ORIGINAL ARTICLE

On connections between Amplitude Flow and Error Reduction for phase retrieval and ptychography

Oleh Melnyk[1,2]

Abstract

In this paper, we consider two iterative algorithms for the phase retrieval problem: the well-known Error Reduction method and the Amplitude Flow algorithm, which performs minimization of the amplitude-based squared loss via the gradient descent. We show that Error Reduction can be interpreted as a scaled gradient method applied to minimize the same amplitude-based squared loss, which allows to establish its convergence properties. Moreover, we show that for a class of measurement scenarios, such as ptychography, both methods have the same computational complexity and sometimes even coincide.

✉ Oleh Melnyk
  oleh.melnyk@helmholtz-muenchen.de; oleh.melnyk@tum.de

1 Mathematical Imaging and Data Analysis, Helmholtz Center Munich, Ingolstaedter Landstrasse 1, Neuherberg 85764, Germany

2 Department of Mathematics, Technical University of Munich, Boltzmannstr. 3, Garching bei München 85748, Germany

1 Introduction

The phase retrieval problem considers the reconstruction of an unknown $x \in \mathbb{C}^d$ from $m \in \mathbb{N}$ amplitude measurements of the form

$$y_k = |(Ax)_k| + n_k, \quad k = 1, \ldots, m, \tag{1}$$

with $A \in \mathbb{C}^{m \times d}$ denoting the measurement matrix and $n \in \mathbb{R}^m$ being noise.

It has many applications such as crystallography [1], noncrystalline materials [2–4] and optical imaging [5], where the goal is to recover the specimen from its diffraction patterns obtained by illumination with penetrating light, e.g., x-rays or electron beam.

One such application is ptychography [6, 7], where the inference on the object of interest is based on a collection of far-field diffraction patterns, each obtained by an illumination of a small region of the specimen. As the regions overlap, it produces the surplus information, which allows for unique identification of the object up to a global phase factor from ptychographic measurements.

Since the introduction of the phase retrieval problem to the mathematical community, many approaches have been developed in order to reconstruct the specimen. The spectrum of methods includes alternating projections methods [8–14], gradient-based minimization [15–19], semidefinite [20–24] and linear programming [25, 26], direct methods [27–31], and many more.

One of the longstanding favored algorithms is Error Reduction (ER), which was introduced in 1972 by Gerchberg and Saxton [8]. Later contributions [10], [32] and [12] classified ER as an alternating projections technique and supplemented it with the detailed analysis on the convergence and also provided an interpretation of the algorithm as a projected gradient method. The version of ER algorithm was also studied as a gradient flow method in continuous setting [33].

Another algorithm, which became popular in recent years, is Amplitude Flow (AF) [17, 19]. It performs the first order optimization applied to the amplitude-based squared loss

$$\mathcal{L}_2(z) = \sum_{k=1}^{m} ||(Az)_k| - y_k|^2, \quad z \in \mathbb{C}^d. \tag{2}$$

AF is well-understood for randomized measurements scenarios [17], where the matrix A is random. It as well possesses the convergence guarantees for arbitrary measurement scenarios [19].

In this paper we connect these two methods by representing ER as a scaled gradient method for the minimization of the amplitude-based squared loss $\mathcal{L}_2$. It allows to establish convergence rate of the ER algorithm, which to our knowledge has never observed in the literature. Furthermore, the scaled gradient representation provides the equivalence between the set of fixed points of two methods. Lastly, we consider ER and AF in application to the ptychographic measurements and show that both methods exhibit the same computational complexity and in special cases even coincide.

The paper is structured in the following way. In Section 2 we provide the reader with necessary notation and detailed overview of the ER and AF algorithms. Our

contribution is then presented in Section 3 and proved in Section 4. Finally, the paper is summarized by a short conclusion.

2 Notation and Preliminaries

2.1 Definitions

Throughout the paper, we will use the short notation $[a] = \{1, 2, \ldots, a\}$ for index sets. The complex unit is denoted by i. The complex conjugate of $\alpha \in \mathbb{C}$ is given by $\bar{\alpha}$. The transpose and complex conjugate transpose of a vector v or a matrix B are denoted by v^t, v^* and B^T, B^*, respectively. The Euclidean norm of a vector $v \in \mathbb{C}^a$ is given by $\|v\|_2 := \left[\sum_{j=1}^{a} |v_j|^2 \right]^{1/2}$. We say that a matrix $B \in \mathbb{C}^{b \times a}$, $b \geq a$ is injective if for all pairs of vectors $u, v \in \mathbb{C}^a$ with $u \neq v$ it holds that $Bu \neq Bv$. The injectivity of B is equivalent to the condition $\text{rank}(B) = a$. We will also denote the image of B as

$$\text{im}(B) := \{Bv \in \mathbb{C}^b : v \in \mathbb{C}^a\}.$$

For a square full rank matrix B its inverse is given by B^{-1}. A matrix $B \in \mathbb{C}^{b \times a}$, $b \geq a$ is called orthogonal if it satisfies $B^* B = I$, where I denotes the identity matrix. A square orthogonal matrix $B \in \mathbb{C}^{a \times a}$ is a unitary matrix and its inverse is $B^{-1} = B^*$.

The projection of $u \in \mathbb{C}^a$ onto a set $\mathcal{S} \subseteq \mathbb{C}^a$ is an element $\tilde{u} \in \mathcal{S}$, such that $\|u - \tilde{u}\|_2 \leq \|u - v\|_2$ for all $v \in \mathcal{S}$. An operator, which maps u to $\tilde{u}$ is called the projection operator onto $\mathcal{S}$. In general, $\tilde{u}$ is not unique, however, in case when $\mathcal{S}$ is a non-empty closed convex set, $\tilde{u}$ can be uniquely identified [34].

For a matrix $B \in \mathbb{C}^{b \times a}$ of rank r its singular value decomposition is given by

$$B = U \Sigma V^*,$$

where $U \in \mathbb{C}^{b \times r}$, $V \in \mathbb{C}^{a \times r}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{r \times r}$ is an invertible diagonal matrix with diagonal entries $\sigma_j(B) > 0$, $j \in [r]$, sorted in decreasing order. The values $\sigma_j(B)$ are also referred to as the singular values of B. The largest singular value $\sigma_1(B)$ equals to the spectral norm of B defined as

$$\|B\| := \max_{v \in \mathbb{C}^a, \|v\|=1} \|Bv\|_2.$$

Using the singular value decomposition, the Moore-Penrose pseudoinverse of B is defined as

$$B^\dagger := V \Sigma^{-1} U^*.$$

For an injective matrix $B \in \mathbb{C}^{b \times a}$, $b \geq a$, its pseudoinverse $B^\dagger$ can be expressed as

$$B^\dagger = (B^* B)^{-1} B^*. \tag{3}$$

It satisfies

$$B^\dagger B = I \text{ and } BB^\dagger \text{ is a projection operator onto the set } \operatorname{im}(B). \tag{4}$$

For a vector $v \in \mathbb{C}^a$, the diagonal matrix $\operatorname{diag}(v) \in \mathbb{C}^{a \times a}$ is formed by placing the entries of the vector v onto main diagonal, so that for $k, j \in [a]$ it holds that

$$\operatorname{diag}(v)_{k,j} := \begin{cases} v_k & k = j, \\ 0 & k \neq j. \end{cases}$$

The discrete Fourier transform is given by a matrix $F \in \mathbb{C}^{d \times d}$ with the entries

$$F_{k,j} = e^{2\pi i (k-1)(j-1)/d}, \quad k, j \in [d], \tag{5}$$

and satisfies equality

$$F^* F = dI. \tag{6}$$

The family of the circular shift matrices $S_s \in \mathbb{C}^{d \times d}$, $s \in \mathbb{Z}$ is defined by its action for all vectors $v \in \mathbb{C}^d$ as

$$(S_s v)_j := v_{(j-1-s \bmod d)+1}, \quad j \in [d]. \tag{7}$$

For the description of the computational complexity of algorithms, we use notation $\mathcal{O}(n)$ for the order of operations, meaning that at most cn operations are required for some constant $c > 0$.

For a function $f : \mathbb{C} \to \mathbb{C}$ and a vector $v \in \mathbb{C}^a$, the notation $f(v)$ will denote the entrywise application of the function f. For a vector $v \in \mathbb{C}^a$ and number $\alpha \in \mathbb{C}$ by $v + \alpha$ we will denote vector in $\mathbb{C}^a$ with entries $v_k + \alpha$, $k \in [a]$. For instance, using this notation we can rewrite the measurements (1) as

$$y = |Ax| + n.$$

2.2 Phase retrieval

In the context of the phase retrieval problem, it is convenient to refer to the spaces $\mathbb{C}^d$ and $\mathbb{C}^m$ as *object and measurement spaces*, respectively. If the phases of the measurements Ax were known, the problem would be the classical recovery from linear measurements, which in general is only possible if the dimension of the measurement space m is at least as large as the dimension of the object space d. Since the phases are lost, the number of required measurements is even higher and, hence, we will assume that $m \geq d$. It is known that $m \geq 4d - 4$ measurements are sufficient for the unique reconstruction of x when A is generic [35] and $m \geq cd$ with constant $c \geq 1$ when A is random [23, 24]. By the unique reconstruction of x, we understand the identification of x up to a global phase factor αx for any $\alpha \in \mathbb{C}$, $|\alpha| = 1$, since it holds that

$$|Ax| = |A\alpha x|.$$

The unique reconstruction of x up to a global phase is equivalent to the unique identification of the set $\{\alpha x : |\alpha| = 1\}$ or, in other words, the function $\{\alpha x : |\alpha| = 1\} \mapsto |Ax|$ is injective. One of the necessary conditions for the unique recovery is the injectivity of the matrix A. If A is not injective then there exist two vectors $u, v \in \mathbb{C}^d$, such that $u \neq v$ and $Au = Av$. Consequently, $|A(u - v)|_j = 0$ and it is not possible to distinguish $u - v$ and the zero vector from the measurements. The injectivity of the matrix A will be the main assumption for our results in Section 3. The injectivity of A, however, is not sufficient for unique recovery. A counterexample is $A = F$, which is injective by (6), but it is well-known that there are multiple objects satisfying the same measurements $|Fx|$ [36].

2.3 Error Reduction

The Error Reduction (ER) is an iterative algorithm for the phase retrieval problem. It considers an initial guess $z^0 \in \mathbb{C}^d$ in the object space and is given by the iterations

$$z^{t+1} = A^\dagger \operatorname{diag}\left(\frac{y}{|Az^t|}\right) Az^t, \quad t \geq 0. \tag{ER}$$

The iterations are repeated until the fixed point is reached, so that $z^{t+1} = z^t$. For $T \in \mathbb{N}$ iterations of ER $\mathcal{O}(md^2 + Tmd)$ operations are required, where $\mathcal{O}(md^2)$ operations are needed to compute the pseudoinverse and $\mathcal{O}(md)$ operations are performed per iteration.

Let us consider iterates in the measurement space $u^t := Az^t, t \geq 0$ for which the update of ER reads as

$$u^{t+1} = Az^{t+1} = AA^\dagger \operatorname{diag}\left(\frac{y}{|Az^t|}\right) Az^t = AA^\dagger \operatorname{diag}\left(\frac{y}{|u^t|}\right) u^t.$$

In this form, $\operatorname{diag}\left(\frac{y}{|u^t|}\right) u^t$ is the projection of u^t onto the set

$$\mathcal{M} := \{u \in \mathbb{C}^m : |u| = y\}.$$

Moreover, the set $\mathcal{M}$ can be viewed as a product of one-dimensional sets

$$\mathcal{M}_k := \{\alpha \in \mathbb{C} : |\alpha| = y_k\}, \quad k \in [m],$$

and, thus, the projection onto $\mathcal{M}$ is performed by projecting each coordinate u_k^t onto corresponding set $\mathcal{M}_k$. The second step is to apply $AA^\dagger$, which is by (3), the projection operator onto $\operatorname{im}(A)$. Therefore, ER first projects onto $\mathcal{M}$, where the measurements are satisfied. Then, the resulting point is projected onto $\operatorname{im}(A)$. The sequential projections onto $\mathcal{M}$ and $\operatorname{im}(A)$ allows for an interpretation of ER as an alternating projection scheme. If $\mathcal{M}$ was a convex set, then ER would converge to the intersection of two sets [37]. However, due to non-convexity of $\mathcal{M}$, convergence of u^t to the intersection of the sets is not guaranteed, which is a known problem of the ER algorithm. We note

that, when A allows for unique recovery and noise is absent, intersection of $\mathcal{M}$ and $\text{im}(A)$ is given by $\{\alpha x : |\alpha| = 1\}$ [12].

Another complication arising from the non-convexity of $\mathcal{M}$ is the non-uniqueness of the projection onto $\mathcal{M}$. Let $y_k \neq 0$ and consider the projection of $\alpha \in \mathbb{C}$ onto $\mathcal{M}_k$. If α is non-zero, the closest point in $\mathcal{M}_k$ is given by $y_k \cdot \alpha/|\alpha|$ [12, Lemma 3.15a]. If $\alpha = 0$, all points in $\mathcal{M}_k$ have the same distance to 0 and any of them can be used as a projection. In the literature, it is resolved by setting the projection either to y_k or $y_k e^{i\varphi}$ for a randomly selected angle $\varphi \in [0, 2\pi)$. In this paper, we will instead map 0 to 0, which is not precisely the projection, but can be interpreted as an average of all possible projections

$$0 = \frac{1}{2\pi} \int_0^{2\pi} y_k e^{i\varphi} d\varphi.$$

Therefore, whenever $(Az^t)_k = 0$ we set $(Az^t)_k/|(Az^t)_k| = 0$.

The ER algorithm can also be interpreted as a projected gradient method [12, Section 3.8] applied to solve the minimization problem

$$\min_{u \in \text{im}(A)} \||u| - y\|_2^2. \tag{8}$$

We note, that substituting Az for u, $z \in \mathbb{C}^d$ leads to an unconstrained minimization of the amplitude-based objective (2), which suggests that ER can be interpreted as a gradient method applied to the function $\mathcal{L}_2$.

It is known that in the absence of noise if an initial guess z^0 is chosen sufficiently close to the set $\{\alpha x : |\alpha| = 1\}$, the ER algorithm will converge to a point in this set [12, Theorem 3.16]. In general, ER does not converge globally to $\{\alpha x : |\alpha| = 1\}$ [12, p.830]. If the loss $\mathcal{L}_2$ is differentiable at z^t, the ER iteration will not increase the value of $\mathcal{L}_2$, i.e., $\mathcal{L}_2(z^{t+1}) \leq \mathcal{L}_2(z^t)$ [12, 38].

For the initialization z^0 of ER, the polarization method can be used [12, 39, 40]. It constructs a matrix containing the estimates of $\text{sgn}(Ax_k)\overline{\text{sgn}(Ax_\ell)}, k, \ell \in [m]$ from the measurements and recovers $\text{sgn}(Ax_k)$ by solving the phase synchronization problem [41–44].

2.4 Amplitude Flow

The Amplitude Flow algorithm (AF) considers the gradient-based optimization of the amplitude-based objective (2). The algorithm is based on the Wirtinger derivatives, which are discussed in greater detail in Section 4.2, while in this section we superficially define the gradient in order to avoid lengthy derivations.

Given an initial guess $z^0 \in \mathbb{C}^d$, AF is based on the iterations

$$z^{t+1} = z^t - \mu_t \nabla \mathcal{L}_2(z^t), \quad t \geq 0, \tag{AF}$$

where $\mu_t > 0$ denotes the so-called learning rate and $\nabla \mathcal{L}_2$ is the generalized Wirtinger gradient of $\mathcal{L}_2$ given by

$$\nabla \mathcal{L}_2(z) = A^* \left[I - \text{diag}\left(\frac{y}{|Az|} \right) \right] Az.$$

Similarly to ER, we treat the case $(Az)_k = 0$ by setting $(Az)_k/|(Az)_k| = 0$. The iteration process is continued until the gradient $\nabla \mathcal{L}_2(z^t)$ vanishes, which is equivalent to reaching the fixed point $z^{t+1} = z^t$. Originally, AF was derived and analyzed for random Gaussian measurements without noise [17]. For such A, it is possible to construct good starting point z^0 via spectral initialization [15] or null initialization [45], such that AF admits linear convergence rate to the set of true solutions $\{\alpha x : |\alpha| = 1\}$. In general, for any choice of the measurement matrix A, the following convergence results have been established in [19].

Theorem 1 ([19, Theorem 1]) *Consider measurements y of the form* (1)*. Let $0 < \mu_t \leq \|A\|^{-2}$ and $z^0 \in \mathbb{C}^d$ be arbitrary. Then, for iterates $\{z^t\}_{t \geq 0}$ defined by AF we have*

$$\mathcal{L}_2(z^t) \geq \mathcal{L}_2(z^{t+1}) \text{ for all } t \geq 0,$$
$$\|z^{t+1} - z^t\|_2 \to 0, \ t \to \infty,$$

and

$$\min_{t=0,\dots,T-1} \|z^{t+1} - z^t\|_2^2 \leq \frac{\mathcal{L}_2(z^0)}{\|A\|^2 T} \text{ for all } T \geq 0.$$

Unlike the randomized scenario, the general case only guarantees convergence to a fixed point with sublinear rate. Therefore, the initialization z^0 is crucial for the convergence to global minimum. For a non-random A, e.g., in case of ptychography, an outcome of the direct (non-iterative) method [30] is a good starting point. Furthermore, with sufficiently good initialization AF can achieve linear convergence rate [46] even for non-random measurements.

As the proof of Theorem 1 resembles the proof of Theorem 3 below, we provide the sketch of proof for Theorem 1 in Remark 10 in Section 4.2.

The computational complexity of AF for $T \in \mathbb{N}$ iterations is given by $\mathcal{O}(Tmd)$ operations. If the learning rate is chosen to be $\mu_t = \|A\|^{-2}$, the computation of the spectral norm can be done with additional $\mathcal{O}(md)$ operations by performing the fixed number of the power method iterations. More precisely, for $K \in \mathbb{N}$ and random initialization v^0, iterates $v^k = A^* A v^{k-1}/\|A^* A v^{k-1}\|_2$, $k \in [K]$, are computed and $\|Av^K\|_2$ is used as an estimate of $\|A\|$.

3 Results

As it was briefly mentioned in Section 2.3, ER can be linked to the minimization of the amplitude-based objective (2). We formalize this intuition in the next lemma.

Lemma 2 *Let A be injective. Then, ER is a scaled gradient method with iterations given by*

$$z^{t+1} = z^t - (A^*A)^{-1}\nabla\mathcal{L}_2(z^t), \quad t \geq 0.$$

We emphasize that the result of Lemma 2 is only true for all $z \in \mathbb{C}^d$, if the ambiguity $0/0$ in the iteration of ER is defined as 0.

The reinterpretation of ER as a scaled gradient method allows to analyze convergence of the algorithm similarly to AF, which leads to an analogue of Theorem 1.

Theorem 3 *Consider the phase retrieval measurements y of the form* (1) *with injective matrix A. Let $z^0 \in \mathbb{C}^d$ be arbitrary. Then, for iterates $\{z^t\}_{t\geq 0}$ given by ER we have*

$$\mathcal{L}_2(z^t) \geq \mathcal{L}_2(z^{t+1}) \text{ for all } t \geq 0,$$
$$\|z^{t+1} - z^t\|_2 \to 0, \ t \to \infty,$$

and

$$\min_{t=0,\dots,T-1} \|z^{t+1} - z^t\|_2^2 \leq \frac{\mathcal{L}_2(z^0)}{T\sigma_d^2(A)} \text{ for all } T \geq 0,$$

where $\sigma_d(A)$ denotes the smallest singular value of the matrix A.

Theorem 3 guarantees that no matter how noisy the measurements are, ER will always converge to a fixed point and the convergence rate is sublinear. However, even in the absence of noise, it does not guarantee the global convergence to a point in the set $\{\alpha x : |\alpha| = 1\}$. We note that for cases $A = F$ and A corresponding to ptychography (see (10) below), the convergence of ER to a fixed point was shown in [10] and [47], respectively. However, the convergence rate was not derived. Comparing Theorem 3 to Theorem 1, we observe that the constant in the convergence rate of ER is worse by $\sigma_1^2(A)/\sigma_d^2(A)$ compared to AF.

A further consequence of Lemma 2 is the equality of the fixed-point sets of both algorithms.

Corollary 4 *Let A be injective. Then, $z \in \mathbb{C}^d$ is a fixed point of ER if and only if z is the fixed point of AF.*

We note that Corollary 4 does not imply that given the same initial guess z^0, both algorithms will necessarily converge to the same fixed point.

By Theorem 1 and Theorem 3, both algorithms seem to be comparable in terms of convergence rate and by Corollary 4 in terms of critical points. However, for $T \in \mathbb{N}$ iterations of ER $\mathcal{O}(md^2 + Tmd)$ operations are required, while AF only needs $\mathcal{O}(Tmd)$ operations and, thus, in general ER is considerably slower in terms of computation complexity. The next corollary shows, that this difference is less significant in cases where the columns of A are orthogonal.

Corollary 5 *Let*

$$A^*A = \text{diag}(v) \text{ for some } v \in \mathbb{R}^d \text{ with } v_\ell > 0, \ \ell \in [d]. \tag{9}$$

Then, for $T \in \mathbb{N}$ iterations both algorithms ER and AF require $\mathcal{O}(Tmd)$ operations.
Furthermore, if $A^*A = cI$, for some $c > 0$, then the iteration of ER coincides with the iteration of AF for the learning rate $\mu_t = \|A\|^{-2}$.

While condition (9) may seem restrictive, it, in fact, holds in many practical applications. For instance, the equivalence of both algorithms was observed for the recovery from Fourier magnitudes ($A = F$) in [10]. Another application of interest is ptychography, for which the measurement matrix A is given by

$$A = \begin{bmatrix} F \operatorname{diag}(S_{s_1} w) \\ \vdots \\ F \operatorname{diag}(S_{s_r} w) \end{bmatrix}, \tag{10}$$

where the vector $w \in \mathbb{C}^d$ denotes the distribution of the light in the illuminated region and $s_1, \ldots, s_r \in [d]$, $r \le d$, are unique positions of the regions. Matrices F and S_{s_j} are given by (5) and (7), respectively. When $r = d$ and $s_j = j$, $j \in [d]$, the matrix A is also known as the discrete Short-Time Fourier transform (STFT) with window w.

The next corollary shows that condition (9) and, consequently, the results of Corollary 5 also hold for ptychographic measurements.

Corollary 6 *Consider measurements of the form* (1) *with ptychographic measurement matrix A as in* (10). *Then, $A^*A = \operatorname{diag}(v)$, where the vector v has entries*

$$v_\ell = d \sum_{j \in [r]} |(S_{s_j} w)_\ell|^2,$$

for all $\ell \in [d]$. The matrix A is injective if and only if $v_\ell > 0$ for all $\ell \in [d]$. Furthermore, if A is the STFT matrix, the vector v has entries $v_\ell = d\|w\|_2^2$ for all $\ell \in [d]$. Consequently, the results of Corollary 5 apply for ptychographic measurements.

In order to illustrate the result of Corollary 6, we perform a numerical reconstructions of randomly generated $x \in \mathbb{C}^d$, $d = 256$ with both AF and ER. In the first case, A is chosen to be the STFT matrix with the window

$$w_j = \begin{cases} \exp\left(-\frac{(j-8.5)^2}{12.8} + i\frac{\pi(j-8.5)^2}{12.8}\right), & j \in [32], \\ 0, & j \notin [32]. \end{cases}$$

In the second case A is given by (10) with the same window and positions $s_j = 16j$ and $j \in [d/16]$. The measurements are additionally corrupted by Poisson noise such that signal-to-noise ratio $10 \log_{10}\left(\frac{\||Ax|^2\|_2^2}{\|y^2 - |Ax|^2\|_2^2}\right)$ is approximately 45. Figure 1a shows the values $\mathcal{L}_2(z^t)$ for 500 iterations of the algorithms starting from a random initialization z^0. Note that for the STFT matrix A, AF and ER coincide as predicted by Corollary 6, while this is no longer true for A as in (10). Despite producing different reconstructions, the runtime of the algorithms in Figure 1b is almost the same, which is in line with Corollary 5.
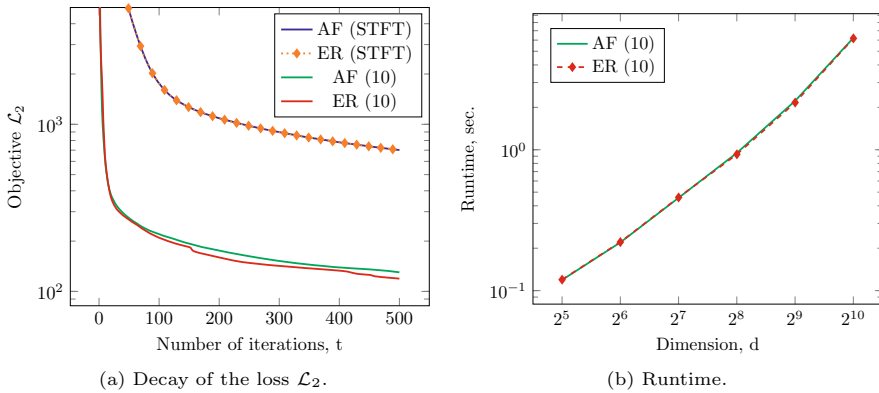
(a) Decay of the loss $\mathcal{L}_2$.   (b) Runtime.

Fig. 1 Numerical visualization of Corollary 5 and Corollary 6

Finally, we consider a scenario when the object is supported on $J \subseteq [d]$. Then, we can rewrite the the measurement model as

$$y = |AE_J x_J| + n,$$

where x_J is the vector containing entries of x in J and E_J is a linear embedding operator, which maps x_J to x. In such case, the results above apply for new measurement matrix $\tilde{A} = AE_J$.

4 Proofs

4.1 Proofs of Lemma 2 and corollaries

We will start with the proof of Lemma 2.

Proof of Lemma 2 By the assumption, A is injective and, thus, identities (3) and (4) hold true. Therefore, the iteration of ER can be rewritten as

$$
\begin{aligned}
z^{t+1} &= A^\dagger \operatorname{diag}\left(\frac{y}{|Az^t|}\right) Az^t = A^\dagger A z^t - A^\dagger A z^t + A^\dagger \operatorname{diag}\left(\frac{y}{|Az^t|}\right) Az^t \\
&= z^t - A^\dagger \left[I - \operatorname{diag}\left(\frac{y}{|Az^t|}\right)\right] Az^t = z^t - (A^*A)^{-1} A^* \left[I - \operatorname{diag}\left(\frac{y}{|Az^t|}\right)\right] Az^t \\
&= z^t - (A^*A)^{-1} \nabla \mathcal{L}_2(z^t).
\end{aligned}
$$

$\square$

Using the result of Lemma 2, we deduce Corollary 4 and Corollary 5.

Proof of Corollary 4 Let $z \in \mathbb{C}^d$ be a fixed point of ER. By Lemma 2, we have that

$$z = z - (A^*A)^{-1} \nabla \mathcal{L}_2(z),$$

which is equivalent to

$$(A^*A)^{-1}\nabla\mathcal{L}_2(z) = 0.$$

Since A is injective and $(A^*A)^{-1}$ exists, the obtained equality holds if and only if $\nabla\mathcal{L}_2(z) = 0$, so that z is the fixed point of AF. □

Proof of Corollary 5 Using the condition (9), we obtain $(A^*A)^{-1} = \mathrm{diag}(1/v)$. Consequently, by Lemma 2, the iteration of ER is given by

$$z^{t+1} = z^t - (A^*A)^{-1}\nabla\mathcal{L}_2(z^t) = z^t - \mathrm{diag}(1/v)\nabla\mathcal{L}_2(z^t).$$

The computation of the gradient requires $\mathcal{O}(md)$ operations. Both the multiplication with $\mathrm{diag}(1/v)$ and the difference can be done in $\mathcal{O}(d)$ operations. Therefore, the total number of operations for a single iteration of ER is given by $\mathcal{O}(md + d) = \mathcal{O}(md)$, which is the same order of operations as for a single iteration of AF. Furthermore, evaluation of v requires additional $\mathcal{O}(md)$ operations. We also note that $\|A\|^2 = \max_{\ell \in [d]} |v_\ell|$ and the computation of the learning rate is done in $\mathcal{O}(d)$ operations. Therefore, both algorithms have total complexity of $\mathcal{O}(Tmd)$ for T iterations.

If $A^*A = cI$, then

$$\|A\|^2 = \|A^*A\| = \|cI\| = c \text{ and } (A^*A)^{-1} = c^{-1}I = \|A\|^{-2}I.$$

Hence, using Lemma 2 for the iteration of ER we have

$$z^{t+1} = z^t - (A^*A)^{-1}\nabla\mathcal{L}_2(z^t) = z^t - \|A\|^{-2}\nabla\mathcal{L}_2(z^t),$$

which is precisely the iteration of AF with $\mu_t = \|A\|^{-2}$. □

The last corollary is the result of direct computations similar to the equation (12) in [19].

Proof of Corollary 6 We compute the product A^*A by using the representation (10),

$$A^*A = \sum_{j=1}^{r}(F\,\mathrm{diag}(S_{s_j}w))^*(F\,\mathrm{diag}(S_{s_j}w)) = \sum_{j=1}^{r}\mathrm{diag}^*(S_{s_j}w)F^*F\,\mathrm{diag}(S_{s_j}w).$$

Next, we use (6) and $\mathrm{diag}^*(S_{s_j}w) = \mathrm{diag}(\overline{S_{s_j}w})$ to obtain

$$A^*A = \sum_{j=1}^{r}d\,\mathrm{diag}(\overline{S_{s_j}w})\,\mathrm{diag}(S_{s_j}w) = \sum_{j=1}^{r}\mathrm{diag}(d|S_{s_j}w|^2)$$

$$= \mathrm{diag}\left(d\sum_{j=1}^{r}|S_{s_j}w|^2\right) = \mathrm{diag}(v).$$

The matrix A is injective if and only if A^*A is invertible, and the diagonal matrix is invertible when all its entries are non-zero. Since $v_\ell = d \sum_{j=1}^r |S_{s_j} w_\ell|^2 \geq 0, \ell \in [d]$, the injectivity of A is equivalent to $v_\ell > 0$ for all $\ell \in [d]$.

If A is the STFT matrix, then $s_j = j$ for all $j \in [d]$ and the entries of the vector v further simplify to

$$v_\ell = d \sum_{j=1}^d |S_{s_j} w_\ell|^2 = d \sum_{j=1}^d |w_{(\ell-j-1) \bmod d+1}|^2, \quad \ell \in [d].$$

Changing the order of summation yields

$$v_\ell = d \sum_{j=1}^d |w_j|^2 = d \|w\|_2^2,$$

for all $\ell \in [d]$, which concludes the proof. $\square$

4.2 Proof of Theorem 3

The proof of Theorem 3 is based on Wirtinger derivatives [48]. Let us recall some basic facts about Wirtinger derivatives based on [49, 50]. A function $f : \mathbb{C} \mapsto \mathbb{C}$ can be viewed as a function of two real variables, the real and imaginary parts of the argument $z = \alpha + i\beta$. The function f is said to be differentiable in real sense if the derivatives with respect to α and β exist.

Then, the Wirtinger derivatives are defined as

$$\frac{\partial f}{\partial z} := \frac{1}{2} \frac{\partial f}{\partial \alpha} - \frac{i}{2} \frac{\partial f}{\partial \beta}, \quad \frac{\partial f}{\partial \bar{z}} := \frac{1}{2} \frac{\partial f}{\partial \alpha} + \frac{i}{2} \frac{\partial f}{\partial \beta},$$

which is nothing, but a change of the coordinate system to conjugate coordinates. In this sense, we treat function f as a function of z and $\bar{z}$ instead of α and β.

As an example consider $f(z) = z = \alpha + i\beta$. Its Wirtinger derivatives are

$$\frac{\partial z}{\partial z} = \frac{1}{2} \frac{\partial(\alpha + i\beta)}{\partial \alpha} - \frac{i}{2} \frac{\partial(\alpha + i\beta)}{\partial \beta} = \frac{1}{2} - \frac{i^2}{2} = 1 \text{ and } \frac{\partial z}{\partial \bar{z}} = 0,$$

which implies that $\bar{z}$ can be treated as a constant when the derivative with respect to z is computed and vice versa.

Similar to the real analysis of multivariate functions, Wirtinger derivatives are extended for $f : \mathbb{C}^d \mapsto \mathbb{C}$, that is for $z \in \mathbb{C}^d$ they are given by

$$\frac{\partial f}{\partial z} = \left(\frac{\partial f}{\partial z_1}, \dots, \frac{\partial f}{\partial z_d} \right) \quad \text{and} \quad \frac{\partial f}{\partial \bar{z}} = \left(\frac{\partial f}{\partial \bar{z}_1}, \dots, \frac{\partial f}{\partial \bar{z}_d} \right).$$

The computation of Wirtinger derivatives is analogous to the standard real analysis as the arithmetic operations and the chain rule extends to the complex case. For

Wirtinger derivatives it also holds that

$$\overline{\frac{\partial f}{\partial z}} = \frac{\partial \bar{f}}{\partial \bar{z}} \quad \text{and} \quad \overline{\frac{\partial f}{\partial \bar{z}}} = \frac{\partial \bar{f}}{\partial z}, \tag{11}$$

for any differentiable function f.

The Wirtinger derivatives are particularly useful for optimization of real-valued functions of complex variables. Let $f : \mathbb{C}^d \mapsto \mathbb{R}$ be a differentiable real-valued function. Its differential can be presented in the form of Wirtinger derivatives as

$$df = \frac{\partial f}{\partial z} dz + \frac{\partial f}{\partial \bar{z}} d\bar{z}.$$

Since f is real-valued, by (11), it holds that

$$\overline{\frac{\partial f}{\partial z}} = \frac{\partial f}{\partial \bar{z}},$$

and the differential simplifies to

$$df = 2 \operatorname{Re} \left(\frac{\partial f}{\partial z} dz \right).$$

It is maximal, when dz is a scaled version of $\overline{\frac{\partial f}{\partial z}} = \frac{\partial f}{\partial \bar{z}}$ and, thus, $\frac{\partial f}{\partial \bar{z}}$ gives the direction of the steepest ascent. Moreover, the critical points of f are those, where derivative with respect to $\bar{z}$ vanishes. For this reason, the gradient of f is defined as

$$\nabla f := \left(\frac{\partial f}{\partial \bar{z}} \right)^T = \left(\frac{\partial f}{\partial z} \right)^*.$$

In our analysis, we would also need the Wirtinger version of the second order Taylor's approximation theorem in integral form. That is for all twice continuously differentiable functions $f : \mathbb{C}^d \mapsto \mathbb{R}$ and all $z, v \in \mathbb{C}^d$ it holds that

$$f(z+v) = f(z) + \begin{bmatrix} \nabla f \\ \overline{\nabla f} \end{bmatrix}^* \begin{bmatrix} v \\ \bar{v} \end{bmatrix} + \begin{bmatrix} v \\ \bar{v} \end{bmatrix}^* \int_0^1 (1-s)\nabla^2 f(z+sv)ds \begin{bmatrix} v \\ \bar{v} \end{bmatrix}, \tag{12}$$

where $\nabla^2 f$ denotes the Hessian matrix

$$\nabla^2 f = \begin{bmatrix} \nabla^2_{z,z} f & \nabla^2_{\bar{z},z} f \\ \hline \nabla^2_{\bar{z},z} f & \nabla^2_{z,z} f \end{bmatrix},$$

and its components are given by

$$\nabla^2_{z,z} f = \frac{\partial}{\partial z} \nabla f = \frac{\partial}{\partial z}\left(\frac{\partial f}{\partial z}\right)^* \quad \text{and} \quad \nabla^2_{\bar{z},z} f = \frac{\partial}{\partial \bar{z}} \nabla f = \frac{\partial}{\partial \bar{z}}\left(\frac{\partial f}{\partial z}\right)^*.$$

For further information on Wirtinger calculus, we refer reader to [49, 50].

Let us go back to the amplitude-based objective (2). We rewrite it as

$$\mathcal{L}_2(z) = \sum_{k=1}^{m} |\sqrt{|(Az)_k|^2} - \sqrt{y_k^2}|^2. \tag{13}$$

Since $\sqrt{\cdot}$ is not differentiable at 0, $\mathcal{L}_2$ is not differentiable on $\mathbb{C}^d$. Hence, the gradient of $\mathcal{L}_2$ is not properly defined for points z with $(Az)_k = 0$ for some $k \in [m]$. In order to overcome this issue, we consider the following smoothed version of (13),

$$\mathcal{L}_{2,\varepsilon}(z) := \sum_{k=1}^{m} |\sqrt{|(Az)_k|^2 + \varepsilon} - \sqrt{y_k^2 + \varepsilon}|^2, \tag{14}$$

where $\varepsilon > 0$. The function $\mathcal{L}_{2,\varepsilon}$ possesses some useful properties. Firstly, $\mathcal{L}_{2,\varepsilon}$ is continuous in ε and we have

$$\mathcal{L}_2(z) = \lim_{\varepsilon \to 0+} \mathcal{L}_{2,\varepsilon}(z).$$

Secondly, we can compute the gradient of $\mathcal{L}_{2,\varepsilon}$ everywhere and properly define the generalized gradient of $\mathcal{L}_2$ as the limit of gradients as parameter ε vanishes.

Lemma 7 *The function $\mathcal{L}_{2,\varepsilon}$ is continuously differentiable with the gradient given by*

$$\nabla \mathcal{L}_{2,\varepsilon}(z) = A^* \left[I - \mathrm{diag}\left(\frac{\sqrt{y^2 + \varepsilon}}{\sqrt{|Az|^2 + \varepsilon}} \right) \right] Az, \quad z \in \mathbb{C}^d.$$

Furthermore, the generalized gradient of $\mathcal{L}_2$ is given by the pointwise limit

$$\nabla \mathcal{L}_2(z) := \lim_{\varepsilon \to 0+} \nabla \mathcal{L}_{2,\varepsilon}(z).$$

Proof Denote by a_k the conjugate of the k-th row of the matrix A, so that $(Az)_k = a_k^* z$. Then, a single summand of $\mathcal{L}_{2,\varepsilon}$ is given by

$$f_k(z) := |\sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon} - \sqrt{y_k^2 + \varepsilon}|^2, \quad k \in [m].$$

The gradient of f_k can be evaluated by the chain rule. We get

$$\nabla f_k(z) = \left[\frac{\partial f_k}{\partial \bar{z}}(z)\right]^T = \left[\frac{\partial |\sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon} - \sqrt{y_k^2 + \varepsilon}|^2}{\partial \sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon} - \sqrt{y_k^2 + \varepsilon}}\right.$$

$$\left. \cdot \frac{\partial \sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon} - \sqrt{y_k^2 + \varepsilon}}{\partial z^T \bar{a}_k a_k^T \bar{z} + \varepsilon} \cdot \frac{\partial z^T \bar{a}_k a_k^T \bar{z} + \varepsilon}{\partial \bar{z}}\right]^T$$

$$= 2\left(\sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon} - \sqrt{y_k^2 + \varepsilon}\right) \frac{1}{2\sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon}} \left[z^T \bar{a}_k a_k^T\right]^T$$

$$= \left(1 - \frac{\sqrt{y_k^2 + \varepsilon}}{\sqrt{z^T \bar{a}_k a_k^T \bar{z} + \varepsilon}}\right) a_k a_k^* z = \left(1 - \frac{\sqrt{y_k^2 + \varepsilon}}{\sqrt{|(Az)_k|^2 + \varepsilon}}\right) (Az)_k a_k.$$

Then, by the linearity of derivatives,

$$\nabla \mathcal{L}_{2,\varepsilon}(z) = \sum_{k=1}^{m} \nabla f_k(z) = \sum_{k=1}^{m} \left(1 - \frac{\sqrt{y_k^2 + \varepsilon}}{\sqrt{|(Az)_k|^2 + \varepsilon}}\right) (Az)_k a_k$$

$$= A^* \left[I - \text{diag}\left(\frac{\sqrt{y^2 + \varepsilon}}{\sqrt{|Az|^2 + \varepsilon}}\right)\right] Az.$$

For the generalized gradient of $\mathcal{L}_2$ we consider two cases. If $(Az)_k \neq 0$ for all $k \in [m]$, then $\sqrt{y_k^2 + \varepsilon}/\sqrt{|(Az)_k|^2 + \varepsilon} \to y_k/|(Az)_k|$, $\varepsilon \to 0+$, for all $k \in [m]$. Note that in this case, $\mathcal{L}_2$ is differentiable at z and its gradient coincides with the limit of $\nabla \mathcal{L}_{2,\varepsilon}(z)$ as $\varepsilon \to 0+$. On the other hand, if $(Az)_k = 0$ for some $k \in [m]$, it holds that

$$\frac{\sqrt{y_k^2 + \varepsilon}}{\sqrt{|(Az)_k|^2 + \varepsilon}} (Az)_k = \frac{\sqrt{y_k^2 + \varepsilon}}{\sqrt{0 + \varepsilon}} \cdot 0 = 0 \to 0 = \frac{y_k}{|(Az)_k|} (Az)_k, \quad \varepsilon \to 0+,$$

with ambiguity $0/0$ resolved as 0. $\square$

The last property concerns the Hessian matrix of $\mathcal{L}_{2,\varepsilon}$.

Lemma 8 [19] *The function $\mathcal{L}_{2,\varepsilon}$ is twice continuously differentiable and its Hessian matrix satisfies*

$$\begin{bmatrix} v \\ \bar{v} \end{bmatrix}^* \nabla^2 \mathcal{L}_{2,\varepsilon}(z) ds \begin{bmatrix} v \\ \bar{v} \end{bmatrix} \leq 2v^* A^* A v, \quad \text{for all } z, v \in \mathbb{C}^d, \varepsilon > 0.$$

Proof See computations on pages 27-28 of [19]. $\square$

Remark 9 The convergence of gradient descent to a critical point is often studied [51] under the assumption that the function f is L-smooth with $L \geq 0$. For twice continuously differentiable functions L-smoothness is equivalent to the inequality

$$-2L\|v\|_2^2 \leq \begin{bmatrix} v \\ \bar{v} \end{bmatrix}^* \nabla^2 f(z)ds \begin{bmatrix} v \\ \bar{v} \end{bmatrix} \leq 2L\|v\|_2^2, \quad \text{for all } z, v \in \mathbb{C}^d.$$

In fact, only the upper bound is sufficient to establish convergence of gradient decent. In our case, its stronger version is given by Lemma 8.

Now, we are equipped for the proof of Theorem 3.

Proof of Theorem 3 In view of Lemma 2, let us consider the smoothed step of the ER algorithm

$$z_\varepsilon^+ := z - (A^*A)^{-1}\nabla\mathcal{L}_{2,\varepsilon}(z).$$

Note that $(A^*A)^{-1}$ exists due to injectivity of A.

We first show that the single step of the smoothed Error Reduction works for minimization of $\mathcal{L}_{2,\varepsilon}$ and then we take the pointwise limits to obtain desired result for $\mathcal{L}_2$. In order to derive that in each iteration step the objective does not increase, we apply Taylor's theorem (12) with arbitrary $z \in \mathbb{C}^d$ and $v = -(A^*A)^{-1}\nabla\mathcal{L}_{2,\varepsilon}(z)$. We note that by Lemma 8, the integral in (12) is bounded as

$$\int_0^1 (1-s) \begin{bmatrix} v \\ \bar{v} \end{bmatrix}^* \nabla^2\mathcal{L}_{2,\varepsilon}(z+sv) \begin{bmatrix} v \\ \bar{v} \end{bmatrix} ds \leq 2v^*A^*Av \int_0^1 (1-s)ds = v^*A^*Av.$$

Hence, by (12), we have

$$\begin{aligned}
\mathcal{L}_{2,\varepsilon}(z_\varepsilon^+) &\leq \mathcal{L}_{2,\varepsilon}(z) - 2[\nabla\mathcal{L}_{2,\varepsilon}(z)]^*(A^*A)^{-1}\nabla\mathcal{L}_{2,\varepsilon}(z) \\
&\quad + [\nabla\mathcal{L}_{2,\varepsilon}(z)]^*((A^*A)^{-1})^*(A^*A)(A^*A)^{-1}\nabla\mathcal{L}_{2,\varepsilon}(z) \\
&= \mathcal{L}_{2,\varepsilon}(z) - [\nabla\mathcal{L}_{2,\varepsilon}(z)]^*(A^*A)^{-1}\nabla\mathcal{L}_{2,\varepsilon}(z),
\end{aligned}$$

where we used that $((A^*A)^{-1})^* = ((A^*A)^*)^{-1} = (A^*A)^{-1}$. Lemma 7 gives

$$z_\varepsilon^+ \to z^+ := z - (A^*A)^{-1}\nabla\mathcal{L}_2(z), \quad \varepsilon \to 0+,$$

and, thus, taking the limit $\varepsilon \to 0+$ yields

$$\mathcal{L}_2(z^+) \leq \mathcal{L}_2(z) - [\nabla\mathcal{L}_2(z)]^*(A^*A)^{-1}\nabla\mathcal{L}_2(z).$$

Selecting z as iterates z^t of ER, we obtain

$$\mathcal{L}_2(z^{t+1}) \leq \mathcal{L}_2(z^t) - [\nabla\mathcal{L}_2(z^t)]^*(A^*A)^{-1}\nabla\mathcal{L}_2(z^t). \tag{15}$$

Since A is injective, its singular value decomposition is given by $A = U\Sigma V^*$ with orthogonal $U \in \mathbb{C}^{m \times d}$, unitary $V \in \mathbb{C}^{d \times d}$ and invertible diagonal matrix $\Sigma \in \mathbb{C}^{d \times d}$. Then,

$$(A^*A)^{-1} = (V\Sigma^2 V^*)^{-1} = (V^*)^{-1}\Sigma^{-2}V^{-1} = V\Sigma^{-2}V^* = (V\Sigma^{-1})(V\Sigma^{-1})^* \tag{16}$$

is the singular value decomposition of $(A^*A)^{-1}$. From this representation we deduce that

$$[\nabla\mathcal{L}_2(z^t)]^*(A^*A)^{-1}\nabla\mathcal{L}_2(z^t) = \|(V\Sigma^{-1})^*\nabla\mathcal{L}_2(z^t)\|_2^2 \geq 0.$$

Thus, by (15),

$$\mathcal{L}_2(z^{t+1}) \leq \mathcal{L}_2(z) - [\nabla\mathcal{L}_2(z)]^*(A^*A)^{-1}\nabla\mathcal{L}_2(z) \leq \mathcal{L}_2(z^t),$$

which shows the first statement of Theorem 3.

In order to prove the remaining statements of Theorem 3, we need to link the decay of the objective to the iterates. By Lemma 2, we have that

$$\|z^{t+1} - z^t\|_2^2 = \|(A^*A)^{-1}\nabla\mathcal{L}_2(z^t)\|_2^2 = [\nabla\mathcal{L}_2(z^t)]^*(A^*A)^{-1}(A^*A)^{-1}\nabla\mathcal{L}_2(z^t).$$

Using (16) and the definition of the spectral norm, the squared distance between the iterates can be bounded as

$$\begin{aligned}
\|z^{t+1} - z^t\|_2^2 &= (\Sigma^{-1}V^*\nabla\mathcal{L}_2(z^t))^*\Sigma^{-2}(\Sigma^{-1}V^*\nabla\mathcal{L}_2(z^t)) \\
&= \|\Sigma^{-1}(\Sigma^{-1}V^*\nabla\mathcal{L}_2(z^t))\|_2^2 \leq \|\Sigma^{-1}\|^2\|\Sigma^{-1}V^*\nabla\mathcal{L}_2(z^t)\|_2^2 \\
&= \sigma_1^2(\Sigma^{-1})[\nabla\mathcal{L}_2(z^t)]^*V\Sigma^{-1}\Sigma^{-1}V^*\nabla\mathcal{L}_2(z^t) \\
&= \sigma_d^{-2}(A)[\nabla\mathcal{L}_2(z^t)]^*(A^*A)^{-1}\nabla\mathcal{L}_2(z^t).
\end{aligned}$$

Next, we sum up the norms for $T \in \mathbb{N}$ iterations of ER and apply (15) to obtain

$$\begin{aligned}
\sum_{t=0}^{T-1}\|z^{t+1} - z^t\|_2^2 &\leq \sigma_d^{-2}(A)\sum_{t=0}^{T-1}[\nabla\mathcal{L}_2(z^t)]^*(A^*A)^{-1}\nabla\mathcal{L}_2(z^t) \\
&\leq \sigma_d^{-2}(A)\sum_{t=0}^{T-1}\left[\mathcal{L}_2(z^t) - \mathcal{L}_2(z^{t+1})\right] \\
&= \sigma_d^{-2}(A)\left[\mathcal{L}_2(z^0) - \mathcal{L}_2(z^T)\right] \leq \sigma_d^{-2}(A)\mathcal{L}_2(z^0),
\end{aligned}$$

where in the last line we used that $\mathcal{L}_2(z) \geq 0$ for all $z \in \mathbb{C}^d$. This implies that the partial sum of the series $\sum_{t=0}^{\infty}\|z^{t+1} - z^t\|_2^2$ is bounded and, thus, the series is convergent. Consequently, summands converge to zero, that is $\|z^{t+1} - z^t\|_2^2 \to 0, t \to$

∞. Furthermore, we have

$$\min_{t=0,\ldots,T-1} \|z^{t+1} - z^t\|_2^2 \leq \frac{1}{T} \sum_{t=0}^{T-1} \|z^{t+1} - z^t\|_2^2 \leq \frac{\mathcal{L}_2(z^0)}{\sigma_d^2(A)T},$$

which concludes the proof. □

Remark 10 Proof of Theorem 1 follows the same logic. By Taylor's theorem (12) with $v = \mu_t \mathcal{L}_2(z)$, the analogue of inequality (15) is established. Since the learning rate is a positive constant, it further implies that the norm of the gradient converges to zero with the desired speed similarly to the proof of Theorem 3.

5 Conclusion

In this paper we established the understanding of the Error Reduction algorithm as a scaled gradient method and derived its convergence rate. Furthermore, it was shown that in practical scenarios, Error Reduction has the same computational complexity as the Amplitude Flow method and the two algorithms coincide in some cases.

In the future, we plan to expand our analysis for the Hybrid Input-Output method [9] and extended Ptychograpic Iterative Engine [52] used for the problem of blind ptychography.

Declarations

Conflict of interest The author has neither the conflict of interests nor relevant financial or non-financial interests to disclose.

References

1. Liu, Z.C., Xu, R., Dong, Y.H.: Phase retrieval in protein crystallography Section A foundations of crystallography. Acta. Cryst **68**(Pt 2), 256–265 (2012). https://doi.org/10.1107/S0108767311053815

2. Shapiro, D., Thibault, P., Beetz, T., Elser, V., Howells, M., Jacobsen, C., Kirz, J., Lima, E., Miao, H., Neiman, A.M., Sayre, D.: Biological imaging by soft X-ray diffraction microscopy. Proc. Natl. Acad. Sci. USA **102**(43), 15343–15346 (2005). https://doi.org/10.1073/pnas.0503305102

3. Thibault, P., Elser, V., Jacobsen, C., Shapiro, D., Sayre, D.: Reconstruction of a yeast cell from X-ray diffraction data Section A foundations of crystallography. Acta. Cryst. **62**(Pt 4), 248–261 (2006). https://doi.org/10.1107/S0108767306016515

4. Miao, J., Ishikawa, T., Shen, Q., Earnest, T.: Extending X-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes. Ann. Rev. Phys. Chem. **59**, 387–410 (2008). https://doi.org/10.1146/annurev.physchem.59.032607.093642

5. Shechtman, Y., Eldar, Y.C., Cohen, O., Chapman, H.N., Miao, J., Segev, M.: Phase retrieval with application to optical imaging: A contemporary overview. IEEE Signal Proc. Mag. **32**(3), 87–109 (2015). https://doi.org/10.1109/MSP.2014.2352673

6. Pfeiffer, F.: X-ray ptychography. Nat. Photonics **12**(1), 9–17 (2018). https://doi.org/10.1038/s41566-017-0072-5

7. Chen, Z., Jiang, Y., Shao, Y.T., Holtz, M.E., Odstrčil, M., Guizar-Sicairos, M., Hanke, I., Ganschow, S., Schlom, D.G., Muller, D.A.: Electron ptychography achieves atomic-resolution limits set by lattice vibrations. Science **372**(6544), 826–831 (2021). https://doi.org/10.1126/science.abg2533

8. Gerchberg, R.W., Saxton, W.O.: A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik **35**, 237 (1972)

9. Fienup, J.R.: Reconstruction of an object from the modulus of its Fourier transform. Optics Lett. **3**(1), 27–29 (1978). https://doi.org/10.1364/ol.3.000027

10. Fienup, J.R.: Phase retrieval algorithms: A comparison. Appl. Opt. **21**(15), 2758–2769 (1982). https://doi.org/10.1364/AO.21.002758

11. Wen, Z., Yang, C., Liu, X., Marchesini, S.: Alternating direction methods for classical and ptychographic phase retrieval. Inverse Problems **28**(11), 115,010 (2012). https://doi.org/10.1088/0266-5611/28/11/115010

12. Marchesini, S., Tu, Y.C., Wu, H.T.: Alternating projection, ptychographic imaging and phase synchronization. Appl. Comput. Harmon. Anal. **41**(3), 815–851 (2016). https://doi.org/10.1016/j.acha.2015.06.005

13. Chang, H., Lou, Y., Duan, Y., Marchesini, S.: Total variation-based phase retrieval for Poisson noise removal. SIAM J. Imag. Sci. **11**(1), 24–55 (2018). https://doi.org/10.1137/16M1103270

14. Fannjiang, A., Zhang, Z.: Fixed point analysis of Douglas-Rachford splitting for ptychography and phase retrieval. SIAM J. Imag. Sci. **13**(2), 609–650 (2020). https://doi.org/10.1137/19M128781X

15. Candes, E.J., Li, X., Soltanolkotabi, M.: Phase retrieval via Wirtinger Flow: Theory and algorithms. IEEE Trans. Inform. Theory **61**(4), 1985–2007 (2015). https://doi.org/10.1109/TIT.2015.2399924

16. Chen, Y., Candes, E.J.: Solving random quadratic systems of equations is nearly as easy as solving linear systems. NIPS (2015)

17. Wang, G., Giannakis, G.B., Eldar, Y.C.: Solving systems of random quadratic equations via Truncated Amplitude Flow. IEEE Trans Inform Theory **64**(2), 773–794 (2018). https://doi.org/10.1109/TIT.2017.2756858

18. Wang, G., Giannakis, G.B., Saad, Y., Chen, J.: Phase retrieval via Reweighted Amplitude Flow. IEEE Trans. Signal Proc. p. 1 (2018). https://doi.org/10.1109/TSP.2018.2818077

19. Xu, R., Soltanolkotabi, M., Haldar, J.P., Unglaub, W., Zusman, J., Levi, A.F.J., Leahy, R.M.: Accelerated Wirtinger Flow: A fast algorithm for ptychography. arXiv:1806.05546

20. Candès, E.J., Eldar, Y.C., Strohmer, T., Voroninski, V.: Phase retrieval via matrix completion. SIAM J. Imag. Sci. **6**(1), 199–225 (2013). https://doi.org/10.1137/110848074

21. Candès, E.J., Strohmer, T., Voroninski, V.: Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. Commun. Pure Appl. Math. **66**(8), 1241–1274 (2013). https://doi.org/10.1002/cpa.21432

22. Kueng, R., Rauhut, H., Terstiege, U.: Low rank matrix recovery from rank one measurements. Appl. Comput. Harm. Anal. **42**(1), 88–116 (2017). https://doi.org/10.1016/j.acha.2015.07.007

23. Kabanava, M., Kueng, R., Rauhut, H., Terstiege, U.: Stable low-rank matrix recovery via null space properties. Inform. Infer. J. IMA **5**(4), 405–441 (2016). https://doi.org/10.1093/imaiai/iaw014

24. Krahmer, F., Kümmerle, C., Melnyk, O.: On the robustness of noise-blind low-rank recovery from rank-one measurements. Linear Algebra Appl. **652**, 37–81 (2022) https://doi.org/10.1016/j.laa.2022.07.002. https://www.sciencedirect.com/science/article/pii/S0024379522002609

25. Goldstein, T., Studer, C.: Phasemax: Convex phase retrieval via Basis Pursuit. IEEE Trans. Inform. Theory **64**(4), 2675–2689 (2018). https://doi.org/10.1109/TIT.2018.2800768
26. Ghods, R., Lan, A.S., Goldstein, T., Studer, C.: in 52nd Annual Conference on Information Sciences and Systems (CISS) (IEEE, Princeton, NJ, 2018), pp. 1–6. https://doi.org/10.1109/CISS.2018.8362270
27. Chapman, H.N.: Phase-retrieval X-ray microscopy by Wigner-distribution deconvolution. Ultramicroscopy **66**(3–4), 153–172 (1996). https://doi.org/10.1016/S0304-3991(96)00084-8
28. Iwen, M.A., Viswanathan, A., Wang, Y.: Fast phase retrieval from local correlation measurements. SIAM J. Imag. Sci. **9**(4), 1655–1688 (2016). https://doi.org/10.1137/15M1053761
29. Iwen, M.A., Preskitt, B., Saab, R., Viswanathan, A.: Phase retrieval from local measurements: Improved robustness via eigenvector-based angular synchronization. Appl. Comput. Harm. Anal. **48**(1), 415–444 (2020). https://doi.org/10.1016/j.acha.2018.06.004
30. Forstner, A., Krahmer, F., Melnyk, O., Sissouno, N.: Well-conditioned ptychographic imaging via lost subspace completion. Inverse Problems **36**(10), 105,009 (2020). https://doi.org/10.1088/1361-6420/abaf3a
31. Perlmutter, M., Merhi, S., Viswanathan, A., Iwen, M.: Inverting spectrogram measurements via aliased Wigner distribution deconvolution and angular synchronization. Inform. Infer. J. IMA (2020). https://doi.org/10.1093/imaiai/iaaa023
32. Levi, A., Stark, H.: in ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing (Institute of Electrical and Electronics Engineers, San Diego, CA, USA, 1984), pp. 88–91. https://doi.org/10.1109/ICASSP.1984.1172785
33. Tsipenyuk, A.: Variational approach to Fourier phase retrieval. Doctoral Thesis
34. Aubin, J.P.: Applied Functional Analysis, 2nd edn. Pure and applied mathematics (John Wiley & Sons, Inc, Hoboken, NJ, USA, 2000). https://doi.org/10.1002/9781118032725. https://onlinelibrary.wiley.com/doi/book/10.1002/9781118032725
35. Conca, A., Edidin, D., Hering, M., Vinzant, C.: An algebraic characterization of injectivity in phase retrieval. Appl. Comput. Harmon. Anal. **38**(2), 346–356 (2015). https://doi.org/10.1016/j.acha.2014.06.005
36. Beinert, R., Plonka, G.: Ambiguities in one-dimensional discrete phase retrieval from Fourier magnitudes. J. Fourier Anal. Appl. **21**(6), 1169–1198 (2015). https://doi.org/10.1007/s00041-015-9405-2
37. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**(3), 367–426 (1996). https://doi.org/10.1137/S0036144593251710
38. Qian, J., Yang, C., Schirotzek, A., Maia, F., Marchesini, S.: Efficient Algorithms for Ptychographic Phase Retrieval. Inverse Problems and Applications. Contemp. Math **615**, 261–280 (2014)
39. Alexeev, B., Bandeira, A.S., Fickus, M., Mixon, D.G.: Phase retrieval with polarization. SIAM J. Imag. Sci. **7**(1), 35–66 (2014). https://doi.org/10.1137/12089939X
40. Pfander, G.E., Salanevich, P.: Robust phase retrieval algorithm for time-frequency structured measurements. SIAM J. Imag. Sci. **12**(2), 736–761 (2019). https://doi.org/10.1137/18M1205522
41. Singer, A.: Angular synchronization by eigenvectors and semidefinite programming. Appl. Comput. Harmon. Anal. **30**(1), 20–36 (2011). https://doi.org/10.1016/j.acha.2010.02.001
42. Boumal, N.: Nonconvex phase synchronization. SIAM J. Optim. **26**(4), 2355–2377 (2016). https://doi.org/10.1137/16M105808X
43. Bandeira, A.S., Boumal, N., Singer, A.: Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. Math. Programm. **163**(1–2), 145–167 (2017). https://doi.org/10.1007/s10107-016-1059-6
44. Filbir, F., Krahmer, F., Melnyk, O.: On recovery guarantees for angular synchronization. J. Fourier Anal. Appl. **27**(2) (2021). https://doi.org/10.1007/s00041-021-09834-1
45. Chen, P., Fannjiang, A., Liu, G.R.: Phase retrieval by linear algebra. SIAM J. Matrix Anal. Appl. **38**(3), 854–868 (2017). https://doi.org/10.1137/16M1107747
46. Bendory, T., Eldar, Y.C., Boumal, N.: Non-convex phase retrieval from STFT measurements. IEEE Trans. Inform. Theory **64**(1), 467–484 (2018). https://doi.org/10.1109/TIT.2017.2745623
47. Griffin, D., Lim, J.: Signal estimation from modified Short-Time Fourier transform. IEEE Trans. Acoustics Speech Signal Proc. **32**(2), 236–243 (1984). https://doi.org/10.1109/TASSP.1984.1164317
48. Wirtinger, W.: Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen. Math. Annalen **97**(1), 357–375 (1927). https://doi.org/10.1007/BF01447872
49. Hunger, R.: An introduction to complex differentials and complex differentiability (2008). https://mediatum.ub.tum.de/doc/631019/631019.pdf

50. Bouboulis, P.: Wirtinger's calculus in general Hilbert spaces. arXiv:1005.5170
51. Beck, A.: First-order methods in optimization (Society for Industrial and Applied Mathematics. Philadelphia (2017). https://doi.org/10.1137/1.9781611974997
52. Maiden, A.M., Rodenburg, J.M.: An improved ptychographical phase retrieval algorithm for diffractive imaging. Ultramicroscopy **109**(10), 1256–1262 (2009). https://doi.org/10.1016/j.ultramic.2009.05.012