2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton) | 979-8-3503-2814-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/Allerton58177.2023.10313514

Stochastic Control under Correlated Disturbances

Faraz Farahvash
Electrical and Computer Engineering
Cornell University
ff227@cornell.edu

Ao Tang
Electrical and Computer Engineering
Cornell University
atang@cornell.edu

Abstract—**Stochastic control is a fundamental concept in control theory. Traditionally, it is assumed that the system disturbances are independent. This assumption does not always hold. In this paper, we look into cases where the disturbances are correlated. We provide a modified version of the Dynamic Programming algorithm to find optimal policies without the independence assumption. Then, we look into the classic LQG problem but with correlated disturbances. We show that the optimal policy is closely related to the trajectory-following LQG problem. The trajectory in this problem is defined by estimating future disturbances and is time-varying. We provide numerical examples to illustrate the theoretical results.**

I. Introduction

The stochastic control problem is a fundamental control concept defined as $x_{t+1} = f_t(x_t, u_t, w_t)$, where x_t, u_t and w_t represent state, control, and disturbance, respectively. It is assumed that the disturbances are independent of each other [1], [2].

Although this assumption simplifies the discussion and development of optimal policies, it is not always accurate. In some physical systems, the disturbances may be correlated. For example, when the vehicle's dynamics are described in the time domain, the process noise should be time-correlated. That's because the surface that the vehicle is moving on is smoothly changing. Thus, the disturbances depending on the surface are correlated [3].

There have been a variety of relevant works where the independence assumption does not hold. The simplest model is where we assume that the noise is dependent on the previous noise and an independent disturbance through a deterministic function [1]. Another method for solving stochastic control problems where noises are not independent is H_∞ method where the cost performance is just dependent on the noise energy [4], [5], [6]. There are also works where future disturbances can be estimated [7], [8], [9].

In this paper, we will analyze the stochastic control problem where disturbances are not independent. First, we start by deriving a Dynamic Programming algorithm for correlated disturbances. Then, we will tackle the Linear Quadratic Gaussian (LQG) control problem. We show that optimal control in this case is closely related to the trajectory-following LQG control. The trajectory in this problem is defined using our estimate of future disturbances and changes over time (contrary to the conventional trajectory-following problem). This work differs from [7]- [9] because disturbances are estimated using the previous disturbances and the covariance matrix.

The rest of this paper is organized as follows: In section II, we briefly summarize the classical stochastic control concepts; Section III presents the main theoretical results including correlated disturbance DP and LQG. In section IV, two methods for finding suboptimal policies are introduced. Section V illustrates the theoretical results through two numerical examples and section VI concludes the paper.

II. Independent Disturbances

In this section, we provide a review of the stochastic control concepts with independent disturbances.

A. General Problem

We define the problem as below [1], [2]:

$$x_{t+1} = f_t(x_t, u_t, w_t)$$
$$y_t = h_t(x_t, v_t)$$

Where x_t is the state (x_0 is the first state and is a random variable), u_t is the input, and w_t is input disturbance or error at time t. y_t is the observation or output and v_k is the unknown measurement error or noise. Also, we assume a finite horizon T for the problem.

Remark. *We assume that the random variables $x_0, w_0, \ldots, w_{T-1}, v_0, \ldots, v_T$ are all independent.*

A feasible policy is defined as $\pi = \{\pi_0, \pi_1, \ldots, \pi_{T-1}\}$. The set of all possible policies is called Π. We define the expected cost associated with policy π as follows:

$$J(\pi) := E[\sum_{t=0}^{T-1} c_t(x_t^\pi, u_t^\pi) + c_T(x_T^\pi)]$$

This work was supported by National Science Foundation (NSF). Grant number: 2133481.

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.

where c_t is called the immediate cost and c_T is called the terminal cost. We are looking for the minimum expected cost and the associated optimal policy for it:

$$J(\pi^*) = J^* = Inf\{J(\pi)|\pi \in \Pi\}$$

B. Dynamic Programming

In the case of complete information ($y_t = x_t$), the optimal policy is Markovian which is defined below.

Definition. *A policy π is called Markovian if π_t only depends on x_t. We call the set of Markovian policies Π_M.*

Dynamic Programming is a method used to find the optimal policy shown in algorithm 1.

Algorithm 1 Dynamic Programming (DP)

1: Define $V_T^*(x_T) := c_T(x_T)$
2: t=T-1
3: **while** $t \geq 0$ **do**
4: Let:

$$f(x_t, u_t) := c_t(x_t, u_t) + \mathbb{E}_{w_t}[V_{t+1}(x_{t+1})|x_t] \quad (1)$$

5: Then:

$$\pi_t^*(x_t) = \arg \inf_{u \in U} f(x_t, u_t)$$

6: And:

$$V_t^*(x_t) = \inf_{u \in U} f(x_t, u_t)$$

7: $t = t - 1$.
8: **end while**

C. Linear Quadratic Gaussian Problem (LQG)

Consider the complete information system below:

$$x_{t+1} = Ax_t + Bu_t + w_t \quad (2)$$

Where $w_t \sim N(0, \Sigma_w)$ are independent Gaussian random variables. The cost function is:

$$J(\pi) = \mathbb{E}[x_T' Q x_T + \sum_{t=0}^{T-1} (x_t' Q x_t + u_t' R u_t)] \quad (3)$$

Where Q is positive semi-definite and R is positive definite. With some assumptions on observability and controllability, we would have the following theorem.

Theorem 1. *For the system dynamics 2 and the cost function 3, The optimal policy will be*

$$\pi_t(x) = L_t x \quad (4)$$

Where:

$$L_t = -(B' K_{t+1} B + R)^{-1} B' K_{t+1} A \quad (5)$$

and

$$\begin{aligned} K_t &= Q + A'(K_{t+1} - K_{t+1} B (B' K_{t+1} B + R)^{-1} B' K_{t+1}) A \\ K_T &= Q \end{aligned} \quad (6)$$

Furthermore, the optimal cost will be:

$$J^* = \mathbb{E}[x_0' K_0 x_0] + \sum_{t=1}^{T} Tr(K_t \Sigma_w) \quad (7)$$

D. Trajectory following LQG

In the traditional LQG problem, the cost function is designed to keep both actions and states close to zero. But, we can modify the cost function to capture a trajectory we want to follow [10]. Let $\bar{x} = (\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_T)$ be the trajectory we want to follow. Also define $\tilde{x}_t := x_t - \bar{x}_t$. The cost function will be:

$$J(\pi) = \mathbb{E}[\tilde{x}_T' Q \tilde{x}_T + \sum_{t=0}^{T-1} (\tilde{x}_t' Q \tilde{x}_T + u_t' R u_t)] \quad (8)$$

This cost function is designed to reflect the desire to keep the system state close to the given trajectory. The optimal policy here is similar to the regular LQG (linear policy) but it includes an intercept. The next theorem summarizes the results.

Theorem 2. *For the system dynamics 2 and the cost function 8, The optimal policy will be*

$$\pi_t(x) = L_t x + M_t, \quad (9)$$

where L_t and K_t are defined the same as equations 5 and 6, respectively. The intercept (M_t) is:

$$M_t := (R + B' K_{t+1} B)^{-1} B' C_{t+1}, \quad (10)$$

where

$$\begin{aligned} C_t &= (A + BL_t)' C_{t+1} + Q \bar{x}_t, \\ C_T &= Q \bar{x}_T, \end{aligned} \quad (11)$$

Furthermore, the optimal cost will be:

$$J^* = \mathbb{E}[x_0' K_0 x_0] - 2C_0' \mathbb{E}[x_0] + D_0 + \sum_{t=1}^{T} Tr(K_t \Sigma_w) \quad (12)$$

Where D_t is defined as:

$$\begin{aligned} D_t &= D_{t+1} + \bar{x}_t' Q \bar{x}_t - C_{t+1}' B (R + B' K_{t+1} B)^{-1} B C_{t+1}, \\ D_T &= \bar{x}_T' Q \bar{x}_T. \end{aligned} \quad (13)$$

III. Correlated Disturbances

We are interested in finding optimal policies for controlling stochastic systems with correlated disturbances.

A. Dynamic Programming with Correlated Disturbances

We update the conventional DP algorithm to consider disturbance correlation. Define $x^{t,\pi}$ as all the states up to time t under policy π. We similarly define $u^{t,\pi}$ and w^t. We will assume that the disturbances and states are observed completely. We provide the following lemmas.

Lemma 1. *For any set $A \subseteq R^n$:*

$$Pr(x_{t+1}^\pi \in A | x^{t,\pi}, u^{t,\pi}, w^{t-1}) = Pr(x_{t+1}^\pi \in A | x_t^\pi, u_t^\pi, w^{t-1})$$

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.

Proof. Plugging in the formula for x_{t+1}:

$$\begin{aligned}
&Pr(x^{\pi}_{t+1} \in A | x^{t,\pi}, u^{t,\pi}, w^{t-1}) \\
&= Pr(f_t(x^{\pi}_t, u^{\pi}_t, w_t) \in A | x^{t,\pi}, u^{t,\pi}, w^{t-1}) \\
&= Pr(f_t(x^{\pi}_t, u^{\pi}_t, w_t) \in A | x^{\pi}_t, u^{\pi}_t, x^{t-1,\pi}, u^{t-1,\pi}, w^{t-1}) \\
&= Pr(f_t(x^{\pi}_t, u^{\pi}_t, w_t) \in A | x^{\pi}_t, u^{\pi}_t, u^{t-1,\pi}, w^{t-1}) \\
&= Pr(f_t(x^{\pi}_t, u^{\pi}_t, w_t) \in A | x^{\pi}_t, u^{\pi}_t, w^{t-1})
\end{aligned}$$

Where the third equality comes from the fact that $x^{t-1,\pi}$ is a function of $u^{t-1,\pi}, w^{t-1}$. And the final equality comes from the fact that w_t conditioned on previous disturbances is independent of the previous inputs. □

Definition. *A policy π is called Pseudo-Markovian if π_t only depends on x_t, and w^{t-1}. We call the set of Pseudo-Markovian policies Π_{PM}. Also, note that $\Pi_M \subseteq \Pi_{PM}$.*

Definition. *The updated cost-to-go function is defined as:*

$$J^{\pi}_t(x^{t,\pi}, w^{t-1}) = \mathbb{E}[c_T(x^{\pi}_T) + \sum_{k=t}^{T-1} c_k(x^{\pi}_k, u^{\pi}_k) | x^{t,\pi}, w^{t-1}]$$

Lemma 2. *Let $\pi \in \Pi_{PM}$ and Define these functions recursively:*

$$\begin{aligned}
&V^{\pi}_T(x_T, w^{T-1}) := c_T(x_T) \\
&V^{\pi}_t(x_t, w^{t-1}) := c_t(x_t, u^{\pi}_t) + \mathbb{E}_{w_t}[V^{\pi}_{t+1}(x^{\pi}_{t+1}) | x_t, w^{t-1}]
\end{aligned}$$

Then $V^{\pi}_t(x^{\pi}_t, w^{t-1}) = J^{\pi}_t(x^{t,\pi}, w^{t-1})$.

Proof. We prove this using backward induction. The base step ($t = T$) is true because:

$$V^{\pi}_T(x^{\pi}_T, w^{T-1}) = c_T(x^{\pi}_T) = J^{\pi}_T(x^{T,\pi}, w^{T-1})$$

For the induction step, assume that $V^{\pi}_{t+1}(x^{\pi}_{t+1}) = J^{\pi}_{t+1}$. We have:

$$\begin{aligned}
&V^{\pi}_t(x^{\pi}_t, w^{t-1}) \\
&= c_t(x_t, u^{\pi}_t) + \mathbb{E}_{w_t}[V^{\pi}_{t+1}(x^{\pi}_{t+1}, w^t) | x^{\pi}_t, w^{t-1}] \\
&= \mathbb{E}_{w_t}[c_t(x_t, u^{\pi}_t) + V^{\pi}_{t+1}(x^{\pi}_{t+1}, w^t) | x^{\pi}_t, w^{t-1}] \\
&= \mathbb{E}_{w_t}[c_t(x_t, u^{\pi}_t) + J^{\pi}_{t+1}(x^{t+1,\pi}, w^t) | x^{\pi}_t, w^{t-1}] \\
&= \mathbb{E}_{w_t}[\mathbb{E}[c_T(x^{\pi}_T) + \sum_{k=t}^{T-1} c_k(x^{\pi}_k, u^{\pi}_k) | x^{t+1,\pi}, w^t] | x^{\pi}_t, w^{t-1}] \\
&= \mathbb{E}_{w_t}[\mathbb{E}[c_T(x^{\pi}_T) + \sum_{k=t}^{T-1} c_k(x^{\pi}_k, u^{\pi}_k) | x^{\pi}_{t+1}, x^{\pi}_t, w^t] | x^{\pi}_t, w^{t-1}] \\
&= \mathbb{E}[c_T(x^{\pi}_T) + \sum_{k=t}^{T-1} c_k(x^{\pi}_k, u^{\pi}_k) | x^{\pi}_t, w^{t-1}] \\
&= \mathbb{E}[c_T(x^{\pi}_T) + \sum_{k=t}^{T-1} c_k(x^{\pi}_k, u^{\pi}_k) | x^{t,\pi}, w^{t-1}] \\
&= J^{\pi}_t(x^{t,\pi}, w^{t-1})
\end{aligned}$$

Where the third equality comes from the induction hypothesis. In the fifth and seventh equalities, the Pseudo-Markovian property is used. The sixth equality comes from the law of iterated expectation. □

Using the lemmas above and the Comparison Principle, we can show that the following algorithm provides an optimal policy. (The proof is analogous to the uncorrelated disturbance case)

Algorithm 2 Correlated Disturbance DP

1: Define $V^*_T(x_T, w^{T-1}) := c_T(x_T)$
2: t=T-1
3: **while** $t \geq 0$ **do**
4: Let:

$$f(x_t, u_t, w^{t-1}) := c_t(x_t, u_t) + \mathbb{E}_{w_t}[V_{t+1}(x_{t+1}) | x_t, w^{t-1}] \tag{14}$$

5: Then:

$$\pi^*_t(x_t, w^{t-1}) = \arg \inf_{u \in U} f(x_t, u_t, w^{t-1})$$

6: And:

$$V^*_t(x_t, w^{t-1}) = \inf_{u \in U} f(x_t, u_t, w^{t-1})$$

7: $t = t - 1$.
8: **end while**

Remark. *Comparing equations 1 and 14, we will get that the procedure for finding the optimal policy in correlated disturbances case is very similar to the independent case. The only difference is, in order to find the optimal policy, at each step, we should condition the expected value on previous disturbances as well.*

Remark. *We can extend this algorithm to the incomplete information case. In order to do that the estimator needs to both estimate x_t and all previous disturbances (i.e. w^{t-1}).*

Now, that we know the procedure for finding the optimal policy, we will tackle the LQG problem in the next section.

B. LQG with correlated disturbances (CorLQG)

In this section, we will go back to the LQG problem presented in II-C but with correlated disturbances. We will assume that disturbances are zero-mean and have a covariance matrix Σ. More specifically:

$$w = \begin{bmatrix} w_0 \\ w_1 \\ \cdot \\ \cdot \\ \cdot \\ w_{T-1} \end{bmatrix} \sim \mathcal{N}(0, \Sigma) \tag{15}$$

Before moving on to presenting the optimal control for this system, we provide the following famous lemma [1].

Lemma 3. *Let X, and Y be jointly Gaussian random variables ($X \sim \mathcal{N}(\bar{X}, \Sigma_X)$, $Y \sim \mathcal{N}(\bar{Y}, \Sigma_Y)$), and $cov(X, Y) = \Sigma_{XY}$,*

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.

then random variable Z_y which is defined as $Z \sim (X|Y = y)$ is also Gaussian, i.e. $Z_y \sim \mathcal{N}(\bar{Z}_y, \Sigma_{Z_y})$, where:

$$\bar{Z}_y = \bar{X} + \Sigma_{XY}\Sigma_Y^{-1}(y - \bar{Y}) \tag{16}$$

$$\Sigma_{Z_y} = \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma'_{XY} \tag{17}$$

Now, the lemma above provides two important insights:

1) Note that in the lemma 3, Σ_{Z_y} is independent of the value y. Thus, at time t, the covariance matrix of w_t given the previous disturbances (i.e. w^{t-1}) is independent from the values of w^{t-1}. Thus, we can define Σ_t as the covariance matrix of w_t given the previous disturbances.
2) Also, at time t, the expected value of the future disturbances given previous ones can be computed using equation 16. Define:

$$\bar{w}_{t_1,t_2}(w^{t_2-1}) = \mathbb{E}[w_{t_1}|w^{t_2-1}]. \tag{18}$$

Let $\bar{w}^{t_1,t_2}(w^{t_2-1}) = (\bar{w}_{t_1,t_2}(w^{t_2-1}), \dots, \bar{w}_{T-1,t_2}(w^{t_2-1}))$. Now, at each time, $-\bar{w}^{t_2,t_2}(w^{t_2-1})$ act as a trajectory we want to follow. Thus, we see a connection between this stochastic system and the trajectory following LQG presented in II-D. The main distinction is that the trajectory is changing over time. We will formalize the optimal policy in the following theorem.

Theorem 3. *For the system dynamics 2 and the cost function 3, and disturbances described in equation 15. The optimal policy will be:*

$$\pi_t(x, w^{t-1}) = L_t x + M_t(\bar{w}^{t,t}(w^{t-1})), \tag{19}$$

where L_t and K_t are defined the same as equation 5 and 6, respectively. Let:

$$F_t(w^{t-1}) = C_{t+1}(\bar{w}^{t+1,t}(w^{t-1})) + K_{t+1}\bar{w}_{t,t}(w^{t-1}),$$

where

$$\begin{aligned} &C_t(\bar{w}^{t,t}(w^{t-1})) = (A + BL_t)' F_t(w^{t-1}), \\ &C_T = 0, \end{aligned} \tag{20}$$

Then the intercept (M_t) will be:

$$M_t(\bar{w}^{t,t}(w^{t-1})) := -(R + B'K_{t+1}B)^{-1}B'F_t(w^{t-1}), \tag{21}$$

To summarize, theorem 3, argues that the optimal input should be a linear function where the slope is the same as the regular LQG, and the intercept is selected based on our estimate for future disturbances.

Remark. *Similar to the regular LQG, the certainty equivalence principle holds here (With a slight update where we replace $\mathbb{E}[w_k]$ with $\mathbb{E}[w_k|w^{t-1}]$ at time t).*

Before proving the theorem, we provide the following lemmas.

Lemma 4. *Let, $\bar{W}^{t_1}(w^{t_2-1})$ be defined as:*

$$\bar{W}^{t_1}(w^{t_2-1}) = \begin{bmatrix} \bar{w}_{t_1,t_2}(w^{t_2-1}) \\ \vdots \\ \bar{w}_{T-1,t_2}(w^{t_2-1}) \end{bmatrix}$$

Then, we can write $C_t(\bar{w}^{t,t}(w^{t-1}))$ defined at 20, as $Z_t\bar{W}^t(w^{t-1})$ where Z_t is the appropriate matrix.

Proof. We prove this using backward induction. The base case is trivial.
Now, assume $C_{t+1}(\bar{w}^{t+1,t+1}(w^t)) = Z_{t+1}\bar{W}^{t+1}(w^t)$. Then we have:

$$\begin{aligned} &C_t(\bar{w}^{t,t}(w^{t-1})) \\ &= (A + BL_t)' F_t(w^{t-1}) \\ &= (A + BL_t)'(Z_{t+1}\bar{W}^{t+1}(w^{t-1}) + K_{t+1}\bar{w}_{t,t}(w^{t-1})) \end{aligned}$$

Thus, we write:

$$C_t(\bar{w}^{t,t}(w^{t-1})) = (A + BL_t)' \left[K_{t+1}, Z_{t+1}\right] \bar{W}^t(w^{t-1}) \tag{22}$$

Thus, the induction is complete. □

Lemma 5.

$$\mathbb{E}[C_{t+1}(\bar{w}^{t+1,t+1}(w^t))|w^{t-1}] = C_{t+1}(\bar{w}^{t+1,t}(w^{t-1}))$$

Proof. Using lemma 4, we have:

$$\begin{aligned} \mathbb{E}[C_{t+1}(\bar{w}^{t+1,t+1}(w^t))|w^{t-1}] &= \mathbb{E}[Z_{t+1}\bar{W}^{t+1}(w^t)|w^{t-1}] \\ &= Z_{t+1}\mathbb{E}[\bar{W}^{t+1}(w^t)|w^{t-1}] \\ &= Z_{t+1}\bar{W}^{t+1}(w^{t-1}) \\ &= C_{t+1}(\bar{w}^{t+1,t}(w^{t-1})) \end{aligned}$$

Where the third equality comes from the law of iterated expectation. □

Now we are ready to prove theorem 3.

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.

proof of theorem 3. We make the following claim for the optimal value function:

$$V_t^*(x, w^{t-1}) = x' K_t x + 2C_t'(\bar{w}^{t,t}(w^{t-1}))x + D_t(w^{t-1}) \tag{23}$$

We prove this claim using backward induction. The base case is true. Now, assume the claim holds for t+1. Then, the value function at stage t is:

$$\begin{aligned}
&V_t^*(x, w^{t-1}) \\
&= \inf_u \left\{ x'Qx + u'Ru + \mathbb{E}_{w_t}[V_{t+1}^*(Ax + Bu + w_t)|x, w^{t-1}] \right\} \\
&= \inf_u \{x'Qx + u'Ru + \mathbb{E}_{w_t}[(Ax + Bu + w_t)' K_{t+1}(Ax + Bu + w_t) + D_{t+1}(w^t) + 2C_{t+1}'((\bar{w}^{t+1,t+1}(w^t))(Ax + Bu + w_t)|x, w^{t-1}]\} \\
&= \inf_u \{x'(Q + A'K_{t+1}A)x + u'(R + B'K_{t+1}A)u + 2x'A'K_{t+1}Bu \\
&+ \mathbb{E}_{w_t}[2(Ax + Bu)'[K_{t+1}w_t + C_{t+1}'((\bar{w}^{t+1,t+1}(w^t)] + D_{t+1}(w^t) + w_t'K_{t+1}w_t + 2C_{t+1}'((\bar{w}^{t+1,t+1}(w^t))w_t|w^{t-1}]\}
\end{aligned}$$

Now, we can use lemma 5,

$$\begin{aligned}
&= \inf_u \{x'(Q + A'K_{t+1}A)x + u'(R + B'K_{t+1}A)u + 2x'A'K_{t+1}Bu + 2(Ax + Bu)'[K_{t+1}\bar{w}_{t,t}(w^{t-1}) + C_{t+1}'((\bar{w}^{t+1,t}(w^t)] \\
&+ \mathbb{E}_{w_t}[D_{t+1}(w^t) + w_t'K_{t+1}w_t + 2C_{t+1}'((\bar{w}^{t+1,t+1}(w^t))w_t|w^{t-1}]\}
\end{aligned}$$

Now, the last line only depends on w^{t-1}. Call it $S_t(w^{t-1})$. We will have:

$$V_t^*(x, w^{t-1}) = S_t(w^{t-1}) + x'(Q + A'K_{t+1}A)x + 2F_t'(w^{t-1})Ax + \inf_u \{u'(R + B'K_{t+1}A)u + 2[F_t'(w^{t-1}) + x'A'K_{t+1}]Bu\}$$

From the strict convexity of the function to infimize, we can use the first-order optimality condition to deduce the optimal input:

$$u^* = -(R + B'K_{t+1}A)^{-1}B'(F_t(w^{t-1}) + K_{t+1}Ax)$$

Which is the same as equation 19. Finally, by plugging in the optimal control we will see that:

$$V_t^*(x, w^{t-1}) = x'K_t x + 2C_t'(\bar{w}^{t,t}(w^{t-1}))x + D_t(w^{t-1})$$

Where K_t and C_t, come from equations 6 and 20, respectively. Thus the induction is complete. □

IV. Approximate Methods

In this section, we will present two approximate methods to find suboptimal controllers for the control system with correlated disturbances.

A. K-step ahead trajectory

In the LQG problem, to compute C_t in theorem 3, we need to compute Z_t and $\bar{W}^t(w^{t-1})$ at step t (lemma 4). Assume, $x_k \in \mathbf{R}^n$ and. Note that Z_t has a dimension $n(T - t - 1) \times n$.
Thus, for large horizons computing the intercept becomes inefficient.

To ameliorate this problem, we will restrict ourselves to estimating the next K disturbances only. In other words, at time t, we will assume that w_l for $l > t + K$ conditioned on the previous disturbances are still zero-mean.

Remark. *In the special case where $K = 1$, we can define:*

$$\tilde{W}^t(w^{t-1}) = \left[\bar{w}_{t,t}(w^{t-1}), 0, \ldots, 0\right]'$$

Thus, replacing $\bar{W}^t$ with $\tilde{W}^t$ in equation 22, we will have:

$$\begin{aligned}
\tilde{C}_t(\bar{w}^{t,t}(w^{t-1})) &= (A + BL_t)' \left[K_{t+1}, Z_{t+1}\right] \tilde{W}^t(w^{t-1}) \\
&= (A + BL_t)' K_{t+1}\bar{w}_{t,t}(w^{t-1})
\end{aligned}$$

We can replace the C_t in theorem 3 with $\tilde{C}_t$ for this method.

Remark. *Notice that in order to compute $\bar{w}_{t,t}(w^{t-1})$, we need to compute the inverse of a $tn \times tn$ matrix. We can use the most recent L disturbances to estimate it sub-optimally.*

B. Model Predictive Methods

Another useful method for solving stochastic control problems where disturbances are not independent is the CE-MPC method which reduces the problem to a deterministic problem at each time. Implementing this controller in cases with correlated disturbances is worthwhile as by using the estimation that we have for future disturbances and solving the deterministic problem each time we can get a good suboptimal controller. Also, note that this method is optimal for the LQG problem described in III-B.

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.

V. Numerical Results

In this section, we illustrate our theoretical results through two numerical examples. In section V-A, we consider a simple control system where all the random variables are scalars. We will compute the gain by taking the correlation of disturbances into consideration. [9]. In section V-B, we consider a widely-used two-dimensional robot control system but with correlated disturbances [11], [12], [13].

A. Scalar LQG

To see the benefit of taking the correlation into consideration, we will explore a simple one-dimensional problem. We have $T = 10$, $x_0 \sim \mathcal{N}(0,1)$, and $A = B = Q = R = 1$. We also assume:

$$cov(w_t, w_k) = \begin{cases} 1, & t = k \\ 0.9, & t \neq k \end{cases}$$

We will consider two different policies; regular LQG (which doesn't take the correlation into consideration) and CorLQG with 1 step ahead trajectory. It can be shown that the cost of regular LQG is equal to 17.042 while the CorLQG achieves a cost of 5.4758. This shows a 67 percent decrease in cost.

Note that if we plug in the covariance matrix of this problem into equations of lemma 3, we will get that the variance of w_t conditioned on previous disturbances is less than or equal to 0.19. However, the cost of CorLQG is more than 19 percent of the regular LQG. This happens for two reasons. First, the initial random variables (x_0 and w_0) still have a variance equal to 1. Second, as the disturbances are positively correlated CorLQG anticipating the next disturbance, may take a bigger step. Thus, for example, if we change the covariance to -0.9 the cost of corLQG will drop to 3.6211.

B. Two-dimensional Robot Control

In this example, we will consider a 2-D robot movement control system where T=40 and:

$$A = \begin{bmatrix} 1 & 0 & 0.2 & 0 \\ 0 & 1 & 0 & 0.2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.2 & 0 \\ 0 & 0.2 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, R = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix},$$

We will also assume that the covariance matrix of disturbances is as follows:

$$\Sigma_{w_t,w_t} = \begin{bmatrix} 1 & 0.5 & 0.2 & 0.2 \\ 0.5 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.5 \\ 0.2 & 0.2 & 0.5 & 1 \end{bmatrix}, \Sigma_{w_t,w_k} = 0.8\Sigma_{w_t,w_t}, \forall k \neq t$$

We will implement the same policies as the previous part. We run the experiment 10000 times and we see that the LQG has a cost of 658.926 compared to a cost of 625.556 of CorLQG. Figure 1, provides a cost histogram for the 10000 experiments.

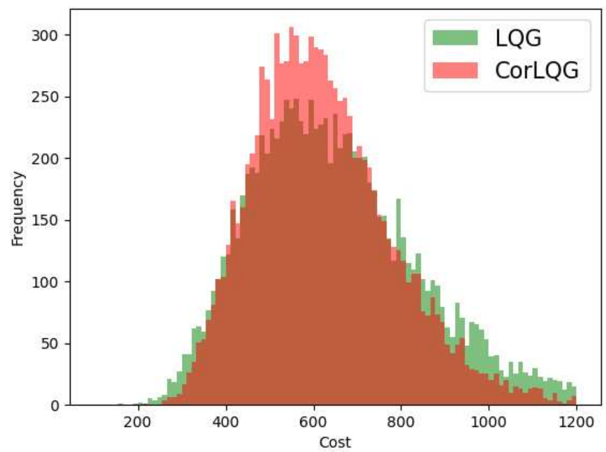

Fig. 1. Cost Histogram of two policies (regular LQG and CorLQG)

VI. Conclusion

In this paper, we looked at stochastic control in cases where the disturbances are correlated. First, we updated the Dynamic Programming algorithm to capture the correlation between disturbances. Then, we found the optimal policy of the LQG problem and showed that it has a connection to the trajectory-following LQG problem. The trajectory is determined by estimating future disturbances using previous disturbances. We called the optimal policy CorLQG. Finally, two examples are provided to further illustrate the theoretical results.

References

[1] P. R. Kumar and P. Varaiya, *Stochastic Systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2015.

[2] D. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. 01 1995.

[3] Y. Wang, R. Su, and B. Wang, "Optimal control of interconnected systems with time-correlated noises: Application to vehicle platoon," *Automatica*, vol. 137, p. 110018, 2022.

[4] K. Uchida and M. Fujita, "Finite horizon h/sup infinity / control problems with terminal penalties," *IEEE Transactions on Automatic Control*, vol. 37, no. 11, pp. 1762–1767, 1992.

[5] G. Goel and B. Hassibi, "Regret-optimal estimation and control," 2021.

[6] G. Goel and B. Hassibi, "Online estimation and control with optimal pathlength regret," in *Proceedings of The 4th Annual Learning for Dynamics and Control Conference* (R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, and M. Kochenderfer, eds.), vol. 168 of *Proceedings of Machine Learning Research*, pp. 404–414, PMLR, 23–24 Jun 2022.

[7] J. Cheng, M. Pavone, S. Katti, S. Chinchali, and A. Tang, "Data sharing and compression for cooperative networked control," 09 2021.

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.

[8] A. K. Singh and B. C. Pal, “An extended linear quadratic regulator for lti systems with exogenous inputs,” *Automatica*, vol. 76, pp. 10–16, 2017.
[9] J. Cheng and A. Tang, “Linear-quadratic-gaussian control with time-varying disturbance forecast,” in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 479–484, 2022.
[10] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Feher/Prentice Hall Digital and, Prentice Hall, 1996.
[11] T. Li, R. Yang, G. Qu, G. Shi, C. Yu, A. Wierman, and S. Low, “Robustness and consistency in linear quadratic control with predictions,” 06 2021.
[12] C. Yu, G. Shi, S.-J. Chung, Y. Yue, and A. Wierman, “Competitive control with delayed imperfect information,” 2022.
[13] Y. Li, X. Chen, and N. Li, “Online optimal control with linear dynamics and predictions: Algorithms and regret analysis,” 06 2019.

Authorized licensed use limited to: Cornell University Library. Downloaded on August 31,2024 at 06:02:51 UTC from IEEE Xplore. Restrictions apply.