COMPUTATIONAL STOCHASTIC DYNAMIC PROGRAMMING PROBLEMS: GROUNDWATER QUALITY REMEDIATION *

Floyd B. Hanson and John J. Westman

Laboratory for Advanced Computing
University of Illinois at Chicago
851 S. Morgan; M/C 249
Chicago, IL 60607-7045
hanson@math.uic.edu

Abstract

The general objective is the development of supercomputing algorithms for the optimal feedback control of larger scale, continuous time, nonlinear, Markov stochastic dynamical systems. The numerical procedures are based on PDE methods such as the finite element or difference methods for stochastic dynamic programming, as well as other advanced numerical methods. The algorithms have been implemented on the Cray vector multiprocessors and massively parallel Connection Machines. These implementations have utilized advanced supercomputing techniques such as parallelization, vectorization and data structures and decompositions. Problems in 5 state space dimensions have been solved. Large dimensions are required by some applications, such as groundwater remediation and resource management. The principal application focus here is the remediation for groundwater quality through the control of pumping policies when the groundwater is subject to uncertain introduction of contaminants.

Keywords: *Stochastic Dynamic Programming, Computational Control, Groundwater Remediation Management Application*

1. Introduction

Parallel and vector computation permit the solution of larger and more realistic problems than would otherwise be possible. The computational demands of stochastic dynamical programming are well known and are called the *Curse of Dimensionality*. The numerical PDE methods have been modified for the nonstandard characteristics of the PDE of stochastic dynamic programming. In order to manage the large computational requirements, high performance supercomputers have been employed [5, 24, 25, 23]. For instance, the problems with up to 5 states and 32 mesh points per state have been successfully solved using finite difference methods [24, 25, 14, 23] on both Cray and Connection Machines. Larger problems are possible with recent hardware and software advances.

The finite element method has computational and memory advantages. This method requires a smaller number of nodes than the corresponding finite difference method of similar accuracy. We have shown [6] that the finite element method not only helps to alleviate Bellman's *Curse of Dimensionality* in dynamical programming computations by permitting the solution of higher dimension problems, but also saving supercomputer storage requirements.

The general aim to develop a fast and efficient parallel computational algorithms and data structures for optimal feedback control of large scale, continuous time, nonlinear, stochastic dynamical systems. Since the finite element procedure requires formulation of the mesh data structures, it is desirable to study the mapping from the problem conceptual structure to the machine configuration for either Cray or Connection Machine computational models [24].

However, the computational treatment of Poisson noise is a particularly unique feature of this research. The numerical approach directly treats the partial differential equation of stochastic dynamic programming. Results give the optimal feedback control variables and the expected optimal performance index in terms of state

*This work has been supported in part by National Science Foundation's Computational Mathematics Program Grant DMS 93–1107, with supercomputer support from Los Alamos National Laboratory's Advanced Computing Laboratory, National Center for Supercomputing Applications, and Pittsburgh Supercomputer Center. This paper is published in *Proceedings of 33th IEEE Conference on Decision and Control* **1**, pp. 455-460, December 1994.

variables and time.

The particular application that will be investigated in this research is a resource management for the remediation of groundwater quality in an underground aquifer. Remediation consists of finding an optimal pumping policy to reduce the effects of uncertain introduction of contaminants to the aquifer. The contamination is assumed to happen as a rare event according to a Poisson process. The optimal policy is to minimize pumping and treatment costs. The optimal control problem is constrained by the dispersion of the hydraulic heads and the diffusion of the contaminant concentration in the groundwater, while Darcy's law governs the fluid velocity of the slow flow through the porous media of the aquifer. The spatial dependence of these partial derivative (distributed parameter) constraints are lumped together by finite element averaging over the slab geometry of the aquifer to help reduce the complexity of the dynamic programming step by restricted only to the state space of the heads and contaminant concentrations.

The model could be further generalized to include a white Gaussian noise (see Gihman and Skorohod [9]for the general theory of Markov processes in continuous time, i.e., for combined Gaussian and Poisson noise, but see [2] and [22] for the special Gaussian examples). This would represent the low level or background fluctuations in the hydraulic heads and contaminant concentrations due to random recharge or leakage, for example. If Gaussian noises were present, additional assumptions would be necessary, for convenience, on the interaction between the Poisson and Gaussian noises. There is far more literature on Gaussian noise than on Poisson noise, undoubtedly due to the fact that the smoothness of the Gaussian or Wiener process makes the analysis more tractable in contrast to the discontinuities of the Poisson process. However, the Poisson process makes a more suitable model for large random fluctuations [11], such as a sudden increase in the contaminant concentration. The Gaussian noise has been omitted in this paper in order to concentrate on the more significant dynamics of the large random fluctuations associated with a Poisson noise model.

2. Computational Stochastic Dynamic Programming

Both computational and memory requirements grow exponentially with the number of states[12], and this is a characteristic of *Grand Challenge* level problems [10]. The methods developed in this research would be useful for many applications in aeronautics, biology, chemistry, computers, engineering, finance, physics and resource models. In a general way, our formulation treats many of the special problems considered as important in the report the NSF sponsored Panel on the *Future Directions in Control Theory* [8], such as computational control, nonlinear control, feedback control and stochastic control systems.

The finite element method is suitable for the parallel computational data structures. Although the finite element method is more computationally intensive, current research has indicated that the Crays and the massively parallel Connection Machines have the potential to permit the cost-effective solutions for large scale problems, by reducing memory demands in favor of increased computational demands.

Our model is a continuous time, stochastic generalization of the discrete time, deterministic optimal control model of Shoemaker and co-workers [3, 4]. The foundations and practice of groundwater models are thoroughly covered by Bear [1]. The generalized model is a state-time stochastic control model that we have been working with, but with more complicated constraints. Thus the spatial finite elements avoid the distributed parameter (PDE) nature of the dynamics.

The optimal control problem is for the minimized treatment and pumping costs,

$$\overline{J}^{*}(x,t) = \min_{U}[\operatorname*{mean}_{P}[J[X,U,t]|X(t)=x, U(t)=u] \quad (1)$$

where the cost objective has the form

$$\begin{aligned} J[X(t),U(t),t] &= \int_t^{t_f} \sum_{i_p=1}^{\hat{m}} U_{i_p}(t) * (a_1 \qquad (2) \\ &+ a_2 * (h_r - h_{i_p}(t)))dt, \end{aligned}$$

$X(t) = \{h(t), c(t)\}$ is the state $\hat{n}$-vector of hydraulic heads and contaminant concentrations, $U(t) = q(t)$ is the control $\hat{m}$-vector of pumping rates and the pumps are at nodes i_p. The ground level reference head is h_r. The constants $\{a_1, a_2\}$ are the treatment and pumping costs coefficients, averaged over spatial coordinates.

The finite element averaged dynamical constraint is the stochastic differential equation

$$\begin{aligned} dX(t) &= E^{-1}[X(t),U(t)] * (A[X(t)] + B[U(t)] \quad (3) \\ &+ S[t])dt, \end{aligned}$$

where E, A, B are finite element matrices depending on $\{X(t), U(t)\}$, and

$$S[t]dt = S_0(t)dt + H(t)dP(t), \quad (4)$$

with $P(t)$ being the Poisson noise that models the uncertain introduction of the contaminant with amplitude $H(t)$.

Since the vector state and control have domain constraints, the pollution concentrations are constrained by

$$c_j(t_f) \le C_{\max}$$

on observation (wells) nodes at the final time, t_f, while the pumping control rates are constrained by

$$U(t) \geq 0 \text{ for all } t \in (0, t_f),$$

and

$$|\sum_{i=1}^{m} U_i(t)| \leq U_{\max} \text{ for all } t \in (0, t_f).$$

These constraints ensure that the final water contaminant concentrations satisfy government groundwater quality regulations, that the pumping rates are positive, and that the sum of the individual pumping rates are bounded, respectively.

The state vector consists of the Galerkin finite element spatial nodal values relative to the finite element basis or shape functions $\{w_i(\xi, \eta) | i = 1, \hat{n}\}$, so that the spatially dependent state vector has the approximation

$$\begin{aligned} \overline{X}(\xi, \eta, t) &\equiv \{\overline{h}, \overline{c}\}(\xi, \eta, t) \\ &\approx \widehat{X}(\xi, \eta, t) \equiv \sum_{i=1}^{\hat{n}} X_i(t) * w_i(\xi, \eta) \end{aligned} \tag{5}$$

at 2-dimensional spatial point (ξ, η) in the aquifer, where $\hat{n}$ is the number of spatial nodes. It is assumed that the problem is averaged over the vertical coordinate $z = \zeta$ value (alternatively, this can also be thought of the *essentially horizontal aquifer* assumption of Bear [1]). The flow transport for the vertically averaged head $\overline{h}(\xi, \eta, t)$, due to the continuity of mass, satisfies the dispersion (diffusion) equation

$$bS_s \frac{\partial \overline{h}}{\partial t} = \nabla \cdot (T \nabla \overline{h}) + \sum_{i_p=1}^{\hat{m}} \overline{q}_{i_p} \cdot \delta(\xi - \xi_{i_p}, \eta - \eta_{i_p}), \tag{6}$$

where b is the aquifer thickness, S_s is the storativity of the aquifer, T is the transmissivity of the aquifer, $\overline{q}_i$ is the pumping rate for the i_pth pump located at the pump node (ξ_{i_p}, η_{i_p}). For low speed flow, the momentum balance can be reduced to Darcy's law [1],

$$\overline{v} = -K \cdot \nabla \overline{h}, \tag{7}$$

for Darcy's velocity $\overline{v}$, where K is the hydraulic conductivity ten

The contaminant transport for the vertically averaged concentration $\overline{c}(\xi, \eta, t)$ satisfies the dispersion-diffusion equation

$$\begin{aligned} \theta b R_d \frac{\partial \overline{c}}{\partial t} &= \nabla \cdot (b \widetilde{D} \nabla \overline{c}) - b \overline{v} \cdot \nabla (\overline{c}) \\ &+ \widetilde{H}(t) P'(t) \cdot \delta(\xi - \xi_{j_c}, \eta - \eta_{j_c}), \end{aligned} \tag{8}$$

where θ is the effective porosity of the aquifer, R_d is the retardation coefficient, and $\widetilde{D}$ is the hydrodynamic dispersion tensor includes both convective dispersion and molecular diffusion. The polluting contaminant enters the aquifer according to the rare events of a Poisson process $P'(t)$ with amplitude $\widetilde{H}(t)$ at node j_p. The goal is to remediate the effects of the contaminant once it is observed by means of observation wells.

Initially, the state variables for the head and concentration are given:

$$\overline{h}(\xi, \eta, 0) = \overline{h}_0(\xi, \eta), \quad \overline{c}(\xi, \eta, 0) = \overline{c}_0(\xi, \eta). \tag{9}$$

The state boundary conditions are assumed to be of the general linear type

$$\begin{aligned} \alpha_h e_n \cdot \nabla \overline{h} &= \beta_h (\overline{h} - \overline{h}_b) + \gamma_h, \\ \alpha_c e_n \cdot \nabla \overline{c} &= \beta_c (\overline{c} - \overline{c}_b) + \gamma_c, \end{aligned}$$

for all boundary points $(\xi, \eta) = (\xi_b, \eta_b)$, where e_n is the unit normal to the boundary, the $\{\alpha_j, \beta_j, \gamma_j\}$ are constants for $j = h$ or c, and the $\{\overline{h}_b, \overline{c}_b\}$ are boundary values for the state variables. Application of finite elements to the state dispersion-diffusion equations for $\{\overline{h}, \overline{c}\}$, and Darcy's velocity for $\overline{v}$, leads to the above stochastic differential equation (3) for the state node vector $X(t)$ above with the coefficient matrices $\{E, A, B\}$.

Applying the *Principle of Optimality* to the optimal stochastic control problem leads to the partial differential equation of Stochastic Dynamic Programming, which is the actual target of our computational methods,

$$\begin{aligned} 0 &= \frac{\partial \overline{J}^*}{\partial t}(x, t) + \lambda [\overline{J}^*(x + H(x, t), t) \\ &- \overline{J}^*(x, t)] + \widetilde{S}(x, t), \end{aligned} \tag{10}$$

where $\overline{J}^* = \overline{J}^*(x, t)$ is the optimal, expected value of the objective cost functional, $x = [x_i]_{\hat{n} \times 1}$ is the state vector, and $u = [u_i(x, t)]_{\hat{m} \times 1}$ is the corresponding feedback control vector that is the output argument of the optimal control term $\widetilde{S}$. Of particular interest are the larger state spaces, when the state space has more than several dimensions. Note that the state variable x plays the role of a space variable in the partial differential equation.

$H(t) = [H_i(t)]_{\hat{n} \times 1}$ is the vector of the Poisson noise amplitude coefficient and λ is the rate of Poisson jumps. Here, Poisson noise is used for modeling the behavior of the underlying dynamical system in an uncertain aquifer environment with large random fluctuations in pollution levels.

The essential part, and usually nonlinear part, of the control problem is contained in the optimal control switching term

$$\widetilde{S}(x, t) = \min_u [\widetilde{C}(x, t, u) + F^T(x, t, u) \nabla \overline{J}^*(x, t)] , \tag{11}$$

which can also contribute nonsmoothness, when a control switches to a constraint. Here $F(x, t, u) = [F_i]_{\hat{n} \times 1} =$

$E^{-1}[x,u](A[x]+B[u]+S_0[t])$ is the deterministic nonlinearity vector and $\widetilde{C}(x,t,u)=\sum_{i_p} u_{i_p}\cdot(a_1+a_2\cdot(h_r-h_{i_p}))$ is the The nonlinearity in the cost function means that the switching term will in general be nonlinear in the solution gradient

$$\nabla\overline{J}^*(x,t)=\left(\frac{\partial\overline{J}^*}{\partial x_1},\frac{\partial\overline{J}^*}{\partial x_2},\cdots,\frac{\partial\overline{J}^*}{\partial x_{\hat{n}}}\right)(x,t). \tag{12}$$

The Bellman Equation of stochastic dynamic programming follows from a general form of the *Itô* chain rule for continuous time Markov processes (see Gihman and Skorohod [9]).

For the finite element computational approach, let $\Phi(x)=[\phi_i]_{M\times 1}$ denote a set of M Galerkin basis functions for the state space. The Galerkin approximation in state space is

$$\overline{J}^*(x,t)\ \approx\ \sum_{j=1}^{M} V_j(t)\cdot\phi_j(x). \tag{13}$$

We define the inner product,

$$(\psi,\phi)=\int_D \psi(x,\cdot)\cdot\phi(x,\cdot)dx$$

for the bounded state domain D. The variational formulation of the Bellman functional PDE is then given as

$$\begin{aligned} 0 \ &= \ \int_D \psi(x)\cdot\left[\frac{\partial\overline{J}^*}{\partial t}(x,t)\right. &(14)\\ &+\ \left.\lambda\cdot[\overline{J}^*(x+H(x,t),t)-\overline{J}^*(x,t)]+\widetilde{S}(x,t)\right]dx, \end{aligned}$$

with respect to the test function $\psi(x)$. Direct substitution of the discrete time Galerkin approximation for $\overline{J}^*(x,t)$, as well as using the M Galerkin basis functions for ψ, leads to the form

$$\begin{aligned} 0 \ &=\sum_{j=1}^{M}\frac{dV_j}{dt}(t)\int_D\phi_i\phi_j dx+\int_D\phi_i S^{(M)}(x,t)dx \quad (15)\\ &+\sum_{j=1}^{M}V_j(t)\lambda\int_D\phi_i\left[\phi_j(x+H(x,t_k),t_k)-\phi_j(x,t_k)\right]dx, \end{aligned}$$

for $i=1,2,\cdots,M$, where

$$\begin{aligned} \widetilde{S}(x,t)\ &\approx\ S^{(M)}(x,t)=\min_u[\widetilde{C}(x,t,u) &(16)\\ &+\ \sum_{j=1}^{M}V_j(t)F^T(x,t,u)\nabla\phi_j(x)]\ , \end{aligned}$$

which can be solved explicitly for quadratic terms in the case of quadratic costs [13, 20, 21]. Use of a multilinear product basis over hyper-rectangular elements should lead to local state space integral errors on the order of $O(||\Delta x||^{\hat{n}+2})$ for an $\hat{n}$-dimensional state space.

Mathur and Johnsson [19] present a data parallel algorithm for the finite element method on the Connection Machine CM-2 using virtual processors to represent an unassembled nodal points. The performance for the elemental stiffness matrix computation is in the range 1.6-1.9 GFLOPS for the CM-2 and they recently gave a modification for the CM-5 [15]. Dantas de Melo and coworkers [7] use Cray vectorization and multitasking techniques on discrete dynamic programming control problems.

The extrapolated, predictor-corrector, Crank-Nicholson method is applied to above equations [13]. We have examined the application of finite element method to the stochastic optimal control problem in lower dimensions such as $\hat{n}=1$ and $\hat{n}=2$ on NCSA Cray-2 supercomputing system. A general and robust technique for time-space mesh selection gives promise of robust convergence to the numerical scheme [20, 21]. Our current results indicate that the finite element method is stable enough for the hyperbolic form (when the Gaussian noise is zero). The finite element method requires less mesh points than the finite difference method, permitting the solution of larger problems [6].

In stochastic optimal control problems, most computations can be classified into the operations between large scale arrays of same size, operations between large scale arrays of different sizes, as well as operations between one large scale array and another small scale vector. The decomposition technique and the broadcasting technique help to modify these operations, so that they are executed only with the same size objects and are consequently more efficient for data parallel computations. In the finite difference method, we have attained good performance in both timing and storage requirements by using these techniques. We also expect to obtain the good results in the finite element applications.

Another avenue of research is the search for *purely parallel algorithms* for the stochastic dynamic programming problems in continuous time. We have taken an interest in the work of Shoemaker and co-workers [17, 3, 4] on parallelizing *differential (discrete) dynamic programming* problems. Another possibly fruitful approach is the research of Luus [18] on continuous time deterministic dynamic programming problems, where it has been possible to reduce dynamic programming dimensionality demands. Kushner and co-workers [16] have another alternative approach using the *Markov Chain Approximation*. Several researchers in the area of computational stochastic control have expressed an interest in a comparison of the above methods.

3. Massively Parallel Computational Results

The authors plan to have results on the new generation Connection Machine CM-5 for the presentation of this paper, when preliminary tests of CM-5 code conversions have been made. The stochastic programming code has been well developed for resource models on the earlier Connection Machines CM-2 and CM-200 [24, 25, 14, 23].

4. Conclusions

In summary, the model for advanced computation of large scale groundwater remediation models has been developed as a stochastic optimal control problem in continuous time, which is approached through stochastic dynamic programming. Techniques such a predictor-corrector methods are used to handle the nonlinear cost terms and robust mesh selection methods permit corrector convergence. These methods should prove useful for large scale groundwater remediation applications, and subsequently as a test for other *grand and national challenge* applications.

5. References

[1] J. Bear and A. Verruijt, **Modeling Groundwater Flow and Pollution**, Kluwer Academic Publishers, Boston, 1987.

[2] A. E. Bryson and Y. Ho, **Applied Optimal Control**, Ginn, Waltham, 1969.

[3] L.-C. Chang, C. A. Shoemaker, and P. L.-F. Liu, *Optimal time-varying pumping rates for groundwater remediation: Application of a constrained optimal control algorithm*, **Water Resources Research**, vol. 28, No. 12, pp. 3157-3173, Dec. 1992.

[4] T. Culver and C. A. Shoemaker *Optimal control for groundwater remediation by differential dynamic programming with quasi-Newton approximations*, **Water Resources Research**, vol. 29, No. 4, pp. 823-831, April 1993.

[5] S.-L. Chung and F. B. Hanson, *Parallel optimizations for computational stochastic dynamic programming*, in **Proc. 1990 Int. Conf. on Parallel Processing, vol. III Algorithms and Applications**, P.-C. Yew, Editor, pp. 254-260, Aug. 1990.

[6] S.-L. Chung, F. B. Hanson and H. H. Xu, *Parallel stochastic dynamic programming: Finite element methods*, **Lin. Alg. Applic.**, vol. 172, pp. 197-218, July 1992.

[7] J. Dantas de Melo, J. L. Calvet and J. M. Garcia, *Vectorization and multitasking of dynamic programming in control: Experiments on a Cray-2*, **Parallel Computing, vol. 13**, 261-269, 1990.

[8] **Future Directions in Control Theory: A Mathematical Perspective**, W. H. Fleming, Chairman, Society for Industrial and Applied Mathematics, Philadelphia, 1988.

[9] I. I. Gihman and A. V. Skorohod, **Controlled Stochastic Processes**, Springer-Verlag, New York, 1979.

[10] **Grand Challenges: High Performance Computing and Communications**, Report by the Committee on Physical, Mathematical, and Engineering Sciences, Federal Coordinating Council for Science, Engineering, and Technology, c/o National Science Foundation, CISE, Washington, DC, 1991.

[11] F. B. Hanson, *Bioeconomic model of the Lake Michigan alewife fishery*, **Can. J. Fish. Aquat. Sci.**, vol. 44, suppl. 2, pp. 298-305, 1987.

[12] F. B. Hanson, *Stochastic dynamic programming: Advanced computing constructs*, **Proc. 28th IEEE Conf. on Decision and Control, vol. I**, pp. 901-903, Dec. 1989.

[13] F. B. Hanson, *Computational dynamic programming on a vector multiprocessor*, **IEEE Trans. Automatic Control 36(4)**, pp. 507-511, Apr. 1991.

[14] F. B. Hanson, D. J. Jarvis and H. H. Xu, *Applications of FORALL-Formed Computations in Large Scale Stochastic Dynamic Programming*, **Proc. Scalable High Perf. Comp. Conf.: SHPCC-92**, IEEE Computer Society, pp. 182-185, April 1992.

[15] Z. Johan, K. K. Mathur, S. L. Johnsson, *An efficient communication strategy for finite element methods on the connection machine CM-5 system*, **TMC Technical Report No. 256**, (file tmc-256.300dpi.ps by anonymous ftp tmc.com) August 6, 1993.

[16] H. J. Kushner and P. G. Dupuis, **Numerical Methods for Stochastic Control in Continuous**, Springer-Verlag, New York, 1992.

[17] L.-Z. Liao and C. A. Shoemaker, *Convergence in unconstrained discrete-time differential dynamic programming*, **IEEE Trans. Automatic Control**, vol. 36(6), pp. 692-706, June 1991.

[18] R. Luus, *Application of dynamic programming to higher-dimensional non-linear optimal control problems*, **Int. J. Control**, vol. 52(1), pp. 239-250, July 1990.

[19] K. K. Mathur and S. L. Johnsson, *The finite element method on a data parallel computing system*, **Int. J. High Speed Computing, vol. 1, No.1**, pp. 29-44, 1989.

[20] K. Naimipour and F. B. Hanson, *Convergence of a numerical method for the Bellman equation of stochastic optimal control with quadratic costs and constrained control*, **Dynamics and Control**, vol. 3(3), pp.237-259, July 1993.

[21] F. B. Hanson and K. Naimipour, *Convergence of numerical method for multistate stochastic dynamic programming*, **Proc. Int. Federation of Automatic Control 12th World Congress**, vol. 9, pp. 501-504, July 1993.

[22] R. F. Stengel, **Stochastic Optimal Control**, Wiley, New York, 1986.

[23] H. H. Xu, D. J. Jarvis and F. B. Hanson, *Parallel Data Vault Methods for Larger Scale Stochastic Dynamic Programming*, in **Proc. 1992 Amer. Control Conf.**, vol. 1, pp. 142-146, June 1992.

[24] H. H. Xu, F. Hanson and S. L. Chung, *Optimal data parallel methods for the stochastic dynamical programming*, in **Proc. 1991 Int. Conf. Parallel Processing, vol. III: Algorithms & Applications**, K. So, Editor, pp. 142-146, Aug. 1991.

[25] H. H. Xu, F. Hanson and S. L. Chung, *Parallel solutions of dimensionality problems for the stochastic dynamical programming*, in **Proc. 30th IEEE Conf. Decision and Control**, vol. 2, pp. 1717-1723, Dec. 1991.