Trajectory planning under a Stochastic Uncertainty

Ulf T. Jönsson
Optimization and Systems Theory
Royal Institute of Technology
SE-100 44 Stockholm, Sweden

Clyde Martin
Department of Mathematics and Statistics
Texas Tech University
Lubbock, Texas, USA

Yishao Zhou
Department of Mathematics
Stockholm University
SE-106 91 Stockholm, Sweden

Abstract

A trajectory planning problem for linear stochastic differential equations is considered in this paper. The aim is to control the system such that the expected value of the output interpolates given points at given times while the variance and an integral cost of the control effort are minimized. The solution is obtained by using the dynamic programming technique, which allows for several possible generalizations that are also considered. The results of this paper can be used for control of systems with a stochastic growth rate. This is frequently the case in investment problems and in biomathematics.

1 Introduction

We are concerned with controlling the state or some measured quantity of the state to given values at discrete times. That is given a system of the form

$$\dot{x} = f(x) + u_1 g_1(x) + \cdots + u_k g_k(x)$$

and

$$y = Cx$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^p$ we look for a system of controls that drive the output through or close to a prescribed sequence of set points

$$\{y(t_i) = \alpha_i : \quad i = 1, \cdots, N\}.$$

In this paper we specialize this problem to one in which the system is linear but the evolution equation is stochastic. Such systems occur in a variety of settings. We consider a rather complete example of a model that contains variable interest rates in Section 3. A second example, considered in Section 4, is trajectory planning for a feedback linearized mobile robot. Here the combined effect of surface irregularities, friction, and other disturbances is modeled as a multiplicative stochastic disturbance. Another set of problems occurs in biomathematics when considering the chemostat. There are extensive references on physical, biological and mathematical background of the chemostat available in the literature, see for example [5]. In many situations the nutrient content of

the chemostat is not deterministic and a better model would be obtained by replacing the linear growth rate of the standard models by a random variable. This variable reflects the rate at which the nutrient is consumed and/or naturally degrades. This rate is a function of temperature and many other variables that are not modelled and hence could and should be taken to be stochastic. It is beyond the scope of this paper to consider a complete analysis of such systems but a model in this form fits the theory described in this paper.

The specific topic of this paper is the problem of designing a feedback control for a linear stochastic differential equation in such a way that the mean value of a specific output interpolates given points at given times at the same time as the variance and the expected value of an integral quadratic cost along the trajectory are minimized. By minimizing the variance at the interpolation points we minimize effect of uncertainty due to the diffusion term in the system equation. By using a new notation we obtain compact formulas for the design equations. The problem is related to several methods based on optimal control for interpolating splines that recently have been proposed in the literature, see for example [4, 3, 12]. Apart from considering stochastic models, this paper distinguishes itself by using a dynamic programming approach which leads to feedback solutions of the corresponding control problem.

In Section 5 we discuss a generalization, where we optimize over a finite set of possible interpolation points at each time instance. This problem reminds of the standard shortest path problem, see e.g. [1], but is complicated by the dynamic nature which make the cost corresponding to each node dependent on future interpolation points as well as the initial condition. The complexity of this problem can be reduced by considering suboptimal techniques such as discretization or tree pruning. We briefly review some related work in [2, 6, 8, 9].

2 Stochastic Trajectory Planning

We consider a path planning problem where a stochastic disturbance enters the differential equation. As our basic problem we consider

$$J_0(x_0) = \min_{u \in \mathcal{M}(t_0, t_N)} \sum_{k=0}^{N-1} w_{k+1} \mathrm{Var}^2\{CX(t_{k+1})\} + E^{t_0, x_0}\left\{\int_0^{t_N} |u|^2 dt\right\}$$

$$\text{subject to} \quad \begin{cases} dX = (AX + Bu)dt + G(X)dZ, \quad X(t_0) = x_0 \\ E^{t_0, x_0}\{CX(t_{k+1})\} = \alpha_{k+1}, \quad k = 0, \dots, N-1 \end{cases}$$

(2.1)

where $w_k \geq 0$ and (A, B) is a controllable pair. The last term in the stochastic differential equation corresponds to a multiplicative noise on the state vector defined by a m-dimensional Brownian motion Z_t. In other words, Z_t is a process with zero mean and independent increments, i.e. $E\{dZ_t dZ_s^T\} = \delta(t - s)I dt$, where δ denotes the dirac function. We assume that $G(x)$ is a linear function on the form

$$G(x) = \sum_{i=1}^n x_i G_i$$

where $G_k \in \mathbb{R}^{n \times m}$. In (2.1), E^{t_0, x_0} is the expectation operator given that $X(t_0) = x_0$ and $\mathrm{Var}^2\{CX(t_k)\} = E^{t_0, x_0}\{(CX(t_k) - E^{t_0, x_0}\{CX(t_k)\})^2\}$.

We consider optimization over all Markov controls, i.e. feedback controls on the form $u(t, \omega) = \mu(t, X(t, \omega))$. We let $\mathcal{M}(t_0, t_N)$ denote the set of Markov controls on an interval (t_0, t_N). It can be shown that our optimal solution also is optimal over all $\mathcal{F}^t$ adapted processes, where $\mathcal{F}_t$ denotes the σ algebra generated by Z_s for $s \leq t$, see e.g. [10].

It turns out that there will be linear and constant terms in the value function due to the variance term in the objective function. It is therefore no essential addition in complexity to consider a more general path planning problem, where we allow the dynamics to be time-varying and different from stage to stage. We also consider integral costs that penalizes the state vector. This gives the following generalization of (2.1).

$$J_0(x_0) = \min_{u \in \mathcal{M}(t_0, t_N)} E^{t_0, x_0} \left\{ \sum_{k=0}^{N-1} \left(w_{k+1} |C_1 X(t_{k+1}) - \beta_{k+1}|^2 + \int_{t_k}^{t_{k+1}} \sigma_k(t, X, u) dt \right) \right\}$$

$$\text{subject to} \quad \begin{cases} dX = (A_k(t)X + B_k(t)u)dt + G_k(t, X)dZ, \ t \in [t_k, t_{k+1}], \ X(t_0) = x_0 \\ E^{t_0, x_0}\{C_2 X(t_{k+1})\} = \alpha_{k+1}, \quad k = 0, \dots, N-1 \end{cases}$$

$$(2.2)$$

where

$$\sigma_k(t, x, u) = x^T Q_k(t)x + 2x^T S_k(t)u + u^T R_k(t)u + 2q_k(t)^T x + 2r_k(t)^T u + \varrho_k(t)$$

and

$$G_k(t, x) = \sum_{i=1}^{n} x_i G_{k,i}(t)$$

and where everything else is defined in a similar way as in (2.1). Here $A_k, B_k, Q_k, \dots, \rho_k$ and $G_{k,i}$ are piecewise continuous functions of time and all pairs $(A_k(t), B_k(t))$ are assumed to be completely controllable and the cost is strictly convex, which is the case if for all $t \in [t_k, t_{k+1}]$, $k = 0, \dots, N-1$, we have

$$R_k(t) > 0 \quad \text{and} \quad \begin{bmatrix} Q_k(t) & S_k(t) \\ S_k(t)^T & R_k(t) \end{bmatrix} \geq 0 \tag{2.3}$$

Note that if $C_1 = C_2 = C$ and $\alpha_k = \beta_k$ then $E^{t_0, x_0}\{|CX(t_k) - \beta_k|^2\} = \text{Var}^2\{CX(t_k)\}$.

Let us define the cost-to-go functions

$$J_k(x) = \min_{u \in \mathcal{M}(t_k, t_N)} E^{t_k, x} \left\{ \sum_{i=k}^{N-1} \left(w_{i+1} |C_1 X(t_{i+1}) - \beta_{i+1}|^2 + \int_{t_i}^{t_{i+1}} \sigma_i(t, X, u) dt \right) \right\}$$

$$\text{subject to} \quad \begin{cases} dX = (A_i(t)X + B_i(t)u)dt + G_i(t, X)dZ, \ t \in [t_i, t_{i+1}], \ X(t_k) = x \\ E^{t_0, x_0}\{C_2 X(t_{i+1})\} = \alpha_{i+1}, \quad i = k, \dots, N-1 \end{cases}$$

$$J_N(x) = 0$$

$$(2.4)$$

Due to the Markov property of the stochastic differential equation we can use dynamic programming to solve (2.1). We next state two propositions and then the main result that solves (2.1). The first states the dynamic programming recursion.

3

Proposition 2.1. *The optimal cost satisfies the following dynamic programming equation.*

$$J_k(x) = \min_{u \in \mathcal{M}(t_k, t_{k+1})} E^{t_k, x} \left\{ \int_{t_k}^{t_{k+1}} \sigma_k(t, X, u) dt + w_{k+1} |C_1 X(t_{k+1}) - \beta_{k+1}|^2 + J_{k+1}(X(t_{k+1})) \right\}$$

$$\text{subject to} \quad \begin{cases} dX = (A_k(t)X + B_k(t)u)dt + G_k(t, X)dZ, \quad X(t_k) = x \\ E^{t_k, x} \{C_2 X(t_{k+1})\} = \alpha_{k+1} \end{cases}$$

$$J_N(x) = 0$$

The next proposition states the solution to the following stochastic optimal control problem. It shows that $J_k(x)$ is a quadratic function which is instrumental in solving the dynamic programming iteration.

$$V(x_0, \alpha, t_0, t_f) = \min_{u \in \mathcal{M}(t_0, t_f)} E^{t_0, x_0} \left\{ \int_{t_0}^{t_f} \sigma(t, X(t), u(t)) dt + X(t_f)^T Q_0 X(t_f) + 2q_0^T X(t_f) + \varrho_0 \right\}$$

$$\text{subject to} \quad \begin{cases} dX = (A(t)X + B(t)u)dt + G(t, X)dZ, \quad X(t_0) = x_0 \\ E^{t_0, x_0} \{CX(t_f)\} = \alpha \end{cases}$$

$$(2.5)$$

where

$$\sigma(t, x, u) = x^T Q(t)x + 2x^T S(t)u + u^T R(t)u + 2q(t)^T x + 2r(t)^T u + \varrho(t)$$

and where Q, R, and S satisfies the conditions in (2.3).

Proposition 2.2. *We have*

$$V(x_0, \alpha, t_0, t_f) = x_0^T P(t_0)x_0 + 2p(t_0)^T x_0 + \rho(t_0) + (N(t_0)^T x_0 + m(t_0))^T W(t_0)^{-1}(N(t_0)^T x_0 + m(t_0))$$

where

$$\begin{aligned}
&\dot{P} + A^T P + PA + Q + \Pi(P) = (PB + S)R^{-1}(PB + S)^T, \quad P(t_f) = Q_0 \\
&\dot{N} + (A - BR^{-1}(B^T P + S^T))^T N = 0, \quad N(t_f) = C^T \\
&\dot{p} + (A - BR^{-1}(B^T P + S^T))^T p + q = (PB + S)R^{-1}r, \quad p(t_f) = q_0 \\
&\dot{W} + NBR^{-1}B^T N = 0, \quad W(t_f) = 0 \\
&\dot{m} = N^T BR^{-1}(B^T p + r), \quad m(t_f) = -\alpha \\
&\dot{\rho} + \varrho = (r + B^T p)^T R^{-1}(r + B^T p), \quad \rho(t_f) = \varrho_0
\end{aligned} \quad (2.6)$$

and where $\Pi(P)$ is a linear matrix function with elements $\Pi(P)_{k,l} = \frac{1}{2}\text{tr}(G_k^T PG_l)$. The optimal control is

$$u^* = -R^{-1}(B^T P + S^T)X - R^{-1}B^T Nv - R^{-1}(B^T p + r)$$

with $v = -W(t_0)^{-1}(N(t_0)^T x_0 + m(t_0))$.

Remark 2.1. *If Q, R and S satisfies a condition analogous to (2.3), then the linearly perturbed Riccati equation in (2.6) has an absolutely continuous unique positive semidefinite solution [11]. All other differential equations in (2.6) are linear with bounded piecewise continuous coefficients, which ensure existence of unique absolutely continuous solutions. Note also that we have $W(t) > 0$ for $t \in [t_0, t_f)$ by the complete controllability of $(A(t), B(t))$.*

Proof. The proof is done by Lagrangian relaxation of the linear constraint. See Appendix A for the complete details. □

In order to obtain a compact notation we introduce

$$
\begin{aligned}
&\widehat{Q}_0 = \begin{bmatrix} Q_0 & q_0 \\ q_0^T & \varrho_0 \end{bmatrix}, \quad \widehat{Q} = \begin{bmatrix} Q & q \\ q^T & \varrho \end{bmatrix}, \quad \widehat{S} = \begin{bmatrix} S \\ r^T \end{bmatrix}, \\
&\widehat{P} = \begin{bmatrix} P & p \\ p^T & \rho \end{bmatrix}, \quad \widehat{N} = \begin{bmatrix} N \\ m^T \end{bmatrix}, \quad \widehat{W} = W, \quad \widehat{R} = R \\
&\widehat{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \quad \widehat{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad \widehat{C}(\alpha) = \begin{bmatrix} C & -\alpha \end{bmatrix}, \\
&\hat{G}(t,x) = \begin{bmatrix} G(t,x) \\ 0 \end{bmatrix}, \quad \hat{\Pi}(\widehat{P}) = \begin{bmatrix} \Pi(P) & 0 \\ 0 & 0 \end{bmatrix}.
\end{aligned}
\tag{2.7}
$$

If we finally let $\widehat{X} = \begin{bmatrix} X^T & 1 \end{bmatrix}^T$ and $\widehat{x}_0 = \begin{bmatrix} x_0^T & 1 \end{bmatrix}^T$ then the optimization problem (2.5) can be written

$$
V(x_0, \alpha, t_0, t_f) = \min_{u \in \mathcal{M}(t_0, t_f)} E^{t_0, \widehat{x}_0} \left\{ \int_{t_0}^{t_f} [\widehat{X}^T \widehat{Q} \widehat{X} + 2\widehat{X}^T \widehat{S} u + u^T \widehat{R} u] dt + \widehat{X}(t_f)^T \widehat{Q}_0 \widehat{X}(t_f) \right\}
$$

$$
\text{subject to} \quad \begin{cases} d\widehat{X} = (\widehat{A}(t)\widehat{X} + \widehat{B}(t)u)dt + \hat{G}(t,X)dZ, \quad \widehat{X}(t_0) = \widehat{x}_0 \\ E^{t_0, \widehat{x}_0} \{ \widehat{C}(\alpha)\widehat{X}(t_f) \} = 0 \end{cases}
\tag{2.8}
$$

and the optimal solution can be written

$$
V(x_0, \alpha, t_0, t_f) = \widehat{x}_0^T \left[\widehat{P}(t_0) + \widehat{N}(t_0)\widehat{W}(t_0)^{-1}\widehat{N}(t_0)^T \right] \widehat{x}_0
$$

where

$$
\dot{\widehat{P}} + \widehat{A}^T \widehat{P} + \widehat{P}\widehat{A} + \widehat{Q} + \hat{\Pi}(\widehat{P}) = (\widehat{P}\widehat{B} + \widehat{S})\widehat{R}^{-1}(\widehat{P}\widehat{B} + \widehat{S})^T, \quad \widehat{P}(t_f) = \widehat{Q}_0
$$

$$
\dot{\widehat{N}} + (\widehat{A} - \widehat{B}\widehat{R}^{-1}(\widehat{B}^T\widehat{P} + \widehat{S}^T))^T \widehat{N} = 0, \quad \widehat{N}(t_f) = \widehat{C}(\alpha)^T
$$

$$
\dot{\widehat{W}} + \widehat{N}^T \widehat{B}\widehat{R}^{-1}\widehat{B}^T \widehat{N} = 0, \quad \widehat{W}(t_f) = 0
$$

The optimal control is

$$
u^* = -\widehat{R}^{-1}(\widehat{B}^T\widehat{P} + \widehat{S})\widehat{x} - \widehat{R}^{-1}\widehat{B}^T \widehat{N}\widehat{W}(t_0)^{-1}\widehat{N}(t_0)^T\widehat{x}_0
$$

There is an equivalent feedback form given as $u^* = -\widehat{R}^{-1}(\widehat{B}^T(\widehat{P} + \widehat{N}\widehat{W}^{-1}\widehat{N}^T) + \widehat{S}^T)\widehat{X}$.

We can now state the solution of the general stochastic trajectory planning problem in (2.2).

Proposition 2.3. *Consider the optimal control problem in (2.2), where the condition (2.3) holds. The optimal Markov control in each time interval $[t_k, t_{k+1}]$ is (all variables are defined analogously with (2.7))*

$$
u^*(t) = -\widehat{R}_k(t)^{-1}(\widehat{B}_k(t)^T\widehat{P}_k(t) + \widehat{S}_k(t)^T)\widehat{X}(t) - \widehat{R}_k(t)^{-1}\widehat{B}_k(t)^T\widehat{N}_k(t)\widehat{W}_k(t_k)^{-1}\widehat{N}_k(t_k)^T\widehat{X}(t_k)
$$

where for $k = N-1, \ldots, 0$

$$\dot{\widehat{P}}_k + \widehat{A}_k^T \widehat{P}_k + \widehat{P}_k \widehat{A}_k + \widehat{Q}_k + \hat{\Pi}_k(\widehat{P}_k) = (\widehat{P}_k \widehat{B}_k + \widehat{S}_k) \widehat{R}_k^{-1} (\widehat{P}_k \widehat{B}_k + \widehat{S}_k)^T,$$

$$\dot{\widehat{N}}_k + (\widehat{A}_k - \widehat{B}_k \widehat{R}_k^{-1} (\widehat{B}_k^T \widehat{P}_k + \widehat{S}_k^T))^T \widehat{N}_k = 0, \quad \widehat{N}_k(t_{k+1}) = \widehat{C}_2(\alpha_{k+1})^T,$$

$$\dot{\widehat{W}}_k + \widehat{N}_k^T \widehat{B}_k \widehat{R}_k^{-1} \widehat{B}_k^T \widehat{N}_k = 0, \quad \widehat{W}_k(t_{k+1}) = 0,$$

and where $\widehat{P}_{N-1}(t_N) = w_N \widehat{C}_1(\beta_N)^T \widehat{C}_1(\beta_N)$ and for $k = N-2, N-3, \ldots, 0$

$$\widehat{P}_k(t_{k+1}) = \widehat{P}_{k+1}(t_{k+1}) + \widehat{N}_{k+1}(t_{k+1}) \widehat{W}_{k+1}(t_{k+1})^{-1} \widehat{N}_{k+1}(t_{k+1})^T + w_{k+1} \widehat{C}_1(\beta_{k+1})^T \widehat{C}_1(\beta_{k+1}).$$

The cost-to-go is

$$J_k(x) = \widehat{x}^T \left[\widehat{P}_k(t_k) + \widehat{N}_k(t_k) \widehat{W}_k(t_k)^{-1} \widehat{N}_k(t_k)^T \right] \widehat{x}.$$

Proof. By dynamic programming. $\qquad\square$

The formulation of Proposition 2.3 in the compact notation (2.7) gives appealing formulas but in a numerical implementation it is more efficient to perform computations using a system on the form (2.6). For the basic trajectory planning in (2.1) this reduces to the following result

Corollary 2.1. *Consider the optimal control problem in (2.1), where the pair (A, B) is controllable. The optimal Markov control in each time interval $[t_k, t_{k+1}]$ is*

$$u^*(t) = -R^{-1}B^T P_k(t)X(t) - R^{-1}B^T N_k(t)v_k - R^{-1}B^T p_k(t)$$

with $v_k = -W_k(t_k)^{-1}(N_k(t_k)^T X(t_k) + m_k(t_k))$, where

$$\begin{aligned}
&\dot{P}_k + A^T P_k + P_k A + \Pi(P_k) = P_k B R^{-1} B^T P_k, \\
&\dot{N}_k + (A - B R^{-1} B^T P_k)^T N_k = 0, \\
&\dot{p}_k + (A - B R^{-1} B^T P_k)^T p_k = 0, \\
&\dot{W}_k + N_k^T B R^{-1} B^T N_k = 0, \\
&\dot{m}_k = N_k^T B R^{-1} B^T p_k, \\
&\dot{\rho}_k = p_k^T B R^{-1} B^T p_k,
\end{aligned} \qquad (2.9)$$

where $\Pi(P_k)$ is a linear matrix function with elements $\Pi(P)_{k,l} = \frac{1}{2}\mathrm{tr}(G_k^T P G_l)$ and the boundary conditions are

$$P_k(t_{k+1}) = P_{k+1}(t_{k+1}) + N_{k+1}(t_{k+1}) W_{k+1}(t_{k+1})^{-1} N_{k+1}(t_{k+1})^T + w_{k+1} C^T C$$

$$p_k(t_{k+1}) = p_{k+1}(t_{k+1}) + N_{k+1}(t_{k+1}) W_{k+1}(t_{k+1})^{-1} m_{k+1}(t_{k+1}) + w_{k+1} \alpha_{k+1}^T C^T$$

$$\rho_k(t_{k+1}) = \rho_{k+1}(t_{k+1}) + m_{k+1}(t_{k+1})^T W_{k+1}(t_{k+1})^{-1} m_{k+1}(t_{k+1}) + w_{k+1} \alpha_{k+1}^T \alpha_{k+1}$$

and $N_k(t_{k+1}) = C^T$, $m_k(t_{k+1}) = -\alpha_{k+1}$ and $W_k(t_{k+1}) = 0$.
The optimal cost-to-go is

$$J_k(x) = x^T P_k(t_k)x + 2p_k(t_k)^T x + \rho_k(t_k) + (N_k(t_k)^T x + m_k(t_k))^T W_k(t_k)^{-1}(N_k(t_k)^T x + m_k(t_k))$$

3   A Homeowners Investment Problem

We consider a very simple problem arising in portfolio optimization. For most people their home is the most dominant feature of the portfolio. In this example we consider the problem of controlling the investment in the home when interest rates fluctuate and inflation, either positive or negative, occurs. In this example we are aware that in addition to making monthly payments the homeowner will usually make improvements to the home thus increasing the value over and above the increase caused by inflation, some features of the home deprecate lowering the value of the home and the homeowner has the option to borrow against the value of the home at several points during the course of ownership. We model the value of the home in terms of two linear stochastic differential equations and allow the noise to enter as a multiple of the state. The variables are categorized in Table 3.

$P(t)$	is the principal remaining on the loan at time t
P_0	initial purchase price
$V(t)$	the actual market value of the home at time t
$u_1(t)$	the amount of cash either paid toward the principal and interest or withdrawn from the home in terms of equity loans
$u_2(t)$	money applied to the value of the house in terms of improvements or taken form the value of the house through the sale of assets
$dZ(t)$	a Brownian motion processes representing inflation and fluctuation in interest
i	nominal annual interest rate

Table 1: Variables

We begin by considering the discrete time case. The fundamental law of payments can be expressed as the principal at time $n + 1$ is equal to the principal at time n plus interest drawn during the time period plus or minus the payment, i.e.,

$$P_{n+1} = (1 + i_n)P_n \pm Q_n$$

where P_n denotes principal, i_n denotes the interest in the nth period and Q_n denotes the nth payment. Now using the notation from the table we have

$$P(t + h) = P(t) + h(i(t)P(t) - u_1(t) + u_2(t)), \quad P(0) = P_0 \tag{3.10}$$

where $u_1(t)$ is rate of payment. The control $u_2(t)$ represents money borrowed from the bank to make improvements. This variable enters into the equation for the value of the house. We divide

the interest rate into two parts. The first is a nominal rate and the second is "random" motion of the rate. For the sake of simplicity we assume that this rate is equal to the rate of inflation. We can think of $i(t) = i + Z(t+h) - Z(t)$ where $Z(t+h) - Z(t)$ is an increment of Brownian motion.

The second part of the model is the market value of the home. Here we have

$$V(t+h) = V(t) + hu_2(t) + (Z(t+h) - Z(t))V(t), \quad V(0) = P_0 \tag{3.11}$$

We allow both u_1 and u_2 to be either positive or negative. If u_1 is negative we consider this to be money that is withdrawn from the value of the home as an equity loan. If u_2 is negative we consider this to be money that is achieved through the sale of assets associated with home. Any such money will be assumed to be applied to the principal. This is not a real restriction since we can immediately borrow this money in terms of an equity loan. We can pass to the limit to obtain a continuous time model.

For the continuation of this model we take the following equations describing the evolution of the value of the home.

$$
\begin{aligned}
dP(t) &= iP(t)dt + (u_2(t) - u_1(t))dt + P(t)dZ(t), \quad P(0) = P_0 \tag{3.12} \\
dV(t) &= V(t)dZ(t) + u_2(t)dt, \quad V(0) = P_0 \tag{3.13}
\end{aligned}
$$

In the first equation we see that $dZ(t)$ can denote the variation that occurs in interest rates. This is of course an approximation of reality. It is unlikely that interest rates would ever become negative and in most home loans there is a provision to bound from above the interest rate. The incorporation of these bounds complicates the problem and obscures the application that we wish to exhibit. In the second equation $dZ(t)$ denotes the effect of inflation on the value of the home. At one time it would not have seemed reasonable to have negative inflation. However during the 1990s and into the beginning of the 21 century we have seen negative inflation in Japan. To use the same rate for inflation and interest fluctuation is not entirely correct but for the purpose of an example to support the theory developed in this paper it is adequate.

The variable of primary interest is a linear combination of $P(t)$ and $V(t)$. We let

$$Y(t) = V(t) - P(t) \tag{3.14}$$

this variable represents the cash value of the home. A reasonable objective is to ask that $Y(t)$ should double during the thirty years of the standard loan. That is we state as one objective that

$$E\{Y(30)\} = 2P_0$$

where $E\{\cdot\}$ denotes expected value. In general we might expect that $y(t)$ should follow a curve such as

$$2P_0(e^{\beta t} - 1)$$

where the variable β is chosen so that $e^{30\beta} - 1 = 2$. We will for the purpose of the example ask that the following constraints be met.

$$E\{Y(10k)\} = 2P_0(e^{10k\beta} - 1), \quad k = 1, 2, 3 \tag{3.15}$$

We want to find an investment strategy that minimizes the financial activity as well the variance of the constraints (3.15). The problem that we consider is hence on the following form.

Problem 3.1.

$$\min_u E\left\{\int_0^{30}(u_1(t)^2 + u_2(t)^2)dt + \sum_{k=1}^{3} w_k \text{Var}^2(Y(10k))\right\}$$

subject to the constraints of equation (3.12), (3.13) and (3.15).

We consider the case when $w_k = 0.01$, $i = 0.075$ and $P_0 = 10^5$\$. In Figure 1 we show a stochastic simulation of the result. In the upper left plot we show a simulation of the cost value Y. The upper right plot shows an estimate of the expected value $E\{Y(t)\}$ obtained by averaging over 1000 simulations. The lower left plot shows the controls u_1, u_2 (u_1 in solid line and u_2 in dashed line) and the lower right plot shows the P and V (P in solid line and V in dashed line).
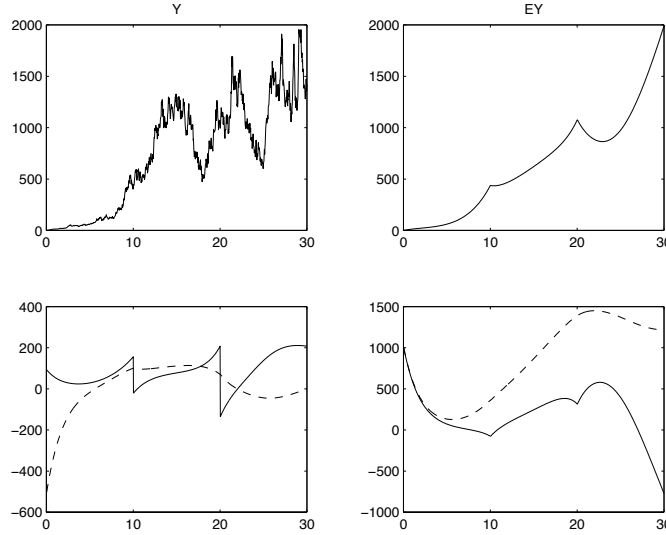

Figure 1: Stochastic simulation of the solution to the investment problem.

4   Path Planning for Mobile Robot

We consider the problem of steering a robot from rest at an initial condition $(-5, 1)$ to rest at the final position $(-1, 5)$ in such a way that it stays inside the corridor in the upper left part of Figure 2. We assume the linear dynamics

$$\ddot{y} = u + w$$

where w is a noise signal that take into account friction, irregularities in the floor, and other error sources. The linear dynamics can be obtained after feedback linearization of realistic robot models [7]. If we let the noise be modeled as a Brownian motion, then the state space realization of the robot dynamics becomes the stochastic system

$$dX = (AX + Bu)dt + G(X)dW$$
$$Y = CX$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad G(x) = \begin{bmatrix} 0 & 0 \\ x_2 & 0 \\ 0 & 0 \\ 0 & x_4 \end{bmatrix}.$$
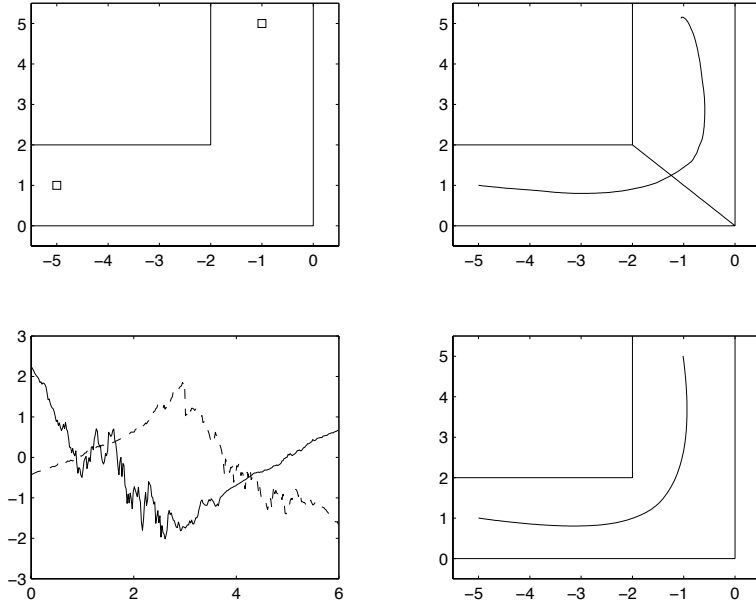
Let us use the design equation

Figure 2: The upper left figure shows the initial and final positions for the robot. The upper right figure shows one realization of the optimal path of the robot and the lower right figure shows the corresponding control signals, where u_1 corresponds to the solid line and u_2 is the dashed line. The weights $w_1 = 7$ and $w_2 = 1$ were used in the optimization problem (4.16). Finally, the lower right figure shows an estimate of the expected path obtained by averaging over 100 stochastic simulations.

$$\min E^{0,x_0} \left\{ w_1 |C_1 X(3) - \alpha_1|^2 + w_2 |C_3 X(6) - \alpha_3|^2 + \int_0^6 |u|^2 dt \right\}$$

$$\text{subj. to} \quad \begin{cases} dX = (AX + Bu)dt + G(X)dW, \ X(0) = x_0 \\ E^{0,x_0} \{C_2 X(3)\} = \alpha_2, \ E^{0,x_0} \{C_3 X(6)\} = \alpha_3 \end{cases} \tag{4.16}$$

where

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \alpha_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$C_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}, \quad \alpha_2 = 0$$

$$C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} -1 \\ 0 \\ 5 \\ 0 \end{bmatrix}.$$
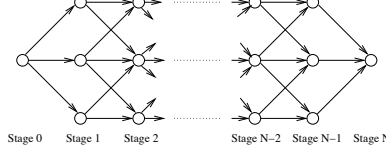
Stage 0   Stage 1   Stage 2   Stage N–2 Stage N–1 Stage N

Figure 3: We consider optimization of the path from a given initial state at stage 1 to a given node α_N at stage N. The path must pass trough one node in the graph at each stage.

The idea behind the optimization problem is to divide the control problem into two stages. First, we steer the robot to the switching surface $C_2 x = \alpha_2$ in such a way that the expected deviation from the point $C_1 x = \alpha_1$ is small. Then in the next stage we steer to the final position $x = \alpha_3$ in such a way that the variance of the deviation from this point is small. The integral cost is introduced to keep the expected control energy small. With the weights $w_1 = 7$ and $w_2 = 1$ we get the result in Figure 2. We see from the lower plot that the expected path of the robot stays well inside the corridor as desired. It is possible to push the trajectory further toward the middle of the corridor by adding an integral cost $E^{0,x_0} \left\{ \int_0^6 q|Cx - y_0(t)|^2 dt \right\}$, where $q \geq 0$ and $y_0(t)$ is some nominal trajectory near the middle of the corridor. The corresponding optimization problem still belongs to the problem class considered in this paper.

5   Route optimization

We can also consider trajectory planning problems where we at each time instant can choose to interpolate one out of several points. This is illustrated in Figure 3, where we have fixed initial and terminal nodes but can choose between several different routes between these two nodes. For consistency with the previous discussion we assume that the initial node is the initial state vector, i.e., $\alpha_0 = x_0$. At all other time instants we have a number of possible nodes, each corresponding to a subspace on the form $\{x \in \mathbb{R}^n : Cx = \alpha_{k_j}\}$. The nodes at each stage are thus defined by the set $S_k = \{\alpha_{k_1}, \dots, \alpha_{k_{N_k}}\}$. We define a set-valued function $\mathcal{N}_k : S_k \mapsto \mathcal{P}(S_{k+1})$ ($\mathcal{P}(S_{k+1})$ is the set of all subsets of S_{k+1}) that defines the possible successors of each node. This means that $\mathcal{N}_k(Cx_k) \subset S_{k+1}$ if $Cx_k \in S_k$ and $\mathcal{N}_k(Cx_k) = \emptyset$ otherwise. We can now generalize (2.1) and (2.2) to the case when we not only optimize a continuous time trajectory interpolating a given number of interpolation points (nodes) but also optimize over the set of nodes at each time instant. Let us consider

$$J_0(x_0) = \min_{u \in \mathcal{M}(t_0,t_f)} \sum_{k=0}^{N-1} \left(w_{k+1} \text{Var}^2\{CX(t_{k+1})\} + E^{t_0,x_0} \left\{ \int_{t_k}^{t_{k+1}} \sigma_k(t, X, u) dt \right\} \right)$$

$$\text{subject to} \quad \begin{cases} dX = (A_k X + B_k u)dt + G_k(X)dZ, \ t \in [t_k, t_{k+1}], \ X(0) = x_0 \\ E^{t_0,x_0}\{CX(t_{k+1})\} \in \mathcal{N}_k(E^{t_0,x_0}\{CX(t_k)\}), \quad k = 0, \dots, N-1 \end{cases}$$

(5.17)

For the investment example in the Section 3 this would mean that we can choose between a number of selected cash values in the intermediate time instances. If we let[1]

$$\mathcal{U}_k(x_k, \alpha_{k+1}) = \left\{ u \in \mathcal{M}(t_k, t_{k+1}) : \begin{cases} dX = (A_k X + B_k u)dt + G_k(X)dZ, & X(t_k) = x_k \\ E^{t_k, x_k}\{CX(t_{k+1})\} = \alpha_{k+1} \end{cases} \right\}$$

$$\mathcal{C}_k(u, \alpha_{k+1}) = \int_{t_k}^{t_{k+1}} \sigma_k(t, X, u)dt + w_{k+1}|CX(t_{k+1}) - \alpha_{k+1}|^2$$

then the dynamic programming iteration for (5.17) can be written as

$$J_k(x_k) = \min_{\alpha_{k+1} \in \mathcal{N}_k(Cx_k)} \min_{u \in \mathcal{U}_k(x_k, \alpha_{k+1})} E^{t_k, x_k}\{\mathcal{C}_k(u, \alpha_{k+1}) + J_{k+1}(X(t_{k+1}))\}.$$

Note that if $Cx_k \notin S_k$ then $\mathcal{N}_k(Cx_k) = \emptyset$ and the optimal objective is unbounded. The solution of the inner optimization can be constructed as in Proposition 2.3. We notice that this value function is completely determined by the node for the case when $C = I$, i.e. when each node is a point in the state space of the dynamical system. For this case the route optimization can be performed by discrete dynamic programming iteration over the discrete state space S_k at each stage. Once we know the optimal route, i.e., the optimal $\alpha_k \in S_k$, $k = 1, \ldots, N$, then we can construct the optimal control as in Proposition 2.3.

For the general case when C has full row rank $m < n$, then the problem is much harder because $J_k(x_k)$ is not completely determined by the choice of nodes at stage $k, \ldots, N$. The problem is then computationally very expensive for large N. One way to reduce the computational complexity is to discretize the state vector in each node. In this way a suboptimal solution can be obtained by discrete dynamic programming. In essense this means that we optimize over a finite collection of predefined trajectories, which is an approach that have been discussed for trajectory planning in, for example, [2, 6]. Another way to reduce the complexity is to prune the search tree in some effisient way in order to find an optimal (or a near optimal) sequence of nodes. In order to discuss this problem we introduce some new notation. We let $\bar{\alpha}_k = (\alpha_k, \ldots, \alpha_{N-1}, \alpha_N)$, where $\alpha_n \in S_n$, denotes an admissible node sequence from stage k to the terminal fixed node α_N, and let

$$\mathcal{A}_k = \left\{ \bar{\alpha}_k \in \Pi_{n=k}^N S_n : \alpha_{n+1} \in \mathcal{N}(\alpha_n) \right\}$$

be the set of all admissible node sequences from stage $k+1$ to the terminal node α_N. Similarly

$$\mathcal{A}_k(\alpha_k) = \left\{ \bar{\alpha}_{k+1} \in \Pi_{n=k+1}^N S_n : \alpha_{n+1} \in \mathcal{N}(\alpha_n) \right\}$$

is the set of all sequences of nodes from stage k to stage N with the kth node fixed to be α_k. In the same way as above, the value function in a particular node $\alpha_k = Cx_k$ at stage k can be computed as

$$J_k(x_k) = \min_{\bar{\alpha}_{k+1} \in \mathcal{A}_k(Cx_k)} J_k(x_k, \bar{\alpha}_{k+1})$$

We would like to remove as many node sequences as possible from $\mathcal{A}_k$ during the dynamic programming recursion. Let $\mathcal{A}_{k,O}(\alpha_k)$ denote sequences that are candidates for optimality (the index O

---

[1]In $\mathcal{C}_k(u, \alpha_{k+1})$ we implicitly assume that $X(t)$ satisfies the stochastic differential equation in the definition of $\mathcal{U}_k(x_k, \alpha_{k+1})$.

stands for the open candidates and is a standard notation). Then at each node during the dynamic programming recursion we take the following two steps

(i) $\mathcal{A}_{k,O}(\alpha_k) = \{(\alpha_{k+1}, \bar{\alpha}_{k+2}) : \bar{\alpha}_{k+2} \in \mathcal{A}_{k+1,O}(\alpha_{k+1}); \alpha_{k+1} \in \mathcal{N}(\alpha_k)\}.$

(ii) Remove branches from $\mathcal{A}_{k,O}(\alpha_k)$ that cannot correspond to the optimal node sequence.

This gives us the value function $J_k(x_k) = \min_{\bar{\alpha}_{k+1} \in \mathcal{A}_{k,O}(\alpha_k)} J_k(x_k, \bar{\alpha}_{k+1})$. The second step is not easy since each function $J_k(x_k, \bar{\alpha}_{k+1})$ depends on the particular node sequence $\bar{\alpha}_{k+1} \in \mathcal{A}_{k+1}$ but also on the the state vector at node k which depends on prior stages and in particular the initial condition. An alternative idea, discussed in for example [8, 9], is to replace step (ii) above with

$(ii)'$ Remove branches from $\mathcal{A}_{k,O}(\alpha_k)$ that are unlikely to be the optimal node sequence.

This generally give suboptimal solutions but the computational complexity may be reduced significantly. Some appealing and promising ideas for implementing $(ii)'$ was suggested in [8, 9]. There they consider optimization of switching sequences for LQR control of switched linear systems. This is mathematically closely related to the route optimization considered in this section and their results can be adapted to the problem discussed in this section.

References

[1] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 1995.

[2] M. . Branicky, T. . Johansen, I. Petersen and E. Frazzoli On-line Techniques for Behavioral Programming In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 1840-1845, Sydney, Australia, 2000.

[3] P. Crouch and J. W. Jackson. Dynamic interpolation for linear systems. In *Proc. 29th IEEE Conf. on Decision and Control*, pages 2312–2314, Hawaii, 1990.

[4] M. Egerstedt and C. Martin. Optimal trajectory planning and smoothing splines. *Automatica*, 37(7):1057–1064, July 2000.

[5] S.F. Ellermeyer, S.S. Pilyugin and R. Redheffer. Persistence Criteria for a Chemostat with variable Nutrient Input Journal of differential Equations, **171**,132–147 (2001)

[6] E. Frazzoli, M. A. Dahleh and E. Feron A Maneuver-Based Hybrid Control Architecture for Autonomous Vehicle Motion Planning In Software Enabled Control: Information Technology for Dynamical Systems, G. Balas and T. Samad editors, 2002. To appear.

[7] J. Lawton, R. Beard and B. Young. A Decentralized Approach To Formation Maneuvers To appear in *IEEE Transactions on Robotics and Automation*.

[8] B. Lincoln and B. Bernhardsson. Efficient pruning of search trees in LQR control of switched linear systems. In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 1828–1833, Sydney, Australia, 2000.

[9] B. Lincoln and A. Rantzer. Optimizing linear system switching. In *Proceedings of the 40th IEEE Conference on Decision and Control*, pages 2063–2068, Orlando, FLA, USA, 2001.

[10] B. Øksendal. *Stochastic Differential Equations, An Introduction with Applications*. Springer, Berlin Heidelberg, fifth edition, 1998.

[11] W. M. Wonham. On a matrix Riccati equation of stochastic control. *Siam J. Control Optim.*, 6(4):681–697, 1968.

[12] Z. Zhang, J. Tomlinson, and C. Martin. Splines and linear control theory. *Acta Math. Appl.*, 49:1–34, 1997.

Appendix A: Proof of Proposition 2.2

We use the compact notation introduced in (2.7) and (2.8) after the proposition. If we apply the Lagrange multiplier rule to (2.8) then we obtain

$$V(x_0, \alpha, t_0, t_f) = \max_{\lambda} \min_{u \in \mathcal{M}(t_0, t_f)} E^{t_0, \widehat{x}_0} \Big\{ \int_0^{t_f} [\widehat{X}^T \widehat{Q} \widehat{X} + 2\widehat{X}^T \widehat{S} u + u^T \widehat{R} u] dt$$
$$+ \widehat{X}(t_f)^T \widehat{Q}_0 \widehat{X}(t_f) + 2\lambda^T \widehat{C}(\alpha) \widehat{X}(t_f) \Big\}$$
$$\text{subject to} \quad d\widehat{X} = (\widehat{A}(t)\widehat{X} + \widehat{B}(t)u)dt + \widehat{G}(t, X)dZ, \quad \widehat{X}(t_0) = \widehat{x}_0$$

The solution to the inner optimization can be obtained from the Hamilton-Jacobi-Bellman Equation (HJBE)

$$-\frac{\partial V}{\partial t} = \min_{u \in \mathbb{R}^m} \left\{ \widehat{x}^T \widehat{Q} \widehat{x} + 2\widehat{x}^T \widehat{S} u + u^T \widehat{R} u + \frac{\partial V}{\partial \widehat{x}}^T (\widehat{A}\widehat{x} + \widehat{B}u) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} \right\}$$
$$V(x, t_f) = \widehat{x}^T \widehat{Q}_0 \widehat{x} + 2\lambda^T \widehat{C}(\alpha) \widehat{x}$$

where

$$a_{ij} = (\sigma\sigma^T)_{ij} = \left((\sum_{k=1}^n x_k G_k)(\sum_{l=1}^n x_l G_l^T) \right)_{ij} = \sum_{k,l=1}^n x_k x_l G_{ki} G_{lj}^T,$$

where G_{ki} is the ith row of G_k. With the value function $V(\widehat{x}, t) = \widehat{x}^T \bar{P}(t)\widehat{x}$ we get

$$\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{1}{2} \sum_{k=1}^n \sum_{l=1}^n x_k x_l \text{tr}(G_k^T P G_l) = \widehat{x}^T \widehat{\Pi}(\bar{P})\widehat{x},$$

If we plug the ansatz $V(\widehat{x}, t) = \widehat{x}^T \bar{P}(t)\widehat{x}$ into the HJBE we get the optimal control $u = -\widehat{R}^{-1}(\widehat{B}^T \bar{P} + \widehat{S})\widehat{x}$, and that the following Riccati equation must hold

$$\dot{\bar{P}} + \widehat{A}^T \bar{P} + \bar{P}\widehat{A} + \widehat{Q} + \widehat{\Pi}(\bar{P}) = (\bar{P}\widehat{B} + \widehat{S})\widehat{R}^{-1}(\bar{P}\widehat{B} + \widehat{S})^T,$$

with boundary condition $\bar{P}(t_f) = \widehat{Q}_0 + \lambda^T \widehat{C} e_{n+1} + e_{n+1}^T \widehat{C}^T \lambda$. The optimal cost becomes

$$V(x_0, \alpha, t_0, t_f) = \max_\lambda \widehat{x}_0^T \bar{P}(\lambda, t_0) \widehat{x}_0 \tag{5.18}$$

To perform the optimization with respect to λ we need to obtain an explicit expression of $\widehat{P}$ as a function of λ. It turns out that we can use

$$\bar{P} = \widehat{P} + \widehat{N} \lambda e_{n+1}^T + e_{n+1} \lambda^T \widehat{N}^T - \lambda^T \widehat{W} \lambda e_{n+1} e_{n+1}^T$$

where $\widehat{P}, \widehat{N}$ and $\widehat{W}$ are given in (2.7). This follows since

$$\begin{aligned}
\bar{P}(t_f) &= \widehat{P}(t_f) + \widehat{N}(t_f) \lambda e_{n+1}^T + e_{n+1} \lambda^T \widehat{N}(t_f)^T + \lambda^T \widehat{W}(t_f) \lambda e_{n+1} e_{n+1}^T \\
&= \widehat{Q}_0 + \lambda^T \widehat{C} e_{n+1} + e_{n+1}^T \widehat{C}^T \lambda
\end{aligned}$$

and

$$\begin{aligned}
\dot{\bar{P}} &= -\widehat{A}^T \widehat{P} - \widehat{P} \widehat{A} - \widehat{Q} - \widehat{\Pi}(\bar{P}) + (\widehat{P}\widehat{B} + \widehat{S})\widehat{R}^{-1}(\widehat{P}\widehat{B} + \widehat{S})^T - (\widehat{A} - \widehat{B}\widehat{R}^{-1}(\widehat{B}^T \widehat{P} + \widehat{S}^T))^T \widehat{N} \lambda e_{n+1}^T \\
&\quad - e_{n+1} \lambda^T \widehat{N}^T (\widehat{A} - \widehat{B}\widehat{R}^{-1}(\widehat{B}^T \widehat{P} + \widehat{S}^T)) - \lambda^T \widehat{N}^T \widehat{B}\widehat{R}^{-1}\widehat{B}^T \widehat{N} \lambda e_{n+1} e_{n+1}^T \\
&\qquad\qquad\qquad = -\widehat{A}^T \bar{P} - \bar{P}\widehat{A} - \widehat{Q} - \widehat{\Pi}(\widehat{P}) + (\bar{P}\widehat{B} + \widehat{S})\widehat{R}^{-1}(\bar{P}\widehat{B} + \widehat{S})^T
\end{aligned}$$

which follows since

$$\begin{aligned}
\widehat{A}^T(e_{n+1} \lambda^T \widehat{N}^T + \lambda^T \widehat{W} \lambda e_{n+1} e_{n+1}^T) &= 0, \\
\widehat{B}^T(e_{n+1} \lambda^T \widehat{N}^T + \lambda^T \widehat{W} \lambda e_{n+1} e_{n+1}^T) &= 0.
\end{aligned}$$

The optimization in (5.18) thus becomes

$$\begin{aligned}
\max_\lambda \widehat{x}_0^T (\widehat{P}(t_0) &+ \widehat{N}(t_0) \lambda e_{n+1}^T + e_{n+1} \lambda^T \widehat{N}(t_0)^T - \lambda^T \widehat{W}(t_0) \lambda e_{n+1} e_{n+1}^T) \widehat{x}_0 \\
&= \widehat{x}_0^T (\widehat{P}(t_0) + \widehat{N}(t_0) \widehat{W}(t_0)^{-1} \widehat{N}(t_0)^T) \widehat{x}_0
\end{aligned}$$

and the optimal Lagrange multiplier is $\lambda = \widehat{W}(t_0)^{-1} \widehat{N}(t_0)^T \widehat{x}_0$. Finally, the optimal control becomes

$$\begin{aligned}
u &= -\widehat{R}^{-1}(\widehat{B}^T(\widehat{P} + \widehat{N}\widehat{W}(t_0)^{-1}\widehat{N}(t_0)^T \widehat{x}_0 e_{n+1^T}) + \widehat{S})\widehat{x} \\
&= -\widehat{R}^{-1}(\widehat{B}^T \widehat{P} + \widehat{S})\widehat{x} - \widehat{R}^{-1}\widehat{B}^T \widehat{N}\widehat{W}(t_0)^{-1}\widehat{N}(t_0)^T \widehat{x}_0
\end{aligned}$$

Considering the optimal control problem from the "initial point" $(t, x(t))$ gives the following feedback formulation of the optimal control $u = -\widehat{R}^{-1}(\widehat{B}^T(\widehat{P} + \widehat{N}\widehat{W}^{-1}\widehat{N}^T) + \widehat{S}^T)\widehat{X}$. We note that under the conditions of the proposition there exists solutions $\bar{P}$ and $\widehat{P}$ to the Riccati equations that are involved in the proof. Indeed, the special structure of the system matrices imply that only the upper left blocks of $\bar{P}$ and $\widehat{P}$ satisfy nonlinear equations, which are identical to the first equation in (2.6). The other blocks corresponds to p and ρ in (2.6). Hence, the existence of $\bar{P}$ and $\widehat{P}$ follows from Remark 2.1.