Applied Mathematics, 2014, 5, 691-706
Published Online March 2014 in SciRes. http://www.scirp.org/journal/am
http://dx.doi.org/10.4236/am.2014.54067

Scientific Research

A Consecutive Quasilinearization Method for the Optimal Boundary Control of Semilinear Parabolic Equations

Mohammad Dehghan Nayyeri*, Ali Vahidian Kamyad

Applied Mathematics Department, Ferdowsi University of Mashhad, Mashhad, Iran
Email: *mu_de324@stu-mail.um.ac.ir, mudenayyeri@gmail.com, kamyad@math.um.ac.ir

Received 15 December 2013; revised 15 January 2014; accepted 23 January 2014

Abstract

Optimal boundary control of semilinear parabolic equations requires efficient solution methods in applications. Solution methods bypass the nonlinearity in different approaches. One approach can be quasilinearization (QL) but its applicability is locally in time. Nonetheless, consecutive applications of it can form a new method which is applicable globally in time. Dividing the control problem equivalently into many finite consecutive control subproblems they can be solved consecutively by a QL method. The proposed QL method for each subproblem constructs an infinite sequence of linear-quadratic optimal boundary control problems. These problems have solutions which converge to *any* optimal solutions of the subproblem. This implies the uniqueness of optimal solution to the subproblem. By merging solutions to the subproblems, the solution of original control problem is obtained and its uniqueness is concluded. This uniqueness result is new. The proposed consecutive quasilinearization method is numerically stable with convergence order at least linear. Its consecutive feature prevents large scale computations and increases machine applicability. Its applicability for globalization of locally convergent methods makes it attractive for designing fast hybrid solution methods with global convergence.

Keywords

Quasilinearization; Optimal Boundary Control; SQP Methods; Semilinear Parabolic PDE's

1. Introduction

The solution methods for the optimal control of nonlinear systems pass from nonlinearity to linearity in different

*Corresponding author.

How to cite this paper: Nayyeri, M.D. and Kamyad, A.V. (2014) A Consecutive Quasilinearization Method for the Optimal Boundary Control of Semilinear Parabolic Equations. *Applied Mathematics*, **5**, 691-706.
http://dx.doi.org/10.4236/am.2014.54067

approaches. For example the gradient methods modify iteratively the previous approximate solution by linearly seeking a suitable direction thorough solving a linear problem [1]. The SQP methods seek the optimal solution by linearizing the optimality systems using some version of Newton's method [1] [2]. Our approach in this respect is to linearize the state equation through a quasilinearization method.

Quasilinearization method for nonlinear equations has its origin in the theory of dynamic programming and has important features in common with Newton's method especially its form [3]. For a formal explanation let Y and Z be ordered Banach spaces, $L:Y\to Z$ be a bounded linear operator and $N:Y\to Z$ be a nonlinear differentiable operator. Consider the equation

$$Ly = N(y). \tag{1.1}$$

In the convex case N, Equation (1.1) can be written as

$$Ly = N(y) = \max_{v\in Y} N(v) + N'(v)(y-v), \tag{1.2}$$

where the right hand side is quasilinear. Then starting from $y_0 \in Y$ thorough the quasilinearization method, a sequence of linear equations is defined

$$L(y_{n+1}) = N(y_n) + N'(y_n)(y_{n+1} - y_n), \tag{1.3}$$

which produces the sequence of approximate solutions $\{y_n\}$ in Y, converging to y, the solution of (1.2) or (1.1); see [3] [4]. This method has the following features. 1) $\{y_n\}$ is monotonic. This stems from positivity and inverse positivity of L and L^{-1}. 2) the convergence is globally in the sense that y_0 can be any lower solution of (1.1), *i.e.* $Ly_0 \le N(y_0)$. 3) The rate of convergence is quadratic. For details on these features refer to [3] [5]. There are some extensions, refinements and generalizations to the quasilinearization method which preserve the above features but relax the convexity assumption on N; for a complete survey see [5]-[7]. Quasilinearization method has intimate connection with the theory of positive and monotone operators, maximum operation and differential inequalities; confer [8], Sec. 4.33; [4] [9].

In order to introduce the proposed consecutive quasilinearization method for optimal control problems let U be a Banach space, $J:Y\times U\to\mathbb{R}$ be a functional and $B:Y\to U$ be a bounded linear boundary operator. Consider the following optimal boundary control problem:

$$\begin{aligned} &\min_{y,u} J(y(t),u(t)) \\ &Ly(t) = N(y(t)), \\ &By(t) = u(t), u\in U\subseteq \mathbb{U}, \end{aligned} \tag{1.4}$$

where t belongs to a time interval $[0,T]$. For $v\in Y$ consider the following approximation to (1.4):

$$\begin{aligned} &\min_{y,u} J(y(t),u(t)) \\ &Ly(t) = N(v(t)) + N'(v(t))(y(t)-v(t)), \\ &By(t) = u(t), u\in U\subseteq \mathbb{U}. \end{aligned} \tag{1.5}$$

Starting from y_0, let (y_n,u_n) be the optimal solution of (1.5) with $v = y_{n-1}$. Then the sequence $\{(y_n,u_n)\}$ converges to the solution of (1.4) with the following features: 1) The convergence is occurred for $T < T_1$, for some $T_1 > 0$. 2) The convergence is globally in the sense that y_0 can be chosen any element in a subspace of Y. 3) The rate of convergence is at least linear but it is not necessarily super-linear or quadratic. Here the sequence $\{y_n\}$ or $\{u_n\}$ is not necessarily monotonic even when N is convex or concave. For the case $T > T_1$ the optimal control problem is decomposed into many finite optimal control subproblems each on a time interval with length less than some T_2 and then the above method be applied to each of them consecutively. Here $T_2 < T_1$ is such that the stability is preserved.

The optimal boundary control problem which is investigated has the standard quadratic objective of tracking type and a state constraint comprised of a semilinear parabolic equation with mixed boundary type. For such control problems, due to lack of convexity of the solution set, there is no general uniqueness result based on the optimality theory of optimal control problems [1] [2] [10]. However, a uniqueness result for such problems is obtained here as a by-product of the convergence of proposed consecutive quasilinearization method.

The organization of paper is as follows. Section 2 introduces the state equation and some estimates concerning solution of linear initial-boundary value problems. Section 3 proves the existence of an optimal solution. Section 4 introduces the quasilinearization method and proves its convergence for $T < T_1$. Section 5 explains how to apply the quasilinearization method consecutively to the optimal boundary control problem when $T > T_1$. Also the uniqueness of optimal solution is stated there. In Section 6 the error and stability analysis of consecutive quasilinearization method is investigated. Section 7 presents a numerical example concerning the obtained results.

2. The State Equation

Let Ω be an open bounded domain in $\mathbb{R}^k$, $k \ge 2$, with boundary $\partial\Omega$ of class $C^{2,\beta}$ for some $\beta \in (0,1]$. Let $T > 0$, $Q = (0,T)\times\Omega$ and $\Sigma = [0,T]\times\partial\Omega$. Consider the control system described by the semilinear parabolic initial-boundary value problem:

$$\begin{aligned} &y_t(t,x) + Ay(t,x) = f(t,x,y(t,x)), (t,x)\in Q,\\ &\partial_{\nu A} y(t,x) + cy(t,x) = u(t,x), (t,x)\in\Sigma,\\ &y(0,x) = y_0(x), x\in\Omega, \end{aligned} \tag{1.6}$$

where $y(t,x)$ and $u(t,x)$ are, respectively, the state and the distributed control of system, $f(t,x,y(t,x))$ is the system nonlinearity and

$$\partial_{\nu A} y(t,x) = \sum_{i,j=1}^{k} a_{ij}(x)\partial_i y(t,x)\nu_i(x), (t,x)\in\Sigma,$$

is the normal derivative of y associated with A wherein $\nu = (\nu_1 \cdots \nu_k)$ is the outward unit normal to $\partial\Omega$.

The following assumptions are imposed on the system and data:

- (A1) A is a secod order differential operator in divergence form:

$$Ay(t,x) := -\sum_{i,j=1}^{k} \partial_j\left(a_{ij}(x)\partial_i y(t,x)\right), \tag{1.7}$$

where $a_{ij} \in C^{1,\beta}(\Omega)$ and A is uniformly elliptic, i.e. for every $\xi = (\xi_1,\cdots,\xi_k)' \in \mathbb{R}^k$,

$$\sum_{i,j=1}^{k} a_{ij}(x)\xi_i\xi_j \ge \mu|\xi|^2, x\in\Omega, \tag{1.8}$$

for some $\mu > 0$. Also it is considered $c \in L^\infty(\Sigma)$.

- (A2) $f : Q\times\mathbb{R}\to\mathbb{R}$ satisfies Caratheodory's condition, *i.e.* f is measurable on Q and continuous on $\mathbb{R}$, and the Nemytskii's operator $F : L^2(Q)\to L^2(Q)$, defined by $Fy := f(t,x,y(t,x)), (t,x)\in Q, y\in L^2(Q)$, is bounded and continuous. A sufficient condition for that is

$$|f(t,x,\xi)| \le C|\xi| + k(t,x), \text{a.e.}(t,x)\in Q, \xi\in\mathbb{R}, \tag{1.9}$$

for some $k \in L^2(Q)$, Theorem 3.2 in [11].

- (A3) f is twice continously differentiable with respect to y and

$$\begin{aligned} &\left|f_y(t,x,\xi)\right| \le M_y, \text{a.e.}(t,x)\in Q, \xi\in\mathbb{R},\\ &\left|f_{yy}(t,x,\xi)\right| \le M_{yy}, \text{a.e.}(t,x)\in Q, \xi\in\mathbb{R}, \end{aligned}$$

for constants M_y, $M_{yy} > 0$.

The standard function spaces $H := L^2(\Omega)$, $V := H^1(\Omega)$ and $V' := H^1(\Omega)'$ are used in the paper. Identifying H with its dual H' results in the evolution triples $V \hookrightarrow H \hookrightarrow V'$, where the embeddings are dense, continuous and compact. The standard solution space of parabolic problems and its norm is defined as

$$\begin{aligned} &W := \left\{y\in L^2(0,T;V) \,\middle|\, y_t \in L^2(0,T;V')\right\},\\ &\|y\|_W := \|y\|_{L^2(0,T;V)} + \|y_t\|_{L^2(0,T;V')}. \end{aligned}$$

The continuous embeddings $W \hookrightarrow C([0,T];H)$ and $W \hookrightarrow L^2(0,T;L^2(Q)) \equiv L^2(Q)$ are well-known and the latter is compact. For detailed definitions and properties of the above spaces refer to [11]-[14].

The bilinear form associated with Equation (1.6) is defined as follows:

$$B[y,\phi] := \sum_{i,j=1}^{k} \int_Q a_{ij} \partial_i y \partial_j \phi \mathrm{dxdt} + \int_\Sigma c y \phi \mathrm{dxdt}, \quad y,\phi \in L^2(0,T;V). \tag{1.10}$$

By Assumption (A1) the coefficients in (0.10) are bounded. This results in the boundedness of $B[\cdot,\cdot]$. Let $\langle \cdot,\cdot \rangle$ denotes the duality pairing between $L^2(0,T;V)$ and its dual $L^2(0,T;V')$, and $(\cdot,\cdot)$ and $(\cdot,\cdot)_{L^2(\Sigma)}$ denote respectively the inner product of $L^2(Q)$ and $L^2(\Sigma)$. Then

Definition 1 *Let* $y_0 \in H$. *For* $u \in L^2(\Sigma)$ *a function* $y \in W$ *is called a weak solution of* (0.6) *if* $y(0) = y_0$ *and*

$$\langle y_t, \phi \rangle + B[y,\phi] = (Fy,\phi) + (u,\phi)_{L^2(\Sigma)},$$

for all $\phi \in L^2(0,T;V)$.

Next theorem under weaker assumptions is proved in Theorem 3.1 of [10].

Theorem 1 *Let* $y_0 \in L^\infty(\Omega)$. *Then under Assumptions* (A1)-(A3) *for every* $u \in L^\infty(\Sigma)$ *problem* (0.6) *admits a unique weak solution* $y_u \in W \cap L^\infty(Q)$.

In the following sections the linear initial-boundary value problems of the type below are used:

$$\begin{aligned} & y_t + Ay = dy + h, \text{ in } Q, \\ & \partial_{\nu A} y + cy = g, \text{ on } \Sigma, \\ & y(0) = y_0, \text{ in } \Omega, \end{aligned} \tag{1.11}$$

where $d \in L^\infty(Q)$, $h \in L^2(Q)$, $c \in L^\infty(\Sigma)$, $g \in L^2(\Sigma)$ and $y_0 \in H$. Define the family of bilinear forms $B[t;\cdot,\cdot]: V \times V \to \mathbb{R}$, a.e. $t \in [0,T]$, by

$$B[t;w,v] := \sum_{i,j=1}^{k} \int_\Omega a_{ij} \partial_i w \partial_j v \mathrm{dx} + \int_{\partial\Omega} c(t,\cdot) wv \mathrm{dx}, \quad w,v \in V, \text{ a.e. } t \in [0,T]. \tag{1.12}$$

By Assumption (A1) the coefficients in (0.12) are bounded for a.e. t. Thus $B[t;\cdot,\cdot]$ is bounded on $V \times V$ for a.e. $t \in [0,T]$. Let $\langle \cdot,\cdot \rangle_{V',V}$ be the duality pairing between V and its dual V', and $(\cdot,\cdot)_H$ and $(\cdot,\cdot)_{L^2(\partial\Omega)}$ be the inner product of $H = L^2(\Omega)$ and $L^2(\partial\Omega)$.

Definition 2 *A function* $y \in W$ *is called a weak solution of* (0.11) *if* $y(0) = y_0$ *and*

$$\langle y_t, v \rangle_{V',V} + B[t;y,v] = (dy,v)_H + (h,v)_H + (g,v)_{L^2(\partial\Omega)},$$

for all $v \in V$ *and a.e.* $t \in [0,T]$.

Norm estimates concerning solution of problem (1.11) are common in the literature of linear initial-boundary value problems [12]-[14]. Next theorem states some of them clarifying the time dependency quality of their constants. Its proof has been included due to lack of suitable reference, on the best of our knowledge, for the form stated here.

Theorem 2 *The initial-boundary value problem* (0.11) *has a unique weak solution* $y \in W$ *with norm estimates*

$$\|y\|_W \le C_W(T)\left(\|y_0\|_H + \|h\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)}\right), \tag{1.13}$$

$$\|y\|_{L^2(Q)} \le C(T)\left(\|y_0\|_H + \|h\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)}\right), \tag{1.14}$$

$$\|y\|_{C([0,T];H)} \le C_0(T)\left(\|y_0\|_H + \|h\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)}\right), \tag{1.15}$$

where $C_W(T)$ *and* $C_0(T)$ *are bounded when* T *varies boundedly and* $\lim_{T\to 0^+} C(T) = 0^+$. *If* $h \in L^\infty(Q)$, $g \in L^\infty(\Sigma)$ *and* $y_0 \in L^\infty(\Omega)$ *then* $y \in L^\infty(Q)$.

Proof 1 *By Theorem* 5.3 *in* Ch. III of [15] *for some* $K > 0$ *the following estimate exists*:

$$\|v\|_H^2 \le K\|v\|_{L^2(\partial\Omega)}^2 + K^2\|\nabla v\|_H^2,\ v \in V.$$

Using the Garding inequality (Proposition 22.45 in [14]) or the elliptic energy estimates (Sec. 6.2.2, Theorem 2 in [12]) it is obtained

$$\beta\|v\|_V \le B[t;v,v] + (dv,v)_H + \gamma\|v\|_H, v \in V, \tag{1.16}$$

for a.e. $t \in [0,T]$, where $\beta > 0$ and $\gamma \ge 0$. Then the existence of a unique weak solution $y \in W$ of (1.11) which satisfies the estimate (1.13) is deduced using a Galerkin procedure (Proposition 23.30 in [14]).

To obtain estimates (1.14) and (1.15) let y be the weak solution of (1.11). Then Definition 2 with $v = y(t)$ yields

$$\langle y_t(t), y(t)\rangle_{V',V} + B[t;y(t),y(t)] = (\mathrm{d}y(t),y(t))_H + (f,y(t))_H + (g,y(t))_{L^2(\partial\Omega)}, \text{a.e.}\, t \in [0,T]. \tag{1.17}$$

Furthermore,

$$\begin{aligned}
&\langle y_t(t), y(t)\rangle_{V',V} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|y(t)\|_H^2\right), \text{ a.e. } t \in [0,T],\\
&\left|(h(t),y(t))_H\right| \le \frac{1}{2}\left(\|h(t)\|_H^2 + \|y(t)\|_H^2\right), \text{ a.e. } t \in [0,T],\\
&\left|(g(t),y(t))_{L^2(\partial\Omega)}\right| \le \frac{1}{2}\left(\varepsilon^{-1}\|g(t)\|_{L^2(\partial\Omega)}^2 + \varepsilon\|y(t)\|_{L^2(\partial\Omega)}^2\right), \text{ a.e. } t \in [0,T], \varepsilon > 0,
\end{aligned} \tag{1.18}$$

where the equation is proved in Ch. III, Proposition 2.1 [11] and the inequalities are obtained by Cauchy's inequality. The continuous embedding $V \hookrightarrow L^2(\partial\Omega)$ yields $\|v\|_{L^2(\partial\Omega)} \le C_V\|v\|_V$, $v \in V$, for some $C_V > 0$, Ch. II, Theorem 3.3 in [11].

Consequently, (0.16)-(0.18) with $v = y(t)$ yield

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\|y(t)\|_H^2\right) + 2(\beta - \varepsilon C_V)\|y(t)\|_V^2\\
&\le 2\left(\langle y_t(t), y(t)\rangle_{V',V} + B[t;y(t),y(t)] - (\mathrm{d}y(t),y(t))_H + \gamma\|y(t)\|_H^2\right)\\
&\le \|h(t)\|_H^2 + (1+2\gamma)\|y(t)\|_H^2 + \varepsilon^{-1}\|g(t)\|_{L^2(\partial\Omega)}^2,
\end{aligned} \tag{1.19}$$

for a.e. $t \in [0,T]$. Now let $\eta(t) = \|y(t)\|_H^2$ and $\xi(t) = \|h(t)\|_H^2 + \varepsilon^{-1}\|g(t)\|_{L^2(\partial\Omega)}^2$. Then (1.19) implies

$$\eta'(t) \le (1+2\gamma)\eta(t) + \xi(t), \text{ a.e. } t \in [0,T],$$

and the differential form of Gronwall's inequality (Appendix B2 [12]) yields

$$\eta(t) \le \mathrm{e}^{(1+2\gamma)t}\left(\eta(0) + \int_0^t \xi(s)\,\mathrm{d}s\right), t \in [0,T].$$

Since $\eta(0) = \|y(0)\|_H^2 = \|y_0\|_H^2$ it is obtained

$$\|y(t)\|_H^2 \le \mathrm{e}^{(1+2\gamma)t}\left(\|y_0\|_H^2 + \|h\|_{L^2(Q)}^2 + \varepsilon^{-1}\|g(t)\|_{L^2(\Sigma)}^2\right), t \in [0,T]. \tag{1.20}$$

Integrating (1.20) from 0 to T, results

$$\|y\|_{L^2(Q)}^2 \le \frac{\mathrm{e}^{(1+2\gamma)T} - 1}{\varepsilon(1+2\gamma)}\left(\|y_0\|_H^2 + \|h\|_{L^2(Q)}^2 + \|g(t)\|_{L^2(\Sigma)}^2\right).$$

By employing the inequality $\sqrt{a^2+b^2} \le |a| + |b|$, the estimate (1.14) with $C(T) = \sqrt{\frac{\mathrm{e}^{(1+2\gamma)T} - 1}{\varepsilon(1+2\gamma)}}$ is concluded and $\lim_{T\to 0^+} C(T) = 0^+$.

The estimate (1.15) is a consequence of (1.20) with $C_0(T)=\varepsilon^{-1/2}\mathrm{e}^{(1+2\gamma)T/2}$. The last assertion of Theorem is proved in Proposition 3.3 of [10].

We also meet the backward form of problem (1.11), *i.e*. the linear final-boundary value problem,

$$\begin{aligned}&-p_t+Ap=dp+h,\ \text{in } Q,\\&\partial_{\nu A}p+cp=g,\ \text{on } \Sigma,\\&p(T)=p_f,\ \text{in } \Omega,\end{aligned}\tag{1.21}$$

with $d\in L^\infty(Q)$, $h\in L^2(Q)$, $c\in L^\infty(\Sigma)$, $g\in L^2(\Sigma)$ and $p_f\in H$. All results of Theorem 2 are valid for (0.21).

Theorem 3 *The initial-boundary value problem* (1.21) *has a unique weak solution* $p\in W$ *with norm estimates*

$$\|p\|_W\le C_W(T)\left(\|p_f\|_H+\|h\|_{L^2(Q)}+\|g\|_{L^2(\Sigma)}\right),\tag{1.22}$$

$$\|p\|_{L^2(Q)}\le C(T)\left(\|p_f\|_H+\|h\|_{L^2(Q)}+\|g\|_{L^2(\Sigma)}\right),\tag{1.23}$$

$$\|p\|_{C([0,T];H)}\le C_0(T)\left(\|p_f\|_H+\|h\|_{L^2(Q)}+\|g\|_{L^2(\Sigma)}\right),\tag{1.24}$$

where $C_W(T)$ *and* $C_0(T)$ *are bounded when* T *varies boundedly and* $\lim_{T\to0^+}C(T)=0^+$. *If* $h\in L^\infty(Q)$, $g\in L^\infty(\Sigma)$ *and* $p_f\in L^\infty(\Omega)$ *then* $p\in L^\infty(Q)$.

Proof 2 *The substitution* $w(t)=p(T-t)$ *in* (0.21) *yields the following equivalent problem to the problem* (0.11) *in the forward form*

$$\begin{aligned}&w_t+Aw=d(T-t)w+h(T-t),\ \text{in } Q,\\&\partial_{\nu A}w+c(T-t)w=g(T-t),\ \text{on } \Sigma,\\&w(0)=p_f,\ \text{in } \Omega.\end{aligned}\tag{1.25}$$

Problem (1.25) satisfies all the assumptions which problem (1.11) satisfies. Therefore the assertions of Theorem 2 and the estimates (1.13)-(1.15) are valid for w. Since $\|w(t)\|_X=\|p(T-t)\|_X=\|p(t)\|_X$, when X is one of spaces $L^2(Q)$, W, $C([0,T];H)$ or $L^\infty(Q)$, the assertions of theorem and the estimates (1.22)-(1.24) are verified.

3. The Optimality System

Let $U_{ad}=\{u\in L^2(\Sigma)\,|\,a\le u\le b\}$ with $a,b\in L^2(\Sigma)\cap L^\infty(\Sigma)$. Consider the following control problem

$$\begin{aligned}&\min_{u\in U_{ad}}J(y,u):=\frac{1}{2}\|y-y_d\|^2_{L^2(Q)}+\frac{\alpha}{2}\|u\|^2_{L^2(\Sigma)}\\&y\text{satisfies (1.6),}\end{aligned}\tag{1.26}$$

where $\alpha>0$ and $y_d\in L^\infty(Q)$.

Theorem 4 *Under Assumptions* (A1)-(A3) *the optimal control problem* (0.26) *has an optimal solution* (y,u) *in* $W\times U_{ad}$.

Proof 3 *By Theorem* 1 *the optimal control problem* (0.26) *is feasible. Also the Nemytskii operator* $Fy=f(t,x,y(t,x))$, $(t,x)\in Q$, as an operator from $L^2(0,T;V)$ into $L^2(0,T;V')$ is completely continuous. Because, when $y_n\rightharpoonup y$, weakly in $L^2(0,T;V)$, the compact embedding $L^2(0,T;V)\hookrightarrow\hookrightarrow L^2(0,T;H)\equiv L^2(Q)$ yields $y_n\to y$, strongly in $L^2(Q)$. Consequently, the continuity of $F:L^2(Q)\to L^2(Q)$ and the continuous embedding $L^2(Q)\equiv L^2(0,T;H)\hookrightarrow L^2(0,T;V')$ results in $Fy_n\to Fy$, strongly in $L^2(0,T;V')$, (confer Assumption (A2)).

Thus, the existence of an optimal solution $(y,u)\in W\times U_{ad}$ for the problem (1.26) can be deduced from Theorem 1.45 of [1].

Theorem 5 *A necessary condition for* $(y,u)\in W\times U_{ad}$ *be a local solution* (*or an optimal solution*) *of problem* (0.26) *is that there exists* $p\in W$ *such that*

$$y_t(t,x) + Ay(t,x) = f(t,x,y(t,x)), \text{ in } Q, \tag{1.27}$$

$$\begin{aligned} &\partial_{\nu A} y(t,x) + cy(t,x) = u(t,x), \text{ in } \Sigma, \\ &y(0,x) = y_0(x), \text{ in } \Omega, \\ &-p_t(t,x) + Ap(t,x) = f_y(t,x,y(t,x))\,p(t,x) - (y(t,x) - y_d(t,x)), \text{ in } Q, \end{aligned} \tag{1.28}$$

$$\begin{aligned} &\partial_{\nu A} p(t,x) + cp(t,x) = 0, \text{ in } \Sigma, \\ &p(T,x) = 0, \text{ in } \Omega, \\ &(\alpha u(t,x) - p(t,x), u(t,x) - v(t,x))_{L^2(\Sigma)} \ge 0, \; u, v \in U_{ad}. \end{aligned} \tag{1.29}$$

Proof 4 Corollary 1.3 in [1] (or Theorem 1.48 in [1]).

Theorem 6 *Any solution* (y, p, u) *of the optimality system* (1.27)-(1.29) *belongs to*

$$L^\infty(Q) \times L^\infty(Q) \times L^\infty(\Sigma).$$

Proof 5 *As* U_{ad} *is bounded, utilizing Theorem* 2.1 (*or Theorem* 3.1 *in* [10]) *it is deduced* $y \in L^\infty(Q)$. *Then Theorem* 2.1 *in* [10] *yields* $p \in L^\infty(Q)$.

Lemma 1 *The optimality condition* (0.29) *can be written in the equivalent form bellow*:

$$\alpha u(t,x) - p(t,x) \begin{cases} \ge 0, & a(t,x) = u(t,x), \\ = 0, & a(t,x) < u(t,x) < b(t,x), \\ \le 0, & u(t,x) = b(t,x), \end{cases} \tag{1.30}$$

for a.e. $(t,x) \in \Sigma$.

Proof 6 *Refer to Lemma* 1.12 in [1].

Corollary 1 *Let* $(u_i, p_i) \in U_{ad} \times W$, $i = 1,2$, *satisfy* (0.30). *Then*

$$|\alpha u_1(t,x) - \alpha u_2(t,x)| \le |p_1(t,x) - p_2(t,x)|, \text{a.e.}(t,x) \in \Sigma. \tag{1.31}$$

(u_i 's and p_i 's do not necessarily satisfy an optimality system).

Proof 7 *Let* $(t,x) \in \Sigma$ *and* (u_i, p_i), $i = 1,2$, *satisfy* (0.30) *at* (t,x). *Then one of the three cases below occurs for* (u_2, p_2) *at* (t,x),

$$\begin{aligned} &\alpha u_2(t,x) = \alpha a(t,x) \ge p_2(t,x), \\ &\alpha u_2(t,x) = p_2(t,x), \\ &p_2(t,x) \ge \alpha b(t,x) = \alpha u_2(t,x). \end{aligned}$$

Similarly one of the three such cases occurs for (u_1, p_1) at (t,x). Let the first case be occurred for (u_1, p_1) at (t,x). Then one of the three cases below must be considered for $\alpha u_1 - \alpha u_2$ at (t,x),

$$\begin{aligned} &0 = \alpha a(t,x) - \alpha a(t,x) = \alpha u_1(t,x) - \alpha u_2(t,x) \le p_1(t,x) - p_2(t,x), \\ &0 \ge \alpha a(t,x) - \alpha u_2(t,x) = \alpha u_1(t,x) - \alpha u_2(t,x) \ge p_1(t,x) - p_2(t,x), \\ &0 \ge \alpha a(t,x) - \alpha b(t,x) = \alpha u_1(t,x) - \alpha u_2(t,x) \ge p_1(t,x) - p_2(t,x). \end{aligned}$$

As you see each of the three cases above satisfies (0.31) at (t,x). In a similar argument for each of the two other cases of (u_1, p_1) at (t,x), three relations as the above can be written proving that each of them satisfy (0.31) at (t,x).

4. The Quasilinearization Method

Consider problem (1.26) under Assumptions (A1)-(A3). We investigate instead of the optimality system (1.27)-(1.29) the following one wherein the optimality condition (1.29) has been replaced by its equivalent form (1.30), confer Lemma 1:

$$y_t(t,x)+Ay(t,x)=f\left(t,x,y(t,x)\right),\ \text{in } Q, \tag{1.32}$$

$$\begin{aligned}&\partial_{\nu A}y(t,x)+cy(t,x)=u(t,x),\ \text{on } \Sigma,\\ &y(0,x)=y_0(x),\ \text{in } \Omega,\end{aligned} \tag{1.33}$$

$$-p_t(t,x)+Ap(t,x)=f_y\left(t,x,y(t,x)\right)p(t,x)-\left(y(t,x)-y_d(t,x)\right),\ \text{in } Q,$$

$$\begin{aligned}&\partial_{\nu A}p(t,x)+cp(t,x)=0,\ \text{in } \Sigma,\\ &p(T,x)=0,\ \text{in } \Omega,\end{aligned} \tag{1.34}$$

$$\alpha u(t,x)-p(t,x)\begin{cases}\geq 0, a(t,x)=u(t,x)\\ =0, a(t,x)<u(t,x)<b(t,x),\ \text{on } \Sigma.\\ \leq 0, u(t,x)=b(t,x)\end{cases}$$

By Theorem 5 and Theorem 4 optimality system (1.32)-(1.34) has at least one solution.

Theorem 7 *Let* $(y,p,u)\in W\times W\times U_{ad}$ *be a solution of optimality system* (0.32)-(0.34). *Then there exists a sequence* $\left\{\left(y^n,p^n,u^n\right)\right\}_{n=1}^{\infty}$ *in* $W\times W\times U_{ad}$ *whose elements are the unique solution of the following linear optimality systems and there exists* $T_1>0$ *such that this sequence converges, at least linearly, to* (y,p,u) *when* $T\leq T_1$. *As a consequence when* $T\leq T_1$ *optimality system* (1.32)-(1.34) *has a unique solution.*

$$\begin{aligned}&y_t^n(t,x)+Ay^n(t,x)\\ &=f\left(t,x,y^{n-1}(t,x)\right)+f_y\left(t,x,y^{n-1}(t,x)\right)\left(y^n(t,x)-y^{n-1}(t,x)\right),\ \text{in } Q,\\ &\partial_{\nu A}y^n(t,x)+cy^n(t,x)=u^n(t,x),\ \text{on } \Sigma,\\ &y^n(0,x)=y_0(x),\ \text{in } \Omega,\end{aligned} \tag{1.35}$$

$$\begin{aligned}&-p_t^n(t,x)+Ap^n(t,x)\\ &=f_y\left(t,x,y^{n-1}(t,x)\right)p^n(t,x)-\left(y^n(t,x)-y_d(t,x)\right),\ \text{in } Q,\\ &\partial_{\nu A}p^n(t,x)+cp^n(t,x)=0,\ \text{on } \Sigma,\\ &p^n(T,x)=0,\ \text{in } \Omega,\end{aligned} \tag{1.36}$$

$$\alpha u^n(t,x)-p^n(t,x)\begin{cases}\geq 0, a(t,x)=u^n(t,x)\\ =0, a(t,x)<u^n(t,x)<b(t,x),\ \text{on } \Sigma.\\ \leq 0, u^n(t,x)=b(t,x)\end{cases} \tag{1.37}$$

Proof 8 *About the existence of sequence* $\left\{\left(y^n,p^n,u^n\right)\right\}_{n=1}^{\infty}$ *in* $W\times W\times U_{ad}$, *note that* (0.35)-(0.37) *is the optimality system of following linear-quadratic optimal control problem*

$$\min_{u^n\in U_{ad}} J\left(y^n,u^n\right):=\frac{1}{2}\left\|y^n-y_d\right\|_{L^2(Q)}^2+\frac{\alpha}{2}\left\|u^n\right\|_{L^2(\Sigma)}^2,\ y^n \text{ satisfies } (1.35), \tag{1.38}$$

which has a unique optimal solution (Theorem 1.43 [1]). Then the optimality theory for linear-quadratic optimal control problems yields the existence of a unique solution $\left(y^n,p^n,u^n\right)$ in $W\times W\times U_{ad}$ of the system (1.35)-(1.37) when $y^{n-1}\in L^2(Q)\cap L^\infty(Q)$, confer Sections 1.5-1.7 in [1]. Referring to Theorems 2 and 3 it is deduced y^n and $p^n\in L^\infty(Q)$.

Now let $(y,p,u)\in W\times W\times U_{ad}$ be a solutions of the optimality systems (1.32)-(1.34). Define

$$Y^n:=y^n-y,\quad U^n:=u^n-u,\quad P^n:=p^n-p. \tag{1.39}$$

Then (1.32) and (1.35) and the mean value theorem yield

$$\begin{aligned}
&Y_t^n(t)+AY^n = f\left(t,x,y^{n-1}\right)-f\left(t,x,y\right)+f_y\left(t,x,y^{n-1}\right)\left(Y^n-Y^{n-1}\right)\\
&= f_y\left(t,x,\eta\right)Y^{n-1}+f_y\left(t,x,y^{n-1}\right)\left(Y^n-Y^{n-1}\right), \text{ in } Q,\\
&\partial_{\nu A}Y^n+cY^n=U^n, \text{ on } \Sigma,\\
&Y^n(0)=0, \text{ in } \Omega,
\end{aligned} \tag{1.40}$$

where $\eta(t,x)$ lies between $y^{n-1}(t,x)$ and $y(t,x)$, $(t,x)\in Q$. By Assumption (A3), $f_y\left(t,x,y^{n-1}\right)\in L^\infty(Q)$. Thus considering (1.40) as the linear problem (1.11) with $d=f_y\left(t,x,y^{n-1}\right)$, it is concluded by Theorem 2,

$$\left\|Y^n\right\|_{L^2(Q)}\le C(T)\left(\left\|\left(f_y\left(t,x,\eta\right)-f_y\left(t,x,y^{n-1}\right)\right)Y^{n-1}\right\|_{L^2(Q)}+\left\|U^n\right\|_{L^2(\Sigma)}\right). \tag{1.41}$$

Also (1.33) and (1.36) yield

$$\begin{aligned}
&-P_t^n(t)+AP^n = f\left(t,x,y^{n-1}\right)p^n-f_y\left(t,x,y\right)p-Y^n\\
&= f_y\left(t,x,y^{n-1}\right)P^n+\left(f_y\left(t,x,y^{n-1}\right)-f_y\left(t,x,y\right)\right)p-Y^n, \text{ in } Q.\\
&\partial_{\nu A}P^n+cP^n=0, \text{ on } \Sigma,\\
&P^n(T)=0, \text{ in } \Omega.
\end{aligned} \tag{1.42}$$

Since $f_y\left(t,x,y^{n-1}\right)\in L^\infty(Q)$, considering (1.42) as the linear problem (1.21) with $d=f_y\left(t,x,y^{n-1}\right)$, it is concluded by Theorem 3,

$$\left\|P^n\right\|_W\le C_W(T)\left(\left\|\left(f_y\left(t,x,y^{n-1}\right)-f_y\left(t,x,y\right)\right)p\right\|_{L^2(Q)}+\left\|Y^n\right\|_{L^2(Q)}\right). \tag{1.43}$$

Referring to Theorem 4, p belongs to $L^\infty(Q)$. Consequently employing Assumption (A3) and the mean value theorem, it is obtained

$$\left\|\left(f_y\left(t,x,y^{n-1}\right)-f_y\left(t,x,y\right)\right)p\right\|_{L^2(Q)}\le C_1\left\|Y^{n-1}\right\|_{L^2(Q)}, \tag{1.44}$$

where $C_1:=\left\|f_{yy}\left(t,x,\gamma\right)p\right\|_{L^\infty(Q)}\le M_{yy}\left\|p\right\|_{L^\infty(Q)}$ in which $\gamma(t,x)$ lies between $y^{n-1}(t,x)$ and $y(t,x)$, $(t,x)\in Q$. Therefore (1.43) yields

$$\left\|P^n\right\|_W\le C_W(T)\left(C_1\left\|Y^{n-1}\right\|_{L^2(Q)}+\left\|Y^n\right\|_{L^2(Q)}\right). \tag{1.45}$$

Owing to Corollary 1 and the continuous embeddings $W\hookrightarrow L^2(0,T;V)\hookrightarrow L^2(\Sigma)$,

$$\alpha\left\|U^n\right\|_{L^2(\Sigma)}\le\left\|P^n\right\|_{L^2(\Sigma)}\le C_\Sigma\left\|P^n\right\|_W. \tag{1.46}$$

Now combining (1.41), (1.45) and (1.46) results in

$$\left\|Y^n\right\|_{L^2(Q)}\le C(T)\left(2M_y\left\|Y^{n-1}\right\|_{L^2(Q)}+\alpha^{-1}C_\Sigma C_W(T)\left(C_1\left\|Y^{n-1}\right\|_{L^2(Q)}+\left\|Y^n\right\|_{L^2(Q)}\right)\right)$$

Consequently, it is obtained

$$\left\|Y^n\right\|_{L^2(Q)}\le C_2(T)\left\|Y^{n-1}\right\|_{L^2(Q)}, \tag{1.47}$$

wherein

$$C_2(T)=\frac{\left(2M_y+\alpha^{-1}C_\Sigma C_W(T)C_1\right)C(T)}{1-\alpha^{-1}C_\Sigma C_W(T)C(T)}. \tag{1.48}$$

Referring to Theorem 2, $C(T)\to 0^+$ when $T\to 0^+$ and $C_W(T)$ is bounded. Consequently, there exists

$T_1 > 0$ such that for $0 < T \le T_1$ the denumerator in (1.48) be positive and $0 < C_2(T) \lneqq 1$. This yields the convergence of Y^n to zero in $L^2(Q) \equiv L^2(0,T;H)$ for $T \le T_1$, thereby the convergence of P^n to zero in $W(0,T) = W$ for $T \le T_1$ via (1.45) and the convergence of U^n to zero in $L^2(\Sigma) \equiv L^2(0,T;L^2(\partial\Omega))$ for $T \le T_1$ via (1.46). The estimate (1.13) in Theorem 2 for the initial boundary value problem (1.40) yields

$$\begin{aligned}\|Y^n\|_W &\le C_W(T)\Big(\big\| f_y(t,x,\eta)Y^{n-1} + f_y(t,x,y^{n-1})(Y^n - Y^{n-1})\big\|_{L^2(Q)} + \|U^n\|_{L^2(\Sigma)}\Big) \\ &\le C_W(T)\Big(M_{yy}\|Y^{n-1}\|_{L^2(Q)} + M_y\|Y^n\|_{L^2(Q)} + \|U^n\|_{L^2(\Sigma)}\Big),\end{aligned} \tag{1.49}$$

where the second inequality is obtained using the mean value theorem. Consequently, the convergence of Y^n to zero in W for $T \le T_1$ is obtained. Referring to (0.47), the convergence of Y^n in $L^2(Q)$ is at least linear whereby the convergence of P^n in W and U^n in $L^2(\Sigma)$ will be at least linear for $T \le T_1$, confer (1.45) and (1.46). Then it is concluded from the estimates (1.49) that the convergence rate of Y^n to zero in W is at least linear for $T \le T_1$.

The sequence $\{(y^n, p^n, u^n)\}_{n=1}^{\infty}$ produced by (1.35)-(1.37) is independent from (y,p,u) and converges to it in $W \times W \times U_{ad}$. As (y,p,u) can be any solution of optimality system (1.32)-(1.34) this is impossible except optimality system (1.32)-(1.34) has only one solution.

The next two corollaries are used in the error analysis in Section 6.

Corollary 2 *Under assumptions of Theorem* 7 *there exists* T_2, $0 < T_2 \le T_1$, *such that for* $T \le T_2$ *the following estimate is valid*

$$\|Y^n\|_{C(0,T;H)} \le C_3(T)C_2(T)^{n-1}\|Y^0\|_{L^2(Q)}, \tag{1.50}$$

where Y^n, $C_2(T)$ and $C_3(T)$ are as in (1.39), (1.48) and (1.53), respectively.

Proof 9 *The proof follows the lines of proof of Theorem* 7. *As* Y^n *satisfies* (0.40) *the estimate* (0.15) *in Theorem* 2 *yields*

$$\|Y^n\|_{C([0,T];H)} \le C_0(T)\Big(2M_y\|Y\|^{n-1}{}_{L^2(Q)} + \|U^n\|_{L^2(\Sigma)}\Big). \tag{1.51}$$

Next employing the estimates (1.45) and (1.46) result in

$$\|Y^n\|_{C([0,T];H)} \le C_0(T)\Big(2M_y\|Y^{n-1}\|_{L^2(Q)} + \alpha^{-1}C_\Sigma C_W(T)\big(C_1\|Y^{n-1}\|_{L^2(Q)} + \|Y^n\|_{L^2(Q)}\big)\Big).$$

As $\|Y^n\|_{L^2(Q)} \le T\|Y^n\|_{C([0,T];H)}$ it is deduced from the above inequality

$$\|Y^n\|_{C([0,T];H)} \le C_3(T)\|Y^{n-1}\|_{L^2(Q)}, \tag{1.52}$$

wherein

$$C_3(T) = \frac{\big(2M_y + \alpha^{-1}C_\Sigma C_W(T)C_1\big)C_0(T)}{1 - \alpha^{-1}C_\Sigma C_W(T)C_0(T)T}. \tag{1.53}$$

Referring to Theorem 2, $C_W(T)$ and $C_0(T)$ are bounded when $T \to 0^+$ whereby there exists $T_0 > 0$ such that the denumerator in (1.53) is positive for $0 < T \le T_0$. Set $T_2 = \min\{T_0, T_1\}$ with T_1 being determined in Theorem 7. Then (1.50) is obtained from (1.52) by repeatedly employing the estimate (1.47).

Corollary 3 *Suppose in the quasilinearization method in Theorem* 7 *instead of the accurate initial value* y_0 *the approximate initial value* $y_0 + \mathcal{E}$ *is used. Let* T_2 *be as in Corollary* 2. *Then for* $T \le T_2$ *the following estimate is valid*

$$\|Y^n\|_{C([0,T];H)} \le C_3(T)C_2(T)^{n-1}\|Y^0\|_{L^2(Q)} + C_5(T)\|\mathcal{E}\|_H, \tag{1.54}$$

where $C_2(T)$, $C_3(T)$ *and* $C_5(T)$ *are as in* (1.48), (1.53) *and* (1.59), *respectively*.

Proof 10 *The proof follows the lines of proof of Theorem* 7. *As* Y^n *satisfies* (1.40) *in* Q *with*

$Y^n(0) = y^n(0) - y(0) = y_0 + \mathcal{E} - y_0 = \mathcal{E}$ *the estimate* (1.15) *in Theorem* 2 *yields*

$$\|Y^n\|_{C([0,T];H)} \le C_0(T)\left(\|\mathcal{E}\|_H + 2M_y\|Y^{n-1}\|_{L^2(Q)} + \|U^n\|_{L^2(\Sigma)}\right). \tag{1.55}$$

Next employing the estimate (1.45) and (1.46) result in

$$\|Y^n\|_{C([0,T];H)} \le C_0(T)\left(\|\mathcal{E}\|_H + 2M_y\|Y^{n-1}\|_{L^2(Q)} + \alpha^{-1}C_\Sigma C_W(T)\left(C_1\|Y^{n-1}\|_{L^2(Q)} + \|Y^n\|_{L^2(Q)}\right)\right).$$

As $\|Y^n\|_{L^2(Q)} \le T\|Y^n\|_{C([0,T];H)}$, choosing $T_2 > 0$ as in Corollary 2, (0.55) for $T \le T_2$ yields

$$\|Y^n\|_{C([0,T];H)} \le C_3(T)\|Y^{n-1}\|_{L^2(Q)} + C_4(T)\|\mathcal{E}\|_H, \tag{1.56}$$

where $C_4(T) = \dfrac{C_0(T)}{1-\alpha^{-1}C_\Sigma C_W C_0(T)T}$.

Now in order to conclude (0.54) we need an estimate like (1.47). (1.47) is for the case $Y^n(0) = 0$ here $Y^n(0) = \mathcal{E}$. Such an estimate is obtained following the lines which (1.47) obtained. As Y^n satisfies (1.40) with $Y^n(0) = \mathcal{E}$, the estimate (1.14) in Theorem 2 yields

$$\|Y^n\|_{L^2(Q)} \le C(T)\left(\|\mathcal{E}\|_H + 2M_y\|Y^{n-1}\|_{L^2(Q)} + \|U^n\|_{L^2(\Sigma)}\right).$$

Then employing the estimates (1.45) and (1.46) result in

$$\|Y^n\|_{L^2(Q)} \le C_2(T)\|Y^{n-1}\|_{L^2(Q)} + \frac{C(T)}{1-\alpha^{-1}C_\Sigma C_W C(T)}\|\mathcal{E}\|_H,$$

where $T \le T_1$, T_1 being determined after (1.48). Referring to (1.48), without loss of generality, it is considered $C_2(T) > \dfrac{C(T)}{1-\alpha^{-1}C_\Sigma C_W(T)C(T)}$ whereby it is obtained

$$\|Y^n\|_{L^2(Q)} \le C_2(T)\left(\|Y^{n-1}\|_{L^2(Q)} + \|\mathcal{E}\|_H\right). \tag{1.57}$$

Employing repeatedly (0.57) yields

$$\begin{aligned}\|Y^n\|_{L^2(Q)} &\le C_2(T)^n\|Y^0\|_{L^2(Q)} + \left(C_2(T)^n + C_2(T)^{n-1} + \cdots + C_2(T)\right)\|\mathcal{E}\|_H,\\ &= C_2(T)^n\|Y^0\|_{L^2(Q)} + \frac{1-C_2(T)^n}{1-C_2(T)}C_2(T)\|\mathcal{E}\|_H,\\ &\le C_2(T)^n\|Y^0\|_{L^2(Q)} + \frac{C_2(T)}{1-C_2(T)}\|\mathcal{E}\|_H,\end{aligned} \tag{1.58}$$

where the last inequality is obtained from $0 < C_2(T) < 1$ for $T \le T_1$. Now utilizing (1.58) in (1.56) results in (1.54) with

$$C_5(T) = C_3(T)\left(\frac{C_2(T)}{1-C_2(T)} + C_4(T)\right). \tag{1.59}$$

5. Application to the Optimal Boundary Control Problems and the Uniqueness

The proposed quasilinearization method in Theorem 7 is convergent on the time intervals $[0,T]$ for $T \le T_1$, T_1 being determined in Theorem 7. In order to apply the quasilinearization method to the optimal control problem (1.26) up to an arbitrary final time T it is possible to decompose the problem into many finite optimal control problems each on an interval with length less than T_1. In order to follow such an approach let $T_2 \le T_1$ [1]

[1]In order to preserve the stability, T_2 is chosen as in Corollary 2 (also confer Section 6).

and $T = mT_2$ for some $m \in \mathbb{N}$. Let $t_i := i \times T_2$, $Q_i := (t_{i-1}, t_i) \times \Omega$ and $\Sigma_i := [t_{i-1}, t_i] \times \partial\Omega$, $i := 0 \cdots m$. Let X be a Banach space. Then $L^2(0,T;X)$ is normisomorphic to $\prod_{i=1}^m L^2(t_{i-1}, t_i; X)$ through the isomorphism

$$\begin{aligned} &L^2(0,T;X) \ni y \to (y_1 \cdots y_i \cdots y_m) \in \prod_{i=1}^m L^2(t_{i-1}, t_i; X), \\ &y_i(t) := y(t), t \in (t_{i-1}, t_i), i := 1 \cdots m, \\ &\|y\|_{L^2(0,T;X)} = \sum_{i=1}^m \|y_i\|_{L^2(t_{i-1}, t_i; X)}. \end{aligned}$$

Replacing X by H yields that $L^2(Q) \equiv L^2(0,T;H)$ be normisomorphic to $\prod_{i=1}^m L^2(Q_i) \equiv \prod_{i=1}^m L^2(t_{i-1}, t_i; H)$ with the norm identity $\|y\|_{L^2(Q)} = \sum_{i=1}^m \|y_i\|_{L^2(Q_i)}$, and replacing X by V yields that W be normisomorphic to the closed subspace W_c of $\prod_{i=1}^m W_i$ with the norm identity $\|y\|_W = \sum_{i=1}^m \|y_i\|_{W_i}$, where

$$\begin{aligned} &W = \left\{ y \in L^2(0,T;V) \middle| y_t \in L^2(0,T;V') \right\}, \\ &W_i = \left\{ y_i \in L^2(t_{i-1}, t_i; V) \middle| y_{it} \in L^2(t_{i-1}, t_i; V') \right\}, \\ &W_c = \left\{ (y_1 \cdots y_i \cdots y_m) \in \prod_{i=1}^m W_i \middle| y_{i-1}(t_{i-1}) = y_i(t_{i-1}), i = 2 \cdots m \right\}. \end{aligned}$$

Thus, if $y \in W$ satisfies the initial-boundary value problem (0.6) then $y_i \in W_i$, $y_i \in W_i, i := 1 \cdots m$, satisfy consecutively the following initial-boundary value problems and vice versa:

$$\begin{aligned} &y_{it}(t,x) + Ay_i(t,x) = f(t,x,y_i(t,x)), (t,x) \in Q_i, \\ &\partial_{\nu A} y_i(t,x) + cy_i(t,x) = u_i(t,x), (t,x) \in \Sigma_i, \\ &y_i(t_{i-1}, x) = y_{i-1}(x), x \in \Omega, \end{aligned} \tag{1.60}$$

wherein $y_{i-1} := y_{i-1}(t_{i-1}) = y(t_{i-1})$. Consequently, the optimal control problem (1.26) is equivalent to the consecutive optimal control subproblems

$$\min_{u_i \in U_{ad}^i} J_i(y_i, u_i) := \frac{1}{2} \left\| y_i - y_d^i \right\|_{L^2(Q_i)}^2 + \frac{\alpha}{2} \|u_i\|_{L^2(Q_i)}^2, y_i \text{ satisfies } (1.60), \tag{1.61}$$

wherein $U_{ad}^i = \left\{ u_{|\Sigma_i} \middle| u \in U_{ad} \right\}$. Therefore, solving the optimal control problem (1.26) is equivalent to consecutively solving the optimal control subproblems (1.61). Furthermore, the proposed quasilinearization method in Theorem 7 is applicable to each optimal control subproblem in (1.61). In fact the substitution $t \rightarrowtail t_{i-1} + t$ in the i-th subproblem in (1.61) transforms it into an equivalent problem on the time interval $[0, t_i - t_{i-1}] = [0, T_2]$ whereby the quasilinearization method will be applicable to it.

Moreover, as a consequence of Theorem 7 the solution of optimality system of i-th subproblem in (1.61) is unique. Thus, in view of Theorem's 4 and 5, each subproblem in (1.61) has a unique optimal solution. Consequently by the equivalence between problem (1.26) and consecutive subproblems (1.61) it can be stated

Theorem 8 *Optimal boundary control problem* (1.26) *under Assumptions* (A1)-(A3) *has unique optimal boundary control solution and optimal state solution.*

Note that the uniqueness could not be established thorough the optimality theory of optimal control problems which was used for stating the existence in Section 3. This is due to lack of convexity of the solution set of problem (1.26).

An issue concerning the above consecutive process is the relation between (y, p, u), the solution of optimality system of problem (1.26), and (y_i, p_i, u_i), the solution of optimality system of i-th subproblem in (1.61). (y, p, u) satisfies (1.27)-(1.29) on Q and (y_i, p_i, u_i) satisfies

$$y_{it}(t,x)+Ay_i(t,x)=f\left(t,x,y_i(t,x)\right),\ \text{in } Q_i, \tag{1.62}$$

$$\begin{aligned}&\partial_{\nu A}y_i(t,x)+cy_i(t,x)=u_i(t,x),\ \text{on } \Sigma_i\\&y_i(0,x)=y_0(x),\ \text{in } \Omega,\\&-p_{it}(t,x)+Ap_i(t,x)=f_y\left(t,x,y_i(t,x)\right)p_i(t,x)-\left(y_i(t,x)-y_d^i(t,x)\right),\ \text{in } Q_i,\end{aligned} \tag{1.63}$$

$$\begin{aligned}&\partial_{\nu A}p_i(t,x)+cp_i(t,x)=0,\ \text{on } \Sigma_i,\\&p_i(T,x)=0,\ \text{in } \Omega,\\&\left\langle \alpha u_i(t,x)-p(t,x),u_i(t,x)-v(t,x)\right\rangle_{L^2(\Sigma_i)}\geq 0,\ \ u_i,v\in U_{ad}^i.\end{aligned} \tag{1.64}$$

In view of Theorem's 8, 4 and 5 optimality system of problem (1.26) has a unique solution. Consequenty comparing (1.61)-(1.64) with (1.26)-(1.29) it is concluded that $J(y,u)=\sum_{i=1}^{m}J_i(y_i,u_i)$ and $y_i(t)=y(t)$, $t\in(t_{i-1},t_i)$, and $u_i(t)=u(t)$, $t\in(t_{i-1},t_i)$. But there is not a similar relation between the costates p and p_i's, since p_i satisfies (1.63) and $p_i(t_i)=0$, but $p(t_i)$ is not necessarily zero; confer (1.28). Also it is not possible in general to construct p from p_i's; however, after obtaining y_i's, p can be computed from (1.28).

6. Error Analysis

By the consecutive quasilinearization method in Section 5, the optimal control problem (1.26) is solved through m consecutive optimal control subproblems (1.61). Each subproblem is solved by the quasilinearization method in Theorem 7 which is an iterative method with infinite iterations. In applications it is implemented up to a finite iterations, thereby producing error. Consequently, during solving each subproblem there exists an error production and an error propagation.

Let (y,p,u) be the solution of optimality system (1.32)-(1.34), (y_i,p_i,u_i) be the solution of *i*-th optimality system (0.62)-(0.64) and $\left(y_i^n,p_i^n,u_i^n\right)$ be the solution provided by the quasilinearization method at iteration *n* for the *i*-th optimality system, *i.e.* one which satisfies (1.35)-(1.37) on Q_i. For the first subproblem the quasilinearization method starts with the accurate initial value $y_1^n(0)=y_1(0)=y(0)=y_0$ and it is terminated after *N* iteration with the final value $y_1^N(t_1)$. The error equals to $Y_1^N(t_1)=y_1^N(t_1)-y_1(0)$. As $t_1\leq T_2$ and the initial value is accurate, Corollary 2 with $T=t_1$ yields

$$\left\|Y_1^N(t_1)\right\|_H\leq\left\|Y_1^N\right\|_{C(0,T;H)}\leq C_3(T)C_2(T)^{N-1}\left\|Y_1^0\right\|_{L^2(Q_1)}. \tag{1.65}$$

For the *i*-th subproblem on $[t_{i-1},t_i]$, $i:=2\cdots m$, the quasilinearization method starts with the approximate initial value $y_i^n(t_{i-1})=y_i(t_{i-1})+Y_{i-1}^N(t_{i-1})$ and it is terminated after *N* iteration with final value $y_i^N(t_i)$. The error of final value equals to $Y_i^N(t_i)=y_i^N(t_i)-y_i(t_i)$. Next, we estimate this error.

The substitution $t\to t_{i-1}+t$ in the *i*-th subproblem in (1.61) transforms it into an equivalent problem on the time interval $[0,t_i-t_{i-1}]=[0,T_2]$. Setting $T=T_2$ and utilizing Corollary 3 for the equivalent problem, yields the estimate (1.54) with $y_0+\mathcal{E}=y_i(t_{i-1})+Y_{i-1}^N(t_{i-1})$. Then utilizing the reverse substitution $t_{i-1}+t\to t$ results in the estimate

$$\begin{aligned}\left\|Y_i^N(t_i)\right\|_H&\leq\left\|Y_i^N\right\|_{C([t_{i-1},t_i];H)}\\&\leq C_3(T)C_2(T)^{N-1}\left\|Y_i^0\right\|_{L^2(Q_i)}+C_5(T)\left\|Y_{i-1}^N(t_{i-1})\right\|_H.\end{aligned} \tag{1.66}$$

Now beginning from $i=m$ down to $i=1$, repeatedly employing (1.66) results in

$$
\begin{aligned}
&\left\|Y_m^N(t_m)\right\|_H \le \left\|Y_m^N\right\|_{C([t_{m-1},t_m];H)} \\
&\le C_3(T)C_2(T)^{N-1}\left\|Y_m^0\right\|_{L^2(Q_m)} + C_5(T)\left\|Y_{m-1}^N(t_{m-1})\right\|_H \\
&\le C_3(T)C_2(T)^{N-1}\left\|Y_m^0\right\|_{L^2(Q_m)} + C_5(T)C_3(T)C_2(T)^{N-1}\left\|Y_{m-1}^0\right\|_{L^2(Q_{m-1})} + C_5(T)^2\left\|Y_{m-2}^N(t_{m-2})\right\|_H \\
&\le C_3(T)C_2(T)^{N-1}\left(\left\|Y_m^0\right\|_{L^2(Q_m)} + C_5(T)\left\|Y_{m-1}^0\right\|_{L^2(Q_{m-1})} + C_5(T)^2\left\|Y_{m-2}^0\right\|_{L^2(Q_{m-2})} + \cdots + C_5(T)^{m-2}\left\|Y_2^0\right\|_{L^2(Q_2)}\right) \\
&\quad + C_5(T)^{m-1}\left\|Y_1^0(t_1)\right\|_H \\
&\le C_3(T)C_2(T)^{N-1}\left(1 + C_5(T) + \cdots + C_5(T)^{m-1}\right)\left\|Y^0\right\|_{L^2(Q)},
\end{aligned}
$$

where the last inequality is obtained by $\left\|Y^0\right\|_{L^2(Q)} = \sum_{i=1}^m \left\|Y_i^0\right\|_{L^2(Q_i)}$ and the estimate (1.65). Consequently,

$$
\left\|Y_m^N(t_m)\right\|_H \le \left\|Y_m^N\right\|_{C([t_{m-1},t_m];H)} \le C_3(T)\frac{C_5(T)^m - 1}{C_5(T) - 1}\left\|Y^0\right\|_{L^2(Q)} C_2(T)^{N-1}. \tag{1.67}
$$

Note that $Y_m^N(t_m)$ presents the accumulated error consists of the production errors and the propagation errors in the consecutive implementation of *m* quasilinearization method, when the implementation is up to *N* iteration on each subproblem. In the estimate (1.67) the term $C_3(T)\frac{C_5(T)^m - 1}{C_5(T) - 1}\left\|Y^0\right\|_{L^2(Q)}$ is independent from *N* and $0 < C_2(T) < 1$; confer (1.48) and thereafter. Since *m* is fixed, by increasing the number of iterations *N*, the total accumulated error $Y_m^N(t_m)$ tends to zero in *H*. Therefore, the proposed consecutive quasilinearization method in Section 5 is stable. Furthermore, $C_5(T) > 1$ for $T > 0$, although $C_5(T)$ and $C_3(T)$ decrease when *T* decrease (or *m* increase). Consequently it may a trade off be necessary between size of *m* (the number of subproblems) and *N* (the number of required iterations in the implementation of quasilinearization method) in order to have the desired total error in the consecutive quasilinearization method.

7. Numerical Example

A typical example is presented reflecting the obtained results in the previous sections in applications. Consider the optimal control problem (1.26) with the following data: $\Omega = (0,1)\times(0,1)$, $Q = (0,1)\times\Omega$, $\alpha = 0.1$, $a_{1,1} = a_{2,2} = 1$, $a_{1,2} = a_{2,1} = 0$, $c = 1$, $f(t, x_1, x_2, y) = 5\exp(-y^2)$, $y_0(x_1, x_2) = 4x_1(1-x_1)x_2(1-x_2)$, $y_d = 1$, $a = -1$, $b = 1$. Setting $T_2 = 1/m$, the consecutive quasilinearization method is implemented on the *m* consecutive subproblems (1.61) with the optimality systems (1.62)-(1.64). The corresponding states y_i^n, costates p_i^n and controls u_i^n are approximated by the elements and boundary elements of continuous linear finite element spaces on $\overline{Q}_i = [t_{i-1}, t_i]\times\overline{\Omega}$ with $h_{FEM} = 1/20$ and $h_t = t_i - t_{i-1} = T_2$, *i.e.* without discretization of time. The linear optimality systems (1.62)-(1.64) are solved by the semismooth Newton's method [16] or Section 2.5 in [1], and the implementation is done with MATLAB software. **Table 1** presents the values of

$$
\mathcal{E}(y) = \left\|y_4^n(t_4) - y_4^{n-1}(t_4)\right\|_{L^2(\Omega)},
$$

$$
\mathcal{E}(p) = \left\|p_4^n(t_4) - p_4^{n-1}(t_4)\right\|_{L^2(\Omega)}
$$

and

$$
\mathcal{E}(u) = \left\|u_4^n(t_4) - u_4^{n-1}(t_4)\right\|_{L^2(\partial\Omega)}.
$$

These values present at least a linear rate of convergence in the quasilinearization method as it was deduced from (1.45)-(1.47).

Table 2 presents the optimal objective values of problem when the consecutive quasilinearization method is implemented with different number of subproblems but fixed number of iterations in each quasilinearization

Table 1. The difference between iterations in the quasilineariztion method for the forth subproblem at $t = t_4$ when $m = 15$ and the number of iterations is $N = 10$.

n	1	2	3	4	5	6	7	8	9	10
$\mathcal{E}(y)$	1.2e+02	1.1e-01	6.4e-05	4.2e-08	7.9e-11	1.8e-13	3.7e-16	0.0e-00	0.0e-00	0.0e-00
$\mathcal{E}(p)$	8.1e-02	3.3e-03	1.8e-06	1.2e-9	1.8e-12	4.9e-15	3.1e-18	0.0e-00	0.0e-00	0.0e-00
$\mathcal{E}(u)$	3.9e-01	9.5e-03	5.9e-06	3.0e-09	6.6e-12	1.3e-14	4.1e-17	0.0e-00	0.0e-00	0.0e-00

Table 2. The optimal objective values with different number of subproblems, m, but fixed number of iterations in the quasilinearization method, *i.e.* $N = 5$.

m	5	10	15	20	25	30	50	100
$J = \sum_{i=1}^{m} J_i$	0.243547	0.278541	0.294925	0.304289	0.310307	0.314484	0.323194	0.329938

method, *i.e.* with different m's and fixed N. As $T_2 = 1/m$ is in some sense the step size of time discretization, its increment yields more accurate approximation to the optimal objective value.

8. Conclusions

A consecutive quasilinearization method was proposed for the optimal boundary control problems with quadratic objective of tracking type and a semilinear parabolic equation with mixed boundary as the state constraint; cf. (1.26) and (1.32). The proposed method divides the control problem equivalently into many finite consecutive subproblems through partitioning the time interval into subintervals; cf. Section 5 and (1.61). Then subproblems are solved consecutively by a quasilinearization method (hence the name of proposed method). Finally the optimal solution of control problem is obtained by consecutively merging optimal solutions of subproblems. The quasilinearization method for each subproblem constructs an infinite sequence of linear-quadratic optimal boundary control problems of form (1.38). The sequence of solutions to the optimality systems of these linear problems converges to *any* solutions of the optimality system of subproblem; confer Theorem 7 and Section 5. This implies the uniqueness of solution to the optimality system of a subproblem, hence the uniqueness of optimal solution to the original control problem; confer Theorem 8. This uniqueness result is new, on the best of our knowledge, in the class of optimal control problems with state constraint of semilinear parabolic equation type.

The convergence of quasilinearization method for each subproblem depends on the time interval length of the subproblem, T_2, and there is a bound on T_2 which the convergence occurs, $T_2 \le T_1$, T_1 being determined in Theorem 7. In comparison with methods which require the fully discretization of original control problem, cf. Chapter 2 in [1], [2] and [17], T_2 can be considered as the time discretization step length. In this view the consecutive feature of proposed method replaces the large scale computations in fully discrete methods by the consecutive small scale computations in the subproblems, hence increasing the machine applicability of method. Specially in quasilinearization method in solving the sequence of linear-quadratic control problems the time discretization can be avoided by choosing T_2 enough small , cf. Section 7.

In comparison with superlinear methods which are locally convergent, as different versions of Newton's method and/or Lagrange-SQP methods (Chapter 2 in [1], and [2]), the consecutive quasilinearization method is globally convergent and its convergence order is at least linear, cf. Theorem 7. For example **Table 1** of Section 7 presents a cubic convergence rate. Thereby the consecutive quasilinearization method is very suitable for the globalization of locally convergent methods by applying it to find a starting solution for those methods.

The quasilinearization method for subproblems has infinite iterations, but in applications it is implemented up to a finite iteration. Therefore its consecutive application on the subproblems produces and propagates errors. However choosing $T_2 \le T_1$ guarantees the numerical stability, cf. Section 6.

The imposed boundedness assumptions on the nonlinearity of problem and the admissible controls are necessary for the convergence proof, cf. Assumption (A3), Section 3 and proof of Theorem 7. As the investigated control problem here also has optimal solution with much weaker boundedness assumptions, cf. [10], application of consecutive quasilinearization method in this case requires new convergence proof.

References

[1] Hinze, M., Pinnau, R., Ulbrich, M. and Ulbrich, S. (2009) Optimization with PDE Constraints. Springer, Berlin.

[2] Trötzsch, F. (1999) On the Lagrange-Newton-SQP Method for the Optimal Control of Semilinear Parabolic Equations. *SIAM Journal on Control and Optimization*, **38**, 294-312. http://dx.doi.org/10.1137/S0363012998341423

[3] Bellman, R. and Kalaba, R. (1965) Quasilinearization and Nonlinear Boundary-Value Problems. American Elsevier Publishing Company, New York.

[4] Bellman, R. (1955) Functional Equations in the Theory of Dynamic Programming. V. Positivity and Quasi-Linearity. *Proceedings of the National Academy of Sciences of the United States of America*, **41**, 743-746. http://dx.doi.org/10.1073/pnas.41.10.743

[5] Buica, A. and Precup, R. (2002) Abstract Generalized Quasilinearization Method for Coincidences. *Nonlinear Studies*, **9**, 371-386.

[6] Lakshmikantham, V. and Vatsala, A.S. (1995) Generalized Quasilinearization for Nonlinear Problems. Kluwer Academic Publishers, Dordrecht.

[7] Carl, S. and Lakshmikantham, V. (2002) Generalized Quasilinearization and Semilinear Parabolic Problems. *Nonlinear Analysis*, **48**, 947-960. http://dx.doi.org/10.1016/S0362-546X(00)00225-X

[8] Beckenbach, E.F. and Bellman, R. (1961) Inequalities. Springer-Verlag, Berlin.

[9] Kalaba, R. (1959) On Nonlinear Differential Equations, the Maximum Operation, and Monotone Convergence. *Journal of Mathematics and Mechanics*, **8**, 519-574.

[10] Raymond, J.P. and Zidani, H. (1999) Hamiltonian Pontryagin's Principles for Control Problems Governed by Semilinear Parabolic Equations. *Applied Mathematics and Optimization*, **39**, 143-177. http://dx.doi.org/10.1007/s002459900102

[11] Showalter, R.E. (1997) Monotone Operators in Banach Spaces and Nonlinear Partial Differential Equations. American Mathematical Society, Providence.

[12] Evans, L.C. (1998) Partial Differential Equations. American Mathematical Society, Providence.

[13] Ladyzenskaja, O.A., Solonnikov, V.A. and Ural'ceva, N.N. (1968) Linear and Quasi-Linear Equations of Parabolic Type. American Mathematical Society, Providence.

[14] Zeidler, E. (1990) Nonlinear Functional Analysis and Its Applications Vol II/A: Linear Monotone Operators. Springer, Berlin.

[15] Showalter, R.E. (1994) Hilbert Space Methods for Partial Differential Equations. Electronic Journal of Differential Equations, Monograph 01, Dover Publications, Mineola.

[16] Ulbrich, M. (2011) Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces. SIAM. http://dx.doi.org/10.1137/1.9781611970692

[17] Neittaanmaki, P. and Tiba, D. (1994) Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms, and Applications. Marcel Dekker Inc., New York.