AN EIGENVALUE APPROACH TO INFINITE-HORIZON OPTIMAL CONTROL

Per Rutquist * **Claes Breitholtz** ** **Torsten Wik** *

* *Volvo Technology Corporation*
** *Chalmers University of Technology*

Abstract: A method for finding optimal control policies for first order state-constrained, stochastic dynamic systems in continuous time is presented. The method relies on solution of the Hamilton-Jacobi-Bellman equation, which includes a diffusion term related to the stochastic disturbance in the model. A variable transformation is applied that turns the infinite-horizon optimal control problem into a linear eigenvalue problem in state-space. The method is demonstrated on a buffer control problem for a fuel cell-supercapacitor system. The obtained closed-form solution explains the shape of previous heuristically found control laws for this type of problem.

1. INTRODUCTION

Dynamic programming (Bellman, 1957) has received a lot of attention in the last decade (Bertsekas and Shreve, 1996), especially in the field of economics (Judd, 1998; Wendell and Stein, 2004). It is a powerful tool that over the years has led to the solution of many optimal control problems (Dorato *et al.*, 1995), such as the infinite-time linear quadratic control, which is one of the most employed synthesis methods for linear systems. As the cost of computational hardware decreases, dynamic programming is becoming important in numerous applications.

Energy management in hybrid electric vehicles is one of these, where the problem can be formulated as a stochastic dynamic programming (Rutquist, 2002). The problem is a nonlinear buffer control problem that is different from most other control problems in that the goal is not to keep the state, or buffer level, close to a set point. (In fact, a buffer whose state is kept constant is no longer a buffer!) The objective is instead to keep the level within bounds with as small control effort as possible. We will solve this problem analytically and numerically with a method that can be generalized to a wider class of optimal control problems.

The classical LQ control problems handled disturbances but required linear systems as well as unconstrained inputs and state space. Many steps have been taken since the first studies towards finding closed form solutions or numerically tractable methods to determine finite or infinite time optimal control for nonlinear systems with stochastic disturbances and input and state-space constraints.

Within the context of model predictive control (MPC) recent advances have been made. Manousiouthakis and Chimelewski (2002) have described a method to determine the constrained infinite time nonlinear optimal control problem for the case when there are no disturbances and the system is linear with respect to control inputs. Their method relies on the use of state dependent Ricatti equation (SDRE) approach introduced by Huang and Lu (1996).

Bemporad *et al.* (2002) have derived an algorithm for determining an explicit piece-wise linear control law for the discrete time constrained LQ-problem. The system must be linear and they suggest unmeasured disturbances to be dealt with by standard augmentation of the state vector, linear observer estimation and inclusion in the control law.

Here, we address the problem of finding an optimal control law for state-constrained nonlinear systems with stochastic disturbances. However, we are limited to first order systems that are input affine and on the form,

$$dx = f(x)dt + k(x)udt + s(x)dw, \quad x \in \Omega \tag{1}$$

where u is the control input and dw is white noise. It will be shown that for this special class of problems, the stochastic Hamilton-Jacobi-Bellman equation for the infinite time optimal control problem can be transformed into a Sturm-Liouville linear eigenvalue problem. Such problems are easily solved numerically with standard solvers. In some cases analytical solutions can be found, as will be shown in the derivation of a closed form optimal control law for buffer control in a fuel cell-super capacitor system.

2. PROBLEM FORMULATION

We associate with the system (1) a cost function c, which is quadratic in u

$$c\left(x,u\right) = c_0\left(x\right) + c_1\left(x\right)u + c_2\left(x\right)u^2 \,. \tag{2}$$

Since c is function of a random variable, we define the cost functional as an expected value

$$V(x,t) = E\left\{ \int_t^T c\left(x,u\right)\mathrm{d}\tau + c_f\left(x\left(T\right)\right) \middle| \, x\left(t\right) = x \right\}, \tag{3}$$

where T is the time horizon, c_f is a final cost which depends on the final state, and c is the instantaneous cost function which depends on state and control. The finite-horizon optimal control problem is to find the $u = u(x,t)$ that minimizes this cost functional while obeying the state constraint.

The infinite-horizon problem arises if we let T approach infinity. Since that would make the above integral to diverge we will instead use the cost functional

$$\lambda = \lim_{T\to\infty} \frac{V}{T} \,. \tag{4}$$

We know that the optimal control policy is stationary because of the Markov property of the process, so we thus seek the $u = u(x)$ that minimizes λ.

In order to be certain that the limit in (4) exists and is independent of the initial state $x(0)$, we assume that there exists a stationary distribution for the process (given the stationary control policy that we are considering). This requires, roughly, that the expected time for the process to diffuse from any given state to within ϵ of another is finite, which should be true for almost any engineering stochastic optimal control problem.

In one dimension, the state constraint can be written as

$$p\left(x \le x_{\min}\right) = 0 \tag{5a}$$

$$p\left(x_{\max} \le x\right) = 0 \,. \tag{5b}$$

(Here p denotes probability.) The disturbance $\mathrm{d}w$ is unbounded, so if $s(x_{min}) > 0$ and $s(x_{max}) > 0$ the above implies that

$$f\left(x_{\min}\right) + k\left(x_{\min}\right)u\left(x_{\min}\right) = \infty \tag{6a}$$

$$f\left(x_{\max}\right) + k\left(x_{\max}\right)u\left(x_{\max}\right) = -\infty \,, \tag{6b}$$

which in turn impose conditions on u on the boundaries of (1).

With the existence of a stationary distribution, equation (4) can be rewritten as

$$\lambda = E\left\{c\left(x, u\left(x\right)\right)\right\} \,, \tag{7}$$

so the optimization problem is completely defined by the functions f, k, s, and c plus the boundary conditions.

It is worth noting that we have not assumed that there exists any particular stationary point or otherwise preferred value of x.

3. THE HAMILTON-JACOBI-BELLMAN EQUATION

The Hamilton-Jacobi-Bellman (HJB) equation is central to optimal control theory. In the finite-horizon case for the above system it can be written as

$$\begin{aligned} -\frac{\partial}{\partial t}V\left(x,t\right) = \left[f\left(x\right) + k\left(x\right)u\right]\frac{\partial}{\partial x}V\left(x,t\right) \\ + \frac{1}{2}s^2\left(x\right)\frac{\partial^2}{\partial x^2}V\left(x,t\right) + c_0\left(x\right) \\ + c_1\left(x\right)u + c_2\left(x\right)u^2 \,. \end{aligned} \tag{8}$$

(See for example (Dorato *et al.*, 1995) for a more detailed discussion on the HJB equation.) The solution to the optimal control problem can be found by choosing u so as to minimize $-\partial V/\partial t$ locally, and integrating the above equation backwards in time.

For the infinite-horizon problem this is not directly applicable, because there is no final cost to serve as initial condition when integrating backwards in time. Also, as stated above, the integral diverges for infinite time.

Having already assumed that there exists a stationary distribution, resulting in an expected cost per unit time. The initial state cannot change this expectancy, but only add an expected one-time cost. We can therefore write

$$V\left(x,t\right) = V\left(x\right) - \lambda\, t \,. \tag{9}$$

We will now try to solve for the time-independent function $V(x)$.

4. EIGENVALUE METHOD (VARIABLE SUBSTITUTION)

The optimal u at each point is given explicitly by setting the derivative of (8) with respect to u equal to zero, yielding

$$u = -\frac{k(x)\,\frac{\partial}{\partial x}V(x,t) + c_1(x)}{2\,c_2(x)}\,. \tag{10}$$

All variables (excluding the constant λ) are now functions of x, but neither of t nor of u. The x-dependency will henceforth be dropped, and the prime symbol will be understood to denote x-derivatives.

The HJB equation (8) combined with the steady state condition (9), and the optimal u from (10) gives

$$\begin{aligned}\lambda = &-\frac{k^2\,(V')^2}{4\,c_2} + \frac{s^2 V''}{2} + fV' \\ &-\frac{kc_1V'}{2\,c_2} + c_0 - \frac{{c_1}^2}{4\,c_2}\,.\end{aligned} \tag{11}$$

We will assume that $c_2 > 0$ for all x, i.e. it is costly to use large $|u|$. We will also assume that $k \neq 0$ for all x, i.e. the control signal always has some effect on the state. (It is in fact not difficult to solve the problem even if $k = 0$ on some subset of the domain, but this will not be discussed here due to space limitations.)

Let W be a twice differentiable function of x that satisfies

$$V' = -\frac{BW'}{W}\,, \tag{12}$$

where B is an arbitrary strictly positive, piecewise continuous and bounded function of x, which we will specify shortly.

Substituting this into (11) yields

$$\begin{aligned}\lambda\,W = &-\frac{s^2BW''}{2} - \frac{k^2B^2\,(W')^2}{4\,c_2W} + \frac{s^2B\,(W')^2}{2\,W} \\ &-\frac{s^2\,(B')\,W'}{2} - fBW' + \frac{kc_1BW'}{2\,c_2} \\ &+c_0W - \frac{{c_1}^2W}{4\,c_2}\,.\end{aligned} \tag{13}$$

The above has the form of an eigenvalue problem except for the two nonlinear terms of the type $(W')^2/W$. However, if we choose B as

$$B = \frac{2\,s^2c_2}{k^2} \tag{14}$$

the troublesome terms cancel, and we can rewrite the above as

$$\lambda\,W = \alpha\,W'' + \beta\,W' + \gamma\,W\,, \tag{15}$$

where α, β and γ are functions of x that are given by:

$$\alpha = -\frac{s^4c_2}{k^2}\,, \tag{16}$$

$$\beta = \frac{2\,s^3c_2s'}{k^2} - \frac{2\,fs^2c_2}{k^2} + \frac{c_1s^2}{k} + \alpha'\,, \tag{17}$$

and

$$\gamma = c_0 - \frac{{c_1}^2}{4\,c_2}\,. \tag{18}$$

Given homogeneous boundary conditions, we are left with the linear eigenvalue problem (15) that can easily be solved using standard analytic and/or numeric techniques.

Since λ is the cost rate that should be minimized, we solve for the smallest eigenvalue. (From Sturm-Liouville theory we know that there exists a smallest eigenvalue for this problem.)

The optimal control policy is then computed directly from W by combining (10), (12) and (14) into

$$u = \frac{s^2W'}{kW} - \frac{c_1}{2\,c_2}\,. \tag{19}$$

4.1 Boundary Conditions

The boundary conditions will, as mentioned above, often be of the type $u = \infty$, which by (10) translates to

$$V'(x_{min}) = -\infty \tag{20}$$

This would seem like a problem, since infinity is typically not something that numerical algorithms handle well. However, after the variable transformation this becomes

$$W(x_{\min}) = 0\,, \tag{21}$$

which is a homogeneous condition that is ideally suited for eigenvalue problem solvers.

In practice it may sometimes be desirable to specify u at the boundary. This gives

$$Wu = \frac{s^2W'}{k} - \frac{c_1W}{2\,c_2}\,, \tag{22}$$

which is also a homogenous condition.

5. EXAMPLE: SUPERCAPACITOR CONTROL

One major advantage of a hybrid electric vehicle compared to a conventional vehicle is the possibility to store energy. Developing control strategies for this energy storage is an important task, which has received a lot of attention in recent years (Paganelli *et al.*, 2001; Rodatz *et al.*, 2004). The performance measure that is used is generally related to fuel consumption over some specified drive-cycle.

If the exact drive-cycle (speed as function of time) is known beforehand, then the optimal control can be computed explicitly as a function of time. This is rarely the case, and will not be further discussed here.

Much more common is to build heuristic control laws that give a control signal as a function of state and then use simulations on a drive-cycle to find proper heuristics. This second kind of strategy is more difficult to construct as it requires an engineer to do many modifications and simulations. It also runs the risk of overfitting to the simulated drive cycle that it was developed on, resulting in poor performance in practice.

5.1 *Model*

We consider a system consisting of a primary power source (for instance a fuel cell) and an energy store (supercapacitor). It provides electric power to some time-varying load.

The power electronics are controlled by a micro processor that ensures that the power balance is always satisfied, that is:

$$P_{primary} + P_{supercap} = P_{load}\,, \tag{23}$$

where P [W] denotes net power. Positive signs on all terms indicate a power flow from the battery and supercapacitor to the load.

Typically the losses can be described as a function of the power output of the primary power source. For the purposes of this example we assume that the following equation can be used to describe the losses:

$$\frac{P_{loss}}{P_{primary}} = C_1 + C_2 P_{primary} \tag{24}$$

The supercapacitor is simply modelled as a lossless energy storage. Its state x [J] reflects how much energy it has stored.

$$x = \int -P_{supercap}\,\mathrm{d}t \quad 0 < x < Q \tag{25}$$

where Q [J] denotes the amount of energy that can be stored in the supercapacitor.

We model the load as a constant average load plus a white noise.

$$P_{load}\,\mathrm{d}t = P_{avg}\,\mathrm{d}t + S\mathrm{d}w\,, \tag{26}$$

where P_{avg} [W] is the mean load, and $\mathrm{d}w$ is a white noise with zero mean and unity variance. S [W s$^{1/2}$] is a measure of the noise amplitude. It can be estimated from a measured load curve by replacing expectancy by average in the formula

$$S = \sqrt{\frac{1}{T}\mathrm{E}\left\{\left(\int_0^T (P_{load} - P_{avg})\,\mathrm{d}t\right)^2\right\}} \tag{27}$$

For a truly white noise, the same S should be obtained for any value of T. In practice, the noise will not contain infinitely high frequencies, and T should be chosen sufficiently large compared to the time constants of the process.

5.2 *Control*

We define u such that we can write the primary power output as

$$P_{primary} = P_{avg} + u\,. \tag{28}$$

Combining equation (24) with (28) yields

$$P_{loss} = C_2 u^2 + (C_1 + 2C_2 P_{avg})\,u + (C_1 + C_2 P_{avg})P_{avg} \tag{29}$$

Since u is defined to have zero mean, the expectancy of the middle term is zero, and we can leave it out to simplify our calculations. Similarly, the last term is, independent of both x and u, so it has no importance for the optimal control. For the cost function (2) we can then choose $c_0(x) = 0$, $c_1(x) = 0$, and $c_2(x) = C_2$ (Note that we would arrive at the same optimal control policy using $c_0(x) = (C_1 + C_2 P_{avg})P_{avg}$ and $c_1(x) = C_1 + 2C_2 P_{avg}$)

Combining equations (23), (25), (26) and (28) we obtain the stochastic differential equation

$$\mathrm{d}x = u\mathrm{d}t - S\mathrm{d}w \quad , \tag{30}$$

which in terms of (1) translates to $f(x) = 0$, $k(x) = 1$, and $s(x) = S$.

Inserting the above into the HJB equation yields

$$-\frac{\partial}{\partial t}V(x,t) = u\frac{\partial}{\partial x}V(x,t) + \frac{S^2 \frac{\partial^2}{\partial x^2}V(x,t)}{2} + C_2 u^2\,. \tag{31}$$

Substituting the optimal u gives

$$-\frac{\partial}{\partial t}V(x,t) = -\frac{\left(\frac{\partial}{\partial x}V(x,t)\right)^2}{4\,C_2} + \frac{S^2 \frac{\partial^2}{\partial x^2}V(x,t)}{2}\,. \tag{32}$$

Applying the variable substitution and choosing B as

$$B = 2\,S^2 C_2 \tag{33}$$

yields

$$\lambda\,W = -S^4 C_2 W''\,. \tag{34}$$

The boundary conditions are

$$W(0) = 0 \tag{35a}$$

$$W(Q) = 0\,. \tag{35b}$$

5.3 *Analytic Solution*

The eigenvalue problem (34), (35) has the analytic solutions

$$W = \sin\left(x\sqrt{\frac{\lambda}{S^4 C_2}}\right)\,, \tag{36}$$

where

$$\lambda = \frac{N^2 \pi^2 C_2 S^4}{Q^2}\,. \tag{37}$$

N can be any positive integer. The smallest possible λ is obtained for $N = 1$.

Inserting this into the expression for the optimal control policy (19) gives the tangent function

$$u = -\frac{\pi S^2}{Q}\tan\left(\frac{\pi x}{Q} - \frac{\pi}{2}\right)\,. \tag{38}$$

In this particular case an analytic solution for V is can easily be obtained from W by simple integration of (12).

5.4 *Numeric Solution*

A numeric solution can be found using the finite element method. Several software packages exist that implement the method, and the user has but to specify the partitioning of space, and the order of piecewise polynomials that will approximate the solution. We want the solution W to be twice differentiable, and this can be accomplished using piecewise third order polynomials.

The solution to equation (34) was computed, setting $S = 1$, $C_2 = 1$, and $Q = 1$. For this example we divided the state-space into three subintervals. With third order polynomials and continuous first derivative we then get six degrees of freedom, so we needed to find the smallest generalized eigenvector to a pair of six-by-six matrices. The lowest eigenvalue turns out to be $\lambda = 9.8699$, which is not far from $\pi^2 \approx 9.8696$, the true answer according to (37).

After computing W, the formula (19) gives the actual control policy

$$u = \begin{cases} \dfrac{3\,x^2 + 0.153\,x - 0.768}{x^3 + 0.077\,x^2 - 0.768\,x} & 0 < x \leq \dfrac{1}{3} \\[2ex] \dfrac{2\,x - 1}{x^2 - x + 0.039} & \dfrac{1}{3} < x \leq \dfrac{2}{3} \\[2ex] \dfrac{3\,x^2 - 6.153\,x + 2.384}{x^3 - 3.077\,x^2 + 2.384\,x - 0.307} & \dfrac{2}{3} < x < 1 \end{cases} . \tag{39}$$

A piecewise rational function is efficient and easy to implement in computer code.

6. DISCUSSION AND CONCLUSIONS

We have shown how a simplified battery-supercapacitor control problem can be solved in a systematic way, yielding a closed-form solution.

6.1 *Extension to higher dimension*

The method presented here can be extended to problems where x and u are vectors in some technically interesting cases. However, due to the computational cost of solving the eigenvalue problem, it is limited to problems of low order.

6.2 *Comparison to Linear Feedback Control*

It is interesting to note that the LQR method, despite its popularity, fails when applied to this buffer control problem. This is because the cost function is independent of x ($c(x, u) = c(u)$). [1] Solving for the optimal control policy then gives us $u \equiv 0$. This control policy is obviously only admissible if $s \equiv 0$. (It is a well-known fact that for an LQR-problem where the state is directly measurable, the solution is independent of the magnitude of the disturbance—so-called *certainty equivalence.*)

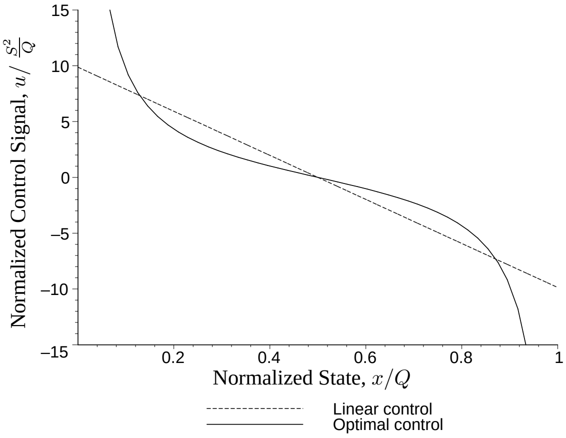

Figure 1. The optimal control law, and a linear control law with the same expected cost.

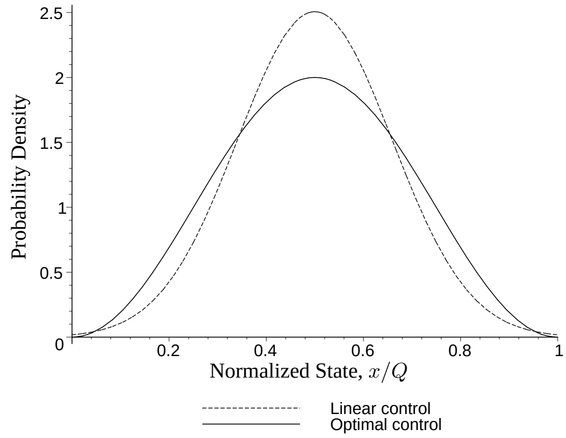

Figure 2. Probability densities. The optimal control law yields a wider peak, which means that more of the energy buffer is being utilized.

Figure 1 shows the optimal control law, for the case where $S = 1$ and $Q = 1$. Also depicted is a linear control law that has the same expected cost, if there is no penalty for violating the state constraints.

In Figure 2, the probability density functions for the state x in the stationary processes are shown. Linear feedback control results in a classical Gaussian distribution. It is centered around the setpoint and has a variance that is proportional the noise variance divided by the controller gain. The tails of the distribution extend to infinity in both directions, so mathematically it is impossible to find a linear control law that respects the constraints.

In the case with the optimal control signal, the probability density function has the shape of a sine-squared curve. While the shape is visually similar to the Gaussian curve, there are two important differences. First,

[1] Another view is that the cost function has an x-dependency in that it is zero on the interior of the state-space, and infinite on its boundary. This is obviously not something that can be approximated well by a quadratic function

the peak is wider, which means that a larger part of the buffer is used relatively often. This is important since it makes little sense to pay for energy storage capacity that is almost never used. Second, the probability goes to zero at both ends of the interval, so the state constraints are never violated.

6.3 *Complex Models and Numeric Solutions*

The model presented here is very simple, in order to admit a simple closed form solution. A more realistic model would include losses in the supercapacitor, power electronics, etc. The efficiency of the primary power source as function of load would also be more complex. The method demonstrated here cannot handle cost functions that depend arbitrarily on u. However, since $c(x, u)$ can be arbitrarily nonlinear in x and u is a function of x, it is quite straightforward to solve the problem iteratively.

The control signal in (38) becomes infinite as x approaches zero or Q. A realistic control signal would be limited in magnitude and so would the disturbance. If the control action can always overcome the disturbance, then we can still find a solution to the problem. If the disturbance can overcome the control, then the problem has no solution; there will always be a trade-off between the control cost and the probability of violating the constraints.

The state-space can be augmented to incorporate dynamics in both the load and the primary power source, yielding a higher-dimensional problem. For instance, it has been shown in (Rutquist, 2002) how a drive cycle can be converted into a two-state (speed and acceleration) stochastic process that can be used for computation of optimal stationary control laws using dynamic programming. The control law obtained from such a computation includes the states of the stochastic process. It will therefore take into account the fact that a vehicle near its maximum speed is more likely to go through a period of braking than acceleration.

6.4 *Applications to Buffer Control in General*

The fact that a proportional controller (with a suitably chosen gain) is optimal for a cost function $c(x, u) = ax^2 + bu^2$ has contributed to the popularity of P, PI, PID, PIDF and similar controllers. The controller presented here in the buffer control problem can be thought of as a proportional controller scaled by the tangent function. It is optimal for the cost function that is quadratic on an interval and infinite outside that interval. This leads us to conjecture that most buffer systems that are today using linear feedback control would benefit from incorporation of the tangent function into the controller.

In fact, many buffer systems are already using nonlinear control functions similar to (38). For example, in (Paganelli *et al.*, 2001) a heuristic function for state-of-charge control is depicted. The resemblance to the tangent function is striking.

REFERENCES

Bellman, Richard (1957). *Dynamic Programming*. Princeton University Press.

Bemporad, Alberto, Manfred Morari, Vivek Dua and Efstratios N. Pistikopoulos (2002). The explicit linear quadratic regulator for constrained systems. *Automatica* **38**, 3–20.

Bertsekas, Dimitri P. and Steven E. Shreve (1996). *Stocastic Optimal Control: The Discrete-Time Case*. Athena Scientific.

Dorato, Peter, Chaouki Abdallalh and Vito Cerone (1995). *Linear-Quadratic Control, An Introduction*. Prentice-Hall.

Huang, Y. and W. M. Lu (1996). Nonlinear optimal control: Alternatives to hamilton-jacobi equation. In: *Proceedings of the 35th Conference on Desicion and Control*. Kobe, Japan. pp. 3942–3947.

Judd, Kenneth L. (1998). *Numerical Methods in Economics*. The MIT Press.

Manousiouthakis, V. and D. J. Chimelewski (2002). On constrained infinite-time nonlinear optimal control. *Chemical Engineering Science* **57**, 105–114.

Paganelli, Gino, Gabriele Ercole, Avra Brahma, Yann Guezennec and Giorgio Rizzoni (2001). General supervisory control policy for the energy optimization of charge-sustaining hybrid electric vehicles. *JSAE Review*.

Rodatz, P., G. Paganelli, A. Sciarretta and L. Guzzella (2004). Optimal power management of an experimental fuel cell/supercapacitor-powered hybrid vehicle. *Control Engineering Practice*.

Rutquist, Per (2002). Optimal control for the energy storage in a hybrid electric vehicle. In: *Proc. 19th International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium and Exhibition (EVS19)*. Busan, Korea. pp. 1133–1139.

Wendell, H. F. and J. L. Stein (2004). Stochastic optimal control, international finance and debt. *Journal of banking and finance* **28**, 979–996.