J Math Chem (2010) 48:28–37
DOI 10.1007/s10910-009-9614-5

ORIGINAL PAPER

Initial guess of the solution of dynamic optimization of chemical processes

L. Bayón · J. M. Grau · M. M. Ruiz · P. M. Suárez

Received: 3 August 2009 / Accepted: 28 September 2009 / Published online: 11 October 2009

Abstract Numerical methods and software packages for solving dynamic optimization or optimal control problems require a suitable initial estimation of the solution. This paper focuses on problems that arise in chemical processes described by complex dynamics. We present a very simple method, based on Pontryagin's Minimum Principle, to obtain an initial guess for the solution. Our method presents numerous advantages: it is very easy to programme, it allows a wide range of problems to be addressed, the computation time is very short, the initial guess is very close to the solution and is attracted to a global minimum.

Keywords Optimal control problem · Chemical processes · Pontryagin's minimum principle

Mathematics Subject Classification (2000) MSC 49K24 · MSC 92E20

1 Introduction

Obtaining the accurate solution of optimal control problems is a crucial aspect in many areas of applied science. In this paper we shall focus especially on problems that arise in chemical engineering. In most cases, these systems are described by a nonlinear set of differential-algebraic equations. These complex dynamics lead to several difficulties, such as, for instance, that of often presenting multiple local minima.

There is a vast array of numerical methods and software packages for numerically solving dynamic optimization or optimal control problems, such as: SOCS [1], RIOTS_95 [2], DIRCOL [3], MISER3 [4], MINOPT [5], NDOT [6], DIDO [7],

L. Bayón (✉) · J. M. Grau · M. M. Ruiz · P. M. Suárez
Department of Mathematics, University of Oviedo, Campus de Viesques, 33203 Gijón, Spain
e-mail: bayon@uniovi.es

GPOCS [8], DYNOPT [9] or Bryson's Matlab code [10]. Unfortunately, these packages require an initial guess of the solution to start the iterations; the algorithms converges reasonably efficiently with a good initial guess for a solution.

For example, direct methods employing nonlinear programming such as Sequential Quadratic Programming (SQP) are frequently used as a numerical method for optimal control problems. Although they have advantages in terms of computational robustness and their usefulness for practical problems, it is usually difficult to choose appropriate initial solutions. Therefore, an estimate of the optimal solution is first obtained using, for instance, Genetic Algorithms (GA). This is a classical statement in which the problem is solved in two steps: first with relaxed constraints, then using the obtained result to find the optimal solution.

Sometimes, however, in cases when convergence is not obtained, it is likely that the initial guess is such that convergence to the true solution is impossible. In such cases, users of these packages are accordingly advised to try different sets of initial guesses. Besides, as the complexity of the system increases, the specification of a suitable initial guess can become troublesome.

In this paper, we shall concentrate on the special structure that appears in numerous problems of chemical reactors; more specifically, in the nonlinear continuous stirred tank reactor (CSTR). We shall present a very simple method to obtain an initial guess for the solution for this complex system. The method is based on Pontryagin's minimum principle (PMP). Moreover, we shall show that the initial guess is very close to the solution for the tested chemical processes and that our initial guess is attracted to a global minimum. We shall show that the theory allows us to address a wide range of problems: constrained, unconstrained, nondifferentiable, etc., while employing a very short computation time in all cases.

The paper is organized as follows. Section 2 presents the mathematical background of our work. We shall set out our problem in terms of optimal control in continuous time, using the Lagrange-type functional. We shall use PMP and we obtain a fundamental result, which is the basis for constructing the algorithm that leads to the determining of the initial guess. In Sect. 3, we present several examples of chemical processes in which the structure of the problem allows us to apply the above result. In Sect. 3.1, we first consider the nonlinear continuous stirred tank reactor (CSTR) unconstrained. We generalize the previous example in Sect. 3.2, considering the constrained case with bounded control, and present a nondifferentiable case in Sect. 3.3. Finally, Sect. 4 summarizes the main conclusions of our research.

2 Mathematical formulation

A standard Lagrange type optimal control problem (OCP) can be mathematically formulated as follows:

$$\min_{\mathbf{u}(t)} I = \int_0^{t_f} F\left(t, \mathbf{x}(t), \mathbf{u}(t)\right) \mathrm{d}t \tag{1}$$

subject to satisfying:

$$\dot{\mathbf{x}}(t) = f\,(t, \mathbf{x}(t), \mathbf{u}(t)) \tag{2}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{3}$$

$$\mathbf{u}(t) \in U(t), \quad 0 \le t \le t_f \tag{4}$$

where I is the performance index, F is an objective function, $\mathbf{x} = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}^n$ is the *state vector*, with initial conditions $\mathbf{x}_0$, $\mathbf{u} = (u_1(t), \ldots, u_m(t)) \in \mathbb{R}^m$ is the *control vector* bounded by $\mathbf{u}_{\min}$ and $\mathbf{u}_{\max}$, U denotes the set of admissible control values, and t is the operation time that starts from 0 and ends at t_f. The *state variables* (or simply the *states*) must satisfy the *state equation* (2) with given initial conditions (3). In this statement, we consider the final instant to be fixed and the final state to be free (we do not consider terminal constraints: $\Psi(\mathbf{x}(t_f), t_f) = 0$). The OCP is referred to as a *constrained* OCP because constraints are imposed on the controls, apart from the dynamic equation (2). Some applications include further conditions such as control constraints: $C(\mathbf{x}(t), \mathbf{u}(t)) \le 0$, state constraints: $S(\mathbf{x}(t)) \le 0$, or interior point conditions, although these are not considered in the present paper.

Let H be the Hamiltonian function associated with the problem

$$H(t, \mathbf{x}, \mathbf{u}, \lambda) = F\,(t, \mathbf{x}, \mathbf{u}) + \lambda \cdot f\,(t, \mathbf{x}, \mathbf{u}) \tag{5}$$

where $\lambda = (\lambda_1(t), \ldots, \lambda_n(t)) \in \mathbb{R}^n$ is called the *costate vector*. The classical approach involves the use of PMP [10], which results in a two-point boundary value problem (TPBVP). In order for $\mathbf{u} \in U$ to be optimal, a nontrivial function λ must necessarily exist, such that for almost every $t \in [0, t_f]$

$$\dot{\mathbf{x}} = H_\lambda = f \tag{6}$$

$$\dot{\lambda} = -H_{\mathbf{x}} \tag{7}$$

$$H(t, \mathbf{x}, \mathbf{u}, \lambda) = \min_{\mathbf{v}(t) \in U} H(t, \mathbf{x}, \mathbf{v}, \lambda) \tag{8}$$

$$\mathbf{x}(0) = \mathbf{x}_0; \lambda(t_f) = \mathbf{0} \tag{9}$$

In this paper, we deal with various chemical models whose dynamic equations present a particular structure (we present the two-dimensional case for the sake of simplicity):

$$\min_{u_1(t)} I = \int_0^{t_f} F\,(x_1(t), x_2(t), u_1(t))\,\mathrm{d}t \tag{10}$$

$$\dot{x_1}(t) = f_1\,(x_1(t), x_2(t), u_1(t)) \tag{11}$$

$$\dot{x_2}(t) = f_2\,(x_1(t), x_2(t)) \tag{12}$$

The principal characteristic of this system is the absence of the control u_2 in the Eqs. (10–12). In previous papers [11,12], the authors presented a very simple method that is able to solve, for a given x_2, the problem formed by Eqs. (10, 11). We now adapt

this method to obtain a initial guess for the solution of the system (10–12). The idea consists in constructing x_1 in an approximate and similar way to how it is constructed in [11,12] and in simultaneously constructing x_2 using Euler's (or Euler's improved) method in (12). In the discretization process, the values of x_2 obtained at the prior nodes are used to calculate x_1 at each node, and the values obtained for x_1 are used to calculate x_2. For the sake of simplicity, in what follows we denote x_1 with x, f_1 with f and u_1 with u.

In virtue of PMP and Eq. (7), there exists a piecewise C^1 function λ (costate variable) that satisfies:

$$\dot{\lambda}(t) = -H_x = -F_x - \lambda(t) \cdot f_x \tag{13}$$

From (13), it follows that

$$\lambda(t) = \left[K - \int_0^t F_x e^{\int_0^s f_x \mathrm{d}z} \mathrm{d}s \right] e^{-\int_0^t f_x \mathrm{d}s} \tag{14}$$

denoting $K = \lambda(0)$. From (8), it follows that for each t, $u(t)$ minimizes H. Hence, in accordance with the Kuhn–Tucker Theorem, for each t, there exists two real non negative numbers, α and β, such that $u(t)$ is a critical point of

$$F^*(u) = F + \lambda(t) \cdot f + \alpha \cdot (u_{\min} - u) + \beta \cdot (u - u_{\max}) \tag{15}$$

it being verified that if $u > u_{\min}$, then $\alpha = 0$ and if $u < u_{\max}$, then $\beta = 0$. We hence have $\dot{F^*} = 0$ and the following cases:

Case (1) $u_{\min} < u < u_{\max}$. In this case, $\alpha = \beta = 0$ and hence

$$F_u + \lambda(t) \cdot f_u = 0 \tag{16}$$

It can be seen that from $\dot{x}(t) = f\,(x(t), x_2(t), u(t))$, we easily obtain

$$\frac{F_u}{f_u} = F_{\dot{x}} \tag{17}$$

From (14) and (16), we have

$$K = -F_{\dot{x}} \cdot e^{\int_0^t f_x ds} + \int_0^t F_x \cdot e^{\int_0^s f_x \mathrm{d}z} \mathrm{d}s \tag{18}$$

If we denote $\mathbb{Y}(t)$ the second member of the above equation, the following relation is fulfilled

$$\mathbb{Y}(t) = K \tag{19}$$

We term $\mathbb{Y}(t)$ the *coordination function* and (19) the *coordination equation*.

Case (2) $u = u_{\max}$, then $\beta \geq 0$ and $\alpha = 0$. In this case, by analogous reasoning, we have

$$\mathbb{Y}(t) \geq K \tag{20}$$

Case (3) $u = u_{\min}$, then $\alpha \geq 0$ and $\beta = 0$. In this case, by analogous reasoning, we have

$$\mathbb{Y}(t) \leq K \tag{21}$$

Thus, the problem consists in finding for each K the function that satisfies (3), the condition (19), and from among these functions, the one that satisfies $\lambda(t_f) = 0$. From the computational point of view, the construction can be performed with the same procedure as the simple shooting method, with the use of a discretized version of the coordination equation (18). When the values obtained do not obey the constraints (4), we force the solution to belong to the boundary until the moment that is established by conditions (20, 21). Therefore, the method that we have developed to obtain x_1 is based on the use of an integral form of the Euler equation, combined with the simple shooting method. Besides, we calculate x_2 as described above and hence obtain the initial guess of the problem. In the next section, we shall see the excellent behavior of our approach by means of various examples.

3 Examples

In this section, we present several examples of chemical processes in which the particular structure of the problem allows us to apply the result obtained in the previous section. We analyze three cases. In Example 3.1, we first consider the nonlinear CSTR as being unconstrained. We generalize the previous example in Example 3.2, considering the constrained case with bounded control. Finally, in Example 3.3, we present a nondifferentiable case.

3.1 Unconstrained CSTR

Let us consider the system consisting of the dynamic optimization of a first-order irreversible chemical reaction carried out under non-isothermal conditions in a CSTR, as first formulated by Aris and Amundson [13] and Lapidus and Luus [14].

The equations describing the chemical reactor are:

$$\frac{dx_1}{dt} = -(2+u)(x_1 + 0.25) + (x_2 + 0.5)\exp\left(\frac{25x_1}{x_1+2}\right) \tag{22}$$

$$\frac{dx_2}{dt} = 0.5 - x_2 - (x_2 + 0.5)\exp\left(\frac{25x_1}{x_1+2}\right) \tag{23}$$

The control variable $u(t)$ depends on the opening of the valve and represents the manipulation of the flow-rate of the cooling fluid, which is inserted in the reactor through a coil. Here $x_1(t)$ represents the deviation from the dimensionless steady-state temperature and $x_2(t)$ represents the deviation from dimensionless steady-state concentration. In this section, we consider the case in which the control u is unbounded and the initial conditions $x_1(0) = 0.09$ and $x_2(0) = 0.09$ are used. The optimal control problem is to determine that u in the time interval $0 \leq t < t_f$ which will minimize the quadratic performance index

$$I = \int_0^{t_f} (x_1^2 + x_2^2 + 0.1u^2)\mathrm{d}t \tag{24}$$

subject to the nonlinear dynamic constraints, where the dimensionless final time t_f is specified as 0.78.

Using a control vector iteration procedure based on Pontryagin's maximum principle, Luus and Cormack [15] showed that there exists a local optimum of $I = 0.244425$ and a global optimum of $I = 0.133094$. This optimal control problem provides a good test problem for optimization procedures and is a member of the list of benchmark problems proposed in the handbook of test problems in local and global optimization [16]. It has been used by Luus [17] to evaluate his iterative dynamic programming (IDP) algorithm and by Luus and Galli [18] to examine the multiplicity of solutions. Dynamic programming [19,20] is a powerful method for solving optimization problems, but has a number of drawbacks that limit its use to solving problems of low dimension. To overcome these limitations, Luus suggested using it in an iterative fashion.

Ali et al. [21] solved this problem using eight stochastic global optimization algorithms: control random search, simulated annealing and clustering, the results obtained varying between $I = 0.135$ and $I = 0.245$. The CPU time used was quite long, in some case more than 2, 382.

Lopez-Cruz et al. [22] compare three methods: the first-order gradient method, the Iterative dynamic programming (IDP) and evolutionary algorithms (EA). Their conclusions are most interesting. First, they show that the convergence of the first-order gradient algorithm to the local or global optimum depends on the initial values for the control. Second, to solve this problem by means of IDP, the time interval $[0, t_f]$ was discretized in only $N = 13$ time intervals; the best minimum obtained was $I = 0.1355$ and the CPU time used (measured on a Pentium III, 700 MHz PC) was quite long: 600–800 s. Moreover IDP may still converge to the local optimum. Finally, four variants of EA were implemented and evaluated. Two are based on the Breeder genetic algorithm and two are differential evolution algorithms. The minimum obtained with EA varies from $I = 0.1358$ to $I = 0.1449$ and CPU time varies from 134 to 451 s.

We apply our simple method and present the results below. The minimum value of $I = 0.1334$ was obtained very rapidly. The computation time (measured on a Pentium IV, 3.4 GHz PC) for 15 iterations, with a discretization of 78 subintervals, was 2.56 s. The resulting initial guess for the optimal control policy is given in Fig. 1 and for the state trajectories in Fig. 2.

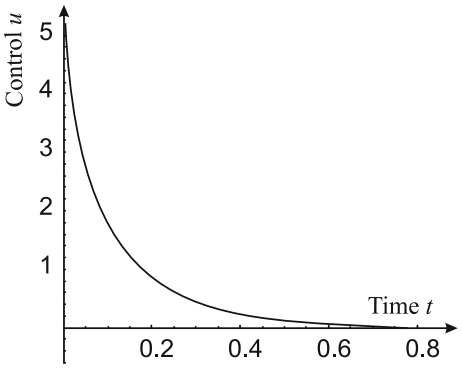


Fig. 1 Optimal control

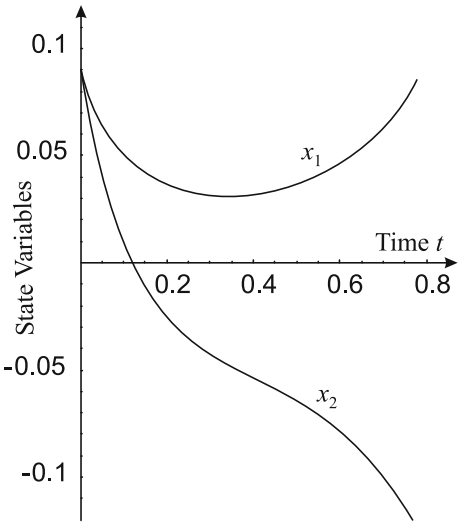


Fig. 2 Trajectories of the state variables

3.2 Constrained CSTR

This example consists of the same dynamic equations and performance index as in Example 3.1

$$I = \int_0^{t_f} (x_1^2 + x_2^2 + 0.1u^2)\mathrm{d}t \tag{25}$$

$$\frac{dx_1}{dt} = -(2+u)(x_1 + 0.25) + (x_2 + 0.5)\exp\left(\frac{25x_1}{x_1+2}\right) \tag{26}$$

$$\frac{dx_2}{dt} = 0.5 - x_2 - (x_2 + 0.5)\exp\left(\frac{25x_1}{x_1+2}\right) \tag{27}$$

though here the initial condition is given (see [14]) by

Fig. 3 Optimal control

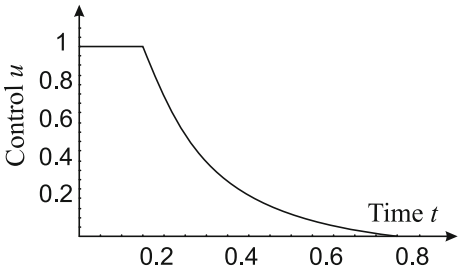
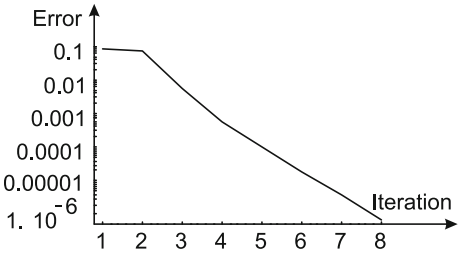


Fig. 4 Convergence profile



$$\mathbf{x}(0) = [\,0.05\ 0\,]^T \tag{28}$$

In [23], the control signal is bounded between upper and lower levels

$$-1 \leq u(t) \leq 1 \tag{29}$$

and using the technique called Integrated System Optimization and Parameter Estimation (ISOPE), Becerra obtained a minimum of $I = 0.028953$.
Our method provides an initial guess of the solution with a minimum value of $I = 0.02945$. The constrained optimal control is shown in Fig. 3. As can be seen in Fig. 4, the algorithm shows rapid convergence. In the example, 8 iterations were found to be sufficient to obtain the prescribed error (10^{-6}) in the equation $\lambda(t_f) = 0$. The variation in relative error in absolute value with iterations is shown in Fig. 4. The computation time for 8 iterations, with a discretization of 78 subintervals, was 1.48 s.

3.3 Nondifferentiable system

We now consider a difficult optimal control problem studied by Thomopoulos and Papadakis [24], who showed difficulties of convergence using several optimization algorithms. The system was also studied by Banga and Seider [25], who proposed a stochastic algorithm: the integrated controlled random search (ICRS), and by Luus [26] using iterative dynamic programming (IDP). The system is described by

$$\frac{dx_1}{dt} = -x_1 - x_2 + u + 100[H(t-0.5) - H(t-0.6)] \tag{30}$$

$$\frac{dx_2}{dt} = x_1 \tag{31}$$

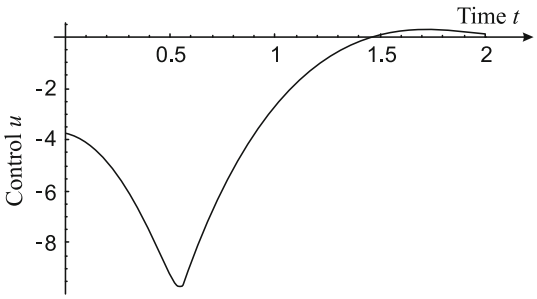


Fig. 5 Optimal control

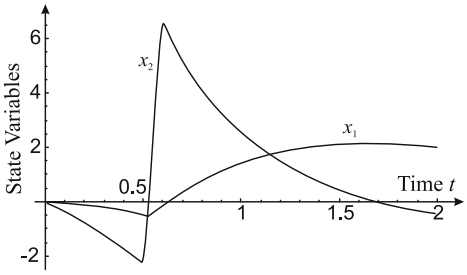


Fig. 6 Trajectories of the state variables

Equation (30) contains a disturbance term in the form of a rectangular pulse of magnitude 100 from $t = 0.5$ until $t = 0.6$. The optimal control problem is to find the control u in the time interval $0 \leq t < t_f$ so that the performance index

$$I = \int_0^{t_f} (5x_1^2 + 2.5x_2^2 + 0.5u^2)\mathrm{d}t \tag{32}$$

is minimized. All these authors consider a final time $t_f = 2$ s and initial conditions: $x_1(0) = 0$; $x_2(0) = 0$.

Using an exponential smoothing technique, [24] overcame the discontinuity problem and obtained a minimum value $I = 58.538$. Using IDP, Luus [26] reported a minimum of $I = 58.18$. Using the ICRS algorithm, Banga [25] found an optimal control with a minimum of 58.13.

In the following figures, we present our approximation to the optimal control (Fig. 5) and to the state trajectories (Fig. 6). The reader can compare each approximation with the values given in the references. Our method provides an initial guess of the solution with a minimum value of $I = 67.67$. The computation time, with a discretization of 1,000 subintervals, was 3.89 s. It can be seen that the rectangular pulse in the time interval from 0.5 to 0.6 causes a very rapid increase in the state variable x_2 during this time interval.

4 Conclusions

The initial estimation of the solution can greatly influence the accuracy of the results obtained from the numerical methods and software packages used to solve dynamic optimization or optimal control problems numerically and, in some cases, can even result in divergence of the process. In the present paper, we use the solution based on PMP theory to provide an initial guess for the solution of chemical processes. An estimate of the optimal solution (with a small value of tolerance) is obtained first using our method. This solution can be then used as the initial guess for another numerical method. This combined method can produce excellent results. Our method presents numerous advantages: it is very easy to programme, the theory allows us to address a wide range of problems: constrained, unconstrained, nondifferentiable, etc., the computation time is very short compared with classical approaches, the initial guess is very close to the solution and is attracted to a global minimum.

References

1. J.T. Betts, http://www.boeing.com/phantom/socs/index.html
2. A.L. Schwartz, E. Polak, Y.Q. Chen, http://www.accesscom.com/~adam/RIOTS/
3. O. von Stryk, http://www.sim.informatik.tu-darmstadt.de/sw/dircol.html.en
4. L.S. Jennings, M.E. Fisher, K.L.Teo, C.J. Goh, http://www.cado.uwa.edu.au/miser/
5. C.A. Schweiger, Ch.A. Floudas, http://titan.princeton.edu/MINOPT/minopt.html
6. M.G. Garcia, E. Balsa-Canto, A.A. Alonso, J.R. Banga, Comput. Aided Chem. Eng. **20**, 121 (2005)
7. Q. Gong, W. Kang, I.M. Ross, IEEE Trans. Auto. Cont. **51**(7), 1115 (2006)
8. A.V. Rao, *User's Manual for GPOCS* (TOMLAB, Gainesville, FL, 2007)
9. M. Cizniar, D. Salhi, M. Fikar, M.A. Latifi, http://www.kirp.chtf.stuba.sk/~fikar/research/dynopt/dynopt.htm
10. A.E. Bryson, *Dynamic Optimization* (Addison-Wesley, New York, 1999)
11. L. Bayon, J.M. Grau, M.M. Ruiz, P.M. Suarez, IMA J. Math. Control I **22**(4), 377 (2005)
12. L. Bayon, J.M. Grau, M.M. Ruiz, P.M. Suarez, Appl. Math. Comput. **184**(1), 12 (2007)
13. R. Aris, N.R. Amundson, Chem. Eng. Sci. **7**, 121 (1958)
14. L. Lapidus, R. Luus, *Optimal Control of Engineering Processes* (Blaisdell, Waltham, Mass, 1967)
15. R. Luus, D.E. Cormack, Can. J. Chem. Eng. **50**, 309 (1972)
16. C.A. Floudas et al., *Handbook of Test Problems in Local and Global Optimization* (Kluwer Academic, The Netherlands, 1996)
17. R. Luus, *Iterative Dynamic Programming* (Chapman & Hall/CRC Press, Boca Raton, FL, 2000)
18. R. Luus, M. Galli, Hung. J. Ind. Chem. **19**, 55 (1991)
19. M.S. Santos, J. Vigo-Aguiar, Appl. Math. Let T **9**(4), 41 (1996)
20. M.S. Santos, J. Vigo-Aguiar, Econometrica **66**(2), 409 (1998)
21. M.M. Ali, C. Storey, A. Torn, J. Optim. Theory Appl. **95**(3), 545 (1997)
22. I.L. Lopez Cruz, L.G. Van Willigenburg, G. Van Straten, Appl. Soft Comput. **3**(2), 97 (2003)
23. V.M. Becerra, Development and applications of novel optimal control algorithms, PhD Thesis. City University, London, UK, 1994
24. S.C.A. Thomopoulos, I.N.M. Papadakis. in *Proceedings of the 1991 American Control Conference*, Boston, MA, pp. 2419–2422, 1991
25. J.R. Banga, W.D. Seider, *Global Optimisation of Chemical Processes Using Stochastic Algorithms* (Princeton University, Princeton, NJ, 1995)
26. R. Luus, Ind. Eng. Chem. Res. **32**, 859 (1993)