INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 4, Number 3-3, Pages 625–647

SHOOTING METHODS FOR NUMERICAL SOLUTIONS OF EXACT CONTROLLABILITY PROBLEMS CONSTRAINED BY LINEAR AND SEMILINEAR 2-D WAVE EQUATIONS

SUNG-DAE YANG

This paper is dedicated to Max Gunzburger on the occasion of his 60th birthday

Abstract. Numerical solutions of exact controllability problems for linear and semilinear 2-d wave equations with distributed controls are studied. Exact controllability problems can be solved by the corresponding optimal control problems. The optimal control problem is reformulated as a system of equations (an optimality system) that consists of an initial value problem for the underlying (linear or semilinear) wave equation and a terminal value problem for the adjoint wave equation. The discretized optimality system is solved by a shooting method. The convergence properties of the numerical shooting method in the context of exact controllability are illustrated through computational experiments.

Key Words. Controllability, finite difference method, distributed control, optimal control, parallel computation, shooting method, wave equation.

1. Introduction

In this paper, we consider an optimal distributed control approach for solving the exact distributed controllability problem for two-dimensional linear or semilinear wave equations defined on a time interval $(0,T)$, and spatial domain Ω in $\mathbb{R}^2$. The exact distributed controllability problem we consider is to seek a distributed control f in $L^2((0,T)\times\Omega)$ and a corresponding state u such that the following system of equations hold:

$$
\begin{cases}
u_{tt} - \Delta u + \Psi(u) = f \quad \text{in } Q \equiv (0,T)\times\Omega\,, \\
u|_{t=0} = w \quad \text{and} \quad u_t|_{t=0} = z \quad \text{in } \Omega\,, \\
u|_{t=T} = W \quad \text{and} \quad u_t|_{t=T} = Z \quad \text{in } \Omega\,, \\
u|_{\partial\Omega} = 0 \quad \text{in } (0,T)\,,
\end{cases}
\tag{1.1}
$$

where w and z are given initial conditions defined on Ω, $W \in L^2(\Omega)$ and $Z \in H^{-1}(\Omega)$ are prescribed terminal conditions, f in $L^2((0,T)\times\Omega)$ is the distributed control, and $\Psi(u)$ is a given function on $\mathbb{R}$.

The exact boundary controllability problems are well known for linear and semilinear cases; see e.g.,[5, 14, 15, 17, 18, 20, 21, 23, 24]. In these problems there are basically two classes of computational methods in the literature. The first class is known Hilbert Uniqueness Method (HUM); see, e.g., [9, 11, 14, 16, 22]. The approximate solutions obtained by the HUM-based methods in general do not seem

Received by the editors January 20, 2006.
2000 *Mathematics Subject Classification.* 93B40, 35L05, 65M06.

to converge (even in a weak sense) to the exact solutions as the temporal and spatial grid sizes tend to zero. Methods of regularization including Tychonoff regularization and filtering that result in convergent approximations were introduced in those papers on HUM-based methods. The second class of computational methods for boundary controllability of the linear wave equation was those based on the method proposed in [10]. One solves a discrete optimization problem that involves the minimization of the discrete boundary L^2 norm subject to the undetermined linear system of equations formed by the discretization of the wave equation and the initial and terminal conditions. This approach was implemented in [8]. The computational results demonstrated the convergence of the discrete solutions when the exact minimum boundary L^2 norm solution is smooth. In the generic case of a non-smooth exact minimum boundary L^2 norm solution the computational results of [8] exhibited at least a weak L^2 convergence of the discrete solutions.

In this paper we develop an alternate numerical method which allows us to apply distributed or boundary control to the exact controllability problems. Ultimately we test the exact boundary controllability problems, but it is beyond the work, and we will present the result in a separate paper. The results in [19] were limited to the one dimensional case. In this paper, we extend those results to the two dimensional case.

We will study numerical methods for optimal control and controllability problems associated with the linear and semilinear wave equations. We are particularly interested in investigating the relevancy and applicability of high performance computing (HPC) for these problems. As a prototype example of optimal control problems for the wave equations we consider the following distributed optimal control problem: choose a control f and a corresponding u such that the pair (u, f) minimizes the cost functional

$$\begin{aligned}\mathcal{J}(u,f) =& \frac{\alpha}{2}\int_0^T\int_\Omega K(u)\,d\mathbf{x}\,dt + \frac{\beta}{2}\int_\Omega \Phi_1(u(T,\mathbf{x}))\,d\mathbf{x} + \frac{\gamma}{2}\int_\Omega \Phi_2(u_t(T,\mathbf{x}))\,d\mathbf{x}\\ &+ \frac{1}{2}\int_0^T\int_\Omega |f|^2\,d\mathbf{x}\,dt\end{aligned} \tag{1.2}$$

subject to the wave equation

$$\begin{cases} u_{tt} - \Delta u + \Psi(u) = f \quad \text{in } Q \equiv (0,T)\times\Omega\,,\\ u|_{\partial\Omega} = 0, \quad \text{in } (0,T)\,,\\ u(0,\mathbf{x}) = w(\mathbf{x}) \text{ and } u_t(0,\mathbf{x}) = z(\mathbf{x}) \quad \text{in } \Omega\,. \end{cases} \tag{1.3}$$

Here Ω is a bounded spatial domain in $\mathbb{R}^d$ ($d = 1$ or 2 or 3) with a boundary $\partial\Omega$; f is a distributed control and u is the corresponding state. Also, K, Φ and Ψ are C^1 mappings (for instance, we may choose $K(u) = (u-U)^2$, $\Psi(u) = 0$, $\Psi(u) = u^3 - u$ and $\Psi(u) = e^u$, $\Phi_1(u) = (u(T,\mathbf{x})-W)^2$, $\Phi_2(u) = (u_t(T,\mathbf{x})-Z)^2$, where U, W, Z are target functions). Moreover we assume that initial conditions w and z are smooth enough to be well defined the given problem; see e.g.,[4]. Also we suppose that nonlinearity $\Psi(u)$ does not alter the regularity of the solution in the wave equation.

Of particular interest to us is the case of large α, β and γ; our computational experiments of the proposed numerical method will be performed exclusively for this case. Our interest in this case stems from the fact that the optimal control problem can be viewed as an approximation to the exact distributed controllability problem (1.1).

Such control problems are classical ones in the control theory literature; see, e.g., [12] for the linear case and [13] for the nonlinear case regarding the existence of optimal solutions as well as the existence of a Lagrange multiplier ξ satisfying the optimality system of equations. However, numerical methods for finding discrete (e.g., finite element and/or finite difference) solutions of the optimality system are largely limited to gradient type methods which are sequential in nature and generally require many iterations for convergence. The optimality system involves both initial value and terminal value problems at $t = 0$ and $t = T$ and thus cannot be solved by marching in time. Direct solutions of the discrete optimality system, of course, are bound to be expensive computationally in 2 or 3 spatial dimensions since the problem is $(d + 1)$ dimensional (where d is the spatial dimension.)

The computational algorithms we propose here are based on shooting methods for two-point boundary value problems for ordinary differential equations (ODEs); see, e.g., [2, 3, 6, 7]. The algorithms we propose are well suited for implementation on a parallel computing platform such as a massive cluster of cheap processors.

The rest of this paper is organized as follows. In Section 2 we establish the equivalence between the limit of optimal solutions and the minimum distributed L^2 norm exact controller; this justifies the use of the optimal control approach for solving the exact controllability problem. In Section 3 we formally derive the optimality system of equations for the optimal control problem and discuss the shooting algorithm for solving the optimality system. In Section 4 we state the discrete version of the shooting algorithm for solving the discrete optimality system. Finally in Sections 5 we present computations of certain concrete controllability problems by the shooting method for solving optimal control problems.

2. The solution of the exact controllability problem as the limit of optimal control solutions

We can consider the exact distributed controllability problem (1.1) as the limit of a sequence of optimal control problem. Under suitable assumptions on f, and using Lagrange multiplier rules, we have the corresponding optimal control problem:

$$\text{minimize (1.2) with respect to the control } f \text{ subject to (1.3).} \tag{2.4}$$

The solution to the constrain equations (1.3) is understood in the following weak sense: for any $v \in C^2([0,T]; H^2(\Omega) \cap H_0^1(\Omega))$,

$$\begin{aligned} &\int_0^T \int_\Omega u(v_{tt} - \Delta v)\, d\mathbf{x}\, dt + \int_0^T \int_\Omega [\Psi(u) - f] v\, d\mathbf{x}\, dt + \int_\Omega v|_{t=T} Z(\mathbf{x})\, d\mathbf{x} \\ &- \int_\Omega v|_{t=0} z\, d\mathbf{x} - \int_\Omega W(\mathbf{x})(\partial_t v)|_{t=T}\, d\mathbf{x} + \int_\Omega (w\partial_t v)|_{t=0}\, d\mathbf{x} = 0\,. \end{aligned} \tag{2.5}$$

In this section we establish the equivalence between the limit of optimal solutions and the minimum distributed L^2 norm exact controller. We will show that if $\alpha \to \infty$, $\beta \to \infty$ and $\gamma \to \infty$, then the corresponding optimal solution $(\widehat{u}_{\alpha\beta\gamma}, \widehat{f}_{\alpha\beta\gamma})$ converges weakly to the minimum distributed L^2 norm solution of the exact distributed controllability problem (1.1). The same is also true in the discrete case.

Theorem 2.1. *Assume that the exact distributed controllability problem* (1.1) *admits a unique minimum distributed L^2 norm solution* $(u_{\mathrm{ex}}, f_{\mathrm{ex}})$. *Assume that for every $(\alpha, \beta, \gamma) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ (where $\mathbb{R}_+$ is the set of all positive real numbers,) there exists a solution $(u_{\alpha\beta\gamma}, f_{\alpha\beta\gamma})$ to the optimal control problem* (2.4). *Then*

$$\|f_{\alpha\beta\gamma}\|_{L^2(Q)} \le \|f_{\mathrm{ex}}\|_{L^2(Q)} \qquad \forall\, (\alpha, \beta, \gamma) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+\,. \tag{2.6}$$

Assume, in addition, that for a sequence $\{(\alpha_n, \beta_n, \gamma_n)\}$ *satisfying* $\alpha_n \to \infty$, $\beta_n \to \infty$ *and* $\gamma_n \to \infty$,

$$\begin{aligned} &u_{\alpha_n\beta_n\gamma_n} \rightharpoonup \overline{u} \text{ in } L^2(Q) \quad \text{and} \\ &\Psi(u_{\alpha_n\beta_n\gamma_n}) \rightharpoonup \Psi(\overline{u}) \text{ in } L^2(0,T;[H^2(\Omega)\cap H_0^1(\Omega)]^*)\,. \end{aligned} \tag{2.7}$$

Then

$$f_{\alpha_n\beta_n\gamma_n} \rightharpoonup f_{\text{ex}} \text{ in } L^2(Q) \text{ and } u_{\alpha_n\beta_n\gamma_n} \rightharpoonup u_{\text{ex}} \text{ in } L^2(Q) \text{ as } n \to \infty\,. \tag{2.8}$$

Furthermore, if (2.7) *holds for every sequence* $\{(\alpha_n, \beta_n, \gamma_n)\}$ *satisfying* $\alpha_n \to \infty$, $\beta_n \to \infty$ *and* $\gamma_n \to \infty$, *then*

$$f_{\alpha\beta\gamma} \rightharpoonup f_{\text{ex}} \text{ in } L^2(Q) \text{ and } u_{\alpha\beta\gamma} \rightharpoonup u_{\text{ex}} \text{ in } L^2(Q) \text{ as } \alpha,\beta,\gamma \to \infty\,. \tag{2.9}$$

Proof. Since $(u_{\alpha\beta\gamma}, f_{\alpha\beta\gamma})$ is an optimal solution, we have that

$$\begin{aligned} &\frac{\alpha}{2}\|u_{\alpha\beta\gamma} - U\|_{L^2(Q)} + \frac{\beta}{2}\|u_{\alpha\beta\gamma}(T) - W\|_{L^2(\Omega)} + \frac{\gamma}{2}\|\partial_t u_{\alpha\beta\gamma}(T) - Z\|_{H^{-1}(\Omega)} \\ &+ \frac{1}{2}\|f_{\alpha\beta\gamma}\|_{L^2(Q)} = \mathcal{J}(u_{\alpha\beta\gamma}, f_{\alpha\beta\gamma}) \le \mathcal{J}(u_{\text{ex}}, f_{\text{ex}}) = \frac{1}{2}\|f_{\text{ex}}\|_{L^2(Q)} \end{aligned}$$

so that (2.6) holds,

$$u_{\alpha\beta\gamma}|_{t=T} \to W \text{ in } L^2(\Omega) \quad \text{and} \quad (\partial_t u_{\alpha\beta\gamma})|_{t=T} \to Z \text{ in } H^{-1}(\Omega)\,, \tag{2.10}$$

as $\alpha, \beta, \gamma \to \infty$. Let $\{(\alpha_n, \beta_n, \gamma_n)\}$ be the sequence in (2.7). Estimate (2.6) implies that a subsequence of $\{(\alpha_n, \beta_n, \gamma_n)\}$, denoted by the same, satisfies

$$f_{\alpha_n\beta_n\gamma_n} \rightharpoonup \overline{f} \text{ in } L^2(Q) \quad \text{and} \quad \|\overline{f}\|_{L^2(Q)} \le \|f_{\text{ex}}\|_{L^2(Q)}\,. \tag{2.11}$$

$(u_{\alpha\beta\gamma}, f_{\alpha\beta\gamma})$ satisfies the initial value problem in the weak form:

$$\begin{aligned} &\int_0^T\int_\Omega u_{\alpha\beta\gamma}(v_{tt} - \Delta v)\,d\mathbf{x}\,dt + \int_0^T\int_\Omega [\Psi(u_{\alpha\beta\gamma}) - f_{\alpha\beta\gamma}]v\,d\mathbf{x}\,dt \\ &+ \int_\Omega (v\partial_t u_{\alpha\beta\gamma})|_{t=T}\,d\mathbf{x} - \int_\Omega v|_{t=0} z\,d\mathbf{x} - \int_\Omega (u_{\alpha\beta\gamma}\partial_t v)|_{t=T}\,d\mathbf{x} \\ &+ \int_\Omega (w\partial_t v)|_{t=0}\,d\mathbf{x} = 0 \qquad \forall\, v \in C^2([0,T]; H^2(\Omega)\cap H_0^1(\Omega))\,. \end{aligned} \tag{2.12}$$

Passing to the limit in (2.12) as $\alpha, \beta, \gamma \to \infty$ and using relations (2.10) and (2.11), we obtain:

$$\begin{aligned} &\int_0^T\int_\Omega \overline{u}(v_{tt} - \Delta v)\,d\mathbf{x}\,dt + \int_0^T\int_\Omega [\Psi(\overline{u}) - \overline{f}]v\,d\mathbf{x}\,dt \\ &+ \int_\Omega v|_{t=T} Z(\mathbf{x})\,d\mathbf{x} - \int_\Omega v|_{t=0} z\,d\mathbf{x} - \int_\Omega W(\mathbf{x})(\partial_t v)|_{t=T}\,d\mathbf{x} \\ &+ \int_\Omega (w\partial_t v)|_{t=0}\,d\mathbf{x} = 0 \qquad \forall\, v \in C^2([0,T]; H^2(\Omega)\cap H_0^1(\Omega))\,. \end{aligned}$$

The last relation and (2.11) imply that $(\overline{u}, \overline{f})$ is a minimum distributed L^2 norm solution to the exact controllability problem (1.1). Hence, $\overline{u} = u_{\text{ex}}$ and $\overline{f} = f_{\text{ex}}$ so that (2.8) and (2.9) follows from (2.7) and (2.11). □

Remark 2.2. *If the wave equation is linear, i.e.,* $\Psi = 0$, *then assumption* (2.7) *is redundant and* (2.9) *is guaranteed to hold. Indeed,* (2.12) *implies the boundedness of* $\{\|u_{\alpha\beta\gamma}\|_{L^2(Q)}\}$ *which in turn yields* (2.7). *The uniqueness of a solution for the linear wave equation implies* (2.7) *holds for an arbitrary sequence* $\{(\alpha_n, \beta_n, \gamma_n)\}$.

Theorem 2.3. *Assume that* i) *for every* $(\alpha,\beta,\gamma)\in\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}_+$ *there exists a solution* $(u_{\alpha\beta\gamma},f_{\alpha\beta\gamma})$ *to the optimal control problem* (2.4)*;*

ii) *the limit terminal conditions hold:*

$$u_{\alpha\beta\gamma}|_{t=T}\to W \text{ in } L^2(\Omega) \text{ and } (\partial_t u_{\alpha\beta\gamma})|_{t=T}\to Z \text{ in } H^{-1}(\Omega) \tag{2.13}$$

as $\alpha,\beta,\gamma\to\infty$ *;*

iii) *the optimal solution* $(u_{\alpha\beta\gamma},f_{\alpha\beta\gamma})$ *satisfies the weak limit conditions as* $\alpha,\beta,\gamma\to\infty$*:*

$$f_{\alpha\beta\gamma}\rightharpoonup\overline{f} \text{ in } L^2(Q)\,,\quad u_{\alpha\beta\gamma}\rightharpoonup\overline{u} \text{ in } L^2(Q)\,, \tag{2.14}$$

and

$$\Psi(u_{\alpha\beta\gamma})\rightharpoonup\Psi(\overline{u}) \text{ in } L^2(0,T;[H^2(\Omega)\cap H_0^1(\Omega)]^*) \tag{2.15}$$

for some $\overline{f}\in L^2(Q)$ *and* $\overline{u}\in L^2(Q)$*. Then* $(\overline{u},\overline{f})$ *is a solution to the exact distributed controllability problem* (1.1) *with* $\overline{f}$ *satisfying the minimum distributed* L^2 *norm property. Furthermore, if the solution to* (1.1) *admits a unique solution* $(u_{\rm ex},f_{\rm ex})$*, then*

$$f_{\alpha\beta\gamma}\rightharpoonup f_{\rm ex} \text{ in } L^2(Q) \text{ and } u_{\alpha\beta\gamma}\rightharpoonup u_{\rm ex} \text{ in } L^2(Q) \text{ as } \alpha,\beta,\gamma\to\infty\,. \tag{2.16}$$

Proof. $(u_{\alpha\beta\gamma},f_{\alpha\beta\gamma})$ satisfies (2.12). Passing to the limit in that equation as $\alpha,\beta,\gamma\to\infty$ and using relations (2.13), (2.14) and (2.15) we obtain:

$$\begin{aligned}&\int_0^T\int_\Omega\overline{u}(v_{tt}-\Delta v)\,d\mathbf{x}\,dt+\int_0^T\int_\Omega[\Psi(\overline{u})-\overline{f}]v\,d\mathbf{x}\,dt\\&+\int_\Omega v|_{t=T}Z(\mathbf{x})\,d\mathbf{x}-\int_\Omega v|_{t=0}z\,d\mathbf{x}-\int_\Omega W(\mathbf{x})(\partial_t v)|_{t=T}\,d\mathbf{x}\\&+\int_\Omega(w\partial_t v)|_{t=0}\,d\mathbf{x}=0\qquad\forall\,v\in C^2([0,T];H^2(\Omega)\cap H_0^1(\Omega))\,.\end{aligned}$$

This implies that $(\overline{u},\overline{f})$ is a solution to the exact distributed controllability problem (1.1).

To prove that $\overline{f}$ satisfies the minimum distributed L^2 norm property, we proceeds as follows. Let $(u_{\rm ex},f_{\rm ex})$ denote an exact minimum distributed L^2 norm solution to the exact controllability problem (1.1). Since $(u_{\alpha\beta\gamma},f_{\alpha\beta\gamma})$ is an optimal solution, we have that

$$\begin{aligned}&\frac{\alpha}{2}\|u_{\alpha\beta\gamma}-U\|^2_{L^2(Q)}+\frac{\beta}{2}\|u_{\alpha\beta\gamma}(T)-W\|^2_{L^2(\Omega)}+\frac{\gamma}{2}\|\partial_t u_{\alpha\beta\gamma}(T)-Z\|^2_{H^{-1}(\Omega)}\\&+\frac{1}{2}\|f_{\alpha\beta\gamma}\|^2_{L^2(Q)}=\mathcal{J}(u_{\alpha\beta\gamma},f_{\alpha\beta\gamma})\le\mathcal{J}(u_{\rm ex},f_{\rm ex})=\frac{1}{2}\|f_{\rm ex}\|^2_{L^2(Q)}\end{aligned}$$

so that

$$\|f_{\alpha\beta\gamma}\|^2_{L^2(Q)}\le\|f_{\rm ex}\|^2_{L^2(Q)}\,.$$

Passing to the limit in the last estimate we obtain

$$\|\overline{f}\|^2_{L^2(Q)}\le\|f_{\rm ex}\|^2_{L^2(Q)}\,. \tag{2.17}$$

Hence we conclude that $(\overline{u},\overline{f})$ is a minimum distributed L^2 norm solution to the exact distributed controllability problem (1.1).

Furthermore, if the exact controllability problem (1.1) admits a unique minimum distributed L^2 norm solution $(u_{\rm ex},f_{\rm ex})$, then $(\overline{u},\overline{f})=(u_{\rm ex},f_{\rm ex})$ and (2.16) follows from assumption (2.14). □

Remark 2.4. *If the wave equation is linear, i.e.,* $\Psi=0$*, then assumptions i) and* (2.15) *are redundant.*

Remark 2.5. *Assumptions ii) and iii) hold if $f_{\alpha\beta\gamma}$ and $u_{\alpha\beta\gamma}$ converges pointwise as $\alpha, \beta, \gamma \to \infty$.*

Remark 2.6. *A practical implication of Theorem 2.3 is that one can prove the exact controllability problem for semilinear wave equations by examining the behavior of a sequence of optimal solutions (recall that exact controllability problem was proved only for some special classes of semilinear wave equations.) If we have found a sequence of optimal control solutions $\{(u_{\alpha_n\beta_n\gamma_n}, f_{\alpha_n\beta_n\gamma_n})\}$ where $\alpha_n, \beta_n, \gamma_n \to \infty$ and this sequence appears to satisfy the convergence assumptions ii) and iii), then we can confidently conclude that the underlying semilinear wave equation is exactly controllable and the optimal solution $(u_{\alpha_n\beta_n\gamma_n}, f_{\alpha_n\beta_n\gamma_n})$ when n is large provides a good approximation to the minimum distributed L^2 norm exact controller $(u_{\rm ex}, f_{\rm ex})$.*

3. An optimality system of equations and a continuous shooting method

Under suitable assumptions on f and through the use of Lagrange multiplier rules, the optimal control problem

$$\text{minimize (1.2) with respect to the control } f \text{ subject to (1.3)} \tag{3.18}$$

may be converted into the following system of equations from which an optimal solution may be determined:

$$\left\{\begin{array}{l} u_{tt} - \Delta u + \Psi(u) = f \quad \text{in } Q \equiv (0,T)\times\Omega\,, \\ u|_{\partial\Omega} = 0 \quad \text{in } (0,T)\,, \\ u(0,\mathbf{x}) = w(\mathbf{x}) \text{ and } u_t(0,\mathbf{x}) = z(\mathbf{x}) \quad \text{in } \Omega\,, \\ \xi_{tt} - \Delta\xi + [\Psi'(u)]^*\xi = \frac{\alpha}{2}K'(u) \quad \text{in } Q\,, \\ \xi|_{\partial\Omega} = 0 \quad \text{in } \Omega\,, \\ \xi(T,\mathbf{x}) = -\frac{\gamma}{2}A^{-1}(\Phi_2'(u_t(T,\mathbf{x}))) \quad \text{in } \Omega\,, \\ \xi_t(T,\mathbf{x}) = -\frac{\beta}{2}\Phi_1'(u(T,\mathbf{x})) \quad \text{in } \Omega\,, \\ f + \xi = 0 \quad \text{in } Q\,, \end{array}\right. \tag{3.19}$$

where the elliptic operator $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined by $Av = \Delta v$ for all $v \in H_0^1(\Omega)$. By eliminating f in the system the optimality system may be simplified as

$$\left\{\begin{array}{l} u_{tt} - \Delta u + \Psi(u) = -\xi \quad \text{in } Q \equiv (0,T)\times\Omega\,, \\ u|_{\partial\Omega} = 0 \quad \text{in } (0,T)\,, \\ u(0,\mathbf{x}) = w(\mathbf{x}) \text{ and } u_t(0,\mathbf{x}) = z(\mathbf{x}) \quad \text{in } \Omega\,, \\ \xi_{tt} - \Delta\xi + [\Psi'(u)]^*\xi = \frac{\alpha}{2}K'(u) \quad \text{in } Q\,, \\ \xi|_{\partial\Omega} = 0 \quad \text{in } \Omega\,, \\ \xi(T,\mathbf{x}) = -\frac{\gamma}{2}A^{-1}(\Phi_2(u_t(T,\mathbf{x}))) \quad \text{in } \Omega\,, \\ \xi_t(T,\mathbf{x}) = -\frac{\beta}{2}\Phi_1'(u(T,\mathbf{x})) \quad \text{in } \Omega\,. \end{array}\right. \tag{3.20}$$

Derivations and justifications of optimality systems are discussed in [12] for the linear case and in [13] for the semilinear case.

The computational algorithm we propose in the paper is a shooting method for solving the optimality system of equations. The basic idea for a shooting method is to convert the solution of an initial-terminal value problem into that of a purely initial value problem (IVP); see, e.g., [2] for a discussion of shooting methods for systems of ordinary differential equations. The IVP corresponding to the optimality system (3.20) is described by

$$
\left\{
\begin{array}{l}
u_{tt} - \Delta u + \Psi(u) = -\xi \quad \text{in } Q \equiv (0,T) \times \Omega\,, \\
u|_{\partial\Omega} = 0 \quad \text{in } (0,T)\,, \\
u(0,\mathbf{x}) = w(\mathbf{x}) \text{ and } u_t(0,\mathbf{x}) = z(\mathbf{x}) \quad \text{in } \Omega\,, \\
\xi_{tt} - \Delta\xi + [\Psi'(u)]^*\xi = \frac{\alpha}{2}K'(u) \quad \text{in } Q\,, \\
\xi|_{\partial\Omega} = 0 \quad \text{in } \Omega\,, \\
\xi(0,\mathbf{x}) = \omega(\mathbf{x}) \text{ and } \xi_t(0,\mathbf{x}) = \theta(\mathbf{x}) \quad \text{in } \Omega\,,
\end{array}
\right.
\tag{3.21}
$$

with unknown initial values ω and θ. Then the goal is to choose ω and θ such that the solution (u,ξ) of the IVP (3.21) satisfies the terminal conditions

$$
\begin{aligned}
F_1(\omega,\theta) &\equiv \Delta\xi(T,x) + \frac{\gamma}{2}\Phi_2'(u_t(T,x)) = 0\,, \\
F_2(\omega,\theta) &\equiv \xi_t(T,x) + \frac{\beta}{2}\Phi_1'(u(T,x)) = 0\,.
\end{aligned}
\tag{3.22}
$$

A shooting method for solving (3.20) consists of the following main steps:

choose initial guesses ω and θ;
for $iter = 1, 2, \cdots, max_{iter}$
  solve for (u,ξ) from the IVP (3.21)
  update ω and θ:
  $(\omega^{\text{new}}, \theta^{\text{new}}) = (\omega,\theta) - [F'(\omega,\theta)]^{-1}F(\omega,\theta)$;
  if $F(\omega^{\text{new}}, \theta^{\text{new}}) = \mathbf{0}$, stop; otherwise, set $(\omega,\theta) = (\omega^{\text{new}}, \theta^{\text{new}})$.

A criterion for updating (ω,θ) can be derived from the terminal conditions (3.22). A method for solving the nonlinear system (3.22) (as a system for the unknowns ω and θ) will yield an updating formula and here we invoke the well-known Newton's method to do so. Also, a discrete version of the algorithm must be used in actual implementations. For ease of exposition we describe in detail a particular problem. We apply a shooting algorithm based on Newton's method with finite difference discretizations in two space dimension. A discussion of Newton's method for an infinite dimensional nonlinear system can be found in many functional analysis textbooks, and for the suitable assumption convergence of Newton iteration for the optimality system is guaranteed.

4. The discrete shooting methods for 2D control problems

Assume $\Omega = [0,X] \times [0,Y] \subset \mathbb{R}^2$. The basic idea for a shooting method in 2D is exactly the same as in 1D;see, e.g., [19]. In two dimensional case, we replace u_{xx} by Δu. Lagrange multiplier rules provide the same optimality system of equations as in the previous section except the replacement part. Thus the IVP corresponding to the optimality system (3.20) is described by

$$
\left\{
\begin{array}{l}
u_{tt} - \Delta u + \Psi(u) = -\xi \quad \text{in } Q \equiv (0,T) \times ([0,X] \times [0,Y])\,, \\
u|_{\partial\Omega} = 0 \quad \text{in } (0,T)\,, \\
u(0,\mathbf{x}) = w(\mathbf{x}) \text{ and } u_t(0,\mathbf{x}) = z(\mathbf{x}) \quad \text{in } [0,X] \times [0,Y]\,, \\
\xi_{tt} - \Delta\xi + [\Psi'(u)]^*\xi = \frac{\alpha}{2}K'(u) \quad \text{in } Q\,, \\
\xi|_{\partial\Omega} = 0 \quad \text{in } [0,X] \times [0,Y]\,, \\
\xi(0,\mathbf{x}) = \omega(\mathbf{x}) \text{ and } \xi_t(0,\mathbf{x}) = \theta(\mathbf{x}) \quad \text{in } [0,X] \times [0,Y]\,,
\end{array}
\right.
\tag{4.23}
$$

with unknown initial values ω and θ. Then the goal is to choose ω and θ such that the solution (u,ξ) of the IVP (4.23) satisfies the terminal conditions

$$\xi(T,\mathbf{x}) = -\frac{\gamma}{2}A^{-1}(\Phi_2'(u_t(T,\mathbf{x}))) \text{ and } \xi_t(T,\mathbf{x}) = -\frac{\beta}{2}\Phi_1'(u(T,\mathbf{x}))\,. \tag{4.24}$$

In two dimension, we discretize the spatial interval $[0,X]\times[0,Y]$ into $0=x_0<x_1<x_2<x_3<\cdots<x_{I+1}=X$, $0=y_0<y_1<y_2<y_3<\cdots<y_{J+1}=Y$ with a uniform spacing $h_x=X/(I+1)$, $h_y=Y/(J+1)$ respectively, and we divide the time horizon $[0,T]$ into $0=t_1<t_2<t_3<\cdots<t_N=T$ with a uniform time step length $\delta=T/(N-1)$. We use the central difference scheme to approximate the initial value problem (4.23): For $i=1,2,\cdots,I$, $j=1,2,\cdots,J$,

$$\begin{aligned}
&u_{ij}^1 = w_{ij}, \quad u_{ij}^2 = w_{ij}+\delta z_{ij}\,, \quad \xi_{ij}^1=\omega_{ij}, \quad \xi_{ij}^2=\xi_{ij}^1+\delta\theta_{ij}\,;\\
&u_{ij}^{n+1} = -u_{ij}^{n-1}+\lambda_x(u_{i-1,j}^n+u_{i+1,j}^n)+2(1-\lambda_x-\lambda_y)u_{ij}^n\\
&\qquad\quad +\lambda_y(u_{i,j-1}^n+u_{i,j+1}^n)-\delta^2\xi_{ij}^n-\delta^2\Psi(u_{ij}^n),\\
&\xi_{ij}^{n+1} = -\xi_{ij}^{n-1}+\lambda_x(\xi_{i-1,j}^n+\xi_{i+1,j}^n)+2(1-\lambda_x-\lambda_y)\xi_{ij}^n\\
&\qquad\quad +\lambda_y(\xi_{i,j-1}^n+\xi_{i,j+1}^n)+\delta^2\frac{\alpha}{2}K(u_{ij}^n)-\delta^2[\Psi'(u_{ij}^n)]^*\xi_{ij}^n,
\end{aligned} \tag{4.25}$$

where $\lambda_x=(\delta/h_x)^2$, $\lambda_y=(\delta/h_y)^2$ (we also use the convention that $u_{ij}^n=\xi_{ij}^n=0$ if $i=0$ or $I+1$ or $j=0$ or $j=J+1$.) A discrete shooting method concentrates on the discrete terminal conditions: For $i=1,2,\cdots,I$, $j=1,2,\cdots,J$,

$$\begin{aligned}
&F_{2(ij)^*-1} \equiv \frac{\xi_{ij}^N-\xi_{ij}^{N-1}}{\delta}+\frac{\beta}{2}\Phi_1'(u_{ij}^N)=0\,,\\
&F_{2(ij)^*} \equiv \Big(\frac{\xi_{i-1,j}^N-2\xi_{i,j}^N+\xi_{i+1,j}^N}{h_x^2}+\frac{\xi_{i,j-1}^N-2\xi_{i,j}^N+\xi_{i,j+1}^N}{h_y^2}\Big)\\
&\qquad\quad +\frac{\gamma}{2}\Phi_2'\Big(\frac{u_{ij}^N-u_{ij}^{N-1}}{\delta}\Big)=0\,,
\end{aligned} \tag{4.26}$$

where $(ij)^*$ is a reordering of the nodes with respect to X,Y except boundary points.

Let $\omega(\mathbf{x})\equiv\{\omega_1,\omega_2,\cdots,\omega_{IJ^*}\}=\{\omega_{11},\omega_{21},\cdots,\omega_{IJ}\}$, and $\theta(\mathbf{x})\equiv\{\theta_1,\theta_2,\cdots,\theta_{IJ^*}\}=\{\theta_{11},\theta_{21},\cdots,\theta_{IJ}\}$ where $IJ^*=I\cdot J$. By denoting

$$\begin{aligned}
q_{(ij)^*k}^n &= q_{(ij)^*k}^n(\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*}) = \frac{\partial u_{ij}^n}{\partial\omega_k},\\
r_{(ij)^*k}^n &= r_{(ij)^*k}^n(\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*}) = \frac{\partial u_{ij}^n}{\partial\theta_k},\\
\rho_{(ij)^*k}^n &= \rho_{(ij)^*k}^n(\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*}) = \frac{\partial \xi_{ij}^n}{\partial\omega_k}\,,\\
\tau_{(ij)^*k}^n &= \tau_{(ij)^*k}^n(\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*}) = \frac{\partial \xi_{ij}^n}{\partial\theta_k}\,,
\end{aligned}$$

we may write Newton's iteration formula as

$$\begin{aligned}
(\omega_1^{\text{new}},&\theta_1^{\text{new}},\omega_2^{\text{new}},\theta_2^{\text{new}},\cdots,\omega_{IJ^*}^{\text{new}},\theta_{IJ^*}^{\text{new}})^T = (\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*})^T\\
&-[F'(\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*})]^{-1}F(\omega_1,\theta_1,\omega_2,\theta_2,\cdots,\omega_{IJ^*},\theta_{IJ^*})
\end{aligned}$$

where the vector F and Jacobian matrix $J = F'$ are defined by

$$F_{2(ij)^*-1} \equiv \frac{\xi_{ij}^N - \xi_{ij}^{N-1}}{\delta} + \frac{\beta}{2}\Phi_1'(u_{ij}^N) = 0\,,$$
$$F_{2(ij)^*} \equiv (\frac{\xi_{i-1,j}^N - 2\xi_{ij}^N + \xi_{i+1,j}^N}{h_x^2} + \frac{\xi_{i,j-1}^N - 2\xi_{ij}^N + \xi_{i,j+1}^N}{h_y^2}) + \frac{\gamma}{2}\Phi_2'(\frac{u_{ij}^N - u_{ij}^{N-1}}{\delta}) = 0\,,$$

$$J_{2(ij)^*-1,2k-1} = \frac{\rho_{(ij)_k^*}^N - \rho_{(ij)_k^*}^{N-1}}{\delta} + \frac{\beta}{2}\Phi_1''(u_{ij}^N) q_{(ij)_k^*}^N\,,$$
$$J_{2(ij)^*-1,2k} = \frac{\tau_{(ij)_k^*}^N - \tau_{(ij)_k^*}^{N-1}}{\delta} + \frac{\beta}{2}\Phi_1''(u_{ij}^N) r_{(ij)_k^*}^N\,,$$
$$\begin{aligned} J_{2(ij)^*,2k-1} = {} & (\frac{\rho_{(i-1,j)^*}^N - 2\rho_{(ij)^*}^N + \rho_{(i+1,j)^*}^N}{h_x^2} + \frac{\rho_{(i,j-1)^*}^N - 2\rho_{(ij)^*}^N + \rho_{(i,j+1)^*}^N}{h_y^2}) \\ & + \frac{\gamma}{2\delta}\Phi_2''(\frac{u_{ij}^N - u_{ij}^{N-1}}{\delta})(q_{(ij)_k^*}^N - q_{(ij)_k^*}^{N-1})\,, \end{aligned}$$
$$\begin{aligned} J_{2(ij)^*,2k} = {} & (\frac{\tau_{(i-1,j)^*}^N - 2\tau_{(ij)^*}^N + \tau_{(i+1,j)^*}^N}{h_x^2} + \frac{\tau_{(i,j-1)^*}^N - 2\tau_{(ij)^*}^N + \tau_{(i,j+1)^*}^N}{h_y^2}) \\ & + \frac{\gamma}{2\delta}\Phi_2''(\frac{u_{ij}^N - u_{ij}^{N-1}}{\delta})(r_{(ij)_k^*}^N - r_{(ij)_k^*}^{N-1})\,. \end{aligned}$$

Moreover, by differentiating (4.25) with respect to ω_k and θ_k we obtain the equations for determining $q_{(ij)^*k}$, $r_{(ij)^*k}$, $\rho_{(ij)^*k}$ and $\tau_{(ij)^*k}$:

$$\begin{aligned} & q_{(ij)^*k}^1 = 0, \quad q_{(ij)^*k}^2 = 0\,, \quad r_{(ij)^*k}^1 = 0\,, \quad r_{(ij)^*k}^2 = 0\,, \\ & \rho_{(ij)^*k}^1 = \delta_{(ij)^*k}\,, \quad \rho_{(ij)^*k}^2 = \delta_{(ij)^*k}\,, \qquad\qquad i,j = 1,2,\cdots,I,J\,; \\ & \tau_{(ij)^*k}^1 = 0\,, \quad \tau_{(ij)^*k}^2 = \delta\delta_{(ij)^*k}\,, \end{aligned}$$

$$\begin{aligned} q_{(ij)^*k}^{n+1} = {} & -q_{(ij)^*k}^{n-1} + \lambda_x(q_{(i-1,j)^*k}^n + q_{(i+1,j)^*k}^n) + 2(1-\lambda_x-\lambda_y) q_{(ij)^*k}^n \\ & + \lambda_y(q_{(i,j-1)^*k}^n + q_{(i,j+1)^*k}^n) - \delta^2\rho_{(ij)^*k}^n \\ & - \delta^2\Psi'(u_{ij}^n) q_{(ij)^*k}^n, \qquad i,j = 1,2,\cdots,I,J\,; \end{aligned}$$

$$\begin{aligned} r_{(ij)^*k}^{n+1} = {} & -r_{(ij)^*k}^{n-1} + \lambda_x(r_{(i-1,j)^*k}^n + r_{(i+1,j)^*k}^n) + 2(1-\lambda_x-\lambda_y) r_{(ij)^*k}^n \\ & + \lambda_y(r_{(i,j-1)^*k}^n + r_{(i,j+1)^*k}^n) - \delta^2\tau_{(ij)^*k}^n \\ & - \delta^2\Psi'(u_{ij}^n) r_{(ij)^*k}^n, \qquad i,j = 1,2,\cdots,I,J\,; \end{aligned}$$

$$\begin{aligned} \rho_{(ij)^*k}^{n+1} = {} & -\rho_{(ij)^*k}^{n-1} + \lambda_x(\rho_{(i-1,j)^*k}^n + \rho_{(i+1,j)^*k}^n) + 2(1-\lambda_x-\lambda_y)\rho_{(ij)^*k}^n \\ & + \lambda_y(\rho_{(i,j-1)^*k}^n + \rho_{(i,j+1)^*k}^n) + \delta^2\frac{\alpha}{2}K'(u_{ij}^n) q_{(ij)^*k}^N - \delta^2[\Psi'(u_i^n)]^*\rho_{(ij)_k^*}^n \\ & - \delta^2[\Psi''(u_i^n) q_{(ij)^*k}^N]^*\xi_{ij}^n, \qquad i,j = 1,2,\cdots,I,J\,; \end{aligned}$$

$$\begin{aligned} \tau_{(ij)^*k}^{n+1} = {} & -\tau_{(ij)^*k}^{n-1} + \lambda_x(\tau_{(i-1,j)^*k}^n + \tau_{(i+1,j)^*k}^n) + 2(1-\lambda_x-\lambda_y)\tau_{(ij)^*k}^n \\ & + \lambda_y(\tau_{(i,j-1)^*k}^n + \tau_{(i,j+1)^*k}^n) + \delta^2\frac{\alpha}{2}K'(u_{ij}^n) r_{(ij)^*k}^N - \delta^2[\Psi'(u_{ij}^n)]^*\rho_{(ij)^*k}^n \\ & - \delta^2[\Psi''(u_{ij}^n) r_{(ij)^*k}^N]^*\xi_{ij}^n, \qquad i,j = 1,2,\cdots,I,J\,; \end{aligned}$$

where δ_{ij} is the Chronecker delta. Thus, we have the following Newton's-method-based shooting algorithm:

Algorithm 2 – Newton method based shooting algorithm with Euler discretizations for distributed optimal control problems

choose initial guesses ω_{ij} and θ_{ij}, $i,j=1,2,\cdots,I,J$;
% set initial conditions for u and ξ
for $i,j=1,2,\cdots,I,J$

$$u_{ij}^1 = w_{ij}, \quad u_{ij}^2 = w_{ij} + \delta z_{ij},$$
$$\xi_{ij}^1 = \omega_{ij}, \quad \xi_{ij}^2 = \xi_{ij}^1 + \delta\theta_{ij};$$

% set initial conditions for $q_{(ij)^*k}, r_{(ij)^*k}, \rho_{(ij)^*k}, \tau_{(ij)^*k}$
for $k=1,2,\cdots,(IJ)^*$
for $i=1,2,\cdots,I$
for $j=1,2,\cdots,J$

$$q_{(ij)^*k}^1 = 0, \quad q_{(ij)^*k}^2 = 0, \quad r_{(ij)^*k}^1 = 0, \quad r_{(ij)^*k}^2 = 0,$$
$$\rho_{(ij)^*k}^1 = 0, \quad \rho_{(ij)^*k}^2 = 0, \quad \tau_{(ij)^*k}^1 = 0, \quad \tau_{(ij)^*k}^2 = 0,$$
$$\rho_{kk}^1 = 1, \quad \rho_{kk}^2 = 1, \quad \tau_{kk}^2 = \delta;$$

% Newton iterations
for $m=1,2,\cdots,M$
% solve for (u,ξ)
for $n=2,3,\cdots,N-1$

$$\begin{aligned}u_{ij}^{n+1} &= -u_{ij}^{n-1} + \lambda_x(u_{i-1,j}^n + u_{i+1,j}^n) + 2(1-\lambda_x-\lambda_y)u_{ij}^n \\ &\quad +\lambda_y(u_{i,j-1}^n + u_{i,j+1}^n),\\ \xi_{ij}^{n+1} &= -\xi_{ij}^{n-1} + \lambda_x(\xi_{i-1,j}^n + \xi_{i+1,j}^n) + 2(1-\lambda_x-\lambda_y)\xi_{ij}^n \\ &\quad +\lambda_y(\xi_{i,j-1}^n + \xi_{i,j+1}^n) + \delta^2\tfrac{\alpha}{2}K(u_{ij}^n) - \delta^2[\Psi'(u_{ij}^n)]^*\xi_{ij}^n;\end{aligned}$$

% solve for q,r,ρ,τ
for $j=1,2,\cdots,I$
for $n=2,3,\cdots,N-1$
for $i=2,\cdots,N-1$

$$\begin{aligned}q_{(ij)^*k}^{n+1} &= -q_{(ij)^*k}^{n-1} + \lambda_x(q_{(i-1,j)^*k}^n + q_{(i+1,j)^*k}^n) \\ &\quad +2(1-\lambda_x-\lambda_y)q_{(ij)^*k}^n + \lambda_y(q_{(i,j-1)^*k}^n + q_{(i,j+1)^*k}^n) \\ &\quad -\delta^2\rho_{(ij)^*k}^n - \delta^2\Psi'(u_{ij}^n)q_{(ij)^*k}^n,\\ r_{(ij)^*k}^{n+1} &= -r_{(ij)^*k}^{n-1} + \lambda_x(r_{(i-1,j)^*k}^n + r_{(i+1,j)^*k}^n) \\ &\quad +2(1-\lambda_x-\lambda_y)r_{(ij)^*k}^n + \lambda_y(r_{(i,j-1)^*k}^n + r_{(i,j+1)^*k}^n) \\ &\quad -\delta^2\tau_{(ij)^*k}^n - \delta^2\Psi'(u_{ij}^n)r_{(ij)^*k}^n,\\ \rho_{(ij)^*k}^{n+1} &= -\rho_{(ij)}^{n-1} + \lambda_x(\rho_{(i-1,j)^*k}^n + \rho_{(i+1,j)^*k}^n) \\ &\quad +2(1-\lambda_x-\lambda_y)\rho_{(ij)^*k}^n + \lambda_y(\rho_{(i,j-1)^*k}^n + \rho_{(i,j+1)^*k}^n) \\ &\quad +\delta^2\tfrac{\alpha}{2}K'(u_{ij}^n)q_{(ij)^*k}^N - \delta^2[\Psi'(u_i^n)]^*\rho_{ij}^n - \delta^2[\Psi''(u_i^n)q_{ij}^N]^*\xi_{ij}^n,\\ \tau_{(ij)^*k}^{n+1} &= -\tau_{(ij)^*k}^{n-1} + \lambda_x(\tau_{(i-1,j)^*k}^n + \tau_{(i+1,j)^*k}^n) \\ &\quad +2(1-\lambda_x-\lambda_y)\tau_{(ij)^*k}^n + \lambda_y(\tau_{(i,j-1)^*k}^n + \tau_{(i,j+1)^*k}^n) \\ &\quad +\delta^2\tfrac{\alpha}{2}K'(u_{ij}^n)r_{(ij)^*k}^N - \delta^2[\Psi'(u_{ij}^n)]^*\rho_{ij}^n - \delta^2[\Psi''(u_{ij}^n)r_i^N]^*\xi_{ij}^n;\end{aligned}$$

% (we need to build into the algorithm the following:
$q_0^n = r_0^n = \rho_0^n = \tau_0^n = 0,$
$q_{I+1}^n = r_{I+1}^n = \rho_{I+1}^n = \tau_{I+1}^n = 0.)$

% evaluate F and F'
for $i, j = 1, 2, \cdots, I, J$

$$F_{2(ij)^*-1} \equiv \frac{\xi_{ij}^N - \xi_{ij}^{N-1}}{\delta} + \frac{\beta}{2}\Phi_1'(u_{ij}^N),$$

$$F_{2(ij)^*} \equiv \left(\frac{\xi_{i-1,j}^N - 2\xi_{ij}^N + \xi_{i+1,j}^N}{h_x^2} + \frac{\xi_{i,j-1}^N - 2\xi_{ij}^N + \xi_{i,j+1}^N}{h_y^2}\right) + \frac{\gamma}{2}\Phi_2'\left(\frac{u_{ij}^N - u_{ij}^{N-1}}{\delta}\right);$$

for $j = 1, 2, \cdots, I$

$$J_{2(ij)^*-1,2k-1} = \frac{\rho_{(ij)^*k}^N - \rho_{(ij)^*k}^{N-1}}{\delta} + \frac{\beta}{2}\Phi_1''(u_{ij}^N) q_{(ij)^*k}^N,$$

$$J_{2(ij)^*-1,2k} = \frac{\tau_{(ij)^*k}^N - \tau_{(ij)^*k}^{N-1}}{\delta} + \frac{\beta}{2}\Phi_1''(u_{ij}^N) r_{(ij)^*k}^N,$$

$$J_{2(ij)^*,2k-1} = \left(\frac{\tau_{(i-1,j)^*}^N - 2\tau_{(ij)^*}^N + \tau_{(i+1,j)^*}^N}{h_x^2} + \frac{\tau_{(i,j-1)^*}^N - 2\tau_{(ij)^*}^N + \tau_{(i,j+1)^*}^N}{h_y^2}\right) + \frac{\gamma}{2\delta}\Phi_2''\left(\frac{u_{ij}^N - u_{ij}^{N-1}}{\delta}\right)\left(q_{(ij)_k^*}^N - q_{(ij)_k^*}^{N-1}\right),$$

$$J_{2(ij)^*,2k} = \left(\frac{\tau_{(i-1,j)^*}^N - 2\tau_{(ij)^*}^N + \tau_{(i+1,j)^*}^N}{h_x^2} + \frac{\tau_{(i,j-1)^*}^N - 2\tau_{(ij)^*}^N + \tau_{(i,j+1)^*}^N}{h_y^2}\right) + \frac{\gamma}{2\delta}\Phi_2''\left(\frac{u_{ij}^N - u_{ij}^{N-1}}{\delta}\right)\left(r_{(ij)_k^*}^N - r_{(ij)_k^*}^{N-1}\right);$$

solve $J\mathbf{c} = -F$ by Gaussian eliminations;
for $i, j = 1, 2, \cdots, I, J$

$\omega_{ij}^{\text{new}} = \omega_{ij} + c_{2(ij)^*-1},$
$\theta_{ij}^{\text{new}} = \theta_{ij} + c_{2(ij)^*};$

if $\max_{ij} |\omega_{ij}^{\text{new}} - \omega_{ij}| + \max_{ij} |\theta_{ij}^{\text{new}} - \theta_{ij}| <$ tol, stop;
otherwise, reset $\omega_{ij} = \omega_{ij}^{\text{new}}$ and $\theta_{ij} = \theta_{ij}^{\text{new}}$, $i, j = 1, 2, \cdots, I, J$;

As in the continuous case, we have the following convergence result for the shooting algorithm. This follows from standard convergence results for Newton's method applied to finite dimensional systems of nonlinear equations.

Remark 4.1. *The algorithms we propose are well suited for implementations on a parallel computing platform such as a massive cluster of processors. The shooting algorithms of this chapter can be regarded as a generalization of their counterpart for systems of ODE (see, e.g.,* [2]*.) There has been a substantial literature on the parallelization of shooting methods for ODEs* [3, 6, 7]*; these cited results will be helpful in parallelizing the shooting algorithms of this chapter.*

5. 2-d computational experiments for controllability of the wave equation

We will apply Algorithm 1 to the special cases of exact controllability problems with generic target functions. In other words, we will approximate the exact controllability problems for the wave equation by optimal control problems. We will test our algorithm with a smooth example (i.e., the continuous minimum L^2 norm controller f and the corresponding state u are smooth) and with a generic example. The convergence properties of the numerical shooting method in the context of the exact controllability problems are illustrated through computational experiments.

We consider the examples of the following types. $K(u) = 0$ is a form for the distributed target, $\Phi_1(u) = (u(T,\mathbf{x}) - W)^2$ and $\Phi_2(u) = (u_t(T,\mathbf{x}) - Z)^2$ are for the terminal time ones. Therefore we seek the pair (u, f) that minimizes the cost functional

$$\begin{aligned}\mathcal{J}(u,f) =& \frac{\beta}{2}\int_\Omega |u(T,\mathbf{x}) - W(\mathbf{x})|^2\,d\mathbf{x} + \frac{\gamma}{2}\int_\Omega |u_t(T,\mathbf{x}) - Z(\mathbf{x})|^2\,d\mathbf{x} \\ &+ \frac{1}{2}\int_0^T\int_\Omega |f(t,\mathbf{x})|^2\,d\mathbf{x}\,dt\end{aligned} \tag{5.27}$$

subject to the wave equation

$$\begin{cases} u_{tt} - \Delta u + \Psi(u) = f & \text{in } Q \equiv (0,T)\times\Omega\,, \\ u|_{\partial\Omega} = 0, \quad \text{in } (0,T)\,, \\ u(0,\mathbf{x}) = w(\mathbf{x}) \text{ and } u_t(0,\mathbf{x}) = z(\mathbf{x}) & \text{in } \Omega\,, \end{cases} \tag{5.28}$$

where $\Omega \subset \mathbb{R}^2$.

5.1. Examples of the linear cases. We choose $Q = [0,1]\times([0,1]\times[0,1])$ in the linear cases(i.e. $\Psi(u) = 0$) in the examples I, II, and consider two sets of $C^\infty(\overline{\Omega})$ initial data:

$$\begin{array}{lll} 1. \quad w = 1 & \text{and} & z = 0\,, \\ 2. \quad w = x(x-1)y(y-1) & \text{and} & z = 0\,, \end{array} \tag{5.29}$$

and two sets of the target functions in the linear examples I, II are

$$\begin{array}{rlll} \text{I.} & W = x(x-1)y(y-1)\cos(1) & \text{and} & Z = x(x-1)y(y-1)\sin(1)\,, \\ \text{II.} & W = 1 & \text{and} & Z = 0\,. \end{array} \tag{5.30}$$

The linear examples I and II have the initial data in (5.29) and the corresponding target functions in (5.30) respectively. Computational results carried out for $\delta = 1/8, 1/16, 1/32, 1/64$ and $h_x = h_y = \delta/2$ with an increasing sequence of $\beta = \gamma = 1, 10^2, 10^4, 10^6$. Numerical results are illustrated in figures 1,3, 4 for example I and 2, 5, 6 for example II. In order to visualize the convergence of our method as control parameters tend to infinity, we provide two kinds of plots of the approximations;one for the three dimensional graph, one for the snapshot. The three dimensional graphs are given in figures 1, 2. For the snapshots, we provide several graphs with fixed y-coordinates in order to observe the computational results easily, so called snap-shot. At $t = T$, we present three graphs with y-coordinates 0.25, 0.5, and 0.75 from left to right. It seems that our method produces (pointwise) convergent approximations for example I, II *without the need for regularization,* but the approximations are not in general convergent in a pointwise sense. Of course, approximations that do not converge in a pointwise sense may be of little practical use, even if they converge in a root mean square sense. Also, note that the results obtained by our method behave very similarly to those obtained in [8]. Finally we simply denote the approximate solutions and target functions by superscript h. For example, $u^h(T,\cdot), u^h_t(T,\cdot)$ mean the approximations of a given optimal control problem at the terminal time with $h = h_x = h_y$ and $\delta = 2h$. Similarly W^h, Z^h will stand for the approximations of the target functions at the given mesh size.

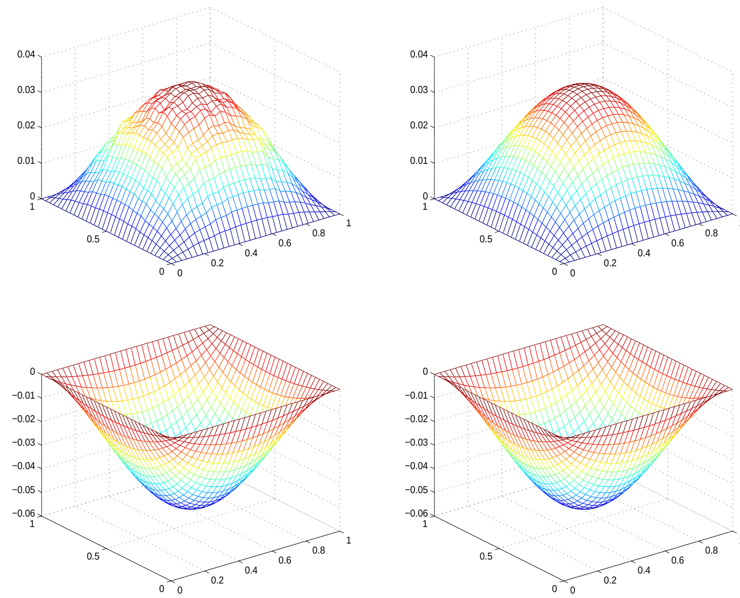

FIGURE 1. Images of approximations for linear example I with initial data 1 in (5.29);$u^h(T,\cdot), u_t^h(T,\cdot)$(left) compared with targets W^h, Z^h(right);$\delta = 1/64, h_x = h_y = \delta/2$ and $\beta = \gamma = 10^6$ ..

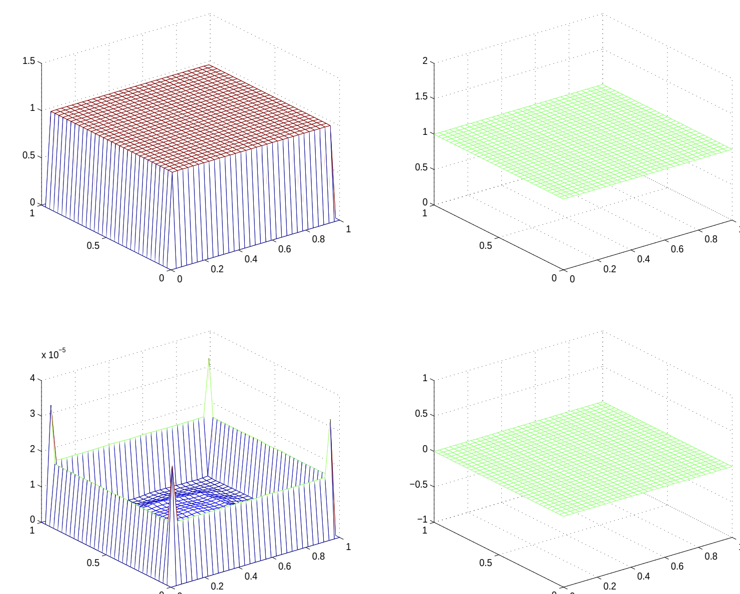

FIGURE 2. Images of approximations for linear example II with initial data 2 in (5.29);$(u^h(T,\cdot), u_t^h(T,\cdot)$(left) compared with targets W^h, Z^h(right);$\delta = 1/64, h_x = h_y = \delta/2$ and $\beta = \gamma = 10^6$ ..

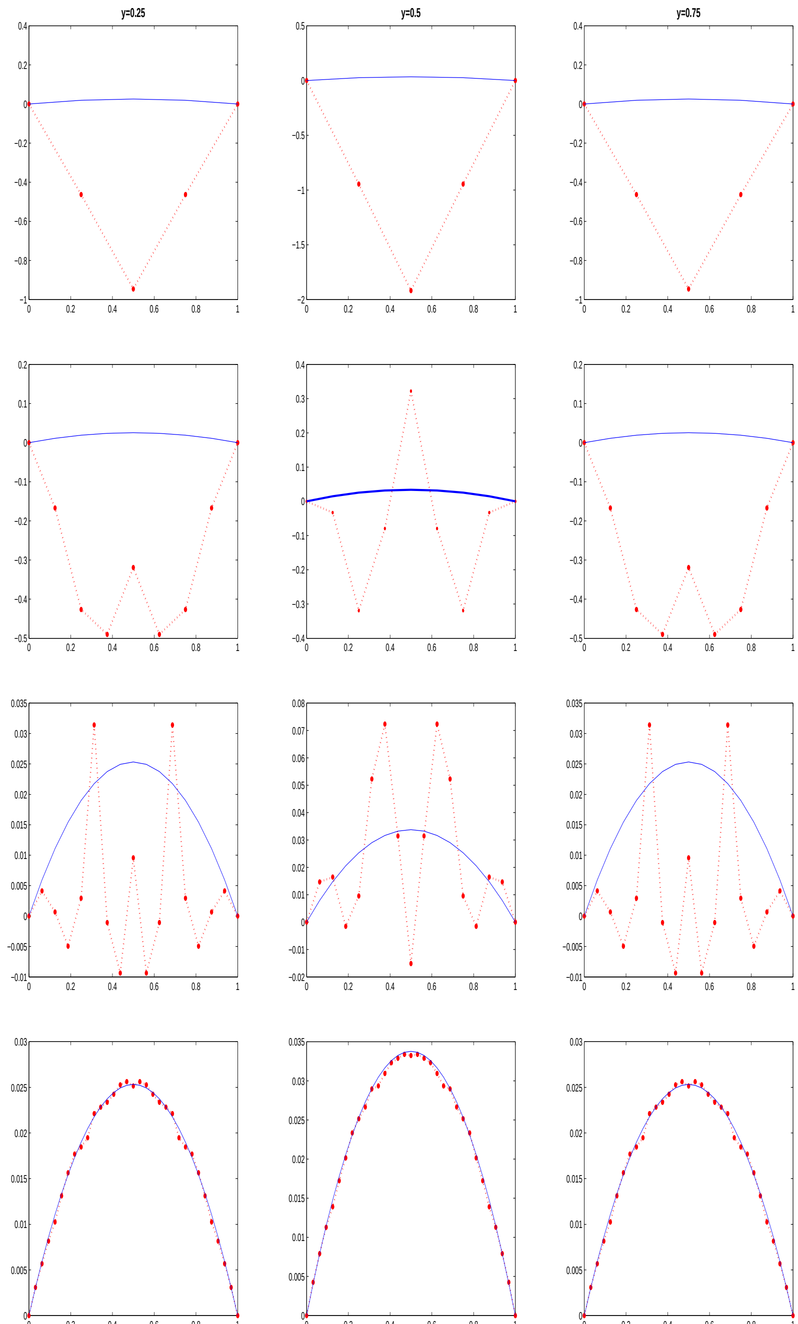

FIGURE 3. Image of $u^h(T,\cdot), W^h$ for linear example I with initial data 1 in (5.29); each row corresponds to $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$ from top to bottom respectively;$\gamma = \beta$, $h_x = h_y = \delta/2$; .: optimal solution $u^h(T,\cdot)$ $-$: target function W^h.

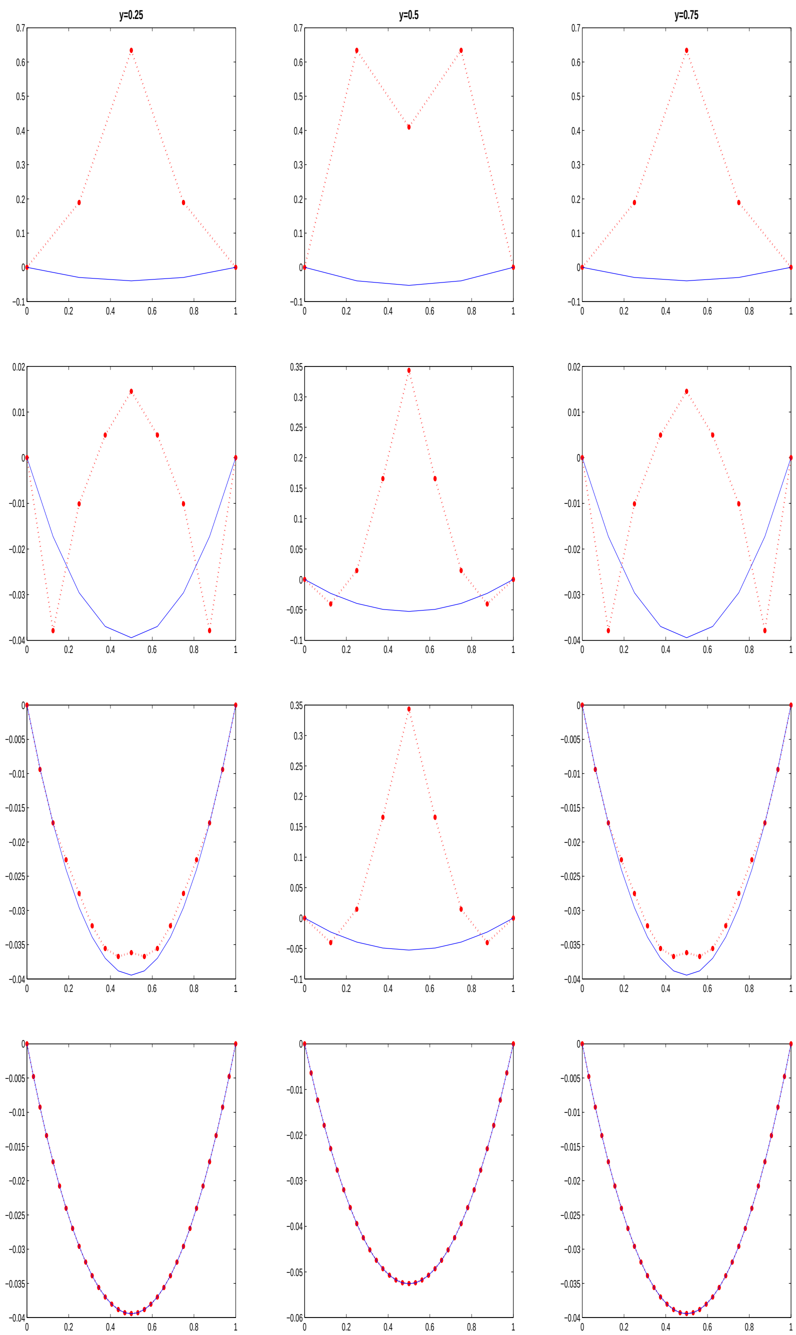

FIGURE 4. Image of $u_t^h(T,\cdot), Z^h$ for linear example I with initial data 1 in (5.29); each row corresponds to $(\delta,\beta) = (1/8, 1), (1/16, 10^2), (1/32, 10^4), (1/64, 10^6)$ from top to bottom respectively;$\gamma = \beta$, $h_x = h_y = \delta/2$; .: optimal solution $u_t^h(T,\cdot)$ $-$: target function Z^h.

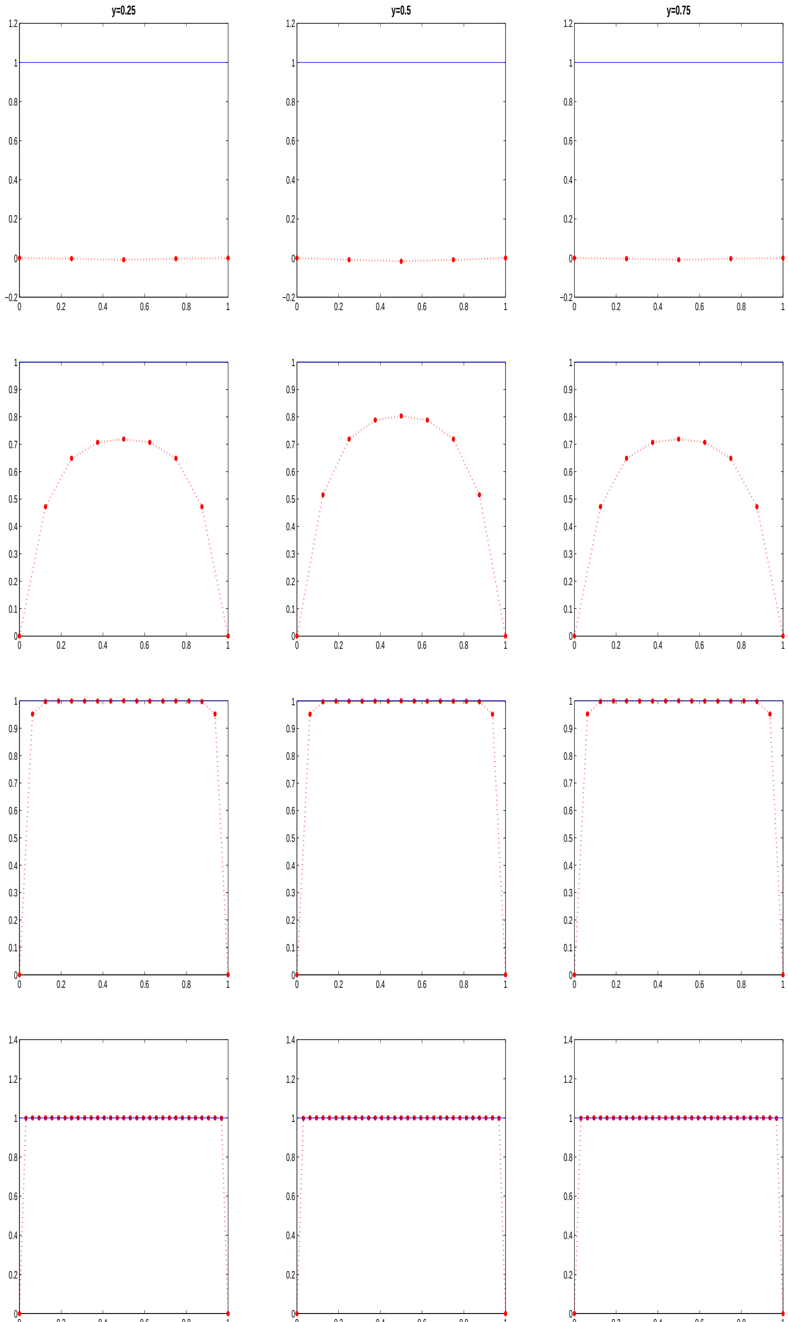

FIGURE 5. Image of $u^h(T,\cdot), W^h$ for linear example II with initial data 2 in (5.29); each row corresponds to $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$ from top to bottom respectively;$\gamma = \beta$, $h_x = h_y = \delta/2$; .: optimal solution $u^h(T,\cdot)$ $-$: target function W^h.

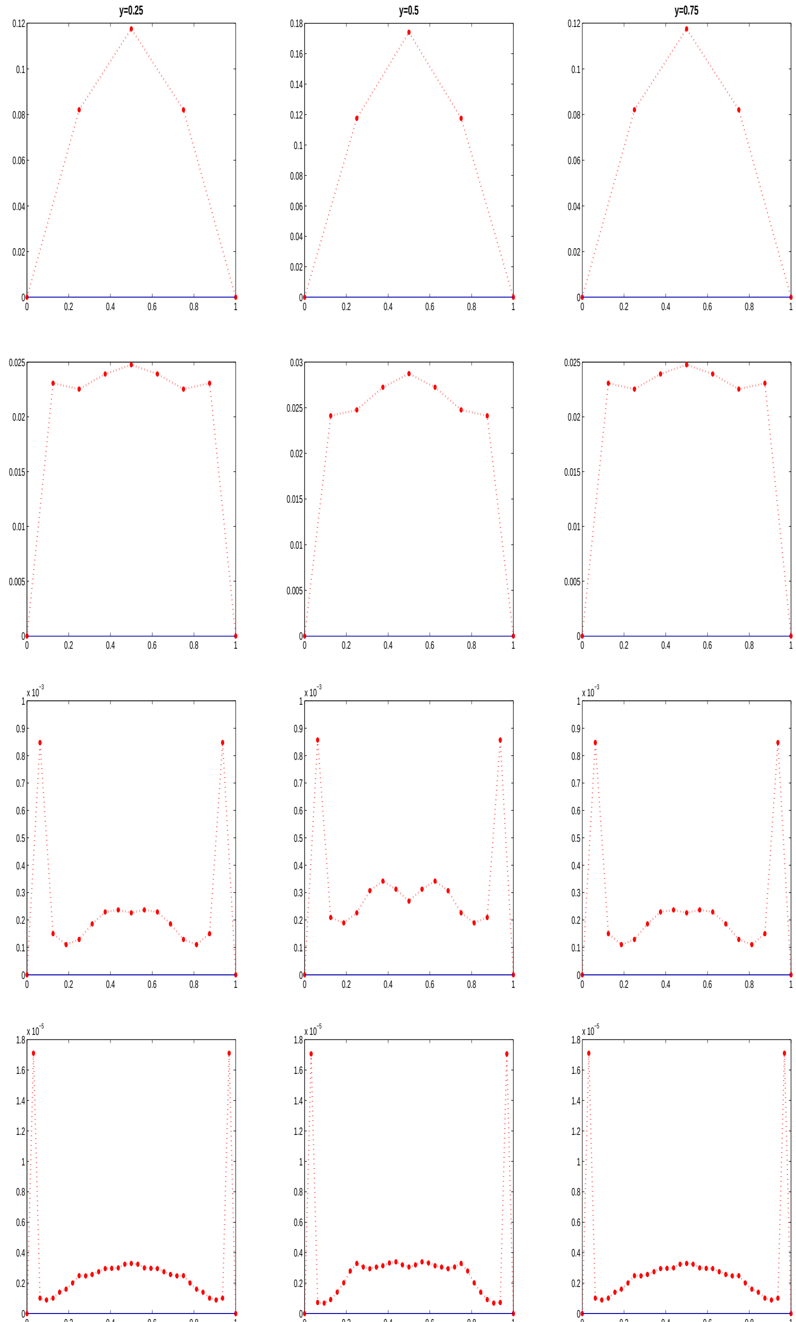

FIGURE 6. Image of $u_t^h(T,\cdot), Z^h$ for linear example II with initial data 2 in (5.29); each row corresponds to $(\delta, \beta) = (1/8, 1), (1/16, 10^2), (1/32, 10^4), (1/64, 10^6)$ from top to bottom respectively;$\gamma = \beta$, $h_x = h_y = \delta/2$; .: optimal solution $u_t^h(T,\cdot)$ $-$: target functions Z^h.

5.2. Examples of the semilinear cases. We will again apply Algorithm 1 to the special case of $W = 0$, $Z = 0$, $\alpha = 0$ and $\beta, \gamma >> 1$ in the semilinear examples I, II for the experiments of the null controllability, and to the generic case of target functions with the same values of the parameters in the semilinear example III. If the nonlinear term $\Psi(u)$ satisfies a certain property such as being asymptotically linear or superlinear, then the exact controllability problem of system (1.1) can be solvable;see, e.g., [23]. In this section, we examine the performance of our method for the asymptotically linear and superlinear cases.

We choose $Q = [0, 2] \times ([0, 1] \times [0, 1])$ in the semilinear examples I, II, III, and consider three sets of semilinear term $\Psi(u)$:

$$\begin{array}{rl} \text{I.} & \Psi(u) = u \ln^2(u^2 + 1) \\ \text{II.} & \Psi(u) = \sin u \\ \text{III.} & \Psi(u) = u^3 - u \end{array} \tag{5.31}$$

associated with two sets of $C^\infty(\overline{\Omega})$ initial data:

$$\begin{array}{llll} 1. & w = 0 & \text{and} & z = x(x-1)y(1-y)\,, \\ 2. & w = x(x-1)y(y-1) & \text{and} & z = 0\,. \end{array} \tag{5.32}$$

The target functions in the semilinear examples I, II are

$$W = 0, \quad Z = 0\,,$$

and in the semilinear example III are

$$W = x(x-1)y(y-1)\cos(1), \quad Z = x(x-1)y(y-1)\sin(1)\,.$$

Note that we choose $T = 2$ for existence of control in the problems; see e.g., [14, 15, 20, 21, 22, 23, 24]. As the linear examples, computational results carried out for $\delta = 1/8, 1/16, 1/32, 1/64$ and $h_x = h_y = \delta/2$ with an increasing sequence of $\beta = \gamma = 1, 10^2, 10^4, 10^6$ for the semilinear cases I, II, III. Numerical results are illustrated in figures 7,8 for example I, in figures 9, 10 for example II, and in figures 11, 12 for example III. Figures 7- 12 show that the computational results for semilinear cases in term of the terminal matchings W, Z at $t = T$ when control parameters β and γ are 10^6 in the full distributed controllability problems provide the good approximations in each case. This is the similar behavior as the linear case, i.e. the shooting method produces convergent (in $L^2(Q)$) approximations for semilinear example I, II and III *without the need for regularization* but the approximations are not in general convergent in a pointwise sense. Of course, approximations that do not converge in a pointwise sense may be of little practical use, even if they converge in a root mean square sense.

Remark 5.1. *Note that for the linear cases, the number of iterations of the shooting methods is just one since the Newton method finds the exact solution for one iteration, but in practice it is about 2 or 3, according to the tolerance and the accuracy of the machines we used. The semilinear cases are different and we need more iterations than the linear cases. However the number of iterations for semilinear cases is not more than 4 iterations in the experiments when we apply the shooting algorithm to the given examples.*

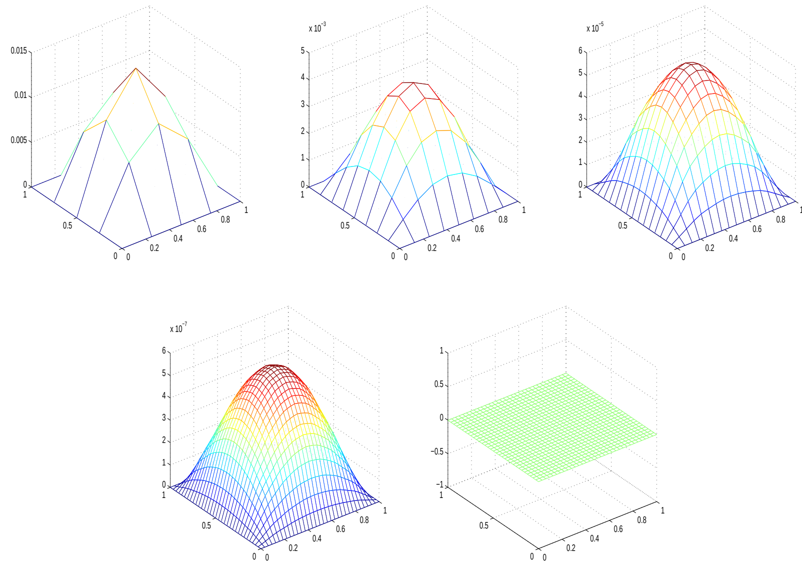

FIGURE 7. Images of sequence of $u^h(T,\cdot)$ and target W^h for semilinear example I with initial data 1 in (5.31) with $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$;$\gamma = \beta$, $h_x = h_y = \delta/2$; $\Psi(u) = u\ln^2(u^2+1)$.

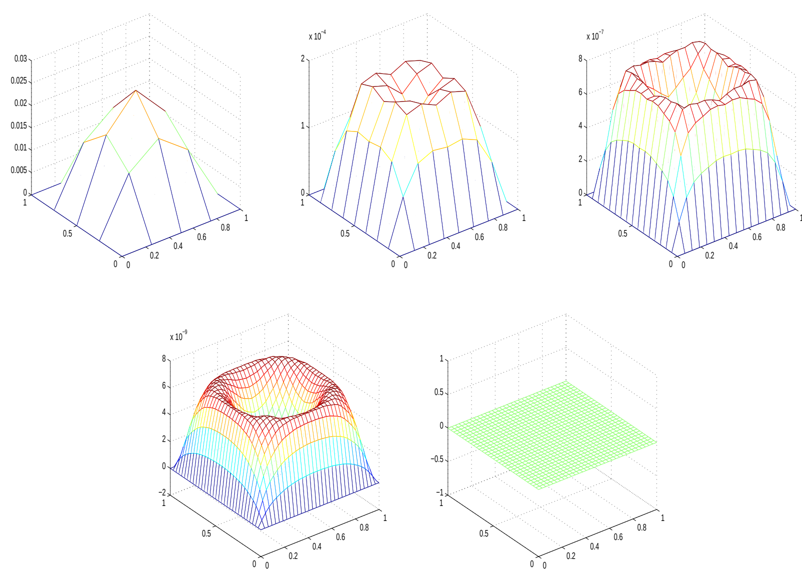

FIGURE 8. Images of sequence of $u_t^h(T,\cdot)$ and target Z^h for semilinear example I with initial data 1 in (5.31) with $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$;$\gamma = \beta$, $h_x = h_y = \delta/2$; $\Psi(u) = u\ln^2(u^2+1)$.

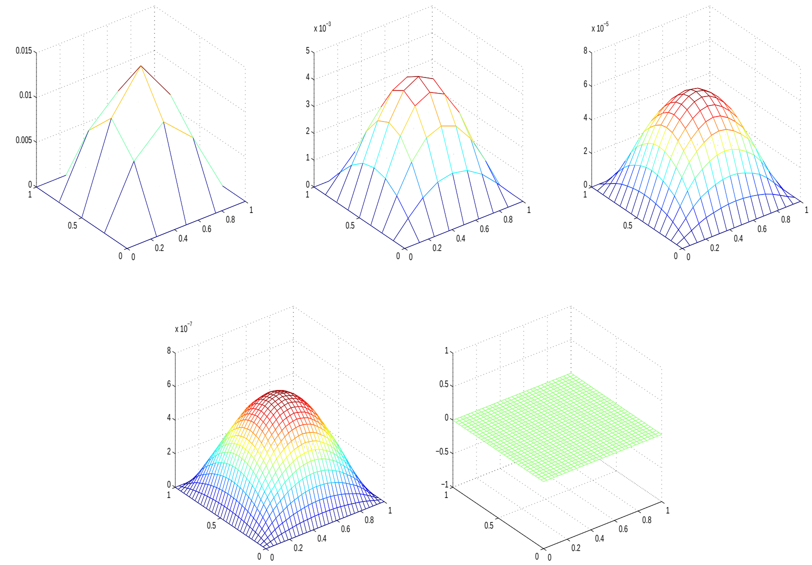

FIGURE 9. Images of sequence of $u^h(T,\cdot)$ and target W^h for semilinear example II with initial data 1 in (5.31) with $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$;$\gamma=\beta,\ h_x=h_y=\delta/2$; $\Psi(u)=\sin u$.

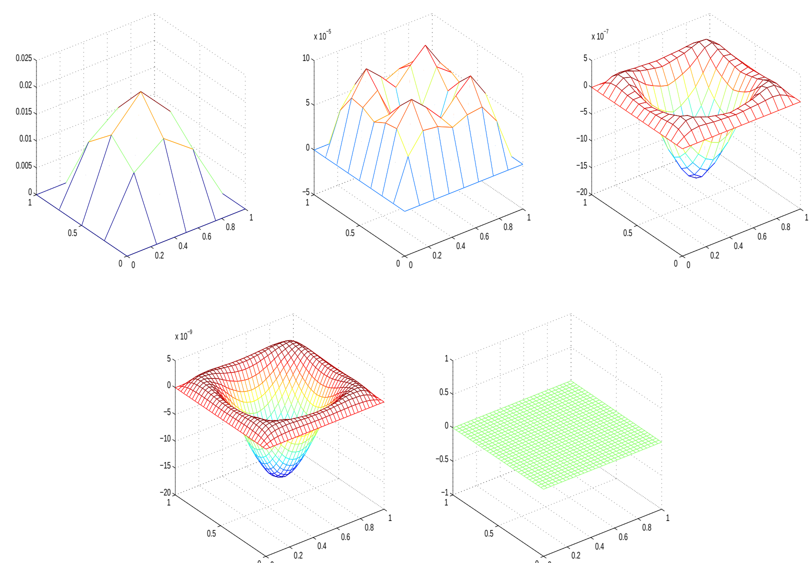

FIGURE 10. Images of sequence of $u_t^h(T,\cdot)$ and target Z^h for semilinear example II with initial data 1 in (5.31) with $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$;$\gamma=\beta,\ h_x=h_y=\delta/2$; $\Psi(u)=\sin u$.

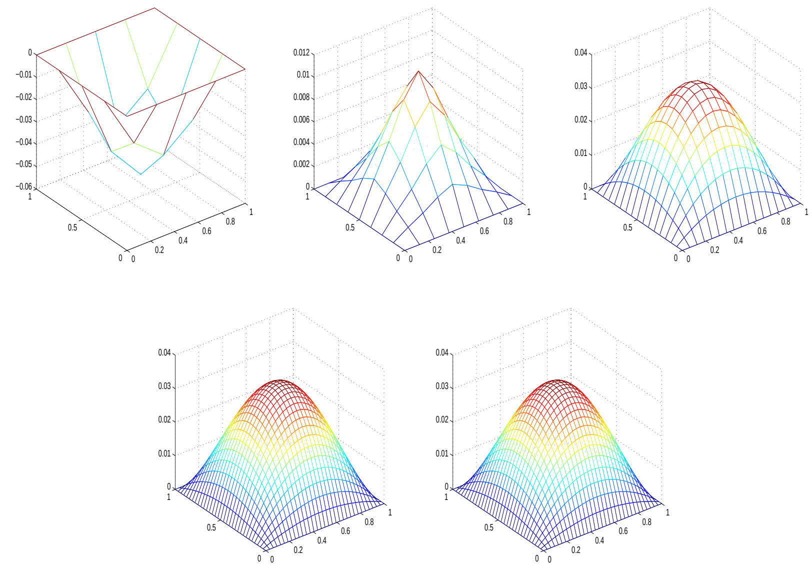

FIGURE 11. Images of sequence of $u^h(T,\cdot)$ and target W^h for semilinear example III with initial data 2 in (5.31) with $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$;$\gamma=\beta$, $h_x = h_y = \delta/2$; $\Psi(u) = u^3 - u$.

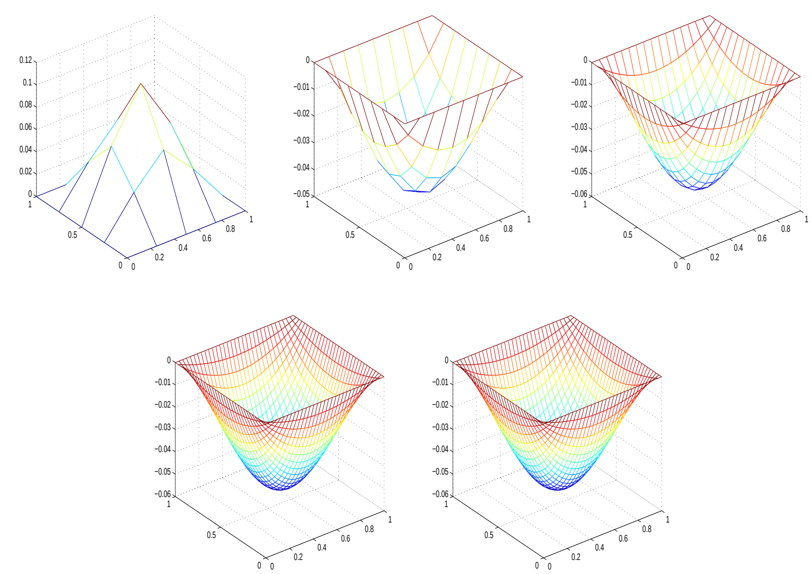

FIGURE 12. Images of sequence of $u_t^h(T,\cdot)$ and target Z^h for semilinear example III with initial data 2 in (5.31) with $(\delta,\beta) = (1/8,1),(1/16,10^2),(1/32,10^4),(1/64,10^6)$;$\gamma=\beta$, $h_x = h_y = \delta/2$; $\Psi(u) = u^3 - u$.

6. Conclusion and remarks

In the paper, we discussed and successfully implemented shooting methods for solving optimal control problems constrained by linear wave equations, semilinear wave equations. The shooting algorithms for optimal control problems were also utilized effectively to find approximate solutions to the exact controllability problems for these equations. Distributed controls in two dimensional case were treated. The boundary control cases will be discussed in a separate paper. The convergence of the algorithms were numerically demonstrated when the smooth target functions are given.

However, a host of issues regarding this algorithm still need be addressed in future work; these include other control objectives, a thorough study of parallel implementations and a analysis of computing complexity, rigorous numerical analysis. Moreover, we may consider generalizations to control other types of equations. For example, we consider *equations of linear and nonlinear elasticity.* Wave equations of the form (1.3) are special cases of PDE systems modeling elastic materials and structures. It is of significant practical interest to study optimal control problems for elasticity. We will attempt to extend the results of the tasks we applied into numerical solutions of control problems for elasticity. We are confident about the successes of research on such problems in one or two space dimensions. We are currently working on these issues and will present the results in separate papers.

Acknowledgments

The author gratefully acknowledges the support of School of Computational Science during the development of the research described in this paper.

References

[1] R. Adams, Sobolev Spaces, Academic, New York, 1975.
[2] U. Ascher, R.M.M. Mattheij and R.D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, Prentice-Hall, 1988.
[3] W.H. Enright and K. Chow, Distributed parallel shooting for BVODEs, Proceedings of High Performance Computing, 1998.
[4] L. Evans, Partial Differential Equations, Amer. Math. Soc., Providence, RI, 1998.
[5] O. Yu. Emanuilov, Boundary controllability of semilinear evolution equations, Russian Math. Surveys, **44**, no. 3, 1989, pp. 183-184.
[6] K. Jackson, A Survey of Parallel Numerical Methods for Initial Value Problems for Ordinary Differential Equations, IEEE Transactions on Magnetics, **27**, 1991, pp. 3792-3797.
[7] K. Jackson and S. Norsett, The potential for parallelism in Runge-Kutta methods. Part 1: RK formulas in standard form, SIAM J. Numer. Anal., **32**, 1995, pp. 49-82.
[8] L.-L. Ju, M. Gunzburger and L.S. Hou, Approximation of exact boundary controllability problems for the 1-D wave equation by optimization-based methods, Contemp. Math., **330**, 2003, pp. 133-153.
[9] R. Glowinski, C.-H. Li, and J.-L. Lions, A numerical approach to the exact boundary controllability of the wave equation (I). Dirichlet controls: description of the numerical methods, Japan J. Appl. Math., **7** 1990, pp. 1-76.
[10] M. Gunzburger and R. Nicolaides, An algorithm for the boundary control of the wave equation, Appl. Math. Lett., **2** 1989, pp. 225-228.
[11] J. A. Infante and E. Zuazua, Boundary observability for the space semi-discretizations of the 1-D wave equation, Mathematical Modelling abd Numerical Analysis, Vol. 33, No. 2, 1999, pp. 407-438.
[12] J.-L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, New York, 1971.
[13] J.-L. Lions, Control of Distributed Singular Systems, Bordas, Paris, 1985.
[14] J.-L. Lions, Exact controllability and stabilization and perturbations for distributed systems, SIAM Review, **30** 1988, pp. 1-68.

[15] J.-L. Lions, Controlabilité Exacte, Perturbation et Stabilisation des Systèmes Destribués, Vol. 1 and 2, Masson, Paris, 1988.

[16] M. Negreanu and E. Zuazua, Uniform boundary controllability of a discrete 1-D wave equation, preprint, 2002.

[17] D. Russell, Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions, SIAM Review, **20** 1978, pp. 639-739.

[18] D. Russell, A unified boundary controllability theory, Studies in Appl. Math., **52** 1973, pp. 189-211.

[19] Yang. Sung-Dae, Shooting Methods for Numerical Solutions of Exact Controllability Problems Constrained by Linear and Semilinear Wave Equations with Local Distributed Controls, Applied Mathematics and Computation, in press.

[20] E. Zuazua, Exact controllability for the semilinear wave equation, J. Math. Pures Appl., **69**, 1990, pp. 1-31.

[21] E. Zuazua, Exact controllability for semilinear wave equations in one space dimension, Ann. Inst. H. Poincar Anal. Non Linaire., **10**, 1993, pp. 109-129.

[22] E. Zuazua, Boundary observability for the finite difference space semi-discretizations of the 2-D wave equation in the square, J. Math. Pures Appl., **78**, 1999, pp. 523-563.

[23] E. Zuazua, Optimal and Approximate Control of Finite-Difference Approximation Schemes for the 1-D Wave Equation, Rendiconti di Matematica, Serie VIII, Vol. **24**, Tomo II, 2004, pp. 201-237.

[24] E. Zuazua, Propagation, Observation, Control and Numerical Approximation of Waves approximated by finite difference methods, SIAM Review , **47 (2)**, 2005, pp. 197-243.

School of Computational Science, Florida State University, Tallahassee, FL 32306, USA
E-mail: `sdyang@scs.fsu.edu`
URL: `http://www.scs.fsu.edu/~sdyang/`