Maximum Entropy Optimal Density Control of Discrete-Time Linear Systems and Schrödinger Bridges

Kaito Ito, Member, IEEE, and Kenji Kashima, Senior Member, IEEE

Abstract—We consider an entropy-regularized version of optimal density control of deterministic discrete-time linear systems. Entropy regularization, or a maximum entropy (MaxEnt) method for optimal control, has attracted much attention especially in reinforcement learning due to its many advantages, such as a natural exploration strategy. Despite the merits, high-entropy control policies induced by the regularization introduce probabilistic uncertainty into systems, which severely limits the applicability of MaxEnt optimal control to safety-critical systems. To remedy this situation, we impose a Gaussian density constraint at a specified time on the MaxEnt optimal control to directly control state uncertainty. Specifically, we derive the explicit form of the MaxEnt optimal density control. In addition, we also consider the case where density constraints are replaced by fixed-point constraints. Then, we characterize the associated state process as a pinned process, which is a generalization of the Brownian bridge to linear systems. Finally, we reveal that the MaxEnt optimal density control gives the so-called Schrödinger bridge associated with a discrete-time linear system.

Index Terms—Maximum entropy (MaxEnt), optimal control, Schrödinger bridge, stochastic control.

I. INTRODUCTION

In this article, we address the problem of steering the state of a deterministic discrete-time linear system to a desired distribution over a finite horizon while minimizing the energy cost with an entropy regularization term for control policies. Entropy-regularized optimal control is referred to as maximum entropy (MaxEnt) optimal control and has a close connection to Kullback–Leibler (KL) control [1], [2], [3]. There has been considerable interest in MaxEnt optimal control especially in reinforcement learning (RL) [4], [5], [6], [7]. This is because the entropy regularization offers many advantages, such as performing good exploration for RL [4], robustness against disturbances [8], and equivalence between the MaxEnt optimal control problem and an inference problem [6], to name a few.

Although the entropy regularization brings many benefits, the resulting high-entropy policy increases the uncertainty of the state of dynamics, which severely limits the applicability of MaxEnt optimal control to safety-critical systems. Therefore, it is important to tame state uncertainty due to the stochasticity of policies. A straightforward approach to achieve this is to impose a hard constraint in the state distribution at a specified time. Steering the state of a dynamical system to a desired distribution without (explicit) entropy regularization has been addressed in the literature. Especially when the state distribution has a density function, this kind of problem is called a density control problem.

Literature review: The most closely related work to this article is [9], which studies the problem of steering a Gaussian initial density of a continuous-time linear stochastic system to a final one with minimum energy. In particular, the optimal policy is derived in explicit form, and moreover, the optimal state process is shown to be a Schrödinger bridge associated with a linear system. The Schrödinger bridge problem, introduced by Schrödinger [10], is known to be equivalent to an entropy-regularized optimal transport problem, which has received renewed attention in data sciences [11]. A substantial literature is devoted to reveal the relationships between stochastic optimal control and Schrödinger bridges in a continuous-time setting [12], [13], [14], [15], [16], [17]. For a discrete time and space setting, the counterpart of the density control and Schrödinger bridge problems is explored in [18].

On the other hand, in [19], [20], and [21], the density control, or covariance steering problem is investigated for discrete-time linear systems. Bakolas [19], [20] employed a convex relaxation technique to solve a covariance steering problem with quadratic cost and a constraint on the state and the control input. Goldshtein and Tsiontsas [21] focused on covariance steering with only quadratic input cost and propose an efficient numerical scheme without relaxation for computing the solution. However, unlike the continuous-time case, the obtained optimal process is not a Schrödinger bridge associated with the discrete-time linear system. Indeed, in [22], the Schrödinger bridge for the discrete-time linear system driven by noise acting on all components of the state vector is derived. Nevertheless, the existence and an implementable form of the optimal control that replicates the
bridge are missing, and it is not possible to obtain the bridge by a linear feedback law.

Other scenarios for density control and covariance steering have been explored, including nonlinear systems, nonquadratic costs, non-Gaussian densities [23], [24], [25], [26], [27], chance constraints [28], and Wasserstein terminal cost instead of a final density constraint [29], [30]. Finally, we refer the reader to [31] for an extensive review of this area.

**Contributions:** The main contributions of this article are listed as follows.

1) We analyze the MaxEnt optimal density control problem for deterministic discrete-time linear systems whose initial density and target density are both Gaussian. Specifically, we show the existence and uniqueness of the optimal policy and then derive its explicit form via two Lyapunov difference equations coupled through their boundary values.

2) As a limiting case, we consider the above problem where density constraints are replaced by point constraints and show that the induced state process is a pinned process of a linear system driven by Gaussian noise. The pinned process is a generalization of the so-called Brownian bridge; see Section V for the definition. This result is the discrete-time counterpart of [32].

3) Based on the above two results, we reveal that the optimal state process of our density control problem is a Schrödinger bridge associated with the discrete-time linear system. This result does not require the invertibility of noise intensity supposed in [22].

It is worth mentioning that this study is the first one that solves a density control problem with entropy regularization, to the best of authors’ knowledge.

**Organization:** The rest of this article is organized as follows. In Section II, we briefly review the MaxEnt optimal control of linear systems without density constraints. In Section III, we provide the problem formulation and derive the coupled Lyapunov equations, which are crucial ingredients in our analysis. The existence, uniqueness, and explicit form of the optimal policy are given in Section IV. In Section V, we consider the MaxEnt point-to-point steering and characterize the associated state process as a pinned process. In Section VI, we reveal that the MaxEnt optimal density control induces the Schrödinger bridge associated with the linear system. In Section VII, we present illustrative examples. Finally, Section VIII concludes this article.

**Notations:** Let $\mathbb{R}$ denote the set of real numbers and $\mathbb{N}_{>0}$ denote the set of positive integers. The set of integers $\{k, k+1, \ldots, l\}$ is denoted by $[k,l]$. Denote by $S^n$ the set of all real symmetric $n \times n$ matrices. For matrices $A, B \in S^n$, we write $A \succ B$ (respectively, $A \prec B$) if $A - B$ is positive (respectively, negative) definite. Similarly, $A \succeq B$ means that $A - B$ is positive semidefinite. For $A > 0$ (respectively, $A \succeq 0$), $A^{1/2}$ denotes the unique positive definite (respectively, semidefinite) square root. The identity matrix is denoted by $I$, and its dimension depends on the context. The Moore–Penrose inverse of a matrix $A \in \mathbb{R}^{m \times n}$ is denoted by $A^\dagger \in \mathbb{R}^{n \times m}$. The transpose of the inverse of an invertible matrix $A$ is denoted by $A^{-\dagger}$.

**II. BRIEF INTRODUCTION TO MAXENT LINEAR QUADRATIC OPTIMAL CONTROL**

We start with a brief introduction to the MaxEnt optimal control of linear systems with quadratic costs. The problem formulation is given as follows.

**Problem 1:** Find a policy $\pi = \{\pi_k\}_{k=0}^{N-1}$ that solves

$$
\text{minimize } \mathbb{E}\left[ \frac{1}{2} (x_N - \bar{x}_N)^\top F(x_N - \bar{x}_N) + \sum_{k=0}^{N-1} \left( \frac{1}{2} \|u_k\|^2 - \varepsilon H(\pi_k(\cdot|x_k)) \right) \right] \\
\text{subject to } x_{k+1} = A_k x_k + B_k u_k \\
\hspace{2cm} u_k \sim \pi_k(\cdot|x_k) \text{ given } x_k = x
$$

where $\varepsilon > 0$, $x_k, \bar{x}_N \in \mathbb{R}^n, u_k \in \mathbb{R}^m, F \in S^n$, and $N \in \mathbb{Z}_{>0}$ is the terminal time. The initial state $x_0$ has a finite second moment, and $\bar{x}_N$ is a constant. A stochastic policy $\pi_k(\cdot|x)$ denotes the conditional density of $u_k$ given $x_k = x$, and $H(\pi_k(\cdot|x)) := -\int_{\mathbb{R}^m} \pi_k(u|x) \log \pi_k(u|x) du$ denotes the entropy of $\pi_k(\cdot|x)$. For larger values of the regularization parameter $\varepsilon$, the entropy of the optimal policy becomes higher.

Under the invertibility of $A_k$, define the state-transition matrix

$$
\Phi(k,l) := \begin{cases} 
A_{k-1} A_{k-2} \cdots A_l, & k > l \\
I, & k = l \\
A_{k-l}^{-1} A_{k-l-1}^{-1} \cdots A_l^{-1}, & k < l.
\end{cases}
$$

Similar to the conventional linear quadratic regulator (LQR) problem [33], the optimal policy of Problem 1 can be obtained in a closed form. The proof is shown in Appendix A.

**Proposition 1:** Assume that $\Pi_k \in S^n$ satisfies $I + B_k^\dagger \Pi_{k+1} B_k > 0$ for any $k \in [0, N - 1]$ and is a solution to the following Riccati difference equation:

$$
\Pi_k = A_k^\dagger \Pi_{k+1} A_k - A_k^\dagger \Pi_{k+1} B_k (I + B_k^\dagger \Pi_{k+1} B_k)^{-1} \times B_k^\dagger \Pi_{k+1} A_k, \ k \in [0, N - 1]
$$

$$
\Pi_N = F.
$$
Assume further that $A_k$ is invertible for any $k \in [0, N - 1]$. Then, the unique optimal policy of Problem 1 is given by

$$\pi_k^*(u|x) = N(u|-(I + B_k^T \Pi_{k+1} B_k)^{-1} B_k^T \Pi_{k+1} A_k (x - \Phi(k,N)\bar{x}_N), \epsilon(I + B_k^T \Pi_{k+1} B_k)^{-1})$$

$$k \in [0, N - 1], x \in \mathbb{R}^n, u \in \mathbb{R}^m. \quad (6)$$

In addition, if $F \succeq 0$, then the solution $\{\Pi_k\}$ to (4) and (5) satisfies $\Pi_k \succeq 0$ for any $k \in [0, N]$. Furthermore, when $\bar{x}_N = 0$, policy (6) with $\Phi(k,N)\bar{x}_N = 0$ is still optimal without the invertibility of $A_k$.

The mean of the optimal policy (6) coincides with the LQ optimal controller. In other words, the optimal policy is the LQ optimal controller perturbed by independent additive Gaussian noise with zero mean and covariance matrix $\epsilon(I + B_k^T \Pi_{k+1} B_k)^{-1}$, which explicitly shows the effect of the entropy regularization.

Remark 1: In Problem 1, without loss of generality, we may assume $\epsilon = 1$. To see this, let $u_k^\epsilon := u_k/\sqrt{\epsilon}$ and $\pi_k^\epsilon(\cdot|x)$ be the conditional density of $u_k^\epsilon$ given $x_k = x$. Then,

$$H(\pi_k^\epsilon(\cdot|x)) = H(\pi_k(\cdot|x)) - \frac{m}{2} \log \epsilon \quad \forall x \in \mathbb{R}^n.$$ 

See e.g., [34, Sec. 8.6]. The cost (1) is rewritten as

$$\mathbb{E}\left[\frac{1}{2\epsilon}(x_N - \bar{x}_N)^T F(x_N - \bar{x}_N) + \sum_{k=0}^{N-1} \frac{1}{2}\|u_k^\epsilon\|^2 - H(\pi_k^\epsilon(\cdot|x)) - \frac{m}{2} \log \epsilon\right].$$

Therefore, by replacing $B_k, F$ by $\sqrt{\epsilon}B_k, F/\epsilon$, respectively, Problem 1 reduces to the case with $\epsilon = 1$. \hfill \Box

### III. Problem Formulation and Preliminary Analysis

Now, we formulate the problem of steering the state of a linear system to a desired distribution. Specifically, in this article, we consider the following optimal control problem.

**Problem 2 (MaxEnt optimal density control problem):** Find a policy $\pi = \{\pi_k\}_{k=0}^{N-1}$ that solves

$$\begin{align*}
\min_{\pi} & \quad \mathbb{E}\left[\sum_{k=0}^{N-1} \frac{1}{2}\|u_k\|^2 - \epsilon H(\pi_k(\cdot|x_k))\right] \\
\text{subject to} & \quad x_{k+1} = A_k x_k + B_k u_k \\
& \quad u_k \sim \pi_k(\cdot|x) \text{ given } x_k = x \\
& \quad x_0 \sim N(0, \Sigma_0), x_N \sim N(0, \Sigma_N)
\end{align*}$$

$$k \in [0, N - 1], x \in \mathbb{R}^n, u \in \mathbb{R}^m. \quad (7)$$

where $\epsilon > 0$, $\Sigma_0, \Sigma_N > 0$, $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and $N \in \mathbb{Z}_{>0}$.

The general case (referred to as Problem 2') where the density constraints (9) are replaced by nonzero mean Gaussian distributions

$$x_0 \sim N(\mu_0, \Sigma_0), x_N \sim N(\mu_N, \Sigma_N)$$

(10) is also of interest. Later, we will see that this can be solved by combining the optimal covariance steering obtained as the solution to Problem 2 and optimal mean steering.

To tackle Problem 2, let us go back to Problem 1 with $\bar{x}_N = 0$ (referred to as Problem 1'). Problem 1' has a terminal cost instead of a constraint on the density of the final state. Assume that $\Pi_k \in \mathcal{S}^n$ satisfies $I + B_k^T \Pi_{k+1} B_k \succ 0$ for any $k \in [0, N - 1]$ and is a solution to the Riccati equation (4) with (5). By Proposition 1, the unique optimal policy of Problem 1' is given by

$$\pi_k^*(u|x) = N(u|-(I + B_k^T \Pi_{k+1} B_k)^{-1} B_k^T \Pi_{k+1} A_k x, \epsilon(I + B_k^T \Pi_{k+1} B_k)^{-1})$$

$$k \in [0, N - 1], x \in \mathbb{R}^n, u \in \mathbb{R}^m. \quad (11)$$

System (8) driven by policy (11) is given by

$$x_{k+1} = \bar{A}_k x_k + B_k w_k^\epsilon, \quad w_k^\epsilon \sim N(0, \epsilon(I + B_k^T \Pi_{k+1} B_k)^{-1})$$

(12)

where $\{w_k^\epsilon\}$ is an independent sequence and

$$\bar{A}_k := A_k - B_k(I + B_k^T \Pi_{k+1} B_k)^{-1} B_k^T \Pi_{k+1} A_k$$

By the linearity of (12) and the Gaussianity of $w_k^\epsilon$, if $x_0$ follows a Gaussian distribution, then for any $k \in [1, N]$, $x_k$ has a Gaussian distribution. Suppose that $\Sigma_k := \mathbb{E}[x_k^\epsilon(x_k^\epsilon)^T]$ satisfies $\Sigma_0 = \Sigma_0, \Sigma_N = \Sigma_N$. Then, policy (11) is the unique optimal solution of Problem 2. In fact, for any policy satisfying (9), the terminal cost $\mathbb{E}[x_N^\epsilon F x_N^\epsilon]/2$ for Problem 1' takes the same value. Thus, if policy (11) is not the unique optimal solution of Problem 2, it contradicts the optimality and uniqueness of (11) for Problem 1'. By (12), $\Sigma_k$ evolves as

$$\Sigma_{k+1} = \bar{A}_k \Sigma_k \bar{A}_k^T + \epsilon B_k(I + B_k^T \Pi_{k+1} B_k)^{-1} B_k^T.$$ 

(13)

Therefore, if $\Sigma_0 \succ 0$ and $A_k$ is invertible for any $k \in [0, N - 1]$, it holds $\Sigma_k \succ 0$ for any $k \in [1, N]$. Henceforth, we assume the invertibility of $A_k$.

Now, inspired by [9], we introduce $H_k := \epsilon \Sigma_k^{-1} - \Pi_k$. Assume that $\Pi_k$ and $H_k$ are invertible on the time interval $[0, N]$. Noting that

$$\Pi_k + H_k \sim (\Pi_k^{-1} - \Pi_k^{-1} \Pi_k^{-1} H_k^{-1})^{-1} \Pi_k^{-1}$$

we obtain

$$\Pi_k^{-1} + H_k^{-1} = \left(\Pi_k^{-1} - \frac{1}{\epsilon} \Pi_k \Sigma_k \Pi_k\right)^{-1}.$$ 

In addition, by using (4) and (13), and noting that $A_k^T \Pi_{k+1} \bar{A}_k = \Pi_k$, we get

$$A_k^T (\Pi_{k+1}^{-1} + H_k^{-1}) A_k^T = \left(A_k^T \left(\Pi_k^{-1} + H_k^{-1}\right) A_k\right)^{-1}$$

$$= \left(\Pi_k^{-1} - \frac{1}{\epsilon} \Pi_k \Sigma_k \Pi_k\right)^{-1} = \Pi_k^{-1} + H_k^{-1}.$$ 

Hence, it holds

$$\Pi_k^{-1} + H_k^{-1} = A_k (\Pi_k^{-1} + H_k^{-1}) A_k^T.$$ (14)
Moreover, the Riccati equation (4) is rewritten as
\[ \Pi_{k+1}^{-1} = A_k \Pi_k^{-1} A_k^\top - B_k B_k^\top. \] (15)

By (14) and (15), it holds
\[ H_{k+1}^{-1} = A_k H_k^{-1} A_k^\top + B_k B_k^\top. \] (16)

Therefore, \( P_k := H_k^{-1} \) and \( Q_k := \Pi_k^{-1} \) satisfy the Lyapunov difference equations
\[ \begin{align*}
P_{k+1} &= A_k P_k A_k^\top + B_k B_k^\top \quad \text{(17a)} \\
Q_{k+1} &= A_k Q_k A_k^\top - B_k B_k^\top \quad \text{(17b)}
\end{align*} \]
for \( k \in [0, N-1] \), and the boundary conditions \( \Sigma_0 = \Sigma_0, \Sigma_N = \Sigma_N \) are written as
\[ \begin{align*}
\varepsilon \Sigma_0^{-1} &= P_0^{-1} + Q_0^{-1} \quad \text{(18a)} \\
\varepsilon \Sigma_N^{-1} &= P_N^{-1} + Q_N^{-1}. \quad \text{(18b)}
\end{align*} \]

In summary, we obtain the following proposition.

Proposition 2: Assume that for any \( k \in [0, N-1] \), \( A_k \) is invertible. Assume further that \( P_k \) and \( Q_k \) satisfy (17a) and (17b) with the boundary conditions (18a) and (18b) and are invertible on \([0, N]\), and that it holds \( I + B_k^\top Q_{k+1}^{-1} B_k > 0 \) for any \( k \in [0, N-1] \). Then, policy \( (11) \) with \( \Pi_k = Q_k^{-1} \) is the unique optimal policy of Problem 2.

Remark 2: Since Proposition 1 does not require \( x_0 \) to be Gaussian, the argument in this section still holds when the density constraints \((9)\) are replaced by the constraints on mean and covariance
\[ \mathbb{E}[x_0] = 0, \quad \mathbb{E}[x_0 x_0^\top] = \Sigma_0, \quad \mathbb{E}[x_N] = 0, \quad \mathbb{E}[x_N x_N^\top] = \Sigma_N. \] (19)

That is, the optimal policy given in Proposition 2 is also optimal for Problem 2 whose constraints \((9)\) are replaced by \((19)\). For nonzero mean constraints, see Corollary 2 in Section IV.

IV. SOLUTION TO MAX\( \varepsilon \)ENT OPTIMAL DENSITY CONTROL PROBLEM

In this section, we analyze the Lyapunov equations (17a) and (17b) with the boundary conditions (18a) and (18b). For the analysis, we introduce the reachability Gramian
\[ G_r(k_1, k_0) := \sum_{k=k_0}^{k_1-1} \Phi(k_1, k+1) B_k B_k^\top \Phi(k_1, k+1)^\top, \quad k_0 < k_1 \] (20)
and the controllability Gramian
\[ G_c(k_1, k_0) := \sum_{k=k_0}^{k_1-1} \Phi(k, k+1) B_k B_k^\top \Phi(k, k+1)^\top, \quad k_0 < k_1. \] (21)

Note that since
\[ G_c(k_1, k_0) = \Phi(k_0, k_1) G_c(k_1, k_0) \Phi(k_0, k_1)^\top \] (22)
if \( G_r(k_1, k_0) \) is invertible, \( G_c(k_1, k_0) \) is also invertible. Now, we provide the solutions to (17a) and (17b) with (18a) and (18b). The proof is shown in Appendix B.

Proposition 3: Assume that for any \( k \in [0, N-1] \), \( A_k \) is invertible and there exists \( k_r \in [1, N] \) such that \( G_r(k, 0) \) is invertible for any \( k \in [k_r, N] \) and \( G_c(N, k) \) is invertible for any \( k \in [0, k-r-1] \). Assume further that for
\[ \begin{align*}
S_0 &:= \frac{1}{\varepsilon} G_c(N, 0)^{-\frac{1}{2}} \Sigma_0 G_c(N, 0)^{-\frac{1}{2}} \quad \text{(23)} \\
S_N &:= \frac{1}{\varepsilon} G_c(N, 0)^{-\frac{1}{2}} \Phi(0, N) \Sigma_N \Phi(0, N)^\top G_c(N, 0)^{-\frac{1}{2}} \quad \text{(24)}
\end{align*} \]
the following two matrices are invertible:
\[ \begin{align*}
F(S_0, S_N) &:= S_0 + \frac{1}{2} I - \left( S_0^\frac{1}{2} S_N S_0^\frac{1}{2} + \frac{1}{4} I \right)^\frac{1}{2} \\
B(S_0, S_N) &:= -S_0 + \frac{1}{2} I + \left( S_0^\frac{1}{2} S_N S_0^\frac{1}{2} + \frac{1}{4} I \right)^\frac{1}{2}. \quad \text{(25)}
\end{align*} \]

Then, (17a) and (17b) with the boundary conditions (18a) and (18b) have two sets of solutions \((P_{\pm,k}, Q_{\pm,k})\), \( k \in [0, N] \), specified by
\[ \begin{align*}
Q_{\pm,0} &= G_c(N, 0)^{\frac{1}{2}} S_0^\frac{1}{2} \left( S_0 + \frac{1}{2} I \pm \left( S_0^\frac{1}{2} S_N S_0^\frac{1}{2} + \frac{1}{4} I \right)^\frac{1}{2} \right)^{-1} \\
&\quad \times S_0^\frac{1}{2} G_c(N, 0)^{\frac{1}{2}} \quad \text{(27)} \\
P_{\pm,0} &= (\varepsilon \Sigma_0^{-1} - Q_{\pm,0}^{-1})^{-1}. \quad \text{(28)}
\end{align*} \]

In addition, the two sets of solutions \((P_{\pm,k}, Q_{\pm,k})\) have the following properties.

i) \( P_{+,k} \) and \( Q_{-,k} \) are both invertible on \([0, N]\), and for any \( k \in [0, N-1] \), it holds \( I + B_k^\top Q_{-,k+1}^{-1} B_k > 0 \).

ii) If \( Q_{+,k} \) is invertible on \([0, N]\), there exists \( s \in [0, N-1] \) such that \( I + B_k^\top Q_{-,s+1}^{-1} B_k \) is not positive definite.

Proposition 3 says that although a pair of Lyapunov equations (17) with (18) has two sets of solutions, only the pair \((P_{-,k}, Q_{-,k})\) is qualified for the construction of the optimal policy based on Proposition 2. In summary, we come to the main result of this section.

Theorem 1: Suppose that the assumptions of Proposition 3 are satisfied. Then, the unique optimal policy of Problem 2 is given by
\[ \pi_k^*(u|x) = N(u) - (I + B_k^\top Q_{-,k+1}^{-1} B_k)^{-1} B_k^\top Q_{-,k+1}^{-1} A_k x, \]
\[ \varepsilon (I + B_k^\top Q_{-,k+1}^{-1} B_k)^{-1} \]
\[ k \in [0, N-1], \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m \] (29)
where \( Q_{-,k} \) is a solution to (17b) with the initial value \( Q_{-,0} \) in (27).

We now give some remarks on the assumptions in Proposition 3.

Remark 3: In many situations, \( A_k \) is expected to be invertible. For example, if system (8) is obtained by a zero-order
hold discretization of a continuous-time system, $A_k$ is always invertible.

**Remark 4:** For a time-invariant system $A_k \equiv A$, $B_k \equiv B$, the reachability Gramian $G_r(k_1, k_0)$ is given by

$$G_r(k_1 - k_0) := \sum_{k=0}^{k_1-k_0-1} A^k B B^T (A^T)^k.$$

Then, the invertibility assumption on the reachability Gramian in Proposition 3 means that there exists $k_r \in [1, N]$ such that $G_{r, TT}(k)$ is invertible for any $k \in [k_r, N]$ and $G_{r, TT}(N - k)$ is invertible for any $k \in [0, k_r - 1]$. In addition, it is well known that if system $(A, B)$ is reachable, there exists $k_r \in [1, n]$ such that $G_{r, TT}(k)$ is invertible for any $k \geq k_r$. Recall that $n$ denotes the dimension of the state space. Therefore, when $N \geq k_r$ and $N - k_r + 1 \geq k_r$, that is, when $N \geq 2k_r - 1$, the invertibility assumption on the reachability Gramian holds. In particular, taking a terminal time $N \geq 2n - 1$ always ensures the invertibility.

**Remark 5:** Note first that since $F(S_0, S_N)$ and $B(S_0, S_N)$ are continuous in the parameters $\{A_k\}_k, \{B_k\}_k, \Sigma_0, r, \Sigma_N, N$, and $\Sigma_N$, they become singular only in exceptional cases. In order to consider the implication of the invertibility assumption of $F(S_0, S_N)$ and $B(S_0, S_N)$, we consider the following two conditions:

$$\Sigma_{k} = \Phi(N,0)\Sigma_{0}\Phi(N,0)^T + \varepsilon G_r(N,0) \quad (30)$$

$$\Sigma_{0} = \Phi(0,0)\Sigma_{N}\Phi(0,0)^T + \varepsilon G_c(N,0). \quad (31)$$

Under (30), $F(S_0, S_N) = 0$ and under (31), $B(S_0, S_N) = 0$. When (30) holds, policy (11) with the terminal weight $F = 0$, i.e., $\Pi_k \equiv 0$, $u_k \sim \mathcal{N}(0, \varepsilon I)\forall k$, steers the state from $\mathcal{N}(0, \Sigma_0)$ to $\mathcal{N}(0, \Sigma_N)$. Hence, by the same argument as in Section III, this policy is optimal for Problem 2 although $\Pi_k$ is not invertible. On the other hand, consider the time-reversed system of (8)

$$x_{k-1} = A_{k-1}^{-1}x_{k-2} - A_{k-1}^{-1}B_k u_{k-1}, \quad k = N - 1, \ldots, 0$$

$$x_{0} \sim \mathcal{N}(0, \Sigma_{N}).$$

Then, condition (31) indicates that an independent Gaussian noise process $u_{k-1} \sim \mathcal{N}(0, \varepsilon I)$ steers $x_{k-1}$ from $\mathcal{N}(0, \Sigma_{N})$ to $\mathcal{N}(0, \Sigma_0)$.

When the invertibility assumption is not fulfilled, an analysis directly via the Riccati equation (4), rather than the Lyapunov equations (17), will be appropriate as in [35]. This analysis is beyond the scope of this article. We would like to notice that the continuous-time counterpart of Proposition 3 [9, Prop. 4] also requires the invertibility of $F(S_0, S_N)$ and $B(S_0, S_N)$, which is not mentioned in the statement.

Next, we extend the result of Theorem 1 to Problem 2' with general mean distributions (10). To this end, let us decompose $u_k$ as $u_k = \bar{u}_k + \hat{u}_k$ where $\bar{u}_k := \mathbb{E}[u_k]$. Then, the evolution of the mean $\mu_k := \mathbb{E}[x_k]$ is governed by

$$\mu_{k+1} = A_k \mu_k + B_k \bar{u}_k, \quad k \in [0, N - 1]$$

which implies that $\mu_k$ depends only on the deterministic input sequence $\{\bar{u}_k\}$. On the other hand, the stochastic control process $\hat{u}_k$ affects $\hat{x}_k := x_k - \mu_k$ as

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k \hat{u}_k, \quad k \in [0, N - 1].$$

Let $\hat{\pi}_k(\cdot | \hat{x})$ be the conditional density function of $\hat{u}_k$ given $\hat{x}_k = \hat{x}$. Since for any fixed $\{\bar{u}_k\}$,

$$\pi_k(\bar{u}_k + \hat{u} | x) = \hat{\pi}_k(\bar{u}_k, x - \mu_k)$$

$$\forall k \in [0, N - 1] \quad \forall x \in \mathbb{R}^n \quad \forall \hat{u} \in \mathbb{R}^m \quad (32)$$

it follows that for all $k \in [0, N - 1]$ and $x \in \mathbb{R}^n$,

$$H(\pi_k(\cdot | x)) = -\int_{\mathbb{R}^m} \pi_k(\bar{u}_k + \hat{u} | x) \log \pi_k(\bar{u}_k + \hat{u} | x) d\hat{u}$$

$$= -\int_{\mathbb{R}^m} \hat{\pi}_k(\bar{u}_k | x - \mu_k) \log \hat{\pi}_k(\bar{u}_k | x - \mu_k) d\hat{u}$$

$$= H(\hat{\pi}_k(\cdot | x - \mu_k)). \quad (33)$$

On the other hand, we have

$$\mathbb{E}[\|u_k\|^2] = \|\bar{u}_k\|^2 + \mathbb{E}[\|\hat{u}_k\|^2]. \quad (34)$$

For the shifted state $\bar{x}_k$, the density constraints (10) are rewritten as

$$\bar{x}_0 \sim \mathcal{N}(0, \Sigma_0), \quad \bar{x}_N \sim \mathcal{N}(0, \Sigma_N). \quad (35)$$

By (33)–(35), Problem 2' can be decomposed into mean steering and covariance steering as follows.

**Problem 3:** Find a deterministic control process $\bar{u} = \{\bar{u}_k\}_{k=0}^{N-1}$ and a policy $\hat{\pi} = \{\hat{\pi}_k\}_{k=0}^{N-1}$ that solve

$$\min_{\bar{u}, \hat{\pi}} \quad \sum_{k=0}^{N-1} \frac{1}{2} \|\bar{u}_k\|^2$$

$$+ \mathbb{E}\left[\sum_{k=0}^{N-1} \left(\frac{1}{2} \|\bar{u}_k\|^2 - \varepsilon H(\hat{\pi}_k(\cdot | \bar{x}_k))\right)\right]$$

subject to

$$\mu_{k+1} = A_k \mu_k + B_k \bar{u}_k$$

$$\bar{u}_0 = \bar{u}_0, \quad \mu_N = \bar{u}_N$$

$$\bar{x}_{k+1} = A_k \bar{x}_k + B_k \bar{u}_k$$

$$\bar{x}_0 \sim \mathcal{N}(0, \Sigma_0), \quad \bar{x}_N \sim \mathcal{N}(0, \Sigma_N). \quad (36)$$

Noting that $\{\bar{u}_k\}$ is relevant only for the first term of (36), the unique optimal solution $\{\bar{u}_k\}$ to Problem 3 is given by

$$\bar{u}_k = B_k^\top \Phi(N, k + 1)^\top G_r(N, 0)^{-1}(\mu_N - \Phi(N, 0)^\top \bar{u}_0). \quad (37)$$

See also (44) in Section V. Moreover, by Theorem 1, the optimal policy $\hat{\pi}^* = \{\hat{\pi}_k^*\}$ is given by

$$\hat{\pi}_k^*(\bar{u} | \bar{x}) = \mathcal{N}(\bar{u} - (I + B_k^\top Q_{-k+1}^{-1} B_k)^{-1} B_k^\top Q_{-k+1}^{-1} A_k \bar{x},$$

$$\varepsilon (I + B_k^\top Q_{-k+1}^{-1} B_k)^{-1}).$$

Finally, by (32), we obtain the optimal policy of Problem 2' as follows.

**Corollary 1:** Suppose that the assumptions of Proposition 3 are satisfied. Let $Q_{-k}$ be a solution to (17b) with the initial value $Q_{-0}$ in (27). Then, the unique optimal policy of Problem 2' is
given by
\[
\pi_k^c(u|x) = \mathcal{N}(u) - (I + B_k^\top Q_{k+1}^{-1} B_k)^{-1} B_k^\top Q_{k+1}^{-1} A_k \\
\times (x - \mu_k^s) + \bar{u}_k^s \varepsilon(I + B_k^T Q_{k+1}^{-1} B_k)^{-1}
\]
for any \( k \in [0, N - 1], \) \( x \in \mathbb{R}^n, u \in \mathbb{R}^m \)

(38)

where \( \{ \bar{u}_k^s \} \) is given by (37) and
\[
\mu_k^s := \begin{cases} \\
\Phi(k, 0) \bar{\mu}_o + \sum_{s=0}^{k-1} \Phi(k, s+1) B_s \bar{u}_s^s & k \in [1, N - 1] \\
\bar{\mu}_0, & k = 0.
\end{cases}
\]

(39)

Finally, we consider an optimal covariance steering problem, where the initial and final densities are not necessarily Gaussian. By Remark 2 and the decomposition into mean steering and covariance steering given in Problem 3, we obtain the following result as a straightforward consequence of Corollary 1.

**Corollary 2:** Suppose that the assumptions of Proposition 3 are satisfied. Then, the unique optimal policy of Problem 2 whose constraints (9) are replaced by
\[
\mathbb{E}[x_0] = \bar{\mu}_o, \quad \mathbb{E}[(x_0 - \bar{\mu}_o)(x_0 - \bar{\mu}_o)^\top] = \Sigma_0
\]
\[
\mathbb{E}[x_N] = \bar{\mu}_N, \quad \mathbb{E}[(x_N - \bar{\mu}_N)(x_N - \bar{\mu}_N)^\top] = \Sigma_N
\]

(40)

(41)

is given by (38).

V. POINT-TO-POINT STEERING OF DISCRETE-TIME LINEAR SYSTEMS

In the previous section, we discussed the density-to-density transfer over linear systems. In this section, we investigate the point-to-point steering of linear systems with entropy regularization. This can be seen as the limiting case where the variances of the initial and target distributions go to zero.

Intuitively, the optimal policy for the point-to-point steering can be obtained by using an infinitely large terminal weight \( F \) in Problem 1. With this in mind, let us consider Problem 1 with \( F = I / \delta, \) \( \delta > 0 \). In the rest of this article, we assume \( \varepsilon = 1 \); see Remark 1. In addition, we assume the invertibility of \( A_k \) on \([0, N - 1]\). Note that if \( G_c(N, k+1) \) is invertible, as \( \delta \downarrow 0 \), the optimal policy \( \pi_k^c(u|x) \) in (6) converges to
\[
\mathcal{N}(u) - (I + B_k^\top Q_{k+1}^{-1} B_k)^{-1} B_k^\top Q_{k+1}^{-1} A_k(x - \Phi(k, N)\bar{x}_N), \\
(I + B_k^\top Q_{k+1}^{-1} B_k)^{-1}
\]

where
\[
Q_{k+1} = A_k Q_{k+1} A_k^\top - B_k B_k^\top
\]
\[
Q_N = 0.
\]

(43a)

(43b)

However, this is no longer true for time \( k \) when \( Q_{k+1} = G_c(N, k+1) \) is not invertible. Instead of (42), to obtain an expression that is valid even when \( Q_{k+1} \) is not invertible, recall that the mean of (6) coincides with the form of the LQ optimal controller. In particular, by taking \( \delta \downarrow 0 \), the associated LQ optimal control from time \( k \) to the terminal time \( N \) given the current state \( x_k = \bar{x}_k \) converges to the solution of the following fixed end-point deterministic optimal control problem:
\[
\text{minimize}_{\{u_s\}_{s=k}^{N-1}} \sum_{s=k}^{N-1} \|u_s\|^2
\]

subject to \( x_{s+1} = A_s x_s + B_s u_s, \) \( s \in [k, N - 1] \)
\[
x_k = \bar{x}_k, \ x_N = \bar{x}_N.
\]

(44)

Since we have
\[
x_N = \Phi(N, k) x_k \\
+ [\Phi(N, k + 1) B_k \Phi(N, k + 2) B_{k+1} \cdots B_{N-1}] \times [u_k; \cdots; u_{N-1}]
\]

the optimal solution to (44) is
\[
[u_k; \cdots; u_{N-1}]
\]

\[
= [\Phi(N, k + 1) B_k \Phi(N, k + 2) B_{k+1} \cdots B_{N-1}] \times \Phi(N, k)\bar{x}_k
\]

which immediately gives the optimal feedback control law at time \( k \) given the current state \( x_k \)
\[
u_k = -B_k^\top \Phi(N, k + 1) \times G_c(N, k)^\top \Phi(N, k) (x_k - \Phi(k, N)\bar{x}_N).
\]

In other words, as \( \delta \downarrow 0 \), the mean of the MaxEnt optimal policy (6) converges to
\[
-B_k^\top \Phi(N, k + 1)^\top G_c(N, k)^\top \Phi(N, k) (x - \Phi(k, N)\bar{x}_N).
\]

(45)

On the other hand, let \( Q_{\delta, k} \) be the solution to (17b) with \( Q_N = F^{-1} = \delta I \). Then,
\[
Q_{\delta, k} = \delta \Phi(k, N)\Phi(k, N)^\top + G_c(N, k).
\]

For the covariance matrix of the MaxEnt optimal policy (6), we have
\[
(I + B_k^\top Q_{\delta, k+1}^{-1} B_k)^{-1} = I - B_k^\top (Q_{\delta, k+1}^{-1} + B_k B_k^\top)^{-1} B_k
\]
\[
= I - B_k^\top \Phi(N, k + 1)^\top (G_c(N, k) + \delta I)^{-1} \Phi(N, k + 1) B_k
\]

where we used (17b) and (22). In addition, it holds
\[
\lim_{\delta \downarrow 0} [\Phi(N, k + 1) B_k \cdots B_{N-1}]^\top (G_c(N, k) + \delta I)^{-1} = [\Phi(N, k + 1) B_k \cdots B_{N-1}]^\top
\]

\[
[\Phi(N, k + 1) B_k \cdots B_{N-1}]^\top G_c(N, k)
\]

where we used the fact \( \lim_{\delta \downarrow 0} X_k \times (X_k^\top + \delta I)^{-1} = X_k^\top \) for any \( X_k \in \mathbb{R}^{m \times n} \) [36, Ch. 3, Ex. 25]. Therefore, we obtain
\[
\lim_{\delta \downarrow 0} [I + B_k^\top Q_{\delta, k+1}^{-1} B_k]^{-1}
\]
\[
= I - B_k^\top \Phi(N, k + 1)^\top G_c(N, k)^\top \Phi(N, k + 1) B_k.
\]

(46)
In summary, we obtain the optimal policy for $\delta \downarrow 0$

$$N \left( u | - B_k^\dagger \Phi(N, k + 1)^\top G_r(N, k)^\dagger \Phi(N, k) \right)$$

$$\times (x - \Phi(k, N) x_N),$$

$$I - B_k^\dagger \Phi(N, k + 1)^\top G_r(N, k)^\dagger \Phi(N, k + 1) B_k \right). \tag{47}$$

The state process driven by (47) with a fixed initial state $\bar{x}_0 \in \mathbb{R}^n$ follows

$$x_{k+1}^* = \hat{A}_k x_k^* + B_k B_k^\top \Phi(N, k + 1)^\top G_r(N, k)^\dagger x_N + B_k w_k^* \tag{48}$$

$$x_0^* = \bar{x}_0, \text{ a.s.} \tag{49}$$

where $\hat{A}_k := (I - B_k B_k^\top \Phi(N, k + 1)^\top G_r(N, k)^\dagger \Phi(N, k + 1) B_k)$.

Especially when $k = N - 1$, (46) is the orthogonal projection matrix onto the null space of $B_{N-1}$. Hence, if $B_{N-1} \neq 0$, the covariance (46) at $k = N - 1$ is degenerate, and therefore, the negative entropy of (47) at $k = N - 1$ and the cost (1) diverge to infinity. For this reason, control policy (47) cannot be characterized in terms of optimality. Nevertheless, policy (47) is still implementable with finite energy. Moreover, in the following subsections, we see that the corresponding state process $\{x_k^*\}$ can be characterized as a pinned process of the associated linear system defined as follows.

**Definition 1 (32):** Let $\bar{x}_0, x_N \in \mathbb{R}^n$ and consider the noise-driven system

$$\bar{x}_{k+1} = \hat{A}_k \bar{x}_k + B_k w_k, \ w_k \sim N(0, I)$$

$$\bar{x}_0 = \bar{x}_0, \text{ a.s.} \tag{50}$$

where $\{w_k\}$ is an independent sequence. Then, a stochastic process $\{x_k\}_{k=0}^N$ is said to be a pinned process associated with (50) from $\bar{x}_0$ to $x_N$ if the joint distribution of $\{x_k\}_{k=0}^N$ coincides with the conditional distribution of $\{\bar{x}_k\}_{k=0}^N$ given $\bar{x}_N = x_N$. \hfill \Diamond

**A. Steering From the Zero Initial State to the Zero Terminal State**

First, we deal with the case where two end points are given by $\bar{x}_0 = 0$ and $x_N = 0$. Then, by linearity of (50), $\{\bar{x}_k\}_{k=0}^N$ follows a zero-mean Gaussian distribution. Therefore, it suffices to consider only the second moment of the conditional distribution of $\{\bar{x}_k\}$ given $\bar{x}_N = 0$. The covariance matrix of $[\bar{x}_k; \bar{x}_s; x_N] \ (k < s)$ is given by

$$\begin{bmatrix}
    P_k & P_k \Phi(s, k)^\top & P_k \Phi(N, k)^\top \\
    \Phi(s, k) P_k & P_s & P_s \Phi(N, s)^\top \\
    \Phi(N, k) P_k & \Phi(N, s) P_s & P_N
\end{bmatrix} \tag{51}$$

where $P_k := \mathbb{E}[\bar{x}_k \bar{x}_k^\top]$ satisfies

$$P_{k+1} = A_k P_k A_k^\top + B_k B_k^\top, \ P_0 = 0. \tag{52}$$

Taking the Schur complement of (51), we obtain the covariance of the conditional distribution of $[\bar{x}_k; \bar{x}_s]$ given $\bar{x}_N = 0$ as follows:

$$\begin{bmatrix}
    \hat{P}(k, k) & \hat{P}(k, s) \\
    \hat{P}(k, s)^\top & \hat{P}(s, s)
\end{bmatrix}$$

where

$$\hat{P}(k, s) := P_k \Phi(s, k)^\top - P_k \Phi(N, k)^\top P_N^{-1} \Phi(N, s) P_s.$$

Let $P^\bullet(s, k) := \mathbb{E}[x_k^* (x_s^*)^\top]$. By showing $P^\bullet(s, k) = \hat{P}(k, s)$ for $k \leq s$, we obtain the following proposition. The proof is given in Appendix C.

**Proposition 4:** Assume that $P_N = G_r(N, 0)$ is invertible, and $A_k$ is invertible for any $k \in [0, N - 1]$. Then, the state process $\{x_k^*\}_{k=0}^N$ following (48) and (49) with $\bar{x}_0 = 0, x_N = 0$ is a pinned process associated with (50) from $\bar{x}_0$ to $x_N = 0$. \hfill \Diamond

**B. Steering Between Arbitrary Boundary Points**

Next, we consider pinned processes between the general boundary points $\bar{x}_0 = \bar{x}_0$ and $x_N = x_N$. The analysis of the second moment is same as for Proposition 4. Hence, we investigate only the first moment. Note that since

$$\mathbb{E}[\bar{x}_k] = \Phi(k, 0) \bar{x}_0$$

it holds

$$\ell_k := \mathbb{E}[\bar{x}_k | \bar{x}_N = x_N]$$

$$= \phi(k, 0) \bar{x}_0 + P_k \Phi(N, k)^\top P_N^{-1} (x_N - \Phi(N, 0) \bar{x}_0).$$

From (48), by showing

$$\ell_{k+1} = \hat{A}_k \ell_k + B_k \Phi(N, k + 1)^\top G_r(N, k)^\dagger x_N \tag{53}$$

we obtain the main result of this section. The proof is given in Appendix C.

**Theorem 2:** Assume that $G_r(N, 0)$ is invertible, and $A_k$ is invertible for any $k \in [0, N - 1]$. Then, for any $\bar{x}_0, x_N \in \mathbb{R}^n$, the state process $\{x_k^*\}_{k=0}^N$ following (48) and (49) is a pinned process associated with (50) from $\bar{x}_0$ to $x_N$. \hfill \Diamond

**VI. SCHRÖDINGER BRIDGES OF DISCRETE-TIME LINEAR SYSTEMS**

In this section, we explore the connection between the MaxEnt optimal density control and Schrödinger bridges. Given a reference stochastic process with finite horizon $N$ and two marginals, the Schrödinger bridge problem seeks to find the closest process to the reference in terms of the KL divergence such that its marginals at times $k = 0$ and $k = N$ coincide with the given marginals; see e.g., [37] and [38] for reviews. Specifically, in this article, the reference process and marginals are given by the noise-driven linear system (50) and the Gaussian distributions (9), respectively. Then, the Schrödinger bridge problem is formulated as follows.

**Problem 4 (Schrödinger bridge problem):** Let $\{\chi_k\}_{k=0}^N$ be the solution to the linear system (50) with the initial state $\bar{x}_0 \sim N(0, \Sigma_0), \Sigma_0 > 0$ and denote by $\mathbb{P}_{0,N}$ the probability distribution of $\{\bar{x}_k\}_{k=0}^N$ on $\chi := (\mathbb{R}^n)^{N+1}$. Then, find a probability
distribution $\hat{P}_{0:N}$ on $X$ that solves

$$\text{minimize}_{\hat{P}_{0:N} \in \Pi(\Sigma_0, \Sigma_N)} D_{KL}\left(\hat{P}_{0:N} \parallel \hat{P}_{0:N}\right)$$

(54)

where $\hat{P}_{0:N} \in \Pi(\Sigma_0, \Sigma_N)$ means that the marginals of $\hat{P}_{0:N}$ at $k = 0$ and $k = N$ are equal to $\mathcal{N}(0, \Sigma_0)$ and $\mathcal{N}(0, \Sigma_N)$, $\Sigma_N > 0$, respectively.

The optimal solution $\hat{P}_{0:N}$ to the above problem is referred to as the Schrödinger bridge associated with (50) from $\mathcal{N}(0, \Sigma_0)$ to $\mathcal{N}(0, \Sigma_N)$. In this article, we also refer to the corresponding stochastic process $\{\tilde{x}_k\}_{k=0}^N \sim \hat{P}_{0:N}$ as the Schrödinger bridge. Since the KL divergence $D_{KL}(\cdot || \cdot)$ is strictly convex, and the constraint set $\Pi(\Sigma_0, \Sigma_N)$ is a convex subset of the vector space of all bounded measures on $X$, the Schrödinger bridge problem is a strictly convex problem. Hence, Problem 4 admits at most one solution.

Denote by $\hat{P}_{0:N}$ and $\hat{P}_{1:N-1|0:N}$ the joint distribution of $(\tilde{x}_0, \tilde{x}_N)$ and the conditional distribution of $\{\tilde{x}_k\}_{k=0}^{N-1}$ given $(\tilde{x}_0, \tilde{x}_N)$, respectively. The same notation is used for $P$. For the analysis of Problem 4, we employ the decomposition of the KL divergence [39, Th. 2]

$$D_{KL}\left(\hat{P}_{0:N} \parallel \hat{P}_{0:N}\right) = D_{KL}\left(P_{0:N} \parallel \hat{P}_{0:N}\right) + \int D_{KL}(\hat{P}_{1:N-1|0,N}(\cdot | \tilde{x}_0 = \tilde{x}_0, \tilde{x}_N = \tilde{x}_N))\hat{P}_{0,N}(d\tilde{x}_0 d\tilde{x}_N).$$

(55)

The first term depends only on the joint distribution of $(\tilde{x}_0, \tilde{x}_N)$. On the other hand, by taking $\hat{P}_{1:N-1|0,N}(\cdot | \tilde{x}_0 = \tilde{x}_0, \tilde{x}_N = \tilde{x}_N) = \hat{P}_{1:N-1|0,N}(\cdot | \tilde{x}_0 = \tilde{x}_0, \tilde{x}_N = \tilde{x}_N)$, the second term attains its minimum value 0. This implies that for any $\tilde{x}_0, \tilde{x}_N \in \mathbb{R}^n$, the Schrödinger bridge $\{\tilde{x}_k\}$ given $\tilde{x}_0 = \tilde{x}_0, \tilde{x}_N = \tilde{x}_N$ is a pinned process associated with (50) from $\tilde{x}_0 = \tilde{x}_0$ to $\tilde{x}_N = \tilde{x}_N$.

In [22], the optimal solution to Problem 4 is derived only for the restricted case where $B_k$ has full row rank. In the following theorem, for general $B_k$, we reveal that the Max-Ent optimal state process $\{x^*_k\}_{k=0}^N$ of Problem 2 is actually an optimal solution to Problem 4. The proof is given in Appendix D.

**Theorem 3:** Suppose that the assumptions of Proposition 3 are satisfied. Consider the optimal state process $\{x^*_k\}_{k=0}^N$ of Problem 2 with $\varepsilon = 1$ following

$$x^*_{k+1} = A_{Q,k} x^*_k + B_{Q,k} w_k, \quad w_k \sim \mathcal{N}(0, I)$$

(56)

where

$$A_{Q,k} := A_k - B_k (I + B_k^T Q^{-1}_{k-1} B_k)^{-1} B_k^T Q^{-1}_{k-1} A_k$$

$$B_{Q,k} := B_k (I + B_k^T Q^{-1}_{k-1} B_k)^{-1/2}.$$  

Then, the probability distribution $\hat{P}_{0:N}^*$ of $\{x^*_k\}_{k=0}^N$ is the unique optimal solution to Problem 4.

![Fig. 1.](image)

**VII. ILLUSTRATIVE EXAMPLES**

In this section, two examples illustrate the obtained results. The system is given by

$$A_k = 
\begin{bmatrix}
0.9 & 0.1 \\
0.05 & 1.2
\end{bmatrix}, \quad B_k = 
\begin{bmatrix}
0 & 0.22 \\
0 & 0
\end{bmatrix}^T \quad \forall k.$$  

(57)

The initial density $\mathcal{N}(0, \Sigma_0)$ and the target density $\mathcal{N}(0, \Sigma_N)$ are determined by

$$\Sigma_0 = 
\begin{bmatrix}
7 & 3 \\
3 & 5
\end{bmatrix}, \quad \Sigma_N = 0.3 I$$

(58)

and set $N = 50$. Then, the assumptions of Proposition 3 and Theorem 2 are fulfilled. In Fig. 1, the sample paths of the optimal state process and the optimal control process obtained in Theorem 1 are shown for different $\varepsilon$. The colored lines indicate the samples, and the black ellipses in Fig. 1(a) and (b) are the $3\sigma$ covariance ellipses for $\mathcal{N}(0, \Sigma_0)$ and $\mathcal{N}(0, \Sigma_N)$ given by

$$\left\{ x \in \mathbb{R}^2 : x^T \Sigma^{-1}_k x = 3^2 \right\}, \quad k = 0, N.$$  

(59)
As can be seen, the optimal control for ε = 1.0 has a larger variance than for ε = 0.02 resulting in higher entropy of the policy. The same is true for the state processes. Despite the difference in entropy, the state is steered to the target distribution in both cases. Note that for ε = 1.0, the eigenvalues of $Q^{-1}_N$, which corresponds to the terminal weight matrix $F$ in Problem 1, are given by $\{-45.81, 3.33\}$. This clarifies that it is not enough to consider only positive semidefinite terminal weight matrices for constructing the MaxEnt optimal density control.

Next, we describe the point-to-point steering for (57) with $\bar{x}_0 = [-2 \ 4]^\top$, $\bar{x}_N = [1 \ 0]^\top$. Fig. 2 depicts the samples of the pinned process $\{x_n^*\}$ following (48) and the corresponding control process given by (47). It can be seen that all the samples of the state are transferred to the target state $\bar{x}_N$ in spite of the fluctuations along the paths.

VIII. CONCLUSION

In this article, we analyzed MaxEnt optimal density control of deterministic discrete-time linear systems. First, we revealed the existence and uniqueness of the optimal solution, and provided the explicit construction of the optimal policy. Next, we considered the point-to-point steering and characterized the associated state process as a pinned process of the noise-driven linear system. Finally, we showed that the MaxEnt optimal density control for a deterministic discrete-time system, the input and the noise induced by entropy regularization are injected into the system through the same channel. Therefore, an optimization based approach [19], [20], rather than the one based on Lyapunov equations, would work for MaxEnt optimal density control of stochastic systems.

APPENDIX A

PROOF OF PROPOSITION 1

Similar to the conventional optimal control, the dynamic programming approach can be applied to MaxEnt optimal control [6]. Let us introduce the value function associated with Problem 1

\[
V(k, x) := \inf_{\{\pi_s\}_{s=1}^{N+1}} \mathbb{E} \left[ \frac{1}{2} (x_N - \bar{x}_N)^\top F (x_N - \bar{x}_N) + \sum_{s=k}^{N-1} \left( \frac{1}{2} \|u_s\|^2 - \varepsilon \mathbb{H}(\pi_s(x_s)) \right) \right] \quad \text{subject to (2)}
\]

\[
k \in [0, N - 1], \quad x \in \mathbb{R}^n
\]

It is known that $V$ satisfies the soft Bellman equation

\[
V(k, x) = -\varepsilon \log \left[ \int_{\mathbb{R}^m} \exp \left( -\frac{1}{\varepsilon} Q_k(x, u) \right) \, du \right],
\]

\[
k \in [0, N - 1], \quad x \in \mathbb{R}^n
\]

(60)

where the Q-function is defined by

\[
Q_k(x, u) := \frac{1}{2} \|u\|^2 + V(k+1, A_k x + B_k u),
\]

\[
x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m
\]

Then, the unique optimal policy is given by

\[
\pi_k^*(u|x) \propto \exp \left( -\frac{1}{\varepsilon} Q_k(x, u) \right), \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m
\]

(61)

In what follows, we calculate the Q-function analytically. For conciseness, we drop the subscripts of $A_k$, $B_k$. First, for $k = N - 1$, we have

\[
Q_{N-1}(x, u) = \frac{1}{2} \|u + (I + B^\top \Pi_N B)^{-1} B^\top \Pi_N (Ax - \bar{x}_N)\|^2_{I + B^\top \Pi_N B} + \frac{1}{2} \|x - A^{-1} \bar{x}_N\|^2_{\Pi_N - 1 (x - A^{-1} \bar{x}_N)}
\]

where we used the invertibility of $I + B^\top \Pi_N B$ by assumption. By combining this with (61), we obtain policy (6) for $k = N -
1. On the other hand, note that under $I + B^\top \Pi_N B > 0$, for any $\mu \in \mathbb{R}^m$

$$
\int_{\mathbb{R}^m} \exp \left( -\frac{1}{2} \| u - \mu \|^2 (I + B^\top \Pi_N B) / \varepsilon \right) \, du \\
= \sqrt{(2\pi)^m \det(\varepsilon(I + B^\top \Pi_N B)^{-1})}.
$$

Then, by (60), a straightforward calculation yields

$$
V(N - 1, x) = \frac{1}{2} (x - \Phi(N - 1, N)\tilde{x}_N)^\top \Pi_{N-1} (x - \Phi(N - 1, N)\tilde{x}_N) \\
- \varepsilon \log \sqrt{(2\pi)^m \det(\varepsilon(I + B^\top \Pi_N B)^{-1})}
$$

where the first term takes the same form as $V(N, x)$ and the second term does not depend on $x$. In addition, if $F \geq 0$, then $\Pi_N$ has the square root $\Pi_N^{1/2}$, and by (4), $\Pi_{N-1}$ can be written as

$$
\Pi_{N-1} = A^\top \Pi_N^{1/2} (I + \Pi_N^{1/2}BB^\top \Pi_N^{1/2})^{-1} \Pi_N^{1/2} A \succeq 0.
$$

Moreover, it is obvious that the above argument holds when $\tilde{x}_N = 0$ and $A_{N-1}$ is not invertible. Finally, by applying the same procedure as above for $k = N - 2, \ldots, 0$, we obtain the desired result.

**APPENDIX B**

**PROOF OF PROPOSITION 3**

First, introduce the change of variables $\xi_k := G_c(N, 0)^{-1/2}\Phi(0, k)x_k$. Then, system (8) is transformed into

$$
\xi_{k+1} = \xi_k + G_c(N, 0)^{-1/2}\Phi(0, k+1)B_ku_k.
$$

We will prove the statement in this new set of coordinates and then turn back to the original set of coordinates at the end. The Lyapunov equations associated with the transformed system (62) are given by

$$
P_{new,k+1} = P_{new,k} + B_{new,k}B_{new,k}^\top, \quad Q_{new,k+1} = Q_{new,k} - B_{new,k}B_{new,k}^\top.
$$

The relationship between $Q_{new,k}$ and $Q_k$ is as follows:

$$
Q_{new,k} = G_c(N, 0)^{-1/2}\Phi(0, k)Q_k\Phi(0, k)^\top G_c(N, 0)^{-1/2}. \quad (64)
$$

Indeed, substituting (64) into $Q_{new,k+1} - Q_{new,k}$ yields

$$
Q_{new,k+1} - Q_{new,k} = G_c(N, 0)^{-1/2}\Phi(0, k) \\
\times (A_k^{-1}Q_{k+1}A_k^{-\top} - Q_k)\Phi(0, k)^\top G_c(N, 0)^{-1/2} \\
= -G_c(N, 0)^{-1/2}\Phi(0, k+1)B_kB_k^\top\Phi(0, k+1)^\top G_c(N, 0)^{-1/2} \\
= -B_{new,k}B_{new,k}^\top
$$

which coincides with (63b). The controllability Gramian and the reachability Gramian corresponding to $\xi_k$ are given by

$$
G_{c,new}(k_1, k_0) = G_{t,new}(k_1, k_0) := \sum_{k=k_0}^{k_1-1} B_{new,k}B_{new,k}^\top
$$

satisfying $G_{c,new}(N, 0) = G_{t,new}(N, 0) = I$. The initial and final covariance matrices for $\xi_0$ and $\xi_N$ are given by

$$
\Sigma_{new,0} := G_c(N, 0)^{-1/2}\Sigma_0G_c(N, 0)^{-1/2} = \varepsilon S_0 \\
\Sigma_{new,N} := G_c(N, 0)^{-1/2}\Phi(0, N)\Sigma_0\Phi(0, N)^\top G_c(N, 0)^{-1/2} = \varepsilon S_N.
$$

In what follows, to simplify notation, we omit the subscript “new.” By substituting

$$
P_N = P_0 + I, \quad Q_N = Q_0 - I
$$

into the boundary conditions (18a) and (18b), similar to the proof of [9, Prop. 4], we obtain the two sets of initial values

$$
Q_{\pm,0} = S_0^\pm \left( S_0 + \frac{1}{2} I \pm \left( S_0^2 S_N S_0^2 + \frac{1}{4} I \right)^{1/2} \right)^{-1} S_0^2 \\
P_{\pm,0} = (S_0^{-1} - Q_{\pm,0}^{-1}).
$$

Here, the invertibility of $F(S_0, S_N)$ and $B(S_0, S_N)$ guarantees the invertibility of $Q_{-,0}$ and $P_{-,0}$. Next, we show that $Q_{-,k} = Q_{-,0} - G_r(k, 0)$ is invertible on $[0, N]$. Note that formally

$$
(Q_{-,0} - G_r(k, 0))^{-1} = -G_r(k, 0)^{-1} \\
- G_r(k, 0)^{-1} (Q_{-,0}^{-1} - G_r(k, 0)^{-1})^{-1} G_r(k, 0)^{-1} \\
= -G_r(k, 0)^{-1} - G_r(k, 0)^{-1} S_0^2 \left[ S_0 + \frac{1}{2} I \right] \\
- \left( S_0^2 S_N S_0^2 + \frac{1}{4} I \right)^{1/2} - S_0^2 G_r(k, 0)^{-1} S_0^2 \\
S_0^2 G_r(k, 0)^{-1}.
$$

By assumption, for any $k \in [k_r, N]$, $G_r(k, 0)$ is invertible. The term in the square brackets obviously attains its maximum at

$$

\frac{1}{2} I - \left( S_0^2 S_N S_0^2 + \frac{1}{4} I \right)^{1/2} < 0.
$$

Therefore, the term in the brackets is invertible for any $k \in [k_r, N]$. This implies that $Q_{-,k}$ has the inverse matrix (66). On the other hand, $Q_{-,k}$ also admits the expression $Q_{-,k} = Q_{-,0} + G_r(k, N)$. Hence, the inverse matrix is formally

$$
(Q_{-,0} + G_r(k, N))^{-1} = G_r(k, N)^{-1} \\
- G_r(k, N)^{-1} (Q_{-,0}^{-1} + G_r(k, N)^{-1})^{-1} G_r(k, N)^{-1} \\
= G_r(N, k)^{-1} - G_r(k, N)^{-1} \\
\times ((Q_{-,0} - I)^{-1} + G_r(k, N)^{-1})^{-1} G_r(k, N)^{-1} \\
= G_r(N, k)^{-1} - G_r(k, N)^{-1} S_0^2 \left[ -S_0 + \left( -\frac{1}{2} I + (S_0^2 S_N S_0^2 + \frac{1}{4} I) \right)^{1/2} \right]^{-1} S_0^2 G_r(k, N)^{-1}.
$$

(67)
Then, by the same argument as for the time interval \([k_r, N]\), \(Q_{r-k}\) is also invertible on \([0, k_r - 1]\). Similarly, it can be shown that \(P_{r-k}\) is invertible on \([0, N]\).

In the next step, we prove that \(I + B_k^\top Q_{r-k+1} B_k > 0\) for any \(k \in [0, N - 1]\). Note that

\[
(I + B_k^\top Q_{r-k+1} B_k)^{-1} = I - B_k^\top (Q_{r-k+1} + B_k B_k^\top)^{-1} B_k
\]

Hence, it suffices to show that \(I - B_k^\top Q_{r-k+1} B_k > 0\). By (66), for \(k \in [k_r, N - 1]\), \(N \geq 2\), we have

\[
I - B_k^\top Q_{r-k+1} B_k = I + B_k^\top G_r(k, 0)^{-1} B_k
\]

\[
+ B_k^\top G_r(k, 0)^{-1} S_0^\frac{1}{2} \left( S_0 + \frac{1}{2} I - \left( S_0^\frac{1}{2} S_N S_0^\frac{1}{2} + \frac{1}{4} I \right)^\frac{1}{2} \right)^\frac{1}{2}
\]

\[
- S_0^\frac{1}{2} G_r(k, 0)^{-1} S_0^\frac{1}{2} G_r(k, 0)^{-1} B_k.
\]

Since the expression in the square brackets is negative definite on \([k_r, N - 1]\), it holds for sufficiently small \(\delta \in (0, 1)\)

\[
I - B_k^\top Q_{r-k+1} B_k > I + B_k^\top (G_r(k, 0) - (1 - \delta)^{-1} I)^{-1} B_k
\]

\[
= I - B_k^\top \left( \sum_{s=k}^{N-1} B_s B_s^\top + ((1 - \delta)^{-1} - 1) I \right)^{-1} B_k.
\]

In addition, it holds that

\[
\left[ I - B_k^\top \left( \sum_{s=k}^{N-1} B_s B_s^\top + (1 - \delta)^{-1} I \right)^{-1} B_k \right]^{-1}
\]

\[
= I + B_k^\top \left( \sum_{s=k+1}^{N-1} B_s B_s^\top + (1 - \delta)^{-1} I \right)^{-1} B_k > 0.
\]

Consequently, we obtain \(I + B_k^\top Q_{r-k+1} B_k > 0\) for \(k \in [k_r, N - 1]\). Noting also that \(Q_{r-k}\) is given by (67) for \(k \in [0, k_r - 1]\), and by the same argument above, it can be shown that \(I + B_k^\top Q_{r-k+1} B_k > 0\) for \(k \in [0, k_r - 1]\).

Next, we show property ii). Note that since \(0 < Q_{r+0} < I\), it holds \(Q_{r+n} < 0\) by (65). Thus, by (63b), there exists \(s \in [0, N - 1]\) such that \(Q_{r+s} > 0\) and \(Q_{r+s+1}\) are not positive definite. By assumption, \(Q_{r+s+1}\) is invertible and

\[
Q_{r+s+1}^{-1} = (Q_{r+s} - B_s B_s^\top)^{-1}
\]

\[
= Q_{r+s}^{-1} + Q_{r+s}^{-1} B_s (I - B_s^\top Q_{r+s}^{-1} B_s)^{-1} B_s^\top Q_{r+s}^{-1}.
\]



Now assume \(I - B_s^\top Q_{r+s}^{-1} B_s > 0\). Then, by \(Q_{r+s}^{-1} > 0\) and (69), we have \(Q_{r+s+1}^{-1} > 0\), which contradicts the fact that \(Q_{r+s+1}\) is not positive definite. Combining this with \(I + B_s^\top Q_{r+s+1} B_s = (I - B_s^\top Q_{r+s+1} B_s)^{-1}\), we conclude ii).

Finally, by employing the relationship (64), we obtain the desired result.

### Appendix C

#### PROOFS OF PROPOSITION 4 AND THEOREM 2

**A. Proof of Proposition 4**

For simplicity, we write \(P_k^* := P^*(k, k), \tilde{P}_k := \tilde{P}(k, k), \Phi_k := \Phi(N, k), G_{r,k} := G_r(N, k), G_{c,k} := G_c(N, k).\) First, we verify \(P_k^* = \tilde{P}_k\). Note that

\[
P_{k+1}^* = A_k P_k^* A_k^\top + B_k B_k^\top - B_k B_k^\top \Phi_{k+1} G_{r,k}^\top \Phi_{k+1}^\top B_k B_k^\top P_0^* = 0
\]

and \(\tilde{P}_0 = 0\). Thereafter, we show that \(\tilde{P}_k\) satisfies

\[
\tilde{P}_{k+1} = A_k \tilde{P}_k A_k^\top + B_k B_k^\top - B_k B_k^\top \Phi_{k+1} G_{r,k}^\top \Phi_{k+1}^\top B_k B_k^\top
\]

(70)

which implies \(P_k^* = \tilde{P}_k\).

Let \(T_k := P_k + Q_k\). Since \(T_{k+1} = A_{k+1} A_k^\top\) by (52) and (43a), it holds \(T_k = \Phi(0) T_0 \Phi(k, 0)^\top\). In addition, by the invertibility of \(A_k\) and \(T_N = P_N = G_{r,0}\), \(T_k\) is invertible for any \(k \in [0, N]\). Now, we have

\[
T_k \Phi_k^\top P_N^{-1} \Phi_k = \Phi(k, 0) T_0 \Phi(0, 0)^\top \Phi_0^\top P_N^{-1} \Phi_k
\]

\[
= \Phi(k, N) \Phi(N, 0) T_0 \Phi(N, 0)^\top P_N^{-1} \Phi_k
\]

\[
= \Phi(k, N) T_N P_N^{-1} \Phi(N, k) = I.
\]

Hence, it holds

\[
\Phi_k^\top P_N^{-1} \Phi_k = (P_k + Q_k)^{-1}.
\]

(71)

By using this, we obtain

\[
\tilde{P}_{k+1} = P_{k+1} + (P_{k+1} + Q_{k+1})^{-1} P_{k+1}
\]

\[
= A_k P_k A_k^\top + B_k B_k^\top - A_k P_k A_k^\top + B_k B_k^\top
\]

\[
\times (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} (A_k P_k A_k^\top + B_k B_k^\top)
\]

(72)

On the other hand, for the right-hand side of (70), it holds

\[
\tilde{A}_k \tilde{P}_k \tilde{A}_k^\top = (I - B_k B_k^\top \Phi_{k+1} G_{r,k}^\top \Phi_{k+1})
\]

\[
\times (A_k P_k A_k^\top - A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} (A_k P_k A_k^\top + B_k B_k^\top)
\]

\[
\times (I - B_k B_k^\top \Phi_{k+1} G_{r,k}^\top \Phi_{k+1})^\top
\]

\[
= A_k P_k A_k^\top - A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} (A_k P_k A_k^\top + B_k B_k^\top
\]

\[
- L_1 - L_2 + L_2
\]

(73)
where
\[
L_1 := B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top + (I - (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1}A_k P_k A_k^\top) \times \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top \\
L_2 := B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k+1 \\
\times (A_k P_k A_k^\top - A_k Q_k A_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k P_k A_k^\top) \\
\times \Phi_k^\top G_{r,k}^\top \Phi_k+1 B_kB_k^\top.
\]

Here, we have
\[
L_1 = B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k G_{r,k}^\top \Phi_k^\top \Phi_k A_k^\top \\
\times (I - (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1}A_k P_k A_k^\top) \\
- B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top \\
\times (I - (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1}A_k P_k A_k^\top) \\
= B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k P_k A_k^\top.
\]

Moreover, it follows from
\[
A_k Q_k A_k^\top = \Phi(k + 1, N)G_{r,k} \Phi(k + 1, N)^\top
\]
that
\[
L_1 = B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top \\
\times (I - (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1}A_k P_k A_k^\top) \\
- B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top \\
\times (I - (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1}A_k P_k A_k^\top) \\
= B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k A_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k P_k A_k^\top.
\]

It is known that if ran(Y^\top) \supseteq ran(X^\top), then Y = X X^\top Y \{41\}, Th. 3.11\}, and we have
\[
\text{ran}(G_{r,k}) = \text{ran}((\Phi_k + B_k \Phi_k + B_{k+2} + \cdots + B_{N-1}) \\
\supseteq \text{ran}(\Phi_k + B_k) = \text{ran}(\Phi_k + B_k B_k^\top \Phi_k^\top).
\]

Hence, we have
\[
B_kB_k^\top \Phi_k+1 \Phi_k^\top G_{r,k} \Phi_k = B_kB_k^\top \Phi_k^\top
\]
which yields
\[
L_1 = B_kB_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k P_k A_k^\top.
\]

By a similar argument,
\[
L_2 := B_kB_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k+1 A_k Q_k A_k^\top \\
\times (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k P_k A_k^\top \Phi_k^\top G_{r,k}^\top \Phi_k+1 B_kB_k^\top \\
= B_kB_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k P_k A_k^\top \\
\times \Phi_k^\top G_{r,k}^\top \Phi_k+1 B_kB_k^\top \\
= B_kB_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} (A_k P_k A_k^\top + A_k Q_k A_k^\top) \\
\times \Phi_k^\top G_{r,k}^\top \Phi_k+1 B_kB_k^\top \\
- B_kB_k^\top (A_k P_k A_k^\top + A_k Q_k A_k^\top)^{-1} A_k Q_k A_k^\top.
\]

By (72), (73), (76), and (77), we obtain (70), and thus, P^*_k = \tilde{P}_k.
Next, we show P^*(k, s) = \tilde{P}(k, s) for k \leq s. Denote by \tilde{P}(k, l) the state-transition matrix for \tilde{A}_k as in (3). Then,
\[
P^*(k, s) = \mathbb{E} \left[ x_k^* \left( \tilde{\Phi}(s, k) \phi_k + \sum_{l=k}^{s-1} \tilde{\Phi}(s, l+1) B_{l+1} \phi_l^* \right) \right]
\]
\[
= P^*(k, k) \tilde{\Phi}(s, k)^\top, k \leq s.
\]

Therefore, it holds
\[
P^*(k, s + 1) = P^*(k, s) \tilde{A}_k^\top, k \leq s.
\]

In what follows, we show that the same relationship
\[
\tilde{P}(k, s + 1) = \tilde{P}(k, s) \tilde{A}_k^\top, k \leq s
\]
also holds. The left-hand side of (79) is written as
\[
\tilde{P}(k, s + 1) = P_k \Phi(s + 1, k) \tilde{A}_k^\top - P_k \Phi(s + 1, k) \tilde{A}_k^\top \\
- P_k \Phi(s + 1, k) \tilde{A}_k^\top (A_s P_s A_s^\top + A_s Q_s A_s^\top)^{-1} B_s B_s^\top \\
\]
where in the last line we used (71). On the other hand, the right-hand side of (79) is
\[
\tilde{P}(k, s) \tilde{A}_k^\top = \left( P_k \Phi(s, k) \tilde{A}_k^\top - P_k \Phi(s, k) \tilde{A}_k^\top \right) A_s^\top \\
\times (I - \phi_{k+1} G_{r,s}^\top \phi_{s+1} B_{s+1}^\top) \\
= P_k \Phi(s + 1, k)^\top - P_k \Phi(s + 1, k)^\top (A_s P_s A_s^\top + A_s Q_s A_s^\top)^{-1} B_s B_s^\top
\]
where
\[
L_2 := P_k \Phi(s + 1, k)^\top (A_s P_s A_s^\top + A_s Q_s A_s^\top)^{-1} A_s P_s A_s^\top \\
\times \phi_{k+1} G_{r,s}^\top \phi_{s+1} B_{s+1}^\top \\
= P_k \Phi(s + 1, k)^\top (A_s P_s A_s^\top + A_s Q_s A_s^\top)^{-1} B_s B_s^\top \\
- P_k \Phi(s + 1, k)^\top (A_s P_s A_s^\top + A_s Q_s A_s^\top)^{-1} B_s B_s^\top
\]
By (80) and (81), we obtain (79). Finally, by combining (78) and (79) with P^*_k = \tilde{P}_k, we arrive at P^*(k, s) = \tilde{P}(k, s) for k \leq s, which completes the proof.
B. Proof of Theorem 2

Here, we show (53). First, by using (71), we obtain
\[
\ell_{k+1} = \Phi(k + 1, 0)\bar{x}_0 + (A_kP_kA_k^T + B_kB_k^*)A_k^{-1}
\times (P_k + Q_k)^{-1}\Phi(k, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[= \Phi(k + 1, 0)\bar{x}_0 + A_kP_k(P_k + Q_k)^{-1}\Phi(k, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[+ B_kB_k^*(A_kP_kA_k^T + A_kQ_kA_k^*)^{-1}\Phi(k + 1, N)
\times (\bar{x}_N - \Phi(N, 0)\bar{x}_0).
\]
(82)

On the other hand, for the right-hand side of (53), it holds
\[
\hat{A}_k\ell_k = \Phi(k + 1, 0)\bar{x}_0 - B_kB_k^*\Phi(N, k + 1)G_t(N, k)^T
\times \Phi(N, 0)\bar{x}_0 + L_4 + L_5
\]
(83)

where
\[
L_4 := A_kP_k\Phi(N, k)^T\Sigma_0^{-1}(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[
L_5 := -B_kB_k^*\Phi(N, k)^T\Phi(N, k + 1)G_t(N, k)^T
\times A_kP_k(P_k + Q_k)^{-1}\Phi(k, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0).
\]
(84)

Here, we have
\[
L_4 = A_kP_k(P_k + Q_k)^{-1}\Phi(k, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[= -B_kB_k^*\Phi(N, k)^T\Phi(N, k + 1)G_t(N, k)^T
\times A_kP_k(P_k + Q_k)^{-1}\Phi(k + 1, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[
= -B_kB_k^*\Phi(N, k)^T\Phi(N, k + 1)G_t(N, k)^T
\times A_kP_k(P_k + Q_k)^{-1}\Phi(k + 1, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[= -B_kB_k^*\Phi(N, k)^T\Phi(N, k + 1)G_t(N, k)^T
\times A_kP_k(P_k + Q_k)^{-1}\Phi(k + 1, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
\[= -B_kB_k^*\Phi(N, k)^T\Phi(N, k + 1)G_t(N, k)^T
\times A_kP_k(P_k + Q_k)^{-1}\Phi(k + 1, N)(\bar{x}_N - \Phi(N, 0)\bar{x}_0)
\]
(85)

where in the last line we used (74) and (75). By (82)–(85), we get (53).

APPENDIX D

PROOF OF THEOREM 3

By expression (55), it suffices to prove the following two statements.

i) Consider the problem of minimizing $D_{KL}(\hat{P}_{0,N}||\bar{P}_{0,N})$ with respect to $\bar{P}_{0,N}$ under the constraint that the marginals of $\bar{P}_{0,N}$ at $k = 0$ and $k = N$ are equal to $N(0, \Sigma_0)$ and $N(0, \Sigma_N)$, respectively. Then, the joint distribution $\bar{P}_{0,N}$ of $(x_0, x_N)$ is an optimal solution to this problem.

ii) The conditional distributions of $(x_k^*)$ and $(\bar{x}_k)$ satisfy
\[
\bar{P}_{1:N-1|0,N}(\cdot|x_0^* = \bar{x}_0, x_N^* = \bar{x}_N) = \bar{P}_{1:N-1|0,N}(\cdot|x_0 = \bar{x}_0, x_N = \bar{x}_N) \quad \forall \bar{x}_0, \bar{x}_N \in \mathbb{R}^n.
\]
(86)

Proof of part i): For brevity, we denote $N_{\Sigma}(\cdot) := N(\cdot|0, \Sigma)$. Note that for any $\Sigma, \Xi > 0$, the KL divergence between $n$-variate Gaussian distributions $N_{\Sigma_1}, N_{\Xi}$ is given by
\[
D_{KL}(N_{\Sigma}||N_{\Xi}) = \frac{1}{2}(\log \det(\Xi) - \log \det(\Sigma) + \text{tr}(\Xi^{-1}\Sigma) - n).
\]
(87)

Now, let $p_{\Sigma}$ be the density function of a distribution with covariance matrix $\Sigma$, which is not necessarily Gaussian. Then, it is known that given covariance matrices $\Sigma$ and $\Xi$, the density $p_{\Sigma}$ that minimizes $D_{KL}(p_{\Sigma}||p_{\Xi})$ is the Gaussian distribution $N_{\Sigma}$; see the proof of [9, Th. 11].

Now, we introduce
\[
\Sigma_{0,0} := \begin{bmatrix} \Sigma_0 & Y^T \\ Y & \Sigma_N \end{bmatrix}, \quad Y \in \mathbb{R}^{n \times n}
\]
\[
\Xi_{0,0} := \begin{bmatrix} \Xi & \Sigma_0 \Phi(N, 0)^T \\ \Phi(N, 0)\Sigma_0 & \Xi_N \end{bmatrix}
\]
\[\Xi_0 := \Sigma_0, \Xi_k := \Phi(k, 0)\Sigma_0\Phi(k, 0)^T + G_t(k, 0), \quad k \in [1, N].
\]

Then, $\Xi_{0,0}$ is the covariance of the Gaussian distribution $\bar{P}_{0,N}$. In the light of the above description, the minimization of $D_{KL}(\bar{P}_{0,N}||\bar{P}_{0,N})$ is equivalent to the minimization of $D_{KL}(N_{\Sigma_{0,0}}||N_{\Xi_{0,0}})$ with respect to $Y$. By (87), this is further equivalent to the minimization of
\[
\text{tr}(\Xi_{0,0}^{-1}\Sigma_{0,0}) - \log \det(\Sigma_{0,0}),
\]
(88)

with respect to $Y$. Noting that
\[
\Xi_{0,0}^{-1} = \begin{bmatrix} \Sigma_{0,0}^{-1} + \Phi_0^T G_{t,0}^{-1}\Phi_0 & -\Phi_0^T G_{t,0}^{-1} \\ -G_{t,0}^{-1}\Phi_0^T & G_{t,0}^{-1} \end{bmatrix}
\]
we have
\[
\Xi_{0,0}^{-1} = \begin{bmatrix} -\Sigma_{0,0}^{-1} + \Phi_0^T G_{t,0}^{-1}\Phi_0 & -\Phi_0^T G_{t,0}^{-1} \\ -G_{t,0}^{-1}\Phi_0^T & G_{t,0}^{-1} \end{bmatrix}
\]
where we introduced the abbreviations $\Phi_0 := \Phi(N, 0)$ and $G_{t,0} := G_t(N, 0)$. Then, (88) is written as
\[
\text{tr}\left((\Sigma_{0,0}^{-1} + \Phi_0^T G_{t,0}^{-1}\Phi_0)\Sigma_0 - \Phi_0^T G_{t,0}^{-1}Y - Y^T G_{t,0}^{-1}\Phi_0
\]
\[+ G_{t,0}^{-1}\Sigma_N\right) - \log \det(\Sigma_0) - \log \det(\Sigma_N - Y\Sigma_N^{-1}Y^T)
\]
In what follows, we consider maximizing
\[
f(Y) := \log \det(\Sigma_N - Y\Sigma_N^{-1}Y^T) + 2\text{tr}(\Phi_0^T G_{t,0}^{-1}Y)
\]
(89)
which is concave. By taking the derivative of $f$, the necessary and sufficient condition for $Y$ to be the maximizer of $f(Y)$ is obtained as
\[
-2\Sigma_N^{-1}Y^T(\Sigma_N - Y\Sigma_N^{-1}Y^T)^{-1} + 2\Phi_0^T G_{t,0}^{-1} = 0.
\]
(90)
For notational simplicity, we omit the subscript “$-$” of $Q_{-k}$. Define the state-transition matrix $Q_{k}(k, l)$ for $A_{Q,k}$ like in (3). Then, $\bar{P}_{0,N}$ is Gaussian with covariance matrix
\[
\begin{bmatrix} \Sigma_0 & \Sigma_0\Phi(Q(N, 0))^T \\ \Phi_0(Q(N, 0))^T & \Sigma_N \end{bmatrix}
\].
Hence, it suffices to show that $Y = \Phi Q(N,0)\Sigma_0$ satisfies the condition \((90)\), which is equivalent to
\[
\Phi Q(N,0)^\top (\Sigma_N - \Phi Q(N,0)\Sigma_0\Phi Q(N,0)^\top)^{-1} \\
= \Phi(N,0)^\top G_r(N,0)^{-1} \\
= \Phi(N,0)^\top (\Xi_N - \Phi(N,0)\Sigma_0\Phi(N,0)^\top)^{-1}.
\]
Taking the inverse of both sides of \((91)\) yields
\[
\Sigma_N\Phi Q(0,N)^\top - \Phi Q(N,0)\Sigma_0 = \Xi_N\Phi(0,N)^\top - \Phi(0,N)\Sigma_0.
\]
Hereafter, we show that
\[
\Sigma_k\Phi Q(0,k)^\top - \Phi Q(k,0)\Sigma_0 = \Xi_k\Phi(0,k)^\top - \Phi(k,0)\Sigma_0 \\
\forall k \in [0, N]
\]
where $\Sigma_k = \mathbb{E}[x_k^\top(x_k^\top)]^\top$, that is,
\[
\Sigma_{k+1} = A_{k+1}\Sigma_k A_{k+1}^\top + B_k(I + B_k^\top Q_{k+1}^{-1}B_k)^{-1}B_k^\top.
\]
For this purpose, we define
\[
F_{1,k} := \Sigma_k\Phi Q(0,k)^\top - \Phi Q(k,0)\Sigma_0 \\
F_{2,k} := \Xi_k\Phi(0,k)^\top - \Phi(k,0)\Sigma_0 \\
F_{3,k} := \Phi Q(0,k)^\top - \Phi(0,k)^\top
\]
and show $F_{2,k} = F_{3,k}$ and $F_{1,k} = F_{3,k}$, which imply \((92)\).

First, noting that $F_{2,0} = F_{3,0}$, we check that $F_{2,k}$ and $F_{3,k}$ satisfy the same difference equation, which means $F_{2,k} = F_{3,k}$ for all $k \in [0, N]$. For $F_{2,k}$, we have
\[
F_{2,k+1} = (\Phi(k+1,0)\Sigma_0\Phi(k+1,0)^\top + G_r(k+1,0)) \\
\times \Phi(0,k+1)^\top - \Phi(k+1,0)\Sigma_0 \\
= G_r(k+1,0)\Phi(0,k+1)^\top \\
= (A_k G_r(k,0)A_k^\top + B_k B_k^\top)\Phi(0,k+1)^\top \\
= A_k F_{2,k} + B_k B_k^\top\Phi(0,k+1)^\top.
\]
On the other hand, for $F_{3,k}$, it holds
\[
F_{3,k+1} = Q_{k+1}\Phi Q(0,k+1)^\top - (A_k Q_k A_k^\top - B_k B_k^\top) \\
\times \Phi(0,k+1)^\top \\
= (Q_{k+1} + B_k B_k^\top)A_k^\top\Phi Q(0,k)^\top - A_k Q_k\Phi(0,k)^\top \\
+ B_k B_k^\top\Phi(0,k+1)^\top \\
= A_k F_{3,k} + B_k B_k^\top\Phi(0,k+1)^\top
\]
where we used $A_k Q_k = (I + B_k B_k^\top Q_{k+1}^{-1})^{-1}A_k$. Hence, we obtain $F_{2,k} = F_{3,k}$.

Next, define $M_k := Q_k^{-1}(F_{3,k} - F_{1,k})$. Then,
\[
M_{k+1} = (I - Q_k^{-1}\Sigma_k)\Phi Q(0,k+1)^\top \\
+ Q_k^{-1}\Phi Q(0,k+1)^\top - \Phi(0,k+1)^\top.
\]
The first term is written as
\[
(I - Q_k^{-1}\Sigma_k)\Phi Q(0,k+1)^\top \\
= (A_k^{-1} - ((Q_{k+1} + B_k B_k^\top)^{-1}A_k)\Sigma_k
\]
\[
+ Q_{k+1}^{-1}B_k(I + B_k^\top Q_{k+1}^{-1}B_k)^{-1}B_k^\top A_k^{-1})\Phi Q(0,k)^\top.
\]
Here, noting that
\[
A_k^{-1} = (Q_{k+1} - B_k(I + B_k^\top Q_{k+1}^{-1}B_k)^{-1}B_k^\top)^{-1}Q_{k+1}A_k^{-1}
\]
we have
\[
Q_{k+1}^{-1}B_k(I + B_k^\top Q_{k+1}^{-1}B_k)^{-1}B_k^{-1}A_k^{-1} = -A_k^{-1} + A_k^{-1}.
\]
Thus, \((95)\) is written as
\[
(I - Q_{k+1}^{-1}\Sigma_k)\Phi Q(0,k+1)^\top \\
= (-A_k Q_k^{-1}A_k^{-1} - A_k^{-1})\Phi Q(0,k)^\top \\
= A_k^{-1}(I - Q_k^{-1}\Sigma_k)\Phi Q(0,k)^\top.
\]
For the second term of the right-hand side of \((94)\), we have
\[
Q_{k+1}^{-1}Q_k(I + B_k B_k^\top Q_{k+1}^{-1}B_k)^{-1}A_k^{-1} \Phi Q(0,k)^\top \\
= A_k^{-1}(I - Q_k^{-1}\Sigma_k)\Phi Q(0,k)^\top.
\]
It follows from \((94)\), \((96)\), and \((97)\) that
\[
M_{k+1} = A_k^{-1}M_k.
\]
Combining this with $M_0 = 0$, we obtain $M_k = 0$ for all $k \in [0, N]$, that is, $F_{1,k} = F_{3,k}$ for all $k \in [0, N]$. In summary, we get \((92)\), which implies \(i\).

**Proof of part ii):** By Theorem 2, it suffices to show that for any end points $x_0, x_N \in \mathbb{R}^n$, the pinned process associated with \((50)\) coincides with the one associated with \((56)\). In other words, we verify that for any $k \in [0, N - 1]$,
\[
B_k B_k^\top\Phi(N, k + 1)^\top G_r(N, k)^\dagger \\
= B_k B_k^\top\Phi Q(N, k + 1)^\top G_r(N, k)^\dagger \\
\Lambda_k := (I - B_k B_k^\top)\Phi Q(N, k + 1)^\top G_r(N, k)^\dagger \\
\times \Phi Q(N, k + 1)^\top B_k B_k^\top \\
= B_k B_k^\top\Phi Q(N, k + 1)^\top G_r(N, k)^\dagger \\
\times \Phi Q(N, k + 1)^\top B_k B_k^\top =: \Lambda_k.
\]
Here, $G_r Q_x$ are the reachability Gramian and the controllability Gramian where $\Phi$ and $B_k$ in \((20)\) and \((21)\) are replaced by $\Phi Q$ and $B Q_k$, respectively.

First, note that
\[
\Phi(N, k + 1)B_k B_k^\top\Phi(N, k + 1)^\top G_r(N, k)^\dagger \\
= \Phi(N, k + 1)B_k B_k^\top A_k^{-1}(G_r(N, k)^\dagger)^\dagger.
\]
Indeed, since it holds $\text{ran}(G_r(N, k)) \supseteq \text{ran}(\Phi(N, k + 1)B_k B_k^\top\Phi(N, k + 1)^\top)$, there exists $W \in \mathbb{R}^{n \times n}$ such that
\[
\Phi(N, k + 1)B_k B_k^\top\Phi(N, k + 1)^\top = G_r(N, k)W \\
= W^\top G_r(N, k)
\]
and we have
\[
\Phi(N, k + 1)B_kB_k^\top \Phi(N, k + 1)^\top G_r(N, k) + W^\top G_r(N, k)G_r(N, k)^\dagger
\]
\[
= W^\top G_r(N, k)G_r(N, k)^\dagger
\]
\[
\Phi(N, k + 1)B_kB_k^\top A_k^\top \left( G_r(N, k)\Phi(N, k)^\top \right)^\dagger
\]
\[
= W^\top G_r(N, k)\Phi(N, k)^\top \left( G_r(N, k)\Phi(N, k)^\top \right)^\dagger.
\]
Here, \( G_r(N, k)G_r(N, k)^\dagger \) is the orthogonal projection matrix onto \( \text{ran}(G_r(N, k)) \), and \( G_r(N, k)\Phi(N, k)^\top \left( G_r(N, k)\Phi(N, k)^\top \right)^\dagger \) is also the orthogonal projection onto \( \text{ran}(G_r(N, k)\Phi(N, k)^\top) = \text{ran}(G_r(N, k)) \). Thus, by the uniqueness of the orthogonal projection, we obtain (101).

By using (22) and (101), the left-hand side of (98) is equal to
\[
B_kB_k^\top A_k^\top \left( \Phi(N, k)G_c(N, k) \right)^\dagger.
\]
Similarly, the right-hand side of (98) is written as
\[
B_{Q,k}B_{Q,k}^\top A_{Q,k}^\top \left( \Phi_Q(N, k)G_Qc(N, k) \right)^\dagger
\]
where we used \( B_{Q,k}B_{Q,k}^\top A_{Q,k}^\top = B_kB_k^\top A_k^\top \). For notational simplicity, we write \( R_{1,k} := G_c(N, k), R_{2,k} := G_Qc(N, k) \) and set \( R_{1,N} = R_{2,N} = 0 \). By (102) and (103), to verify (98), it suffices to check that
\[
\Phi(N, k)R_{1,k} - \Phi_Q(N, k)R_{2,k} = 0.
\]
Since it holds
\[
A_{Q,k} = (I + B_kB_k^\top Q_k^{-1})^{-1}A_k = Q_kA_k^{-1}Q_k^{-1}
\]
we have \( \Phi_Q(N, k) = Q_k\Phi(N, k)^\top Q_k^{-1} \), and therefore,
\[
\Phi(N, k)R_{1,k} - \Phi_Q(N, k)R_{2,k}
\]
\[
= \Phi(N, k) \left( R_{1,k} - \Phi(N, k)^\top Q_k^{-1}R_{2,k} \right).
\]
Thus, (104) is equivalent to
\[
J_k := R_{1,k} + R_{1,k}Q_k^{-1}R_{2,k} - R_{2,k} = 0.
\]
Hereafter, we show that
\[
J_{k+1} = A_kJ_kA_{Q,k}, \; k \in [0, N - 1].
\]
The above relationship with \( J_N = 0 \) and the invertibility of \( A_k \) and \( A_{Q,k} \) imply (104), which concludes (98). For the left-hand side of (105), it holds
\[
J_{k+1} = A_kR_{1,k}A_k^\top - B_kB_k^\top + A_kR_{1,k}Q_k^{-1}R_{2,k}A_{Q,k}
\]
\[
- A_kR_{1,k}Q_k^{-1}A_k^{-1}B_kB_k^\top - B_kB_k^\top A_k^{-1}Q_k^{-1}R_{2,k}A_{Q,k}
\]
\[
+ B_kB_k^\top A_k^{-1}Q_k^{-1}A_k^\top B_kB_k^\top - A_kR_{2,k}A_{Q,k} + B_{Q,k}B_{Q,k}^\top.
\]
Here, we have
\[
B_kB_k^\top A_k^{-1}Q_k^{-1}R_{2,k}A_{Q,k} = B_{Q,k}B_{Q,k}^\top A_{Q,k}^{-1}Q_k^{-1}R_{2,k}A_{Q,k}
\]
\[
= B_{Q,k}B_{Q,k}^\top Q_k^{-1}A_kR_{2,k}A_{Q,k}^{-1}k
\]
\[
= (A_k - A_{Q,k})R_{2,k}A_{Q,k}^{-1}k
\]
where in the last line we used \( A_{Q,k} = A_k - B_{Q,k}B_{Q,k}^\top \).

On the other hand,
\[
B_kB_k^\top A_k^{-1}Q_k^{-1}B_kB_k^\top
\]
\[
= B_kB_k^\top (Q_k^{-1} + B_kB_k^\top)^{-1}B_kB_k^\top
\]
\[
= -B_k(I - B_k^\top Q_k^{-1}B_k^{-1})B_k^\top
\]
\[
= B_kB_k^\top - B_k(I + B_k^\top Q_k^{-1}B_k^{-1})B_k^\top
\]
\[
= B_kB_k^\top - B_kQ_k^{-1}B_k^\top.
\]

Moreover,
\[
A_kR_{1,k}A_k^{-1}Q_k^{-1}A_k^\top B_kB_k^\top
\]
\[
= A_kR_{1,k}A_k^{-1} - A_kR_{1,k}Q_k^{-1}A_k^\top B_{Q,k}B_{Q,k}^\top
\]
\[
= A_kR_{1,k}A_k^{-1}_k - A_kR_{1,k}A_k^{-1}_kB_{Q,k}B_{Q,k}^\top
\]
\[
= A_kR_{1,k}A_k^{-1}_k.
\]

By (106), (107), (108), and (109), it holds
\[
J_{k+1} = A_kR_{1,k}A_k^{-1}Q_k^{-1}R_{2,k}A_{Q,k} - A_kR_{2,k}A_{Q,k}
\]
\[
= A_kJ_kA_{Q,k}.
\]

Hence, we come to (98).

By using (98), we get (99) as follows:
\[
\hat{A}_{Q,k} = \left[I - B_kB_k^\top A_k^{-1} \left( \Phi(N, k)G_c(N, k) \right)^\dagger \right]A_{Q,k}
\]
\[
= A_k - B_{Q,k}B_{Q,k}^\top Q_k^{-1}_kA_k - B_kB_k^\top A_k^{-1}_k
\]
\[
\times \left( \Phi(N, k)G_c(N, k) \right)^\dagger \Phi(N, k)^\top Q_k^{-1}_k \Phi(N, k)^\top Q_k^{-1}_k
\]
\[
= A_k - B_{Q,k}B_{Q,k}^\top Q_k^{-1}_kA_k - B_kB_k^\top A_k^{-1}_k
\]
\[
\times \left( \Phi(N, k)G_c(N, k) \right)^\dagger \Phi(N, k) \left(I - G_c(N, k)Q_k^{-1}_k \right)
\]
\[
= A_k - B_{Q,k}B_{Q,k}^\top Q_k^{-1}_kA_k - B_kB_k^\top A_k^{-1}_k
\]
\[
\times \left( \Phi(N, k)G_c(N, k) \right)^\dagger \Phi(N, k) + B_kB_k^\top A_k^{-1}_k
\]
\[
= \hat{A}_k.
\]

Finally, by using (98) and (108), we show (100) as
\[
A_{Q,k} = B_{Q,k}B_{Q,k}^\top - B_kB_k^\top \left( \Phi(N, k)G_c(N, k) \right)^\dagger
\]
\[
\times \Phi_Q(N, k + 1)B_{Q,k}B_{Q,k}^\top
\]
\[
= B_kB_k^\top - B_kB_k^\top A_k^{-1}_kB_kB_k^\top - B_kB_k^\top A_k^{-1}_k
\]
\[
\times \left( \Phi(N, k)G_c(N, k) \right)^\dagger \Phi(N, k) \left(I - G_c(N, k)Q_k^{-1}_k \right)
\]
\[
\times A_k^{-1}_kB_kB_k^\top
\]
\[
= B_kB_k^\top
\]
\[
- B_kB_k^\top A_k^{-1}_k \left( \Phi(N, k)G_c(N, k) \right)^\dagger \Phi(N, k + 1)B_kB_k^\top
\]
\[
= \hat{A}_k.
\]

In summary, we obtain ii). The expression (55) with i) and ii) completes the proof.
REFERENCES

[1] E. Todorov, “Efficient computation of optimal actions,” in Proc. Nat. Acad. Sci., vol. 106, no. 28, pp. 11478–11483, 2009.

[2] É. Garrabé and G. Russo, “Probabilistic design of optimal sequential decision-making algorithms in learning and control,” Ann. Rev. Control, vol. 54, pp. 81–102, 2022.

[3] K. Ito and K. Kashima, “Kallback–Leibler control for discrete-time nonlinear systems on continuous spaces,” SICE J. Control, Meas., Syst. Integration, vol. 15, no. 2, pp. 119–129, 2022.

[4] T. Haarnoja, H. Tang, P. Abbeel, and S. Levine, “Reinforcement learning with deep energy-based policies,” in Proc. Int. Conf. Mach. Learn., 2017, pp. 1352–1361.

[5] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in Proc. Int. Conf. Mach. Learn., 2018, pp. 1861–1870.

[6] S. Levine, “Reinforcement learning and control as probabilistic inference: Tutorial and review,” 2018, arXiv:1805.09099.

[7] J. Ho and S. Ermon, “Generative adversarial imitation learning,” in Proc. Neural Inf. Process. Syst., 2016.

[8] B. Eybenbach and S. Levine, “Maximum entropy RL (provably) solves some robust RL problems,” in Proc. Int. Conf. Learn. Representations, 2022. [Online]. Available: https://openreview.net/forum?id=P5SAD3cuaA2

[9] Y. Chen, T. T. Georgiou, and M. Pavon, “Optimal steering of a linear stochastic system to a final probability distribution, part I,” IEEE Trans. Autom. Control, vol. 61, no. 5, pp. 1158–1169, May 2016.

[10] E. Bakolas, “Sur la théorie relativiste de l’électron et l’interprétation de la mécanique quantique,” Annales de l’institut Henri Poincaré, vol. 2, no. 4, pp. 269–310, 1932.

[11] G. Peyré and M. Cuturi, “Computational optimal transport: With applications,” Found. Trends Mach. Learn., vol. 11, no. 5/6, pp. 355–607, 2019.

[12] Y. Chen, T. T. Georgiou, and M. Pavon, “On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint,” J. Optim. Theory Appl., vol. 169, no. 2, pp. 671–691, 2016.

[13] A. Beghi, “Continuous-time Gauss–Markov processes with fixed reciprocal dynamics,” J. Math. Syst. Estimation Control, vol. 7, no. 3, pp. 343–366, 1997.

[14] A. Wolkolbinger, “A simplified variational characterization of Schrödinger processes,” J. Math. Phys., vol. 30, no. 12, pp. 2943–2946, 1989.

[15] A. Blaquiere, “Controllability of a Fokker–Planck equation, the Schrödinger system, and a related stochastic optimal control (revised version),” Dyn. Control, vol. 2, no. 3, pp. 235–253, 1992.

[16] P. Dai Pra, “A stochastic control approach to reciprocal diffusion processes,” Appl. Math. Optim., vol. 23, no. 1, pp. 313–329, 1991.

[17] P. Dai Pra and M. Pavon, “On the Markov processes of Schrödinger, the Feynman–Kac formula and stochastic control,” in Realization and Modeling in System Theory. Boston, MA, USA: Birkhäuser, 1990, pp. 497–504.

[18] Y. Chen, T. T. Georgiou, M. Pavon, and A. Tannenbaum, “Robust transport over networks,” IEEE Trans. Autom. Control, vol. 62, no. 9, pp. 4675–4682, Sep. 2017.

[19] E. Bakolas, “Optimal covariance control for discrete-time stochastic linear systems subject to constraints,” in Proc. IEEE 55th Conf. Decis. Control, 2016, pp. 1153–1158.

[20] E. Bakolas, “Finite-horizon covariance control for discrete-time stochastic linear systems subject to input constraints,” Automatica, vol. 91, pp. 61–68, 2018.

[21] M. Goldstein and P. Tsioptas, “Finite-horizon covariance control of linear time-varying systems,” in Proc. IEEE 56th Ann. Conf. Decis. Control, 2017, pp. 3606–3611.

[22] A. Beghi, “On the relative entropy of discrete-time Markov processes with given end-point densities,” IEEE Trans. Inf. Theory, vol. 42, no. 5, pp. 1529–1535, Sep. 1996.

[23] Z. Xi, Z. Cao, E. Theodorou, and Y. Chen, “Nonlinear covariance control via dynamic programming,” in Proc. IEEE Amer. Control Conf., 2020, pp. 3571–3576.

[24] K. F. Caluya and A. Halder, “Wasserstein proximal algorithms for the Schrödinger bridge problem: Density control with nonlinear drift,” IEEE Trans. Autom. Control, vol. 67, no. 3, pp. 1163–1178, Mar. 2022.

[25] H. Yu, Z. Chen, and Y. Chen, “Covariance steering for nonlinear control-affine systems,” 2021, arXiv:2108.09530.

[26] E. Bakolas and A. Tsolovikos, “Greedy finite-horizon covariance steering for discrete-time stochastic nonlinear systems based on the unscented transform,” in Proc. IEEE Amer. Control Conf., 2020, pp. 5595–3600.

[27] K. Elamvazuthi, P. Grover, and S. Berman, “Optimal transport over deterministic discrete-time nonlinear systems using stochastic feedback laws,” IEEE Control Syst. Lett., vol. 3, no. 1, pp. 168–173, Jan. 2019.

[28] K. Okamoto, M. Goldstein, and P. Tsioptas, “Optimal covariance control for stochastic systems under chance constraints,” IEEE Control Syst. Lett., vol. 2, no. 2, pp. 266–271, Apr. 2018.

[29] A. Halder and E. D. B. Wendel, “Finite horizon linear quadratic Gaussian density regulator with Wasserstein terminal cost,” in Proc. IEEE Amer. Control Conf., 2016, pp. 7249–7254.

[30] I. M. Balci and E. Bakolas, “Covariance steering of discrete-time stochastic linear systems based on Wasserstein distance terminal cost,” IEEE Control Syst. Lett., vol. 5, no. 6, pp. 2000–2005, Dec. 2021.

[31] Y. Chen, T. T. Georgiou, and M. Pavon, “Optimal transport in systems and control,” Ann. Rev. Control Robot., Autom. Syst., vol. 4, pp. 89–113, 2021.

[32] Y. Chen and T. Georgiou, “Stochastic bridges of linear systems,” IEEE Trans. Autom. Control, vol. 61, no. 2, pp. 526–531, Feb. 2016.

[33] F. L. Lewis, D. Vrabie, and V. L. Syrmos, Optimal Control, Hoboken, NJ, USA: Wiley, 2012.

[34] T. M. Cover and J. A. Thomas, Elements of Information Theory, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[35] Y. Chen, T. T. Georgiou, and M. Pavon, “Optimal steering of a linear stochastic system to a final probability distribution—Part III,” IEEE Trans. Autom. Control, vol. 63, no. 9, pp. 3112–3118, Sep. 2018.

[36] A. Ben-Israel and T. N. Greville, Generalized Inverses: Theory and Applications, vol. 15, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.

[37] C. Léonard, “A survey of the Schrödinger problem and some of its connections with optimal transport,” Discrete Continuous Dynamical Syst., vol. 34, no. 4, pp. 1533–1574, 2014.

[38] Y. Chen, T. T. Georgiou, and M. Pavon, “Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge,” SIAM Rev., vol. 63, no. 2, pp. 249–313, 2021.

[39] C. Léonard, “Some Properties of Path Measures,” in Séminaire De Probabilités XLVI. Berlin, Germany: Springer, 2014, pp. 207–230.

[40] Y. Chen, T. T. Georgiou, and M. Pavon, “Optimal steering of a linear stochastic system to a final probability distribution, part II,” IEEE Trans. Autom. Control, vol. 61, no. 5, pp. 1170–1180, May 2016.

[41] H. Yanai, K. Takeuchi, and Y. Takane, Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition. Berlin, Germany: Springer, 2011.

Kaito Ito (Member, IEEE) received the bachelor’s degree in engineering, and the master’s and doctoral degrees in informatics from Kyoto University, Kyoto, Japan, in 2017, 2019, and 2022, respectively.

He is currently an Assistant Professor with the School of Computing, Tokyo Institute of Technology, Yokohama, Japan. His research interests include stochastic control, large-scale systems, and statistical learning.

Kenji Kashima (Senior Member, IEEE) received the doctoral degree in informatics from Kyoto University, Kyoto, Japan, in 2005.

He was with Tokyo Institute of Technology, Tokyo, Japan, Universität Stuttgart, Stuttgart, Germany, Osaka University, Osaka, Japan. In 2013, he joined Kyoto University, where he is currently an Associate Professor. His research interests include control and learning theory for complex (large scale, stochastic, and networked) dynamical systems, as well as its applications.

Dr. Kashima was an Associate Editor for IEEE TRANSACTIONS OF AUTOMATIC CONTROL, IEEE CONTROL SYSTEMS LETTERS, IEEE CSS CONFERENCE Editorial Board, and Asian Journal of Control.