Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 21, No. 1, pp. 17–29, 2025

A STUDY ON PARAMETERIZED EM ALGORITHM FOR LINEAR IMAGE RECONSTRUCTION PROBLEM AND ITS CONVERGENT PROPERTY

Caifang Wang*, Xiaobin Xia and Dan Gao

Abstract: EM algorithm is a widely used reconstruction method for linear imaging problem. It can be derived from the minimization problem of Kullback-Leibler divergence. Considering the Kullback-Leibler divergence is sensitive to noise due to part of small boundary measurements, a parameter is introduced into Kullback-Leibler divergence and a novel iteration method is derived using Karush-Kuhn-Tucker condition. Consequently, the convergent property of the method is proved based on hidden data and two subproblems. Preliminary numerical experiments are taken to analyze convergent performances of the derived iterative method. Compared with the traditional EM algorithm, the parameterized algorithm can achieve a relatively better reconstruction effect and faster convergent speed by adjusting the parameter.

Key words: *nonlinear optimization, Kullback-Leibler divergence, iterative reconstruction method, convergent property*

Mathematics Subject Classification: *65K10, 15A29, 78M50*

---

1 Introduction

In the field of medical imaging, linear imaging problem is widely concerned because of its wide application, relatively simple imaging model and various reconstruction algorithms [19]. In the aspect of application, many well-known problems such as X-ray computed tomography [18, 22] and radiation therapy treatment planning [15], can be described by linear equations models and then converted to different optimization problems [2].

If the measured data consists with the Gaussian noise model, a series Landweber schemes can be derived based on different weighted least-square functionals. SART, ART, Cimmino's algorithm, DWE and CAV are typical Landweber type and block Landweber algorithms [5, 10]. The choice of relaxation coefficient has a significant influence on the convergence behavior of Landweber scheme. Whether it is a consistent problem or an inconsistent problem, the sufficient convergence conditions for sequential block-iterative version and simultaneous block-iterative version Landweber scheme can be achieved by adjusting the relaxation coefficient [10]. Then the necessary and sufficient convergence conditions for general Landweber schemes are first derived in [14]. Converting the linear reconstruction problem to convex feasibility problems, some general methods, such as projection onto convex set

*Corresponding Author. The authors are partially supported by Shanghai University Science and Technology Innovation Action Plan Local University Capacity Building Project (23010502100).

 DOI: https://doi.org/10.61208/pjo-2023-023

(POCS) [12], subgradient projection algorithms [21] and Richardson-Lucy method [3] can also be used to solve related problems.

Some statistical algorithms are used in linear image reconstruction. The maximum likelihood (ML) estimation criterion, which solves the problem of maximum likelihood estimation of incomplete data is proposed to solve the linear imaging problem [16]. Assuming that the measured data consists with the Poisson noise model, the expectation maximization (EM) algorithm is developed in [17]. Furthermore, the ML-EM method is used to estimate the attenuation coefficient by combining the ML method with the EM algorithm [4,11]. In order to accelerate the convergent speed of the EM algorithm, some scholars divide the data into different ordered subsets (OS) and propose different OS-EM algorithms [8]. The OS-EM is extended into an infinite dimensional space in [6]. In addition, some regularization methods are used to overcome the ill posed reconstruction problem with Poisson data [7].

In fact, the EM algorithm for linear systems can be obtained by optimizing the Kullback-Leibler divergence (KL distance, or cross-entropy) between the measurement data and the forward problem data. The KL divergence is first introduced by Kullback and Leibler in [13]. Then the convergence properties of EM algorithm by optimizing KL divergence are proved by the authors [1, 9]. Considering that the KL divergence is particularly sensitive where the measurement data is very small, we introduce a parameter λ in KL divergence and measure the difference between the measurements and the forward problem data with the λ-KL divergence in this paper. We then derive an iterative formula for the linear system using Karush-Kuhn-Tucker condition of the objective function. By introducing hidden data, we transform the original problem into two optimization subproblems and obtain the convergence results of the iterative algorithm. In order to further verify the convergence performance of the algorithm, we simulate the scanning mode of fan beam CT and generate simulation measurement data of Shepp-Logan phantom. The correctness and validity of the iterative formula are verified by these data.

The rest of the paper is organized as follows. In Section 2, we make some basic assumptions about the linear system and introduce the concept and properties of λ-KL divergence. We then derive an iterative method based on the λ-KL divergence and provide its convergence property in Section 3 and Section 4, respectively. In Section 5, we consider numerical test to validate the performances of the iteration method. In Section 6, we present the conclusions.

2 Preliminaries

Throughout this paper, we mainly investigate a parameterized iterative reconstruction method for imaging problem

$$Ax = b, \tag{2.1}$$

where $A \in \mathbb{R}^{m\times n}$ is the imaging matrix, $b \in \mathbb{R}^m$ is the detected data and $x \in \mathbb{R}^n$ is the unknown image.

For convenience of description, we recall some notations and the definition for KL divergence first. Let $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m \big|\, x_i \geq 0, i = 1, 2, \ldots, m\}$, $\mathbb{R}^m_{++} = \{x \in \mathbb{R}^m \big|\, x_i > 0, i = 1, 2, \ldots, m\}$. For $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_m)^{\mathrm{T}} \in \mathbb{R}^m_{++}$ and $\beta = (\beta_1, \beta_2, \cdots, \beta_m)^{\mathrm{T}} \in \mathbb{R}^m_{++}$, the KL divergence is defined as [9]

$$KL(\alpha, \beta) = \sum_{i=1}^{m} \left(\alpha_i \log \frac{\alpha_i}{\beta_i} + \beta_i - \alpha_i \right). \tag{2.2}$$

Let $\lambda \in (0,1)$ be a constant. Similarly to λ-divergence [20], we define a new measure of divergence as

$$\begin{aligned} d_\lambda(\alpha,\beta) &= \lambda KL(\alpha,\lambda\alpha+(1-\lambda)\beta)+(1-\lambda)KL(\beta,\lambda\alpha+(1-\lambda)\beta) \\ &= \sum_{i=1}^{m}\left[\lambda\alpha_i\log\frac{\alpha_i}{\lambda\alpha_i+(1-\lambda)\beta_i}+(1-\lambda)\beta_i\log\frac{\beta_i}{\lambda\alpha_i+(1-\lambda)\beta_i}\right]. \end{aligned} \tag{2.3}$$

The same definitions and notations can be applied to $\mathbb{R}^n_{++}$ analogously.

3 Iteration formula for linear mathematical problem

In the following, we first make some assumptions for the system matrix and measurements. Then we provide an iteration method for the image reconstruction problem. The properties of the iteration operator and sequence are placed at the end of this section.

Assume that $A \in \mathbb{R}^{m\times n}_+$ satisfies $\sum_{j=1}^n a_{ij} > 0$ for all $i = 1,2,\dots,m$, and $\sum_{i=1}^m a_{ij} = 1$ for all $j = 1,2,\cdots,n$. $b \in \mathbb{R}^m_{++}$ satisfies $\sum_{i=1}^m b_i = 1$. These assumptions can be simply achieved by the method of substituting variables.

According to the property of KL divergence, the imaging problem can be solved by minimizing the function

$$F(x) = \begin{cases} d_\lambda(b, Ax), & x\in\Gamma, \\ +\infty, & \text{otherwise}, \end{cases} \tag{3.1}$$

where $\Gamma = \{x \in \mathbb{R}^n_+ \big| Ax \succ 0\}$. Then the optimization problem can be solved by the iteration method

$$\begin{cases} x^{(0)} \in \mathbb{R}^n_{++}, \\ x^{(k+1)} = P(x^{(k)}), \quad \text{for } \ k = 0,1,\dots. \end{cases} \tag{3.2}$$

Here the projection $P : \Gamma \to \Gamma$ is defined as

$$P(x)_j = x_j \prod_{i=1}^{m}\left[\lambda\frac{b_i}{\sum_{j=1}^n a_{ij}x_j}+(1-\lambda)\right]^{a_{ij}}, \quad \text{for } \ j=1,2,\dots,n. \tag{3.3}$$

For the projection P and the sequence $\{x^{(k)}\}$ generated by the above method, the following propositions can be obtained.

Proposition 3.1. $\{x\in\Gamma\big|P(x)=x\} \supseteq \{x\in\Gamma\big|x=\arg\min F(x)\}$.

Proof. It is easy to verify that Γ is a convex set. $d_\lambda(b,y)$ is convex with respect to y since $\varphi_\lambda(t)$ is strict convex. Composition with a linear transform $y = Ax$, $F(x)$ is also convex and it attains the minimum at those points in Γ. Then those points in the set of minimizers of F satisfy the first order optimality conditions for $F(x)$ such that $x \geq 0$. That means

$$\frac{\partial F}{\partial x_j} = (1-\lambda)\sum_{i=1}^m a_{ij}\log\frac{\sum_{j=1}^n a_{ij}x_j}{\lambda b_i+(1-\lambda)\sum_{j=1}^n a_{ij}x_j} = \mu_j, \tag{3.4}$$

$$x_j,\mu_j \geq 0, \tag{3.5}$$

$$x_j\mu_j = 0, \tag{3.6}$$

for $j = 1, 2, \ldots, n$. If $x_j \neq 0$, according to stationarity in Eq. (3.4), we obtain

$$\prod_{i=1}^{m} \left[\lambda \frac{b_i}{\sum_{j=1}^{n} a_{ij} x_j} + (1 - \lambda) \right]^{a_{ij}} = 1. \tag{3.7}$$

Therefore, whether x_j is zero or not, we always have

$$x_j \prod_{i=1}^{m} \left[\lambda \frac{b_i}{\sum_{j=1}^{n} a_{ij} x_j} + (1 - \lambda) \right]^{a_{ij}} = x_j. \tag{3.8}$$

Finally, we get the conclusion that the set of minimizer of $F(x)$ is contained in the set of fixed point of P. □

Proposition 3.2. *The nonnegative sequence $\{x^{(k)}\}$ generated by Eq. (3.2) is bounded by*

$$\sum_{j=1}^{n} x_j^{(k+1)} \leq [1 - (1 - \lambda)^{k+1}] \sum_{i=1}^{m} b_i + (1 - \lambda)^{k+1} \sum_{j=1}^{n} x_j^{(0)}. \tag{3.9}$$

In particular, if $\sum_{j=1}^{n} x_j^{(0)} = 1$, then $\sum_{j=1}^{n} x_j^{(k+1)} \leq 1$.

Proof. According to the iteration (3.2), we obtain

$$x_j^{(k+1)} = x_j^{(k)} \exp \left\{ \sum_{i=1}^{m} a_{ij} \log \left[\frac{\lambda b_i}{\sum_{s=1}^{n} a_{is} x_s^{(k)}} + (1 - \lambda) \right] \right\}. \tag{3.10}$$

Using the convexity of the exp function, we then have

$$x_j^{(k+1)} \leq x_j^{(k)} \left\{ \sum_{i=1}^{m} a_{ij} \left[\frac{\lambda b_i}{\sum_{s=1}^{n} a_{is} x_s^{(k)}} + (1 - \lambda) \right] \right\}. \tag{3.11}$$

Summing on j

$$\begin{aligned}
\sum_{j=1}^{n} x_j^{(k+1)} &\leq \sum_{j=1}^{n} x_j^{(k)} \left\{ \sum_{i=1}^{m} a_{ij} \left[\frac{\lambda b_i}{\sum_{s=1}^{n} a_{is} x_s^{(k)}} + (1 - \lambda) \right] \right\} \\
&= \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n} a_{ij} x_j^{(k)} \left[\frac{\lambda b_i}{\sum_{s=1}^{n} a_{is} x_s^{(k)}} + (1 - \lambda) \right] \right\} \\
&= \sum_{i=1}^{m} \left[\lambda b_i + (1 - \lambda) \sum_{j=1}^{n} a_{ij} x_j^{(k)} \right] \\
&= \lambda \sum_{i=1}^{m} b_i + (1 - \lambda) \sum_{j=1}^{n} x_j^{(k)}.
\end{aligned} \tag{3.12}$$

Using recursive expression above, we get

$$\begin{aligned}
\sum_{j=1}^{n} x_j^{(k+1)} &\leq \lambda \sum_{i=1}^{m} b_i + (1-\lambda) \left[\lambda \sum_{i=1}^{m} b_i + (1-\lambda) \sum_{j=1}^{n} x_j^{(k-1)} \right] \\
&\vdots \\
&\leq \left[1 + (1-\lambda) + \cdots + (1-\lambda)^k \right] \lambda \sum_{i=1}^{m} b_i + (1-\lambda)^{k+1} \sum_{j=1}^{n} x_j^{(0)} \\
&= [1 - (1-\lambda)^{k+1}] \sum_{i=1}^{m} b_i + (1-\lambda)^{k+1} \sum_{j=1}^{n} x_j^{(0)}. \qquad (3.13)
\end{aligned}$$

When $\sum_{j=1}^{n} x^{(0)} = 1$ and $\sum_{i=1}^{m} b_i = 1$, we have $\sum_{j=1}^{n} x^{(k+1)} \leq 1$. □

4 Hidden data and convergence analysis

In this section, we divide the optimization problem into two subproblems and then prove the convergent property of the iteration method.

Define an mn-dimensional hidden vector $q(x)$ with entries $(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$

$$q_{ij} = q_{ij}(x) = a_{ij} x_j. \qquad (4.1)$$

Meanwhile, let r be an mn-dimensional vector with entries

$$r_{ij} \geq 0, \quad \text{and} \quad \sum_{j=1}^{n} r_{ij} = b_i. \qquad (4.2)$$

Similar to $F(x)$, we introduce an alternating function with respect to hidden data as follows

$$d_\lambda(r, q(x)) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left[\lambda r_{ij} \log \frac{r_{ij}}{\lambda r_{ij} + (1-\lambda) q_{ij}} + (1-\lambda) q_{ij} \log \frac{q_{ij}}{\lambda r_{ij} + (1-\lambda) q_{ij}} \right]. \qquad (4.3)$$

The convergent property of the iteration method (3.2) can be obtained by solving the following two subproblems with some constrains.

1. Find minimizer of $d_\lambda(r, q(x^{(k)}))$ with respect to r;
2. Find minimizer of $d_\lambda(r(x^{(k)}), q(x))$ with respect to $q(x)$.

Lemma 4.1. *The minimizer of the problem $d_\lambda(r, q(x^{(k)}))$ with respect to r with the constraint $\sum_{j=1}^{n} r_{ij} = b_i$ $(i = 1, 2, \ldots, m)$ is*

$$r_{ij}(x^{(k)}) = \frac{b_i}{\sum_{j=1}^{n} q_{ij}(x^{(k)})} q_{ij}(x^{(k)}). \qquad (4.4)$$

Proof. The Lagrange function for the optimal problem is

$$L(r, \mu) = d_\lambda(r, q(x^{(k)})) + \sum_{i=1}^{m} \mu_i \left(\sum_{j=1}^{n} r_{ij} - b_i \right). \qquad (4.5)$$

According to stationarity condition

$$\frac{\partial L}{\partial r_{ij}} = \lambda \log \frac{r_{ij}}{\lambda r_{ij} + (1-\lambda) q_{ij}(x^{(k)})} + \mu_i = 0, \tag{4.6}$$

we have

$$r_{ij} = \mathrm{e}^{\frac{\mu_i}{-\lambda}} \left[\lambda r_{ij} + (1-\lambda) q_{ij}(x^{(k)}) \right]. \tag{4.7}$$

Summing over j and using the primal feasibility condition $\sum_{j=1}^n r_{ij} = b_i$, we then have

$$\mathrm{e}^{\frac{\mu_i}{-\lambda}} = \frac{b_i}{\lambda b_i + (1-\lambda) \sum_{j=1}^n q_{ij}(x^{(k)})}. \tag{4.8}$$

By substituting Eq. (4.8) for Eq. (4.7), we know that

$$r_{ij}(x^{(k)}) = \frac{b_i}{\sum_{j=1}^n q_{ij}(x^{(k)})} q_{ij}(x^{(k)}) \tag{4.9}$$

is the minimizer of the optimal problem. □

For simplicity, we use the following notations. Let

$$\begin{aligned} q_{ij}^k &= q_{ij}(x^{(k)}), && (4.10)\\ T_i^k &= \frac{b_i}{\sum_{j=1}^n q_{ij}(x_j^{(k)})} = \frac{b_i}{\sum_{j=1}^n q_{ij}^k}, && (4.11)\\ r_{ij}^{k+1} &= r_{ij}(x^{(k)}) = T_i^k q_{ij}^k, && (4.12)\\ S_j^k &= \prod_{i=1}^m \left[\lambda \frac{b_i}{\sum_{j=1}^n q_{ij}(x_j^{(k)})} + (1-\lambda) \right]^{a_{ij}} = \prod_{i=1}^m \left[\lambda T_i^k + (1-\lambda) \right]^{a_{ij}}. && (4.13) \end{aligned}$$

Then

$$x_j^{(k+1)} = S_j^k x_j^{(k)} \quad \text{and} \quad q_{ij}^{k+1} = S_j^k q_{ij}^k. \tag{4.14}$$

Next, we fix $r(x^{(k)})$ and consider the optimization problem

$$\min_{x\in\Gamma} \quad f(x) = d_\lambda(r(x^{(k)}), q(x)). \tag{4.15}$$

Lemma 4.2. $d_\lambda(r(x^{(k)}), q(x^{(k+1)})) \le d_\lambda(r(x^{(k)}), q(x^{(k)}))$.

Proof. Let $r_j^{k+1} = (r_{1j}^{k+1}, \cdots, r_{mj}^{k+1})^{\mathrm{T}}$ and $q_j^k = (q_{1j}^k, \cdots, q_{mj}^k)^{\mathrm{T}}$ for $j = 1, 2, \ldots, n$. Then

$$\begin{aligned} d_\lambda(r^{k+1}, q(x^{(k)})) &= \sum_{j=1}^n d_\lambda(r_j^{k+1}, q_j^k), && (4.16)\\ d_\lambda(r^{k+1}, q(x^{(k+1)})) &= \sum_{j=1}^n d_\lambda(r_j^{k+1}, S_j^k q_j^k). && (4.17) \end{aligned}$$

In order to get the conclusion of the lemma, we only have to prove that

$$d_\lambda(r_j^{k+1}, S_j^k q_j^k) \le d_\lambda(r_j^{k+1}, q_j^k). \tag{4.18}$$

We can easily show that $d_\lambda(r_j^{k+1}, tq_j^k)$ is convex with respect to t. Let t^* be the minimum point of $d_\lambda(r_j^{k+1}, tq_j^k)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} d_\lambda(r_j^{k+1}, tq_j^k)\Big|_{t=t^*} = \sum_{i=1}^{m}(1-\lambda)q_{ij}^k \log\frac{t^*}{\lambda T_i^k + (1-\lambda)t^*} = 0. \tag{4.19}$$

Since $\sum_{i=1}^m a_{ij} = 1$, we have

$$\prod_{i=1}^{m}\left[\lambda\frac{T_i^k}{t^*} + (1-\lambda)\right]^{a_{ij}} = 1. \tag{4.20}$$

Next we show that S_j^k is between 1 and t^* in the following three cases.

If $t^* = 1$, then

$$S_j^k = \prod_{i=1}^{m}\left[\lambda T_i^k + (1-\lambda)\right]^{a_{ij}} = 1. \tag{4.21}$$

If $t^* > 1$, then

$$\prod_{i=1}^{m}\left[\lambda\frac{T_i^k}{t^*} + (1-\lambda)\right]^{a_{ij}} < \prod_{i=1}^{m}\left[\lambda T_i^k + (1-\lambda)\right]^{a_{ij}} < \prod_{i=1}^{m}\left[\lambda T_i^k + (1-\lambda)t^*\right]^{a_{ij}}. \tag{4.22}$$

If $t^* < 1$, then

$$\prod_{i=1}^{m}\left[\lambda\frac{T_i^k}{t^*} + (1-\lambda)\right]^{a_{ij}} > \prod_{i=1}^{m}\left[\lambda T_i^k + (1-\lambda)\right]^{a_{ij}} > \prod_{i=1}^{m}\left[\lambda T_i^k + (1-\lambda)t^*\right]^{a_{ij}}. \tag{4.23}$$

Using the convexity of $d_\lambda(r_j^{k+1}, tq_j^k)$, we have $d_\lambda(r_j^{k+1}, S_j^k q_j^k) \le d_\lambda(r_j^{k+1}, q_j^k)$. Consequently, with Eqs. (4.16) and (4.17), we get the proof of the lemma.

□

Lemma 4.3. $d_\lambda(x^{(k+1)}, x^{(k)}) \le d_\lambda(r(x^{(k)}), q(x^{(k)})) - d_\lambda(r(x^{(k)}), q(x^{(k+1)}))$.

Proof. Using the same notations as Lemma 4.2, we only have to prove

$$d_\lambda(S_j^k x_j^{(k)}, x_j^{(k)}) \le d_\lambda(r_j^{k+1}, q_j^k) - d_\lambda(r_j^{k+1}, S_j^k q_j^k) \tag{4.24}$$

for those $x_j^{(k)} > 0$. With simple calculation, inequality (4.24) is equivalent to

$$\begin{aligned}\sum_{i=1}^{m} a_{ij}\{&[\lambda T_i^k + (1-\lambda)S_j^k]\log[\lambda T_i^k + (1-\lambda)S_j^k] - [\lambda T_i^k + (1-\lambda)]\log[\lambda T_i^k + (1-\lambda)]\}\\ &\ge S_j^k \log S_j^k - [\lambda S_j^k + (1-\lambda)]\log[\lambda S_j^k + (1-\lambda)].\end{aligned} \tag{4.25}$$

In the following, we use three cases to prove the above inequality.

If $S_j^k = \prod_{i=1}^m(\lambda T_i^k + (1-\lambda))^{a_{ij}} = 1$, (4.25) is obviously true.

If $S_j^k = \prod_{i=1}^m (\lambda T_i^k + (1-\lambda))^{a_{ij}} > 1$, consider the function

$$U_i(t) = [\lambda T_i^k + (1-\lambda)t] \log[\lambda T_i^k + (1-\lambda)t], \tag{4.26}$$

and we can obtain that the left-hand-side of inequality (4.25) equals to $\sum_{i=1}^m a_{ij}[U_i(S_j^k) - U_i(1)]$. Applying the Lagrange mean value theorem to the function (4.26) we get

$$\begin{aligned}
\sum_{i=1}^m a_{ij}[U_i(S_j^k) - U_i(1)] &= \sum_{i=1}^m a_{ij} U_i'(\xi_i)(S_j^k - 1) \qquad (1 \le \xi_i \le S_j^k) \\
&= \sum_{i=1}^m a_{ij}(1-\lambda)(S_j^k - 1)[\log(\lambda T_i^k + (1-\lambda)\xi_i) + 1] \\
&= (1-\lambda)(S_j^k - 1)\left\{\log \prod_{i=1}^m \left[(\lambda T_i^k + (1-\lambda)\xi_i)^{a_{ij}}\right] + 1\right\} \\
&\ge (1-\lambda)(S_j^k - 1)(\log S_j^k + 1).
\end{aligned} \tag{4.27}$$

On the other hand, consider the function

$$V(t) = [\lambda S_j^k + (1-\lambda)t] \log[\lambda S_j^k + (1-\lambda)t]. \tag{4.28}$$

Then the right-hand-side of inequality equals to $V(S_j^k) - V(1)$ and

$$\begin{aligned}
V(S_j^k) - V(1) &= V'(\eta)(S_j^k - 1) \qquad (1 \le \eta \le S_j^k) \\
&= (1-\lambda)(S_j^k - 1)\{\log[\lambda S_j^k + (1-\lambda)\eta] + 1\} \\
&\le (1-\lambda)(S_j^k - 1)(\log S_j^k + 1).
\end{aligned} \tag{4.29}$$

Thus, the inequality (4.25) can be derived by combining inequalities (4.27) and (4.29).

If $0 < S_j^k < 1$, we can use the same strategy in the above case and get the conclusion (4.25).

□

Theorem 4.4. $\lim\limits_{k\to\infty} d_\lambda(x^{(k+1)}, x^{(k)}) = 0.$

Proof. According to Lemma 4.1, we can conclude that

$$d_\lambda(r(x^{(k+1)}), q(x^{(k+1)})) \le d_\lambda(r(x^{(k)}), q(x^{(k+1)})). \tag{4.30}$$

Using Lemma 4.3 and Eq. (4.30),

$$\begin{aligned}
0 \le d_\lambda(x^{(k+1)}, x^{(k)}) &\le d_\lambda(r(x^{(k)}), q(x^{(k)})) - d_\lambda(r(x^{(k)}), q(x^{(k+1)})) \\
&\le d_\lambda(r(x^{(k)}), q(x^{(k)})) - d_\lambda(r(x^{(k+1)}), q(x^{(k+1)})).
\end{aligned} \tag{4.31}$$

Thus $\{d_\lambda(r(x^{(k)}), q(x^{(k)}))\}$ is a decreasing sequence with nonnegative quantities. Consequently, the difference $d_\lambda(r(x^{(k)}), q(x^{(k)})) - d_\lambda(r(x^{(k+1)}), q(x^{(k+1)})) \to 0$ as $k \to \infty$. This means $d_\lambda(x^{(k)}, x^{(k+1)}) \to 0$ as $k \to \infty$. □

Theorem 4.5. *The sequence $\{x^{(k)}\}$ generated by Eq. (3.2) is convergent.*

Proof. According to Proposition 3.2, we know that the sequence is contained in a bounded set. Thus the nonnegative sequence $\{x_j^{(k)}\}$ is bounded for $j = 1, 2, \ldots, n$. Besides, $\{x^{(k)}\}$ has subsequential limit points. Let $\{x^{(k_s)}\}$ be a subsequence of $\{x^{(k)}\}$ such that $\lim\limits_{k_s \to \infty} x^{(k_s)} = x^*$. According to Theorem 4.4, we have

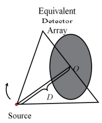
$$d_\lambda(x^{(k_s+1)}, x^{(k_s)}) = \sum_{j=1}^{n} \Big[\lambda(S_j^{k_s} x_j^{(k_s)}) \log(S_j^{k_s} x_j^{(k_s)}) + (1-\lambda) x_j^{(k_s)} \log x_j^{(k_s)} - \Big(\lambda(S_j^{k_s} x_j^{(k_s)}) + (1-\lambda) x_j^{(k_s)}\Big) \log \Big(\lambda(S_j^{k_s} x_j^{(k_s)}) + (1-\lambda) x_j^{(k_s)}\Big) \Big] \quad \to 0, \tag{4.32}$$

as $k_s \to \infty$. Using the convexity of function $\varphi(t) = t \log t$, we have $S_j^{k_s} \to 1$ if $x_j^* \neq 0$, or $S_j^k x_j^{(k_s)} \to 0$ if $x_j^* = 0$. Thus $x_j^{(k_s+1)} = S_j^{k_s} x_j^{(k_s)} \to x_j^*$ for $j = 1, 2, \ldots, n$. Consequently, x^* is the limit point of the sequence $\{x^{(k)}\}$. □

5 Numerical Test

In this section, we use a fan-beam X-ray CT reconstruction problem of low-contract Shepp-Logan phantom to test the performance of the iteration formula (3.2). The details of numerical phantom and imaging system are shown in Table 1.

Table 1: Numerical phantom and imaging system

Low-contrast Shepp-Logan 	radius of phantom r: 10 cm resolution of reconstructed image 256×256
Fan-beam X-ray CT 	rotation radius of source R: 20 cm detector array: 256 sampling number of source half turn: 360
Imaging matrix $\tilde{A} = (\tilde{a}_{ij})_{m \times n}$	$\tilde{a}_{ij}$ represents the length of the intersection line between the i-th ray and the j-th pixel
Measurements $\tilde{b} = (\tilde{b}_i)_{m \times 1}$	$\tilde{b}_i$ is the line integral of the ith ray with noise

The system matrix $\tilde{A}$ and right-hand side $\tilde{b}$ of the reconstruction problem do not have the property of normalization. By simple variables substitution, the iterative formula for attenuation $\tilde{x}^{(k+1)}$ can be expressed as

$$\tilde{x}_j^{(k+1)} = \tilde{x}_j^{(k)} \prod_{i=1}^{m} \left[\lambda \frac{\tilde{b}_i}{\sum_{j=1}^{n} \tilde{a}_{ij} \tilde{x}_j^{(k)}} + (1-\lambda) \right]^{\tilde{a}_{ij} / \sum_{i=1}^{m} \tilde{a}_{ij}}, \tag{5.1}$$

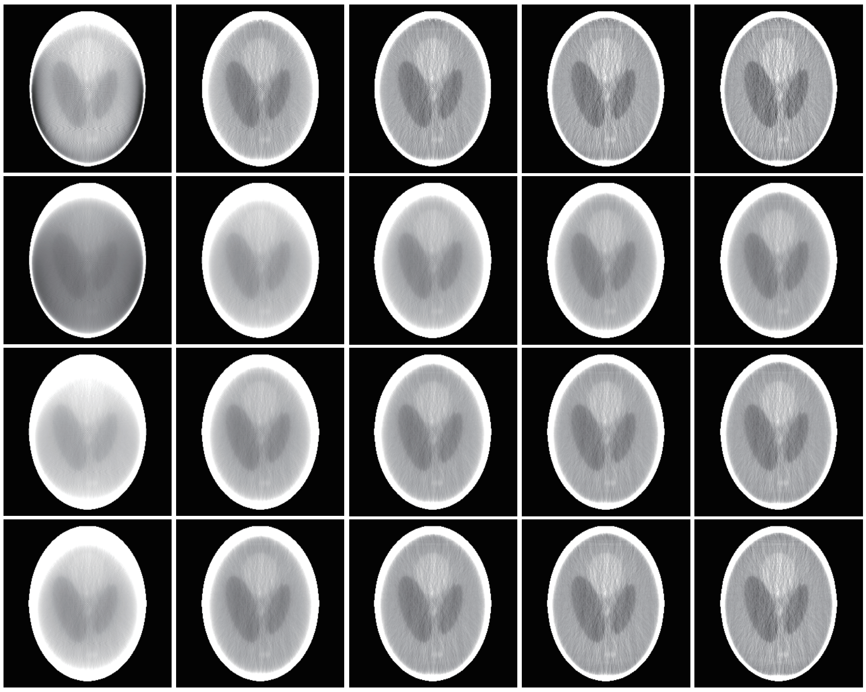
for $j = 1, 2, \ldots, N$.

EM

$\lambda = 0.3$

$\lambda = 0.5$

$\lambda = 0.7$



,

Figure 1: Reconstructed images of different iterations

The deviation between reconstructed image $\tilde{x}^{(k)}$ and phantom $\tilde{x}$ can be represented by the root mean squared error (RMSE) of the iteration sequence

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(\tilde{x}_j^{(k)} - \tilde{x}_j)^2}. \tag{5.2}$$

The first experiment verifies whether the iterative formula can run correctly. The reconstructed images of traditional EM algorithm and our method are shown in Figure 1. Here $\lambda = 0.3, 0.5, 0.7$. From left to right are the iterative results of steps 10, 20, 30, 40 and 50, respectively. The results show that all the sequences converge to phantom successfully. However, with the continuous iteration, the smoothness of the reconstructed image slightly decreases.

The second experiment is taken to test the reconstruction efficiency. Let λ be 0.3, 0.5, 0.7 and 0.9 respectively. We compared our iteration method with traditional EM algorithm and provide the RMSEs of different iterations in Figure 2. The figure shows that when $\lambda \geq 0.5$, the RMSEs of the iterative sequences generated by our algorithm decrease faster than the RMSE of traditional EM algorithm. However, the phenomenon of semi convergence appears with the increase of iteration steps. The larger the λ, the more obvious this phenomenon is. Thus the problem of stopping criterion needs further study. Besides, for different imaging systems, the impact of λ on the convergence will be further carried out in the future work.

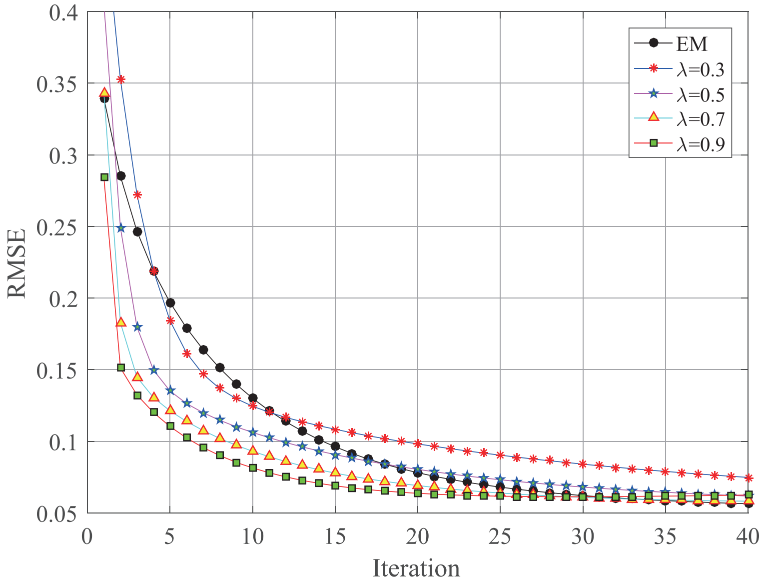


Figure 2: RMSEs plotted against iterations

$\boxed{6}$ Conclusion

In this paper, we focus on an iterative reconstruction method for image reconstruction problem $Ax = b$. Based on the Kullback-Leibler divergence, we introduce a parameter λ and use the λ-Kullback-Leibler divergence $d_\lambda(b, Ax)$ to measure the similarity between the observed data b and forward data Ax. We solve the linear reconstruction problem through a new iterated method which is derived using Karush-Kuhn-Tucker condition of the optimization problem. Afterward, we define two mn-dimensional hidden vectors and split the original optimization problem into two subproblems. The convergent property of the iteration sequence is proved based on the subproblems. To verify the convergence performance of the iteration formula, we reconstruct the Shepp-Logan model of fan-beam X-ray CT problem. Reconstructed images as well as the RMSEs of the derived method are compared with those of traditional EM algorithm. Preliminary numerical experiments show that our iterative formula is feasible and effective.

References

[1] C. Byrne, Iterative image reconstruction algorithms based on cross-entropy minimization, *IEEE Trans. Image Process.* 2 (1993) 96–103.

[2] Y. Censor and G. T. Herman, On some optimization techniques in image reconstruction from projections, *Appl. Numer. Math.* 3 (1987) 365–391.

[3] D. Chen, H. Xiao and J. Xu, An improved Richardson-Lucy iterative algorithm for C-scan image restoration and inclusion size measurement, *Ultrasonics.* 91 (2019) 103–113.

[4] A. Gaitanis, G. Kontaxakis, G. Panayiotakis, G. Spyrou and G. Tzanakos, The role of the updating coefficient of the ML-EM algorithm in PET image reconstruction, *Eur. J. Nucl. Med. Mol. Imaging.* 33 (2006): S314.

[5] X. Guo, Convergence studies on block iterative algorithms for image reconstruction, *Appl. Math. Comput.* 273 (2016) 525–534.

[6] M. Haltmeier, A. Leitão and E. Resmerita, On regularization methods of EM-Kaczmarz type, *Inverse Probl.* 25 (2009): 075008.

[7] T. Hohage and F. Werner, Inverse problems with Poisson data: statistical regularization theory, applications and algorithms, *Inverse Probl.* 32 (2016): 093001.

[8] M. H. Hudson and R. S. Larkin, Accelerated image reconstruction using ordered subsets of projection data, *IEEE Trans. Med. Imaging.* 13 (1994) 601–609.

[9] A. N. Iusem, A short convergence proof of the EM algorithm for a specific Poisson model, *Braz. J. Probab. Stat.* 6 (1992) 57–67.

[10] M. Jiang and G. Wang, Convergence studies on iterative algorithms for image reconstruction, *IEEE Trans. Med. Imaging.* 22 (2003) 569–579.

[11] R. Kasai, Y. Yamaguchi, T. Kojima and T. Yoshinaga, Hybrid of ML-EM and MART algorithms for X-ray CT image reconstruction, *Basic Clin. Pharmacol. Toxicol.* 125 (2019): 26.

[12] D. Kim, D. Lee, H. Kim, Z. Chao, M. Lee and H. Kim, Image restoration based on projection onto convex sets algorithm for beam modulation CT acquisition, *Radiat. Phys. Chem.* 158 (2019) 199–204.

[13] S. Kullback and R. A. Leibler, On information and sufficiency, *The Annals of Mathematical Statistics.* 22 (1951) 79–86.

[14] G. Qu, C. Wang and M. Jiang, Necessary and sufficient convergence conditions for algebraic image reconstruction algorithms, *IEEE Trans. Image Process.* 18 (2009) 435–440.

[15] N. T. A. Nguyen, G. Charron, D. Blais and D. Roberge, Turn down the noise-a blinded evaluation of iterative image reconstruction in radiation therapy computed tomography simulation, *Pract. Radiat. Oncol.* 5 (2015) E393–E400.

[16] A. J. Rockmore and A. Macovski, A maximum likelihood approach to emission image reconstruction from projections, *IEEE Trans. Nucl. Sci.* 23 (1976) 1428–1432.

[17] L. A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Trans. Med. Imaging.* 1 (1982) 113–122.

[18] K. Shimomura, M. Hirose, T. Higashino and Y. Takahashi, Three-dimensional iterative multislice reconstruction for ptychographic X-ray computed tomography, *Opt. Express.* 26 (2018) 31199–31208.

[19] X. Tang, E. A. Krupinski, H. Xie and A. E. Stillman, On the data acquisition, image reconstruction, cone beam artifacts, and their suppression in axial MDCT and CBCT - A review, *Med. Phys.* 45 (2018) E761–E782.

[20] F. Topsøe, Some inequalities for information divergence and related measures of discrimination, *IEEE Trans. Inf. Theory.* 46 (2000) 1602–1609.

[21] C. Wang, Properties of subgradient projection iteration when applying to linear imaging system, *Optim. Lett.* 13 (2019) 1285–1297.

[22] H. Xie, T. Xia, Z. Tian, X. Zheng and X. Zhang, A least squares support vector regression coupled linear reconstruction algorithm for ECT, *Flow Meas. Instrum.* 77 (2021): 101874.

Manuscript received 11 January 2023
revised 21 July 2023
accepted for publication 21 July 2023

Caifang Wang
College of Sciences, Shanghai Maritime University
Shanghai 201306, China
E-mail address: cfwang@shmtu.edu.cn

Xiaobin Xia
College of Sciences, Shanghai Maritime University
Shanghai 201306, China
E-mail address: 202031010003@stu.shmtu.edu.cn

Dan Gao
College of Sciences, Shanghai Maritime University
Shanghai 201306, China
E-mail address: unrivalled_gd@163.com