First-Optimize-Then-Discretize Strategy for the Parabolic PDE Constrained Optimization Problem With Application to the Reheating Furnace

ZHI YANG [1,2], MING LIU[1], AND XIAOCHUAN LUO[3], (Member, IEEE)

[1] School of Mechanical and Automotive Engineering, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China
[2] Shandong Institute of Mechanical Design and Research, Jinan 250353, China
[3] College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Corresponding author: Xiaochuan Luo (luoxch@mail.neu.edu.cn)

ABSTRACT Each slab entering to the reheating furnace has an optimal and unique reheating curve. The process of obtaining the optimal reheating curve is to solve the typical Partial differential equations (PDE) constrained optimization problem. Obviously, the solution of optimization problem is determined by both the precision of the mathematical PDE model and the numerical method. Firstly, the more accurate mathematical PDE model, in which some key parameters are reconsidered as temperature-dependent, is built for the reheating furnace. Secondly, the first-optimize-then-discretize approach is introduced to solve this PDE-constrained optimization problem. The analysis of the Fréchet gradient of the cost functional is given and we can prove the gradient is Lipschitz continuous. Then, an improved conjugate gradient method is proposed to solve this problem. Finally, numerical simulations and experiment examples are given and analyzed. The results can prove the effectiveness of the proposed strategy.

INDEX TERMS Reheating furnace, PDE-constrained optimization problem, first-optimize-then-discretize, Lipschitz continuous, improved conjugate gradient algorithm.

I. INTRODUCTION

The reheating furnace is one of the key equipments in the steel industry production line. It is used for the reheat treatment of steel slabs until they are capable of being shaped in the hot rolling mill, as shown in Fig.1. Steel slabs have to go through several thermal, mechanical, or thermomechanical processes. Each slab entering to the furnace has an unique and ideal reheating curve(IRC) in theory [1]. When the temperature of slab rises along the optimal slab reheating curve, the best reheating performance of the slab can be ensured and the heat consumption is least [2]. Thus, it is the request of drawing an optimal slab heating curve as the reference trajectory to meet with the requirements of production and

The associate editor coordinating the review of this manuscript and approving it for publication was Guido Lombardi.

technology. However, the reheating curve is a function of slab-depending factors (such as the reheating strategies of furnace, the slab thickness, steel grade, the required final discharging temperature, the reheating time and so on). The problem is the selection of the optimal reheating curve from the many possibilities [3]. Obviously, the key for this issue is to solve the typical Partial differential equations (PDE) constrained optimization problem.

A. EXITING METHODS

In general, the PDE-constrained optimization problems are widely used in the scientific and engineering applications including optimal design, control and parameter identification, etc [4]. For the extremely complicated problems in real life applications, the PDE-constrained optimization problems are determined by both the precision of the mathematical
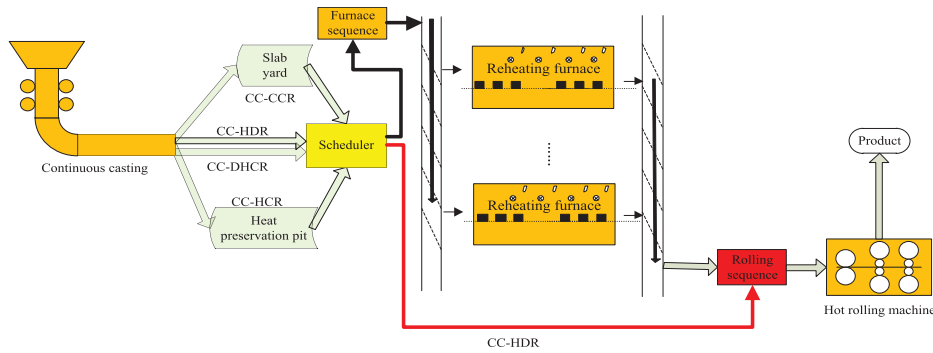
FIGURE 1. The main process of the hot rolling line for the steel slabs.

PDE model and the numerical methods. Thus, a more accurate mathematical PDE model is built for the reheating furnace, and then the numerical approach is given to solve this PDE-constrained optimization problem. Generally, current numerical methods for this class of optimization problems fall into either first-discretize-then-optimize (FDTO) approach or first-optimize-then-discretize (FOTD) approach [5]. On the one hand, the FDTO approach is that we discretize the PDE and the cost functional first and then obtain the first order optimality conditions or Karush-Kuhn-Tucker (KKT) conditions. For instance, the FDTO approach combined with an artificial neural network model is proposed to solve a PDE constrained optimization problem in the paper [6]. On the other hand, the FOTD approach needs to derive the infinite-dimensional first order conditions and then select an appropriate discretization method. Compared with the traditional FDTO approach [7], the FOTD approach allows the calculation to be solved quickly and accurately, and thus achieves a significant reduction in calculation time.

With the rapid development of technologies, more and more researchers are using the FOTD approach and different algorithms to solve the PDE-constrained optimization problems for various research fields. For instance, Wang *et al.* [8] use the FOTD approach to solve a two-dimensional parabolic PDE-constrained optimal control problem, which is applied for the cooling process of steel billets in continuous casting secondary cooling zones. Luo *et al.* [9] use the FOTD strategy to solve the 2-dimensional PDE optimal control problem to obtain the reference values of the optimal furnace zone temperatures for the reheating furnace. Chen *et al.* [4] follow the FOTD approach to solve a coupled system of PDEs, which is derived from the continuous first order optimality conditions. Rezazadeh *et al.* [10] use the FOTD approach and the space-time spectral collocation method to solve the parabolic constrained optimal control problem. It is proved that the accuracy of numerical solution through this method is much higher than the classical numerical solutions. Güttel *et al.* [11] use the FOTD approach to for solve the time-dependent PDE-constrained optimization problems. Cipolla *et al.* [12] use the FOTD approach to address the numerical solution for two Fraction PDE constrained optimization problems.

Both theoretical and experimental analysis of the problem are carried out. Liu and Wang [13] analyze the convergence of several FOTD and FDTO algorithms for solving elliptic PDE-constrained optimal control problems. Several new theoretical conclusions on the convergence of both FOTD and FDTO algorithms are obtained with some elementary error analysis techniques.

The paper is constituted as follows. In Section II, the mathematical PDE model and PDE constrained optimization problem for the reheating furnace are briefly summarized. In Section III, the proposed PDE constrained optimization problem is solved by the FOTD approach. It is proved that the gradient of the cost functional can be written via the weak solution of the adjoint equation. Then, the Lipschitz continuous for the gradient of functional $J(w)$ is proved and an improved conjugate gradient algorithm is proposed. Afterwards, numerical simulations and experiment examples are given and analyzed in Section IV, which demonstrate the performance of the proposed method. Finally, conclusions are presented in Section V.

II. MATHEMATICAL MODELS
A. MATHEMATICAL PDE MODEL
The most important heat transfer process in the furnace is the heat from the furnace gas transferred to the steel slab. Here, the 1-dimensional transient nonlinear heat conduction model is built for the steel slabs. The mathematical description of slab in the furnace can be defined as follows:

$$\rho c\left(T\right) \frac{\partial T}{\partial t}\left(y, t\right) = \frac{\partial}{\partial y}\left(\lambda\left(T\right) \frac{\partial T}{\partial y}\left(y, t\right)\right), t \in [t_0, t_1] \quad (1)$$

where ρ is the density, 7850 (kg/m^3). The symbols $c(T)$ and $\lambda(T)$ are the specific heat ($J/(kg.K)$) and the thermal conductivity ($W/(m.K)$), respectively. The symbol t_0 is the time when the slab enters into the furnace and t_1 is the time when the slab is discharged from the furnace.

To solve the heat transfer problem (1), the initial charging temperature of slab and the nonlinear boundary condition according to the Stefan-Boltzmann law should be given. Therefore, the mathematical PDE model for the reheating

TABLE 1. The thermal properties for A3 steel slab.

Slab Type	Specific heat $c\,(T)$	a	Heat conductivity $\lambda\,(T)$
A3	$408.7 + 199.4\left(\frac{T}{1000}\right)^1 + 810.9e^{-a\lvert T-786\rvert}$	$\dfrac{0.0099}{0.0261}$	$55.58 - \dfrac{31.23}{ch[0.258\frac{T-935}{100}]}$

furnace is shown as following:

$$\rho c\,(T)\frac{\partial T}{\partial t}(y,t) - \frac{\partial}{\partial y}\left(\lambda\,(T)\frac{\partial T}{\partial y}(y,t)\right) = 0,\ (y,t)\in\Omega_{yt},$$

$$T\,(y,t_0) - T_0(y) = 0,$$

$$-\lambda\,(T)\frac{\partial T}{\partial y}\left(\pm\frac{l}{2},t\right)\mp q^{\pm}\,(t) = 0.$$

$$q^{-}\,(t) = \sigma\varepsilon^{-}\,(t)\left[b^4\,(t) - T^4\left(-\frac{l}{2},t\right)\right],$$

$$q^{+}\,(t) = \sigma\varepsilon^{+}\,(t)\left[u^4\,(t) - T^4\left(\frac{l}{2},t\right)\right]. \tag{2}$$

Here, $\Omega_{yt} := \left[-\frac{l}{2},\frac{l}{2}\right]\times[t_0,t_1]$, T_0 is the slab's charging temperature (K), l is the thickness of the slab (m), and σ is the Stefan-Boltzmann constant, $5.67*10^{-8}\,(W/m^2.K^4)$. The symbols $\varepsilon^{\pm}\,(t)$ are the top and bottom total heat exchange factors of the slab. The symbols $q^{\pm}\,(t)$ define the top and bottom heat flux density of the slab. Besides, $b(t)$ and $u(t)$ are the furnace temperatures at the bottom and upper surface of the slab.

To improve the precision of the proposed PDE model, some key parameters (the material parameters (thermal conductivity, the specific heat) and the total heat exchange factors) needs to be improved. Firstly, the choice of material parameters (the specific heat c and the thermal conductivity λ) is closely related to the accuracy of the solution of the PDE model. In previous studies, the material parameters of slabs are always given constant. However, according to the paper [14], the specific heat c and the thermal conductivity λ may explicitly depend on the location (x,y,z) or on the current local temperature $T(x,y,z,t)$ or both. In this paper, $c = c\,(T)$, $\lambda = \lambda\,(T)$ are the function of the slab's temperature. For the usual steel slabs, the calculation formulas can be found in some published researches. According to the paper [15], the c and λ of the A3 steel slab are given in Table. 1.

Secondly, the total heat exchange factors in the PDE models are also given as different constant in different papers ([16], [17], etc.). For instance, in the paper [16], the total heat exchange factors are given as follow: $\varepsilon^{\pm} = \frac{1}{(\varepsilon_s^{\pm}+\varepsilon_w^{\pm})/(\varepsilon^{\pm}\varepsilon_w^{\pm})-1}$. Here, ε^{-}, ε^{+} are $0.65, 0.75$, respectively. And $\varepsilon_w^{-} = \varepsilon_w^{+} = 0.7$. Meanwhile, in the paper [17], they are obtained by the another different equation: $\varepsilon = \frac{\varepsilon_g\varepsilon_s(1+\varphi_{ws}(1-\varepsilon_g))}{\varepsilon_g+\varphi_{ws}(1-\varepsilon_g)[\varepsilon_s+\varepsilon_g(1-\varepsilon_g)]}$. In contrast to the foregoing analysis, the results in our previous work [18] are used here as the identified total heat exchange factors $\varepsilon^{\pm}\,(t)$ along the length of reheating furnace. And we refer the readers to this reference [18] for further details. In Section IV-A, some simulation results are given to prove that the proposed PDE model provides more accurate results.

To isolate the nonlinear material characteristics in (2), a transformation function of the temperature is given.

The transformation law is $\tilde{T}\,(T) = \tilde{T}_0 + \frac{1}{\tilde{c}_0}\int_{T_0}^{T} c\,(\tau)\,d\tau$. It is a nonlinear, time-invariant and bijective function as demonstrated in ([16], [19]). Here $\tilde{c}_0 = c\,(T_0)$, which is the value of the specific heat capacity at the reference temperature T_0, and $\tilde{T}_0 = T_0$. Then, utilization of transformation law into (2), the new PDE equations are given as follows:

$$\rho\tilde{c}_0\frac{\partial\tilde{T}}{\partial t}(y,t) - \frac{\partial}{\partial y}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\tilde{T}}{\partial y}(y,t)\right) = 0,$$

$$\tilde{T}\,(y,t_0) - T_0(y) = 0,$$

$$-\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\tilde{T}}{\partial y}\left(\pm\frac{l}{2},t\right)\mp q^{\pm}\,(t) = 0.$$

$$q^{-}\,(t) = \sigma\varepsilon^{-}\,(t)\left[b^4\,(t) - T^4\left(-\frac{l}{2},t\right)\right],$$

$$q^{+}\,(t) = \sigma\varepsilon^{+}\,(t)\left[u^4\,(t) - T^4\left(\frac{l}{2},t\right)\right]. \tag{3}$$

Here, $\tilde{\lambda}\left(\tilde{T}\right) = \frac{\lambda(\tilde{T})\tilde{c}_0}{c(\tilde{T})}$. Obviously, $q^{\pm}\,(t)$ still define the heat flux density, which does not depend on y.

B. PDE-CONSTRAINED OPTIMIZATION PROBLEM

From the paper [20], it is concluded that an ideal reheating curve can be obtained based on the following conditions: 1) meeting technological conditions; 2) the fuel consumption on the heating process is minimum. If the above two demands are satisfied, neither overheated nor underheated curve will be obtained while maximizing furnace efficiency. In other words, the reference trajectory for the slab in the furnace can be obtained by solving the following optimal problem:

$$\min J\,(w) = \int_{-\frac{l}{2}}^{\frac{l}{2}}\left(\tilde{T}\,(y,t_1;w) - T^*\right)^2 dy$$

$$+\alpha\int_{t_0}^{t_1}\left(u^2\,(t) + b^2\,(t)\right)dt \tag{4}$$

$$st.\rho\tilde{c}_0\frac{\partial\tilde{T}}{\partial t}(y,t) - \frac{\partial}{\partial y}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\tilde{T}}{\partial y}(y,t)\right) = 0,$$

$$\tilde{T}\,(y,t_0) - T_0(y) = 0,$$

$$-\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\tilde{T}}{\partial y}\left(\pm\frac{l}{2},t\right)\mp q^{\pm}\,(t) = 0.$$

$$q^{-}\,(t) = \sigma\varepsilon^{-}\,(t)\left[b^4\,(t) - T^4\left(-\frac{l}{2},t\right)\right],$$

$$q^{+}\,(t) = \sigma\varepsilon^{+}\,(t)\left[u^4\,(t) - T^4\left(\frac{l}{2},t\right)\right]. \tag{5}$$

Here, $w = \{u,b\}\in W$, which represents the control variable.

III. NUMERICAL OPTIMIZATION METHOD BY FOTD APPROACH

Numerical optimization methods such as gradient method, conjugate gradient method, and quasi-Newton method are crucial to solve PDE-constrained optimization problems. The convergence of many numerical optimization algorithms depends on Lipschitz continuity of the gradient of the cost functional. Essentially, there are two methods: the FDTO approach with the sensitivity method, and the FOTD approach with adjoint method [21]. In this section, the FOTD approach by adjoint method is introduced to obtain the Fréchet gradient of the cost functional. Afterwards, the Lipschitz continuous is also proved. Finally, an improved conjugate gradient method, which has good convergence performance, is proposed for the FOTD approach.

A. THE FIRST VARIATION OF THE COST FUNCTIONAL AND ADJOINT EQUATION

Let $w = \{u, b\} \in W$ and $w + \Delta w := \{u + \Delta u, b + \Delta b\} \in W$ be control functions. The first variation $\Delta J(w)$ of the cost functional (4) is

$$\Delta J(w) = J(w + \Delta w) - J(w). \tag{6}$$

Then, it can be rewritten as following:

$$\Delta J(w) = 2 \int_{-\frac{l}{2}}^{\frac{l}{2}} \left(\tilde{T}(y, t_1; w) - T^* \right) \Delta \tilde{T}(y, t_1; w) \, dy$$
$$+ \int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta T^2(y, t_1; w) \, dy + 2\alpha \|\Delta w\|_W^1 + \alpha \|\Delta w\|_W^2 \tag{7}$$

where, $\|\Delta w\|_W^1 := \int_{t_0}^{t_1} u(t) \Delta u(t) \, dt + \int_{t_0}^{t_1} b(t) \Delta b(t) \, dt$, $\|\Delta w\|_W^2 := \int_{t_0}^{t_1} \Delta u^2(t) \, dt + \int_{t_0}^{t_1} \Delta b^2(t) \, dt$.

Obviously, the gradient of last two terms of equation (7) is Lipschitz continuous for the Δw, only the first two terms of (7) need to be considered.

Then, the Lagrange function based approach [21] is used to derive the adjoint derivative. The Lagrange function is obtained and shown as:

$$L\left(\tilde{T}, p, w\right)$$
$$= J(w) + \int_{-\frac{l}{2}}^{\frac{l}{2}} \left(\tilde{T}(y, t_0) - T_0(y) \right) p(y, t_0) \, dy$$
$$+ \int_{t_0}^{t_1} \int_{-\frac{l}{2}}^{\frac{l}{2}} \left(\frac{\partial \tilde{T}}{\partial t}(y, t) - \frac{\partial}{\partial y} \left(\frac{\tilde{\lambda}(\tilde{T})}{\rho \tilde{c}_0} \frac{\partial \tilde{T}}{\partial y}(y, t) \right) \right)$$
$$p(y, t) \, dy dt$$
$$- \int_{t_0}^{t_1} \left(\frac{\tilde{\lambda}(\tilde{T})}{\rho \tilde{c}_0} \frac{\partial \tilde{T}}{\partial y}\left(-\frac{l}{2}, t\right) + \frac{q^-(t)}{\rho \tilde{c}_0} \right) p\left(-\frac{l}{2}, t\right) dt$$
$$+ \int_{t_0}^{t_1} \left(-\frac{\tilde{\lambda}(\tilde{T})}{\rho \tilde{c}_0} \frac{\partial \tilde{T}}{\partial y}\left(\frac{l}{2}, t\right) + \frac{q^+(t)}{\rho \tilde{c}_0} \right) p\left(\frac{l}{2}, t\right) dt \tag{8}$$

According to the book [21] by M. Hinze, namely such that the equation $\Delta L = 0$ for $\tilde{T}$, this is nothing else but the adjoint equation. Thus, it is easy to obtain the following adjoint equation (9).

$$\frac{\partial p}{\partial t}(y, t; w) + \frac{\partial}{\partial y} \left(\frac{\tilde{\lambda}(\tilde{T})}{\rho \tilde{c}_0} \frac{\partial p}{\partial y}(y, t; w) \right) = 0,$$
$$2\left(\tilde{T}(y, t_1) - T^* \right) + p(y, t_1) = 0,$$
$$\tilde{\lambda}(\tilde{T}) \frac{\partial p}{\partial y}\left(-\frac{l}{2}, t\right) - 4\sigma \varepsilon^-(t) T^3\left(-\frac{l}{2}, t\right) p\left(-\frac{l}{2}, t\right) = 0,$$
$$\tilde{\lambda}(\tilde{T}) \frac{\partial p}{\partial y}\left(\frac{l}{2}, t\right) + 4\sigma \varepsilon^+(t) T^3\left(\frac{l}{2}, t\right) p\left(\frac{l}{2}, t\right) = 0. \tag{9}$$

Besides, in order to prove that the gradient of the cost function is Lipschitz continuous, the following equations should be given.

Denote by $\tilde{T}(y, t_1; w)$, $\tilde{T}(y, t_1; w + \Delta w)$ are the corresponding solutions of problem (4). Then $\Delta \tilde{T}(y, t_1; w) := \tilde{T}(y, t_1; w + \Delta w) - \tilde{T}(y, t_1; w)$ will be the weak solution of the following parabolic problem:

$$\rho \tilde{c}_0 \frac{\partial \Delta \tilde{T}}{\partial t}(y, t; w) - \frac{\partial}{\partial y} \left(\tilde{\lambda}(T) \frac{\partial \Delta \tilde{T}}{\partial y}(y, t; w) \right) = 0,$$
$$\Delta \tilde{T}(y, t_0) = 0,$$
$$-\tilde{\lambda}(\tilde{T}) \frac{\partial \Delta \tilde{T}}{\partial y}\left(-\frac{l}{2}, t\right) = 4\sigma \varepsilon^-(t) \left(b^3(t) \Delta b(t) \right.$$
$$\left. -T^3\left(-\frac{l}{2}, t\right) \Delta \tilde{T}\left(-\frac{l}{2}, t\right) \right),$$
$$\tilde{\lambda}(\tilde{T}) \frac{\partial \Delta \tilde{T}}{\partial y}\left(\frac{l}{2}, t\right) = 4\sigma \varepsilon^+(t) \left(u^3(t) \Delta u(t) \right.$$
$$\left. -T^3\left(\frac{l}{2}, t\right) \Delta \tilde{T}\left(\frac{l}{2}, t\right) \right). \tag{10}$$

B. Fréchet GRADIENT OF THE COST FUNCTIONAL

Now, we begin to prove that the first two terms of (7) is Lipschitz continuous for the Δw. Some Lemmas and Theorem are given and proved as following.

Firstly, the **Lemma** 1 is given for the first terms of (7).

Lemma 1: Let $w(t)$, $w(t) + \Delta w(t)$ be given functions. If $\tilde{T}(y, t; w)$ is the corresponding solution of the direct problem(4), and $p(y, t; w)$ is the solution of the adjoint equation (9), then for all $\Delta w(t)$, the following identity holds:

Proof: Let us first consider the left-hand side of (11), as shown at the bottom of the next page. According to the boundary conditions of (9), the left-hand side of (11) can be rewritten as the following equation.

Afterwards, based on the boundary conditions in (10) and (9), the right-hand side of (12), as shown at the bottom of the next page, can be transformed as the equation (13) at page 6.

Lemma 3.1 is proved.

Secondly, by considering the second term of right-hand of (7) and the definition of Fréchet-differential, the **Lemma** 2 is then given and proved as follows.

Lemma 2: Let $\Delta \tilde{T} = \Delta \tilde{T}(y, t; w) \in H^{1,0}(\Omega_{yt})$ be the solution of the parabolic problem (10), corresponding to a given $w \in W$. Then there exists a constant $L_1 > 0$, so that the following equation can be obtained:

$$\int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta \tilde{T}^2(y, t_1) \, dy \leq L_1 \|\Delta w\|_W^2 \tag{14}$$

where $\|\Delta w\|_W^2 := \int_{t_0}^{t_1} \Delta u^2(t) \, dt + \int_{t_0}^{t_1} \Delta b^2(t) \, dt$, and the constants L_1 is defined as: $L_1 = \max \left\{ \frac{4\sigma}{\gamma \rho \tilde{c}_0} \varepsilon_{1*}^2 u_*^6, \frac{4\sigma}{\gamma \rho \tilde{c}_0} \varepsilon_{2*}^2 b_*^6 \right\}$.
Here, $0 < \gamma \leq \min \left\{ 2\varepsilon_{1*} \left(\tilde{T}_{1*} \right)^3, 2\varepsilon_{2*} \left(\tilde{T}_{2*} \right)^3 \right\}$. And $\varepsilon_{1*} = \min_{\Omega_t} \varepsilon^+(t), \varepsilon_{2*} = \min_{\Omega_t} \varepsilon^-(t); \tilde{T}_{1*} = \min_{\Omega_{yt}} \tilde{T} \left(\frac{l}{2}, t \right),$
$b_* = \min_{\Omega_t} b(t), u_* = \min_{\Omega_t} u(t), \tilde{T}_{2*} = \min_{\Omega_{yt}} \tilde{T} \left(-\frac{l}{2}, t \right).$

Proof. Multiply both sides of $\frac{\partial \Delta \tilde{T}}{\partial t}(y, t) = \frac{1}{\rho \tilde{c}_0} \frac{\partial}{\partial y} \left(\tilde{\lambda} \left(\tilde{T} \right) \frac{\partial \Delta \tilde{T}}{\partial y}(y, t) \right)$ by $\Delta \tilde{T}$, integrating over Ω_{yt}, we can obtain the following equation:

$$\frac{1}{2} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \left(\Delta \tilde{T}(y, t) \right)^2 dt dy$$

$$= \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \left(\frac{\partial \Delta \tilde{T}}{\partial t}(y, t) \Delta \tilde{T}(y, t) \right) dt dy$$

$$= \frac{1}{\rho \tilde{c}_0} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \left(\frac{\partial}{\partial y} \left(\tilde{\lambda} \left(\tilde{T} \right) \frac{\partial \Delta \tilde{T}}{\partial y}(y, t) \right) \Delta \tilde{T}(y, t) \right) dt dy \tag{15}$$

On the one hand, using $\Delta \tilde{T}(y, t_0) = 0$, the left-hand of (15) can be transferred as:

$$\frac{1}{2} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \left(\Delta \tilde{T}(y, t) \right)^2 dt dy = \frac{1}{2} \int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta \tilde{T}^2(y, t_1) \, dy \tag{16}$$

On the other hand, the following energy identity can be obtained for the right-hand of (15):

$$\frac{1}{\rho \tilde{c}_0} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \left(\frac{\partial}{\partial y} \left(\tilde{\lambda} \left(\tilde{T} \right) \frac{\partial \Delta \tilde{T}}{\partial y}(y, t) \right) \Delta \tilde{T}(y, t) \right) dt dy$$

$$= \frac{1}{\rho \tilde{c}_0} \int_{t_0}^{t_1} \left(\tilde{\lambda} \left(\tilde{T} \right) \frac{\partial \Delta \tilde{T}}{\partial y}(y, t) \Delta \tilde{T}(y, t) \right) \Big|_{-\frac{l}{2}}^{\frac{l}{2}} dt$$

$$- \frac{1}{\rho \tilde{c}_0} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \left(\tilde{\lambda} \left(\tilde{T} \right) \left(\frac{\partial \Delta \tilde{T}}{\partial y}(y, t) \right)^2 \right) dt dy \tag{17}$$

The first part of the right-hand of (17) can be simplified as:
Afterwards, integrating the (16) into the left-hand of (15), and introducing both (17) and (18), as shown at page 6, into the right-hand of (15), we can obtain the following equation.

Notice, each part of left hand side of equation (19), as shown at page 6, is positive. Next, the ε-inequality $\alpha\beta \leq \varepsilon \frac{\alpha^2}{2} + \frac{\beta^2}{2\varepsilon}, \forall \alpha, \beta \in R, \forall \varepsilon > 0$ is introduced into the right-hand side integrals of this identity (19), we can obtain the following estimate:

Then, combining (20), as shown at page 6, and (19), we can obtain:

Here, requiring that $\gamma \leq \min \left\{ 2\varepsilon_{1*} \left(\tilde{T}_{1*} \right)^3, 2\varepsilon_{2*} \left(\tilde{T}_{2*} \right)^3 \right\}$, and $\tilde{T}_{1*} = \min_{\Omega_{yt}} \tilde{T} \left(\frac{l}{2}, t \right) > 0, \tilde{T}_{2*} = \min_{\Omega_{yt}} \tilde{T} \left(-\frac{l}{2}, t \right) > 0;$
$\varepsilon_{1*} = \min_{\Omega_t} \varepsilon^+(t) > 0, \varepsilon_{2*} = \min_{\Omega_t} \varepsilon^-(t) > 0.$

$$2 \int_{-\frac{l}{2}}^{\frac{l}{2}} \left(\tilde{T}(y, t_1) - T^* \right) \Delta \tilde{T}(y, t_1) \, dy = \frac{4\sigma}{\rho \tilde{c}_0} \int_{t_0}^{t_1} \left[\varepsilon^+(t) u^3(t) p \left(\frac{l}{2}, t \right) \Delta u(t) + \varepsilon^-(t) b^3(t) p \left(-\frac{l}{2}, t \right) \Delta b(t) \right] dt \tag{11}$$

$$2 \int_{-\frac{l}{2}}^{\frac{l}{2}} \left(\tilde{T}(y, t_1) - T^* \right) \Delta \tilde{T}(y, t_1) \, dy$$

$$= - \int_{-\frac{l}{2}}^{\frac{l}{2}} p(y, t_1) \Delta \tilde{T}(y, t_1) \, dy$$

$$= - \left(\int_{-\frac{l}{2}}^{\frac{l}{2}} p(y, t_1) \Delta \tilde{T}(y, t_1) \, dy - \int_{-\frac{l}{2}}^{\frac{l}{2}} p(y, t_0) \Delta \tilde{T}(y, t_0) \, dy \right)$$

$$= - \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{\partial}{\partial t} \left(p(y, t) \Delta \tilde{T}(y, t) \right) dy = - \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{\partial p}{\partial t}(y, t) \Delta \tilde{T}(y, t) \, dy - \int_{-\frac{l}{2}}^{\frac{l}{2}} \frac{\partial \Delta \tilde{T}}{\partial t}(y, t) p(y, t) \, dy$$

$$= \frac{1}{\rho \tilde{c}_0} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \frac{\partial}{\partial y} \left(\tilde{\lambda} \left(\tilde{T} \right) \frac{\partial p}{\partial y}(y, t) \right) \Delta \tilde{T}(y, t) \, dt dy - \frac{1}{\rho \tilde{c}_0} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \frac{\partial}{\partial y} \left(\tilde{\lambda} \left(\tilde{T} \right) \frac{\partial \Delta \tilde{T}}{\partial y}(y, t) \right) p(y, t) \, dt dy \tag{12}$$

$$\frac{1}{\rho\tilde{c}_0}\int_{-\frac{l}{2}}^{\frac{l}{2}}\int_{t_0}^{t_1}\frac{\partial}{\partial y}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial p}{\partial y}(y,t)\right)\Delta\tilde{T}(y,t)\,dtdy - \frac{1}{\rho\tilde{c}_0}\int_{-\frac{l}{2}}^{\frac{l}{2}}\int_{t_0}^{t_1}\frac{\partial}{\partial y}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\Delta\tilde{T}}{\partial y}(y,t)\right)p(y,t)\,dtdy$$

$$= \frac{1}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial p}{\partial y}\left(\frac{l}{2},t\right)\Delta\tilde{T}\left(\frac{l}{2},t\right) - \tilde{\lambda}\left(\tilde{T}\right)\frac{\partial p}{\partial y}\left(-\frac{l}{2},t\right)\Delta\tilde{T}\left(-\frac{l}{2},t\right)\right)dt$$

$$\quad - \frac{1}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\Delta\tilde{T}}{\partial y}\left(\frac{l}{2},t\right)p\left(\frac{l}{2},t\right) - \tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\Delta\tilde{T}}{\partial y}\left(-\frac{l}{2},t\right)p\left(-\frac{l}{2},t\right)\right)dt$$

$$= \frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\varepsilon^+(t)\,u^3(t)\,p\left(\frac{l}{2},t\right)\Delta u(t) + \varepsilon^-(t)\,b^3(t)\,p\left(-\frac{l}{2},t\right)\Delta b(t)\right]dt \tag{13}$$

$$\frac{1}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left(\tilde{\lambda}\left(\tilde{T}\right)\frac{\partial\Delta\tilde{T}}{\partial y}(y,t)\,\Delta\tilde{T}(y,t)\right)\Big|_{-\frac{l}{2}}^{\frac{l}{2}}\,dt$$

$$= \int_{t_0}^{t_1}\left(\tilde{\lambda}(T)\frac{\partial\Delta T}{\partial y}\left(\frac{l}{2},t\right)\Delta T\left(\frac{l}{2},t\right)\right)dt - \int_{t_0}^{t_1}\left(\tilde{\lambda}(T)\frac{\partial\Delta T}{\partial y}\left(-\frac{l}{2},t\right)\Delta T\left(-\frac{l}{2},t\right)\right)dt$$

$$= \frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\varepsilon^+(t)\,u^3(t)\,\Delta u(t)\,\Delta\tilde{T}\left(\frac{l}{2},t\right) + \varepsilon^-(t)\,b^3(t)\,\Delta b(t)\,\Delta\tilde{T}\left(-\frac{l}{2},t\right)\right]dt$$

$$\quad - \frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\varepsilon^+(t)\,\tilde{T}^3\left(\frac{l}{2},t\right)\Delta\tilde{T}^2\left(\frac{l}{2},t\right) + \varepsilon^-(t)\,\tilde{T}^3\left(-\frac{l}{2},t\right)\Delta\tilde{T}^2\left(-\frac{l}{2},t\right)\right]dt \tag{18}$$

$$\frac{1}{\rho\tilde{c}_0}\int_{-\frac{l}{2}}^{\frac{l}{2}}\int_{t_0}^{t_1}\left(\tilde{\lambda}\left(\tilde{T}\right)\left(\frac{\partial\Delta\tilde{T}}{\partial y}(y,t)\right)^2\right)dtdy + \frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\varepsilon^+(t)\,\tilde{T}^3\left(\frac{l}{2},t\right)\Delta\tilde{T}^2\left(\frac{l}{2},t\right) + \varepsilon^-(t)\,\tilde{T}^3\left(-\frac{l}{2},t\right)\Delta\tilde{T}^2\left(-\frac{l}{2},t\right)\right]dt$$

$$+ \frac{1}{2}\int_{-\frac{l}{2}}^{\frac{l}{2}}\Delta\tilde{T}^2(y,t_1)\,dy = \frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\varepsilon^+(t)\,u^3(t)\,\Delta u(t)\,\Delta\tilde{T}\left(\frac{l}{2},t\right) + \varepsilon^-(t)\,b^3(t)\,\Delta b(t)\,\Delta\tilde{T}\left(-\frac{l}{2},t\right)\right]dt \tag{19}$$

$$\frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\varepsilon^+(t)\,u^3(t)\,\Delta u(t)\,\Delta\tilde{T}\left(\frac{l}{2},t\right) + \varepsilon^-(t)\,b^3(t)\,\Delta b(t)\,\Delta\tilde{T}\left(-\frac{l}{2},t\right)\right]dt$$

$$\leq \frac{2\sigma}{\gamma\rho\tilde{c}_0}\int_{t_0}^{t_1}\left(\left(\varepsilon^+(t)\right)^2 u^6(t)\,\Delta u^2(t) + \left(\varepsilon^-(t)\right)^2 b^6(t)\,\Delta b^2(t)\right)dt + \frac{2\sigma\gamma}{\rho\tilde{c}_0}\int_0^t\left(\Delta\tilde{T}^2\left(\frac{l}{2},t\right) + \Delta\tilde{T}^2\left(-\frac{l}{2},t\right)\right)dt \tag{20}$$

$$\frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\left(\varepsilon^+(t)\,\tilde{T}^3\left(\frac{l}{2},t\right) - \frac{\gamma}{2}\right)\Delta\tilde{T}^2\left(\frac{l}{2},t\right)\right]dt + \frac{4\sigma}{\rho\tilde{c}_0}\int_{t_0}^{t_1}\left[\left(\varepsilon^-(t)\,\tilde{T}^3\left(-\frac{l}{2},t\right) - \frac{\gamma}{2}\right)\Delta\tilde{T}^2\left(-\frac{l}{2},t\right)\right]dt$$

$$+ \frac{1}{2}\int_{-\frac{l}{2}}^{\frac{l}{2}}\Delta\tilde{T}^2(y,t_1)\,dy + \frac{1}{\rho\tilde{c}_0}\int_{-\frac{l}{2}}^{\frac{l}{2}}\int_{t_0}^{t_1}\left(\tilde{\lambda}\left(\tilde{T}\right)\left(\frac{\partial\Delta\tilde{T}}{\partial y}(y,t)\right)^2\right)dtdy$$

$$\leq \frac{2\sigma}{\gamma\rho\tilde{c}_0}\int_{t_0}^{t_1}\left(\left(\varepsilon^+(t)\right)^2 u^6(t)\,\Delta u^2(t) + \left(\varepsilon^-(t)\right)^2 b^6(t)\,\Delta b^2(t)\right)dt \tag{21}$$

Finally, with the parameters $\gamma > 0$, we obtain the following estimate:

$$\int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta \tilde{T}^2 (y, t_1) \, dy \le L_1 \|\Delta w\|_W^2 \qquad (22)$$

with $L_1 = \max \left\{ \frac{4\sigma}{\gamma \rho \tilde{c}_0} \varepsilon_{1*}^2 u_*^6, \frac{4\sigma}{\gamma \rho \tilde{c}_0} \varepsilon_{2*}^2 b_*^6 \right\}$. Here,

$$b_* = \min_{\Omega_t} b(t) > 0,$$
$$u_* = \min_{\Omega_t} u(t) > 0.$$

The Lemma 2 is proved. Thus, the second integral term in right-hand of (7) is obviously bounded by the term $o\left(\|\Delta w\|_W^2\right)$.

Finally, by definition of Fréchet -differential and equations (7), (11) and (22), we can conclude the **Theorem** 1 as following.

Theorem 1: Let lemmas 1, 2 hold. Then the cost functional is Fréchet -differential $J(w) \in C^1(W)$. Moreover, the Fréchet derivative of the cost functional $J(w)$ at $w \in W$ can be defined by the solution of the adjoint equation $p(y, t) \in H^{1.0}(\Omega_{yt})$ as the equation (23).

C. LIPSCHITZ CONTINUOUS

The main distinguished feature of the FOTD approach is that the gradient $J'(w)$ of the cost functional can be calculated by the adjoint equation (9). This suggests an idea of construction of the iteration method for approximate solution of minimization problem (4) based on the gradient formula (23), shown at the bottom of the page.

Using lemmas in the previous section, we can show that the gradient $J'(w)$ of the cost functional is Lipschitz continuous. We can also estimate the Lipschitz constant by solving the original and adjoint equations.

Lemma 3: Let condition of Theorem 1 hold. Then there is a constant L to make the following inequality hold:

$$\left\| J'(w + \Delta w) - J'(w) \right\|_W^2 \le L \|\Delta w\|_W^2, \qquad (24)$$

and the Lipschitz constant $L > 0$ is defined L, as shown at the bottom of the page.

The brief proof of this lemma is reported in the Appendix V.

Algorithm 1 Algorithm 1: ICG method for FOTD approach.

Begin

1: Set parameters $\phi = 0.5 \in (0, 1)$, $s = 0.2 \in (0, 1)$, $\delta = 1e^{-2}$, $\alpha = eps = 1e^{-6}$, $k = 0$. Choose an initial point $w_0 = \{u(t), b(t)\}$.

2: Solve the original parabolic PDE (4) and the adjoint PDE (9) based on the input value w_k, then we can obtain the $T_k(y, t)$ and $p_k(y, t)$.

3: Calculate the gradient of the cost functional g_k based on the (23), If $|g_k| \le eps$, then stop the algorithm, and give the final value $w^*(t) = w_{k+1}(t)$.

4: Evaluate the search direction d_k by equations $d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k g_{k-1} - \theta_k y_{k-1}, & k > 0, \end{cases}$ and $\beta_k = \frac{g_k y_{k-1}}{\|g_{k-1}\|^2}$, $\theta_k = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$, $y_{k-1} = g_k - g_{k-1}$..

5: Determine a step length α_k, where α_k satisfies the modified Armijo-type line search conditions.

$$J(w_k + \alpha_k d_k) \le J(w_k) - \delta \alpha_k^2 \| d_k \|^2,$$
$$\alpha_k = max\{\phi s^m, m = 0, 1, 2, \ldots\}. \qquad (26)$$

6: Update $w_{k+1} = w_k + \alpha_k d_k$, $k := k + 1$ and go to Step 2.

end

D. AN IMPROVED CONJUGATE GRADIENT ALGORITHM

Several optimization methods by FDTO approach have been extensively developed to solve the PDE-constrained optimization problem, such as steepest descent (SD) method [22], gradient method [22] and so on.

The basic concept of these optimization methods is to update w for each iteration via the gradient g_k of cost function or the value of cost function. For instance, the gradient g_k by FDTO approach with the sensitivity method [21] can be obtained in terms of the Jacobian matrix: $g_k = J'(w) = \begin{bmatrix} \frac{\partial J}{\partial w_1} & \frac{\partial J}{\partial w_2} & \cdots & \frac{\partial J}{\partial w_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial J}{\partial T} \frac{\partial T}{\partial w_1} & \frac{\partial J}{\partial T} \frac{\partial T}{\partial w_2} & \cdots & \frac{\partial J}{\partial T} \frac{\partial T}{\partial w_n} \end{bmatrix}$, where, w is the discrete form of $w(t)$: $w = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}$, and n is total number of discrete points. The first term $\frac{\partial J}{\partial T}$ is easy to obtain, the problem is to compute the second term $\frac{\partial T}{\partial w_i}$.
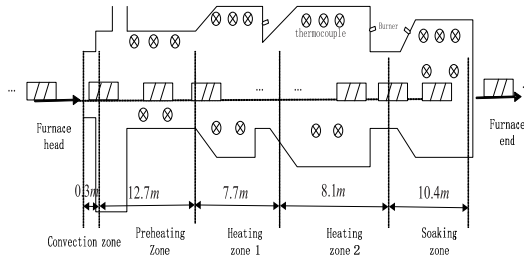
$$J'(w) = \left\{ 2\alpha u(t) + \frac{4\sigma}{\rho \tilde{c}_0} \varepsilon^+(t) u^3(t) p\left(\frac{l}{2}, t\right), 2\alpha b(t) + \frac{4\sigma}{\rho \tilde{c}_0} \varepsilon^-(t) b^3(t) p\left(-\frac{l}{2}, t\right) \right\}. \qquad (23)$$

$$where, \|J'(w + \Delta w) - J(w)\|_W^2 := \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \int_{t_0}^{t_1} \left[\left(\varepsilon^+(t)\right)^2 u^6(t) \Delta p^2\left(\frac{l}{2}, t; w\right) + \left(\varepsilon^-(t)\right)^2 b^6(t) \Delta p^2\left(-\frac{l}{2}, t; w\right) \right] dt$$

$$+ 2\alpha \int_{t_0}^{t_1} \left[\Delta u^2(t) + \Delta b^2(t) \right] dt + \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \int_{t_0}^{t_1} \left[\left(\varepsilon^+(t)\right)^2 p^2\left(\frac{l}{2}, t\right) u^4(t) \Delta u^2(t) + \left(\varepsilon^-(t)\right)^2 p^2\left(-\frac{l}{2}, t\right) b^4(t) \Delta b^2(t) \right] dt$$

$$(25)$$

$$L = \max \left\{ 2\alpha + \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{1*}^2 p_{1*}^4 u_*^4 + \frac{32\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{1*}^2 u_*^6 L_1, 2\alpha + \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{2*}^2 p_{2*}^4 b_*^4 + \frac{32\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{2*}^2 b_*^6 L_1 \right\}.$$

FIGURE 2. The structure of a walking beam reheating furnace.

The relationship between T and w_i is governed by nonlinear PDEs, and the analytic form of $\frac{\partial T}{\partial w_i}$ does not exist. Thus, the partial derivation at a given point can be approximated as: $\frac{\partial T}{\partial w_i} = \frac{T(w+\epsilon \mathbf{e}_i)-T(w)}{\epsilon}$. Here, ϵ is a small positive scalar and $\mathbf{e}_i$ is a vector, whose elements are all '0' expect for only one '1' in the ith position. In general, this process requires evaluation of T based on w and the n perturbed points $T(w+\epsilon \mathbf{e}_i)$. Therefore, $n+1$ times solution of nonlinear PDEs are required to calculate the Jacobian matrix. This is an exceedingly large computation task.

In comparison, the gradient $J'(w)$ of the cost functional by the FOTD approach can be written as the equation (23), which is described via the weak solution of the adjoint equation. To obtain the fast convergence speed and the high accuracy, an improved conjugate gradient algorithm with sufficient descent property is proposed for numerical solution. The flowchart is shown in Algorithm 1.

IV. EXPERIMENT RESULTS

In order to prove the reliability and effectiveness of the proposed method, some simulations are performed and shown in the following sections. All simulation computations in the following are worked on a standard PC (3.4 GHz, 8 GB RAM) and performed in "MATLAB" (R2018b) software.

A. VERIFY THE PROPOSED MATHEMATICAL MODEL

In this section, the experiments are given to determine whether the proposed model(PModel) in Section II can provide reliable results. Here, the constant mathematical model(CModel), in which material parameters are constant, is introduced here as another comparative object. Because, the CModel is often used by most researchers for the convenience of computing.

Firstly, the data measured by thermocouples experiments from the real reheating furnace are given to verify the proposed mathematical model(PModel). The measured data used in the experiment examples are obtained from a walking beam (WB) slab reheating furnace in Angang Group Corporation Limited. The basic structure of the walking-beam type reheating furnace is indicated in Fig. 2. It has the effective length 40 m. The effective lengths of each zone are 0.3 m, 12.7 m, 7.7 m, 8.1 m and 10.4 m, respectively.

As shown in Fig. 3, the data acquisition unit is placed in a water cooled box, which is wrapped with insulation material.

FIGURE 3. The state of the test slab before charging into the furnace.

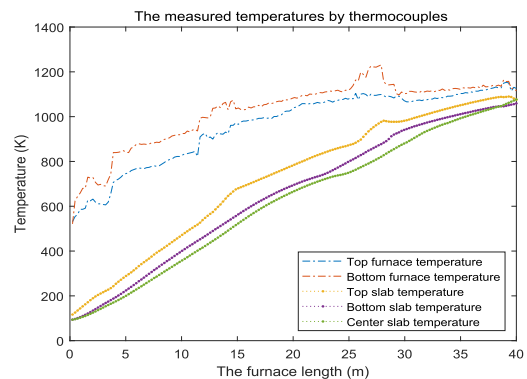**FIGURE 4.** The state of the test slab after discharging out from the furnace.

FIGURE 5. The measured temperatures by thermocouples from the test slab.

The thermocouples (Type K, $\phi = 15mm$) are assembled in the trial slab at different locations to record the slab temperatures every 20 seconds. The data saved in the temperature recorder are exported when the test slab with the mounted thermocouples is discharged from the reheating furnace, as shown in Fig. 4.

Finally, the measured top, center and bottom temperatures of test slab and the corresponding furnace temperatures are plotted in Fig. 5.

The simulation results are compared and shown in Fig. 6, 7 and 8. Here, some definitions are given in advance. The symbols $T_m^{t/c/b}$ are the top, center and bottom measured
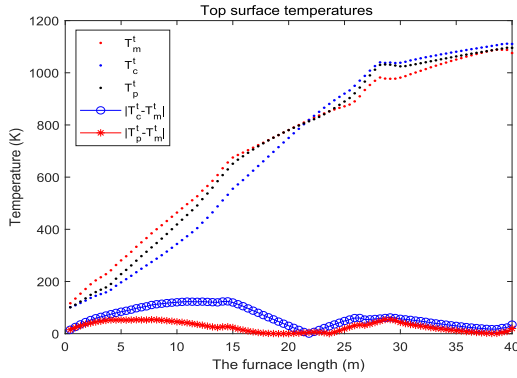
FIGURE 6. The top surface temperature curves and corresponding temperature differences.
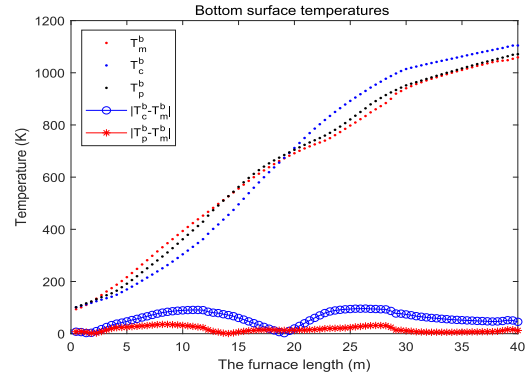

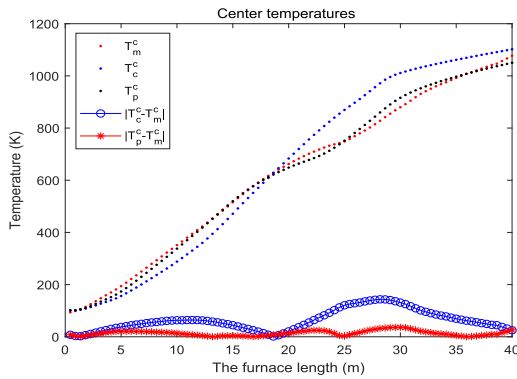**FIGURE 8.** The bottom surface temperature curves and corresponding temperature differences.

FIGURE 7. The center temperature curves and corresponding temperature differences.

temperatures obtained by the thermocouples from the test slab. The symbols $T_c^{t/c/b}$ are the top, center and bottom calculated temperatures obtained by solving the CModel base on the corresponding furnace temperatures in in Fig. 5. The symbols $T_p^{t/c/b}$ are the top, center and bottom calculated temperatures obtained by solving the PModel. Thus, $|T_c^{t/c/b} - T_m^{t/c/b}|$ represent the top, center and bottom temperature deviations between the measured value and the calculated values by the CModel. $|T_p^{t/c/b} - T_m^{t/c/b}|$ represent the top, center and bottom different temperatures between the measured slab temperature and the calculated values by the PModel.

In Fig. 6, the black dotted curve, which represents top surface temperature trajectory from the PModel, is almost the same with the red dotted curve from the measured thermocouple. However, the blue dotted curves obtained by the CModel deviates significantly from the red dotted curve in the range of 8-18 m. In Fig. 6, it can be obtained that the maximum top temperature deviations of $|T_c^t - T_m^t|$ is 123.82, which is much higher than 53.98 form $|T_p^t - T_m^t|$.

In Fig. 7, the blue dotted curve, which represents center temperature trajectory from the CModel, is deviates significantly from the red dotted curve from the measured thermocouple. Here, the blue dotted curve in the range of 10-20 m is obviously lower than red dotted curve, while

it appears obvious convex hull phenomenon in the range of 25-30 m. On the contrary, the black dotted curve follows closely along the red dotted curve. The maximum center temperature deviations of $|T_c^c - T_m^c|$ and $|T_p^c - T_m^c|$ results are 143.59 and 36.16 respectively.

The bottom surface temperature trajectories in Fig. 8 are almost same as the center temperature trajectories in Fig. 7. Here, the maximum bottom temperature deviations of $|T_c^b - T_m^b|$ and $|T_p^b - T_m^b|$ results are 96.18 and 35.55 respectively.

Finally, three figures of this experiment cab prove that the proposed mathematical model(PModel) in Section II can provide more accuracy results.

B. REFERENCE TRAJECTORY EXPERIMENTS
In order to compare the performance of different optimization strategies, two different comparative experiments are listed as follow. Notice, the PDE model used in the following experiments is the proposed mathematical model (PModel).

1) FOTD APPROACH VS FDTO APPROACH
Here, the first-discretize-then-optimize(FDTO) approach is introduced as the comparative object. In the FDTO approach, the explicit finite difference method [23], which has second-order accuracy in space and first order accuracy in time, is used to discretize the PDE model and the optimization problem, then the KKT condition is obtained. The improved conjugate gradient method in Section III-C is also used here for numerical solution. The numerical values of the computational results are given in Fig. 9 and Table 2.

Combing Fig. 9 and Table 2, it is clear that the FDTO approach can almost obtain the same value of the final cost functional J as the FOTD approach. The iteration steps is 21 for the FDTO approach and 22 for the FOTD approach. However, the simulation time of FDTO approach is 370.52 seconds, which is a hundred times larger than 3.70 seconds by the FOTD approach. As mentioned above, $n + 1$ times solution of nonlinear PDEs are required to calculate the Jacobian matrix for the FDTO approach. This is an exceedingly large computation task. Thus, the FDTO
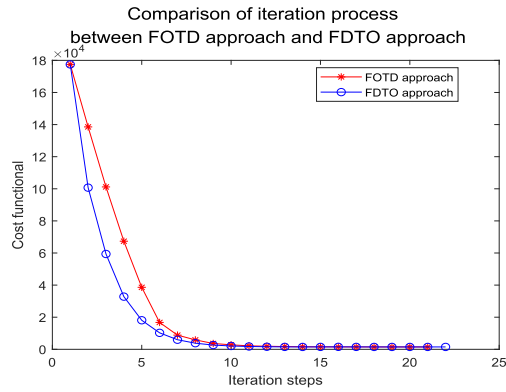
FIGURE 9. Comparison of iteration process between FOTD approach and FDTO approach.

TABLE 2. Comparison of simulation performance between FOTD approach and FDTO approach.

	simulation time (s)	iteration steps	Final cost functional (J)
FOTD	3.70	21	1481.78
FDTO	370.52	22	1488.16

FIGURE 10. Comparison of iteration process by two different gradient method.

approach cost almost a hundred times larger in simulation time than the FOTD approach.

2) STEEPEST descent(SD) METHOD VS IMPROVED CONJUGATE gradient(ICG) METHOD

In order to prove the effectiveness of the improved conjugate gradient method, the steepest descent method is introduced as the comparative object. Obviously, the FOTD approach has better performance than FDTO approach from the above experiment. Thus, the two different gradient method are compared by applying the same FOTD approach in this experiment. The numerical values of the computational results are given in Fig. 10, 11 and Table 3.

Obviously, the iteration behavior (value of the cost functional J) of the gradient methods consists of three same phases: the initial phase of rapid decrease but very short duration, the second phase of slow decrease, and the third phase of almost constant behavior. At the initial phase, the improved
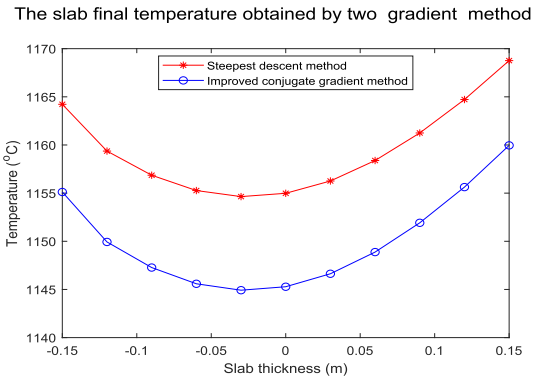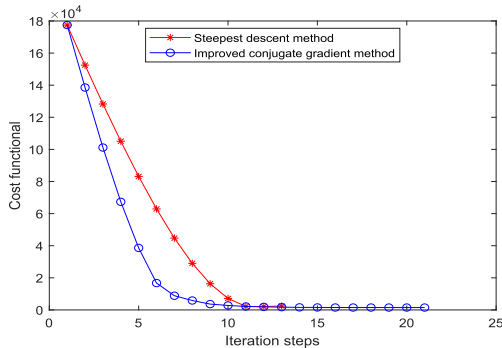

FIGURE 11. The slab's final temperature obtained by two gradient method.

TABLE 3. Comparison of simulation performance between ICG method and SD method.

	simulation time (s)	iteration steps	Final cost functional (J)
SD	8.09	13	2463.48
ICG	3.70	21	1481.78

conjugate gradient method has more rapid decrease than the steepest descent method as shown in Fig. 10. At the second phase, the improved conjugate gradient method seems somehow to be the poor relation of the steepest descent method. However, at the third phase, the final cost functional J converges to 1481.78 by the improved conjugate gradient method, which performs much better than the steepest descent method (2463.48). This can also be seen in Fig. 11. The slab's final temperatures obtained by the improved conjugate gradient method are centered better on the desired discharging temperature $T^* = 1150°C$. In general, the two gradient methods have different convergent performances. The improved conjugate gradient method has sufficient descent property and has better convergent result than the Steepest descent method. Finally, we can conclude that the improved conjugate gradient algorithm has a better performance.

V. CONCLUSION
In order to solve PDE-constrained optimization problem for the reheating furnace, this paper applies first-optimization-then-discrete method to this optimal problem. More accurate mathematical PDE model is built for the reheating furnace. Then, we prove that the gradient of cost function is Lipschitz continuous. Base on this, we propose an improved conjugate algorithm to solve this problem. Finally, some experiments are given and the simulation results prove that the proposed algorithm is a good choice.

APPENDIX A
PROOF OF LEMMA 3
Proof. According to the formula (9), it is easy to obtain that $\Delta p = p(y, t; w + \Delta w) - p(y, t; w)$ will be the weak solution

$$2 \int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta \tilde{T}^2 (y, t_1) \, dy$$

$$= \frac{1}{2} \int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta p^2 (y, t_0) \, dy + \frac{1}{\rho \tilde{c}_0} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{t_0}^{t_1} \tilde{\lambda} (T) \left(\frac{\partial \Delta p}{\partial y} (y, t) \right)^2 dt dy$$

$$+ 4\sigma \int_{t_0}^{t_1} \varepsilon^+ (t) T^3 \left(\frac{l}{2}, t \right) \Delta p^2 \left(\frac{l}{2}, t \right) dt + 12\sigma \int_{t_0}^{t_1} \varepsilon^+ (t) p \left(\frac{l}{2}, t \right) T^2 \left(\frac{l}{2}, t \right) \Delta T \left(\frac{l}{2}, t \right) \Delta p \left(\frac{l}{2}, t \right) dt$$

$$+ 4\sigma \int_{t_0}^{t_1} \varepsilon^- (t) T^3 \left(-\frac{l}{2}, t \right) \Delta p^2 \left(-\frac{l}{2}, t \right) dt + 12\sigma \int_{t_0}^{t_1} \varepsilon^- (t) p \left(-\frac{l}{2}, t \right) T^2 \left(-\frac{l}{2}, t \right) \Delta T \left(-\frac{l}{2}, t \right) \Delta p \left(\frac{-l}{2}, t \right) dt \qquad (28)$$

$$\int_{t_0}^{t_1} \Delta p^2 \left(-\frac{l}{2}, t \right) dt \leq 2 \int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta \tilde{T}^2 (y, t_1) \, dy, \quad \int_{t_0}^{t_1} \Delta p^2 \left(\frac{l}{2}, t \right) dt \leq 2 \int_{-\frac{l}{2}}^{\frac{l}{2}} \Delta \tilde{T}^2 (y, t_1) \, dy \qquad (29)$$

$$\frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \int_{t_0}^{t_1} \left[\left(\varepsilon^+ (t) \right)^2 u^6 (t) \Delta p^2 \left(\frac{l}{2}, t; w \right) + \left(\varepsilon^- (t) \right)^2 b^6 (t) \Delta p^2 \left(-\frac{l}{2}, t; w \right) \right] dt \leq L_2 \| \Delta w \|_W^2 \qquad (30)$$

$$2\alpha \int_{t_0}^{t_1} \left[\Delta u^2 (t) + \Delta b^2 (t) \right] dt + \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \int_{t_0}^{t_1} \left[\left(\varepsilon^+ (t) \right)^2 p^2 \left(\frac{l}{2}, t \right) u^4 (t) \Delta u^2 (t) + \left(\varepsilon^- (t) \right)^2 p^2 \left(-\frac{l}{2}, t \right) b^4 (t) \Delta b^2 (t) \right] dt$$

$$\leqslant L_3 \| \Delta w \|_W^2, \qquad (31)$$

of the following parabolic problem (27):

$$\frac{\partial \Delta p}{\partial t} (y, t) + \frac{\partial}{\partial y} \left(\frac{\tilde{\lambda} (\tilde{T})}{\rho \tilde{c}_0} \frac{\partial \Delta p}{\partial y} (y, t) \right) = 0,$$

$$2\Delta \tilde{T} (y, t_1) + \Delta p (y, t_1) = 0,$$

$$\tilde{\lambda} (T) \frac{\partial \Delta p}{\partial y} \left(-\frac{l}{2}, t \right) - 4\sigma \varepsilon^- (t) T^3 \left(-\frac{l}{2}, t \right) \Delta p \left(-\frac{l}{2}, t \right)$$

$$-12\sigma \varepsilon^- (t) p \left(-\frac{l}{2}, t \right) T^2 \left(-\frac{l}{2}, t \right) \Delta T \left(-\frac{l}{2}, t \right) = 0,$$

$$\tilde{\lambda} (T) \frac{\partial \Delta p}{\partial y} \left(\frac{l}{2}, t \right) + 4\sigma \varepsilon^+ (t) T^3 \left(\frac{l}{2}, t \right) \Delta p \left(\frac{l}{2}, t \right)$$

$$+12\sigma \varepsilon^+ (t) p \left(\frac{l}{2}, t \right) T^2 \left(\frac{l}{2}, t \right) \Delta T \left(\frac{l}{2}, t \right) = 0. \qquad (27)$$

We multiply both sides of (27) by $\Delta p (y, t; w)$, integrating over Ω_{yt}, and use the boundary conditions of the (27), the following identity can be obtained:

This identity (28) implies the following two inequalities in (29), as shown at the top of the page.

Then, from the Lemma 3.2, we can conclude in (30), as shown at the top of the page.

where $L_2 = \max \left\{ \frac{32\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{1*}^2 u_*^6 L_1, \frac{32\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{2*}^2 b_*^6 L_1 \right\}$.

Now, we can consider the second and third part of the right hand side of (25), as shown at the bottom of page 7. It is easy to obtain the following result in (31), as shown at the top of the page.

where, $L_3 = \max \left\{ 2\alpha + \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{1*}^2 p_{1*}^2 u_*^4, 2\alpha + \frac{16\sigma^2}{\rho^2 \tilde{c}_0^2} \varepsilon_{2*}^2 p_{2*}^2 b_*^4 \right\}$. Here, $p_{1*} = \min_{\Omega_{yt}} p \left(\frac{l}{2}, t \right) > 0, p_{2*} = \min_{\Omega_{yt}} p \left(-\frac{l}{2}, t \right) > 0$. Finally, combining (30) and (31), the Lemma 3 is proved.

REFERENCES

[1] X. M. Nguyen, F. Lawayeb, P. Rodriguez-Ayerbe, D. Dumur, and A. Mouchette, "Nonlinear model predictive control of steel slab walking-beam reheating furnace based on a numerical model," in *Proc. IEEE Conf. Control Appl. (CCA)*, Oct. 2014, pp. 191–196.

[2] J. Sun, L. Liu, and B. Chen, "Study on the optimization control strategy of reheating furnace temperature," in *Proc. 2nd Int. Conf. Appl. Robot. Power Ind. (CARPI)*, Sep. 2012, pp. 252–255.

[3] G. R. Hurd, J. Kaufman, H. C. Wu, J. Ward, and E. Rodriguez, "Process control and automation systems advancements for reheat furnaces," *Iron Steel Technol.*, vol. 8, no. 9, pp. 67–74, 2011.

[4] H. Chen and Q. Huang, "Preconditioned iterative method for boundary value method discretizations of a parabolic optimal control problem," *Calcolo*, vol. 57, no. 1, pp. 1–25, Mar. 2020.

[5] B. Li, J. Liu, and M. Xiao, "A new multigrid method for unconstrained parabolic optimal control problems," *J. Comput. Appl. Math.*, vol. 326, pp. 358–373, Dec. 2017.

[6] M. Darehmiraki, A. Rezazadeh, and A. Ahmadian, "An artificial neural network-based method for the optimal control problem governed by the fractional parabolic equation," *Numer. Methods Partial Differ. Equ.*, vol. 37, no. 3, pp. 2296–2316, May 2021.

[7] Z. Yang and X. Luo, "Optimal set values of zone modeling in the simulation of a walking beam type reheating furnace on the steady-state operating regime," *Appl. Thermal Eng.*, vol. 101, pp. 191–201, May 2016.

[8] Y. Wang, X. Luo, Y. Yu, and H. Cui, "Optimal control of two-dimensional parabolic partial differential equations with application to steel billets cooling in continuous casting secondary cooling zone," *Optim. Control Appl. Methods*, vol. 37, no. 6, pp. 1314–1328, Nov. 2016.

[9] X. Luo and Z. Yang, "Dual strategy for 2-dimensional PDE optimal control problem in the reheating furnace," *Optim. Control Appl. Methods*, vol. 39, no. 2, pp. 981–996, Mar. 2018.

[10] A. Rezazadeh, M. Mahmoudi, and M. Darehmiraki, "Space–time spectral collocation method for one-dimensional PDE constrained optimisation," *Int. J. Control*, vol. 93, no. 5, pp. 1231–1241, May 2020.

[11] S. Güttel and J. W. Pearson, "A rational deferred correction approach to parabolic optimal control problems," *IMA J. Numer. Anal.*, vol. 38, no. 4, pp. 1861–1892, Oct. 2018.

[12] S. Cipolla and F. Durastante, "Fractional PDE constrained optimization: An optimize-then-discretize approach with L-BFGS and approximate inverse preconditioning," *Appl. Numer. Math.*, vol. 123, pp. 43–57, Jan. 2018.

[13] J. Liu and Z. Wang, "Non-commutative discretize-then-optimize algorithms for elliptic PDE-constrained optimal control problems," *J. Comput. Appl. Math.*, vol. 362, pp. 596–613, Dec. 2019.

[14] A. Steinböck, *Model-Based Control and Optimization of a Continuous Slab Reheating Furnace*, vol. 12. Aachen, Germany: Shaker Verlag, 2011.

[15] R. B. Bird, "Transport phenomena," *Appl. Mech. Rev.*, vol. 55, no. 1, pp. R1–R4, 2002.

[16] A. Steinboeck, D. Wild, T. Kiefer, and A. Kugi, "A fast simulation method for 1D heat conduction," *Math. Comput. Simul.*, vol. 82, no. 3, pp. 392–403, Nov. 2011.

[17] K. Chen, H. L. Ke, L. He, and Y. H. Peng, "A novel numerical model for billet reheating furnace," *Ironmaking Steelmaking*, vol. 44, no. 6, pp. 395–402, Jul. 2017.

[18] X. Luo and Z. Yang, "A new approach for estimation of total heat exchange factor in reheating furnace by solving an inverse heat conduction problem," *Int. J. Heat Mass Transf.*, vol. 112, pp. 1062–1071, Sep. 2017.

[19] H. D. Baehr and K. Stephan, *Heat and Mass Transfer*. Berlin, Germany: Springer, 2006.

[20] C. Wanli, K. Ning, and W. Daohong, "Dynamic furnace temperature setting research on combustion system of rolling mill reheating furnace," *Energy Procedia*, vol. 66, pp. 217–220, Jan. 2015.

[21] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization With PDE Constraints*, vol. 23. Dordrecht, The Netherlands: Springer, 2008.

[22] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *Traces and Emergence of Nonlinear Programming*. Basel, Switzerland: Springer, 2014, pp. 247–258.

[23] G. W. Recktenwald, "Finite-difference approximations to the heat equation," *Mech. Eng.*, vol. 10, no. 1, pp. 1–27, 2004.

MING LIU received the B.S. degree from the Shandong University of Science and Technology, Qingdao, China, in 2019. He is currently pursuing the M.S. degree with the Qilu University of Technology (Shandong Academy of Sciences), Jinan, China.

His current research interests include modeling and optimization for the complex industrial systems, parallel programming and applications, and intelligent optimization methods.

ZHI YANG received the B.S. degree from Shandong University, Jinan, China, in 2009, the M.S. degree from the Shenyang University of Chemical Technology, Shenyang, China, in 2013, and the Ph.D. degree from Northeastern University, Shenyang, in 2019.

He is currently a Lecturer with the Qilu University of Technology (Shandong Academy of Sciences), Jinan. His current research interests include modeling and optimization for the complex industrial systems, parallel programming and applications, and intelligent optimization methods.
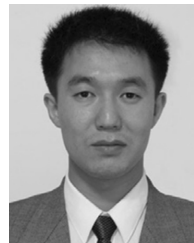
XIAOCHUAN LUO (Member, IEEE) received the M.Eng. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 1999 and 2002, respectively.

From 2002 to 2004, he held a postdoctoral position with the University of Technology of Troyes, Troyes, France. He is currently an Associate Professor with Northeastern University, Shenyang, China. His current research interests include modeling and optimization of processing industry manufacturing systems, production planning and scheduling, and optimization algorithms.

Mr. Luo is also a Peer-Reviewed Expert of the National Natural Science Foundation of China, an IEEE Reviewer, *IJPR*, *EJIE*, and other international journal reviewers. He was selected into the 2008 Ministry of Education's New Century Excellent Talents Support Program, and he was selected into the "Million Talents Project" in Liaoning Province, in 2009.

• • •