Ph.D. Dissertation Defense

Statistical Design of Position-Encoded 3D Microarrays

Pinaki Sarder
Advisor: Prof. Arye Nehorai

Department of Electrical and Systems Engineering
Washington University in St. Louis
Outline

• Background
• Performance analysis
• Statistical design
• Estimation
• Results
• Future work
• Other research
Background
Existing 3D Microarrays

Background:

- Microarrays detect, identify, and quantify targets (proteins, antibodies, mRNA etc.) in a solution [1].
- Conventional microarrays are 2D.

3D microarrays:

- Contain multiple polystyrene microspheres (5-6 μm in diameter) sensitive to individual targets [1], [2].
- Microspheres are randomly placed on a substrate [2].
- Microspheres employ color-coded quantum-dots (QDs) [3].
- Color-codes identify targets.

Existing 3D Microarrays (Cont.)

- Advantages of 3D microarrays:
  - High sensitivity.
  - Directional binding.
  - Fast reaction (high surface-to-volume ratio).

- However, 2D microarrays have position encoding:
  - Avoid target identification errors.

- Potential applications:
  - Versatile screening.
  - Drug discovery.
  - Gene sequencing.
  - Environmental monitoring.
  - Home-tests (pregnancy tests, glucose meters, cholesterol meters, etc.).
Target-Captured Microsphere

- Receptors on the microsphere react to target molecules.
- Nanospheres:
  - Diameter $\sim$100nm.
  - Indicate reaction (yes-no).
  - Enhance detection.
Microarray Solution Flow

To create reaction, the target solution:

- Flows tangentially.
- Encounters all the microspheres.
How Does 3D Microarray Work?

Targets:

- Antibody, protein, mRNA, etc.
- Free of labels (e.g., fluorescent dyes).
- Unknown concentrations.

Inferences:

- Identify targets using the light spectra from the microsphere QDs.
- Estimate target concentrations using the light levels from the nanosphere QDs.
How Does 3D Microarray Work? (Cont.)

Adding nanospheres:

- Detection is label-free. Hence, target molecule structures and chemical properties are not modified.
- QDs do not suffer from saturation or photo-disintegration under intense illumination. Hence, the detection sensitivity enhances.
To obtain the measurement:

- **Ultra-violet (UV) light** is applied to excite the QDs.
- **A fluorescence microscope** (e.g., non-confocal) and an image sensor (e.g., CCD/CMOS) capture cross-section images of the microarray.
Microscopy

Non-confocal fluorescence microscope.
Ideal Microsphere and Nanosphere Images

Microsphere QD light:
- Spherical source.
- Cross-section images result in 2D discs.

Nanosphere QD light around each microsphere:
- Spherical-shell source.
- Cross-section images result in 2D rings.

2D cross-section intensity image (ideal).
- QD light from the microsphere.
- QD light from the nanospheres.
Real Microscope Image

Focal-plane quantum-dot intensity image of microspheres without targets.
Drawbacks of Existing 3D Microarrays

- Random placement of microspheres results in:
  - Inefficient packing.
  - Complex image processing.
  - Suboptimal detectibility (close proximity of microspheres).

- Microsphere QD color spectra are noisy which results in error in identifying targets.

- Sensors require high cooling in imaging, hence expensive.

Optical cross-talking in target-concentration profiles of two neighboring microspheres (schematic).
Our Research

Goal:

- Design a compact 3D microarray device without target identification errors.
- Minimize the error in estimating the target concentrations.

Our solution:

- Place the microspheres at controllable positions, to avoid the identification error.
- Use statistical analysis to select the optimal (minimal) distance between microspheres for a desired imaging performance, at a given temperature.

We:

- Analyze the statistical performance using posterior Cramér-Rao bound.
- Select the optimal distance.
- Estimate the image parameter using maximum-likelihood.
Proposed 3D Microarray Layout

Positioning of microspheres is realized using microfluidics or magnetics.

- Position encoding (no QD encoding of the microspheres).
- Error-free target identifiability.
- Higher sensitivity.
- Efficient packing.
- Simplified imaging.

Uniform 2D grid layout of the microarray.
Statistical Performance Analysis
Performance Analysis Overview

- **Performance measure:** Error in estimating the unknown target concentrations from two neighboring microspheres.

- **Analysis** through computing the posterior Cramér-Rao bound (PCRB) on this error.
Measurement Model

Block diagram of the imaging.

\[
\tilde{s}(x, y, z; \theta) = \int \int \int h(x - \tilde{x}, y - \tilde{y}, z, \tilde{z}) s(\tilde{x}, \tilde{y}, \tilde{z}; \theta) d\tilde{x} d\tilde{y} d\tilde{z}[4].
\]

- \( h(x, y, z) \): Microscope PSF.
- \( s(x, y, z, \theta) \): Target-concentration profile.
- \( \theta \): Unknown target-concentration levels.

Measurement Model (Cont.)

Measurement:

\[ g(x, y, z; \theta) = g_{\text{ccd}}(x, y, z; \theta) + w_b(x, y, z). \]

- \( \beta g_{\text{ccd}}(x, y, z; \theta) \sim \mathcal{P}(\beta \tilde{s}(x, y, z; \theta)) \).
- \( \beta \) : Photon conversion factor of the image sensor.
- \( \mathcal{P}(\eta) \) : Poisson random variable with mean-rate \( \eta \).
- \( w_b(x, y, z) \) : Background (thermal) noise, i.i.d. from voxel to voxel.
  - Gaussian distributed with mean zero and variance \( \sigma_b^2 \).
  - Independent of \( g_{\text{ccd}}(x, y, z; \theta) \) at each voxel.
- \( x \in \{x_1, x_2, \ldots, x_K\} \), \( y \in \{y_1, y_2, \ldots, y_L\} \), and \( z \in \{z_1, z_2, \ldots, z_M\} \).
Measurement Model (Cont.)

• Assuming high signal-to-noise ratio (SNR) imaging of QDs:

\[
g(x, y, z; \theta) = g_{\text{ccd}}(x, y, z; \theta) + w_b(x, y, z),
\]

\[
\approx \tilde{s}(x, y, z; \theta) + w_p(x, y, z; \theta) + w_b(x, y, z),
\]

where \( w_p(x, y, z; \theta) \) is:

- Gaussian distributed with zero mean and variance \( \tilde{s}(x, y, z; \theta)/\beta \).
- Independent from voxel to voxel.

• Measurement in vector form:

\[
g = \tilde{s} + w_p + w_b.
\]
Object Model: Full Shell

- Assumptions:
  - Microspheres are completely surrounded by intended targets.
  - Targets are attached to nanospheres.

- Nanosphere QD lights from two neighboring microspheres, in shell shapes:

\[
s(x, y, z; \theta) = s_{sh}(x, y, z; \theta_1) + s_{sh}(x - d, y, z; \theta_2),
\]

where (for \( i = 1, 2 \))

\[
s_{sh}(x, y, z; \theta_i) = \begin{cases} 
\theta_i & \text{if } r_1 \leq \sqrt{x^2 + y^2 + z^2} \leq r_2, \\
0 & \text{otherwise.}
\end{cases}
\]

- \( d \): Distance between shells.
- \( \theta = [\theta_1, \theta_2]^T \): Unknown parameters.
- \( \theta_i \sim \text{Uniform}(0, \theta_{\text{MAX}}) \): Target-concentration level per measured voxel.
- \( \theta_{\text{MAX}} \): User-chosen constant.
Ideal Nanosphere QD Light Images: Full Shell

Full shell:

Nanospheres

Microsphere

2D cross-section intensity image (ideal).

Full-shell shape fits:

- High target concentration.
- Sufficiently long sensing duration.
Object Model: Sparse Shell

- **Assumptions:** Same as for full shell, except that microspheres are partially surrounded by intended targets.

- **Nanosphere QD lights from two neighboring microspheres, in shell shapes:**

\[
s(x, y, z; \theta) = s_{sh}(x, y, z; \theta_1) + s_{sh}(x - d, y, z; \theta_2),
\]

where (for \( i = 1, 2 \))

\[
s_{sh}(x, y, z; \theta_i) = \begin{cases} 
\theta_i(x, y, z) & \text{if } r_1 \leq \sqrt{x^2 + y^2 + z^2} \leq r_2, \\
0 & \text{otherwise}.
\end{cases}
\]

- \( d \): Distance between shells.
- \( \theta = [\theta_1^T, \theta_2^T]^T \): Unknown parameters.
- \( \theta_i \): Vector form of \( \theta_i(x, y, z) \) from the measured voxels.
- \( \theta_i(\cdot) \sim \text{Exp}(\tau) \) [5]: Sparse target-concentration level in a measured voxel.
- \( \tau \): Inversely proportional to sparsity.

Ideal Nanosphere QD Light Images: Sparse Shell

Sparse shell:

Nanospheres

Microsphere

2D cross-section intensity image (ideal).

Sparse-shell shape fits:

- Low target concentration.
- Short sensing duration.
Microscope PSF Model

Microscope PSF model (for a point source at a depth $\tilde{z}$ in the specimen) [6]:

$$h(x, y, z) = \left| \int_0^1 J_0(2\pi N_a \alpha \sqrt{x^2 + y^2}/M' \lambda) \exp(j2\pi \psi(z, \tilde{z}, N_a, \alpha, n_o, n_s)/\lambda) \alpha d\alpha \right|^2.$$

- $J_0(\cdot)$ : Bessel function of the first kind.
- $N_a$ : Microscope numerical aperture.
- $\alpha$ : Normalized radius in the back focal plane.
- $M'$ : Lens magnification.
- $\lambda$ : QD emission wavelength.
- $\psi(\cdot)$ : Optical path difference function.
- $n_o$ and $n_s$ : Refractive indexes of the immersion oil and specimen.

Posterior Cramér-Rao Bound

- Joint data-parameter log-likelihood:

\[
\log p_{G,\Theta}(g, \theta) \propto - \sum_x \sum_y \sum_z \left[ \frac{(g(\cdot) - \tilde{s}(\cdot))^2}{2\left(\frac{\tilde{s}(\cdot)}{\beta} + \sigma_b^2\right)} + \frac{1}{2} \log \left(\frac{\tilde{s}(\cdot)}{\beta} + \sigma_b^2\right) \right]
\]

\[
+ \log(p_{\Theta}(\theta)).
\]

- Information matrix \( J \) has elements \( J_{i'j'} = \mathbb{E} \left[ - \frac{\partial^2 \log p_{G,\Theta}(g, \theta)}{\partial \theta_{i'} \partial \theta_{j'}} \right] \).

- PCRBs on estimation errors correspond to the diagonal entries of \( J^{-1} \) [7].

Statistical Design
Design Approach

• **Goal:** Use the statistical performance of the imaging to select the minimal distance between the microspheres, for a desired estimation error and a given temperature.

• **Design variables:**
  - Distance between the microspheres.
  - Temperature in imaging.

• **Performance measure:** \( p = \text{Trace}(\text{PCRB}) \).

• **Trade offs:** Higher temperature (i.e., reducing cost) and/or smaller distance vs. higher estimation accuracy.
Sensor Temperature Effect

- Estimation error depends on the image sensor temperature $T$.

- Background noise variance $\sigma_b^2$ varies with $T$ [8]:

$$\sigma_b^2(T) = B \exp(-E_g/2k_B T),$$

where
- $B$ : Constant.
- $E_g$ : Bandgap of the detector material.
- $k_B$ : Boltzmann constant.

Minimal Distance Selection
As we increase the distance between the microspheres:

- Their lights do not interfere with each other.
- The performance measure flattens.
- PCRB matrix becomes block-diagonal.
Minimal Distance Selection

Goal: For a given temperature, select the minimal distance at which the performance measure starts to flatten.

Approach:
- Substitute estimates of $B$ and $\beta$ in the performance measure.
- Fit a parametric curve to the estimated performance measure.
- Find the minimal distance at which the performance measure starts to flatten using least-squares estimation.
Performance Measure: Parametric Model

- Parametric model of the performance measure as a function of $d$:

$$p(d) = c \exp(-\rho d)I_{[0,d_0)}(d) + p_0,$$

where

- $I_{[0,d_0)}(d)$ is an indicator function:

$$I_{[0,d_0)}(d) = \begin{cases} 1 & \text{if } 0 \leq d < d_0, \\ 0 & \text{otherwise}. \end{cases}$$

- $c$, $\rho$, $d_0$, and $p_0$ are unknown parameters,

- Performance measure starts to flatten at $d_0$. 
Performance Parametric Model Example

Example of the parametric model for the performance measure.

\[ p(d) = c \exp(-\rho d) + p_0 \]

- \( c = 2328.7 \)
- \( p_0 = 0.1866 \)
- \( \rho = 1.1765 \times 10^6 \)
- \( d_0 = 10.8919 \mu m \)
Selecting the Minimal Distance

- Compute $p(d)$ at increasing values of $d$ at $d_1 \leq d_2 \leq \cdots \leq d_N$.

- In vector form:

$$ p = A(\zeta)x + e, $$

where

- $p = [p(d_1), p(d_2), \ldots, p(d_N)]^T$.
- $A(\zeta)$: Matrix with $j^{th}$ row as $[\exp(-\rho d_j)I_{[0,d_0]}(d_j), 1]$.
- $e$: Error vector.
- $\zeta = [\rho, d_0]^T$.
- $x = [c, p_0]^T$.
- $\zeta, x$: Unknown parameters.
Selecting the Minimal Distance (Cont.)

• Least-squares estimates [9]:

\[ \hat{\zeta} = \arg \max_{\zeta} \{ p^T \Pi(\zeta) p \}, \]

\[ \hat{x} = [A^T(\hat{\zeta}) A(\hat{\zeta})]^{-1} A^T(\hat{\zeta}) p, \]

where

\[ \Pi(\zeta) = A(\zeta) [A^T(\zeta) A(\zeta)]^{-1} A^T(\zeta). \]

• Select the minimal distance as \( \hat{d}_0 \).

Estimating $B$ and $\beta$
Estimation Overview

**Goal:** Estimate $B$ and $\beta$ using a pilot experiment with existing 3D microarray.

**Approach:**
- Image multiple QD-embedded microspheres [10] at temperature $T_0$.
- Estimate $B$ using the noise-only sections of the captured image.
- Estimate $\beta$ from each microsphere image.

Measurement Model

Measurement from a single microsphere:

\[ g(x, y, z; \theta, \beta, \sigma_b^2) = \int_{\tilde{x}} \int_{\tilde{y}} \int_{\tilde{z}} h(x - \tilde{x}, y - \tilde{y}, z, \tilde{z}) s(\tilde{x}, \tilde{y}, \tilde{z}; \theta) d\tilde{x} d\tilde{y} d\tilde{z} \]

\[ + w_p(x, y, z; \theta, \beta) + w_b(x, y, z; \sigma_b^2). \]

- \( s(x, y, z; \theta) \): QD intensity profile of a single microsphere, in spherical shape,

\[ s(x, y, z; \theta) = \begin{cases} \theta & \text{if } \sqrt{x^2 + y^2 + z^2} \leq r, \\ 0 & \text{otherwise}. \end{cases} \]

- \( \theta \): Microsphere QD intensity level per voxel.
- \( r \): Radius of the microsphere.
- \( \beta, \sigma_b^2, \) and \( \theta \) (nuisance): Unknown parameters.
- \( x \in \{x_1, x_2, \ldots, x_K\}, y \in \{y_1, y_2, \ldots, y_L\}, \) and \( z \in \{z_1, z_2, \ldots, z_M\}. \)
Estimating $\theta$

- **Goal:** Estimate $\theta$ from each microsphere image $U_{n'} (\forall n' \in \{1, 2, \ldots, N'\})$.

- **Assumption:** Large $\beta$ (highly sensitive imaging).

- Hence, neglecting $w_p(\cdot)$:

$$g(x, y, z; \theta, \sigma^2_b) \approx \theta \int \int \int h(x - \tilde{x}, y - \tilde{y}, z, \tilde{z}) f(\tilde{x}, \tilde{y}, \tilde{z}) d\tilde{x}d\tilde{y}d\tilde{z} + w_b(x, y, z; \sigma^2_b),$$

where $f(x, y, z) = s(x, y, z; \theta)/\theta$.

- Define, $\tilde{s}'(x, y, z) = \int \int \int h(x - \tilde{x}, y - \tilde{y}, z, \tilde{z}) f(\tilde{x}, \tilde{y}, \tilde{z}) d\tilde{x}d\tilde{y}d\tilde{z}$.

- In vector form:

$$g = \theta \tilde{s}' + w_b,$$

- $\tilde{s}'$ and $w_b$ are the vector forms of $\tilde{s}'(\cdot)$ and $w_b(\cdot)$, respectively.
Estimating $\theta$ (Cont.)

- Log-likelihood:

$$C_0(\theta, \sigma_b^2) = -\frac{KLM}{2} \ln \sigma_b^2 - \frac{||g - \theta \tilde{s}'||^2}{2\sigma_b^2}.$$ 

- Maximum likelihood (ML) estimates [11]:

$$\hat{\theta} = \left[ \tilde{s}'^T \tilde{s}' \right]^{-1} \tilde{s}'^T g,$$

$$\hat{\sigma}_b^2 = (KLM)^{-1} g^T P_{\tilde{s}} \tilde{s}' g,$$

where

$$P_{\tilde{s}} = I - \tilde{s}' \left[ \tilde{s}'^T \tilde{s}' \right]^{-1} \tilde{s}'^T.$$

Estimating $B$

Estimate $B$ using the estimate of $\sigma_d^2$:

- Estimate $\sigma_d^2$ from the noise only sections of the captured image, employing the classical ML estimation method discussed in [12, Ch. 6].
  - Use a large number of measurement to ensure consistency.

- Compute $B$ by fitting the estimated $\sigma_d^2$ with $B \exp(-E_g/2k_B T_0)$.

Estimating $\beta$

- **Goal:** Estimate $\beta$ from each microsphere image $U_{n'} \ (\forall n' \in \{1, 2, \ldots, N'\})$.

- **Method-of-moments estimate:**

$$\hat{\beta} = \frac{\sum_x \sum_y \sum_z \tilde{s}(x, y, z; \hat{\theta})}{\sum_x \sum_y \sum_z \left[ (g(x, y, z) - \tilde{s}(x, y, z; \hat{\theta}))^2 - \hat{\sigma}_b^2 \right]},$$

where

- $\hat{\theta}$: Estimate of $\theta$.
- $\hat{\sigma}_b^2$: Estimate of $\sigma_b^2$.

- **Note:** Conventionally $\beta$ is computed using detector’s physical parameters [12].

Estimation Summary

Capturing Image $U$

Individual object images $U_{n'} \approx g_{n'}(x, y, z)$ ($\forall n' \in \{1, 2, \ldots, N'\}$)

Estimating $\sigma_b^2$

Estimating $\theta$

Maximum likelihood estimation

$\hat{\theta}_{n'}$ $
\begin{bmatrix}
\hat{\theta}_{n'} & \forall n' \in \{1, 2, \ldots, N'\} \\
\hat{\beta}_{n'} & \forall n' \in \{1, 2, \ldots, N'\}
\end{bmatrix}$

Estimating $\beta$

Method of moments estimation

Estimating $B$

Maximum likelihood estimation

$\hat{B}$

Note: The estimated histograms of the imaging parameters from multiple microsphere images can be employed as prior information for any next step estimation using maximum a posteriori (MAP) techniques.
Results
Estimating $B$, $\beta$, and $\theta$ Using Existing Microarray

- We imaged microspheres ($r = 2.5 \mu m$), placed on a PDMS substrate.
- Used non-confocal microscope for image acquisition, with (partial) set-up:
  - $N_a = 1.3$, $M' = 40X$, $\lambda = 535nm$, $n_o = 1.334$, $n_s = 1.3$.
  - Resolution: $\Delta z = 1 \mu m$ and $\Delta x = \Delta y = 0.654 \mu m$.
  - Temperature: $T_0 = 10^0C$.

Microspheres (without targets) image using focal-plane quantum-dot intensity.
Estimating $B$, $\beta$, and $\theta$ Using Real Data

Histograms of the estimated $\beta$ and $\theta$ from 65 individual microsphere images.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimated values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>$7.29 \times 10^6$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>305 (median)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0053 (median)</td>
</tr>
</tbody>
</table>
Numerical Design Set-Up: Full-Shell Case

Set-up:

- Object: QD light intensity from two target-captured microspheres.
  - \( r_1 = 2.774 \mu m \) and \( r_2 = 2.874 \mu m \).
  - Protein targets (diameter 250nm).
  - \( \theta_{\text{MAX}} \approx 0.0053 \).

- Microscope parameters: \( N_a = 1.3 \), \( M' = 40X \), \( \lambda = 535\text{nm} \), \( n_o = 1.334 \), \( n_s = 1.3 \).
- CCD parameter: \( \beta = 305 \) (median of the estimates).
- Noise level \( \sigma_b^2 \): \( B = 7.29 \times 10^6 \), \( E_g = 1.15\text{eV (Si)} \), \( T \approx -10 \text{ to } 20^0\text{C} \).
Effect of Microspheres’ Distance on Performance

Result at $T = 0^\circ C$.

Observation: Optimal (minimal) distance is $17 \mu m$. 
Effect of Maximum Light Level on Design

Result at \(T = 0^0\text{C}\) for varying \(\theta_{\text{MAX}}\).

Observation: Optimal distance is robust with respect to the maximum possible target-concentration level.
Effect of Temperature on Performance

Result at $d = 13\mu$m.

Observation: Performance degrades with higher temperature at a fixed distance.
Distance and Temperature Effects on Performance

Performance as a function of temperature and distance.
Numerical Design Set-Up: Sparse-Shell Case

Set-up:

- Object: QD light intensity from two target-captured microspheres.
  - \( r_1 = 2.774\,\mu m \) and \( r_2 = 2.874\,\mu m \).
  - Protein targets (diameter 250nm).
  - \( \tau = 1 \) (more sparsity) or \( \tau = 5 \) (less sparsity).
- Microscope parameters: \( N_a = 1.3, M' = 40X, \lambda = 535nm, n_o = 1.334, n_s = 1.3 \).
- CCD parameter: \( \beta = 305 \) (median of the estimates).
- Noise level \( \sigma_b^2 : B = 7.29 \times 10^6, E_g = 1.15\text{eV (Si)}, T \approx -10 \text{ to } 20^\circ\text{C} \).
Effect of Microspheres’ Distance on Performance

Result at $T = 10^0C$ with $\tau = 1$.

Observation: Optimal (minimal) distance is $11\mu m$. 
Effect of Sparsity on Design

Result at $T = -10^0 \text{C}$ with $\tau = 1$ and $\tau = 5$.

Observation: Optimal distance is robust with respect to the sparsity level.
Effect of Temperature on Performance

Result at $d = 7.5\,\mu m$ with $\tau = 1$.

Observation: Performance degrades with higher temperature at a fixed distance.
Distance and Temperature Effects on Performance

Performance as a function of temperature and distance with $\tau = 1$. 
Implementing the Microarrays
Overview

**Goal:** Implement the proposed microarray based on the proposed statistical design.

**Microfluidic approach:**

- Fabricate PDMS traps (in a microfluidic channel) using lithography.
- Separate the traps based on the computed minimal microsphere distance.
- Place microspheres in traps using microfluidic hydrodynamic trapping.

**Acknowledgement:** This implementation is done by our collaborators Drs. Axel Scherer and Zhenyu Li of the Caltech Nanofabrication Group.
Microfluidic Hydrodynamic Trapping

Microsphere trapping device based on a microfluidic hydrodynamic trapping (separations are not yet optimal).

- Microspheres in liquid flow from the inlet to outlet through vertically aligned microfluidic channels containing the traps.
- Microspheres are trapped in the determined positions.
Trapping:

- Fluidic resistance through P1 is smaller than P2.
- Hence, microsphere moves through P1 into the empty trap.
Bypassing: Second microsphere bypasses first trap through P2 and fills the second trap.
Microsphere trapping through P2.
Real Trapping Demonstration
Conclusion

• We proposed a 3D microarray with position-controlled microspheres, and discussed its advantages.

• This research is multidisciplinary, involving optimal statistical design of the device, and implementing using microfluidics.

• We derived the posterior Cramér-Rao bounds on the errors in estimating the target concentrations, for uniform or sparse profiles.

• Estimated the unknown imaging parameters using maximum-likelihood.

• Showed quantitatively the effects of distance and temperature on the imaging performance.

• Computed the optimal distance between the microspheres for a given temperature.
Future Work

• Implement the proposed position-encoded 3D microarray:
  – Use minimal microsphere distance.
  – Optimize microfluidic hydrodynamic trapping.
  – Employ chambers with micromechanical valves for position encoding.
  – Load chambers with respective microspheres sensitive to specific targets.

• Derive tighter bounds for low SNR (sparse shell models).
Other Research
Topics

- cDNA microarray image segmentation.
- Sparse gene regulatory network (GRN) estimation.
- Gene reachability analysis.
cDNA Microarray Image Segmentation
Overview

**Goal:** Segment and estimate gene signals for the noisy cDNA microarray images.

**Research [14]:**
- Model cDNA microarray spots using parametric ellipses and circles.
- Segment the foreground signals and estimate the noisy spot images using the Gibbs sampling based Markov Chain Monte Carlo (MCMC) method.

**Clinical application:** Identify diseases by detecting gene signals of the protozoan human gut parasite *Entamoeba histolytica*.

---

Results and Future Work

(a) Noisy cDNA microarray spot image. (b) Estimation result using MCMC.

Spots were printed by the Washington University Microarray Core Facility [15].

Future work:

• Develop a fast algorithm.
• Apply our analysis to clinical data and experiments.

Sparse GRN Estimation
Overview

Goal: Estimate sparse GRN from time-series microarray datasets.

Research [16]:

• Estimate regulatory relationships between genes using a Bayesian linear regression method [17] for sparse parameter vectors.
• Obtain a consensus GRN using our proposed method, a correlation-coefficient method [18], and a database method [19].

Clinical applications: Diagnose diseases and discover drugs.

Results and Future Work

• **Result:** Demonstrated that our proposed method outperforms or performs competitively with NIR (TSNI) [20], [21] and BANJO [22], the two well-established approaches in the community.

• **Result:** Generated four biologically meaningful and consensus sub-networks for a human buffy-coat microarray expression profile dataset of ventilator-associated pneumonia (VAP).

• **Future work:** Analyze the VAP sub-networks using conventional molecular analysis.


Gene Reachability Analysis
Overview

Goal: Analyze gene reachability in complex gene co-expression (CoE) networks.

Research [23]:

• Compute the average connections per gene in a gene CoE network, inspired by the Google Page-Rank algorithm [24].
• Modify the Google Page-Rank algorithm [24], based on this average.
• Compute this average as eight for human and three to seven for yeast. (These numbers agree well with other published results [25], [26].)
• Analyze the gene reachability in gene CoE networks, and cluster genes.

Clinical application: Select genes.

Acknowledgement

- **Advisor:** Prof. Arye Nehorai (ESE, WUSTL)

- **Committee Members:**
  - Prof. Hiro Mukai (ESE, WUSTL)
  - Prof. Jr-Shin Li (ESE, WUSTL)
  - Prof. Richard Martin Arthur (ESE, WUSTL)
  - Prof. Dibyen Majumdar (Stats, UIC)
  - Prof. Samuel Achileefu (Radiology, WUSTL)

- **Collaborators:**
  - Dr. Zhenyu Li (EE, Caltech)
  - Dr. Paul Davis (Biology, U Penn)
  - Prof. Samuel L. Stanley (Stoney Brook)
  - Dr. J. Perren Cobb (Anesthesia, Critical Care and Pain Medicine, MGH)
  - Mr. William Schierding (Genome Center, WUSTL)
  - Prof. Weixiong Zhang (CSE, WUSTL)
Publications

Book chapter:


Journal papers:


Conference papers:


Thank You!