Pathwise turnpike and dissipativity results for discrete-time stochastic linear-quadratic optimal control problems

Jonas Schießl, Ruchuan Ou, Timm Faulwasser, Michael Heinrich Baumann, and Lars Grüne

Abstract— We investigate pathwise turnpike behavior of discrete-time stochastic linear-quadratic optimal control problems. Our analysis is based on a novel strict dissipativity notion for such problems, in which a stationary stochastic process replaces the optimal steady state of the deterministic setting. The analytical findings are illustrated by a numerical example.

I. INTRODUCTION

Turnpike properties are a valuable feature of optimal control problems (OCPs). In deterministic problems, the most common variant of this phenomenon describes that for increasing horizon lengths, the optimal trajectories spend most of their time near an optimal steady state. This means that the turnpike property can be seen as a finite-horizon variant of the usual asymptotic stability property. It was first observed by von Neumann and Ramsey, see [19], [22], but it is still a topic of recent research, particularly due to its relation to model predictive control, see [3], [9]. A closely related concept is strict dissipativity. For deterministic problems the connection between strict dissipativity and the turnpike property is well studied, see [4], [8], [10].

While turnpike phenomena are rather well understood in the deterministic setting, more theoretical work is needed when uncertainties enter the dynamics and the setting becomes stochastic. Some of the challenges here are the definition of the stochastic counterpart to the optimal steady state and the question of which objects we should work within the theoretical analysis, given that in numerical experiments one can observe turnpike properties concerning different objects like distributions, moments, or sample paths of the stochastic system, see [17].

The contribution of this paper is twofold: On the one hand we show a pathwise turnpike phenomenon in probability for stochastic systems in which the individual paths do not converge to a steady state but are subject to continued excitation by an additive disturbance. As we will see in Theorem 3, this pathwise turnpike behavior states that there is a specific stationary process such that with high probability the paths of

near-optimal processes stay close to corresponding paths of the stationary process, except for a number of time instances independent of the optimization horizon. This distinguishes our results from existing ones, e.g., in [12], [14], [21], which either consider turnpike phenomena in distribution rather than pathwise or assume that the stochastic system is L^2-stabilizable, in contrast to our setting. On the other hand, we provide the—to the best of the authors' knowledge—first dissipativity-based analysis of a pathwise turnpike property for stochastic optimal control problems.

We investigate this connection in the discrete-time linear-quadratic setting. We note that our strict dissipativity notion differs from the one recently introduced in [6], because our notion cannot be reformulated in the sense of the underlying probability measures. In fact, a dissipativity notion in the sense of probability measures would not be strong enough to conclude the pathwise turnpike property in probability that we achieve in this paper, cf. Remark 1 and Section IV.

The remainder of the paper is structured as follows: Section II introduces the considered problem formulation and recalls the concepts of turnpike and dissipativity in the deterministic setting. Section III shows that a pair of stationary stochastic processes fulfilling a suitable optimality criterion can replace the deterministic optimal steady state. Moreover, we show in Section IV that a time-varying dissipativity notion implies a pathwise turnpike property for this stationary pair. We also present an illustrative example in Section V and summarize our results in Section VI.

II. SETTING AND PRELIMINARIES

Before investigating stochastic turnpike properties, we introduce our problem set-up and briefly recall the basic concepts of dissipativity and turnpike in the deterministic case.

A. Problem formulation

For a stabilizable pair (A, B), we consider linear stochastic systems of the form

$$X(k + 1) = AX(k) + BU(k) + W(k), \quad X(0) = X_0 \quad (1)$$

where at each time step $k \in \mathbb{N}_0$, $X(k) \in L^2(\Omega, \mathcal{F}_k, \mathbb{P}, \mathbb{R}^n)$, $U(k) \in L^2(\Omega, \mathcal{F}_k, \mathbb{P}, \mathbb{R}^l)$, and $W(k) \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^n)$. Here Ω is the set of realizations, $\mathbb{P}$ is the probability measure, $\mathcal{F}$ is a σ-algebra, and $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$ is a filtration following the usual hypotheses of [18]. For our purpose, we choose $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$ as the smallest filtration such that X is an adapted process, i.e.

$$\mathcal{F}_k = \sigma(X(0), \ldots, X(k)), \quad k \in \mathbb{N}_0.$$

This choice of the stochastic filtration induces a causality requirement, which ensures that the control action $U(k)$ at time k only depends on the sequence of past disturbances $\{W(k)\}_{k=0,\ldots,k-1}$ and not on future events. We refer to [5], [18] for more details on stochastic filtrations.

Further, for all $k \in \mathbb{N}_0$ we assume that $W(k) \sim \mathcal{N}(0, \Sigma_W)$ are *i.i.d* Gaussian random variables which are independent of $X(k)$ and $U(k)$ and which have zero mean and covariance matrix $\Sigma_W \in \mathbb{R}^{n \times n}$. For a given initial value $X_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$ and control $U(\cdot)$ we denote the solution of system (1) by $X_U(\cdot, X_0)$, or short by $X(\cdot)$ if the initial value and the control are unambiguous. Note, that the solution $X_U(\cdot, X_0)$ also depends on the disturbance $W(\cdot)$. However, for the sake of readability, we do not highlight this in our notation.

For two matrices $R \in \mathbb{R}^{l \times l}$, $Q \in \mathbb{R}^{n \times n}$ with $R > 0$, $Q \geq 0$ and $(A, Q^{1/2})$ detectable, the stage cost is given by

$$\ell(X, U) := \mathbb{E}\left[\|X\|_Q^2 + \|U\|_R^2\right]. \tag{2}$$

The stochastic optimal control problem (stochastic OCP) under consideration is

$$\min_{U(\cdot)} J_N(X_0, U) := \sum_{k=0}^{N-1} \ell(X(k), U(k)) \tag{3}$$

subject to (1). The solution of problem (3) is well known, see [2], and given by $U^*(k) = K_N(k) X_{U^*}(k, X_0)$ where

$$K_N(k) := -\left[R + B^T P_N(k+1) B\right]^{-1} B^T P_N(k+1) A$$

and $P_N(k)$ is the solution of the backward Riccati iteration (or Riccati difference equation)

$$\begin{aligned} P_N(k) =& A^T P_N(k+1) A + Q - A^T P_N(k+1) B \\ & \times [R + B^T P_N(k+1) B]^{-1} B^T P_N(k+1) A \end{aligned} \tag{4}$$

with the terminal condition $P_N(N) = 0$.

B. Deterministic Dissipativity and Turnpike

Let us denote the deterministic counterpart to system (1) by

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0 \tag{5}$$

and the corresponding solution analogously by $x_u(\cdot, x_0)$. It is easy to see that $(x^s, u^s) = (0, 0)$ solves the optimization problem

$$\min_{x,u} \ell(x, u) = \|x\|_Q^2 + \|u\|_R^2 \quad \text{s.t. } x = Ax + Bu. \tag{6}$$

Therefore, we call $(x^s, u^s) = (0, 0)$ the optimal steady state of the deterministic linear-quadratic optimal control problem. Next, we formalize the strict dissipativity property for the deterministic problem.

Definition 1: The deterministic linear-quadratic optimal control problem is called *strictly (x, u)-dissipative* at (x^s, u^s) if there exists a storage function $\lambda : \mathbb{R}^n \to \mathbb{R}$ bounded from below and a function $\rho \in \mathcal{K}_\infty$, i.e., $\rho : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ continuous, strictly increasing and unbounded with $\rho(0) = 0$, such that

$$\begin{aligned} \ell(x, u) - \ell(x^s, u^s) + \lambda(x) - \lambda(Ax + Bu) \\ \geq \rho(\|(x - x^s, u - u^s)\|) \end{aligned} \tag{7}$$

holds for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^l$.

The following lemma, a slight modification of [7, Theorem 5.3], shows that strict dissipativity implies turnpike-like behavior for the deterministic case.

Lemma 1: Assume strict (x, u)-dissipativity at (x^s, u^s). Then for each $x_0 \in \mathbb{R}^n$ there exists a constant $C \in \mathbb{R}$ such that for each $\delta > 0$, each control sequence $u(\cdot)$ satisfying $J_N(x_0, u) \leq N\ell(x^s, u^s) + \delta$ and each $\varepsilon > 0$ the value $Q_\varepsilon := \#\{k \in \{0, \ldots, N-1\} \mid \|(x_u(k, x_0) - x^s, u(k) - u^s)\| \leq \varepsilon\}$ satisfies the inequality $Q_\varepsilon \geq N - (\delta + C)/\rho(\varepsilon)$.

We note that this result holds for much more general optimal control problems than specified above. In the setting from above, we have $\ell(x^s, u^s) = 0$ and thus we will see the turnpike phenomenon for all solutions $x_u(k, x_0)$ for which $J_N(x_0, u)$ is uniformly bounded in N. We note that although in the sequel we do not change the structure of the cost function, we will not have that the cost vanishes in the stochastic counterpart of the turnpike steady state (x^s, u^s) that we identify below.

III. STATIONARY SOLUTIONS AND OVERTAKING OPTIMALITY

Our aim is to transfer the results from the deterministic to the stochastic setting. For that, we have to consider how we can receive an equivalent formulation of (6) to characterize the optimal steady state for system (1). We note that, for any $U \in L^2(\Omega, \mathcal{F}_k, \mathbb{P}, \mathbb{R}^l)$, the condition

$$X = AX + BU + W \tag{8}$$

cannot be satisfied as U may not depend on W according to the underlying filtration. Thus, a random variable, which is constant in time, is not a suitable candidate for the stochastic counterpart of the optimal steady state. However, we can keep the distribution of such a process constant, instead, and use the following definition of stationary processes.

Definition 2: A pair of state and control processes $(X^s(\cdot), U^s(\cdot))$ is called stationary for system (1) if there is a stationary distribution ϱ^s such that

$$X^s(k) \sim \varrho^s \quad \text{and} \quad X^s(k+1) = AX^s(k) + BU^s(k) + W(k)$$

for all $k \in \mathbb{N}_0$.

Remark 1 (Stationarity in probability measures):
The probability measure defining a stationary distribution is called an invariant measure, see [16]. Hence, an alternative approach to the above stationarity concept could be to switch entirely to the set of underlying probability measures and to conduct the analysis there, see [6]. However, this approach is limited in that all information about the single realization paths of the solutions is lost – although numerical simulations suggest a turnpike phenomenon for the paths, as well, see [17].

As shown in [15], every control of the form $U(k) := LX(k)$ with $A + BL$ Schur-stable yields an invariant probability measure and, therefore, such a stationary pair. So we need an additional optimality criterion to characterize optimal stationary pairs along with the optimal L. Since we aim at formulating this criterion regarding the stochastic

processes, which are not constant over time, we cannot minimize the stage cost (2) for one single time step. Instead, we must include the entire evolution of $X(\cdot)$ and $U(\cdot)$ in the optimization process. For this purpose, let us first derive a reformulation of the cost.

Lemma 2: Let P be the unique positive semidefinite solution of the discrete-time algebraic Riccati equation

$$P = A^T P A + Q - A^T P B [R + B^T P B]^{-1} B^T P A, \quad (9)$$

and set $K := -\left[R + B^T P B\right]^{-1} B^T P A$. Then for every $N \in \mathbb{N}$ the cost (3) can be rewritten as

$$J_N(X_0, U) = \sum_{k=0}^{N-1} \mathbb{E}\left[\|U(k) - KX(k)\|_{\widetilde{R}}^2\right] + \mathbb{E}\left[\|X_0\|_P^2\right]$$
$$- \mathbb{E}\left[\|X(N)\|_P^2\right] + \sum_{k=0}^{N-1} \mathbb{E}\left[\|W(k)\|_P^2\right]$$

with $\widetilde{R} := R + B^T P B$ symmetric and positive definite.

Proof: Since P is the constant solution of the Riccati iteration (4) with terminal condition $P_N(N) = P$, this lemma is a direct consequence of [20, Chapter 8, Lemma 6.1] and the stochastic independence of $X(k)$ and $W(k)$ as well as $U(k)$ and $W(k)$. ∎

Lemma 2 reveals that for every control $U(\cdot)$, the cost grows unbounded if we let N go to infinity. That means we cannot directly solve the infinite horizon stochastic OCP. Thus, we use the following concept of *overtaking optimality*, originally introduced by [1], [23]. Moreover, it was already used to analyze turnpike behavior for time-varying discrete-time optimal control problems in deterministic settings, see [11].

Definition 3: Let X_0 be a fixed initial condition. Then a control $U_1(\cdot)$ overtakes $U_2(\cdot)$ if there is a time N_0 such that

$$J_N(X_0, U_1) < J_N(X_0, U_2), \quad \text{for all } N > N_0.$$

For a set of controls $\mathcal{U}$, a control $U^*(\cdot) \in \mathcal{U}$ is called overtaking optimal on $\mathcal{U}$ if it overtakes every other control $U(\cdot) \in \mathcal{U}$.

Note that this definition implies the uniqueness of the overtaking optimal control if it exists. For the continuous-time case, it has been shown in [13] that the feedback law obtained from the solution of the algebraic Riccati equation defines an overtaking optimal control. We adapt this proof and show that the claim also holds for the discrete-time case.

Theorem 1: For every initial condition $X_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$ there exists a unique overtaking optimal control $U^*(\cdot)$ on the set

$$\mathcal{U} = \{U(\cdot) \mid \exists \alpha > 0 : \mathbb{E}\left[\|X_U(k, X_0)\|^2\right] \leq \alpha \text{ and}$$
$$U(k) \in L^2(\Omega, \mathcal{F}_k, \mathbb{P}, \mathbb{R}^l) \text{ for all } k \in \mathbb{N}_0\}$$

given by the feedback law

$$U^*(k) = KX^*(k), \quad X^*(\cdot) := X_{U^*}(\cdot, X_0) \quad (10)$$

where $K := -\left[R + B^T P B\right]^{-1} B^T P A$ and P is the positive semidefinite symmetric solution of the discrete-time algebraic Riccati equation (9). Further, the following stronger

version of overtaking optimality holds: For every $U(\cdot) \in \mathcal{U}$ with $U(\cdot) \neq U^*(\cdot)$ there is a time N_0 and an $\varepsilon > 0$ such that

$$J_N(X_0, U^*) + \varepsilon < J_N(X_0, U)$$

for all $N \geq N_0$.

Proof: First we have to verify that $U^*(\cdot) = KX^*(\cdot) \in \mathcal{U}$. Since $U^*(\cdot)$ only depends on the current state, it is clear that the filtration condition is fulfilled and thus $U^*(k) \in L^2(\Omega, \mathcal{F}_k, \mathbb{P}, \mathbb{R}^l)$ holds for all $k \in \mathbb{N}_0$. So it remains to show that $\mathbb{E}[\|X^*(k)\|^2]$ is bounded for all $k \in \mathbb{N}_0$. We can calculate the distribution of $X^*(k) \sim \mathcal{N}(\mu(k), \Sigma(k))$ directly by the equations

$$\begin{aligned} \mu(k+1) &= A_K \mu(k) \\ \Sigma(k+1) &= A_K \Sigma(k) A_K^T + \Sigma_W \end{aligned} \quad (11)$$

with the initial condition $(\mu(0), \Sigma(0)) = (\mu_0, \Sigma_0)$ and the Schur-stable matrix $A_K := A + BK$. Therefore, we know that $\mu(k)$ and $\Sigma(k)$ converge because of the Schur-stability of A_K and, thus, we get that

$$\mathbb{E}\left[\|X(k)\|^2\right] = \sum_{i=0}^{n} \mathbb{E}\left[X_i(k)^2\right] = \mu(k)^T \mu(k) + Tr(\Sigma(k))$$

converges, too, which especially implies the boundedness of $\mathbb{E}[\|X^*(k)\|^2]$.

Now, since we know that $U^*(\cdot) \in \mathcal{U}$, we can write every

$$U(k) = KX(k) + V(k), \quad \text{for all } k \in \mathbb{N}_0$$

with $V(\cdot) \in \mathcal{U}$. If we use Lemma 2 to calculate the cost of $U(\cdot)$ we get

$$J_N(X_0, U) = \sum_{k=0}^{N-1} \mathbb{E}\left[\|V(k)\|_{\widetilde{R}}^2\right] + \mathbb{E}\left[\|X_0\|_P^2\right]$$
$$- \mathbb{E}\left[\|X(N)\|_P^2\right] + \sum_{k=0}^{N-1} \mathbb{E}\left[\|W(k)\|_P^2\right]$$

and, thus,

$$J_N(X_0, U) - J_N(X_0, U^*)$$
$$= \sum_{k=0}^{N-1} \mathbb{E}\left[\|V(k)\|_{\widetilde{R}}^2\right] - \mathbb{E}\left[\|X(N)\|_P^2\right] + \mathbb{E}\left[\|X^*(N)\|_P^2\right].$$

We consider two cases to prove that $U^*(\cdot)$ is overtaking optimal. First assume $\lim_{N \to \infty} \sum_{k=0}^{N-1} \mathbb{E}[\|V(k)\|_{\widetilde{R}}^2] = \infty$. Then we get by the boundedness of $\mathbb{E}[\|X(T)\|^2]$ and the nonnegativity of $\mathbb{E}[\|X^*(T)\|_P^2]$ that there exists a time $N_0 > 0$ and a constant $\alpha_0 > 0$ such that

$$J_N(X_0, U) - J_N(X_0, U^*) \geq \sum_{k=0}^{N-1} \mathbb{E}\left[\|V(k)\|_{\widetilde{R}}^2\right] - \alpha_0 > \varepsilon$$

for all $N > N_0$, i.e., U^* is overtaking optimal.

Let us now consider the second possibility

$$0 \leq \lim_{N \to \infty} \sum_{k=0}^{N-1} \mathbb{E}\left[\|V(k)\|_{\widetilde{R}}^2\right] =: \gamma < \infty. \quad (12)$$

First, we assume $\gamma = 0$. Then we can conclude that $\mathbb{E}[\|V(k)\|_{\tilde{R}}^2] = 0$ for all $k \in \mathbb{N}_0$ and thus $V(\cdot) = 0$ almost surely, which implies $U(\cdot) = U^*(\cdot)$. Therefore, we can w.l.o.g. assume that $\gamma > 0$, which implies that there exists a constant $\tilde{\varepsilon} > 0$ and a time $N_1 > 0$ such that

$$\sum_{k=0}^{N-1} \mathbb{E}\left[\|V(k)\|_{\tilde{R}}^2\right] > \tilde{\varepsilon} \quad \text{for all } N \geq N_1. \quad (13)$$

Next, we show that

$$\lim_{N\to\infty} \left[\mathbb{E}\left[\|X(N)\|_P^2\right] - \mathbb{E}\left[\|X^*(N)\|_P^2\right]\right] = 0 \quad (14)$$

for which by the equality $x^T P x = tr(P x x^T)$ it is sufficient to prove that

$$\lim_{N\to\infty}\left[\mathbb{E}\left[X(N)X(N)^T\right] - \mathbb{E}\left[X^*(N)X^*(N)^T\right]\right] = 0.$$

The equations satisfied by $X^*(\cdot)$ and $X(\cdot)$ are

$$X^*(k+1) = A_K X^*(k) + W(k)$$
$$X(k+1) = A_K X(k) + BV(k) + W(k)$$

and thus, the solution $X(\cdot)$ is given by

$$X(k) = X^*(k) + \sum_{j=0}^{k-1} A_K^j BV(k-1-j). \quad (15)$$

From equation (15) we get

$$\mathbb{E}\left[X(k)X(k)^T\right] = \mathbb{E}\left[X^*(k)X^*(k)^T\right]$$
$$+ \mathbb{E}\left[X^*(k)\tilde{V}(k)^T\right] + \mathbb{E}\left[\tilde{V}(k)X^*(k)^T\right] + \mathbb{E}\left[\tilde{V}(k)\tilde{V}(k)^T\right]$$

with $\tilde{V}(k) := \sum_{j=0}^{k-1} A_K^j BV(k-1-j)$. It follows from the Cauchy-Schwarz inequality that there is a constant $\kappa > 0$ such that any two random variables Z_1 and Z_2 satisfy

$$\left\|\mathbb{E}\left[Z_1 Z_2^T\right]\right\| \leq \kappa \left(\mathbb{E}\left[\|Z_1\|^2\right]\right)^{1/2} \left(\mathbb{E}\left[\|Z_2\|^2\right]\right)^{1/2}. \quad (16)$$

By inequality (16), it will be sufficient to prove that $\mathbb{E}[\|\tilde{V}(k)\|^2] \to 0$ as $k \to \infty$ in order to establish $\mathbb{E}[X(k)X(k)^T - X^*(k)X^*(k)^T] \to 0$. It holds that

$$\mathbb{E}\left[\|\tilde{V}(k)\|^2\right] \leq \|B\|^2 \sum_{j=0}^{k-1} \|A_K^j\|^2 \mathbb{E}\left[\|V(k-1-j)\|^2\right]$$

which converges to zero for $k \to \infty$ because $\|A_K^j\|^2$ and $\mathbb{E}[\|V(j)\|^2]$ are summable by Schur-stability of A_K and equation (12), respectively. This proves (14), and thus there exist a time $N_0 \geq N_1$ such that

$$J_N(X_0, U) - J_N(X_0, U^*)$$
$$= \sum_{k=0}^{N-1} \mathbb{E}\left[\|V(k)\|_{\tilde{R}}^2\right] + \mathbb{E}\left[\|X^*(N)\|_P^2\right] - \mathbb{E}\left[\|X(N)\|_P^2\right]$$
$$> \tilde{\varepsilon} - \frac{\tilde{\varepsilon}}{2} = \frac{\tilde{\varepsilon}}{2} := \varepsilon > 0$$

for all $N > N_0$, which concludes the proof. ∎

Note that the restriction to the set $\mathcal{U}$ in Theorem 1 is not limiting, since we are only interested in optimal stationary pairs and every control defining a stationary pair must lie in $\mathcal{U}$ to provide an invariant measure. Next, we show that with a proper choice of X_0 the overtaking optimal control process from Theorem 1 defines a stationary pair in the sense of Definition 2.

Lemma 3: For system (1), there exists a stationary pair with control $U_*^s(\cdot) = KX_*^s(\cdot)$, where K is the feedback law generating the overtaking optimal control from Theorem 1, together with the state process

$$X_*^s(k+1) = (A + BK)X_*^s(k) + W(k)$$

and initial distribution $X_*^s(0) \sim \mathcal{N}(\mu_*^s, \Sigma_*^s)$. Here $\mu_*^s = 0$ and Σ_*^s is the solution of the Lyapunov equation

$$\Sigma_*^s = (A + BK)\Sigma_*^s(A + BK)^T + \Sigma_W.$$

Proof: Since $A + BK$ is Schur-stable, by [15] the stationary state process $X_*^s(\cdot)$ exists. From equation (11) it is clear that the corresponding invariant measure is given by $\mathcal{N}(\mu_*^s, \Sigma_*^s)$ where μ_*^s and Σ_*^s must satisfy the equations

$$\mu_*^s = A_K \mu_*^s, \quad \Sigma_*^s = A_K \Sigma_*^s A_K^T + \Sigma_W. \quad ∎$$

IV. STOCHASTIC DISSIPATIVITY AND TURNPIKE

As we have characterized the optimal stationary pair for the considered stochastic OCP, we now show that this pair satisfies a stochastic extension of strict dissipativity. In the resulting dissipativity inequality, the distance measure on the right-hand side depends on the exact realization of the random variables rather than only on their distributions, cf. (7). That means our distance measure is not a dissimilarity measure in the sense of [6], and, thus, our concept of stochastic dissipativity differs from that one presented there. Specifically, we construct a modified cost function and a corresponding storage function in several steps in a series of lemmas, starting with the stage cost from Lemma 2.

Lemma 4: For every control $U(\cdot)$ and corresponding state $X(\cdot) = X_U(\cdot, X_0)$, the equality

$$\hat{\ell}(X(k), U(k)) := \ell(X(k), U(k)) - \ell(X_*^s(k), U_*^s(k))$$
$$+ \hat{\lambda}(X(k)) - \hat{\lambda}(X(k+1)) = \mathbb{E}\left[\|U(k) - KX(k)\|_{\tilde{R}}^2\right]$$

holds with $\hat{\lambda}(X(k)) = -\mathbb{E}[\|X(k)\|_P^2]$.

Proof: Using Lemma 2 with $N = 1$ we get

$$\ell(X(k), U(k)) = \mathbb{E}\left[\|U(k) - KX(k)\|_{\tilde{R}}^2\right] + \mathbb{E}\left[\|X(k)\|_P^2\right]$$
$$- \mathbb{E}\left[\|X(k+1)\|_P^2\right] + \mathbb{E}\left[\|W(k)\|_P^2\right]$$

and $\ell(X_*^s(k), U_*^s(k)) = \mathbb{E}[\|W(k)\|_P^2]$. This results in

$$\hat{\ell}(X(k), U(k)) = \ell(X(k), U(k))$$
$$- \mathbb{E}\left[\|W(k)\|_P^2\right] - \mathbb{E}\left[\|X(k)\|_P^2\right] + \mathbb{E}\left[\|X(k+1)\|_P^2\right]$$
$$= \mathbb{E}\left[\|U(k) - KX(k)\|_{\tilde{R}}^2\right]$$

which proves the claim. ∎

If we see the control $U(\cdot)$ as a kind of strategy $\pi(\cdot)$ of the form $U(k) = \pi(X(k))$ then Lemma 4 could already be interpreted as a dissipativity equation with respect to the strategy. However, since we are interested in deriving a

stochastic version of (x, u)-dissipativity, we further modify this equation, leading to the following lemma. For simplicity, we denote in the following the difference between a stochastic state or control process and the optimal stationary process by $\widetilde{X}(\cdot) := X(\cdot) - X_*^s(\cdot)$, respectively, by $\widetilde{U}(\cdot) := U(\cdot) - U_*^s(\cdot)$.

Lemma 5: For every control $U(\cdot)$ it holds that

$$\bar{\ell}(X(k), U(k)) := \ell(X(k), U(k)) - \ell(X_*^s(k), U_*^s(k)) \\ + \bar{\lambda}(k, X(k)) - \bar{\lambda}(k+1, X(k+1)) \\ = \mathbb{E}\left[\|X(k) - X_*^s(k)\|_Q^2 + \|U(k) - U_*^s(k)\|_R^2 \right]$$

for every $k \in \mathbb{N}_0$. Here, $\bar{\lambda}$ is given by

$$\bar{\lambda}(k, X(k)) := -\mathbb{E}\left[\|X(k)\|_P^2 - \|X(k) - X_*^s(k)\|_P^2 \right].$$

Proof: First, we observe that

$$\mathbb{E}\left[\|U(k) - KX(k)\|_{\hat{R}}^2 \right] \\ = \mathbb{E}\left[\|(U(k) - U_*^s(k)) - K(X(k) - X_*^s(k))\|_{\hat{R}}^2 \right]$$

holds with $U_*^s(k) = KX_*^s(k)$. Thus, we obtain $\hat{\ell}(X(k), U(k)) = \hat{\ell}(\widetilde{X}(k), \widetilde{U}(k))$ by Lemma 4 and therefore

$$\mathbb{E}\left[\|\widetilde{X}(k)\|_Q^2 + \|\widetilde{U}(k)\|_R^2 \right] = \ell(\widetilde{X}(k), \widetilde{U}(k)) \\ = \hat{\ell}(X(k), U(k)) + \ell(X_*^s(k), U_*^s(k)) \\ - \hat{\lambda}(\widetilde{X}(k)) + \hat{\lambda}\left(A\widetilde{X}(k) + B\widetilde{U}(k) + W(k)\right).$$

Further, because $(X(k), U(k))$ as well as $(U_*^s(k), X_*^s(k))$ are stochastically independent of $W(k)$ we get

$$\hat{\lambda}\left(A\widetilde{X}(k) + B\widetilde{U}(k) + W(k)\right) = -\mathbb{E}\left[\|W(k)\|_P^2 \right] \\ - \mathbb{E}\left[\|AX(k) + BU(k) - AX_*^s(k) - BU_*^s(k)\|_P^2 \right] \\ = -\mathbb{E}\left[\|\widetilde{X}(k+1)\|_P^2 \right] - \ell(X_*^s(k), U_*^s(k)).$$

Putting all these equations together, we finally get

$$\mathbb{E}\left[\|X(k) - X_*^s(k)\|_Q^2 + \|U(k) - U_*^s(k)\|_R^2 \right] \\ = \hat{\ell}(X(k), U(k)) + \mathbb{E}\left[\|X(k) - X_*^s(k)\|_P^2 \right] \\ - \mathbb{E}\left[\|X(k+1) - X_*^s(k+1)\|_P^2 \right] = \bar{\ell}(X(k), U(k)). \quad \blacksquare$$

With a final modification of the stage cost we can replace the positive semidefinite matrix Q in the weight of the state with a positive definite matrix. This leads to the next lemma.

Lemma 6: There exists symmetric and positive definite matrices $S \in \mathbb{R}^{n \times n}$, $H \in \mathbb{R}^{(n+l) \times (n+l)}$, such that for every control $U(\cdot)$ it holds that

$$\tilde{\ell}(X(k), U(k)) := \ell(X(k), U(k)) - \ell(X_*^s(k), U_*^s(k)) \\ + \tilde{\lambda}(k, X(k)) - \tilde{\lambda}(k+1, X(k+1)) \\ = \mathbb{E}\left[\|(X(k) - X_*^s(k), U(k) - U_*^s(k))\|_H^2 \right]$$

with $\tilde{\lambda}(k, X(k)) := \bar{\lambda}(k, X(k)) + \mathbb{E}\left[\|X(k) - X_*^s(k)\|_S^2 \right]$.

Proof: Since $(A, Q^{1/2})$ is detectable, we know by [8, Lemma 5.4] that there is $\tilde{S} \in \mathbb{R}^{n \times n}$ symmetric and positive definite satisfying the matrix inequality $Q + \tilde{S} - A^T\tilde{S}A > 0$.

For a given $\gamma \in (0, 1]$, set $\widetilde{S}_\gamma := \gamma\widetilde{S}$ and $Q_\gamma := Q + \widetilde{S}_\gamma - A^T\widetilde{S}_\gamma A$. Then following the calculation of [8, Lemma 4.1] we get

$$\ell(\widetilde{X}(k), \widetilde{U}(k)) + \mathbb{E}\left[\|\widetilde{X}(k)\|_{\widetilde{S}_\gamma}^2 \right] - \mathbb{E}\left[\|\widetilde{X}(k+1)\|_{\widetilde{S}_\gamma}^2 \right] \\ = \bar{\ell}(X(k), U(k)) + \mathbb{E}\left[\|\widetilde{X}(k)\|_{\widetilde{S}_\gamma}^2 \right] \\ - \mathbb{E}\left[\|A\widetilde{X}(k) + B\widetilde{U}(k)\|_{\widetilde{S}_\gamma}^2 \right] = \mathbb{E}\left[\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2 \right]$$

with

$$H := \frac{1}{2}\begin{pmatrix} 2Q_\gamma & \gamma E \\ \gamma E & 2R_\gamma \end{pmatrix},$$

$R_\gamma := R - B^T\widetilde{S}_\gamma B$ and $E := -A^T\widetilde{S}B - B^T\widetilde{S}A$. Using the Schur complement we can show, that H is positive definite for a sufficient small $\tilde{\gamma} \in (0, 1]$, see [8], which proves the lemma with $S = \widetilde{S}_{\tilde{\gamma}} > 0$. $\quad \blacksquare$

Note that,

$$\tilde{\lambda}(k, X(k)) = \mathbb{E}\left[\|X(k) - X_*^s(k)\|_{P+S}^2 - \|X(k)\|_P^2 \right] \\ = \mathbb{E}\left[\|X(k)\|_S^2 - 2X(k)^T(P+S)X_*^s(k) + \|X_*^s(k)\|_{P+S}^2 \right]$$

is bounded from below since $S > 0$ and $P \geq 0$ and thus, Lemma 6 delivers a dissipativity equation equivalently to Definition 1. Therefore, we can show a turnpike property for the stochastic case analogous to Lemma 1.

Theorem 2: For each $X_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$, there exists a constant $C \in \mathbb{R}$ such that for each $\delta > 0$, each control process $U(\cdot)$ satisfying $J_N(X_0, U) \leq \delta + \sum_{k=0}^{N-1} \ell(X_*^s(k), U_*^s(k))$ and each $\varepsilon > 0$ the value

$$Q_\varepsilon := \#\left\{ k \in \{0, \dots, N-1\} \mid \mathbb{E}\left[\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2\right] \leq \varepsilon \right\}$$

satisfies the inequality $Q_\varepsilon \geq N - (\delta + C)/\varepsilon$ for all $N \in \mathbb{N}$.

Proof: The proof follows the same arguments as [7, Theorem 5.3]. Set $C := \tilde{\lambda}(0, X_0) - M$ where $M \in \mathbb{R}$ is a lower bound on $\tilde{\lambda}$. Then for $\tilde{J}_N(X_0, U) := \sum_{k=0}^{N-1} \tilde{\ell}(X(k), U(k))$ we get

$$\tilde{J}_N(X_0, U) = J_N(X_0, U) - \sum_{k=0}^{N-1} \ell(X_*^s(k), U_*^s(k)) \\ + \tilde{\lambda}(0, X(0)) - \tilde{\lambda}(N, X(N)) \leq \delta + C. \tag{17}$$

Now assume that $Q_\varepsilon < N - (\delta + C)/\varepsilon$. This means there is a set $\bar{\mathcal{M}} \subset \{0, \dots, N-1\}$ of $N - Q_\varepsilon > (\delta + C)/\varepsilon$ time instants such that $\mathbb{E}[\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2] \geq \varepsilon$ for all $k \in \bar{\mathcal{M}}$. Using Lemma 6, this implies $\tilde{J}_N(X_0, U) = \sum_{k=0}^{N-1} \mathbb{E}[\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2] > \delta + C$, which contradicts (17) and, thus, proves the theorem. $\quad \blacksquare$

It is not immediately obvious what Theorem 2 means for the single realization paths of $X(\cdot)$ and $U(\cdot)$. What we cannot conclude is an almost sure turnpike property of the paths. However, what we can conclude is a turnpike property in probability. This means the probability that a realization is not near $(X_*^s(\cdot), U_*^s(\cdot))$ is small in an appropriate probabilistic sense, except for a number of time instances independent of N. The following theorem formalizes this result.

Theorem 3: For each $X_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$, there exists a constant $C \in \mathbb{R}$ such that for each $\delta > 0$, each control process $U(\cdot)$ satisfying $J_N(X_0, U) \leq \delta + \sum_{k=0}^{N-1} \ell(X_*^s(k), U_*^s(k))$ and each $\epsilon, \eta > 0$ the value

$$P_{\epsilon,\eta} := \#\left\{k \in \mathcal{M}_N \mid \mathbb{P}\left(\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2 \geq \epsilon\right) \leq \eta\right\}$$

satisfies $P_{\epsilon,\eta} \geq N - (\delta + C)/(\epsilon\eta)$ for all $N \in \mathbb{N}$, where $\mathcal{M}_N := \{0, \ldots, N-1\}$.

Proof: Using the Markov inequality, we get

$$\mathbb{P}\left(\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2 \geq \epsilon\right) \leq \frac{1}{\epsilon}\mathbb{E}\left[\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2\right]. \quad (18)$$

Further, by Theorem 2, we know that there are at least $N - (\delta + C)/(\epsilon\eta)$ time instants for which $\mathbb{E}[\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2] \leq \epsilon\eta$. Using equation (18), this gives $\mathbb{P}\left(\|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2 \geq \epsilon\right) \leq \eta$ for all these time instants and, thus, proves the claim. ∎

Note that in contrast to the classical turnpike approaches, we have to fix two constants in Theorem 3 where ϵ defines how close we want to be to the realization of the stationary pair and η defines the desired probability bound for this.

V. NUMERICAL EXAMPLE

Consider the one-dimensional stochastic OCP

$$\min_{U(\cdot)} \sum_{k=0}^{N-1} \mathbb{E}\left[X(k)^2 + 5U(k)^2\right] \quad (19)$$

$$\text{s.t. } X(k+1) = 1.2X(k) + U(k) + W(k)$$

with initial condition $X(0) = X_0 \sim \mathcal{N}(3, 1.5)$ and $W(k) \sim \mathcal{N}(0, 10)$ for all $k = 0, \ldots, N-1$. Then the corresponding stationary pair from Lemma 3 is defined by the feedback $K \approx -0.558$ and the initial distribution $X_*^s(0) \sim \mathcal{N}(0, 17)$. To observe the pathwise turnpike property from Theorem 3, we fix a realization of $W(\cdot)$ given by $w = \{w_k\}$ and simulate the states and controls according to the equation (19) and the optimal stationary pair with $W(k) = w_k$ and random initial values following the desired distributions. Figure 1 shows the noise-sequence w and the resulting turnpike behavior of the realization. Further, Table I shows the number of points $P_{0.25}^w := \#\{k \in \mathcal{M}_N \mid \|(\widetilde{X}(k), \widetilde{U}(k))\|_H^2 \geq 0.25\}$ outside the $\varepsilon = 0.25$ neighborhood of the stationary pair for the paths from Figure 1 for the fixed realization w of the noise. It can be observed that the value $P_{0.25}^w$ is uniformly bounded by 12, which further illustrates the pathwise turnpike property of this realization.

N	10	15	20	25	30	35	40	45	50
$P_{0.25}^w$	10	9	11	12	9	10	12	10	12

TABLE I

NUMBER OF POINTS $P_{0.25}^w$ FOR DIFFERENT HORIZONS N.

VI. SUMMARY

This paper presented pathwise turnpike results in probability based on strict dissipativity for discrete-time stochastic linear-quadratic optimal control problems. Compared to the deterministic setting, a stationary stochastic process replaces the optimal steady state. We have illustrated the theoretical results considering a one-dimensional example.
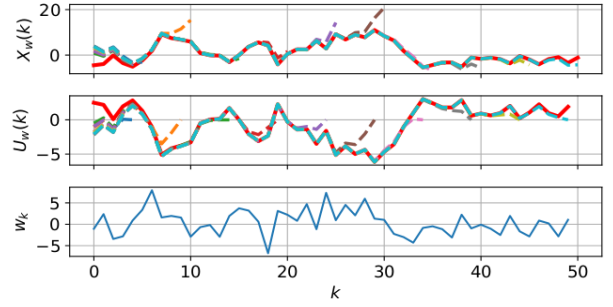

Fig. 1. Fixed realization w of the noise and corresponding realizations of the optimal solutions $(X_N^*(\cdot), U_N^*(\cdot))$ from (19) on different horizons N (dashed) and the optimal stationary pair $(X_*^s(\cdot), U_*^s(\cdot))$ (solid red).

REFERENCES

[1] H. Atsumi. Neoclassical growth and the efficient program of capital accumulation. *Rev. Econ. Stud.*, 32(2):127—136, 1965.

[2] P. E. Caines and D. Q. Mayne. On the discrete time matrix Riccati equation of optimal control. *Internat. J. Control*, 12:785–794, 1970.

[3] T. Faulwasser and L. Grüne. Turnpike properties in optimal control. In *Numerical Control: Part A*, pages 367–400. Elsevier, 2022.

[4] T. Faulwasser, M. Korda, C. N. Jones, and D. Bonvin. On turnpike and dissipativity properties of continuous-time optimal control problems. *Automatica*, 81:297–304, jul 2017.

[5] B. Fristedt and L. Gray. *A Modern Approach to Probability Theory*. Birkhäuser Boston, 1997.

[6] S. Gros and M. Zanon. Economic MPC of markov decision processes: Dissipativity in undiscounted infinite-horizon optimal control. *Automatica*, 146:110602, dec 2022.

[7] L. Grüne. Economic receding horizon control without terminal constraints. *Automatica*, 49(3):725–734, mar 2013.

[8] L. Grüne and R. Guglielmi. Turnpike properties and strict dissipativity for discrete time linear quadratic optimal control problems. *SIAM J. Control Optim.*, 56(2):1282–1302, jan 2018.

[9] L. Grüne, C. M. Kellett, and S. R. Weller. On the relation between turnpike properties for finite and infinite horizon optimal control problems. *J. Optim. Theory Appl.*, 173(3):727–745, mar 2017.

[10] L. Grüne and M. A. Müller. On the relation between strict dissipativity and turnpike properties. *Systems Control Lett.*, 90:45–53, apr 2016.

[11] L. Grüne, S. Pirkelmann, and M. Stieler. Strict dissipativity implies turnpike behavior for time-varying discrete time optimal control problems. In *Lecture Notes in Economics and Mathematical Systems*, pages 195–218. Springer International Publishing, 2018.

[12] V. Kolokoltsov and W. Yang. Turnpike theorems for markov games. *Dyn. Games Appl.*, 2(3):294–312, may 2012.

[13] A. Leizarowitz. Infinite horizon stochastic regulation and tracking with the overtaking criterion. *Stochastics*, 22(2):117–150, oct 1987.

[14] R. Marimon. Stochastic turnpike property and stationary equilibrium. *J. Econom. Theory*, 47(2):282–306, apr 1989.

[15] S. P. Meyn. Ergodic theorems for discrete time stochastic systems using a stochastic lyapunov function. *SIAM J. Control Optim.*, 27(6):1409–1439, nov 1989.

[16] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer London, 1993.

[17] R. Ou, M. H. Baumann, L. Grüne, and T. Faulwasser. A simulation study on turnpikes in stochastic LQ optimal control. *IFAC-PapersOnLine*, 54(3):516–521, 2021.

[18] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer Berlin Heidelberg, 2005.

[19] F. P. Ramsey. A mathematical theory of saving. *The Economic Journal*, 38(152):543, dec 1928.

[20] K. J. Åström. *Introduction to Stochastic Control Theory (Mathematics in Science and Engineering, Volume 70)*. Academic Press, 1970.

[21] J. Sun, H. Wang, and J. Yong. Turnpike properties for stochastic linear-quadratic optimal control problems. *Chinese Ann. Math. Ser. B*, 43(6):999–1022, nov 2022.

[22] J. von Neumann. A model of general economic equilibrium. *Rev. Econ. Stud.*, 13(1):1, 1945.

[23] C. C. von Weizsäcker. Existence of optimal programs of accumulation for an infinite time horizon. *Rev. Econ. Stud.*, 32:85–104, 1965.