# OPTIMAL CONTROL OF DISTRIBUTED PROCESSES USING REDUCED ORDER MODELS

**Eva Balsa-Canto*, Antonio A. Alonso† and Julio R. Banga***

* Chemical Engineering Lab. Instituto de Investigaciones Marinas (CSIC), Eduardo Cabello, 6, 36308-Vigo, Spain.
e-mail: eva@iim.csic.es, julio@iim.csic.es

† Department of Chemical Engineering, Universidad de Vigo, 36200-Vigo, Spain.
e-mail:aalvarez@uvigo.es

**Keywords:** Optimal control, distributed parameter processes, proper orthogonal decomposition, reduced order models.

## Abstract

The open loop optimal control (dynamic optimization) of distributed parameter systems is considered here. These problems are usually solved by the Control Vector Parameterization (CVP) approach, which transforms the original dynamic optimization method into an outer nonlinear programming problem, which requires the solution of an inner initial value problem (IVP). The solution of this IVP (set of partial and ordinary differential equations) for each function evaluation is usually very demanding in terms of computation time, thus efficient numerical techniques are necessary in order to reduce the overall computational effort of the CVP optimizations.

In this work, the use of low order models, obtained by Galërking projection on a set of proper orthogonal functions, is presented as a very efficient alternative for the rapid solution of this class of problems.

## 1 Introduction

Dynamic optimization (also called open loop optimal control) involves the calculation of time varying control profiles that optimize (minimize or maximize) a desired objective functional subject to the system dynamics. The increasing detail of process models leads to consider the spatial variation of properties within the process. Typical examples include polymerisation reactors, crystallisation units and thermal processing of bioproducts, among others. The resulting mathematical models consist of sets of partial differential equations (PDEs) and/or ordinary differential and algebraic equations (DAEs), expressing the physical laws of conservation of mass, energy and momentum.

Although several strategies have been presented in the literature for the solution of optimal control problems (see review in [Vas93]) we will use here the control vector parameterization approach (CVP), which proceeds dividing the duration of the process into a number of control intervals and approximating the control function using a low order polynomial. The decision variables in the optimization process are the polynomial coefficients. The solution of the master NLP involves the solution of the initial value problem (IVP) (system of PDAES, in our case), for each function evaluation. Therefore, it is of the highest interest to derive methods which permit the solution of the IVP of complex systems in short computation times.

In general, the numerical solution of partial differential equations involves two steps: first, the original infinite dimension problem (PDE) is transformed into a finite dimension problem, i.e. a set of DAEs, and second, the resultant system is then solved by a standard initial value problem (IVP) solver.

Different techniques can be applied to transform the original PDEs into DAEs (step 1)(see reviews in [Sch94] and [Oh95]). However, some of these techniques result in large scale DAE systems which are not always convenient for dynamic optimization, as the demand for computational resources increases rapidly with the number of equations (discretization level) and the stiffnes of the DAE system.

Alternative approaches consist of making use of reduced order models which are typically much less expensive to compute. Reduced order models have been widely applied in the process engineering area. Most of the studies focus on identification, observer construction and the development of robust control schemes ([ABS00], [Bur93], [SK98]), estimation of parameters ([PKC98]) and stationary optimization ([TAZ98], [BC99]).

In this work, the use of low order models, obtained by Galërkin projection on a set of proper orthogonal functions, is presented as a very efficient alternative for the rapid solution of the dynamic optimization of distributed parameter processes. A set of examples is considered to show the efficiency of this new approach.

## 2 Statement of optimal control problem

The general dynamic optimization (open loop optimal control) problem considered here can be stated as follows:

Find $\mathbf{c(t)}$ to minimize:

$$J[\mathbf{u}, \mathbf{y}, \mathbf{c}, \mathbf{v}, \xi, \mathbf{t}] \tag{1}$$

Subject to:

$$\Psi(\mathbf{u}, \mathbf{u_\xi}, \mathbf{u_{\xi\xi}}, \mathbf{u_t}, \mathbf{y}, \dot{\mathbf{y}}, ..., \mathbf{c}, \mathbf{v}, \xi, \mathbf{t}) = \mathbf{0} \tag{2}$$

with $\xi \in \mathbf{\Omega} \subset \mathbf{R^3}$ the space variables, $\mathbf{u}(\xi, \mathbf{t}) \in U \subset R^\nu$ the state variables depending on the position and time, $\mathbf{y(t)} \in Y \subset R^\mu$ the subset of state variables which are position invariant, $\mathbf{u_\xi} = \partial\mathbf{u}/\partial\xi$, $\mathbf{u_{\xi\xi}} = \partial^\mathbf{2}\mathbf{u}/\partial\xi^\mathbf{2}$, $\mathbf{u_t} = \partial\mathbf{u}/\partial\mathbf{t}$, $\dot{\mathbf{y}} = d\mathbf{y}/\mathbf{dt}$, $\mathbf{c} \in \mathbf{C} \subset \mathbf{R}^\sigma$ the control variables and $\mathbf{v} \in \mathbf{V} \subset \mathbf{R}^\rho$ the time invariant parameters. Appropriate initial and boundary conditions must be imposed to determine exactly the mathematical solution of Eqn. 2.

Path and point constraints may be imposed to the system to express conditions related to safe or proper operation. These path constraints must be satisfied over the entire domain of interest, and are stated as follows:

$$\mathbf{h}(\mathbf{u}(\xi, \mathbf{t}), \mathbf{y(t)}, \mathbf{c(t)}, \mathbf{v}, \xi, \mathbf{t}) = \mathbf{0} \tag{3}$$

$$\mathbf{g}(\mathbf{u}(\xi, \mathbf{t}), \mathbf{y(t)}, \mathbf{c(t)}, \mathbf{v}, \xi, \mathbf{t}) \leq \mathbf{0} \tag{4}$$

As it has been mentioned, this problem can be solved using the CVP approach.

## 3 Initial value problem solution

As it was mentioned in the introduction, numerical solution of partial differential equations involves the transformation of the original PDEs into a set of DAEs, and the solution of the latest by a standard initial value problem (IVP) solver.

The main idea is to project the infinite dimension solution into a number of spatial functions locally or globally defined. When the domain of the basis functions is localized in space, the PDE is transformed into a large scale, usually stiff, DAE system. Examples of this type of strategies are, the numerical method of lines (NMOL) ([Sch91]), the finite difference, finite elements and finite volume approaches. This type of approaches has been followed for the dynamic optimization by several authors ([BAMS94], [PRG$^+$96], [BS97]). However, computation times increase rapidly with the size and the stiffness of the DAE system.

In contrast with these methods which make use of locally defined functions, alternative approaches exploit global, orthogonal trial functions. The so called spectral methods involve representing the solution to the problem as a truncated series of known, smooth functions of independent variables. These functions may correspond to the eigenfunctions of the spatial operator (see for example, [Smo83]), Legendre Polynomials, Chevishev polynomials (see for example, [ST95]) or more general sets such as those employed in the so-called proper orthogonal decomposition techniques ([BHL93], [Sir87]). These type of methods result in low order models thus in important reductions on computation time in the overall optimization process.

In this work, we propose the use of the proper orthogonal decomposition strategy to compute a set of empirical eigenfunctions. These empirical eigenfunctions are subsequently used as basis functions in a Galërkin method, resulting in a low order system of ODEs which will allow the efficient dynamic optimization of the distributed process.

### 3.1 The Galërkin method

This method is based on the separation of variables approach and attempts to find an approximate solution in the form of a truncated series expansion given by:

$$\tilde{u}(\xi, t) = \sum_{i=1}^{N} a_i(t)\phi_i(\xi) \tag{5}$$

where $\phi_i(\xi)$ are known trial functions. The trial functions form an orthonormal basis for the appropriate function space and are chosen to satisfy the original PDEs and the boundary conditions. The $a_i(t)$ are the time dependent coefficients selected to assure that the original PDE is satisfied as closely as possible. This is achieved by minimizing the error produced by using the expression (5) instead of the exact solution, i.e. the residual must be orthogonal to the trial functions in the expansion allowing the calculation of the time dependent coefficients $a_i(t)$.

Mathematically, suppose that the original system in equation (2) can be written as follows:

$$u_t = \Lambda(u) \tag{6}$$

where $\Lambda(.)$ is the spatial operator and appropriate boundary and initial conditions were included. The residual can be then defined as:

$$r(\xi, t) = \tilde{u}_t - \Lambda(\tilde{u}) \tag{7}$$

Forcing the residual to be orthogonal to a number of trial functions:

$$< r(\xi, t), \phi_i(\xi) >= 0 \quad i = 1, ..., N \tag{8}$$

where the following notation has been introduced:

Inner product: $< f, g >= \int_\Omega fg d\Omega$

Substituting the equation (5) into equation (7), applying the orthogonality condition (8) and using the orthonormality property of the trial functions we obtain:

$$\dot{a}_i(t) = \int_\Omega \Lambda(\sum_{i=1}^{N} a_i(t)\phi_i(\xi))\phi_i(\xi) d\Omega \tag{9}$$

Therefore the reduced order model consists of a set of N ordinary differential equations, with the initial conditions given by:

$$a_i(0) = \int_\Omega u(\xi,0)\phi_i(\xi)d\Omega \quad i = 1,...,N \tag{10}$$

### 3.2 Optimal basis set computation

As it was mentioned before, standard basis usually consist of sets of polynomials or transcendental functions. However we consider here the use of the proper orthogonal decomposition (POD) method, also known as Karhunen-Loève expansion, which yields a basis with two desirable properties:

1. The basis elements are ordered from the one with the highest average energy to that with the lowest.
2. This basis is optimal in the sense that the first terms contain, in average, more energy than the same number of terms from any other basis ([HLB96]).

Only the algebraic description of the POD method is presented here, more details can be found in [BHL93] and [Sir87].

The first step in the model reduction method, consists of simulating the high order model, or developing a set of experiments, under a sufficiently diverse set of operating conditions to collect a set of data, $S_d = \{u_i\}_{i=1}^{l}$, which describes the behaviour of the process. An optimal set of globally defined trial functions is calculated from these data through the use of the Karhunen-Loève expansion. The low k-dimensional set $S_{op}^k = \phi_j{}_{j=1}^{k}$, with $k \leq n$, is defined as the set of orthonormal vectors which maximize the average projection of the data set $S_d$.

$$\min_{\phi_j} J(\phi_j, u) \tag{11}$$

with:

$$J(\phi_j,u) = \frac{1}{\ell}\sum_{j=1}^{\ell}(u_j^T - w_j^T)(u_j - w_j) - \sum_{j=1}^{k}\lambda_j(\phi_j^T\phi_j - 1) \tag{12}$$

where $w_j$ represents the projection vector associated to the data $u_j$, being of the form:

$$w_j = \Phi c_j \tag{13}$$

with $\Phi$ a matrix:

$$\Phi = \left[\begin{array}{ccc} \phi_1 & \vdots & \phi_k \end{array}\right] \tag{14}$$

and $c_j \in R^k$ being a coefficient vector satisfying $c_j = \Phi^T u_j$. The solution of this optimization problem leads to the following set of equations to be satisfied for each element in $S_{op}^k$:

$$R\phi_j = \lambda_j\phi_j \tag{15}$$

$$R = \frac{1}{\ell}\sum_{j=1}^{\ell} u_j u_j^T \tag{16}$$

In this way, any element $u_j$ of the data set $S_d$ can be expressed as:

$$u_j = \Phi c_j + \varepsilon \tag{17}$$

where $\varepsilon$ is an error vector, orthogonal to $S_{op}^k$, that indicates the distance at which each data point lies from the low dimensional projection plane. The average distance of the data set to $S_{op}^k$ is computed as:

$$D_{av}^2 = \frac{1}{\ell}\sum_{i=1}^{\ell} d_i^2 = \frac{1}{\ell}\sum_{j=1}^{\ell} u_j^T u_j - \sum_{i=1}^{k}\lambda_i \tag{18}$$

Note that $k = l$ implies $D_{av} = 0$, i.e. $S_{op}^\ell$ and $S_d$ coincides, and therefore:

$$\sum_{i=1}^{\ell}\lambda_i = \frac{1}{\ell}\sum_{j=1}^{\ell} u_j^T u_j \tag{19}$$

so that the eigenvalues provide a measure of how close the data are to the optimal set, or equivalently, of how much *energy* (understood as the inner product of $u_i$) is captured by the low dimensional representation.

## 4 Case Studies

In this section the solution of two cases studies is considered. To begin with, a very simple, one-dimensional case regarding the uniform heating of a slab, is considered. Next, a quite complex, three-dimensional case, regarding the sterilization of a pre-packaged food is also considered.

### 4.1 Temperature control of a slab.

We first consider the simple problem of adjusting the temperature distribution of a slab of length L=1, to a desired temperature $T_{set} = 0.2$ using the external temperature ($T_{out}$) as the control variable. The system is described in terms of the heat conduction equation subject to boundary conditions expressing the flux of heat to the slab. The statement is [Ram94]:

Find $T_{out}(t)$ to minimize:

$$J = \int_0^L \left(T_{set} - T(x,t_f)\right)^2 dx \tag{20}$$

subject to:

$$T_t = T_{xx} \tag{21}$$

with the initial and boundary conditions:

$$\begin{array}{c} T(x,0) = 0 \\ T_x(0,t) = K(T - T_{out}), T_x(L,t) = 0 \end{array} \tag{22}$$

Regarding to the reduced order model calculation, a set of data has been obtained through the simulation of the

process using the NMOL. The *full* model obtained with NMOL consists of 11 ODEs. The *reduced* order model was obtained following the steps described in section 3.2, resulting in a 3-ODE model, containing the 99.99% of the energy.

The dynamic optimization problem was solved with ICRS/DS, ([BC87] and [BS96]). Both models, *full* and *reduced*, were used for the optimization for the sake of comparison and the control variable was approximated using a piecewise linear approximation with 6 variable length elements. Results obtained are reported in Table 4.1, where $J$ is the performance index value, $T_{max}$ and $T_{min}$ are the maximum and minimum temperature values at $t_f$, and $T_{CPU}$ is the computation time, in seconds, on a PC Pentium III/450 MHz.

**Table 4.1:** Summary of results for the case study 1.

| *Model* | $J$ | $T_{max}$ | $T_{min}$ | $T_{CPU}$ |
|---|---|---|---|---|
| *Reduced* | $2.432 \times 10^{-6}$ | 0.2040 | 0.1978 | 0.5 |
| *Full* | $4.2971 \times 10^{-6}$ | 0.2027 | 0.1969 | 2.3 |

Note that using the *reduced* model a better result was obtained more than 4 times faster than with the *full* model.

## 4.2 Industrial sterilization of canned foods.

This problem has been studied by many researches (see recent review by [Dur97]). In [BMGC91] the formulation and solution of several optimal control problems using different types of performance indexes and constraints was considered. Here we consider the statement of the optimal control problem of maximizing the final retention of a nutrient, $ret_N\,(t_f)$, with constraints on the final microbiological lethality, $F_c(t_f)$, and the final temperature in the hottest point, using the retort temperature, $T_{retort}(t)$, as the control variable as considered by [BMGC91]. We will study the particular case of a cylindrical container of volume $V_T$ (radius $R$ and height $2L$) filled with a conduction heated food. Due to symmetry considerations it is sufficient to model the heat transfer and kinetic phenomena in the half plane of the cylinder.

The mathematical statement of the problem is as follows:
Find $T_{retort}(t)$ to maximize:

$$J = \frac{1}{V_T}\int_0^{V_T} \exp\left(\frac{-\ln 10}{D_N}\int_0^{t_f}\exp\left(\frac{T(r,z,t)-T_N}{Z_N}\ln 10\right)dt\right)dV \tag{23}$$

subject to:

$$\frac{\partial T}{\partial t} = \alpha\left(\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}+\frac{\partial^2 T}{\partial z^2}\right) \tag{24}$$

with the boundary and initial conditions:

$$T(R,z,t) = T_{retort}(t) \quad T(r,L,t) = T_{retort}(t) \tag{25}$$

$$\frac{\partial T}{\partial r}(0,z,t) = 0 \qquad \tfrac{\partial T}{\partial z}(r,0,t) = 0 \tag{26}$$

$$T(r,z,0) = T_0 \tag{27}$$

Two final time constraints are considered, one for the hottest point inside the container:

$$T_c(r,z,t=t_f) \le T_0 \qquad \forall r,z \in V_T \tag{28}$$

and a second constraint on the lethality associated with the point with minimum lethality (critical point):

$$F_c(t=t_f) \ge F_{c,D} \tag{29}$$

where $F_{c,D} = 876\ s$ in this case, and $F_c$ is the microbial lethality at the final time:

$$F_c(t=t_f) = \int_0^{t_f} 10^{\frac{T_c(t)-T_M}{Z_M}}dt \tag{30}$$

As we consider a conduction-heated product, the governing equation is Fourier's second law, equation (24), with initial condition given by equation (27), and boundary conditions represented by equations (26). These boundary conditions can be easily changed to take into account arbitrary heat transfer coefficients. The thermal destruction of microorganism spores and the thermal degradation of nutrients are assumed to follow pseudo first-order reaction kinetics:

$$\frac{dC_M(t)}{dt} = -(\frac{\ln 10}{D_{M,ref}})C_M(t)\exp(\frac{T(r,z,t)-T_{M,ref}}{z_{M,ref}}) \tag{31}$$

$$\frac{dC_N(t)}{dt} = -(\frac{\ln 10}{D_{N,ref}})C_N(t)\exp(\frac{T(r,z,t)-T_{N,ref}}{z_{N,ref}}) \tag{32}$$

which are described by two kinetic parameters, $D$ and $Z$. Subscript $M$ refers to microorganism spores, while $N$ is used for the nutrients.

Regarding the reduced order model calculation, a set of data was obtained through the simulation of the nominal process model using the numerical method of lines. The reduced order model was obtained following the steps described in section 3.2, resulting in a 5-ODE model which describes the temperature distribution inside the load. The empirical eigenfunctions are shown in Figure 1. In order to test the reduced model a randomly generated control profile was used for the simulation of both, full and reduced, models, and a comparison is presented in Figure 2, where the time evolution of the temperature in a number of points inside the load has been plotted, showing very good agreement.

The computation of the performance index and the lethality at the critical point were incorporated to the reduced model as two additional ODEs.

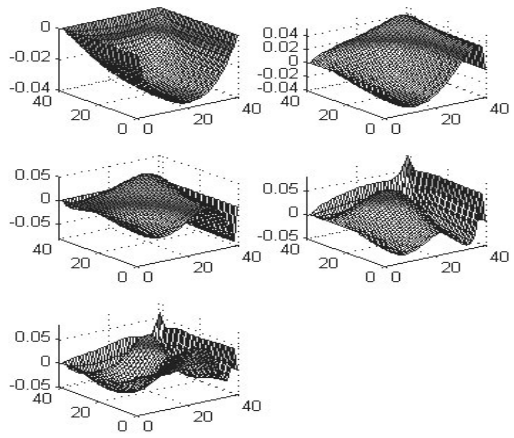

Figure 1: Empirical eigenfunctions $\phi_1, \ldots, \phi_5$.

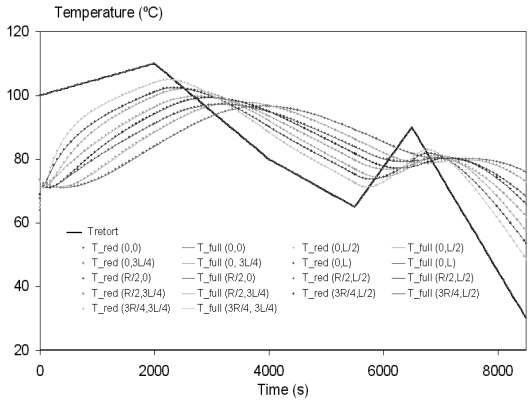


Figure 2: Comparison of the full and the reduced order models.

The optimal control problem was solved using ICRS/DS with the control profile approximated by a piecewise linear approximation, considering both the *reduced* and the *full* model, obtained with the NMOL (consisting of 243 equations).

The optimization of both models resulted in a best objective of $J = 0.47$ (47% of nutrient retention), within the specified constraints, which is in agreement with [BMGC91]. The optimal control profile and selected state variables are shown in Figure 3. The computation time required for the optimization of the full model was 5 *min* whereas for the reduced order model it only required 14.6 *s*. Thus speedups of around 20 or even more can be achieved with the use of reduced models for this type of problems.

Remark also that an objective value within the 98% of the best obtained may be achieved in less than 5 *s* with the reduced model, which makes this technique very promising for its use in a model predictive control framework.

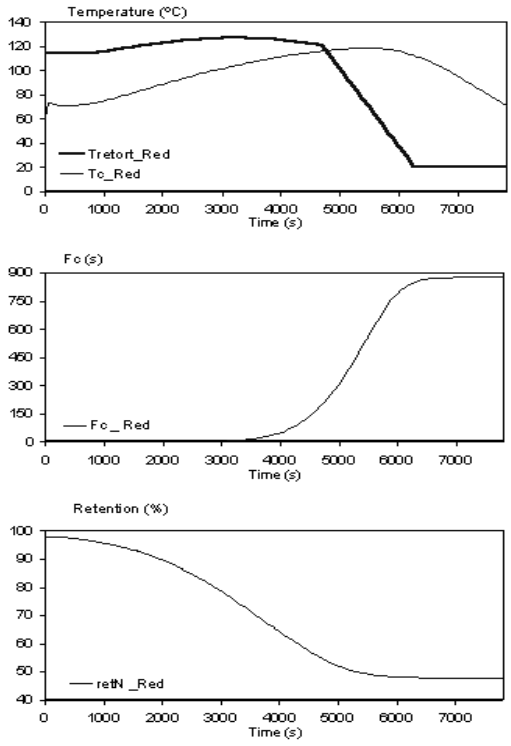


Figure 3: Optimal control profile and selected state variables.

# 5 Conclusions

The use of the proper orthogonal decomposition technique was successfully applied to obtain low order models, with very good predictive characteristics, which allow the rapid solution of the dynamic optimization of distributed processes.

*Full* order models, obtained with the NMOL technique, were compared with *reduced* ones for the solution of two dynamic optimization case studies. As a result, very significant speedups in the overall optimization process were obtained, achieving the same or even better objective results.

This new approach allows a very fast approximation to the optimal control of distributed process systems. This property will allow the application of this technique for real time optimization or model predictive control.

**Acknowledgments:** This work was supported in part by the EU (project FAIR CT96-1192) and the Spanish Government (CICyT project ALI97-1939-CE). Author Balsa-Canto thanks the Diputación Provincial de Pontevedra, Spain, for a pre-doctoral fellowship.

## References


[ABS00] A. A. Alonso, J. R. Banga, and I. Sánchez. Passive control design for distributed parameter process systems: Theory and applications. *AICHE J.*, 46(8):1593–1606, 2000.

[BAMS94] J. R. Banga, A. A. Alonso, R. I. P. Martín, and R. P. Singh. Optimal control of heat and mass transfer in food and bioproducts processing. *Comp. & Chem. Eng.*, 18:S699–S705, 1994.

[BC87] J. R. Banga and J.J. Casares. ICRS: Application to a wastewater treatment plant model. In *Process Optimisation*, IChemE Symposium Series No. 100, pages 183–192. Pergamon Press, Oxford, UK, 1987.

[BC99] E. Bendersky and P.D. Christofides. A computationally efficient method for optimization of transport-reaction processes. *Comp. & Chem. Eng.*, 23(s):447–450, 1999.

[BHL93] G. Berkooz, P. Holmes, and L. Lumley. The Proper Orthogonal Decomposition in the analysis of turbulent flows. *Ann. Rev. Fluid Mech.*, 25:539–375, 1993.

[BMGC91] J.R. Banga, R.P. Martin, J.M. Gallardo, and J.J. Casares. Opimization of thermal processing of conduction-heated canned foods: Study of several objective functions. *J. Food Engineering*, 14(1):25–51, 1991.

[BS96] J. R. Banga and W. D. Seider. Global optimization of chemical processes using stochastic algorithms. *In "State of the Art in Global Optimization", C. A. Floudas and P. M. Pardalos (Eds.), Kluwer Academic Pub., Dordrecht, The Netherlands*, pages 563–583, 1996.

[BS97] M. Blatt and K. Schittkowski. PDECON: A FORTRAN code for solving control problems based on ordinary, algebraic and partial differential equations. Technical report, Dep. of Mathematics, University of Bayreuth, 1997.

[Bur93] J. B. Burl. Reduced-order system identification using the Karhunen-Loève transform. *Int. J. of Modelling & Simulation*, 13(4):183–188, 1993.

[Dur97] T. D. Durance. Improving canned food quality with variable retort temperature processes. *Trends in Food Sci. & Technol.*, 8(4):113–118, 1997.

[HLB96] P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge University Press, 1996.

[Oh95] M. Oh. *Modelling and Simulation of Combined Lumped and Distributed Processes.* PhD thesis, Imperial College, University of London, London, U.K., 1995.

[PKC98] H. M Park, T.H. Kim, and D.H. Cho. Estimation of parameters in flow reactors using the Karhunen-Loève decomposition. *Comput. & Chem. Eng.*, 23:109–123, 1998.

[PRG+96] L. Petzold, J. B. Rosen, P.E. Gill, L.O. Jay, and K. Park. Numerical optimal control of parabolic PDEs using DASOPT. IMA workshop on large scale optimization. 1996.

[Ram94] W. F. Ramirez. *Process Control and Identification.* Academic Press, Inc., San Diego, USA, 1994.

[Sch91] W.E. Schiesser. *The Numerical Method of Lines.* Academic Press, New York, 1991.

[Sch94] W. E. Schiesser. *Computational Mathematics in Engineering and Applied Science: ODEs, DAEs and PDEs.* CRC Press, Inc., 1994.

[Sir87] L. Sirovich. Turbulence and the dynamics of coherent structures. Part I: Coherent structures. *Quaterly of Appl. Math.*, pages 561–571, 1987.

[SK98] S. Y. Shvartsman and I. G. Kevrikidis. Nonlinear model reduction for control of distributed systems: A computer–assisted study. *AICHE J*, 44(7):1579–1595, 1998.

[Smo83] J. Smoller. *Shock Waves and Redaction-Diffusion Equations.* Springer-Verlag, NewYork, 1983.

[ST95] J. Shen and R. Temam. Nonlinear Galërkin Method using Chebyshev and Legendre Polynomials. I. The one-dimensional case. *Society for Ind. and Appl. Math.*, 32(1):215–234, 1995.

[TAZ98] A. Theodoropoulou, R. A. Adomaitis, and E. Zafiriou. Model reduction for optimization of rapid thermal chemical vapor deposition systems. *IEEE Trans. on Semiconduct. Manufact.*, 11(1):85–98, 1998.

[Vas93] V. S. Vassiliadis. *Computational Solution of Dynamic Optimization Problems with General Differential-Algebraic Constraints.* PhD thesis, Imperial College, University of London, London, U.K., July 1993.