

570g Adaptive Optimization of Noisy Black-Box Functions Inherent in Microscopic Models

Eddie Davis and Marianthi Ierapetritou

Traditionally, system level optimization is performed on a macroscopic continuum level description of the physical system that presumes the accuracy of closed-form equations. However, for micro/nano applications and other systems where exact constitutive relations are unknown or do not provide a good representation of the system's behavior, a microscopic or molecular level description can be used as a realistic representation of the system physics. Furthermore, even though system analysis is usually performed at the macroscopic level, recently researchers (Makeev et al., 2002) have shown that the details of the model description can have a profound effect on these analyses. Since closed-form equations are unavailable for modeling and optimization at a fine resolution, coarse time-steppers can instead be used. Coarse time-steppers estimate a discrete-time input-output mapping generated by closed-form equations (if they were available) using kinetic Monte Carlo simulations. Yet even though time-steppers provide a bridge between microscopic/stochastic simulations and computational optimization (Bindal et al., 2004, Kevrekidis et al, 2003), running microscopic simulations is computationally very expensive and therefore the use of standard optimization techniques can be prohibitive. Thus, there exists a need for the development of robust algorithms capable of economically optimizing these inherently noisy systems. Different approaches have been employed towards the optimization of noisy functions (Kelley, 1999). Global zero-order optimization methods such as genetic and evolutionary algorithms (Storn et al, 1995) can provide optimal solutions. Local zero-order methods utilizing function evaluations on sequences of simplices (Nelder-Mead, multilevel search, and implicit filtering) yield optima with reasonable accuracy (Choi & Kelley, 2000, Huyer and Neumaier, 1999). Although convergence bounds can be established with rigorous higher-order methods, their main disadvantage is that they can become trapped in artificial local optima introduced due to the presence of noise. An approach that is proposed to overcome this problem is optimisation of response surface models (Myers & Montgomery, 2002). In this approach, the stochastic function is approximated by a low-order polynomial using the input-output response mapping, which is then used for optimization, most commonly by steepest descent. The use of steepest descent in order to accelerate convergence performs wells when in the neighborhood of the optimum, but far away from it, slow movement may not justify the computational overhead in building the response surfaces. Although much of the work has focused on improving local optimization of response surfaces, global methods have also been studied (Jones, 2001, Zilinskas, 1991). Let a stochastic system consisting of N particles at the microscopic level describe the set of macroscopic properties x . Considering the optimization of an objective function $g(x, N)$, if the error in the function evaluation is assumed to be independent of x , then this function can be described as $g(x, N) = f(x) + \sigma(N)$, where $\sigma(N)$ is the stochastic error associated with the size of the microscale model. The variance of the stochastic error at any system size N can be given as $\sigma^2(N) = \text{Var}(\{g_i(x, N)\} | i = 1 \dots k)$ for k microscale simulations around the same initial condition. As N increases, the variance decreases. The goal is to obtain $g(x, N)$ to within acceptable accuracy while at the same time minimizing the computational effort in order to perform system-level optimization. In this work we employ three iterative methods for the optimization of noisy functions arising from the description of systems at microscopic level. The first method determines an optimal difference interval with which to compute numerical gradients in order to formulate the optimization problem, which is then solved using nonlinear programming (NLP) solvers. An improvement of this approach is also studied whereby the level of microscopic detail is gradually increased upon approaching the optimum, reducing the computational cost. The second method moves towards the optimum region by selecting the best optimum among iterative simplices. Once this area is identified, refinement of the optimum occurs via movement by steepest descent of response surfaces. In the third technique, the response surface functionality is used within SQP to overcome the limitation of initial fixed local movement of the second approach. A potential new iterate is the optimum of the local function extrapolated over the entire feasible domain. Movement to this point occurs only when this vector is proven to provide

improvement in the objective function over the current candidate solutions. A five-species reaction system is selected in order to investigate the effectiveness of the proposed optimization strategies. System information is obtained by making function calls on two input species. The objective function is given in terms of two non-input species at steady-state. The equivalent of a function call using closed-form equations can be made using time-steppers whereby input species concentrations are first specified at a macroscopic level and “lifted” to a microscopic level. Using Monte Carlo simulations, the system evolves to steady-state after a long time horizon. The steady-state solution vectors are then “restricted” back to macroscopic level concentrations, after which the objective function can be evaluated. The deterministic solution of the system yields both a local and a global optimum. It is shown that all proposed strategies successfully attain one of the optima. In the algorithm utilizing optimal difference intervals to formulate and solve the NLP, 50% computational savings is obtained when the level of system detail is gradually increased to finer resolution as the optimum is neared, over that of when the system size remains fixed throughout. However, the solution error increases from 0.04% to 0.2%. Optimization of response surfaces using either simplicial/steepest descent or SQP achieves an error of 0.1% - 1%. The Simplex/Steepest Descent method requires on average 30% fewer function calls than the SQP-based RSM optimization algorithm because the initial design for each iteration in the latter method requires more sample points. The simplex design for the former method not only uses fewer points in each iteration, but because the size of the local region also remains unchanged from one iteration to the next, any common points from earlier iterations do not need to be recalculated. The comparable performance of the SQP-based algorithm to that of the remaining methods in determining the optimum indicates the competitiveness of this strategy. For the case study, the distance that a random starting point needs to cover to converge to either the local or global optimum turns out to be roughly the same. For other functions in which a global optimum exists near the limits of the design region, the SQP-based algorithm is expected to provide faster convergence to the optimum in contrast to the simplex/steepest descent method as the ability to move in a better derivative-based overall direction will quickly overcome the slower movement of sequential simplices. Our existing work focuses on the use of SQP to optimize non-interpolating response surfaces based on a minimum number of sample points. Compared to the computational effort needed for the method utilizing an optimal difference interval and to earlier results reported in Bindal et. al. (2004), the simplex/steepest descent algorithms are more efficient by two orders of magnitude. It is found that a tradeoff exists between the quality of the solution and the number of function calls required. Algorithmic efficiency in obtaining the optimum can be increased through parallelization of computations by assigning one microscopic simulation to each CPU. The next step is to continue integrating these methodologies in order to develop more efficient local optimization approaches for noisy functions. This work is expected to serve as a cornerstone towards developing a global optimization framework for the same class of noisy functions. **References** [1] Bindal, A., M.G. Ierapetritou, S. Balakrishnan, A. Makeev, I. Kevrekidis and A. Armaou, Submitted for publication, Chem. Eng. Sci., 2004. [2] Choi, T. D. and C.T. Kelley, 2000, SIAM Journal of Optimization 10, 4, 1149. [3] Kevrekidis, I.G., C.W. Gear, J.M. Hyman, P.G. Kevrekidis, O. Runborg, and C. Theodoropoulos, 2003, Comm. Math. Sci., 1, 715. [4] Huyer, W. and A. Neumaier, 1999, J. Global Optimization, 14, 331. [5] Jones, D., 2001, J. of Global Optimization, 21, 345. [6] Jones, D., M. Schonlau, and W. Welch, 1998, J. of Global Optimization, 13, 455. [7] Kelley, C.T., 1999, Iterative Methods for Optimization, Volume 2, SIAM. [8] Makeev, A., D. Maroudas, & I. Kevrekidis, 2002, J. Chem. Physics, 116, 23. [9] Myers, R. and D. Montgomery, 2002, Response Surface Methodology. [10] Storn, R. and Price, K., 1995, Differential Evolution – A Simple And Efficient Adaptive Scheme For Global Optimization Over Continuous Spaces, Technical Report, International Computer Science Institute. [11] Zilinskas, 1992, J. of Global Optimization, 2, 145.