

# Denoising and Covariance Estimation of Cryo-EM Images

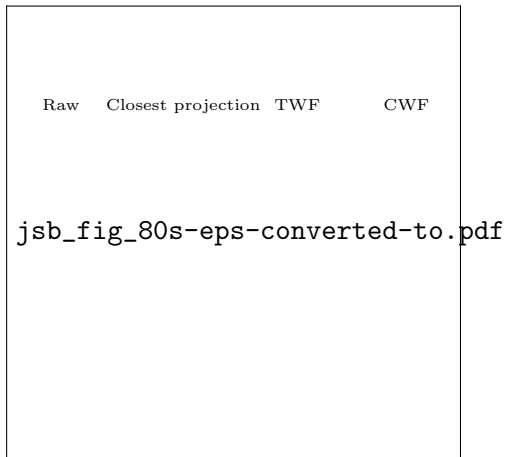
Tejal Bhamre

Princeton University

*Joint work with Teng Zhang (UCF) and Amit Singer (Princeton)*

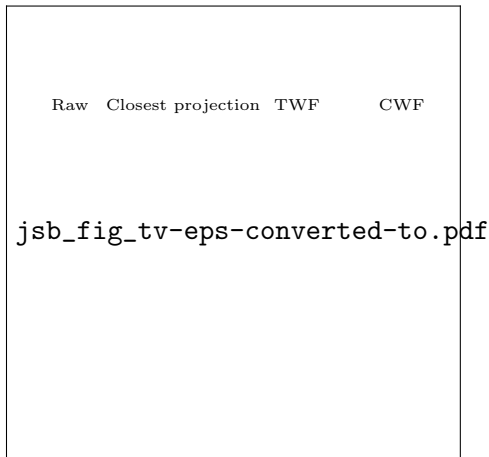
June 10, 2016

# Experimental data - 80S ribosome



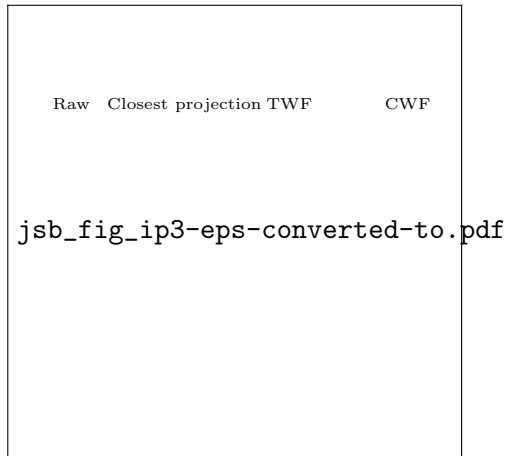
- FALCON II 4k×4k direct electron detector
- 105247 motion corrected, picked particle images of 360×360 pixels

# Experimental data - TRPV1



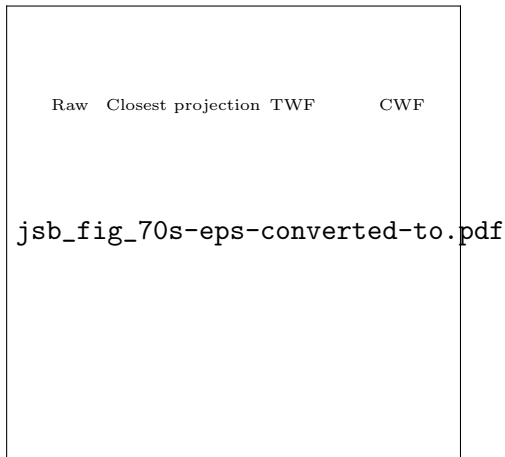
- K2 direct electron detector
- 35645 motion corrected, picked particle images of  $256 \times 256$  pixels

# Experimental data -IP<sub>3</sub>R1



- Gatan 4k×4k CCD
- 37382 picked particle images of 256×256 pixels

# Experimental data - 70S ribosome



- TVIPS TEMCAM-F415 (4k x 4k) CCD
- 216517 picked particle images of  $250 \times 250$  pixels

# Cryo-EM Pipeline

- Particle picking from micrographs
- Preprocessing
- 2D classification (Class averaging): inspect underlying particles, estimate viewing angles
- 3D classification
- 3D refinement

# Cryo-EM Pipeline

- Particle picking from micrographs
- **Preprocessing**
- 2D classification (Class averaging): inspect underlying particles, estimate viewing angles
- 3D classification
- 3D refinement

# Cryo-EM Pipeline

- Particle picking from micrographs
- **Preprocessing**
- **2D classification (Class averaging): inspect underlying particles, estimate viewing angles**
- 3D classification
- 3D refinement



# Cryo-EM Pipeline

- Particle picking from micrographs
- **Preprocessing**
- **2D classification (Class averaging): inspect underlying particles, estimate viewing angles**
- 3D classification
- 3D refinement

# Cryo-EM Pipeline

- Particle picking from micrographs
- **Preprocessing**
- **2D classification (Class averaging): inspect underlying particles, estimate viewing angles**
- 3D classification
- 3D refinement

# Motivation

- Visualization of underlying particles **without class averaging**
- Image restoration (CTF correction and denoising) in a **single step**
- **Outlier detection**

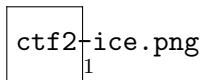
# Motivation

- Visualization of underlying particles **without class averaging**
- Image restoration (CTF correction and denoising) in a **single step**
- **Outlier detection**

# Motivation

- Visualization of underlying particles **without class averaging**
- Image restoration (CTF correction and denoising) in a **single step**
- **Outlier detection**

# CTF Correction

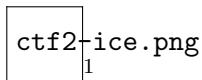


- CTF suppresses information and inverts contrast
- **Challenge:** CTF cannot be trivially inverted (zero crossings)
- Information lost from one defocus group could be recovered from another group that has different zero crossings.

---

<sup>1</sup><http://www.bio.brandeis.edu/~shaikh/lab/ctf.htm>

# CTF Correction

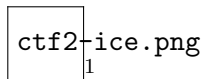


- CTF suppresses information and inverts contrast
- **Challenge:** CTF cannot be trivially inverted (zero crossings)
- Information lost from one defocus group could be recovered from another group that has different zero crossings.

---

<sup>1</sup><http://www.bio.brandeis.edu/~shaikh/lab/ctf.htm>

# CTF Correction



- CTF suppresses information and inverts contrast
- **Challenge:** CTF cannot be trivially inverted (zero crossings)
- Information lost from one defocus group could be recovered from another group that has different zero crossings.

---

<sup>1</sup><http://www.bio.brandeis.edu/~shaikh/lab/ctf.htm>



# CTF Correction

- Do nothing (Denial)
- Correct Fourier phases but not amplitudes
- Correct both Fourier phases and amplitudes

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**
  - Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
  - Preserves noise statistics
  - Data adaptive basis: eigenvectors of the sample covariance matrix
  - Phase flipping corrects only phases
- **Traditional Wiener Filtering (TWF):**
  - Corrects both phases and amplitudes
  - Requires prior estimation of the spectral signal to noise ratio (SSNR)
  - Cannot restore information at zero crossings of the CTF
  - Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)



# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**
  - Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
  - Preserves noise statistics
  - Data adaptive basis: eigenvectors of the sample covariance matrix
  - Phase flipping corrects only phases
- **Traditional Wiener Filtering (TWF):**
  - Corrects both phases and amplitudes
  - Requires prior estimation of the spectral signal to noise ratio (SSNR)
  - Cannot restore information at zero crossings of the CTF
  - Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**
  - Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
  - Preserves noise statistics
  - Data adaptive basis: eigenvectors of the sample covariance matrix
  - Phase flipping corrects only phases
- **Traditional Wiener Filtering (TWF):**
  - Corrects both phases and amplitudes
  - Requires prior estimation of the spectral signal to noise ratio (SSNR)
  - Cannot restore information at zero crossings of the CTF
  - Not in a data adaptive basis (restricted to Fourier basis)

# Current Image Restoration Techniques

- **Phase flipping + steerable PCA (sPCA):**

- Flip sign of the Fourier coefficients at frequencies for which the CTF is negative
- Preserves noise statistics
- Data adaptive basis: eigenvectors of the sample covariance matrix
- Phase flipping corrects only phases

- **Traditional Wiener Filtering (TWF):**

- Corrects both phases and amplitudes
- Requires prior estimation of the spectral signal to noise ratio (SSNR)
- Cannot restore information at zero crossings of the CTF
- Not in a data adaptive basis (restricted to Fourier basis)

# Covariance Wiener Filtering (CWF)

- Estimate the CTF-corrected covariance matrix of the underlying clean 2D projection images
- Wiener filtering to solve the image restoration deconvolution problem
- No averaging, act on each image separately
- CTF correction and denoising in a single step

# Covariance Wiener Filtering (CWF)

- Estimate the CTF-corrected covariance matrix of the underlying clean 2D projection images
- Wiener filtering to solve the image restoration deconvolution problem
- No averaging, act on each image separately
- CTF correction and denoising in a single step

# Covariance Wiener Filtering (CWF)

- Estimate the CTF-corrected covariance matrix of the underlying clean 2D projection images
- Wiener filtering to solve the image restoration deconvolution problem
- No averaging, act on each image separately
- CTF correction and denoising in a single step

# Covariance Wiener Filtering (CWF)

- Estimate the CTF-corrected covariance matrix of the underlying clean 2D projection images
- Wiener filtering to solve the image restoration deconvolution problem
- No averaging, act on each image separately
- CTF correction and denoising in a single step

# Covariance Wiener Filtering (CWF)

Table : Comparison of CTF Correction/Denoising Methods

Property	Phaseflip + sPCA	TWF	CWF
Applicable at preliminary stage	✓	✓	✓
Data dependent basis	✓	✗	✓
Correct both phases and amplitudes	✗	✓	✓
CTF corrected covariance estimate	✗	✗	✓



# Fourier-Bessel Steerable Basis

- The population covariance matrix  $\Sigma$  must be invariant under in-plane rotation of the projection images
- Block diagonal in any steerable basis in which the basis elements are outer products of radial functions and angular Fourier modes <sup>2</sup>
- CTF and the whitening filter are also block diagonal in the Fourier Bessel basis (radial isotropy)
- Suffices to estimate each diagonal block of  $\Sigma$ , corresponding to the angular frequency  $k$ , separately
- Nearly unitary transformation

---

<sup>2</sup>Fast Steerable Principal Component Analysis, Z. Zhao and Y. Shkolnisky and A. Singer, IEEE Transactions on Computational Imaging

# Fourier-Bessel Steerable Basis

- The population covariance matrix  $\Sigma$  must be invariant under in-plane rotation of the projection images
- Block diagonal in any steerable basis in which the basis elements are outer products of radial functions and angular Fourier modes <sup>2</sup>
- CTF and the whitening filter are also block diagonal in the Fourier Bessel basis (radial isotropy)
- Suffices to estimate each diagonal block of  $\Sigma$ , corresponding to the angular frequency  $k$ , separately
- Nearly unitary transformation

---

<sup>2</sup>Fast Steerable Principal Component Analysis, Z. Zhao and Y. Shkolnisky and A. Singer, IEEE Transactions on Computational Imaging

# Fourier-Bessel Steerable Basis

- The population covariance matrix  $\Sigma$  must be invariant under in-plane rotation of the projection images
- Block diagonal in any steerable basis in which the basis elements are outer products of radial functions and angular Fourier modes <sup>2</sup>
- CTF and the whitening filter are also block diagonal in the Fourier Bessel basis (radial isotropy)
- Suffices to estimate each diagonal block of  $\Sigma$ , corresponding to the angular frequency  $k$ , separately
- Nearly unitary transformation

---

<sup>2</sup>Fast Steerable Principal Component Analysis, Z. Zhao and Y. Shkolnisky and A. Singer, IEEE Transactions on Computational Imaging

# Fourier-Bessel Steerable Basis

- The population covariance matrix  $\Sigma$  must be invariant under in-plane rotation of the projection images
- Block diagonal in any steerable basis in which the basis elements are outer products of radial functions and angular Fourier modes <sup>2</sup>
- CTF and the whitening filter are also block diagonal in the Fourier Bessel basis (radial isotropy)
- Suffices to estimate each diagonal block of  $\Sigma$ , corresponding to the angular frequency  $k$ , separately
- Nearly unitary transformation

---

<sup>2</sup>Fast Steerable Principal Component Analysis, Z. Zhao and Y. Shkolnisky and A. Singer, IEEE Transactions on Computational Imaging

# Fourier-Bessel Steerable Basis

- The population covariance matrix  $\Sigma$  must be invariant under in-plane rotation of the projection images
- Block diagonal in any steerable basis in which the basis elements are outer products of radial functions and angular Fourier modes <sup>2</sup>
- CTF and the whitening filter are also block diagonal in the Fourier Bessel basis (radial isotropy)
- Suffices to estimate each diagonal block of  $\Sigma$ , corresponding to the angular frequency  $k$ , separately
- Nearly unitary transformation

---

<sup>2</sup>Fast Steerable Principal Component Analysis, Z. Zhao and Y. Shkolnisky and A. Singer, IEEE Transactions on Computational Imaging

# Computational Complexity: Timings

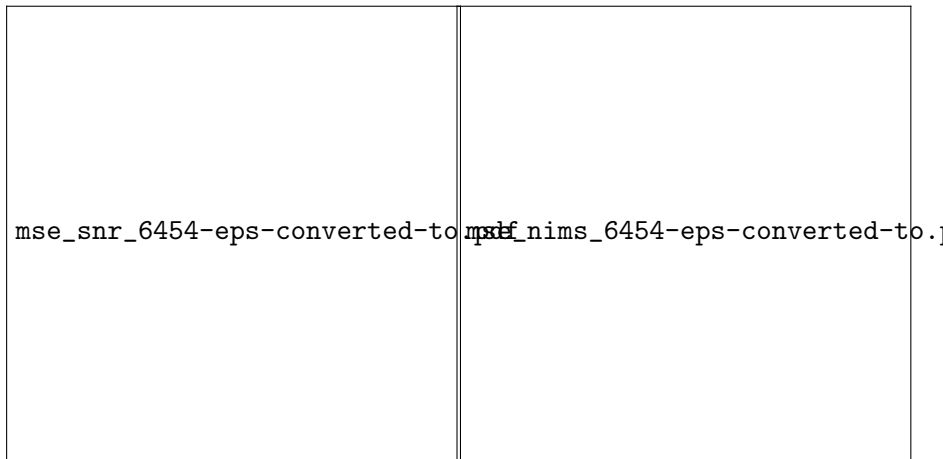
$n$  images of size  $L \times L$

UNIX environment with 60 cores, running at 2.3 GHz, with total RAM of 1.5TB

Table : Timing in seconds

Dataset	$L$	$n$	Basis coeffs	CWF
TRPV1	256	35645	312s	574s
80s	360	30000	731s	385s
IP3R1	256	37382	429s	589s
70s	250	99979	1174s	113s

# Relative error of estimated clean images



(a)

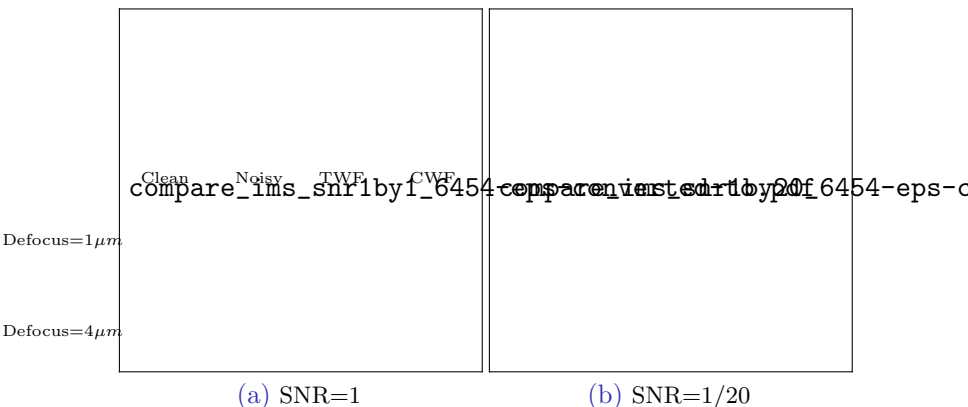
(b)

# Relative error of estimated covariance

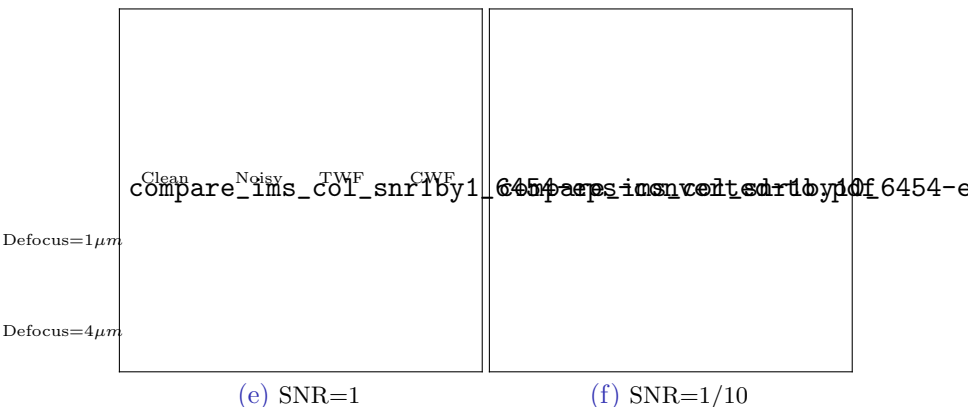
cwf\_shrinkage\_compare-eps-converted-to.pdf



# Simulations with white noise: 80S ribosome (EMDB-6454)



# Simulations with colored noise: 80S ribosome (EMDB-6454)



# Outlier Detection

- Significant amount of time is spent on discarding outliers by visual inspection after the particle picking
- CWF: automatic way to classify picked particles
- Specimen particles at various depths in the ice layer: acquired projection images can have different contrasts

- 

$$Y_i = \alpha_i A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (1)$$

- Absorb  $\alpha$  into  $\mathbf{X}$  and estimate  $\alpha_i X_i$
- Outlier images typically have low contrast after denoising: linear classifier after CWF

# Outlier Detection

- Significant amount of time is spent on discarding outliers by visual inspection after the particle picking
- CWF: automatic way to classify picked particles
- Specimen particles at various depths in the ice layer: acquired projection images can have different contrasts

- 

$$Y_i = \alpha_i A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (1)$$

- Absorb  $\alpha$  into  $\mathbf{X}$  and estimate  $\alpha_i X_i$
- Outlier images typically have low contrast after denoising: linear classifier after CWF

# Outlier Detection

- Significant amount of time is spent on discarding outliers by visual inspection after the particle picking
- CWF: automatic way to classify picked particles
- Specimen particles at various depths in the ice layer: acquired projection images can have different contrasts

•

$$Y_i = \alpha_i A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (1)$$

- Absorb  $\alpha$  into  $\mathbf{X}$  and estimate  $\alpha_i X_i$
- Outlier images typically have low contrast after denoising: linear classifier after CWF

# Outlier Detection

- Significant amount of time is spent on discarding outliers by visual inspection after the particle picking
- CWF: automatic way to classify picked particles
- Specimen particles at various depths in the ice layer: acquired projection images can have different contrasts

- 

$$Y_i = \alpha_i A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (1)$$

- Absorb  $\alpha$  into  $\mathbf{X}$  and estimate  $\alpha_i X_i$
- Outlier images typically have low contrast after denoising: linear classifier after CWF

# Outlier Detection

- Significant amount of time is spent on discarding outliers by visual inspection after the particle picking
- CWF: automatic way to classify picked particles
- Specimen particles at various depths in the ice layer: acquired projection images can have different contrasts

- 

$$Y_i = \alpha_i A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (1)$$

- Absorb  $\alpha$  into  $\mathbf{X}$  and estimate  $\alpha_i X_i$
- Outlier images typically have low contrast after denoising: linear classifier after CWF

# Outlier Detection

- Significant amount of time is spent on discarding outliers by visual inspection after the particle picking
- CWF: automatic way to classify picked particles
- Specimen particles at various depths in the ice layer: acquired projection images can have different contrasts

- 

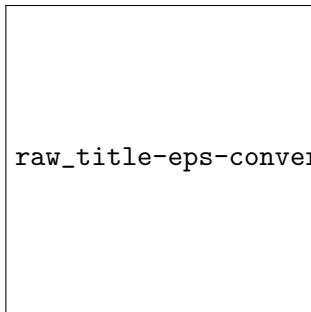
$$Y_i = \alpha_i A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (1)$$

- Absorb  $\alpha$  into  $\mathbf{X}$  and estimate  $\alpha_i X_i$
- Outlier images typically have low contrast after denoising: linear classifier after CWF

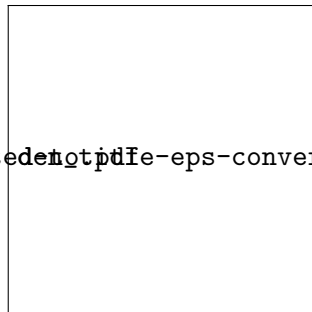


# Outlier Detection: 80S ribosome (EMDB-6454)

SNR=1/20  $\alpha \in [0.75, 1.5]$  10% images are pure noise



(h)



(i)



# Current and Future Work

- Better class averages
- Covariance matrix: Orthogonal Replacement using Kam's method
- 3D reconstruction from denoised images, without class averaging

# Current and Future Work

- Better class averages
- Covariance matrix: Orthogonal Replacement using Kam's method
- 3D reconstruction from denoised images, without class averaging

# Current and Future Work

- Better class averages
- Covariance matrix: Orthogonal Replacement using Kam's method
- 3D reconstruction from denoised images, without class averaging

# Resources

- Code: [https://github.com/PrincetonUniversity/cwf\\_denoise](https://github.com/PrincetonUniversity/cwf_denoise)
- Paper: Journal of Structural Biology:  
10.1016/j.jsb.2016.04.013

# The Model: Real space

Linear, weak phase approximation

$$y_i = a_i * x_i + \epsilon_i, \quad i = 1, 2, \dots, n \quad (2)$$

$n$ : number of images

$*$ : convolution operation

$y_i$ : noisy, CTF filtered  $i$ 'th image in real space

$x_i$ : underlying clean projection image in real space

$a_i$ : the point spread function of the microscope

$\epsilon_i$ : additive Gaussian noise that corrupts the image

# The Model: Fourier space

$$Y_i = A_i X_i + \xi_i, \quad i = 1, 2, \dots, n \quad (3)$$

$A_i$ : diagonal operator, whose diagonal consists of the Fourier transform of the point spread function

$X_1, \dots, X_n$ : vectors in  $\mathbb{C}^p$ , where  $p$  is the number of pixels  
 i.i.d. samples from a distribution with mean  $\mathbb{E}[\mathbf{X}] = \mu$  and covariance  
 $\mathbb{E}[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T] = \Sigma$

**“All models are wrong but some are useful” - George Box**

# The Model

$$\mathbb{E}[\mathbf{Y}_i] = A_i \mathbb{E}[\mathbf{X}_i], \quad i = 1, 2, \dots, n. \quad (4)$$

$$\begin{aligned} \mathbb{E}[(\mathbf{Y}_i - \mathbb{E}[\mathbf{Y}_i])(\mathbf{Y}_i - \mathbb{E}[\mathbf{Y}_i])^T] &= \mathbb{E}[A_i(\mathbf{X}_i - \mu)(\mathbf{X}_i - \mu)^T A_i^T] + \sigma^2 I \\ &= A_i \Sigma A_i^T + \sigma^2 I. \end{aligned} \quad (5)$$

Relates the second order statistics of the noisy images with the population covariance  $\Sigma$  of the clean images



# Mean Estimation

$$\hat{\mu} = \arg \min_{\mu} \sum_{i=1}^n \|(Y_i - A_i \mu)\|_2^2 + \lambda \|\mu\|_2^2 \quad (6)$$

$$\hat{\mu} = \left( \sum_{i=1}^n A_i^T A_i + \lambda I \right)^{-1} \left( \sum_{i=1}^n A_i^T Y_i \right). \quad (7)$$

# Covariance Estimation

$$\begin{aligned}
 \hat{\Sigma} &= \arg \min_{\Sigma} \sum_{i=1}^n \|(Y_i - \mathbb{E}[\mathbf{Y}_i])(Y_i - \mathbb{E}[\mathbf{Y}_i])^T - (A_i \Sigma A_i^T + \sigma^2 I)\|_F^2 \\
 &= \arg \min_{\Sigma} \sum_{i=1}^n \|A_i \Sigma A_i^T + \sigma^2 I - C_i\|_F^2
 \end{aligned} \tag{8}$$

where  $C_i = (Y_i - A_i \mu)(Y_i - A_i \mu)^T$  and  $\|\cdot\|_F$  is the Frobenius matrix norm.

# Solving using Conjugate Gradient

System of linear equations for the elements of the matrix  $\hat{\Sigma}$

$$\sum_{i=1}^n A_i^T A_i \hat{\Sigma} A_i^T A_i = \sum_{i=1}^n A_i^T C_i A_i - \sum_{i=1}^n \sigma^2 A_i^T A_i \quad (9)$$

$$L(\hat{\Sigma}) = B \quad (10)$$

where  $L : \mathbb{R}^{p \times p} \rightarrow \mathbb{R}^{p \times p}$  is the linear operator acting on  $\hat{\Sigma}$  defined by the left hand side of eqn. ??, and  $B$  is the right hand side.

- Direct inversion of this linear system is slow for large image sizes
- Applying  $L$  only involves matrix multiplications: fast!
- Conjugate gradient

# Eigenvalue Thresholding

- $L(\hat{\Sigma})$  is a PSD matrix whenever  $\hat{\Sigma}$  is PSD (as a sum of PSD matrices)
- $B$  may not necessarily be PSD due to finite sample fluctuations ( $n$  is finite)
- Project  $B$  onto the cone of PSD matrices
- Compute the spectral decomposition of  $B$  and set all negative eigenvalues to 0 (eigenvalue thresholding)
- $n \gg p$

# Eigenvalue Thresholding

- $L(\hat{\Sigma})$  is a PSD matrix whenever  $\hat{\Sigma}$  is PSD (as a sum of PSD matrices)
- $B$  may not necessarily be PSD due to finite sample fluctuations ( $n$  is finite)
- Project  $B$  onto the cone of PSD matrices
- Compute the spectral decomposition of  $B$  and set all negative eigenvalues to 0 (eigenvalue thresholding)
- $n \gg p$

# Eigenvalue Thresholding

- $L(\hat{\Sigma})$  is a PSD matrix whenever  $\hat{\Sigma}$  is PSD (as a sum of PSD matrices)
- $B$  may not necessarily be PSD due to finite sample fluctuations ( $n$  is finite)
- Project  $B$  onto the cone of PSD matrices
- Compute the spectral decomposition of  $B$  and set all negative eigenvalues to 0 (eigenvalue thresholding)
- $n \gg p$

# Eigenvalue Thresholding

- $L(\hat{\Sigma})$  is a PSD matrix whenever  $\hat{\Sigma}$  is PSD (as a sum of PSD matrices)
- $B$  may not necessarily be PSD due to finite sample fluctuations ( $n$  is finite)
- Project  $B$  onto the cone of PSD matrices
- Compute the spectral decomposition of  $B$  and set all negative eigenvalues to 0 (eigenvalue thresholding)
- $n \gg p$

# Eigenvalue Thresholding

- $L(\hat{\Sigma})$  is a PSD matrix whenever  $\hat{\Sigma}$  is PSD (as a sum of PSD matrices)
- $B$  may not necessarily be PSD due to finite sample fluctuations ( $n$  is finite)
- Project  $B$  onto the cone of PSD matrices
- Compute the spectral decomposition of  $B$  and set all negative eigenvalues to 0 (eigenvalue thresholding)
- $n \gg p$



# Eigenvalue Shrinkage: Spiked Covariance Model

Analyze  $B$  when  $X_i = 0$  for all  $i$  (input images are white noise images)

- 

$$M = \sum_{i=1}^n A_i^T C_i A_i = \sum_{i=1}^n A_i^T Y_i Y_i^T A_i. \quad (11)$$

- $\mathbb{E}[M] = \sigma^2 \sum_{i=1}^n A_i^T A_i$  and  $B = M - \mathbb{E}[M]$
- $S = (\mathbb{E}[M])^{1/2}$ , i.e.  $S$  is PSD and  $\mathbb{E}[M] = S^2$
- 

$$S^{-1} L(\hat{\Sigma}) S^{-1} = S^{-1} (M - \mathbb{E}[M]) S^{-1} = S^{-1} M S^{-1} - I. \quad (12)$$

# Eigenvalue Shrinkage: Spiked Covariance Model

- $S^{-1}MS^{-1}$  can be viewed as a sample covariance matrix of  $n$  vectors in  $\mathbb{R}^p$  whose population covariance is the identity matrix
- Eigenvalues corresponding to the signal can only be detected if they reside outside of the support of the Marčenko Pastur (MP) distribution
- Kritchman Nadler (KN) rank estimation to determine the number of eigenvalues corresponding to the signal <sup>3</sup>
- Apply operator norm eigenvalue shrinkage procedure (Donoho et al.) to those eigenvalues, while setting all other eigenvalues to 0 <sup>4</sup>

---

<sup>3</sup>Determining the number of components in a factor model from limited noisy data, Shira Kritchman and Boaz Nadler, Chemometrics and Intelligent Laboratory Systems

<sup>4</sup>Optimal shrinkage of eigenvalues in the spiked covariance model, David Donoho, Matan Gavish and I M Johnstone, [arxiv.org/abs/1311.0851](https://arxiv.org/abs/1311.0851)

# Eigenvalue Shrinkage: Spiked Covariance Model

- $S^{-1}MS^{-1}$  can be viewed as a sample covariance matrix of  $n$  vectors in  $\mathbb{R}^p$  whose population covariance is the identity matrix
- Eigenvalues corresponding to the signal can only be detected if they reside outside of the support of the Marčenko Pastur (MP) distribution
- Kritchman Nadler (KN) rank estimation to determine the number of eigenvalues corresponding to the signal <sup>3</sup>
- Apply operator norm eigenvalue shrinkage procedure (Donoho et al.) to those eigenvalues, while setting all other eigenvalues to 0 <sup>4</sup>

---

<sup>3</sup>Determining the number of components in a factor model from limited noisy data, Shira Kritchman and Boaz Nadler, Chemometrics and Intelligent Laboratory Systems

<sup>4</sup>Optimal shrinkage of eigenvalues in the spiked covariance model, David Donoho, Matan Gavish and I M Johnstone, [arxiv.org/abs/1311.0851](https://arxiv.org/abs/1311.0851)

# Eigenvalue Shrinkage: Spiked Covariance Model

- $S^{-1}MS^{-1}$  can be viewed as a sample covariance matrix of  $n$  vectors in  $\mathbb{R}^p$  whose population covariance is the identity matrix
- Eigenvalues corresponding to the signal can only be detected if they reside outside of the support of the Marčenko Pastur (MP) distribution
- Kritchman Nadler (KN) rank estimation to determine the number of eigenvalues corresponding to the signal <sup>3</sup>
- Apply operator norm eigenvalue shrinkage procedure (Donoho et al.) to those eigenvalues, while setting all other eigenvalues to 0 <sup>4</sup>

---

<sup>3</sup>Determining the number of components in a factor model from limited noisy data, Shira Kritchman and Boaz Nadler, Chemometrics and Intelligent Laboratory Systems

<sup>4</sup>Optimal shrinkage of eigenvalues in the spiked covariance model, David Donoho, Matan Gavish and I M Johnstone, [arxiv.org/abs/1311.0851](https://arxiv.org/abs/1311.0851)

# Eigenvalue Shrinkage: Spiked Covariance Model

- $S^{-1}MS^{-1}$  can be viewed as a sample covariance matrix of  $n$  vectors in  $\mathbb{R}^p$  whose population covariance is the identity matrix
- Eigenvalues corresponding to the signal can only be detected if they reside outside of the support of the Marčenko Pastur (MP) distribution
- Kritchman Nadler (KN) rank estimation to determine the number of eigenvalues corresponding to the signal <sup>3</sup>
- Apply operator norm eigenvalue shrinkage procedure (Donoho et al.) to those eigenvalues, while setting all other eigenvalues to 0 <sup>4</sup>

---

<sup>3</sup>Determining the number of components in a factor model from limited noisy data, Shira Kritchman and Boaz Nadler, Chemometrics and Intelligent Laboratory Systems

<sup>4</sup>Optimal shrinkage of eigenvalues in the spiked covariance model, David Donoho, Matan Gavish and I M Johnstone, [arxiv.org/abs/1311.0851](https://arxiv.org/abs/1311.0851)

# Wiener Filtering

- White noise: estimate  $X_i$  as

$$\hat{X}_i = (I - H_i A_i) \hat{\mu} + H_i Y_i \quad (13)$$

where  $H_i = \hat{\Sigma} A_i^T (A_i \hat{\Sigma} A_i^T + \sigma^2 I)^{-1}$  is the linear Wiener filter

- Colored noise: estimate  $X_i$  as

$$\hat{X}_i = (I - H_i W A_i) \hat{\mu} + H_i Y_i \quad (14)$$

with  $H_i = \hat{\Sigma} A_i^T W^T (W A_i \hat{\Sigma} A_i^T W^T + \sigma^2 I)^{-1}$

# Computational Complexity

$O(TDL^4 + nL^3)$ , where  $T$  is the number of conjugate gradient iterations

- $D$  defocus groups with  $d_i$  images in group  $i$
- Images of size  $L \times L$
- $n$  images

# Acknowledgement

- NIGMS, AFOSR, Simons Foundation, Moore Foundation Data-Driven Discovery Investigator Award
- Maofu Liao, Sjors Scheres, Irina Serysheva and Joachim Frank
- Yoel Shkolniskly, Fred Sigworth, Zhizhen Zhao, Joakim Andén