3.36pt

Fixed Point Algorithms for Phase Retrieval and Ptychography

Albert Fannjiang

University of California, Davis

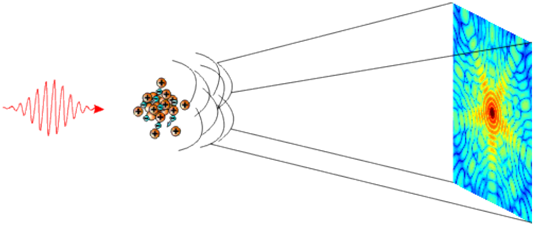
Mathematics of Imaging Workshop:
Variational Methods and Optimization in Imaging
IHP, Paris, February 4th-8th 2019

Collaborators: Pengwen Chen (NCHU), Gi-Ren Liu (NCKU), Zheqing Zhang (UCD)

Outline

- Introduction
- Alternating projection for feasibility
- Douglas-Rachford splitting/ADMM
- Convergence analysis
- Initialization methods
- Blind ptychoraphy
- Conclusion

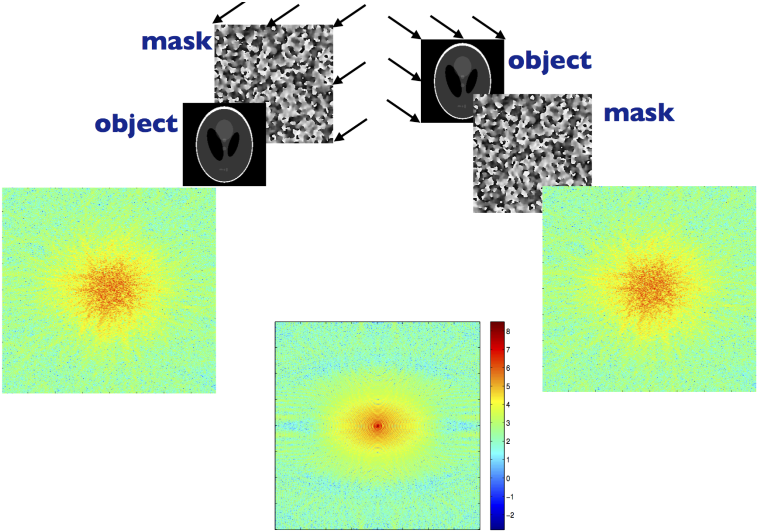
Phase retrieval



- X-ray crystallography: von Laue, Bragg etc. since 1912.
- Non-periodic structures: Gerchberg, Saxton, Fienup etc since 1972, delay due to low SNR.
- Nonlinear signal model: data $=$ diffraction pattern $= |\mathcal{F}(f)|^2$

$$\mathcal{F} = \text{Fourier transform}, \qquad |\cdot| = \text{componentwise modulus.}$$

Coded diffraction pattern



Alternating projections

Nonconvex feasibility

- Masking μ + propagation $\mathcal{F}$ + intensity measurement:

$$\text{coded diffraction pattern} = |\mathcal{F}(f \odot \mu)|^2.$$

- F (2012): Uniqueness with probability one

$$\begin{array}{rll} & b = |Ax|, & x \in \mathcal{X} \\ \text{(1 mask)} & \mathcal{X} = \mathbb{R}^n, & A = \Phi\ \mathrm{diag}(\mu) \\ \text{(2 masks)} & \mathcal{X} = \mathbb{C}^n, & A = \begin{bmatrix} \Phi\ \mathrm{diag}(\mu_1) \\ \Phi\ \mathrm{diag}(\mu_2) \end{bmatrix} \end{array}$$

- Non-convex feasibility:

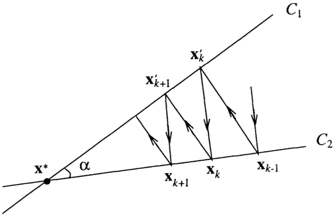
$$\begin{array}{rcl} \text{Find}\quad \hat{y} & \in & A\mathcal{X} \cap \mathcal{Y} \\ \mathcal{Y} & := & \{y \in \mathbb{C}^N : |y| = b\} \end{array}$$

Intersection of N-dim torus $\mathcal{Y}$ and n- or $2n$-dim subspace $A\mathcal{X}$

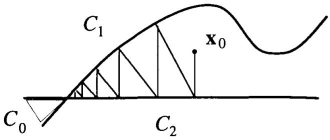
Alternating projections

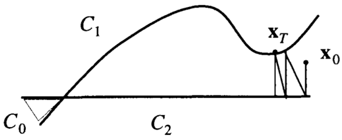
von Neuman 1933

Cheney-Goldstein 1959
Bregman 1965



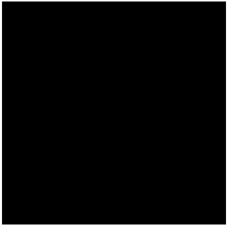
Non convex: local convergence?



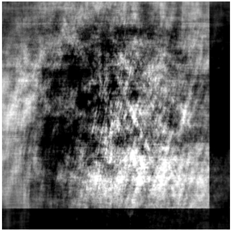


Coded vs plain diffraction pattern

(a) coded; 40 iter



(b) error



(c) plain;1000 iter

(d) error

- AP: real-valued Cameraman with one diffraction pattern.
- Plain diffraction pattern allows ambiguities such as translation, twin-image which are forbidden by the presence of a random mask.

Douglas-Rachford splitting

Alternating minimization

Minimization with a sum of two objective functions

$$\arg\min_u K(u) + \mathcal{L}(v), \quad u = v$$

where

$$\begin{aligned} K &= \text{Indicator function of} \quad \{Ax : x \in \mathbb{C}^n\} \\ \mathcal{L}(v) &= \sum_i |v[i]|^2 - b^2[i] \ln |v[i]|^2 \quad \text{(Poisson log-likelihood).} \end{aligned}$$

- Projection onto $K = AA^\dagger u$.
- Linear constraint $u = v$.
- $\mathcal{L}$ has a simple asymptotic form

Gaussian log-likelihood

- High SNR: Gaussian distribution with variance = mean: $\frac{e^{-(b^2-\lambda)^2/(2\lambda)}}{\sqrt{2\pi\lambda}}$.
- Gaussian log-likelihood: $\lambda = |v|^2$

$$\sum_j \ln|v[j]| + \frac{1}{2}\left|\frac{b^2[j]}{|v[j]|} - |v[j]|\right|^2 \longrightarrow \mathcal{L}$$

- In the vicinity of b, we make the substitution

$$\frac{b[j]}{|v[j]|} \to 1, \quad \ln|v[j]| \to \ln\sqrt{b[j]}$$

to obtain

$$\text{const.} + \frac{1}{2}\sum_j |b[j] - |v[j]||^2 \longrightarrow \mathcal{L}$$

which is the smoothest of the 3 functions.

Alternating projections revisited

- Hard constraint $u = v$

$$\arg\min_u K(u) + \mathcal{L}(u) \quad = \quad \arg\min_x \mathcal{L}(u), \qquad u = Ax$$

where

$$\begin{aligned} K &= \text{Indicator function of} \quad \{Ax : x \in \mathbb{C}^n\} \\ \mathcal{L}(u) &= \frac{1}{2}\|b - |u|\|^2 \quad \text{(Gaussian log-likelihood).} \end{aligned}$$

- $\mathcal{L}$ non-smooth where b vanishes.
- AP = gradient descent with unit stepsize: $x^{k+1} = x^k - \nabla\mathcal{L}(x^k)$.
- Wirtinger flow = gradient descent with

$$\mathcal{L} = \frac{1}{2}\||Ax|^2 - b\|^2 \quad \text{(additive i.i.d. Gaussian noise).}$$

Proximal optimality

- Proximity operators are generalization of projections:

$$\begin{aligned}\mathrm{prox}_{\mathcal{L}/\rho}(u) &= \arg\min_x \mathcal{L}(x) + \frac{\rho}{2}\|x-u\|^2 \\ \mathrm{prox}_{K/\rho}(u) &= AA^\dagger u.\end{aligned}$$

  For simplicity, set $\rho = 1$.
- Proximal reflectors $R_{\mathcal{L}} = 2\ \mathrm{prox}_{\mathcal{L}} - I, \quad R_K = 2\ \mathrm{prox}_K - I$
- Proximal optimality:

$$0 \in \partial\mathcal{L}(x) + \partial K(x) \quad \text{iff} \quad \xi = R_{\mathcal{L}} R_K(\xi), \quad x = \mathrm{prox}_K(\xi)$$

Proximal optimality: proof

- Let $\eta = R_K(\xi)$. Then $\xi = R_{\mathcal{L}}(\eta)$.
- Also $\zeta := \frac{1}{2}(\xi + \eta) = \mathrm{prox}_{\mathcal{L}}(\eta) = \mathrm{prox}_K(\xi)$. Equivalently

$$\xi \in \partial K(\zeta) + \zeta, \quad \eta \in \partial \mathcal{L}(\zeta) + \zeta$$

- Adding the two equations: $0 \in \partial K(\zeta) + \partial \mathcal{L}(\zeta)$.
- Finally $\zeta = \mathrm{prox}_K(\xi)$ is a stationary point.

Douglas-Rachford splitting (DRS)

- Optimality leads to Peaceman-Rachford splitting: $z^{k+1} = R_{\mathcal{L}/\rho} R_{K/\rho}(z^k)$.
- DRS $z^{l+1} = \frac{1}{2} z^l + \frac{1}{2} R_{\mathcal{L}/\rho} R_{K/\rho}(z^l)$: for $l = 1, 2, 3 \cdots$

$$
\begin{aligned}
y^{l+1} &= \mathrm{prox}_{K/\rho}(u^l); \\
z^{l+1} &= \mathrm{prox}_{\mathcal{L}/\rho}(2y^{l+1} - u^l) \\
u^{l+1} &= u^l + z^{l+1} - y^{l+1}.
\end{aligned}
$$

- $\gamma = 1/\rho =$ stepsize; $\rho = 0$ the classical DR algorithm.
- Alternating Direction Method of Multipliers (ADMM) applied to the dual problem

$$
\max_{\lambda} \min_{y,z} \mathcal{L}^*(y) + K^*(-A^* z) + \langle \lambda, y - A^* z \rangle + \frac{\rho}{2} \|A^* z - y\|^2
$$

DRS map

- Object update: $f = A^\dagger u^\infty$ where u^∞ is the terminal value of

$$\begin{aligned} u^{l+1} &= \frac{1}{\rho+1}u^l + \frac{\rho-1}{\rho+1}Pu^l + \frac{1}{\rho+1}b\odot\operatorname{sgn}\big(2Pu^l - u^l\big) \\ &= \frac{1}{2}u^l + \frac{\rho-1}{2(\rho+1)}Ru^l + \frac{1}{\rho+1}b\odot\operatorname{sgn}\big(Ru^l\big) \end{aligned}$$

  where $P = AA^\dagger$ is the orthogonal projection onto the range of A and $R = 2P - I$ is the corresponding reflector.

- $\rho = 0$: the classical Douglas-Rachford algorithm

$$\begin{aligned} u^{l+1} &= \frac{1}{2}u^l - \frac{1}{2}Ru^l u^l + b\odot\operatorname{sgn}\big(Ru^l\big) \\ &= u^l - Pu^l + b\odot\operatorname{sgn}\big(Ru^l\big). \end{aligned}$$

Convergence analysis

Convergence analysis

- Lewis-Malick (2008): local linear convergence of AP for transversally intersecting smooth manifolds.
- Lewis-Luke-Malick (2009): transversal intersection $\longrightarrow$ linearly regular intersection (LRI).
- Aragoón-Borwein (2012): global convergence of DR ($\rho = 0$) for intersection of a line and a circle.
- Hesse-Luke (2013): local geometric convergence of DR ($\rho = 0$) for LRI of an affine set and a super-regular set.
- Li-Pong (2016):
  - $\rightarrow$ $\mathcal{L}$ has uniformly Lipschitz gradient (ULG).
  - $\rightarrow$ DRS with ρ sufficiently large, depending on Lipschitz constant.
  - $\rightarrow$ Global convergence: cluster point = stationary point.
  - $\rightarrow$ Local geometric convergence for semi-algebraic case.

K and $\mathcal{L}$ don't have ULG and optimal performance is with $\rho \sim 1$.

- Candes et at. (2015): global convergence of Wirtinger flow with spectral initialization.

Fixed point equation

- Fixed point equation

$$u = \frac{1}{2}u + \frac{\rho-1}{2(\rho+1)} R_\infty u + \frac{1}{\rho+1} b \odot \operatorname{sgn}\left(R_\infty u\right)$$

- The differential map is given by $\Omega J_A(\eta)$ where

$$\begin{aligned} J_A(\eta) &= CC^\dagger\eta - \frac{1}{1+\rho}\Big[\Re(2CC^\dagger\eta - \eta) \\ &\quad + \imath\big(I - \mathrm{diag}(b/|Ru|)\big)\Im\left(2CC^\dagger\eta - \eta\right)\Big] \end{aligned}$$

where

$$\Omega = \mathrm{diag}(\operatorname{sgn}(Ru)), \quad C = \Omega^* A.$$

Fixed point analysis

Two randomly coded diffraction patterns:

- F (2012) – intersection $\sim S^1$ (arbitrary phase factor).
- Chen & F (2016) – DR ($\rho = 0$) fixed points u take the form

$$\begin{aligned} u &= e^{\mathrm{i}\theta}(b+r)\odot \mathrm{sgn}(Af), \quad r\in\mathbb{R}^N, \quad b+r\geq 0 \\ \Longrightarrow \mathrm{sgn}(u) &= \theta + \mathrm{sgn}(Af) \end{aligned}$$

  where r is a real null vector of $A^\dagger \mathrm{diag}[\mathrm{sgn}(Af)]$
  $\Longrightarrow$ DR fixed point set has real dimension $N-n$.
- Chen, F & Liu (2016) – AP based on the hard constraint $u = v$

$$\text{AP fixed point } x_*: \quad \|Ax_*\| = \|Af\| \quad \text{iff} \quad x_* = \alpha f, \quad |\alpha| = 1.$$

Spectral gap and linear convergence rate

J_A can be analyzed by the eigen-structure of

$$H := \begin{bmatrix} \Re[A^\dagger \Omega] \\ \Im[A^\dagger \Omega] \end{bmatrix}, \quad \Omega = \mathrm{diag}(\mathrm{sgn}(Af)).$$

- $\|J_A(\eta)\| = \|\eta\|$ occurs at $\eta = \pm \mathrm{i} b$.
- Linear convergence rate is related to the spectral gap of H.
- One randomly coded diffraction pattern:
  - → Chen & F (2016) – the differential map at Af has the largest singular value 1 corresponding to the constant phase and a positive spectral gap $\Longrightarrow$ the true solution is an attractor (local linear convergence).
  - → F & Zhang (2018) – the differential map at any DR fixed point has a spectral radius $= 1$.
  - → Chen, F & Liu (2016) – same for AP (parallel or serial).

DRS fixed points

Proposition

Let u be a fixed point and $f_\infty := A^\dagger u$.
(i) $\rho \geq 1$: If $\|J_A(\eta)\|_2 \leq \|\eta\|_2$ then $|\mathcal{F}(\mu, f_\infty)| = b$.
(ii) $\rho \geq 0$: If $|\mathcal{F}(\mu, f_\infty)| = b$ then $\|J_A(\eta)\|_2 \leq \|\eta\|_2$. where the equality holds iff η parallels $\imath b$.

Summary:

- DRS ($\rho \geq 1$) fixed point is linearly stable iff it is a true solution
- DR ($\rho = 0$) introduces harmless, stable fixed points.
- AP likely introduces spurious nonsolution fixed points.
- Linear convergence rate:

$$\text{Serial AP} < \text{parallel AP} \sim \text{DRS } (\rho = 1) < \text{DR } (\rho = 0).$$

Initialization

Initialization by feature extraction

$b = |Af|$ where $A \in \mathbb{C}^{N\times n}$ is the measurement matrix.

Feature: two sets of signals, weak and strong.

- Weak signals selected by a threshold τ, i.e. $b_i \leq \tau$, $i \in I$.
- $x_{\text{null}} :=$ ground state of A_I.

Isometry: $\|Ax\|^2 = \|A_I x\|^2 + \|A_{I_c} x\|^2 = \|x\|^2 \Longrightarrow$

$$\begin{aligned} x_{\text{null}} &= \arg\min \left\{ \|A_I x\|^2 : \|x\| = \|f\| \right\} \\ &= \arg\max \left\{ \|A_{I_c} x\|^2 : \|x\| = \|f\| \right\} \end{aligned}$$

solved by the power method efficiently.

Non-isometry $\Longrightarrow$ QR: $A = QR$

Null vector algorithm

Let $\mathbf{1}_c$ be the characteristic function of the complementary index I_c with $|I_c| = \gamma N$.

Algorithm 1: The null vector method

Random initialization: $x_1 = x_{\text{rand}}$
Loop:
for $k = 1 : k_{\max} - 1$ **do**
  $x'_k \leftarrow A(\mathbf{1}_c \odot A^* x_k)$;
  $x_{k+1} \leftarrow \left[x'_k\right]_{\mathcal{X}} / \| \left[x'_k\right]_{\mathcal{X}} \|$
end
Output: $x_{\text{null}} = x_{k_{\max}}$.

Algorithm 2: The spectral vector method

Random initialization: $x_1 = x_{\text{rand}}$
Loop:
for $k = 1 : k_{\max} - 1$ **do**
  $x'_k \leftarrow A(|b|^2 \odot A^* x_k)$;
  $x_{k+1} \leftarrow \left[x'_k\right]_{\mathcal{X}} / \| \left[x'_k\right]_{\mathcal{X}} \|$;
end
Output: $x_{\text{spec}} = x_{k_{\max}}$.

Netrapalli-Jain-Sanghavi 2015

Truncated spectral vector

$$x_{\text{t-spec}} = \arg\max_{\|x\|=1} \left\| A \left(\mathbf{1}_\tau \odot |b|^2 \odot A^* x \right) \right\|$$

$$\{i : |A^* x(i)| \le \tau \|b\|\}$$

Candes-Chen 2015

Performance guarantee: Gaussian case

Theorem (Chen-F.-Liu 2016)

Let A be drawn from the $n \times N$ standard complex Gaussian ensemble. Let

$$\sigma := |I|/N < 1, \quad \nu = n/|I| < 1.$$

Then for any $x_0 \in \mathbb{C}^n$ the following error bound

$$\|x_0 x_0^* - x_{\text{null}} x_{\text{null}}^*\|^2 \quad \leq \quad c_0 \sigma \|x_0\|^4$$

holds with probability at least

$$1 - 5 \exp\left(-c_1 |I|^2/N\right) - 4 \exp(-c_2 n).$$

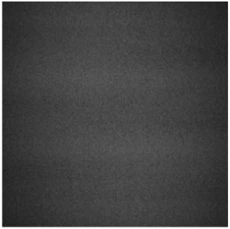
- Non-asymptotic estimate: $n < |I| < N < |I|^2, \quad L = N/n$

$$|I| = N^\alpha n^{1-\alpha} \Longrightarrow \mathrm{RE} \sim L^{(\alpha-1)/2}, \ \alpha \in [1/2, 1)$$

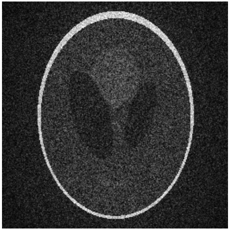
2 CDPs, $|I| = \sqrt{nN}$.

Uniqueness of phase retrieval with 2 CDPs (F. 2012).

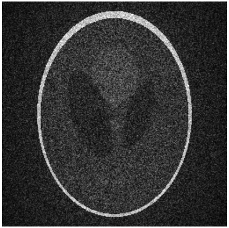
(e) phantom



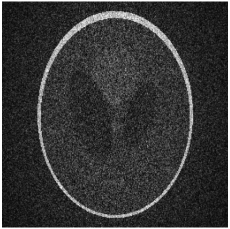
(f) Spectral vector



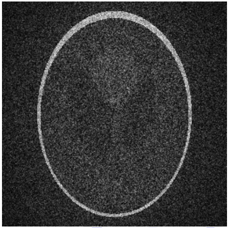
(g) Null vector



(h) NSR=10%

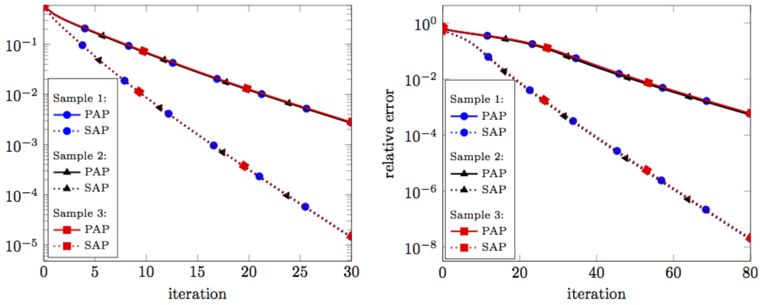


(i) NSR=15%



(j) NSR=20%

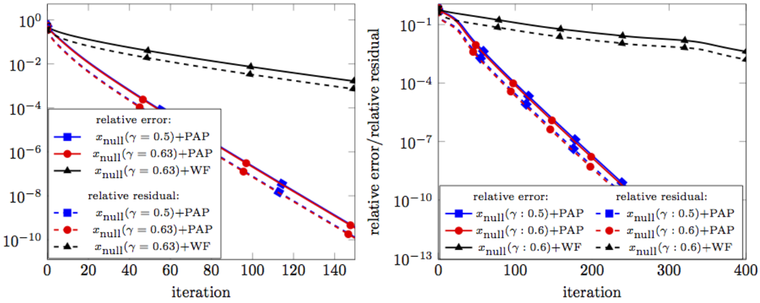
Experiments: with null initialization



(a) RSCB (b) RPP

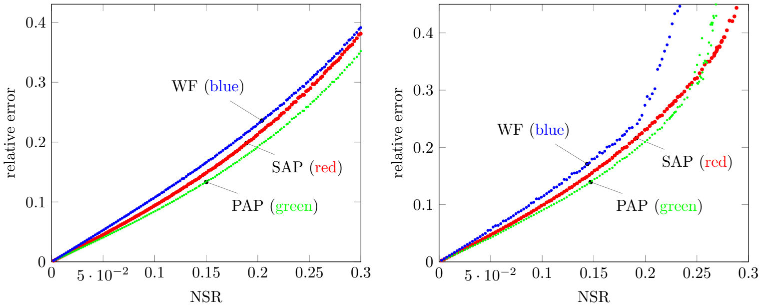
- PAP: two diffraction patterns used in parallel
- SAP: two diffraction patterns used in serial

Comparison with Wirtinger flow



(a) RSCB (b) RPP

Complex Gaussian noise



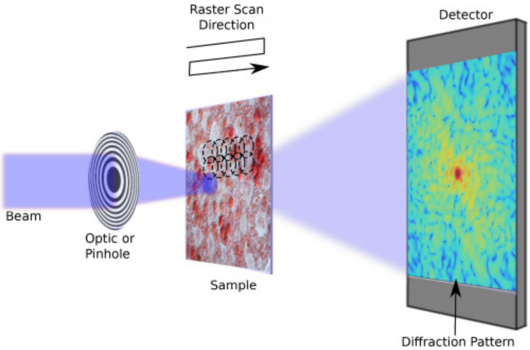
(a) RSCB (b) RPP

- $b = |Af + \text{complex Gaussian noise}|$
- NSR $=$ noise/signal

Blind ptychography

Ptychography: extended objects

Hoppe (1969), Nellist-Rodenburg (95), Faulkner-Rodenburg (04, 05). , Thibault *et al.* (08, 09)



- Inverse problem with shifted *windowed* Fourier intensities.
- Unlimited, extended objects: structural biology, materials science etc.

Linear phase ambiguity

Consider the probe and object estimates

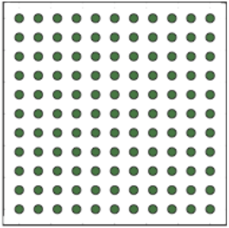
$$\begin{aligned} \nu^0(\mathbf{n}) &= \mu^0(\mathbf{n}) \exp(-\mathrm{i}a - \mathrm{i}\mathbf{w} \cdot \mathbf{n}), \quad \mathbf{n} \in \mathcal{M}^0 \\ g(\mathbf{n}) &= f(\mathbf{n}) \exp(\mathrm{i}b + \mathrm{i}\mathbf{w} \cdot \mathbf{n}), \quad \mathbf{n} \in \mathbb{Z}_n^2 \end{aligned}$$

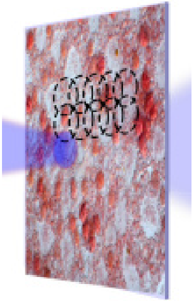
for any $a, b \in \mathbb{R}$ and $\mathbf{w} \in \mathbb{R}^2$. We have all $\mathbf{n} \in \mathcal{M}^{\mathbf{t}}, \mathbf{t} \in \mathcal{T}$

$$\nu^{\mathbf{t}}(\mathbf{n}) g^{\mathbf{t}}(\mathbf{n}) = \mu^{\mathbf{t}}(\mathbf{n}) f^{\mathbf{t}}(\mathbf{n}) \exp(\mathrm{i}(b-a)) \exp(\mathrm{i}\mathbf{w} \cdot \mathbf{t}).$$

Raster scan pathology

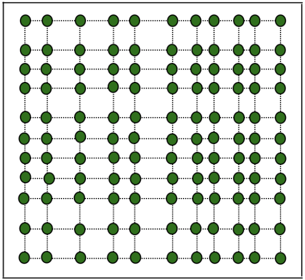
Raster scan: $\mathbf{t}_{kl} = \tau(k, l), k, l \in \mathbb{Z}$ where τ is the step size.
$\mathcal{M} = \mathbb{Z}_n^2$, $\mathcal{M}^0 = \mathbb{Z}_m^2$, $n > m$, with the periodic boundary condition.

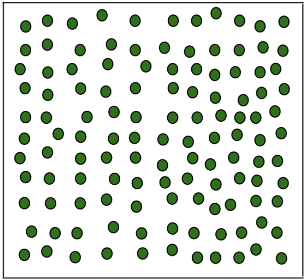




Mixing schemes

- **Partial perturbation** $\mathbf{t}_{kl} = \tau(k,l) + (\delta^1_k, \delta^2_l)$.
- **Full perturbation** $\mathbf{t}_{kl} = \tau(k,l) + (\delta^1_{kl}, \delta^2_{kl})$.

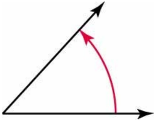




Mask phase constraint (MPC)

- μ^0: independent phases with range $\geq \pi$.
- ν^0 satisfies MPC if $\nu_0(\mathbf{n})$ and $\mu^0(\mathbf{n})$ form an acute angle

$$|\arg[\nu^0(\mathbf{n})/\mu^0(\mathbf{n})]| < \pi/2$$



Global uniqueness

Theorem (F 2018)

Suppose f does not vanish in $\mathbb{Z}_n^2$. Let $a_j^i = 2\delta_{j+1}^i - \delta_j^i - \delta_{j+2}^i$ and let $\{\delta_{j_k}^i\}$ be the subset of perturbations satisfying $\gcd_{j_k}\{|a_{j_k}^i|\} = 1, \quad i = 1, 2,$ *and*

$$2\tau \le m - \max_{i=1,2}\{\delta_{j_k+2}^i - \delta_{j_k}^i\} \quad \textit{(Overlap > 50\%)}$$

$$\max_{i=1,2}[|a_{j_k}^i| + \max_{k'}\{\delta_{k'+1}^i - \delta_{k'}^i\}] \le m - \tau$$

$$\delta_{j_k+1}^i - \delta_{j_k+2}^i \le \tau \le m - 1 + \delta_{j_k+1}^i - \delta_{j_k+2}^i.$$

Then APA and SF are the only ambiguities, i.e. for some explicit $\mathbf{r}$

$$\begin{aligned} g(\mathbf{n})/f(\mathbf{n}) &= \alpha^{-1}(0)\exp(\mathrm{i}\mathbf{n}\cdot\mathbf{r}), \\ \nu^0(\mathbf{n})/\mu^0(\mathbf{n}) &= \alpha(0)\exp(\mathrm{i}\phi(0) - \mathrm{i}\mathbf{n}\cdot\mathbf{r}) \\ \theta_{kl} &= \theta_{00} + \mathbf{t}_{kl}\cdot\mathbf{r}. \end{aligned}$$

Initialization with mask phase constraint

- Mask/probe initialization

$$\mu_1(\mathbf{n}) = \mu^0(\mathbf{n}) \exp\left[\mathrm{i}\phi(\mathbf{n})\right],$$

  where $\phi(\mathbf{n})$ i.i.d. uniform on $(-\pi/2, \pi/2)$
  Relative error of the mask estimate

$$\sqrt{\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} |e^{\mathrm{i}\phi} - 1|^2 d\phi} = \sqrt{2(1 - \frac{2}{\pi})} \approx 0.8525$$

- Object initialization: f_1 = constant or random phase object.

Alternating minimization

$|\mathcal{F}(\mu, f)| = b$: the ptychographic data. Define $A_k h := \mathcal{F}(\mu_k, h)$, $B_k \eta := \mathcal{F}(\eta, f_{k+1})$. We have $A_k f_{j+1} = B_j \mu_k$.

1. Initial guess μ_1.
2. Update the object estimate $\quad f_{k+1} = \operatorname*{argmin}_{g \in \mathbb{C}^{n\times n}} \mathcal{L}(A_k^* g)$
3. Update the probe estimate $\quad \mu_{k+1} = \operatorname*{argmin}_{\nu \in \mathbb{C}^{m\times m}} \mathcal{L}(B_k^* \nu)$
4. Terminate when $\| |B_k^* \mu_{k+1}| - b \|$ is less than tolerance or stagnates. If not, go back to step 2 with $k \to k+1$.

Fixed point algorithm with $\rho = 1$

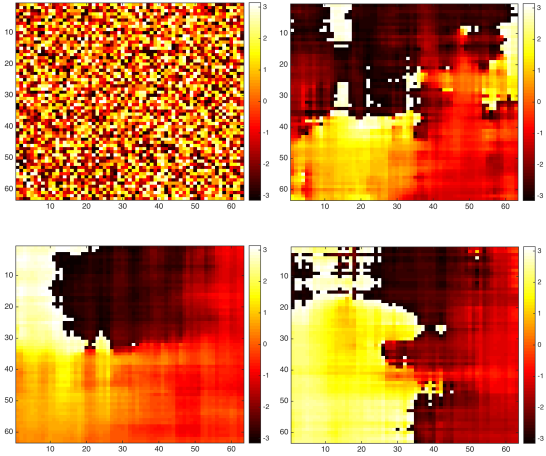
- $\rho = 1$
- Reflectors: $R_k = 2P_k - I, S_k = 2Q_k - I$.
- Gaussian:

$$
\begin{aligned}
u_k^{l+1} &= \frac{1}{2}u_k^l + \frac{1}{2}b \odot \operatorname{sgn}\left(R_k u_k^l\right) \\
v_k^{l+1} &= \frac{1}{2}v_k^l + \frac{1}{2}b \odot \operatorname{sgn}\left(S_k v_k^l\right).
\end{aligned}
$$

- Poisson:

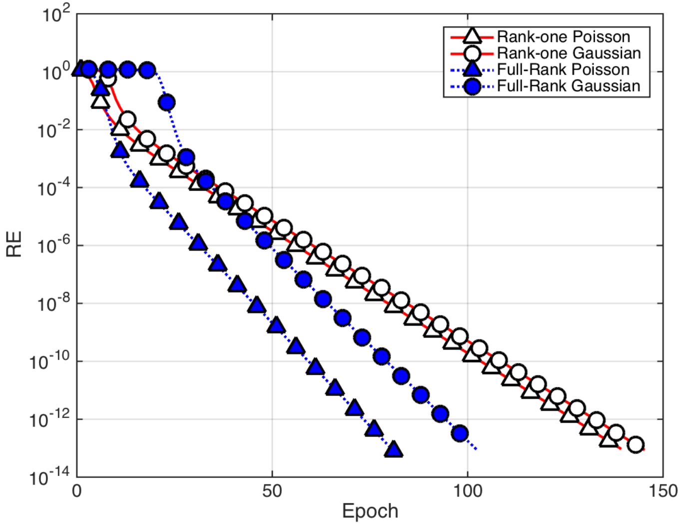
$$
\begin{aligned}
u_k^{l+1} &= \frac{1}{2}u_k^l - \frac{1}{3}R_k u_k^l + \frac{1}{6}\sqrt{|R_k u_k^l|^2 + 24b^2} \odot \operatorname{sgn}\left(R_k u_k^l\right) \\
v_k^{l+1} &= \frac{1}{2}v_k^l - \frac{1}{3}S_k v_k^l + \frac{1}{6}\sqrt{|S_k v_k^l|^2 + 24b^2} \odot \operatorname{sgn}\left(S_k v_k^l\right).
\end{aligned}
$$

Masks

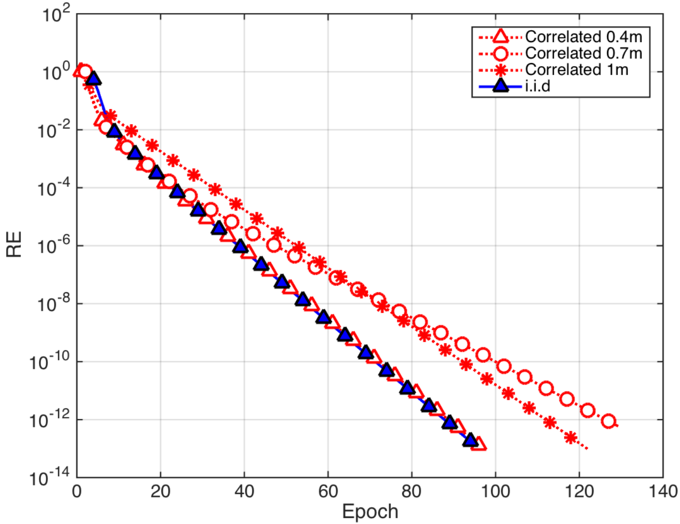


correlation length $c = 0, 0.4m, 0.7m, 1m$

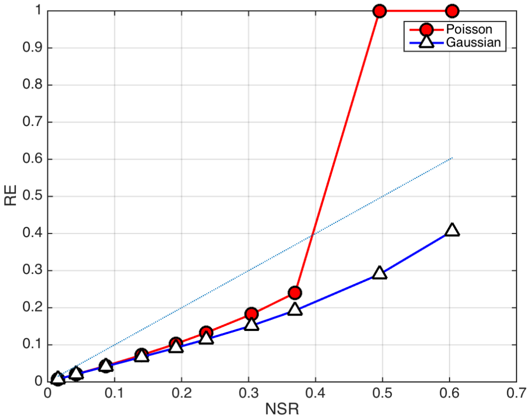
Rank-one vs. full-rank



Independent vs. correlated mask



Poisson noise



- Photon counting noise: $b^2 =$ Poisson r.v. with mean $= |Af|^2$.
- Gaussian log-likelihood outperforms Poisson log-likelihood.

Conclusion

1. Disorder can better condition measurement schemes: random mask, random perturbation to raster scan
2. Analytical and statistical considerations can guide our way to a better objective function
3. Fixed point analysis can help determine parameters or select algorithms
4. Initialization by feature extraction

Thank you!

References

1. F (2012), "Absolute uniqueness of phase retrieval with random illumination," *Inverse Problems* **28** 075008.
2. Netrapalli, Jain & Sanghavi (2015) "Phase retrieval using alternating minimization," *IEEE Transactions on Signal Processing* **63** 4814-4826.
3. Chen, F. & Liu (2017) "Phase retrieval by linear algebra". *SIAM J. Matrix Anal. Appl.* **38** 854-868.
4. Chen, F & Liu (2018) "Phase retrieval with one or two diffraction patterns by alternating projections of the null vector". *J. Fourier Anal. Appl.* **24** 719-758.
5. Chen & F. (2018) "Fourier phase retrieval with a single mask by Douglas-Rachford Algorithm," *Appl. Comput. Harm. Anal.***44** (2018) 665-69.
6. Chen & F (2017), "Coded-aperture ptychography: Uniqueness and reconstruction," *Inverse Problems* **34**, 025003.
7. F & Chen (2018), "Blind ptychography: Uniqueness & ambiguities," arXiv: 1806.02674.
8. F & Zhang (2018), "Blind Ptychography by Douglas-Rachford Splitting," arXiv: 1809.00962
9. F (2018): "Raster Grid Pathology and the Cure" arXiv: 1810.00852