

# A generalization for optimized phase retrieval algorithms

Daniel E. Adams,<sup>1\*</sup> Leigh S. Martin,<sup>1</sup> Matthew D. Seaberg, Dennis F. Gardner, Henry C. Kapteyn, and Margaret M. Murnane

*JILA, University of Colorado Boulder, Colorado 80309-0440, USA*

<sup>1</sup>*These authors contributed equally to this work*

[\\*daniel.e.adams@gmail.com](mailto:daniel.e.adams@gmail.com)

**Abstract:** In this work, we demonstrate an improved method for iterative phase retrieval with application to coherent diffractive imaging. By introducing additional operations inside the support term of existing iterated projection algorithms, we demonstrate improved convergence speed, higher success rate and, in some cases, improved reconstruction quality. New algorithms take a particularly simple form with the introduction of a generalized projection-based reflector. Numerical simulations verify that these new algorithms surpass the current standards without adding complexity to the reconstruction process. Thus the introduction of this new class of algorithms offers a new array of methods for efficiently deconvolving intricate data.

© 2012 Optical Society of America

**OCIS codes:** (100.5070) Phase retrieval; (110.1758) Computational imaging; (110.7440) X-ray imaging; (170.0180) Microscopy.

---

## References and links

1. J. Miao and R. Sandberg, "Coherent x-ray diffraction imaging," *J. Sel. Top. Quantum Electron.* **18**, 399–410 (2012).
2. C. La-O-Vorakiat, E. Turgut, C. Teale, H. Kapteyn, M. Murnane, S. Mathias, M. Aeschlimann, C. Schneider, J. Shaw, H. Nembach, and T. Silva, "Ultrafast demagnetization measurements using extreme ultraviolet light: comparison of electronic and magnetic contributions," *Phys. Rev. X* **2**, 1–7 (2012).
3. S. Mathias, C. La-O-Vorakiat, P. Grychtol, P. Granitzka, E. Turgut, J. M. Shaw, R. Adam, H. T. Nembach, M. E. Siemens, S. Eich, C. M. Schneider, T. J. Silva, M. Aeschlimann, M. M. Murnane, and H. C. Kapteyn, "Probing the timescale of the exchange interaction in a ferromagnetic alloy," *Proc. Nat. Acad. Sci. USA* pp. 1–6 (2012).
4. J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature* **400**, 342–344 (1999).
5. C. a. Larabell and K. a. Nugent, "Imaging cellular architecture with x-rays." *Curr. Opin. Struct. Biol.* **20**, 623–631 (2010).
6. H. Jiang, C. Song, C. Chen, R. Xu, K. Raines, B. Fahimian, C.-h. Lu, T. Lee, A. Nakashima, J. Urano, and Others, "Quantitative 3D imaging of whole, unstained cells by using x-ray diffraction microscopy," *Proc. Nat. Acad. Sci. USA* **107**, 11234 (2010).
7. K. Giewekemeyer, P. Thibault, S. Kalbfleisch, A. Beerlink, C. Kewish, M. Dierolf, F. Pfeiffer, and T. Salditt, "Quantitative biological imaging by ptychographic x-ray diffraction microscopy," *Proc. Nat. Acad. Sci. USA* **107**, 529 (2010).
8. I. Robinson and R. Harder, "Coherent x-ray diffraction imaging of strain at the nanoscale," *Nature Mater.* **8**, 291–298 (2009).
9. M. D. Seaberg, D. E. Adams, E. L. Townsend, D. A. Raymondson, W. F. Schlotter, Y. Liu, C. S. Menoni, L. Rong, C.-C. Chen, J. Miao, H. C. Kapteyn, and M. M. Murnane, "Ultrahigh 22 nm resolution coherent diffractive imaging using a desktop 13 nm high harmonic source." *Opt. Express* **19**, 22470–22479 (2011).
10. D. J. Brady, M. E. Gehm, R. a. Stack, D. L. Marks, D. S. Kittle, D. R. Golish, E. M. Vera, and S. D. Feller, "Multiscale gigapixel photography." *Nature* **486**, 386–389 (2012).

11. S. Marchesini, "Invited article: a [corrected] unified evaluation of iterative projection algorithms for phase retrieval." *Rev. Sci. Instrum.* **78**, 011301 (2007).
12. S. Marchesini, "Phase retrieval and saddle-point optimization." *J. Opt. Soc. Am. A* **24**, 3289–3296 (2007).
13. D. Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse Probl.* **21**, 37 (2005).
14. J. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Opt. Lett.* **3**, 27–29 (1978).
15. J. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
16. R. Gerchberg and W. O. Saxton, "Phase determination from Image and diffraction plane pictures in the electron microscope," *Optik* **34**, 275–283 (1971).
17. R. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35** (1972).
18. H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization," *J. Opt. Soc. Am. A* vol. **19**, 1334–1345 (2002).
19. V. Elser, "Phase retrieval by iterated projections," *J. Opt. Soc. Am. A* **20**, 40–55 (2003).
20. H. H. Bauschke, P. L. Combettes, and D. Luke, "Finding best approximation pairs relative to two closed convex sets in Hilbert spaces," *J. Approx. Theory* **127**, 178–192 (2004).
21. A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," *J. Opt. Soc. Am. A* **1**, 932 (1984).
22. K. Goebel, *Topics in metric fixed point theory* (Cambridge University Press, 1990).
23. Y. Censor, and S. A. Zenios, *Parallel Optimization: Theory, Algorithms, and Applications*. (Oxford University Press: 1997).
24. D. F. Gardner, B. Zhang, M. D. Seaberg, L. S. Martin, D. E. Adams, F. Salmassi, E. Gullikson, H. Kapteyn, and M. Murnane, "High numerical aperture reflection mode coherent diffraction microscopy using off-axis apertured illumination," *Opt. Express* **20**, 19050 (2012).
25. A. D'Alfonso, A. Morgan, A. Martin, H. Quiney, and L. Allen, "Fast deterministic approach to exit-wave reconstruction," *Phys. Rev. A* **85**, 1–9 (2012).
26. D. Luke, "Local linear convergence of approximate projections onto regularized sets," *Nonlinear Analysis: Theory, Methods & Applications* (2011).

## 1. Introduction

The last decade has seen significant advances in x-ray sources. Third generation synchrotrons, fourth generation x-ray free electron lasers (XFELs) and high harmonic generation (HHG) sources are currently paving the way for powerful new characterization techniques with unparalleled temporal and spatial resolution [1–3]. EUV light and higher energy soft x-rays are ideal probes for nano- and atomic-scale systems due to their inherent elemental contrast and ability to penetrate thick samples using novel imaging techniques. The last decade has also seen great progress in coherent diffraction imaging (CDI) as an alternative to more traditional zone-plate or mirror based imaging techniques [4]. In CDI, a plane wave illuminates an isolated specimen (i.e. the entire finite sample is illuminated by the wave) and the diffracted light is detected in the far-field or so-called Fraunhofer zone. In this case the diffracted light represents the scaled, complex Fourier transform of the scattering electronic potential. In practice however, only the diffracted intensity can be measured, resulting in the loss of phase information. The goal of iterated phasing algorithms is to solve for the missing phases, ultimately revealing details about the diffracting electronic potential. Recent work using CDI has explored phase contrast inspection of biological systems [5–7] strain fields in nano-particles [8] and demonstrated that high numerical aperture imaging with near-wavelength resolution is possible [9]. These diverse applications of CDI show great promise for widespread biological and industrial imaging. Moreover, the necessity for computational phase retrieval casts CDI into a broader class of inverse problems, which are interesting to the mathematical community as well.

However, a drawback of CDI is that computation time for image reconstruction is a substantial bottleneck, especially considering that most common imaging modalities offer direct imaging or involve non-iterative post processing. Some illustrative examples include applications in the semiconductor industry, which will require the imaging of massive areas, and the emerging use of use of gigapixel detectors [10]. Numerous algorithms have been proposed with the intent of accelerating the reconstruction process [11]. Some succeed in reducing the number

of iterations necessary, but occasionally at the expense of computational simplicity, resulting in a slower algorithm [12]. Keeping in mind the importance of convergence speed, we use physical intuition in this paper to build on the concept of iterated, generalized projections that quickly push an initial random guess toward convergence without introducing additional steps. We provide an outline for deriving augmented fixed point equations and hence new iterative phasing algorithms, which we call generalized interior feedback (GIF). These algorithms have application in CDI, and are motivated by the convex analysis employed by Luke [13] and the linear systems approach used by Fienup [14, 15]. We also show through extensive numerical studies that GIF algorithms consistently outperform unmodified algorithms in terms of convergence rate, and even offer improved reconstruction quality in some cases.

## 2. Fixed point equations and generalized interior feedback

In CDI, the phase retrieval problem was originally solved using iterated projections onto constraint subspaces, motivated by the pioneering work of Gerchberg and Saxton [16, 17]. The diffraction amplitudes  $m$  serve as one constraint and if some element  $u \in \mathcal{L} : \mathbb{Z}^N \rightarrow \mathbb{R}$  satisfies this constraint then we say  $u \in M \equiv \{v \in \mathcal{L} : |\mathcal{F}v| = m\}$ , where  $\mathcal{F}$  is the discrete Fourier transform and  $\mathcal{L}$  is the Hilbert space of square integrable functions. Upon discretely sampling the scattered light at a rate higher than the Nyquist frequency, we may assume that the electronic potential is isolated. The latter condition indicates that the electronic potential  $u$  (or in some cases the exit surface wave leaving the specimen) is supported on some set  $D$  i.e.  $\text{supp}(u) \subset D \subset \mathbb{Z}^N$ . Further, we write  $CD$  as the compliment of  $D$  and  $1_{CD}$  which takes on a value of 0 on  $D$  and 1 elsewhere so that  $S = \{u \in \mathcal{L} : u \cdot 1_{CD} = 0\}$ , where we have defined  $\mathbb{Z}^N$  as the sample domain,  $\mathbb{R}$  is the reals and  $\text{supp}(u)$  is the support of  $u$ . If we also require that  $u$  is non-negative i.e.  $u \rightarrow u_*$ , then we require that  $\mathbb{R} \rightarrow \mathbb{R}_+$ , which leads to the definition  $S_+ = \{u_* \in \mathcal{L} : u_*(x) \geq 0 \forall x \in \mathbb{Z}^N\}$ .

Even with these constraints, the phase retrieval problem in CDI was intractable until Fienup introduced the hybrid input-output (HIO) algorithm [15], which was sufficiently fast to enable real-world applications. Historically, Bauschke showed a relationship between Dykstra's algorithm and the Basic Input-Output of Fienup and Fienup's Hybrid Input-Output algorithm and the Douglas-Rachford algorithm [18]. Elser showed that the HIO is a specific instance of the difference map, a general class of iterative projection algorithms. These algorithms gain their speed-up over previous methods by searching off of the constraint subspaces. Specifically, Bauschke showed that the difference map moves past near intersections of the subspaces in search of global minima [19]. Although this helps prevent stagnation near local minima, it means that the HIO may diverge from the global minimum if the constraint subspaces do not intersect. This behavior often becomes prominent in phase retrieval when noise is present in the diffraction pattern.

The relaxed averaged alternating reflections (RAAR) algorithm [13] also demonstrates rapid convergence and avoidance of local minima, making it practical for phase retrieval. In contrast, however, it can be shown that RAAR has fixed points that are related to points in one constraint subspace nearest the other [20]. This means that RAAR searches for the best approximate solution to the phase problem when an exact solution does not exist. This behavior is highly advantageous when no exact solution to the phase retrieval problem exists. Whereas HIO (and when the non-negativity constraint is correctly implemented, the Hybrid Projection Reflection (HPR) algorithm [20]) only have fixed points near intersections, RAAR explores near approaches in search of a solution that optimizes the amplitude and support constraints.

### 2.1. Mathematical preliminaries

We present the problem of phase retrieval as a general feasibility problem, that is:

$$\text{find } u \in S \cap M \quad (1)$$

Phase retrieval algorithms proceed by iteratively applying a projection-based operator to a random initial guess  $u$ , which ultimately drives the current iterate toward a fixed point. We define the projection  $P_M(u)$ ,  $u \in \mathcal{L}$  onto the set  $M$ :

$$P_M(u) = \mathcal{F}^{-1}\hat{v} \quad \hat{v}(\xi) = \begin{cases} m(\xi) \frac{\mathcal{F}u(\xi)}{|\mathcal{F}u(\xi)|} & \text{if } \mathcal{F}u(\xi) \neq 0 \\ m(\xi) & \text{otherwise,} \end{cases} \quad (2)$$

$\hat{v}$  is one of the multivalued Fourier domain projections and  $m$  is the measured Fourier modulus. Additionally we define the projection onto the set  $D \subset \mathbb{Z}^N$  with  $D$  being compact on  $\mathbb{R}^N$   $P_S(u)$ ,  $u \in \mathcal{L}$  as:

$$(\forall x \in \mathbb{Z}^N) \quad (P_S(u))(x) = \begin{cases} u(x) & \text{if } x \in D \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

In the case where we insist the function  $u \rightarrow u_* \in \mathbb{R}_+$  and  $S \rightarrow S_+$ , the conditional statement is modified to read  $\max\{0, u(x)\}$  if  $x \in D$  (the otherwise case is left unchanged).

As the name suggests, RAAR is based on reflection operators, which are defined  $R_C = 2P_C - I$ . We now introduce the generalized reflector of an arbitrary projector,  $P_C$  as:

$$R_C^\gamma \equiv (1 + \gamma)P_C - \gamma I \quad (4)$$

It is worth noting some properties of generalized reflector  $R_C^\gamma$  that are summarized in Fig. 1. For values of  $\gamma < 1$ , the generalized reflector is the relaxation parameter of Levi and Stark [21],

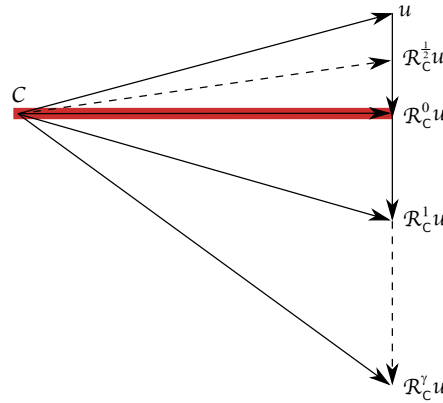


Fig. 1. Generalized reflector, where  $\mathcal{C}$  is the set around which the guess  $u$  is being reflected. For  $\gamma < 1$  the reflector is a relaxation parameter,  $\gamma = 1$  recovers the standard reflector and  $\gamma > 1$  chooses the depth of reflection.

which appears in many iterative phase retrieval algorithms including the difference map.  $\gamma = 0$  is simply a projection,  $\gamma = 1$  produces the standard reflection  $R_C$  as in [22, 23] while  $\gamma > 1$  essentially chooses the reflection depth.

## 2.2. Generalized interior feedback

We now state the RAAR algorithm without further justification and develop the concept of generalized interior feedback (GIF). RAAR can be written as a fixed point equation of the form  $u_{n+1} = \mathcal{V}(\mathcal{T}, \beta_n) u_n$ , where, in general,  $\mathcal{T}$  is a generic operator that can include any type of averaging of reflectors and projectors.  $\mathcal{V}(\mathcal{T}, \beta_n)$  is an operator that includes some kind of relaxation strategy through the use of the relaxation parameter  $\beta_n$ , which is modified at the  $n^{\text{th}}$  iteration. Using this notation we see that the RAAR algorithm, usually written as the conditional statement,

$$(\forall x \in \mathbb{Z}^N) \quad u_{n+1}(x) = \begin{cases} (P_M(u_n))(x) & \text{if } x \in D \\ \beta_n u_n + (1 - 2\beta_n)(P_M(u_n))(x) & \text{otherwise,} \end{cases} \quad (5)$$

may be written as a fixed point algorithm

$$u_{n+1} = \left(\frac{1}{2}\beta_n(R_S R_M + I) + (1 - \beta_n)P_M\right)(u_n), \quad (6)$$

where  $\mathcal{T} \equiv \frac{1}{2}(R_S R_M + I)$  and the algorithm is relaxed by the term  $(1 - \beta_n)P_M$ . Upon expanding the support reflector  $R_S$ , Eq. (6) becomes,

$$u_{n+1} = \left(\frac{1}{2}\beta_n(2P_S(R_M) - R_M + I) - (1 + \beta_n)P_M\right)(u_n). \quad (7)$$

Fienup and others related iterative phase retrieval to control theory and motivated several developments with the concept of feedback. Most algorithms only include feedback terms outside the support, which push object-domain amplitudes to zero. Here, we consider the addition of a general feedback term of the form  $\alpha \mathcal{E}(P_M, I)$  inside the support. In the case where the feasibility problem is consistent, we require the feedback term to vanish as the algorithm approaches a fixed point. This leads us to a specific form of the feedback as  $\alpha \mathcal{E}(I - P_M)$ . Inserting this form in the interior of the support in Eq. (7)

$$u_{n+1} = \left(\frac{1}{2}\beta_n(2P_S(R_M - \alpha \mathcal{E}(I - P_M)) - R_M + I) - (1 + \beta_n)P_M\right)(u_n), \quad (8)$$

and, for example, if we choose the especially simple form for  $\mathcal{E}(I - P_M) \equiv I - P_M$  and collect terms, we arrive at

$$u_{n+1} = \frac{1}{2}\beta_n(R_S R_M^\gamma + I)u_n + \frac{1}{2}\beta_n(R_M^\gamma - R_M)u_n + (1 - \beta_n P_M)u_n, \quad (9)$$

where we have defined  $\gamma \equiv \alpha + 1$ . It is now instructive to write Eq. (9) in the conditional form preferred by the optics community. For any arbitrary, real signal  $z$  we can write the positive and negative parts as  $z^+ = \max\{0, z\}$  and  $z^- = z - z^+$ . Returning to the more general form of GIF, setting the scaling factor  $\alpha$  to unity and enforcing positivity on  $u \rightarrow u_* \in \mathbb{R}_+$  Eq. (8) is,

$$u_{n+1} = (1_D \cdot [P_M - \beta \mathcal{E}(I - P_M)] + 1_{CD} \cdot [\beta I + (I - 2\beta)P_M] + 1_D \cdot [R_M - \mathcal{E}(I - P_M)]^-(u_n). \quad (10)$$

Then there are three cases to consider. If  $x \in D$  &  $R_M \geq \mathcal{E}(I - P_M)$ , then the right hand side of Eq. (10) reduces to  $P_M - \beta \mathcal{E}(I - P_M)$ . If  $x \in D$  &  $R_M \leq \mathcal{E}(I - P_M)$  then this term becomes

$\beta I + (I - 2\beta)P_M$ . Finally, if  $x \notin D$  then it reads  $\beta I + (I - 2\beta)P_M$ . Phrased in a conditional statement Eq. (10) is  $(\forall x \in \mathbb{Z}^N)$ ,

$$u_{n+1}(x) = \begin{cases} ((P_M(u_n) - \beta \mathcal{E}(I - P_M)u_n)(x) & \text{if } x \in D \text{ and} \\ (R_M - \mathcal{E}(I - P_M))u_n(x) \geq 0 & \\ \beta_n u_n(x) + (1 - 2\beta_n)(P_M(u_n))(x) & \text{otherwise.} \end{cases} \quad (11)$$

If we again choose a simple form for  $\mathcal{E}(I - P_M) = I - P_M$ ,  $(\forall x \in \mathbb{Z}^N)$  we arrive at,

$$u_{n+1}(x) = \begin{cases} (R_M^{\beta_n}(u_n))(x) & \text{if } x \in D \text{ and} \\ (R_M - I + P_M)u_n(x) \geq 0 & \\ \beta_n u_n(x) + (1 - 2\beta_n)(P_M(u_n))(x) & \text{otherwise.} \end{cases} \quad (12)$$

### 3. Numerical results

Heuristically, appendix A demonstrates which elements GIF-RAAR converges to if it converges, and provides insight concerning the behavior of GIF-RAAR during phase retrieval. However, the analysis found in appendix A does not bear on issues such as convergence rate or behavior in a non-convex setting. To study these aspects, we explore the algorithm's properties in the context of coherent diffraction imaging. In an ideal setting, the sets  $M$  and  $S+$  intersect, and thus the fixed point of GIF-RAAR is the set of elements that exactly satisfy both constraints. Real data sets contain noise, which eliminates exact intersections and dramatically changes the nature of the problem.

We consider both noiseless and noisy data. In the latter case, white Gaussian noise is added to the square of the Fourier amplitudes of the object to simulate detector read-out noise and background light, which are usually the dominant contributions. Results on experimental data are also given at the end of this section. Figure 2 shows the test image along with the support. These algorithms may be used both with and without a non-negativity constraint. For objects that contain a phase angle greater than  $\pi/2$ , the typical non-negativity constraint cannot be used. Simulations consider both cases. When not using non-negativity, a phase profile shown in Fig. 2 is added to the image and a tighter support is used to compensate for the missing constraint.

Table 1 shows results from a series of reconstructions without non-negativity. Each case consists of one of four different signal to noise ratios and one of three algorithms. Each case is comprised of 100 individual runs starting with a random phase guess and then run for 3000 iterations. The translation of each reconstruction is determined by measuring the root mean square difference between the object and the reconstruction during a pixel-by-pixel raster scan. The real-space error is then the lowest error value given by the scan. A cutoff of 5% reciprocal-space error defines the convergence condition. Runs that do not reach this cutoff are considered unsuccessful and are not included in calculations of the convergence iteration or real-space error.

In this setting, HIO surpasses RAAR and GIF-RAAR except in the presence of a large amount of noise. However, it is important to realize that we have only considered the case of  $\beta = 0.9$ , and that each algorithm behaves differently with respect to this parameter. We make these comparisons because this choice of  $\beta$  is most common within the optics community regardless of algorithm, but note that the most meaningful contrasts are between GIF algorithms and their standard counterparts. In general, our research has consistently shown that RAAR outperforms HPR (HIO with the correct non-negativity constraint) when non-negativity can be used, but that HIO outperforms RAAR without this constraint. This important trend is reflected in the following data.

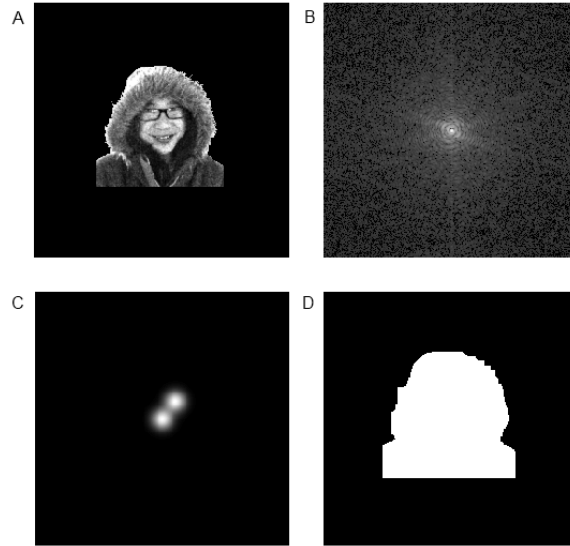


Fig. 2. (a) The object used for reconstruction. (b) Simulated diffraction pattern in the presence of noise raised to the quarter power for visibility. (c) The phase profile added to reconstructions when the non-negativity constraint was omitted. (d) Support used for reconstruction when the non-negativity constraint could be used. A similar but tighter support (not shown) was used in simulations without non-negativity.

Notice in table 1 that although RAAR converges more quickly than GIF-RAAR, its real-space error appears not to be noise limited, but is fixed at 20% regardless of the noise level. On the contrary, GIF-RAAR consistently reaches a significantly lower real-space error, which shows the expected correlation with noise level. GIF-RAAR also achieves a much higher success rate, near 100% for all cases. We found that GIF-HIO performed slightly better for  $\alpha = 0.9$ , and thus used this value in reconstructions. GIF-HIO converges more quickly than HIO, but with a consistently larger real-space error. With a signal-to-noise ratio of 11 (the smallest used), GIF-HIO fails to converge after 3000 iterations.

RAAR was originally developed with the non-negativity constraint implemented. Table 2 shows improvement resulting from its use. GIF-RAAR surpasses unmodified RAAR in all categories. Once again, RAAR has a lower but still useful success rate. GIF-HPR also surpasses its unmodified counterpart for higher signal-to-noise ratios, though in the presence of large quantities of noise its convergence rate drops substantially. In this simulation, GIF algorithms surpass their standard counterparts' real-space error and convergence rate in all but this case. The most dramatic improvement is the order of magnitude lower real-space error achieved by GIF-RAAR over RAAR after the allotted number of iterations.

Reconstruction quality depends strongly on many details of the problem in intricate and subtle ways. To ensure that the results above carry over to real-world applications, we present

Table 1. Reconstruction quality, convergence speed and success rate without the use of the non-negativity constraint.

Measure	SNR	HIO $\alpha = 0$	HIO $\alpha = 0.9$	RAAR $\alpha = 0$	RAAR $\alpha = 1$
Convergence Iteration	$\infty$	$76 \pm 4$	$49 \pm 3$	$210 \pm 50$	$230 \pm 60$
	400	$74 \pm 4$	$49 \pm 2$	$150 \pm 30$	$230 \pm 60$
	81	$73 \pm 4$	$50 \pm 2$	$170 \pm 40$	$170 \pm 40$
	11	$330 \pm 20$	$> 3000$	$180 \pm 40$	$280 \pm 50$
R-Space Error	$\infty$	$(24 \pm 6) \times 10^{-8}$	$(4 \pm 2) \times 10^{-7}$	$0.20 \pm 0.01$	$0.075 \pm 0.008$
	400	$0.0091 \pm 0.0001$	$0.0135 \pm 0.0001$	$0.20 \pm 0.01$	$0.094 \pm 0.009$
	81	$0.0247 \pm 0.0003$	$0.0327 \pm 0.0002$	$0.20 \pm 0.01$	$0.092 \pm 0.009$
	11	$0.138 \pm 0.002$		$0.20 \pm 0.01$	$0.127 \pm 0.007$
Success Rate	$\infty$	100%	100%	67%	100%
	400	100%	100%	62%	100%
	81	100%	100%	62%	99%
	11	100%	0%	52%	97%

Table 2. Reconstruction quality, convergence speed and success rate with the use of the non-negativity constraint.

Measure	SNR	HPR $\alpha = 0$	HPR $\alpha = 0.9$	RAAR $\alpha = 0$	RAAR $\alpha = 1$
Convergence Iteration	$\infty$	$66 \pm 2$	$47 \pm 1$	$90 \pm 10$	$61 \pm 5$
	400	$67 \pm 2$	$46.9 \pm 0.9$	$100 \pm 30$	$70 \pm 8$
	81	$66 \pm 2$	$47.4 \pm 0.8$	$88 \pm 8$	$80 \pm 11$
	11	$180 \pm 20$	$1500 \pm 800$	$180 \pm 30$	$77 \pm 6$
R-Space Error	$\infty$	$0.0521 \pm 0.0008$	$0.0372 \pm 0.0005$	$0.076 \pm 0.007$	$0.007 \pm 0.002$
	400	$0.0633 \pm 0.0006$	$0.0489 \pm 0.0005$	$0.091 \pm 0.009$	$0.009 \pm 0.002$
	81	$0.0821 \pm 0.0007$	$0.0660 \pm 0.0005$	$0.10 \pm 0.01$	$0.013 \pm 0.002$
	11	$0.158 \pm 0.002$	$0.175 \pm 0.009$	$0.111 \pm 0.006$	$0.074 \pm 0.006$
Success Rate	$\infty$	100%	100%	77%	100%
	400	100%	100%	76%	100%
	81	100%	100%	77%	100%
	11	100%	3%	76%	100%

reconstructions on a set of experimental data. Figure 3 shows a representative diffraction pattern and reconstruction using the GIF-RAAR algorithm. The object was a negative 1951 USAF test target imaged using CDI, an optical laser ( $\lambda = 633nm$ ) and the apertured illumination technique [24]. A significant amount of readout noise is present in the data. Panel (a) of Fig. 3 is a diffraction pattern collected using the apertured illumination technique. Panel (b) shows a representative GIF-RAAR reconstruction, averaging 1000 iterations after the algorithm converged. Also shown in panel (b) is a dashed line indicating the border of the support used to reconstruct the pattern and a photograph of the test pattern for comparison.

Panel (c) in Fig. 3 shows the relative error for both GIF-RAAR and RAAR highlighting the plateau common in RAAR reconstructions. This particular comparison is motivated by the consistent improvement GIF seems to offer to RAAR, based on the data in table 2. As in most cases, even a modest amount of averaging after the algorithm has converged significantly improves reconstructed image quality. The overall reconstruction quality and relative reconstruction quality between GIF-RAAR and RAAR depend heavily on the conditions of the problem. For example, changing the support size dramatically changes the convergence rates. In this test, GIF-RAAR



always converges faster than RAAR, at times by 500 iterations and under other conditions only 50. In some cases, RAAR achieves a slightly lower k-space error than GIF-RAAR. However, the visual quality of the resulting reconstructions is identical.

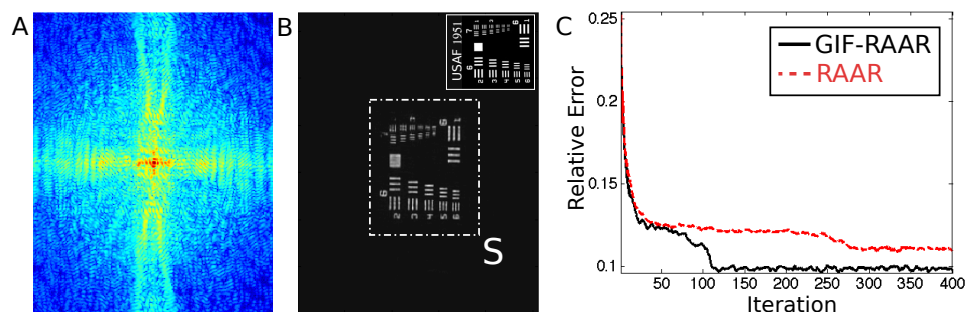


Fig. 3. (a) Diffraction pattern data raised to the quarter power for visibility. (b) Reconstructed object with dashed line showing the boundary of the support,  $S$  used for both GIF-RAAR and RAAR. The inset shows a traditional microscope image of the test pattern for comparison (Thorlabs R3L3S1N - Negative 1951 USAF Test Target, 3'' x 3''). (c) Relative error as a function of iteration number for both GIF-RAAR and RAAR respectively.

#### 4. Conclusion

Phase retrieval has proven to be a subtle but very important problem both physically and mathematically. Although practical implementations of phase retrieval have been available since the 1970's, development of optimal algorithms continues to this day [13, 25]. Non-convex optimization problems remain an advanced field of study within pure mathematics, and while some algorithms have been well-characterized [26], the exact convergence properties of most algorithms remain elusive. In a practical setting, phase retrieval of experimental data often relies on the use of a combination of methods and empirical observation which are used to determine optimal parameter values. The introduction of GIF-RAAR provides a tool which, in all of our numerical simulations, outperforms the most commonly used methods while maintaining the computational speed of iterative fast Fourier transform fixed point algorithms. Furthermore, our simulations demonstrate the advantage of introducing feedback within the object domain in general, suggesting that similar benefit may be gained from incorporation into other methods. Only a relatively simple internal feedback mechanism has been analyzed here, and other types of terms may exhibit advantageous properties. Finally, we suggest that this internal feedback mechanism may provide a conduit relating RAAR type algorithms to a more generalized difference map.

#### A. Appendix

By design, Eq. (12) strongly resembles RAAR. However, the internal feedback mechanism modifies the update rule inside the support to be the generalized reflector instead of a projection. The update rule for the so-called GIF-RAAR algorithm depends on the pointwise sign of the reflector as in RAAR but additionally depends on the pointwise sign of the feedback term. The question then arises; what, if anything, is fundamentally different about the new GIF-RAAR algorithm? In the present section we examine the fixed points of the GIF-RAAR algorithm and find them to be the same as those of RAAR in the convex case. Although a convex analysis only

provides heuristics for the phase retrieval problem, which is non-convex, such analyses are in general much simpler and have predicted the important properties of RAAR [13]. Non-convex behavior of RAAR has been analyzed in reference [26].

The following proof is an extension of that given in ref. [20]. Let  $A$  and  $B$  be two closed convex subsets of the Hilbert space  $\mathcal{L}$  of square integrable functions, and replace to  $S_+$  and  $M$  respectively.  $P_A$  and  $P_B$  are projections onto these sets.  $\mathcal{L}$  has an inner product  $\langle \cdot, \cdot \rangle$  and norm  $\|\cdot\|$ . Analogous to RAAR, we define the operator  $T_*^\gamma$  as

$$T_*^\gamma \equiv \frac{1}{2}(R_A R_B^\gamma + I) + \frac{1}{2}(R_B^\gamma - R_B)$$

or in simplified form,

$$T_*^\gamma = P_A R_A^\gamma - P_B + I, \quad (13)$$

GIF-RAAR is then the  $\beta$ -weighted average between  $T_*^\gamma$  and  $P_B$ :

$$\mathcal{V}(T_*^\gamma, \beta) \equiv \beta T_*^\gamma + (1 - \beta)P_B \quad (14)$$

Let  $F$  be the set of points in  $B$  nearest  $A$  and let  $E$  be the set of points in  $A$  nearest  $B$  as shown in Fig. 4. We now wish to show that the fixed points of GIF-RAAR correspond to those of RAAR, which are given by:

$$u \in F - \frac{\beta}{1 - \beta}g \quad (15)$$

or

$$u = P_B u - \frac{\beta}{1 - \beta}g \quad (16)$$

here,  $g$  is the gap vector, which is the shortest vector between  $A$  and  $B$ .  $g$  is rigorously defined by considering the set  $\mathcal{D} = \overline{B - A}$ , which is the closure of the set of all vectors starting in  $A$  and ending in  $B$ .  $g$  may be written as  $g = P_{\mathcal{D}}(0)$ . Notice that even when the problem is ill posed so that  $A$  and  $B$  have no intersection, RAAR still has well defined fixed points which are related to the points in one set nearest the other. This feature makes RAAR advantageous over other algorithms.

With the definition  $f = P_B u$ , recall that  $P_A(2f - u) = P_A(2P_B u - u) = P_A R_B u$ . To show the fixed points of GIF-RAAR, begin by rewriting Eq. (13) as

$$(\forall u \in \mathcal{L}) \quad u - T_*^\gamma u = -P_A R_B^\gamma u + f, \quad (17)$$

and for our choice of  $u$

$$\beta T_*^\gamma u + (1 - \beta)P_B u = u, \quad (18)$$

which may be rewritten as

$$u - T_*^\gamma u = \frac{(1 - \beta)}{\beta}(P_B u - u) = -\frac{1 - \beta}{\beta}y. \quad (19)$$

where we have defined  $y = P_B u - u$ . Using Eq. (17) we obtain

$$P_A R_B^\gamma u = f + \frac{1 - \beta}{\beta}y. \quad (20)$$

Because  $A$  is nonempty, closed and convex,  $\forall a \in A \quad \langle a - P_A x, x - P_A x \rangle \leq 0$  for any  $x$ . Letting

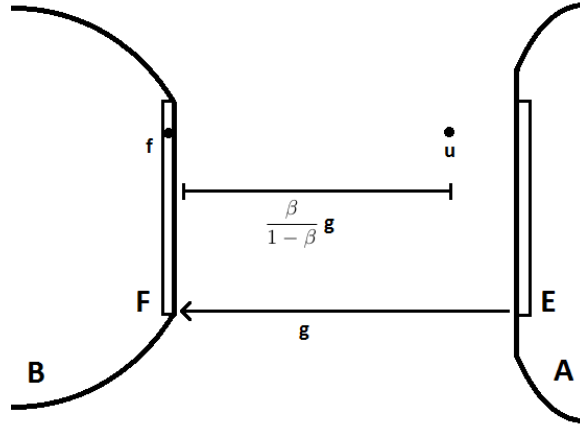


Fig. 4. Definition of relevant quantities for convex analysis, considering two sets in  $\mathbb{R}^2$  for visualization. A and B are convex sets in  $\mathcal{L}$ , E and F are the points in A and B closest to B and A, respectively. The gap vector  $g$  is effectively the shortest distance between A and B.

$x = R_B^\gamma u$ , we can write

$$(\forall a \in A) \quad \langle a - P_A R_B^\gamma u, R_B^\gamma u - P_A R_B^\gamma u \rangle \leq 0, \quad (21)$$

and using the previous results

$$\left\langle (a - f) - \frac{1 - \beta}{\beta} y, -(\gamma + \frac{1 - \beta}{\beta}) y \right\rangle \leq 0. \quad (22)$$

Rearranging gives the result

$$\langle -a + f, y \rangle + \frac{1 - \beta}{\beta} \|y\|^2 \leq 0. \quad (23)$$

On the other hand, since B is also nonempty, closed and convex, with  $f = P_B u$  we have  $\forall b \in B$ ,

$$\langle b - f, y \rangle \leq 0 \quad (24)$$

and upon combining results

$$\langle b - f, y \rangle + \langle -a + f, y \rangle + \frac{1 - \beta}{\beta} \|y\|^2 \leq 0, \quad (25)$$

simplifying and noting that  $\beta$  is positive definite and so is  $\|y\|^2$  we have

$$\langle b - a, y \rangle \leq -\frac{1 - \beta}{\beta} \|y\|^2 \leq 0. \quad (26)$$

Now, taking a sequence  $a_0, a_1, a_2, \dots$  in A and a sequence  $b_0, b_1, b_2, \dots$  in B such that  $g_n = b_n - a_n \rightarrow g$  then we have

$$(\forall n \in \mathbb{N}) \quad \langle g_n, y \rangle \leq -\frac{1 - \beta}{\beta} \|y\|^2. \quad (27)$$

Taking the limit and using the Cauchy-Schwartz inequality yields

$$\|y\| \leq \frac{\beta}{1-\beta} \|g\|. \quad (28)$$

Conversely  $u = (\beta T_*^\gamma u + (1-\beta)P_B u) = 0$ , is also  $y = \beta(P_A R_B^\gamma u - f) + \beta y$ , and solving for  $y$ ,

$$y = \frac{\beta}{1-\beta} (P_A R_B^\gamma u - f). \quad (29)$$

Taking the  $l_2$  norm gives

$$\|y\| = \frac{\beta}{1-\beta} \|P_A R_B^\gamma u - f\|. \quad (30)$$

Using the previous result

$$\|y\| = \frac{\beta}{1-\beta} \|P_A R_B^\gamma u - f\| \geq \frac{\beta}{1-\beta} \|g\|$$

hence

$$\|y\| = \frac{\beta}{1-\beta} \|g\|, \quad (31)$$

and taking the limit gives  $y = -\frac{\beta}{1-\beta} g$ ,  $\implies u = f - \frac{\beta}{1-\beta} g$ , where we have appropriately chosen this  $u \in \text{Fix}\{\mathcal{V}(T_*^\gamma, \beta)\}$ . We may now write

$$u = F - \frac{\beta}{1-\beta} g \quad (32)$$

To show the equality  $\text{Fix } \mathcal{V}(T_*^\gamma, \beta) = F - \frac{\beta}{1-\beta} g$ , we show mutual containment of the sets  $F - \frac{\beta}{1-\beta} g$  and  $\text{Fix } \mathcal{V}(T_*^\gamma, \beta)$ . The following proof is a close adaptation of that given in reference [20], and shows that if  $u$  is of the form given in Eq. (16), then it is fixed by GIF-RAAR, so that  $\mathcal{V}(T_*^\gamma, \beta)u = u$ .

We use the concept of a cone, which is similar to a linear vector space, but is instead closed only under multiplication by positive scalars. We define the normal cone map of a convex set  $A$  as

$$N_A : x \mapsto \begin{cases} \{u : \forall a \in A, \langle a - x, u \rangle \leq 0\} & x \in A \\ \emptyset & \text{otherwise} \end{cases} \quad (33)$$

This may be thought of as the set of elements that, when added to some  $x \in A$  are projected back to  $x$  by  $P_A$ , i.e.  $x = P_A(x+u) \Leftrightarrow u \in N_A(x)$ .

Let  $u = f - \frac{\beta}{1-\beta} g$  and  $e = f - g$ , where  $f = P_B u$ . As  $g = P_{\mathcal{D}}(0) = P_{\mathcal{D}}(g - g)$ , it follows that  $-g \in N_{\mathcal{D}}(g)$ . From the definitions of  $E$ ,  $F$  and  $\mathcal{D}$ , it follows that  $N_{\mathcal{D}}(g) = N_B(f) \cap (-N_A(e))$ . By closure of cones, for  $0 \leq \beta < 1$ ,  $-\frac{\beta}{1-\beta} g \in N_{\mathcal{D}}(g) \subset N_B(f)$ . It follows that

$$\begin{aligned} P_B u &= P_B \left( f - \frac{\beta}{1-\beta} g \right) = f \\ R_B^\gamma u &= f + \frac{\gamma\beta}{1-\beta} g \end{aligned} \quad (34)$$

As  $g \in N_A(e)$  and  $f = e + g$ , it follows that

$$\begin{aligned}
 P_A R_B^\gamma u &= P_A \left( f + \frac{\gamma\beta}{1-\beta} g \right) \\
 &= P_A \left( e + \left( 1 + \frac{\gamma\beta}{1-\beta} \right) g \right) \\
 &= e
 \end{aligned} \tag{35}$$

So long as  $\left( 1 + \frac{\gamma\beta}{1-\beta} \right) \geq 0$ . Acting  $\mathcal{V}(T_*^\gamma, \beta)$  on  $u$ , we find

$$\begin{aligned}
 \mathcal{V}(T_*^\gamma, \beta) &= \beta T_*^\gamma u + (1-\beta) P_B u \\
 &= \beta (I - P_B + P_A R_B^\gamma) u + (1-\beta) P_B u \\
 &= \beta (u - g) + (1-\beta) f \\
 &= u
 \end{aligned} \tag{36}$$

or  $u \in \text{Fix } \mathcal{V}(T_*^\gamma, \beta)$ . Thus, the fixed points of GIF-RAAR coincide with those of RAAR.

### Acknowledgments

We gratefully acknowledge fruitful conversations with Dr. D. Russell Luke, Institut für Numerische und Angewandte Mathematik Universität Göttingen and Dr. Keith A. Kearnes, University of Colorado at Boulder. We would also like to extend our gratitude to Moritz Barkowski, Technische Universität Kaiserslautern. The authors also gratefully acknowledge support from a National Security Science and Engineering Faculty Fellowship and from the National Science Foundation Engineering Research Center in EUV Science and Technology.