# Stochastic Feedback Controller Design Considering the Dual Effect

Jinwhan Kim* and Stephen Rock†

*Aerospace Robotics Laboratory, Stanford University, Stanford, CA, 94305, USA*

**A new approach for the problem of nonlinear stochastic control is presented. A typical procedure for stochastic control takes a two step approach by separating estimation and control - estimate the system state using noise corrupted measurements, and control the system based on the estimated state. However, for nonlinear stochastic control systems and adaptive control systems with uncertain parameters, control and estimation are often not separable, since control inputs can affect not only the system state, but also the quality of state estimation. Dynamic programming and search-based methods are general solution techniques for solving this type control problem, however, those solution techniques are often infeasible for real-time applications due to the huge computational requirements even for small problems. In this paper, a new suboptimal control algorithm is presented, which can be implemented as an online feedback controller for a stochastic control problem of moderate dimension. For an application example, automated docking of a nonholonomic vehicle using a bearing-only sensor is considered. The results from a series of numerical simulations are shown.**

## I. Introduction

Most dynamic systems operate in the presence of uncertainties such as environmental disturbance and measurement noise. Typical control strategies are based on the certainty equivalence (CE) assumptions that the choice of control actions does not affect nor is it affected by the estimation uncertainty. This leads to a two-step approach - estimate the state of the dynamic system from noise corrupted measurements using a standard state estimator such as the Kalman filter, and calculate control inputs assuming that those estimated states are true, using a conventional control technique which is valid for deterministic systems. This approach often works quite well, and is in fact optimal in the sense of minimum mean square error (MMSE) if the system is observable and the LQG assumptions apply. That is, the system is *linear*, the cost criterion is *quadratic*, and the noise is *Gaussian*. However, the assumptions may not hold for a control system that involves nonlinearities, or for an adaptive control system that contains uncertain system parameters in its dynamics.

To achieve better performance for these types of control systems, so-called *dual effects* need to be considered. This concept was discussed by Feldbaum.[1] He pointed out that the control input has two purposes that may conflict with each other. One is to achieve the control goal of system stabilization, and the other is to help learning about any unknown parameters and/or the state of the system.[2] When these conflicting control goals are closely coupled and not separable, the quality of estimation affects the quality of control, and vice versa. Dynamic programming and search-based approaches are the general solution techniques for solving this type of control problem. However, in many cases, those solution techniques are practically of little use, since the computational burden even for a small problem is commonly prohibitive.[3,4]

In this study, a new stochastic control algorithm is suggested that incorporates a cost incurred by the system uncertainty into the performance index and applies a linear quadratic control technique, which provides a *linear* feedback control law with the *quadratic* performance index considering the *dual* features -

*Ph.D. Candidate, Department of Aeronautics and Astronautics, Durand Building, 496 Lomita Mall, Stanford, CA 94305, Student Memeber AIAA.

†Professor, Department of Aeronautics and Astronautics, Durand Building, 496 Lomita Mall, Stanford, CA 94305, Fellow AIAA.

this will be abbreviated hence as LQD. It can be implemented as a real-time controller even for a stochastic control problem of moderate dimension through a systematic design procedure.

The performance of this new approach is demonstrated by considering the docking control of a nonholonomic vehicle using a bearing only sensor. The problem has two competing control objectives - the vehicle needs to keep its line-of-sight (LOS) angle close to zero, and at the same time, must swerve from the LOS direction to improve observability of the target location. For this type of problem, CE controllers perform poorly. (e.g. when the vehicle is initially aligned to the LOS line to the target) A series of simulations are performed to check the validity of the suggested approach.

## II. Problem Statement : Optimal Stochastic Control

The general problem considered in this paper is a nonlinear stochastic system described by following equations of states, and measurements

$$\dot{x} = f(x,u) + w \tag{1}$$

$$z = h(x) + v \tag{2}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the control input vector, and $z \in \mathbb{R}^o$ is the measurement vector. $w \in \mathbb{R}^n$, $v \in \mathbb{R}^o$ are the process noise, and the measurement noise vector, respectively. Both of them are random sequences with known probability distributions.

For stochastic optimization, the performance index to be minimized is defined as

$$J = E\Big[ \int_0^{t_f} L(x,u;t)\; dt \Big] \tag{3}$$

where $L$ is a loss function, and the expectation is taken with respect to all underlying random variables. For convenience of mathematical description and numerical implementation, the discrete-time version of Eq. (3) is generally preferred. It has the form

$$J = E\Big[ L_N(x_N) + \sum_{k=0}^{N-1} L_k(x_k,u_k) \Big] \tag{4}$$

The solution to this problem is of the form $u_k^*$ for $k=0,1,\cdots,N-1$, where $u_k^*$ is a control input sequence that minimizes Eq. (4). In principle, this problem can be solved as a dynamic programming problem. That is, let $V_k$ be the minimum cost-to-go from $k$. Then the principle of optimality gives the backward recursive equation called the Bellman equation.

$$V_k = \min_{u_k} \Big\{ E\Big[ L_k(x_k,u_k) + V_{k+1} \;\Big|\; \mathcal{I}_k \Big] \Big\} \tag{5}$$

$$\text{for } k = N-1,\, N-2\,, \cdots ,\, 0$$

with the boundary condition given by

$$V_N = E\Big[ L_N(x_N,u_N) \;\Big|\; \mathcal{I}_N \Big] \tag{6}$$

Note that the expectation is conditioned on $\mathcal{I}_k$ in Eq. (5) which represents currently available information including the set of measurements and control inputs up to $k$. The optimal control $u_k^*$ can be obtained by solving this dynamic programming problem formulated as Eqs. (5) and (6). However, numerical difficulties due to *the curse of dimensionality* make this problem practically unsolvable even for simple cases.[3]

## III. Previous Approaches

Since Feldbaum first discussed the concept of dual control in his seminal papers in 1960 and 1961, a considerable amount of research has been performed on this topic. There are some simple problems for which optimal solutions have been obtained. A four state Markov chain probelm was solved by Sternby,[5] and a first-order system with two possible gain values was treated by Bernhardsson.[6] Numerical solutions for some very simple adaptive dual control problems have also been calculated.[7] However, they are exceptional cases, and problems of a reasonable size are practically unsolvable either analytically or numerically.

Various suboptimal techniques that posses some dual features have also been developed for adaptive and/or stochastic control applications. For more specific physical applications, missile interception using angle-only measurements has attracted some attention, and several dual-control guidance laws have been suggested.[8,9] Those techniques provide approximate solutions by adding perturbation signals, constraining the variance of the estimates, modifying the cost function, approximating a value function with a finite parameter sets, etc.[a] There are two relative common and intuitive approaches among them. The first is to add a *probing input*, $u_a$ to improve the observability of the system state. That is

$$u_{new} = u_o + u_a \tag{7}$$

where $u_0$ is the solution to a deterministic or cautious control problem and $u_a$ is an input that is designed separately based on a set of heuristics. The advantage of this approach is that it is simple. The disadvantage is that there is no systematic way of designing the $u_a$, and the overall performance depends heavily on the heuristics chosen.

The second method is to form a cost function that is the weighted sum of a *standard* cost $J_0$ (e.g. linear quadratic control) and a term that penalizes the system uncertainty, $J_a$. That is

$$J_{new} = J_o + \lambda J_a \tag{8}$$

Often a scalar function of the error covariance is used for $J_a$. The advantage of this approach is that it lends itself well to numerical search techniques. However, solutions are typically limited since the search space grows exponentially with respect to look-ahead time. Also, there is no systematic means of choosing the weighting parameter, $\lambda$, and the resulting cost function might have little relevance to the true cost defined in Eq. (3).

# IV. Suboptimal Stochastic Control

Proposed here is a suboptimal technique that offers an efficient and systematic method to generate an augmented cost function and to calculate probing inputs while minimizing a quadratic form of the cost defined in Eq. (3). The approach yields a feedback control solution of the form

$$u = -K_x\hat{x} - K_y y \tag{9}$$

where $\hat{x}$ is the estimated state calculated using a standard extended Kalman filter (EKF), and $y$ is a new state vector associated with the system uncertainties. The term $K_y y$ provides the inputs required to reduce the uncertainty, while the term $K_x\hat{x}$ drives the system to its desired state.

Details of this approach are presented below. Section A develops the performance index. Section B defines an augmented state vector, $\xi$ which includes the system state and and *uncertainty state*, and rewrites the performance index in terms of this new state vector. Sections C and D describe how the system state and the uncertainty state are propagated. Section E summarizes the resulting control logic.

## A. Performance Index

To define the quadratic cost function that will be used for optimization, consider the random variable $x$ to be represented as a cloud of particles $x_i, i = 1, 2, \cdots, N$ which are distributed in state space as in Figure 1. Given this the cost can be calculated as the sum of sequences of the distances between the desired states and each particles. The cost reflects both the distance between those particles and the desired state, and how severely scattered they are. The quality of control at a particular time can be expressed using the following quadratic expression.

$$\begin{aligned} \delta J &= \frac{1}{N}\sum_{i=1}^{N}\Big\{(x_i - x^d)^T Q_c (x_i - x^d) + u^T R_c u\Big\} \\ &= E\Big[(x - x^d)^T Q_c (x - x^d) + u^T R_c u\Big] \end{aligned} \tag{10}$$

[a]The overview of various suboptimal techniques are available in Ref. 10 and 11.

where $x^d$ is a desired state. $Q_c$ and $R_c$ are weighting matrices. The average operator $E[\,\cdot\,]$ works on samples only, not on time. Equation (10) can be rewritten as follows.

$$\begin{aligned} \delta J &= (\hat{x}-x^d)^T Q_c(\hat{x}-x^d) + u^T R_c u + E\left[(x-\hat{x})^T Q_c (x-\hat{x})\right] \\ &= (\hat{x}-x^d)^T Q_c(\hat{x}-x^d) + u^T R_c u + tr\{Q_c P_e\} \end{aligned} \tag{11}$$

with $E[x] = \hat{x}$, and $P_e = \tilde{x}\tilde{x}^T = (\hat{x}-x)(\hat{x}-x)^T$. The expression above is in a familiar quadratic form, and it allows that existing mathematical tools to be used with minor modification. The error covariance $P_e$ is a positive semi-definite matrix, thus there exists a matrix $S$ that satisfies $P_e = S\,S^T$. However, matrix square-roots are not unique in general.[b] To secure uniqueness, $S$ is kept as a lower triangular matrix whose diagonal elements are positive. By introducing a simplifying assumption that $Q_c$ is a diagonal matrix,[c] a more compact expression can be obtained. The last term in Eq. (11) is now expressed as

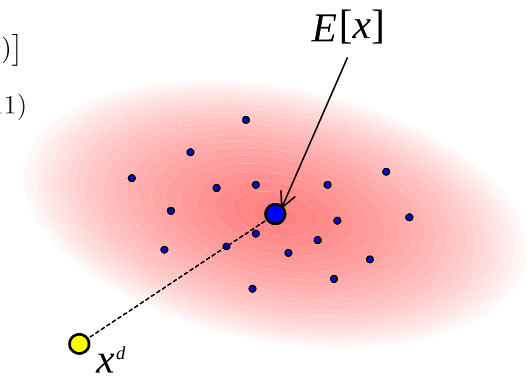



**Figure 1. A cloud of particles representing a distribution of the random variable $x$ at a specific time**

$$tr\{Q_c P_e\} = tr\{Q_c\,S\,S^T\} = \sum_{j=1}^{n}\left\{q_{jj}\left\{\sum_{k=1}^{j} s_{jk}^2\right\}\right\} \tag{12}$$

where $Q_e(i,i) = q_{ii}$ and $S(i,j) = s_{ij}$. Equation (11) can now be rewritten as

$$\delta J = (\hat{x}-x^d)^T Q_c(\hat{x}-x^d) + u^T R_c u + \sum_{j=1}^{n}\left\{q_{jj}\left\{\sum_{k=1}^{j} s_{jk}^2\right\}\right\} \tag{13}$$

This is the quadratic function of the estimated state deviation $(\hat{x}-x^d)$, the control input $u$, and the new variable $s$ associated with the square-root of the error covariance matrix. Finally, by integrating Eq. (13) along time, we get the performance index to be minimized for stochastic optimization.

$$J = \int_0^{t_f}\left\{(\hat{x}-x^d)^T Q_c(\hat{x}-x^d) + u^T R_c u + \sum_{j=1}^{n}\left\{q_{jj}\sum_{k=1}^{j} s_{jk}^2\right\}\right\} dt \tag{14}$$

As a limiting case, when perfect estimation is achieved under no uncertainty ( $\hat{x}\to x,\ s_{ij}\to 0$ ), Eq. (14) reduces to the quite familiar quadratic performance index for the deterministic system.

$$J = \int_0^{t_f}\left\{(x-x^d)^T Q_c(x-x^d) + u^T R_c u\right\} dt \tag{15}$$

### B. State Augmentation

The integrand of the performance index in Eq. (14) involves the sum of two groups of terms. The first is the quadratic function of the state deviation and the control input consumption, which is quite familiar from linear quadratic control. The second involves the matrix elements of the square-root of the error covariance matrix, $S$. By representing the elements in a vector form, a simpler expression for the performance index can be obtained, which is advantageous for subsequent mathematical operations. The matrix $S$ has been set to be in a lower triangular form as

$$S = \begin{bmatrix} s_{11} & 0 & 0 & \cdots & 0 \\ s_{21} & s_{22} & 0 & \cdots & 0 \\ s_{31} & s_{32} & s_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ s_{n1} & s_{n2} & s_{n3} & \cdots & s_{nn} \end{bmatrix} \tag{16}$$

---

[b] When $S$ is a square-root matrix of $P$, then so is $ST$. Only if $T$ is an arbitrary orthonormal matrix.

[c] This is not a required assumption. $Q_c$ only needs to be real symmetric positive semi-definite.

Let us define $y$ as the single column vector that contains all the nonzero elements in $S$, and call this new vector an uncertainty state vector, since it contains the information on the internal status of the system uncertainty.

$$y = [s_{11} \quad s_{21} \quad s_{22} \quad s_{31} \quad \cdots \quad s_{nn}]^T \tag{17}$$

where $y \in \mathbb{R}^{n(n+1)/2}$. The quadratic performance index in Eq. (14) can be expressed as shown below. For descriptive convenience, the desired state $x^d$ is set to zero hereafter without loss of generality.

$$J = \int_0^{t_f} \left\{ \hat{x}^T Q_c \hat{x} + u^T R_c u + y^T W_c y \right\} dt \tag{18}$$

where

$$W_c = \begin{bmatrix} q_{11} & 0 & 0 & 0 & \cdots & 0 \\ 0 & q_{22} & 0 & 0 & \cdots & 0 \\ 0 & 0 & q_{22} & 0 & \cdots & 0 \\ 0 & 0 & 0 & q_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & q_{nn} \end{bmatrix} \tag{19}$$

and $W_c \in \mathbb{R}^{\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}}$. If $Q_c$ is not a diagonal matrix, neither is $W_c$. However, it still can be represented as a quadratic expression shown in Eq. (18).[d]

A more compact expression can be obtained through state augmentation. Denote the augmented state vector as

$$\xi = \begin{bmatrix} x \\ y \end{bmatrix} \tag{20}$$

where $\xi \in \mathbb{R}^{n(n+3)/2}$. The dimension $n(n+3)/2$ corresponds to the minimum number of states for representing the stochastic system under the Gaussian distribution assumption. The system has $n$ state variables, and the error covariance matrix which is symmetric has $n(n+1)/2$ independent variables. Equation (18) can be rewritten as follows.

$$J = \int_0^{t_f} \left\{ \xi^T Q_a \xi + u^T R_c u \right\}_{x=\hat{x}} dt \tag{21}$$

where

$$Q_a = \left[ \begin{array}{c|c} Q_c & 0 \\ \hline 0 & W_c \end{array} \right]$$

## C. Extended Kalman Filter

The control law defined in Eq. (9) requires the propagation of states $\hat{x}$ and $y$. To calculate $\hat{x}$, the EKF is used here (alternative derivations are possible) because nonlinearities are assumed to be present in the state or/and measurement equation. The random sequences $w$ in Eq. (1) and $v$ in Eq. (2) are assumed to follow zero mean Gaussian distributions given by $w \sim \mathcal{N}(0, Q_e)$ and $v \sim \mathcal{N}(0, R_e)$. The differential equations for updating state estimate $\hat{x}$ and error covariance matrix $P_e$ are

$$\dot{\hat{x}} = f(\hat{x}, u) + P_e H^T R_e^{-1} \big(z - h(\hat{x})\big) \tag{22}$$

$$\dot{P}_e = F P_e + P_e F^T + Q_e - P_e H^T R_e^{-1} H P_e \tag{23}$$

where

$$F = \left. \frac{\partial f(x,u)}{\partial x} \right|_{x=\hat{x}}$$

$$H = \left. \frac{\partial h(x)}{\partial x} \right|_{x=\hat{x}}$$

[d] According to the derivation, the weighting matrix $W_c$ is determined once $Q_c$ is given. However, the independency between those two matrices may provide additional degrees of freedom for a controller designer. For example, more weight can be put on estimation rather than on control, or vice versa.

Note : Since uncertainties are associated with the square-root of the error covariance matrix, the use of the square-root Kalman filter algorithm for the state estimation is appropriate. It is algebraically equivalent to the standard Kalman filter. Details are given in the next section.

### D. Propagation of the Uncertainty States

To propagate $y$, note that the first-order time derivative of the augmented state state vector is

$$\dot{\xi} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \triangleq \begin{bmatrix} f \\ g \end{bmatrix} = \eta(\xi, u) \tag{24}$$

where

$$g = [\dot{s}_{11} \quad \dot{s}_{21} \quad \dot{s}_{22} \quad \dot{s}_{31} \quad \cdots \quad \dot{s}_{nn}]^T$$

and

$$\dot{s}_{ij} = \dot{S}(i,j)$$

As seen above, $g$ is simply an enumeration of the first-order time derivatives of the square-root error covariance matrix. With the definition $P = SS^T$, Eq. (23) becomes

$$\begin{aligned} \dot{P}_e &= \dot{S}S^T + S\dot{S}^T \\ &= \Big(F - \frac{1}{2} SS^T H^T R_c^{-1} H\Big) SS^T + \frac{1}{2} Q_c S^{-T} S^T \\ &\quad + SS^T \Big(F^T - \frac{1}{2} H^T R_c^{-1} H SS^T\Big) + \frac{1}{2} S^{-T} S^T Q_c \end{aligned} \tag{25}$$

The matrix differential equation that is satisfied by the particular solution is given by

$$\dot{S}_e = \Big(F - \frac{1}{2} SS^T H^T R_c^{-1} H\Big) S + \frac{1}{2} Q_c S^{-T} \tag{26}$$

However, careful treatment is required when propagating $S$ along time, since $\dot{S}$ may not be lower triangular even though $S$ is. An efficient method for computing $\dot{S}$ in a lower triangular form has been suggested by Morf et al.,[12] which guarantees the uniqueness of $S$. $\dot{S}$ in a lower triangular can be expressed in the following form.

$$\dot{S} = S\,[M]^{+/2} \tag{27}$$

where

$$\begin{aligned} M &\triangleq S^{-1}\big(\dot{S}S^T + S\dot{S}^T\big)S^{-T} \\ &= S^{-1}\dot{S} + \dot{S}^T S^{-T} \\ &= S^{-1}FS + S^T F^T S^{-T} - S^T H^T R_c^{-1} H\, S + S^{-1} Q_c\, S^{-T} \end{aligned} \tag{28}$$

and the operator $[\,\cdot\,]^{+/2}$ is defined by

$$[\,U\,]^{+/2}_{ij} = \begin{cases} \frac{1}{2} u_{ij} & \text{for } i = j \\ u_{ij} & \text{for } i < j \\ 0 & \text{otherwise} \end{cases}$$

for an arbitrary matrix $U$.

Because the uncertainty states are obtained from the square-root of the error covariance matrix, it is more effective and convenient to use the square-root filter formulation than the standard Kalman filter formulation. This substitution brings a couple of advantages. First, the increase in computational load is minimal, because few redundant filter-related calculations are required. Second, the square-root filter is known to have better properties in accuracy and stability than the standard Kalman filter.[13]

### E. Suboptimal Controller

Finally, the optimization problem can be summarized as :

Find the optimal control input history $u^*(t)$ for $0 \leq t < t_f$ that satisfies

$$u^* = \underset{u}{\operatorname{argmin}} \int_0^{t_f} \left\{ \xi^T Q_a \xi + u^T R_c u \right\} dt \tag{29}$$

subject to

$$\dot{\xi} = \eta(\xi, u) \tag{30}$$

Due to the approximation involved in the covariance propagation step and the complexity of its state-space representation, it is a difficult task to calculate the optimal solution to the problem. Several techniques are available to obtain suboptimal solutions for the nonlinear optimal control problem. In this study, state-feedback control using an infinite horizon steady-state gain is used. Suppose that the original differential Riccati equation has a limiting solution $P_c$ which satisfies the algebraic Reccati equation (ARE) shown below.

$$0 = A^T P_c + P_c A + Q_a - P_c B R_c^{-1} B^T P_c \tag{31}$$

where

$$\begin{aligned} A &= \left. \frac{\partial \eta(\xi, u)}{\partial \xi} \right|_{x=\hat{x}} \\ B &= \left. \frac{\partial \eta(\xi, u)}{\partial u} \right|_{x=\hat{x}} \end{aligned}$$

To derive matrices $A$ and $B$ analytically may be complicated and cumbersome, and it would generally be convenient to calculate a numerical Jacobian instead. By solving Eq. (31), the steady state Kalman control gain matrix is obtained, which is given by

$$K = R_c^{-1} B^T P_c \tag{32}$$

The gain matrix $K$ is calculated at every time step, and this provides a time-varying controller. The suboptimal control law is state feedback

$$u = K\,(\xi - \xi_{ss})\big|_{x=\hat{x}} = K_x\,\hat{x} + K_y\,(y - y_{ss}) \tag{33}$$

where

$$K = [\,K_x \quad K_y\,]$$

and $y_{ss}$ is obtained from the steady-state solution of Eq. (23). The equation above is composed of two distinctive feedback terms. The first feedback term is mainly for achieving the control goal of stabilization. This is the typical feedback control law for the deterministic system as it is. The second term provides an additional feedback input driven by the uncertainty state. This additional input decreases uncertainties of the control system in a direct manner, and improves the control performance indirectly. Consequently, by taking the dual features into consideration at the same time, the overall performance of the stochastic control system is expected to be improved.

## V. Application Example : Docking Using a Bearing-Only Sensor

For the system configuration, a nonholonomic vehicle with a bearing-only sensor is considered. The control goal is to drive a vehicle to a given target with a sufficient certainty and accuracy of the vehicle's relative position to the target.

## A. State-Space Representation

The state-space representation of a nonholonomic vehicle with a bearing only sensor in polar coordinates[e] is represented as

$$\dot{x} = \begin{bmatrix} \dot{\rho} \\ \dot{\gamma} \\ \dot{\psi} \\ \dot{V} \end{bmatrix} = \begin{bmatrix} -V\cos\gamma \\ (V/\rho)\sin\gamma - u_1 \\ u_1 \\ u_2 \end{bmatrix} + w \quad (34)$$

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \psi \\ \gamma \end{bmatrix} + v \quad (35)$$

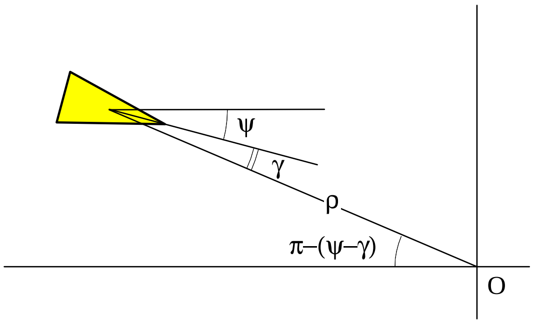


**Figure 2. Polar coordinates**

where $\rho$ and $\gamma$ are the relative range and bearing, respectively. $\psi$ is the yaw angle, and $V$ is the longitudinal velocity of the vehicle in the body-fixed coordinate frame. $u_1$ is a yaw rate control input, and $u_2$ is a linear acceleration control input which is set to zero here for simplifying the problem. The vehicle is assumed to have a curvature constraint, and this limits the magnitude of the yaw rate command input as $|u_1| \leq V/R_{min}$, where $R_{min}$ is the minimum turning radius of the vehicle.

## B. Simulation Results : CE controller vs. LQD controller

The performance of the CE controller is compared with that of the LQD controller through numerical simulations. Those two controllers are designed through basically the same procedure except for their performance indices to be minimized. For the LQD controller, Eq. (21) with the given $Q_c$ and its corresponding $W_c$ is considered. Whereas, for the CE controller, the same performance index with the degenerate case with the same $Q_c$ and $W_c = 0$ is considered, which can also be interpreted as the case that the estimation is perfect.

For the simulations, the vehicle is assumed to start approaching a given distance away from the target, and the onboard bearing-only sensor is initially aligned to the target feature. A coarse initial position estimate and the corresponding error covariance are given to the state estimator. The simulation settings are $V = 1$, $R_{min} = 2$, $x_0 = [10,\ 0,\ 0,\ 1]^T$, $P_0 = diag\{5^2,\ (3°)^2,\ (3°)^2,\ 0.1^2\}$, $Q_c = diag\{9,\ 9,\ 1,\ 0\}$, $R_c = diag\{100,\ 100\}$, $Q_e = diag\{1e-6,\ 1e-6,\ 0,\ 0\}$, and $R_e = diag\{0.0012,\ 0.0012\}$.

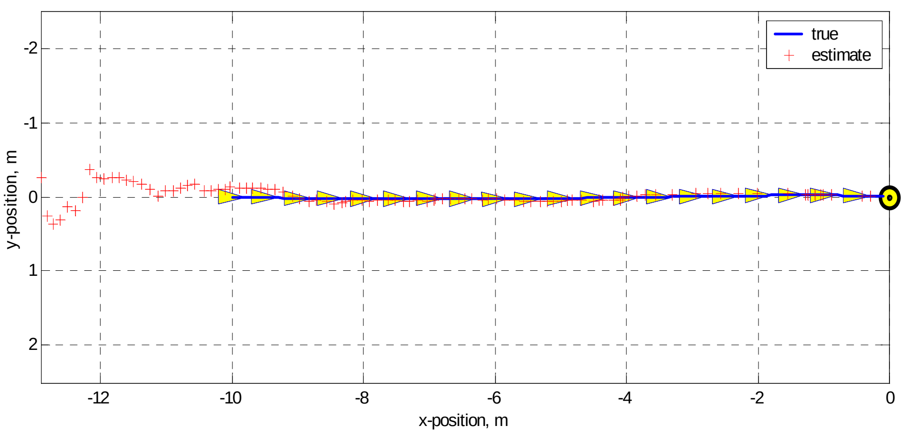
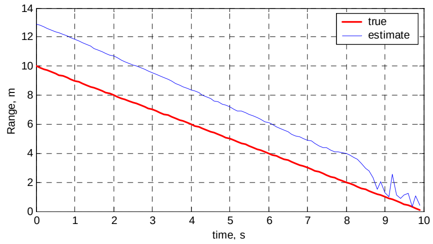
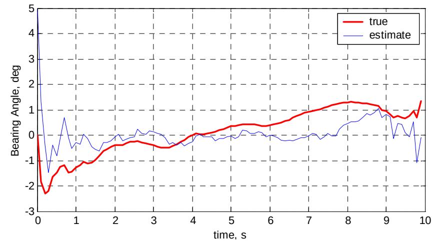


**Figure 3. CE control : Vehicle trajectory**

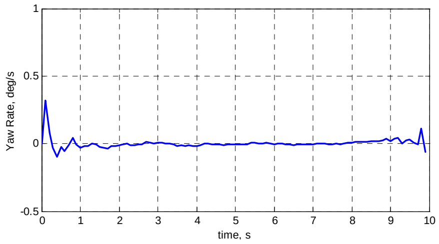
[e]For the nonlinear estimation problems associated with range and/or bearing measurements, it is known that the estimation in the polar coordinate system is better in accuracy and stability than that in the Cartesian coordinate system.



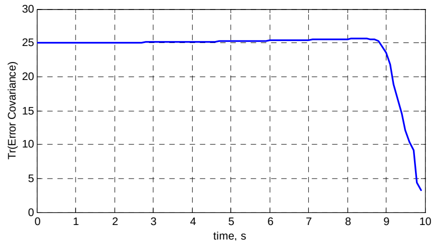

(a) Range : $\rho$




(b) Bearing angle : $\gamma$




(c) Yaw Rate control input : $u_2$




(d) Uncertainty measure : $Tr(P_e)$

**Figure 4. CE control : Time history of range, bearing angle, control input, and uncertainty measure**

Figures 3 and 4 present results for the CE controller. The bearing angle $\gamma$ can be estimated fairly accurately, because it is directly measured. However, little improvement is observed for the range estimate until the vehicle gets very close to the target. The sudden drop of the range estimate error occurs about a meter before the collision. This does not allow sufficient time nor space to take any subsequent actions for preventing collision, and may cause operational difficulties especially for this type of nonholonomic vehicle which has limited maneuverability.

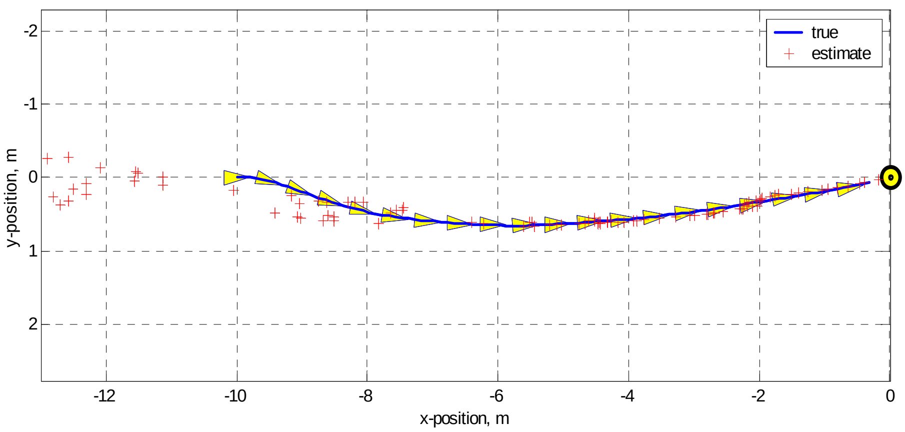
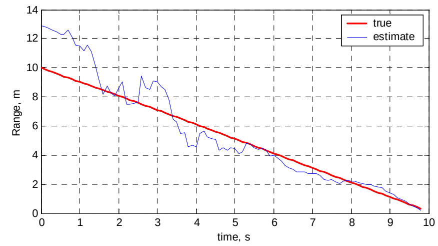
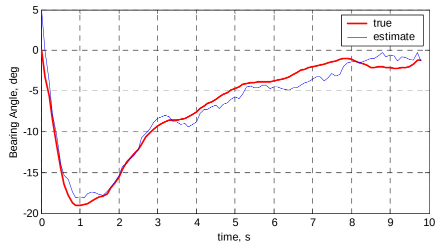



**Figure 5. LQD control : Vehicle trajectory**

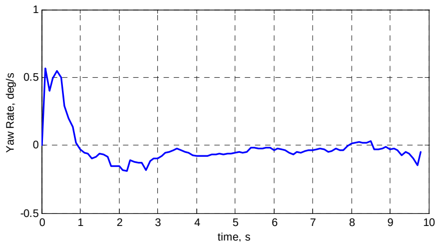
The results for the LQD controller are shown in Figures 5 and 6. The vehicle now takes a curved path to achieve better observability. Comparing the simulation results by the LQD controller in Figure 6 with the results by the CE controller in Figure 4, it is clear that the LQD controller performance is superior in reducing relative position uncertainties.



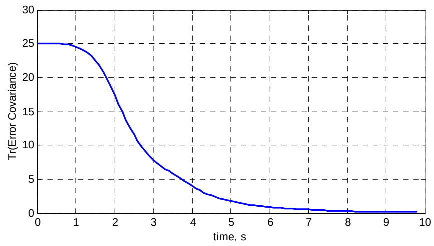

(a) Range : $\rho$




(b) Bearing angle : $\gamma$




(c) Yaw Rate control input : $u_2$




(d) Uncertainty measure : $Tr(P_e)$

**Figure 6. LQD control : Time history of range, bearing angle, control input, and uncertainty measure**

## C. Performance Analysis

Two controllers using search-based optimization techniques have also been compared to the LQD controller. For the comparison, their performances are evaluated through a series of deterministic and stochastic simulations.

### 1. *Deterministic Analysis*

To evaluate controller performance deterministically, the controller is run under zero process and zero measurement noises with perfect initial state estimate. However, the initial error covariance is set to be a non-zero value to make the dual properties apparent. In addition to the LQD controller, two different types of search-based controllers are considered for comparative performance evaluation.

The first controller for the comparison adopts a look-ahead search.

$$u^* = \underset{u_m,\cdots,u_{m+p-1}}{\arg\min} \left\{ E\left[ \sum_{k=m}^{m+p} L_k(x_k, u_k) \,\middle|\, \mathcal{I}_m \right] \right\} \tag{36}$$

where $p$ is the look-ahead parameter. The simplest form of this kind is obtained by setting $p = 1$, which is called one-step-look-ahead search. As $p$ increases, the problem soon becomes intractable. In this study, for each finite look-ahead time interval, constant control inputs are assumed to be maintained. For a specific optimization technique, the gradient descent method with a variable step-size is used.

The second controller for the comparison exploits exhaustive search.

$$u^* = \underset{u_0,\cdots,u_{N-1}}{\arg\min} \left\{ E\left[ L_N(x_N) + \sum_{k=0}^{N-1} L_k(x_k, u_k) \,\middle|\, \mathcal{I}_N \right] \right\} \tag{37}$$

The exhaustive search requires huge computational power, thus cannot be used as a real-time controller. But, a control trajectory which is closer to the optimum is expected to be found through this naive approach. To keep computation time manageable, the search is performed for the first few seconds during which the choice of control input is relatively critical. For the remaining time, proportional navigation (PN) guidance, which is widely used as a terminal controller, is applied. The same simulation conditions in Section B are applied.

**Table 1. Performance index and runtime cost**

| Controller | Cost $J$ | Runtime |
|---|---|---|
| (1) LQD | 3837.9 | 0.83 |
| (2-1) Look-ahead search ($p = 1$) | 5205.2 | 0.64 |
| (2-2) Look-ahead search ($p = 5$) | 4851.0 | 1.48 |
| (2-3) Look-ahead search ($p = 10$) | 4475.8 | 3.14 |
| (2-4) Look-ahead search ($p = 20$) | 3958.5 | 7.81 |
| (2-5) Look-ahead search ($p = 30$) | 3887.7 | 12.47 |
| (3-1) Exhaustive search ($t_{\mathsf{sec}} = 0.5, n_{\mathsf{sec}} = 4$) | 3822.8 | 327.5 |
| (3-2) Exhaustive search ($t_{\mathsf{sec}} = 1.0, n_{\mathsf{sec}} = 4$) | 3828.2 | 332.7 |
| (3-3) Exhaustive search ($t_{\mathsf{sec}} = 0.7, n_{\mathsf{sec}} = 5$) | 3797.1 | 3433.2 |
| (3-4) Exhaustive search ($t_{\mathsf{sec}} = 0.5, n_{\mathsf{sec}} = 6$) | 3816.0 | 33795.3 |
| (3-5) Exhaustive search ($t_{\mathsf{sec}} = 0.7, n_{\mathsf{sec}} = 6$) | 3817.3 | 34243.2 |

In Table 1, $p$ represents the look-ahead step parameter. $t_{\mathsf{sec}}$ and $n_{\mathsf{sec}}$ are the parameters for the exhaustive search, which mean the duration of each time section, the number of time sections, respectively.

According to the results in 1, the performance of the LQD controller is found to be quite satisfactory considering the runtime cost in comparison with the results obtained through other search-based methods.

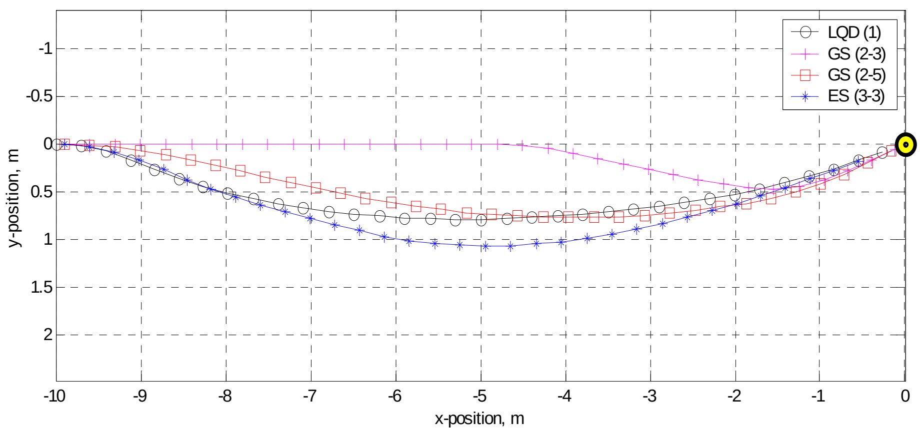



**Figure 7. Trajectories generated by various controller**

As shown in Figure 7, the vehicle moves away from the centerline to reduce the position uncertainties due to the lack of range information. Note that the trajectory by the LQD controller is very similar to that by the time-consuming exhaustive search, especially for the first few seconds on which the search effort is concentrated. This provides some evidence that the LQD solution might be effective even in the early stage.

## *2. Stochastic Analysis*

To evaluate the performance under uncertainties, Monte-Carlo simulations are performed for the cases with the LQD controller and the look-ahead search controller. The exhaustive search is not considered for the stochastic analysis due to its prohibitive computational requirement. Two sets of simulations are performed for different initial position uncertainties, and each simulation case is run for a total of 100 different random seeds. The results are shown in Table 2 and 3.

The statistics are obtained after removing a small number of outliers which are caused by filter divergence. This numerical instability is believed to be the limitation of EKF rather than the LQD algorithm, since those

**Table 2. Cost and range error statistics with $P_0(1,1) = 5.0^2$**

| Controller | $E[J]$ | $E[\tilde{\rho}]$ |
|---|---|---|
| (1) LQD | 3631.9 | 1.244 |
| (2-1) Look-ahead search ($p = 1$) | 4202.0 | 4.264 |
| (2-2) Look-ahead search ($p = 5$) | 4101.3 | 3.625 |
| (2-3) Look-ahead search ($p = 10$) | 3869.3 | 3.339 |
| (2-4) Look-ahead search ($p = 20$) | 3653.2 | 1.439 |
| (2-5) Look-ahead search ($p = 30$) | 3694.1 | 1.381 |

**Table 3. Cost and range error statistics with $P_0(1,1) = 2.0^2$**

| Controller | $E[J]$ | $E[\tilde{\rho}]$ |
|---|---|---|
| (1) LQD | 3175.6 | 0.652 |
| (2-1) Look-ahead search ($p = 1$) | 3203.2 | 1.480 |
| (2-2) Look-ahead search ($p = 5$) | 3190.1 | 1.440 |
| (2-3) Look-ahead search ($p = 10$) | 3189.6 | 1.475 |
| (2-4) Look-ahead search ($p = 20$) | 3181.0 | 0.974 |
| (2-5) Look-ahead search ($p = 30$) | 3172.5 | 0.864 |

outliers exist for both LQD and look-ahead search.

In addition to the computational advantage in runtime cost as shown in Table 1, the stochastic simulation results in Table 2 and 3 prove that LQD works satisfactorily in minimizing the defined performance index as well, compared to the search method with various look-ahead parameters.[f]

There may exist some other search-based techniques such as brute-force or A* search which may outperform LQD in minimizing the performance index. However, the proposed LQD technique is a very effective design tool for dual control considering computational efficiency and real-time capability.

## VI. Summary and Conclusions

In this study, a new stochastic control algorithm has been introduced, which can take the dual features into consideration. The augmented state vector and the quadratic performance index considering system uncertainties have been defined, which yield a familiar optimal control formulation. Finally, a state feedback control law has been presented as an approximate solution of the nonlinear stochastic dual control problem.

The effectiveness of the suggested approach is demostrated through a series of simulations. Satisfactory performance has been observed from the simulation results comparing to other search-based suboptimal approaches. The real-time capability of the new technique has also been verified. The proposed algorithm can be implemented as an online controller on practical dual control applications of moderate dimensions, which has generally not been possible for conventional search-based techniques. Another crucial advantage of the suggested technique lies in its systematic approach which is free from heuristics or ad-hoc methods.

## References


[1]Feldbaum, A. A., "Dual Control Theory I-IV," *Automation Remote Control*, Vol. 21, 22, 1960-61, pp. 874–880, 1033–1039, 1–12, 109–121.

[2]Tse, E., Bar-Shalom, Y., and L. Meier, I., "Wide-Sense Adaptive Dual Control for Nonlinear Stochastic Systems," *IEEE Trans. Automatic Control*, Vol. 18, No. 2, April 1973, pp. 98–108.

[3]Kirk, D. E., *Optimal Control Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1970.


[f]Longer look-ahead time is expected to provide better controller performance at the price of increased computational cost. However, it is not always guaranteed here, because, in this study, constant control inputs are assumed to be maintained over each look-ahead time interval for computational convenience.

[4]Bryson, A. E. and Ho, Y. C., *Applied Optimal Control*, Hemisphere, New York, 1975.

[5]Sternby, J., "A Simple Dual Control Problem with an Analytical Solutions," *IEEE Trans. Automatic Control*, Vol. 21, No. 6, Dec. 1976, pp. 840–844.

[6]Bernhardsson, B., "Dual Control of a First-Order Sytem with Two Possible Gains," *International Journal of Adaptive Control and Signal Processing*, Vol. 3, 1989, pp. 15–22.

[7]Åström and Wittenmark, B., *Adaptive Control, 2nd Ed.*, Addison Wesley, Reading, Massachusetts, 1995.

[8]Casler Jr., R. J., "Dual-Control Guidance Strategy for Homing Interceptors Taking Angle-Only Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 1, 1978, pp. 63–70.

[9]Hull, D. G., Speyer, J. L., and Burris, D. B., "Linear-Quadratic Guidance Law for Dual Control of Homing Missiles," *Journal of Guidance, Control, and Dynamics*, Vol. 3, 1990, pp. 137–144.

[10]Wittenmark, B., "Adaptive Dual Control Methods: An Overview," *5th IFAC Symposium on Adaptive Systems in Control and Signal Processing*, Budapest, Hungary, 1995, pp. 62–72.

[11]Unbehauen, H., "Adaptive Dual Control Systems: A Survey," *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000*, ASSPCC, Lake Louise, Canada, 2000, pp. 171–180.

[12]Morf, M., Levy, B., and Kailath, T., "Square-Root Algorithms for the Continuous-Time Linear Least-Square Estimation Problem," *IEEE Trans. Automatic Control*, Vol. 23, No. 5, Oct. 1978, pp. 907–911.

[13]Maybeck, P. S., *Stochastic Models, Estimation, and Control, Vol. 1*, Academic Press, New York, 1979.