Phase Retrieval

Ruicheng Ao 1900012179

February 28, 2022

1 Problem description

One popular formulation of the phase retrieval problem is solving a system of quadratic equations in the form

$$y_i = |\langle a_i, z\rangle|^2, i = 1, 2, \ldots, m, \tag{1}$$

where $z \in \mathcal{C}^n$ is the decision variable, $a_i \in \mathcal{C}^n$ are known sampling vectors, $\langle a_i, z\rangle$ is the inner product between a_i and z, $|\cdot|$ denotes the norm and $y_i \in \mathbb{R}$ are the observed measurements. The relationship can be reformulated into minimal square error problem

$$\min_{z\in\mathcal{C}^n} \quad f(z) = \tfrac{1}{2m}\textstyle\sum_{r=1}^m (|\langle a_r, z\rangle|^2 - y_r)^2. \tag{2}$$

In this report, we mainly focus on the algorithm, which starts with a careful initialization obtained by means of eigenvector and iteratively applies novel update rules proposed in [1] that solves the problem (2) efficiently.

2 Algorithm description

The algorithm chooses carefully the initial point z^0 via a spectral method so that the iterates $\{z^t\}_{t=0}^{\infty}$ converge to the solution with high probability. To speak more specifically, z^0 is chosen as the leading eigenvector of the positive definite Hermitian matrix $\sum_{r=1}^m y_r a_r a_r^*$. The framework is given in Algorithm 1.

Algorithm 1 Wirtinger Flow: Initialization

Input Observation $A = [a_1, a_2, \ldots, a_m] \in \mathcal{C}^{n\times m}, y = [y_1, y_2, \ldots, y_m] \in \mathbb{R}^m$, the number of iterations T

1: Set $\lambda = \min\{\sqrt{n\frac{\sum_{r=1}^m y_r}{\sum_{r=1}^m \|a_r\|_2^2}}, \sqrt{\frac{\sum_{r=1}^m y_r}{\mathrm{numel}(y)}}\}$
2: **for** $t = 1, 2, \ldots, T$ **do**
3:     Update $z^0 = A\,\mathrm{diag}(y)A^*z^0$
4:     Normalize $z^0 = z^0/\|z^0\|_2^2$
5: **end for**
6: Normalize $z^0 = \lambda z^0$
7: **return** z^0

After initializing the iterate z^0, a vanilla gradient descent method is applied in [1], namely,

$$\begin{aligned} z^{t+1} &= z^t - \frac{\mu_t}{\|z^0\|^2}\nabla f(z^t) \\ &= z^t - \frac{\mu_t}{\|z^0\|^2}(\frac{1}{m}\sum_{r=1}^m(|a_r^* z| - y_r)a_r a_r^* z), \end{aligned} \tag{3}$$

μ	$n = 64$				
	0.02	0.05	0.1	0.2	0.3
$m = 400$	$5.79e-05$	$6.86e-10$	$2.22e-15$	$3.63e-16$	$2.53e-16$
$m = 800$	$1.70e-04$	$3.37e-09$	$2.34e-15$	$4.52e-16$	$2.80e-16$
$m = 1200$	$3.57e-05$	$4.36e-10$	$2.52e-15$	$3.22e-16$	$3.12e-16$
μ	$n = 256$				
	0.02	0.05	0.1	0.2	0.3
$m = 400$	$1.55e-04$	$5.48e-09$	$2.49e-15$	$4.42e-16$	$3.71e-16$
$m = 800$	$5.77e-05$	$9.52e-10$	$1.92e-15$	$3.69e-16$	$3.94e-16$
$m = 1200$	$1.13e+00$	$1.14e+00$	$1.14e+00$	$1.14e+00$	$1.14e+00$
μ	$n = 256$				
	0.02	0.05	0.1	0.2	0.3
$m = 400$	$6.08e-02$	$9.47e-05$	$2.05e-09$	$8.84e-16$	$5.35e-16$
$m = 800$	$3.24e-03$	$8.42e-06$	$1.11e-09$	$6.04e-16$	$4.85e-16$
$m = 1200$	$1.85e-03$	$4.58e-06$	$3.66e-10$	$6.12e-16$	$4.89e-16$

Table 1: Relative error after 3000 iterations by Wirtinger Flow on dataset Gaussian

which is a form of steepest descent and the parameter μ_t is the step size. Note nonetheless that the feective step size is inversely proportional to the magnitude of the initial guess.

3 Numerical experiments

We test the introduced algorithm on two datasets for 1D Wirtinger Flow in [1], namely, Gaussian dataset and coded diffraction patterns (CDP) dataset to show the efficiency. For Gaussian dataset, we test the case when $n = 64, 128, 256$ and $m = 400, 800, 1200$. For CDP dataset, we choose $L = 3, 6, 12$ and $n = 128, 256, 512$. The step size is chosen to be $\mu_t = \min\{\mu, \exp(-t/\tau_0)\}$, where τ is selected from the set $\{0.02, 0.05, 0.1, 0.2, 0.3\}$ for comparison and $\tau_0 = 330$. The number of power iterations in initialization is selected to be 50 as the same as [1]. The numerical results are shown as below. Table 1 and 2 give the relative error in ℓ_2 norm with the primitive data after 3000 iterations. Figures 1-18 display the iterative curves on different datasets and parameters.

Observation of the results tells that if the dataset is too small, i.e., m is small, the iteration can converge to different restoration of the primitive data, which can be interpreted by the over-determined system of MSE problem. [1] has proved that Wirtinger Flow converges with high probability when $m \geq c_0 n \log n$, where c_0 is a constant. On the other hand, $z \to |Az|$ is injective when $m = 4n$. Such theorems suggest that large m is required to gurantee the restoration of primitive data. However, it is helpful for restoration by setting the stepsize smaller.Secondly, given that the iterates finally converge to the primitive data, larger step size results in faster convergence. Anyway, our implementation of Wirtinger Flow attains very good results both in speed and accuracy.

References

[1] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, 2015.

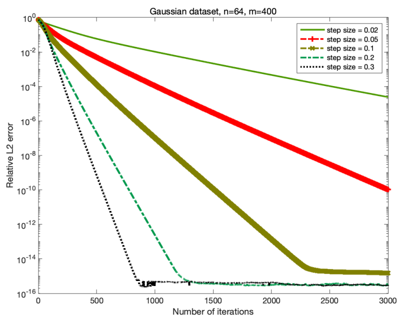

Figure 1: $m = 400$

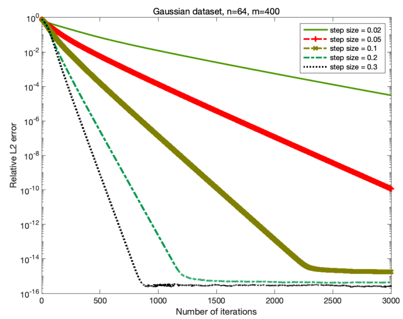

Figure 2: $m = 800$

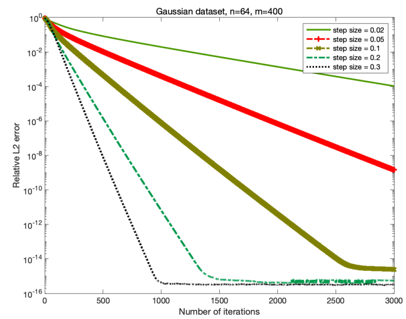

Figure 3: $m = 1200$

Results on Gaussian dataset when $n = 64$

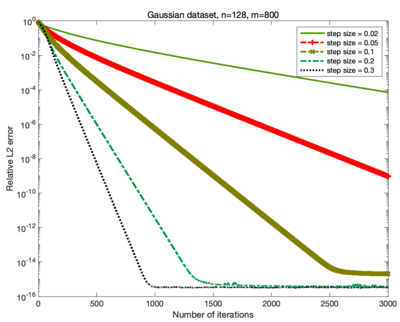

Figure 4: $m = 400$

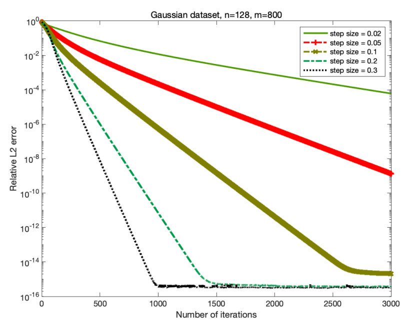

Figure 5: $m = 800$

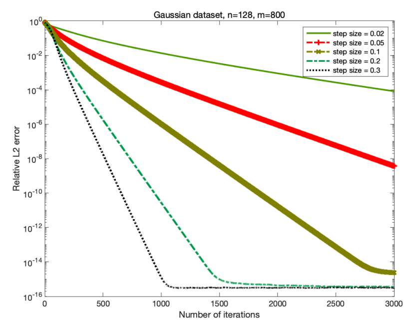

Figure 6: $m = 1200$

Results on Gaussian dataset when $n = 128$

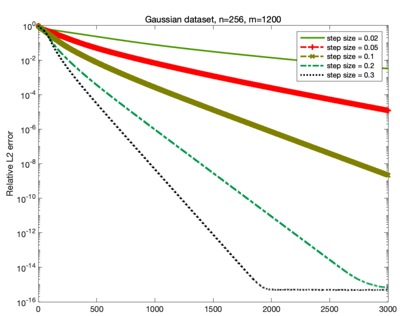

Figure 7: $m = 400$

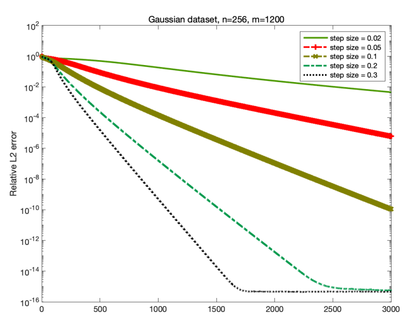

Figure 8: $m = 800$

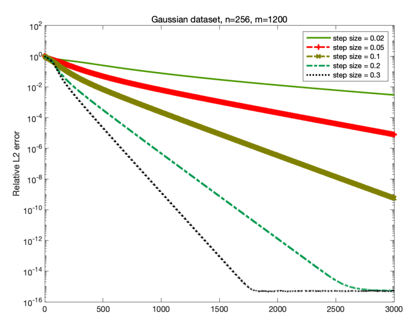

Figure 9: $m = 1200$

Results on Gaussian dataset when $n = 256$

μ	$n = 128$				
	0.02	0.05	0.1	0.2	0.3
$L = 3$	$4.33e - 01$	$9.37e - 02$	$5.57e - 03$	$5.65e - 05$	$7.22e - 07$
$L = 6$	$6.09e - 01$	$5.52e - 01$	$3.87e - 01$	$1.48e - 03$	$5.11e - 02$
$L = 12$	$1.89e - 01$	$1.42e - 02$	$7.40e - 04$	$4.44e - 06$	$1.22e - 01$
μ	$n = 256$				
	0.02	0.05	0.1	0.2	0.3
$L = 3$	$2.04e - 04$	$7.76e - 09$	$2.88e - 15$	$5.29e - 16$	$2.87e - 16$
$L = 6$	$1.43e - 04$	$6.87e - 09$	$2.83e - 15$	$3.98e - 16$	$2.88e - 16$
$L = 12$	$1.22e - 04$	$5.21e - 09$	$2.71e - 15$	$3.21e - 16$	$2.77e - 16$
μ	$n = 512$				
	0.02	0.05	0.1	0.2	0.3
$L = 3$	$3.14e - 08$	$3.25e - 15$	$1.09e - 15$	$3.47e - 16$	$2.08e - 16$
$L = 6$	$2.33e - 08$	$3.12e - 15$	$1.04e - 15$	$3.20e - 16$	$2.28e - 16$
$L = 12$	$4.56e - 08$	$3.36e - 15$	$1.16e - 15$	$3.28e - 16$	$2.30e - 16$

Table 2: Relative error after 3000 iterations by Wirtinger Flow on dataset CDP

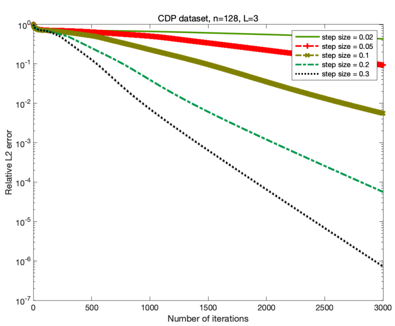

Figure 10: $L = 3$

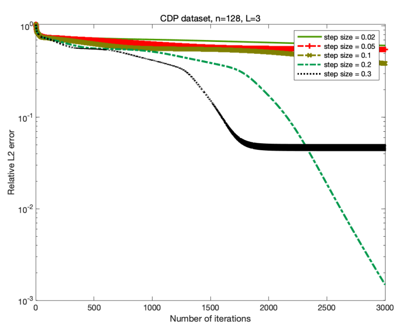

Figure 11: $L = 6$

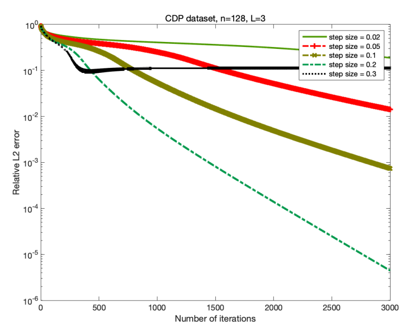

Figure 12: $L = 12$

Results on CDP dataset when $n = 128$

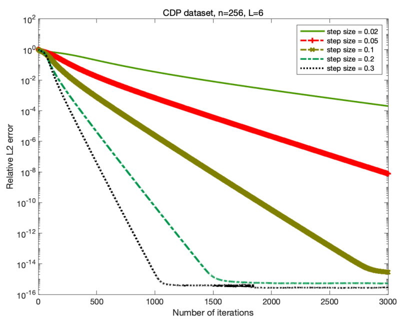

Figure 13: $L = 3$

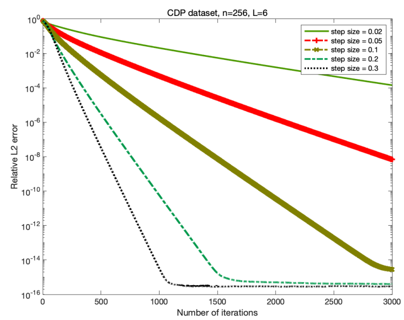

Figure 14: $L = 6$

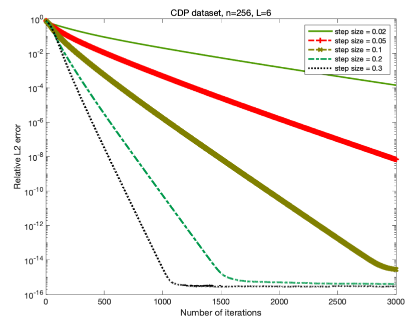

Figure 15: $L = 12$

Results on CDP dataset when $n = 256$

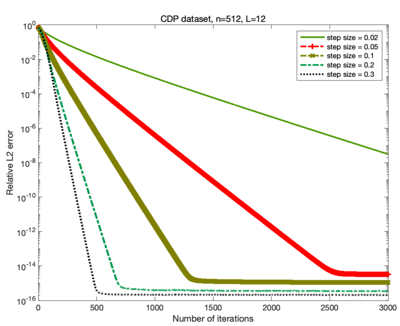

Figure 16: $L = 3$

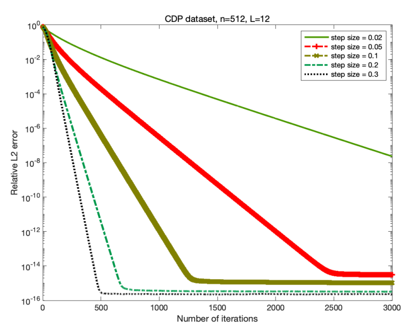

Figure 17: $L = 6$

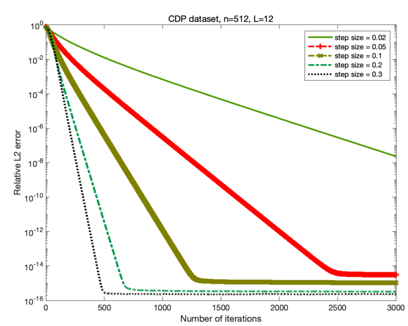

Figure 18: $L = 12$

Results on CDP dataset when $n = 512$