

Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation

Floyd B. Hanson
University of Illinois
Chicago, Illinois, USA

Chapter 12 Applications in Mathematical Biology

Copyright © 2006 by the Society for Industrial and Applied Mathematics.

May 24, 2006

Chapter 12

Applications in Mathematical Biology and Medicine

Despite assertions in both the lay and the professional literature, it is now known that normal physiology is anything but “regular.” ...

Loss of event-to-event variability occurs during normal aging and also occurs pathologically in critical illness.

—Dr. Timothy G. Buchman (2004) [45].

*Mathematics Is Biology’s Next Microscope, Only Better;
Biology Is Mathematics’ Next Physics, Only Better*

—Joel E. Cohen (2004) [58].

The application to optimal harvesting in uncertain environments is made in the presence of both background Gaussian noise and catastrophic jump events. Many problems in nature exhibit random effects and undergo catastrophic changes for which the stochastic calculus of continuous Wiener processes alone is inadequate.

12.1 Stochastic Bioeconomics: Optimal Harvesting Applications

For deterministic problems of optimal harvesting of renewable resources, the seminal reference by C. W. Clark is *Mathematical Bioeconomics: The Optimal Management of Renewable Resources* [56]. The book is nicely self-contained with introduction to the necessary economics, calculus of variations and optimal control theory. An excellent survey of stochastic bio-economics is given by Anderson and Sutinen in [9].

In this chapter, examples of optimal harvesting problems in random environments are illustrated. The first application is optimal harvesting with random population fluctuations [238]. A second application is optimal harvesting with random population fluctuations, but also with price fluctuations [115], so is a two-

dimensional state generalization of the first application.

12.1.1 Optimal Harvesting of Logistically Growing Population Undergoing Random Jumps

This problem of natural logistic growth of a renewable resource subject to random disasters and bonanzas was treated by Ryan and Hanson [238]. The parameter data was motivated by the **boom and bust** characteristics of Antarctic pelagic whaling at the time as studied by Clark and Lamberson [57]. The problem is summarized in the notation of this book, so for more information the reader should refer to [238].

Let $X(t)$ be the amount of biomass (mass of the biological species) of the harvested species at time t with stochastic dynamics given by

$$dX(t) = X(t)(r(1 - X(t)/K) - qU(t)) dt + X(t) \sum_{i=1}^{n_p} \nu_i dP_i(t), \quad (12.1)$$

where $X(0) = x_0 > 0$ is the initial biomass, $r > 0$ is the constant intrinsic growth rate and $K > 0$ is the constant biomass carrying capacity that reflects the size of the population that the environment can support in absence of harvesting and other factors. Hence, the natural growth function $f(x) = rx(1 - x/K)$ is called the logistic function since as $x \rightarrow K$ a saturation effect due to crowding limits growth. Under the assumption of linear harvesting, the rate of harvesting is $H(t) = h(X(t), U(t)) = qU(t)X(t)$, where $U(t) \geq 0$ is the harvesting effort or rate and also the control variable, while $q > 0$ is called the **catchability coefficient** and is a measure of the efficiency of the harvest. The population suffers from rare random jumps from various sources for $i = 1:n_p$ linear in the biomass $X(t)$ with proportions $-1 < \nu_i$. The negative values $-1 < \nu_i < 0$ denote disastrous effects but limited by a lower bound so that the population will not be wiped out by a single disaster, while the positive values $\nu_i > 0$ denote bonanzas or beneficial effects. The randomness of the jumps is modeled by a set of n_p Poisson processes $P_i(t)$ with common infinitesimal means and variances

$$E[dP_i(t)] = \lambda_i dt = \text{Var}[dP_i(t)],$$

for $i = 1:n_p$, where $\lambda_i > 0$ is the i th jump rate. The actual jump at the j th jump time $t_{i,j}$ of the i th Poisson process is given in jump notation by

$$[X](t_{i,j}) \equiv X(t_{i,j}^+) - X(t_{i,j}^-) = \nu_i X(t_{i,j}^-).$$

The motivation for the multitude of jump terms in (12.1) is that large random fluctuations can come from many causes, like climatic changes, over-fishing and epi-zootics (see [208, 135, 245, 238], for instance).

In [237], Ryan and Hanson treated the optimal harvesting case where the natural growth of the biomass is exponential with jumps, i.e., $1/K = 0$ and the natural growth function is linear, $f(x) = rx$, so the overall growth of $X(t)$ is exponential with harvesting and jumps. The model (12.1) is a pure jump model with logistic drift because the focus is on the effects of jumps on the harvesting

bioeconomics, although diffusion terms could have been easily added to the model. For $r > 0$ with no harvesting and jumps, the exponential model $dX(t) = rX(t)dt$ leads to unbounded exponential growth, while the logistic model $dX(t) = rX(t)(1 - X(t)/K)dt$ leads to saturated growth as $X(t) \rightarrow K^-$ if $x_0 < K$ or limiting decay as $X(t) \rightarrow K^+$ if $x_0 > K$. The density dependent (nonlinear) jump case is treated by Hanson and Ryan in [113].

The economic value of the harvest, starting at time t with biomass x and ending at the final time t_f , is given by the expected, discounted present value,

$$\bar{V}[X, U](x, t) = \mathbb{E} \left[\int_t^{t_f} e^{-\delta s} h(X(s), U(s)) R(X(s), U(s)) ds \right] \quad (12.2)$$

$$X(t) = x, U(t) = u,$$

where δ is the continuous, inflation-corrected discount rate with discounting starting at $t = 0$ and $\exp(-\delta t)$ is the discount factor which accounts for the opportunity costs of investing money elsewhere in a secure investment. The instantaneous net harvest revenue per unit harvest is

$$R(x, u) = (ph(x, u) - C(u)) / h(x, u).$$

It can be assumed that $x > 0$ and $u > 0$ to avoid dividing by zero, but the net revenue always appears in the product form $h(x, u)R(x, u)$, so the divide check is not needed. The price of a unit of a harvested biomass ($h = qux$) is p and

$$C(u) = c_1u + c_2u^2$$

is the total cost of the harvesting effort when the biomass or stock size is x given that $c_2 > 0$ so that $C(u)$ is a genuine quadratic. Note that $C(u)$ is assumed to be quadratic in the effort, which suggests that the effort is more costly the larger it becomes. In the case of fisheries, this might mean that more inefficient fishing boats or less experienced fisherman are used as the fishing effort $U(t)$ increases. The effort is assumed to be bounded, i.e., constrained, so that

$$0 \leq U^{(\min)} \leq U(t) \leq U^{(\max)} < \infty \quad (12.3)$$

and the objective is to seek the maximum, expected current value

$$v^*(x, t) = \max_U [\bar{V}[X, U](x, t)].$$

Thus, the goal is to calculate optimal value $\bar{V}^*(x, t)$ and the optimal feedback control or effort

$$u^*(x, t) = \operatorname{argmax}_U [\bar{V}[X, U](x, t)]$$

for $0 \leq t < t_f$. However, the initial optimal expected, current value $\bar{V}^*(x, 0)$ is the optimal expected, discounted present value of future revenues.

In order to facilitate the application of the Hamilton-Jacobi-Bellman (HJB) equation theorem 7.3 to the discounted current value form in (12.2) with the so-called **cost function** $C(\mathbf{x}, \mathbf{u}, t) = \exp(-\delta t)h(\mathbf{x}, u)R(\mathbf{x}, u)$ here, the discount factor $\exp(-\delta t)$ can be removed in the pseudo-Hamiltonian by converting from the present value $v^*(x, t)$ of Chapter 7 to the current value $\mathcal{V}^*(x, t)$ by the transformation

$$v^*(x, t) = \exp(-\delta t)\mathcal{V}^*(x, t).$$

Thus, $v_t^*(x, t) = \exp(-\delta t)(\mathcal{V}_t^*(x, t) - \delta t \mathcal{V}^*(x, t))$, where v_t^* and \mathcal{V}_t^* are the partial derivatives of the value functions with respect to time. Note that initially both value functions coincide, $v^*(x, 0) = \mathcal{V}^*(x, 0)$.

Kamien and Schwartz [155] define the difference between the present and current value in terms of the present and current value Hamiltonians. The current value Hamiltonian $\hat{\mathcal{H}}(x, u, t)$ is related to the present value Hamiltonian $\tilde{\mathcal{H}}(x, u, t)$ by

$$\begin{aligned} \hat{\mathcal{H}}(x, u, t) &\equiv e^{+\delta t}\tilde{\mathcal{H}}(x, u, t) \\ &= (pqxu - c_1u - c_2u^2) + (rx(1 - x/K) - qux)\mathcal{V}_x^*(x, t) \\ &\quad + \sum_{i=1}^{n_p} \lambda_i(t) (\mathcal{V}^*(x + \nu_i x, t) - \mathcal{V}^*(x, t)), \end{aligned}$$

cancelling out the discount factor $\exp(-\delta t)$. Further, separating out the control terms, the HJBE is

$$\begin{aligned} 0 &= \mathcal{V}_t^*(x, t) - \delta \mathcal{V}^*(x, t) + rx(1 - x/K)\mathcal{V}_x^*(x, t) + \hat{\mathcal{S}}^*(x, t) \\ &\quad + \sum_{i=1}^{n_p} \lambda_i(t) (\mathcal{V}^*(x + \nu_i x, t) - \mathcal{V}^*(x, t)), \end{aligned} \tag{12.4}$$

where the control switching term contains all control terms in the quadratic form:

$$\hat{\mathcal{S}}(x, u, t) \equiv ((p - \mathcal{V}_x^*(x, t))qx - c_1 - c_2u)u,$$

including only the control dependent terms. The interior critical point of $\hat{\mathcal{S}}(x, u, t)$ with respect to the control u is the regular optimal control,

$$u^{(\text{reg})}(x, t) = \frac{0.5}{c_2} ((p - \mathcal{V}_x^*(x, t))qx - c_1), \tag{12.5}$$

since $c_2 > 0$, with the regular control being easily computed in terms of the gradient $\mathcal{V}_x^*(x, t)$ due to the quadratic cost assumption. As in the case of many applications, the control is constrained like in (12.3), so the constrained optimal control is the composite **bang-regular-bang** control function,

$$u^*(x, t) = \left\{ \begin{array}{ll} U^{(\text{min})}, & u^{(\text{reg})}(x, t) \leq U^{(\text{min})} \\ u^{(\text{reg})}(x, t), & U^{(\text{min})} \leq u^{(\text{reg})}(x, t) \leq U^{(\text{max})} \\ U^{(\text{max})}, & U^{(\text{max})} \leq u^{(\text{reg})}(x, t) \end{array} \right\}. \tag{12.6}$$

Consequently, the optimal control switch term is

$$\widehat{\mathcal{S}}^*(x, t) \equiv \widehat{\mathcal{S}}(x, u^*(x, t), t) = c_2 u^*(x, t) \left(2u^{(\text{reg})}(x, t) - u^*(x, t) \right),$$

after some algebraic manipulations. When $u^{(\text{reg})}(x, t)$ is within the constraints (12.3), the switch term will be quadratic in $u^{(\text{reg})}(x, t)$, i.e., $\mathcal{S}^*(x, t) = c_2 (u^{(\text{reg})})^2(x, t)$, and consequently quadratic in the value gradient $\mathcal{V}_x^*(x, t)$, so the PDE of stochastic dynamic programming will be PDE with a quadratic nonlinearity. The gradient $\mathcal{V}_x^*(x, t)$ is the so-called **shadow price** [56] for the way it directly modifies the price p in (12.5) and represents the expected value of future harvests [56]. The PDE is also a partial differential-difference equation, since the discrete Poisson jumps lead to difference terms in (12.4) rather than the mark integral over difference terms as more generally presented in Chapter 7.

The final condition for the backward HJB equation is $\mathcal{V}^*(x, t_f) = 0$ for $x > 0$ in absence of salvage or terminal costs. Thus, the final regular control or effort at $t = t_f$ is given by

$$u^{(\text{reg})}(x, t_f) = (pqx - c_1)/c_2 = c_1(x - x_f)/(2c_2x_f)$$

where $x_f \equiv c_1/(pq)$ is also the deterministic equilibrium stock value x_∞ [56]. However, in this stochastic case, if $c_1 \neq 0$, the final minimum control switch point is

$$x_{f, \min} = x_f \left(1 + 2c_2 U^{(\min)}/c_1 \right)$$

and the final maximum control switch point is

$$x_{f, \max} = x_f \left(1 + 2c_2 U^{(\max)}/c_1 \right).$$

As the biomass approaches extinction levels, $X(t) \rightarrow 0^+$, the rate of change $dX(t)$ (12.1) vanishes along with it, but the net revenue $R(x, u)$ should have become negative since costs dominate at low biomass. Hence, it will be assumed in addition that $R(x, u) \geq 0$, i.e., replacing $R(x, u)$ by $\max[R(x, u), 0]$, so that the extinction boundary condition is

$$\mathcal{V}^*(0^+, t) = 0$$

for $0 < t < t_f$.

A very reasonable approximation to the solution can be obtained from the **quasi-deterministic approximation**,

$$\begin{aligned} dX^{(\text{qdet})}(t) &\equiv \text{E} \left[dX(t) \mid X(t) = X^{(\text{qdet})}(t), U(t) = U^{(\text{qdet})}(t) \right] \\ &= r^{(\text{qdet})} X^{(\text{qdet})}(t) \left(1 - X^{(\text{qdet})}(t)/K^{(\text{qdet})} \right) dt, \end{aligned}$$

where $r^{(\text{qdet})} \equiv r + \sum_{i=1}^{n_p} \lambda_i \nu_i$ and $K^{(\text{qdet})} \equiv Kr^{(\text{qdet})}/r$, comprising an approximate logistic model. For this simplified model, the HJBE will no longer have difference terms since the jumps have been averaged out, but the optimal control will still be of the form (12.6).

Due to the complexity of the PDE, numerical methods are needed to approximate the solution. The HJBE can be solved with PDE-oriented finite difference methods [108] or the probability-oriented Markov chain approximation [175]. The finite difference method requires a sufficiently small mesh ratio for a comparison regular parabolic PDE [108] in the jump-diffusion case, while the Markov chain approximation, if the Markov chain probabilities are properly constructed, automatically comes with a weak convergence property [175]. For the current application in [238] and also in [237], Hanson and Ryan used the PDE-oriented finite difference method of [108] with predictor-corrector procedures to iterate on the non-linear terms and precision-preserving interpolation to approximate the jump terms by values at neighboring finite difference nodes. Both methods are variations of the finite difference method and are summarized in Chapter 9 in Sections 9.1 for the PDE-oriented method and 9.2 for the Markov chain approximation, respectively.

The primary bio-economic parameters used in [238] come from [57], i.e., r , K , q , p and c_1 , while other parameters like δ , t_f , λ_i and ν_i are reasonable estimates. Many of these estimated parameters were subjected to sensitivity tests in [238] in the many numerical results presented there. Some of the parameters are now obsolete, since whaling is no longer permitted in many countries or else highly restricted. Interest and discount rates are much lower now than then. However, significant sensitivity in u^* and V^* was found to the parameters δ , c_2 and $\lambda_i\nu_i/r$ for both a bonanza-dominated case with $\lambda_i\nu_i/r = 2\delta_{i,1}$ and a disaster-dominated case with $\lambda_i\nu_i/r = -0.5\delta_{i,2}$, where here $\delta_{i,j}$ is the Kronecker delta. In particular, in the **cheap control** limit as $c_2 \rightarrow 0^+$, the **bang-regular-bang** control law approaches a **bang-bang** control law in absence of a regular control component.

12.1.2 Optimal Harvesting with Both Price and Population Random Dynamics

The optimal harvesting problem, under joint population and price fluctuations in a random jump environment of Hanson and Ryan [115], is also an example of a two-dimensional state problem. Here, the problem is briefly summarized in the notation of this book. For general introduction to stochastic resource modeling, the reader can consult Anderson and Sutinen [9] or Mangel [189].

Again, let $X_1(t)$ be the amount of biomass (mass of the biological species) of the harvested species population at time t with stochastic dynamics consisting of logistic deterministic dynamics, discrete Poisson jumps and now with background, stochastic diffusion,

$$dX_1(t) = X_1(t) \left((r_1(1 - X_1(t)/K_1) - q_1U_1(t)) dt + \sigma_1 dW_1(t) + \sum_{i=1}^{n_1} \nu_{i,1} dP_{i,1}(t) \right), \quad (12.7)$$

$X_1(0) = x_{1,0} > 0$, where the extra subscript 1 designates population parameters or processes, i.e., the essential biological component of the bio-economic process. The

diffusion process $\sigma_1 dW_1(t)$ satisfies zero mean and $\sigma_1^2 dt$ variance properties, with $\sigma_1 > 0$ assumed. For the Poisson processes, $\nu_{i,1} > -1$ and $\lambda_{i,1} > 0$ for $i = 1:n_1$.

The economic process or price process $p(t)$ depends on the time-dependent bio-mass harvest rate $H(t) = h(X_1(t), U_1(t)) = q_1 U_1(t) X_1(t)$ and other stochastic processes. Since on the average $p(t)$ decreases as $h(t)$ increases [115] following **supply-demand principles**, the price is assumed to satisfy

$$p(t) = (p_0/H(t) + p_1) X_2(t), \tag{12.8}$$

where p_0 is a constant supply-demand coefficient such that $p(t)H(t)$ is the gross harvest return, p_1 is the constant price per unit harvested bio-mass coefficient and $X_2(t)$ is a fluctuating inflationary factor [115] satisfying the SDE

$$dX_2(t) = X_2(t) \left(r_2 dt + \sigma_2 dW_2(t) + \sum_{i=1}^{n_2} \nu_{i,2} dP_{i,2}(t) \right), \tag{12.9}$$

$X_2(0) = x_{2,0} > 0$, where the extra subscript 2 designates parameters and processes in the price process SDE, $\sigma_2 > 0$, $\nu_{i,2} > -1$ and $\lambda_{i,2} > 0$ for $i = 1:n_2$. It is assumed that all primary stochastic processes, $P_{i,1}(t)$, $P_{i,2}(t)$, $W_1(t)$ and $W_2(t)$, are pair-wise independent.

The economic value of the harvest, starting at time t with biomass x_1 and ending at the final time t_f , is given by the expected, discounted current value,

$$\bar{V}[\mathbf{X}, U_1](\mathbf{x}, t) = \mathbb{E} \left[\int_t^{t_f} e^{-\delta s} H(s) R(\mathbf{X}(s), U_1(s)) ds \mid \mathbf{X}(t) = \mathbf{x}, U_1(t) = u_1 \right], \tag{12.10}$$

where δ is the continuous, nominal discount rate, uncorrected by inflation since inflation is included in $X_2(t)$, with discounting starting at t . The random vector state is $\mathbf{X}(t) = [X_1(t) \ X_2(t)]^\top$ and $\mathbf{x} = [x_1 \ x_2]^\top$ is a sampled vector state, such that

$$R(\mathbf{x}, u_1) = ((p_0 + p_1 h(x_1, u_1)) x_2 - C(u_1)) / h(x_1, u_1)$$

is the instantaneous net harvest revenue rate per unit biomass. It can be assumed that $x_1 > 0$ and $u_1 > 0$ to avoid dividing by zero, but the net revenue always appears in the product form $h(x_1, u_1) R(\mathbf{x}, u_1)$ so the zero check is unneeded. The price of a harvested biomass unit is p and

$$C(u_1) = c_1 u_1 + c_2 u_1^2$$

is the total cost of the harvesting effort given that $c_2 > 0$ so that $C(u_1)$ is a genuine quadratic.

The effort control constraints are again assumed to be

$$0 \leq U_1^{(\min)} \leq U_1(t) \leq U_1^{(\max)} < \infty, \tag{12.11}$$

while the objective is to seek the maximum, expected current value

$$v^*(\mathbf{x}, t) = \max_{U_1} [\bar{V}[\mathbf{X}, U_1](\mathbf{x}, t)]$$

and the optimal feedback effort control

$$u_1^*(\mathbf{x}, t) = \operatorname{argmax}_{U_1} [\bar{V}[\mathbf{X}, U_1](\mathbf{x}, t)]$$

for $0 \leq t < t_f$. Again, the present values $v^*(\mathbf{x}, t)$ are transformed present values $\mathcal{V}^*(\mathbf{x}, t)$,

$$v^*(\mathbf{x}, t) = \exp(-\delta t)\mathcal{V}^*(\mathbf{x}, t).$$

The present value Hamiltonian $\tilde{\mathcal{H}}(\mathbf{x}, u_1, t)$ related to the current value Hamiltonian $\hat{\mathcal{H}}(\mathbf{x}, u_1, t)$ is

$$\begin{aligned} \tilde{\mathcal{H}}(\mathbf{x}, u_1, t) &\equiv e^{+\delta t}\hat{\mathcal{H}}(\mathbf{x}, u_1, t) \\ &= (p_0 + p_1q_1u_1x_1)x_2 - c_1u_1 - c_2u_1^2 \\ &\quad + (r_1x_1(1 - x_1/K_1) - q_1u_1x_1)\mathcal{V}_{x_1}^*(\mathbf{x}, t) + r_2x_2\mathcal{V}_{x_2}^*(\mathbf{x}, t) \\ &\quad + \frac{1}{2}\sigma_1^2x_1^2\mathcal{V}_{x_1,x_1}^*(\mathbf{x}, t) + \frac{1}{2}\sigma_2^2x_2^2\mathcal{V}_{x_2,x_2}^*(\mathbf{x}, t) \\ &\quad + \sum_{i=1}^{n_1} \lambda_{i,1}(t) (\mathcal{V}^*(1 + \nu_{i,1})x_1, x_2, t) - \mathcal{V}^*(\mathbf{x}, t) \\ &\quad + \sum_{i=1}^{n_2} \lambda_{i,2}(t) (\mathcal{V}^*(x_1, (1 + \nu_{i,2})x_2, t) - \mathcal{V}^*(\mathbf{x}, t)) . \end{aligned}$$

Upon cancelling out the discount factor $\exp(-\delta t)$ and separating out the control dependence from the Hamiltonian, the HJB equation is

$$\begin{aligned} 0 &= \mathcal{V}_t^*(\mathbf{x}, t) - \delta\mathcal{V}^*(\mathbf{x}, t) + p_0x_2 + r_1x_1(1 - x_1/K_1)\mathcal{V}_{x_1}^*(\mathbf{x}, t) + r_2x_2\mathcal{V}_{x_2}^*(\mathbf{x}, t) \\ &\quad + \frac{1}{2}\sigma_1^2x_1^2\mathcal{V}_{x_1,x_1}^*(\mathbf{x}, t) + \frac{1}{2}\sigma_2^2x_2^2\mathcal{V}_{x_2,x_2}^*(\mathbf{x}, t) \\ &\quad + \sum_{i=1}^{n_1} \lambda_{i,1}(t) (\mathcal{V}^*((1 + \nu_{i,1})x_1, x_2, t) - \mathcal{V}^*(\mathbf{x}, t)) \\ &\quad + \sum_{i=1}^{n_2} \lambda_{i,2}(t) (\mathcal{V}^*(x_1, (1 + \nu_{i,2})x_2, t) - \mathcal{V}^*(\mathbf{x}, t)) \\ &\quad + \hat{\mathcal{S}}^*(\mathbf{x}, t) \end{aligned} \tag{12.12}$$

where the control switching term has the quadratic form:

$$\hat{\mathcal{S}}(\mathbf{x}, u_1, t) \equiv p_1q_1u_1x_1x_2 - c_1u_1 - c_2u_1^2 - q_1u_1x_1\mathcal{V}_{x_1}^*(\mathbf{x}, t),$$

including only the control dependent terms. The interior critical point of $\hat{\mathcal{S}}(x, u, t)$ with respect to the control u is the regular optimal control,

$$u_1^{(\text{reg})}(\mathbf{x}, t) = \frac{0.5}{c_2} ((p_1x_2 - \mathcal{V}_{x_1}^*(\mathbf{x}, t)) q_1x_1 - c_1), \tag{12.13}$$

since $c_2 > 0$, with the regular control being easily computed in terms of the gradient $\mathcal{V}_x^*(\mathbf{x}, t)$ due to the quadratic cost assumption. As in the case of many applications, the control is constrained like in (12.11), so the constrained optimal control is the composite **bang-regular-bang** control function,

$$u_1^*(\mathbf{x}, t) = \left\{ \begin{array}{ll} U_1^{(\min)}, & u_1^{(\text{reg})}(\mathbf{x}, t) \leq U_1^{(\min)} \\ u_1^{(\text{reg})}(\mathbf{x}, t), & U_1^{(\min)} \leq u_1^{(\text{reg})}(\mathbf{x}, t) \leq U_1^{(\max)} \\ U_1^{(\max)}, & U_1^{(\max)} \leq u_1^{(\text{reg})}(\mathbf{x}, t) \end{array} \right\}. \quad (12.14)$$

The temporal side condition for the backward HJBE (12.12) is the final condition $\mathcal{V}^*(\mathbf{x}, t_f) = 0$ in absence of any terminal conditions for the first quadrant of state space and the natural corner condition

$$\mathcal{V}^*(0, 0, t) = - \left(c_1 + c_2 U_1^{(\min)} \right) U_1^{(\min)} (1 - \exp(-\delta(t_f - t))) / \delta$$

at the origin $(0, 0)$ for $0 < t < t_f$, since $U_1^{(\min)} \geq 0$. On the edge $(x_1, 0)$ for $x_1 > 0$, the boundary condition is similar to solving the pure jump optimal resource HJBE of Subsection 12.1.1 except that there is an additional diffusion term. On the edge $(0, x_2)$ for $x_2 > 0$, the boundary condition also involves solving an even less similar HJBE since in this case the deterministic inflationary growth is exponential rather than logistic.

In [115], data of the International Pacific Halibut Commission (IPHC) annual reports [145] are used since the catch and price data were readily available over a long period of time. Other data came from Clark [56]. The hybrid extrapolated-predictor-corrector Crank-Nicolson method similar to that described in the previous section and in Section 9.1 was used. The major result was that large inflationary increases had a very strong effect on the optimal return but not on the optimal effort.

Another multidimensional optimal harvesting problem can be found in the Lake Michigan salmon-alewife predator-prey model of Hanson in [105], where the alewife suffered large scale die-offs every several years. The model was also mixed economically, since the salmon are fished recreationally while the alewife were fished in a commercial fishery, now disbanded. Multidimensional visualization and parallel processing for renewable resources are developed by developed Practico et al. [224] and Hanson et al. in [112].

12.2 Stochastic Biomedical Applications

Variability plays an important role in medicine. Discussing critical care, Buchman [45] emphasizes that variability is **normal** for the individual patient and that illness is often accompanied by loss of individual variability. For instance, Boker et al. [36] find a variable ventilator improved lung function during surgery and recovery more than a controlled constant ventilator. Priplata et al. [227] find that input noise enhances balance, particularly for the elderly. Ashkenazy et al. [14] present a stochastic model to portray the variation in an individual's gait showing that

variability changes with maturation and aging. Moss et al. [207] find increased sensitivity in detecting threshold levels with stochastic noise for stochastic resonance to occur for nonlinear neural systems during information processing.

Swan [256] presents many applications of optimal control to biomedicine in his book, but the emphasis is on deterministic compartment or ODE models. One chapter is on cancer therapy control and another is on drug administration control. Murray's [209, 210] two volumes on models of mathematical biology has information on cancer and other models, but no real optimal control models.

According to Steel [250], and Goldie and Coldman [99] stochastic effects play a important role in the stages of development of cancer, the subsequent growth and the invasiveness of tumors or the more liquid lymphomas. Mutations can be induced by environmental chemical agents or ionizing radiation, while spontaneous mutations are more rare, usually without obvious cause [99].

12.2.1 Diffusion Approximation of Tumor Growth and Tumor Doubling Time Application

Tumor Growth Branching Process

Sometimes approximating a discrete stochastic process by a diffusion process can be useful. Hanson and Tier [117] present an example for branching process for modeling the growth tumor cells. This discrete model is then approximated as a diffusion process for the purposes of analysis and computation.

Let B_i be the branching process, in the i th generation for $i = 1, 2, 3, \dots$, such that there are three possible transitions in the time interval $(t, t + \Delta t)$ for generation i ,

$$B_i = \left\{ \begin{array}{l} 0, \text{ if cell death} \\ 1, \text{ if no cell change} \\ 2, \text{ if cell division} \end{array} \right\}, \quad (12.15)$$

similar to a birth-death model, but with a middle state of no change. This yields a total cancer cell count change from $N(t)$ at time t to

$$N(t + \Delta t) = \sum_{i=1}^{N(t)} B_i, \quad (12.16)$$

with the cell count change in $(t, t + \Delta t)$ being

$$\Delta N(t) = \sum_{i=1}^{N(t)} B_i - N(t) = \sum_{i=1}^{N(t)} (B_i - 1).$$

The B_i are assumed to be independent, identically distributed (IID) random variables with basic conditional moments that are dependent on $N(t)$, i.e., density dependent,

$$E[B_i | N(t) = n] = m(n; \Delta t)$$

and

$$\text{Var}[B_i | N(t) = n] = v(n; \Delta t).$$

The higher moments

$$\text{E}[(B_i - m(N(t), \Delta t))^k | N(t) = n] = m_k(n; \Delta t),$$

will also be needed to demonstrate that they will be small for $k \geq 3$.

Thus, the basic conditional moments of the tumor cell count change $\Delta N(t)$ are

$$\text{E}[\Delta N(t) | N(t) = n] = \sum_{i=1}^n \text{E}[B_i | N(t) = n] - n = n(m(n; \Delta t) - 1)$$

and

$$\begin{aligned} \text{Var}[\Delta N(t) | N(t) = n] &= \text{E}[(\Delta N(t) - n(m(n; \Delta t) - 1))^2 | N(t) = n] \\ &= \text{E}\left[\left(\sum_{i=1}^n (B_i - m(n; \Delta t))\right)^2 \mid N(t) = n\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \text{E}[(B_i - m(n; \Delta t))(B_j - m(n; \Delta t)) \mid N(t) = n] \\ &= \sum_{i=1}^n \text{E}[(B_i - m(n; \Delta t))^2 \mid N(t) = n] = nv(n; \Delta t), \end{aligned}$$

where the usual diagonalization technique has been used to apply the IID property of the B_i .

Diffusion Approximation of the Tumor Branching Process

Using some additional assumptions, a diffusion approximation will be constructed. Suppose T is some reference generation time such as the threshold for detection so

$$\tau = t/T$$

is a new scaled time and let a new scaled stochastic process be

$$X(\tau) = N(t)/T,$$

since the tumor will grow roughly as the number of generations. In order, for the model to be consistent with these scalings, the basic moments have to be refined so that the changes in $X(\tau)$ are small for small changes in τ . *The basic idea of the diffusion approximation is that it will not work well unless the order of the state changes are the same order as the time changes, i.e., $\Delta X(\tau) = O(\Delta\tau)$.* Hence, let the infinitesimal mean be of the *near-replacement* form,

$$m(N(t), \Delta t) = 1 + (m_1(X(\tau)) + o(1))\Delta\tau \text{ as } \Delta\tau \rightarrow 0,$$

where $m_1(x)$ is a function to be specified, and let the infinitesimal variance be of the form

$$v(N(t), \Delta t) = (v_1(X(\tau)) + o(1))T\Delta\tau \text{ as } \Delta\tau \rightarrow 0,$$

where $v_1(x)$ is a function to be specified. In addition, the higher moments should satisfy the form,

$$m_k(N(t), \Delta t) = o(\Delta\tau) \text{ as } \Delta\tau \rightarrow 0.$$

First for the diffusion approximation, the infinitesimal moments of $X(\tau)$, with $\Delta X(\tau) = \Delta N(t)/T$ are computed as in (8.64-8.65),

$$\begin{aligned} \mu(x) &= \lim_{\Delta\tau \rightarrow 0} \frac{E[\Delta X(\tau) | X(\tau) = x]}{\Delta\tau} = \lim_{\Delta\tau \rightarrow 0} \frac{E[\Delta N(T\tau)/T | N(T\tau) = Tx]}{\Delta\tau} \\ &= \lim_{\Delta\tau \rightarrow 0} \frac{x(m(Tx, T\Delta\tau) - 1)}{\Delta\tau} = \lim_{\Delta\tau \rightarrow 0} [x(m_1(x) + o(1))] = xm_1(x) \end{aligned} \quad (12.17)$$

and

$$\begin{aligned} \sigma^2(x) = \sigma_{1,1}(x) &= \lim_{\Delta\tau \rightarrow 0} \frac{\text{Var}[\Delta X(\tau) | X(\tau) = x]}{\Delta\tau} \\ &= \lim_{\Delta\tau \rightarrow 0} [x(v_1(x) + o(1))] = xv_1(x). \end{aligned} \quad (12.18)$$

In addition, the higher moment condition (8.66) when $k = 3$ is used (since any $k \geq 3$ can be used) in place of the continuity condition (8.63) due to the Chebyshev inequality (8.67). Hence,

$$\begin{aligned} \lim_{\Delta\tau \rightarrow 0} \frac{E[|\Delta X(\tau) - x(m(Tx, T\Delta\tau) - 1)|^3 | X(\tau) = x]}{\Delta\tau} &= \lim_{\Delta\tau \rightarrow 0} \frac{n \cdot m_3(Tx, T\Delta\tau)}{T^3\Delta\tau} \\ &= \lim_{\Delta\tau \rightarrow 0} \frac{x \cdot o(\Delta\tau)}{T^2\Delta\tau} = 0, \end{aligned}$$

completing the verification of the diffusion approximation and going substantially beyond the justification in [117].

For this particular application, the deterministic growth is chosen to be the **Gompertz growth model** [250, 99]

$$\mu(x) = xm_1(x) = \mu_1 x \ln(K/x), \quad (12.19)$$

where μ_1 is a constant growth coefficient and K is the carrying capacity or saturation level. Note that the Gompertz growth is singular as $x \rightarrow 0^+$, in that its derivative is unbounded as $x \rightarrow 0^+$, since $d(\mu_1 x \ln(k/x))/dx = -\mu_1 \ln(ex/K) \rightarrow +\infty$. However, the Gompertz model is often used in analyzing cancer experiments, although other models are also used, such as the simpler exponential growth model in shorter clinical trials [99]. In addition, the infinitesimal variance is taken to be purely linear, i.e.,

$$\sigma^2(x) = xv_1(x) = \sigma_1 x,$$

where $\sigma_1 > 0$ is a constant.

In summary, the backward operator in this time homogeneous case is

$$\mathcal{B}_{x_0}[u](x_0) = \frac{1}{2}\sigma_1 x u''(x_0) + \mu_1 x_0 \ln(K/x_0) u'(x_0). \quad (12.20)$$

Expected Tumor Doubling Time

The interest here is in the tumor doubling time, so suppose the tumor start is at the observed size c and then find the time it takes the tumor to double in size to $X(t) = 2c$. However, due to the stochastic nature of cancer, the tumor could become extinct, $X(t) = 0$ before it doubles in size. Hence, the proper problem is one of conditional probabilities, with the condition that the tumor doubles before it becomes extinct.

First consider the exit time at $2c$ starting at the general size $x_0 > 0$ at time t_0 ,

$$\tau_e^{(2c)}(x_0, t_0) = \inf_t [t \mid X(t) = 2c, X(s) \in (0^+, 2c), t_0 \leq s < t, X(t_0) = x_0], \quad (12.21)$$

so the backward formulation of Subsection 8.7.1 can be used with variable x_0 , here with $b = 2c$. Again let the exit time distribution function be

$$\Phi_{\tau_e^{(2c)}(x_0, t_0)}(t) = \text{Prob}[\tau_e^{(2c)}(x_0, t_0) < t]$$

with corresponding density $\phi_{\tau_e^{(2c)}(x_0, t_0)}(t)$ and let the ultimate probability of exit at $X(t) = 2c$ be

$$\Phi_e^{(2c)}(x_0, t_0) = \int_0^\infty \phi_{\tau_e^{(2c)}(x_0, t_0)}(t) dt.$$

Consequently, the final answer will be the expected doubling time

$$\Phi_{\text{dbl}}(c) = \Phi_e^{(2c)}(c, 0),$$

eventually using the initial values $x_0 = c$ and $t_0 = 0$.

Now let $u = u_0(x_0) = \Phi_e^{(2c)}(x_0, 0)$ and this satisfies the homogeneous backward equation

$$\mathcal{B}_{x_0}[u_0](x_0) = \frac{1}{2}\sigma_1 x u_0''(x_0) + \mu_1 x_0 \ln(K/x_0) u_0'(x_0) = 0, \quad (12.22)$$

from (12.20) in particular and (8.59) in general, but with boundary conditions,

$$u_0(0^+) = 0 \quad \text{and} \quad u_0(2c) = 1,$$

since an exit at $X(0) = 0^+$ is excluded under the conditioning and an exit at $X(t) = 2c$ is a certain conditional exit. Eq. (12.22) is integrable in u and $x_0 > 0$ by using an integrating factor or its inverse called the **Wronskian** (also called the diffusion **scale density**),

$$\begin{aligned} W(x_0) &\equiv \exp\left(-2 \int^{x_0} \frac{\mu(x)}{\sigma^2(x)} dx\right) = \exp\left(-2 \frac{\mu_1}{\sigma_1} \int^{x_0} \ln\left(\frac{K}{x}\right) dx\right) \\ &= \exp\left(-2 \frac{\mu_1 x_0}{\sigma_1} \ln\left(\frac{K}{e x_0}\right)\right) = (\beta_1 x_0)^{\gamma_1 x_0} \end{aligned} \quad (12.23)$$

here for the Gompertz model, where $\gamma_1 = 2\mu_1/\sigma_1$ and $\beta_1 = e/K > 0$. Thus, (12.22) simplifies to

$$(u'_0/W)'(x_0) = 0.$$

Thus, after two integrations and boundary condition substitutions lead to the solution of the boundary value problem,

$$\Phi_e^{(2c)}(x_0, 0) = u_0(x_0) = \frac{\int_{0^+}^{x_0} W(x)dx}{\int_{0^+}^{2c} W(y)dy}. \tag{12.24}$$

Since as $x \rightarrow 0^+$, $W(x) = (\beta_1 x)^{\gamma_1 x} \sim 1 + \gamma_1 x \ln(\beta_1 x)$ and then

$$\int_{0^+}^x dyW(y) \sim x + 0.5\gamma_1 x^2(\ln(\beta_1 x) - 0.5),$$

$W(x)$ is integrable as $x \rightarrow 0^+$ so that (12.24) is well defined, all other points on $(0, 2c)$ being obviously regular or non-singular points. Thus, setting $x_0 = c$ as the initial size gives the ultimate probability of a tumor doubling in size, $\Phi_e^{(2c)}(c, 0)$. More results by way of numerical and asymptotic approximations are given in [117].

The expected doubling time from (8.61) is

$$T_e^{(2c)}(c) = M_e^{(2c)}(c)/\Phi_e^{(2c)}(c, 0), \tag{12.25}$$

normalizing the first moment from (8.60), which here is

$$M_e^{(2c)}(x_0) \equiv \int_0^{+\infty} t\phi_{\tau_e^{(2c)}(x_0, 0)}(t)dt,$$

for general initial size x_0 and satisfying the backward equation from (8.62)

$$\mathcal{B}_{x_0} [M_e^{(2c)}](x_0) = -\Phi_e^{(2c)}(x_0, 0). \tag{12.26}$$

The backward equation for the moment is easier to solve than the one derived for the expected time quotient (12.25) since the quotient leads to a much more complicated equation. The boundary conditions are homogeneous,

$$M_e^{(2c)}(0^+) = 0 \text{ and } M_e^{(2c)}(2c) = 0,$$

but for different reasons, the first because 0^+ is the excluded exit and the second because it means an instant exit.

The solution can again utilize the Wronskian as a reciprocal integrating factor, such that

$$(u'_0/W)'(x_0) = -2V(x_0)u_0(x_0),$$

where

$$V(x) \equiv \frac{1}{\sigma^2(x)W(x)} = \frac{1}{\sigma_1 x (\beta_1 x)^{\gamma_1}},$$

here for the Gompertz model, is called the **speed density**. As $x \rightarrow 0^+$,

$$V(x) \sim \frac{1}{\sigma_1 x} (1 - \gamma_1 x \ln(\beta_1 x)),$$

so that for $0 < \epsilon \ll x \ll 1$,

$$\int_{\epsilon}^x dy V(y) \sim \sigma^{-1}(\ln(x/\epsilon) + \gamma_1 \epsilon \ln(\beta_1 \epsilon) + 1)$$

and is not integrable as $\epsilon \rightarrow 0^+$. The integrability of both $W(x)$ and $V(x)$, as well as that of some other functions, plays role in the Feller classification of boundaries for the Kolmogorov equations in one-dimension [31, 158]. Since a boundary is called a **regular boundary** if both $W(x)$ and $V(x)$ are integrable as the boundary point is approached from the interior of the domain, then 0^+ is a non-regular or **singular boundary** [158].

After two integrations, substitution of the boundary conditions to eliminate constants of integration and some manipulation of the integral forms, the solution of (12.26) can be written in the form

$$\begin{aligned} u_1(x_0) = & 2(1 - u_0(x_0)) \int_{0^+}^{x_0} dy W(y) \int_y^{2c} dz V(z) u_0(z) \\ & - 2u_0(x_0) \int_{x_0}^{2c} dy W(y) \int_y^{2c} dz V(z) u_0(z), \end{aligned} \quad (12.27)$$

provided the integrals exist. Letting $x_0 = c$, the expected doubling time is given by the formula in (12.25) or more simply by.

$$T_c^{(2c)}(c) = u_1(c)/u_0(c).$$

The multiple integral form of the solution (12.27) is too complicated to analyze further here, but additional numerical and asymptotic results are given in the paper of Hanson and Tier [117], including deterministic results. The application in [117] is based upon Fortner plasmacytoma data of Simpson-Herren and Lloyd [244]. The presentation here is somewhat different since it needed to be consistent with the notation and analytical formulation of this text.

Related formulation and results for other optimal stopping problem are some extinction problems for stochastic populations. They are examined for both diffusion in [116] and Poisson noise in [119, 121].

12.2.2 Optimal Drug Delivery to Brain PDE Model

In many applications, the control problem is formulated in terms of partial differential equations (PDEs), not ordinary differential equations (ODEs), since the problem depends on spatial variations and not just time variations. The ODE driven control problem is usually called **lumped parameter control**, sometimes arising from **compartmental models** lumping the spatial variables so that a PDE is not used, while the PDE driven control model is called **distributed parameter control**. The parameters in this latter case refer to the spatial variables in the background of the control problem. The mathematical background to this problem can be found in Section 6.5 or in Gunzberger [101] in much more detail for flow problems.

Cancer drug delivery to eliminate or reduce tumors is usually based upon expensive sets of experiments using animal and later human subjects to determine a fixed dose size and dose period to fit general patient, tumor and drug characteristics. Brain tumors can be very invasive and deadly, especially gliomas [257, 210]. When possible, the most of the mass of the tumor is removed (also called resectioned), but drug chemotherapy or radiotherapy is used in an attempt to kill any remaining cancer cells, including mobile metastases [80]. Gliomas can also be very diffusive [210], so reaction-diffusion equations may be used to model the drug delivery to the brain [257, 210, 92]. However, these reaction-diffusion investigations are only studies of the behavior of the solutions. No control of the drug delivery is involved. In this subsection, the paper of Chakrabarty and Hanson [48] on the control of reaction diffusion equations for optimal drug delivery to the brain is briefly summarized.

Optimal Control Problem for Drug Delivery Reaction-Diffusion Equations

Consider a reaction-diffusion model of a three-state system consisting of tumor cells, normal cells and cancer drug concentration in a brain. Let $y_1(\mathbf{x}, t)$ be the density of remaining tumor cells, $y_2(\mathbf{x}, t)$ be the density of normal cells and $y_3(\mathbf{x}, t)$ be the concentration of the drug at time t in time horizon $[0, t_f]$ and position \mathbf{x} in the brain domain \mathcal{D}_x . Let $\mathbf{y}(\mathbf{x}, t) = [y_3(\mathbf{x}, t)]_{3 \times 1}$ be the global state vector.

The tumor cell density satisfies the coupled reaction-diffusion equation

$$\frac{\partial y_1}{\partial t}(\mathbf{x}, t) = D_1 \nabla_x^2 [y_1](\mathbf{x}, t) + a_1 y_1 g_1(y_1) - (\alpha_{1,2} y_2 + \kappa_{1,3} y_3) y_1 \quad (12.28)$$

and the normal cells satisfy a similar equation

$$\frac{\partial y_2}{\partial t}(\mathbf{x}, t) = D_2 \nabla_x^2 [y_2](\mathbf{x}, t) + a_2 y_2 g_2(y_2) - (\alpha_{2,1} y_1 + \kappa_{2,3} y_3) y_2, \quad (12.29)$$

where D_i is the diffusion coefficient for the i th state, the $a_i y_i g_i(y_i)$ is the growth law for the i th state, the interaction coefficient $\alpha_{i,j} > 0$ signifies a constant death rate of tissue of state i due to tissue state j and the coefficient $\kappa_{i,3} > 0$ denotes a constant death rate due to the drug. For concreteness, the growth terms are taken to be logistic, i.e., $a_i y_i g_i(y_i) = a_i y_i (1 - y_i / K_i)$, where a_i is a constant intrinsic growth coefficient and $K_i > 0$ is a constant carrying-capacity or saturation level. Thus, there can be a strong interaction between the tumor and normal tissues, but the drug interaction is uni-directional. The drug concentration $y_3(\mathbf{x}, t)$ diffuses, gets absorbed and is controlled according to this reaction diffusion equation,

$$\frac{\partial y_3}{\partial t}(\mathbf{x}, t) = D_3 \nabla_x^2 [y_3](\mathbf{x}, t) + a_3 y_3 g_3(y_3) + u(\mathbf{x}, t), \quad (12.30)$$

where $a_3 y_3 g_3(y_3)$ is the drug absorption loss term and $u(\mathbf{x}, t)$ is the drug input control variable. For simplicity, the absorption term is taken to be exponential decay, so $a_3 y_3 g_3(y_3) = a_3 y_3$, where $a_3 < 0$ is the negative of the absorption coefficient and is assumed constant.

The vector reaction-diffusion PDE form merging (12.28,12.29,12.30) corresponding to (6.137) is

$$\frac{\partial \mathbf{y}}{\partial t}(\mathbf{x}, t) = D \nabla_x^2 [\mathbf{y}](\mathbf{x}, t) + \mathbf{B}(\mathbf{y}(\mathbf{x}, t), \mathbf{x}, t) + A \mathbf{u}(\mathbf{x}, t), \quad (12.31)$$

where $D = [D_i \delta_{i,j}]_{3 \times 3}$ is the diffusion coefficient,

$$\begin{aligned} \mathbf{B}(\mathbf{y}(\mathbf{x}, t), \mathbf{x}, t) = & (a_1 y_1 (1 - y_1) - (\alpha_{1,2} y_2 + \kappa_{1,3} y_3) y_2) \mathbf{e}_1 \mathbf{e}_1^\top \\ & + (a_2 y_2 (1 - y_2) - (\alpha_{2,1} y_1 + \kappa_{2,3} y_3) y_2) \mathbf{e}_2 \mathbf{e}_2^\top \\ & + a_3 y_1 \mathbf{e}_3 \mathbf{e}_3^\top \end{aligned} \quad (12.32)$$

is the bilinear reaction term with unit vectors $\mathbf{e}_k = [\delta_{i,k}]_{3 \times 1}$ for $k = 1:3$, $A = \mathbf{e}_3 \mathbf{e}_3^\top$ is the unit drug control coefficient and the drift term does not appear since $C \equiv 0$ here. The initial conditions for the vector PDE (12.31) is the vector

$$\mathbf{y}(\mathbf{x}, 0) = \mathbf{y}_0(\mathbf{x}), \text{ for } \mathbf{x} \in \mathcal{D}_x \quad (12.33)$$

and the boundary condition is a no-flux condition,

$$-(\hat{\mathbf{n}}^\top \nabla_x)[\mathbf{y}](\mathbf{x}, t) = \mathbf{0}, \quad (12.34)$$

where $\hat{\mathbf{n}} = \hat{\mathbf{n}}(\mathbf{x}, t)$ is the outward normal to the boundary $\partial \mathcal{D}_x$.

An objective in space-time is the minimization of the quadratic costs form,

$$\begin{aligned} V[\mathbf{y}, \mathbf{u}] = & \frac{1}{2} \int_{t_0}^{t_f} dt \int_{\mathcal{D}_x} d\mathbf{x} (\mathbf{y}^\top Q \mathbf{y} + (\mathbf{u} - \mathbf{u}_0)^\top R (\mathbf{u} - \mathbf{u}_0))(\mathbf{x}, t) \\ & + \frac{1}{2} \int_{\mathcal{D}_x} d\mathbf{x} (\mathbf{y}^\top S \mathbf{y})(\mathbf{x}, t_f), \end{aligned} \quad (12.35)$$

which is a slight variation in the control of the form (6.138), where the quadratic coefficients are $R = r_3 \mathbf{e}_3 \mathbf{e}_3^\top$ for the tumor burden cost, $S = s_1 \mathbf{e}_1 \mathbf{e}_1^\top$ for the drug delivery costs and $Q = q_1 \mathbf{e}_1 \mathbf{e}_1^\top + q_3 \mathbf{e}_3 \mathbf{e}_3^\top$ for the terminal costs, while the target threshold control value is $\mathbf{u}_0 = u_{0,3} \mathbf{e}_3$.

Hamiltonian Variational Formulation

The optimization problem above has three sets of constraints: the dynamics (12.31), the initial condition (12.33) and the boundary condition (12.34), so requires three Lagrange multipliers $\boldsymbol{\lambda}(\mathbf{x}, t)$, $\boldsymbol{\mu}(\mathbf{x}, t)$ and $\boldsymbol{\nu}(\mathbf{x})$ (without t since $t = 0$ for the initial condition), respectively, to form the **pseudo-Hamiltonian** as in (6.139),

$$\begin{aligned} \mathcal{H}(\mathbf{y}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}) = & V[\mathbf{y}, \mathbf{u}] + \int_{t_0}^{t_f} dt \int_{\mathcal{D}_x} d\mathbf{x} \boldsymbol{\lambda}^\top (\mathbf{y}_t - D \nabla_x^2 [\mathbf{y}] - \mathbf{B} - A \mathbf{u})(\mathbf{x}, t) \\ & + \int_{t_0}^{t_f} dt \int_{\partial \mathcal{D}_x} d\Gamma \boldsymbol{\mu}^\top (-\hat{\mathbf{n}}^\top \nabla_x)[\mathbf{y}](\mathbf{x}, t) \\ & + \int_{\mathcal{D}_x} d\mathbf{x} \boldsymbol{\nu}^\top (\mathbf{y}(\mathbf{x}, 0^+) - \mathbf{y}_0(\mathbf{x})). \end{aligned} \quad (12.36)$$

The main idea is that the Lagrange multipliers extend the three-vector state space to an extended six-vector state space

$$\mathbf{z}(\mathbf{x}, t) \equiv \{\mathbf{y}(\mathbf{x}, t), \mathbf{u}(\mathbf{x}, t), \boldsymbol{\lambda}(\mathbf{x}, t), \boldsymbol{\mu}(\mathbf{x}, t), \boldsymbol{\nu}(\mathbf{x})\}$$

to make the variations $\delta\mathbf{z}(\mathbf{x}, t)$ about $zbf^*(\mathbf{x}, t)$ in the extended objective systematic. Hence,

$$\mathcal{H}(\mathbf{z}^*(\mathbf{x}, t) + \delta\mathbf{z}(\mathbf{x}, t)) = \mathcal{H}(\mathbf{z}^*(\mathbf{x}, t)) + \delta\mathcal{H}(\mathbf{z}^*(\mathbf{x}, t), \delta\mathbf{z}(\mathbf{x}, t)) + O(|\delta\mathbf{z}|^2(\mathbf{x}, t)),$$

assuming that $\mathbf{z}^*(\mathbf{x}, t)$ exists and is a unique optimal solution under sufficient differentiability assumptions on $\mathcal{H}(\mathbf{z}(\mathbf{x}, t))$. Critical to these assumptions is that the perturbation of the nonlinear reaction term $\mathbf{B}(\mathbf{y}, \mathbf{x}, t)$ has a quadratic approximation, but that is trivial for this application since \mathbf{B} is quadratic in \mathbf{y} .

Skipping the details contained in Subsection 6.5.2, something very similar to the first variation $\delta\mathcal{H}(\mathbf{z}^*(\mathbf{x}, t), \delta\mathbf{z}(\mathbf{x}, t))$ in (6.141) is found. Setting the coefficients of $\delta\boldsymbol{\lambda}^\top(\mathbf{x}, t)$, $\delta\boldsymbol{\nu}^\top(\mathbf{x})$ and $\delta\boldsymbol{\mu}^\top(\mathbf{x}, t)$ (only for $\mathbf{x} \in \mathcal{D}_x$), respectively, to zero confirms that the PDE (12.31), initial condition (12.33) and boundary condition (12.34) hold with the optimal state $\mathbf{y}^*(\mathbf{x}, t)$ replacing for the state $\mathbf{y}(\mathbf{x}, t)$ of the original problem.

The final-boundary value PDE problem for the optimal adjoint state $\boldsymbol{\lambda}^*(\mathbf{x}, t)$ comes from setting the coefficients for $\delta\mathbf{y}^\top(\mathbf{x}, t_f)$, $\delta\mathbf{y}^\top(\mathbf{x}, t_f)$ and $\delta\mathbf{y}^\top(\mathbf{x}, t)$ (only for $\mathbf{x} \in \mathcal{D}_x$), respectively, to zero, producing

$$(\boldsymbol{\lambda}_t^* + \nabla_x^2[D\boldsymbol{\lambda}^*] - \nabla_y[\mathbf{B}^\top]^* \boldsymbol{\lambda}^* - Q\mathbf{y}^*)(\mathbf{x}, t) = \mathbf{0}, \quad \mathbf{x} \in \mathcal{D}_x, \quad t \in [0, t_f], \quad (12.37)$$

with final condition,

$$(\boldsymbol{\lambda}^* + S\mathbf{y}^*)(\mathbf{x}, t_f) = \mathbf{0}, \quad \mathbf{x} \in \mathcal{D}_x, \quad (12.38)$$

and boundary condition

$$(\hat{\mathbf{n}}^\top \nabla_x)[D\boldsymbol{\lambda}^*](\mathbf{x}, t) = \mathbf{0}, \quad \mathbf{x} \in \partial\mathcal{D}_x, \quad t \in (0, t_f), \quad (12.39)$$

which is the corresponding no-flux condition in backward form.

Setting the coefficient of $\delta\mathbf{u}(\mathbf{x}, t)$ to zero leads to

$$R(\mathbf{u}^*(\mathbf{x}, t) - \mathbf{u}_0^*(\mathbf{x}, t)) = A^\top \boldsymbol{\lambda}(\mathbf{x}, t),$$

which reduces to

$$u_3^*(\mathbf{x}, t) = u_{0,3}^*(\mathbf{x}, t) + \lambda_3^*(\mathbf{x}, t)/r_3, \quad \mathbf{x} \in \mathcal{D}_x, \quad t \in [0, t_f], \quad (12.40)$$

There are other optimality conditions that interrelate the Lagrange multipliers,

$$\boldsymbol{\nu}^*(\mathbf{x}) = \boldsymbol{\lambda}^*(\mathbf{x}, 0^+) \text{ for } \mathbf{x} \in \mathcal{D}_x$$

and

$$\boldsymbol{\mu}^*(\mathbf{x}, t) = D\boldsymbol{\lambda}^*(\mathbf{x}, t), \quad \mathbf{x} \in \partial\mathcal{D}_x, \quad t \in [0, t_f],$$

which will not be needed in the computations.

Forward-Backward Computational Iterations

The presence of nonlinear reaction terms in the forward state equation (12.31) using $\mathbf{y}^*(\mathbf{x}, t)$ with $\mathbf{u}^*(\mathbf{x}, t)$ and in the corresponding backward co-state equation (12.37) for $\boldsymbol{\lambda}^*(\mathbf{x}, t)$ make computational methods essential. The computational method of Chakrabarty and Hanson [48, 49, 50] employs a forward state integration of (12.31) and a backward integration of (12.37) with sufficient iterations until the norm of the iteration difference is small enough. The forward equation (12.31) is independent of the co-state $\boldsymbol{\lambda}^*(\mathbf{x}, t)$ but depends on the optimal control $\mathbf{u}^*(\mathbf{x}, t)$ which is a critical objective to be determined, so a starting guess for $\mathbf{u}^*(\mathbf{x}, t)$ is needed to start the forward integration, until a backward itegration generates a better guess using (12.40). On the other hand, the backward equation (12.37) depends strongly on the state distribution $\mathbf{y}^*(\mathbf{x}, t)$ as well as on its final values from (12.38), so that iterations, each consisting of a **double-shot** of both a forward iteration followed by a backward iteration, are required for reasonable accuracy. This double shot method is similar to the **opposite directions** multiple shooting method of Hackbusch [103] for parabolic equations. Gunzberger [101] calls many such methods **one-shot** methods and give more rigorous justification of them.

In order to keep the computational presentation manageable, let the forward and backward PDEs be represented in the more compact notation:

$$\begin{aligned} \mathbf{y}_t^*(\mathbf{x}, t) &= \mathcal{F}(\mathbf{x}, t, \mathbf{y}^*(\mathbf{x}, t), \mathbf{u}^*(\mathbf{x}, t)), \\ \mathbf{0} &= \boldsymbol{\lambda}_t^*(\mathbf{x}, t) + \mathcal{G}(\mathbf{x}, t, \boldsymbol{\lambda}^*(\mathbf{x}, t), \mathbf{y}^*(\mathbf{x}, t)), \end{aligned}$$

with general vector functions \mathcal{F} and \mathcal{G} for the forward and backward equations, respectively. Let the space vector \mathbf{x} be replaced by the discrete representation,

$$\mathbf{x}_j \equiv [x_{j_i,1} + (j_i - 1) \cdot \Delta x_i]_{3 \times 1}$$

where Δx_i is the step size in the i th direction, $\mathbf{j} = [j_i]_{3 \times 1}$ where, $j_i = 1:M_i$ nodes per direction, $i = 1:3$. Let the time t be replaced by the forward discretization

$$t_k \equiv k\Delta t,$$

for $k = 0:K$ time steps where Δt is the forward time step size, $t_0 = 0$ and $t_K = t_f$. The backward discrete time will be of the form $t_k^{(b)} \equiv t_f - k\Delta t = (K - k)\Delta t = t_{K-k}$. The corresponding discretization of the dependent vectors will be

$$\mathbf{y}(\mathbf{x}_j, t_k) \simeq \mathbf{Y}_{\mathbf{j},k}, \boldsymbol{\lambda}(\mathbf{x}_j, t_k) \simeq \boldsymbol{\Lambda}_{\mathbf{j},k} \text{ and } \mathbf{u}(\mathbf{x}_j, t_k) \simeq \mathbf{U}_{\mathbf{j},k}.$$

The numerical procedure used is the Crank-Nicolson method for second order accuracy in both space and time, but modified with additional extrapolation, prediction and correction techniques to accommodate nonlinear terms and multi-dimensions. The forward and backward discrete versions are written,

$$\mathbf{Y}_{\mathbf{j},k+1}^{(\gamma+1,\ell)} = \mathbf{Y}_{\mathbf{j},k}^{(\ell)} + \Delta t \mathcal{F}_{\mathbf{j},k+0.5}^{(\gamma,\ell)}, \quad (12.41)$$

$$\boldsymbol{\Lambda}_{\mathbf{j},k-1}^{(\gamma+1,\ell)} = \boldsymbol{\Lambda}_{\mathbf{j},k}^{(\ell)} + \Delta t \mathcal{G}_{\mathbf{j},k-0.5}^{(\gamma,\ell)}, \quad (12.42)$$

for $\gamma = 0:n_c$ corrections ($\gamma = 0$ is the prediction step) in each time step k until a relative stopping criterion for corrections in the tumor cell state component $Y_{1,\mathbf{j},k+1}^{(\gamma+1,\ell)}$ is satisfied,

$$\left\| Y_{1,\mathbf{j},k+1}^{(\gamma+1,\ell)} - Y_{1,\mathbf{j},k+1}^{(\gamma,\ell)} \right\| < \text{tol}_y \left\| Y_{1,\mathbf{j},k+1}^{(\gamma,\ell)} \right\| \quad (12.43)$$

for every state index \mathbf{j} , for $k = 0 : K - 1$ and during all double shot iterations $\ell = 1:L$, provided $\|Y_{1,\mathbf{j},k+1}^{(\gamma,\ell)}\| \neq 0$. The general notation means that

$$\mathcal{F}_{\mathbf{j},k+0.5}^{(\gamma,\ell)} = \mathcal{F} \left(\mathbf{x}_{\mathbf{j}}, t_{k+0.5}, \mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma,\ell)}, \mathbf{U}_{\mathbf{j},k+0.5}^{(\gamma,\ell)} \right)$$

and similarly for $\mathcal{G}_{\mathbf{j},k-0.5}^{(\gamma,\ell)}$. The relative tolerance in $Y_{1,\mathbf{j},k}^{(\gamma,\ell)}$ is tol_y . The Crank-Nicolson midpoint values are ordinarily approximated by the average,

$$\mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma,\ell)} \simeq 0.5 \left(\mathbf{Y}_{\mathbf{j},k+1}^{(\gamma,\ell)} + \mathbf{Y}_{\mathbf{j},k}^{(\gamma,\ell)} \right)$$

for $k = 0 : K - 1$ and

$$\mathbf{\Lambda}_{\mathbf{j},k-0.5}^{(\gamma,\ell)} \simeq 0.5 \left(\mathbf{\Lambda}_{\mathbf{j},k-1}^{(\gamma,\ell)} + \mathbf{\Lambda}_{\mathbf{j},k}^{(\gamma,\ell)} \right),$$

for $k = K : -1 : 1$, where $\mathbf{Y}_{\mathbf{j},k}^{(\ell)}$ and $\mathbf{\Lambda}_{\mathbf{j},k}^{(\ell)}$ are the final corrections for each time step k given shot ℓ , consistent with the second order Crank-Nicolson accuracy and implicitness reduction. A similar form is used for $\mathbf{U}_{\mathbf{j},k+0.5}^{(\gamma,\ell)}$. Second order central finite differences are used for all derivatives and based upon $\mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma,\ell)}$ or $\mathbf{\Lambda}_{\mathbf{j},k-0.5}^{(\gamma,\ell)}$.

The final stopping criterion for the convergence of the double shot iterations $\ell = 2:L$ is the pair of norms,

$$\left\| U_{3,\mathbf{j},k}^{(\ell)} - U_{3,\mathbf{j},k}^{(\ell-1)} \right\| < \text{tol}_u \left\| U_{3,\mathbf{j},k}^{(\ell-1)} \right\| \quad \text{and} \quad \left\| Y_{1,\mathbf{j},k}^{(\ell)} - Y_{1,\mathbf{j},k}^{(\ell-1)} \right\| < \text{tol}_y \left\| Y_{1,\mathbf{j},k}^{(\ell-1)} \right\|, \quad (12.44)$$

where the norm is over all \mathbf{j} and k , for $\ell = 2:L$ until satisfied, provided $\|U_{3,\mathbf{j},k}^{(\ell-1)}\| \neq 0$ and $\|Y_{1,\mathbf{j},k}^{(\ell-1)}\| \neq 0$, where $\text{tol}_u > 0$ and $\text{tol}_y > 0$ are some specified tolerances.

The treatment of the bilinear reaction term (12.32) requires careful consideration to accommodate the usual linear framework of the Crank-Nicolson method. Since this term has the pure bilinear form,

$$\mathbf{B}(\mathbf{y}, \mathbf{x}, t) = \widehat{B}(\mathbf{y})\mathbf{y},$$

in this application, this quasi-linear approximation is very appropriate

$$\widehat{B} \left(\mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma,\ell)} \right) \mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma,\ell)} \simeq \widehat{B} \left(\mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma-1,\ell)} \right) \mathbf{Y}_{\mathbf{j},k+0.5}^{(\gamma,\ell)},$$

in the forward equation for corrections $\gamma \geq 1$ and time steps $k \geq 1$.

Another special treatment needed is that of the no-flux boundary condition since central differences are inappropriate at the boundary, but backward and for-

ward differences of the same second order accuracy work very well, e.g.,

$$\mathbf{0} = -((\hat{\mathbf{n}}^\top \nabla_x)[\mathbf{Y}^*])_{\mathbf{j},k}^{(\gamma,\ell)} \simeq -\frac{(3\mathbf{Y}_{\mathbf{j},k}^{(\gamma,\ell)} - 4\mathbf{Y}_{\mathbf{j}-\hat{\mathbf{n}},k}^{(\ell)} + \mathbf{Y}_{\mathbf{j}-2\hat{\mathbf{n}},k}^{(\gamma,\ell)})}{(2|\hat{\mathbf{n}}^\top \Delta \mathbf{x}|)},$$

$$\mathbf{0} = ((\hat{\mathbf{n}}^\top \nabla_x)[(\mathbf{\Lambda})^*])_{\mathbf{j},k}^{(\gamma,\ell)} \simeq +\frac{(3\mathbf{\Lambda}_{\mathbf{j},k}^{(\gamma,\ell)} - 4\mathbf{\Lambda}_{\mathbf{j}-\hat{\mathbf{n}},k}^{(\gamma,\ell)} + \mathbf{\Lambda}_{\mathbf{j}-2\hat{\mathbf{n}},k}^{(\gamma,\ell)})}{(2|\hat{\mathbf{n}}^\top \Delta \mathbf{x}|)},$$

respectively, where $\hat{\mathbf{n}} \equiv \hat{\mathbf{n}}_{\mathbf{j},k}$, $\Delta \mathbf{x} = [\Delta x_i]_{3 \times 1} > \mathbf{0}$ and, e.g.,

$$\mathbf{Y}_{\mathbf{j}-\hat{\mathbf{n}},k}^{(\gamma,\ell)} = \mathbf{Y}^{(\gamma,\ell)}(\mathbf{x}_{\mathbf{j}} - |\hat{\mathbf{n}}^\top \Delta \mathbf{x}| \hat{\mathbf{n}}, t_k).$$

A sample output of the computations in Fig. 12.1 shows significant decrease in tumor size in one space dimension for a five day drug treatment trial. For information on the parameters used see Chakrabarty and Hanson [48]. For the corresponding two-dimensional space model of drug delivery see [49].

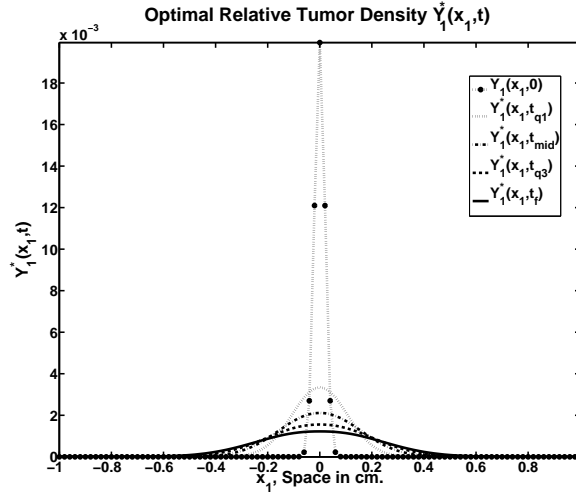


Figure 12.1. Optimal tumor density $Y_1^*(x_1, t)$ in the one-dimensional case with time as a parameter rounded at quartile values $\{0, t_{q1} = t_f/4, t_{mid} = t_f/2, t_{q3} = 3t_f/4, t_f\}$, where $t_f = 5$ days. The total tumor density integral is reduced by 29% in the 5-day simulated drug treatment trial.

Suggested References for Further Reading

- Anderson and Sutinen, 1984 [9].
- Ashkenazy et al., 2002 [14].
- Bharucha-Reid, 1960 [31].

- Boker et al., 2004 [36].
- Buchman, 2004 [45].
- Chakrabarty and Hanson, 2005a [48], 2005b [49] and 2006 [50].
- Clark, 1976 and 1990 [56].
- Clark and Lamberson, 1982 [57].
- Engelhard, 2000 [80].
- Goel and Dyn, 1974 [98].
- Goldie and Coldman, 1998 [99].
- Gunzberger, 2003 [101].
- Hackbusch, 1978 [103].
- Hanson, 1987 [105] and 1996 [108].
- Hanson et al., 1993 [112].
- Hanson and Ryan, 1988 [113] and 1998 [115].
- Hanson and Tier, 1981 [116] and 1982 [117].
- Hanson and Tuckwell, 1978 [118], 1981 [119], 1983 [120] and 1997 [121].
- Hennemuth et al., 1980 [135].
- Karlin and Taylor, 1975 [157, I] and 1981 [158, II].
- Kamien and Schwartz, 1981 [155].
- Kushner and Dupuis, 2001 [175].
- Ludwig, 1974 [183] and [184].
- Mangel, 1985 [189].
- Moss et al., 2004 [207].
- Murray, 2002 [209, I] and 2003 [210, II].
- Nisbet and Gurney, 1982 [215].
- Practico et al., 1992 [224].
- Priplata et al., 2002 [227].
- Ryan and Hanson, 1958 [237] and 1986 [238].
- Simpson-Herren and Lloyd, 1970 [244].

12.2. Stochastic Biomedical Applications

- Steel, 1977 [250].
- Swan, 1984 [256].
- Swanson, 1999 [257].
- Taylor and Karlin, 1998 [260].
- Tuckwell, 1989 [264] and 1995 [265].

DRAFT