Super-resolution multi-reference alignment

Tamir Bendory (Tel Aviv University, EE)

June 29, 2020

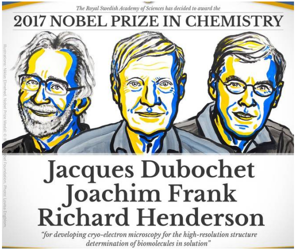
Joint work with:
Ariel Jaffe (Yale University, Applied math)
Will Leeb (University of Minnesota, Math)
Nir Sharon (Tel Aviv University, Applied math)
Amit Singer (Princeton University, Math&PACM)

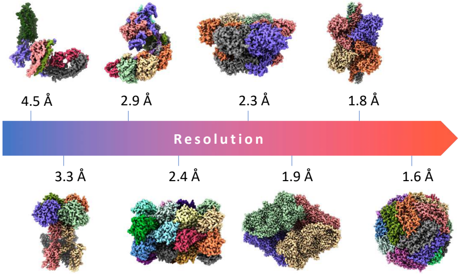
Outline

1. Motivation
2. Problem formulation
3. Main results
4. Computational considerations
5. Future work

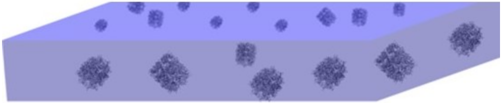
Cryo-electron microscopy

Single particle cryo-electron microscopy (cryo-EM) is an emerging technology for structure determination of biological molecules (e.g., viruses, proteins).

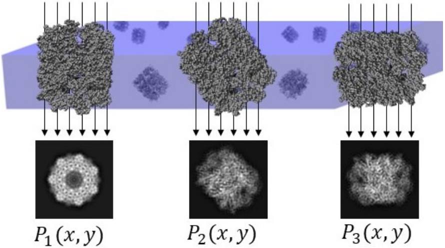




Cryo-EM mathematical model



Cryo-EM mathematical model

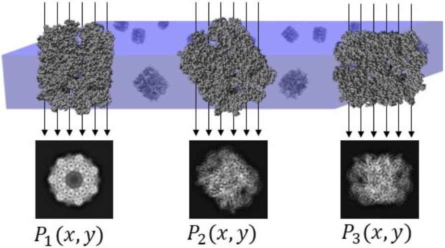


$$P_i = STR_iX + \text{noise}, \qquad R_i \in SO(3)$$

$$P_i = \text{sampling}(\text{projection}(\text{rotation}(X))) + \text{noise}$$

The images of E. coli 50S ribosomal subunit were provided by Dr. Fred Sigworth, Yale Medical School.

Cryo-EM mathematical model



$$P_i = STR_iX + \text{noise}, \qquad R_i \in SO(3)$$

$$P_i = \text{sampling}(\text{projection}(\text{rotation}(X))) + \text{noise}$$

Experimental images:

The images of E. coli 50S ribosomal subunit were provided by Dr. Fred Sigworth, Yale Medical School.

Resolution limits of cryo-EM

"Folk Theorem": Shannon-Nyquist sampling theorem implies that the resolution of any estimate of the 3-D structure $\hat{X}$ is limited by the resolution of the 2-D projection images (dictated by the detectors acquiring the data):

$$\text{Resolution}(\hat{X}) \leq \text{Resolution}(P_i)$$

Resolution limits of cryo-EM

"Folk Theorem": Shannon-Nyquist sampling theorem implies that the resolution of any estimate of the 3-D structure $\hat{X}$ is limited by the resolution of the 2-D projection images (dictated by the detectors acquiring the data):

$$\text{Resolution}(\hat{X}) \leq \text{Resolution}(P_i)$$

Is this correct? **Obviously not.** Shannon-Nyquist concerns a single observation.

Resolution limits of cryo-EM

"Folk Theorem": Shannon-Nyquist sampling theorem implies that the resolution of any estimate of the 3-D structure $\hat{X}$ is limited by the resolution of the 2-D projection images (dictated by the detectors acquiring the data):

$$\text{Resolution}(\hat{X}) \leq \text{Resolution}(P_i)$$

Is this correct? **Obviously not.** Shannon-Nyquist concerns a single observation.

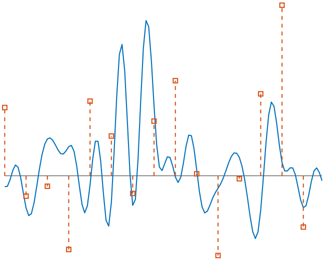
Can the resolution of the estimated 3-D structure surpass the resolution of the 2-D projection images?

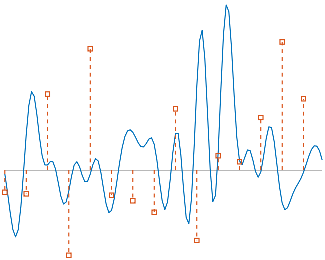
Problem formulation (toy model for cryo-EM)

Problem: Estimate a signal in $x \in \mathbb{R}^M$ from its circularly shifted, sampled, noisy copies

$$y_i = SR_{t_i}x + \varepsilon_i, \quad i = 1, \ldots, N,$$
$$t \sim \text{Uni}[0, M-1], \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I),$$

where S is a sampling operator that selects L equally-spaced samples.





Problem formulation (toy model for cryo-EM)

Problem: Estimate a signal in $x \in \mathbb{R}^M$ from its circularly shifted, sampled, noisy copies

$$y_i = SR_{t_i}x + \varepsilon_i, \quad i = 1, \ldots, N,$$
$$t \sim \text{Uni}[0, M-1], \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I),$$

where S is a sampling operator that selects L equally-spaced samples.

Problem formulation (toy model for cryo-EM)

Problem: Estimate a signal in $x \in \mathbb{R}^M$ from its circularly shifted, sampled, noisy copies

$$y_i = SR_{t_i}x + \varepsilon_i, \quad i = 1, \dots, N,$$
$$t \sim \text{Uni}[0, M-1], \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I),$$

where S is a sampling operator that selects L equally-spaced samples.

We wish to understand how many samples L we need to collect to estimate x accurately as a function of M, N, and σ.

Problem formulation (toy model for cryo-EM)

Problem: Estimate a signal in $x \in \mathbb{R}^M$ from its circularly shifted, sampled, noisy copies

$$y_i = SR_{t_i}x + \varepsilon_i, \quad i = 1, \dots, N,$$
$$t \sim \text{Uni}[0, M-1], \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I),$$

where S is a sampling operator that selects L equally-spaced samples.

We wish to understand how many samples L we need to collect to estimate x accurately as a function of M, N, and σ.

Super-resolution interpretation: We refer to x as the "high-resolution" signal with bandwidth proportional to M, and to $y_1, \dots, y_N$ as the "low-resolution" observations.

Problem formulation (toy model for cryo-EM)

Problem: Estimate a signal in $x \in \mathbb{R}^M$ from its circularly shifted, sampled, noisy copies

$$y_i = SR_{t_i}x + \varepsilon_i, \quad i = 1, \ldots, N,$$
$$t \sim \text{Uni}[0, M-1], \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I),$$

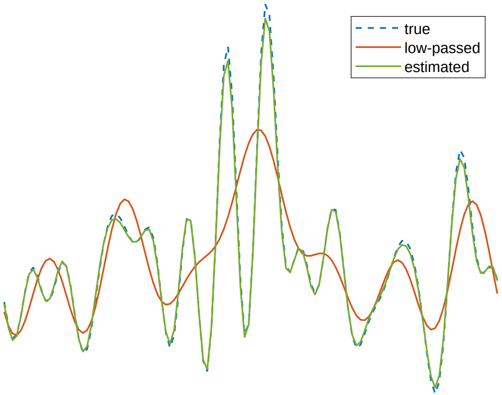
where S is a sampling operator that selects L equally-spaced samples.

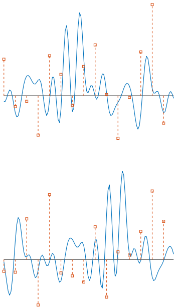
We wish to understand how many samples L we need to collect to estimate x accurately as a function of M, N, and σ.

Super-resolution interpretation: We refer to x as the "high-resolution" signal with bandwidth proportional to M, and to $y_1, \ldots, y_N$ as the "low-resolution" observations.

This problem is an instance of the **multi-reference alignment** model.

Super-resolution example





Sampling rate = (Nyquist rate)/2, SNR = 1, $N = 10^4$

Main result

In the low SNR regime $\sigma \to \infty$, the signal can be identified if:

Main result

In the low SNR regime $\sigma \to \infty$, the signal can be identified if:

- $N/\sigma^6 \to \infty$ (necessary condition even if $L = M$)
- $L \geq \sqrt{6M}$.

Main result

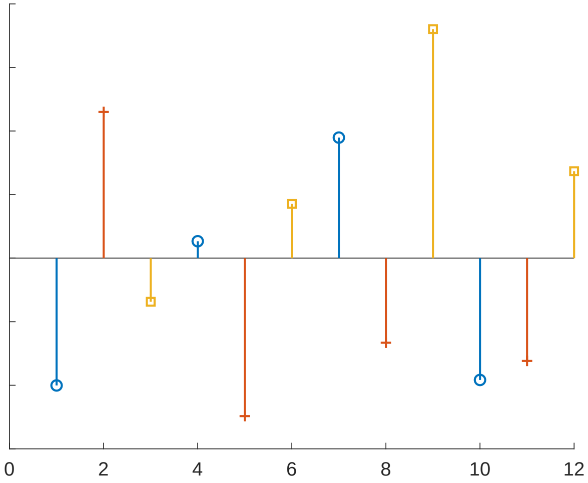
In the low SNR regime $\sigma \to \infty$, the signal can be identified if:

- $N/\sigma^6 \to \infty$ (necessary condition even if $L = M$)

- $L \geq \sqrt{6M}$.

Informally: **one can square the resolution.**

Example: $M = 12, L = 4, K = M/L = 3$



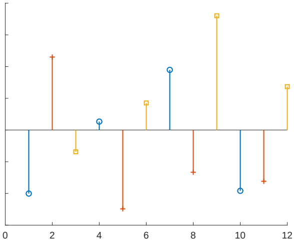
Proof strategy

The model

$$y = SR_t x + \varepsilon \quad t \sim \text{Uni}[0, M-1],$$

is equivalent to

$$y = R_\ell x_k + \varepsilon, \quad \ell \sim \text{Uni}[0, L-1], \quad k \sim \text{Uni}[0, K-1].$$



Proof strategy (Likelihood)

The model

$$y = SR_t x + \varepsilon \quad t \sim \text{Uni}[0, M-1],$$

is equivalent to

$$y = R_\ell x_k + \varepsilon, \quad \ell \sim \text{Uni}[0, L-1], \quad k \sim \text{Uni}[0, K-1].$$

This model is called **heterogeneous multi-reference alignment**.

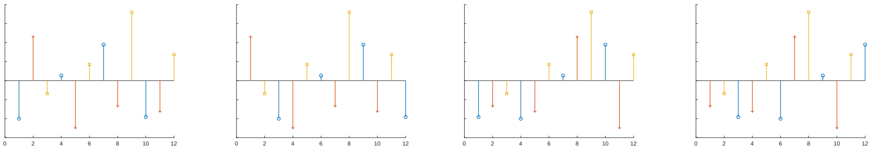
The likelihood function (of a single observation) is then given

$$p(y|x) = \frac{1}{M} \sum_{t=0}^{L-1} \sum_{k=0}^{K-1} \mathcal{N}(R_\ell x_k, \sigma^2 I)$$

Proof strategy (Likelihood)

Likelihood:

$$p(y|x) = \frac{1}{M} \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \mathcal{N}(R_\ell x_k, \sigma^2 I)$$

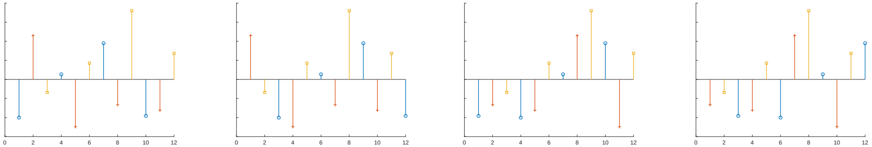


All signals in the orbit $Gx := \{gx | g \in G\}$—where G is a subgroup of the permutation group—have the same likelihood.

Proof strategy (Likelihood)

Likelihood:

$$p(y|x) = \frac{1}{M} \sum_{\ell=0}^{L-1} \sum_{k=0}^{K-1} \mathcal{N}(R_\ell x_k, \sigma^2 I)$$



All signals in the orbit $Gx := \{gx|g \in G\}$—where G is a subgroup of the permutation group—have the same likelihood.

Conclusion: The likelihood does not determine x uniquely, only the orbit Gx. We must assume a prior on the signal.

Proof strategy (Method of moments)

Likelihood:

$$p(y|x) = \frac{1}{M} \sum_{k=0}^{K-1} \sum_{\ell=0}^{L-1} \mathcal{N}(R_\ell x_k, \sigma^2 I)$$

[1] Bandeira, Blum-Smith, Kileel, Perry, Weed, Wein. "Estimation under group actions: recovering orbits from invariants." 2018

Proof strategy (Method of moments)

Likelihood:

$$p(y|x) = \frac{1}{M} \sum_{k=0}^{K-1} \sum_{\ell=0}^{L-1} \mathcal{N}(R_\ell x_k, \sigma^2 I)$$

Computing the moments of y is equivalent to averaging over the moments of the sub-signals $x_0, \ldots, x_{K-1}$:

$$M_y^q = \frac{1}{K} \sum_{k=0}^{K-1} M_{x_k}^q$$

[1] Bandeira, Blum-Smith, Kileel, Perry, Weed, Wein. "Estimation under group actions: recovering orbits from invariants." 2018

Proof strategy (Method of moments)

Likelihood:

$$p(y|x) = \frac{1}{M} \sum_{k=0}^{K-1} \sum_{\ell=0}^{L-1} \mathcal{N}(R_\ell x_k, \sigma^2 I)$$

Computing the moments of y is equivalent to averaging over the moments of the sub-signals $x_0, \ldots, x_{K-1}$:

$$M_y^q = \frac{1}{K} \sum_{k=0}^{K-1} M_{x_k}^q$$

It has been shown that M_y^3 (having $O(L^2)$ entries) determines Gx as long as $K \leq L/6$[1], implying

$$K = \frac{M}{L} \leq L/6 \quad \Rightarrow \quad M \leq L^2/6$$

[1] Bandeira, Blum-Smith, Kileel, Perry, Weed, Wein. "Estimation under group actions: recovering orbits from invariants." 2018

Proof strategy (Method of moments)

Conclusion: If the third moment can be estimated accurately, then one can recover the orbit Gx as long as $L \geq \sqrt{6M}$.

Proof strategy (Method of moments)

Conclusion: If the third moment can be estimated accurately, then one can recover the orbit Gx as long as $L \geq \sqrt{6M}$.

How many observations are required to estimate the third moment:

Proof strategy (Method of moments)

Conclusion: If the third moment can be estimated accurately, then one can recover the orbit Gx as long as $L \geq \sqrt{6M}$.

How many observations are required to estimate the third moment:

- In the low SNR regime, $N/\sigma^6 \to \infty$
  (a necessary condition for any algorithm [Bandiera et al., '17; Abbe et al., '18])

Proof strategy (Method of moments)

Conclusion: If the third moment can be estimated accurately, then one can recover the orbit Gx as long as $L \geq \sqrt{6M}$.

How many observations are required to estimate the third moment:

- In the low SNR regime, $N/\sigma^6 \to \infty$
  (a necessary condition for any algorithm [Bandiera et al., '17; Abbe et al., '18])

- In the high SNR regime, $N \approx K \log K$ (in expectation)

Proof strategy (Last stage)

So far:

- From the likelihood function, one can only recover the orbit Gx.

- Given the third moment, Gx can be estimated as long as $L \geq \sqrt{6M}$.

Proof strategy (Last stage)

So far:

- From the likelihood function, one can only recover the orbit Gx.
- Given the third moment, Gx can be estimated as long as $L \geq \sqrt{6M}$.

Last stage: Given almost any Gaussian prior on the signal, there is a unique signal in Gx that achieves the maximum of the posterior distribution (MAP).

Computational considerations

Our theoretical analysis suggests a two-stage procedure:

- identifying the orbit Gx;
- identifying x from Gx using a prior on x.

Computational considerations

Our theoretical analysis suggests a two-stage procedure:

- identifying the orbit Gx;
- identifying x from Gx using a prior on x.

We devised an expectation-maximization algorithm to maximize the posterior distribution directly (likelihood+prior)

Computational considerations

Our theoretical analysis suggests a two-stage procedure:

- identifying the orbit Gx;
- identifying x from Gx using a prior on x.

We devised an expectation-maximization algorithm to maximize the posterior distribution directly (likelihood+prior)

Can we achieve the theoretical bound $L \approx M^{1/2}$? **No**.

Computational considerations

Our theoretical analysis suggests a two-stage procedure:

- identifying the orbit Gx;
- identifying x from Gx using a prior on x.

We devised an expectation-maximization algorithm to maximize the posterior distribution directly (likelihood+prior)

Can we achieve the theoretical bound $L \approx M^{1/2}$? **No**.

Previous works postulated a **statistical-computational gap**: efficient algorithms can recover the orbit Gx only when $L \approx M^{2/3}$.

Computational considerations

Our theoretical analysis suggests a two-stage procedure:

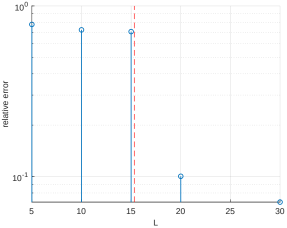
- identifying the orbit Gx;
- identifying x from Gx using a prior on x.

We devised an expectation-maximization algorithm to maximize the posterior distribution directly (likelihood+prior)

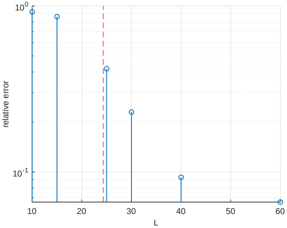
Can we achieve the theoretical bound $L \approx M^{1/2}$? **No**.

Previous works postulated a **statistical-computational gap**: efficient algorithms can recover the orbit Gx only when $L \approx M^{2/3}$.

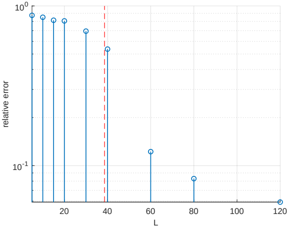
Our task is significantly harder, and thus empirically we need $L > M^{2/3}$.



(a) $M = 60$



(b) $M = 120$



(c) $M = 240$

$\mathrm{SNR} = 5, N = 10^3$, red vertical line indicates $L = M^{2/3}$

Future work

1. Super-resolution of continuous setups, multi-dimensional signals, and cryo-EM

2. Sampling theory in low SNR environments using moments (characterizing the interplay between M, L, N, σ)

3. Statistical-computational gaps

Thanks for your attention!