Reconstruction of small molecular structures using cryo-EM

Tamir Bendory

November 16, 2022

Electrical Engineering, Tel Aviv University
https://www.tau.ac.il/~bendory

Outline

1. Introduction
2. Autocorrelation analysis
3. Approximate expectation-maximization

Outline

1. Introduction
2. Autocorrelation analysis
3. Approximate expectation-maximization

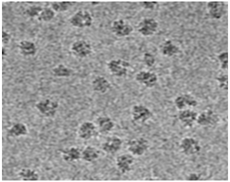
Small molecules and SNR

Common belief: Small molecules cannot be reconstructed using cryo-EM.

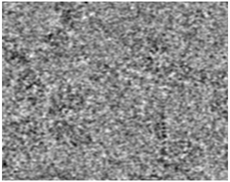
Small molecules and SNR

Common belief: Small molecules cannot be reconstructed using cryo-EM.

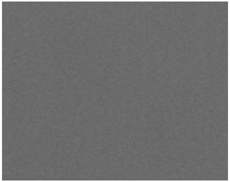
Why? Small molecular structures induce low SNR



EMPIAR 10028
4MDa



EMPIAR 10061
465 KDa

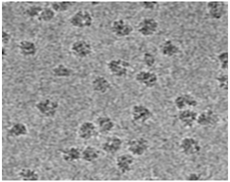


EMPIAR 10249
82KDa

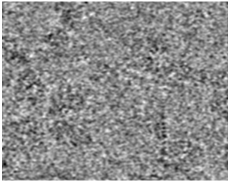
Small molecules and SNR

Common belief: Small molecules cannot be reconstructed using cryo-EM.

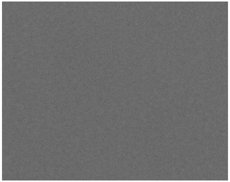
Why? Small molecular structures induce low SNR



EMPIAR 10028
4MDa



EMPIAR 10061
465 KDa



EMPIAR 10249
82KDa

Reasoning:

small molecules $\Rightarrow$ low SNR $\Rightarrow$ detection fails $\Rightarrow$ reconstruction fails

Can we estimate small molecules?

For example, Richard Henderson claimed that below ~ 40 kDa, detection (and thus recovery) is impossible [Henderson, '95, ~ 1250 citations].

Can we estimate small molecules?

For example, Richard Henderson claimed that below ~ 40 kDa, detection (and thus recovery) is impossible [Henderson, '95, ~ 1250 citations].

Joachim Frank voices a similar observation in his 2017 Nobel Prize lecture.

Can we estimate small molecules?

For example, Richard Henderson claimed that below ~ 40 kDa, detection (and thus recovery) is impossible [Henderson, '95, ~ 1250 citations].

Joachim Frank voices a similar observation in his 2017 Nobel Prize lecture.

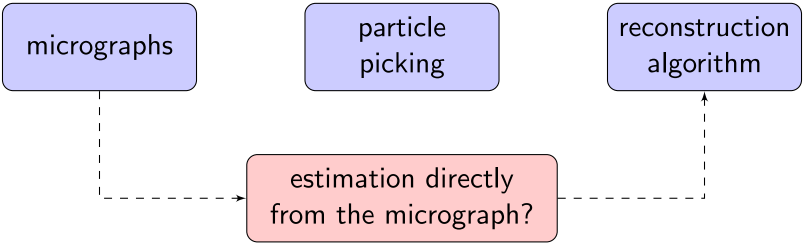
But there is a gap: If reconstruction is possible without detection (particle picking), then the impossibility of detection does not necessarily translate into impossibility of reconstruction.

Can we estimate small molecules?

For example, Richard Henderson claimed that below $\sim$ 40 kDa, detection (and thus recovery) is impossible [Henderson, '95, $\sim$ 1250 citations].

Joachim Frank voices a similar observation in his 2017 Nobel Prize lecture.

But there is a gap: If reconstruction is possible without detection (particle picking), then the impossibility of detection does not necessarily translate into impossibility of reconstruction.



Estimation theory perspective

- Assume we represent the 3-D structure with L parameters and collect N particle projections (observations).

Estimation theory perspective

- Assume we represent the 3-D structure with L parameters and collect N particle projections (observations).
- Each particle is associated with 5 pose parameters: 3-D rotation and 2-D location.

Estimation theory perspective

- Assume we represent the 3-D structure with L parameters and collect N particle projections (observations).
- Each particle is associated with 5 pose parameters: 3-D rotation and 2-D location.
- If we aim to estimate the structure and the pose parameters (as in older cryo-EM algorithms), the number of parameters is $L + 5N$, namely, increases linearly with the number of projections.

Estimation theory perspective

- Assume we represent the 3-D structure with L parameters and collect N particle projections (observations).
- Each particle is associated with 5 pose parameters: 3-D rotation and 2-D location.
- If we aim to estimate the structure and the pose parameters (as in older cryo-EM algorithms), the number of parameters is $L + 5N$, namely, increases linearly with the number of projections.
- In this case, the existence of a consistent estimator is not guaranteed.

  Examples:

Estimation theory perspective

- Assume we represent the 3-D structure with L parameters and collect N particle projections (observations).
- Each particle is associated with 5 pose parameters: 3-D rotation and 2-D location.
- If we aim to estimate the structure and the pose parameters (as in older cryo-EM algorithms), the number of parameters is $L + 5N$, namely, increases linearly with the number of projections.
- In this case, the existence of a consistent estimator is not guaranteed.

  Examples:

  - Neyman-Scott paradox

Estimation theory perspective

- Assume we represent the 3-D structure with L parameters and collect N particle projections (observations).
- Each particle is associated with 5 pose parameters: 3-D rotation and 2-D location.
- If we aim to estimate the structure and the pose parameters (as in older cryo-EM algorithms), the number of parameters is $L + 5N$, namely, increases linearly with the number of projections.
- In this case, the existence of a consistent estimator is not guaranteed.

  Examples:
  - Neyman-Scott paradox
  - The Cramer-Rao bound of multi-image alignment is proportional to the noise level, and independent of the number of observations [Aguerrebere et al., '16]

Estimation theory perspective

- **Solution**: Marginalize over the pose (nuisance) parameters, and estimate only the fixed number of L parameters describing the structure. Maximum likelihood is consistent!

Estimation theory perspective

- **Solution**: Marginalize over the pose (nuisance) parameters, and estimate only the fixed number of L parameters describing the structure. Maximum likelihood is consistent!

- Note that current approaches in cryo-EM are hybrid: they marginalize over the rotations, but estimate the locations. Overall, these methods estimate $2N + L$ parameters and thus are not necessarily consistent. In particular, they cannot work at very low SNR.

Estimation theory perspective

- **Solution**: Marginalize over the pose (nuisance) parameters, and estimate only the fixed number of L parameters describing the structure. Maximum likelihood is consistent!

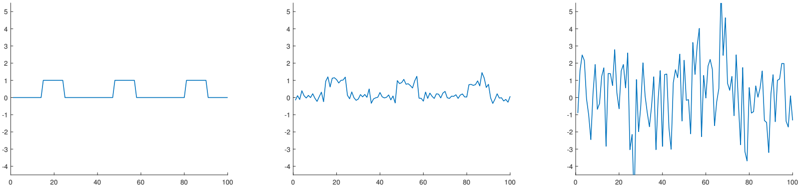
- Note that current approaches in cryo-EM are hybrid: they marginalize over the rotations, but estimate the locations. Overall, these methods estimate $2N + L$ parameters and thus are not necessarily consistent. In particular, they cannot work at very low SNR.

- We will develop methods to marginalize over all pose parameters, allowing estimation in extremely low SNR.

Simplified model for cryo-EM (multi-target detection)

Problem: Multiple occurrences of x are embedded at random locations in a noisy measurement y

Goal: Estimating x from y (the locations are nuisance variables)

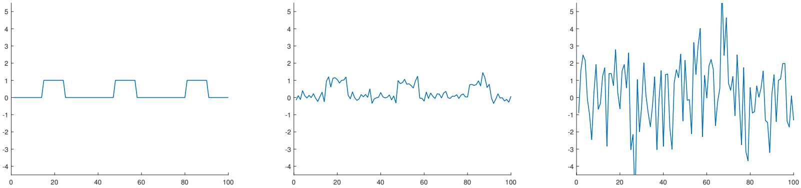


(a) $\sigma = 0$ (b) $\sigma = 0.2$ (c) $\sigma = 2$

Simplified model for cryo-EM (multi-target detection)

Problem: Multiple occurrences of x are embedded at random locations in a noisy measurement y

Goal: Estimating x from y (the locations are nuisance variables)



(a) $\sigma = 0$ (b) $\sigma = 0.2$ (c) $\sigma = 2$

Estimation in low SNR:

- Autocorrelation analysis
- Approximate expectation-maximization

Outline

Autocorrelation analysis

Suppose that the distribution of y is parametrized by x. The goal is to estimate x from y.

Recipe:

1. Derive the expected autocorrelations
2. Estimate the autocorrelations from the data
3. Solve the (polynomial) system of equations

$$a_y^1 = \frac{1}{N}\sum_i y[i] \approx p_1(x)$$

$$a_y^2[\ell] = \frac{1}{N}\sum_i y[i]y[i+\ell] \approx p_2(x)$$

$$a_y^3[\ell_1, \ell_2] = \frac{1}{N}\sum_i y[i]y[i+\ell_1]y[i+\ell_2] \approx p_3(x)$$

Autocorrelation analysis

Suppose that the distribution of y is parametrized by x. The goal is to estimate x from y.

Recipe:

1. Derive the expected autocorrelations
2. Estimate the autocorrelations from the data
3. Solve the (polynomial) system of equations

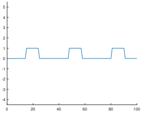
$$a_y^1 = \frac{1}{N}\sum_i y[i] \approx p_1(x)$$

$$a_y^2[\ell] = \frac{1}{N}\sum_i y[i]y[i+\ell] \approx p_2(x)$$

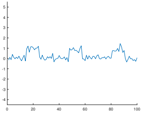
$$a_y^3[\ell_1,\ell_2] = \frac{1}{N}\sum_i y[i]y[i+\ell_1]y[i+\ell_2] \approx p_3(x)$$

Properties: Simple, requires only one pass over the data, parallelizable, consistent, not statistically efficient

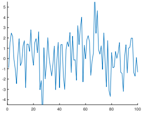
Autocorrelation analysis for multi-target detection



(a) $\sigma = 0$



(b) $\sigma = 0.2$



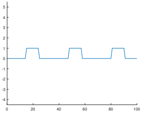
(c) $\sigma = 2$

If any two signals are separated by at least $(L-1)$ entries, then:

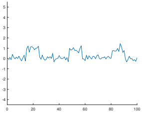
$$\lim_{N\to\infty} a_y^q = \gamma a_x^q, \quad q = 1, 2, 3, \ldots,$$

where $\gamma \in [0, 1]$ is a density parameter.

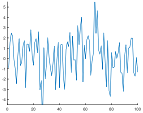
Autocorrelation analysis for multi-target detection



(a) $\sigma = 0$



(b) $\sigma = 0.2$



(c) $\sigma = 2$

If any two signals are separated by at least $(L-1)$ entries, then:

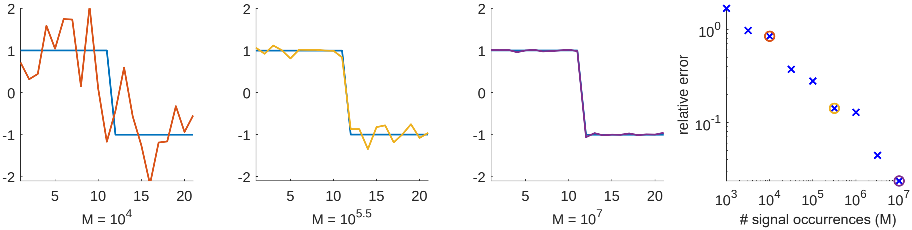
$$\lim_{N\to\infty} a_y^q = \gamma a_x^q, \quad q = 1, 2, 3, \ldots,$$

where $\gamma \in [0, 1]$ is a density parameter.

Theorem (informal)

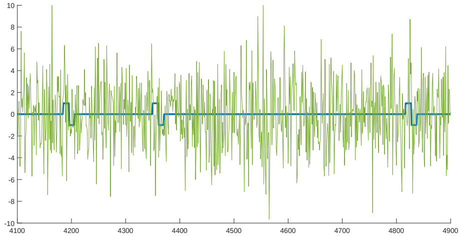
The signal x is determined uniquely from a_y^3. Namely, the signal x is determined, in any SNR level, without intermediate detection, if $N \gg \sigma^6$.

Numerical experiments



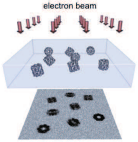
Details:

γ and σ are unknown
Recovery by least squares
$\sigma = 3$
Micrograph size $= 10M(2L - 1)$
Relative error $\gamma = 4.8\%, 4\%, 1.2\%$



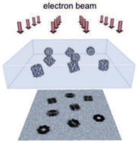
Application to cryo-EM

- We aim at estimating the 3-D volume directly from the micrograph.



Application to cryo-EM

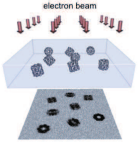
- We aim at estimating the 3-D volume directly from the micrograph.



- An L-bandlimited 3-D volume is described by $\sim L^3$ parameters.

Application to cryo-EM

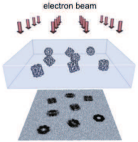
- We aim at estimating the 3-D volume directly from the micrograph.



- An L-bandlimited 3-D volume is described by $\sim L^3$ parameters.

- We scan the micrographs with a sliding window of size $L \times L$.

Application to cryo-EM



- We aim at estimating the 3-D volume directly from the micrograph.
- An L-bandlimited 3-D volume is described by $\sim L^3$ parameters.
- We scan the micrographs with a sliding window of size $L \times L$.
- We compute the first three autocorrelations of each window with respect to the center point and average over all windows.

Application to cryo-EM

- The autocorrelations of the micrographs converge to scaled versions of the volume's autocorrelations:

$$\lim_{N\to\infty} a_y^1 = \gamma \left\langle a^1_{P_\omega(x)} \right\rangle_{\omega\in SO(3)},$$
$$\lim_{N\to\infty} a_y^2[\ell_1,\ell_2] = \gamma \left\langle a^2_{P_\omega(x)}[\ell_1,\ell_2] \right\rangle_{\omega\in SO(3)},$$
$$\lim_{N\to\infty} a_y^3[\ell_1,\ell_2;\ell_3,\ell_4] = \gamma \left\langle a^3_{P_\omega(x)}[\ell_1,\ell_2;\ell_3,\ell_4] \right\rangle_{\omega\in SO(3)}.$$

Application to cryo-EM

- The autocorrelations of the micrographs converge to scaled versions of the volume's autocorrelations:

$$\lim_{N\to\infty} a_y^1 = \gamma \left\langle a^1_{P_\omega(x)} \right\rangle_{\omega\in SO(3)},$$
$$\lim_{N\to\infty} a_y^2[\ell_1,\ell_2] = \gamma \left\langle a^2_{P_\omega(x)}[\ell_1,\ell_2] \right\rangle_{\omega\in SO(3)},$$
$$\lim_{N\to\infty} a_y^3[\ell_1,\ell_2;\ell_3,\ell_4] = \gamma \left\langle a^3_{P_\omega(x)}[\ell_1,\ell_2;\ell_3,\ell_4] \right\rangle_{\omega\in SO(3)}.$$

- No detection is required, and thus small molecular are (in principle) within reach.

Application to cryo-EM

- The autocorrelations of the micrographs converge to scaled versions of the volume's autocorrelations:

$$\lim_{N\to\infty} a_y^1 = \gamma \left\langle a^1_{P_\omega(x)} \right\rangle_{\omega\in SO(3)},$$
$$\lim_{N\to\infty} a_y^2[\ell_1,\ell_2] = \gamma \left\langle a^2_{P_\omega(x)}[\ell_1,\ell_2] \right\rangle_{\omega\in SO(3)},$$
$$\lim_{N\to\infty} a_y^3[\ell_1,\ell_2;\ell_3,\ell_4] = \gamma \left\langle a^3_{P_\omega(x)}[\ell_1,\ell_2;\ell_3,\ell_4] \right\rangle_{\omega\in SO(3)}.$$

- No detection is required, and thus small molecular are (in principle) within reach.
- The third-order autocorrelation contains $\sim L^3$ independent cubic equations (rather than L^4) that can be related to the $\sim L^3$ coefficients of the volume.

Application to cryo-EM

- We estimate the volume's coefficients and γ by least-squares.

Application to cryo-EM

- We estimate the volume's coefficients and γ by least-squares.

- All technical details appear in:

  "Toward single particle reconstruction without particle picking: Breaking the detection limit". T. Bendory, N. Boumal, W. Leeb, E. Levin, A. Singer. Available at *arXiv preprint arXiv:1810.00226*.

Application to cryo-EM

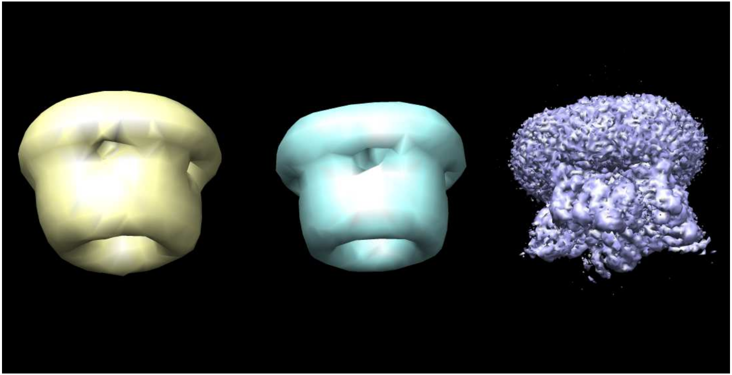
- We estimate the volume's coefficients and γ by least-squares.

- All technical details appear in:

  "Toward single particle reconstruction without particle picking: Breaking the detection limit". T. Bendory, N. Boumal, W. Leeb, E. Levin, A. Singer. Available at *arXiv preprint arXiv:1810.00226*.

- Unfortunately, the mapping is highly ill-conditioned, preventing stable recovery from noisy data.

Recovery from clean autocorrelations



estimated structure (yellow), low-resolution structure (blue), high-resolution structure (purple)

TRPV1, the low-resolution molecule ($L = 5$) was down-sampled from 192^3 to 20^3 pixels

Perspective

- A possible solution to the ill-conditioning: Prior! (Future work)

Perspective

- A possible solution to the ill-conditioning: Prior! (Future work)
- Further ingredients can be readily included in the model, such as CTF and densely packed particle images [Kreymer and Bendory, '22].

Perspective

- A possible solution to the ill-conditioning: Prior! (Future work)
- Further ingredients can be readily included in the model, such as CTF and densely packed particle images [Kreymer and Bendory, '22].
- A recent paper [Lan et al., '22] applied the technique to a random conical tilt reconstruction (averaging only over in-plane rotations) demonstrating improved numerical results.

Perspective

- A possible solution to the ill-conditioning: Prior! (Future work)
- Further ingredients can be readily included in the model, such as CTF and densely packed particle images [Kreymer and Bendory, '22].
- A recent paper [Lan et al., '22] applied the technique to a random conical tilt reconstruction (averaging only over in-plane rotations) demonstrating improved numerical results.
- The method is highly efficient, and thus, perhaps, it can be used for additional tasks when the SNR is very low. For example, to generate templates for particle picking.

Perspective

- A possible solution to the ill-conditioning: Prior! (Future work)
- Further ingredients can be readily included in the model, such as CTF and densely packed particle images [Kreymer and Bendory, '22].
- A recent paper [Lan et al., '22] applied the technique to a random conical tilt reconstruction (averaging only over in-plane rotations) demonstrating improved numerical results.
- The method is highly efficient, and thus, perhaps, it can be used for additional tasks when the SNR is very low. For example, to generate templates for particle picking.
- Perhaps we should consider an alternative computational method?

Outline

1. Introduction
2. Autocorrelation analysis
3. Approximate expectation-maximization

Expectation-maximization (EM)

- EM is a general algorithm to find a local maximum of a likelihood function (or posterior distribution) with nuisance variables.

Expectation-maximization (EM)

- EM is a general algorithm to find a local maximum of a likelihood function (or posterior distribution) with nuisance variables.
- I will focus on models of the form:

$$y_i = L_{\theta_i} x + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2 I), \quad i = 1, \ldots, N,$$

where L_θ is a linear operator acting on the signal x, parameterized by a random variable $\theta \in \Theta$.

Expectation-maximization (EM)

- EM is a general algorithm to find a local maximum of a likelihood function (or posterior distribution) with nuisance variables.
- I will focus on models of the form:

$$y_i = L_{\theta_i} x + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2 I), \quad i = 1, \dots, N,$$

  where L_θ is a linear operator acting on the signal x, parameterized by a random variable $\theta \in \Theta$.
- The goal is to estimate x from $\mathbf{y} := y_1, \dots, y_N$, where $\theta_1, \dots, \theta_N$ are the nuisance variables.

Expectation-maximization (EM)

- EM is a general algorithm to find a local maximum of a likelihood function (or posterior distribution) with nuisance variables.
- I will focus on models of the form:

$$y_i = L_{\theta_i} x + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2 I), \quad i = 1, \ldots, N,$$

  where L_θ is a linear operator acting on the signal x, parameterized by a random variable $\theta \in \Theta$.
- The goal is to estimate x from $\mathbf{y} := y_1, \ldots, y_N$, where $\theta_1, \ldots, \theta_N$ are the nuisance variables.
- The likelihood function is given by

$$p(\mathbf{y}; x) = \frac{1}{(2\pi\sigma^2)^{(M/2)}} \prod_{i=1}^{N} \sum_{\theta_\ell \in \Theta} p(\theta_\ell) e^{-\frac{1}{2\sigma^2} \|y_i - L_{\theta_\ell} x\|}$$

Expectation-maximization (EM)

- We first write the Q function:

$$Q(x|x_t) = \mathbb{E}_{\theta|\mathbf{y},x_t}\left\{\log p(x|\mathbf{y},\theta)\right\} \propto \sum_{i=1}^{N}\sum_{\theta_\ell\in\Theta} w_{i,\ell}\|y_i - L_{\theta_\ell}x\|^2,$$

where

$$w_{i,\ell} = p(\theta=\theta_\ell|y_i,x_t) \propto e^{-\frac{1}{2\sigma^2}\|y_i - L_{\theta_\ell}x_t\|}.$$

Expectation-maximization (EM)

- We first write the Q function:

$$Q(x|x_t) = \mathbb{E}_{\theta|\mathbf{y},x_t}\left\{\log p(x|\mathbf{y},\theta)\right\} \propto \sum_{i=1}^{N}\sum_{\theta_\ell\in\Theta} w_{i,\ell}\|y_i - L_{\theta_\ell}x\|^2,$$

where

$$w_{i,\ell} = p(\theta=\theta_\ell|y_i,x_t) \propto e^{-\frac{1}{2\sigma^2}\|y_i - L_{\theta_\ell}x_t\|}.$$

- We apply two steps iteratively:

Expectation-maximization (EM)

- We first write the Q function:

$$Q(x|x_t) = \mathbb{E}_{\theta|\mathbf{y},x_t}\left\{\log p(x|\mathbf{y},\theta)\right\} \propto \sum_{i=1}^{N}\sum_{\theta_\ell\in\Theta} w_{i,\ell}\|y_i - L_{\theta_\ell}x\|^2,$$

  where

$$w_{i,\ell} = p(\theta=\theta_\ell|y_i,x_t) \propto e^{-\frac{1}{2\sigma^2}\|y_i - L_{\theta_\ell}x_t\|}.$$

- We apply two steps iteratively:
  - In the E-step, we compute the weights $w_{i,\ell}$.

Expectation-maximization (EM)

- We first write the Q function:

$$Q(x|x_t) = \mathbb{E}_{\theta|\mathbf{y},x_t}\left\{\log p(x|\mathbf{y},\theta)\right\} \propto \sum_{i=1}^{N}\sum_{\theta_\ell\in\Theta} w_{i,\ell}\|y_i - L_{\theta_\ell}x\|^2,$$

  where

$$w_{i,\ell} = p(\theta=\theta_\ell|y_i,x_t) \propto e^{-\frac{1}{2\sigma^2}\|y_i - L_{\theta_\ell}x_t\|}.$$

- We apply two steps iteratively:
  - In the E-step, we compute the weights $w_{i,\ell}$.
  - In the M-step, we update $x_{t+1} = \arg\max Q(x|x_t)$ by solving a linear system of equations.

EM for cryo-EM

- EM was first introduced to cryo-EM by Fred Sigworth in 1998, and is by now the most popular algorithm for structure refinement.

EM for cryo-EM

- EM was first introduced to cryo-EM by Fred Sigworth in 1998, and is by now the most popular algorithm for structure refinement.

- The standard strategy is to first locate and extract the particle projections, and then apply EM, where Θ is the space of 3-D rotations and small 2-D translations.

EM for cryo-EM

- EM was first introduced to cryo-EM by Fred Sigworth in 1998, and is by now the most popular algorithm for structure refinement.

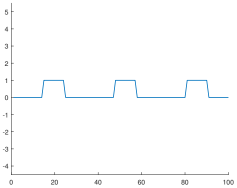
- The standard strategy is to first locate and extract the particle projections, and then apply EM, where Θ is the space of 3-D rotations and small 2-D translations.

- However, if the molecular structure is small, the SNR drops, and we cannot locate the particle images reliably. Thus, this paradigm fails.

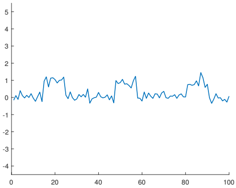
EM for cryo-EM

- EM was first introduced to cryo-EM by Fred Sigworth in 1998, and is by now the most popular algorithm for structure refinement.
- The standard strategy is to first locate and extract the particle projections, and then apply EM, where Θ is the space of 3-D rotations and small 2-D translations.
- However, if the molecular structure is small, the SNR drops, and we cannot locate the particle images reliably. Thus, this paradigm fails.
- Can we apply EM for structure recovery directly from the micrograph?

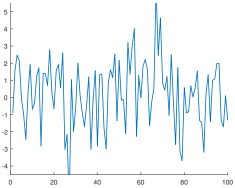
- Recall the multi target detection model, where multiple copies of a target signal occur at unknown locations in a long noisy measurement.



(a) $\sigma = 0$



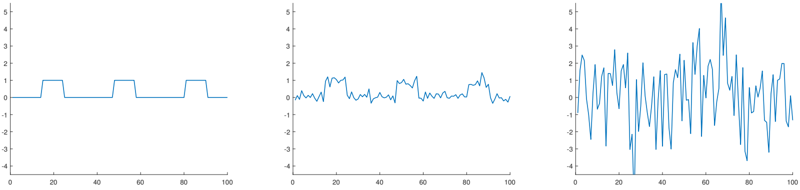
(b) $\sigma = 0.2$



(c) $\sigma = 2$

Approximate EM for multi-target detection

- Recall the multi target detection model, where multiple copies of a target signal occur at unknown locations in a long noisy measurement.

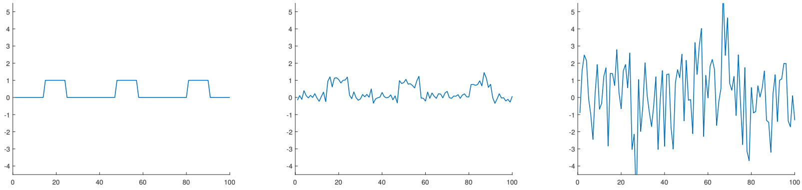


(a) $\sigma = 0$ (b) $\sigma = 0.2$ (c) $\sigma = 2$

- Assuming we know the number of signal occurrences K, the E-step requires computing probabilities for all $\sim \binom{N}{K}$ possible configurations.

Approximate EM for multi-target detection

- Recall the multi target detection model, where multiple copies of a target signal occur at unknown locations in a long noisy measurement.



(a) $\sigma = 0$ (b) $\sigma = 0.2$ (c) $\sigma = 2$

- Assuming we know the number of signal occurrences K, the E-step requires computing probabilities for all $\sim \binom{N}{K}$ possible configurations.

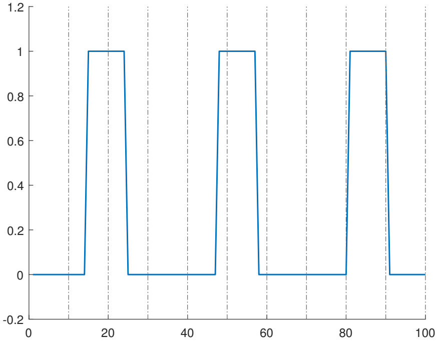
- Therefore, EM is intractable.

Approximate EM for multi-target detection

- In the approximate EM, we divide the measurement into N/L non-overlapping patches, and assume they are independent.

Approximate EM for multi-target detection

- In the approximate EM, we divide the measurement into N/L non-overlapping patches, and assume they are independent.
- Each patch may contain a full signal, no signal, or a part of the signal.



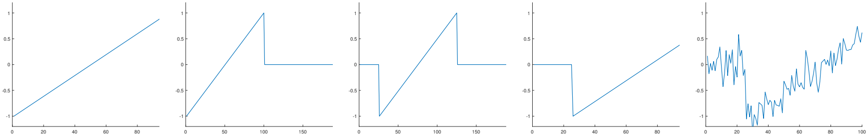
Approximate EM for multi-target detection

We wish to maximize the approximate likelihood function $\prod_i p(y_i|x)$ where

$$y_i = CR_{\theta_i}Zx + \varepsilon_i$$

$$\text{patch}_i = \text{cropping} \circ \text{circular shift}_i \circ \text{padding} \circ x + \varepsilon_i$$

Shift by 25 entries:



(a) signal (b) padding (c) shifting (d) cropping (e) patch

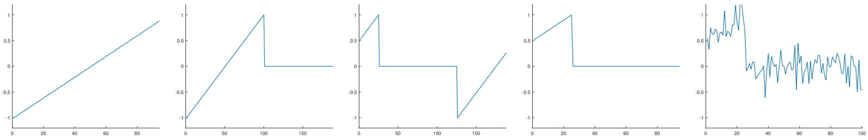
Approximate EM for multi-target detection

We wish to maximize the approximate likelihood function $\prod_i p(y_i|x)$ where

$$y_i = CR_{\theta_i}Zx + \varepsilon_i$$

$$\text{patch}_i = \text{cropping} \circ \text{circular shift}_i \circ \text{padding} \circ x + \varepsilon_i$$

Shift by 125 entries:



(a) signal (b) padding (c) shifting (d) cropping (e) patch

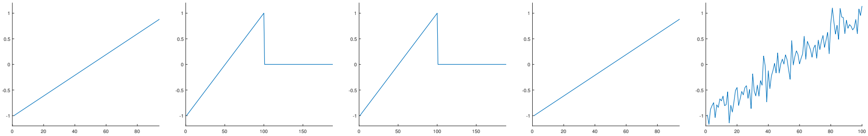
Approximate EM for multi-target detection

We wish to maximize the approximate likelihood function $\prod_i p(y_i|x)$ where

$$y_i = CR_{\theta_i}Zx + \varepsilon_i$$

$$\text{patch}_i = \text{cropping} \circ \text{circular shift}_i \circ \text{padding} \circ x + \varepsilon_i$$

Shift by 0 entries (full signal):



(a) signal (b) padding (c) shifting (d) cropping (e) patch

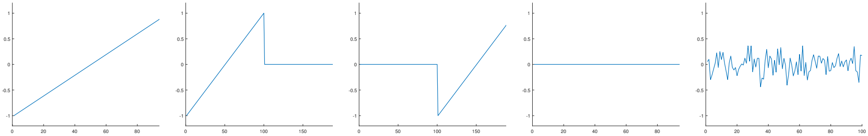
Approximate EM for multi-target detection

We wish to maximize the approximate likelihood function $\prod_i p(y_i|x)$ where

$$y_i = CR_{\theta_i} Zx + \varepsilon_i$$

$$\text{patch}_i = \text{cropping} \circ \text{circular shift}_i \circ \text{padding} \circ x + \varepsilon_i$$

Shift by 100 entries (no signal):



(a) signal (b) padding (c) shifting (d) cropping (e) patch

Approximate EM for multi-target detection

- We then apply EM to the model

$$y_i = CR_{\theta_i}Zx + \varepsilon_i = L_{\theta_i}x + \varepsilon_i,$$

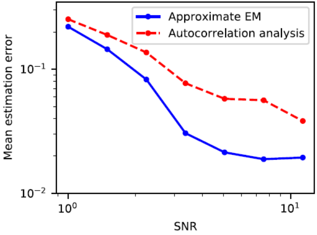
assuming all observations are independent, where the circular shifts are the nuisance variables.

Approximate EM for multi-target detection

- We then apply EM to the model

$$y_i = CR_{\theta_i}Zx + \varepsilon_i = L_{\theta_i}x + \varepsilon_i,$$

  assuming all observations are independent, where the circular shifts are the nuisance variables.
- An example from [Kreymer et al., '22]:

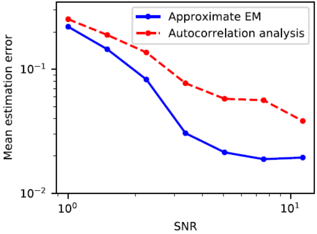


Approximate EM for multi-target detection

- We then apply EM to the model

$$y_i = CR_{\theta_i} Zx + \varepsilon_i = L_{\theta_i} x + \varepsilon_i,$$

  assuming all observations are independent, where the circular shifts are the nuisance variables.
- An example from [Kreymer et al., '22]:



- The statistical model can be extended to account for densely packed signals, where a patch may contain two signals [Lan et al., '20].

- We model a micrograph by

$$\mathcal{I}[\vec{\ell}] = \sum_i P_{\omega_i}(x)[\vec{\ell} - \vec{\ell}_i] + \varepsilon[\vec{\ell}],$$

  where $P_{\omega_i}(x)$ denotes the tomographic projection obtained from viewing direction $\omega_i \in SO(3)$.

Approximate EM for cryo-EM

- We model a micrograph by

$$\mathcal{I}[\vec{\ell}] = \sum_i P_{\omega_i}(x)[\vec{\ell} - \vec{\ell}_i] + \varepsilon[\vec{\ell}],$$

  where $P_{\omega_i}(x)$ denotes the tomographic projection obtained from viewing direction $\omega_i \in SO(3)$.

- We assume that the Fourier transform of the volume $\hat{x}$ may be finitely expanded by the Fourier-Bessel expansion

$$\hat{x}(ck, \theta, \varphi) = \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} \sum_{s=1}^{S(\ell)} x_{\ell,m,s} Y_\ell^m(\theta, \varphi) j_{\ell,s}(k), \quad k \leq 1,$$

  where c is the bandlimit, Y_ℓ^m are spherical harmonics, and $j_{\ell,s}$ is the normalized spherical Bessel function.

Approximate EM for cryo-EM

- Then, each projection image is equal to

$$P_\omega(\hat{x})(ck,\varphi) = \sum_{\ell,m,m',s} x_{\ell,m,s} D^{\ell}_{m',m}(\omega) Y_\ell^{m'}\left(\frac{\pi}{2},\varphi\right) j_{\ell,s}(k),$$

- Then, each projection image is equal to

$$P_\omega(\hat{x})(ck,\varphi) = \sum_{\ell,m,m',s} x_{\ell,m,s} D^{\ell}_{m',m}(\omega) Y_\ell^{m'}\left(\frac{\pi}{2},\varphi\right) j_{\ell,s}(k),$$

- Now we can easily modify the principles of the 1-D approximate EM to 2-D.

Approximate EM for cryo-EM

- Then, each projection image is equal to

$$P_\omega(\hat{x})(ck,\varphi) = \sum_{\ell,m,m',s} x_{\ell,m,s} D^{\ell}_{m',m}(\omega) Y_{\ell}^{m'}\left(\frac{\pi}{2},\varphi\right) j_{\ell,s}(k),$$

- Now we can easily modify the principles of the 1-D approximate EM to 2-D.

- In particular, each 2-D patch is modeled as

$$y_i = CR_{\ell_i} Z P_{\omega_i}(x) + \varepsilon_i$$

Approximate EM for cryo-EM

- Then, each projection image is equal to

$$P_{\omega}(\hat{x})(ck, \varphi) = \sum_{\ell,m,m',s} x_{\ell,m,s} D^{\ell}_{m',m}(\omega) Y^{m'}_{\ell}\left(\frac{\pi}{2}, \varphi\right) j_{\ell,s}(k),$$

- Now we can easily modify the principles of the 1-D approximate EM to 2-D.

- In particular, each 2-D patch is modeled as

$$y_i = CR_{\ell_i} Z P_{\omega_i}(x) + \varepsilon_i$$

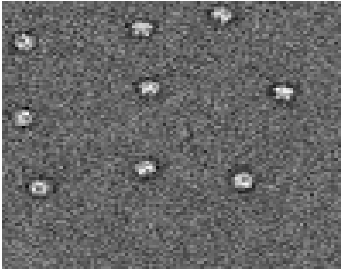
- All technical details appear in a manuscript in preparation by S. Kreymer, A. Singer, and T. Bendory.

Numerical results

- Volumes were downsampled to $11 \times 11 \times 11$ voxels, and expanded to $L = 10$.

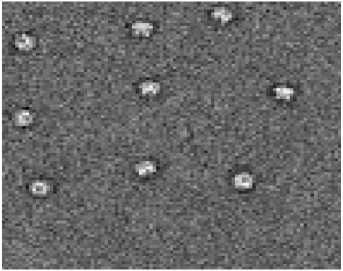
Numerical results

- Volumes were downsampled to $11 \times 11 \times 11$ voxels, and expanded to $L = 10$.
- SNR=3.5 (currently working on data sets with lower SNR levels)



Numerical results

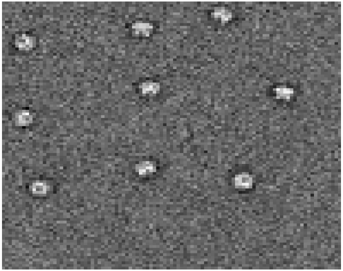
- Volumes were downsampled to $11 \times 11 \times 11$ voxels, and expanded to $L = 10$.
- SNR=3.5 (currently working on data sets with lower SNR levels)



- ~ 120000 projections

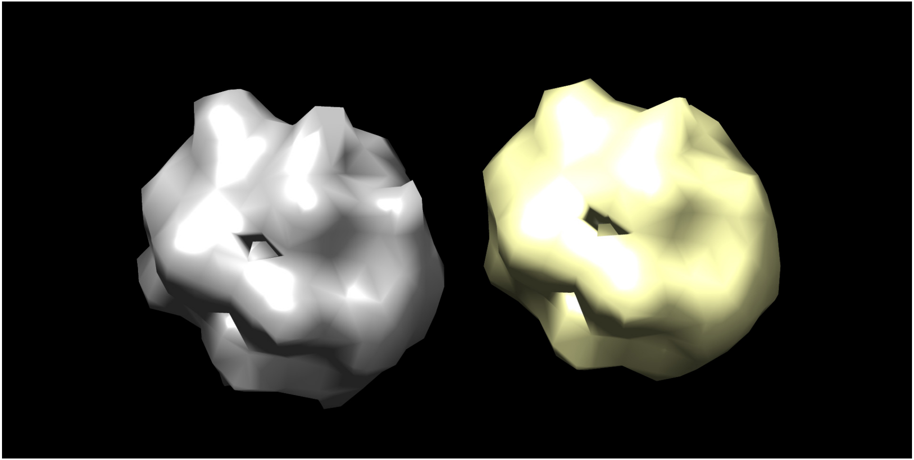
Numerical results

- Volumes were downsampled to $11 \times 11 \times 11$ voxels, and expanded to $L = 10$.
- SNR=3.5 (currently working on data sets with lower SNR levels)



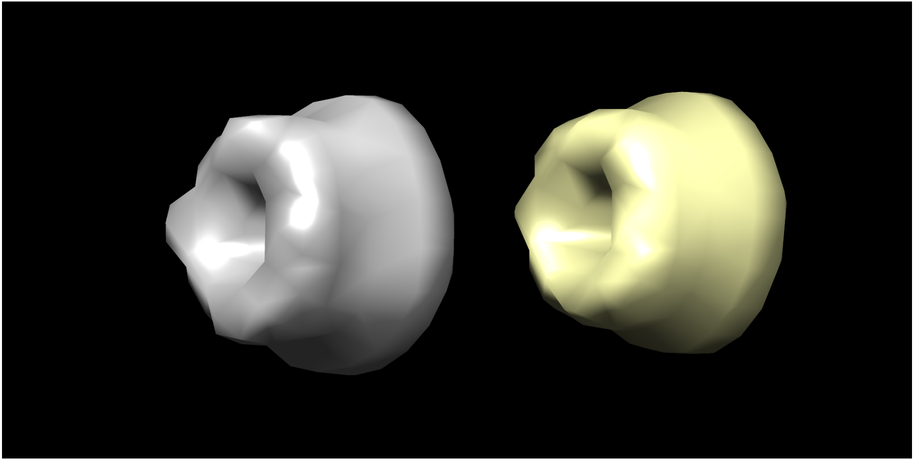
- $\sim$ 120000 projections
- $\sim$ 30 batch EM iterations

Shepp-Logan



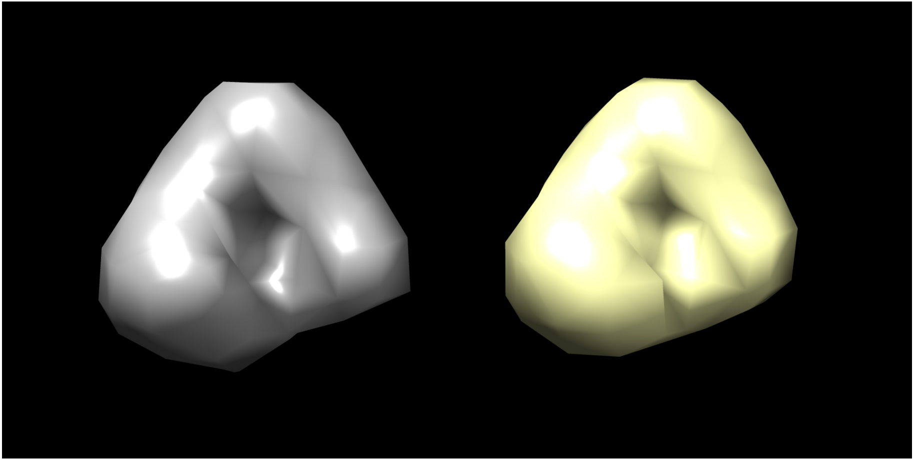
Ground truth in gray, estimate in yellow

TRPV1



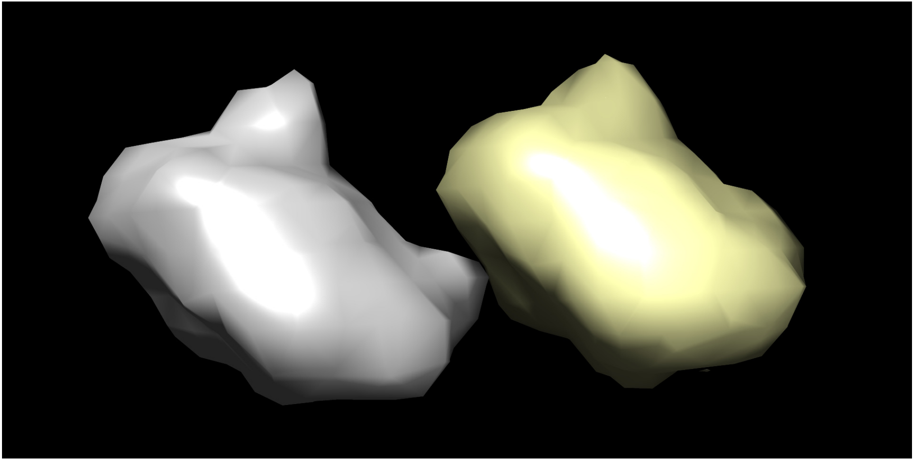
Ground truth in gray, estimate in yellow

Plasmodium falciparum 80S ribosome



Ground truth in gray, estimate in yellow

Bovine Pancreatic Trypsin Inhibitor (BPTI) mutant



Ground truth in gray, estimate in yellow

Conclusions and future work

- We have discussed two methods to recover molecular structures directly from micrographs. We hope it will pave the way to recover small molecular structures using cryo-EM.

Conclusions and future work

- We have discussed two methods to recover molecular structures directly from micrographs. We hope it will pave the way to recover small molecular structures using cryo-EM.

- Autocorrelation analysis is computationally efficient but (currently) provides low-resolution estimates. Next step: designing priors.

Conclusions and future work

- We have discussed two methods to recover molecular structures directly from micrographs. We hope it will pave the way to recover small molecular structures using cryo-EM.
- Autocorrelation analysis is computationally efficient but (currently) provides low-resolution estimates. Next step: designing priors.
- Approximate expectation-maximization provides high resolution recoveries for moderate SNR levels. Next steps: acceleration and designing priors.

Conclusions and future work

- We have discussed two methods to recover molecular structures directly from micrographs. We hope it will pave the way to recover small molecular structures using cryo-EM.
- Autocorrelation analysis is computationally efficient but (currently) provides low-resolution estimates. Next step: designing priors.
- Approximate expectation-maximization provides high resolution recoveries for moderate SNR levels. Next steps: acceleration and designing priors.
- Theoretical analysis: Sample complexity analysis and analysis of the EM iterations.

Conclusions and future work

- We have discussed two methods to recover molecular structures directly from micrographs. We hope it will pave the way to recover small molecular structures using cryo-EM.
- Autocorrelation analysis is computationally efficient but (currently) provides low-resolution estimates. Next step: designing priors.
- Approximate expectation-maximization provides high resolution recoveries for moderate SNR levels. Next steps: acceleration and designing priors.
- Theoretical analysis: Sample complexity analysis and analysis of the EM iterations.
- Alternative computational schemes such as CryoGAN [Gupta et al., '21] and dynamic programming.

Thanks for your attention!