A VARIATIONAL EXPECTATION-MAXIMIZATION METHOD FOR THE INVERSE BLACK BODY RADIATION PROBLEM*

Jiantao Cheng and Tie Zhou

School of Mathematical Sciences, Peking University, Beijing 100871, China

Email: jiantaocheng@gmail.com, tie.zhou@pku.edu.cn

Abstract

The inverse black body radiation problem, which is to reconstruct the area temperature distribution from the measurement of power spectrum distribution, is a well-known ill-posed problem. In this paper, a variational expectation-maximization (EM) method is developed and its convergence is studied. Numerical experiments demonstrate that the variational EM method is more efficient and accurate than the traditional methods, including the Tikhonov regularization method, the Landweber method and the conjugate gradient method.

1. Introduction

The inverse black body radiation (BBR) problem is to determine the area temperature distribution subject to the total power spectral measurements of its radiation. The first formulation of the problem was proposed by Bojarski [1] by using the Laplace transform and an iterative process. Since then, several kinds of methods have been developed, see, e.g., [2–6].

Mathematically, BBR is an inherently ill-posed problem since it belongs to the Fredholm integral equation of the first kind. Sun and Jaggard [2] and Dou and Hodgson [4] used the Tikhonov regularization technique to overcome the ill-posedness, but they did not choose rule to fix the regularization parameter. Li and Xiao [6] applied the Morozov discrepancy technique to determine the parameter. However, the method works well only when the measurement error is known beforehand. On the other hand, Dou and Hodgson [4] introduced the potential function and applied the entropy method to study the problem. Recently, Li [5] presented some numerical results making use of the conjugate gradient method.

In this work, we propose a variational EM method for BBR problem, which is a variant of well-known EM algorithm [7,8]. Its convergence study is given in the appendix. We compare our method with three traditional methods: the Tikhonov regularization method, the Landweber method and the conjugate gradient method. Numerical experiments demonstrate that the proposed method is more efficient and accurate than the traditional methods.

The organization of the paper is as follows. In Section 2, we introduce the BBR problem and its mathematical formulas. After reviewing the traditional methods in Section 3, we propose a variational EM method for the BBR problem in Section 4. In Section 5, we discuss some relevant issues in the numerical computation. In Section 6, numerical implementation is provided. Finally, the convergence of the variational EM method is studied in the appendix.

---

2. Black Body Radiation Problem

Given the area temperature distribution $a(T)$, the total radiated power spectrum $W(v)$ can be expressed by the Planck's law as

$$W(v) = \frac{2hv^3}{c^2} \int_0^\infty \frac{a(T)}{e^{hv/kT} - 1} dT, \tag{2.1}$$

where h is Planck's constant, k is Boltzman's constant, and c is the velocity of light [1]. Set

$$k(v, T) = \frac{2hv^3}{c^2(e^{hv/kT} - 1)}.$$

Consequently,

$$W(v) = \int_0^\infty k(v, T)a(T)dT. \tag{2.2}$$

In practice, the range of T usually goes from 100K to 1000K, and v goes from 0 Hz to 2×10^{14} Hz. In the following, we set $T \in [T_1, T_2]$ and $v \in [V_1, V_2]$, where

$$T_1 = 100\text{K}, \ T_2 = 800\text{K}, \ V_1 = 0 \text{ Hz}, \ V_2 = 2 \times 10^{14} \text{ Hz}. \tag{2.3}$$

Therefore, the problem can be transformed into a standard Fredholm integral equation of the first kind

$$(Ka)(v) = \int_{T_1}^{T_2} k(v, T)a(T)dT = W(v). \tag{2.4}$$

Here, K is the first kind of integral operator from $L^2[T_1, T_2]$ to $L^2[V_1, V_2]$. Given the limited and sometimes noisy power spectrum $W(v)$, the problem becomes how to determine the area temperature distribution $a(T)$ from Eq. (2.4).

3. Traditional Methods

According to the mathematical theories of the inverse problem [9], we implement three kinds of traditional methods for the BBR problem: the Tikhonov regularization method, the Landweber method and the conjugate gradient method.

The dual operator K^* of K needs to be known for implementing these methods. According to the definition of dual operators, we have

$$(K^*\phi)(T) = \int_{V_1}^{V_2} k(v, T)\phi(v)dv, \tag{3.1}$$

where K^* is an operator from $L^2[V_1, V_2]$ to $L^2[T_1, T_2]$.

The Simpson's quadrature rule is used to obtain the numerical evaluation of K. We discretize the interested domain as

$$\begin{aligned}
v_i &= V_1 + i/M * (V_2 - V_1), \quad i = 0, 1, \cdots, M, \\
T_j &= T_1 + j/N * (T_2 - T_1), \quad j = 0, 1, \cdots, N,
\end{aligned} \tag{3.2}$$

where N and M are even natural numbers. We replace $(Ka)(v_i)$ by

$$\sum_{j=0}^N w_j k(v_i, T_j)a(T_j), \quad w_j = \begin{cases} \dfrac{1}{3N}, & j = 0 \text{ or } N, \\ \dfrac{4}{3N}, & j = 1, 3, \cdots, N-1, \\ \dfrac{2}{3N}, & j = 2, 4, \cdots, N-2, \end{cases} \tag{3.3}$$

with the corresponding matrix A

$$[A_{ij}]_{M+1,N+1} = [w_j k(v_i, T_j)]_{M+1,N+1}, \tag{3.4}$$

and replace $(K^*\phi)(T_j)$ by

$$\sum_{l=0}^{M} w_l^g k(v_l, T_j)\phi(v_l), \quad w_l^g = \begin{cases} \dfrac{1}{3M}, & l = 0 \text{ or } M, \\ \dfrac{4}{3M}, & l = 1, 3, \cdots, M-1, \\ \dfrac{2}{3M}, & l = 2, 4, \cdots, M-2, \end{cases} \tag{3.5}$$

with the corresponding matrix A^g

$$[A_{jl}^g]_{N+1,M+1} = [w_l^g k(v_l, T_j)]_{N+1,M+1}. \tag{3.6}$$

Define

$$W^\delta = W + \delta W_e \zeta.$$

This implies that W^δ is the perturbation of the true value W (W_e is the average of W in $[V_1, V_2]$; ζ is a Gaussian distribution with mean 0 and standard deviation 1; δ is the noise control term).

3.1. Tikhonov regularization method

To use the Tikhonov regularization method [9], we need to determine a solution $a^\alpha \in L^2[T_1, T_2]$ that minimizes the Tikhonov functional

$$J_\alpha(a) = \|Ka - W\|^2 + \alpha\|a\|^2, \quad a \in L^2[T_1, T_2], \tag{3.7}$$

which leads to

$$\alpha a^\alpha + K^* K a^\alpha = K^* W. \tag{3.8}$$

Moreover, the discretized Tikhonov equation is

$$\alpha a^{\alpha,\delta} + A^g * A a^{\alpha,\delta} = A^g W^\delta. \tag{3.9}$$

With a proper regularization parameter α, appropriate numerical results can be obtained. The discrepancy principle of Morzov method can be used to choose the parameter α, but this strategy depends on the choice of δ. Unfortunately, in practice we do not know the exact error δ beforehand. In the following numerical implementation, we choose α by numerical tests.

3.2. Landweber method

Landweber [10] and Bialy [11] suggested to rewrite the equation $Kx = y$ in the form $x = (I - wK^*K)x + wK^*W$ for some relaxation parameter $w > 0$. For the BBR problem, the continuous Landweber method is

$$a^0 := 0 \quad \text{and} \quad a^m = (I - wK^*K)a^{m-1} + wK^*W,$$

and the discrete Landweber method is

$$a^0 := 0 \quad \text{and} \quad a^m = (I - wA^g * A)a^{m-1} + wA^g W.$$

The convergence condition for the relaxation parameter w depends on the operator norm $\|K^*K\| = \|K\|^2$, e.g., $0 < w < 1/\|K\|^2$; see, e.g., [9] pp. 42-44 for more details. The Landweber method has also been used to solve linearized equation for inverse scattering problems recently [12].

3.3. Conjugate gradient method

The standard text [9] provides the following derivative process. Define the functional

$$f(a) := \|Ka - W\|^2, \quad a \in L^2[T_1, T_2]. \tag{3.10}$$

We abbreviate

$$\nabla f(a) = 2K^*(Ka - W) \in L^2[T_1, T_2]$$

and note that $\nabla f(a)$ is indeed the Riesz representation of the Fréchet derivative $f'(a)$ of f at a. Since both $L^2[T_1, T_2]$ and $L^2[V_1, V_2]$ are Hilbert spaces, this leads to the following CG algorithm.

3.3.1. CG Algorithm for BBR

Step 1: Initialization.

1. Compute the integral

$$W = \int_{T_1}^{T_2} a(T)k(v, T)dT. \tag{3.11}$$

Set $a^0 = 0, \; n = 0$.

2. If $K^*W = 0$, stop.

3. Set

$$p^0 = -K^*W = -\frac{1}{2}\nabla f(a^0). \tag{3.12}$$

Step 2: For $n \geq 0$, do the following iteration until the convergence criteria is satisfied.

1. Set

$$t_n = \frac{(Ka^n - W, Kp^n)}{\|Kp^n\|^2}, \quad a^{n+1} = a^n - t_np^n. \tag{3.13}$$

2. If $K^*(Ka^{n+1} - W) = 0$, stop.

3. Set

$$\gamma_n = \frac{\|K^*(Ka^{n+1} - W)\|^2}{\|K^*(Ka^n - W)\|^2}, \quad p^{n+1} = K^*(Ka^{n+1} - W) + \gamma_np^n. \tag{3.14}$$

Step 3: The reconstructed temperature distribution is given by a^{n+1}.

4. Variational EM Method

From the formulation in the article [8], we define

$$F[a] = \int_{V_1}^{V_2} \{W \log K[a] - K[a]\} \, dv, \tag{4.1}$$

which is a generalized form of the log likelihood function when the measured data W is subject to the Poissonian distribution. From the integral equation (2.1), it follows that $K[a] > 0$ if

$a > 0$. According to the physical condition [1], the solution a to the BBR problem should be nonnegative if it exists. Therefore, W must be positive from Eq. (2.2). In the following, we assume that $W > 0$ and perform the optimization

$$\underset{a \geq 0}{\arg\max} \, F[a]. \tag{4.2}$$

We first assume that $a > 0$ is a maximizer of F. The case of $a \geq 0$ can be handled as the limiting case. We need to find the Fréchet derivative of F. Let

$$f(t) = F[a + tb], \quad \text{for } t \text{ around } 0, \tag{4.3}$$

where b is an arbitrary bounded function of $L^2[T_1, T_2]$. Then we have

$$\begin{aligned}\frac{d}{dt} f(t)|_{t=0} &= \int_{V_1}^{V_2} \left\{ W \frac{1}{K[a]} - 1 \right\} K[b] \, dv \\ &= \int_{T_1}^{T_2} K^* \left[\frac{W}{K[a]} - 1 \right] b \, dT.\end{aligned}$$

Hence, the Fréchet derivative of F is

$$F'[a] = K^* \left[\frac{W}{K[a]} - 1 \right]. \tag{4.4}$$

If $a > 0$ is a solution of (4.2), it follows that $F'[a] = 0$. The general case of $a \geq 0$ is given by the Kuhn-Tucker condition [13]:

$$a \cdot K^* \left[\frac{W}{K[a]} - 1 \right] = 0. \tag{4.5}$$

Let $q_1 = K^*[1]$, i.e.,

$$q_1 = \int_{V_1}^{V_2} \frac{2hv^3}{c^2(e^{hv/kT} - 1)} dv. \tag{4.6}$$

It follows from the property of the integral equation (2.1) that $0 < q_1 < C$, where C is a positive constant. The Kuhn-Tucker condition (4.5) can be rewritten as

$$a = \frac{1}{q_1} a \cdot K^* \left[\frac{W}{K[a]} \right]. \tag{4.7}$$

Then, we obtain the following variational EM formula

$$a^{n+1} = \frac{1}{q_1} a^n \cdot K^* \left[\frac{W}{K[a^n]} \right]. \tag{4.8}$$

In summary, the variational EM algorithm is formulated in the following subsection.

4.1. Variational EM algorithm for BBR

Step 1: Initialization.

    1. Compute the following integral

$$W = \int_{T_1}^{T_2} k(v,T)a(T)dT. \tag{4.9}$$

    2. Choose an initial a^0.

Step 2: For $n \geq 0$, do the following iteration until the convergence criteria is satisfied.

    1. Compute the following integral

$$p^n = \int_{T_1}^{T_2} a^n(T)k(v,T)dT. \tag{4.10}$$

    2. Set $\phi^n = W/p^n$.

    3. Compute the following integral

$$q^n = \int_{V_1}^{V_2} k(v,T)\phi^n(v)dv. \tag{4.11}$$

    4. Set $a^{n+1} = a^n \cdot q^n/q_1$.

Step 3: The reconstructed temperature distribution is given by a^{n+1}.

5. Some Relevant Numerical Issues

5.1. Compute integrals

Since the domains $[T_1, T_2]$ and $[V_1, V_2]$ are bounded, we just compute the values at those points which belong to (3.2), i.e.,

$$\{(v_i, T_j), i = 0, 1, 2, \cdots M, j = 0, 1, 2, \cdots ., N\}.$$

The inner product of functions in $[T_1, T_2]$ and $[V_1, V_2]$

$$(p_1, p_2)_{L^2[V_1,V_2]} = \int_{V_1}^{V_2} p_1 p_2 dv, \quad p_1, p_2 \in L^2[V_1, V_2], \tag{5.1}$$

$$(q_1, q_2)_{L^2[T_1,T_2]} = \int_{T_1}^{T_2} q_1 q_2 dT, \quad q_1, q_2 \in L^2[T_1, T_2] \tag{5.2}$$

can be computed based on the above points by the Simpson's rule. The computation of the norm in the space $L^2[T_1, T_2]$ or $L^2[V_1, V_2]$ follows the same process.

5.2. Choice of a^0 for variational EM algorithm

From the variational EM formula (4.8), it is known that $a^{n+1}(T) = 0$ if $a^n(T) = 0$, $n \in \{0, 1, 2, \cdots\}$. Therefore, we choose the initial guess $a^0(T) > 0$. In the following experiment, we set $a^0(T) = 0.1$. According to our numerical experiences, the numerical results have no difference for $a^0(T) > 0$.
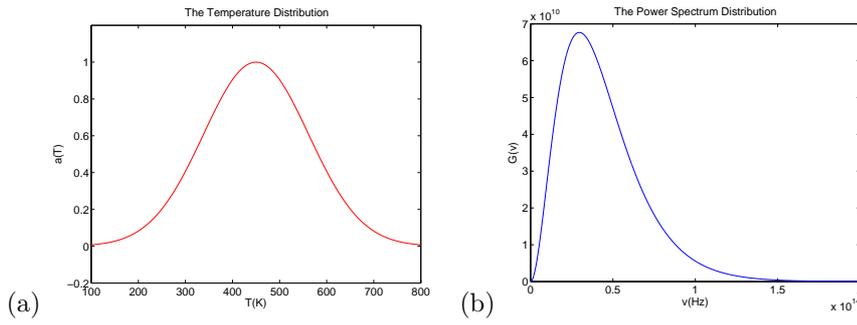
Fig. 6.1. The first numerical example: (a) is the temperature distribution, and (b) is the power spectrum distribution.
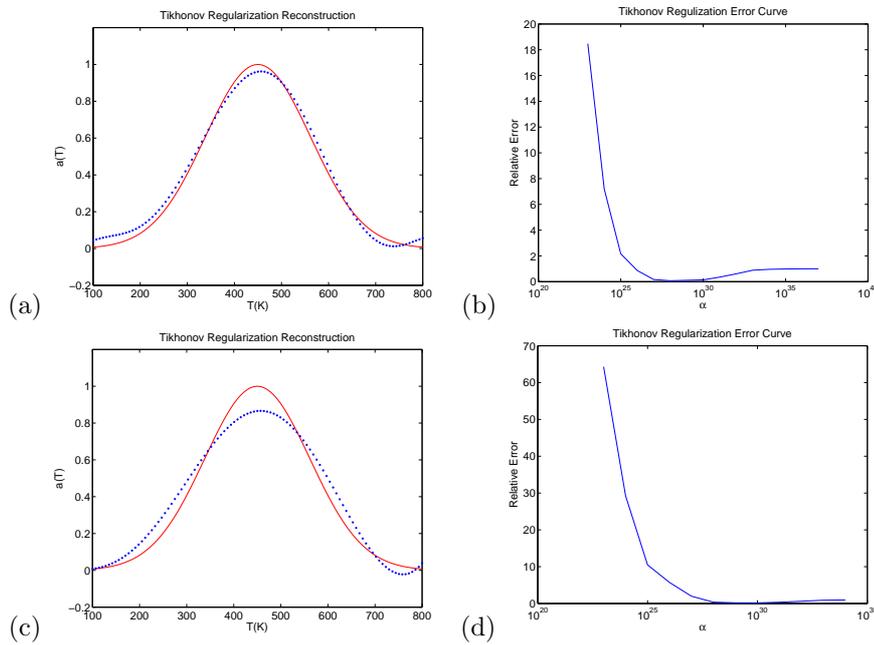

Fig. 6.2. Tikhonov regularization reconstruction: (a) and (c) show the exact (solid line) and the best reconstructed (dotted line) temperature distributions under 1% noise and 10% noise, respectively. (b) and (d) are the error curves according to α.

5.3. Avoid inverse crime

To avoid the notorious Inverse Crime [14], we employ different quadrature points to compute the integral (2.2) in modeling and inversion. In the modeling (e.g., (4.9) and (3.11)), we use the Simpson's rule to compute W with $10N$ points in the T direction, while N points in the T direction are used for the inversion (see, e.g., (4.10)).

5.4. Convergence criteria

The convergence criteria for all algorithms may include (1) when the iteration number n exceeds an assumed maximum number; (2) when the successive incremental $|a^{n+1} - a^n|$ is smaller than an given error tolerance.
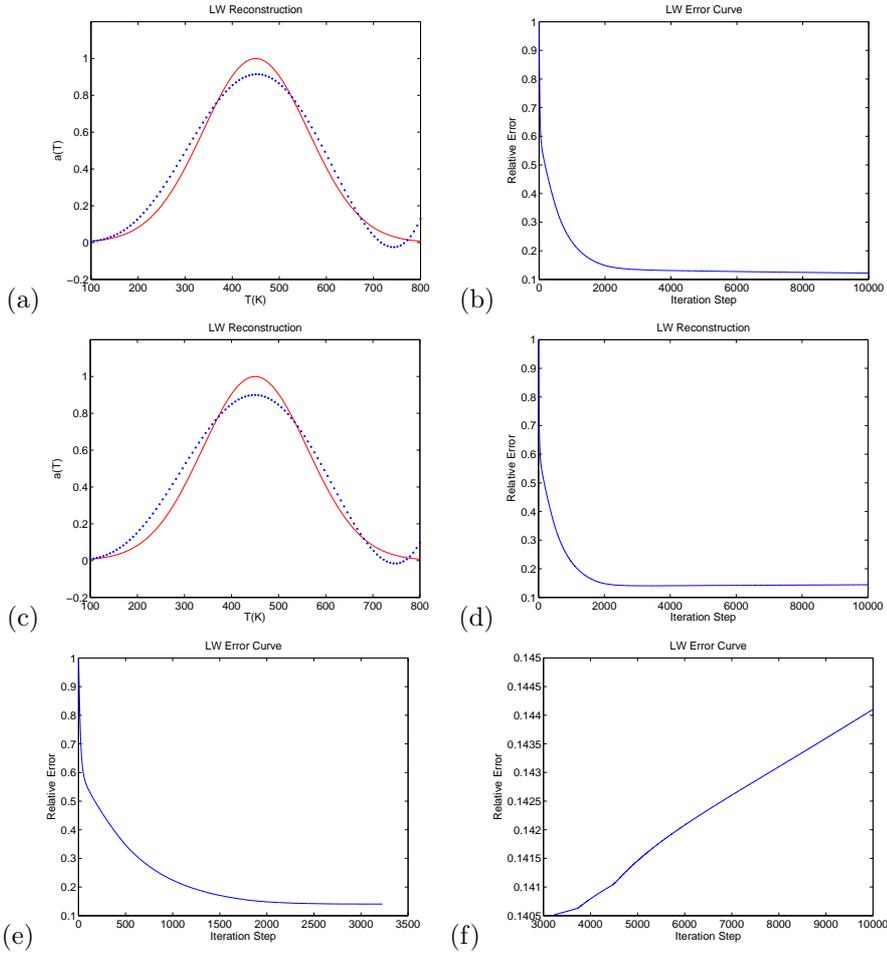
Fig. 6.3. The Landweber method reconstruction: (a) and (c) are exact (solid line) and the best reconstructed (dotted line) temperature distributions under 1% noise and 10% noise, respectively. (b) and (d) show the error curves against the iteration steps. (e) gives the error curve when $k = 1, 2, \cdots, 3173$ under 10% noise; and (f) shows the error curve when $k = 3173, \cdots, 10000$ under 10% noise.

6. Numerical Experiments

To verify the proposed algorithm, we test two numerical examples. In the first example, the exact area temperature distribution as shown in Fig. 6.1(a) is of the form

$$a(T) = e^{-(T-450)^2/25000}, \quad 100 \text{ K} \leq T \leq 800 \text{ K}, \tag{6.1}$$

and we set $M = 1000$ and $N = 100$ for the discretization (3.2).

For the simulation, the power spectrum function $W(v)$ is computed by the Simpson's rule (2.2), which is shown in Fig. 6.1(b). The Gaussian noise is added, i.e.,

$$W^\delta = W + \delta W_e \zeta, \tag{6.2}$$

where W_e is the average of W in $[V_1, V_2]$; ζ is a normal distribution with mean 0 and standard deviation 1; δ is the noise control term.
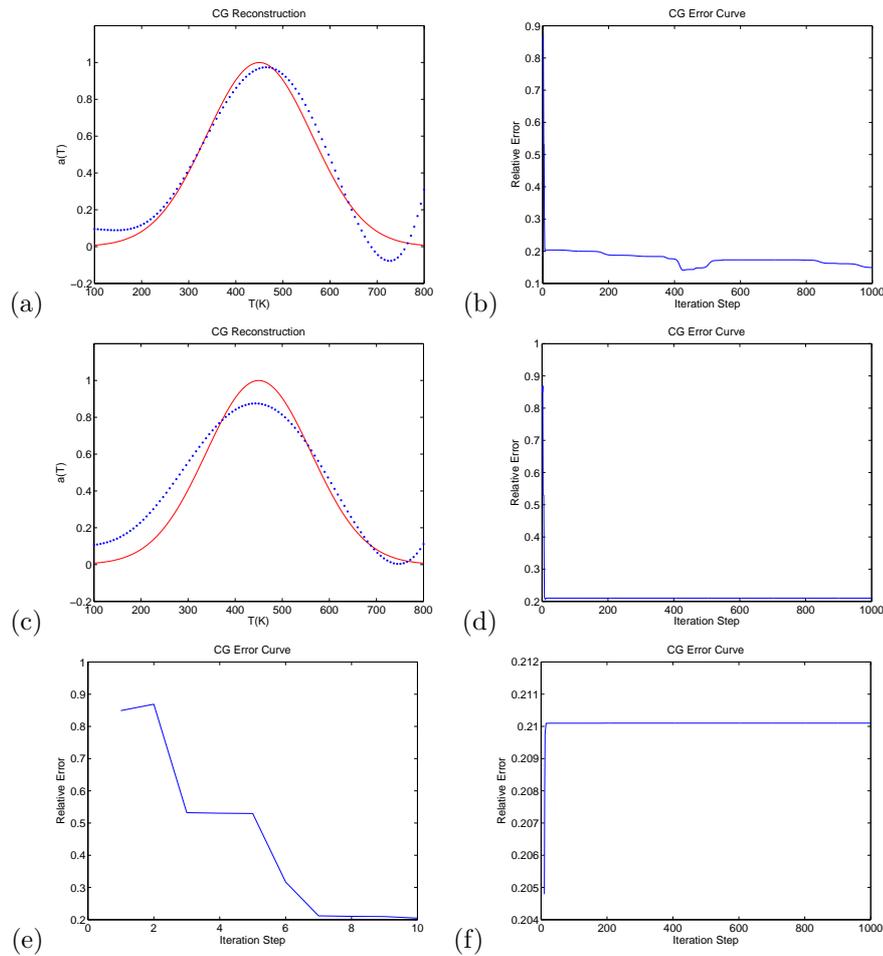
Fig. 6.4. Conjugate gradient method reconstruction: a) and (c) are exact (solid line) and the best reconstructed (dotted line) temperature distributions under 1% noise and 10% noise, respectively. (b) and (d) show the error curves against the iteration steps. (e) gives the error curve when $k = 1, 2, \cdots, 10$ under 10% noise; and (f) shows the error curve when $k = 10, \cdots, 10000$ under 10% noise.

We let $\delta = 0.01, 0.1$, i.e., add $1\%, 10\%$ Gaussian distribution noise. The following figures show the performances of the four methods: the Tikhonov regularization method in Fig. 6.2, the Landweber (LW) method in Fig. 6.3, the conjugate gradient (CG) method in Fig. 6.4 and the EM method in Fig. 6.5.

When $\delta = 0.01$, i.e., adding 1% Gaussian noise, we choose $\alpha = 1.0 \times 10^{28}$ in the Tikhonov regularization method; in this case the minimum relative error is 0.075. The other three iterative methods (LW, CG, EM) give the minimum relative errors $0.1221, 0.141, 0.016$ when the iteration step k are $10000, 426, 756$, respectively. According to the regularization theories [9], we realize that the iteration of Landweber can be continued if smaller errors are desired.

When $\delta = 0.1$, i.e., adding 10% Gaussian noise, we choose $\alpha = 1.0 \times 10^{29}$ in the Tikhonov regularization method; in this case the minimum relative error is 0.151. The other three iterative methods (LW, CG, EM) give the minimum relative error $0.141, 0.205, 0.0439$ when the iteration step k are $3173, 10, 164$, respectively.
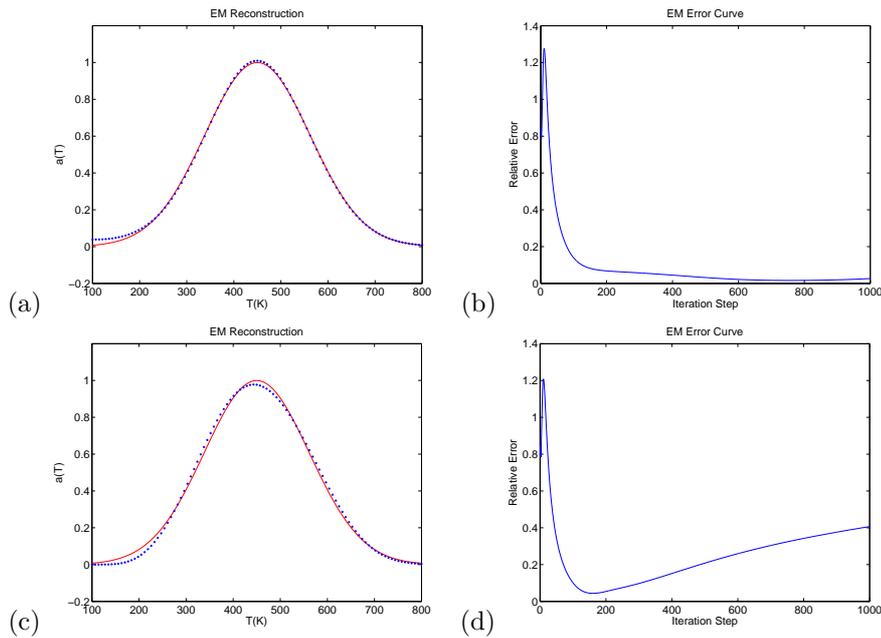
Fig. 6.5. EM method reconstruction: (a) and (c) are exact (solid line) and the best reconstructed (dotted line) temperature distribution under 1% noise and 10% noise, respectively. (b) and (d) show the error curves against the iteration steps.
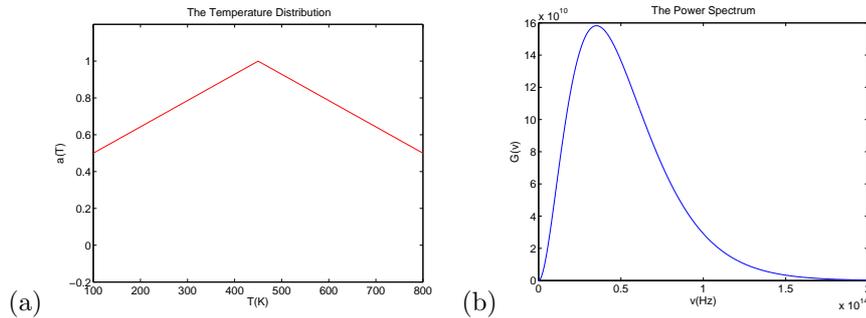

Fig. 6.6. The second numerical example: (a) is the temperature distribution and (b) is the power spectrum distribution.

In these numerical experiments, all the methods seem to give satisfied results. However, the Landweber method needs more iteration steps than other methods. For the Tikhonov regularization method, we must choose a proper parameter; otherwise the errors may become large. The Landweber method also requires to choose proper relaxation parameters. The CG method and the EM method are parameters free and both converge very fast. On the other hand, all the traditional methods conform to the regularization stragies [9], and the variational EM method follows them, too.

Checking the efficiency of all the methods, we find that the Tikhonov regularization method is the fastest since it solves the equation system only once if the parameter is chosen. The other three are iterative methods, and the computational time depends on the iteration steps used. It is observed that the CG method and the EM method converge faster than the LW method.
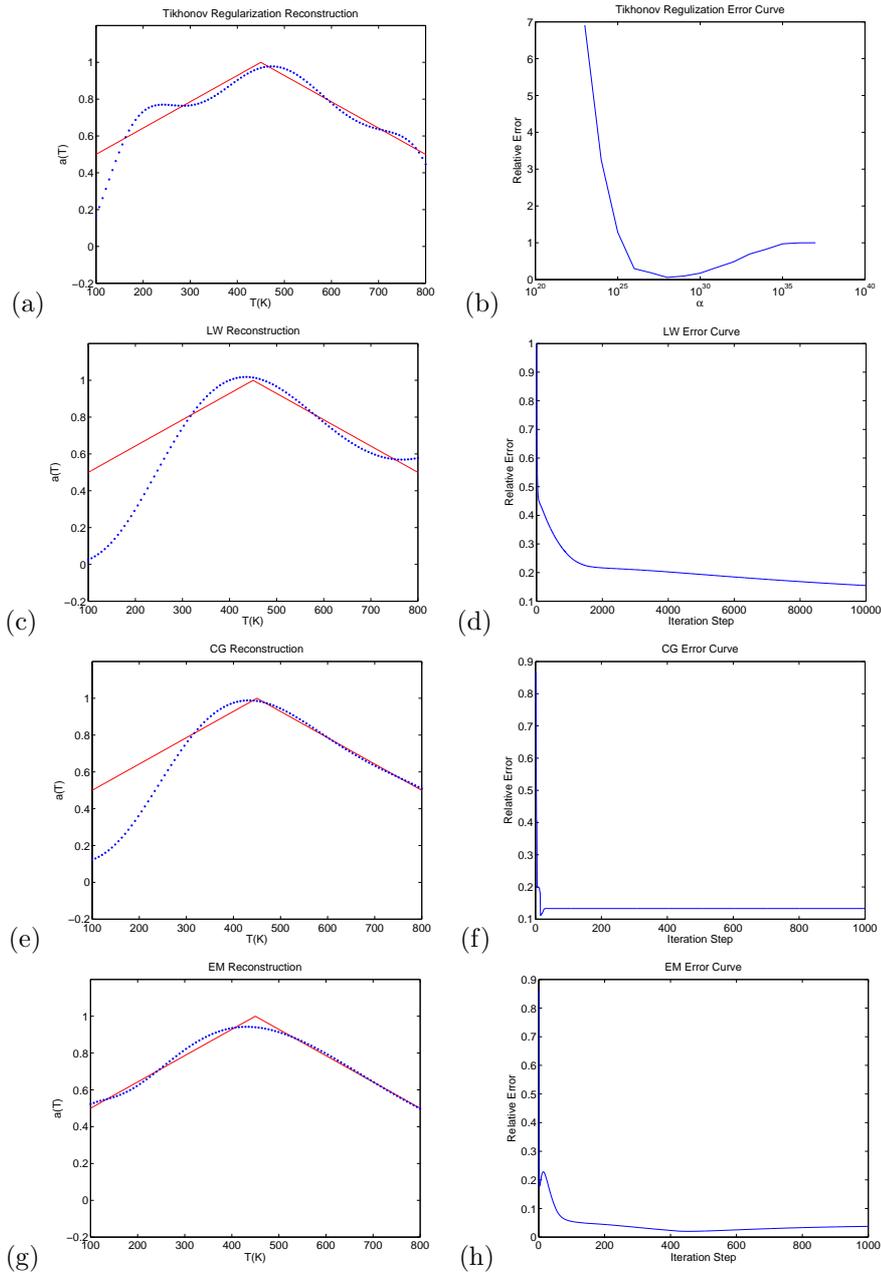
Fig. 6.7. Reconstruction: (a), (c), (e) and (g) are exact (solid line) and the best reconstructed (dotted line) temperature distributions under 1% noise; (b) is the error curve according to α; (d), (f) and (h) are the error curves according to iteration steps.

More importantly, the EM algorithm has higher accuracy than the CG method.

In the second example, we consider a case of a non-smooth area temperature whose distribution is shown in Fig. 6.6(a) and is of the form

$$a(T) = 1 - \frac{1}{700}|T - 450|, \quad 100 \text{ K} \leq T \leq 800 \text{ K}. \tag{6.3}$$

We also set $M = 1000$ and $N = 100$ in the discretization (3.2). The noise term δ is set to be 0.01, i.e., we add 1% Gaussian noise.

Fig. 6.7 shows the reconstruction figures and the error curves of the four methods considered. The Tikhonov regularization method arrives its minimum error 0.059 when $\alpha = 1.0 \times 10^{28}$. The other three iterative methods (LW, CG, EM) reach to the minimum relative errors $0.155, 0.118, 0.020$ when the iteration step k are 10000, 15, 455, respectively. The variational EM method also gives the best reconstruction results; and similar conclusions with the first example can be drawn.

7. Conclusion

In this work, we have proposed a new method for the BBR problem and evaluated its feasibility by numerical experiments. The variational EM method has been shown to be more efficient and accurate than the traditional methods. We point out that the variational EM method can only solve the Fredholm integral equations of the first kind that have a positive kernel and non-negative solutions. We will investigate theoretical properties of the variational EM method in our future work.

Appendix: Convergence Study of EM Method

Using the convergence proof of the EM method given in [13], we study the convergence of the variational EM method.

Theorem A.1. *Let $a^0 > 0$. If a^n of (4.8) has a convergence subsequence, then a^n converges to a maximizer of (4.1).*

Proof. We first define a similar Kullback-Leibler distance

$$L_1(a(T), b(T)) = \int_{T_1}^{T_2} K^*[1](a(T) \log \frac{a(T)}{b(T)} + b(T) - a(T))dT, \tag{A.1}$$

$$L_2(w(v), u(v)) = \int_{V_1}^{V_2} (w(v) \log \frac{w(v)}{u(v)} + u(v) - w(v))dv. \tag{A.2}$$

In the first step of the proof, we show that

$$F(a^{n+1}) \geq F(a^n), \ k = 1, 2, \cdots. \tag{A.3}$$

We start out from the identity

$$\int_{V_1}^{V_2} W \log(K[a]) = \int_{V_1}^{V_2} W \int_{T_1}^{T_2} \frac{k(v,T)h}{K[h]} \left[\log(k(v,T)a) - \log\left(\frac{k(v,T)a}{K[a]}\right) \right] dT dv, \tag{A.4}$$

where $h(T) \geq 0$ is a temperature distribution. Since for $n > 0$

$$\int_{V_1}^{V_2} K[a^{n+1}]dv = \int_{T_1}^{T_2} a^{n+1} \cdot K^*[1]dT$$

$$= \int_{T_1}^{T_2} \frac{1}{K^*[1]}a^n \cdot K^*\left[\frac{W}{K[a^n]}\right] \cdot K^*[1]dT = \int_{V_1}^{V_2} W(v)dv, \tag{A.5}$$

we have

$$F(a^{n+1}) - F(a^n) = \int_{V_1}^{V_2} W \log(K[a^{n+1}]) - \int_{V_1}^{V_2} W \log(K[a^n]). \tag{A.6}$$

Applying the identity with $h = a^n$ and $a = a^n, a^{n+1}$, respectively, we obtain

$$F(a^{n+1}) - F(a^n)$$

$$= \int_{V_1}^{V_2} W \int_{T_1}^{T_2} \sigma^n \left[\log(k(v,T)a^{n+1}) - \log\left(\sigma^{n+1}\right)\right] dT dv$$

$$- \int_{V_1}^{V_2} W \int_{T_1}^{T_2} \sigma^n \left[\log(k(v,T)a^n) - \log\left(\sigma^n\right)\right] dT dv$$

$$= \int_{V_1}^{V_2} W \int_{T_1}^{T_2} \sigma^n \left[\log\left(\frac{a^{n+1}}{a^n}\right) - \log\left(\sigma^{n+1}/\sigma^n\right)\right] dT dv$$

$$= \int_{T_1}^{T_2} K^*[1]a^{n+1} \log\left(\frac{a^{n+1}}{a^n}\right) dT - \int_{V_1}^{V_2} W \int_{T_1}^{T_2} \sigma^n \log\left(\sigma^{n+1}/\sigma^n\right) dT dv, \tag{A.7}$$

where $\sigma^n = k(v,T)a^n/K[a^n]$. Applying Jensen's inequality to the double integrals, we obtain

$$F(a^{n+1}) - F(a^n) \geq \int_{T_1}^{T_2} K^*[1]a^{n+1} \log\left(\frac{a^{n+1}}{a^n}\right) - \int_{V_1}^{V_2} W \log \int_{T_1}^{T_2} \frac{k(v,T)a^{n+1}}{K[a^{n+1}]}$$

$$= L_1(a^n, a^{n+1}) - \int_{V_1}^{V_2} W \log 1 \geq 0.$$

This proves (A.3). In the second step, we show that for each limit point a^* of the sequence a^n we have

$$L_1(a^*, a^{n+1}) \leq L_1(a^*, a^n). \tag{A.8}$$

Here we define

$$k_1(v,T) = K^*[1]\frac{k(v,T)W/K[a^*]}{K^*(W/K[a^*])}, \quad k_2(v,T) = K^*[1]\frac{k(v,T)W/K[a^n]}{K^*(W/K[a^n])}. \tag{A.9}$$

Then, we have

$$\int_{V_1}^{V_2} k_1 = \int_{V_1}^{V_2} k_2 = K^*[1]$$

and

$$0 \leq \int_{T_1}^{T_2} a^* L_2(k_1, k_2) = \int_{T_1}^{T_2} a^* \int_{V_1}^{V_2} k_1 \log \frac{k_1}{k_2}$$

$$= \int_{T_1}^{T_2} a^* \int_{V_1}^{V_2} K^*[1]\frac{k(v,T)W/K[a^*]}{K^*(W/K[a^*])} \log \frac{K[a^n]K^*(W/K[a^n])}{K[a^*]K^*(W/K[a^*])}$$

$$= \int_{T_1}^{T_2} a^* \int_{V_1}^{V_2} K^*[1]\frac{k(v,T)W/K[a^*]}{K^*(W/K[a^*])} \log \frac{K[a^n]a^{n+1}}{K[a^*]a^n}. \tag{A.10}$$

Since

$$a^* = \frac{1}{q_1} a^* \cdot K^* \left[\frac{W}{K[a^*]} \right], \tag{A.11}$$

we obtain

$$K^*[1] = K^* \left[\frac{W}{K[a^*]} \right]. \tag{A.12}$$

It follows that

$$
\begin{aligned}
0 &\leq \int_{T_1}^{T_2} a^* \int_{V_1}^{V_2} \frac{k(v,T)W}{K[a^*]} \log \frac{K[a^n] a^{n+1}}{K[a^*] a^n} \\
&= \int_{T_1}^{T_2} a^* \int_{V_1}^{V_2} \frac{k(v,T)W}{K[a^*]} \left(\log \frac{K[a^n]}{K[a^*]} + \log \frac{a^{n+1}}{a^n} \right) \\
&= \int_{V_1}^{V_2} W \log \frac{K[a^n]}{K[a^*]} - \int_{T_1}^{T_2} \Lambda^*[1] a^* \log \frac{a^{n+1}}{a^n} \\
&= F(a^n) - F(a^*) + L_1(a^*, a^n) - L_1(a^*, a^{n+1}). \tag{A.13}
\end{aligned}
$$

Now (A.8) follows from (A.3).

In the final step, if we take a^* as the limit of a subsequence a^{n_s}, then $L_1(a^*, a^{n_s}) \to 0$ as $s \to \infty$. Since $L_1(a^*, a^n)$ is nonincreasing, we have $L_1(a^*, a^n) \to 0$ as $n \to \infty$. From the property of L_1, it follows that $a^n \to a^*$. In order to show that a^* is a maximizer of F, we check the Kuhn-Tucker conditions. They are obviously satisfied for $a^*(T) > 0$, since

$$K^* \left[\frac{W}{K[a^*]} - 1 \right](T) = 0.$$

For $a^*(T) = 0$, we have

$$a^{n+1}(T) = \frac{a^0}{(K^*[1])^n} \left(K^*[\frac{W}{K[a^0]}] \right) \cdots \left(K^*[\frac{W}{K[a^n]}] \right)(T) \to 0 \tag{A.14}$$

as $n \to \infty$. Since

$$\left(K^*[\frac{W}{K[a^n]}] \right) \to \left(K^*[\frac{W}{K[a^*]}] \right),$$

this is only possible if $\left(K^*[\frac{W}{K[a^*]}] \right) \leq K^*[1]$, i.e., the Kuhn-Tucker conditions are satisfied for $a^*(T) = 0$, too. $\qquad\square$

References

[1] N. Bojarski, Inverse black body radiation, *IEEE Trans. Antennas Propag.*, **30** (1982), 778-780.

[2] X. Sun and D. Jaggard, The inverse black body radiation problem: a regularization solution, *J. Appl. Phys.*, **62** (1987), 4382-4386.

[3] N. Chen and G. Li, Theoretical investigation on the inverse black body radiation problem, *IEEE T. Antenn. Propag.*, **38** (1990), 1287-1290.

[4] L. Dou and R. Hodgson, Maximum entropy method in inverse black body radiation problem, *J. Appl. Phys.*, **71** (1992), 3159-3163.

[5] H. Li, Solution of inverse black body radiation problem with conjugate gradient method, *IEEE T. Antenn. Propag.*, **53** (2005), 1840-1842.

[6] C. Li and T. Xiao, The fast and stable algorithm for the numerical inversion of black body radiation, *Chinese J. Comput. Phys.*, **19** (2002), 121-125, (in Chinese).

[7] M. Jiang and G. Wang, Image reconstruction for bioluminescence tomography, In *Proceedings of SPIE*, **5535** (2004), 335-351.

[8] M. Jiang, T. Zhou, J. Cheng, W. Cong and G. Wang, Image reconstruction for bioluminescence tomography from partial measurement, *Opt. Express.*, **15** (2007), 11,095-11,116.

[9] A. Kirsch, An Introduction to the Mathematical Theory of Inverse Problems, Springer-Verlag, New York, 1999.

[10] L. Landweber, An iteration formula for fredholm integral equations of the first kind, *Am. J. Math.*, **73** (1951), 615-624.

[11] H. Bialy, Iterative brhsnfluh linearer funkionalgleichungen, *Arch. Rat. Mech. Anal.*, **4** (1959), 166.

[12] G. Bao and P. Li, Inverse medium scattering problems in near-field optics, *J. Comput. Math.*, **25** (2007), 252-265.

[13] F. Natterer and F. Wübbeling, Mathematical Methods in Image Reconstruction, SIAM, Philadelphia, PA, 2001.

[14] D. Colton and R. Kress, Inverse Acoustic and Elctromagnetic Scattering Theory, 2nd ed., Springer-Verlag, Berlin, 1998.