



The Optimal Solution of a Non-Convex State-Dependent LQR Problem and Its Applications

Xudan Xu^{1,2*}, J. Jim Zhu², Ping Zhang¹

¹ School of Automation Science and Electrical Engineering, Beihang University, Beijing, China, ² School of Electrical Engineering and Computer Science, Ohio University, Athens, Ohio, United States of America

Abstract

This paper studies a Non-convex State-dependent Linear Quadratic Regulator (NSLQR) problem, in which the control penalty weighting matrix R in the performance index is state-dependent. A necessary and sufficient condition for the optimal solution is established with a rigorous proof by Euler-Lagrange Equation. It is found that the optimal solution of the NSLQR problem can be obtained by solving a Pseudo-Differential-Riccati-Equation (PDRE) simultaneously with the closed-loop system equation. A Comparison Theorem for the PDRE is given to facilitate solution methods for the PDRE. A linear time-variant system is employed as an example in simulation to verify the proposed optimal solution. As a non-trivial application, a goal pursuit process in psychology is modeled as a NSLQR problem and two typical goal pursuit behaviors found in human and animals are reproduced using different control weighting $R(x)$. It is found that these two behaviors save control energy and cause less stress over Conventional Control Behavior typified by the LQR control with a constant control weighting R , in situations where only the goal discrepancy at the terminal time is of concern, such as in Marathon races and target hitting missions.

Citation: Xu X, Zhu JJ, Zhang P (2014) The Optimal Solution of a Non-Convex State-Dependent LQR Problem and Its Applications. PLoS ONE 9(4): e94925. doi:10.1371/journal.pone.0094925

Editor: Francesco Pappalardo, University of Catania, Italy

Received: October 23, 2013; **Accepted:** March 21, 2014; **Published:** April 18, 2014

Copyright: © 2014 Xu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The author Xudan Xu would like to thank the Chinese Scholar Council for the financial support during her stay at the Ohio University as a Visiting Research Scholar. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: xx505111@ohio.edu

Introduction

1.1 Problem Definition

In this paper, we seek an optimal control law $u = k(x, t)$, for which the performance index

$$\begin{aligned} J(x(t), u(t)) &= \int_{t_0}^{t_f} L(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ &= \frac{1}{2} \int_{t_0}^{t_f} (x^T(t) Q(t) x(t) + u^T(t) R(x(t)) u(t)) dt \\ &\quad + \frac{1}{2} x^T(t_f) S(t_f) x(t_f) \end{aligned} \quad (1)$$

is minimized along the associated closed-loop system trajectory of the Linear Time-variant (LTV) system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \quad x(t_0) = x_0 \quad (2)$$

where $u(t) \in \mathbb{R}^m$ is the control input, $x(t) \in \mathbb{R}^n$ is the system state, t_0 is the starting time, t_f is the terminal time and x_0 is the initial value of $x(t)$ at time t_0 . The coefficients $A(t), Q(t), S(t_f) \in \mathbb{R}^{n \times n}$, $B(t) \in \mathbb{R}^{n \times m}$, $R(x(t)) \in \mathbb{R}^{m \times m}$. To simplify notation, the dependence of variables on t is omitted when no confusion will be introduced in the rest of the paper. It is assumed that A, B, Q are continuous in t , $R(x)$ is differentiable with respect to x , and $\frac{\partial R}{\partial x}$ is bounded.

The coefficients Q and S are positive semi-definite symmetric matrices for all $t \in [t_0, t_f]$, and $R(x)$ is a positive definite symmetric matrix for all $x(t), t \in [t_0, t_f]$. Additional conditions on $R(x)$ will be imposed in order to obtain the sufficiency for optimality.

It is noted that when the state-dependent matrix $R(x)$ in Eq (1) is replaced by a time-dependent matrix $R(t)$, the performance index J is quadratic and convex in both x and u , and Eq (1) and (2) constitute the standard Linear Quadratic Regulator (LQR) problem. The classical LQR theory provides a mature way to find an optimal control law for such a convex quadratic performance index. However, the state-dependent coefficient $R(x)$ in Eq (1) renders the performance index in the problem no longer convex in both x and u , which makes the LQR theory inapplicable here. However, the formalism of the LQR theory is still useful. Therefore, we denote the problem defined above as a Non-convex State-dependent LQR (NSLQR) problem. The associated Riccati Equation of the NSLQR problem is named as Pseudo-Differential-Riccati-Equation (PDRE). In this paper, a necessary and sufficient condition for the optimal solution of the defined NSLQR problem is presented, with an additional condition on $R(x)$, and the optimality of the solution is proven with Euler-Lagrange Equation. The PDRE is also studied to obtain the optimal solution and a theorem is given to estimate the solution of the PDRE.

1.2 Related Work

A similar problem has been studied in the context of State-Dependent Riccati Equation (SDRE) control strategy since mid-90's. The strategy, proposed by Pearson[1] and expanded by Wernli and Cook[2], was independently studied by Mracek and Cloutier[3] in 1998. Friedland[4], Salamci and Gokbilen[5], Cimen et.al[6,7] also contributed to the existence of solutions as well as the properties of optimality and stability. In the SDRE strategy, a nonlinear system is “factored” into the product of the state vector and the state-dependent matrix-valued function in the form

$$\begin{aligned}\dot{x}(t) &= f(x(t)) + g(x(t))u(t) = A^*(x(t))x(t) + B^*(x(t))u(t) \\ x(t_0) &= x_0\end{aligned}\quad (3)$$

which is a linear structure having state-dependent coefficients. Borrowing the LQR theory, the SDRE strategy postulates an approximately optimal feedback control law as

$$u(x) = -R^{*-1}(x)B^{*T}(x)P^*(x)x(t) \quad (4)$$

for the performance index

$$J^* = \frac{1}{2} \int_{t_0}^{\infty} (x^T(t)Q^*(x)x(t) + u^T(t)R^*(x)u(t))dt \quad (5)$$

where $P^*(x)$ is the solution of an algebraic Riccati Equation (RE) as

$$\begin{aligned}A^{*T}(x)P^*(x) + P^*(x)A^*(x) \\ - P^*(x)B^*(x)R^{*-1}(x)B^{*T}(x)P^*(x) + Q^*(x) = 0\end{aligned}\quad (6)$$

This strategy has been applied in a wide variety of nonlinear control applications, such as autopilot design [8,9], integrated guidance and control design[10], etc.

However, only the necessary condition for optimality has been studied in the SDRE control strategy, and it cannot always be established. So the optimality of the control law in the SDRE control strategy cannot be guaranteed. Since a simplified algebraic RE is employed to obtain P^* instead of a differential RE, the application of the SDRE control strategy is limited to Slowly Time-Varying and Weakly State-dependent Systems. Moreover, even though the SDRE strategy has been used in many applications, in most cases, the coefficients R^* and Q^* in the performance index J^* are constant instead of state-dependent, as shown in its formulation[7].

The NSLQR problem defined in this paper focuses on the state-dependent $R(x)$ and the time-dependent $Q(t)$ in the performance index and starts with the LTV systems. The optimality of the solution is validated by a rigorous proof with Euler-Lagrange Equation. The solution can be obtained by solving a PDRE associated with the problem. The work is a special case of the SDRE control strategy, but with rigorous mathematical proof. It could be considered as a theoretical support for the SDRE control strategy.

On another aspect, the solution of the optimal LTV problem is usually obtained through numerical approximation approaches, which can be roughly classified into offline and online methods. The offline method usually pre-computes solutions and stores them for fast online look-up [11,12]. Since the computation grows exponentially with the size of the control problem, offline methods

are normally used in small- and medium-size applications. The most prominent online methods are active set[13] and interior point method[14]. The method of active set performs well in large-size cases even though its convergence rate is unknown. For the interior point method, the reported lower iteration number is larger than the practically observable number. In Ref. [15] and [16], a fast gradient method is introduced to help calculating the lower iteration bound for a quadratic LTV optimal problem with input constraints. Though the work listed above is mainly about the optimal problem with a time-dependent R , the formalism is still applicable when developing the numerical solution for the defined NSLQR problem.

1.3 Application Background

The NSLQR problem discussed in this paper can be applied to model a psychological goal pursuit process, as a non-trivial example.

Psychologists observe that there are two different behaviors when intelligent creatures pursue a goal. One is the Goal-Gradient Behavior (GGB) [17–20], in which the control effort to reach a goal increases monotonically with the proximity to the desired end state, such as the predator stalking behavior and the deadline beating behavior. Fig. 1 (a) and (b) give the normalized goal discrepancy and control effort of the GGB. As it is shown, with a monotonically increasing control energy, as the goal is approached the discrepancy reaches zero faster at the end of the process than it does at the beginning. The other is the Stuck-in-the-Middle Behavior (SMB) [21], in which the control effort to reach a goal is high at the beginning of the goal pursuit and when the desired end state is in sight, but it is maintained at a low level in between, such as what athletes do in Marathon. Part (c) and (d) in Fig. 1 show typically the SMB where the goal discrepancy decreases faster at the two ends than it is in the middle of the goal pursuit process and the control effort is maintained at a low level in the middle.

Both the GGB and the SMB are different from the Conventional Control Behavior (CCB) found in an engineering control system, as shown in part (e) and (f) in Fig. 1. For the CCB, the control effort is proportional to the goal discrepancy, so the effort decreases with proximity to the desired end. The purpose of this paper is to study which one of these three behaviors is the best. Some computational models of the GGB have been proposed based on psychological interpretation [22,23]. In this paper, a single-task goal pursuit process is modeled as a NSLQR problem and the three behaviors are reproduced for comparison, facilitating “a deeper understanding of mathematical characterizations of principles of adaptive intelligence”[24] instead of psychological interpretation.

In the sequel, Section 2 presents the necessary and sufficient condition for the optimality of the solution to the NSLQR problem. Section 3 analyzes the solution of a PDRE involved in the NSLQR problem and presents a Comparison Theorem. Section 4 verifies the feasibility of the NSLQR theory with a LTV system and applies the NSLQR to model a goal pursuit process. The numerical simulation results are presented to demonstrate that the GGB and SMB save control energy and cause less stress over the CCB in some applications. Conclusion and Future Work are presented in Section 5.

Analysis of the Optimality of the Solution

In this section, the main result of this paper is presented: the necessary and sufficient condition of the optimality of the solution to the NSLQR problem defined in Eq (1) and (2). Before that, an Optimality Lemma is introduced first. The Optimality Lemma

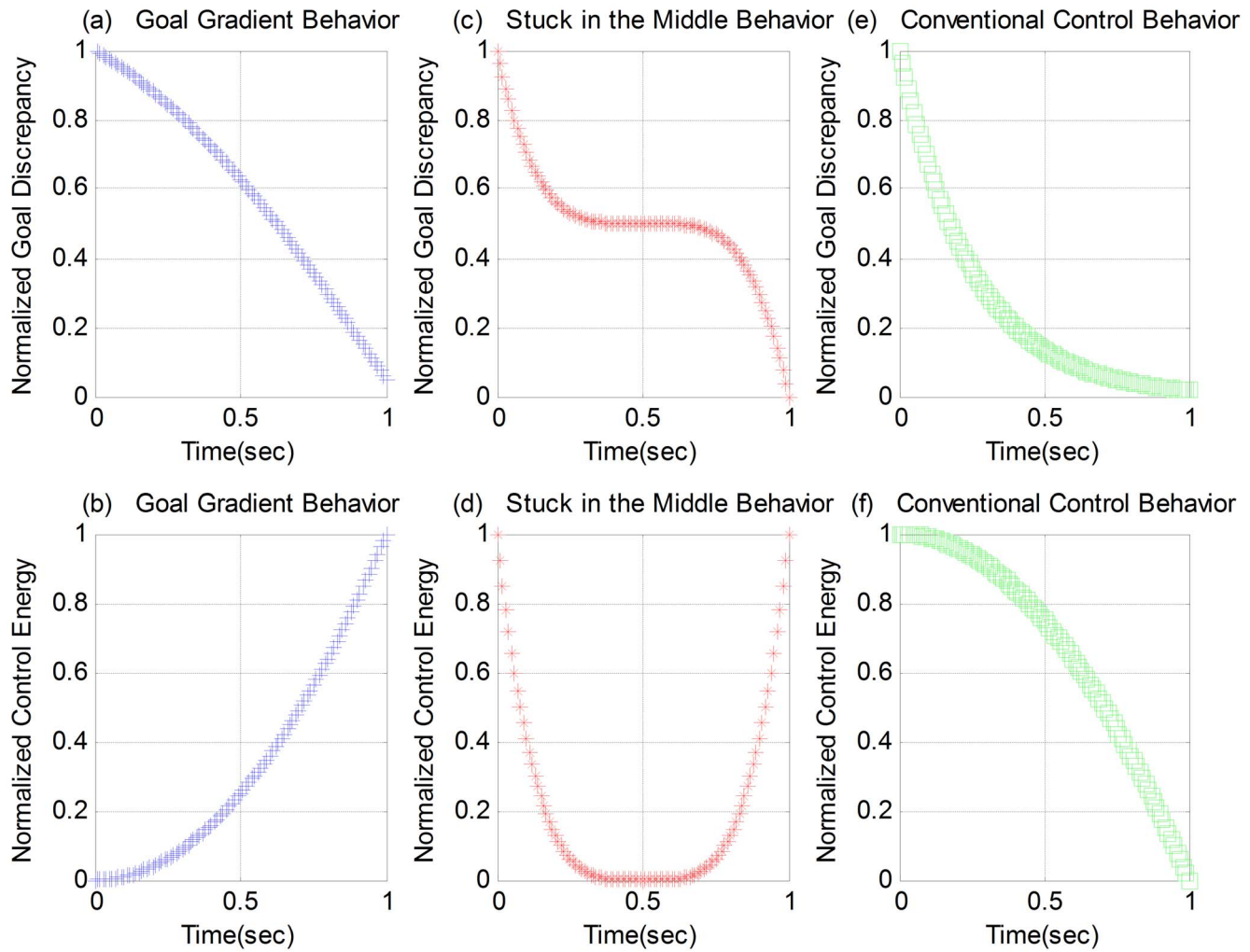


Figure 1. Three Different Behaviors: GGB, SMB and CCB.

doi:10.1371/journal.pone.0094925.g001

discusses a more general optimal problem, compared with the NSLQR problem. To differ from the performance index J in the NSLQR, we denote the performance index in the Lemma as J' with an associated general $L'(x(t), u(t), t)$. The associated augmented performance index is defined as \tilde{J}' in Eq (10), to distinguish it from the \tilde{J} in the NSLQR problem. In this paper, the header “~” indicates the associated augmented performance index, as presented on Page 379 in Ref [25].

Lemma 1 For the problem of finding an optimal control law $u(t)$, for which the performance index

$$J'(x(t), u(t)) = \int_{t_0}^{t_f} L'(x(t), u(t), t) dt + \phi'(x(t_f), t_f) \quad (7)$$

is minimized along the associated closed-loop system trajectory $x(t)$ of the LTV system (2), with a fixed starting time t_0 and terminal time t_f , to simplify notation, define

$$z(t) = [x^T(t), u^T(t)]^T \quad (8)$$

$$\begin{aligned} f'(z(t), \dot{z}(t), t) &= h'(x(t), u(t), \dot{x}(t), \dot{u}(t), t) \\ &= L'(x(t), u(t), t) + \lambda'(t)^T (A(t)x(t) + B(t)u(t) - \dot{x}(t)) \end{aligned} \quad (9)$$

and an augmented performance index as

$$\tilde{J}'(x(t), u(t)) = \int_{t_0}^{t_f} f'(z(t), \dot{z}(t), t) dt + \phi'(x(t_f), t_f) \quad (10)$$

where $\lambda'(t)$ is the Euler-Lagrange multiplier, with the boundary condition as

$$\begin{aligned} \frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta z \Big|_{z=z^o, t=t_f} - \frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta z \Big|_{z=z^o, t=t_0} \\ + \frac{\partial \phi'}{\partial x_{t_f}}(x_{t_f}, t_f) \delta x_{t_f} = 0 \end{aligned} \quad (11)$$

Then the point $z^o = [(x^o)^T, (u^o)^T]^T$ that satisfies Euler-Lagrange Equation

$$\frac{\partial f'}{\partial z}(z, \dot{z}, t) \Big|_{z^o} = \frac{d}{dt} \left(\frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \right) \Big|_{z^o} \quad (12)$$

is the optimal solution if the augmented performance index $\tilde{J}'(x, u)$ is strictly convex in x , uniformly in u ; and strictly convex in u , uniformly in x .

Here, the superscript “o” indicates the optimal solution.

Proof 1 It can be proven that $\tilde{J}'(x, u)$ is equivalent to $J'(x, u)$ in that they have the same minimizing function, if it exists [25]. With Euler-Lagrange Equation (12), the variation of $\tilde{J}'(x, u)$ at z^o can be written as

$$\begin{aligned} \delta \tilde{J}'(x^o, u^o) &= \int_{t_0}^{t_f} \left(\frac{\partial f'}{\partial z} \delta z + \frac{\partial f'}{\partial \dot{z}} \delta \dot{z} \right) \Big|_{z^o} dt + \frac{\partial \phi'}{\partial x_{t_f}}(x_{t_f}, t_f) \delta x_{t_f} \\ &= \int_{t_0}^{t_f} \left(\frac{d}{dt} \frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta z + \frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta \dot{z} \right) \Big|_{z^o} dt \\ &\quad + \frac{\partial \phi'}{\partial x_{t_f}}(x_{t_f}, t_f) \delta x_{t_f} \\ &= \int_{t_0}^{t_f} \frac{d}{dt} \left(\frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta z \right) \Big|_{z^o} dt + \frac{\partial \phi'}{\partial x_{t_f}}(x_{t_f}, t_f) \delta x_{t_f} \\ &= \frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta z \Big|_{z^o, t=t_f} - \frac{\partial f'}{\partial \dot{z}}(z, \dot{z}, t) \delta z \Big|_{z^o, t=t_0} \\ &\quad + \frac{\partial \phi'}{\partial x_{t_f}}(x_{t_f}, t_f) \delta x_{t_f} \end{aligned} \quad (13)$$

To make the point z^o be the optimal solution, the variation $\delta \tilde{J}'(x, u)$ is supposed to be 0 at z^o , which leads to Eq (11). However, the point z^o satisfying Eq (11) and (12) can either be an extreme point or a saddle point of the $\tilde{J}'(x, u)$. Now we prove that, under the convexity constraint stated in Lemma 1, the point z^o is a minimum point by contradiction.

With $\tilde{J}'(x, u)$ being strictly convex in x , uniformly in u , it follows that $x^o(t)$ satisfying Eq (12) is a minimum point of $\tilde{J}'(x, u)$ with respect to x , uniformly in u . Then we have

$$\tilde{J}'(x, u) > \tilde{J}'(x^o, u) \quad (14)$$

for every $u(t), t \in [t_0, t_f]$ and every $x(t) \neq x^o(t)$. Similarly, we have

$$\tilde{J}'(x, u) > \tilde{J}'(x, u^o) \quad (15)$$

for every $x(t), t \in [t_0, t_f]$ and every $u(t) \neq u^o(t)$.

Assume that the point z^o is a maximum point of $\tilde{J}'(x, u)$. Then since $\tilde{J}'(x, u)$ is continuous, there exists a $x(t) \neq x^o(t)$, such that

$$\tilde{J}'(x, u^o) \leq \tilde{J}'(x^o, u^o) \quad (16)$$

which contradicts Eq (14). Thus z^o cannot be a maximum point of $\tilde{J}'(x, u)$. Now assume that z^o is a saddle point of $\tilde{J}'(x, u)$. Then by continuity of $\tilde{J}'(x, u)$, there must be a $z^1 = [(x^1)^T, (u^1)^T]^T \neq z^o$ and a $z^2 = [(x^2)^T, (u^2)^T]^T \neq z^o$ such that

$$\tilde{J}'(x^1, u^1) \leq \tilde{J}'(x^o, u^o) \leq \tilde{J}'(x^2, u^2) \quad (17)$$

Then with Eq (15), we have

$$\tilde{J}'(x^1, u^o) \leq \tilde{J}'(x^1, u^1) \leq \tilde{J}'(x^o, u^o) \quad (18)$$

The equality holds if and only if $x^1 = x^o$ and $u^1 = u^o$, which contradicts the definition of z^1 . The inequality contradicts Eq (14). Thus the point z^o cannot be a saddle point of $\tilde{J}'(x, u)$ either. Then the point z^o must be a minimum point of $\tilde{J}'(x, u)$. So the solution (x^o, u^o) minimizes the performance index J' in (7).

From the proof above, it can be said that the classical LQR is a special case of Lemma 1. In the classical LQR theory, the sufficiency of the optimality of the solution obtained from Euler-Lagrange Equation is guaranteed by the convexity of the augmented performance index \tilde{J} in its arguments (x, u) [25]. However, in the NSLQR, \tilde{J} is no longer convex in (x, u) because of the state-dependent $R(x)$. The theorem below shows that the solution of Euler-Lagrange Equation is still optimal for the NSLQR problem with a constraint on $R(x)$.

Theorem 1 Under the convexity constraint that the function $l(x, u) = u^T R(x) u$ is a strictly convex function in x , uniformly in u , the state feedback control law

$$u^o(t) = k(x^o, t) = -R^{-1}(x^o) B^T(t) P(t) x^o(t) \quad (19)$$

for the NSLQR problem defined in Eq (1) and (2), and the associated closed-loop system trajectory $x^o(t)$ as

$$\dot{x}^o(t) = [A(t) - B(t) R^{-1}(x^o) B^T(t) P(t)] x^o(t) \quad x^o(t_0) = x_0 \quad (20)$$

minimizes the performance index (1) if and only if the $n \times n$ matrix $P(t)$ satisfies the PDRE as

$$\begin{aligned} -\dot{P}(t) x^o(t) &= A^T(t) P(t) x^o(t) + P(t) A(t) x^o(t) \\ &\quad - P(t) B(t) R^{-1}(x^o) B^T(t) P(t) x^o(t) \\ &\quad + Q(t) x^o(t) + M(x^o, u^o) \end{aligned} \quad (21)$$

with

$$P(t_f) = S(t_f) \quad (22)$$

where the column vector

$$M(x^o, u^o) = M(x, u)|_{(x^o, u^o)} = \frac{1}{2} \frac{\partial (u^T R(x) u)}{\partial x} \Big|_{(x^o, u^o)} \quad (23)$$

Here, the explicit $x(t)$ on both sides of the equation can be eliminated for some sharpened $R(x)$. One example is discussed in Section 3.

This theorem provides an optimal solution, which is similar to that of the classical LQR theory, for the NSLQR problem. However, the PDRE (21) is with an additional term “ $M(x^o, u^o)$ ”, compared with the standard Riccati Equation in the LQR theory. This term comes from the derivative of the state-dependent $R(x)$ with respect to x in the Euler-Lagrange Equation, as detailed in Proof 2. Theorem 1 can be proven with Lemma 1 as follows.

Proof 2 To simplify notation, define

$$\begin{aligned} f(z(t), \dot{z}(t), t) &= h(x(t), u(t), \dot{x}(t), \dot{u}(t), t) \\ &= L(x(t), u(t), t) + \lambda(t)^T (A(t)x(t) \\ &\quad + B(t)u(t) - \dot{x}(t)) \end{aligned} \quad (24)$$

where

$$z(t) = [x^T(t), u^T(t)]^T \quad (25)$$

and $\lambda(t)$ is the Euler-Lagrange multiplier. The admissible directions of $z(t)$ are denoted as

$$v(t) = [\xi^T(t), \mu^T(t)]^T \quad (26)$$

with $\xi(t_0) = 0$, $\xi(t_f)$, $\mu(t_0)$, and $\mu(t_f)$ being free, since the initial value of $x(t)$ is fixed. The augmented performance index is

$$\tilde{J}(x(t), u(t)) = \int_{t_0}^{t_f} f(z(t), \dot{z}(t), t) dt + \phi(x(t_f), t_f) \quad (27)$$

Similarly, $\tilde{J}(x, u)$ is equivalent to $J(x, u)$ in that they have the same minimizing function, if it exists [25]. Now we need to prove that $x^o(t)$, $u^o(t)$ defined in Eq (19) and (20) minimizes the augmented performance index (27) if and only if $P(t)$ satisfies the PDRE (21) and $P(t_f) = S(t_f)$.

We start with the necessary condition. From Euler-Lagrange Equation, it is known that for a point $z^o = [(x^o)^T, (u^o)^T]^T$ to be an optimal solution, the necessary condition is

$$\frac{\partial f}{\partial z}(z, \dot{z}, t)|_{z=z^o} = \frac{d}{dt} \left(\frac{\partial f}{\partial \dot{z}}(z, \dot{z}, t) \right) |_{z=z^o} \quad (28)$$

with a boundary condition, as discussed later. For a state-dependent $R(x)$

$$\begin{aligned} \frac{\partial f}{\partial z}(z, \dot{z}, t)|_{z=z^o} &= \left[\frac{\partial h}{\partial x}(x, u, \dot{x}, \dot{u}, t), \frac{\partial h}{\partial u}(x, u, \dot{x}, \dot{u}, t) \right] |_{(x^o, u^o)} \\ &= [x^T(t)Q(t) + \lambda^T(t)A(t) + \frac{1}{2} \left(\frac{\partial(u^T R(x)u)}{\partial x} \right)^T, \\ &\quad u^T(t)R(x) + \lambda^T(t)B(t)] |_{(x^o, u^o)} \end{aligned} \quad (29)$$

$$\begin{aligned} \frac{\partial f}{\partial \dot{z}}(z, \dot{z}, t)|_{z=z^o} &= \left[\frac{\partial h}{\partial \dot{x}}(x, u, \dot{x}, \dot{u}, t), \frac{\partial h}{\partial \dot{u}}(x, u, \dot{x}, \dot{u}, t) \right] |_{(x^o, u^o)} \\ &= [-\lambda^T(t), 0] |_{(x^o, u^o)} \end{aligned} \quad (30)$$

To simplify notation, the column vector $\frac{1}{2} \frac{\partial(u^T R(x)u)}{\partial x}$ is denoted as $M(x, u)$. Substituting the two equations above into Eq (28), we have

$$(x^o)^T(t)Q(t) + \lambda^T(t)A(t) + M^T(x^o, u^o) = -\dot{\lambda}^T(t) \quad (31)$$

$$(u^o)^T(t)R(x^o) + \lambda^T(t)B(t) = 0 \quad (32)$$

which leads to a control law as

$$u^o(t) = -R^{-1}(x^o)B^T(t)\lambda(t) \quad (33)$$

with $\lambda(t)$ satisfying

$$-\dot{\lambda}(t) = A(t)^T \lambda(t) + Q(t)x^o(t) + M(x^o, u^o) \quad (34)$$

The variation of $\tilde{J}(x, u)$ at the point $z^o = [(x^o)^T, (u^o)^T]^T$ can be written as

$$\begin{aligned} \delta \tilde{J}((x^o, u^o); (\xi, \mu)) &= \int_{t_0}^{t_f} \left(\frac{\partial f}{\partial z} \delta z + \frac{\partial f}{\partial \dot{z}} \delta \dot{z} \right) |_{z^o} dt \\ &\quad + \frac{\partial \phi(x(t_f), t_f)}{\partial z} |_{z^o} \delta z(t_f) \\ &= \int_{t_0}^{t_f} ([x^T(t)Q(t) + \lambda^T(t)A(t) + M^T(x, t), \\ &\quad u^T(t)R(x) + \lambda^T(t)B(t)] \begin{bmatrix} \xi(t) \\ \mu(t) \end{bmatrix} \\ &\quad + [-\lambda^T(t), 0] \begin{bmatrix} \dot{\xi}(t) \\ \dot{\mu}(t) \end{bmatrix}) |_{(x^o, u^o)} dt \\ &\quad + x^T(t_f)S(t_f)\xi(t_f) \\ &= \int_{t_0}^{t_f} ([-\dot{\lambda}^T(t), 0] \begin{bmatrix} \xi(t) \\ \mu(t) \end{bmatrix} + [-\lambda^T(t), 0] \\ &\quad \begin{bmatrix} \dot{\xi}(t) \\ \dot{\mu}(t) \end{bmatrix}) dt + x^T(t_f)S(t_f)\xi(t_f) \\ &= -\lambda^T(t_f)\xi(t_f) + \lambda(t_0)^T \xi(t_0) \\ &\quad + x^T(t_f)S(t_f)\xi(t_f) \\ &= (x^T(t_f)S(t_f) - \lambda^T(t_f))\xi(t_f) \end{aligned} \quad (35)$$

To minimize $\tilde{J}(x(t), u(t))$, we need to achieve $\delta \tilde{J}((x, u); (\xi, \mu)) = 0$ for all admissible directions. Since $\xi(t_0) = 0$ and $\xi(t_f)$ is free, the terminal value of λ needs to satisfy

$$\lambda(t_f) = S(t_f)x(t_f) \quad (36)$$

We choose $\lambda(t)$ such that it is linearly related to $x(t)$ through $\lambda(t) = P(t)x(t)$. Then the boundary condition Eq (36) becomes

$$P(t_f) = S(t_f) \quad (37)$$

Substituting the assumption $\lambda(t) = P(t)x(t)$ into Eq (33) and (34), we obtain a control law as

$$u^o(t) = k(x^o, t) = K(t)x^o(t) = -R^{-1}(x^o)B^T(t)P(t)x^o(t) \quad (38)$$

where $K(t) = -R^{-1}(x^o)B^T(t)P(t)$, with the associated closed-loop system as

$$\dot{x}^o(t) = [A(t) - B(t)R^{-1}(x^o)B^T(t)P(t)]x^o(t) \quad x^o(t_0) = x_0 \quad (39)$$

where $P(t)$ satisfies

$$\begin{aligned} -\dot{P}(t)x^o(t) &= A^T(t)P(t)x^o(t) + P(t)A(t)x^o(t) \\ &\quad - P(t)B(t)R^{-1}(x^o)B^T(t)P(t)x^o(t) \\ &\quad + Q(t)x^o(t) + M(x^o, u^o) \end{aligned} \quad (40)$$

and $P(t_f) = S(t_f)$.

Now we prove that the solution (x^o, u^o) from the necessary condition of optimality also satisfies the sufficiency condition using Lemma 1.

First, it is easy to verify that the solution (x^o, u^o) in Eq (38) and (39) with $P(t_f) = S(t_f)$ satisfies Eq (11) and (12) in Lemma 1. Now we prove that $\tilde{J}(x, u)$ in Theorem 1 is strictly convex in u , uniformly in x . Considering an arbitrary fixed x , we set $\xi(t) = 0$, then the variation of $\tilde{J}(x, u)$ in u is

$$\begin{aligned} \delta((x, u); (0, \mu)) &= \int_{t_0}^{t_f} \left(\frac{\partial}{\partial u} \mu + \frac{\partial}{\partial u} \dot{\mu} \right) dt \\ &= \int_{t_0}^{t_f} (u^T R(x) \mu + \lambda^T B \mu) dt \end{aligned} \quad (41)$$

In terms of Eq (41), we have

$$\begin{aligned} \tilde{J}((x, u); (0, \mu)) - \tilde{J}(x, u) &= \left(\int_{t_0}^{t_f} f(z + [0, \mu^T]^T, \dot{z} + [0, \dot{\mu}^T]^T, t) \right. \\ &\quad \left. dt + \phi(x(t_f), t_f) \right) - \left(\int_{t_0}^{t_f} f(z, \dot{z}, t) dt + \phi(x(t_f), t_f) \right) \\ &= \int_{t_0}^{t_f} (f(z + [0, \mu^T]^T, \dot{z} + [0, \dot{\mu}^T]^T, t) \\ &\quad - f(z, \dot{z}, t)) dt \\ &= \int_{t_0}^{t_f} (L(x, u + \mu, t) - L(x, u, t) \\ &\quad + \lambda^T B \mu) dt \\ &= \int_{t_0}^{t_f} \left(\frac{1}{2} ((u + \mu)^T R(x) (u + \mu) \right. \\ &\quad \left. - u^T R(x) u) + \lambda^T B \mu \right) dt \\ &= \int_{t_0}^{t_f} (u^T R(x) \mu + \lambda^T B \mu) dt \\ &\quad + \frac{1}{2} \int_{t_0}^{t_f} \mu^T R(x) \mu dt \\ &= \delta \tilde{J}((x, u); (0, \mu)) \\ &\quad + \frac{1}{2} \int_{t_0}^{t_f} \mu^T R(x) \mu dt \\ &\geq \delta \tilde{J}((x, u); (0, \mu)) \end{aligned} \quad (42)$$

Since $R(x) > 0$, the equality holds when and only when $\mu(t) = 0$. So $\tilde{J}(x, u)$ is strictly convex in u , uniformly in x .

Now we prove that $\tilde{J}(x, u)$ is strictly convex in x , uniformly in u . Setting $\mu(t) = 0$, the variation of $\tilde{J}(x, u)$ in x is written as

$$\begin{aligned} \delta \tilde{J}((x, u); (\xi, 0)) &= \int_{t_0}^{t_f} \left(\frac{\partial \tilde{J}}{\partial x} \xi + \frac{\partial \tilde{J}}{\partial \dot{x}} \dot{\xi} \right) dt \\ &= \int_{t_0}^{t_f} ((x^T(t) Q(t) + \lambda^T(t) A(t) \\ &\quad + M^T(x, u)) \xi(t) - \lambda^T(t) \dot{\xi}(t)) dt \\ &\quad + x(t_f)^T S(t_f) \xi(t_f) \end{aligned} \quad (43)$$

where $M(x, u) = \frac{1}{2} \frac{\partial (u^T R(x) u)}{\partial x}$. In term of Eq (43), we obtain

$$\begin{aligned} \tilde{J}((x, u); (\xi, 0)) - \tilde{J}(x, u) &= \left(\int_{t_0}^{t_f} f(z + [\xi^T, 0]^T, \dot{z} + [\dot{\xi}^T, 0]^T, t) dt \right. \\ &\quad \left. + \phi(x(t_f) + \xi(t_f), t_f) \right) \\ &\quad - \left(\int_{t_0}^{t_f} f(z, \dot{z}, t) dt + \phi(x(t_f), t_f) \right) \\ &= \int_{t_0}^{t_f} (f(z + [\xi^T, 0]^T, \dot{z} + [\dot{\xi}^T, 0]^T, t) \\ &\quad - f(z, \dot{z}, t)) dt + x(t_f)^T S(t_f) \xi(t_f) \\ &\quad + \frac{1}{2} \xi(t_f)^T S(t_f) \xi(t_f) \\ &= \int_{t_0}^{t_f} (L(x + \xi, u, t) - L(x, u, t) \\ &\quad + \lambda^T (A \xi - \dot{\xi})) dt + x(t_f)^T S(t_f) \xi(t_f) \\ &\quad + \frac{1}{2} \xi(t_f)^T S(t_f) \xi(t_f) \\ &= \int_{t_0}^{t_f} ((x^T(t) Q(t) + \lambda^T(t) A(t) \\ &\quad + M^T(x, u)) \xi(t) - \lambda^T(t) \dot{\xi}(t)) dt \\ &\quad + x(t_f)^T S(t_f) \xi(t_f) \\ &\quad + \frac{1}{2} \int_{t_0}^{t_f} (\xi^T Q \xi + u^T (R(x + \xi) \\ &\quad - R(x)) u - 2 M^T \xi) dt \\ &\quad + \frac{1}{2} \xi(t_f)^T S(t_f) \xi(t_f) \end{aligned} \quad (44)$$

Since the function $l(x, u) = u^T R(x) u$ is a strictly convex function in x , uniformly in u as it is stated in Theorem 1, we have

$$u^T R(x + \xi) u - u^T R(x) u \geq 2 M^T \xi \quad (45)$$

Then

$$\begin{aligned} \tilde{J}((x, u); (\xi, 0)) - \tilde{J}(x, u) &\geq \delta \tilde{J}((x, u); (\xi, 0)) \\ &\quad + \frac{1}{2} \int_{t_0}^{t_f} \xi^T Q \xi dt + \frac{1}{2} \xi(t_f)^T S(t_f) \xi(t_f) \geq \delta \tilde{J}((x, u); (\xi, 0)) \end{aligned} \quad (46)$$

the equality holds when and only when $\xi(t)=0$. So $\tilde{J}(x,u)$ is strictly convex in x , uniformly in u .

From the analysis above, it holds that the $u^o(t)$ and $x^o(t)$ defined in (19) and (20) with the boundary condition (22) satisfy (11) and (12) in Lemma 1. In addition, the augmented performance index $\tilde{J}(x,u)$ is strictly convex in x , u separately, uniformly in the other one. We conclude that x^o and u^o minimize the performance index J in Eq (1) for the NSLQR problem. So $u^o(t)$ defined in (19) is the optimal control law for the NSLQR problem.

From the proof of Theorem 1 and Lemma 1, we obtain a corollary for the defined NSLQR problem with a general $R(x)$ as following.

Corollary 1 For the optimal problem defined in (1) and (2) with $R(x)>0$ and free boundary values of $x(t_0)$ and $x(t_f)$, the $u^o(t)$ and $x^o(t)$ defined in (19) and (20) with $P(t)$ satisfying (21) and (22) is either a minimum or a saddle for the performance index J in (1).

It follows from Eq (42) that \tilde{J} is strictly convex in u , uniformly in x if $R(x)>0$. It then follows from the proof of Lemma 1 that $\tilde{J}(x^o, u^o)$ cannot be a maximum. Detailed proof is omitted.

Remark: The corollary signifies that the optimal control law (19) gives the minimum cost for a particular $x(t_0)$ among all possible control laws for a general $R(x)>0$, since u^o is the minimum point of $J(x,u)$ with respect to u . However, the cost may be lowered if a different $x(t)$ trajectory with a different $x(t_0)$ is chosen since (x^o, u^o) can be a saddle of J . So for a given $x(t_0)$, the control law (19) gives the optimal solution.

From the analysis, it can be said that, for the NSLQR problem with a general $R(x)>0$, the optimal solution (x^o, u^o) needs to be evaluated for every specific $x(t_0)$. Whereas in the classical LQR problem or the SDRE problem, the optimal solution can be explicitly written as a function of t or x , uniformly for any $x(t_0)$. For the NSLQR to have such uniform solutions, $R(x)$ has to satisfy the additional constraint that the function $l(x,u)=u^T R(x)u$ is strictly convex in x , uniformly in u .

Analysis of the Solution for PDRE

3.1 A Sharpened $R(x)$

The PDRE (21) of the NSLQR problem is different from the RE in the SDRE control strategy literature in

1. There is an additional term $M(x,u)$ in the PDRE (21), which is derived from the derivative of the state-dependent $R(x)$ with respect to x ;
2. It is a differential RE instead of an algebraic RE;
3. The system state $x(t)$ appears on both sides of the equation.

The way to solve the SDRE is not applicable for the PDRE. To obtain the optimal solution of the NSLQR, it is necessary to investigate the solution of the PDRE (21). In this section, a sharpened $R(x)$ is studied as an example.

As stated in Theorem 1, the function $l(x,u)$ needs to be strictly convex in x , a quadratic function of x seems to be a reasonable choice for the matrix R . So consider

$$R(x) = x^T R_0 x I_{m \times m} + R_1 \quad (47)$$

where R_0 is a $n \times n$ symmetric, positive semi-definite constant matrix, R_1 is a $m \times m$ symmetric, positive definite constant matrix and $I_{m \times m}$ is the $m \times m$ identity matrix. The term R_1 guarantees that $R(x)$ is invertible. Then

$$\begin{aligned} M(x,u) &= \frac{1}{2} \frac{\partial(u^T R(x)u)}{\partial x} \\ &= \frac{1}{2} \frac{\partial(u^T (x^T R_0 x I_{m \times m} + R_1)u)}{\partial x} \\ &= u^T u R_0 x \end{aligned} \quad (48)$$

and the PDRE (21) becomes

$$\begin{aligned} -\dot{P}(t)x^o(t) &= A^T(t)P(t)x^o(t) + P(t)A(t)x^o(t) \\ &\quad - P(t)B(t)R^{-1}(x^o)B^T(t)P(t)x^o(t) + Q(t)x^o(t) + u^{oT}u^o R_0 x^o(t) \\ \Rightarrow -\dot{P}(t) &= A^T(t)P(t) + P(t)A(t) \\ &\quad - P(t)B(t)R^{-1}(x^o)B^T(t)P(t) + Q(t) + u^{oT}u^o R_0 \\ &= A^T(t)P(t) + P(t)A(t) - P(t)B(t)R^{-1}(x^o)B^T(t)P(t) + Q(t) \\ &\quad + x^{oT}(t)P^T(t)B(t)R^{-1}(x^o)R^{-1}(x^o)B^T(t)P(t)x^o(t)R_0 \end{aligned} \quad (49)$$

which can be denoted as

$$\begin{aligned} -\dot{P}(t) &= A^T(t)P(t) + P(t)A(t) \\ &\quad - P(t)B(t)R^{-1}(x^o)B^T(t)P(t) + Q'(x^o, t) \end{aligned} \quad (50)$$

where $Q'(x^o, t) = Q(t) + x^{oT}(t)P^T(t)B(t)R^{-1}(x^o)R^{-1}(x^o)B^T(t)P(t)x^o(t)R_0 \geq Q(t) \geq 0$, and $x^o(t)$ satisfies the closed-loop system (20).

Remark: The PDRE (50) is coupled with closed-loop system (20). In classical LQR theory, system state $x(t)$ and the solution of Riccati Equation $P(t)$ can be obtained through a $2n$ -dimensional Hamiltonian matrix [25] or by decoupling the system plant and the Riccati Equation. However, in the NSLQR problem, the $2n$ -dimensional Hamiltonian matrix is not linear any more, and the decoupling is not applicable. The PDRE (50) has to be solved together with the closed-loop system (20).

To generalize the results, the PDRE (50) can be rewritten into a general form as:

$$\dot{P}(t) + A^T(t)P(t) + P(t)A(t) + G(x,t) - P(t)V(x,t)P(t) = 0 \quad (51)$$

with a given terminal value $P(t_f)$, and $x(t)$ is a continuous single-valued function of t . The matrices $G(x,t)$, $V(x,t)$ are positive semi-definite, symmetric, and continuous in both arguments.

For convenience of reference, the PDRE (51) is denoted as $\Re(P)=0$ in the sequel. The dependence of the variables on t is omitted, for example, $x(t_m)$ is denoted as x_m . The time argument of matrices and vectors are omitted when no misunderstanding is introduced. For instance, G_m denotes the value of matrix $G(x_m, t_m)$.

3.2 A Comparison Theorem for the PDRE

The propositions and theorem introduced below are derived from Proposition 7 and 8 in [26]. In [26], similar results are developed for time-dependent Riccati Equations with initial values. Now, a Comparison Theorem for the PDRE (51) with a terminal value is given. Before presenting the theorem, four propositions need to be established first.

Proposition 1 Let $P(t)$ be a symmetric solution of the PDRE (51) on $D = T \times X$, where $T = [t_0, t_f]$, $X = \{x(t) | \forall t \in T, x(t) \in \mathbb{R}^n\}$.

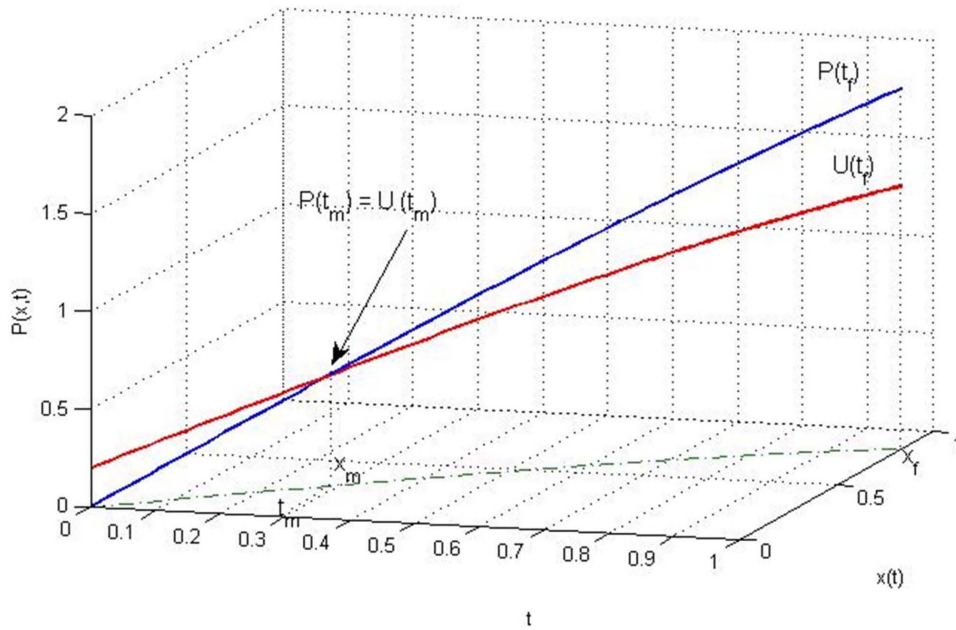


Figure 2. Trajectories of $P(t)$ and $U(t)$.
doi:10.1371/journal.pone.0094925.g002

- (1) If $U(t)$ is a symmetric solution of the inequality $\Re(U) > 0$ on D such that $U(t_f) < P(t_f)$, then $U(t) < P(t)$ on D .
- (2) If $W(t)$ is a symmetric solution of the inequality $\Re(W) < 0$ on D such that $W(t_f) > P(t_f)$, then $W(t) > P(t)$ on D .

Proof 3 Since the matrices and vectors in the PDRE (51) are continuous in the arguments, $P(t)$, $U(t)$ and $W(t)$ are continuous in t . We start with part (1) of the proposition.

Suppose that $U(t) < P(t)$ does not hold on D , there must be a time $t_m \in [t_0, t_f]$ by the Mean Value Theorem, such that

$$U(t_m) = P(t_m) \quad (52)$$

and

$$U(t) < P(t) \quad (53)$$

for any $t \in (t_m, t_f]$, as shown in Fig. 2. Let

$$\phi(t) = \rho^T (P(t) - U(t)) \rho \quad t \in [t_m, t_f] \quad (54)$$

where ρ is a non-zero constant vector. Then

$$\phi(t) \geq 0 \quad t \in [t_m, t_f] \quad (55)$$

and the equality holds if and only if $t = t_m$.

$$\begin{aligned} \dot{\phi}(t_m) &= \rho^T (\dot{P}(t_m) - \dot{U}(t_m)) \rho \\ &= \rho^T \dot{P}(t_m) \rho - \rho^T \dot{U}(t_m) \rho \\ &= \rho^T \Re(P_m) \rho - \rho^T \Re(U_m) \rho \\ &= 0 - \rho^T \Re(U_m) \rho < 0 \end{aligned} \quad (56)$$

Therefore, $\phi(t) < 0$ in a right neighborhood of t_m , which contradicts Eq (55). So $U(t) < P(t)$ holds on D .

Part (2) can be proven similarly.

Proposition 2 Let $P(t)$ be a symmetric solution of the PDRE (51) on $D = T \times X$, where $T = [t_0, t_f]$, $X = \{x(t) | \forall t \in T, x(t) \in \mathbb{R}^n\}$.

- (1) If $U(t)$ is a symmetric solution of the inequality $\Re(U) > 0$ on D such that $U(t_f) \leq P(t_f)$, then $U(t) \leq P(t)$ on D .
- (2) If $W(t)$ is a symmetric solution of the inequality $\Re(W) < 0$ on D such that $W(t_f) \geq P(t_f)$, then $W(t) \geq P(t)$ on D .

Proof 4 The inequality part of the proposition has been proven in Proposition 1. The following part is the proof for the case in which $U(t_f) = P(t_f)$. We start with part (1).

As it is discussed in Proof 3, $P(t)$, $U(t)$ and $W(t)$ are continuous in t . Assume that $U(t) \leq P(t)$ does not hold on D when $U(t_f) = P(t_f)$, there must be a period $[t_m, t_n] \in [t_0, t_f]$, such that

$$U(t) > P(t) \quad (57)$$

for any $t \in [t_m, t_n]$, and

$$U(t_n) = P(t_n) \quad (58)$$

Let

$$\phi(t) = \rho^T (P(t) - U(t))\rho \quad t \in [t_m, t_n] \quad (59)$$

where ρ is a non-zero constant vector. Then

$$\phi(t) \leq 0 \quad t \in [t_m, t_n] \quad (60)$$

and the equality holds if and only if $t = t_n$.

$$\begin{aligned} \dot{\phi}(t_n) &= \rho^T (\dot{P}(t_n) - \dot{U}(t_n))\rho \\ &= \rho^T \dot{P}(t_n)\rho - \rho^T \dot{U}(t_n)\rho \\ &= \rho^T \Re(P_n)\rho - \rho^T \Re(U_n)\rho \\ &= 0 - \rho^T \Re(U_n)\rho < 0 \end{aligned} \quad (61)$$

Therefore, $\phi(t) > 0$ in a left neighborhood of t_n , which contradicts Eq (60). So $U(t) \leq P(t)$ holds on D .

Part (2) can be proven similarly.

Proposition 3 Let $P(t)$ be a symmetric solution of the PDRE (51) on $D = T \times X$, where $T = [t_0, t_f]$, $X = \{x(t) | \forall t \in T, x(t) \in \mathbb{R}^n\}$.

- (1) If $U(t)$ is a symmetric solution of the inequality $\Re(U) \geq 0$ on D such that $U(t_f) < P(t_f)$, then $U(t) < P(t)$ on D .
- (2) If $W(t)$ is a symmetric solution of the inequality $\Re(W) \leq 0$ on D such that $W(t_f) > P(t_f)$, then $W(t) > P(t)$ on D .

Proposition 4 Let $P(t)$ be a symmetric solution of the PDRE (51) on $D = T \times X$, where $T = [t_0, t_f]$, $X = \{x(t) | \forall t \in T, x(t) \in \mathbb{R}^n\}$.

- (1) If $U(t)$ is a symmetric solution of the inequality $\Re(U) \geq 0$ on D such that $U(t_f) \leq P(t_f)$, then $U(t) \leq P(t)$ on D .
- (2) If $W(t)$ is a symmetric solution of the inequality $\Re(W) \leq 0$ on D such that $W(t_f) \geq P(t_f)$, then $W(t) \geq P(t)$ on D .

The proofs of these two propositions are similar to those of Propositions 1 and 2.

These four propositions give a boundary estimation for the solution of the PDRE. Based on the four propositions discussed above, a Comparison Theorem is introduced for the PDRE (51).

Theorem 2 For $i \in \{1, 2\}$, let $P_i(t)$ be the solution of the PDRE

$$-\dot{P}_i(t) = A_i^T(t)P_i(t) + P_i(t)A_i(t) + G_i(x, t) - P_i(t)V_i(x, t)P_i(t) \quad (62)$$

on $D = T \times X$, where $T = [t_0, t_f]$, $X = \{x(t) | \forall t \in T, x(t) \in \mathbb{R}^n\}$. If $P_1(t_f) \leq P_2(t_f)$, and

$$\begin{bmatrix} G_2(x, t) & A_2^T(t) \\ A_2(t) & -V_2(x, t) \end{bmatrix} \geq \begin{bmatrix} G_1(x, t) & A_1^T(t) \\ A_1(t) & -V_1(x, t) \end{bmatrix} \quad (t, x) \in D \quad (63)$$

Then $P_1(t) \leq P_2(t)$ on D .

Proof 5 Let $P_0(t) = P_2(t) - P_1(t)$, then $P_0(t_f) \geq 0$.

$$\begin{aligned} \dot{P}_0(t) &= \dot{P}_2(t) - \dot{P}_1(t) \\ &= -A_2^T(t)P_2(t) - P_2(t)A_2(t) - G_2(x, t) \\ &\quad + P_2(t)V_2(x, t)P_2(t) - \{-A_1^T(t)P_1(t) - P_1(t)A_1(t) \\ &\quad - G_1(x, t) + P_1(t)V_1(x, t)P_1(t)\} \\ &= (P_2 - P_1)V_2(P_2 - P_1) + P_1V_2(P_2 - P_1) \\ &\quad + (P_2 - P_1)V_2P_1 + P_1(V_2 - V_1)P_1 \\ &\quad - A_2^T(P_2 - P_1) - (P_2 - P_1)A_2 - (A_2 - A_1)^T P_1 \\ &\quad - P_1(A_2 - A_1) - (G_2 - G_1) \\ &= -\tilde{A}^T P_0 - P_0 \tilde{A} - [I_{n \times n}, P_1] \\ &\quad \begin{bmatrix} G_2 - G_1 & A_2^T - A_1^T \\ A_2 - A_1 & -V_2 - (-V_1) \end{bmatrix} \begin{bmatrix} I_{n \times n} \\ P_1 \end{bmatrix} \end{aligned} \quad (64)$$

where $I_{n \times n}$ is the $n \times n$ identity matrix and $\tilde{A} = A_2 - V_2P_1 - \frac{1}{2}V_2P_0$.

Since

$$[I_{n \times n}, P_1] \begin{bmatrix} G_2 - G_1 & A_2^T - A_1^T \\ A_2 - A_1 & -V_2 - (-V_1) \end{bmatrix} \begin{bmatrix} I_{n \times n} \\ P_1 \end{bmatrix} \geq 0 \quad (65)$$

for the PDRE $\Re(P_0)$ defined as

$$\begin{aligned} \dot{P}_0(x, t) &= \\ &= -\tilde{A}^T P_0 - P_0 \tilde{A} - [I, P_1] \begin{bmatrix} G_2 - G_1 & A_2^T - A_1^T \\ A_2 - A_1 & -V_2 - (-V_1) \end{bmatrix} \begin{bmatrix} I \\ P_1 \end{bmatrix} \end{aligned} \quad (66)$$

we have $\Re(P_0) = 0$, $\Re(0) \geq 0$, $P_0(t_f) \geq 0$. It is readily concluded that $P_0(t) = P_2(t) - P_1(t) \geq 0$ for all $(t, x) \in D$ by Proposition 4. So $P_1(t) \leq P_2(t)$ on D .

3.3 Application of Comparison Theorem

For the PDRE (50), it is readily verified that $Q'(x^o, t) = Q(t) + u^{oT}(t)u^o(t)R_0 \geq Q(t)$ and $B(t)R^{-1}(x^o)B^T(t) > 0$. Then we have

$$\begin{bmatrix} Q'(x^o, t) & A^T(t) \\ A(t) & -B(t)R^{-1}(x^o)B^T(t) \end{bmatrix} \geq \begin{bmatrix} Q(t) & A^T(t) \\ A(t) & -B(t)R^{-1}(x^o)B^T(t) \end{bmatrix} \quad (t, x) \in D \quad (67)$$

and

$$\begin{bmatrix} Q'(x^o, t) & A^T(t) \\ A(t) & 0 \end{bmatrix} \geq \begin{bmatrix} Q'(x^o, t) & A^T(t) \\ A(t) & -B(t)R^{-1}(x^o)B^T(t) \end{bmatrix} \quad (t, x) \in D \quad (68)$$

With Theorem 2 in Section 3.2, it follows that the solution $P(t)$ of

Table 1. The Values of Parameters in Simulation Case 01.

	x_0	x_f	P_0	P_f	$u(t)$	$P(t)$	J_1
<i>Optimal Sol.</i>	1	0.413	0.5	$3.7e-15$	$u^o(t)$	$PDRE()$	0.237653
<i>Riccati Per.</i>	1	0.417	0.5	$3.3e-08$	$u^o(t)$	$Diff.RE^*$	0.237672
<i>Control Per.</i>	1	0.421	0.5	$3.7e-15$	$u^o(t)+0.01$	$PDRE()$	0.237695

*Diff. RE: $-\dot{P}(t) = A^T(t)P(t) + P(t)A(t) - P(t)B(t)R^{-1}(t)B^T(t)P(t) + Q(t)$.
doi:10.1371/journal.pone.0094925.t001

the PDRE (50) satisfies

$$P_1(t) \leq P(t) \leq P_2(t) \quad (69)$$

where $P_1(t)$ is the solution of

$$\begin{aligned} -\dot{P}_1(t) &= A^T(t)P_1(t) + P_1(t)A(t) \\ &\quad - P_1(t)B(t)R^{-1}(x)B^T(t)P_1(t) + Q(t) \\ P_1(t_f) &= S(t_f) \end{aligned} \quad (70)$$

and $P_2(t)$ is the solution of

$$\begin{aligned} -\dot{P}_2(t) &= A^T(t)P_2(t) + P_2(t)A(t) + Q'(x,t) \\ P_2(t_f) &= S(t_f) \end{aligned} \quad (71)$$

These two equations are differential REs, which are similar to the algebraic RE in the SDRE control strategy. The analysis shows that the solution of a PDRE can be estimated by the solutions of two differential REs. Thus the methods of solving a differential RE can be borrowed to facilitate the solution to a PDRE, such as determining the initial value for $P(t)$.

Simulation

In this section, the NSLQR problem studied above is applied to two specific simulation cases to first verify the optimality numerically and then to model the goal pursuit process introduced in the section Introduction, so that the three goal pursuit behaviors, GGB, SMB and CCB are reproduced for further studying.

The NSLQR problem is, technically, a Two-Point Boundary Value (TPBV) Problem, since the initial value of the system plant (20) x_0 and the terminal value of the PDRE (21) P_f are known. A shooting method is employed in the simulation with sacent iteration to solve this TPBV problem [27]. The convergence of the method is slightly slower than being second-order with the chosen $R(x)$, as it is discussed in [28] and [29]. The solver of “ode4 Runge-Kutta” is chosen in the simulation with a fixed step of 0.01 second. The simulation error threshold is set as 10^{-6} .

4.1 Numerical Verification of the NSLQR Optimality

4.1.1 Simulation Model. In this section, we consider a specific optimal problem of seeking a control law $u(t)$ to minimize the given performance index

$$J_1 = \frac{1}{2} \int_{t_0=0}^{t_f=1} (1 \cdot x^2(t) + 4 \cdot x^2(t) \cdot u^2(t)) dt + \frac{1}{2} \cdot 0 \cdot x^2(t_f) \quad (72)$$

along the first-order LTV system as

$$\dot{x}(t) = \frac{-1}{t+1} \cdot x(t) + 1 \cdot u(t) \quad x(t_0) = 1 \quad (73)$$

with $A = \frac{-1}{t+1}$, $B = 1$, $Q = 1$, $S(t_f) = 0$ and $R = 4x^2(t)$.

Three forms of control law are considered. The first one is the optimal control law $u^o(t)$ defined in Eq (19) with the $P(t)$ from the PDRE (21). The second control law is the same $u^o(t)$ in Eq (19) but with the $P(t)$ from a standard Differential Riccati Equation, which has no additional term of $M(x,u)$. The third control law is the same $u^o(t)$ in Eq (19) with a perturbation of 0.01. To differentiate with each other, the three control laws are named as “Optimal Solution”, “Riccati Perturbation” and “Control Perturbation”, respectively. Table 1 summarizes the parameter details used in this simulation case as well as the values of performance index J_1 . Fig. 3 gives the system behaviors with the three different control laws.

4.1.2 Discussion of Results. From Fig. 3, it is valid to say that the three control laws, with the same function of $R(x)$, all bring the system state a stable behavior. However, the optimal control law $u^o(t)$ supplies the minimal performance index J_1 , as shown in Table 1. Even though it cannot be said with great confidence that the control law $u^o(t)$ is optimal since it is difficult to verify infinite numerical examples. It is still proved that the classical LQR solution, which is the “Riccati Perturbation” case, provides a greater value of J_1 than the optimal control law $u^o(t)$ does, so the classical LQR theory is not applicable to the NSLQR problem anymore. Moreover, the NSLQR theory does provide the minimal value of J_1 among all three control laws, which verifies the optimality of the NSLQR theory at some degree.

4.2 Application of the NSLQR to Goal Pursuit Processes

4.2.1 Modeling Goal Pursuit Behaviors. From a psychological perspective, the system state $x(t)$ in the NSLQR problem represents the *goal discrepancy*. The parameter $A(t)$ in system model (2) represents the *goal attraction*. For a constant A , all the eigenvalues having negative real parts (asymptotically stable) means the goal is attractive; if the eigenvalues are with non-positive real parts and those eigenvalues with zero real parts are simple (marginal stable), then the goal is neutral; otherwise (unstable), the goal is repulsive. Similar interpretations apply to time-varying $A(t)$, where the asymptotic stability, marginal stability and instability can be interpreted as attractive, neutral and repulsive goals. The input $u(t)$ represents the level of *control effort*, while the parameter $B(t)$ is treated as *control effectiveness*. In the performance index J , the weighting coefficient $Q(t)$ functions as *goal discrepancy penalty*. A greater value of $\|Q\|$ results in less discrepancy; and the weighting coefficient $S(t_f)$ is known as *terminal penalty*. A greater value of $\|S\|$ leads to smaller terminal goal discrepancy; the weighting coefficient $R(x)$ is *control energy penalty*, which depends on goal discrepancy. A greater value of $\|R\|$ means less control energy expenditure is allowed.

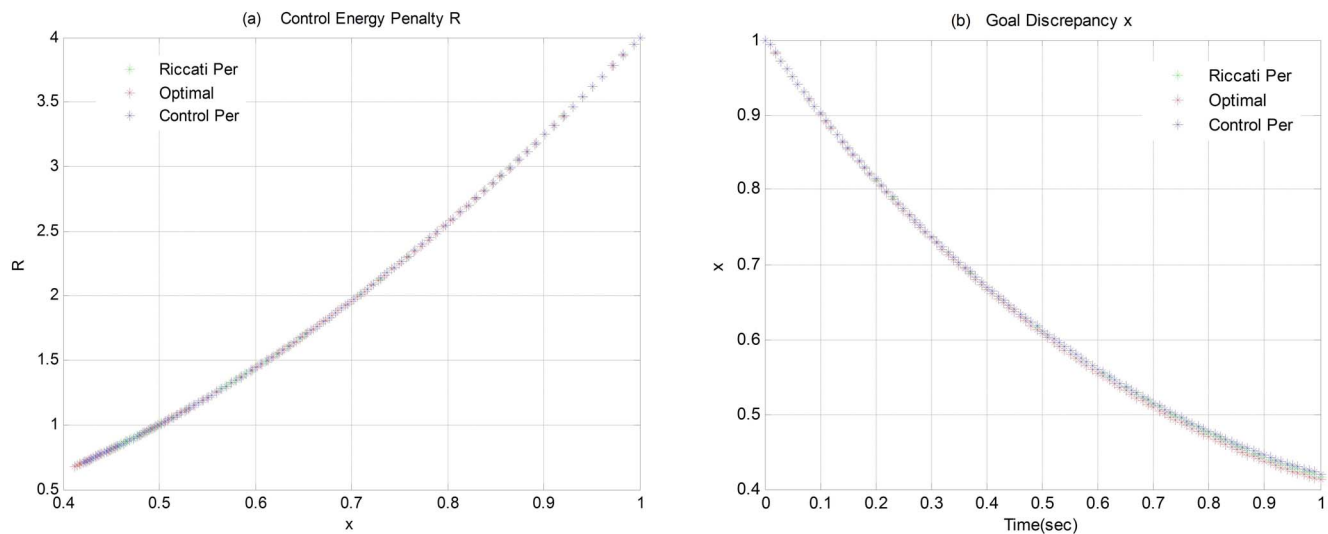


Figure 3. System Behavior of Simulation Case 01: a) Control Energy Penalty R vs. Goal Discrepancy x ; b) Goal Discrepancy x .
doi:10.1371/journal.pone.0094925.g003

4.2.2 Simulation Model. As an initial study, we consider a first-order linear goal attainment process in the simulation case. The process is set as

$$\dot{x}(t) = 0 \cdot x(t) + 1 \cdot u(t) \quad x(t_0) = 1 \quad (74)$$

which is a neutrally attractive goal process. The parameter $R(x)$ is the concern in the current simulation case study. Based on the analysis above, we hypothesize that the GGB can be produced by $R(x)$ that is a monotonically increasing function of $\|x\|$; the SMB can be produced by $R(x)$ that is a hump function of $\|x\|$; and the CCB can be produced by a constant R . Table 2 gives the value of the parameters used in the simulation. The simulation results are presented in Table 3 and Fig. 4. From Table 2, it can be seen that the choice of $R(x)$ in the GGB satisfies the convexity constraint of the function $l(x, u)$, so the solution is optimal for the GGB. Even though the convexity constraint is not satisfied in the SMB, from Corollary 1, it still can be said that the solution is optimal since $x(t_0)$ is fixed. For a meaningful comparison, the parameters are adjusted such that three behaviors achieve roughly equal terminal values, as the values of x_f shown in Table 3.

From Fig. 4 Part d), it can be seen that, the NSLQR method has successfully modeled the goal pursuit process with three behaviors. Part a) and Part b) show the control energy penalty matrix R vs time t and goal discrepancy x , respectively. The hypothesis is validated that when R is a monotonically increasing function of

$\|x\|$, the goal pursuit process can exhibit the GGB; when R is a hump function of $\|x\|$, the goal pursuit process can exhibit the SMB; and when R is a constant, the goal pursuit process can exhibit the CCB. Part e) and Part f) present the control effort u vs time t and goal discrepancy x , respectively. For the CCB, the control effort increases as it approaches the goal; for the GGB, the control effort decreases with goal discrepancy; and for the SMB, the control effort is higher at two ends than in the middle.

Table 3 lists some selected norms of the goal discrepancy $x(t)$, and the control effort $u(t)$. It shows that, with the same initial value x_0 and terminal value x_f , the CCB features the least accumulated error $\|x\|_1$, but consumes the most control energy $\|u\|_2$ and suffers from the highest stress level $\|u\|_\infty$. The SMB consumes the least control energy and suffers from the lowest stress level at the price of a higher accumulated error. The CCB is in between of these two.

4.2.3 Discussion of Results. Based on the simulation results above, it is concluded that in pursuing a goal with a finite terminal time (deadline), the GGB and the SMB behaviors may save control energy and reduce stress level over the CCB. However, the CCB has the least accumulated error. So the GGB and SMB may be beneficial in applications where only the level of goal attainment at the terminal time is of concern, such as a deadline beating process. However, the GGB or SMB would not be a preferred choice when the goal needs to be maintained over long time or needs to be approached smoothly.

Table 2. The Values of Parameters in Simulation Case 02.

	Param.	GGB	SMB	CCB
Goal Discrepancy Penalty	Q	1	1	23.5
Terminal Penalty	S	10	10	0
Control Energy Penalty	R	$3.7(x^2 + 0.1)$	$2.2(0.5 - (x - 0.5)^2)$	2
Simulation Error Threshold	E	0.000001	0.000001	0.000001
Initial Value of $x(t)$	x_0	1	1	1
Initial Value of $P(t)$	P_0	2.831	0.903	6.841

doi:10.1371/journal.pone.0094925.t002

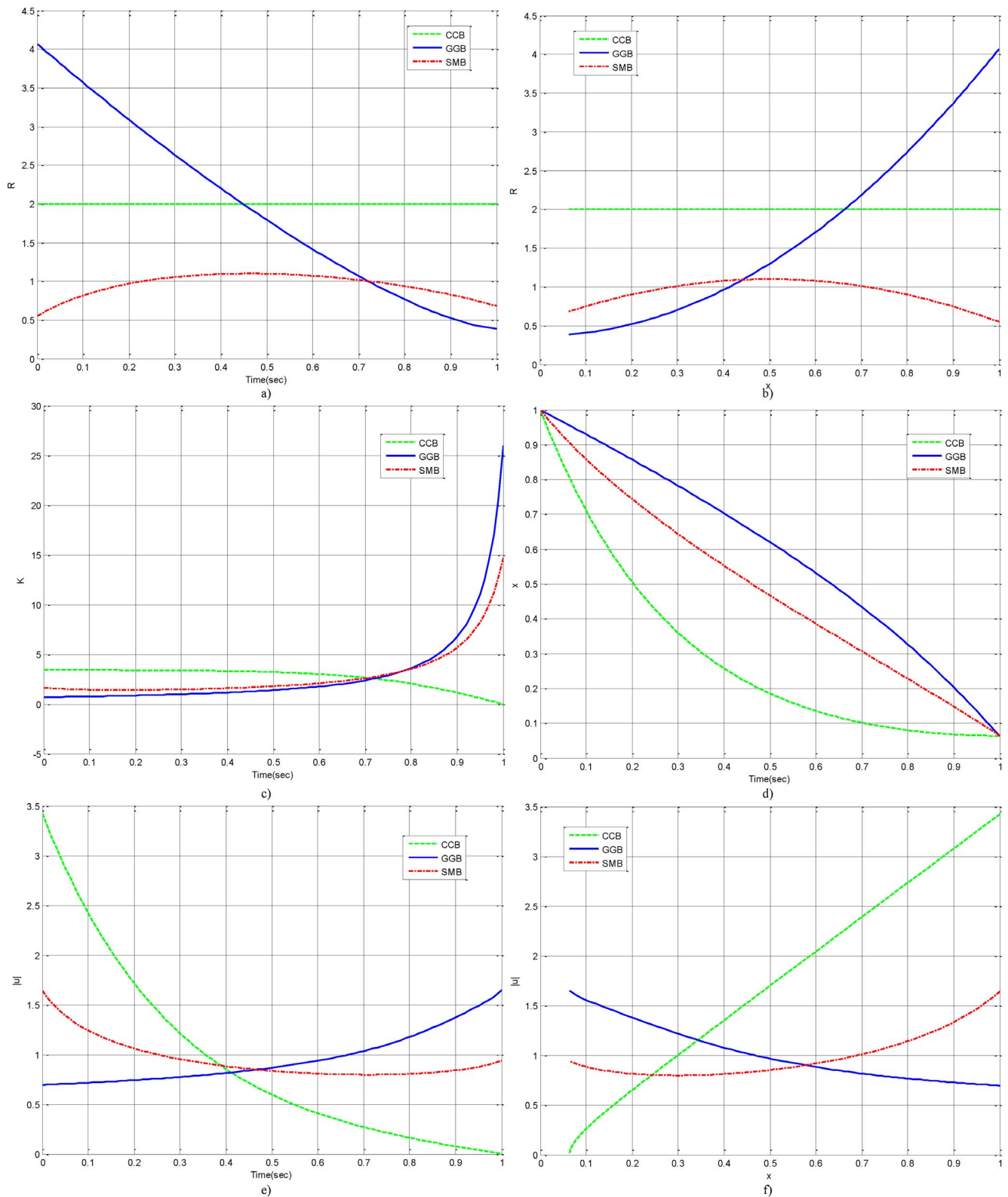


Figure 4. Three Goal Pursuit Behaviors: a) Control Energy Penalty R vs. Time; b) Control Energy Penalty R vs. Goal Discrepancy x ; c) Feedback Gain K ; d) Goal Discrepancy x ; e) Control Effort u vs. Time; f) Control Effort u vs. Goal Discrepancy x .
doi:10.1371/journal.pone.0094925.g004

Table 3. Simulation Results of Three Goal Pursuit Behaviors.

	<i>Param.</i>	<i>GGB</i>	<i>SMB</i>	<i>CCB</i>
Terminal Value of $x(t)$	x_f	0.064	0.064	0.064
Terminal Value of $P(t)$	P_f	10.000	10.000	0.000
Accumulated Error	$\ x\ _1$	59.847	49.183	29.680
Control Action	$\ u\ _1$	96.937	95.770	95.180
Control Energy	$\ u\ _2$	9.977	9.730	13.207
Stress Level	$\ u\ _\infty$	1.651	1.643	3.421

doi:10.1371/journal.pone.0094925.t003

Conclusion and Future Work

In this paper, a necessary and sufficient Optimality Theorem with rigorous mathematical proof for the NSLQR problem with a convexity constraint on $R(x)$ is presented. It is also argued that for given $x(t_0)$, the NSLQR gives an optimal solution. A Comparison Theorem for the solution of the PDRE with a general form for the NSLQR problem is presented as well. In the simulation, the NSLQR is first applied to a first-order LTV system to verify the proposed theory. The NSLQR is then used to model two psychological behaviors (GGB and SMB) in goal pursuit processes identified from psychology, along with the typical behavior of engineering control systems (CCB) by employing different control energy weighting $R(x)$. The simulation results show that the NSLQR modeling method can reproduce the three goal pursuit behaviors and the psychological goal pursuit behaviors can be more beneficial than the CCB in terms of energy saving and stress reduction in applications where only the goal discrepancy at the terminal time is of concern, such as in Marathon race, animal stalking, beating a deadline or hitting a target.

In this paper only some scalar cases of the goal pursuit process are studied; studies of the multi-variable cases are the next steps of our work. In the current study, the parameter $R(x)$ is selected to reproduce the goal pursuit behaviors. Similar results should be achievable with a state-dependent goal discrepancy weighting

$Q(x)$, which would be more akin to intuitive psychological tendency to employing the GGB and SMB strategies in terminal goal pursuit processes; whereas the control weighting modeling is more akin to conscious choice of the GGB and SMB for its energy saving and stress reduction benefits. Since for the NSLQR problem the PDRE has to be solved simultaneously with the closed-loop system, it is a TPBV problem. An inherent difficulty in this TPBV problem is how to determine the initial value of $P(t)$. An Approaching-Horizon algorithm based on a shooting method is developed to address this problem, which will be presented in detail in a separate paper.

Acknowledgments

The authors would sincerely thank Dr. J. Vancouver at the Department of Psychology, Ohio University for the psychology interpretations and inspiring discussions, and Dr. D. Lawrence at the School of Electrical Engineering and Computer Science, Ohio University for the technical discussions of the optimality. They would also thank Dr. Chen Wei at Beihang University for her valuable suggestions to polish the manuscript.

Author Contributions

Conceived and designed the experiments: XX JZ PZ. Performed the experiments: XX. Analyzed the data: XX. Contributed reagents/materials/analysis tools: XX. Wrote the paper: XX JZ PZ.

References

- Pearson J (1962) Approximation methods in optimal control. *Journal of Electronics and Control*: 453–469.
- Cook G, Wernli A (1975) Suboptimal control for the nonlinear quadratic regulator problem. *Automatica*: 75–84.
- Mracek C, Cloutier J (1998) Control designs for the nonlinear benchmark problem via the statedependent riccati equation method. *International Journal of Robust and Nonlinear Control*: 401–433.
- Friedland B (2007) Quasi-optimum control and the sdre method. In: *Proc. of the IFAC Symposium on Automatic Control in Aerospace*. Toulouse, France.
- Salamci M, Gokbilen B (2007) Sdre missile autopilot design using sliding mode control with moving sliding surfaces. In: *Proc. of the IFAC Symposium on Automatic Control in Aerospace*. Toulouse, France.
- Cimen T, McCaffrey D, Harrison R, Banks S (2007) Asymptotically optimal nonlinear filtering. In: *Proc. of the IFAC Symposium on Automatic Control in Aerospace*. Toulouse, France.
- Cimen T (2007) State-dependent riccati equation (sdre) control: A survey. In: *Proc. of the IFAC Symposium on Automatic Control in Aerospace*. Toulouse, France.
- Menon P, Ohlmeyer E (2004) Computer-aided synthesis of nonlinear autopilots for missiles. *Nonlinear Studies*: 173–198.
- Mracek C (2007) Sdre autopilot for dual controlled missiles. In: *Proc. of the IFAC Symposium on Automatic Control in Aerospace*. Toulouse, France.
- Vaddi S, Menon P, Ohlmeyer E (2007) Numerical sdre approach for missile integrated guidance - control. In: *Proc. of the AIAA Guidance, Navigation, and Control Conference*. South Carolina, USA.
- Bemporad A, Morari M, Dua V, Pistikopoulos E (2002) The explicit linear quadratic regulator for constrained systems. *Automatica* 38: 3C20.
- Alessio A, Bemporad A (2009) A survey on explicit model predictive control. *Nonlinear Model Predictive Control*: 345C369.
- Ferreau H, Bock H, Diehl M (2008) An online active set strategy to overcome the limitations of explicit mpc. *International Journal of Robust Nonlinear Control* 18: 816C830.
- Nesterov Y, Nemirovskii A (1994) *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA, US: SIAM: Society for Industrial and Applied Mathematics.
- Richter S, Jones GN, Morari M (2012) Computational complexity certification for real-time mpc with input constraints based on the fast gradient method. *IEEE Transactions on Automatic Control* 57: 1391–1403.
- Richter S (2012) *Computational Complexity Certification of Gradient Methods for Real-Time Model Predictive Control*. Ph.D, Swiss Federal Institute of Technology.
- Hull C (1932) *The goal gradient hypothesis and maze learning*. Yale University: 25–43.
- Kivets R, Urminsky O, Zheng Y (2006) The goal-gradient hypothesis resurrected: Purchase acceleration, illusory goal progress and customer retention. *Journal of Marketing Research* XIII: 39–58.
- Nunes J, Dreze X (2006) The endowed progress effect: How artificial advancement increases effort. *Journal of Consumer Research* 32: 504–512.
- Petz J, Wilson A, Strahan E (2009) So far away: The role of subjective temporal distance to future goals in motivation and behavior. *Social Cognition* 27: 475–495.
- Bonczzi A, Brendl C, Angelis M (2011) Stuck in the middle: The psychophysics of goal pursuit. *Psychological Science*.

22. Vancouver J, Weinhardt J, Schmidt A (2010) A formal, computational theory of multiple-goal pursuit: Integrating goal-choice and goal-striving processes. *Perceptual and Motor Skills* 95: 985–1008.
23. Liu S (2011) Engineering modeling of psychological goal-gradient effect phenomena. Master's project report, School of Electrical Engineering and Computer Science, Ohio University, Athens, OH, USA.
24. Air Force Office of Scientific Research and Arlington VA (2011) AFOSR Broad Agency Announcement (BAA-AFOSR-2011-01).
25. Williams R, Lawrence D (2007) *Linear State-Space Control Systems*. USA: John Wiley and Sons.
26. Coppel W (1971) *Disconjugacy* (Lecture Notes in Mathematics). USA: Springer.
27. Meade D, Haran B, White R (1996) The shooting technique for the solution of two-point boundary value problems. *BVPs, Shooting Method, and Chemical Engineering*: 1–8.
28. Koch O, Weinmüller E (2001) The convergence of shooting methods for singular boundary value problems. *Mathematicas of Computation* 72: 289–305.
29. *Two Point Boundary Value Problems*, chapter *Fundamental Problems and Methods*. pp. 3–41.