

Equation-free, coarse-grained computational optimization using timesteppers

Aditya Bindal^a, Marianthi G. Ierapetritou^a, Suhrid Balakrishnan^a, Antonios Armaou^b,
Alexei G. Makeev^c, Ioannis G. Kevrekidis^{d,*}

^aDepartment of Chemical and Biochemical Engineering, Rutgers University, Piscataway, NJ 08854, USA

^bChemical Engineering Department, The Pennsylvania State University, University Park, PA 16802, USA

^cDepartment of Computational Mathematics and Cybernetics, Moscow State University, Moscow, 119899, Russia

^dDepartments of Chemical Engineering, PACM and Mathematics, Princeton University, Princeton, NJ 08544, USA

Received 5 January 2005; received in revised form 16 June 2005; accepted 28 June 2005

Available online 19 September 2005

Abstract

System level optimization computations for engineering problems are typically based on continuum level, macroscopic system descriptions, obtained using accurate closures. In many cases, however, including micro/nanoscale systems, the best available description is a fine scale (atomistic, stochastic or agent-based) model for which accurate, coarse-grained, system level descriptions are not known. The recently introduced equation-free approach [Theodoropoulos, K., Qian, Y.-H., Kevrekidis, I.G., 2000. “Coarse” stability and bifurcation analysis using timesteppers: a reaction diffusion example. *Proceedings of the National Academy of Sciences* 97, 9840–9843; Gear, C.W., Kevrekidis, I.G., Theodoropoulos, C., 2002. ‘Coarse’ integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods. *Computers and Chemical Engineering* 26, 941–963; Kevrekidis, I.G., Gear, C.W., Hummer, G., 2004. Equation-free: the computer-assisted analysis of complex, multiscale systems. *A.I.Ch.E. Journal* 50, 1346–1354; Kevrekidis, I.G., Gear, C.W., Hyman, J.M., Kevrekidis, P.G., Runborg, O., Theodoropoulos, K., 2003. Equation-free multiscale computation: enabling microscopic simulators to perform system-level tasks. *Communications in Mathematical Sciences* 1, 715–762] provides a computational bridge between the underlying microscopic process model and system level numerical computations. In this paper, we employ the equation-free approach to perform system level optimization by acting directly on microscopic/stochastic models. The approach substitutes the evaluation of closed form macroscopic equations with the design and execution of appropriately initialized short bursts of fine scale simulation; processing the simulation results yields estimates of the quantities (residuals, actions of Jacobians and Hessians) required for continuum computations. We illustrate the combination of “coarse timesteppers” with standard (both local and global) optimization techniques. The efficiency of alternative optimization formulations is compared; we see that it can be enhanced by exploiting a separation of time-scales in the system dynamics. The approach constitutes a computational “wrapper” around microscopic/stochastic simulators; yet it can also be wrapped around legacy continuum dynamic simulators.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Optimization; Numerical analysis; Equation-free; Timestepper

1. Introduction

The dynamics of many physicochemical systems occur across different length scales, often classified as microscopic, mesoscopic and macroscopic. Traditionally, when

analyzing and designing engineering systems we are interested at the coarse-grained, macroscopic, systems level behavior and performance; the models that are used for these tasks are evolution equations (e.g., mass, momentum and energy balances) closed through constitutive equations (e.g., Newton’s law of viscosity, chemical kinetics). At the atomistic, *fine* scale, the dynamics involve the evolution of interacting entities (molecules); yet the macroscopic equations

* Corresponding author. Tel.: +1 609 258 2818; fax: +1 609 258 0211.

E-mail address: yannis@Princeton.edu (I.G. Kevrekidis).

are written in terms of macroscopic *observables*: typically a few lower-order moments of the evolving microscopic entity distributions, such as concentration or density (zeroth moments) or momentum fields (the first moment of the distribution of molecules over velocities). Techniques of continuum numerical analysis are applied to these continuum equations to analyze and optimize features of the system behavior at the macroscopic, coarse-grained level.

In contemporary engineering modeling we are faced with many new systems for which the coarse-grained equations conceptually exist, but are simply not known; accurate closures—which embody the effects of the detailed, microscopic evolution on the coarse-grained behavior—are not available in closed form. A *fine scale*, molecular, “microscopic” model of the system is available (e.g., kinetic Monte Carlo (kMC), lattice-Boltzmann, molecular dynamics or Brownian dynamics), and we have reasons to believe that “emergent”, coarse-grained dynamic behavior arises from this model, even though no accurate closed form description of this behavior is available. A similar situation arises in the engineering study of nanoscopic systems in which fluctuations are inherently important. Here, the “coarse-grained” equations would be equations for the statistics of ensembles of replicas of such systems (e.g., expectations); we are again missing the closures giving rise to accurate deterministic equations for these statistics—the quantities that we want to design, control or optimize.

Over the last few years, we have been developing an *equation-free* computational framework for the study of coarse-grained dynamics of such “effectively simple” systems—systems for whose behavior coarse-grained equations conceptually exist but are not available in closed form (Theodoropoulos et al., 2000; Gear et al., 2002; Makeev et al., 2002a; Kevrekidis et al., 2003). This approach constitutes a bridge between traditional, continuum numerical analysis and fine scale (atomistic/stochastic) simulation. The simple underlying idea is to substitute function and derivative evaluations of the (unavailable in closed form) coarse-grained model with the design and execution of short bursts of computational experimentation with the fine scale simulator. The results of these simulations are processed using standard systems theory tools (estimation, filtering, variance reduction) to produce *on demand* estimates of the macroscopic quantities (residuals, action of slow Jacobians, Hessians) required for continuum numerical analysis. Equation-free codes are thus two-tiered (possibly multitiered) codes; at the top level lie traditional continuum numerical algorithms (in our case, optimization algorithms); yet the subroutine calls to a macroscopic model are substituted with calls to the *coarse timestepper*: an input–output map between coarse-grained, macroscopic observables at time t and their values at time $t + \delta t$ (the result of evolving the system with the given initial conditions for time δt). Because of the ease of prescribing arbitrary initial conditions in a computer experiment (compared to physical experimentation) and by using short bursts of

microscopic simulation from nearby initial conditions, we can estimate the action of Jacobians of the unavailable macroscopic equation; more generally, using *matrix-free iterative linear algebra techniques* we can implement linear and non-linear equation solvers, eigensolvers, etc., as computational wrappers around coarse timesteppers (Shroff and Keller, 1993; Christodoulou and Scriven, 1988; Saad and Schultz, 1986). Many continuum numerical tasks (integration, steady-state computation, stability and bifurcation analysis, controller design, and, as we will illustrate in this paper, optimization) can be performed wrapping this computational enabling technology around microscopic simulators (Kevrekidis et al., 2003; Siettos et al., 2003a; Armaou et al., 2004; Armaou, 2005; Siettos et al., 2005; Varshney and Armaou, 2005); for an application to dynamic optimization (see Armaou and Kevrekidis, 2005).

This “wrapper” methodology is common computational practice in optimization; in effect, it corresponds to developing a response surface for the problem in question using (computational) experiments with what we more or less consider to be a black box code; the “twist” in this case is that the black box macroscopic input–output relation is obtained through appropriately designed *fine scale* computations. There is both long tradition and extensive current research on wrapper-type methods for (both local and global) optimization, where the essential call is to a “black box” subroutine evaluating the objective function (Eldred et al., 2003; Kolda et al., 2003; Biegler and Grossmann, 2004; Grossmann and Biegler, 2004). In this paper, we will discuss and illustrate the combination of coarse timesteppers with such optimization methodologies.

It is interesting that the same computational technology that we wrap here around coarse timesteppers can also be wrapped around macroscopic dynamic “legacy codes” used as black boxes, in effect enabling a black box dynamic simulator to perform tasks that it has not been directly designed for; see, for example, the recursive projection method of Shroff and Keller (1993), turning a dynamic simulator into a fixed point algorithm. Both coarse and deterministic “legacy” timesteppers inherently contain effects of the dynamics of the process; we will see that when a separation of time-scales and an associated “slow manifold” exists in the system long-term dynamics, it can be exploited to enhance the computations by effectively optimizing a reduced problem on this slow manifold (see, the examples and discussion in Siettos et al., 2003c).

The paper is organized as follows: Section 2 contains a brief introduction to the different aspects of the problem, while alternative optimization formulations are listed in Section 3. Our motivating kMC example is discussed in Section 4 illustrating the significance of the proposed optimization approach for fine scale models. An illustrative case study is then discussed in Section 5 for a problem for which both a kMC description and an accurate macroscopic, ordinary differential equation (ODE)-based description are available. Standard optimization methods are applied in conjunction

with both deterministic and coarse timesteppers (using the ODE and the kMC model, respectively). The results for different optimization formulations, optimization strategies and reduction through time-scale separation are presented. Our motivating kMC problem is then revisited, followed by conclusions in Section 6.

2. Optimization using timesteppers

2.1. Timesteppers

A timestepper is an input–output, discrete-time model of the evolution of a process; its input is a prescribed initial condition representing the state of the system at time t , while its output is the state of the system at time $t + \delta t$, resulting from the evolution of the system for time δt . If we do have a closed, deterministic evolution equation describing the process, the timestepper contains a call to an appropriate integration code, which evolves the state—typically with adaptive time-steps and error control—and returns the value of the state at the *time reporting horizon* δt . If we believe that such an equation exists, but do not have it available in closed form, we can obtain the same result through the *fine scale* realization of the process through the following steps:

- The *lifting* step, where the available macroscopic initial conditions are used to create an ensemble of (macroscopically) consistent microscopic ones (e.g., microscopic distributions conditioned on their lower moments).
- The *evolution* step, where the “fine scale” evolution scheme is used to evolve these microscopic initial conditions; and
- the *restriction* step, where the macroscopic variables at time $t + \delta t$ are computed from the evolved microscopic final conditions (e.g., computation of moments, possibly averaging over the simulation ensemble).

These lifting and restriction steps are crucial elements of the equation-free approach.

Constructing distributions conditioned on various statistics may be relatively straightforward (e.g., creating a distribution with a given mean and standard deviation) or quite intricate (e.g., creating distributions of points on a lattice with several prescribed high-order moments, such as pair or triplet probabilities, or even pair-correlation functions). Techniques for attempting such initializations exist and are successfully used (Torquato, 2001). Clearly, lifting is not a “one-to-one” mapping—there are many more degrees of freedom in the fine system state, and many initializations consistent with a few macroscopic observables exist. The knowledge (already assumed) that we can write a closed macroscopic model based on only a few observables suggests that, even if these additional degrees of freedom are initialized “inaccurately”, a short direct transient will equilibrate them (slave them) to the few, slow macroscopic vari-

ables (Kevrekidis et al., 2003). This brief “healing” period brings the system close to the slow manifold, establishing the missing closure; that is why we often refer to equation-free computation as a “*closure on demand*” approach. It is always desirable to “lift” (initialize) as close to the closure (the slow manifold) as possible; this becomes especially important if the chosen macroscopic observables are adequate for parameterizing the slow manifold but contain both fast and slow components. This is discussed in more detail in Gear and Kevrekidis (2005), Gear et al. (2004). Indeed, the “preparatory” step bringing the state close to the slow manifold may involve constrained evolution, reminiscent of procedures like SHAKE or umbrella sampling in computational chemistry (Ryckaert et al., 1977; Torrie and Valleau, 1974).

The restriction step may sometimes be relatively simple to implement (e.g., computing the average of a distribution); yet estimating the expectation of fluctuating processes, especially in problems that are distributed in space, may require sophisticated maximum likelihood techniques (see, for example, the Thermodynamic Field Estimator of Yip and coworkers, Li et al., 1998). A more detailed description of these issues can be found in earlier works (Gear et al., 2002; Siettos et al., 2000; Makeev et al., 2002b).

2.2. Some optimization facts

In traditional engineering/industrial modeling, the cost functions are usually obtained from models cast in terms of macroscopic variables, such as product concentrations, temperature, pressure, etc. We are often interested in optimizing a cost function that depends on *steady-state* process outputs, given a range of input parameters. Depending on the nature of the objective function (differentiability, smoothness, convexity), different approaches are utilized to obtain the (global) optimal solution. In this paper, we are interested in “wrapping” standard continuum optimization algorithms around macroscopic input–output system descriptions obtained through a microscopic/stochastic model; the state variables are typically low-order moments of atomistically/stochastically evolving distributions. The outputs of the model and the objective function are thus expected to be noisy; therefore, derivative information of the objective function with respect to the variables is non-trivial to estimate accurately, and can affect the performance of the algorithms. Moreover, fluctuations may lead to several apparent local minima in the objective function, which are not of interest, and the presence of noise might “trap” conventional gradient-based methods.

Approaches to the problem of optimization of noisy functions in the literature can be traced back to either the response surface methodology or to the field of stochastic approximation. The response surface methodology is based on the work of Box and Wilson (1951) who were concerned with minimizing an unknown quadratic objective function perturbed by random noise of constant strength. They

proposed constructing a local linear or quadratic model of the objective function by performing experiments in the neighborhood of the current iterate, and taking a step in the direction of steepest descent as derived from this model. According to Trosset and Torczon (1997), the experimental points considered in these studies can be used as patterns in pattern search methods. Response surface methodology, thus, is a direct precursor of methods such as those of Hooke and Jeeves (1961), Spendley et al. (1962), Nelder and Mead (1965), Torczon (1997), Humphrey and Wilson (2000), and Anderson and Ferris (2001). Surface response ideas have been used also in the area of global optimization where the objective function is built based on a set of sample points. Jones (2001) and Jones et al. (1998) provide an overview of global optimization methods based on this idea.

Stochastic approximation on the other hand dates back to work of Robbins and Monro (1951) and Kiefer and Wolfowitz (1952). The latter authors suggested the use of finite differencing for obtaining an approximation to the gradient of an unknown, noisy function, and proceed in the direction of this (approximate) gradient. The *implicit filtering* algorithm of hyperlinkbib26Gilmore and Kelley (1994) and the simultaneous perturbation stochastic approximation algorithm of Spall (1992, 2000) are derived from this approach. Stochastic methods such as simulated annealing, or genetic algorithms have been successful in locating good suboptimal solutions (Srinivas and Patnaik, 1994). While these methods are relatively simple to implement, their convergence depends on a number of user-defined parameters that cannot be easily evaluated. Nevertheless, some of the recent strategies which have received considerable attention and have been found to be effective for noisy expensive functions include differential evolution (DE) (Storn and Price, 1995), multicoordinate search (Huyer and Neumaier, 1998), radial basis function (Guttman, 2001; Bjorkman and Holmstrom, 2001), implicit filtering (Choi and Kelley, 2000; Gilmore and Kelley, 1994) and DIRECT methods (Jones et al., 1993; Kelley, 1999). A deterministic global optimization framework has been recently proposed (Meyer et al., 2002) for the case of constrained non-linear programming problems in which some of the constraints are not available in closed form. Our objective is not to evaluate these algorithms; it is rather to illustrate how it is possible to link such algorithms to coarse timesteppers and perform coarse-grained optimization using the underlying atomistic/stochastic simulators directly.

2.3. On multiple time-scales

Many chemical processes involve evolution phenomena occurring at different time-scales (from molecular reactive and unreactive collisions to the evolution of a chemical concentration in a stirred reactor). Transport and mixing, and their interaction with reactive dynamics also introduce different evolution time-scales. At the macroscopic level, a set of simultaneous reactions, depending on the activation en-

ergy for each elementary step, may exhibit several response time-scales; often these time-scales are separated by orders of magnitude. When such a large time-scale separation exists, the long-term system dynamics may become attracted to a slow manifold in phase space, on which lie the steady states (stable and unstable), limit cycles and other attractors of the system. This separation of scales can be quantified by inspecting, for example, the spectrum of the linearized dynamics around the system steady states; its signature will be a gap between a group of “slow” eigenvalues (close to the imaginary axis) and a group of “fast” stable ones.

Commonly used approaches that take advantage of separation of time-scales for model reduction/computational acceleration purposes include the quasi-steady-state approximation (QSSA), the method of intrinsic low-dimensional manifolds (ILDM), and computational singular perturbation (CSP). The main idea of these approaches is to generate an appropriate (non-stiff) approximation of a dynamical system by projecting it on an appropriate set of basis vectors so that the contribution of fast modes can be neglected. These methods have been used in the reduction of chemical mechanisms (Turanyi et al., 1996; Lovas et al., 2000) and identification of the slow manifold using geometric considerations (Davis and Skodje, 1999). In CSP (Lam, 1993; Lam and Goussis, 1994) successive approximations are generated to match the initial conditions to the dynamics of the slow manifold. CSP successively refines a set of basis vectors for non-linear problems, uncoupling the fast and the slow modes. Zagaris et al. (2004) showed, in a detailed study, that the successive application of the CSP algorithm generates, order by order, the asymptotic expansion of the slow manifold. The application of the CSP ideas to reduce partial differential equations (PDEs) has been explored by Valorani et al. (2003), in order to enhance the understanding and the accuracy of the qualitative description of flow behavior. More generally, the long-term dynamics of dissipative PDEs involving diffusion, viscosity and/or heat conduction, are often characterized by low dimensionality. The theory of inertial manifolds, leading to approximate inertial manifold algorithms, also effectively exploits time-scale separation to obtain reduced models (Nicolaenko et al., 1989; Temam, 1988; Foias et al., 1988; Jolly et al., 1990; Christofides, 2001).

In our case, the ability to model atomistically/stochastically evolving *fine scale* dynamics through a macroscopic closed equation in terms of only a few variables is also, clearly, a reduction process. This reduction is based again on the separation of time-scales; typically the low-order moments of the simulated distribution evolve over slow time-scales, while the higher-order moments can be thought of relaxing fast to functionals of the low-order, slow ones. A clear illustrative example is the case of molecular simulations of Newtonian fluid: even if the stresses are not initialized to be proportional to velocity gradients, it will typically take only a few collision times for them to relax to being proportional to velocity gradients—to approach the “slow manifold” described through the Newtonian viscosity closure.

In all the above cases—whether in reduced chemical kinetics models, approximate inertial forms or coarse timesteppers, separation of time-scales effectively allows us to accurately reduce the model, so that the steady state and optimization computations are performed in a reduced space.

3. Optimization framework

Optimization is performed on the coarse-grained variables. In particular, the model presented to the optimization software is a discrete time, input–output model (a timestepper). This is done both in a case where the appropriate macroscopic equation is known (so that “black box” deterministic timestepper results are compared with coarse timestepper results obtained using a stochastic simulation) and in a case where the known macroscopic closures are inaccurate.

Assuming that the macroscopic system behavior can be described by the following ODE system:

$$\frac{dx}{dt} = f(x, p),$$

where x are the (observable, macroscopic) variables and p the system parameters, the optimization problem is then defined as minimization of $F(x^*; p)$, where x^* is the steady-state solution of this ODE system and $F(x^*; p)$ is the chosen objective function. This can be formulated as:

- (1) A constrained optimization problem, considering the steady-state equations as constraints:

$$\begin{aligned} \min_{x,p} \quad & F(x; p), \\ \text{s.t.} \quad & f(x, p) = 0. \end{aligned} \quad (1)$$

In this formulation the optimization variables include both the parameters and the variables of the model.

- (2) An unconstrained optimization utilizing a penalty function:

$$\min_{x,p} F(x; p) + \lambda f(x, p), \quad (2)$$

where λ is a penalty parameter the value of which should be judiciously chosen. Although the number of optimization variables remains the same, unconstrained problems may, in some cases, be easier to solve given a good choice for λ (Fletcher, 1981); see also the more general review (Forsgren et al., 2002).

- (3) An unconstrained optimization, evaluating the system steady state independently:

$$\min_p F(x^*; p), \quad (3)$$

where $f(x^*, p) = 0$. In this case a non-linear solver is required to locate the steady state of the system.

3.1. Timesteppers and coarse timesteppers

The same system level optimization performed using the closed form equations can be implemented using timesteppers. The objective function is thus defined on the solutions of the fixed point equation:

$$\Phi(x^*, p) - x^* = 0, \quad (4)$$

where $\Phi(x, p)$ represents the output from the timestepper after a (relatively short) reporting time horizon and x^* denotes the stationary values of the model variables. The continuum dynamic model itself can be used as a “legacy code” timestepper.

For the coarse timestepper, fine scale (e.g., kMC) simulations through the “lift-evolve-restrict” approach are used to estimate the expected value of the system observables after the time-reporting horizon. The same constrained and unconstrained optimization formulations described above can be solved with x^* determined through the solution of the fixed point equation using the (possibly coarse) timestepper $\Phi(x, p)$. In particular:

- Constrained optimization

$$\begin{aligned} \min_{x,p} \quad & F(x; p), \\ \text{s.t.} \quad & \Phi(x, p) - x = 0. \end{aligned} \quad (5)$$

- Penalty-based unconstrained optimization

$$\min_{x,p} F(x; p) + \lambda[\Phi(x, p) - x]. \quad (6)$$

- Decoupled unconstrained optimization

$$\min_p F[\Phi(x^*, p); p]. \quad (7)$$

4. Motivating example

To illustrate the significance of atomistic/stochastic modeling and optimization around it, let us consider a lattice-gas model of the catalytic CO oxidation reaction on a $N \times N$ square lattice, representing a single metal crystal surface (a Pt surface), as described in Makeev et al. (2002b). Our objective is to optimize the (stationary) reaction rate for this process as a function of operating process parameters (specifically, the oxygen gas phase pressure). The model considers six elementary steps, involving two types of adsorbed species in the reaction mechanism. Each lattice site is characterized by the local occupation numbers of CO and O adsorbed species. The model takes into account the lateral interactions between CO_{ads} at the nearest neighbor (NN) lattice sites. Also, it is assumed that the migration rates of CO_{ads} and O_{ads} significantly exceed the rates of all other

elementary processes such as adsorption, desorption and reaction of CO₂ formation. The time evolution of the reaction system can be described by the chemical master equation. In general, this equation cannot be solved directly; therefore, one has to either use some approximations in order to derive macroscopic evolution equations, or to implement a kMC simulator which can calculate the solution of the lattice-gas model with high accuracy. This motivating example has been chosen because *in certain parameter regimes* it is possible to obtain accurate macroscopic closures, while in other regimes—which, as we will see below, critically affect our optimization results—these explicit closures become inaccurate.

Assuming that the adsorbed layer is randomly well-mixed, the reaction can be approximated at the macroscopic level by means of the mean-field approximation (MFA). This simplest approximation gives two ODEs for the CO_{ads} and O_{ads} surface coverages:

$$\begin{aligned}\frac{d\theta_1}{dt} &= \alpha\theta_0 - \gamma\theta_1 \exp\left[\frac{4\theta_1\varepsilon}{RT}\right] - R_{\text{CO}_2}, \\ \frac{d\theta_2}{dt} &= 4\beta\theta_0^2 - R_{\text{CO}_2},\end{aligned}\quad (8)$$

where $R_{\text{CO}_2} = 4k_r\theta_1\theta_2 \exp[3\theta_1\varepsilon/(RT)]$ is the reaction rate; θ_1 and θ_2 is the CO_{ads} and O_{ads} coverage, respectively; $\theta_0 = 1 - \theta_1 - \theta_2$ accounts for empty sites; ε is the energetic parameter of lateral interactions between NN CO_{ads} species (negative for repulsive interactions); T is the absolute temperature and R the ideal gas constant. The dimensionless constants α , β and γ are associated with the rates of adsorption of CO, dissociative adsorption of oxygen, and desorption of CO_{ads}, respectively. In addition, k_r represents the rate constant of the surface reaction between adsorbed CO and the adsorbed O atoms to give CO₂ which immediately desorbs upon formation. MFA becomes increasingly inaccurate for strong interactions and high adsorbate coverages, because it neglects the spatial correlations in the non-ideal adsorbed overlayer.

In a more refined modeling step, we may employ the so-called *quasi-chemical approximation* (QCA) introducing the pair probabilities, $g_{ij} = g_{ji}$. The QCA consists of a set of five coupled differential-algebraic equations:

$$\frac{d\theta_1}{dt} = \alpha\theta_0 - \gamma\theta_1(S_1)^4 - R_{\text{CO}_2},$$

$$\frac{d\theta_2}{dt} = 4\beta g_{00} - R_{\text{CO}_2},$$

$$r_{11}g_{11}g_{00} = (g_{01})^2,$$

$$r_{22}g_{22}g_{00} = (g_{02})^2,$$

$$r_{12}g_{12}g_{00} = g_{01}g_{02},$$

$$R_{\text{CO}_2} = 4k_r g_{12}(S_1 S_2)^3,$$

$$\sum_j g_{ij} = \theta_i \quad (i, j = 0, 1, 2),$$

$$r_{ij} = r_{ji} = \exp[-\varepsilon_{ij}/(RT)] \quad (i, j = 1, 2),$$

$$S_i = (g_{i0} + r_{i1}g_{i1} + r_{i2}g_{i2})\theta_i \quad (i, j = 1, 2). \quad (9)$$

Here, $\varepsilon_{ij} = \varepsilon_{ji}$ are the energetic parameters of lateral interactions between NN adsorbed particles. In our case, $\varepsilon_{11} = \varepsilon < 0$, while all other interaction energies are equal to zero. The independent variables are two surface coverages and three pair probabilities, for example, g_{11} , g_{22} and g_{12} .

Because of lateral interactions, the rates of chemical reactions and diffusion depend on the local environment; as a result, an ordered adsorbed overlayer may form on the lattice. Repulsive CO_{ads}–CO_{ads} lateral interactions give rise to a so-called $c(2 \times 2)$ structure which corresponds to an Ising anti-ferromagnet. Neither MFA nor QCA are able to model these effects correctly at high coverages; therefore, kMC simulations were implemented to calculate the “correct” stationary state coverages and reaction rate. Our implementation of kMC simulations as well as the algorithms for timestepper-based coarse-grained kMC (C-kMC) bifurcation computations have been described in previous publication (Makeev et al., 2002b); they are very briefly reviewed here.

The C-kMC algorithm assumes that the (expected) dynamics of the lattice-gas model can *in principle* be described on a macroscopic level by a system of “coarse” ODEs for average coverages Θ :

$$\frac{d\Theta}{dt} = F(\Theta),$$

where the function F is unavailable in closed form. All higher-order moments of the distributions (e.g., pair probabilities) do not appear in the deterministic equation; the model *closes* with coverages only. This is a significant statement—it is easy to see that *over very short time-scales*, say of the order of a few kMC events, this statement cannot be true; the details of initializing the lattice—i.e., not only coverages, but pair probabilities, triplet probabilities etc.,—should be important for the instantaneous reaction rate. Yet over longer time-scales, due to a separation of time-scales in the evolution of adatom distributions, higher-order correlation functions are quickly slaved to the coverages because of the fast mobility of adsorbed species. This implies that the evolution of the moments of the adatom distributions can be considered as a singularly perturbed system; over longer time-scales the dynamics can be successfully modeled by a *reduced* model—a continuous-time ODE for the evolution of only coverages, the zeroth moments of adatom distributions. Yet—as we will see below—the true optimum cannot be accurately captured by neither of the closed form explicit reduced models

(MFA and QCA); instead, we had to resort to C-kMC computations.

Since we want to optimize the *stationary* reaction rate, we are clearly interested in finding stationary states. The corresponding equations have to be solved at each (feasible) optimization step or must be satisfied upon convergence; we briefly describe how to solve them here before we consider using them in an optimization framework. The C-kMC algorithm for the location of stationary states contains the lifting, evolution and restriction steps as described above. Through the “lift-evolve-restrict” procedure and exploiting the Newton–Raphson iterative method with numerical derivatives to solve the equation $\Theta - \Phi(\Theta) = 0$, the C-kMC algorithm calculates both stable and unstable stationary solutions of the “unavailable” fixed point system of equations $F(\Theta) = 0$.

The following parameter values are considered in this paper: $\alpha = 1.6 \text{ s}^{-1}$, $\gamma = 0.001 \text{ s}^{-1}$, $k_r = 1 \text{ s}^{-1}$, $\varepsilon = -2 \text{ kcal/mol}$, $T = 500 \text{ K}$. To calculate a stationary reaction rate of the lattice-gas model, N_{run} different equilibrated lattices, obtained by a Metropolis algorithm for the fixed stationary CO and O coverages, were used for averaging. Figs. 1(a) and (b) show bifurcation diagrams with respect to the parameter β , whose variation corresponds to variation of the gas-phase pressure of oxygen. The stationary CO coverages are shown in Fig. 1(a), while Fig. 1(b) gives the stationary values of reaction rate, R_{CO_2} . The bifurcation diagrams for MFA, QCA and C-kMC are overlaid in the figures. They have been obtained by combining the (coarse-grained) fixed point algorithm described above with arclength continuation. It is clear that both MFA and QCA become inaccurate at high CO coverages, and as a result are not useful for optimization under such conditions; C-kMC on the other hand, does capture the correct long-term kMC dynamics both at low and at high CO coverages.

5. An illustrative case study

5.1. Problem Definition

Before we proceed with the optimization of our motivating problem, we demonstrate equation-free optimization computations using timesteppers in a test problem for which the microscopic computations are simpler and much faster. For this problem a good macroscopic closed model exists in the regime of interest, so that we can compare the results obtained through coarse timestepping with those obtained through accurate closures; we selected a simple chemical reaction mechanism example. In the deterministic setting, the reactions are assumed to take place in an ideal CSTR and the reaction network we consider is a modification of the Fugitt and Hawkins mechanism (Floudas and Pardalos, 1990). The reaction network involves five species and

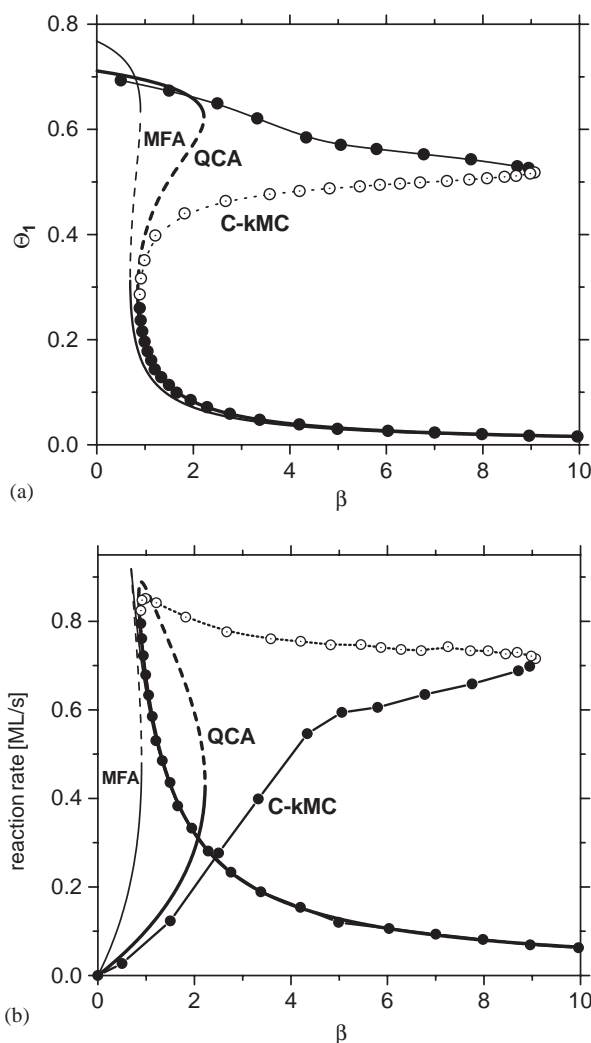
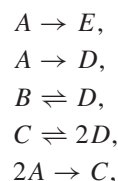


Fig. 1. Bifurcation diagrams of the CO oxidation mechanism obtained via pseudo-arclength continuation using MFA (thin lines), QCA (thick lines) and C-kMC (circles joined by interpolation curves) with β as the bifurcation parameter. Parameters of the C-kMC algorithm: $\tau = 0.001$, 200×200 lattice, $N_{\text{run}} = 10^5$. (a) Solid (dashed) lines show stable (unstable) stationary state branches of CO coverage, θ_1 . (b) Solid (dashed) lines show stable (unstable) stationary state branches of the reaction rate, R_{CO_2} .

reactions as follows:



where the rate constants are $k_1^f = 3.33384 \text{ s}^{-1}$, $k_2^f = 0.26687 \text{ s}^{-1}$, $k_3^f = 0.29425 \text{ s}^{-1}$, $k_3^r = 0.18957 \text{ s}^{-1}$, $k_4^f = 0.011932 \text{ s}^{-1}$, $k_4^r = 0.18957 \text{ L}(\text{mol s})^{-1}$ and $k_5^f = 0.009598 \text{ L}(\text{mol s})^{-1}$. It is assumed that only species A and C enter the CSTR, which has a residence time of $\tau = 12.5 \text{ s}$. The range of permissible inlet concentrations for species A and

C are $C_A^0 \in [3.0, 30.0]$ mol/L and $C_C^0 \in [0.0, 10.0]$ mol/L, respectively. The system of ODEs describing the dynamic behavior of the CSTR is as follows:

$$\begin{aligned}\frac{dC_A}{dt} &= \frac{1}{\tau} (C_A^0 - C_A) - k_1^f C_A - k_2^f C_A - k_5^f C_A^2, \\ \frac{dC_B}{dt} &= \frac{1}{\tau} (-C_B) - k_3^f C_B + k_3^r C_D, \\ \frac{dC_C}{dt} &= \frac{1}{\tau} (C_C^0 - C_C) - k_4^f C_C + 0.5k_4^r C_D^2 \\ &\quad + 0.5k_5^f C_A^2, \\ \frac{dC_D}{dt} &= \frac{1}{\tau} (-C_D) + k_2^f C_A + k_3^f C_B - k_3^r C_D \\ &\quad + 2k_4^f C_C - k_4^r C_D^2, \\ \frac{dC_E}{dt} &= \frac{1}{\tau} (-C_E) + k_1^f C_A.\end{aligned}\quad (10)$$

The objective function involves the steady-state concentration values of species C and D , namely C_C^* and C_D^* , respectively, as follows:

$$F(x, y) = 4(x - 0.6)^2 + 4(y - 0.4)^2 + \sin^3(\pi x) + 0.4, \quad (11)$$

where

$$x = 0.1428C_C^* - 0.357C_D^*, \quad (12)$$

$$y = -0.1428C_C^* + 2.857C_D^* - 1.0. \quad (13)$$

Selecting the values of initial conditions and utilizing the linear transformation to the steady-state values of C_C^* and C_D^* , results in $x, y \in (0, 1)$. Figs. 2(a) and (b) present a three-dimensional and a contour graph of $F(x, y)$ as function of the variables x and y , respectively; we observe the presence of a local and a global minimum.

We also considered a stochastic implementation of the same kinetic scheme using the stochastic simulation algorithm (SSA) (Gillespie, 1976, 1992). In the SSA, the extensivity parameter Ω relates the numbers N_p of molecules of species p ($p = A, B, C, D, E$) to the concentrations $C_p = N_p/\Omega$, and controls the total number of particles that can be present in the system. For the CSTR, in addition to the seven chemical reaction events, events accounting for input and output to the reactor are also taken into account. For example, the transition probability of species A for inflow to the reactor is equal to $(\Omega)(C_A^0)/\tau$, while the transition probability for outflow from the reactor is equal to N_A/τ . Overall, 14 different elementary events are accounted for in the SSA. In the limit of a system with a large number of particles ($\Omega \rightarrow \infty$), SSA and integration of deterministic ODEs give effectively the same results.

5.2. Computational results

The decoupled unconstrained optimization problem of Section 3 using the steady states of Eq. (10) was solved to identify the optimal solutions and served as a benchmark for

optimization with timesteppers. The parameters were randomly initialized. For each call to the objective function the steady state was computed using Newton's method. Table 1 shows the two local solutions obtained for different initial conditions. For comparison purposes, we integrated the deterministic ODE system to form the deterministic timestepper which we used as a "black box" legacy dynamic code; the corresponding coarse timestepper was implemented through averaging several realizations of the SSA algorithm for "the same" extent of macroscopic time. An important numerical issue with the stochastic timestepper is that of noise, which is inherent in the lifting process and the stochastic simulation; this affects both the objective function evaluation and the coarse derivative evaluations. The level of noise depends on the physical system size (number of particles in our simulation) and the number of copies in the kMC simulation ensemble. Asymptotically, an SSA simulation will approach the "corresponding" ODE for large enough system sizes. For validation purposes, we utilize system sizes (number of particles) and ensemble sizes (number of copies) for which the coarse timestepper yields results close to the deterministic ones; more particles and more ensemble copies had comparable effects in reducing fluctuations for this particular SSA coarse timestepper output. Table 2 illustrates this in terms of the objective function and the stationary states of the system for 100 runs using the same macroscopic initial conditions.

Fluctuations due to noise are critical in the estimation of partial derivatives for derivative-based optimization algorithms and for the selection of "scales" (maximum distance along a search direction) for direct search algorithms. If the magnitude of these optimization "scales" is chosen too small, grossly erroneous searches and/or increased computational time with no apparent improvement in the search can result; the reader may refer to Armaou and Kevrekidis (2005) for a discussion of these issues in a dynamic optimization context.

5.3. Local optimization

Three different algorithms for local optimization of the above problem using both ODE-based and coarse timesteppers were investigated—the Nelder–Mead (NM) algorithm and the Hooke–Jeeves (HJ) algorithm, as well as gradient-based sequential quadratic programming (SQP). NM maintains a simplex of points while HJ evaluates the objective function values on a stencil; both search for the optimal solution based purely on objective function evaluations. SQP, on the other hand, iteratively solves a quadratic approximation of the problem requiring explicit local Jacobian information. Centered finite differences are used to numerically estimate this Jacobian. In our simulations, the tolerance for the Newton iteration was set to 5×10^{-4} . The optimization scales used for the HJ algorithm were $[1, 2^{-1}, 2^{-2}, \dots, 2^{-8}]$. The stencil size was successively reduced to the lowest scale and the optimal solution found

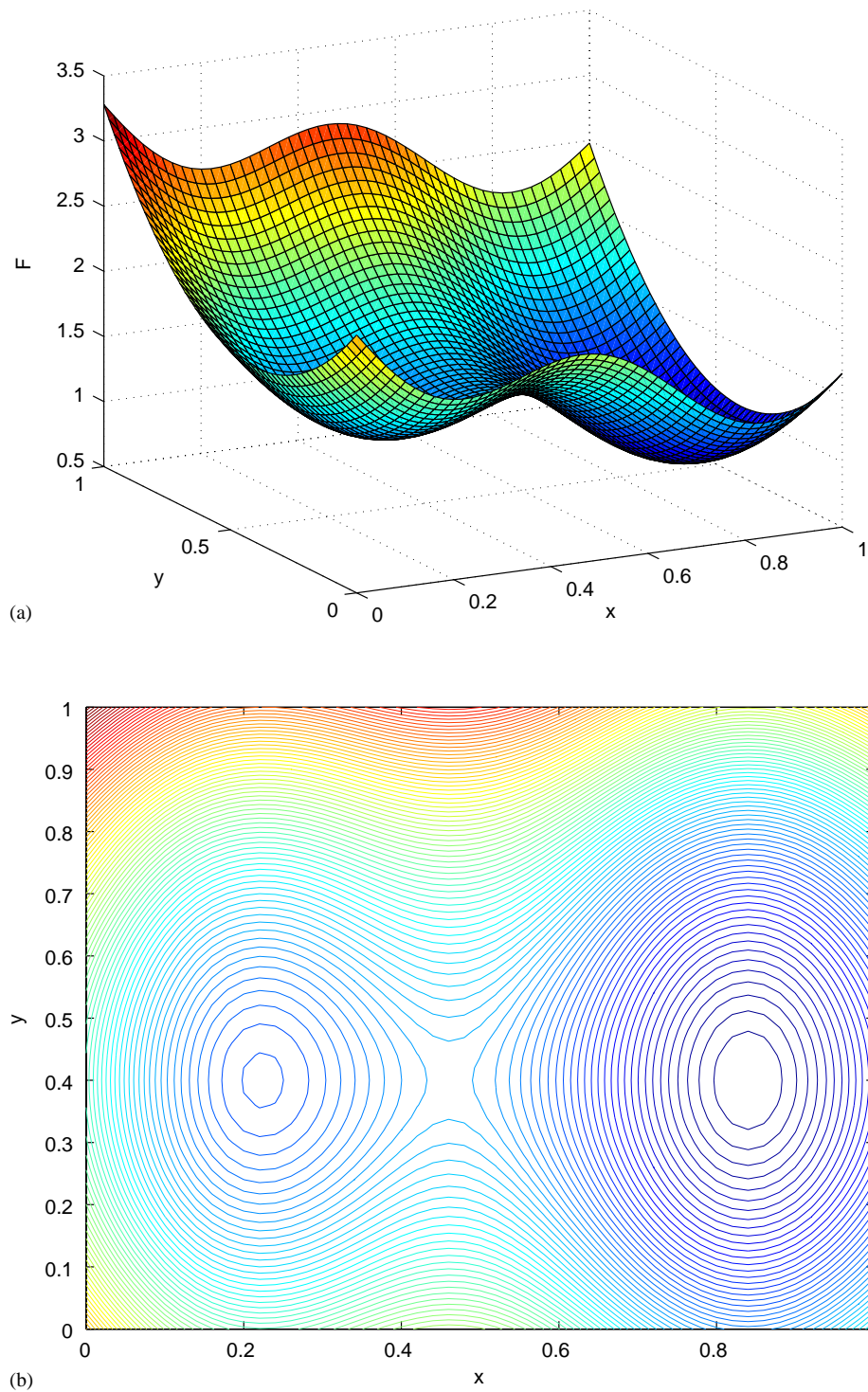


Fig. 2. Profile of objective function of Eq. (11), $F(x, y)$, as a function of x and y : (a) three-dimensional plot; (b) contour plot.

was used to initialize another search before declaring convergence “over all scales” (Kelley, 1999). Tables 3–5 illustrate the performance of NM, HJ and SQP, respectively. Note that as the number of particles increases, the optimal solution found using the coarse timestepper approaches

that of ODE-based one. The number of objective function calls was found to be comparable for SQP and NM; HJ, however, necessitated a larger number of function calls, due to the restarting of the optimization and the stencil scales used.

Table 1
Optimal solutions to the kinetic problem

	Local optimum	Global optimum
Starting point [C_A^0, C_C^0]	[15.0, 5.0]	[8.0, 6.0]
Steady state [C_C^*, C_D^*]	[3.192, 0.649]	[8.121, 0.895]
Optimal solution [C_A^{opt}, C_C^{opt}]	[13.202, 3.163]	[10.117, 8.378]
Objective value	1.23641	0.74221
Eigenvalues of linearization	−0.64398 −0.21787 −0.08000 −0.08000 −3.68621	−0.69891 −0.24433 −0.08000 −0.08000 −3.68904

5.4. Global optimization

Two global optimization strategies were investigated for finding the global optimum of the kinetic problem using both the deterministic timestepper and the coarse-grained, kMC-based one. A variant of DIRECT called multilevel coordinate search (MCS) algorithm (Huyer and Neumaier, 1998) was employed as a deterministic global strategy and a modified genetic algorithm called DE was used as a heuristic global strategy (Storn and Price, 1995). MCS divides the search space into hyper-rectangles where it searches for the local

optima, and keeps a “shopping basket” of hyper-rectangles based on the function values at the vertices/center where it continues the local search by further subdivision. This method guarantees an optimal solution for a deterministic system based on exhaustive search; no such guarantee can be provided for stochastic models since the optimality conditions at the optima are not defined for noisy functions. The choice of the scale parameter in such a strategy would be the size of the hyper-rectangle in the search process, which is estimated by the number of times the initial box has been subdivided. For our simulations, initially the parameter space was divided into boxes with function evaluations at the corners and midpoint, and a maximum of 20 subdivision levels were allowed, with successive local searches in the smaller boxes. DE, on the other hand, is a stochastic method which uses a randomly generated initial population drawn from within the bounding box. It generates a new set of trial vectors by taking the difference of two randomly selected population vectors and adding it to a third randomly selected vector, referred to as the mutation process. The selection is based on the objective function value. The number of population vectors chosen is 10 with a maximum of 50 mutation iterations. The coefficient parameter of the differential (of two randomly selected vectors) can be considered as a “scale parameter” for optimization of noisy functions.

Table 2
Effect of system size and copies on kMC results' variance

Model	No. of particles	No. of copies	Mean [C_C^*, C_D^*]	Mean objective	Std. dev. objective
Legacy	–	–	(4.8699, 0.7077)	1.4728	–
kMC	10^5	1	(4.8666, 0.7074)	1.4742	0.0123
kMC	10^4	10	(4.8664, 0.7062)	1.4742	0.0124
kMC	10^3	10^2	(4.8715, 0.7077)	1.4740	0.0110

Table 3
Nelder–Mead search results

Model	No. of particles	No. of copies	Objective value	Solution [C_A^{opt}, C_C^{opt}]	Time (s)	Obj. fun. calls
SS	–	–	0.7422	(10.122, 8.379)	2.0	44
Legacy	–	–	0.7422	(10.115, 8.378)	38.5	44
kMC	10^4	5	0.7432	(9.505, 8.460)	793.9	56
kMC	10^4	10	0.7424	(10.449, 8.382)	1947.5	67
kMC	10^5	5	0.7422	(10.353, 8.378)	10111.1	69
kMC	10^5	10	0.7422	(10.133, 8.385)	22101.8	75

Table 4
Hooke–Jeeves search results

Model	No. of particles	No. of copies	Objective value	Solution [C_A^{opt}, C_C^{opt}]	Time (s)	Obj. fun. calls
SS	–	–	0.7422	(10.121, 8.379)	4.97	87
Legacy	–	–	0.7422	(10.117, 8.378)	111.26	80
kMC	10^4	10	0.7422	(10.054, 8.339)	2670.5	150
kMC	10^5	5	0.7425	(10.593, 8.410)	19908.5	216
kMC	10^5	10	0.7422	(10.363, 8.382)	25926.1	194
kMC	10^6	5	0.7422	(10.207, 8.383)	162540.0	113

Table 5
Sequential quadratic programming search results

Model	No. of particles	No. of copies	Objective value	Solution [C_A^{opt} , C_C^{opt}]	Time (s)	Obj. fun. calls
SS	–	–	0.7422	(10.121, 8.379)	0.005	50
Legacy	–	–	0.7422	(10.117, 8.378)	0.24	56
kMC	10^5	1	0.7471	(10.327, 8.478)	1321	69
kMC	5×10^5	1	0.7428	(10.185, 8.428)	6544	61
kMC	10^6	1	0.7422	(10.144, 8.377)	11637	56

Table 6
Multilevel coordinate search results

Model	No. of particles	No. of copies	Objective value	Solution [C_A^{opt} , C_C^{opt}]	Time (s)	Obj. fun. calls
SS	–	–	0.7422	(10.119, 8.378)	5.9	142
Legacy	–	–	0.7422	(10.119, 8.378)	117.0	129
kMC	10^4	5	0.7422	(10.482, 8.367)	1878.6	246
kMC	10^4	10	0.7423	(10.359, 8.399)	3368.5	200
kMC	10^5	5	0.7422	(10.307, 8.374)	17473	201
kMC	10^5	10	0.7422	(10.098, 8.396)	18966	109

Table 7
Differential evolution search results

Model	No. of particles	No. of copies	Objective value	Solution [C_A^{opt} , C_C^{opt}]	Time (s)
SS	–	–	0.7422	(10.119, 8.379)	17.9
Legacy	–	–	0.7422	(10.118, 8.378)	434.8
kMC	10^4	5	0.7558	(9.286, 8.371)	4065.4
kMC	10^4	10	0.7423	(10.264, 8.371)	8048.1
kMC	10^5	5	0.7422	(10.011, 8.371)	41425.8
kMC	10^5	10	0.7422	(10.120, 8.373)	83801.7

In our optimization runs a value of 0.8 for all system sizes performed consistently well. Tables 6 and 7 present the optimal solution for different initial system sizes using MCS and DE, respectively. As with local optimization techniques, using increased number of particles results in an optimal solution that approaches that of the ODE-based timestepper. The number of objective function calls for MCS was found to be higher than the local methods, since this approach searches at various locations over the entire domain. The number of function calls for DE was fixed at 500.

5.5. Reduced optimization

As discussed in Section 3, the complexity of the optimization problem can be reduced when separation of time-scales prevails in the dynamics of the problem under consideration. This is qualitatively analogous to using (when appropriate) a QSSA in a high-order model, reducing it to a lower-order one, and performing the optimization computations on the reduced model. If the evolution equations are available they can be used to construct explicit reduced-order models (e.g., Kumar et al., 1998; Prud'homme et al., 2002); this would be

followed by optimization based on the reduced model for steady or dynamic process operation (e.g., Bendersky and Christofides, 2000; Armaou and Christofides, 2002; Otto et al., 1997). When we have a “black box” simulator, the dynamic equations are not explicitly available; as a result we cannot analytically reduce them to a closed, accurate, simpler model. Since timesteppers naturally contain dynamic information about the short-term system evolution, it is possible to use them for “on line” model reduction without explicitly constructing a reduced model. If the long-term dynamics of the system of interest quickly approach a lower-dimensional slow manifold (the graph of a function over a reduced set of variables in phase space), a short integration with the detailed simulator will bring a transient close to this manifold and the corresponding reduced description. Thus, one can use the direct short integration that constitutes part of the timestepper to estimate the dynamics (and the steady-states, in which we are interested) of the full system *close to* this slow manifold, effectively reducing the number of degrees of freedom we need to search in for steady-states during optimization. Our kinetic ODE problem is clearly a good candidate for time-scale-based model reduction: see

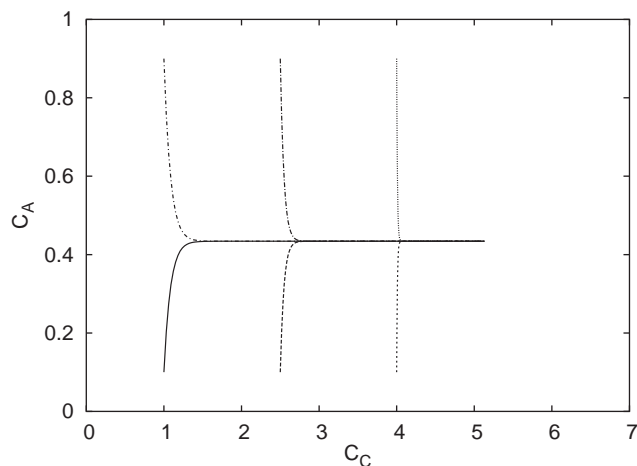


Fig. 3. C_A – C_C phase portrait for system of Eq. (10), during evolution to steady state.

the linearized stability analysis at the steady-state for typical conditions shown in Table 1. This eigen-analysis suggests that variables C_A and C_E will exhibit a relatively fast transient response as compared to other variables. To illustrate the difference in time-scales, a plot exemplifying the two-scale response of C_A as it approaches the slow manifold is shown in Fig. 3. The deterministic ODE process model was integrated for 200 s and the concentration of C_A is plotted as a function of C_C . It can be seen from the plot that C_A approaches a slow manifold in phase space rapidly—from different initial conditions—and then slowly evolves on it towards the ultimate steady-state.

A reduced optimization problem was then solved for the steady-state concentrations of B , C and D . In an optimization context this translates to a reduction in the number of variables for the steady-state calculation for feasible (decoupled) computations, or, alternatively, a reduction in the number of constraints in the constrained optimization formulation with timesteppers. The performance of the method should of course be strongly dependent on the reporting horizon of the timestepper. For successful reduced optimization runs, the reporting horizon of the timesteppers has to be greater than the largest relaxation time corresponding to the fast variables, which, for the case study chosen, was 1.4 s. If the reporting horizon of the timestepper is smaller than the relaxation time of the fast variables, then the fast variables will not have relaxed on the slow manifold, and no reduction in possible; a sufficiently long reporting time horizon of 2 s was chosen for all the calculations here.

The adaptive estimation of this “longest fast variable” relaxation time using timesteppers is discussed in more detail in Siettos et al. (2003b); it involves matrix-free iterative linear algebra eigen-computations, performed on demand to identify a gap between local fast and slow system dynamics. These computations, similar to adaptive time-step or mesh computations in traditional continuum numerical analysis, should be performed periodically to test whether

Table 8
Sequential quadratic programming search results with time-scale separation

Model	Optimization method	Number variables	Time (s)	Obj. fun. calls	Timestepper calls
Legacy	Constrained	7	0.0074	131	269
Legacy-ts	Constrained	5	0.0062	89	139
Legacy	Penalty based	7	0.0100	475	475
Legacy-ts	Penalty based	5	0.0088	194	194
Legacy	Decoupled	2	0.04	52	706
Legacy-ts	Decoupled	2	0.03	51	544

the model can be reduced, and to what number of variables. It is suggested in Siettos et al. (2003b) that the *eigenvectors* corresponding to the slowest among the group of “fast relaxation times” can help detect which of the fast variables to retain as a slow one if we need to augment the dimension of the reduced model. The comparison of results for reduced optimization with optimization in the full space is shown in Table 8. The computational time and number of function calls are averaged for 500 runs from random initial conditions in the variable space. The optimization variable space for the constrained and penalty-based formulation was reduced from 7 to 5. This reduction led here to computational savings of approximately 20%. As can be seen from the reduction to the number of timestepper calls, time-scale separation and subsequent reduction of optimization space dimension may improve the computational efficiency when function calls are very “expensive” computationally (e.g., when using coarse timesteppers). Also, for large-scale engineering problems described by legacy simulators, this separation of time-scales may be useful to exploit in optimization-based wrappers. This subject will be revisited in the discussion below.

5.6. Coarse-grained optimization for the CO oxidation model

We now return to our motivating example, and consider the following optimization problem: maximize the stationary reaction rate of the lattice-gas model with respect to the parameter β (i.e., the gas-phase pressure of oxygen, a parameter one can realistically manipulate in experiments). Fig. 4 shows that deterministic models (MFA and QCA) are not capable of capturing the maximal stationary reaction rate correctly. This is, clearly, because the maximum occurs under conditions (high enough coverages) where these models are not accurately capturing the stationary states. We use the C-kMC algorithm in order to accurately find the stationary solution of the problem with the maximal reaction rate. In a simple feasible optimization scheme, prescribing the oxygen gas phase pressure (the parameter β) we find the stationary states for this value of β , compute the corresponding stationary reaction rate $R_{\text{CO}_2,ss}(\beta)$ and then simply solve $dR_{\text{CO}_2,ss}/d\beta = 0$, using Newton–Raphson with derivatives

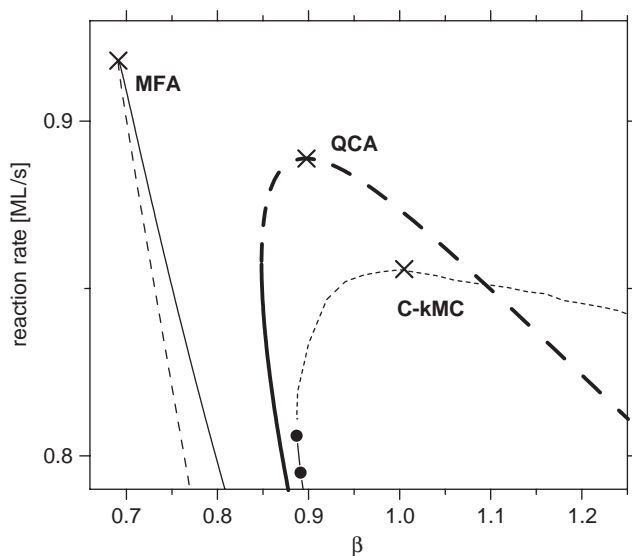


Fig. 4. Detail of Fig. 1(b). Crosses mark the maximum CO oxidation rate, R_{CO_2} , computed using MFA, QCA and C-kMC.

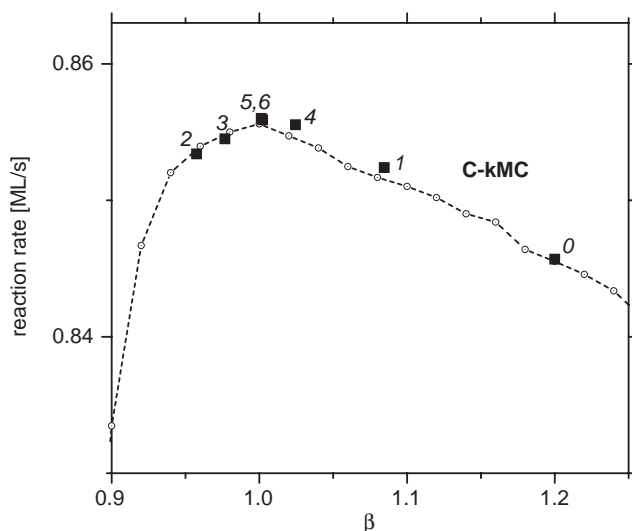


Fig. 5. Convergence of coarse Newton–Raphson method to a local maximum of the stationary reaction rate. Squares and digits represent the sequence of iterations. A perturbation $\delta = 2 \times 10^{-2}$ was used to calculate the numerical derivatives of R_{CO_2} . Parameters of the C-kMC algorithm: $\tau = 0.001$, 500×500 lattice, $N_{\text{run}} = 10^5$.

estimated using the timestepper. Note that coefficients for finite difference estimates of derivatives in the presence of noise have been discussed in Drews et al. (2003).

Fig. 5 demonstrates the convergence of iterations to a local maximum of $R_{\text{CO}_2,ss}(\beta)$ starting from $\beta_0 = 1.2$. It was found that a value of $\beta \approx 1$ provides the maximal rate of CO_2 production at stationary conditions. For these calculations, we used a 500×500 lattice with $N_{\text{run}} = 10^5$ and $\delta = 2 \times 10^{-2}$. The dashed curve in Fig. 5 presents the reaction rate calculated with lower accuracy (using a 200×200 lattice with $N_{\text{run}} = 2 \times 10^5$) for various fixed values of β .

Note that the maximum occurs in a regime with three steady states, and, even more interestingly, very close to the turning point, apparently on the unstable branch. This has interesting implications; for example, feedback control must be used to stabilize the corresponding stationary state if it is unstable; see Siettos et al. (2003a), Armaou et al. (2004), Armaou (2005), Siettos et al. (2005), for an equation-free approach to coarse controller design-based microscopic/atomistic process models. Direct kMC simulation would simply not be able to find this maximum, since it lies on an unstable stationary state branch, which is not accessible to direct dynamic simulation.

6. Conclusions

In this work, we have illustrated the application of optimization search algorithms using timesteppers on microscopic/macroscale input–output black box models. Both gradient-based and simplex-based algorithms were successfully applied for local optimization of deterministic and stochastic models. DE and MCS strategies were also successfully used for global optimization. In all cases the optimal solution was found to approach the solution obtained using a deterministic, ODE-based, timestepper as the system size was increased. For the SSA fine scale simulator, the coarse timestepper results are expected to approach the “corresponding ODE” ones at sufficiently large system sizes. For most microscopic/stochastic fine scale models however (e.g., for nanoscale applications, such as the field emitter tip oscillations of Imbihl and coworkers (Suchorski et al., 1998)) closed form deterministic equations for the statistics of ensembles of experiments are not available, and variance reduction techniques are required to make coarse timestepper-based optimization practical. Finding the appropriate statistics (observables) of the fine scale simulation that become the state variables of the coarse-grained model is an important issue. For the examples discussed in this paper these observables were known (concentrations, coverages, pair probabilities); the discovery of the appropriate observables based on data processing techniques is the subject of intense current work (see, e.g., Belkin and Niyogi, 2003; Nadler et al., 2005).

When accurate, closed form coarse-level models are available, it is obvious that they should be used for optimization instead of the kMC ones. When coarse timestepper-based optimization becomes necessary, the wall clock time can be reduced using distributed computing: each processor evolves a different realization of the same coarse initial condition. It was found that constrained or penalty-based methods may become computationally more efficient than the unconstrained decoupled formulation. Different formulations (e.g., constrained vs. unconstrained) of the optimization problem may provide advantages depending on the problem features; exploiting the time-scale separation of the unavailable macroscopic equations, along with modern

matrix-free linear algebra techniques, seems to hold promise for the coarse computational optimization of what we call “effectively simple” systems (Kevrekidis et al., 2004).

Acknowledgements

Financial support from the Air Force Office of Scientific Research (Dynamics and Control), National Science Foundation, ITR, NSF/CTS 0456657 and Office of Naval Research Contract no. 0014-03-1-0207, and Pennsylvania State University, Chemical Engineering department, is gratefully acknowledged.

References

- Anderson, E.J., Ferris, M.C., 2001. A direct search algorithm for optimization with noisy function evaluation. *SIAM Journal on optimization* 11 (3), 837–857.
- Armaou, A., 2005. Output feedback control of dissipative distributed processes via microscopic simulations. *Computers and Chemical Engineering* 29, 771–782.
- Armaou, A., Christofides, P.D., 2002. Dynamic optimization of dissipative PDE systems using nonlinear order reduction. *Chemical Engineering Science* 57, 5083–5114.
- Armaou, A., Kevrekidis, I.G., 2005. Equation-free optimal switching policies for bistable reacting systems using coarse time-steppers. *International Journal of Robust and Nonlinear Control*, in press.
- Armaou, A., Siettos, C.I., Kevrekidis, I.G., 2004. Time-steppers and ‘coarse’ control of distributed microscopic processes. *International Journal of Robust and Nonlinear Control* 14, 89–111.
- Belkin, M., Niyogi, P., 2003. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation* 15, 1373–1396.
- Bendersky, E., Christofides, P.D., 2000. Optimization of transport-reaction processes using nonlinear model reduction. *Chemical Engineering Science* 55, 4349–4366.
- Biegler, L.T., Grossmann, I.E., 2004. Retrospective on optimization. *Computers and Chemical Engineering* 28, 1169–1192.
- Bjorkman, M., Holmstrom, K., 2001. Global optimization of costly non-convex functions using radial basis functions. Technical Report, Center for Mathematical Modeling, Department of Mathematics and Physics, Malardalen University, Sweden.
- Box, G.E.P., Wilson, K.B., 1951. On the experimental attainment of optimum conditions. *Journal of Royal Statistical Society, Series B* 13 (1), 1–38.
- Choi, T.D., Kelley, C.T., 2000. Superlinear convergence and implicit filtering. *SIAM Journal of Optimization* 10 (4), 1149–1162.
- Christodoulou, K.N., Scriven, L.E., 1988. Finding leading modes of a viscous free surface flow: an asymmetric generalized eigenproblem. *Journal of Scientific Computing* 3, 355–405.
- Christofides, P.D., 2001. *Nonlinear and Robust Control of Partial Differential Equation Systems: Methods and Applications to Transport-Reaction Processes*. Birkhäuser, New York.
- Davis, M., Skodje, R.T., 1999. Geometric investigation of low-dimensional manifolds in systems approaching equilibrium. *Journal of Chemical Physics* 111, 859–874.
- Drews, T.O., Braatz, R.D., Alkire, R.C., 2003. Parameter sensitivity analysis of Monte Carlo simulations of copper electrodeposition with multiple additives. *Journal of Electronic Society* 150, C807–C812.
- Eldred, M.S., Giunta, A.A., van Bloemen Waanders, B.G., Wojtkiewicz, S.F., Hart, W.E., Alleva, M.P., 2003. Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis. Version 3.1 Users Manual SAND2001-3796, Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185-0847.
- Fletcher, R., 1981. *Practical Methods of Optimization*, vol. 2. Wiley, Toronto, New York.
- Floudas, C.A., Pardalos, P.M., 1990. *A Collection of Test Problems for Constrained Global Optimization Algorithms*. Springer, Berlin, New York.
- Foias, C., Jolly, M., Kevrekidis, I.G., Sell, G.R., Titi, E., 1988. On the computation of inertial manifolds. *Physics Letters A* 131, 433–446.
- Forsgren, A., Gill, P.E., Wright, M.H., 2002. Interior methods for nonlinear optimization. *SIAM Review* 44, 525–597.
- Gear, C.W., Kevrekidis, I.G., 2005. Constraint-defined manifolds: a legacy-code approach to low-dimensional computations. *Journal of Scientific Computing*, in press.
- Gear, C.W., Kevrekidis, I.G., Theodoropoulos, C., 2002. ‘Coarse’ integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods. *Computers and Chemical Engineering* 26, 941–963.
- Gear, C.W., Kaper, T.J., Kevrekidis, I.G., Zagaris, A., 2004. Projecting on a slow manifold: singularly perturbed systems and legacy codes. SIADS, submitted, can also be found at arXiv.org/Physics/0405074.
- Gillespie, D.T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* 22, 403–434.
- Gillespie, D.T., 1992. A rigorous derivation of the chemical master equation. *Physica A* 188, 404–425.
- Gilmore, P., Kelley, C.T., 1994. An implicit filtering algorithm for optimization of functions with many local minima. Technical Report, North Carolina State University, Center for Research in Scientific Computation and Department of Mathematics, Raleigh, NC.
- Grossmann, I.E., Biegler, L.T., 2004. Part II. Future perspective on optimization. *Computers and Chemical Engineering* 28, 1193–1218.
- Guttman, H.M., 2001. A radial basis function method for global optimization. *Journal of Global Optimization* 19, 201–227.
- Hooke, R., Jeeves, T.A., 1961. Direct search solution of numerical and statistical problems. *Journal of the Association for Computing Machinery* 8, 212–219.
- Humphrey, D.G., Wilson, J.R., 2000. A revised simplex search procedure for stochastic simulation response surface optimization. *INFORMS Journal on Computing* 12 (4), 272–283.
- Huyer, W., Neumaier, A., 1998. Global optimization by multilevel coordinate search. Technical Report, Institut für Mathematik, Universität Wien, Austria.
- Jolly, M.S., Kevrekidis, I.G., Titi, E.S., 1990. Approximate inertial manifolds for the Kuramoto–Sivashinsky equation: analysis and computations. *Physica D* 44, 38–60.
- Jones, D.R., 2001. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization* 21, 345–383.
- Jones, D.R., Perttunen, C.D., Stuckman, B.E., 1993. Lipschitzian optimization without the Lipschitz constant. *Journal of Optimization Theory and Application* 79 (1), 157–181.
- Jones, D.R., Schonlau, M., Welch, W.L., 1998. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* 13, 455–492.
- Kelley, C.T., 1999. *Iterative Methods for Optimization*, *Frontiers in Applied Mathematics*, vol. 18. SIAM, Philadelphia, USA.
- Kevrekidis, I.G., Gear, C.W., Hyman, J.M., Kevrekidis, P.G., Runborg, O., Theodoropoulos, K., 2003. Equation-free multiscale computation: enabling microscopic simulators to perform system-level tasks. *Communications in Mathematical Sciences* 1, 715–762.
- Kevrekidis, I.G., Gear, C.W., Hummer, G., 2004. Equation-free: the computer-assisted analysis of complex, multiscale systems. *A.I.Ch.E. Journal* 50, 1346–1354.
- Kiefer, J., Wolfowitz, J., 1952. Stochastic estimation of a regression function. *Annals of Mathematical Statistics* 23, 462–466.
- Kolda, T.G., Lewis, R.M., Torczon, V., 2003. Optimization by direct search: new perspectives on some classical and modern methods. *SIAM Review* 45, 385–482.

- Kumar, A., Christofides, P.D., Daoutidis, P., 1998. Singular perturbation modeling of nonlinear processes with non-explicit time-scale multiplicity. *Chemical Engineering Science* 53, 1491–1504.
- Lam, S.H., 1993. Using csp to understand complex chemical kinetics. *Combustion Science and Technology* 89, 375–404.
- Lam, S.H., Goussis, D.A., 1994. The csp method of simplifying kinetics. *International Journal of Chemical Kinetics* 26, 461–486.
- Li, J., Liao, D.Y., Yip, S., 1998. Coupling continuum to molecular dynamics simulation: reflecting particle method and the field estimator. *Physics Review E* 57, 7259–7267.
- Lovas, T., Nilsson, D., Mauss, F., 2000. Automatic reduction procedure for chemical mechanisms applied to premixed methane/air flames. *Proceedings of the Combustion Institute*, pp. 1809–1815.
- Makeev, A.G., Maroudas, D., Kevrekidis, I.G., 2002a. Coarse stability and bifurcation analysis using stochastic simulator: kinetic Monte Carlo examples. *Journal of Chemical Physics* 116, 10083–10091.
- Makeev, A.G., Maroudas, D., Panagiotopoulos, A.Z., Kevrekidis, I.G., 2002b. Coarse bifurcation analysis of kinetic Monte Carlo simulations: a lattice gas model with lateral interactions. *Journal of Chemical Physics* 117, 8229–8240.
- Meyer, C.A., Floudas, C.A., Neumaier, A., 2002. Global optimization with nonfactorable constraints. *Industrial Engineering Chemistry Research* 41, 6413–6424.
- Nadler, B., Lafon, S., Coifman, R.C., Kevrekidis, I.G., 2005. Diffusion maps, spectral clustering and the reaction coordinates of dynamical systems. *Applied and Computational Harmonic Analysis*.
- Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. *The Computer Journal* 7, 308–313.
- Nicolaenko, B., Foias, C., Temam, R., Constantin, P., 1989. *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations*, Applied Mathematical Sciences, vol. 70. Springer, New York, NY.
- Otto, J., Paraschivoiu, M., Yesilyurt, S., Patera, A.T., 1997. Bayesian-validated computer-simulation surrogates for optimization and design: error estimates and applications. *Mathematics and Computers in Simulation* 44, 347–367.
- Prud'homme, C., Rovas, D.V., Veroy, K., Machiels, L., Maday, Y., Patera, A.T., Turinici, G., 2002. Reliable real-time solution of parametrized partial differential equations: reduced-basis output bound methods. *A.S.M.E. Transactions on Journal of Fluids Engineering-Transactions of the ASME* 124, 70–80.
- Robbins, H., Monro, S., 1951. A stochastic approximation method. *Annals of Mathematical Statistics* 29, 400–407.
- Ryckaert, J.P., Ciccotti, G., Berendsen, H., 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *Journal of Computational Physics* 23, 237.
- Saad, Y., Schultz, M.H., 1986. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal of Scientific and Statistical Computing* 7, 856–869.
- Shroff, G.M., Keller, H.B., 1993. Stabilization of unstable procedures: the recursive projection method. *SIAM Journal of Numerical Analysis* 30, 1099–1120.
- Siettos, C.T., Qian, Y., Kevrekidis, I.G., 2000. Coarse stability and bifurcation analysis using time-steppers: a reaction diffusion example. *Proceedings of the National Academy of Science of the United States of America* 97 (18), 9840–9843.
- Siettos, C.I., Armaou, A., Makeev, A.G., Kevrekidis, I.G., 2003a. Microscopic/stochastic timesteppers and coarse control: a kinetic Monte Carlo example. *A.I.Ch.E. Journal* 49, 1922–1926.
- Siettos, C.I., Graham, M., Kevrekidis, I.G., 2003b. Coarse Brownian dynamics for nematic liquid crystals: bifurcation, projective integration and control via stochastic simulation. *Journal of Chemical Physics* 118, 10149–10156.
- Siettos, C.I., Pantelides, C.C., Kevrekidis, I.G., 2003c. Enabling dynamic process simulators to perform alternative tasks: A time-stepper based toolkit for computer-aided analysis. *Industrial Engineering Chemistry Research* 42, 6795–6801.
- Siettos, C.I., Kevrekidis, I.G., Kazantzis, N., 2005. An equation-free approach to nonlinear control: feedback linearization with pole-placement. *International Journal of Bifurcations and Chaos*, in press.
- Spall, J.C., 1992. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control* 37, 332–341.
- Spall, J.C., 2000. Adaptive stochastic approximation by simultaneous perturbation method. *IEEE Transactions on Automatic Control* 45 (10), 1839–1853.
- Spendley, W., Hext, G.R., Himsforth, F.R., 1962. Sequential application of simplex designs in optimization and evolutionary operation. *Technometrics* 4, 441–461.
- Srinivas, M., Patnaik, L.M., 1994. Genetic algorithms: a survey. *Computer* 27, 17–27.
- Storn, R., Price, K., 1995. Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces. *Technical Report*, International Computer Science Institute.
- Suchorski, Y., Imbihl, R., Medvedev, V.K., 1998. Compatibility of field emitter studies of oscillating surface reactions with single crystal measurements: catalytic CO oxidation on Pt. *Surface Science* 401, 392.
- Temam, R., 1988. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Applied Mathematical Sciences, vol. 68. Springer, New York, NY.
- Theodoropoulos, K., Qian, Y.-H., Kevrekidis, I.G., 2000. “Coarse” stability and bifurcation analysis using timesteppers: a reaction diffusion example. *Proceedings of the National Academy of Sciences* 97, 9840–9843.
- Torczon, V., 1997. On the convergence of pattern search algorithms. *SIAM Journal of Optimization* 7, 1–25.
- Torquato, S., 2001. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*. Springer, New York, NY, USA.
- Torrie, G.M., Valleau, J.P., 1974. Monte Carlo free energy estimates using non-Boltzmann sampling: application to the sub-critical Lennard Jones fluid. *Chemical Physics Letters* 28, 578–581.
- Trosset, M.W., Torczon, V., 1997. Numerical optimization using computer experiments. *Technical Report*, Institute for Computing Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23681-2199.
- Turanyi, T., Hughes, K., Pilling, M., Tomlin, A., 1996. Kinalc: program for the analysis of reaction mechanisms. *Technical Report*, Combustion simulations at the University of Leeds.
- Valorani, M., Goussis, D.A., Najm, H.N., 2003. CSP analysis of a transient flame-vortex interaction: time scales and manifolds. *Combustion and Flame* 134, 35–53.
- Varshney, A., Armaou, A., 2005. Multiscale optimization using hybrid PDE/kMC process systems with application to thin film growth. *Chemical Engineering Science* 60, 6780–6794.
- Zagaris, A., Kaper, H.G., Kaper, T.J., 2004. Analysis of the computational singular perturbation reduction method for chemical kinetics. *Journal of Nonlinear Science* 14, 59.