Sparsity in Partially Controllable Linear Systems

Yonathan Efroni [1] **Sham Kakade** [1] **Akshay Krishnamurthy** [1] **Cyril Zhang** [1]

Abstract

A fundamental concept in control theory is that of controllability, where any system state can be reached through an appropriate choice of control inputs. Indeed, a large body of classical and modern approaches are designed for controllable linear dynamical systems. However, in practice, we often encounter systems in which a large set of state variables evolve exogenously and independently of the control inputs; such systems are only *partially controllable*. The focus of this work is on a large class of partially controllable linear dynamical systems, specified by an underlying sparsity pattern. Our main results establish structural conditions and finite-sample guarantees for learning to control such systems. In particular, our structural results characterize those state variables which are irrelevant for optimal control, an analysis which departs from classical control techniques. Our algorithmic results adapt techniques from high-dimensional statistics—specifically soft-thresholding and semiparametric least-squares—to exploit the underlying sparsity pattern in order to obtain finite-sample guarantees that significantly improve over those based on certainty-equivalence. We also corroborate these theoretical improvements over certainty-equivalent control through a simulation study.

1. Introduction

A recurring theme in modern sequential decision making and control applications is the presence of high-dimensional signals containing much irrelevant information. Operating on raw signals provides flexibility to learn much higher-quality policies than what may be expressed using hand-engineered inputs or features, but it poses new challenges for reinforcement learning (RL) and control. In the context of controls, high-dimensionality inevitably leads to many state variables that do not affect and cannot be affected by the controller inputs. Hence, these state variables are irrelevant for optimal control. In this work, we consider the question of how to efficiently learn to control partially controllable systems, while ignoring these irrelevant variables.

Example 1 (Turbine Orientation (Stanfel et al., 2020))**.** *Consider the problem of learning to orient turbines in a wind farm in response to sensor measurements of wind speed and direction. To learn a high-quality controller that can anticipate local wind patterns, it is desirable to collect measurements from a broad region. However geographical features such as mountains and valleys may render some of these measurements irrelevant for the control task, although this may not be known to the system designer in advance. As such, we would like our controller to efficiently learn to ignore these irrelevant sensors while relying on the relevant ones for decision making.*

Systems like this contain two challenging elements for learning to control. First, a large part of the system state — namely the wind speed and direction at all locations — is completely *uncontrollable*, as the wind turbines negligibly affect weather patterns. Rather, the controller must react to these state variables even though they cannot be controlled. Second, some of the uncontrollable variables may be completely *irrelevant*, meaning they have no bearing on the optimal control decisions. To complicate matters, which variables are controllable, uncontrollable, and irrelevant must be learned, ideally in a sample-efficient manner.

In the broader literature, there are two well-studied approaches for addressing high dimensionality. One approach is through feature engineering or the use of kernel machines, while the other exploits sparsity to recover certain low-dimensional structural information. Both approaches have been utilized in the context of decision making, the former via dimension-free linear control (Perdomo et al., 2021) and the Kernelized Nonlinear Regulator (Deisenroth and Rasmussen, 2011; Mania et al., 2020; Kakade et al., 2020), and the latter both in RL (Agarwal et al., 2020; Hao et al., 2021) and some works on continuous control (Fattahi and Sojoudi, 2018; Wang and Yang, 2020; Sun et al., 2020). This work contributes to the latter line of work on structure recovery in continuous control.

[*]Equal contribution [1]Microsoft Research New York, NY. Correspondence to: Yonathan Efroni <jonathan.efroni@gmail.com>.

Proceedings of the 39^{th} International Conference on Machine Learning, Baltimore, Maryland, USA, PMLR 162, 2022.

Our focus is on establishing non-asymptotic guarantees for learning to control in high-dimensional partially controllable systems like the wind farm example described above. We focus our attention on the problem of learning the linear quadratic regulator (LQR) in which the majority of the state variables are irrelevant.

Technical Overview. Deferring further details and technical motivation to subsequent sections, we present a brief overview of the setup and results. Consider a dynamical system of the form $x_{t+1} = Ax_t + Bu_t + \xi_t$ where $x_t \in \mathbb{R}^d$ is the system state, $u_t \in \mathbb{R}^{d_u}$ is the controller input, and ξ_t is a (stochastic) disturbance. The system is said to be *controllable* if, in expectation, any system state can be reached through an appropriate choice of a deterministic control sequence (Formally, this condition is equivalent to the controllability matrix being full rank. See Section 3). When such a condition does not hold, we call the system *partially controllable*. For such systems, it is well known that there exists an invertible transformation of the state variables, such that the system can be rewritten with dynamics of the form (Klamka, 1963; Sontag, 2013):

$$A = \begin{bmatrix} A_1 & A_{12}^{\mathrm{PC}} \\ 0 & A_2^{\mathrm{PC}} \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}. \tag{1}$$

Here the first block of coordinates corresponds to the controllable subsystem. On the other hand, the second block of *uncontrollable* coordinates cannot be affected by the control inputs (due to that $B_2 = 0$, although it can affect the controllable subsystem (if $A_{12}^{\mathrm{PC}} \neq 0$) (Klamka, 1963; Zhou et al., 1996; Sontag, 2013).

To capture the presence of *irrelevant* state variables that do not affect the controllable subsystem, we consider a dynamical system that is more structured than (1). In our setting, which we call the partially controllable linear-quadratic (PC-LQ) control problem, the system admits the block structure:

$$A = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix}. \tag{2}$$

To capture the irrelevance of state variables, our main learnability results will assume that the underlying dynamics of the system are determined by an (A, B) in this form, up to a permutation of the coordinates (see below for more discussion about this assumption). As we shall see, the first two blocks make up the relevant part of the system, while the third block of coordinates are irrelevant (in the sense that if we condition on knowing the values of the coordinates in blocks 1 and 2, then the state variables in block 3 provide no further information with regards to predicting the controllable coordinates in block 1, which, as we shall see, is what is required for optimal control). We are particularly interested in the high-dimensional regime where $A_1 \in \mathbb{R}^{s_c \times s_c}$, $A_2 \in \mathbb{R}^{s_e \times s_e}$ and $s_c + s_e := s \ll d$.

Our Contributions. Our first theorem is a structural result characterizing which state variables are irrelevant for optimal control. The result pertains to all problems equivalent to PC-LQ control, and is proven via an invariance argument. When specialized to PC-LQ control, the theorem verifies that the third block of state variables can be ignored by the optimal controller (while it is clear that the optimal value function *depends* on block three). This structural result and our assumption that the relevant subsystem (blocks one and two) comprises few state variables, shows that the optimal policy is "sparse": it is determined by poly(s) parameters, although neither the system dynamics A nor the optimal value function are sparse matrices.

Relying on the characterization of the relevant state variables for optimal control we turn to the main contribution of our work. We derive two algorithms that incorporate ideas from high-dimensional statistics to efficiently estimate only the relevant parts of the system dynamics. In Table 1 on page 3, we summarize the main results of the paper and compare with guarantees for certainty-equivalent control. We study two settings that differ only in their assumptions on the distribution of the starting state x_0. In the first setting (labeled "diagonal" in Table 1 on page 3), we assume that x_0 is sampled such that $\mathbb{E}[x_0] = 0$ and $\mathbb{E}[x_0 x_0^\top]$ is a diagonal matrix. In this case, we show that our algorithm learns a near-optimal control with a *nearly-dimension-free* rate: the sample complexity scales polynomially with the sparsity s and action dimension d_u, but only logarithmically with the ambient dimension d.

The second setting generalizes the diagonal case to only require that x_0 has strictly positive definite (PD) covariance. Here our algorithm incurs a lower order polynomial dependence on the ambient dimension d. In particular, for $d^2 \leq (s^2 + d_u s)/\epsilon$ this lower order term is dominated by the leading term, which yields the same sample complexity as in the diagonal case. In both settings, our bounds compare quite favorably to certainty-equivalent control, which incurs a poly$(d)/\epsilon$ leading order dependence. For the second setting, our algorithmic approach relies on a reduction to a semi-parametric least squares estimation (Chernozhukov et al., 2016; 2018a; Foster and Syrgkanis, 2019). We provide a new result (see Proposition 9), which might be of independent interest, for the semi-parametric least squares estimation algorithm for the linear case.

2. Preliminaries and Notation

Linear-Quadratic Control. A linear-quadratic (LQ) control problem is specified by a tuple of matrices $L = (A, B, Q, R)$. The state $x \in \mathbb{R}^d$ evolves according to $x_{t+1} = Ax_t + Bu_t + \xi_t$ where $u \in \mathbb{R}^{d_u}$ is the input to the system and ξ_t is i.i.d. noise. The cost conditioning on the first observation to be x_1 is given by

Covariance Matrix	Estimation Algorithm	Sample Complexity
Positive Definite	Least-Squares	$\tilde{O}\left(\frac{poly(d,d_u)}{\epsilon}\right)$
Diagonal	Second-Moment Product	$\tilde{O}\left(\frac{s^2+d_u s}{\epsilon}\right)$
Positive Definite	Semiparametric Least-Squares	$\tilde{O}\left(\frac{s^2+d_u s}{\epsilon}+\frac{\sqrt{(s^2+d_u s)d}}{\epsilon^{0.5}}\right)$

Table 1. Sample complexity results for learning a near-optimal controller in the PC-LQ setting. Our results, highlighted in gray, compare favorably with the classical least-squares/certainty-equivalent control when the relevant subsystem has dimensionality $s \ll d$. We assume the third, irrelevant, block of (2) is stable in L_∞ norm (Assumption 1). In $\tilde{O}(\cdot)$ we only keep polynomial dependence in ϵ, d, s, and d_u. See Appendix A for a thorough summary.

$J(x_1, \{u_t\}_{t\geq 1}) = \mathbb{E}\left[\sum_{t\geq 1} x_t^\top Q x_t + u_t R u_t \mid x_1\right]$ and $J(\{u_t\}_{t\geq 1}) = \mathbb{E}\left[J(x_1, \{u_t\}_{t\geq 1})\right]$. The cost matrices are assumed to be positive-semi definite, $Q \succcurlyeq 0, R \succ 0$. The task is to find the policy that minimizes $J\left(\{u_t\}_{t\geq 1}\right)$. It is well-known that the optimal controller, the linear quadratic regulator (LQR), of such a system is linear in the state vector, $u_t = K_\star x_t$, and the optimal value from any x_1 is given by $J_\star(x_1) = x_1^\top P_\star x_1$, where $P_\star$ is the solution of the Riccati equation and $K_\star = (R + B^T P_\star B)^{-1} B^\top P_\star A$. With some abuse of notation we let $J(K)$ be the expected cost when following taking actions according to $u = Kx$.

In this work, we assume that $R = I_{d_u}$, and write $L = (A, B, Q)$ for short. This can be obtained by rotating $u \to R^{-1/2}u$, which is valid since $R \succ 0$. We also assume the system is stabilizable, which means that there exists a matrix $K \in \mathbb{R}^{d_u \times d}$ such that $\rho(A + BK) < 1$, where $\rho(X) = \max\{|\lambda_i(X)|\}_i$ is the spectral radius of X and $\lambda_i(X)$ refers to the eigenvalues. Furthermore, we denote $A_{\max} = \max_{i,j\in[d]} |A(i,j)|$ and $B_{\max} = \max_{i\in[d],k\in[d_u]} |B(i,k)|$.

Notation. We denote by $K_\star(L)$ as the optimal policy of L. We let $[n] = \{1, .., n\}$. Given two ordered lists $\mathcal{I}_1$ and $\mathcal{I}_2$ we let $\mathcal{I}_2/\mathcal{I}_1 = \{x \in \mathcal{I}_2 | x \notin \mathcal{I}_1\}$ denote their difference. Furthermore, given a vector $x \in \mathbb{R}^d$ and a list $\mathcal{I}$ with entries in $[d]$ we let $x(\mathcal{I})$ denote the vector in $\mathbb{R}^{|\mathcal{I}|}$ which contains the coordinates of $\mathcal{I}$, $x(\mathcal{I}) = \begin{bmatrix} x(\mathcal{I}(1)) & \cdots & x(\mathcal{I}(|\mathcal{I}|)) \end{bmatrix}$. We denote I_d as the identity matrix of dimension d. The spectral/L_2 norm of a matrix is denoted by $||A||_{\text{op}}$ and the Frobenius norm by $||A||_F$. We use $O(X)$ to refer to a quantity that depends on X up to constants, and denote $a \vee b = \max(a, b)$. For a square matrix $A \in \mathbb{R}^{d\times d}$ we denote $\text{size}(A) = d$.

3. The Partially Controllable Linear-Quadratic Control Problem

In this section we formally define the LQ problem we analyze and later derive sample complexity results. We focus on an LQ problem that consists of a partially controllable system and define an explicit notion of irrelevant state variables. Specifically, we establish that these state variables are irrelevant for optimally control this system, and, for that reason, we say the optimal controller of such a system is sparse.

A linear system is said to be *partially controllable* if the controllability matrix $\mathcal{G} = \begin{bmatrix} B & AB & \cdots & A^d B \end{bmatrix}$ is not of a full rank, that is $\text{rank}(\mathcal{G}) = s_c < d$ (e.g., Sontag (2013)). For an LQ problem in such a system, there exists a linear transformation T that transforms the system and cost function to obtain an equivalent LQ control problem $\tilde{L} = (\tilde{A}, \tilde{B}, \tilde{Q})$ with the block structure of (1). This representation reveals that the second block of coordinates A_2^{PC} cannot be affected by the controller inputs. As such, one might hope that A_{12}^{PC} and A_2^{PC} are not required for optimal control. Unfortunately, this is not the case, as we show in the next simple example. Even when $\text{rank}(\mathcal{G}) = 1$ and $Q = I_d$, the optimal policy may depend on the full dynamics of the uncontrollable subsystem (see Appendix C for detailed analysis).

Example 2 (Necessity of uncontrollable dynamics for optimal control)**.** *Let $\rho \in \mathbb{R}^{d-1}, ||\rho||_\infty < 1$,*

$$A_\rho = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \rho_1 & 0 & \cdots & 0 \\ & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \rho_{d-1} \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, Q = I_d,$$

Let $L_\rho = (A_\rho, B, I_d)$ be a stabilizable LQ problem. Then, $K^\star(L_\rho)$ is a function of ρ.

The example highlights that, without further structure, the optimal policy may depend on $\Omega(d)$ parameters of the transition dynamics A even though only a small portion of the system is controllable. Intuitively, this occurs because the uncontrollable system interacts with the controllable one through matrix A_{12}^{PC} in (1), so the optimal controller must plan for and react to the uncontrollable state.

On the other hand, there are many systems in which some uncontrollable state variables *do not* affect the controllable ones whatsoever. The following model captures this scenario; we refer to this model as a Partially Controllable Linear Quadratic (PC-LQ) control problem.[1]

$$A = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}, \; B = \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix}, \; Q = I_d, \tag{3}$$

where $A_1 \in \mathbb{R}^{s_c \times s_c}, A_2 \in \mathbb{R}^{s_e \times s_e}, A_3^{d-s \times d-s}, B_1 \in \mathbb{R}^{s_c \times d_u}$ and $s = s_e + s_c$. The linear system in a PC-LQ problem[2] can be decomposed into three components: a controllable system, an uncontrollable relevant system, and an uncontrollable irrelevant system, where the latter has no interaction with the controllable system. These are the first, second, and third blocks on the diagonal, respectively. Furthermore, A_{12} is a coupling that allows the uncontrollable relevant dynamics to affect the controllable ones, and A_{32} is a coupling that allows the uncontrollable relevant system to affect the irrelevant one. Observe that any LQ control problem can be written in the form of (3), for some s_c and s_e, where, for a general stable system, with no uncontrollable irrelevant dynamics, $s_c + s_e = d$.

If the PC-LQ has $s < d$, then there are variables that are essential for modeling the dynamics that are superfluous for optimal control. Indeed, as we show in the next result, the optimal policy of any PC-LQ problem does not depend on the entire transition dynamics, specifically, the optimal controller is insensitive to the dynamics of the uncontrollable irrelevant subsystem (blocks A_3 and A_{32}). On the other hand, this subsystem can exhibit a very complex temporal structure, so it is important for dynamics modeling/certainty equivalence. Thus, even though the dynamics matrix A is not a low-dimensional object, when $s \ll d$, it is thus apt to say that the optimal policy of a PC-LQ is low-dimensional. The following result explores two invariance properties of the optimal controller in a PC-LQ problem under cost and dynamics transformation (see Appendix D for the proof).

Theorem 1 (Invariance of Optimal Policy for PC-LQ)**.** *Consider the following PC-LQ problems (as in equation* (3)*):*

1. *Let $L_1 = (A, B, I_d), L_2 = (A, B, I_{1+})$ be PC-LQ problems in stabilizable systems with similar dynamics. Let I_{1+} be a diagonal matrix such that (i) if $i \in [d]$ is a coordinate of the first block then $I_{1+}(i,i) = 1$, and, (ii) for any other $i \in [d]$, $I_{1+}(i,i) \in \{0,1\}$.*
2. *Let $L_1 = (A, B, I_d), L_2 = (\bar{A}, B, I_d)$ be PC-LQ problems in stabilizable systems such that*

$$A = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}, \; \bar{A} = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & \bar{A}_{32} & \bar{A}_3 \end{bmatrix}.$$

Then, for both (1) *and* (2)*, the optimal policy of L_1 and L_2 is equal, i.e., $K^*(L_1) = K^*(L_2)$.*

Of course, since $Q = I_d$, the optimal value functions for L_1 and L_2 will – in general – be quite different. Since the uncontrollable blocks A_3 and A_{32} of a PC-LQ are irrelevant to optimally control it, we refer to both of the block as the *irrelevant blocks* from this point onward. This highlights the fact that the LQR of a PC-LQ is sparse: it does not depend on the parameters of the irrelevant blocks.

3.1. Characterization via controllability and the relevant disturbances matrices

A natural question is to understand when a system is equivalent to a PC-LQ with an irrelevant subsystem. The next result provides a characterization of PC-LQ in terms of the controllability matrix and a new object that we call the *relevant disturbances matrix*. Recall that any LQ problem with controllability index s_c can be rotated into the form (1). For brevity, denote $X_{12} = A_{12}^{\mathrm{PC}}$ and $X_2 = A_2^{\mathrm{PC}}$. Let the relevant disturbances matrix using this representation be

$$\mathcal{RD} = \begin{bmatrix} X_{12}^\top & X_2^T X_{12}^\top & \cdots & (X_2^T)^{d-s_c} X_{12}^\top \end{bmatrix}. \tag{4}$$

Then, we have the following structural characterization of a PC-LQ through the controllability and relevant disturbances Krylov matrices (see Appendix E for the proof).

Proposition 2 (Controllability characterization of PC-LQ)**.** *If L has controllability index s_c and rank$(\mathcal{RD}) = s_e$ then $L = (A, B, I_d)$ is rotationally equivalent to* (3)*.*

3.2. Characterization via minimal invariant subspaces

We next characterize a PC-LQ via the notion of minimal invariant subspaces. This characterization is more useful for our subsequent algorithmic development. Minimal invariant subspaces (w.r.t., an initial subspace) are formalized in the next definition.

Definition 3 (Minimal invariant subspace w.r.t. another subspace, e.g., (Basile and Marro, 1992))**.** *Let K be a subspace and $A \in \mathbb{R}^{n \times n}$. Subspace V is an* invariant subspace *of A w.r.t. K if (i), $K \subseteq V$, and (ii) $AV \subset V$. V is the* minimal invariant subspace *of A w.r.t. K if (i) and (ii) hold and V is the subspace with the smallest dimension that satisfies both (i) and (ii).*

That is, the minimal invariant subspace of A w.r.t. K is the smallest subspace that contains K and is closed/invariant under the action of A, meaning that $Av \subset V$ for any $v \in V$. In Appendix F we show that the minimal invariant subspace is always unique, and, thus, it is always well defined.

The next result shows that the first and second blocks of a partially controllable system can be expressed in terms of two minimal invariant subspaces. This yields a simple

[1] Note that the results in this section apply to any system that is rotationally equivalent to (3).

[2] For brevity, we will henceforth use "a PC-LQ" to stand for "a PC-LQ control problem".

Algorithm 1 Learning Optimal Policy of PC-LQ

1: **Require:** $\epsilon, \delta > 0$.
2: **Define:** $\mathrm{STh}_\epsilon(x) = \mathbb{1}\{|x| > \epsilon\}(x - \mathrm{sign}(x)\epsilon)$.
3: Get $\widehat{A}$ and $\widehat{B}$, an (ϵ, δ) element-wise estimates of A and B, respectively.
4: Soft threshold the empirical estimates element-wise, $\bar{B} = \mathrm{Th}_\epsilon(\widehat{B}), \bar{A} = \mathrm{Th}_\epsilon(\widehat{A})$.
5: **Return:** Optimal policy of $\bar{L} = (\bar{A}, \bar{B}, I)$.

algebraic characterization of the relevant components of the system, which we will use to develop algorithms (see Appendix E for the proof).

Proposition 4 (PC-LQ and Minimal Invariant Subspaces)**.** *An LQ problem is equivalent to PC-LQ* (3) *if and only if there exist projection matrices with* $\mathrm{rank}(P_B) \leq \mathrm{rank}(P_c) \leq \mathrm{rank}(P_r)$ *where*

1. *P_c is an invariant subspace of A w.r.t. P_B and* $\mathrm{rank}\,(P_c) = s_c$*,*

2. *P_r is an invariant subspace of $(I - P_c)A^\top$ w.r.t. P_c and* $\mathrm{rank}\,(P_r) = s_c + s_e = s$*,*

such that A, B can be written as

$$A = P_c A P_c + P_r A(P_r - P_c) + (I - P_r)A(I - P_c),$$
$$B = P_B B.$$

The subspaces P_c and P_r are the minimal invariant subspaces if and only if the controllability matrix is of rank s_c and the relevant disturbances matrix is of rank s_e.

With the above notation, the subspace P_c represents the first block of (3), and P_r represents the first two blocks which are generally required for optimally control a PC-LQ. The matrix $(I - P_r)A(I - P_c)$ represents the irrelevant blocks of a PC-LQ which we can safely ignore by Theorem 1.

4. Learning Sparse LQRs in Partially Controllable Systems

We now turn to our main question and focus on the learnability of optimal policy in PC-LQ. We assume that the model is transformed to be in the form of (3), so it is *axis-aligned* up to permutations, i.e., the irrelevant state variables are not a-priori known to the algorithm designer. We further assume $\mathrm{size}(A_1) + \mathrm{size}(A_2) = s_e + s_c = s \ll d$. Of course, as we have discussed, the dynamics matrix A itself *is not sparse*, but the optimal policy of such system, the LQR, is sparse. Theorem 1 establishes the LQR depends only on $O(poly(s))$ parameters. Thus, we hope for sample complexity guarantees that scale primarily with the intrinsic dimension s, rather than the ambient dimension d.

Remark 5 (Axis-aligned assumption)**.** *The axis-aligned assumption is a natural extension of the sparsity assumption made in sparse regression literature (e.g., (Wainwright, 2019), Chapter 7). In control problems, this assumption may be satisfied when the state variables x arise from physical measurements. In this case, axis-alignment corresponds to negligible coupling between different state variables that represent measurements in different locations (as elaborated in Example 1). Furthermore, all the results generalize naturally when the rotation for which the LQ problem can be written as* (3) *is known. We comment that asymptotic dimension-free bounds for system identifications without the axis-aligned assumptions are impossible, due to the need to learn the rotation matrix. We leave it as an interesting future question to study whether asymptotic dimension-free bounds are possible for general PC-LQ problems.*

By Proposition 4 the optimal controller is insensitive to errors in $(I - P_r)A(I - P_c)$, corresponding to block 3 of the dynamics matrix. However, to take advantage of this, we must first identify the zero pattern of the matrix A. More formally, we seek estimates $(\bar{A}, \bar{B})$ of the dynamics satisfying the following *no false positive* property:

$$\forall i, j \in [d],\ k \in [d_u] :\ A(i,j) = 0 \Rightarrow \bar{A}(i,j) = 0, \text{and}$$
$$B(i,k) = 0 \Rightarrow \bar{B}(i,k) = 0. \tag{5}$$

Indeed, in the presence of such a condition, we can ensure that there is no interaction between the relevant and irrelevant parts of the system in the *estimated model*, so that $(\bar{A}, \bar{B})$ is a PC-LQ with a similar block structure to the true dynamics.

A natural way to obtain estimates of (A, B) that satisfy (5) is to perform *soft-thresholding* on an entrywise accurate initial estimate. Note that the soft-thresholding operation does not introduce much additional error. Since many options are available for obtaining the initial estimate, we formalize this via an oracle that we call the *entrywise estimate*. In Section 5, we instantiate this oracle with two different procedures and analyze their sample complexity.

Definition 6 (Entrywise estimator)**.** *We say that $\widehat{X}$ is an (ϵ, δ) entrywise estimator of a matrix $X \in \mathbb{R}^{d_1 \times d_2}$ if with probability at least $1 - \delta$ we have $\max_{i,j} |\widehat{X}(i,j) - X(i,j)| \leq \epsilon$.*

Given access to such an oracle, Algorithm 1 learns an optimal policy in a PC-LQ problem. First, it estimates (A, B) via the entrywise estimator, to obtain $(\widehat{A}, \widehat{B})$. Second, it applies a soft-thresholding to these estimates to get $\bar{A}, \bar{B}$. Finally, it returns the optimal policy of the LQ problem $\bar{L} = (\bar{A}, \bar{B}, I)$.

For the analysis, we require a technical assumption on the L_∞ stability of the irrelevant subsystem A_3.

Assumption 1 (L_∞-stability of irrelevant dynamics)**.** *A_3 is L_∞ stable:* $\max_i \sum_j |A_3(i,j)| = ||A_3||_\infty < 1$.

In addition, our guarantee scales with the operator norm of the optimal value function for the *relevant* subsystem only. Formally, let $L_{1:2} = (A_{1:2}, B_{1:2}, I_{1:2})$ be an s-dimensional LQ problem defined by the first two blocks of (3) and let $P_{\star,1:2}$ be the solution to the Ricatti equation for this system. The guarantee is given as follows (see Appendix G for the proof).

Theorem 7 (Learning the PC-LQ)**.** *Fix $\epsilon, \delta > 0$. Assume access to an entrywise estimator of (A, B) with parameters $(\sqrt{\epsilon/(2s(s+d_u))}, \delta)$, and that Assumption 1 holds. Then, if $\epsilon < 1/||P_{\star,1:2}||_{\text{op}}^{10}$, with probability greater than $1-\delta$ Algorithm 1 outputs a policy $\bar{K}$ such that*

$$J_\star(\bar{K}) \leq J_\star + O(||P_{\star,1:2}||_{\text{op}}^{8}\epsilon).$$

To prove this result we utilize the machinery of Theorem 1, Proposition 4, the perturbation result of (Simchowitz and Foster, 2020), and the no-false positive property of the estimated model.

5. Sample Complexity for Entrywise Estimation

We now instantiate two entrywise estimators and establish their sample complexity guarantees in two settings. First, when the initial state x_0 has a diagonal covariance matrix, we show that a simple second-moment estimator suffices. In the more general setting where the initial state x_0 has PD covariance, we develop an estimator based on semiparametric least-squares. The first estimator has better sample complexity guarantees, while the second estimator is more general.

5.1. Diagonal covariance matrix

When the initial state x_0 has a diagonal covariance matrix, we analyze a simple second-moment estimator. Specifically we estimate the model with

$$\widehat{A} = \frac{1}{N\sigma_0^2}\sum_n x_{1,n}x_{0,n}^\top, \quad \text{and} \quad \widehat{B} = \frac{1}{N}\sum_n x_{1,n}u_{0,n}^\top, \tag{6}$$

given N partial trajectories $\{(x_{0,i}, u_{0,i}, x_{1,i})\}_{i=1}^N$ where $u_{0,i} \sim \mathcal{N}(0, I_{d_u})$. For this estimator we prove the following (see Appendix H.1 for a proof):

Proposition 8 (Entrywise estimation with diagonal covariance)**.** *Assume that $x_0 \sim \mathcal{N}(0, \sigma_0 I_d)$ and that Assumption 1 holds. Denote $\sigma_{\text{eff}} = 1 + A_{\max}\sqrt{s} + \left(1 + B_{\max}\sqrt{d_u}\right)((\sigma/\sigma_0) \vee \sigma)$. Then, given $N = O\left(\frac{\log(\frac{d}{\delta})\sigma_{\text{eff}}^2}{\epsilon^2}\right)$ samples (6) is an entrywise estimator of (A, B) with parameters (ϵ, δ).*

Combining with Theorem 7, we obtain the first shaded row of Table 1 on page 3.

5.2. Positive definite covariance matrix

For the second setting, we only assume that the covariance of x_0 is PD. This, more general setting, is of importance since the stationary measure of a policy may be quite complex, and, in particular, it may induce correlations between the irrelevant and relevant blocks (see Appendix B for further discussion on the need to handle general covariance matrices). In this case, the least-squares estimator of A yields a guarantee in the Frobenius norm, which can be translated into an entrywise estimate. However, the sample complexity of this approach scales as $\text{poly}(d)/\epsilon^2$, which is too large for our purposes. Instead of using classical least-squares, our approach is based on a reduction to *semiparametric least-squares* (Chernozhukov et al., 2016; 2018b;a; Foster and Syrgkanis, 2019), which, as we will see, results in a sample complexity of $1/\epsilon^2 + d/\epsilon$ for entrywise estimation. Observe that here the ambient dimension only appears in the lower order term.

The main idea is as follows: Suppose we wish to learn the (i,j)-th entry of A and assume we have (x_1, x_0) sample pairs from the model $x_1 = Ax_0 + \xi$ where ξ is a zero-mean σ sub-gaussian vector. Then, for any $i \in [d]$,

$$x_1(i) = A(i,j)x_0(j) + \langle A(i,[d]/j), x_0([d]/j)\rangle + \xi_i. \tag{7}$$

If the first and second terms on the RHS were uncorrelated, then a linear regression of $x_1(i)$ onto $x_0(j)$ would yield an unbiased estimate of $A(i,j)$. Unfortunately, these two terms are correlated under our assumptions, so least-squares may be biased. To remedy this, we attempt to decorrelate the two terms using a two-stage regression procedure. The first stage involves high dimensional regression problems, but these errors ultimately only appear in the lower order terms.

Since our results for this problem may be of independent interest, we next study a generalization of the model in (7) and explain the estimator in detail. As a corollary, we obtain a sample complexity guarantee for the entrywise estimator for the PC-LQ.

Semiparametric least-squares. As a generalization of (7), assume that $x \sim \mathcal{N}(0, \Sigma)$ where $\lambda_{\min}(\Sigma) > 0$ and $x \in \mathbb{R}^d$ and let

$$y = \langle w_\star, x_1\rangle + \langle e_\star, x_2\rangle + \xi \tag{8}$$

where $w_\star, x_1 \in \mathbb{R}^{d_w}, e_\star, x_2 \in \mathbb{R}^{d_e}$, $x = [x_1, x_2]^\top$ and ξ is σ sub-Gaussian. Furthermore, $\Sigma = \mathbb{E}\left[x_2 x_2^T\right]$, and Σ/Σ_2 is the Schur complement. By observing tuples sampled from

Algorithm 2 Semiparametric Least Squares

1: **Require:** Dataset $\mathcal{D} = \{(x_{1,n}, x_{0,n})\}_{n=1}^{2N}$ row and column indices $i, j \in [d]$.

2: Reduction to semiparametric LS: $\mathcal{D}_{SP} = \{(y_n, z_{1,n}, z_{2,n})\}_{n=1}^{N}$ where

$$y_n = x_{1,n}(i),\ z_{1,n} = x_{0,n}(j),\ z_{2,n} = x_{0,n}([d]/j).$$

3: Estimate cross correlation $\widehat{L} = \left(\sum_{n=1}^{N} z_{1,n} z_{2,n}^\top\right)\left(\sum_{n=1}^{N} z_{2,n} z_{2,n}^\top\right)^\dagger$.

4: Estimate conditional output $\widehat{c} = \left(\sum_{n=1}^{N} z_{2,n} z_{2,n}^\top\right)^\dagger \left(\sum_{n=1}^{N} y_n z_{2,n}\right)$.

5: Set $\widehat{A}(i,j) = \left(\sum_{n=N+1}^{2N} (z_{1,n} - \widehat{L} z_{2,n})(z_{1,n} - \widehat{L} z_{2,n})^\top\right)^\dagger \left(\sum_{n=N+1}^{2N} (y_n - \langle \widehat{c}, z_{2,n}\rangle)(z_{1,n} - \widehat{L} z_{2,n})\right)$.

6: **Output:** $\widehat{A}(i,j)$

this model $\{y_n, x_{1,n}, x_{2,n}\}_{n=1}^{N}$ we wish to estimate only $w_\star$. To do so, we first estimate $L_\star \in \mathbb{R}^{d_w \times d_w}$ and $c_\star \in \mathbb{R}^{d_e}$, that relate x_2 to the conditional expectation $\mathbb{E}[x_1|x_2]$ and $\mathbb{E}[y|x_2]$, with N samples via standard least-squares. Due to the model Gaussian assumption, it holds that

$$\mathbb{E}[x_1|x_2] = L_\star x_2, \qquad \mathbb{E}[y|x_2] = c_\star^T x_2.$$

When access to exact estimates of these quantities is given, we show in Appendix H.2.1, that the model (8) can be 'orthogonalized' and written as

$$y = \langle w_\star, x_1 - L_\star x_2\rangle + \langle c_\star, x_2\rangle,$$

where $\mathbb{E}[(x_1 - L_\star x_2)x_2^\top] = 0$, so that the two terms on the right hand side are uncorrelated, unlike in the original model. Thus, given estimates $\widehat{L}_N, \widehat{c}_N$, we regress $y - \langle \widehat{c}_N, x_2\rangle$ onto $(x_1 - \widehat{L}_N x_2)$ to get an estimate of $w_\star$. See Algorithm 2 for a description of the algorithm. In the next result, we show that this estimator has leading order error scaling with d_w and only a lower order error term scaling with d_e. Furthermore, we get a minimal dependence in $\lambda_{\min}(\Sigma)$, with similar scaling as in usual OLS analysis (Hsu et al., 2012b) (see Appendix H.2.2 for proof).

Proposition 9 (Semiparametric Least-Squares)**.** *Let $\delta \in (0, e^{-1})$. Consider model (8) and assume that Σ is PD. Denote $\sigma_c^2 = ||w_\star||_{\Sigma/\Sigma_2}^2 + \sigma^2$. Then, if $N \geq O\left(\left(\sigma_c^2/\lambda_{\min}(\Sigma)\right) \vee 1\right) d d_w \log\left(\frac{d}{\delta}\right)$, with probability $1-\delta$, the semiparametric LS estimator $\hat{w}$ of $w_\star$ satisfies*

$$||w_\star - \widehat{w}||_2 \leq O\left(\sqrt{\frac{\sigma^2 d_w \log\left(\frac{1}{\delta}\right)}{N\lambda_{\min}(\Sigma)}} + \frac{\left(\sigma_c^2 \vee \sigma_c\right) d d_w \log\left(\frac{d_w}{\delta}\right)}{N\sqrt{\lambda_{\min}(\Sigma)}}\right).$$

Returning to the PC-LQ setting, we obtain an entrywise estimator for A by applying the semiparametric LS approach on each pair $(i,j) \in [d]^2$. To estimate B, since we can sample u_0 with a diagonal covariance, we can apply the results for the diagonal covariance case. We summarize the sample complexity for entrywise estimation in the next corollary (see Appendix H.2 for proof).

Corollary 10 (Element-wise Estimate, PD Covariance)**.** *Assume $x_0 \sim \mathcal{N}(0, \Sigma)$ and that $\lambda_{\min}(\Sigma) > 0$. Denote $\sigma_c^2 = A_{\max}^2 \lambda_{\max}(\Sigma) + \sigma^2$. Then, if $N \geq O\left(\left(\sigma_c^2/\lambda_{\min}(\Sigma) \vee 1\right) d \log\left(\frac{d}{\delta}\right)\right)$, and $N = O\left(\frac{\sigma^2 \log\left(\frac{d}{\delta}\right)}{\epsilon^2 \lambda_{\min}(\Sigma)} + \frac{d\left(\sigma_c^2 \vee \sigma_c\right)\log\left(\frac{d}{\delta}\right)}{\epsilon\sqrt{\lambda_{\min}(\Sigma)}}\right)$, then the semiparametric LS yields an entrywise estimate of A with parameters (ϵ, δ).*

Combining with Theorem 7, we obtain the second shaded row of Table 1 on page 3.

6. Experiments

We present a proof-of-concept empirical study, to demonstrate the end-to-end statistical advantages of leveraging sparsity in the LQR of a PC-LQ. We generate synthetic systems with marginally stable controllable blocks; the task is to learn a stabilizing controller K (such that $\rho(A + BK) < 1$) from finite samples, in the presence of many irrelevant state coordinates (letting d increase, while holding s and d_u constant). We compare Algorithm 1 with the certainty-equivalent controller obtained from the ordinary least-squares (OLS) estimator for the system's dynamics.

Synthetic PC-LQ problems were generated with i.i.d. standard Gaussian entries (for all $A_1, A_2, A_3, A_{12}, A_{32}, B_1$); the diagonal blocks were normalized by their top singular values so that $\rho(A_1) = 1$, and $\rho(A_2) = \rho(A_3) = 0.9$. We computed $\bar{A}$ from the minimum-norm N-sample OLS estimator, as well as the soft-thresholded semiparametric least-squares estimator from Algorithm 1 (with $\epsilon = 0.1$), and obtained certainty-equivalent controllers $\bar{K}$ by solving the Riccati equation with $\bar{L} := (\bar{A}, B)$. Over 100 trials in each setting, we recorded the fraction of times $\bar{K}$ stabilized the system ($\rho(A + B\bar{K}) < 1$, and $J(\bar{K}) \leq 1.1 \cdot J(K^*)$).

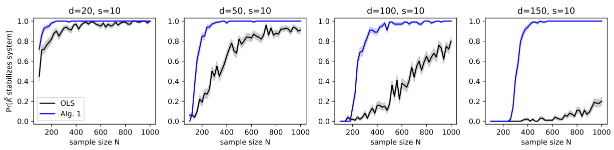

Figure 1. Empirical comparison of Algorithm 1 with OLS for stabilizing a marginally stable PC-LQ. As the number of irrelevant features increases, the sample complexity of the sparsity-leveraging estimator grows much more slowly. Success frequencies (with standard deviations from the normal approximation) are measured over 100 trials.

Figure 1 summarizes our findings: keeping the relevant dimensions fixed ($s_c = s_e = 5, d_u = 1$) and allowing d to grow, the sample complexity of stabilizing the system exhibits a far milder dependence on the ambient dimension d when using our estimator. A complete description of the experimental protocol is given in Appendix J.

7. Related Work

Partial controllability in control theory. The notion of controllability and partial controllability has been well studied from many different aspects in both classical and modern control theory (Kalman, 1963; Lin, 1974; Glover and Silverman, 1976; Jurdjevic and Quinn, 1978; Zhou et al., 1996; Bashirov et al., 2007; Sontag, 2013), as well as, the relation between controllability and invariant subspaces (Klamka, 1963; Basile and Marro, 1992). In Section 3, we characterize which parts of a PC-LQ are not needed for optimal control. To the best of our knowledge, such characterization does not exist in previous literature. One may interpret the results of Section 3 as an extension of Kalman's canonical decomposition. That is, we further decompose the uncontrollable and observable system (see Kalman (1963), Page 165) into relevant and irrelevant parts for optimal control.

Structural results in LQ. Recently, there has been a surge of interest in the learnability of LQ (Abbasi-Yadkori and Szepesvári, 2011; Dean et al., 2019; Sarkar and Rakhlin, 2019; Cohen et al., 2019; Mania et al., 2019; Simchowitz and Foster, 2020; Cassel et al., 2020; Tsiamis and Pappas, 2021). However, learning in the presence of structural properties of an LQ has been, to large extent, unexplored. Closely related to our work is the problem studied in (Fattahi and Sojoudi, 2018; Fattahi et al., 2019). There, the authors considered an LQ problem in which the dynamics itself has a sparse structure. Specifically, the dynamics was assumed to have some sparse block structure such that all elements in each block are simultaneously zero or non-zero. We do not put any such restriction on a PC-LQ. Moreover, in our case, the transition matrix A need not be a sparse matrix, and may have $\Omega(d^2)$ non-zero elements. The sparsity utilized in our work is *sparsity of the optimal controller* and not of the dynamics itself. We also comment that in (Fattahi and Sojoudi, 2018; Fattahi et al., 2019) additional assumptions were made, which are not satisfied in our setting. First, the authors assume a mutual-incoherence condition on the covariance matrix. Additionally, it is assumed that $A(i, j), B(i, j) \geq \gamma > 0$, i.e., that there is a minimal value for the entries of the dynamics. These assumptions are crucial for identification of the non-zero entries; assumptions we do not make in this work (see Appendix B for further discussion on the structure of the covariance matrix in our setting). That is, we recover a near optimal policy without the need to recover the true block structure.

Another related work is the work of (Wang and Yang, 2020), where the authors assumed the dynamics is of low rank and fully controllable. We do not make such an assumption and allow for uncontrollable part to affect the controllable part. Lastly, in (Sun et al., 2020), the authors analyzed system identification via low-rank Hankel matrix estimation. Observe that Hankel based techniques only enable the recovery of the controllable parts of the system, as they are based on a function of $A^n B$. However, to optimally control a stable system, knowledge of the relevant uncontrollable process is also needed (see Example 2).

8. Summary and Future Work

In this work, we studied structural and learnability aspects of the PC-LQ. We characterized an invariance property of the LQR of a PC-LQ. This revealed that the optimal controller of such systems is, in fact, a low-dimensional object. Then, given an entrywise estimator, we showed that the sample complexity of learning an axis-aligned PC-LQ has only a mild dependence on the ambient dimension, scaling primarily with the dimensionality/sparsity of the optimal controller.

The results presented in this work opens several interesting future research avenues. First, we believe it would be interesting to study additional invariance properties of optimal policies of other control and RL problems. As stressed in this work, invariances of the optimal controller can yield statistical improvements for learning in such models. More broadly, is there a general way to characterize such invariances? Second, in this work, we assumed the PC-LQ model is sparse, or, axis-aligned. A natural question would be to study the learnability of such a model when the system is not axis-aligned, and understand the nature of possible sample complexity improvements in such systems? Lastly, extending our results to a single trajectory setting is of interest, and may require developing new tools for semiparametric least-squares analysis.

References

Abbasi-Yadkori, Y. and Szepesvári, C. (2011). Regret bounds for the adaptive control of linear quadratic systems. In *Proceedings of the 24th Annual Conference on Learning Theory*. JMLR Workshop and Conference Proceedings.

Agarwal, A., Kakade, S., Krishnamurthy, A., and Sun, W. (2020). Flambe: Structural complexity and representation learning of low rank mdps. *Advances in Neural Information Processing Systems*.

Bashirov, A. E., Mahmudov, N., Şemı, N., and Etıkan, H. (2007). Partial controllability concepts. *International Journal of Control*.

Basile, G. and Marro, G. (1992). *Controlled and conditioned invariants in linear system theory*. Prentice Hall Englewood Cliffs, NJ.

Cassel, A., Cohen, A., and Koren, T. (2020). Logarithmic regret for learning linear quadratic regulators efficiently. In *International Conference on Machine Learning*, pages 1328–1337. PMLR.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018a). Double/debiased machine learning for treatment and structural parameters.

Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., and Robins, J. M. (2016). Locally robust semiparametric estimation. *arXiv preprint arXiv:1608.00033*.

Chernozhukov, V., Nekipelov, D. N., Semenova, V., and Syrgkanis, V. (2018b). Plug-in regularized estimation of high-dimensional parameters in nonlinear semiparametric models. Technical report, cemmap working paper.

Cohen, A., Koren, T., and Mansour, Y. (2019). Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret. In *International Conference on Machine Learning*. PMLR.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. (2019). On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*.

Deisenroth, M. and Rasmussen, C. E. (2011). Pilco: A model-based and data-efficient approach to policy search. In *Proceedings of the 28th International Conference on machine learning (ICML-11)*. Citeseer.

Fattahi, S., Matni, N., and Sojoudi, S. (2019). Learning sparse dynamical systems from a single sample trajectory. In *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE.

Fattahi, S. and Sojoudi, S. (2018). Sample complexity of sparse system identification problem. *arXiv preprint arXiv:1803.07753*.

Foster, D. J. and Syrgkanis, V. (2019). Orthogonal statistical learning. *arXiv preprint arXiv:1901.09036*.

Glover, K. and Silverman, L. (1976). Characterization of structural controllability. *IEEE Transactions on Automatic control*.

Hao, B., Lattimore, T., Szepesvári, C., and Wang, M. (2021). Online sparse reinforcement learning. In *International Conference on Artificial Intelligence and Statistics*. PMLR.

Hsu, D., Kakade, S., Zhang, T., et al. (2012a). A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*.

Hsu, D., Kakade, S. M., and Zhang, T. (2012b). Random design analysis of ridge regression. In *Conference on learning theory*. JMLR Workshop and Conference Proceedings.

Jurdjevic, V. and Quinn, J. P. (1978). Controllability and stability. *Journal of differential equations*.

Kakade, S., Krishnamurthy, A., Lowrey, K., Ohnishi, M., and Sun, W. (2020). Information theoretic regret bounds for online nonlinear control. *arXiv preprint arXiv:2006.12466*.

Kalman, R. E. (1963). Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*.

Klamka, J. (1963). Controllability of linear dynamical systems. *Contrib. Theory Differ. Equ*, pages 189–213.

Lancaster, P. and Rodman, L. (1995). *Algebraic riccati equations*. Clarendon press.

Lin, C.-T. (1974). Structural controllability. *IEEE Transactions on Automatic Control*.

Mania, H., Jordan, M. I., and Recht, B. (2020). Active learning for nonlinear system identification with guarantees. *arXiv preprint arXiv:2006.10277*.

Mania, H., Tu, S., and Recht, B. (2019). Certainty equivalence is efficient for linear quadratic control. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*.

Perdomo, J. C., Simchowitz, M., Agarwal, A., and Bartlett, P. (2021). Towards a dimension-free understanding of adaptive linear control. *arXiv preprint arXiv:2103.10620*.

Sarkar, T. and Rakhlin, A. (2019). Near optimal finite time identification of arbitrary linear dynamical systems. In *International Conference on Machine Learning*. PMLR.

Simchowitz, M. and Foster, D. (2020). Naive exploration is optimal for online lqr. In *International Conference on Machine Learning*. PMLR.

Smith, R. L. (1992). Some interlacing properties of the schur complement of a hermitian matrix. *Linear algebra and its applications*.

Sontag, E. D. (2013). *Mathematical control theory: deterministic finite dimensional systems*. Springer Science & Business Media.

Stanfel, P., Johnson, K., Bay, C. J., and King, J. (2020). A distributed reinforcement learning yaw control approach for wind farm energy capture maximization. In *2020 American Control Conference (ACC)*. IEEE.

Sun, Y., Oymak, S., and Fazel, M. (2020). Finite sample system identification: improved rates and the role of regularization.

Tropp, J. A. (2012). User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*.

Tsiamis, A. and Pappas, G. J. (2021). Linear systems can be hard to learn. *arXiv preprint arXiv:2104.01120*.

Vershynin, R. (2010). Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*. Cambridge University Press.

Wang, T. and Yang, L. F. (2020). Episodic linear quadratic regulators with low-rank transitions. *arXiv preprint arXiv:2011.01568*.

Zhou, K., Doyle, J. C., Glover, K., et al. (1996). *Robust and optimal control*, volume 40.