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]

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

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

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

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 $\epsilon < 1/||P_{\star,1:2}||_{\mathrm{op}}^{10}$, with probability greater than $1-\delta$ Algorithm 1 outputs a policy $\bar{K}$ such that*

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

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 $\sigma_{\mathrm{eff}} = 1 + A_{\max}\sqrt{s} + \left(1 + B_{\max}\sqrt{d_u}\right)((\sigma/\sigma_0) \vee \sigma)$. Then, given $N = O\left(\frac{\log\left(\frac{d}{\delta}\right)\sigma_{\mathrm{eff}}^2}{\epsilon^2}\right)$ 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 $\mathrm{poly}(d)/\epsilon^2$, 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 $\widehat{A}(i,j) = \left(\sum_{n=N+1}^{2N} (z_{1,n} - \widehat{L} z_{2,n})(z_{1,n} - \widehat{L} z_{2,n})^\top\right)^\dagger \left(\sum_{n=N+1}^{2N} \left(y_n - \langle \widehat{c}, z_{2,n}\rangle\right)(z_{1,n} - \widehat{L} z_{2,n})\right)$.
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 $\sigma_c^2 = ||w_\star||^2_{\Sigma/\Sigma_2} + \sigma^2$. 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 $J(\bar{K}) \leq 1.1 \cdot J(K^*)$).

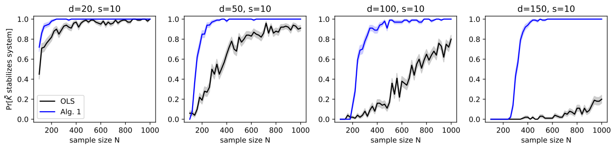

Figure 1. Empirical comparison of Algorithm 1 with OLS for stabilizing a marginally stable PC-LQ. As the number of irrelevant features increases, the sample complexity of the sparsity-leveraging estimator grows much more slowly. Success frequencies (with standard deviations from the normal approximation)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

Algorithm 3 Oracle via second-moments

1: **Require:** $N > 0$, $\sigma_0 > 0$
2: Sample $\mathcal{D} = \{(x_{0,n}, x_{1,n})\}_{n=1}^N$
3: Set $\hat{A} = \frac{1}{N\sigma_0^2}\sum_{n=1}^N x_{1,n}x_{0,n}^\top$
4: **Output:** $\hat{A}$

Algorithm 4 Oracle via Semiparametric LS

1: **Require:** $N > 0$
2: Sample $\mathcal{D} = \{(x_{0,n}, x_{1,n})\}_{n=1}^{2N}$
3: **for** $i, j \in [d]$ **do**
4: Estimate $\widehat{A}(i,j)$ via semiparametric LS
5: for indices (i,j), Algorithm 5
6: **end for**
7: **Output:** $\widehat{A}$

A. Summary of Sample Complexity Results

In Section 4, we study the performance of Algorithm 1, which assumes an oracle access to an $(\sqrt{\epsilon/2s(s+d_u)}), \delta)$ element-wise estimate of the dynamics (A, B). Given such access, Theorem 7 establishes a near-optimal performance guarantee of Algorithm 1.

Then, in Section 5 we study the sample complexity of the assumed (ϵ, δ) element-wise estimate for two settings: when x_0 has a diagonal covariance (Section 5.1) and when x_0 has a PD covariance (Section 5.2).

By combining these together, we get, as a corollary, the sample complexity of the two algorithms considered in this work. That is, when Algorithm 1 is instantiated with Algorithm 3 or with Algorithm 4. We now formally give these corollaries, which Table 1 on page 3 summarizes, for completeness.

Corollary 11 (Learning PC-LQ with second-moment Estimate)**.** *Let the assumptions of Proposition 8 and Theorem 7 hold. Then, given*

$$N = O\left(\frac{\log\left(\frac{d}{\delta}\right)\left(s^2 + d_u s\right)\sigma_{\text{eff}}^2 \vee 1}{\epsilon}\right)$$

samples, the optimal policy $\bar{K}$ of the returned model of Algorithm 1 is at most ϵ suboptimal, where

$$J_\star(\bar{K}) \le J_\star + O\left(||P_{\star,1:2}||^8\epsilon\right).$$

Proof. By Proposition 8, given such amount of samples, the second-moment estimate is an $\left(O(\sqrt{\epsilon/s(s+d_u)}), \delta\right)$ element-wise estimate of the dynamics matrix. Applying Theorem 7 implies the result. □

Corollary 12 (Learning PC-LQ with semiparametric Least Square Estimate)**.** *Let the assumptions of Proposition 9 and Theorem 7 hold. Then, given*

$$N = O\left(\frac{\sigma^2(s^2 + sd_u)\log\left(\frac{d}{\delta}\right)}{\epsilon\lambda_{\min}(\Sigma)} + \frac{d\left(A_{\max}^2\lambda_{\max}(\Sigma) + \sigma^2\right)\sqrt{s^2 + sd_u}\log\left(\frac{d}{\delta}\right)}{\sqrt{\epsilon\lambda_{\min}(\Sigma)}}\right)$$

samples, the optimal policy $\bar{K}$ of the returned model of Algorithm 1 is at most ϵ suboptimal, where

$$J_\star(\bar{K}) \le J_\star + O\left(||P_{\star,1:2}||^8\epsilon\right).$$

Proof. By Proposition 9, given such amount of samples, the semiparametric LS estimate is an $\left(O(\sqrt{\epsilon/s(s+d_u)}), \delta\right)$ element-wise estimate of the dynamics matrix. Applying Theorem 7 implies the result. □

B. Comment on the Structure of Covariance Matrix

In Section 5.2 we devised an entrywise estimator given a general covariance matrix. We further elaborate why this is needed for general PC-LQ problems. We consider two cases, (1) that the sampling policy does not depend on the state variables of the third block, and (2) when they may depend on the state variables of the third block.

Case 1: sampling policy does not depend on the state variables of the third block. In this case, assuming the noise is Gaussian with a diagonal covariance matrix, the covariance matrix has the following block structure

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{12}^\top & \Sigma_{22} & \Sigma_{23} \\ 0 & \Sigma_{23}^\top & \Sigma_{33} \end{bmatrix}.$$

That is, there is a coupling between the state-variables on the second and third block.

Case 2. sampling policy depends on the state variables of the third block. In this case, the covariance matrix may take an arbitrary shape. That is, if the sampling policy is a function of the state variables of the third block, the covariance matrix might have non-zero off-diagonal in the PC-LQ model. Indeed, in lack of prior information on the identity of the non-controllable and non-relevant state variables, the sampling policy may depend on these state variables.

C. Counterexample with a General Uncontrollable System

Consider an LQR model $L_\rho = (A_\rho, B, Q)$ where $|\rho| < 1$

$$A_\rho = \begin{bmatrix} 1 & 1 \\ 0 & \rho \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{9}$$

See that by Theorem 1 the optimal policy of this LQR and the LQR with a modified cost $Q = I_d$ is invariant. See that only the first coordinate of this system is controllable. For simplicity of analysis, we consider $L_\rho = (A_\rho, B, Q)$.

Let $P_{\rho,\star}$ be the solution of the Riccati equation. Then, the optimal policy is then given by

$$K_{x,\star} = (R + B^\top P_{\rho,\star} B)^{-1} (B^\top P_{\rho,\star} A).$$

In Appendix C.1 we solve the Riccati equation, in closed form, and show that

$$P_{\rho,1} = \frac{1+\sqrt{5}}{2}, \text{ and } P_{\rho,12} = \frac{P_1}{P_1^2 - \rho}.$$

This implies that the optimal policy takes the following form,

$$K_{\rho,\star} = \frac{1}{1+P_1} \begin{bmatrix} P_1 & P_1 + \rho P_{\rho,12} \end{bmatrix} = \frac{P_1}{1+P_1} \begin{bmatrix} 1 & \frac{P_1^2}{P_1^2 - \rho} \end{bmatrix}.$$

Observe that since $P_1 > 1$ and $|\rho| < 1$ this object is well defined.

The above implies that for $\rho_1 \neq \rho_2$ it holds that

$$K_{\rho_1,\star} \neq K_{\rho_2,\star}. \tag{10}$$

Hence, the optimal policy is a function of ρ.

Extending the Construction to Arbitrary Dimension To extend the argument to arbitrary dimension consider the d dimensional deterministic LQR problem $L_\rho = (A_\rho, B, Q)$

$$A_\rho = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \rho(1) & 0 & \cdots & 0 \\ 0 & 0 & \rho(2) & 0 & 0 \\ & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \rho(d-1) \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & & 0 \end{bmatrix},$$

where $|\rho| < 1$. As before, the optimal policy of L_ρ and the LQR system $L_\rho = (A_\rho, B, I_d)$ is invariant by Theorem 1: only the first coordinate of this system is controllable. For simplicity we analyze L_ρ.

Observe that if a state variable is initialized as $x_0(i) = 0$ for any $i \in \{2, .., d\}$ then it remains zero, no matter which action is applied, since these coordinates are uncontrollable. Furthermore, since $K_{\rho,\star} \in \mathbb{R}^{d_u \times d}$ induces an optimal policy *for any* state variable, it induces an optimal policy for any such initial state.

Observe that if we initialize the state variable as $x_0(i) = \mathbb{1}\{i = i_0\}$ for some $i_0 \in \{2, .., d\}$ the system is effectively equivalent to the 2-dimensional system of Appendix C.1. For this two dimensional system, we show the optimal controller is a function of ρ, see (10). This establishes the fact that for any two different vectors $\rho_1 \neq \rho_2$ the optimal policy of $L_{\rho_1} = (A_{\rho_1}, B, I_d)$ and $L_{\rho_2} = (A_{\rho_2}, B, I_d)$ is different.

C.1. Solving the Riccati Equation

The Riccati equation, for the above systems, has the following form.

$$\begin{aligned} P_{\rho,\star} &= A_\rho^\top P_{\rho,\star} A_\rho + Q - (B^T P_{\rho,\star} A_\rho)^\top (R + B^T P B)^{-1} B^T P_{\rho,\star} A_\rho \\ &= \begin{bmatrix} P_{\rho,1} & P_{\rho,1} + \rho P_{\rho,12} \\ P_{\rho,1} + \rho P_{\rho,12} & P_{\rho,1} + 2\rho P_{\rho,12} + \rho^2 P_{\rho,2} \end{bmatrix} + I_2 - \frac{1}{1+P_{\rho,1}} \begin{bmatrix} P_{\rho,1}^2 & P_{\rho,1}(P_{\rho,1} + \rho P_{\rho,12}) \\ P_{\rho,1}(P_{\rho,1} + \rho P_{\rho,12}) & (P_{\rho,1} + \rho P_{\rho,12})^2 \end{bmatrix}. \end{aligned}$$

Solving for $P_{\rho,1}$. We solve the Riccati equation for its $(1,1)$ entry. For this entry, we get

$$P_{\rho,1}^2 - P_{\rho,1} - 1 = 0.$$

Solving for $P_{\rho,1}$ we get two solutions, independently of the value of x.

$$P_{\rho,1} \equiv P_1 = \frac{1 \pm \sqrt{5}}{2}.$$

Eventually, we will show that only a single solution is valid among the two.

Solving for $P_{\rho,12}$. We solve the Riccati equation for its $(1,2)$ entry (or, equivalently $(2,1)$). For this entry, we get

$$P_{\rho,12} = \frac{P_1}{1 + P_1 - \rho} = \frac{P_1}{P_1^2 - \rho}. \tag{11}$$

Solving for $P_{\rho,2}$. Finally, we solve the Riccati equation for its $(2,2)$ entry. For this entry, we get

$$P_{\rho,2} = \frac{\left(P_1^5 - P_1\rho^2 - P_1^4\right)/(1-\rho^2)}{(P_1^2 - \rho)^2} \tag{12}$$

Picking a solution. Observe that the eigenvalues of a 2×2 matrix are

$$\lambda_{\pm} = \frac{\operatorname{tr}(A) \pm \sqrt{\operatorname{tr}(A)^2 - 4\det(A)}}{2}. \tag{13}$$

We now show that $P_1 = \frac{1+\sqrt{5}}{2}$ is a PSD solution whereas $P_1 = \frac{1-\sqrt{5}}{2}$ induces a non-PSD $P_{\rho,\star}$.

$P_1 = \frac{1+\sqrt{5}}{2}$ is a PSD solution. We check that $\det(P_{\rho,\star}) \geq 0$ and $\operatorname{tr}(P_{\rho,\star}) \geq 0$. This implies that $P_1 = \frac{1+\sqrt{5}}{2}$ is a PSD solution by (13). We show that $\det(P_{\rho,\star}) \geq 0$. Since $P_{\rho,\star}$ is symmetric, this condition is equivalent to

$$\begin{aligned} &\left(P_1^6 - P_1^2\rho^2 - P_1^5\right)/(1-\rho^2) \geq P_1^2 \\ \iff &P_1^6 - P_1^2\rho^2 - P_1^5 \geq P_1^2(1-\rho^2) \\ \iff &P_1^4 - P_1^3 \geq 1, \end{aligned}$$

which holds since $P_1^4 - P_1^3 \geq 2.6$. We show that $\operatorname{tr}(P_{\rho,\star}) \geq 0$. To show that, it suffices to check that

$$P_1 + \frac{\left(P_1^5 - P_1\rho^2 - P_1^4\right)/(1-\rho^2)}{(P_1^2 - \rho)^2} \geq 0.$$

Since $P_1^5 - P_1\rho^2 - P_1^4 \geq 0$ for $\rho \in (-1,1)$ and $P_1 = (1+\sqrt{5})/2$ we get that $\operatorname{tr}(P_{\rho,\star}) \geq 0$. Hence, $P_1 = \frac{1+\sqrt{5}}{2}$ induces a PSD solution.

$P_1 = \frac{1-\sqrt{5}}{2}$ is not a PSD solution. We show that for this solution, either $\det(P_{\rho,\star}) < 0$ or $\operatorname{tr} P_{\rho,\star} < 0$. This implies, by (13) that the matrix has a negative eigenvalues and thus it is not a PSD matrix. This contradicts the fact $P_{\rho,\star}$ is PSD.

By the above calculation, and since $P_{\rho,\star}$ is symmetric, it holds that

$$\det(P_{\rho,\star}) = P_1 P_{\rho,2} - P_{\rho,12}^2.$$

To show that $\det(P_{\rho,\star}) < 0$ it suffices to show

$$\begin{aligned} &\left(P_1^6 - P_1^2\rho^2 - P_1^5\right)/(1-\rho^2) < P_1^2 \\ \iff &P_1^6 - P_1^2\rho^2 - P_1^5 < P_1^2(1-\rho^2) \\ \iff &P_1^4 - P_1^3 < 1, \end{aligned}$$

which always holds since $P_1^4 - P_1^3 < 0.4$. Thus, $\det(P_{\rho,\star}) < 0$ for $P_1 = \frac{1+\sqrt{5}}{2}$ which implies this solution should be eliminated.

D. Invariance of Optimal Policy of a PC-LQ

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*

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

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

Proof. **First statement.** First, we show that for any fixed and stable policy K, the difference in values between L_1 and L_2 does not depend on the policy K when the cost is transformed $I_d \to I_{1+}$. Fix K which stabilizes A_1 and $x \in \mathbb{R}^d$. We calculate the difference $J_{L_2,K}(x) - J_{L_1,K}(x)$ and show it does not depend on K. It holds that

$$J_{L_2,K}(x) - J_{L_1,K}(x) = \mathbb{E}\left[\sum_{t\geq 1} c_2(x_t) - c_1(x_t) | x_1 = x; K\right] = \mathbb{E}\left[\sum_{t\geq 1}\sum_{i\in\mathcal{I}} ||x_t(i)||_2^2 \mid x_1 = x; K\right],$$

where $\mathcal{I}$ is the set of coordinates for which the diagonal of I_{1+} is zero. That is,

$$\mathcal{I} = \{i \in [d] : I_{1+}(i,i) = 0\}.$$

Observe that for any coordinate $i \in \mathcal{I}$ the state variable $x_t(i)$ is in either the second or third blocks of (3), the coordinates that corresponds to uncontrollable state variables. Thus, $x_t(i)$ for any $i \in \mathcal{I}$ is not affected by the policy K (see Lemma 14). This implies that for any $x \in \mathbb{R}^d$,

$$J_{L_2,K}(x) - J_{L_1,K}(x) = \mathbb{E}\left[\sum_{t\geq 1}\sum_{i\in\mathcal{I}} ||x_t(i)||_2^2 \mid x_1 = x; K\right] = C,$$

i.e., the difference $J_{L_2,K}(x) - J_{L_1,K}(x)$ is constant. This implies that for any x,

$$\arg\min_K J_{L_1,K}(x) = \arg\min_K J_{L_2,K}(x).$$

Hence, the policy $u(x) = K^*(L_1)x$ which is optimal for L_1 is also optimal for L_2.

Second statement. Combining the first statement together with Lemma 13 we prove the claim. That is, consider two alternative PC-LQ problems, $\tilde{L}_1 = (A, B, I_1), \tilde{L}_2 = (\bar{A}, B, I_1)$ where I_1 is diagonal such that

$$I_1(i,i) = \mathbb{I}\{\textit{i belongs to the first block}\}.$$

By Lemma 13 it holds that

$$K^*(L_1) = K^*(\tilde{L}_1), \textit{and}, K^*(L_2) = K^*(\tilde{L}_2).$$

Then, by the first statement it holds that

$$K^*(\tilde{L}_1) = K^*(\tilde{L}_2).$$

Combining the two relations concludes the proof. ∎

Lemma 13 (Invariance of Optimal Policy Under Model Transformation)**.** *Consider the following LQR problems,* $L_1 = (A, B, I_1), L_2 = (\bar{A}, B, I_1)$ *where the dynamics are given by*

$$A = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}, \; \bar{A} = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & \bar{A}_{32} & \bar{A}_3 \end{bmatrix}, \; B = \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix},$$

and,

$$I_1 = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then, the optimal policy of the two models is similar, i.e., $K^*(L_1) = K^*(L_2)$.

To prove this result, we consider the run of the policy iteration algorithm on both L_1 and L_2 for a specific initialization. We show, that there is a conserved structure on both L_1 and L_2 by which we conclude that $P^\star(L_1) = P^\star(L_2)$ by applying Theorem 15. The formal proof is given as follows.

Proof. **Step 1.** Verifying conditions of Theorem 15. First, see that $\tilde{P} = 0$ satisfies the linear inequality in the requirement of Theorem 15. We now show there exists a stable policy for both L_1, L_2.
Let $K_0 \in \mathbb{R}^{d_u \times s_c}$ be a stable policy for (A_1, B_1). Indeed, since (A_1, B_1) are controllable, such policy exists. We first claim that $A + BK_0^E$ where $K_0^E = \begin{bmatrix} K_0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{d_u \times d_x}$ is a stable policy To prove this claim, observe that due to the block structure of $A + BK_0^E$ it holds that

$$\det(A + BK_0^E - \lambda I) = \det(A_1 + B_1 K_0 - \lambda I_1) \det(A_2 - \lambda I_2) \det(A_3 - \lambda I_3),$$

thus, λ is an eigenvalue of $A + BK_0^E$ if and only if it is an eigenvalue of either $A_1 + B_1 K_0$, A_2 or A_3. Since all of these systems are stable, i.e., every eigenvalue is smaller than one, then $A + BK_0^E$ is also stable. Furthermore, since $\bar{A}_3$ is also assumed to be stable, then, by similar reasoning, $\bar{A} + BK_0^E$ is stable.

Step 2. Applying policy iteration on both L_1 and L_2 with the initialized $K_0^{(E)}$. We now apply the policy iteration algorithm on both L_1 and L_2, where we initialize both from $K_0^{(E)}$. Let $K_i^{(1)}, K_i^{(2)}$ be the policies obtained at the i^{th} iteration when running policy iteration on L_1 and L_2, respectively.
The following claim is established via induction: for any iteration $i \geq 0$, it holds that

$$K_i^{(1)} = K_i^{(2)} = [K_{i,1}, \; K_{i,2}, \; 0],$$

i.e., the policy does not depend on the third block. Due to the convergence of policy iteration to the optimal policy, this result will conclude the proof.
Base case. Holds due to the initialization $K_0^{(1)} = K_0^{(2)} = K_0^E$.
Inductive step. Assume the claim holds until the $(i-1)^{th}$ iteration. We prove it holds for the i^{th} iteration. Since the policies at the $(i-1)^{th}$ iteration are equal and does not depend on the third block by the induction hypothesis (the third block is zero) it holds that

$$P_{i-1}^{(1)} = P_{i-1}^{(2)} = P_{i-1} = \begin{bmatrix} P_1 & P_{12} & 0 \\ P_{12} & P_2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

by Lemma 16 for some P_1, P_{12}, P_2. The policy at the i^{th} iteration is given by

$$K_i^{(1)} = (R + B^\top P_{i-1} B) B^T P_{i-1} A$$
$$K_i^{(2)} = (R + B^\top P_{i-1} B) B^T P_{i-1} \bar{A}$$

By a direct calculation due to the form of P_{i-1}, it can be observed that $K_i^{(1)} = K_i^{(2)} = \begin{bmatrix} K_{i,1} & K_{i,2} & 0 \end{bmatrix}$ for some $K_{i,1}, K_{i,2}$, that is, $K_i^{(1)}$ and $K_i^{(2)}$ are equal and both do not depend on the third block. Hence, the induction step is proven, and the lemma follows. □

Lemma 14. *Let $x_t(i)$ be a state vector where i belongs either to the second or third blocks of a PC-LQ. That is, state vector of the uncontrollable coordinates. Then, for any policy K it holds that*

$$\mathbb{E}\left[\sum_{t\geq 1} x_t(i)^2 \mid x_1 = x; K\right] = C,$$

that is, it does not depend on the policy K.

Proof. First, observe that any power n of a block matrix is given by

$$A^{n-1} = \begin{bmatrix} A_1 & X_2 \\ 0 & X_3 \end{bmatrix}^n = \begin{bmatrix} A_1^n & Poly(X_2, X_3, A_1) \\ 0 & X_3^n \end{bmatrix}, \tag{14}$$

where $Poly(X_2, X_3, A_1)$ is some polynomial of the matrices X_2, X_3, A_1. See that the full state vector of any fixed policy K is give by

$$x_t = (A - BK)^t x_0 + \sum_{\tau=1}^{t} (A - BK)^{t-\tau} \xi_\tau,$$

where ξ_τ is an independent i.i.d. and zero mean random vector. Let e_i be a one-hot vector with $e_i(i) = 1$ and zero elsewhere. Due to (14), and since the first block is the only controllable block, we get that

$$e_i^\top (A - BK)^n = e_i^\top \begin{bmatrix} (A_1 - B_1K_1)^n & Poly(X_2, X_3, A_1, B_1K_{12}) \\ 0 & X_3^n \end{bmatrix} = e_i^\top \begin{bmatrix} 0 & 0 \\ 0 & X_3^n \end{bmatrix}, \tag{15}$$

since

$$BK = \begin{bmatrix} B_1K_1 & B_1K_{12} \\ 0 & 0 \end{bmatrix},$$

and since $e_i(j) = 0$ for all coordinates j of the first block. Combining the above we get that

$$\begin{aligned} x_t(i) &= e_i^\top (A - BK)^t x_0 + \sum_{\tau=1}^{t} e_i^\top (A - BK)^{t-\tau} \xi_\tau \\ &= e_i^\top \begin{bmatrix} 0 & 0 \\ 0 & X_3^t \end{bmatrix} x_0 + \sum_{\tau=1}^{t} T \begin{bmatrix} 0 & 0 \\ 0 & X_3^{t-\tau} \end{bmatrix} \xi_\tau \\ &= e_i^\top A_0^t x_0 + \sum_{\tau=1}^{t} A_0^{t-\tau} \xi_\tau, \end{aligned}$$

where

$$A_0 = \begin{bmatrix} 0 & 0 \\ 0 & X_3 \end{bmatrix}$$

does not depend on K. Thus,

$$\mathbb{E}[x_t(i)^2 | K] = \mathbb{E}\left[\left(e_i^\top A_0^t x_0 + \sum_{\tau=1}^{t} A_0^{t-\tau} \xi_\tau\right)^2 | K\right] = C$$

does not depend on the policy K, since ξ_τ is i.i.d. and has the same distribution for all K. $\square$

D.1. Useful Results

Theorem 15 (Asymptotic Convergence of Policy Iteration for LQR, e.g., (Lancaster and Rodman, 1995), Theorem 13.1.1.)**.** *Assume that (A, B) are stabilizable, R invertible, and assume that there is an hermitian solution $\tilde{P}$ to the linear matrix inequality*

$$P \leq A^\top PA + Q - (B^T PA)^\top (R + B^T PB)^{-1} B^T PA,$$

for which $R + B^\top \tilde{P} B > 0$. Then, there exists a unique solution $P^\star$ to the Riccati equation

$$P = A^\top PA + Q - (B^T PA)^\top (R + B^T PB)^{-1} B^T PA, \tag{16}$$

such that $P^\star \geq P$ for all the solutions of (16). Furthermore, the Policy Iteration procedure in which we initialize (K_0, P_{K_0}) with some stable policy and update

$$K_i = (R + B^\top P_{i-1} B) B^T P_{i-1} A, \quad P_i = \sum_{t \geq 0} ((A + BK_i)^\top)^t (Q + K_i^\top R K_i)(A + BK_i)^t$$

converges to $P^\star$.

Lemma 16. *Lee $L = (A, B, I_1)$ where*

$$A = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}, \; B = \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix}, \; I_1 = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Assume that a policy K induces a stable closed loop policy and does not depend on the third block. Then,

$$P_K = \begin{bmatrix} P_1 & P_{12} & 0 \\ P_{12} & P_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

See that if the policy K does have a non-zero component in the third block P_K might have non-zero components at the third row and third column.

Proof. Observe that by the model assumption

$$A + BK = \begin{bmatrix} A_1 + B_1 K_1 & A_{12} + B_1 K_2 & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}.$$

Taking this matrix to some power $t > 0$ we get that

$$(A + BK)^t = \begin{bmatrix} (A_1 + B_1 K_1)^t & V_{1,t}(A_{12}, B_1, K_2, A_2) & 0 \\ 0 & A_2^t & 0 \\ 0 & V_{2,t}(A_2, A_{32}, A_3) & A_3^t \end{bmatrix}, \tag{17}$$

where $V_{1,t}, V_{2,t}$ is some polynomial in its arguments.
We now apply the previous calculation to prove the result. The matrix P_K satisfies the Lyapunov relation

$$\begin{aligned} P_K &= \sum_{t \geq 0} \left((A + BK)^t\right)^\top (I_1 + K^\top RK) \left((A + BK)^t\right) \\ &= \sum_{t \geq 0} \left((I_1 + K^\top RK)^{1/2} (A + BK)^t\right)^\top \left((I_1 + K^\top RK)^{1/2} (A + BK)^t\right). \end{aligned}$$

By a direct computation and by plugging the form of (17) , we see that the matrix $(I_1 + K^\top RK)^{1/2}(A + BK)^t$ have zero elements at the third row and column, that is

$$(I_1 + K^\top RK)^{1/2}(A + BK)^t = \begin{bmatrix} Y_1 & Y_{12} & 0 \\ Y_{21} & Y_2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

for some Y_1, Y_{12}, Y_{21}, Y_2. This also implies that $((I_1 + K^\top RK)^{1/2}(A + BK)^t)^\top (I_1 + K^\top RK)^{1/2}(A + BK)^t$ have zero elements at the third row and column, and, hence,

$$P_K = \sum_{t \geq 0} \left((I_1 + K^\top RK)^{1/2} (A + BK)^t\right)^\top \left((I_1 + K^\top RK)^{1/2} (A + BK)^t\right),$$

have zero elements at the third row an column as well. □

E. Structural Properties of PC-LQ Problems

The following lemma is well known, and is used to properly define the notion of controllable subspace, e.g. (Klamka, 1963; Basile and Marro, 1992; Zhou et al., 1996; Sontag, 2013).

Lemma 17 (E.g., (Sontag, 2013), Lemma 3.3.3.)**.** *Let $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ and*

$$\mathcal{G} = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}.$$

Then, if $\operatorname{rank}(\mathcal{G}) = r < n$ then there exists an invertible transformation such that the matrices $\tilde{A} = TAT^{-1}, \tilde{B} = TB$ have the block structure

$$\tilde{A} = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \tag{18}$$

where $A_1 \in \mathbb{R}^{r\times r}, A_3 \in \mathbb{R}^{d-r\times d-r}$. Conversely, if A and B are given by (18) then $\operatorname{rank}(\mathcal{G}) \le r$.

Proposition 18 (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).*

Proof. By Lemma 17, it holds that the controllable subpace is of rank $\le s_c$ if and only if there exists an invertible transformation U_1 such that

$$T_1AT_1^{-1} = \begin{bmatrix} A_1 & X_2 \\ 0 & X_3 \end{bmatrix}, \quad T_1B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

Apply this transformation and consider the relevant disturbances matrix (4)

$$\mathcal{RU} = \begin{bmatrix} X_2^T & X_3^TX_2^\top & \cdots & (X_3^T)^{d-s_c}X_2^\top \end{bmatrix}.$$

By Lemma 17, while plugging $X_2^\top = B, X_3^\top = A$, it holds that $\operatorname{rank}(\mathcal{RU}) \le s_e$ if and only if there exists an invertible transformation $T_2 \in \mathbb{R}^{d-s_c\times d-s_c}$ such that

$$\begin{aligned} \bar{T}_2X_3^\top\bar{T}_2^{-1} &= \begin{bmatrix} A_2^\top & A_{32}^T \\ 0 & A_3^\top \end{bmatrix}, \quad \bar{T}_2X_2^\top = \begin{bmatrix} A_{12}^\top \\ 0 \end{bmatrix} \\ \iff T_2^{-1}X_3T_2 &= \begin{bmatrix} A_2 & 0 \\ A_{32} & A_3 \end{bmatrix}, \quad X_2T_2 = \begin{bmatrix} A_{12} & 0 \end{bmatrix} \end{aligned} \tag{19}$$

where $T_2 = \bar{T}_2^\top$. Define an invertible transformation extended to $\mathbb{R}^d$,

$$T_3 = \begin{bmatrix} I & 0 \\ 0 & T_2^{-1} \end{bmatrix}, \quad T_3^{-1} = \begin{bmatrix} I & 0 \\ 0 & T_2 \end{bmatrix}.$$

Then, the concatenation T_3T_1 yields the result since,

$$\begin{aligned} T_3T_1AT_1^{-1}T_2^{-1} &= T_3\begin{bmatrix} A_1 & X_2 \\ 0 & X_3 \end{bmatrix}T_3^{-1} \\ &= \begin{bmatrix} A_1 & X_2T_2 \\ 0 & T_2^{-1}X_3T_2 \end{bmatrix} \\ &= \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix}, \end{aligned}$$

where the last relation holds by (19). □

We now prove Proposition 4. This proposition gives an alternative characterization of a PC-LQ relatively to Proposition 2. Specifically, Proposition 4 characterizes a PC-LQ by invariant and minimal invariant subspaces (which we review in

Appendix F) instead of relying on the notion of the controllability matrix and the relevant disturbances matrix. Before supplying with the proof observe that if V is an invariant subspace of A with $\dim(s)$ then, A can be written as

$$A = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \tag{20}$$

in the basis were the first s coordinates span V.

Proposition 19 (PC-LQ and Minimal Invariant Subspaces)**.** *An LQ problem is equivalent to PC-LQ* (3) *if and only if there exist projection matrices with* $\operatorname{rank}(P_B) \leq \operatorname{rank}(P_c) \leq \operatorname{rank}(P_r)$ *where*

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

such that A, B can be written as

$$\begin{aligned} A &= P_c A P_c + P_r A (P_r - P_c) + (I - P_r) A (I - P_c), \\ B &= P_B B. \end{aligned}$$

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.

Given the definition of minimal invariant subspace, the proof is straightforward.

Proof. $\rightarrow$**.** If an LQR is equivalent to a PC-LQ then, there exists some basis such that the dynamics of L is given as

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

It can be observed that, alternatively, in this basis, we can write

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

where $P_B \subseteq P_c \subseteq P_r$, and P_B is the projection on the coordinates on which B has non-zero rows, P_c is a projection on the coordinate of block A_1, and P_r is a projection on the coordinates of the first two blocks. Then, rotating to the original basis does not change this representation.

$\leftarrow$**.** First, rotate B such that P_B is diagonal. In these coordinates,

$$B = \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix}.$$

Since $P_B \subseteq P_c$ it can be jointly diagonalized with P_B. Thus, in this basis, since P_c is an invariant subspace, we can write A as

$$A = \begin{bmatrix} A_1 & X_2 \\ 0 & X_3 \end{bmatrix}, \textit{thus, } (I - P_c)A^\top = \begin{bmatrix} 0 & 0 \\ X_2^\top & X_3^\top \end{bmatrix}.$$

Since $P_c \subseteq P_r$ it can be jointly diagonalized with P_c by a matrix

$$U = \begin{bmatrix} I & 0 \\ 0 & \tilde{U} \end{bmatrix},$$

where $\tilde{U}$ is orthogonal matrix. In this basic, by applying the transformation $(I - P_c)A^\top \rightarrow U(I - P_c)A^\top U^\top$, it holds that

$$(I - P_c)A^\top = \begin{bmatrix} 0 & 0 \\ \tilde{U} X_2^\top & \tilde{U} X_3^\top \tilde{U}^\top \end{bmatrix}.$$

Furthermore, since P_r is an invariant subspace, it must hold that

$$(I - P_c)A^\top = \begin{bmatrix} 0 & 0 & 0 \\ A_{32}^\top & A_2^\top & A_{12}^\top \\ 0 & 0 & A_3^\top \end{bmatrix},$$

since, otherwise, P_r is not an invariant subspace. Lastly, observe that then we can write

$$(I - P_c)A^\top = P_r(I - P_c)A^T P_r + (I - P_r)(I - P_c)A^T P_r + 0 = (P_r - P_c)A^\top P_r + (I - P_c)A^\top (I - P_r),$$

using the fact that $P_c P_r = P_r P_c = P_c$ since $P_c \subseteq P_r$. Combining the above, we get that

$$A = P_c A P_c + P_r A(P_r - P_c) + (I - P_r)A(I - P_c) = \begin{bmatrix} A_1 & A_{12} & 0 \\ 0 & A_2 & 0 \\ 0 & A_{32} & A_3 \end{bmatrix},$$

as we needed to show.

Minimal representation. The last part of the proposition is a corollary of Lemma 23. This lemma establishes that the minimal invariant subspace of A w.r.t. P_B and the span of of the Krylov matrix

$$\mathcal{G} = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}.$$

is equal. $\square$

F. Invariant Subspace and Minimal Invariant Subspace

An invariant subspace V of a matrix $A \in \mathbb{R}^{n\times n}$ satisfies the following definition.

Definition 20 (Invariant Subspace, e.g., (Basile and Marro, 1992), Section 3.2)**.** *Let $A \in \mathbb{R}^{n\times n}$, and V be a subspace of $\mathbb{R}^n$. We say that V is an invariant subspace of A if $AV \subseteq V$.*

Instead of relying on the common definition of invariant subspace (see Definition 20) we give an equivalent and algebraic characterization for this notion. This allows for our proofs to have a more algebraic nature which we found simpler in several proofs along this work.

Proposition 21 (Equivalent Property of Invariant Subspace)**.** *Let $A \in \mathbb{R}^{n\times n}$, V be a subspace of $\mathbb{R}^n$ and $P_V \in \mathbb{R}^{n\times n}$ be the orthogonal projection onto V. The subspace V is an invariant subspace w.r.t. A if and only if $AP_V = P_V AP_V$.*

Proof. **Definition 20 $\rightarrow$ Proposition 21.** Assume that V satisfies Definition 20. We show that it also satisfies Proposition 21. Let P_V be an orthogonal projection on the subspace V, which implies that $P_V = UU^\top$ for some $U \in \mathbb{R}^{N\times \dim(V)}$ with orthogonal columns. Furthermore, let $\{u_i\}_{i=1}^{\dim(V)}$ be the set of orthogonal columns. Since $u_i \in V$ it holds for each u_i, since Definition 20 holds, that

$$\begin{aligned} &Au_i \subseteq V\\ \iff &Au_i = P_V Au_i\\ &\rightarrow Au_iu_i^\top = P_V Au_iu_i^\top. \end{aligned} \tag{21}$$

Summing on all $\dim(V)$ equations we conclude the proof of this part since,

$$\begin{aligned} AP_V &= \sum_{i=1}^{\dim(V)} Au_iu_i^\top && (P_V = \textstyle\sum_{i=1}^{\dim(V)} u_iu_i^\top)\\ &= \sum_{i=1}^{\dim(V)} P_V Au_iu_i^\top && \text{(Equation (21))}\\ &= P_V A \sum_{i=1}^{\dim(V)} u_iu_i^\top = P_V AP_V. && (P_V = \textstyle\sum_{i=1}^{\dim(V)} u_iu_i^\top) \end{aligned}$$

Proposition 21 $\rightarrow$ Definition 20. Assume that V satisfies Proposition 21. We show it also satisfies Definition 20. Observe that by Proposition 21 it holds that $P_V AP_V = AP_V$. Multiplying this relation by any $v \in V$ from both sides we get.

$$AP_V v = P_V AP_V v.$$

Furthermore, by Lemma 24 for any $v \in V$ it holds that $P_V v = v$. Thus,

$$Av = P_V Av.$$

This also implies that $Av \in V$, since, by Lemma 24, any vector that satisfies $P_V u = u$ is contained within V, thus, $Av \in V$ since $P_V(Av) = Av$. $\square$

The notion of minimal invariant subspace is given in Definition 3. For such a definition to be valid, one needs to show that the minimal subspace is unique. The following result establishes this fact. That is, the minimal invariant subspace is unique, and, thus, it is a well defined notion; there are no two minimal invariant subspaces of A w.r.t. a subspace K.

Proposition 22 (Minimal Invariant Subspace is Unique)**.** *Let K be a subspace and $A \in \mathbb{R}^{n\times n}$. If V_1 and V_2 are both minimal invariant subspaces of A w.r.t. K then $V_1 = V_2$.*

Proof. Assume that $V_1 \neq V_2$ and both are minimal invariant subspaces of A w.r.t. K. We show there exists a smaller invariant subspace then both V_1 and V_2, and, thus, get a contradiction to the assumption $V_1 \neq V_2$.

By the requirement (3) of Definition 3 $\dim(V_1) = \dim(V_2)$. Thus, the subspaces V_1/V_2 and V_1/V_2 are non-empty. Furthermore, since both are invariant subspaces it holds that

$$P_{V_1} A P_{V_1} = A P_{V_1}, \textit{and}, \; P_{V_2} A P_{V_2} = A P_{V_2}. \tag{22}$$

Let $P_{V_1} = P_{V_1 \cap V_2} + P_{V_1/V_2}$ and $P_{V_2} = P_{V_1 \cap V_2} + P_{V_2/V_1}$. First, by multiplying the first and second relations of (22) by $P_{V_1 \cap V_2}$ from the right and using $P_{V_1 \cap V_2} P_{V_2/V_1} = P_{V_1 \cap V_2} P_{V_1/V_2} = 0$, since the subspaces are orthogonal, we get

$$\begin{aligned} P_{V_1} A P_{V_1 \cap V_2} &= A P_{V_1 \cap V_2} \\ P_{V_2} A P_{V_1 \cap V_2} &= A P_{V_1 \cap V_2} \end{aligned} \tag{23}$$

which implies that

$$P_{V_1} A P_{V_1 \cap V_2} = P_{V_2} A P_{V_1 \cap V_2}.$$

Multiplying this relation by P_{V_1/V_2} from the left and using $P_{V_1/V_2} P_{V_1} = P^2_{V_1/V_2} = P_{V_1/V_2}$ and $P_{V_1/V_2} P_{V_2} = 0$ we get

$$P_{V_1/V_2} A P_{V_1 \cap V_2} = 0. \tag{24}$$

This relation, together with (23) implies the following.

$$\begin{aligned} A P_{V_1 \cap V_2} &= P_{V_1} A P_{V_1 \cap V_2} && \text{(Equation (23))} \\ &= (P_{V_1 \cap V_2} + P_{V_1/V_2}) A P_{V_1 \cap V_2} \\ &= P_{V_1 \cap V_2} A P_{V_1 \cap V_2}. && \text{(Linearity and Equation (24))} \end{aligned}$$

These relations imply that

$$A P_{V_1 \cap V_2} = P_{V_1 \cap V_2} A P_{V_1 \cap V_2},$$

i.e., $V_1 \cap V_2$ is an invariant subspace of A w.r.t. to K. Observe that $K \subseteq V_1 \cap V_2$: both V_1 and V_2 includes the subspace K, by definition, and, thus, their intersection includes K. Hence, we found an invariant subspace of A, $V_1 \cap V_1$, that includes K, and is strictly smaller than V_1, since V_1/V_2 is non empty. Since $\dim(V_1) = \dim(V_2)$ it also implies that $V_1 \cap V_2$ has smaller dimension than V_2 as well. This implies a contradiction, since we assumed that V_1 and V_2 are minimal subspace of A w.r.t. K. □

The next result establishes a relation between the minimal invariant subspace w.r.t. an initial subspace and the span of a Krylov matrix. This allows us to draw a correspondence between the notion of minimal invariant subspace and, e.g., controllable subspace.

Lemma 23 (Equivalent of the Span of Krlyov Matrices and Minimal Invariant Subspace)**.** *Let $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ and*

$$\mathcal{G} = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}.$$

Then, the span of $\mathcal{G}$ and the minimal invariant subspace of A w.r.t. B are equal.

Proof. Let $B = U\Lambda V^\top$ be the SVD decomposition of B. Then, let $P_B = UU^\top$ and let V_B be the span of B.

$$\mathcal{G}_{P_B} = \begin{bmatrix} P_B & AP_B & \cdots & A^{n-1}P_B \end{bmatrix},$$

and observe that the span of $\mathcal{G}$ and $\mathcal{G}_{P_B}$ is equal. Let V_c and V_m be the span of $\mathcal{G}_{P_B}$ and the span of the minimal invariant subspace of A w.r.t. V_B.

$V_c \subseteq V_m$**.** Since the minimal invariant subspace is an invariant subspace w.r.t. P_B (that is $P_m P_B = P_B$) it satisfies that

$$AP_B = P_m A P_B, \; \textit{and} \; AP_m = P_m A P_m.$$

This implies that for any n

$$A^{n-1}P_B = A^{n-1}P_m A P_m P_B = A^{n-2} P_m A P_m A P_B = \cdots (P_m A P_m)^n P_B^n.$$

Hence,

$$\mathcal{G}_{P_B} = \begin{bmatrix} P_B & AP_B & \cdots & A^{n-1}P_B \end{bmatrix} = P_m \mathcal{G}_{P_B},$$

which implies that $V_c \subseteq V_m$.

$V_m \subseteq V_c$**.** Since V_c is the span of $\mathcal{G}_{P_B}$ it holds that

$$\mathcal{G}_{P_B} = P_c \mathcal{G}_{P_B} \tag{25}$$

This relation implies that for all $i \in [n-1] \cup \{0\}$

$$A^i P_B = P_c A^i P_B. \tag{26}$$

Observe that by the Cayley Hamilton theorem, the n^{th} power can be written as the following sum, for some set of coefficients

$$\begin{aligned} A^n P_B &= \sum_{i=0}^{n-1} \alpha_i A^i P_B \\ &= \sum_{i=0}^{n-1} \alpha_i P_c A^i P_B && \text{(By (25))} \\ &= P_c \sum_{i=0}^{n-1} \alpha_i A^i P_B = P_c A^n P_B && (27) \end{aligned}$$

Thus, together with (26), we get that for all $i \in [n] \cup \{0\}$

$$A^i P_B = P_c A^i P_B. \tag{28}$$

Observe that

$$\begin{aligned} A\mathcal{G}_{P_B} &= AP_c \mathcal{G}_{P_B} && \text{(By (25))} \\ &= \begin{bmatrix} AP_B & A^2P_B & \cdots & A^nP_B \end{bmatrix} \\ &= P_c \begin{bmatrix} AP_B & A^2P_B & \cdots & A^nP_B \end{bmatrix} && \text{(By (28))} \\ &= P_c A \begin{bmatrix} P_B & AP_B & \cdots & A^{n-1}P_B \end{bmatrix} \\ &= P_c A \mathcal{G}_{P_B}. \end{aligned}$$

Since $\mathcal{G}_{P_B} \mathcal{G}_{P_B}^\dagger = U^\top \Sigma^{1/2} V^T V \Sigma^{-1/2} U^\top = UU^\top = P_c$ the above relation implies that

$$AP_c = P_c A P_c, \tag{29}$$

by multiplying by $\mathcal{G}_{P_B}^\dagger$ from the RHS. By Proposition 21 this suggests that P_c is an invariant subspace.

From the above, we get that V_c is an invariant subspace. Furthermore, due to the form of $\mathcal{G}_{P_B}$, it must contain the span of B, V_B. Since the minimal invariant subspace V_m is the smallest subspace that contains V_B and is an invariant subspace w.r.t. A, we get that $V_m \subseteq V_c$. This conclude the proof since it holds that $V_m \subseteq V_c$ and $V_c \subseteq V_m$, which implies that $V_m = V_c$. □

F.1. Linear Algebra Facts

Lemma 24. *Let P_V be an orthogonal projection onto V. Then, $v \in V$ if and only if*

$$P_V v = v.$$

Proof. $\rightarrow$**.** We prove that $P_V v = v$ implies that $v \in V$. Write $P_V = UU^\top$ where U is a matrix with orthonormal columns $\{u_i\}_{i=1}^{\dim(V)}$ and u_i span V. With this notation, $P_V v = v$ implies that

$$v = \sum_{i=1}^{\dim(V)} \langle u_i, v \rangle u_i,$$

hence, v is in the span of V since we can write it as $v = \sum_{i=1}^{\dim(V)} \alpha_i u_i$ and $\{u_i\}_{i=1}^{\dim(V)}$ span V.

$\leftarrow$. We prove that if $v \in V$ then $P_V v = v$. Since $v \in V$ then it can be written as a linear combination of $\{u_i\}_{i=1}^{\dim(V)}$,

$$v = \sum_{i=1}^{\dim(V)} \alpha_i u_i.$$

Since $P_V u_i = u_i$ and by the linearity of orthogonal projection we conclude the proof since

$$P_V v = \sum_{i=1}^{\dim(V)} \alpha_i P_V u_i = \sum_{i=1}^{\dim(V)} \alpha_i u_i = v. \qquad \square$$

G. Learning Sparse LQRs in Partially Controllable Systems

We now establish the correctness of Algorithm 1 given an (ϵ, δ) element-wise oracle (see Definition 6).

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 $\epsilon < 1/||P_{\star,1:2}||_{\mathrm{op}}^{10}$, with probability greater than $1-\delta$ Algorithm 1 outputs a policy $\bar{K}$ such that*

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

Proof. **Consequence of thresholded estimation.** Assume that $\widehat{A}$ and $\widehat{B}$ is an (ϵ, δ) entrywise estimator Definition 6 of A and B, and condition on the event it satisfies the entrywise estimation property. Then, the soft thresholded matrices $(\bar{A}, \bar{B})$ of $(\widehat{A}, \widehat{B})$ satisfy that

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

By this property, and since the true dynamics is of the form given in Proposition 4, the estimates $\bar{A}, \bar{B}$ can be written as follows

$$\bar{A} = P_{\mathcal{I}_c}\bar{A}P_{\mathcal{I}_c} + P_{\mathcal{I}_r}\bar{A}(P_{\mathcal{I}_r} - P_{\mathcal{I}_c}) + (I - P_{\mathcal{I}_r})\bar{A}(I - P_{\mathcal{I}_c}), \quad \bar{B} = P_{\mathcal{I}_B}\bar{B}. \tag{30}$$

Invariance argument for estimated system. Let $K_\star(\bar{L})$ be the optimal policy of the LQR system $\bar{L} = \left(\bar{A}, \bar{B}, I_d\right)$. This LQR system is also a PC-LQ system by comparing (30) and the form supplied in Proposition 4. Observe that $\bar{L}$ is a stabilizable PC-LQ.

1. The system that contains the first two blocks of $\bar{L}$ is stabilizable by utilizing the perturbation result of (Simchowitz and Foster, 2020) as we formally establish below in $(*stb)$.

2. The uncontrolled and non-relevant system $(I - P_{\mathcal{I}_r})\bar{A}(I - P_{\mathcal{I}_r})$ is stable since

$$||(I - P_{\mathcal{I}_r})\bar{A}(I - P_{\mathcal{I}_r})||_\infty \leq ||(I - P_{\mathcal{I}_r})A(I - P_{\mathcal{I}_r})||_\infty \leq 1.$$

   The first inequality holds due the soft thresholding which implies that $|\bar{A}(i,j)| \leq |A(i,j)|$ which leads to the inequality. The second inequality holds by Assumption 1. Since $\rho((I - P_{\mathcal{I}_r})\bar{A}(I - P_{\mathcal{I}_r})) \leq ||(I - P_{\mathcal{I}_r})\bar{A}(I - P_{\mathcal{I}_r})||_\infty \leq 1$ we get that uncontrolled and non-relevant is stable.

By the first and second statement of Theorem 1 the optimal policy is invariant under a change in the dynamics and cost. Let $\bar{L}_{\mathrm{inv}} = \left(\bar{A}_{\mathrm{inv}}, \bar{B}, I_{1:2}\right)$

$$\bar{A}_{\mathrm{inv}} = P_{\mathcal{I}_c}\bar{A}P_{\mathcal{I}_c} + P_{\mathcal{I}_r}\bar{A}(P_{\mathcal{I}_r} - P_{\mathcal{I}_c}), \quad \bar{B} = P_{\mathcal{I}_B}\bar{B},$$

that is, when we set $(I - P_{\mathcal{I}_r})\bar{A}(I - P_{\mathcal{I}_c}) = 0$, and the cost

$$I_{1:2} = \begin{bmatrix} I_{s_c} & 0 & 0 \\ 0 & I_{s_e} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

that is, we set the cost of the third block to zero ($I_{1:2}$ is a subset of I_{1+} defined in Theorem 1). By Theorem 1 it holds that

$$K_\star(\bar{L}) = K_\star(\bar{L}_{\mathrm{inv}}). \tag{31}$$

Invariance argument for the true system. By again applying the first and second statement of Theorem 1, we get that the optimal policy of the true system is invariant when transforming it to the LQR system $L_{\mathrm{inv}} = (A_{\mathrm{inv}}, B, I_{1:2})$ where

$$A_{\mathrm{inv}} = P_{\mathcal{I}_c}AP_{\mathcal{I}_c} + P_{\mathcal{I}_r}A(P_{\mathcal{I}_r} - P_{\mathcal{I}_c}), \quad B = P_{\mathcal{I}_B}B.$$

That is,

$$K_\star(L) = K_\star(L_{\text{inv}}). \tag{32}$$

Perturbation result on invariant systems. We now apply a perturbation result of (Simchowitz and Foster, 2020), Theorem 5 (which we partially restate in Theorem 25 for convenience) on the invariant systems $\bar{L}_{\text{inv}}$ and L_{inv}. First, observe that for both $\bar{L}_{\text{inv}}, L_{\text{inv}}$ the optimal value has the following form

$$P = \begin{bmatrix} P_1 & P_{12} & 0 \\ P_{12} & P_2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

since the cost of the third block is zero $I_{1:2}$, and the dynamics of the third row and column is zero on the invariant systems L_{inv} and $\bar{L}_{\text{inv}}$. Thus, we can eliminate the third row and third columns of the LQR systems $\bar{L}_{\text{inv}}$ L_{inv} and apply a perturbation bound on the smaller system. Let $\bar{L}_{\text{inv},1:2}, L_{\text{inv},1:2}$ be this restriction.

Observe that the errors of $\bar{L}_{\text{inv},1:2}$ relatively to $L_{\text{inv},1:2}$ scales with $\sqrt{s(s+d_u)}\epsilon$, i.e.,

$$\begin{aligned}
&||\bar{A}_{\text{inv},1:2} - A_{\text{inv},1:2}||_F = ||\bar{A}_{P_{\mathcal{I}_c}}(\bar{A}-A)P_{\mathcal{I}_c} + P_{\mathcal{I}_r}(\bar{A}-A)(P_{\mathcal{I}_r} - P_{\mathcal{I}_c})||_F \le \sqrt{2}s\epsilon \\
&||\bar{B}_{1:2} - B_{1:2}||_F \le \sqrt{2sd_u}\epsilon,
\end{aligned}$$

where the $\sqrt{2}$ factor comes from the soft thresholding operations together with the (ϵ,δ) element-wise estimation of (A,B). Setting $\epsilon = \sqrt{\epsilon'/2s(s+d_u)}$ in the element-wise estimation of (A,B), and renaming ϵ' as ϵ, we get that

$$\begin{aligned}
&\max\left\{||\bar{A}_{\text{inv},1:2} - A_{\text{inv},1:2}||_F, ||\bar{B}-B||_F\right\} \le \sqrt{\epsilon} \\
&\max\left\{||\bar{A}_{\text{inv},1:2} - A_{\text{inv},1:2}||_{\text{op}}, ||\bar{B}-B||_{\text{op}}\right\} \le \max\left\{||\bar{A}_{\text{inv},1:2} - A_{\text{inv},1:2}||_F, ||\bar{B}-B||_F\right\} \le \sqrt{\epsilon},
\end{aligned}$$

since for any matrix $||A||_{\text{op}} \le ||A||_F$.

By Theorem 5 of (Simchowitz and Foster, 2020) (see Theorem 25) we get that if $\sqrt{\epsilon} \le 54||P_\star(L_{\text{inv},1:2})||_{\text{op}}^5$ then the optimal policies of $\bar{L}_{\text{inv},1:2}$ and $L_{\text{inv},1:2}$ are close and both system are stabilizable (specifically, the first two block of the estimated system $\bar{L}_{\text{inv},1:2}$ is stable as was needed to show in $(*stb)$). That is,

$$J_\star(K_\star(\bar{L}_{\text{inv},1:2}); L_{\text{inv},1:2}) \le J_\star(K_\star(L_{\text{inv},1:2}); L_{\text{inv},1:2}) + 2C_{\text{est}}(A,B)\epsilon, \tag{33}$$

where $C_{\text{est}}(A,B) = 142||P_\star(L_{\text{inv}})||_{\text{op}}^8$.

Since the optimal policies of the system L, $L_{\text{inv},1:2}$ and $\bar{L}$, $\bar{L}_{\text{inv},1:2}$ is invariant, the above implies that,

$$\begin{aligned}
J_\star(K_\star(\bar{L}); L_{\text{inv},1:2}) &= J_\star(K_\star(\bar{L}_{\text{inv},1:2}); L_{\text{inv},1:2}) && \text{(By (31))} \\
&\le J_\star(K_\star(L_{\text{inv},1:2}); L_{\text{inv},1:2}) + 2C_{\text{est}}(A,B)\epsilon && \text{(By (33))} \\
&= J_\star(K_\star(L); L_{\text{inv},1:2}) + 2C_{\text{est}}(A,B)\epsilon. && \text{(By (32))}
\end{aligned}$$

Lastly, since the difference in values between the invariant and original system is a constant, that does not depend on the policy, by the first statement of Theorem 1, it holds that

$$\begin{aligned}
J_\star(K_\star(\bar{L}); L_{\text{inv},1:2}) &= J_\star(K_\star(\bar{L}); L) + C = J_\star(K_\star(\bar{L})) + C \\
J_\star(K_\star(L); L_{\text{inv},1:2}) &= J_\star(K_\star(L); L) + C = J_\star + C.
\end{aligned}$$

Combining the above yields that

$$J_\star(K_\star(\bar{L})) \le J_\star + 2C_{\text{est}}(A,B)\epsilon.$$

□

Theorem 25 ((Simchowitz and Foster, 2020), Theorem 5)**.** *Let $L = (A, B, I_d)$ be a stabilizable system. Given an alternative pair of matrices $\bar{L} = (\bar{A}, \bar{B}, I_d)$, for each $\circ \in \{\text{op}, F\}$ define $\epsilon_\circ = \max\left\{||A - \widehat{A}||_\circ, ||B - \widehat{B}||_\circ\right\}$. Then, if $\epsilon_{\text{op}} \le 1/54||P_\star||_{\text{op}}^5$*

$$J_\star(K_\star(\bar{L})) \le J_\star + C_{\text{est}} J\epsilon_F^2,$$

where $C_{\text{est}} = 142||P_\star||^8$.

H. Learning Element-wise Estimates of a Matrix

H.1. Diagonal Covariance Matrix

In this section, we analyze the sample complexity of of obtaining an (ϵ, δ) element-wise good estimate of a matrix assuming that the covariance matrix of x_0 is diagonal. Before presenting the results, we define the quantity $\gamma_{d,\delta} \equiv \sqrt{Cd \log\left(\frac{1}{\delta}\right)}$, where $C = ||x_i||_\psi$ is the subgaussian parameters of x_i and is bounded for Gaussian distribution by a constant. This quantity controls the concentration of the covariance estimation and is defined in Lemma 42

Proposition 26 (Entrywise estimation with diagonal covariance)**.** *Assume that $x_0 \sim \mathcal{N}(0, \sigma_0 I_d)$ and that Assumption 1 holds. Denote $\sigma_{\text{eff}} = 1 + A_{\max}\sqrt{s} + \left(1 + B_{\max}\sqrt{d_u}\right)\left((\sigma/\sigma_0) \vee \sigma\right)$. Then, given $N = O\left(\frac{\log\left(\frac{d}{\delta}\right)\sigma_{\text{eff}}^2}{\epsilon^2}\right)$ samples (6) is an entrywise estimator of (A, B) with parameters (ϵ, δ).*

This result is a direct corollary of Lemma 27 as we now show.

Proof. Observe that we apply random inputs of the form $u_0 \sim \mathcal{N}(0, I_d)$, that $x_0 \sim \mathcal{N}(0, \sigma_0 I_d)$ and that ξ is σ subgaussian. Thus,

$$x_1 = Ax_0 + Bu_0 + \xi. \tag{34}$$

Estimation of A. The estimator of A is given by

$$\widehat{A} = \frac{1}{N\sigma_0^2}\sum_n x_{1,n}x_{0,n}^T,$$

where, by (34)

$$x_{1,n} = Ax_{0,n} + \xi_{B,n}$$

and $\xi_{B,n} = \xi_0 + Bu_{0,n}$. Observe that u_0 and x_0 are i.i.d., and, for any i, $\xi_{B,n}(i)$ is a zero mean σ_B sub gaussian noise where

$$\sigma_B = \sqrt{\sigma^2 + ||B(i,\cdot)||_2^2} \leq \sqrt{\sigma^2 + B_{\max}^2 d_u}$$

Applying Lemma 27 directly implies that

$$\mathbb{P}\left(\forall i,j \in [d] : \left|\hat{A}(i,j) - A(i,j)\right| \geq \frac{5\gamma_{2,\delta/(6d^2)}(\sigma_B + \sigma_0 \max_i ||A(i,\cdot)||_2}{\sigma_0\sqrt{N}}\right) \leq \delta.$$

Estimation of B. The analysis is similar to the first part. The estimator of B is given by

$$\widehat{B} = \frac{1}{N}\sum_n x_{1,n}u_{0,n}^T.$$

By (34), we see that $x_{1,n}$ can be written as

$$x_{1,n} = Bu_{0,n} + \xi_{A,n},$$

and $\xi_{A,n} = \xi_0 + Ax_{0,n}$. Since x_0 and u_0 are i.i.d., and for any i it holds that $\xi_{A,n}(i)$ is zero mean σ_A sub gaussian noise where

$$\sigma_A = \sqrt{\sigma^2 + ||A(i,\cdot)||_2^2} \leq \sqrt{\sigma^2 + A_{\max}^2 s + 1},$$

since if i is in the third block it holds that $\sum_j (A_3(i,j))^2 \leq \left(\sum_j |A_3(i,j)|\right) \leq 1$ by Assumption 1. Applying Lemma 27 directly implies that

$$\mathbb{P}\left(\forall i \in [d], j \in [d_u] : \left|\hat{B}(i,j) - B(i,j)\right| \geq \frac{5\gamma_{2,\delta/(6d^2)}(\sigma_A + \max_i ||B(i,\cdot)||_2}{\sqrt{N}}\right) \leq \delta.$$

Taking a union bound concludes the proof. $\square$

Lemma 27 (Elementwise Convergence of second-moment Based Estimation)**.** *Let $\epsilon, \delta > 0$. Let the plug-in estimator of A be given as*

$$\widehat{A} = \frac{1}{\sigma_0^2 N}\sum_{i=1}^{N} x_{1,n} x_{0,n}^\top,$$

where $x_1 = Ax_0 + \xi$, $x_0 \sim \mathcal{N}(0, \sigma_0^2 I)$ and for any $i \in [d]$ it holds that $\xi(i)$ is σ sub gaussian. Then,

$$\mathbb{P}\left(\forall i,j \in [d] : |\widehat{A}(i,j) - A(i,j)| \geq \frac{5\gamma_{2,\delta/(6d^2)}(\sigma + \sigma_0||A(i,\cdot)||_2)}{\sigma_0\sqrt{N}}\right) \leq \delta.$$

Proof. Observe that

$$\widehat{A} = A + \underbrace{\frac{1}{N}\sum_n A\left(\frac{x_{0,n}x_{0,n}^\top}{\sigma_0^2} - I\right)}_{(i)} + \underbrace{\frac{1}{\sigma_0^2 N}\sum_n \xi_n x_{0,n}^\top}_{(ii)} \tag{35}$$

We get a point-wise bound for each one of the terms to conclude the proof.
Term (i)**.** Let $z_n = \frac{1}{\sigma_0}x_{0,n}$ and observe it is $\mathcal{N}(0, I_d)$ gaussian random vector. Fix $i, j \in [d]$. It holds that the i, j entry of term (i) can be written as follows.

$$\left[\frac{1}{N}\sum_n A(z_n z_n^\top - I)\right]_{ij} \tag{36}$$

$$= \underbrace{\frac{1}{N}\sum_n \sum_{m\neq j} A(i,m)z_n(m)z_n(j)}_{(i)} + \underbrace{\frac{1}{N}\sum_i A(i,j)(z_n(j)^2 - 1)}_{(ii)}\,. \tag{37}$$

To bound the first term of (37), we write it as follows

$$\frac{1}{N}\sum_n \sum_{m\neq k} A(i,m)z_n(m)z_n(j) = \frac{1}{N}\sum_n \langle A(i,[d]/j), z([d]/j)\rangle\, z_n(j).$$

Observe that $\mathbb{E}[\langle A(i,[d]/j), z([d]/j)\rangle\, z_n(j)] = \mathbb{E}[\langle A(i,[d]/j), z([d]/j)\rangle]\mathbb{E}[z_n(j)]$ since the first term the vector $z([d]/j)$ does not contain $z_n(j)$, and, thus, the two are independent. By Lemma 28 we get that with probability at least $1-\delta$ it holds that

$$\left|\frac{1}{N}\sum_n \sum_{m\neq j} A(i,m)x_{0,n}(m)x_{0,n}(j)\right| = \left|\frac{1}{N}\sum_n \langle A(i,[d]/j), z_n([d]/k)\rangle\, z_n(j)\right| \leq \frac{3||A(i,[d]/j)||_2\gamma_{2,\delta/3}}{\sqrt{N}},$$

for $N \geq \gamma_{2,\delta/3}$.

We bound the second term of (37) by directly applying Lemma 42 by which

$$\left|\frac{1}{N}\sum_n (z_n(j)^2 - 1)\right| \leq \frac{\gamma_{1,\delta}\,|A(i,j)|}{\sqrt{N}},$$

with probability greater than $1-\delta$ for $N \geq \gamma_{1,\delta}$. By taking the union bound on the two events and on all $i, j \in [d]$ we get that for all $i, j \in [d]$

$$\left|A_{kl}\left[\frac{1}{N}\sum_n A(z_n z_n^\top - I)\right]_{ij}\right| \leq \frac{3\gamma_{2,\delta/(6d^2)}(|A(i,j)| + ||A(i,[d]/j)||_2}{\sqrt{N}} \leq \frac{5\gamma_{2,\delta/(6d^2)}||A(i,\cdot)||_2}{\sqrt{N}},$$

with probability greater than $1-\delta$. The last inequality follows from Jensen's inequality,

$$a_i + \sqrt{\sum_{j\neq i} a_j^2} = \sqrt{a_i^2} + \sqrt{\sum_{j\neq i} a_j^2} \leq \sqrt{2}\sqrt{\sum_i a_j^2},$$

since $\sqrt{c_1}+\sqrt{c_2} \leq \sqrt{2}\sqrt{c_1+c_2}$.

Term (ii). See that $|\frac{1}{N}\sum_n \xi_n(i)x_{0,n}(j)|$ can be bounded by a direct application of Lemma 43 since $\xi, x_{0,n}$ are independent. Specifically, with probability greater than $1-\delta$ it holds that

$$\left|\frac{1}{N\sigma_0^2}\sum_n \xi_n(i)x_{0,n}(j)\right| \leq \frac{3\sigma\gamma_{2,\delta/(6d^2)}}{\sigma_0\sqrt{N}}$$

for all $i, j \in [d]$ by applying the union bound.

Combining the two bounds. By a union bound on the events by which terms (i) and (ii) are bounded, we get that for $N \geq \gamma_d$ it holds that

$$\mathbb{P}\left(\forall i, j \in [d] : \left|\widehat{A}(i,j) - A(i,j)\right| \geq \frac{5\gamma_{2,\delta/(6d^2)}(\sigma + \sigma_0||A(i,\cdot)||_2)}{\sigma_0\sqrt{N}}\right) \leq \delta.$$

□

Lemma 28. *Let $\{x_n\}_{n=1}^N$ be an i.i.d. vector such that $x_n \sim \mathcal{N}(0,\sigma_x I_d)$, and let $\{y_n\}_{n=1}^N$ be i.i.d. σ_y subgaussian, zero mean, random variables and assume that $N \geq \gamma_{2,\delta/3}$. Let $a \in \mathbb{R}^d$. Then,*

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_n \langle a, x_n\rangle y_n\right| \geq \frac{3\sigma_x\sigma_y||a||_2\gamma_{2,\delta/3}}{\sqrt{N}}\right) \leq \delta.$$

with probability greater than $1-\delta$.

This result is a direct application of Lemma 43 as we now show.

Proof. Observe that $z_{1,n} = \frac{1}{\sigma_x||a||_2}\langle a, x_n\rangle, z_{2,n} = \frac{1}{\sigma_y}n_i$ are both 1-sub-gaussian random variable with zero mean. Thus, to prove this result we can bound

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_n \langle a, x_n\rangle y_n\right| \geq \frac{3\sigma_x\sigma_y||a||_2\gamma_{2,\delta/3}}{\sqrt{N}}\right) = \mathbb{P}\left(\left|\frac{1}{N}\sum_i z_{1,n}z_{2,n}\right| \geq \frac{3\gamma_{2,\delta/3}}{\sqrt{N}}\right),$$

where both $z_{1,n}$ and $z_{2,n}$ are independent. Thus, we can apply Lemma 43 while setting $d_1 = d_2 = 1$ and conclude the proof. □

Algorithm 5 Semiparametric Least Squares

1: **Require:** Number of samples $N > 0$, row and column indices $i, j \in [d]$
2: Sample $\{(y_n, x_{1,n}, x_{2,n})\}_{n=1}^{2N}$
3: Estimate cross correlation $\widehat{L} = \left(\sum_{n=1}^{N} x_{1,n} x_{2,n}^\top\right)\left(\sum_{n=1}^{N} x_{2,n} x_{2,n}^\top\right)^\dagger$
4: Estimate conditional output $\widehat{c} = \left(\sum_{n=1}^{N} x_{2,n} x_{2,n}^\top\right)^\dagger \left(\sum_{n=1}^{N} y_n x_{2,n}\right)$
5: Estimate $\widehat{w}$ through plug-in

$$\widehat{w} = \left(\sum_{n=N+1}^{2N} (x_{1,n} - \widehat{L}x_{2,n})(x_{1,n} - \widehat{L}x_{2,n})^\top\right)^\dagger \left(\sum_{n=N+1}^{2N} \left(y_n - \langle \widehat{c}, x_{2,n}\rangle\right)(x_{1,n} - \widehat{L}x_{2,n})\right)$$

6: **Output:** $\widehat{w}$

H.2. Positive Definite Covariance Matrix

We now analyze the sample complexity of obtaining an element-wise good estimation of a matrix assuming that the covariance matrix of x_0 is PD. This result is a corollary of a careful semiparametric LS analysis we supply in the next section.

Corollary 29 (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).*

Proof. Fix an $i, j \in [d]$. For any such i, j we can estimate $A(i,j)$ via a semiparametric LS where the model is

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

Applying Proposition 9 and setting $d_w = 1, d_e = d-1 \leq d$ and $|w_\star| = |A(i,j)| \leq A_{\max}$ yields the bound for any any fixed $i, j \in [d]$. Applying the union bound on all $i, j \in [d]$ concludes the proof for estimating matrix A. □

H.2.1. Semiparametric Least Squares for Linear Model

Consider the following model

$$y = \langle w_\star, x_1\rangle + \langle e_\star, x_2\rangle + \epsilon, \tag{38}$$

where $x_1 \in \mathbb{R}^{d_w}, x_2 \in \mathbb{R}^{d_e}$ and ϵ is a zero mean σ sub-gaussian noise. Furthermore, assume that the covariance matrix of $[x_1 \quad x_2]$ is PD, that is $\Sigma = \mathbb{E}\left[[x_1 \quad x_2]^\top [x_1 \quad x_2]\right]$ is PD. Our goal is to recover $w_\star$ by accessing tuples of $\{y, x_1, x_2\}$, and, to achieve improved rates relatively to estimation of the entire vector $[w_\star \quad e_\star]$.

Observe that would x_1, x_2 be uncorrelated, LS regression of y given x_1 achieves our goal. With this observation, a natural first step would be to orthogonalize the model as we now show. Since $x = (x_1, x_2)$ is normally distributed it holds that

$$\mathbb{E}[x_1 | x_2] = \Sigma_{12}\Sigma_2^{-1} x_2 \equiv L_\star x_2, \tag{39}$$

where $L_\star \in \mathbb{R}^{d_w \times d_e}$, from which we get, by linearity of expectation, that

$$\mathbb{E}[y|x_2] = \langle w_\star, L_\star x_2\rangle + \langle e_\star, x_2\rangle \equiv \langle c_\star, x_2\rangle\,, \tag{40}$$

where $c_\star = L_\star^T w_\star + e_\star$. Using this, the model (38) can be written as follows.

$$\begin{aligned} y &= \langle w_\star, x_1\rangle + \langle e_\star, x_2\rangle + \epsilon \\ &= \langle w_\star, (x_1 - L_\star x_2)\rangle + \langle c_\star, x_2\rangle + \epsilon. \end{aligned} \tag{41}$$

Unlike in (38) where the features are not orthogonal $\mathbb{E}[x_1 x_2^\top] = \Sigma_{12}\Sigma_2^{-1} \neq 0$, in this new representation, the features are orthogonal since

$$\mathbb{E}[(x_1 - L_\star x_2)x_2^\top] = 0, \tag{42}$$

by construction. Thus, if we define $z_1 = x_1 - L_\star x_2$ and $z_2 = x_2$ we get that (41) is given by

$$y = \langle w_\star, z_1 \rangle + \langle c_\star, z_2 \rangle + \epsilon,$$

where z_1, z_2 are orthogonal and their covariance matrix is given by

$$\begin{aligned} &\mathrm{Cov}_{z_1} = \mathbb{E}[(x_1 - L_\star x_2)(x_1 - L_\star x_2)^\top] = \Sigma_1 - \Sigma_{12}\Sigma_2^{-1}\Sigma_{12}^\top \equiv \Sigma/\Sigma_2, \\ &\mathrm{Cov}_{z_2} = \Sigma_2, \end{aligned} \tag{43}$$

where Σ/Σ_2 is the known as the Schur complement.

Importantly, would we be given $L_\star$ and $c_\star$, we can get an unbiased estimate of $w_\star$ using the data set $\{(y_n, x_{1,n}, x_{2,n})\}_{n=1}^N$ through an ordinary least-squares approach,

$$\widehat{w} = \left(\sum_{n=N+1}^{2N} (x_{1,n} - L_\star x_{2,n})(x_{1,n} - L_\star x_{2,n})^\top \right)^\dagger \left(\sum_{n=N+1}^{2N} (y_n - \langle c_\star, x_{2,n} \rangle)(x_{1,n} - L_\star x_{2,n}) \right). \tag{44}$$

It can be shown that $\mathbb{E}[\widehat{w}] = w_\star$ when the design matrix $V_N = \sum_{n=1}^N (x_{1,n} - L_\star x_{2,n})(x_{1,n} - L_\star x_{2,n})^\top$ is PD. This fact, motivates us to study the *finite* sample performance of this approach when both $L_\star$ and $c_\star$ are estimated from data (Algorithm 5). In the next section, we study this estimator without any assumption besides of positive minimal eigenvalue of the covariance matrix of $x = (x_1, x_2)$.

H.2.2. Finite Sample Analysis: Semiparametric LS

We are now ready to analyze the performance of Algorithm 5. Relying on the OLS (44), Algorithm 5 splits the data in two, with the first dataset it estimates $L_\star$ and $c_\star$. With the second dataset, it solves the OLS (44) in which the exact $L_\star$ and $c_\star$ are replaced by their estimators.

The following lemma establishes a finite performance guarantee of Algorithm 5. Importantly, we see that there's only a lower order dependence in d_e which we suffer due to the need to estimate $L_\star$ and $c_\star$.

Proposition 30 (Semiparametric Least-Squares)**.** *Let $\delta \in (0, e^{-1})$. Consider model (8) and assume that Σ is PD. Denote $\sigma_c^2 = ||w_\star||^2_{\Sigma/\Sigma_2} + \sigma^2$. 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*

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

Overview of the analysis of Proposition 9 . We decompose the error into three terms in (48). The first term is of dimension s (as oppose to d) and is bounded via standard concentrations for least-squares (Hsu et al., 2012a). The second and third terms are errors we suffer due to in-exact estimation of $L_\star$ and $c_\star$.

Importantly, we bound the errors in the estimates of $L_\star$ and $c_\star$ in weighted norms. Specifically, we show we can bound

$$||\Sigma/\Sigma_2^{-1/2}(\widehat{L} - L_\star)\Sigma_2^{1/2}||_{\mathrm{op}} \textit{ and } ||\Sigma_2^{1/2}(\widehat{c} - c_\star)||_2$$

by a term which is independent of minimal eigenvalues of Σ or Σ/Σ_2. With this at hand, and by further careful analysis, we show, that the second and third terms in (48) can be bounded by terms that are independent of minimal eigenvalues. The final result follows by relating the minimal and eigenvalues of Σ/Σ_2 to the of Σ, supplied in (Smith, 1992).

Proof. The OLS solution $\widehat{w}$ satisfies the following relation

$$
\begin{aligned}
&\sum_{n=N+1}^{2N} \frac{1}{N} \left[\left(x_{1,n} - \widehat{L} x_{2,n} \right) \left(x_{1,n} - \widehat{L} x_{2,n} \right) \right]^\top (w_\star - \widehat{w}) \\
&= \sum_{n=N+1}^{2N} \frac{1}{N} \left(x_{1,n} - \widehat{L} x_{2,n} \right) \left(y_n - \left\langle w_\star, x_{1,n} - \widehat{L} x_{2,n} \right\rangle - \langle \widehat{c}, x_{2,n} \rangle \right). \qquad (45)
\end{aligned}
$$

Let

$$
\widehat{z}_{1,n}(\widehat{L}) = x_{1,n} - \widehat{L} x_{2,n},\ z_{1,n} = x_{1,n} - L_\star x_{2,n},\ z_{2,n} = x_{2,n}
$$

and define the design matrix as

$$
V_N = \frac{1}{N} \sum_{n=N+1}^{2N} \widehat{z}_{1,n}(\widehat{L}) \widehat{z}_{1,n}(\widehat{L})^\top.
$$

By multiplying both sides of this relation by $\Sigma/\Sigma_2(\widehat{L})^{-1/2} = \mathbb{E}[\widehat{z}_{1,n}(\widehat{L})\widehat{z}_{1,n}(\widehat{L})^\top]$ and by some additional algebraic manipulations, it can be shown that (45) implies that

$$
\begin{aligned}
&\Sigma/\Sigma_2(\widehat{L})^{-1/2} V_N (w_\star - \widehat{w}) \\
&= \Sigma/\Sigma_2(\widehat{L})^{-1/2} \sum_{i=1}^{N} \frac{1}{N} \left(x_{1,n} - \widehat{L} x_{2,n} \right) \left(y_i - \langle w_\star, x_{1,n} - L_\star x_{2,n} \rangle - \langle c_\star, x_{2,n} \rangle \right) \\
&\quad + \Sigma/\Sigma_2(\widehat{L})^{-1/2} \sum_{i=1}^{N} \frac{1}{N} \left(x_{1,n} - L_\star x_{2,n} \right) \left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n} \right\rangle \\
&\quad + \Sigma/\Sigma_2(\widehat{L})^{-1/2} \sum_{i=1}^{N} \frac{1}{N} \left((\widehat{L} - L_\star) x_{2,n} \right) \left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n} \right\rangle \\
&= \Sigma/\Sigma_2(\widehat{L})^{-1/2} \sum_{i=1}^{N} \frac{1}{N} \widehat{z}_{1,n}(\widehat{L}) \epsilon_n \\
&\quad + \Sigma/\Sigma_2(\widehat{L})^{-1/2} \sum_{i=1}^{N} \frac{1}{N} z_{1,n} \left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, z_{2,n} \right\rangle \\
&\quad + \Sigma/\Sigma_2(\widehat{L})^{-1/2} \sum_{i=1}^{N} \frac{1}{N} \left((\widehat{L} - L_\star) z_{2,n} \right) \left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, z_{2,n} \right\rangle \qquad (46)
\end{aligned}
$$

where the last relation holds since $y_i - \langle w_\star, x_{1,n} - L_\star x_{2,n} \rangle - \langle c_\star, x_{2,n} \rangle = \epsilon_n$ due to the model assumption (41). Observe we obtained a vector equality of the form

$$
a = b_1 + b_2 + b_3
$$

where $a, b_1, b_2, b_3 \in \mathbb{R}^{d_w}$. This equality implies that

$$
||a||_2 = ||b_1 + b_2 + b_3||_2 \le ||b_1||_2 + ||b_2||_2 + ||b_3||_2 \qquad (47)
$$

due to the triangle inequality. Hence, the vector equality in (46) together with (47) implies that

$$\begin{aligned}
&||\Sigma/\Sigma_2(\widehat{L})^{-1/2}V_N(\widehat{w}-w_\star)||_2\\
&=\underbrace{||\sum_{i=1}^N \frac{1}{N}\widehat{z}_{1,n}(\widehat{L})\epsilon_n||_{\Sigma/\Sigma_2(\widehat{L})^{-1}}}_{(i)}\\
&+\underbrace{||\sum_{i=1}^N \frac{1}{N}z_{1,n}\left\langle(c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star, z_{2,n}\right\rangle||_{\Sigma/\Sigma_2(\widehat{L})^{-1}}}_{(ii)}\\
&+\underbrace{||\sum_{i=1}^N \frac{1}{N}\left((\widehat{L}-L_\star)z_{2,n}\right)\left\langle(c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star, z_{2,n}\right\rangle||_{\Sigma/\Sigma_2(\widehat{L})^{-1}}}_{(iii)}.
\end{aligned} \tag{48}$$

We bound each one of these terms by Lemma 33, Lemma 34 and Lemma 35. We verify the conditions of these lemmas hold.

1. $N\geq 9\gamma_{d,\delta}^2\geq\gamma_{d,\delta}$, by assumption and since $\delta\in(0,e^{-1})$ (see that $9\gamma_{d,\delta}^2=\Theta(d\log(\frac{1}{\delta})))$.
2. By Lemma 31

$$\begin{aligned}
&||\Sigma_2^{1/2}((c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star)||\\
&\leq||\Sigma_2^{1/2}((c_\star-\widehat{c})||_2||\Sigma_2^{1/2}(\widehat{L}-L_\star)^\top\Sigma/\Sigma_2^{-1/2}||_{\mathrm{op}}||\Sigma_2^{1/2}w_\star||_2\\
&\leq 10\sigma_c\sqrt{\frac{dd_w\log\left(\frac{d_w}{\delta}\right)}{N}}=\Delta
\end{aligned}$$

   and by Lemma 32

$$||\left(\Sigma/\Sigma_2\right)^{-1/2}(\widehat{L}-L_\star)\Sigma_2^{1/2}||_{\mathrm{op}}\leq 5\sigma_c\sqrt{\frac{dd_w\log\left(\frac{d_w}{\delta}\right)}{N}}=\Delta_L=\Delta/2.$$

   with probability greater than $1-\delta$. Thus, we approximate $c_\star, L_\star$ in the scaled norms by the covariance matrices Σ_2 and Σ/Σ_2.
3. Since $\Delta_L^2=25\sigma_c^2dd_w\log\left(\frac{d_w}{\delta}\right)/N$ it holds that for $N\geq 50\sigma_c^2dd_w\log\left(\frac{d_w}{\delta}\right)/\lambda_{\min}(\Sigma)$ the covariance matrix $\Sigma/\Sigma_2(\widehat{L})$ is PD, and specifically,

$$\lambda_{\min}(\Sigma/\Sigma_2(\widehat{L}))\geq\lambda_{\min}(\Sigma/\Sigma_2)/2>\lambda_{\min}(\Sigma)/2>0, \tag{49}$$

   where the first relation holds by Lemma 36 while setting $\Delta_L^2=\lambda_{\min}(\Sigma/\Sigma_2)/2$, and the second relation by standard fact on the Schur complement of a matrix (see (Smith, 1992), Theorem 5).
4. For $N\geq 50\sigma_c^2dd_w\log\left(\frac{d_w}{\delta}\right)/\lambda_{\min}(\Sigma/\Sigma_2)$ by the third relation of Lemma 36 we get that $||\left(\Sigma/\Sigma_2(\widehat{L})\right)^{-1/2}\Sigma/\Sigma_2^{1/2}||_{\mathrm{op}}\leq\sqrt{2}$.

Observe that by taking

$$N\geq O\left(\left(\sigma_c^2/\lambda_{\min}(\Sigma)\vee 1\right)dd_w\log\left(\frac{d_w}{\delta}\right)\right), \tag{50}$$

we satisfy all the requirements on the sample size.

Applying the union bound on all the above and scaling $\delta\leftarrow\delta/3$ we get that all the events hold with probability greater than $1-\delta$. We refer to this event as the first good event $\mathcal{G}_1$. We can now apply Lemma 33, Lemma 34 and Lemma 35 and bound (48) conditioning on $\mathcal{G}_1$. By applying these lemmas and using the union bound we get that with probability greater than $1-\delta$

$$
\begin{aligned}
&||\Sigma/\Sigma_2(\widehat{L})^{-1/2}V_N(\widehat{w}-w_\star)||_2 \\
&\leq 3\sigma\sqrt{\frac{d_w\log\left(\frac{6}{\delta}\right)}{N}}+\frac{5\Delta\gamma_{d,\delta/9}}{\sqrt{N}}+\frac{5\Delta^2/2\gamma_{d,\delta/9}}{\sqrt{N}}+\Delta^2 \\
&\leq 3\sigma\sqrt{\frac{d_w\log\left(\frac{6}{\delta}\right)}{N}}+50\gamma_{d,\delta/9}\sigma_c\sqrt{\frac{d_w d\log\left(\frac{6d_w}{\delta}\right)}{N^2}}+\frac{250\gamma_{d,\delta/9}\sigma_c^2 dd_w\log\left(\frac{6d_w}{\delta}\right)}{N^{3/2}}+\frac{100\sigma_c^2 dd_w\log\left(\frac{6d_w}{\delta}\right)}{N}
\end{aligned}
\tag{51}
$$

by plugging the form of Δ, Δ_L and using $d_e+d_w=d$.

Finally, we translate this bound to a bound on $||\widehat{w}-w_\star||_{\Sigma/\Sigma_2(\widehat{L})}$ by applying Lemma 39. We now verify the conditions of this lemma.

1. The matrix $\Sigma/\Sigma_2(\widehat{L})$ is PD by (49).
2. The empirical covariance is concentrated around the true one,

$$
||\left(\Sigma/\Sigma_2(\hat{L})\right)^{-1/2}V_N\left(\Sigma/\Sigma_2(\hat{L})\right)^{-1/2}-I||_{\mathrm{op}}\leq\frac{\gamma_{d_e,\delta}}{\sqrt{N}}\leq\frac{1}{3},
\tag{52}
$$

with probability greater than $1-\delta$ by Lemma 42, and the second inequality holds since $N\geq 9\gamma_{d,\delta}^2\geq 9\gamma_{d_e,\delta}^2$.

Applying Lemma 39 with $c=1/3$, while using the bound in (51) we get

$$
\begin{aligned}
&||(\widehat{w}-w_\star)||_{\Sigma/\Sigma_2(\widehat{L})} \\
&\leq 6\sigma\sqrt{\frac{d_w\log\left(\frac{6}{\delta}\right)}{N}}+\frac{80\gamma_{d,\delta/9}\sigma_c\sqrt{d_w d\log\left(\frac{6d_w}{\delta}\right)}}{N}+\frac{400\gamma_{d,\delta/9}\sigma_c^2 dd_w\log\left(\frac{6d_w}{\delta}\right)}{N^{3/2}}+\frac{150\sigma_c^2 dd_w\log\left(\frac{6d_w}{\delta}\right)}{N}.
\end{aligned}
$$

Furthermore, observe that

$$
\lambda_{\min}(\Sigma/\Sigma_2(\widehat{L}))\geq\frac{1}{2}\lambda_{\min}(\Sigma/\Sigma_2)\geq\lambda_{\min}(\Sigma)
$$

where the first relation holds by, and the second by identifies of the Schur complement of a PD matrix (49) and the second relation by (Smith, 1992), Theorem 5. Thus,

$$
\begin{aligned}
&||(\widehat{w}-w_\star)||_2\leq\frac{\sqrt{2}}{\sqrt{\lambda_{\min}(\Sigma)}}||(\widehat{w}-w_\star)||_{\Sigma/\Sigma_2(\widehat{L})} \\
&\leq\frac{1}{\sqrt{\lambda_{\min}(\Sigma)}}\left(9\sqrt{\frac{\sigma^2 d_w\log\left(\frac{1}{\delta}\right)}{N}}+\frac{400\gamma_{d,\delta/9}\sigma_c^2 dd_w\log\left(\frac{d_w}{\delta}\right)}{N^{3/2}}+\frac{230\left(\sigma_c^2\vee\sigma_c\right)dd_w\log\left(\frac{9d_w}{\delta}\right)}{N}\right).
\end{aligned}
\tag{53}
$$

Lastly, by the choice of N given in (50) and the definition of $\gamma_{d,\delta/9}=O(\sqrt{d\log\left(\frac{1}{\delta}\right)})$ (see Lemma 42) it holds that,

$$
\frac{\gamma_{d,\delta}}{\sqrt{N}}\leq O(1).
$$

Thus, the last two term of (53) are related by a multiplicative constant factor. This concludes the proof. □

H.2.3. Analysis of the First Phase Errors

Lemma 31 (Sample Complexity of Learning $c_\star$)**.** *Let $\delta\in(0,e^{-1})$ and let $\Sigma_2=\mathbb{E}[x_2x_2^\top]$. Assume that $N\geq 9\gamma_{d_e,\delta}^2$. Then, with probability greater than $1-\delta$ it holds that*

$$
||\Sigma_2^{1/2}(\widehat{c}-c_\star)||_2\leq 5\sigma_c\sqrt{\frac{d_e\log\left(\frac{2}{\delta}\right)}{N}},
$$

where $\sigma_c^2=||w_\star||_{\Sigma/\Sigma_2}^2+\sigma^2$.

Proof. This result is a direct application of Proposition 41 which establishes performance guarantee on the OLS. We show that $y = \langle c_\star, x_2 \rangle + \epsilon$ to apply this result. See that

$$y_n = \mathbb{E}[y|x_{2,n}] + \mathbb{E}[y|x_{2,n}] - y_n = \langle c_\star, x_2 \rangle + \epsilon_{c,n}$$

where $\epsilon_{c,n} = y_n - \langle c_\star, x_{2,n} \rangle$ is a $\sigma_c = \sqrt{||w_\star||^2_{\Sigma/\Sigma_2} + \sigma^2}$ sub gaussian, zero mean random variable. Indeed,

$$\mathbb{E}[\epsilon_{c,n}] = \mathbb{E}[y_n - \langle c_\star, x_{2,n} \rangle] = \mathbb{E}[y_n - \mathbb{E}[y_n|x_{2,n}]] = 0.$$

To see it is a σ_c sub gaussian observe that

$$\epsilon_{c,n} = y_n - \langle c_\star, x_{2,n} \rangle = ||x_{1,n} - L_\star x_{2,n}, w_\star|| + \epsilon_n.$$

Thus, and due to the independence of ϵ and (x_1, x_2) we get

$$\begin{aligned}
&Var(\epsilon_{c,n}) = Var(\langle x_{1,n} - L_\star x_{2,n}, w_\star \rangle) + Var(\epsilon_n) \\
&= Var(\left\langle (\Sigma/\Sigma_2)^{-1/2} (x_{1,n} - L_\star x_{2,n}), (\Sigma/\Sigma_2)^{1/2} w_\star \right\rangle) + \sigma^2 \\
&= ||\Sigma/\Sigma_2^{1/2} w_\star||^2 + \sigma^2 = ||w_\star||^2_{\Sigma/\Sigma_2} + \sigma^2.
\end{aligned}$$

Thus, the claim follows from Proposition 41. $\square$

Lemma 32 (Sample Complexity of Learning $L_\star$)**.** *Let $\delta \in (0, e^{-1})$ and let $\Sigma_2 = \mathbb{E}[x_2 x_2^\top]$. Assume that $N \geq 9\gamma^2_{d_e,\delta}$. Let the OLS estimate of $L_\star$ be*

$$\widehat{L} = \frac{1}{N_1} \sum_{n=1}^{N_1} x_{1,n} x_{2,n}^\top V^\dagger_{N_1,x_2}.$$

Then, with probability greater than $1-\delta$ it holds that

$$||\Sigma/\Sigma_2^{-1/2} (\widehat{L} - L_\star) \Sigma_2^{1/2}||_{\text{op}} \leq 5\sqrt{\frac{d_w d_e \log\left(\frac{2d_w}{\delta}\right)}{N}}.$$

Proof. We apply the concentration result on the OLS estimator, Proposition 41. To see it is applicable, we reduce this problem to a single parameter estimation. First, bound the operator norm by the Frobenius norm. Let $e_i \in \mathbb{R}^{d_w}$ be a one hot vector with one at its i^{th} entry. Then,

$$||\Sigma/\Sigma_2^{-1/2} (\widehat{L} - L_\star) \Sigma_2^{1/2}||^2_{\text{op}} \leq ||\Sigma/\Sigma_2^{-1/2} (\widehat{L} - L_\star) \Sigma_2^{1/2}||^2_F = \sum_{i=1}^{d_w} ||e_i^\top \left(\Sigma/\Sigma_2^{-1/2} (\widehat{L} - L_\star) \Sigma_2^{1/2} \right)||^2_2. \tag{54}$$

Observe that the following vector equality holds by the model assumption (39).

$$x_1 = \mathbb{E}[x_1|x_2] + x_1 - \mathbb{E}[x_1|x_2] = L_\star x_2 + (x_1 - L_\star x_2), \tag{55}$$

where $\mathbb{E}[x_1 - L_\star x_2] = 0,\ \mathrm{Cov}(x_1 - L_\star x_2) = \Sigma/\Sigma_2$ (see (43)). Multiplying (55) from the left by $e_i^\top (\Sigma/\Sigma_2)^{-1/2}$, we get that for any $i \in [d_w]$

$$y_{n,i} \equiv e_i^\top (\Sigma/\Sigma_2)^{-1/2} x_{1,n} = e_i^\top (\Sigma/\Sigma_2)^{-1/2} L_\star x_2 + e_i^\top \epsilon_n = \langle \beta_i, x_{2,n} \rangle + \epsilon_{n,i},$$

where

$$\begin{aligned}
&\beta_i = e_i^\top (\Sigma/\Sigma_2)^{-1/2} L_\star, \\
&\epsilon_{n,i} = e_i^\top (\Sigma/\Sigma_2)^{-1/2} (x_1 - L_\star x_2),
\end{aligned}$$

and $\epsilon_{n,i}$ is zero mean with a unit variance, since,

$$\begin{aligned}&\mathbb{E}[\epsilon_{n,i}] = e_i^\top \left(\Sigma/\Sigma_2\right)^{-1/2}\mathbb{E}[x_1 - L_\star x_2] = 0\\&Var(\epsilon_n) = e_i^\top \left(\Sigma/\Sigma_2\right)^{-1/2}\left(\Sigma/\Sigma_2\right)\left(\Sigma/\Sigma_2\right)^{-1/2} e_i = 1.\end{aligned} \tag{56}$$

Observe that the ordinary least square estimator of $\beta_i^\top$ is given by the following equivalent forms

$$\begin{aligned}\widehat{\beta}_i^\top &= \sum_{n=1}^N y_{n,i} x_{2,n}^\top V_{N,2}^\dagger\\&= \sum_{n=1}^N e_i^\top \left(\Sigma/\Sigma_2\right)^{-1/2} x_{1,n} x_{2,n}^\top V_{N,2}^\dagger\\&= e_i^\top \left(\Sigma/\Sigma_2\right)^{-1/2} \widehat{L}_N.\end{aligned}$$

By applying the concentration result for OLS, Proposition 41, and applying the union bound, we get that for all $i \in [d_w]$, assuming $N \geq 9\gamma_{d,\delta}^2$

$$||e_i^\top \left(\Sigma/\Sigma_2\right)^{-1/2}(\widehat{L}_N - L_\star)|| = ||(\widehat{\beta}_i - \beta_i)^\top|| \leq 5\sigma_i^2 \frac{d_e \log\left(\frac{d_w}{\delta}\right)}{N} \leq 5\frac{d_e \log\left(\frac{d_w}{\delta}\right)}{N}.$$

Thus,

$$\begin{aligned}(54) &= \sum_{i=1}^{d_w} ||e_i^\top \left(\Sigma/\Sigma_2^{-1/2}(\widehat{L} - L_\star)\Sigma_2^{1/2}\right)||_2^2\\&\leq \sum_{i=1}^{d_w} \frac{25\sigma_{i,l} d_e \log\left(\frac{d_w}{\delta}\right)}{N}\\&= \frac{25 d_w d_e \log\left(\frac{d_w}{\delta}\right)}{N}, && \text{(By (56) } \sigma_{i,l} = 1 \text{ for all } i \in [d_w])\end{aligned}$$

which concludes the proof. ☐

H.2.4. Analysis of the Second Phase Errors

Lemma 33 (Bound on First Term of Proposition 9)**.** *Let $\delta \in (0, e^{-1})$. Assume that $\Sigma/\Sigma_2(\widehat{L})$ is invertible. Then, with probability greater then $1-\delta$ it holds that*

$$||\frac{1}{N}\sum_n \left(x_{1,n} - \widehat{L}x_{2,n}\right)\epsilon_n||_{\Sigma/\Sigma_2(\widehat{L})^{-1}} \leq 3\sigma\sqrt{\frac{d_w \log\left(\frac{1}{\delta}\right)}{N}}.$$

Proof. This term cab be directly bounded by applying Lemma 38. Let $z_n = \frac{1}{\sqrt{N}}(x_{1,n} - \widehat{L}x_{2,n})$ and define $Z_N \in \mathbb{R}^{N\times d_w}$ as the matrix with z_i as its rows. Observe that with this notation $V_N = \sum_n z_n z_n^\top = Z_N^\top Z_N$. Furthermore, define $\xi_N \in \mathbb{R}^N$ as a vector with ϵ_n in its rows. The following relations hold

$$||\frac{1}{N}\sum_n \left(x_{1,n} - \widehat{L}x_{2,n}\right)\epsilon_n||^2_{\Sigma/\Sigma_2(\widehat{L})^{-1}} = ||\Sigma/\Sigma_2(\widehat{L})^{-1/2} Z_N^\top \xi_N||_2.$$

With this form of writing $||\frac{1}{N}\sum_n \left(x_{1,n} - \widehat{L}x_{2,n}\right)\epsilon_n||^2_{\Sigma/\Sigma_2(\widehat{L})^{-1}}$, we see it can be bounded by applying Lemma 38. ☐

Lemma 34 (Bound on Second Term of Proposition 9)**.** *Let $\delta \in (0, e^{-1})$. Assume the following holds.*

1. $N \geq \gamma_{d,\delta}$.

2. $\Sigma/\Sigma_2(\widehat{L})$ *is invertible and* $||\Sigma/\Sigma_2(\widehat{L})^{-1/2}\Sigma/\Sigma_2^{1/2}||_{\mathrm{op}} \leq \sqrt{2}$.

3. $||\Sigma_2^{1/2}((c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star)|| \leq \Delta$.

Then, with probability greater then $1-\delta$ *it holds that*

$$||\frac{1}{N}\sum_{i=1}^{N}(x_{1,n} - L_\star x_{2,n})\left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n}\right\rangle ||_{\Sigma/\Sigma_2(\widehat{L})^{-1}} \leq \frac{5\Delta\gamma_{d,\delta/3}}{\sqrt{N}}.$$

Proof. First, observe that the following relation hold

$$\begin{aligned}
&||\frac{1}{N}\sum_{i=1}^{N}(x_{1,n} - L_\star x_{2,n})\left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n}\right\rangle ||_{\Sigma/\Sigma_2(\widehat{L})^{-1}} \\
&\leq ||\frac{1}{N}\sum_{i=1}^{N}(x_{1,n} - L_\star x_{2,n})\left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n}\right\rangle ||_{\Sigma/\Sigma_2^{-1}}||\Sigma/\Sigma_2(\widehat{L})^{-1/2}\Sigma/\Sigma_2^{1/2}|| \\
&\leq \sqrt{2}||\frac{1}{N}\sum_{i=1}^{N}(x_{1,n} - L_\star x_{2,n})\left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n}\right\rangle ||_{\Sigma/\Sigma_2^{-1}} && \text{(By assumption)} \\
&\leq \frac{\sqrt{2}}{N}||\sum_{i=1}^{N}(x_{1,n} - L_\star x_{2,n})\left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n}\right\rangle ||_{\Sigma/\Sigma_2^{-1}} \\
&= \frac{\sqrt{2}\Delta}{N}||\sum_{i=1}^{N}\Sigma/\Sigma_2^{-1/2}(x_{1,n} - L_\star x_{2,n})(\Sigma_2^{-1/2}x_{2,n})^\top||_{\mathrm{op}}, && (57)
\end{aligned}$$

where the last relation holds since $||Ab||_2^2 \leq ||b||_2^2||A||_{\mathrm{op}}^2$, and since $||\Sigma_2^{1/2}((c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star)||_2 \leq \Delta$ by assumption. We now bound (57).

Let $z_1 = \Sigma/\Sigma_2(x_{1,n} - L_\star x_{2,n})$ and $z_2 = \Sigma_2^{-1/2}x_{2,n}$ and observe that $z_1 \sim \mathcal{N}(0, I_{d_w})$, $z_2 \sim \mathcal{N}(0, I_{d_e})$ and are independent random vectors since $\mathbb{E}[z_1 z_2^\top] = 0$ (see (42)). Thus, by Lemma 43, for $N \geq \gamma_{d,\delta/3}$, the following bound holds with probability greater than $1-\delta$

$$(57) = \frac{\sqrt{2}\Delta}{N}||\sum_{i=1}^{N} z_{1,n}z_{2,n}^\top||_{\mathrm{op}} \leq \frac{5\Delta\gamma_{d,\delta/3}}{\sqrt{N}}.$$

□

Lemma 35 (Bound on Third Term of Proposition 9)**.** *Let* $\delta \in (0, e^{-1})$. *Assume the following holds.*

1. $N \geq \gamma_{d,\delta/3}$

2. $\Sigma/\Sigma_2(\widehat{L})$ *is invertible and* $||\Sigma/\Sigma_2(\widehat{L})^{-1/2}\Sigma/\Sigma_2^{1/2}||_{\mathrm{op}} \leq \sqrt{2}$.

3. $||\Sigma_2^{1/2}((c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star)|| \leq \Delta$, *and* $||\left(\Sigma/\Sigma_2\right)^{-1/2}(\widehat{L} - L_\star)\Sigma_2^{1/2}||_{\mathrm{op}} \leq \Delta_L$.

Then, with probability greater than $1-\delta$, *it holds that*

$$||\frac{1}{N}\sum_{i=1}^{N}\left((\widehat{L} - L_\star)x_{2,n}\right)\left\langle (c_\star - \widehat{c}) + (\widehat{L} - L_\star)^\top w_\star, x_{2,n}\right\rangle ||_{\Sigma/\Sigma_2(\widehat{L})^{-1}} \leq \frac{5\Delta_L\Delta\gamma_{d_e,\delta/3}}{\sqrt{N}} + \sqrt{2}\Delta_L\Delta.$$

Proof. The following relations hold.

$$
\begin{aligned}
&||\frac{1}{N}\sum_{i=1}^{N}\left((\widehat{L}-L_\star)x_{2,n}\right)\left\langle(c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star, x_{2,n}\right\rangle||_{\Sigma/\Sigma_2(\widehat{L})^{-1}}\\
&\le ||\frac{1}{N}\sum_{i=1}^{N}\left((\widehat{L}-L_\star)x_{2,n}\right)\left\langle(c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star, x_{2,n}\right\rangle||_{\Sigma/\Sigma_2^{-1}}||\Sigma/\Sigma_2(\widehat{L})^{-1/2}\Sigma/\Sigma_2^{1/2}||_{\text{op}}\\
&\le \sqrt{2}||\frac{1}{N}\sum_{i=1}^{N}\left((\widehat{L}-L_\star)x_{2,n}\right)\left\langle(c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star, x_{2,n}\right\rangle||_{\Sigma/\Sigma_2^{-1}} \qquad \text{(By assumption)}\\
&= \sqrt{2}||\left(\Sigma/\Sigma_2^{-1/2}(\widehat{L}-L_\star)\Sigma_2^{1/2}\right)\frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_2^{-1/2}x_{2,n}\right)\left(\Sigma_2^{-1/2}x_{2,n}\right)^\top\left(\Sigma_2^{1/2}(c_\star-\widehat{c})+(\widehat{L}-L_\star)^\top w_\star\right)||_2\\
&\le \sqrt{2}\Delta_L\Delta||\sum_{i=1}^{N}\left(\Sigma_2^{-1/2}x_{2,n}\right)\left(\Sigma_2^{-1/2}x_{2,n}\right)^\top||_{\text{op}}\\
&\le \sqrt{2}\Delta_L\Delta||\sum_{i=1}^{N}\left(\Sigma_2^{-1/2}x_{2,n}\right)\left(\Sigma_2^{-1/2}x_{2,n}\right)^\top - I_{d_e}||_{\text{op}}+\sqrt{2}\Delta_L\Delta\\
&\le \frac{5\Delta_L\Delta\gamma_{d_e,\delta/3}}{\sqrt{N}}+\sqrt{2}\Delta_L\Delta.
\end{aligned}
$$

where the last relation holds with probability greater than $1-\delta$ by the concentration of the empirical covariance matrix (see Lemma 42) for $N\ge\gamma_{d,\delta}\ge\gamma_{d_e,\delta}$, while observing that $\Sigma_2^{-1/2}x_{2,n}\sim\mathcal{N}(0,I_{d_e})$. □

Lemma 36 (Lower Order Bound on the Design Matrix of the Perturbed Least Square)**.** *Let* $||(L_\star-\widehat{L})\Sigma_2^{1/2}||_{\text{op}}\le\Delta_L$. *Furthermore,*

$$
\begin{aligned}
\Sigma/\Sigma_2 &= \mathbb{E}\left[(x_1-L_\star x_2)(x_1-L_\star x_2)^\top\right]\\
\Sigma/\Sigma_2(\widehat{L}) &= \mathbb{E}\left[(x_1-\widehat{L}x_2)(x_1-\widehat{L}x_2)^\top\right].
\end{aligned}
$$

Then, the following relations hold.

1. $||\Sigma/\Sigma_2-\Sigma/\Sigma_2(\widehat{L})||_{\text{op}}\le\Delta_L^2$.
2. $\lambda_{\min}(\Sigma/\Sigma_2(\widehat{L}))\ge\lambda_{\min}(\Sigma/\Sigma_2)-\Delta_L^2$.
3. *Assuming* $\left(\Delta_L^2/\lambda_{\min}(\Sigma/\Sigma_2)\right)<1$ *then*

$$
||\left(\Sigma/\Sigma_2(\widehat{L})\right)^{-1/2}\Sigma/\Sigma_2^{1/2}||_{\text{op}}\le\sqrt{(1-(\Delta_L^2/\lambda_{\min}(\Sigma/\Sigma_2)))^{-1}}.
$$

Proof. **First relation.** By definition,

$$
\begin{aligned}
&\Sigma/\Sigma_2-\Sigma/\Sigma_2(\widehat{L}) = \mathbb{E}[(x_1-L_\star x_2)(x_1-L_\star x_2)^\top]-\mathbb{E}[(x_1-\widehat{L}x_2)(x_1-\widehat{L}x_2)^\top]\\
&= \mathbb{E}[(x_1-L_\star x_2)(x_1-L_\star x_2)^\top]-\mathbb{E}[(x_1-L_\star x_2-(\widehat{L}-L_\star)x_2)(x_1-L_\star x_2-(\widehat{L}-L_\star)x_2)^\top] \qquad (58)
\end{aligned}
$$

Since $(a+b)(a+b)^\top=aa^\top+bb^\top+ab^\top+ba^\top$, defining $a=x_1-L_\star x_2, b=-(\widehat{L}-L_\star)x_2$, we get that

$$
\begin{aligned}
(58) &= -\mathbb{E}[(x_1-L_\star x_2)((\widehat{L}-L_\star)x_2)^\top]-\mathbb{E}[(\widehat{L}-L_\star)x_2)(x_1-L_\star x_2)^\top]+(\widehat{L}-L_\star)\mathbb{E}[x_2x_2^\top](\widehat{L}-L_\star)^\top\\
&= (\widehat{L}-L_\star)\Sigma_2(\widehat{L}-L_\star)^\top,
\end{aligned}
$$

since

$$
\mathbb{E}[(x_1-L_\star x_2)((\widehat{L}-L_\star)x_2)^\top]=\mathbb{E}[\mathbb{E}[(x_1-L_\star x_2)|x_2](\widehat{L}-L_\star)x_2)^\top]=\mathbb{E}[(L_\star-\widehat{L})x_2((L_\star-L_\star)x_2)^\top]=0,
$$

and, similarly, $\mathbb{E}[((\widehat{L}-L_\star)x_2)(x_1-L_\star x_2)^\top]=0$. We get that

$$\begin{aligned}
&||\Sigma/\Sigma_2-\Sigma/\Sigma_2(\widehat{L})||_{\text{op}}=||(\widehat{L}-L_\star)\Sigma_2(\widehat{L}-L_\star)||_{\text{op}}\\
&=||((\widehat{L}-L_\star)\Sigma_2^{1/2})((\widehat{L}-L_\star)\Sigma_2^{1/2})^\top||_{\text{op}}\\
&\le ||((\widehat{L}-L_\star)\Sigma_2^{1/2})||_{\text{op}}^2\le \Delta_L^2.
\end{aligned}$$

Second relation. Direct application of Weyl's inequality.

Third relation. Let $\Delta=\Sigma/\Sigma_2^{-1/2}\left(\Sigma/\Sigma_2(\widehat{L})-\Sigma/\Sigma_2\right)\Sigma/\Sigma_2^{-1/2}$. Observe that

$$\begin{aligned}
||\Delta||_{\text{op}} &\le \frac{1}{\lambda_{\min}(\Sigma/\Sigma_2)}||\Sigma/\Sigma_2(\widehat{L})-\Sigma/\Sigma_2||_{\text{op}} && (||\cdot||_{\text{op}}\text{ is submultiplicative})\\
&\le \frac{\Delta_L^2}{\lambda_{\min}(\Sigma/\Sigma_2)}. && (\text{First relation of the lemma})
\end{aligned}$$

Thus, since $||\Delta||_{\text{op}}\le\frac{\Delta_L^2}{\lambda_{\min}(\Sigma/\Sigma_2)}<1$ by assumption, we can apply Lemma 40 and conclude the proof. □

H.2.5. Least Square General Results and Tools

Theorem 37 ((Hsu et al., 2012a), Theorem 1)**.** *Let $A\in\mathbb{R}^{m\times N}$ be a matrix, and let $K=A^\top A$. Suppose tat $\xi=(\epsilon_1,..,\epsilon_N)$ is a zero-mean and σ sub-gaussian vector such that for some $\sigma\ge 0$ it holds that $\mathbb{E}[\exp(v^\top\xi)]\le\exp\left(||v||_2^2\sigma^2/2\right)$ for all $v\in\mathbb{R}^N$. Then, for any $t>0$,*

$$\Pr\left(||A\xi||_2^2\ge\sigma^2\left(\operatorname{tr}(K)+2\sqrt{\operatorname{tr}(K^2)\,t}+2||K||t\right)\right)\le\exp(-t)\,.$$

The following lemma will be useful for our analysis.

Lemma 38 (Noise Concentration for the OLS)**.** *Let $X_N\in\mathbb{R}^{N\times d}$ be a matrix with $\left\{\frac{1}{\sqrt{N}}x_n\right\}_{n=1}^N$ in its rows, $V_N=\frac{1}{N}\sum_{n=1}^N x_nx_n^\top=X_N^TX_N$ and $\xi_N\in\mathbb{R}^N$ be a matrix with $\{\epsilon_n\}_{n=1}^N$ in its rows. Furthermore, assume that (i) ϵ_n and x_n are independent, (ii) $||\Sigma^{-1/2}V_N^{1/2}||_{\text{op}}^2\le 3/2$, and (iii) $\Sigma=\mathbb{E}[x_nx_n^\top]$ is PD. Then, with probability greater than $1-\delta$ it holds that*

$$\frac{1}{\sqrt{N}}||\Sigma^{-1/2}X_N^\top\xi_N||_2\le 3\sigma\sqrt{\frac{d\log\left(\frac{1}{\delta}\right)}{N}}.$$

Proof. Let $P_{X_N}\in\mathbb{R}^{d\times N}$ be the projection on the column space of $X_N^\top$. Observe that, by definition, $X_N^\top=P_{X_N}X_N^\top$ and $P_{X_N}=V_N^{1/2}(V_N^\dagger)^{1/2}$, and thus

$$\begin{aligned}
(62) &=\frac{1}{\sqrt{N}}||\Sigma^{-1/2}X_N^\top\xi_N||_2\\
&=\frac{1}{\sqrt{N}}||\Sigma^{-1/2}V_N^{1/2}(V_N^\dagger)^{1/2}X_N^\top\xi_N||_2\\
&\le\frac{1}{\sqrt{N}}||\Sigma^{-1/2}V_N^{1/2}||_{\text{op}}||X_N^\top\xi_N||_{V_N^\dagger}.
\end{aligned}\tag{59}$$

The first term is bounded by $3/2$ by assumption. We now bound the second term in (59). It holds that

$$||X_N^\top\xi_N||_{V_N^\dagger}^2=\frac{1}{N}\xi_N^\top X_N(X_N^TX_N)^\dagger X_N^\top\xi_N.\tag{60}$$

Furthermore, it can be verified that $\operatorname{tr}\left(X_N(X_N^TX_N)^\dagger X_N^\top\right)\le d$ and $||X_N(X_N^TX_N)^\dagger X_N^\top||_{\text{op}}\le 1$. Hence, Theorem 37 of (Hsu et al., 2012a), is applicable. Applying this result and assuming $\delta\in(0,e^{-1})$, we get

$$||X_N^\top\xi_N||_{V_N^\dagger}\le 2\sigma\sqrt{\frac{d\log\left(\frac{1}{\delta}\right)}{N}}.\tag{61}$$

with probability greater than $1-\delta$. This concludes the proof of the lemma. □

The following lemma allows us to translate performance of the OLS under the empirical design matrix to the performance under the expected empirical design matrix, i.e., the covariance matrix.

Lemma 39 (Translating Empirical to Expected Performance)**.** *Let $w \in \mathbb{R}^d$ be a vector, $V_N = \frac{1}{N}\sum_{n=1}^{N} x_n x_n^\top$ and $\Sigma = \mathbb{E}[xx^\top]$ be a PD matrix. Let $\Delta = \Sigma^{-1/2} V_N \Sigma^{-1/2} - I$ and assume that*

1. $||\Sigma^{-1/2} V_N w|| \leq \epsilon$.
2. $||\Sigma^{-1/2} V_N \Sigma^{-1/2} - I||_{\text{op}} \leq c < 1$.

Then, it holds that $||w||_\Sigma \leq \frac{\epsilon}{1-c}$.

Proof. We prove this result by standard analysis and by applying the assumptions. The following relations hold.

$$\begin{aligned}
||w||_\Sigma &= ||\Sigma^{1/2} w||_2 \\
&\leq ||\Delta \Sigma^{1/2} w||_2 + ||\Sigma^{-1/2} V_N \Sigma^{-1/2} \Sigma^{1/2} w||_2 && \text{(Triangle inequality)} \\
&= ||\Delta \Sigma^{1/2} w||_2 + ||\Sigma^{-1/2} V_N w||_2 \\
&\leq ||\Delta \Sigma^{1/2} w||_2 + \epsilon && \text{(By assumption)} \\
&\leq ||\Delta||_{\text{op}} ||\Sigma^{1/2} w||_2 + \epsilon && \text{(Submultiplicative property of norm)} \\
&= c||w||_\Sigma + \epsilon. && \text{(By assumption)}
\end{aligned}$$

Rearranging yields the result. □

Lemma 40 (Relative Spectral Norm Error, (Hsu et al., 2012b), Lemma 3)**.** *Let Σ_1, Σ_2 be a PD matrices. Let $\Delta = \Sigma_1^{-1/2}(\Sigma_2 - \Sigma_1)\Sigma_1^{-1/2}$. If $||\Delta|| < 1$ then*

$$||\Sigma_1^{1/2} \Sigma_2^{-1/2}||_{\text{op}}^2 = ||\Sigma_2^{-1/2} \Sigma_1^{1/2}||_{\text{op}}^2 = ||\Sigma_1^{1/2} \Sigma_2^{-1} \Sigma_1^{1/2}||_{\text{op}} \leq \frac{1}{1 - ||\Delta||_{\text{op}}}.$$

Proof. The first equality follows from the fact that $||A||_{\text{op}} = \lambda_{\max}(A^T A) = \lambda_{\max}(A^T A) = ||A^\top||_{\text{op}}$. The second equality follows from the fact that $||A^T A|| = \lambda_{\max}((A^T A)^2) = \lambda_{\max}(A^T A)^2 = ||A||^2$. The third inequality is proved in (Hsu et al., 2012b), Lemma 3. □

Proposition 41 (Ordinary Least Squares: In-Distribution Error)**.** *Let $\delta \in (0, e^{-1})$. Let $\{x_n, y_n\}_{n=1}^N$ be a data set where $x_n \in \mathbb{R}^d, y_n \in \mathbb{R}$, and assume that $y_n = \langle x_n, \beta_\star \rangle + \epsilon_n$ where ϵ_n is zero mean and σ sub-gaussian. Let $\Sigma = \mathbb{E}[xx^\top]$. Let $\widehat{\beta}$ be the solution of the ordinary least square objective*

$$\widehat{\beta} = V_N^\dagger X_N^\top Y_N$$

Then, assuming that $N \geq 9\gamma_{d,\delta}^2$ where $\gamma_{d,\delta}$ is defined in Lemma 42, with probability greater than $1 - \delta$

$$||\beta_\star - \widehat{\beta}||_\Sigma \leq 5\sigma \sqrt{\frac{d \log\left(\frac{2}{\delta}\right)}{N}}.$$

and $||\Sigma^{-1/2} V_N \Sigma^{-1/2} - I||_{\text{op}} \leq \frac{1}{3}$.

Proof. Let $X_N \in \mathbb{R}^{N \times d}$ be a matrix with $\left\{\frac{1}{\sqrt{N}} x_n\right\}_{n=1}^N$ in its rows, $V_N = \frac{1}{N}\sum_{n=1}^N x_n x_n^\top = X_N^T X_N$ and $\xi_N \in \mathbb{R}^N$ be a matrix with $\{\epsilon_n\}_{n=1}^N$ in its rows. Let $\widehat{\beta}$ be the OLS, i.e., it is the minimal norm solution that satisfies

$$V_N(\beta_\star - \widehat{\beta}) = \frac{1}{\sqrt{N}} X_N^\top \xi_N.$$

Multiply both sides of the above equation by $\Sigma^{-1/2}$ (where Σ is PD by assumption). We get that

$$\Sigma^{-1/2} V_N(\beta_\star - \widehat{\beta}) = \Sigma^{-1/2} X_N^\top \xi_N.$$

Taking the norm of both sides and by the triangle inequality we get

$$||\Sigma^{-1/2} V_N(\beta_\star - \widehat{\beta})||_2 = \frac{1}{\sqrt{N}} ||\Sigma^{-1/2} X_N^\top \xi_N||. \tag{62}$$

We bound this term by applying Lemma 36. To apply this result we bound, with high probability, $||\Sigma^{-1/2} V_N^{1/2}||_{\text{op}}^2 \leq \frac{3}{2}$ by the concentration of the empirical covariance matrix. It holds that

$$\begin{aligned}
&||\Sigma^{-1/2} V_N^{1/2}||_{\text{op}}^2 = ||\Sigma^{-1/2} V_N \Sigma^{-1/2}||_{\text{op}} \\
&\leq 1 + ||\Sigma^{-1/2} V_N \Sigma^{-1/2} - I||_{\text{op}} \\
&= 1 + ||\frac{1}{N} \sum_{n=1}^{N} (\Sigma^{-1/2} x_n)(\Sigma^{-1/2} x_n)^\top - I||_{\text{op}} \\
&\leq 1 + \frac{\gamma_{d,\delta}}{\sqrt{N}},
\end{aligned} \tag{63}$$

where the last relation holds with probability greater than $1 - \delta$ by Lemma 42 since $\Sigma^{-1/2} x_n \sim \mathcal{N}(0, I_d)$. Thus, if $N \geq 9\gamma_{d,\delta}^2$ it holds that $||\Sigma^{-1/2} V_N^{1/2}||_{\text{op}}^2 \leq \frac{3}{2}$. Hence, conditioning on this event, Lemma 36 is applicable. Thus, with probability greater than $1 - 2\delta$

$$||\Sigma^{-1/2} V_N^{1/2}||_{\text{op}}^2 = (62) \leq 3\sigma \sqrt{\frac{d \log\left(\frac{1}{\delta}\right)}{N}}. \tag{64}$$

We now translate this bound to a bound w.r.t. $||\beta_\star - \widehat{\beta}||_\Sigma$. To do so, we apply Lemma 39. Observe that (i) Σ is PD by assumption, (ii) conditioning on the good event $||\Sigma^{-1/2} V_N(\beta_\star - \widehat{\beta})||_2$ is bounded in (64), and, (iii) conditioning on the good event $||\Sigma^{-1/2} V_N \Sigma^{-1/2} - I|| \leq \frac{1}{3}$ (63) and by the choice of N. Thus, by Lemma 39 and setting $c = 1/3$, we get that conditioning on the good event that holds with probability greater than $1 - 2\delta$, for $N \geq 9\gamma_{d,\delta}^2$,

$$||\beta_\star - \widehat{\beta}||_\Sigma \leq 5\sigma \sqrt{\frac{d \log\left(\frac{1}{\delta}\right)}{N}}.$$

□

I. Matrix Concentration Results

Lemma 42 (Covariance Estimation for Sub-Gaussian Distributions, Corollary 5.50, (Vershynin, 2010) and Remark 5.51)**.** *Let $\delta \in (0, e^{-1})$. Consider a sub-gaussian distribution in $\mathbb{R}^d$ with covariance $\mathbb{E}[xx^\top] = I$. Let $\Sigma_N = \frac{1}{N}\sum_{n=1}^N x_n x_n^\top$ be the empirical covariance matrix. Then, with probability greater then $1-\delta$ it holds that*

$$||\Sigma_N - I||_{\text{op}} \le \frac{\gamma_{d,\delta}}{\sqrt{N}}.$$

for

$$N \ge \gamma_{d,\delta} \equiv \sqrt{Cd \log\left(\frac{1}{\delta}\right)}$$

where $C = C_K$ depends only on the sub-gaussian norm $K = ||x_i||_{\psi_2}$ and C is an absolute constant if $x \sim \mathcal{N}(0, I_D)$.

Lemma 43. *Let $z_1 \sim \mathcal{N}(0, I_{d_1}), z_2 \sim \mathcal{N}(0, I_{d_2})$ be independent random variables and $d = d_1 + d_2$. Assume that $N \ge \gamma_{d,\delta/3}$ where $\gamma_{d,\delta}$ is defined in Lemma 42. Then,*

$$\mathbb{P}\left(||\frac{1}{N}\sum_{n=1}^N z_{1,n} z_{2,n}^\top||_{\text{op}} \le \frac{3\gamma_{d,\delta/3}}{\sqrt{N}}\right) \ge 1-\delta.$$

Proof. Let $E_n = z_{1,n} z_{2,n}^\top$. Define the hermitian dilation of E to be

$$H_n = \begin{bmatrix} 0 & E_n \\ E_n^\top & 0 \end{bmatrix},$$

and see that (e.g., (Tropp, 2012), section 2.6)

$$||\frac{1}{N}\sum_n E_n||_{\text{op}} = \lambda_{\max}\left(\frac{1}{N}\sum_n H_n\right). \tag{65}$$

Hence, instead of bounding the first we can bound the latter.

Let $H_{1,n}, H_{2,n}, H_{f,n} \in \mathbb{R}^{d\times d}$ be defined as follows,

$$H_{1,n} = \begin{bmatrix} z_{1,n} z_{1,n}^\top & 0 \\ 0 & 0 \end{bmatrix}, \quad H_{2,n} = \begin{bmatrix} 0 & 0 \\ 0 & z_{2,n} z_{2,n}^\top \end{bmatrix}, \quad H_{f,n} = \begin{bmatrix} z_{1,n} z_{1,n}^\top & z_{1,n} z_{2,n}^\top \\ z_{2,n} z_{1,n}^\top & z_{2,n} z_{2,n}^\top \end{bmatrix}.$$

With these definitions we get that

$$H_n = H_{f,n} - H_{1,n} - H_{2,n}.$$

Furthermore, since $\mathbb{E}[H_n] = 0$ due to the independence of z_1 and z_2, and since they are assumed to be zero mean, we get that

$$\begin{aligned}\frac{1}{N}\sum_n H_n &= \frac{1}{N}\sum_n H_{f,n} - H_{1,n} - H_{2,n} \\ &= \frac{1}{N}\sum_n H_{f,n} - \mathbb{E}[H_{f,n}] - \frac{1}{N}\sum_n H_{1,n} - \mathbb{E}[H_{1,n}] - \frac{1}{N}\sum_n H_{2,n} - \mathbb{E}[H_{1,n}]. \\ & \qquad \text{(Since } \mathbb{E}[H_n] = \mathbb{E}[H_{f,n} - H_{1,n} - H_{2,n}] = 0)\end{aligned}$$

Hence, we can bound $\lambda_{\max}\left(\frac{1}{N}\sum_n H_n\right)$ by the following sum

$$\lambda_{\max}\left(\frac{1}{N}\sum_n H_n\right) \tag{66}$$

$$\le \lambda_{\max}\left(\frac{1}{N}\sum_n H_{f,n} - \mathbb{E}[H_{f,n}]\right) + \lambda_{\max}\left(\frac{1}{N}\sum_n H_{1,n} - \mathbb{E}[H_{1,n}]\right) + \lambda_{\max}\left(\frac{1}{N}\sum_n H_{2,n} - \mathbb{E}[H_{1,n}]\right) \tag{67}$$

since $\lambda_{\max}(A+B) \leq \lambda_{\max}(A) + \lambda_{\max}(B)$[3].

Observe that each one of the three summands is the deviation of the empirical covariance from its average. Applying Lemma 42 and by applying the union bound we get that with probability greater than $1-3\delta$, and assuming that $N \geq \frac{\gamma_{d,\delta}}{\sqrt{N}}$

$$(67) \leq \left(||\Sigma||\gamma_{d,\delta} + ||\Sigma_1||\gamma_{d_1,\delta} + ||\Sigma_2||\gamma_{d_2,\delta}\right)/\sqrt{N} \leq 3||\Sigma||\gamma_{d,\delta}/\sqrt{N} = 3\gamma_{d,\delta}/\sqrt{N},$$

where the last relation holds since $\Sigma_1, \Sigma_2 \leq \Sigma = I$ and since $\gamma_{d,\delta}$ is increasing in d. Finally, setting $\delta \leftarrow \delta/3$ yields the result. $\square$

[3]E.g., by using the variational form of maximal eigenvalue and since $\max_{x:||x||_2=1}(x^T Ax + x^T Bx) \leq \max_{x:||x||_2=1} x^T Ax + \max_{x:||x||_2=1} x^T Bx$

J. Experiment Details

In this section, we complete the details for the experimental setup outlined in Section 6.

Synthetic PC-LQs. We constructed a family of PC-LQ problems, parameterized by (s_c, s_e, d_u, d). The diagonal blocks $A_1 \in \mathbb{R}^{s_c \times s_c}$, $A_2 \in \mathbb{R}^{s_e \times s_e}$, $A_3 \in \mathbb{R}^{d-s_c-s_e, d-s_c-s_e}$ were generated by sampling each entry from $\mathcal{N}(0,1)$, dividing by the spectral radius (i.e. the largest modulus of complex eigenvalues), then multiplying by the desired spectral radius. We set $\rho(A_1) = 1$, to make the controllable part of the system marginally stable, and set $\rho(A_2) = \rho(A_3) = 0.9$. The matrices A_{12}, A_{32}, and B_1 were obtained by sampling each entry from $\mathcal{N}(0,1)$. Finally, for the LQR cost matrices, we selected $Q = I_{1+}$ and $R = I_{d_u}$.

System identification. Two system identification methods for estimating A were compared: ordinary least squares regression from x_1 onto x_0 (with the least-Frobenius norm solution), and the soft-thresholded semiparametric least squares estimator from Algorithm 1, with a choice of $\epsilon = 0.1$.[4] Fixing a sample size N, we sampled all $x_0 \sim \mathcal{N}(0, I)$ i.i.d., and $x_1 = Ax_0 + \eta_0$, where $\eta_0 \sim \mathcal{N}(0, I)$.

Certainty-equivalent control. We plugged these $(\hat{A}, B)$ into SciPy's discrete algebraic Riccati equation solver, which outputs the fixed-solution solution $P_\star$ under the nominal dynamics; then, the LQR cost of the derived controller on the true system was measured; if this was finite and within a factor of 1.1 of the optimal cost on the true dynamics, we called this trial (indexed by an independent sample) a success: the learned controller stabilized this marginally stable system.

We varied the sample size N between 100 and 1000 in increments of 20, and varied $d \in \{20, 50, 100, 150\}$, fixing $s_c = s_e = 5, d_u = 1$. Figure 1 shows the fraction of successful trials over 100 repetitions; error bars show normal approximation-derived standard deviations. All experiments took around 2 hours on a single 2.3 GHz Intel i7 CPU machine.

[4] With these synthetic systems, Algorithm 1 performed similarly with the thresholded OLS estimator.