Quadratic Optimal Control of Switched Linear Stochastic Systems

Wei Zhang[a], Jianghai Hu[b], Jianming Lian[c]

[a] *Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94709*
[b] *School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907*
[c] *Center for Advanced Power Systems, Florida State University, Tallahassee, FL 32310*

Abstract

This paper studies the quadratic optimal control problem for the discrete-time switched linear stochastic system with nonautonomous subsystems perturbed by Gaussian random noises. The goal is to jointly design a deterministic switching sequence and a continuous feedback law to minimize the expectation of a finite-horizon quadratic cost function. Both the value function and the optimal control strategy are characterized analytically. A numerical relaxation framework is developed to efficiently compute a control strategy with a guaranteed performance upper bound. It is also proved that by choosing the relaxation parameter sufficiently small, the performance of the resulting control strategy can be made arbitrarily close to the optimal one.

Keywords: Hybrid Systems, LQG, uncertain switched system

1. Introduction

The study of switched systems has gained considerable momentum in recent years, resulting in numerous methods and tools for their analysis and design [4, 7, 13, 17]. A majority of these methods are based on deterministic models for subsystem dynamics. However, uncertainty is ubiquitous in realistic system models. Roughly speaking, there are two ways to handle uncertainties. The first one is to estimate *a priori* conservative bounds for the uncertainties and design the control strategy with an acceptable performance even in the worst-case scenario. Such a viewpoint is predominant in the previous studies of switched uncertain systems. For example, in [10, 14], some convex conditions are established for quadratic and asymptotic stabilizability of switched linear systems with bounded parametric uncertainties. In [18], an L_2 controller synthesis method is proposed for switched linear systems with both parametric

Email addresses: weizhang@eecs.berkeley.edu (Wei Zhang), jianghai@purdue.edu (Jianghai Hu), lian@caps.fsu.edu (Jianming Lian)

uncertainties and exogenous disturbances. A robust controller synthesis method is proposed in [16] for a special class of switched nonlinear systems with general norm-bounded disturbances. On the other hand, instead of considering the worst case, one can also describe uncertainties using stochastic models, and design a control strategy to meet certain design criteria in the probabilistic sense. In [12], the switched linear Gaussian system is studied and some convex conditions are derived for the existence of a finite-path-dependent state-feedback controller which achieves certain required expected output-regulation level.

This paper focuses on the discrete-time switched linear Gaussian systems (SLGS) with nonautonomous subsystems perturbed by Gaussian random noises. The goal is to jointly design a deterministic switching sequence and a continuous feedback law to minimize the expectation of a finite-horizon quadratic cost function along the closed-loop trajectory. When making the control decision, we assume that the continuous control has access to the noisy measurements of the continuous state, while the switching control is determined at time zero based on the covariance of the initial continuous state. It is worth mentioning that this problem is related to but substantially different from the quadratic optimal control problem of the Markovian jump linear system (MJLS) that has been extensively studied in the literature [6, 8, 9]. For the MJLS, the switching sequence is modeled as a Markov chain that evolves according to some given transition probability matrix, while for the SLGS, the switching sequence is a control variable that can be chosen freely to minimize the cost function.

In [15], a pruning algorithm was proposed to compute a suboptimal switching sequence and a continuous control law for a SLGS. Our contribution beyond [15] is mainly twofold. Firstly, analytical characterizations are derived for both the value function and the optimal control strategy using the *switched Riccati sets*, which was originally introduced to study the switched LQR problem [19]. These analytical expressions reveal several important properties of the underlying problem. It is shown that the optimal continuous control law at each time step is a linear function of the continuous state, the optimal switching control at each time step depends only on the covariance matrix of the system state, and the value function is a piecewise affine function of the covariance matrix of the initial state. Secondly, a relaxation scheme is developed to efficiently compute a suboptimal hybrid-control strategy. It is shown that by choosing a sufficiently small relaxation parameter, the cost associated with the resulting control strategy can be made arbitrarily close to the optimal one.

This paper is organized as follows. The quadratic optimal control problem to be studied is formulated in Section 2. Its value function and optimal control strategy are derived analytically in Section 3 and some related properties are also discussed. In Section 4, a numerical relaxation framework is proposed and a performance upper bound of the resulting controller is derived. Several simulation examples are given in Section 5.

Notations: Let n, d, M, p and N be some positive integers. Denote by $\mathbb{Z}_+$ the set of positive integers, by $\mathcal{A}$ the set of positive semidefinite (p.s.d.) matrices, by $2^{\mathcal{A}}$ the power set of $\mathcal{A}$, and by I_d and I_n the identity matrices of dimension d and n, respectively. Let $\mathbb{M} \triangleq \{1, \ldots, M\}$ be the set of subsystem

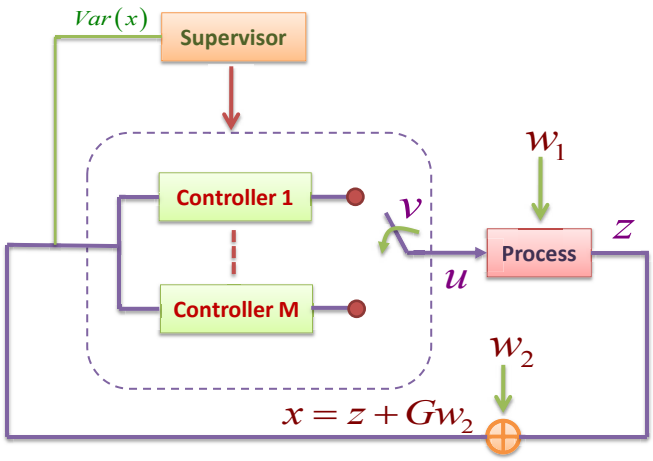

Figure 1: A supervisory control example.

indices, $T_N \triangleq \{0, \ldots, N-1\}$ be the control horizon, $\|\cdot\|$ be the Euclidean norm of a given vector as well as the matrix norm induced by the Euclidean norm, and $\lambda_{\min}(\cdot)$ be the smallest eigenvalue of a p.s.d. matrix. We write $P \succ 0$ ($P \succeq 0$) if P is strictly positive definite (semidefinite).

2. Problem Formulation

2.1. A Motivating Example

As a motivating example, we consider a supervisory control problem as shown in Fig. 1. Suppose that the process dynamics and the measurement equation are given by:

$$\begin{cases} z(t+1) = Az(t) + Bu(t) + Fw_1(t) \\ x(t) = z(t) + Gw_2(t), \end{cases}$$

where A, B, F and G are constant matrices of appropriate dimensions, and $w_1(t)$ and $w_2(t)$ are two independent standard Gaussian processes on $\mathbb{R}^d$. Suppose that M different linear controllers, characterized by the feedback gains $\{K_i\}_{i=1}^M$, are given *a priori*. On top of the linear controllers, there is a supervisory unit that determines which linear controller is being used at each time instant. If controller $v(t)$ is selected at time t, then the overall dynamics of the measured state $x(t)$ is given by

$$x(t+1) = \left(A + BK_{v(t)}\right) x(t) + Gw_2(t+1) + Fw_1(t) - AGw_2(t).$$

Let $\tilde{A}_i = A + BK_i$, $i \in \mathbb{M}$ and denote by $\tilde{F}$ the square root of the covariance matrix of $Gw_2(t+1) + Fw_1(t) - AGw_2(t)$. Then the above dynamical system

is equivalent to

$$x(t+1) = \tilde{A}_{v(t)}x(t) + \tilde{F}\tilde{w}(t),$$

for some standard Gaussian process $\tilde{w}(t)$ on $\mathbb{R}^d$.

Due to the randomness of the measurement, the switching sequence should be selected based on some statistical information of the system instead of the particular noisy measurement of the system state. Since the closed-loop system is a linear Gaussian process, the covariance matrix of $x(t)$ captures all the probabilistic behavior of the system. Therefore, an interesting problem is to design a switching sequence based on the current state covariance matrix to minimize the expectation of a quadratic cost function over a certain look-ahead horizon. This is a special case of the general problem to be introduced in the next subsection.

2.2. General Problem Statement

Suppose the state measurement of a dynamical process can be described by the following discrete-time switched linear stochastic system:

$$x(t+1) = A_{v(t)}x(t) + B_{v(t)}u(t) + F_{v(t)}w(t), \tag{1}$$

where at each time $t \in T_N$, $x(t) \in \mathbb{R}^n$ is the continuous state, $u(t) \in \mathbb{R}^p$ is the continuous control, $v(t) \in \mathbb{M}$ is the switching control that determines the discrete mode of the system, and $w(t) \in \mathbb{R}^d$ is the disturbance. For each $i \in \mathbb{M}$, A_i, B_i and F_i are constant matrices of appropriate dimensions. Let $\Sigma_i = F_i F_i^T$ for each $i \in \mathbb{M}$ and define

$$\mathbf{u} \triangleq (u(0), \ldots, u(N-1)) \text{ and } \mathbf{v} \triangleq (v(0), \ldots, v(N-1)).$$

We assume that $\{w(t)\}_{t\in T_N}$ is a d-dimensional standard Gaussian process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and the initial state $x(0)$ is a zero-mean Gaussian on $(\Omega, \mathcal{F}, \mathbb{P})$ that is independent of $w(t)$ for any $t \in T_N$, with covariance matrix ϕ. Let $\{\mathcal{F}_t\}_{t\in T_N}$ be the filtration associated with the measured state sequence $\{x(t)\}_{t\in T_N}$. Denote by $\mathbb{E}(\cdot)$ the expectation with respect to the probability measure $\mathbb{P}$.

The goal is to regulate the system state with a reasonable control cost subject to the random disturbances. When making the control decision, we assume that the continuous control sequence $\{u(t)\}_{t\in T_N}$ is adapted to the filtration $\{\mathcal{F}_t\}_{t\in T_N}$, i.e., $u(t)$ is measurable with respect to $\mathcal{F}_t$, for each $t \in T_N$, while the switching control sequence $\{v(t)\}_{t\in T_N}$ is deterministic. In other words, the continuous control has full access to the measured continuous state $x(t)$, while the switching control is determined deterministically at time $t = 0$. These assumptions are particularly useful when we have a group of feedback controllers and try to jointly design both their feedback laws and their deterministic scheduling to improve the regulation performance. Denote by $\mathcal{U}^N$ the set of all the R^p-valued sequences that are adapted to the filtration $\{\mathcal{F}_t\}_{t\in T_N}$ and by $\mathbb{M}^N$ the set of all the possible deterministic switching sequences of length N. The performance

of the sequences $\mathbf{u} \in \mathcal{U}^N$ and $\mathbf{v} \in \mathbb{M}^N$ can be measured by the following cost function:

$$J_N(\phi; \mathbf{v}, \mathbf{u}) = \mathbb{E}\big[x(N)^T Q_f x(N) + \sum_{t=0}^{N-1} \big(x(t)^T Q_{v(t)} x(t) + u(t)^T R_{v(t)} u(t)\big)\big], \quad (2)$$

where $\{Q_i\}_{i\in\mathbb{M}}$, $\{R_i\}_{i\in\mathbb{M}}$, and Q_f are weighting matrices of appropriate dimensions. Denote $\lambda_Q^- = \min_{i\in\mathbb{M}}\{\lambda_{\min}(Q_i)\}$ and $\lambda_R^- = \min_{i\in\mathbb{M}}\{\lambda_{\min}(R_i)\}$. We assume that the weighting matrices are chosen so that $\lambda_Q^- > 0$ and $\lambda_R^- > 0$. The goal of this paper is to solve the following problem.

Problem 1. *Find* $\mathbf{v} \in \mathbb{M}^N$ *and* $\mathbf{u} \in \mathcal{U}^N$ *to minimize* $J_N(\phi; \mathbf{v}, \mathbf{u})$ *subject to the dynamical equation (1) with initial distribution* $x(0) \sim \mathcal{N}(0, \phi)$.

3. Analytical Solution Using Dynamic Programming

Problem 1 is a Markov decision problem that can be solved using dynamic programming. The value function associated with this problem is determined by the initial covariance matrix instead of some particular value of the initial state. For a generic initial covariance $\phi \in \mathcal{A}$ and a nonnegative integer $k \in \mathbb{Z}_+$, the k-horizon value function is defined by

$$V_k(\phi) = \min_{\mathbf{u}\in\mathcal{U}^k,\, \mathbf{v}\in\mathbb{M}^k} J_k(\phi; \mathbf{u}, \mathbf{v}). \quad (3)$$

The goal of this section is to derive analytical expressions for the N-horizon value function $V_N(\phi)$ and the corresponding optimal controls $(\mathbf{u}^*, \mathbf{v}^*)$.

3.1. Analytical Value Function

Let $\rho_i : \mathcal{A} \to \mathcal{A}$ denote the *Riccati mapping* of subsystem $i \in \mathbb{M}$, i.e.,

$$\rho_i(P) = Q_i + A_i^T P A_i - A_i^T P B_i (R_i + B_i^T P B_i)^{-1} B_i^T P A_i, \, \forall P \in \mathcal{A}. \quad (4)$$

When $\mathbf{v} \in \mathbb{M}^N$ is fixed, system (1) reduces to a stochastic linear time-varying system. By the standard LQG theory [1], the N-horizon value function of Problem 1 in this special case is $V_N(\phi) = \sum_{t=0}^N t\, r(P_t \phi)$, where $\{P_t\}_{t=0}^N$ is a sequence of p.s.d. matrices generated by the Riccati recursion:

$$P_t = \rho_{\mathbf{v}(t)}(P_{t+1}), t \in T_N, \text{ with } P_0 = Q_f.$$

Following a similar idea, we can use the Riccati mapping to compute the general value function V_N by moving backward in time. Starting from the terminal weighting matrix Q_f, after one iteration, we will have M matrices $\{\rho_i(Q_f)\}_{i\in\mathbb{M}}$ depending on which subsystem is used at time $N-1$. Each of these M matrices will again be mapped to another set of M matrices in the next iteration depending on the subsystem used at time $N-2$. Therefore, if we consider all the possible switching sequences, instead of having a matrix-valued iteration as in the traditional linear system case, we will have a set-valued iteration that produces a sequence of sets.

Definition 1 (Switched Riccati Mapping). *The mapping* $\rho_{\mathbb{M}} : 2^{\mathcal{A}} \to 2^{\mathcal{A}}$ *defined by:* $\rho_{\mathbb{M}}(\mathcal{H}) = \{\rho_i(P) : i \in \mathbb{M} \text{ and } P \in \mathcal{H}\}$, $\forall \mathcal{H} \in 2^{\mathcal{A}}$, *is called the* Switched Riccati Mapping (SRM) *associated with system (1).*

Definition 2 (Switched Riccati Set). *The sequence of sets* $\{\mathcal{S}_k\}_{k=0}^{N}$ *generated iteratively by* $\mathcal{S}_{k+1} = \rho_{\mathbb{M}}(\mathcal{S}_k)$ *with* $\mathcal{S}_0 = \{Q_f\}$ *is called the* Switched Riccati Sets (SRS) *associated with system (1).*

The sequence of switched Riccati sets always starts from the singleton $\{Q_f\}$ and evolves according to the switched Riccati mapping. It is easy to see that each matrix in $\mathcal{S}_N$ corresponds to a matrix generated by the Riccati mapping along a particular N-horizon switching sequence. As proved in [20], the optimal strategy and the value function of the discrete-time switched LQR (DSLQR) problem are exactly characterized by the switched Riccati sets. Following a similar idea, we can also characterize the value function of Problem 1 using a variant of the switched Riccati set. Let $\mathcal{H}_N$ be the set of ordered pairs of matrices defined recursively by:

$$\begin{aligned} \mathcal{H}_0 = \{(Q_f, 0)\}, \text{ and } \mathcal{H}_{k+1} = h(\mathcal{H}_k),\ k \in T_N \\ \text{with} \quad h(\mathcal{H}_k) \triangleq \{\big(\rho_i(P), \Gamma + P\Sigma_i\big) : i \in \mathbb{M}, (P, \Gamma) \in \mathcal{H}_k\}. \end{aligned} \tag{5}$$

The sets $\{\mathcal{H}_k\}_{k=0}^{N}$ defined above are called the *characteristic sets* associated with Problem 1 as they completely characterize the value function and the optimal strategy of Problem 1.

Remark 1. *The characteristic set* $\mathcal{H}_k$ *differs from the SRS* $\mathcal{S}_k$ *in its additional component* Γ *that captures the accumulated cost over the past* k *steps.*

Theorem 1. *The* N*-horizon value function of Problem 1 is*

$$V_N(\phi) = \min_{(P,\Gamma)\in\mathcal{H}_N} tr(P\phi + \Gamma). \tag{6}$$

In addition, if $u^*(0)$ *and* $v^*(0)$ *are the optimal continuous control and the optimal switching control at time* $t = 0$ *of the* N*-horizon problem, then*

$$\begin{aligned} u^*(0) = \mu_N^*(x;\phi) \triangleq -K_{\nu_N^*(\phi)}(P_N^*(\phi)) \cdot x(0), \text{ and } v^*(0) \triangleq \nu_N^*(\phi), \\ \text{where} \quad (\nu_N^*(\phi), P_N^*(\phi), \Gamma_N^*(\phi)) = \operatorname*{arg\,min}_{i\in\mathbb{M},(P,\Gamma)\in\mathcal{H}_{N-1}} tr(\rho_i(P)\phi + \Gamma + P\Sigma_i), \end{aligned} \tag{7}$$

and $K_\cdot(\cdot)$ *denotes the Kalman gain function defined as*

PROOF. The theorem can be proved using induction. When $N = 1$, we have

$$\begin{aligned} V_1(\phi) = \min_{v\in\mathbb{M},u\in\mathcal{U}} \mathbb{E}\big[x(0)^T Q_v x(0) + u^T R_v u \\ + (A_v x(0) + B_v u + F_v w(0))^T Q_f (A_v x(0) + B_v u + F_v w(0))\big], \end{aligned} \tag{8}$$

where $\mathcal{U}$ denotes the set of all the $\mathbb{R}^p$-valued random variables that are measurable with respect to $\mathcal{F}_0$. Since $w(0)$ is independent of $x(0)$, it can be easily derived that for each $v \in \mathbb{M}$, the optimal continuous control should be $u^*(0) = -K_v(Q_f)x(0)$. Substituting this into the above equation and simplifying the expression yields

$$\begin{aligned} V_1(\phi) &= \min_{v\in\mathbb{M}} \mathbb{E}\big[x(0)^T \rho_v(Q_f)x(0) + w(0)^T F_v^T Q_f F_v w(0)\big] \\ &= \min_{v\in\mathbb{M}} t\,r\big(\rho_v(Q_f)\phi + Q_f\Sigma_v\big). \end{aligned}$$

Since $\mathcal{H}_0 = \{(Q_f, 0)\}$ and $\mathcal{H}_1 = \{(\rho_i(Q_f), Q_f\Sigma_i)\}_{i\in\mathbb{M}}$, we have that $V_1(\phi) = \min_{(P,\Gamma)\in\mathcal{H}_1} t\,r(P\phi+\Gamma)$ and the continuous and the discrete controls that achieve this minimum are given by equation (7). Therefore, the results hold for $N = 1$.

We now suppose that the results hold for a general positive integer N. Let $v^*(0)$ and $u^*(0)$ be the optimal switching control and the optimal continuous control, respectively, at time $t = 0$ of an $(N + 1)$-horizon problem. Define

$$x^*(1) = A_{v^*(0)}x(0) + B_{v^*(0)}u^*(0) + F_{v^*(0)}w(0).$$

By a standard result of LQG theory, we know that for each fixed switching control sequence, the optimal system trajectory is a zero-mean Gaussian process. In particular, under the optimal deterministic switching sequence v^*, the optimal state $x^*(1)$ at time $t = 1$ is a zero-mean Gaussian random variable with covariance $\Phi^*(1) = \mathbb{E}\left(x^*(1)\cdot x^*(1)^T\right)$. By the Bellman's principle of optimality, we know that

$$\begin{aligned} &V_{N+1}(\phi) = \mathbb{E}\big[x(0)^T Q_{v^*(0)}x(0) + u^*(0)^T R_{v^*(0)}u^*(0) + V_N(\Phi^*(1))\big] \\ &= \min_{(P,\Gamma)\in\mathcal{H}_N} \Big\{\mathbb{E}\big[x(0)^T Q_{v^*(0)}x(0) + u^*(0)^T R_{v^*(0)}u^*(0)\big] + t\,r(P\Phi^*(1) + \Gamma)\Big\} \\ &= \min_{(P,\Gamma)\in\mathcal{H}_N} \Big\{t\,r\left(\Gamma + Q_{v^*(0)}\phi\right) + \mathbb{E}\big[u^*(0)^T R_{v^*(0)}u^*(0) + x^*(1)Px^*(1)\big]\Big\} \\ &= \min_{(P,\Gamma)\in\mathcal{H}_N} \Big\{t\,r\big(\Gamma + Q_{v^*(0)}\phi + \Sigma_{v^*(0)}P + A_{v^*(0)}^T P A_{v^*(0)}\phi\big) \\ &\quad + \mathbb{E}\big[u^*(0)^T(R_{v^*(0)} + B_{v^*(0)}^T P B_{v^*(0)})u^*(0) + 2x(0)^T A_{v^*(0)}^T P B_{v^*(0)}u^*(0)\big]\Big\}. \end{aligned}$$

Clearly, $u^*(0)$ should minimize the quadratic term inside the square bracket at the last line, which leads to $u^*(0) = -K_{v^*(0)}(P)x(0)$. Substituting this into the above equation, we obtain

$$V_{N+1}(\phi) = \min_{(P,\Gamma)\in\mathcal{H}_N} t\,r(\rho_{v^*(0)}(P)\phi + \Gamma + P\Sigma_{v^*(0)}).$$

We shall choose (P, Γ) and $v^*(0)$ to minimize the quantity inside the parentheses of the above equation, which implies that the optimal discrete and continuous controls at time $t = 0$ are as defined in (7). In addition, by the definition of $\mathcal{H}_{N+1}$, we also have $V_{N+1}(\phi) = \min_{(P,\Gamma)\in\mathcal{H}_{N+1}} t\,r(P\phi + \Gamma)$. □

Remark 2. *It is worth mentioning that the value function V_N can also be characterized through the enumeration over all the switching control sequences $\mathbf{v} \in \mathbb{M}^N$. The above theorem properly transforms the enumeration over the switching sequences in $\mathbb{M}^N$ to the enumeration over the pairs of matrices in $\mathcal{H}_N$. It will become clear in Section 4 that the expression given by (6) is more convenient for the analysis and the efficient computation of the suboptimal solutions of Problem 1.*

Remark 3. *Theorem 1 reveals several important properties of Problem 1. Firstly, as the trace operation $\mathrm{t}\, r(\cdot)$ is a linear operator, the value function V_N is a piecewise affine function on the semi-definite cone $\mathcal{A}$. Secondly, the linearity*[1] *of $u^*(0) = \mu_N(x(0); \phi)$ in $x(0)$ implies that the optimal trajectory $\{x^*(t)\}_{t\in T_N}$ is a zero-mean Gaussian process. Finally, the optimal switching control at any time $t \in T_N$ depends only on the covariance matrix $\Phi^*(t) = \mathbb{E}(x^*(t)x^*(t)^T)$, while the continuous control, which has full access to the random continuous state, depends on both the covariance $\Phi^*(t)$ and the value of $x^*(t)$.*

3.2. Feedback Policy

It is often beneficial to express the control action at a particular time step as a feedback law that is determined by certain measurements. Such an expression depends largely on the information available for making the control decision. For Problem 1, as mentioned in Remark 3, the optimal switching control depends only on the covariance of the state, while the optimal continuous control depends on both the covariance and the value of the state. This motivates us to consider the feedback laws defined as follows.

For each $k = 1, \ldots, N$, let $\mu_k : \mathbb{R}^n \times \mathcal{A} \to \mathbb{R}^p$ and $\nu_k : \mathcal{A} \to \mathbb{M}$ be the *continuous-control law* and the *switching-control law*, respectively, applied at time $t = N - k$, i.e., with k steps to go. Denote by $\mathcal{L}$ the set of all the continuous-control laws $\mu(x, \phi)$ that are linear in x and by $\mathcal{V}$ the set of all the switching-control laws. Define the *hybrid-control law* at time $t = N - k$ as $\xi_k \triangleq (\mu_k, \nu_k) : \mathbb{R}^n \times \mathcal{A} \to \mathbb{R}^p \times \mathbb{M}$. Denote by $\pi_N = \{\xi_N, \ldots, \xi_1\}$ the N-horizon hybrid-control policy. According to Theorem 1, to study Problem 1, it suffices to consider only the policies whose continuous-control law $\mu_k \in \mathcal{L}$ for all $k = 1, \ldots, N$. Denote by Π_N the set of all *such* N-horizon control policies. A policy $\pi_N \in \Pi_N$ governs the evolution of the system as follows:

$$\begin{cases} x(t+1) = A_{\nu_k(\Phi(t))}x(t) + B_{\nu_k(\Phi(t))}\mu_k(x(t), \Phi(t)) + F_{\nu_k(\Phi(t))}w(t) \\ \Phi(t+1) = \mathbb{E}(x(t+1)x(t+1)^T), \end{cases} \tag{9}$$

for $t \in T_N$ and $k = N - t$. The cost associated with a policy $\pi_N \in \Pi_N$ with

[1] The linearity of the optimal control is not guaranteed even for some LQG type of problems [11]. It depends largely on the information structure available for making the optimal decision.

initial covariance ϕ is

$$J_{\pi_N}(\phi) = \mathbb{E}\Big[x(N)^T Q_f x(N) + \sum_{t=0}^{N-1} \Big(x(t)^T Q_{\nu_k(\Phi(t))} x(t) + \mu_k(x(t), \Phi(t))^T R_{\nu_k(\Phi(t))} \mu_k(x(t), \Phi(t))\Big)\Big],$$

where $k = N - t$ and $x(t)$ is the closed-loop trajectory of (9) driven by π_N with initial covariance ϕ.

Theorem 1 implies that solving Problem 1 is equivalent to finding a policy $\pi_N \in \Pi_N$ that minimizes $J_{\pi_N}(\phi)$ and the value function V_N can also be written as $V_N(\phi) = \min_{\pi_N \in \Pi_N} J_{\pi_N}(\phi)$. Furthermore, the optimal policy that achieves the minimum cost can also be directly obtained from Theorem 1.

Corollary 1. *For each $k = 1, \ldots, N$, let μ_k^* and ν_k^* be defined by (7) with N replaced by k. Then the policy $\pi_N^* = \{(\mu_N^*, \nu_N^*), \ldots, (\mu_1^*, \nu_1^*)\}$ is the optimal policy that achieves the minimum cost $V_N(\phi)$ for any initial covariance $\phi \in \mathcal{A}$.*

The notion of feedback laws and polices introduced in this subsection allows us to describe the future evolution of the system only based on the value and covariance matrix of the current state, which in turn enables a recursive characterization of the value function through the so-called *one-stage* value iteration.

$$\begin{aligned} V_{k+1}(\phi) &= \min_{\mu \in \mathcal{L}, \nu \in \mathcal{V}} \Big\{\mathbb{E}_x(x, \mu(x, \phi), \nu(\phi)) + V_k(\Phi_{\mu,\nu}(1; \phi))\Big\} \\ &\triangleq \min_{\mu \in \mathcal{L}, \nu \in \mathcal{V}} \Big\{L(x, \mu(x, \phi), \nu(\phi)) + V_k(\Phi_{\mu,\nu}(1; \phi))\Big\}, \forall \phi \in \mathcal{A}, \end{aligned} \tag{10}$$

where $x \sim \mathcal{N}(0, \phi)$, $\Phi_{\mu,\nu}(1; \phi)$ is the covariance of $A_{\nu(\phi)}x + B_{\nu(\phi)}\mu(x, \phi)$, and $L \triangleq \mathbb{E}_x$ simply denotes the expectation operation with respect to the distribution of x. Note that the operator L can also be viewed as the running cost function of the value iteration.

4. Efficient Suboptimal Solution

According to (10), at iteration k, the value functions and the optimal feedback laws at all the future iterations only depend on the current value function V_k that is completely characterized by the set $\mathcal{H}_k$. As k increases, since $|\mathcal{H}_k|$ grows exponentially fast, it becomes increasingly challenging to obtain an exact representation of V_k. A natural way of simplifying the computation is to ignore some less important pairs in $\mathcal{H}_k$ to obtain an approximate representation of V_k. This approximation should be close enough to V_k so that its effect on the future value iterations is negligible. The goal of this section is to develop an efficient way to compute a performance-guaranteed approximation of V_k and to derive an upper bound for the cost associated with the policy generated by the approximate value functions.

4.1. Equivalent Subsets and Suboptimal Strategy

To formalize the above idea, we introduce some definitions. Let $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_0]^T \in \mathbb{R}^2_+$ be the *numerical relaxation vector*. Each relaxation vector corresponds to an *affine error function* $g^{\boldsymbol{\epsilon}} : \mathcal{A} \to [0, \infty)$ defined as:

$$g^{\boldsymbol{\epsilon}}(\phi) = \epsilon_1 \, t\, r(\phi) + \epsilon_0. \tag{11}$$

Hence, for any $\phi \in \mathcal{A}$, the error function $g^{\boldsymbol{\epsilon}}(\phi) \to 0$ as $\|\boldsymbol{\epsilon}\| \to 0$. In the rest of this section, $\mathcal{H}$ denotes a generic set of matrix pairs $(P, \Gamma) \in \mathcal{A} \times \mathcal{A}$.

Definition 3. *A pair of matrices $(\tilde{P}, \tilde{\Gamma}) \in \mathcal{H}$ is called $\boldsymbol{\epsilon}$-redundant with respect to $\mathcal{H}$ if*

$$\min_{(P,\Gamma)\in\mathcal{H}\setminus\{(\tilde{P},\tilde{\Gamma})\}} t\, r(P\phi + \Gamma) \le g^{\boldsymbol{\epsilon}}(\phi) + \min_{(P,\Gamma)\in\mathcal{H}} t\, r\big(P\phi + \Gamma\big),\ \forall \phi \in \mathcal{A}.$$

The following lemma provides an equivalent definition of the $\boldsymbol{\epsilon}$-redundancy.

Lemma 1. *A pair $(\tilde{P}, \tilde{\Gamma}) \in \mathcal{H}$ is $\boldsymbol{\epsilon}$-redundant if and only if for any $\phi \in \mathcal{A}$, there always exists another pair $(P, \Gamma) \in \mathcal{H} \setminus \{(\tilde{P}, \tilde{\Gamma})\}$ such that $t\, r(\tilde{P}\phi + \tilde{\Gamma}) \ge t\, r(P\phi + \Gamma) - g^{\boldsymbol{\epsilon}}(\phi)$.*

PROOF. Straightforward.

Definition 4 ($\boldsymbol{\epsilon}$-ES). *The set $\mathcal{H}^{\boldsymbol{\epsilon}}$ is called an $\boldsymbol{\epsilon}$-equivalent subset ($\boldsymbol{\epsilon}$-ES) of $\mathcal{H}$ if $\mathcal{H}^{\boldsymbol{\epsilon}} \subseteq \mathcal{H}$ and*

$$\min_{(P,\Gamma)\in\mathcal{H}^{\boldsymbol{\epsilon}}} t\, r(P\phi + \Gamma) \le g^{\boldsymbol{\epsilon}}(\phi) + \min_{(P,\Gamma)\in\mathcal{H}} t\, r(P\phi + \Gamma),\ \forall \phi \in \mathcal{A}.$$

The error introduced by removing the $\boldsymbol{\epsilon}$-redundant pairs is no larger than $g^{\boldsymbol{\epsilon}}(\phi)$. To simplify the computation, at iteration step k, we shall remove as many $\boldsymbol{\epsilon}$-redundant pairs in $\mathcal{H}_k$ as possible. This idea is feasible only when we have an efficient way to test whether a pair in $\mathcal{H}_k$ is $\boldsymbol{\epsilon}$-redundant or not. A well-known fact on symmetric matrices given in the following lemma is useful for developing such a test.

Lemma 2. *For any P_1, P_2 and $\phi \in \mathcal{A}$, if $P_1 \succeq P_2$, then $t\, r(P_1\phi) \ge t\, r(P_2\phi)$.*

With the above lemma, we are able to provide a sufficient condition to test the $\boldsymbol{\epsilon}$-redundancy.

Lemma 3 (Convex Redundancy Test). *A pair $(\tilde{P}, \tilde{\Gamma}) \in \mathcal{H}$ is $\boldsymbol{\epsilon}$-redundant in $\mathcal{H}$ if there exist nonnegative constants $\{\alpha_i\}_{i=1}^{l-1}$ such that*

$$\sum_{i=1}^{l-1} \alpha_i = 1, \quad \begin{bmatrix} \tilde{P} + \epsilon_1 I_n & 0 \\ 0 & t\, r(\tilde{\Gamma}) + \epsilon_0 \end{bmatrix} \succeq \sum_{i=1}^{l-1} \alpha_i \begin{bmatrix} P^{(i)} & 0 \\ 0 & t\, r(\Gamma^{(i)}) \end{bmatrix}, \tag{12}$$

where $l = |\mathcal{H}|$ and $\{(P^{(i)}, \Gamma^{(i)}\}_{i=1}^{l-1}$ is an enumeration of $\mathcal{H} \setminus \{(\tilde{P}, \tilde{\Gamma})\}$.

PROOF. Let $\phi \in \mathcal{A}$ be arbitrary and suppose that the condition in (12) holds. By Lemma 2, we know that

$$t\,r(\tilde{P}\phi) + \epsilon_1 t\,r(\phi) \geq \sum_{i=1}^{l-1} \alpha_i \cdot t\,r(P^{(i)}\phi), \text{ and } t\,r(\tilde{\Gamma}) + \epsilon_0 \geq \sum_{i=1}^{l-1} \alpha_i \cdot t\,r(\Gamma^{(i)}).$$

Hence,

$$t\,r(\tilde{P}\phi + \tilde{\Gamma}) + g^{\boldsymbol{\epsilon}}(\phi) \geq \sum_{i=1}^{l-1} \alpha_i \cdot t\,r(P^{(i)}\phi + \Gamma^{(i)}) \geq t\,r(P^{(i_\phi)}\phi + \Gamma^{(i_\phi)}),$$

for some $i_\phi \in \{1, \ldots, l-1\}$. Since ϕ is arbitrary, the desired result follows from Lemma 1. □

Algorithm 1 $\left[Algo_{\boldsymbol{\epsilon}}(\mathcal{H}_k^{\epsilon})\right]$

Set $\mathcal{H}_k^{\boldsymbol{\epsilon}} = \emptyset$.
for each $(P, \Gamma) \in \mathcal{H}_k$ **do**
  if (P, Γ) does NOT satisfies the condition in Lemma 3 with respect to $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ **then**
    $\mathcal{H}_k^{\boldsymbol{\epsilon}} = \mathcal{H}_k^{\boldsymbol{\epsilon}} \cup \{(P, \Gamma)\}$;
  end if
end for
Return $\mathcal{H}_k^{\boldsymbol{\epsilon}}$.

Checking the condition in Lemma 3 is an LMI feasibility problem [3] and can be easily solved through convex optimization [2]. Based on this lemma, an efficient algorithm (Algorithm 1) is developed to compute an $\boldsymbol{\epsilon}$-ES of $\mathcal{H}_k$ for any $k \in T_N$. Denote by $Algo_{\boldsymbol{\epsilon}}(\mathcal{H}_k)$ the set returned by this algorithm. The algorithm can be applied after each iteration in (5) to generate a sequence of $\boldsymbol{\epsilon}$-relaxed characteristic sets:

$$\mathcal{H}_0^{\boldsymbol{\epsilon}} = \{(Q_f, 0)\} \text{ and } \mathcal{H}_{k+1}^{\boldsymbol{\epsilon}} = Algo_{\boldsymbol{\epsilon}}(h(\mathcal{H}_k^{\boldsymbol{\epsilon}})),\ k \in T_N. \tag{13}$$

The sets $\{\mathcal{H}_k^{\boldsymbol{\epsilon}}\}_{k=0}^N$ generated above typically contain much fewer pairs of matrices than $\{\mathcal{H}_k\}_{k=0}^N$ and are thus much easier to deal with. Intuitively, when $\|\boldsymbol{\epsilon}\|$ is small, $\{\mathcal{H}_k^{\boldsymbol{\epsilon}}\}_{k=0}^N$ can also be used in place of $\{\mathcal{H}_k\}_{k=0}^N$ to define a control strategy whose performance is close to the optimal one.

Definition 5. *For each positive integer $k = 1, \ldots, N$, let $V_k^{\boldsymbol{\epsilon}}$, $\mu_k^{\boldsymbol{\epsilon}}$ and $\nu_k^{\boldsymbol{\epsilon}}$ be defined by equation (6) and (7) with $\mathcal{H}_N$ replaced by $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ and let $\xi_k^{\boldsymbol{\epsilon}} = (\mu_k^{\boldsymbol{\epsilon}}, \nu_k^{\boldsymbol{\epsilon}})$. Define $\pi_N^{\boldsymbol{\epsilon}} = \{\xi_N^{\boldsymbol{\epsilon}}, \ldots, \xi_1^{\epsilon}\}$. The function $\xi_k^{\boldsymbol{\epsilon}}$ is called the hybrid-feedback law generated by the characteristic set $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ and the policy $\pi_N^{\boldsymbol{\epsilon}}$ is called the feedback policy generated by the sets $\{\mathcal{H}_k^{\boldsymbol{\epsilon}}\}_{k=0}^N$.*

In the next subsection, we will show that the relaxed feedback policy $\pi_N^{\boldsymbol{\epsilon}}$ is suboptimal in the sense that by choosing $\|\boldsymbol{\epsilon}\|$ small enough, the cost $J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi)$ can be made arbitrarily close to the optimal one $V_N(\phi)$ for each $\phi \in \mathcal{A}$. Before doing this, we use the following example to demonstrate the simplicity of iteration (13).

Example 1.

$$\begin{aligned} &A_1 = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}, \\ &B_1 = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}, B_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, B_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \\ &F_1 = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}, F_2 = \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix}, F_3 = \begin{bmatrix} 0.3 \\ 0 \end{bmatrix}, \\ &Q_f = 0, Q_i = I_2,\ R_i = 1,\ i = 1, 2, 3, N = 30. \end{aligned} \tag{14}$$

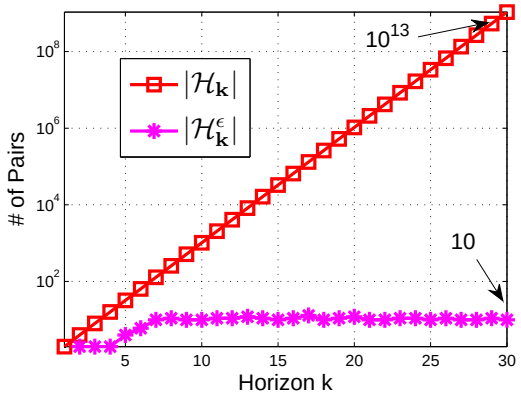

Figure 2: Complexity growth of $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ for Example 1 with $\boldsymbol{\epsilon} = (10^{-2}, 10^{-2})$.

As shown in Fig. 2, a direct computation of $\mathcal{H}_k$ will result in a combinatorial complexity at the order of 10^{13}. However, using the relaxed iteration (13) with relaxation vector $\boldsymbol{\epsilon} = (10^{-2}, 10^{-2})$, the number of pairs in $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ grows much more slowly and saturates at 10 after only a few iterations, which testifies to the simplicity for computing $\pi_N^{\boldsymbol{\epsilon}}$.

4.2. Performance analysis

The goal of this subsection is to derive an upper bound for the cost $J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi)$ and to assure that this upper bound can be made arbitrarily close to the optimal cost $V_N(\phi)$ by choosing $\boldsymbol{\epsilon}$ sufficiently small. To achieve this goal, it is beneficial to decompose each iteration of (13) into two steps, where in the first step, the set $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ is mapped to a larger set $\tilde{\mathcal{H}}_{k+1}^{\boldsymbol{\epsilon}} \triangleq h(\mathcal{H}_k^{\boldsymbol{\epsilon}})$, while in the second step, $Algo_{\boldsymbol{\epsilon}}$ is applied to remove the $\boldsymbol{\epsilon}$-redundant pairs in $\tilde{\mathcal{H}}_k^{\boldsymbol{\epsilon}}$ which yields the set $\mathcal{H}_{k+1}^{\boldsymbol{\epsilon}}$. By

Definition 5, $V_k^{\boldsymbol{\epsilon}}$ and $V_{k+1}^{\boldsymbol{\epsilon}}$ are the functions defined based on the sets $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ and $\mathcal{H}_{k+1}^{\boldsymbol{\epsilon}}$, respectively. We shall also introduce a function defined based on the set $\tilde{\mathcal{H}}_{k+1}^{\boldsymbol{\epsilon}}$ which is the key for connecting the two steps mentioned above.

For each $k = 1, \ldots, N$, define

$$\tilde{V}_k^{\boldsymbol{\epsilon}}(\phi) = \min_{(P,\Gamma)\in\tilde{\mathcal{H}}_k^{\epsilon}} t\, r(P\phi + \Gamma),\ \forall \phi \in \mathcal{A}. \tag{15}$$

Denote by $\{\Phi_{\pi_N^*}(t;\phi)\}_{t=0}^N$ and $\{\Phi_{\pi_N^{\boldsymbol{\epsilon}}}(t;\phi)\}_{t=0}^N$ the covariance matrices of system (9) driven by π_N^* and $\pi_N^{\boldsymbol{\epsilon}}$, respectively, with initial covariance ϕ, where π_N^* and $\pi_N^{\boldsymbol{\epsilon}}$ are the N-horizon polices defined in Corollary 1 and Definition 5, respectively. Following the same argument as in the proof of Theorem 1, it can be easily verified that for each $k = 1, \ldots N$,

$$\tilde{V}_k^{\boldsymbol{\epsilon}}(\phi) = L(x, \mu_k^{\boldsymbol{\epsilon}}(x,\phi), \nu_k^{\boldsymbol{\epsilon}}(\phi)) + V_{k-1}^{\boldsymbol{\epsilon}}(\Phi_{\pi_k^{\boldsymbol{\epsilon}}}(1;\phi)), \tag{16}$$

where $\xi_k^{\boldsymbol{\epsilon}} = (\mu_k^{\boldsymbol{\epsilon}}, \nu_k^{\boldsymbol{\epsilon}})$ is defined in Definition 5, $\pi_k^{\boldsymbol{\epsilon}} = \{\xi_k^{\boldsymbol{\epsilon}}, \ldots, \xi_1^{\boldsymbol{\epsilon}}\}$ and $x \sim \mathcal{N}(0,\phi)$. Since $H_k^{\boldsymbol{\epsilon}} = Algo_{\boldsymbol{\epsilon}}(\tilde{\mathcal{H}}_k^{\boldsymbol{\epsilon}})$, by Definition 4 we have

$$\tilde{V}_k^{\boldsymbol{\epsilon}}(\phi) \le V_k^{\boldsymbol{\epsilon}}(\phi) \le \tilde{V}_k^{\boldsymbol{\epsilon}}(\phi) + g^{\boldsymbol{\epsilon}}(\phi). \tag{17}$$

The reason for introducing $\{\tilde{V}_k^{\boldsymbol{\epsilon}}\}_{k=0}^N$ is that the cost $J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi)$ is upper bounded by $\tilde{V}_N^{\boldsymbol{\epsilon}}(\phi)$.

Lemma 4. *$J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi) \le \tilde{V}_N^{\boldsymbol{\epsilon}}(\phi)$ for any $\phi \in \mathcal{A}$.*

PROOF. For simplicity, for each $t \in T_N$, let $\hat{\Phi}(t) = \Phi_{\pi_N^{\boldsymbol{\epsilon}}}(t;\phi)$, $\hat{x}(t) \sim \mathcal{N}(0, \hat{\Phi}(t))$, $\hat{u}(t) = \mu_{N-t}^{\boldsymbol{\epsilon}}(\hat{x}(t), \hat{\Phi}(t))$ and $\hat{v}(t) = \nu_{N-t}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t))$. According to (16), we have

$$\begin{aligned}
&J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi) = \sum_{t=0}^{N-1} L(\hat{x}(t), \hat{u}(t), \hat{v}(t)) + \mathbb{E}(\hat{x}(N)^T Q_f \hat{x}(N)) \\
=&\sum_{t=0}^{N-1} \big[\tilde{V}_{N-t}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t)) - V_{N-(t+1)}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t+1))\big] + t\, r(\hat{\Phi}(N) Q_f) \\
=&\tilde{V}_N^{\boldsymbol{\epsilon}}(\phi) + \sum_{t=1}^{N-1} \big[\tilde{V}_{N-t}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t)) - V_{N-t}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t))\big] - V_0^{\boldsymbol{\epsilon}}(\hat{\Phi}(N)) + t\, r\big(\hat{\Phi}(N) Q_f\big).
\end{aligned} \tag{18}$$

Recall that $\mathcal{H}_0 = \mathcal{H}_0^{\boldsymbol{\epsilon}} = \{(Q_f, 0)\}$. Thus, $V_0^{\boldsymbol{\epsilon}}(\hat{\Phi}(N)) = t\, r\big(\hat{\Phi}(N) Q_f\big)$. In addition, by (17), we know $\tilde{V}_{N-t}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t)) - V_{N-t}^{\boldsymbol{\epsilon}}(\hat{\Phi}(t)) \le 0$ for each $t = 1, \ldots, N-1$. The desired result then follows from (18). □

Now our goal is shifted to finding an upper bound for $\tilde{V}_N^{\boldsymbol{\epsilon}}(\phi)$.

Lemma 5. *For any $\phi \in \mathcal{A}$, let $\{\Phi_{\pi_N^*}(t;\phi)\}_{t=0}^N$ be the covariance matrices along the optimal trajectory starting with initial covariance ϕ. Then*

$$\tilde{V}_N^{\epsilon}(\phi) \le V_N^{\epsilon}(\phi) \le V_N(\phi) + \sum_{t=0}^{N-1} g^{\boldsymbol{\epsilon}}(\Phi_{\pi_N^*}(t;\phi)). \tag{19}$$

PROOF. See Appendix A.

By the above lemma, for any fixed N and ϕ, $\tilde{V}_N^{\boldsymbol{\epsilon}}(\phi)$ can be made arbitrarily close to the optimal cost $V_N(\phi)$ by choosing $\|\boldsymbol{\epsilon}\|$ sufficiently small, which together with Lemma 4 implies the suboptimality of $\pi_N^{\boldsymbol{\epsilon}}$.

Theorem 2. *The N-horizon policy $\pi_N^{\boldsymbol{\epsilon}}$ is suboptimal with performance $J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi) \leq V_N + \sum_{t=0}^{N-1} g^{\boldsymbol{\epsilon}}(\Phi_{\pi_N^*}(t;\phi))$.*

4.3. Further Discussion for Large Horizon N

Since a nontrivial disturbance enters the system at each time step, the optimal cumulative cost $V_N(\phi)$ will grow unbounded for any $\phi \in \mathcal{A}$ as N increases. In other words, every N-horizon feedback policy will eventually result in an infinite cumulative cost as N approaches infinity. A more proper cost function for the large or infinite horizon problem is the *average-per-stage cost function*. For any N-horizon policy $\pi_N \in \Pi_N$, the average-per-stage cost function $\bar{J}_{\pi_N}$ is defined as

$$\bar{J}_{\pi_N}(\phi) = \tfrac{1}{N} J_{\pi_N}(\phi),\ \forall \phi \in \mathcal{A}.$$

Accordingly, we can define $\bar{V}_N(\phi) = \min_{\pi_N \in \Pi_N} \bar{J}_{\pi_N}(\phi)$, $\forall \phi \in \mathcal{A}$. Clearly, we have $\bar{V}_N(\phi) = \frac{1}{N} V_N(\phi)$ for any $\phi \in \mathcal{A}$. For any finite N, the cost functions J_{π_N} and $\bar{J}_{\pi_N}$ will result in the same optimal strategy. However, as N increases, the average-per-stage cost $\bar{J}_{\pi_N}$ provides a more consistent criterion for comparing different policies.

It follows directly from Theorem 2 that for any integer $N \geq 1$ and any performance tolerance $\delta > 0$, there always exists a constant $\hat{\epsilon}_N > 0$ such that $\bar{J}_{\pi_N^{\boldsymbol{\epsilon}}}(\phi) - \bar{V}_N(\phi) < \delta$ for all $\|\boldsymbol{\epsilon}\| \leq \hat{\epsilon}_N$. The limitation of such a result is that $\hat{\epsilon}_N$ may decay to zero rapidly as N increases, which in turn indicates a fast growth of the complexity to retain the same tolerance δ. The goal of this subsection is to further show that under some mild conditions, there exists a constant $\hat{\epsilon}$, independent of N, that can guarantee the δ suboptimality of $\pi_N^{\boldsymbol{\epsilon}}$ under the average-per-stage cost criterion. To achieve this goal, the following stabilizability assumption is introduced.

(A1) $\exists i \in \mathbb{M}$, such that(A_i, B_i) is stabilizable.

Lemma 6. *Under (A1), there exist finite positive constants β_1^-, β_1^+ and β_0^- and β_0^+ such that*

$$\beta_1^- t\, r(\phi) + \beta_0^- N \leq V_N(\phi) \leq \beta_1^+ t\, r(\phi) + \beta_0^+ N, \quad \forall N \in \mathbb{Z}_+, \forall \phi \in \mathcal{A}. \tag{20}$$

PROOF. See Appendix B. □

Notice that

$$V_N(\phi) \geq \sum_{t=0}^{N-1} t\, r(\Phi_{\pi_N^*}(t;\phi) \cdot Q_{\hat{v}(t)}) \geq \lambda_Q^- \sum_{t=0}^{N-1} t\, r(\Phi_{\pi_N^*}(t;\phi)), \tag{21}$$

where $\hat{v}(t)$ denotes the optimal closed-loop switching sequence corresponding to the initial covariance ϕ, and the first inequality follows by ignoring the control cost. Since $\lambda_Q^- > 0$, it follows from (20) and (21) that there exist positive constants $\alpha_1 < \infty$ and $\alpha_2 < \infty$ such that

$$\sum_{t=0}^{N-1} g^{\boldsymbol{\epsilon}}(\Phi_{\pi_N^*}(t;\phi)) \leq \|\boldsymbol{\epsilon}\|(\alpha_1 t\, r(\phi)+\alpha_2 N), \quad \forall \phi \in \mathcal{A}, \boldsymbol{\epsilon} \in \mathbb{R}_+^2 \text{ and } N \in \mathbb{Z}_+.$$

This inequality together with Theorem 2 implies a uniform performance upper bound in terms of the average-per-stage cost.

Theorem 3. *Under (A1), there exist positive constants $\alpha_1 < \infty$ and $\alpha_2 < \infty$ such that*

$$\bar{J}_{\pi_N^{\boldsymbol{\epsilon}}}(\phi) \leq \bar{V}_N(\phi) + \|\boldsymbol{\epsilon}\| \left(\frac{\alpha_1}{N} t\, r(\phi) + \alpha_2\right), \quad \forall \phi \in \mathcal{A} \text{ and } \forall N \in \mathbb{Z}_+. \tag{22}$$

Remark 4. *Theorem 3 implies that for smalle $\|\boldsymbol{\epsilon}\|$, the policy $\pi_N^{\boldsymbol{\epsilon}}$ performs uniformly well for all horizon N. Furthermore, as N increases, the effect of the initial covariance ϕ on the error bound also gets attenuated. Therefore, for large N and small $\|\boldsymbol{\epsilon}\|$, the the performance of $\pi_N^{\boldsymbol{\epsilon}}$ will be very close to the optimal one regardless of the initial covariance of the state.*

Remark 5. *It can be proved that the result in Theorem 3 still holds when none of the subsystem is stabilizable but the whole switched linear stochastic system is exponentially stabilizable in absence of the random perturbations. The proof of this result is rather technical and hence omitted here. See Example 2 of Section 5 for a numerical illustration.*

5. Numerical Examples

5.1. Example 1 Revisited

Consider the example given by (14) with a larger horizon $N = 50$. To solve this problem, we introduce the relaxation vector $\boldsymbol{\epsilon} = (0.01, 0.01)$. Iteration (13) is then used to compute the relaxed characteristic sets $\{\mathcal{H}_k^{\boldsymbol{\epsilon}}\}_{k=0}^N$. This computation can be done rather efficiently as shown in Fig. 2. These sets define the feedback laws $\{\xi_k^{\boldsymbol{\epsilon}} = (\mu_k^{\boldsymbol{\epsilon}}, \nu_k^{\boldsymbol{\epsilon}})\}_{k=1}^N$ that constitute the suboptimal policy $\pi_N^{\boldsymbol{\epsilon}}$. Under this policy, the continuous feedback gain and the switching control applied at any time $t \in T_N$ can be easily computed based on the current state covariance matrix. Suppose that at time t the state covariance matrix is $\Phi(t)$, the switching control is v and the feedback gain is K, respectively, then the cost incurred at this time step is simply

$$\mathbb{E}\big[x(t)^T Q_v x(t) + (-Kx(t))^T R_v (-Kx(t))\big] = t\, r\big(\Phi(t)(Q_v + K R_v K^T)\big).$$

In addition, the covariance matrix at time $t+1$ will be $\Phi(t+1) = (A_v - B_v K)\Phi(t)(A_v - B_v K)^T + \Sigma_v$. Using these equations, the average per stage

cost associated with the policy $\pi_N^{\boldsymbol{\epsilon}}$ for the initial covariance $\phi = [4, 5; 5, 8]$ is computed. This cost is compared in Table 1 with the average-per-stage costs of three standard LQG controllers associated with the 3 subsystems. The result clearly indicates that the cost can be dramatically reduced by properly switching among different subsystems. In addition, 100 realizations of the noisy closed-loop state trajectory are shown in Fig. 3. It can be seen that the proposed policy can properly regulate the state to a small neighborhood around the origin despite the persistent disturbance of the system. Furthermore, to show how the relaxation parameter $\boldsymbol{\epsilon}$ affects our algorithm, the cost and the complexity of the policy $\pi_N^{\boldsymbol{\epsilon}}$ under 5 different values of $\boldsymbol{\epsilon}$ are listed in Table 2. We observe that increasing the relaxation parameter significantly reduces the complexity of our algorithm, but only slightly degrades the performance of the obtained policy.

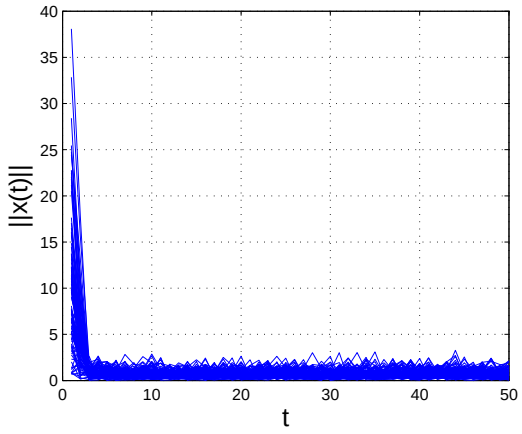

Figure 3: 100 realizations of the random state trajectory driven by $\pi_N^{\boldsymbol{\epsilon}}$ starting with initial covariance $\phi = [4, 5; 5, 8]$ for Example 1.

Table 1: Average-per-stage Cost of Different Controllers for Example 1

$\pi_N^{\boldsymbol{\epsilon}}$	LQG_1	LQG_2	LQG_3
0.2687	1.4224	1.6454	1.8114

Table 2: Performance of $\pi_N^{\boldsymbol{\epsilon}}$ under different $\boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}_1 = [1, 1]$ and $\phi = [4, 5; 5, 8]$

$\boldsymbol{\epsilon}$	$0.005\boldsymbol{\epsilon}_1$	$0.01\boldsymbol{\epsilon}_1$	$0.05\boldsymbol{\epsilon}_1$	$0.1\boldsymbol{\epsilon}_1$	$0.2\boldsymbol{\epsilon}_1$
$J_{\pi_N^{\boldsymbol{\epsilon}}}(\phi)$	0.2687	0.2687	0.2738	0.2928	0.2928
$\lvert\mathcal{H}_N^{\boldsymbol{\epsilon}}\rvert$	14	10	4	2	2

5.2. Example 2 with Unstabilizable Subsystems

We now consider a more challenging example with the following system matrices:

$$\begin{aligned}
&A_1 = \begin{bmatrix} 2 & 0 \\ 1 & 1.5 \end{bmatrix}, A_2 = \begin{bmatrix} 1.5 & 1 \\ 0 & 2 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \\
&B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, B_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\
&F_1 = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}, F_2 = \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix}, F_3 = \begin{bmatrix} 0.3 \\ 0 \end{bmatrix}, \\
&Q_f = 0, Q_i = I_2,\ R_i = 1,\ i = 1, 2, 3, N = 50.
\end{aligned} \tag{23}$$

Clearly, (A_i, B_i) is not stabilizable for any $i = 1, 2, 3$. Nevertheless, as pointed out in Remark 5, $\pi_N^{\boldsymbol{\epsilon}}$ is also suboptimal and will produce a finite average-per-stage cost for any horizon N as long as the overall switched system is exponentially stabilizable in absence of the random disturbances.

Following a similar procedure as described in the last example, the policy $\pi_N^{\boldsymbol{\epsilon}}$ with $\boldsymbol{\epsilon} = (0.01, 0.01)$ is computed. The number of pairs in $\mathcal{H}_k^{\boldsymbol{\epsilon}}$ saturates around 20 after 10 iterations. The average-per-stage cost of $\pi_N^{\boldsymbol{\epsilon}}$ for initial covariance $\phi = [4, 5; 5, 8]$ is computed to be $\bar{J}_{\pi_N^{\boldsymbol{\epsilon}}}(\phi) = 1.1121$, while the average-per-stage cost of the classical LQG controller for any subsystem is exponentially large at the order of 10^{10} for this particular horizon $N = 50$. Furthermore, as the horizon N increases, $\bar{J}_{\pi_N^{\boldsymbol{\epsilon}}}(\phi)$ remains almost the same, while the cost of the classical LQG controller for any subsystem keeps increasing exponentially fast. In Fig. 4, the random realizations of the state trajectory with initial distribution $\mathcal{N}(0, \phi)$ controlled by $\pi_N^{\boldsymbol{\epsilon}}$ are plotted. As observed in the figure, the closed-loop trajectory has a small overshoot at the beginning but then quickly decays to a small neighborhood around the origin. The switching sequence generated by the policy $\pi_N^{\boldsymbol{\epsilon}}$ for this particular initial covariance ϕ is $v = \{1, 2, 3, 2, 2, 3, 2, 2, 3, \ldots, 2, 2, 3, 2, 3\}$, which is almost periodic except at the beginning and the end of the control horizon. This switching sequence together with the properly designed feedback gains achieves the regulation performance as shown in Fig. 4, which is not possible to obtain using the classical LQG controller associated with any individual subsystem.

6. Conclusion

The finite-horizon quadratic optimal control problem for the discrete-time switched linear Gaussian system has been studied. Both the value function and the optimal control strategy have been characterized analytically. The value function is proved to be a piecewise affine function of the covariance matrix of the initial state and the optimal control law at each time step depends jointly on the random continuous state and its covariance matrix. A numerical relaxation framework is developed to simplify the computation of the suboptimal

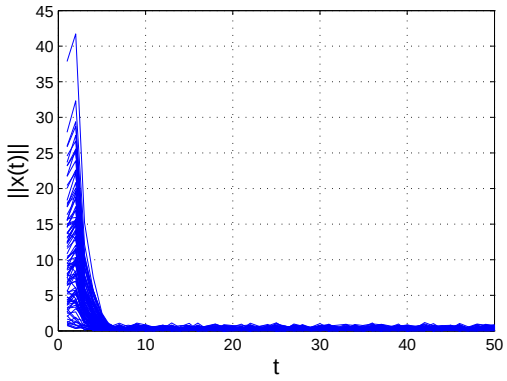

Figure 4: 100 realizations of the random state trajectory driven by $\pi_N^{\boldsymbol{\epsilon}}$ starting with initial covariance $\phi = [4, 5; 5, 8]$ for Example 2.

control strategy. It has also been proved that the performance of the suboptimal strategy can be made arbitrarily close to the optimal one by choosing the relaxation parameter small enough. The results of this paper can be used to solve many important problems of switched linear Gaussian systems, such as the stabilization problem, regulation problem and the supervisory control problem.

Appendix A. Proof of Lemma 5

When $N = 1$, since $\mathcal{H}_1^{\boldsymbol{\epsilon}} = Algo_{\boldsymbol{\epsilon}}(\mathcal{H}_1)$, the second inequality in (19) follows directly from the definitions of V_N^{ϵ} and the $\boldsymbol{\epsilon}$-ES. Suppose the result holds for $N = k$ for some $k \in \mathbb{Z}_+$. By (10) and (16), we have

$$\tilde{V}_{k+1}^{\boldsymbol{\epsilon}}(\phi) = \min_{\mu\in\mathcal{L},\nu\in\mathcal{V}} \Big\{ L(x(0), \mu(x(0), \phi), \nu(\phi)) + V_k^{\boldsymbol{\epsilon}}(\Phi_{\mu,\nu}(1; \phi)) \Big\}$$

$$\leq \min_{\mu\in\mathcal{L},\nu\in\mathcal{V}} \left\{ L(x(0), \mu(x(0), \phi), \nu(\phi)) + V_k(\Phi_{\mu,\nu}(1; \phi)) + \sum_{t=0}^{k-1} g^{\boldsymbol{\epsilon}}(\Phi_{\pi_k^*}(t; \Phi_{\mu,\nu}(1; \phi))) \right\}$$

If $(\mu, \nu) = (\mu_{k+1}^*, \nu_{k+1}^*)$, then the first two terms in the bracket of the last line becomes exactly $V_{k+1}(\phi)$, and in this case, by the Bellman's principle of optimality, for each $t = 0, \ldots, k-1$,

$$g^{\boldsymbol{\epsilon}}(\Phi_{\pi_k^*}(t; \Phi_{\mu,\nu}(1; \phi))) = g^{\boldsymbol{\epsilon}}(\Phi_{\pi_{k+1}^*}(t+1; \phi)).$$

Since $(\mu_{k+1}^*, \nu_{k+1}^*)$ is just one choice for (μ, ν), we have

$$\tilde{V}_{k+1}^{\boldsymbol{\epsilon}}(\phi) \leq V_{k+1}(\phi) + \sum_{t=0}^{k-1} g^{\boldsymbol{\epsilon}}(\Phi_{\pi_{k+1}^*}(t+1; \phi)) \leq V_{k+1}(\phi) + \sum_{t=1}^{k} g^{\boldsymbol{\epsilon}}(\Phi_{\pi_{k+1}^*}(t; \phi)).$$

Then, the desired result follows from (17). □

Appendix B. Proof of Lemma 6

Define $J_N^*(\phi;\mathbf{v}) \triangleq \min_{\mathbf{u}\in\mathcal{U}^N} J_N(\phi;\mathbf{u},\mathbf{v})$. Suppose that (A_i, B_i) is stabilizable. Let $\boldsymbol{v}(t) = i$, for all $t \in T_N$. Let $P_N^{\mathbf{v}}$ be the matrix generated by the Riccati mapping of subsystem i, i.e., $P_{k+1}^{\mathbf{v}} = \rho_i(P_k^{\mathbf{v}})$ for $k \in T_N$. By the standard LQR theory [5], there exists a p.s.d. matrix P^* such that $\|P_N^{\boldsymbol{v}}\| \to \|P^*\|$ as $N \to \infty$. Therefore, there must exist a finite constant β^+ such that $P_N^{\boldsymbol{v}} \prec \beta^+ I_n$, for any $N \in \mathbb{Z}_+$. Thus, by the definitions of $V_N(\phi)$ and $J_N^*(\phi;\mathbf{v})$, we have

$$V_N(\phi) \le J_N^*(\phi;\boldsymbol{v}) \le \beta^+ t\, r(\phi) + N\beta^+ \max_{i\in\mathbb{M}} t\, r(\Sigma_i),$$

which implies the second inequality in (20). Furthermore, since $Q_i \succ 0$ for each $i \in \mathbb{M}$, there must exist a constant $\beta^- > 0$ such that $Q_i \succ \beta^- I_n$ for all $i \in \mathbb{M}$. Then, it follows from (4) that $P_N^{\boldsymbol{v}} \succ \beta^- I_n$ for any $N \ge 1$ and $\boldsymbol{v} \in \mathbb{M}^N$. This together with Lemma 2 yields the first inequality in (20).

References

[1] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, 2nd edition, 2001.

[2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004. ISBN 0521833787.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, Philadelphia, PA, Jun 1994.

[4] M. S. Branicky. Multiple Lyapunov functions and other analysis tools for switched and hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):475–482, 1998.

[5] P. E. Caines and D. Q. Mayne. On the discrete time matrix Riccati equation of optimal control. *Internation Journal of Control*, 12(5):785–794, 1970.

[6] O. L. V. Costa and E. F. Tuesta. Finite horizon quadratic optimal control and a separation principle for markovian jump linear systems. *IEEE Transactions on Automatic Control*, 48(10):1836–1842, Oct. 2003.

[7] R. DeCarlo, M. Branicky, S. Pettersson, and B. Lennartson. Perspectives and results on the stability and stabilizability of hybrid systems. *Proceedings of IEEE, Special Issue on Hybrid Systems*, 88(7):1069–1082, 2000.

[8] V. Dragan, T. Morozan, and A. Stoica. H_2 optimal control for linear stochastic systems. *Automatica*, 40(7):1103–1113, Jan. 2004.

[9] Y. Ji and H. J. Chizeck. Controllability, stabilizability, and continuous-time markovian jump linear quadratic control. *IEEE Transactions on Automatic Control*, 35(7):777–788, Jul. 1990.

[10] Z. Ji, L. Wang, and G. Xie. Quadratic stabilization of uncertain discrete-time switched systems via output feedback. *Circuits, systems, and signal processing*, 24(6):733–751, 2005.

[11] J. T. Lee, E. Lau, and Y. C. Ho. The witsenhausen counterexample: A hierarchical search approach for nonconvex optimization problems. *IEEE Transactions on Automatic Control*, 46(3):382–397, Mar. 2001.

[12] J.-W. Lee and P. P. Khargonekar. Optimal output regulation for discrete-time switched and markovian jump linear systems. *SIAM Journal on Control and Optimization*, 47(1):40–72, Jan. 2008.

[13] D. Liberzon. *Switching in Systems and Control*. Systems & Control: Foundations & Applications. Birkhauser, Boston, 2003.

[14] H. Lin and P. J. Antsaklis. Switching stabilizability for continuous-time uncertain switched linear systems. *IEEE Transactions on Automatic Control*, 52(4):633–646, Apr. 2007.

[15] B. Lincoln and B. Bernhardsson. LQR optimization of linear system switching. *IEEE Transactions on Automatic Control*, 47(10):1701– 1705, Oct. 2002.

[16] Y. Orlov. Finite time stability and robust control synthesis of uncertain switched systems. *SIAM Journal on Control and Optimization*, 43(4):1253–1271, 2004.

[17] A. V. Savkin and R. J. Evans. *Hybrid Dynamical Systems: Controller and Sensor Switching Problems*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2001.

[18] D. Xie, L. Wang, F. Hao, and G. Xie. LMI approach to L_2-gain analysis and control synthesis of uncertain switched systems. *IEE Proceedings on Control Theory and Application*, 151(1):21–28, Jan. 2004.

[19] W. Zhang and J. Hu. On Optimal Quadratic Regulation for Discrete-Time Switched Linear Systems. In *International Workshop on Hybrid Systems: Computation and Control*, pages 584–597, St Louis,MO,USA, Apr. 2008.

[20] W. Zhang, A. Abate, and J. Hu. Efficient suboptimal solutions of switched LQR problems. In *Proceedings of the American Control Conference*, St. Louis, MO, Jun. 2009.