}}(x, y) = d(x, y, T)$, 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 secretion over time.
A Physical Model for Biomedical Assays (Modeling I)

Relevant quantities for the assay are

- A density of bound particles \( d(x, y, t) \geq 0 \), where the image will be \( d_{\text{obs}}(x, y) = d(x, y, T) \), 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 secretion over time.

\[
\frac{\partial}{\partial t} c = D \Delta c , \\
\frac{\partial}{\partial t} d = \kappa_a c \big|_{z=0} - \kappa d d , \\
- D \frac{\partial}{\partial z} c \big|_{z=0} = s - \frac{\partial d}{\partial t} .
\]

This physical model was presented before, also for ELISPOT and Fluorospot.
Relevant quantities for the assay are

- A density of bound particles \( d(x, y, t) \geq 0 \), where the image will be \( d_{\text{obs}}(x, y) = d(x, y, T) \), 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 secretion over time.

\[
\begin{align*}
\frac{\partial}{\partial t} c &= D \Delta c , \\
\frac{\partial}{\partial t} d &= \kappa_a c \bigg|_{z=0} - \kappa_d d , \\
- D \frac{\partial}{\partial z} c \bigg|_{z=0} &= s - \frac{\partial d}{\partial t} .
\end{align*}
\]

This physical model was presented before, also for ELISPOT and Fluorospot.
A Physical Model for Biomedical Assays (Modeling I)

Relevant quantities for the assay are

- A density of bound particles \( d(x, y, t) \geq 0 \), where the image will be \( d_{\text{obs}}(x, y) = d(x, y, T) \), 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 secretion over time.

\[
\begin{align*}
\frac{\partial}{\partial t} c &= D\Delta c, \\
\frac{\partial}{\partial t} d &= \kappa_a c|_{z=0} - \kappa_d d, \\
-D\frac{\partial}{\partial z} c|_{z=0} &= s - \frac{\partial}{\partial t} c.
\end{align*}
\]

This physical model was presented before, also for ELISPOT and Fluorospot.
We consider the image observation $d_{\text{obs}} \in D_+$, with $D = L^2(\mathbb{R}^2)$ and prove that

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

with $a \in A_+$ and $A \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\text{max}} = \sqrt{2D T}$, and
An Observation Model for Biomedical Assays (I) (Modeling II)

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

$$d_{\text{obs}}(x, y) = \int_0^{\sigma_{\text{max}}} G_\sigma a_\sigma \, d\sigma,$$

with $a \in A_+$ and $A \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\text{max}} = \sqrt{2DT}$, and
An Observation Model for Biomedical Assays (I) (Modeling II)

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

$$d_{\text{obs}}(x, y) = Aa,$$

we call $A$ the diffusion operator, with $a \in \mathcal{A}_+$ and $\mathcal{A} \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\text{max}} = \sqrt{2DT}$, and
We consider the image observation $d_{\text{obs}} \in \mathcal{D}_+$, with $\mathcal{D} = L^2(\mathbb{R}^2)$ and prove that

$$d_{\text{obs}}(x, y) = Aa,$$

we call $A$ the \textit{diffusion} operator,

with $a \in \mathcal{A}_+$ and $\mathcal{A} \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\text{max}} = \sqrt{2DT}$, and

- $a(x, y, \sigma)$ is an equivalent of $s(x, y, t)$ where the effect of adsorption and desorption have been summarized.

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

- $a(x, y, \sigma)$ preserves all the spatial information in $s(x, y, t)$. 
The modeling result: The image $d_{\text{obs}} \in D_+$ can be expressed as

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

**How?**

- Independence of Brownian motion in $x$, $y$ and $z$. 

- Adsorption ($\kappa_a$) and desorption ($\kappa_d$) only regulated by $z$-movement.

- $x$- and $y$-movements only depend on $\tau$, total time in Brownian motion. In particular, according to Green function for 2D diffusion, $g_{\sqrt{2}D\tau}(x,y)$.

- $\phi(\tau,t)$ summarizes the effect of adsorption and desorption onto the time in free motion $\tau$ for each time of final adsorption $t$.

- Change variables to those significative to $x$- and $y$-movement, $\sigma = \sqrt{2D\tau}$. 

The modeling result: The image $d_{\text{obs}} \in \mathcal{D}_+$ can be expressed as

$$d_{\text{obs}} = \int_{0}^{\sigma_{\text{max}}} G_{\sigma} a_{\sigma} \, d\sigma.$$ 

How?

- Independence of Brownian motion in $x$, $y$ and $z$.
- Adsorption ($\kappa_a$) and desorption ($\kappa_d$) only regulated by $z$-movement.
The modeling result: The image \( d_{\text{obs}} \in \mathcal{D}_+ \) can be expressed as
\[
d_{\text{obs}} = \int_0^{\sigma_{\text{max}}} G_{\sigma} a_{\sigma} d\sigma.
\]

How?

- Independence of Brownian motion in \( x, y \) and \( z \).
- Adsorption (\( \kappa_a \)) and desorption (\( \kappa_d \)) only regulated by \( z \)-movement.
- \( x \)- and \( y \)-movements only depend on \( \tau \), total time in Brownian motion. In particular, according to Green function for 2D diffusion, \( g_{\sqrt{2D\tau}}(x, y) \).
The modeling result: The image $d_{\text{obs}} \in \mathcal{D}_+$ can be expressed as

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

How?

- Independence of Brownian motion in $x$, $y$ and $z$.
- Adsorption ($\kappa_a$) and desorption ($\kappa_d$) only regulated by $z$-movement.
- $x$- and $y$-movements only depend on $\tau$, total time in Brownian motion. In particular, according to Green function for 2D diffusion, $g_{\sqrt{2D\tau}}(x, y)$.
- $\phi(\tau, t)$ summarizes the effect of adsorption and desorption onto the time in free motion $\tau$ for each time of final adsorption $t$. 

Change variables to those significative to $x$- and $y$-movement, $\sigma = \sqrt{2D\tau}$. 
The modeling result: The image $d_{\text{obs}} \in \mathcal{D}_+$ can be expressed as

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

How?

- Independence of Brownian motion in $x$, $y$ and $z$.
- Adsorption ($\kappa_a$) and desorption ($\kappa_d$) only regulated by $z$-movement.
- $x$- and $y$-movements only depend on $\tau$, 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 $\tau$ for each time of final adsorption $t$.
- Change variables to those significant to $x$- and $y$-movement, $\sigma = \sqrt{2D\tau}$. 
The modeling result: The image $d_{\text{obs}} \in \mathcal{D}_+$ can be expressed as

$$d_{\text{obs}} = \int_0^{\sigma_{\text{max}}} G_\sigma a_\sigma d\sigma.$$
The modeling result: The image $d_{\text{obs}} \in \mathcal{D}_+$ can be expressed as

$$d_{\text{obs}} = \int_0^{\sigma_{\text{max}}} G_\sigma a_\sigma d\sigma.$$
Functional Inverse Diffusion (Optimization I)

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

$$\min_{a \in A} \left[ \|Aa - d_{\text{obs}}\|_D^2 + \delta_{A_+}(a) + \lambda \int_{\mathbb{R}^2} \left( \int_0^{\sigma_{\text{max}}} a^2(x, y, \sigma) \, d\sigma \right)^{\frac{1}{2}} \, dx \, dy \right],$$

Proximal Optimization
Functional Inverse Diffusion (Optimization I)

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

\[
\min_{a \in \mathcal{A}} \left[ \| Aa - d_{\text{obs}} \|_D^2 + \delta_{\mathcal{A}_+}(a) + \lambda \left\| a_{r} \right\|_{L^2(\mathbb{R}_+)} \left\| a_{r} \right\|_{L^1(\mathbb{R}^2)} \right],
\]

with \( r = (x, y) \).

Proximal Optimization
We have $d_{\text{obs}} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{a \in A} \left[ \|Aa - d_{\text{obs}}\|_D^2 + \delta_{A_+}(a) + \lambda \left\| r \right\|_{L^2(\mathbb{R}^+)} \left\| r \right\|_{L^1(\mathbb{R}^2)} \right],$$

with $r = (x, y)$.

Proximal Optimization

- How do we solve this optimization problem? Can it be solved?
Functional Inverse Diffusion (Optimization I)

We have \( d_{\text{obs}} \in D_+ \) and want to recover \( a \in A_+ \). We propose the (non-smooth, constrained) convex problem

\[
\min_{a \in A} \left[ \| Aa - d_{\text{obs}} \|_D^2 + \delta_{A_+}(a) + \lambda \left\| a r \right\|_{L^2(R^+)} \left\| r \right\|_{L^1(R^2)} \right],
\]

with \( r = (x, y) \).

Proximal Optimization

- How do we solve this optimization problem? Can it be solved?
- Three terms, two non-smooth (with known prox), one smooth (with non-trivial but manageable gradient). Convex problem, but existence and unicity not given (function spaces).
We have $d_{\text{obs}} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{a \in A} \left[ \| Aa - d_{\text{obs}} \|_D^2 + \delta_{A_+}(a) + \lambda \left\| ar \right\|_{L^2(\mathbb{R}^+)} \left\| ar \right\|_{L^1(\mathbb{R}^2)} \right],$$

with $r = (x, y)$.

**Proximal Optimization**

- How do we solve this optimization problem? Can it be solved?
- Three terms, two non-smooth (with known prox), one smooth (with non-trivial but manageable gradient). Convex problem, but existence and unicity not given (function spaces).
- Do we need forward-backward primal-dual splitting?
We have \( d_{\text{obs}} \in D_+ \) and want to recover \( a \in A_+ \). We propose the (non-smooth, constrained) convex problem

\[
\min_{a \in A} \left[ \| Aa - d_{\text{obs}} \|_D^2 + \delta_{A_+}(a) + \lambda \left\| a_r \right\|_{L^2(\mathbb{R}^+)} \left\| a_r \right\|_{L^1(\mathbb{R}^2)} \right],
\]

with \( r = (x, y) \).

How do we solve this optimization problem? Can it be solved?

Three terms, two non-smooth (with known prox), one smooth (with non-trivial but manageable gradient). Convex problem, but existence and unicity not given (function spaces).

Do we need forward-backward primal-dual splitting? No. Not if we can find the prox of the sum of the two non-smooth terms. It is faster (Pustelnik and Condat, 2017).
Functional Inverse Diffusion (Optimization I)

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

$$\min_{a \in A} \left[ \| Aa - d_{\text{obs}} \|_D^2 + \delta_{A_+}(a) + \lambda \left\| [a_r]_{L^2(\mathbb{R}^+)} \right\|_{L^1(\mathbb{R}^2)} \right],$$

with $r = (x, y)$.

Proximal Optimization

- Do we need forward-backward primal-dual splitting? No. Not if we can find the prox of the sum of the two non-smooth terms. It is faster (Pustelnik and Condat, 2017).
- We showed that the prox of the non-negative group-sparsity regularizer is

$$p_r = [a_r]_+ \left( 1 - \frac{\gamma \lambda}{\left\| [a_r]_{+} \right\|_{L^2([0,\sigma_{\text{max}}])}} \right)_+.$$
Functional Inverse Diffusion - APG algorithm (Optimization II)

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

**Ensure:** A solution $a_{opt} \in \mathcal{A}_+$

1. $b^{(0)} \leftarrow a^{(0)}$, $i \leftarrow 0$
2. **repeat**
   3. $i \leftarrow i + 1$, $\alpha \leftarrow \frac{t(i-1)-1}{t(i)}$
   4. $a^{(i)} \leftarrow b^{(i-1)} - \sigma_{\text{max}}^{-1} A^* \left( A b^{(i-1)} - d_{obs} \right)$
   5. **for all** $r \in \mathbb{R}^2$ **do**
      6. $a_r^{(i)} \leftarrow \begin{bmatrix} a_r^{(i)} \end{bmatrix} + \left( 1 - \frac{(2\sigma_{\text{max}})^{-1} \lambda}{\|a_r^{(i)}\|_{L^2([0,\sigma_{\text{max}}])}} \right) +$
   7. **end for**
   8. $b^{(i)} \leftarrow a^{(i)} + \alpha \left( a^{(i)} - a^{(i-1)} \right)$
9. **until** convergence
10. $a_{opt} \leftarrow a^{(i)}$

Sequences of $t(i)$ can be chosen as (Bech and Teboulle, 2009) or as (Chambolle and Dossal, 2015).
Discretization

- Spatial grid given by camera sensor

Sensor’s grid

$\text{supp } (\mu)$

$[0, \sigma_{\text{max}}]$
Discretization

- Spatial grid given by camera sensor
- \( \sigma \)-grid with different levels of detail

\[ \text{supp} \left( \mu \right) \]

sensor’s grid

\[ [0, \sigma_{\text{max}}] \]
Discretization

- Spatial grid given by camera sensor
- $\sigma$-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)

$\text{supp}(\mu) \subset [0, \sigma_{\text{max}}]$
Discretization

- Spatial grid given by camera sensor
- \(\sigma\)-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators

The typical size of the variable \(a[m,n,k]\) to recover will be \(2048^2 \times 6 = 25 \cdot 10^6\). Different kernel approximations are considered.
Discretization

- Spatial grid given by camera sensor
- \(\sigma\)-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators
- Resulting algorithm can be reasoned as discrete APG

Supp(\(\mu\))

Sensor’s grid

\([0, \sigma_{\text{max}}]\)

Different kernel approximations are considered
Discretization

- Spatial grid given by camera sensor
- $\sigma$-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators
- Resulting algorithm can be reasoned as discrete APG
- The typical size of the variable $a[m, n, k]$ to recover will be $2048^2 \times 6 = 25 \cdot 10^6$
Discretization

- Spatial grid given by camera sensor
- $\sigma$-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators
- Resulting algorithm can be reasoned as discrete APG
- The typical size of the variable $a[m, n, k]$ to recover will be $2048^2 \times 6 = 25 \cdot 10^6$
- Different kernel approximations are considered
Evaluation on Synthetic Data

Besides thorough human testing on real data, we can evaluate our approach on synthetic data. To evaluate the location accuracy, we run 10000 iterations of the algorithm, find spatial maxima and threshold them optimally, and, defining a tolerance of $\Delta = 3$ pix we compute the detection metrics

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

Example

$\times$ : Real cells
Evaluation on Synthetic Data

Besides thorough human testing on real data, we can evaluate our approach on synthetic data. To evaluate the location accuracy, we run 10000 iterations of the algorithm, find spatial maxima and threshold them optimally, and, defining a tolerance of $\Delta = 3$ pix we compute the detection metrics

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

Example

$x$: Real cells  
$+$: Detections
Evaluation on Synthetic Data

Besides thorough human testing on real data, we can evaluate our approach on synthetic data. To evaluate the location accuracy, we run 10000 iterations of the algorithm, find spatial maxima and threshold them optimally, and, defining a tolerance of $\Delta = 3 \text{ pix}$, we compute the detection metrics

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

Example

$\times$: Real cells  
$\dagger$: Detections
Results on Synthetic Data (I)

F1-Scores ($\lambda: 0.50$, Noise Level: 3, $\lambda_d: 0.00$)

512 × 512 noisy images with noise equivalent to 6-bit quantization.
True positions (orange triangles) and detections (yellow circles).

Pixels’ contr. to the regularizer, i.e.,
$$\sqrt{\int a^2(x, y, \sigma)d\sigma}.$$
Detection results (yellow circles) and human labeling (orange squares). F1-Score relative to human, 0.9 (whole image).
SpotNet - Learned iterations for faster inverse problems

Based on the learned gradient descent of (Gregor and LeCun, 2010), recently explored by (Giryes, Eldar et al., 2018).
Based on the learned gradient descent of (Gregor and LeCun, 2010), recently explored by (Giryes, Eldar et al., 2018).
SpotNet - Learned iterations for faster inverse problems

\[ a^{(i)} \leftarrow \left[ b^{(i-1)} - \sigma_{\text{max}}^{-1} A^* \left( Ab^{(i-1)} - d_{\text{obs}} \right) \right] \]

\[ a^{(i)} \leftarrow \varphi_{\lambda} \left( a^{(i)} \right) \]
\[ a(i) \leftarrow \left[ b(i-1) - \sigma_{\text{max}}^{-1} A^* \left( Ab^{(i-1)} - d_{\text{obs}} \right) \right] \]

\[ a(i) \leftarrow \varphi_{\lambda} \left( a(i) \right) \]

\[ b(i) \leftarrow a(i) + \alpha \left( a(i) - a(i-1) \right) \]
SpotNet - Learned iterations for faster inverse problems

\[ a^{(i)} \leftarrow \left[ b^{(i-1)} - \sigma_{\text{max}}^{-1} A^* \left( A b^{(i-1)} - d_{\text{obs}} \right) \right] \]

\[ a^{(i)} \leftarrow \varphi_{\lambda} \left( a^{(i)} \right) \]

\[ b^{(i)} \leftarrow a^{(i)} + \alpha \left( a^{(i)} - a^{(i-1)} \right) \]

Based on the learned gradient descent of (Gregor and LeCun, 2010), recently explored by (Giryes, Eldar et al., 2018).
Results for SpotNet with $L = 3$ and smaller kernels

(a) SpotNet
(b) ConvNet
(c) MSE on 150 test images

- Evaluation of SpotNet and a generic ConvNet on $\text{MSE}\{\hat{a}\}$.
- Trained on 7 images with 1250 cells.
Evaluation of SpotNet and a generic ConvNet on F1 score as above.

Trained on 7 images with 1250 cells.
Thank you

Please, feel free to ask questions.

January 14, 2019 at CosmoStat