Two-Dimensional Tomography from Noisy Projection Tilt Series Taken at Unknown View Angles with Non-Uniform Distribution

Lingda Wang

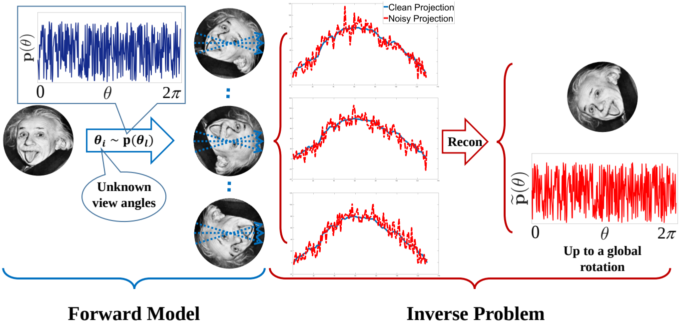
Department of Electrical and Computer Engineering
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

The 26th IEEE International Conference on Image Processing (ICIP 2019)

September 23rd 2019

Joint work with Zhizhen Zhao

Motivation & Intro: A Statistical Estimation Problem



- Unknown 2-D object f, and unknown non-uniform view angle distribution $\mathbf{p}(\theta)$.
- Observation model: $\mathbf{y}_{i,\kappa} = \mathcal{P}_{\theta_i+\kappa\alpha}(f) + \mathbf{n}_{i,\kappa}$, $\forall i \in [N]$, $|\kappa| \leq K$.
- Parameters of interest: f and $\mathbf{p}(\theta)$.

Related Works and Problems

- Existing methods focus on estimating view angles.
  - S. Basu and Y. Bresler[1]:
    - View angle ordering via nearest neighbor;
    - Joint maximum likelihood refinement.
  - A. Singer and H. Wu[2]:
    - Denoising (e.g., linear Wiener filtering and graph denoising);
    - Diffusion maps for view angle ordering.
  - Cons: Poor performance under low SNR; Computationally inefficient with large number of projections.

[1] Samit Basu and Yoram Bresler, Feasibility of tomography with unknown view angles, IEEE Transactions on Image Processing 9 (2000), no.6, 1107-1122.

[2] Amit Singer and H-T Wu, Two-dimensional tomography from noisy projections taken at unknown random directions, SIAM journal on imaging sciences 6 (2013), no.1, 136-175.

Related Works and Problems

- Existing methods focus on estimating view angles.
  - S. Basu and Y. Bresler[1]:
    - View angle ordering via nearest neighbor;
    - Joint maximum likelihood refinement.
  - A. Singer and H. Wu[2]:
    - Denoising (e.g., linear Wiener filtering and graph denoising);
    - Diffusion maps for view angle ordering.
  - Cons: Poor performance under low SNR; Computationally inefficient with large number of projections.
- Related problems:
  - 1-D: Multi-segment reconstruction (MSR).
  - 3-D: Sub-tomogram averaging (STA) in cryo-ET and single-particle reconstruction (SPR) in cryo-EM.

[1] Samit Basu and Yoram Bresler, Feasibility of tomography with unknown view angles, IEEE Transactions on Image Processing 9 (2000), no.6, 1107-1122.

[2] Amit Singer and H-T Wu, Two-dimensional tomography from noisy projections taken at unknown random directions, SIAM journal on imaging sciences 6 (2013), no.1, 136-175.

Proposed Method

- Method of Moments (MoM):
  - Conjecture: first and second order moments may contain sufficient information for the recovery.
  - Difficulty: algorithm by directly inverting the first and second order moments is unavailable with unknown non-uniform view angle distribution $\mathbf{p}(\theta)$.
  - Solution: use nonconvex optimization methods to solve constrained weighted nonlinear least squares of first and second order moments.

Proposed Method

- Method of Moments (MoM):
  - Conjecture: first and second order moments may contain sufficient information for the recovery.
  - Difficulty: algorithm by directly inverting the first and second order moments is unavailable with unknown non-uniform view angle distribution $\mathbf{p}(\theta)$.
  - Solution: use nonconvex optimization methods to solve constrained weighted nonlinear least squares of first and second order moments.
- Pros: lower computational complexity. (N : # of data)
  - MoM: $\mathcal{O}(N)$ for moments (once) and $\mathcal{O}(1)$ for each iteration.
  - MLE (e.g., EM): $\mathcal{O}(N)$ for each iteration.

Moment Features

- Fourier domain: $\mathcal{F}(f)(\xi,\theta) \approx \sum_{k=-k_{\max}}^{k_{\max}} \sum_{q=1}^{q_k} a_{k,q}\psi_c^{k,q}(\xi,\theta)$.
- Fourier slice theorem: $\widehat{\mathbf{y}}_{i,\kappa}[\xi_j] = \mathcal{F}(f)(\xi_j,\theta_i+\kappa\alpha) + \widehat{\mathbf{n}}_{i,\kappa}[\xi_j]$.
- First order moment:

$$\begin{aligned}\boldsymbol{\mu}[j;\kappa] &= \sum_{k=-k_{\max}}^{k_{\max}} \sum_{q=1}^{q_k} \sum_{l=0}^{n_\theta-1} a_{k,q}\psi_c^{k,q}(\xi_j,\phi_l+\kappa\alpha)\,\mathbf{p}[l] \\ &= \boldsymbol{\Psi}\left(\mathbf{a}\circ\mathbf{g}(\widehat{\mathbf{p}})\right)[j;\kappa].\end{aligned}$$

- Second order moment:

$$\begin{aligned}\mathbf{C}[j_1;\kappa_1,j_2;\kappa_2] &= \sum_{k_1=-k_{\max}}^{k_{\max}} \sum_{k_2=-k_{\max}}^{k_{\max}} \sum_{q_1=1}^{q_{k_1}} \sum_{q_2=1}^{q_{k_2}} a_{k_1,q_1}\overline{a_{k_2,q_2}} \\ &\quad\times \psi_c^{k_1,q_1}(\xi_{j_1},\kappa_1\alpha)\,\overline{\psi_c^{k_2,q_2}(\xi_{j_2},\kappa_2\alpha)}\,\widehat{\mathbf{p}}[k_2-k_1] \\ &= \left(\boldsymbol{\Psi}\left(\mathbf{a}\mathbf{a}^*\circ\mathbf{H}(\widehat{\mathbf{p}})\right)\boldsymbol{\Psi}^*\right)[j_1;\kappa_1,j_2;\kappa_2].\end{aligned}$$

Moment Features

- Unbiased empirical estimators:

$$\widetilde{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^{N} \widehat{\mathbf{y}}_i, \quad \widetilde{\mathbf{C}} = \frac{1}{N}\sum_{i=1}^{N} \widehat{\mathbf{y}}_i \widehat{\mathbf{y}}_i^* - \widehat{\boldsymbol{\Sigma}}.$$

- Constrained weighted nonlinear least squares:

$$(\widetilde{\mathbf{a}}, \widetilde{\mathbf{p}}) = \arg\min_{\mathbf{a},\mathbf{p}} \begin{array}{l} \frac{\lambda_1}{2} \left\| \boldsymbol{\Psi} \left(\mathbf{a} \circ \mathbf{g}(\widehat{\mathbf{p}}) \right) - \widetilde{\boldsymbol{\mu}} \right\|_w^2 \\ + \frac{\lambda_2}{2} \left\| \boldsymbol{\Psi} \left(\mathbf{a}\mathbf{a}^* \circ \mathbf{H}(\widehat{\mathbf{p}}) \right) \boldsymbol{\Psi}^* - \widetilde{\mathbf{C}} \right\|_W^2, \end{array}$$
$$\text{s.t.} \quad \mathbf{p} \geq \mathbf{0} \quad \text{and} \quad \mathbf{1}^\top \mathbf{p} = 1.$$

- Gradient based methods (e.g., gradient descent and trust region) do not work well.

An Alternating Direction Method of Multiplier (ADMM) Approach

- Reformulation and relaxation: split **a** into **a** and **z**, and relax the positive constraint on **p**.

$$
\begin{aligned}
(\widetilde{\mathbf{a}}, \widetilde{\mathbf{p}}) = \arg\min_{\mathbf{a},\mathbf{p}} \;& \frac{\lambda_1}{2} \left\| \boldsymbol{\Psi} \left(\mathbf{a} \circ \mathbf{g}(\widehat{\mathbf{p}}) \right) - \widetilde{\boldsymbol{\mu}} \right\|_w^2 \\
& + \frac{\lambda_2}{2} \left\| \boldsymbol{\Psi} \left(\mathbf{a}\mathbf{a}^* \circ \mathbf{H}(\widehat{\mathbf{p}}) \right) \boldsymbol{\Psi}^* - \widetilde{\mathbf{C}} \right\|_W^2, \\
\text{s.t.} \quad & \mathbf{p} \geq \mathbf{0} \quad \text{and} \quad \mathbf{1}^\top \mathbf{p} = 1.
\end{aligned}
$$

An Alternating Direction Method of Multiplier (ADMM) Approach

- Reformulation and relaxation: split **a** into **a** and **z**, and relax the positive constraint on **p**.

$$
\begin{aligned}
(\widetilde{\mathbf{a}}, \widetilde{\mathbf{z}}, \widetilde{\mathbf{p}}) = \arg\min_{\mathbf{a},\mathbf{z},\mathbf{p}} \; & \tfrac{\lambda_1}{2}\|\boldsymbol{\Psi}\left(\mathbf{a} \circ \mathbf{g}(\widehat{\mathbf{p}})\right) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \tfrac{\lambda_1}{2}\|\boldsymbol{\Psi}\left(\mathbf{z} \circ \mathbf{g}(\widehat{\mathbf{p}})\right) - \widetilde{\boldsymbol{\mu}}\|_w^2 \\
& + \tfrac{\lambda_2}{2}\|\boldsymbol{\Psi}\left(\mathbf{a}\mathbf{z}^* \circ \mathbf{H}(\widehat{\mathbf{p}})\right)\boldsymbol{\Psi}^* - \widetilde{\mathbf{C}}\|_W^2 \\
\text{s.t.} \quad & \mathbf{a} = \mathbf{z}, \quad \mathbf{p} \geq \mathbf{0} \quad \text{and} \quad \mathbf{1}^\top \mathbf{p} = 1.
\end{aligned}
$$

An Alternating Direction Method of Multiplier (ADMM) Approach

- Reformulation and relaxation: split **a** into **a** and **z**, and relax the positive constraint on **p**.

$$(\widetilde{\mathbf{a}}, \widetilde{\mathbf{z}}, \widetilde{\mathbf{p}}) = \arg\min_{\mathbf{a},\mathbf{z},\mathbf{p}} \frac{\lambda_1}{2}\|\boldsymbol{\Psi}(\mathbf{a}\circ\mathbf{g}(\widehat{\mathbf{p}})) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \frac{\lambda_1}{2}\|\boldsymbol{\Psi}(\mathbf{z}\circ\mathbf{g}(\widehat{\mathbf{p}})) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \frac{\lambda_2}{2}\|\boldsymbol{\Psi}(\mathbf{a}\mathbf{z}^*\circ\mathbf{H}(\widehat{\mathbf{p}}))\boldsymbol{\Psi}^* - \widetilde{\mathbf{C}}\|_W^2$$

$$\text{s.t.} \quad \mathbf{a} = \mathbf{z}, \quad \cancel{\mathbf{p} \succeq \mathbf{0}} \quad \text{and} \quad \mathbf{1}^\top\mathbf{p} = 1.$$

An Alternating Direction Method of Multiplier (ADMM) Approach

- Reformulation and relaxation: split **a** into **a** and **z**, and relax the positive constraint on **p**.

$$(\widetilde{\mathbf{a}}, \widetilde{\mathbf{z}}, \widetilde{\mathbf{p}}) = \arg\min_{\mathbf{a},\mathbf{z},\mathbf{p}} \frac{\lambda_1}{2}\|\boldsymbol{\Psi}(\mathbf{a} \circ \mathbf{g}(\widehat{\mathbf{p}})) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \frac{\lambda_1}{2}\|\boldsymbol{\Psi}(\mathbf{z} \circ \mathbf{g}(\widehat{\mathbf{p}})) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \frac{\lambda_2}{2}\|\boldsymbol{\Psi}(\mathbf{a}\mathbf{z}^* \circ \mathbf{H}(\widehat{\mathbf{p}}))\boldsymbol{\Psi}^* - \widetilde{\mathbf{C}}\|_W^2$$

$$\text{s.t.} \quad \mathbf{a} = \mathbf{z}, \quad \cancel{\mathbf{p} \succeq \mathbf{0}} \quad \text{and} \quad \mathbf{1}^\top \mathbf{p} = 1.$$

- Augmented Lagrangian:

$$\mathcal{L}(\mathbf{a}, \mathbf{z}, \mathbf{p}; \mathbf{s}) = \frac{\lambda_1}{2}\|\boldsymbol{\Psi}(\mathbf{a} \circ \mathbf{g}(\widehat{\mathbf{p}})) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \frac{\lambda_1}{2}\|\boldsymbol{\Psi}(\mathbf{z} \circ \mathbf{g}(\widehat{\mathbf{p}})) - \widetilde{\boldsymbol{\mu}}\|_w^2 + \frac{\lambda_2}{2}\left\|\boldsymbol{\Psi}(\mathbf{a}\mathbf{z}^* \circ \mathbf{H}(\widehat{\mathbf{p}}))\boldsymbol{\Psi}^* - \widetilde{\mathbf{C}}\right\|_W^2 + \frac{\rho}{2}\|\mathbf{a} - \mathbf{z} + \mathbf{s}\|_2^2.$$

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.
  - $\mathbf{a}^{(t+1)} = \arg\min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{z}^{(t)}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.
  - $\mathbf{a}^{(t+1)} = \arg\min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{z}^{(t)}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.
  - $\mathbf{a}^{(t+1)} = \arg\min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{z}^{(t)}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{p}^{(t+1)} = \arg\min_{\mathbf{p}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}^{(t+1)}, \mathbf{p}; \mathbf{s}^{(t)})$

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.
  - $\mathbf{a}^{(t+1)} = \arg\min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{z}^{(t)}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{p}^{(t+1)} = \arg\min_{\mathbf{p}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}^{(t+1)}, \mathbf{p}; \mathbf{s}^{(t)})$
  - $\mathbf{s}^{(t+1)} = \mathbf{s}^{(t)} + \mathbf{a}^{(t+1)} - \mathbf{z}^{(t+1)}$

An ADMM Approach

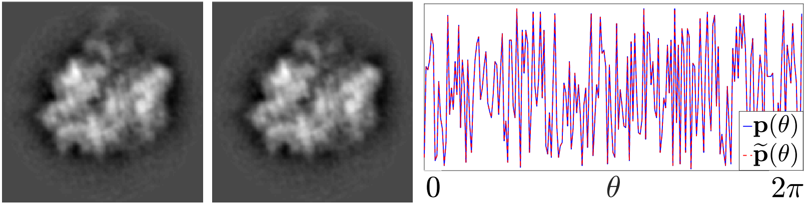
- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.
  - $\mathbf{a}^{(t+1)} = \arg\min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{z}^{(t)}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{p}^{(t+1)} = \arg\min_{\mathbf{p}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}^{(t+1)}, \mathbf{p}; \mathbf{s}^{(t)})$
  - $\mathbf{s}^{(t+1)} = \mathbf{s}^{(t)} + \mathbf{a}^{(t+1)} - \mathbf{z}^{(t+1)}$
- Output $\mathbf{a}^{(t)}, \mathbf{p}^{(t)}$.

An ADMM Approach

- Initialization: random initialization of $\mathbf{a}^{(0)}, \mathbf{z}^{(0)}, \mathbf{p}^{(0)}$, and $\mathbf{s}^{(0)} = \mathbf{0}$.
- Iterations: alternates between the primal updates of variables $\mathbf{a}$, $\mathbf{z}$, and $\mathbf{p}$, and the dual update for $\mathbf{s}$ until convergence.
  - $\mathbf{a}^{(t+1)} = \arg\min_{\mathbf{a}} \mathcal{L}(\mathbf{a}, \mathbf{z}^{(t)}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{z}^{(t+1)} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}, \mathbf{p}^{(t)}; \mathbf{s}^{(t)})$
  - $\mathbf{p}^{(t+1)} = \arg\min_{\mathbf{p}} \mathcal{L}(\mathbf{a}^{(t+1)}, \mathbf{z}^{(t+1)}, \mathbf{p}; \mathbf{s}^{(t)})$
  - $\mathbf{s}^{(t+1)} = \mathbf{s}^{(t)} + \mathbf{a}^{(t+1)} - \mathbf{z}^{(t+1)}$
- Output $\mathbf{a}^{(t)}, \mathbf{p}^{(t)}$.
- Remark: each update can be realized by solving simple least squares.

Numerical Results: Clean Case

- Parameters: $N = 10000$, $\alpha = 1.5$ deg , $|\kappa| = 13$, $c = 0.3$, $\lambda_1 = 1$, $\lambda_2 = 0.5$, and $\rho = 1$.
- Exact recovery on a projection image of 70S ribsome (up to a rotation).
- Perfect match of view angle distribution (up to a rotation).



(a) Original (b) Reconstructed (c) $\widetilde{\mathbf{p}}(\theta)$

Numerical Results: Noisy Case

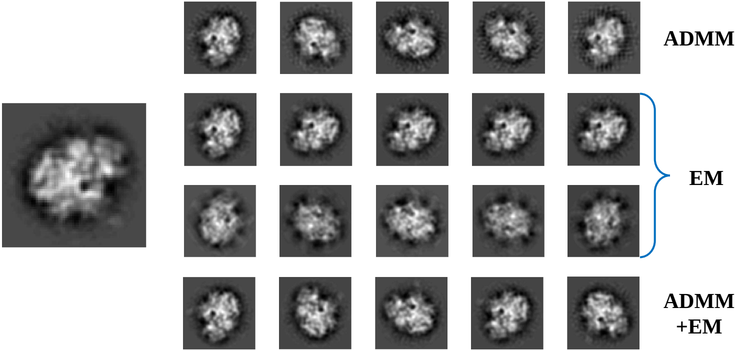
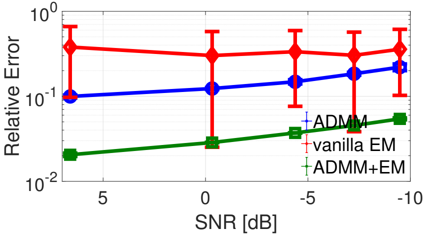


Figure: SNR [dB] = 6.61, -0.32, -4.38, -7.25, -9.49.

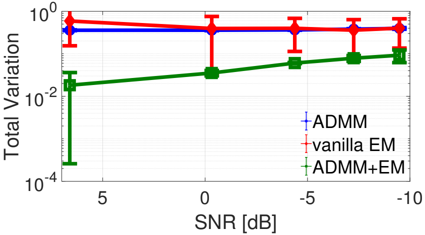
- EM algorithm for maximum marginalized log-likelihood estimation:

$$\max_{\mathbf{a},\mathbf{p}} \sum_{i=1}^{N} \ln P\left(\widehat{\mathbf{y}}_i \middle| \mathbf{a}, \mathbf{p}\right) \quad \text{s.t.} \quad \mathbf{p} \geq \mathbf{0} \quad \text{and} \quad \mathbf{1}^\top \mathbf{p} = 1.$$

Numerical Results: Noisy Case



(a) Image Reconstruction



(b) View Distribution

- Parameters: $N = 10000$, $\alpha = 3.8$ deg , $|\kappa| = 13$, $c = 0.3$, $\lambda_1 = 1$, $\lambda_2 = 5$, and $\rho = 1$.
- Results over 20 independent experiments.
- Performance: ADMM+EM>ADMM>EM.

Thank you!

- Our paper *Two-dimensional Tomography from Noisy Projection Tilt Series Taken at Unknown View Angles with Non-uniform Distribution* is available online at:
  https://ieeexplore.ieee.org/document/8803755

- Our codes are available online at:
  https://github.com/LingdaWang/2D_TOMO_ICIP2019