When deep denoising meets iterative phase retrieval Supplementary Material

In this supplementary material, we provide proofs on the proximal operators used in our algorithms and show how ADMM (Boyd et al., 2011) with indicator functions coincides with Hybrid-Input-Output (HIO) (Fienup, 1982) and Hybrid-Projection-Reflection (HPR) (Bauschke et al., 2003).

1 Proximal operators

We consider two proximal operators for Fourier phase retrieval: the squared error of Fourier amplitudes and regularization by denoising (RED) coupled with additional object-space constraints.

1. $R(x) = \bar{I}_{\mathcal{C}}(x) + \frac{\lambda}{2}\langle x, x - D(x)\rangle$

   Let D be the denoiser used in RED and $\mathcal{C}$ be the set of signals satisfying the additional constraints provided, where we assume that the denoiser D is (locally) homogeneous with symmetric Jacobian (Romano et al., 2017) and $\mathcal{C}$ is a convex set. For any $\tau > 0$, if $v^+ = \text{prox}_{\tau R}(v)$, then the first-order optimality condition gives

$$\begin{aligned} v^+ &= \underset{x\in\mathbb{R}^n}{\text{armin}}\; \tau R(x) + \frac{1}{2}\|v - x\|^2 \\ \Rightarrow &\; \tau(\partial \bar{I}_{\mathcal{C}}(v^+) + \lambda(v^+ - D(v^+)) + v^+ - v = 0 \\ \Leftrightarrow v^+ &= \left(I + \frac{\tau}{1+\lambda\tau}\partial \bar{I}_{\mathcal{C}}\right)^{-1}\left(\frac{v + \lambda\tau D(v^+)}{1+\lambda\tau}\right) \\ \Leftrightarrow v^+ &= \Pi_{\mathcal{C}}\left(\frac{v + \lambda\tau D(v^+)}{1+\lambda\tau}\right) \end{aligned} \tag{S1}$$

   where $\partial \bar{I}_{\mathcal{C}}$ is the subgradient of the indicator function and the last equality follows by noting that the resolvent of $\partial \bar{I}_{\mathcal{C}}$ is the projection $\Pi_{\mathcal{C}}$ onto $\mathcal{C}$ (Ryu & Boyd, 2016).

2. $f(z) = \frac{1}{2}\|y - |Fz|\|^2$

   Let F be the (normalized) discrete Fourier transform and y be the measured Fourier amplitude, which is non-negative. For simplicity, we consider 1D signals only (the conclusion holds for any dimension). Using the overhead symbol $\hat{\cdot}$ to denote the signal after Fourier transform, Parseval's theorem gives

$$\begin{aligned} x^+ &= \text{prox}_{\tau f}(x) = \text{argmin}_z\, \frac{\tau}{2}\|y - |Fz|\|_2^2 + \frac{1}{2}\|x - z\|^2 \\ \Leftrightarrow \widehat{x^+} &= \text{argmin}_{\hat{z}}\, \frac{\tau}{2}\|y - |\hat{z}|\|_2^2 + \frac{1}{2}\|\hat{x} - \hat{z}\|^2 \\ &= \text{argmin}_{\hat{z}}\, \frac{1}{2}\sum_k \tau(|\hat{z}[k]| - y[k])^2 + |\hat{z}[k] - \hat{x}[k]|^2 \end{aligned} \tag{S2}$$

   It was noticed in (Wen et al., 2012) that the solution is

$$\widehat{x^+}[k] = \frac{\tau}{\tau+1} y[k]\frac{\hat{x}[k]}{|\hat{x}[k]|} + \frac{1}{\tau+1}\hat{x}[k] \quad \forall k \tag{S3}$$

   which follows from the first-order optimality condition. Here, we provide an alternative proof that this solution is the global minimum.

We start by using the triangle inequality $|\hat{z}[k]-\hat{x}[k]|^2 \geq (|\hat{z}[k]|-|\hat{x}[k]|)^2$ to give the lower bound

$$\min_{\hat{z}} \sum_k \tau(|\hat{z}[k]|-y[k])^2 + |\hat{z}[k]-\hat{x}[k]|^2 \geq \min_{\hat{z}} \sum_k \tau(|\hat{z}[k]|-y[k])^2 + (|\hat{z}[k]|-|\hat{x}[k]|)^2 \tag{S4}$$

Equality between the right- and left-hand sides is achieved when

$$\Re(\overline{\hat{z}[k]}\hat{x}[k]) = |\hat{z}[k]\hat{x}[k]| \quad \forall k \tag{S5}$$

i.e., when the complex phase $\angle\hat{z}[k] = \angle\hat{x}[k]$ ($\angle\hat{z}[k]$ can be arbitrary if $\hat{x}[k]=0$). As the right-hand side is convex on $|\hat{z}[k]|$, the minimum is achieved when

$$|\hat{z}[k]| = \frac{\tau y[k] + |\hat{x}[k]|}{\tau+1} \quad \forall k \tag{S6}$$

as $y[k], |\hat{x}[k]| \geq 0$. Therefore, if x^+ minimizes (S2), then for all k,

$$\begin{aligned} \widehat{x^+}[k] &= \frac{\tau y[k] + |\hat{x}[k]|}{\tau+1}\exp(i\angle\hat{x}[k]) \\ &= \frac{\tau}{\tau+1}y[k]\exp(i\angle\hat{x}[k]) + \frac{1}{\tau+1}|\hat{x}[k]|\exp(i\angle\hat{x}[k]) \\ &= \frac{\tau}{\tau+1}y[k]\frac{\hat{x}[k]}{|\hat{x}[k]|} + \frac{1}{\tau+1}\hat{x}[k] \end{aligned} \tag{S7}$$

Performing an inverse Fourier transform gives (26) in the main text:

$$x^+ = \frac{\tau}{1+\tau}\Pi_{\mathcal{M}}(x) + \frac{1}{\tau+1}x \tag{S8}$$

2 Equivalence between ADMM and HIO\HPR

Let x_0 be the ground truth and S and $\tilde{S}$ be the support for x_0 and the extended support for padded $\tilde{x}_0 = P_{mn}x_0$, respectively.

If there is additional information about the signal support, e.g. an estimation γ such that $S \subseteq \gamma$, then the relation $\tilde{S} \subseteq \tilde{\gamma}$ holds for the extended support as well. For example, if we use the same vectorization as in the main text, such that

$$\tilde{x} = P_{mn}x = \begin{bmatrix} x \\ 0_{m-n} \end{bmatrix} \tag{S9}$$

then we will have $S = \tilde{S}$ and $\gamma = \tilde{\gamma}$. Define subset $\mathcal{S}$ for the signals satisfying the given support constraint,

$$\mathcal{S} := \{x \in \mathbb{C}^n \mid x_i = 0\ \forall i \notin \gamma\} \tag{S10}$$

The projection onto $\mathcal{S}$ is

$$\Pi_{\mathcal{S}}(x)_i = \begin{cases} x_i & \text{if } i \in \gamma \\ 0 & \text{otherwise} \end{cases} \tag{S11}$$

and similarly for $\tilde{\mathcal{S}} := \{x \in \mathbb{C}^m \mid x_i = 0\ \forall i \notin \tilde{\gamma}\}$ on the extended support.

According to (Bauschke et al., 2002), HIO with $\beta = 1$ can be written as

$$\tilde{x}^{k+1} = \Pi_{\tilde{\mathcal{S}}}(2\Pi_{\mathcal{M}}(\tilde{x}^k) - \tilde{x}^k) - \Pi_{\mathcal{M}}(\tilde{x}^k) + \tilde{x}^k \tag{S12}$$

We now relate this to the optimization of FPR with the support constraint

$$\begin{aligned} &\underset{x\in\mathbb{C}^n, z\in\mathbb{C}^m}{\text{minimize}}\ \bar{I}_{\mathcal{M}}(z) + \bar{I}_{\mathcal{S}}(x) \\ &\text{subject to } z = O_{mn}x \end{aligned} \tag{S13}$$

With $\tilde{x} = O_{mn}x$, this can be rewritten as

$$\begin{aligned} &\underset{\tilde{x}, z\in\mathbb{C}^m}{\text{minimize}}\ \bar{I}_{\mathcal{M}}(z) + \bar{I}_{\tilde{\mathcal{S}}}(\tilde{x}) \\ &\text{subject to } z = \tilde{x} \end{aligned} \tag{S14}$$

for which ADMM gives the update rule as

$$
\begin{aligned}
\tilde{x}^{k+1} &= \Pi_{\tilde{\mathcal{S}}}(z^k + u^k) \\
z^{k+1} &= \Pi_{\mathcal{M}}(\tilde{x}^{k+1} - u^k) \\
u^{k+1} &= u^k + z^{k+1} - \tilde{x}^{k+1}
\end{aligned} \tag{S15}
$$

As in (Wen et al., 2012), the updates for $m^{k+1} = \tilde{x}^{k+1} - u^k$ are given by

$$
\begin{aligned}
m^{k+2} &= \tilde{x}^{k+2} - u^{k+1} \\
&= \Pi_{\tilde{\mathcal{S}}}(2\Pi_{\mathcal{M}}(m^{k+1}) - m^{k+1}) - \Pi_{\mathcal{M}}(m^{k+1}) + m^{k+1}
\end{aligned} \tag{S16}
$$

which coincides with (S12).

Next, we denote $\mathcal{S}_+$ as the set containing signals which not only satisfy the support constraint but also have non-negative elements in the real part:

$$
\mathcal{S}_+ := \{x \in \mathbb{C}^n \mid x_i = 0 \ \forall i \notin \gamma \text{ and } \Re(x_i) \geq 0 \ \forall i\} \tag{S17}
$$

The projection onto $\mathcal{S}_+$ is

$$
\Pi_{\mathcal{S}_+}(x) = \Pi_{Re_+}\left(\Pi_{\mathcal{S}}(x)\right) \tag{S18}
$$

with Π_{Re_+} being the element-wise projection

$$
\Pi_{Re_+}(x)_i = \begin{cases} x_i & \text{if } \Re(x_i) \geq 0 \\ i\Im(x_i) & \text{otherwise} \end{cases} \tag{S19}
$$

Changing $\mathcal{S}$ to $\mathcal{S}_+$ in (S14) and repeating (S15) to (S16) gives the recursion for m^{k+1} as

$$
m^{k+2} = \Pi_{\tilde{\mathcal{S}}_+}(2\Pi_{\mathcal{M}}(m^{k+1}) - m^{k+1}) - \Pi_{\mathcal{M}}(m^{k+1}) + m^{k+1} \tag{S20}
$$

which coincides with HPR with $\beta = 1$ (Bauschke et al., 2003).

References

Bauschke, H. H., Combettes, P. L., and Luke, D. R. Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization. *J. Opt. Soc. Am. A*, 19(7):1334–1345, Jul 2002.

Bauschke, H. H., Combettes, P. L., and Luke, D. R. Hybrid projection–reflection method for phase retrieval. *J. Opt. Soc. Am. A*, 20(6):1025–1034, Jun 2003.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

Fienup, J. R. Phase retrieval algorithms: a comparison. *Appl. Opt.*, 21(15):2758–2769, Aug 1982.

Romano, Y., Elad, M., and Milanfar, P. The little engine that could: Regularization by denoising (red). *SIAM Journal on Imaging Sciences*, 10(4):1804–1844, 2017.

Ryu, E. K. and Boyd, S. Primer on monotone operator methods. *Appl. Comput. Math*, 15(1):3–43, 2016.

Wen, Z., Yang, C., Liu, X., and Marchesini, S. Alternating direction methods for classical and ptychographic phase retrieval. *Inverse Problems*, 28(11):115010, oct 2012.