Stochastic Control: Hybrid Systems, Switching Diffusions, and Simulation-Based Methods

Steve Marcus
University of Maryland, College Park

November 1, 2009

Switching Diffusions

- Interesting and useful class of hybrid systems
  - Continuous diffusion component, discrete jump component
- Joint work with M. Ghosh and A. Arapostathis
  - *SIAM Journal on Control and Optimization*, **31**, September 1993.
  - *IEEE Trans. Automatic Control*, **40**, November 1995.
  - *SIAM J. Control and Optimization*, **35**, 1997, 1952-1988.
- Much work in hybrid stochastic systems builds on and adds to this work (later in the talk)
- Recent work on simulation-based methods
  - H.S. Chang, M.C. Fu, J. Hu, and S.I. Marcus, *Simulation-based Algorithms for Markov Decision Processes*, Springer-Verlag, 2007.
  -

Summary of Results

- Well-posedness and Markov properties (under general assumptions) by writing as SDEs driven by Brownian motion process and Poisson random measure
  - Deterministic systems: problems with well-posedness and singularities at boundaries – noise "smooths" these effects
- Occupation measures => optimize in space of measures => linear programming approach
- Compactness, convexity, extremal points
- Optimality of Markov (non-randomized) control law
  - In a larger class of control laws
- Dynamic programming (HJB) equation

A Simplified Example

- One machine producing a single commodity
- Demand = d > 0
- S(t) takes values in {0,1}
- 0: down; 1: functional. Generator:

$$\begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & \lambda_1 \end{bmatrix}$$

- The inventory equation is

$$dX(t) = (u(t) - d)\,dt + \sigma\, dW(t)$$

- Production constraint: $u(t) = 0, S(t) = 0; u(t) \in [0, R], S(t) = 1$
- The cost c(X) is convex, Lipschitz and asymptotically unbounded.

Example: Switching LQG

- Let *S(t)* be a (continuous time) Markov chain taking values in S = {1,2,…,N} with generator $\Lambda = [\lambda_{ij}]$ such that $\lambda_{ij} > 0, \quad i \neq j$
- Let *X(t)* be given by

$$dX(t) = [A(S(t))X(t) + B(S(t))u(t)]dt + \sigma(S(t))dW(t)$$

- The instantaneous cost function *c(x,i,u)* is given by

$$c(x,i,u) = C(i)x^2 + D(i)u^2,$$

where $C(i) > 0$, $D(i) > 0$ for each *i*.

- The cost is

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T [C(S(t))X^2(t) + D(S(t))u^2(t)]dt$$

Mathematical Model

S = {1,2,…,N}, U: compact

$$(X(t), S(t)) \in \Re^d \times S$$

$$dX(t) = b(X(t), S(t), u(t))dt + \sigma(X(t), S(t))dW(t)$$

$$P(S(t+\delta t) = j \mid S(t) = i, X(s), S(s), u(s), s \le t)$$

$$= \lambda_{ij}(X(t), u(t))\delta t + O(\delta t), i \ne j,$$

$$\lambda_{ij} \ge 0, i \ne j, \sum_j \lambda_{ij} = 0.$$

Admissible control: $u(.)$ is U-valued nonanticipative process

Markov control: $u(\mathrm{t}) = \nu(X(\mathrm{t}), S(\mathrm{t}))$

Relaxed control: $u(.)$ is P(U)-valued

Mathematical Model (cont.)

- A function $h : \Re^d \times S \times U \times \Re \to \Re$ can be defined so that "*S(t)* has generator $[\lambda_{ij}]$" in the switching diffusion process (*X*(.),*S*(.)):

$$dX(t) = b(X(t), S(t), u(t))dt + \sigma(X(t), S(t))dW(t);$$

$$dS(t) = \int_{\Re} h(X(t), S(t-), v(t), z)\, p(dt, dz)$$

for $t \geq 0$ with $X(0) = X_0, S(0) = S_0$

- $W(.) = [W_1(.), \ldots, W_d(.)]^T$ is a standard Wiener process
- *p* (*dt*,*dz*) is a Poisson random measure independent of *W*(.) with intensity $dt \times m(dz)$, where *m* is Lebesgue measure
- *p* (.,.), *W*(.), X_0 and S_0 are independent
- *u* (.) is a *U*-valued "nonanticipative" process

Markovian Properties

Define

$$L^{\mathrm{u}} f(x,i) = L_i^u f(x,i) + \sum_{j=i}^{N} \lambda_{ij}(x,u) f(x, j)$$

where

$$L_i^u f(x,i) = \frac{1}{2}\sum_{j,k=1}^{d} a_{jk}(x,i)\frac{\partial^2 f(x,i)}{\partial x_j \partial x_k} + \sum_{j=1}^{d} b_j(x,i,u)\frac{\partial f(x,i)}{\partial x_j}$$

$$a_{jk}(x,i) = \sum_{l=1}^{d} \sigma_{jl}(x,i)\sigma_{kl}(x,i)$$

- **Theorem**: Under a Markov policy u, SDE admits an a.s. unique strong solution such that (X(.), S(.)) is Feller process w/ gen. L^{u}.

Cost Models

- **Cost function**

$$c : \Re^d \times S \times U \rightarrow \Re$$

- **Discounted Cost:** $\alpha > 0$

$$J_\alpha(u(.), x, i) = E_{x,i}^{u(.)} \int_0^\infty \int_U e^{-\alpha t} c(X(t), S(t), y) u(t)(dy) dt.$$

$$J_\alpha(x, i) = \inf_{u(.)} J_\alpha(u(.), x, i)$$

- **Average Cost**

$$J(u(.), x) = \limsup_{T \to \infty} \frac{1}{T} \int_0^T \int_U c(X(t), S(t), y) u(t)(dy) dt$$

- **Objective**: To find a Markov control which is optimal

A Linear Programming Approach

- **Discounted occupation measure**

For u(.) relaxed control, define $\nu_\alpha[u] \in P(\Re^d \times S \times U)$ by

$$\sum_i \int_{\Re^d \times U} f(x,i,y)\nu_\alpha[u](dx,\{i\},dy) = \alpha E^u_{x,i} \int_0^\infty e^{-\alpha t} \int_U f(X(t),S(t),y)u(t)(dy)dt$$

$$M_1 = \{\nu_\alpha[u] : u(.) \text{ is relaxed control}\}$$

$$M_2 = \{\nu_\alpha[u] : u(.) \text{ is Markov relaxed control}\}$$

$$M_3 = \{\nu_\alpha[u] : u(.) \text{ is Markov control}\}$$

$$J_\alpha(u(.),x,i) = \alpha^{-1} \sum_i \int_{\Re^d \times U} c(x,i,y)\nu_\alpha[u](dx,\{i\},dy)$$

--linear over M_1

Results

- **Theorem:** $M_1 = M_2$. M_2 is compact and convex and $M_2^e \subset M_3$. $M_2^e =$ the set of extreme points of M_2.

- **Theorem:** There exists a Markov control v which is discounted cost optimal for any initial condition.

Hamilton-Jacobi-Bellman (DP) Equations

For $u \in U$, let

$$L^u f(x,i) = L_i^u f(x,i) + \sum_{j=1}^{N} \lambda_{ij}(x,u) f(x,j)$$

$$L_i^u f(x,i) = \frac{1}{2} \sum_{j,k=1}^{d} a_{jk}(x,i) \frac{\partial^2 f(x,i)}{\partial x_j \partial x_k} + \sum_{j=1}^{d} b_j(x,i,u) \frac{\partial f(x,i)}{\partial x_j}$$

- HJB equation for DC problem is

$$\inf_{u \in U} [L^u \phi(x,i) + c(x,i,u)] = \alpha \phi(x,i) \qquad (*)$$

- **Theorem:** The DC value fct. $V_\alpha(x,i)$ is the unique solution of (*). A Markov control *v* is DC optimal if and only if it realizes the pointwise infimum in (*).

Manufacturing Example

α-discounted HJB Equations

$$\begin{bmatrix} \frac{\sigma^2}{2} V''_\alpha(x,0) \; - \; dV'_\alpha(x,0) \\ \frac{\sigma^2}{2} V''_\alpha(x,1) \; + \; \min_{u\in[0,R]} \{(u-d)V'_\alpha(x,i)\} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} c(x)$$

$$= \begin{bmatrix} \alpha+\lambda_0 & -\lambda_0 \\ -\lambda_1 & \alpha+\lambda_1 \end{bmatrix} \begin{bmatrix} V_\alpha(x,0) \\ V_\alpha(x,1) \end{bmatrix}.$$

$V_\alpha(x,i)$ is convex in x for each i. Hence $\exists x^*$ such that

$$\begin{aligned} V'_\alpha(x,1) &\le 0 \text{ for } x \le x^* \\ V'_\alpha(x,1) &\ge 0 \text{ for } x \ge x^*. \end{aligned}$$

Thus an optimal control:

$$v(x,0) = 0, \; v(x,1) \; = \begin{cases} R & \text{if } x < x^* \\ d & \text{if } x = x^* \\ 0 & \text{if } x > x^*. \end{cases}$$

Note: No singular situation to the presence of noise.

Average Cost Problem

- To minimize pathwise (long-run) average cost

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T c(X(s),S(s),u(X(s),S(s)))ds$$

- Need stability (positive recurrence or ergodicity)
  - Difficult: interaction of discrete & continuous components
  - If, for each i, the diffusion is positive recurrent and the parametric Markov chain is ergodic, the hybrid system is not necessarily stable
    - Switching between two positive recurrent processes can result in a process that isn't

Average Cost Problem

- Mathematics are *much* more complicated, but can prove similar results under appropriate conditions
  - Optimality of stationary Markov nonrandomized control law
  - Hamilton-Jacobi-Bellman (dynamic programming) equations
  - Stability of the optimal control law if there is some stable Markov nonrandomized control law

More Recent Stochastic Hybrid Systems Models

- Many models, some simpler, some generalizations
- Lygeros, Sastry, Pappas, Ghosh, Bagchi, Koutsokos
- Simplifications
  - Piecewise deterministic systems
  - Jump linear stochastic systems
- Generalizations
  - Resets (controlled and uncontrolled)
  - Jumps in continuous state

Composition, Computation, Model Checking

- Koutsokos (2008)
  - Verification of reachability
  - Lygeros GSHS models
  - Kushner finite state Markov chain approximations
- Julius & Pappas (2009)
  - Approximation & verification of stochastic hybrid systems
  - Focus on jump linear stochastic systems
  - Uses approximate bisimulation (Girard and Pappas)

Simulation-Based Methods for Markov Decision Processes

- Motivation
  - Unknown random transitions/costs and/or much easier to simulate than to build MDP model

- Examples: capacity expansion in semiconductor fab, "transitions" involve complex simulation of entire fab; biological systems

DP Notation

- state x, action a
- reward R(x,a)
- value function V(x), Q-function Q(x,a)
- discount factor γ
- policy π

- Goal: maximize $\Sigma_t \gamma^t R(x_t, \pi_t(x_t))$

Simulation-Based Setting

- setting:
  - transition probabilities not explicitly known, but can be easily simulated;
  - finite horizon
- targeted at problems with
  - huge state spaces
  - limited simulation budget
  - Goal: estimate optimal value function efficiently (simulation-based value iteration)

- ADAPTIVE SAMPLING: multi-armed bandit models to decide which actions to sample

Main Ideas

- Value function estimated based on simulated trees
- Objective: which action to sample next (simulate to generate next sampled state)
- Trade off between exploitation and exploration: choose action that maximizes

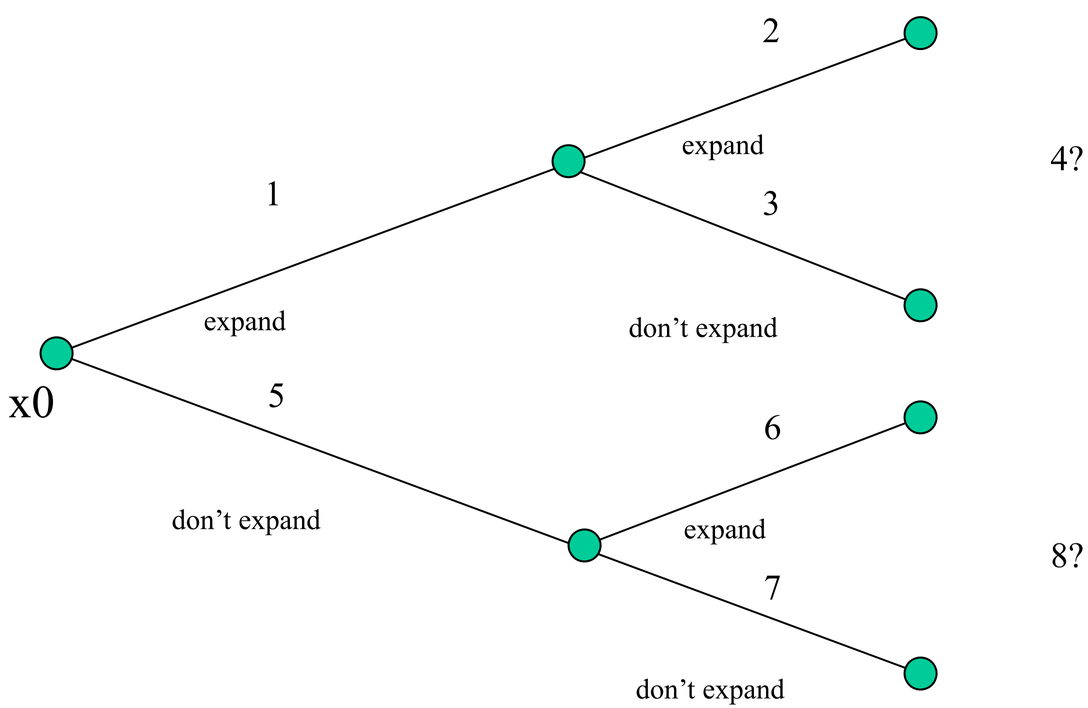
$$\hat{Q}_i(x,a) + k\sqrt{\frac{\ln n_i}{n_{a,i}^x}}$$

stage i samples thus far (total, state/action specific)

Simple Illustrative Example

Simulated Tree for two stages, two actions

samples per state in stage: $N_1=2$, $N_2=3$



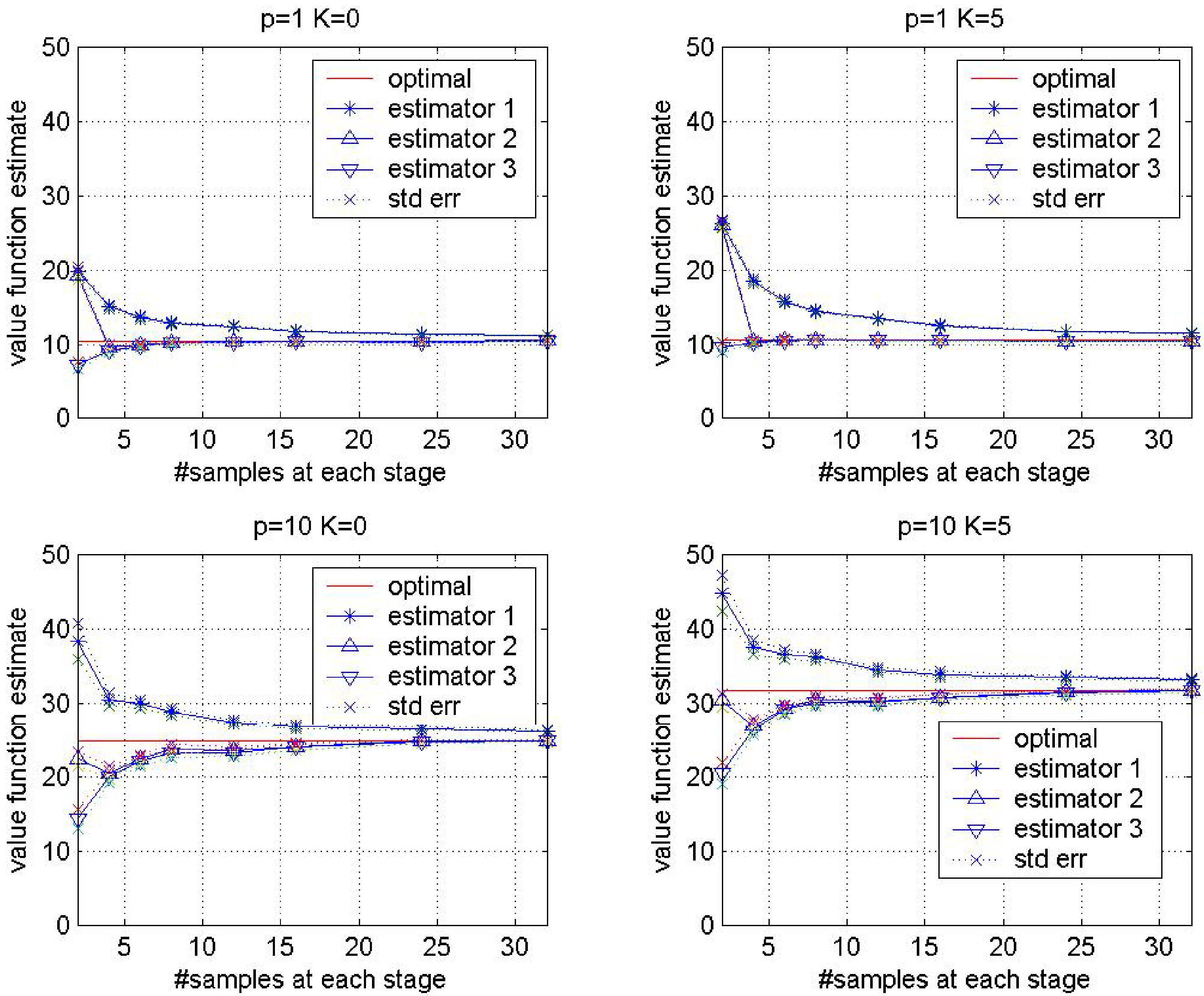
$$\hat{Q}_i(x,a) + k\sqrt{\frac{\ln n_i}{n_{a,i}^x}}$$

Nodes represent simulated state reached from simulation, numbers indicate sequence of simulations carried out

Results

- provable convergence with bounded rate (bias)
- complexity $O(N^H)$ (N = total # simulations)
  vs. backwards induction $O(H|A||X|^2)$
  - independent of size of state space X (action space A)
  - exponential in horizon length H

Inventory Control Example



Future Directions

- With Rance Cleaveland (stochastic hybrid systems)
  - Composition
  - Control, approximate bisimulation, model checking
    - Special case of deterministic discrete controller, continuous stochastic plant
- With Ed Clarke, Sumit Jha, et. al?
  - Statistical model checking with nondeterministic processes, Markov decision processes, hybrid stochastic systems
- Which models important in CMACS applications?