Undersampled Phase Retrieval with Outliers: Supplementary Material

Daniel S. Weller, Member, IEEE, Ayelet Pnueli, Gilad Divon, Ori Radzyner, Yonina C. Eldar, Fellow, IEEE, and Jeffrey A. Fessler, Fellow, IEEE

Here we provide additional mathematical details for the $u$ update step with different data fit terms and additional experimental support related to the paper [1].

I. UPDATING $u$: SQUARED-MAGNITUDE MEASUREMENTS, LAPLACE DATA FIT TERM

In this case, $f(\cdot) = (\cdot)$, and $q = 2$. When $y_m < 0$, $f_+(u)$ is always greater than $f_-(u)$, so the solution is always the minimizer of $f_+(u)$. Otherwise, we must consider all three cases.

Let $d = [Ax + b]_m$, $s$ represent the appropriate choice of $s_m$ or $\bar{s}_m$, $\eta = \mu/2$, and drop the subscripts. Writing out $f_+(u)$ and $f_-(u)$,

$$f_+(u) = \eta |u - d|^2 + |u|^2 - y,$$

$$f_-(u) = \eta |u - d|^2 + y + |s|^2 - 2|s|\Re\{ue^{-i\eta}\}.$$

The function $f_+(u)$ is quadratic in $u$; differentiating yields

$$\frac{df_+(u)}{du} = 2\eta(u - d) + 2u.$$

Thus, $f_+(u)$ is minimized by $u_+ = \frac{\eta}{1 + \eta} d$.

The function $f_-(u)$ is also a quadratic, so

$$\frac{df_-(u)}{du} = 2\eta(u - d) - 2s,$$

which set to zero yields the minimizer $u_- = \frac{s}{\eta} + d$.

The minimization of $f_+(u)$ or $f_-(u)$ along the curve on which both functions are equal-valued, involves parameterizing this curve and minimizing $f_+(u)$ as a function of this parameter. These functions are equal when $|u|^2 - y = y + |s|^2 - 2|s|\Re\{ue^{-i\eta}\}$, which corresponds to the circle $|u + s|^2 = 2(y + |s|^2)$. The parameterization then corresponds to the angle along the circle; call it $\theta$. The curve of interest is $(u + s) = \sqrt{2(y + |s|^2)}e^{i\theta}$.

Incorporating this parameterization into $f_+(u)$ yields

$$f_+(u(\theta)) = -2\sqrt{2(y + |s|^2)}\Re\{(1 + \eta)s + \eta d)e^{-i\theta}\}$$

+ constants,

which is minimized when $\theta = \angle((1 + \eta)s + \eta d)$. So, $u_\pm = \sqrt{2(y + |s|^2)}e^{i\angle((1 + \eta)s + \eta d) - s}$.

DSW was funded by National Institutes of Health (NIH) grant F32 EB015914. JAF is funded in part by NIH grant U01 EB018753 and an equipment donation from Intel. YCE is funded in part by Israel Science Foundation Grant 170/10, SRC, and Intel Collaborative Research Institute for Computational Intelligence.

DSW is with the Charles L. Brown Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904 USA (email: dweller@virginia.edu). AP was with, and GD, OR, and YCE are with the Electrical Engineering Department, Technion, Israel Institute of Technology, Haifa 32000, Israel (emails: ayeltpnueli@gmail.com, giladd44@gmail.com, radzy@campus.technion.ac.il, yonina@ee.technion.ac.il). JAF is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (email: fessler@umich.edu).
II. UPDATING $u$: SQUARED-MAGNITUDE MEASUREMENTS, GAUSSIAN DATA FIT TERM

In this case, $f(\cdot) = (\cdot)^2$, and $q = 2$. Again, as with the Laplace data fit term, when $y_m < 0$, $f_+(u) > f_-(u)$, so we always minimize $f_+(u)$. Otherwise, we consider all three cases.

Again, let $d = [Ae^{i\theta_1} + b]^m$, $s$ represent the appropriate choice of $s_m$ or $\bar{s}_m$, $\eta = \mu/2$, and drop the subscripts. Writing out $f_+(u)$ and $f_-(u)$,

$$
f_+(u) = \eta |u - d|^2 + (|u|^2 - y)^2,
$$
$$
f_-(u) = \eta |u - d|^2 + (y + |s|^2 - 2|s|\text{Re}\{ue^{-i\zeta}\})^2.
$$

Writing $f_+(u)$ in terms of the magnitude $|u|$ and phase $\angle u$ of $u$,

$$
f_+(u) = \eta |u|^2 + \eta |d|^2 - 2\eta |u||d| \cos(\angle u - \angle d)
+ |u|^4 - 2y|u|^2 + y^2,
$$

which is clearly minimized when $\angle u = \angle d$, when $\cos(\angle u - \angle d) = 1$. Then, $f_+(u)$ becomes a quartic equation in $|u|$, which has the derivative

$$
\frac{df_+(u)}{d|u|} = 4|u|^3 + (2\eta - 4y)|u| - 2\eta |d|.
$$

The function $f_+(u)$ is minimized either when the derivative is zero or when $|u| = 0$. The depressed cubic equation will have between zero and three nonnegative real roots, which can be found analytically. Note that if there are three positive real roots, since the cubic must be increasing below the least positive root, the derivative at $|u| = 0$ is negative, and the fourth candidate point $|u| = 0$ cannot be the global minimum. The minimizer $u_+$ is the candidate point with minimum function value $f_+(u)$, multiplied by $e^{i\zeta d}$.

Finding a minimum of $f_-(u)$ is straightforward. Define $\bar{u} = ue^{-i\zeta s}$, and $\bar{d} = de^{-i\zeta s}$. Then,

$$
f_-(\bar{u}) = \eta |\bar{u} - \bar{d}|^2 + (y + |s|^2 - 2|s|\text{Re}\{\bar{u}\})^2.
$$

Separating the real and imaginary parts, we observe

$$
f_-(\bar{u}) = \eta (\text{Re}\{\bar{u}\} - \text{Re}\{\bar{d}\})^2 + \eta (\text{Im}\{\bar{u}\} - \text{Im}\{\bar{d}\})^2
+ (y + |s|^2 - 2|s|\text{Re}\{\bar{d}\})^2,
$$

which is clearly minimized when $\text{Im}\{\bar{u}\} = \text{Im}\{\bar{d}\}$. The real component is quadratic in $\text{Re}\{\bar{u}\}$, so differentiating with respect to $\text{Re}\{\bar{u}\}$ yields

$$
\frac{df_-(\bar{u})}{d\text{Re}\{\bar{u}\}} = 2\eta (\text{Re}\{\bar{u}\} - \text{Re}\{\bar{d}\})
+ 4|s|(2|s|\text{Re}\{\bar{u}\} - (y + |s|^2)),
$$

which is minimized at

$$
\text{Re}\{\bar{u}\} = \frac{\eta \text{Re}\{\bar{d}\} + 2|s|(y + |s|^2)}{\eta + 4|s|^2}.
$$

This closed form solution yields

$$
u_+ = (\text{Re}\{\bar{u}\} + i\text{Im}\{\bar{u}\})e^{i\zeta s}.
$$

Minimizing $f_+(u)$ along the curve $f_+(u) = f_-(u)$ requires parameterizing the curve. Again, define $\bar{u} = ue^{-i\zeta s}$, $\bar{d} = de^{-i\zeta s}$, and $s = |s|$. Note that $f_-(\bar{u}; s, y) = |\bar{s} - \bar{u}|^2 + (y - |\bar{u}|^2)$, where the latter term equals $B = h_+(\bar{u}; y)$.

Along the curve $f_+(\bar{u}) = f_-(\bar{u})$, $B^2 = (B + |\bar{s} - \bar{u}|^2)^2$, which is true when $s = \bar{u}$, or when $|\bar{s} - \bar{u}|^2 = -2B = 2(|\bar{u}|^2 - y)$. For this second case to yield a nontrivial solution requires $B < 0$, which corresponds to $|\bar{u}|^2 > y$.

Rearranging terms yields our familiar circle $|\bar{u} + \bar{s}|^2 = 2(y + \bar{s}^2)$ from the Laplace distribution case. Plugging our angular parameterization $\bar{u} = c_0 e^{i\theta} - s$, where $c_0 = \sqrt{2(y + \bar{s}^2)}$, into $f_+(\bar{u})$ yields

$$
f_+(\bar{u}(\theta)) = (|c_0 e^{i\theta} - \bar{s}|^2 - y)^2 + \eta |c_0 e^{i\theta} - \bar{s} - \bar{d}|^2
= (c_0^2 - 2c_0 \text{Re}\{e^{i\theta} \bar{s}\} + \bar{s}^2 - y)^2
+ \eta (c_0^2 + |\bar{s} + \bar{d}|^2 - 2c_0 \text{Re}\{e^{i\theta} (\bar{s} + \bar{d})^*\}).
$$
Let \( c_1 = c_0^2 + \bar{s}^2 - y \), and \( c_2 = c_0^2 + |\bar{s} + \bar{d}|^2 \), so
\[
\begin{align*}
f_\pm(\bar{u}(\theta)) &= (c_1 - 2c_0\text{Re}\{e^{i\theta}\bar{s}^*\})^2 \\
&\quad + \eta(c_2 - 2c_0\text{Re}\{e^{i\theta}(\bar{s} + \bar{d})^*\}) \\
&\quad = (2c_0)^2\text{Re}\{e^{i\theta}\bar{s}^*\}^2 \\
&\quad - 2c_0\text{Re}\{e^{i\theta}(2c_1\bar{s} + \eta(\bar{s} + \bar{d}))^*\} + c_1^2 + \eta c_2.
\end{align*}
\]

For convenience, let \( r_1 = 2c_0\bar{s} \), and \( r_2 \) and \( \alpha \) be the magnitude and phase of \( 2c_0(2c_1\bar{s} + \eta(\bar{s} + \bar{d})) \). Differentiating with respect to \( \theta \),
\[
\frac{df_+(\bar{u}(\theta))}{d\theta} = r_2 \sin(\theta - \alpha) - 2r_1^2 \sin \theta \cos \theta.
\]

Setting the derivative equal to zero,
\[
\frac{r_2}{r_1} \sin(\theta - \alpha) = \sin(2\theta).
\]

Defining \( \xi \) such that \( \theta = 2\arctan \xi \), we have \( \sin \theta = \sin(2\arctan \xi) = \frac{2\xi}{1+\xi^2} \), and \( \cos \theta = \cos(2\arctan \xi) = \frac{1-\xi^2}{1+\xi^2} \). Thus,
\[
\begin{align*}
\sin(2\theta) &= \frac{2\xi(1-\xi^2)}{1+\xi^2}, \\
\sin(\theta - \alpha) &= \frac{2\xi \cos \alpha - (1-\xi^2) \sin \alpha}{1+\xi^2}.
\end{align*}
\]

Substituting,
\[
\begin{align*}
0 &= \frac{r_2}{r_1} \left( 2\xi \cos \alpha - (1-\xi^2) \sin \alpha \right) \left( 1 + \xi^2 \right) - 4\xi (1-\xi^2) \\
&= \frac{r_2}{r_1} \left( 2\xi \cos \alpha + 2\xi^3 \cos \alpha - \sin \alpha + \xi \sin \alpha \right) \\
&\quad - 4\xi (1-\xi^2) \\
&= \left( \frac{r_2}{r_1} \sin \alpha \right) \xi^4 + \left( 2\frac{r_2}{r_1} \cos \alpha + 4 \right) \xi^3 \\
&\quad + \left( 2\frac{r_2}{r_1} \cos \alpha - 4 \right) \xi - \frac{r_2}{r_1} \sin \alpha.
\end{align*}
\]

This quartic equation can be solved analytically; the real root that corresponds to \( \theta \) with the minimum \( f_+(\bar{u}(\theta)) \) is used to generate \( u_\pm = (c_0 e^{i\theta} - \bar{s}) e^{i\pm \xi} \), which is valid as long as \( |u_\pm|^2 > y \). Also, one must consider \( \theta = \pm \pi \), which correspond to \( \xi = \pm \infty \), in case either extreme point minimizes \( f_+(\bar{u}(\theta)) \).

### III. Additional Monte Carlo (1D) Simulations

In [1], we ran 50-trial Monte Carlo simulations to characterize the reconstruction quality of the proposed and competing methods. Those simulations employed 128-element signal vectors with sparsities \( K \) ranging from 3 to 8, and sampled noisy squared-magnitude measurements of these signals with 5 outliers and either Gaussian or Laplace noise (both 40 dB SNR). In addition to the median squared error values reported in the paper, we provide mean squared error values (still via PSER, in dB) in Figures 1–2 for the four methods. Note the proposed method still outperforms the competing methods, but the trend as \( K/N \) or \( M/N \) varies appears much less stable. This instability versus the median value is due to outliers in reconstruction quality where the best minimum identified did not correlate with the true signal.

The paper also explores trends in reconstruction quality as a function of the number of outliers and additive noise SNR, demonstrating that the proposed method achieves significantly greater median PSER than competing methods over a wide range of outliers and noise levels. Here, we include similar results for alternate sparsity levels \( K \) demonstrating similar advantages as outliers increase (Figure 3) and as noise levels change (Figure 4). The trends in mean PSER values (not shown) are similar, with the same variability depicted in the trends in mean squared error shown for measurements and sparsity in Figures 1 and 2.

In addition to all these experiments comparing against GESPAR, PR-GAMP, and L1-Fienup, the proposed method is compared against the compressive phase retrieval (CPRL) method in [2], for a length-64 one-dimensional signal. The CPRP implementation from http://users.isy.liu.se/en/tr/ohlsson/code.html uses the standard CVX toolbox from http://cvxr.com/cvx/ with included semidefinite program solver SDPT3. This solver uses > 17 GB of memory.
Fig. 1. The mean PSERs for 50 trials reconstructed using GESPAR, the proposed method, PR-GAMP, and L1-Fienup, for a range of measurement ($M/N$) and sparsity fractions ($K/N$), for measurements with 40 dB SNR Gaussian noise and 5 outliers.

Fig. 2. The mean PSERs for 50 trials reconstructed using GESPAR, the proposed method, PR-GAMP, and L1-Fienup, for a range of measurement ($M/N$) and sparsity fractions ($K/N$), for measurements with 40 dB SNR Laplace noise and 5 outliers.

Fig. 3. The PSER of 50 trials reconstructed using GESPAR, the proposed method, PR-GAMP, and L1-Fienup, for a range of measurement ($M/N$) and outliers, for $K = 5$ and measurements with 40 dB SNR Gaussian noise. The top and bottom rows display results for outliers with ranges $[1, \sqrt{2}]$ and $[1, 2]$, respectively.
for a length-128 signal, necessitating a smaller problem for this experiment. This simulation also uses a different sensing matrix $A$, with a random Gaussian matrix multiplying the DFT. Tailoring the error bound $\epsilon$ to the true error in the measurements, and hand-tuning the best regularization parameter $\lambda$ for CPRL’s 1-norm sparsity term, the CPRL method is run for a range of sparsities $K$ and measurements $M$ corrupted with 40 dB SNR Gaussian noise and $0 - 2$ outliers. The median squared error is compared against the proposed method for the same signals, and the results are shown in Figure 5. The proposed method remains robust in the presence of outliers, while the compressive matrix lifting method does not.

IV. ADDITIONAL IMAGE COMPARISONS (2D)

The image comparisons in [1] demonstrate the superior reconstruction quality of the proposed method when the number of outliers is a sizable fraction (1%) of the measurements. The reconstructions in Figure 6 portray the degradation in image quality of the competing methods, as opposed to the consistent quality of the proposed method, as the number of outliers rises from negligible ($0.001\% = 2$) to more significant ($0.1\% = 132$). Although each set of results corresponds to one trial, the proposed method consistently outputs the star of David phantom, with only nominal gain differences in a few discs. The GESPAR method actually succeeds the best of any competing method, reproducing the star of David shape with only mild attenuation errors when few outliers are present. However, when outliers are more significant, GESPAR’s performance degrades noticeably. The PR-GAMP method tends to fail to reconstruct the phantom, although the image for 0.001% outliers appears to bear a faint resemblance to the phantom. The $L_1$-Fienup method performs reasonably well when outliers are negligible, but it appears less stable than GESPAR or the proposed method as the number of outliers increases. As in the paper, these reconstructions of the $N = 512 \times 512$ star of David image (inspired by related work [3]) are all from $M = N/2 = 131,072$ measurements with outliers and 60 dB SNR additive white Gaussian noise, and the reconstructions all use the same $512 \times 512$ atom dictionary of discs 21 pixels wide as the synthesis transform. The same parameter value $\beta = 0.3$ used for the proposed method with 1% outliers in the paper is used here as well, suggesting that additional parameter tuning is unnecessary over a wide range of outliers. Again, the competing methods are run for at least as many (often many more) initializations and computations as the proposed method for fairness.

REFERENCES


Fig. 5. The median PSERs for 50 trials reconstructed using compressive matrix lifting phase retrieval (CPRL) and the proposed method for a range of measurement ($M/N$) and sparsity fractions ($K/N$), for measurements with 40 dB SNR Gaussian noise and 0 – 2 outliers. Note: signal length is $N = 64$ to avoid CPRL memory issues.
Fig. 6. The reconstruction (with regularization parameter $\beta$ from the experiment in the paper – no additional tuning) for the proposed method is compared against competing methods with the optimal (true) values of $\beta_{sf}$ or $K$. These images are shown for the $512 \times 512$-pixel star of David phantom, from $M = N/2$ measurements, with 60 dB AWGN noise and outliers ranging from 0.001% to 0.1% of the measurements.