Performance Evaluation of FISTA with Constant Inertial Parameter

Kaito Kameda*, Ryo Hayakawa*, Kazunori Hayashi†, and Youji Iiguni*
* Osaka University, Osaka, Japan
E-mail: kameda@sip.sys.es.osaka-u.ac.jp
E-mail: rhayakawa@sys.es.osaka-u.ac.jp
E-mail: iiguni@sys.es.osaka-u.ac.jp
† Kyoto University, Kyoto, Japan
E-mail:hayashi.kazunori.4w@kyoto-u.ac.jp

***Abstract*—Compressed sensing is a framework for reconstructing sparse signals from their underdetermined linear measurements. In this paper, we discuss compressed sensing algorithms appropriate for the implementation with optical circuits, which have attracted attention as a technology that can perform vector-matrix products at high speed with low power consumption. Although the fast iterative shrinkage thresholding algorithm (FISTA) is one of the basic compressed sensing algorithms, updating the inertial parameters is difficult to implement in optical circuits. We thus propose an algorithm by replacing the inertial parameter of FISTA with a constant and evaluate its performance via computer simulations. Simulation results show that the proposed algorithm with an appropriate inertial parameter can achieve comparable performance to the original FISTA.**

I. Introduction

Compressed sensing [1] is a framework for recovering sparse signals from a small amount of observed data. Since various signals in engineering have the sparsity, compressed sensing has many applications in image processing, wireless communications, control engineering, and other fields [2][3].

One of the basic compressed sensing algorithms is the iterative shrinkage thresholding algorithm (ISTA) [4]. ISTA is an iterative algorithm based on convex optimization and it iteratively updates the estimate of the sparse unknown vector to obtain a final estimate. The fast iterative shrinkage thresholding algorithm (FISTA) [5] has been proposed as an accelerated version of ISTA. FISTA reduces the number of iterations for the convergence compared to ISTA by introducing an inertial parameter in the update equations and updating the estimate based on the last two estimates.

One way to reduce latency and power consumption of reconstructions in compressed sensing is the use of faster and more power-efficient computers. However, since the performance of the complementary metal oxide semiconductor (CMOS) computer is reaching the limit [6][7], we need to take another approach to achieve much lower latency and power consumption than current CMOS-based computers.

This work was supported in part by JST, CREST, Grant Number JPMJCR21C3, Japan and JSPS KAKENHI Grant Number JP20K23324.

For the fast and efficient computation, analog computing with optical circuits [6][7] have attracted much attention in photonics. Unlike conventional electronic circuits, optical circuits use photons instead of electrons as information carriers. Optical circuits can thus perform some operations such as vector-matrix products with lower latency and power consumption than conventional electronic circuits. However, since there are limitations on the operations that can be performed with optical circuits, we cannot necessarily implement all algorithms with optical circuits.

In this paper, we discuss compressed sensing algorithms appropriate for optical circuits to achieve low latency and power consumption of the reconstruction. Since the original FISTA requires the division by a dynamic variables to update the inertial parameter, it is difficult to implement in optical circuits. We thus propose FISTA with constant inertial parameters, reffered to as constant inertial FISTA (CIFISTA), as a compressed sensing algorithm suitable for the implementation in optical circuits. To determine the possible range of the constant inertial parameter, we obtain the range of the inertial parameters in the original FISTA. Moreover, we evaluate the performance of CIFISTA by computer simulations. The simulation results show that the value of the appropriate constant inertial parameter depends on the problem setup and other parameters of the algorithm. Furthermore, the results also show that the proposed CIFISTA can achieve similar performance to the original FISTA when we choose an appropriate value of the constant inertial parameter.

II. Compressed Sensing via FISTA

A. Compressed Sensing

The basic problem setup for compressed sensing is to estimate a sparse unknown vector $\boldsymbol{x}^* \in \mathbb{R}^N$ from a known linear measurement vector

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{e} \in \mathbb{R}^M \ (N \geq M), \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{M \times N}$ is a known measurement matrix, and $\boldsymbol{e} \in \mathbb{R}^M$ is the measurement noise.

The estimation accuracy of the unknown vector depends on the relationship between N and M. In the absence of

Algorithm 1 ISTA

Require: $\gamma > 0$, $\boldsymbol{x}_0 \in \mathbb{R}^N$.
  $t = 0$
  repeat
    $\boldsymbol{u}_{t+1} = \boldsymbol{x}_t - \gamma \boldsymbol{A}^\top (\boldsymbol{A}\boldsymbol{x}_t - \boldsymbol{y})$
    $\boldsymbol{x}_{t+1} = S_{\gamma\lambda}(\boldsymbol{u}_{t+1})$
    $t = t + 1$
  until convergence;

measurement noise, i.e., $\boldsymbol{e} = \boldsymbol{0}$, a necessary and sufficient condition for the perfect estimation is $\mathrm{rank}\, \boldsymbol{A} = N$ in general. On the other hand, when $M < N$, that is, when the number of variables N is larger than the number of measurements M, the solution cannot be uniquely determined and it is difficult to correctly identify $\boldsymbol{x}^*$. However, taking advantage of the fact that the unknown vector $\boldsymbol{x}^*$ is sparse, it may be possible to completely reconstruct the unknown vector $\boldsymbol{x}^*$ from $\boldsymbol{y}$ even when $M < N$ [1].

B. ℓ_1-ℓ_2 reconstruction

One way to estimate the sparse unknown vector $\boldsymbol{x}^*$ from the noisy measurement $\boldsymbol{y}$ in (1) is to compute the solution of the ℓ_1-ℓ_2 optimization problem

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^N}{\arg\min} \left\{ \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \lambda \|\boldsymbol{x}\|_1 \right\} \tag{2}$$

as the estimate. Here,

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{N} |x_i| \tag{3}$$

and

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{N} x_i^2} \tag{4}$$

are the ℓ_1 norm and the ℓ_2 norm of the vector $\boldsymbol{x} = [x_1 \cdots x_N]^\top$, respectively, where $(\cdot)^\top$ denotes the transpose. The first term $\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2$ of the objective function of (2) is the objective function used in the least squares method. The optimization problem (2) can be interpreted as a sparse estimator while minimizing the difference between $\boldsymbol{y}$ and $\boldsymbol{A}\boldsymbol{x}$. The parameter λ (> 0) determines how much the sparsity of the estimate is emphasized. When λ is large, the optimization problem gives priority to the sparsity of the estimate over the smallness of the difference between $\boldsymbol{y}$ and $\boldsymbol{A}\boldsymbol{x}$. Conversely, when λ is small, priority is given to reducing the difference between $\boldsymbol{y}$ and $\boldsymbol{A}\boldsymbol{x}$ rather than to making the estimate sparse.

C. ISTA

ISTA is an algorithm for the optimization problem (2), and is shown in Algorithm 1. We define the soft thresholding function $S_\alpha(v)$ as

$$S_\alpha(v) \triangleq \begin{cases} v - \alpha & (v \geq \alpha) \\ 0 & (-\alpha < v < \alpha), \\ v + \alpha & (v \leq -\alpha) \end{cases} \tag{5}$$

Algorithm 2 FISTA

Require: $\gamma > 0$, $\boldsymbol{z}_0 = \boldsymbol{x}_0 \in \mathbb{R}^N$.
  $t = 0$, $s_0 = 1$
  repeat
    $\boldsymbol{u}_{t+1} = \boldsymbol{z}_t - \gamma \boldsymbol{A}^\top (\boldsymbol{A}\boldsymbol{z}_t - \boldsymbol{y})$
    $\boldsymbol{x}_{t+1} = S_{\gamma\lambda}(\boldsymbol{u}_{t+1})$
    $s_{t+1} = \frac{1+\sqrt{1+4s_t^2}}{2}$
    $a_{t+1} = \frac{s_t - 1}{s_{t+1}}$
    $\boldsymbol{z}_{t+1} = \boldsymbol{x}_{t+1} + a_{t+1}(\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)$
    $t = t + 1$
  until convergence;

where $\alpha > 0$ and $v \in \mathbb{R}$. When the input of the soft thresholding function is a vector, it is defined as the element-wise operator.

As shown in [9], ISTA converges when γ satisfies

$$0 < \gamma \leq \frac{1}{\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})}, \tag{6}$$

where $\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})$ is the maximum absolute value of the eigenvalues of the matrix $\boldsymbol{A}^\top \boldsymbol{A}$, i.e., spectral radius of $\boldsymbol{A}^\top \boldsymbol{A}$.

Let $f(\boldsymbol{x})$ be the objective function of the optimization problem (2). The error $f(\boldsymbol{x}_t) - f(\hat{\boldsymbol{x}})$ at step t of ISTA satisfies

$$f(\boldsymbol{x}_t) - f(\hat{\boldsymbol{x}}) \leq \frac{\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}\|_2^2}{2t}, \tag{7}$$

which means that the order of the convergence of the objective function is $O(1/t)$ [5].

D. FISTA

FISTA is a fast algorithm based on ISTA. The update equations of FISTA is shown in Algorithm 2.

While ISTA computes $\boldsymbol{u}_{t+1}$ using $\boldsymbol{x}_{t+1}$, FISTA computes $\boldsymbol{u}_t$ on the basis of $\boldsymbol{z}_t$ by introducing an acceleration formula $\boldsymbol{z}_{t+1} = \boldsymbol{x}_{t+1} + a_{t+1}(\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)$ with inertial parameter a_{t+1}. This acceleration enables us to use the last two estimates $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t-1}$ for the calculation of $\boldsymbol{x}_{t+1}$. By this acceleration, the error of the objective function $f(\boldsymbol{x}_t) - f(\hat{\boldsymbol{x}})$ at the t-step of FISTA satisfies

$$f(\boldsymbol{x}_t) - f(\hat{\boldsymbol{x}}) \leq \frac{2\lambda_{\max}(\boldsymbol{A}^\top \boldsymbol{A})\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}\|_2^2}{(t+1)^2}. \tag{8}$$

This means that the order of the convergence of FISTA is $O(1/t^2)$, which is faster than that of ISTA [5].

III. FISTA WITH CONSTANT INERTIAL PARAMETER

In this section, we describe FISTA in the complex domain and the proposed CIFISTA, which is more appropriate for optical circuits.

Algorithm 3 FISTA in the complex domain

Require: $\gamma > 0$, $\boldsymbol{z}_0 = \boldsymbol{x}_0 \in \mathbb{C}^N$.
$t = 0$, $s_0 = 1$
repeat
$\boldsymbol{u}_{t+1} = \boldsymbol{z}_t - \gamma \boldsymbol{A}^{\mathsf{H}}(\boldsymbol{A}\boldsymbol{z}_t - \boldsymbol{y})$
$\boldsymbol{x}_{t+1} = T_{\gamma\lambda}(\boldsymbol{u}_{t+1})$
$s_{t+1} = \frac{1+\sqrt{1+4s_t^2}}{2}$
$a_{t+1} = \frac{s_t - 1}{s_{t+1}}$
$\boldsymbol{z}_{t+1} = \boldsymbol{x}_{t+1} + a_{t+1}(\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)$
$t = t + 1$
until convergence;

A. FISTA in the Complex Domain

Since optical circuits can handle complex numbers, in this and the subsequent sections, we consider the compressed sensing problem in the complex domain. Especially, we estimate a complex sparse vector $\boldsymbol{x}^* \in \mathbb{C}^N (N \geq M)$ from a known measurement vector

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{e} \in \mathbb{C}^M, \tag{9}$$

where $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ is the measurement matrix and $\boldsymbol{e} \in \mathbb{C}^M$ is the measurement noise.

For the extension of FISTA to the complex domain, we need to define a soft thresholding function for complex numbers. In this paper, the soft thresholding function for a complex number is defined as

$$T_\alpha(v) \triangleq \begin{cases} (|v| - \alpha)\dfrac{v}{|v|} & (|v| \geq \alpha) \\ 0 & (|v| < \alpha) \end{cases} \tag{10}$$

as in [10], where $\alpha > 0$ and $v \in \mathbb{C}$. The algorithm of FISTA in the complex domain with the soft thresholding function (10) is shown in Algorithm 3, where $(\cdot)^{\mathsf{H}}$ denotes the complex conjugate transpose.

B. Proposed Constant Inertial FISTA

One of the problems in executing FISTA with optical circuits is the difficulty of the update of a_{t+1}, which requires division by the dynamic variable in s_{t+1}. Therefore, the acceleration formula of $\boldsymbol{z}_{t+1}$ required for faster convergence than ISTA is difficult to implement in optical circuits. We thus replace the inertial parameter a_{t+1} of the original FISTA with a constant c and propose the resultant algorithm with the pseudo acceleration formula as an approach suitable for optical circuits.

The following holds for the inertial parameter a_{t+1} in the original FISTA.

Lemma 1. For any t, we have

$$0 \leq a_{t+1} < 1 \tag{11}$$

and

$$\lim_{t\to\infty} a_{t+1} = 1. \tag{12}$$

Proof: From $s_{t+1} = \frac{1+\sqrt{1+4s_t^2}}{2}$, we have

$$s_t < \frac{1+\sqrt{1+4s_t^2}}{2} = s_{t+1}. \tag{13}$$

From this and $s_0 = 1$, we have

$$1 \leq s_t < s_{t+1}. \tag{14}$$

Subtracting 1 from (14), we have

$$0 \leq s_t - 1 < s_{t+1} - 1 < s_{t+1}. \tag{15}$$

Dividing (15) by s_{t+1} yields

$$0 \leq \frac{s_t - 1}{s_{t+1}} < 1. \tag{16}$$

From $a_{t+1} = \frac{s_t - 1}{s_{t+1}}$ and (16), we obtain

$$0 \leq a_{t+1} < 1. \tag{17}$$

Next, to show

$$\lim_{t\to\infty} s_t = \infty, \tag{18}$$

we use mathematical induction on t to prove that

$$\frac{\sqrt{t}}{2} < s_t \tag{19}$$

holds for any natural number t. It is clear that (19) holds when $t = 0$. Assuming that it holds for $t = k$, it also holds for $t = k + 1$ since

$$s_{k+1} = \frac{1+\sqrt{1+4s_k^2}}{2} > \frac{1+\sqrt{1+k}}{2} > \frac{\sqrt{1+k}}{2}. \tag{20}$$

Therefore, (19) holds for any natural number t. Furthermore, since

$$\lim_{t\to\infty} \frac{\sqrt{t}}{2} = \infty, \tag{21}$$

we have

$$\lim_{t\to\infty} s_t = \infty. \tag{22}$$

Therefore,

$$\begin{aligned} \lim_{t\to\infty} a_{t+1} &= \lim_{t\to\infty} \frac{s_t - 1}{s_{t+1}} \\ &= \lim_{t\to\infty} \frac{s_t - 1}{\frac{1+\sqrt{1+4s_t^2}}{2}} \\ &= \lim_{t\to\infty} \frac{(s_t - 1)(\sqrt{1+4s_t^2} - 1)}{2s_t^2} \\ &= \lim_{t\to\infty} \left(1 - \frac{1}{s_t}\right)\left(\sqrt{\frac{1}{4s_t^2} + 1} - \frac{1}{2s_t}\right) \\ &= 1 \end{aligned} \tag{23}$$

holds. ■

Lemma 1 shows that the range of the inertial parameter a_{t+1} is $0 \leq a_{t+1} < 1$ in the original FISTA. Fig. 1 shows the transition of the inertial parameter a_{t+1}. From Fig. 1, we can see that the inertial parameter a_{t+1} actually increases monotonically in the range from 0 to 1. From the above

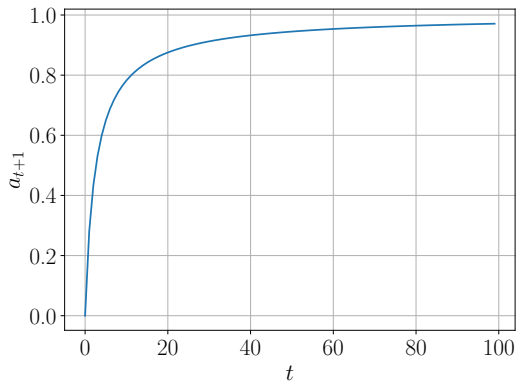


Fig. 1: Inertial parameters of the original FISTA

Algorithm 4 CIFISTA

Require: $\gamma > 0$, $0 \le c \le 1$, $\boldsymbol{z}_0 = \boldsymbol{x}_0 \in \mathbb{C}^N$.
  $t = 0$
  repeat
    $\boldsymbol{u}_{t+1} = \boldsymbol{z}_t - \gamma \boldsymbol{A}^{\mathsf{H}}(\boldsymbol{A}\boldsymbol{z}_t - \boldsymbol{y})$
    $\boldsymbol{x}_{t+1} = T_{\gamma\lambda}(\boldsymbol{u}_{t+1})$
    $\boldsymbol{z}_{t+1} = \boldsymbol{x}_{t+1} + c(\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)$
    $t = t + 1$
  until convergence;

discussion, we set the range of the constant inertial parameter c to $0 \le c \le 1$ in this paper.

The proposed algorithm named constant inertial FISTA (CIFISTA) is shown in Algorithm 4. Note that CIFISTA is equivalent to ISTA in the complex domain when $c = 0$.

IV. Computer Simulations

In this section, we evaluate the number of iterations for the convergence of the proposed CIFISTA and the optimal value of c for various situations by computer simulations.

A. Simulation Settings

The nonzero components of the complex sparse unknown vectors $\boldsymbol{x}^* \in \mathbb{C}^N$ are assumed to follow an independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian distribution with mean 0 and variance 1. The number of nonzero components of $\boldsymbol{x}^*$ is K ($K \le M$). Each element of noise vector $\boldsymbol{e} \in \mathbb{C}^M$ is assumed to be an i.i.d. circularly symmetric complex Gaussian noise with mean 0 and variance $\sigma_{\boldsymbol{e}}^2$. Let $\sigma_{\boldsymbol{x}^*}^2$ be the variance of each component of $\boldsymbol{x}^*$, and define the signal to noise ratio (SNR) as

$$10\log_{10}\frac{\sigma_{\boldsymbol{x}^*}^2}{\sigma_{\boldsymbol{e}}^2} = 10\log_{10}\frac{K}{N\sigma_{\boldsymbol{e}}^2}. \tag{24}$$

The initial values of FISTA and CIFISTA are $\boldsymbol{x}_0 = \boldsymbol{0}$ and $\boldsymbol{z}_0 = \boldsymbol{0}$. In the simulations, we have judged the convergence is achieved when

$$\frac{1}{N}\|\boldsymbol{x}_{t-i} - \boldsymbol{x}_{t-1-i}\|_2^2 < 10^{-14},\ i = 0, 1, 2, 3 \tag{25}$$

is satisfied.

B. Simulation Results for Partial DFT Measurement Matrix

Since the optical circuit is suitable for computing the product of a discrete Fourier transform (DFT) matrix and a vector, this section evaluates the performance when a partial DFT matrix is used as the measurement matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$. The parameter γ must satisfy

$$0 < \gamma \le \frac{1}{\lambda_{\max}(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})} \tag{26}$$

for the convergence of FISTA in the complex domain. Since $\boldsymbol{A}$ is a partial DFT matrix, $\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}$ is the identity matrix and its eigenvalues are all 1. From the fact that the non-zero eigenvalues of $\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}$ coincide with those of $\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}$, the eigenvalues of $\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}$ are 0 or 1. Therefore, $\lambda_{\max}(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}) = 1$ and hence the original FISTA converges when $0 < \gamma \le 1$.

In the simulations, we consider different settings in Table I. In settings 2–7, one variable is changed from setting 1 as follows:

- setting 2: $N = 250$
- setting 3: $M = 400$
- setting 4: $K = 50$
- setting 5: $\lambda = 0.008$
- setting 6: $\gamma = 0.099$
- setting 7: SNR is 25 dB.

In this section, we first evaluate the performance of CIFISTA for various values of the constant inertial parameter c. We then obtain the best value of c that minimizes the number of iterations for the convergence for various sparsity ratios and measurement ratios.

1) Performance for Various Values of c: We have compared the performance of CIFISTA for several values of the constant inertial parameter c. In Fig. 2, we show the mean squared error (MSE)

$$\frac{1}{N}\|\boldsymbol{x}_t - \boldsymbol{x}^*\|_2^2 \tag{27}$$

at each iteration of CIFISTA and FISTA for setting 1. Fig. 2 shows that the number of iterations for the convergence depends on the value of the constant inertial parameter c. We can also see that the MSE at the convergence of CIFISTA is the same as that of FISTA.

For setting 1, $c_{\text{best}} = 0.64599$ is the best value of c in terms of the number of iterations for the convergence. The MSE at each iteration with $c = c_{\text{best}} = 0.64599$ is shown in Fig. 3. The figure shows that CIFISTA can achieve almost the same performance as the original FISTA if the value of the constant inertial parameter c is appropriately determined.

2) Performance for Other Settings: To evaluate the performance of CIFISTA for different settings, we compute c_{best} for settings 2–7 via simulations. The results are shown in Table I. Table I shows that c_{best}'s for settings 2–7 are different from c_{best} for setting 1, indicating that c_{best} depends on the values of N, M, K, λ, γ, and SNR.

Fig. 4 shows the MSE at each iteration of CIFISTA with c_{best} and FISTA for settings 2–7. From the figure, we can see that CIFISTA can achieve almost the same performance as FISTA

TABLE I: Settings in simulation

	N	M	K	λ	γ	SNR [dB]	c_{best}
setting 1	500	200	25	0.020	0.990	15	0.64599
setting 2	250	200	25	0.020	0.990	15	0.54765
setting 3	500	400	25	0.020	0.990	15	0.45464
setting 4	500	200	50	0.020	0.990	15	0.75984
setting 5	500	200	25	0.008	0.990	15	0.80808
setting 6	500	200	25	0.020	0.099	15	0.87255
setting 7	500	200	25	0.020	0.990	25	0.53745

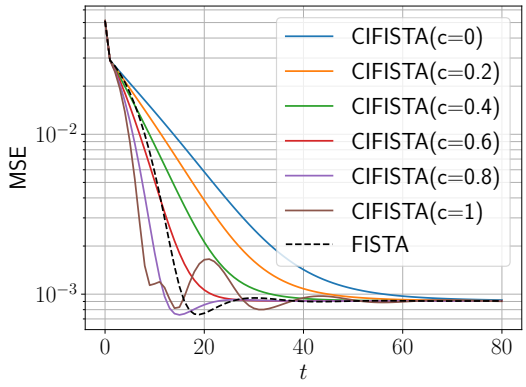


Fig. 2: MSE at each iteration of CIFISTA and FISTA for setting 1

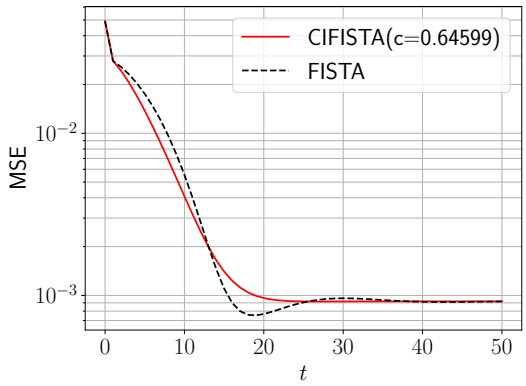


Fig. 3: MSE at each iteration of CIFISTA and FISTA with c_{best} of setting 1

if the value of the constant inertial parameter c is appropriately chosen according to the setting.

3) c_{best} for Various Sparsity Ratios and Measurement Ratios: We have evaluated c_{best} for various values of sparsity ratio, measurement ratio, and SNR. We set $\lambda = \lambda_{\text{best}}$ and $\gamma = 1/\lambda_{\max}(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}) \times 0.99 = 0.99$, where λ_{best} is the value that minimizes the MSE at convergence. The sparsity ratio is the fraction of non-zero components of the complex sparse unknown vector $\boldsymbol{x}^*$ and is defined as K/N. The measurement ratio is the ratio of the number of rows to the number of columns in the measurement matrix $\boldsymbol{A}$, defined as M/N [8]. In the simulations in this section, we set $N = 500$.

First, λ_{best} was obtained by simulation for various sparsity and measurement ratios. Fig. 5 shows the four results for SNRs of 10 dB, 15 dB, 20 dB, and 25 dB. From Fig. 5, we can see that λ_{best} increases as the SNR decreases, i.e., as the noise power increases. This is because the measurement $\boldsymbol{y}$ becomes unreliable as the noise power increases. In such a case, the data fidelity term $\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2$ in the objective function of the optimization problem (2) does not work well, and hence the estimation accuracy is better when more weight is given to the regularization term $\|\boldsymbol{x}\|_1$.

Fig. 6 shows the value of c_{best} for various sparsity ratios and measurement ratios obtained by simulation. From Fig. 6, c_{best} varies in a layered manner, and it can be seen that c_{best} becomes smaller as the measurement ratio value increases and the sparsity ratio value decreases. Furthermore, regardless of the SNR, c_{best} is almost the same when the sparsity ratio is small and the measurement ratio is large.

C. Simulation Results for i.i.d. Gaussian Measurement Matrix

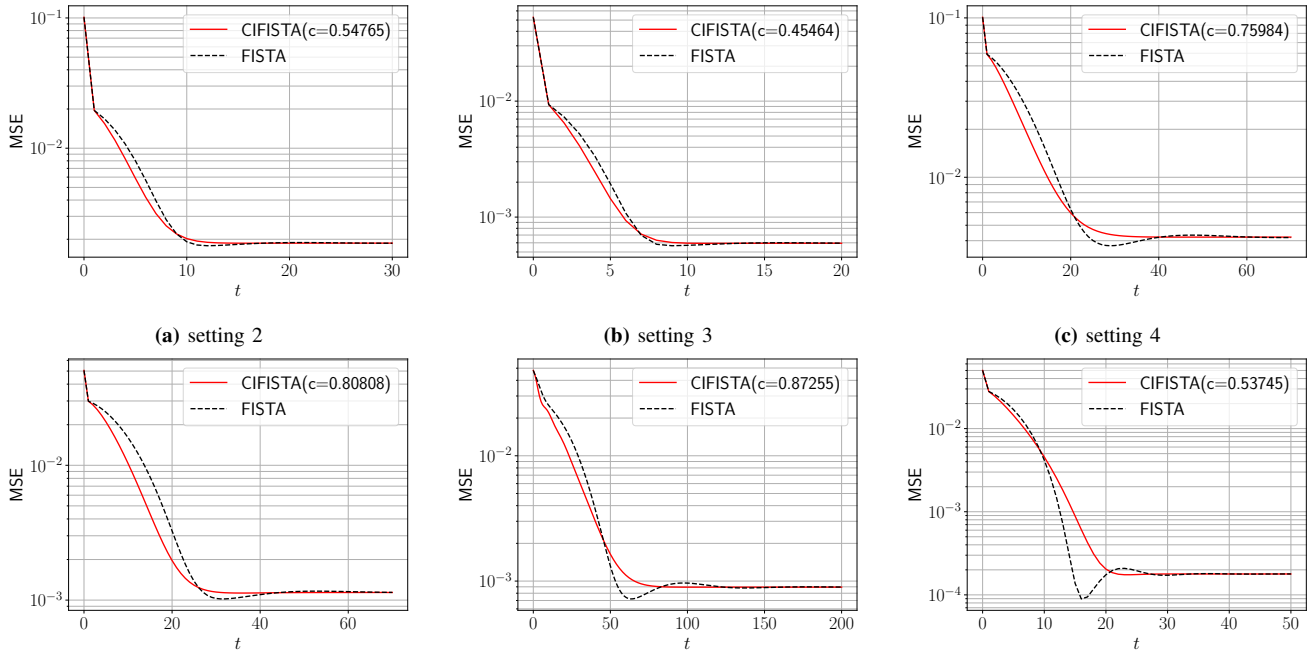
We evaluate the performance of CIFISTA for the i.i.d. Gaussian measurement matrix with mean 0 and variance $1/N$. The values of the parameters λ and γ are selected as $\lambda = \lambda_{\text{best}}$ and $\gamma = 1/\lambda_{\max}(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A}) \times 0.99$. c_{best} is obtained for various values of sparsity ratio, measurement ratio, and SNR. In the simulations in this section, we set $N = 500$.

Fig. 7 shows the value of λ_{best} for various sparsity ratios and measurement ratios at SNR of 10 dB, 15 dB, 20 dB, and 25 dB obtained by simulation. From Fig. 7, we can see that λ_{best} increases as the SNR decreases. Comparing with Fig. 5, we can see that the values are almost the same. In other words, λ_{best} is almost the same when the measurement matrix is the partial DFT matrix and the i.i.d. Gaussian matrix.

Fig. 8 shows the value of c_{best} for various sparsity ratios and measurement ratios at SNRs of 10 dB, 15 dB, 20 dB, and 25 dB obtained by simulation. The figure shows that c_{best} is almost the same when the sparsity ratio is small and the measurement ratio is large, regardless of the SNR. Compared to the case with the partial DFT matrix in Fig. 6, c_{best} is larger for the i.i.d. Gaussian measurement matrix.

V. Conclusions

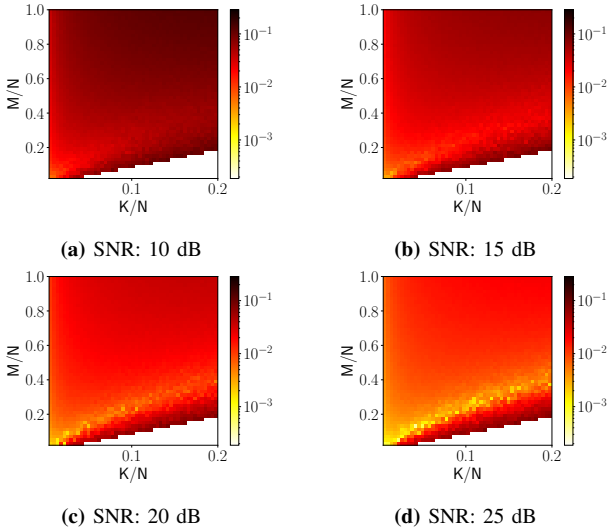
In this paper, we have proposed CIFISTA as a compressed sensing algorithm appropriate for the implementation in optical circuits. Unlike FISTA, which involves the division by dynamic variables to update inertial parameters, CIFISTA has constant inertial parameters and is therefore considered to be suitable for execution in optical circuits, where division by dynamic variables is hard to realize. Simulation results show that the appropriate value of the constant inertial parameter in CIFISTA depends on the value of the other parameters such



(a) setting 2 **(b)** setting 3 **(c)** setting 4

(d) setting 5 **(e)** setting 6 **(f)** setting 7

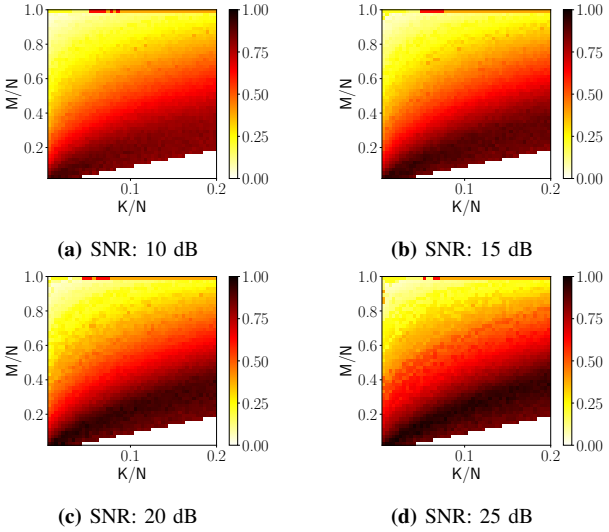
Fig. 4: MSE at each iteration of CIFISTA and FISTA



(a) SNR: 10 dB **(b)** SNR: 15 dB

(c) SNR: 20 dB **(d)** SNR: 25 dB

Fig. 5: λ_{best} for various sparsity ratios and measurement ratios ($\boldsymbol{A}$: partial DFT matrix)



(a) SNR: 10 dB **(b)** SNR: 15 dB

(c) SNR: 20 dB **(d)** SNR: 25 dB

Fig. 6: c_{best} for various sparsity ratios and measurement ratios ($\boldsymbol{A}$: partial DFT matrix)

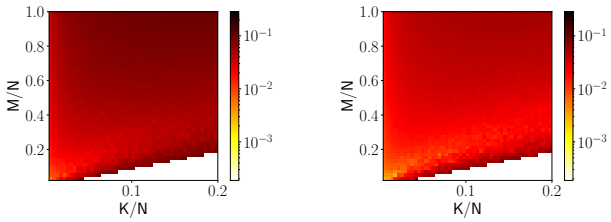
as the sparsity and the measurement ratio. When we use the appropriate value of the constant inertial parameter, CIFISTA can achieve almost the same performance as the original FISTA. From the results, we also observe that the appropriate value of the constant inertial parameter also depends on the structure of the measurement matrix $\boldsymbol{A}$. Future work includes the investigation of soft thresholding functions suitable for the computation in optical circuits, the theoretical evaluation of the order of convergence for CIFISTA, and the proper choice of the constant inertial parameter.

REFERENCES

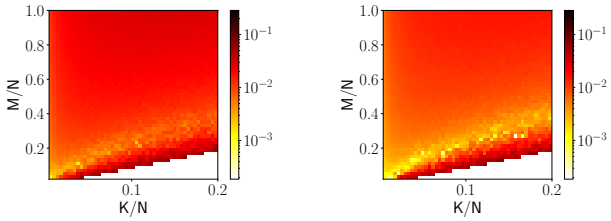
[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," *IEICE Transactions on Communications*, vol. E96–B, no. 3, pp. 685–712, Mar. 2013.

[3] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, Mar. 2008.

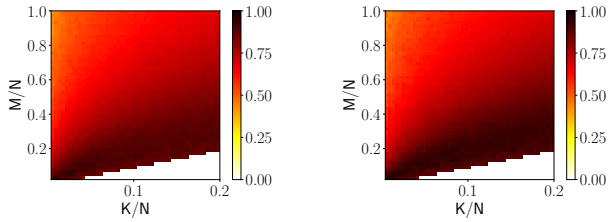


(a) SNR: 10 dB **(b)** SNR: 15 dB

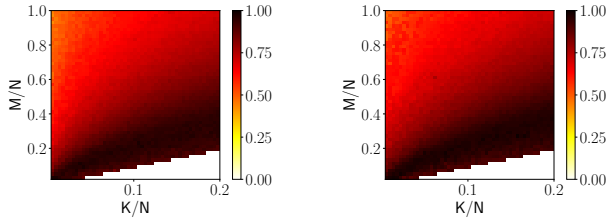


(c) SNR: 20 dB **(d)** SNR: 25 dB

Fig. 7: λ_{best} for various sparsity ratios and measurement ratios ($\boldsymbol{A}$: i.i.d. Gaussian matrix)



(a) SNR: 10 dB **(b)** SNR: 15 dB



(c) SNR: 20 dB **(d)** SNR: 25 dB

Fig. 8: c_{best} for various sparsity ratios and measurement ratios ($\boldsymbol{A}$: i.i.d. Gaussian matrix)

[4] I. Daubechies, M. Defrise, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no, 11, pp. 1413–1457, 2004.

[5] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 188–202, 2009.

[6] K. Kitayama, M. Notomi, M. Naruse, K. Inoue, S. Kawakami, and A. Uchida, "Novel frontier of photonics for data processing–Photonic accelerator," in *Proceedings of the International Conference on Rebooting Computing (ICRC)*, pp. 95–101, Dec. 2020.

[7] K. Nozaki, S. Matsuo, T. Fujii, K. Takeda, A. Shinya, E. Kuramochi, and M. Notomi, "Femto-farad optoelectronic integration demonstrating energy-saving signal conversion and nonlinear functions," *Nature Photonics*, vol. 13, pp. 454–459, 2019.

[8] Z. Yang and C. Zhang, "Sparsity-undersampling tradeoff of compressed sensing in the complex domain," in *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, May. 2011.

[9] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal-recovery ploblem," in *Convex optimization*, Cambrige University Press, 2010.

[10] M. V. W. Zibetti, E. S. Helou, R. R. Regatte, and G. T. Herman, "Monotone FISTA with variable acceleration for compressed sensing magnetic resonance imaging," *IEEE Transactions on Computational Imaging*, vol. 5, no. 1, pp. 109–119, Mar. 2019.