Dynamic Optimization of Dissipative PDE Systems Using Approximate Inertial Manifolds *

Antonios Armaou and Panagiotis D. Christofides
Department of Chemical Engineering
University of California, Los Angeles, CA 90095-1592

Abstract

In this work, we propose a computationally efficient method for the solution of *dynamic* constraint optimization problems arising in the context of spatially-distributed processes governed by highly-dissipative nonlinear partial differential equations (PDEs). The method is based on spatial discretization using combination of the method of weighted residuals with spatially-global basis functions and approximate inertial manifolds. We use the Kuramoto-Sivashinsky equation, a model of wavy behavior, to demonstrate the implementation and evaluate the effectiveness of the proposed optimization method.

1. Introduction

The standard approach for the solution of the dynamic nonlinear programs (NLPs) with dissipative PDE constraints involves the discretization of the spatial and temporal domain using finite-element or finite-difference techniques and subsequently the solution of the resulting large-scale nonlinear program (NLP) using optimization techniques for sparse NLPs, such as reduced gradient and reduced successive quadratic programming methods (see, for example, [1,2]). The main limitation of this approach is that the nonlinear program resulting from the temporal and spatial discretization is typically of very high-order (in order to compute the optimal solution with the desired accuracy), and thus, it cannot be efficiently solved. The reason for which this approach may be computationally inefficient is that a brute force discretization with finite differences/elements does not account for the inherent characteristics of the PDE equality constraints.

To overcome this limitation, we recently employed [3] a new approach to the solution of steady-state optimization problems arising in the context transport-reaction processes (described by parabolic PDEs) which is based on spatial discretization using the method of weighted residuals with empirical eigenfunctions as basis functions. The empirical eigenfunctions are constructed by applying Karhunen-Loève (K-L) expansion to process solution data. This approach to spatial discretization takes into consideration the presence of dominant spatial patterns in the solution of the parabolic PDEs and leads to reduced-order NLPs that can be solved significantly faster compared to NLPs resulting from spatial discretization using the finite-difference method (see [3] for detailed comparisons).

In this work, we propose a computationally efficient method for the solution of *dynamic* constraint optimization problems arising in the context of spatially-distributed processes governed by highly-dissipative nonlinear partial differential equations (PDEs). The method is based on spatial discretization using combination of the method of weighted residuals with spatially-global basis functions and approximate inertial manifolds. The proposed method accounts for the fact that the dominant dynamics of highly-dissipative PDE systems are low-dimensional in nature and lead to approximate optimization problems that are of significantly-lower order compared to the ones obtained from spatial discretization using finite-difference and finite-element techniques, and thus, they can be solved with significantly smaller computational demand. We employ backward finite-differences (implicit Euler) to perform temporal discretization and solve the nonlinear programs resulting from temporal and spatial discretization using reduced gradient techniques (MINOS). We apply the proposed optimization method to the Kuramoto-Sivashinsky equation.

*Financial support for this work from the National Science Foundation, CTS-0002626, is gratefully acknowledged.

2. Formulation of the optimization problem

We focus on spatially-distributed processes modeled by highly dissipative PDE systems with the following state-space description:

$$\frac{\partial x}{\partial t} \quad = \quad \mathcal{A}(x) + f(t,x,d) \tag{1}$$

subject to the mixed-type boundary condition and the initial condition:

$$g(x, \frac{dx}{d\eta}, \ldots, \frac{d^{n_o-1}x}{d\eta^{n_o-1}}) = 0, \quad on\ \Gamma, \quad x(z,0) = x_0(z) \tag{2}$$

In the above PDE system, $x(z,t) \in \mathbb{R}^n$ denotes the vector of state variables, $t \in [0, t_f]$ is the time (t_f is the terminal time), $z = [z_1, z_2, z_3] \in \Omega \subset \mathbb{R}^3$ is the vector of spatial coordinates, Ω is the domain of definition of the process and Γ its boundary. $\mathcal{A}(x)$ is a dissipative, possibly nonlinear, spatial differential operator which includes higher-order spatial derivatives, $f(t,x,d)$ is a nonlinear, possibly time-varying, vector function which is assumed to be sufficiently smooth with respect to its arguments, $d(t) \in \mathbb{R}^p$ is the vector of design variables which are assumed to be piecewise continuous functions of time, $g(x, \frac{dx}{d\eta}, \ldots, \frac{d^{n_o-1}x}{d\eta^{n_o-1}})$ is a nonlinear vector function which is assumed to be sufficiently smooth (n_o is the order of the PDE of Eq.1), $\left.\frac{dx}{d\eta}\right|_{\Gamma}$ denotes the derivative in the direction perpendicular to the boundary and $x_0(z)$ is a smooth vector function of z. A general optimization problem for the system of Eqs.1-2 can be formulated as follows:

$$\begin{aligned} & min \int_0^{t_f} \int_{\Omega} G(x(z,t), d(t)) dz dt \\ & -\frac{\partial x}{\partial t} + \mathcal{A}(x) + f(t,x,d) = 0, \\ & x(z,0) = x_0(z), \quad g(x, \frac{dx}{d\eta}, \ldots, \frac{d^{n_o-1}x}{d\eta^{n_o-1}}) = 0 \quad on\ \Gamma \\ & g(x,d) \leq 0, \quad \forall\ z \in \Omega, \quad t \in [0, t_f] \end{aligned} \tag{3}$$

where $\int_0^{t_f} \int_{\Omega} G(x,d) dz dt$ is the objective function and $g(x,d)$ is the vector of inequality constraints which may include bounds on the state and design variables. Both $G(x,d)$ and $g(x,d)$ are assumed to be sufficiently smooth functions of their arguments. We focus on the computation of a local optimum.

3. Method of weighted residuals

In this subsection, we derive finite-dimensional approximations of the infinite-dimensional nonlinear program of Eq.3 by using the method of weighted residuals. To simplify the notation, we consider the optimization program of Eq.3 with $n = 1$. We initially expand the solution $x(z,t)$ in an infinite series in terms of a complete set of basis functions $\phi_k(z)$:

$$x(z,t) \quad = \quad \sum_{k=1}^{\infty} a_k(t) \phi_k(z) \tag{4}$$

where $a_k(t)$ are time-varying coefficients. Substituting the expansion of Eq.4 into Eq.3, multiplying the PDE and the inequality constraints with the weighting functions, $\psi_\nu(z)$, integrating over the entire spatial domain, and truncating the series expansion of $x(z,t)$ up to order N and keeping the first N equations (i.e. $\nu = 1, \ldots, N$), the infinite-dimensional program of Eq.3 reduces to the following one with ODE equality constraints, where the optimization parameters are the design variables $d(t)$ and the time varying coefficients $a_{kN}(t)$:

$$
\begin{aligned}
&min \int_0^{t_f} \int_\Omega G(\sum_{k=1}^{N} a_{kN}(t)\phi_k(z), d)dzdt \\
&-\sum_{k=1}^{N} \dot{a}_{kN}(\int_\Omega \psi_\nu(z)\phi_k(z)dz) + \int_\Omega \psi_\nu(z)\mathcal{A}(\sum_{k=1}^{N} a_{kN}(t)\phi_k(z))dz \\
&+\int_\Omega \psi_\nu(z) f(t, \sum_{k=1}^{N} a_{kN}(t)\phi_k(z), d)dz = 0 \\
&\int_\Omega \psi_\nu(z) g(\sum_{k=1}^{N} a_{kN}\phi_k(z), d)dz \leq 0
\end{aligned} \tag{5}
$$

where $a_{kN}(t)$ is the approximation of $a_k(t)$ obtained by an N-th order truncation. From Eq.5, it is clear that the form of the algebraic equality and inequality depends on the choice of the weighting functions, as well as on N. Owing to the smoothness of the functions $G(x,d)$, $\mathcal{A}(x)$, $f(t,x,d)$, $g(x,d)$ and the completeness of the set of basis functions, $\phi_k(z)$, the nonlinear program of Eq.5 is a well-defined approximation of the infinite-dimensional program Eq.3 in the sense that its optimal solution converges to the optimal solution of the program of Eq.3 as $N \to \infty$.

4. Approximate inertial manifolds

In this section, we propose an approach to the solution of the program of Eq.3 which is based on combination of the method of weighted residuals with the concept of approximate inertial manifolds. Following the derivation of a large-scale discretization of the PDE system using the method of weighted residuals (consider the dynamic nonlinear program of Eq.5 with N large), the central idea is to use the concept of inertial manifold to reduce the ODE system that describes the dynamics of the higher-order (fast) eigenmodes to an algebraic one. To present this procedure and simplify the notation, we rewrite the program of Eq.5 in the following form:

$$
min \int_0^{t_f} G(a_N, d)dt, \;\; \dot{a}_N = \tilde{f}(a_N, d), \;\; \tilde{g}(a_N, d) \leq 0 \tag{6}
$$

where $a_N(t) = [a_{1N} \; \cdots \; a_{kN}]^T$ is the vector of the time-varying coefficients of the basis eigenfunctions and the explicit form of $\tilde{f}$ and $\tilde{g}$ can be directly derived from Eq.5. We consider the optimization program of Eq.6 and let $a_s(t)$ be the vector of the modes that are associated with the dominant dynamics of the PDE system and $a_f(t)$ be the modes that are associated with dynamics that decay very fast but are important in terms of capturing the long-time behavior of the PDE. Using this decomposition, the dynamic nonlinear program of Eq.6 can be written as:

$$
min \int_0^{t_f} G(a_s, a_f, d)dt, \quad \dot{a}_s = \tilde{f}_s(a_s, a_f, d), \quad \dot{a}_f = \tilde{f}_f(a_s, a_f, d), \quad \tilde{g}(a_s, a_f, d) \leq 0 \tag{7}
$$

Since the dynamics of the fast modes decay very fast, we can formally set the time-derivative of a_f equal to zero to obtain the following approximate program:

$$
min \int_0^{t_f} G(a_s, a_f, d)dt, \quad \dot{a}_s = \tilde{f}_s(a_s, a_f, d), \quad 0 = \tilde{f}_f(a_s, a_f, d), \quad \tilde{g}(a_s, a_f, d) \leq 0 \tag{8}
$$

In the case of highly dissipative PDE systems, a rigorous justification of the above approximation can be obtained through the concepts of inertial manifold and approximate inertial manifolds (see, for example, [4] for applications of this approach to feedback control of parabolic PDE systems).

Remark 1: When there is a need to capture the evolution of the fast transients, one can complement the dynamic nonlinear program of Eq.8 with an approximation of the dynamic nonlinear program of Eq.7 that captures its behavior in the short-time interval, $[0, \tau_f]$, needed for the dynamics of the fast modes to settle.

This approximate nonlinear program can be used to compute $d(t)$ in the interval, $[0, \tau_f]$, and has the following form:

$$min \int_0^{\tau_f} G(a_s(0), a_f(t), d)dt, \;\; \dot{a}_f = \tilde{f}_f(a_s(0), a_f(t), d), \;\; \tilde{g}(a_s(0), a_f(t), d) \leq 0 \tag{9}$$

With the above formulation, one can solve the above program to compute $d(t)$ in the interval, $[0, \tau_f]$, and then solve the nonlinear program of Eq.8 to compute $d(t)$ in the interval, $[\tau_f, t_f]$. This type of two-time-scale decomposition may be useful when the initial conditions associated with the fast modes are far from the equilibrium manifold and therefore, the approximation $0 = \tilde{f}_f(a_s, a_f, d)$ is not valid for short times. The accuracy of this two-time-scale decomposition of the nonlinear program of Eq.7 improves as the separation between the slow and fast modes increases (which can be always accomplished by increasing the number of modes included in the slow set).

Remark 2: The resulting dynamic nonlinear programs resulting from the above discretization include equality constraints that constitute a low-order system of coupled ordinary differential equations and algebraic equations, and can be then solved with combination of standard temporal discretization and nonlinear programming techniques. We employ backward finite-differences (implicit Euler) to perform temporal discretization and solve the nonlinear programs resulting from temporal and spatial discretization using reduced gradient techniques (MINOS).

5. Application to Kuramoto-Sivashinsky equation

In this section, we present an application of the proposed optimization method to the Kuramoto-Sivashinsky equation with distributed actuation:

$$\frac{\partial x}{\partial t} = -\nu \frac{\partial^4 x}{\partial z^4} - \frac{\partial^2 x}{\partial z^2} - x\frac{\partial x}{\partial z} + b(z)u(t) \tag{10}$$

subject to the periodic boundary conditions and the initial condition:

$$\frac{\partial^j x}{\partial z^j}(-\pi, t) = \frac{\partial^j x}{\partial z^j}(+\pi, t), \quad j = 0, \ldots, 3, \quad x(z, 0) = x_0(z) \tag{11}$$

where $x(z,t)$ is the state of the system, $z \in [-\pi, \pi]$ is the spatial coordinate, t is the time and 2π is the length of the spatial domain, ν is the instability parameter, $x_0(z)$ is the initial condition, $u(t)]$ is the magnitude of the actuation, and $b(z)$ is the actuator distribution function. A direct computation of the solution of the above eigenvalue problem for the spatial differential operator yields $\lambda_0 = 0$ with $\psi_0(z) = \frac{1}{\sqrt{2\pi}}$, and $\lambda_n = -\nu n^4 + n^2$ (λ_n is an eigenvalue of multiplicity two) with eigenfunctions $\phi_n(z) = \frac{1}{\sqrt{\pi}}\sin(nz)$ and $\psi_n(z) = \frac{1}{\sqrt{\pi}}\cos(nz)$ for $n = 1, \ldots, \infty$. From the expression for the eigenvalues, it follows that for a fixed value of $\nu > 0$ the number of unstable eigenvalues of $\mathcal{A}$ is finite and the distance between two consecutive eigenvalues increases as n increases. This implies that for a fixed value of $\nu > 0$, the dominant dynamics of the KSE can be described by a finite-dimensional system. We take $\nu = 0.12$ in which case the spatially-uniform steady state $x(z,t) = 0$ is unstable, and the state moves to a spatially-nonuniform steady-state Therefore, we formulate the optimization problem as the one of computing an optimal input trajectory $u(t)$ for the actuators such that a meaningful cost that includes penalty on the process response and the control action is minimized, in the presence of constraints in the magnitude of the actuation. Mathematically, this optimization problem is as follows:

$$\begin{aligned} min\ J &= \int_0^{t_f}\int_{-\pi}^{\pi} (w_s x^2 + w_u u^2)\, dz\, dt \\ \frac{\partial x}{\partial t} &= -\nu\frac{\partial^4 x}{\partial z^4} - \frac{\partial^2 x}{\partial z^2} - x\frac{\partial x}{\partial z} + b(z)u(t) \\ \frac{\partial^j x}{\partial z^j}(-\pi, t) &= \frac{\partial^j x}{\partial z^j}(+\pi, t), j = 0, \ldots, 3, \;\; x(z,0) = x_0(z), \;\; |u(t)| \leq M \end{aligned}$$

where $M = 3.0$, $w_s = 100$ and $w_u = 20$. We took $x_0 = 0.5\sum_{i=1}^{3} sin(iz) + 1.5\sum_{i=4}^{6} sin(iz)$, and assumed that two actuators with $b_1(z) = \delta(z + 0.5\pi)$ and $b_2(z) = \delta(z - 0.5\pi)$ (point actuation applied at $z = -0.5\pi$ and $z = 0.5\pi$) are available. We initially tried to compute an optimal solution to the above problem by performing spatial discretization using Galerkin's method with the eigenfunctions of the spatial differential operator (sinusoidal functions) as basis functions and temporal discretization using implicit Euler. Optimal solution profiles of $u(t)$ were computed for different numbers of basis functions (in all these cases the step of temporal discretization was appropriately adjusted to guarantee numerical stability of the temporal integration). Figure 1 shows solution profiles of $u(t)$, for N=3,4,5,6. Clearly, these profiles show that convergence to a single optimal profile improves as more basis functions are used. However, even for $N = 6$, convergence of the input $u_1(t)$ to the optimal profile has not been obtained, and it is clear that a higher-order discretization is needed to obtain a convergent profile for $u_1(t)$. To be able to achieve convergence with a low-order approximation, we subsequently used the proposed combination of Galerkin's method with approximate inertial manifolds to solve the optimization problem for the same initial condition. Figure 2 (top plots) shows solution profiles of $u(t)$ for different orders of approximation. Clearly, these profiles converge to a single optimal profile; note the small difference (especially for small times) in the optimal solution profiles for $u_1(t)$ and $u_2(t)$ between the $(3, 5)$ and $(5, 5)$ dynamic nonlinear programs. Figure 2 (bottom plot) shows the profile of the state $x(z, t)$ under the $u(t)$ obtained from the solution of the $(5, 5)$ dynamic nonlinear program. It is clear that the optimal input profile leads to operation of the process close to the spatially uniform steady-state at a finite time. The time needed to solve the optimization problem using this approach is the fraction of the time needed to solve this problem when spatial discretization is performed using finite-differences or through a high-order linear Galerkin truncation. Finally, to further improve the accuracy of the $(3, 5)$ optimization program for small times, the proposed combination of Galerkin's method with modified approximate inertial manifolds, formulation of Eqs.8-9, was used. Figure 3 (top plots) shows optimal solution profiles of $u(t)$ for different orders of approximation. Clearly, the convergence properties have been substantially improved for small times and the $(3, 5)$ and $(5, 5)$ optimization programs give identical results. Figure 3 (bottom plot) shows the profile of the state $x(z, t)$. The optimal input profile again leads to operation of the process close to the spatially uniform steady-state at a finite time. In summary, we clearly see that the combination of Galerkin's method and approximate inertial manifolds, as well as the two-time-scale modification of Eq.9 lead to improved results (in terms of the order of the convergent approximation and the time needed to obtain the optimal solution), compared to discretization using linear Galerkin's method. The reason is that the chosen initial condition excites higher-order modes of the KSE that cannot be captured with a low-order ODE approximation.

6. References

1. Vasantharajan, S., J. Viswanathan, L. T. Biegler, "Reduced successive quadratic programming implementation for large-scale optimization problems with smaller degrees of freedom," *Comp. & Chem. Eng.*, **14**, 907-915, 1990.
2. Biegler, L. T., J. Nocedal and C. Schmid, "A Reduced Hessian Method for Large-Scale Constrained Optimization," *SIAM Journal of Optimization*, **5**, 314-347, 1995.
3. Bendersky, E. and P. D. Christofides, "Optimization of Transport-Reaction Processes Using Nonlinear Model Reduction," *Chem. Eng. Sci.*, **55**, 4349-4366, 2000.
4. Christofides, P. D., "Nonlinear and Robust Control of Partial Differential Equation Systems: Methods and Applications to Transport-Reaction Processes", Birkhäuser, Boston, 2001.

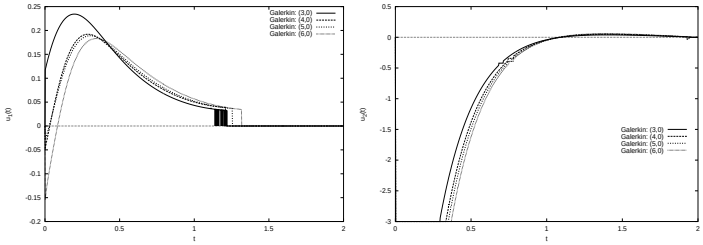

Figure 1: Profiles of $u(t)$ in the case of using Galerkin's method with sinusoidal basis functions.

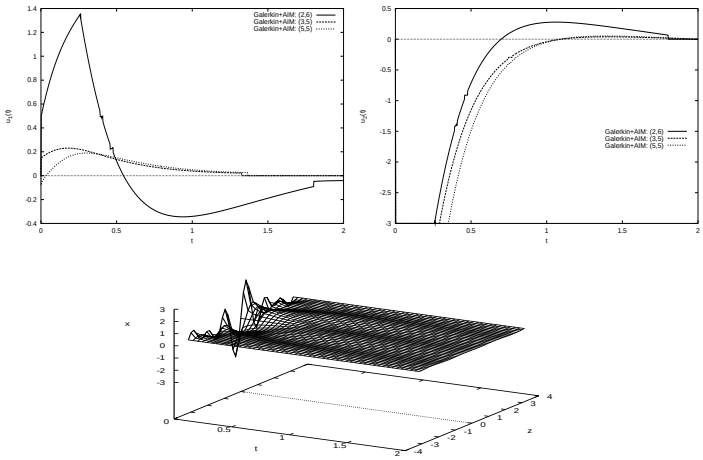

Figure 2: Top two plots: Profiles of $u(t)$ in the case of spatial discretization using Galerkin's method with approximate inertial manifolds. Bottom plot: Profile of the state of the KSE for optimal $u(t)$ (Galerkin+AIM (5,5)).

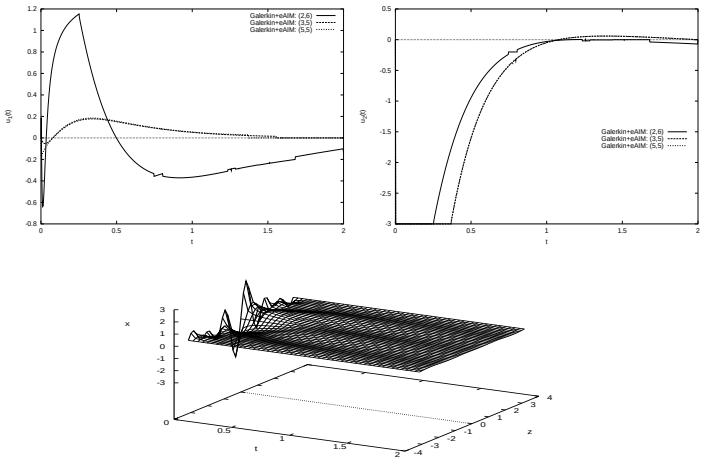

Figure 3: Top two plots: Profiles of $u(t)$ in the case of spatial discretization using Galerkin's method with modified approximate inertial manifolds. Bottom plot: Profile of the state of the KSE for optimal $u(t)$ (Galerkin+eAIM (5,5)).